Protein precipitates, aggregation kinetics and membrane protein receptors characterized by solid-state NMR

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

vorgelegt von

Manuel Etzkorn

aus Aachen

Göttingen

2008

This thesis is a product of the collaboration between the Max-Planck Institute for biophysical chemistry and the Georg-August-University, Göttingen.

Kurzzusammenfassung

Die vorliegende Arbeit befasst sich mit der Untersuchung nicht löslicher, biomolekularer Systeme mit Hilfe der Kernmagnetischen-Resonanzspektroskopie (NMR). Einen Schwerpunkt bildet die Erkundung neuer Anwendungsgebiete wie die Messung an aus Lösung ausgefallenen Proteinen und die zweidimensionale Festkörper-Echtzeitspektroskopie. Des Weiteren wurden die Grenzen bekannter Anwendungsgebiete, insbesondere im Bereich der Membranproteine, ausgelotet und durch neue Analysemethoden erweitert. Im Einzelnen konnte ein atomares Bild des präzipitierten Zustandes des Crh-Proteins entwickelt werden. Ausgehend von diesen Ergebnissen konnte erstmalig Proteinaggregation, induziert durch eine temperaturbedingte strukturelle Umwandlung des Crh-Präzipitats, in Echtzeit aufgenommen und kinetisch analysiert werden.

Im Weiteren wurde das System der zwei Membranproteine SRII und HtrII untersucht. Neben Untersuchungen an isoliertem SRII führten Messungen an dem Proteinkomplex der beiden Proteine zu einem erweiterten Bild der SRII/HtrII-Bindungsfläche. Zusätzlich wurde die Funktionsweise des SRII/HtrII-Komplexes mit Hilfe von Spektren nach Lichtaktivierung untersucht.

Abschließend wurde durch Kombination verschiedener Techniken eine strukturelle Untersuchung des Multidomänen-Membranproteins DcuS im besonderen Hinblick auf dessen Funktion vorgenommen. Hierzu wurden erstmalig die spektroskopischen Daten einer nicht löslichen Domäne im Vergleich zu einer computergestützten Strukturvorhersage analysiert. Die erhaltenen Ergebnisse, sowie Informationen vorausgegangener Untersuchungen ermöglichten die Einführung eines konsistenten Modells der Signalweiterleitung. Vergleiche mit ähnlichen Systemen deuten eine Allgemeingültigkeit des vorgeschlagenen Mechanismus an.

Abstract

This work addresses the investigation of insoluble, biomolecular systems using nuclear magnetic resonance (NMR). Introducing novel areas of applications an atomic picture of the precipitated state of the Crh protein as well as kinetic information of protein aggregation using real time solid-state NMR could be obtained.

Additionally the SRII/HtrII membrane protein system was studied. Experimental results obtained on SRII and on SRII in complex with HtrII led to the identification of an extended binding mode between the two membrane proteins. Measurements on the light activated state of the complex were carried out to investigate the mechanism of signal transduction.

Functional aspects were also investigated for the multidomain membrane protein DcuS. Therefore spectra obtained on an isolated insoluble domain were compared to *in silico* structure prediction data. The results were combined with previous data of an additional soluble domain as well as mutagenesis data. A self consistent model of signal transduction, which may represent a general mechanism present also in similar systems, could be proposed.

Contents

Contents

1 Chapter 1
Introduction

1.1 Scientific background and motivation

Biological life is intimately related to (inter)actions of proteins. Protein function itself is diverse and in general associated with its three dimensional structure [1]. A variety of experimental techniques to determine protein structure exist. Measured by the number of structures solved so far, X-ray crystallography ($\backsim$43000) and solution-state NMR ($\backsim$7200) have proven to be most successful. However some protein classes are difficult to access with these techniques. Two of these classes form the basis of the presented work and will be introduced in the following.

1.1.1 Membrane proteins

Reflecting the inherent challenge in structure determination of membrane proteins, a tremendous gap exists between the occurrence of membrane embedded proteins as encoded in the genome ($\backsim$25%) or as targeted by drugs ($\backsim$50%) [2] and their actual representation in the data base of known protein structures [3] ($<$1%).

In general, each cell is constantly sensing its environment and will react on specific signals. A key aspect in a fundamental understanding of cellular signaling is the characterization of molecular structure and dynamics of membrane proteins. One example

connected to this thesis and essential for prokaryotic bacteria are the so called two component systems, which sense extra cellular stimuli such as light (see chapters 6 and 7) or nutrition (see chapter 8) and regulate gene expression to account for specific environmental conditions. In humans membrane proteins are responsible for e.g. sensing taste [4], transmitting pain [5] and initiating cell apoptosis [6], a critical objective in cancer research [7].

Due to the outstanding importance of membrane proteins a huge effort is made to optimize strategies for resolving structural information. Indeed, in analogy to soluble proteins, it is believed that the number of known membrane protein structures will increase exponentially [8]. Although the actual number of known structures ($\backsim$157) is already considerably smaller than the predicted value ($>$210) after four years. However, besides the needs of *de novo* structure determination of full size membrane proteins, additional challenges emerge, two of which are closely connected to the presented work.

The first relates to the question of structural integrity of already solved membrane protein structures. In particular missing or artificially introduced interactions in the crystal lattice or membrane mimicking detergent can potentially influence structural features. As a result interpretation of structural aspects can be misleading in the context of a native membrane environment.

Another aspect concerns a large fraction of membrane proteins, which are composed of several interacting protein domains. Important information about such multidomain membrane related proteins can be obtained by studying the isolated, often soluble domains. However, an experimental need to probe the validity of structural attributes in the context of the full length membrane embedded protein is evident [9].

1.1.2 Protein aggregates and precipitates

A second class of proteins, which is difficult to access by X-ray diffraction or solution-state NMR are protein aggregates. A number of serious diseases such as Alzheimer's, Parkinson's and Huntington's disease (and others see e.g. [10]) is associated with aggregation of specific proteins. In particular early aggregation intermediates and an understanding of the aggregation kinetics itself was identified as a key aspect in developing

useful therapeutic treatments [11]. Often those, in general fibrillar aggregates, do not form crystals and their molecular weight is too large for high resolution structural studies in solution [12].

Notably, a variety of globular proteins is unstable in solution and additionally does not form larger crystals. In contrast to most protein aggregates, protein precipitation can in principle stabilize native protein structure. As long as the native structural features are largely retained, their characterization can benefit from an experimental technique applicable to protein precipitates.

1.2 Overview and outline of this thesis

The inherent capability of solid-state NMR (ssNMR) for probing non-crystalline systems, such as protein aggregates [13] and membrane proteins [14], render this technique in particular useful to approach the challenges pointed out above. In the following, after a general introduction into the basic theory (chapter 2) and spectroscopic methodologies (chapter 3), explicit contributions of ssNMR are presented.

The first aspect is discussed in chapter 4 and concerns protein structure in the precipitated state. The lack of long range order and the per definition insoluble nature of protein precipitates render ssNMR the only technique, which can report on overall structure with atomic resolution. It is shown that precipitation at the isoelectric point (pI) does not significantly alter the three dimensional structure of the protein Crh in respect to its crystalline appearance. Additionally, for the first time protein aggregation could be followed using two dimensional real time ssNMR spectroscopy, which led to the detection of a partially unfolded state prior to protein aggregation that is characterized by an increased β-strand content.

Based on the spectral footprints detected during the aggregation process, chapter 5 reports a detailed theoretical analysis of the folding transition and shows that real time ssNMR can contribute to the kinetic analysis of protein aggregation.

The following three chapters deal with the structural investigation of two component systems. First the limits of an ssNMR based *de novo* structural investigation of a seven

transmembrane helical (7 TM) protein receptor were explored (chapter 6). Combination of new ssNMR methodology, optimized sample labeling and high field spectrometers enabled spectral assignments in large fragments of the receptor. Additionally dynamical features and aspects of the membrane topology could be probed.

Following these results complex formation of the receptor to its cognate transducer molecule as well as complex activation was investigated (chapter 7). The results were analyzed in reference to crystallographic information. In spite of large similarities, differences were found which predominantly occur in regions that experience close intermolecular contacts within the crystal lattice. Besides the biological relevance, this study extended the size limits of accessible uniformly labeled proteins compared to previous ssNMR work.

Finally the last project (chapter 8) involves the investigation of a multidomain membrane embedded sensory histidine-kinase, for which only the structure of the sensory domain was available. This study benefits in particular from the results obtained in chapters 4 and 6. In detail, protein precipitation enabled the structural investigation of a protein domain that could not be studied using X-ray or solution-state NMR. A novel approach that combines ssNMR with comparative modeling was introduced, which significantly facilitated the structural characterization of the precipitated domain. Additionally, the validity of structural aspects of the isolated domains compared to their appearance in the multidomain protein was investigated using ssNMR on a 40kDa membrane embedded protein. Based on these data a model for intra-cellular signal transduction, regulating kinase activity, could be proposed. A mutagenesis study, carried out in a collaborating lab, could additionally corroborate the presented functional interpretation, revealing that ssNMR offers a spectroscopic means for structural investigation of large membrane proteins in close reference to biological function.

2 Chapter 2
Basic principles of (solid-state) NMR

Nuclear magnetic resonance (NMR) is based on the effect of an external magnetic field on a magnetic momentum of an atomic nucleus. For a proper understanding of the methods and results presented in this work, knowledge of the interactions between nuclei and external fields as well as between the nuclei themselves is fundamental. Therefore in the following chapter the basic concepts of NMR as well as their quantum mechanical description will be discussed. Experimental aspects and methodologies explicitly used throughout the presented work will be introduced in chapter 3.

2.1 Origin of the signal

The nuclear spin is composed of the intrinsic spins of the protons and the neutrons as well as of their orbital angular momenta. Only nuclei which do not have an even number of protons and neutrons experience a magnetic momentum $\neq 0$. Its first experimental detection was achieved by Rabi and Cohen for the sodium nucleus [15][1]. The magnetic momentum $\hat{\boldsymbol{\mu}}$ of a nuclear spin $\hat{\boldsymbol{I}}$ can be written as:

$$\hat{\boldsymbol{\mu}} = \gamma \hat{\boldsymbol{I}} \tag{2.1}$$

In equation (2.1) γ is called the gyromagnetic ratio, a constant characteristic for a

---

[1]Ref. [16] nicely reviews the 'milestones' related to the discovery of the spin.

specific nucleus and $\hat{\boldsymbol{I}}$ is the spin angular momentum operator (see supporting figure on page 110).

2.2 External spin interactions

2.2.1 Zeeman splitting

A magnetic momentum in a static magnetic fields leads to a separation of the possible energy levels of the spin eigenstates. In this work only nuclei with spin quantum number $I = \frac{1}{2}$ were investigated and the static magnetic field is referred to as $\boldsymbol{B_0} = (0, 0, B_0)$. The resulting nuclear magnetic energy can be expressed by the Hamilton operator of the Zeeman interaction:

$$\hat{\mathcal{H}}_Z = -\gamma \hat{I}_z B_0 \tag{2.2}$$

The term $-\gamma B_0$ is also called Larmor frequency of the spin.

2.2.2 Radio frequency perturbation

Radio frequency (r.f.) generating coils are in general implemented with their symmetry axis perpendicular to $\boldsymbol{B_0}$. The resulting current induced, time dependent field $\boldsymbol{B_{rf}}(t)$, is transversal and normally several orders of magnitude smaller than $\boldsymbol{B_0}$. In analogy to equation (2.2) the underlying Hamilton operator reads as follows:

$$\hat{\mathcal{H}}_{rf} = -\gamma \hat{\boldsymbol{I}} \boldsymbol{B_{rf}}(t) \tag{2.3}$$

A transversal r.f. field along the x-axis of amplitude $2B_{rf}$, oscillation carrier frequency ω_c and phase ϕ of the form:

$$\boldsymbol{B_{rf}}(t) = 2B_{rf} \cos(\omega_c t + \phi) \cdot \boldsymbol{e}_x \tag{2.4}$$

can be divided in two counter-rotating components in the xy-plane. Only the component rotating in the same sense as the spin precision has a significant effect on the motion of the spin. Hence, equation (2.3) can be written as:

$$\hat{\mathcal{H}}_{rf} = \gamma B_{rf} \left[\cos(\omega_c t + \phi) \hat{I}_x + \sin(\omega_c t + \phi) \hat{I}_y \right] \tag{2.5}$$

2.3 Anisotropy in NMR

In general the interactions a magnetic momentum experiences within a NMR sample are anisotropic i.e. dependent on the orientation in respect to the external magnetic field. The major difference between NMR spectroscopy of (smaller) proteins in solution and in a more solid phase (e.g. crystalline, aggregated, membrane embedded, bound to large complexes, etc.) is related to the contributions from these anisotropic interactions. In solution the molecular tumbling dynamically averages out anisotropic components whereas the orientation dependence of the interactions is retained in solid-state NMR.

Anisotropic interactions can most conveniently be described using a second-rank tensor $\mathcal{A}$ taking the orientation of the interaction into account.

$$\mathcal{A} = \begin{pmatrix} a_{xx} & a_{xy} & a_{xz} \\ a_{yx} & a_{yy} & a_{yz} \\ a_{zx} & a_{zy} & a_{zz} \end{pmatrix} \quad (2.6)$$
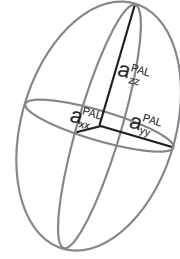
For the interactions considered here a special coordinate system (see section 2.5.4) can be found were $\mathcal{A}$ is diagonal, with the eigenvalues a_{xx}, a_{yy} and a_{zz}. Such a tensor can be illustrated as shown in figure 2.1. Alternatively, $\mathcal{A}$ can be split into an isotropic (a_{iso}), an anisotropic (δ) and an asymmetric (η) component:

Figure 2.1: Representation of a second-rank tensor as ellipsoid.

$$\begin{aligned} a_{iso} &= \frac{1}{3}(a_{xx} + a_{yy} + a_{zz}) & (2.7) \\ \delta &= a_{zz} - a_{iso} & (2.8) \\ \eta &= \frac{a_{yy} - a_{xx}}{\delta}, & (2.9) \end{aligned}$$

where $|a_{zz} - a_{iso}| \geq |a_{xx} - a_{iso}| \geq |a_{yy} - a_{iso}|$.

2.4 Internal spin interactions

In addition to the interactions described above, the magnetic momentum of a nucleus also experience magnetic and electric fields originating from the sample itself. The most

import internal interactions for spin-1/2 nuclei will be discussed in the following.

2.4.1 Chemical shielding

According to the Zeeman splitting (2.2) nuclei with the same gyromagnetic ratio γ, would have exactly the same resonance frequency. However, an essential aspect in NMR spectroscopy is that the chemical environment can influence the local magnetic field a nucleus senses and hence change its resonance frequency. This effect arises from induced circular currents in the electron clouds, which themselves generate a small local magnetic field. These so called chemical shielding interaction is anisotropic since it depends on the shape and orientation of the electron cloud in respect to the magnetic field $\boldsymbol{B_0}$. The local field $\boldsymbol{B_l}$ can hence be expressed as a superposition of the shielding field $\boldsymbol{B_{cs}}$ and $\boldsymbol{B_0}$:

$$\boldsymbol{B_l} = \boldsymbol{B_0} + \boldsymbol{B_{cs}} = (\boldsymbol{\mathcal{I}} + \boldsymbol{\sigma}) \cdot \boldsymbol{B_0} \tag{2.10}$$

Where $\boldsymbol{\sigma}$ is the chemical shielding tensor and $\boldsymbol{\mathcal{I}}$ is the identity matrix. The resulting Hamiltonian reads as follows:

$$\hat{\mathcal{H}}_{cs} = -\gamma \hat{\boldsymbol{I}} \boldsymbol{\sigma} \boldsymbol{B_0}. \tag{2.11}$$

2.4.2 Dipolar coupling

The magnetic momentum of a spin itself generates a magnetic field, which acts on the magnetic fields of the surrounding nuclei and vice versa. In contrast to the previous interaction this so called dipole-dipole interactions is independent of the strength of $\boldsymbol{B_0}$. The Hamiltonian is of the form:

$$\hat{\mathcal{H}}_D = \hat{\boldsymbol{I}}_{\boldsymbol{1}} \boldsymbol{\mathcal{D}} \hat{\boldsymbol{I}}_{\boldsymbol{2}} \tag{2.12}$$

According to classical mechanics the underlying energy is well defined [17] and directly follows from the correspondence principle:

$$\hat{\mathcal{H}}_D = -\frac{\mu_0}{4\pi} \frac{\gamma_1 \gamma_2}{r_{12}^3} \left(3 \left(\hat{\boldsymbol{I}}_1 \cdot \boldsymbol{e}_{12} \right) \left(\hat{\boldsymbol{I}}_2 \cdot \boldsymbol{e}_{12} \right) - \hat{\boldsymbol{I}}_1 \cdot \hat{\boldsymbol{I}}_2 \right) \tag{2.13}$$

Here r_{12} is the spin-spin distance, $\boldsymbol{e}_{12}$ is the unit vector parallel to the line joining the two nuclei and μ_0 is a constant.

2.4.3 J-coupling

The magnetic moments of nuclei which are linked via a chemical bond are additionally affected by indirect magnetic interactions through the involvement of the electrons. Unlike the direct dipolar interactions discussed in 2.4.2 these so called J-couplings normally can be reduced to an isotropic component. However, introducing the J-coupling tensor $\boldsymbol{\mathcal{J}}$ the general form of the Hamiltonian (in units of Hz) reads as follows:

$$\hat{\mathcal{H}}_J = 2\pi \hat{\boldsymbol{I_1}} \boldsymbol{\mathcal{J}} \hat{\boldsymbol{I_2}} \tag{2.14}$$

2.5 Useful mathematical simplifications

2.5.1 Secular approximation

Since in general for spin-1/2 nuclei the Zeeman interaction is by far the strongest, the other interactions can be treated as a small perturbation of the system. Under this assumption a particular useful, so called secular or high field approximation can be applied. According to perturbation theory, only the parts of the Hamiltonians that commute with $\hat{\mathcal{H}}_Z$ (2.2) contribute to the total energy.

2.5.2 Frames of reference

The description of the discussed interaction can significantly benefit from the choice of an adequate frame of reference. Two such reference frames, namely the *laboratory frame* (LAB) (i.e. the 'regular' space) and the *principle axis frame* (PAF) (i.e. the frame in which the interaction tensor is of diagonal form), have already been used. Other important frames are the *molecular frame* or the *rotor frame*, which describe an interaction fixed to the molecule or an rotating cylinder, respectively. Of particular interest is also the *rotating frame* (ROT) which is defined by constant rotation (normally equal to the lamor frequency) in respect to the LAB.

In general the transformation of an operator in the new frame with respect to the old frame is of the form:

$$\hat{A}^{new} = \hat{R}^{-1} \hat{A}^{old} \hat{R} \tag{2.15}$$

Where $\hat{R}(\alpha, \beta, \gamma)$ is an operator which performs a rotation by the Euler angles α, β, γ:

$$\hat{R}(\alpha, \beta, \gamma) = e^{-i\alpha\hat{L}_z} e^{-i\beta\hat{L}_y} e^{-i\gamma\hat{L}_z} \tag{2.16}$$

A transformation of spatial coordinates (vide infra) is in general given by $\hat{\boldsymbol{L}} = \hat{\boldsymbol{J}}$ ($\hat{\boldsymbol{J}}$ is the angular momentum operator), whereas $\hat{\boldsymbol{L}} = \hat{\boldsymbol{I}}$ in spin space.

2.5.3 Irreducible spherical tensors

The rotation of an operator as given in equation (2.16) can lead to complicate expressions using regular Cartesian coordinates. Alternatively the so called irreducible spherical representation of a tensor which components fulfill the following transformation property:

$$\hat{A}_{l,m}^{new} = \hat{R}^{-1} \hat{A}_{l,m}^{old} \hat{R} = \sum_{m'=-l}^{l} \hat{A}_{l,m'}^{old} D_{mm'}^{l}(\alpha, \beta, \gamma) \tag{2.17}$$

can be used. $D_{mm'}^{l}$ are elements of the *Wigner rotation matrix* [18]. Dividing the effect of an interaction $\boldsymbol{\mathcal{A}}$ on a spin system in an irreducible spherical tensor operator $\hat{\mathscr{T}}$, containing the spin information and an irreducible spherical tensor $\mathscr{A}$, which expresses the orientation dependence of the interaction an alternative way of expressing the resulting Hamiltonian is given by:

$$\hat{\mathcal{H}}_A = \hat{\boldsymbol{I}}_1 \boldsymbol{\mathcal{A}} \hat{\boldsymbol{I}}_2 = \sum_{l=0}^{2} \sum_{=-l}^{l} (-1)^m \mathscr{A}_{l,-m} \hat{\mathscr{T}}_{l,m} \tag{2.18}$$

Where $\mathscr{A}$ and $\hat{\mathscr{T}}$ depend on the specific interaction as for example given in [19].

2.5.4 Secular Hamiltonians in the rotating frame

Making use of the listed formalisms the description of the interaction becomes simplified. For example the secular parts of the interaction Hamilton operators can be expressed in the rotating frame in Cartesian coordinates according to:

$$\hat{\mathcal{H}}_{rf} = \gamma B_{rf} \left(\hat{I}_x \cos\phi + \hat{I}_y \sin\phi \right) \tag{2.19}$$

$$\hat{\mathcal{H}}_{cs} = \omega_0 \sigma_{iso} + \frac{1}{2}\omega_0 \delta \left(3\cos^2\theta - 1 + \eta\sin^2\theta\cos 2\varphi \right) \hat{I}_z \tag{2.20}$$

$$\hat{\mathcal{H}}_D^{homo} = -d_{12}\frac{1}{2} \left(3\cos^2\theta - 1 \right) \left(3\hat{I}_{z1}\hat{I}_{z2} - \hat{\boldsymbol{I}}_1\hat{\boldsymbol{I}}_2 \right) \tag{2.21}$$

$$\hat{\mathcal{H}}_D^{hetero} = -d_{12} \left(3\cos^2\theta - 1 \right) \hat{I}_{z1}\hat{I}_{z2} \tag{2.22}$$

Where σ_{iso}, δ and η are the components of $\boldsymbol{\sigma}^{PAF}$ according to equations (2.7)-(2.8). σ_{iso} is the isotropic chemical shift relative to the Lamor frequency ω_0 and the polar angles θ and φ define the orientation of $\boldsymbol{B}_0$ in the respective PAF. Additionally the dipolar coupling constant:

$$d_{12} := -\frac{\mu_0}{4\pi} \frac{\gamma_1 \gamma_2}{r_{12}^3} \tag{2.23}$$

was used in equation (2.21) and equation (2.22).

Relevant irreducible spherical components. Analogue to the representation in Cartesian coordinates, most of the components of the irreducible spherical description introduced in 2.5.3 can be neglected. The remaining components are:

$$\hat{\mathcal{H}}_{cs} = \mathscr{A}_{00}^{PAF} \cdot \hat{\mathscr{T}}_{00} + \mathscr{A}_{20}^{PAF} \cdot \hat{\mathscr{T}}_{20} \tag{2.24}$$

$$\hat{\mathcal{H}}_D = \mathscr{A}_{20}^{PAF} \cdot \hat{\mathscr{T}}20 \tag{2.25}$$

2.6 Quantum mechanical treatment

2.6.1 Density matrix

Instead of treating the large number of single spins individually, the theoretical description of an NMR experiment is in general evaluated using a density operator of the form:

$$\hat{\rho} := \sum_k p_k \left| \psi_k \right\rangle \left\langle \psi_k \right| \tag{2.26}$$

where p_k is the relative population of the (ensemble of) state ψ_k. The matrix representation of $\hat{\rho}$ is called density matrix, whose diagonal elements refer to the populations of the involved states whereas the off diagonal elements refer to the coherences between these states. Measurement of the operator $\hat{Q}$ that represents a physical observable Q is given by its expectation value according to:

$$\left\langle \hat{Q} \right\rangle = \mathrm{Tr} \left(\hat{\rho} \hat{Q} \right) \tag{2.27}$$

2.6.2 Spin dynamics

The time evolution of the density operator under the effect of a Hamilton operator $\hat{\mathcal{H}}$ is given by the Louiville-von Neumann equation:

$$\frac{d}{dt}\hat{\rho}(t) = -[H(t), \hat{\rho}(t)] \qquad (2.28)$$

Which can be solved for a time independent $\hat{\mathcal{H}}$ by introducing a propagator $\hat{U}(t) = e^{-i\hat{\mathcal{H}}t}$:

$$\hat{\rho}(t) = \hat{U}(t)\hat{\rho}(0)\hat{U}^{-1}(t) \qquad (2.29)$$

NMR pulse sequences normally render the Hamiltonian time dependent. A useful approach is to assume piecewise constant Hamiltonians to solve equation (2.28). This enables the numerical description of polarisation transfer behaviour (e.g. by using the program GAMMA [20]). The results can be compared to experimental data as used in chapter 7.2.3 to e.g. determine the strength of a coupling.

3 | Chapter 3
Experimental techniques and applied methodology

3.1 Magic-angle spinning
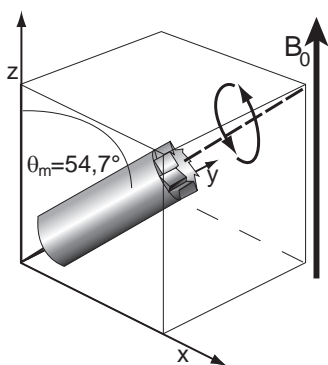
Figure 3.1: Schematic representation of the sample rotation around the axis determined by the space diagonal (MAS). θ_m is known as the magic-angle.

As discussed in chapter 2.3 anisotropic components of the introduced interactions result in a large variety of occurring resonance frequencies for a single spin depending on its orientation relative to the magnetic field. A very useful tool in removing most anisotropic effects is the so called *magic-angle spinning* (MAS) [21, 22], which applies mechanical sample rotation with frequency ω_r around a fixed axis that is oriented along the space diagonal of the laboratory frame (see figure 3.1). The rotation renders the space part of the interaction Hamiltonian time dependent and if ω_r is sufficiently fast, reduces the interaction tensor to its isotropic value. However, interactions that result in resonance frequency offsets which are close to ω_r are not completely cancelled. This effect that can be used to establish specific magnetisation transfer (as presented in [23] and used at several stages

during this work). Additionally spinning sidebands (see figure 3.2) can occur. In principle the latter can be used to obtain information about the shape and magnitude of the underlying interaction (see figure 3.2), here however they are considered as spectral artefacts and ω_r was selected to minimize spectral overlap with unwanted spinning sidebands.
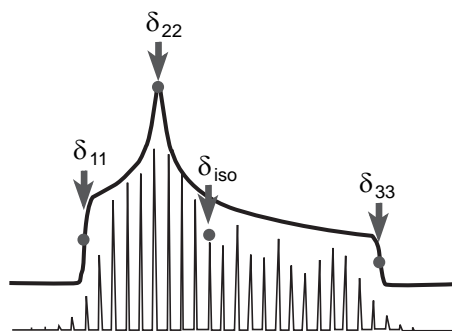
Figure 3.2: Powder pattern spectrum for an asymmetric CSA tensor (simulated data; upper spectrum). MAS leads to spinning sidebands at a difference of ω_r to the isotropic value (simulated data; lower spectrum).

For the samples investigated throughout this work spinning frequencies of $\omega_r \approx 10$ kHz were applied. In general the spinning speed should exceed the strength of the anisotropic interaction by at least a factor of 3 to completely cancel it out. Hence the applied spinning speed is not sufficient to sufficiently reduce the strong ^{1}H-^{1}H or ^{1}H-^{13}C dipolar couplings, which can be significantly stronger[1]. In particular the ^{1}H-^{13}C couplings (20kHz) lead to increase linewidth in ^{13}C detected spectra. Therefore additional heteronuclear proton decoupling pulse sequences [24, 25] were applied using field strength in the order of 80 kHz to effectively reduce the remaining couplings and hence increase the resolution.

3.2 Fourier transform (2D) NMR spectroscopy

The signal detected in NMR arises in general from the transversal components of the net magnetisation, which is given by the superposition of the magnetic moments of all involved spins. The time evolution of the components of this macroscopic quantity can be described using the *Bloch equations* [26]. In this view oscillating net magnetisation of the macroscopic excited state induces a small electromotive force in a detector coil surrounding the sample, which decays over time as the magnetic moments relax to equilibrium. The resulting voltage can be measured as a function of time and is referred to as free induction decay (FID). It contains the information about all excited spins and

---

[1]The strength of the dipolar coupling between two protons at a distance of 1 Å is approximately 120kHz.

Using this value the strength of the dipolar coupling between two nuclei can be approximated easily with knowledge of their gyromagnetic constants and their interatomic distance using equation (2.23).

is in general analysed after Fourier transformation (FT) from time into frequency space.

An important step in NMR was the development of multidimensional spectroscopy [27]. The basic concept of two dimensional (2D) spectroscopy is depicted in figure 3.3a. After a preparation step, which creates a distinct state, an evolution period is added. A mixing step follows during which designated interactions are allowed to correlate the spins to each other. The information is finally read out in the detection step, which consists of the regular acquisition of the FID. Fourier transformation of the direct signal lead to a modulated series of the one dimensional spectra for increasing incremental time steps in the evolution period (figure 3.3b), which also can be Fourier transformed. Hence potential cross peaks in the resulting 2D spectrum (figure 3.3c) are caused by frequency modulations of one spin by its interaction partner(s).
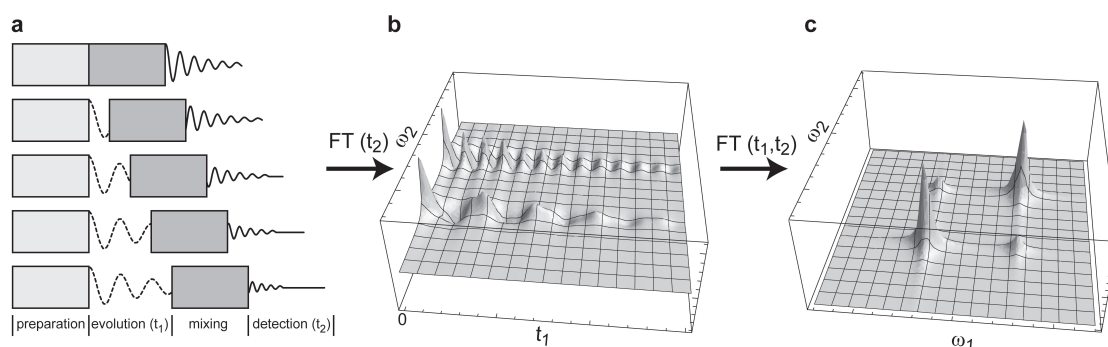
Figure 3.3: (a) Basic elements of 2D NMR spectroscopy. (b) FT of the direct time signal (t_2) lead to oscillating 1D slices. Here two coupled nuclei (of the same kind) with distinct resonance frequencies are shown. (c) 2D spectrum after double FT in t_2 and t_1. Signal of the individual spins result in two diagonal peaks, their coupling in weaker cross peaks.

3.2.1 Signal averaging, referencing and processing

A critical factor in ssNMR is the signal to noise ratio ($\frac{S}{N}$). Signal averaging, i.e. the repetition and summation of the identical experiment, allows to improve $\frac{S}{N}$ since the signal increases linearly with the number of repetitions (n), whereas the noise only

scales proportionally to $\sqrt{n}$. Hence:

$$\frac{S}{N} \propto \frac{n}{\sqrt{n}} = \sqrt{n} \tag{3.1}$$

Notably, equation (3.1) only holds true if the initial conditions for each single experiment are the same. Therefore the system has to relax into equilibrium before the start of the next one. This is typically archived, for the biological samples investigated, after a few (1-3) seconds. Consequently a large number of repetitions steps (as it was necessary at several stages of the presented work) can lead to total experimental times of several days.

For reasons of simplification and a better comparison of data acquired at different magnetic field strength, the detected resonance frequency is normally expressed as the difference to a reference frequency ω_{ref}. For convenience this so called chemical shift (δ) is given in units of parts per million (ppm).

$$\delta(\omega, \omega_{\text{ref}}) := 10^6 \cdot \frac{\omega - \omega_{\text{ref}}}{\omega_{\text{ref}}} \tag{3.2}$$

All experimental data in the presented work are referenced to adamantane (^{13}C) and AGG (^{15}N) as described in [28, 29].

Knowledge of characteristic features of the FID enables the use of so called window functions after acquisition of the data to artificially increase resolution and to decrease noise contributions. The functions used to process the data presented in the following are given in the supporting information on page 156.

3.3 Magnetization transfer schemes

Measurements of distances with atomic resolution and the detection of molecular dynamics on different timescales are probably the most important applications of NMR spectroscopy. In this work both, the measurement of distances and of dynamics is strongly correlated to the detection of spin dipolar couplings. Hereinafter experimental setups and pulse sequences used to measure these interactions will be briefly discussed. However, without knowledge of the precise resonance frequencies of the spins of interest no reliable information can be obtained. Therefore experiments which were used to obtain resonance assignments will be described first.

With the exception of decoupling sequences, in principle all experiments applied during the presented work can be reduced to magnetization transfer or manipulation steps comprising one or a combination of few fundamental r.f. pulse elements. These pulse elements comprise:

- The $\frac{\pi}{2}$- or π-pulse: The simplest form of r.f. interference flips the magnetization by the respective angle. $\frac{\pi}{2}$-pulses are often used to change the magnetization between longitudinal and transversal (as in direct excitation or longitudinal mixing), whereas π-pulses are used here to refocus magnetization.

- Cross polarization (CP): The simultaneous, continuous irradiation with a carrier frequency close to the Larmor frequencies of two different spin species (ω_0^X, ω_0^Y) enables heteronuclear magnetization transfer. This transfer is based on the local dipolar coupling of the involved spins and is achieved by adjusting the field strength of the applied r.f. pulses (ω_1^X, ω_1^Y) according to the Hartman-Hahn condition [30]. Under MAS with spinning speed ω_r this is given by:

$$\omega_1^X \pm \omega_1^Y = n\omega_r, \ n = \pm 1, \pm 2 \tag{3.3}$$

  Useful application of CP steps is the polarization enhancement of spin species with lower gyromagnetic constant [31] and preparation of a desired state (spectral editing) [32].

- Multiple quantum excitation: A large number of different methods for excitation and reconversion of multiple quantum coherences is available. Here the SPC5 [33] sequence was used to generate double quantum (2Q) coherence to measure homonuclear dipolar couplings. The principle is based on symmetry considerations and requires a series of rotor synchronized pulses.

- INEPT (Insensitive Nuclei Enhanced by Polarization Transfer): A combination of $\frac{\pi}{2}$ and π pulses applied simultaneous on two spin species allows heteronuclear magnetization transfer based on J-couplings [34, 35]. Like the CP, it can be used to enhance polarization or for spectral editing.

- TOBSY (total through-bond correlation spectroscopy): Rotor synchronized pulse sequence to establish homonuclear transfer based on J-couplings [36].

3 Experimental techniques and applied methodology

The implementation of the r.f. pulses and parameters used for acquisition of the signal is incorporated into small text based scripts called pulse programs. The pulse programs used throughout this work contain detailed information about the applied pulses and their respective phase information and can be found in the supporting information on pages 144-154.

3.3.1 Experiments for resonance assignments

Knowledge of the resonance frequency of individual spins is in general a prerequisite for a detailed NMR investigation of a molecule. In ssNMR the resonance assignment process is complicated by the relative broad peaks and low sensitivity, which lead to spectral overlap and limits the acquisition of higher order multidimensional spectra. As a result the number of proteins which could at least be partially assigned is still limited [37, 38, 39, 40, 35, 41, 42, 43, 44, 14, 45]. However, in uniformly labeled proteins a general approach exists, which combines intra residue transfer for the identification of the spin system with (specific) inter residue transfer for the connection to sequential neighbors (e.g. [46]). Here intra residue (^{13}C,^{13}C) transfer was obtained using proton driven spin diffusion (PDSD) with short mixing times, double quantum spectroscopy or specific ^{15}N $-$ ^{13}Cα [32] transfer. Sequential transfer was established using specific ^{15}N $-$ ^{13}C' transfer or spin diffusion under weak coupling conditions (SDWC) [23]. Due to sensitivity considerations predominantly 2D spectroscopy was applied.

Secondary chemical shifts

The resonance frequency of specific spins connected to the protein backbone is strongly correlated with the secondary structure of the protein [47, 28]. Hence sequential resonance assignments already contain informations about the local secondary structure of a protein. In this respect the calculation of the so called secondary chemical shifts ($\Delta\delta$), i.e. the difference between the detected resonance frequency (ω_{exp}^X) to the value reported on average for the spin in the specific amino acid as given in the Biological Magnetic

Secondary structure element	expected secondary chemical shift
α-helical	$\Delta\delta_{C\alpha} > 0$
	$\Delta\delta_{C\beta} < 0$
	$\Delta\Delta\delta_{C\alpha,C\beta} < 0$
β-strand	$\Delta\delta_{C\alpha} < 0$
	$\Delta\delta_{C\beta} > 0$
	$\Delta\Delta\delta_{C\alpha,C\beta} < 0$
random coil	$\Delta\delta_{C\alpha} \approx 0$
	$\Delta\delta_{C\beta} \approx 0$
	$\Delta\Delta\delta_{C\alpha,C\beta} \approx 0$

Table 3.1: Statistically expected values for the secondary chemical shift in the different secondary structure elements.

Resonance Bank (BMRB) [48] (ω^X_{BMRB}), turned out to be useful:

$$\Delta\delta \quad = \quad \omega^X_{exp} - \omega^X_{\text{BMRB}} \tag{3.4}$$

$$\Delta\Delta\delta_{C\alpha,C\beta} \quad = \quad \left(\omega^{C\alpha}_{exp} - \omega^{C\alpha}_{\text{BMRB}}\right) - \left(\omega^{C\beta}_{exp} - \omega^{C\beta}_{\text{BMRB}}\right) \tag{3.5}$$

During this work especially the secondary chemical shifts for $^{13}C\alpha$ and $^{13}C\beta$ and their combination ($\Delta\Delta\delta_{C\alpha,C\beta}$) were used to identify the secondary structure elements. Table 3.1 summarizes the underlying correlation.

3.3.2 Spin couplings as probe for spin distances and dynamics

Distance detection

As evident from equation (2.23) the strength of the dipolar coupling strongly depends on the distance between the involved spins. Hence measurement of the dipolar coupling strength permits in principle an accurate detection of inter atomic distances. However in uniformly labeled samples multispin effects complicate the detailed detection of a single distance.

Alternatively one can transfer magnetization uniformly and read out cross correlations between two spins which arise from their couplings. During the initial rate of the mag-

netization transfer, the transfer efficiency strongly depends on the underlying coupling strength. Thus a cross peak appearance indicates close interatomic proximity and hence the selection of the time the spins are allowed to exchange magnetization (mixing time) and/or the evaluation of cross peak amplitudes enable simultaneous detection of a large number of interatomic distances. During this work distances were measured using the CHHC or NHHC experiment [49, 50, 43], which indirectly probe proton-proton distances.

Dynamically reduced dipolar couplings

As described in 2.4.2 the dipolar coupling is anisotropic. Hence its strength depends additionally on the relative orientation of the involved spins in respect to the magnetic field. Thus on the sample average, molecular reorientation on the timescale of the inverse coupling strength reduces the effective coupling.

In chapter 7.2.3 the strength of homonuclear (^{13}C,^{13}C) couplings was investigated. A simple form to describe reduced couplings is the introduction of a factor $s \in [0, 1]$, which is called order parameter and scales the dipolar couplings according to:

$$d_{12}^{ef} = s \cdot d_{12} \tag{3.6}$$

Using equation (3.6) the order parameter for specific spin pairs can be calculated by fitting simulated data to their experimental magnetization buildup curve [51].

Another aspect of reduced dipolar couplings becomes evident in the extreme case of vanishing dipolar couplings due to molecular motion. Under these condition the transversal relaxation rate (T_2) is sufficient long to allow for INEPT based transfer without additional decoupling sequences. In general this is the case for smaller proteins in solution, but it was also detected for terminal ends of membrane embedded proteins [35]. In chapters 6 and 7 it is shown that also membrane exterior loops of larger membrane receptors can be sufficiently mobile to enable INEPT based detection. Hence spectral editing by selection of the transfer pathway determined by local dynamics is possible. The resulting sets of spectra are complimentary and therefore reduce the spectral overlap and assignment possibilities in each of them.

3.3.3 Water edited spectroscopy

Making use of the long T_2 times of protons in the bulk water T_2-filtered spectroscopy [52] enables the possibility to transfer magnetization from the water to proteins (see also chapter 6.2.2). This is especially useful for membrane proteins, since it offers a way to determine the membrane topology of the protein [53]. Based on water edited 2D experiments recorded on protein fibrils [54], here existing pulse sequences were optimized for membrane proteins by implementation of a selective pulse on the water resonance to additionally decrease contributions arising from mobile lipid protons. In general this technique allows generating a 2 dimensional correlation map of the water-protein interface, if the assignment is known. Alternatively, if the topology is partially known it can also assist in the assignment process by excitation of a reduced number of spins.

3.3.4 Reverse labeling

A powerful approach to reduce spectral overlap in proteins was identified to be reverse labeling [55, 56], i.e. the addition of a large amount of a selected set of not labeled amino acids during protein expression. This approach allows removing contribution from specific amino acid types. In chapter 6 it is shown that reverse labeling of high abundant hydrophobic residues as e.g. valine or leucine, significantly resolves spectral crowding in helical membrane proteins.

3.4 Protein structure prediction

Since more than four decades it is known that the three dimensional structure of a protein is inherently linked to its primary sequence [57]. The driving force behind this is the tendency of each system to adopt a configuration which reflects the global minimum of the free-energy function [58]. These two results explicitly define the *de novo* structure prediction problem, that has been a challenging field of research since then.

Today the computational methods remain far from solving this problem and achieving the high-resolution structures available from X-ray crystallography and NMR [59]. However, recent progress in predicting structures of small to medium size proteins us-

ing comparative modeling has shown to be remarkably useful [60, 61, 62] and can be combined with NMR [63, 64] and X-ray data [65] to cross-validate the (experimental or predicted) data and/or to speed up the structure calculation process.

3.4.1 Comparative modeling

Analysis of available experimental determined protein structures has shown that proteins with similar amino acid sequences often also adopt similar tertiary structure. Therefore, the easiest and also the most accurate way to predict a protein structure is to use sequence alignment tools and to build the structure based on sequence relatives [66]. Such an approach is called comparative modeling. Since in most cases the sequence relatives and the target protein belong to the same functional family, i.e. they are homologues of each other, comparative modeling is also referred to as homology modeling.

The web-server *Robetta* [67, 68, 69] and the *Zhang-Server (I-Tasser)* [70, 71, 72] are among the top rated structure prediction servers that are available [60, 70]. During this work both servers were used to develop a structural model of a 118 residue protein of unknown 3D structure. In chapter 8.2.1 the models obtained from the *Robetta* server are discussed in more detail. However, the results obtained from the Zhang server are very similar and support the structural model (see supporting figure on page 111 for a comparison of the structures from both servers). The underlying steps, exemplified for the *Robetta* server, will be briefly discussed in the following[2]. The *Robetta* server is based on the *Rosetta* algorithm [73, 74]. Starting from the primary sequence of the target protein, prior to 3D structural prediction, a domain parsing step using the Ginzu method [67, 75] is applied to identify domain boundaries. The actual process of 3D modeling is based on the Ginzu outcome and can be classified into the following five major steps, which themselves consist of different methods and often rely on the usage of other programs (see table 3.2 for a list of the most important programs incorporated into the Robetta server).

1. **Homologue detection:** Depending on the level of correlation, different existing

---

[2]Note that both servers are fully automated and in principle can be treated by the user as a *black box* allowing for a broad applicability.

programs for homologue scanning are applied to identify homologues (also referred to as parents) from a given database.

2. **Structure alignment:** Robetta applies its own algorithm (k*sync) [73] for alignment of structural features of the parent to the target sequence. The algorithm weights the outcome of established sequence alignment and secondary structure prediction tools and additionally incorporates the occurrence of obligate elements.

3. **Ensemble generation:** For up to five different parents a large ensemble of structures is generated by varying distinct parameters in the scoring function or the programs used. The ensemble is evaluated to find the best matches.

4. **Loop modeling:** Loop regions were modeled using the Rosetta *de novo* fragment-replacement approach.

5. **Molecular dynamics refinement:** In a final step a basic molecular dynamic energy minimization is applied to relax the backbone and side-chain atoms to more favorable energy values.

A detailed list of the energy terms of the Rosetta energy function used in steps 4. and 5. can be found in [76]. If the first step fails in finding a homologue, the Rosetta *de novo* algorithm is applied. The programs/methods applied during the first three individual steps are:

Step of the modeling process	Implemented (external) program / method
Homologue detection	PSI-BLAST [77, 78]
Structure alignment	PSI-BLAST [77, 78]
	StrAD-Stack [73] (based on 3D-Pair [79]
	and Taylor's method [80])
	PsiPred [81]; DSSP [82]
Ensemble generation	all above plus:
	SAM-T99 [83, 84] and JUFO [85]

Table 3.2: Overview of individual (external) components implemented on the Robetta server.

Chapter 4

4 An atomic picture of insoluble protein folding intermediates

4.1 Summary

This chapter demonstrates how two-dimensional solid-state NMR can be used to investigate structure and dynamics of insoluble folding states. The study was carried out on the model system *catabolite repression histidine-containing phosphocarrier protein* (Crh) which is known to form domain-swapped dimers [86]. Starting from the protein precipitated at its pI, conformational changes due to a modest temperature increase were investigated at the level of individual residues and in real-time. As compared to the crystalline state, Crh pI-precipitates exhibited a higher degree of molecular mobility for several regions of the protein. A rigidly intact center was observed including a subset of residues of the hydrophobic core. Raising the temperature by 13 K to 282 K created a partially unfolded intermediate state that was converted into β-sheet-rich aggregates that are mostly of spherical character according to electron microscopy. Residue-by-residue analysis indicated that two out of three helices in aggregated Crh underwent major structural rearrangements while the third helix was preserved. Residues in the hinge region exhibited major chemical-shift changes, indicating that the domain swap was not conserved in the aggregated form. This study provides direct evidence that

protein aggregates of a domain-swapped protein retain a significant fraction of native secondary structure and demonstrates that solid-state NMR can be used to directly monitor slow molecular folding events.

4.2 Introduction

Understanding the ways in which proteins or protein fragments self-assemble into ordered, insoluble aggregates associated with amyloid diseases is of paramount biophysical and medical interest[12]. Delineating the mechanisms of protein folding and aggregation and their possible control by pharmacological intervention requires the identification of key species involved in folding and aggregation and the precise definition of their roles in each pathway. The final, and possibly also earlier, intermediate protein states are often insoluble. This complicates the application of well-established techniques to study biomolecules at atomic resolution, that is, solution-state NMR and X-ray crystallography. Solid-state NMR has recently made significant progress in studying nonnative protein states [87, 88, 89, 56, 90] and protein folding [91, 92, 93] at the molecular level. Indeed, ssNMR spectroscopy can be performed in fully fibrillized or in freeze-trapped biomolecules; here it is shown that ssNMR also offers a spectroscopic means to follow the underlying protein folding events in real-time. While NMR has already successfully been applied for protein folding in solution [94], the following chapter demonstrates how high-resolution ssNMR methods can be used to study conformational transitions between insoluble native and non-native states of a protein.

This study focuses on the Crh protein from *B. Subtilis*. While its specific function remains elusive [86], Crh has been shown to exist in remarkably different conformations. In solution, Crh forms a mixture of monomers and dimers in a slowly (i.e. time scale of hours) exchanging equilibrium [95]. The monomer structure was resolved by solution-state NMR [96]. X-ray crystallography revealed a 3D domain-swapped dimer structure formed by β1-strand swapping of two monomers [97]. Domain swapping has been forwarded as a mechanism for protein oligomerization [98, 99, 100, 101, 102], possibly involving the native protein state [102], and has been considered as a possible means for functional regulation [101, 103, 104]. It may also be closely associated with misfolding

and aggregation, as has been suggested in the context of human prion protein [99] and amyloidogenic human cystatin C [105]. A variety of biophysical methods, computational studies [106, 107, 108] and solution-state NMR [109], as well as mutagenesis [110] have been employed to elucidate structural details associated with folding of a monomeric protein into a domain-swapped form. However, direct structural information about insoluble intermediate and fully aggregated folding states and their respective lifetimes has been lacking.

Previously, it was shown by ssNMR that Crh domain-swapped dimers are already present in the microcrystalline state [37, 111]. Here, the spectroscopic investigation of conformational states obtained by precipitation of Crh at a pH of 6.5 corresponding to its isoelectric point, at which its solubility is minimal, is described. In the following this preparation is refered to as pI-precipitated Crh.

ssNMR data obtained on the precipitate are consistent with an overall domain-swapped structure in which major protein segments are destabilized and only part of the hydrophobic core is rigidly intact. It is shown that a folding event can be induced by a modest temperature increase, which is monitored by 2D ssNMR in real-time. Analysis of the corresponding 2D ssNMR correlation spectra acquired before, during, and after the folding event is consistent with a folding pathway starting at a dynamically destabilized domain-swapped dimer. Upon temperature increase, a partially unfolded intermediate state was created that directly led to protein aggregates not only rich in β-sheet secondary structure but also preserving significant α-helical segments. This study

- reveals that molecular dynamics may play an important role in protein aggregation,

- provides direct evidence for a partially unfolded intermediate state in a domain-swapped dimer, and

- demonstrates the possibility to structurally analyze aggregated protein states in a site-resolved manner.
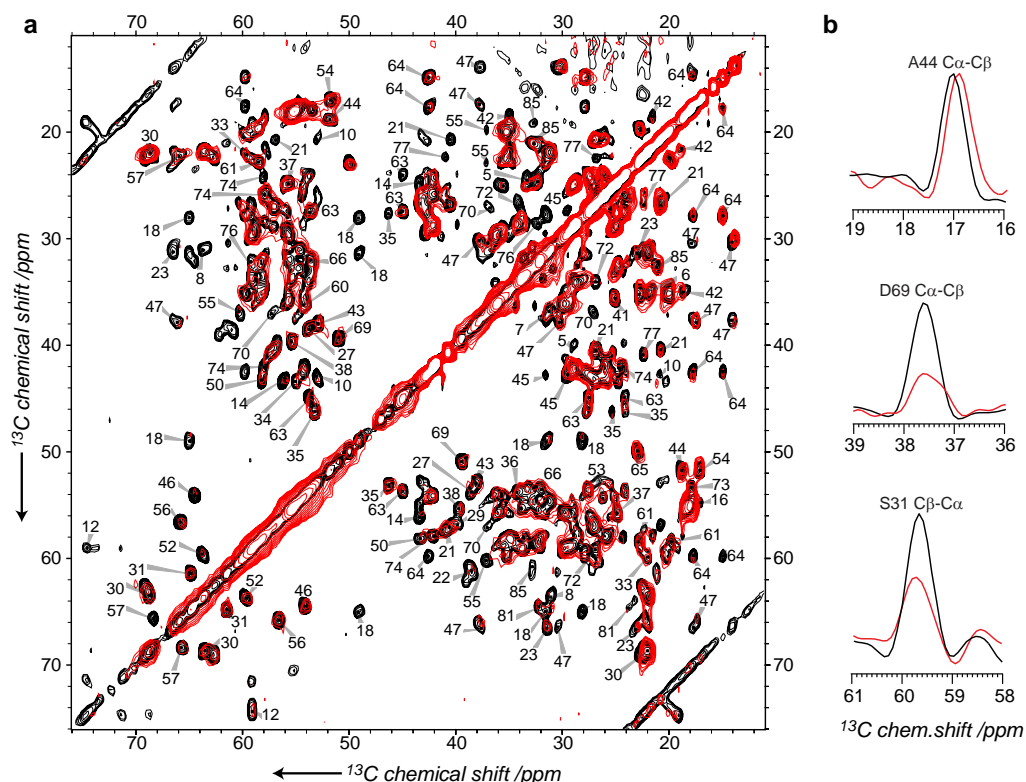
Figure 4.1: Comparison of 2D ^{13}C-^{13}C PDSD spectra (aliphatic region) obtained on microcrystalline (black) and pI-precipitated (red) Crh (a). Extracts shown in (b) exemplify variations in Cα-Cβ cross-peak amplitudes for the indicated residues. Both spectra were taken under identical experimental conditions. Assignments are according to ref [37].

4.3 Results

4.3.1 Comparison of microcrystalline/pI-precipitated Crh

In figure 4.1, a homonuclear (^{13}C,^{13}C) proton driven spin diffusion (PDSD) spectra of Crh pI-precipitate at 269 K (red) is compared to results using a microcrystalline prepa-ration [37] (black). The spectra clearly show that the general chemical-shift pattern is conserved and ^{13}C line widths are comparable for most signals. Similar observations were made for 2D ^{13}C-^{15}N heteronuclear correlation spectra (see supporting figure on page 112). Even without sequential assignments for the precipitate, chemical-shift changes can be traced for most residues, and no new cross peaks are observed. The domain swap

involving β-strand 1 can still be identified by the characteristic chemical shift of Thr12 located in the hinge region (figure 4.1a). The high similarity of chemical shifts observed between the microcrystalline Crh and the precipitated form indicates that the global fold of Crh is conserved in the precipitate.

