Department Chemie

der Technischen Universität München

Structure and Exchange Dynamics of the Apical Domains of the Thermosome and Human Saposins by NMR Spectroscopy

Michael John

Vollständiger Abdruck der von der Fakultät 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:	Univ.-Prof. Dr. H. Kessler
	Univ.-Prof. Dr. W. Domcke
	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.

dedicated to my parents

In theory, there is no difference
between theory and practice.
In practice, however, there is.
- *Unknown*

Abstract

NMR spectroscopy provides a versatile and indispensable methodology for the characterisation of molecular structure and dynamics in modern biochemical research. A part of the present work focuses on fast structural predictions, derived from major NMR parameters and from their similarity between sequence-related proteins. In two isolated apical domains of a large archaeal chaperonin, residues whose conformation was influenced by crystal packing interactions were shown to be highly disordered in solution. For a small class of homologous proteins, the human saposins, residual dipolar couplings were used to validate their common fold. Many proteins are subject to chemical and conformational exchange processes, which are often essential for their function, but may render the molecule unamenable for structural investigations, as in the case of the GM2 activator protein. Several NMR-spectroscopic methods were employed to study hydrogen- and slow conformational exchange kinetics as well as its dependence on pH and temperature in saposin D.

Zusammenfassung

Die NMR-Spektroskopie stellt der modernen biochemischen Forschung ein vielseitiges methodisches Arsenal zur Charakterisierung von molekularer Struktur und Dynamik zur Verfügung. Ein Teil der vorliegenden Arbeit befasst sich mit schnellen strukturellen Vorhersagen aus wichtigen NMR-Parametern sowie deren Ähnlichkeit in sequenziell verwandten Proteinen. Anhand von zwei isolierten apikalen Domänen eines grossen archäalen Chaperonins konnte gezeigt werden, dass einige Aminosäurereste, im Gegensatz zum Kristall, in Lösung eine ungeordnete Konformation einnehmen. In einer kleinen Klasse von homologen Proteinen, den humanen Saposinen, wurden dipolare Restkopplungen zur Validierung der einheitlichen Faltung verwendet. Viele Proteine unterliegen chemischen oder konformativen und für ihre Funktion essentiellen Austauschprozessen, durch die sie jedoch für eine strukturelle Charakterisierung unzugänglich werden, wie zum Beispiel der GM2-Aktivator. Verschiedene NMR-spektroskopische Methoden wurden eingesetzt, um Amid- und Konformationsaustausch sowie dessen pH- und Temperaturabhängigkeit in Saposin D zu untersuchen.

Parts of this thesis have already been published:

Expression of the GM2-activator protein in the methylotrophic yeast *Pichia pastoris*, purification, isotopic labelling and biophysical characterisation. M. Wendeler, J. Hoernschemeyer, M. John, N. Werth, M. Schoeniger, T. Lemm, R. Hartmann, H. Kessler and K. Sandhoff (2004) *Protein Expr. Purif.* **34**, 147-157**.**

Backbone ^{1}H, ^{15}N and ^{13}C resonance assignments of α-ADT and β-ADT. M. John, M. Heller, M. Coles, G. Bosch, W. Baumeister and H. Kessler (2004) *J. Biomol. NMR* **29**, 209-210.

NMR studies on the substrate-binding domains of the thermosome: structural plasticity in the protrusion region. M. Heller, M. John, M. Coles, G. Bosch, W. Baumeister and H. Kessler (2004) *J. Mol. Biol.* **336,** 717-729.

Acknowledgement

The work presented in this thesis was done under the guidance of Prof. Dr. Horst Kessler at the Institute for Organic Chemistry and Biochemistry of the Technical University of Munich, Germany, in the period between July 2000 and June 2004. I would like to thank my supervisor Prof. Dr. Horst Kessler for excellent research facilities, his confidence and the freedom to pursue the projects with own ideas, helpful discussions and unrestricted support in all concerns.

My further thank goes to:

● the former and present staff of the "NCE" Markus Heller, Melina Haupt, the "kangaroo" Dr. Murray Coles, Dr. Tammo Diercks, Dr. "Vinocento" Vincent Truffault, "Sarav" Saravanakumar Narayanan, Zhongjing Chen, Dr. Louis "XIV" Patard and Jochen Klages for the great atmosphere, but also for steady scientific exchange.

● the "old guys" Dr. "Gustav" Gerd Gemmecker, Dr. Julien Furrer, Dr. "Bulu" Burkard Luy, Dr. "Tosh" Thomas Schulte-Herbrüggen and Dr. Bernd "Greif" Reif for their patience being confronted with my demand for discussions.

● the "knechts" Andreas Brandl, Sebastian Ahrendt, Stefanie Gasteyer, Tim Bartels, Florian Leiss and Tobias Rossteuscher for their fresh enthusiasm.

● again Markus for carefully reading the manuscript (while preparing his own thesis !) and helpful suggestions

● the team for spectrometers and computers Dr. Rainer Haessner, Monika Goede and Alex Frenzel for continuously setting up the latest versions (before one is used to the old ones).

● the "synthies" Dr. Martin Sukopp, Armin Modlinger, Dominik Heckmann, Axel Meyer and Timo Weide for making their facilities and knowledge available and the proof that an interaction between chemists and spectroscopists is indeed possible.

● Mandar Deshmukh and Mrunal Pendke for an unforgettable Indian wedding, loads of Indian food and help in preparing Xmas cookies.

● the little mermaid Cindie Kehlet for some sunny coffee breaks.

● the secretaries Beate Diaw, Marianne Machule and Evelyn Bruckmaier for their professional work.

● the sometimes obstinate fellows Laurel, Hardy, Pat and Chef for their company and entertainment during lonely evenings.

● all further group members for a wonderful time in the lab, on conferences, in the mountains and at the "Wies´n".

● Dr. Gundula Bosch and Prof. Dr. Wolfgang Baumeister from the MPI for Biochemistry in Martinsried, Germany for calling the legend „ADT" into being, and shipping of samples on the spot.

● Michaela Wendeler and Prof. Dr. Konrad Sandhoff from the University of Bonn for the delivery of tons of activators and saposins.

● Prof. Dr. Ruth Gschwind, University of Bonn, for lightening the fire of NMR spectroscopy inside me.

● the SOMA for efficient relaxation at low frequencies.

● my parents for their unlimited support.

Contents

Abbreviations

αADT	*α-chain ADT*
ADC	*analogue-digital converter*
ADT	*apical domain of the thermosome*
AK	*adenosine kinase*
AM	*amplitude modulation*
APS	*ammonium peroxodisulfate*
ATP	*adenosine triphosphate*
βADT	*β-chain ADT*
BMRB	*Biological magnetic resonance bank*
BPP	*bipolar pulse pair*
BPTI	*bovine pancreatic trypsin inhibitor*
CCT	*chaperonin containing TCP*
cDNA	*complementary deoxyribonucleic acid*
CITY	*computer-optimised homonuclear tocsy*
CMC	*critical micelle concentration*
COSY	*correlation spectroscopy*
CPMG	*Carr-Purcell-Meiboom-Gill (spin-lock sequence)*
CSA	*chemical shift anisotropy*
CSI	*chemical shift index*
CSM	*chemical shift modulation*
CT	*constant time*
CTS	*coherence transfer step*
CTAB	*cetyltrimethylammonium bromide*
CW	*continuous wave*
D	*dimensional*
DD	*dipole-dipole*
DHPC	*1,2-dihexanoyl-sn-glycero-3-phosphocholine*
DIPSI	*decoupling in the presence of scalar couplings*
DMPC	*1,2-dimyristoyl-sn-glycero-3-phosphocholine*
DMSO	*dimethylsulfoxide*
DQ	*double quantum*
DSS	*2,2-dimethyl-2-silapentane-5-sulfonic acid*

DSTE	*double stimulated echo*
ESI-Q-TOF	*electrospray ionisation quadrupole time-of-flight*
EXSY	*exchange spectroscopy*
FHSQC	*fast HSQC*
FID	*free induction decay*
FT	*fourier transform*
FWHH	*full width at half height (=linewidth)*
GalNAc	*N-acetyl galactosamine*
GARP	*globally optimised alternating phase rectangular pulses*
GATE	*gradient tailored (suppression)*
GlcNAc	*N-acetyl glucosamine*
GM2	*mono sialic acid ganglioside no. 2*
GM2AP	*GM2 activator protein*
GroEL	
GROMOS	*Groningen molecular simulation package*
GSL	*glycosphingolipid*
HD	*hydrodynamic*
hetNOE	*heteronuclear NOE*
HMQC	*heteronuclear multiple quantum correlation*
HSQC	*heteronuclear single quantum correlation*
HX	*hydrogen exchange*
INEPT	*insensitive nuclei enhancing polarisation transfer*
IPAP	*in-phase and anti-phase*
LED	*longitudinal eddy current delay*
LP	*linear prediction*
Man	*mannose*
MD	*molecular dynamics*
MEXICO	*measurement of exchange rates in isotopically labelled compounds*
NeuNAc	*N-acetyl neuraminic acid (=sialic acid)*
Ni-NTA	*nickel nitrilotriacetic acid resin*
NMR	*nuclear magnetic resonance* or *no meaningful results*
NOE(SY)	*nuclear overhauser enhancement (spectroscopy)*
PAS	*principal axis system*
PDB	*brookhaven protein data bank*

PDLA	*poly-D,L-alanine*
PF	*protection factor*
PFG	*pulsed field gradient*
PISA	*polar index slant angle*
PM	*phase modulation*
PTH	*peptidyl tRNA hydrolase*
RDC	*residual dipolar coupling*
RF	*radio frequency*
RMSD	*root mean squared deviation*
ROE(SY)	*rotating frame NOE(SY)*
RT	*real time*
SAG	*strained polyacrylamide gel*
SAP	*sphingolipid activator protein*
SCT	*semi-constant time*
SDS	*sodium dodecyl sulphate*
SE	*spin echo*
SOMA	*sinfonic orchestra munich-andechs*
SQ	*single quantum*
STE	*stimulated echo*
TCP	*T-complex polypeptide*
TMEDA	*tetramethylethylendiamine*
TMSCl	*trimethylsilyl chloride*
TOCSY	*total correlation spectroscopy*
TPPI	*time proportional phase incrementation*
TRiC	*TCP ring complex*
tRNA	*transfer ribonucleic acid*
TROSY	*transverse relaxation optimised spectroscopy*
TSP	*3-(trimethylsilyl)-propionic acid*
TXI	*triple resonance inverse detection*
WEX	*water exchange*
wt	*wild type*
ZQ	*zero quantum*

Introduction and scope of the work

In the last few decades, biological sciences have been revolutionised by a growing understanding of physiological processes on a molecular level, hallmarked by the decipherment of the genetic code and the discovery of the flow of information from genes to proteins. Constructed from a few relatively simple building blocks, proteins have been found to form a tremendous plethora of structures and functions, ranging from huge complexes like the thermosome, which aids in the folding of other proteins, to small enzyme cofactors like the saposins. Appreciating the role of biomolecules in biochemical pathways and the ability to encounter diseases resulting from potential functional defects requires knowledge about their three dimensional structure as well as the exact mechanism of action.

Two major experimental techniques are currently able to accomplish this task down to atomic resolution, and by now, almost 26,000 protein, nucleic acid and carbohydrate structures have been solved by X-ray crystallography and NMR spectroscopy and deposited in the Brookhaven Protein Data Bank. On the other hand, recent progress has been achieved in supplementary “low resolution” methods as cryo-electron microscopy or approaches that extract information from data base searches or theoretical simulations. Today there is a growing awareness that biological sciences will benefit most from a close and fruitful cooperation, in which each method focuses on its individual superiorities.

Since 1945, when the first weak radio frequency responses from atomic nuclei in paraffin and water were observed, NMR has become a powerful tool for a diversity of applications, ranging from medicine, organic and biochemistry to material sciences and quantum computing. This development has been accompanied by technical advances in magnetic field strength and homogeneity, probe sensitivity and data processing, but even more by methodological milestones such as Fourier transform spectroscopy, multidimensional and heteronuclear spectroscopy, relaxation optimised sequences, residual dipolar couplings, pulsed field gradients and spatial imaging.

In biomolecular NMR, uniformly or selectively isotopically labelled recombinant macromolecules of up to 80 kDa are currently studied. However, whereas information about secondary structure and the fold may be collected rather rapidly, the resonance assignment and accumulation of distance, dihedral angle and orientational restraints required for solving complete structures often remains a tedious procedure. Currently, huge efforts are made to automate single steps of the structure determination process. In return, NMR is highly flexible when studying molecular interactions, and spin relaxation provides a powerful tool for

detecting motional and exchange processes on various timescales and in solution, where molecules are able to reveal their full dynamic properties. In particular, conformational exchange processes are a rapidly growing subject of study, and many of them have meanwhile been attributed to biological function.

On the other hand, like the lack of stability, exchange processes may heavily deteriorate spectra and make a structural characterisation of even small proteins by NMR virtually impossible. Since basic knowledge of the effects of exchange to NMR spectra is absolutely essential for the discussion of the underlying processes, they will be discussed in **chapter 1**. Hydrogen exchange of labile protons, albeit having less biological significance, is conceptually so closely related to conformational exchange and encountered in several projects in this thesis that it will be focussed on in the second half of this chapter.

Two dimensional heteronuclear correlation-, especially ^{15}N-HSQC spectroscopy, is the workhorse for assessing a new protein´s stability and amenability for further studies, for estimating structural similarities, and for detecting and mapping weak and strong ligand binding. Affording high resolution within relatively short experimental time, it is the most widely used detection scheme for measurements on a per residue basis, and occurs as building block in a variety of double- and triple resonance experiments. A few basic principles along with important parameters will be discussed in **chapter 2**, using the experience and examples from own experimental work.

Chapter 3 illustrates, how NMR of two small homologous 17 kDa domains can contribute to the study of the archaeal 930 kDa thermosome from *Thermoplasma acidophilum*, without having to solve structures completely *de novo*. X-ray structures were available of the complete complex as well as both isolated apical domains (ADTs), but there was reasonable suspicion that the most important feature, a 20 Å long helical protrusion was strongly influenced by crystal packing interactions. This project was accomplished in collaboration with Markus Heller, who focussed on ^{15}N-relaxation and the dynamical properties of the protrusions.

Likewise, the human GM2 activator protein, which is an essential cofactor for the lysosomal degradation of GM2 to GM3 and the subject of **chapter 4**, had been crystallised before, but under neutral pH conditions, where it shows no biological activity. At acidic lysosomal, i.e. active pH, however, the protein suffered from poor NMR performance and low stability and did not allow more than an initial characterisation in terms of aggregation state, secondary structure content and binding of GM1.

Functionally closely related to, but genetically and structurally distinct from the GM2 activator, the saposins are rather small, but yet poorly understood homologous 11 kDa glycoproteins. Along with the ADTs, the four representatives of this class are prime examples to show how sequence- and structural similarity is reflected in the appearance of NMR spectra.

Unfortunately, the pool of interesting proteins with good NMR performance is increasingly grazed or liable to scientific competition. For example, promising attempts to solve the structure of only ^{15}N-labelled saposin C at pH 7 were foiled by a successful structural analysis of the doubly labelled protein in another laboratory. Exchange processes or aggregation, leading to much lower spectral quality, render the other saposins or saposin C at pH 4 less suitable for a complete structure determination. In **chapter 5**, only their initial characterisation and, if possible, ^{15}N-HSQC assignment is presented.

One of the most significant recents impacts on NMR spectroscopy was the development of weakly aligning media in combination with methods for efficient extraction and interpretation of the dipolar coupling information. Not all media are equally qualified for a given biomolecule, and after a rather theoretical introduction into the subject, experimental work on different alignment media for the saposins is described in **chapter 6**. Finally, the data obtained for three saposins is compared, and structural conclusions are drawn.

Chapter 7 is exclusively dedicated to the slow conformational exchange process in saposin D, which could be studied over a wide temperature range due to the enormous thermal stability of the protein. Kinetic data was obtained using exchange of longitudinal and transverse magnetisation, and discussed in terms of the involved enthalpies and entropies.

Since no extensive synthetic or biochemical laboratory work, which is not described elsewhere, was done in this thesis, it is abstained from a separate chapter on material and methods. Tables of unpublished NMR data as well as novel implemented pulse programs are given in **chapter 9**.

1 Exchange processes in NMR

1.1 Motion and exchange

One of the most powerful features of NMR spectroscopy is its sensitivity to motional processes at timescales, which are most relevant in biology and chemistry (**Fig. 1.1.1**). Generally, motion renders the nuclear spin hamiltonian *time dependent* and thus affects the spectrum and relaxation properties of spins.[1]

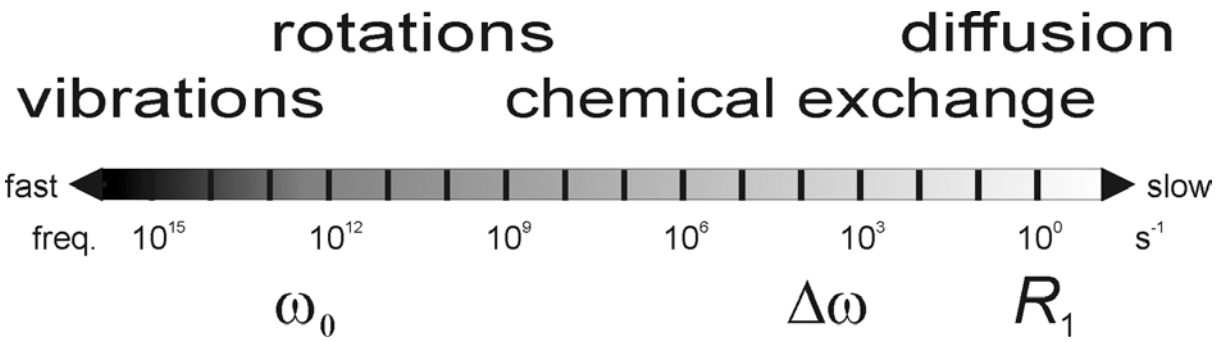


Fig. 1.1.1: Frequency scale of motional processes and their positions relative to the Larmor frequency ω_0, the amplitude of chemical shift modulation (CSM) $\Delta\omega$ (both in rad s^{-1}) and spin-lattice relaxation rate R_1.

On the fast side, the very end of the observable scale is marked by motions that are faster than the Larmor frequency ω_0, because in this case the dynamic hamiltonian can no longer *a priori* be distinguished from a static one, expressed in terms of averaged distances and angles. Using the static, *snapshot* structural picture provided by ultrafast methods like X-ray and neutron diffraction requires the scaling of spin interactions with an *order parameter S*, which accommodates vibrations and fast rotations (e.g. of methyl groups).

Motions in the nano-picosecond regime, comprising global molecular rotations, interdomain motion, flexibility of side chains and loops, and helix-coil transitions, average *secular spin interactions* (i.e. those which affect the NMR spectra), but not the non-secular ones. For example, all dipolar contributions to the spin hamiltonian average to zero in isotropic solution, their nanosecond random fluctuation does, however, provide the spectral density required for spin-lattice- and cross relaxation. Today a large theoretical body is established to extract amplitudes (i.e. order parameters) and timescales of fast motions from R_1 and R_2 relaxation rates and heteronuclear NOEs.[2-4] In this thesis, these will play only a minor role.

Exchange processes, which are even slower in the micro- to millisecond timescale, start to affect the lineshapes, because the time dependence of the spin hamiltonian (i.e. *chemical shift modulation*, *CSM*) becomes a source of transverse relaxation.[5] Exchange is classified as *fast*, *intermediate* or *slow*, if its associated rate constant k_{ex} is larger, in the range of, or smaller than the spectral timescale, which is defined by the amplitude $\Delta\omega$ of the hamiltonian fluctuations, expressed in units of frequency. Classically, *chemical exchange* requires breaking and formation of chemical bonds, but today the term is usually applied to all kinds of conformational transitions and hindered rotations like amide bond isomerisations and slow aromatic ring flips.[6] After introducing a few basic definitions, which will be referred to in later chapters, lineshapes in exchanging systems, i.e. exchange of transverse magnetisation under *free precession* will be discussed in section 1.2. The formalism will be extended to the evolution of transverse magnetisation under *CPMG* conditions in section 1.3.

Finally, in the case of very slow exchange, associated with high energy barriers or extremely low temperatures, simple NMR spectra can no longer be distinguished from mixtures of non-exchanging components. Nevertheless, if k_{ex} does not greatly fall below the spin-lattice relaxation rate R_1 of involved nuclei, the process may still be detected by exchange of *longitudinal* magnetisation, which will be the subject of section 1.4. Very recently, progress has been achieved in extending the memory time of nuclear spins beyond T_1 ($=1/R_1$),[7] opening the possibility to study even slower processes.

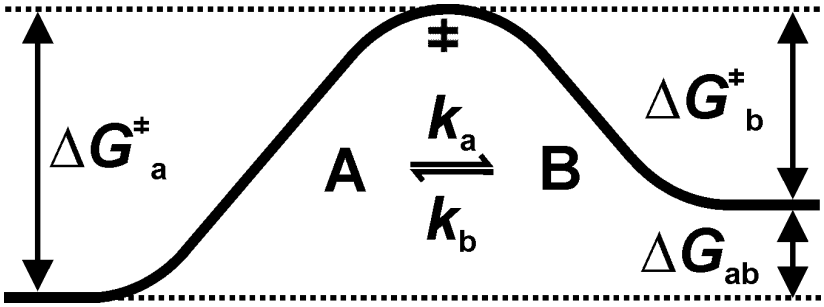


Fig. 1.1.2: Reaction profile of a two site (A, B) exchange equilibrium with transition state (≠), forward and backward rate constants k_a and k_b and differences in Gibbs free enthalpy ΔG.

The great advantage and power of NMR spectroscopy is, that all motional processes mentioned so far can be studied under equilibrium conditions and do not require complicated devices for mixing and time resolved measuring. Of course, NMR is also capable of dealing with non-equilibrium systems with virtually no lower limit of k_{ex}, a prominent example being H→D exchange of amide protons in D_2O. Hydrogen exchange (HX) with protic solvents is an example of a two site exchange with highly skewed populations, which will be discussed in detail in sections 1.5 to 1.9.

Consider an equilibrium between two species A and B with associated forward and backward first order exchange rate constants k_a and k_b (**Fig. 1.1.2**), which is more conveniently characterised by the overall exchange rate constant $k_{ex} = k_a + k_b$ and the fractional populations $p_a = k_b/k_{ex}$ and $p_b = k_a/k_{ex}$ ($p_a + p_b = 1$). In highly skewed populations ($p_a \gg p_b$), as in the case of hydrogen exchange between labile protons in a macromolecule and the bulk water, $k_{ex} \approx k_b$ is essentially a pseudo-zero order off-rate from that particular site. The equilibrium constant $K = k_a/k_b = p_b/p_a$ is associated with the difference in free enthalpy ΔG_{ab} of B and A:

$$K = \exp\left(\frac{-\Delta G_{ab}}{RT}\right) = \exp\left(\frac{-\Delta H_{ab}}{RT} + \frac{\Delta S_{ab}}{R}\right) \tag{1.1.1}$$

Similarly, the temperature dependence of k_a (and k_b) may be calculated from *Arrhenius* law in combination with the theory of the activated complex:[8]

$$k_a = \frac{kT}{h}\exp\left(\frac{-\Delta G^{\neq}_a}{RT}\right) = \frac{kT}{h}\exp\left(\frac{-\Delta H^{\neq}_a}{RT} + \frac{\Delta S^{\neq}_a}{R}\right) \tag{1.1.2}$$

$\Delta G^{\neq}_a$ is the Gibbs free enthalpy of the transition state, which contains an enthalpic term $\Delta H^{\neq}_a$ as well as an entropic contribution $-T\Delta S^{\neq}_a$, if the transition state requires a significant degree of order or tolerates disorder in the system. Let a certain spin have a precession frequency of ω_a in molecules belonging to state A and ω_b for state B, then their difference is given by $\Delta\omega_{ab} = \omega_b - \omega_a$ and the exchange process may be classified as:

$$\begin{aligned} &\textit{slow} && k_{ex} \ll \Delta\omega_{ab} \\ &\textit{intermediate} && k_{ex} \approx \Delta\omega_{ab} \\ &\text{and } \textit{fast} && k_{ex} \gg \Delta\omega_{ab} \end{aligned} \tag{1.1.3}$$

on the spectral timescale of that particular spin. Note that a single exchange process in a molecule may for different spins give rise to different regimes and lineshapes, which depend on the static magnetic field. The term *slow* should not be confused with a complicated process of conversion from A to B involving, for example, several transition states. In fact, the transition can be regarded as instantaneous event, and *slow* refers simply to a low probability of such a jump in either direction to occur within unit time.

1.2 Exchange of transverse magnetisation (free precession)

In order to understand the effect of k_{ex} on the lineshape, it is instructive to consider a *symmetrical* two site exchange process with equal populations $p_a = p_b = 0.5$ and rate constants $k_a = k_b = k_{ex}/2$. Starting from state A, transverse magnetisation evolves, i.e. oscillates in the rotating frame with $\Omega_a = \omega_a - \omega_0$, where ω_0 is the spectrometer reference (= transmitter) frequency, until a transition occurs to state B, and the signal continues oscillating with $\Omega_b = \omega_b - \omega_0$, etc. (**Fig. 1.2.1**).

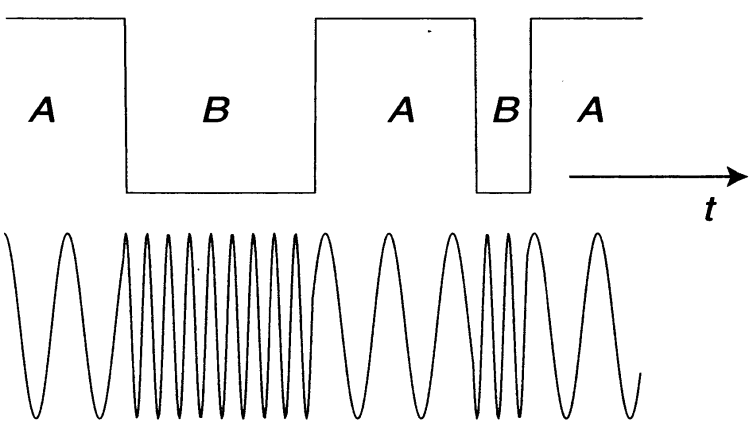


Fig. 1.2.1: Evolution of transverse magnetisation in a single molecule under the influence of symmetrical two-site exchange. Since the lifetimes of states A and B between the transitions are long enough to allow the magnetisation to evolve with their distinct frequencies, this example corresponds to a *slow* exchange process. Figure taken from [1].

Of course, the exact time point of the transitions in an individual molecule is unpredictable and subject to statistical variations among the molecules. The gradual loss of synchronisation of transverse magnetisation in an ensemble of state A molecules may be visualised by summing up simulated random trajectories for only a few molecules (**Fig. 1.2.2**). For low values of k_{ex}, the dephasing is monoexponential with time constant $k_{ex}/2$ (k_a in an asymmetric equilibrium), in the absence of further transverse relaxation resulting in a Lorentzian frequency domain signal at Ω_a with linewidth k_{ex} (in rad s^{-1}). Likewise, all molecules starting from state B give rise to a Lorentzian signal at Ω_b.

The peak positions start to move towards each other, if k_{ex} enters the intermediate exchange regime, where signals often are already so broad, that they are no longer visible in the spectrum. After the *crossover* point $k_{ex} = \Delta\omega_{ab}$ (**Fig. 1.2.2 C**), the transitions become so frequent, that spins do no longer accumulate larger phase differences, and the averaging of Ω_a and Ω_b becomes more and more complete. This results in a *motionally narrowed* single peak at the average position $\Omega_{av} = 1/2\ (\Omega_a \pm \Omega_b)$ (in the asymmetric case weighted with the populations p_a and p_b).

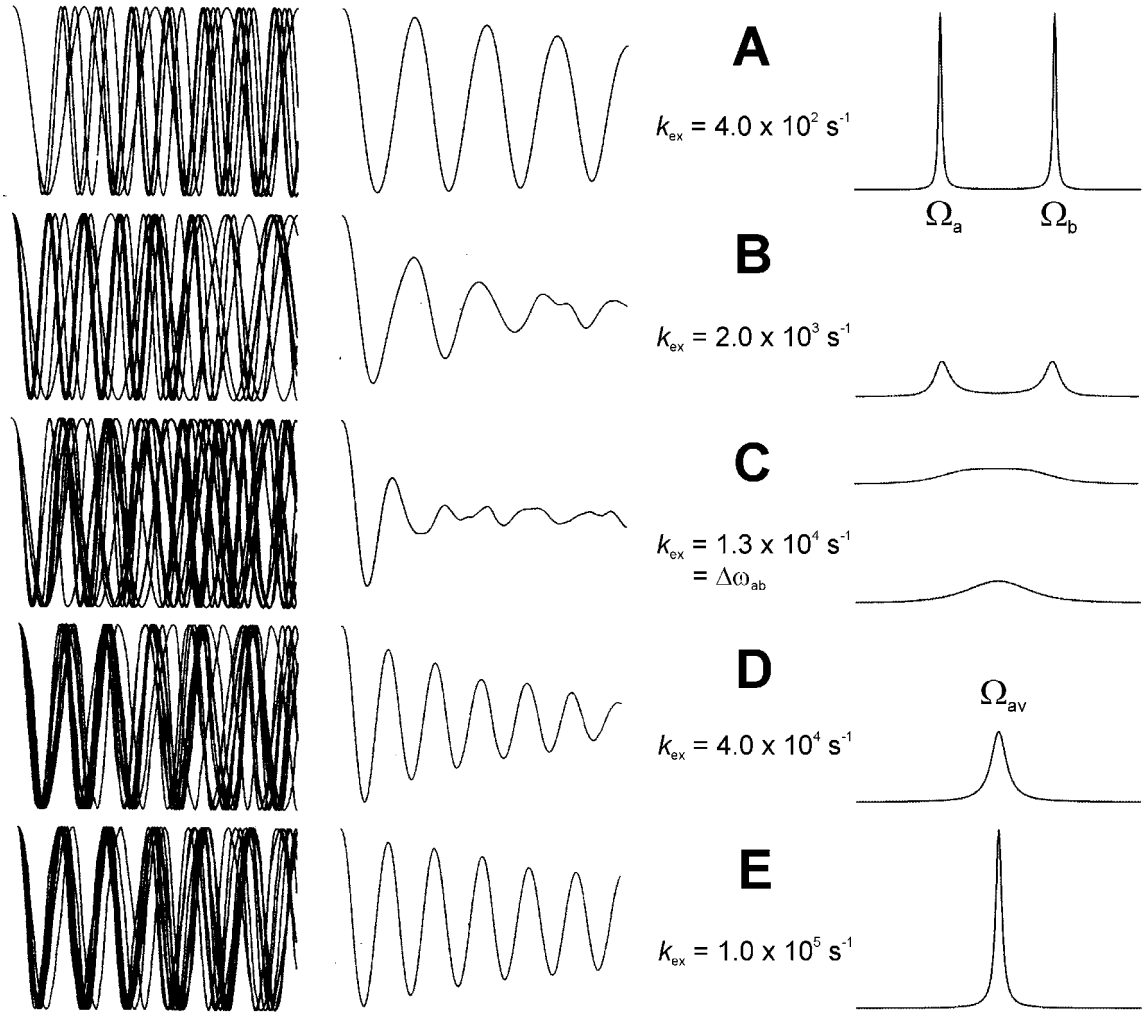


Fig. 1.2.2: Simulations of precessing transverse magnetisation in an ensemble of 20 molecules, all starting in state A, with fixed $\Delta\omega_{ab}$ ($\Omega_a/2\pi$ = 2 kHz, $\Omega_b/2\pi$ = 4 kHz) and variable k_{ex}. **Left column:** superposition of individual trajectories spanning 1 ms. **Middle column:** summed transverse magnetisation. **Right column:** spectra resulting after Fourier transformation with equal amount of molecules starting from A and B. The intermediate exchange crossover (C) is defined as $k_{ex} = \Delta\omega_{ab}$. Coalescence (maximal broad signal) occurs before at $k_{ex} = \Delta\omega_{ab}/\sqrt{2}$. Figure partially taken from [1].

Assuming that the spin topology in the exchanging species is the same, the overall density operator $\boldsymbol{\rho}$ of the system is given by:

$$\boldsymbol{\rho} = p_a\boldsymbol{\rho}_a + p_b\boldsymbol{\rho}_b \qquad \begin{pmatrix} \rho^{\alpha} & \rho^{+} \\ \rho^{-} & \rho^{\beta} \end{pmatrix} = p_a \begin{pmatrix} \rho^{\alpha}{}_a & \rho^{+}{}_a \\ \rho^{-}{}_a & \rho^{\beta}{}_a \end{pmatrix} + p_b \begin{pmatrix} \rho^{\alpha}{}_b & \rho^{+}{}_b \\ \rho^{-}{}_b & \rho^{\beta}{}_b \end{pmatrix} \tag{1.2.1}$$

where $\boldsymbol{\rho}_a$ and $\boldsymbol{\rho}_b$ are averaged only over the molecules, which are in state A and B at a particular point of time, respectively. Each matrix element ρ^{α}, ρ^{β}, ρ^{-} and ρ^{+} of the matrix representation of $\boldsymbol{\rho}$ in the eigenbase of $\mathbf{I}_z$ of a single I = ½ spin may be decomposed in

analogy, and all eight elements representing the individual states evolve according to inherent dynamics (in the rotating frame) as well as to the exchange process. For example, (−1) quantum coherence in molecules A obeys the equation of motion:

$$\frac{d}{dt}\rho^{-}_{a} = (i\Omega_a - R_{2a})\rho^{-}_{a} - k_a\rho^{-}_{a} + k_b\rho^{-}_{b} \tag{1.2.2}$$

Together with an analogous equation for ρ^{-}_{b} this may be combined to:

$$\frac{d}{dt}\begin{pmatrix}\rho^{-}_{a}\\ \rho^{-}_{b}\end{pmatrix} = \begin{pmatrix} i\Omega_a - R_{2a} - k_a & k_b \\ k_a & i\Omega_b - R_{2b} - k_b\end{pmatrix}\begin{pmatrix}\rho^{-}_{a}\\ \rho^{-}_{b}\end{pmatrix} = \mathbf{L}\boldsymbol{\rho}^{-} \tag{1.2.3}$$

The time evolution of $\boldsymbol{\rho}^{-}$ is governed by the propagator $\mathbf{Q} = \exp(\mathbf{L}t)$, which is calculated by a unitary transformation $\mathbf{D} = \mathbf{V}^{-1}\mathbf{L}\mathbf{V}$ and subsequent integration, as described below in section 1.4. For a symmetrical equilibrium with identical transverse relaxation rates $R_{2a} = R_{2b} = R_2$ one obtains the following signals after Fourier transformation:

$$S(\Omega) = \frac{1}{2}\left(1 - \frac{\mathrm{i}k_{ex}}{2P}\right)\mathcal{L}\left(\Omega_{av} + P, R_2 + \frac{k_{ex}}{2}\right) + \frac{1}{2}\left(1 + \frac{\mathrm{i}k_{ex}}{2P}\right)\mathcal{L}\left(\Omega_{av} - P, R_2 + \frac{k_{ex}}{2}\right) \quad k_{ex}<\Delta\omega_{ab} \tag{1.2.4}$$

$$S(\Omega) = \frac{1}{2}\left(1 + \frac{k_{ex}}{2P}\right)\mathcal{L}\left(\Omega_{av}, R_2 + \frac{k_{ex}}{2} - P\right) + \frac{1}{2}\left(1 - \frac{k_{ex}}{2P}\right)\mathcal{L}\left(\Omega_{av}, R_2 + \frac{k_{ex}}{2} + P\right) \quad k_{ex}>\Delta\omega_{ab} \tag{1.2.5}$$

with $P = \sqrt{\left|k_{ex}^{\,2} - \Delta\omega_{ab}^{\,2}\right|}/2$ and $\Omega_{av} = (\Omega_a + \Omega_b)/2$

Here, $\mathcal{L}(\Omega, R_2)$ denotes a Lorentzian which is centred on Ω and has a linewidth of $2R_2$, as directly evident from equation (1.2.4) in the absence of exchange ($k_{ex} = 0$), where the overall signal consists of 2 lines at $\Omega_{av} + \Delta\omega_{ab}/2$ (= Ω_b) and $\Omega_{av} - \Delta\omega_{ab}/2$ (= Ω_a). With increasing k_{ex}, P decreases, and the lines, being phase twisted by the imaginary term and broadened by k_{ex}, move towards each other. At the crossover $k_{ex} = \Delta\omega_{ab}$, $P = 0$, and both Lorentzians become identical, and remain centred on Ω_{av} for all $k_{ex} > \Delta\omega_{ab}$. However, because P is now dominated by $k_{ex}/2$, the first term of equation (1.2.5) represents a line with increasing intensity and decreasing linewidth, whereas for the second term it is just the other way round, and it may finally be neglected. Coalescence occurs at $k_{ex} = \Delta\omega_{ab}/\sqrt{2}$ (**Fig. 1.2.2**), when both (twisted) Lorentzians are still separated by $2P = \Delta\omega_{ab}/\sqrt{2}$, but no longer resolved (for absorptive Lorentzians this occurs at $\Delta\omega_{ab}$ = half linewidth). In systems with large intrinsic R_2 like biomolecules, coalescence is shifted to even lower k_{ex}.

1.3 Exchange of transverse magnetisation (CPMG conditions)

As described in the previous section, CSM with rate constants in the order of $\Delta\omega_{ab}$ results in increased free precession linewidths, i.e. an exchange contribution R_{ex} superimposed onto the intrinsic transverse relaxation rate R_2 of the affected nuclei:

$$R_{ex} = k_a = p_b k_{ex} \qquad \text{for } k_{ex} \ll \Delta\omega_{ab} \tag{1.3.1}$$

$$R_{ex} = \frac{4 p_a p_b \Delta\omega_{ab}^{\ 2}}{k_{ex}} \qquad \text{for } k_{ex} \gg \Delta\omega_{ab} \tag{1.3.2}$$

In order to remove artefacts from *J*-couplings and an imperfect static magnetic field homogeneity, R_2 is usually measured from the exponential signal decay after a CPMG train of variable length. In the presence of exchange, however, the apparent (or effective) transverse relaxation rate R_2^{eff} is no longer $R_2 + R_{ex}$ as for free precession, but becomes a function of the 180° puls spacing $2\tau_{CP}$ owing to partial refocussing of $\Delta\omega_{ab}$. This has been exploited to identify exchange from the *dispersion* of $R_2^{eff}(\nu_{CP})$ with variable CPMG field $\nu_{CP} = 1/(4\tau_{CP})$ and constant relaxation delay,[9] and full analytical expressions for the profile $R_2^{eff}(\nu_{CP})$ as function of $\Delta\omega_{ab}$, k_{ex} and p_a under the assumption of two site exchange have been derived.[10, 11] Since CSM relaxation can be fully suppressed in the fast pulsing limit ($\nu_{CP} \gg \Delta\omega_{ab}/2\pi$), R_{ex} is usually defined as the difference $R_2^{eff}(\nu_{CP}\to 0) - R_2^{eff}(\nu_{CP}\to\infty)$. Under slow exchange conditions, $R_2^{eff}(\nu_{CP})$ shows oscillations in the region of slow pulsing (low ν_{CP})(**Fig. 1.3.1**).

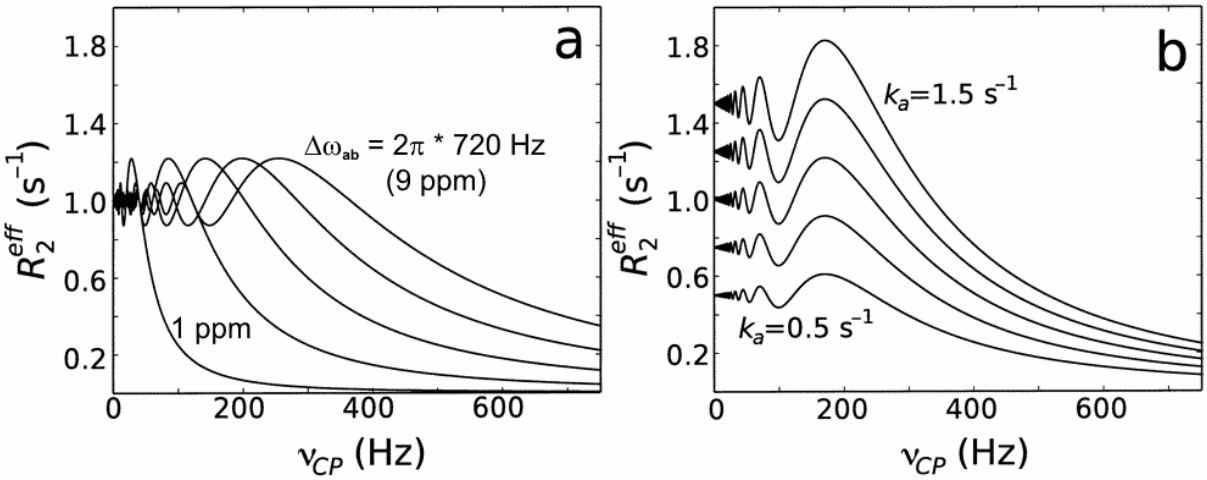


Fig. 1.3.1: Simulated CPMG dispersion profiles for component A in a two-site slow exchange showing only the exchange contribution (i.e. $R_{2a} = R_{2b} = 0$) **a**: Effect of $\Delta\omega_{ab}$, corresponding to a ^{15}N chemical shift difference at B_0 = 18.8 T (from the left to the right) of 1, 3, 5, 7 and 9 ppm or increasing B_0, with constant forward rate $k_a = R_{ex}$ = 1 s^{-1}. **b**: Effect of the forward rate constant k_a (= R_{ex}), increasing from the bottom to the top: 0.5, 0.75, 1.0, 1.25 and 1.5 s^{-1}, with constant $\Delta\omega_{ab} = 2\pi$ * 480 Hz (6 ppm). Figure taken from [12].

For site A, these oscillations have been shown to depend only on the forward rate k_a and the product of $\Delta\omega_{ab}$ and the CPMG delay τ_{CP}:[12]

$$R_2^{eff} = R_{2a} + k_a - k_a \frac{\sin(\Delta\omega_{ab}\tau_{CP})}{\Delta\omega_{ab}\tau_{CP}} \tag{1.3.3}$$

which drops to the intrinsic rate R_{2a} in the fast pulsing limit. At $\nu_{CP} = 0$ (free precession), the actual transfer of transverse magnetisation between A and B within one precession cycle $1/\Delta\omega_{ab}$ averages to zero, and the effect of exchange is simply an increased dephasing on A. In contrast, refocusing pulses may lead to a net transfer from B to A, which can increase or decrease the amount of A transverse magnetisation, resulting in a lower or higher effective transverse relaxation rate (**Fig. 1.3.2**).

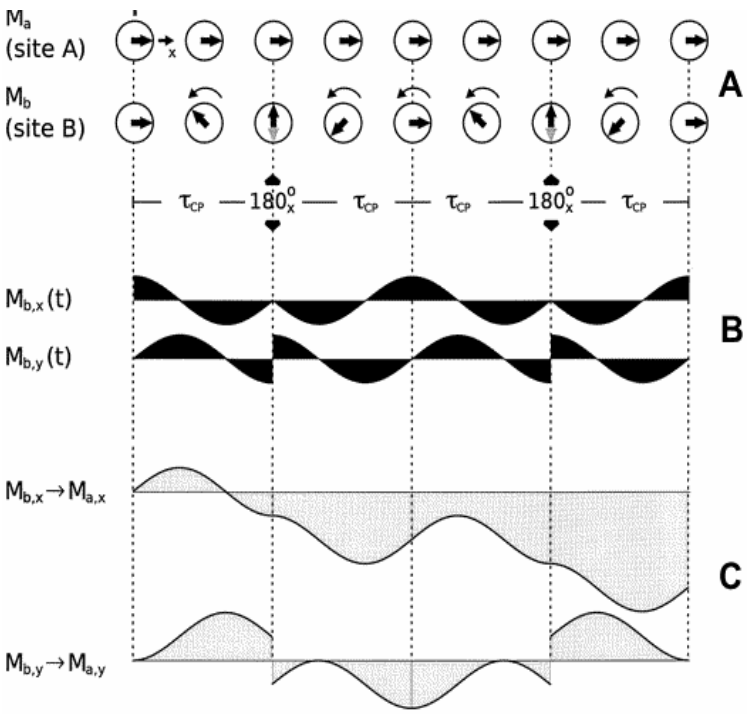


Fig. 1.3.2: Scheme illustrating the transfer of transverse magnetisation from B to A through two CPMG cycles τ_{CP}-180°$_x$-τ_{CP}. Spin A is assumed to be on-resonant (Ω_a = 0), whereas B evolves with $\Omega_b = \Delta\omega_{ab}$. **A**: rotating magnetisation vectors $\mathbf{M}_a$ and $\mathbf{M}_b$, starting from x-magnetisation. **B**: trajectory of x- and y- components of $\mathbf{M}_b$, $M_{b,x}(t) = \langle \mathbf{I}_x \rangle_b = \mathrm{tr}(\boldsymbol{\rho}_b \mathbf{I}_x)$ and $M_{b,y}(t) = \langle \mathbf{I}_y \rangle_b = \mathrm{tr}(\boldsymbol{\rho}_b \mathbf{I}_y)$ for $\Delta\omega_{ab}\tau_{CP} = 3\pi/2$ ($\nu_{CP} = \Delta\omega_{ab}/(3*2\pi)$). **C**: net amount of magnetisation accumulated on A due to transfer from B (= integral of trajecories). Fig. taken from [12].

Hence, the peculiar situation arises, that R_2^{eff} can be higher at nonzero CPMG field strength than in the free precession limit. For the first (going from right to left in Fig. 1.3.1) and highest maximum of R_2^{eff} in the dispersion curve at $\Delta\omega_{ab}\tau_{CP} = 3\pi/2$, the contribution of magnetisation transfer makes up $2/(3\pi)k_a$ ($\approx 0.21\ k_a$). In contrast, at integral multiples of $\Delta\omega_{ab}\tau_{CP}/\pi$, the sinc term in equation 1.3.2 vanishes and the system behaves as under free precession. At low B_0, or if the chemical shift difference between A and B is small, the oscillations may not be observed, since R_2^{eff} cannot be sampled at arbitrarily small ν_{CP} due to sensitivity limitations.

1.4 Exchange of longitudinal magnetisation

If k_{ex} is very slow, lineshapes are merely affected, since the contribution k_a (k_b) to the transverse relaxation rates is much smaller compared to intrinsic transverse relaxation R_{2a} (R_{2b}) or contributions from unresolved *J*-couplings. Such dynamic processes may still be studied by NMR using exchange of longitudinal magnetisation $\langle \mathbf{I}_z \rangle = p_a \langle \mathbf{I}_z \rangle_a + p_b \langle \mathbf{I}_z \rangle_b = p_a(\rho^{\alpha}_a - \rho^{\beta}_a) + p_b(\rho^{\alpha}_b - \rho^{\beta}_b)$, which is part of the overall density operator as defined in equation 1.2.1. There is no oscillation in the evolution of longitudinal magnetisation, giving rise to a relatively simple equation of motion (neglecting that $\langle \mathbf{I}_z \rangle$ actually drives back to a nonzero thermal equilibrium, e.g. by an appropriate phase cycle):

$$\frac{d}{dt}\begin{pmatrix} \langle \mathbf{I}_z \rangle_a \\ \langle \mathbf{I}_z \rangle_b \end{pmatrix} = \underbrace{\begin{pmatrix} -R_{1a} - k_a & k_b \\ k_a & -R_{1b} - k_b \end{pmatrix}}_{\mathbf{L}} \begin{pmatrix} \langle \mathbf{I}_z \rangle_a \\ \langle \mathbf{I}_z \rangle_b \end{pmatrix} \tag{1.4.1}$$

This system of coupled differential equations is solved by diagonalising **L** according to $\mathbf{D} = \mathbf{V}^{-1}\mathbf{L}\mathbf{V}$, where the unitary matrix **V** contains the eigenvectors, and the diagonal matrix **D** the eigenvalues of **L**:

$$\mathbf{V} = \frac{1}{\sqrt{k_{ex}}}\begin{pmatrix} k_b & 1 \\ k_a & 1 \end{pmatrix} \qquad \mathbf{D} = \mathbf{V}^{-1}\mathbf{L}\mathbf{V} = \begin{pmatrix} -R_1 & 0 \\ 0 & -R_1 + k_{ex} \end{pmatrix} \tag{1.4.2}$$

assuming equal intrinsic longitudinal relaxation rates $R_{1a} = R_{1b} = R_1$. Subsequent exponentiation of $\mathbf{L}t$ is now straightforward to yield the propagator **Q**:

$$\mathbf{Q} = \exp(\mathbf{L}t) = \mathbf{V}(\mathbf{V}^{-1}\exp\mathbf{L}t\mathbf{V})\mathbf{V}^{-1} = \mathbf{V}\exp(\mathbf{V}^{-1}\mathbf{L}t\mathbf{V})\mathbf{V}^{-1} = \mathbf{V}\exp(\mathbf{D}t)\mathbf{V}^{-1} \tag{1.4.3}$$

The matrix elements of **Q** indicate, which fraction of the initial polarisations, e.g. $\langle I_z \rangle_a(0)$, are retained on A (q_{aa}) and transferred to B (q_{ab}) during the time period *t*, respectively, and vice versa (equation 1.4.4):

$$\begin{pmatrix} \langle \mathbf{I}_z \rangle_a(t) \\ \langle \mathbf{I}_z \rangle_b(t) \end{pmatrix} = \underbrace{\begin{pmatrix} q_{aa} & q_{ba} \\ q_{ab} & q_{bb} \end{pmatrix}}_{\mathbf{Q}} \begin{pmatrix} \langle \mathbf{I}_z \rangle_a(0) \\ \langle \mathbf{I}_z \rangle_b(0) \end{pmatrix} = \frac{1}{k_{ex}} \begin{pmatrix} k_b e^{-R_1 t} + k_a e^{-(R_1 + k_{ex})t} & k_b e^{-R_1 t} - k_b e^{-(R_1 + k_{ex})t} \\ k_a e^{-R_1 t} - k_a e^{-(R_1 + k_{ex})t} & k_a e^{-R_1 t} + k_b e^{-(R_1 + k_{ex})t} \end{pmatrix} \begin{pmatrix} \langle \mathbf{I}_z \rangle_a(t) \\ \langle \mathbf{I}_z \rangle_b(t) \end{pmatrix}$$

They are experimentally accessible, if magnetisation is frequency labelled before and after the exchange takes place during a mixing period, as in 2D exchange spectroscopy (EXSY),[13] where q_{aa} and q_{bb} produce diagonal or *auto* signals, whereas q_{ab} and q_{ba} give rise to *cross* signals correlating the exchanging sites. Starting from full thermal polarisations, and assuming equal treatment during preparation, terms originating from $\langle \mathbf{I}_z \rangle_a(0)$ (q_{aa} and q_{ab}) and

$\langle \mathbf{I}_z \rangle_b(0)$ (q_{bb} and q_{ba}) may be scaled with p_a and p_b, respectively, as shown in the simulations (**Fig. 1.4.1**). Under these conditions, i.e. if z-magnetisation represents true concentrations, $p_a q_{ab} = p_b q_{ba}$, since in the equilibrium the number of exchanging molecules in the forward and backward reaction is identical.

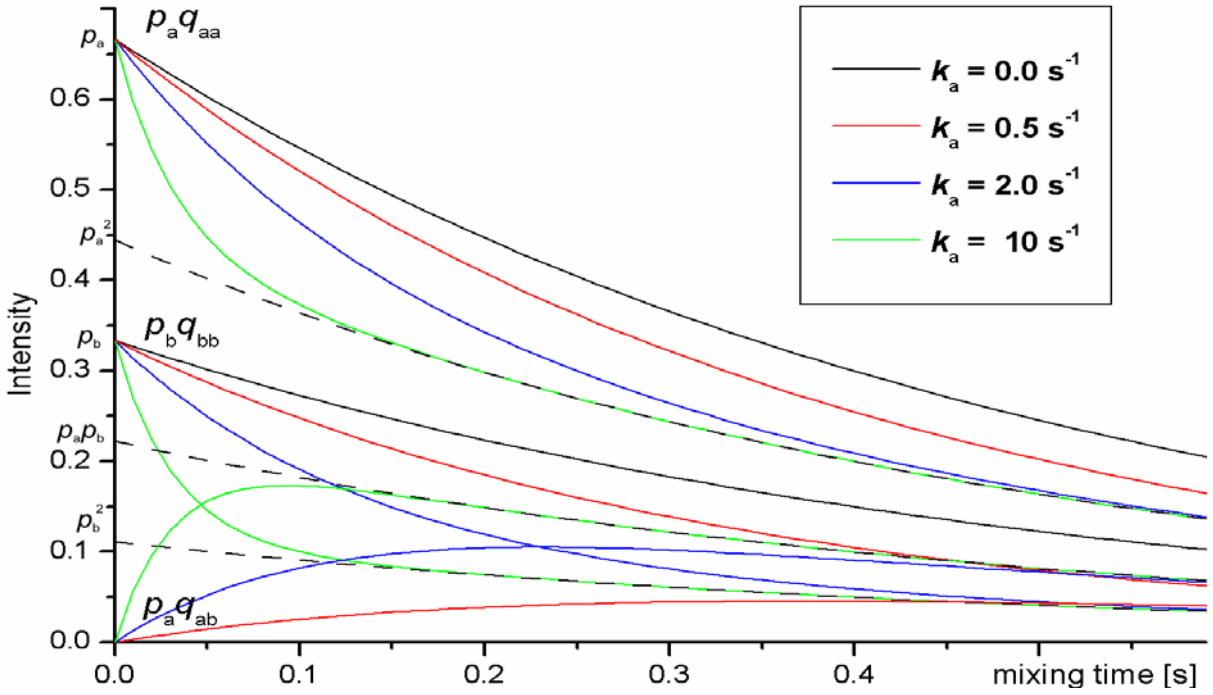


Fig. 1.4.1: Simulation of the time evolution of matrix **Q** elements according to equation 1.4.4 with scaling as mentioned in the text. For all curves, $R_1 = 2\ s^{-1}$ and $k_b = 2\ k_a$ ($p_a = 2/3$, $p_b = 1/3$). The black dashed lines are the $k_{ex} \gg R_1$ asymptotic plateaus for q_{aa}, q_{ab} and q_{bb} decaying with $\exp(-R_1 t)$.

In the absence of exchange, both exponentials cancel q_{ab} and q_{ba} to zero, and q_{aa} and q_{bb} follow a monoexponential decay with the time constant R_1. Therefore, exchange is principally not detectable for $k_{ex} \ll R_1$, because the build-up of cross peaks requires long mixing periods, where the magnetisation has already decayed owing to spin lattice relaxation. In the other extreme ($k_{ex} \gg R_1$), the second exponential term, which is the same for all four peaks (apart from opposite sign for the diagonal and cross signals), governs the initial behaviour of the curves: $p_a q_{aa}$ and $p_b q_{bb}$ quickly drop to plateaus at p_a^2 and p_b^2, the drop being much more pronounced for the minor species of an asymmetric equilibrium. In contrast, $p_a q_{ab}$ and $p_b q_{ba}$ rise to a plateau at $p_a p_b$, which is remarkably situated between the other two, and all plateaus slowly relax according to $\exp(R_1 t)$. Cross-relaxation in a 2-spin system is completely analogous, if in this case $k_a = k_b$ is replaced by cross relaxation-, and $R_1 + k_a$ by the auto (= leakage) relaxation rates, respectively.

1.5 Hydrogen exchange chemistry

Discussing hydrogen exchange (HX) in macromolecules in terms of structural biology requires the knowledge of the underlying chemical events, upon which structural and experimental influences are superimposed.[14, 15] Consider a process, in which a catalyst molecule and the exchanging site (either of the two is donor *AH* and acceptor *B*), collide with a rate constant k_D, form a rapidly equilibrating (K_C) H-bridged complex, which dissociates again with the rate constant k_{-D} (equation 1.5.1). Owing to the large excess of water molecules in aqueous solution, final de- or reprotonation by H_2O, thus re-establishing the educt molecules, is always fast.

$$AH + B \underset{k_{-D}}{\overset{k_D}{\rightleftharpoons}} (AH\cdots B \overset{K_C}{\rightleftharpoons} A\cdots HB) \underset{k_D}{\overset{k_{-D}}{\rightleftharpoons}} A + HB \qquad (1.5.1)$$

In free solution, the collision rate k_D is diffusion-limited and depends on the catalyst concentration [Cat], whereas the equilibrium within the complex K_C determines the fraction of successful collisions F_C to yield the overall rate k_{tr} for proton transfer:

$$k_{tr} = k_D F_C \qquad k_D = 10^{10}\ [\mathrm{Cat}]\ \mathrm{M^{-1}s^{-1}} \qquad F_C = \frac{K_C}{1+K_C} \qquad K_C = 10^{(pK_B - pK_A)} \qquad (1.5.2)$$

When transfer of a proton is energetically downhill, from a stronger to a weaker acid ($pK_B > pK_A$), F_C approaches unity (**Fig. 1.5.1**), and the exchange proceeds (in the absence of other slowing factors) at the maximum possible rate k_D.

A acid	base	pK
-CO-NH-	-CO-N^--	18.5
H_2O	OH^-	15.7
lysine ζ-NH_3^+	lysine ζ-NH_2	10.0
imidazole-H^+	imidazole	7.0
H_3O^+	H_2O	-1.7
-C(OH)=NH^+-	-CO-NH-	-12.5

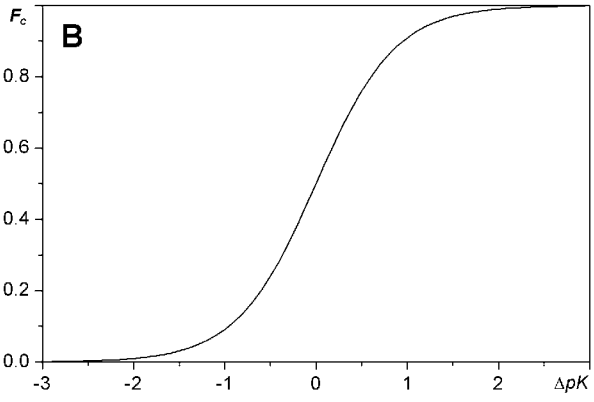


Fig. 1.5.1: A: pK of selected acid-base pairs. At pH = pK the concentrations of acid and base are equal. **B:** F_C as a funtion of $\Delta pK = pK_B - pK_A$ (for a proton transfer from A to B).

For example, in a 0.2 M aqueous solution of imidazole at pH 7.0, corresponding to a 0.1 M base catalyst, exchange of lysine ζ-NH_3^+ protons proceeds at $k_{tr} = 10^6$ s^{-1} ($F_C = 10^{-3}$). Although less efficient, this is much faster than OH^--ion catalysis ($k_{tr} = 10^3$ s^{-1}, $F_C \approx 1$) for these conditions, simply due to the low concentration of hydroxide ions. In contrast, deprotonation of a peptide amide group to the *imidate anion* is energetically strongly unfavourable and therefore faster by OH^--ion catalysis ($k_{tr} = 10^{0.2}$ s^{-1}) compared to catalysis by imidazole ($k_{tr} = 10^{-2.5}$ s^{-1}). Generally, exchange of amide protons by buffer or salt ion catalysis in protein solutions is negligible, and the *intrinsic* HX-rate k_{int} can be expressed as sum of H_3O^+ and OH^- catalysis only:

$$k_{int} = k_{tr,H} + k_{tr,OH} = k_H [H_3O^+] + k_{OH} [OH^-] \qquad (1.5.3)$$

k_H (= 0.41 $M^{-1}s^{-1}$ for poly-D,L-alanine at 20 °C and low salt conditions) is much smaller than k_{OH} (= 1.49*10^8 $M^{-1}s^{-1}$), because protonation of an amide group by H_3O^+ is energetically uphill by 11 decades! The minimum of k_{int} is located at a pH, where acid- and base-catalysis are equally fast, and from there k_{int} increases 10-fold per pH unit in both directions, giving rise to typical V-shaped profiles (**Fig. 1.5.2**).

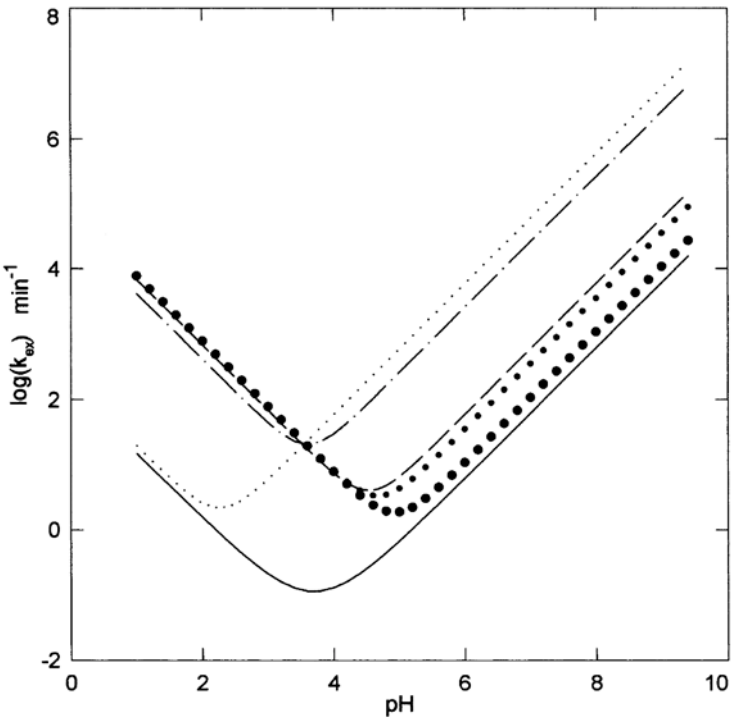


Fig. 1.5.2: HX-profile of log(k_{int}) in min^{-1} versus pH at 25 °C for nitrogen-bound protons. poly-D,L-alanine NH (solid line), glutamine ε-NH_2 (bold and medium dots for Z and E protons), tryptophan ε-NH (dashed line), arginine ε-NH (small dots) and arginine η-NH_2 (dash-dots). Figure taken from [15].

At conditions of biological relevance, acid catalysis as well as a pH-independent contribution of catalysis by neutral water molecules is negligible.

1.6 Hydrogen exchange in peptides

In peptides, the intrinsic rate constants k_H and k_{OH} of acid and base catalysed amide hydrogen exchange are sensitive to the neighbouring side chains, i.e. they depend on the primary sequence.[16] Surprisingly, side chains of residues (i−1) and (i) have been found to exert simply additive effects on $\log_{10}k_H$ and $\log_{10}k_{OH}$, which allows to calculate intrinsic exchange rates for arbitrary dipeptide pairs (i−1,i) by reference to a suitable alanine peptide like poly-D,L-alanine (see above) and correction factors *a* and *b* given by **Table 1.6.1**:[17]

$$\log_{10} k_H\,(i-1,i) = \log_{10} k_H\,(Ala,Ala) + a(i-1) + a(i) \qquad (1.6.1)$$

$$\log_{10} k_{OH}\,(i-1,i) = \log_{10} k_{OH}(Ala,Ala) + b(i-1) + b(i) \qquad (1.6.2)$$

Table 1.6.1: Dipeptide sequence specific correction terms *a*(i−1), *a*(i) (acid catalysis), *b*(i−1) and *b*(i) (base catalysis) for the calculation of the intrinsic HX rate of residue (i) according to equations 1.5.3, 1.6.1 and 1.6.2. Positive (negative) charges need to be considered for pH <(>) *pK* of the respective acidic (basic) side chain: pK(Arg) ≈ 12, pK(Lys) ≈ 9, pK(His) ≈ 7, pK(Glu) ≈ pK(Asp) ≈ 5. When His is neutral (pH > 7), acid catalysis is too slow compared to base catalysis to be measured.

Xaa	*a*(i-1)	*a*(i)	*b*(i-1)	*b*(i)	Xaa	*a*(i-1)	*a*(i)	*b*(i-1)	*b*(i)
Ala	+0.00	+0.00	+0.00	+0.00	Leu	–0.13	–0.57	–0.21	–0.58
Arg^+	–0.32	–0.59	+0.22	+0.08	Lys^+	–0.29	–0.56	+0.12	–0.04
Asn	–0.13	–0.58	+0.32	+0.49	Met	–0.28	–0.64	+0.11	–0.01
Asp^-	+0.58	+0.90	–0.18	–0.30	Phe	–0.43	–0.52	+0.06	–0.24
Asp	–0.12	–0.90	+0.60	+0.69	Pro (t)	–0.19	-	–0.24	-
Cys	–0.46	–0.54	+0.55	+0.62	Pro (c)	–0.85	-	+0.60	-
Cys_2	–0.58	–0.74	+0.46	+0.55	Ser	–0.39	–0.44	+0.30	+0.37
Gly	+0.22	–0.22	+0.17	+0.27	Thr	–0.47	–0.79	+0.20	–0.07
Gln	–0.27	–0.47	+0.20	+0.06	Trp	–0.44	–0.40	–0.11	–0.41
Glu^-	+0.31	–0.90	–0.15	–0.51	Tyr	–0.37	–0.41	+0.05	–0.27
Glu	–0.27	–0.60	+0.39	+0.24	Val	–0.30	–0.74	–0.14	–0.70
His	-	-	+0.14	–0.10	$Nter^+$	–1.32	-	1.62	-
His^+	–0.51	–0.80	+0.83	+0.80	$Cter^-$	-	+0.96	-	–1.80
Ile	–0.59	–0.91	–0.23	–0.73	Cter	-	+0.05	-	-

Both acid and base catalysis are significantly slower in hydrophobic residues due to steric blocking of the amide proton, which corresponds to a downward shift of the V-shaped exchange profile of **Fig. 1.5.2**. Electron-withdrawing groups in the side chain exert an opposite inductive effect onto the basicity and acidity of the peptide group and therefore shift the minimum of the V to lower pH. Additionally, positively and negatively charges residues or the peptide termini can electrostatically influence local catalyst concentrations, and logarithmic rates of respective residues have been shown to scale linearly with the square root of ionic strength.[18] Efficient base catalysis of His^+, which is approximately 50 % present at pH 7.0, is the reason why amide protons of the unstructured histidine tag are usually only visible under very acidic conditions.

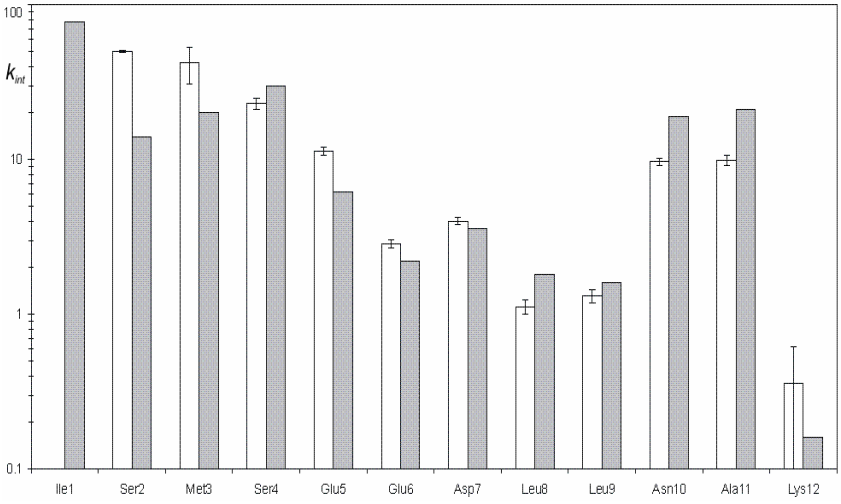


Fig. 1.6.1: Measured (white bars with error bars) and predicted (grey bars) rates of H→H exchange in an unstructured dodecapeptide at 20 °C, 0.5 M KCl and pH 7.0.[19] Ile1 is not visible in spectra due to fast exchange, and for Ser2, a lower limit of 50 s^{-1} was estimated from rates determined at lower temperature.

Furthermore, intrinsic HX rates may be calculated for arbitrary temperatures, if Arrhenius behaviour is assumed for k_H and k_{OH}:

$$k_{H/OH} \propto \exp\left(-\frac{E_a}{RT}\right) \tag{1.6.3}$$

Activation energies E_a have been determined as 14 and 17 kcal/mol for acid and base catalysis, respectively, corresponding to an approximately threefold acceleration per 10 °C increment.[17] In order to compare k_{int} for H→D with H→H exchange, small isotope corrections need to be applied.[20] Thus predicted values of k_{int} have proven a fairly good approximation of experimental hydrogen exchange in unstructured peptides, except for the N- and C- termini (**Fig. 1.6.1**).[19]

1.7 Slowing of intrinsic rates in proteins

In folded proteins, HX-rates can be slowed down by many decades compared to the free peptide k_{int}, and some amide groups have been shown to resist H→D exchange for years.[21] It is now widely accepted, that *intramolecular H-bonding* of an amide proton blocks the formation of the transfer encounter complex, and exchange requires structural fluctuations, which imply breaking of that respective H-bond. According to a model proposed by Linderstrøm-Lang early in 1957,[22] hydrogen exchange of a particular amide in globular proteins is preceded by an equilibrium of closed and open states (**Fig. 1.7.1 A**).

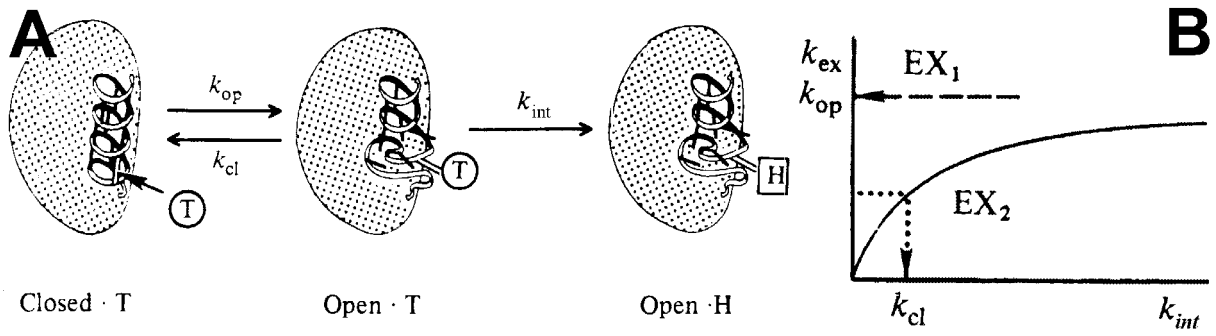


Fig. 1.7.1: A: Schematic illustration of a local unfolding equilibrium Closed-T → Open-T with forward and backward rate constants k_{op} and k_{cl}, which is required for T(ritium)→H(proton) amide exchange (k_{int}) of an α-helical residue. **B:** Plot of overall exchange rate k_{ex} versus k_{int}, as given by equation (1.7.1). Fig. taken from [14].

