TECHNISCHE UNIVERSITÄT MÜNCHEN
Department Chemie
Lehrstuhl II
für Organische Chemie

Protein Dynamics of ADT and KdpBN by NMR Spectroscopy

Markus Heller

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Dr. J. Buchner
Prüfer der Dissertation:	1. Univ.-Prof. Dr. H. Kessler
	2. Hon.-Prof. Dr. W. Baumeister
	3. Priv.-Doz. Dr. G. Gemmecker

Die Dissertation wurde am 30. 06. 2004 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 27. 07. 2004 angenommen.

To my parents.

A theory is something
nobody believes, except the
person who made it.
An experiment is something
everybody believes, except
the person who made it.

(Unknown)

Acknowledgement

The work presented in this thesis was prepared from February 2001 until June 2004 in the group of Prof. Dr. Horst Kessler at the Department of Chemistry of the Technical University of Munich, Germany.

I would like to thank my supervisor Prof. Dr. Kessler for giving me the opportunity to join his group after doing my diploma thesis at Hoffmann LaRoche, for the excellent research facilities, for unrestricted support, helpful discussions, and for giving me lots of personal freedom.

Of course there are a lot of other people I would like to thank:

- The staff members of the NCE: Michael John, Dr. Murray Coles, Melina Haupt, and Jochen Klages for the great atmosphere and ambitious science.
- Michael John for critical reading of my thesis and providing constructive suggestions while being busy with preparing his own manuscript.
- Dr. Gustav Gemmecker, Dr. Murray Coles, and especially Michael John for lots of useful discussions.
- My student "Knechts" Gerold Probadnik, Peter Kaden, Alexandra Rost, and Robert Huber for their commitment.
- The system administrators Dr. Rainer Haessner, Alexander Frenzel, and Monika Goede for their support and for keeping the LAN alive.

- Dr. Rainer Haessner and Dr. Gustav Gem(m)ecker for their support in NMR spectrometer hardware concerns.
- Dr. Martin Sukopp for synthesizing pA_4 and for the good collaboration on a couple of small peptides.
- Prof. Dr. Bernd Reif and Dr. Maggy Hologne for their collaboration in solid-state NMR.
- Dr. Gundula Bosch for preparing loads of ADT samples and delivering the treasures.
- The secretaries Beate Diaw, Marianne Machule, and Evelyn Bruckmaier for their professional work.
- Gerd "Geha" Hauser and Michael John for a nice and relaxing road trip across Florida.
- Dr. Hans Senn and Dr. Alfred Ross (Hoffmann LaRoche AG, Basel) for continuous support way beyond my diploma thesis.
- All other group members for a wonderful time inside and outside the lab.
- The LaTeX community represented by `(de.)comp.text.tex`; especially Axel Sommerfeldt, Donald Arseneau, and Walter Schmidt for their help.
- Last, but definetely not least, "Geha" for his very obliging supervision during the advanced practical NMR course. ;-)

I am indebted to my family, especially to my parents, for their endless support over the years. All this would not have been possible but for them.

Thank you.

Parts of this thesis have been published:

M. Heller, M. John, M. Coles, G. Bosch, W. Baumeister, H. Kessler, "NMR Studies on the Substrate-binding Domains of the Thermosome: Structural Plasticity in the Protrusion Region", *J. Mol. Biol.* **2004**, *336*, 717–729.

M. John, M. Heller, M. Coles, G. Bosch, W. Baumeister, H. Kessler, "Letter to the Editor: Backbone ^{1}H, ^{15}N and ^{13}C Resonance Assignments of α-ADT and β-ADT", *J. Biomol. NMR* **2004**, *29*, 209–210.

Contents

List of Figures

List of Tables

Abbreviations and Symbols

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
Å	Ångstrøm
ADT	Apical domain of the thermosome
B_0	Static magnetic field strength
χ^2	Weighting function in a least-squares fit
c	CSA coupling constant
$C(t)$	Rotational correlation function
CPMG	Carr-Purcell-Meiboom-Gill spin-echo sequence
CPP	Cyclic pentapeptide
CSA	Chemical shift anisotropy
$C_\mathrm{i}(t)$	Correlation function of the internal motion
$C_\mathrm{o}(t)$	Correlation function of the overall motion
D	Rotational diffusion tensor
d	Dipole–dipole coupling constant
DD	Dipole–dipole interaction
δ_iso	Isotropic chemical shift
$D_\parallel$	Parallel component of the rotational diffusion tensor
$D_\perp$	Perpendicular component of the rotational diffusion tensor
FMF	FASTModelfree
Γ	Sum squared error
γ	Gyromagnetic ratio
hetNOE	Heteronuclear Overhauser effect
HSQC	Heteronuclear single quantum correlation
Hz	Hertz

INEPT	Insensitive nuclei enhanced by polarization transfer
J	Scalar coupling constant
$J(\omega)$	Orientational spectral density function
k_B	Boltzmann constant
k_{ex}	Exchange rate constant
KdpBN	Nucleotide binding domain of Kdp
kHz	Kilohertz
LS	Lipari-Szabo
µs	Microsecond
mM	Millimolar
ms	Millisecond
MF	Model-free
MHz	Megahertz
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
ns	Nanosecond
ω	Larmor frequency
$\Delta\omega$	Chemical shift difference between sites A and B
ω_{eff}	Effective field strength
p_a, p_b	Populations of sites A and B
pm	Picometer
ppm	Parts per million (10^{-6})
ps	Picosecond
R_1	Longitudinal or spin-lattice relaxation rate constant
R_2	Transverse or spin-spin relaxation rate constant
$R_{1\rho}$	Transverse relaxation rate constant in the rotating frame
r_{NH}	Bond length between N and H
R_a, R_b	Transverse relaxation rate constants for sites A and B
R_2^{eff}	Effective transverse relaxation rate
R_{ex}	Exchange contribution to the relaxation rate constant
R.F.	Radio frequency
REDOR	Rotational echo double resonance
S^2	Generalized squared order parameter of the internal motion
S_f^2	Generalized squared order parameter of the fast internal motion
$\Delta\sigma$	Chemical shift anisotropy

σ_{xx}, σ_{yy}, σ_{zz}	Principal components of the CSA tensor
S_s^2	Generalized squared order parameter of the slow internal motion
s	Second
ssNMR	Solid-state NMR
τ_c	Rotational correlation time
T	Temperature
T_1	Longitudinal or spin-lattice relaxation time constant
T_2	Transverse or spin-spin relaxation time constant
$T_{1\rho}$	Transverse relaxation time constant in the rotating frame
τ_{cp}	Delay between two 180° pulses in a CMPG pulse train
τ_f	Correlation time of the fast internal motion
τ_i	Correlation time of the internal motion
τ_s	Correlation time of the slow internal motion
T	Tesla

Scope of this Work

This thesis is dedicated to the characterization of protein backbone dynamics by means of ^{15}N spin relaxation data. The introduction in chapter 1 gives an overview about the topic "molecular dynamics from NMR spin relaxation" and provides the most important literature references. Since the field of NMR and especially the characterization of dynamic processes is growing continuously, the introduction represents only a snap-shot and will soon be out-dated.

Chapter 2 describes the theory of ^{15}N NMR spin relaxation including cross-correlation effects. NMR experiments for measuring spin relaxation rates are discussed. Using this framework, the characterization of fast internal motions according to the model-free approach is introduced in a pictorial way, but still providing the most common equations.

Chapter 3 addresses the analysis of a two-site chemical exchange process on a millisecond time scale using relaxation dispersion data. In addition to the theoretical background, the constant relaxation time CPMG experiment is explained. Furthermore, the extraction of the exchange parameters from the experimental data is described. Although statistical methods are mentioned in the latter two chapters, they are not discussed in detail; the reader is referred to statistical textbooks for a comprehensive description.

Based on the preceeding theoretical explanations, chapter 4 demonstrates in detail how NMR relaxation data can be analyzed using self-written scripts. Every step is tackled: generation of peak lists, reformatting of lists, creating input files for specialized software, data fitting, including a description of all relevant files. The model-free analysis as well as the analysis of relaxation dispersion data are demonstrated using two examples. This framework is intended to pro-

vide a guideline, thereby facilitating the analysis of relaxation data. In addition, two complete examples are given in appendix C.

Using the techniques introduced in the preceeding chapters, dynamic processes in two different proteins were probed. Intrinsic disorder was revealed in the helical protrusions of the apical domains of the thermosome from *Thermoplasma acidophilum* (chapter 5). In chapter 6, the effects of AMP-PNP binding to the nucleotide binding domain of the P-type ATPase Kdp on internal motions on fast (pico- to nanosecond) and slow (millisecond) time scales were investigated.

Chapter 7 describes how solution- and solid-state NMR techniques can be combined advantageously to investigate hydrogen bonds in a small cyclic pentapeptide.

Solvent suppression in NMR experiments with signal detection on aliphatic protons is a critical issue on high-field spectrometers and becomes even more important if a cryogenic probe is to be used. Chapter 8 demonstrates how good water suppression can be achieved for selected NMR experiments implemented on a 600 MHz spectrometer equipped with a cryo probe.

Chapter 1

Introduction

Since its discovery in the 1940s, Nuclear Magnetic Resonance (NMR) Spectroscopy and Magnetic Resonance Imaging (MRI) have become powerful, interdisciplinary methods. A brief historical review would reveal as many as nine Nobel Prize laureates. *Isador I. Rabi* developed the resonance method for recording the magnetic properties of atomic nuclei. In 1944, he was awarded the Nobel Prize in Physics. The NMR phenomenon was first demonstrated for protons by *Felix Bloch* and *Edward M. Purcell* in 1946; six years later, they too were awarded the Nobel Prize in Physics. After years of continuous development, *Richard Ernst* received the Nobel Prize in Chemistry for his fundamental contributions to the NMR methology. A decade later, *Kurt Wüthrich* shared the Nobel Price in Chemistry for determining the three dimensional structure of biological macromolecules in solution. In 2003, the Nobel Prize in Physiology and Medicine was awarded jointly to *Paul C. Lauterbur* and *Sir Peter Mansfield* for their pioneering contributions to enabling the use of magnetic resonance in medical imaging. All these contributions could not have been achieved without

superconducting materials. In 2003, *Alexeij A. Abrikosow* and *Vitalij L. Ginzburg* were each awarded one third of the Physics Nobel Price for their contributions to the theory of superconductors.

Today, nuclear magnetic resonance is a method with a large variety of applications. It is used by organic chemists to control reactions, analyze product mixtures and determine the structure of organic compounds. The pharmaceutical industry has discovered NMR as an invaluable tool, stimulated by the "SAR-by-NMR" approach, the structure-activity relationship by nuclear magnetic resonance.[1] Since then, NMR is used in several instances of the drug development process, and new techniques have emerged (for recent reviews see refs[2,3,4,5]); a promising approach is the combination of NMR and *in-silico* screening.[6]

Investigations of biological macromolecules, including structure determination and characterization of molecular motions, represent a large field in NMR. NMR spectroscopy of proteins has made a tremendous progress during the last decade, profiting from continuously improved hardware and methodological development. About ten years ago, the upper limit of molecular weight for proteins amenable to NMR studies was approximately 15 kDa. Recently, backbone and Ile, Leu, and Val sidechain methyl assignments as well as ^{15}N backbone relaxation data have been reported for malate synthase G, a single-domain, 723-residue protein with a molecular mass of 82 kDa;[7,8] slow internal motions of proteins with molecular weights of 53 and 82 kDa have also been investigated.[9,10] This illustrates the enormous improvement and potential in the field of biomolecular NMR spectroscopy. The present work focusses on the characterization of molecular motion in proteins using backbone ^{15}N relaxation data.

Why study molecular dynamics? In the past decade, a large number of three dimensional (3D) structures of biological macromolecules, especially proteins, have been solved by X-ray crystallography and multidimensional NMR techniques. Although this effort has provided a wealth of information on protein

architecture, it has become evident that the picture of a static structure alone is not sufficient to explain the modes of action of a protein. A vast number of hints point to the biological importance of dynamic processes, especially on slower time scales. Protein function strongly depends on changes in the 3D structure in response to specific molecular interactions.[11] For example, access of a ligand to the active site of an enzyme may require conformational rearrangements. Enzyme catalysis and ligand off-rates have been measured to be on the order of 10–10^5 s^{-1},[12] and folding rates for small globular proteins are in the range of 10^{-1}–10^5 s^{-1}.[13]

NMR spectroscopy has been successfully used to measure protein folding rates,[14,15] and for direct observation of protein-ligand interaction kinetics.[16] Furthermore, the insights heteronuclear NMR relaxation measurements can provide on the role of protein motions in molecular recognition have been reviewed.[17] But not only slow motions on a milli- to microsecond time scale are of interest. Much faster motions on a nano- to picosecond time scale seem to be associated with entropy in the folded state of a protein (see ref[18] and references cited therein). Fast internal motions of proteins have the potential to report on the number of states that are accessible to each single site in a protein and thus can act as an "entropy meter". Although a lot of open questions remain, it seems to be clear that proteins indeed have a considerable amount of residual entropy and that changes in functional states are often associated with redistributions of this entropy.[18]

Why use NMR? One of the most distinct advantages of NMR over any other method is its capability of probing molecular motions with atomic resolution. In combination with isotope labeling techniques (see, for example ref[19,20]), virtually every single atom in a protein is accessible. NMR spin relaxation measurements provide information on amplitudes and time scales of internal motions, albeit care has to be taken when interpreting the results in terms of motional models, since a number of models may be consistent with the parameters derived from relaxation rate constants.[21,22]

Solid-state NMR (ssNMR) relaxation experiments are not affected by overall rotational diffusion, hence additional information about molecular dynamics is gained from anisotropic quadrupolar and chemical shift interactions, which are averaged to zero in solution. However, ssNMR usually requires incorporation of expensive specific labels, thus limiting the number of sites that can be studied.

X-ray crystallographic *B*-factors are sensitive to the mean square displacements of heavy atoms due to thermal motions. Although *B*-factors are available at nearly all heavy atom positions, they do not provide information about the time scale of the thermal motions and are furthermore subject to static disorder, crystal packing effects and refinement protocols. Although qualitative agreement between squared order parameters derived from NMR relaxation data and crystallographic *B*-factors is often observed, quantitative correlations are weak.[23] Diffusive X-ray scattering provides additional information on correlated motions of heavy atoms, while information on amplitudes, correlation times and correlation lengths for fluctuations of hydrogen atoms is obtainable from incoherent quasi-elastic neutron scattering (IQNS). However, no theoretical methods for comparing data derived from NMR relaxation, diffusive X-ray scattering and IQNS have been reported until today.

Time-resolved fluorescence anisotropy decay is sensitive to overall rotational diffusion and internal motions on time scales comparable to the fluorescence lifetime of the fluorophore. The analysis of the anisotropy decay is similar to the analysis of NMR relaxation data in many respects, and the model-free formalism (see section 2.3 on page 29) is equally applicable. However, spectroscopy of intrinsic fluorophores in proteins is restricted to the aromatic amino acids Tyr and Trp, and a quantitative analysis even requires a unique fluorophore.

Molecular dynamics (MD) simulations provide details on the dynamic behaviour of proteins to the atomic level.[24] After the first MD simulation of a protein about 25 years ago,[25] they have become increasingly useful since force fields have improved and computers became more powerful. In principle, all

information that can be obtained from NMR spin relaxation data is contained in a MD trajectory. Comparison between order parameters derived from MD and NMR indicates that for rigid proteins, nearly quantitative agreement can be obtained, although care has to be taken while performing the comparison. With increasing computational power, the major drawback of MD simulations is remedied: the limited length of the trajectories. With trajectories much longer than the rotational correlation time, new informations about the coupling between internal and rotational motions should be obtainable.

All methods mentioned here have their own advantages, but also their disadvantages. A prerequisite for X-ray crystallography are crystals of high quality with good diffraction properties, which may be very tedious to prepare. Solid-state NMR and time-resolved fluorescence spectroscopy are limited to only a few (if not a single) sites, although rapid progress is made in the field of ssNMR. Nuclear magnetic resonance in the liquid state enables the scientist to probe internal motions occuring on various time scales at an atomic level under nearly physiological conditions.[26,27] The combination of NMR relaxation data with results from MD simulations is likely to contribute to the better understanding of the relationship between dynamics and functions of proteins.

What Information Does NMR Relaxation Provide? In principle: everything, depending on the questions to be adressed and the amount of money that can be spent. Fast internal motions can be probed by laboratory frame relaxation measurements of spins whose relaxation mechanisms are associated with a bond vector. The data are most commonly analyzed using the "model-free" approach introduced by Lipari and Szabo (chapter 2.3),[21,22] although application of this formalism implies strong restrictions . Informations about internal motions faster than overall tumbling are obtained, such as spatial restrictions and effective or internal correlation times (for recent reviews see refs[17,28,29]). While these fast motions seem to be related to the residual entropy of folded proteins,[18] they also provide information on the flexibility of the molecule under study. NMR relaxation in the laboratory frame is sensitive to *rotations* of a

given bond vector around a *perpendicular* axis; *translations* of the same vector remain unnoticed.[27]

In order to obtain a picture of internal motions as complete as possible, it is thus useful to investigate the relaxation properties of different nuclei. However, in most cases, only ^{15}N laboratory frame relaxation is used to characterize motions of protein backbones; this has become virtually routine, and several programs for convinient data analysis are available.[30,31,32,33] Several reasons are responsible for this development: isotope labeling with ^{15}N is relatively cheap; furthermore, the amide moiety represents an almost ideal two spin AX system (this situation can be improved if deuteration of all non-exchangeable protons is feasible); hence, data analysis is straightforward.

Things become more complicated if ^{13}C nuclei are the subject of interest. If only a uniformly labeled ^{13}C-sample is available, additional scalar couplings, dipole–dipole interactions, and cross-correlation effects between the carbons render NMR experiments and data analysis complicated.[34,35] Despite of this, interesting results were obtained using ^{15}N in combination with ^{13}C relaxation. Anisotropic motions of the peptide plane have been detected, which could not be identified from ^{15}N relaxation data alone.[36,37] The development of new labeling strategies had—and still has—a tremendous impact on the studies of sidechain dynamics in proteins.[19,20] For the studies of ^{13}C relaxation, CH_n groups can be converted into CHD_{n-1} moieties *via* deuteration,[38] thus circumventing some of the difficulties mentioned above. Further simplification can be achieved with labeling schemes yielding alternating $^{12}C - ^{13}C - ^{12}C$ patterns to suppress the large $^1J_{CC}$ coupling.[39] Quite recently, deuterium relaxation was discovered as a reporter of sidechain dynamics in proteins. In these studies, deuterons of ^{13}CHD methylene and $^{13}CH_2D$ methyl groups are used as reporters;[40,41] this approach has been used to investigate sidechain dynamics in a number of different proteins.

Motions slower than overall molecular tumbling (on a ms–µs time scale) can be characterized using transverse relaxation in the laboratory (R_2) or rotating frame ($R_{1\rho}$); this topic has been reviewed recently.[17,42,43,27] If a chemical or

conformational exchange process alters the magnetic environment of a nucleus, then the resulting time dependence of the resonance frequencies leads to contributions to transverse relaxation,[44,45,46,47,48] often referred to as "R_{ex}". R_{ex} is characterized by measuring the dependence of the transverse relaxation rate as a function of the effective field strength. Two methods exist: determination of R_2 using Carr–Purcell–Meiboom–Gill based experiments or determination of $R_{1\rho}$ by spin-lock experiments.[49,50] The introduction of a ^{15}N relaxation experiment employing an off-resonance radio frequency (R.F.) field has enabled studies of exchange phenomena on a µs time scale.[51] Improvements of this experiment, including the use of weak R.F. fields, have made possible the characterization of exchange processes on time scales that could not be addressed so far.[52] Three years later, a modified CPMG-type spin-echo experiment—also referred to as "relaxation dispersion experiment"—for monitoring motions on a millisecond time scale was introduced and has been improved continously.[53,54] Recently, a detailed characterization of a disulfide bond isomerization in BPTI (basic pancreatic trypsin inhibitor) was achieved using a combination of these experiments and chemical shift modeling.[55]

Relaxation dispersion experiments have been applied to ^{13}C nuclei as well. For uniformly ^{13}C-enriched samples, complications arise especially from the large one-bond ^{13}C–^{13}C coupling constant, and the first study was restricted to methionine sidechain methyl groups in a cavity mutant of T4 lysozyme.[56] Using $^{13}CH_3$-pyruvate as the sole carbon source, the applicability of CPMG dispersion spectroscopy was extended to a large number of Val, Leu, Ile, Ala, and Met sideschain methyl groups in the same protein.[57] Using band-selective pulses, R_{ex} contributions to transverse CO relaxation can also be measured in uniformly ^{13}C labeled samples,[58] as well as for amide protons;[59] although in the latter case, deuteration of all non-exchangeable protons is a great benefit. Combining dispersion experiments of amide proton and nitrogen as well as carbonyl carbons allows for a detailed characterization of ms–µs time scale motion of protein backbone nuclei, which in turn yields information about slow dynamics of hydrogen bonding networks in proteins.[58] Efforts in en-

hancing the accuracy of the measurements and the parameters derived from the data have also been made during the last years. While the CPMG dispersion experiments mentioned above were applied to single quantum coherences, methods for measuring the decay of multiquantum coherences have been introduced.[60,61,62] When data from single and multiple quantum experiments are combined, a more quantitative picture is obtained, making it possible to distinguish between a two-site and more a complicated exchange process.

Chapter 2

Fast Internal Motions

This chapter explains the basics of ^{15}N NMR spin relaxation and focusses on the methods that have been used in this work, i.e. characterization of molecular dynamics faster than the overall molecular tumbling using the model-free approach and ^{15}N relaxation in the laboratory frame. A detailed description of NMR relaxation theory is far beyond the scope of this work; this chapter should rather be seen as a collection of the basic principles and literature references. A number of reviews on the topic of spin relaxation have appeared recently,[27, 34, 29, 63, 28, 17] and the appropriate chapters in *Protein NMR Spectroscopy – Principles and Practice* and *Spin Dynamics – Basics of Nuclear Magnetic Resonance* provide a pictorial derivation of the theory.[64, 65] A detailed description of the model-free analysis of backbone ^{15}N NMR relaxation data for a small number of residues is given in chapter 4.

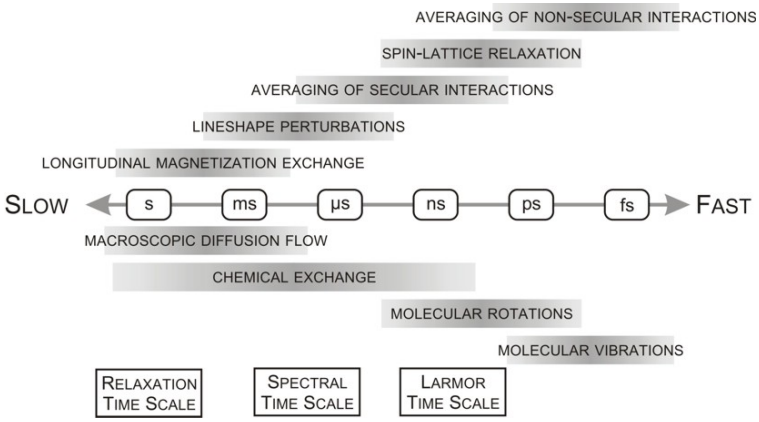


FIGURE 2.1: Motional time scales and their association with different NMR phenomena. Figure adapted from *Spin Dynamics*.[65]

2.1 NMR Relaxation of Spin-$\frac{1}{2}$-Nuclei

Protein dynamics are complex and difficult to analyze, since a variety of motions occur on different time scales. Figure 2.1 illustrates the time scales of a variety of motions relevant for biomolecular NMR. The effect of the motional processes depend on their relationship to three characteristic time scales of a nuclear spin system, as illustrated in figure 2.1.

- The *Larmor time scale* indicates the time required for a spin to precess through 1 radian in the magnetic field. The time scale τ_{Larmor} is defined as

  $$\omega\tau_{\text{Larmor}} \sim 1$$

  where ω is the Larmor frequency of the spins. Consider as an example the Larmor frequency of the spin as $\omega/2\pi = 600$ MHz, then τ_{Larmor} is approximately 0.26 ns.

- The *spectral time scale* is given by the inverse width of the NMR spectrum. Consider a two spin system with the chemical shifts of the nuclei being Ω_1 and Ω_2. If the chemical shift interactions are dominant, the spectral time scale τ_{spect} or more precisely, the *chemical shift time scale* τ_{shift} is given by

  $$|\Omega_1 - \Omega_2|\tau_{\text{shift}} \sim 1$$

  A chemical shift difference of two ^{13}C nuclei of 5 ppm in a static magnetic field B_0 of 14.1 T would define a chemical shift time scale of ~ 0.2 ms.

- The *relaxation time scale* indicates the order of the spin-lattice relaxation time constant T_1. For proteins, this is usually of the order of seconds.

All time scales depicted in figure 2.1 are accessible to NMR experiments. Macroscopic diffusion is related to the transverse diffusion coefficient of a molecule and can be probed by NMR experiments using pulsed field gradients (PFGs). This is, however, not a method based on relaxation; hence it is not described here. Molecular dynamics occuring on a very slow time scale on the order of several seconds lead to longitudinal magnetization exchange and can be quantified using exchange spectroscopy.[66] As can be seen in figure 2.1, motions on a ms–µs time scale lead to lineshape perturbations and thus affect transverse relaxation in the laboratory or rotating frame.[42] Motions on time scales faster than nanoseconds are usually characterized by measuring longitudinal and transverse relaxation rates and interpreting them in terms of the model-free formalism,[21,22] or using reduced spectral density mapping.[67]

2.1.1 Spin-$\frac{1}{2}$-Nuclei in an External Magnetic Field

Nuclei with a nuclear spin $I \neq 0$ interact with magnetic fields. In the absence of a magnetic field, all $2I + 1$ energy levels are degenerate. The application of an external magnetic field B_0 leads to a splitting of the energy levels and thus to a loss of degeneracy. This effect is known as the *Zeemann effect*. A nucleus with a spin $I = \frac{1}{2}$ has two energy levels α and β. For a system of two coupled spin-$\frac{1}{2}$

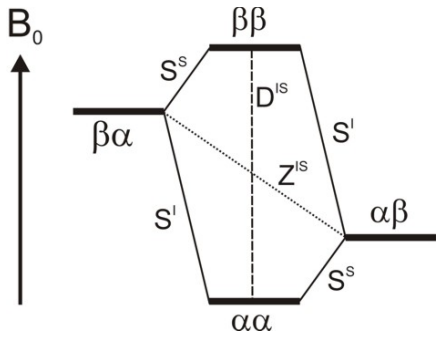


FIGURE 2.2: Energy levels of a system of two coupled spin-$\frac{1}{2}$ nuclei I and S with $\gamma_{I/S} > 0$. The straight lines indicate single-quantum transitions, while the dashed and dotted lines represent double and zero-quantum transitions, respectively.

nuclei I and S, the Zeeman effect leads to four energy levels as illustrated in figure 2.2. The energy levels are coupled to each other *via transitions* denoted S^I, S^S, D^{IS}, and Z^{IS}. A transition involving a change of the spin state of only one spin (e.g. $\alpha\alpha \leftrightarrow \beta\alpha$, S^I) is called *single quantum transition*. *Multi quantum transitions* are associated with transitions of both spins. The transition $\alpha\alpha \leftrightarrow \beta\beta$, D^{IS}, is referred to as *double quantum* or *flip-flip transition*. On the other hand, the transition $\beta\alpha \leftrightarrow \alpha\beta$, Z^{IS}, is called *zero quantum* or *flip-flop transition*. The transition frequencies as well as the corresponding transition probabilities are given in table 2.1. The latter can be used to describe relaxation rates (see section 2.1.4).

If nuclear spins are undisturbed for a long time in a magnetic field, they reach a state of *thermal equilibrium*. This implies that all coherences are absent and that the populations follow the *Boltzmann distribution* at the given temperature. The process during which the system returns to its thermal equilibrium is called *relaxation*. Unlike in optical spectroscopy, spontaneous as well as stimulated emission have negligible influence on NMR relaxation. Instead, nuclear spin relaxation is a consequence of coupling of the spin system to the *surround-*

Table 2.1: Transitions in a system of two coupled nuclei I and S with a spin of $\frac{1}{2}$ in an external magnetic field B_0 as depicted in figure 2.2. The transition probabilities are proportional to the spectral density function $J(\omega)$; these are introduced in section 2.1.3.

transition	transition frequency	transition propability[a]
Z^{IS}	$\omega_I - \omega_S$	$W_0^{IS} = c^2 J(\omega_I - \omega_S)/24$
S^I	$\omega_I = \gamma_I B_0$	$W_1^I = c^2 J(\omega_I)/16$
S^S	$\omega_S = \gamma_S B_0$	$W_1^S = c^2 J(\omega_S)/16$
D^{IS}	$\omega_I + \omega_S$	$W_2^{IS} = c^2 J(\omega_I + \omega_S)/4$

[a] For the dipolar interaction, c is defined as $c = \sqrt{6}\,(\mu_0/4\pi)\,\hbar\gamma_I\gamma_S r_{IS}^{-3}$.

ings or *lattice*. The lattice influences the local magnetic fields at the nuclei and therefore couples the spin system to the lattice. Stochastic Brownian motions of molecules in solution (see below) render these variations time-dependent. The field variations can be decomposed into components parallel and perpendicular to the static B_0 field. *Transverse* components of the stochastic local field lead to *nonadiabatic* contributions to relaxation. These contributions lead to transitions between energy states and thus allow for energy transfer between the spin system and the lattice. This energy exchange brings the system back to the thermal equilibrium. Components of the stochastic local field *parallel* to the static field cause random fluctuations of the Larmor frequencies of the spins. Thus, these *adiabatic* contributions to relaxation lead to a loss of coherence.

2.1.2 Relaxation Mechanisms

As discussed above, relaxation of nuclei with a spin of $\frac{1}{2}$ is caused by fluctuations in the local magnetic field at the site of the spins. Let us consider the direct *dipole–dipole* interaction between two adjacent spins in the same molecule, e.g. a ^{15}N–^{1}H spin pair in the backbone of a protein. Every dipole has its own local dipolar field. Depending on the orientation of the ^{15}N–^{1}H bond vector with

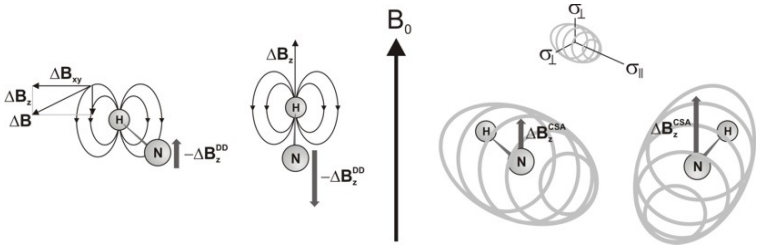


FIGURE 2.3: The main relaxation mechanisms for spin-$\frac{1}{2}$ nuclei. LEFT: The dipolar field of a nucleus leads to fluctuations in the local magnetic field of an adjacent nucleus due to molecular tumbling. RIGHT: Modulation of the local magnetic field due the chemical shift anisotropy. The local fields at the ^{15}N nucleus are symbolized by dark grey arrows.

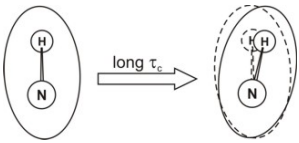
respect to the static magnetic field, the dipolar field of the proton influences the local field of the nitrogen (see figure 2.3). Any random change in the orientation of the bond vector will cause fluctuations in the local magnetic field at the ^{15}N spin and thus lead to a relaxation process. Note that the local dipolar fields at ^{1}H and ^{15}N have opposite signs ($\gamma_{\mathrm{H}} > 0$, whereas $\gamma_{\mathrm{N}} < 0$).

Chemical shifts are the results of electron motions induced by the external magnetic field. These motions of electrons generate secondary magnetic fields which can enhance or weaken the main static field. The slightly different local magnetic fields for each nucleus lead to different Larmor frequencies and thus to different chemical shifts. Generally, these local fields are orientation-dependent, i.e. anisotropic, and provide the basis of to the *chemical shift anisotropy CSA*. The CSA tensor can be described by three principal components, σ_{xx}, σ_{yy} and σ_{zz}. For ^{15}N, the CSA tensor has axial symmetry and is oriented approximately colinear with the bond vector (see figure 2.3). Changes in the orientation of the bond vector with respect to the external field cause fluctuations in the local magnetic field of the nitrogen, which in turn lead to relaxation processes.

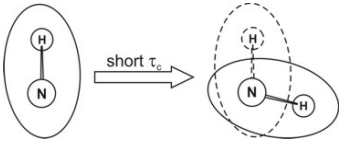
CSA represents an important relaxation mechanism only for nuclei with a large range of chemical shifts; thus, CSA contributions to the relaxation of protons are negligible. In biomolecular NMR spectroscopy, CSA relaxation is important for aromatic and carbonyl ^{13}C as well as for ^{15}N and ^{31}P nuclei. The relaxation rate has a quadratic dependence on B_0; therefore, use of high magnetic field strengths does not necessarily improve the sensitivity. For large molecules at high magnetic fields, relaxation interference between dipole–dipole and CSA relaxation mechanisms occurs, which forms the basis of transverse relaxation optimized spectroscopy (TROSY).[68,69] Similar to the CSA mechanism, *spin rotation* of methyl groups can also lead to fluctuations in local magnetic fields. The usual order of importance of relaxation mechanisms for spin-$\frac{1}{2}$ nuclei is dipole–dipole > chemical shift anisotropy.

2.1.3 Correlation and Spectral Density Functions

So far, the direct dipole–dipole interaction and the CSA have been discussed as mechanisms leading to flucuations in the local magnetic field at the site of a nucleus. It has been shown that these local fields depend on the orientation of the ^{15}N–^{1}H bond vector with respect to the external field B_0. Consider a ^{15}N–^{1}H spin pair with a *fixed* orientation with respect to a molecular frame of reference. The orientation of the ^{15}N–^{1}H bond vector changes as the molecules tumbles in solution due to Brownian motion. The magnitude of the change depends on how fast the molecule tumbles. As an example, consider the orientation of the bond vector at time t and at a time $t + \delta$. For a large molecule which rotates slow, the orientation at $t + \delta$ is very similar to the orientation at time t: both orientations are correlated to a high degree.



On the other hand, if the molecule tumbles fast, the bond vector orientations at time t and $t + \delta$ are very different. They are not correlated to each other any more:



This loss of correlation can be described by a *correlation function* $C(t)$. For isotropic rotational diffusion of a spherical top, $C(t)$ is given as

$$C(t) = A_C \cdot e^{-\frac{t}{\tau_c}} \tag{2.1}$$

where the normalization constant A_C equals $\frac{1}{5}$ and τ_c is the rotational correlation time of the molecule. For the assumptions made for equation 2.1, the rotational correlation time is related to the hydrodynamic properties *via* Stoke's law as:

$$\tau_c = \frac{4\pi\eta_W r_H^3}{3k_B T} \tag{2.2}$$

in which η_W is the viscosity of the solvent, r_H is the effective hydrodynamic radius of the solute, k_B is the Boltzmann constant, and T is the temperature. Large values of τ_c correspond to slow tumbling (large molecules, low temperatures), whereas small τ_c indicate fast tumbling (small molecules, high temperature). Generally, raising the temperature results in smaller correlation times. Fourier transformation of the correlation function yields the corresponding *spectral density function* $J(\omega)$:

$$J(\omega) = A_J \cdot \frac{\tau_c}{1 + \omega^2\tau_c^2} \tag{2.3}$$

which respresents a *Lorentzian* function. For the correlation function given in equation 2.1 with $A_C = \frac{1}{5}$, the normalization constant A_J equals $\frac{2}{5}$. As illustrated in figure 2.4, short correlation times lead to broad spectral densities and *vice versa*.

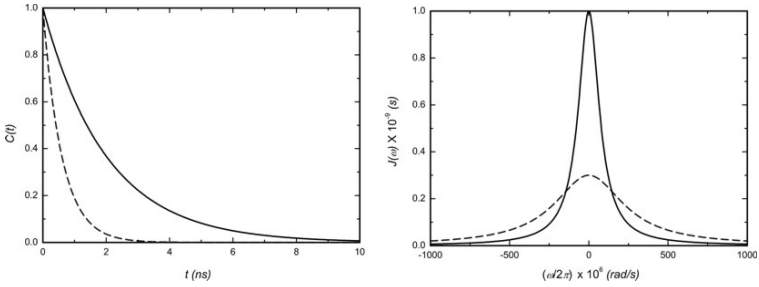


Figure 2.4: Correlation functions (Left) and spectral densities (Right) illustrating the relation between the correlation time τ_c and $J(\omega)$. Solid lines correspond to $\tau_c = 2\,\text{ns}$, dashed lines represent $\tau_c = 0.6\,\text{ns}$. The correlation function were calculated using equation 2.1; for the spectral densities, equation 2.3 was used. Note that both equations were scaled to 1.

2.1.4 Longitudinal Relaxation

Spin-lattice or *longitudinal relaxation* is the process of spin populations returning to their Boltzmann equilibrium. Spin-lattice relaxation is characterized by a time constant T_1 or its reciprocal R_1, the spin-lattice relaxation rate. The longitudinal relaxation rate is given by:

$$\begin{aligned} R_1 &= R_1^{\text{DD}} + R_1^{\text{CSA}} \\ &= \frac{d^2}{4}\left[J(\omega_\text{H} - \omega_\text{X}) + 3J(\omega_\text{X}) + 6J(\omega_\text{H} + \omega_\text{X})\right] + c^2 J(\omega_\text{X}) \end{aligned} \tag{2.4}$$

in which the dipolar coupling constant d and the CSA coupling constant c are given as

$$d = \frac{\mu_0 \hbar \gamma_\text{H} \gamma_\text{X}}{4\pi r_\text{XH}^3} \quad \text{and} \quad c = \frac{1}{\sqrt{3}}\omega_\text{X}\Delta\sigma \tag{2.5}$$

where μ_0 is the permeability of the free space; $\hbar$ is Planck's constant divided by 2π; γ_H and γ_X are the gyromagnetic ratios of ^{1}H and the X spin (in our case ^{15}N), respectively; r_XH is the X–H bond length; ω_H and ω_X are the Larmor frequencies of the ^{1}H and X spins, respectively; and $\Delta\sigma$ is the chemical shift

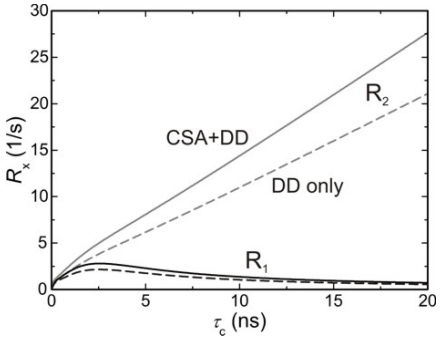


FIGURE 2.5: Longitudinal (black) and transverse (grey) relaxation rates for a ^{15}N–^{1}H spin pair, calculated with equations 2.4 and 2.6. Calculations were perfomed assuming $B_0 = 14.1$ T. In the calculation of the solid curve, CSA and dipole–dipole interactions were considered; CSA contributions were omitted for calculation of the dashed curve.

anisotropy of the X spin, assuming an axially symmetric chemical shift tensor. In this work, an effective ^{15}N–^{1}H distance of 1.02 Å (corrected for librations) and a nitrogen CSA of -160 ppm were used. Figure 2.5 shows plots of R_1 *vs.* τ_c. Note that longitudinal relaxation is slow in the *slow tumbling* limit ($|\omega\tau_c| \gg 1$) and in the *extreme narrowing* limit ($|\omega\tau_c| \ll 1$).

2.1.5 Transverse Relaxation

The decay of coherences is called *spin-spin* or *transverse relaxation*, characterized by the time constant T_2; the transverse relaxation rate is given by R_2.

$$\begin{aligned} R_2 &= R_2^{\mathrm{DD}} + R_2^{\mathrm{CSA}} + R_{\mathrm{ex}} \\ &= \frac{d^2}{8}\left[4J(0) + J(\omega_{\mathrm{H}} - \omega_{\mathrm{X}}) + 3J(\omega_{\mathrm{X}}) + 6J(\omega_{\mathrm{H}}) + 6J(\omega_{\mathrm{H}} + \omega_{\mathrm{X}})\right] \\ &\quad + \frac{c^2}{6}\left[4J(0) + 3J(\omega_{\mathrm{X}})\right] + R_{\mathrm{ex}} \end{aligned} \tag{2.6}$$

where all constants are the same as defined above. R_{ex} is an additional contribution to transverse relaxation due to chemical exchange and is discussed in detail in chapter 3. Note the contribution of the spectral density at zero frequency $J(0)$. Plots of R_2 *vs.* τ_c for a ^{15}N–^{1}H spin pair are shown in figure 2.5. In contrast to the longitudinal relaxation rate constant, R_2 increases monotonically with increasing τ_c.

2.1.6 The Heteronuclear NOE

Application of a weak radio frequency (R.F.) field at the resonance frequency of a spin I for a sufficient long time affects the longitudinal magnetization of another spin X in spatial proximity. This effect is called the *steady state nuclear Overhauser effect, steady state NOE*. The steady state heteronuclear NOE (hetNOE) can be described quantitatively by

$$NOE = 1 + \frac{d^2}{4\,R_1}\frac{\gamma_X}{\gamma_H}\left[6\,J\left(\omega_H + \omega_X\right) - J\left(\omega_H - \omega_X\right)\right] \tag{2.7}$$

Figure 2.6 shows a plot of the NOE *vs.* the correlation time calculated for a ^{15}N–^{1}H spin pair. Note the change in sign: for pure dipole–dipole interactions, the theoretical limits for extreme narrowing and slow tumbling are -3.93 and 0.78, respectively.

2.1.7 Effects of Cross-Correlation Between Dipole–Dipole and CSA Relaxation Mechanisms on Relaxation Rates

The dipole–dipole interaction constant d and the strength of the chemical shift anisotropy interaction c (equation 2.5) have the following values for a ^{15}N–^{1}H spin pair at a B_0 field of 14.1 T (assuming a distance of 1.02 Å and a CSA of -160 ppm): $d = -72.1 \times 10^3\,s^{-1}$ and $c = -34.8 \times 10^3\,s^{-1}$. Both values are of the same order of magnitude and especially for higher correlation times, the CSA contributions become more important (see figure 2.5). It has been shown

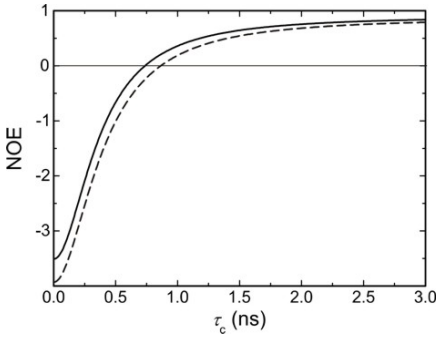


FIGURE 2.6: Plot of the heteronuclear NOE for a ^{15}N–^{1}H spin pair, calculated with equation 2.7 for $B_0 = 14.1$ T. In the calculation of the solid curve, CSA and dipole–dipole interactions were considered; CSA contributions were omitted for calculation of the dashed curve.

earlier that *cross-correlation effects* can have significant effects on the relaxation rates derived from NMR measurements unless precautions are taken.[70,71]

Goldman has shown that transverse relaxation of the multiplet components i and j of spin A in an AX system is given by:[72]

$$\frac{d}{dt}M^i_{\text{tr}} = -(\lambda+\eta)\,M^i_{\text{tr}} \quad \text{and} \quad \frac{d}{dt}M^j_{\text{tr}} = -(\lambda-\eta)\,M^j_{\text{tr}} \tag{2.8}$$

where M^i_{tr} and M^j_{tr} are the magnetizations associated with multiplet components i and j; λ and η are the auto and cross-correlated relaxation rates, respectively. The decay of net transverse magnetization is given by the sum over all multiplet components and is thus proportional to the sum of their exponentials:

$$\begin{aligned} M_{\text{tr}}(t) &= M^i_{\text{tr}}(t) + M^j_{\text{tr}}(t) \\ &= 0.5A(0)\left\{\exp\left[-(\lambda+\eta)\,t\right] + \exp\left[-(\lambda-\eta)\,t\right]\right\} \end{aligned} \tag{2.9}$$

Equation 2.9 predicts that transverse relaxation of A magnetization is *biexponential* which leads to a serious overestimation of R_2. Similar expressions have

been derived for longitudinal relaxation.[72]

For a quantitative analysis of relaxation data, it is of utmost importance to suppress these cross-correlation effects; if not, all parameters derived from this data would be erroneous. How can suppression of these effects be achieved? In a ^{15}N–^{1}H spin pair, a flip of the ^{1}H spin interconverts the multiplet components of the ^{15}N spin and thus averages the different relaxation rates of both components. If the spin flip rate is large compared to the relaxation rate of the fastest relaxing multiplet component, then cross-correlation effects are suppressed.[71] The spin flips can either be the result of random fluctuations in the local magnetic field at the site of the ^{1}H spin, or may be introduced artificially by applying 180° proton pulses at an appropriate rate, as is usually done in NMR experiments for measuring relaxation rates.

It should be noted that cross-correlated relaxation rates are not affected by chemical exchange and thus provide a means of identifying residues subject to exchange processes (see section 3.3).

