APPENDIX A **128**

List of Publications

This thesis is based on the following publications:

1. Jehle S, van Rossum BJ, Stout JR, Noguchi SR, Falber K, Rehbein K, Oschkinat H, Klevit RE, and Rajagopal P (2008) αB-Crystallin: A hybrid solid-solution state NMR investigation reveals structural aspects of the heterogeneous oligomer. (Submitted)

2. Scholz I, Jehle S, Schmieder P, Hiller M, Eisenmenger F, Oschkinat H, and van Rossum BJ (2007) J-deconvolution using maximum entropy reconstruction applied to 13C-13C solid-state cross-polarization magic-angle-spinning NMR of proteins. *J Am Chem Soc,* **129,** 6682-6683.

3. Jehle S, Hiller M, Rehbein K, Diehl A, Oschkinat H, and van Rossum BJ (2006) Spectral editing: selection of methyl groups in multidimensional solid-state magic-angle spinning NMR. *J Biomol NMR,* **36,** 169-177.

Further Publications:

4. Hiller M, Higman VA, Jehle S, van Rossum BJ, Kuhlbrandt W, and Oschkinat H (2008) [2,3-(13)C]-labeling of aromatic residues--getting a head start in the magic-angle-spinning NMR assignment of membrane proteins. *J Am Chem Soc,* **130,** 408-409.

5. Balayssac S, Bertini I, Falber K, Fragai M, Jehle S, Lelli M, Luchinat C, Oschkinat H, and Yeo KJ (2007) Solid-state NMR of matrix metalloproteinase 12: an approach complementary to solution NMR. *Chembiochem,* **8,** 486-489.

6. Jehle S, Rehbein K, Diehl A, and van Rossum BJ (2006) Amino-acid selective experiments on uniformly 13C and 15N labeled proteins by MAS NMR: Filtering of lysines and arginines. *J Magn Reson,* **183,** 324-328.

7. Fuchs S, Otto H, Jehle S, Henklein P, and Schluter AD (2005) Fluorescent dendrimers with a peptide cathepsin B cleavage site for drug delivery applications. *Chem Commun (Camb),* 1830-1832.

Talks at Conferences:

Jehle S, Rajagopal P, Oschkinat H, van Rossum BJ (2007) Chemical shift analysis and structural studies of human αB-crystallin by MAS solid-state NMR. *16th Triennial Conference for the International Society of Magnetic Resonance.* October 14-19, 2007; Kenting (Taiwan).

Abbreviations

1,3G-αB	1,3-^{13}C-glycerol, ^{15}N labelled αB-crystallin
2G-αB	2-^{13}C-glycerol, ^{15}N labelled αB-crystallin
1D, 2D, 3D	one dimensional, two dimensional, three dimensional
αA	αA-crystallin
αB	αB-crystallin
Å	Ångstrøm (1 Å = 0.1 nm)
CP	cross polarization
CSA	chemical shift anisotropy
DARR	Dipolar Assisted Rotational Resonance
FID	free induction decay
γ	gyromagnetic ratio
Hz	Hertz
kDa	kilo Dalton
MAS	magic-angle spinning
MjHSP16.5	Methanococcus jannashii heat shock protein 16.5
NMR	nuclear magnetic resonance
PDSD	proton-driven spin diffusion
PFMW-αB	αB-crystallin sample that has proline, phenylalanine, methionine and tryptophane ^{13}C, ^{15}N labelled
ppm	parts per million
rf	radio frequency
RFDR	Radio Frequency Driven Recoupling
sCP	specific CP
SH3	Src-homology 3
TPPI	Two pulse phase incrementation
TPPM	Two-pulse phase modulation
wHSP16.9	wheat heat shock protein 16.9

1. General Introduction

Advances in structural biology have been strongly supported by the success of three leading techniques: X-ray crystallography, solution NMR spectroscopy and electron microscopy. In recent years, solid-state NMR spectroscopy has matured and will extend these tools in structural biology (1). Today, structure determination of non-crystalline solid protein preparations has been made possible through rapid progress of solid-state NMR methodology for biological systems. Castellani *et. al.* have solved and refined a first structure of a microcrystalline protein by using only solid-state MAS NMR spectroscopy (2-4). Its success relies on the ability to access systems that are difficult to crystallise or that form large heterogeneous complexes and insoluble aggregates. Examples are WW-fibrils and the HET-s prion protein that have been studied using solid-state NMR (5;6). Ligands bound to insoluble receptors form another target for solid-state NMR, i.e. Lange and coworkers solved the structure of a toxin, which binds to a potassium channel (7).