Since, in a strictly basic or acidic regime, k_{int} is proportional to the catalyst concentration, the combined kinetics is analogous to the common Michaelis-Menton formulation, except second and first order processes occur in the reverse order. For an equilibrium strongly biased towards the closed state, i.e. $k_{op} \ll k_{cl}$ the overall HX-rate k_{ex} simplifies to:

$$k_{ex} = \frac{k_{op} k_{int}}{k_{cl} + k_{int}} \qquad (1.7.1)$$

The relative size of rates k_{cl} and k_{int} determines the regime of hydrogen exchange, which is commonly referred with the apparent order of the overall exchange reaction (**Fig. 1.7.1 B**).[23]

$k_{cl} \ll k_{int}$: $k_{ex} = k_{op}$ EX1 (1.7.2)

$k_{cl} \gg k_{int}$: $k_{ex} = K_{op}\, k_{int}$ $K_{op} = k_{op}/k_{cl}$ EX2 (1.7.3)

In the EX1 regime, the amide proton exchanges every time the H-bond is broken by a local structural fluctuation, before it is re-closed, which is typically after *milliseconds* or shorter.[24]

Since k_{int} is relatively slow under conditions, where proteins retain native conformations (below pH ≈ 8), amide exchange from folded proteins mostly follows pH-dependent EX2 kinetics. K_{op} defines the time fraction, in which the amide proton is not H-bonded, i.e. accessible to chemical exchange, and is often estimated by the inverse of the protection factor *PF*:

$$PF = k_{int}/k_{ex} = 1/K_{op} \tag{1.7.4}$$

PF is defined as the extend, to which the experimentally determined exchange rate k_{ex} is suppressed relative to the exchange rate k_{int} of a non-H-bonded amide in an unstructured peptide having the same dipeptide sequence at the same pH and temperature. Hence, the relation between *PF* and K_{op} assumes, that the open state is equivalent to a random coil state, which may not be the case. For example, for completely unprotected amides at the surface of globular proteins k_{ex} (= k_{int}) may differ from random coil values owing to the given geometric features of the local environment. K_{op} yields quantitative evaluation of an apparent free energy associated with the exchange limiting backbone fluctuation:

$$\Delta G = -RT \ln(K_{op}) \tag{1.7.5}$$

According to thermodynamics, all possible conformations in a protein are populated with their relative energies, and "open" and "closed" states may actually summarise a large amount of HX-competent and HX-protected conformations connected by a variety of modes and timescales. Extremely slow protons ($K_{op} \approx 10^{10}$) in the core β-sheet of BPTI require global unfolding as the energetically highest of the excited states.[25]

In alternative HX-models, k_{ex} has been proposed to depend on the penetration of solvent and catalyst to the exchanging amide, and thus predicted to be reduced especially in the protein core.[26] However, for thermodynamic reasons, amides in the core are always also H-bonded, and *solvent penetration* cannot explain reduced HX-rates in solvent accessible, but H-bonded amides like surface helices.[27]

It has been postulated that acid-catalysed hydrogen exchange should occur via *imidic acid mechanism*, in which the amide oxygen is protonated because of its larger basicity compared to the nitrogen.[28] However, in proteins this mechanism requires fluctuations freeing H-bonds of both the amide carbonyl and proton, and N-protonation may compete with O-protonation for strongly H-bonded carbonyls.[29] For amides in the protein core that are connected to the solvent via a chain of H-bonds, exchange may principally occur by a relayed imidic acid pathway, in which H_3O^+ does not have to reach the buried amide itself.[30]

1.8 Measuring hydrogen exchange

The simplest method to determine HX rates of amide protons is lyophylising re-dissolving the protein in D_2O and observing the decay of individual resonances by a series of 2D NMR experiments like ^{15}N-HSQC or TOCSY. This *exchange out* is, however, limited to protons with $k_{ex} < 10^{-3}$ s^{-1}, which are not yet fully exchanged, before the first experiment is recorded (typically 15-20 minutes). In principle, rates of faster protons may at least be estimated from the signal decay during acquisition of the 2D experiment, which becomes manifest as line broadening in the indirect dimension.[31] If slow protons of interest can be resolved in 1D spectra, e.g. because most others have already disappeared (*exchange editing*), the limit may be extended to $k_{ex} \approx 10^{-1}$ s^{-1} using ultrafast transfer lines.[32] Vice versa, under conditions, where amide exchange is generally very slow, the fastest protons may be resolved by *exchange in* from completely exchange-deuterated protein in H_2O.[33]

D→H exchange trapping has successfully been employed to identify structured regions within folding intermediates, which are too short lived to be studied by other methods.[34] Here, a fully exchange deuterated, denatured protein, kept under pH conditions where hydrogen exchange with the ^{1}H-solvent is insignificant, is subject to refolding conditions at time point zero. After a variable refolding time τ_f, a short high pH pulse allows unprotected amides to exchange completely with ^{1}H, and folding proceeds in the absence of exchange. In ^{1}H- (or ^{15}N-HSQC) spectra of the native protein, signals of amides, which have gained a significant solvent protection within τ_f will be weakened relative to those, which remain exposed at that stage of the folding process.

In the intermediate exchange regime, linewidths of ^{1}H resonances may be directly converted into HX rates up to $k_{ex} \approx 10^3$ s^{-1}.[29] Faster exchange, as it occurs for solvent exposed amides at pH > 7 and high temperatures (> 40 °C), is no longer detectable by NMR spectroscopy, since the protons usually do not give rise to signals above noise level. In modern high field spectrometers, amides enter the fast exchange regime at $k_{ex} \approx 10^4$ s^{-1}, where they become finally indistinguishable from the solvent resonance. For this reason, most lysine ζ-NH_3, arginine η-NH_2 and histidine aromatic HN protons, which are rarely H-bonded, do not appear in spectra of proteins under biologically relevant conditions. The same holds true for serine, threonine and tyrosine OH, whose intrinsic chemical shifts are even closer to the water frequency.[35] Nevertheless, in highly concentrated protein solutions, intermediate regime hydrogen exchange of these individual moieties was shown to have a significant effect on the transverse relaxation rate of the H_2O resonance itself (**Figure 1.8.1**).[36]

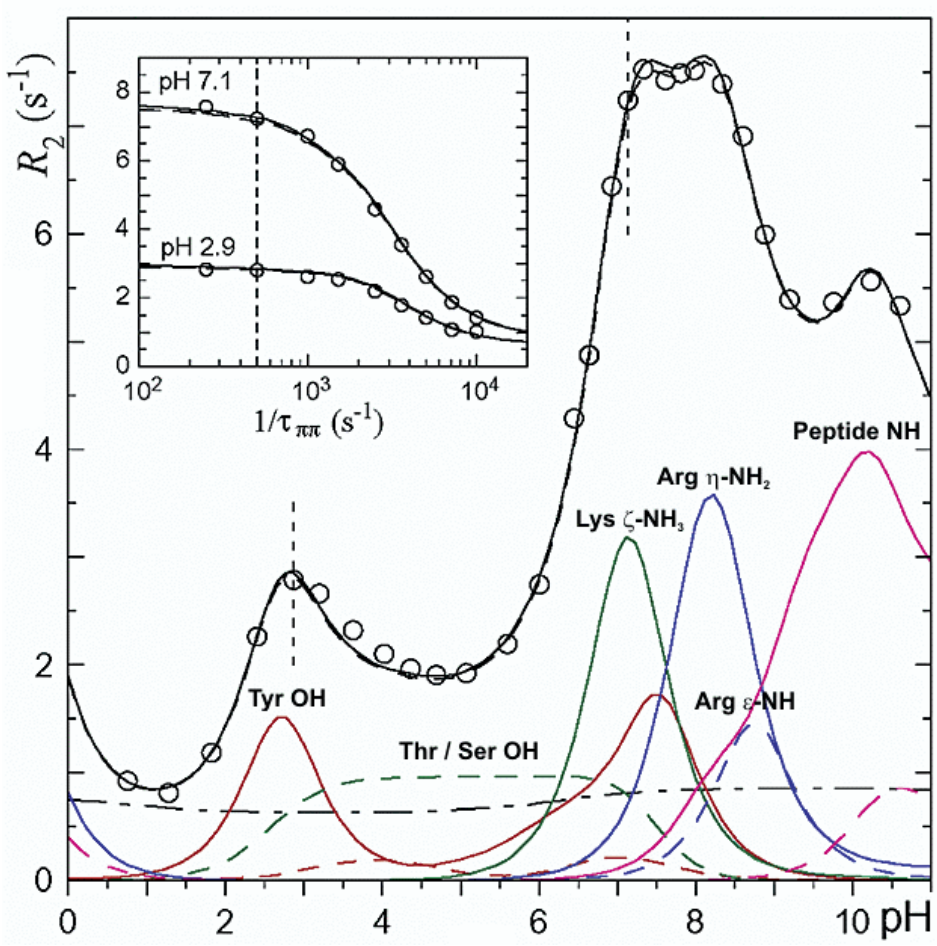


Fig.1.8.1: Transverse relaxation of the water resonance as a function of pH in a 39 mM (!) solution of BPTI. The open circles represents data from CPMG decay, in the inset from CPMG dispersion. The black solid line is back calculated from a fit summing contributions of different moieties. These are denoted above the individual maxima, where the respective protons are in the intermediate exchange with H_2O. Figure taken from [36].

For amides exchanging on the order of ${}^1J_{NH}$ ($k_{ex} \approx 10^2\ s^{-1}$), already the loss of phase coherence during INEPT transfer, being proportional to $\exp(-k_{ex}/{}^1J_{NH})$, leads to a significant signal attenuation.[37] Residues affected this way often escape ^{15}N-HSQC and H^N-detected triple resonance schemes and require additional circumstantial CH-detected experiments. The use of heteronuclear Hartmann-Hahn polarisation transfer, rather than INEPT, has been shown to yield significantly enhanced sensitivity in cases of fast hydrogen exchange.[38]

Hydrogen exchange of exposed amides with good ^{15}N-HSQC performance, but too fast for time resolved measurements is usually characterised by transfer of polarisation from or to the the bulk water after a selective perturbation. In its classical application, attenuation of ^{15}N-HSQC resonances upon presaturation of the water may be measured,[39] this, however, requires knowledge of the individual intrinsic proton spin-lattice relaxation rates R_{1H}. Several methods (MEXICO,[40] WEX-FHQC,[41] CLEANEX-PM [42]) have been developed, which aim at measuring the recovery of saturated amide magnetisation from exchange with the non-saturated solvent. Their principle will be discussed in the following section.

1.9 Exchange rates from polarisation transfer

The individual, time dependent polarisations of amide protons are coupled to a bath of nonzero bulk water magnetisation, which can be assumed to be constant due to the large excess of water protons. Every time a single proton exchanges between the bulk water and a particular amide group, on average the ”incoming“ proton transfers the water magnetisation FM_∞ and the leaving proton carries the current amide magnetisation M(t) with it. M in this context denotes the expectation value of the operator $\mathbf{I}_z$ in the given density matrix ρ as calculated from $M = \langle \mathbf{I}_z \rangle = \mathrm{tr}(\boldsymbol{\rho}\mathbf{I}_z)$. The factor F takes into account that water may be not be 100 % restored to the Boltzmann equilibrium M_∞ before the mixing period due to partial saturation, RF inhomogeneity, pulse imperfections and radiation damping, or it is inverted deliberately. However, given the slow spin lattice relaxation of water protons, it can be assumed to be constant over typical exchange periods of up to 200 ms. The magnetisation build-up on the amide group dM(t)/dt is proportional to the probability k_{ex} for an exchange process to happen per unit time:

$$\frac{dM(t)}{dt} = k_{ex}(FM_\infty - M(t)) \tag{1.9.1}$$

In order to solve this differential equation, the inhomogenous part FM_∞ needs to be added to the solution $M = M_0 \exp(-k_{ex}t)$ of the homogenous equation dM/dt + $k_{ex}M$ = 0. If amides are completely saturated at t = 0, the boundary condition $M(0) = M_0 + FM_\infty = 0$ yields:

$$M(t) = FM_\infty(1 - \exp(-k_{ex}t)) \tag{1.9.2}$$

After an initial linear build-up regime $M = FM_\infty k_{ex}t$ this function has, as expected, a steady state plateau at $M = FM_\infty$. $M(t)$ may be read out and assigned to individual protons by a 1D-watergate or ^{15}N-HSQC detection scheme.

A serious limitation of simply measuring this recovery is, that evidently saturated amide protons are subject to spin lattice relaxation during the mixing period, which may wrongly pretend hydrogen exchange. Longitudinal relaxation (rate constant R_{1H}) is always driven by the difference to the full Boltzmann polarisation M_∞ and contributes to the equation of motion:

$$\frac{dM(t)}{dt} = k_{ex}(FM_\infty - M(t)) + R_{1H}(M_\infty - M(t)) = (-k_{ex} - R_{1H})M + (Fk_{ex} + R_{1H})M_\infty \tag{1.9.3}$$

which is solved in analogy to above steps:

$$M(t) = M_0 e^{-(k_{ex}+R_{1H})t} + \frac{Fk_{ex}+R_{1H}}{k_{ex}+R_{1H}} M_\infty = \frac{Fk_{ex}+R_{1H}}{k_{ex}+R_{1H}} M_\infty (1 - e^{-(k_{ex}+R_{1H})t}) \quad (1.9.4)$$

where the last step enforces $M(0) = 0$. Now the steady state depends on F as well as the relative sizes of k_{ex} and R_1, whereas the initial slope is given by $M_\infty(Fk_{ex}+R_{1H})$. For $F \approx 1$, spin lattice relaxation and exchange act constructively and cannot be separated. Instead, if water is flipped to −z ($F \approx -1$) prior to the exchange period, both processes try to drive the magnetisation in the opposite sense and the initial slope and the plateau depend on the difference of k_{ex} and R_{1H}.

The current experimental schemes separate k_{ex} and R_{1H} by a phase cycle, in which the water magnetisation is alternating flipped to +z and −z, with the receiver phase adjusted accordingly. The final signal, averaged over two respective transients is then:

$$M(t) = \frac{1}{2}\left\{ \frac{Fk_{ex}+R_{1H}}{k_{ex}+R_{1H}} - \frac{-Fk_{ex}+R_{1H}}{k_{ex}+R_{1H}} \right\} M_\infty (1 - e^{-(k_{ex}+R_{1H})t}) = \frac{Fk_{ex}}{k_{ex}+R_{1H}} M_\infty (1 - e^{-(k_{ex}+R_{1H})t}) \quad (1.9.5)$$

Where F is now an average (positive) scaling factor of the water steady state polarisation. Obviously, the influence of amide proton longitudinal relaxation cannot simply be removed by phase cycling, which is a consequence of the nonlinear dependence of the relaxation response dM/dt to M, as already pointed out by Gemmecker *et al.*[40] Only in the initial build-up regime, where exchange causes only minor changes in the polarisation of amide protons, $M = FM_\infty k_{ex} t$ is virtually unperturbed by spin lattice relaxation. After all, this linear regime can extend to long mixing times (≈ 200 ms) for those residues, where k_{ex} is in the order of R_1 (≈ 1 s^{-1}). Usually, a series of experiments with variable mixing periods is performed, and the measured intensities are divided to M_∞ obtained from the reference ^{15}N-HSQC experiment. Individual k_{ex} should be extracted from the data points, where the polarisation increase is still nearly linear. F may be calibrated from measuring the water signal after the mixing period or from the plateau of residues with $k_{ex} \gg R_{1H}$.

Further problems arise from the presence of cross-relaxation during the mixing period, which may be principally suppressed using conditions, where $\omega\tau_c \approx 1$ is met,[19] or by NOE-ROE cancellation (*clean mixing*) in the slow tumbling limit.[42] In order to keep polarisation buildup by intramolecular NOE effects as small as possible, all protein protons have to be saturated prior to mixing by heteronuclear ^{15}N and ^{13}C filtering.[40, 43] A variant of MEXICO applicable to only ^{15}N-labelled proteins, NewMEXICO (see also section 7.8), uses radiation damping (RD) of the water signal within 20-40 ms after an initial π/2 pulse as an active element to

discriminate between water and carbon bound protons.[44] Before the mixing period, all amide protons are again completely saturated by a ^{15}N filter. Although the RD-"pulse" has been shown to be highly selective,[45] H^{α} protons resonating exactly at the water frequency may be restored to the z-axis and transfer polarisation to nearby H^{N} protons via NOE. The same is valid for OH protons, which are usually either merged with the bulk water signal (fast exchange) or at least exchange quickly during mixing and give rise to exchange-relayed NOE effects. Secondary effects including cross relaxation between adjacent amide protons (with strongly different k_{ex}) or between amide and aliphatic protons may be neglected in the initial buildup regime. Summing up these cross-relaxation side effects, a systematic error and therefore lower limit of approximately 0.5 s^{-1} seems reasonable for the determination of individual HX rates.

2 Heteronuclear correlation spectroscopy

2.1 Coherence transfer

Correlating the Larmor precession frequencies of covalently attached spins in heteronuclear multidimensional NMR spectroscopy requires coherences to be efficiently transferred through a network of *J*-couplings, for example in the protein backbone or sidechains. Insensitive nuclei with low gyromagnetic ratio γ like ^{15}N are often only accessible *via* such transfer steps from a nucleus with much higher thermal polarisation like ^{1}H. Formally, the term *coherence* denotes non-diagonal elements of the density operator, represented in the eigenbase of $\mathbf{I}_z$, as for example in the case of an ensemble of non-interacting $I = ½$ spins:[1, 46, 47]

$$\boldsymbol{\rho}(t) = \begin{pmatrix} \rho^{\alpha}(t) & \rho^{+}(t) \\ \rho^{-}(t) & \rho^{\beta}(t) \end{pmatrix} = \rho^{\alpha}(t)\mathbf{I}^{\alpha} + \rho^{\beta}(t)\mathbf{I}^{\beta} + \rho^{+}(t)\mathbf{I}^{+} + \rho^{-}(t)\mathbf{I}^{-} \qquad (2.1.1)$$

where $\mathbf{I}^{+}$ etc. are (stationary) elementary spin operators and ρ^{+} and ρ^{-} are called (time dependent) (+1)- and (−1)-quantum coherences, respectively. Physically, coherence requires the existence of spins which are in a superposition of α and β quantum states and additionally partially aligned in the xy-plane over the ensemble. The degree of spin alignment is given by the magnitude, and its direction by the phase of the complex numbers ρ^{+} and ρ^{-}. The classical cartesian magnetisation vector components,[48] often also called coherences, are quantum mechanical *expectation values* of spin operators, which may be always calculated from the density according to: $M_x = \langle \mathbf{I}_x \rangle = \mathrm{tr}(\boldsymbol{\rho}\mathbf{I}_x) = ½(\rho^{+} + \rho^{-})$, etc. Usually prepared from the thermal equilibrium polarisation M_z, under free Larmor precession in the rotating frame (+1)- and (−1)-quantum coherences evolve with negative and positive frequencies, respectively:[49]

$$M_z \xrightarrow{\frac{\pi}{2}\mathbf{I}_y} M_x = \frac{1}{2}(\rho^{+} + \rho^{-}) \xrightarrow{\Omega_I \mathbf{I}_z t} \frac{1}{2}(\rho^{+} e^{-i\Omega_I t} + \rho^{-} e^{+i\Omega_I t}) \qquad (2.1.2)$$

All terms carry the Boltzmann-factor $\hbar\gamma_I B_0/2kT$, which is proportional to the gyromagnetic ratio γ_I and the static magnetic field B_0. The NMR signal induced in the coil depends on the one hand on the bulk magnetic moment in the sample (proportional to the number of spins and γ_I), and on the other hand on the Larmor frequency ω_I (= $-\gamma_I B_0$). A property of the quadrature detection scheme during acquisition is that exclusively (−1)-quantum coherence is detected.

In a two-spin system IS, I- and S-spin polarisations may exist correlated, and the evolution of spin density can be efficiently treated using operator *products*. For example, a non vanishing expectation value $\langle 2\mathbf{I}_x\mathbf{S}_z\rangle$ describes a situation in which neither I-spins have a net alignment along x nor S-spins along z, but in a single spin pair there is a certain nonzero probability for S to be polarised in the z-direction, if I is polarised along x, and vice versa. Cartesian antiphase I-spin coherence $\langle 2\mathbf{I}_x\mathbf{S}_z\rangle$ is completely converted into antiphase S-spin coherence $\langle 2\mathbf{I}_z\mathbf{S}_x\rangle$ (and back to $\langle 2\mathbf{I}_x\mathbf{S}_z\rangle$) by simultaneous application of $\pi/2$-pulses on both RF-channels with phase *y*. Using such *coherence transfer steps* (*CTS*), the Boltzmann factor associated with the originally excited nucleus may be propagated throughout the whole pulse sequence, thus enabling indirect detection of low-sensitivity nuclei like ^{15}N.

In contrast, pure (+1)-quantum antiphase coherence $\langle \mathbf{I}_z\mathbf{S}^+\rangle$ is transferred not only to I-spin single quantum, but also double and zero quantum terms by a simple CTS, because $\mathbf{S}^+ = \mathbf{S}_x + i\mathbf{S}_y$ contains both orthogonal cartesian operators. Also, both SQ terms $\langle \mathbf{I}^+\mathbf{S}_z\rangle$ and $\langle \mathbf{I}^-\mathbf{S}_z\rangle$ are excited, but each only with half amplitude (**Fig. 2.1.1 A**).

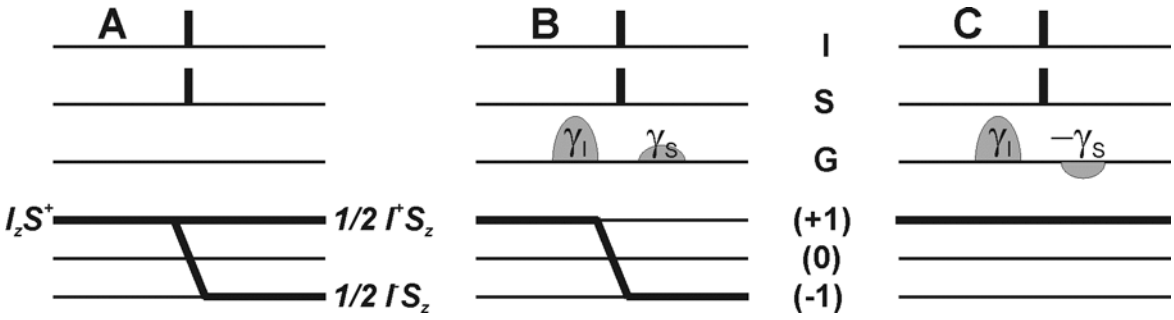


Fig. 2.1.1: Coherence level diagrams[50] for CTS from S to I by two simultaneous $\pi/2$ pulses on both RF-channels. **A:** without gradients. **B:** echo pathway. **C:** antiecho pathway. Brackets are omitted for simplicity.

After the application of a pulsed field gradient (PFG) *G* for a period τ_G, $\langle \mathbf{I}_z\mathbf{S}^+\rangle$ acquires a (*z*-dependent) phase factor $\exp(-i\gamma_S Gz\tau_G)$, which in combination with a second, rephasing gradient pulse after the CTS may be used to select either $\frac{1}{2}\langle \mathbf{I}^+\mathbf{S}_z\rangle$ (**Fig. 2.1.1 B**) or $\frac{1}{2}\langle \mathbf{I}^-\mathbf{S}_z\rangle$ (**Fig. 2.1.1 C**). Pathways with inversion and retention of coherence order are, in analogy with homonuclear refocussing, called *echo* and *antiecho* pathways,[51] and they identically apply for transfer of the adjoint coherence $\langle \mathbf{I}_z\mathbf{S}^-\rangle$. Antiphase coherences like $\langle \mathbf{I}_z\mathbf{S}^+\rangle$ are readily converted into the respective inphase coherences $\langle \mathbf{S}^+\rangle$, and *vice versa*, by J_{IS}-evolution periods with a length of $1/(2J_{IS})$.

2.2 The HSQC experiment

HSQC, the fundamental phase cycled version of which is shown in **Fig. 2.2.1**, is the most widely used experiment for correlating frequencies of a sensitive (I) and an insensitive (S) nucleus through the heteronuclear scalar coupling constant J_{IS}.[52]

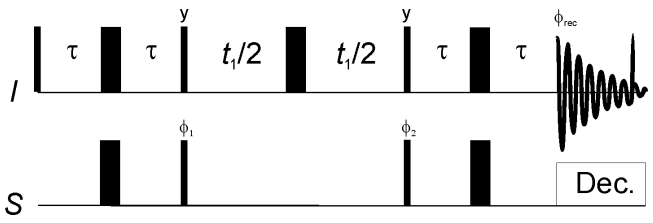


Fig. 2.2.1: Pulse scheme of the fundamental HSQC experiment. The delay τ is set to $1/(4J_{IS})$ and decoupling of the S-spins during acquisition may be accomplished with a GARP sequence. Cycling of ϕ_1 (= y,-y) ensures that all protons not bound to S are cancelled, and in order to suppress DQ and ZQ terms generated by the second CTS, a four step phase cycle (ϕ_2 = 2(y),2(−y) and ϕ_{rec} = y,2(−y),y) needs to be applied. The phase of all other RF-pulses is x unless stated otherwise. Quadrature detection is achieved by the States method.[53]

Through an INEPT element,[54] single quantum (SQ) $\langle 2\mathbf{I}_z\mathbf{S}_x\rangle$ coherence is generated, which is allowed to evolve under $\omega_S\mathbf{S}_z$ during a variable period t_1 and transferred back to I-spin coherence, which is ultimately detected during t_2. Only one of the cartesian products $\langle 2\mathbf{I}_z\mathbf{S}_x\rangle$ and $\langle 2\mathbf{I}_z\mathbf{S}_y\rangle$ present at the end of the t_1-evolution period can be transferred back to observable proton magnetisation, leading to an *amplitude modulation* (AM) of the final time domain signal. Quadrature detection in F_1 (*States method*[53]) is achieved by interleaved sampling of *cosine* and *sine* data points (by shifting ϕ_1) in the odd and even serial FIDs, respectively, to yield the following interferograms:

$$f_{cos}(t_1,t_2) = \cos(\Omega_S t_1)\exp(i\Omega_I t_2) \qquad \text{detection of} \quad \langle \mathbf{S}_x\rangle = \tfrac{1}{2}\ (\langle \mathbf{S}^+\rangle + \langle \mathbf{S}^-\rangle) \tag{2.2.1}$$

$$f_{sin}(t_1,t_2) = \sin(\Omega_S t_1)\exp(i\Omega_I t_2) \qquad \text{detection of} \quad \langle \mathbf{S}_y\rangle = \tfrac{1}{2}i\ (-\langle \mathbf{S}^+\rangle + \langle \mathbf{S}^-\rangle)$$

Other schemes like *TPPI*[55] and *States-TPPI* are closely related. Although some readjusting of the receiver gain to unwanted signals does not severely affect the sensitivity of an experiment, a desired signal never rises above noise level, if it falls below the digitiser resolution (=1/dynamic range). Thus, in order to make the scheme of Fig. 2.2.1 applicable to aqueous protein solutions, the water signal, which exceeds the protein resonances typically by 10^5, needs to be actively suppressed before reaching the receiver.[56, 57]

A simple method is *presaturation* of the water polarisation by a long selective irradiation of a few mW at the water frequency, which, however, also saturates fast exchanging amide protons and thus makes them unobservable (**Fig. 2.2.2 A**).

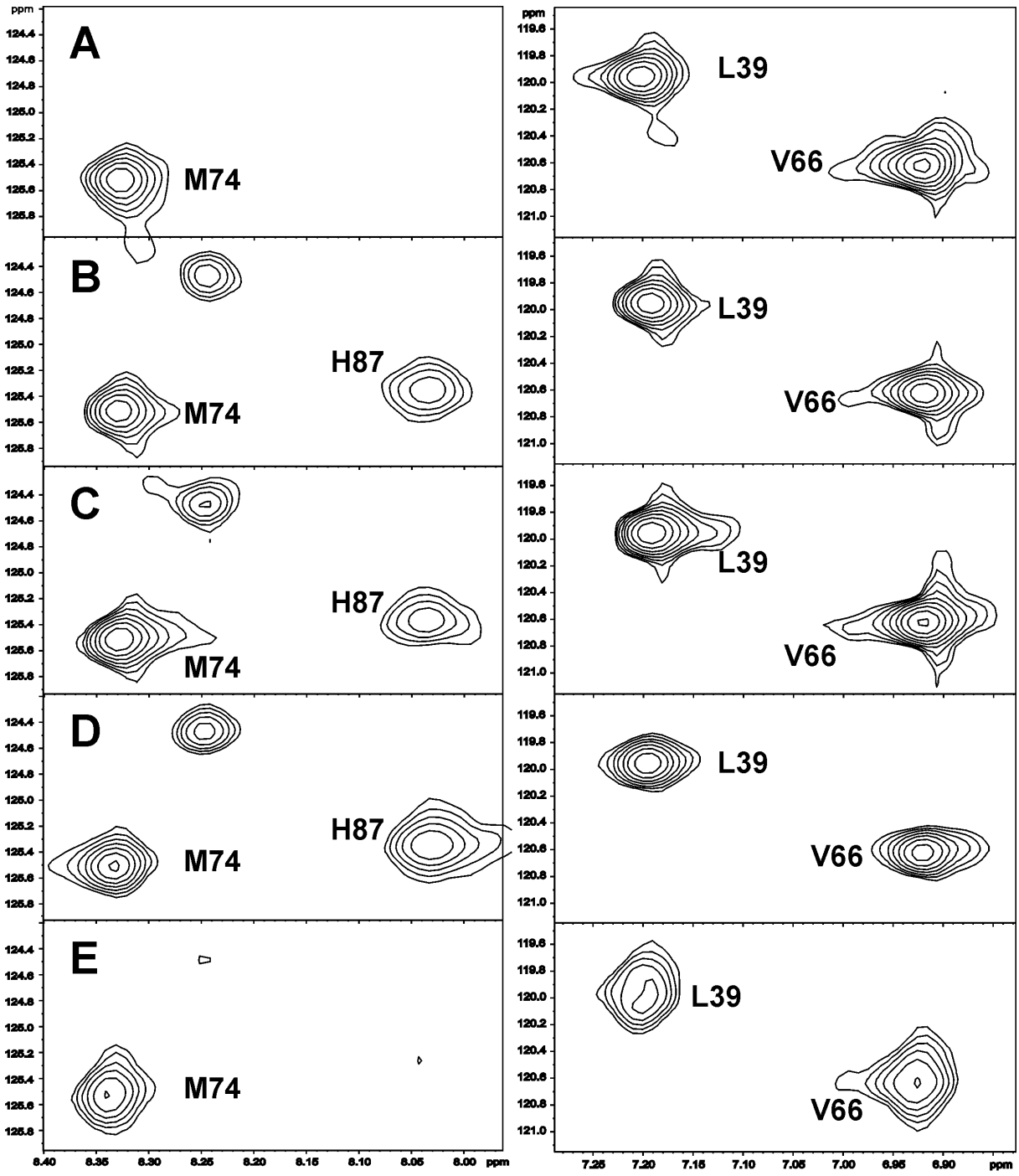


Fig. 2.2.2: Comparison of ^{15}N correlation spectra for different amide resonances in saposin C at pH 7 and 37 °C. H87 undergoes fast solvent exchange, whereas V66 is closest to the water frequency. Spectral parameters and experiment time are identical. **A:** HSQC with 2s presaturation. **B:** gradient-selected HSQC (see section 2.3). **C:** gradient-selected sensitivity-enhanced HSQC. **D:** FHSQC with 3-9-19 watergate sequence. **E:** HMQC

In gradient-selected experiments (see section 2.3), water magnetisation is not rephased by the second PFG and thus very efficiently supressed without the use of selective pulses (**Fig. 2.2.2 B** and **C**). Also, a homonuclear gradient echo may be combined with a band-selective refocussing π-pulse on the desired protons, whose inversion profile sharply drops to zero at the water frequency. Binomial (*WATERGATE*) sequences,[58, 59] which achieve good selectivity in a relatively short period, have been developed and incorporated into the HSQC scheme (**Fig. 2.2.3**).

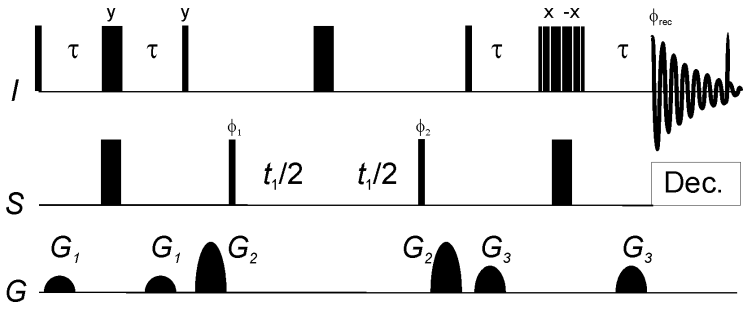


Fig. 2.2.3: FHSCQ with solvent supression by a binomial 3-9-19 sequence in combination with flip-back. Pulse phases and delays are as in Fig. 2.2.1. Gradients are variable and may be optimised for water suppression.

This scheme has been called *fast*-HSQC (*FHSQC*)[60] and is preferably employed if transverse relaxation of protons is fast due to the molecular size or conformational or hydrogen exchange (**Fig. 2.2.2 D**). As in virtually all modern amide detected experiments, water is flipped back to the z-axis prior to acquisition in order to avoid partial saturation of the water resonance.[61] Since the water magnetisation can behave rather unpedictable during delays due to radiation damping,[62] it is essential to guide it carefully through the sequence by use of selective pulses and de- and rephasing by PFGs. In FHSQC, partial supression of protons resonating close to the water frequency may occur, therefore the scheme cannot be used for the detection of H^{α} protons.

Instead of using SQ-evolution (HSQC), the heteronuclear frequency Ω_S may be evolved alternatively from DQ ($\Omega_S+\Omega_I$) and ZQ ($\Omega_S-\Omega_I$) coherences during periods of equal length (*HMQC*).[63] In terms of their relaxation properties, multiple quantum coherences have been shown to be superior for certain spin systems in large proteins.[64] For ^{15}N correlations in small to medium proteins, however, HMQC is usually less sensitive due to the modulation of the t_1-time domain signal with homonuclear J_{HH} couplings, leading to rather broad resonances (**Fig. 2.2.2 E**).

2.3 Gradient selection

Alternatively to phase cycling, the desired coherence pathway can be selected by a pair of PFG, placed shortly after the t_1-evolution period and prior to acquisition, in a single scan. Because $\langle \mathbf{S}^+ \rangle$ and $\langle \mathbf{S}^- \rangle$ evolve with opposite frequencies, *echo*- and *antiecho* coherence selection yields different interferograms, which are in analogy to the States scheme recorded and stored separately:

$$f_{\text{echo}}(t_1,t_2) = \tfrac{1}{2} \exp(-\mathrm{i}\Omega_S t_1) \exp(\mathrm{i}\Omega_I t_2) \qquad \text{detection of} \quad \tfrac{1}{2}\langle \mathbf{S}^+ \rangle \qquad (2.3.1)$$

$$f_{\text{antiecho}}(t_1,t_2) = \tfrac{1}{2} \exp(+\mathrm{i}\Omega_S t_1) \exp(\mathrm{i}\Omega_I t_2) \qquad \text{detection of} \quad \tfrac{1}{2}\langle \mathbf{S}^- \rangle$$

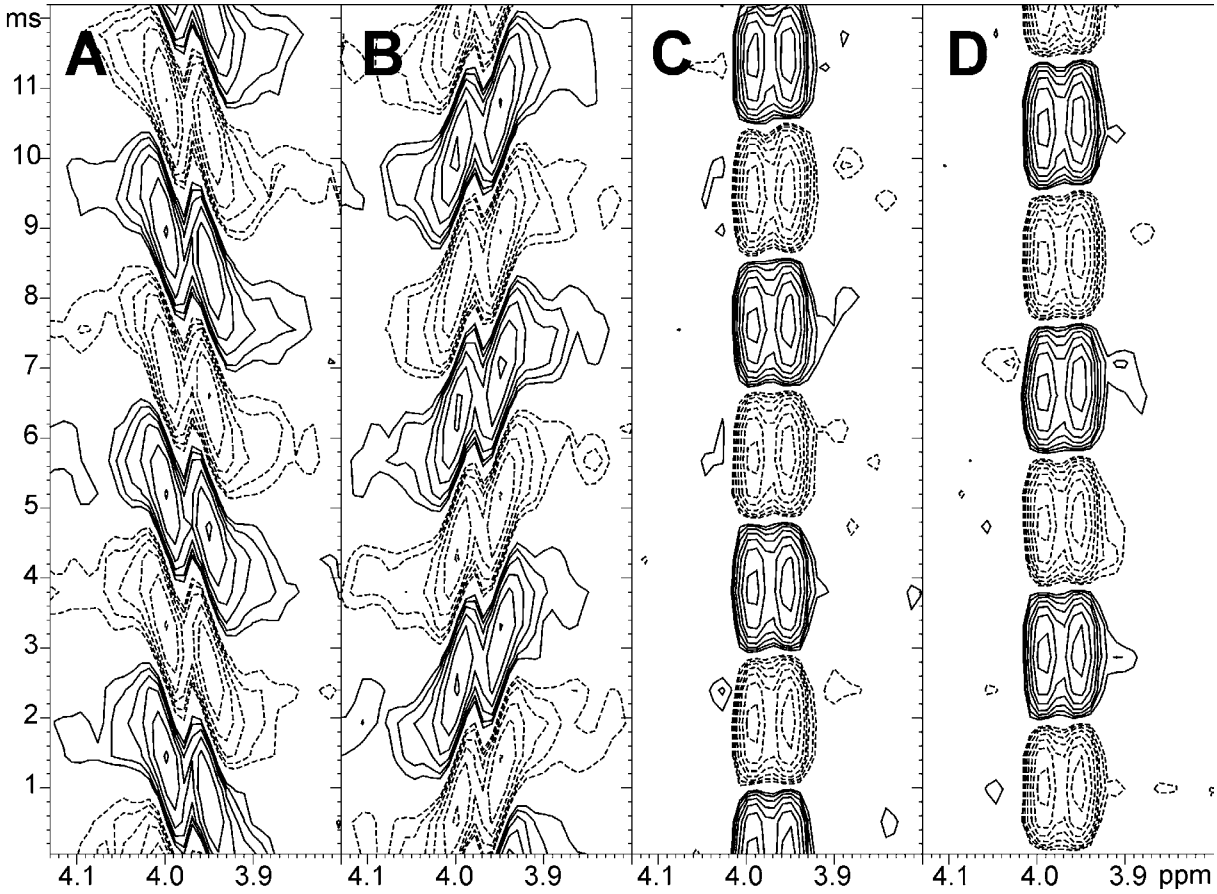


Fig. 2.3.1: F_2-Fourier transformed interferograms from a gradient selected ^{13}C-HSQC experiment with usual interleaved echo/antiecho selection, showing a ^{1}H doublet of dioncophillin. t_{aq} = 12.2 ms corresponds to the 32th t_1-increment (the 64th experiment). **A:** f_{echo}, constructed from experiments number 1, 3, 5, etc. **B:** f_{antiecho}, constructed from experiments number 2, 4, 6, etc. **C:** i(f_{echo}+ f_{antiecho}), corresponding to (f_{echo} + f_{antiecho}) and 90° phase shift, and **D:** (f_{antiecho} − f_{echo}). The signal in **A** and **B** can be imagined as coils turning in opposite directions with parallel alignment at time points with $\Omega_S t_1/\pi = 0, 2, 4$, etc. and antiparallel alignment at time points with $\Omega_S t_1/\pi = 1, 3, 5$, etc. **C** and **D** correspond exactly to the States data $f_{\cos}$ and $f_{\sin}$, except that the noise is also added.

The factor ½ arises because only half of $\langle \mathbf{S}^{\pm} \rangle$ is actually transferred, as discussed section 2.1. From these *phase modulated* data, States data can be constructed simply by addition $f_{cos} = f_{echo} + f_{antiecho}$ and subtraction $f_{cos} = i(f_{echo} - f_{antiecho})$, thereby enabling quadrature detection in F_1 (**Fig. 2.3.1**).

While exactly retaining the desired signal, adding two subspectra also increases the noise level by a factor of √2, making the echo/antiecho method less sensitive than the conventional HSQC.[65, 66] Signal is principally lost if field gradients are combined with CTS by RF-pulses, which select cartesian rather than coherence operators. To overcome this drawback, an elegant method has been introduced by Rance [67] and Kay.[68] *Sensitivity enhancement* makes use of two othogonal CTSs, which allow to convert both cartesian components $\langle 2\mathbf{I}_z\mathbf{S}_x \rangle$ and $\langle 2\mathbf{I}_z\mathbf{S}_y \rangle$ into observable I-spin magnetisation (**Fig. 2.3.2**).

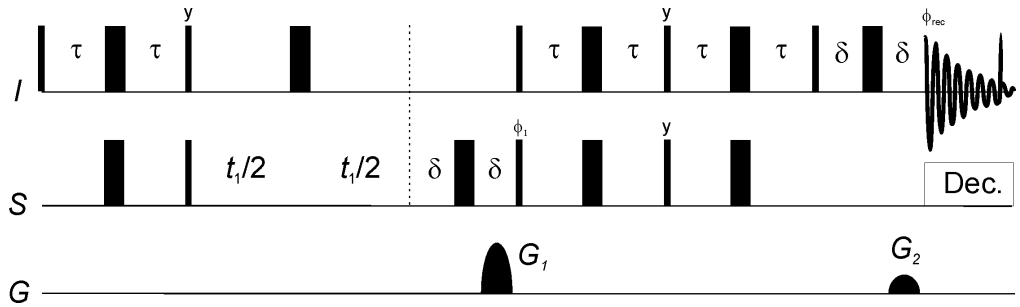


Fig. 2.3.2: Sensitivity-enhanced HSQC with coherence selection by field gradients, disregarding solvent suppression. G_2 is adjusted to the relative larmor frequencies of I and S spins. δ is just long enough to accommodate a gradient pulse (e.g. 800 ms) with a subsequent ring down delay (e.g. 200 ms). For the echo pathway ϕ_1 = x, and G_1 and G_2 have opposite sign (for positive γ_S), for the antiecho pathway ϕ_1 and G_2 are inverted with constant ϕ_{rec} (= x). The phase of all other pulses is x, unless stated otherwise, and parameters are the same as in Fig. 2.2.1.

In terms of coherences, either full $\langle \mathbf{S}^{+} \rangle$ (echo) or $\langle \mathbf{S}^{-} \rangle$ (antiecho), which is present during t_1, is transferred to $\langle \mathbf{I}^{-} \rangle$, as controlled by the phase ϕ_1 of the first transfer pulse on S. In this scheme, no signal is discarded by the additional employment of gradients, only their relative sign has to match the selection done by ϕ_1, Again, echo and antiecho subspectra are recorded interleaved, but now, each t_1-increment contains the full phase-modulated signal. As summarised in **Fig. 2.3.3**, the result is an enhancement of sensitivity by a factor of 2 compared to the simple gradient-seclected scheme, or √2 compared to the conventional HSQC.

In practice, the gain of √2 is barely reached, because the sensitivity-enhanced scheme contains additional pulses and is longer at least by 1/(2*J*), for the gradient selected version by another 4δ. For larger proteins with fast transverse relaxation and IS-groups which are subject to conformational or solvent exchange, this can turn into a severe drawback. On the other hand, in gradient selected schemes the receiver does not have to be adjusted to strong undesired signals from solvent, buffer or, in the case of ^{13}C at natural abundance, the 100 fold excess of ^{12}C bound protons.

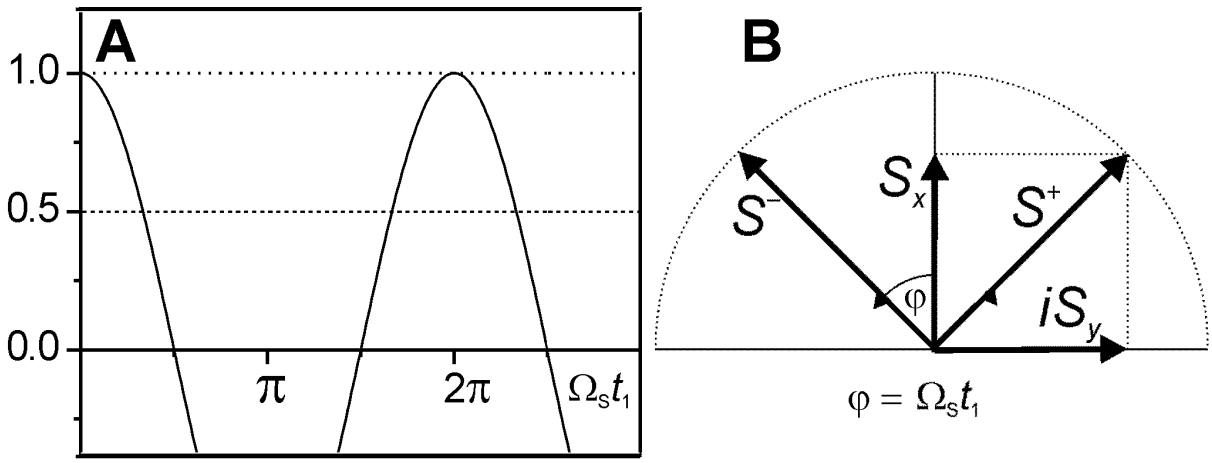


Fig. 2.3.3: A: Cosine amplitude-modulated signal detected by the States method (solid line) vs. signal with constant amplitude, but modulated phase as detected by the echo/antiecho method, with simple gradient selection (dashed line) and sensitivity enhancement (dotted line). The root mean square of the cosine function is √(1/2). **B:** Vector diagram of S-spin coherences. Echo and antiecho methods detect the counter-rotating vectors $\langle \mathbf{S}^+ \rangle$ and $\langle \mathbf{S}^- \rangle$, respectively, whereas the States scheme detects projections onto the axes, which are on average √2 smaller. Brackets are omitted for simplicity.

In triple resonance experiments, full phase modulated transfer is principally possible for each CTS, but the gain is usually outbalanced by relaxation losses for small *J*-couplings, i.e. long *J*-evolution periods. Furthermore, if a pair of de- and rephasing PFGs is used to select coherences, the gradients should be applied as close in time as possible in order to avoid signal decay by translational diffusion. Sensitivity enhancement may also suffer from signal phase distortions, which result from differential relaxation of the spin terms involved in the CTS, and which may severely deteriorate the correlation spectrum.

2.4 Real time and constant time evolution

Although alternative approaches using selective excitations are currently experiencing a revival,[69-71] most commonly frequency labelling of nuclei requires the excitation of coherences, in which the respective spins are polarised in the transverse plane. For example, the HSCQ schemes shown in the previous sections excite (and finally transfer) $\langle \mathbf{I}_z\mathbf{S}^+ \rangle$ and $\langle \mathbf{I}_z\mathbf{S}^- \rangle$ coherences, which then evolve under a Hamiltonian containing the chemical shift interaction of S nuclei (**Fig. 2.4.1 A**). All these schemes have in common that the sampling of individual t_1 data points is done during a coherent period τ_{co} (**Fig. 2.4.1 B**), whose length is exactly t_1, i.e. in *real time* (*RT*). During this period, the coherence decays exponentially with a rate R, which includes in- and antiphase transverse relaxation and inhomogeneity of B_0, thus giving rise to a frequency domain signal which is proportional to the area below the decay curve up to the acquisition time t_{aq} (**Fig. 2.4.1 C**). *J*-coupling to I and further heteronuclei may, if not explicitly desired, always be refocussed by the application of appropriate π-pulses on the respective RF channel.

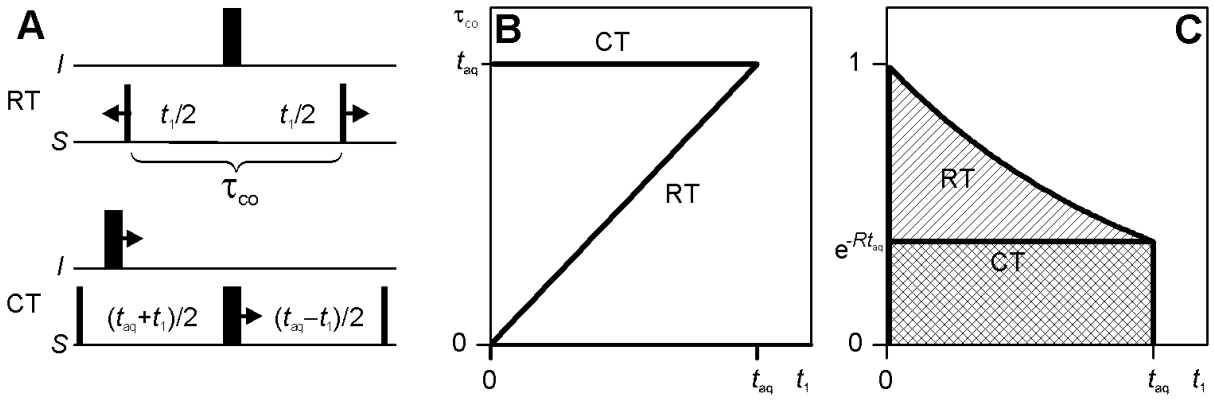


Fig. 2.4.1: t_1-evolution. **A:** Pulse sequence elements for real time and constant time evolution. **B:** Coherent period $\tau_{co}(t_1)$ required to sample t_1 data points. **C:** Enveloping (scaled) time domain signal $f(t_1) = \exp(-R\tau_{co})$. The frequency domain signal is proportional to the integral of $f(t_1)$ from $t_1 = 0$ to $t_1 = t_{aq}$ (hatched).

Alternatively, τ_{co} may be kept constant (= *constant time, CT*), and the phase $\Omega_S t_1$ may be evolved by shifting a π-pulse on the S channel may be shifted away from the centre by $t_1/2$.[72] In order to achieve the desired spectral resolution $1/t_{aq}$, $\tau_{co} = t_{aq}$ has to be chosen, otherwise the π-pulse reaches the end of the constant time before at $t_1 = \tau_{co}$. This results in a strongly reduced frequency domain signal, which is further lowered by apodisation to zero, unless resonance side bands are tolerated.

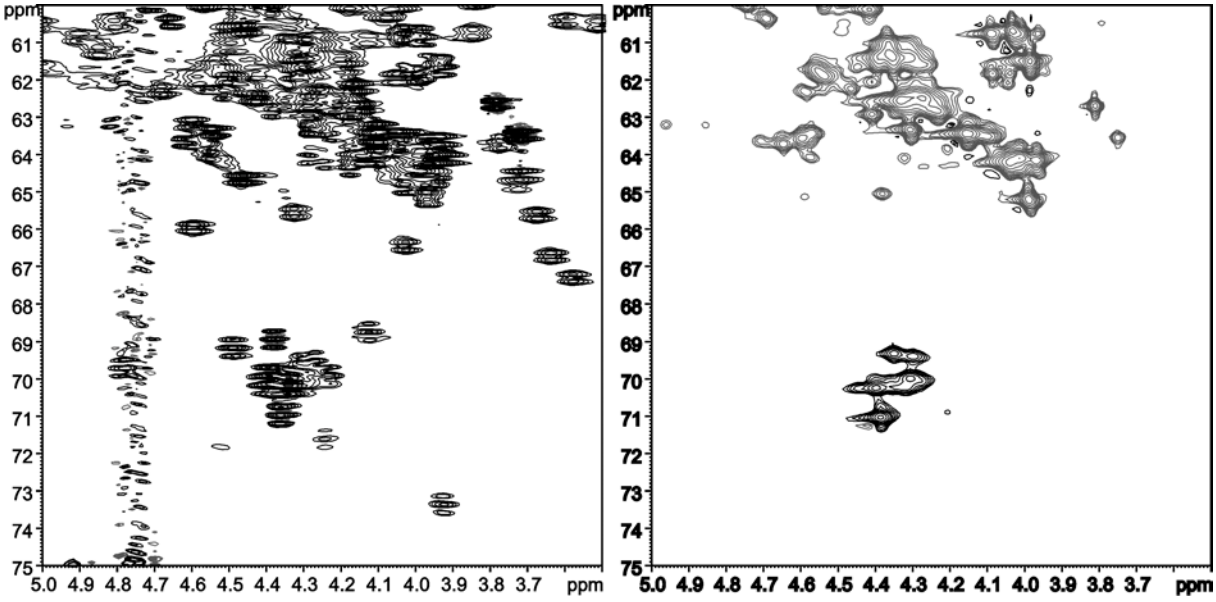


Fig. 2.4.2: Identical cutouts from ^{13}C-HSQC spectra of αADT. **A:** Real time t_1-evolution, during which $^1J_{CC}$ (≈ 35 H) fully evolves. Doublets indicate C^{α} atoms and C^{β} atoms of serine, whereas pseudo-triplets are observed for threonine C^{β} atoms. The indirect acquisition time is t_{aq} = 41 ms (LP to 62 ms). **B:** Constant time t_1-evolution with τ_{co} = 28 ms = $1/J_{CC}$ (n = 1). The phase of the resonances is given by the multiplicity of attached non-carbonyl carbons (grey = negative, C^{α}, and black = positive, C^{β} of threonine). The reduced sensitivity is apparent from the lack of signals compared to A.

The reduced sensitivity of HSQC with CT chemical shift evolution restricts its application to cases where homonuclear $^1J_{CC}$ (≈ 35 Hz) multiplicity needs to be suppressed in uniformly ^{13}C-labelled proteins by chosing $\tau_{co} = n/^1J_{CC}$. (**Fig. 2.4.2**).[73]

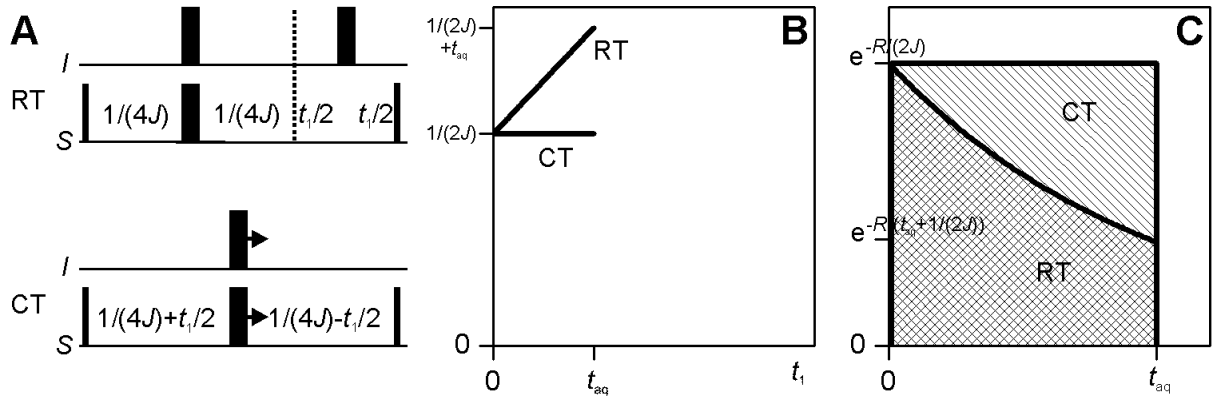


Fig. 2.4.3: Combined t_1- and J-evolution in the case of $t_{aq} < 1/(2J)$, here $t_{aq} = 1/(4J)$. **A:** Pulse sequence elements for concatenated J-and RT-t_1-evolution periods and t_1-evolution integrated into a CT-J-evolution. **B:** Overall coherent period $\tau_{co}(t_1)$ required to evolve J and sample t_1 data points. **C:** Enveloping (scaled) time domain signal as described in Fig. 2.6.1.

In triple resonance schemes, chemical shift evolution periods are often directly pre- or succeeded by *J*-evolution periods (= 1/(2*J*)), which are required to prepare coherence transfer and which can extend to several tens of milliseconds for small values of *J*. Since in 3D experiments t_{aq} for the indirect dimensions is usually small, t_1 (or t_2) evolution may be integrated as constant time in the *J*-evolution period rather than concatenated, provided that $t_{aq} < 1/(2J)$ (**Fig. 2.4.3**).

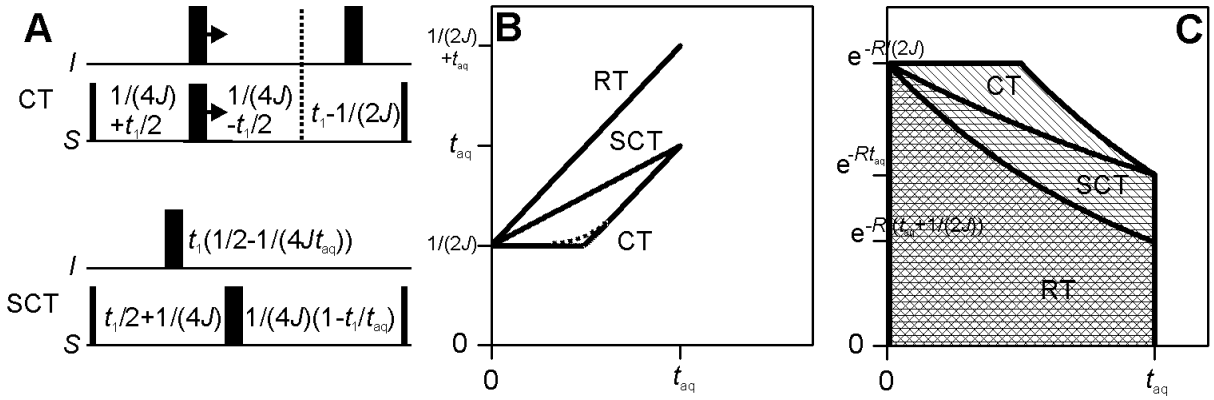


Fig. 2.4.4: Combined t_1- and *J*-evolution in the case of $t_{aq} > 1/(2J)$, here $t_{aq} = 1/J$. **A:** Pulse sequence elements for RT-extended CT-*J*- and t_1-evolution and SCT evolution. The pure RT scheme is not shown. **B:** Overall coherent period $\tau_{co}(t_1)$ required to evolve *J* and sample t_1 data points. The slope of the SCT scheme is $1-1/(2Jt_{aq})$. Nonlinear incrementation is indicated by a dotted curve. **C:** Enveloping (scaled) time domain signal as described in Fig. 2.6.1.

This simple method is no longer viable for $t_{aq} > 1/(2J)$, i.e. if high spectral resolution is desired or the coupling to evolve is large ($^1J_{NH}$, $^1J_{CH}$ or $^1J_{COCA}$). In this case, the first t_1 data points may be fully integrated into the CT of length 1/(2*J*), but the rest of them needs to be sampled by an appended RT. Alternatively, the overall coherent period τ_{co} may be allowed to grow linearly in t_1 right from the beginning to exactly reach the final value $\tau_{co} = t_{aq}$ at $t_1 = t_{aq}$ (**Fig. 2.4.4**). The slope $1-1/(2Jt_{aq})$ of this *semi-constant time* (*SCT*)[74-76] growth is between zero (constant time) and unity (real time), depending on *J* and the chosen t_{aq}, and the π-pulses on S and its *J*-coupled partner have to be shifted with a corresponding scaled increment. If the spectrometer software allows a more flexible handling of delays than just incrementation by a fixed period, a curve $\tau_{co}(t_1)$ may be chosen which approaches the CT curve without showing its kink. By using this *nonlinear incrementation*, one should gain up to 10 % sensitivity compared to the conventional SCT scheme.

2.5 Optimising sensitivity

Of course, getting the maximum sensitivity (= signal to noise ratio per unit experiment time) out of the protein solution in the NMR tube requires careful calibration of all pulses and the appropriate choice of solvent suppression, transfer and evolution schemes as described in the previous sections. For large proteins, shortening of the INEPT transfer delay τ to 70-80 % of $1/(4J)$ or use of the heteronuclear NOE for ^{13}C polarisation enhancement [77] may improve the sensitivity. Furthermore, if a full recovery is not essential for the experiment, the recycle delay τ_{rec} may be optimised with respect to the longitudinal relaxation rate R_{1H} of the excited protons. Assuming that all protons are saturated prior to acquisition, their recovery during the recycle delay is given by:

$$1-\exp(-R_{1H}\tau_{rec}) \quad = \quad 1-\exp(-\lambda) \tag{2.5.1}$$

In most NMR experiments, the recycle delay determines the length of one transient, and therefore the number of transients within fixed experiment time is approximately inverse proportional to τ_{rec}, and hence to λ. Considering that signal and noise accumulate directly and with the square root of the number of transients, respectively, the sensitivity follows the solid curve shown in **Fig. 2.5.1**. For most rigid protons in proteins $R_{1H} \approx$ 1-1.5 s^{-1} is rather insensitive to the molecular size, temperature and spectrometer field, thus, a recycle delay of 1s is close to the optimum in most cases.

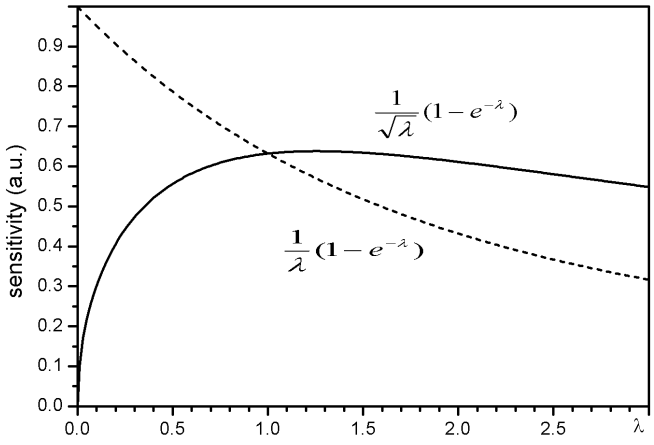


Fig. 2.5.1: Solid curve: Sensitivity as function of $\lambda = R_{1H}\tau_{rec}$ (the recycle delay, expressed in units of $1/R_{1H}$). The same curve is obtained for $\lambda = Rt_{aq}$ (the acquisition period, expressed in units of FID decay time $1/R$), if the the experiment time is independent of t_{aq} (direct dimension). The maximum is at $\tau_{rec} \approx 1.25$. Dashed curve: Sensitivity as function of $\lambda = Rt_{aq}$, if experiment time is proportional to t_{aq} (indirect dimensions).

When resolution is not an issue in the NMR experiment, the question arises to what extend sampling of real time t_1 (or t_2) increments still contributes to the signal/noise ratio. If apodisation is neglected, signal accumulates according to the integral of the FID envelope from zero to t_{aq}, which is given by:

$$\int_0^{t_{aq}} \exp(-Rt)dt = \frac{1}{R}(1-\exp(-Rt_{aq})) \propto (1-\exp(-\lambda)) \qquad (2.5.2)$$

whereas white noise accumulates with $\sqrt{t_{aq}}$. The ratio of both, describing the sensitivity in cases where experiment time is independent of t_{aq}, i.e. in the direct acquisition dimension, adopts the same functional form as for the recycle delay, if now $\lambda = Rt_{aq}$ is substituted. In indirect dimensions, however, experiment time is directly proportional to t_{aq}. In other words, for a fixed experiment time extensive sampling replaces transient repeats, which needs to be accounted for by another factor $1/\sqrt{t_{aq}}$. As shown by the dashed curve in **Fig. 2.5.1**, sensitivity continues to increase with decreasing the indirect acquisition times, and the actually chosen values of t_{aq} rather depend on the desired resolution.

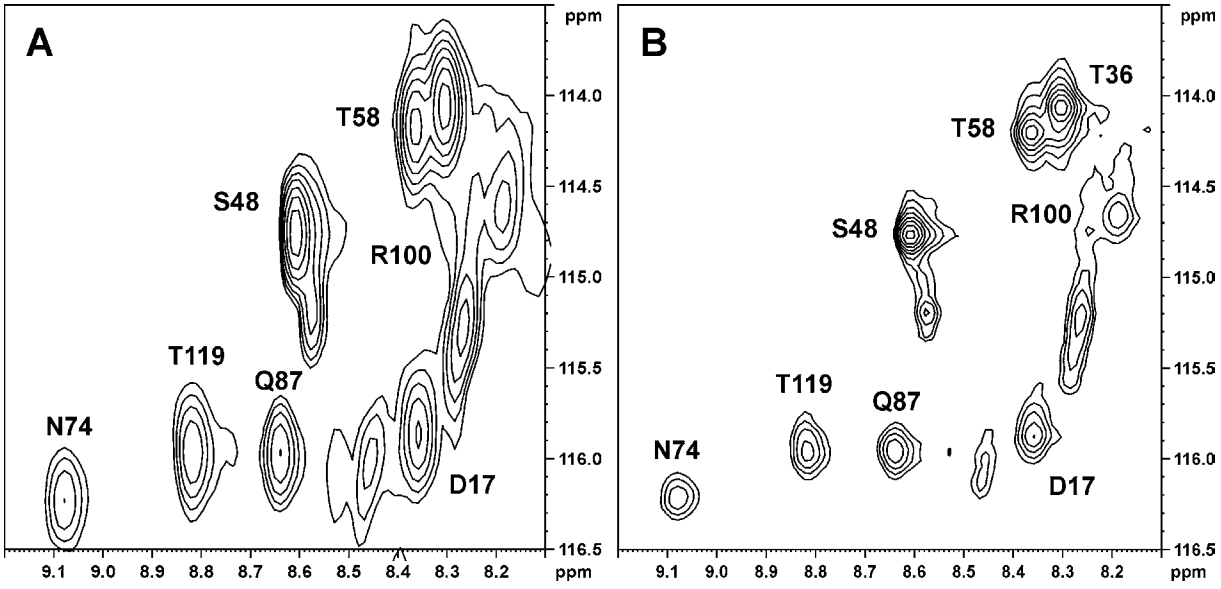


Fig. 2.5.2: Cutouts from ^{15}N-HSQC spectra of αADT at 42 °C and 14.1 T with acquisition times t_{aq} in the ^{15}N-dimension of 32 ms (**A**) and 128 ms (**B**). Ser48 belongs to the most flexible residues in the protein with the smallest ^{15}N-R_2. The respective resonances sharpen and gain in intensity most with increasing t_{aq}.

For $t_{aq} \gg 1/R$ also the signal resolution becomes limited by the intrinsic resonance linewidths rather than t_{aq}. This may be exploited to quickly identify residues with high internal ns-flexibility (low ^{15}N-R_2) from their sharp and intense resonances in a single ^{15}N-HSQC spectrum with high ^{15}N-t_{aq} (**Fig. 2.5.2**). However, care has to be taken in the presence of protein fragments or impurities, since their signals may start to compete with rather broad protein resonances in increasingly resolved spectra.

2.6 Temperature and pH

According to hydrodynamic theory, the rotational correlation time τ_c of a rigid sphere with hydrodynamic radius r_h may be calculated according to:

$$\tau_c = \frac{4\pi r_h^3 \eta}{3kT} \tag{2.6.1}$$

where η is the viscosity of the solvent. With increasing temperature, τ_c in aqueous solution becomes smaller, in parts because of its explicit dependence on T and because the water viscosity is a steadily decreasing function of T (**Fig. 2.6.1**).

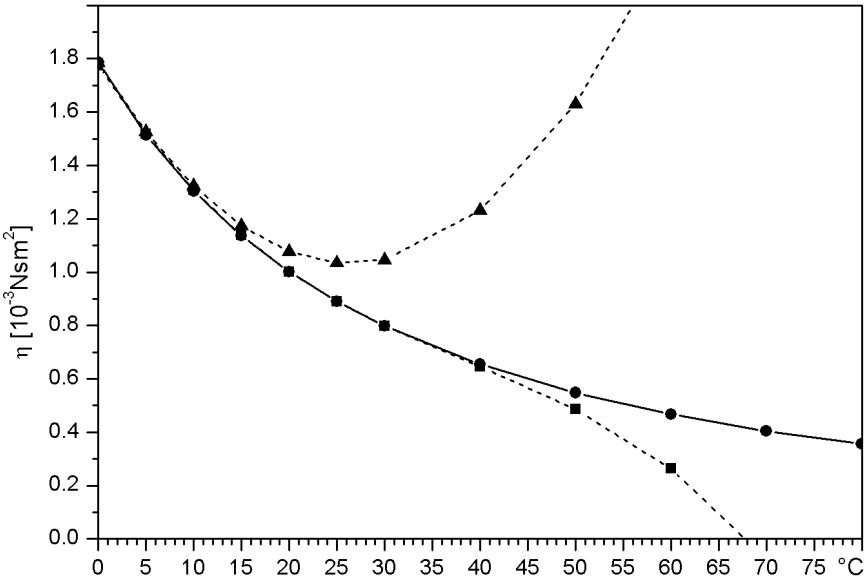


Fig. 2.6.1: Viscosity of H_2O as function of temperature ϑ. Circles with solid line represent experimental values, whereas the dashed lines represent 2nd (triangles) and 3rd (squares) order polynomial approximations: $\eta/10^{-6} = 1775.3 - 56.5\vartheta + 1.0751\vartheta^2 - 0.0092222\vartheta^3$.

As a consequence, resonances generally sharpen and gain in intensity, by far over-compensating unfavourable thermal polarisations and thermal noise in the coil. (**Fig. 2.6.2 A**) However, apart from limitations due to the protein´s thermal stability, the sensitivity for individual residues in ^{15}N-HSQC spectra may be severely affected by temperature dependent hydrogen exchange up to the complete loss of signal (**Fig. 2.6.2 B**).

As discussed in section 1.5, hydrogen exchange rates are strongly pH dependent, and decreasing the pH by one unit compensates an increase in temperature by approximately 30 °C. When lowering the pH is not possible, one needs to clarify whether the loss of some resonances is justified by a significant removal of overlap in others. A further option is the assignment of resonances affected by hydrogen exchange or overlap by a second data set at different temperature.

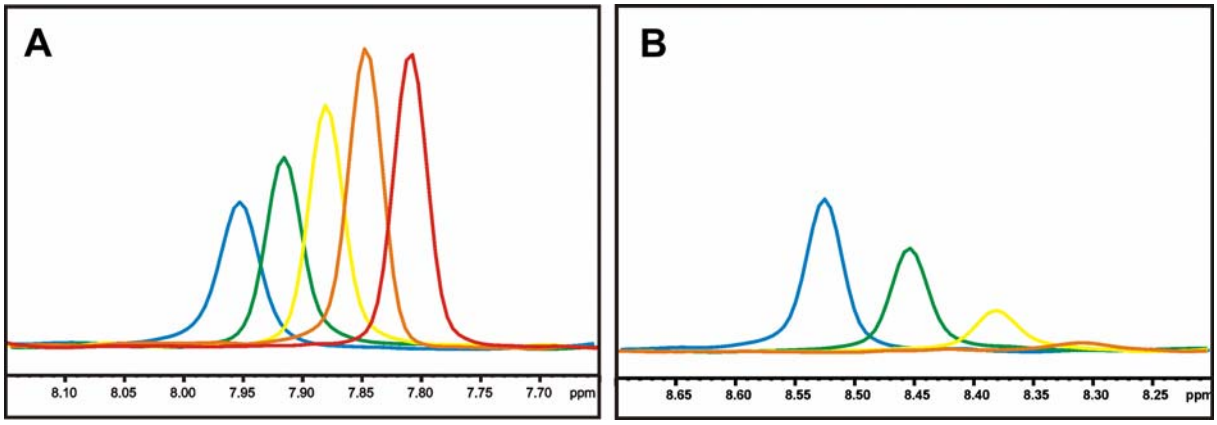


Fig. 2.6.2: 1D slices (^{1}H dimension) of ^{15}N-HSQC spectra of saposin D as a function of temperature: 17 ° (blue), 27 °C (green), 37 °C (yellow), 47 °C (orange) and 57 °C (red). **A:** Signal of Ile75. **B:** Unassigned signal at $\delta(^{15}\mathrm{N}) \approx 115$ ppm, which is strongly subject to hydrogen exchange.

It needs to be emphasised that the comparison of spectra recorded at different temperatures, pH values and solvent conditions requires careful referencing of chemical shifts. Whereas in organic solvents this is conveniently done for ^{1}H and ^{13}C by use of the inert internal standard TMS, analogous water-soluble compounds suffer from a pH- and temperature dependence of their characteristic chemical shift. In this regard, DSS has been found to have the best properties and recommended as universal chemical shift standard.[78] If not available, TSP may be used instead, if its frequency is slightly corrected with respect to DSS (**Fig. 2.6.3**). In order to exclude possible interactions with the biomolecules under study, TSP has been used externally in the present work. Once the zero-ppm ^{1}H frequeny is known for a given magnetic field, chemical shifts of other nuclei are easily referenced using tabulated fixed frequency ratios.[78]

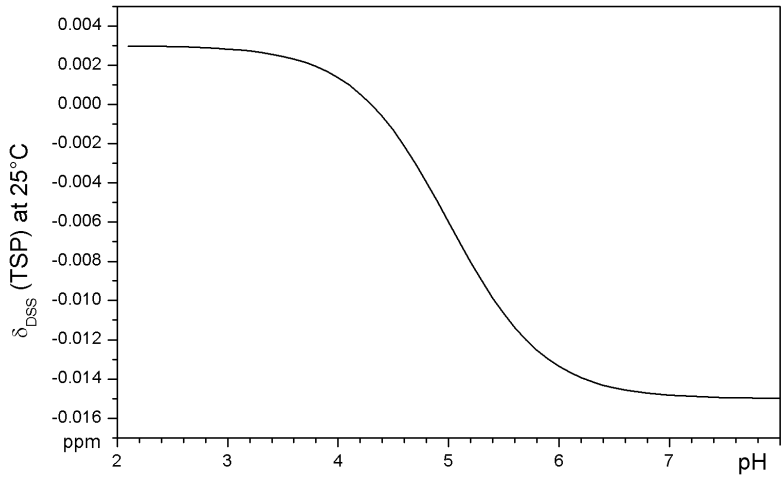


Fig. 2.6.3: Chemical shift of methyl protons of TSP relative to DSS as a function of pH at 25 °C, according to $\delta_{DSS}(\mathrm{TSP}) = -0.019\mathrm{ppm}(1+10^{5.0-\mathrm{pH}})^{-1}$. [78] The methyl protons of DSS are pH-independent in the range 2-11. $\delta_{DSS}(\mathrm{TMS}) = 0.048$ ppm

3 Studies of the apical domains of the thermosome

3.1 The thermosome

Chaperonins are large, barrel-like, double-ring assemblies, which play an essential role in the recognition and ATP-dependent refolding of misfolded substrates, e.g. upon heat shock. Whereas the mechanism of action of the bacterial GroEL/GroES complex as a representative of the so called *Group I* chaperonins is well studied in both biochemical and structural terms,[79] detailed information about the archaeal/eukaryotic *Group II* chaperonins has been gained only in the past decade. The archaeal 930 kDa hexadecameric chaperonin, usually referred to as "thermosome", is composed of two homologous alternating subunits α and β,[80] each forming three distinct domains connected by flexible hinges: An equatorial or ATPase domain, an intermediate and an apical or substrate binding domain.