2.2 NMR Experiments for Measuring Relaxation Rates

The pulse sequences described in this section are versions of published two dimensional (2D) experiments modified to achieve optimal water suppression on conventional as well as on cryogenic probes.[73] A short description of the relevant product operators is given here, while the sequences are explained in more detail below. The experiments for measuring relaxation rates are based on HSQC-type (heteronuclear single quantum correlation) experiments modified by adding a relaxation period, during which the operator of interest is allowed to relax. A schematic diagram illustrating the following building blocks is shown in figure 2.7.

1. *Preparation*. In case of the heteronuclear NOE experiment, preparation of the desired spin density operator is achieved by R.F. irradiation of the protons; if longitudinal or transverse relaxation rate constants are of inter-

est, the desired operator is created using a refocussed INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) step.[74]

2. *Relaxation*. After preparation, the desired spin density operator is allowed to relax during the relaxation period T. In most experiments, T is varied in a time-series of 2D spectra; when the heteronuclear NOE is measured, the relaxation period is omitted.

3. *Frequency labeling*. The chemical shifts of the heteronuclei are recorded to generate the indirect dimension t_1 of the 2D spectrum.

4. *Mixing* and *acquisition*. The relaxation-encoded, frequency-labeled coherence is transferred back to protons using a refocussed INEPT or semi-constant time transfer step and detected during t_2.[75,76]

The similarity of the sequences based on these bulding blocks is also illustrated in figure 2.8 by grey dashed boxes.

The evolution of the initial proton polarization during the refocussed INEPT transfer in the pulse sequences shown in figure 2.8 for measuring nitrogen R_1 and R_2 can be summarized as follows:

$$H_z \xrightarrow{\frac{\pi}{2}\,{}^{H}_{x}} -H_y \xrightarrow[2\delta J_{HN}\pi]{\pi^H_y,\,\pi^N_x} 2H_xN_z \xrightarrow{\frac{\pi}{2}\,{}^{H}_{y},\,\frac{\pi}{2}\,{}^{N}_{x}} 2H_zN_y \xrightarrow[2\delta' J_{HN}\pi]{\pi^H_y,\,\pi^N_x} N_x$$

At point *a*, the operator N_x is subsequently allowed to relax with the transverse relaxation rate constant during the delay T, or is converted into $\pm N_z$ prior to T by application of a 90° pulse, if longitudinal relaxation is of interest. After the relaxation period, the ^{15}N chemical shift is recorded starting from point *b* and

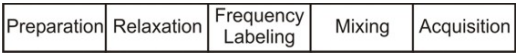


FIGURE 2.7: Block diagram of NMR experiments for measuring relaxation rates.

magnetization is transferred back to protons for detection:

$$\begin{aligned} N_x \exp(-TR_x) &\xrightarrow[2\delta' J_{HN}\pi,\, t_1]{\pi_x^H,\, \pi_x^N} 2H_z N_y \exp(-TR_x)\cos(\omega_N t_1) \\ &\xrightarrow{\frac{\pi}{2}{}_x^H,\, \frac{\pi}{2}{}_x^N} -2H_y N_z \exp(-TR_x)\cos(\omega_N t_1) \\ &\xrightarrow[2\delta J_{HN}\pi]{\pi_y^H,\, \pi_x^N} H_x \exp(-TR_x)\cos(\omega_N t_1) \end{aligned}$$

where $R_x = R_1$ or R_2. In contrast to the experiments for measuring relaxation rates, the heteronuclear NOE experiment lacks a relaxation period and starts with in-phase nitrogen magnetization.

2.2.1 R_1 Experiment

The pulse sequence used for measuring longitudinal ^{15}N relaxation rates is shown in figure 2.8. It can be briefly described as follows: The first 90° pulse on nitrogen followed by a gradient destroys all natural ^{15}N magnetization. Coherence is created on protons and converted into in-phase nitrogen coherence N_x at point *a* using a refocussed INEPT step.[74] Any residual transverse magnetization on protons is purged by the 90° pulse which also alignes the water magnetization along *z*.

The first 90° pulse on nitrogen with phase ϕ_2 generates N_z. It is important that this pulse is phase cycled in order to average longitudinal relaxation from $+N_z$ and $-N_z$. The operator $\pm N_z$ is allowed to relax during the time $T = n \cdot 2\tau + 4\tau$, where n is chosen such that the maximum relaxation delay is on the order of T_1. The 180° pulses applied on the proton channel every 3 ms suppress interference between dipole–dipole and CSA relaxation mechanisms by inversion of the ^{1}H spins.[71] The relaxation period is flanked by two gradients that dephase all unwanted magnetization. Prior to the 90° pulse on nitrogen at point *b*, the relevant magnetization is given by $N_z \exp(-T R_1)$.

Transverse nitrogen magentization N_x is generated at point *b*, which is simultaneously frequency labeled with the ^{15}N chemical shift and converted into

$2H_zN_y \cos(\omega_N t_1) \exp(-T R_1)$ anti-phase magnetization using a semi-constant time period.

At point *c*, longitudinal two-spin order $2H_zN_z$ is present and all transverse magnetization is purged by a gradient. Water-selective pulses (grey) ensure that the water magnetization is kept along *z*. The anti-phase term is transferred back to protons and refocussed to in-phase magnetization using a re-INEPT step in combination with a WATERGATE (water suppression by gradient tailored excitation) sequence that dephases any residual water magnetization.[77] Prior to detection, the magnetization is given by $H_x \cos(\omega_N t_1) \exp(-T R_1)$.

2.2.2 R_2 Experiment

The experiment used for measuring transverse ^{15}N relaxation is identical to the R_1 sequence with exception of the relaxation period (see figure 2.8). Similarly, natural nitrogen magnetization is purged by a 90° pulse followed by a gradient. Coherence is generated on protons and transferred to in-phase ^{15}N magnetization using a refocussed INEPT to yield N_x at point *a*. The 90° pulse on protons aligns the water magnetization along *z* and purges any residual *y* magnetization of other protons.

After point *a*, transverse nitrogen magnetization is allowed to relax during a CPMG (Carr Purcell Meiboom Gill) sequence for a time $T = n \cdot 16\tau$ with the maximum relaxation delay being on the order of T_2.[78,79] It is of *utmost importance* to ensure that the delay τ during the CPMG pulse train is small compared to the one-bond *J*-coupling between ^{15}N and ^{1}H ($\tau \ll {}^1J_{HN}$); otherwise, anti-phase coherence contributes to relaxation and thus renders the data unusable. 180° pulses on protons are applied every $5 - 10$ ms at the peak of a spin-echo to average the relaxation rates of the individual multiplet components and hence suppress cross-correlation effects.[71] After this period, the magnetization is given as $N_x \exp(-T R_2)$.

The use of a *z* filter at point *b* leads to improved lineshapes in the indirect dimension and allows axial peaks to be shifted to the edges of the spectrum. After

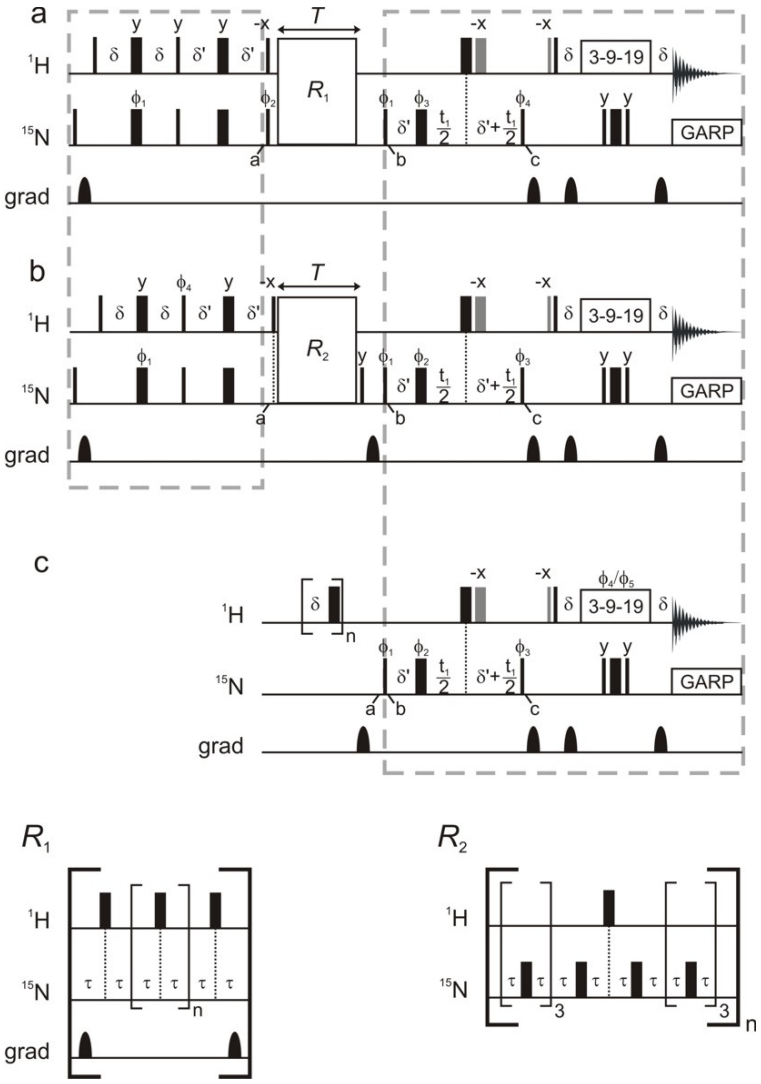
frequency labeling, anti-phase magnetization is transferred back to observable proton in-phase magnetization as described above. The relevant operator prior to detection is given by $H_x \cos(\omega_N t_1) \exp(-T R_2)$.

2.2.3 Heteronuclear NOE Experiment

Measurement of the $\{^1H\}^{15}N$ NOE is not trivial due to chemical exchange between amide and water protons.[83,84,85,86] Any error in the hetNOE will translate into errors of motional order parameters (see section 2.3) and may thus lead to misinterpretations of molecular dynamics. A comprehensive study of the "traditional" approach described here has been published.[87] Based on these results, a relaxation delay of 5 s in combination with saturation for 3 s was applied in this work. Idiyatullin *et al*. have proposed a different approach for measuring the hetNOE with improved accuracy in the presence of amide proton exchange with the solvent;[88] however, this approach is not common in the literature and has therefore not been used.

The heteronuclear NOE is calculated as the ratio of signal intensities in the NOE experiment *with* saturation of the amide protons and the signal intensities in the reference experiment *without* saturation; therefore, two experiments have to be acquired. In the NOE experiment, saturation of the amide protons is achieved using a train of 120° pulses with the carrier offset set to the center of the amide region;[81] the pulse length is chosen in order to achieve a null excitation at the water resonance. As discussed in the literature, accidential saturation of water protons must be avoided, since this would render the NOE values erroneous.[83,85,89,88] This part is replaced by a delay of equal length in the reference experiment or, more advantageously, the same pulse train is used with the carrier set off-resonance to ensure that the same amount of heat energy is transferred into the sample during both experiments. After the saturation period, a purge gradient is applied to destroy any transverse magnetization.

At point *a*, pure natural nitrogen polarization N_z is present in the case of the reference experiment; in the NOE experiment, the amount of N_z is affected


a
b
c
^{1}H
^{15}N
grad
3-9-19
GARP
R_1
R_2
T
δ
δ'
τ
n
3

by the hetNOE. Transverse ^{15}N magnetization is excited, and the remainder of the experiment is similar to the experiments described before and will thus not be explained again. It should be kept in mind that these experiments are rather insensitive, since coherence is excited directly on nitrogen and not transferred from protons; the sensitivity loss compared to a HSQC-type experiment is proportional to $(\gamma_H/\gamma_N)^{-1}$ for a ^{15}N–^{1}H spin pair.

2.2.4 Data Extraction and Error Estimation

The motional parameters that describe the internal dynamics of a protein are derived from fitting relaxation rates to spectral density functions (see section 2.3). The relaxation rates as well as hetNOE values in turn are derived from signal intensities, and are thus subject to "experimental variations". In order to assess the reliability of the fitted motional parameters, the precision of the relaxation rates has to be estimated. More detailed discussions on error analysis

---

Figure 2.8: Pulse sequences for measuring longitudinal (a), transverse (b) ^{15}N relaxation rates, and the $\{^{1}H\}^{15}N$ NOE (c). The pulse elements during the relaxation periods for measuring R_1 and R_2 are shown at the bottom. Narrow and wide bars indicate pulses with a flip angle of 90° and 180°, respectively. The grey pulses correspond to water-selective pulses with a Gaussian shape and a length of 2 ms; water suppression was achieved using a WATERGATE sequence.[77] Delays are $\delta = 2.2$ ms, $\delta' = 2.7$ ms and $\tau = 450$ µs. Decoupling during acquisition is achieved using a GARP sequence.[80] In the R_1 experiment (a), the following phase cycle is applied: $\phi_1 = 4(y), 4(-y)$; $\phi_2 = y, y, -y, -y$; $\phi_3 = x, y, -x, -y$; $\phi_4 = 8(x), 8(-x)$; $\phi_{rec} = x, 2(-x), x, -x, 2(x), -x, -x, 2(x), -x, x, 2(-x), x$. The phase cycle for the R_2 experiment (b) is $\phi_1 = 2(y), 2(-y)$; $\phi_2 = x, y, -x, -y$; $\phi_3 = 4(-x), 4(x)$; $\phi_4 = 4(y), 4(-y)$; $\phi_{rec} = x, 2(-x), x$. The phase cycle of the NOE experiment (c) is $\phi_1 = 2(y), 2(-y)$; $\phi_2 = x, y, -x, -y$; $\phi_3 = 4(-x), 4(x)$; $\phi_{rec} = x, 2(-x), x - x, 2(x), -x$. Saturation of protons is achieved using a train of 120° pulses centered in the middle of the amide region with a R.F. amplitude of 5 kHz, separated by a delay of 5 ms.[81] In the reference experiment, the pulsetrain is substituted by a delay of equal length. Gradient pulses have a sine shape and a duration of 1 ms. Gradient strengths should be optimized for best water suppression. Quadrature detection in the indirect dimension is achieved using the States method.[82]

of NMR relaxation data are given in the literature;[84,34] in this section, the error estimation used in the present work is explained briefly.

For longitudinal and transverse relaxation, the decay of signal intensities is fitted to an exponential decay:

$$I(T) = I_0 \exp(-T R_x) \tag{2.10}$$

where $I(T)$ and I_0 are the intensities of a given peak at a relaxation delay T and at $T = 0$, respectively; and $R_x = R_1$ or R_2. Both I_0 and R_x are variational parameters. Usually, several (8–12) time points per relaxation curve are used to determine a relaxation rate. In addition to these points, duplicate experiments are recorded for 2–3 relaxation delays. Using these duplicate points, experimental uncertainties of peak intensities can be estimated.[33,84,57] When the Monte-Carlo approach is applied, a large number of synthetic data sets (≈ 100) is created where random noise is added to the experimental values, i.e. to the peak intensities. This is achieved by drawing random numbers from Gaussian distributions centered on the experimental values (mean $= 0$) with standard deviations given by the experimental uncertainties. These data sets are fitted to equation 2.10 and the final reported rates are the means of the ensemble with the uncertainties given by the standard deviations of the ensemble.

The heteronuclear NOE is calculated as the ratio of signal intensities with saturation (the NOE spectrum) and without saturation (the reference spectrum) of protons:

$$NOE = \frac{I^{\text{sat}}}{I^{\text{ref}}} \tag{2.11}$$

where I^{sat} and I^{ref} are the intensities of a peak in the NOE and the reference spectrum, respectively. Two seperate sets of NOE experiments are recorded, and the final hetNOE values and uncertainties are taken to be the mean and the standard deviation of the two sets.

2.3 The Model-Free Approach

The model-free approach – in the further course of this work, abbreviated as "MF" – was introduced by *G. Lipari* and *A. Szabo* in 1982 and later extended by *G. M. Clore* and coworkers and is the most common way to analyze NMR relaxation data.[21,22,90] It allows characterization of internal motions on time scales *faster* than the overall molecular tumbling utilizing the dependence of the longitudinal and transverse relaxation rates R_1 and R_2 and the heteronuclear NOE on the spectral density function $J(\omega)$. The original approach introduces two parameters for the description of NMR relaxation data, a *generalized squared order parameter* S^2 and an *internal correlation time* τ_i. Since the spectral density function of this formalism is derived without invoking a model or any assumptions on the kind of motions and S^2 and τ_i are defined in a model-independent way, the approach is referred to as "model-free".

2.3.1 Theory

Correlation and Spectral Density Functions

Let us consider a ^{15}N–^{1}H spin pair in a protein whose overall motion can be described by a single correlation time. In contrast to section 2.1.3, the orientation of the bond vector is *not* fixed with respect to a molecular frame of reference. Rather, it changes due to internal motions. Assuming that the overall and internal motions are *independent*, the total correlation function is given as

$$C(t) = C_o(t)\, C_i(t) \tag{2.12}$$

where the indices o and i refer to overall and internal motions, respectively. It should be emphasized that the *independence of overall and internal motions* is the *fundamental assumption* of the MF-approach. Especially in proteins where large parts are involved in slow motions and thus affect the molecular shape, overall and internal motions are not independent. It has been shown that in

such cases, data from multiple static magnetic fields enable identification of these large scale motions.[91] A structural mode coupling approach with dynamical coupling between global rotational diffusion and internal motions has also been proposed;[92] in cases where the decoupling assumption cannot be made, a recently published protocol may be used to characterize internal motions on a nanosecond time scale and to determine rotational correlation times independent of the time scale of the internal motions.[93]

For isotropic overall motion, $C_o(t)$ is given by equation 2.1 with $A_C = \frac{1}{5}$. The internal correlation function can be expressed as

$$C_i(t) = S^2 + \left(1 - S^2\right) e^{-\frac{t}{\tau_i}} \tag{2.13}$$

where τ_i is the correlation time and S^2 is the squared order parameter of the internal motion. S^2 describes the spatial restriction of the motion with two limiting values as illustrated in figure 2.9. Note that in this section, the model-free formalism is introduced using a ^{15}N–^{1}H spin pair as an example. Therefore, the term *"internal motions"* refers to motions of the ^{15}N–^{1}H *bond vector* relative to a fixed molecular frame of reference. In the case $S^2 \rightarrow 1$, internal motions of the bond vector are said to be *restricted*, and relaxation is governed by *global* motion; if $S^2 \rightarrow 0$, the *unrestricted* internal motions describe the relaxation.

The squared order parameter allows a simple geometrical interpretation depending on a particular motional model. For the wobbing-in-a-cone model, S^2 is related to the semi-cone angle θ as $S^2 = [0.5 \cos\theta (1 + \cos\theta)]^2$.[94,95] Other motional models are rotation-on-a-cone and the Gaussian axial fluctuation model.[96] Inserting equations 2.1 and 2.13 into equation 2.12 yields

$$C(t) = \frac{1}{5} e^{-\frac{t}{\tau_c}} \cdot \left[S^2 + \left(1 - S^2\right) e^{-\frac{t}{\tau_i}}\right] \tag{2.14}$$

with a Fourier transformation leading to the corresponding spectral density function

$$J(\omega) = \frac{2}{5}\left[\frac{S^2 \tau_c}{1 + \omega^2 \tau_c^2} + \frac{\left(1 - S^2\right)\tau'}{1 + \omega^2 \tau'^2}\right] \tag{2.15}$$

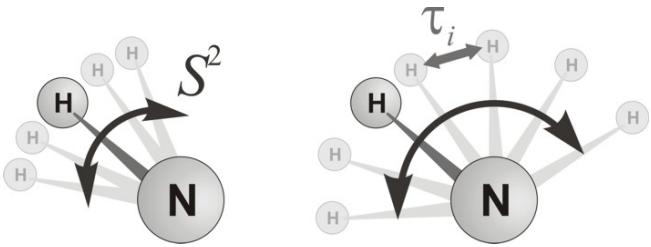
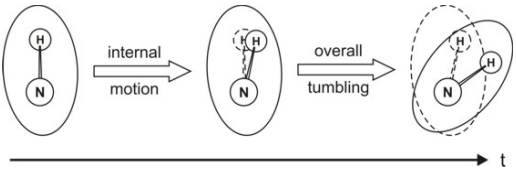


Figure 2.9: Illustration of S^2 and τ_i. S^2 describes the spatial restriction of the motion, in this case the motion of a ^{15}N–^1H bond vector. The time scale of the motion is given by τ_i. Left: Highly restricted motion, $S^2 \rightarrow 1$. Right: Largely unrestricted motion, $S^2 \rightarrow 0$.

where τ' is related to the rotational and internal correlation times according to $\tau'^{-1} = \tau_c^{-1} + \tau_i^{-1}$. When the internal motion is slow compared to overall molecular tumbling ($\tau_i \gg \tau_c$), then $\tau' \approx \tau_c$, and the spectral density is given by $J(\omega)_{global}$. In contrast, if the internal motion is faster than rotational correlation ($\tau_i \ll \tau_c$), then $\tau' \approx \tau_i$ and the spectral density function is scaled by S^2: $J(\omega) = S^2 J(\omega)_{global}$.

In the latter case, $C(t)$ rapidly decays to a plateau S^2 with a time constant τ_i due to internal motions as depicted in figure 2.10. With increasing time, global motions take over and C decays according to the overall correlation time τ_c. This is illustrated schematically below:



The ^{15}N–^1H bond vector reorients *fast* due to restricted internal motion and *slow* due to the overall tumbling.

In some cases, it is necessary to introduce an additional motion in order to explain the experimental relaxation data.[90] This is achieved in the extended

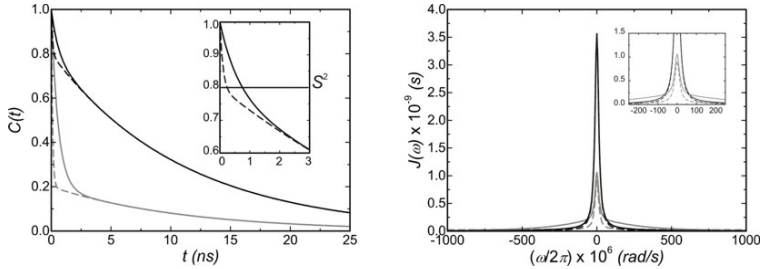


FIGURE 2.10: Model-free correlation functions (LEFT) and corresponding spectral density functions (RIGHT) for two parameter sets. Black curves $S^2 = 0.8$, grey curves $S^2 = 0.2$; solid lines $\tau_i = 0.6\,\text{ns}$, dashed lines $\tau_i = 0.1\,\text{ns}$. For all calculations, $\tau_c = 11\,\text{ns}$ was assumed.

Lipari-Szabo formalism by parametrizing the correlation function of the internal motions as

$$C_i(t) = C_f(t) \cdot C_s(t) = S^2 + \left(1 - S_f^2\right) e^{-\frac{t}{\tau_f}} + \left(S_f^2 - S^2\right) e^{-\frac{t}{\tau_s}} \tag{2.16}$$

where $S^2 = S_f^2 S_s^2$. S_f^2 and S_s^2 are the squared order parameters of the fast and slow internal motion, respectively; τ_f and τ_s are the corresponding correlation times. A simple model for the extended Lipari-Szabo formalism is illustrated in figure 2.11. The motions of ^{15}N–^{1}H bond vector are described by a two-site jump,[97] where the slower motion corresponds to a jump of the bond vector between two sites, while the faster motion is represented by free diffusion within two axially symmetric cones. If both sites are populated equally, S_s^2 is related to the angle between both cones, ϕ, as follows: $S_s^2 = [1 + 3\cos^2(\phi)]/4$.[97]

Motions described by a generalized order parameter occur on the ns – ps time scale. Hence, $\tau_f \ll \tau_s \ll \tau_c$. The full spectral density function is given by

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_c}{1+\omega^2\tau_c^2} + \frac{\left(1 - S_f^2\right)\tau_f'}{1+\omega^2\tau_f'^2} + \frac{\left(S_f^2 - S^2\right)\tau_s'}{1+\omega^2\tau_s'^2}\right] \tag{2.17}$$

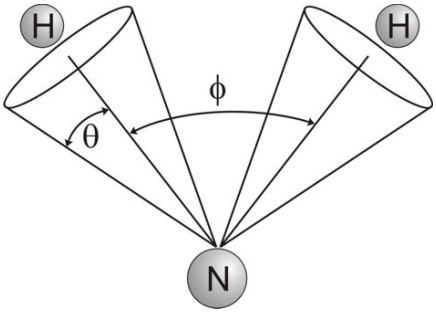


Figure 2.11: Schematic representation of the two-site-jump model for a ^{15}N–^{1}H bond vector. The semiangle of the cone is given by θ, while the angle between the cone axes is described by ϕ.

where $\tau_k' = \tau_k\tau_c/(\tau_k + \tau_c)$ and k = f or s; S^2 is defined as above. Note that for $S_f^2 = 1$ equation 2.17 reduces to equation 2.15. In the case of an axially symmetric diffusion tensor, equation 2.17 becomes

$$J(\omega) = \frac{2}{5}\sum_{j=1}^{3} A_j \left[\frac{S^2\tau_j}{1+\omega^2\tau_j^2} + \frac{\left(1-S_f^2\right)\tau_f'}{1+\omega^2\tau_f'^2} + \frac{\left(S_f^2-S^2\right)\tau_s'}{1+\omega^2\tau_s'^2}\right] \tag{2.18}$$

with similar parameter definitions as for isotropic diffusion, except that $\tau_1^{-1} = 6D_\perp$, $\tau_2^{-1} = D_\parallel + 5D_\perp$, $\tau_3^{-1} = 4D_\parallel + 2D_\perp$, $\tau_k' = \tau_j\tau_k/(\tau_j + \tau_k)$, where k is either f or s and $j = 1, 2,$ or 3; $A_1 = \left(3\cos^2\theta - 1\right)^2/4$, $A_2 = 3\sin^2\theta\cos^2\theta$, $A_3 = (3/4)\sin^4\theta$, and θ is the angle between the ^{15}N–^{1}H bond vector and the unique axis of the diffusion tensor. For the completely anisotropic case, $J(\omega)$ is given by the sum of five terms.[98]

The Diffusion Tensor

NMR relaxation depends on the tumbling of the molecules in solution. A large number of proteins studied so far have approximately spherical globular

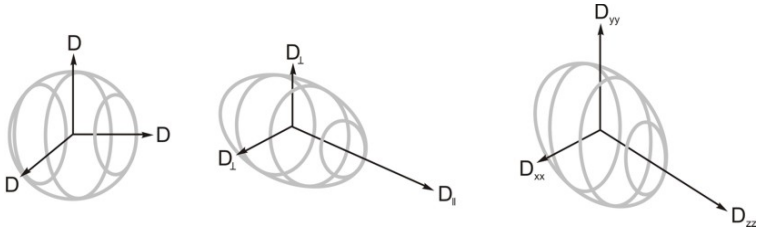


FIGURE 2.12: Illustration of rotational diffusion tensors. LEFT: For a globular spheric shape, the tumbling is isotropic. MIDDLE: Cigar-shaped molecules are characterized by an axially symmetric diffusion tensor. RIGHT: The tumbling of an asymmetric top is described by an asymmetric rotational diffusion tensor.

shapes and thus, isotropic overall rotational diffusion was assumed. However, it has been recognized early that anisotropic rotational diffusion has profound effects on spin relaxation and therefore on the interpretation of experimental NMR relaxation data.[98] Consequently, a detailed knowledge of the rotational diffusion tensor is essential for the analysis of intramolecular motions in non-spherical proteins. Furthermore, rotational diffusion anisotropy can provide additional information, e.g. on the conformation of multidomain proteins.[99]

Basically, the rotational diffusion tensor describes how a molecule "behaves" in solution, i.e. whether it tumbles as a globular sphere or something different. In the case of a globular, spheric shape, the tumbling of the protein—and hence the diffusion tensor—is *isotropic*; that is, the tumbling is the same for all directions. On the other hand, if rotational diffusion is *anisotropic*, the molecular tumbling is described by three diffusion coefficients as illustrated in figure 2.12. For anisotropic tumbling, two cases can be distinguished: if all three components have different magnitudes, the tensor is completely anisotropic; if two components are of similar size, the diffusion tensor is *axially symmetric*. In the latter case, the molecule tumbles either as a *prolate* rotor, where $D_{xx} \approx D_{yy} \approx D_\perp$ (shown in the middle of figure 2.12), or as an *oblate* rotor with $D_{yy} \approx D_{zz} \approx D_\perp$.

2.3.2 Model definitions

In order to extract the motional parameters described in the previous section, the experimental data have to be fitted against the equations defining the relaxation rates (equations 2.4–2.7) using the appropriate forms of the spectral density (equation 2.15 or 2.17), or versions of these modified to account for axially symmetric or anisotropic tumbling.

Fitting the experimental data using the spectral density function of the extended Lipari-Szabo model for isotropic tumbling requires at most six parameters: the five parameters given in equation 2.17 and R_{ex} (see chapter 3). In principle, all six parameters can be fitted if enough data are available. In most cases, only three experimental parameters are available: the longitudinal and transverse relaxation rates and the heteronuclear NOE. Hence, a maximum of four parameters can be extracted from this data (see section 2.3.3). An approach established in the literature over several years uses five different models with a maximum number of three parameters plus the overall rotational diffusion tensor. These models are described in the following; an overview is given in table 2.2.

Model 1 and 3 Model 1 is the most simple model of all and requires only one parameter: the squared order parameter S^2. For this model, the internal motions are assumed to be very fast, with the correlation time for the internal motion $\tau_i \ll \tau_c$, and the spectral density function is given by equation 2.19. In the case of chemical exchange as an additional source of relaxation, R_{ex} is introduced as second fit parameter in model 3.

Model 2 and 4 Model 2 is sometimes referred to as the "classical" Lipari-Szabo. Here, τ_i is relaxation active and the spectral density function is defined by equation 2.15. Again, R_{ex} is introduced in the case of chemical exchange to yield model 4.

TABLE 2.2: Models and parameters for fitting relaxation data to the model-free spectral densitiy functions.[a]

Model	fitted parameter(s)
1	S^2
2	S^2, $\tau_i = \tau_f$
3	S^2, R_{ex}
4	S^2, $\tau_i = \tau_f$, R_{ex}
5	S_f^2, S^2, $\tau_i = \tau_s$

[a] The overall rotational diffusion tensor is fitted in addition.

Model 5 The extended Lipari-Szabo model includes a fast and a slow internal motion with their internal correlation times differing by at least one order of magnitude; the spectral density function is given by equation 2.17. The motions can be described by diffusion- or wobbing-in-a-cone as fast and a two-site jump as slow motion.[97]

2.3.3 Data Analysis

Estimation of the Overall Rotational Diffusion Tensor

Before fitting of the relaxation data to the Lipari-Szabo spectral density functions can be performed, the rotational diffusion tensor has to be estimated. Since all relaxation rates calculated during the fitting process depend on the diffusion tensor, this estimation should be as precise as possible. Two methods for determining the diffusion tensor have been established in the literaure: analysis of local diffusion coefficients using local correlation times,[99] or direct fitting of the R_2/R_1 ratios for ^{15}N–^{1}H bond vectors with highly restricted internal motions.[100,101]

If internal motions are very fast ($\tau_i \ll \tau_c$), and R_2 is not enhanced by chemical exchange, $J(\omega)$ becomes independent of τ_i:[102, 100, 101]

$$J(\omega) = \frac{2}{5} \frac{S^2 \tau_c}{\left(1 + \omega^2 \tau_c^2\right)} \tag{2.19}$$

which represents a simplified form of equation 2.15. Hence, R_2/R_1 is independent of S^2, and the parameters describing the diffusion tensor can be extracted. A detailed discussion of determining the rotational diffusion tensor using the R_2/R_1 ratio is given by Lee *et al.*[103]

The effect of the bond vector reorientation with respect to the rotational diffusion tensor can be explained for the axially symmetric case as follows. Rotation around the long axis of the tensor is faster than rotation around a perpendicular axis (see section 2.3.1). Transverse relaxation therefore depends on the orientation of the bond vector in the diffusion frame. ^{15}N–^{1}H vectors aligned parallel to the long axis of the diffusion tensor are not reoriented by rotations around this axis and are thus characterized by faster transverse relaxation.

Consider a protein with axially symmetric rotational diffusion and three helices as illustrated in figure 2.13. Helix B is aligned parallel to the long axis; hence, the bond vectors have a faster transverse relaxation compared to those in helices A and C. Rotational diffusion anisotropy is clearly evident from the plot shown in figure 2.13: all vectors oriented approximately parallel to the long axis of the tensor (helix B) have higher R_2/R_1 ratios due to a slower reorientation of their ^{15}N–^{1}H bond vectors.

As mentioned above, care has to be taken to exclude residues subject to chemical exchange (see section 3.3), since for these residues, transverse relaxation is enhanced as well. A method for the discrimination between motional anisotropy and chemical exchange using the product R_2R_1 instead of the ratio R_2/R_1 has been proposed recently.[104]

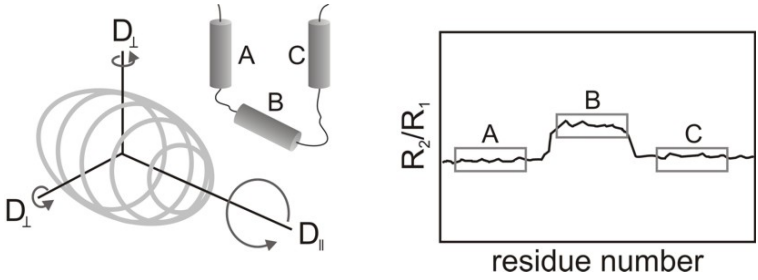


FIGURE 2.13: Estimation of the diffusion tensor anisotropy using the R_2/R_1 ratio. Residues with ^{15}N–^{1}H bond vectors oriented parallel to $D_{\parallel}$ (helix B, LEFT) are readily identified in a plot of R_2/R_1 *vs*. the residue number (RIGHT).

Model Selection, Diffusion Tensor Optimization and Error Analysis

Fitting of relaxation data according to the MF-approach is most commonly performed using TENSOR or MODELFREE,[31,32,33] the latter being used in this work. Generally, model selection is based on statistical methods such as Monte-Carlo simulations and *F*-Tests. A detailed description of these methods is beyond the scope of this work; the interested reader is referred to statistical textbooks. Briefly, the fitting procedure can be described as follows. Optimization of parameters is generally achieved by minimizing a *target* or *weighting function*. The target function is given by the *sum squared error* Γ:

$$\Gamma(N,n) = \frac{\left(par_{\text{exp}} - par_{\text{back}}\right)^2}{\sigma^2_{par_{\text{exp}}}} \tag{2.20}$$

where par_{exp} and par_{back} are the experimental and back-calculated values (also referred to as the fitted values) of the parameter *par*, respectively; $\sigma_{par_{\text{exp}}}$ is the experimental uncertainty of par_{exp}; N is the total number of data points and n is the number of parameters in the fitting model. The back-calculated parameters are changed such as to minimize Γ. If Γ is less than the *α-critical value* of a

simulated distribution of Γ obtained from Monte-Carlo simulations centered on par_{back}, the fit is said to be satisfactory and the fitting model is accepted.

During the MF-analysis of relaxation data, more than one parameter has to be optimized. The appropriate weighting function to be minimized, χ^2, can be written as the sum of the Γ of the individual relaxation rates:

$$\begin{aligned}\chi^2 &= \sum_{i=1}^{N} \Gamma(i) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{M}\left\{ \frac{\left(R_{1ij} - \overline{R}_{1ij}\right)^2}{\sigma^2_{R_{1ij}}} + \frac{\left(R_{2ij} - \overline{R}_{2ij}\right)^2}{\sigma^2_{R_{2ij}}} + \frac{\left(NOE_{ij} - \overline{NOE}_{ij}\right)^2}{\sigma^2_{NOE_{ij}}} \right\}\end{aligned} \tag{2.21}$$

in which R_{1ij}, R_{2ij}, NOE_{ij} are the experimental relaxation parameters for the ith spin and jth static magnetic field; $\overline{R}_{1ij}$, $\overline{R}_{2ij}$, $\overline{NOE}_{ij}$ are the corresponding fitted or back-calculated values, and $\sigma^2_{R_{1ij}}$, $\sigma^2_{R_{2ij}}$, $\sigma^2_{NOE_{ij}}$ are the experimental uncertainties of the relaxation parameters. The total number of spins to be analyzed is given by N, and M is the number of static magnetic fields for which relaxation data is available.

If two models with a different number of parameters are to be compared, the *F-Test* is commonly used:

$$F = \frac{(N-n)\left[\Gamma_{(N,m)} - \Gamma_{(N,n)}\right]}{(n-m)\,\Gamma_{(N,n)}} \tag{2.22}$$

In this equation, n and m are the number of parameters in the complicated and simple model, respectively. The improvement in Γ by using a more complicated model is statistically significant—i.e. the model is accepted—if F is larger than the α-critical value of a simulated distribution.

Uncertainties in the fitted parameters are usually estimated by Monte-Carlo simulations. A large number of synthetic data sets is created ($\approx$ 300–500), where the values of relaxation rate constants are obtained by adding a random

noise to either the experimental or back-calculated values of the relaxation rates. The noise terms are taken from Gaussian distributions with a mean equal to zero and standard deviations given by the experimental uncertainties. These data sets are then fitted to the MF spectral density functions, and the final reported values and uncertainties of the relaxation rate constants are the mean and standard deviations of the ensemble.

In the initial stage of data fitting, the parameters describing the diffusion tensor are kept constant at their initial guesses. Relaxation data are fitted against model 1 for all spins. Those residues with Γ_i smaller than the α-critical value are assigned to model 1. Next, all residues for which model 1 could not reproduce the data satisfactory are fitted against models with two variable parameters. At this stage, model selection is based on *F*-statistics. Residues neither assigned to model 2 nor 3, but with a Γ_i slightly larger than the cut-off for model 1 are assigned to model 1 without testing against models 4 and 5. Spins that are still not assigned to one of the models 1–3 are tested against models 4 and 5. Usually, data from only one static magnetic field is available; hence, *F*-testing cannot be performed, and models 4 or 5 are accepted only if $\Gamma_i = 0$. If data from more than one static fields is available, *F*-statistics can be performed for models 4 and 5 as well. This process of model selection has been established in the literature and is illustrated in figure 2.14.[32]

After the model selection, the overall diffusion tensor is optimized together with the motional parameters of each spin. However, since model selection is based on initial guesses of the diffusion tensor, there is no guarantee that the model assignment indeed represents the best fit. To overcome this issue, the optimized diffusion tensor parameters are used as initial guesses for a repetition of the model assignment. If the optimum diffusion tensor parameters had been used as initial guesses, the correct models would have been assigned for each spin and optimization of the diffusion tensor would yield the same parameters. Thus, by iterating model assignment and diffusion tensor optimization, it is possible to converge to a global minimum for both the internal motional parameters for each spin as well as for the global diffusion tensor.

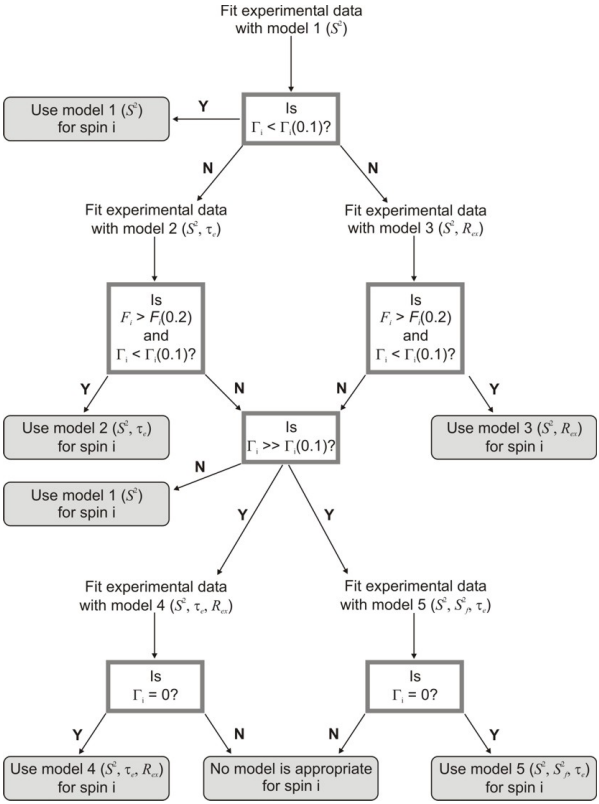


Figure 2.14: Model selection process during the model-free analysis of NMR relaxation data for R_1, R_2, and heteronuclear NOE data acquired at a single B_0 as proposed by Mandel *et al*.[32]

Chapter 3

Slow Internal Motions

The model-free approach discussed in the previous chapter is capable of characterizing internal dynamics that occur on time scales faster than overall molecular tumbling. However, dynamic processes on time scales slower than overall tumbling are important for the functions of proteins, including recognition, allostery and catalysis.[12, 105] As illustrated in figure 2.1, motions on a ms to µs time scale modulate the magnetic environment of a nucleus and cause lineshape perturbations.

In general, there are two different methods of determining the characteristic parameters. (1) Analysis of lineshapes as function of temperature.[45, 106, 107, 108] This method is, from an experimental point of view, the most simple, because it only requires one dimensional lineshapes. However, the need to sample a wide range of temperatures and difficulties arising from signal overlap in larger molecules limit this method to the characterization of ms–µs time scale dynamics of small, highly stable proteins. (2) Analysis of the exchange contribution R_{ex} to the transverse relaxation rate in the laboratory or rotating frame as a

function of the effective field strength ω_{eff}; the transverse relaxation rates can be measured using a spin-lock experiment ($R_{1\rho}$) or a Carr–Purcell–Meiboom–Gill (CPMG) sequence (R_2).[49,50] This method usually requires the acquisition of two dimensional spectra at only one temperature. Thus, it is well suited for characterizing slow dynamics even in larger, less temperature-stable proteins. Nevertheless, measuring R_{ex} as a function of the effective field strength also has drawbacks. The maximum effective field strength is limited by sample heating and amplifier duty cycles. Conventional $R_{1\rho}$ experiments can only be performed for spins with no or only small offsets with respect to the transmitter frequency. Recently, a method has been proposed that allows to map chemical exchange in proteins with high molecular weights (> 50 kDa).[109]

In this work, CPMG spin echo experiments were used to analyse ms–µs time scale dynamics of an 18 kDa protein (see chapter 6) and are therefore discussed in detail. It should be mentioned, however, that analysis of relaxation rates derived from spin-lock experiments is very similar.[110,111,112]

3.1 Two-Site Chemical Exchange

Chemical exchange processes can involve a number of different states, rendering the mathematical description complicated. Although a system with at least three different species has been characterized recently,[55] a two-site chemical exchange process is considered in the following. Exchange between two sites can be described as

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B \tag{3.1}$$

in which the exchange rate constant k_{ex} is defined as

$$k_{\text{ex}} = k_1 + k_{-1} = \frac{k_1}{p_{\text{b}}} = \frac{k_{-1}}{p_{\text{a}}} \tag{3.2}$$

where p_{a} and p_{b} are the equilibrium populations of site A and B, respectively; $p_{\text{a}} + p_{\text{b}} = 1$; k_1 is the forward first-order kinetic rate constant and k_{-1} is the

reverse first-order kinetic rate constant.[113] Site A and B are assumed to have distinct chemical shifts ω_a and ω_b. The frequency difference between the chemical shifts of the two sites is $\Delta\omega = |\omega_a - \omega_b| = |\gamma\Delta\sigma B_0|$, in which γ is the gyromagnetic ratio of the exchanging spins, $\Delta\sigma$ is the difference in chemical shielding of the two sites, and B_0 is the static magnetic field strength. The transverse relaxation rates for sites A and B are denoted R_a and R_b, respectively.

Chemical exchange in general leads to a contribution to transverse relaxation. The exact effect depends critically on the *exchange regime* of the process with respect to the NMR chemical shift time scale given by $\Delta\omega$:

slow exchange	$k_{ex}/\Delta\omega < 1$
intermediate exchange	$k_{ex}/\Delta\omega \approx 1$
fast exchange	$k_{ex}/\Delta\omega > 1$

In the regime of slow exchange, two distinct signals are observed at resonance frequencies ω_a and ω_b, respectively, with the intensities being proportional to the relative populations of the sites, if transverse relaxation is equal for both sites. When the exchange process approaches the intermediate regime, the signals coalesce. Fast exchange is characterized by the presence of a single resonance with the averaged frequency $\omega_a p_a + \omega_b p_b$. This behaviour is illustrated in figure 3.1.

Unfortunately, the site populations are highly unequal in many cases of interest.[114] If, for example, site A is more stable than site B by only $2\,k_B T$, then $p_a = 0.88$ and $p_b = 0.12$. In the limit of slow exchange, the resonance at ω_b is of lower intensity and significantly broader compared to the resonance at ω_a and may thus be undetectable. Consequently, observation of a single exchange-broadened line does *not* indicate that the exchange process is fast on the NMR chemical shift time scale.