The most remarkable difference between the spectra of microcrystalline and precipitated Crh lies in the signal intensities. As exemplified in figure 4.1b, the ssNMR line width is comparable, for most cases, even when signal intensities are largely attenuated. For these residues, static disorder (leading to inhomogenous broadening) is not the source of the observed signal loss, and variations in signal intensities can be attributed to variations in molecular mobility. For residues with resolved cross signals, the relative intensity ratio I_m/I_p of ssNMR signals obtained on microcrystalline and pI-precipitated Crh was determined (see figure 4.1 and supporting figure on page 113). High ratios indicate a higher degree of local mobility in the pI-precipitated sample, which is a priori not constrained by a crystal lattice. These parameters are mapped on the ribbon diagram of the domain-swapped crystal structure of Crh in figure 4.2a (PDB code 1mu4 [97]). The degree of signal attenuation ranges from green (rigid, $I_m/I_p \approx 1$) to red (strongest dynamics, $I_m/I_p > 15$). Residues not identified are given in white. In figure 4.2b, assigned protein residues exhibiting sizable dynamics ($I_m/I_p > 5$ in a) on the backbone (purple) or side-chain (pink) level are shown in surface mode. The attenuation of hydrophilic side-chain signals from residues located at the surface comes without surprise. Interestingly, many of the backbone signals, and also signals from residues being part of the hydrophobic core of Crh, are affected. Moreover, significant cross signal attenuation is observed around the active site (residues Lys45, Ser46, and Ile47) and for several residues in helix A (residues Ala16, Pro18, Leu21, Val23). The most affected secondary structure element is, however, the β-strand 1a, which forms the hinge region in the domain-swapped dimer (residues Thr12-Leu14). Another region where dynamic variations concentrate is the end of β-strand 4 and helix C are able to undergo important structural changes [112]. Indeed, a destabilization in this part of the protein in the Hpr F29W mutant leads to a domain-swapped form of HPr, involving the C-terminal part of the protein, including β-strand 4 and helix C. The Asp69 residue, equally conserved in Crh, and located in the loop between strand 4 and helix C, plays a central role in the

folding mechanism of Hpr [113]. In Crh, the Asp69 side chain is involved in hydrogen bonds with the amide and hydroxyl protons of residues Thr30 and Ser31, as well as with Ser31 Hγ. Both Asp69 and Ser31 residues show attenuated cross-peak intensities, indicating a destabilization of this important contact.

In figure 4.2b, the empty pocket in the center of each monomer reveals a dynamically stable subset of the hydrophobic core illustrated in cyan in figure 4.2c. Among these residues, only two are found on the N- or C-terminal part of the protein. On the other hand, residues forming part of the classical hydrophobic core and situated in the N-terminus, including Val8, Leu10, and Leu14, or in the C-terminal β-strand 4 and helix C, like Leu63, Ala73, Leu74, and Val81, are not part of this rigid core. Notably, these observations are in line with Eisenberg et al. [114], who postulated that any protein can swap domains under appropriate conditions, as long as the terminal domain of the protein is unconstrained.

4.3.2 Investigation of the aggregation process of Crh pI-precipitates

To further analyze the conformational space adopted by a domain-swapped protein, 2D ssNMR spectra immediately after a temperature rise from 269 to 282 K were recorded. Figure 4.3 shows spectra of the U[^{13}C,^{15}N] Crh precipitate recorded before (red), after the temperature increase (black), and 1 day afterward (blue). Interestingly, chemical-shift differences can be detected in several regions of the ssNMR spectrum, which can be most easily followed for the $C\alpha/C\beta$ spectral region involving Thr and Ser residues. As visible from the extracted region in figure 4.3b, degenerate Thr $C\beta/C\gamma2$ cross signal around 69/22 ppm are shifting along the ω_2 dimension to lower field, resulting in a significant chemical-shift difference of about 2.5 ppm for the Thr $C\beta$ atoms. Even stronger variations are seen for $C\alpha/C\gamma$ correlations that exhibit resonance shifts of 5 ppm. The Thr cross-peaks around 70 ppm present in the 2D spectrum recorded immediately after the temperature increase (black) clearly indicate that not only the initial and final folding states are present in the spectrum. The corresponding partially resolved spectral "footprints" strongly suggest the presence of at least one transition state that differs from the initial and final state and exhibits a lifetime comparable to
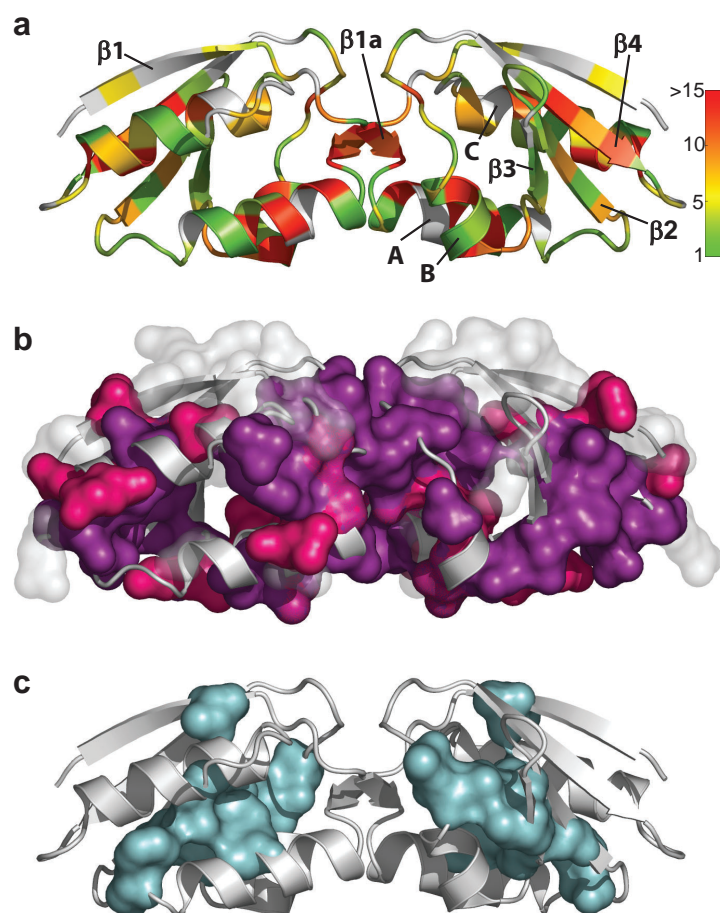
Figure 4.2: (a) X-ray structure of the Crh domain-swapped dimer (PDB code 1mu4C [97]) with color gradients corresponding to the observed cross signal attenuation in spectra of pI-precipitated Crh as compared to the spectrum of the microcrystalline sample, ranging from strong ($I_m/I_p > 15$, red) to no ($I_m/I_p = 15$, green) attenuation. Not observed: white. The largest attenuation detected for a resonance belonging to a specific residue is reported. Helices A-C and β-strands $\beta 1$-$\beta 4$ are indicated. In (b), residues that exhibit attenuated cross signals ($I_m/I_p > 5$) for resonances involving ^{13}C'-^{13}Cα and ^{13}Cα-^{13}Cβ cross signals are shown in (surface-mode) purple, and cross signals involving side-chain carbons are in pink. Residues for which no information could be obtained due to signal overlap are colored in transparent gray. In (c), the static residues of the hydrophobic core are shown in cyan. All other protein segments are given in white.

a significant fraction of the total experiment time (see chapter 5 for details). Indeed, protein conformations and folding intermediates different from the initial state will only

be detectable in real-time if their resonances are distinct from those of the initial state and if the rate constants for their formation and disappearance result in a significant population of a given species on the time scale of our ssNMR experiment. In this case, solid-state NMR signals from the different species in the frequency domain (F1) spectrum after double Fourier transformation will be convoluted by their kinetic profile in the folding reaction. These findings are corroborated by additional $C\alpha/C\beta$ cross-peak signals for Ser, Thr, and Pro, which exhibit chemical shifts typically associated with coil conformations in soluble proteins. (A detailed analysis of the real-time folding kintetics can be found in chapter 5.)
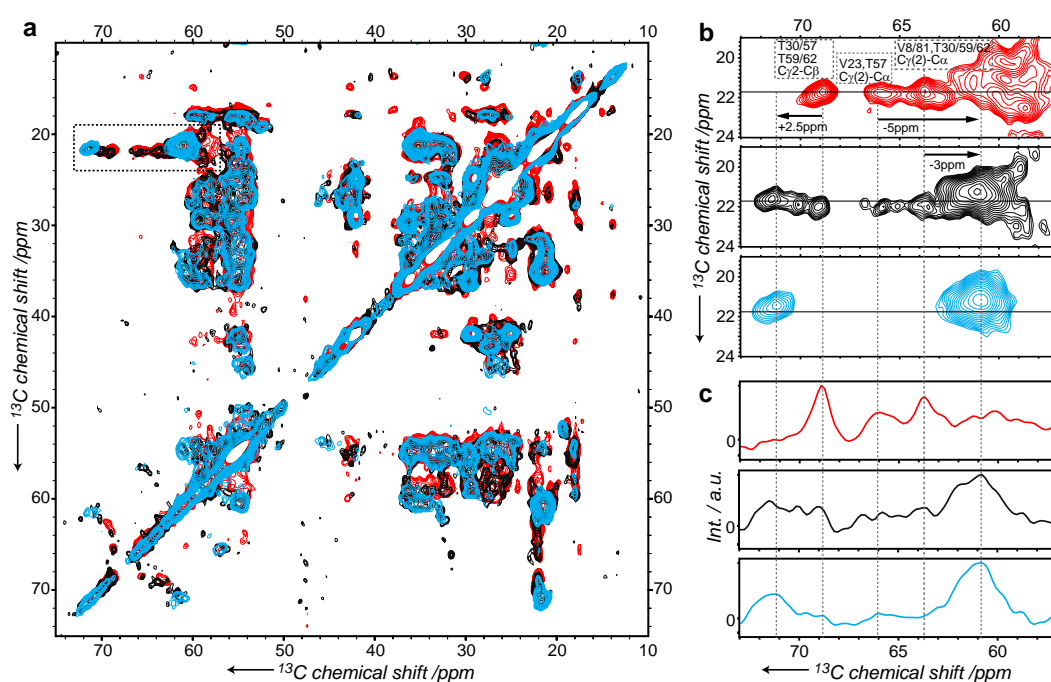
Figure 4.3: Time-resolved 2D ssNMR experiments.(a) Overlay of three spectra: red, pI-precipitated Crh (T = 269 K); black, recorded immediately after a temperature increase to T = 282 K; blue, 1 day after temperature raise (T = 282 K). Spectra of the highlighted region are shown separately in (b). In (c), cross-peak intensities of the frequency indicated by horizontal lines (in (b)) are plotted. The spectra were recorded at 800 MHz, at a spinning frequency of 12.5 kHz. Magnetization transfer was achieved using spin diffusion under weak-coupling conditions [23] with a mixing time of 150 ms.

In the final state (third panel, figure 4.3b), maximum intensity is found for the Thr $C\beta$ residues around 71.5 ppm. The signals corresponding to the Val8 and Val23 $C\alpha/C\gamma$,

and also to the overlapping Thr30/59/62 $C\alpha/C\gamma2$ and Thr57 $C\alpha/C\gamma2$, respectively, are shifted upfield by about 3 and 5 ppm. Only very moderate intensity remains around 66 ppm, as shown in figure 4.3c, panel three. The observed chemical-shift changes remained the same in subsequent spectra recorded several weeks later (see, e.g. figure 4.4 for a spectrum recorded 3 weeks after the transition). An upfield shift in $C\alpha$ chemical shifts, combined with a downfield shift in $C\beta$ hemical shifts, is indicative of an increase in β-strand conformation after the temperature increase. The corresponding Crh form was investigated by electron microscopy, which reveals mostly spherical aggregates (vide infra). Such aggregates have also been observed for precursor forms of amyloid-like aggregates [115].
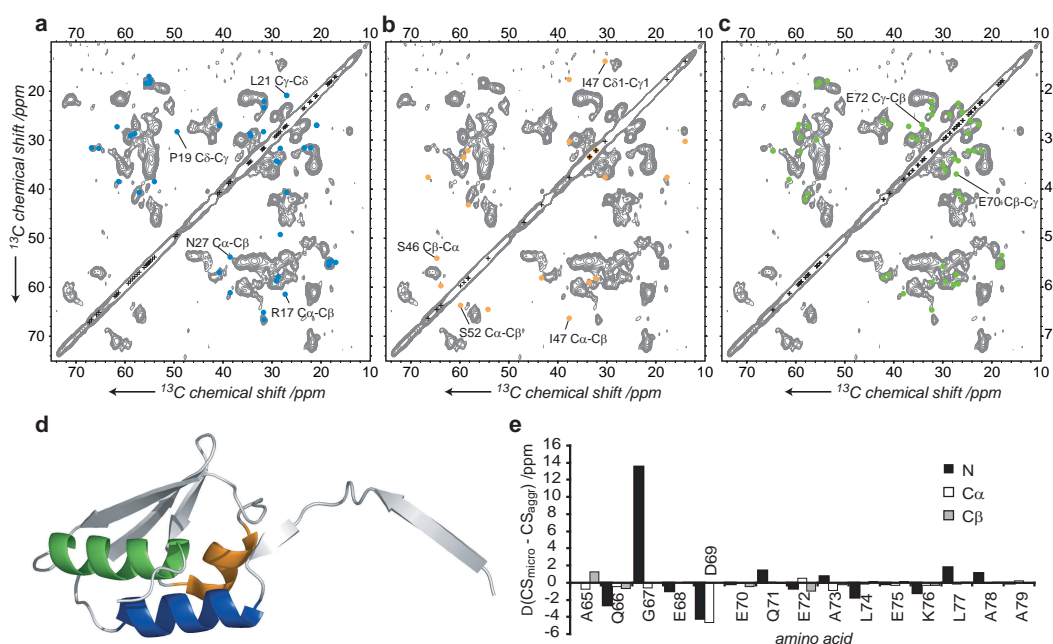
Figure 4.4: 2D $^{13}C - ^{13}C$ PDSD spectrum of aggregated Crh. The spectrum was recorded at 600 MHz ^{1}H frequency, at a spinning speed of 10.5 kHz and at 274 K. PDSD with a mixing time of 20 ms was used for magnetization transfer. In (a), one-bond correlations for helix A assuming a microcrystalline dimer structure are shown in blue; in (b) and (c), the corresponding signal sets are given for helices B and C in yellow and green, respectively. (d) Location of the three helices A (blue), B (orange), and C (green) in one Crh monomer as seen in the crystal structure [97]. (e) Chemical-shift changes $\Delta(CS_{micro} - CS_{aggr})$ observed in NCACX and NCOCX spectra (relative to ref[37]) for the sequential walk concerning residues 65-79. Note that helix C starts at residue E70.

4.3.3 Structural analysis of the aggregated form

Although the ^{13}C line width seen for correlations in the aggregated state is comparable to spectra recorded before the temperature-induced conformational transition, sequential resonance assignments are difficult to obtain for the final state. Instead, the chemical shifts of the individual residues in microcrystalline Crh were compared to data obtained on the aggregated form. Disappearing signals indicate either increased dynamics or conformational changes. Speaking in favor for the later, additional $C\alpha/C\beta$ cross signals with typical β-strand chemical shifts as well as a general increase in β-strand conformation was observed. Some α-helical secondary structure is, however, still present, as exemplified by the occurrence of α-helical chemical shifts, clearly observable for Ala residues (see below). Because the induced temperature increase is moderate, it is assumed that these shifts reflect α-helical protein segments already present before the transition. Correspondingly, the ssNMR analysis of the microcrystalline Crh sample could be used as a reference for determination of α-helical segments in the aggregate. In particular, it is examined which α-helices are conserved in the aggregated form of the protein. Figure 4.4a-c shows the PDSD spectrum of the aggregated form, as compared to resonance assignments obtained for the microcrystalline protein for helix A (figure 4.4a, blue), helix B (figure 4.4b, orange), and helix C (figure 4.4c, green), respectively.

For 33% of residues in helix A (Ala16-Asn27) and 43% residues in helix B (Ser46-Ser52), at least one of the corresponding signals is no longer observed in Crh aggregates. On the contrary, all signals could be identified for helix C (residues Glu70-Gln82), with the exception of only two side-chain resonances. The high conservation of cross signals from helix C is also confirmed in NCACX and NCOCACX spectra (see supporting figure on page 114). Figure 4.4e shows the chemical-shift changes between aggregated and microcrystalline Crh identified by a sequential walk from Ala65 to Ala79. These changes are in general small except for Gly67, which exhibits an upfield ^{15}N shift of more than 10 ppm. In the domain-swapped Crh fold, Gly67 NH forms a hydrogen bond to Glu70 $O\epsilon$, stabilizing the turn between β-strand 4 and helix C. In addition, substantial chemical-shift changes were seen for Asp69, which forms important hydrogen bonds to residues 30 and 31, as discussed above. These findings indicate that, while helix C remains

largely intact upon aggregation, the 3D fold of the protein observed in its crystalline or pI-precipitated forms is no longer conserved in the Crh aggregates.

In addition, an overall reduction in helix content of Crh aggregates is reflected in the α-helical Ala signal intensities that were reduced from 70% in Crh microcrystals to 40% in the aggregate. Concomitantly, an increase in β-sheet and random-coil content of 20% and 10%, respectively is observed(see supporting figure on page 115). This increase in overall β-sheet content complicated an investigation analogous to the study regarding the three α-helices for the β- strand region of Crh aggregate. Nevertheless, it is found that at least parts of β-strands β3 and β4 of Crh were preserved. Both segments, along with helix C, are indicated in figure 4.5 on the Crh X-ray structure in blue (lower panel). Interestingly, these segments are all located far from the hinge region. In addition, figure 4.5 shows in red protein residues that exhibit chemical-shift alterations, meaning that the X-ray structure shown in the figure is no longer conserved in these regions upon aggregation. Remarkably, residues with isolated chemical shifts and located in loops or turns, such as Thr30, Ser31, Asp38, Ala54, Thr57/59, and Asp69, were clearly no longer observed at their initial chemical shift, indicating that the three-dimensional organization is profoundly altered, even of the conserved secondary structure elements. Finally, the outlier chemical shift for Thr12 $C\alpha/C\beta$, which is indicative for dimer formation, was not detected in the spectra of the aggregated form. This supports the conclusion that the conformation of the hinge region is substantially altered in the aggregated form and that protein regions essential for stabilization of the 3D fold of the Crh domain-swapped dimer have undergone major conformational changes.

4.4 Discussion

It was shown that ssNMR allows one to study the structure and identify the dynamics in a site-resolved manner in insoluble folding states. Hence, ssNMR not only allows to follow conformational transitions at specific molecular sites [116] but, as demonstrated here, also enables the recording of time-dependent conformational changes in real time and at the level of individual residues in fully labeled proteins. Our data reveal that Crh pI-precipitates exhibit a higher degree of molecular mobility as compared to the
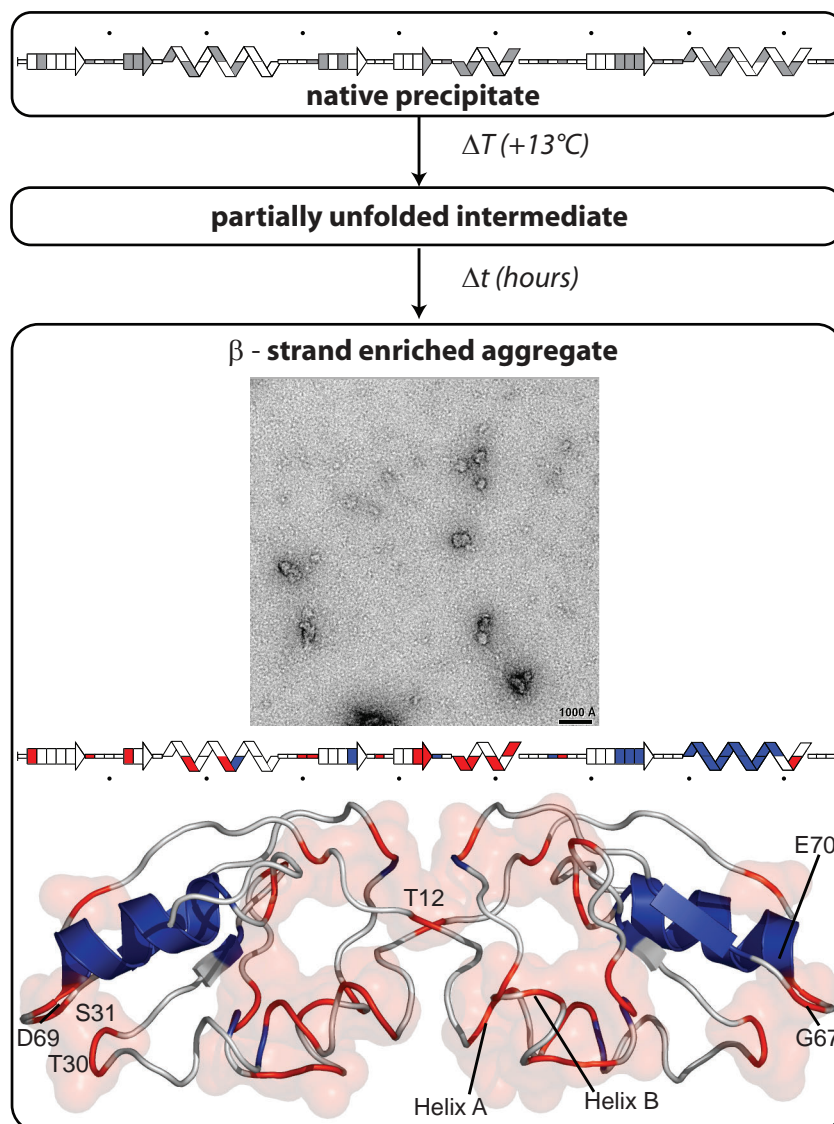
Figure 4.5: Comparison of the results of the spectroscopic analysis of Crh pI-precipitates before and after temperature increase to those obtained for the Crh aggregates. In the upper panel, increased protein dynamics in the native pI-precipitate are indicated in gray according to figure 4.2a. For Crh aggregates (lower panel, EM micrograph), red color-coding refers to chemical-shift changes and/or increased molecular dynamics. Secondary-structure elements, which can still be identified in the aggregate, are given in blue.

crystalline state, with a core that is rigidly intact. Destabilization of this conformation by a moderate temperature increase was directly followed by 2D ssNMR. Accordingly,

a partially unfolded intermediate state was created, which led to β-strand rich protein aggregates containing a remarkable fraction of α-helical segments.

The presented observations confirm the previously postulated [117] presence of intermediate states for domain-swapped proteins. Crh aggregates do not contain native domain-swapped proteins, which speaks against a run-away domain swap [118, 119] as a mechanism for oligomer formation in this protein. The conversion of a destabilized native protein state to a β-sheet rich aggregate was previously seen for a fibril-forming protein [120], where the aggregation process occurred without the need for dissolution and renucleation of the aggregates. This is compatible with the here observed "solid" intermediate states during the conformational transition in Crh, where aggregates are formed by a reorganization of the initial pI-precipitates, possibly assisted by the higher degree of flexibility observed in the precipitate. Dynamic behavior was also observed in a native-like fibril precursor of a naturally amyloidogenic protein [121], and the increased flexibility was linked to the interconversion of the native state and a folding intermediate.

Recently, ssNMR methods have shown their immense potential to study 3D molecular structure and dynamics (see, e.g. refs [122, 123, 124]) using a single isotope-labeled sample. Application of these methods will offer additional means to characterize structure and dynamics and in combination with the results presented here should be applied to study protein refolding events that take place on a slow time scale and in dense molecular systems, as observed in vitro and in vivo for many neurodegenerative diseases.

4.5 Material and methods

4.5.1 Sample Preparation

Microcrystalline Crh was prepared as described previously [37]. The pI-precipitated Crh sample was obtained by changing the pH of a protein solution (20 mg/mL) to the isoelectric pH (pI = 6.5) by addition of HCl 1 M. Precipitation occurred instantaneously, and the sample was left in the freezer for 5 days to complete the process. The precipitate was transferred into a 4 mm CRAMPS rotor, and the rotor cap was sealed.

4.5.2 Electron Microscopy (EM)

Crh aggregates were diluted 1:50 with NH_4CO_3 buffer and prepared on a glow discharged carbon foil. After being stained with 1% uranyl acetate, the samples were evaluated with a CM 200 transmission electron microscope (FEI, Eindhoven, and The Netherlands). Pictures were taken with a TemCam F415A camera (TVIPS, Gauting, Germany) at 20 000-fold magnification.

4.5.3 Solid-State NMR Spectroscopy and Data Analysis

All NMR experiments were conduced using 4 mm triple-resonance probeheads at static magnetic fields of 18.8 and 14.1 T corresponding to 800 and 600 MHz ^{1}H resonance frequency (Bruker Biospin, Karlsruhe/Germany). Broadband (^{1}H,[13]C) and chemical-shift selective (^{15}N,[13]C) Hartmann-Hahn cross polarization (SPECIFIC-CP [32]) steps were applied using SPINAL64 [25] proton decoupling with radio frequency fields of 75-90 kHz during evolution (t_1), detection (t_2), and N-C transfer. For intra-residue C-C transfer, PDSD (proton-driven spin diffusion) with a mixing time of 20 ms was used. Sequential ($Cx_i - Cy_{i\pm1}$) magnetization transfer was achieved using spin diffusion under weak coupling conditions [23] with mixing times between 80 and 150 ms. All spectra were processed using QSINE window functions in F1 and F2 and analyzed with Sparky version 3.110 (T. D. Goddard, D. G. Kneller, University of California). Protein structures were visualized using Pymol (DeLano Scientific, San Carlos, CA).

Chapter 5

5 Kinetic analysis of protein aggregation monitored by real-time ssNMR

5.1 Introduction

Protein aggregation has become recognized as an important aspect of the protein folding landscape. This process not only interfere with protein expression and recovery assays used in biotechnology, but also impacts the every day live of cells and organisms and is associated with a variety of human diseases [125]. For the latter it is not exactly clear how abnormal protein aggregation can lead to cellular dysfunction nor which structural species is involved in the pathology of these so called amyloid diseases [126]. Recent results indicate that in particular the early oligomeric intermediates have significant toxicity [127]. A detailed understanding of the kinetics and thermodynamics involved in protein aggregation has hence been proposed to be a key step for the development of useful therapeutic treatments [11].

The methods to characterize aggregation kinetics so far comprise predominantly biochemical and biophysical approaches such as gel filtration, sedimentation assays, binding of fluorescent markers, AFM imaging, dynamic light scattering or circular dichroism that

report on the oligomerisation state or the overall secondary structure content of the protein [126, 125]. Real-time solution-state NMR has shown to be useful to follow protein folding at the level of individual residues [128, 129]. However since larger oligeromeric species escape detection, a kinetic analysis of protein aggregation is largely restricted to the observation of the disappearance of NMR signal in solution. In principal solid-state NMR (ssNMR) offers the complementary spectroscopic means and time-resolved 1D ssNMR has been used to monitor the buildup of signal during protein aggregation [116]. Due to the limited spectral resolution 1D ssNMR has to be combined with specific labeling techniques to offer site-specific resolution.

In this chapter it is shown how real-time 2D solid-state NMR can contribute to the kinetic and thermodynamic understanding of protein aggregation by offering site-resolved information on uniformly labeled proteins. The presented analysis is based on experimental results obtained on the Crh protein as presented in chapter 4.3.2.

5.2 Results and discussion

5.2.1 Initial considerations

In NMR spectroscopy the detected signal is in general a superposition of resonance frequencies of spins that experience a different local environment. In a molecule that occurs in two different conformations A and B (e.g. unfolded and folded), a single spin might have two distinct resonance frequencies (ω_A and ω_B). If the populations of the two conformations (P_A and P_B) on the sample average are conserved during the experimental time the signal in time space $s(t)$ and in frequency space $S(\omega)$ is given by:

$$s(t) \quad = \quad a_A \cdot e^{(2\pi i \omega_A - \lambda_A)t} + a_B \cdot e^{(2\pi i \omega_B - \lambda_B)t} \tag{5.1}$$

$$S(\omega) \quad = \quad \int_0^\infty \left[a_A \cdot e^{(2\pi i \omega_A - \lambda_A)t} + a_B \cdot e^{(2\pi i \omega_B - \lambda_B)t} \right] \cdot e^{-2\pi i \omega t} dt \tag{5.2}$$

The conventional absorbtion NMR spectrum is the real part of equation (5.2), which is the sum of two Lorentzian functions $L_n(\omega)$ in frequency space where $\lambda_{A,B}$ is determined

by the transversal relaxation time of the spin in the given conformation (see figure 5.1):

$$\mathrm{Re}\left(S(\omega)\right) = \sum_n a_n L_n \qquad n = A, B \qquad (5.3)$$

$$L_n(\omega) = \frac{\lambda_n}{\lambda_n^2 + 4\pi^2 \left(\omega - \omega_n\right)^2} \qquad (5.4)$$

In this static case the peak volume reports on the relative population of the two conformations (in the following referred to as state A and state B).
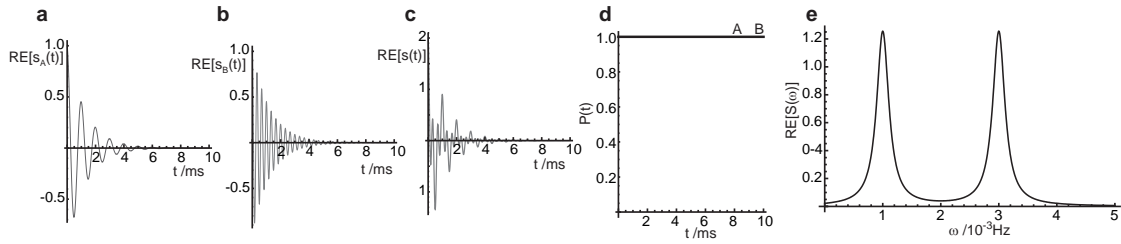

Figure 5.1: One spin in two different conformations. (a) $\mathrm{Re}(s_A(t))$, (b) $\mathrm{Re}(s_B(t))$, (c) $\mathrm{Re}(s(t))$ and (d) $\mathrm{Re}(S(\omega))$ Parameters used are: $a_A = a_B = 1$; $\omega_A = 1 \cdot 10^{-3}$Hz ; $\omega_B = 3 \cdot 10^{-3}$Hz ; $\lambda_A = \lambda_B = 0.8 \cdot 10^{-3}$s^{-1}.

Two-state transition

If at first only state A is present and a transition of state A to state B is initiated at the beginning of the experiment, equation (5.2) has to be modified to:

$$\begin{aligned}
S(\omega) &= \int_0^\infty [a_A \cdot e^{(-kt)} \cdot e^{(2\pi i\omega_A - \lambda_A)t} \\
&\quad + a_B \cdot \left(1 - e^{(-kt)}\right) \cdot e^{(2\pi i\omega_B - \lambda_B)t}] \cdot e^{-2\pi i\omega t} dt
\end{aligned} \qquad (5.5)$$

Here a transition of the form: $P_A(t) = e^{-kt}$ and $P_B(t) = 1 - e^{-kt}$, was assumed. The decrease of P_A is mathematical equivalent to a faster decay of the signal and leads to a peak broadening after Fourier transformation (FT), whereas the buildup of signal for state B results in a baseline distortion as shown in figure 5.2. It is evident that the peak intensity and shape in real-time NMR strongly depends on the rate constant k of the transition. Fitting simulated spectra to the experimental data hence offers a way to determine k [128, 130].
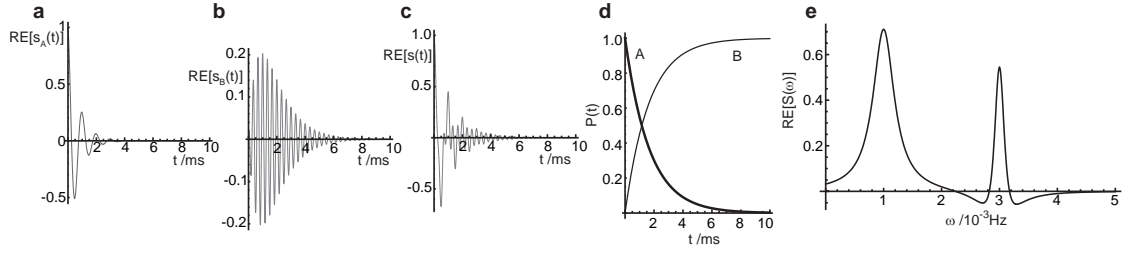
Figure 5.2: Same as in figure 5.1, but assuming a two-state transition of the form given in equation (5.5).

Three-state transition

If state B is not the final but an intermediate state that directly undergoes a transition to a state C equation equation (5.5) for the resulting three state transition $A \to B \to C$ reads as follows:

$$
\begin{aligned}
S(\omega) \quad = \quad & \int_0^\infty [a_A \cdot e^{(-k_1 t)} \cdot e^{(2\pi i \omega_A - \lambda_A)t} + a_B \cdot \left(1 - e^{(-k_1 t)}\right) e^{(-k_2 t)} \cdot e^{(2\pi i \omega_B - \lambda_B)t} \\
& + a_C \cdot \left(1 - e^{(-k_1 t)}\right)\left(1 - e^{(-k_2 t)}\right) \cdot e^{(2\pi i \omega_C - \lambda_C)t_2}] \cdot e^{-2\pi i \omega t} dt
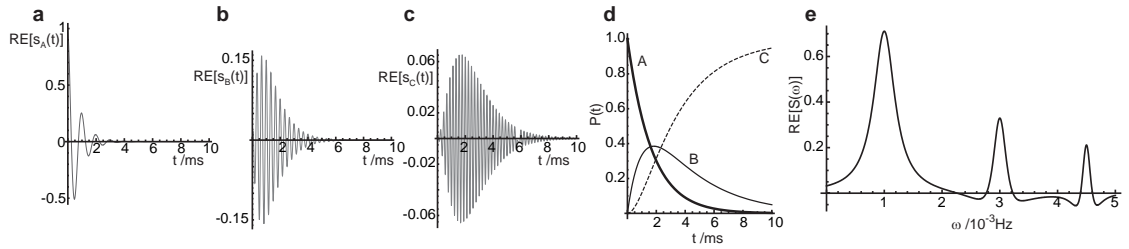\end{aligned}
\qquad (5.6)
$$

Figure 5.3: Same as in figure 5.1, but assuming a three-state transition of the form given in equation (5.6).

Three-state transition in 2D spectroscopy

Under normal circumstances time frame, resolution and sensitivity offered by a single 1D NMR experiment are insufficient for the investigation of protein aggregation. A

series of 1D spectra with sufficient repetition steps recorded during the transition (time-resolved NMR spectroscopy) can increase sensitivity and might offer an appropriate time frame, however as discussed above site resolved resolution or even the detection of an intermediate species is difficult to achieve. Instead multidimensional spectroscopy can be used to follow the transition during the evolution of an indirect dimension (real-time NMR spectroscopy). Under the assumption that the transition during repetitions of a single time step is negligible small, the spectrum after Fourier transformation in the direct dimension remains unaffected, whereas the peak shape in the indirect dimension is modified as discussed above. Figure 5.4 shows such a theoretical three state transition detected in real time during the acquisition of a 2D spectrum after FT in the direct (t_2) and indirect (t_1) dimension. The underlying equation reads as:

$$
\begin{aligned}
S(\omega_1, \omega_2) &= \int_0^\infty [a_A \cdot e^{(-k_1 t_1)} \cdot e^{\left(2\pi i \omega_{A,1} - \lambda_{A,1}\right) t_1} \cdot L_A \\
&+ a_B \cdot \left(1 - e^{(-k_1 t_1)}\right) e^{(-k_2 t_1)} \cdot e^{\left(2\pi i \omega_{B,1} - \lambda_{B,1}\right) t_1} \cdot L_B \\
&+ a_C \cdot \left(1 - e^{(-k_1 t_1)}\right) \left(1 - e^{(-k_2 t_1)}\right) \cdot e^{\left(2\pi i \omega_{C,1} - \lambda_{C,1}\right) t_1} \cdot L_C] \\
&\cdot e^{-2\pi i \omega_1 t_1} dt_1
\end{aligned}
\tag{5.7}
$$

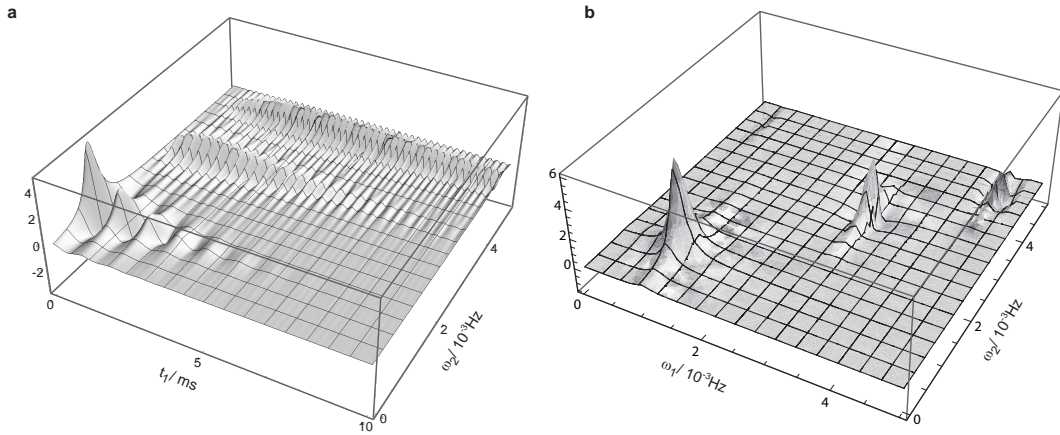Note that equation equation (5.7) already includes the Fourier transformtion in the

Figure 5.4: Three-state transition after Fourier transformation of the direct (t_2) dimension(a) and after double Fourier transformation (b).

direct dimension (t_2), which led to the Loretzian functions $L_n(\omega_2)$ as given in equation (5.4).

5.2.2 Crh aggregation in a classical three-state folding transition

As described in chapter 4.3.2 protein aggregation of the Crh protein induced by a modest temperature increase could be detected in real time during a 2D (^{13}C,^{13}C) ssNMR experiment. The transition could be best followed by cross signal in the Thr Cγ2-Cβ region, which revealed a resolved peak for the initial (natively folded), an intermediate (partially unfolded) and an final (aggregated) state (see figure 4.3). For the analysis of real-time solution-state NMR data, cross sections of the 2D spectra are usually fitted to theoretically, simulated peakshapes [128, 130] or direct analytical solution [131, 129]. Due to increased spectral overlap in ssNMR, here, based on equation equation (5.7), a Mathematica script was developed (see supporting figure on page 117) to calculate the difference of a full theoretical 2D cross peak pattern to the experimental data. Free theoretical parameters such as $\lambda_{n_{(1,2)}}$ and $\omega_{n_{(1,2)}}$ were chosen from the line width and position of the peak maxima in the experimental spectrum (see also Methods)[1]. Figure
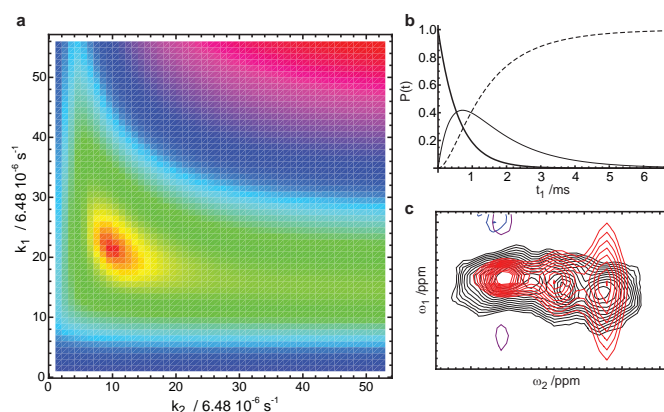
Figure 5.5: (a) Difference map between experimental and theoretical calculated Thr Cγ2-Cβ cross signal in dependence on both rate constants. Population profile (b) and comparison of spectral extract (c) of the resulting best fit. In (c) positive/negative contours of the experimental spectrum are given in black/blue, whereas theoretical data are given in red/purple.

5.5a shows the resulting difference map as a function of the two rate constants k_1 for unfolding and k_2 for aggregation. The resulting kinetic profile and spectral comparison

---

[1] Amplitudes of cross signal was chosen identical for all three states, which uses the assumption that transfer efficiency was the same in all stages of the transition.

of the best fit is given in figure 5.5b,c respectively. The timescale of the simulated data
has to be corrected for the difference of the indirect acquisition time and the time of the
experiment, according to:

$$t_{\text{effective}} = \frac{t_{\text{total,t}_1}}{t_{\text{total,experiment}}} \cdot t \approx 8.1 \cdot 10^{-8} \cdot t \tag{5.8}$$

The global minimum in figure 5.5a is given for the effective rate constants $k_1 = 1.36 \cdot 10^{-4}s^{-1}$ and $k_2 = 0.58 \cdot 10^{-4}s^{-1}$. According to figure 5.5c even the best fit shows some
considerable differences to the experimental data. Therefore calculation of the error
margin is not attempted here. However, since the detected transition predominantly
interferes with the timescale of acquisition of the indirect dimension, the $(^{13}C,^{13}C)$ spec-
trum is asymmetric and cross peaks on the other side of the diagonal offer an independent
set of information.
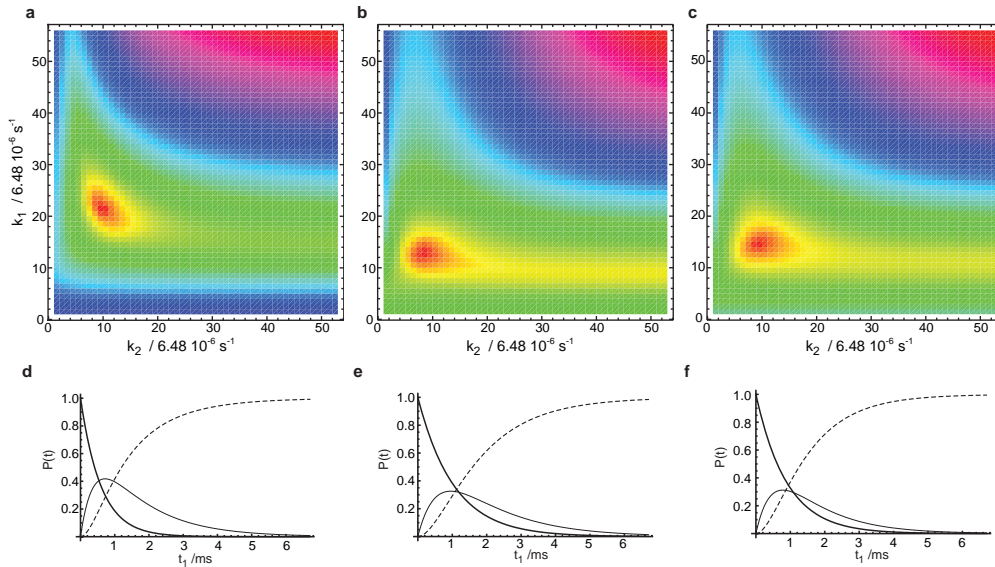
Figure 5.6: Difference maps for Thr Cγ2-Cβ (a), Thr Cβ-Cγ2 (b) and normalized sum of both (c).
(d-f) show corresponding kinetic profiles.

Figure 5.6 compares the difference maps of the Thr Cγ2-Cβ (a) and Thr Cβ-Cγ2
(b) cross correlations. In the latter case the minimum is found for $k_1 = 0.77 \cdot 10^{-4}s^{-1}$
and $k_2 = 0.52 \cdot 10^{-4}s^{-1}$. Combination of data from both sides of the diagonal leads to
the difference map shown in figure 5.6c. The respective population profiles according

to the best fit rate constants are given in figure 5.6d-e and allow for an estimation of the accuracy of the method. The rate constants obtained from figure 5.6c are $k_1 = 0.91 \cdot 10^{-4} s^{-1}$ and $k_2 = 0.65 \cdot 10^{-4} s^{-1}$.

Figure 5.7: Comparison of theoretical and experimental spectral extracts of the combined best fit (figure 5.6c). (Colors as in figure 5.5c.)

The corresponding theoretical spectra are compared to the experimental data in figure 5.7a,b. Notably, characteristic features such as line broadening for the initial and baseline 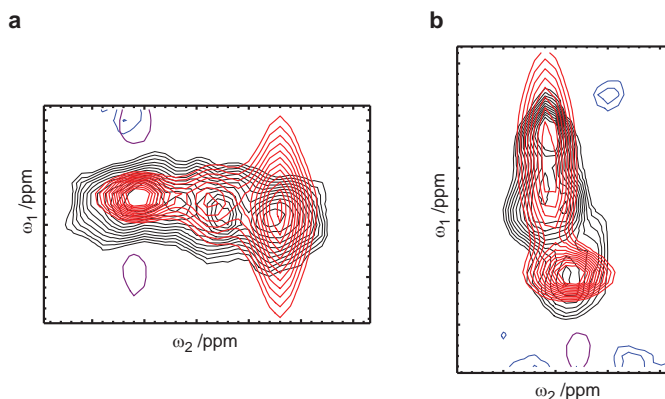distortions for the final state are significantly less reproduced by the experimental data than expected from the considered single exponential transition.

5.2.3 Crh in a classical aggregation scenario

In fact the single exponential transition used to describe protein folding [128] is in general not sufficient to kinetically describe an aggregation process since it does not account for a nucleation step. Instead several mechanisms and their mathematical formalisms have been suggested to describe experimental determined data of protein aggregation (see e.g. refs. [132, 133, 134, 135, 136, 137]). It was also shown that the measurement of a population profile is not sufficient to reveal the underlying mechanism [132]. In the following the "Finke-Watzky" (F-W) model of protein aggregation [138] was selected to theoretically descripe the kinetic profile. The F-W model assumes a 2-step mechanism involving a (slow) continuous nucleation with rate constant k_{a1}, followed by a (fast) autocatalytic surface growth with the rate constant k_{a2}, and offers the most basic

discription, which as shown recently [133] is still able to quantitatively account for all previous published results representative on the field of neurological protein aggregation. The most convenient form for analyzing the experimental data was suggested [133] to be:

$$[C]_t = [B]_0 - \frac{\frac{k_{a1}}{k_{a2}} + [B]_0}{1 + \frac{k_{a1}}{k_{a2}[B]_0}e^{(k_{a1}+k_{a2}[B]_0)t}} \tag{5.9}$$

Where equation (5.9) only describes the aggregation process from the (unfolded) intermediate to the final aggregated state (B→C). In general the transition A→B should be largely unaffected be the mechanism of aggregation, since it most likely reflects protein unfolding. Since the experimental ssNMR data contain three species in the real time spectrum the following time dependence of the relative populations was used to fit the experimental data:

$$P_A(t) = e^{-k_1 t} \tag{5.10}$$

$$P_B(t) = \left(1 - e^{-k_1 t}\right) \cdot \left(\frac{\frac{k_{a1}}{k_{a2}} + 1}{1 + \frac{k_{a1}}{k_{a2}}e^{(k_{a1}+k_{a2})t}}\right) \tag{5.11}$$

$$P_C(t) = \left(1 - e^{-k_1 t}\right) \cdot \left(1 - \frac{\frac{k_{a1}}{k_{a2}} + 1}{1 + \frac{k_{a1}}{k_{a2}}e^{(k_{a1}+k_{a2})t}}\right) \tag{5.12}$$

Figure 5.8 summarizes the results obtained with this model. The three-dimensional contour maps shown in figure 5.8a-c, visualize areas which are within 5% (red), 10% (pale yellow) and 15% (white) deviation from the minimum difference between experimental and simulated cross section of Thr Cγ2-Cβ (a), Thr Cβ-Cγ2 (b) and both (c). The resulting population profile of the minimum is shown in (d-f) respectively. Figure 5.8g and h compares the best fit (according to c and f) to the experimental data.

Notably, a comparison of figure 5.7a,b, i.e. the best fit according to a single exponential three-state transition, to figure 5.8g,h revelas that the assumption of a nucleation step here did not improve the fitting. Indeed the line broadening of the initial state is predominantly related to the unfolding mechanism and as evident from figure 5.8 not reproduced even by the best fit of the profiles represented by equations (5.10)-(5.12).
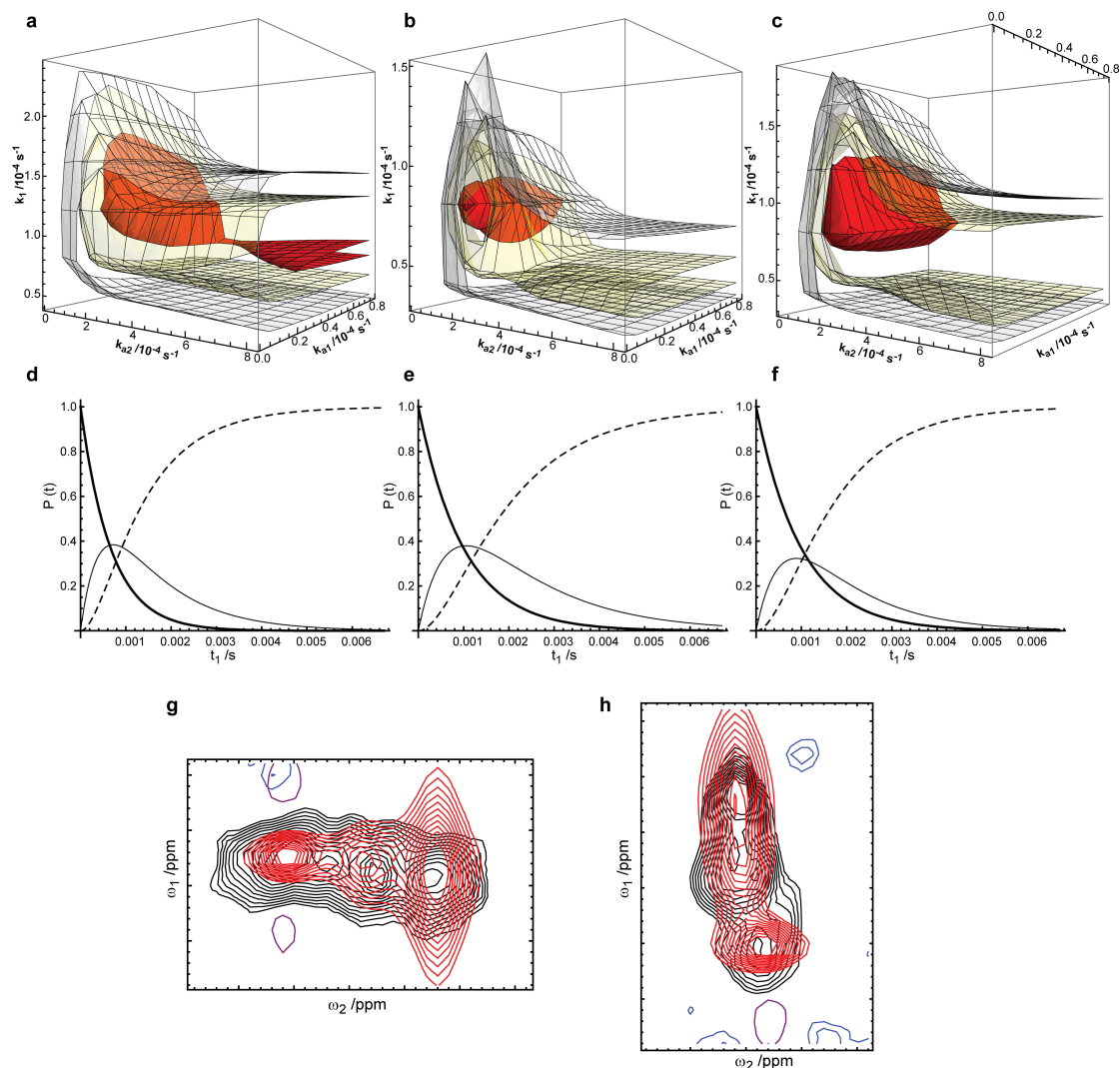
Figure 5.8: Results of kinetic analysis using the F-W model. (a-c) show three-dimensional contour maps of the difference between experimental and simulated data depending on the three parameters k_1, k_{a1} and k_{a2}. Red, yellow and white contour level indicate areas within 5%, 10% and 15% deviation to the best fit. The results for the Thr Cγ2-Cβ (a) / Thr Cβ-Cγ2 cross section are shown in (a) /(b), whereas (c) shows the sum of both. The kinetic profiles of the best fit are given in (d-f) respectively. (g) and (h) compare the experimental data to the best fit according to (c and f). Color code as in figure 5.5c.