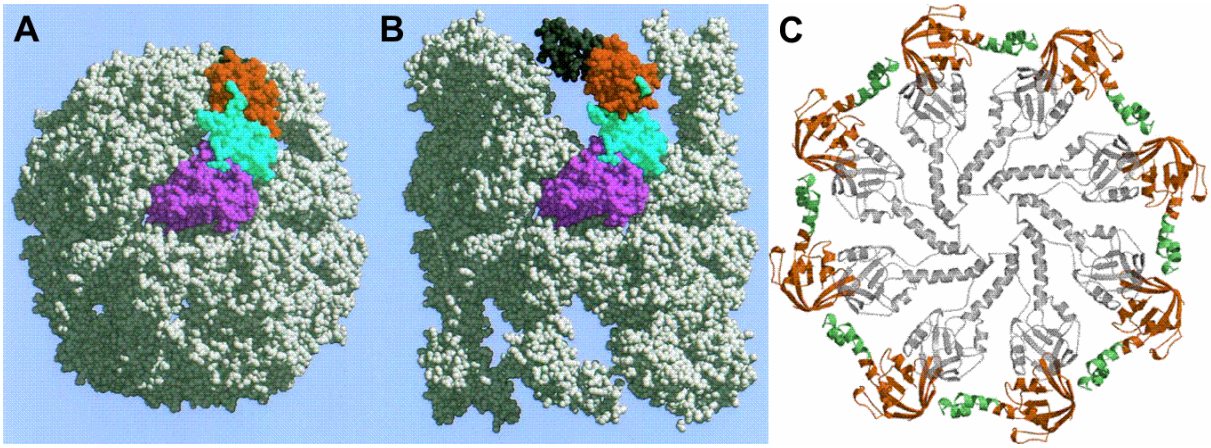


Fig. 3.1.1:A: Closed thermosome as determined by X-ray crystallography.[81] **B:** Open form, modelled from cryo-electron tomography.[82] Equatorial, intermediate and apical domains of one subunit are coloured in purple, cyan and red, respectively, whereas the protrusion is shown in black. **C:** Top view of the iris-type aperture formed by eight apical domains in the closed (grey) and open form (red with green protrusions). Figure taken from [83].

So far, the only structural information of higher resolution about the different mechanistic states in the reaction cycle of the thermosome [84] was obtained from X-ray crystallography and cryo-electron tomography, showing the thermosome in a closed [81] and open [82] conformation, respectively (**Fig. 3.1.1**). An iris-type aperture made up from helical segments, protruding from the apical domains and forming an eight-membered intersubunit β-barrel, was suggested to control access to and exit from the folding cavity.

The helical protrusions, which are completely absent in Group I chaperonins, display a highly conserved pattern of residues in all Group II chaperonins (**Fig. 3.1.2**). Large hydrophobic patches identified in the protrusions of the thermosome are assumed to play some role in substrate binding, although no natural substrates have been identified yet. In contrast, sequence positions in the globularly folded part of the apical domains which are involved in substrate binding in Group I chaperonins [85] are mainly hydrophilic in Group II.

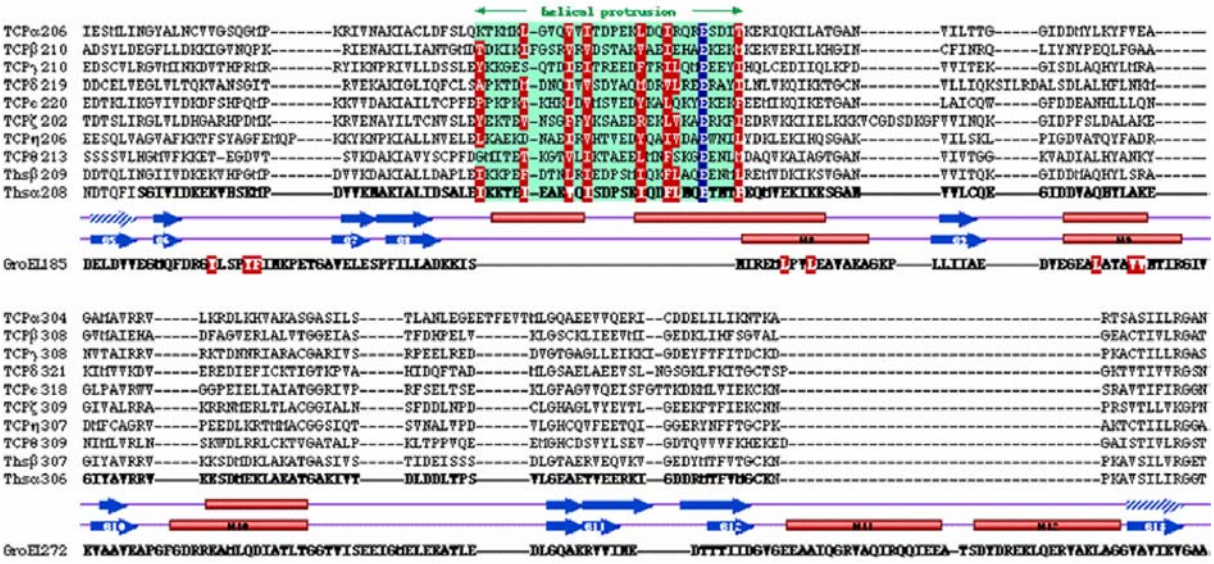


Fig. 3.1.2: Sequence alignment of apical domains of the eight different subunits of the eukaryotic CCT-TCP chaperonin,[86] the α and β subunits of the archaeal thermosome and bacterial GroEL (GroupI). The secondary structures of αADT and GroEL are shown below, with blue arrows representing β-strands and red cylinders α-helices. The protrusion region is shaded green, with conserved hydrophobic residues emphasised in red, and a highly conserved glutamate in blue. The hydrophobic residues which are involved in substrate binding in GroEL [87] are also shaded red, they are mainly hydrophilic in GroupII chaperonins. Figure taken from [88].

To facilitate comparisons between the two apical domains of the thermosome, αADT and βADT, residue numbering for the complete chains will be changed into sequence positions within the ADTs for all following sections. For example Ser214 in the α-subunit is now "Ser2", and Asn215 in the β-subunit "Asn2". Since all preparations of ADTs have been performed in *E.coli*, the first residue is always a methionine.

In contrast, the numbering of secondary structure elements in the following discussion corresponds to the elements found in the closed thermosome. Thus, the β-strands in the apical domains are termed S9-S19, and the α-helices H10-H12.

3.2 Structures of the apical domains

Recently, crystal structures of both ADTs from *Thermosplasma acidophilum* have been solved as isolated 17 kDa constructs consisting of 153 residues, of which 101 are conserved.[88, 89] Not surprisingly, in a globularly folded region, which is made up from a β-sandwich with two orthogonal sheets enclosing a helix (H12), the two structures can be superimposed with a backbone RMSD of only 0.4 Å (**Fig. 3.2.1**). β-strands S9, S10, the first half of S11, S17, S18 and S19 are arranged in an antiparallel β-sheet, in which the βI-hairpin between S18 and S19 adjoints to the intermediate domain. Bulges consisting of two neighbouring hydrophilic side chains pointing towards the surface separate S10 from S11 (Lys20, Asn/Asp21) and S17 from S18 (residues Glu133, Thr/Arg134) respectively. In contrast to that, the second half of S11, S14, S15 and S16 forms a parallel β-sheet in the very core of the apical domain. Against this sheet, two amphiphilic helices are packed, H11 and H10, the latter of which is N-terminally extended to the 20 Å long protrusion. The backside of H10 is stabilised by a long loop (sequentially following H12), which forms part of the inter-subunit interface in the closed complex and is anchored mainly by hydrophobic interactions involving as many as 4 leucines.

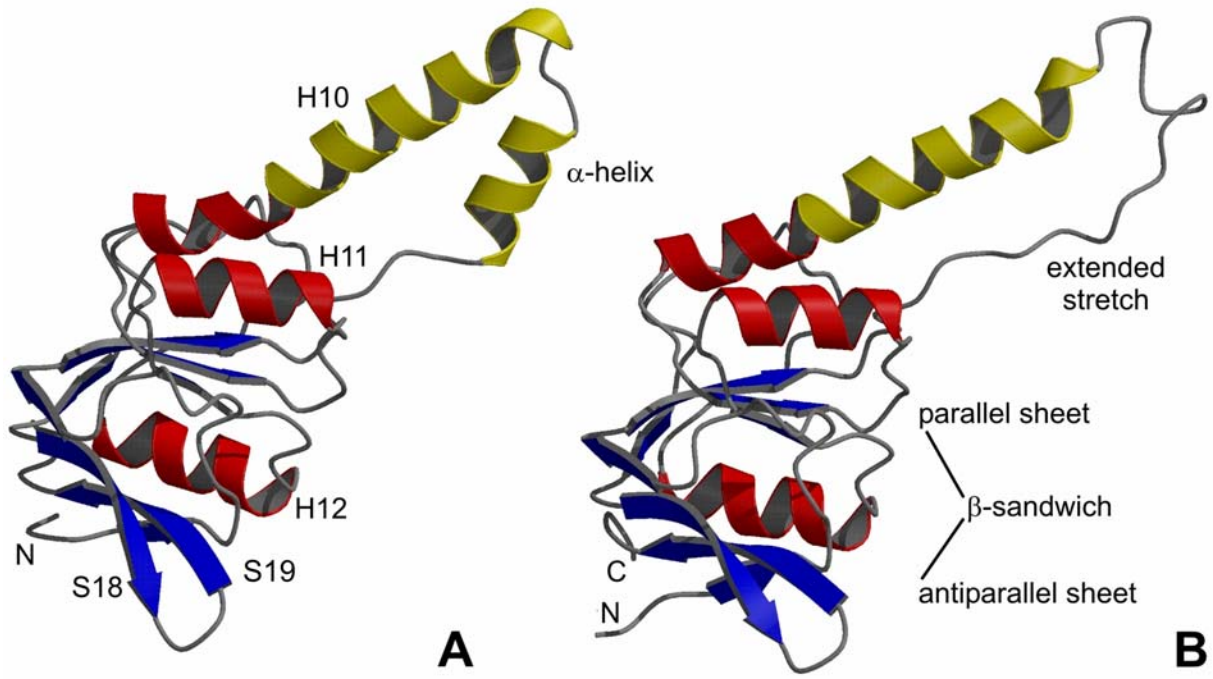


Fig. 3.2.1: Ribbon diagrams of the crystal structures of αADT (**A**, PDB-code 1ASS [88]) and βADT (**B**, PDB-code 1E0R [89]). β-strands and α-helices of the globulary folded parts of the proteins are coloured blue and red, respectively. The protrusion is shown in yellow.

Differences were found in the N-proximal, i.e. outgoing half of the protrusion, and none of the structures resembled the conformation of this region in the closed thermosome. In αADT, after a short stretch of residues (Glu32-Lys34) emerging from the globular part, Lys35-Gln43 form a small α-helix, whereas the respective residues in βADT continue rather extended. The lack of helical winding up makes the βADT protrusion approximately 3-4 Å longer, before both sequences conduct a sharp βI turn (with Pro47 in i+1 position), which merges into the 3_{10}-capping of H10. In the closed thermosome, the whole N-proximal stretch bends off by more than 90° at Phe/Ile38 and the following residues are involved in forming the eightfold β-barrel with residues from the neighbouring subunits.

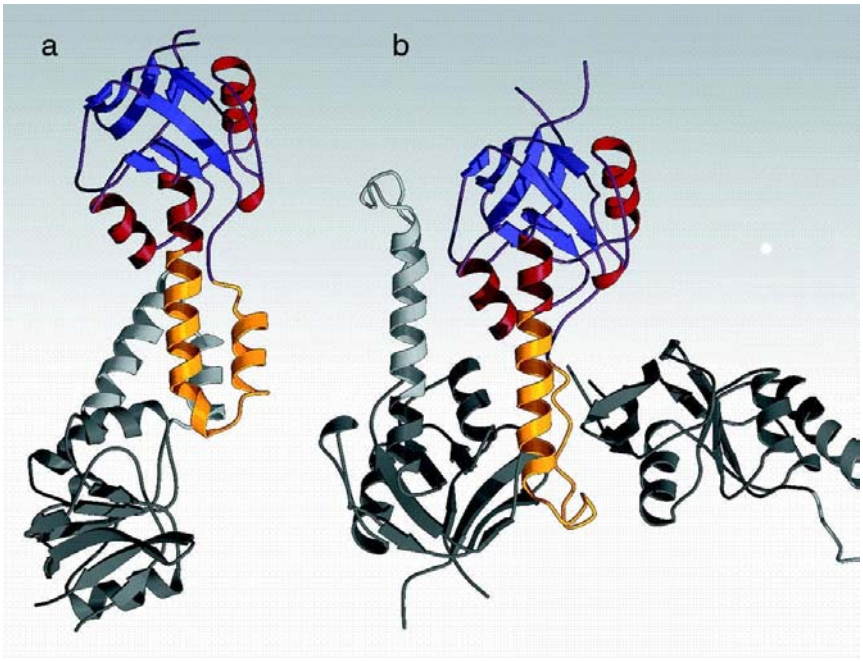


Fig. 3.2.2: Packing of apical domains in their respective crystals. In αADT (**a**), the protrusions of two symmetry-related molecules form a 4-helix bundle, whereas in βADT (**b**), the protrusions are tighly packed against the globular parts of two adjacent molecules. Colour coding is as in Fig. 3.2.1. Figure taken from [89].

Considering the arrangement of the ADTs in their crystal lattice, the above-mentioned conformations in the protrusion may be purely artificially induced by intermolecular packing interactions (**Fig. 3.2.2**), which may energetically compete with intramolecular interactions. The absence of tight packing has been shown to lead to a complete loss of electron density for the N-proximal half of the γCCT protrusion,[90] suggesting it is highly flexible, but without specifying a detailed picture.

With this background, solution NMR spectroscopy provides the most suitable tools to investigate the physiological conformational and dynamic properties of the ADTs, in particular the protrusion regions. Bearing in mind the cryo-electron model with spatially far separated apical domains, the behaviour of isolated domains in solution is likely to represent the situation in the open state of the thermosome.

3.3 ^{15}N-HSQC spectra

Uniformly ^{15}N and ^{15}N,^{13}C labelled samples (≈ 1 mM) of the 153-residue constructs of αADT and βADT were prepared as described elsewhere [91, 92] in the group of W. Baumeister at the MPI for Biochemistry in Martinsried, Germany. The pH (50 mM phosphate buffer) was adjusted to 5.5 for αADT and to 7.8 for βADT, since below pH 7.8 βADT (pI 6.3) was subject to aggregation. ^{15}N-HSQC resolution was only satisfactory at a temperature of 42 °C, where sample lifetimes were approximately 2-3 weeks, during which ^{15}N-HSQC spectra especially for αADT heavily deteriorated (**Fig.3.3.1**).

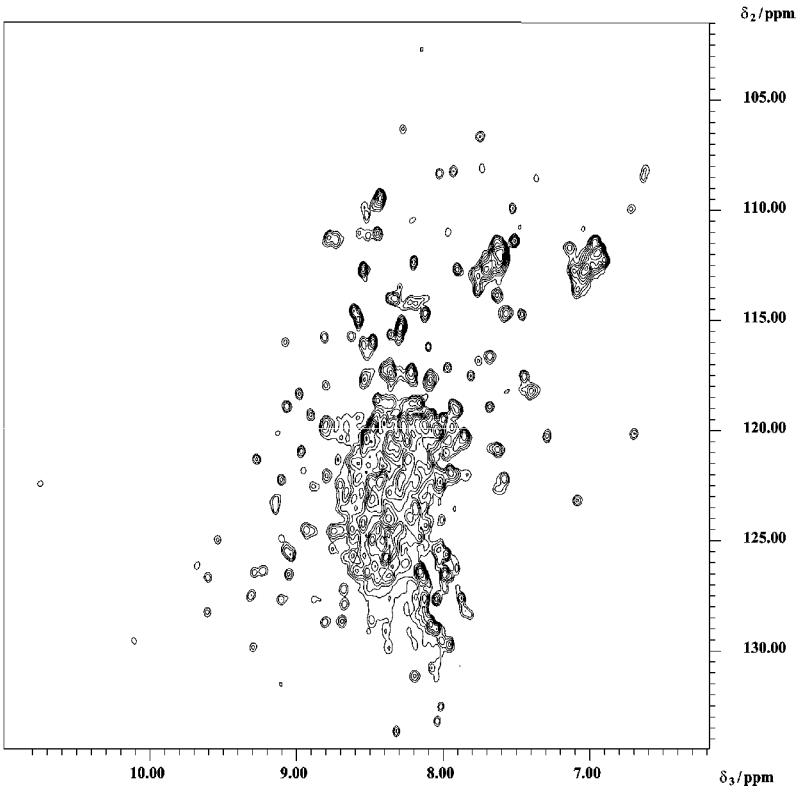
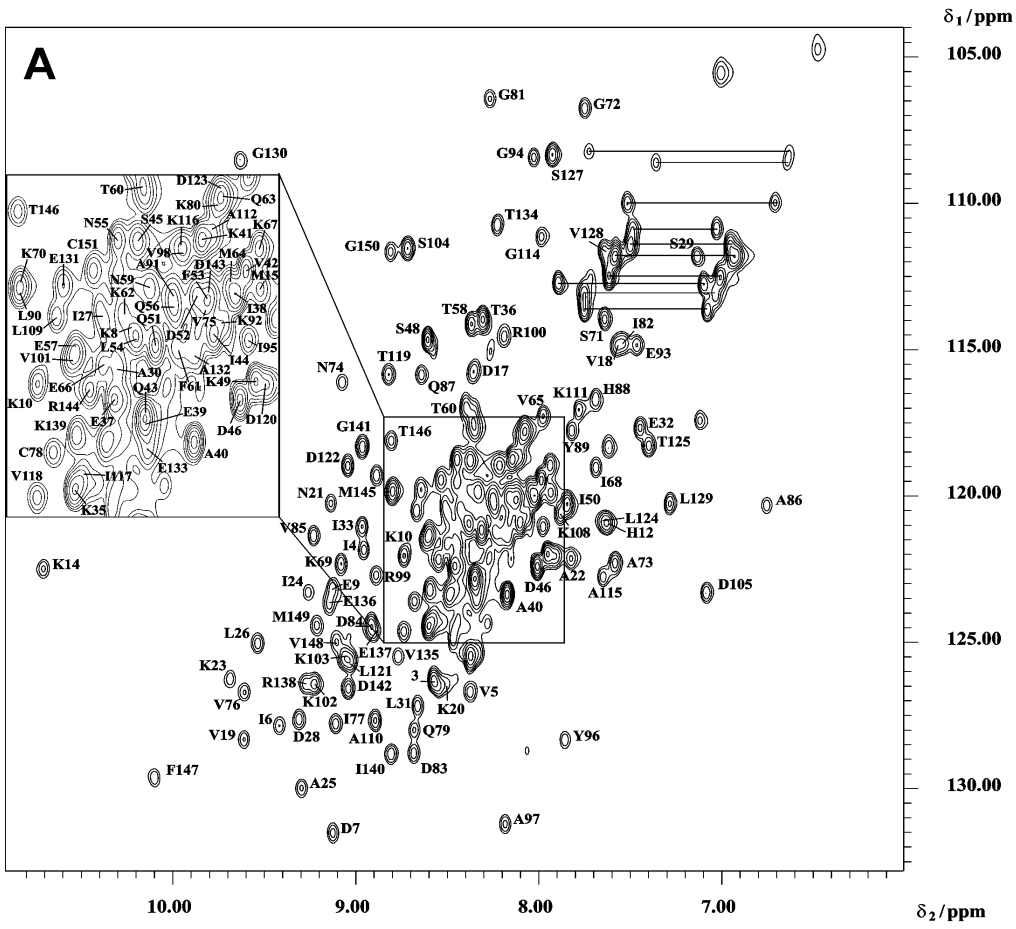
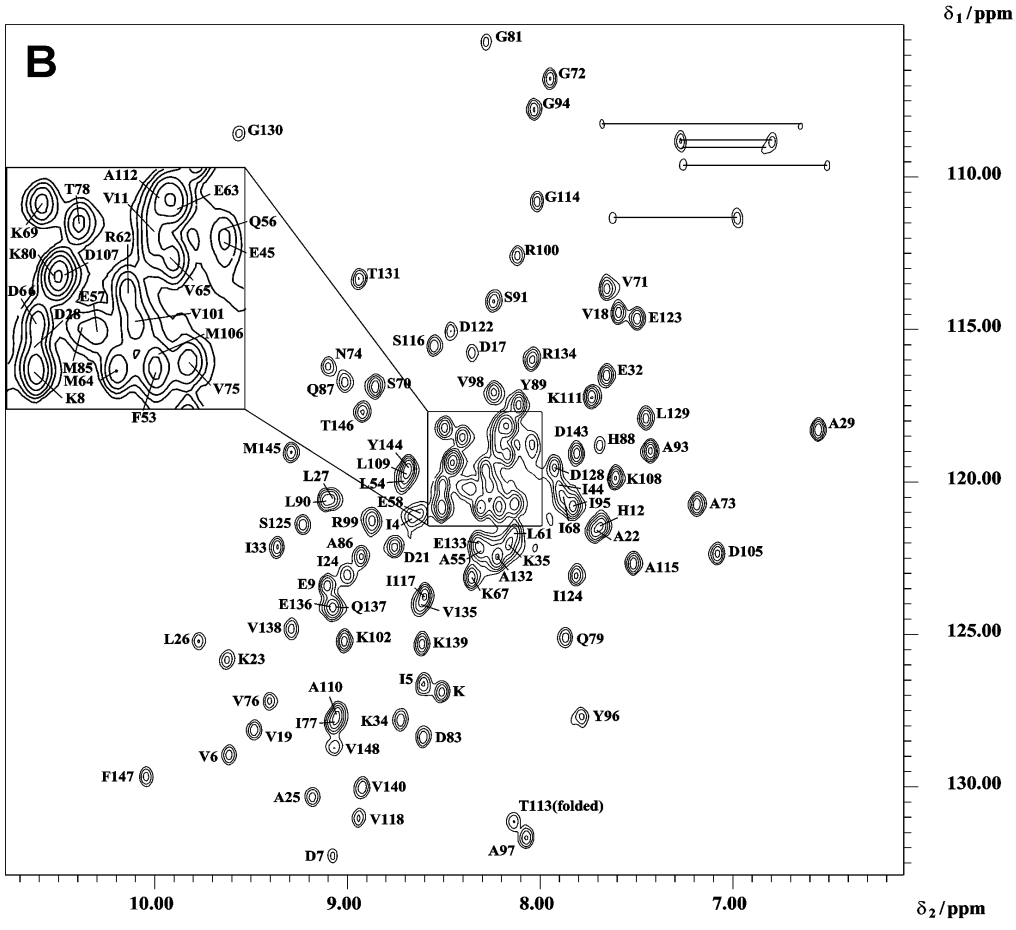


Fig. 3.3.1: ^{15}N-HSQC spectrum of αADT after ~3 weeks at 42 °C. All NMR experiments were performed in 90 % H_2O and 10 % D_2O in Shigemi tubes at 42 °C and 14.1 T.

Fig. 3.3.2 (following page): Annotated ^{15}N-HSQC spectra of αADT (**A**, pH 5.5) and βADT (**B**, pH 7.8) at 42 °C and 14.1 T. The regions of strong overlap are expanded for the purpose of clarity. Side chain NH_2 groups of asparagine and glutamine residues are connected by horizontal lines.

At these elevated temperature, also fast solvent exchange at pH 7.8 became a major issue, as apparent from the strongly reduced number of visible amide resonances for βADT compared to αADT (**Fig. 3.3.2**). Using the methodology described in the following section, in the α-subunit domain all except the three N- and two C-terminal of the 150 non-proline resonances could be assigned, but only 110 out of 148 non-prolines in the β-subunit homologue. For example, the conserved residue Lys14, which gives rise to a signal with exceptional downfield H^N shift, was only observed in βADT upon lowering the temperature. Almost identical resonance positions of conserved residues already provide strong evidence for structural similarity, as it was seen in the crystal structures.


A
δ1/ppm
105.00
110.00
115.00
120.00
125.00
130.00
10.00
9.00
8.00
7.00
δ2/ppm


B
δ1/ppm
110.00
115.00
120.00
125.00
130.00
10.00
9.00
8.00
7.00
δ2/ppm
T113(folded)

3.4 Backbone assignment

The concepts of coherence transfer and heteronuclear shift-editing discussed in the previous chapter may be readily extended to a third and fourth dimension and include both carbon and nitrogen nuclei, provided they are simultaneously isotopically enriched in the molecule. In proteins, this allows sequential connectivities to be established by using exclusively the network of relatively large and structure-independent one bond ^1J-couplings with relatively short transfer times in the backbone and side chains (**Fig. 3.4.1**).[93]

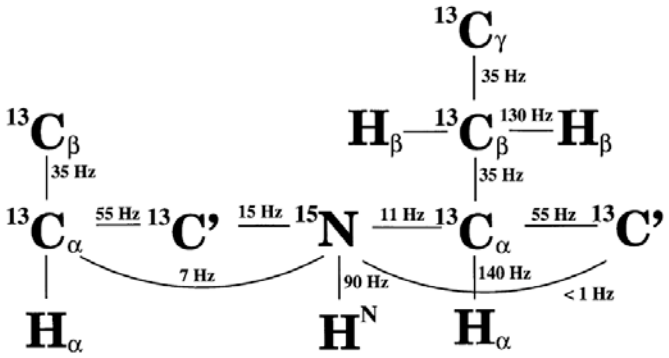


Fig. 3.4.1: Typical values of one bond ^1J-couplings between ^{1}H, ^{13}C and ^{15}N nuclei in protein backbone. Figure taken from [94].

For medium-sized (15-30 kDa) proteins like αADT and βADT, the most efficient and therefore most commonly used 3D triple resonance schemes for backbone assignment use the final detection of ^{15}N (ω_2) and H^N (ω_3) nuclei, i.e. a ^{15}N-HSQC-type projection.[94] In the remaining (ω_1) dimension, magnetisation is evolved with respect to the frequency of C′, C^α, C^β, H^α and H^β nuclei of residues (i) and/or (i–1) (**Table 3.4.1**). Different transfer periods and transverse relaxation behaviour of the spins involved are responsible for the varying sensitivity of the individual experiments and may favour an *out and back* ("HN...") or *transfer* ("...NH") design. Above 20-25 kDa, TROSY variants of these schemes, in combination with fractional or full deuteration, may be used.[95]

For the sequential assignment of αADT and βADT, PASTA-[96] pseudo residue lists were generated by picking approximately 200 resonances in ^{15}N-HSQC spectra with moderate ^{15}N acquisition times (32 ms) of the fresh doubly labelled samples. Additional ^{15}N-HSQC spectra with inverted phase for NH_2 groups and with large ^{15}N acquisition times (128 ms) helped to identify asparagine and glutamine side chain resonances and in case of resonance overlap. After recording the experiments indicated in **Table 3.4.1**, ω_1-frequencies were transferred from the individual spectra to the program PASTA and corrected manually.

Table 3.4.1: Common ^{15}N-H^N detected triple resonance experiments for the sequential assignment of backbone resonances of medium-sized proteins. The experiments used for αADT and βADT are indicated in bold, using spectral windows of –1.5-10.9 ppm for H^N and H^α/H^β (in HNHA and HNHB), 101.5-134.5 ppm for ^{15}N, 170-180 ppm for C´, 42-68 ppm for C^α, 10-76 ppm for C^α/C^β, and 1-6.5 ppm for H^α/H^β (in HBHA(CO)NH) nuclei. Experiments were recorded at 42 °C and a magnetic field of 14.1 T with 72*36*512 complex data points, 16 transients and total acquisition times of 2-2.5 d each.

res.	C´	C^α	C^β	H^α	H^β
(i–1)	**HNCO**	**HNCA**	**HNCACB**	HN(CA)HA	**HBHA(CO)NH**
	HN(CA)CO	**HNCACB**	**CBCA(CO)NH**	HN(COCA)HA	
		HN(CO)CA		**HBHA(CO)NH**	
		CBCA(CO)NH			
(i)	HN(CA)CO	**HNCA**	**HNCACB**	**HNHA**	**HNHB**
		HNCACB		HN(CA)HA	

PASTA was then employed to globally optimise sequential connections based on penalties for non-matching pairs of shifts. C^α(i–1) occured in both HNCA and CBCA(CO)NH, and the remaining resonances in these spectra were assigned to C^α(i) (usually the stronger one) and C^β(i–1), respectively. The low- and high frequency resonances of HBHA(CO)NH were assigned to H^α(i–1) and H^β(i–1) except for serine and threonine residues, whereas H^α(i) and H^β(i) shifts were known from HNHA and HNHB spectra on ^{15}N-labelled samples. Care had to be taken in some cases where a close proximity of $H^{\beta2}$ and $H^{\beta3}$ was resolved in HBHA(CO)NH, but gave rise to a single maximum at average frequency in HNHB due to its larger ω_1 spectral width.

Measurement of C^β(i) shifts turned out to be most critical as it relied on the less sensitive HNCACB, a modification of HNCA, in which two $^1J_{C\alpha C\beta}$-evolution periods on the fast relaxing C^α spin are inserted on either side of t_1. A simulation of the transfer amplitude as a function of the length of this period shows that the corresponding delays have to be chosen very carefully (**Fig. 3.4.2 A**). Still, C^β(i) could be identified only in about half of the residues, fortunately often in flexible protein regions, where the dispersion of C^β(i) is actually required to remove sequential ambiguities due to the reduced dispersion of other nuclei. For residues with known C^α(i) and H^α(i) shifts, assignment of C^β(i) was completed by recording an additional (H)CCH-COSY experiment.

Similarly, C´(i) frequencies, which were found to be directly inaccessible due to the low inherent sensitivity of HN(CA)CO, were measured indirectly using HCACO.[97] In this experiment, different maxima for $^1J_{C\alpha C'}$-transfer (**Fig. 3.4.2 B**) were chosen for out and back steps, because the latter was combined with a constant time evolution of the C^α frequency.

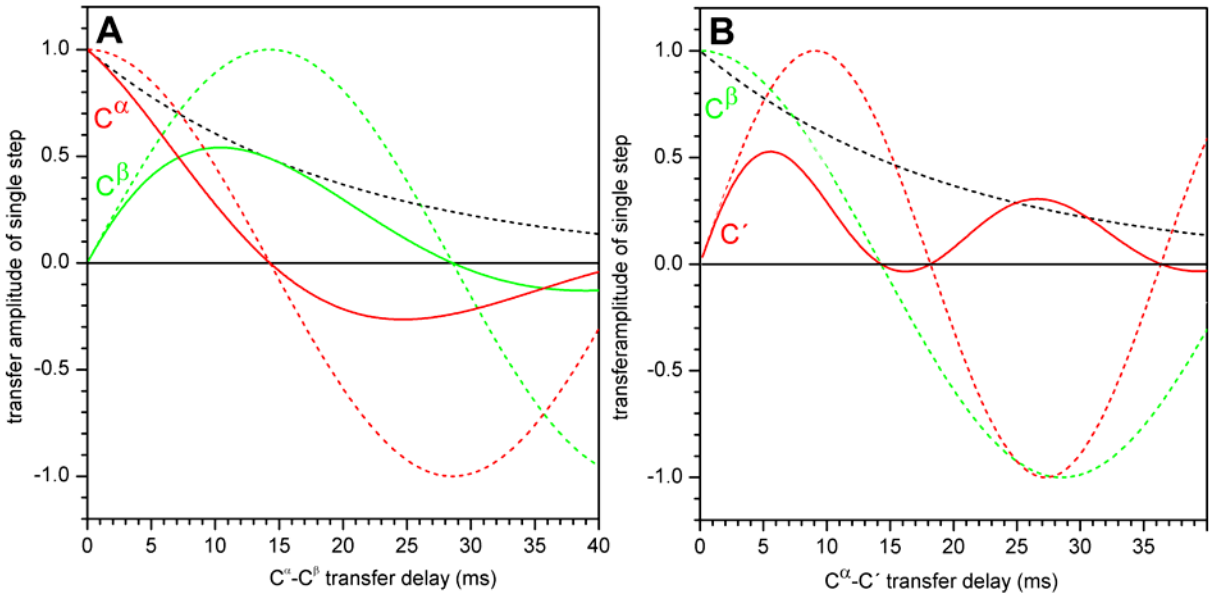


Fig. 3.4.2: Transfer amplitudes for coherence transfer from C^α to C^β (**A**) and to C´ (**B**), assuming an average C^α transverse decay (dashed black lines) with R_{av} = 50 s^{-1}, which is typical for the size of the ADTs. In **A,** the sine- (transferred magnetisation) and cosine (retained magnetisation) oscillations ($^1J_{C\alpha C\beta}$ = 35 Hz) are indicated with green and red dashed lines, respectively. Relaxation corrected curves are shown as solid lines. At the C^β transfer optimum (≈ 11 ms), retained C^α magnetisation is weak, whereas at a delay of 7 ms, which was chosen for the HNCACB, both nuclei are expected to have equal intensity. Note that for out and back transfer, the shown amplitudes have to be squared, giving an overall sensitivity of approximately 25 % compared to HNCA. In **B**, $^1J_{C\alpha C\beta}$ is now passive and leads to an undesired modulation indicated as dashed green line. The active transfer oscillation ($^1J_{C\alpha C'}$ = 55 Hz) is shown in dashed red, and the relaxation corrected product in solid red. Chosen delays for out and back in HCACO were 6 and 27 ms, respectively.

Amide-detected triple resonance experiments have the severe drawback that residues with fast exchanging amide protons escape this ordinary sequential assignment procedure, as obvious from the lack of ^{15}N-HSQC signals for 38 residues in βADT mentioned above. The application of alternative CH-detecting experiments providing sequential information (like HCAN and HCA(CO)N) to βADT, failed, however, due to their low inherent sensitivity and resolution. Backbone resonance assignments have been published [92] and deposited in the BioMag-ResBank (http://www.bmrb.wisc.edu) under BMRB accession numbers 5930 (αADT) and 5936 (βADT) and are therefore not listed in chapter 9.

3.5 Chemical shifts

Since experimental chemical shifts δ_{exp} of H^{α}, C^{α}, C^{β} and C′ nuclei are sensitive to the local peptide conformation, they may be used to identify secondary structure elements.[98] ^{15}N chemical shifts are less frequently used for this analysis, since they have been shown to strongly depend on the neighbouring residues.[99] Provided that shifts have been calibrated carefully (here done with external TSP for ^{1}H and indirect referencing for the other nuclei),[78] this is usually accomplished by first calculating the *secondary* chemical shift δ_{sec}:

$$\delta_{sec}(i) = \delta_{exp}(i) - \delta_{rc}(i) \qquad (3.5.1)$$

where δ_{rc} are tabulated *random coil* values from small peptide models.[100] Sequence plots of $\delta_{sec}(i)$ in αADT and βADT are shown in **Fig. 3.5.2** for H^{α}, C^{α} and C′. If this number exceeds or falls below a certain interval around δ_{rc}, a *chemical shift index* (CSI) (+1) or (−1) is assigned to that particular nucleus, and helices/strands are identified from opposite CSI being constant over at least three to four residues.[101] Finally, individual values for H^{α}, C^{α}, C^{β} and C′ nuclei are combined to the *consensus* CSI, which is shown in **Fig. 3.5.1** for αADT. In the globular part of the protein, an excellent agreement with the secondary structure elements in the crystal is obtained, proving the protein adopts the same fold in solution.

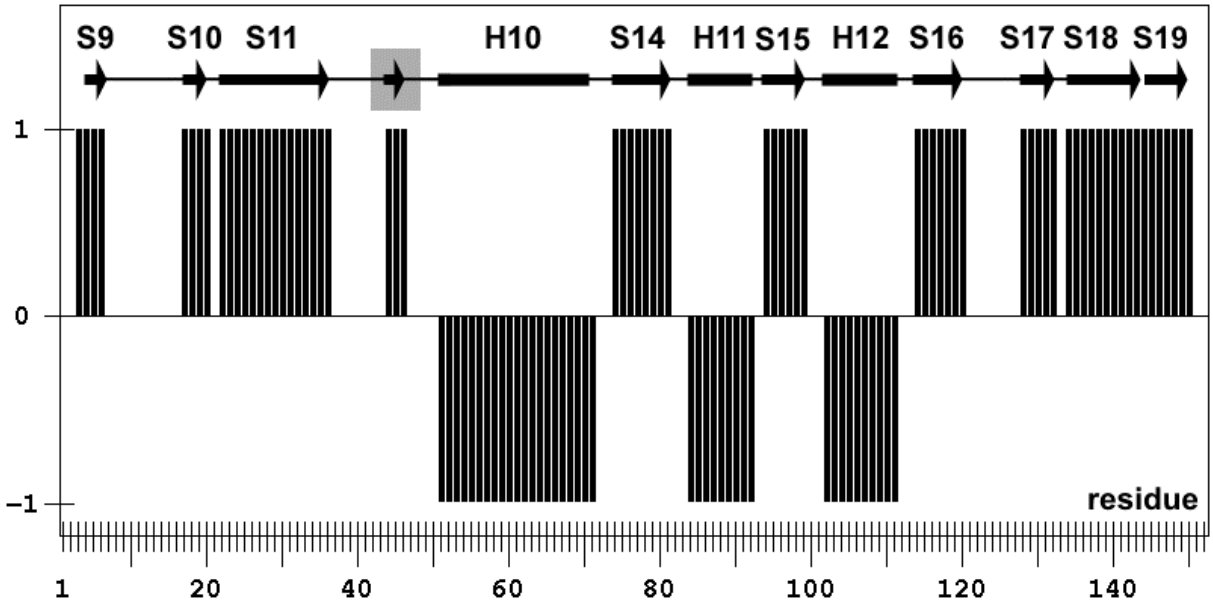


Fig. 3.5.1: Consensus CSI for αADT with secondary structure elements from the crystal (PDB-code 1ASS) shown above. Instead of the small helix Lys35-Gln43, the CSI indicates an extended stretch for residues Ile44-Asp46, which is marked by a grey-shadowed arrow.

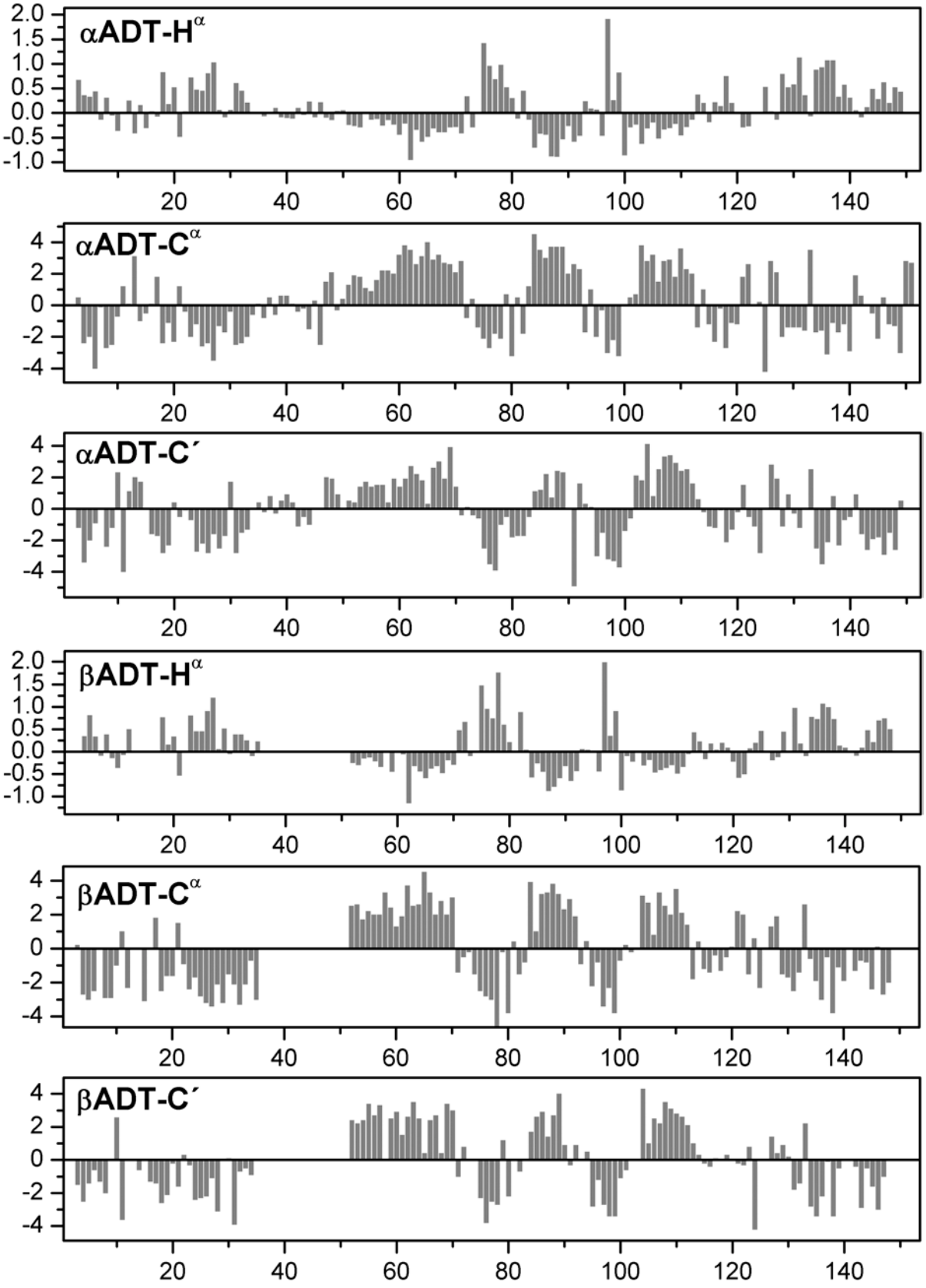


Fig. 3.5.2: Sequence plot of secondary chemical shifts $\delta_{sec}(i)$ of H^{α}, C^{α} and C´nuclei of αADT and βADT (random coil shifts taken from [100]). Significantly negative values for H^{α}, and simultaneously positive values for C^{α} and C´ over a stretch of at least 4 residues indicate α-helices. Residues in β-strands show the opposite behaviour.

As apparent from **Fig. 3.5.2** and **Fig. 3.5.3**, the secondary chemical shifts among both homologues are highly correlated, leading to virtually coinciding absolute shifts for conserved residues. This phenomenon is expected for proteins sharing the same fold and has recently been exploited for structural predictions.[102]

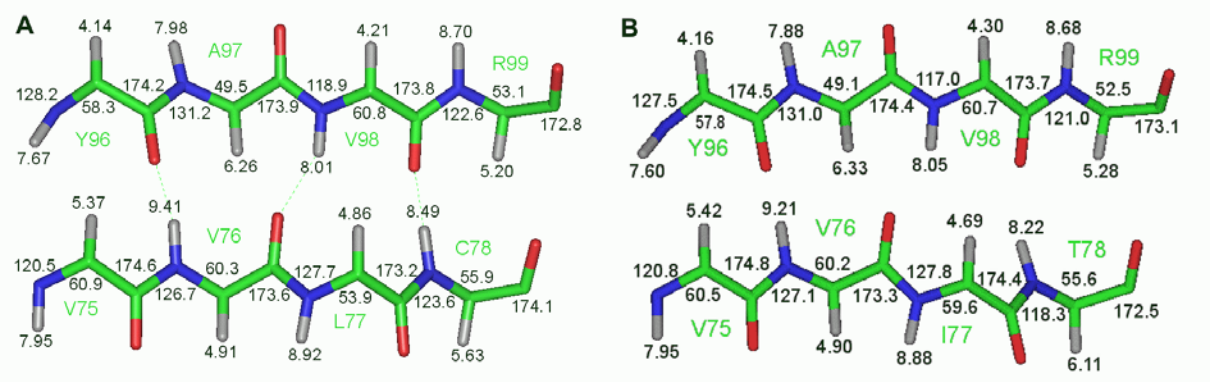


Fig. 3.5.3: Backbone stick diagrams of strands S14 (bottom) and S15 (top) in the central parallel β-sheet of αADT (**A**) and βADT (**B**). Carbon, nitrogen, oxygen and hydrogen atoms are represented in green, blue, red and grey, respectively, and labelled with their experimental chemical shifts.

In both proteins, the H^{α} proton of residue Ala97 resonates at extremely low field, and a database search in the BMRB [103] showed that $\delta_{sec} > 1.7$ for H^{α} occurs almost exclusively in parallel β-sheets. Owing to steric restrictions, the geometry of H-bonds between amide protons and the carboxyl group of the opposite strand is less favourable than in antiparallel sheets. In the ADTs, S14 is slightly shifted with respect to S15, thereby shortening the H-bonds Val76H^{N}-Tyr96CO and Cys/Thr78H^{N}-Val98CO (**Fig. 3.5.3**). As a result, Val76CO is moved away from its H-bonding donor Val98H^{N} and approaches Ala97H^{α}. Considering, that the conformations of residues in this core of the protein are likely to be the same in solution, the electronic influence of this carbonyl may explain the unusually high frequency of Ala97H^{α}. Theoretical considerations also suggested the presence of a large $^{3h}J_{C\alpha C'}$ coupling (across H^{α}), which have recently been shown to reach about half the size of $^{3h}J_{NC'}$.[104]

Given the obviously high degree of structural correlation between the globular parts of the two ADTs (and to the crystals) and the lack of assignment in the βADT protrusion it was decided to proceed with a further backbone-conformational analysis only for αADT. For the same reason, side-chain assignments and a complete structure determination was not envisaged, since it would have required unreasonable amounts of doubly labelled protein.

3.6 Scalar couplings $^3J_{HNH\alpha}$

In the HNHA experiment, the intensities I_d and I_c of diagonal ($\omega_1 = \omega_3$, H^N) and cross (H^α) peaks of residue (i) are modulated with the sine and cosine of the coupling constant $^3J_{HNH\alpha}(i)$, which evolves during the transfer delay τ (square for out and back):[105]

$$\frac{I_c(i)}{I_d(i)} = \frac{\sin^2(\pi {}^3J_{HNH\alpha}(i)\tau)}{\cos^2(\pi {}^3J_{HNH\alpha}(i)\tau)} \quad \Leftrightarrow \quad {}^3J_{HNH\alpha}(i) = \frac{\tan^{-1}\sqrt{\frac{I_c(i)}{I_d(i)}}}{\pi\tau} \qquad (3.6.1)$$

The coupling constants determined this way for αADT were scaled with another factor 1.1 to account for the different relaxation behaviour of H^N and H^α nuclei (**Fig. 3.6.1**). They were compared to values back-calculated from the backbone angle Φ(i) in the crystal structure (1ASS) via the Karplus relation:[106]

$$^3J_{HNH\alpha}(i) = 6.4\cos^2(\Phi - 60°) - 1.4\cos(\Phi - 60°) + 1.9 \qquad (3.6.2)$$

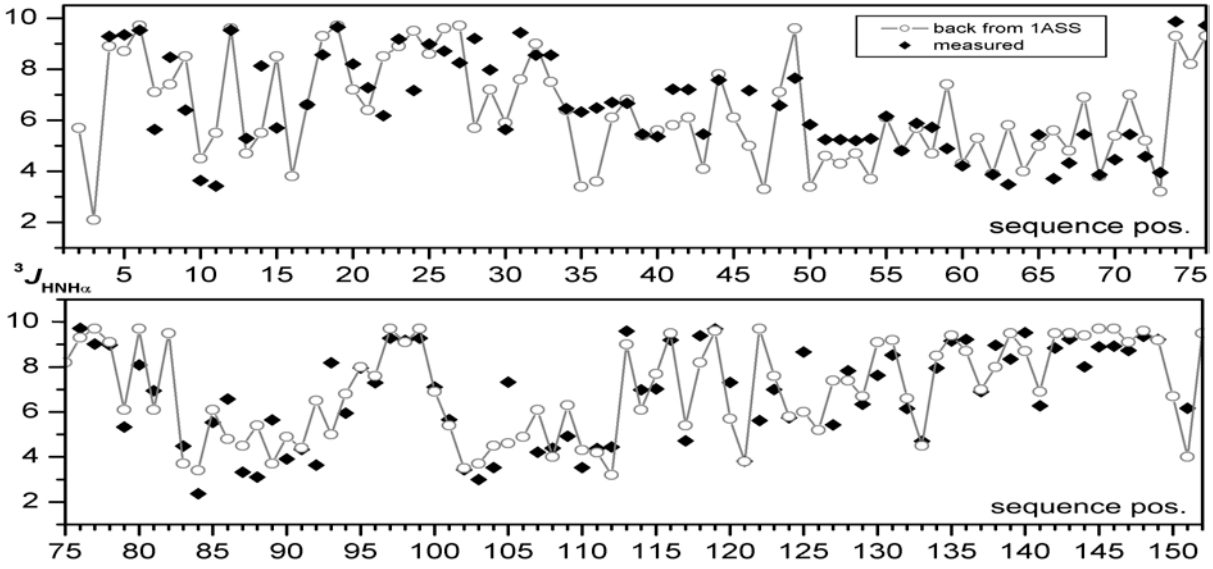


Fig. 3.6.1: Plot of $^3J_{HNH\alpha}$ in αADT from HNHA (diamonds) and back-calculated from 1ASS (circles and line). Values above 8 Hz and below 6 Hz indicate extended and helical conformation of the residues.

A good overall correlation, albeit with local deviations, is obtained. The amphiphilicity of the second half of H10 (Gln63-Ala73) and H11 (Asp84-Glu93) is visible from the helical periodicity of $^3J_{HNH\alpha}$, whereas H12 (Lys103-Thr113), which is sandwiched between the two β-sheets, is not amphiphilic. In residues Lys35-Gln43, experimental couplings are higher than predicted from the small helix in the crystal, and like in the residues following up to Ile50, they are typical of unstructured peptides.

3.7 Backbone NOE pattern

The pattern of sequential backbone NOEs $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$ in αADT was analysed from 3D single (*HNH*-)[107] and doubly (*NNH*-)[108] ^{15}N-edited NOESY spectra. In β-strands $d_{\alpha N}(i,i+1) \approx 2.2$ Å is usually short, giving rise to a strong NOE correlation, whereas in α-helices $d_{NN}(i,i+1) \approx 2.8$ Å is short and several intermediate range NOEs up to $d_{\alpha N}(i,i+4)$ and $d_{NN}(i,i+4)$ may be observed. Secondary structure elements of αADT, identified this way, are shown in **Fig. 3.7.1**. The topology, i.e. the 3D arrangement of β-strands was established from 19 long range $d_{\alpha N}(i,j)$ and 13 long range $d_{NN}(i,j)$ correlations.

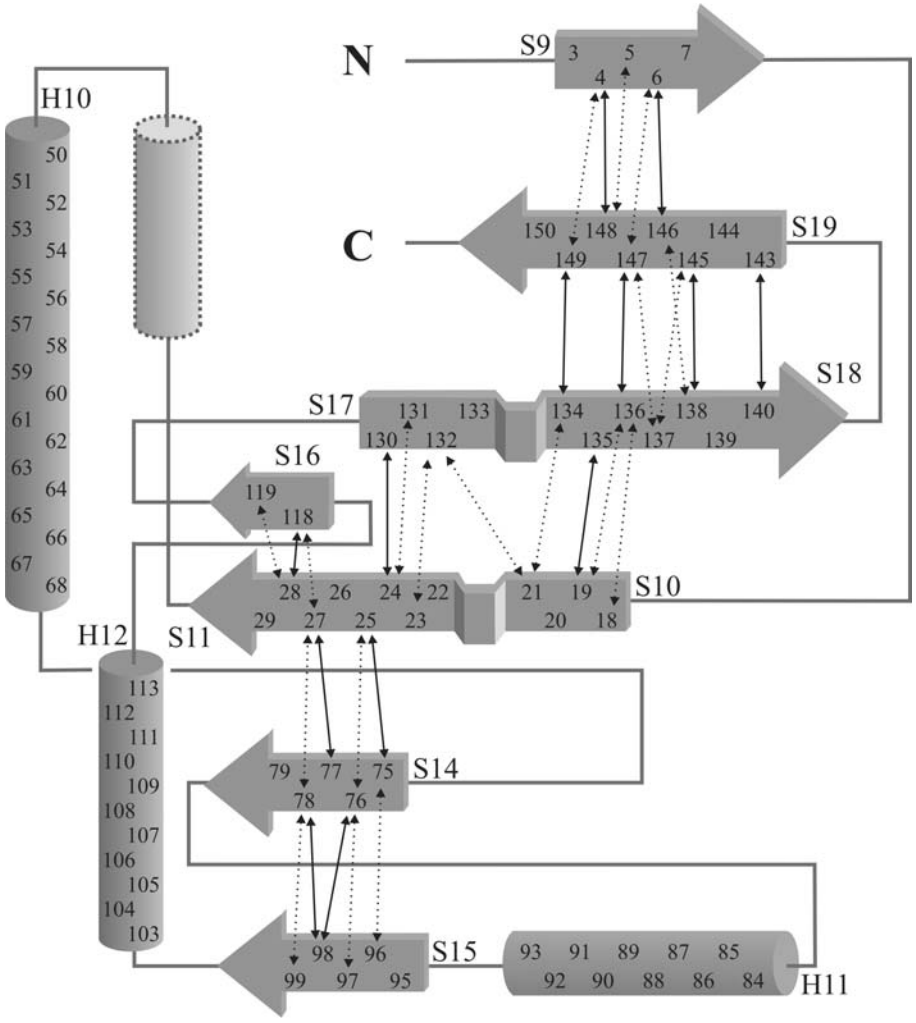


Fig. 3.7.1: Topology diagram of aADT as constructed from short, intermediate, and long range ($d_{\alpha N}(i,j)$ as dotted arrows and $d_{NN}(i,j)$ as solid arrows) backbone NOE correlations. Strands and helices are shown as arrows and cylinders, respectively. S10 and S11 as well as S17 and S18 are separated by β-bulges with two hydrophilic side chains (Lys20/Asn21 and Glu133/Thr134) on the same side of the sheet. The dotted helix Lys35-Gln43 is only found in the crystal.

As expected, and in agreement with chemical shifts and $^3J_{HNH\alpha}$ data, the first two residues (Glu32-Ile33) protruding from the globular core, and less pronounced also Lys34 and Lys35, show the characterisitics of an extended conformation. For the following residues (as for almost all residues in the protrusion), $d_{NN}(i,i+1)$ NOE correlations are visible, but estimated cross relaxation rates were much lower than normally observed in α-helices. Since also the secondary shifts are close to zero, it can be concluded that the small α-helix Lys35-Gln43 found in the crystal is not significantly populated in solution.

In contrast, there is some evidence that Ile44-Asp46 forms a rather extended stretch, which was also observed in the consensus CSI (**Fig. 3.5.1**). For residues Gln51-Thr60 the secondary structure indicators indeed predict a helix, albeit less pronounced than for the residues which follow, or for the other helices. This may be interpreted as partial fraying of the N-terminal half of H10 by fast helix-coil transitions, thereby averaging chemical shifts, coupling constants and cross relaxation rates. As in the crystals, residues Asp46-Ile50 form a βI-turn (induced by Pro47), which could be modelled from the NOE data (**Fig. 3.7.2**). Possibly, this turn confers some rigidity to the otherwise rather unstructured N-proximal half of the protrusion, which may be necessary for substrate binding.

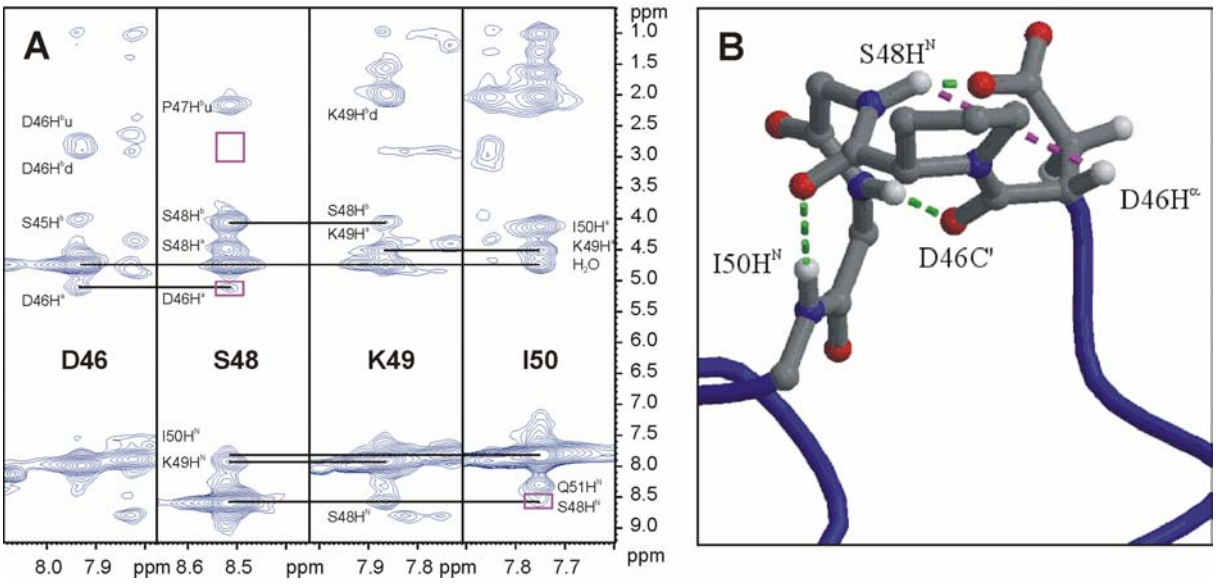


Fig. 3.7.2: A: ω_1-strip plots for residues Asp46-Ile50 (residue 47 is a proline) in αADT taken from HNH-NOESY. Sequential NOE correlations are marked with horizontal solid lines, additionally the correlations Ser48H^N-Ile50H^N and Asp46H$^\alpha$-Ser48H^N, but not Asp46H$^\beta$-Ser48H^N can be seen (purple box). **B:** βI-turn at the tip of the protrusion in ball-and-stick, modelled from this NOE data, with carbon, nitrogen, oxygen and hydrogen atoms shown in grey, blue, red and white, respectively. The H-bonding pattern is shown as green dashed lines. The side chain of Asp46 is turned with respect to the crystal and is likely to form an H-bond with Ser48H^N.

3.8 Amide exchange

The 28 residues from Ile33 through Thr60 of the protrusion suffer most from the aforementioned assignment problems under basic (pH 7.8) conditions and 42 °C. Indeed, 18 out of 38 unassigned residues, i.e. those with high HX rates, in βADT are located in the protrusion region, pointing out extreme solvent exposure (see section 1.7). In αADT at pH 5.5, where intrinsic hydrogen exchange is slower by a factor of 200 and k_{int} is on the order of a few s^{-1}, all protrusion backbone amide resonances were visible, albeit with poor chemical shift dispersion. HX rates were obtained by the NewMEXICO experiment (sections 1.9 and 7.8)[44] for the whole protein except for residues affected by ^{15}N-HSQC overlap. **Fig. 3.8.1** shows a good correlation between k_{ex} in αADT and unobservable residues in βADT, but also a correlation with amides which are not H-bonded in the crystal structure of αADT.

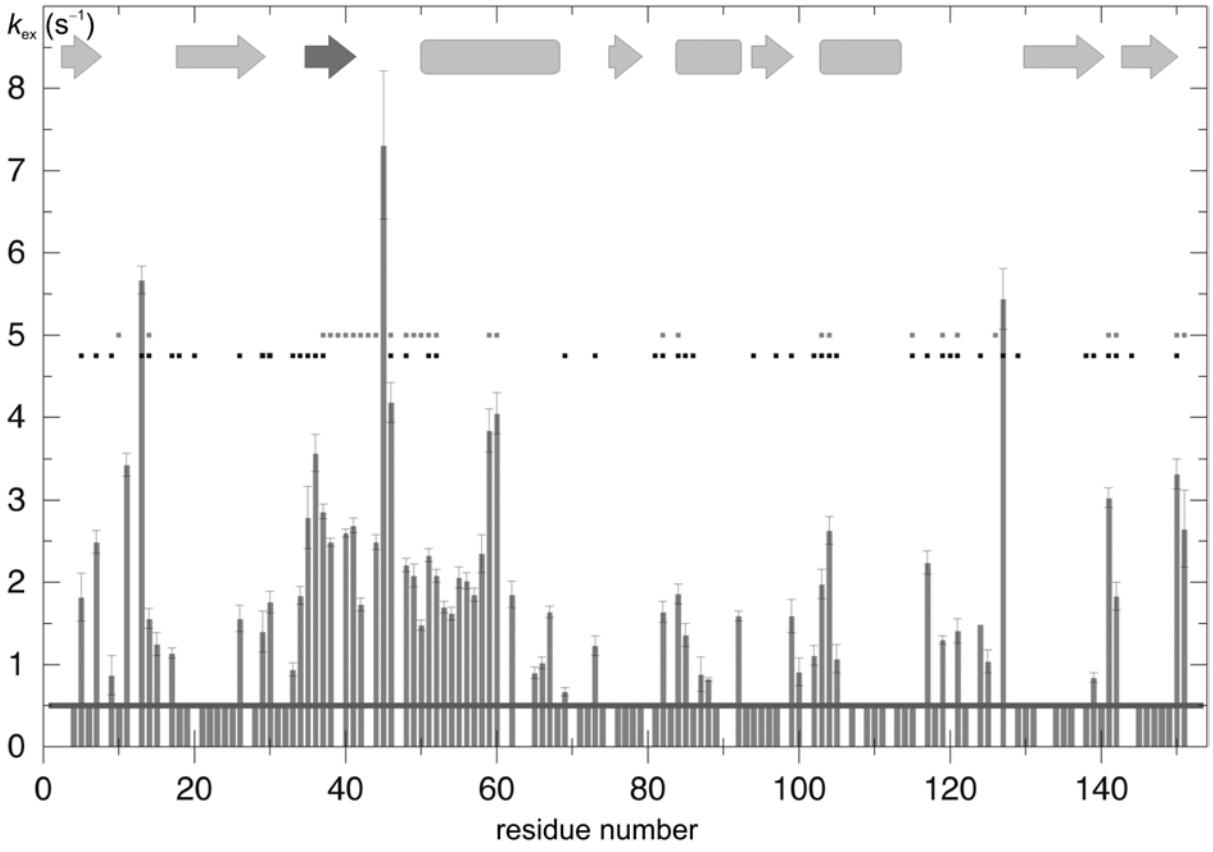


Fig. 3.8.1: Experimental amide proton exchange rates kex of αADT, measured by the NewMEXICO experiment. Bars at 0.5 s^{-1} indicate that for these residues the measured k_{ex} is below this value, i.e. under the detection limit of the experiment. The secondary structure is shown above, with the small helix Lys35-Gln43 shown as dark-grey arrow. Black squares represent non H-bonded amide protons in the crystal structure of αADT, while grey squares indicate residues not observable in βADT.

For example, in the edging strand S9, HX is slow for Ile4H^N and Ile6H^N, which are H-bonded to carbonyl acceptors in S19, but fast for Val5H^N and Asp7H^N, which point towards the solvent. Also, the long loop between H12 and S17 forms a large ring, thereby leaving a hole around Leu26H^N, which indeed shows fast hydrogen exchange as the only amide in S11. In the globular part, k_{ex} is low in most of the secondary structure elements except for the first turns of H11 and H12, and the βI turn between S18 and S19, where also residues Ile82-Met85, Lys103-Ser104 and Gly141-Glu142 remained unassigned in βADT.

In βADT, although the peptide chain emerges from the globular core through Glu32 and Ile33, their amide protons seem to be protected from the solvent and are still visible at pH 7.8. This might be inferred from the vicinity of potential side chain carboxylate H-bond acceptors for these residues. Similarly, Lys34, the second last residue visible in the N-proximal part of the βADT protrusion, faces a completely conserved glutamate residue (Glu57) located on the opposite branch of the protrusion in helix H10. Lys35 is followed by a proline and an unassigned gap of further 16 residues, until Lys52. In the αADT crystal structure, the helix Lys35-Gln43 protects the amide protons of Ile38-Ser45, for which no reduced HX rates were observed in solution. k_{ex} rather follows the calculated k_{int} with protection factors scattering around 1 as expected for a completely disordered peptide segment (**Fig. 3.8.2**).

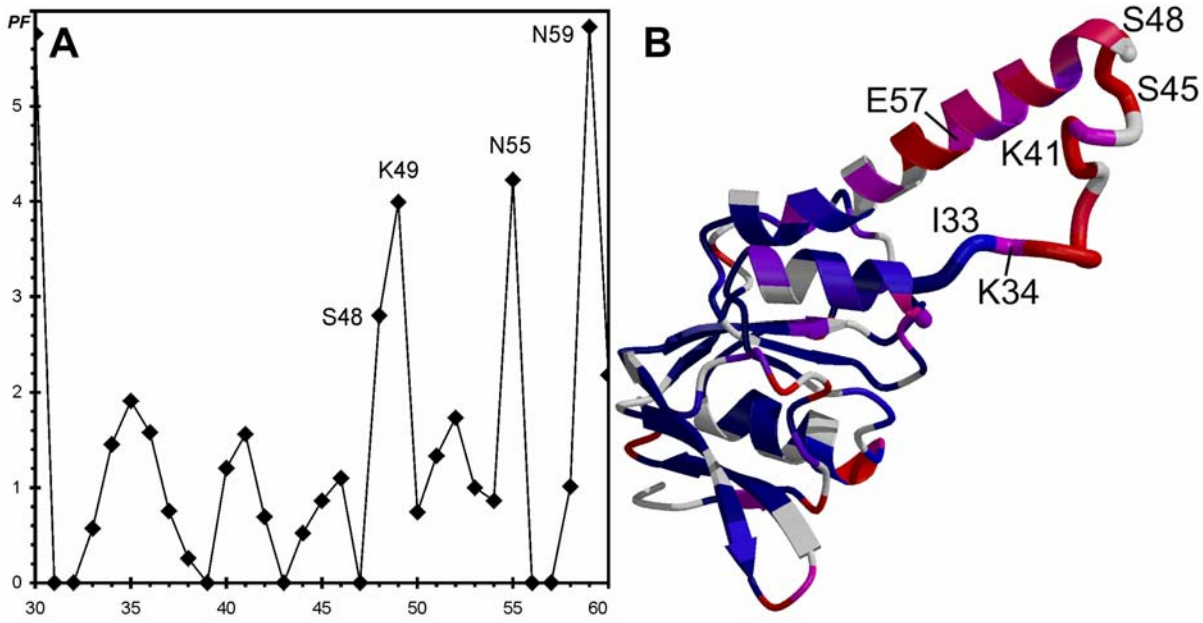


Fig. 3.8.2: A: $PF = k_{int}/k_{ex}$ for the protrusion region of αADT. **B:** Amide exchange rates mapped onto the structure of αADT. The color code is ramped from dark blue ($k_{ex} < 0.5\ s^{-1}$) to red ($k_{ex} > 2.5\ s^{-1}$). Prolines and residues affected by spectral overlap are shown in light grey.

Interestingly, amide exchange even remains fast throughout what is the second and third turn of the H10 N-cap in αADT crystal structures suggesting a temporary helical unwinding. Also for βADT, at least partial N-terminal uncapping of H10 seems possible as could be concluded from strong positive exchange peaks at the ω_1 water frequency in the HNH-NOESY spectrum for the respective amides. In contrast to that, pure cross relaxation with surface water has been shown to give rise to negative peaks (extreme narrowing) due to its generally sub-nanosecond residence time.[109] *PF* is larger (k_{ex} is reduced) for Ser48 and Lys49 in the βI turn at the tip of the protrusion, and for Asn55 and Asn59, both of which face the inner side of H10 and may therefore indicate asymmetry in the unwinding process.

At the very end of H10 in αADT, the amide protons of Ser71, Gly72 and Ala73 show strong NOE correlations to a proton resonating at 5.53 ppm, which is neither nitrogen nor carbon bound. Obviously, the polar side chain of Ser71 (which is Val71 in βADT) is forced towards the interior, leading to an unusual protection of Ser71OH from the solvent and drastically reduced hydrogen exchange.

3.9 Titration with trifluoroethanol

Trifluoroethanol is known to strongly stabilise α-helical structures in peptides and proteins [110] and should therefore probe the propensity of Lys35-Gln43 in αADT to form a helix in the absence of crystal packing effects. A series of ^{15}N-HSQC spectra with increasing concentrations of TFE showed that despite strong precipitation at only 7 % (v/v), the resonance positions of these residues were indeed strongly affected (**Fig. 3.9.1**).

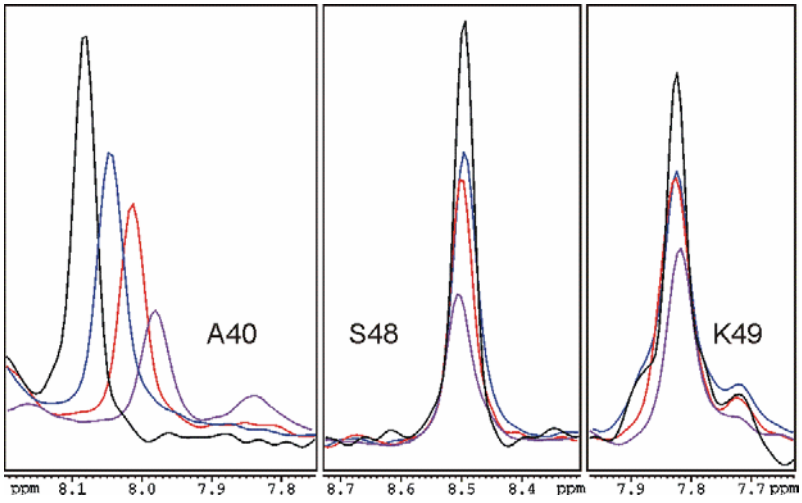


Fig. 3.9.1: 1D slices from 15N-HSQC of αADT recorded at various TFE concentrations: black 0 % (reference), blue 3 %, red 5 %, purple 7 % v/v TFE. The insensitivity of the amides of Ser48 and Lys49 to the TFE concentration suggests that these residues are protected from the solvent.

Large H^N and ^{15}N signal shifts were also observed for most other solvent exposed residues, for example in H10 (**Fig. 3.9.2**) and may therefore be attributed to pure solvation effects. Residues Ser48 and Lys49 in the βI turn at the tip of the protrusion, which also showed reduced HX rates, were nearly unaffected by TFE.

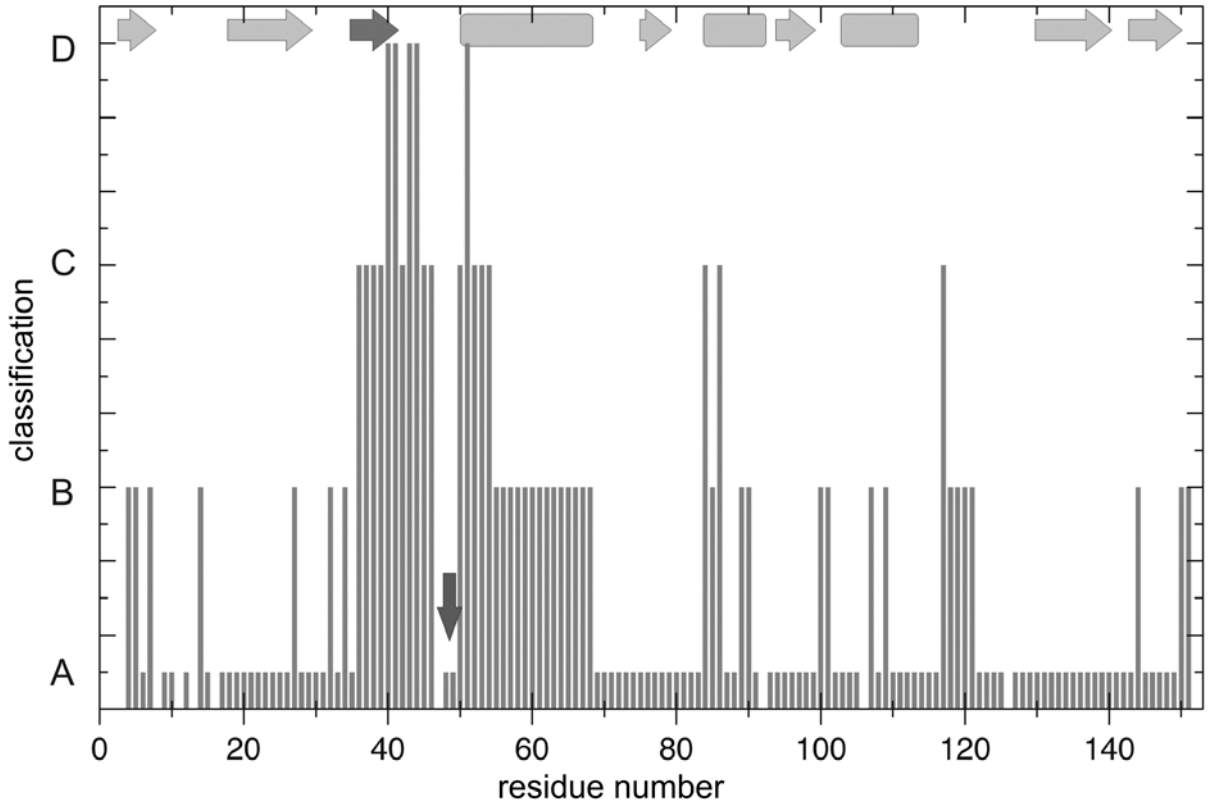


Fig. 3.9.2: Results of the TFE titration for αADT. The chemical shift changes of the amide groups are normalised to 1 % v/v TFE, classified and plotted versus the sequence number. A: not significant, B: < 10 ppb, C: between 10 and 20 ppb, D: > 20 ppb. The vertical grey arrow indicates the position of Ser48 and Lys49. Secondary structure elements are illustrated. The dark grey arrow symbolises the β-strand found in the closed thermosome.