Such "difficult" biological systems remained so far a challenge, and combinations of EM, X-ray crystallography, solution NMR and/or solid-state NMR might be required for a complete structural description. For instance, in 2007 a part of the spliceosomal complex, consisting of RNA and several proteins was solved by using a combination of X-ray crystallography and solution NMR spectroscopy (8). Al-Amoudi *et al.* used electron microscopy and a previously solved X-ray structure, to describe the molecular architecture of cadherins in native epidermal desmosomes (9).

Another group of "difficult" proteins are membrane proteins, for which 30 % of the proteinogenic genes code (10). They are generally difficult to crystallise and mostly insoluble in water without detergents. Small membrane proteins (< ~50 kDa) may be investigated using solution NMR (11). However, most membrane proteins have multiple domains and are frequently organised as large multimeric complexes of several hundred kDa of size. Solution NMR can be used to study the soluble extra- and intra-cellular domains and solid-state NMR to study the insoluble membrane-integrated domain (12). Electron microscopy can yield information on the assembly of the complete complex. Docking programmes, such as HADDOCK (13) may be used

to calculate a structure of the complete complex by using the structures of the components and experimental information of the interaction sites.

The aim of this thesis is the structure determination of a heterogeneous human small heat-shock protein (sHSP), αB crystallin, and method development for solid-state NMR. Heterogeneous oligomeric systems and membrane proteins are mostly of high molecular weight and show very crowded NMR spectra, resulting in reduced resolution. Therefore, it is desirable to develop methods that enhance the resolution of NMR spectra.

A hybrid solid-solution-state NMR study is used to investigate a human small heat-shock protein (sHSP), αB-crystallin (αB), that forms heterogeneous oligomers of several hundred kDa of size, and which does not crystallise. Human sHSPs are involved in many diseases that are connected to cellular stress and missfolded proteins and thus interesting pharmacological targets. The outer membrane protein G (OmpG) of *E. coli* was used as a realistic model system for membrane proteins to develop methods for resolution enhancement in solid-state NMR. In chapter 3 a method for Methyl-Filtering and *J*-decoupling is presented.

1.1 Quantum mechanical basics of NMR

Nuclei with a spin quantum number unequal to zero possess a quantum-mechanical property that is called nuclear spin. By this, these nuclei interact with a magnetic field. Under the influence of an external magnetic field B_0 a nuclear spin behaves like a tiny magnet (figure 1-1).

figure 1-1 – In an external magnetic field, each nucleus in a molecule can have two possible orientations: parallel or antiparallel to the external magnetic field B_0. In a quantum mechanical

description, these two orientations correspond to two discrete energy levels.

For spin ½ nuclei, the orientation may be parallel or antiparallel to the external magnetic field. The energy difference between the two states induced by B_0, is given by equation 1:

$$\Delta E = \frac{\gamma h B_0}{2\pi},$$

equation 1-1

where h is Planck's constant, γ is the gyromagnetic ratio and B_0 the external magnetic field. By introducing the Bohr condition $(\Delta E = h\upsilon)$, the energy difference υ_0 between the two levels can be written in frequency units as:

$$\upsilon_0 2\pi = \gamma B_0 = \omega_0.$$

equation 1-2

In the thermodynamic equilibrium, spins are nutating around the field B_0, with the Larmor frequency ω_0 which correlates with the energy difference between the parallel and antiparallel state. The population of the energy levels is described by the Boltzmann equation (equation 1-3), the lower energetic level is slightly preferred, and hence, slightly more populated (14):

$$\frac{N_{\text{upper}}}{N_{\text{lower}}} = e^{\frac{-\Delta E}{kT}} = e^{\frac{-h\upsilon_0}{kT}}.$$

equation 1-3 – Boltzmann Distribution

In this expression, T represents the absolute temperature in Kelvin and k the Boltzmann constant. There is a slight excess population of spins with orientation parallel to the external field, due to the lower energy of this state. This is defined as z-magnetisation.

Radio frequency (rf) pulses applied from an axis perpendicular to B_0, with a frequency close to the Larmor frequency ω_0, lead to a disturbance of the thermodynamic equilibrium and redirect the magnetisation according to Lenz´ law. A pulse which turns the macroscopic z-magnetisation into the orthogonal *xy*-plane, is called a 90° pulse. Following such a pulse, both energy levels are equally populated. By introducing a frame that rotates with the Larmor frequency, the spins appear static in this so-called rotating frame. There are, however, small differences in the nutation frequencies of different kinds of spins due to interactions other than with the B_0 field. For instance, in the rotating frame, the spins are evolving under their chemical shifts, which depend on the gyromagnetic ratio and the amount of shielding by the chemical environment (see chapter 1.2).