5.2.4 Crh in downhill unfolding

In protein (un)folding also some other experimental results exist that could not very well be described with a single exponential transition [139]. Phenomenologically, these findings were related to a small or absent energy barrier separating the initial and final state [140, 141, 142]. The absence of a distinct energy barrier can lead to a variety of possible transitions, each reflected by a different rate constant. This process is called downhill folding and multiple or stretched exponential functions can be used to account for the occuring effects [143, 141]. Since the latter reduce the effect to a single additional parameter b $\in \{0, 1\}$, a stretched exponential function is used to describe the protein unfolding in the following. The resulting population profiles read as follows:

$$P_A(t) = e^{-(k_1 t)^b} \tag{5.13}$$

$$P_B(t) = \left(1 - e^{-(k_1 t)^b}\right) \cdot \left(\frac{\frac{k_{a1}}{k_{a2}} + 1}{1 + \frac{k_{a1}}{k_{a2}} e^{(k_{a1} + k_{a2})t}}\right) \tag{5.14}$$

$$P_C(t) = \left(1 - e^{-(k_1 t)^b}\right) \cdot \left(1 - \frac{\frac{k_{a1}}{k_{a2}} + 1}{1 + \frac{k_{a1}}{k_{a2}} e^{(k_{a1} + k_{a2})t}}\right) \tag{5.15}$$

Equations (5.13)-(5.15) already contain four free parameters. The results of fitting these model to the experimental spectrum are shown in figure 5.9a as a series of 3D contour maps depending on the parameter b. Each 3D plot depends on the same three parameters as already described in figure 5.8. It is clearly evident that the simulated cross correlation gets closer to the experimental data for decreasing values of b. Plotting the minimal difference to the experimental data for the values of b taken into account (figure 5.8b) reveals that the global minimum is found for $b = 0.1$. However, also values up to $b = 0.4$ may reproduce the experimental spectrum within an error of 5%. Figure 5.9c-e show the population profile as well as the resulting peak pattern of the best fit ($b = 0.1$; $k_1 = 2.03 \cdot 10^{-4} s^{-1}$; $k_{a1} = 0.78 \cdot 10^{-4} s^{-1}$; $k_{a2} = 0.81 \cdot 10^{-5} s^{-1}$).

Indeed the downhill unfolding scenario could explain the missing line broadening in the experimental data of state A. Notably, a complete downhill scenario as suggested for transthyretin aggregation [126] could also explain the data (see supporting figure on page 116). However, since the experimental data analysed here exclusively report on the population profile and additional experimental data on the effect of concentration of native Crh and of seeding with pre-aggregated Crh is lacking, no discrimination be-
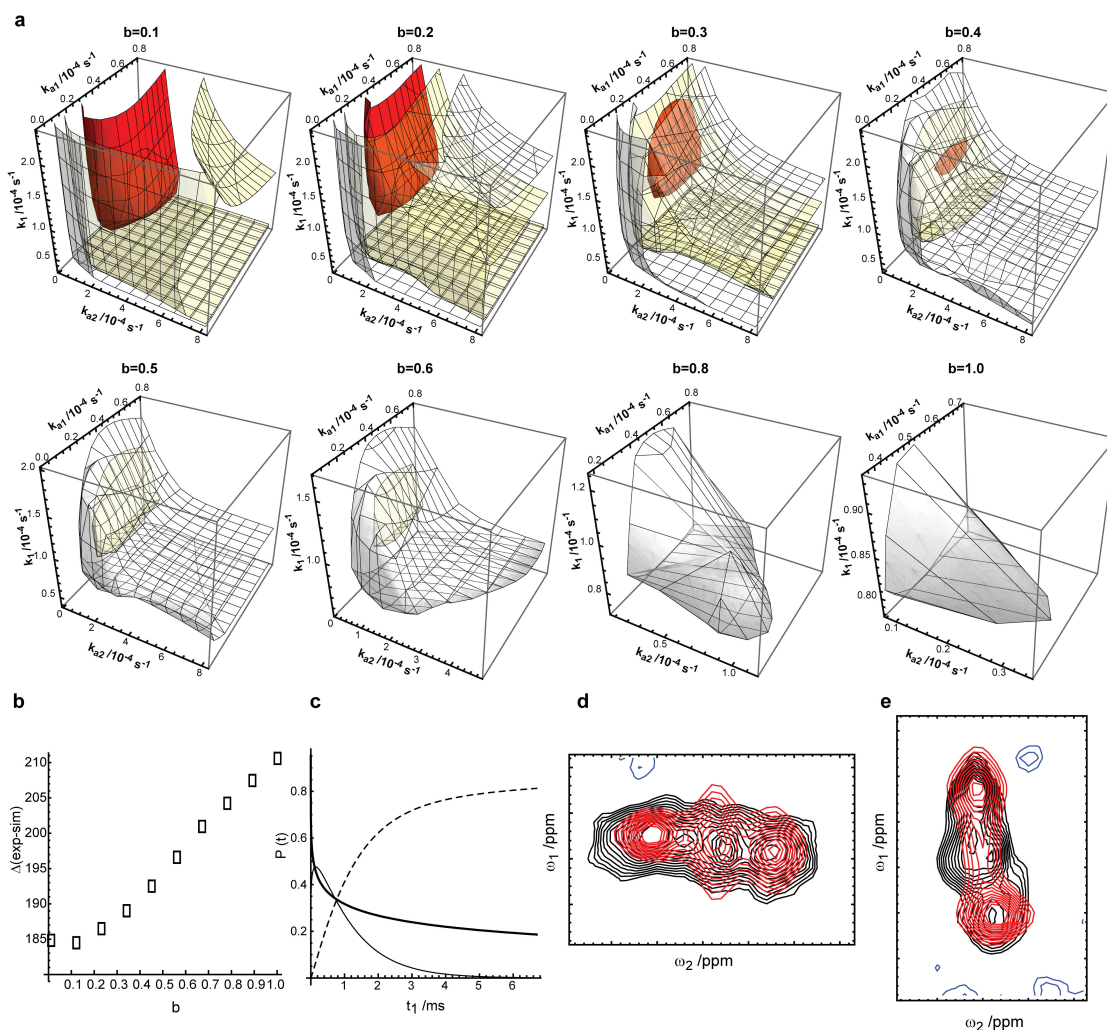
Figure 5.9: Kinetic analysis assuming downhill unfolding and the F-W model for aggregation. (a) shows for indicated values of b the same 3D counter maps as given in figure 5.8a-c. The minimal difference found to the experimental depending on b is plotted in (b). (c) shows the population profile of the global minimum and (d+e) the corresponding simulated peak pattern.

tween downhill aggregation, i.e. barrier free growth of aggregates with the monomer as the critical nucleus size, and the F-W model, i.e. continuous nucleation followed by a autocatalytic surface growth, is possible [132, 126].

5.3 Conclusion

Manifold methods exist that allow studying protein aggregation kinetics. However, most of these techniques rely on detection of global factors or indirect mechanisms (e.g. the β-strand content in Thioflavin T binding or CD spectroscopy) and may lack the formation and kinetic properties of intermediate species. Here it is shown that ssNMR can contribute to kinetic analysis of protein aggregation by offering site resolved resolution, as well as the simultaneous detection of several different states including intermediate states, which might be an important target to interfere with the aggregation [11]. The range of observable states can be modified by a combination of different sets of polarization transfer mechanisms (dipolar or scalar) [35] or simply recording after direct excitation [116]. Here, the analysis was purely based on ssNMR cross signals from a single kinetic transition to investigate the potential of the method. Rate constants could be extracted using a conventional single exponential three-state transition as well as a conventional aggregation mechanism. No significant difference between the two approaches could be detected, suggesting that the formation of a nucleus is not a significant step in the aggregation of the Crh protein.

Remaining differences to the experimental data additionally suggest that a single exponential transition for the initial unfolding step does not suffice to properly describe the detected process. Instead a stretched exponential function, as found in downhill folding [139, 141], significantly improves the reproduction of experimental data. Note that a sigmoidal transition for the unfolding step is counterproductive in explaining the peak shape, whereas a complete downhill scenario offers an additional possible mechanism. The limited sensitivity and resolution combined with the lack of additional information and the degree of freedom offered by the underlying four free parameters does not allow for a more detailed analysis of the mechanism and complicates the exact determination of the involved rate constants. At this point the combination with other methods to determine at least one of these free parameters should significantly improve the interpretation.

In summary, the presented approach offers the possibility to obtain site resolved information directly and from a large variety of possible (intermediate) states. The resulting

information is in general complimentary and should be combined with additional methods to extend the kinetic and thermodynamic understanding of the aggregation process.

5.4 Methods

Mathematica version 6.0.1 was used to fit the experimental data to simulated peak pattern by numerically integrate (FT) $S(t_1, \omega_2)$ using the discussed population profiles. The applied script is attached in the supporting information on page 117. 'SymbolicProcessing' was switched off to speed up the integration process. The experimental spectrum was processed using exponential line broadening of 50Hz to increase signal to noise. The underlying window function was also implemented in the simulations before FT in the indirect dimension. The effect in the direct dimension was neglected. Additionally the transfer efficiency in all three states as well as the behaviour of the three involved Thr residues was assumed to be the same.

6

Secondary structure, dynamics and topology of a seven-helix receptor

6.1 Introduction

Understanding the molecular details of signal transduction across the cell membrane is intimately related to the investigation of conformational changes and molecular flexibility of membrane proteins at atomic resolution. For example, X-ray crystallography has provided important insight into the structural details of color-sensitive phototaxis. Upon light excitation, sensory rhodopsin II (SRII) triggers via a tightly bound transducer protein (HtrII) a signal transduction chain homologous to the two-component system of eubacterial chemotaxis (see figure 6.1 for details). Sensory rhodopsins are seven-helix (A-G) membrane proteins containing as cofactor retinal which is bound to a lysine residue on helix G via a protonated Schiff base. The color of this complex ($\lambda_{max} = 490nm$) is regulated by specific protein-retinal interactions. On light excitation the all-trans chromophore isomerizes to a 13-*cis* configuration, thereby triggering conformational changes, which lead to the activation of the cognate transducer [144]. 3D structures were not only obtained of SRII in free [145, 146, 147] and transducer bound form [148] but also photocycle intermediates K and late M were recently characterized on atomic level [149]. These structures provide snapshots of the structural rearrangements

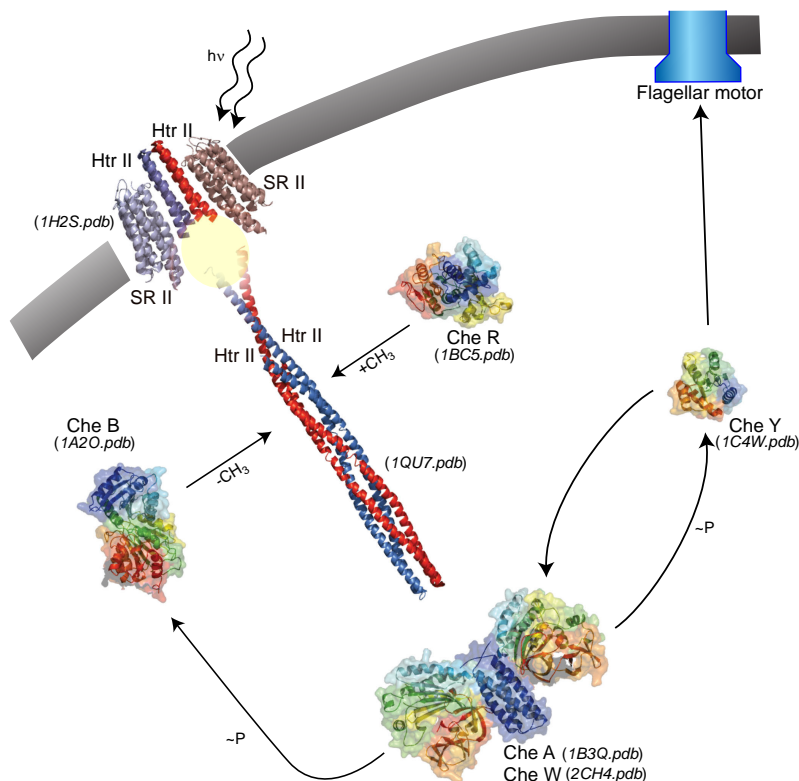in SRII and concomitant changes in the transducer upon light activation.

Figure 6.1: Phototactic two-component signalling cascade induced by the SRII/HtrII system. CheA, a histidine kinase, is bound via the adapter molecule CheW to the cytoplasmic tip of the transducer. Ligth sensing by the receptor is transfer to and along the transducer and increases the autophospory- lation activity of CheA.In the next step one of the two response regulators, CheY or CheB is phospory- lated by CheA. Phosporylated CheY is a switch factor of the falgellar motor increasing the reversal frequency. To adapt to constant stimulus and retain the ability to sense altering gradients of stimulus the level of methylation modulates the signalling state in an inverse manner. Here phosphorylation of CheB acts against the constitutively active methyltransferase CheR, decreasing the methylation level and hence the attractant signalling state [150]. (*Picture adopted from* [144])

For a full thermodynamic characterization of signal transduction across the cell mem- brane, not only intermediate structures but also information on their relationship to the intrinsic protein dynamics is hence of particular interest. In principle, the B-factor (atomic displacement parameter) in crystal structures is regarded as an important indi- cator of protein flexibility [151]. In addition to molecular dynamics, these parameters

are, however, sensitive to resolution of the structure, packing effects and the refinement procedure [152]. Additionally, information about protein flexibility was obtained from the comparison of X-ray structures to kinetic data or computational results [153, 154] and from conducting solution-state NMR experiments [155, 156, 157]. These studies provide evidence that intrinsic flexibility of proteins in the unbound state can correlate with conformational changes associated with protein-protein interactions [158]. In this chapter it is discussed how solid-state NMR (ssNMR) can be used to obtain site-resolved information on structure and dynamics of membrane protein receptors under near physiological conditions. Following these initial studies chapter 7 shows effects related to complex formation.

Solid-state NMR under magic-angle spinning (MAS) conditions has previously been successfully applied to investigate ligand binding to membrane proteins [159, 160, 161, 162, 163, 14]. However, the determination of entire 3D structures of larger membrane-embedded proteins has been complicated by the length of the amino acid sequence, the high repetitiveness of hydrophobic residues, and the limited spectral resolution arising from the dominant influence of a single type of regular secondary structure (α-helix or β-sheet) [164, 165, 166]. In principle, spectral crowding can be reduced by advanced labeling approaches [165, 166, 167], but these may be precluded by low protein expression levels and/or the costs of labeled starting materials.

In the following it is shown that uniform isotope labeling (with ^{13}C and ^{15}N), which is straightforward in many cell-based expression systems, plus the addition of a well-selected set of unlabeled amino acids (reverse labeling) [55, 56], and the application of ssNMR spectroscopic methods that separate the signals of mobile, static, and water-exposed protein segments can be used to investigate the structure and topology of an entire seven-helix receptor. The experiments are carried out on a single isotope-labeled sample in a native membrane environment. The approach is demonstrated for sensory rhodopsin II from *Natronomonas pharaonis* (NpSRII).

6.2 Results

6.2.1 Secondary structure

Figure 6.2a shows a ^{13}C-^{13}C spin diffusion (SD) [168] spectrum of an NpSRII sample that is uniformly labeled, with the exception of the four dominant residue types valine, leucine, phenylalanine, and tyrosine, which occur in natural abundance (U[^{13}C,^{15}N \(V,L,F,Y)] NpSRII). The spectra were recorded in proteoliposomes at a ^{1}H resonance frequency of 800 MHz. The observed ^{13}C line width of 0.7 ppm is not only indicative of a well-folded membrane protein, but also confirms that high-resolution ssNMR spectra can be obtained for membrane proteins in their natural lipid environment. The reverse labeling of four dominant residue types, which account for 34% of the entire amino acid sequence, significantly improves spectral resolution, compared to that in studies of a uniformly labeled sample [164]. This aspect is particularly striking in figure 6.2b where a 2D SD spectrum recorded under weak coupling conditions (SDWC) [23] reveals a variety of sequential correlations, which are characterized by signal intensities that are significantly higher than the noise level (see supporting figure on page 120). These correlations reduce the level of ambiguity introduced by the occurrence of shorter isotope-labeled amino acid stretches in U[^{13}C,^{15}N\(V,L,F,Y)] NpSRII (compared to a fully labeled sample).
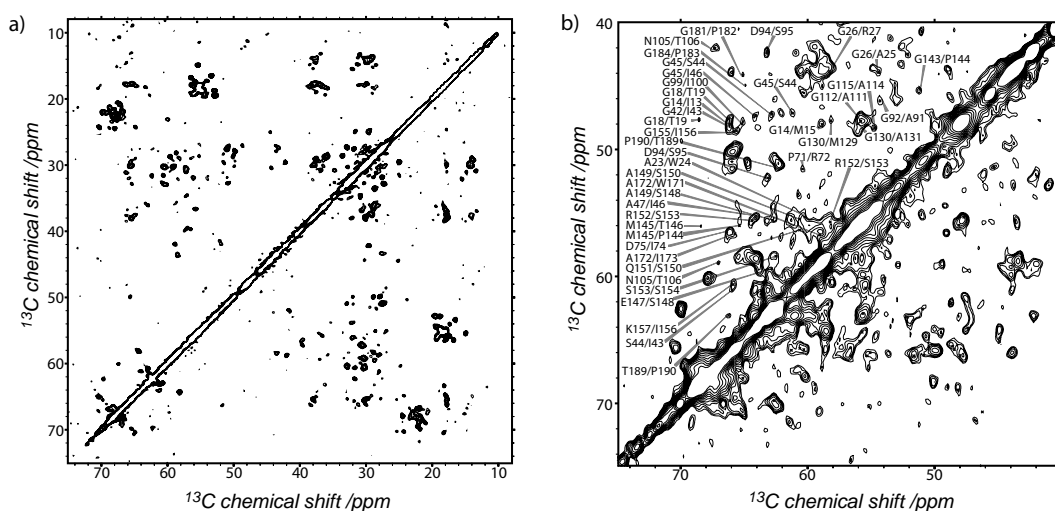
Figure 6.2: Dipolar ^{13}C-^{13}C correlation spectra of U[^{13}C,^{15}N\(V,L,F,Y)] NpSRII in proteoliposomes. (a) ^{13}C-^{13}C SD spectrum (mixing time 15 ms). (b) ^{13}C,^{13}C SDWC spectrum [23] (mixing time 150 ms); sequential Cα,Cα and Cα,Cβ correlations are indicated.

The correlations were combined with the results of ^{15}N-^{13}C dipolar (through-space) transfer experiments [46] (see supporting figure on page 121) to derive sequential resonance assignments, which were classified into three levels of reliability (see supporting figure on page 118 for the assignment list and following text for a more detailed discussion of the assignment quality). Figure 6.3 shows an example of an "sequential-walk" for residues in the E-F loop region. Based on all assigned residues, conformation-dependent chemical shifts could be defined, which revealed a mainly α-helical conformation for the transmembrane segments of NpSRII (see figure 6.4 for a secondary chemical shift plot).

6.2.2 Membrane topology

To further characterize the supramolecular assembly of NpSRII in native membranes, an water-edited ^{13}C-^{13}C correlation experiment was performed. The corresponding ssNMR pulse sequence relies on the possibility of polarization exchange between mobile protons from water and protons from the protein complex [52]. Firstly, the signals of the mobile water protons can be selected using a relaxation filter. During a subsequent mixing time, polarization transfer to the immobilized biomolecule can take place. For short mixing times, the resulting polarization-transfer characteristics in the spectra are sensitive to the distance between a given nuclear spin in the interior of the molecular complex and the surrounding water environment [53].

For the purpose of these studies, the existing pulse schemes were extended by an additional ^{13}C-^{13}C mixing unit, which permits to record a 2D correlation map of all detectable protein resonances for a given diffusion time (see figure 6.5).

The water-edited ^{13}C-^{13}C spectrum shown in green in figure 6.6 contains only a subset of the correlations observed in figure 6.2a). Under the selected experimental conditions, with a ^{1}H-^{1}H mixing time of 4ms, polarization transfer is largely confined to the protein-water interface of NpSRII. Because rigid components are selected by the dipolar polarization transfer, signal sets are observed for immobilized protein segments, such as Glu 147, Ser 153, and Thr 189. Notably, the absence of cross peaks for Asp 94 in the water-edited ^{13}C-^{13}C correlation experiment suggests that the peptide loop (L3) connecting helices C and D is embedded in the membrane. These correlations allow the
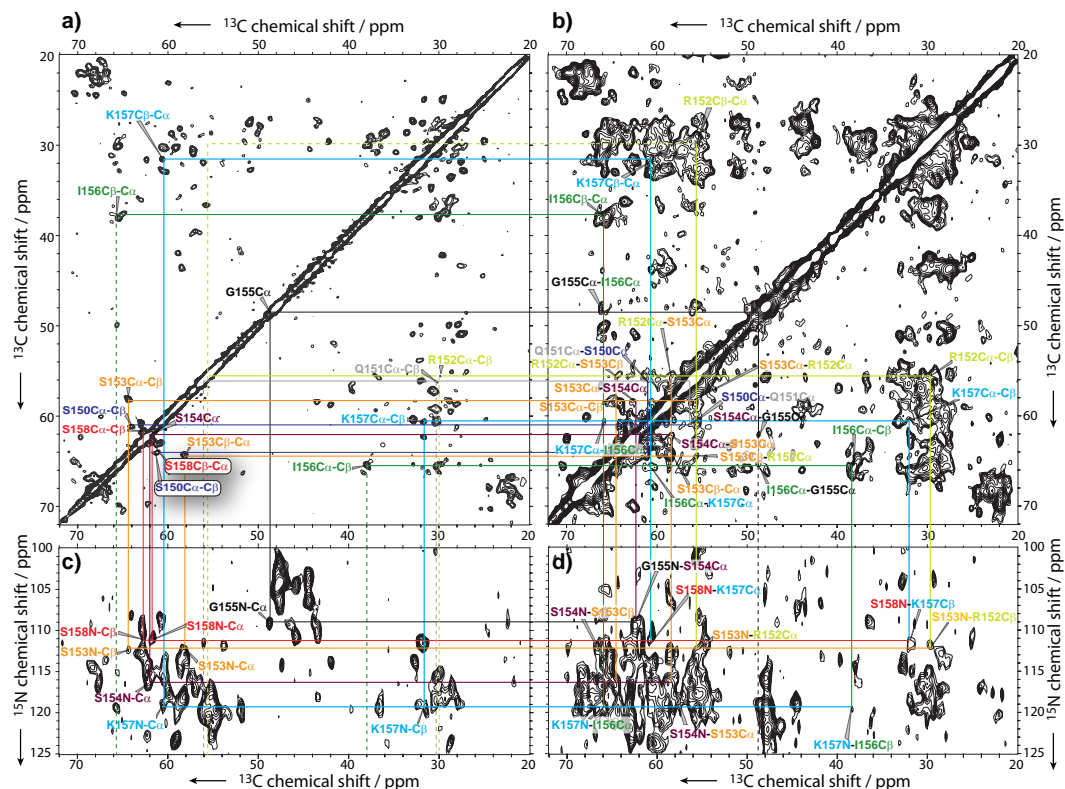
Figure 6.3: Sequential assignment of the amino-acid stretch Ser150-Ser158 using PDSD spectra with $t_{CC} = 15ms$ (a) and $t_{CC} = 150ms$ (b) (b under weak coupling conditions [23]).3]). N-C correlation spectra recorded under SPECIFIC transfer conditions [32] are shown in c) (NCACX) and d)(NCOCX). Note that the spectra in a) and c) only contain intra-residue transfer; d) only contains sequential $(N_i - C_{i-1})$ correlations whereas b) contains intra- and inter-residue $(C_i - C_{i\pm1})$ crosspeaks. Resonances of each amino acid of the considered stretch are characterized by a specific color. Vertical and horizontal lines exemplify a sequential walk within the considered amino-acid stretch. In the case of spectral overlap in ^{15}N and ^{13}Cα backbone resonances, side-chain information was used. Variations in cross peak amplitudes are likely caused by differential molecular mobility.

known crystal structure of NpSRII to be positioned relative to the lipid-water interface (see below).

Figure 6.4: Secondary chemical shifts [28] $\Delta\delta_i = (\delta C\alpha_i - \delta C\alpha_{BMRB}) - (\delta C\beta_i - \delta C\beta_{BMRB})$ for the assigned residues in U[^{13}C,^{15}N\(V,L,F,Y)]NpSRII. (If Cβ is not present (Gly) or not assigned, only Cα shifts were used.) Positive values are indicative of α-helical segments, strong negati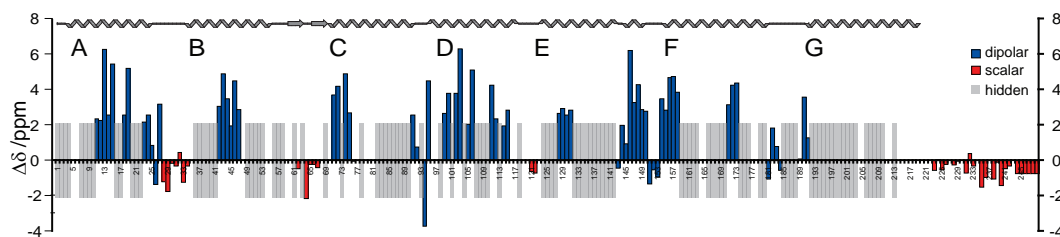ve values are consistent with β-strands. For comparison, a secondary structure plot as seen by X-ray crystallography (PDB:1h68) is given at the top of the figure.
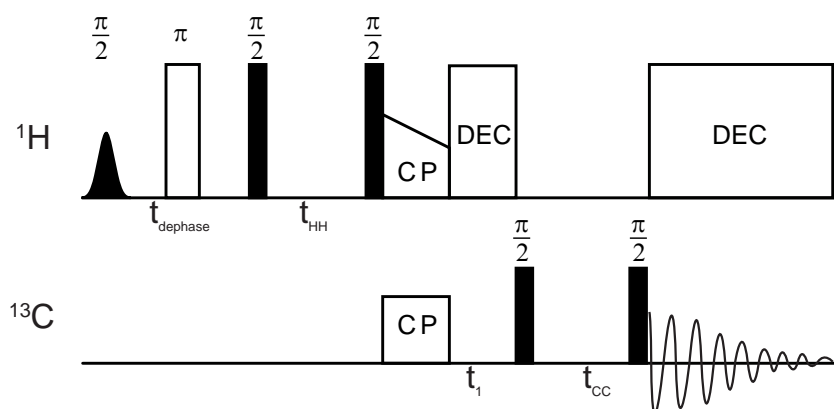
Figure 6.5: The pulse scheme employs polarization transfer from mobile water to the semi-solid environment [52]. After an initial Gaussian 90° pulse on resonance with the water ^{1}H signal, a ^{1}H T_2 filter (including a dephasing time and a 180° refocusing pulse) leads to the selection of mobile water ^{1}H magnetization. A subsequent longitudinal spin diffusion unit (T_{HH}) leads to polarization transfer to protein protons. As shown elsewhere [53], the rate of transfer is, in the initial rate regime, sensitive to the relative distance between protein spin and bilayer surface. The corresponding protein residues are subsequently read out after a cross polarization step to ^{13}C followed by a (^{13}C,^{13}C) spin-diffusion unit and recorded in two spectral dimensions.

6.2.3 Intrinsic molecular dynamics

The absence of larger fractions of receptor loops in the dipolar spectra shown in figure 6.2 may either be due to static structural heterogeneity or fast molecular motion. To
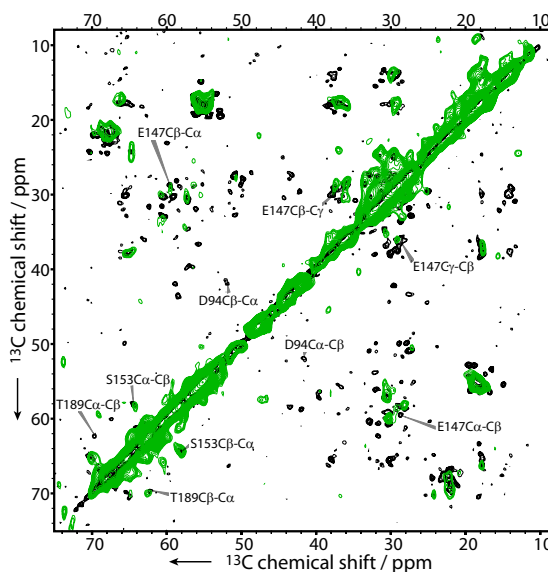
Figure 6.6: Comparison of a water-edited ^{13}C-^{13}C correlation spectrum (green; ^{1}H-^{1}H mixing time 4 ms) to spectrum in figure 6.2a (black); correlations discussed in the text are indicated.

discriminate between the two mechanisms, a series of 2D correlation experiments under MAS conditions was recorded using scalar (through-bond) transfer units, which detect only mobile protein segments [35]. In figure 6.7a, the spectra recorded under ^{1}H-^{13}C INEPT (red) [34, 35] and ^{1}H-^{13}C INEPT-TOBSY (black) [36, 35] transfer are super-imposed. Using these spectra, amino acid specific assignments can readily be made. To obtain sequential resonance assignments, NCA and NCOCA correlation experiments based on scalar coupling were performed (figure 6.7b+c). Similarly to spectra based on dipolar transfer (figure 6.2), the resolution in both the ^{15}N and the ^{13}C dimensions of the NCA and NCOCA spectra is well below 1 ppm and, thus, permits sequential assignments in a variety of protein segments. Note that in figure 6.7b the correlations involving glycine residues (blue) are negative, owing to the specific transfer characteristics within methylene groups. Taking into account the reverse isotope-labeling pattern of our NpSRII sample, the scalar transfer experiments unambiguously reveal that the C terminus, from residue 223 onwards, is mobile. Additional correlations in the 2D spectra suggest that other protein segments also exhibit high mobility. According to our data, these segments involve loops A-B (L1), B-C (L2), and D-E (L4). Indeed, the corre-

sponding secondary chemical shifts of these residues are largely of random-coil character (see figure 6.4 (red bars)), indicating fast structural rearrangements in the protein on a ns-to-μs time scale.

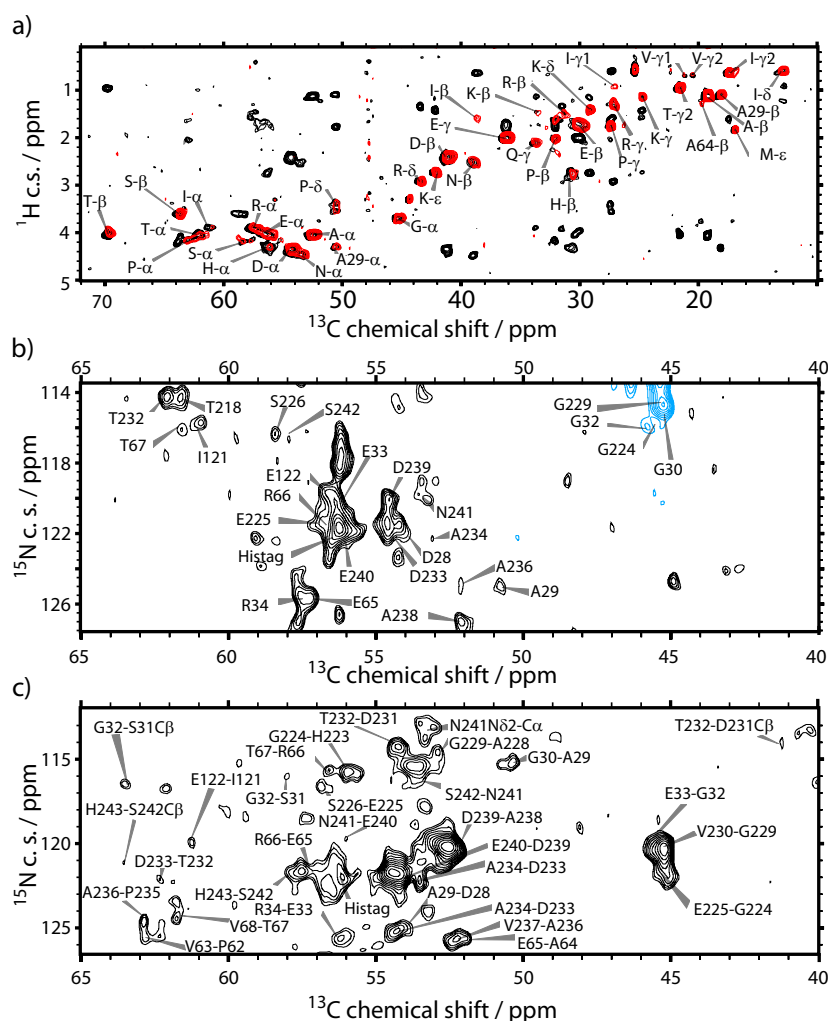**Figure 6.7:** Scalar ^{1}H-^{13}C and ^{15}N-^{13}C correlation spectra of U[^{13}C,^{15}N\(V,L,F,Y)]NpSRII in proteoliposomes. (a) Comparison of a ^{1}H-^{13}C INEPT spectrum (red) with a ^{1}H-^{13}C INEPT-TOBSY spectrum (black). (b) NCA spectrum [35]; correlations involving glycine residues are indicated in blue. (c) NCOCA spectrum [35]. In all cases, mobile protein segments are detected.

6.3 Discussion

Figure 6.8 summarizes the ssNMR spectroscopic results on a schematic of the primary sequence and in a representation of the 3D crystal structure of NpSRII (PDB 1H68) [147]. Immobilized protein segments that are largely of α-helical character are indicated in blue. In total, resonances were assigned for 98 amino acids, that is, 73% of the sequentially assignable receptor residues. Our data suggest that in NpSRII, reconstituted in purple membrane lipids, only three out of six peptide loops, as well as the C terminus, exhibit sizable dynamics. Furthermore, information from the water-edited ^{13}C-^{13}C correlation experiment was used to position the protein in a model membrane (see figure 6.8 b).

In summary, it could be demonstrated that high-resolution ssNMR spectroscopy can be used to study the secondary structure, dynamics, and membrane topology of an entire seven-helix receptor in a native membrane environment. The structural accuracy could be further improved by combining the spectroscopic results for several differently reverse-labeled protein samples, possibly in the context of 3D NMR spectroscopy. The data presented here provide a basis for further ssNMR spectroscopic studies, for example, the use of indirectly detected proton-proton contacts [14], to assemble the 3D structure of NpSRII in a native membrane environment (see also supporting information for initial feasibility studies).

Complemented by other biophysical methods [148, 149, 169], ssNMR spectroscopic studies, such as those described herein, may provide important insight into the structural details associated with signal transduction in NpSRII or other membrane proteins.

6.4 Materials and methods

U[^{13}C,^{15}N\(V,L,F,Y)] NpSRII was expressed in *E. coli* grown in ^{13}C labeled glucose and $^{15}N NH_4Cl$. Unlabeled amino acids were added in concentrations of 1 mmol (leucine, phenylalanine, and tyrosine) or 4 mmol (valine). Proteoliposomes were prepared as described in reference [170]. All NMR experiments were conducted using 4-mm triple-resonance (^1H,^{13}C,^{15}N) probeheads at static magnetic fields of 18.8 and 14.1 T, corresponding to ^1H resonance frequencies of 800 and 600 MHz (Bruker Biospin, Karlsruhe).
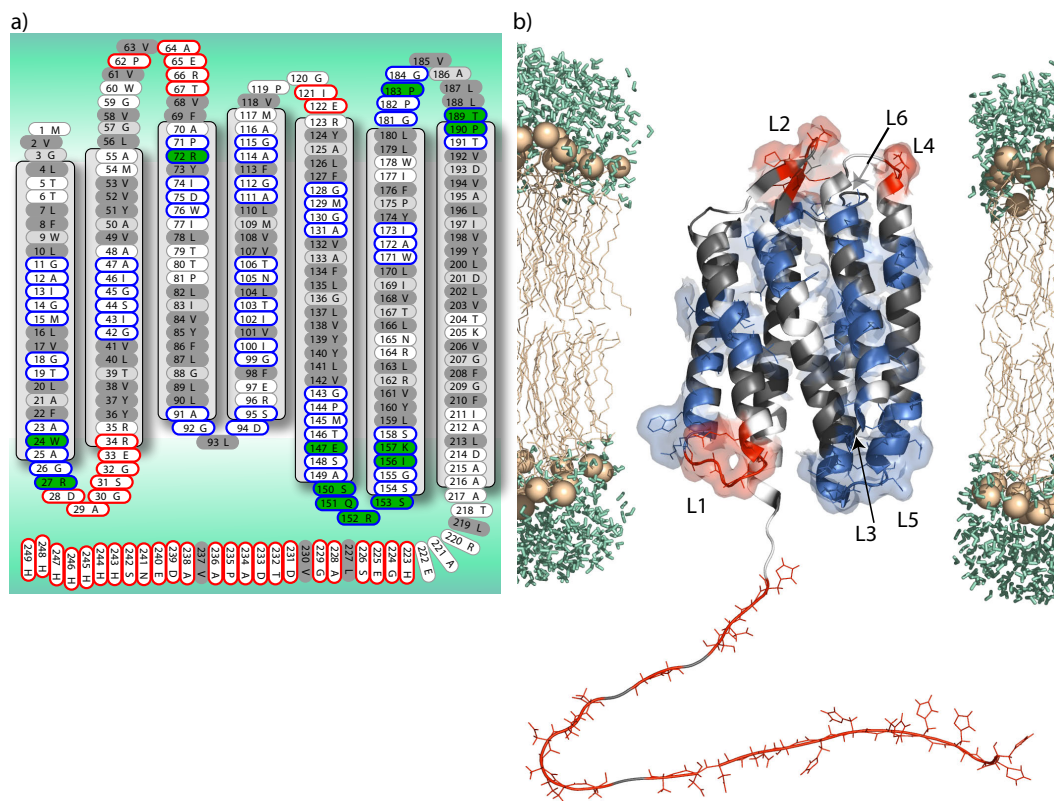
Figure 6.8: Summary of the ssNMR spectroscopic results on U[^{13}C,^{15}N\(V,L,F,Y)]NpSRII in prote-oliposomes, mapped onto a) a schematic of the primary sequence, in an aqueous environment, and b) a representation of the 3D crystal structure of NpSRII [147] in a lipid-bilayer environment. In both cases, mobile residues are indicated in red, and rigid protein segments in blue. In (a), the residues identified in the water-edited ^{13}C-^{13}C correlation spectrum are indicated in green, residues not labeled in the U[^{13}C,^{15}N\(V,L,F,Y)] NpSRII sample in dark gray, and residues that cannot be assigned sequentially, as a result of the reverse labeling, in light gray. In both cases, residues exhibiting fast motion are indicated in red.

Dipolar transfer experiments involved broad-band ^{1}H,^{13}C and chemical-shift-selective [46] ^{15}N,^{13}C cross-polarization (CP) schemes. SPINAL64 proton decoupling [25] was applied during the dipolar correlation experiments using radio-frequency fields of 75-90 kHz. Sequential ^{15}N,^{13}C resonance assignments were made by combining 2D NCACX and NCOCX [46] results with the results of (^{13}C,^{13}C) correlation experiments performed under weak coupling conditions [23]. MAS rates of $8 - 12.5 kHz$ and probe tempera-

tures of $260 - 278K$ were used. Mobile protein segments were investigated by NMR spectroscopic methods described elsewhere [35]. Water-edited (^{13}C,^{13}C) correlation experiments were conducted using a ^{1}H relaxation filter of 1 ms and a ^{1}H,^{1}H mixing time of 4 ms.

7 Chapter 7

Complex formation of the sensory rhodopsin - transducer system

7.1 Introduction

While increasing progress is being made to characterize protein sequence and structure found in a cellular environment, a detailed understanding of which aspects determine membrane protein function and a principal understanding of intermolecular communication between integral membrane proteins is currently limited. The NpSRII/NpHtrII complex offers an ideal system to study such aspects since, according to crystallographic data, its intermolecular interactions are found predominantly within the membrane integral parts (see figure 7.1).

Moreover, a striking characteristic of macrobial rhodopsins in general is their diverse function [150]. Some are light-driven transporters, such as the proton pumps bacteriorhodopsin (BR), the Cl-Ion pump halorhodopsin (HR) [171, 172] and proteorhodopsin in marine bacteria [173]. Others are light sensors, such as the sensory rhodopsins SRI and SRII, which were found in different bacteria [174, 175, 176, 177]. Additional light-gated ion channels (channelrhodopsins; ChR1 and ChR2) were found [178, 179]. Protein function can be changed from a photo signaler to an (ion) transporter by simply removing the transducer molecule [180, 181, 182]. In addition, it was shown that mutation

of three residues in BR is sufficient to efficiently relay the retinal photoisomerization signal to the SRII integral membrane transducer (HtrII) and induce robust phototaxis responses [183].

Is protein sequence and structural homology hence the only determinant for protein function or are additional aspects, such as molecular plasticity, relevant? Interestingly, the structures available for NpSRII (free [146, 147]; transducer-bound ground [148] and light activated [149] state) not only deviate in several aspects from other biophysical evidence [184] in reference to the light induced structural changes in the NpSRII-HtrII systems, but they also speak against significant changes in terms of receptor structure (see figure 7.1) or dynamics, as evidenced by the crystallographic B-factors, upon complex formation (see figure 7.2). However, a view along the crystal lattice reveals that crystal packing may significantly impact a detailed structural and motional analysis in specific areas of interest. Based on the results obtained in terms of protein structure and dynamics of isolated NpSRII (presented in chapter 6), structural and dynamical changes associated with the formation of an NpSRII/NpHtrII protein complex were subsequently investigated in lipid bilayers using ssNMR and will be presented in this chapter.

Distinct changes in molecular conformation and protein dynamics upon complex formation were found. In addition, the receptor-transducer interface was studied on the basis of intermolecular proton-proton contacts. In accordance with other biophysical data [185, 186] it is shown that interaction between the two membrane proteins extends farther than what was seen in protein crystals. 2D (^{13}C,^{13}C) spectroscopy was used to investigate receptor function. In a light activated, freeze trapped sample chemical shift changes could be identified for residues in helix E,F and G as well as in the retinal pocket.

7.2 Results and discussion

Comparing the crystallographic data of the free [147] and transducer bound [148] NpSRII as shown in Figure 7.1, only minor structural changes are expected for NpSRII upon complex formation. However, the chemical shift of a spin is in general very sensitive to local modifications in its chemical surroundings and hence numerous chemical shift
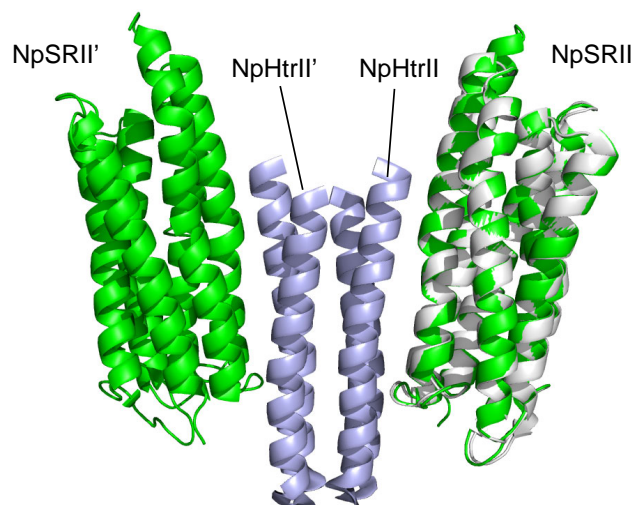
Figure 7.1: X-ray structure of the dimeric NpSRII/NpHtrII complex [148]. On the right hand side the X-ray structure of isolated NpSRII (white) [147] is aligned to the complexed NpSRII (green); NpHtrII is shown in blue.

changes between the free and bound states of the interaction partners have been used to identify binding interfaces in solution [187, 188, 189]. Thus far, ssNMR studies of molecular complexation mainly involved smaller peptide - protein interactions (see, e.g. [161, 14, 190, 191, 192]).

7.2.1 Detecting structural changes at an atomic level (Proof of principle)

To investigate whether minor structural modifications can also be detected in the context of a 250 amino-acid membrane protein receptor, a mutant (D75N) of NpSRII was used. This mutant has been shown to exhibit minimal visible light absorption changes during its photocycle, but mediates normal phototaxis responses [193, 194]. Notably, Asp75 contains the counter ion of the protonated Schiff base and is the proton acceptor from 13-cis retinal. Figure 7.3 compares a ssNMR 2D (^{13}C,^{13}C) correlation spectra obtained on reverse labeled WT NpSRII in complex with non-labeled NpHtrII, i.e. U[^{13}C,^{15}N\(V,L,F,Y)]NpSRII/NpHtrII and on D75N U[^{13}C,^{15}N\(V,L,F,Y)]NpSRII/NpHtrII. Overall, the cross peak pattern is largely identical thus confirming that protein secondary and tertiary structure is largely unperturbed by the mutation. The data are identical for free
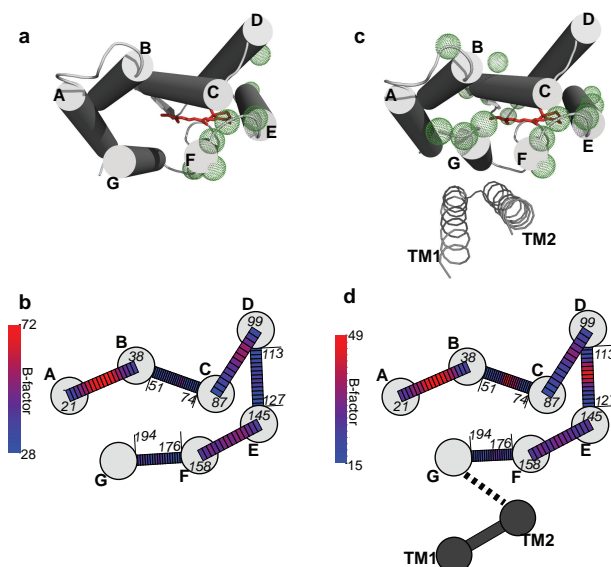
Figure 7.2: Top view on the structures of free (a+b) and complexed (c+d) NpSRII. In a+c helices (A-G) are shown as grey cylinders. Green spheres indicate $C\alpha$ atoms with close intermolecular contacts ($d_{C\alpha_i C\alpha'_j} < 6\text{Å}$) within the crystal lattice. The retinal is shown in red. b+d illustrate the loop 'mobility' as seen by the crystallographic B-factors. Starting and ending residues for the considered loops are indicated. The color of the $C\alpha$ B-factor of each residue is plotted according to the scale shown on the left hand side. Grey circles represent cytoplasmic end of each helix as shown in a+c.

NpSRII and D75N-NpSRII (see supporting figure on page 122).

Closer inspections (figure 7.3b,c for complexed and free NpSRII, respectively) reveal chemical-shift changes between WT (black) and D75N NpSRII (blue) for Thr79 $C\alpha$-$C\beta$ close to the retinal, suggesting a rearrangement of the hydrogen bonding network. According to the X-ray structure of wild-type NpSRII, the side chain of Asp75 forms two hydrogen bonds; the first with Thr79, the second with coordinated water 402. (Since the hydrogen atom positions are missing in the crystallographic data, the oxygen distances are plotted in figure 7.3d.) A peak shift detected for Thr79 (which could not be assigned in the WT) indicates that its hydrogen bond to the Asn75 mutant is not conserved, which suggests that the NH_2 group of Asn75 points towards Thr79 and hence the hydrogen bond to water 402 should largely be unaffected.

These findings are well in line with very recent crystallographic data of D75N NpSRII
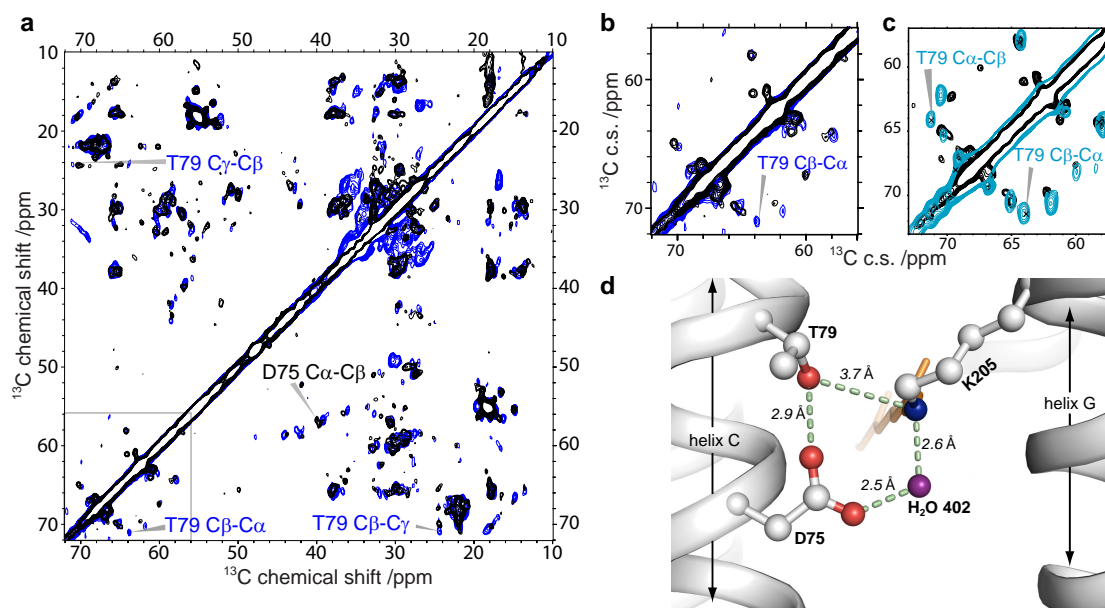
Figure 7.3: a) Comparision between (^{13}C,^{13}C) SD spectra of WT NpSRII/NpHtrII (black) and D75N NpSRII/NpHtrII (blue). Spectal extract illustrating Thr Cα-Cβ region in (b) WT NpSRII/NpHtrII (black) vs D75N NpSRII/NpHtrII (blue) and (c) WT NpSRII (black) vs D75N NpSRII (light blue). (d) Structure of the mutation side in WT NpSRII. The retinal is shown in yellow; the Schiff base nitrogen in blue; the oxygens of Thr79 and Asp75 in red, that of the water 402 in purple. Selected distances are highlighted in green.

(J. Labahn, personal communication) and hence confirm the power of this ssNMR-based strategy to monitor the potentially very specific structural changes associated with complexation at the level of individual residues in a seven-helix membrane receptor.

7.2.2 Chemical shift changes upon complex formation

Figure 7.4 compares (^{13}C,^{13}C) spectra of free NpSRII (black) and NpSRII in complex with NpHtrII (green). In both cases NpSRII was reverse labeled (U[^{13}C,^{15}N\(V,L,F,Y)]); in the complexed sample the NpHtrII was not labeled. The overall cross peak pattern is again largely conserved, reflecting structural similarity as expected. On the other hand, distinct chemical shift changes are seen. Based on the assignments of free NpSRII, several shift changes can readily be identified. In order to provide independent support
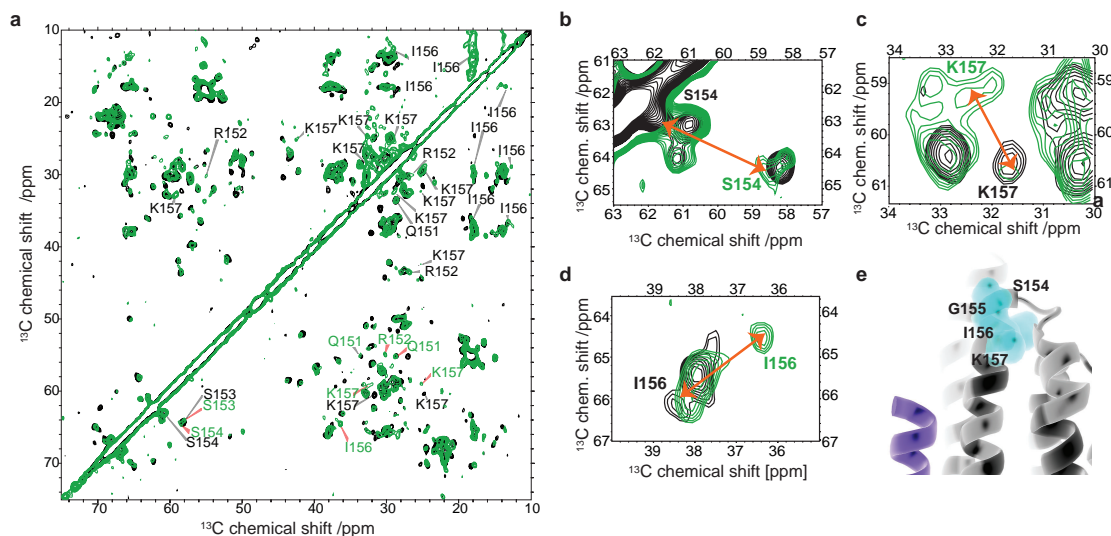
Figure 7.4: Chemical-shift changes seen upon complex formation on the receptor. a) $(^{13}C,^{13}C)$ SD spectra of free NpSRII (black) and NpSRII/HtrII complex (green). b-d) show extracts for residues in the EF-loop highlighted in e (purple helix represents C-terminal end of NpHtrII as seen in the crystal).

for the assignments of shifted peaks in complexed NpSRII, additional $(^{13}C,^{13}C)$ and $(^{15}N,^{13}C)$ correlation spectra were recorded (see supporting figure on page 124). As visible in figure 7.4c for the Lys157 $C\alpha$-$C\beta$ correlation, residual signal intensity at the position found in the free receptor is still present although a general peak shift could be observed. This suggests that either a small fraction of NpSRII is not bound to NpHtrII or that the bound configuration includes a smaller population which samples the free state. The spectra also contain shifted peaks that could not be identified sequentially due to overlap. These shift perturbations can still be assigned at an amino-acid specific level and reveal additional interactions for at least one isoleucine, arginine and proline, respectively.