No substantial shifts of H^α resonances in the protrusion was, however, observed in a HNHA spectrum recorded at 7 % TFE. In agreement with the other data, significant tendency towards helix formation may therefore be excluded and the occurrence of rigid secondary structure elements in this part of the protrusion seems unlikely. Similar results were obtained from a ^{15}N relaxation analysis of the ADTs, which showed the protrusions as intrinsically disordered elements.[111]

4 The GM2 activator protein

4.1 Degradation of glycosphingolipids

Glycosphingolipids (GSLs) are ubiquitous constituents of eukaryotic plasma membranes and, with their oligosaccharide head groups protruding into the extracellular space, form part of the cell´s *glycocalix*.[112] The catabolic pathway of these complex lipids takes place on the surface of vesicles in the acidic milieu of the *lysosome*, where specific *exohydrolases* sequentially cleave off the terminal monosaccharide moieties from the non-reducing end. In the case of GSLs with rather short oligosaccharide head groups, this requires the assistance of small non-enzymatic cofactors, glycoproteins, which mediate the interaction between the membrane-embedded lipid substrate and the water-soluble hydrolases.[113] Blockage of any of these steps, caused by a functional defect in either the enzyme or the catalyst, leads to accumulation of the respective GSL and severe storage diseases (**Figs. 4.1.1**).

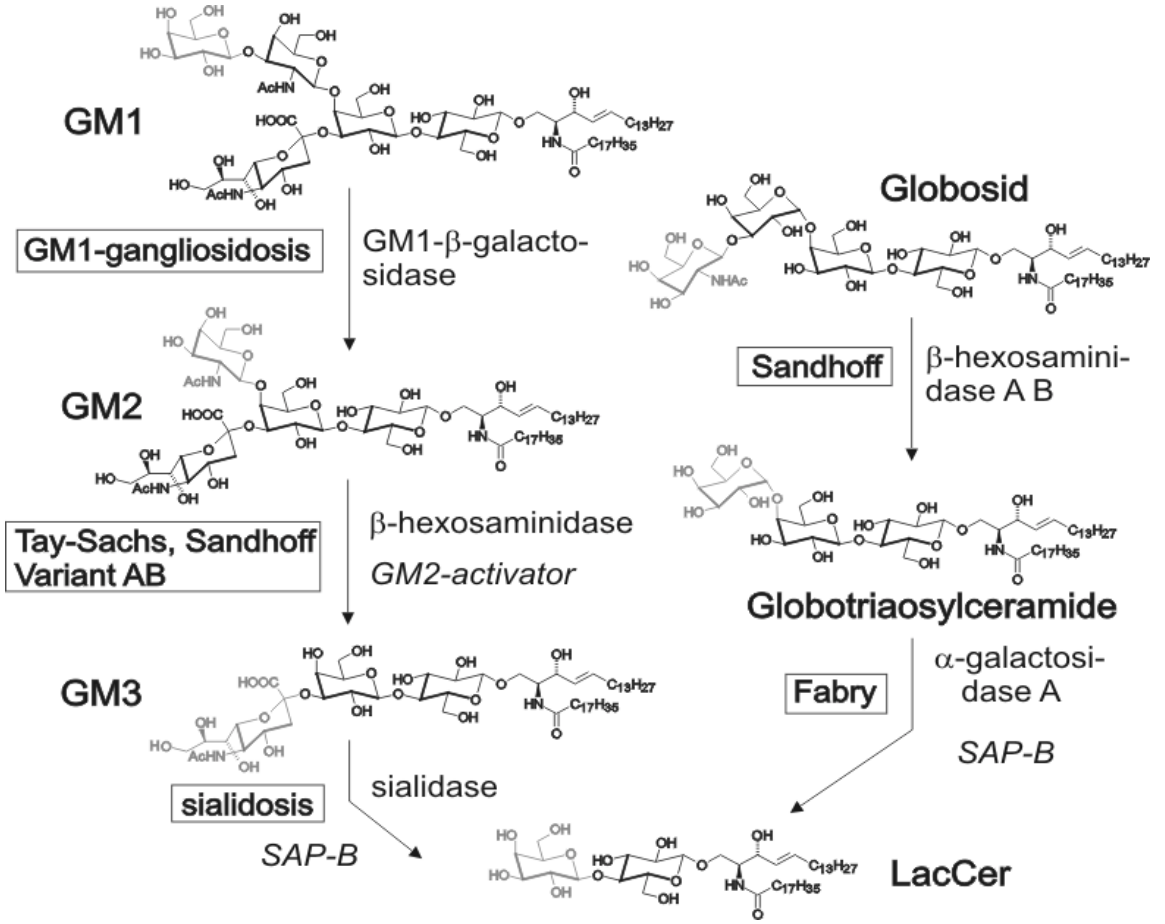


Fig. 4.1.1: Catabolic pathways of GSLs in the lysosome (part I). Molecules are labelled in bold letters, monosaccharide moieties cleaved by the respective enzymes in grey, cofactors in italics and the storage diseases related to a functional defect in either of the two in black with boxes.

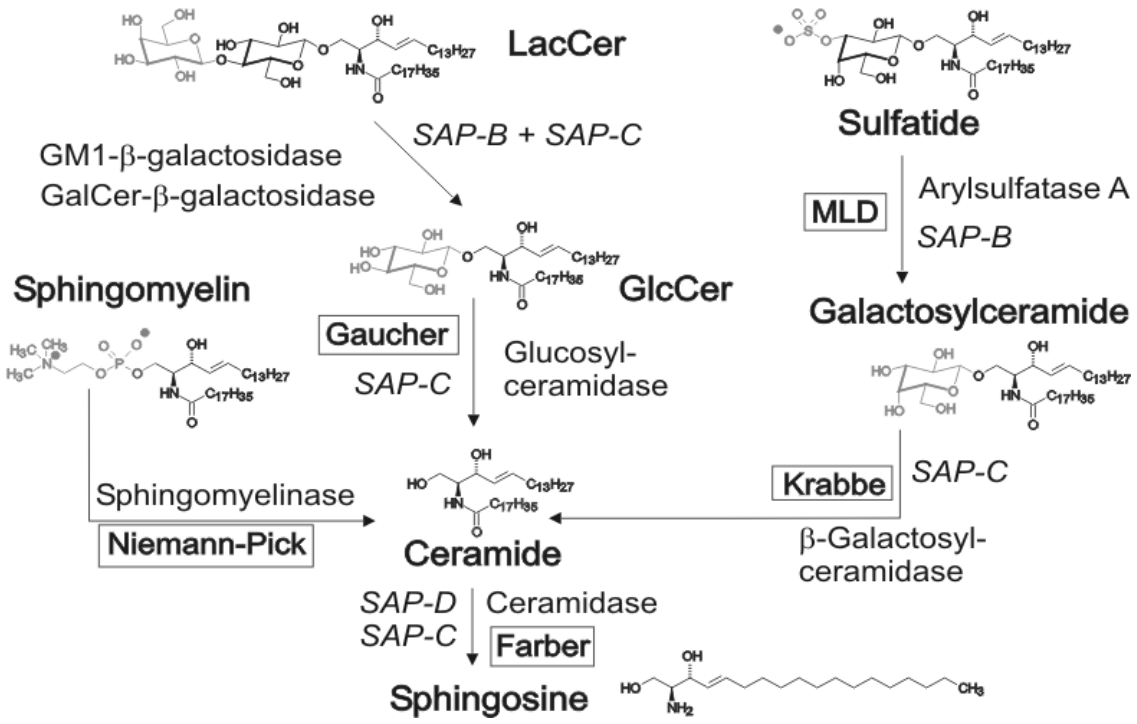


Fig. 4.1.1 (cont.): Catabolic pathways of GSLs in the lysosome.

To date, five of these so called *sphingolipid activator proteins* are known: the four saposins (SAPs) A-D, which are proteolytically generated from a single precursor, and the GM2 activator protein (GM2AP), which is coded by a separate gene. The latter serves as an essential cofactor in the degradation of GM2, a member of the special class of sialic acid containing GSLs, which are called *gangliosides* and are highly prevalent in the central nervous system. Its physiological significance is illustrated by the occurence of a fatal neurological storage disorder, the AB variant of GM2 gangliosidosis.[114] According to the *liftase model*, GM2AP recognises, complexes and lifts GM2 out of the lipid bilayer, thereby presenting it to the enzyme β-hexosaminidase A for hydrolysis (**Fig. 4.1.2**).[115]

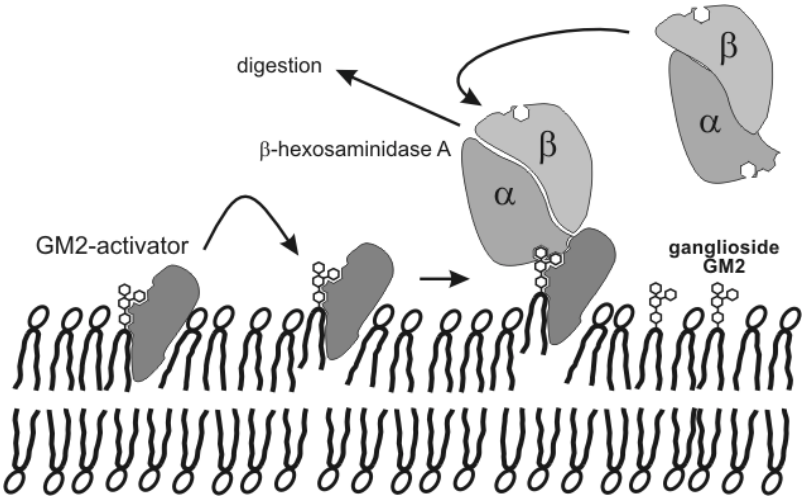


Fig. 4.1.2: Liftase model. GM2-activator protein binds to ganglioside GM2 and lifts it up from the (vesikel-) membrane, thereby making it accessible to β-hexosaminidase A. The enzyme, a α/β heterodimer, undergoes a conformational change upon interaction with GM2AP.

4.2 The GM2 activator protein

In mammalian cells, GM2AP is first synthesised as a 193 amino acid precursor, which is then glycosylated at Asn63 and modified along the pathway to the lysosome: In the endoplasmatic reticulum, the translocation (pre) sequence is cleaved off after Ala23, and after reaching the lysosome the pro-sequence (His24-Leu31) is removed (**Fig. 4.2.1**).[116] Of the five well documented mutations found in gangliosidosis AB (AB denotes intact α- and β- chains of β-hexosaminidase) variant patients, two are truncated proteins and the others are: ΔK88, C138R and R169P.[117]

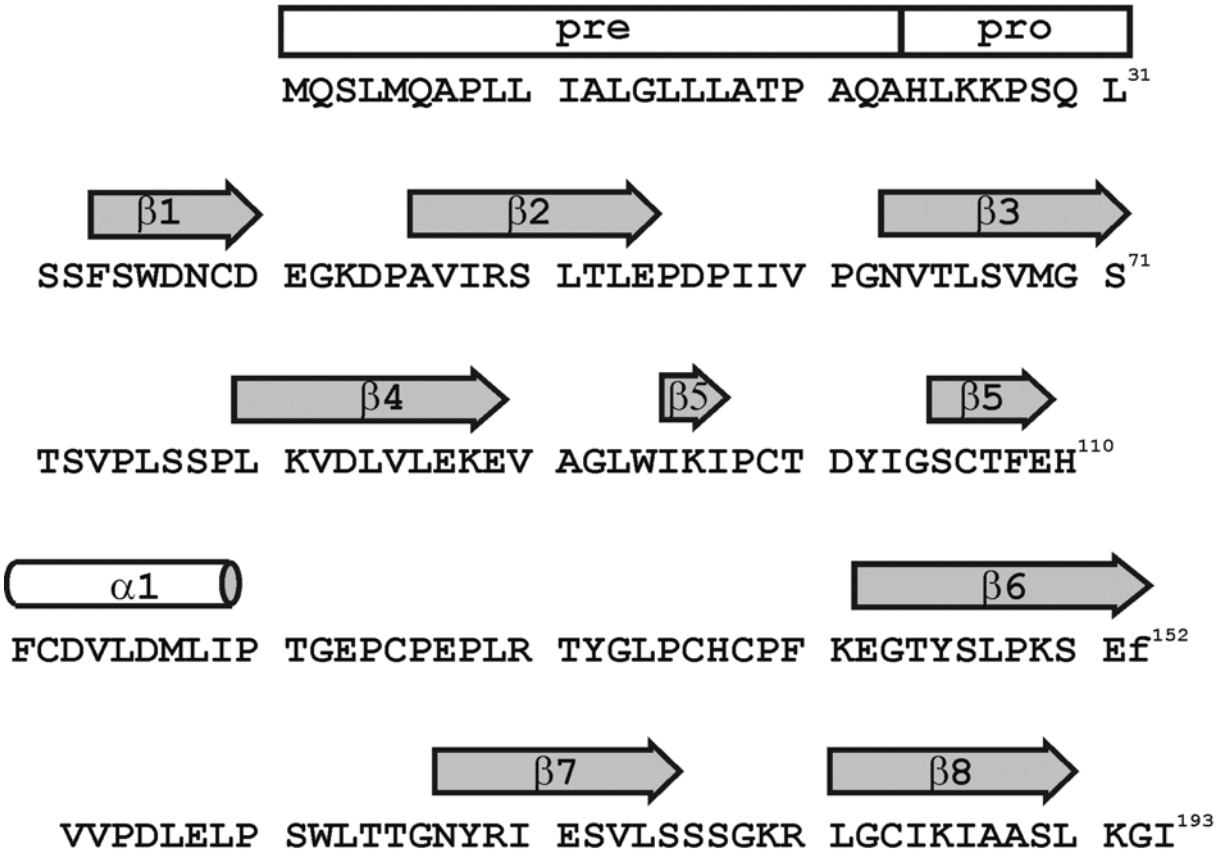


Fig. 4.2.1: Sequence of the full human wt-GM2 activator pre-pro-protein. Pre- and pro- segments are marked above the sequence, and the secondary structure is indicated by arrows (β-strands) and a cylinders (α-helix).[118] In the NMR studies the mature sequence (Ser32-Leu193) including an N-terminal tag EAEAYV originating from the expression vector and a C-terminal tag RH_6 for purification was used.

The crystal structure of recombinant, unglycosylated human wt-GM2AP expressed in *E.coli* revealed a novel β-cup fold, whose main feature is an eight-stranded anti-parallel β-sheet with a central hydrophobic cavity (**Fig. 4.2.2**).[118] Two of the four disulfide bonds form

remarkably short loops: Cys99-Cys106, which also interrupts strand β5, and Cys125-Cys136, a third disulfide Cys112-Cys138 stabilises the only short α-helix in the protein. Most recently, the structural analysis of lipid complexes of GM2AP confirmed the occurence of different conformers in the assymmetric unit, as well as different lipid binding modes within the apolar pocket.[119] In this second paper, also a model of membrane binding was proposed, involving two of the three tryptophan residues, Trp94 and Trp162.

Despite these structural data and a wealth of biochemical information, the mechanism of GM2AP action at the water-lipid interface is far from being understood. Moreover, since the activity of GM2AP was shown to drop sharply from the maximum at pH 4.2,[120] the crystallisation performed at pH 7 and 5.5 and in the absence of glycosylation may not reflect the biologically most active conformation. For a better understanding of the activity of GM2AP, structural data of the protein in solution from NMR spectroscopy under close-to-in-vivo conditions were highly desirable.

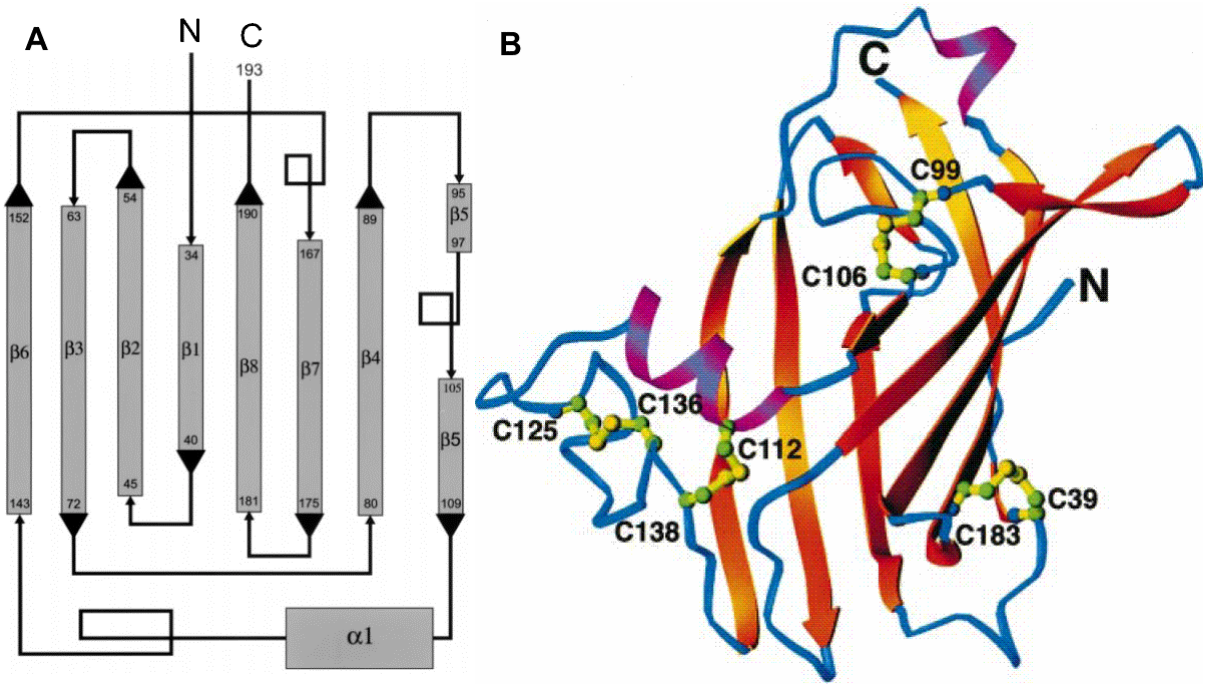


Fig. 4.2.2: A: Topology diagram of secondary structure, showing β-strands as arrows, labelled with first and last residue and the α-helix as cylinder. The short loops formed by disulfide bonds are also displayed. **B:** Ribbon model of the tertiary structure of the GM2 activator protein as determined by X-ray crystallography. β-strands are shown in orange, the α-helix in purple, and loop regions in blue. The positions of the four disulfide bonds are shown as ball-and-stick with green balls representing C^{α} and C^{β} carbons, and yellow balls representing sulfur atoms. Both figures were taken from [118].

4.3 Expression in *Pichia pastoris* and purification

Previous attempts to produce isotopically labelled GM2AP in *E.coli* after refolding from inclusion bodies led to unfavourable aggregation behaviour at concentrations required for NMR spectroscopy. Glycosylation was expected to enhance the solubility of the protein, but requires eukaryotic hosts like insect cells, which may not be cultivated in minimal media and thus make isotope labelling extremely costly. Methylotropic yeasts like *Pichia pastoris* are now widely recognised as efficient expression systems that combine the ease of genetic manipulation with the ability of eukaryotic processing, folding, and post-translational modifications.[121] All expression and purification work of GM2AP was done in the laboratory of Prof. K. Sandhoff at the Kekule-Institute for Organic Chemistry and Biochemistry, University of Bonn, Germany, and is described in detail elsewhere.[122]

Briefly, the cDNA of the mature GM2AP (starting with S32) with C-terminal RH_6 tag was ligated into the expression vector pPIC9K and fused to the *S.cerevisiae* derived α-factor signal sequence, under control of the alcohol oxidase (AOX) 1 promoter. Transformed *Pichia pastoris* cells (GS115 line) were grown at 30 °C to an OD_{600} of 3-6 in a modified FM22 medium containing 2 g/L ^{15}N-ammonium sulfate. The medium was also optimised for an optional efficient isotopic labelling with ^{13}C-glucose instead of ^{13}C-glycerol as the sole carbon source. Expression of GM2AP was induced by resuspension of the cells in the desired methanol medium, and the directly secreted protein was isolated after 3 days from the supernatant.

Sequencing of the N-terminus by ESI-MS/MS fragmentation analysis after tryptic digestion confirmed the cleavage of the α-factor except for two remaining glutamate-alanine repeats. After Ni-NTA chromatography the protein was essentially pure, except for the presence of 5-10 strongly bound lipid molecules per molecule GM2AP. Because the solutions obtained after concentration turned out highly viscous and not suitable for NMR analysis, lipids and aggregates were removed by reverse phase and gel filtration chromatography, respectively. The specific activity was found to correspond to the one reported for GM2AP isolated from human kidney.

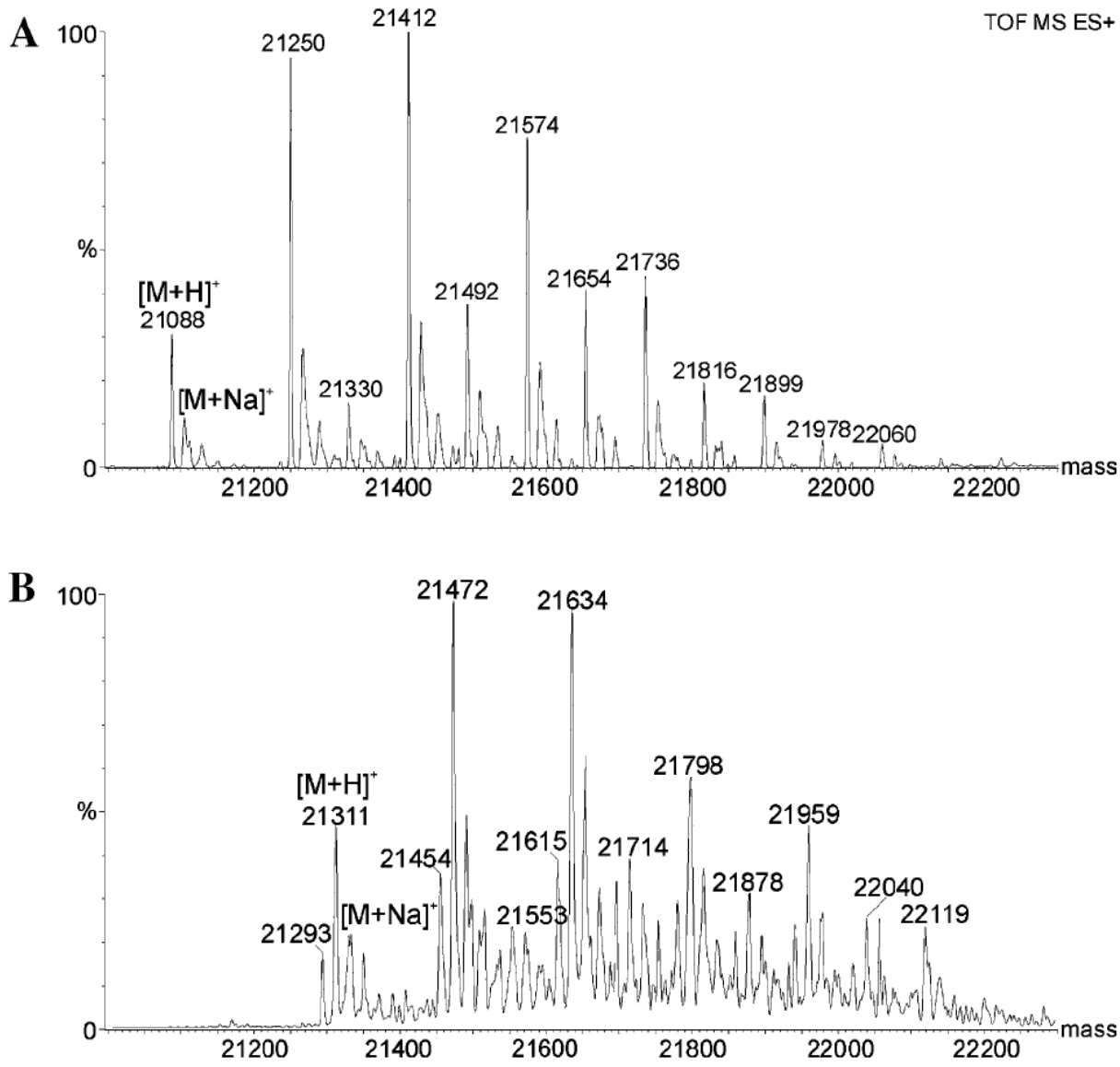


Fig. 4.3.1: ESI-Q-TOF mass spectra of unlabelled (**A**) and ^{15}N-labelled (**B**) purified GM2AP expressed in *Pichia pastoris*. The average mass for the recombinant mature protein $[M+H]^+$ bearing the oligosaccharide $GlcNAc_2Man_9$, which is shown on the right, is 21090 Da, for its uniformly ^{15}N-labelled analogon 21314.5 Da. All peaks at higher masses with increments of 162 Da represent GM2AP bearing the moieties $GlcNAc_2Man_{10\text{-}16}$. Figure taken from [122].

ESI-Q-TOF mass spectrometry revealed that GM2AP is glycosylated with a glycan structure, which is frequently found on glycoproteins expressed in *Pichia pastoris* (**Fig. 4.3.1**). This so-called *high mannose* type consists of two units GlcNAc and a branched chain of 9-16 α1,2, α1,3 and α1,6 glycosidic linked mannose moieties with an overall mass of 1866-3000 Da. ^{15}N labelling of the 224 nitrogen atoms increased the experimentally determined mass by 222 Da and therefore can be considered nearly complete (> 98 %).

4.4 Stability and amenability for NMR spectroscopy

Samples of ^{15}N-labelled GM2AP were prepared in acetate buffer (50 mM, pH 4.5) and phosphate buffer (50 mM, pH 7.0) and at the limit of solubility (11 mg/mL corresponding to 480 µM). In ^{15}N-HSQC spectra, recorded in the temperature range of 27 °C to 57 °C, H^N protons are well frequency dispersed, indicative of a folded protein (**Fig. 4.4.1**). In all spectra the H^ε resonances of the three tryptophan residues Trp36, Trp94 and Trp162 could be identified. At pH 4.5, one pair of asparagine δ-NH_2 resonances (Asn38 or Asn167, Asn63 is the site of glycosylation) and all five arginine ε-NH protons are visible at δ(^{15}N) ~ 85 ppm, but only two at pH 7.0 (data not shown).

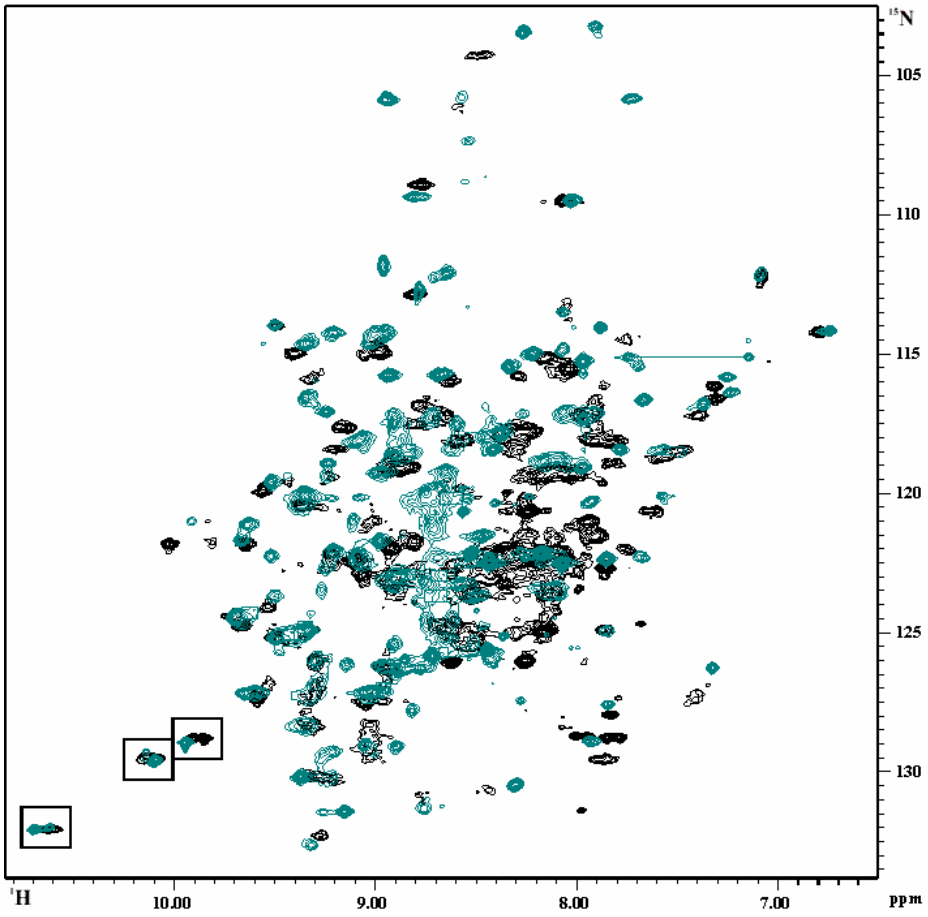


Fig. 4.4.1: Overlay of ^{15}N-HSQC spectra of GM2AP at 42 °C and 14.1 T in 50 mM acetate (pH 4.5, blue) and 50 mM phosphate (pH 7.0, black). Acquisition times were 85.0 ms and 70.7 ms in the ^{15}N and ^{1}H dimensions, respectively. Tryptophan ε-NH are marked with boxes, and the asparagine δ-NH_2 group with a horizontal bar.

Some peaks are clearly recognisable from coinciding resonance frequencies in the ^{15}N-HSQC spectra at acidic and neutral buffer, which is a strong hint that the overall fold and at least some structural motifs are conserved. It remains a matter of speculation, however, whether the differences in the remaining signals are due to true conformational changes, locally differently charged aspartate, glutamate and histidine side chains or simply fast solvent exchange expected at 42 °C and pH 7.0.

Besides, there is massive evidence for conformational exchange processes within the molecule, which are apparently even more pronounced at neutral pH. In the pH 4.5 spectra, the overall number of backbone resonances exceed the number of non-proline residues of 159 in the sequence, and several signal pairs can be attributed to different conformations in slow exchange. For example, both tryptophan ε-NH protons at $\delta(^1H)$ = 10.7 and 10.1 ppm are split into two resonances, albeit with different relative populations (**Fig. 4.4.2**).

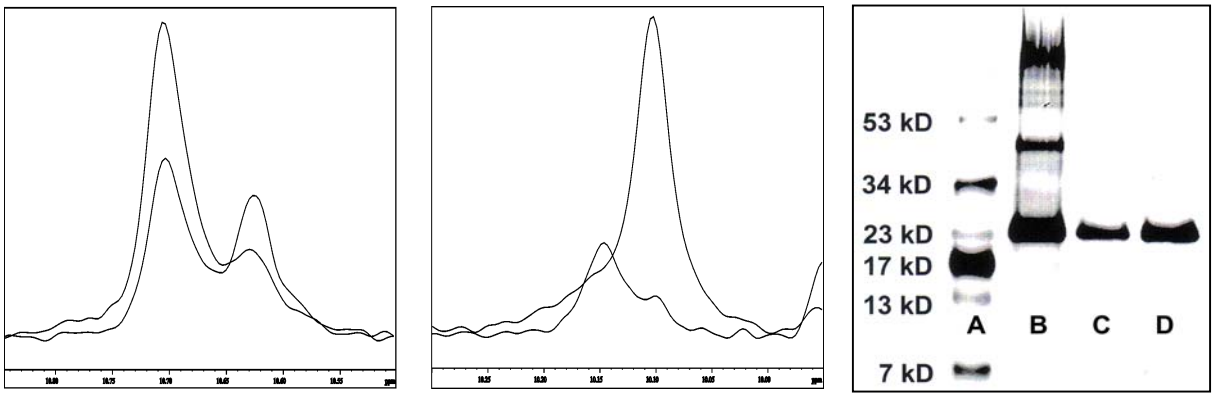


Fig. 4.4.2: Left and middle: Cross sections of pairs of trp ε-H resonances in the ^{15}N-HSQC spectrum of Fig. 4.4.1 Relative populations are 2:1 (left) and 4:1 (right). **Right:** SDS gels with silver staining showing molecular weight standards (trace **A**), precipitate in the NMR tube (**B**), supernatant in the NMR tube (**C**) and GM2AP before heating in the magnet (**D**). The precipitate comprises dimers and higher molecular associates.

Many resonances show temperature- and field dependent line broadening leading to a wide distribution of signal intensities (intermediate exchange), from extremely sharp and intense signals (loops and termini) to ones which are virtually indistinguishable from the noise. In the region of random coil H^N frequencies many signals suffer from severe overlap, which could not be removed using ^{15}N-TROSY techniques [123] and 17.5 T magnetic field.

At elevated temperatures, broadening occurred due to fast hydrogen exchange of amide protons at pH 7.0, but the sample remained stable. On the other hand, under acidic conditions hydrogen exchange is not an issue, but close to the molecule´s pI of 5.4, precipitation of protein aggregates (**Fig. 4.4.2**) was observed. Despite treatment of the NMR-tubes with TMSCl in toluene prior to use, the amount of precipitate comprised as much as 30 % of the concentration within 3-4 days at 42 °C, being less in dilute samples. The presence of lipids (omitting the RP chromatography purification step), significantly increased the lifetime of GM2AP, but again at the cost of spectral resolution. The instability of highly in contrast to moderately purified GM2AP was long back pointed out by Li *et al*.[124]

4.5 Diffusion and ^{15}N-relaxation

Molecules in solution are not fixed in their positions, they rather perform a thermally driven random translational motion, which for spherical particles in an isotropic environment of viscosity η can be described by a single parameter, the translational diffusion coefficient D_t. According to the Stokes-Einstein relation D_t is related to the hydrodynamic radius r_h, i.e. the size of the hydrated particle, which in turn depends on the molar volume $\bar{v}\,M$ ($\bar{v}$ is the mean specific volume and M is the molar mass).

$$D_t = \frac{kT}{6\pi\eta r_h} \qquad r_h = \sqrt[3]{\frac{3\bar{v}M}{4\pi N_A}} \tag{4.5.1}$$

Pulsed field gradients (PFG) are now widely used to measure translational diffusion coefficients D_t and thereby obtain an estimate of the size and association state of macromolecules in solution. The underlying idea is the incomplete rephasing of coherence by an opposite pair of gradient pulses, separated by a delay T_d, due to movement of the molecule along the gradient axis (**Fig. 4.5.1**).

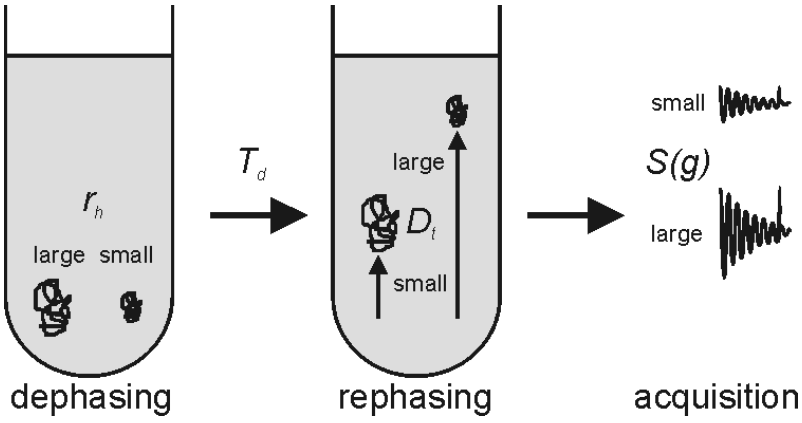


Fig. 4.5.1: Principle of PFG-NMR diffusion measurements: After coherence dephasing by a PFG molecules undergo diffusion with D_t, which depends on the size r_h. After a period T_d, rephrasing of coherence is incomplete and leads to a gradient depending signal $S(G)$.

Recorded as a function of the gradient strength G or its length δ, the final signal follows a Gauss-type behaviour, from which D_t may be extracted:

$$S(G) = S_0 \exp(-D_t (G\gamma\delta)^2 T_d) \tag{4.5.2}$$

Several improvements of the original Hahn gradient spin echo (PFG-SE) sequence [125] have been developed: The difference in longitudinal and transverse spin relaxation times of macromolecules can be exploited, if T_d is implemented during a period of longitudinal magne-

tisation (stimulated echo, PFG-STE).[126] To address the problem of gradient ring down, after the rephasing gradient a second period of longitudinal magnetisation (longitudinal eddy current delay) is inserted (STE-LED).[127] Disturbing fields induced at interfaces with non-continuous susceptibility (in an NMR tube)are cancelled out if using a bipolar pair of gradient pulses, separated by a π-pulse, instead of a simple gradient pulse (BPP-LED).[128] At elevated temperatures, temperature gradients within the sample (especially if large sample volumes are used) can lead to convection overrating diffusion. Coherence dephasing by a flow of constant velocity along the gradient axis (convection) is supressed by the use of a double stimulated echo (DSTE) scheme with selection of coherence order with inverted sign.[129]

The main drawback of measuring translational rather than rotational diffusion is that D_t is inverse proportional to only the cubic root of, whereas the rotational diffusion tensor is directly inverse proportional to the molecular size. On the other hand, unlike spin relaxation, translational diffusion is insensitive to conformational exchange as long as the shape of the molecule is not strongly affected. Given the exchange processes in GM2AP apparent from the ^{15}N-HSQC spectra, D_t should therefore be a more reliable measure of the association state.

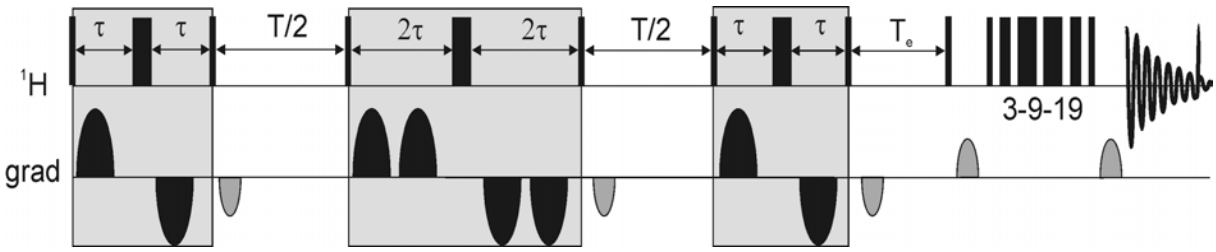


Fig. 4.5.2: Pulse sequence used for the measurement of diffusion coefficients with additional water suppression by a 3-9-19 binomial sequence.[129] The 2 ms sine-shaped gradient pulses in the shaded boxes were linearly incremented in 32 experiments, and a 4 ms ring down delay was applied. T was usually set to 150 ms for macromolecules and 25 ms for a control experiment for measurement of the diffusion coefficient of H_2O (also omitting the water suppression scheme), and T_e was 10 ms. D_t was determined by a least squares fit of the intensities $S(G)$ of five spectral maxima to equation 4.5.2, where in this pulse sequence $T_d = T + 8/3\delta + 3(\tau-\delta)$, and G is the absolute gradient strength in Gauss/cm, averaged over the sine shape.

Using the sequence of **Fig. 4.5.2**, for solutions of 0.5 mM and 0.25 mM the measured diffusion coefficients at 42 °C were very similar (D_t = 1.38 ± 0.03 * 10^{-10} m^2/s versus D_t = 1.40 ± 0.05 * 10^{-10} m^2/s), arguing against unspecific association of GM2AP. A hydrodynamic simulation with HYDRONMR [130] using a model of the native mature GM2 activator protein

(17.6 kDa) with an atomic bead size of 3.2 Å predicted a Stokes radius of r_h = 23.5 Å and D_t = 1.59 * 10^{-10} m^2/s, which is 14 % more than the experimental value. However, a stoichiometric dimerisation is unlikely, since for a dimer D_t < 1.25 * 10^{-10} m^2/s would be expected. The discrepancy may be understood considering the higher molecular weight of the experimental system of 22-23 kDa, and assuming that the flexible N- and C- termini and the glycan chain have a larger mass specific contribution to translational friction than the compactly folded protein. Apart from that, the hydrodynamic calculation used the tabulated viscosity of pure H_2O, whereas concentrated protein solutions for NMR spectroscopy are slightly more viscous. This is due to mutual friction of macromolecules, which can fill up several percent of the solution volume. For example, from the molar volume 0.029 m^3/mol of the hydrodynamic model GM2AP and the experimental concentration 480 µM a nearly 1.5 % fractional volume of the hydrated protein can be calculated.

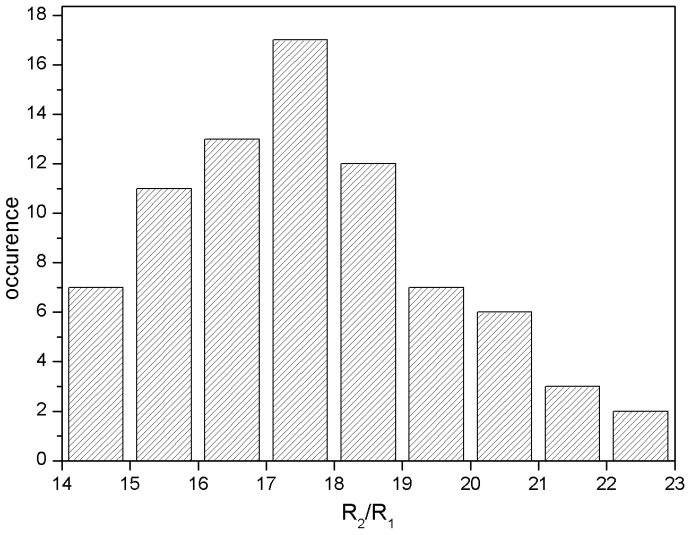


Fig. 4.5.3: Distribution of ^{15}N-R_2/R_1 at 14.1 T and 42 °C for 78 residues of GM2AP. Relaxation rates and the ^{15}N{^{1}H} NOE were measured for 109 ^{15}N-HSQC-resolved peaks using standard pulse schemes [131] and data analysis with DASHA.[132] 20 residues with errors in either rate larger than 10 % were excluded, as well as 11 residues with ^{15}N{^{1}H} NOE < 0.7 indicative of internal mobility.

Similar results were obtained from measuring the ratio of ^{15}N-R_2 and R_1 relaxation rates for a large number of amide groups. It is important to note that the mean R_2/R_1 only allows for estimating an upper limit of the molecular size, since the individual R_{ex} contributions to R_2 are not known prior to knowledge of the rotational diffusion tensor. Referring to the quality of ^{15}N-HSQC spectra, however, large R_{ex} were expected for many residues, which might explain the broad distribution of R_2/R_1 (**Fig. 4.5.3**). Neglecting R_{ex} and anisotropic tumbling, the average R_2/R_1 of 17.5 yields a molecular correlation time of $\tau_c \approx 13$ ns, as compared to a harmonic mean of $\tau_c = 8.1$ ns from hydrodynamic modelling. Again, the discrepancy is likely to be caused by the incompleteness of the model rather than dimerisation of GM2AP, assuming τ_c (dimer) $\approx 2\,\tau_c$ (monomer).

4.6 Secondary structure

Once resonances of backbone nuclei are assigned, secondary structure elements in the sequence may be predicted using $\delta_{sec}(i)$ of H^{α}, C^{α}, C^{β} and C', $^{3}J_{HNH\alpha}$, $d_{\alpha N}(i,i+1)$ and $d_{NN}(i,i+1)$, as described in sections 3.5 to 3.7. For GM2AP resonance assignment seemed virtually impossible with reasonable amounts of isotopically labelled protein, given the poor ^{15}N-HSQC quality and short lifetime at least under the chosen conditions. However, the relative occurrence of secondary structure elements may be estimated from distributions of above parameters, which, of course, may be biased towards residues with good ^{15}N-HSQC resolution.

For the distribution of $^{3}J_{HNH\alpha}$ (**Fig. 4.6.1 A**), 125 diagonal intensities I_d were directly measured in the HNHA spectrum, and further 24 were estimated from contour levels in the case of overlap. Similarly, 118 intensities I_c of H^{α} cross signals were measured, 19 intensities below the peak picking threshold were estimated, and 12 intensities of invisible cross peaks were assigned to the noise level. Glycine residues, arginine ε-NH, tryptophan ε-NH and asparagine δ-NH_2 groups were excluded from the analysis.

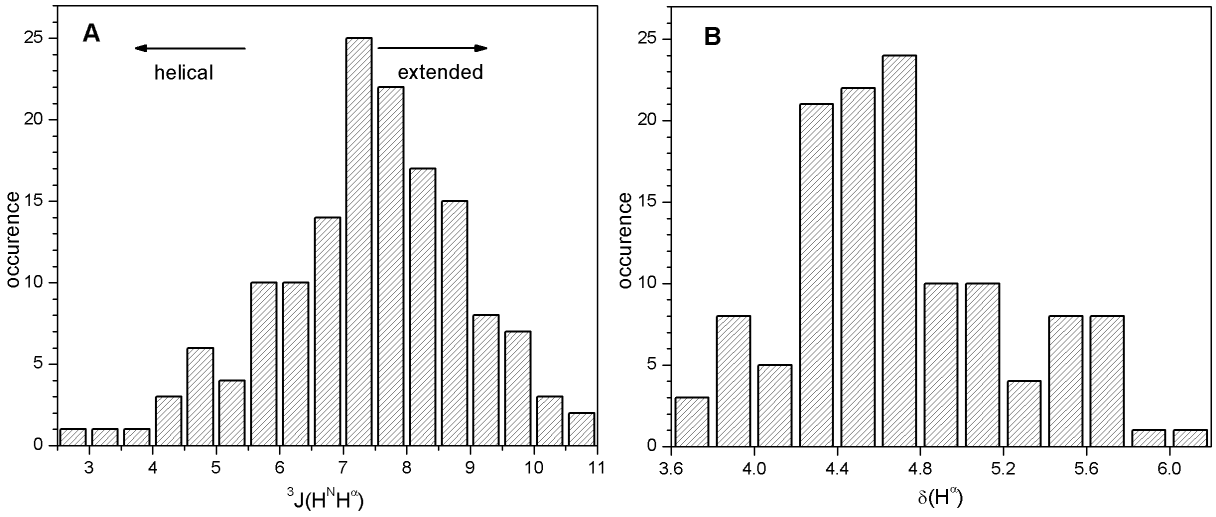


Fig. 4.6.1: Distributions of coupling constants $^{3}J_{HNH\alpha}$ (**A**) and absolute chemical shifts $\delta(H^{\alpha})$ (**B**) for 149 residues with ^{15}N-HSQC-resolved resonances. The HNHA experiment [Vuister93] was acquired at 14.1 T with acquisition times of 9.9 ms, 15.4 ms and 70.7 ms in the $\omega_1(H^{\alpha})$, $\omega_2(^{15}N)$ and $\omega_3(H^{N})$ dimensions and a total experiment time of approximately four days. Coupling constants were calculated according to $^{3}J_{HNH\alpha} = \arctan(I_c/I_d)^{0.5}/(2\pi*0.9*d_3)$, where d_3 (= 13 ms) was the overall evolution period of $^{3}J_{HNH\alpha}$, and 0.9 is a scaling factor taking into account different relaxation behaviour of H^{N} and H^{α} protons.

From the same spectrum, a distribution of absolute chemical shifts $\delta_{exp}(H^{\alpha})$ was obtained (**Fig. 4.6.1 B**). Albeit less significant than secondary chemical shifts $\delta_{sec}(H^{\alpha})$, they support the estimate of 10 % helical and 50 % extended residues from $^{3}J_{HNH\alpha}$. For comparison, in the crystal structure [118] 7 % of the residues are located in helices (plus another 6 % have typical helical conformation), and 42 % (plus 32 %) in β-sheets.

Without ^{15}N-HSQC assignment, is principally impossible to identify sequential NOEs $d_{NN}(i,i+1)$ from HNH- and NNH-NOESY experiments due to the presence of intermediate and long range correlations. However, $d_{\alpha N}(i,i+1)$ may be identified for residue (i+1), because the frequency of the intraresidual $H^{\alpha}(i+1)$ is known from the HNHA correlation. Neglecting the interference with $d_{\beta N}(i,i+1)$ or $d_{\beta N}(i,i)$ in serine and threonine, $d_{\alpha N}(i,i+1)$ stronger than $d_{\alpha N}(i+1,i+1)$, which is nearly secondary structure independent, is a strong indicator for an extended conformation of residue (i). In GM2AP, this was observed for 63 residues out of 149 in a HNH-NOESY spectrum (80 ms mixing),[44] thus confirming a predominance of residues in extended conformation in GM2AP. In the remaining 86 residues, no strong $d_{\alpha N}$ correlation was found, or $d_{\alpha N}(i,i+1)$ and $d_{\alpha N}(i+1,i+1)$ correlations were overlapped (data not shown).

4.7 Binding of GM1

At physiological pH, the presence of sialic acid confers at least partially a negative charge on the bulky oligosaccharide head group of gangliosides, which form micelles in aqueous solution at concentrations above CMCs of typically 10^{-5} to 10^{-4} M.[133] For example, micelles of GM1 have been reported to contain approximately 160 molecules, rendering the micelles a mass of about 250 kDa and a hydrodynamic radius of $r_h \approx 57$ Å.[134] As therefore expected, the ^{1}H-NMR spectrum of a solution of GM1 shows very broad resonances (**Fig. 4.7.1**).

Since no separate signal set was observed for free monomeric GM1, either its concentration (the CMC) is too low or the equilibrium between free and micelle bound GM1 is in the fast exchange regime averaging chemical shifts and linewidths. Diffusion measurements suffered from poor signal noise ratio, but yielded a single diffusion coefficient $D_t = 0.7 \pm 0.1 * 10^{-10}$ m^2/s, which is in good agreement with a large species of about 200 kDa size.

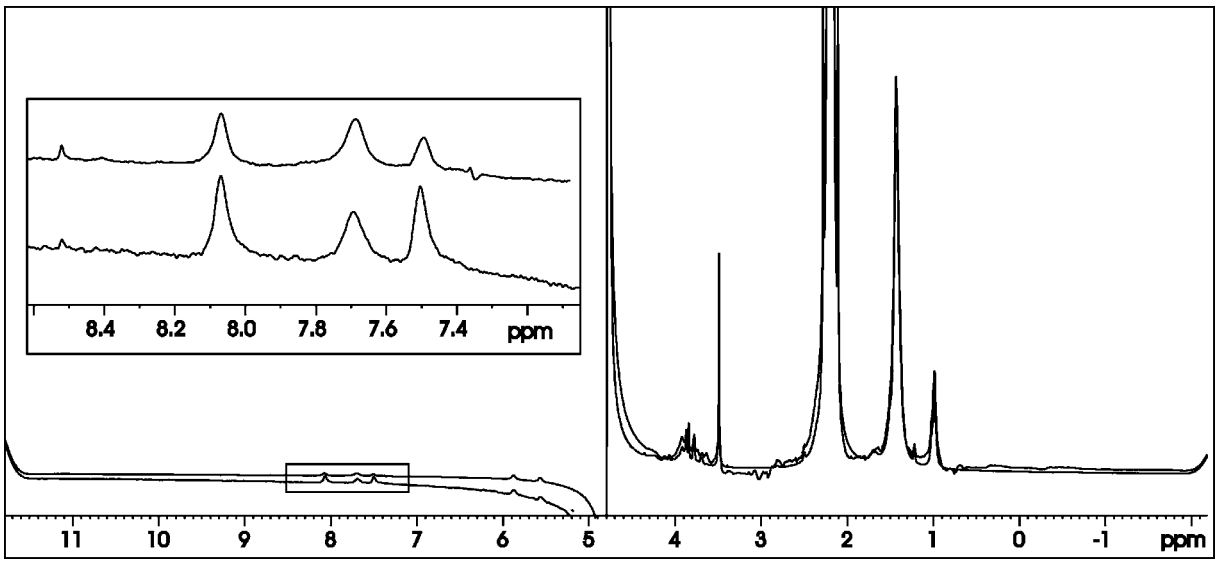


Fig. 4.7.1: ^{1}H-NMR spectra of ganglioside GM1, 480 µM in H_2O, buffered with 50 mM acetate, pH 4.5 at 42 °C and 750 MHz after 128 transients. Water suppression was achieved by presaturation (upper spectrum) and a 3-9-19 binomial watergate sequence (lower spectrum), both techniques introduce a phase distortion of the residual water signal. The strong signal at 2.16 ppm belongs to the acetate methyl group. In GM1, only the olefinic protons of sphingosine and the three amide protons are resolved. In contrast to the resonance of the sphingosine amide proton (FWHH ≈ 50 Hz) the NeuNAc and GalNAc amide peaks (FWHH ≈ 35Hz) are sharper due to flexibility in the oligosaccharide moiety, but weaken upon water presaturation, indicative of higher solvent accessibility.

Gangliosides GM1 and GM2 have been shown to bind equally well to GM2AP in a 1:1 complex with a dissociation constant of $K_D = 3.5$ µM,[135] but sphingolipids lacking sialic acid bind much weaker.[136] Measurements of K_D proved difficult due the inability to immobilise gangliosides without affecting activator binding and the fact, that GM2AP can actually accommodate higher amounts of lipids, albeit much less specific. At NMR relevant concentrations (480 µM) of an equimolar mixture (**Fig. 4.7.2**), > 90 % of GM1 molecules are expected be in the bound state.

Apart from a higher amount of carbohydrate bound protons (3-5.5 ppm) the ^{1}H-NMR spectrum of the 1:1 complex of GM2AP and GM1 is very similar to the one of the pure GM2AP and no signal could be clearly attributed to GM1. Upon diffusion- and $T_{1\rho}$- editing the complex spectrum remained unaffected and was simply scaled with respect to the gradient strength and $T_{1\rho}$-delay. From diffusion measurements, the obtained value for $D_t = 1.40 \pm 0.05 * 10^{-10}$ m^2/s is virtually identical to the free GM2AP. If larger amounts of GM1 micelles were present in solution, the appearance of diffusion edited spectra would have changed towards the spectrum of GM1 micelles, given that their diffusion coefficients differ by a factor of two.

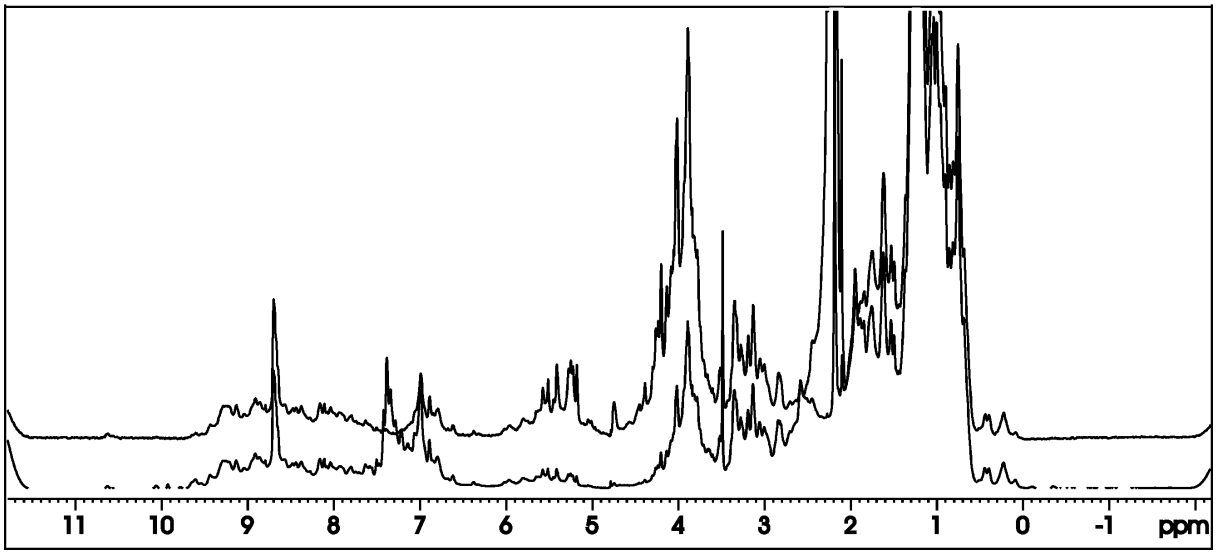


Fig. 4.7.2: ^{1}H-NMR spectra of pure GM2AP (lower spectrum) and GM2AP in the presence of an equimolar amount of ganglioside GM1 (upper spectrum) at 42 °C and 14.1 T. Both concentrations (480 µM) and buffer conditions (50 mM acetate, pH 4.5) were identical. Water suppression was achieved by a binomial 3-9-19 watergate scheme, and the buffer signal was suppressed by additional weak presaturation.

Rather strong binding could be also deduced from drastic changes in the appearance of ^{15}N-HSQC spectra in the absence and presence of GM1 (data not shown), which are much stronger than the perturbations caused by the variation of pH (**Fig. 4.4.1**). It is worth recalling the difficulties in removal of bound lipids during the purification of GM2AP, and the fact that residual electron density was found in the binding pocket during the X-ray structure refinement.[118] Despite an apparent slight improval of the protein´s thermal stability, the quality of the ^{15}N-HSQC further decreased in the presence of GM1, which might reflect the different lipid binding modes within the cavity.[119]

5 Characterisation of human saposins

5.1 The saposin family

Originally derived from *sphingolipid activator protein*, the term *saposins* (or short *SAPs*) today denote four proteins, which are functionally related to, but genetically and structurally distinct from the GM2 activator protein.[137] Essential for the *in vivo* degradation of GSLs with even shorter oligosaccharide head groups, the specificity and mode of activation differs among the individual saposins A to D. For example, saposin B facilitates the hydrolysis of the sulfate group from sulfatide by arylsulfatase A in the formation of galactosylceramide,[138] which is then catabolised to ceramide in a reaction activated by saposin C (**Fig. 4.1.2**).

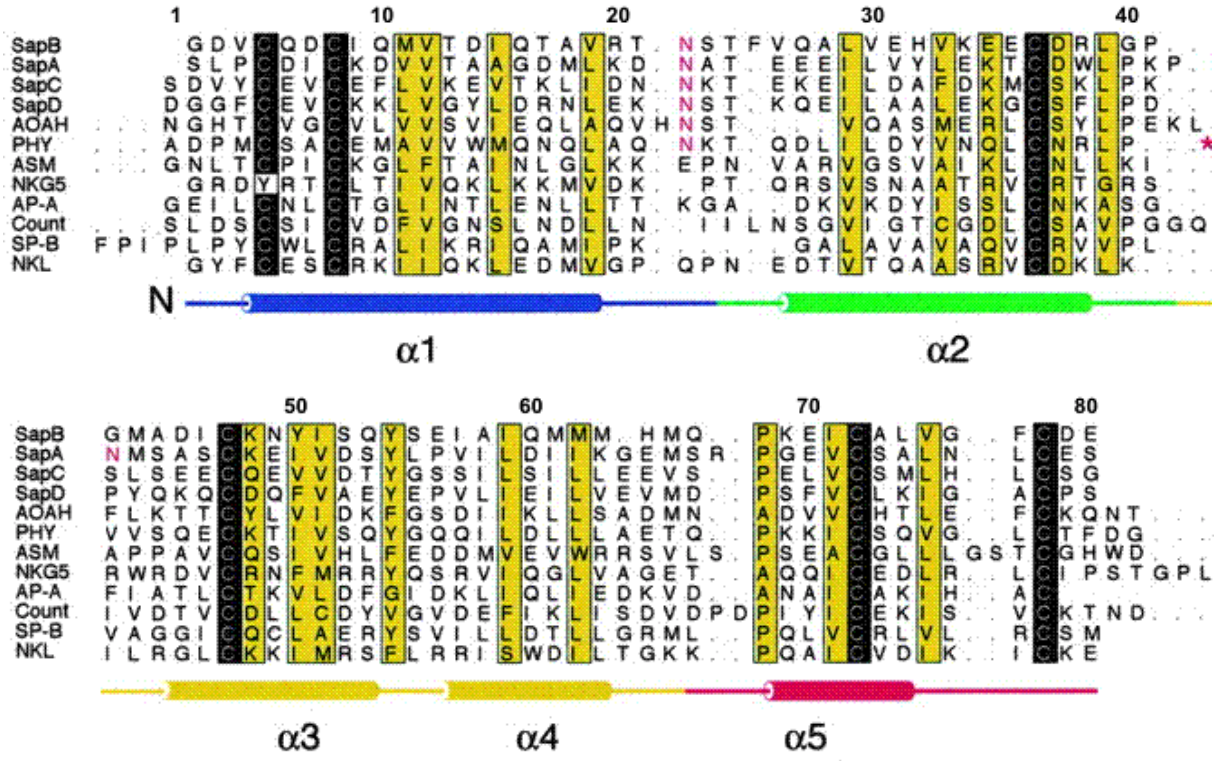


Fig. 5.1.1: Multiple alignment of representative saposin-like proteins from different functional families: Human saposins (Sap) A-D, acyloxyacyl hydrolase (AOAH), plant phytepsin (PHY), acid sphingomyelinase (ASM), granulysin (NKG5), pore forming amoebapore A (AP-A), Countinin (Count) surfactant-associated protein B (SP-B) and porcine NK-lysin (NKL). Leading and trailing dots indicate cases where the saposin motif is a sub-sequence of a larger protein. Note that in the phytepsin *swaposin* domain, the sequence block VVSQ...TFDG precedes the block ADPM...NRLP, as marked by a red asterisk. Sequence numbering is according to saposin C. The highly conserved pattern of cysteine residues forming three disulfide bonds Cys5-Cys78, Cys8-Cys72 and Cys36-Cys47 is shaded in black. Other conserved residues are shaded yellow, and experimentally determined glycosylation sites (Asn22) are indicated with red lettering. The five helices of NK-lysin are shown as cylinders. Figure taken from [139].

Saposin D has been shown to stimulate sphingomyelinase activity. In analogy to GM2AP (chapter 4), a functional defect due to mutation in one saposin molecule results in a specific phenotype of pathological GSL storage, which is only marginally bypassed by the remaining intact saposins.[112] In spite of that, saposins A-D with a size of approximately 11-12 kDa share a high sequence similarity including six strictly conserved cysteine residues and a common glycosylation site (**Fig. 5.1.1**).

The saposin motif does not only occur in a family of distinct sequence-related homologues from different functional families, but also as a sub-sequence of larger proteins, where it has been shown to adopt the same fold.[140] In some plant aspartic proteinases the sequential order of the first and second half of the saposin domain is inverted, which provided one of the first evidence for the occurrence of circular permutation during protein evolution.[141] These *swaposin domains* are easily rationalised by a peptide link between the spatially close N- and C-termini and a break between helices α2 and α3 producing new N- and C-termini. (Proximity of amino- and carboxy-termini is observed in most proteins, and it is a strict requirement for domain insertions.) Bearing in mind, that also the four saposins are produced from the single precursor pro-saposin by cleavage of linker segments, it seems likely, that both saposin and swaposin domains have evolved from a similar pro-saposin-like gene (**Fig. 5.1.2**).

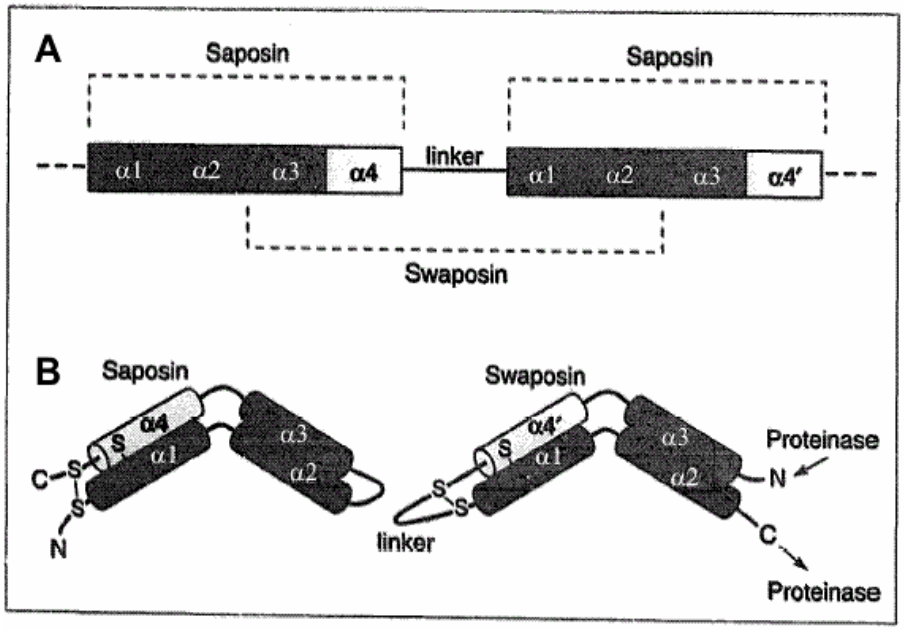


Fig. 5.1.2: A: How saposin homologues and swaposins might have evolved from a similar pro-saposin-like gene. **B:** Schematic topology of saposin and swaposin domains including disulfide bonds. In both panels helices α1-α4 correspond to the helices actually found in the structures of NK-lysin and saposin C, and short helix α5 is omitted. Figure taken from [141].

5.2 Structures of members of the saposin family

The first member of the saposin family, for which a 3D structure has been solved by NMR spectroscopy, was porcine *NK-lysin* (PDB-code 1NKL),[142] and it was the only structure available of a distinct homologue, when the work of this dissertation was started. NK-lysin is an unglycosylated effector polypeptide of T- and natural killer (NK)-cells present in CD2+, CD4+ and CD8+ cells with marked antibacterial activity and the capability to lyse tumour cell lines. The fold is a relatively compact monomer made up from five α-helices, which has also been seen in crystals of the vacuolar-targeting *swaposin* domain of *pro-phytepsin* (PDB-code 1QDM).[140] Against a long helix a1 a slightly bend helix α2 and α3 are packed from one side, and helices α4 and α5 from the other side (**Fig. 5.2.1 A**). Despite lack of sequence similarity and stabilising disulfide bonds, a similar and extremely stable fold is also adopted by the antimicrobial, 70 residue cyclic peptide *bacteriocin AS-48* (PDB-code 1E68).[143]

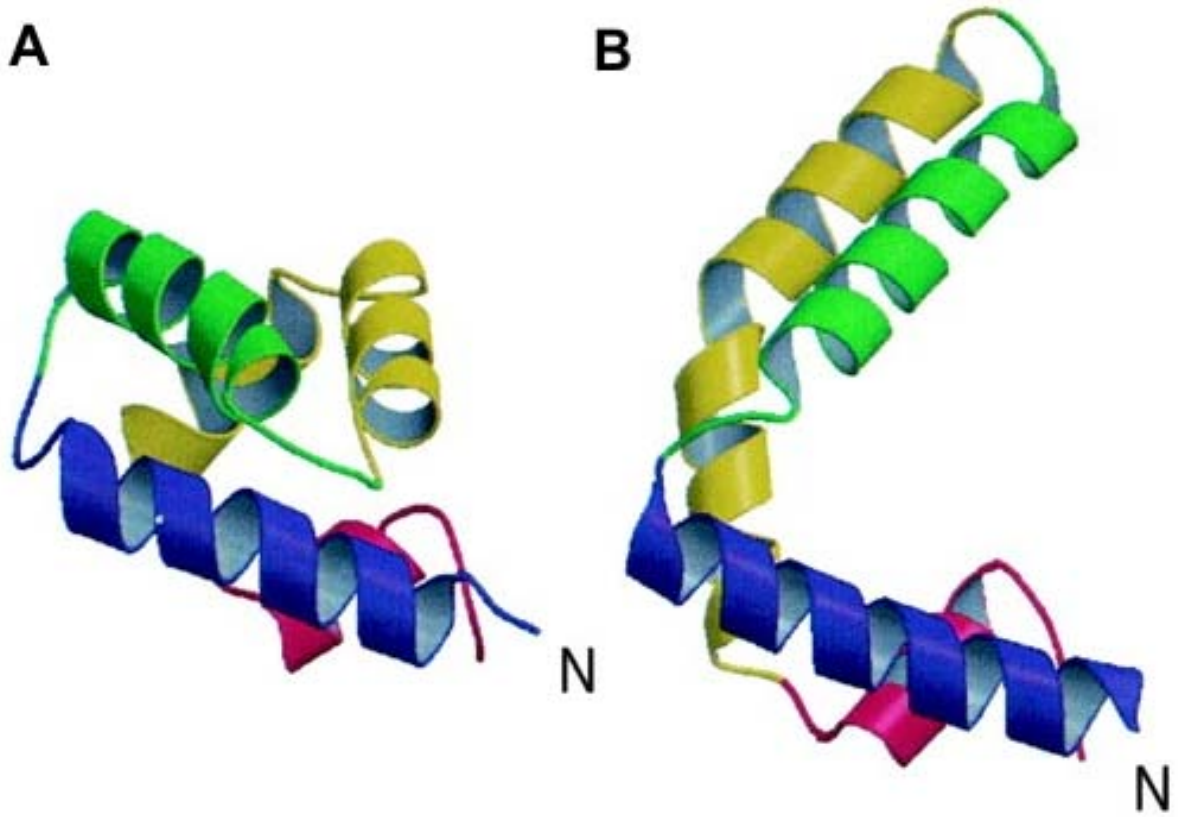


Fig. 5.2.1: Structures of members of the saposin family as ribbon diagrams. **A:** Compact *monomeric* fold of NK-lysin (PDB-code 1NKL) [Liepinsh97] comprising helices α1 (blue), α2 (green), α3, α4 (both yellow) and α5 (red). **B:** V-shaped chain B of the homodimeric, but asymmetric *dimer* of saposin B (PDB-code 1N69) [Ahn03] with the same colour coding. The N- termini are marked. Figure taken from [139].

Recently, crystallisation of *saposin B* at pH 5.8 and in the absence of glycosylation has revealed a homodimeric structure forming a shell around a large hydrophobic cavity.[139] Although the secondary structure is similar to the known compact monomeric members of the saposin family, the helices are repacked in a different tertiary arrangement making up two V-shaped monomers. The differences are largely caused by two points of variability: a less sharp turn between α1 and α2, and a merely minor kink in α3 instead of two approximately orthogonal helices α3 and α4 (**Fig. 5.2.1 B**).

Very recently, the solution structure of likewise unglycosylated *saposin C* has been determined by NMR spectroscopy at pH 6.8, and it shows nearly exactly the monomeric fold expected from NK-lysin.[144] Although both proteins possess membrane binding capability, their surface electrostatic potential is remarkably different, being highly negative for saposin C and mostly positive for NK-lysin. Within the saposins, the theoretical pI varies from 4.2 (saposin A) to 4.7 (saposin D), with a difference of negatively (aspartate and glutamate) and positively (lysine and arginine) charged residues of 8 and 5, respectively. A C-terminal hexa-histidine tag increases the theoretical pI by approximately one unit and may therefore significantly influence intermolecular repulsion and propensity of dimerisation. In contrast, NK-lysin is strongly basic with a pI of 9.2 and an inverted ratio of charged residues and positively charged residues at numerous positions, where the saposins carry conserved negatively or uncharged residues.

Similar to GM2AP, maximum activity of saposins occurs under acidic conditions, and lipid [145] and membrane [146] binding properties have been reported to strongly vary with pH and assigned to conformational variability. Comparative studies of saposins A-D in solution at varying pH by NMR spectroscopy should give further insight into the structural aspects of substrate and pH specificity of saposin activity. One important question was, whether the capability of dimerisation observed for saposin B is a feature inherent to all saposins, and under which conditions it occurs. All NMR data presented in later sections will be discussed in terms of protein models derived from existing structural data. Therefore, the following two sections give an overview of modern comparative protein modelling and its application to saposins.

5.3 Comparative protein modelling

Proteins from different sources and sometimes diverse biological functions can have similar sequences, and it is generally accepted that high sequence similarity is reflected by distinct structure similarity. Indeed, the root mean square deviation (RMSD) of C^{α} coordinates for protein cores sharing 50 % residue identity is expected to be around 1 Å. The idea, that structure is conserved to a much greater extend than sequence and that there is a limited number of backbone motifs [147] often allows to predict the structure of a given *target* sequence by comparing it to proteins (= *templates*), whose structure has explicitly been determined by X-ray crystallography or NMR spectroscopy. Predictive methods have gained much interest, since the number of solved 3D structures increases only slowly compared to the rate of sequencing novel cDNAs. It has to be emphasised, however, that the results of predictive methods do by no means provide the precision of experimental structures and should be regarded as *low-resolution models* guiding the design of actual experiments.

In principle, a model can be built for any query sequence by moving it position by position through the structure of a known protein expected to share the same fold, and computing the thermodynamically most favourable sum of pair-wise interactions between residues.[148] However, the current requirements of these *threading* methods in terms of both hardware and expertise have proven to be obstacles to most structural biologists. With the increasing amount of structural information available, it becomes increasingly likely that high sequence similarities are detected in a database search at least for sections of the target. The first and crucial step of modern modelling methods is therefore the identification of templates and optimising the sequence alignment, followed by model building and refinement.