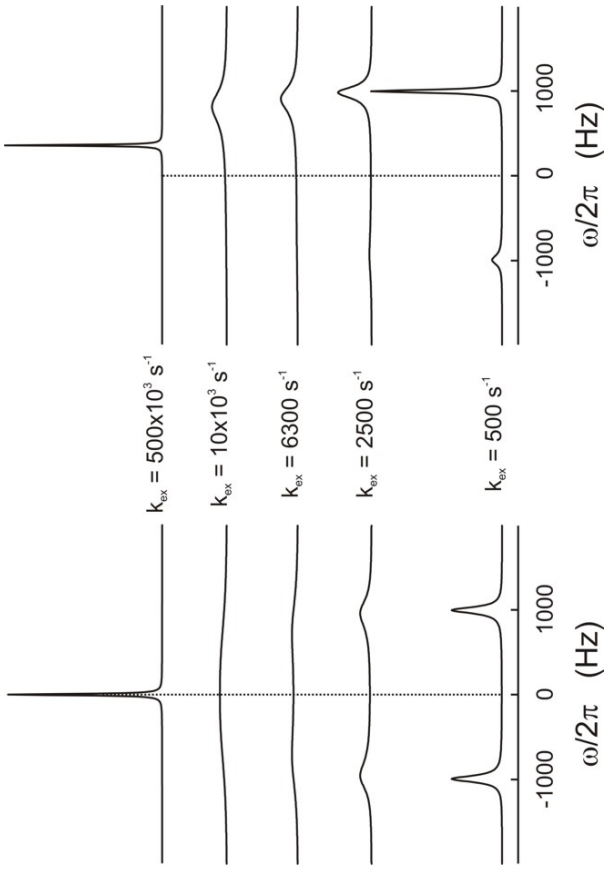


FIGURE 3.1: Calculated spectra for a two-site exchange. LEFT: Symmetric exchange with equally populated sites ($p_a = 0.5 = p_b$). RIGHT: Asymmetric exchange with skewed populations ($p_a = 0.8$ and $p_b = 0.2$). Calculations were performed using $\omega_a/2\pi = +1\,\text{kHz}$ and $\omega_b/2\pi = -1\,\text{kHz}$; the transverse relaxation rate constant was assumed to be equal for both sites ($R_a = R_b = 10\,\text{Hz}$).

3.2 Transverse Relaxation and Chemical Exchange

The contribution of a chemical exchange process to transverse relaxation, R_{ex}, can be measured using a CPMG spin-echo experiment. A general expression of $R_2(1/\tau_{cp})$, the phenomenological transverse relaxation rate of site A, as a function of the delay τ_{cp} between two successive 180° pulses, is given by:[115,116,113]

$$R_2\left(1/\tau_{cp}\right) = 0.5\left(R_a + R_b + k_{ex} - \frac{1}{\tau_{cp}}\cosh^{-1}\left[D_+\cosh\left(\eta_+\right) - D_-\cos\left(\eta_-\right)\right]\right) \tag{3.3}$$

where

$$D_\pm = 0.5\left[\pm 1 + \frac{\psi + 2\Delta\omega^2}{\left(\psi^2 + \zeta^2\right)^{1/2}}\right] \tag{3.4}$$

$$\eta_\pm = \frac{\tau_{cp}}{\sqrt{2}\left[\pm\psi + \left(\psi^2 + \zeta^2\right)^{1/2}\right]^{1/2}} \tag{3.5}$$

$$\psi = \left(R_a - R_b - p_a k_{ex} + p_b k_{ex}\right)^2 - \Delta\omega^2 + 4p_a p_b k_{ex}^2 \tag{3.6}$$

$$\zeta = 2\Delta\omega\left(R_a - R_b - p_a k_{ex} + p_b k_{ex}\right) \tag{3.7}$$

The variation of $R_2(1/\tau_{cp})$ as a function of $1/\tau_{cp}$ is called *relaxation dispersion*. If $R_2(1/\tau_{cp})$ is plotted *vs.* $1/\tau_{cp}$, relaxation dispersion profiles as illustrated in figure 3.2 are obtained for spins undergoing chemical exchange. In the absence of exchange, the transverse relaxation rate is independent of τ_{cp} and flat lines are obtained in the plot.

The contribution of a chemical exchange process to the transverse relaxation rate constant is defined by the difference between the apparent relaxation rates

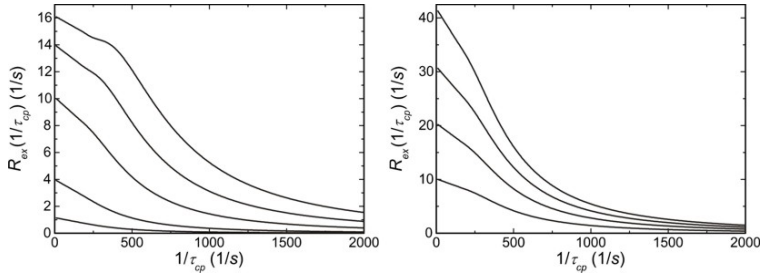


FIGURE 3.2: Dispersion profiles calculated using equation 3.3. For both panels, $k_{ex} = 1000\ s^{-1}$, and the intrinsic transverse relaxation rates for both sites were assumed to be equal and set to zero ($R_a = R_b = 0$). LEFT: The profiles were calculated for $p_a = 0.98$ and $\Delta\omega = 250$, 500, 1000, 1500, and 2000 rad $\cdot\ s^{-1}$ (from bottom to top). RIGHT: Effect of p_a on R_{ex}. Calculations were performed assuming $k_{ex} = 1000\ s^{-1}$, and $\Delta\omega = 1000$ rad $\cdot\ s^{-1}$. The population of site A was varied from $p_a = 0.92$, 0.94, 0.96 to 0.98 from top to bottom.

in the limit of slow and fast pulsing. For site A, R_{ex} is given by:[113]

$$R_{ex} = \Delta R_2\left(0,\infty\right) = R_2\left(1/\tau_{cp} \to 0\right) - R_2\left(1/\tau_{cp} \to \infty\right) = 0.5\left\{\left(\psi + \Delta\omega^2\right)^{1/2} - \frac{1}{\sqrt{2}}\left[\psi + \left(\psi^2 + \zeta^2\right)^{1/2}\right]^{1/2}\right\} \quad (3.8)$$

Provided that the relative site populations are highly skewed, $p_a \gg p_b$, and assuming $R_a = R_b$, equation 3.8 can be simplified to:

$$R_{ex} \approx \frac{p_a p_b k_{ex}}{1 + \left(k_{ex}/\Delta\omega\right)^2} \quad (3.9)$$

It should be noted that R_{ex} in the context of relaxation dispersion is *not* equivalent to the R_{ex} term used in the MF approach. R_{ex} is explicitly defined (equation 3.8) and measured as a function of τ_{cp} in the former case, whereas in the case of model-free analysis of relaxation data, an R_{ex} contribution is introduced when the back-calculated R_2 is smaller than the experimentally observed value.

The source of relaxation dispersion becomes evident from the dependence of R_{ex} on $\Delta\omega$, as given by equation 3.8. Consider the key element of a CPMG pulse train represented by $\tau_{cp}/2 - \pi - \tau_{cp} - \pi - \tau_{cp}/2$, where π corresponds to a pulse with a flip angle of 180°.[78,79,50] Application of this sequence to transverse magnetization refocusses the chemical shift evolution. The degree of refocussing depends on the pulse spacing and thus on τ_{cp}: in the slow pulsing limit ($1/\tau_{cp} \to 0$), the chemical shift evolution is not refocussed and the full R_{ex} contribution is retained. In contrast, the chemical shift evolution is refocussed in the fast pulsing limit ($1/\tau_{cp} \to \infty$), and R_{ex} tends to zero. If $R_2(1/\tau_{cp})$ is sampled for a range of τ_{cp} values, the relaxation dispersion profiles illustrated in figure 3.2 are obtained.

Equation 3.3 is, however, very complex and the dependence of $R_2(1/\tau_{cp})$ on τ_{cp} is not obvious. A simple function valid for all time scales has been introduced by Ishima and Torchia.[110] If the populations are highly skewed ($p_a \gg p_b$), the effective transverse relaxation rate of the higher populated site is given by

$$R_2\left(1/\tau_{cp}\right) = R_2\left(1/\tau_{cp} \to \infty\right) + p_a p_b \Delta\omega^2 k_{ex} / \left[k_{ex}^2 + \left(p_a^2 \Delta\omega^4 + 144/\tau_{cp}^4\right)^{1/2}\right] \tag{3.10}$$

In the limit of fast exchange, equation 3.3 reduces to the fast-exchange formula:[50]

$$R_2\left(1/\tau_{cp}\right) = R_2\left(1/\tau_{cp} \to \infty\right) + R_{ex}\left[1 - \frac{2}{k_{ex}\tau_{cp}} \tanh\left(k_{ex}\tau_{cp}/2\right)\right] \tag{3.11}$$

The values of R_{ex} in the limits of very slow and very fast exchange were shown to be:[117,67,118,110]

$$R_{ex} = p_b k_{ex} \qquad (k_{ex}/\Delta\omega \to 0) \tag{3.12}$$

$$R_{ex} = p_a p_b \Delta\omega^2 / k_{ex} \qquad (k_{ex}/\Delta\omega \to \infty) \tag{3.13}$$

3.3 Identification of Chemical Exchange

Identifying residues subject to chemical exchange is usually the first step in characterizing any kinetic process. Provided that p_b is large and exchange is slow, two resonances are observed for each site. These are most commonly identified during sequence specific residue assignment and confirmed using a N_z experiment.[119] In cases where exchange is in the intermediate or fast regime or if $p_a \gg p_b$, only one resonance with exchange linebroadening is observed.

How can be decided from a single resonance whether or not the spin is subject to chemical exchange? Several methods have been proposed that aim at the determination of R_2^0, the intrinsic transverse relaxation rate in the *absence* of chemical exchange (see ref.[42] and references cited therein). Once R_2^0 is known, exchanging residues can be identified through the relation $R_2 = R_2^0 + R_{ex}$. R_2^0 can be determined using the model-free formalism (see section 2.3), the magnetic field dependence of R_2 to determine $J(0)$,[67,118,120] or dipole–dipole/CSA relaxation interference rate constants that are not affected by chemical exchange (see subsection 2.1.7).[120,121] However, all these methods require the analysis of several relaxation rates.

Three simple, yet robust methods for the identification of chemical exchange using only data from HSQC or CPMG dispersion experiments are listed below.

1. In the most simple case, the lineshape of a signal indicates exchange broadening. During the assignment process, residues displaying signals with extreme broad linewidth(s) in one (two) dimension(s) are readily identified.

2. A second approach becomes obvious when considering the dependence of R_{ex} on τ_{cp}. Two spectra using the pulse sequence shown in figure 3.4 are recorded using a small and a large value of τ_{cp}. A difference in the apparent relaxation rate constants indicates chemical exchange (see equation 3.8 and figure 3.2).

3. When complete dispersion profiles have been recorded, the data are fitted to a flat line and to the function for fast exchange. If, based on statistical methods, the fast exchange equation yields a significant improvement over the flat line, then the residue is subject to chemical exchange.

3.4 Determination of the Exchange Regime

Plots of R_{ex} as a function of k_{ex} calculated using equation 3.8 and assuming $R_a = R_b$ are presented on the left panel of figure 3.3. As is clearly evident, a given value of R_{ex} can arise from a chemical exchange process that is either fast or slow on the NMR chemical shift time scale. Therefore, a single measured value of R_{ex} does not suffice to ascertain the exchange regime. In some cases, the temperature dependence of R_{ex} can be used to determine the relevant exchange regime, since increasing the temperature will also increase $k_{ex}/\Delta\omega$ through the Arrhenius equation.[122] However, changes in temperature also affect the site populations through the Boltzmann equation, might affect $\Delta\omega$ through temperature-dependent conformational changes and can hamper the investigation of temperature-labile proteins.

A better method to determine the relevant exchange regime is to analyze the dependence of R_{ex} on the static magnetic field.[123] The fractional change in chemical exchange broadening, $\partial R_{ex}/R_{ex}$, is given for small changes in the static magnetic field, $\partial B_0/B_0$, as

$$\frac{\partial R_{ex}}{R_{ex}} = \alpha \frac{\partial B_0}{B_0} \tag{3.14}$$

where the scaling factor α is defined as

$$\alpha = \frac{\mathrm{d}\ \ln R_{ex}}{\mathrm{d}\ \ln \Delta\omega} \tag{3.15}$$

For highly skewed populations ($p_a \gg p_b$), α depends only on $k_{ex}/\Delta\omega$ and therefore defines the chemical shift time scale:

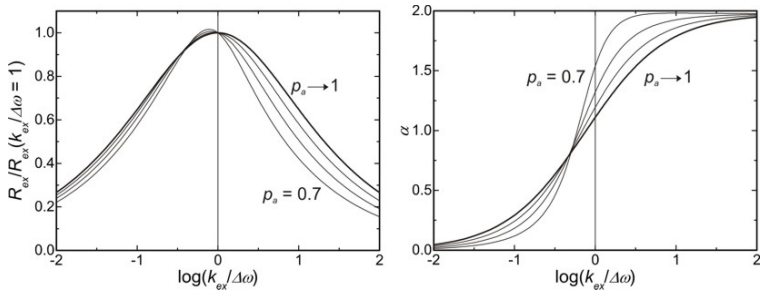


FIGURE 3.3: LEFT: Plots of R_{ex} as a function of the exchange rate constant k_{ex} calculated using equation 3.8. Values of R_{ex} are normalized by the value of R_{ex} at coalescence ($k_{ex}/\Delta\omega = 1$). The x-axis in both panels is normalized by $\Delta\omega$. RIGHT: Plots of α as a funtion of k_{ex} calculated using equation 3.16. Values of p_a range from $p_a = 0.7$ to $p_a \rightarrow 1$ (heavy lines), $\Delta\omega$ was assumed to be 6280 rad s^{-1} (1 kHz), and $R_a = R_b$.

$$\begin{aligned} 0 \leq \alpha < 1 \quad & \text{slow exchange} \\ \alpha = 1 \quad & \text{intermediate exchange} \\ 1 < \alpha \leq 2 \quad & \text{fast exchange} \end{aligned}$$

A numerical approximation of equation 3.15 is given by

$$\alpha = \left(\frac{B_{02} + B_{01}}{B_{02} - B_{01}}\right)\left(\frac{R_{ex2} - R_{ex1}}{R_{ex2} + R_{ex1}}\right), \tag{3.16}$$

in which R_{ex1} and R_{ex2} correspond to the lower field, B_{01}, and the higher field, B_{02}, respectively. Using equation 3.16 in combination with dispersion data from two static magnetic fields defines the exchange regime unambiguously, as illustrated on the right panel of figure 3.3.

3.5 The Constant Relaxation Time CPMG Experiment

Figure 3.4 shows the NMR experiment used for acquisition of relaxation dispersion data. It is based on published schemes with modifications to ensure optimal water suppression on cryogenic probes.[54,53] The evolution of magne-

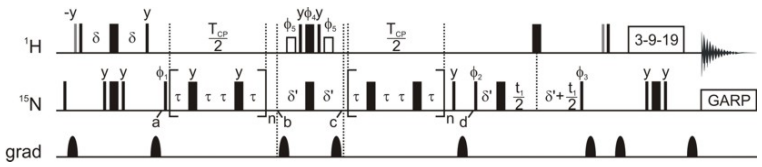


Figure 3.4: Pulse sequence for measuring CPMG dispersion profiles.[54,53] Narrow and wide bars indicate pulses with a flip angle of 90° and 180°, respectively. The open boxes represent 1.6 ms water selective 90° proton pulses to ensure that water magnetization is kept along z during both relaxation periods. The grey pulses correspond to water-selective pulses with a Gaussian shape and a length of 2 ms; water suppression was achieved using a WATERGATE sequence.[77] Delays are $\delta = 2.2\,\text{ms}$, $\delta' = 2.7\,\text{ms}$ and $T_{\text{CP}} = 80\,\text{ms}$. Decoupling during acquisition is achieved using a GARP sequence.[80] The phase cycle is $\phi_1 = x, -x$; $\phi_2 = 2(y), 2(-y)$; $\phi_3 = 4(x), 4(-x)$; $\phi_4 = 2(x), 2(-x)$; $\phi_5 = \phi_4$; $\phi_{\text{rec}} = x, 2(-x), x, -x, 2(x), -x$. Gradient pulses have a sine shape and a duration of 1 ms. Gradient strengths were optimized for best water suppression. Quadrature detection in the indirect dimension is achieved using the States method.[82] Note that $\tau_{\text{cp}} = 2\tau$.

tization during the experiment is briefly described in the following. Transverse proton magnetization is created and converted into longitudinal two spin order $2H_zN_z$ at point a. A 90° pulse applied to ^{15}N creates anti-phase coherence $-2H_zN_y$, which is allowed to relax during the delay $0.5 \cdot T_{\text{CP}}$. A CPMG pulse train is used; however, no attempts to suppress cross-correlation effects are made at this point. Between b and c, anti-phase coherence $-2H_zN_y$ is converted into in-phase N_x coherence. The composite 180° pulse on protons in the middle between b and c suppresses dipole–dipole/CSA relaxation mechanism interferences. During the second relaxation delay $0.5 \cdot T_{\text{CP}}$, N_x is allowed to relax. After this period, a z filter is applied and starting from point d, the remainder of the sequence is similar to those discussed in section 2.2.

The apparent transverse relaxation rate constant measured with any CPMG experiment is given by

$$R_2\left(1/\tau_{\text{cp}}\right) = \epsilon R_{\text{in}} + (1-\epsilon)R_{\text{anti}} + R_{\text{ex}} \tag{3.17}$$

where R_{in} and R_{anti} are the relaxation rate constants for in-phase and anti-phase coherences averaged over each state, and $0 \leq \epsilon \leq 1$ reflects the averaging due to the evolution of scalar couplings during τ_{cp}.[124,125,71] Note that τ_{cp} is equal to 2τ in figure 3.4. When fast internal motions are to be characterized by the MF-formalism, the transverse relaxation rate of in-phase nitrogen magnetization is of interest. In the R_2 sequence (see figure 2.8), τ is usually on the order of $400 - 500\,\mu s$, and thus, $\tau_{cp} < (4\,J_{NH})^{-1}$. In this case, $\epsilon \approx 1$ and the measured relaxation rate constant is associated with $N_{x/y}$.[126] Broadband proton decoupling during T_{CP} renders $\epsilon = 1$, but accentuates the decay of transverse magnetization.[125,71,127] The sequence shown in figure 3.4 explicitly averages in-phase and anti-phase magnetization and ensures that $\epsilon = 0.5$ for all values of τ_{cp}.[53] Thus, equation 3.17 can be rewritten as

$$R_2\left(1/\tau_{cp}\right) = \overline{R} + R_{ex} \tag{3.18}$$

where $\overline{R} = 0.5\,(R_{in} + R_{anti})$.

If exchange contributions to transverse relaxation are to be measured using a non-constant relaxation time experiment, complete decay curves have to be recorded for each τ_{cp} and have to be fitted to an exponential decay to yield the effective transverse relaxation rates R_2^{eff}.[111,53] This procedure translates into a large amount of instrument time. In contrast, in the constant relaxation time experiment, a certain number of ^{15}N refocussing pulses is used between points *a* to *b* and *c* to *d*. Different effective field strengths are created by variation of the number of refocussing pulses and thus τ_{cp}, while the net relaxation delay T_{CP} remains constant. The signal intensities are converted into effective transverse relaxation rates *via*

$$R_2^{eff}\left(1/\tau_{cp}\right) = -\frac{1}{T_{CP}} \ln\left[\frac{I\left(1/\tau_{cp}\right)}{I_0}\right] \tag{3.19}$$

where $I\left(1/\tau_{cp}\right)$ and I_0 are the intensities of corresponding signals in a relaxation and a reference experiment, respectively; and T_{CP} is the duration of the

constant relaxation delay.[128] For the reference spectrum, both CPMG periods (*a* to *b* and *c* to *d*) are omitted. A plot of R_2^{eff} *vs.* $1/\tau_{\text{cp}}$ yields relaxation profiles as illustrated in figure 3.2. Using the constant relaxation time experiment, 10–15 experiments per profile suffice; in contrast, typically 50 experiments have to be recorded using the non-constant relaxation time approach.

3.6 Extracting the Exchange Parameters

The exchange parameters are obtained by non-linear curve fitting of the dispersion profiles to the fast exchange equation (equation 3.11) or the general expression for a two-site chemical exchange (equation 3.3). As discussed in the previous section, data from at least two different static magnetic fields are necessary to define the correct exchange regime. Extracting the data for every B_0 field individually can lead to erroneous interpretations, as was also mentioned before. An easy and straightforward way to overcome this problem is to fit both data sets *simultaneously* to both fields in a so-called *global* fit. Recall the parameters describing the apparent transverse relaxation rate for site A in a system exchanging between two sites: the relative site populations p_{a} and p_{b}, the difference in the resonance frequencies $\Delta\omega$, and the exchange rate constant k_{ex}. The only parameter depending on the static magnetic field is $\Delta\omega$, and $\Delta\omega$ can be easily scaled according to $\Delta\omega_{02} = B_{02}/B_{01} \cdot \Delta\omega_{01}$. The target function to be minimized, χ^2_{global}, is given as the sum of the χ^2 for each static field:

$$\chi^2_{\text{global}} = \chi^2_{B_{01}} + \chi^2_{B_{02}} \tag{3.20}$$

The advantage of using global fits in the analysis of dispersion data is illustrated in figure 3.5 for Val423, a residue in the nucleotide binding domain of the P-type ATPase Kdp (see chapter 6). When both profiles are fitted independently, excellent fits are obtained using equation 3.11 at 600 and 3.3 at 750 MHz with reduced χ^2_{red} values of 0.36 and 0.33, respectively. These results lead to the conclusion that the exchange process is fast on the NMR chemical shift time

scale at 600 MHz, but in the slow to intermediate exchange regime at 750 MHz. When the fast-exchange equation is used for a global analysis, it becomes clear that the process is outside the fast exchange limit; hence, the general expression for a two site exchange process has to be used.

A common approach to fit relaxation dispersion data can be summarized as follows:[54,56,57] Each profile is fitted to a flat line and to the fast-exchange equation. *F*-statistics are used to assess whether statistical improvements of the fits were obtained by using the fast-exchange equation. Residues passing the *F*-test are involved in a chemical exchange process and are subsequently fitted to the general expression. Again, *F*-statistics are used to decide which model describes the data best. Note that *global fits* are used in all cases. Uncertainties in the fitted parameters are estimated using Monte-Carlo simulations.

When the process is outside the fast exchange regime, k_{ex}, p_a, p_b, and $\Delta\omega$ are obtained from a fit of the data to equation 3.3 valid for all time scales. If the exchange process is fast on the NMR chemical shift time scale, only R_{ex} and k_{ex} are obtainable from non-linear curve fitting of the dispersion profiles to equation 3.11.[112,111,42] In addition to the parameters mentioned, the transverse relaxation rates for infinite small values of τ_{cp}, $R_2\left(1/\tau_{cp} \rightarrow \infty\right)$, are fitted in addition for each B_0 field.

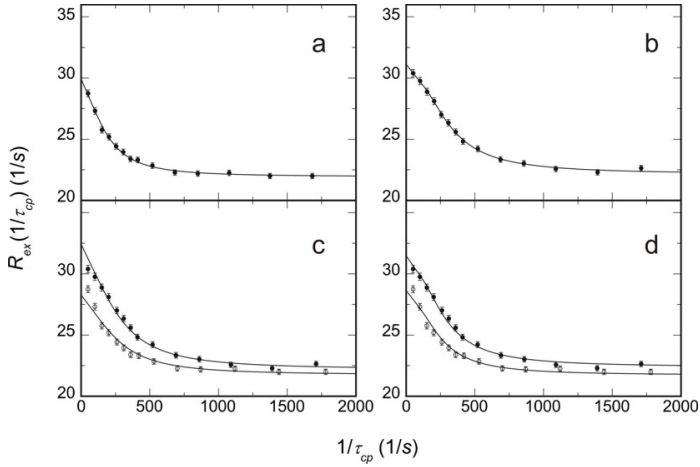


Figure 3.5: Relaxation dispersion profiles for Val423 in the nucleotide binding domain of a P-type ATPase (see chapter 6) recorded at 600 MHz (a) and 750 MHz (b). Data were fitted independently to the fast-exchange equation (a) and to the general two site exchange equation (b) yielding excellent fits: $\chi^2_{\mathrm{red}} = 0.36$ and 0.33 at 600 MHz and 750 MHz, respectively. Values of $k_{\mathrm{ex}} = 612\,\mathrm{s}^{-1}$ and $706\,\mathrm{s}^{-1}$ are obtained. A global fit of the data to the fast-exchange equation indicates that the exchange process is outside the fast exchange regime (c). Use of the general two site exchange equation yields $k_{\mathrm{ex}} = 630\,\mathrm{s}^{-1}$ and a reduction of χ^2_{red} from 4.61 to 3.46 (d).

Chapter 4

Tutorial for the Analysis of NMR Relaxation Data

A number of programs suitable for analyzing relaxation data have been coded by other groups and are available *via* the world wide web (see below). However, these programs require data input in files in special formats. Creating and reformatting such input files manually is a tediuos work; furthermore, it is prone to errors. Most computer operating systems (in this thesis, LINUX was used) support scripting languages like AWK or PERL, which are well suited for such reformatting tasks.

During the course of the present work, a number of AWK or shell scripts have been implemented for the purpose of creating such input files. In the case of other specific requirements, the scripts can be easily modified. All scripts have been commented in detail and it should be possible to use them without any further description; nevertheless, this chapter illustrates the procedure of analyzing ^{15}N relaxation data according to the MF-approach and the analysis of

relaxation dispersion data using these scripts. Furthermore, all relevant files of other programs are explained, and the complete examples including all necessary files are given in appendix C.

4.1 Obtaining Software

All software used can be obtained from the following web pages:

- SPARKY
  `http://www.cgl.ucsf.edu/home/sparky/`

- PDBINERTIA, TM_EST, R2R1_DIFFUSION, QUADRIC_DIFFUSION
  `http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/`

- MODELFREE, CURVEFIT
  `http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/`

- FASTMODELFREE, SPARKY2RATE
  `http://xbeams.chem.yale.edu/~loria/software.htm`

- SCILAB
  `http://scilabsoft.inria.fr/`

- FMINUIT
  `http://www.fis.unipr.it/ãllodi/Fminuit/Fminuit_intro.html/`

- XMGRACE
  `http://plasma-gate.weizmann.ac.il/Grace/`

The scripts implemented during the course of this work are available upon request.

4.2 Model-Free Analysis

As a prerequisite for this procedure, the NMR spectra including the assignment have to be displayed in Sparky. This is readily accomplished as described in the Sparky manual and is therefore not described here. Peak lists from programs other than sparky can be converted into Sparky format using awk. An example of a Sparky peaklist format is shown below:

```
assignmentN-HN delta(N) delta(H)
R317N-HN 120.83 8.27
```

Once *all* spectra to be analyzed have been loaded (e.g. for determining R_1, 10 spectra corresponding to 10 relaxation delays *plus* duplicate measurements), the command `rh`, which is described in the Sparky manual, is used. The resulting file contains the fitted relaxation time constant, an estimated error, and the peak intensities for all relaxation delays. The error estimation done by Sparky is based on the signal-to-noise ratio, not on duplicate experiments. The perl script sparky2rate calculates errors using duplicate experiments and interpolates the uncertainties for all time points. As input file, the output from `rh` is used. Relaxation rates are fitted using CurveFit, and uncertainties in the fitted rates are estimated from either Jackknife or Monte-Carlo simulations. Two parameters have to be defined in `sparky2rate.pl` by the user:

```
$curvefit = "/usr/local/bin/curvefit/curvefit";
$curvefitcall = "$curvefit -jack -grid";
```

The path to the CurveFit binary is given by `$curvefit`, and options for calling the program are specified in `$curvefitcall`. In this case, `-jack` causes CurveFit to use Jackknife simulations for error estimation; `-grid` invokes a grid search for the initial guess of the relaxation rate (see the CurveFit manual for more details). The script is invoked with the following command:

`sparky2rate.pl` ⟨*rh-file*⟩

For example, fitting R_1 data is done by typing

```
sparky2rate.pl T1_600.rh
```

The fitted rates are written to `T1_600.rh.rates` and `T1_600.rh.rates.xmgr` in the format

```
residue   rate   error
```

`T1_600.rh.rates.xmgr` contains only the residue number, while the amino acid type is included in `T1_600.rh.rates`. The interpolated errors are stored in `errors.xmgr`; plots of the fitting results for each residue are saved in the directory `T1_rh.600.dat` and can be visualized by using XMGRACE:

`xmgrace -type xydy` ⟨*filename*⟩

If the fits are to be checked manually (recommended), all files ending with ⟨*extension*⟩ can be displayed successively by issuing

`display.csh` ⟨*extension*⟩

The next residue is displayed after XMGRACE has been closed.

4.2.1 Estimation of the Rotational Diffusion Tensor

Prior to estimating the rotational correlation time, the R_2/R_1 ratio has to be calculated, and residues either subject to chemical exchange or undergoing fast internal motions have to be excluded. All this is performed by the script `create_r2r1.csh`:

`create_r2r1.csh` ⟨*R1_file*⟩ ⟨*R2_file*⟩ ⟨*NOE_file*⟩ ⟨*NOE_cutoff*⟩

where ⟨*NOE_cutoff*⟩ represents a cutoff: residue with NOEs smaller than this value will be excluded from the estimation of τ_c. The output includes five files: `R1.av` and `R2.av` contain the sum and the average of the relaxation rates, as well as the number of residues. The sequence number, values of R_2/R_1 and uncertainties as well as a flag are listed in `R2R1.results`:

```
381 11.5061  0.7247 0
451  2.4568  0.0371 2
```

The flag indicates whether or not the residue is accepted for estimating τ_c; possible values are:

```
0 accepted
1 chemical exchange & low NOE
-1 chemical exchange
2 low NOE
```

Residue numbers and flags are also stored in `exclude.results`:

```
381 0 accepted!
451 2 low NOE
```

The input file which contains only accepted residues is called `R2R1.input`. Before the diffusion tensor can be estimated, the center of mass of the molecular structure should be translated to the coordinate origin; this is achieved with PDBINERTIA:

`pdbinertia -rt` ⟨*input_pdb*⟩ ⟨*output_pdb*⟩

The option `-rt` causes PDBINERTIA to write the translated and rotated coordinates to ⟨*output_pdb*⟩. It might be useful to pipe the statistics from the screen into a file, for example:

```
pdbinertia -rt kdp.pdb kdp_rot.pdb > pdbinertia.log
```

As described in section 2.3.1, two possibilities for estimating the rotational diffusion tensor exist. R2R1_DIFFUSION can be used for a direct estimation based on the R_2/R_1 ratio; however, this program is very sensitive to the initial guess of τ_c. Furthermore, R2R1_DIFFUSION is not capable of analyzing a fully anisotropic tensor. In order to use R2R1_DIFFUSION, a control file containing the following informations has to be created:

```
# of residues   spectrometer field (MHz)   #simulations
Diso   Dratio   theta   phi
Dratio_min   Dratio_max   #steps
```

```
name of R2/R1 file
name of input pdb file
name of output pdb file
```

where `#simulations` are the number of simulations for statistical analysis, `Diso` is the isotropic rotational diffusion coefficient (see equation 4.1), `Dratio` is given by $D_{\parallel}/D_{\perp}$, and `theta` and `phi` are the Euler angles describing the orientation of the diffusion frame. `Diso` is the only required entry; `Dratio`, `theta`, and `phi` can be set to 1.0, 0.0, and 0.0, respectively. `Dratio_min` and `Dratio_max` set the limits for grid searching on `Dratio`; the number of grid searching steps is given by `#steps`.

Note that R2R1_DIFFUSION and QUADRIC_CONFUSION (see below) use the isotropic rotational diffusion coefficient instead of the rotational correlation time. Both parameters are related to each other according to

$$\tau_c = \frac{1}{6D_{\text{iso}}} \tag{4.1}$$

An example control file looks like

```
# example control file for R2R1_diffusion
5 600.13 100
1.0e+9 1.1 20 0
1.0 1.5 20
R2R1.input
kdp_rot.pdb
kdp_axsym_r2r1.pdb
```

R2R1_DIFFUSION is called by issuing

`r2r1_diffusion` ⟨*control_file*⟩

and the output is piped into a file:

```
r2r1_diffusion r2r1_diff.ctrl > r2r1_diff.log
```

Sections containing the results for the isotropic and axially symmetric diffusion model are indicated in the output (an example for the analysis of five residues can be found in the appendix). For the axially symmetric model, the structure is rotated to the diffusion frame. It should be mentioned, however, that five residues are *not* enough for a precise estimation of the rotational diffusion tensor.

Use of Quadric_Confusion is recommended if a discrimination between axially symmetric and complete anisotropic tumbling has to be made. For the quadric analysis, local effective correlation times are necessary; these are calculated from the R_2/R_1 ratios using R2R1_tm. The output should be piped into a file. For example:

```
r2r1_tm > tm.dat
```

The program prompts for the following information:

```
Input nucleus (13C or 15N): 15N
The default bond length and csa are: 1.0200 A and  -160.00 ppm
Enter new values or <RETURN> to accept:
Input the proton spectrometer field used for R1 (MHz): 600.13
Input the proton spectrometer field used for R2 (MHz): 600.13
Input initial guess for tm (ns): 11
Input number of residues to fit: 5
Input file name for R2/R1 data: R2R1.input
```

The output, in this example `tm.dat`, has the following content:

```
# R2/R1 Analysis:
#
# nucleus:   15N
# bond length (A):     1.0200
# CSA (ppm):    -160.00
#
#residue      tm     dtm
318   11.7887      0.1131
```

```
321   11.5759      0.1075
333   10.8482      0.1608
345   15.0847      0.2807
381   10.4297      0.3395
```

QUADRIC_CONFUSION is called with the command

`quadric_diffusion` ⟨*control_file*⟩

where the ⟨*control_file*⟩ contains these informations:

```
Dratio_min   Dratio_max   #steps
number of atom types   'atom_1'   ...   'atom_n'
name of tm-file for atom_1
...
name of tm-file for atom_n
name of input pdb-file
name of output pdb-file (axially symmetric tensor)
name of output pdb-file (anisotropic model)
```

where `Dratio_min`, `Dratio_max`, and `#steps` have the same meanings as above. The atom types have to be enclosed in single quotes, e.g. 'N' or 'CA'; the output of R2R1_TM provides the `tm-file`.

Note that no initial guess for τ_c has to be given, since local effective correlation times are used. In our example, the ⟨*control_file*⟩ has the following content:

```
# example control file for quadric_diffusion
0.1 1.2 40
1 'N'
tm.dat
kdp_rot.pdb
kdp_ax_qdr.pdb
kdp_an_qdr.pdb
```

QUADRIC_CONFUSION is invoked with the command

```
quadric_diffusion quadric.ctrl > quadric.log
```

Sections containing the results for the isotropic, axially symmetric and fully anisotropic model are indicated in the output file; an example for the analysis for five residues is given in the appendix. It should be emphasized that the rotational diffusion tensor *cannot* be determined reliably from such a small number of ^{15}N–^{1}H vectors; this example just illustrates the fitting procedure.

4.2.2 Model-free Analysis using FASTModelfree

Analyzing relaxation data with Modelfree requires a vast number of different input files and extensive user interaction, especially during the process of model assignment. This is not only time consuming, but also prone to errors. The Perl script FASTModelfree, in the remainder of the work abbreviated FMF, reduces user interaction to a minimum, since every step is perfomed automatically: creation of input files, model assignment etc. Furthermore, a detailed protocol is printed which enables the user to trace every single step during the optimization process. The use of FMF in combination with Modelfree for the analysis of ^{15}N backbone relaxation data is described in the following section.

The `FMF.config` File

A prerequisite for using FMF is the configuration file `FMF.config`, containing all relevant information and parameters. `FMF.config` can be created using a graphical user interface by typing the command

```
setupFMF
```

or by editing the file manually. The configuration file is explained in detail in the following (the complete file is listed in the appendix).

```
tensor isotropic
```

Definition of the diffusion tensor. Possible options are `isotropic` and `axial symmetric`; the fully anisotropic model is not yet implemented in Modelfree.

`cutoff 0.95`

Confidence limit for χ^2-testing; a commonly used value is 0.95. The fit of any model to the relaxation data of a given spin is tested using the residual sum squared error, Γ_i^{fit} (see equation 2.20). Simulated distributions of Γ_i^{sim} are obtained from Monte-Carlo simulations. The confidence limit or confidence region contains a certain percentage of the total probability distribution. The meaning of the confidence region can be stated loosely as: "There is a 95 % chance that the true parameters fall within this region around the fitted ones". If $\Gamma_i^{\text{fit}} < \Gamma_i^{\text{sim}}(\texttt{cutoff})$, the model adequately describes the data and is accepted.

`Fcutoff 0.80`

Confidence limit for *F*-testing; a common value is 0.80. *F*-statistics are used for the comparison of models with a different number of parameters, e.g. models 1 and 2. If a residue could not be assigned to model 1 based on Γ_i, but model 2 could be accepted, it has to be verified if this reducion in Γ_i is statistically significant. A simple reduction of the sum squared error, $\Gamma_i^{\text{fit}}(\text{M2}) < \Gamma_i^{\text{fit}}(\text{M1})$, does not indicate an improved fit; it is a mere consequence of the introduction of an additional parameter. The *F*-value is calculated according to equation 2.22 for the fitted and simulated Γ_i for model 2, and the more complicated model is only accepted if $F^{\text{fit}} > F^{\text{sim}}(\texttt{Fcutoff})$.

`optimize Yes`

Determines whether the diffusion tensor is optimized; general setting is `Yes`.

`maxloop 10`

Maximum number of runs. The analysis will be stopped if convergence of the diffusion tensor is achieved or `maxloop` runs have been performed.

`almost1 20`

Defines a new cutoff value for Γ_i. If a residue could be neither assigned to model 1, 2, or 3, it is nevertheless assigned to model 1 if $\Gamma_i <$ `almost1` (see figure 2.14). Typically, a value of 20 is used.

```
S2cutoff 0.7
```

Lower limit of S^2. The optimization of the diffusion tensor, especially the anisotropy, is improved if residues with a larger degree of motional freedom are excluded from the analysis. For relaxation data acquired at a B_0 field of 14.1 T, a value of 0.7 is commonly used.

```
seed 1975
```

Integer for the random number generator necessary for Monte-Carlo simulations. Usually, the `seed` is taken from the system clock.

```
numsim 100
```

Number of Monte-Carlo simulations. Generally, 300–500 are statistically satisfactory, although a smaller number might be used for initial runs.

```
jobname iso
```

Name of the job; used for output files.

```
gamma -2.71
rNH 1.02
N15CSA -160
```

Constants for the spin pair for which data is analyzed.

```
tm 11.6
tmMin 9.0
tmMax 14.0
tmGrid 40
tmConv 0.001
```

These parameters define the initial guess for the rotational correlation time (`tm`), as well as upper and lower bounds (`tmMin, tmMax`) and number of steps for grid searching (`tmGrid`). Convergence of τ_c is controlled with `tmConv`: τ_c is regarded as converged if its value changes by less than `tmConv` between two successive runs.

```
Dratio 1.22
DratioMin 1.0
DratioMax 1.3
DratioGrid 10
DratioConv 0.001
```

Same as above, but only used if `tensor` is set to `axially symmetric`. Two similar blocks for the polar angles ϕ and θ have to be provided.

```
model1only No
```

If set to `Yes`, the diffusion tensor is optimized using only spins assigned to model 1. In some cases, e.g. large number of spins or multiple fields, authors have reported a significant improvement in speed. Generally, `model1only` is set to `No`.

```
mpdb
file{0}{R1} ../r1.600
file{0}{R2} ../r2.600
file{0}{NOE} ../nhnoe.600
```

Names including the path variable of the coordinate file, R_1, R_2, and hetNOE data files. A pdb coordinate file is only needed for an axially symmetric diffusion tensor and is created as described above using either R2R1_DIFFUSION or QUADRIC_CONFUSION.

```
file{0}{field} 600
```

Proton resonance frequency in MHz of the field at which the data were acquired. In cases of multiple static fields, additional lines have to be added, for example `file{1}{field} 750`.

Running FASTModelfree

Prior to running the FMF PERL script, the path to the MODELFREE binary file has to be specified at the beginning of FMF:

```
$modelfree = "../modelfree_v415";
```

A complete MF analysis usually takes 8–10 hours depending on `numsim` and is usually run over night. Therefore, it is helpful to run the analysis in the background and pipe the output into a file (e.g. `example.log`):

`fastMF >` ⟨*log_file*⟩ `&`

FASTModelfree Output

FMF creates a number of files needed for running Modelfree; they are explained in detail in the Modelfree manual and will not be discussed here. These files include ⟨*jobname*⟩`.MF*, mfinput.*, mfmodel.*`, and `mfout.*`. The other files are explained in the following.

The file ⟨*jobname*⟩`.log` details the model assignment and diffusion tensor data for each iteration. In addition, the changes of the diffusion tensor parameters as well as possible convergencies are shown. The optimized parameters of each iteration `x` are stored in ⟨*jobname*⟩`.iterx.par`. In case of an axially symmetric diffusion tensor, the coordinate system of the original structure file is rotated to the principal axis system of the diffusion tensor; these coordinates are saved in ⟨*jobname*⟩`.x.pdb`.

If FMF has been invoked with the command shown above, detailed information about the model assignment and diffusion tensor optimization are stored in `example_iso.log`. This output is explained using the MF analysis for the five residues from the examples above with the results of the diffusion tensor analysis obtained from R2R1_diffusion: an isotropic rotational diffusion tensor and an initial estimate of the correlation time $\tau_c = 11.6\,\text{ns}$.

For each Modelfree run, the input file controlling the optimization process is printed.

```
 ***** Starting Model 1 *****
mfinput File Listing
optimization   tval
```

The type of weighting function applied to the relaxation parameters is controlled by `optimization`. Reciprocal uncertainties are used if `optimization` is set to `tval`, in this case, χ^2 is calculated using equation 2.21. Reciprocal squares of the relaxation parameters are used instead of reciprocal uncertainties if `optimization` is set to `frac`. In most cases, `optimization` should be set to `tval`.

```
seed   -1975
```

The seed for the random number generator.

```
search    grid
```

Controls whether a grid search is performed for the internal motional parameters; options: `grid`, `none`.

```
diffusion    isotropic    none
      fix
```

This section defines the diffusion model (`isotropic`, `axial`). The tensor is kept `fix`, and no grid searching is done (`none`), since model selection has to be completed before optimizing the diffusion tensor.

```
simulations   pred   100   0.000
```

Control of the Monte-Carlo simulations. `pred` causes MODELFREE to simulate data sets by adding random noise on the fitted relaxation parameters; if set to ⟨*expr*⟩, the random noise is added to the experimental values. 100 simulations are performed in this example; usually 300–500 are satisfactory. The last option controls the trim of the data; this should normally be set to zero.

```
selection   none
      fix
```

No selection based on *F*-tests needs to be done here, since only model 1 is fitted.

```
fields    1   600.000
tm   0.116E+02   1   2   0.900E+01   0.140E+02   40
```

The ^{1}H resonance frequency of the field at which the relaxation data were acquired, and initial guess and grid search control for `tm`; this is obsolete at this stage, since the diffusion tensor is kept fixed.

```
Starting initial grid search for model 1
Initial chi-square:   0.8243E+06
Best chi-square:   0.6261E+06
Starting fix isotropic diffusion optimization for model 1
Total Chi-Square:    0.622E+06
tm:     11.600
Starting   100 Simulations
 Simulation    10
 Simulation    20
 Simulation    30
 Simulation    40
 Simulation    50
 Simulation    60
 Simulation    70
 Simulation    80
 Simulation    90
 Simulation   100
```

A grid search for the internal motional parameters (in this case only S^2) is done, followed by optimization and Monte-Carlo simulations.

```
Starting final output
Results for Model 1:
318   117.911   6.542
321   2.393   5.357***
333   18.201   5.794
345   527.692   5.185
381   35.358   7.062
451   620807.562   5.535
```

These are the results for fitting the data to model 1. The first column contains the residue number, followed by Γ_i^{fit} and $\Gamma_i^{\text{sim}}(0.95)$ (remember that `cutoff` was set to 0.95 in the example `FMF.config` file). Residues that were adequately fit by model 1, i.e. $\Gamma_i^{\text{fit}} < \Gamma_i^{\text{sim}}(0.95)$, are indicated by `***`.

```
Finished Model 1
Current model assignments:
Model 1 spins:
321
Model 2 spins:
Model 3 spins:
Model 4 spins:
Model 5 spins:
Unassigned spins:
318 333 345 381 451
```

Before fitting of model 2 is started, the current model assignment is shown.

```
 ***** Starting Model 2 *****
mfinput File Listing
[...]
```

The relaxation data are fitted to model 2. This optimization yields Γ_i^{fit} and $\Gamma_i^{\text{sim}}(0.95)$.

```
mfinput.2 File Listing
optimization              tval
[...]
selection                ftest
```

If $\Gamma_i^{\text{fit}} < \Gamma_i^{\text{sim}}(0.95)$ for any of the residues, an F-test has to be performed. A synthetic data set is created using model 1 and subsequently fitted with model 2, and F_i^{fit} and F_i^{sim} are calculated.

```
[...]
```

```
Starting final output
Results for Model 2:
318 118.255 6.722 Ftest: 0.000e+00 0.000e+00
333 18.068 5.045 Ftest: 0.000e+00 0.000e+00
345 527.692 7.237 Ftest: 0.000e+00 0.000e+00
381 1.999 3.378* Ftest: 1.669e+01 5.829e+00***
451 4509.825 3.288 Ftest: 0.000e+00 0.000e+00
```

The first three columns of the output shown above are the same as for model 1: residue number, Γ_i^{fit} and $\Gamma_i^{\text{sim}}(0.95)$. Residues for which $\Gamma_i^{\text{fit}} < \Gamma_i^{\text{sim}}(0.95)$ are indicated by `*`. The last two columns represent F_i^{fit} and F_i^{sim}, respectively. Residues which are assigned to model 2, i.e. $F_i^{\text{fit}} > F_i^{\text{sim}}(0.80)$, are marked by `***`.

```
Finished Model 2
Current model assignments:
Model 1 spins:
321
Model 2 spins:
381
Model 3 spins:
Model 4 spins:
Model 5 spins:
Unassigned spins:
318 333 345 451
```

The actual model assignments are given before model 3 is tested; the process is exactly the same as described for model 2.

```
[...]
Finished Model 3
Current model assignments:
[...]
Unassigned spins where Model 1 SSE < 20:
```

```
Spin 333: SSE 18.2009 Old Cutoff 5.7940 New Cutoff 20
```

This is the "almost 1" selection. Residues that could neither be fitted adequately by models 1–3 are assigned to model 1 if $\Gamma_i^{\text{fit}} < 20$. All residues that have not been assigned to a model at this stage are fitted against model 4.

```
 ***** Starting Model 4 *****
[...]
selection   none
```

Model 4 has three fit parameters (S^2, τ_i, and R_{ex}) for three experimental parameters. Thus, the number of degrees of freedom is zero and F-testing cannot be performed.

```
[...]
Starting final output
Results for Model 4
318 1.1825e+02 7.2236e+00
345 0.0000e+00 4.6257e+00***
451 4.6032e+03 1.4371e+00
```

Model 4 is only accepted if $\Gamma_i^{\text{fit}} = 0$; in this example, Glu345 is assigned to model 4. The same procedure is repeated for model 5.

```
Finished Model 5
Current model assignments:
[...]
Remaining unassigned spins:
Residue 318 SSEs:
Model 1: 117.9109 6.5417
Model 2: 118.2547 6.7220
Model 3: 115.0846 3.5663
Model 4: 118.2547 7.2236
Model 5: 118.2547 7.2633
```

Residues which could not be assigned to any model are displayed, along with Γ_i^{fit} and $\Gamma_i^{\text{sim}}(0.95)$.

```
Removing undesired spins for diffusion tensor optimization
Throwing out 451, low S2 = 0.080
Model assignments for final optimization:
```

Before the diffusion tensor is optimized, residues with $S^2 < 0.7$ are excluded and the model assignment for the final optimization is shown.

```
mfinput.3 File Listing
[...]
diffusion            isotropic           none
   powell           grid           1
```

The diffusion tensor is optimized using Brent's implementation of Powell's method for multidimensional minimization; the same method is used during Monte-Carlo simulations.

```
Starting initial grid search for model 1
Initial chi-square:   0.1886E+04
Best chi-square:   0.2200E+03
```

Grid searching is now performed for all parameters including the diffusion tensor.

```
Starting powell isotropic diffusion optimization for model 1
tm:    11.6000     X2:   0.2259E+02
tm:    10.6069     X2:   0.1045E+03
tm:    12.5167     X2:   0.1518E+03
tm:    11.4558     X2:   0.1746E+02
tm:    11.3999     X2:   0.1681E+02
tm:    11.3795     X2:   0.1677E+02
tm:    11.3852     X2:   0.1677E+02
tm:    11.3738     X2:   0.1677E+02
```

```
Total Chi-Square:    0.168E+02
tm:     11.380
Starting   100 Simulations
[...]
Starting final output
```

After optimization of the diffusion tensor, the final results are stored. The internal motional parameters are written to `iso.iter1.par`, and the current model assignment, diffusion tensor, and changes of the latter are written to `iso.log`.

```
==================================================
Iteration 1
Tensor: tm 11.6
[model assignment]
Delta tm: 0.2200 Converged? NO
==================================================
Iteration 2
Tensor: tm 11.380
Model 1 spins:
321 333
Model 2 spins:
381
Model 3 spins:

Model 4 spins:
345
Model 5 spins:
451
Unassigned spins:
318
Delta tm: 0.0000 Converged? YES
```

This procedure of model assignment and diffusion tensor optimization repre-

sents *one* iteration and is repeated until `maxloop` is reached or the diffusion has tensor converged (`tmConv` < 0.001), which is the case for the example after the second iteration as indicated in `iso.log`.