Overall, most chemical shift perturbations can be explained by the binding interface as seen in the crystal structure. However, the ssNMR data reveal that also residues located in the loop connecting helix E and F of the receptor are subject to chemical shift perturbations. Crystallographic data of the binding interface in this region are missing, since the last 32 C-terminal residues of the transducer are not ordered in the crystallized

complex most likely due to packing interactions [148] (see figure 7.4e and supporting figure on page 125 for crystal packing). The ssNMR data hence suggest that in a lipid environment these missing residues interact with the receptor loop.

In general the chemical shift perturbations detected for the EF-loop could also be explained by allosteric static or dynamic interactions induced by other parts of the transducer. To distinguish between direct or indirect, dynamical or static interactions different experimental setups were used, which will be discussed in the following.

7.2.3 Changes in receptor mobility

In chapter 6 it was shown that T_2-filtered, scalar based (^{1}H,^{13}C) correlation experiments provide a sensitive means to study mobile protein segments of NpSRII under MAS conditions. Similar experiments were conducted here to probe changes in receptor dynamics upon complex formation. As discussed in 3.3.2 these experiments are sensitive to protein segments that exhibit molecular motion which is comparable to or faster than the inverse of the strongest (^{1}H,^{1}H) dipolar coupling (i.e. $10^{-5}s$). In figure 7.5a (^{1}H,^{13}C)-INEPT-based correlation spectra are shown for free (black) and complexed (green) NpSRII. Overall, a significant reduction in signal intensity of dynamic loops is observed. In particular, reduced dynamics for the BC-loop, as visible by reduced signal for Ala64, and for the DE-loop, as exemplified for Ile121, are seen. According to the X-ray structure both loops are not directly involved in the binding interface and show increased B-factors when comparing the crystallographic data from the free and NpHtrII bound NpSRII (figure 7.5b+d). A possible explanation for the disagreement between the increased B-factors and the decreased mobility measured by ssNMR could be alterations in the crystal-packing contacts as shown in figure 7.2a+c.

To investigate receptor motion in the kHz regime, dipolar (^{13}C,^{13}C) double-quantum correlation spectroscopy was used. For this purpose Cα-Cβ double-quantum buildup curves, which report on the effective dipolar couplings between the involved spins, were recorded by measuring a set of 2D (2Q,1Q) spectra on the free and complex sample varying the double quantum contact time [51]. The corresponding buildup characteristics for resolved residues in the EF-loop are shown in figure 7.5b-d for free (black)
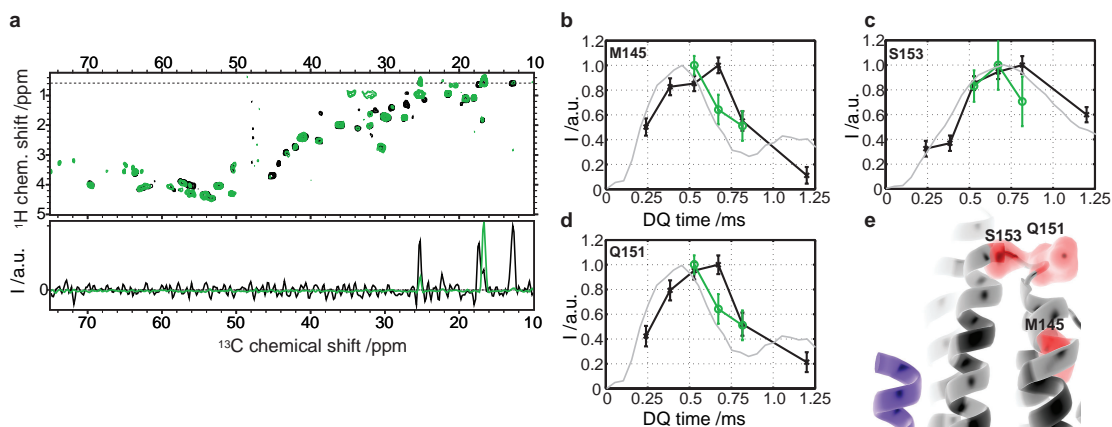
Figure 7.5: Analysis of changes in fast and medium-scale dynamics of NpSRII upon binding to NpHtrII (free black, complex green). a) Reduction in signal intensity of highly dynamic loops. Signal reduction is indicative of reduced dynamics for BC- and DE-loops, as exemplified for Ile121 in DE-loop. b-e) visualizes dynamical changes of the EF-loop in NpSRII upon NpHtrII binding. $C\alpha$-$C\beta$ double-quantum buildup curves for M145, Q151 and S153 respectively of the free (black) and complexed (green) NpSRII. Grey lines represent simulated behavior for the given spin system without dynamics ($S = 1$).

and transducer bound (green) NpSRII, and reveal dynamical changes in NpSRII upon NpHtrII binding for M145, Q151 and S153. (Experimental spectra are shown in the supporting figure on page 128.) Grey lines in figure 7.5b-d represent simulated behavior for the given spin system without dynamics (order parameter $S = 1$), considering the presence of (and chemical shift difference to) a $C\gamma$ spin as discussed in [51], which are close to the spin dynamics seen in case of the complex. In contrast, experimental data obtained for the free receptor are best described by dipolar order parameters ranging between 0.8 and 0.9. Within the given resolution of the double quantum experiment, residues outside the EF-loop show no reduced order parameter neither for the complex nor for the free NpSRII (see supporting figure on page 129).

Dynamical changes for the EF-loop upon complex formation are again not reproduced in the crystallographic data. Notably, close crystal contacts are present for several residues in the EF-loop in crystallographic data of free as well as complexed NpSRII (see figure 7.2a,c). The ssNMR data regarding molecular dynamics strongly corroborate

the interaction suggested by the chemical shift perturbation for the EF-loop and suggests that the transducer molecule has a stabilizing effect on the receptor.

7.2.4 The direct binding interface of the receptor

Dynamical and structural modifications in regions which are, according to crystallographic data, outside of the binding interface could in principle also be explained by indirect interactions induced by the contacts present in the crystal. To directly detect the molecular NpHtrII-NpSRII interface, a heterogeneous (^{15}N:^{13}C) labeled sample was used. Here the transducer molecule was uniformly but exclusively ^{15}N labeled, whereas the receptor was uniformly but exclusively ^{13}C labeled. For a better comparison to the previous data the four residues (Leu,Phe,Tyr,Val) were additionally reverse labeled in the receptor. As shown elsewhere [111] intermolecular proton-proton contacts in heterogeneous (^{15}N:^{13}C) labeled samples can be detected indirectly using the NHHC experiment [49]. In the case of NpSRII-NpHtrII in liposomes the expected contact interface accounts only for a small fraction of the total sample volume. Hence the intermolecular transfer efficiency is rather limited and does not allow for a two-dimensional detected correlation map. Instead, figure 7.6a+b shows a one dimensional NHHC spectrum of U-^{15}N NpHtrII-U[^{13}C\(F,L,V,Y)] NpSRII (black). The proton-proton contact time was set to $500\mu s$ to allow for a detection of direct contacts up to 5Å and to increase sensitivity. Notably reverse labeling reduces the number of receptor residues that can contribute to the NHHC signal in the interface. This significantly simplifies the spectral interpretation of the one dimensional experiment.

Figure 7.6c shows the residues which are, according to the crystallographic data, expected to contribute to the signal under the applied experimental conditions. A Mathematica script (see supporting information on page 126) was used to generate a predicted spectrum based on proton-proton distances measured in the X-ray structure [148]. Since during the relative long proton-proton contact time relay mechanisms can interfere with the distance information encoded in the peak intensity, a detailed analysis between the peak amplitudes and the corresponding distance was not attempted here. Instead based on the list of possible excited frequencies in the receptor a superposition of the cor-
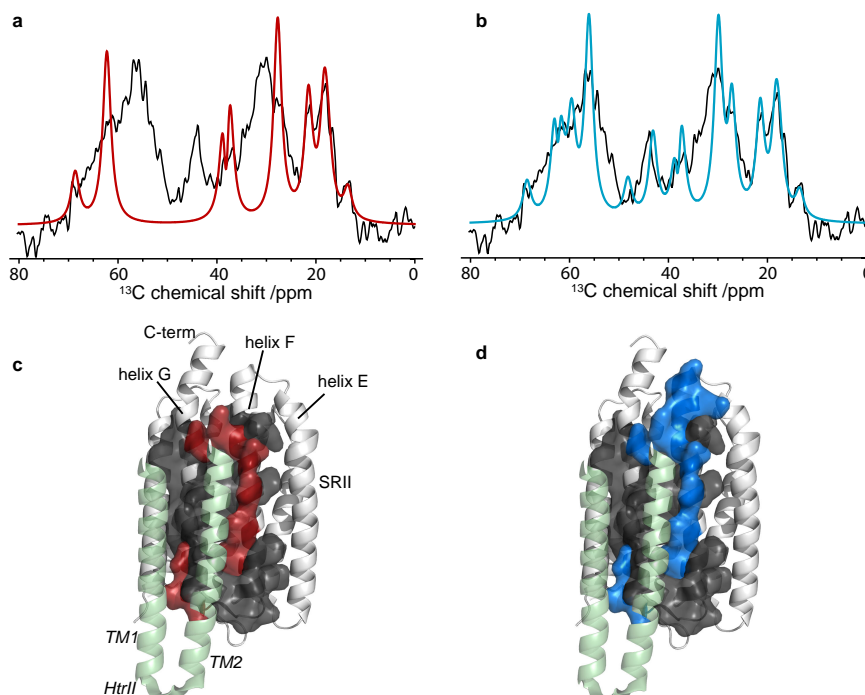
Figure 7.6: a+b) Experimental NHHC spectrum recorded on U-[^{15}N] NpHtrII-U[^{13}C\(F,L,V,Y)] NpSRII (black). Theoretical calculated spectra of intermolecular interfaces as suggested in c+d are overlaid in red and blue respectively. Dark grey residues in c+d are not labeled.

responding Lorentzian functions was fitted to the experimental data. The best fit for the molecular interface as seen in the crystal is given in figure 7.6a in red. Due to the high abundance of isoleucines in the crystal interface the methyl region (10-30ppm) can be nicely explained. However, owing to the large sidechains of the involved residues only very few backbone interactions are predicted, which lead to a large disagreement in the Cα region (45-60ppm).As discussed above, relay mechanisms may interfere with the chosen distance cut off and influence the presented results. The effect of the considered distance was investigated by additional calculations (see supporting figure on page 127) confirming the experimental interpretation.

Assuming a direct role of the EF-loop in the binding interface, residues highlighted in figure 7.6d, which would face towards an extended transducer, could additionally contribute to the NHHC spectrum. Figure 7.6b shows in blue the best fit of frequencies involved in such an elongated binding interface to the experimental data. The signifi-

cant improvement compared to the data based purely on the crystal structure strongly suggests close contacts between the transducer and residues in the EF-loop region of the receptor.

7.2.5 A structural model of the NpSRII-NpHtrII complex
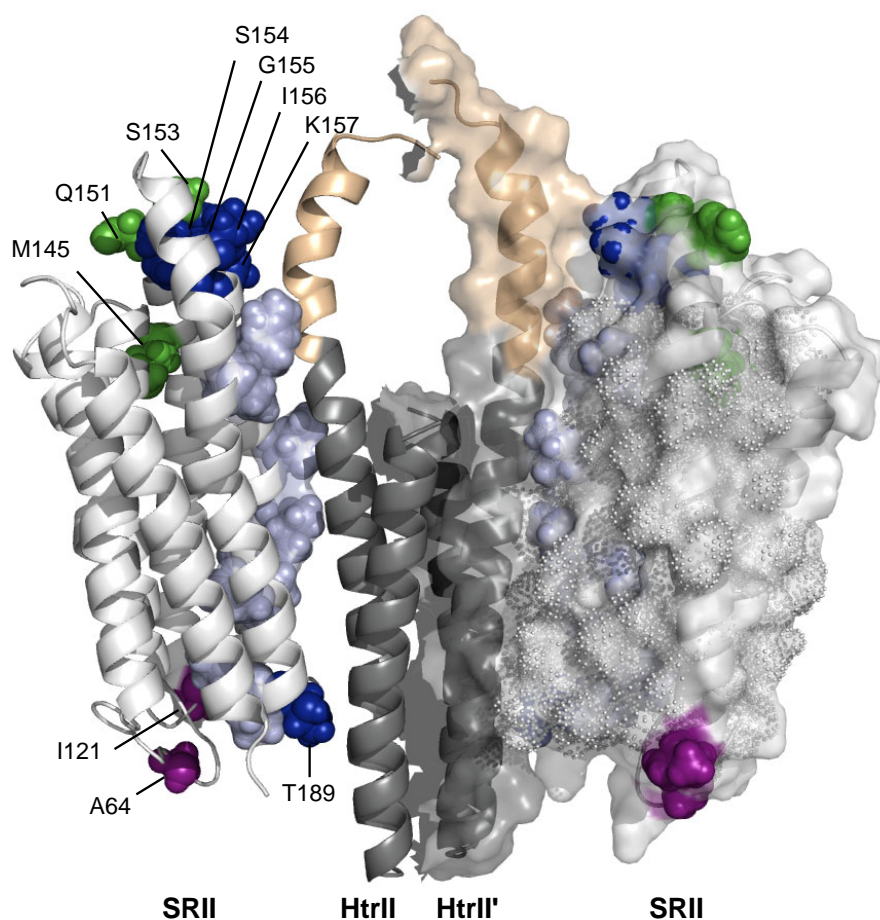
Figure 7.7: Structural model illustrating the ssNMR results. Highlighted HtrII region was calculated using CNS as discribed in the text. Grey region is as seen in the X-ray structure. Light blue and blue residues indicated residues with potential and identified chemical shift changes respectively. Green residues show increased order parameters upon HtrII binding. High mobility observed in the free receptor is not conserved in the complex for purple residues. Complex on the right hand side is given in surface mode and shows dots for not labeled residues.

Perturbations identified in the resonance frequencies, reduced dynamics observed for distinct residues and polarization transfer along the direct intermolecular interface consistently imply an elongated binding mode of the transducer molecule. This mainly leads to additional interactions involving receptor residues of the EF-loop. These findings are in good agreement with mutagenesis data, which reveal a critical role of Ser154 in complex formation [186] and EPR studies that measured intermolecular NpHtrII-NpHtrII and NpSRII-NpHtrII distances in doubly spin labeled samples [185]. Assuming a conservative error margin of 15 Å compensating for the relative long distance between the paramagnetic center of the MTSL tag and the protein backbone, the EPR distances were combined with the ssNMR data to develop a structural model of the NpSRII-NpHtrII interface. (*Model calculations of the complex contributed by K. Seidel.*) Here the ssNMR data were incorporated using arbitrary constrains based on the chemical shift changes detected and additional ssNMR results obtained on the NpHtrII molecule. Figure 7.7 shows the resulting structural model which, based on the X-ray structure, is also consistent with EPR and ssNMR data.

7.2.6 Light activation

Controversial results of the mechanism responsible for signal transduction of NpSRII to the NpHtrII molecule exist. In analogy to the rotary motion of helix F in bacteriorhodospin (BR), which was detected by cryo-electron microscopy [195], EPR spectroscopy [196, 197] and X-ray structural analysis [198, 199] it was concluded that a similar motion also triggers the activation of the transducer inducing a rotating or screw-like motion of TM2. Indeed EPR spectroscopy [200] and X-ray data [149] could detect the expected conformational change for TM2. However, X-ray and EPR data differ in the interpretation of the initiating signal of the receptor, which is responsible for TM2 rotation. Whereas X-ray structures of freeze trapped activated states show no significant conformational changes for helix F in NpSRII [149] but instead a shift of helix G with respect to helix F, initial and additional time resolved EPR data suggest a rotating motion of helix F [200, 184] (see supporting figure on page 130 for a comparision of SRII and BR in terms of atom position and torsion angle changes associated with activation).

Since the observed photocycle in the NpSRII/NpHtrII crystals largly agrees with the one in liposomes, structural changes associated with the retinal pocket appear to be valid in the crystal structure. However, the discussed packing artifacts may also interfere with signal transduction along helix F whereas EPR data were recorded under near physiological conditions but may suffer from artificially induced specific rearrangement of the spin label site chain [184]. SsNMR offers a spectroscopic means to study activation of a native receptor in a lipid bilayer environment. Previously, ssNMR has been used for the investigation of activation in rhodopsins using specific labeling of the chromophore, the Schiff-base or selected residues e.g. [201, 202, 203, 204, 205, 206, 207, 208, 209]. With only the specific labels, one spectral dimension for the encoding of resonance frequencies was in general sufficient for these studies. In the following the possibilities and limitations of an ssNMR structural study of uniformly labeled NpSRII activation are discussed.

Overall, two approaches are feasible to characterize the light activated state of a protein by ssNMR. The first is the use of a transparent rotor (e.g. sapphire) allowing the excitation of the sample inside the probehead using laser light transmitted via a fiber glass wire [201, 202]. The second is to freeze trap the light activated state of interest outside the magnet and insert and measure under sufficient low temperatures to conserve the active form. The first approach enables the investigation under similar conditions as used for the previous study. However, initial experiments revealed that the stability of the sapphire rotors used is very limited and led to the loss of two rotors including samples. Therefore in the following results obtained using the freeze trap approach are presented.

Figure 7.8 shows in black a ^{15}N CP spectrum of a NpSRII sample of the ground state which was labeled only on the $N\zeta$ position of the lysine side chain. The spectral assignment can be readily made, since only two lysines (K157 and K205) are present in NpSRII. The upfield shift of 27ppm is consistent with the statistical average found in the BMRB [48], whereas the lowfield shift of 180ppm agrees very well with experimental results obtained for the Schiff-base forming Lys in BR using a similar setup [203, 210]. The peak at around 120ppm is the natural abundance signal from the protein backbone. In figure 7.8 (red), the spectrum is shown under the same experimental setup, but recorded after light excitation outside of the magnet and freeze trapping using liquid
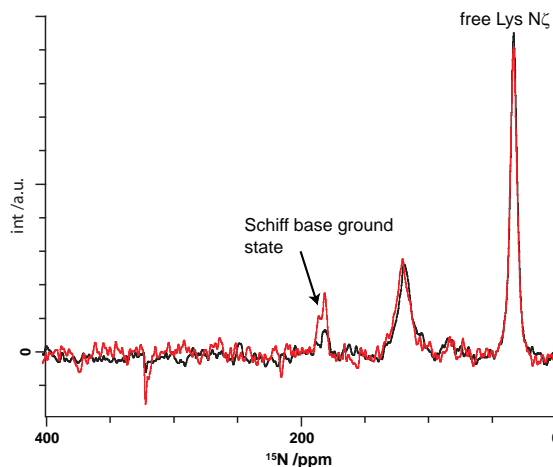
Figure 7.8: ^{1}H-^{15}N CP spectrum of [^{15}Nζ]Lys-NpSRII/NpHtrII in the ground state (black) and in light excited and freeze trapped state (red).

nitrogen (see supporting figure on page 131 for a more detailed description). Whereas the signal from the free Lys sidechain and the protein backbone shows the same intensity, signal arising from Lys205 $N\zeta$ is reduced by more than a factor of two. According to [210] signal from the activated state should be shifted to around 290ppm. Since in the M-state the Schiff base nitrogen gets deprotonated transfer efficiency of the cross polarization to the nitrogen gets reduced. Together with the generally low sensitivity the absence of a clear peak indicative for the M-state could be explained. However, the significant reduction in signal characteristic for the ground state strongly suggests the presence of a large fraction of an excited state. This is further supported by a characteristic color change of the sample, which is indicative for the M state (see also supporting figure on page 131).

Following these test experiments, a light activated sample of U[^{13}C,^{15}N\(V,L,F,Y)] NpSRII/NpHtrII was prepared and freeze trapped. Figure 7.9 a-d show (^{13}C,^{13}C) spectra of this sample in the activated state (b+d) and after storage at room temperature for 30min (a+c). The spectra in figure 7.9a+b are recorded using a mixing time $t_{CC} = 40ms$ to obtain intra residue correlations. In figure 7.9c+d a mixing time of $t_{CC} = 350ms$ was used to obtain sequential correlations even in the absence of weak coupling conditions.
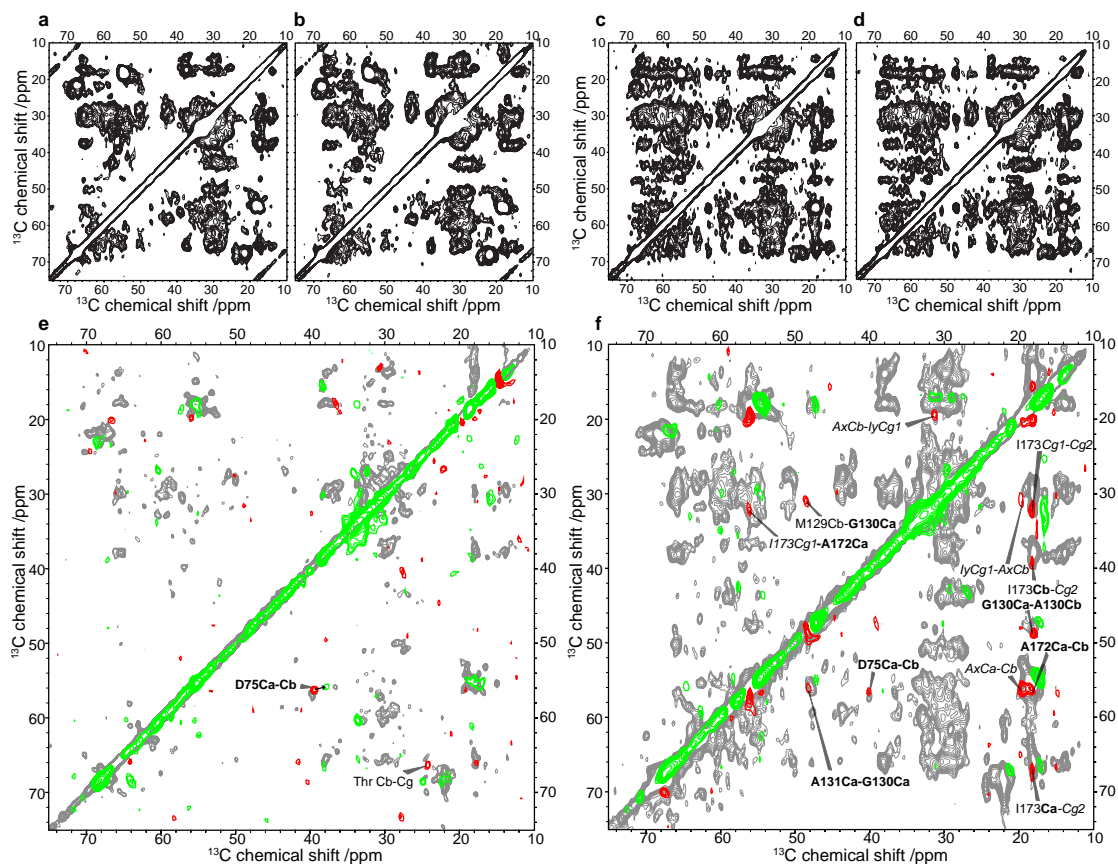
Figure 7.9: a-d (^{13}C,^{13}C) PDSD spectra using mixing times of $t_{CC} = 40ms$ (a+b) and $t_{CC} = 350ms$ (c+d) of ground (a+c) and light activated (b+d) U[^{13}C,^{15}N\(V,L,F,Y)]NpSRII/NpHtrII. e+f show difference spectra of spectrum a-b (in e) and c-d (in f). Red contours show positive values (signal that was stronger in the ground state) whereas negative values are given in green. Grey spectra in e and f are the same as shown in figure 7.4a (green) and figure A.13a (green) respectively and serve as reverence. As exemplified for D75Cα−Cβ a chemical shift change from the ground to the excited state can be readily identified. In f bold letters indicate peak assignment found in NpSRII and italic letters indicate spins which were not assigned before, but are most consistent with the data.

A comparison with the spectra recorded before reveals that the sample preparation (e.g. sample grinding) and/or the experimental conditions (e.g. low temperature, shock freezing) used for the freeze trapped samples reduces spectral resolution and hence complicate the spectral interpretation. Since the spectra of ground and excited state were recorded on the same sample using the same experimental setup ,the most straight forward way

of identifying spectral changes is to subtract one spectrum from the other. Such difference spectra (ground state − excited state) are shown in figure 7.9e+f. Signal that was present in the ground state and is shifted in the activated state will give positive values (shown in red), whereas the new positions of the signal should result in a negative peak (shown in green). Note that although the number of scans is identical for the respective spectra, the signal in the excited state is slightly stronger, leading to a predominant negative diagonal. Therefore the positive and negative contour levels were adjusted individually to match the actual noise level of the cross peaks. For a better comparison, the difference spectrum is superimposed on the (^{13}C,^{13}C) spectrum shown in figure 7.4a (red) used for chemical shift identification in the complex. Such a comparison readily leads to the identification of Asp75 Cα-Cβ, which shows a strong positive peak at the position assigned in the ground state of NpSRII. The close negative cross peak suggests that especially Asp75 Cβ is shifted by about 2ppm in the activated state. As already discussed in 7.2.1, the side chain of Asp75 is the proton acceptor of the Schiff base upon light activation. The identified chemical shift change confirms that two dimensional ss-NMR spectroscopy enables the possibility to investigate light induced structural changes in uniformly (reverse) labeled 7TM helices membrane receptors. Additional negative and positive cross peaks in figure 7.9e indicate structural changes involving at least one Thr, Ala and Ile residue. To sequentially identify the remaining chemical shift perturbations the spectra in figure 7.9c+d were recorded using a mixing time of $t_{CC} = 350$ms. Indeed their difference spectrum shown in figure 7.9f not only reproduces the intra residual peaks identified in figure 7.9e, but also enables the sequential assignment of perturbations for Gly130-Ala131 and strongly suggest effects for Ala172-Ile173 and an additional Ala-Ile pair. Interestingly Gly130 and Ala131 are located in helix E facing towards residues Ala172 and Ile 173 in helix F (see figure 7.10). The additional Ala-Ile pair does not fit to any assigned correlation. The only unassigned Ale-Ile pair is Ala212-Ile211 in helix G, which is according to the model shown in figure 7.7 close to the proposed molecular interface involving the E-F loop and TM2. Light induced changes in this area could be related to the change in hydrogen bond formation of Arg162 to D214 observed in the crystal [149]. As visible in figure 7.10, Gly130 and Ala131 are in close proximity to the retinal, hence a direct influence of the isomerisation or a coupled effect due to helix G
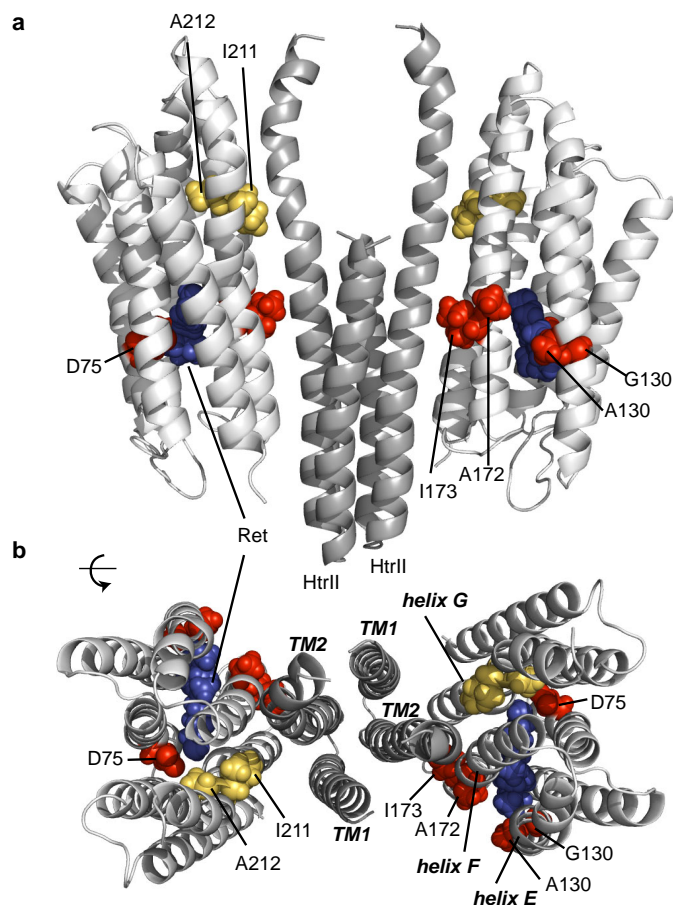
Figure 7.10: Structural model of the NpSRII/NpHtrII complex as shown in figure 7.7, but highlighting the chemical shift changes induced by light activation. Residues colored in red could sequentially be identified. Residues displayed in orange are most consistent with the data. Retinal is given in blue.

movement could explain the chemical shift changes. However, according to the crystallographic data helix E is largely unaffected by the movement of helix G. Therefore the chemical shift perturbations identified for residues in helix E,F and G could most easily be explained by a rotation of helix F, which would affect both neighboring helices, even in the absence of direct structural modifications in helix E (see also supporting figure on page 130.

7.3 Conclusion

This chapter demonstrates that ssNMR offers a spectroscopic means to study a reverse but uniformly labeled 2x 44kDa membrane protein complexes in lipid bilayers. Here the study was carried out in reference to existing X-ray structures, and effects of a membrane environment were investigated and yielded complementary information to the present structures. Not only structural and dynamical changes could be identified upon complex formation of NpSRII to its cognate transducer, but also light induced changes associated with receptor function could be followed using 2D (^{13}C,^{13}C) spectroscopy. In terms of complexation the ssNMR data suggest an elongated binding mode of the transducer involving the EF-loop of the receptor, which is in agreement with data reported by EPR [185] and mutagenesis [186]. Dynamical changes indicate that the receptor is stabilized upon complex formation and that intrinsic medium size flexibility in the transducer free EF-loop coincides with the binding interface. Both dynamical and structural changes identified with ssNMR are only partially reproduced by the crystallographic data and imply that the disagreement is strongest in regions with close intermolecular contacts in the crystal. Spectra recorded on freeze trapped activated states of the complex and data recorded on the D75N mutant show structural changes related to the retinal binding pocket. These changes and the perturbations associated with light activation found for residues in helix F are in line with crystallographic data. Unexpected chemical shift changes in helix E reveal the value of an uniform labeling approach and speak in favor for a rotation of helix F during activation.

8 | Chapter 8

Structural aspects of signal transduction in the histidine-kinase DcuS

8.1 Introduction

Bacteria are equipped with membrane integral sensors for rapid response to changing environmental conditions. The most common sensors systems are two component systems consisting of a sensor kinase and a response regulator triggering the cellular response which is mostly a change in gene expression [211, 212]. Most of the sensor kinases are membrane integral to allow for direct interaction with environmental stimuli. A big subgroup is represented by the periplasmic or extracellular sensing His kinases (HisKA)[211]. DcuS, the C_4-dicarboxylate sensor of Escherichia coli, is a member of these HisKA [213, 214, 215]. For realization of transmembrane sensing, DcuS is a multidomain protein with functional domains in the periplasm, within the membrane, and the cytoplasm. Periplasmic signal perception is achieved by a Per-Arnt-Sim domain (PAS_P). A membrane integral domain consisting of two transmembrane helices transmits the signal to a cytoplasmic region (figure 8.1). The latter comprises a second PAS domain (PAS_C) and the C-terminal transmitter or kinase region consisting of conserved

DHp (dimerization and HisP-transfer) and catalytic (HATPase) domains. PAS domains are common in membrane integral His kinases and can serve as (periplasmic) sensing domains such as in the prototypical (EnvZ/PhoQ/VirA) (see e.g. ref. [211, 212] for review) and the CitA/DcuS-like sensor kinases [215, 216]. The function of the cytoplasmic PAS domains, though present in about 33% of all membrane integral His kinases [212], is largely unknown. Multidomain periplasmic sensing His kinases have not been accessible to structural analysis of the full-length proteins and structural information could so far only be obtained from isolated domains that do not contain transmembrane segments (see e.g. [211, 212]). In general, the characterization of complex biological molecules including multi-domain protein complexes benefits from a combination of biophysical approaches [217, 218]. For example, both X-ray crystallography [218] and solution-state NMR [219] has successfully been incorporated into the structural study of multi-domain proteins and protein complexes. In parallel, computational biology provides increasing possibilities to determine molecular structure in reference to experimental data obtained using electron microscopy [220], X-ray crystallography [65] and NMR [64, 221, 222]. Combination of NMR and computational methods is believed to improve speed and accuracy of the structure determination process of small to medium-size proteins in solution. Still experimental validation of *in silico* predicted structures for proteins of unknown structure is mandatory [59, 223, 65]. In a membrane setting, such strategies are furthermore complicated on the experimental level by molecular size, intrinsic flexibility or the influence of the surrounding cell membrane [224, 225, 226]. While isolated membrane protein domains may be accessible to structure prediction and experimental analysis, the lack of domain interaction, of molecular dynamics, and of the surrounding cell membrane may influence the functional interpretation of structural aspects in the full-length protein.

In the following it is shown that combination of solid-state NMR (ssNMR) experiments with structural modeling routines provides a general strategy to study individual as well as membrane-embedded DcuS constructs (figure 8.1) in close reference to protein function. In a first set of experiments it is shown how ssNMR data can be rapidly analyzed using structural models derived from ROBETTA [69] to probe the 3D molecular structure of the isolated DcuS-[PAS$_C$] domain. Examination of ssNMR chemical shifts, peak

volumes and interatomic distances corroborates the validity of the modeled structures, refines areas of uncertainty and suggests that regions which are poorly defined in the structure prediction reflect intrinsic protein disorder. In a next stage, these results along with previous solution-state NMR data of the periplasmic domain [227] provide a valuable reference for a ssNMR-based structural investigation of the membrane-embedded multidomain construct. Compared to the structural findings for the individual domains, the membrane-anchored DcuS-[PAS_P/$TM_{1,2}$/PAS_C] construct exhibits structural alterations that are largely restricted to the terminal ends of the periplasmic domain. In addition, these results suggest that a disordered N-terminal helix identified in the isolated cytoplasmic PAS domain is also present in the membrane integral multidomain construct. Mutation of residues in this protein domain specifically influences function, suggesting a relationship between protein plasticity and activity. Hence, even in the absence of high-resolution structural information for the full-length construct, combination of experimental solid- and solution-state NMR data with *in silico* prediction provides insight into the functional aspects of a multidomain membrane protein.

8.2 Results

8.2.1 Structural characterization of DcuS-[PAS_C]

The primary sequence of DcuS-[PAS_C] is given in figure 8.2a. Attempts to solve its three-dimensional structure by solution-state NMR and X-ray crystallography thus far were unsuccessful. Alternatively, as shown in chapter 4 and ref. [228] protein precipitation allows for ssNMR-based structural studies under native-like conditions . Hence, two isotope-labeled samples of DcuS-[PAS_C] using different precipitation protocols and labeling schemes were prepared. The first sample (referred to as U-PAS_C) was uniformly ^{13}C and ^{15}N labeled and precipitated using Polyethylenglycol (PEG). Sample two (R-PAS_C) was uniformly labeled with ^{13}C and ^{15}N, but contained for reverse labeling [55, 56] the six amino acids Phe, Ile, Lys, Leu, Arg and Val in natural abundance (i.e. U[^{13}C,^{15}N \(F,I,K,L,R,V]) and was precipitated by changing the pH to the isoelectric point (here pH=4.6). As described elsewhere [37], a *de-novo* structural analysis of ssNMR DcuS-[PAS_C] samples would require, in the first place, sequential resonance assignments along
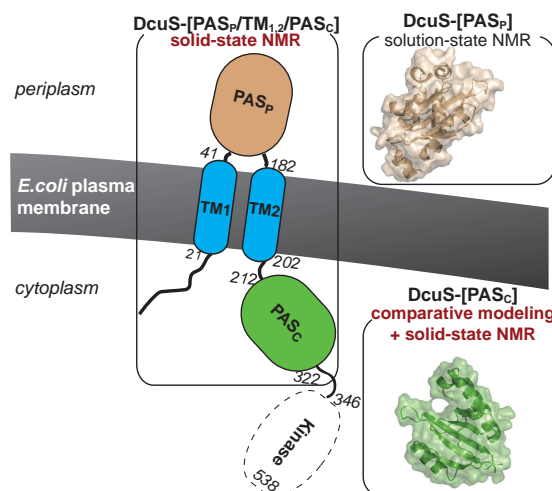
Figure 8.1: Model of the *E. coli* C_4-dicarboxylate sensor kinase DcuS inside the cytoplasmic membrane with PAS_P, $TM_{1,2}$ PAS_C and kinase domains. Isolated domains like the periplasmic domain (DcuS-[PAS_P]) are accessible to solution-state NMR (1ojg.pdb20) [227] and/or X-ray crystallography. Solid-state NMR offers a way to study membrane embedded multidomain constructs. In combination with comparative modeling, solid-state NMR may be also the method of choice for the study of insoluble and poorly crystallizable domains, which is the case for the isolated cytoplasmic PAS domain (DcuS-[PAS_C]).

the protein sequence. However, the three-dimensional structures of Per-Arnt-Sim (PAS) domains are in spite of a low sequence identity highly conserved [229]. To facilitate the ssNMR-based structural analysis comparative models that resulted from usage of the fully automated ROBETTA server [69] were generated. In general, PAS domains comprise four structural attributes [230]: (i) an N-terminal cap ($\alpha 1$), (ii) a PAS core ($\beta 1$, $\beta 2$, $\alpha 2$, $\alpha 3$, $\alpha 4$), (iii) a helical linker connecting PAS core and β-scaffold ($\alpha 5$), and (iv) the β-scaffold ($\beta 3$, $\beta 4$, $\beta 5$). The modeled structures of DcuS-[PAS_C] converged to a specific PAS-fold with an α-helical N-terminal cap. The orientation of the N-cap differs most strongly among the five modeled structures (see supporting figure on page 132 for an alignment of the modeled structures). In the context of ssNMR experiments, such structural disorder can lead to signal attenuation, changes in resonance frequency and impacts the ssNMR measurement of through-space distances. In line with earlier work [228] the ssNMR data are discussed in the following in close reference to ssNMR reso-

nance frequencies (i), cross-peak amplitudes (ii) and proton-proton distances detected indirectly (iii).
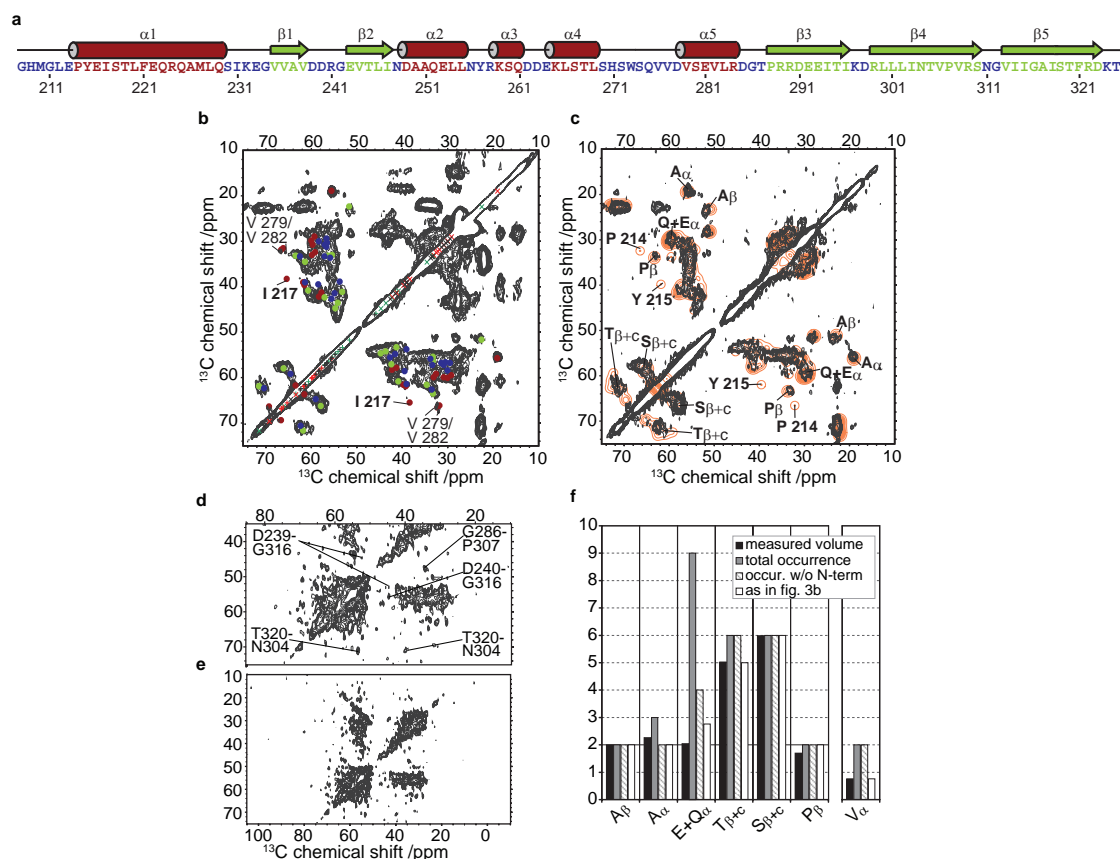
Figure 8.2: Experimental results obtained on isolated DcuS-[PAS$_C$]. (a) DcuS-[PAS$_C$] primary sequence and secondary structure according to PSIPRED [81]. (b) (^{13}C,^{13}C) SD spectrum of U-PAS$_C$. Circles indicate Cα-Cβ peak position based on the predicted secondary structure (according to ref.[231]; color code as in a). (c) (^{13}C,^{13}C) SD spectrum of R-PAS$_C$ (black). A simulated spectrum based on the peak positions according to shiftX [232] of the first modeled structure is superimposed. Specific residues and amino acid types used during the analysis are labeled. (d) CHHC spectral extract of R-PAS$_C$. Long range contacts that are unambiguous based on the ROBETTA models are indicated. The full spectrum is shown in (e) and shows the noise level of the experiment. (f) Cα-Cβ cross peak volume analysis for regions with spectral overlap but residue type specific resolution. The bars represent: black = experimental determined peak volume (normalized); grey = total abundance in DcuS-[PAS$_C$]; line pattern= abundance without residues in N-terminal helix; white = contribution for each residue as exemplified in figure 8.3b (see also text for details).

(i) ssNMR resonance frequencies: Figure 8.2b compares an experimental (^{13}C,^{13}C) spin diffusion (SD) spectrum of U-PAS$_C$ to the expected Cα-Cα peak position (given as colored circles) assuming a secondary structure as predicted by PSIPRED [81] (peak positions are taken from ref.[231]). In general Cα and Cβ resonance frequencies are very sensitive to the local secondary structure of a protein [233]. With the exception of Ile217 that is not present in the spectrum of U-PAS$_C$, the remarkable agreement between predicted and measured cross-peak signals strongly supports the validity of the predicted secondary structure. To reduce spectral overlap, we recorded the same (^{13}C,^{13}C) SD spectrum using the R-PAS$_C$ sample (figure 8.2c). This approach facilitates the analysis of spectral regions characteristic for Tyr, Pro, Glu and Gln (C$alpha$-Cβ) correlations. Overall, the signal pattern agrees with data seen for U-PAS$_C$, suggesting that the two precipitation protocols lead to a similar 3D fold of DcuS-[PAS$_C$]. Based on the comparative models a prediction of the peak position using the program shiftX [232] is possible (figure 8.2c, orange). Again the experimental spectrum is in very good agreement to the predicted correlation pattern. In addition to Ile217, figure 8.2c suggests the absence of cross signals for Tyr215 and Pro214 in the experimental spectrum. Notably, all missing residues are located in the N-terminal helix. For further analysis, a (^{13}C,^{13}C) spectrum revealing sequential correlations [23] was recorded (see supporting figure on page 133). Knowledge of the amino acid sequence and the secondary structure, readily leads to pair-wise assignments in several segments of the protein. On the other hand, multiple peaks for a single correlation could be identified for a few residues, e.g. Ser310, Asn311 and Gly312 (see supporting figure on page 133). Peak splitting is in general indicative of polymorphism, i.e. the occurrence of a residue in different configurations. Interestingly, previous molecular dynamic simulations [234, 235] have shown that among the investigated PAS domains and with the exception of the N-terminal cap, flexibility (reflected by the Cα positional shift) was highest for residues connecting $\beta4$ and $\beta5$ (corresponding to residues around Gly312 in DcuS-[PAS$_C$]). Interestingly, the sequential correlation spectrum (supporting figure on page 133) also revealed weak helical Pro Cα-Cδ cross signal suggesting residual α-helical propensity in the N-terminal cap. Figure 8.3a summarizes the results based on the analysis of the peak position on the average ROBETTA structure. The radius of the ribbon is scaled according to the residue-specific rmsd between

all five predicted structures.

(ii) ssNMR cross-peak amplitudes: Within the errors of the predicted peak position, the theoretical spectrum (figure 8.2c, orange) also contains information about the expected peak height in regions of sizable spectral overlap. Comparing the expected and experimentally detected peak intensities one finds major differences in the region characteristic for helical Glu and Gln ($C\alpha$-$C\beta$) cross signals (see also supporting figure on page 134 for a difference spectrum). A more detailed analysis of the peak volume (figure 8.2f) supports this finding and reveals that best agreement between experimental and predicted peak volumes is obtained assuming a mobile and/or unstructured N-terminal helix. In addition resolved but weak helical Val $C\alpha$-$C\beta$ peaks in U-PAS$_C$ and the remaining difference in expected and measured peak volume point to reduced contribution for residues located in the PAS core and the helical linker, indicative for reduced order parameters. The most consistent model is shown in figure 8.3b.

(iii) ssNMR proton-proton distances: As an additional independent means to study the tertiary structure of DcuS-[PAS$_C$] we subsequently conducted a series of CHHC correlation experiments [49] on R-PAS$_C$ and U-PAS$_C$ that indirectly probe short proton-proton contacts (figure 8.2d,e and supporting figure on page 137). For the fully labeled sample, a short ^{1}H,^{1}H contact time t_{HH} of $100s$ was used to detect only the shortest contacts (< 3Å). Since the number of visible contacts is highly reduced in the reverse labeled sample a longer contact time of $t_{HH} = 400s$ could be employed. The resolution of the spectra combined with the limited resonance assignment accuracy based on shiftX for larger fractions of the protein still complicates unambiguous *de novo* assignments of specific long range contacts. However, assuming the validity of the modeled PAS fold, which is further supported by our ssNMR data presented in (i) and (ii), the spectrum shown in figure 8.2d contains long range distance correlations in isolated regions that are unambiguous based on the predicted fold. The highlighted correlations (figure 8.2d and figure 8.3c,d) indicate that DcuS-[PAS$_C$] precipitates adopt a stable overall fold that closely resembles the modeled structure.

Following the manual identification of correlations based on the predicted structure and the approach used in ref.[42], an automated protocol was developed to assemble a list of constraints which are most consistent with all ssNMR data. The result of

a conventional X-PLOR structure calculation based on these constraints is shown in figure 8.3c,d (see methods, ref.[42] and supporting figures on page 137 and 135 for more details).
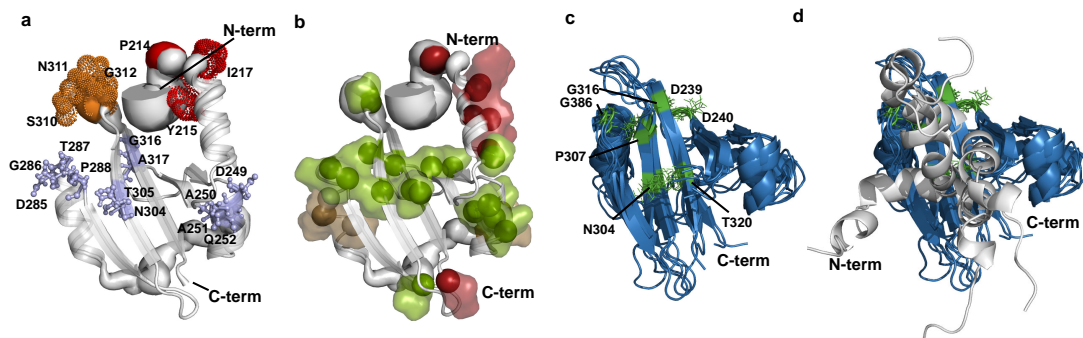
Figure 8.3: Graphical representation of the ssNMR experimental results on DcuS-[PAS$_C$]. (a) and (b) show comparative models of DcuS-[PAS$_C$] as obtained by ROBETTA[69]. Ribbon radius represents the Cα rmsd between the 5 resulting structures. (a) highlights results of the ssNMR analysis of the peak position. Red residues are missing, orange residues show presumably multiple configurations and blue residues show isolated single strong sequential cross peaks indicative for a well folded region. (b) shows results based on the peak volume. The color code represents the most consistent contribution of the respective residue to the peak volume of an only residue specific resolved cross peak. (color code: green/brown/red=full/reduced/no contribution). (c) and (d) Superposition of the 5 lowest energy structures of DcuS-[PAS$_C$] back calculated from an extended strain (see text and supporting figures on page 137 and 135 for details). Residues identified in figure 8.2d are highlighted in green. The N-terminal cap (residues 208-229) is not shown c and indicated in white in d.

Analysis of ssNMR peak positions, amplitudes and CHHC correlations hence lead to a self-consistent view of the 3D fold of DcuS-[PAS$_C$]: A well defined core, predominantly formed by the β-sheet and an unstructured/mobile N-terminal part involving mainly residues 208-229. Disordered protein regions identified by ssNMR correlate with PAS segments exhibiting enhanced rmsd values according to the ROBETTA models.

8.2.2 Structural characterization of DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$]

To study the effect of domain interactions and membrane anchoring, solid-state NMR measurements on the 41.4 kDa multidomain DcuS construct shown in figure 8.1 (DcuS-
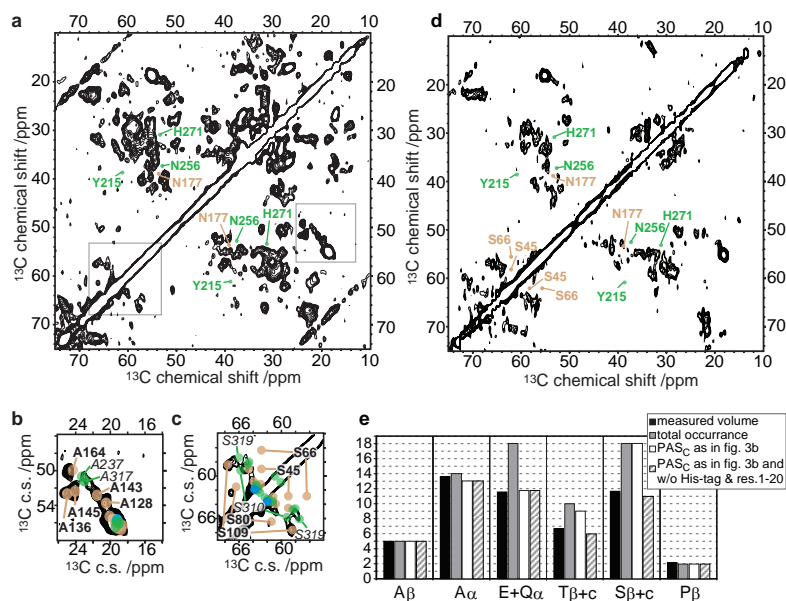
Figure 8.4: SsNMR results obtained on DcuS-PAS$_P$/TM$_{1,2}$/PAS$_C$]. (a) (^{13}C,^{13}C) SD spectrum of membrane embedded U-DcuS. (b) Magnification of the Ala Cα-Cβ and (c) Ser Cα-Cβ region. Resonance assignments as reported for soluble DcuS-[PAS$_P$] [227] are shown as brown circles. Green and blue circles are predicted peak positions for residues in the PAS$_C$ and TM$_{1,2}$ domains respectively. (d) (^{13}C,^{13}C) SD spectrum of R-DcuS. (e) Peak volume specific analysis of R-DcuS as discussed in the text.