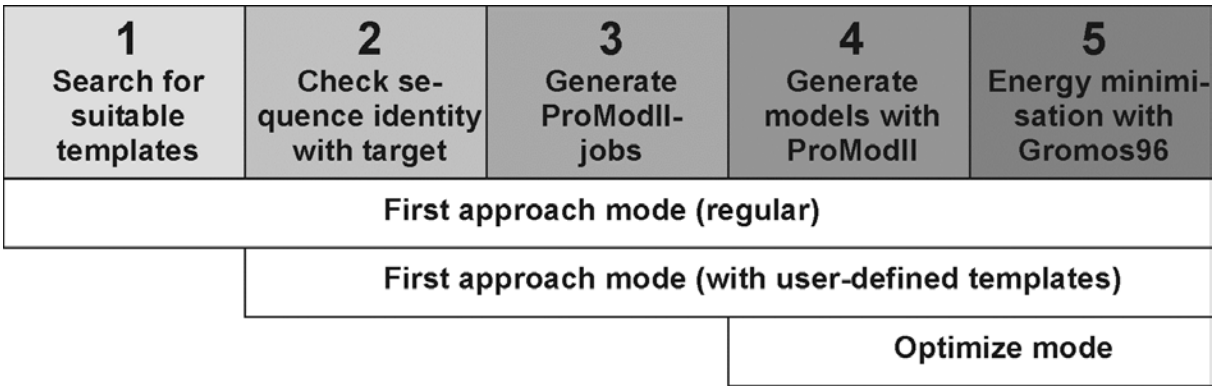


Fig. 5.3.1: Scheme of operation of the SWISS-MODEL server.[149]

One of the powerful servers meanwhile available for fully automated comparative protein modelling is the *SWISS-MODEL* server (**Fig. 5.3.1**).[149] The target sequence is submitted via a web interface, and the results of the modelling procedure including coordinates of the final model in the PDB format is returned via e-mail. First, a *BLAST* search against *ExPdb*, the sequence database derived from the Brookhaven Protein Data Bank, selects sequences yielding a poisson unlikelyhood probability $P(N) < 10^{-5}$ and sharing at least 25 % sequence identity in a > 20 residue subsequence. $P(N)$ specifies the probability that the actual alignment score between two sequences of length N is generated by chance alone.

A framework for the atoms in the backbone is constructed by averaging the positions in the selected template structures, weighted with the local degree of sequence identity, by the program *ProModII*. Non conserved loops are added by defining stems of usually two to three residues in both directions, followed by a selection of database motifs, for which the C^{α} positions match the already constructed framework within a given cut-off RMSD. Conformations of conserved side chains are directly taken over from the best template, and for the others the most favourable rotamers in terms of van-der-Waals exclusion are accepted. Idealisation of bond geometry and removal of unfavourable non-bonded contacts is performed by a force field energy minimisation for a limited number of steps to avoid excessive structural drifts.

5.4 Modelling of the saposins

Throughout this thesis residue numbering for all saposins is adopted to saposin C (**Fig. 5.1.1**), which occupies the sequence positions Ser311 ("Ser1") to Gly390 ("Gly80") of the common precursor pro-saposin. Saposin D, from Asp405 ("Asp1") to Ser484 ("Ser80") can be aligned to saposin C without gaps, whereas in saposin A (Ser60 = "Ser2" to Ser140 = "Ser80") a proline after "Lys41" and an arginine after "Ser67" are inserted, and in saposin B (Gly195 = "Gly2" to Glu273 = "Glu80") a phenylalanine is inserted after "Thr24", and position "64" is missing. Sequence identity within the saposins ranges from 15 % (B-C) to 39 % (A-C), increasing to 47 % and 60 %, if amino acids with similar properties are considered (**Fig. 5.4.1**).

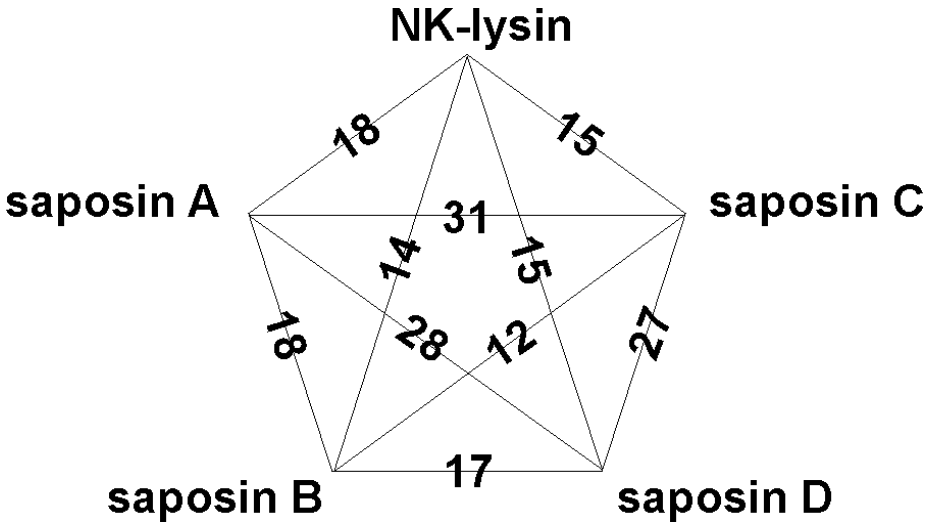


Fig. 5.4.1: Diagram elucidating the sequence identity of saposins A-D and NK-lysin, referring to a total of 80 amino acids. Non conserved residues in the pairwise alignment with similar properties are not considered.

Before the structures of saposin C (PDB-code 1M12) and saposin B (1N69) were accessible in the protein data bank, NK-lysin (1NKL) was the only distinct homologous protein with known structure. Back then, saposin query sequences had to be adapted to NK-lysin in positions believed to be of only minor concern in the structure, in order not to be rejected by the SWISS-MODEL server in advance. In saposin C, for example, the necessary substitutions were: D2G, E69Q, S79K and G80E. Even with these modifications the sequence alignment remained incorrect around residue Pro40, which is a gap in NK-lysin, and all respective residues were shifted by one position in the model structure, thereby disrupting the disulfide bond Cys36-Cys47. Therefore Pro40 was removed in the query, yielding a correct ungapped alignment, but, of course a model with this residue missing. Finally, residues substituted prior to submission were corrected with a suitable structure editor.

```
TARGET   1       SLPCDIC KDVVTAAGDM LKDNATEEEI LVYLEKTCDW LPKPNMSASC
1m12A    1    S--DVYCEVC EFLVKEVTKL IDNNKTEKEI LDAFDKMCSK LPKS-LSEEC
                   . *..*   .*     . . .* ** ** *  ..* *   ***. .*  *
TARGET             hhhhh hhhhhhhhhh hh    hhhh hhhh             hhhh
1m12A              hhhhh hhhhhhhhhh hh    hhhh hhhh              hhh

TARGET   48   KEIVDSYLPV ILDIIKGEMS RPGEVCSALN LCESLQ -
1m12A    48   QEVVDTYGSS ILSILLEEVS -PELVCSMLH LCSGLVPR
               *.**.* .  ** *.  *.*  *  *** *. ** .*
TARGET         hhhhh   hh hhhhh          hhhh
1m12A          hhhhh   hh hhhhh        hhhhhh
```

Fig. 5.4.2: *BLASTP* sequence alignment of saposin A (target) to saposin C (template). Identical positions and amino acids with similar properties are marked by stars and points, respectively. Predicted target and template secondary structures are indicated (h = helix).

Since sequence similarity is much higher within the saposins, current direct queries of saposin A and D including the C-terminal tag RH_6 yield correct alignment with saposin C without further modifications (**Fig. 5.4.2, Table 5.4.1**). Saposin B was selected as template for neither protein. No extra loops had to be constructed for saposin D, and for saposin A two loops around the asparagine following Pro40 and arginine following Ser67 were successfully built using Pro40-Ser44 and Glu65-Gly69 as anchoring residues, respectively. In all cases energy minimisation was done using the GROMOS96 force field with 200 cycles of steepest descend and 300 cycles of conjugate gradient. All model structures were checked by overlaying target and template, hydrogen atoms were added and named according to the IUPAC standard (HN, HA, ...), and residue numbering was changed, if necessary, to match saposin C.

Table 5.4.1: Saposin models generated with the SWISS-MODEL FirstApproach mode.

target	template	sequence identity	P(N) of alignment	name of model
saposinC(mod)	1NKL	27 %	$2*10^{-6}$	*SapC_1nkl_smh*
saposin A	1M12	39 %	$8*10^{-11}$	*SapA_1m12_smh*
saposin D	1M12	34 %	$3*10^{-10}$	*SapD_1m12_smh*

5.5 Backbone assignment strategy

According to the present standard, proteins of 11-12 kDa size require ^{15}N, ^{13}C doubly labelled material for efficient backbone assignment using the triple resonance experiments described in chapter 3. The older *Wüthrich* strategy [150] to use exclusively NOE based $d_{NN}(i,i+1)$ and $d_{\alpha N}(i,i+1)$ correlations for sequentially connecting the amino acid spin systems is not straightforward, because these correlations are often weak and difficult to distinguish from intermediate and long range correlations. Furthermore, considering the high degeneracy in a set of only pairs of proton frequencies, in completely unlabelled proteins this work can become very tedious.

However, in α-helices the steadily short $d_{NN}(i,i+1)$ distances of about 2.8 Å usually provide two strong sequential NOEs in the strip of a given amide group. In principle, in a purely α-helical protein, labelled with ^{15}N, complete ^{15}N-HSQC assignment is possible, using only a single 4D doubly ^{15}N-edited NOESY experiment.[151] Alternatively, if the resonances are sufficiently spread in the ^{15}N-HSQC plane, the correlations to residues (i−1) and (i+1) can be observed separately in the ω_1-dimensions of single (HNH)[107] and double (NNH)[108] ^{15}N-edited 3D-NOESY experiments (**Fig. 5.5.1**).

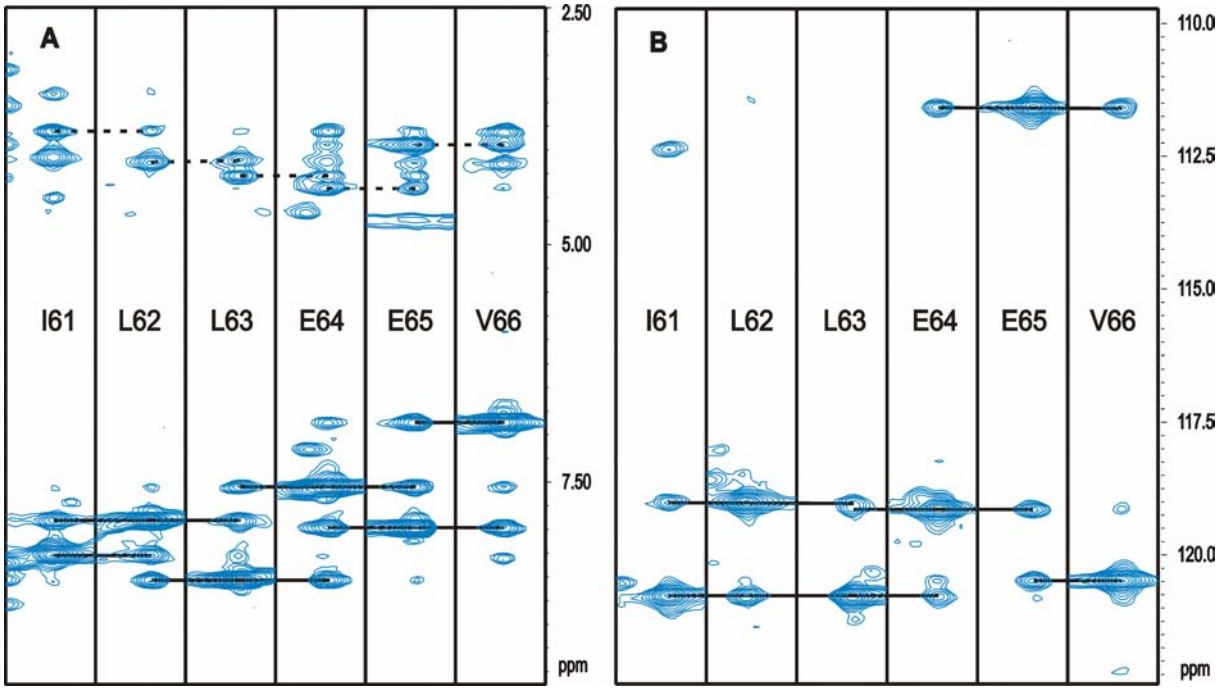


Fig. 5.5.1: ω_1-strip plots of HNH-NOESY (**A**) and NNH-NOESY (**B**) experiments (120 ms mixing) for the helical segment I61-V66 of saposin C. Straight and dashed horizontal lines connect sequential $d_{NN}(i,i+1)$ and $d_{\alpha N}(i,i+1)$ correlations. Acquisition times were 21.5 ms (HNH-NOESY) and 123.6 ms (NNH-NOESY) in ω_1, 34.8 ms (LP to 61.8 ms) in ω_2 (^{15}N) and 86.0 ms in ω_3 (H^N) dimensions.

After a few residues have been successfully connected to a segment, its direction in the amino acid sequence can be fixed using $d_{\alpha N}(i-1,i)$, which is much stronger than $d_{N\alpha}(i,i+1)$, for a residue where $H^{\alpha}(i)$ and $H^{\alpha}(i-1)$ are not overlapped. If not obvious from the HNH-NOESY spectrum, the HNHA spectrum provides the necessary $H^{\alpha}(i)$ resonance frequencies.[105] Mapping of the segment to the sequence is achieved by identifying potential amino acids by comparison of side chain resonance frequencies, obtained from a HNH-TOCSY,[152] with tabulated random coil values.[100] In many cases the H^{β} resonances need to be distinguished from other side chain protons (or from H^{α} protons for serine and threonine residues) by another HNHB experiment.[153] The remaining, mostly non-helical residues can be found via intermediate and long-range NOE correlations, with the exception of extremely fast solvent exchanging residues, for which frequency labelled magnetisation is usually quenched during the mixing times.

Table 5.5.1: Resonance assignment of saposins, specifying the fraction of non-proline residues (excl. RH_6 tag) with assignment of ^{15}N, H^N and H^{α}. NOESY and (CleanCITY-)TOCSY[154] mixing periods were 120 and 80 ms, respectively. Because of overlap and a considerable number of fast slvent exchanging protons in saposin D, assignment at 37 °C was supplemented by another set of experiments recorded at 17 °C, and the assignment was transferred to 37 °C by ^{15}N-HSQC spectra in steps of 5 °C.

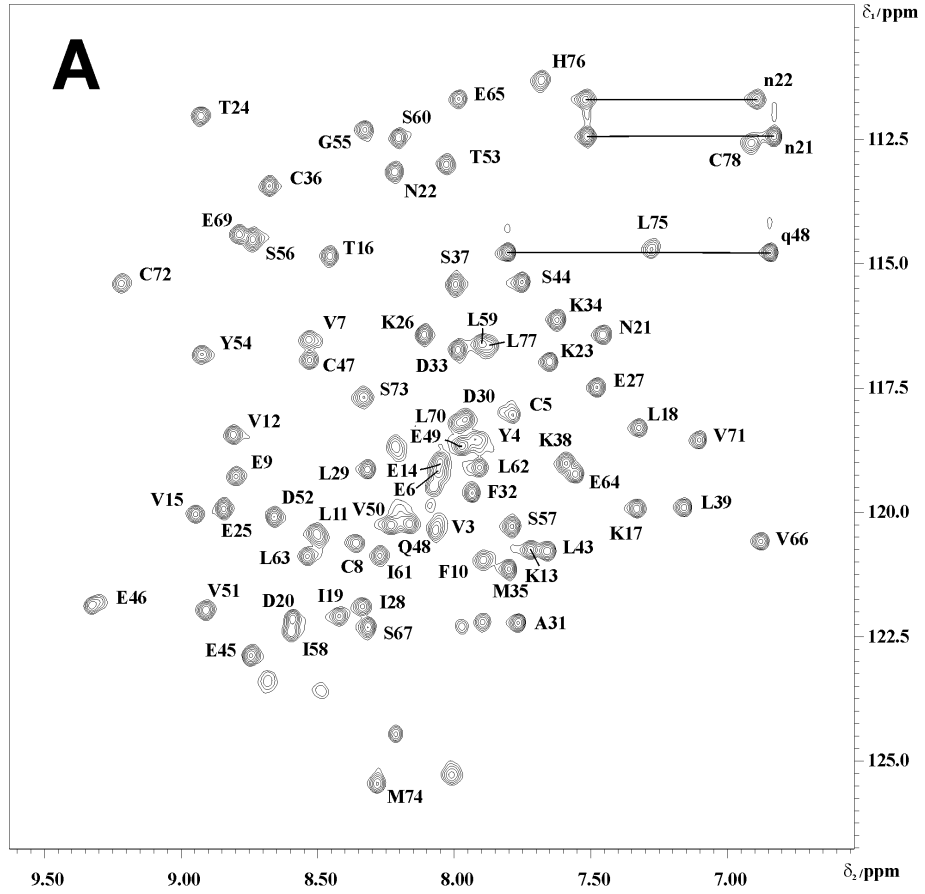
protein	buffer(50 mM)	temperature[°C]	field [T]	fraction
saposin A	phosphate pH7	37	14.1	69 / 78
saposin C	phosphate pH7	37	14.1	72 / 78
saposin C	acetate pH4	17 (monomer)	14.1	77 / 78
saposin C	acetate pH4	57 (dimer)	14.1	68 / 78
saposin D	phosphate pH7	37 (17)	11.7 (14.1)	70 / 77

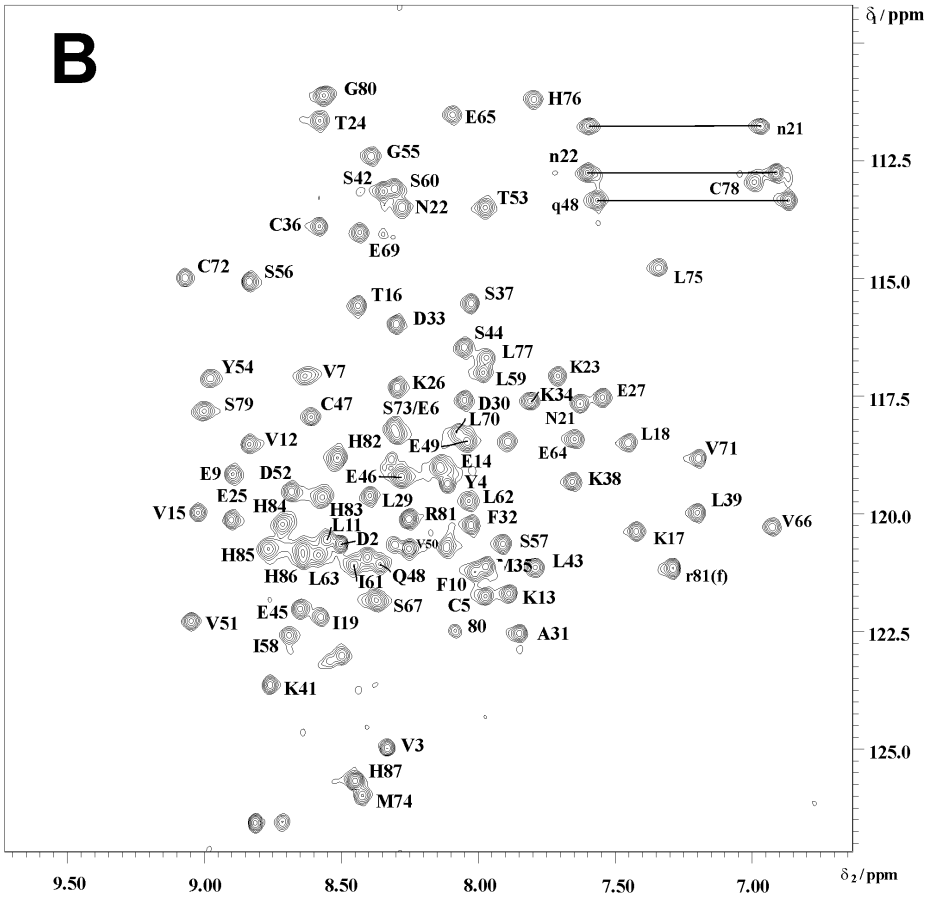
This strategy has so far been applied to ^{15}N-labelled samples of saposins A, C and D at pH 7.0, and to saposin C at pH 4.0, samples of which were provided by the group of K. Sandhoff. The protein sequences as in **Fig. 5.1.1** were furnished with the C-terminal tag RH_6, expressed in *Pichia pastoris* cells and purified under conditions similar to those described in section 4.3 for GM2AP.[122] For each sample, the total required measurement time for five 3D spectra was about one week and ^{15}N-HSQC assignment was usually nearly complete within one further week (**Table 5.5.1** and **Tables 8.1** to **8.6**).

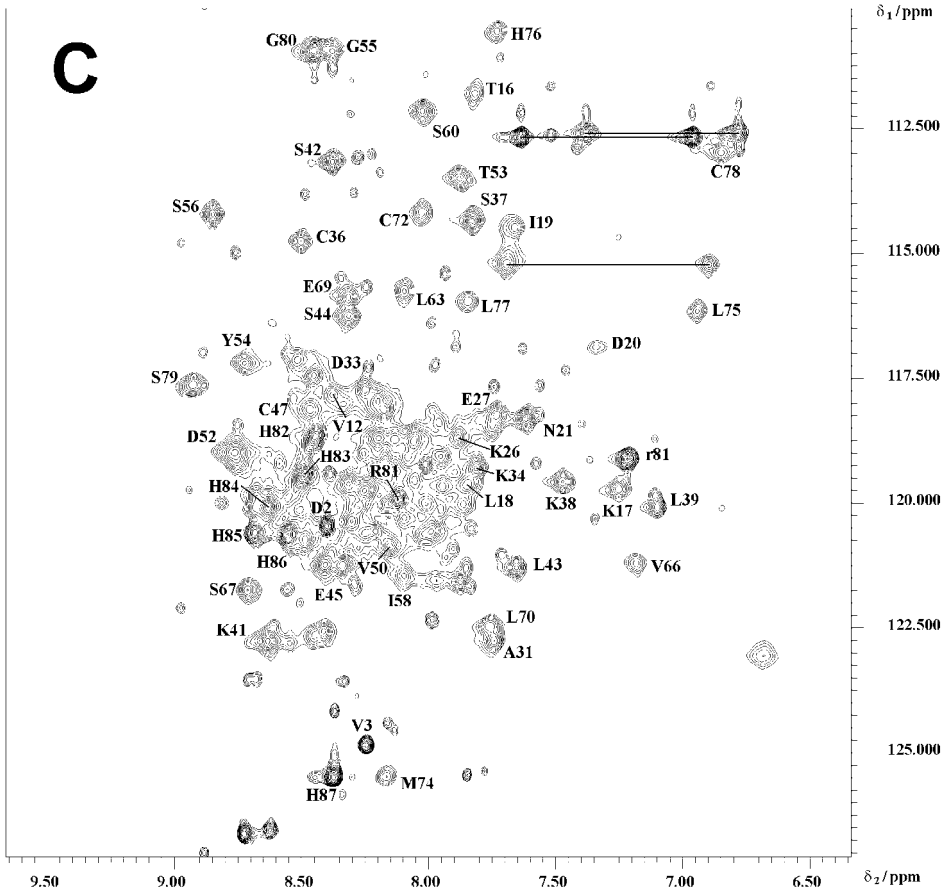
5.6 Saposin C at pH 7

Stabilised by three disulfide bonds, all saposins are thermally extremely robust proteins at pH 7.0 with absolutely constant ^{15}N-HSQC spectra after heating to 57 °C for several hours, 37° C for several weeks or storage in the fridge for now more than one year. Saposin C is readily concentrated up to ~4 mM without showing significant propensity to unspecific association, as concluded from unchanged linewidths of proton resonances during the concentration process. Typical of a purely α-helical protein is the appearance of the ^{15}N-HSQC spectrum (**Fig. 5.6.1 A**), where all amide peaks resonate within less than 14.5 ppm in the ^{15}N-, and 2.0 ppm in the H^N-dimensions. Peak resolution is satisfactory already at 17 °C and further improved at 37° C, which was kept fixed in all following measurements as the maximum temperature without major effects imposed by hydrogen exchange. All three side chain NH_2 groups of Asn21, Asn22 ans Gln48 are visible and could be assigned using the HNH-NOESY spectrum, and the H^ε proton of Arg81 disappears beyond 27° C. The ^{15}N-HSQC compares well with a spectrum reported by de Alba *et al.*[144] (**Fig. 5.6.1 D**).

The secondary structure was analysed in the usual manner (sections 3.5 to 3.7) using the H^α secondary chemical shifts, $^3J_{HNH\alpha}$ coupling constants (both from the HNHA experiment) and sequential $d_{NN}(i,i+1)$ and $d_{\alpha N}(i,i+1)$ NOEs (from the HNH-NOESY experiment). Similar to NK-lysin, five α-helices (Tyr4-Asn21, Glu25-Leu39, Ser44-Tyr54, Ser56-Glu64 and Glu69-Leu75) with rather short connecting loops were found (**Fig. 5.6.2**). From NewMEXICO experiments, fast amide exchange could be confirmed for residues Val3, Tyr4, Glu25, Glu45, Ser56 and Ser 57, which are expected either in the loops or in the first turn of each helix. Further exchanging H^N protons are the ones of Asp33, Lys34, Ser37 and Lys38 located in α2, where also the secondary structure indicators predicted a kink between residues Lys34 and Met35. Except for the termini Ser1-Asp2 and Ser79-His87 the only unassigned non-proline residues were Lys41 and Ser42 in the turn between α2 and α3, which are believed to be highly solvent exposed. As can be seen from ^{15}N R_1 and R_2 relaxation rates and the ^{15}N{^{1}H}hetNOE (**Table 8.4**), the molecular tumbling is only slightly anisotropic (R_2/R_1 = 4.4 ± 0.1 in α1 vs. 4.2 ± 0.1 in α3) with little mobility in the loop regions, constituting a compact, spherical fold. A similar behaviour of the relaxation parameters has been found by de Alba *et al.* [144] and was therefore not further analysed.


A
δ1/ppm
δ2/ppm
9.50
9.00
8.50
8.00
7.50
7.00
112.5
115.0
117.5
120.0
122.5
125.0
H76
n22
E65
T24
S60
G55
T53
n21
C78
C36
N22
E69
L75
q48
S56
T16
S37
C72
S44
K34
V7
K26
L59
N21
Y54
L77
K23
C47
D33
E27
S73
D30
C5
V12
L70
L18
E49
Y4
K38
V71
E9
L29
E14
L62
E64
E6
F32
V15
D52
L39
L11
V50
S57
E25
V3
K17
V66
L63
Q48
L43
C8
I61
F10
K13
E46
V51
D20
I19
I28
M35
S67
A31
E45
I58
M74


B
δ1 / ppm
δ2 / ppm
9.50
9.00
8.50
8.00
7.50
7.00
112.5
115.0
117.5
120.0
122.5
125.0
G80
H76
T24
E65
n21
G55
n22
S42
S60
C78
T53
N22
q48
C36
E69
C72
S56
L75
T16
S37
D33
S44
L77
Y54
V7
K23
K26
L59
K34
E27
S79
C47
S73/E6
D30
N21
L70
V12
H82
E49
L18
E64
V71
E9
D52
E46
E14
Y4
K38
E25
H84
L29
L62
L39
V15
H83
R81
F32
V66
L11
K17
H85
D2
V50
S57
L43
M35
H86
L63
I61
Q48
F10
r81(f)
E45
S67
C5
K13
V51
I19
80
A31
I58
K41
V3
H87
M74


C
G80
G55
H76
T16
S60
S42
T53
C78
S56
S37
C72
I19
C36
E69
L63
L77
L75
S44
Y54
D20
D33
S79
E27
C47
V12
N21
D52
H82
K26
r81
H83
R81
K34
H84
L18
K38
L39
D2
K17
H85
H86
V50
L43
V66
S67
E45
I58
L70
K41
A31
V3
M74
H87
δ1/ppm
112.500
115.000
117.500
120.000
122.500
125.000
9.50
9.00
8.50
8.00
7.50
7.00
6.50
δ2/ppm

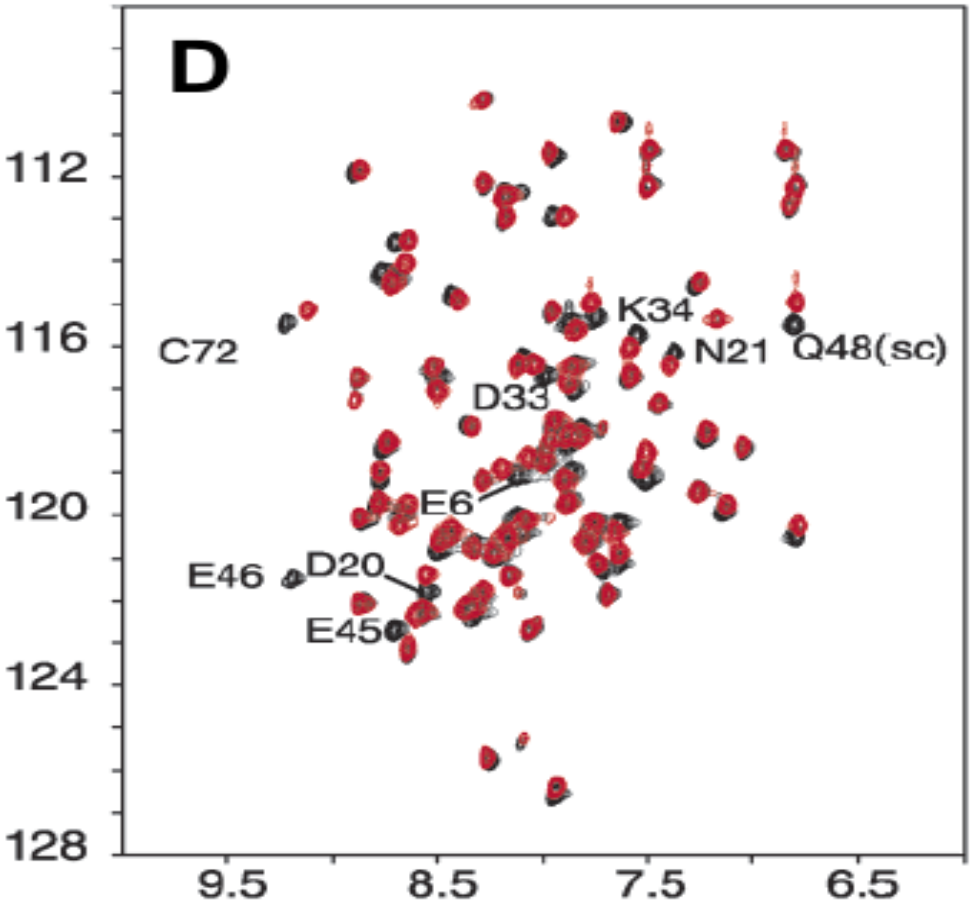

D
112
116
120
124
128
C72
K34
N21
Q48(sc)
D33
E6
E46
D20
E45
9.5
8.5
7.5
6.5

Fig. 5.6.1 (previous two pages): Assigned ^{15}N-HSQC spectra of saposin C at 37 °C and 14.1 T, recorded with acquisition times of 61.8 ms (extended to 92.7 ms) and 86.0 ms in the ^{15}N and H^N dimensions, respectively. **A:** 1.2 mM in 50 mM phosphate, pH 7.0. **B:** 1.2 mM in 50 mM acetate, pH 4.0, directly after heating to 37°C. **C:** same sample as in B, but after several hours at 37 °C. **D:** comparison with published overlay of ^{15}N-HSQC spectra at pH 6.8 (black) and 5.4 (red).[144]

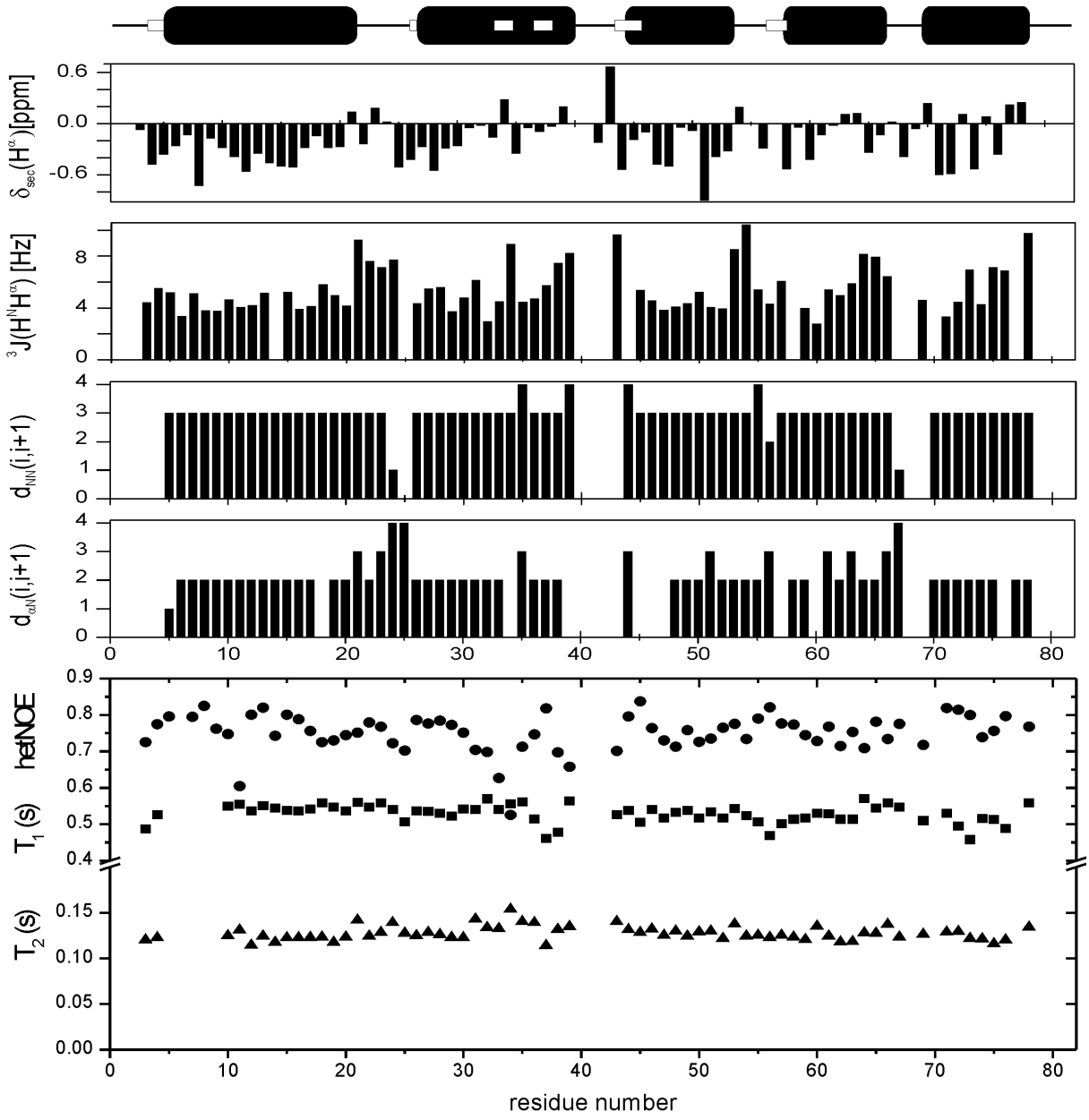


Fig. 5.6.2: Sequence plot of secondary chemical shifts $\delta_{sec}(H^{\alpha})$, coupling constants $^3J_{HNH\alpha}$ (calculated according to section 3.6), $d_{NN}(i,i+1)$ and $d_{\alpha N}(i,i+1)$ NOEs, both classified into 1 (weak, < 5 Å) 2 (medium, < 4 Å) 3 (strong, < 3 Å) and 4 (very strong, < 2.2 Å), ^{15}N-R_1 and -R_2 relaxation rates and the ^{15}N{^{1}H}hetNOE. The five helices of NK-lysin are shown as cylinders on top with white bars denoting residues with fast (k_{ex} > 1 s^{-1}) amide exchange.

5.7 Saposin C at pH 4

In the ^{15}N-HSQC spectrum of a pH 4 sample, recorded directly after inserting it into the magnet at 37 °C (**Fig. 5.6.1 B**), signals no longer suffer from fast hydrogen exchange, so that now all non-proline residues from Asp2 to His87 become visible and were successfully assigned. Although pH 7 and pH 4 spectra can not be overlayed, there is good evidence for an identical structure because of a close correlation between peak positions for most residues. Indeed, the strongest pH dependence of resonance frequencies was observed for the amide groups of Cys5, Thr24, Glu46 and Glu69, which are likely to be H-bonded to the side chain carboxylate groups of Asp2, Glu27, Glu46 and Glu69, respectively. Since in saposin C most carboxylates have been shown to titrate with a pK_a in the range of 5-5.5,[144] they are certainly in a differently charged state at pH 4 and 7.

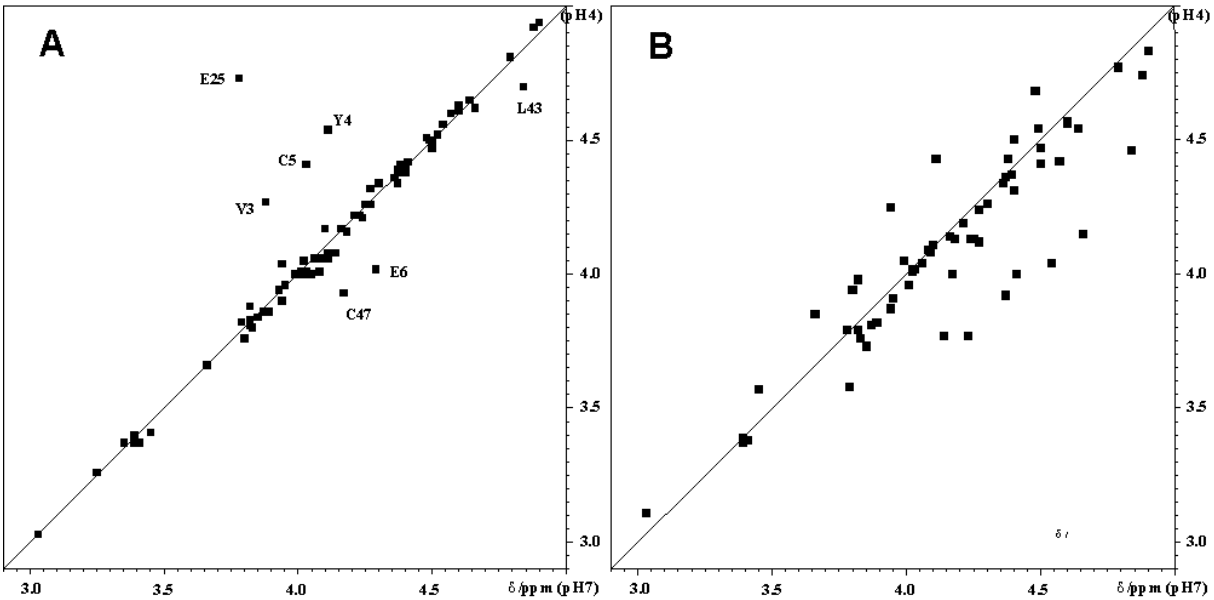


Fig. 5.7.1: Comparison of H^{α} chemical shifts in saposin C. **A:** pH 4.0 and 17 °C versus pH 7.0 and 37 °C. **B:** pH 4.0 and 17 °C (monomer) versus pH 4.0 and 57 °C (dimer).

This is supported by a similarity of $\delta(H^{\alpha})$ chemical shifts (**Fig. 5.7.1 A**), which are much less prone to small perturbations of electrostatic interactions than those of H^N and ^{15}N. Variations observed at the N-terminus, in Glu25, Leu43 and Cys47 may as well be a local conformational perturbation induced by the different temperatures at which $\delta(H^{\alpha})$ were measured. After a few hours at 37° C the appearance of the ^{15}N-HSQC spectrum completely changes, and only 10-15 % of the intensity of the original peaks is retained, whereas new and much broader signals emerge (**Fig. 5.6.1 C**). This happens faster with increasing temperature,

but could not be reversed upon cooling in the fridge even after several weeks. ^{15}N-HSQC assignment was done at 17 °C for the first and 57 °C for the second species, where both slowly degrade within about two weeks, and transferred to 37 °C in steps of 5 °C. Molecular correlation times τ_c of the "fresh" protein and at pH 7 (**Table 5.7.1**) are similar, monomeric, albeit 40-50 % larger than reported [144] and modelled for an untagged and unglycosylated species (see section 4.5). At higher concentrations rather unspecific association leads to an apparent increase of τ_c and linewidths in an otherwise identical ^{15}N-HSQC spectrum. In contrast, a nearly doubled τ_c for the second species in combination with the remarkable changes in the ^{15}N-HSQC spectrum can only be explained by specific dimerisation.

Table 5.7.1: Molecular size M_r and correlation times τ_c of saposin C at 37 °C, estimated from ^{15}N-R_1 and R_2 (averaged over all residues). The results are compared with the reported value [144] and with HYDONMR-simulations [130] of monomeric NK-lysin [142] and the ab-dimer of saposin B.[139]

concentr.	1.2 mM	1.2 mM	1.2 mM	2.1 mM	0.25 mM	HDmodel	HDmodel
Buffer	pH 7.0	pH 4.0 [a]	pH 4.0 [b]	pH 4.0 [a]	pH 6.8	1NKL [d]	1N69 [d]
M_r [kDa]	12-13	12-13	24-26	12-13	9	9	18
τ_c [ns]	5.2 ± 0.2	5.5 ± 0.5	9.5 ± 0.5	7.0 ± 0.5	3.6 [c]	3.6	7.2

[a] directly after heating to 37 °C. [b] several h after heating to 37° C. [c] corrected for solvent viscosity from 25 °C to 37 °C by a factor of 0.78. [d] 3.2 Å atomic bead radius.

Apart from an absolutely identical ^{15}N-HSQC signal pattern produced by the histidine tag, most amide resonances of residues in the „monomer" and „dimer" are found in close vicinity, and their H^{α} resonance frequencies coincide (**Fig. 5.7.1 B**). However, there are large chemical shift perturbations for the second half of α1, the loop connecting α1 and α2, and helices α4 and α5, and the N-terminal half of α1 could not be assigned due to strong overlap or broad lines. These sites are too far apart from each other to be explained by a single dimerisation interface without major rearrangement of helices. On the other hand, for a V-shaped dimer like in saposin B, strong chemical shift perturbations would be expected between α3 and α4, which were not observed. Here, an approach using RDCs of H^N-N vectors and comparison with different structural models could be useful for the determination of relative helix orientations. Acid triggered dimerisation has not yet been reported for saposin C and may, far beyond simple neutralisation of the surface charge, play an important role in its *in vivo* membrane and GSL binding activity.

5.8 Saposins A and B

The ^{15}N-HSQC spectra of saposin A at pH 7.0 (**Fig. 5.8.1**) showed symptoms of conformational heterogeneity, with most resonances affiliating to slow and intermediate exchange regimes, and best spectrum quality at lower field strength. One mode of conformational flip processes is likely to be similar to the one in saposin D (chapter 7), since a second signal set with pronounced differences in the resonance frequencies was observed for the same residues (Ser44, Ala45 and Ser46). The molecular correlation time of τ_c = 5.4 ns, calculated from the ratio R_2/R_1 = 4.0 of average ^{15}N relaxation rates at 37 ° and 14.1 T, compares with the monomeric species of saposin C.

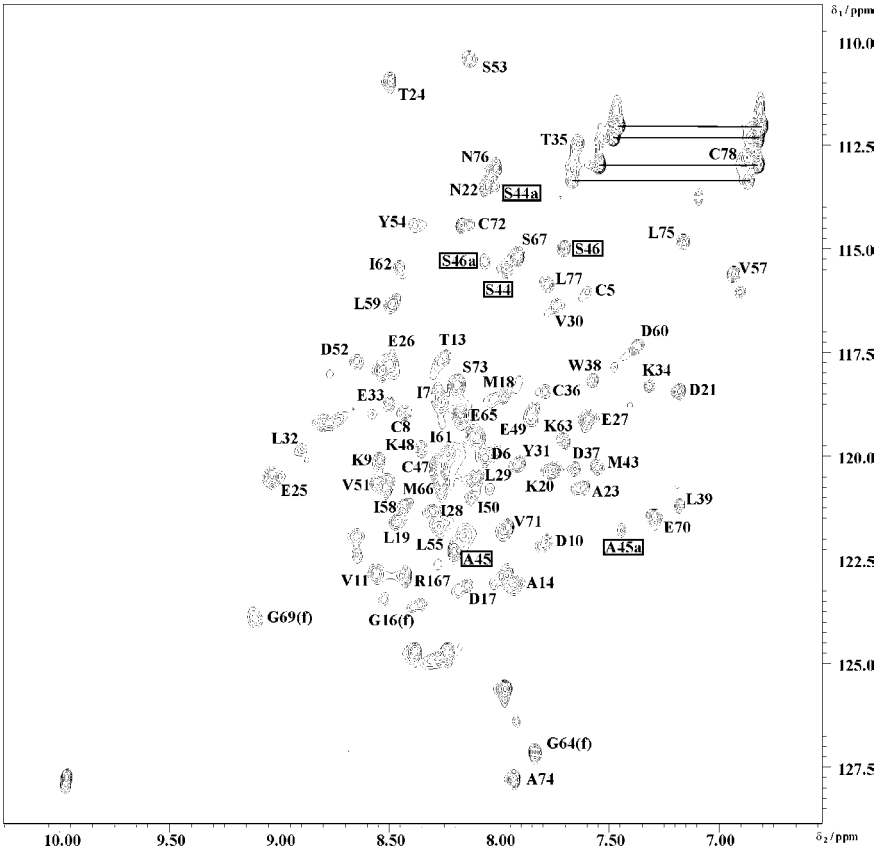


Fig. 5.8.1: ^{15}N-HSQC spectrum of saposin A, 0.6 mM in 50 mM phosphate, pH 7.0, at 37 ° and 14.1 T.

Saposin A is the only saposin bearing a Trp residue, whose aromatic ε-NH group appears in the lower left corner. Amide groups of Ser44, Ala45 and Ser46 are marked with boxes, and folded Gly residues with (f).

At pH 4.0, saposin A is strongly aggregated and slowly precipitated from solution even at concentrations as low as 0.5 mM, where the only resolved peaks in the ^{15}N-HSQC spectra were those from the RH_6 tag. Saposin B was sufficiently thermally stable under acidic and neutral conditions, but yielded spectra with extremely broad lines, again the only exception being the RH_6 tag at pH 4.0. Similar to the GM2 activator, this protein might undergo large conformational exchange processes, which cannot be pushed into fast or slow exchange limits in the available range of temperature and field strength, and was not further analysed.

5.9 Methyl groups

The easy the backbone assignment by NOE correlations is in helical proteins, the difficult it is to obtain information about their arrangement in a three dimensional fold. This is because the backbones even of adjacent helices are too far separated in space to give rise to an extensive long-range NOE network of H^N and H^α protons. On the other hand, assignment of valuable NOE correlations including side chains generally requires expensive ^{13}C-labelled material for heteronuclear editing, except for a limited amount of protons with unique resonance frequencies. In saposin C, at neutral pH combing unlimited stability with excellent spectra, only 32 such long range correlations could be identified from 3D HNH-NOESY and highly resolved 900 MHz 2D NOESY experiments without *a priori* knowledge of the structure.

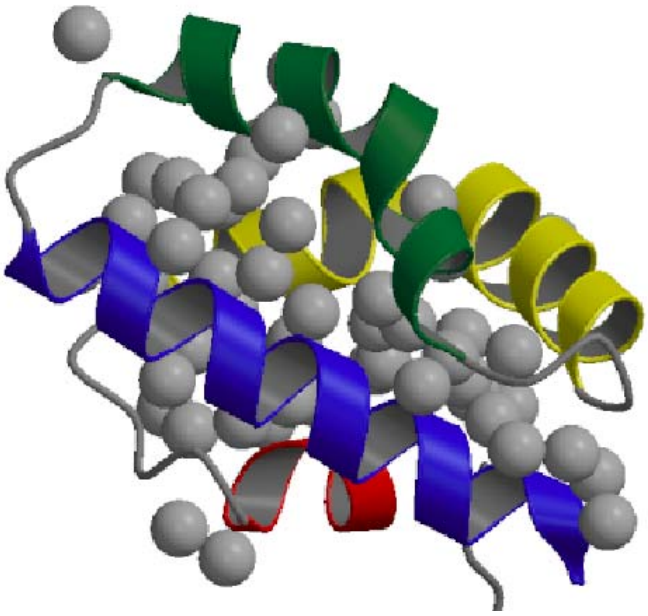


Fig. 5.9.1: Ribbon diagram of saposin C (PDB-code 1M12) with colour coding according to Fig. 5.2.1. Methyl groups are shown in CPK representation with a radius of 1.3 Å.

A new strategy for conformational studies, which looked promising for the saposins, combined comparative protein modelling, RDCs in various orienting media and the NOE-network of methyl groups in the protein core (**Fig. 5.9.1**). All saposins are very hydrophobic polypeptides, for example, saposin C comprises as much as 51 methyls from a wealth of valine, leucine and isoleucine residues. With the sensitivity of a 900 MHz cryoprobe system and the sharp resonance lines of methyls stemming from three protons, ^{13}C-HSQC and 3D ^{13}C edited (HCH-)NOESY spectroscopy at natural abundance of ^{13}C seemed realistic for a 4 mM sample of saposin C. In contrast, NOE information from hydrophilic side chains, most of which emerge into the solvent and are therefore highly flexible, is much less valuable to define the protein fold.

In saposin C, overall side chain proton assignment (see **table 8.1**) was completed up to 97 % and 100 % for methyl groups using 3D HNH-TOCSY and 2D TOCSY experiments. Diastereospecific assignment of H^{β} protons and χ_1 dihedral angles were obtained for 45 % of all residues by measuring $d_{N\beta}(i,i)$ (HNH-NOESY), $d_{\alpha\beta}(i,i)$ (NOESY), $^3J_{NH\beta}$ (HNHB) and the multiplet structure of H^{β} protons (NOESY) (**Figure 5.9.2**).

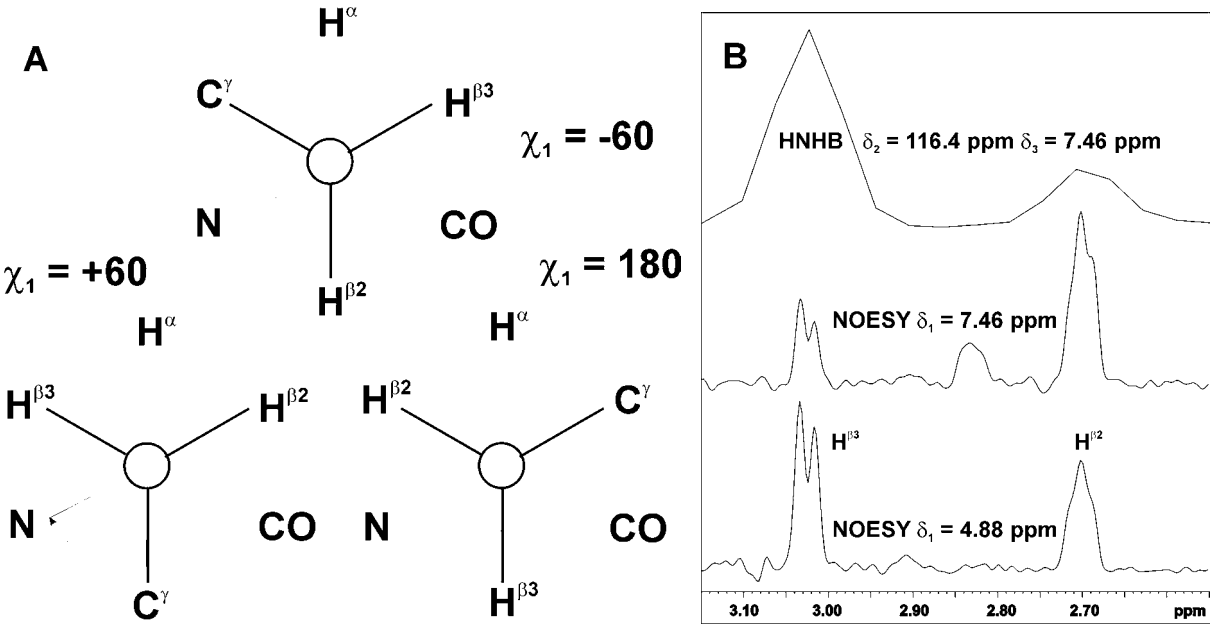


Fig. 5.9.2: A: Staggered χ_1-conformers. For $\chi_1 = -60°$, $^3J_{NH\beta2} = 0.5$ Hz, $^3J_{NH\beta3} = 5.5$ Hz, $^3J_{H\alpha H\beta2} = 13$ Hz, $^3J_{H\alpha H\beta3} = 4$ Hz, $d_{\alpha\beta2} = 3.05$ Å, $d_{\alpha\beta3} = 2.50$ Å and $d_{N\beta2} < d_{N\beta3}$. **B:** 1D ω_1-cross section of 3D HNHB spectrum, recorded at 14.1 T with an acquisition time (t_1) of 21.5 ms, at the ^{15}N and H^N frequencies of Asn21 and ω_2-cross sections of NOESY spectrum, recorded at 21.1 T with an acquisition time (t_2) of 114.7 ms and ^{15}N decoupling, at the H^N and H^{α} frequencies of Asn21. The 3D HNHB is an indicator for $^3J_{NH\beta}$, but cannot provide quantitative results. Large/small values of $^3J_{H\alpha H\beta}$ split the resonances of $H^{\beta2}$ and $H^{\beta2}$ in a pseudo-triplet and doublet, respectively. In agreement with the NOESY intensities, a value of $\chi_1 = -60°$ is obtained for Asn21.

In the natural abundance ^{13}C-HSQC spectrum (**Fig. 5.9.3 A**), all methyl signals of methionine (2), threonine (4), alanine (1) and isoleucine (4+4), as well as 13 methyls from 8 valine and 19 from 11 leucine residues could be assigned primarily on the basis of their ^1H frequencies. In the case of $\delta(^1H)$-degeneracy well visible intraresidual NOE correlations in the *HCH*-NOESY spectrum (**Fig. 5.9.3 B**) to H^{α}, H^{β} and H^{γ} protons with known resonance frequencies were used.

Unfortunately, due to low inherent sensitivity, long range NOE data turned out to be too sparse for constructing a network of methyl groups, which would require a multiple of the measurement time even on a cryo-system at 900 MHz. After the solution structure of saposin C was published, a limited number of expected correlations just rising above noise level were reproduced from the experiment. Thus, ^{13}C-edited experiments were not applied to other saposin samples.

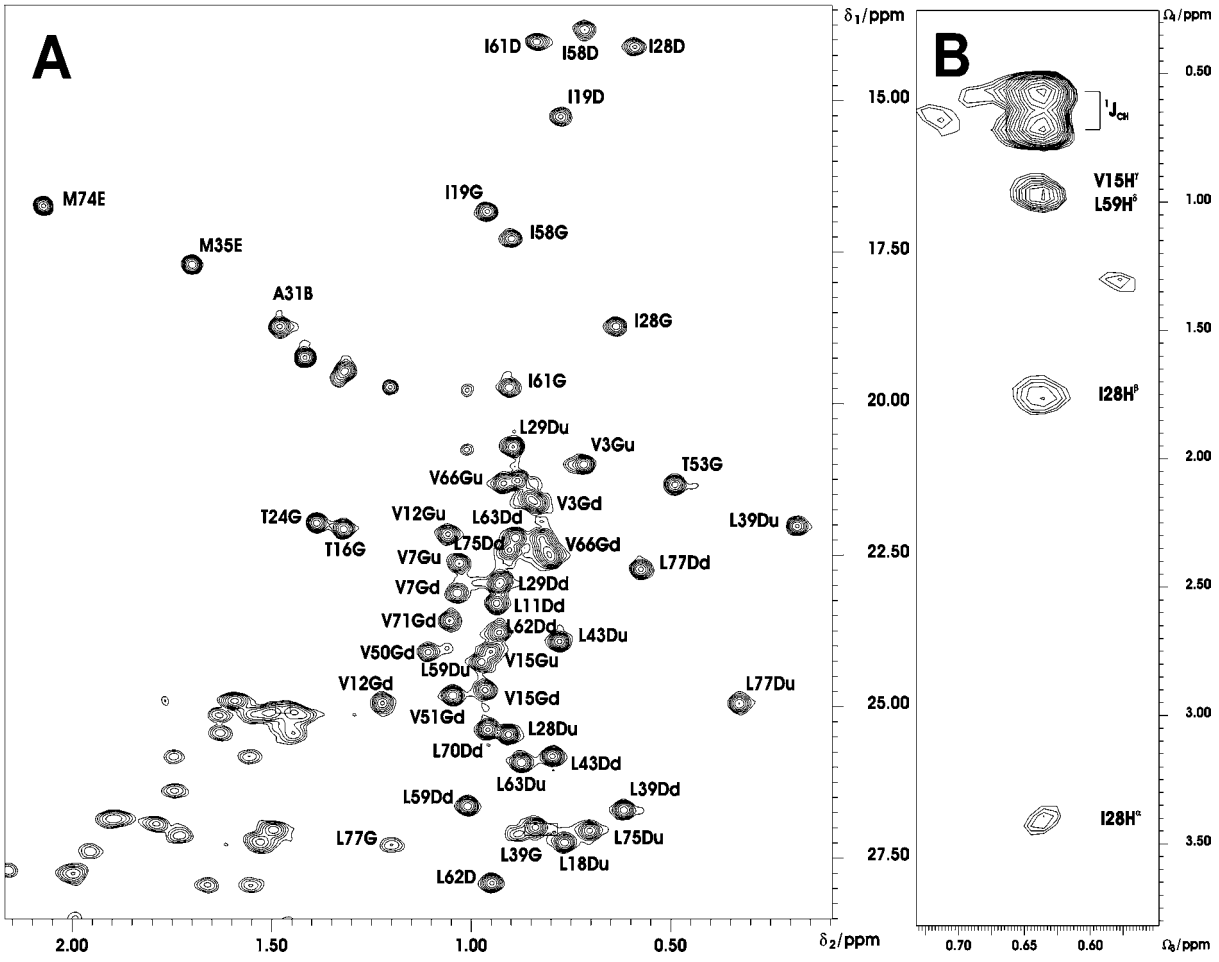


Fig. 5.9.3: A: Methyl region of the ^{13}C-HSQC spectrum of saposin C (pH 7.0, 4 mM), recorded at 37 ° on a 900 MHz cryoprobe at natural abundance ^{13}C. Acquisition times were 18.9 ms (LP to 29.6 ms) and 57.3 ms for ^{13}C and ^{1}H dimensions, respectively, and 64 transients with a total measurement time of 9 h were used. The peaks are labelled with residue number, B, G, D and E for β-, γ-, δ-, and ε-correlations and u/d for upfield/downfield in the ^{1}H dimension, since methyl groups of valines and leucines were not assigned diastereospecifically. **B:** ω_1-strip from the 3D *HCH*-NOESY spectrum at ω_2 and ω_3 of the γ-methyl group of Ile28, recorded under the same conditions. Protons were not decoupled from ^{13}C during t_1, leading to a $^{1}J_{CH}$ splitting of the diagonal signals. 16 transients were used for a total measurement time of approximately 3 days.

6 Studies of saposins by residual dipolar couplings

6.1 The alignment tensor

The *dipolar interaction* between two nuclei I and S with gyromagnetic ratio γ_I and γ_S and internuclear distance r_{IS} forms the basis of observable RDCs and is given by:[155]

$$\mathscr{H}_D(t) = -\frac{\mu_0\hbar}{8\pi}\gamma_I\gamma_S\left\langle r_{IS}^{-3}\right\rangle\left(3\cos^2\vartheta(t)-1\right)\ 2\mathbf{I}_Z\mathbf{S}_Z = -b_{IS}\left(3\cos^2\vartheta(t)-1\right)\ 2\mathbf{I}_Z\mathbf{S}_Z \qquad (6.1.1)$$

where the angle brackets denote the average of r_{IS}^{-3} due to bond stretching vibrations, and additionally $\mathscr{H}_D$ is scaled down by an order parameter due to fast bending motions (= librations). The time independent terms can be combined to b_{IS}, and time dependence is only governed by the angle ϑ between internuclear vector and external magnetic field due to molecular tumbling. Commonly, the observed *residual dipolar coupling* (*RDC*) D_{IS}, which simply results from the time (or ensemble) average of $\mathscr{H}_D$, is expressed in terms of a global *alignment tensor* **A** and the orientation of the bond vector with respect to the *principal axis system* (*PAS*) of this tensor:

$$D_{IS} = b_{IS}\left\{A_a(3\cos^2\theta-1)+\frac{3}{2}A_r\sin^2\theta\cos(2\phi)\right\} \qquad (6.1.2)$$

Here, the spherical coordinates θ and ϕ represent the angle between r_{IS} and the z-axis, and the azimuth of the projection of r_{IS} onto the xy-plane as measured from the x-axis. The *axial* A_a and *rhombic* A_r components of **A** are related to the Cartesian tensor components A_{zz}, A_{yy} and A_{xx} (in the PAS all elements A_{ij} with $i \neq j$ vanish) as follows:

$$\begin{aligned} A_a &= A_{zz} - \frac{1}{2}(A_{xx}+A_{yy}) & A_{zz} &= \frac{2}{3}A_a \\ A_r &= A_{xx}-A_{yy} & A_{yy} &= -\frac{1}{3}A_a-\frac{1}{2}A_r \\ 0 &= A_{zz}+A_{xx}+A_{yy} & A_{xx} &= -\frac{1}{3}A_a+\frac{1}{2}A_r \end{aligned} \qquad (6.1.3)$$

A is traceless (bottom left equation), which may be physically interpreted in terms of two orthogonal bond vectors that may not be preferentially oriented in the same given direction, and its properties are further elucidated in **Table 6.1.1**.

Table 6.1.1: Examples of alignment tensors of different *rhombicity* A_r/A_a. In reality the tensor components are in the order of 10^{-4} to 10^{-3}, meaning that on average one out of 10^3 to 10^4 molecules is fully aligned. The ratio A_r/A_a cannot exceed 2/3, since higher rhombicity violates the general convention $|A_{zz}| \geq |A_{yy}| \geq |A_{xx}|$. In this case (row 5), the z- and y- axes have to be swapped, and A_a and A_r change sign. Similarly, the y- and x-axes can be swapped so that A_a and A_r always possess the same sign (row 6).

A_a	A_r	A_{zz}	A_{yy}	A_{xx}	**Fig. 6.1.1**
1	0	2/3	−1/3	−1/3	A
−1	0	−2/3	1/3	1/3	-
1	1/3	2/3	−1/2	−1/6	B
1	2/3	2/3	−2/3	0	C
not: 1	1	2/3	−5/6	1/6	-
rather: −5/4	−1/2	−5/6	2/3	1/6	
not: 1	−1/3	2/3	−1/6	−1/2	-
rather: 1	1/3	2/3	−1/2	−1/6	

In order to visualise the dependence of D_{IS} on A_r/A_a, it is instructive to plot D_{IS} as a function of spherical coordinates θ and ϕ (equation 6.1.2) on the surface of a sphere (**Fig. 6.1.1**).

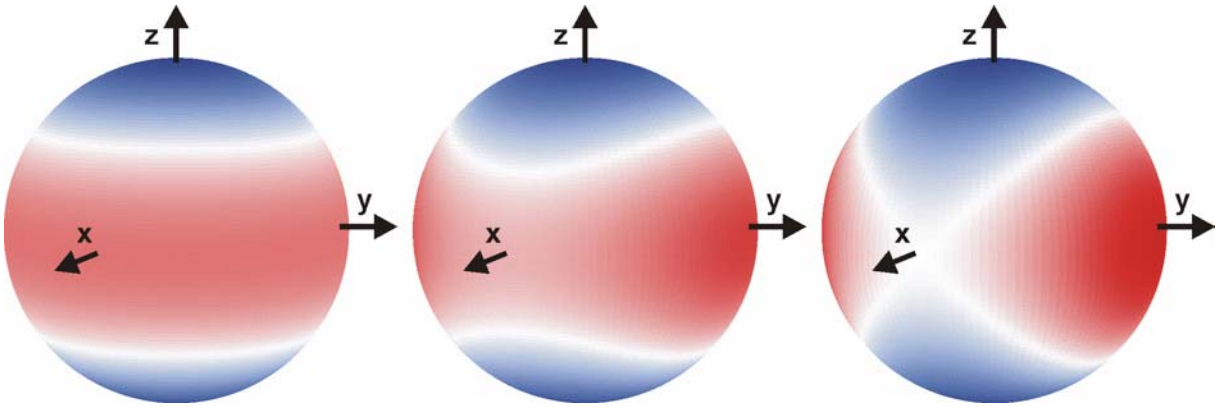


Fig. 6.1.1: Graphical representations of alignment tensors with different rhombicities. **A:** A_r/A_a = 0 (axial symmetry), **B:** A_r/A_a = 1/3 and **C:** A_r/A_a = 2/3 (maximum rhombicity). Vector orientations with the same colour intensity (red: negative, blue: positive, for positive gyromagnetic ratios and A_a) give rise to the same residual dipolar coupling. The white circles (D_{IS} = 0) in the axial symmetric tensor correspond to θ = 54.7° (magic angle). The pictures were generated with *Mathematica4* (*Wolfram Research Inc., Champaign, IL, USA*).

6.2 Residual dipolar couplings in proteins

Residual dipolar couplings (RDCs) are nowadays widely used for the refinement of NOE derived structures of biological macromolecules.[156-158] In contrast to *local* distance restraints imposed by NOEs, they are able to fix bond vectors with respect to a *global* molecular frame and therefore over large distances. RDCs have proven to be particularly powerful in the determination of the global structure of strongly anisotropic molecules like oligonucleotides, and of relative domain and subunit orientations in larger molecular assemblies, when NOE data between the domains is sparse.

However, attempts to obtain protein structures using exclusively RDC data have remained difficult, owing to the low accuracy of measured dipolar couplings other than for H-X, their degeneracy with respect to the bond vector orientation and insensitivity to translations.[159, 160] Furthermore, an incomplete cancellation of the dipolar interaction between two nuclei in solution, leading to observable RDC, requires either a strong intrinsic magnetic susceptibility [161] or the presence a weakly orienting, liquid crystalline phase,[162] which is not always compatible with the molecule under study.

Table 6.2.1: Typical one-, two- and > two-bond internuclear distances taken from the PDB-database and maximum residual dipolar couplings $2b_{IS}$ (for $\theta = 0$ and full alignment) according to equation 6.1.1 in the absence of internal motions for a number of combinations IS occurring in proteins. The listed two bond distances are nearly secondary structure independent, whereas the listed three (and four) bond distances depend on one or more dihedral angles ([a] α-helix, [b] β-strand) Only the distance $H^N(i)C^\alpha(i-1)$ is nearly constant. [c] j denotes a hydrogen bond acceptor carbonyl group. [d] distance between E- and Z-protons within the primary amide groups of asparagine and glutamine.

one bond	r_{IS}[Å]	$2b_{IS}$ [kHz]	two bond	r_{IS}[Å]	$2b_{IS}$ [kHz]	>two bond	r_{IS}[Å]	$2b_{IS}$ [kHz]
$H^N(i)N(i)$	1.03	−22.3	$H^N(i)C^\alpha(i)$	2.16	6.0	$H^N(i)H^\alpha(i)$	~3.0	8.9
$N(i)C^\alpha(i)$	1.45	−2.0	$H^N(i)C'(i-1)$	2.05	7.0	$H^N(i)H^\alpha(i-1)$[a]	~3.6	5.1
$N(i)C'(i-1)$	1.33	−2.6	$H^\alpha(i)C'(i)$	2.18	5.8	$H^N(i)H^\alpha(i-1)$[b]	~2.2	22.5
$C^\alpha(i)C'(i)$	1.50	4.5	$H^\alpha(i)N(i)$	2.09	−2.7	$H^N(i)C^\alpha(i-1)$	2.57	3.6
$H^\alpha(i)C^\alpha(i)$	1.09	46.6	$C^\alpha(i)C'(i-1)$	2.39	1.1	$H^N(i)C'(j)$ [c]	~2.8	2.8
$C^\alpha(i)$-$C^\beta(i)$	1.52	4.3	H^N_E-H^N_Z [d]	1.78	42.6			

Most commonly, the dipolar coupling contribution between two nuclei is extracted as difference of apparent doublet splitting in *oriented* versus *isotropic* solutions. Virtually all methods employ direct coherence transfer between I and S and therefore require $|D_{IS}|$ to be smaller than $|J_{IS}|$, the latter of which also should be uniform, as it is usually the case for one bond couplings. If RDCs are to be measured between nuclei separated by more than one chemical bond, r_{IS} has to be precisely known, since it depends on often non-uniform bond and dihedral angles (**Table 6.2.1**). With increasing degree of orientation, i.e. liquid crystal concentration, the RDC information becomes more accurate. On the other hand the different environment may influence the molecular structure and lead to broader lines because of slower rotational diffusion and unresolved passive dipolar couplings.

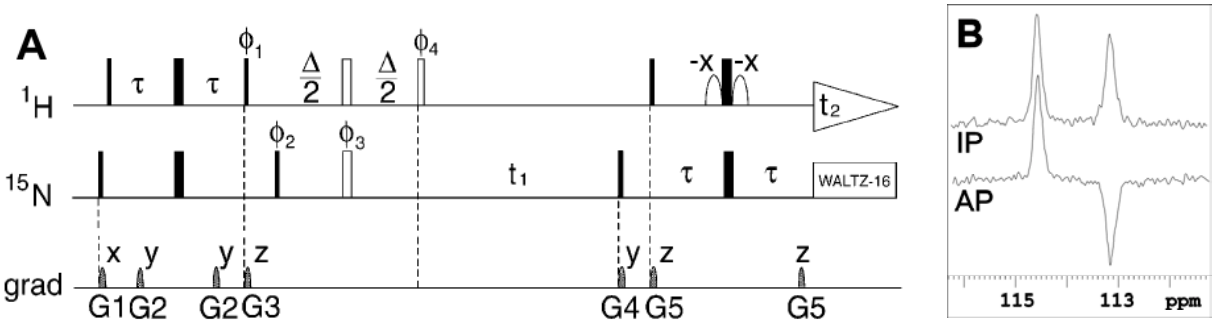


Fig. 6.2.1: A: Pulse scheme of the the IPAP-[^{1}H,^{15}N]-HSQC experiment for the measurement of $^1J_{NH}$.[163] The π-pulse denoted by an open bar is applied only in the AP-experiment and ϕ_1 = y,-y; ϕ_2 = 2(x),2(−x) for IP and 2(y),2(−y) for AP; ϕ_3 = 4(x),4(y), ϕ_4 = 8(x),8(−x) and ϕ_{rec} = x,2(−x),x,−x,2(x),−x. IP and AP sub-spectra are recorded in an interleaved manner and stored separately. Figure taken from [163]. **B**: ^{15}N-cross sections of IP and AP sub-spectra. After processing, the IP spectrum needs to be rescaled to account for the slightly faster relaxation in the AP experiment.

By far the most commonly used spin pair for the measurement of RDCs is N-H^N, because it requires only ^{15}N labelled material, and the isotropic scalar coupling ($^1J_{NH}$ = −93 Hz) is large and nearly secondary structure independent. Furthermore, the ^{15}N spins are well frequency-dispersed and among the slowest relaxing nuclei in proteins, which allows to measure ^{15}N frequency splittings with high accuracy. In small proteins with little overlap $^1J_{NH}$ can be extracted directly from a ^{15}N-HSQC spectrum without proton decoupling during t_1-evolution. A more sophisticated method is to record two separate data sets with *in-phase* (*IP*) and *anti-phase* (*AP*) doublet structure and cancel one of the doublet components by both adding and subtracting the frequency domain data (**Fig. 6.2.1**).[163] When assigning peaks, it should be considered, that peaks may not only be shifted by $^1J_{NH}/2$, but also because of chemical shift anisotropy and solvent effects.

From a set of measured couplings $D_{exp}(i)$ and a structure with given bond vectors, the alignment tensor components A_a and A_r and the respective set of back calculated couplings $D_{back}(i)$ are determined by a least squares fitting procedure to equation (6.1.2), minimising the target function χ^2_n:

$$\chi^2{}_n = \sum_i^n \frac{(D_{exp}(i) - D_{back}(i))^2}{\sigma_D(i)^2}$$

(6.2.1) which depends on the individual experimental errors $\sigma_D(i)$. Error estimates of the tensor parameters can be obtained by Monte-Carlo-type simulations, where usually several hundreds of data sets are generated by random variation of $D_{back}(i)$ within $\sigma_D(i)$. The experimental data represents a purely statistical variation of a given tensor, if the experimental χ^2_n is well located within the simulated χ^2_n-distribution, as often defined by a cutoff value including 95 % of the simulations. For homogenous errors $\sigma_D(i)$ this is equivalent to the criterion of χ^2_n being approximately equal to number of couplings, i.e. average deviations not exceeding $\sigma_D(i)$. Lower correlation between the sets $D_{exp}(i)$ and $D_{back}(i)$, having its origin in either an underestimation of $\sigma_D(i)$, a systematic error in the measurement of $D_{exp}(i)$ or inadequate structural data, is often assessed by the Pearson correlation coefficient R^2:

$$R^2 = \frac{(\sum_i (D_{exp}(i) - \overline{D_{exp}})(D_{back}(i) - \overline{D_{back}}))^2}{\sum_i (D_{exp}(i) - \overline{D_{exp}})^2 \sum_i (D_{back}(i) - \overline{D_{back}})^2} \qquad (6.2.2)$$

For an alignment tensor that is based upon an unrefined structure, the individual $D_{exp}(i)$ may be used to constrain the IS bond vector orientations with respect to this tensor in a structure refinement. The power of this method can be enhanced by using different orienting media, thus exploiting the variety of steric and electrostatic interactions between the medium and the biomolecule. Independent tensors do principally allow for an unambiguous determination of bond vector orientations in 3D space, whereas different concentrations of a single alignment medium do not add to the information.