Note that output files from Modelfree are overwritten during the next iteration; only the *last* generation is retained. These output files are not explained here, since FMF extracts all relevant informations and stores them in `iso.iter1.par` and `iso.log`, respectively.

Extraction of the Optimized Parameters

The optimized values for the rotational diffusion tensor can be extracted from `mfout.final` by typing

```
get_tensor.awk
```

which prints the *simulated* value including the uncertainty for the correlation time and, in the case of axial symmetry, `Dratio`, `theta`, and `phi`.

For the internal motional parameters, a similar script has been implemented that extracts the data from `mfout.final`:

`get_all` ⟨*path_to_script*⟩

The resulting output does not contain residues that were excluded from the final optimization due to higher flexibility. If data for *all* residues are desired, which is usually the case, the following command should be used:

`get_all_iter` ⟨*name_of_iteration_file*⟩ ⟨*path_to_script*⟩

The difference between both scripts is explained now: In the example from this section, Gly451 was excluded from the final optimization due to a S^2 smaller than `S2cutoff`. The output of

```
get_all ./
```

contains only data for Glu321, Ile333, Glu345, and Ser381, whereas

```
get_all_iter iso.iter2.par ./
```

includes Gly451. The output is stored in `parameter.dat` for `get_all` and in ⟨*name_of_iteration_file*⟩`.parameter` for `get_all_iter`.

4.3 Relaxation Dispersion

Although a program for fitting CPMG relaxation dispersion data called CPMGFIT is available from the Palmer-group, it has not been used in this work for several reasons. First, when data from multiple fields are fitted globally, the $R_2\left(1/\tau_{cp} \rightarrow \infty\right)$ are assumed to be independent of B_0. This assumption is valid if the data were acquired using TROSY-type sequences; if conventional pulse schemes had been used, this assumption is often violated. Second, holding an arbitrary parameter at a constant value during optimization is not possible with the available binaries. The source code can be tweaked to keep the exchange rate constant fix, but this is a rather inconvenient procedure. Third, the program is written in FORTRAN, and changing the source code therefore requires knowledge of this language.

Instead, scripts for fitting relaxation dispersion data to a flat line (no exchange), the fast exchange approximation (equation 3.11), and the equation valid for all time scales (equation 3.3) have been implemented for SCILAB. SCILAB's scripting language is easy to learn, hence it is possible to modify the scripts according to personal needs. It should be noted that all scripts were implemented for *global fitting* of two static fields, i.e. data from two B_0 fields have to be acquired.

4.3.1 Creating the Input Files

For creating the input files, several AWK and shell scripts are used. In the first step, the signal intensities are read on a per-residue basis from peak lists along with length the corresponding relaxation period (T_{CP} in figure 3.4 on page 53). The signal intensities should be in the fourth column, but the script can be changed to read any other column. An example of a SPARKY peak list format

including signal intensities in the fourth column is given below:

```
Assignment   w1   w2   Data Height
```

In this work, every point of the dispersion curve was measured twice in order to estimate the experimental uncertainties. Therefore, two peak lists exist for each value of τ_{cp}, and these lists should be named according to the following scheme (see the example given below): `filename_l6.dat`, where `filename` is a unique identifier for a single series of experiments; l6 is a loop counter defining τ_{cp} in the pulse sequence used in this work. In order to calculate τ_{cp}, T_{CP}, l6, and p3 have to be given. Note that the fitting scripts require $1/(\tau_{cp})$ as x-values; this conversion is performed automatically.

The script `dispersion_sp.awk` converts the peak lists into a single file named `R2eff.dat` containing all residues in the following format

```
Residue   1/(2tcp)   R2eff   uncertainty
```

and is invoked with the command

`dispersion_sp.awk` ⟨*files1*⟩ ⟨*files2*⟩ ⟨*numpts*⟩ ⟨*pulselength*⟩

where ⟨*numpts*⟩ refers to the number of τ_{cp}-values recorded *plus* the reference experiment, ⟨*pulselength*⟩ is the ^{15}N 90° pulse length in µs; the other arguments have been described above. A complete example where 14 values of τ_{cp} plus a reference experiment have been recorded looks like this:

```
./dispersion_sp.awk disp1600_"*".list disp2600_"*".list 15 52
```

In the next step, residues to be analyzed have to be extracted from `R2eff.dat` and stored in seperate files. This is achieved by typing the following command in a shell:

`extract_R2eff.awk` ⟨*R2eff.dat*⟩ ⟨*residue_file*⟩ ⟨*extension*⟩

in which ⟨*residue_file*⟩ is the name of a file containing the residues to be analyzed, one residue per line:

```
R350
R403
D447
```

and ⟨*extension*⟩ is the extension of the output files, for example

```
./extract_R2eff.awk R2eff.dat residues.dat 600
```

The resulting output files `R350.600`, `R403.600`, and `D447.600` have the same format as `R2eff.dat`, but contain only data of one residue. Note that both steps, i.e. invoking `dispersion_sp.awk` and `extract_R2eff.awk`, have to be repeated for the data of the second static magnetic field. Since `./dispersion_sp.awk` will overwrite an existing `R2eff.dat` file, it is necessary to rename `R2eff.dat`.

SCILAB requires the data in *rows*, the residue number followed by the `R2eff`-values sorted according to *increasing* $1/(\tau_{cp})$:

```
350 28.3826 27.1166 26.2409   ...
```

This conversion is perfomed by running `reformat.csh`:

`reformat.csh` ⟨*extension*⟩ ⟨*outpath*⟩

where ⟨*extension*⟩ refers to the input files and ⟨*outpath*⟩ is the path variable where the output files are to be written. To continue with the example:

```
./reformat.csh 600 ./
```

causes `reformat.csh` to convert all files with the extension ".600" in the current directory; the output written is to the same path. The output comprises `R350.600.mtx`, `R403.600.mtx`, and `D447.600.mtx`. The data of all residues are also written to `all_res.600.mtx`. The format of these files is the same as shown above.

4.3.2 Using Scilab

The scripts for each function are comprised of two files: one file containing the actual script, e.g. `fast.sce`, and a fit function, `fitfun_fast.sci`. Note

that these and three additional files, `fminuit.scimex`, `load_fminuit.sce`, and `mkscale.sci`, have to be copied to the directory containing the scripts. The first step is to edit the script, e.g. `fast.sce`, to provide path variables, B_0 fields, $1/(\tau_{cp})$-values, and errors. The relevant section is shown below:

```
//---------------------------------------------------------
// user input
//---------------------------------------------------------
// enter input files (including path) here
fnam1 = './all_res.750.mtx';     // field 1
fnam2 = './all_res.600.mtx';     // field 2
// enter path to output file
// this is one file for all residues
outpath = './fast/';
// enter name of output file
outname = 'fast.out';
outfile = outpath+outname;
// enter path to xmgrace files
gracepath = './grace/fast/data/';
// define the static fields
B01 = 17.62; // field 1
B02 = 14.09; // field 2
// define the x-values
tcp1 = [50.21 100.83 ...];  // field 1
tcp2 = [50.18 100.73 ...];  // field 2
// define the errors
err1 = 0.02 * [1 1 ...]; // error for field 1, constant value
err2 = 0.02 * [1 1 ...]; // error for field 2, constant value
// If errors on a per-residue and per-tcp basis are desired,
// create a file of the following format:
// residue number err(x1) err(x2) ... err(xn)
```

```
// for each field and process them as the data files below
//---------------------------------------------------------
// end of user input
//---------------------------------------------------------
```

Most of the information the user has to provide should be self-explaining, but a few notes will be made here.

- If only one residue is to be analyzed, e.g. Asp447, the input files should be defined as `D447.mtx.600` and `D447.mtx.750`.
- The x-values are the values of $1/(\tau_{cp})$ in *increasing* order.
- In the current versions of the scripts, a "global" error is assumed for all values of $R_2\left(1/\tau_{cp}\right)$: the uncertainties obtained from duplicate experiments resulted in errors significantly smaller than 1 %; χ^2-fitting with errors this small is not possible. Therefore, a uniform error of 1–2 % is assumed for all $R_2\left(1/\tau_{cp}\right)$.
- Initial guesses for the fitting parameters are defined in the scripts. A good convergence is obtained from a large variety of initial guesses; hence, it should not be necessary to change these.
- A common way of presenting relaxation dispersion data are plots of dispersion profiles. Unfortunately, the default graphical output from SCILAB does not meet high-quality standards. For producing plots, the use of XMGRACE is recommended (see section 4.3.4). The path where SCILAB stores the raw data for these plots has to be provided by the user as `gracepath`.

SCILAB is started by typing `scilab` in a shell on a LINUX computer. The fitting scripts are invoked by typing

`exec` ⟨*script_name*⟩

in the main Scilab window, e.g. `exec fast.sce` for fitting data against the fast-exchange approximation. After the parameters for each residue have been optimized, the experimental values as well as the fitted curves are displayed in a graphic window; fitting will continue after pressing the ⟨`Enter`⟩ key in the main window.

The main output is written to the directory `outpath`; in the example, this is `./fast`. Output files are `350.res`, `403.res`, `447.res`, and `fast.out`. The former contain the fitting results for each residue, in this case fitted to the fast-exchange approximation:

```
Results for residue 350:
========================
kex  :     768.9355 +/-  63.8930
Rex  :      11.8784 +/-   0.5533
R20_1:      23.0437 +/-   0.2302
R20_2:      21.4316 +/-   0.1751
Chisq:       8.5764
```

Note that Monte-Carlo or Jackknife simulations have not been implemented so far; estimated uncertainties of the fitted parameters are derived from the covariance matrix.

The results for all residues are also saved in `fast.out` in a format that allows data to be extracted by simple scripts:

```
res kex kex_err Rex Rex_err R20_1 err R20_2 err chisq
```

If the data were fitted to the equation valid for all time scales, the files would contain the parameters given in the example below:

```
Results for residue 350:
========================
kex  :     739.9969 +/-  65.0777
pa   :       0.8750 +/-   0.1227
dw   :     291.3490 +/- 127.0809
```

```
R20_1:      23.0677 +/-   0.1236
R20_2:      21.4263 +/-   0.0879
Rex_1:      11.6639
Rex_2:       7.6619
alpha:       1.8602
Chisq:      34.0866
```

Values for `Rex_1`, `Rex_2`, and `alpha` are calculated using equation 3.8 and 3.16, respectively.

4.3.3 Statistical Evaluation of the Results

Once the data have been fitted against two different models, *F*-testing has to be used to assess whether the introduction of the more complicated model is statistically justified. This can be accomplished using the script `F.csh`. It is recommended to copy `R2.out`, `fast.out`, and `full.out` into one directory; the script should be called from there. Two path variables have to be defined in the script itself:

```
### CHANGE THESE PATHS ACCORDING TO YOUR SETTINGS ###
set DATA_PATH = ./
set SCRIPT_PATH = ./
#####################################################
```

with `DATA_PATH` being the path to the `.out` files; the path to the script executable is specified by `SCRIPT_PATH`. The command line for invoking the script is

`F.csh` ⟨*comp*⟩ ⟨*simp*⟩ ⟨*numpts*⟩ ⟨*par_comp*⟩ ⟨*par_simp*⟩

where ⟨*comp*⟩ and ⟨*simp*⟩ correspond to the complicated and simple model, i.e. `R2`, `fast`, or `full`, respectively; ⟨*numpts*⟩ is the number of experimental points; ⟨*par_comp*⟩ and ⟨*par_simp*⟩ are the number of fitted parameters of the complicated and fitted models. The command for testing `R2` against `fast` is:

```
F.csh fast R2 28 4 2
```

The results are printed on the screen by default, but it is useful to pipe the output into a file. The output contains four plus one optional column:

```
residue   X2.simp   X2.comp   F
```

Column two and three contain the χ^2 values of the simple and complicated model, the F-value is given in column four. Large positive values for F indicate that use of the more complicated model is statistically justified; this is indicated by `#` or `###` in the optional column. Since no simulated values for F are obtainable, visual inspection of the results is recommended. For the three residues of the example, the output of `F.csh fast R2 28 4 2` is

```
350       2691.43            8.58          3753.82        ###
403         29.00            7.25            36.00
447       5257.20           24.59          2553.85        ###
```

indicating that Arg350 and Asp447 are subject to chemical exchange.

4.3.4 Visualization of the Fits

As mentioned before, high quality plots of fitted data are preferrably produced using dedicated programs; in this case, Xmgrace is used. This section describes how input files for Xmgrace can be generated.

For each residue analyzed, four files are saved:

```
350.res.exp1
350.res.exp2
350.res.fit1
350.res.fit2
```

These contain the experimental and fitted data for both B_0 fields. To create the input files for Xmgrace, `grace.csh` is used from the directory where the results of the fits (e.g. 350.res) are stored:

`grace.csh` ⟨*data_path*⟩ ⟨*output_path*⟩

where ⟨*data_path*⟩ is the path to the files listed above and ⟨*output_path*⟩ defines where the output is written. It is recommended to copy a `header` file containing options for XMGRACE to produce plots of the desired appearance to ⟨*output_path*⟩. An example of such a file is given in the appendix; for a description of available options, the reader is referred to the XMGRACE manual. Here is an example of how to use `grace.csh`, issued from `./fast`:

```
grace.csh ../grace/fast/data ../grace/fast
```

This command creates `350.res.xmgr`, `403.res.xmgr`, and `447.res.xmgr` in the directory `../grace/fast`. The plots can be visualized using

`xmgrace -type xydy` ⟨*filename*⟩

Figure 4.1 shows examples obtained with the output of `grace.csh` and XMGRACE. An alternative, e.g. to verify the results for all residues, is available which displays all files ending with ⟨*extension*⟩ successively:

`display.csh` ⟨*extension*⟩

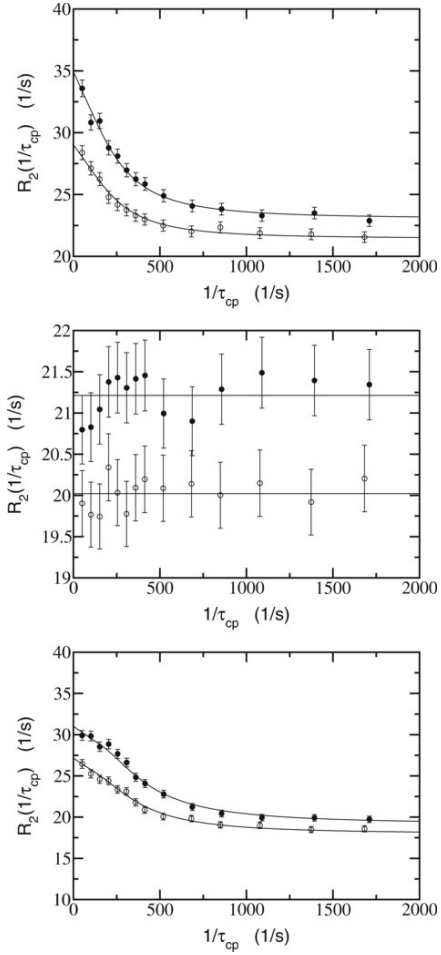


Figure 4.1: Plots of dispersion profiles for Arg360, Arg403, and Asp447 as obtained using the header file shown in the appendix.

Chapter 5

Backbone Motions in the Apical Domains of the Thermosome

This project was conducted in cooperation with the group of Prof. W. Baumeister, Max-Planck Institut für Biochemie, Martinsried. Protein samples were prepared by Dr. G. Bosch.

5.1 Biological Background

Chaperonins, a particular class of chaperons, reversibly bind and refold non-natively folded substrates in an ATP-dependent manner. They are characterized by a multimeric toroidal structure,[129] capturing their substrates in the interior of the torus, thus separating it from the cytosol and preventing non-productive interactions.[130] The refolding process is promoted by a sequence of large conformational rearrangements, which in turn are driven by energy derived from ATP hydrolysis.

Chaperonins can be divided into two groups. Group I chaperonins are generally found in bacteria and eukaryotic organelles of bacterial origin, although there is recent evidence for exceptions.[131] GroEL from *Escherichia coli* is the most prominent member of group I chaperonins.[132, 133, 134] The second, distantly related family, representing the group II chaperonins, occurs in archaea and the eukaryotic cytosol.[135, 136, 137] The thermosome of *Thermoplasma acidophilum* is a representative of the group II chaperonins. Although both classes presumably share a common ancestor, their sequence identity is only 20–25 %, suggesting that they have evolved independently for a long time.[138, 139, 140] This is reflected in a number of distinct features for both groups. The bacterial group I chaperonins are toroids with a sevenfold symmetry that require a co-chaperonin for proper function. In contrast, archaeal-eukaryotic group II chaperonins form eight or nine-membered rings and function independent of a general co-chaperonin.

Both groups feature a common domain organization: an equatorial domain containing the ATPase site, an apical domain responsible for substrate recognition, and an intermediate domain serving as a flexible hinge between the other two domains.[141, 142, 143] In contrast to GroEL, the thermosome is an octameric protein comprising four α and four β subunits. The most distinct difference between the two families is found in the apical domains, at the periphery of the complex: the crystal structures of the isolated α- and β-apical domains of the thermosome (ADT) revealed a helical protrusion with an approximate length of 25 Å,[144, 145, 146] which was also found in an eukaryotic homologue, TriC / CCT.[147] While conserved among all members of the group II family, this protrusion is absent in group I chaperonins.

This difference has an important functional implication: while closure of the central cavity is mediated by the co-chaperonin in members of the group I family, this is achieved by an iris-type aperture in group II chaperonins (see figure 5.1). Their apical domains rotate about 70° and move inward by approximately 2 nm, and the N-terminal parts of the protrusions form an eight-membered β-barrel plug.[145]

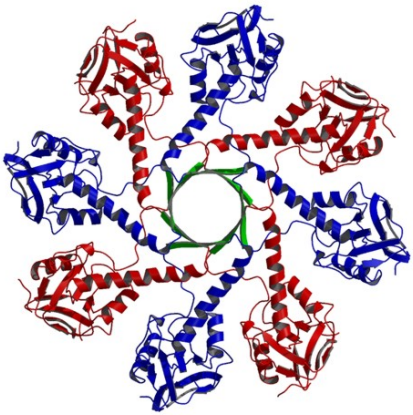


Figure 5.1: Top-view of the apical domains in the closed thermosome. The α- and β-subunits are shown in red and blue, respectively, while the β-strands forming the lid are colored green. This figure and other figures containing molecular graphics were prepared using Molscript and Raster3D.[148,149]

It is interesting to note that a segment of the protrusion adopts different conformations in the X-ray structures of isolated α-ADT and β-ADT (see figure 5.2).[144,146] This has been attributed to crystal packing effects; however, no electron density was found for the corresponding region in TriC / CCT,[147] thus suggesting mobility with possible functional importance. Furthermore, while this segment is α-helical in the isolated form of α-ADT, it participates in the formation of the β-barrel plug in the closed thermosome, thus representing a "chamaeleon sequence". Since the protrusion represents a putative substrate binding site and the presence of regular secondary structure most certainly affects binding, the question arises whether this segment is structured in solution.

Investigation of substrate binding to group I chaperonins has located the binding site between helices H8 and H9 (H11 and H12 in ADT) at a predomi-

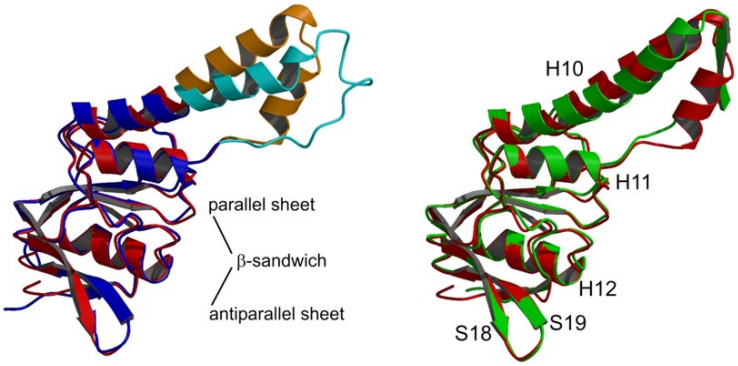


Figure 5.2: Left: Superposition of the crystal structures of the isolated α-apical domain (red) and the isolated β-apical domain (blue). The protrusion is shown in orange for α- and cyan for β-ADT. Right: The crystal structure of α-ADT in the closed thermosome (green) superimposed on the structure of the isolated α-domain (red).

---

nantly hydrophobic surface,[150,151] albeit recent studies on GroEL provide evidence for a participation of residues located deeper in the central channel.[152] In group II chaperonins, the corresponding area in the cleft between helices H11 and H12 is rather hydophilic; however, a hydrophobic surface is found on the protrusions facing the center of the cavity, well oriented for interactions with potential substrates.

Unfortunately, no conclusive data on the location of the substrate binding site in group II chaperonins is available to date. Therefore, both apical domains of the *T. acidophilum* thermosome have been investigated as 17 kDa constructs consisting of 153 residues using heteronuclear NMR spectroscopy.

5.2 Results

5.2.1 Resonance Assignment

Initial HSQC spectra at variable temperatures have shown that the optimal spectral resolution for both domains was achieved at a temperature of 42 °C. The pH value of the α-ADT sample was 5.5; however, the β-domain could only be obtained at pH = 7.8, due to precipitation when the domain's pI (6.3) is approached. At this combination of pH and temperature, intrinsic exchange rates of more than $1000\,s^{-1}$ broaden the signals of solvent exposed amide protons beyond the detection limit of HSQC and H^N-detected triple resonance experiments, which is evident in figure 5.3. Consequently, 142 of the 150 non-proline H^N and ^{15}N backbone resonances for α-ADT and only 110 of 148 for the β-domain could be assigned.

Sequential connectivities were assigned using a collection of standard 3D NMR experiments (see ref[153] and references cited therein). The chemical shifts of amide ^1H and ^{15}N as well as C^α and C' resonances have been deposited in the BioMag-ResBank (`http://www.bmrb.wisc.edu`) under BMRB accession numbers 5930 (α-ADT) and 5936 (β-ADT).[154]

5.2.2 Topology of the Globular Part

An excellent correlation between the secondary chemical shifts of α-ADT and β-ADT was found, indicating a virtually identical structure of both domains in solution (see figure 5.4). It is well known that residues in homologuous proteins display similar secondary chemical shifts; this phenomenon has been used for structure predictions of proteins.[155] Both apical domains share a 63 % sequence identity. However, weaker correlations of the secondary chemical shifts are observed for some residues. In most cases, these outliers reflect non-conserved sidechain properties (e.g. Ser/Ala29, Ser/Val71, Val/Asp128, Val/Arg138) and might hence be attributed to local conformational variations.

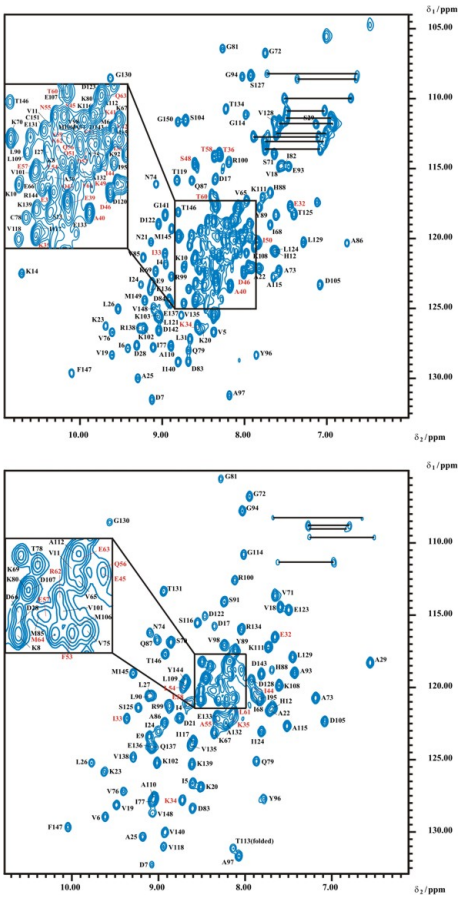


FIGURE 5.3: Annotated ^{1}H–^{15}N-HSQC spectra of α-ADT (TOP) and β-ADT (BOTTOM), both recorded at a temperature of 42 °C and a proton frequency of 600 MHz. Residues located in the protrusions are indicated with red labels. The regions of strong overlap are expanded for sake of clarity; the amide sidechain resonances of Asn and Gln residues are connectd by horizontal lines.

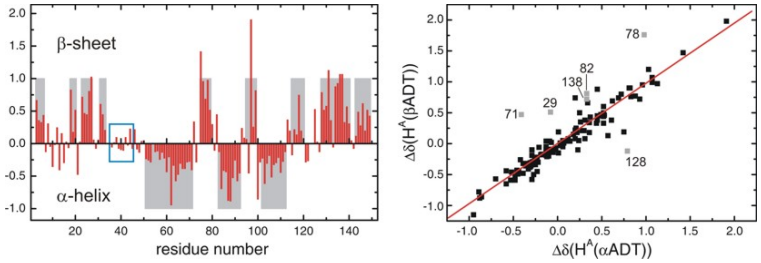


Figure 5.4: Left: Secondary H^{α} chemical shifts of α-ADT (red bars); the consensus CSI (C', C^{α}, C^{β}, H^{α}) is shown as grey outlined bars. Residues being part of the small α-helix in the crystal structure of the isolated α-domain as well as residues forming the β-strand in the closed thermosome are indicated by the blue box. Right: Correlation between the secondary H^{α} chemical shifts of α-ADT (horizontal axis) and β-ADT (vertical axis). Non-conserved residues with large deviations from the diagonal are colored grey and labeled with their sequence number.

Since resonance assignment was more complete for α-ADT, a detailed analysis of secondary structure and topology using ${}^{3}J_{H^{N}H^{\alpha}}$ coupling constants, the chemical shift index (CSI),[156,157] as well as characteristic H^{N} and H^{α} nuclear Overhauser effect (NOE) patterns was performed for the α-domain only. The results are in good agreement with the X-ray structure and are summarized in the following. β-strands S9, S10, the first part of S11, S17, S18, and S19 form an antiparallel β-sheet. Bulges consisting of two neighboring hydrophilic sidechains separate S10 from S11 and S17 from S19. In the core of the globular part, the second half of S11, S14, S15, and S16 form a parallel β-sheet. Two helices, H10 and H11, are amphiphilic, emphasized by a periodicity of ϕ angles, whereas H12 is embedded in the β-sandwich.

All of the residues located in secondary structure elements show very low amide hydrogen exchange rates (data not shown), with the exceptions of the first turns of helices H11 (Ile82–Met85) and H12 (Lys103–Ser104); these residues could not be assigned for β-ADT. Due to the good agreement between the topology of α-ADT and the topology found in the crystal, a complete struc-

ture determination including sidechain resonance and NOE assignment was not performed.

5.2.3 Overall Molecular Tumbling

Already in early stages of the analysis, comparing NMR spectra of the two apical domains revealed larger linewidths for β-ADT, pointing to a higher apparent molecular weight.

A possible oligomerization of the β-domain was supported by a first estimation of the rotational diffusion tensor, obtained from ^{15}N R_2/R_1 ratios on a per-residue basis assuming an axially symmetric diffusion tensor (see table 5.1). For α-ADT, the estimated $\tau_c = 9.00 \pm 0.08$ ns is in good agreement with a correlation time of 8.9 ns calculated from rigid body hydrodynamic modeling and is therefore consistent with a monomer. The correlation time for the β-domain was calculated to be 9.2 ns, which agrees well with 8.9 ns obtained for α-ADT, but not with $\tau_c = 13.31 \pm 0.09$ ns obtained for β-ADT from the experimental data. As evident from the F-values, no statistically significant improvent of the fits is achieved when a completely anisotropic diffusion tensor is used.

In order to assess whether the observed oligomerization is concentration-dependent, a 1 mM β-ADT sample was diluted by a factor of approximately three. Almost immediately after dilution, several new peaks arising from fragments were identified in the ^{15}N-HSQC; after two days, no properly folded protein could be detected.

Thus restricted by sample lifetime, only estimates of ^{15}N transverse relaxation rate constants could be obtained. R_2 rates for α-ADT $\left(10.3\,\mathrm{s}^{-1}\right)$, the diluted $\left(10.6\,\mathrm{s}^{-1}\right)$, and the concentrated β-ADT $\left(16.7\,\mathrm{s}^{-1}\right)$ sample clearly indicate a concentration-dependent, i.e. rather unspecific oligomerization of the β-domain. Indeed, β-ADT seems to exihibit some tendency to form lower oligomeric associates under low ionic strength conditions, as detected by analytical gel filtration. On the basis of these results, model-free analysis of backbone ^{15}N relaxation data was performed only for α-ADT, since the MF-analysis

Table 5.1: Diffusion tensor analysis for α- and β-ADT.

	τ_c (ns)[a]		D_{ratio}[b]		F	
Tensor[c]	α-ADT	β-ADT	α-ADT	β-ADT	α-ADT	β-ADT
isotropic	9.11 ± 0.02	13.22 ± 0.04	—	—	—	—
axially symmetric[d]	9.00 ± 0.08	13.31 ± 0.09	1.26 ± 0.09	1.42 ± 0.10	22.28	20.85
anisotropic[e]	9.03 ± 0.02	13.35 ± 0.05	1.31 ± 0.02	1.45 ± 0.03	0.03	1.82
optimized MF results[f]	9.10 ± 0.03	—	1.51 ± 0.02	—	—	—
hydrodyn. calc.[g]	8.90	9.22	1.78	1.92	—	—

[a] Rotational correlation time obtained from the relation $\tau_c = 1/(6D_{iso})$.

[b] Ratio of the components of the diffusion tensor.

[c] Anisotropy of the diffusion tensor.

[d] $D_{ratio} = D_{\parallel}/D_{\perp}$.

[e] $D_{ratio} = 2D_{zz}/(D_{xx} + D_{yy})$.

[f] Optimized model-free results.

[g] Results from hydrodynamic calculations.

for an axially symmetric diffusion tensor requires structural information in form of a pdb coordinate file.

During the course of diffusion tensor optimization, τ_c for α-ADT changed only slightly to $\tau_c = 9.10 \pm 0.03$ ns. As expected, the long axis of the diffusion tensor points towards the protrusion, approximately parallel to H10. The optimized ratio of the parallel and perpendicular components, $D_\parallel / D_\perp$, is 1.51 ± 0.02, compared to 1.78 obtained from the hydrodynamic calculation. This discrepancy can be explained by a dynamic behaviour of the protrusion, given that motions with large amplitudes alter the shape of the molecule and hence influence $D_\parallel / D_\perp$. It should be mentioned that the apical domains of the thermosome may represent a limiting case for the applicability of the MF formalism, since overall molecular tumbling and internal motions of the protrusion might be correlated (cf. page 29).

5.2.4 ^{15}N Backbone Motions

Analysis of ^{15}N relaxation data for 91 residues was performed using the MF approach. The upper panel of figure 5.5 shows a plot of the squared order parameter S^2 against the sequence number of α-ADT. Motions in the globular part of the protein are largely restricted, with an average squared order parameter $\langle S^2 \rangle$ of 0.85 ± 0.01 for all secondary structure elements except for the protrusion. Assuming free diffusion within a cone as motional model, this value of $\langle S^2 \rangle$ corresponds to a semi-cone angle of approximately 19° (see page 30). Moreover, the majority of residues were assigned to model 1 (cf. section 2.3.2). Residues with decreased S^2 in the globular part, indicating higher flexibility, are located in loop regions or in regions that are in contact with the intermediate domain in the complete thermosome.

In contrast to the globular part, the protrusion region exhibits a more complex dynamic behaviour. Noteably, all residues from Lys34 throughout Asn59 were assigned to model 5 (see lower panel of figure 5.5), with two distinct dynamic processes on different time scales: a picosecond mobility corresponding

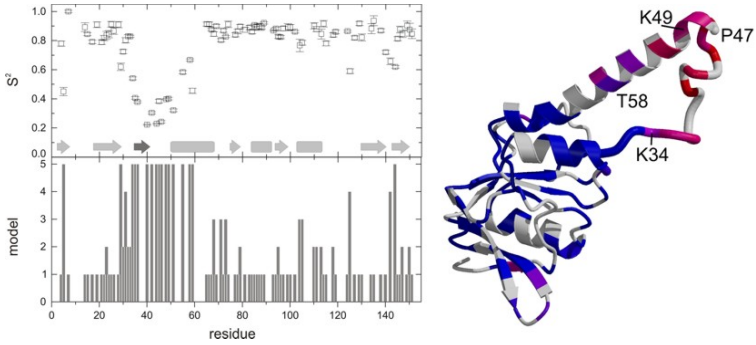


Figure 5.5: Left, upper panel: Plot of S^2 *vs.* the sequence number of α-ADT. Lower panel: Model selection *vs.* the sequence position for α-ADT. Secondary structure elements are indicated as bars and arrows for α-helices and β-strands. The dark grey arrow symbolizes the additional β-strand found in the closed thermosome. Right: S^2 mapped onto the structure of α-ADT. Some residues are indicated for orientation. The color ramp ranges from blue ($S^2 = 0.8$) to red ($S^2 = 0.2$).

to fast librations of the ^{15}N–^{1}H bond vector, and a slow nanosecond mobility implying dihedral transitions and therefore conformational variability.

Similar to the secondary structure elements, the fast motion in the protrusion is rather restricted with an average S_f^2 of 0.74 ± 0.01, corresponding to a semi-cone angle θ of roughly 25°, if the fast motion is represented by free diffusion in two axially symmetric cones. However, the values of S_s^2 decrease when moving outward from the globular part towards the tip of the protrusion (see figure 5.6), indicating that mobility on a ns time scale is less restricted. For the two-site jump model depicted in figure 2.11, ϕ increases from 39° for Lys34 ($S_s^2 = 0.70 \pm 0.01$) to 67° for Asp46 ($S_s^2 = 0.34 \pm 0.02$). In addition to Asp46, Ala40 and Ile44 display extremely low S_s^2 values of 0.31 ± 0.01 and 0.30 ± 0.01, corresponding to ϕ-angles of approximately 75°.

From the tip of the protrusion, ϕ decreases again from 60° for Gln51 ($S_s^2 = 0.43 \pm 0.01$) to 29° for Thr58 ($S_s^2 = 0.64 \pm 0.04$), consistent with a fraying of H10

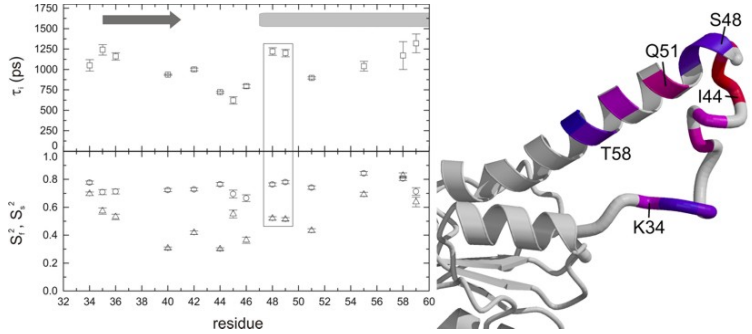


FIGURE 5.6: LEFT, UPPER PANEL: Plot of $\tau_i = \tau_s$ *vs.* the sequence number of α-ADT. LOWER PANEL: Squared order parameters of the fast and slow motion plotted *vs.* the sequence position for α-ADT. Ser48 and Lys49 are highlighted with a box. The dark grey arrow symbolizes the additional β-strand found in the closed thermosome, while H10 is indicated by the light grey bar. RIGHT: Detailed view of slow internal correlation times τ_s mapped onto the structure of α-ADT. The color ramp ranges from blue ($\tau_s = 1400$ ns) to red ($\tau_s = 600$ ns).

in the first turns; this fraying is confirmed by amide hydrogen exchange rates (data not shown). The C-terminal part of H10 is stabilized by interactions with H11 and the long loop between H12 and S17, and thus behaves like a rigid secondary structure element. This is reflected in high S^2 values of approximately 0.9 found for residues Val65–Ser71.

The pattern of S_s^2 and ϕ values observed for residues Lys34–Asn59 indicates that the ^{15}N–^{1}H bond vectors in the proximal part of the protrusion sample a wide conformational space. Therefore, a helix, as found in the crystal structure for residues Lys35–Gln43, with virtually no drop in the *B*-factors, is either short-lived on the NMR chemical shift time scale or not present at all.

The tip of the protrusion itself is rather restricted, as indicated by higher S_s^2 values for Ser48 ($S_s^2 = 0.52 \pm 0.01$) and Lys49 ($S_s^2 = 0.51 \pm 0.01$). These residues are located in the $i + 2$ and $i + 3$ positions of a βI-turn; hence, their ^{15}N–^{1}H bond vectors are fixed by hydrogen bonds. The described pattern of

mobility becomes more evident when looking at the correlation times for the slow internal motion τ_s (see figure 5.6). The values of τ_s decrease from the N and C-terminal residues towards the tip of the protrusion, with significantly higher τ_s for Ser48 and Lys49.

Although a detailed motional model for the protrusion region cannot be given, these results indicate a flexible protrusion with partial unwinding of the N-terminal part of H10.

5.3 Discussion

5.3.1 X-ray *versus* NMR Data

It should be emphasized that *no* structure calculations for the apical domains were performed during the course of the present work, mainly for two reasons: First, sample stability was considered too low for lengthy NOESY experiments, and second, all secondary structure indicators as well as the NOE-derived topology for the globular part were in excellent agreement with the crystal structure.

For the protrusion region, the main subject of interest, the NMR data are at variance with the corresponding crystal structures. In the X-ray structure of the isolated α-domain, Lys35–Asn43 form a small α-helix, presumably stabilized by formation of a four-helix bundle between symmetry-related molecules. The corresponding segment in β-ADT forms an extended β-stretch, also stabilized by packing interactions. In the closed thermosome, these residues form a β-stretch contributing to the circularly closed β-sheet forming the lid (see figure 5.1). In contrast, no evidence for regular secondary structure elements for this segment was found in solution, although the existence of a low-populated helix cannot be excluded completely. As mentioned earlier, β-ADT could not be analyzed to the same extend than α-ADT; however, since the protrusion region is highly conserved between both subunits, it is assumed that this region in both domains behaves identically in solution. Significant variances between

crystal and solution structures have been described in the literature. For example, a comparison of X-ray and NMR structures of the cystein-proteinase inhibitor chicken egg cystatin revealed a highly conserved segment being clearly α-helical in the crystal, but largely disordered in solution.[158, 159, 160]

5.3.2 Intrinsic Disorder and Flexibility

The present work revealed that the protrusion regions of the apical domains of the thermosome and group II chaperonins show an extraordinary high degree of conformational flexibility. For instance, helix H10 undergoes partial unwinding in the absence of crystal packing interactions. This is in good agreement with the crystal structure of the CCTγ apical domain, where an increase of *B*-factors from the base to the tip of the protrusion was observed.[147] In this protein, high mobility in the N-terminal stem led to a complete loss of electron density for residue Lys34–Asp49. Flexibility in the CCTγ apical domain might even be more pronounced than in arachaeal homologues, since the DPSK/DPSM βI-turn motif found in archaea is replaced by the extremely charged sequence REED.

In the apical domains of the thermosome, the Pro-induced turn is found in all crystal structures as well as in solution. A sequence alignment of group II chaperonins has shown that all considered archaeal,[131] but none of the eukaryotic protrusions (with the execption of CCTα) feature a motif with β-turn propensity.[139, 144, 161, 162]

Structural plasticity has been assigned a crucial role in substrate binding by GroEL. The substrate binding groove adopts different conformations in the unliganded form as well as in in the complex with various peptide ligands.[151] Elevated *B*-factors of the surrounding structure elements further support a structural plasticity of this region.[141, 163, 164] The present NMR studies confirm a structural plasticity of the protrusion region, thereby suggesting intrinsic disorder. This is especially true for the N-terminal part in the protrusion of the thermosome, and possibly of all group II chaperonins. Intrinsic disorder is rec-

ognized to play an important functional role, in particular for molecular recognition.[165,166]

5.3.3 Implications for Substrate Binding

While flexibility and intrinsic disorder are likely to facilitate interactions with a wide range of substrates, there is another remarkably feature in the protrusions: In the crystal structure of isolated β-ADT, the sidechain carboxy group of Glu57, a completely conserved residue in all group II chaperonins, forms a hydrogen bond to the amide group of the highly conserved Lys34 on the opposite branch of the protrusion. This interaction contributes to a plinth-like stabilization of the base of the N-terminal part of the protrusion, which would be expected to be flexible without this hydrogen bond. This is most probably the reason why resonances up to Lys35 could be assigned for β-ADT. NMR relaxation data for the α-apical domain show that Ile33 and, to a lesser extend, Lys34 are more rigid than the remainder of the N-terminal stem of the protrusion. These observations suggest the existence of a similar stabilization of the proximal protrusion base in solution.

Strikingly, sequence alignments of a number of archaeal and eukaryotic protrusions have revealed that similar contacts are possible for all proteins examined.[139,144,161,162,131] In the open state of the chaperonin, the protrusions are exposed at the periphery of the complex like antennas and are well separated from each other; hence, it is assumed that the structure of the apical domains in solutions is most relevant to the open state. The large, hydrophobic surface patches previously found in group II chaperonins would be completely exposed to the surrounding medium in order to recognize and bind a wide range of non-natively folded substrates.[144] In addition, the conserved βI-turn at the tip of the protrusion might further assist substrate binding.[167]

A number of conserved residues in the apical domains of group II chaperonins have been implicated in substrate binding.[140,144,147,168,169,170] Residues that are accessible from the interior of the torus are located in three regions: in

the protrusions, preceeding and in the first turns of helices H11 and H12, or in the loop connecting strands S9 and S10; they comprise hydrophilic as well as hydrophobic residues. It seems plausible that more than one substrate binding site exists, and that implicated residues are involved at different stages of the binding event and folding cycle.