[PAS$_P$/TM$_{1,2}$/PAS$_C$]) were performed after reconstitution in *E. coli* lipids. Figure 8.4a shows a (^{13}C,^{13}C) SD spectrum of uniformly ^{13}C and ^{15}N labeled DcuS-[PAS$_P$/TM$_{1,2}$ /PAS$_C$] (U-DcuS). The spectrum is characteristic for a well folded protein with well dispersed ssNMR signals and ^{13}C line width below 1 ppm. As exemplified in figure 8.4b for the Ala Cα-Cβ cross peaks, the peak positions assigned for the sensory domain in solution [227] (brown circles) fit remarkably well to the experimental spectrum of the membrane-embedded construct. This strongly indicates that the structure of the periplasmic domain found in solution is to a large extent conserved (see also supporting figure on page 139 for an analysis of all residues). Within the given resolution two assignments (Ser45 and Ser66 Cα-Cβ) seem to be remarkably changed (figure 8.4c). Ser45 is the first N-terminal residue found in solution of isolate PAS$_P$. Ser66 is part of the loop following the N-terminal helix in the periplasmic domain. The observed change

for Ser45 suggests that the connection of the terminal ends to the transmembrane parts indeed modifies the local structure. The changes for Ser66 could be explained by a different orientation of the N-terminal helix or by dimer formation which is expected according to X-ray data on the homologous domain of CitA [236], especially affecting the N-terminal helix as well as the following loop. However, ssNMR signals of the remaining residues highlighted in figure 8.4c fit again very well to the spectrum. To reduce spectral overlap and to facilitate spectroscopic comparison to results obtained on isolated DcuS-[PASC], additional ssNMR spectra were recorded using a sample of DcuS-[PAS_P/$TM_{1,2}$/PAS_C](R-DcuS) with the same reverse labeling as for R-PAS_C. Figure 8.4d shows the corresponding (^{13}C,^{13}C) SD spectrum. Reduced overlap in R-DcuS facilitates sequential assignments for several residues throughout the protein sequence (see supporting figure on page 139). For example, secondary chemical shifts and amplitudes for residues assigned in TM_1 and TM_2 (S26, A27, Y40, G190, M191, G197, T198 and C199) are in line with a well ordered α-helical configuration within the membrane as illustrated in figure 8.5. In addition, a comparative analysis of the spectra of U- and R-DcuS (see supporting figure on page 139) suggest, additional missing signal sets (see labels figure 8.4a,d) in both PAS domains. Due to increased spectral overlap and the presence of lipid signals (as discussed in supporting figure on page 139) their interpretation is less reliable. Nevertheless, residues affected in the cytoplasmic PAS_C domain relate to Tyr215 in the N-terminal cap and Asn256 and His271 in the PAS-core. For both regions, reduced signal intensity was already observed for the isolated domain (see figure 8.2 and figure 8.3). In the PAS_P domain signals assigned for Asn177 in solution seem to be shifted, indicative of structural alterations in the PAS_P/TM_2 connecting region.

In figure 8.5a, residues that were identified based on well isolated sequential or intra-residue cross peaks are highlighted in dark and light blue, respectively. Residues indicated in yellow are found in overlapping regions, but cross peak positions are consistent with spectra in the isolated domains. Residues not consistent with solution-state NMR data on PAS_P or fully structured/ordered PAS_C as discussed above are given in purple. In a corresponding structural model (shown in figure 8.5b), chemical-shift perturbations found for the PAS_P/TM_1 connecting region, which is significantly bent in solution, suggest a different relative orientation of the TM domains to each other, which is in
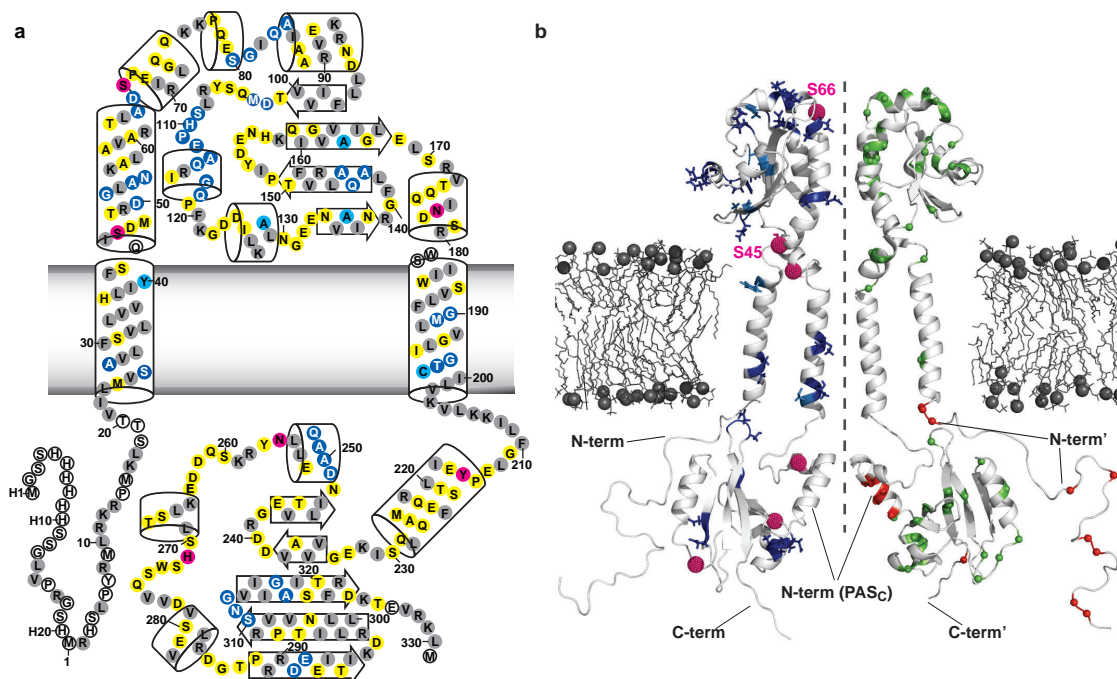
Figure 8.5: Graphical representation of the ssNMR experimental results on DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$]. (a) Schematic illustration of the multidomain, membrane embedded, 351aa construct. Color coding reports on consistency of ssNMR signal sets in U- and R-DcuS compared to values observed or predicted for the isolated domains. Dark/light blue residues have isolated inter/intra residue correlations where expected. Yellow residues are in overlapping regions, but are consistent with the spectra. Peak positions for residues indicated in purple are not present at the expected position (see also supporting figure on page 139). Grey/white residues are not labeled in R-DcuS or not incorporated in experimental/modeled data and are not considered here. (b) Structural model build from the isolated domains. Blue and purple residues (as in a) are highlighted on the left handed structure. Results of the volume analysis are mapped on the right structural model. As described in the text, red residues are unstructured and or mobile and do not contribute to the signal. Green residues are fully present. Orientation of the periplasmic domains of both monomers is adjusted according to the X-ray structure of the homologous CitA dimer [236]. Note that the relative orientation of the TM domains is likely an artifact of the bended terminal helices of the DcuS-[PAS$_P$] solution structure.

agreement with crystallographic data on CitA [236]. Note that, as in figure 8.5a, blue and purple residues are highlighted on the left handed structure in figure 8.5b. The observed absence of cross-peak signals for residues in the cytoplasmic PAS$_C$ domain

(figure 8.4 and figure 8.5) already suggest, that the disorder found in the isolated domain is also present upon membrane anchoring. While the presence of the other domains in the ssNMR spectra interferes with a further investigation of the characteristic peak positions described above, the large abundance of Glu and Gln in the N-terminal part of the PAS_C domain still enables the analysis of the peak amplitudes. We note that DcuS-$[TM_{1,2}/PAS_C]$ and DcuS-$[PAS_P/TM_{1,2}]$ constructs formed no stable membrane integral product and could not be used for ssNMR studies. Indeed, the measured peak volume (figure 8.4e; black bars) does not agree with the abundance of helical Glu and Gln residues in DcuS-$[PAS_P/TM_{1,2}/PAS_C]$ (figure 8.4e; grey bars) if secondary structure elements are used as indicated in figure 8.5a. The most plausible explanation for the observed difference found for Glu and Gln residues is structural disorder in PAS_C. In fact, a contribution of PAS_C, as seen in the isolated domain (figure 8.3b), leads to good agreement between measured and predicted signal intensities (figure 8.4e; white bars) suggesting that the disorder found for the N-terminal helix of the cytoplasmic PAS_C domain is not restrained by the connection to the transmembrane helix.

Notably, a similar discrepancy is found for Thr and Ser residues with resonance frequencies typically found for coil or β-strand conformations. These differences can be readily attributed to the presence of a disordered N terminus which is equivalent to the absence of the first 20 residues plus the 20 residue His-tag (figure 8.4e, line pattern) in our ssNMR spectra. As in figure 8.3b, results of the volume analysis are mapped on the structural model on the right (figure 8.5b) where red residues are unstructured and or mobile and do not contribute to the signal while green residues are fully present.

8.2.3 Structure-function relationship in the cytoplasmic PAS domain

The function of DcuS, or the capability for autophosphorylation and transcriptional activation of target genes can be tested in vivo by measuring the expression of the *dcuB'-'lacZ* reporter gene. Expression of *dcuB'-'lacZ* depends on DcuS and reports the functional state of DcuS in the DcuSR two component system [215, 237, 238]. To investigate the functional role of the cytsoplasmic PAS_C domain for gene expression a mutagenesis study was performed.(*Mutagenesis data are contributed by P.Dünnwald and*

Prof. G. Unden.) figure 8.6a displays the activities of the mutants in the cytoplasmic PAS$_C$ domain without stimulus (grey bar) and in the activated state in the presence of fumarate (black bars). Most prominent effects are found for replacement of Asn248 by Ala, Asp or Gly and of Asn304 by Asp which caused high constitutive activity of *dcuB'-'lacZ* expression, whereas wild-type and other DcuS mutations (including N248S) required C$_4$-dicarboxylates for full induction.

How can these findings be related to our spectroscopic view of a disordered N-terminal helix? A structural analysis of the dimeric crystal structure of AvNifL-PAS [239] which was identified by ROBETTA to be most homologous to DcuS-[PAS$_C$] offers additional insight. In figure 8.6b-d, two versions of the modeled structure of DcuS-[PAS$_C$] (blue and green) are aligned to the dimeric crystal structure AvNifL-PAS [239]. In the dimeric model of DcuS-[PAS$_C$], the N-terminal helix plays a crucial role for the dimer interface and the two mutation sites (indicated as red sticks on both monomers) that lead to C$_4$-dicarboxylate independent activation are found near to the *alpha*-helical N-terminal cap. Residue Asn248 is conserved and forms in both proteins intramolecular hydrogen bonds to the N-terminal helix, which to a large extent determine the orientation of this helix (figure 8.6c and supporting figure on page 135). Removal of these interactions, as in the case of N248A, N248G and N248D leads to protein activation, which is not observed for the N248S mutation that can partially restore hydrogen bonding. Notably, residue Asn248 is the most conserved residue in PAS domains [230] and within PAS$_C$ of the DcuS/CitA family. The corresponding residue (Asn34) of the aerotaxis sensor Aer of *E. coli* is required for signal transduction as well [240], although the protein is a chemotaxis sensor with completely different domain composition. According to the arrangement shown in figure 8.6b,d the second mutation site (Asn304) of one monomer reveals close intermolecular proximity to the side chain of Lys232 of the second monomer. Since Lys232 is not conserved in the homolog, crystallographic data of a possible contact are missing. However, the effects of an Asn304 to Asp304 mutation could very well be explained by an intermolecular interaction with the Lys232 side chain.

Comparison to the ssNMR data suggest that protein activation of DcuS is intrinsically related to increased molecular disorder as observed for the N-terminal helix in the membrane anchored and precipitated DcuS-[PAS$_C$]. In this respect the data obtained

on isolated WT DcuS-[PAS$_C$] imply that additional factors are necessary to stabilize a dimer interface as found for AvNifL-PAS. Indeed, recent biochemical results for the system *in vivo* have shown that DcuS is permanently in the active state, even in the absence of C$_4$-dicarboxylates, when the fumarate/succinate antiporter DcuB of anaerobic fumarate respiration [241, 216] that functions as an additional signal input site of DcuS (Kleefeld *et al.* submitted for publication), is lacking. It is hence tempting to assume that DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$] studied here most likely reflects the activated state and that mutations in the *in-vivo* case lead to destabilization of a potential PAS$_C$ dimer interface as a prerequisite for activation of kinase activity.
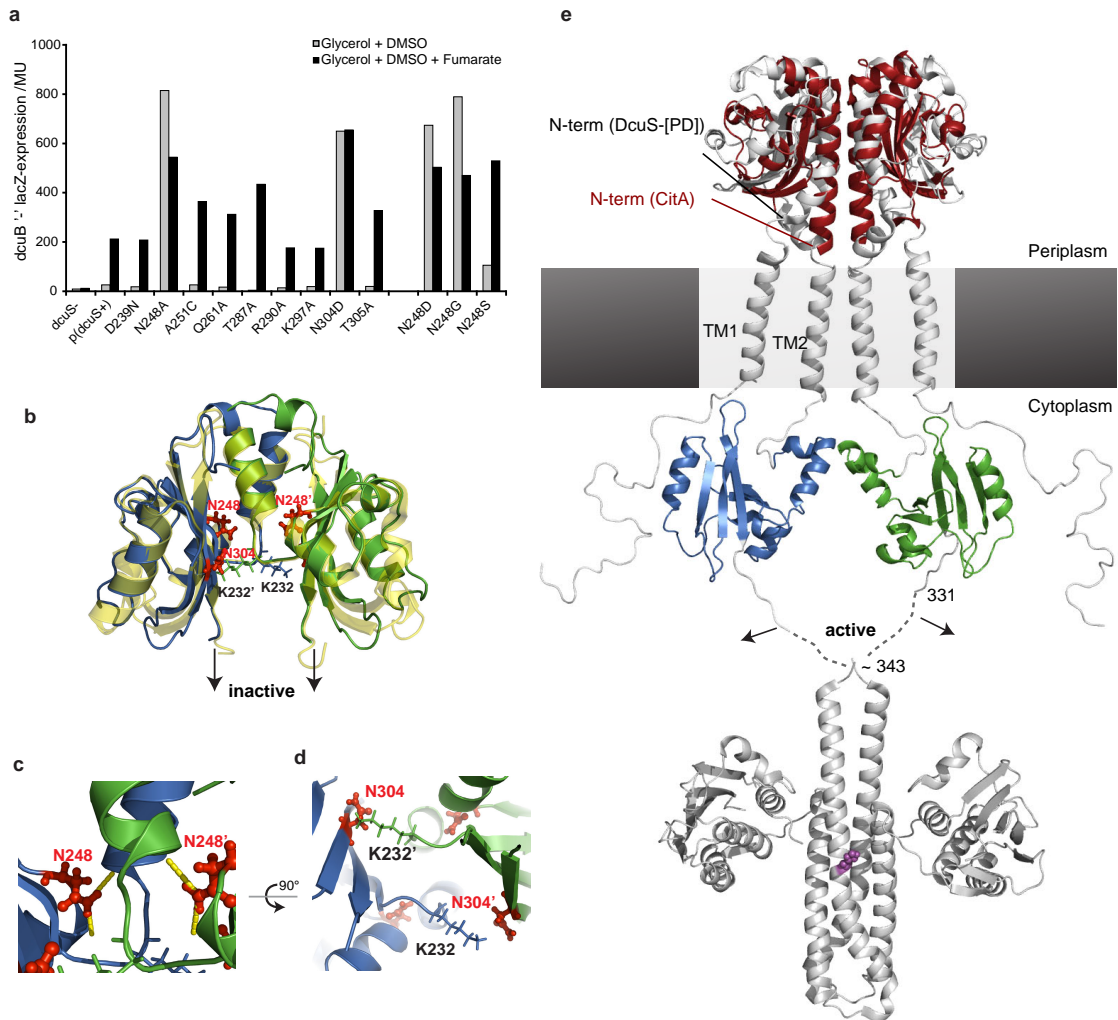
Figure 8.6: Functional effects of mutations in the cytoplasmic PAS domain and their potential structural consequences. (a) *In vivo* activation (*dcuB-lacZ*-expression levels) of *E. coli* IMW260 (*dcuS* deficient) and *dcuS* on plasmid (wild-type or mutated). Bars represent expression of *dcuB-lacZ* in IMW260 with a control plasmid, and a plasmid encoding wild-type (DcuS+) and mutant forms of DcuS after activation without (grey) and with (black) fumarate. (b) Two versions of the first ROBETTA model (blue and green) are aligned according to the dimeric crystal structure of AvNifL [239] (transparent yellow). Mutation sites that show activation in the absence of stimulus are shown in red and highlighted in c,d. (c) Asn248 forming hydrogen bonds to the N-terminal helix. (d) Asn304 exposes in this model close proximity to a lysine side chain of the other monomer. (e) Structural model of PAS$_C$ mediated signal transduction in full-length DcuS. On the periplasmic side two momomeric solution-state NMR structures of DcuS-[PAS$_P$] [227] (white) are aligned according to the X-ray structure of dimeric CitA [236] (red). As discussed in the text the relative orientation of the TM domains is probably an artifact. Combining the mutagenesis, comparative modeling and ssNMR data a plausible model of signal transduction involving the cytoplasmic PAS domain is shown in the lower half in e. As illustrated dimer disassembly of the PAS$_C$ domains upon activation/mutation would significantly alter their C-terminal extensions (see also b) and have a considerable effect to the consecutive kinase domains (here 2C2A.pdb [242]; phosph-acceptor residue His260 is highlighted in purple.)

8.3 Discussion

Understanding the mechanisms by which membrane-embedded sensor kinases recognize signals and regulate kinase activity is a key issue in prokaryotic signal transduction. While structural information has been obtained for individual protein domains, in particular regarding the periplasmic sensing domain of DcuS [227] and the cytoplasmic kinase domain of a membrane bound sensor His kinase [242], structural insight into multidomain, membrane-embedded sensor kinases has thus far been limited.

Here, solid-state NMR was applied to study individual segments (DcuS-[PAS_C]) as well as membrane integral (DcuS-[PAS_P/$TM_{1,2}$/PAS_C]) constructs of DcuS in close relation to functional data. For the isolated PAS_C domain, combination of ssNMR data with structure prediction resulted in the first experimentally validated 3D structure of a precipitated protein that is characterized by a well defined core with an N-terminal cap that is structurally/dynamically disordered. Using solution and solid-state NMR data obtained on the isolated DcuS-[PAS_P] and DcuS-[PAS_C], respectively, a structural ssNMR analysis of the membrane embedded construct was possible. This approach led to resonance assignments in each protein domain, including the transmembrane region. Compared to the isolated domains, our data suggest that the 3D protein domain structures are largely conserved in the membrane-embedded construct. However, chemical shift perturbations point to structural alterations in the PAS_P-TM connecting region and the disordered state of the N-terminal cap found in precipitated DcuS-[PAS_C] is conserved in the presence of the membrane-embedded TM_2 helix. Previous studies have shown that the N-terminal cap in PAS domains exhibits low sequence conservation and can be removed from an isolated PAS domain of the PYP protein without perturbing the rest of the structure [235]. Still, in concordance with the multidomain membrane related aerotaxis receptor Aer a critical functional role of the N-terminal cap was found [243]. In our case, mutations around the N-terminal cap of DcuS-[PAS_C] strongly affect protein activity. These findings identify the PAS_C domain and its intrinsically disordered N-terminal helix as a critical mediator in transmembrane signaling and protein activation. In this process, the side chain of Asn248 largely determines the orientation of the N-terminal helix (figure 8.6b,c and supporting figure on page 135). Mutation of

Asn248 which transforms DcuS to the constitutive active state of the kinase, can lead to an N-terminal helix that is not longer anchored to the rest of the PAS_C domain. This non-restrained form of $\alpha1$ in PAS_C potentially mimics the active configuration obtained in wild-type DcuS upon binding of the stimulus by PAS_P.

According to X-ray, AvNifL-PAS [239], which was used for comparative modeling of DcuS-[PAS_C], forms dimers involving the N-terminal helix (figure 8.6b). Indeed, PAS domains are known to form dimers and exhibit multiple protein-protein interactions [244]. In particular, dimer formation with other PAS domains is believed to represent a near-universal feature of signaling to downstream domains [230] and was recently structurally identified for KinA PAS-A [Lee et al., Biochemistry 2008, in press]. Based on the comparative modeling a similar dimer as found for AvNifL is expected for DcuS-PAS_C.

Mutation analysis (Dünnwald & Unden, unpublished) and search for cofactors [245] gave no indication for PAS_C functioning as an additional signal input site. Mutagenesis, however, strongly suggests the relevance of the Asn248 and Asn304 side chains for dimer formation and that interactions at the dimer interface are necessary to form the inactive state. In AvNifL-PAS, the N-terminal helix is further stabilized by additional hydrogen bonds which are not conserved in DcuS-PAS_C. These differences suggest that (isolated) DcuS-[PAS_C] is intrinsically disordered and that the inactive state of DcuS *in vivo* requires the presence of additional protein factors (as found in ref. Kleefeld *et al.* submitted for publication) for stabilization. Under such conditions, the initiating signal from the sensor could be a rotation- or piston-like motion of TM_2 as suggested for the CitA system [236]. Shortening of the linker connecting TM_2 and PAS_C domain could disrupt dimer contacts and conformation, and consequently have a significant impact on the conformation and activity of the subsequent kinase domains (as illustrated in figure 8.6e).

In summary, this work suggests that structural disorder in the cytoplasmic DcuS-PAS_C domain plays an important role in signal transduction to the kinase domain and regulation of kinase activity. The combination of ssNMR with structural data obtained on individual domains using *in silico* prediction routines or solution-state NMR represents a general approach to study partially disordered (membrane) proteins. While such

systems are difficult to access by other biophysical methods, this presented strategy may provide a powerful tool to relate protein structure to function in a membrane setting.

8.4 Methods

8.4.1 Sample preparation

DcuS-[PAS$_C$]. The cDNA encoding fragment 211-325 of *E. coli* DcuS-[PAS$_C$] was cloned into a modified pET16b vector containing a TEV cleavage site. The construct transformed into the *E. coli* strain Bl21(DE3) and expressed at 30 °C. Expression of the uniformly labeled sample (U-PAS$_C$) was performed in M9 minimal medium with $^{15}NH_4Cl$ and $^{13}C_6$-D-glucose as the sole nitrogen and carbon sources, respectively. For the uniformly labeled sample containing several reverse-labeled residue types, the minimal medium was supplemented with 1 mM unlabeled Phe, Ile, Lys, Leu and Arg and 4 mM Val. After cell lysis by three passages through a French Press at 1200 psi the protein was purified by Ni-NTA chromatography. The N-terminal His-tag was cleaved off with TEV protease, leaving behind the N-terminal additional amino acids Gly, His and Met due to cloning. The protein was further purified by gel filtration on a Superdex 75 HiLoad 16/60 column. The purified protein was dialyzed against 20 mM HEPES, pH 7.0 and 50 mM NaCl. The protein was concentrated to 13 mg/ml with a 10 kDa MWCO Vivascience ultrafiltration concentrator. U-PAS$_C$ was precipitated from this solution by adding an equal volume of 25% PEG 2000 dissolved in the same buffer, followed by concentrating the sample in to about half its initial volume in a centrifugal evaporator [246]. R-PAS$_C$ was precipitated by gradually adjusting the pH of the protein solution with hydrochloric acid to 4.6, the theoretical isoelectric point of the 211-325 DcuS fragment.

DcuS - Genetic Methods. Standard genetic methods were carried out according to ref.[247]. Plasmid and strains are given in Supplementary Table 1. For over-expression of dcuS for ssNMR, a truncated version of the gene was amplified using primers dcuStmr and dcuSPAS3 (supporting figure on page 140) from *E. coli* AN387 genomic DNA, introducing a stop codon (TAA) at aa position 333 of native DcuS. The PCR product was cloned into pET28a (Novagen) behind the T7/lac promotor into *Nde*I and *Xho*I sites. The resulting plasmid pMW309 codes for DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$] and an N-terminal His$_6$-tag. DcuB could, thus far, not be overproduced or isolated, and the ssNMR study was carried out in the absence of DcuB. For site directed mutagenesis the QuikChange kit (Stratagene) and plasmid pMW181 containing the intact *dcuS* gene served as the template for site-directed mutagenesis by PCR using primers of Table shown in supporting figure on page 141. Plasmids with verified mutations were transformed into *E. coli*

IMW260 with an insertionally inactivated chromosomal dcuS gene and a *dcuB´-´lacZ* reporter gene fusion. For assays on the function of the dcuS mutations, the bacteria were grown under anoxic conditions in enriched M9 mineral medium with glycerol (50 mM), dimethyl sulfoxide (DMSO; 20 mM) as an electron acceptor and fumarate (20 mM) as effector. Exponentially growing bacteria (OD_{578} of 0.5 to 0.8) were assayed for β-galactosidase activity [248] in four replicates each from four independent growth experiments [237, 238].

Work with DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$]. Overexpression of DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$] was performed in *E. coli* C43(DE3) [249] containing plasmid pMW309. For preparation of uniformly labelled protein *E. coli* C43(DE3) pMW309 was grown in M9 medium47 without NH$_4$Cl, but with 7 mM [^{13}C]glucose and 1 g/l [^{13}C]NH$_4$Cl, at 30°C under aerobic conditions. For reverse labelling, amino acids phenylalanine, isoleucine, leucine, arginine, lysine (1 mM each) and valine (4 mM) were included in the medium. At an OD_{578} of 0.6 - 0.8 expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 4 to 5 h. Cells were harvested and membranes were prepared [216]. The membranes were homogenised in buffer 1 (50 mM TrisHCl pH 7.7, 0.5 M NaCl, 10 % glycerol, 10 mM imidazole). Membrane proteins were solubilised by a modification of a published method [216] using detergent Empigen BB at a final concentration of 2 % (w/v) by gently stirring the mixture on ice for 30 min. After centrifugation (300,000g, 50 min), the supernatant was put on a Ni^{2+}-NTA column (2 ml, Qiagen) equilibrated in buffer 1 and washed with 20 vol buffer 2 (buffer 1 with 20 mM imidazole and 0.04 % LDAO). Bound protein was eluted with 10 x 1 ml buffer 3 (buffer 2 with 500 mM imidazole). Fractions with His$_6$-DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$] of sufficient purity (> 90 %) were pooled. For reconstitution of DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$] liposomes were prepared from *E. coli* phospholipids (polar lipid extract, 20 mg/ml in chloroform, Avanti Polar Lipids) [216]. Liposomes were destabilised by addition of Triton X-100 at an effective detergent:lipid ratio of 2.5 (ref.[250]). Purified DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$] (5.5 mg) was added at a phospholipid:protein ratio of 10:1 (mg/mg) and mixed by gentle agitation for 10 to 15 min at 20°C. For every mg of Triton X-100, 3 x 5 mg degassed Bio-Beads were added to remove the detergent. The mixture was incubated for 2 hours at room temperature under gentle stirring. Then again 5 mg Bio-Beads / mg detergent were added and incubated overnight at 4°C. Fresh Bio-Beads were added and incubated for 1 to 2 h at 20°C. The supernatant was removed and centrifugated (300.000xg, 50 min). Proteoliposomes were washed twice (300,000xg, 10 min) and resuspended in 50 mM TrisHCl pH 7.7 to a volume of 400 μl (5.5 mg protein), frozen in liquid N$_2$ and thawed at room temperature for three cycles before freezing and storage at -80°C.

8.4.2 Solid-state NMR

All solid-state NMR experiments were conducted using 4 mm triple-resonance (^{1}H,^{13}C,^{15}N) probeheads at static magnetic fields of 18.8 T and 14.1 T corresponding to 800 MHz and 600 MHz ^{1}H resonance frequencies (Bruker Biospin, Karlsruhe/Germany). (^{13}C,^{13}C) SD experiments were measured in non frozen sample conditions at effective temperatures between 1-10°C. (^{13}C,^{13}C) mixing times of t_{CC}=15ms were used to detect intraresidue spin systems. Sequential resonance assignments were obtained using (^{13}C,^{13}C) correlation experiments performed under weak coupling conditions [23] using t_{CC}=150ms. Total experiment time for one spectrum was in the order of one (DcuS-[PAS$_C$]) to three (DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$]) days. CHHC experiments were carried out using short CP contact times (100 μs for R-PAS$_C$ and 120 μs for U-PAS$_C$) bracketing the proton-proton mixing time (t_{HH}=400 μs and t_{HH}=100 μs respectively). R-PAS$_C$ was measured in the frozen state to increase signal to noise, U-PAS$_C$ was measured at 10°C. Total experiment time for each spectrum was about 7 days at 800 MHz. SPINAL64 proton decoupling [25] was applied using radio-frequency fields of 75-90 kHz and MAS rates between 8 and 12.5 kHz.

8.4.3 Comparative modeling and structure calculations

Primary sequence as present in the DcuS-[PAS$_C$] domain samples was submitted to the RO-BETTA server. No further manual intervention was done. Only one fragment was used and aligned to pdb: 2gj3 (ref.[239]). Structure calculations were performed with the X-PLOR NIH software package [251] using torsion angle dynamics (TAD) protocol [252]. The force field constant for the TAD calculations were taken from parallhdg5.3.pro [253]. Standard values of ϕ and ψ backbone dihedral angles for residues predicted by PSIPRED as α-helix and β-strand were included into structural calculation with a tolerance of 25°. For the fully labelled sample, distance restraints were incorporated with an upper limit of 3.0 Å and for the reverse labelled sample an upper bound of 4.5 Å was used. The calculation started with 3,000 cycles of TAD at 50,000 K, followed by 3,000 cycles of TAD with increasing values of interatomic repulsion while cooling to 1,000 K, and subsequent 7,000 cycles of molecular dynamics in Cartesian space while cooling to 300 K and then 1,000 cycles of final Powell energy minimization. The force constant were set to 300, 300, 300 and 150 kcal/mol for the distance restraints, and 100, 200, 250 and 300 kcal/mol for dihedral restraints during the four stages of structure calculations. A soft square potential was used for distance restraints and a square well potential was used for dihedral restraints.

9

Summary, conclusion and outlook

Protein precipitates. One focus of the presented work was the investigation of protein precipitates. Using precipitation at the isoelectric point it could be demonstrated that native protein structure for the model protein Crh was largely retained. Based on these results a similar analysis enabled the structural characterization of DcuS-[PAS$_C$], a partially disordered protein domain. It can be speculated that the intrinsic plasticity identified by ssNMR for specific protein fragments is responsible for an unstable behavior in solution and bad crystallization properties (as observed before). The intrinsic disorder could be related to protein function which suggests that artificially increased protein stability, enforcing crystallization or solubility, may lack important aspects of the system. In this respect future structural studies on unstable proteins may in general benefit from an ssNMR based investigation of their precipitated state.

Protein aggregation. Temperature induced aggregation of precipitated Crh could be monitored in real time using 2D ssNMR spectroscopy. In addition to the initial, native conformation and the β-strand enriched aggregated state, the spectrum also revealed the presence of a partially unfolded state. A detailed kinetic analysis of the aggregation process showed that the experimental data could not be reproduced adequately assuming a single exponential folding transition and suggested the appearance of a barrier free unfolding step. This project revealed that ssNMR provides a side resolved spectroscopic means, which is complimentary to most established methods, to follow protein

aggregation.

ssNMR & comparative modeling. Previous studies have shown that computational protein modeling can improve the quality and speed of solution-state NMR [64, 221, 222] and X-ray crystallographic [65] studies of small to medium size proteins. Here the combination of comparative modeling and ssNMR was introduced, which resulted in the first experimental validated structure of a precipitated protein. Intrinsic disorder, limited sensitivity and resolution, combined with the relative large size of the protein reduced the accuracy of the resulting structure. However this approach can, analogous to the fitting of structural models to experimental data obtained using electron cryo-microscopy (cryo-EM) [220, 254], become a general routine in the structure elucidation process. The quality of the resulting structures strongly depends on the underlying experimental and computational data of the system and can be improved by using e.g. different labeling schemes. Further work should be done to automatically implement and optimize ssNMR data according to structural modeling (or vice versa) and to determine the actual resolution offered by this approach. Due to the prerequisites offered by the ssNMR data the resolution should in principle be between the typical resolution of cryo-EM data ($\backsim$3-10 Å) and the high resolution obtainable with conventional NMR spectroscopic techniques (<1 Å). Notably and promisingly in terms of future applications, the presented structure of DcuS-[PAS$_C$] is the first (and largest) structure of a globular protein solved with ssNMR, which had no reference structure determined by another experimental technique.

Membrane proteins I. Regarding the initial objective of the characterization of membrane related proteins, several aspects could be investigated. For the sensory rhodopsin system a *de novo* structural investigation was carried out. Secondary structure, molecular dynamics and membrane topology could be probed. Due to the large number of possible cross contacts, a three dimensional *de novo* structure determination was not attempted. In principle computational modeling could also simplify structural interpretation of ssNMR data obtained on membrane proteins; however in this context comparative modeling is less reliable because of the limited number of reference structures in the protein data bank [3]. Initial attempts to access these challenge have already been carried out [255]. In particular the presence of several structures related to the rhodopsin

family renders this system an excellent test system for further studies concerning the structure elucidation of membrane proteins based on ssNMR, probably in combination with computational approaches.

Membrane proteins II. In terms of complex formation of sensory rhodopsin and its transducer molecule, novel insights into the binding interface could be obtained. This study benefits on the one side from the existence of crystallographic information and on the other side from the possibility to study a uniformly labeled system in a lipid bilayer setting. It could be shown that the crystallographic reference lacks important features in regions that experience close intermolecular contacts induced be the crystal lattice. Additional ssNMR data obtained on the activated state suggested that also functional aspects associated with receptor activation are affected by the crystal lattice. The sample condition used for acquiring ssNMR data on the activated state led to decreased resolution. In principle an optimized procedure (possibly lacking the graining step) could improve spectroscopic results of the activation process.

Membrane proteins III. Crystallization of multidomain membrane embedded proteins is particularly difficult. The results presented for the DcuS protein exemplified that ssNMR can offer structural and dynamical insights into such systems. The integrity of structural information obtained on the isolated periplasmic domain [227] and a cytoplasmic domain was evaluated in reference to data recorded on a multidomain fragment containing both domains as well as their helical membrane embedded connections. While most segments are conserved in the membrane embedded construct, structural alterations were identified in the direct inter-domain connecting regions. In addition the ssNMR data imply that the structural disorder found for the isolated cytoplasmic domain was conserved upon membrane connection. Supported by other (independent) experimental results this study suggests a mechanism for intra cellular signaling and regulation of kinase activity. Since homologous domains occur in about $\frac{1}{3}$ of all membrane embedded sensory kinases and the identified key residues are most conserved among them, the proposed model might potentially be a general mechanism of signal transduction.

10

List of publications

Initial work, side projects and the chapters directly presented in this thesis are related to the following publications:

1. Ader C, Schneider R, Seidel K, Etzkorn M and Baldus, M: Magic-angle-spinning NMR spectroscopy applied to small molecules and peptides in lipid bilayers. *Biochem Soc Trans* 2007, 35, 991-995.

2. Ader C, Schneider R, Seidel K, Etzkorn M, Becker S and Baldus M: Structural rearrangements of membrane proteins probed by water-edited solid-state NMR spectroscopy *manuscript submitted*

3. Bardiaux B, Favier A, Etzkorn M, Baldus M, Böckmann A, Nilges M and Malliavin TE: Simultaneous use of liquid and solid-state NMR to study the conformational landscape of the Crh protein during oligomerisation and crystallization. *manuscript in preparation*

4. Etzkorn M, Böckmann A and Baldus M: Kinetic analysis of protein aggregation monitored by real-time 2D solid-state NMR. *manuscript in preparation*

5. Etzkorn M, Böckmann A, Penin F, Riedel D and Baldus M: Characterization of folding intermediates of a domain-swapped protein by solid-state NMR spectroscopy. *J Amer Chem Soc* 2007, 129, 169-175.

6. Etzkorn M, Böckmann A, Lange A and Baldus M: Probing molecular interfaces

using 2D magic-angle-spinning NMR on protein mixtures with different uniform labeling. *J Amer Chem Soc* 2004, 126, 14746-14751.

7. Etzkorn M, Kneuper H, Dünnwald P, Vijaian V, Griesinger C, Becker S, Unden G and Baldus M: Plasticity of the PAS domain and a potential role for signal transduction in the histidine kinase DcuS. *Nat Struct Mol Biol* 2008, *in press*

8. Etzkorn M, Martell S, Andronesi OC, Seidel K, Engelhard M and Baldus M: Secondary structure and topology of a seven-helix receptor in native membranes studied by solid-state NMR. *Angew Chem Int Ed* 2007, 46, 459-462.

9. Etzkorn M, Seidel K, Lee L, Martell S, Engelhard M and Baldus M: Complex formation and activation of the Sensory Rhodopsin - Transducer system monitored by solid-state NMR. *manuscript in preparation*

10. Gardiennet C, Loquet A, Etzkorn M, Heise H, Baldus M and Böckmann A: Structural constraints for the Crh protein from solid-state NMR. *J Biomol NMR* 2008, 40, 239-250.

11. Heise H, Seidel K, Etzkorn M, Becker S and Baldus M: 3D MAS Solid-state NMR spectroscopy for resonance assignment and structure elucidation of proteins. Novel pulse schemes and sensitivity considerations. *J Magn Reson* 2005, 173, 64-74.

12. Seidel, K, Etzkorn M, Heise H, Becker S and Baldus M: High-resolution solid-state NMR studies on uniformly [^{13}C, ^{15}N] labeled Ubiquitin. *ChemBioChem* 2005, 6, 1638-1647.

13. Seidel K, Etzkorn M, Sonnenberg L, Griesinger C, Sebald A and Baldus M: Studying 3D structure and dynamics by high-resolution solid-state NMR: Application to L-Tyrosine-Ethylester. *J Phys Chem A* 2005, 109, 2436-2442.

14. Schneider R, Seidel K, Etzkorn M and Baldus M: Molecular motion detected by double-quantum (^{13}C,^{13}C) solid-state NMR spectroscopy. *manuscript in preparation*

A Supporting Material

A.1 Supplemental Material for chapter 2

The spin operator

In analogy to the angular momentum operator, the spin operator $\hat{\boldsymbol{I}}$ can be defined by the commutation relationship of its components according to:

$$\left[\hat{I}_i, \hat{I}_j\right] = i\hbar\epsilon_{ijk}\hat{I}_k$$

Where ϵ_{ijk} is the Levi-Civita-Tensor. The two quantum numbers l and m with:

$$\hat{\boldsymbol{I}}^2 \left|lm\right\rangle = \hbar^2 l(l+1)\left|lm\right\rangle$$
$$\hat{I}_z \left|lm\right\rangle = \hbar\left|lm\right\rangle$$

desripe the respective spin states.

Basis set for spin-1/2 nuclei

The following *Pauli matrices* together with the unitimatrix form a complet basis set for spin-1/2 nuclei:

$$\hat{I}_x = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \ \hat{I}_y = \begin{pmatrix} 0 & -\frac{i}{2} \\ \frac{i}{2} & 0 \end{pmatrix}, \ \hat{I}_z = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}.$$

A.2 Supplemental Material for chapter 3
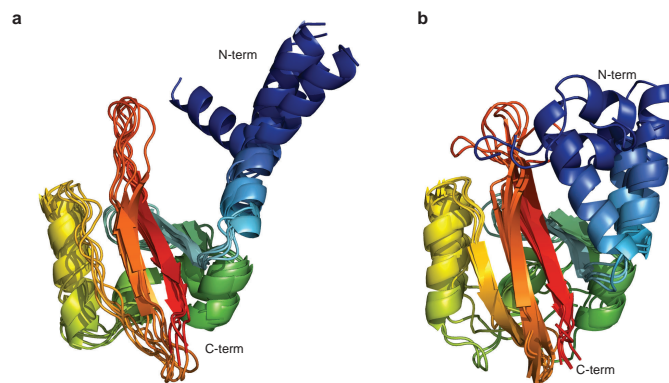
Figure A.1: Comparison between structural models obtained from I-Tasser (a) and Robetta (b) using the same primary sequence (DcuS-[PAS$_C$]).

A.3 Supplemental Material for chapter 4

Figure A.2: (a) Extract of the carbonyl region of the proton driven spin diffusion spectrum shown in figure 4.1. (b)[15]N,[13]C NCA correlation spectra of microcrystalline (black) and precipitated (red) Crh. The spectra were recorded at 800 MHz ^{1}H resonance frequency, 269 K sample temperature and an MAS rate of 10.5 kHz. Assignments are shown according to [37].
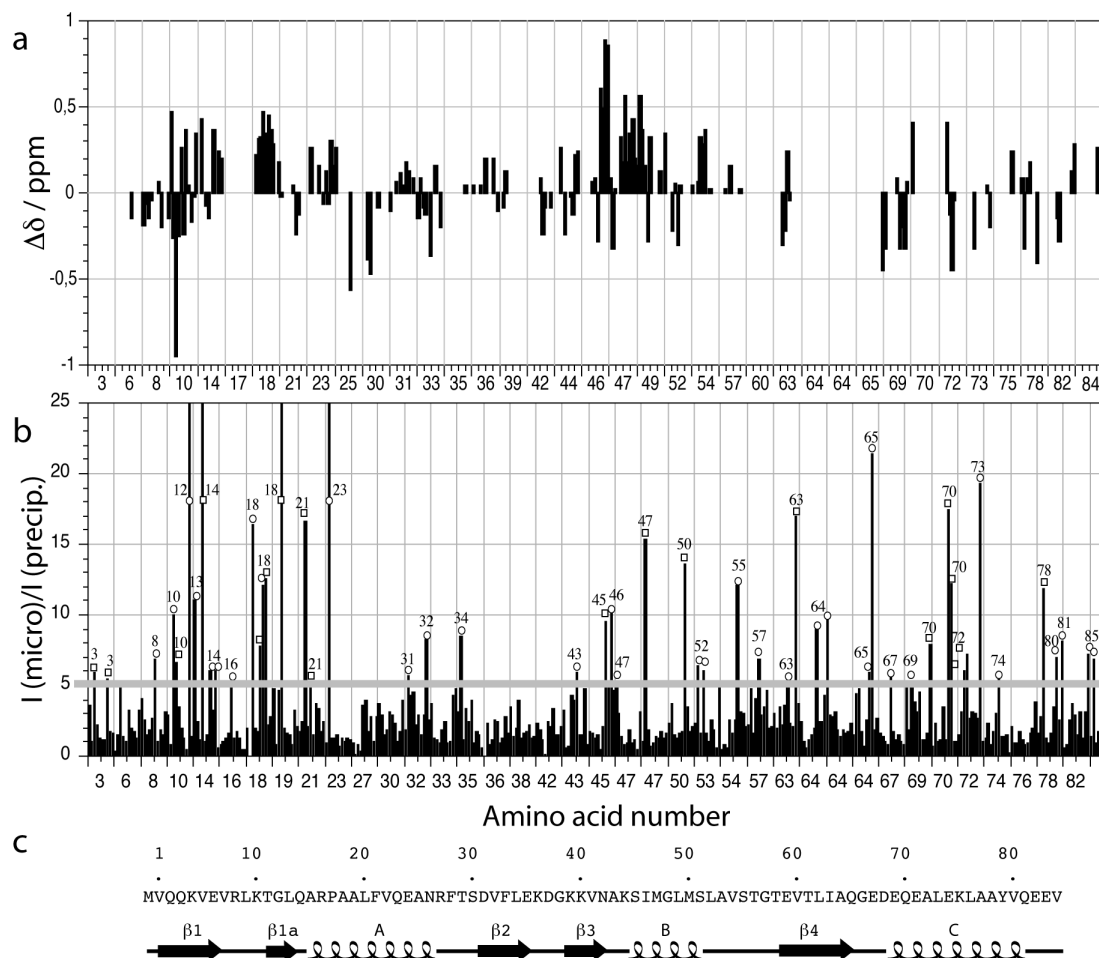
Figure A.3: (a) Carbon chemical-shift differences identified in the spectra of microcrystalline and precipitated Crh $\Delta\delta = 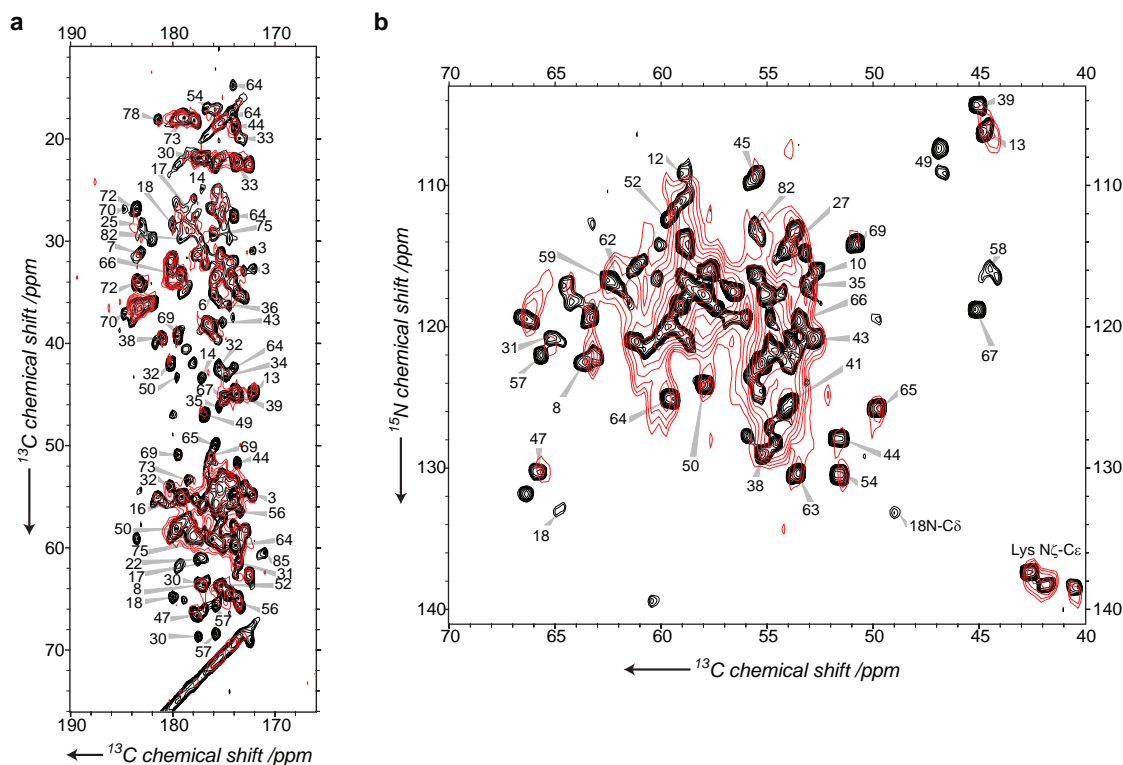(\delta_{micro} - \delta_{precip})$. (b) Ratio of peak intensities (heights) measured on microcrystalline and pI-precipitated Crh $\left(\frac{I_m}{I_p}\right)$ using the 2D spectra shown in figure 4.1. Circles indicate signals involving $C\alpha$-C' or $C\alpha$-Cβ carbons, and squares relate to side-chain resonances. Peak intensities were evaluated using the SPARKY software. Intensities were measured for resolved cross signals. (c) Primary sequence and secondary structure of Crh in its dimeric form.
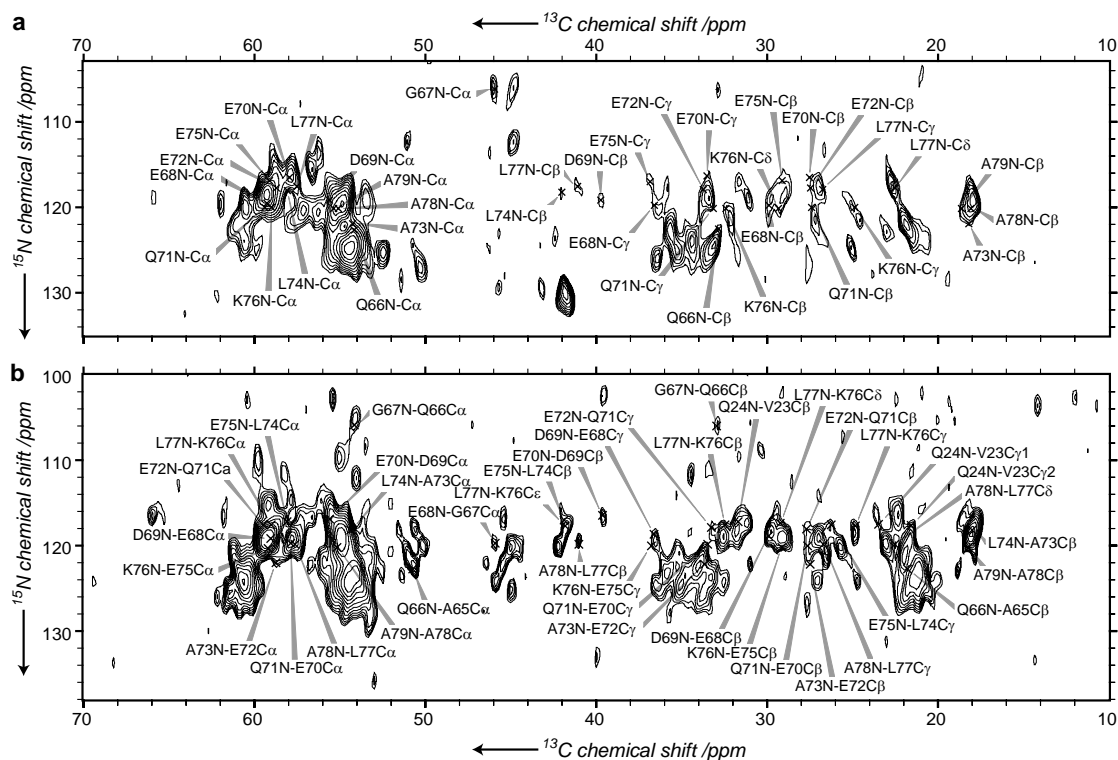
Figure A.4: NCACX (a) and NCOCX (b) spectra of the aggregated form of Crh. The indicated labels correspond to the sequential walk shown in figure 4.5e. Spectra were recorded at 800 MHz ^{1}H resonance frequency, 269 K sample temperature and an MAS rate of 10.5 kHz.
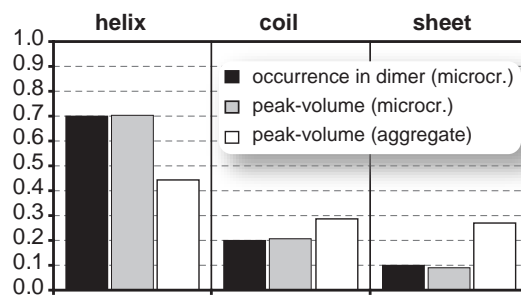
Figure A.5: Occurrence of Alanines in the different secondary structure elements, and relative signal volume measured in the spectra of microcrystalline and aggregated Crh.