Considering the high efficiency of ^{15}N-HSQC assignment in combination with sparseness of long range NOE data described in chapter 5, the fold of helical proteins may be efficiently analysed by RDCs, if possible, in different alignment media. The following sections give a short introduction to some of the most commonly used media and their application to solutions of saposin C, and the rest of the chapter is dedicated to the analysis of RDC data obtained for saposins C, D and A.

6.3 Phospholipid bicelles

Bicelles are disc-shaped lipid assemblies with a thickness of ~ 40 Å and a diameter of several hundred Å, which are known to be fully oriented in magnetic fields even at low concentrations.[164] Easily prepared from DMPC/DHPC in approximately 3:1 molar ratio, they have been used for weak alignment of macromolecules with great success.[162] These phospholipids are electrostatically neutral over a wide pH range (**Fig. 6.3.1**) and thus are applicable to *positively* (pH < pI) as well as *negatively* (pH > pI) charged biomolecules. Electrostatic interactions between protein and the bicelles can even be tuned to some degree by the addition of small amounts (10 % of DHPC) of charged lipids like CTAB or SDS.[165] The overall degree of alignment present in solution depends on the lipid concentration and can readily be monitored by the *quadrupolar* splitting of the solvent (90% H_2O, 10% D_2O) ^2H signal.[166]

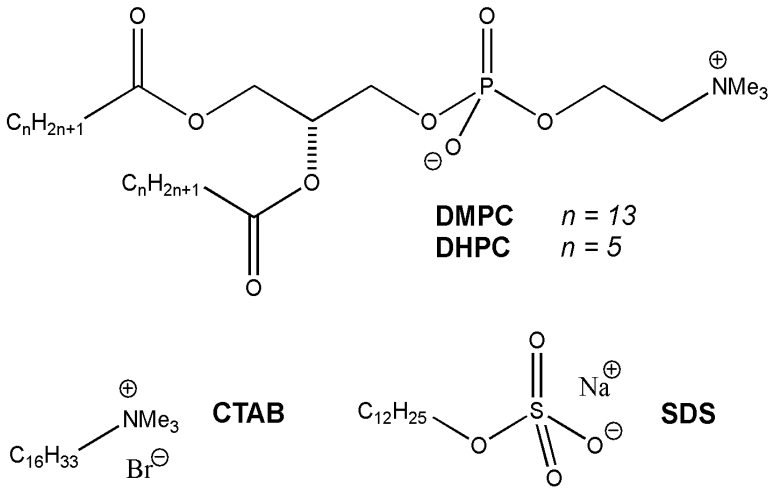


Fig. 6.3.1: Lipids for the preparation of bicelles.

DMPC: 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine,

DHPC: 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine,

CTAB: cetyltrimethyl-ammonium bromide,

SDS: sodium dodecyl sulfate.

Since under acidic conditions saposins are well known to bind to membranes, and DMPC and DHPC have been reported to suffer from slow hydrolysis, experiments were done with phosphate buffered (50 mM, pH 7.0) saposin solutions. The same buffer was used for the preparation of a 15% w/v bicelle stock in according to standard procedures.[167] Briefly, in an eppendorf cap (*Eppendorf AG, Hamburg, Germany*), 17 mg (38 µmol) DHPC (*Avanti Polar Lipids Inc., Alabaster, AL, USA*) were completely dissolved in 100 µL cold (0 °C) buffer, added to a cold suspension of 77 mg (114 µmol) DMPC in 200 µL buffer and washed three times with 100 µL buffer. After 15 min of cooling and regular vortexing, the 600 µL stock, having a milky appearance, was kept at 38° C for 30 min. Several cycles of warming and cooling were repeated, until a homogenous and highly viscous phase above room temperature

was obtained, and the stock was stored at 4° C and consumed within three days. For NMR measurements, a shigemi tube (*Shigemi Co., Ltd, Tokyo, Japan*), containing 100 µL bicelle stock + 180 µL protein solution + 20 µL D_2O was transferred quickly from the ice bath into the preheated (37° C) magnet.

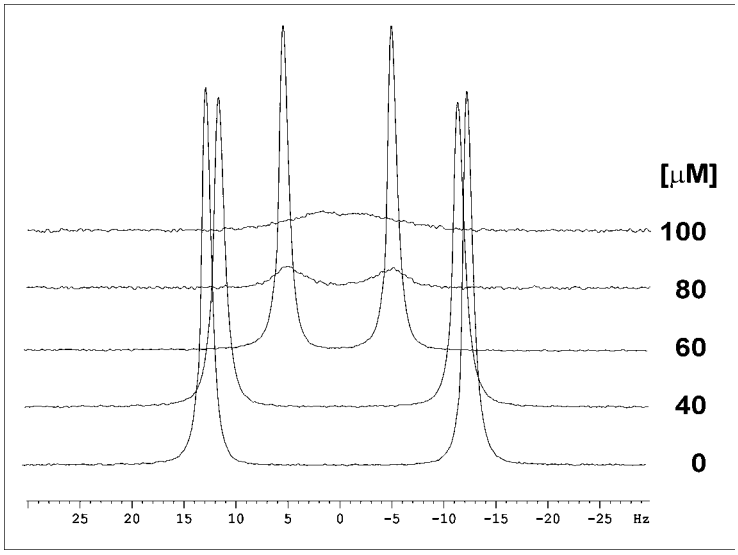


Fig. 6.3.2: Solvent ^2H signal of a 3:1:0.1 mixture of DMPC/DHPC/SDS (total lipid concentration 5% w/v), buffered with 50 mM phosphate, pH 7.0. Saposin C (2 mM in the same buffer) was added stepwise to yield the overall protein concentration as specified.

For both saposins C and D, the result was always a complete collapse of the ^2H splitting, which was analysed in a titration experiment (**Fig.6.3.2**). In the presence of SDS (10 % of DHPC) the liquid crystalline phase remained stable up to 40 µM protein (the slight decrease of the ^2H splitting merely reflects the dilution of the lipids) at 37°C for short time, but not for the time required to collect data with reasonable signal to noise ratio, even on a cryoprobe. At higher protein concentration and irrespective of the total amount of lipids, the ^2H splitting rapidly decreases, the lines broaden, and finally merge to the sharp singlet peak observed for isotropic phases. Simultaneously, the protein shows essentially the isotropic $^1J_{NH}$ splitting within the error limits. Stabilisation of the oriented phase was achieved by variation of the neither the temperature (27-47° C) nor the buffer conditions of the bicelle preparation (50-100 mM phosphate, pH 6-7).

This behaviour can only be explained either by binding of the protein to the membrane-like bicelles, thereby preventing the establishment of a stable liquid crystalline phase, or a selective binding of one of the lipids to the protein, leading to a strong perturbation of the lipid ratio. Since SDS is known to be a strongly denaturating agent, and the amounts of ^{15}N labelled proteins were limited, no further experiments with higher concentration of SDS or different lipid ratios were made.

6.4 Strained polyacrylamide gels

Polyacrylamide, which is linked with N,N´-methylenebisacrylamide, forms elastic, neutral, hydrophilic and chemically extremely inert gels, which have been used in biochemistry laboratories for decades. Since polymerisation can be initiated under mild conditions in the presence of biomacromolecules, the problem of diffusion into the gel is elegantly circumvented. When the gel body is macroscopically stretched or compressed along one axis, also the microscopic environment of the molecule enclosed is no longer isotropic and can confer weak orientation via steric interactions.[168, 169] A convenient and meanwhile commercially available device for introducing strain has been developed (**Fig. 6.4.1**).[170]

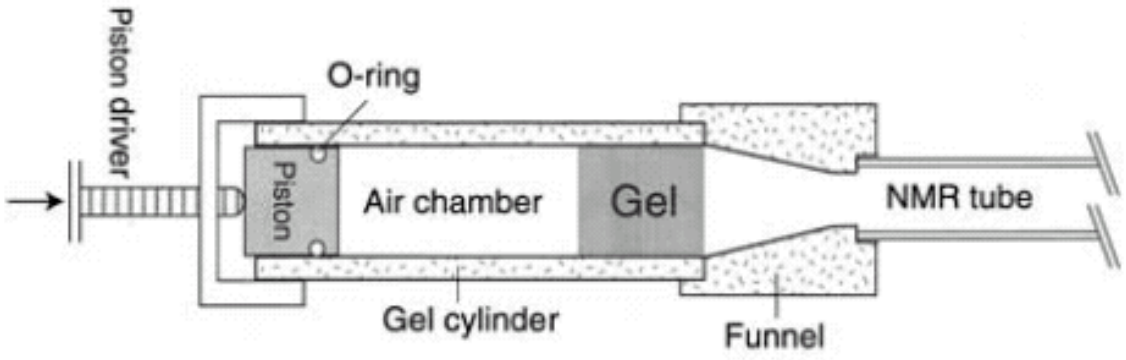


Fig. 6.4.1: Apparatus for straining polyacrylamide gels with most pieces made from teflon. After polymerisation in the cylinder the gel body is squeezed by a piston through a funnel into an NMR tube, the bottom of which has been ground off. The aspect ratio of compression is given by $(R_C/R_N)^3$, where R_C and R_N (= 4.24 mm) are the inner diameters of the cylinder and the NMR tube, respectively, and alignment has been shown to be proportional to $(R_C/R_N)^3-1$. Figure taken from [170].

As described by the authors, gels were prepared from a stock solution containing 36 % w/v acrylamide and 0.94 % w/v N,N´-methylenebisacrylamide, corresponding to a molar ratio of 1:83. For a gel with a total volume of 400 µL, 80 µL of the stock, 270 µL of the protein solution, 40 µL of D_2O and 8 µL of a 10 % aqueous solution of APS were mixed in an eppendorf cap. After the addition of 1 µL TMEDA the solution was quickly transferred to a cylinder for overnight polymerisation. First experiments with 50 mM phosphate buffer (pH 7.0) in a copy of the apparatus (R_C = 5.5 mm), made by the workshop of the institute, showed, that the gel body often broke shortly after insertion into the NMR tube, resulting in complete relaxation of strain and collapse of the ^2H splitting.

Better results were obtained for gels which had been polymerised in a modified blue (1 mL) pipette tip with R_C = 5.1 mm, corresponding to a reduction of the aspect ratio of compression from 2.18 to 1.78, and transferred to the teflon apparatus. To render the glass surface more hydrophobic and facilitate the gliding of the gel within the NMR tube, it was rinsed with propyltrimethoxysilane shortly prior to use. As major remaining problem, a part of the gel regularly remained in the teflon apparatus, and the volume of the transferred gel was too small for a satisfactory field homogeneity.

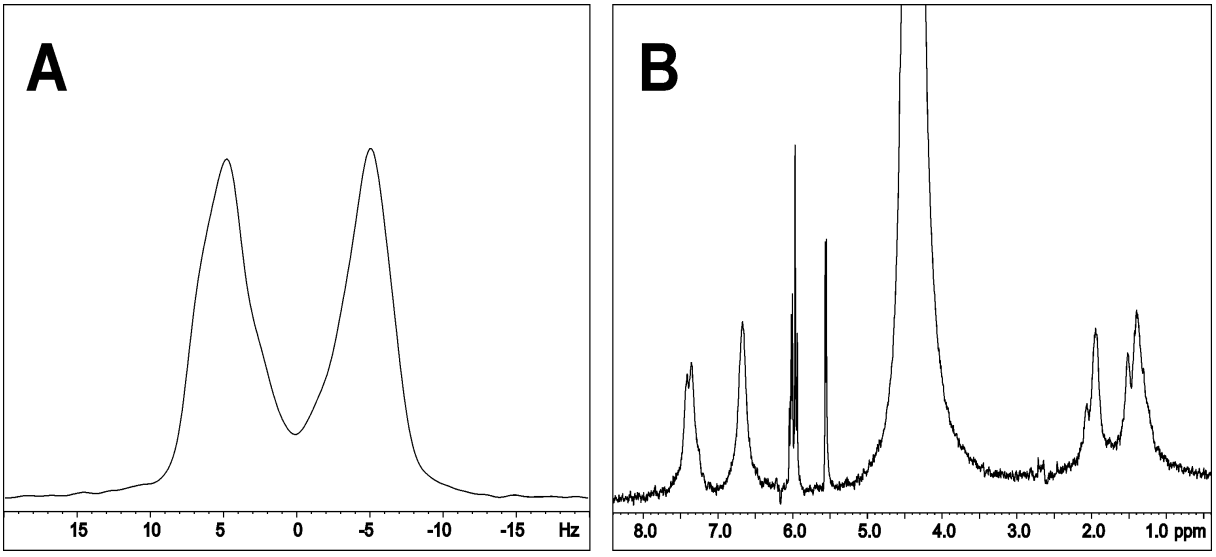


Fig. 6.4.2: A: ^{2}H spectrum of a strained polyacrylamide gel. The ^{2}H splitting, and thus the strain in the sample is not homogenous and ranges from approximately 6-14 Hz, the maxima being separated by 10 Hz. **B:** ^{1}H spectrum of the same sample. The broad signals are from the polymer, and narrow lines from remaining monomeric acrylamide indicate good magnetic field homogeneity.

After successful tests (**Fig. 6.4.2**), the phosphate buffer was replaced by a 1.2 mM solution of saposin C, and the gel body was rinsed with phosphate buffer (pH 7.0) buffer for 30 min to compensate the increase of pH during polymerisation. However, ^{15}N-HSQC spectra of the gels, showed extremely broad resonances, indicative of either strong interactions between protein and gel, or restricted molecular tumbling. Longer rinsing of the gel with buffer to rule out pH effects only resulted in loss of signal due to diffusional leakage. As a result, unless the procedure is extensively optimised, polyacrylamide gels cannot be considered a fast and reliable method of weakly orienting saposins. Electrostatic interactions could be in principle influenced by the use of charged gels from a copolymer of acrylamide-acrylate.[171]

6.5 Pf1 filamentous phage

Pf1 bacteriophage consist of a 7,349-nucleotide single-stranded circular DNA genome, which is packaged at approximately 1:1 ratio of nucleotide and coat protein into a ~ 60 Å diameter by ~ 20,000 Å long particle.[172] The highly negatively charged coat protein forms an α-helical structure, whose network of carbonyl groups is believed to be the source of the phage's large anisotropic magnetic susceptibility, with its long axis aligning parallel to the magnetic field (**Fig. 6.5.1**). Since its first employment to oligonucleotide and protein cosolutes,[173] magnetic alignment of Pf1 phage has been characterised in detail.[174]

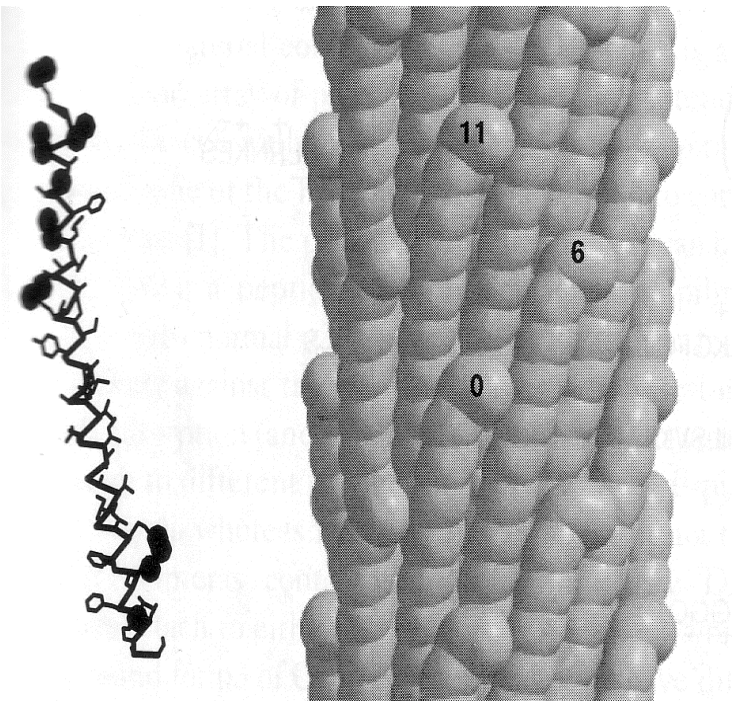


Fig. 6.5.1: Structure of phage Ff protein sheath. One α-helical subunit is shown on the left, with charges indicated by filled circles. The N-terminus (top) contains negatively charged residues facing the outer surface of the phage particle, whereas C-terminal (bottom) positively charged residues at the inner surface neutralise the charge of the DNA core. The assembly of overlapping helices, about 1% of the total length shown on the right, is held together by apolar residues. Fig. taken from [172].

Saposins were oriented by addition of a cold stock solution (52 ± 4 mg/mL) of Pf1-phage (*ASLA Biotech Ltd., Riga, Latvia*) to 250 µL of the protein solution and 50 µL D_2O in a Shigemi NMR tube (*Shigemi Co., Ltd, Tokyo, Japan*). The tube was closed with a conventional cap, sealed with parafilm and shaken rigorously, with regular cooling in order to decrease the macroscopic viscosity. Air bubbles were removed by centrifuging at 1000 rpm for 30 min, and the cap was replaced by the usual Shigemi glass piston. The liquid crystalline phase was stable below ~ 47 °C, and at 4 °C for several months. Compared to the difficulties arising in the other tested orienting media this procedure is extremely straightforward and easy and has so far been applied to phosphate (50 mM, pH 7.0) buffered solutions of saposins

A, C and D. Employment to acetate (50 mM, pH 4.0) buffered protein solutions resulted in precipitation and very broad signals in the NMR spectra and has not been further optimised.

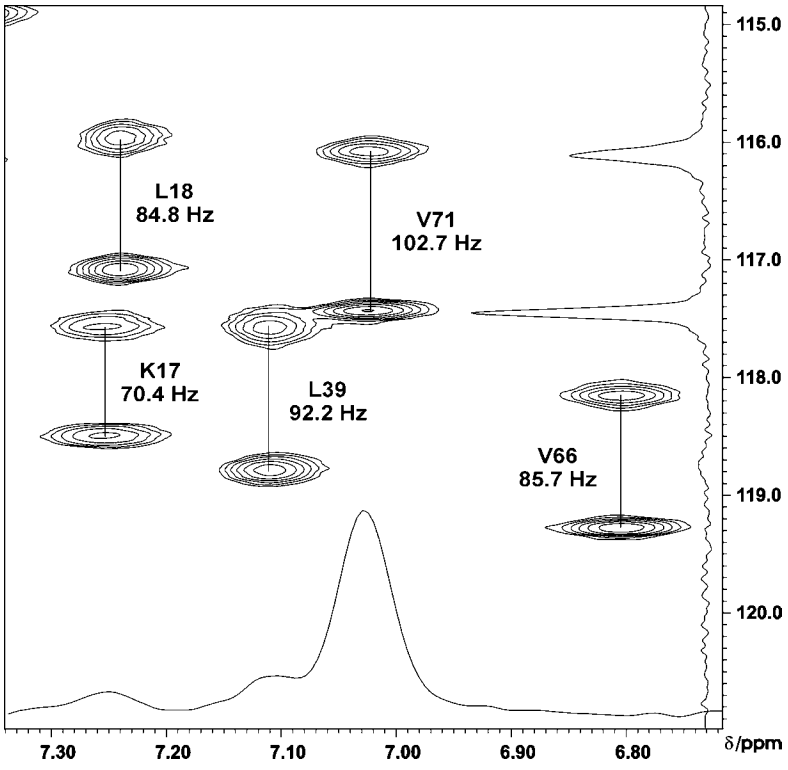


Fig. 6.5.2: Symmetric 465 Hz * 465 Hz cut of IP sub-spectrum of saposin C (13 mg/mL Pf1, 37 °C and 17.6 T). Long acquisition times (160 ms, LP to 250 ms in ^{15}N, 69 ms in ^{1}H) in both dimensions yield linewidths nearly given by the intrinsic transverse relaxation rates and unresolved couplings. The 1D traces are shown for the downfield component of V71. Due to the linewidths (FWHH (^{1}H) = 42 Hz vs. 23 Hz in isotropic solution and FWHH (^{15}N) = 7 and 10 Hz for down- and upfield components, the same as in isotropic solution), measurements of D_{NH} from the indirect dimension is 3-5 times more accurate.

All IPAP-[^{1}H,^{15}N]-HSQC spectra (**Fig. 6.5.2**) were measured at 37 °C and a magnetic field of 17.6 T field, with a maximum t_1 of approximately 160 ms ($\approx T_2$, linear predicted to ~ 250 ms) and a total acquisition time of 1.5 days. Acquisition- and processing- (digitisation of ~ 0.5 Hz in ω_1) parameters were identical for isotropic and oriented phase. The *AP* sub-spectra were added and subtracted from the *IP* sub-spectra (scaled by 0.9) in the frequency domain, to yield *upfield* and *downfield* sub-spectra, which contained less than 1 % contamination from the other component. Peaks assignments were transferred with from ^{15}N-HSQC spectra after a ω_1-shift of −45 Hz (downfield) and +45 Hz (upfield) using the program sparky (*T.D. Goddard and D.G. Kneller, University of California, San Francisco, CA, USA*). Apparent splittings $J_{NH,app}$ were extracted from frequency lists (upfield minus downfield), and RDCs were calculated according to $D_{NH} = J_{NH,app}(\text{oriented}) - J_{NH,app}(\text{isotropic})$, corresponding to D_{NH} being negative for a large doublet splitting. A uniform uncertainty in the final residual dipolar couplings of 1 Hz was assumed.

6.6 Residual dipolar couplings in saposin C

Alignment in saposin C was introduced by stepwise addition of 25 µL of a stock solution of Pf1 phage (52 ± 4 mg/mL) to 250 µL of the protein solution (1.2 mM) and 50 µL of D_2O, as described in the previous section. After four portions (= 13 mg/mL Pf1 phage), corresponding to a ^2H quadrupolar splitting of 13 Hz, the largest deviations in the N-H^N doublet splitting from the isotropic value $^1J_{NH}$ = −92 Hz reached approximately ± 30 Hz. For a large doublet splitting D_{NH} was taken as negative, but the sign is not too important, since apart from an inversion of A_a and A_r it does not change the fit results. Approximately equal maximum positive and negative D_{NH} were indicative of an alignment tensor with high rhombicity, in contrast to the familiar 2:1 amplitude of an axially symmetric tensor.

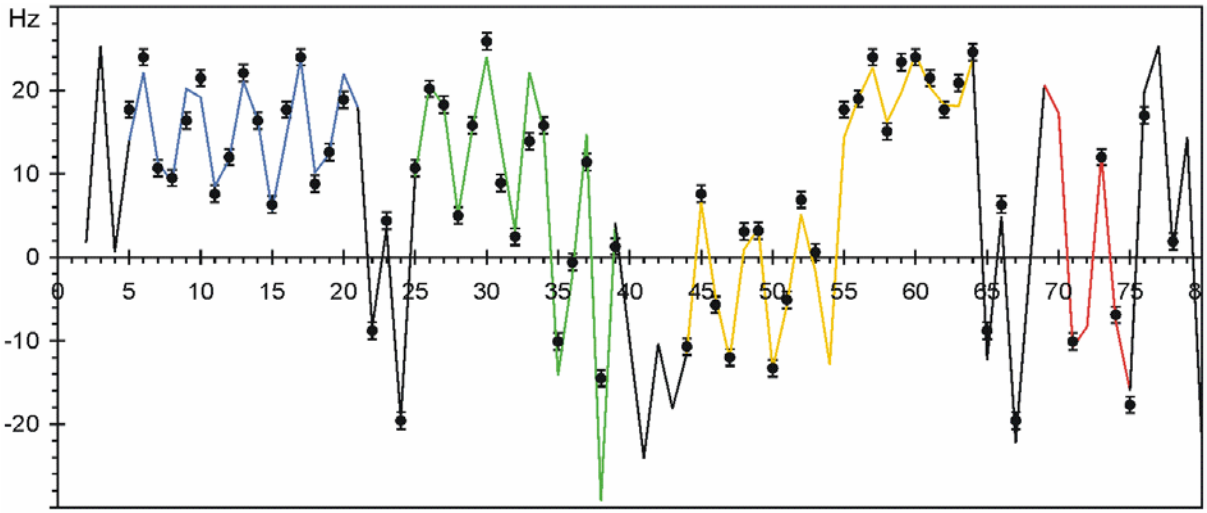


Fig. 6.6.1: Diagram of experimental RDCs versus the sequence number of saposin C in 13 mg/mL Pf1 phage (black circles). A uniform error of 1 Hz was assumed. The solid line in the usual colour coding (Fig. 5.1.1) represents values back calculated from the alignment tensor obtained by fitting 59 couplings to the published structure of saposin C (PDB-code 1M12).[144] Four residues with significant deviation, Ala31, Asp33, Met35 and Lys38, were excluded from the fit. Missing data is due to fast hydrogen exchange or signal overlap.

Plotting D_{NH} against the sequence number (**Fig. 6.6.1**) confirms the existence of five helices proposed in section 5.6, as couplings within one helix oscillate around an average representing the orientation of the helix axis. These oscillations have recently been termed *dipolar waves* [175] and will be discussed further below. α1 is a long, straight helix, whereas α2 shows a strong curvature in the second half, with a more pronounced kink between Lys34 and Met35, as becomes manifest in a strong drift of D_{NH}. α3 and α4 can be distinguished by their orientations, in contrast to the dimeric saposin B, where they are nearly merged.

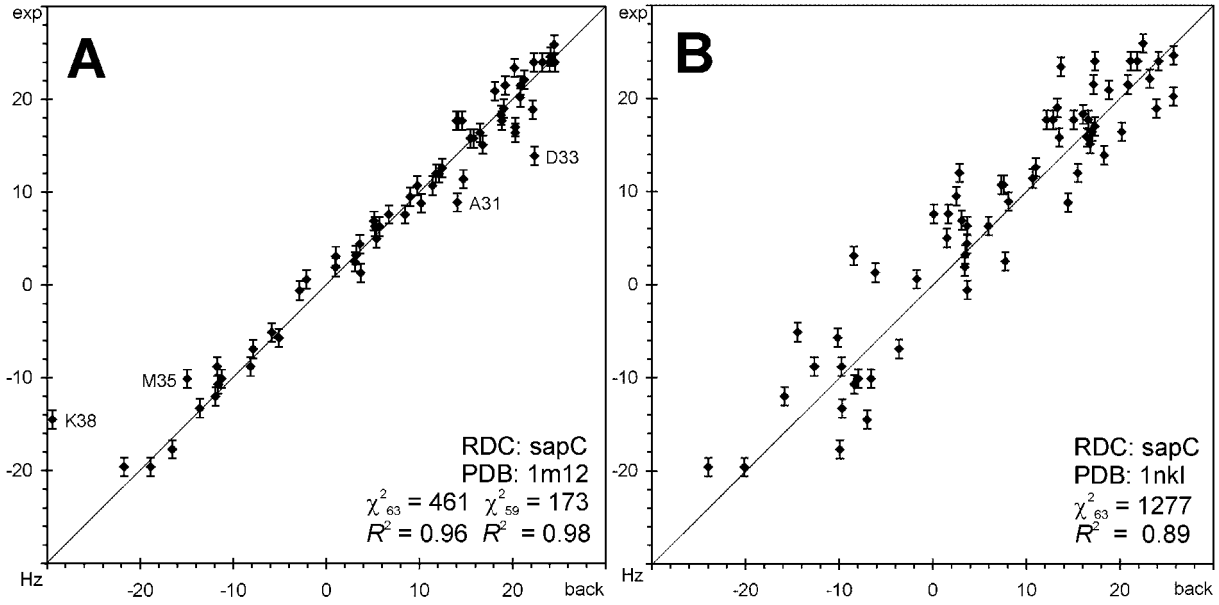


Fig. 6.6.2: Correlation of 63 experimental couplings of saposin C in 13 mg/mL Pf1 phage (vertical axis) and back-calculated (horizontal axis) couplings. **A:** Fit to the structure of saposin C (1M12),[144] excluding couplings from the four labelled residues (> 4 Hz deviation). **B:** Fit to model derived from NK-lysin (*sapC_1nkl_smh*).

When fitting the data to the recently published structure of saposin C,[144] RDCs are reasonably well reproduced with a mean per residue squared deviation of less than 3 Hz^2 (**Fig. 6.6.2 A, Table 8.3**). A few exceptions (Ala31, Asp33, Met35 and Lys38) are located in α2, where also de Alba *et al.* found the largest deviations from the final structure, which may imply exchange processes. Since the structure determination largely relied on $N\text{-}H^N$, $C^\alpha\text{-}H^\alpha$ and C^α-C' vector restraints from Pf1- and pd-phage containing solutions, it is instructive to compare the alignment tensors obtained in the different laboratories (**Table 6.6.1**).

Table 6.6.1: Alignment tensor parameters obtained by fitting different data sets of D_{NH} to the published structure of saposin C.[144] Sets A and B were used as deposited in the PDB databank, after removing residues Val3, Ala31, Asp33, Lys38, Leu39 and all residues following Cys78 from data A, and Ala31, Asp33, Leu39, and all residues following Cys78 from data B. Data C, obtained in this work, was fitted as described in Fig. 6.6.1. Experimental conditions for A and B: 50 mM phosphate (pH 7.0), 25 °C, C-terminus following Gly80: VVVV. Experimental conditions for C: 50 mM phosphate (pH 7.0), 37 °C, C-terminus following Gly80: RH_6.

data	medium	couplings	α	β	γ	$A_a[10^{-3}]$	A_r/A_a	χ^2
A	Pf1	69	6.0°	163.6°	42.3°	1.75	0.31	54
B	pd	51	3.4°	163.7°	42.4°	1.28	0.36	55
C	Pf1	59	35.0°	166.2°	60.8°	1.16	0.50	173

As apparent from the resemblance of Euler angles and rhombicity of tensors A and B, very similar mechanisms govern the alignment of saposin C by Pf1 and pd phage, which therefore may not be regarded as truly independent orienting media. The relative sizes A_a, which correspond to the observation of nearly linearly scaled couplings, are probably the result of different concentrations of phage particles or variable surface charge density (*pI* = 4.0 for Pf1 and *pI* = 4.2 for pd).[176] In contrast, tensor C (from own data) is tilted by more than 45° relative to the others and much more rhombic. This may be due to different buffer and temperature used in this study, but more likely to a large contribution of the hexa-histidine tag, which had been cleaved off by deAlba *etal*. Protein alignment by phage solutions is assumed to be dominated by *electrostatic* interactions, and positively charged surface patches have been reported to strongly interact with highly negatively charged Pf1 phage particles.[176]

Fig 6.6.3: Overlay of the structure *1m12* of saposin C (purple) and the model *SapC_1nkl_smh* (blue), based on NK-lysin, in ribbon representation. The view is approximately as in Fig. 5.9.1. The overall RMSD = 2.5 Å for the backbone atoms N, C^{α} and C′, with the largest deviations occuring in α2, which is more distorted in saposin C. In the model, Lys41 directly follows Leu39 owing to a gap in NK-lysin at Pro40, which has also been removed from the structure for comparison.

Before the structure of saposin C was published, RDC data was fit to a protein model, derived from the structure of NK-lysin (**Fig. 6.6.2 B**). Although there is an overall linear correlation between D_{exp} and D_{back}, individual couplings deviate by 5 Hz and more. With respect to structural similarity of both proteins (**Fig. 6.6.3**), these may be regarded as the result of local perturbations of $N\text{-}H^N$ bond vector orientations rather than a major variation in the fold. However, RDC data proved not sufficient to refine the model without including an adequate amount of long-range NOE information.

6.7 Dipolar waves in saposin C

Dipolar waves describe the periodic variations in the magnitudes of dipolar couplings in the backbone of partially or fully aligned proteins, if plotted versus the residue number.[175, 177, 178] Like *PISA wheels*,[179] they are spectral signatures of the inherent *periodicity* of secondary structure elements: α-helix (periodicity 3.6), π-helix (4.4), 3_{10}-helix (3.0) and β-strand (2.0-2.3). Even in ideal ($\Phi = -65°$, $\Psi = -40°$) α-helices the individual N-H^N vectors deviate by $\delta = 15.8°$ from the helix axis and thus sample discrete points D_{NH} of a sinusoidal oscillation, whose amplitude depends on δ and the overall helical orientation [θ_{av}, ϕ_{av}] and whose phase depends on the rotation ρ_0 of the helix with respect to the molecular frame (**Fig. 6.7.1**).

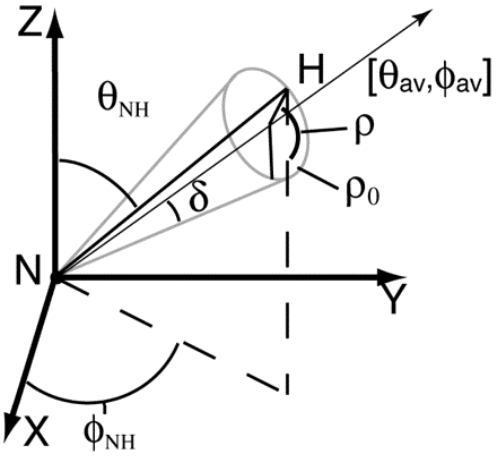


Fig. 6.7.1: A specific N-H^N bond vector (of residue i), which is located at phase position $\rho(i) = \rho_0 + 2\pi*i/3.6$ on a cone tilted at an angle δ away from the helix axis. The latter, represented by an arrow, has a given orientation [θ_{av}, ϕ_{av}] and rotation ρ_0 in the PAS, which characterise the dipolar wave. The individual $D_{NH}(i)$ is a function of the phase ρ of that particular residue (i). Figure taken from [178].

Thus, dipolar waves provide much more information compared to a fictive situation, where only an average D_{NH} within one helix (= D_{av}) is observed for collinear N-H^N bond vectors. For example, in saposin C, D_{av} is misleadingly similar for α1 and α4 (**Fig. 6.6.1**), but the amplitudes of the waves differ by approximately a factor of two. Indeed, both helices are oriented almost perpendicular to each other. Of course, the helices in saposin C are not ideal, but rather amphiphatic and distorted, due to their surface exposure in the small protein. In a canonical model, proposed by Zhou *et al.*,[180] $\Phi(i+1) = -59°$ and $\Psi(i) = -44°$ for residues on the *hydrophobic* face, and $\Phi(i+1) = -66°$ and $\Psi(i) = -41°$ for residues on the *hydrophilic* face. If plotted onto the spherical representation of the alignment tensor, distorted N-H^N vectors become visible as deviations from a circle with radius δ, expected for ideal helices (**Fig. 6.7.2**). Helix α3 is a particularly illustrative example for the origin of dipolar waves, since D_{av} is close to zero, and the $D_{NH} = 0$ line is crossed twice every helix turn, as visible by

a counter-clockwise rotation of the individual vectors (viewed against the helix direction). α4 is kinked off by approximately 50 ° and is situated in a region of strong positive D_{NH} (compare **Fig. 6.6.1**).

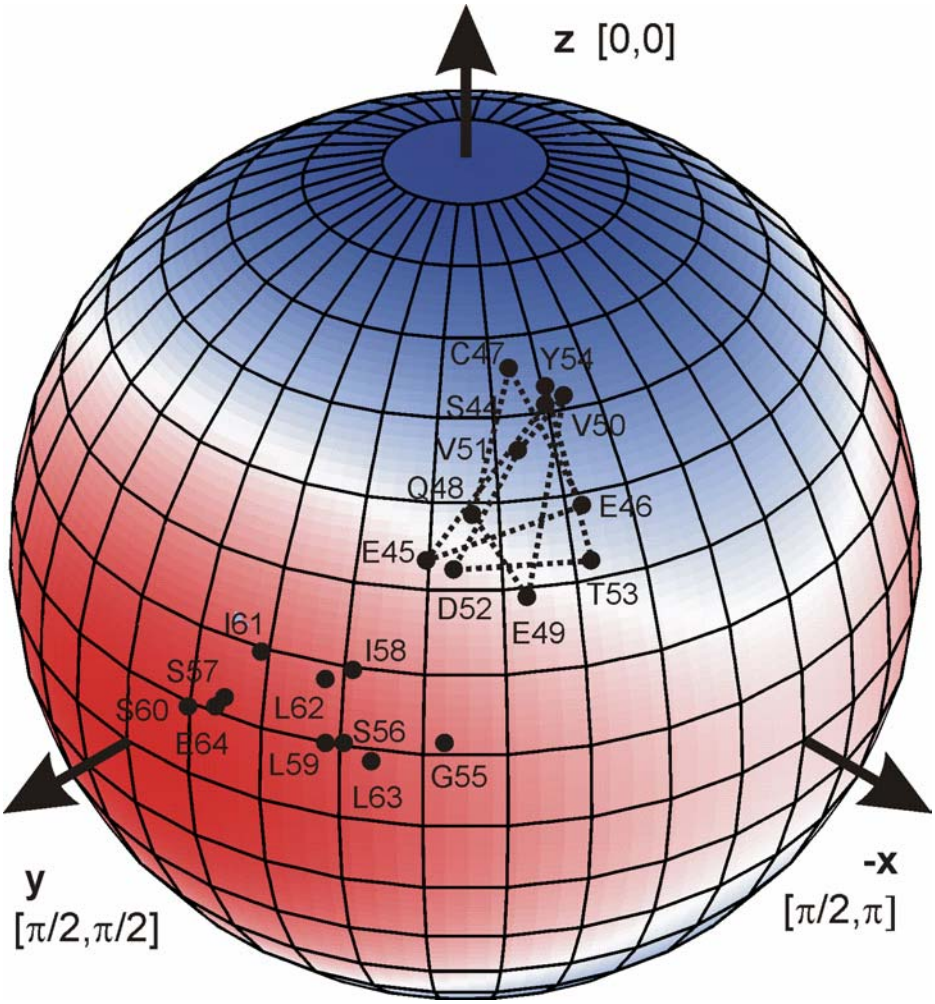


Fig. 6.7.2: Spherical representation of the alignment tensor C from Table 6.6.1. (blue = negative values) with 10° grid size. The N-H^N vectors of helices α3 (residues S44–Y55) and α4 (G55-E64) as defined by the structure of saposin C are plotted on the surface and labelled with sequence position. The angle φ_{12} between two arbitrary vectors $[\phi_1,\theta_1]$ and $[\phi_2,\theta_2]$ in the PAS is given by: $\varphi_{12} = \cos^{-1}(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2\cos(\phi_1-\phi_2))$.

Strictly, the helical dipolar wave is not a single sinusoidal oscillation with the frequency of one helix turn, but contains sometimes neglected weak contributions from a second term with *half* periodicity, as pointed out recently.[181] To understand its physical origin, we only need to consider a helix pointing along the x-axis in the PAS of a highly rhombic alignment tensor, as in **Fig. 6.1.1 C**. The average helical D_{NH} will then be close to zero, but within one helix turn the vectors will twice be passing regions of small positive (when inclined towards z and -z) and negative (y and -y) D_{NH}. However, the small value of δ keeps the amplitude of this oscillation small, which may be observed only in very special cases of helix orientation, when the regular periodic wave almost vanishes.

6.8 Residual dipolar couplings in saposin D and A

Since saposin D, at least under the conditions of study, is a system in two site slow dynamic equilibrium (chapter 7), in principle two separate sets of RDCs can be observed, at least for residues where both conformers are spectrally resolved. It has to be noted, however, that the averaging of chemical shifts and couplings by an *exchange* process principally takes place on different timescales given by the inverse of $\Delta\omega_{ab}$ and ΔJ_{ab}. Depending on the method of extracting RDCs, this may lead to perturbations of the intrinsic dipolar contributions D_a and D_b for both conformers A and B, as further discussed in section 7.7. Since at 37° exchange rates in saposin D are smaller than the experimental error of D_{NH} in rad/s, dynamic effects will be neglected here.

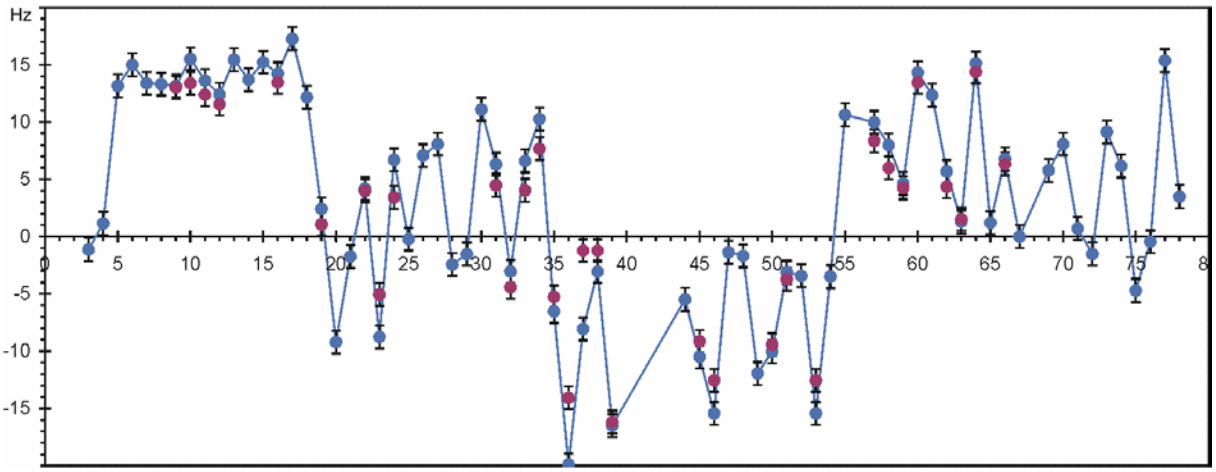


Fig. 6.8.1: Sequence diagram of D_a(i) (blue circles and connecting line) and D_b(i) (magenta circles), measured for saposin D in the presence of 13 mg/mL Pf1 phage (250 µL protein + 50 µL D_2O + 100 µL Pf1 stock). If couplings could not be resolved with respect to a and b conformers, they were assigned to D_a(i).

Where separately measurable, the two sets D_a(i) and D_b(i) of RDCs for major and minor conformer, respectively, closely resemble each other, as expected for an identical overall fold (**Fig. 6.8.1**). The largest deviations (ΔD_{ab} ~ 7 Hz) occurred in Cys36 and Ser37, which are located in the region where also large chemical shift differences were observed. In other residues of this region ΔD_{ab} is much smaller, but this does not necessarily allow conclusions about their relative N-H^N bond vector orientations. Compared to saposin C, the helices appear less pronounced in the RDC data, but also here, a jump clearly distinguishes α3 from α4. D_{NH} is strongly positive with almost no characteristic oscillation in α1, which must therefore be

oriented nearly parallel to a strongly positive axis of the alignment tensor. Referring to the sharp drop in residues Gly3-Phe4 and Leu19-Glu20, this helix appears shorter than in saposin C, which, at least at the N-terminus, may be the result of helical fraying.

Compared to saposins C and D, alignment of saposin A turned out to be much weaker and required approximately twice the liquid crystal concentration for a satisfactory distribution of D_{NH}. Despite the enormous *macroscopic* viscosity of solutions containing 26 mg/mL Pf1 phage, the average ^{15}N-R_2 increases only slightly from 7.4 s^{-1} in isotropic solution to 8.5 s^{-1}. Thus, IPAP-[^{1}H,^{15}N]-HSQC spectra could be measured separately with good resolution in the presence of 7.4, 13 and 26 mg/mL of Pf1 phage (**Fig. 6.8.2)**.

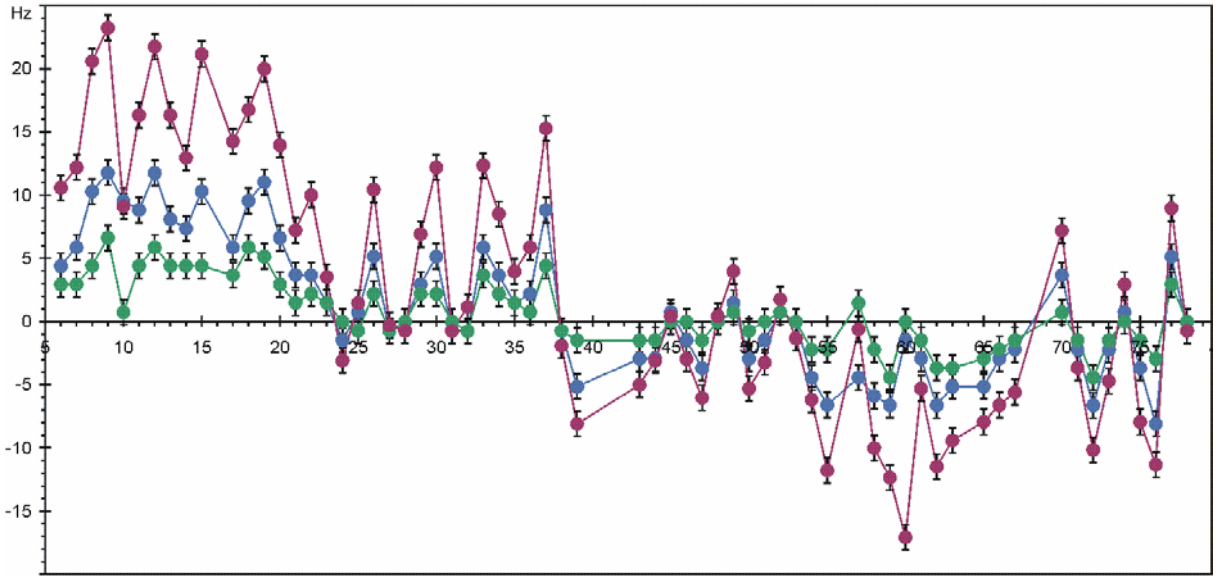


Fig. 6.8.2: Sequence diagram of D_{NH}, measured for saposin A in the presence of 7.4 (green, 250 µL protein + 50 µL D_2O + 50 µL Pf1 stock), 13 (blue, + another 50 µL Pf1 stock) and 26 (magenta, + another 200 µL of Pf1 stock) mg/mL Pf1 phage, corresponding to ^{2}H quadrupolar splittings of 8, 14 and 28 Hz. Despite slightly increased linewidths at 26 mg/mL, a uniform error of 1 Hz was assumed.

D_{NH} increases linearly for almost all residues, as expected for an alignment whose magnitude (A_a) depends on the liquid crystal concentration, whereas rhombicity A_r/A_a and orientation (= Euler angles α, β and γ) remain constant. The pattern of dipolar waves resembles the one in saposin C, but D_{NH} shows an opposite trend in α2 (like in saposin D), and D_{av} in α4 is notably different. It should again be emphasised, that this behaviour does not have to imply a different curvature of α2 or orientation of α4. Owing to non-conserved surface charge distribution among the saposins, their alignment tensors may be completely different and only pretend structural variability.

6.9 Conservation of the saposin fold

If the fold of monomeric saposins A and D is conserved with respect to saposin C under identical conditions (pH 7.0), as expected from high sequence similarity (section 5.4), experimental RDCs should yield high correlations, when fitted to saposin C-based models. This was indeed observed (**Fig. 6.9.1 A** and **Fig. 6.9.2 A**), when excluding a few outliers, almost all of which are located in either the termini (Gly3, Phe4 in saposin D), turns (Ala23, Met43 and Glu65 in saposin A and Glu55, Met66 and Gly76 in saposinD) and helix α2 (Val30, Lys34, Thr35 and Cys36 in saposin A and Gly35 and Ser37 in saposin D). Although the correlation coefficient is much lower than for the saposin C data itself, it is comparable to the fit of this data to a homologous protein, NK-lysin. It may be concluded, that the structural similarity within the saposins is in the order of the variability of Fig. 6.6.3, with stronger perturbations at the respective positions. In comparison, when fitting the data to a different fold, like the dimeric saposin B (**Fig. 6.9.1 B** and **Fig. 6.9.2 B**), the correlation between D_{exp} and D_{back} is poor, and the outliers are distributed over the whole sequence.

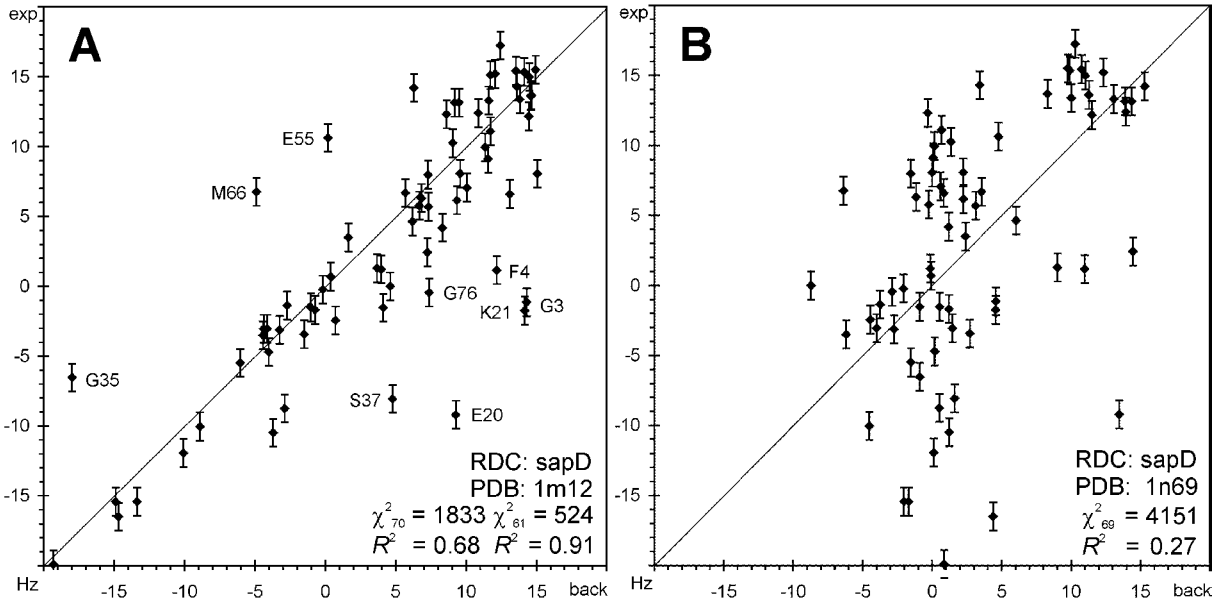


Fig. 6.9.1: Correlation of 63 experimental couplings of saposin D in 13 mg/mL Pf1 phage (vertical axis) and back-calculated (horizontal axis) couplings. **A:** Fit to model derived from saposin C (*sapD_1m12_smh*), excluding couplings from 9 labelled residues deviating by more than 7 Hz ($A_a = 0.87 \cdot 10^{-3}$, $A_r/A_a = 0.36$). **B:** Fit of all couplings to chain C of the dimeric saposin B (1N69), except for Gly64, which is a gap in saposin B.

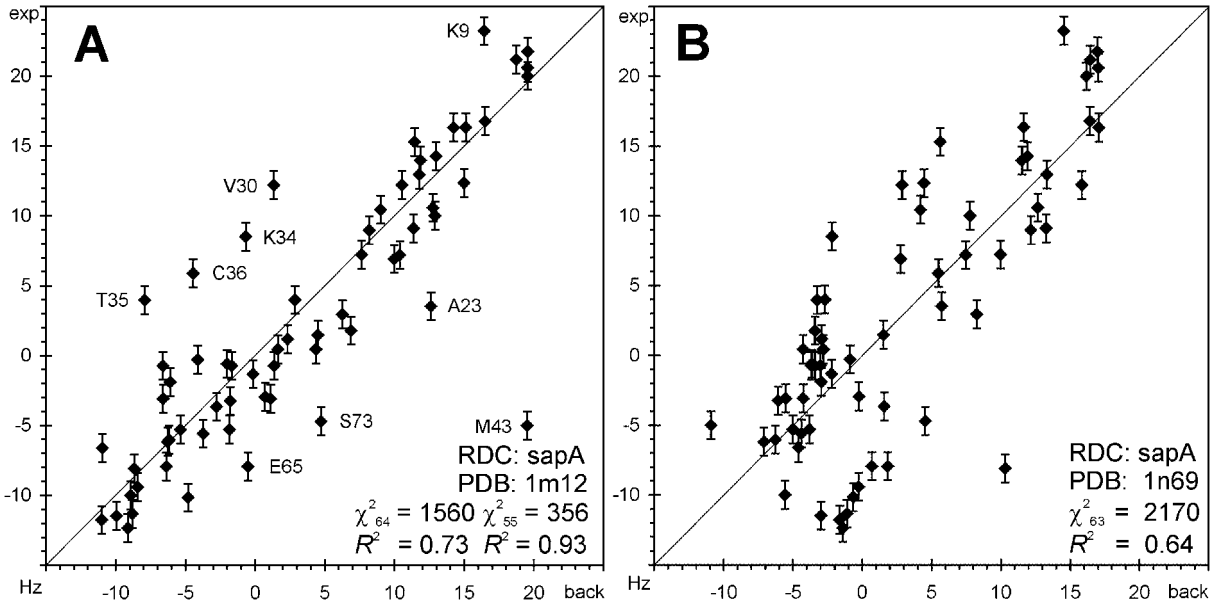


Fig. 6.9.2: Correlation of 64 experimental couplings of saposin A in 26 mg/mL Pfl phage (vertical axis) and back-calculated (vertical axis) couplings. **A:** Fit to model derived from saposin C (*sapA_1m12_smh*), excluding couplings from 9 labelled residues deviating by more than 7 Hz ($A_a = 0.90 \cdot 10^{-3}$, $A_r/A_a = 0.58$). **B:** Fit of all couplings to chain C of the dimeric saposin B (1N69), except for Glu64, which is a gap in saposinB.

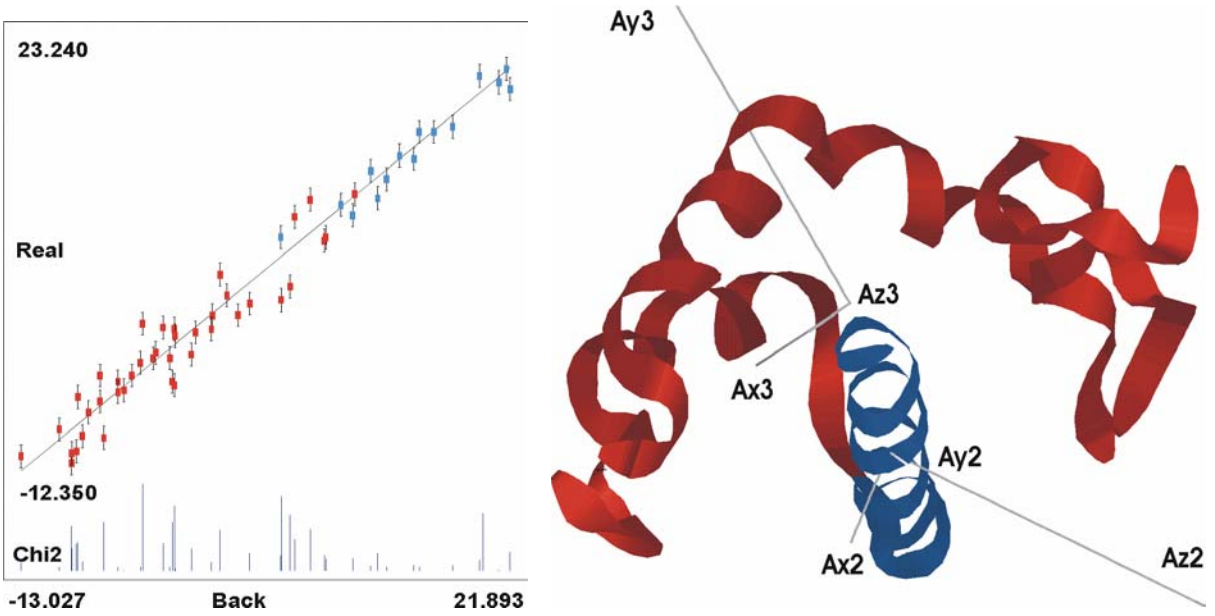


Fig. 6.9.3: Improved fit ($\chi^2 = 203$) of 55 couplings of saposinA (excluding the residues from Fig. 6.9.2 A) to two independent modules: helices α1 (blue, $\chi^2_{15} = 16$) and α2-α5 (red, 40 couplings, $\chi^2_{40} = 187$). **A:** Correlation plot provided by the program MODULE,[182] also showing the squared deviations for the individual residues (Chi2), as well as the lowest and highest values of measured and back-calculated couplings. **B:** View along the long axis (Az3) of the alignment tensor of module α2-α5, which is approximately parallel to the y-axis (Ay2) of the tensor of module α1, corresponding to a swap of axes in highly rhombic tensors (Ax2 and Ax3 are small) as discussed in Table 6.1.1. Az2 is tilted away from Ay3 by 31°, which corresponds to a phase shift in the rotation of α1.

In saposin A, a decrease of χ^2 can be achieved, if $\alpha 1$ is rotated by approximately 30° around its axis in the half-shell formed by the other helices (**Fig. 6.9.3)**, which corresponds to a phase shift of the dipolar coupling wave as discussed in section 6.7. In principle, another decrease is possible if the molecule is further bisected into smaller units. However, since the helices in the saposins are short, deviations in measured RDCs are more likely to originate from sequence-specific helical distortions than from inadequate helix orientations and rotations. Note also, that since the alignment tensor contains five independent parameters, any five vectors can always be fitted to some tensor with essentially zero χ^2.

7 Exchange in saposin D

7.1 ^{15}N-HSQC spectrum

From the number of resonances in the ^{15}N-HSQC spectrum (**Fig. 7.1.2**) it is immediately evident that two (monomeric) species, populated in a ratio of approximately 2:1, are present in solutions of saposin D at pH 7 and 37 °C. Nevertheless, for the majority of residues, either just a single resonance is observed, or two resonances may be distinguished, but these are merged beyond coalescence. This is strongly indicative of minor structural differences, which are locally restricted and do not affect the common overall fold and the particular environment of these amide groups. Similar spectra, showing a coexistence of major and minor conformers, were obtained in samples of the protein at pH 4, which were, however, highly unstable.

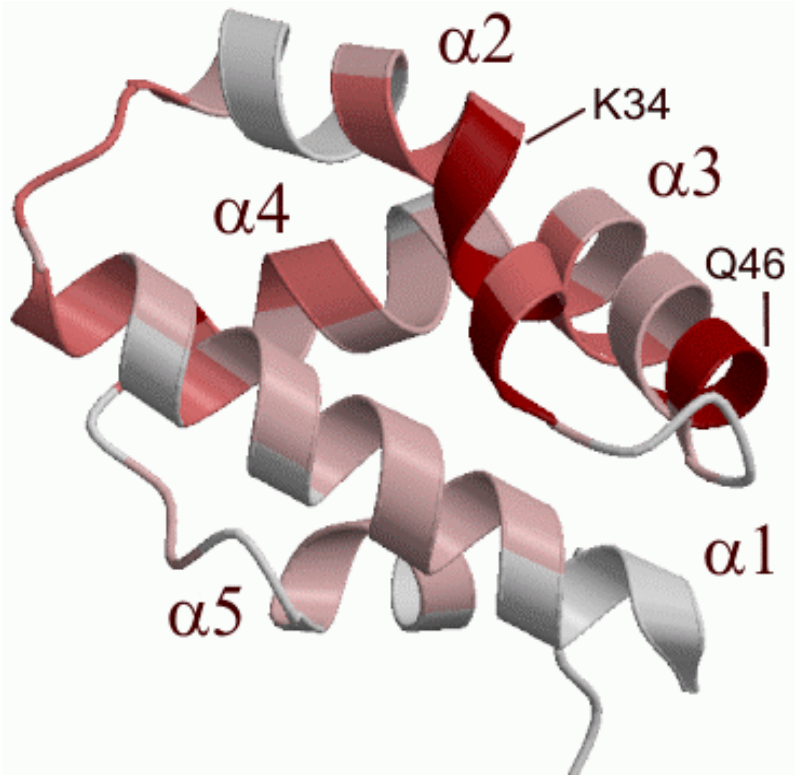


Fig. 7.1.1: ^{15}N-HSQC resonance separation between conformers A and B in saposin D, mapped upon the model *SapD_1m12_smh* (see section 5.4). Colour coding: single resonance or unassigned (grey), beyond coalescence (= single maximum, pink), peaks not separated down to the baseline (red) and well separated (magenta). These definitions depend on the linewidths, chosen acquisition times and magnetic field, which were as stated in Fig. 7.1.2. The figure was generated with *MOLSCRIPT* [183] and *Raster3D*.[184]

In a second group of residues the two amide resonances representing conformers A and B are well frequency separated in one or both dimensions. Mapping onto a structural model of saposin D allocates this group to the inner helical faces of the end of α1 (and the following loop), α2, α3 and α4, with remarkable shift differences occurring in residues Lys34-Cys36, and again in Gln44-Gln46 (**Fig. 7.1.1**).

These two stretches are located in close proximity at the end of α2 and beginning of α3 and are linked by a disulfide bond Cys36-Cys47. Intriguingly, residues following the first stretch (Ser37-Leu39) and Tyr43, which preceeds the second stretch, are much less affected in the ^{15}N-HSQC spectrum. The whole loop region in between, comprising residues Pro40-Pro42, is not traceable, since the amide proton of Lys41 exchanges fast with the solvent.

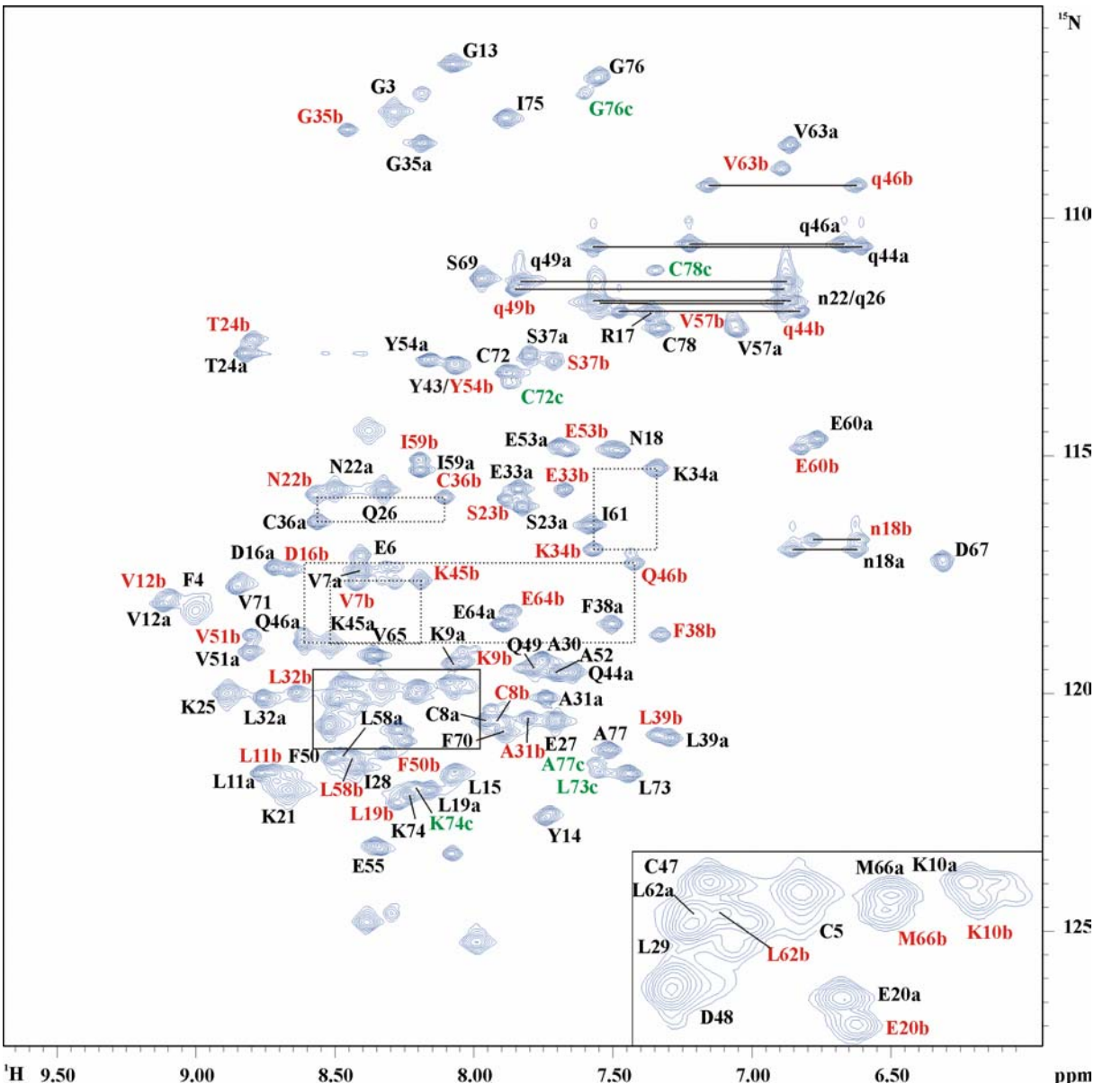


Fig. 7.1.2: ^{15}N-HSQC of saposin D (2 mM, 50 mM phosphate, pH 7.0) at 37 °C and 14.1 T with acquisition times of 99.8 (LP to 149.8) and 86.1 ms in the ^{15}N and H^N dimensions, respectively. Residues are labelled in black with "a" for the major, and red with "b" for the minor conformer. In case of coalescent resonances black labels are used without specification of the conformer. Rectangles illustrate the frequency separations of Lys34, Cys36, Lys45 and Gln46. Peaks from a third species ("c") are labelled green. The insert shows an expansion of the crowded region in the box.

An obvious reason for the coexistence of two conformers in saposin D may be given by a *cis-trans* isomerisation of one of the above mentioned prolines, or both. Due to the sparse assignment and NOE data in this region, this possibility has been neither approved nor disproved so far. Considering the pattern of strongly affected residues, however, an isomerisation of the disulfide bond along with different side chain conformations of the involved cysteines, as it has been observed in BPTI (**Fig. 7.1.3**),[185] seems much more likely. In aqueous solution, no multiple conformers of free cystines can be separately observed by NMR, and from small organic model compounds at −100 °C, the inversion of disulfide bond chirality has been estimated to require only approximately 7 kcal/mol of activation energy.[186] However, under steric constraints imposed by the framework of cyclic peptides[187] or proteins, this barrier may be more than doubled, thereby making the individual conformations accessible to NMR studies at room temperature.

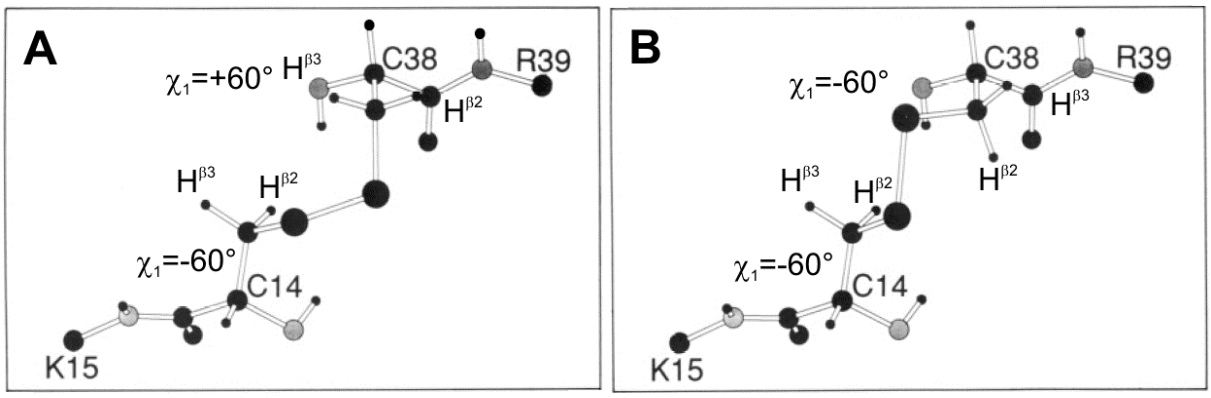


Fig. 7.1.3: Ball-and-stick diagram showing the backbone and side chain conformations of residues Cys14 and Cys38 in BPTI and the chirality of the disulfide bond, if viewed in the direction of the sulphur atoms (large black balls). Medium and small sized balls represent carbon and hydrogen, and grey balls represent nitrogen atoms. H^{β} protons are labelled, together with the respective dihedral angle χ_1. **A:** major conformer with χ_1(C38) = +60° and right handed disulfide chirality. **B:** minor conformer with χ_1(C38) = −60° and left handed disulfide chirality.

Using the methods described in section 5.9, identical values χ_1 = −60° were found for the Cys36-conformers in saposin D, and Cys47 gives rise to only one ^{15}N-HSQC resonance, thus disulfide isomerisation remains a hypothesis. Even in residues where nuclei are strongly affected by the above mentioned chemical shift separations, virtually identical NOE data did not allow to elaborate clear structural differences between the two conformers. Minor perturbations in the RDC data are discussed in sections 6.8 and 7.7. Further studies will have to clarify whether the existence of two states, which seems to be an intrinsic property of saposin D, reflects a biological relevance.

7.2 Assignment

Because for most residues two resonances exist in ^{15}N-HSQC spectra of saposin D, the overlap is strongly increased. For example, the ^{15}N and H^N frequencies of Tyr43a and Tyr54b are exactly degenerate, thereby making the identification of two tyrosine residues from side chain resonances extremely difficult. Using the assignment strategy described in section 5.5, the peculiar situation arose, that for both HNH- and NNH-NOESY experiments often three rather than two strong NOE correlations were observed. This occurs, because longitudinal magnetisation does exchange between the two species during the NOE mixing period, producing peaks with the same sign as those from cross relaxation in the slow tumbling regime. In order to distinguish both effects, mixing sequences have been designed, in which (negative) NOE and (positive) ROE mutually cancel out, leaving only signals from true chemical exchange.[188] However, exchange peaks in the HNH-NOESY spectrum can also be identified from a corresponding peak in the HNH-TOSCY spectrum, since exchange also occurs during spin lock mixing (**Fig. 7.2.1**).

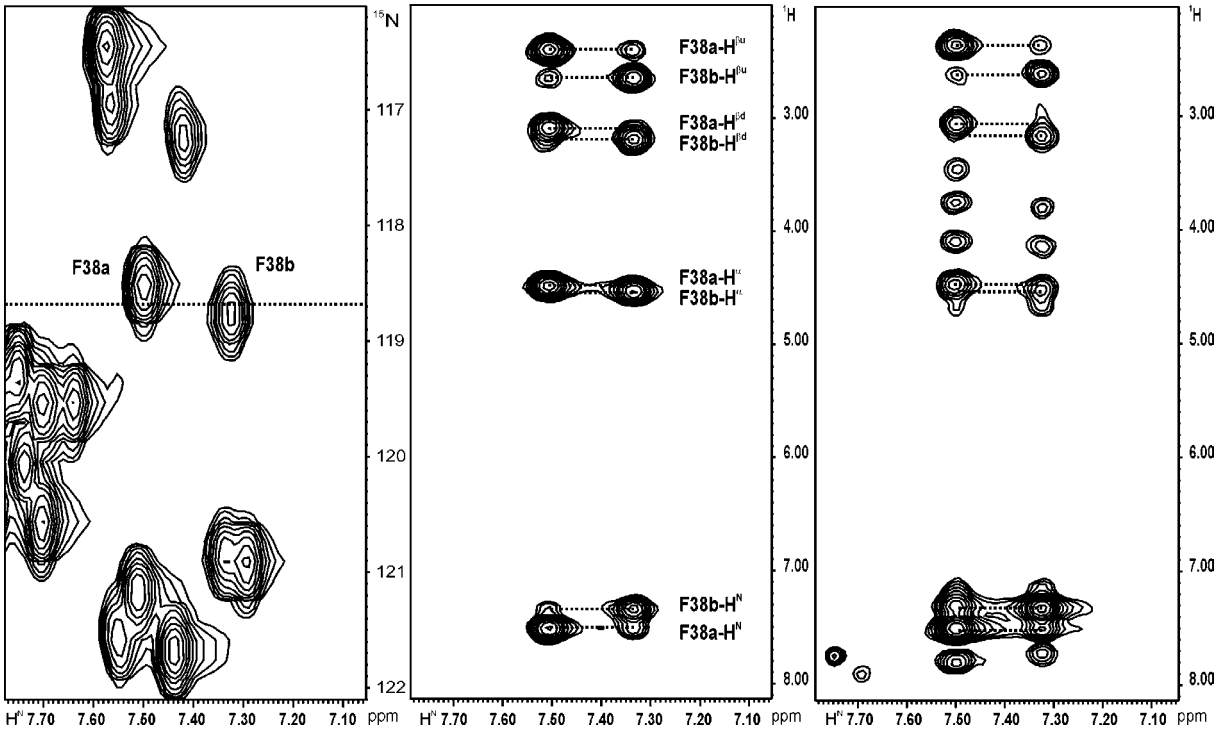


Fig. 7.2.1: A: Section of $\omega_2(^{15}N)$-$\omega_3(H^N)$-projection (^{15}N-HSQC dimensions) of a HNH-TOCSY spectrum of saposin D (90 ms clean-CITY mixing [154]), showing the exchanging resonances F38a and F38b. **B:** $\omega_1(^1H)$-$\omega_3(H^N)$-plane at $\omega_2(^{15}N)$ = 118.7 ppm of this HNH-TOCSY spectrum. **C:** the same plane in a HNH-NOESY spectrum with 120 ms mixing. Dotted lines connect direct and exchange relayed peaks, and the position of the plane is shown as dotted line in A.