Further biochemical as well as structural studies, in particular the investigation of substrate binding of group II chaperonins, will show whether the protrusion is indeed involved in substrate recognition and binding, or whether its only function is to form a lid in order to close the folding chamber.

Chapter 6

Effects of Nucleotide Binding on Backbone Motions in KdpBN

The structure of *apo*-KdpBN and the model of *holo*-KdpBN were determined by M. Haupt and Dr. M. Coles as part of a cooperation with the group of Prof. K. Altendorf, Universität Osnabrück. The protein samples were prepared by Dr. M. Bramkamp.

6.1 Introduction

P-type ATPases are ubiquitous proteins that catalyze the ATP-dependent transport of mono- and divalent cations, including protons, alkali, and heavy metal ions, across membranes.[171] The cytoplasmic loops of P-type ATPases are comprised of three different domains.[172] The P-domain contains an invariant aspartic acid residue which is phosphorylated during ATP hydrolysis. The A-domain mediates dephosphorylation of this Asp, while the N-domain is re-

sponsible for nucleotide binding. In contrast to the P- and A-domains, which are of similar size and fold, the nucleotide binding domains of P-type ATPases differ significantly.

The K^+ transporting KdpFABC complex of *E. coli* has a unique position in all classes of P-type ATPases, regarding both structural and functional aspects. KdpFABC acts as an emergency K^+ uptake system, which is only expressed upon osmotic stress or K^+ limitation and is rapidly degraded when no longer required. Generally, a central catalytic subunit is responsible for ion transport as well as ATP hydrolysis in P-type ATPases. In the KdpFABC complex, phosphorylation and hydrolysis of ATP occurs in the KdpB subunit, whereas K^+ binding and transport are associated with KdpA.[173] A matter of ongoing discussion is how the energy derived from ATP hydrolysis is transferred from KdpB to KdpA.

In our group, the structure of KdpBN, the nucleotide binding domain of KdpB, has been solved recently; furthermore, a detailed model of nucleotide binding using AMP-PNP has been established.[174] KdpBN has a nearly spherical shape and consists of a six-membered, anti-parallel β-sheet flanked by four α-helices on either side and a small 3_{10}-helix (see figure 6.1). Binding of AMP-PNP, a non-hydrolyzable ATP analogue, occurs in the cleft formed by the β-sheet and restricted by helices α1 and α2. A dissociation constant K_d of 1.4 mM indicates weak binding and fast exchange between free and bound forms.

Recently, the structures of the Na^+,K^+-ATPase N-domain in the *apo* and *holo* form with AMP-PNP as ligand have been published.[175] A large-scale conformational rearrangement was found upon ligand binding, resulting in a reorientation of the hinge region between the N- and P-domains. Furthermore, for 29 residues, multiple peaks were found in the NMR spectra, indicating a slow chemical exchange process. No resonances were observed at all for 11 residues. In the ^{15}N,^1H-HSQC spectrum of KdpBN, several peaks are very intense and display sharp lines, while signals of other residues show a field-dependent line-broadening (see figure 6.2). These observations suggest the presence of internal motions on different time scales and were further investigated by analysis of

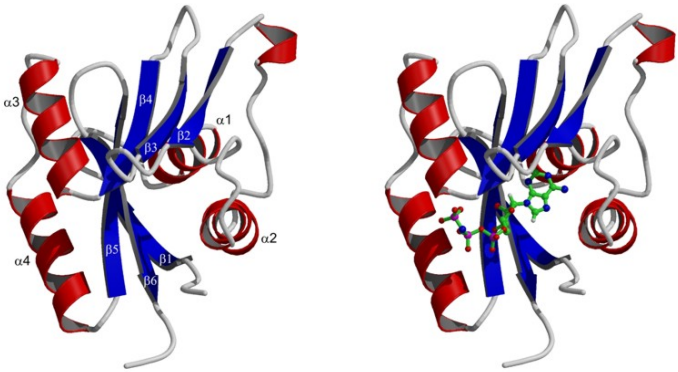


Figure 6.1: Left: NMR structure of the KdpB nucleotide binding domain. Secondary structure elements are indicated. Right: Detailed model of AMP-PNP binding. The binding site is formed by the β-sheet and restricted by helices α1 and α2.

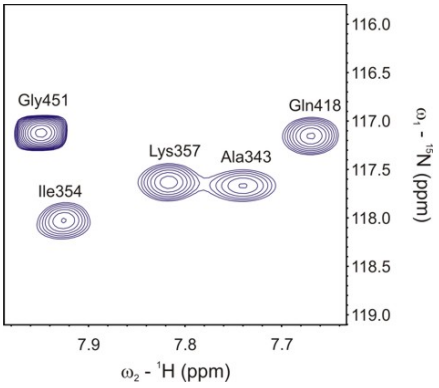


Figure 6.2: Expanded region from the ^{15}N,^{1}H-HSQC of KdpBN illustrating different linewidths in the ^{15}N dimension. Note the smaller linewidth for Ala343 and the very intense peak for Gly451, which is located at the C-terminus.

^{15}N relaxation data according to the MF formalism as well as relaxation dispersion experiments for backbone nitrogen sites.

In order to characterize changes in molecular motions upon ligand binding, ^{15}N R_1, R_2, and hetNOE experiments were recorded in the absence and presence of a five-fold excess of AMP-PNP. Since the structure determination of *holo*-KdpBN is currently in progress, the MF analysis of data for KdpBN in presence of AMP-PNP was perfomed using the coordinates of the *apo*-structure.

6.2 Results

6.2.1 Overall Molecular Tumbling

A plot of the R_2/R_1 ratio for KdpBN in the absence of nucleotide *vs.* the sequence number shows significantly higher values for residues Val398–Ala407 and Thr414–Arg425 (see figure 6.3). These residues are located in helices α3 and α4, which are aligned nearly parallel to each other. Whereas all residues in α4 display elevated R_2/R_1, the ratio drops for residues near the C-terminal end of helix α3, which might be explained by a slight kink. Thus, the R_2/R_1 data suggest that overall tumbling of KdpBN is anisotropic, with the long axis of the diffusion tensor oriented parallel to the axis of helix α4.

The proposed anisotropy was confirmed by an estimation of the diffusion tensor. Overall tumbling was found to be axially symmetric with $D_\parallel$ oriented nearly parallel to the helix axis of α4. Assuming full anisotropy did not result in a significant improvement, as indicated by the F-value (see table 6.1). The optimized results for τ_c and $D_\parallel/D_\perp$ are in good agreement with the estimations. A rotational correlation time of 11 ns at 27 °C indicates that KdpBN tumbles as a monomer at a concentration of ~ 1 mM.

For KdpBN in presence of AMP-PNP, initial estimates of the rotational diffusion tensor resulted in a significantly smaller correlation time: $\tau_c = 10.48 \pm 0.06$ ns, whereas a value of 11.01 ± 0.02 ns was obtained from the MF analysis for *apo*-KdpBN. A closer inspection of the data revealed the R_1 values to

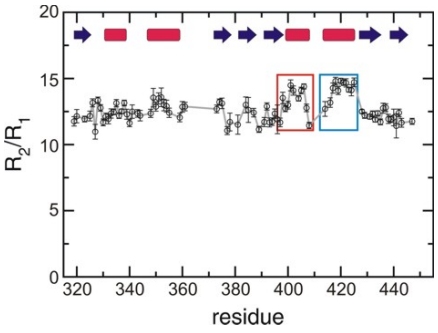


Figure 6.3: Plot of R_2/R_1 *vs*. the sequence position of KdpBN. Helices α3 and α4 are highlighted by red and blue boxes, respectively; secondary structure elements are indidcated by red bars (α-helices) and blue arrows (β-strands).

Table 6.1: Diffusion tensor analysis for KdpBN.

Tensor[a]	τ_c (ns)[b]	D_{ratio}[c]	F
iso	11.12 ± 0.01	—	—
ax-sym[d]	11.18 ± 0.02	1.17 ± 0.01	12.13
aniso[e]	11.19 ± 0.01	1.17 ± 0.01	0.15
opt[f]	11.01 ± 0.02	1.22 ± 0.01	—

[a] Anisotropy of the diffusion tensor.
[b] Rotational correlation time obtained from $\tau_c = 1/(6D_{iso})$.
[c] Ratio of the components of the diffusion tensor.
[d] $D_{ratio} = D_\parallel / D_\perp$.
[e] $D_{ratio} = 2D_{zz} / (D_{xx} + D_{yy})$.
[f] Optimized model-free results.

be increased by approximately 10 %. This discrepancy could be attributed to a 2 °C difference in sample temperature, since data for *apo*- and *holo*-KdpBN were acquired on different spectrometers. The observed difference in temperature translates into a decrease of τ_c by 0.5 ns. Bearing this difference in mind, the estimated correlation time of *holo*-KdpBN is in good agreement with the optimized value for the *apo*-form, and the diffusion tensor was also found to be axially symmetric. MF analysis of the ^{15}N relaxation data for *holo*-KdpBN yielded $\tau_c = 10.43 \pm 0.02$ ns and $D_\parallel / D_\perp = 1.16 \pm 0.01$, which compare favourably to the values obtained for *apo*-KdpBN. However, a slight difference in the orientation of the diffusion tensor between the *apo* and *holo* form is observed; this might reflect possible minor structural changes, or might be a consequence of the small difference in sample temperature, although this seems to be unlikely.

6.2.2 Motions on a Pico- to Nanosecond Time Scale

Internal motions in the secondary structure elements of *apo*-KdpBN are highly restricted, with an average squared order parameter of 0.90 ± 0.01. If free diffusion within a cone is assumed as a motional model, this value of $\langle S^2 \rangle$ corresponds to a semi-cone angle of approximately 15°. Regions with higher internal mobility are found in the loops connecting helix α2 and the 3_{10} helix (Arg363–Leu370), between helices α3 and α4 (Asn408–Phe412), and at the C-terminus (Ile448–Gly451); these residues were all assigned to model 5. When a two-site-jump is assumed to represent the slower motion in addition to the fast diffusion within a cone, the average angle $\langle \phi \rangle$ between both cones is 27° for equal populations; a minimum value for S_s^2 of 0.25 is obtained when $\phi = 90°$. Ala356, a residue in helix α2 assigned to model 5, has a S_s^2 of approximately 0.4, corresponding to $\phi \approx 65°$. Although Ala356 is located in a secondary structure element, its amide bond vector exhibits are large mobility on a slow time scale ($\tau_s = 1591.6 \pm 86.3$ ps). This might be attributed to the fact that helix α2 is slightly bent, thereby weakening the hydrogen bond A356H^N→I352C′. For the C-terminal residues, a two-site-jump does not represent a valid motional mo-

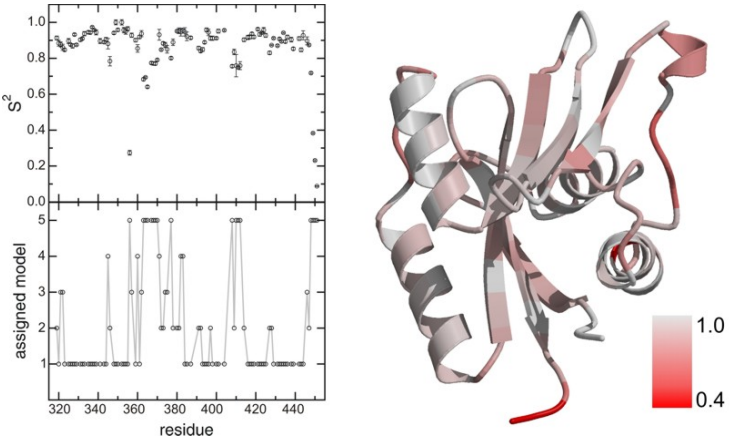


Figure 6.4: Left, upper panel: Plot of S^2 *vs.* the sequence position of KdpBN. Note the unusually slow value for Ala356. Left, lower panel: Model assignment for KdpBN. In addition to the regions mentioned in the text, Ala356 and Phe377 were assigned to model 5. Right: The squared order parameter mapped onto the structure of KdpBN; the color ramp from $S^2 = 1.0$ to 0.4 is indicated.

del; the S^2_s decreases from 0.81 for Ile448 to 0.15 for Gly451, indicating almost unrestricted motions.

As expected, the majority of residues shows no alterations in their relaxation rates upon ligand binding (see figure 6.5; R_1 and ăR_2 were scaled in this figure to account for the 2 °C difference in sample temperature). Nevertheless, a number of sites are affected. With the exceptions of His371 and Gly427, all residues are clustered around the binding site. The changes in relaxation rates point towards a higher mobility of the respective ^{15}N–^{1}H bond vectors; only Thr346 and Gly427 seem to be more restricted in presence of AMP-PNP.

Not all of the differences observed for the relaxation rates are reflected in the squared order parameter. The upper panel of figure 6.6 shows a plot of the differences in S^2 between *apo*- and *holo*-KdpBN, $\Delta S^2 = S^2(\text{apo}) - S^2(\text{holo})$, *vs.*

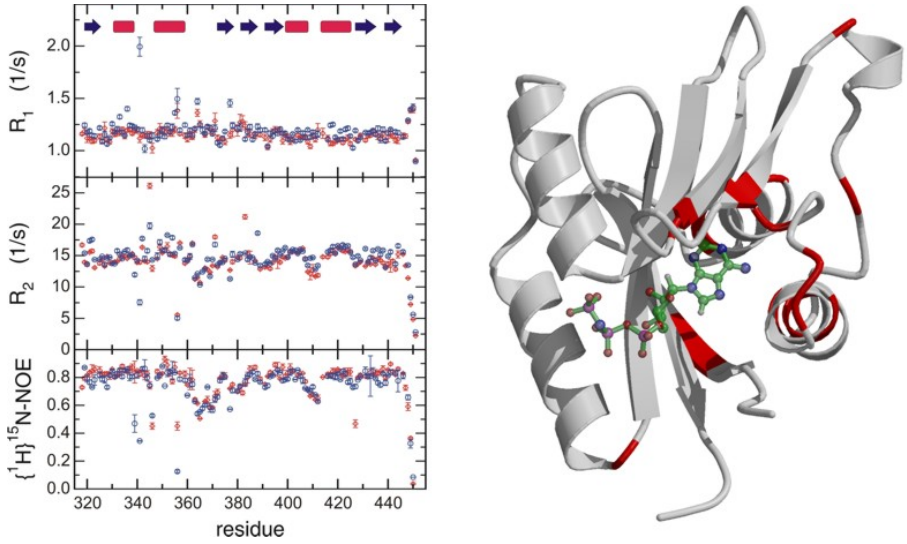


Figure 6.5: Left: Plots of longitudinal (top) and transverse (middle) relaxation rates and the heteronuclear NOE (bottom) for KdpBN in absence (red) and presence (blue) of a five-fold excess AMP-PNP. Note that R_1 and R_2 were scaled for this figure to compensate for the temperature difference. Right: Model of nucleotide binding for KdpBN. Residues with relaxation rates affected by ligand binding are colored red.

the sequence number of KdpBN. It is interesting to note that some residues become more flexible, while motions of other residues seem to be more restricted in the presence of the nucleotide.

For Phe377 and Met383, ΔS^2 is positive, pointing to a higher degree of mobility in the *holo* form, while the internal correlation time for the slow motion is nearly unaltered. In the case of Phe377, an increase in flexibility is clearly evident from the relaxation rates. These residues are located in the loop connecting β-strands $\beta 2$ and $\beta 3$, with Met383 being involved in two hydrogen bonds, M383H$^{\mathrm{N}}\rightarrow$T378C′ and T378H$^{\mathrm{N}}\rightarrow$M383C′. A possible explanation for the increased mobility of Phe377 and Met383 is a weakening of these hydrogen bonds upon interaction with AMP-PNP.

The first residues immediately following helix $\alpha 1$, Leu339 and Ala340 as well as Ser341, also tend to be more flexible in the complex. During MF analysis of the *apo* form data, both Leu339 and S341 have been assigned to model 1. In *holo*-KdpBN, these residues were fitted by model 5 and 4, respectively. Helix $\alpha 1$ is located at the inner end of the binding cleft, with its C-terminal part representing the "rear wall". The sidechain of Leu339 participates in hydrophobic interactions with Ile386 in β-sheet $\beta 3$ in the structure of *apo*-KdpBN; these interactions might be weakened upon nucleotide binding.

For Ala356, a moderate change of S^2 is observed ($\Delta S^2 = 0.05 \pm 0.01$), although the value of the hetNOE in *holo*-KdpBN is significantly lower. While τ_s remains constant, S_f^2 is increased from 0.70 ± 0.03 to 0.83 ± 0.01; in contrast, S_s^2 is reduced from 0.39 ± 0.03 to 0.27 ± 0.01. Assuming a two-site-jump as model for the slow motion, this value is close to the theoretical minimum of 0.25. The decrease in S_s^2 corresponds to an increase of ϕ from 65° to 82°, illustrating the large flexibility of this particular ^{15}N–^{1}H bond vector. A reason for the increased mobility of Ala356 in *holo*-KdpBN might be that the bend in helix $\alpha 2$ is more pronounced in the *holo* form, but this question will be answered once the structure of the complex has been determined.

In contrast, negative values for ΔS^2 are found for Thr346, which is located in one of the two loops surrounding the binding cleft, as well as for residues

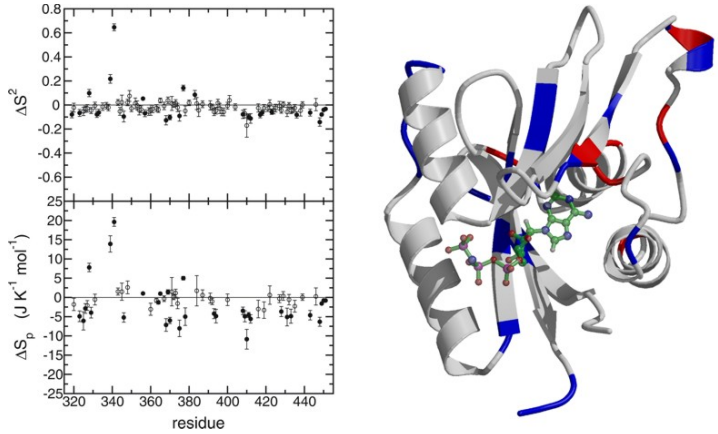


FIGURE 6.6: LEFT: Plots of ΔS^2 (UPPER PANEL) and per-residue $\Delta S_{p,i}$ (LOWER PANEL) *vs.* the sequence number of KdpBN. Changes are considered to be significant if the error is less than half of the corresponding parameter value. In both panels, significant changes are indicated by filled circles. RIGHT: Significant $\Delta S_{p,i}$ mapped onto the structure of KdpBN; residues with positive and negative values for $\Delta S_{p,i}$ are colored red and blue, respectively.

Asn408 through Phe412. Notably, this stretch is one of the more flexible regions in *apo*-KdpBN and was assigned to the extended LS model. The correlation times of the slow internal motion, τ_s, are reduced by a factor of ~ 16 for Gly410 through Phe412. In the other flexible region found in the *apo* form, Arg363 through Leu370, only Gln368 and Leu370 have negative ΔS^2, while the values of τ_s for the *apo* and *holo* form are approximately the same.

Order parameters (S^2) are related to conformational entropy (ΔS_p), since they contain information about the different states accessible to bond vectors. If bond vector motions are described by the diffusion-in-a-cone model, the total conformational entropy difference ΔS_p between two states of a protein is given

by the sum over the individual contributions of all bond vectors i:[176]

$$\Delta S_{\mathrm{p}} = -k_{\mathrm{B}} \sum_i \ln \left[\frac{3 - \sqrt{1 + 8\, S_{\mathrm{B},i}}}{3 - \sqrt{1 + 8\, S_{\mathrm{A},i}}} \right] \tag{6.1}$$

where $S_{\mathrm{A},i}$ and $S_{\mathrm{B},i}$ are the LS order parameters (*not* S^2) of the bond vector i in state A and B, respectively. Equation 6.1 is only valid as long as $S^2 \leq 0.95$. In the lower panel of figure 6.6, a plot of the individual $\Delta S_{\mathrm{p},i}$ *vs.* the sequence number of KdpBN is shown. Significant $\Delta S_{\mathrm{p},i}$ (see caption of figure 6.6) were found for 32 residues, corresponding to an overall change in conformational entropy of $-67.1 \pm 38.8\,\mathrm{J/(K \cdot mol)}$.

6.2.3 Slow Motions on a Millisecond Time Scale

In order to investigate molecular motions on a slow, millisecond time scale, CPMG relaxation dispersion profiles were recorded at static magnetic fields of 14.1 and 17.6 T; representative plots of dispersion profiles are shown in figure 6.7. The data were subsequently fitted to a flat line, the fast exchange approximation (equation 3.11), and to the general expression for a two-site exchange (equation 3.3), and F-statistics were used to decide which equation described the experimental data most appropriately (see section 3.3). Chemical exchange was considered to be significant if $R_{\mathrm{ex}} > 1\,\mathrm{s}^{-1}$. Using this protocol, 42 residues subject to chemical exchange were identified in *apo*-KdpBN, with an average exchange rate constant $\langle k_{\mathrm{ex}} \rangle = 881 \pm 40\,\mathrm{s}^{-1}$. In *holo*-KdpBN, 63 residues were found to be affected by an exchange process with $\langle k_{\mathrm{ex}} \rangle = 414 \pm 36\,\mathrm{s}^{-1}$.

For residues subject to chemical exchange, the data were refitted assuming that all residues are affected by the same process, and the average value of k_{ex} was used. For residues fitted to the full equation, averaged populations of the major populated site $\langle p_{\mathrm{a}} \rangle$ of 0.98 ± 0.01 and 0.96 ± 0.01 were obtained for the *apo* and *holo* form, respectively.

In the absence of ligand, residues with R_{ex} contributions cluster around the nucleotide binding cleft (see figure 6.8). In addition, exchange contributions

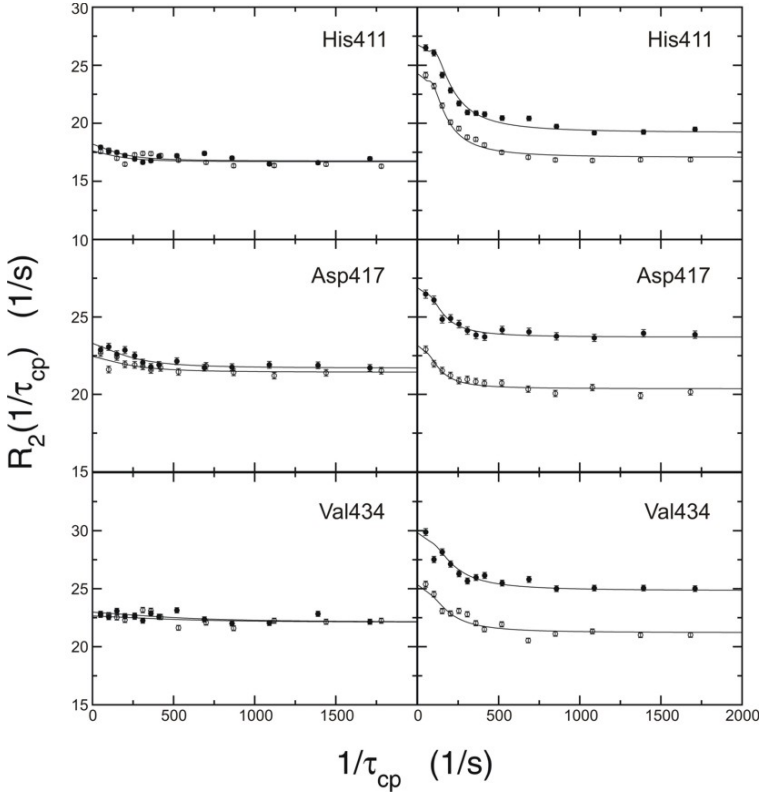


FIGURE 6.7: CPMG relaxation dispersion profiles for selected residues in *apo*- (LEFT) and *holo*-KdpBN (RIGHT). All data were fitted globally. The profiles for His411, Asp417, and Val434 in the absence of ligand were best fitted to the fast-exchange approximation, yielding $R_{ex} = 1.47 \pm 0.23\,s^{-1}$, $1.61 \pm 0.20\,s^{-1}$, and $0.87 \pm 0.18\,s^{-1}$, respectively. Upon interaction with AMP-PNP, all three residues had to be fitted against the general expression for a two-site exchange process, as indicated by α-values of 0.21, 0.59, and 0.84. R_{ex} values obtained for His411, Asp417, and Val434 are $7.57 \pm 1.21\,s^{-1}$, $3.19 \pm 0.56\,s^{-1}$, and $4.98 \pm 0.69\,s^{-1}$, respectively.

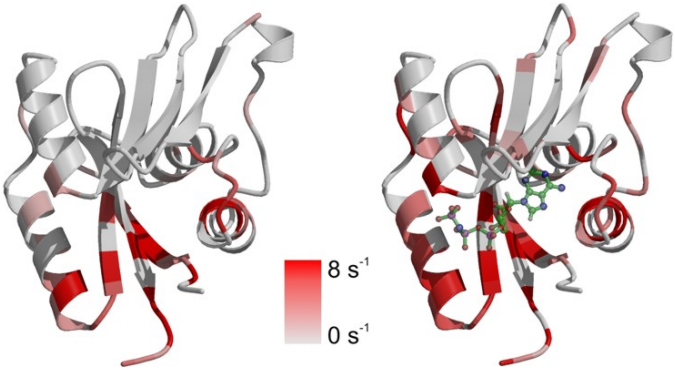


Figure 6.8: R_{ex} values obtained from fits using $\langle k_{ex} \rangle$ mapped onto the structure of *apo*-KdpBN (Left) and the model of *holo*-KdpBN (Right). Note that residues affected by chemical exchange are clustered around the nucleotide binding site in the *apo* form. When nucleotide is bound, several residues distant from the binding cleft are also affected by an exchange process.

are found for a number of sites in helix α4. In the nucleotide binding domain of the Na^+,K^+-ATPase,[175] two signal sets or no resonances at all were observed for a large number of residues located in the long loop preceeding helix α1 and in the N-terminal part of helix α2. In KdpBN, Thr332 and Ala334 in helix α1 as well as Ser341, the last residue of helix α1, show R_{ex} contributions. In the long loop, only Ala325 and Gln326 are affected by chemical exchange. In contrast, virtually all residues in strand β1 are subject to an exchange process. Helix α2 in KdpBN also displays chemical exchange, and multiple signal sets were detected for corresponding residues in the Na^+,K^+-ATPase N-domain. A slow dynamic process is also observed for individual sites after helix α5 (corresponding to α4 in KdpBN), as well as in strands β5 and β6. As is evident from figure 6.8, the corresponding regions in the N-domain of KdpB participate in an exchange process as well.

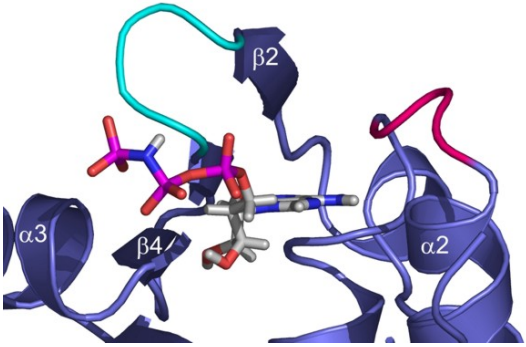


FIGURE 6.9: Detailed view of the nucleotide binding site of KdpBN with loops connecting helices $\alpha 1$ and $\alpha 2$ (red) and strands $\beta 2$ and $\beta 3$ (cyan). Both loops are in direct proximity to the binding cleft and the nucleotide, but only the loop between $\alpha 1$ and $\alpha 2$ (red) is flexible on a millisecond time scale. This figure was created with PYMOL.[177]

For complexed KdpBN, chemical exchange affects a larger number of residues. In addition to residues in proximity to the binding site, the N-terminal winding of helix $\alpha 2$, nearly all residues in helix $\alpha 4$, and the C-terminal part of strand $\beta 5$ show significant R_{ex} contributions.

Interestingly, the loop connecting β-strands $\beta 2$ and $\beta 3$ in direct proximity to the binding site is not subject to chemical exchange (see figure 6.9), neither in the absence nor in the presence of ligand. In the N-domain of the Na^+,K^+-ATPase, no resonances were observed for two residues directly following strand $\beta 2$, indicating excessive linebroadening from chemical exchange occuring on a millisecond time scale. The difference in chemical shift between the two sites A and B, $\Delta\omega$, provides structural information about the exchange process. In some cases, $\Delta\omega$ has been interpreted using chemical shifts for an alternative conformation of the protein, e.g. *apo* or ligated form,[178,179] or using random coil chemical shifts for disordered conformations.[180,14,181] It should be noted that the *sign* of $\Delta\omega$ is important for a correct interpretation in terms of

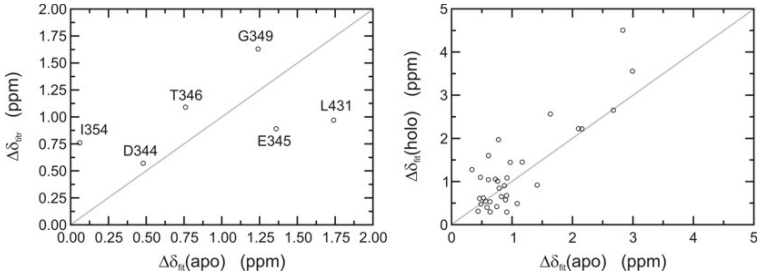


Figure 6.10: Left: Plot of the chemical shift differences obtained from the analysis of relaxation dispersion data for unligated KdpBN *vs*. the chemical shift differences for the protein in presence of a five-fold excess of AMP-PNP. Right: Correlation between the fitted values of $\Delta\omega$ for *apo*- and *holo*-KdpBN. In both panels, a grey diagonal is drawn as a guide.

structural changes; however, it is not obtained from the analysis of relaxation dispersion data, but has to be determined by other means.[181] The sign of $\Delta\omega$ was not determined in this work; nevertheless, the magnitude of $\Delta\omega$ may provide hints on the nature of the exchange process.

The left panel of figure 6.10 shows a plot of differences in chemical shifts obtained from the analysis of relaxation dispersion data for *apo*-KdpBN *vs*. the absolute values of those from titration experiments at a five-fold excess of AMP-PNP. In the case of fast exchange, $\Delta\omega$ was calculated from R_{ex} using the average values given above for k_{ex} and p_a. Although far from being perfect, a weak correlation is discernible, suggesting that conformational rearrangements on a millisecond time scale are necessary for nucleotide binding and are present even in the absence of ligand. Asp344, Glu345, and Thr346 are located in the loop immediately preceeding helix α2, forming some sort of a "lid" over the bindig cleft. This arrangement bears some resemblence to the "ATP-lid" of the Bergerat fold found in the GHKL superfamily of ATPases / kinases.[182] It can be envisioned that access to the nucleotide binding site is, at least partially, controlled by a motion of this lid.

For the differences in chemical shift of *apo*- and *holo*-KdpBN (right panel of figure 6.10), a better correlation is observed, suggesting that the conformations sampled by the exchange process in the *apo* form are not altered by ligand binding. Unfortunately, a structural model of the minor populated state cannot be provided. The small population ($\langle p_b \rangle \approx 3\,\%$) suggests that the minor conformer might represent an "excited" state, probably a partial unfolding of the protein; this will be clarified when the signs of $\Delta\omega$ have been determined for *apo*- and *holo*-KdpBN.

6.3 Discussion

The effects of nucleotide binding on fast and slow time scale backbone dynamics in KdpBN have been investigated by ^{15}N NMR relaxation techniques. Motions on a pico- to nanosecond time scale were characterized using the MF approach. In general, fast motions of the ^{15}N–^{1}H bond vectors are highly restricted in KdpBN with an average S^2 of 0.9 for the secondary structure elements. Higher mobility is found for the loop connecting helix α2 to the 3_{10} helix, for the 3_{10} helix itself, as well as for the loop between helices α3 and α4.

Upon nucleotide binding, S^2 values increase for most of the residues. Smaller S^2 are observed for Phe377 and Met383, located at the C-terminal end of strand β2 and the N-terminal part of strand β3. These strands are connected by a loop stabilized by two hydrogen bonds, M383H$^{N}\rightarrow$T378C′ and T378H$^{N}\rightarrow$M383C′. The aromatic ring of Phe377 is involved in π–π interactions with the purine ring of the nucleotide, possibly implying a rearrangement of the sidechain conformation, which in turn might lead to a weakening of the aforementioned H-bonds *via* changes in backbone conformations. The ^{15}N–^{1}H bond vector of Ala356 is also found to be more flexible in the presence of AMP-PNP, presumably due to a stronger pronounced bend in helix α2 at the postion of Ala356, leading to a loss of the H-bond A356H$^{N}\rightarrow$I352C′.

In contrast, an increase in S^2 is observed for residues Asn408 throughout Phe412, connecting helices α3 and α4. This loop is located on the opposite side

of the β-sheet with respect to the ligand binding site. It has been proposed that changes in coupled dynamic modes might be used to transduce energetic signals to distal regions on a protein.[183,18] Recently, long-range dynamic effects in sidechains have been reported for a PDZ domain and were attributed to a possible allostery.[184] Although a direct correlation to the binding site is not obvious for Asn408–Phe412 in KdpBN, the increased S^2 values of this loop might represent an allosteric signal.

Relaxation dispersion data has been used to probe motions on a millisecond time scale. In the absence of nucleotide, 42 residues were shown to be affected by a chemical exchange process, while 63 exchanging residues were found in *holo*-KdpBN. The regions for which R_{ex} contributions have been identified are in agreement with those for which slow dynamics were suggested in the N-domain of the Na^+,K^+-ATPase,[175] with one exception: whereas the loop between strands β2 and β3 seems to be subject to a slow exchange process in the Na^+,K^+-ATPase; in KdpBN, this loop is not affected by chemical exchange at all. In contrast, the loop connecting helices α1 and α2 displays R_{ex} contributions in the *apo* as well as in the *holo* form.

Assuming a two-site exchange process, the differences in chemical shifts between both states indicate that the process is not altered by ligand binding, although the average exchange rate constant is reduced by a factor of approximately 2. An average population of $\sim 3\,\%$ suggests that the minor state might represent an excited state. However, the signs of the chemical shift differences have not been determined; hence, a more detailed statement about the nature of the exchange process cannot be given.

Changes in S^2 upon conformational transitions, e.g. folding–unfolding of a protein, or upon ligand binding, reflect altered flexibility and hence changes in conformational entropy. It has been shown that order parameters can be used to obtain an upper estimate for the change in entropy or Gibb's free energy as a result of conformational restriction, if motions of the bond vectors are assumed to be uncorrelated to global or other internal motions.[185,186,187,188,189,176] Furthermore, if relaxation data for different temperatures are available, the contri-

butions of ps–ns time scale motions to local heat capacities can be estimated in addition.[190] It should be kept in mind, however, that a comprehensive characterization of conformational entropy requires information about motions of all bond vectors in the backbone and sidechains on time scales faster and slower than overall tumbling, as well as solvation entropy.

For KdpBN, changes in conformational entropy associated with nucleotide binding could be analyzed for 32 residues. ΔS_p is $-67\,\mathrm{J/(K\cdot mol)}$, corresponding to $-T\Delta S_\mathrm{p} \approx 20\,\mathrm{kJ/mol}$ at 27 °C. The positive value of $-T\Delta S_\mathrm{p}$ indicates that nucleotide binding is entropically unfavored. The main interactions between the nucleotide and the protein are a π–π interaction between Phe377 and the purine ring, as well as a cation–π interaction between the sidechain of Lys395 and the purine ring.[174] The energy contributions of these interactions were estimated to be -8 and $-23\,\mathrm{kJ/mol}$, respectively.[191,192] Thus, the "entropy penalty" of ligand binding is overcompensated by the π–π and cation–π interactions.

The measured dissociation constant of the complex,[174] $K_\mathrm{d} \approx 1.4\,\mathrm{mM}$, corresponds to a binding energy of roughly $-16\,\mathrm{kJ}$. Thus, an energy contribution of about $-5\,\mathrm{kJ}$ has to be gained from other sources, such as increased mobility of sidechains on a ps–ns time scale or increased millisecond dynamics. Indeed, more residues are subject to a chemical exchange process in *holo*-KdpBN (63 in the *holo* and 42 in the *apo* form, respectively). In principle, the temperature dependence of the exchange rate constant can be used to obtain information about the thermodynamics of the exchange process. However, relaxation dispersion data have not been recorded at various temperatures so far.

Although the structure determination of *holo*-KdpBN is still in progress, no indications for large conformational rearrangements upon nucleotide binding were found.[174] The results presented in this chapter support this notion, and a more precise interpretation of the backbone dynamics will be possible when the subtle changes in the structures of native KdpBN and the complex are known.

Chapter 7

Hydrogen Bonds in a Small Cyclic Pentapeptide

All solid state NMR experiments were conducted in cooperation with Prof. B. Reif, FMP Berlin. Fitting of CSA tensors from 1D spectra was done by Dr. M. Hologne, FMP Berlin; the cyclic pentapeptides used in this project were synthesized by Dr. M. Sukopp.

7.1 Introduction

One of the main obstacles in the process of drug design is the absence of reliable information on the three dimensional structure of the ligand within a ligand–receptor complex. Thus, it would be extremely useful to have a variety of "conformational templates", i.e. model ligands which should meet at least three requirements: (i) they should be conformationally restricted and possess only one 3D structure, (ii) they shoud be readily accessible synthetically, and

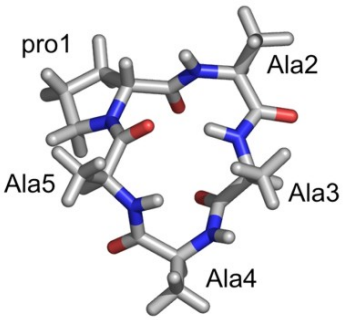


FIGURE 7.1: Structure of pA_4 as published by Kessler *et al.*[199] The βII′-turn with pro1 in the $i+1$ position is characterized by a hydrogen bond $A3H^N \rightarrow A5C'$, while a γ-turn centered about Ala4 would comprise a hydrogen bond between A3C′ and $A5H^N$.

(iii) they should be able to position the amino acid sidechains appropriately for interaction with the receptor.

Excellent candidates for such conformational templates are cyclic pentapeptides (CPPs).[193] During the last two decades, extensive experimental studies of the 3D structures of CPPs using NMR spectroscopy and X-ray crystallography have been performed.[194,193,195] Employment of CPPs as receptor probes was initiated by the Kessler group in the early nineties.[196,197] Based on extensive NMR studies, a "conformational template" of the *aBCDE* type was proposed (in the remainder of the work, D-amino acids are denoted by lower case letters). This template was thought to possess a single conformation characterized by a βII′-turn centered at the *aB* fragment, and a γ-turn at the *D* residue.[198]

It has soon been recognized, however, that even CPPs exist as different conformers, and that conformational averaging fast on the NMR chemical shift time scale leads to averaging of the NMR observables. Protocols for the refinement of structures based on NMR data have been developed, assuming

that the experimental observables arise from ensembles of conformations. In order to identify significant conformational changes fast on the NMR chemical shift time scale, the application of conformational restraints which average differently has been proposed.[200] This method was illustrated using the cyclic pentapeptide cyclo[-pro-Ala2-Ala3-Ala4-Ala5-], abbreviated as pA_4.[199] It was shown that the "upper half" of pA_4 represents a well-defined βII′-turn induced by the D-Pro in the $i+1$ postion. In contrast, four different conformations were found for Ala4, including a γ- and a γ'-turn. The NMR observables can only be described by averaging over an ensemble containing all conformations. Recently, Nikiforovitch *et al*. pointed out that the βII′γ model for CPPs would not be valid if the γ-turn was centered at an amino acid in L-configuration.[201] In order to gain a more detailed insight into the structure of CPPs, pA_4 was investigated using liquid-state as well as solid-state (ssNMR) NMR techniques.

During the past decade, NMR has been discovered as a tool for the characterization of hydrogen bonds in nucleic acids and proteins. Direct experimental evidence for such interactions can be obtained from magnetization transfer *via* scalar couplings between atoms separated by less than the sum of the van der Waals radii. These couplings were first detected in the early nineties by Summers *et al*.[202,203] They observed *J*-couplings ranging from 0.3 to 4 Hz between backbone amide protons hydrogen bonded to the sulphur atoms ligating the metal (^{113}Cd or ^{199}Hg) in metal-substituted rubredoxin from *Pyrococcus furiosus*. Remarkably, *J*-coupling between the protons of an alanyl methyl group adjacent to such an S atom and the metal was also observed. In the late nineties, direct observations of hydrogen bonds (H-bonds) *via* scalar couplings in nucleic acid base pairs using NMR were reported for the first time.[204] These $^{2}J_{NN}$ correlations were detected between the imino ^{15}N atom of the donor base and the acceptor ^{15}N atom of the complementary base; furthermore, additional *J*-couplings were found between the imino hydrogen itself and the H-bond accepting ^{15}N nucleus.[205] These discoveries were almost followed immediately by the observation of H-bonds in proteins *via* interresidue $^{3h}J_{NC'}$ scalar couplings.[206,207] While the ^{h}J-couplings observed between the H-bond donating

and accepting Watson-Crick base pairs were surprisingly large (6–7 Hz), significantly smaller values are found for the N−H·····O=C hydrogen bonds in proteins; they fall in the range of 0.3–0.5 Hz. During the following years, a lot of effort has been made to characterize this phenomenon in greater detail. A number of theoretical articles have been published,[208,209,210] and the correlation between $^{3h}J_{NC'}$ and the hydrogen bond length as well as the temperature dependence of H-bond properties have been described.[211,212] Very recently, the first direct experimental evidence of C^{α}−H·····O=C hydrogen bonds was found in the immunoglobulin binding domain of protein G.[213]

Solid-state NMR can provide accurate distance information with a precision of up to ±0.1 Å. It has been shown that the ^{15}N–^{1}H bond lengths in the side-chain of His · HCl · H_2O differ about 4 pm, which has been attributed to a strong H-bond between the $N^{\delta 1}$ proton and the carboxy-group of a molecule in the next neighbor unit cell.[214] Furthermore, anisotropic interactions that are absent in the liquid state, e.g. the chemical shift anisotropy (CSA), can be measured quantitatively, offering valuable structural information such as the orientation of the amide bond vector with respect to the carbonyl group. In addition, possible distortions of the CSA tensor due to hydrogen bonding can be detected.

7.2 Results and Discussion

7.2.1 Liquid-State NMR

Hydrogen bonds of the N−H·····O=C type in proteins are detected using a long-range (lr) HNCO experiment. The pulse scheme is a regular constant-time (CT) HNCO experiment (see figure 7.2),[215,216] optimized for the detection of small ^{15}N–$^{13}C'$ couplings.[206,207] The transfer amplitude A of the HNCO scheme depicted in figure 7.2 is given by

$$A = \exp\left(-4T\,R_2\right) \cdot \sin^2\left(2T\pi\,J\right) \cdot \prod \cos^2\left(2T\pi\,J'\right) \tag{7.1}$$

where $2T$ is the dephasing time during the INEPT transfer between in-phase N_y^i and anti-phase $2N_x^iC_z'^{(i-1)}$ magnetization, R_2 is the apparent transverse relaxation rate of the ^{15}N nucleus, and the product extends over all passive couplings. If the one-bond coupling between N^i and $C'^{(i-1)}$ is active during $2T$, J is given by $^1J_{NC'}$, while J' represents the much smaller intraresidual $^2J_{NC'}$ and $^{3h}J_{NC'}$ couplings. Both passive couplings are smaller than 1 Hz; hence, $\cos^2(2T\pi J') \approx 1$. Thus, the transfer amplitude for a direct one-bond coupling can be rewritten as

$$A_{\text{direct}} = \exp(-4T R_2) \cdot \sin^2\left(2T\pi\, {}^1J_{NC'}\right) \tag{7.2}$$

In proteins, values for $^1J_{NC'}$ vary between 13 and 17 Hz, with an apparent correlation to the strength of the hydrogen bond.[217] Typically, $2T$ is usually set to values slightly smaller than $1/(2 \cdot {}^1J_{NC'})$, and the direct correlations are obtained. On the other hand, if a very small coupling is active during $2T$, e.g. $^{3h}J_{NC'}$, A is proportional to $\sin^2(2T\pi\, {}^{3h}J_{NC'}) \cdot \cos^2(2T\pi\, {}^1J_{NC'})$. If $2T$ is set to $n/{}^1J_{NC'}$, where n is any integer number, then $\cos^2(2T\pi\, {}^1J_{NC'}) \approx \cos^2(n\pi) = 1$, and the transfer amplitude is described by

$$A_{\text{H-bond}} = \exp(-4T R_2) \cdot \sin^2\left(2T\pi\, {}^{3h}J_{NC'}\right) \tag{7.3}$$

Thus, it is possible to detect a small $^{3h}J_{NC'}$ coupling in the presence of large $^1J_{NC'}$ couplings. In this work, values for $2T$ of 136 ms ($n = 2$) and 272 ms ($n = 4$) were used.