A.4 Supplemental Material for chapter 5
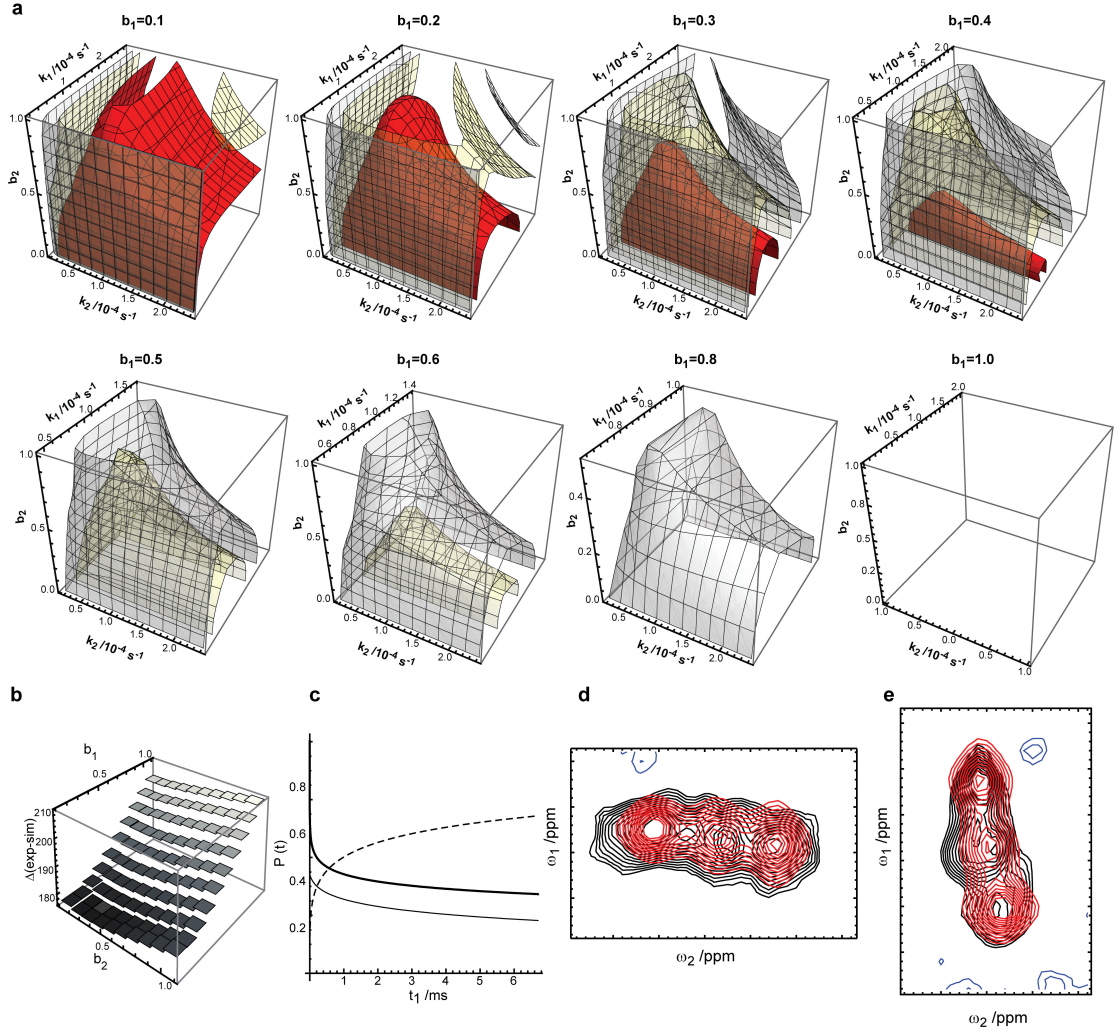
Figure A.6: Complete downhill scenario for the aggregation of Crh. The four free parameters b_1, b_2, k_1 and k_2 were used according to equations (A.1)-(A.3) to describe the kinetic profile. With the exception of the modified parameters the same aspects as shown in figure 5.9 are indicated. Note that in (b) a two dimensional dependence is plotted to account for b_2.

$$P_A(t) = e^{-(k_1 t)^{b_1}} \tag{A.1}$$

$$P_B(t) = \left(1 - e^{-(k_1 t)^b}\right) \cdot e^{-(k_2 t)^{b_2}} \tag{A.2}$$

$$P_C(t) = \left(1 - e^{-(k_1 t)^b}\right) \cdot \left(1 - e^{-(k_2 t)^{b_2}}\right) \tag{A.3}$$

```
α=185* π;
α01= α02= α11= α12= α21= α22= α;
ω01=4422;  ω02=13917;
ω11=4387;  ω12=14082;
ω21=4349;  ω22=14350;
ω1=.
ω2=.
a0=1;a1=1;a2=1;
k1=1;
k2=1500;
LB=50;
pA[t_]:=Exp[-k1*t]
pB[t_]:=(1-Exp[-k1*t])*Exp[-k2*t]
pC[t_]:=(1-Exp[-k1*t])*(1-Exp[-k2*t])
L[ ω_, ω0_, α0_]:=  α0/( α0^2 +4* π^2*(ω- ω0)^2)
ThrCgCbem50=Get["mathThrCgCbem50.txt"];
exspec=ThrCgCbem50;
wdS[ ω1_, ω2_]:=NIntegrate[(a0*pA[t1]*Exp[(2*π*I*ω01- α01-
1/2*  π*LB)*t1]*L[ ω2, ω02, α02] +a1*pB[t1]*Exp[(2* π*I*ω11- α11-
1/2*  π*LB)*t1]*L[ ω2, ω12, α12]+a2*pC[t1]*Exp[(2*π*I*ω21- α21-
1/2*  π*LB)*t1]*L[ ω2, ω22, α22])*Exp[-
2* π*I*  ω1*t1],{t1,0,          ∞},Method   →
{Automatic,"SymbolicProcessing" →0}]
k1min=0;k1max=4400;k1in=80;
k2min=0;k2max=4160;k2in=80;
lengthM1=(k1max-k1min-Mod[(k1max-k1min),k1in])/k1in ;
lengthM2=(k2max-k2min-Mod[(k2max-k2min),k2in])/k2in ;
M=Table[0,{i,lengthM1+1},{j,lengthM2+1}];
parameter={"Parameters used for difference matrix units in
Hz
/s:","alpha",  α/ π,"k1min",k1min,"k1max",k1max,"k1in",k1in,"k2
min",k2min,"k2max",k2max,"k2in",k2in,"em window funct in
t1",LB}
x=1;y=1;
For[k1=k1min,k1<=k1max,k1=k1+k1in,For[k2=k2min,k2<=k2max,k2=
k2+k2in,
 theospec=Table[Re[wdS[ ω1, ω2]],{    ω1,4768,4071,-
33.3},{ω2,14659,13613,-33.8}];
   normdif=exspec/exspec[[11,23]]-theospec/theospec[[11,23]];
M[[y,x]]=Sum[Sqrt[normdif[[i,j]]^2],{i,1,Length[exspec]},{j,
1,Length[exspec[[1]]]}];x=x+1];y=y+1;x=1;Print[k1 "k1"]]
Put[parameter,M,"MThrCgCbem50.txt"]
```

Figure A.7: Mathematica code used for simulating the experimental peak shape. The code is examplified for the simulation of the downhill unfolding and nucleated aggregation scenario. For other scenarios the population profiles have been modified as given in the text.

A.5 Supplemental Material for chapter 6

Residue		N /ppm	Cα /ppm	Cβ /ppm
11	Gly	-	47.7 *	
12	Ala	-	55.4 *	-
13	Ile	119.4 **	66.1 **	36.8 ***
14	Gly	108.3 **	47.9 **	
15	Met	118.6 *	59.0 *	30.4 *
18	Gly	-	47.9 **	
19	Thr	106.0 ***	66.3 ***	68.6 **
23	Ala	-	55.4 *	19.0 *
24	Trp	-	60.2 **	30.0 **
25	Ala	120.1 **	54.3 ***	19.2 **
26	Gly	103.0 **	44.0 **	
27	Arg	118.7 **	58.4 **	29.0 **
28	Asp	122.1 *	53.4 ***	+
29	Ala	125.1 ***	50.6 ***	18.1 ***
30	Gly	115.1 ***	45.2 **	
31	Ser	-	58.0 ***	63.5 ***
32	Gly	116.2 **	45.8 **	
33	Glu	120.3 *	56.2 *	+
34	Arg	125.8 *	57.4 *	31.5 *
42	Gly	-	48.4 *	
43	Ile	115.7 *	65.6 *	37.7 *
44	Ser	114.7 *	61.1 *	62.8 *
45	Gly	114.3 *	47.3 *	
46	Ile	119.4 *	63.8 *	36.3 *
47	Ala	-	55.3 *	18.2 *
62	Pro		62.8 *	+
63	Val	125.6 *		
64	Ala	-	51.9 **	19.8 **
65	Glu	125.7 **	57.2 **	+
66	Arg	121.6 *	56.6 **	+
67	Thr	115.8 **	61.7 **	+
71	Pro	119.2 **	65.6 **	30.4 **
72	Arg	-	60.5 *	30.1 *
74	Ile	-	66.3 **	38.4 **
75	Asp	120.0 **	56.5 **	40.0 **
76	Trp	113.7 *	60.5 *	32.9 *
91	Ala	-	54.3 **	17.5 **
92	Gly	111.8 **	46.1 **	
94	Asp	119.3 **	52.1 ***	42.0 ***
95	Ser	112.8 ***	63.1 ***	-
99	Gly	-	48.0 *	
100	Ile	-	65.2 *	38.4 *
102	Ile	-	65.2 *	38.4 *
103	Thr	113.5 *	66.8 **	68.0 *
105	Asn	114.0 **	58.9 ***	42.0 ***
106	Thr	106.5 ***	67.2 ***	-
111	Ala	-	55.8 *	17.3 *
112	Gly	100.9 *	47.7 *	
114	Ala	-	54.8 *	18.6 *
115	Gly	99.4 *	48.2 *	
121	Ile	115.8 ***	61.2 ***	38.8 **
122	Glu	119.7 ***	56.7 *	+
128	Gly	-	48.0 *	
129	Met	115.4 **	58.7 **	32.6 **
130	Gly	104.0 **	47.9 **	
131	Ala	-	56.0 *	-
143	Gly	-	44.9 *	
144	Pro	140.5 *	65.4 **	31.9 **
145	Met	111.6 **	56.0 **	31.9 **
146	Thr	114.9 *	68.3 **	-
147	Glu	117.9 **	59.4 ***	28.7 **
148	Ser	116.9 *	62.9 **	-
149	Ala	-	55.5 ***	18.4 ***
150	Ser	108.6 **	60.7 **	63.1 *
151	Gln	-	56.3 *	30.3 *
152	Arg	-	55.6 **	29.9 **
153	Ser	112.3 **	58.3 ***	64.4 ***
154	Ser	116.1 **	62.1 ***	-
155	Gly	109.2 **	48.2 **	
156	Ile	-	65.7 **	38.0 **
157	Lys	119.2 **	60.7 ***	31.8 ***
158	Ser	111.2 ***	61.6 **	62.9 *
171	Trp	119.3 *	58.9 **	28.1 **
172	Ala	117.3 *	56.6 **	18.1 *
173	Ile	-	65.6 *	38.2 *
181	Gly	108.6 *	44.3 **	
182	Pro	120.5 *	65.1 *	-
183	Pro	119.6 *	65.1 **	32.8 **
184	Gly	99.9 **	44.8	
189	Thr	111.7 **	62.2 ***	69.6 ***
190	Pro	120.0 ***	66.1 ***	31.0 **
191	Thr	116.9 *	62.4 **	68.6 *
223	His	-	55.9 **	+
224	Gly	115.0 ***	45.4 **	
225	Glu	121.4 *	56.9 *	+
226	Ser	116.5 **	58.4 **	+
228	Ala	-	52.9 **	+
229	Gly	114.6 **	45.3 ***	
231	Asp	-	54.3 ***	41.2 **
232	Thr	114.2 ***	62.3 ***	69.4 **
233	Asp	122.3 ***	54.3 **	+
234	Ala	128.7 *	+	+
235	Pro	-	62.9 ***	32.9 ***
236	Ala	124.6 **	52.2 **	+
238	Ala	127.0 **	52.1 **	+
239	Asp	120.1 **	54.5 **	+
240	Glu	122.3 *	56.0 **	+
241	Asn	119.9 ***	53.5 ***	39.1 ***
242	Ser	116.3 **	58.0 **	63.5 **
243	His	121.6 **	+	+

Figure A.8: Spectral assignment of U[^{13}C,^{15}N \ (V,L,F,Y)] NpSRII in purple membrane lipids under MAS conditions.

Solid-state NMR Resonance assignments. ^{13}C resonances were calibrated using adamantane as an external reference. The upfield resonance of adamantane was set to 31.47 ppm to allow for a direct comparison of the solid-state chemical shifts to solution-state NMR data. Accordingly, ^{15}N resonances were calibrated using the tripeptide AGG as an external reference [28]. All resonance assignments are self-consistent and reflect at least one isolated sequential correlation that is characteristic for each spin system under investigation. The reliability largely depends on the length of each amino acid segment and the spectral overlap and ranges from very good (***), good (**) to reasonable (*). Residues identified with the help of through-bond experiments are given in red; resonances denoted by a '+' symbol in figure A.8 were not assigned sequentially but appear at the amino-acid specific chemical shift in the H-C INEPT spectrum. Note that ^{15}N resonance assignments are also given for Val 63. This observation is due to transaminase reactions that lead to ^{15}N labeling of Val residues even if natural abundance Val is added to the minimal medium (see, e.g. ref.[56] and references therein).
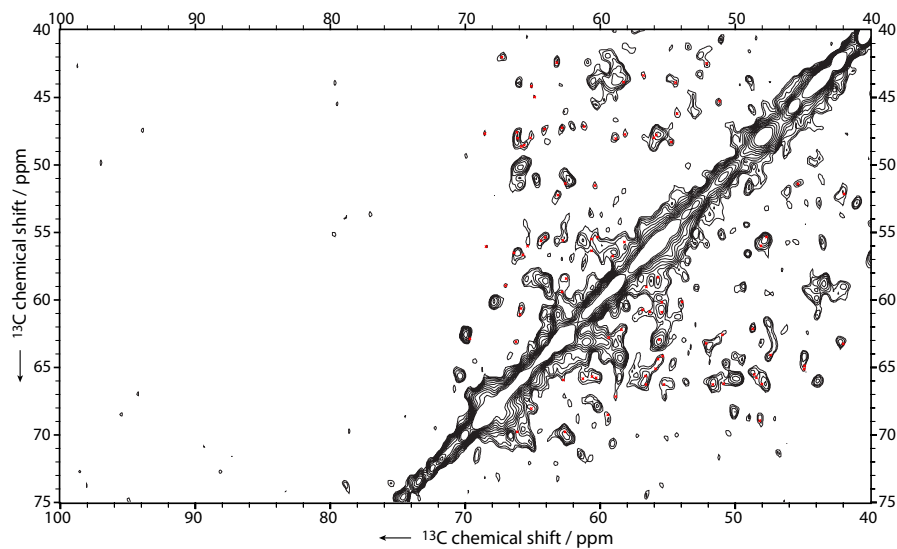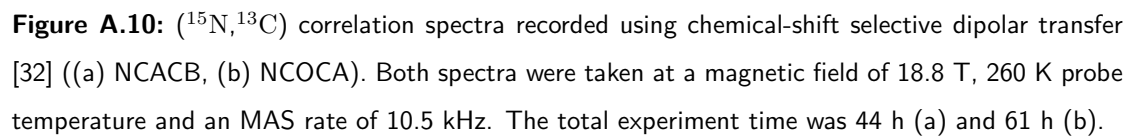
Figure A.9: ((^{13}C,^{13}C) spin-diffusion spectrum recorded under weak-coupling conditions (see figure 6.2b). On the left side of the spectrum (75-100 ppm) a region which should contain no cross signal is shown to visualize the relationship between experimental data and the spectral noise. Contour levels are identical to figure 6.2b. Red dots indicate crosspeaks used for the assignment.

Figure A.10: (^{15}N,^{13}C) correlation spectra recorded using chemical-shift selective dipolar transfer [32] ((a) NCACB, (b) NCOCA). Both spectra were taken at a magnetic field of 18.8 T, 260 K probe temperature and an MAS rate of 10.5 kHz. The total experiment time was 44 h (a) and 61 h (b).

A.6 Supplemental Material for chapter 7
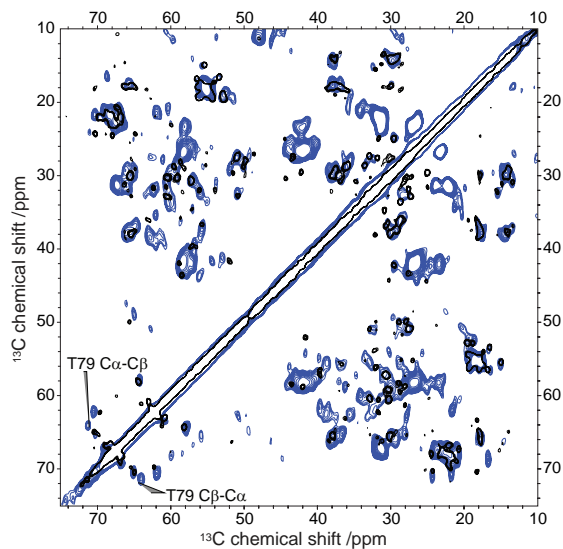
Figure A.11: (^{13}C,^{13}C) SD spectrum of WT U[^{13}C,^{15}N\(V,L,F,Y)]NpSRII (black) and (D75N U[^{13}C,^{15}N\(V,L,F,Y)]NpSRII (blue).
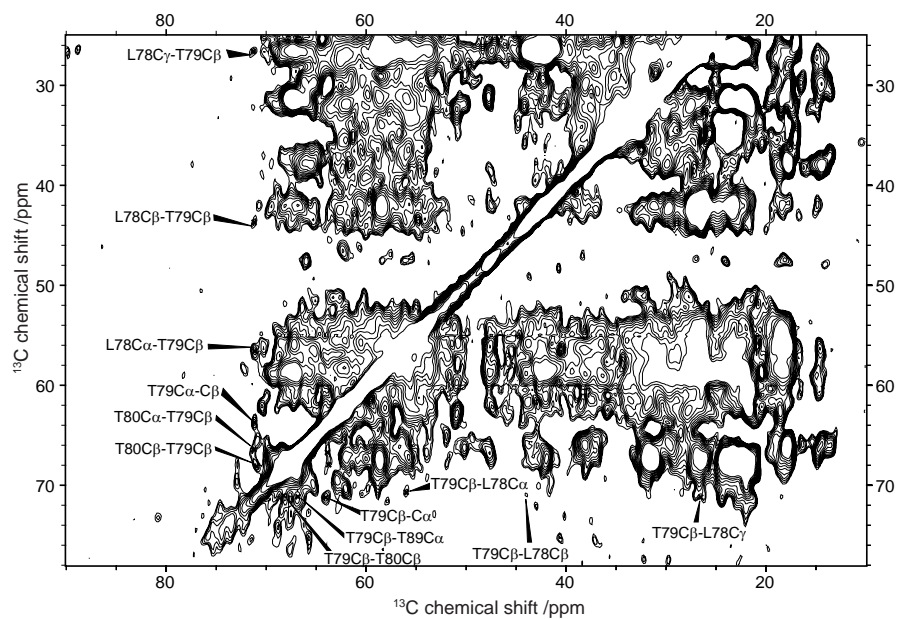
Figure A.12: (^{13}C,^{13}C) SDWC spectrum of U[^{13}C,^{15}N]NpSRII−NpHtrII. Sequential assignment of Thr79 is indicated.
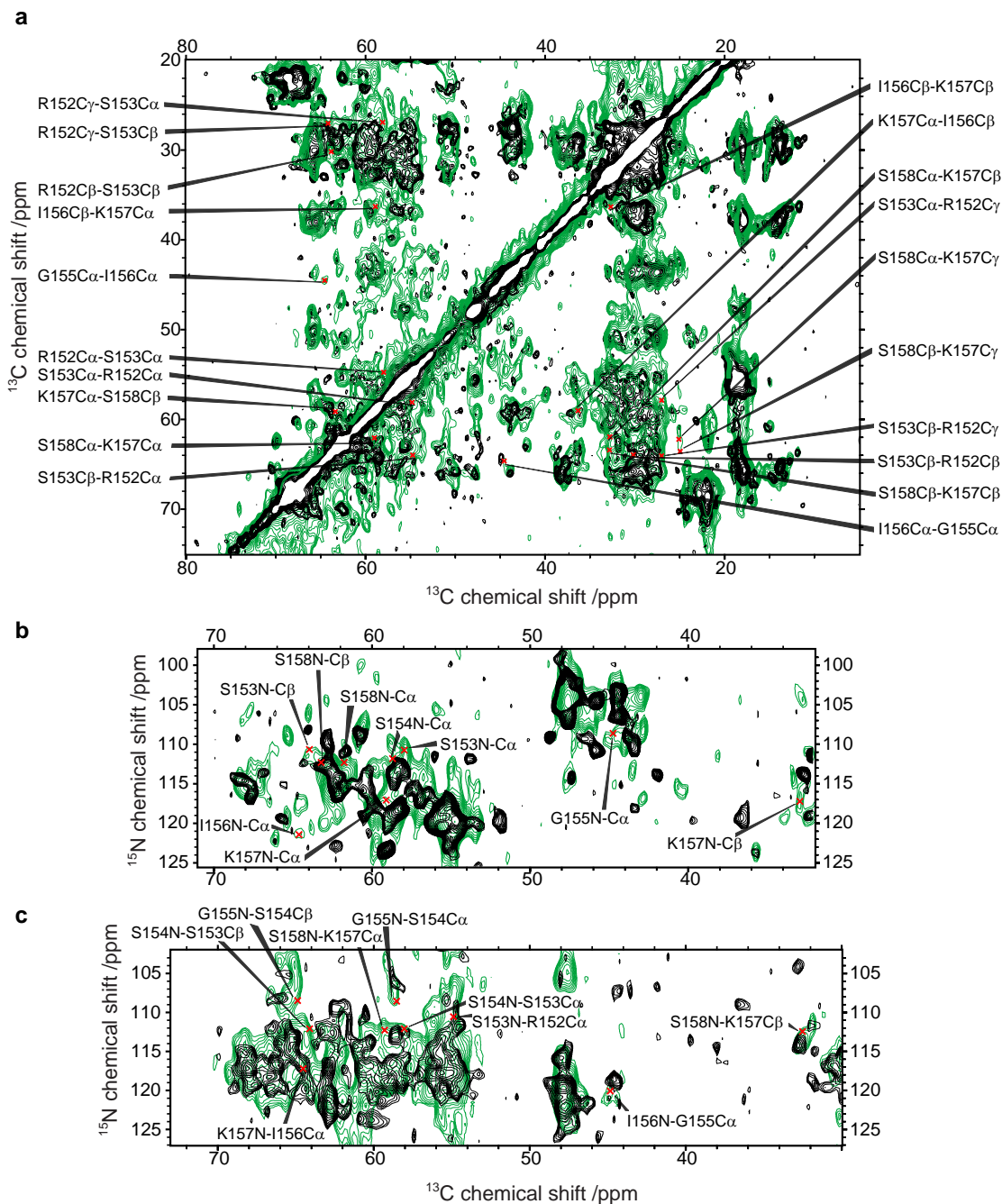
Figure A.13: Spectral extracts of $(^{13}\mathrm{C},^{13}\mathrm{C})$ SDWC (a), NCACB (b) and NCOCA (c) experiments showing sequential assignments of complex NpSRII (green). Spectra recorded under comparable experimental conditions on free NpSRII are shown in balck.
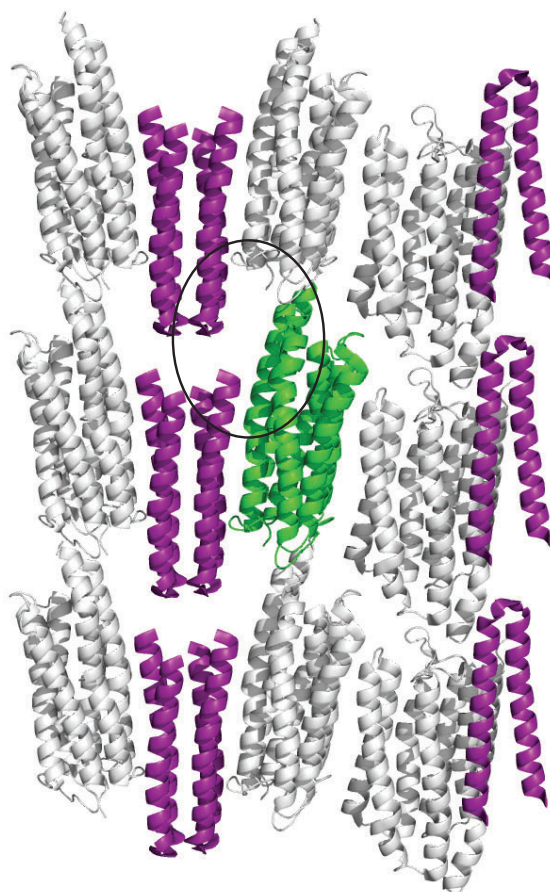
Figure A.14: Crystal packing of NpSRII/HtrII. The terminal ends of the transducer molecule (purple) are in close contact to the transducer molecule of the next layer. Note that the transducer sequence (Htr$_{114}$ used for crystallization comprises about 20 not visible residues on both termini. Region discussed in the text is highlighted.

```
(*Read_correlations_from_X-ray_structure*)
xray = << "shift-dist-xray_single_mat";
(*Read_correlations_from_extended_model*)
ssnmr = << "shift-dist-model16_8_single_reverse_mat";
(*Read_ssNMR_spectrum*)
sr2spec = << "1dspec_mat";
(*Parameters_for _calculation*)
spec = sr2spec; shiftsd = xray; lwfit = 1.`; cut = 7.00;
shiftsfiltered = {};
For[n = 2, n < Length[shiftsd] + 1, n = n + 2,
 If[shiftsd[[n]] <= cut, shiftsfiltered = Append[shiftsfiltered, shiftsd[[n - 1]]]]]
```

$$
loras[i\_, y\_] := \frac{lwfit}{lwfit^2 + (y - shiftsfiltered[[i]])^2}
$$

```
fitpeaks[y_] = Table[loras[n, y], {n, 1, Length[shiftsfiltered] - 1}];

datafit = Fit[spec, fitpeaks[x], x];

pos = {}; val = {};
For[i = 1, i < Length[spec] + 1, i++,
 pos = Append[pos, spec[[i, 1]]]; val = Append[val, spec[[i, 2]]]]
fitval = datafitx[pos];
spectplot = Plot[datafit, {x, 0, 80}, PlotRange ->All, PlotStyle ->{RGBColor[1, 0, 0]}]
diff = Tr[Abs[val - fitval]]


(*Calculate_cutt_off_dependend_difference*)
incr = 0.2; difflist = {};

For[cut = 0, cut < 10.1, cut = cut + incr, shiftsfiltered = {};
 For[n = 2, n < Length[shiftsd] + 1, n = n + 2,
  If[shiftsd[[n]] < cut, shiftsfiltered = Append[shiftsfiltered, shiftsd[[n - 1]]]]];
 loras[i_, y_] := lwfit / (( lwfit) ^2 + (y - shiftsfiltered[[i]]) ^2);
 fitpeaks[y_] = Table[loras[n, y], {n, 1, Length[shiftsfiltered] ? 1}];
 datafit = Fit[spec, fitpeaks[x], x];
 pos = {}; val = {};
 For[i = 1, i < Length[spec] + 1, i++,
  pos = Append[pos, spec[[i, 1]]]; val = Append[val, spec[[i, 2]]]];
 fitval = datafitx[pos];
 difflist = Append[difflist, diff = Tr[Abs[val ? fitval]]]]

ref = Tr[Abs[val]];
diffplot = {};
dist = Table[x, {x, 0, 10, incr}];
For[i = 1, i < Length[difflist], i++,
 diffplot = Append[diffplot, {dist[[i]], difflist[[i]] / ref}]]
ListPlot[diffplot, PlotStyle ->PointSize[0.02], PlotRange ->{{0, 10}, {0, 1.1}}]
```

Figure A.15: Mathematica code used to fit a series of Lorentzian functions to the experimental 1D spectrum.
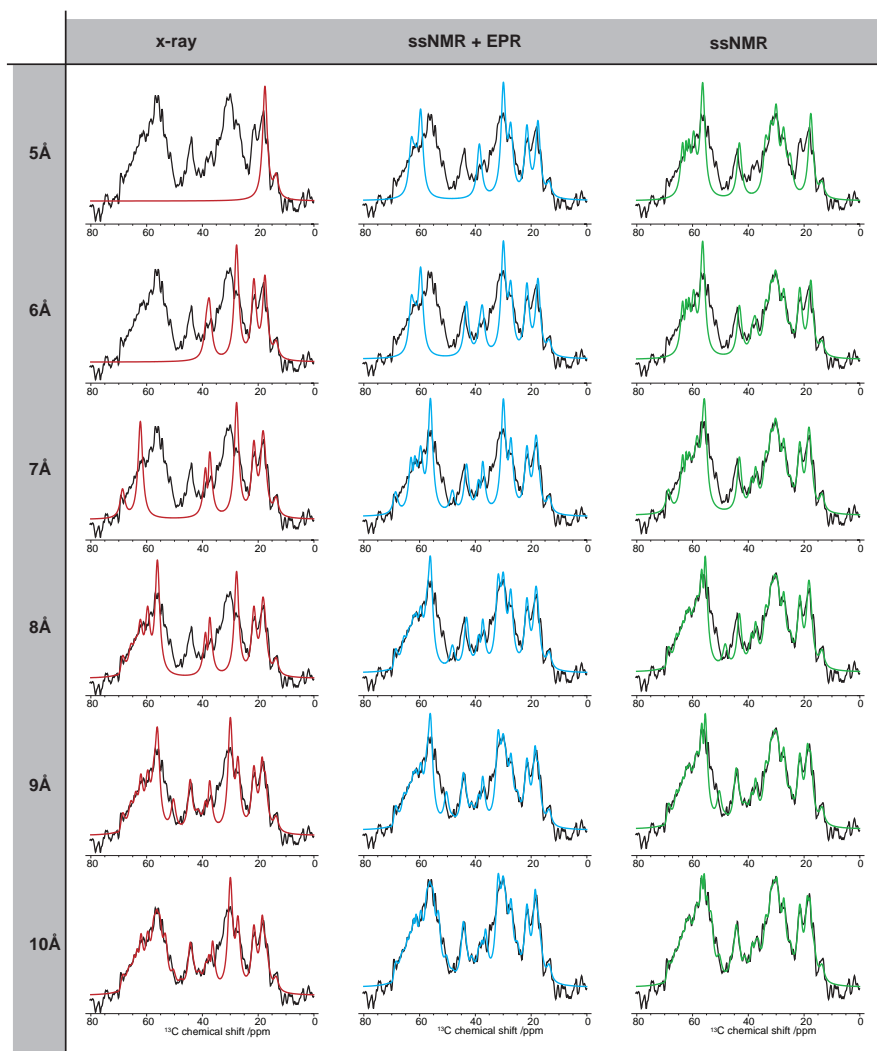
Figure A.16: Effect of the chosen intermolecular proton-proton distance cut off for the Mathemetica fit. The columns in a show the results for proton-proton distances taken from the X-ray structure [148], the model shown in figure 7.7 (which is based on EPR and ssNMR data) and a haddock [256, 257] calculation which was based only on the ssNMR (starting from the xray structure) shown in b (*contributed by K. Seidel*). Increasing the distance more spins are considered and are allowed to contribute to the fit and hence the difference (as shown in c) to the experimental spectrum gets smaller.
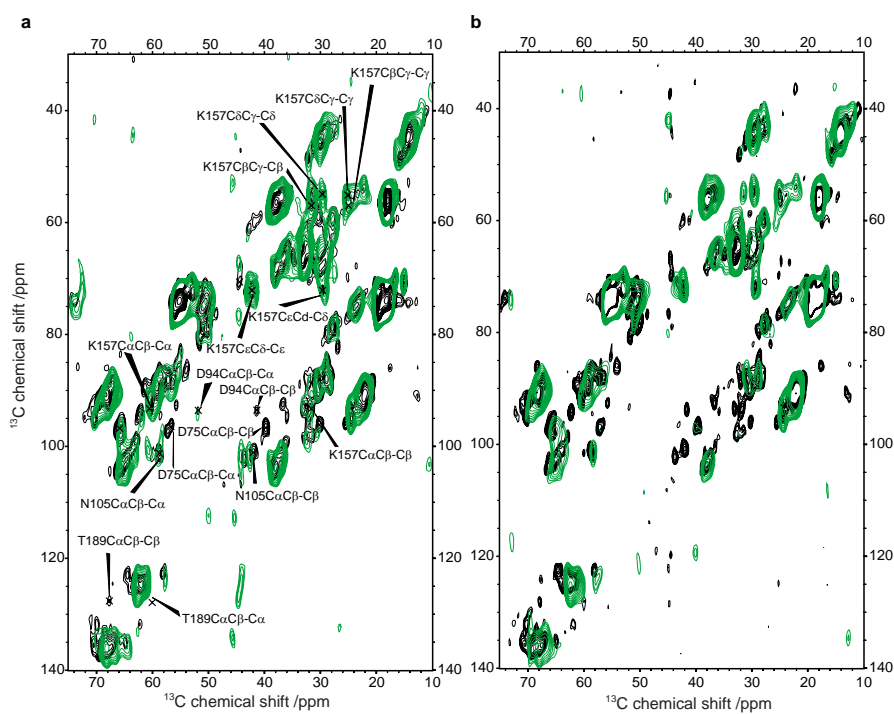
Figure A.17: Two examples of the (2Q,1Q) spectra of free NpSRII (black) and NpSRII/NpHtrII complex (green) as used for the buildup shown in figure 7.5 using a (^{13}C,^{13}C) DQ mixing time of 528 μs (a) and 816 μs (b)
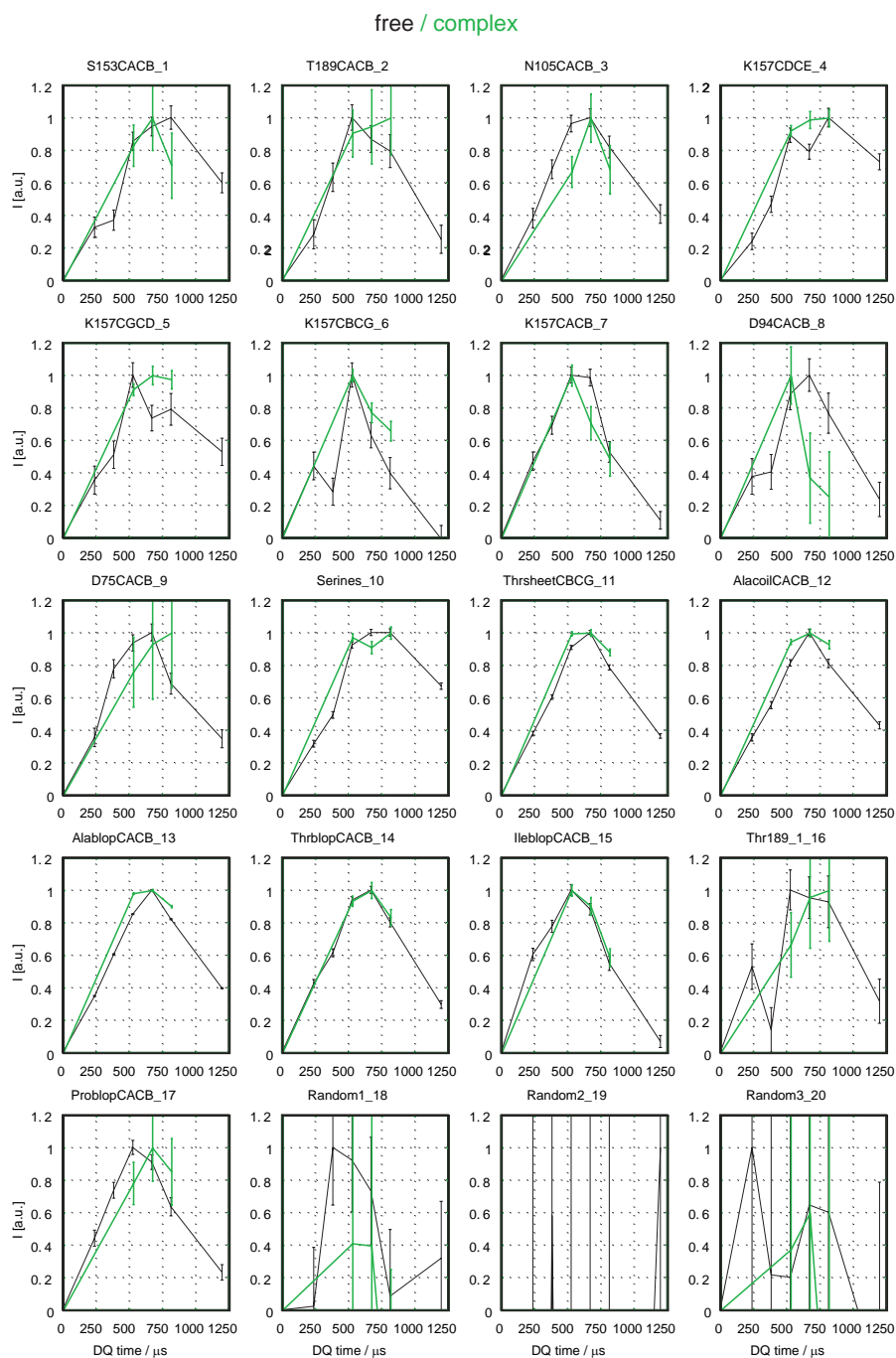
Figure A.18: (2Q,1Q) buildup curves for several resolved or unresolved (blop) peaks shown in figure A.17. The last three buildups were taken from regions with no signal and reflect the accuracy of the error margin.
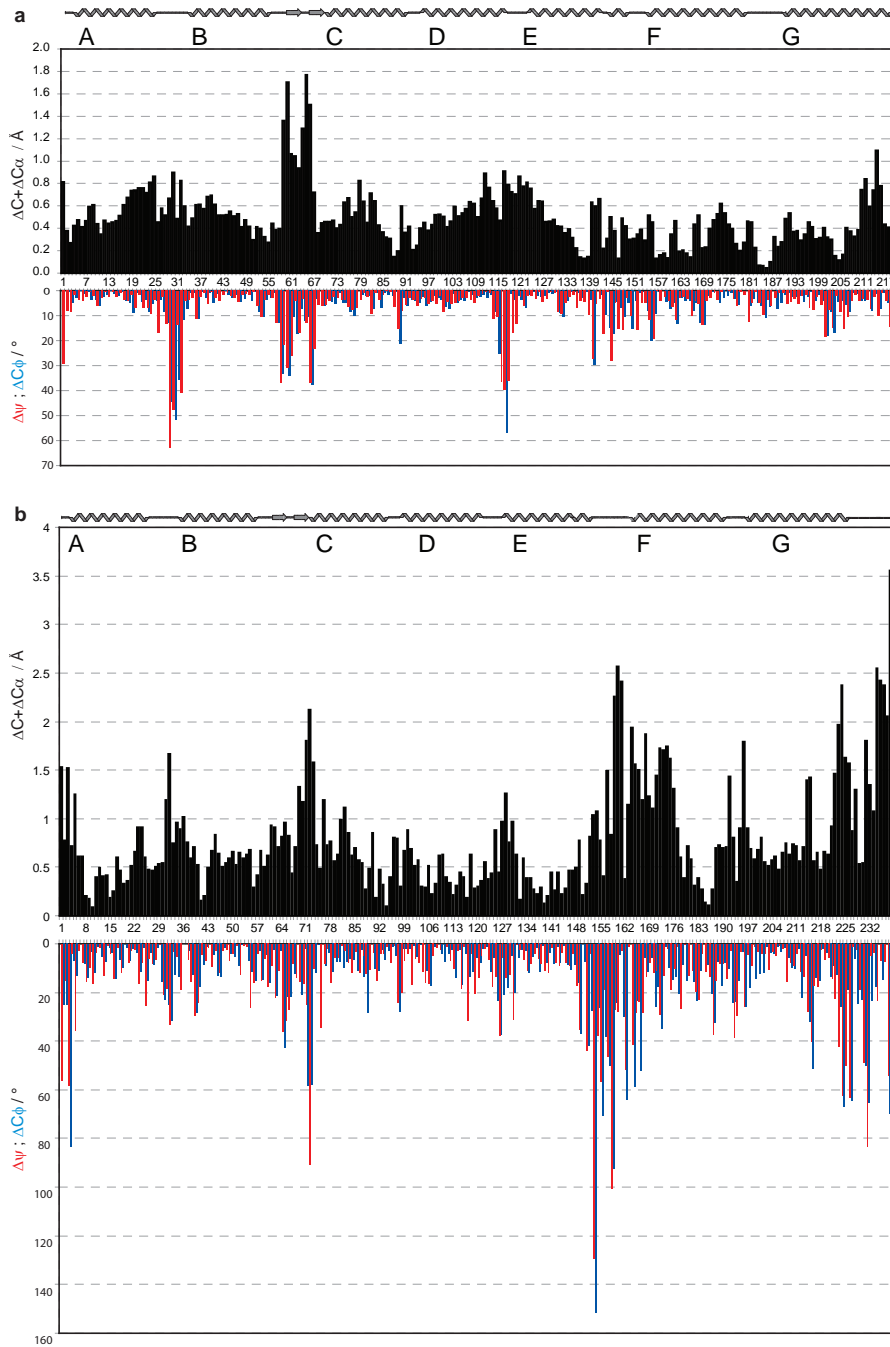
Figure A.19: (a) Upper part shows sum of difference of C' and Cα positions between the ground state (1H2S.pdb [148]) and the m-state (2F95.pdb [149]) in NpSRII. Lower part in (a) reports same difference for the torsion angles ψ and ϕ. (b) Same plot as in (a) but for BR (1CWQ.pdb [199]). Note that scaling in (a) and (b) is identical and that the position of helices is slightly different in the two rhodopsins. Most prominent difference is found for residues in the EF-loop and in the cytoplasmic half of helix F, which undergoes mayor structural alterations in BR but is not affected in NpSRII.
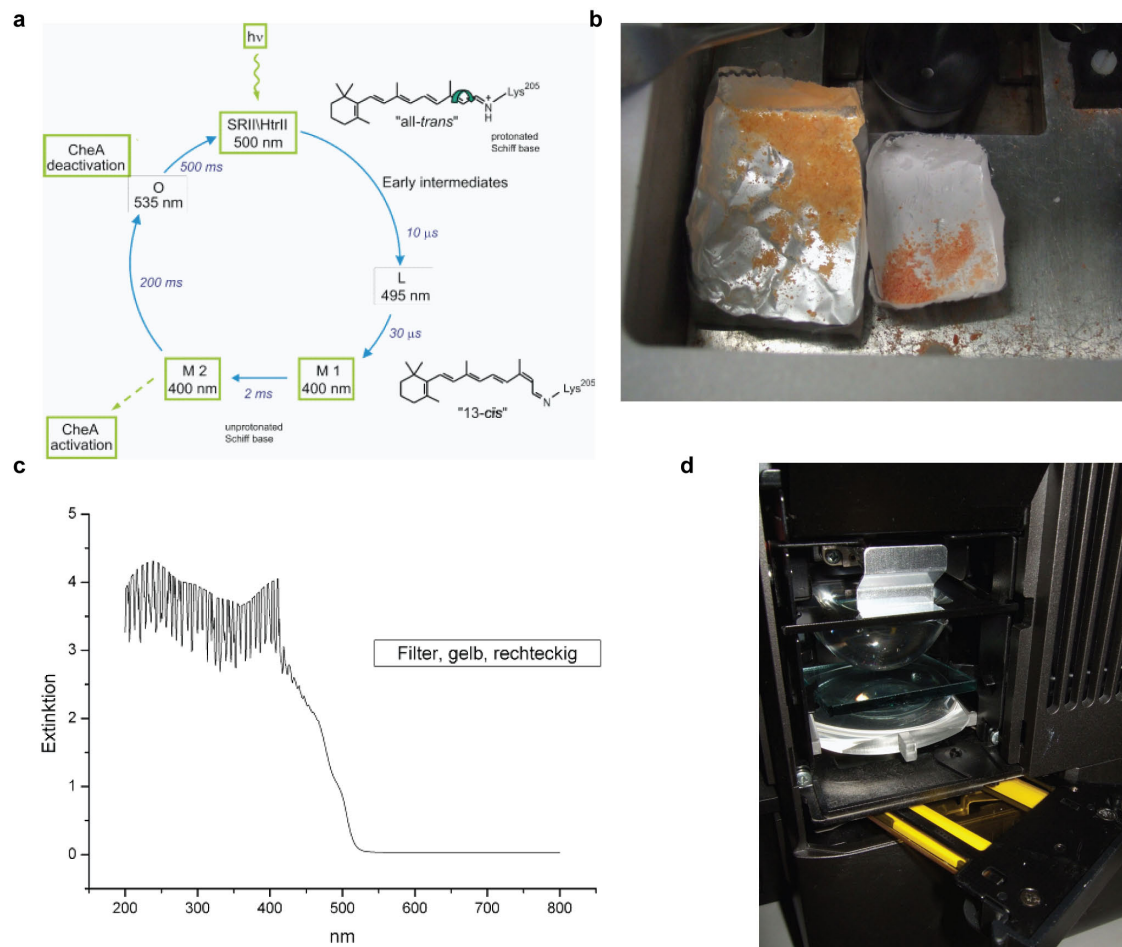
Figure A.20: (a) Photocycle of NpSRII/NpHtrII. (b) Photograph of freeze trapped, light excited, NpSRII/HtrII left box (light yellow) and not excited NpSRII/NpHtrII (orange). To increase the light accessible surface the sample was frozen and ground before excitation. The excitation itself was carried out at around -10°C by putting the powdered sample underneath a projector with a color filter. (c) Shows the excitation spectrum of the filter used. (d) Shows a photograph of the filter insert in the projector. Liquid nitrogen was used to freeze trap the sample while still excited with light. Rotor filling was done at -100°C. Sample transfer was archived by putting the rotor in liquid nitrogen and quickly inserting it inside the precooled magnet.

A.7 Supplemental Material for chapter 8
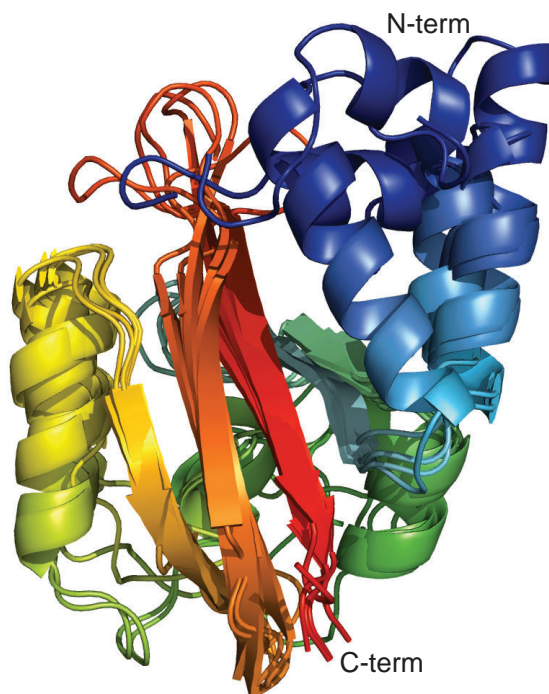
Figure A.21: Alignment of the five resulting ROBETTA structures of DcuS-[PAS$_C$]. N-terminal helix is shown in blue.
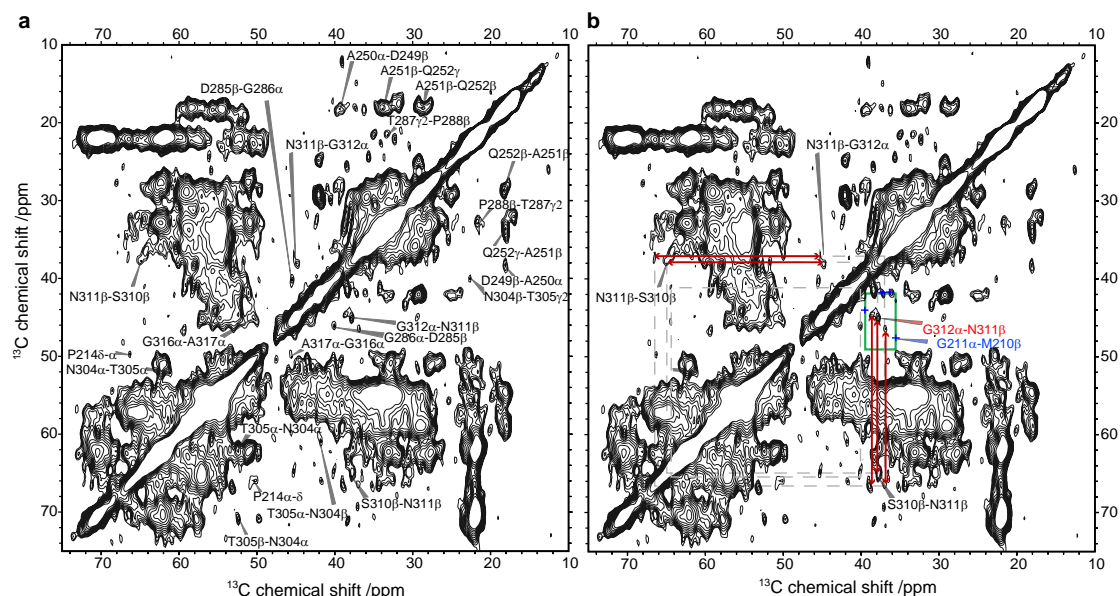
Figure A.22: (a) (^{13}C,^{13}C) SDWC of R-PAS showing sequential assignments. Due to the specific reverse labeling pattern Alanine and Threonine are the only labeled amino acids containing methyls. As a result sequential contacts involving these residues can be identified most effectively. (b) Same spectrum than in a illustrating multiple cross peaks for Asn311-Gly312. The next nearest cross peaks to G312Cα-N311Cβ, which were predicted using shiftX on all five modeled structures, are indicated as blue crosses and determine the border of the green box (in three directions). (Note that G211Cα-M210Cβ is located at the very N-terminal end and therefore is not expected to show significant signal). Hence cross signals within the green box arise, within the error of shiftX, predominantly from the G312Cα-N311Cβ contact. Multiple peaks for one single correlation are indicative for different structural confirmations and have been found in several ssNMR studies before. Additional sequential correlations from N311Cβ to S310Cβ (highlighted with red arrows) can be identified for several peaks as well and corroborate multiple (stable) confirmations of these residues in DcuS-[PAS$_C$] precipitates.