Obviously, in the TOCSY experiment, not only magnetisation exchange of detected H^N protons and transfer through the $^3J_{HH}$-coupled spin system occurs, but also combinations of both, giving rise to completely new peaks in the spectra. For exchange rates k_{ex} in the order of $^3J_{HH}$, principally each member of the spin system of species A is correlated with each member of the spin system of species B. A similar argumentation is valid for the NOESY experiment, if k_{ex} competes with cross relaxation rates σ_{HH} within the dipolar coupled network of both individual species. As pointed out by Otting *et al.*,[185] this enormously blows up the number of resonances and complicates the analysis of spectra (**Fig. 7.2.2**), although the *exchange relayed* correlations are often much weaker.

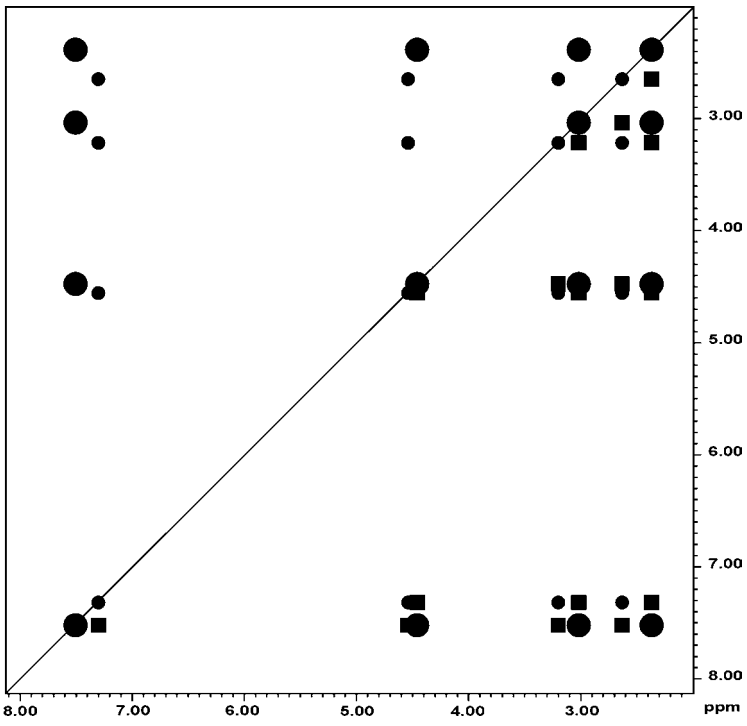


Fig. 7.2.2: Schematic 2D NOESY or TOCSY spectrum of the slowly exchanging $H^NH^{\alpha}H^{\beta 2}H^{\beta 3}$ spin system of F38a/F38b, exemplifying Cys, Asp, Phe, His, Asn, Ser, Trp and Tyr residues. The upper left triangle shows the spectrum in the absence of exchange (e.g. at low temperatures), with large and small circles representing (each 4 diagonal- + 2*6 cross- = 16 overall) signals from the major and minor conformer, respectively. In the lower right triangle (k_{ex} in the order of $^3J_{HH}$ or σ_{HH}) 4 direct + 12 relayed exchange correlations (= overall 32 signals, squares) add to the spectrum.

In practice, the number of resonances is greatly reduced owing to completely or nearly degenerate chemical shifts in the exchanging sites. Since carbon bound protons are generally less sensitive to conformational changes than amide protons, in most residues of saposin D H^{α} and H^{β} frequencies of conformers A and B are very similar, even if the amides are separated. On the other hand, all residues with (nearly) coalescent peaks in the ^{15}N-HSQC spectrum display virtually a single signal set in NOESY and TOCSY spectra. Chemical shift degeneracy was solely an issue for some cases, where the neighbours of one of these residues were not degenerate and gave rise to a double set of strong d_{NN}(i,i+1) correlations.

7.3 The N_z-exchange experiment

Once slow chemical exchange is identified and distinguished from cross-relaxation, it may be quantified by 2D exchange spectroscopy (EXSY),[13] a homonuclear experiment completely identical to 2D NOESY, with a series of mixing times. Its application to biological macromolecules is, however, severely limited by the need to resolve pairs of exchanging protons in both diagonal and cross positions. Especially when cross peaks are close to large diagonal peaks, their detection, not to mention quantification, becomes virtually impossible. Assuming that the slow exchange equilibrium in saposin D affects all residues with the same overall pair of rate constants k_a and k_b, in principle a single proton pair would allow for their extraction, still, such a pair does not exist in the molecule.

To overcome this problem, an approach has been proposed, that exploits the increased chemical shift resolution of heteronuclear correlation experiments by monitoring the exchange of two spin order $\langle 2\mathbf{I}_z\mathbf{S}_z \rangle$.[189] This experiment was modified by mixing on $\langle \mathbf{S}_z \rangle$.[189] (*$N_z$-exchange*), thereby extending the lower limit for exchange rates down to pure ^{15}N longitudinal relaxation (**Fig. 7.3.1**), and applied to the N-terminal SH3 domain of protein drk.[190] It was known that this 59-residue fragment exists in a 2:1 slow dynamic folded-unfolded equilibrium at 20°C.

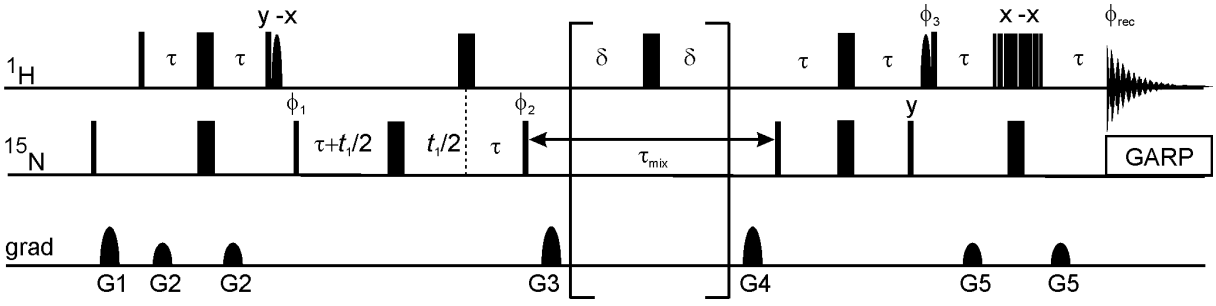


Fig. 7.3.1: Pulse scheme for the measurement of N_z-exchange as used in this work, which is slightly modified from the literature.[190] Narrow and wide bars represent 90° and 180° pulses, respectively, which are applied with phase x, unless indicated otherwise. τ was set to 2.5 ms, and in the mixing period τ_{mix}, consisting of repetitions of the brackets (6 ms) plus gradients G3 and G4 (each 1 ms), 180° pulses on protons were applied every 2δ = 3 ms in order to suppress dipole-CSA cross-correlated relaxation. Water selective 90° pulses were applied with a Gaussian shape and a length of 2 ms, and final water suppression was achieved with a 3-9-19 binomial sequence with a delay of 200 µs between the pulses. Pulse phases: ϕ_1 = x,−x, ϕ_2 = 2(y),2(−y), ϕ_3 = 4(x),4(−x) and ϕ_{rec} = x,2(−x),x,−x,2(x),−x. Relative gradient strength (where 100 corresponds to 33 Gauss/cm): G_1 = 30, G_2 = 15, G_3 = 60, G_4 = 40, G_5 = 20.

The sequence starts with an INEPT polarisation transfer from ^{1}H to ^{15}N spins, and antiphase magnetisation with respect to protons is refocussed in combination with t_1-evolution (which may be implemented as semi-constant time). Subsequently, magnetisation is alternatingly flipped to $+N_z$ and $-N_z$, ensuring its decay to zero rather than the thermal equilibrium polarisation, if the receiver phase follows this pulse. Thus, in analogy to section 1.4, a ^{15}N frequency labelled and population weighted pair of initial polarisations ($\langle \mathbf{N}_z \rangle_a(0)$, $\langle \mathbf{N}_z \rangle_b(0)$) is prepared at time point $t = 0$, which is allowed to evolve under the propagator **Q** to the final pair ($\langle \mathbf{N}_z \rangle_a(\tau_{mix})$, $\langle \mathbf{N}_z \rangle_b(\tau_{mix})$) at time point $t = \tau_{mix}$. This is finally converted into observable transverse proton magnetisation by a refocussed INEPT step and detected as H^N-frequency separated resonances. The corresponding spectra have a ^{15}N-HSQC-like appearance, but apart from just two *auto peaks* for each residue in both conformational states, an additional pair of *exchange peaks* arises from magnetisation transfer between both states, forming a rectangle with the auto peaks (**Fig. 7.3.2**).

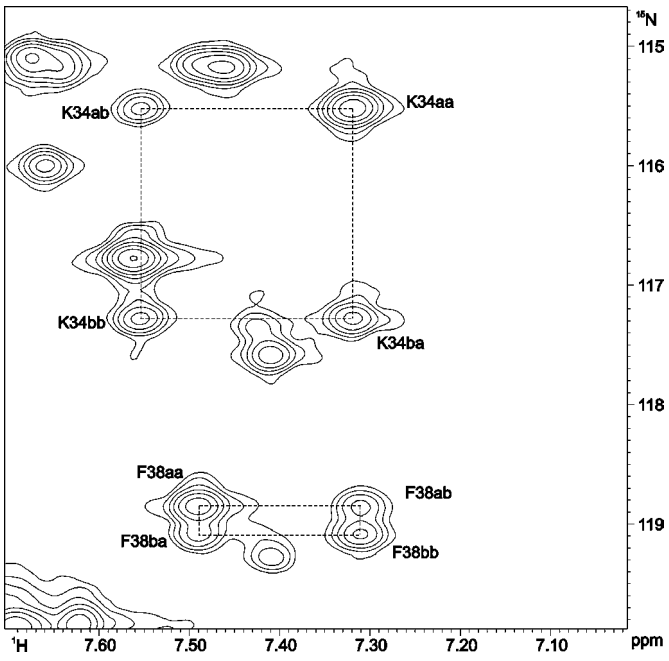


Fig. 7.3.2: N_z-exchange spectrum with single mixing period (τ_{mix} = 363 ms) of saposin D (pH 7, 37° and 14.1 T cryoprobe), which aided in the ^{15}N-HNQC assignment of the protein. Acquisition times were 74.9 (extended to 124.8) and 85.2 ms in the ^{15}N and H^N dimensions, respectively. During τ_{mix} auto signals K34aa and K34bb (corresponding to K34a and K34b in the ^{15}N-HSQC) have decayed to approximately half of the initial intensity. At τ_{mix} = 363 ms, the exchange signals K34ab (exchange from A to B) and K34ba (exchange from B to A) have maximum intensity.

Fitting all four peak volumes as a function of τ_{mix} against the analytical expressions for the matrix elements of **Q** as given in equation 1.4.4 allows to extract ^{15}N longitudinal relaxation (R_1) and exchange (k_{ex}) rates. Whereas in a two site equilibrium one single value for k_{ex} should be obtained, the ^{15}N-R_1 rates vary among the residues, depending on the N-H^N bond vector orientation with respect to the global diffusion tensor and superimposed internal motions. The individual values for ^{15}N-R_1 may be used in a Lipari-Szabo type analysis to study internal protein dynamics.

7.4 Application of N_z-exchange to saposin D

The extreme thermal stability of saposin D at pH 7 allows the exchange process between states A and B to be studied over a wide range of temperatures. An interval of 10 °C was chosen, starting from 17 °C, where increasingly broad lines caused severe overlap, to 57 °C, where solvent evaporation during the experiments became a major issue. For all temperatures, N_z-exchange spectra were recorded (for the pulse program see section 9.7), using the same 10 mixing periods up to τ_{mix} = 722 ms (≈ T_1 at 17 °C or ≈ $2T_1$ at 57 °C). A long recycle delay of 4 s ensured, that all proton spins in states A and B are completely relaxed to their thermal equilibrium before the following transient. In the following, a terminology of two letters behind the residue name will be used, the first denoting the ^{15}N frequency labelled origin, and the second the destination of magnetisation.

Generally, if relaxation losses during the INEPT transfer steps affect A and B differently, the initial and final polarisations $\langle \mathbf{N}_z \rangle_{a/b}(0)$ and $\langle \mathbf{N}_z \rangle_{a/b}(\tau_{mix})$ require further scaling with respect to the initial and final conformational state. However, unlike in the N-terminal SH3 domain of the protein drk, states A and B in saposin D are both folded monomeric proteins, presumably sharing a high structural similarity. It may therefore be assumed, that all relaxation rates (including $R_{1a} = R_{1b} = R_1$ of ^{15}N), and thus coherence transfer efficiencies, of corresponding spins in A and B are identical within experimental error. A series of N_z-exchange spectra with increasing transfer delays $1/(2J_{NH})$, $3/(2J_{NH})$, $5/(2J_{NH})$, as suggested by Tollinger *et al.*,[12] did not reveal differential relaxation behaviour, except Lys45 (see section 7.8).

Differential transverse relaxation (i.e. linewidths) of spins belonging to A and B should also be considered, when $\mathbf{Q}(\tau_{mix})$ is extracted from peak intensities rather than volumes. In particular, this may become significant in an asymmetric equilibrium at high k_{ex}, where the exchange contribution to the linewidths (k_a/π and k_b/π, in units of Hz) is different. In saposin D, a comparison of intensities (= data heights) $I(\tau_{mix})$ with peak volumes $V(\tau_{mix})$ from fitting 2D Gaussian lineshapes above noise level did not reveal significantly different behaviour. Even at 57 °C, linewidths are still dominated by intrinsic transverse relaxation and unresolved scalar couplings, and the influence of differential exchange line broadening may be neglected. Thus, intensities, which are much easier to obtain, were used for all data sets. Of all residues, Lys34 and Cys36 showed the required frequency separation in both ^{15}N and H^N dimensions, and for both residues three out of four peaks (aa, bb and ab) were isolated in the spectra at all temperatures and therefore easily identified and picked (**Fig. 7.4.1**).

At low temperatures, where exchange is largely suppressed, exchange peaks remain weak for all mixing times, and the auto peaks essentially follow monoexponential behaviour. With increasing temperature, the decay of $I_{aa}(\tau_{mix})$ of K34aa becomes faster due to increased R_1, thereby shortening the timescale for exchange, and biexponential. On the other hand, $I_{ab}(\tau_{mix})$ (from K34ab) builds up more efficiently and reaches its maximum at shorter τ_{mix}, before it also finally decays with R_1. Both curves approach each other, because in the case $k_{ex} \gg R_1$ magnetisation is rapidly distributed over the equilibrium. The temperature dependence of $I_{aa}(0)$ is a consequence of proton thermal polarisations, as well as temperature-different relaxation losses during INEPT transfer.

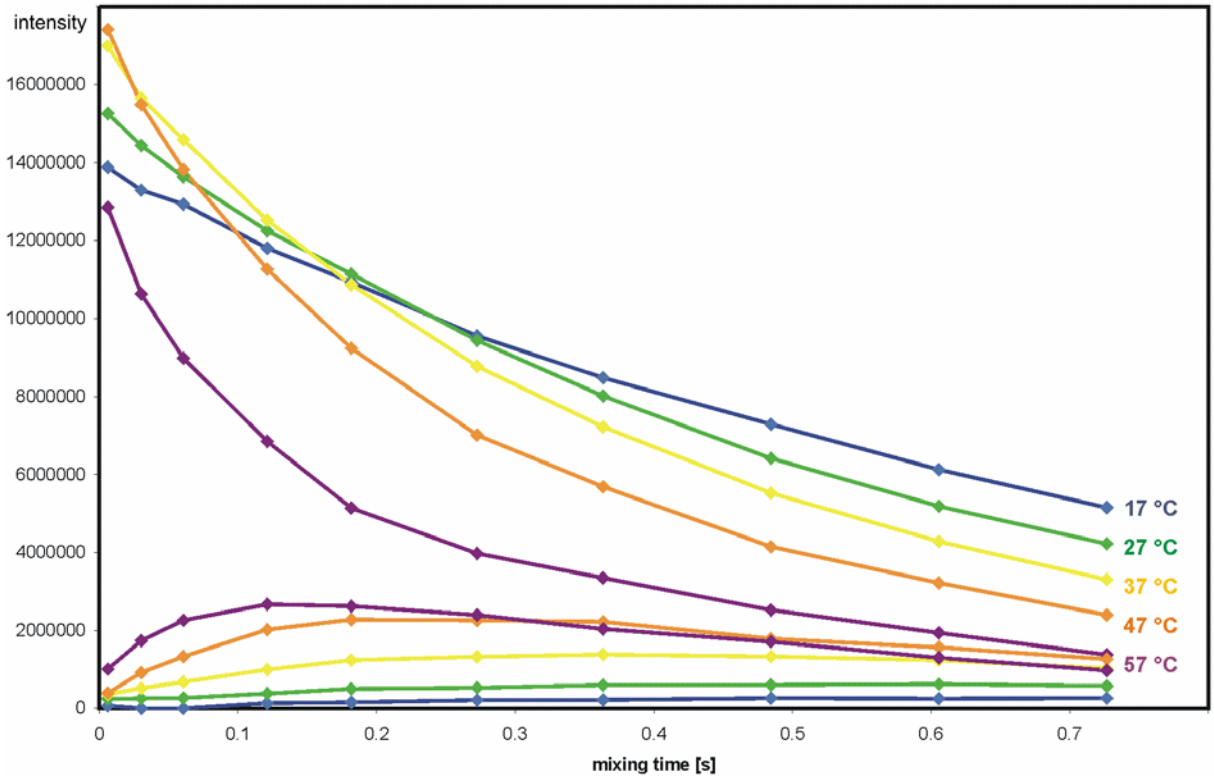


Fig. 7.4.1: N_z-exchange curves of K34aa (auto) and K34ab (exchange) signals, monitored by peak intensities $I_{aa}(\tau_{mix})$ and $I_{ab}(\tau_{mix})$ as function of the N_z-mixing period τ_{mix}. $I_{bb}(\tau_{mix})$ is not shown, but was also used for the analysis. Spectral parameters were used as specified in Fig. 7.3.2. Using 8 transients for each of the 192 t_1-increments and a recycle delay of 4 s, the experimental time was approximately 2 h per single N_z-exchange experiment, and 1 d for one series of τ_{mix}. Mixing times τ_{mix} were 8, 32, 62, 122, 182, 272, 362, 482, 602 and 722 ms. For the sake of clarity, the data points are connected with solid lines, with colours representing the temperatures: 17 °C (blue), 27 °C (green), 37 °C (yellow), 47 °C (orange) and 57 °C (magenta).

At high temperatures, the intercept $I_{ab}(0) \neq 0$ indicates that significant exchange takes place already at zero mixing time, i.e. during the refocussed INEPT transfer of $1/(J_{NH})$, leading to high χ^2 in fitting the data against **Q**. The presence of chemical exchange in the nuclear spin Hamiltonian can principally not be influenced by RF pulses, and therefore needs to be accounted for by an improved formula, which will be briefly derived in the following:

Assuming that the pair of polarisations ($\langle \mathbf{N}_z \rangle_a(t)$, $\langle \mathbf{N}_z \rangle_b(t)$), having evolved under $\mathbf{Q_1}(t)$, is not directly detected at $t = \tau_{mix}$, but rather subject to another propagator $\mathbf{Q_2}(t)$ for a fixed time τ_{off}, the final signal will be:

$$\begin{pmatrix} \langle \mathbf{N}_z \rangle_a (t+\tau_{off}) \\ \langle \mathbf{N}_z \rangle_b (t+\tau_{off}) \end{pmatrix} = \mathbf{Q}_2(\tau_{off})\mathbf{Q}_1(t) \begin{pmatrix} \langle \mathbf{N}_z \rangle_a (0) \\ \langle \mathbf{N}_z \rangle_b (0) \end{pmatrix} = \mathbf{Q}(t) \begin{pmatrix} \langle \mathbf{N}_z \rangle_a (0) \\ \langle \mathbf{N}_z \rangle_b (0) \end{pmatrix} \tag{7.4.1}$$

A further simplification is, that over the re-INEPT period τ_{off} magnetisation is simultaneously subject to some averaged transverse (^{15}N and ^{1}H) relaxation R_{av} (the same for A and B), and chemical exchange k_a and k_b, thus expressed in the Liouvillian operator:

$$\mathbf{L}_2 = \begin{pmatrix} -R_{av} - k_a & k_b \\ k_a & -R_{av} - k_b \end{pmatrix} \tag{7.4.2}$$

$\mathbf{Q}_2(\tau_{off}) = \exp(\mathbf{L}_2 \tau_{off})$ then adopts the same form as $\mathbf{Q}_1$, with matrix elements given in analogy to equation 1.4.4. The matrix product $\mathbf{Q}_2(\tau_{off})\,\mathbf{Q}_1(t)$ is readily shown to be:

$$\mathbf{Q}_2(\tau_{off})\mathbf{Q}_1(t) = \frac{e^{-R_{av}\tau_{off}}}{k_{ex}} \begin{pmatrix} k_b e^{-R_1 t} + k_a e^{-R_1 t - k_{ex}(t+\tau_{off})} & k_b e^{-R_1 t} - k_b e^{-R_1 t - k_{ex}(t+\tau_{off})} \\ k_a e^{-R_1 t} - k_a e^{-R_1 t - k_{ex}(t+\tau_{off})} & k_a e^{-R_1 t} + k_b e^{-R_1 t - k_{ex}(t+\tau_{off})} \end{pmatrix} \tag{7.4.3}$$

This resembles equation 1.4.4 with an additional damping constant $\exp(-R_{av}\tau_{off})$, but now all k_{ex}-terms (not the R_1-terms) depend on $t+\tau_{off}$ instead of t, because exchange takes place during both periods. At $t = 0$ (zero mixing time), the ab and ba matrix elements no longer cancel to zero, but have already accumulated to $k_{a/b}(1-\exp((-k_{ex}\tau_{off}))$, and the aa and bb elements have dropped by a similar factor. With the assumptions made in the previous section, the final signals $I(\tau_{mix})$ after a discrete mixing time τ_{mix} depend on the respective initial polarisation $\langle \mathbf{N}_z \rangle(0)$ and its evolution according to equation 7.6.3:

$$I_{ij}(\tau_{mix}) = \langle \mathbf{N}_z \rangle_i(0)\; q_{ij}(\tau_{mix}) \tag{7.4.4}$$

where $q_{ij}(\tau_{mix})$ with i,j = a,b are the matrix elements of $\mathbf{Q}(\tau_{mix})$ as specified in equation 7.4.3.

7.5 The N_x-exchange experiment

As transverse spin relaxation slows down, owing to faster molecular tumbling at elevated temperatures, rates k_{ex} of conformational exchange start to compete also with transverse ^{15}N relaxation rates R_2. Thus, transverse N_x instead of longitudinal N_z magnetisation may be used during the mixing period for the simultaneous study of ^{15}N-R_2 and k_{ex}. The *N_x-exchange experiment* required merely minor modifications in the pulse scheme for N_z-exchange, and was implemented and tested on the same sample of saposin D (**Fig. 7.5.1** and section 9.8).

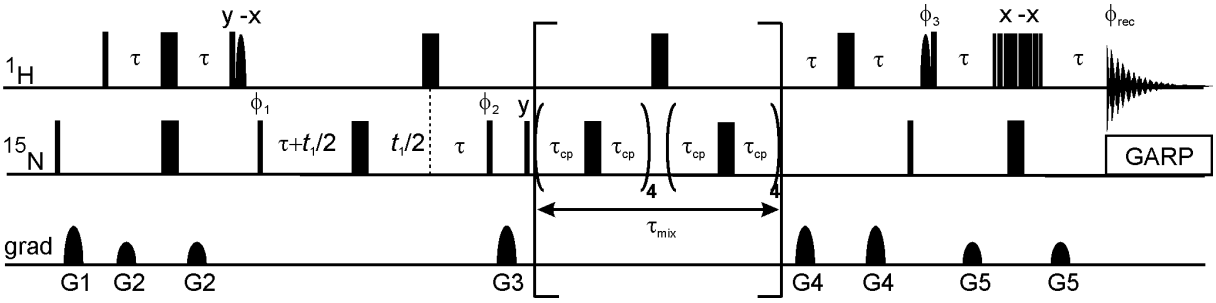


Fig. 7.5.1: Pulse scheme for the measurement of N_x-exchange. Narrow and wide bars represent 90° and 180° pulses, respectively, which are applied with phase x, unless indicated otherwise. τ was set to 2.5 ms and the CPMG mixing period τ_{mix} consisted of repetitions of the cycle in brackets (~8 ms with τ_{CP} = 450 µs). Again, 180° pulses on protons were applied once per cycle to suppress dipole-CSA cross-correlated relaxation. For an experiment with zero mixing time, all pulses in the brackets were replaced by a delay of 3 µs. Water selective 90° pulses were applied with a Gaussian shape and a length of 2 ms, and final water suppression was achieved with a 3-9-19 binomial sequence with a delay of 200 µs between the pulses. Pulse phases: ϕ_1 = x,−x, ϕ_2 = 2(y),2(−y), ϕ_3 = 4(x),4(−x) and ϕ_{rec} = x,2(−x),x,−x,2(x),−x. Relative gradient strength (where 100 corresponds to 33 Gauss/cm): G_1 = 30, G_2 = 15, G_3 = 60, G_4 = 25, G_5 = 20.

Initial polarisation transfer from ^1H to ^{15}N spins and t_1 evolution elements are completely identical, however, longitudinal ^{15}N magnetisation is subject to a z-filter element (ensuring pure phases) and subsequently flipped to the x-axis. After the following CPMG mixing period, through which ^{15}N spins undergo transverse relaxation and chemical exchange, coherence transfer from ^{15}N to ^1H spins is again analogous to the N_z scheme.

As described in section 1.3, during a CPMG pulse train, transverse magnetisation relaxes with an effective rate R_{2eff}, which depends on τ_{CP} and may strongly oscillate, if $\nu_{CP} = 1/(4\tau_{CP})$ is smaller than $\Delta\omega_{ab}/2\pi$. In Saposin D, Lys34 has the largest ^{15}N frequency separation $\Delta\omega_{ab}/2\pi$,

ranging from 130 Hz at 17 °C to 86 Hz at 57 °C and 14.1 T. With the chosen CPMG delay τ_{CP} of 450 µs (corresponding to $\Delta\nu_{CP}$ = 555 Hz), all ^{15}N spins should approximately relax with their intrinsic transverse relaxation rates R_2 without significant exchange contributions. In this particular case, and assuming equal intrinsic relaxation rates $R_{2a} = R_{2b} = R_2$, the equation of motion, and thus the propagator **Q**, resemble the one for polarisations, as given by equation 1.4.4, with R_1 replaced by R_2.

The spectra have the same appearance, and were recorded with the same parameters as for the N_z-exchange, but here the decay of auto and exchange peaks is governed by the ^{15}N transverse relaxation rates. Compared to R_1, R_2 has an inverse temperature dependence, with a rapid decay for low temperatures, during which exchange has no time to build up. Only above 37°, exchange peaks clearly emerge from the noise, and they become more intense with increasing k_{ex} and with the extending $1/R_2$ timescale.

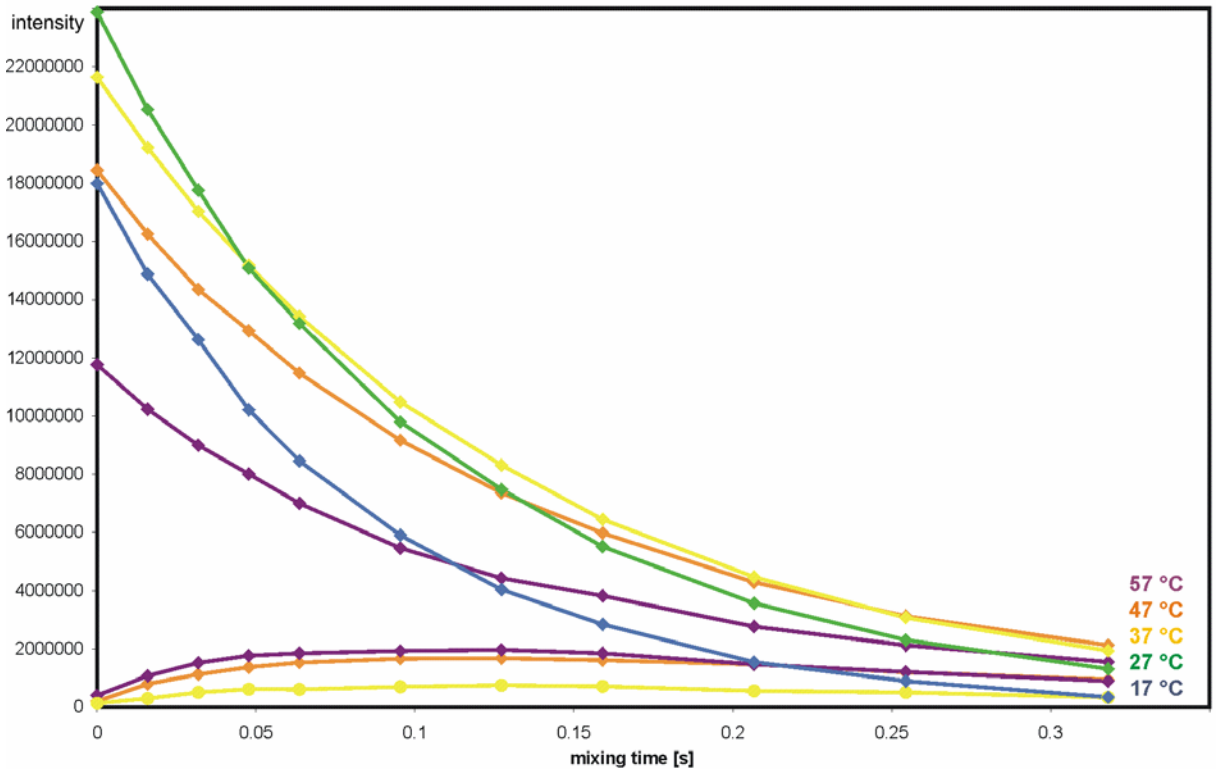


Fig. 7.5.2: N_x-exchange curves of K34aa (auto) and K34ab (exchange) signals, monitored by peak intensities $I_{aa}(\tau_{mix})$ and $I_{ab}(\tau_{mix})$ as function of the N_x-mixing period τ_{mix}. $I_{bb}(\tau_{mix})$ is not shown, but was also used for the analysis. Spectral parameters were exactly as specified in Figures 7.3.2 and 7.4.1. Mixing times τ_{mix} for all temperatures were 0, 16, 32, 48, 64, 96, 128, 160, 208, 256 and 320 ms, the largest τ_{mix} corresponding to approximately 1.5*T_2 at 57 °C and 3*T_2 at 17 °C. Colour coding is as in Fig. 7.4.1.

7.6 Fit results

For residues Lys34 and Cys36, the missing $I_{ba}(\tau_{mix})$ was set to $I_{ab}(\tau_{mix})$, and all four $q_{ij}(\tau_{mix})$ were fitted simultaneously against $\langle \mathbf{N}_z \rangle_a(0)$, $\langle \mathbf{N}_z \rangle_b(0)$, R_1, k_a, k_b and τ_{off} (equations 7.6.3 and 7.6.4). This was done with a script implemented in Matlab (*MathWorks Inc., Natick, MA, USA*), which requires four separate input files for I_{aa}, I_{bb}, I_{ab} and I_{ba} with residues listed in rows, and mixing periods in columns, and which generates a plot for each residue (**Fig. 7.6.1**) and a single file containing results for all residues. The generation and fitting of randomly generated data sets for Monte-Carlo-type simulations and error estimates of the results has not been implemented yet.

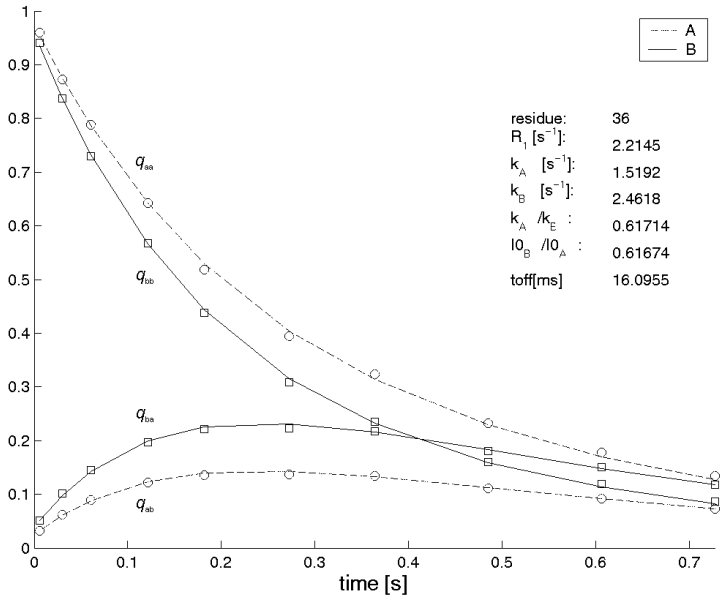


Fig. 7.6.1: Plot of q_{ij} for N_z-exchange data from Cys36 at 47 °C. Squares and circles indicate experimental data heights of magnetisation originating on A (q_{aa} and q_{ab}) and B (q_{bb} and q_{ba}), respectively. Solid and dashed lines represent the fitted curves according to equations 7.6.3 and 7.6.4. The ratio $I0_B/I0_A$ denotes the ratio of initial polarisations $\langle \mathbf{N}_z \rangle_b(0)/\langle \mathbf{N}_z \rangle_a(0)$.

Especially at 47 °C and 57 °C, excellent N_z data fits were obtained with τ_{off} in the order of the expected period of 4τ (≈ 10 ms) for refocussed INEPT transfer, thus τ_{off} = 10 ms was kept fixed for the other temperatures and all N_x data. The equilibrium constant as defined by the ratio k_a/k_b agreed well with the ratio of initial polarisations $\langle \mathbf{N}_z \rangle_b(0)/\langle \mathbf{N}_z \rangle_a(0)$, indicating that the assumptions of fully relaxed conditions and equal relaxation rates apply here. Within experimental error, both residues Lys34 and Cys36 yielded the same exchange, but also R_1 relaxation rates, which may be explained by isotropic molecular tumbling and the absence of further internal motions, as it was found in the homologue saposin C (section 5.6). The relaxation rates also correspond to the quasi monoexponential decay curves obtained for residues with only small frequency separations between A and B.

Exchange rates from N_x data agree with those from N_z data at 37 °C and above, however, as R_2 is approximately 10 % higher for Lys34 at high temperatures, CSM contributions to R_2

may not be fully negligible. Below 37 °C, exchange peaks merely rise above noise level, and only the auto peaks were subject to monoexponential fitting, yielding R_2 and $\langle \mathbf{N}_z \rangle_b(0)/\langle \mathbf{N}_z \rangle_a(0)$. The results for residues Lys34 and Cys36 are summarised in **Table 7.6.1**.

Table 7.6.1: Temperature dependence of ^{15}N-R_1 and R_2 relaxation rates, forward and backward exchange rates k_a and k_b, their ratio, and ratio of $\langle \mathbf{N}_z \rangle(0)$, from fitting N_z and N_x (in brackets) data with fixed τ_{off} = 10 ms.

	R_1 (R_2) [s^{-1}]	k_a [s^{-1}]	k_b [s^{-1}]	k_a/k_b	$\langle \mathbf{N}_z \rangle_b(0)/\langle \mathbf{N}_z \rangle_a(0)$
17 °C Lys34	1.09 (10.96 [a])	0.11 (-)	0.23 (-)	0.46 (-)	0.45 (0.47 [a])
Cys36	1.12 (10.78 [a])	0.09 (-)	0.20 (-)	0.43 (-)	0.42 (0.41 [a])
27 °C Lys34	1.48 (7.96 [a])	0.21 (-)	0.44 (-)	0.48 (-)	0.50 (0.52 [a])
Cys36	1.42 (7.88 [a])	0.23 (-)	0.47 (-)	0.49 (-)	0.50 (0.51 [a])
37 °C Lys34	1.84 (6.16)	0.62 (0.44)	1.18 (1.01)	0.53 (0.44)	0.54 (0.49)
Cys36	1.88 (6.25)	0.63 (0.70)	1.16 (1.40)	0.55 (0.50)	0.55 (0.57)
47 °C Lys34	2.25 (4.95)	1.58 (1.51)	2.69 (2.52)	0.59 (0.60)	0.59 (0.60)
Cys36	2.21 (4.69)	1.52 (1.69)	2.46 (2.74)	0.62 (0.62)	0.62 (0.63)
57 °C Lys34	2.38 (4.50)	3.84 (4.13)	6.07 (6.33)	0.63 (0.65)	0.63 (0.68)
Cys36	2.43 (4.21)	3.79 (3.70)	6.03 (5.89)	0.65 (0.64)	0.64 (0.67)

[a] obtained from monoexponential fitting.

The relaxation rates R_2/R_1 drops from 10 (17 °C) to below 2 (57 °C) for Cys36, corresponding to a decrease in the molecular tumbling time from τ_c = 9.6 ns to 3.1 ns (7.2 ns to 2.9 ns in hydrodynamic simulations). As in saposin C, the glycosylation and the hexa-histidine tag, but also intermolecular interactions or a larger hydration shell may serve as reason for the 10 % increase of the apparent radius of gyration leading to this discrepancy at low temperatures.

Information about the relative energies and entropies of A, B and the transition state between the two may be gained from a semi-logarithmic plot of k_a, k_b and K versus $(1/T)$ (**Fig. 7.6.2**), corresponding to the logarithmic form of equations 1.1.1 and 1.1.2:

$$\ln K = -\frac{\Delta G_{ab}}{RT} = -\frac{\Delta H_{ab}}{RT} + \frac{\Delta S_{ab}}{R} \qquad (7.6.1)$$

$$\ln k_a = -\frac{\Delta G^{\neq}{}_a}{RT} + \ln\frac{kT}{h} = -\frac{\Delta H^{\neq}{}_a}{RT} + \frac{\Delta S^{\neq}{}_a}{R} + \ln\frac{kT}{h} \qquad (7.6.2)$$

Since (ln kT/h) varies only from 29.4 at 17 °C to 29.6 at 57 °C, ln(k_a) and ln(k_b) yield approximately straight lines with the intercepts $\Delta S^{\neq}{}_a/R + 29.5$ and $\Delta S^{\neq}{}_a/R + 29.5$, respectively.

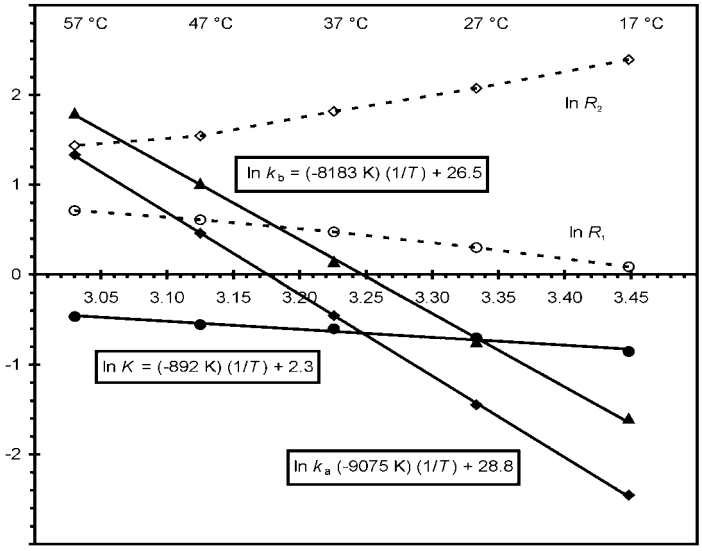


Fig. 7.6.2: Arrhenius plot of ln(k_a) (filled diamonds), ln(k_b) (filled triangles) and ln(K) = ln(k_a/k_b) (filled circles) versus (1/T) from the results for residue Cys36. The straight lines, labelled with their slope and intercept, were obtained by a least squares fit to the data points. ^{15}N R_1 (open circles) and R_2 (open diamonds) relaxation rates are shown for comparison.

The slope of ln(K) corresponds to a characteristic temperature of 892 K and therefore to a positive equilibrium enthalpy of ΔH_{ab} = 1.77 kcal/mol in favour of conformer A, which is approximately the strength of one H-bond. However, the positive intercept 2.3 at (ΔS_{ab} = 4.57 cal/molK) infinite temperature shows that conformer B is slightly more disordered and thus entropically stabilised by −1.42 kcal/mol at 37 °C. An apparent compensation of enthalpy and entropy is often found for kinetic data from biological systems,[191] but this principle has been challenged recently.[192] A full compensation (i.e. ΔG_{ab} = 0) in saposin D is expected at approximately 115 °C and above this temperature, B becomes the major conformer.

From the slopes of ln(k_a) and ln(k_b) activation enthalpies of $\Delta H^{\neq}{}_a$ = 18.02 kcal/mol $\Delta H^{\neq}{}_b$ = 16.25 kcal/mol, respectively, are obtained, which approximately corresponds to the rotation barrier of a peptide bond, but also to the barrier of disulfide isomerisation in BPTI.[193] The intercepts of both curves indicate negative activation entropies (i.e. an ordered transition state) of $\Delta S^{\neq}{}_a$ = −1.4 cal/molK and $\Delta S^{\neq}{}_a$ = −6.0 cal/molK, but these values may be afflicted with a large extrapolation error. The observation of a single predominant conformer in the other saposins A and C, which should in principle be able to undergo the same exchange process, reflects either a strong stabilisation of one state or a low isomerisation barrier, i.e. a fast exchange process. For example, a large variation of kinetic data of BPTI upon a single point mutation has been observed in BPTI.[185]

7.7 Exchange of residual dipolar couplings in saposin D

Separate values D_a and D_b were obtained for many residues in the two conformers A and B, raising the question, how dipolar or general couplings J_a and J_b are averaged under the influence of chemical exchange. In saposin D, apparent *J*-splittings were measured from ^{15}N frequencies, *J*-modulated during t_1-evolution, and RDCs were calculated as difference of these values in weakly aligned and isotropic phase. Thus, one needs to calculate how the system of four coherences, each of which gives rise to a single ^{15}N resonance (**Fig. 7.7.1**), evolves during the t_1-period.

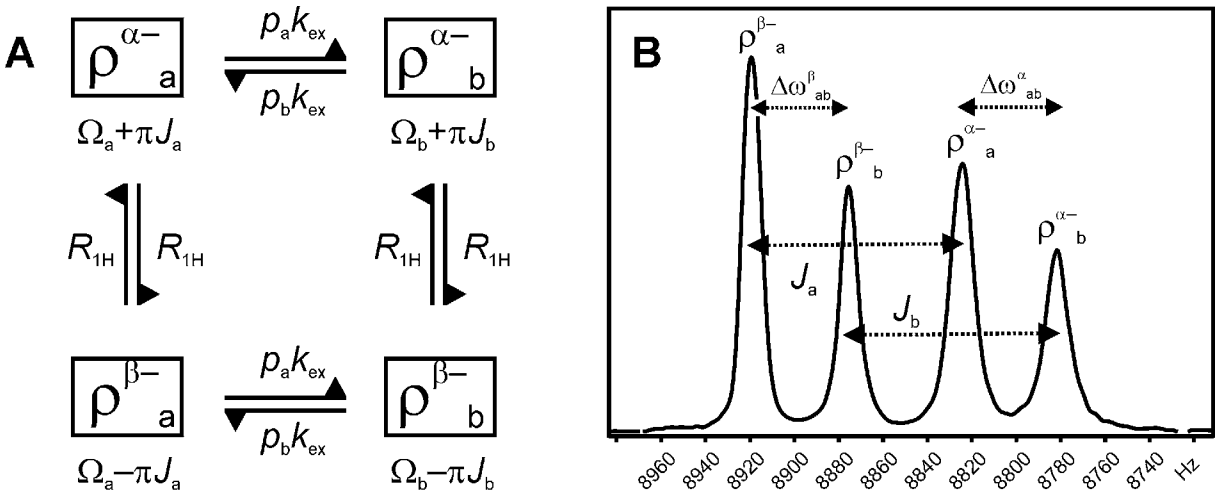


Fig. 7.7.1: A: System of exchanging ^{15}N (–1) quantum coherences of an NH spin pair in the presence of chemical exchange. A slight bias of ^1H spin state populations towards α, corresponding to a difference in R_{1H} for upward and downward transitions, is neglected. **B:** Hypothetical ^{15}N spectrum (in units of Hz) of residue Cys36, generated by adding the ^{15}N cross sections through C36a and C36b of an IP-[^1H,^{15}N]-HSQC subspectrum. Coherences and couplings are labelled as defined in the text.

For example, $\rho^{\alpha-}{}_a$ denotes ^{15}N (–1) quantum coherence with the attached proton in the spin state α, averaged over all molecules in state A, which is connected to $\rho^{\alpha-}{}_b$ with the familiar rate constants $p_bk_{ex} = k_a$ and $p_ak_{ex} = k_b$ of the exchange equilibrium. Furthermore, owing to transitions of the proton spins, $\rho^{\alpha-}{}_a$ equilibrates with $\rho^{\beta-}{}_a$ with the spin lattice relaxation rate of the protons R_{1H}, which can be assumed to be identical for A and B for the same reason as discussed in section 7.1 and is typically in the order of 1 s^{-1}. Only for small $2\pi J_a$ and $2\pi J_b$ (smaller than R_{1H}), the observable ^{15}N splitting is expected to break down, and averaged signals at Ω_a and Ω_b appear, this is, however, never the case for large one-bond couplings.

The rather complicated general solution of this system requires diagonalisation of a 4 x 4 Liouvillian matrix, which is beyond the scope of this text. However, if $k_{ex} \gg R_{1H}$, the proton spin state is retained on the exchange timescale, and the equilibrium of four coherences can be decomposed into two equilibria of two coherences and thus be treated exactly as section 1.2. The only difference is that the exchange regime and thus the lineshape is now effectively governed by ^{1}H spin state dependent frequency differences:

$$\Delta\omega^{\alpha}_{ab} = (\Omega_b + \pi J_b) - (\Omega_a + \pi J_a) = \Delta\omega_{ab} + \pi\Delta J_{ab} \qquad (7.7.1)$$

$$\Delta\omega^{\beta}_{ab} = (\Omega_b - \pi J_b) - (\Omega_a - \pi J_a) = \Delta\omega_{ab} - \pi\Delta J_{ab}$$

Assuming $\Delta J_{ab} = J_b - J_a \approx 0$ in isotropic solution (pure scalar coupling) and $\Delta J_{ab} = \Delta D_{ab}$ in oriented solution, fast exchange averaging of resonance frequencies of one of the dublet components may occur, if $\Delta D_{ab}/2$ is on the order of $\Delta\omega_{ab}./2\pi$. Of the two residues in saposin D with remarkably high $\Delta D_{ab} \approx 6$ Hz, this applies to Ser37 ($\Delta\omega_{ab}/2\pi \approx 5$ Hz at 37 °C), but not to Cys36 ($\Delta\omega_{ab}/2\pi \approx -35$ Hz at 37 °C). Unfortunately, in Ser37, the temperature dependence of RDCs could not be measured owing to overlap with side chain resonances above 37 °C. However, the exceptional thermal stability of saposin D and Pf1 phage allowed the measurement of $\Delta\omega^{\alpha}_{ab}$ and $\Delta\omega^{\beta}_{ab}$ and thus ΔD_{ab} for Cys36 up to the probe´s limit at approximately 72 °C (**Fig. 7.7.2**).

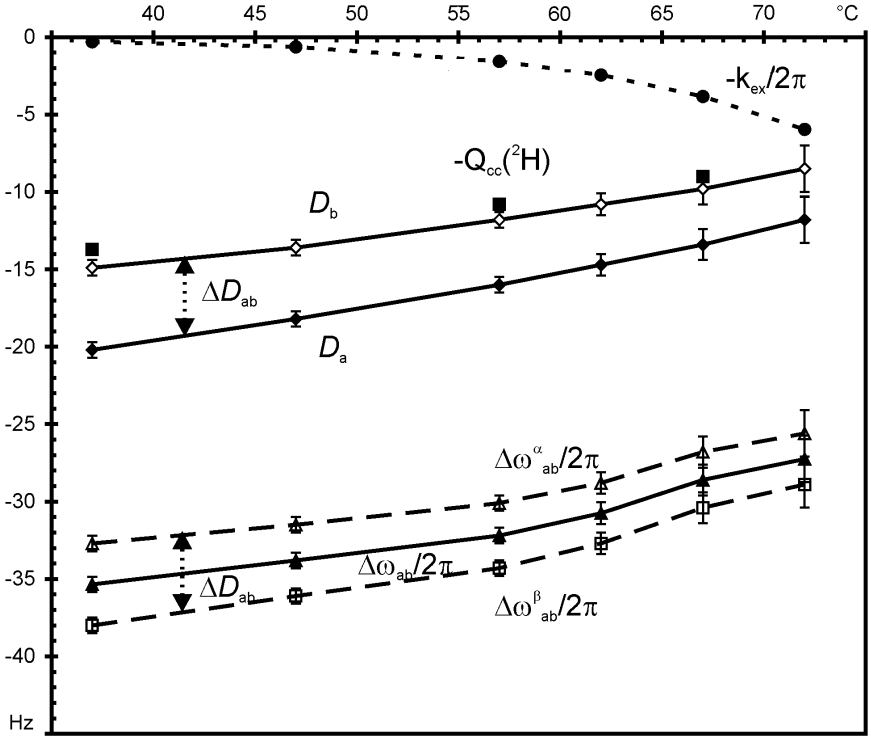


Fig. 7.7.2: Temperature dependence of D_a (filled diamonds) and D_b (open diamonds, data points are connected by solid lines for both), $\Delta\omega^{\alpha}_{ab}$ (filled diamonds) and $\Delta\omega^{\beta}_{ab}$ (open diamonds, connected by dashed lines). $\Delta\omega_{ab}$ is the average of $\Delta\omega^{\alpha}_{ab}$ and $\Delta\omega^{\beta}_{ab}$. ^{2}H quadrupolar splittings Q_{cc} (squares) were measured at 37 °C, 57 °C and 67 °C, and k_{ex} (circles) was extrapolated for temperatures above 57 °C with the results of section 7.6.

D_a and D_b were obtained from t_1-coupled ^{15}N-HSQC experiments by fitting Gaussian lineshapes, assuming temperature-independent isotropic couplings $J_a = J_b = -94.5$ Hz (measured at 37 °C). The results indicate that RDCs are not averaged between states A and B, even though $k_{ex} > 2\pi\Delta D_{ab}$ at high temperatures, and the decrease of D_a, D_b and ΔD_{ab} merely reflects thermally decreased alignment of Pfl phage (**Fig. 7.7.2**). At high temperatures, however, k_{ex} increases towards the intermediate exchange regime of $\Delta\omega^{\alpha}_{ab}$ and $\Delta\omega^{\beta}_{ab}$, therefore broadening resonances and impeding the accurate measurement of frequencies. Besides, intrinsic resonance frequencies may be superimposed by a dynamic shift of resonances, which reaches the order of ΔD_{ab} at approximately $k_{ex} = \sqrt{2}\pi\Delta D_{ab}\Delta\omega_{ab}$,

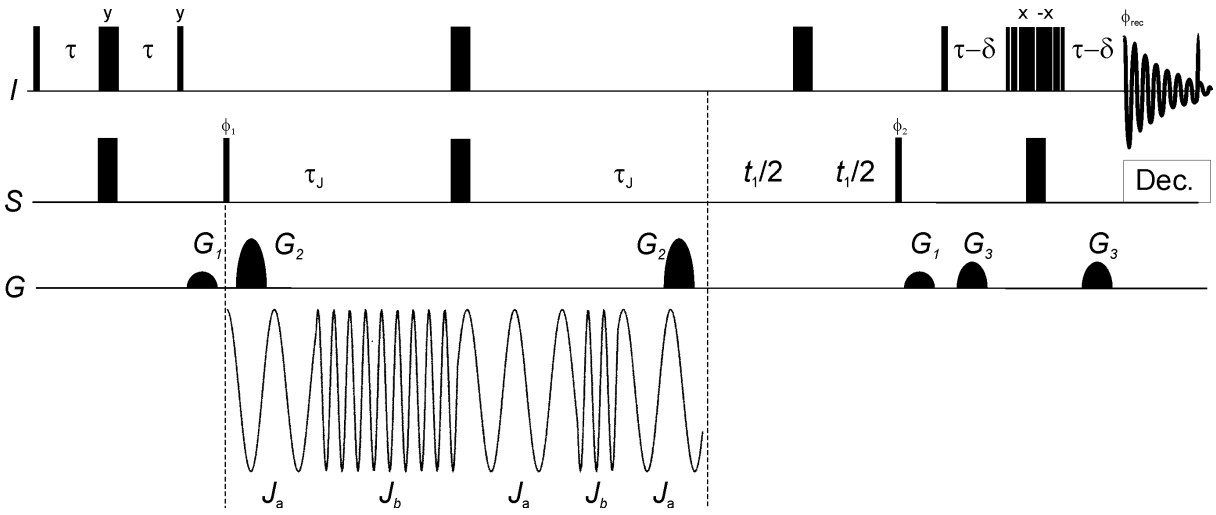


Fig. 7.7.3: Pulse sequence of *J*-modulated ^{15}N-HSQC experiment.[194] Parameters are the same as described in Fig. 2.4.1. For an expected J_{NH}, τ_J is sampled in the region $\tau_J = (n+1)/(2J_{NH})$. Below the sequence the modulation of one particular NH signal with J_a and J_b is shown, which is subject to averaging in the presence of chemical exchange.

The averaging of J_a and J_b is fundamentally different in quantitative-*J*-correlation experiments,[105, 153] in which the amplitude modulation of the signal during a pure *J*-evolution period τ_J is recorded in a series of 2D spectra (**Fig. 7.7.3**). If τ_J is sampled efficiently in an interval around the true $\tau_J = (n+1)/(2J_{NH})$ yielding zero amplitude, RDCs may be obtained with approximately five times higher accuracy compared to conventional frequency splittings.[194] However, chemical exchange during this period acts on $2\pi\Delta J_{ab}$ exactly like on $\Delta\omega_{ab}$ in a chemical shift evolution period (equations 1.2.4 and 1.2.5) and leads to an averaging of J_a and J_a to $J_{av} = p_a J_a + p_b J_a$ for $k_{ex} > 2\pi\Delta J_{ab}$.

7.8 Hydrogen exchange in saposin D

In the ^{15}N-HSQC spectrum of saposin D (**Fig. 7.1.2**), several resonances are significantly broadened in both dimensions owing to hydrogen exchange with the solvent during t_1-evolution and acquisition. One of them is Lys45a, which is expected to be situated in the first turn of a helix (α3), where amide protons are often not H-bonded. Remarkably, no broadening was obvious for its well resolved counterpart Lys45b, indicative of a high protection of this amide in conformer B. In N_z- and N_x-exchange spectra the peaks of Lys45 were broad only in the spectral dimension belonging to the coherences and frequencies of state A (**Fig. 7.8.1**). Due to differential apparent relaxation rates during INEPT transfers, t_1-evolution and acquisition periods, fitting N_z and N_x intensities or volumes to equation 7.4.3 failed for this residue.

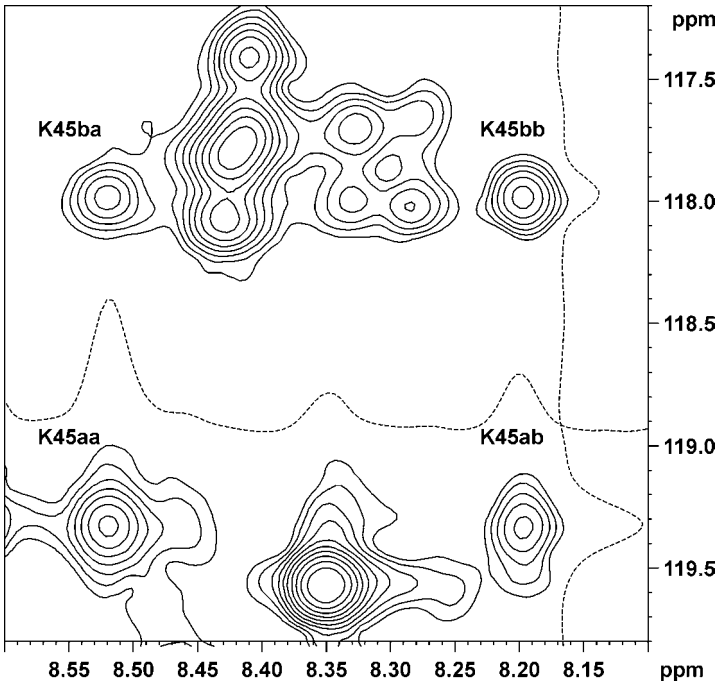


Fig. 7.8.1: Appearance of auto and exchange signals of residue Lys45 in the N_z-exchange spectrum as described in Fig. 7.3.2. 1D-traces, drawn as dashed lines, are taken at the ^{15}N and ^{1}H frequencies of K45aa.

In order to investigate, whether hydrogen exchange of Lys45b requires a conformational transition to the fast exchanging Lys45a (i.e. depends on k_a and k_b), NewMEXICO experiments were applied to saposin D at 37 °C. To overcome deleterious radiation damping effects on cryogenic probes, the existing pulse sequence was slightly modified by including a weak field gradient every 5 ms in the mixing period and an improved readout ^{15}N-HSQC scheme (**Fig. 7.8.2** and section 9.9). A single gradient pulse was not able to keep water along −z for more than approximately 40 ms, leading to a subtraction of exchange-transferred polarisation from transients with water along +z (**Fig. 7.8.3**).

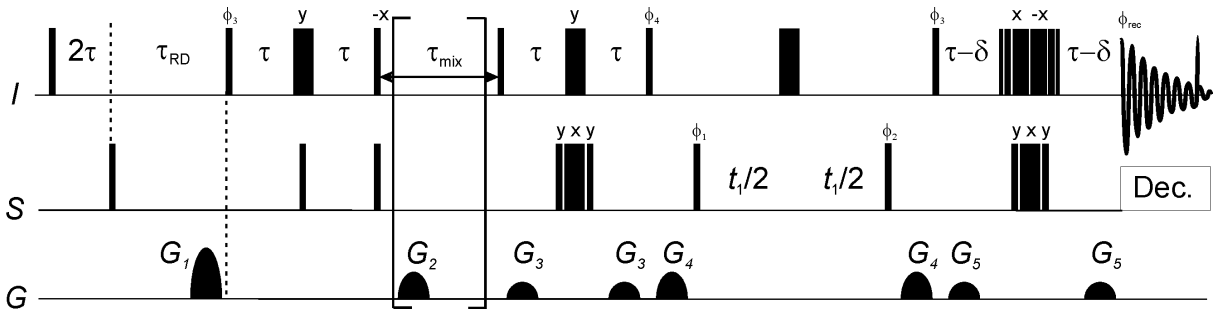


Fig. 7.8.2: Pulse scheme of the modified NewMEXICO experiment.[44] The first two pulses saturate all protons and create ^{1}H-^{15}N multiple quantum coherence. Water z-magnetisation is re-established after typically τ_{RD} = 20 ms on a cryoprobe and τ_{RD} = 40 ms on a conventional probe via radiation damping. Carbon bound protons maintain > 90% saturation during this period. The second filter only saturates amide protons and alternatively flips the water magnetisation to +z and –z (ϕ_3 = 4x,4(–x)). Polarisation is transferred to amide protons during the mixing delay τ_{mix}, implemented as 1...n repetitions of the brackets (5 ms). Apart from proton pulse phase cycling, ensuring that water is along +z before acquisition in all transients, the readout ^{15}N-HSQC corresponds to the scheme of Fig. 2.4.1 and was also used as reference. ϕ_1 = x,–x, ϕ_2 = 2x,2(–x), ϕ_4 = 4y,4(–y), ϕ_{rec} = x,2(–x),x, –x,2x,–x. Relative strength of gradient pulses (1 ms): G_1 = 80, G_2 = 30, G_3 = 15, G_4 = 25, G_5 = 20.

Although amide exchange is indeed fast for Lys45a, the rate k_{ex} = 17.2 s^{-1}, obtained from the initial slope, is only about half the one for the fastest exchanging assigned residue, Phe4 (k_{ex} = 38 s^{-1}) (**Fig. 7.8.3**). For Lys45b, hydrogen exchange is significantly slower (k_{ex} = 4.5 s^{-1}), but still much faster than k_a and k_b, indicating that this proton is exchangeable by structural fluctuations within conformer B, which are distinct from the transition to conformer A.

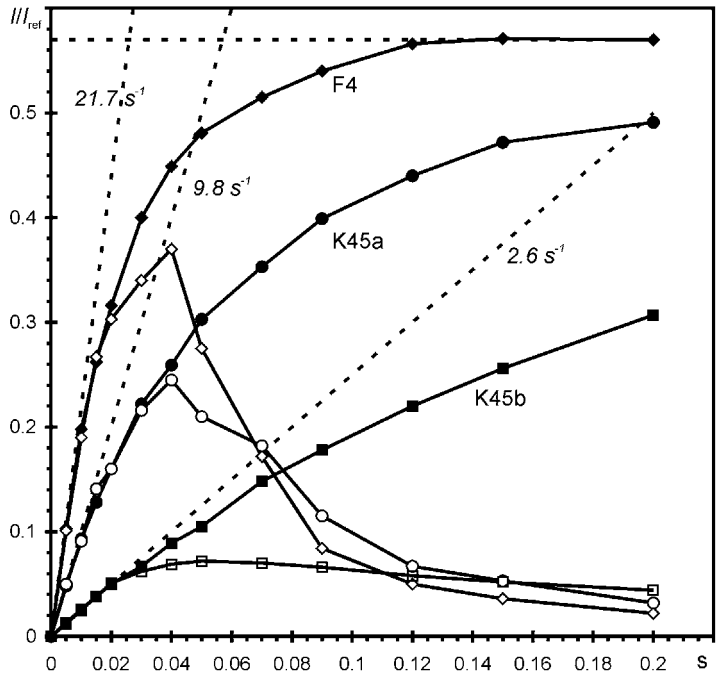


Fig. 7.8.3: $I(\tau_{mix})/I_{ref}$ of NewMEXICO experiments versus τ_{mix} (= 5, 10, 15, 20, 30, 40, 50, 70, 90, 120, 150 and 200 ms) for residues Phe4 (diamonds), Lys45a (circles) and Lys45b (squares). Open and filled symbols connected with solid lines indicate data with single gradient and a series of gradients during the mixing period, respectively. The initial slopes Fk_{ex} are drawn by dashed lines and labelled with their values. F = 0.57 (the average water polarisation) is given by the plateau of Phe4, which is also drawn as dashed line.

8 Summary

In this thesis, different proteins are characterised by NMR spectroscopy in terms of association state as well as structural features and exchange processes in solution. In none of the projects encountered a complete structure determination was feasible, due to either limited sample stability, poor spectral performance or simply obvious agreement of the data with existing structures. Instead, the work focusses on information which is obtained relatively fast, such as secondary structure indicators, residual dipolar couplings, ^{15}N relaxation- and hydrogen exchange rates.

In the project presented in chapter 3, the two isolated 17 kDa apical domains of the 930 kDa thermosome from *Thermoplasma acidophilum*, αADT and βADT, were studied, whose protrusion regions were believed to be strongly influenced by crystal packing interactions. Backbone assignment of both ^{15}N, ^{13}C labelled proteins was accomplished using common triple resonance experiments, unfortunately, βADT suffered from fast hydrogen exchange under basic sample conditions. The ^{15}N-HSQC-pattern as well as chemical shifts of H^{α}, C^{α}, C^{β} and C′ nuclei turned out highly homologous, and further analysis of $^{3}J_{HNH\alpha}$ and $d_{\alpha N}$- and d_{NN}-NOEs in αADT proved that the globular core regions adopt the crystalline fold. In contrast, the N-proximal half of the protrusion seems to be unstructured, rather than forming a short α-helix as in the crystal, and the first turns of H10 in the C-proximal half undergoes helical fraying. Still, amide exchange rates and a ^{15}N-HSQC titration with TFE strongly indicated that the βI-turn at the very tip of the helical protrusion is preserved in solution. ^{15}N backbone relaxation data support the picture of intrinsic disorder in large parts of the ADT protrusions.

The mature GM2 activator protein bearing ^{15}N-labelling and a high-mannose type glycosylation from efficient expression in *Pichia pastoris* cells was subjected to NMR spectroscopy at varying pH (4.5 and 7.0), temperature, concentration and in the presence and absence of lipids (chapter 4). Diffusion- and relaxation studies revealed, that the protein is monomeric in solution with only little tendency to form associates, and a secondary structure analysis confirms a prevalence of β-sheets as found in the crystals. However, conformational heterogeneity and exchange processes at various timescales continued to be obstacles for a detailed NMR spectroscopic characterisation. Before envisaging a structure determination and binding studies at atomic resolution, sample lifetime and ^{15}N-HSQC quality need to be further optimised in a systematic and possibly time consuming process.

Samples of all four human saposins A-D, bearing uniform ^{15}N-labelling, glycosylation and a hexa-histidine tag from purification, were examined by NMR-spectroscopy at variable temperature and pH (chapter5). Although highly homologous and purely α-helical under all conditions of study, the individual proteins show a remarkably different behaviour in solution. For samples showing sufficient spectral resolution, backbone ^{15}N, H^N and H^α chemical shifts were assigned using a combination of five ^{15}N-edited 3D experiments. By far the best spectral quality was obtained for saposin C, whose secondary structure was analysed to yield the five α-helices expected from NK-lysin. However, the determination of the fold by constructing an NOE network of methyl groups suffered from ^{1}H chemical shift degeneracy and low sensitivity of ^{13}C edited experiments at natural abundance and was not further pursued, after the solution structure had been solved by another research group. At variance with the authors, at acidic conditions and elevated temperatures specific dimerisation of saposin C was observed, presumably accompanied by a rearrangement of helices. The other saposins are rather heterogenous in solution, varying from a pH-independent two site equilibrium in saposin D, to a highly and unresolvable dynamic system in saposin B.

A conformational analysis using residual dipolar couplings of saposin A, which is strongly prone to aggregation at pH 4.0, and saposin D is the subject of chapter 6. In order to obtain weakly oriented solutions of human saposins, different media, which are well established for the alignment of biolocical macromolecules, were tested on saposin C at pH 7.0 (chapter 6). Bicelles were prepared from a mixture of phospholipids DHPC and DMPC, but do not form stable oriented phases in the presence of lipid- and membrane-binding saposins, even if a negative surface charge is introduced by addition of SDS. Polyacrylamide gels were efficiently strained with a special device, but spectra of saposin C, oriented this way, were of rather poor quality. Pf1 filamentous phage turned out to be the only medium readily applicable to saposins, and RDC data of saposin A, C and D at pH 7.0 was obtained using the IPAP-[^{1}H,^{15}N]-HSQC experiment. All proteins display the dipolar waves typical for helical proteins, and five helices with different orientations can be identified. In saposin C, the variance with recently published RDC data may be due to the presence of a hexa-histidine tag. Fitting of couplings to different structure models confirmed that saposin A and D share the monomeric fold of saposin C, with structural variations being comparable to the variations between the known structures of saposin C and NK-lysin.

Saposin D exists in a slow dynamic conformational equilibrium, which is reflected in a double signal set in the ^{15}N-HSQC spectra at pH 7 and pH 4 (chapter 7). Exchange between both sites A and B lead to additional crowding by exchange-relayed signals in TOCSY and

NOESY spectra, thus hampering ^{15}N-HSQC assignment and the extraction of structural data. A close similarity of the chemical shifts, NOE, *J*-coupling and RDC information gained so far supports nearly identical structures with a local perturbation by either a proline cis-trans or disulfide bond isomerisation. However, a remarkable amide chemical shift separation for some residues allowed the equilibrium to be studied at various temperatures using heteronuclear longitudinal (N_z) exchange spectroscopy. At higher temperatures, also exchange of heteronuclear transverse (N_x) magnetisation under CPMG mixing was observed and exploited in the determination of forward k_a and backward k_b exchange rates. From their temperature dependance, both positive equilibrium enthalpy ΔH_{ab} = 1.77 kcal/mol and entropy ΔS_{ab} = 4.57 cal/molK were calculated, indicating a partial compensation of enthalpy and entropy. A relatively large difference in the dipolar couplings for residue Cys36 persisted up to 72 °C and is not averaged by the conformational exchange rate. In a second residue, Lys45, different amide exchange rates indicates a different strength of H-bonding and solvent protection in both forms.