2D HNCO spectra of pA_4 with uniformly ^{15}N,^{13}C-labeled Ala residues dissolved in d_6-DMSO are depicted in figure 7.3. In the left panel, a spectrum recorded with $2T$ set to $7/(2 \cdot {}^1J_{NC'}) = 238$ ms is shown. All three expected direct correlations are observed; in addition, a weak peak corresponding to an intraresidue correlation A2H^N–A2C′ *via* the $^2J_{NC'}$ coupling is found. The chemical shifts of A2C′ and A5C′, the hydrogen acceptor in the βII′-turn, differ only by about 0.02 ppm. Thus, in order to differenciate between the $^1J_{NC'}$ and

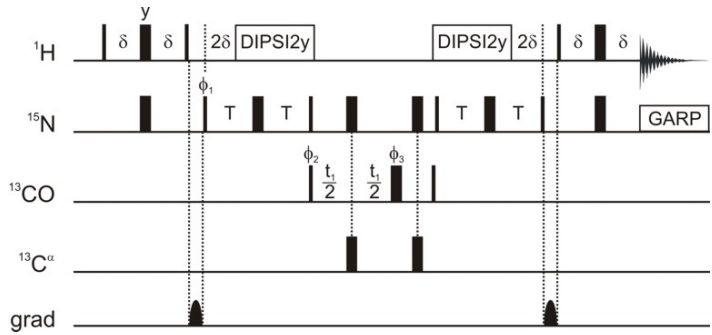


FIGURE 7.2: Pulse scheme of the 2D lr-HNCO experiment used in this work. Narrow and wide bars indicate pulses with a flip angle of 90° and 180°, respectively. All pulses are applied with phase x unless indicated otherwise. Pulses applied on the C′ channel had a square shape and a duration adjusted to achieve a null excitation at the C^{α} frequency and *vice versa*. Decoupling during acquisition is achieved using a GARP sequence;[80] evolution of ^{1}H–^{15}N anti-phase magnetization is suppressed using a DIPSI2 sequence.[218,219] Note that the peptide was dissolved in d_6-DMSO; hence, no solvent suppression was used. The delay δ was set to $1/(4 \cdot {}^1J_{\mathrm{NH}}) \approx 2.7$ ms. For values of T, see text. The following minimal phase cycle was employed: $\phi_1 = 4(x), 4(-x)$; $\phi_2 = x, -x$ $\phi_3 = x, x, y, y$; $\phi_{\mathrm{rec}} = x, 2(-x), x, -x, 2(x), -x$. Quadrature detection in the indirect dimension was achieved using the States method.[82]

the ${}^{3h}J_{\mathrm{NC}'}$ correlation, it is important to completely suppress the direct correlation A3H$^{\mathrm{N}}$–A2C′ in the lr-HNCO experiments. Using a value of 15 Hz as a starting point, $2T$ was set to 133 ms and varied until complete suppression of this direct correlation was achieved. The final value of $2T$ was 136 ms, corresponding to ${}^1J_{\mathrm{NC}'} = 14.7$ Hz. The J values for the other direct correlations are ${}^1J_{\mathrm{NC}'}(\mathrm{A4H^N} - \mathrm{A3C}') = 13.47$ Hz and ${}^1J_{\mathrm{NC}'}(\mathrm{A5H^N} - \mathrm{A4C}') = 13.89$ Hz. These values are approximately 2 Hz smaller than those observed for turns in proteins,[217] a fact that might be attributed to the strained geometry of CPPs. Juranić *et al*. have shown that ${}^1J_{\mathrm{NC}'}$ has the lowest value for a residue in the $i+2$ position of a reverse turn of the β type.[217] This correlation is, however, not observable in pA_4, since pro1 is not isotopically labeled. In contrast, the largest

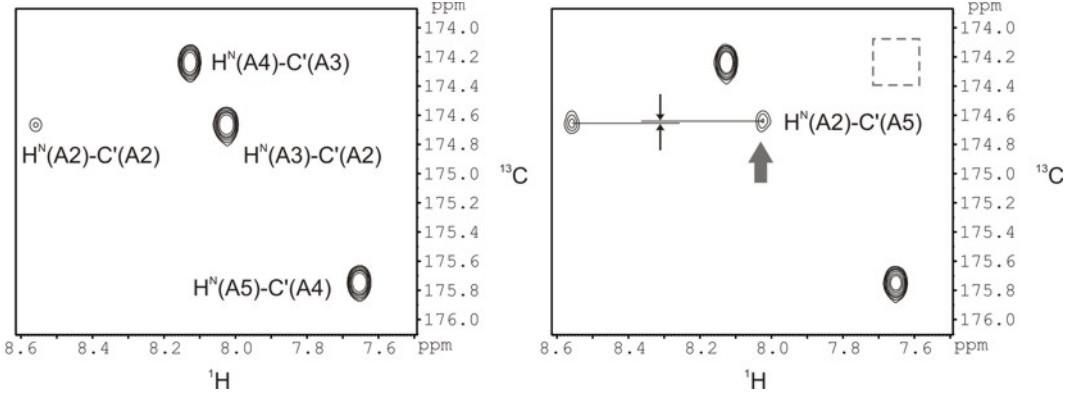


Figure 7.3: Left: HNCO spectrum of pA_4 recorded with $2T = 7/(2 \cdot {}^1J_{NC'}) = 238\,\text{ms}$. All three expected one-bond correlations are observed and labeled with the assignment. Right: Long-range HNCO spectrum with $2T$ set to $4/{}^1J_{NC'} = 272\,\text{ms}$. The peak corresponding to the H-bond A3H^N→A5C′ is indicated with a grey arrow. Note that the difference in chemical shift between the one-bond correlation A3H^N–A2C′ and the peak corresponding to the H-bond A3H^N→A5C′ is only about 0.02 ppm. The dashed box indicates where a peak corresponding to the γ-turn is expected.

$^1J_{NC'}$ is reported for the residue in the $i+3$ position. In pA_4, the largest $^1J_{NC'}$ is indeed observed for A3H^N–A2C′, which is located at the $i+3$ position.

The right panel of figure 7.3 shows a long-range HNCO recorded with $2T = 4/^1J_{NC'} = 272$ ms. Despite the small difference in chemical shift, the cross peak resulting from the H-bond A3H^N→A5C′ is readily identified. The two intense peaks in the right panel of figure 7.3 represent the one-bond correlations between A4H^N–A3C′ and A5H^N–A4C′; these signals are not completely suppressed due to small variations in their individual $^1J_{NC'}$. A cross peak corresponding to a γ-turn at Ala4 was not observed.

Once correlations *via* H-bonds have been identified, the magnitude of $^{3h}J_{NC'}$ can be estimated. In principle, various approaches for obtaining estimates of $^{3h}J_{NC'}$ exist. For example, the size of the coupling can be determined by a quantitative J-correlation technique.[203,206] This approach requires the acquisition of a reference spectrum with a slightly modified pulse scheme; this experiment was not implemented during this work. Alternatively, the magnitude of $^{3h}J_{NC'}$ can be obtained using a quantitative analysis of cross-peak intensities,[220,221] or by fitting peak intensities to transfer amplitudes. The latter method is generally not applicable to proteins due to fast transverse relaxation, since values larger than 200 ms have to be used for $2T$ in order to collect enough data points; for pA_4, two points were recorded.

As a first step, the transverse ^{15}N relaxation rate has to be estimated. For this purpose, three HNCO experiments were acquired with $2T$ set to 102, 170, and 238 ms, respectively, corresponding to $n/(2 \cdot {}^1J_{NC'})$ where $n = 3$, 5, and 7. The intensities of the peak A3H^N–A2C′ were fitted against equation 7.2 to yield $R_2 = 4.43 \pm 0.06$ Hz (see left panel of figure 7.4). The transverse ^{15}N relaxation rate of A3 was also estimated from the ratios of peak intensities for the three possible combinations; the average value of 4.35 ± 0.57 Hz is in good agreement with the result obtained from the fit.

Subsequently, the intensities of the peak A3H^N–A5C′ were fitted to equation 7.3 using $R_2 = 4.43$ Hz. The peak intensities were taken from spectra with $2T$ set to 136 and 272 ms, corresponding to $n/^1J_{NC'}$, where $n = 2$ and 4. The

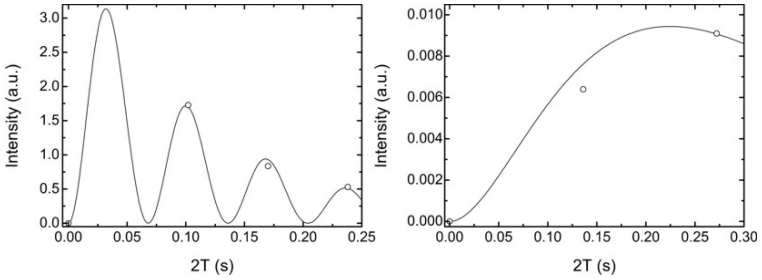


Figure 7.4: Fits of the lr-HNCO data. Open circles correspond to experimental values, while the solid line represents the best fit. Left: Fit of the transverse relaxation rate of A3N against equation 7.2 using peak intensities of the one-bond correlation A3N–A2C′. Right: The value of ${}^{3h}J_{NC'}$ was fitted using intensities of the peak A3N–A5C′.

coupling constant ${}^{3h}J_{NC'}$ obtained from this fit is 0.18 ± 0.08 Hz (see figure 7.4). The large uncertainty of ${}^{3h}J_{NC'}$ (44 %) can be most likely attributed to the fact that the minimum of the least-squares fit function is rather broad and small variations of ${}^{3h}J_{NC'}$ have only little effect on the fit.

In order to verify the results of the fit, ${}^{3h}J_{NC'}$ was estimated by quantitative analysis of cross peak intensities as described earlier.[207] Briefly, two experiments have to be recorded: one with $2T$ set to $n/(2 \cdot {}^{1}J_{NC'})$ for the direct correlation (A), and the other with $2T$ set to $(n+1)/(2 \cdot {}^{1}J_{NC'})$ in order to observe the cross-peak *via* the H-bond (B). The magnitude of ${}^{3h}J_{NC'}$ is derived from the following equation:

$$ {}^{3h}J_{NC'} = \frac{1}{\pi \cdot 2T} \cdot \arcsin\left[\sqrt{\frac{I^{B}}{I^{A}} \exp\left(\frac{R_2}{{}^{1}J_{NC'}}\right)}\right] \tag{7.4} $$

where I^{A} is the intensity of the one-bond N^{i}–$C'^{(i-1)}$ peak and I^{B} is the intensity of the peak *via* the H-bond; $2T$ is the dephasing time from experiment (B). Estimations were performed for spectra with $n = 3$ and 7, yielding a mean value of ${}^{3h}J_{NC'} = 0.17 \pm 0.01$ Hz, which agrees well with the result of the fit.

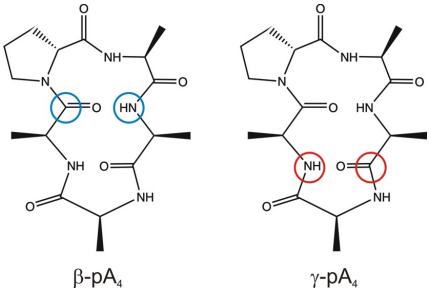


FIGURE 7.5: Sketches of β- (LEFT) and γ-pA$_4$ (RIGHT) used for solid-state NMR experiments. The molecules are only labeled at the H-bond donor and acceptor positions of the βII′-turn (A3N and A5C′, blue circles) as well as A3C′ and A5N (red circles).

7.2.2 Solid-State NMR

In order to derive information from ssNMR data, two selectively labeled peptides have been synthesized (see figure 7.5). In the first molecule, the H-bond donor- and acceptor positions of the βII′-turn, A3N and A5C′, are ^{15}N and ^{13}C labeled; this compound will be referred to as β-pA$_4$. Its counterpart γ-pA$_4$ is labeled at A3C′as well as A5N, the hydrogen donor and acceptor positions of the γ-turn. A detailed theoretical description of ssNMR is far beyond the scope of this work; only a brief introduction will be given here.

In the solid state, orientation-dependent interactions like dipole–dipole (DD) couplings and chemical shift anisotropy (CSA) interactions are not averaged out due to stochastic Brownian motion of the molecules as in liquid-state NMR. These strong interactions lead to a large linebroadening with linewidths on the order of several kHz. The Hamiltonian containing both DD and CSA contributions is dependent on the orientation of the dipole axis as well as the major components of the chemical shielding tensor according to:

$$\mathscr{H} \sim 0.5 \left[3\cos^2(\theta) - 1\right] = \mathrm{P}_2(\cos\theta) \tag{7.5}$$

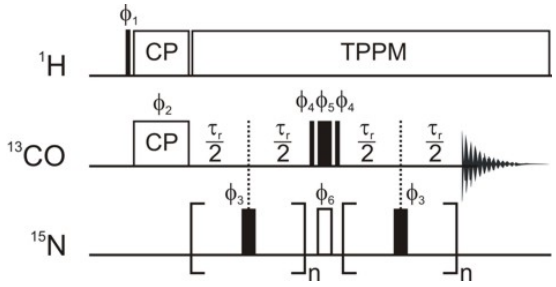


Figure 7.6: Pulse scheme of the REDOR experiment. Narrow and wide bars indicate pulses with a flip angle of 90° and 180°, respectively. All pulses are applied with phase x unless indicated otherwise. Proton decoupling is achieved using the TPPM sequence.[223] The outlined 180° pulse on ^{15}N is only applied in the reference experiment. The delay τ_r was set to the length of one rotor period (100 µs), the maximum dephasing time was 15.1 ms. The following phase cycle was employed: $\phi_1 = y, -y$; $\phi_2 = 2(x), 2(y), 2(-x), 2(-y)$ $\phi_3 = 2(x, y), 2(y, x)$; $\phi_4 = x, x, y, y$; $\phi_5 = y, y, -x, -x$; $\phi_6 = x, x, -x, -x$; $\phi_{\text{rec}} = x, -x, y, -y, -x, x, -y, y$.

The term $P_2(\cos\theta)$ is referred to as *second Legendre Polynomial*. $P_2(\cos\theta)$ vanishes for $\theta = \arccos(1/\sqrt{3}) \approx 54.7°$, the "magic angle". Thus, if the sample is rotated around the magic angle with respect to the external magnetic field, and if the rotation is fast compared to the size of the dipolar or CSA interaction, the linebroadening arising from anisotropic interactions vanishes.

One of the best known and probably most intuitive ssNMR experiments is REDOR (rotational echo double resonance).[222] REDOR allows to measure distances between two heteronuclei, e.g. ^{15}N and ^{13}C. The pulse scheme is depicted in figure 7.6 and can be summarized as follows. After magnetization transfer from ^{1}H to ^{13}C *via* cross-polarization (CP), the magnetization of interest is given by C_x. Subsequently, the dipolar interaction between ^{13}C and ^{15}N is recoupled by applying 180° pulses in the middle of each rotor period. Without the pulses, the Hamilton operator of the dipole–dipole interaction, $\mathscr{H}^{\text{DD}}$, would be averaged out after each rotor period (see top of figure 7.7). If 180° pulses are applied in the middle of each rotor period on ^{13}C *or* ^{15}N, the sign

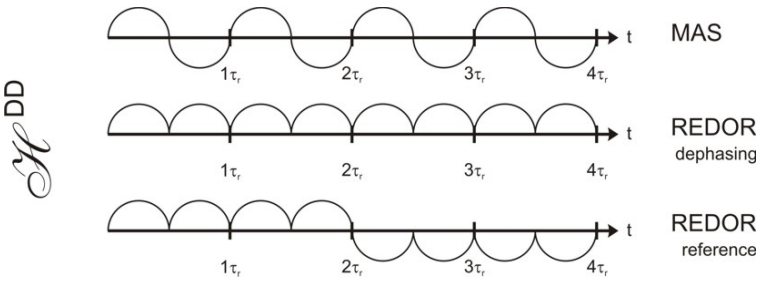


FIGURE 7.7: Evolution of the dipolar Hamiltonian $\mathscr{H}^{DD}$ under MAS (TOP), during the REDOR dephasing experiment (MIDDLE), and the REDOR reference experiment (BOTTOM). Under MAS conditions, the dipole–dipole interaction is averaged out after each rotor period, while it is retained in the dephasing experiment. In the reference experiment, the sign of $\mathscr{H}^{DD}$ is inverted in the middle of the dephasing period; as a result, the dipolar coupling is refocussed at the end of the dephasing time.

of $\mathscr{H}^{DD}$ is inverted and the dipolar coupling is reintroduced (middle of figure 7.7), and a decay of C_x as function of the dephasing time is observed. If relaxation is neglected, the dephasing of the signal can be described as

$$S(t) = \int_0^{2\pi} \mathrm{d}\gamma(t) \int_0^{\pi} \mathrm{d}\beta \sin(\beta) \cos[\Phi(t)] \tag{7.6}$$

where β and $\gamma(t)$ are Euler angles, Φ represents the effective dipolar coupling given by $\Phi = d\sqrt{2}\sin(2\beta)\sin[\gamma(t)]$, and d is defined as in equation 2.5 on page 17.

The observed decay is in fact a result of two mechanisms: (i) transverse relaxation, i.e. C_x decays with R_2, and (ii) dipole–dipole coupling, i.e. C_x is converted into $2C_yN_z$ anti-phase magnetization. This anti-phase magnetization cannot refocus during acquisition and is thus not detectable. In order to seperate contributions arising from relaxation and decay due to diploar coupling, two experiments are recorded. In the dephasing experiment, $\mathscr{H}^{DD}$ is reintroduced and decay due to both mechanisms is measured. In the reference exper-

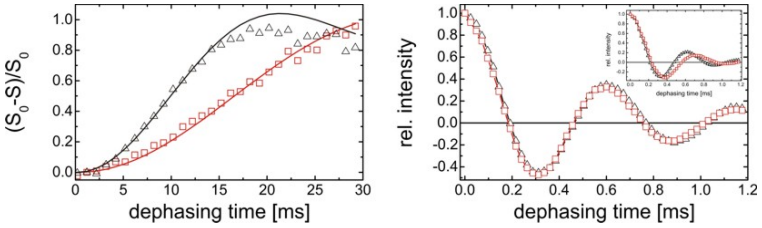


Figure 7.8: Left: Plots of the REDOR dephasing factor *vs*. the dephasing time for β-pA$_4$ (red squares) and γ-pA$_4$ (black triangles). The open symbols represent experimental data points, while the solid lines correspond to the best fit. Right: Plot of the N–H dephasing due to recoupling of the ^{1}H–^{15}N dipolar coupling using the T-MREV sequence for β-pA$_4$ (red squares) and γ-pA$_4$ (black triangles).[224] The inset shows the dephasing curves for the two sidechain NH-groups in His $\cdot$ HCl $\cdot$ H$_2$O. A different dephasing behaviour resulting from an elongation of the N$^{\delta 1}$–H distance of 4 pm is clearly visible.

iment, the sign of $\mathscr{H}^{\mathrm{DD}}$ is inverted in the middle of the dephasing time (see lower panel of figure 7.7). Therefore, the dipolar interaction is averaged out at the end of the dephasing period, and only the decay due to relaxation is obtained. Finally, the distance is extracted from fitting the REDOR dephasing curve against equation 7.6. The dephasing curve is defined as

$$\frac{\Delta S}{S_0} = \frac{S_0 - S}{S_0} \tag{7.7}$$

where S and S_0 are the signal intensities of the dephasing and reference experiment, respectively.

REDOR experiments have been applied to β-pA$_4$ and γ-pA$_4$ at a MAS frequency of 10 kHz; plots of the REDOR dephasing factor *vs*. the dephasing time are shown in figure 7.8. The distances were fitted as 4.03 ± 0.06 Å for A3N–A5C′ and 3.39 ± 0.03 Å for A3C′–A5N. These values are in good agreement with the distances of 4.17 and 3.28 Å found in the structure published by the Kessler group (see table 7.1),[199] but not with the distances found in the structures published by Nikiforovitch *et al*.[201]

TABLE 7.1: Comparison of N–C distances for different pA_4 structures. N1 through N4 are structures from Nikiforovitch *et al.*;[201] NMR corresponds to the structure by Kessler *et al.*[199]

	dA3N–A5C′[a]	dA3C′–A5N[a]
N1	4.75	3.61
N2	5.03	3.79
N3	4.98	3.70
N4	4.74	3.28
NMR	4.17	3.28
REDOR	4.03	3.39

[a] Distances in Å.

In addition to the N–C′ distances, the ^{15}N–H^N distances have been measured for Ala3 and Ala5. It seems intuitive that these distances can be significantly eloganted if the atoms are involved in a strong H-bond. This phenomenon has indeed been reported for ^{15}N–^1H distances in the sidechain of $His \cdot HCl \cdot H_2O$,[214] where an elongation of 4 pm has been attributed to a strong H-bond between the $N^{\delta 1}$ proton and the carboxy-group of a molecule in a neighboring unit cell. In pA_4, the ^{15}N–^1H distances in Ala3 and Ala5 are identical, as is evident from the right panel of figure 7.8.

As an additional source of information, the C′ CSA tensor should be indicative for hydrogen bonding. Figure 7.9 shows experimental and simulated 1D spectra of β-pA_4 and γ-pA_4 at a MAS frequency of 2 kHz. It is obvious that the spectra of both CPPs are different, which is expected if one carbonyl group is involved in a H-bond while the other is not. The carbonyl CSA tensor is given by

$$\mathrm{CSA} = \Delta\sigma \sqrt{1 + \frac{\eta^2}{3}} \tag{7.8}$$

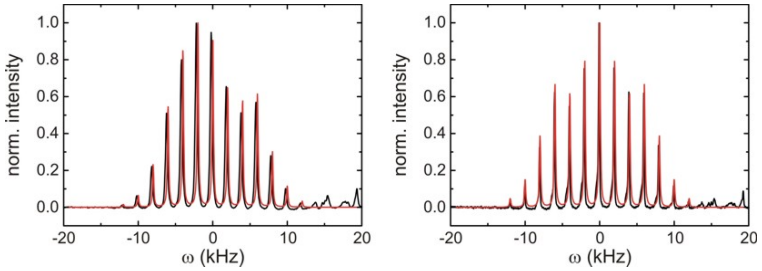


Figure 7.9: Experimental (black) and simulated (red) 1D spectra of β-pA$_4$ (left) and γ-pA$_4$ (right) at a MAS frequency of 2 kHz.

where $\Delta\sigma = \sigma_{zz} - [(\sigma_{yy} + \sigma_{xx})/2]$ and $\eta = (\sigma_{yy} - \sigma_{xx})/(\sigma_{zz} - \delta_{\text{iso}})$; $\Delta\sigma$ is the CSA for a symmetric shift tensor, and η is the asymmetry parameter; δ_{iso} is the isotropic chemical shift. The CSA tensor is illustrated in figure 7.10. The σ_{zz} component encloses an angle α with the N–C′ bond, while σ_{yy} is oriented approximately parallel to the C′–O bond. The plane defined by the atoms forming the peptide bond, N, H$^{\text{N}}$, C$^{\alpha}$, and C′, is in the paper plane, and σ_{xx}, which is oriented perpendicular to the paper plane, points towards the reader.

Different CSA tensors indicate a different electron density at both carbonyl sites. If a carbonyl group is involded in a hydrogen bond, the electron density is

Table 7.2: Components of the CSA tensor for β-pA$_4$ and γ-pA$_4$. As a comparison, values reported by Tycko *et al.* are shown.[225] All values are given in ppm; η has no unit.

	δ_{iso}	$\Delta\sigma$	η	σ_{zz}	σ_{yy}	σ_{xx}
β-pA$_4$	179.7	77.1	0.76	230	172	133
γ-pA$_4$	178.3	80.3	1	232	178	125
Tycko *et al.*[225]	171.0	106.5	0.76	242	182	89

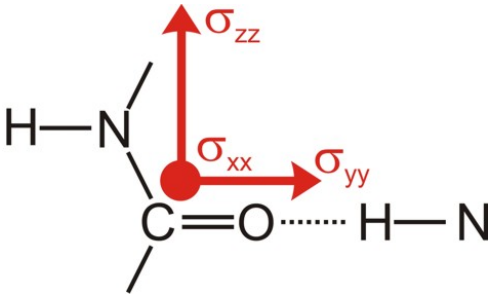


FIGURE 7.10: Components of the carbonyl CSA tensor. The plane defined by the peptide bond lies in the paper plane. σ_{zz} points into the direction of the amide nitrogen and encloses an angle α with the N–C′ bond. σ_{yy} is oriented parallel to the C′–O bond, while σ_{xx} is perpendicular to the paper plane and point towards the reader.

redistributed towards the H-bond donor; therefore, a change in the σ_{yy} component is expected, while a change in σ_{zz} might reflect a different ${}^{1}J_{\mathrm{NC'}}$ coupling constant. For β-pA$_4$, values of 77.05 ppm and 0.76 were obtained for $\Delta\sigma$ and η, respectively. The results for γ-pA$_4$ differ significantly; here, $\Delta\sigma = 80.3$ ppm and a symmetric tensor with $\eta = 1$ were obtained. Literature values of $\Delta\sigma$ vary between 112 and 147 ppm.[226,227] The largest difference is observed for σ_{yy}, which is agreement with the existence of a H-bond. However, the observed value of $\eta = 1$ for γ-pA$_4$ does not correspond to the canonical value. Therefore, it must be assumed that the electron density at A4C′ is distorted as well.

7.3 Conclusion

The results of the present work clearly confirm the existence of a hydrogen bond A3H$^{\mathrm{N}}$→A5C′. In the long-range HNCO spectra, cross peaks at the expected chemical shifts are readily visible, and the magnitude of ${}^{3\mathrm{h}}J_{\mathrm{NC'}}$ could be estimated. However, no peak corresponding to a hydrogen-bond A5N→A3C′ was observed; this fact might be explained with an equilibrium between four

different conformers,[199] in which the population of the conformer featuring the γ-turn is too small to allow for a detection of ${}^{3h}J_{NC'}$.

Evaluation of a REDOR experiment allowed to measure the N–C′ distances for A3N–A5C′ and A3C′–A5N in the solid state. As shown in table 7.1, these values are in good agreement with the distances found in the structure published by Kessler *et al.*,[199] but do not compare to those found in conformers by Nikiforovitch *et al.*[201] As expected, different C′ CSA tensors were obtained from fitting 1D lineshapes of β-pA_4 and γ-pA_4 against equation 7.8, indicating that the electron densities at both carbonyl groups are different. Since the involvement of A5C′ in a H-bond is established from liquid-state NMR data, it is straightforward to conclude that the different CSA tensors are caused by the presence (A5C′) or absence (A3C′) of hydrogen bonding.

Chapter 8

NMR Experiments with Detection on Aliphatic Protons on Cryogenic Probes

8.1 Introduction

In biomolecular NMR spectroscopy, the majority of experiments—especially for backbone assignment—employ signal detection on amide protons (for a review of NMR experiments see[153]). Suppression of the intense water signal, which is approximately $10^5\times$ larger than the signals of the molecules of interest, is usually achieved by gradient selection, water selective pulses as implemented in the WATERGATE scheme,[77] water flip-back,[83] or a combination of these techniques. The application of water selective pulses is feasible due to the large frequency separation ($\sim 2\,\mathrm{kHz}$) between amide and water protons.

There are, however, a number of experiments in which the signal is detected on aliphatic protons, e.g. experiments for sidechain assignment. For these experiments, the use of water selective pulses is not possible, since this would

also affect all protons resonating at or near the water frequency, e.g. H^{α} protons. Hence, the options for water suppression in HC-detected experiments are limited to gradient selection, use of z-gradients to dephase the water magnetization, and spin-lock pulses.

In this chapter, combinations of all these techniques are investigated in terms of water suppression effectiveness on a cryogenic probe equipped with a z-gradient at a magnetic field strength of 14.1 T, corresponding to a ^{1}H resonance frequency of 600 MHz.

8.2 Water Suppression and Radiation Damping

The vast majority of biomolecular NMR studies is preformed in aqueous solutions. The concentration of protons in water is approximately 110 M, compared to 1–2 mM for the macromolecule. Thus, detection of the solute signal in the presence of such a large solvent signal represents a difficult problem, because the dynamic range of the electronic components of the spectrometer hardware is limited. Therefore, the magnitude of the solvent signal has to be reduced. In biomolecular NMR spectroscopy, this is commonly achieved using water suppression techniques such as dephasing of the water magnetization using R.F. (homo-spoil) or gradient pulses, selective excitation of the water resonance, and, to a lesser extend, presaturation of the water resonance with a weak R.F. field during the recycle delay between transient acquisitions (for reviews of solvent suppression techniques see[228,229]). The latter two cannot be used in HC-detected experiments due to the resonance overlap with the water signal.

NMR spectroscopy in aqueous solutions also suffers from *radiation damping* of the solvent signal.[230,231] A detailed theoretical description and diagnostics of radiation damping can be found in the journal *Concepts in Magnetic Resonance*;[232,233] here, only a short introduction will be given. Magnetization precessing in the transverse plane induces a time-dependent current in the coil. This current, in turn, induces an electromagnetic field of the same frequency, which acts like a selective R.F. pulse and rotates the magnetization back to the

$+z$ axis. Radiation damping is the result of the interaction between magnetization and a current and thus represents an *intrinsic* physical phenomenon that theoretically affects *every* NMR signal. In practice, however, only the currents induced by the large solvent magnetization lead to notable effects.

Radiation damping has two detrimental effects for solvent suppression in NMR. First, the water magnetization after application of a 90° pulse will return to the $+z$ axis within tens of milliseconds as a result of radiation damping on a modern NMR probe. This interferes with the expected evolution of the water magnetization and hence spoils solvent suppression schemes. Second, since the relaxation of the water signal is governed by radiation damping, its linewidth depends on the time constant of this process:

$$\Delta\nu_{\mathrm{FWHH}} = \frac{0.8384}{\pi \cdot T_{\mathrm{rd}}} \tag{8.1}$$

where $\Delta\nu_{\mathrm{FWHH}}$ is the full width at half height of the water signal and T_{rd} is the time constant of the radiation damping process.[234] On a well tuned and matched cryo probe at a B_0 of 14.1 T, values of 100 Hz and more for $\Delta\nu_{\mathrm{FWHH}}$ of the water signal are typical and resonances near the water line may be obscured by the broad tails of the residual solvent peak. T_{rd} in turn is given by

$$T_{\mathrm{rd}} = (2\pi\, M_0\, Q\, \gamma)^{-1} \tag{8.2}$$

where M_0 is the equilibrium water magnetization and Q represents the quality factor of the coil.[230] Equation 8.2 illustrates that a reduction of M_0 leads to slower radiation damping and hence to a narrower linewidth of the water signal.

8.3 ^{1}H,^{13}C-HSQC

The constant time (CT) HSQC represents one of the basic ^{1}H,^{13}C correlation experiments (see figure 8.1).[235,236] Usually, it is implemented with gradient se-

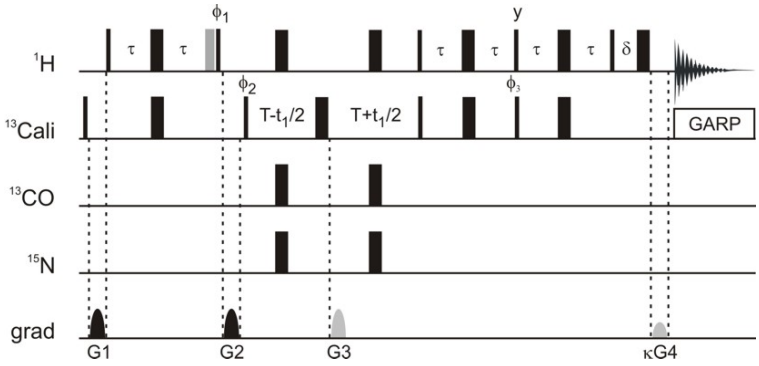


FIGURE 8.1: Pulse scheme of the ^{1}H-^{13}C-CT-HSQC with gradient selection and sensitivity enhancement. In all pulse schemes in this chapter, narrow and wide bars represent pulses with a flip angle of 90° and 180°, respectively, and are applied with phase x unless indicated otherwise. The grey pulse on the proton channel corresponds to a spin-lock pulse with a maximum duration of 1 ms applied with the same R.F. amplitude as the hard pulses. Delay durations are $\tau = 1.3$ ms, $T = 28$ ms, and $\delta = 1$ ms. Gradient durations (ms) and strengths (G/cm) are G1 0.8/32, G2 variable/32, G3 0.8/40, G4 0.8/10. Gradients shaded grey were used for coherence selection. Quadrature detection in F_1 was achieved by inverting κ in concert with ϕ_3.

lection (GS) and sensitivity enhancement (SE).[237,238,239,240,241] This CT-HSQC-GSSE offers reliable performance for samples with protonated buffer on spectrometers with conventional probes.

On a 600 MHz spectrometer equipped with a cryo probe, a proper adjustment of the receiver gain is virtually impossible due to the intense water signal. Incorporation of a spin-lock pulse in the first INEPT step and a z-spoil gradient prior to the carbon excitation pulse suppress the water resonance almost completely (see figures 8.2 and 8.3).

In order to optimize water suppression, the length of the spin-lock pulse and G2 were varied. Figure 8.2 shows HSQC spectra aquired with the pulse sequence depicted above and three different lengths of G2: 1 ms, 2 ms, and 4 ms, respectively. Although longer gradients work more efficiently, a duration

FIGURE 8.2: ^{1}H,^{13}C-HSQC spectra of uniformly ^{15}N,^{13}C-labeled ubiquitin dissolved in protonated buffer recorded without a spin-lock pulse. Gradient durations are 1 ms (LEFT), 2 ms (MIDDLE), and 4 ms (RIGHT).

of 1 ms already effectively suppresses the water signal. For gradients longer than 4 ms water suppression improves only slightly.

Water suppression using spin-lock pulses relies on the dephasing of the water coherence due to B_1 inhomogeneity. This *inhomogeneity* of the R.F. field describes the variation in the amplitude of the B_1 field as a function of position in the sample. As a results, the integral of magnetization *orthogonal* to the phase of the spin-lock pulses tends to zero if the B_1 field is applied sufficiently long. With modern NMR probes, a significant dephasing of the solvent signal is achieved within a few milliseconds.

Since spin-lock pulses dephase coherences othogonal to the R.F. field, they also eliminate artifacts in NMR spectra arising from non-carbon bound protons (e.g. N–H and O–H) or protons bound to ^{12}C. In order to use spin-lock pulses for solvent suppression, orthogonal components of the solvent and solute magnetization have to be created. For isotopically enriched samples, this is readily achieved by creating anti-phase operators, e.g. $2H_xC_z$, using the heteronuclear J-coupling.

In the present work, all proton spin-lock pulses were applied with the same R.F. amplitude as the hard pulses and a maximum duration of 1 ms to avoid damage of the probe. Figure 8.3 shows HSQC spectra acquired with different spin-lock pulse lengths. Surprisingly, a duration of only 50 µs gave the best water suppression. For 250 µs or even 1 µs, the water signal deteriorates the quality of the spectra.

An alternative is to substitute the first INEPT transfer with a saturation of the protons and hence using the heteronuclear NOE. In this case, the water signal would be saturated, and excellent water suppression could be achieved by additional use of z gradients. However, experiments using an INEPT step for the coherence transfer ^{1}H$\rightarrow^{13}$C yield a better sensitivity than experiments exploiting the heteronuclear NOE.

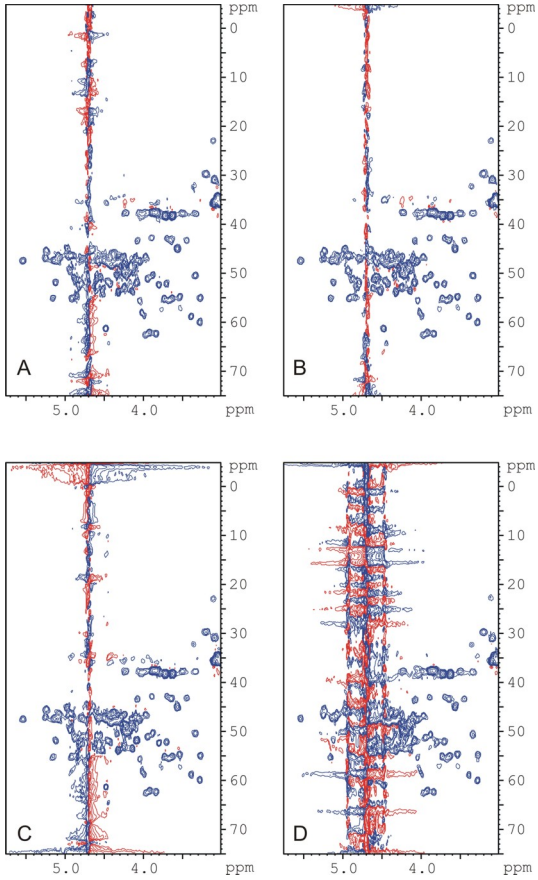


FIGURE 8.3: ^{1}H,^{13}C-HSQC spectra of ubiquitin recorded with the length of G2 set to 1 ms. Spin-lock pulse lengths were 0 µs (**A**), 50 µs (**B**), 250 µs (**C**), and 1 ms (**D**).

8.4 HCACO

The HCACO experiment is useful in cases where HN-detected experiments are hampered due to fast exchange of H^N and solvent protons.[215,242,243,75] Most conveniently, the HCACO experiment is perfomed with the sample dissolved in deuterated buffer in order to reduce the water signal and thus overcome overlap of H^α resonances with water protons.[64] This procedure is, however, not feasible in most cases due to a limited amount of protein. Changing the buffer to deuterated buffer often results in significant losses of sample. Hence, it is desirable to perform the HCACO in protonated buffer. A pulse scheme for this purpose has been published and was optimized here.[244]

Figure 8.4 shows two pulse sequences which differ only by the position of the two proton spin-lock pulses applied with x and y phase, respectively; the water suppression achieved by both sequences is almost identical. For the lower sequence, the water suppression is slightly improved when both spin-lock pulses have the same length (50 µs). A significantly inferior water suppression was observed when the duration of G3 was less than 6 ms. It should be noted that these sequences are similar to the pulse scheme published earlier;[244] in this work, however, gradient selection and significantly shorter spin-lock pulses are used.

The planes from HCACO spectra shown in figure 8.5 demonstrate that this experiment can indeed be performed in protonated buffer on a high field spectrometer equipped with a cryoprobe. Planes taken from the H^α–C^α as well as from the H^α–C' dimension display only a small residual water trace and allow for the detection of peaks that overlap with the water signal. As mentioned above, long durations of the z gradients (G3 and G4 in figure 8.4) have to be chosen in order to achieve a good water suppression. Longer gradient durations translate into sensitivity losses due to relaxation. However, the z gradients are applied when the magnetization of interest is present as polarization. Therefore, these relaxation losses should not be critical. Furthermore, the significantly improved qualitiy of the spectra overcompensates these small losses.

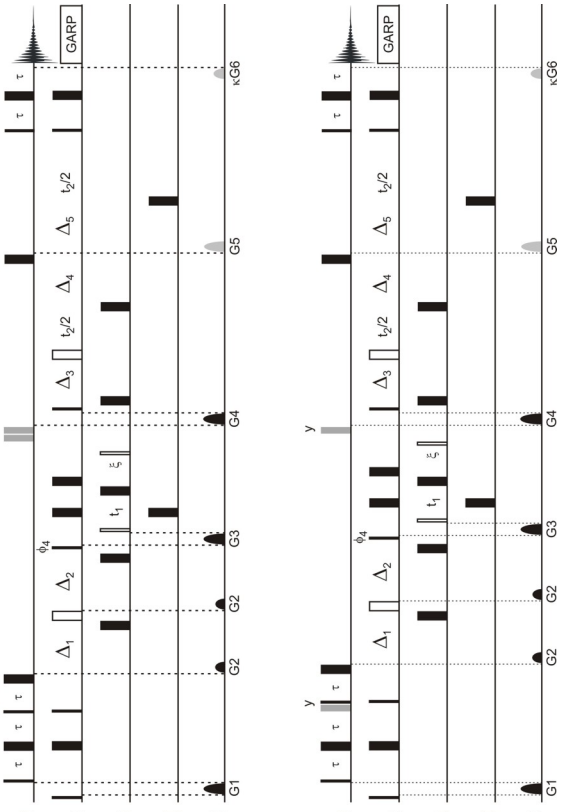


FIGURE 8.4: Pulse schemes of the constant-time HCACO experiment. Both sequences differ only in the position of the first proton spin-lock pulse and yield similar water suppression. Durations of the spin-lock pulses were 50 µs along x and 100 µs along y for the upper sequence. For the lower sequence, a duration of 50 µs for both spin-lock pulses resulted in a slightly better water suppression. For the z-spoil gradients, settings were G3 6/36 and G4 2/32.

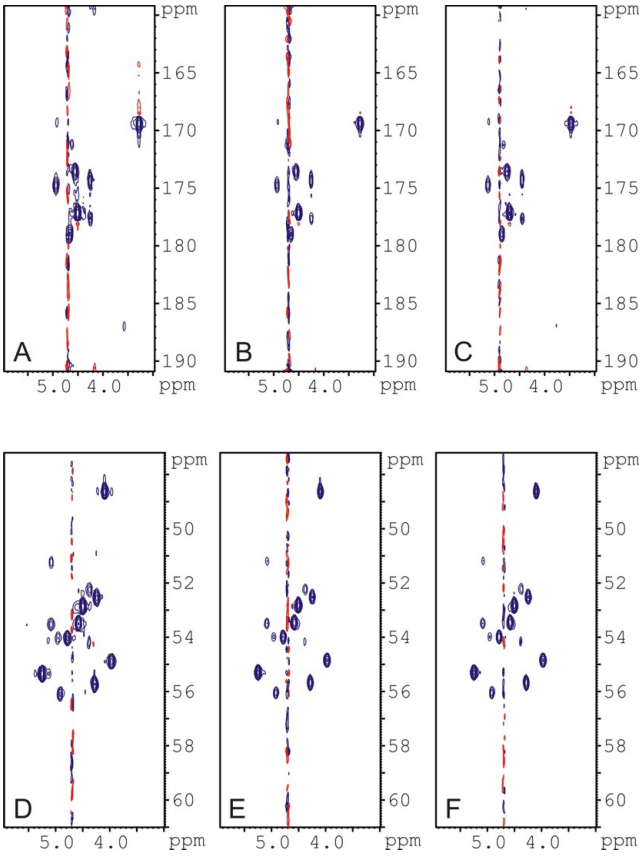


FIGURE 8.5: Planes from HCACO spectra of ubiquitin. Panels **A** and **D** were taken from a spectrum recorded with the upper sequence of figure 8.4. The lengths of the spin-lock pulses were 50 µs along x and 100 µs along y, respectively. The other panels were taken from spectra recorded with the lower pulse scheme of figure 8.4. Panels **B** and **E**: both spin-lock pulses had the same durations as before; panels **C** and **F**: a uniform duration of the spin-lock pulses of 50 µs was used. For all experiments, gradient settings were G3 6/36 and G4 2/32.

8.5 NOESY-HSQC and HSQC-NOESY-HSQC

In protein NMR spectroscopy, the ^{13}C edited NOESY-HSQC as well as the HSQC-NOESY-HSQC represent important experiments,[245,246,247,248,249] since they provide a means to overcome the signal overlap in ^{1}H–^{1}H 2D NOESYs and thus help to identify NOE contacts critical for structure determination of biological macromolecules. Figure 8.6 shows pulse schemes of CT versions of these two experiments.

Radiation damping is expected to be a severe issue for homonuclear proton correlations, especially for higher t_1 increments, since recovery of water z magnetization seriously affects the water suppression perfomance of virtually any homonuclear ^{1}H NMR experiment. Therefore, effective solvent suppression in experiments with an indirect proton evolution time represents a challenge.

If no attempts to suppress radiation damping during the ^{1}H evolution period of a NOESY-HSQC are made, the huge water trace obscures the signals of all other protons resonating near the water resonance. A slighty better quality of the spectra is achieved by insertion of a proton spin-lock pulse into the first INEPT step (panel **D** of figure 8.7). Incorporation of weak gradients with opposite sign during the ^{1}H evolution period improves the water suppression (see upper panel of figure 8.6), although the result still does not represent the optimum.

In the HSQC pulse schemes shown above, the water coherence is efficiently dephased; furthermore, an appropriate phase cycle shifts any residual water signal to the edges of the ^{13}C spectral window. Therefore, the HSQC-projections of both heteronuclear edited NOESYs show only a small water trace (see figure 8.7). In the HSQC-NOESY-HSQC, a better solvent suppression was achieved by omitting all spin-lock pulses. It should be noted that the pulse scheme of the HSQC-NOESY-HSQC shown in figure 8.6 does not represent the best implementation of this experiment. Here, two CT-HSQCs were simply concatenated using a NOE mixing time on protons.

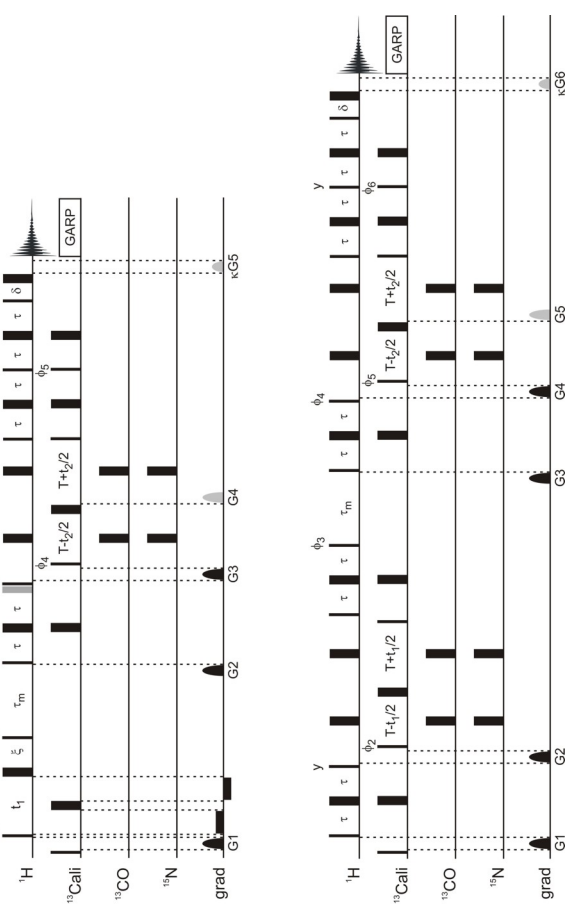


FIGURE 8.6: Pulse schemes of the NOESY-HSQC (TOP) and HSQC-NOESY-HSQC (BOTTOM). In the NOESY-HSQC, acceptable water suppression was achieved using weak gradients of opposite sign during t_1 and a 50 µs spin-lock pulse during the first INEPT step. For the HSQC-NOESY-HSQC, omitting the spin-lock pulse yielded a better water suppression.