Figure A.23: (a),(b) Calculated difference spectra (using Mathematica) of spectra shown in figure 8.2c. Grey contour lines represent simulated spectrum and serve as reference. Color gradient represents difference between experimental and theoretical spectrum, based on the first modeled structure. Bot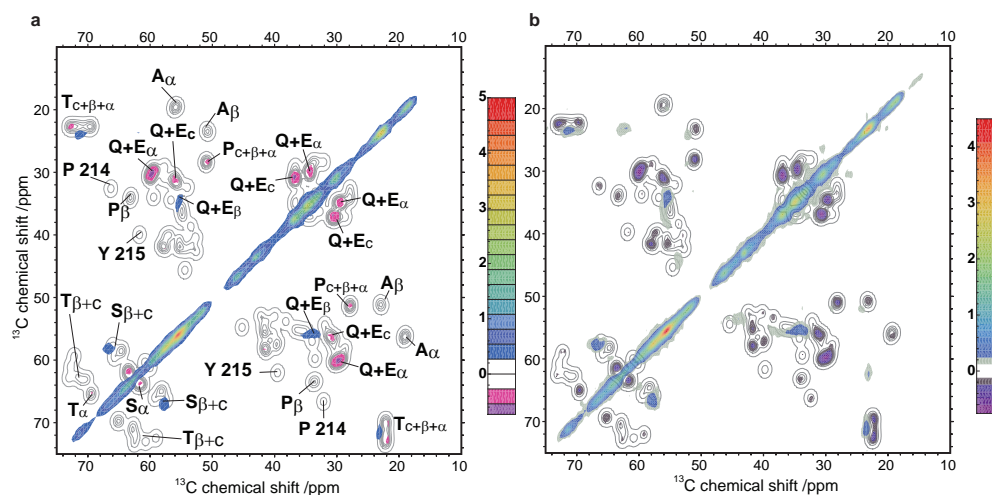h spectra were normalized to their highest peak. Note that diagonal peaks were not taken into account for the simulated spectrum and hence the diagonal to cross peak ratio was chosen to fit the experimental spectrum. In (a) 20 contour levels and a diagonal to cross peak ratio of 5 were used; (b) was plotted using 40 contour levels and a diagonal to cross peak ratio of 4.5. To reduce noise the first positive and negative color was removed. Due to the coarse intervals used in a cross peaks based on one single residue, e.g. Pro214 are not visible. However, a large negative cross peak for helical Glu and Gln $C\alpha$-$C\beta$ that is indicative for a disordered N-terminal helix, dominates the difference spectrum. Since the N-terminal cap of PAS domains was found to adopt different secondary structure elements, missing helical signal for the N terminus in DcuS-[PAS$_C$] could in principle be explained by a still ordered but different, most likely coil, secondary structure. The difference spectrum clearly shows that the region characteristic for coiled Glu and Gln is not increased. In contrast the region characteristic for Glu and Gln in β-strands is slightly increased. However, this could most likely be explained by $C\alpha$-$C\gamma$ correlations, which were not included in the simulation, but predominately affect this region. The side chain $C\beta$-$C\gamma$ cross peaks corroborate this finding. We therefore conclude that the N-terminal cap does not adopt a rigid and ordered configuration.
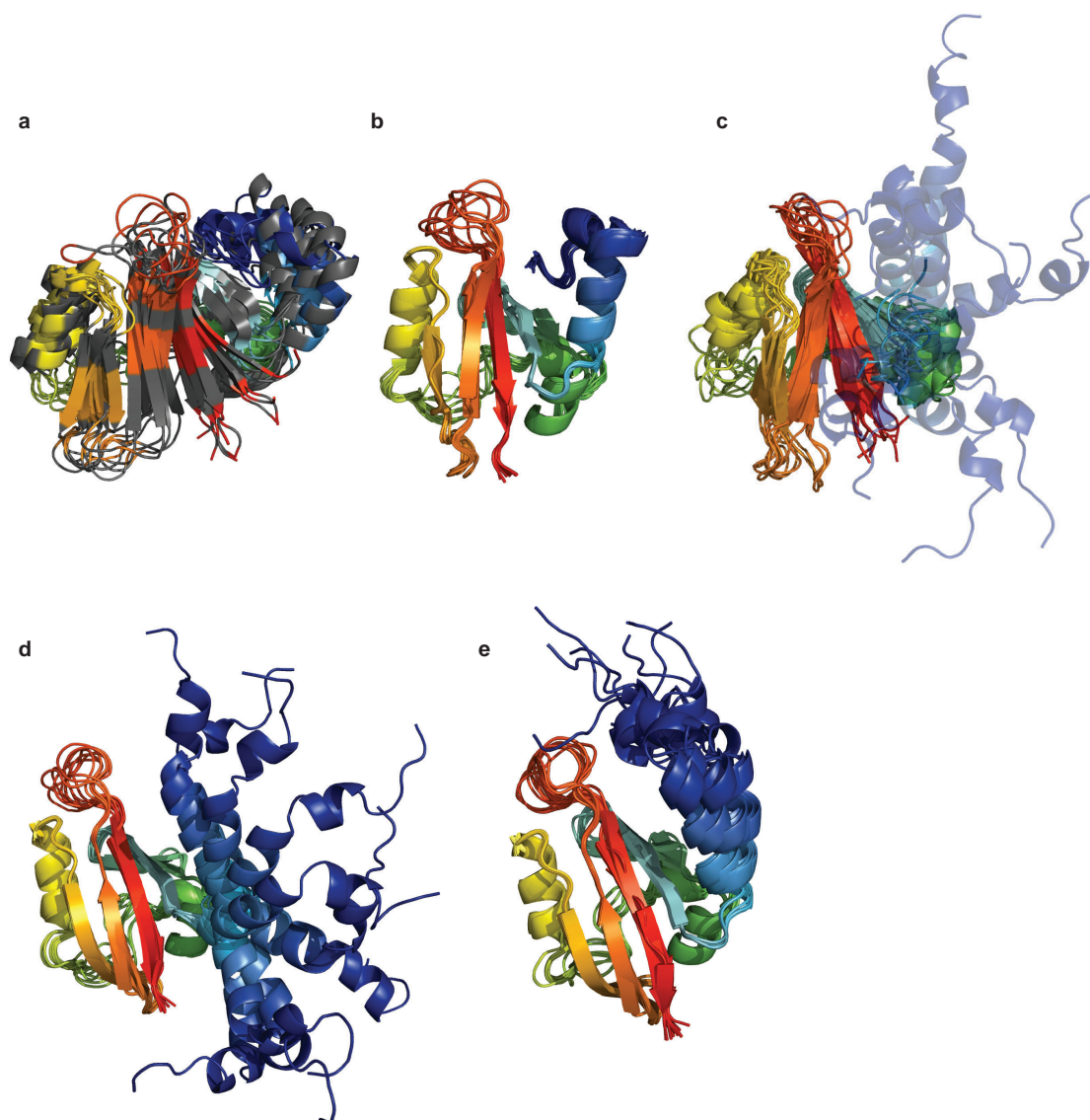
Figure A.24: Results of the structure calculations as described above. (a) based on all correlations up to 4.5 Å present in the spectrum of reverse labeled DcuS-[PAS$_C$]. Residues that are not labeled are shown in grey. (b) based on all correlations in a plus all up to 3 Å present in the spectrum of fully labeled DcuS-[PAS$_C$]. (c) based only on correlations not overlapping with signal present in the CCSD spectrum. First 22 residues are not restrained and are shown transparent. (d,e) 3D DcuS-[PAS$_C$] structural models showing the effect of the hydrogen bonds of Asn248. (d) using restrains as in b, but excluding all restrains for the N-terminal helix (residue G208-Q229). (e) same as in d but including restrains for Q229 (which is the interaction partner of N248). For each picture the ten lowest energy structures are aligned.
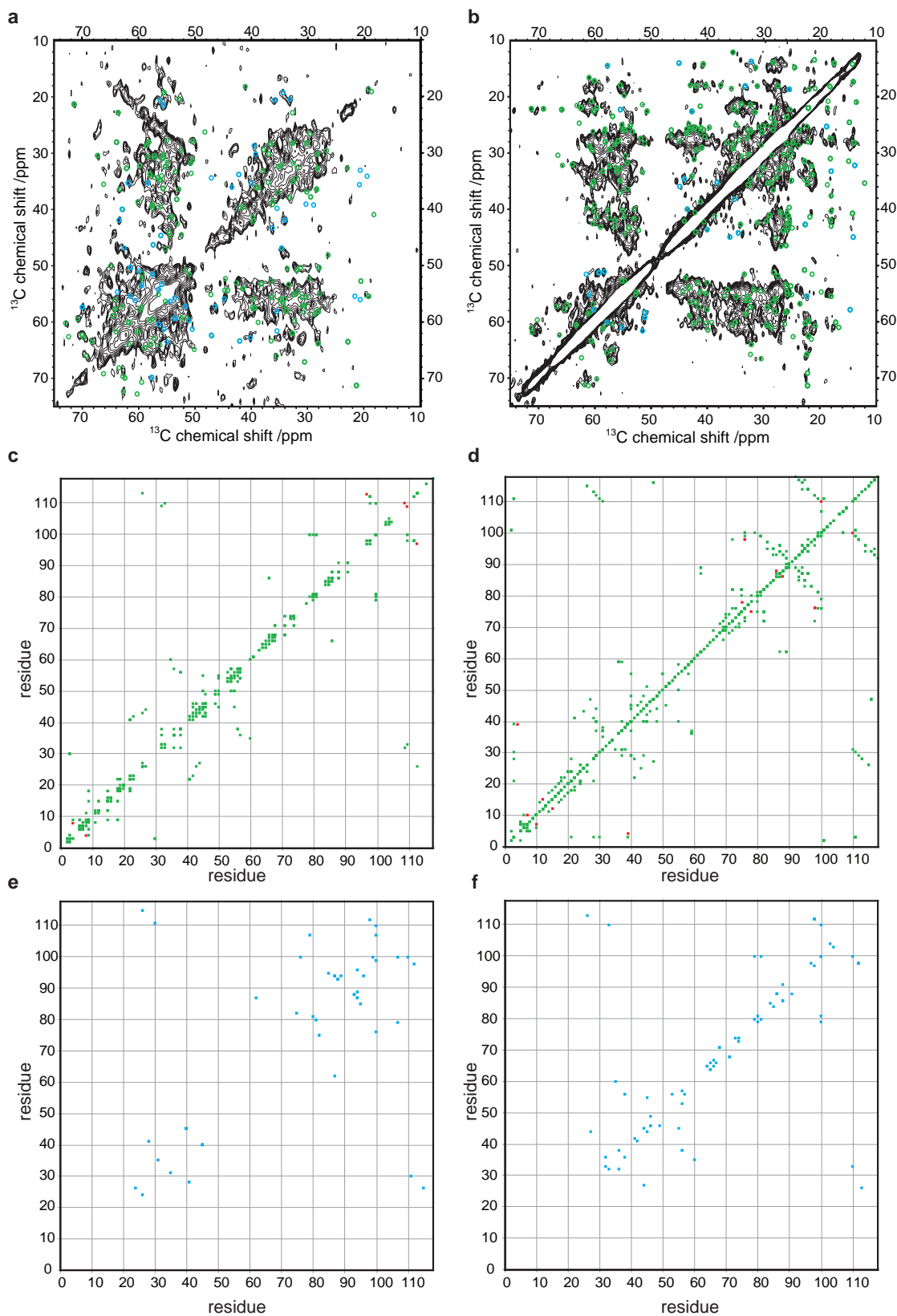
Figure A.25: (a,b) Experimental CHHC spectra of DcuS-[PAS$_C$]. Since the correlations cannot be assigned unequivocally by 2D spectroscopy, correlations predicted from the homolog model were analyzed in detail. According to ref.[49] under the experimental conditions considered here (t_{HH} = 400 / 100μs) interactions involving proton-proton distances up to 4.5/3 Åshould dominate the spectra in a/b respectively. Using an automated procedure proton contacts within this range in the first modeled structure were measured and combined with chemical shift information based predominantly on shiftX predictions to obtain a set of expected cross correlations. These correlations were compared to a manually created peak list for the experimental spectra. To account for the error of shiftX and the line width a shift tolerance of 1.5 ppm was used to automatically group the correlations in present and absent. A peak that was only detected on one side of the diagonal was classified as present. Correlations involving methyls were not treated as missing since their signal intensity is reduced due to the short CP steps in the CHHC experiments. The results for both spectra are summarized by the contact plots shown in (c,d). Predicted correlations that were consistent with the experimental spectrum are shown in green (and overlaid as open green circles on the spectra) correlations involving residues for which no signal could be detected are shown in red. Green contacts in c served as restrains for the structures shown in Supplementary Figure A.24a. Green contacts in c and d were used of the structure in Supplementary Figure A.24b. The low rmsd for the ensemble shown in Supplementary Figure A.24b does not necessarily reflect the accuracy of the presented method, since several predicted long range contacts overlap with trivial (C-C one bond) correlations. To reduce this overlap an additional peak list of the CCSD spectra shown in figure 8.2b,c were created. Contacts which are present in these spectra were hence removed from the constrain list (see also ref.[42]). To additionally account for the results of the ssNMR study regarding the N-terminal helix, correlations involving the first 22 residues were ignored. The remaining correlations are shown as contact plots in e and f and are overlaid as filled blue circles on the respective spectrum in a,b. The resulting structure is shown in figure 8.3c,d and Supplementary Figure A.24c and is most consistent with all ssNMR data.

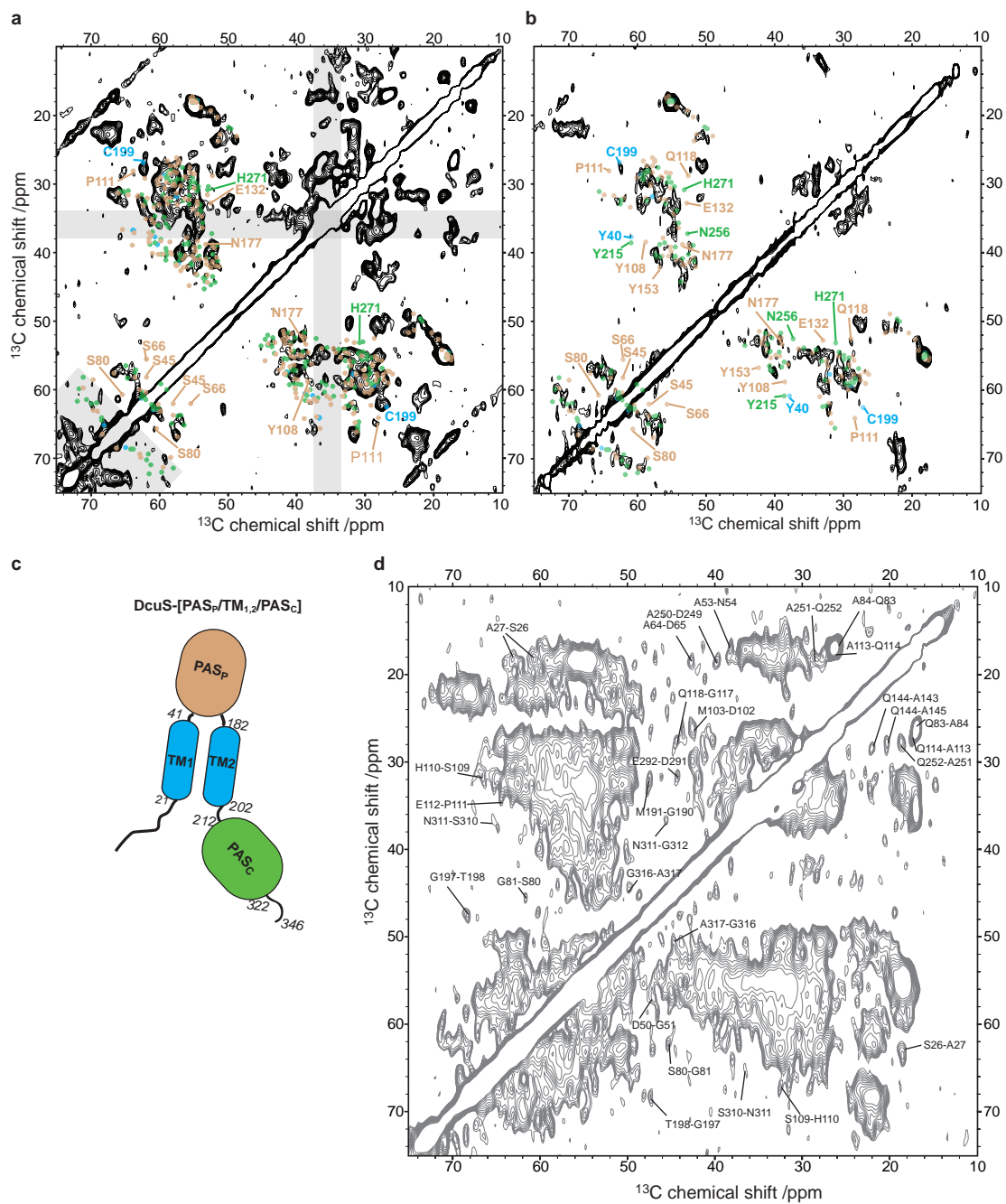**Figure A.26:** (^{13}C,^{13}C) SD spectra of (a) U-[PAS$_P$/TM$_{1,2}$/PAS$_C$] and (b) R-DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$]. Grey areas contain artifacts because of dominant lipid peaks or rotational resonance interference (Thr). Predicted cross peak positions of the domains shown in c are indicated by the respective color. (d) (^{13}C,^{13}C) SDWC of R-DcuS-[PAS$_P$/TM$_{1,2}$/PAS$_C$] showing sequential assignments for each domain. Cross signal predicted to be isolated (based on combination of liquid-state assignments and shiftX predictions) was assigned to the respective correlation. Several of these correlations (as e.g. Ser26-Ala27 in TM$_1$) remain unambiguous even in the absence of any structural assumption and are still consistent with the chemical shift expected from the structural model (e.g. α-helical for Ser26 and Ala27). The color code used in figure 8.5a is based on following criteria of a comparative analysis of the spectra in a,b and d: Dark blue residues have sequential correlations identified in d; Light blue residues have at least one isolated cross peak in a or b close to the expected position. If additionally a sequential peak could be identified the residue was colored dark blue. For purple residues no cross signal is present close to the expected position in a and b. For Ser45 and Ser66 the expected position is well isolated and does not interfere with areas containing spectral artifacts, whereas remaining ambiguities such as potential overlap or interference with spectral arti-facts can not be completely ruled out for Asn177, Tyr215, Asn256 and His271 (see labels in a,b). Grey residues are not labeled in R-DcuS and in addition to white residues, for which no structural information is available, are excluded from the analysis to increase reliability. All other residues were considered as consistent and are colored yellow.

E. coli strain or plasmid	Genotyp	Source of reference
Strains		
IMW237 (wt)	MC4100, [λ(ΦdcuB´-´lacZ)Hyb Amp^r]	Zientz et al. 1998
IMW260	MC4100 but [λ(ΦdcuB´-´lacZ)Hyb Amp^r] dcuS::Cam^r	Zientz et al. 1998
AN387	Wild-type	
C43(DE3)	Spontaneous mutation of BL21(DE3) for overproduction of membrane proteins	Miraux & Walker 1996
Plasmids		
pET28a		Novagen
pMW309	pET28a with *dcuS* (1003 bp *Nde*I/*Xho*I fragment), Kan^R	This work
pMW552	pMW309 with *dcuS* mutant encoding DcuS-PAS-N248D, Kan^R	This work
pMW181	pET28a with *dcuS* (2.2 kb *Xba*I/*Hin*dI fragment), Kan^R	Kneuper et al. 2005
pMW420	pMW181 with *dcuS* mutant encoding DcuS-PAS-D239N	This study
pMW422	pMW181 with *dcuS* mutant encoding DcuS-PAS-K259A	This study
pMW474	pMW181 with *dcuS* mutant encoding DcuS-PAS-A251C	This study
pMW475	pMW181 with *dcuS* mutant encoding DcuS-PAS-T287A	This study
pMW476	pMW181 with *dcuS* mutant encoding DcuS-PAS-R290A	This study
pMW478	pMW181 with *dcuS* mutant encoding DcuS-PAS-K297A	This study
pMW479	pMW181 with *dcuS* mutant encoding DcuS-PAS-T305A	This study
pMW480	pMW181 with *dcuS* mutant encoding DcuS-PAS-S270A	This study
pMW481	pMW181 with *dcuS* mutant encoding DcuS-PAS-W273A	This study
pMW482	pMW181 with *dcuS* mutant encoding DcuS-PAS-Q261A	This study
pMW485	pMW181 with *dcuS* mutant encoding DcuS-PAS-N248A	This study
pMW487	pMW181 with *dcuS* mutant encoding DcuS-PAS-D291A	This study
pMW488	pMW181 with *dcuS* mutant encoding DcuS-PAS-N304D	This study
pMW489	pMW181 with *dcuS* mutant encoding DcuS-PAS-L254T	This study
pMW490	pMW181 with *dcuS* mutant encoding DcuS-PAS-R289A	This study
pMW491	pMW181 with *dcuS* mutant encoding DcuS-PAS-E293Q	This study

Figure A.27: Strains of *Escherichia coli* and plasmids.

Mutation	Sequence (5´→3´)[a]
dcuStmr	CACAAGGAAGCATATGAGACATTC
dcuSPAS3	AGACCCTCGACTTACTGCATC
DcuS-PAS-D239N	GTC GTT GCC GTG <u>AAC</u> GAT CGC GGC GAG GTC
DcuS-PAS-N248A	GTC ACG CTG ATC <u>GCC</u> GAT GCC GCA CAA GAA TTG CTG
DcuS-PAS-N248D	GTC ACG CTG ATC <u>GAC</u> GAT GCCCAA GAA TTG CTG
DcuS-PAS-A251C	GTC ACG CTG ATC AAC GAT GCC <u>TGC</u> CAA GAA TTG CTG
DcuS-PAS-L254T	GAT GCC GCA CAA GAA <u>GCG</u> CTG AAT TAC CGT AAG TCG
DcuS-PAS-K259A	GAA GAG ATT ACG ATT <u>GCA</u> GAC CGG CTA TTA CTG
DcuS-PAS-Q261A	CTG AAT TAC CGT AAG TCG <u>GCG</u> GAC GAT GAG AAA CTG
DcuS-PAS-S270A	CTG TCG ACG CTA <u>GCC</u> CAC TCA TGG TCA CAG
DcuS-PAS-273A	CTA AGC CAC TCA <u>GCG</u> TCA CAG GTG GTA GAT GTC
DcuS-PAS-T287A	GTG TTA CGC GAC GGT <u>GCC</u> CCG CGC CGC GAC
DcuS-PAS-R289A	CGC GAC GGT ACC CCG <u>GCC</u> CGC GAC GAA GAG ATT ACG
DcuS-PAS-R290A	CGC GAC GGT ACC CCG CGC <u>GCC</u> GAC GAA GAG
DcuS-PAS-D291A	GGT ACC CCG CGC CGC <u>AAC</u> GAA GAG ATT ACG
DcuS-PAS-E293Q	CCG CGC CGC GAC GAA <u>CAG</u> ATT ACG ATT AAA GAC CGG
DcuS-PAS-K297A	GAA GAG ATT ACG ATT <u>GCA</u> GAC CGG CTA TTA CTG
DcuS-PAS-N304D	CGG CTA TTA CTG ATC <u>GAC</u> ACC GTT CCG GTG CGC
DcuS-PAS-T305A	CGG CTA TTA CTG ATC AAC <u>GCC</u> GTT CCG GTG CGC
[a] Oligonucleotides for mutagenesis resulting in amino acid exchanges in DcuS are underlined.	

Figure A.28: Site-directed mutagenesis of *dcuS* in the region coding for the cytoplasmatic PAS$_C$ domain of DcuS in plasmid pMW181.

B

Appendix B

Biochemical background

A proteins is a polymer of the different amino acids, thus it is also referred to as polypeptide. In general one discriminates 20 different amino acids (see figure B.1), which share the following design:

- a central carbon atom ($C\alpha$), which is covalently bound to one hydrogen (two in the case of glycine).

- an amino group (NH_3^+)

- a carboxyl group (COO^-)

- and a residue, referred to as side chain, which is specific for each amino acid.

Polypeptides are formed by the so called peptide bound, a covalent bound between the amino group of one amino acid and the carboxyl group of the following (one water molecule gets separated).

The chain of peptide bounds, including the $C\alpha$ atom, is called protein backbone. The sequence of residues attached to the backbone forms the primary structure of a protein. Parts of the backbone can adopt specific conformations which are determined by characteristic torsion angles. The most important of these so called secondary structure elements are the β-strand, an elongated backbone, and the α-helix, a spiral turning backbone. The three dimensional arrangement of all atoms is called the tertiary structure, whereas the structural arrangement of protein domains in respect to each other is

referred to as quaternary structure. A more detailed overview can for example be found in [1].
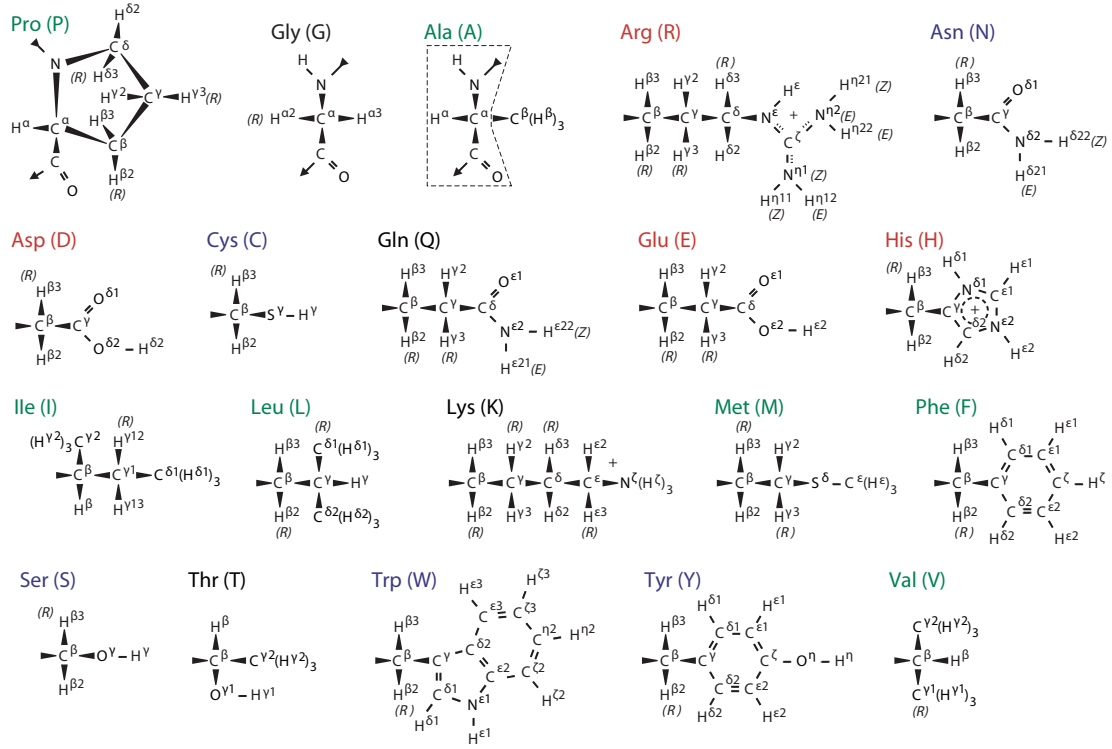
Figure B.1: Natural occuring amino acids. Picture and nomenclature adapted from ref [258].Color code refers to hydrophobic (green), polar (blue) and charged (red) amino acids.

C Appendix C
Applied pulse programs

$\frac{\pi}{2}$-pulse (1D)

```
1  ze
2  d1
   1u pl1:f1
   p1:f1 ph1
;aquisition
   d3
   gosc ph31
   lo to 2 times ns
   100m wr #0
HaltAcqu, 1m
exit


ph1 = 0 1 2 3
ph31= 0 1 2 3
```

Cross polarization (CP)(1D)

```
1  ze
2  d1
   3u pl2:f2
   p2:f2 ph1
   3u pl5:f1 pl6:f2
   (p15 ph2):f1 (p15:sp0 ph0):f2
;aquisition
   3u pl12:f2
   3u cpd2:f2
   10u:f1 ph30
   3u adc ph31
   aq
   1m do:f2
   rcyc=2
   100m wr #0
exit


ph0 = 0
ph1 = 1 1 1 1   3 3 3 3
ph2 = 0 1 2 3
ph3 = 1 2 3 0
ph30= 0
ph31= 0 1 2 3   2 3 0 1
```

Spin diffusion (CCSD)(2D)

```
1  ze
2  d1 do:f2
1u fq=0:f2
;cp
   1u pl2:f2
   p2:f2 ph1
   1u pl5:f1 pl6:f2
   (p15 ph2):f1 (p15:spf0 pl6 ph0):f2
;t1 evolution
   2u pl12:f2
   2u cpd2:f2
   d0
   2u do:f2
;proton-driven spindiffusion
   1u pl1:f1
   p1:f1 ph3
   d11
   p1:f1 ph4
;aquisition
   5u pl12:f2
   5u cpd2:f2
   gosc ph31
   1m do:f2
   lo to 2 times ns
   100m wr #0 if #0 zd
   1m id0
   1m ip2
   lo to 1 times td1
exit

ph0 = 0
ph1 = 1 1 1 1  1 1 1 1  3 3 3 3  3 3 3 3
ph2 = 0 0 0 0  2 2 2 2
ph3 = 1
ph4 = 3 0 1 2
ph31= 0 1 2 3  2 3 0 1  2 3 0 1  0 1 2 3
```

NCC (specific N-CO or N-Cα(2D)

```
1  ze
2  d1
   1u pl2:f2
   p2:f2 ph1
;cp f2 -> f3
   1u pl5:f3 pl6:f2
   (p15 ph2):f3 (p15:spf0 pl6 ph0):f2
;t1 evolution f3
   1u pl12:f2
   1u cpds2:f2
   dwone
   2u pl13:f2
   2u cw:f2
;cp f3 -> f1
   ; 1u fq=cnst20:f1
   1u pl15:f3 pl16:f1
   (p25 ph3):f3 (p25:spf1 pl16 ph4):f1
   ; 1u fq=0:f1
;DARR recoupling
   2u pl1:f1
   2u pl29:f2
   p1:f1 ph5
   d10
   p1:f1 ph7
;aquisition
   5u pl12:f2
   5u cpds2:f2
   gosc ph31
   1m do:f2
   lo to 2 times ns
   100m wr #0 if #0 zd
   "dwone = dwone + d0"
   1m ip2
   lo to 1 times td1
 HaltAcqu, 1m   ; jump address for protection,
exit

 ph0 = 0
 ph1 = 1 1 1 1   3 3 3 3
 ph2 = 0
```

HC-INEPT (2D)

```
ph3 = 0 0 0 0   0 0 0 0   2 2 2 2   2 2 2 2
ph4 = 0 1 2 3
ph5 = 1 2 3 0
ph7 = 0 1 2 3
ph31= 0 1 2 3   2 3 0 1   2 3 0 1   0 1 2 3
```

```
;INEPT delays
"d4=(1/4*cnst2)"
"d3=(1/cnst3*cnst2)"
;Compenstaion delays
"d5=p1-p2"
"d6=d5/2"


1  ze
2  d1
;90 1H
   2u fq=0:f2
   3u pl2:f2 pl1:f1
   p2:f2 ph1
;t1 evolution
   d0
   (p1*2 ph0):f1
   d0
;spinecho
   d4
   (d5 p2*2 ph2):f2 (p1*2 ph4):f1
   d4
;PT
   (d6 p2 ph3):f2 (p1 ph5):f1
;spinecho
   d3
   (d5 p2*2 ph2):f2 (p1*2 ph6):f1
   d3 pl12:f2
;aquisition
   2u cpds2:f2
   gosc ph31
   1m do:f2
   lo to 2 times ns
   100m wr #0 if #0 zd
   1m id0
   1m ip1
   lo to 1 times td1
HaltAcqu, 1m
exit


ph0 = 0
ph1 = 0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph2 = 0 2
```

HCC-INEPT TOBSY (2D)

```
ph3 = 1 1 3 3
ph4 = 0 2
ph5 = 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph6 = 0 2 0 2 1 3 1 3
ph31= 0 0 2 2 1 1 3 3
```

```
;INEPT echo delays
"d4=1/(4*cnst2)"
"d3=1/(cnst3*cnst2)"
;compensation delays
"d5=p1-p2"
"d6=d5/2"
;TOBSY P9(3,1) mixing
;cnst31 = MAS rate
;p11 = 90 pulse for TOBSY
;l5 = 9 - 3 rotor periods
"p11=1s/(24*cnst31)"

1  ze
2  d1
;90 1H
   2u fq=0:f2
   2u fq=cnst20:f1
   2u pl2:f2 pl1:f1
   p2:f2 ph1
;t1 evolution
   d0
   (p1*2 ph0):f1
   d0
;spinecho
   d4
   (d5 p2*2 ph2):f2 (p1*2 ph4):f1
   d4
;PT
   (d6 p2 ph3):f2 (p1 ph5):f1
;spinecho
   d3
   (d5 p2*2 ph2):f2 (p1*2 ph6):f1
   d3
;TOBSY P9(3,1) mixing
   2u pl11:f1
   p11:f1   ph10
3  p11*1:f1 ph7
   p11*4:f1 ph8^
   p11*3:f1 ph7^
   lo to 3 times l5
   p11:f1   ph11
;aquisition
```

```
   2u pl12:f2
   2u cpds2:f2
   2u fq=0:f1
   10u
   gosc ph31
   1m do:f2
   1m rpp7
   1m rpp8
   lo to 2 times ns
   100m wr #0 if #0 zd
   1m id0
   1m ip1
   lo to 1 times td1
HaltAcqu, 1m
exit


ph0 = 0
ph1 = 0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph2 = 0 2
ph3 = 1 1 3 3
ph4 = 0 2
ph5 = 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph6 = 0 2 0 2 1 3 1 3
ph7 = (16384) 0    1820  3641  5461
        7282 9102 10923 12743 14566
ph8 = (16384) 8192 10012 11833 13653
        15474 910  2730  4551  6371
ph10= 1 1 1 1 2 2 2 2 3 3 3 3 0 0 0 0
ph11= 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph31= 0 0 2 2 1 1 3 3
```

NCα-INEPT TOBSY (2D)

```
;cnst2 = J_HN coupling
;cnst4 = J_NC coupling
;cnst3 = 6
;cnst5 = N-C refocusing (can be 4 or bigger, i.e. 6)
;cnst31= H rf field decoupling
;INEPT H-N echo
"d4=(1s)/(4*cnst2)"
"d3=(1s)/(cnst3*cnst2)"
;compensation H-N delay
;p1>p2
"d13=p3-p2"
"d14=d13/2"
;INEPT N-C
"d5=(1s)/(4*cnst4)"
"d6=(1s)/(cnst5*cnst4)"
;compensation N-C delay
;p4>p3
"d15=p4-p3"
"d16=d15/2"
;compensation H-C delay for t1
"d17=p1-p2"
;WALTZ16 or GARP decoupling
;p30 90 pulse
;"cnst31=10000"
;"p30=(1s)/(4*cnst31)"


1  ze
2  d1
   2u fq=0:f2
   2u fq=cnst20:f1
   2u pl1:f1
   2u pl2:f2
   2u pl3:f3
;90 on H
   p2:f2 ph1
;direct INEPT H-N
;spinecho
   d3
   (p3*2 ph2):f3 (d13 p2*2 ph3):f2
   d3
;PT
   (p3 ph4):f3   (d14 p2 ph5):f2
```

```
;spinecho
   d4
   (p3*2 ph6):f3 (d13 p2*2 ph7):f2
   d4
;15N t1 evolution
   d0
   (d17 p2*2 ph8):f2 (p1*2 ph9):f1
   d0
;H dec
   1u pl13:f2
   1u cpds1:f2
;INEPT N-C selective
;spinecho
   d5 pl4:f1
   (d15 p3*2 ph10):f3 (p4*2 ph11):f1
   d5
;PT
   (d16 p3 ph12):f3    (p4 ph13):f1
;spinecho
   d6
   (d15 p3*2 ph14):f3 (p4*2 ph15):f1
   d6
  1u fq=0:f1
;aquisition
  1u pl12:f2
  1u cpds2:f2
  5u
  gosc ph31
  1m do:f2
  lo to 2 times ns
  100m wr #0 if #0 zd
  1m id0
  1m ip2
  1m ip4
  1m ip6
  lo to 1 times td1
HaltAcqu, 1m
exit

ph1  = 1
ph2  = 1
ph3  = 1
ph4  = 0
ph5  = 0 0 2 2
```

```
ph6  = 1
ph7  = 1
ph8  = 0
ph9  = 0
ph10 = 1
ph11 = 1
ph12 = 0 2
ph13 = 0
ph14 = 1
ph15 = 1
ph31 = 0 2 2 0
```

N-CO-Cα-INEPT TOBSY (2D)

```
;hnc-inept-tobsy_2d.me
;Selective HNCO INEPT + CO-CA TOBSY
;cnst2 = J_HN coupling
;cnst4 = J_NC coupling
;cnst3 = 4
;cnst5 = 4
;cnst20= offset for TOBSY
;cnst30= MAS rate
;INEPT H-N echo
"d4=(1s)/(4*cnst2)"
"d3=(1s)/(cnst3*cnst2)"
;compensation H-N delay
;p1>p2
"d13=p3-p2"
"d14=d13/2"
;INEPT N-CO echo
"d5=(1s)/(4*cnst4)"
"d6=(1s)/(cnst5*cnst4)"
;compensation N-CO delay
;p1>p3
"d15=p4-p3"
"d16=d15/2"
;compensation H-C delay for t1
"d17=p1-p2"
;WALTZ16 or GARP decoupling
;p30 90 pulse
;"cnst31=10000"
;"p30=(1s)/(4*cnst31)"
;TOBSY P9(3,1) mixing
;p15 = 90 pulse for TOBSY
"p15=1s/(24*cnst30)"

1  ze
2  d1
   2u fq=0:f2
   2u fq=cnst20:f1
   2u pl1:f1
   2u pl2:f2
   2u pl3:f3
;90 on H
   p2:f2 ph1
;direct INEPT H-N
;spinecho
   d3
   (p3*2 ph2):f3 (d13 p2*2 ph3):f2
   d3
;PT
   (p3 ph4):f3   (d14 p2 ph5):f2
;spinecho
   d4
   (p3*2 ph6):f3 (d13 p2*2 ph7):f2
   d4
;15N t1 evolution
   d0
   (d17 p2*2 ph8):f2 (p1*2 ph9):f1
   d0
;H dec
   1u pl13:f2
   1u cpd1:f2
;INEPT N-CO selective
;spinecho
   d5 pl4:f1
   (d15 p3*2 ph10):f3 (p4*2 ph11):f1
   d5
;PT
   (d16 p3 ph12):f3   (p4 ph13):f1
;spinecho
   d6
   (d15 p3*2 ph14):f3 (p4*2 ph15):f1
   d6 pl1:f1
;TOBSY P9(3,1) mixing
   p1:f1  ph16
   1u fq=0:f1
   1u pl15:f1
   1u do:f2
3  p15*1:f1 ph18
   p15*4:f1 ph19^
   p15*3:f1 ph18^
   lo to 3 times l5
   p15:f1 ph17
;aquisition
   1u pl12:f2
   1u cpds2:f2
   5u
```

```
    gosc ph31

    1m do:f2

    1m rpp18

    1m rpp19

    lo to 2 times ns

    100m wr #0 if #0 zd

    1m id0

    1m ip2

    1m ip4

    1m ip6

    lo to 1 times td1

HaltAcqu, 1m

exit


ph1  = 1

ph2  = 1

ph3  = 1

ph4  = 0

ph5  = 0 0 2 2

ph6  = 1

ph7  = 1

ph8  = 0

ph9  = 0

ph10 = 1

ph11 = 1

ph12 = 0 2

ph13 = 0

ph14 = 1

ph15 = 1

ph16 = 1 1 1 1  3 3 3 3

ph17 = 0 0 0 0  0 0 0 0

       1 1 1 1  1 1 1 1

       2 2 2 2  2 2 2 2

       3 3 3 3  3 3 3 3

ph18 = (16384) 0    1820  3641  5461

       7282 9102 10923 12743 14566

ph19 = (16384) 8192 10012 11833 13653

       15474 910  2730  4551  6371

ph30 = 0

ph31 = 0 2 2 0  2 0 0 2

       1 3 3 1  3 1 1 3

       2 0 0 2  0 2 2 0

       3 1 1 3  1 3 3 1
```

N/CHHC (2D)

```
1  ze

2  d1

;cp to carbons

    3u pl2:f2

    p2:f2 ph1

    3u pl5:f1 pl6:f2

    (p15 ph2):f1 (p15:spf0 pl6 ph0):f2

;t1 evolution

    3u pl12:f2

    3u cpd2:f2

3  d0

    3u do:f2

;cp to protons

    3u pl15:f1 pl16:f2

    (p16 ph5):f1 (p16:spf1 pl16 ph6):f2

;mixing protons

    3u pl2:f2

    p2:f2 ph7

    d5

    p2:f2 ph8

;back to the carbons

    3u pl15:f1 pl16:f2

    (p17 ph10):f1 (p17:spf2 pl16 ph9):f2

;aquisition

    5u pl12:f2

    5u cpd2:f2

    gosc ph31

    1m do:f2

    lo to 2 times ns

    100m wr #0 if #0 zd

    1m id0

    1m ip2

    lo to 1 times td1

exit


ph0 = 0

ph1 = 1 1 1 1  1 1 1 1  1 1 1 1  1 1 1 1

      3 3 3 3  3 3 3 3  3 3 3 3  3 3 3 3

ph2 = 0 2

ph5 = 0

ph6 = 1 1 3 3

ph7 = 2
```

```
ph8 = 0
ph9 = 1
ph10= 0 0 0 0  1 1 1 1  2 2 2 2  3 3 3 3
ph31= 0 2 2 0  1 3 3 1  2 0 0 2  3 1 1 3
      2 0 0 2  3 1 1 3  0 2 2 0  1 3 3 1
```

DQ/SQ SPC5 (2D)

```
1  ze
2  d1 do:f2
    2u fq=0:f2
   1m rpp6
   1m rpp7
   1m rpp8
   1m rpp9
10  0.2u ipp8
    0.2u ipp9
    lo to 10 times l5
;90 on H
   2u pl2:f2
   p2:f2 ph1
;cp H (f2) to X (f1)
   2u pl5:f1 pl6:f2
   (p15 ph2):f1 (p15:spf0 pl6 ph0):f2
;Turn on H decoupling for SPC5
   2u pl13:f2
   2u cpds3:f2
;90 on X
   2u pl1:f1
   p1:f1 ph4
;2QC excitation - SPC5
   2u pl11:f1
3  p11*1:f1 ph6
   p11*4:f1 ph7^
   p11*3:f1 ph6^
   lo to 3 times l5
;2QC reconversion - SPC5
4  p11*1:f1 ph8
   p11*4:f1 ph9^
   p11*3:f1 ph8^
   lo to 4 times l5
;90 on X
   2u pl1:f1
   p1:f1 ph5
;Change H decoupling power
   1u do:f2
   2u pl12:f2
   2u cpds2:f2
;acquisition
   10u
```

```
    gosc ph31
    1m do:f2
    1m ip8*5
    1m ip9*5
    lo to 2 times ns
    100m wr #0
HaltAcqu, 1m
exit


ph0 = 1
ph1 = 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
ph2 = 0
ph4 = 1 1 1 1  1 1 1 1  1 1 1 1  1 1 1 1
      1 1 1 1  1 1 1 1  1 1 1 1  1 1 1 1
      3 3 3 3  3 3 3 3  3 3 3 3  3 3 3 3
      3 3 3 3  3 3 3 3  3 3 3 3  3 3 3 3
ph5 = 0 0 0 0  1 1 1 1  2 2 2 2  3 3 3 3
ph6 = (20)  0  4  8 12 16 10 14 18  2  6
ph7 = (20) 10 14 18  2  6  0  4  8 12 16
ph8 = (20)  0  4  8 12 16 10 14 18  2  6
ph9 = (20) 10 14 18  2  6  0  4  8 12 16
ph30= 0
ph31= 0 2 0 2  1 3 1 3  2 0 2 0  3 1 3 1
      2 0 2 0  3 1 3 1  0 2 0 2  1 3 1 3
      2 0 2 0  3 1 3 1  0 2 0 2  1 3 1 3
      0 2 0 2  1 3 1 3  2 0 2 0  3 1 3 1
```

Water edited CC (2D)

```
1  ze
2  d1 fq=0:f2
;90 degree
   1u pl2:f2
   p2:f2 ph1
;T2 filter
   d6
   p2*2:f2 ph3
   d6
; longitudinal HH mix
  p2:f2 ph5
  d5
  p2:f2 ph6
;cp f2 -> f1
   1u pl5:f1 pl6:f2
   (p15 ph2):f1 (p15:spf0 pl6 ph0):f2
;t1 evolution
   2u pl12:f2
   2u cpd2:f2
   d0
   2u do:f2
;proton-driven spindiffusion
   1u pl1:f1
   p1:f1 ph7
   d11
   p1:f1 ph8

;aquisition
   5u pl12:f2
   5u cpd2:f2
   gosc ph31
   1m do:f2
   lo to 2 times ns
   100m wr #0 if #0 zd
   1m id0
   1m ip2
   lo to 2 times td1
exit

ph0 = 0
ph1 = 1 3
ph2 = 0 0 2 2
```

```
ph3 = 0 2 2 2  2 0 0 0
ph5 = 1
ph6 = 1 1 1 1  3 3 3 3
ph7 = 1
ph8 = 0 0 0 0  0 0 0 0  1 1 1 1  1 1 1 1
      2 2 2 2  2 2 2 2  3 3 3 3  3 3 3 3
ph31= 0 2 2 0  2 0 0 2  1 3 3 1  3 1 1 3
      2 0 0 2  0 2 2 0  3 1 1 3  1 3 3 1
```

D

Appendix D

Experimental parameter

Table D.1 lists the most important experimental parameters used for all spectra presented in chapters 4-8, i.e. B_0, MAS-rate, temperature, total experimental time, duration of key mixing step (if present) and window function with corresponding parameter used for processing.

D Experimental parameter

Spectrum	Pulsprogram	$\omega_0(^1\text{H})/\text{MHz}$	ω_r/Hz	T /°C	$\smile t_{exp}$ /h	t_{mix}	wdf*
fig 4.1 (black)	CCSD	800	10500	-13	14	20ms	Qsine 4
fig 4.1 (red)	CCSD	800	10500	-13	17	20ms	Qsine 4
fig 4.3 (red)	CCSD	800	12500	-13	18	150ms	Qsine 3
fig 4.3 (black/blue)	CCSD	800	12500	0	22.5	150ms	Qsine 3
fig 4.4	CCSD	600	10500	-16	15	20ms	Qsine 4
fig 6.2a	CCSD	800	10500	-13	17	15ms	Qsine 3.5
fig 6.2b	CCSD	800	12500	0	67	150ms	Qsine 2.5
fig 6.3c	NCC	800	10500	-13	44	40ms	Qsine 2.5
fig 6.3d	NCC	800	10500	-13	61	50ms	EM 50
fig 6.6 (green)	H_2O CC	800	10500	-13	85	4ms	EM 150
fig 6.7a (red)	J-HC	600	8333	0	4		Qsine 3.5
fig 6.7a (black)	J-HCC	600	8333	0	44		Qsine 3.5
fig 6.7b	J-NCα	600	8333	0	107		EM 50
fig 6.7c	J-NCC	600	8333	0	125		EM 50
fig 7.3a,b (blue)	CCSD	800	10500	-42	34	15	Qsine 4.2
fig 7.3c (blue)	CCSD	600	11000	-28	20	15	Qsine 3.5
fig 7.4 (green)	CCSD	800	10500	-12	43	15	Qsine 3.5
fig 7.5a (green)	J-HC	600	8333	0	6		Qsine 3.5
fig 7.6	NHHC	800	7000	-43	210	0.5ms	LB 150
fig 7.9a,b	CCSD	600	8500	-43	53	40ms	Qsine 3
fig 7.9c,d	CCSD	600	8500	-43	47	350ms	Qsine 3.5
fig 8.2b	CCSD	800	10500	-15	43	15ms	Qsine 3.5
fig 8.2c	CCSD	800	10500	-15	28	15ms	Qsine 2.5
fig 8.2d,e	CHHC	800	10500	-42	161	0.4ms	EM 80
fig 8.4a,b,c	CCSD	600	7500	-12	21	15ms	Qsine 3
fig 8.4d	CCSD	800	10600	-12	63	15ms	Qsine 2.5

Table D.1: Experimental parameters for spectra shown throughout this work.(* wdf: window function used for processing)

Bibliography

[1] L. Stryer, J.M. Berg, and J.L. Tymoczko. *Biochemistry.* W.H. Freeman & Co, 2002.

[2] Erik Lindahl and Mark Sp Sansom. Membrane proteins: molecular dynamics simulations. *Current Opinion in Structural Biology*, Apr 2008.

[3] Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne. The protein data bank. *Nucleic Acids Research*, 28(1):235–242, January 1, 2000 2000.

[4] J. Chandrashekar, K. L. Mueller, M. A. Hoon, E. Adler, L. Feng, W. Guo, C. S. Zuker, and N. J. Ryba. T2rs function as bitter taste receptors. *Cell*, 100(6):703–711, Mar 2000.

[5] Malin C Lagerström and Helgi B Schiöth. Structural diversity of g protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov*, 7(4):339–357, Apr 2008.

[6] Harald Wajant. The fas signaling pathway: more than a paradigm. *Science*, 296(5573):1635–1636, May 2002.

[7] Bruce Alberts. The promise of cancer research. *Science*, 320(5872):19, Apr 2008.

[8] Stephen H White. The progress of membrane protein structure determination. *Protein Science*, 13(7):1948–1949, Jul 2004.

[9] Matthias Gralle and Sérgio T Ferreira. Structure and functions of the human amyloid precursor protein: the whole is more than the sum of its parts. *Prog Neurobiol*, 82(1):11–32, May 2007.

Bibliography

[10] Christopher M Dobson. Protein aggregation and its consequences for human disease. *Protein and Peptide Letters*, 13(3):219–227, 2006.

[11] Fred E Cohen and Jeffery W Kelly. Therapeutic approaches to protein-misfolding diseases. *Nature*, 426(6968):905–909, Dec 2003.

[12] Christopher M Dobson. Protein folding and misfolding. *Nature*, 426(6968):884–890, Dec 2003.

[13] Christian Wasmer, Adam Lange, Hélène Van Melckebeke, Ansgar B Siemer, Roland Riek, and Beat H Meier. Amyloid fibrils of the het-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science*, 319(5869):1523–1526, Mar 2008.

[14] Adam Lange, Karin Giller, Sönke Hornig, Marie-France Martin-Eauclaire, Olaf Pongs, Stefan Becker, and Marc Baldus. Toxin-induced conformational changes in a potassium channel revealed by solid-state nmr. *Nature*, 440(7086):959–962, Apr 2006.

[15] I. I. Rabi and V. W. Cohen. The nuclear spin of sodium. *Physical Review*, 43(7):0582–0583, April 1933.

[16] Milestones. Physics is set spinning. *Nature Physics*, 4:S5–S20, 2008.

[17] John David Jackson. *Classical Electrodynamics.* John Wiley & Sons, 1962.

[18] D.M. Brink and G.R. Satchler. *Angular Momentum.* Claredon Press , Oxford, 1994.

[19] Melinda J. Duer. *Solid-State NMR spectroscopy.* Blackwell Publishing Ltd, 2004.

[20] S. A. Smith, T. O. Levante, B. H. Meier, and R. R. Ernst. Computer-simulations in magnetic-resonance - an object-oriented programming approach. *Journal of Magnetic Resonance Series A*, 106(1):75–105, Jan 1994.

[21] E. R. Andrew, A. Bradbury, and R. G. Eades. Nuclear magnetic resonance spectra from a crystal rotated at high speed. *Nature*, 182(4650):1659–1659, 1958.

[22] I. J. Lowe. Free induction decays of rotating solids. *Physical Review Letters*, 2(7):285–287, 1959.

[23] K. Seidel, A. Lange, S. Becker, C. E. Hughes, H. Heise, and M. Baldus. Protein solid-state nmr resonance assignments from (c-13, c-13) correlation spectroscopy. *Physical Chemistry Chemical Physics*, 6(22):5090–5093, 2004.

[24] A. E. Bennett, C. M. Rienstra, M. Auger, K. V. Lakshmi, and R. G. Griffin. Heteronuclear decoupling in rotating solids. *Journal of Chemical Physics*, 103(16):6951–6958, Oct 1995.

[25] B. M. Fung, A. K. Khitrin, and K. Ermolaev. An improved broadband decoupling sequence for liquid crystals and solids. *Journal of Magnetic Resonance*, 142(1):97–101, Jan 2000.

[26] F. Bloch, W. W. Hansen, and M. Packard. The nuclear induction experiment. *Physical Review*, 70(7-8):474–485, 1946.

[27] J. Jeener. Pulse pair techniques in high resolution nmr. In *Ampere International Summer School II*, 1971.

[28] S. Luca, D. V. Filippov, J. H. van Boom, H. Oschkinat, H. J. M. de Groot, and M. Baldus. Secondary chemical shifts in immobilized peptides and proteins: A qualitative basis for structure refinement under magic angle spinning. *Journal of Biomolecular NMR*, 20(4):325–331, Aug 2001.

[29] K. Seidel. *Structural characterization of membrane proteins by solid-state NMR spectroscopy*. PhD thesis, University of Goettingen, 2008.

[30] S. R. Hartmann and E. L. Hahn. Nuclear double resonance in rotating frame. *Physical Review*, 128(5):2042–&, 1962.

[31] A. Pines, M. G. Gibby, and J. S. Waugh. Proton-enhanced nmr of dilute spins in solids. *Journal of Chemical Physics*, 59(2):569–590, 1973.

[32] M. Baldus, A. T. Petkova, J. Herzfeld, and R. G. Griffin. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Molecular Physics*, 95(6):1197–1207, Dec 1998.

[33] M. Hohwy, C. M. Rienstra, C. P. Jaroniec, and R. G. Griffin. Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy. *Journal of Chemical Physics*, 110(16):7983–7992, Apr 1999.

Bibliography

[34] G. A. Morris and R. Freeman. Enhancement of nuclear magnetic-resonance signals by polarization transfer. *Journal of the American Chemical Society*, 101(3):760–762, 1979.

[35] Ovidiu C Andronesi, Stefan Becker, Karsten Seidel, Henrike Heise, Howard S Young, and Marc Baldus. Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state nmr spectroscopy. *Journal of the American Chemical Society*, 127(37):12965–12974, Sep 2005.

[36] M. Baldus and B. H. Meier. Total correlation spectroscopy in the solid state. the use of scalar couplings to determine the through-bond connectivity. *Journal of Magnetic Resonance Series A*, 121(1):65–69, Jul 1996.

[37] Anja Böckmann, Adam Lange, Anne Galinier, Sorin Luca, Nicolas Giraud, Michel Juy, Henrike Heise, Roland Montserret, François Penin, and Marc Baldus. Solid state nmr sequential resonance assignments and conformational analysis of the 2x10.4 kda dimeric form of the bacillus subtilis protein crh. *Journal of Biomolecular NMR*, 27(4):323–339, Dec 2003.

[38] Federica Castellani, Barth-Jan van Rossum, Annette Diehl, Kristina Rehbein, and Hartmut Oschkinat. Determination of solid-state nmr structures of proteins by means of three-dimensional 15n-13c-13c dipolar correlation spectroscopy and chemical shift analysis. *Biochemistry*, 42(39):11476–11483, Oct 2003.

[39] Toshimichi Fujiwara, Yasuto Todokoro, Hajime Yanagishita, Midori Tawarayama, Toshiyuki Kohno, Kaori Wakamatsu, and Hideo Akutsu. Signal assignments and chemical-shift structural analysis of uniformly 13c, 15n-labeled peptide, mastoparan-x, by multidimensional solid-state nmr under magic-angle spinning. *Journal of Biomolecular NMR*, 28(4):311–325, Apr 2004.

[40] A. J. van Gammeren, F. B. Hulsbergen, J. G. Hollander, and H. J. M. de Groot. Residual backbone and side-chain c-13 and n-15 resonance assignments of the intrinsic transmembrane light-harvesting 2 protein complex by solid-state magic angle spinning nmr spectroscopy. *Journal of Biomolecular NMR*, 31(4):279–293, Apr 2005.

[41] D. Marulanda, M. L. Tasayco, M. Cataldi, V. Arriaran, and T. Polenova. Reso-

nance assignments and secondary structure analysis of e. coli thioredoxin by magic angle spinning solid-state nmr spectroscopy. *Journal of Physical Chemistry B*, 109(38):18135–18145, September 29, 2005 2005.

[42] Karsten Seidel, Manuel Etzkorn, Henrike Heise, Stefan Becker, and Marc Baldus. High-resolution solid-state nmr studies on uniformly [13c,15n]-labeled ubiquitin. *ChemBioChem*, 6(9):1638–1647, Sep 2005.

[43] Adam Lange, Stefan Becker, Karsten Seidel, Karin Giller, Olaf Pongs, and Marc Baldus. A concept for rapid protein-structure determination by solid-state nmr spectroscopy. *Angewandte Chemie International Edition English*, 44(14):2089–2092, Mar 2005.

[44] W. T. Franks, D. H. Zhou, B. J. Wylie, B. G. Money, D. T. Graesser, H. L. Frericks, G. Sahota, and C. M. Rienstra. Magic-angle spinning solid-state nmr spectroscopy of the beta 1 immunoglobulin binding domain of protein g (gb1): N-15 and c-13 chemical shift assignments and conformational analysis. *Journal of the American Chemical Society*, 127(35):12291–12305, Sep 2005.

[45] Manuel Etzkorn, Swetlana Martell, Ovidiu C Andronesi, Karsten Seidel, Martin Engelhard, and Marc Baldus. Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state nmr spectroscopy. *Angewandte Chemie International Edition English*, 46(3):459–462, 2007.

[46] M. Baldus. Correlation experiments for assignment and structure elucidation of immobilized polypeptides under magic angle spinning. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 41(1-2):1–47, Sep 2002.

[47] D. S. Wishart, B. D. Sykes, and F. M. Richards. Relationship between nuclear-magnetic-resonance chemical-shift and protein secondary structure. *Journal of Molecular Biology*, 222(2):311–333, Nov 1991.

[48] Eldon L Ulrich, Hideo Akutsu, Jurgen F Doreleijers, Yoko Harano, Yannis E Ioannidis, Jundong Lin, Miron Livny, Steve Mading, Dimitri Maziuk, Zachary Miller, Eiichi Nakatani, Christopher F Schulte, David E Tolmie, R. Kent Wenger, Hongyang Yao, and John L Markley. Biomagresbank. *Nucleic Acids Research*, 36(Database issue):D402–D408, Jan 2008.