9 Tables and pulse programs

9.1 Chemical shifts in saposin C at pH 7 and 37 °C

residue	HN	N	HA	HBd	HBu	HG	HD	HE(HZ)
V3	8.07	120.4	3.88		1.95	0.85/0.72		
Y4	7.92	118.6	4.11	3.23	2.94		7.10	6.93
C5	7.79	118.0	4.29	3.92	3.02			
E6	8.06	119.2	4.03	2.21	2.09	2.32		
V7	8.53	116.5	3.82		2.13	1.03/1.03		
C8	8.36	120.6	3.93	3.53	3.15			
E9	8.80	119.3	4.11	2.10	2.04	2.38/2.53		
F10	7.89	121.0	4.38	3.40	3.37		7.27	7.18(7.31)
L11	8.51	120.4	3.79	2.20	1.32	2.19	0.92/0.83	
V12	8.81	118.5	3.39		2.23	1.05/1.22		
K13	7.72	120.8	4.01	2.01	1.96	1.63/1.44	1.75/1.73	3.02
E14	8.05	119.0	3.82	1.96	2.00	1.72		
V15	8.95	120.0	3.45		2.03	0.96/0.95		
T16	8.46	114.9	3.83		4.29	1.32		
K17	7.33	119.9	4.08	1.89	1.89	1.42/1.63	1.66/1.71	2.89
L18	7.32	118.3	4.03	2.00	1.15	1.90	0.80/0.76	
I19	8.42	122.1	3.66		1.92	1.77/0.95	0.77	
D20	8.59	122.2	4.48	2.82	2.69			
N21	7.46	116.4	4.88	3.01	2.69		6.83/7.51	ND:112.4
N22	8.22	113.2	4.49	3.10	2.89		6.89/7.52	ND:111.7
K23	7.65	117.0	4.54	1.71	1.51	1.54/1.45	1.66	2.91
T24	8.93	112.0	4.37		4.71	1.38		
E25	8.85	119.9	3.78		2.09	2.25/2.17		
K26	8.11	116.4	3.94	1.89	1.75	1.49/1.44	1.64	3.03/3.15
E27	7.48	117.5	4.02	2.40	1.90	2.26		
I28	8.34	121.9	3.39		1.77	0.63	0.59	
L29	8.32	119.1	3.89	1.88	1.53	1.79	0.93/0.90	
D30	7.96	118.1	4.50	2.75	2.68			
A31	7.77	122.2	4.30		1.48			
F32	7.93	119.6	4.64	3.39	3.33		7.49	7.05(6.78)
D33	7.98	116.7	4.57	2.88	2.80			
K34	7.62	116.2	4.66	2.09	1.85	1.54/1.51	1.76/1.70	3.04/3.10
M35	7.80	121.1	4.16	2.40	2.00	2.64/2.74		1.69
C36	8.68	113.4	4.60	3.06	2.99			
S37	8.00	115.4	4.40	4.07	4.07			
K38	7.59	119.0	4.36	2.15	2.09	1.55	1.74/1.83	3.06
L39	7.16	119.9	4.37	1.57	1.24	0.88	0.62/0.19	
P40								
K41								
S42								
L43	7.66	120.8	4.84	1.62	1.46	1.53	0.79/0.77	
S44	7.75	115.4	3.95	3.99	3.99			
E45	8.74	122.9	4.10	2.04	2.04	2.29/2.41		
E46	9.33	121.9	4.18	1.99	1.46	2.21		
C47	8.53	116.9	4.17	3.35	2.85			
Q48	8.16	120.2	3.87		2.23	2.56/2.45	NE:114.8	7.80/6.85
E49	7.97	118.7	4.24	2.42	2.28	2.53/2.33		
V50	8.23	120.3	3.85		2.23	0.80/1.10		
V51	8.91	122.0	3.03		2.15	0.83/1.04		

D52	8.66	120.1	4.38	2.86	2.69			
T53	8.03	113.0	4.02		3.78	0.49		
Y54	8.93	116.8	4.79	3.21	2.80		7.23	6.91
G55	8.33	112.3	3.25	HA:3.57				
S56	8.74	114.5	4.21	4.04	3.95			
S57	7.79	120.3	4.50	4.22	4.04			
I58	8.60	122.4	3.41		2.14	1.77/0.90	0.71	
L59	7.90	116.6	4.11	1.55	2.20	1.73	1.00/0.97	
S60	8.20	112.5	4.09	4.04	4.03			
I61	8.27	120.9	3.80		1.81	1.02/0.90	0.83	
L62	7.91	119.1	4.14	2.05	1.33	1.95	0.94/0.92	
L63	8.54	120.9	4.27	2.02	1.59	1.88	0.88/0.87	
E64	7.56	119.2	4.40	2.35	2.28	2.56/2.35		
E65	7.98	111.7	3.94	2.35	2.28	2.18/2.16		
V66	6.88	120.6	3.82		1.67	0.92/0.91		
S67	8.32	122.3	4.52	3.96	3.88			
P68								
E69	8.79	114.4	4.23		2.00	2.17/2.26		
L70	7.98	118.2	4.41	1.77	1.53	1.52	0.96/0.81	
V71	7.10	118.6	3.35		2.11	0.82/1.05		
C72	9.22	115.4	4.06	3.00	2.82			
S73	8.33	117.7	4.60	3.85	3.76			
M74	8.28	125.5	3.99	2.27	2.15	2.54/2.66		2.07
L75	7.28	114.7	4.25	1.75	1.47	1.74	0.90/0.70	
H76	7.68	111.3	4.27	3.51	3.44			Har: 7.19
L77	7.87	116.7	4.39	1.79	1.32	1.20	0.57/0.32	
C78	6.91	112.6	4.90	3.17	2.18			

9.2 Chemical shifts in saposin C at pH 4 and 37 °C

residue	pH4 N	pH4 HN	pH4 HA[a]	pH4d N [b]	pH4dHN	pH4dHA[c]	pH7 HA	HAsec [d]
D2	120.6	8.35	4.27					0.00
V3	125.0	8.20	4.27				3.88	-0.07
F4	119.5	7.97	4.54				4.11	-0.49
C5			4.02				4.29	-0.36
E6			4.41				4.03	-0.26
V7			3.88				3.82	-0.13
C8			3.94				3.93	-0.72
E9	118.8	8.68	4.06				4.11	-0.18
F10	121.1	7.76	4.41				4.38	-0.28
L11	120.4	8.31	3.82	119.8	8.41	3.58	3.79	-0.38
V12	118.4	8.60	3.40	117.1	8.47	3.37	3.39	-0.56
K13	121.5	7.63	4.01	118.9	8.20	3.96	4.01	-0.35
E14	120.7	8.17	3.83	117.9	8.35	3.98	3.82	-0.47
V15	119.8	8.80	3.41	119.9	8.58	3.57	3.45	-0.50
T16	115.5	8.20	3.80	111.6	7.90	3.76	3.83	-0.52
K17	120.0	7.19	4.01	119.4	7.42	4.09	4.08	-0.28
L18	118.1	7.18	4.01	119.0	7.98	4.02	4.03	-0.14
I19	122.1	8.40	3.66	114.8	7.83	3.85	3.66	-0.29
D20			4.51	116.0	7.59	4.68	4.48	-0.28
N21	117.4	7.37	4.92	117.9	7.79	4.74	4.88	0.13
N22	112.7	8.00	4.50	119.0	8.83	4.54	4.49	-0.26
K23	116.8	7.49	4.56	120.2	8.48	4.04	4.54	0.18

T24	111.4	8.37	4.34	116.9	8.27	3.92	4.37	0.02
E25	119.8	8.69	4.73	119.6	8.49	3.79	3.78	-0.51
K26	117.1	8.12	3.90	117.9	7.98	3.87	3.94	-0.42
E27	117.2	7.35	4.00	117.7	7.87	4.02	4.02	-0.27
I28	121.8	8.18	3.37	120.2	8.22	3.39	3.39	-0.56
L29	119.3	8.18	3.86	119.7	8.43	3.82	3.89	-0.28
D30	117.2	7.84	4.49	117.7	8.36	4.41	4.5	-0.26
A31	122.5	7.61	4.34	122.4	7.92	4.26	4.3	-0.05
F32	119.9	7.74	4.65	119.3	8.94	4.54	4.64	-0.02
D33	115.8	8.07	4.60	116.9	8.52	4.42	4.57	-0.19
K34	117.1	7.58	4.62	118.7	7.96	4.15	4.66	0.30
M35	121.0	7.67	4.17	120.1	8.50	4.14	4.16	-0.36
C36	113.6	8.44	4.61	114.5	8.58	4.56	4.6	-0.05
S37	115.5	7.83	4.38	113.7	7.95	4.31	4.4	-0.10
K38	119.0	7.43	4.36	119.2	7.68	4.34	4.36	-0.06
L39	119.7	6.98	4.39	119.7	7.26	4.36	4.37	0.20
P40								0.00
K41	123.6	8.60	4.44	122.3	8.72	4.05		0.00
S42	113.0	8.16	4.22	112.6	8.48	4.17		-0.22
L43	120.9	7.55	4.70	120.7	7.83	4.46	4.84	0.67
S44	116.1	7.77	3.96	115.9	8.48	3.91	3.95	-0.55
E45	121.9	8.48	4.17	120.7	8.41	4.11	4.1	-0.19
E46	119.1	8.10	4.16	118.4	8.05	4.13	4.18	-0.11
C47			3.93	117.8	8.60	4.00	4.17	-0.48
Q48			3.86	119.2	8.45	3.81	3.87	-0.50
E49	118.1	7.81	4.21	118.4	8.19	4.13	4.24	-0.05
V50	120.4	7.99	3.84	120.5	8.34	3.73	3.85	-0.10
V51	122.2	8.83	3.03	121.3	8.86	3.11	3.03	-0.11
D52	119.3	8.47	4.41	118.4	8.87	4.43	4.38	-0.38
T53	113.3	7.74	4.05	112.9	8.07	4.01	4.02	-0.33
Y54	116.9	8.77	4.81	116.7	8.89	4.77	4.79	0.19
G55	112.2	8.14	3.26	110.8	8.55		3.25	0.00
S56	114.8	8.65	4.22	114.0	8.94	4.19	4.21	-0.29
S57	120.4	7.68	4.47	119.0	8.16	4.47	4.5	0.00
I58	122.3	8.47	3.37	121.4	8.38	3.38	3.41	-0.54
L59			4.08	116.9	8.06	4.43	4.11	-0.06
S60	112.9	8.10	4.06	111.8	8.07	4.08	4.09	-0.41
I61	120.7	8.17	3.76	119.0	8.09	3.94	3.8	-0.15
L62	119.2	7.80	4.08	121.1	8.13	3.77	4.14	-0.03
L63	120.9	8.40	4.26	114.5	8.16	4.12	4.27	0.10
E64	118.1	7.39	4.41	118.6	8.05	4.50	4.4	0.11
E65	110.9	7.84	4.04	114.9	7.91	4.25	3.94	-0.35
V66	120.2	6.65	3.81	120.0	7.35	3.79	3.82	-0.13
S67	121.8	8.22	4.52	121.7	8.59	5.01	4.52	0.02
P68			4.00				4.05	-0.39
E69	113.5	8.18	4.22	115.3	8.42	3.77	4.23	-0.06
L70	117.9	7.82	4.42	121.9	7.85	4.00	4.41	0.24
V71	118.4	6.91	3.37	119.2	8.14		3.35	-0.60
C72	114.8	8.89	4.06	113.7	8.16	4.04	4.06	-0.59
S73	118.0	8.14	4.63	119.4	8.71	4.57	4.6	0.10
M74	125.8	8.17	4.00	125.2	8.41	4.05	3.99	-0.53
L75	114.3	7.09	4.26	115.5	7.12	4.13	4.25	0.08
H76	111.1	7.56	4.32	110.2	7.92	4.24	4.27	-0.36
L77	116.1	7.77	4.38	115.7	7.97	4.37	4.39	0.22
C78	112.7	6.72	4.94	112.6	7.02	4.83	4.9	0.25

S79	117.7	8.89	117.1	8.96
G80	111.0	8.39	111.0	8.39
R81	119.9	8.07	119.9	8.07
H82	118.6	8.36	118.6	8.36
H83	119.4	8.39	119.4	8.39
H84	120.0	8.55	120.0	8.55
H85	120.6	8.60	120.6	8.60
H86	120.6	8.48	120.6	8.48
H87	125.5	8.31	125.5	8.31

[a] measured at 17 °C [b] d = dimeric species at pH 4 [c] measured at 57 °C [d] for pH 7

9.3 Scalar and residual dipolar couplings in saposin C at pH 7 and 37 °C

residue	JHNHA	Dexp	Dback [a]	Dlit-Pf1 [b]	Dlit-pd [c]	PhiX [d]	ThetaZ [d]
D2			1.7793	-7.35		1.0633	1.0864
V3	4.42		25.2926	-22.657		0.1145	1.4614
F4	5.52		0.6718	-3.422		0.8789	0.9255
C5	5.17	17.7	13.9001	-19.158	-13.116	0.6785	1.2467
E6	3.34	24	22.1568	-26.778	-21.024	0.3429	1.3668
V7	5.11	10.7	11.4683	-16.916		0.5448	1.0421
C8	3.80	9.5	8.9311	-11.98	-6.516	0.7337	1.0923
E9	3.75	16.4	20.215	-22.072	-15.445	0.5349	1.5229
F10	4.64	21.5	19.1746	-28.033	-20.288	0.3575	1.218
L11	4.06	7.6	8.4836	-12.925	-8.296	0.6269	0.9987
V12	4.19	12	11.7077	-15.1	-11.772	0.6892	1.1618
K13	5.17	22.1	21.1164	-26.047		0.428	1.3943
E14		16.4	16.4852	-21.806		0.4448	1.1578
V15	5.21	6.3	5.606	-9.833	-6.191	0.8006	1.0284
T16	3.90	17.7	14.5277	-20.926	-14.387	0.7251	1.3409
K17	4.10	24	23.7984	-29.668	-21.455	0.3146	1.5102
L18	5.81	8.8	10.1466	-14.605	-10.049	0.6336	1.0587
I19	4.96	12.6	12.4149	-17.955		0.8338	1.3439
D20	4.17	18.9	22.0103	-24.055	-17.193	0.4173	1.4772
N21	9.24		17.9829	-25.281	-19.175	0.3952	1.1876
N22	7.61	-8.8	-8.4147	7.083	5.986	1.2319	0.8439
K23	7.13	4.4	3.5171	-10.683	-5.65	1.2061	1.2592
T24	7.70	-19.6	-19.4107	24.742	17.912	1.4646	0.5836
E25	2.49	10.7	9.6885	-3.895	-4.056	0.6754	1.0736
K26	4.33	20.2	20.5992	-20.154	-15.166	0.4046	1.3211
E27	5.49	18.3	18.1944	-20.312	-15.331	0.4158	1.2115
I28	5.57	5	5.0641	3.675		0.6771	0.9244
L29	3.72	15.8	15.4045	-11.847	-8.257	0.5859	1.2277
D30	4.77	25.9	24.039	-27.528		0.2028	1.3908
A31	6.15	8.9	13.4818	-2.608	-2.512	0.5044	1.0852
F32	2.95	2.5	3.1313	6.452	5.589	0.8165	0.9597
D33	4.49	13.9	22.2129	-17.315	-11.581	0.361	1.4052
K34	8.93	15.8	15.2751	-15.084	-11.334	0.4613	1.1234
M35	4.45	-10.1	-14.1513	20.195	15.139	1.5498	0.7501
C36	4.73	-0.6	-2.1946	1.143	0.751	1.5417	1.1317
S37	5.72	11.4	14.7304	-11.95	-7.988	0.663	1.3039
K38	7.47	-14.5	-29.218	29.164		1.5476	0.1357
L39	8.23	1.3	4.1156	-7.977	-5.111	1.2974	1.3676

P40			-24.0911	40.551		1.3279	0.3736
K41			-10.3522	12.392		1.1569	0.7574
S42	9.66		-18.1419	24.033		1.13	0.4904
L43	2.50	-10.7	-11.4913	26.529	19.016	1.0474	0.6648
S44	5.37	7.6	6.5152	-0.094	0.419	0.6857	0.9749
E45	4.55	-5.7	-4.6793	14.282	11.184	1.0903	0.8923
E46	3.83	-12	-12.2165	26.114		0.9751	0.5968
C47	4.08	3.1	0.9615	10.039		0.8098	0.8881
Q48	4.34	3.2	3.4534	2.726		0.9387	1.0572
E49	5.20	-13.3	-13.3314	28.218		1.1248	0.6546
V50	4.06	-5.1	-5.8304	19.717	13.312	0.9309	0.7642
V51	3.93	6.9	5.1066	3.053	1.646	0.7657	0.9869
D52	8.53	0.6	-1.6788	8.853	8.415	1.0986	0.9909
T53	10.41		-12.8177	28.194	21.819	1.0521	0.6268
Y54	5.40	17.7	14.3128	-23.614	-18.088	0.7611	1.3745
G55	4.30	19	19.0599	-29.325	-22.694	0.5309	1.3817
S56	6.08	24	22.7545	-25.743	-18.814	0.2612	1.3429
S57		15.1	16.2098	-20.514	-16.252	0.5399	1.2247
I58	3.99	23.4	19.7485	-29.608	-21.851	0.5004	1.3829
L59	2.78	24	24.389	-31.693	-24.636	0.1745	1.4035
S60	5.41	21.5	20.3865	-21.788	-17.157	0.3223	1.2492
I61	4.95	17.7	18.2958	-23.07	-18.589	0.4658	1.258
L62	5.87	20.9	18.1267		-24.154	0.5957	1.4062
L63	8.14	24.6	23.7516	-26.263		0.2386	1.3829
E64	7.94	-8.8	-12.2632	10.108	6.441	1.5039	0.8072
E65	6.44	6.3	4.9086	-16.288	-12.277	1.1461	1.285
V66	2.69	-19.6	-22.2346	27.612		1.4247	0.4797
S67	4.58		20.6356	-29.531	-21.863	0.495	1.4682
P68			17.2967	-21.882		0.4323	1.1812
E69	3.32	-10.1	-10.7454	12.625	8.588	0.9899	0.6519
L70	4.45		-8.3377	9.849	5.744	1.0115	0.7378
V71	6.93	12	12.2239	-17.801	-14.228	0.5141	1.0653
C72	4.28	-6.9	-7.6684	8.965	6.002	0.8917	0.6842
S73	7.14	-17.7	-16.0039	21.498	15.969	1.1039	0.5551
M74	6.87	17	19.8572	-28.883	-24.127	0.4891	1.3795
L75			25.3532	-31.487	-22.638	0.1811	1.4705
H76	9.78	1.9	0.7904	9.701		0.772	0.8577
S79			14.3343	15.488	11.319	0.6483	1.2394
G80			-21.1714	-2.928	-2.088	1.2891	0.4703

[a] from structure 1M12 [144] [b] published couplings Pf1 [144] [c] published couplings pd [144]

[d] 1M12 in the PAS of the tensor given by Table 6.6.1C.

9.4 ^{15}N relaxation and hydrogen exchange rates in saposin C at pH 7 and 37 °C

residue	R2	R2err	R1	R1err	hetNOE	hetNerr	HX
V3	8.33	0.11	2.05	0.09	0.73	0.04	2.12
Y4	8.17	0.12	1.90	0.02	0.77	0.04	0.80
C5					0.80	0.04	<0.5
V7					0.79	0.04	0.63
C8					0.83	0.04	<0.5
E9					0.76	0.04	<0.5

F10	8.02	0.05	1.82	0.01	0.75	0.04	0.52
L11	7.64	0.07	1.80	0.01	0.60	0.03	1.28
V12	8.73	0.10	1.86	0.02	0.80	0.04	<0.5
K13	8.05	0.07	1.82	0.01	0.82	0.04	<0.5
E14	8.51	0.06	1.84	0.02	0.74	0.04	0.53
V15	8.17	0.08	1.86	0.02	0.80	0.04	<0.5
T16	8.16	0.08	1.86	0.03	0.79	0.04	<0.5
K17	8.14	0.07	1.85	0.02	0.76	0.04	<0.5
L18	8.12	0.09	1.79	0.01	0.73	0.04	<0.5
I19	8.51	0.05	1.83	0.02	0.73	0.04	0.50
D20	8.10	0.04	1.86	0.01	0.74	0.04	<0.5
N21	7.04	0.05	1.79	0.02	0.75	0.04	<0.5
N22	8.04	0.08	1.83	0.01	0.78	0.04	0.72
K23	7.79	0.08	1.79	0.01	0.77	0.04	<0.5
T24	7.18	0.10	1.85	0.02	0.72	0.04	0.60
E25	7.85	0.06	1.97	0.04	0.70	0.04	1.62
K26	8.01	0.03	1.87	0.01	0.79	0.04	<0.5
E27	7.78	0.04	1.87	0.02	0.78	0.04	0.59
I28	7.93	0.04	1.89	0.02	0.78	0.04	<0.5
L29	8.13	0.07	1.91	0.02	0.77	0.04	<0.5
D30	8.16	0.07	1.84	0.01	0.75	0.04	0.70
A31	6.99	0.10	1.85	0.02	0.70	0.04	<0.5
F32	7.48	0.07	1.76	0.01	0.70	0.03	0.42
D33	7.54	0.05	1.85	0.01	0.63	0.03	1.05
K34	6.51	0.04	1.80	0.04	0.53	0.03	1.34
M35	7.12	0.06	1.78	0.01	0.71	0.04	<0.5
C36	7.19	0.05	1.94	0.02	0.75	0.04	1.08
S37	8.76	0.08	2.16	0.07	0.82	0.04	2.14
K38	7.60	0.07	2.09	0.03	0.70	0.03	0.91
L39	7.42	0.07	1.77	0.01	0.66	0.03	<0.5
L43	7.14	0.05	1.90	0.03	0.70	0.04	0.68
S44	7.61	0.06	1.86	0.02	0.80	0.04	<0.5
E45	7.79	0.06	1.98	0.02	0.84	0.04	1.33
E46	7.58	0.07	1.85	0.02	0.76	0.04	<0.5
C47	7.98	0.06	1.93	0.01	0.73	0.04	0.42
Q48	7.69	0.04	1.88	0.02	0.71	0.04	<0.5
E49	8.03	0.04	1.86	0.01	0.76	0.04	0.65
V50	7.75	0.06	1.93	0.01	0.73	0.04	<0.5
V51	7.69	0.09	1.87	0.01	0.73	0.04	<0.5
D52	8.22	0.04	1.93	0.01	0.76	0.04	<0.5
T53	7.27	0.06	1.84	0.02	0.78	0.04	0.96
Y54	8.04	0.08	1.91	0.02	0.73	0.04	<0.5
G55	7.98	0.06	1.98	0.01	0.79	0.04	<0.5
S56	8.14	0.06	2.13	0.07	0.82	0.04	2.05
S57	7.99	0.06	2.00	0.04	0.78	0.04	1.28
I58	8.10	0.08	1.95	0.02	0.77	0.04	0.50
L59	8.30	0.05	1.93	0.01	0.74	0.04	<0.5
S60	7.40	0.07	1.89	0.01	0.73	0.04	<0.5
I61	8.03	0.09	1.89	0.02	0.77	0.04	<0.5
L62	8.47	0.05	1.95	0.01	0.71	0.04	<0.5
L63	8.46	0.08	1.95	0.02	0.75	0.04	<0.5
E64	7.82	0.05	1.75	0.01	0.71	0.04	<0.5
E65	7.83	0.08	1.84	0.01	0.78	0.04	<0.5
V66	7.28	0.07	1.79	0.01	0.73	0.04	0.53
S67	8.10	0.05	1.82	0.01	0.78	0.04	<0.5

E69	7.91	0.06	1.96	0.02	0.72	0.04	0.99
V71	7.78	0.05	1.88	0.01	0.82	0.04	<0.5
C72	7.73	0.09	2.02	0.02	0.81	0.04	0.73
S73	8.22	0.10	2.18	0.06	0.80	0.04	1.66
M74	8.25	0.10	1.94	0.02	0.74	0.04	0.64
L75	8.62	0.10	1.95	0.03	0.76	0.04	0.56
H76	8.32	0.13	2.05	0.02	0.80	0.04	0.73
C78	7.46	0.08	1.79	0.02	0.77	0.04	0.78

9.5 Chemical shifts and residual dipolar couplings in saposin A at pH 7 and 37 °C

residue	**N**	**HN**	**HA**	**HAsec**	**HB**	**furtherH**	**remark** [a]	**RDC**
C5	116.02	7.60	4.20	-0.34	3.04		HX	
D6	120.04	8.06	4.26	-0.43	2.75		HX	10.58
I7	118.65	8.27	3.86	-0.31	1.74	0.85	HX	12.21
C8	118.96	8.42	3.82	-0.72	3.50/3.00			20.59
K9	120.12	8.54	3.72	-0.61	1.91			23.24
D10	122.05	7.78	4.49	-0.22	2.94/2.71		broad	9.11
V11	122.83	8.56	3.70	-0.42	2.34	1.08/0.87		16.32
V12	119.19	8.80	3.43	-0.69	2.11	0.93		21.76
T13	117.61	8.24	3.90	-0.45	3.92	1.29	broad	16.33
A14	123.10	7.93	4.28	-0.05	1.60			12.95
A15	119.12	8.73	3.97	-0.38	1.38			21.17
G16	123.60	8.36	4.11/3.76	-0.20			broad	
D17	123.14	8.15	4.44	-0.27	2.83/2.63			14.27
M18	118.58	7.98	4.15	-0.33	1.98/1.89	2.70/2.32		16.77
L19	121.59	8.47	4.12	-0.20	1.90/1.66	0.90/0.75		20
K20	120.31	7.75	4.14	-0.19	1.99	1.72/1.56		13.97
D21	118.42	7.18	4.89	0.18	2.95/2.55			7.22
N22	113.47	8.05	4.41	-0.33	3.09/2.94			10.01
A23	120.78	7.62	4.45	0.12	1.29			3.53
T24	110.96	8.50	4.42	0.07	4.21	1.38	HX	-3.09
E25	120.51	9.04	3.75	-0.58	2.03	2.27	HX	1.47
E26	117.92	8.53	4.02	-0.31	2.06/1.99	2.33	HX	10.44
E27	119.12	7.59	3.95	-0.38	2.40	1.92	HX	-0.29
I28	121.32	8.30	3.74	-0.43	1.98	1.02/0.75		-0.73
L29	120.55	8.11	3.89	-0.43	1.89/1.64	1.00		6.91
V30	116.37	7.74	3.75	-0.37	2.10	1.12/1.02		12.21
Y31	120.16	7.91	4.15	-0.40	3.20/3.00	7.02		-0.74
L32	119.81	8.90	3.93	-0.39	2.09	1.35/1.00		1.17
E33	118.73	8.50	3.67	-0.66	2.14/1.97	2.54		12.35
K34	118.30	7.32	3.87	-0.46	1.69/1.44	1.22/2.89		8.52
T35	112.47	7.64	4.00	-0.35	3.26	0.83		3.97
C36	118.46	7.79	4.58	0.04	3.35/2.93			5.88
D37	120.31	7.65	4.55	-0.16	2.67			15.29
W38	118.15	7.57	4.80	0.14	3.45/3.29	9.97/7.16		-1.91
L39	121.17	7.18	4.57	0.35	1.84	1.47/1.00		-8.09
M43	120.28	7.56	4.17	-0.31	2.25	2.66/2.54	HX	-5.01
S44	115.48	7.97	4.22	-0.25	4.01		HX	-3.09
A45	122.29	8.21	4.15	-0.18	1.52		HX	0.44
S46	114.98	7.71	4.32	-0.15	3.98/3.84		HX	-2.95
C47	120.31	8.28	4.01	-0.53	3.53/2.60		HX	-6.03
K48	119.85	8.35	3.79	-0.44	1.88	1.58		0.44

E49	118.88	7.84	4.19	-0.14	2.25	2.42/1.90		3.98
I50	121.01	8.13	3.96	-0.21	2.15	1.16/0.83		-5.3
V51	120.66	8.56	3.41	-0.71	2.16	1.03		-3.24
D52	117.69	8.65	4.37	-0.34	2.81/2.68			1.77
S53	110.42	8.14	4.30	-0.17	3.62/3.11			-1.33
Y54	114.44	8.39	4.97	0.42	3.17/2.98			-6.18
L55	121.71	8.28	4.27	-0.05	2.18/1.89	1.00/0.86		-11.76
P56			4.22	-0.22	2.12	3.79/3.42		
V57	115.60	6.93	4.02	-0.10	2.30	1.17		-0.6
I58	121.20	8.43	3.43	-0.74	2.12	0.84/0.65		-10
L59	116.33	8.48	3.98	-0.34	1.96	1.38/0.87		-12.35
D60	117.34	7.37	4.44	-0.27	2.81			
I61	119.97	8.23	3.80	-0.37	1.94	1.01/0.84		-5.29
I62	115.44	8.45	4.20	0.03	2.06	1.49/0.97		-11.48
K63	119.62	7.70	4.15	-0.18	1.90	1.72/1.50		-9.41
G64	127.16	7.84	4.25/3.85	-0.11			HX	
E65	119.00	8.18	4.30	-0.03	2.24/2.07		HX	-7.94
M66	120.62	8.27	4.52	0.04	2.00/2.90	2.50	HX	-6.61
S67	115.17	7.91	4.52	0.05	3.75		HX	-5.59
R68	122.90	8.42	4.68	0.30	2.04/1.77	3.26	HX	
G69	123.91	9.11	3.93/3.47	-0.53			HX	
E70	121.47	7.30	4.12	-0.19	2.31/2.13		HX	7.2
V71	121.71	7.96	3.58	-0.54	2.14	0.89	HX	-3.66
C72	114.44	8.17	4.04	-0.50	2.94		HX	-10.15
S73	118.27	8.19	4.58	0.11	3.89		HX	-4.71
A74	127.81	7.93	4.19	-0.14	1.56			2.94
L75	114.83	7.16	4.25	-0.07	1.83/1.50	0.78/0.63		-7.94
N76	113.05	8.02	4.47	-0.27	3.15/2.92			-11.33
L77	115.87	7.78	4.40	0.08	1.74/1.32	0.72/0.47		8.97
C78	112.82	6.87	4.82	0.28	3.31/2.37			-0.74

[a] HX = water signal in HNH-NOESY

9.6 Chemical shifts and residual dipolar couplings in saposin D at pH 7 and 37 °C

residue	Na	HNa	HAa	RDCa	Nb	HNb	HNb	RDCb
G3	108.12	8.30		-1.14				
F4	118.61	9.01		1.16				
C5	120.22	8.34		13.15				
E6	117.46	8.42		14.99				
V7	117.79	8.42	3.56	13.38				
C8	120.94	7.97	3.39	13.30	120.93	7.93		
K9	119.72	8.09	3.77	13.15	119.67	8.05		13.01
K10	120.12	8.10	4.08	15.50	120.28	8.08		13.38
L11	122.03	8.77	4.25	13.61	122.02	8.74		12.39
V12	118.47	9.13	3.67	12.40	118.33	9.10		11.55
G13	107.12	8.09	3.92	15.43				
Y14	122.97	7.75	4.11	13.68				
L15	122.01	8.07	3.75	15.21	122.07	8.09		
D16	117.71	8.73	4.36	14.21	117.75	8.68		13.46
R17	112.36	7.37	4.20	17.25				
N18	115.24	7.49	4.70	12.17				
L19	122.39	8.17	4.57	2.43	122.65	8.28		1.07

E20	121.12	8.28		-9.20				
K21	122.39	8.68		-1.74				
N22	116.07	8.51		4.19	116.18	8.58		3.99
S23	116.43	7.84		-8.75	116.27	7.89		-5.05
T24	113.19	8.83	4.47	6.69	112.91	8.81		3.42
K25	120.36	8.90	3.79	-0.23				
Q26	116.08	8.34	4.00	7.07			4.53	
E27	120.94	7.72	4.41	8.07				
I28	121.90	8.43	3.64	-2.44				
L29	120.49	8.50		-1.53				
A30	119.65	7.76	4.19	11.10			4.14	
A31	120.45	7.75		6.31	120.88	7.82		4.48
L32	120.46	8.77	4.06	-3.04	120.33	8.65		-4.41
E33	116.05	7.85		6.60	116.07	7.69		4.03
K34	115.62	7.35	4.60	10.26	117.33	7.58	4.34	7.67
G35	108.78	8.20	3.47	-6.53	108.50	8.46	3.35	-5.25
C36	116.76	8.57	4.51	-19.91	116.23	8.12	4.66	-14.06
S37	113.25	7.81	4.10	-8.06	113.37	7.72	4.17	-1.22
F38	118.90	7.51	4.45	-3.05	119.13	7.34	4.55	-1.22
L39	121.29	7.31	4.59	-16.50	121.23	7.35	4.59	-16.19
Q44	119.92	7.65	3.72	-5.48			3.90	
K45	119.32	8.53	4.24	-10.49	117.99	8.20	4.24	-9.13
Q46	119.29	8.62	4.44	-15.43	117.62	7.43	4.48	-12.54
C47	125.59	8.00		-1.37				
D48	121.05	8.53		-1.68				
Q49	119.81	7.80		-11.94				
F50	121.71	8.52		-10.04	121.60	8.33		-9.43
V51	119.47	8.82		-3.12	119.13	8.81		-3.72
A52	119.91	7.72		-3.43				
E53	115.13	7.70		-15.43	115.22	7.67		-12.54
Y54	113.33	8.17		-3.50	113.43	8.08		
E55	123.55	8.36		10.63				
V57	112.71	7.07	4.28	9.96	112.59	7.07	4.28	8.36
L58	121.66	8.48	4.05	7.98	121.74	8.45	4.05	6.00
I59	115.66	8.20	3.84	4.64	115.44	8.21	3.84	4.26
E60	115.01	6.78	4.20	14.30	115.20	6.83	4.20	13.46
I61	116.82	7.59	4.41	12.32	120.47	8.42		
L62	120.38	8.50	4.81	5.69	109.33	6.90	4.81	4.34
V63	108.82	6.87	3.92	1.29			3.92	1.52
E64	118.90	7.91	4.53	15.12	118.62	7.88	4.53	14.37
V65	119.56	8.37	4.30	1.21				
M66	120.24	8.21		6.77	120.37	8.22		6.32
D67	117.52	6.33		0.00				
S69	111.63	7.98	4.70	5.77				
F70	121.15	7.90	4.24	8.06				
V71	118.11	8.86	3.30	0.69				
C72	113.61	7.89	4.21	-1.52				
L73	122.04	7.45	4.37	9.13				
K74	122.50	8.25	4.45	6.16				
I75	108.25	7.90	4.47	-4.70				
G76	107.40	7.57	4.38	-0.45				
A77	121.54	7.53	4.41	15.35				
C78	112.67	7.34	4.21	3.50				

9.7 Pulse program for the Nz-exchange experiment

```
;mjT1ex.2D                    mj/17/06/2003

;avance-version

;#########################  NOTE  ########################
;2D 15N T1 relaxation and exchange experiments with the
;following features:
;- N is flipped to +z and -z
;- suppression of CSA[N]/DD[NH] cross-correlated relaxation
;- H2O flip-back
;- STATES quadrature detection & WATERGATE
;- good water suppression on a cryoprobe
;- all dimensions decoupled
;#######################################################

;N.A.Farrow, O.Zhang, J.D.Forman-Kay and L.E.Kay(1994)J.Biomol.NMR 4,
;727-734.
define list<gradient> EA=<EA>

;#########################Processing Information ########################
;MC2: STATES
;REVERSE = TRUE in F1
;ATTENTION!!! MAKE SURE TO USE THE 'xfb nc_proc x' COMMAND TO OBTAIN
;EQUAL SCALING FOR ALL SUCCESSIVE SUB-SPECTRA:
;x = NC scaling factor from 'dpp' after processing reference spectrum
;T1mix = d25*l10 (1-120)

;############# Setting Parameters: #################
;*) NS = n*4
;*) DS = any value, but multiples of TD1 recommended!
;*) SFO1: H2O MUST be on-resonant (for water flip-back)
;*) Set nd0 = 2
;*) CPDPRG2 = garp

;############# ASED parameters #################
;in0 = 1/[2*SWH(15N)]
;p1 = 90deg high power on 1H (F1, pl1)
;p3 = 90deg high power on 15N (F2, pl2)
;p16 >=800u gradient pulse
;p17 = 1.8m gradient pulse (gp0)
;pcpd2 >=200u 15N decoupler (pl12,F2)
;d1 >2s long relax.delay (heat equilibration!)
;d4 =< 2.7m (=1/[4J(NH)])
;d14 =< 2.7m (=2.7m for NH2 suppression!)
;d16 >= 150u (gradient recovery)
;d23 >= 1.5m (echo for CSA/DD decoupling)

;############# fixed parameters, do not change ####
"p2=p1*2"
"p4=p3*2"
define delay cen31
define delay cen42
"cen31=(p3-p1)/2"       ;pulse center delay
"cen42=(p4-p2)/2"       ;pulse center delay
"d0=3u"
"d11=10m"               ;delay for phase and buffer incrementation
"d14=2.7m"
"d24=d23-p16-d16"
"d25=d23*4+p2*2"        ;calculated time for 1 loop in T1
define delay d14a       ;compensated d14 J-evolution delays
```

```
define delay d14b
define delay d14c
define delay d4grad
define delay d14grad
"d4grad=d4-p17-d16"
"d14a=d14+p3*0.64-p1-p17-d16"
"d14b=d14+d0+p1"
"d14c=d14-d0-p1-p3*0.64"
"d14grad=d14-p17-d16"
define delay d16c       ;compensated gradient duration
"d16c=p16+d16"
define delay wg
"wg=p1*4.77+d19*10"
define delay cen14
"cen14=(wg-p4*2-6us)/2"
"l10=cnst10"

"l0=(td1/2)"            ;loop counter for quadrature detection
"l9=0"                  ;dummy scan flag

1 ze
    d25
2   d1 setnmr2^0 setnmr0^34^32^33 do:f2
    d11*2
3   d11
    d11 pl2:f2
4   50u setnmr2|0 setnmr0|34|32|33 pl2:f2

;### Destruction of natural 15N magnetisation ###
  (p3 ph0):f2
  p16:gp1*0.3           ;z-spoil
   d16

;### Start of INEPT ###
  (p1 ph0)
  p17:gp0               ;dephasing of residual H2O
    d16
    d4grad pl2:f2       ;H2O in -y
  (cen42 p2 ph1) (p4 ph5):f2
  p17:gp0               ;dephasing of residual H2O
    d16
    d4grad pl2:f2       ;H2O in -y
  (p1 ph1)              ;H2O in -y

;### Start of t1 on 15N ###
  (p3 ph5):f2
    d14b
  (p4 ph4):f2
    d0                  ;H2O in +z
  (p2 ph0):f1
    d0                  ;H2O in -z
    d14c
;### End of t1 time on 15N ###


;### Start of T1 mixing ###
   (p3 ph12):f2         ;flipping N coherence +/-z cancels
     d24                ;15N equilibrium magnetisation as T1 offset
   p16:gp2*0.2          ;z-spoil
     d16                ;H2O in +z
   (p2 ph20)            ;decoupling of CSA/DD cross-correlated relaxation
     d23 ip20           ;using MLEV expansion
15   d23 ip20           ;H2O in -z
```

```
   (p2 ph20)
     d23                ;H2O in +z
     d23
   (p2 ph20)
     d23 ip20           ;H2O in -z
   lo to 15 times l10
     d23 ip20           ;H2O in -z
   (p2 ph20)
     d24                ;H2O in +z
   p16:gp2*0.4          ;z-spoil
     d16
;###  End of T1 mixing  ###

;### Coherence transfer to 15N and J[HN] rephasing ###
  (p3 ph0):f2           ;create 2NyHz
  p17:gp0
    d16
    d14grad             ;H2O in -y
  (cen42 p2 ph1) (p4 ph0):f2  ;rephase to Nx
  p17:gp0
    d16
    d14a                ;H2O in -y
  (p3 ph9):f2
  p16:gp2*0.6           ;z-spoil
    d16
  (p1 ph0):f1
    d24
  p16:gp1
    d16
  (p1*0.231 ph14 d19*2 p1*0.692 ph14 d19*2 p1*1.462 ph14 d19*2 p1*1.462
ph15 d19*2 p1*0.692 ph15 d19*2 p1*0.231 ph15):f1 (cen14 p3 ph1 3u p4 ph0 3u
p3 ph1):f2
  p16:gp1
    d16 pl12:f2
    d24

;### Start of data acquisition ###
  go=2 ph31 cpd2:f2
    d1 do:f2 wr #0 if #0 zd
    d11 setnmr2^0 setnmr0^34^32^33
    d11 ip5
  lo to 3 times 2             ;States loop
;### Increment t1 time ###
    d11 id0                   ;shift axial peaks and H2O
    d11 ip31*2                ;to the edges of the spectrum
  lo to 4 times l0
exit

ph0=0
ph1=1
ph2=2
ph4= 0 1 2 3            ;(N) t1 phase purge 180deg pulse
ph5= 0                  ;(N) t1 excitation pulse, States phase
ph9= {{1}*8}^2          ;(N) invertable read pulse for sens.enh.
ph10={{1}*8}^2          ;(N) read pulse for sens.enh.
ph12=1 1 3 3            ;(N) for cancellation of 15N magnetisation in T1
ph14= {{0}*8}^2         ;watergate
ph15= {{2}*8}^2         ;watergate
ph20=0
ph31={0 2 2 0 0 2 2 0}^2
```

9.8 Pulse program for the Nx-exchange experiment

```
;mjT2ex                        mj/25/2003

;avance-version


;#########################  NOTE  ########################
;2D 15N T2 relaxation and exchange experiments with the
;following features:
;- CPMG spin-echo pulse train
;- suppression of CSA[N]/DD[NH] cross-correlated relaxation
;- H2O flip-back and gradient echoes during INEPT
,- good water suppression on a cryoprobe
;- STATES quadrature detection w/ WATERGATE
;- all dimensions decoupled
;- z-filter before CPMG pulse train enhances t1-phases
;#############################################################

define list<gradient> EA=<EA>

;#########################Processing Information ######################
; MC2: STATES
; REVERSE = TRUE in F1
; ATTENTION!!! MAKE SURE TO USE THE 'xfb nc_proc x' COMMAND TO OBTAIN
; EQUAL SCALING FOR ALL SUCCESSIVE SUB-SPECTRA:
; x = NC scaling factor from 'dpp' after processing reference spectrum
; T2mix = d22*l10 (1-12)

;############# Setting Parameters: #################
;*) NS = n*4
;*) DS = any value, but multiples of TD1 recommended!
;*) SFO1: H2O MUST be on-resonant (for water flip-back)
;*) Set nd0 = 2
;*) CPDPRG2 = garp

;############# ASED parameters ##################
;in0 = 1/[2*SWH(15N)]
;p1 = 90deg high power on 1H (F1, pl1)
;p11 = 2ms 90 deg GAUSS (F1, sp11)
;p12 = 2ms 180 deg GAUSS (F1, sp12=sp11-6dB)
;p3 = 90deg high power on 15N (F2, pl2)
;p16 >=800u gradient pulse
;p17 = 1.8m gradient pulse (gp0)
;pcpd2 >=200u 15N decoupler (pl12,F2)
;d1 >2s long relax.delay (heat equilibration!)
;d4 =< 2.7m (=1/[4J(NH)])
;d14 =< 2.7m (=2.7m for NH2 suppression!)
;d16 >= 150u (gradient recovery)
;d20 ~ 0.45m CPMG echo delay
;     ;NOTE: long d20 minimize sample heating but are
;     ;detrimental due to longer antiphase evolution
;     ;and emphasise conformational exchange processes

;############# fixed parameters, do not change ####
"p2=p1*2"
"p4=p3*2"
define delay cen31
define delay cen42
"cen31=(p3-p1)/2"       ;pulse center delay
"cen42=(p4-p2)/2"       ;pulse center delay
"d0=3u"
```

```
"d11=10m"               ;delay for phase and buffer incrementation
"d14=2.7m"
"d21=d20-p1"            ;corrected CPMG delay
"d22=d20*32+p4*16"      ;duration of 1 CPMG pass
"d24=d4-p16-d16-600u"   ;600u compensate for J-evolution during 3919 seq
define delay d14a       ;compensated d14 J-evolution delays
define delay d14b
define delay d14c
define delay d4grad
define delay d14grad
"d4grad=d4-p17-d16"
"d14a=d14+p3*0.64-p1-p17-d16"
"d14b=d14+d0+p1"
"d14c=d14-d0-p1-p3*0.64"
"d14grad=d14-p17-d16"
define delay d16c       ;compensated gradient duration
"d16c=p16+d16"
define delay wg
"wg=p1*4.77+d19*10"
define delay cen14
"cen14=(wg-p4*2-6us)/2"
"l10=cnst10"

"l0=(td1/2)"            ;loop counter for quadrature detection

1 ze
    d22
2   d1 setnmr2^0 setnmr0^34^32^33 do:f2
    d11*2
3   d11
    d11 pl2:f2
4   50u setnmr2|0 setnmr0|34|32|33 pl2:f2

;### Destruction of natural 15N magnetisation ###
  (p3 ph0):f2
  p16:gp1*0.3           ;z-spoil
  d16

;### Start of INEPT ###
  (p1 ph0)
  p17:gp0               ;dephasing of residual H2O
    d16
    d4grad pl2:f2       ;H2O in -y
  (cen42 p2 ph1) (p4 ph5):f2
    d4grad              ;H2O in -y
  p17:gp0               ;rephasing of residual H2O
    d16
  (p1 ph11)             ;H2O in -y

 ;### Start of t1 on 15N ###
  (p3 ph5):f2
    d14b
  (p4 ph4):f2
    d0
  (p2 ph0):f1           ;H2O in -z
    d14c                ;H2O in +z
    d0
 ;### End of t1 time on 15N ###

 ;###    15N z-filter    ###
  (p3 ph0):f2
  p16:gp2*0.7           ;z-spoil gradient
    d16
```

```
  (p3 ph1):f2

;### Start of mixing (CPMG along x, i.e. parallel to incoming N-coherence!)
10  d20
  (p4 ph0):f2
    d20
  lo to 10 times 3
    d20
  (p4 ph0):f2
    d21                 ;H2O in +z
  (p2 ph0):f1           ;decouple CSA/DD cross-correlated relaxation!!!
    d21                 ;H2O in -z
  (p4 ph0):f2
    d20
11  d20
  (p4 ph0):f2
    d20
  lo to 11 times 6
    d20
  (p4 ph0):f2
    d21                 ;H2O in -z
  (p2 ph2):f1           ;decouple CSA/DD cross-correlated relaxation!!!
    d21                 ;H2O in +z
  (p4 ph0):f2
    d20
12  d20
  (p4 ph0):f2
    d20
  lo to 12 times 2
    d20
  (p4 ph0):f2
    d20
  lo to 10 times l10
;### End of CPMG mixing ###

;### Coherence transfer to 15N and J[HN] rephasing ###
  p17:gp0
    d16
    d14grad             ;H2O in -y
  (cen42 p2 ph1) (p4 ph0):f2
  p17:gp0
    d16
    d14a

  (p3 ph9):f2
  p16:gp2*0.4           ;H2O rephasing
    d16
  (p1 ph0):f1           ;H2O in z
    d24
  p16:gp1
    d16
 (p1*0.231 ph14 d19*2 p1*0.692 ph14 d19*2 p1*1.462 ph14 d19*2 p1*1.462 ph15
d19*2 p1*0.692 ph15 d19*2 p1*0.231 ph15):f1 (cen14 p3 ph1 3u p4 ph0 3u p3
ph1):f2
  p16:gp1
    d16 pl12:f2
    d24

;### Start of data acquisition ###
  go=2 ph31 cpd2:f2
    d1 do:f2 wr #0 if #0 zd
   d11 setnmr2^0 setnmr0^34^32^33
    d11 ip5
```

```
  lo to 3 times 2       ;States loop
;### Increment t1 time ###
    d11 id0             ;shift axial peaks and H2O
    d11 ip31*2          ;to the edges of the spectrum
  lo to 4 times l0
exit

ph0=0
ph1=1
ph2=2
ph3=3
ph4=0 1 2 3             ;(N) t1 pahse purge 180deg pulse
ph5= 1 1 3 3            ;(N) t1 excitation pulse, States phase
ph9= {{2}*4}^2          ;(N) invertable read pulse for sens.enh.
ph11={{1}*4}^2          ;(H) INEPT transfer pulse
ph14= {{0}*8}^2
ph15= {{2}*8}^2
ph31=0 2 2 0
```

9.9 Pulse program for the NewMEXICO experiment

```
;mjnewmexwg.f2               mj/25/11/2003


;#########################  NOTE  ##############################
;fast 2D N,H-HSQC
;using WATERGATE suppression scheme with water flip-back
;BEST possible water suppression,60% H2Oz recovery
;long gradient during tmix ensures water stays in –z for ~200 ms
;all dimensions decoupled
;#############################################################

define list<gradient> EA=<EA>

;### IMPORTANT ####
;H2O must be on-resonant!
;nd0 = 2 (States)

;### Parameters to set up ###
;p1 = 90deg on 1H (pl1,F1)
;pl1 = for p1=90deg
;p3 = 90deg on X (pl2,F2)
;pl2 = for p3=90deg
;pcpd2 = 90deg decoupler on X (pl12,F2)
;pl12 = for pcpd2=90deg
;p16 >= 800u (gradient pulse)
;d0 = 3u (= 5,5m for inverted NH2 groups)
;d1 ~1s (relaxation delay)
;d4 <= 1/4JHX
;d16 >= 150u (gradient recovery)
;d19 ~ DW (WATERGATE delay)
;NOTE: zero intensity reoccurs at offset=1/(2*d19)!!!

;### Calculated Parameters ###
"p2=p1*2"
"p4=p3*2"
define delay wg
define delay cen14
define delay cen214
```

```
define delay d0c
"d0c=d0*2+p2"
"cen214=p4+3u-p1"
"wg=p1*4.77+d19*10"
"cen14=(wg-p4*2-6us)/2"
"d11=10m"
"d12=d10-20u-d16"
"d17=d7-p16-d16"
"d22=(p3-p2)/2"
"d23=(p4-p2)/2"
"d24=d4-p16-d16-500u"   ;600u compensate for J-evolution during 3919

"l3=(td1/2)"

1 ze
    d12
2   d1 do:f2
    d11*2
3   d11*2


4 (p1 ph0)
    d2 pl2:f2           ;H,N-HMQC evolution
  (p3 ph10):f2          ;H,N-HMQC creation
    d17
    p16:gp1
    d16                 ;radiation damping period

  (p1 ph0)              ;create maximum quantum coherence (invisible)
    d2
  (d22 p2 ph0) (p3 ph11):f2
    d2
  (d23 p1 ph0) (p3 ph12):f2;H2O returns to +z, all other magnetization as
MQC

    10u gron1*0.1
    d12
    10u groff
    d16                 ;mixing      time      (transfer      H2O-equilibrium
magnetization)

  (p1 ph0)              ;H2O in -y
    d4 pl2:f2
  (cen214 p2 ph1) (p3 ph13 3u p4 ph3 3u p3 ph13):f2
    d4 setnmr2|0 setnmr0|34|32|33
  (p1 ph1)              ;H2O in -y

  p16:gp1*0.2
  d16

  (p3 ph3):f2
    d0
  (p2 ph1)
    d0
  (p4 ph0):f2
    d0c
  (p3 ph4):f2

  p16:gp1*0.2
  d16

  (p1 ph2)              ;H2O in +z
    d24
```

```
  p16:gp1*0.3
    d16
  (p1*0.231 ph0 d19*2 p1*0.692 ph0 d19*2 p1*1.462 ph0 d19*2 p1*1.462 ph2
d19*2 p1*0.692 ph2 d19*2 p1*0.231 ph2):f1 (cen14 p3 ph1 3u p4 ph0 3u p3
ph1):f2
  p16:gp1*0.3
    d16 pl12:f2
    d24 setnmr2^0 setnmr0^34^32^33
  go=2 ph31 cpds2:f2

    d1 do:f2 wr #0 if #0 zd
    d11 ip3
    d11 ip13
  lo to 3 times 2

    d11 id0
    d11 ip31*2
  lo to 4 times l3
exit


ph0=0
ph1=1
ph2=2
ph3=0 2
ph13=1 3
ph4=0 0 2 2
ph10=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph11=0 0 0 0 2 2 2 2
ph12=0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
     2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
ph31=0 2 2 0
```

10 Bibliography

[1] M.H. Levitt (2001) *Spin Dynamics*, Wiley and Sons, Chichester, UK.

[2] G. Lipari and A. Szabo (1982) *J. Am. Chem. Soc.* **104**, 4546-4559.

[3] N. Tjandra, S.E. Feller, R.W. Pastor and A. Bax (1995) *J. Am. Chem. Soc.* **117**, 12562-12566.

[4] G.M. Clore, A. Szabo, A. Bax, L.E. Kay, P.C. Driscoll and A.M. Gronenborn (1990) *J. Am. Chem. Soc.* **112**, 4989-4991.

[5] A.D. Bain (2003) *Prog. NMR Spectr.* **43**, 63-103.

[6] C.S. Verma, S. Fischer, L.S.D. Caves, G.C.K. Roberts and R.E. Hubbard (1996) *J. Phys. Chem.* **100**, 2510-2518.

[7] M. Caravetta, G. Johannessen and M.H. Levitt, in *ENC*, Asilomar, **2004**.

[8] H. Eyring, H. Geishiweritz and C.E. Sun (1935) *J. Chem. Phys.* **3**, 786-790.

[9] V.Y. Orekhov, K. Pervushin and A.S. Arseniev (1994) *Eur. J. Biochem.* **219**, 887-896.

[10] A. Allerhand and H.S. Gutowsky (1965) *J. Chem. Phys.* **42**, 1587-1591.

[11] J.P. Carver and R.E. Richards (1972) *J. Magn. Reson.* **6**, 89-95.

[12] M. Tollinger, N.R. Skrynnikov, F.A.A. Mulder, J.D. Forman-Kay and L.E. Kay (2001) *J. Am. Chem. Soc.* **123**, 11341-11352.

[13] J. Jeener, B.H. Meier, P. Bachmann and R.R. Ernst (1979) *J. Chem. Phys.* **71**, 4546-4553.

[14] S.W. Englander and N.R. Kallenbach (1983) *Q. Rev. Biophys.* **16**, 521-655.

[15] C.E. Dempsey (2001) *Prog. NMR Spectr.* **39**, 135-170.

[16] R.S. Molday, S.W. Englander and R.G. Kallen (1972) *Biochemistry* **11**, 150-158.

[17] Y. Bai, J.S. Milne, L. Mayne and S.W. Englander (1993) *Proteins* **17**, 75-86.

[18] M. Christoffersen, S. Bolvig and E. Tüchsen (1996) *Biochemistry* **35**, 2309-2315.

[19] S. Koide, W. Jahnke and P.E. Wright (1995) *J. Biomol. NMR* **6**, 306-312.

[20] G.P. Connelly, Y. Bai, M.F. Jeng and S.W. Englander (1993) *Proteins* **17**, 87-92.

[21] K. Wuthrich and G. Wagner (1979) *J. Mol. Biol.* **130**, 1-18.

[22] A. Berger and K. Linderstrom-Lang (1957) *Arch. Biochem. Biophys.* **69**, 106-118.

[23] A. Hvidt and S.O. Nielsen (1966) *Adv. Protein Chem.* **21**, 287-386.

[24] H. Roder, G. Wagner and K. Wüthrich (1985) *Biochemistry* **24**, 7396-7407.

[25] G. Wagner and K. Wüthrich (1979) *J. Mol. Biol.* **130**, 31-37.

[26] C. Woodward, I. Simon and E. Tüchsen (1982) *Mol. Cell. Biochem.* **48**, 135-160.

[27] M.F. Jeng and H.J. Dyson (1995) *Biochemistry* **34**, 611-619.

[28] R.B. Martin and W.C. Hutton (1973) *J. Am. Chem. Soc.* **95**, 4752-4754.

[29] C.L. Perrin (1989) *Acc. Chem. Res.* **22**, 268-275.

[30] E. Tuchsen and C. Woodward (1985) *J. Mol. Biol.* **185**, 421-430.

[31] A.M. Jorgensen, H.B. Olsen, P. Balschmidt and J.J. Led (1996) *J. Mol. Biol.* **257**, 684-699.

[32] A.K. Bhuyan and J.B. Udgaonkar (1998) *Proteins* **30**, 295-308.

[33] C.E. Dempsey (1998) *Biochemistry* **27**, 6893-6901.

[34] S.W. Englander and L. Mayne (1992) *Annu. Rev. Biophys. Biomol. Struct.* **21**, 243-265.

[35] E. Liepinsh, G. Otting and K. Wuthrich (1992) *J. Biomol. NMR* **2**, 447-465.

[36] V.P. Denisov and B. Halle (2002) *J. Am. Chem. Soc.* **124**, 10264-10265.

[37] G.D. Henry and B.D. Sykes (1990) *Biochemistry* **29**, 6303-6313.

[38] K. Zangger and I.M. Armitage (1998) *J. Magn. Reson.* **135**, 70-75.

[39] S. Spera, M. Ikura and A. Bax (1991) *J. Biomol. NMR* **1113**, 155-165.

[40] G. Gemmecker, W. Jahnke and H. Kessler (1993) *J. Am. Chem. Soc.* **115**, 11620-11621.

[41] S. Mori, J.M. Berg and P.C.M. van Zeijl (1996) *J. Biomol. NMR* **7**, 77-82.

[42] T.L. Hwang, P.C. van Zijl and S. Mori (1998) *J. Biomol. NMR* **11**, 221-226.

[43] S. Grzesiek and A. Bax (1993) *J. Biomol. NMR* **3**, 627-638.

[44] W. Jahnke (1994) PhD thesis, TU Munich

[45] A. Bockmann, F. Penin and E. Guittet (1996) *FEBS Lett.* **383**, 191-195.

[46] K. Blum (1981) *Density Matrix Theory and Applications*, Plenum Press, New York.

[47] R.R. Ernst, G. Bodenhausen and A. Wokaun (1987) *Principles of nuclear magnetic resonance in one and two dimensions*, Oxford University Press, New York.

[48] A. Derome (1987) *Modern NMR techniques for chemistry research*, The Pergamon Press, Oxford.

[49] M.H. Levitt (1997) *J. Magn. Reson.* **126**, 164-182.

[50] A.D. Bain (1984) *J. Magn. Reson.* **56**, 418-427.

[51] A.L. Davis, J. Keeler, E.D. Laue and D. Moskau (1992) *J. Magn. Reson.* **98**, 207-216.

[52] G. Bodenhausen and D.J. Ruben (1980) *Chem. Phys. Lett.* **69**, 185-189.

[53] D.J. States, R.A. Haberkorn and D.J. Ruben (1982) *J. Magn. Reson.* **48**, 286-292.

[54] G.A. Morris and R. Freeman (1979) *J. Am. Chem. Soc.* **101**, 760-762.

[55] D. Marion and K. Wüthrich (1983) *Biochem. Biophys. Res. Comm.* **113**, 967-974.

[56] P.J. Hore (1989) *Meth. Enzymol.* **176**, 64-77.

[57] M. Gueron, P. Plateau and M. Decorps (1991) *Prog. NMR Spectr.* **23**, 135-209.

[58] V. Sklenar, M. Piotto, R. Leppik and V. Saudek (1993) *J. Magn. Reson.* **102**, 241-245.

[59] M. Piotto, V. Saudek and V. Sklenar (1992) *J. Biomol. NMR* **2**, 661-665.

[60] S. Mori, C. Abeygunawardana, M.O. Johnson and P.C. van Zijl (1995) *J. Magn. Reson. B* **108**, 94-98.

[61] S. Grzesiek and A. Bax (1993) *J. Am. Chem. Soc.* **115**, 12593-12594.

[62] N.J. Bloembergen and R.V. Pound (1954) *Phys. Rev.* **95**, 8-12.

[63] L. Müller (1979) *J. Am. Chem. Soc.* **101**, 4481- 4484.

[64] V. Tugarinov, P.M. Hwang, J.E. Ollerenshaw and L.E. Kay (2003) *J. Am. Chem. Soc.* **125**, 10420-10428.

[65] J. Keeler and D. Neuhaus (1985) *J. Magn. Reson.* **63**, 454-472.

[66] G. Kontaxis, J. Stonehouse, E.D. Laue and J. Keeler (1994) *J. Magn. Reson. A* **111**, 70-76.

[67] A.G. Palmer, J. Cavanagh, P.E. Wright and M. Rance (1991) *J. Magn. Reson.* **93**, 151-170.

[68] L.E. Kay, P. Keifer and T. Saarinen (1992) *J. Am. Chem. Soc.* **114**, 10663-10664.

[69] E. Kupce and R. Freeman (2003) *J. Magn. Reson.* **162**, 300-310.

[70] E. Kupce and R. Freeman (2003) *J. Magn. Reson.* **162**, 158-165.

[71] E. Kupce, T. Nishida and R. Freeman (2003) *Prog. NMR Spectr.* **42**, 95-122.

[72] J. Santoro and G.C. King (1992) *J. Magn. Reson.* **97**, 202-207.

[73] G.W. Vuister and A. Bax (1992) *J. Magn. Reson.* **98**, 428-435.

[74] S. Grzesiek, J. Anglister and A. Bax (1993) *J. Magn. Reson. B* **101**, 114-119.

[75] S.R. van Doren and E.R.P. Zuiderweg (1994) *J. Magn. Reson. B* **104**, 193-198.

[76] H. Aitio and P. Permi (2000) *J. Magn. Reson.* **143**, 391-396.

[77] J.L. Markley, W.J. Jorsley and M.P. Klein (1971) *J. Chem. Phys.* **55**, 3604-3607.

[78] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley and B.D. Sykes (1995) *J. Biomol. NMR* **6**, 135-140.

[79] P.B. Sigler, Z. Xu, H.S. Rye, S.G. Burston, W.A. Fenton and A.L. Horwich (1998) *Annu. Rev. Biochem.* **67**, 581-608.

[80] M. Nitsch, M. Klumpp, A. Lupas and W. Baumeister (1997) *J. Mol. Biol.* **267**, 142-149.

[81] L. Ditzel, J. Lowe, D. Stock, K.O. Stetter, H. Huber, R. Huber and S. Steinbacher (1998) *Cell* **93**, 125-138.

[82] M. Nitsch, J. Walz, D. Typke, M. Klumpp, L.O. Essen and W. Baumeister (1998) *Nat. Struct. Biol.* **5**, 855-857.

[83] I. Gutsche, L.O. Essen and W. Baumeister (1999) *J. Mol. Biol.* **293**, 295-312.

[84] I. Gutsche, O. Mihalache and W. Baumeister (2000) *J. Mol. Biol.* **300**, 187-196.

[85] Z. Wang, H. Feng, S.J. Landry, J. Maxwell and L.M. Gierasch (1999) *Biochemistry* **38**, 12537-12546.

[86] M.R. Leroux and F.U. Hartl (2000) *Curr. Biol.* **10**, R260-264.

[87] W.A. Fenton, Y. Kashi, K. Furtak and A.L. Horwich (1994) *Nature* **371**, 614-619.

[88] M. Klumpp, W. Baumeister and L.O. Essen (1997) *Cell* **91**, 263-270.

[89] G. Bosch, W. Baumeister and L.O. Essen (2000) *J. Mol. Biol.* **301**, 19-25.

[90] G. Pappenberger, J.A. Wilsher, S.M. Roe, D.J. Counsell, K.R. Willison and L.H. Pearl (2002) *J. Mol. Biol.* **318**, 1367-1379.

[91] M. Heller, M. John, M. Coles, G. Bosch, W. Baumeister and H. Kessler (2004) *J. Mol. Biol.* **336**, 717-729.

[92] M. John, M. Heller, M. Coles, G. Bosch, W. Baumeister and H. Kessler (2003) *J. Biol. NMR* **29**, 209-210.

[93] L.E. Kay, M. Ikura, R. Tschudin and A. Bax (1990) *J. Magn. Reson.* **89**, 496-514.

[94] M. Sattler, J. Schleucher and C. Griesinger (1999) *Prog. NMR Spectr.* **34**, 93-158.

[95] M. Salzmann, G. Wider, K. Pervushin, H. Senn and K. Wüthrich (1999) *J. Am. Chem. Soc.* **121**, 844-848.

[96] M. Leutner, R.M. Gschwind, J. Liermann, C. Schwarz, G. Gemmecker and H. Kessler (1998) *J. Biomol. NMR* **11**, 31-43.

[97] S. Grzesiek and A. Bax (1993) *J. Magn. Reson. B* **102**, 103-106.

[98] D.S. Wishart and B.D. Sykes (1994) *Meth. Enzymol.* **239**, 363-392.

[99] Y. Wang and O. Jardetzky (2004) *J. Biomol. NMR* **28**, 327-340.

[100] D.S. Wishart, C.G. Bigam, A. Holm, R.S. Hodges and B.D. Sykes (1995) *J. Biomol. NMR* **5**, 67-81.

[101] D.S. Wishart, B.D. Sykes and F.M. Richards (1992) *Biochemistry* **31**, 1647-1651.

[102] D.S. Wishart, M.S. Watson, R.F. Boyko and B.D. Sykes (1997) *J. Biomol. NMR* **10**, 329-336.

[103] B.R. Seavey, E.A. Farr, W.M. Westler and J.L. Markley (1991) *J. Biomol. NMR* **1**, 217-236.

[104] F. Cordier, M. Barfield and S. Grzesiek (2003) *J. Am. Chem. Soc.* **125**, 15750-15751.

[105] G.W. Vuister and A. Bax (1993) *J. Am. Chem. Soc.* **115**, 7772-7777.

[106] A. Pardi, M. Billeter and K. Wüthrich (1984) *J Mol Biol* **180**, 741-751.

[107] W. Jahnke, M. Baur, G. Gemmecker and H. Kessler (1995) *J. Magn. Reson. B* **106**, 86-88.

[108] T. Frenkiel, C. Bauer, M.D. Carr, B. Birdsall and J. Feeney (1990) *J. Magn. Reson.* **90**, 420-425.

[109] G. Otting, E. Liepinsh and K. Wüthrich (1991) *Science* **254**, 974-980.

[110] M. Buck (1998) *Q. Rev. Biophys.* **31**, 297-355.

[111] M. Heller (2004) PhD thesis, Tschechische Universitaet Munich

[112] T. Kolter and K. Sandhoff (1999) *Angew. Chem. Int. Ed. Engl.* **38**, 1532-1568.

[113] W. Furst and K. Sandhoff (1992) *Biochim. Biophys. Acta* **1126**, 1-16.

[114] E. Conzelmann and K. Sandhoff (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3979-3983.

[115] E.M. Meier, G. Schwarzmann, W. Fürst and K. Sandhoff (1991) *J. Biol. Chem.* **266**, 1879-1887.

[116] G.J. Glombitza, E. Becker, H.W. Kaiser and K. Sandhoff (1997) *J. Biol. Chem.* **272**, 5199-5207.

[117] D.J. Mahuran (1999) *Biochim. Biophys. Acta* **1455**, 105-138.

[118] C.S. Wright, S.C. Li and F. Rastinejad (2000) *J. Mol. Biol.* **304**, 411-422.

[119] C.S. Wright, Q. Zhao and F. Rastinejad (2003) *J. Mol. Biol.* **331**, 951-964.

[120] U. Bierfreund, T. Lemm, A. Hoffmann, G. Uhlhorn-Dierks, R.A. Childs, C.T. Yuen, T. Feizi and K. Sandhoff (1999) *Neurochem. Res.* **24**, 295-300.

[121] J.L. Cereghino and J.M. Cregg (2000) *FEMS Microbiol. Rev.* **24**, 45-66.

[122] M. Wendeler, J. Hoernschemeyer, M. John, N. Werth, M. Schoeniger, T. Lemm, R. Hartmann, H. Kessler and K. Sandhoff (2004) *Protein Expr. Purif.* **34**, 147-157.

[123] K. Pervushin, R. Riek, G. Wider and K. Wüthrich (1997) *Proc. Natl. Acad. Sci. USA* **94**, 12366-12371.

[124] S.C. Li, Y. Hirabayashi and Y.T. Li (1981) *J. Biol. Chem.* **256**, 6234-6240.

[125] E.O. Stejskal and J.E. Tanner (1965) *J. Chem. Phys.* **41**, 288-292.

[126] J.E. Tanner (1970) *J. Chem. Phys.* **52**, 2523-2526.

[127] S.J. Gibbs and C.S. Johnson (1991) *J. Magn. Reson.* **93**, 395-402.

[128] D. Wu, A. Chen and C.S. Johnson (1995) *J. Magn. Reson. A* **115**, 260-264.

[129] A. Jerschow and N. Müller (1997) *J. Magn. Reson.* **125**, 372-375.

[130] J.G. de la Torre, M.L. Huertas and B. Carrasco (2000) *J. Magn. Reson.* **147**, 138-146.

[131] N.A. Farrow, R. Muhandiram, A.U. Singer, S.M. Pascal, C.M. Kay, G. Gish, S.E. Shoelson, T. Pawson, J.D. Formankay and L.E. Kay (1994) *Biochemistry* **33**, 5984-6003.

[132] V.Y. Orekhov, D.E. Nolde, A.P. Golovanov, D.M. Korzhnev and A.S. Arseniev (1995) *Appl. Magn. Reson.* **9**, 581-588.

[133] H. Yohe and A. Rosenberg (1972) *Chem. Phys. Lipids* **9**, 279-282.

[134] D. Orthaber and O. Glatter (1998) *Chem. Phys. Lipids* **92**, 53-62.

[135] N. SmiljanicGeorgijev, B. Rigat, B. Xie, W. Wang and D.J. Mahuran (1997) *Biochim. Biophys. Acta* **1339**, 192-202.

[136] Y. Hama, Y.T. Li and S.C. Li (1997) *J. Biol. Chem.* **272**, 2828-2833.

[137] Y. Kishimoto, M. Hiraiwa and J.S. Obrien (1992) *J. Lipid Res.* **33**, 1255-1267.

[138] A.L. Fluharty (1995) *Trends Glycosci. Glycotechnol.* **7**, 167-189.

[139] V.E. Ahn, K.F. Faull, J.P. Whitelegge, A.L. Fluharty and G.G. Prive (2003) *Proc. Natl. Acad. Sci. USA* **100**, 38-43.

[140] J. Kervinen, G.J. Tobin, J. Costa, D.S. Waugh, A. Wlodawer and A. Zdanov (1999) *EMBO J.* **18**, 3947-3955.

[141] C.P. Pointing and R.B. Russell (1995) *Trends Biochem. Sci.* **20**, 179-180.

[142] E. Liepinsh, M. Andersson, J.M. Ruysschaert and G. Otting (1997) *Nature Struct. Biol.* **4**, 793-795.

[143] C. Gonzalez, G.M. Langdon, M. Bruix, A. Galvez, E. Valdivia, M. Maqueda and M. Rico (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11221-11226.

[144] E. de Alba, S. Weiler and N. Tjandra (2003) *Biochemistry* **42**, 14729-14740.

[145] S. Soeda, M. Hiraiwa, J.S. Obrien and Y. Kishimoto (1993) *J. Biol. Chem.* **268**, 18519-18523.

[146] A.M. Vaccaro, F. Ciaffoni, M. Tatti, R. Salvioli, A. Barca, D. Tognozzi and C. Scerch (1995) *J. Biol. Chem.* **270**, 30576-30580.

[147] C. Chothia (1992) *Nature* **357**, 543-544.

[148] D.T. Jones and J.M. Thornton (1996) *Curr. Opin. Struct. Biol.* **6**, 210-216.

[149] T. Schwede, J. Kopp, N. Guex and M.C. Peitsch (2003) *Nucleic Acids Res.* **31**, 3381-3385.

[150] K. Wüthrich (1986) *NMR of proteins and nucleic acids*, Wiley-Interscience Publication, John Wiely and Sons, New York.

[151] G.W. Vuister, G.M. Clore, A.M. Gronenborn, R. Powers, D.S. Garrett, R. Tschudin and A. Bax (1993) *J. Magn. Reson. B* **101**, 210-213.

[152] O. Zhang, L.E. Kay, J.P. Olivier and J.D. Forman-Kay (1994) *J. Biomol. NMR* **4**, 845-858.

[153] S.J. Archer, M. Ikura, D. Torchia and A. Bax (1991) *J. Magn. Reson.* **95**, 636-641.

[154] J. Briand and R.R. Ernst (1991) *Chem. Phys. Lett.* **185**, 276-285.

[155] F. Kramer, M.V. Deshmukh, H. Kessler and S.J. Glaser (2004) *Conc. Magn. Reson.* **21**, 10-21.

[156] E. de Alba and N. Tjandra (2002) *Prog. NMR Spectr.* **40**, 175-197.

[157] A. Bax, G. Kontaxis and N. Tjandra (2001) *Meth. Enzymol.* **339**, 127-173.

[158] J.H. Prestegard, H.M. al-Hashimi and J.R. Tolman (2000) *Q. Rev. Biophy.* **33**, 371-424.

[159] F. Delaglio, G. Kontaxis and A. Bax (2000) *J. Am. Chem. Soc.* **122**, 2142-2143.

[160] J.C. Hus, D. Marion and M. Blackledge (2000) *J. Mol. Biol.* **298**, 927-936.

[161] J.R. Tolman, J.M. Flanagan, M.A. Kennedy and J.H. Prestegard (1995) *Proc. Natl. Acad. Sci. USA* **92**, 9279-9283.

[162] N. Tjandra and A. Bax (1997) *Science* **278**, 1111-1114.

[163] M. Ottiger, F. Delaglio and A. Bax (1998) *J. Magn. Reson.* **131**, 373-378.

[164] C.R. Sanders and J.P. Schwonek (1992) *Biochemistry* **31**, 8898-8905.

[165] J.A. Losonczi and J.H. Prestegard (1998) *J. Biomol. NMR* **12**, 447-451.

[166] E. Brunner (2001) *Conc. Magn. Reson.* **13**, 238-259.

[167] M. Ottiger and A. Bax (1998) *J. Biomol. NMR* **12**, 361-372.

[168] H.J. Sass, G. Musco, S.J. Stahl, P.T. Wingfield and S. Grzesiek (2000) *J. Biomol. NMR* **18**, 303-309.

[169] R. Tycko, F.J. Blanco and Y. Ishii (2000) *J. Am. Chem. Soc.* **122**, 9340-9341.

[170] J.J. Chou, S. Gaemers, B. Howder, J.M. Louis and A. Bax (2001) *J. Biomol. NMR* **21**, 377-382.

[171] S. Meier, D. Häussinger and S. Grzesiek (2002) *J. Biomol. NMR* **24**, 351-356.

[172] D.A. Marvin (1998) *Curr. Opin. Struct. Biol.* **8**, 150-158.

[173] M.R. Hansen, L. Mueller and A. Pardi (1998) *Nat. Struct. Biol.* **5**, 1065-1074.

[174] M. Zweckstetter and A. Bax (2001) *J. Biomol. NMR* **20**, 365-377.

[175] M.F. Mesleh, G. Veglia, T.M. deSilva, F.M. Marassi and S.J. Opella (2002) *J. Am. Chem. Soc.* **124**, 4206-4207.

[176] D.D. Ojennus, R.M. Mitton-Fry and D.S. Wuttke (1999) *J. Biomol. NMR* **14**, 175-179.

[177] M.F. Mesleh, S. Lee, G. Veglia, D.S. Thiriot, F.M. Marassi and S.J. Opella (2003) *J. Am. Chem. Soc.* **125**, 8928-8935.

[178] M.F. Mesleh and S.J. Opella (2003) *J. Magn. Reson.* **163**, 288-299.

[179] F.M. Marassi and S.J. Opella (2000) *J. Magn. Reson.* **144**, 150-155.

[180] N.E. Zhou, B.Y. Zhu and B.D. Sykes (1992) *J. Am. Chem. Soc.* **114**, 4320-4326.

[181] A. Mascioni and G. Veglia (2003) *J. Am. Chem. Soc.* **125**, 12520-12526.

[182] P. Dosset, J.C. Hus, D. Marion and M. Blackledge (2001) *J. Biomol. NMR* **20**, 223-231.

[183] P.J. Kraulis (1991) *J. Appl. Cryst.* **24**, 946-950.

[184] E.A. Merritt and M.E.P. Murphy (1994) *Acta Cryst. D* **50**, 869-873.

[185] G. Otting, E. Liepinsh and K. Wüthrich (1993) *Biochemistry* **32**, 3571-3582.

[186] R.R. Fraser, G. Boussard, J.K. Saunders, J.B. Lambert and C.E. Mixan (1971) *J. Am. Chem. Soc.* **93**, 3822-3823.

[187] G. Jung and M. Ottnad (1974) *Angew. Chem. Int. Ed. Engl.* **13**, 818-819.

[188] C. Griesinger, G. Otting, K. Wüthrich and R.R. Ernst (1988) *J. Am. Chem. Soc.* **110**, 7870-7872.

[189] G.T. Montelione and G. Wagner (1989) *J. Am. Chem. Soc.* **111**, 3096-3098.

[190] N.A. Farrow, O. Zhang, J.D. Forman-Kay and L.E. Kay (114) *J. Biomol. NMR* **4**, 727-734.

[191] J.D. Dunitz (1995) *Chem. Biol.* **2**, 709-712.

[192] A. Cornish-Bowden (2002) *J. Biosci.* **27**, 121-126.

[193] M.J. Grey, C. Wang and A.G. Palmer (2003) *J. Am. Chem. Soc.* **125**, 14324-14335.

[194] N. Tjandra, S. Grzesiek and A. Bax (1996) *J. Am. Chem. Soc.* **118**, 6264-6272.