The relaxation to thermodynamic equilibrium after an excitation pulse is governed by two mechanisms, the spin-lattice (T_1) and spin-spin (T_2) relaxation. Both of them result in a decay of the magnetisation, which is recorded as the "Free Induction Decay" (FID). Relaxation occurs due to heat exchange with the lattice (T_1-relaxation), while the interactions with local fields, represented by the other nuclei surrounding the distinct nucleus, is the driving force behind T_2-relaxation.

1.2 Spin interactions

The physical Interactions of nuclei with non-zero magnetic moments (spin-quantum number unequal to zero) are quantum mechanically described by the Hamilton operator. Each possible Interaction is described by a distinct term in the Hamilton operator (equation 4):

$$H = H_Z + H_{RF} + H_{CS} + H_J + H_Q + H_D.$$

equation 1-4

H_Z describes the Zeeman Interaction, which is directly related to the Larmor precession ω_0 in the absence of other interactions (chapter 1.1). Generally, the *z*-axis is selected parallel to B_0 and H_Z can be written as

$$H_Z = -\gamma_I \hbar I_z B_0 \,,$$

equation 1-5

where $\hbar$ is given as $\dfrac{h}{2\pi}$ with h as Planck's constant, B_0 is the z-component of the magnetic field vector $\vec{B}_0$, and I_z is the *z*-component of the spin vector $\vec{I}$. In the frame rotating with ω_0, spins appear static and the Zeeman precession effectively cancels from equation 1.4.

H_{RF} describes the interaction with the magnetic components of a radio-frequency field (rf-field). By irradiation in the *x*-direction in the rotating frame, the corresponding time-independent Hamiltonian is:

$$H_{RF} = -\gamma_I \hbar I_x B_1 \,.$$

equation 1-6

With respect to $\vec{B}_0$, the rf-field is very weak and only induces transitions if applied with a frequency close to the Larmor frequency of an observable nucleus ("on resonance").

H_{CS} describes the chemical shift Hamiltonian. All nuclei of the same species have the same gyromagnetic ratio, and in a given external magnetic field the same Larmor frequency. However, the electron distribution surrounding each nucleus induces small local fields, according to Lenz´ law, that shield the nucleus. This changes the field, observed by the nucleus. These variations in nutation frequency are very small and are expressed as parts-per-million (ppm). They are recorded as the chemical shift. In the laboratory frame with the z-axis parallel to B_0 the chemical shift Hamiltonian is

$$H_{CS} = \gamma_I \hbar I_z \sigma_{zz} B_0 \,,$$

equation 1-7

where σ_{zz} is the shielding tensor. Since the electron distribution is not spherically symmetric, the shielding effect is not equal for all three space directions, hence H_{CS} is dependent on the orientation of the molecule with respect to B_0. The orientation dependence is also known as anisotropy. An angle θ describes the orientation of the anisotropic tensor with respect to B_0 in the z-direction. For instance, the tensor shape for the shielding of carbonyl carbons resembles a "cigar" due to the electronegativity of the oxygen, while methyl carbons have a rather spherical distribution of electrons. Anisotropic tensors are multiplied with a term $(3 \cdot \cos^2 \theta - 1)$ to describe the orientation of the tensor with respect to B_0.

H_J is the Hamiltonian which describes the J-coupling and has the form

$$H_J = J\vec{I}\vec{S} \, .$$

equation 1-8

$\vec{I}$ and $\vec{S}$ represent the spins of the two nuclei and J represents the coupling constant. J-couplings are mediated through the chemical bonds and are weak compared to the much larger dipolar interactions.

H_Q is the quadrupolar interaction Hamiltonian and proportional to the electric field gradient tensor V, containing an anisotropy term $(3\cos^2 \theta - 1)$. H_Q is non-zero only for nuclei with spins greater than ½ and therefore will not be further discussed in this thesis.

Dipolar interactions are of particular interest in solid-state NMR-spectroscopy. They are described by the Hamiltonian H_D. Each spin S generates a local field $\vec{B}_{loc}$ experienced by the spin I:

$$\vec{B}_{loc} = \frac{\vec{\mu}_S}{r_{IS}^3}\left(3\cos^2 \theta_{IS} - 1\right).$$

equation 1-9

Here $\vec{\mu}_S$ is the magnetic moment of spin S and r_{IS} is the internuclear distance. θ_{IS} is the angle made by the vector connecting the two spins and the B_0 field. Thus, the

dipolar interaction is both distance and orientation dependent. The Hamiltonian for the homonuclear dipolar interaction has the form:

$$H_D^{II} = \frac{\mu_0 \gamma_I^2 \hbar^2}{4\pi r_{II}^3} \frac{1}{2} \left(3\cos^2\theta - 1\right)\left(\vec{I}_1\vec{I}_2 - 3I_{z1}I_{z2}\right).$$

equation 1-10

For the heteronuclear dipolar interaction, the Hamiltonian has a simpler form:

$$H_D^{IS} = \frac{\mu_0 \gamma_I \gamma_S \hbar^2}{4\pi r_{IS}^3} \left(3\cos^2\theta - 1\right)\left(I_z S_z\right).$$

equation 1-11

Here μ_0 is the permeability constant in a vacuum and I_z, S_z are the z-components of the spin vectors $\vec{I}$ and $\vec{S}$. For structure research, the most interesting part of these formulas is the distance dependence $\left(\frac{1}{r^3}\right)$, which can be exploited to gain distance information.