Bibliography

[49] Adam Lange, Sorin Luca, and Marc Baldus. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *Journal of the American Chemical Society*, 124(33):9704–9705, Aug 2002.

[50] Adam Lange, Karsten Seidel, Laurent Verdier, Sorin Luca, and Marc Baldus. Analysis of proton-proton transfer dynamics in rotating solids and their use for 3d structure determination. *Journal of the American Chemical Society*, 125(41):12640–12648, Oct 2003.

[51] R. Schneider, K. Seidel, M. Etzkorn, S. Becker, and M. Baldus. Molecular motion detected by double-quantum (13c,13c) solid-state nmr spectroscopy. *mansucript in preparation.*

[52] H. T. Edzes and E. T. Samulski. Cross relaxation and spin diffusion in proton nmr of hydrated collagen. *Nature*, 265(5594):521–523, 1977.

[53] K. K. Kumashiro, K. Schmidt-Rohr, O. J. Murphy, K. L. Ouellette, W. A. Cramer, and L. K. Thompson. A novel tool for probing membrane protein structure: Solid-state nmr with proton spin diffusion and x-nucleus detection. *Journal of the American Chemical Society*, 120(20):5043–5051, May 1998.

[54] H. Heise, M.S. Celej, S. Becker, D. Riedel, A. Pelah, A. Kumar, T.M. Jovin, and M. Baldus. Solid-state nmr reveals structural differences between wild-type and disease-related a53t mutant fibrils of a-synuclein. *submitted*, 2008.

[55] G. W. Vuister, S. J. Kim, C. Wu, and A. Bax. 2d and 3d nmr-study of phenylalanine residues in proteins by reverse isotopic labeling. *Journal of the American Chemical Society*, 116(20):9206–9210, Oct 1994.

[56] Henrike Heise, Wolfgang Hoyer, Stefan Becker, Ovidiu C Andronesi, Dietmar Riedel, and Marc Baldus. Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state nmr. *Proceedings of the National Academy of Sciences of the United States of America*, 102(44):15871–15876, Nov 2005.

[57] C. B. Anfinsen, E. Haber, M. Sela, and F. H. White. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proceedings*

162

of the National Academy of Sciences of the United States of America, 47:1309–1314, Sep 1961.

[58] C. B. Anfinsen. Principles that govern the folding of protein chains. *Science*, 181(96):223–230, Jul 1973.

[59] Ora Schueler-Furman, Chu Wang, Phil Bradley, Kira Misura, and David Baker. Progress in modeling of protein structures and interactions. *Science*, 310(5748):638–642, Oct 2005.

[60] Michael Tress, Iakes Ezkurdia, Osvaldo Graña, Gonzalo López, and Alfonso Valencia. Assessment of predictions submitted for the casp6 comparative modeling category. *Proteins*, 61 Suppl 7:27–45, 2005.

[61] Kira M S Misura, Dylan Chivian, Carol A Rohl, David E Kim, and David Baker. Physically realistic homology models built with rosetta can be more accurate than their templates. *Proceedings of the National Academy of Sciences of the United States of America*, 103(14):5361–5366, Apr 2006.

[62] Philip Bradley, Kira M S Misura, and David Baker. Toward high-resolution de novo structure prediction for small proteins. *Science*, 309(5742):1868–1871, Sep 2005.

[63] Jens Meiler and David Baker. The fumarate sensor dcus: progress in rapid protein fold elucidation by combining protein structure prediction methods with nmr spectroscopy. *Journal of Magnetic Resonance*, 173(2):310–316, Apr 2005.

[64] Jens Meiler and David Baker. Rapid protein fold determination using unassigned nmr data. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26):15404–15409, Dec 2003.

[65] Bin Qian, Srivatsan Raman, Rhiju Das, Philip Bradley, Airlie J McCoy, Randy J Read, and David Baker. High-resolution structure prediction and the crystallographic phase problem. *Nature*, 450(7167):259–264, Nov 2007.

[66] H. Zhang. Protein tertiary structure: Prediction from amino acid sequence. *Encyclopedia of life sciences*, 2002.

[67] Dylan Chivian, David E Kim, Lars Malmström, Philip Bradley, Timothy Robert-

son, Paul Murphy, Charles E M Strauss, Richard Bonneau, Carol A Rohl, and David Baker. Automated prediction of casp-5 structures using the robetta server. *Proteins*, 53 Suppl 6:524–533, 2003.

[68] Dylan Chivian, David E Kim, Lars Malmström, Jack Schonbrun, Carol A Rohl, and David Baker. Prediction of casp6 structures using automated robetta protocols. *Proteins*, 61 Suppl 7:157–166, 2005.

[69] David E Kim, Dylan Chivian, and David Baker. Protein structure prediction and analysis using the robetta server. *Nucleic Acids Research*, 32(Web Server issue):W526–W531, Jul 2004.

[70] Sitao Wu, Jeffrey Skolnick, and Yang Zhang. Ab initio modeling of small proteins by iterative tasser simulations. *BMC Biol*, 5:17, 2007.

[71] Yang Zhang. Template-based modeling and free modeling by i-tasser in casp7. *Proteins*, 69 Suppl 8:108–117, 2007.

[72] Yang Zhang. I-tasser server for protein 3d structure prediction. *BMC Bioinformatics*, 9:40, 2008.

[73] Dylan Chivian and David Baker. Homology modeling using parametric alignment ensemble generation with consensus and energy-based model selection. *Nucleic Acids Research*, 34(17):e112, 2006.

[74] Carol A Rohl, Charlie E M Strauss, Dylan Chivian, and David Baker. Modeling structurally variable regions in homologous proteins with rosetta. *Proteins*, 55(3):656–677, May 2004.

[75] David E Kim, Dylan Chivian, Lars Malmström, and David Baker. Automated prediction of domain boundaries in casp6 targets using ginzu and rosettadom. *Proteins*, 61 Suppl 7:193–200, 2005.

[76] Carol A Rohl, Charlie E M Strauss, Kira M S Misura, and David Baker. Protein structure prediction using rosetta. *Methods Enzymol*, 383:66–93, 2004.

[77] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. Gapped blast and psi-blast: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17):3389–3402, Sep 1997.

[78] A. A. Schäffer, L. Aravind, T. L. Madden, S. Shavirin, J. L. Spouge, Y. I. Wolf, E. V. Koonin, and S. F. Altschul. Improving the accuracy of psi-blast protein database searches with composition-based statistics and other refinements. *Nucleic Acids Research*, 29(14):2994–3005, Jul 2001.

[79] Dariusz Plewczynski, Jakub Pas, Marcin von Grotthuss, and Leszek Rychlewski. 3d-hit: fast structural comparison of proteins. *Appl Bioinformatics*, 1(4):223–225, 2002.

[80] W. R. Taylor. Protein structural domain identification. *Protein Eng*, 12(3):203–216, Mar 1999.

[81] D. T. Jones. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, 292(2):195–202, Sep 1999.

[82] W. Kabsch and C. Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12):2577–2637, Dec 1983.

[83] K. Karplus, C. Barrett, and R. Hughey. Hidden markov models for detecting remote protein homologies. *Bioinformatics*, 14(10):846–856, 1998.

[84] J. Park, K. Karplus, C. Barrett, R. Hughey, D. Haussler, T. Hubbard, and C. Chothia. Sequence comparisons using multiple sequences detect three times as many remote homologues as pairwise methods. *Journal of Molecular Biology*, 284(4):1201–1210, Dec 1998.

[85] J. Meiler, M. Mueller, A. Zeidler, and F. Schmaeschke. Jufo: Secondary structure prediction for proteins. *http://www.jens-meiler.de/*, 2002.

[86] J. Deutscher, A. Galinier, and I. Martin-Verstraete. *Bacillus subtilis and its closest relatives: From Genes to Cells*, pages 129–150. ASM Press: Washington, 2001.

[87] Robert Tycko. Progress towards a molecular-level structural understanding of amyloid fibrils. *Current Opinion in Structural Biology*, 14(1):96–103, Feb 2004.

[88] Christiane Ritter, Marie-Lise Maddelein, Ansgar B Siemer, Thorsten Lührs, Matthias Ernst, Beat H Meier, Sven J Saupe, and Roland Riek. Correlation of structural elements and infectivity of the het-s prion. *Nature*, 435(7043):844–848,

Jun 2005.

[89] Ansgar B Siemer, Christiane Ritter, Matthias Ernst, Roland Riek, and Beat H Meier. High-resolution solid-state nmr spectroscopy of the prion protein het-s in its amyloid conformation. *Angew Chem Int Ed Engl*, 44(16):2441–2444, Apr 2005.

[90] Aneta T Petkova, Yoshitaka Ishii, John J Balbach, Oleg N Antzutkin, Richard D Leapman, Frank Delaglio, and Robert Tycko. A structural model for alzheimer's beta-amyloid fibrils based on experimental constraints from solid state nmr. *Proceedings of the National Academy of Sciences of the United States of America*, 99(26):16742–16747, Dec 2002.

[91] R. H. Havlin and R. Tycko. Probing site-specific conformational distributions in protein folding with solid-state nmr. *Proceedings of the National Academy of Sciences of the United States of America*, 102(9):3284–3289, Mar 2005.

[92] Sandra Chimon and Yoshitaka Ishii. Capturing intermediate structures of alzheimer's beta-amyloid, abeta(1-40), by solid-state nmr spectroscopy. *Journal of the American Chemical Society*, 127(39):13472–13473, Oct 2005.

[93] Henrike Heise, Sorin Luca, Bert L de Groot, Helmut Grubmüller, and Marc Baldus. Probing conformational disorder in neurotensin by two-dimensional solid-state nmr and comparison to molecular dynamics simulations. *Biophysical Journal*, 89(3):2113–2120, Sep 2005.

[94] J. J. Balbach, Y. Ishii, O. N. Antzutkin, R. D. Leapman, N. W. Rizzo, F. Dyda, J. Reed, and R. Tycko. Amyloid fibril formation by a beta(16-22), a seven-residue fragment of the alzheimer's beta-amyloid peptide, and structural characterization by solid state nmr. *Biochemistry*, 39(45):13748–13759, Nov 2000.

[95] F. Penin, A. Favier, R. Montserret, B. Brutscher, J. Deutscher, D. Marion, and D. Galinier. Evidence for a dimerisation state of the bacillus subtilis catabolite repression hpr-like protein, crh. *Journal of Molecular Microbiology and Biotechnology*, 3(3):429–432, Jul 2001.

[96] Adrien Favier, Bernhard Brutscher, Martin Blackledge, Anne Galinier, Josef Deutscher, François Penin, and Dominique Marion. Solution structure and dy-

namics of crh, the bacillus subtilis catabolite repression hpr. *Journal of Molecular Biology*, 317(1):131–144, Mar 2002.

[97] M. Juy, F. Penin, A. Favier, A. Galinier, R. Montserret, R. Haser, J. Deutscher, and A. Bockmann. Dimerization of crh by reversible 3d domain swapping induces structural adjustments to its monomeric homologue hpr. *Journal of Molecular Biology*, 332(4):767–776, Sep 2003.

[98] Frederic Rousseau, Joost W H Schymkowitz, and Laura S Itzhaki. The unfolding story of three-dimensional domain swapping. *Structure*, 11(3):243–251, Mar 2003.

[99] K. J. Knaus, M. Morillas, W. Swietnicki, M. Malone, W. K. Surewicz, and V. C. Yee. Crystal structure of the human prion protein reveals a mechanism for oligomerization. *Nat Struct Biol*, 8(9):770–774, Sep 2001.

[100] John M Louis, In-Ja L Byeon, Ulrich Baxa, and Angela M Gronenborn. The gb1 amyloid fibril: recruitment of the peripheral beta-strands of the domain swapped dimer into the polymeric interface. *Journal of Molecular Biology*, 348(3):687–698, May 2005.

[101] M. J. Bennett, M. P. Schlunegger, and D. Eisenberg. 3d domain swapping: a mechanism for oligomer assembly. *Protein Science*, 4(12):2455–2468, Dec 1995.

[102] Shilpa Sambashivan, Yanshun Liu, Michael R Sawaya, Mari Gingery, and David Eisenberg. Amyloid-like fibrils of ribonuclease a with three-dimensional domain-swapped and native-like structure. *Nature*, 437(7056):266–269, Sep 2005.

[103] M. J. Bennett, S. Choe, and D. Eisenberg. Refined structure of dimeric diphtheria toxin at 2.0 a resolution. *Protein Science*, 3(9):1444–1463, Sep 1994.

[104] Marcia E Newcomer. Protein folding and three-dimensional domain swapping: a strained relationship? *Current Opinion in Structural Biology*, 12(1):48–53, Feb 2002.

[105] R. Janowski, M. Kozak, E. Jankowska, Z. Grzonka, A. Grubb, M. Abrahamson, and M. Jaskolski. Human cystatin c, an amyloidogenic protein, dimerizes through three-dimensional domain swapping. *Nat Struct Biol*, 8(4):316–320, Apr 2001.

[106] Sichun Yang, Samuel S Cho, Yaakov Levy, Margaret S Cheung, Herbert Levine,

Bibliography

Peter G Wolynes, and José N Onuchic. Domain swapping is a consequence of minimal frustration. *Proceedings of the National Academy of Sciences of the United States of America*, 101(38):13786–13791, Sep 2004.

[107] Luciana Esposito and Valerie Daggett. Insight into ribonuclease a domain swapping by molecular dynamics unfolding simulations. *Biochemistry*, 44(9):3358–3368, Mar 2005.

[108] Jorge Chahine and Margaret S Cheung. Computational studies of the reversible domain swapping of p13suc1. *Biophysical Journal*, 89(4):2693–2700, Oct 2005.

[109] In-Ja L Byeon, John M Louis, and Angela M Gronenborn. A captured folding intermediate involved in dimerization and domain-swapping of gb1. *Journal of Molecular Biology*, 340(3):615–625, Jul 2004.

[110] F. Rousseau, J. W. Schymkowitz, H. R. Wilkinson, and L. S. Itzhaki. Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10):5596–5601, May 2001.

[111] Manuel Etzkorn, Anja Böckmann, Adam Lange, and Marc Baldus. Probing molecular interfaces using 2d magic-angle-spinning nmr on protein mixtures with different uniform labeling. *Journal of the American Chemical Society*, 126(45):14746–14751, Nov 2004.

[112] Sudharsan Sridharan, Abbas Razvi, J. Martin Scholtz, and James C Sacchettini. The hpr proteins from the thermophile bacillus stearothermophilus can form domain-swapped dimers. *Journal of Molecular Biology*, 346(3):919–931, Feb 2005.

[113] Jason P Schmittschmitt and J. Martin Scholtz. The side chain of aspartic acid 69 dictates the folding mechanism of bacillus subtilis hpr. *Biochemistry*, 43(5):1360–1368, Feb 2004.

[114] Yanshun Liu and David Eisenberg. 3d domain swapping: as domains continue to swap. *Protein Science*, 11(6):1285–1299, Jun 2002.

[115] Rakez Kayed, Elizabeth Head, Jennifer L Thompson, Theresa M McIntire, Saskia C Milton, Carl W Cotman, and Charles G Glabe. Common structure

of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*, 300(5618):486–489, Apr 2003.

[116] M. Kamihira, A. Naito, S. Tuzi, A. Y. Nosaka, and H. Saitô. Conformational transitions and fibrillation mechanism of human calcitonin as studied by high-resolution solid-state 13c nmr. *Protein Science*, 9(5):867–877, May 2000.

[117] Frederic Rousseau, Joost W H Schymkowitz, Hannah R Wilkinson, and Laura S Itzhaki. Intermediates control domain swapping during folding of p13suc1. *Journal of Biological Chemistry*, 279(9):8368–8377, Feb 2004.

[118] Zhefeng Guo and David Eisenberg. Runaway domain swapping in amyloid-like fibrils of t7 endonuclease i. *Proceedings of the National Academy of Sciences of the United States of America*, 103(21):8042–8047, May 2006.

[119] Rebecca Nelson and David Eisenberg. Recent atomic models of amyloid fibril structure. *Current Opinion in Structural Biology*, 16(2):260–265, Apr 2006.

[120] Georgia Plakoutsi, Francesco Bemporad, Martino Calamai, Niccolò Taddei, Christopher M Dobson, and Fabrizio Chiti. Evidence for a mechanism of amyloid formation involving molecular reorganisation within native-like precursor aggregates. *Journal of Molecular Biology*, 351(4):910–922, Aug 2005.

[121] Thomas R Jahn, Martin J Parker, Steve W Homans, and Sheena E Radford. Amyloid formation under physiological conditions proceeds via a native-like folding intermediate. *Nature Structural & Molecular Biology*, 13(3):195–201, Mar 2006.

[122] Sorin Luca, Henrike Heise, and Marc Baldus. High-resolution solid-state nmr applied to polypeptides and membrane proteins. *Accounts of Chemical Research*, 36(11):858–865, Nov 2003.

[123] N. Giraud, A. Bockmann, A. Lesage, F. Penin, M. Blackledge, and L. Emsley. Site-specific backbone dynamics from a crystalline protein by solid-state nmr spectroscopy. *Journal of the American Chemical Society*, 126(37):11422–11423, Sep 2004.

[124] Ann E McDermott. Structural and dynamic studies of proteins by solid-state nmr spectroscopy: rapid movement forward. *Current Opinion in Structural Biology*,

Bibliography

14(5):554–561, Oct 2004.

[125] Ronald Wetzel. Kinetics and thermodynamics of amyloid fibril assembly. *Accounts of Chemical Research*, 39(9):671–679, Sep 2006.

[126] Amy R Hurshman, Joleen T White, Evan T Powers, and Jeffery W Kelly. Transthyretin aggregation under partially denaturing conditions is a downhill polymerization. *Biochemistry*, 43(23):7365–7381, Jun 2004.

[127] W. L. Klein, G. A. Krafft, and C. E. Finch. Targeting small abeta oligomers: the solution to an alzheimer's disease conundrum? *Trends in Neurosciences*, 24(4):219–224, Apr 2001.

[128] J. Balbach, V. Forge, W. S. Lau, N. A. van Nuland, K. Brew, and C. M. Dobson. Protein folding monitored at individual residues during a two-dimensional nmr experiment. *Science*, 274(5290):1161–1163, Nov 1996.

[129] Markus Zeeb and Jochen Balbach. Protein folding studied by real-time nmr spectroscopy. *Methods*, 34(1):65–74, Sep 2004.

[130] M. Helgstrand, T. Härd, and P. Allard. Simulations of nmr pulse sequences during equilibrium and non-equilibrium chemical exchange. *Journal of Biomolecular NMR*, 18(1):49–63, Sep 2000.

[131] J. Balbach, C. Steegborn, T. Schindler, and F. X. Schmid. A protein folding intermediate of ribonuclease t1 characterized at high resolution by 1d and 2d real-time nmr spectroscopy. *Journal of Molecular Biology*, 285(2):829–842, Jan 1999.

[132] F. Ferrone. Analysis of protein aggregation kinetics. *Methods Enzymol*, 309:256–274, 1999.

[133] Aimee M Morris, Murielle A Watzky, Jeffrey N Agar, and Richard G Finke. Fitting neurological protein aggregation kinetic data via a 2-step, minimal/"ockham's razor" model: the finke-watzky mechanism of nucleation followed by autocatalytic surface growth. *Biochemistry*, 47(8):2413–2427, Feb 2008.

[134] A. Wegner and J. Engel. Kinetics of the cooperative association of actin to actin filaments. *Biophysical Chemistry*, 3(3):215–225, Jul 1975.

[135] C. Frieden and D. W. Goddette. Polymerization of actin and actin-like systems:

170

evaluation of the time course of polymerization in relation to the mechanism. *Biochemistry*, 22(25):5836–5843, Dec 1983.

[136] D. Thusius, P. Dessen, and J. M. Jallon. Mechanisms of bovine liver glutamate dehydrogenase self-association. i. kinetic evidence for a random association of polymer chains. *Journal of Molecular Biology*, 92(3):413–432, Mar 1975.

[137] M. Eigen. Prionics or the kinetic basis of prion diseases. *Biophysical Chemistry*, 63(1):A1–18, Dec 1996.

[138] M. A. Watzky and R. G. Finke. Transition metal nanocluster formation kinetic and mechanistic studies. a new mechanism when hydrogen is the reductant: Slow, continuous nucleation and fast autocatalytic surface growth. *Journal Of The American Chemical Society*, 119(43):10382–10400, October 1997.

[139] J. Sabelko, J. Ervin, and M. Gruebele. Observation of strange kinetics in protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, 96(11):6031–6036, May 1999.

[140] Maria M Garcia-Mira, Mourad Sadqi, Niels Fischer, Jose M Sanchez-Ruiz, and Victor Muñoz. Experimental identification of downhill protein folding. *Science*, 298(5601):2191–2195, Dec 2002.

[141] Hironori K Nakamura, Masaki Sasai, and Mitsunori Takano. Squeezed exponential kinetics to describe a nonglassy downhill folding as observed in a lattice protein model. *Proteins*, 55(1):99–106, Apr 2004.

[142] Szabolcs Osváth, Jobiah J Sabelko, and Martin Gruebele. Tuning the heterogeneous early folding dynamics of phosphoglycerate kinase. *Journal of Molecular Biology*, 333(1):187–199, Oct 2003.

[143] Hairong Ma and Martin Gruebele. Kinetics are probe-dependent during downhill folding of an engineered lambda6-85 protein. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7):2283–2287, Feb 2005.

[144] J. P. Klare, V. I. Gordeliy, J. Labahn, G. Buldt, H. J. Steinhoff, and M. Engelhard. The archaeal sensory rhodopsin ii/transducer complex: a model for transmembrane signal transfer. *FEBS Letters*, 564(3):219–224, Apr 2004.

Bibliography

[145] K. Edman, A. Royant, P. Nollert, C. A. Maxwell, E. Pebay-Peyroula, J. Navarro, R. Neutze, and E. M. Landau. Early structural rearrangements in the photocycle of an integral membrane sensory receptor. *Structure*, 10(4):473–482, Apr 2002.

[146] H. Luecke, B. Schobert, J. K. Lanyi, E. N. Spudich, and J. L. Spudich. Crystal structure of sensory rhodopsin ii at 2.4 angstroms: Insights into color tuning and transducer interaction. *Science*, 293(5534):1499–1503, Aug 2001.

[147] A. Royant, P. Nollert, K. Edman, R. Neutze, E. M. Landau, E. Pebay-Peyroula, and J. Navarro. X-ray structure of sensory rhodopsin ii at 2.1-angstrom resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18):10131–10136, Aug 2001.

[148] V. I. Gordeliy, J. Labahn, R. Moukhametzianov, R. Efremov, J. Granzin, R. Schlesinger, G. Buldt, T. Savopol, A. J. Scheidig, J. P. Klare, and M. Engelhard. Molecular basis of transmembrane signalling by sensory rhodopsin ii-transducer complex. *Nature*, 419(6906):484–487, Oct 2002.

[149] R. Moukhametzianov, J. P. Klare, R. Efremov, C. Baeken, A. Goppner, J. Labahn, M. Engelhard, G. Buldt, and V. I. Gordeliy. Development of the signal in sensory rhodopsin and its transfer to the cognate transducer. *Nature*, 440(7080):115–119, Mar 2006.

[150] Johann Klare, Igor Chizhov, and Martin Engelhard. Microbial rhodopsins: Scaffolds for ion pumps, channels, and sensors. *Results Probl Cell Differ*, Sep 2007.

[151] P. A. Karplus and G. E. Schulz. Prediction of chain flexibility in proteins - a tool for the selection of peptide antigens. *Naturwissenschaften*, 72(4):212–213, 1985.

[152] A. Schlessinger and B. Rost. Protein flexibility and rigidity predicted from sequence. *Proteins-Structure Function And Bioinformatics*, 61(1):115–126, October 2005.

[153] Leo C James, Pietro Roversi, and Dan S Tawfik. Antibody multispecificity mediated by conformational diversity. *Science*, 299(5611):1362–1367, Feb 2003.

[154] Dror Tobi and Ivet Bahar. Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proceedings of the National*

Academy of Sciences of the United States of America, 102(52):18908–18913, Dec 2005.

[155] A. J. Wand. Dynamic activation of protein function: a view emerging from nmr spectroscopy. *Nat Struct Biol*, 8(11):926–931, Nov 2001.

[156] Peter M Hwang, Russell E Bishop, and Lewis E Kay. The integral membrane enzyme pagp alternates between two dynamically distinct states. *Proceedings of the National Academy of Sciences of the United States of America*, 101(26):9618–9623, Jun 2004.

[157] Elan Z Eisenmesser, Oscar Millet, Wladimir Labeikovsky, Dmitry M Korzhnev, Magnus Wolf-Watz, Daryl A Bosco, Jack J Skalicky, Lewis E Kay, and Dorothee Kern. Intrinsic dynamics of an enzyme underlies catalysis. *Nature*, 438(7064):117–121, Nov 2005.

[158] B. Y. Ma, S. Kumar, C. J. Tsai, and R. Nussinov. Folding funnels and binding mechanisms. *Protein Engineering*, 12(9):713–720, September 1999.

[159] F. Creuzet, A. McDermott, R. Gebhard, K. Vanderhoef, M. B. Spijkerassink, J. Herzfeld, J. Lugtenburg, M. H. Levitt, and R. G. Griffin. Determination of membrane-protein structure by rotational resonance nmr - bacteriorhodopsin. *Science*, 251(4995):783–786, Feb 1991.

[160] A. F. L. Creemers, S. Kiihne, P. H. M. Bovee-Geurts, W. J. DeGrip, J. Lugtenburg, and H. J. M. de Groot. H-1 and c-13 mas nmr evidence for pronounced ligand-protein interactions involving the ionone ring of the retinylidene chromophore in rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America*, 99(14):9101–9106, Jul 2002.

[161] Sorin Luca, Jim F White, Awinder K Sohal, Dmitri V Filippov, Jacques H van Boom, Reinhard Grisshammer, and Marc Baldus. The conformation of neurotensin bound to its g protein-coupled receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 100(19):10706–10711, Sep 2003.

[162] A. B. Patel, E. Crocker, M. Eilers, A. Hirshfeld, M. Sheves, and S. O. Smith. Coupling of retinal isomerization to the activation of rhodopsin. *Proceedings of*

Bibliography

the National Academy of Sciences of the United States of America, 101(27):10048–10053, Jul 2004.

[163] L. Krabben, B. J. van Rossum, F. Castellani, E. Bocharov, A. A. Schulga, A. S. Arseniev, C. Weise, F. Hucho, and H. Oschkinata. Towards structure determination of neurotoxin ii bound to nicotinic acetylcholine receptor: a solid-state nmr approach. *FEBS Letters*, 564(3):319–324, Apr 2004.

[164] T. A. Egorova-Zachernyuk, J. Hollander, N. Fraser, P. Gast, A. J. Hoff, R. Cogdell, H. J. de Groot, and M. Baldus. Heteronuclear 2d-correlations in a uniformly [13c, 15n] labeled membrane-protein complex at ultra-high magnetic fields. *Journal of Biomolecular NMR*, 19(3):243–253, Mar 2001.

[165] A. J. van Gammeren, F. B. Hulsbergen, J. G. Hollander, and H. J. M. de Groot. Biosynthetic site-specific c-13 labeling of the light-harvesting 2 protein complex: A model for solid state nmr structure determination of transmembrane proteins. *Journal of Biomolecular NMR*, 30(3):267–274, Nov 2004.

[166] M. Hiller, L. Krabben, K. R. Vinothkumar, F. Castellani, B. J. van Rossum, W. Kuhlbrandt, and H. Oschkinat. Solid-state magic-angle spinning nmr of outer-membrane protein g from escherichia coli. *ChemBioChem*, 6(9):1679–1684, Sep 2005.

[167] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. M. Ono, and P. Guntert. Optimal isotope labelling for nmr protein structure determinations. *Nature*, 440(7080):52–57, Mar 2006.

[168] N. Bloembergen. On the interaction of nuclear spins in a crystalline lattice. *Physica*, 15(3-4):386–426, 1949.

[169] Y. Sudo, H. Okuda, M. Yamabi, Y. Fukuzaki, M. Mishima, N. Kamo, and C. Kojima. Linker region of a halobacterial transducer protein interacts directly with its sensor retinal protein. *Biochemistry*, 44(16):6144–6152, Apr 2005.

[170] J. P. Klare, E. Bordignon, M. Doebber, J. Fitter, J. Kriegsmann, I. Chizhov, H. J. Steinhoff, and M. Engelhard. Effects of solubilization on the structure and function of the sensory rhodopsin ii/transducer complex. *Journal of Molecular*

Biology, 356(5):1207–1221, Mar 2006.

[171] A. Matsuno-Yagi and Y. Mukohata. Two possible roles of bacteriorhodopsin; a comparative study of strains of halobacterium halobium differing in pigmentation. *Biochemical and Biophysical Research Communications*, 78(1):237–243, Sep 1977.

[172] A. Matsuno-Yagi and Y. Mukohata. Atp synthesis linked to light-dependent proton uptake in a rad mutant strain of halobacterium lacking bacteriorhodopsin. *Archives of Biochemistry and Biophysics*, 199(1):297–303, Jan 1980.

[173] O. Béjà, L. Aravind, E. V. Koonin, M. T. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich, and E. F. DeLong. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science*, 289(5486):1902–1906, Sep 2000.

[174] N. A. Dencher and E. Hildebrand. Sensory transduction in halobacterium halobium: retinal protein pigment controls uv-induced behavioral response. *Z Naturforsch [C]*, 34(9-10):841–847, 1979.

[175] E. Hildebrand and N. Dencher. Two photosystems controlling behavioural responses of halobacterium halobium. *Nature*, 257(5521):46–48, Sep 1975.

[176] J. L. Spudich and W. Stoeckenius. Light-regulated retinal-dependent reversible phosphorylation of halobacterium proteins. *Journal of Biological Chemistry*, 255(12):5501–5503, Jun 1980.

[177] W. Sperling and A. Schimz. Photosensory retinal pigments in halobacterium halobium. *Biophys Struct Mech*, 6(2):165–169, 1980.

[178] Georg Nagel, Tanjef Szellas, Wolfram Huhn, Suneel Kateriya, Nona Adeishvili, Peter Berthold, Doris Ollig, Peter Hegemann, and Ernst Bamberg. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proceedings of the National Academy of Sciences of the United States of America*, 100(24):13940–13945, Nov 2003.

[179] Georg Nagel, Doris Ollig, Markus Fuhrmann, Suneel Kateriya, Anna Maria Musti, Ernst Bamberg, and Peter Hegemann. Channelrhodopsin-1: a light-gated proton channel in green algae. *Science*, 296(5577):2395–2398, Jun 2002.

Bibliography

[180] R. A. Bogomolni, W. Stoeckenius, I. Szundi, E. Perozo, K. D. Olson, and J. L. Spudich. Removal of transducer htri allows electrogenic proton translocation by sensory rhodopsin i. *Proceedings of the National Academy of Sciences of the United States of America*, 91(21):10188–10192, Oct 1994.

[181] G. Schmies, M. Engelhard, P. G. Wood, G. Nagel, and E. Bamberg. Electrophysiological characterization of specific interactions between bacterial sensory rhodopsins and their transducers. *Proceedings of the National Academy of Sciences of the United States of America*, 98(4):1555–1559, Feb 2001.

[182] Y. Sudo, M. Iwamoto, K. Shimono, M. Sumi, and N. Kamo. Photo-induced proton transport of pharaonis phoborhodopsin (sensory rhodopsin ii) is ceased by association with the transducer. *Biophysical Journal*, 80(2):916–922, Feb 2001.

[183] Yuki Sudo and John L Spudich. Three strategically placed hydrogen-bonding residues convert a proton pump into a sensory receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 103(44):16129–16134, Oct 2006.

[184] Enrica Bordignon, Johann P Klare, Julia Holterhues, Swetlana Martell, Aliaksei Krasnaberski, Martin Engelhard, and Heinz-Jürgen Steinhoff. Analysis of light-induced conformational changes of natronomonas pharaonis sensory rhodopsin ii by time resolved electron paramagnetic resonance spectroscopy. *Photochemistry and Photobiology*, 83(2):263–272, 2007.

[185] Enrica Bordignon, Johann P Klare, Meike Doebber, Ansgar A Wegener, Swetlana Martell, Martin Engelhard, and Heinz-Jürgen Steinhoff. Structural analysis of a hamp domain: the linker region of the phototransducer in complex with sensory rhodopsin ii. *Journal of Biological Chemistry*, 280(46):38767–38775, Nov 2005.

[186] Chii-Shen Yang, Oleg Sineshchekov, Elena N Spudich, and John L Spudich. The cytoplasmic membrane-proximal domain of the htrii transducer interacts with the e-f loop of photoactivated natronomonas pharaonis sensory rhodopsin ii. *Journal of Biological Chemistry*, 279(41):42970–42976, Oct 2004.

[187] Y. Chen, J. Reizer, M. H. Saier, W. J. Fairbrother, and P. E. Wright. Mapping of the binding interfaces of the proteins of the bacterial phosphotransferase system,

hpr and iiaglc. *Biochemistry*, 32(1):32–37, Jan 1993.

[188] N. A. van Nuland, G. J. Kroon, K. Dijkstra, G. K. Wolters, R. M. Scheek, and G. T. Robillard. The nmr determination of the iia(mtl) binding site on hpr of the escherichia coli phosphoenol pyruvate-dependent phosphotransferase system. *FEBS Letters*, 315(1):11–15, Jan 1993.

[189] Erik R P Zuiderweg. Mapping protein-protein interactions in solution by nmr spectroscopy. *Biochemistry*, 41(1):1–7, Jan 2002.

[190] Marc Baldus. Molecular interactions investigated by multi-dimensional solid-state nmr. *Current Opinion in Structural Biology*, 16(5):618–623, Oct 2006.

[191] Karsten Seidel, Ovidiu Andronesi, Joachim Krebs, Christian Griesinger, Howard Young, Stefan Becker, and Marc Baldus. Structural characterization of ca(2+)-atpase-bound phospholamban in lipid bilayers by solid-state nuclear magnetic resonance (nmr) spectroscopy(,). *Biochemistry*, Mar 2008.

[192] Jakob J Lopez, Arun K Shukla, Christoph Reinhart, Harald Schwalbe, Hartmut Michel, and Clemens Glaubitz. The structure of the neuropeptide bradykinin bound to the human g-protein coupled receptor bradykinin b2 as determined by solid-state nmr spectroscopy. *Angew Chem Int Ed Engl*, 47(9):1668–1671, 2008.

[193] Keiichi Inoue, Jun Sasaki, John L Spudich, and Masahide Terazima. Laser-induced transient grating analysis of dynamics of interaction between sensory rhodopsin ii d75n and the htrii transducer. *Biophysical Journal*, 92(6):2028–2040, Mar 2007.

[194] G. Schmies, B. Lüttenberg, I. Chizhov, M. Engelhard, A. Becker, and E. Bamberg. Sensory rhodopsin ii from the haloalkaliphilic natronobacterium pharaonis: light-activated proton transfer reactions. *Biophysical Journal*, 78(2):967–976, Feb 2000.

[195] S. Subramaniam and R. Henderson. Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature*, 406(6796):653–657, Aug 2000.

[196] T. Rink, M. Pfeiffer, D. Oesterhelt, K. Gerwert, and H. J. Steinhoff. Unraveling photoexcited conformational changes of bacteriorhodopsin by time resolved electron paramagnetic resonance spectroscopy. *Biophysical Journal*, 78(3):1519–1530, Mar 2000.

Bibliography

[197] H. Steinhoff, A. Savitsky, C. Wegener, M. Pfeiffer, M. Plato, and K. Möbius. High-field epr studies of the structure and conformational changes of site-directed spin labeled bacteriorhodopsin. *Biochimica et Biophysica Acta*, 1457(3):253–262, Apr 2000.

[198] H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailler, and J. K. Lanyi. Structural changes in bacteriorhodopsin during ion transport at 2 angstrom resolution. *Science*, 286(5438):255–261, Oct 1999.

[199] H. J. Sass, G. Büldt, R. Gessenich, D. Hehn, D. Neff, R. Schlesinger, J. Berendzen, and P. Ormos. Structural alterations for proton translocation in the m state of wild-type bacteriorhodopsin. *Nature*, 406(6796):649–653, Aug 2000.

[200] A. A. Wegener, J. P. Klare, M. Engelhard, and H. J. Steinhoff. Structural insights into the early steps of receptor-transducer signal transfer in archaeal phototaxis. *EMBO Journal*, 20(19):5312–5319, Oct 2001.

[201] G. Metz, F. Siebert, and M. Engelhard. Asp85 is the only internal aspartic acid that gets protonated in the m intermediate and the purple-to-blue transition of bacteriorhodopsin. a solid-state 13c cp-mas nmr investigation. *FEBS Letters*, 303(2-3):237–241, Jun 1992.

[202] G. Metz, F. Siebert, and M. Engelhard. High-resolution solid state 13c nmr of bacteriorhodopsin: characterization of [4-13c]asp resonances. *Biochemistry*, 31(2):455–462, Jan 1992.

[203] K. V. Lakshmi, M. R. Farrar, J. Raap, J. Lugtenburg, R. G. Griffin, and J. Herzfeld. Solid state 13c and 15n nmr investigations of the n intermediate of bacteriorhodopsin. *Biochemistry*, 33(30):8853–8857, Aug 1994.

[204] V. Copié, A. E. McDermott, K. Beshah, J. C. Williams, M. Spijker-Assink, R. Gebhard, J. Lugtenburg, J. Herzfeld, and R. G. Griffin. Deuterium solid-state nuclear magnetic resonance studies of methyl group dynamics in bacteriorhodopsin and retinal model compounds: evidence for a 6-s-trans chromophore in the protein. *Biochemistry*, 33(11):3280–3286, Mar 1994.

[205] G. Gröbner, I. J. Burnett, C. Glaubitz, G. Choi, A. J. Mason, and A. Watts. Ob-

servations of light-induced structural changes of retinal within rhodopsin. *Nature*, 405(6788):810–813, Jun 2000.

[206] Ashish B Patel, Evan Crocker, Markus Eilers, Amiram Hirshfeld, Mordechai Sheves, and Steven O Smith. Coupling of retinal isomerization to the activation of rhodopsin. *Proceedings of the National Academy of Sciences of the United States of America*, 101(27):10048–10053, Jul 2004.

[207] M. Engelhard, S. Finkler, G. Metz, and F. Siebert. Solid-state c-13-nmr of [(3-c-13)pro]bacteriorhodopsin and [(4-c-13)pro]bacteriorhodopsin - evidence for a flexible segment of the c-terminal tail. *European Journal of Biochemistry*, 235(3):526–533, Feb 1996.

[208] Evan Crocker, Markus Eilers, Shivani Ahuja, Viktor Hornak, Amiram Hirshfeld, Mordechai Sheves, and Steven O Smith. Location of trp265 in metarhodopsin ii: implications for the activation mechanism of the visual receptor rhodopsin. *Journal of Molecular Biology*, 357(1):163–172, Mar 2006.

[209] Michael F Brown, Maarten P Heyn, Constantin Job, Suhkmann Kim, Stephan Moltke, Koji Nakanishi, Alexander A Nevzorov, Andrey V Struts, Gilmar F J Salgado, and Ingrid Wallat. Solid-state 2h nmr spectroscopy of retinal proteins in aligned membranes. *Biochimica et Biophysica Acta*, 1768(12):2979–3000, Dec 2007.

[210] J. G. Hu, B. Q. Sun, M. Bizounok, M. E. Hatcher, J. C. Lansing, J. Raap, P. J. Verdegem, J. Lugtenburg, R. G. Griffin, and J. Herzfeld. Early and late m intermediates in the bacteriorhodopsin photocycle: a solid-state nmr study. *Biochemistry*, 37(22):8088–8096, Jun 1998.

[211] Thorsten Mascher, John D Helmann, and Gottfried Unden. Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol Mol Biol Rev*, 70(4):910–938, Dec 2006.

[212] Hendrik Szurmant, Robert A White, and James A Hoch. Sensor complexes regulating two-component signal transduction. *Current Opinion in Structural Biology*, 17(6):706–715, Dec 2007.

Bibliography

[213] E. Zientz, J. Bongaerts, and G. Unden. Fumarate regulation of gene expression in escherichia coli by the dcusr (dcusr genes) two-component regulatory system. *Journal of Bacteriology*, 180(20):5421–5425, Oct 1998.

[214] P. Golby, S. Davies, D. J. Kelly, J. R. Guest, and S. C. Andrews. Identification and characterization of a two-component sensor-kinase and response-regulator system (dcus-dcur) controlling gene expression in response to c4-dicarboxylates in escherichia coli. *Journal of Bacteriology*, 181(4):1238–1248, Feb 1999.

[215] Ingo G Janausch, Inma Garcia-Moreno, and Gottfried Unden. Function of dcus from escherichia coli as a fumarate-stimulated histidine protein kinase in vitro. *Journal of Biological Chemistry*, 277(42):39809–39814, Oct 2002.

[216] I. G. Janausch, E. Zientz, Q. H. Tran, A. Kröger, and G. Unden. C4-dicarboxylate carriers and sensors in bacteria. *Biochimica et Biophysica Acta*, 1553(1-2):39–56, Jan 2002.

[217] Frank Alber, Svetlana Dokudovskaya, Liesbeth M Veenhoff, Wenzhu Zhang, Julia Kipper, Damien Devos, Adisetyantari Suprapto, Orit Karni-Schmidt, Rosemary Williams, Brian T Chait, Michael P Rout, and Andrej Sali. Determining the architectures of macromolecular assemblies. *Nature*, 450(7170):683–694, Nov 2007.

[218] Michael G Rossmann, Marc C Morais, Petr G Leiman, and Wei Zhang. Combining x-ray crystallography and electron microscopy. *Structure*, 13(3):355–362, Mar 2005.

[219] A. R. Pickford and I. D. Campbell. Nmr studies of modular protein structures and their interactions. *Chemical Reviews*, 104(8):3557–3565, Aug 2004.

[220] Maya Topf and Andrej Sali. Combining electron microscopy and comparative protein structure modeling. *Current Opinion in Structural Biology*, 15(5):578–585, Oct 2005.

[221] Oscar Hur and Kevin Karplus. Methods of translating nmr proton distances into their corresponding heavy atom distances for protein structure prediction with limited experimental data. *Protein Eng Des Sel*, 18(12):597–605, Dec 2005.

[222] Andrea Cavalli, Xavier Salvatella, Christopher M Dobson, and Michele Vendruscolo. Protein structure determination from nmr chemical shifts. *Proceedings of the*

National Academy of Sciences of the United States of America, 104(23):9615–9620, Jun 2007.

[223] Rhiju Das, Bin Qian, Srivatsan Raman, Robert Vernon, James Thompson, Philip Bradley, Sagar Khare, Michael D Tyka, Divya Bhat, Dylan Chivian, David E Kim, William H Sheffler, Lars Malmström, Andrew M Wollacott, Chu Wang, Ingemar Andre, and David Baker. Structure prediction for casp7 targets using extensive all-atom refinement with rosetta@home. *Proteins*, 69 Suppl 8:118–128, 2007.

[224] Tomasz Cierpicki, John H Bushweller, and Zygmunt S Derewenda. Probing the supramodular architecture of a multidomain protein: the structure of syntenin in solution. *Structure*, 13(2):319–327, Feb 2005.

[225] Andrej Sali, Robert Glaeser, Thomas Earnest, and Wolfgang Baumeister. From words to literature in structural proteomics. *Nature*, 422(6928):216–225, Mar 2003.

[226] Anthony G Lee. How lipids affect the activities of integral membrane proteins. *Biochimica et Biophysica Acta*, 1666(1-2):62–87, Nov 2004.

[227] Lucia Pappalardo, Ingo G Janausch, Vinesh Vijayan, Eva Zientz, Jochen Junker, Wolfgang Peti, Markus Zweckstetter, Gottfried Unden, and Christian Griesinger. The nmr structure of the sensory domain of the membranous two-component fumarate sensor (histidine protein kinase) dcus of escherichia coli. *Journal of Biological Chemistry*, 278(40):39185–39188, Oct 2003.

[228] Manuel Etzkorn, Anja Böckmann, François Penin, Dietmar Riedel, and Marc Baldus. Characterization of folding intermediates of a domain-swapped protein by solid-state nmr spectroscopy. *Journal of the American Chemical Society*, 129(1):169–175, Jan 2007.

[229] Kylie J Watts, Mark S Johnson, and Barry L Taylor. Minimal requirements for oxygen sensing by the aerotaxis receptor aer. *Molecular Microbiology*, 59(4):1317–1326, Feb 2006.

[230] B. L. Taylor and I. B. Zhulin. Pas domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev*, 63(2):479–506, Jun 1999.

[231] Yunjun Wang and Oleg Jardetzky. Probability-based protein secondary structure

identification using combined nmr chemical-shift data. *Protein Science*, 11(4):852–861, Apr 2002.

[232] Stephen Neal, Alex M Nip, Haiyan Zhang, and David S Wishart. Rapid and accurate calculation of protein 1h, 13c and 15n chemical shifts. *Journal of Biomolecular NMR*, 26(3):215–240, Jul 2003.

[233] D. S. Wishart and B. D. Sykes. Chemical-shifts as a tool for structure determination. In *Nuclear Magnetic Resonance, Part C*, volume 239 of *Methods in Enzymology*, pages 363–392. Academic Press Inc, San Diego, 1994.

[234] A. Pandini and L. Bonati. Conservation and specialization in pas domain dynamics. *Protein Eng Des Sel*, 18(3):127–137, Mar 2005.

[235] Jocelyne Vreede, Michael A van der Horst, Klaas J Hellingwerf, Wim Crielaard, and Daan M F van Aalten. Pas domains. common structure and common flexibility. *Journal of Biological Chemistry*, 278(20):18434–18439, May 2003.

[236] Madhumati Sevvana, Vinesh Vijayan, Markus Zweckstetter, Stefan Reinelt, Dean R Madden, Regine Herbst-Irmer, George M Sheldrick, Michael Bott, Christian Griesinger, and Stefan Becker. A ligand-induced switch in the periplasmic domain of sensor histidine kinase cita. *Journal of Molecular Biology*, 377(2):512–523, Mar 2008.

[237] Holger Kneuper, Ingo G Janausch, Vinesh Vijayan, Markus Zweckstetter, Verena Bock, Christian Griesinger, and Gottfried Unden. The nature of the stimulus and of the fumarate binding site of the fumarate sensor dcus of escherichia coli. *Journal of Biological Chemistry*, 280(21):20596–20603, May 2005.

[238] J. Krämer, J. D. Fischer, E. Zientz, V. Vijayan, C. Griesinger, A. Lupas, and G. Unden. Citrate sensing by the c4-dicarboxylate/citrate sensor kinase dcus of escherichia coli: binding site and conversion of dcus to a c4-dicarboxylate- or citrate-specific sensor. *Journal of Bacteriology*, 189(11):4290–4298, Jun 2007.

[239] Jason Key, Marco Hefti, Erin B Purcell, and Keith Moffat. Structure of the redox sensor domain of azotobacter vinelandii nifl at atomic resolution: signaling, dimerization, and mechanism. *Biochemistry*, 46(12):3614–3623, Mar 2007.

[240] A. Repik, A. Rebbapragada, M. S. Johnson, J. O. Haznedar, I. B. Zhulin, and B. L. Taylor. Pas domain residues involved in signal transduction by the aer redox sensor of escherichia coli. *Molecular Microbiology*, 36(4):806–816, May 2000.

[241] S. Six, S. C. Andrews, G. Unden, and J. R. Guest. Escherichia coli possesses two homologous anaerobic c4-dicarboxylate membrane transporters (dcua and dcub) distinct from the aerobic dicarboxylate transport system (dct). *Journal of Bacteriology*, 176(21):6470–6478, Nov 1994.

[242] Alberto Marina, Carey D Waldburger, and Wayne A Hendrickson. Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. *EMBO Journal*, 24(24):4247–4259, Dec 2005.

[243] Kylie J Watts, Kirsten Sommer, Sheena L Fry, Mark S Johnson, and Barry L Taylor. Function of the n-terminal cap of the pas domain in signaling by the aerotaxis receptor aer. *Journal of Bacteriology*, 188(6):2154–2162, Mar 2006.

[244] Xiaolei Ma, Nazish Sayed, Padmamalini Baskaran, Annie Beuve, and Focco van den Akker. Pas-mediated dimerization of soluble guanylyl cyclase revealed by signal transduction histidine kinase domain crystal structure. *Journal of Biological Chemistry*, 283(2):1167–1178, Jan 2008.

[245] Aly E Abo-Amer, Jonathan Munn, Kerry Jackson, Murat Aktas, Paul Golby, David J Kelly, and Simon C Andrews. Dna interaction and phosphotransfer of the c4-dicarboxylate-responsive dcus-dcur two-component regulatory system from escherichia coli. *Journal of Bacteriology*, 186(6):1879–1889, Mar 2004.

[246] Rachel W Martin and Kurt W Zilm. Preparation of protein nanocrystals and their characterization by solid state nmr. *Journal of Magnetic Resonance*, 165(1):162–174, Nov 2003.

[247] J. Sambrook and D.W. Russel. *Molecular Cloning*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2001.

[248] J.H. Miller. *A short course in bacterial genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1992.

[249] B. Miroux and J. E. Walker. Over-production of proteins in escherichia coli: mu-

tant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology*, 260(3):289–298, Jul 1996.

[250] J. L. Rigaud, B. Pitard, and D. Levy. Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. *Biochimica et Biophysica Acta*, 1231(3):223–246, Oct 1995.

[251] Charles D Schwieters, John J Kuszewski, Nico Tjandra, and G. Marius Clore. The xplor-nih nmr molecular structure determination package. *Journal of Magnetic Resonance*, 160(1):65–73, Jan 2003.

[252] E. G. Stein, L. M. Rice, and A. T. Brünger. Torsion-angle molecular dynamics as a new efficient tool for nmr structure calculation. *Journal of Magnetic Resonance*, 124(1):154–164, Jan 1997.

[253] Jens P Linge, Mark A Williams, Christian A E M Spronk, Alexandre M J J Bonvin, and Michael Nilges. Refinement of protein structures in explicit solvent. *Proteins*, 50(3):496–506, Feb 2003.

[254] Maya Topf, Matthew L Baker, Bino John, Wah Chiu, and Andrej Sali. Structural characterization of components of protein assemblies by comparative modeling and electron cryo-microscopy. *J Struct Biol*, 149(2):191–203, Feb 2005.

[255] Vladimir Yarov-Yarovoy, Jack Schonbrun, and David Baker. Multipass membrane protein structure prediction using rosetta. *Proteins*, 62(4):1010–1025, Mar 2006.

[256] C. Dominguez, R. Boelens, and Amjj Bonvin. Haddock: A protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, 125(7):1731–1737, Feb 2003.

[257] Sjoerd J de Vries, Aalt D J van Dijk, Mickaël Krzeminski, Mark van Dijk, Aurelien Thureau, Victor Hsu, Tsjerk Wassenaar, and Alexandre M J J Bonvin. Haddock versus haddock: new features and performance of haddock2.0 on the capri targets. *Proteins*, 69(4):726–733, Dec 2007.

[258] J. L. Markley, A. Bax, Y. Arata, C. W. Hilbers, R. Kaptein, B. D. Sykes, P. E. Wright, and K. Wuthrich. Recommendations for the presentation of nmr structures of proteins and nucleic acids - (iupac recommendations 1998). *Pure and Applied*

Acknowledgements

D Acknowledgements

L E B E N S L A U F

Manuel Etzkorn
Geboren am 12.Mai 1979 in Aachen
Deutsche Staatsangehörigkeit

Studium und schulische Ausbildung

Seit Februar 2005	Promotionsstudent und wissenschaftlicher Mitarbeiter der Arbeitsgruppe Festkörper NMR Spektroskopie von Dr. Marc Baldus, Max-Planck Institut für Biophysikalische Chemie, Göttingen
Nov 2004	Physik Diplom, Georg-August Universität, Göttingen
April 2003 – Okt 2004	Diplomarbeit mit dem Thema: „Untersuchung von molekularer Struktur und Dynamik mittels Festkörper-NMR: Neue Methoden und Anwendungen", MPI biophys. Chemie, Göttingen.
WS 2002/2003	Auslandsaufenthalt im Rahmen eines Erasmus-Sokrates Teilstipendiums, Universidad de La Laguna, San Cristobal de La Laguna, Spanien
WS 1999 – WS 2004	Studium der Physik mit Schwerpunkten in der Biophysik und Materialphysik, Georg-August-Universität, Göttingen.
Juli 1998 – August 1999	Zivildienst in der „Jungenbildungs- und Freizeitstätte Heinrich-Rabbich Haus", Essen
Juni 1998	Abitur, Gymnasium-Essen Überruhr

Wissenschaftliche Hilfskraft in der Lehre

SS 2006	Einführung in die Biophysik (Prof. Salditt)
SS2003 – SS 2004	Physikalisches Nebenfach Praktikum für Biologen und Chemiker (PD. Ronning, Dr. Uhrmacher)

Wissenschaftliche Auszeichnungen

Dez 2005	Minerva-Gentner Symposium: „Best scientific student presentation", Eilat, Israel