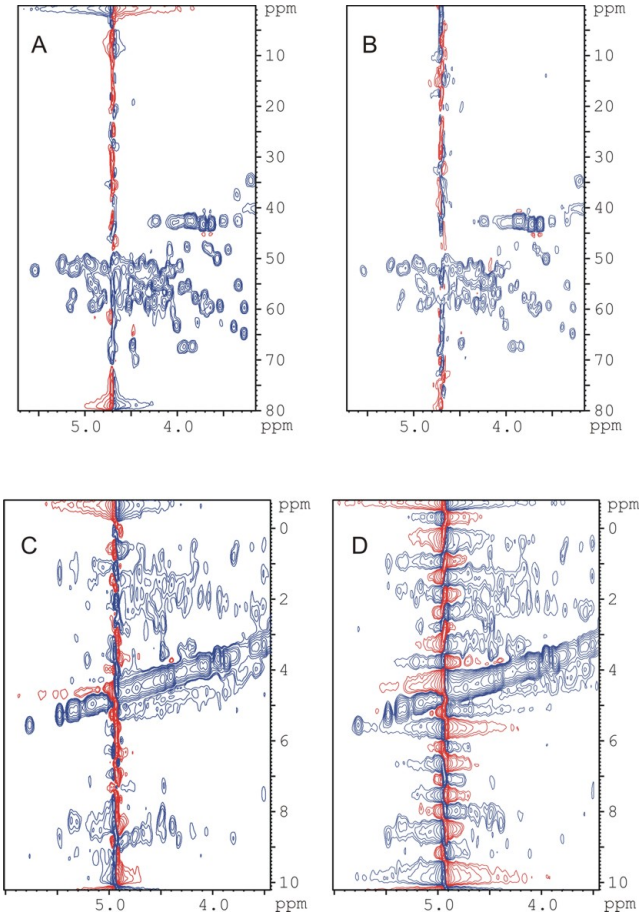


FIGURE 8.7: HSQC-projections of the NOESY-HSQC (**A**) and HSQC-NOESY-HSQC (**B**) recorded with the sequences given in figure 8.6. Both projections display a good water suppression. Panel **C** and **D** show 2D NOESY-HSQCs recorded with weak gradients (**C**, 4 G/cm) and without gradients (**D**) during t_1.

8.6 Conclusion

In this chapter, three standard 2D and 3D HC-detected NMR experiments were optimized for use in combination with a cryogenic probe at high B_0 fields. Water suppression was achieved by application of z gradients as well as spin-lock pulses to efficiently dephase the water coherence. Surprisingly, short spin-lock pulses (50–100 µs) gave a better water suppression than longer durations up to 1 ms. The length of the z gradients necessary for optimal water suppression is on the order of 2 to 6 ms. These rather long durations may lead to intensity losses due to relaxation. However, these losses should not be severe since the magnetization is present as polarization when the gradients are applied. Furthermore, minor losses are overcompensated by the improved spectral qualitiy due to better water suppression.

Chapter 9

Summary

In the present thesis, the main subject of interest is the characterization of protein dynamics by NMR techniques. In the first part, an introduction to theoretical aspects of spin relaxation is given. NMR experiments for measuring spin relaxation rates are discussed, and the model-free analysis of relaxation data is described, which yields information about amplitudes of bond vector motions on a pico- to nanosecond timescale (chapter 2). Motions on a slow, millisecond time scale can be characterized using NMR relaxation dispersion data; this technique is introduced in chapter 3.

Using the previous chapters as a framework, the analysis of relaxation data according to the model-free approach as well as the analysis of relaxation dispersion data are illustrated in detail in chapter 4. During the course of this work, several scripts have been implemented in order to facilitate the analysis of NMR relaxation data. Usually, this analysis is very time-consuming and prone to errors, since a large number of different files have to be created and analyzed.

In the X-ray structures of the isolated α- and β-apical domains of the thermosome of *Thermoplasma acidophilum*, different conformations were observed for a stretch of residues in the proximal part of the helical protrusion. Since this region represents a putative substrate binding region, α- and β-ADT were investigated using heteronuclear NMR experiments. As expected from a 63 % sequence homology, an identical topology was deduced for both domains. This topology is in good agreement with the crystal structures; however, the proximal part of the protrusion was found to be intrinsically disordered. Model-free analysis of ^{15}N backbone relaxation data revealed motions for these residues sampling a wide conformational space that is not expected for motions within an ordered secondary structure element.

In chapter 6, the influence of AMP-PNP binding to the nucleotide binding domain of the P-type ATPase Kdp was investigated. Ligand binding occurs with a dissociation constant of approximately 1 mM, indicating weak binding and a fast exchange between free and bound forms. In the *apo* form, motions of the secondary structure elements were found to be restricted on a fast, pico- to nanosecond time scale, as indicated by high S^2 values, whereas 42 residues were identified to be subject to a chemical exchange process on a millisecond time scale. All exchanging residues are clustered around the nucleotide binding site. In the presence of AMP-PNP, a significant decrease in S^2 was observed for three residues in proximity to the binding site, whereas a loop on the opposite side of KdpBN was found to be more rigid. S^2 is related to the conformational entropy ΔS_p. For 32 residues, significant changes in ΔS_p were obtained, summing up to a slight loss of entropy. This loss is related to fast motions and might thus be compensated for by increased dynamics on slower time scales. Indeed, 63 residues were found to be subject to a millisecond time scale exchange process in *holo*-KdpBN.

Earlier studies revealed that the upper half of the cyclic pentapeptide pA_4 represents a well-defined βII$'$-turn with pro1 in the $i+1$ position, while four different conformations including a γ- and a γ'-turn were found for Ala4. In order to confirm these structural motifs, hydrogen bonds in the cyclic pentapep-

tide pA_4 were probed using a combination of solution- and solid-state NMR techniques. Long-range HNCO experiments showed a signal at the chemical shifts expected for a hydrogen bond A3H^N→A5C′, thereby confirming the βII′-turn; however, no evidence of an additional hydrogen bond A5H^N→A3C′ corresponding to a γ-turn centered at Ala4 was found. The magnitude of the scalar coupling constant across the H-bond, $^{3h}J_{NC'}$, was estimated to be less than 0.2 Hz. In the solid state, N–C′ distances for selectively labeled peptides were measured using the REDOR experiment; the results are in good agreement with the structure published earlier. Furthermore, different CSA tensors were found for Ala3C′ and Ala5C′, indicating that the electron densities at both carbonyl groups are different. These differences were attributed to the participation of Ala5C′ in hydrogen bonding.

Chapter 8 illustrates how efficient water suppression can be achieved for a number of 2D and 3D NMR experiments with detection on carbon-bound protons on a 600 MHz spectrometer equipped with a cryo probe. Appropriate combinations of spin-lock pulses and pulsed field gradients ensured good water suppression in ^{1}H,^{13}C-HSQC, HCACO, and ^{1}H,^{13}C-HSQC-NOESY-HSQC experiments. The solvent suppression in the ^{1}H,^{13}C-NOESY-HSQC could be improved significantly; however, homonuclear proton–proton correlations remain a challenge.

Appendix A

Material and Methods

A.1 Protein Sample Preparation

Protocols for the preparation of NMR samples of α- and β-ADT as well as for KdpBN have been published earlier and are thus not described here.[250, 174]

A.2 NMR Spectroscopy: ^{15}N Backbone Relaxation Data

All NMR experiments for the ADTs were performed at 42 °C; at this temperature, a good spectral quality was obtained for both domains. Sequential assignments were achieved using a collection of standard 3D NMR experiments; for details see.[250, 154] The chemical shifts of amide ^{1}H, ^{15}N, C^{α}, and C′ resonances have been deposited in the BioMag-ResBank (`http://www.bmrb.wisc.edu`) under accession numbers 5930 and 5936 for α-ADT and β-ADT, respectively.[154] Relaxation data for α-ADT was collected in a similar manner as described below; the results of the MF analysis for α-ADT are given in table B.1.

All experiments for KdpBN were collected at 27 °C; the *holo*-KdpBN sample contained a five-fold excess of AMP-PNP. For R_1 and R_2 experiments (see figure 2.8), typically 10 relaxation delays were recorded; in addition, 3 duplicate experiments were acquired in order to estimate the experimental uncertainties of the peak heights. In the indirect ^{15}N dimension, 80 complex points were sampled with 48 transients and a relaxation delay of 2 s between successive scans, resulting in an average aquisition time of approximately 4 h per spectrum. For the heteronuclear NOE, two sets of experiments consisting of spectra with and without saturation of the amide protons were recorded. A relaxation delay between successive scans of 5 s was followed by saturation of the amide protons for 3 s. Saturation of H^N was achieved by a train of 120° square pulses with the carrier centered on the middle of the amide region and the R.F. amplitude set such as to achieve a null excitation at the water resonance. In the reference experiment, the pulse train was substituted by a delay of equal length. 80 complex points with 48 transients resulted in a net aquisition time of roughly 17 h per spectrum.

Relaxation dispersion data were recorded at B_0 fields of 14.1 T and 17.6 T using the pulse scheme depicted in figure 3.4. For each profile, 14 experiments with different values of τ_{cp} ranging from $\sim$ 600 µs to 20 ms and a total relaxation period $T_{CP} = 80$ ms were collected. In addition, a reference experiment lacking the relaxation period was acquired. 48 transients, 80 complex points in the ^{15}N dimension, and a relaxation delay of 2 s correspond to a measurement time of approximately 4 h per spectrum.

Peak intensities were extracted using SPARKY. Relaxation rates were obtained from fitting the intensities to an exponential decay using CURVEFIT and the PERL-script SPARKY2RATE. A minimum uncertainty of 2 % for all rates was assumed. Initial guesses of the rotational diffusion tensor were obtained using either QUADRIC_DIFFUSION or R2R1_DIFFUSION and the PDB coordinate file 1SVJ for *apo*-KdpBN. Highly mobile residues or residues subject to chemical exchange were excluded from the estimation. Model-free analysis of the ^{15}N relaxation data was done using MODELFREE interfaced with FASTMODEL-

free.[33,32,30] The results for the *apo*- and *holo*-form of KdpBN are shown in table B.2 and B.3, respectively.

Analysis of relaxation dispersion data was achieved using self-written scripts for Scilab following a procedure described earlier.[54,56,57] Briefly, for each residue, both profiles were fitted globally to a flat line and the fast-exchange approximation (equation 3.11), and *F*-tests were used to assess whether statistical improvements were achieved by using equation 3.11. Residues passing the *F*-test were subsequently fitted to the general expression for a two-site chemical exchange process (equation 3.3), and *F*-statistics were used to decide which equation describes the data best. Estimations of uncertainties of the fitted parameters are not implemented in the scripts yet; the reported uncertainties are derived from the covariance matrix. The results for *apo*- and *holo*-KdpBN are given in table B.4 and B.5, respectively.

For a detailed description of the analysis of ^{15}N relaxation data and a list of related software, see chapter 4.

A.3 The Cyclic Pentapeptide pA$_4$

All peptides used in this work were synthesized by Dr. M. Sukopp.

Peptide Synthesis

Trityl chloride polystyrene resin (PepChem, Tübingen, Germany) was used for peptide synthesis. All fluorenylmethoxycarbonyl- (Fmoc) amino acids and unprotected amino acids were purchased from NovaBiochem (Schwalbach, Germany), Cambridge Isotope Laboratories (Andover, USA), and Martek (Columbia, USA). All reagents and solvents used during the synthesis were reagent grade or better and were used without further purification. Linear peptides were synthesized by SPPS applying the Fmoc strategy.[251,252] Couplings of unlabeled amino acids were performed twice for 1 hour at room temperature in *N*-methyl-pyrrolidone (NMP) using a 3-fold excess of *O*-(Benzotriazol-1-yl)-*N*, *N*, *N*′, *N*′-tetramethyluronium tetrafluoroborate (TBTU) / *N*-hydroxybenzo-

triazole (HOBt) / Fmoc-amino acid and 2.8 equivalents of DIEA in NMP. Labeled amino acids were coupled with an 1.7-fold excess of TBTU / HOBt / Fmoc-amino acid and with 2.8 equivalents of DIEA in NMP overnight. The Fmoc group was removed by sequential treatment of the resins with an excess of 20 % piperidine in DMF for 5 and 15 minutes, respectively. Cleavage of the peptides was achieved by treatment with 20 % hexafluoroisopropanol (HFIP) in dichloromethane (DCM; 3×10 min) and subsequent evaporation. The linear peptides were cyclized in solution (concentration 1 mM) with 3 equivalents diphenyl phosphoryl azide (DPPA) and 5 equivalents sodium bicarbonate in *N,N*-dimethylformamide (DMF) for 10 hours. After filtration and evaporation of the solvent, the cyclic peptides were purified by HPLC using a Pharmacia Biotech Series 900 with a reverse phase C-18 column (YMC-Pack ODS/A column) using water and ACN with 0.1 % TFA as eluents (gradient: 0–50 %; approximately 12.5 min).

ESI-MS and RP-HPLC-ESI-MS analyses were performed using a LCQ-ESI spectrometer (Finnigan) coupled to a Hewlett Packard HP1100 HPLC system. The masses were 396.2 Da for pA_4 with uniformly ^{13}C,^{15}N-labeled Ala residues and 383.1 Da for the selectively labeled peptides.

Long-range HNCO Experiments

For liquid-state NMR experiments, pA_4 was dissolved in d_6-DMSO to yield a final peptide concentration of 1 mM. The residue-specific assignment provided by the earlier work of Kurz was confirmed.[253] All HNCO experiments were acquired as 2D versions with the evolution time of the indirect dimension on C′; no solvent suppression was used (see figure 7.2). In order to estimate the transverse ^{15}N relaxation rate of Ala3, three experiments were recorded with $2T$ set to 102, 170, and 238 ms, respectively. R_2 was obtained from a fit of the signal intensities to equation 7.2, with $^1J_{NC'} = 14.7$ Hz. The chemical shifts of A2C′ and A5C′ differ only by 0.02 ppm; hence, $2T$ was optimized in order to suppress the one-bond correlation A3H^N–A2C′ as good as possible. In order to estimate the magnitude of $^{3h}J_{NC'}$, 2 lr-HNCO spectra were recorded with $2T$

set to 136 and 272 ms, respectively; 48 complex points in the indirect dimension, a relaxation delay of 2 s, and 160 transients correspond to a measurement time of approximately 8.5 h. Fitting of the peak intensities for the correlation A3H^N–A5C′ to equation 7.2 provided ${}^{3h}J_{NC'}$.

Solid-state NMR Experiments

In 1D ssNMR experiments, a cross polarization contact time of 1 ms was used for the transfer $H^N \rightarrow C'$; the relaxation delay between 2 successive scans was set to 4 s. 1D experiments of β-pA_4 and γ-pA_4 for fitting the CSA tensor were acquired at a MAS frequency of 2 kHz collecting 3072 transients, resulting in an acquisition time of approximately 3.5 h. REDOR and reference experiments were recorded in an interleaved fashion at a MAS frequency of 10 kHz using the pulse scheme shown in figure 7.6. 30 points were collected, with a maximum dephasing time of 29 ms. The acquisition time for a set of REDOR experiments is roughly 32 h, with a relaxation delay of 3 s, 2×30 dephasing times, and 640 transients. N–C′ distances were obtained from fitting the REDOR dephasing curves given by $\Delta S/S_0$ against equation 7.6.

Modeling of the Peptide Structures

All peptide structures were modeled using values for ϕ, ψ, and ω as published by Mierke *et al*.[199] and Nikiforovitch *et al*.[201] The structures were minimized using the Discover module from Insight II (Molecular Simulations Inc.). Minimization was performed using the steepest descent algorithm followed by minimization with the conjugate gradient method.

Appendix B

Results of the ^{15}N Relaxation Data Analysis

In this part of the appendix, tables with the results of the model-free analysis for α-ADT (table B.1), *apo*- and *holo*-KdpBN (table B.2 and B.3), as well as the analysis of CPMG dispersion data for *apo*- and *holo*-KdpBN (table B.4 and B.5) are given.

B.1 Model-free Parameters for α-ADT

TABLE B.1: Results of the model-free analysis for α-ADT.

Residue	Model	S^2	S^2_f	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
I4	1	0.779±0.017				1.157
V5	5	0.450±0.028	0.729±0.019	1556±146		0.000
I6						
D7	1	1.000±0.008				8.910
K14	1	0.892±0.030				1.087
M15	1	0.844±0.010				3.944
D17	1	0.791±0.012				0.825
V19	1	0.909±0.019				3.149
N21	1	0.789±0.016				2.027
A22	1	0.818±0.018				3.380
K23	2	0.836±0.027		35±14		0.918
I24	1	0.844±0.025				4.872
A25	1	0.912±0.019				2.896
L26	1	0.866±0.023				5.903
D28	1	0.909±0.021				0.317
S29	5	0.620±0.025	0.831±0.019	1116±123		0.000
A30	2	0.724±0.014		53±7		0.065
L31	4	0.802±0.018		61±11	3.09±0.35	0.000
E32	2	0.830±0.010		42±11		0.019
I33	2	0.829±0.008		37±12		0.773
K34	5	0.541±0.010	0.777±0.011	1049±70		0.000
K35	5	0.406±0.008	0.709±0.018	1236±64		0.000
T36	5	0.379±0.007	0.713±0.018	1160±45		0.000
A40	5	0.222±0.004	0.723±0.011	936±7		0.000
V42	5	0.304±0.008	0.728±0.011	1000±16		0.000
I44	5	0.229±0.004	0.765±0.013	721±12		0.000
S45	5	0.384±0.013	0.695±0.028	621±45		0.000
D46	5	0.242±0.008	0.665±0.024	794±22		0.000
S48	5	0.397±0.005	0.763±0.012	1215±46		0.000
K49	5	0.401±0.009	0.780±0.011	1202±45		0.000
Q51	5	0.321±0.009	0.741±0.012	896±14		0.000
N55	5	0.583±0.011	0.843±0.013	1045±60		0.000
T58	5	0.667±0.007	0.806±0.015	1172±170		0.000
N59	5	0.455±0.018	0.713±0.027	1316±115		0.000
V65	1	0.913±0.014				3.276

Table B.1: continued.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
E66	1	0.912±0.015				5.942
K67	1	0.875±0.008				4.174
I68	3	0.845±0.015			1.14±0.31	1.073
K69	1	0.899±0.011				2.008
S71	3	0.803±0.011			1.49±0.22	1.981
G72	1	0.869±0.012				2.722
A73	3	0.827±0.007			1.45±0.29	0.097
N74	1	0.912±0.022				1.488
L77	1	0.840±0.014				2.537
C78	1	0.863±0.016				4.825
Q79	2	0.896±0.015		61±29		0.811
G81	1	0.874±0.017				0.694
D83	1	0.889±0.015				0.127
D84	1	0.861±0.015				2.095
V85	1	0.906±0.008				1.010
A86	1	0.891±0.020				2.760
Q87	1	0.899±0.020				3.796
H88	1	0.887±0.010				2.044
Y89	1	0.921±0.007				0.538
E93	1	0.869±0.005				3.316
G94	1	0.882±0.011				1.707
I95	2	0.827±0.012		39±10		0.257
Y96	1	0.824±0.011				0.198
A97	1	0.883±0.015				0.675
R99	1	0.892±0.025				10.236
R100	1	0.891±0.019				1.783
K102	1	0.861±0.009				3.304
S104	3	0.767±0.043			1.76±0.58	0.006
D105	3	0.785±0.016			2.10±0.29	3.189
A110	2	0.901±0.021		77±34		0.514
K111	2	0.879±0.018		48±21		0.939
T113	2	0.845±0.034		75±179		2.734
G114	1	0.916±0.022				1.103
A115	1	0.780±0.016				11.196
V118	2	0.839±0.015		60±14		0.055
T119	1	0.860±0.013				8.048
L124	1	0.864±0.007				3.661

TABLE B.1: continued.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
T125	4	0.590±0.016		32±3	2.04±0.25	0.000
S127	1	0.819±0.013				4.506
L129	1	0.917±0.014				1.999
G130	1	0.821±0.024				1.583
E131						
T134	1	0.882±0.034				0.604
V135	1	0.936±0.034				2.662
R138	1	0.869±0.012				15.689
K139						
I140	1	0.720±0.013				14.016
D142	4	0.658±0.023		21±5	0.82±0.42	0.000
R144	5	0.619±0.009	0.734±0.008	538±57		0.000
M145	1	0.813±0.014				5.210
T146	1	0.876±0.023				8.100
F147	2	0.845±0.024		56±19		0.980
V148						
M149	1	0.903±0.031				0.189
G150	2	0.875±0.044		46±199		0.183
C151	1	0.847±0.033				2.196

B.2 Model-free Parameters for KdpBN

Table B.2: Results of the model-free analysis for *apo*-KdpBN.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
Q318						
A319	2	0.913±0.011		18±8		0.184
S320	1	0.884±0.016				3.223
E321	3	0.877±0.014			2.09±0.24	3.121
F322	3	0.861±0.017			2.46±0.38	0.272
I323	1	0.847±0.008				9.420
A325	1	0.896±0.014				1.190
Q326	1	0.876±0.006				4.142
G327	1	0.869±0.012				4.364
V328	1	0.932±0.006				4.074
D329	1	0.875±0.005				0.463
E330						
K331	1	0.903±0.009				7.480
T332	1	0.910±0.006				2.000
L333	1	0.937±0.012				1.454
A334						
D335	1	0.945±0.008				13.388
A336	1	0.943±0.007				2.759
A337	1	0.973±0.005				18.061
Q338	1	0.957±0.005				3.662
L339	1	0.944±0.014				1.694
A340						
S341	1	0.896±0.014				0.043
L342						
A343	1	0.890±0.006				6.107
D344	1	0.905±0.011				0.207
E345	4	0.881±0.029		32±14	12.61±0.62	0.000
T346	2	0.784±0.026		77±37		0.062
E348	1	0.941±0.007				2.128
G349	1	1.000±0.014				4.352
R350	1	0.958±0.008				2.848
S351						
I352	1	1.000±0.015				7.987
V353	1	0.957±0.015				0.700

TABLE B.2: continued.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
I354	1	0.950±0.016				8.722
L355	1	0.964±0.012				2.601
A356	5	0.274±0.013	0.698±0.027	1592±86		0.000
K357	3	0.928±0.012			2.56±0.23	1.046
R359	1	0.901±0.010				0.056
F360	4	0.855±0.021		24±5	1.17±0.44	0.000
N361	1	0.919±0.009				10.071
L362	3	0.937±0.017			1.77±0.33	0.383
R363	5	0.683±0.008	0.838±0.009	1382±80		0.000
E364	5	0.694±0.004	0.906±0.010	1446±53		0.000
R365	5	0.641±0.007	0.865±0.010	1113±36		0.000
V367	5	0.774±0.006	0.890±0.007	1013±60		0.000
Q368	5	0.771±0.009	0.908±0.011	1188±130		0.000
S369	5	0.770±0.008	0.902±0.014	900±102		0.000
L370	5	0.791±0.006	0.923±0.009	904±65		0.000
H371	4	0.931±0.030		616±324	3.49±0.42	0.000
A372	2	0.848±0.005		41±7		0.892
T373	2	0.883±0.009		12±3		2.444
F374	3	0.874±0.014			1.31±0.23	2.120
V375	3	0.850±0.021			1.24±0.41	0.424
F377	5	0.801±0.007	0.871±0.010	1162±172		0.000
T378	2	0.889±0.018		32±121		1.976
Q380	2	0.951±0.005		175±64		0.879
S381	2	0.953±0.016		760±337		0.142
R382	4	0.946±0.022		531±263	24.60±1.13	0.000
M383	4	0.954±0.009		88±48	6.57±0.30	0.000
S384	1	0.947±0.023				3.184
G385	1	0.922±0.024				3.342
N387	1	0.914±0.008				10.598
D389						
R391	2	0.857±0.010		24±4		0.603
M392	2	0.840±0.007		5±1		0.910
I393	1	0.849±0.011				3.771
R394	1	0.889±0.007				11.770
K395	1	0.958±0.011				18.718
G396	1	0.944±0.018				0.488
S397	2	0.911±0.010		19±4		1.051

Table B.2: continued.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
V398	1	0.912±0.012				4.280
D399						
A400	1	0.912±0.006				2.584
I401	1	0.951±0.011				6.876
R402						
H404	1	0.956±0.005				2.316
V405						
E406						
A407						
N408	5	0.756±0.008	0.835±0.007	1223±109		0.000
G409	2	0.835±0.015		41±5		2.831
G410	5	0.759±0.062	0.867±0.039	1250±283		0.000
H411	5	0.749±0.006	0.836±0.008	830±97		0.000
F412	5	0.759±0.022	0.871±0.023	1085±233		0.000
T414	2	0.903±0.010		11±1		3.511
V416	1	0.916±0.012				1.626
D417	1	0.919±0.011				0.683
Q418	1	0.936±0.008				2.602
K419	1	0.919±0.009				0.957
V420						
D421	1	0.962±0.007				5.069
Q422	1	0.933±0.010				8.704
V423	1	0.942±0.005				0.040
A424	1	0.957±0.014				0.310
R425	1	0.927±0.011				2.060
G427	2	0.830±0.009		105±10		0.719
A428	2	0.873±0.003		9±2		0.007
T429	1	0.912±0.004				1.022
L431	1	0.870±0.005				4.467
V432	1	0.908±0.006				2.325
V433	1	0.894±0.005				0.556
V434	1	0.942±0.003				0.132
E435	1	0.893±0.008				2.350
G436	1	0.918±0.011				0.085
S437						
R438	1	0.904±0.013				3.439
V439	1	0.853±0.008				3.896

TABLE B.2: continued.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
L440						
G441						
V442	1	0.910±0.010				5.678
I443	1	0.847±0.011				1.176
A444	1	0.928±0.025				12.660
K446	3	0.897±0.022			1.50±0.34	0.779
D447	2	0.874±0.005		37±7		0.015
I448	5	0.718±0.005	0.889±0.009	1148±100		0.000
V449	5	0.383±0.004	0.794±0.007	1276±21		0.000
K450	5	0.231±0.004	0.785±0.012	1010±9		0.000
G451	5	0.088±0.002	0.588±0.005	650±2		0.000

TABLE B.3: Results of the model-free analysis for *holo*-KdpBN.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
A319	1	0.992±0.010				7.566
S320	1	0.906±0.013				8.737
E321						
F322						
I323	2	0.913±0.014		23±10		0.652
A325	1	0.949±0.013				1.064
Q326	1	0.911±0.013				7.218
G327	1	0.897±0.013				9.446
V328	4	0.833±0.017		26±5	1.54±0.37	0.000
D329	1	0.921±0.013				3.843
E330						
K331	1	0.909±0.013				11.361
T332	1	0.988±0.012				1.931
L333	1	1.000±0.008				8.158
A334	1	0.938±0.013				1.293
D335	1	0.959±0.014				6.385
A336						
A337	1	0.979±0.013				2.032
Q338	1	0.978±0.013				7.285

Table B.3: continued.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
L339	5	0.727±0.017	0.919±0.013	829±60		0.000
A340	1	0.938±0.014				3.744
S341	4	0.250±0.030		1472±26	1.23±0.35	0.000
L342	1	1.000±0.009				18.116
A343	1	0.869±0.013				9.597
D344	1	0.958±0.014				15.904
E345	4	0.859±0.018		28±7	6.12±0.48	0.000
T346	2	0.880±0.011		153±51		0.412
E348	1	0.920±0.013				13.493
G349	3	0.925±0.019			2.56±0.42	0.143
R350	1	0.986±0.012				9.830
S351						
I352	3	0.986±0.016			2.91±0.43	0.357
V353	1	0.961±0.013				1.162
I354	3	0.993±0.013			2.10±0.39	2.229
L355	1	1.000±0.008				13.709
A356	5	0.221±0.008	0.834±0.013	1125±8		0.000
K357	1	0.997±0.009				12.403
R359	1	0.956±0.014				11.964
F360	2	0.898±0.013		78±17		2.913
N361	1	0.952±0.014				15.781
L362						
R363	5	0.708±0.018	0.864±0.013	1254±95		0.000
E364	2	0.733±0.015		1221±56		0.002
R365	5	0.604±0.015	0.841±0.012	1328±57		0.000
D366	5	0.788±0.019	0.911±0.014	773±101		0.000
V367	5	0.782±0.019	0.905±0.013	898±106		0.000
Q368	5	0.898±0.020	0.964±0.014	1090±245		0.000
S369	5	0.732±0.017	0.888±0.013	1026±78		0.000
L370	2	0.894±0.010		153±56		0.124
H371	4	0.918±0.009		140±37	2.46±0.32	0.000
A372	2	0.846±0.012		50±6		2.005
T373	1	0.867±0.013				10.490
F374	1	0.895±0.013				0.781
V375	1	0.941±0.013				1.579
F377	5	0.660±0.016	0.946±0.013	1461±62		0.000
T378	2	0.938±0.014		788±219		0.007

TABLE B.3: continued.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
Q380	2	0.933±0.011		104±65		0.008
S381						
M383	4	0.869±0.018		21±6	1.84±0.42	0.000
S384	1	0.935±0.013				0.827
G385	1	0.966±0.014				1.662
I386	1	0.913±0.012				3.536
N387	1	0.908±0.013				11.139
I388	3	0.973±0.019			3.45±0.46	0.035
D389						
R391	5	0.860±0.020	0.926±0.014	1094±241		0.000
M392	2	0.858±0.013		26±5		2.867
I393	1	0.907±0.013				7.062
R394	1	0.937±0.013				0.423
K395	1	0.942±0.013				0.892
G396	1	0.973±0.013				10.903
S397	1	0.968±0.014				0.717
V398	1	0.970±0.013				10.967
D399	1	0.987±0.012				7.480
A400	1	0.918±0.014				5.168
I401	3	0.912±0.019			0.90±0.39	0.556
R402	1	0.974±0.013				1.897
R403	1	0.943±0.014				9.430
H404	1	0.973±0.013				3.240
V405	3	0.901±0.018			1.57±0.40	1.086
E406	1	0.971±0.014				9.104
A407	2	0.901±0.012		51±11		1.646
N408	5	0.834±0.019	0.918±0.013	1211±194		0.000
G409	2	0.906±0.014		67±17		0.316
G410	2	0.930±0.013		71±28		2.439
H411	2	0.849±0.012		49±7		2.621
F412	2	0.871±0.013		81±13		0.103
V416	1	0.941±0.014				4.117
D417	1	1.000±0.008				9.123
Q418	1	1.000±0.009				6.533
K419	1	0.945±0.013				1.353
V420	3	0.931±0.019			1.25±0.42	0.013
D421	1	0.990±0.011				7.287

Table B.3: continued.

Residue	Model	S^2	S_f^2	τ_i (ps)	R_{ex} (s^{-1})	Γ_i
Q422	3	0.928±0.018			1.55±0.41	1.227
V423	1	1.000±0.009				0.290
A424	1	1.000±0.008				17.006
R425	1	0.953±0.013				11.684
Q426	4	0.825±0.017		24±4	1.45±0.38	0.000
G427	5	0.836±0.019	0.915±0.014	1399±237		0.000
A428	1	0.917±0.013				10.401
T429	1	0.908±0.012				1.997
L431	3	0.928±0.020			1.73±0.41	1.669
V432	1	0.914±0.013				3.456
V433	1	0.940±0.013				1.444
V434	1	0.983±0.012				3.053
E435	1	0.918±0.014				4.126
G436	1	1.000±0.008				17.627
S437	4	0.871±0.018		29±7	1.57±0.40	0.000
R438	3	0.951±0.020			1.15±0.42	0.036
V439	2	0.852±0.013		20±5		1.327
L440	1	0.947±0.014				0.887
G441	1	0.981±0.012				5.601
I443	1	0.910±0.012				3.060
A444						
K446	3	0.894±0.018			1.78±0.39	0.468
I448	5	0.860±0.020	0.970±0.014	970±140		0.000
V449	5	0.462±0.012	0.874±0.013	1135±20		0.000
K450	5	0.272±0.009	0.834±0.013	1030±9		0.000
G451	5	0.121±0.005	0.608±0.010	659±8		0.000

B.3 Exchange Parameters for KdpBN

Table B.4: Results of the relaxation dispersion data analysis for *apo*-KdpBN.

Residue	k_{ex} (s^{-1})	R_{ex} (s^{-1})	p_a	$\Delta\omega$ (rad $\cdot$ s^{-1})	$R_2(750)$ (s^{-1})	$R_2(600)$ (s^{-1})	α
Q318	1102±124		0.986±0.001	2174±51	21.2±0.3	19.9±0.2	0.484
A319	1224±241		0.995±0.001	1489±239	21.8±0.2	21.9±0.2	0.918
S320	446±73		3.100±0.327	21±0	22.1±0.1	55.8±0.0	0.000
E321	1080±108		0.982±0.001	1205±57	21.3±0.2	19.8±0.1	1.011
F322	775±109		0.985±0.001	1424±49	22.9±0.2	21.6±0.1	0.539
I323	801±91	3.6±0.2			20.6±0.1	20.7±0.1	
A325	1392±420	1.0±0.2			22.6±0.1	22.4±0.1	
Q326	749±215	1.3±0.2			20.6±0.1	20.6±0.1	
G327	1726±629	1.4±0.2			22.9±0.2	22.6±0.1	
D329	2234±934	0.8±0.2			20.3±0.2	19.9±0.1	
T332	669±99	2.5±0.2			20.5±0.1	20.1±0.1	
L333	912±311	0.9±0.2			22.6±0.1	22.3±0.1	
A334	466±142	1.4±0.3			21.6±0.1	21.3±0.1	
L339	662±346	0.4±0.2			21.4±0.1	18.0±0.1	
S341	899±87	4.7±0.2			21.1±0.1	21.6±0.1	
D344	514±110	2.2±0.3			19.0±0.1	18.7±0.1	
E345	1095±154	4.4±0.3			40.2±0.2	34.1±0.1	
T346	1278±85	6.4±0.2			23.4±0.1	15.2±0.1	
G349	1574±59	14.4±0.2			23.2±0.2	23.4±0.1	
R350	772±32	11.8±0.3			23.0±0.1	21.5±0.1	
I352	780±101		0.981±0.002	2386±52	23.9±0.3	22.4±0.2	0.230
V353	816±75	4.8±0.2			23.1±0.1	22.8±0.1	
I354	792±73	3.4±0.2			22.9±0.1	13.8±0.1	
K357	552±105	2.1±0.3			23.3±0.1	23.1±0.1	

Table B.4: continued.

Residue	k_{ex} (s^{-1})	R_{ex} (s^{-1})	p_a	$\Delta\omega$ ($rad \cdot s^{-1}$)	$R_2(750)$ (s^{-1})	$R_2(600)$ (s^{-1})	α
E364	1376±448	1.0±0.1			18.1±0.1	17.4±0.1	
R365	977±207	1.3±0.1			16.2±0.1	15.7±0.1	
H371	633±139	2.1±0.3			30.2±0.1	26.7±0.1	
A372	997±346	0.7±0.1			17.9±0.1	17.3±0.1	
V398	670±98	2.3±0.2			22.7±0.1	22.6±0.1	
D399	623±114	2.3±0.2			20.2±0.1	20.0±0.1	
H411	568±187	1.5±0.2			16.7±0.1	16.7±0.1	
F412	2491±663	2.5±0.3			21.6±0.3	21.1±0.2	
D417	781±173	1.6±0.2			21.7±0.1	21.4±0.1	
K419	603±119	2.2±0.3			22.7±0.1	22.4±0.1	
Q422	818±110	3.0±0.2			21.8±0.1	21.1±0.1	
V423	629±63		0.969±0.004	582±53	22.4±0.1	21.7±0.1	1.212
A424	721±40	8.2±0.3			22.5±0.1	21.3±0.1	
R425	218±188		0.987±0.006	443±72	21.5±0.1	21.0±0.1	0.462
G427	869±67	4.8±0.2			19.5±0.1	18.9±0.1	
A428	794±64	4.4±0.2			18.5±0.1	18.0±0.1	
T429	466±72		0.976±0.003	539±56	20.5±0.1	19.9±0.1	0.978
L431	718±56	5.6±0.2			21.1±0.1	21.0±0.1	
A444	972±81		0.976±0.002	951±56	21.5±0.2	21.4±0.1	1.150
K446	681±78		0.983±0.001	1081±48	22.7±0.1	22.4±0.1	0.665
D447	874±56		0.972±0.003	829±49	19.2±0.1	18.0±0.1	1.185
I448	788±62		0.981±0.001	1356±31	16.5±0.1	14.9±0.1	0.592
V449	1014±108	2.2±0.1			10.8±0.1	10.8±0.0	
K450	1191±77	3.0±0.1			7.9±0.1	7.8±0.0	
G451	672±14	4.3±0.1			3.6±0.0	3.2±0.0	

Table B.5: Results of the relaxation dispersion data analysis for *holo*-KdpBN.

Residue	k_{ex} (s^{-1})	R_{ex} (s^{-1})	p_a	$\Delta\omega$ ($rad \cdot s^{-1}$)	$R_2(750)$ (s^{-1})	$R_2(600)$ (s^{-1})	α
A319	1160±257	2.5±0.2			28.4±0.2	23.2±0.1	
S320	1086±370	1.1±0.2			24.2±0.1	20.8±0.1	
E321	72±45		0.906±0.058	1304±41	22.7±0.1	20.6±0.1	0.006
F322	111±63		0.952±0.027	1627±52	23.9±0.1	22.6±0.1	0.010
I323	269±94		0.982±0.002	471±61	23.1±0.1	20.3±0.1	0.577
A325	530±322		0.997±0.001	1372±341	24.1±0.2	21.2±0.1	0.316
Q326	307±80		0.020±0.005	374±76	22.5±0.1	19.5±0.1	0.924
G327	224±44	5.7±1.0			28.7±0.1	26.8±0.1	
T332	459±59	4.5±0.3			22.2±0.1	19.2±0.1	
L333	827±228	1.4±0.2			24.7±0.1	21.4±0.1	
A334	369±104	1.7±0.3			22.6±0.1	19.7±0.1	
A340	868±286	1.6±0.2			22.7±0.1	19.4±0.1	
S341	23±0		0.500±0.009	374±3	9.3±0.0	8.6±0.0	0.000
L342	833±169	2.6±0.2			27.0±0.1	22.7±0.1	
E345	1177±161	4.4±0.2			30.8±0.2	24.8±0.1	
T346	541±71	3.6±0.3			22.6±0.1	18.6±0.1	
E348	325±58	2.8±0.4			24.7±0.1	20.8±0.1	
G349	90±32		0.912±0.031	1217±34	27.3±0.1	23.9±0.1	0.011
R350	183±47		0.955±0.007	536±30	25.0±0.1	21.8±0.1	0.236
S351	468±33	6.9±0.3			23.8±0.1	20.5±0.1	
I352	202±77		0.952±0.018	2465±63	24.9±0.2	22.7±0.2	0.015
V353	254±59		0.969±0.003	424±42	25.2±0.1	21.8±0.1	0.616
I354	19±7		0.655±0.138	697±58	24.7±0.1	21.6±0.1	0.000
K357	148±101		0.979±0.009	427±57	25.2±0.1	21.2±0.1	0.255
F360	510±67	2.7±0.3			23.0±0.1	19.5±0.1	
L362	205±46	5.1±1.1			26.1±0.1	21.2±0.1	

Table B.5: continued.

Residue	k_{ex} (s^{-1})	R_{ex} (s^{-1})	p_a	$\Delta\omega$ ($rad \cdot s^{-1}$)	$R_2(750)$ (s^{-1})	$R_2(600)$ (s^{-1})	α
E364	545±253	1.4±0.3			20.9±0.1	18.3±0.1	
R365	569±70	3.3±0.2			17.7±0.1	14.9±0.1	
A372	966±268	1.0±0.2			20.2±0.1	16.5±0.1	
T373	1480±440	1.5±0.2			21.8±0.2	18.7±0.1	
F374	1946±426	2.1±0.2			22.5±0.2	19.0±0.1	
M383	432±170	2.4±0.5			25.1±0.1	21.3±0.1	
I388	605±105	3.3±0.3			30.4±0.1	25.3±0.1	
M392	1899±384	2.2±0.2			21.7±0.2	18.3±0.1	
K395	458±161	1.4±0.3			24.7±0.1	21.7±0.1	
G396	2005±476	2.3±0.2			25.3±0.2	21.4±0.2	
V398	224±57		0.962±0.005	586±35	25.4±0.1	22.0±0.1	0.295
D399	176±71		0.953±0.016	818±28	24.2±0.1	20.3±0.1	0.098
R402	753±186	2.3±0.3			24.8±0.1	21.2±0.1	
E406	162±138		0.987±0.005	378±97	25.0±0.1	21.6±0.1	0.372
H411	191±43		0.957±0.006	598±23	19.2±0.1	17.1±0.1	0.210
F412	139±80		0.946±0.029	958±33	28.1±0.1	26.8±0.1	0.044
D417	230±101		0.982±0.003	400±61	23.7±0.1	20.4±0.1	0.585
Q418	458±65	4.1±0.4			23.9±0.1	20.3±0.1	
K419	310±73		0.032±0.012	309±79	24.8±0.1	21.4±0.1	1.135
D421	259±70		0.977±0.003	533±50	24.5±0.1	20.9±0.1	0.451
Q422	242±55		0.971±0.003	560±42	23.5±0.1	20.3±0.1	0.369
V423	87±31		0.913±0.031	958±25	24.1±0.1	21.5±0.1	0.016
A424	16±7		0.654±0.174	871±67	22.1±0.1	19.4±0.1	0.000
R425	283±74		0.026±0.008	324±73	23.2±0.1	20.1±0.1	0.987
G427	237±77		0.977±0.004	578±42	21.4±0.1	18.3±0.1	0.340
A428	260±45		0.970±0.002	497±32	20.4±0.1	17.4±0.1	0.502
T429	120±70		0.934±0.034	704±27	21.9±0.1	19.4±0.1	0.059

Table B.5: continued.

Residue	k_{ex} (s^{-1})	R_{ex} (s^{-1})	p_a	$\Delta\omega$ ($rad \cdot s^{-1}$)	$R_2(750)$ (s^{-1})	$R_2(600)$ (s^{-1})	α
L431	246±53		0.963±0.005	731±33	27.3±0.1	22.4±0.1	0.235
V432	516±117	2.6±0.3			23.5±0.1	20.5±0.1	
V433	606±145	2.6±0.3			24.1±0.1	21.3±0.1	
V434	595±52	6.3±0.3			24.8±0.1	21.3±0.1	
E435	804±130	3.6±0.2			23.2±0.1	20.1±0.1	
V439	1586±401	2.6±0.2			22.1±0.2	19.3±0.1	
L440	2273±709	1.6±0.2			25.5±0.3	22.4±0.2	
I443	818±161	3.0±0.2			23.4±0.1	20.6±0.1	
A444	112±56		0.924±0.037	1083±25	23.6±0.1	21.0±0.1	0.021
K446	56±23		0.855±0.060	1110±30	22.8±0.1	20.6±0.1	0.004
I448	98±37		0.919±0.030	2091±51	19.1±0.1	18.9±0.1	0.004
V449	939±267	1.1±0.1			12.4±0.1	10.9±0.1	
K450	343±36	3.2±0.2			10.2±0.0	9.7±0.0	
G451	227±14		0.968±0.001	441±10	3.8±0.0	3.4±0.0	0.489

Appendix C

Examples

C.1 Model-free Analysis

Fit Relaxation Rates

Required files: Output from the SPARKY command `rh` (`T1_600.rh`, T2_600.rh).

Sequence of commands:

```
sparky2rate.pl T1_600.rh
sparky2rate.pl T2_600.rh
mv T1_600.rh.rates.xmgr R1.600
mv T2_600.rh.rates.xmgr R2.600
```

Calculate R_2/R_1 Ratio

Required files:

- files containing the relaxation rates
  `R1.600`, `R2.600`

- file containing the hetNOE values
  `NHNOE.600`

Command:

```
create_r2r1.csh R1.600 R2.600 NHNOE.600 0.7
```

Rotate PDB File
Required file: PDB coordinates (`kdp.pdb`).

Command:

```
pdbinertia -rt kdp.pdb kdp_rot.pdb > pdbinertia.log
```

Estimate the Rotational Diffusion Tensor using R2R1_Diffusion
Required files:

- control file for R2R1_DIFFUSION
  `r2r1_diff.ctrl`
- file containing the R_2/R_1 values
  `R2R1.input`

Content of `r2r1_diff.ctrl`:

```
# example control file for R2R1_diffusion
5 600.13 100
1.0e+9 1.1 20 0
1.0 1.5 20
R2R1.input
kdp_rot.pdb
kdp_axsym_r2r1.pdb
```

Command:

```
r2r1_diffusion r2r1_diff.ctrl > r2r1_diff.log
```

Estimate the Rotational Diffusion Tensor using Quadric_Diffusion

Required files:

- file containing the R_2/R_1 values
  `R2R1.input`
- file containing the local correlation times
  `tm.dat`
- control file for Quadric_Diffusion
  `quadric.ctrl`

First, calculate local correlation times:

```
r2r1_tm > tm.dat
```

Content of `tm.dat`:

```
# R2/R1 Analysis:
#
# nucleus:   15N
# bond length (A):     1.0200
# CSA (ppm):    -160.00
#
#residue      tm     dtm
318   11.7887      0.1131
321   11.5759      0.1075
333   10.8482      0.1608
345   15.0847      0.2807
381   10.4297      0.3395
```

Next, run Quadric_Diffusion.

Content of `quadric.ctrl`:

```
# example control file for quadric_diffusion
0.1 1.2 40
1 'N'
tm.dat
kdp_rot.pdb
kdp_ax_qdr.pdb
kdp_an_qdr.pdb
```

Command:

```
quadric_diffusion quadric.ctrl > quadric.log
```

Run FASTModelfree

Required files:

- files containing relaxation rates
  `R1.600, R2.600, NHNOE.600`
- pdb coordinate file (only in the case of axially symmetric tumbling)
- FMF configuration file
  `FMF.config`

Create `FMF.config`:

```
setupFMF
```

Content of `FMF.config`:

```
tensor isotropic
cutoff 0.95
Fcutoff 0.80
optimize Yes
maxloop 10
almost1 20
S2cutoff 0.7
```

```
seed 1975
numsim 100
jobname iso
gamma -2.71
rNH 1.02
N15CSA -160
tm 11.6
tmMin 9.0
tmMax 14.0
tmGrid 40
tmConv 0.001
Dratio 1.22
DratioMin 1.0
DratioMax 1.3
DratioGrid 10
DratioConv 0.001
Theta 15
ThetaMin 0
ThetaMax 40
ThetaGrid 5
ThetaConv 0.001
Phi 0
PhiMin 0
PhiMax 360
PhiGrid 20
PhiConv 0.001
model1only No
mpdb kdp.pdb
file{0}{R1} ../r1.600
file{0}{R2} ../r2.600
file{0}{NOE} ../nhnoe.600
```

```
file{0}{field} 600
```

Finally, run FMF:

```
fastFMF > example_iso.log
```

Content of `iso.log`:

```
==================================================
Iteration 1
Tensor: tm 11.6
Model 1 spins:
321 333
Model 2 spins:
381
Model 3 spins:

Model 4 spins:
345
Model 5 spins:
451
Unassigned spins:
318
Delta tm: 0.2200 Converged? NO
==================================================
Iteration 2
Tensor: tm 11.380
Model 1 spins:
321 333
Model 2 spins:
381
Model 3 spins:
```

```
Model 4 spins:
345
Model 5 spins:
451
Unassigned spins:
318
Delta tm: 0.0000 Converged? YES
```

A listing of the FMF protocol file (`example_iso.log`) is not given here due to its length of more than 88 pages.

Extract Optimized Parameters
Required files:

- final output file of FMF
  `mfout.final`
- file containing parameters of iteration 2
  `iso.iter2.par`

Extract the rotational diffusion tensor:

```
get_tensor.awk
```

Extract the final MF-parameters:

```
get_all ./
```

Extract the MF-parameters of iteration 2:

```
get_all_iter iso.iter2.par ./
```

Congratulations, you made it!! Now that you've done a nice model-free analysis of ^{15}N relaxation data, it's about time to relax and have some coffee before you move on to the next quest... :-)

C.2 CPMG Dispersion Data Analysis

Create Scilab input files

Required files:

- peak lists for two B_0 fields
  `disp1600_"*".list`, `disp2600_"*".list`
- file containing residues to be analyzed
  `residues.dat`

Sequence of commands:

```
./dispersion_sp.awk disp1600_"*".list disp2600_"*".list 15 52
./extract_R2eff.awk R2eff.dat residues.dat 600
mv R2eff.dat R2eff.600

./dispersion_sp.awk disp1750_"*".list disp2750_"*".list 15 41
./extract_R2eff.awk R2eff.dat residues.dat 750
mv R2eff.dat R2eff.750

./reformat.csh 600 ./
./reformat.csh 750 ./
```

Resulting files for further use:

```
R350.600.mtx, R403.600.mtx, D447.600.mtx, all_res.600.mtx
R350.750.mtx, R403.750.mtx, D447.750.mtx, all_res.750.mtx
```

Fit Data with Scilab

Required files:

- Data files
  `all_res.600.mtx`, `all_res.750.mtx`

- Scilab scripts
  `R2.sce, fast.sce, full.sce`
- Fit functions
  `fitfun_R2.sce, fitfun_fast.sce, fitfun_full.sce`
- Auxilliary Scilab functions
  `fminuit.scimex, load_fminuit.sce, mkscale.sci`

Required modifications of the fitting scripts are described in section 4.3.2.
Start Scilab:

```
scilab
```

Fit data against a flat line, the fast-exchange approximation and the general expression (type commands in Scilab's main window):

```
exec R2.sce
exec fast.sce
exec full.sce
```

F-testing for chemical exchange (`R2.out`, `fast.out`, and `full.out` need to be copied into one directory)

```
F.csh fast R2 28 4 2
```

yields

```
350      2691.43            8.58         3753.82        ###
403        29.00            7.25           36.00
447      5257.20           24.59         2553.85        ###
```

indicating that Arg350 and Asp447 are subject to chemical exchange. Now, the exchange regime has to be determined:

```
F.csh full fast 28 5 4
```

gets the following results:

```
350         8.58            8.52            0.15
403         7.25            7.25            0.00
447        24.59           15.10           14.45        #
```

A significant improvement when using the general equation instead of the fast-exchange approximation is suggested only for Asp447.

The results of the relaxation dispersion analysis for Arg350, Arg403, and Asp447 can be summarized as follows:

- Arg403 is not subject to chemical exchange.
- Arg350 and Asp447 are affected by chemical exchange.
- Asp447 is in the intermediate exchange regime ($\alpha = 1.15$).

Xmgrace Header Example File

```
@ VIEW XMIN 0.20
@ VIEW XMAX 0.90
@ VIEW YMIN 0.35
@ VIEW YMAX 0.85
@ WORLD XMIN       0.000E+00
@ WORLD XMAX       0.200E+01
@ XAXIS TICK MAJOR       0.500E+00
@ XAXIS TICK MINOR       0.250E+00
@ XAXIS LABEL "1/\f{Symbol}t\f{Helvetica}\scp\N   (1/s)"
@ YAXIS LABEL "R\s2\N(1/\f{Symbol}t\f{Helvetica}\scp\N)   (1/s)"
@ S0 TYPE xydy
@ S0  LINESTYLE 0
@ S0  COLOR 1
@ S0  SYMBOL 1
@ S0  SYMBOL SIZE 0.5
@ S0  SYMBOL FILL 1
```

```
@ S0  SYMBOL COLOR 1
@ S0  ERRORBAR LENGTH 0.5
@ S0  ERRORBAR COLOR 1
@ S1 TYPE xydy
@ S1  LINESTYLE 0
@ S1  COLOR 1
@ S1  SYMBOL 1
@ S1  SYMBOL SIZE 0.5
@ S1  SYMBOL FILL 0
@ S1  SYMBOL COLOR 1
@ S1  ERRORBAR LENGTH 0.5
@ S1  ERRORBAR COLOR 1
@ S2 TYPE xy
@ S2  ERRORBAR LENGTH 0
@ S2  COLOR 1
@ S3 TYPE xy
@ S3  ERRORBAR LENGTH 0
@ S3  COLOR 1
```

Bibliography

[1] Shuker, S. B.; Hadjuk, P. J.; Meadows, R. P.; Fesik, S. W.; *Science* **1996,** *274,* 1531–1534.

[2] Zerbe, O., Ed.; *BioNMR in Drug Research;* volume 16 of *Methods and Principles in Medicinal Chemistry* Wiley: Weinheim, 2003.

[3] Coles, M.; Heller, M.; Kessler, H.; *Drug Discov. Today* **2003,** *8,* 803–810.

[4] Pellechia, M.; Sem, D. H.; Wüthrich, K.; *Nat. Rev. Drug. Discov.* **2002,** *1,* 211–219.

[5] Diercks, T.; Coles, M.; Kessler, H.; *Curr. Opin. Chem. Biol.* **2001,** *5,* 285–291.

[6] Rüdisser, S.; Jahnke, W.; *Comb. Chem. High. Throughput. Screen.* **2002,** *5,* 591–603.

[7] Tugarinov, V.; Muhandrim, R.; Ayed, A.; Kay, L. E.; *J. Am. Chem. Soc.* **2002,** *124,* 10025–10035.

[8] Tugarinov, V.; Kay, L. E.; *J. Am. Chem. Soc.* **2003,** *125,* 13868–13878.

[9] Kempf, J. G.; Jung, J.-Y.; Sampson, N. S.; Loria, J. P.; *J. Am. Chem. Soc.* **2003,** *125,* 12064–12065.

[10] Korzhnev, D. M.; Kloiber, K.; Kanelis, V.; Tugarinov, V.; Kay, L. E.; *J. Am. Chem. Soc.* **2004,** *126,* 3964–3973.

[11] Karplus, M.; McCammon, J. A.; *Annu. Rev. Biochem.* **1983,** *53,* 263–300.

[12] Fersht, A.; *Enzyme Structure and Mechanism;* W. H. Freeman & Co.: New York, 1985.

[13] Jackson, E. S.; *Folding Des.* **1998,** *3,* 81–91.

[14] Vugmeyster, L.; Kroenke, C. D.; Picart, F.; Palmer, A. G.; Raleigh, D. P.; *J. Am. Chem. Soc.* **2000,** *122,* 5387–5388.

[15] Dyson, H. J.; Wright, P. E.; *Annu. Rev. Phys. Chem.* **1996,** *47,* 369–395.

[16] Mittag, T.; Schaffhausen, B.; Günther, U. L.; *Biochemistry* **2003,** *42,* 11128–11136.

[17] Atkinson, R. A.; Kieffer, B.; *Prog. Nucl. Magn. Reson. Spectrosc.* **2004,** *44,* 141–187.

[18] Wand, A. J.; *Nat. Struct. Biol.* **2001,** *8,* 926–931.

[19] Patzelt, H.; Goto, N.; Iwai, H.; Lundstrom, K.; Fernholz, E.; Modern Methods for the Expression of Proteins in Isotopically Enriched Form. In *BioNMR in Drug Research*, Vol. 16; Zerbe, O., Ed.; Wiley: Weinheim, 2003.

[20] Goto, N. K.; Kay, L. E.; *Curr. Opin. Struct. Biol.* **2000,** *10,* 585–592.

[21] Lipari, G.; Szabo, A.; *J. Am. Chem. Soc.* **1982,** *104,* 4546–4559.

[22] Lipari, G.; Szabo, A.; *J. Am. Chem. Soc.* **1982,** *104,* 4559–4570.

[23] Powers, R.; Clore, G. M.; Garret, D. S.; Gronenborn, A. M.; *J. Magn. Reson. B* **1993,** *101,* 325–327.

[24] Case, D. A.; *Acc. Chem. Res.* **2002,** *35,* 325–331.

[25] McCammon, J. A.; Gelin, B. R.; Karplus, M.; *Nature* **1977,** *267,* 585–590.

[26] Palmer, A. G.; *Curr. Opin. Struct. Biol.* **1997,** *7,* 732–737.

[27] Kay, L. E.; *Nat. Struct. Biol.* **1998,** *5,* 513–517.

[28] Frueh, D.; *Prog. Nucl. Magn. Reson. Spectrosc.* **2002,** *41,* 305–324.

[29] Palmer, A. G.; *Annu. Rev. Biophys. Biomol. Struct.* **2001,** *30,* 129–155.

[30] Cole, R.; Loria, J. P.; *J. Biomol. NMR* **2003,** *26,* 203–213.

[31] Dosset, P.; Hus, J.-C.; Blackledge, M.; Marion, D.; *J. Biomol. NMR* **2000,** *16,* 23–28.

[32] Mandel, A. M.; Akke, M.; Palmer, A. G.; *J. Mol. Biol.* **1995,** *246,* 144–163.

[33] Palmer, A. G.; Rance, M.; Wright, P. E.; *J. Am. Chem. Soc.* **1991,** *113,* 4371–4380.

[34] Fischer, M. W. F.; Majumdar, A.; Zuiderweg, E. R. P.; *Prog. Nucl. Magn. Reson. Spectrosc.* **1998,** *33,* 207–272.

[35] Yamazaki, T.; Muhandiram, R.; Kay, L. E.; *J. Am. Chem. Soc.* **1994,** *116,* 8266–8278.

[36] Pang, Y.; Buck, M.; Zuiderweg, E. R. P.; *Biochemistry* **2002,** *41,* 2655–2666.

[37] Fischer, M. W. F.; Zeng, L.; Pang, Y.; Hu, W.; Majumdar, A.; Zuiderweg, E. R. P.; *J. Am. Chem. Soc.* **1997,** *119,* 12629–12642.

[38] Kushlan, D. M.; LeMaster, D. M.; *J. Am. Chem. Soc.* **1993,** *115,* 11026–11027.

[39] Kushlan, D. M.; LeMaster, D. M.; *J. Am. Chem. Soc.* **1996,** *118,* 9255–9264.

[40] Muhandiram, D. R.; Yamazaki, T.; Sykes, B. D.; Kay, L. E.; *J. Am. Chem. Soc.* **1995,** *117,* 11536–11544.

[41] Yang, D. W.; Mittermaier, A.; Mok, Y. K.; Kay, L. E.; *J. Mol. Biol.* **1998,** *276,* 939–954.

[42] Palmer, A. G.; Kroenke, C. D.; Loria, J. P.; *Methods Enzymol.* **2001,** *339,* 204–238.

[43] Cavanagh, J.; Venters, R. A.; *Nat. Struct. Biol.* **2001,** *8,* 912–914.

[44] Sandstrom, J.; *Dynamic NMR Spectroscopy;* Academic Press: London, 1982.

[45] Johnson, C. S. J.; . In *Advances in Magnetic Resonance*, Vol. 1; Waugh, J. S., Ed.; Academic Press: New York, 1965.

[46] Kaplan, J. I.; Fraenkel, G.; *NMR of Chemically Exchanging Systems;* Academic Press: New York, 1980.

[47] Woessner, D. E.; *Concepts Magn. Reson.* **1996,** *8,* 397–421.

[48] Johnston, E. R.; *Concepts Magn. Reson.* **1995,** *7,* 219–242.

[49] Douglass, D. C.; Jones, G. P.; *J. Chem. Phys.* **1966,** *45,* 956–963.

[50] Luz, Z.; Meiboom, S.; *J. Chem. Phys.* **1963,** *39,* 366–370.

[51] Akke, M.; Palmer, A. G.; *J. Am. Chem. Soc.* **1996,** *118,* 911–912.

[52] Massi, F.; Johnson, E.; Wang, C.; Rance, M.; Palmer, A. G.; *J. Am. Chem. Soc.* **2004,** *126,* 2247–2256.

[53] Loria, J. P.; Rance, M.; Palmer, A. G.; *J. Am. Chem. Soc.* **1999,** *121,* 2331–2332.

[54] Tollinger, M.; Skrynnikov, N. R.; Mulder, F. A. A.; Forman-Kay, J. D.; Kay, L. E.; *J. Am. Chem. Soc.* **2001,** *123,* 11341–11352.

[55] Grey, M. J.; Wang, C.; Palmer, A. G.; *J. Am. Chem. Soc.* **2003,** *125,* 14324–14335.

[56] Skrynnikov, N. R.; Mulder, F. A. A.; Hon, B.; Dahlquist, F. W.; Kay, L. E.; *J. Am. Chem. Soc.* **2001,** *123,* 4556–4566.

[57] Mulder, F. A. A.; Hon, B.; Mittermaier, A.; Dahlquist, F. W.; Kay, L. E.; *J. Am. Chem. Soc.* **2002,** *124,* 1443–1451.

[58] Ishima, R.; Barber, J.; Louis, J. M.; Torchia, D. A.; *J. Biomol. NMR* **2004,** *29,* 187–198.

[59] Ishima, R.; Torchia, D. A.; *J. Biomol. NMR* **2003,** *25,* 243–248.

[60] Orekhov, V. Y.; Korzhnev, D. M.; Kay, L. E.; *J. Am. Chem. Soc.* **2004,** *126,* 1886–1891.

[61] Dittmer, J.; Bodenhausen, G.; *J. Am. Chem. Soc.* **2004,** *126,* 1314–1315.

[62] Wist, J.; Frueh, D.; Tolman, J. R.; Bodenhausen, G.; *J. Biomol. NMR* **2004,** *28,* 263–272.

[63] Luginbühl, P.; Wüthrich, K.; *Prog. Nucl. Magn. Reson. Spectrosc.* **2002,** *40,* 199–247.

[64] Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G.; Skelton, N. J.; *Protein NMR Spectroscopy – Principles and Practice;* Academic Press: San Diego, 1996.

[65] Levitt, M. H.; *Spin Dynamics – Basics of Nuclear Magnetic Resonance;* Wiley: New York, 2001.

[66] Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R.; *J. Chem. Phys.* **1979,** *71,* 4546–4553.

[67] Peng, J.; Wagner, G.; *Biochemistry* **1995,** *34,* 16733–16752.

[68] Reif, B.; Hennig, M.; Griesinger, C.; *Science* **1997,** *276,* 1230–1233.

[69] Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K.; *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94,* 12366–12371.

[70] Boyd, J.; Hommel, U.; Campbell, I. D.; *Chem. Phys. Lett.* **1990,** *175,* 477–482.

[71] Kay, L. E.; Nicholson, L. K.; Delaglio, F.; Bax, A.; Torchia, D. A.; *J. Magn. Reson.* **1992,** *97,* 359–375.

[72] Goldman, M.; *J. Magn. Reson.* **1984,** *60,* 60–60.

[73] Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E.; *Biochemistry* **1994,** *33,* 5984–6003.

[74] Morris, G. A.; Freeman, R.; *J. Am. Chem. Soc.* **1979,** *101,* 760–762.

[75] Grzesiek, S.; Bax, A.; *J. Magn. Reson. B* **1993,** *102,* 103–106.

[76] Logan, T. M.; Olejniczak, E. T.; Xu, R. X.; Fesik, S. W.; *J. Biomol. NMR* **1993,** *3,* 225–231.

[77] Piotto, M.; Saudek, V.; Sklenar, V.; *J. Biomol. NMR* **1992,** *2,* 661–665.

[78] Carr, H. Y.; Purcell, E. M.; *Phys. Rev.* **1954,** *94,* 630–638.

[79] Meiboom, S.; Gill, D.; *Rev. Sci. Instrum.* **1958,** *29,* 688–691.

[80] Shaka, A. J.; Barker, P. B.; Freeman, R.; *J. Magn. Reson.* **1985,** *64,* 547–552.

[81] Markley, J. L.; Horsley, W. J.; Klein, M. P.; *J. Chem. Phys.* **1971,** *55,* 3604–3605.

[82] States, D. J.; Haberkorn, R. A.; Ruben, D. J.; *J. Magn. Reson.* **1982,** *48,* 286–292.

[83] Grzesiek, S.; Bax, A.; *J. Am. Chem. Soc.* **1993,** *115,* 12593–12594.

[84] Skelton, N. J.; Palmer, A. G.; Akke, M.; Kördel, J.; Rance, M.; Chazin, W. J.; *J. Magn. Reson. B* **1993,** *102,* 253–264.

[85] Skelton, N. J.; Palmer, A. G.; Akke, M.; Kördel, J.; Rance, M.; Chazin, W. J.; *J. Magn. Reson. B* **1993,** *102,* 308–332.

[86] Baguet, E.; Roby, C.; *J. Magn. Reson.* **1997,** *128,* 149–160.

[87] Renner, C.; Schleicher, M.; Moroder, L.; Holak, T. A.; *J. Biomol. NMR* **2002,** *23,* 23–33.

[88] Idiyatullin, D.; Daragan, V. A.; Mayo, K. H.; *J. Magn. Reson.* **2001,** *333,* 333–333.

[89] Li, Y.-C.; Montelione, G. T.; *J. Magn. Reson. B* **1994,** *105,* 45–51.

[90] Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M.; *J. Am. Chem. Soc.* **1990,** *112,* 4989–4991.

[91] Korzhnev, D. M.; Orekhov, V. Y.; Arseniev, A. S.; *J. Magn. Reson.* **1997,** *127,* 184–191.

[92] Tugarinov, V.; Liang, Z.; Shapiro, Y. E.; Freed, J. H.; Meirovitch, E.; *J. Am. Chem. Soc.* **2001,** *123,* 3055–3063.

[93] Larsson, G.; Martinez, G.; Schleucher, J.; Wijmenga, S. S.; *J. Biomol. NMR* **2003,** *27,* 291–312.

[94] Lipari, G.; Szabo, A.; *Biophys. J.* **1980,** *30,* 489–506.

[95] Lipari, G.; Szabo, A.; *J. Chem. Phys.* **1981,** *75,* 2971–2976.

[96] Brüschweiler, R.; Wright, P. E.; *J. Am. Chem. Soc.* **1994,** *116,* 8426–8427.

[97] Clore, G. M.; Driscoll, P. C.; Wingfield, P. T.; Gronenborn, A. M.; *Biochemistry* **1990,** *29,* 7387–7401.

[98] Woessner, D. E.; *J. Chem. Phys.* **1962,** *37,* 647–654.

[99] Brüschweiler, R.; Liao, X.; Wright, P. E.; *Science* **1995,** *268,* 886–889.

[100] Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A.; *J. Am. Chem. Soc.* **1995,** *117,* 12562–12566.

[101] Zheng, Z.; Czaplicki, J.; Jardetzky, O.; *Biochemistry* **1995,** *34,* 5212–5223.

[102] Kay, L. E.; Torchia, D. A.; Bax, A.; *Biochemistry* **1989,** *28,* 8972–8979.

[103] Lee, L. K.; Rance, M.; Chazin, W. J.; Palmer, A. G.; *J. Biomol. NMR* **1997,** *9,* 287–298.

[104] Kneller, J. M.; Lu, M.; Bracken, C.; *J. Am. Chem. Soc.* **2002,** *124,* 1852–1853.

[105] Creighton, T. E.; *Proteins. Structures and Molecular Properties;* W. H. Freeman & Co.: New York, 1993.

[106] Kessler, H.; *Angew. Chem. Int. Ed. Engl.* **1970,** *9,* 219–235.

[107] Farrar, T. C.; Becker, E. D.; *Pulse and Fourier Transform NMR;* Academic Press: New York, 1971.

[108] Binsch, G.; Kessler, H.; *Angew. Chem. Int. Ed. Engl.* **1980,** *19,* 411–428.

[109] Wang, C.; Rance, M.; Palmer, A. G.; *J. Am. Chem. Soc.* **2003,** *125,* 8968–8969.

[110] Ishima, R.; Torchia, D. A.; *J. Biomol. NMR* **1999,** *14,* 369–372.

[111] Mulder, F. A. A.; van Tilborg, P. J. A.; Kaptein, R.; Boelens, R.; *J. Biomol. NMR* **1993,** *13,* 275–288.

[112] Szyperski, T.; Luginbühl, P.; Otting, G.; Güntert, P.; Wüthrich, K.; *J. Biomol. NMR* **1993,** *3,* 151–164.

[113] Davis, D. G.; Perlman, M. E.; London, R. E.; *J. Magn. Reson. B* **1994,** *104,* 266–275.

[114] Ishima, R.; Wingfield, P. T.; Stahl, S. J.; Kaufmann, J. D.; Torchia, D. A.; *J. Am. Chem. Soc.* **1998,** *120,* 10534–10542.

[115] Carver, J. P.; Richards, R. E.; *J. Chem. Phys.* **1972,** *6,* 89–105.

[116] Jen, J.; *J. Magn. Reson. B* **1978,** *30,* 111–128.

[117] Farrow, N. A.; Zhang, O.; Szabo, A.; Torchia, D. A.; Kay, L. E.; *J. Biomol. NMR* **1995,** *6,* 153–162.

[118] Phan, I. Q. H.; Boyd, J.; Campbell, I. D.; *J. Biomol. NMR* **1996,** *8,* 369–378.

[119] Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E.; *J. Biomol. NMR* **1994,** *4,* 727–734.

[120] Kroenke, C. D.; Rance, M.; Palmer, A. G.; *J. Am. Chem. Soc.* **1999,** *121,* 10119–10125.

[121] Kroenke, C. D.; Loria, J. P.; Lee, L. K.; Rance, M.; Palmer, A. G.; *J. Am. Chem. Soc.* **1998,** *120,* 7905–7915.

[122] Mandel, A. M.; Akke, M.; Palmer, A. G.; *Biochemistry* **1996,** *35,* 16009–16023.

[123] Millet, O.; Loria, J. P.; Kroenke, C. D.; Pons, M.; Palmer, A. G.; *J. Am. Chem. Soc.* **2000,** *122,* 2867–2877.

[124] Gutowsky, H. S.; Vold, R. L.; Wells, E. J.; *J. Chem. Phys.* **1965,** *43,* 4107–4125.

[125] Palmer, A. G.; Skelton, N. J.; Chazin, W. J.; Wright, P. E.; Rance, M.; *Mol. Phys.* **1992,** *75,* 699–711.

[126] Orekhov, V. Y.; Pervushin, K. V.; Arseniev, A. S.; *Eur. J. Biochem.* **1994,** *219,* 887–896.

[127] Shaka, A. J.; Keeler, J.; *Prog. Nucl. Magn. Reson. Spectrosc.* **1987,** *19,* 47–129.

[128] Mulder, F. A. A.; Skrynnikov, N. R.; Hon, B.; Dahlquist, F. W.; Kay, L. E.; *J. Am. Chem. Soc.* **2001,** *123,* 967–975.

[129] Hemmingsen, S. M.; Woolford, C.; van der Vies, S. M.; Tilly, K.; Dennis, D. T.; Georgopoulos, C.; *Nature* **1988,** *333,* 330–334.

[130] Bukau, B.; Horwich, A. L.; *Cell* **1998,** *92,* 351–366.

[131] Klunker, D.; Haas, B.; Hirtreiner, A.; Figueiredo, L.; Naylor, D. J.; Pfeifer, G.; *J. Biol. Chem.* **2003,** *278,* 33256–33267.

[132] Fenton, W. A.; Horwich, A. L.; *Protein Sci.* **1997,** *6,* 743–760.

[133] Sigler, P. B.; Xu, Z. H.; Rye, H. S.; Burston, S. G.; Fenton, W. A.; Horwich, A. L.; *Annu. Rev. Biochem.* **1998,** *67,* 581–608.

[134] Hartl, F. U.; Hayer-Hartl, M.; *Science* **2002,** *295,* 1852–1858.

[135] Willison, K. R.; Horwich, A. L.; Structure and Function of Chaperonins in Archaebacteria and Eukaryotic Cytosol. In *The Chaperonins*; Ellis, R. J., Ed.; Academic Press: San Diego, CA, 1996.

[136] Gutsche, I.; Essen, L. O.; Baumeister, W.; *J. Mol. Biol.* **1999,** *293,* 295–312.

[137] Kubota, H.; *Vitam. Horm.* **2002,** *65,* 313–331.

[138] Gupta, R. S.; *Biochem. Int.* **1990,** *20,* 833–841.

[139] Kubota, H.; Hynes, G.; Carne, A.; Ashworth, A.; Willison, K.; *Curr. Biol.* **1994,** *4,* 89–99.

[140] Archibald, J. M.; Blouin, C.; Doolittle, W. F.; *J. Struct. Biol.* **2001,** *135,* 157–169.

[141] Braig, K.; Otwinowski, Z.; Hedge, R.; Boisvert, D. C.; Joachimiak, A.; Horwich, A. L.; Sigler, P. B.; *Nature* **1994,** *371,* 578–586.

[142] Waldmann, T.; Lupas, A.; Kellermann, J.; Peters, J.; Baumeister, W.; *Biol. Chem. H.-S.* **1995,** *376,* 119–126.

[143] Roseman, A. M.; Chen, S.; White, H.; Braig, K.; Saibil, H. R.; *Cell* **1996,** *87,* 241–251.

[144] Klumpp, M.; Baumeister, W.; Essen, L. O.; *Cell* **1997,** *91,* 263–270.

[145] Ditzel, L.; Lowe, J.; Stock, D.; Stetter, K. O.; Huber, H.; Huber, R.; Steinbacher, S.; *Cell* **1998,** *93,* 125–138.

[146] Bosch, G.; Baumeister, W.; Essen, L. O.; *J. Mol. Biol.* **2000,** *301,* 19–25.

[147] Pappenberger, G.; Wilsher, J. A.; Roe, S. M.; Counsell, D. J.; Willison, K. R.; Pearl, L. H.; *J. Mol. Biol.* **2002,** *318,* 1367–1379.

[148] Kraulis, P. J.; *J. Appl. Crystallogr.* **1991,** *24,* 946–950.

[149] Merrit, E. A.; Murphy, M. E. P.; *Acta Crystallographica. Section D: Biological Crystallography* **1994,** *50,* 869–873.

[150] Fenton, W. A.; Kashi, Y.; Furtak, K.; Horwich, A. L.; *Nature* **1994,** *371,* 614–619.

[151] Feltham, J. L.; Gierasch, L. M.; *Cell* **2000,** *100,* 193–196.

[152] Hua, Q. X.; Dementieva, I. S.; Walsh, M. A.; Hallenga, K.; Weiss, M. A.; Joachimiak, A.; *J. Mol. Biol.* **2001,** *306,* 513–525.

[153] Sattler, M.; Schleucher, J.; Griesinger, C.; *Prog. Nucl. Magn. Reson. Spectrosc.* **1999,** *34,* 93–158.

[154] John, M.; Heller, M.; Coles, M.; Bosch, G.; Baumeister, W.; Kessler, H.; *J. Biomol. NMR* **2004,** *29,* 209–210.

[155] Wishart, D. S.; Watson, M. S.; Boyko, R. F.; Sykes, B. D.; *J. Biomol. NMR* **1997,** *10,* 329–336.

[156] Wishart, D. S.; Sykes, B. D.; Richards, F. M.; *Biochemistry* **1992,** *31,* 1647–1651.

[157] Wishart, D. S.; Sykes, B. D.; *J. Biomol. NMR* **1994,** *4,* 171–180.

[158] Bode, W.; Engh, R.; Musil, D.; Thiele, U.; Huber, R.; Karshikov, A.; Brzin, J.; Kos, J.; Turk, V.; *EMBO J.* **1988,** *7,* 2593–2599.

[159] Dieckmann, T.; Mitschang, L.; Hofman, M.; Kos, J.; Turk, V.; Auerswald, E. A.; Jaenicke, R.; Oschkinat, H.; *J. Mol. Biol.* **1993,** *234,* 1048–1059.

[160] Engh, R. A.; Dieckmann, T.; Bode, W.; Auerswald, E. A.; Turk, V.; Huber, R.; Oschkinat, H.; *J. Mol. Biol.* **1993,** *234,* 1060–1069.

[161] Archibald, J. M.; Logsdon, J. M.; Doolittle, W. F.; *Curr. Biol.* **1999,** *9,* 1053–1056.

[162] Archibald, J. M.; Logsdon, J. M.; Doolittle, W. F.; *Mol. Biol. Evol.* **2000,** *17,* 1456–1466.

[163] Braig, K.; Adams, P. D.; Brünger, A. T.; *Nat. Struct. Biol.* **1995,** *2,* 1083–1094.

[164] Boisvert, D. C.; Wang, J. M.; Otwinowski, Z.; Horwich, A. L.; Sigler, P. B.; *Nat. Struct. Biol.* **1996,** *3,* 170–177.

[165] Dyson, H. J.; Wright, P. E.; *Curr. Opin. Struct. Biol.* **2002,** *12,* 54–60.

[166] Dunker, A. K.; Brown, C. J.; Lawson, J. D.; Iakoucheva, L. M.; Obradovic, Z.; *Biochemistry* **2002,** *41,* 6573–6582.

[167] Rose, G. D.; Gierasch, L. M.; Smith, J. A.; *Adv. Protein Chem.* **1985,** *37,* 1–109.

[168] Miklos, D.; Caplan, S.; Mertens, D.; Hynes, G.; Pitluk, Z.; Kashi, Y.; Harrison-Lavoie, K.; Stevenson, S.; Brown, C.; Barrell, B.; *Proc. Natl. Acad. Sci. U. S. A.* **1994,** *91,* 2743–2747.

[169] Kim, S.; Willison, K. R.; Horwich, A. L.; *Trends Biochem. Sci.* **1994,** *19,* 543–548.

[170] Lin, P.; Cardillo, T. S.; Richard, L. M.; Segel, G. B.; Sherman, F.; *Genetics* **1997,** *147,* 1609–1633.

[171] Perdersen, P. L.; Carafoli, E.; *Trends Biochem. Sci.* **1987,** *12,* 146–150.

[172] Moller, J. V.; Juul, B.; LeMaire, M.; *Biochim. Biophys. Acta – Rev. Biomembr.* **1996,** *1286,* 1–51.

[173] Altendorf, K.; Gassel, M.; Puppe, W.; Mollenkamp, T.; Zeeck, A.; Boddien, C.; Fendler, K.; Bamberg, E.; Drose, S.; *Acta Physiol. Scand. Suppl.* **1998,** *643,* 137–146.

[174] Haupt, M.; Bramkamp, M.; Coles, M.; Altendorf, K.; Kessler, H.; *J. Mol. Biol.,* in press.

[175] Hilge, M.; Siegal, G.; Vuister, G. W.; Güntert, P.; Gloor, S. M.; Abrahams, J. P.; *Nat. Struct. Biol.* **2003,** *10,* 468–474.

[176] Yang, D. W.; Kay, L. E.; *J. Mol. Biol.* **1996,** *263,* 369–382.

[177] DeLano, W. L.; "The PyMOL Molecular Graphics System", DeLano Scientific LLC, San Carlos, CA, USA,.

[178] Evenäs, J.; Malmendal, A.; Akke, M.; *Structure* **2001,** *9,* 185–195.

[179] Ishima, R.; Freedberg, D. I.; Wang, Y.-X.; Louis, J. M.; Torchia, D. A.; *Structure* **1999,** *7,* 1047–1055.

[180] Hill, R. B.; Bracken, C.; DeGrado, W. F.; Palmer, A. G.; *J. Am. Chem. Soc.* **2000,** *122,* 11610–11619.

[181] Skrynnikov, N. R.; Dahlquist, F. W.; Kay, L. E.; *J. Am. Chem. Soc.* **2002,** *124,* 12352–12360.

[182] Dutta, R.; Inouye, M.; *Trends Biochem. Sci.* **2000,** *25,* 24–28.

[183] Cooper, A.; Dryden, D. T. F.; *Eur. Biophys. J. Biophys. Letters* **1984,** *11,* 103–109.

[184] Fuentes, E. J.; Der, C. J.; Lee, A. L.; *J. Mol. Biol.* **2004,** *335,* 1105–1115.

[185] Akke, M.; Brüschweiler, R.; Palmer, A. G.; *J. Am. Chem. Soc.* **1993,** *115,* 9832–9833.

[186] Alexandrescu, A. T.; Rathgeb-Szabo, K.; Rumpel, K.; Jahnke, W.; Schulthess, T.; Kammerer, R. A.; *Protein Sci.* **1998,** *7,* 389–402.

[187] Bremi, T.; Brüschweiler, R.; *J. Am. Chem. Soc.* **1997,** *119,* 6672–6673.

[188] LeMaster, D. M.; *J. Am. Chem. Soc.* **1999,** *121,* 1726–1742.

[189] Li, Z.; Raychaudhuri, S.; Wand, A. J.; *Protein Sci.* **1996,** *5,* 2647–2650.

[190] Yang, D. W.; Mok, Y. K.; Forman-Kay, J. D.; Farrow, N. A.; Kay, L. E.; *J. Mol. Biol.* **1997,** *272,* 790–804.

[191] Boehr, D. D.; Farley, A. R.; Wright, G. D.; Cox, J. R.; *Chem. Biol.* **2002,** *9,* 1209–1217.

[192] Gallivan, J. P.; Dougherty, D. A.; *J. Am. Chem. Soc.* **2000,** *122,* 870–874.

[193] Kessler, H.; *Angew. Chem. Int. Ed. Engl.* **1982,** *21,* 512–523.

[194] Stradley, S. J.; Rizo, J.; Bruch, M. D.; Stroup, A. N.; Gierasch, L. M.; *Biopolymers* **1990,** *29,* 263–287.

[195] Koppitz, M.; Huenges, M.; Gratias, R.; Kessler, H.; Goodman, S. L.; Jonczyk, A.; *Helv. Chim. Acta* **1997,** *80,* 1280–1300.

[196] Aumailley, M.; Gurrath, M.; Müller, G.; Calvete, J.; Timpl, R.; Kessler, H.; *FEBS Lett.* **1991,** *291,* 50–54.

[197] Müller, G.; Gurrath, M.; Kessler, H.; Timpl, R.; *Angew. Chem. Int. Ed. Engl.* **1992,** *31,* 326–328.

[198] Gurrath, M.; Müller, G.; Kessler, H.; Aumailley, M.; Timpl, R.; *Eur. J. Biochem.* **1992,** *210,* 911–921.

[199] Mierke, D. F.; Kurz, M.; Kessler, H.; *J. Am. Chem. Soc.* **1994,** *116,* 1042–1049.

[200] Mierke, D. F.; Kessler, H.; *Biopolymers* **1993,** *33,* 1003–1017.

[201] Nikiforovitch, G. V.; Kövér, K. E.; Zhang, W.-J.; Marshall, G. R.; *J. Am. Chem. Soc.* **2000,** *122,* 3262–3273.

[202] Blake, P. R.; Lee, B.; Summers, M. F.; Adams, M. W.; Park, J. B.; Zhou, Z. H.; Bax, A.; *J. Biomol. NMR* **1992,** *2,* 527–533.

[203] Blake, P. R.; Park, J.-B.; Adams, M. W.; Summers, M. F.; *J. Am. Chem. Soc.* **1992,** *114,* 4931–4933.

[204] Dingley, A. J.; Grzesiek, S.; *J. Am. Chem. Soc.* **1998,** *120,* 8293–8297.

[205] Pervushin, K.; Ono, A.; Fernandez, C.; Szyperski, T.; Kainosho, M.; Wüthrich, K.; *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95,* 14147–14151.

[206] Cordier, F.; Grzesiek, S.; *J. Am. Chem. Soc.* **1999,** *121,* 1601–1602.

[207] Cornilescu, G.; Hu, J.-S.; Bax, A.; *J. Am. Chem. Soc.* **1999,** *121,* 2949–2950.

[208] Arnold, W. D.; Oldfield, E.; *J. Am. Chem. Soc.* **2000,** *122,* 12835–12841.

[209] Bagno, A.; *Angew. Chem. Int. Ed. Engl.* **2000,** *6,* 2925–2930.

[210] Bene, J. E. D.; Perera, S. A.; Barlett, R. J.; *J. Am. Chem. Soc.* **2000,** *122,* 3560–3561.

[211] Cornilescu, G.; Ramirez, B. E.; Frank, M. K.; Clore, G. M.; Gronenborn, A. M.; Bax, A.; *J. Am. Chem. Soc.* **1999,** *121,* 6275–6279.

[212] Cordier, F.; Grzesiek, S.; *J. Mol. Biol.* **2002,** *317,* 739–752.

[213] Cordier, F.; Barfield, M.; Grzesiek, S.; *J. Am. Chem. Soc.* **2003,** *125,* 15750–15751.

[214] Zhao, X.; Sudmeier, J. L.; Bachovchin, W. W.; Levitt, M. H.; *J. Am. Chem. Soc.* **2001,** *123,* 11097–11098.

[215] Kay, L. E.; Ikura, M.; Tschudin, R.; Bax, A.; *J. Magn. Reson.* **1990,** *89,* 496–514.

[216] Grzesiek, S.; Bax, A.; *J. Magn. Reson.* **1992,** *96,* 432–440.

[217] Juranić, N.; Ilich, P. K.; Macura, S.; *J. Am. Chem. Soc.* **1995,** *117,* 405–410.

[218] Shaka, A. J.; Lee, C. J.; Pines, A.; *J. Magn. Reson.* **1988,** *77,* 274–293.

[219] Rucker, A. P.; Shaka, A. J.; *Mol. Phys.* **1989,** *68,* 509–517.

[220] Bax, A.; Max, D.; Zax, D.; *J. Am. Chem. Soc.* **1991,** *114,* 6923–6925.

[221] Bax, A.; Vuister, G. W.; Grzesiek, S.; Delaglio, F.; Whang, A. C.; Tschudin, R.; Zhu, G.; *Methods Enzymol.* **1994,** *239,* 79–105.

[222] Gullion, T.; Schaefer, J.; *J. Magn. Reson.* **1989,** *81,* 196–200.

[223] Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshimi, K. V.; Griffin, R. G.; *J. Chem. Phys.* **1995,** *103,* 6951–6958.

[224] Hohwy, M.; Jaroniec, C. P.; Reif, B.; Rienstra, C. M.; Griffin, R. G.; *J. Am. Chem. Soc.* **2000,** *122,* 3218–3219.

[225] Tycko, R.; Weliky, D. P.; Berger, A. E.; *J. Chem. Phys.* **1996,** *18,* 7915–7930.

[226] Pang, Y.; Zuiderweg, E. R. P.; *J. Am. Chem. Soc.* **2000,** *122,* 4841–4842.

[227] Cornilescu, G.; Bax, A.; *J. Am. Chem. Soc.* **2000,** *122,* 10143–10145.

[228] Hore, P. J.; *Meth. Enzymol.* **1989,** *176,* 64–77.

[229] Guéron, M.; Plateau, P.; Decorps, M.; *Prog. Nucl. Magn. Reson. Spectrosc.* **1991,** *23,* 135–209.

[230] Bloembergen, N. J.; Pound, R. V.; *Phys. Rev* **1954,** *95,* 8–12.

[231] Bloom, S.; *J. Appl. Phys.* **1957,** *28,* 800–805.

[232] Mao, X.-A.; Ye, C.-H.; *Concepts Magn. Reson.* **1997,** *9,* 173–187.

[233] Szántay, C.; Demeter, Á.; *Concepts Magn. Reson.* **1999,** *11,* 121–145.

[234] Mao, X.-A.; Guo, J.-X.; Ye, C.-H.; *Phys. Rev. B* **1994,** *49,* 15702–15711.

[235] Santoro, J.; King, G. C.; *J. Magn. Reson.* **1992,** *97,* 202–207.

[236] Vuister, G. W.; Bax, A.; *J. Magn. Reson.* **1992,** *98,* 428–435.

[237] Cavanagh, J.; Palmer, A. G.; Wright, P. E.; Rance, M.; *J. Magn. Reson.* **1991,** *91,* 429–436.

[238] Palmer, A. G.; Cavanagh, J.; Wright, P. E.; Rance, M.; *J. Magn. Reson.* **1991,** *93,* 429–436.

[239] Kay, L. E.; Keifer, P.; Saarinen, T.; *J. Am. Chem. Soc.* **1992,** *114,* 10663–10665.

[240] Palmer, A. G.; Cavanagh, J.; Byrd, R. A.; Rance, M.; *J. Magn. Reson.* **1992,** *96,* 416–424.

[241] Cavanagh, J.; Rance, M.; *Annu. Rep. NMR Spectrosc.* **1993,** *27,* 1–58.

[242] Powers, R.; Gronenborn, A. M.; Clore, G. M.; Bax, A.; *J. Magn. Reson.* **1991,** *94,* 209–213.

[243] Palmer, A. G.; Fairbrother, W. J.; Cavanagh, J.; Wright, P. E.; Rance, M.; *J. Biomol. NMR* **1992,** *2,* 103–108.

[244] Zhang, W.; Gmeiner, W. H.; *J. Biomol. NMR* **1996,** *7,* 247–250.

[245] Marion, D.; Driscoll, P. C.; Kay, L. E.; Wingfield, P. T.; Bax, A.; *Biochemistry* **1989,** *28,* 6150–6156.

[246] Marion, D.; Kay, L. E.; Sparks, S. W.; Torchia, D. A.; Bax, A.; *J. Am. Chem. Soc.* **1989,** *111,* 1515–1517.

[247] Zuiderweg, E. R. P.; Fesik, S. W.; *Biochemistry* **1989,** *28,* 2387–2391.

[248] Jahnke, W.; Baur, M.; Gemmecker, G.; Kessler, H.; *J. Magn. Reson. B* **1995,** *106,* 86–86.

[249] Frenkiel, T.; Bauer, C.; Carr, M. D.; Birdsall, B.; Feeney, J.; *J. Magn. Reson.* **1990,** *90,* 420–425.

[250] Heller, M.; John, M.; Coles, M.; Bosch, G.; Baumeister, W.; Kessler, H.; *J. Mol. Biol.* **2004,** *336,* 717–729.

[251] Carpino, L. A.; Han, G. Y.; *J. Org. Chem.* **1972,** *37,* 3404–3409.

[252] Fields, G. B.; Noble, R. L.; *Int. J. Peptide Protein Res.* **1990,** *35,* 161–214.

[253] Kurz, M.; *Cyclische Modellpeptide als Template für ein konformationell orientiertes Peptiddesign,* Thesis, Technische Universität München, 1991.