The observation of anisotropic interactions is the major difference between solution and solid-state NMR. In solution, the angle θ is averaged out due to molecular tumbling and, hence, only the isotropic parts in the Hamilton Operators remain. This is not the case for immobilized samples. Especially the strong homonuclear coupled network of protons in solid proteins can lead to non-interpretable spectra. Also, the chemical shift has an isotropic and an anisotropic part in the Hamilton Operator, where the anisotropic part leads to line broadening only for static samples. In the following chapter it will be described how the information of these interactions can be extracted and what has to be done to cancel out unwanted interactions, which lead to line broadening due to their anisotropy.

1.3 Basic tools in solid-state NMR

1.3.1 Magic Angle Spinning (MAS)

Terms in the Hamiltonian containing the factor $(3\cos^2\theta - 1)$ are dependent on the orientation of the anisotropic tensor in the magnetic field. In particular, the chemical shift can be strongly orientation dependent. This leads to line broadening for nuclei with anisotropic electron shielding, as it is the case for carbonyl carbons. The dependence on θ is exploited in MAS NMR. At an angle of 54.7°, the term $(3\cos^2\theta - 1)$ becomes zero. Thus by spinning the sample at the magic angle (figure 1-2) with a rotor frequency ω_r larger than the anisotropic tensors, these interactions are averaged out which leads to sharp lines. In other words, the anisotropic tensors in the numerous molecules change orientation dependent on the spinning frequency and the orientation of the rotor in the magnet. As a rule-of-thumb, the anisotropic interactions are averaged out if the rotation is three times faster as the magnitude of the interaction. The chemical shifts recorded under MAS are thus the isotropic shifts, the same as the shifts recorded in liquid state NMR spectra.
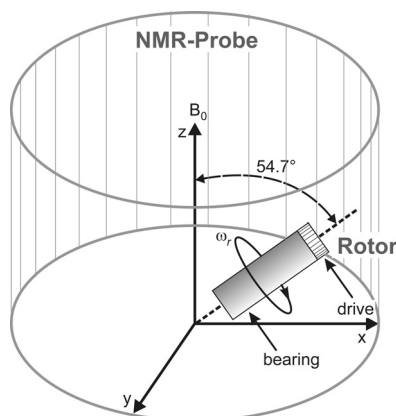
figure 1-2 – spinning at the magic angle. The z-axis represents the direction of $\vec{B}_0$.

The influence of the spinning frequency on the chemical shift anisotropy is shown in figure 1-3.
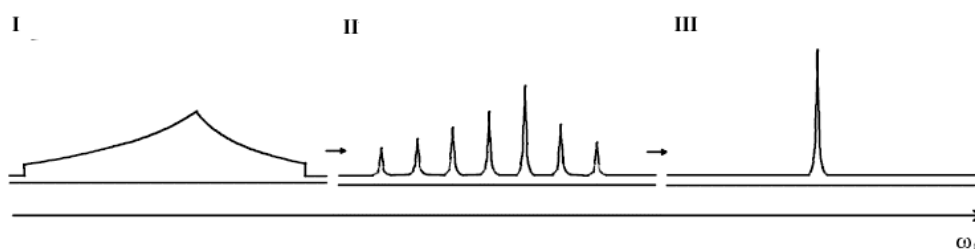
figure 1-3 - dependence of the line shape on the spinning frequency ω_r.

In figure 1-3 I, a solid-state NMR spectrum of a static sample is shown, while from II to III the rotor frequency is increased. For slow MAS (II), spinning sidebands appear at multiples of the rotation frequency ω_r, while at fast MAS (III), the isotropic chemical shift is observed and the chemical shift anisotropy is averaged out. In MAS solid-state NMR of proteins, the spectroscopist faces a tradeoff between sample stability and MAS frequency. That means that *g*-forces and heating might damage the sample at high ω_r. Typically, for 4 mm rotors, spinning frequencies between 8,000 and 12,500 Hz form a good compromise between sample stability and obliteration of anisotropic interactions.

1.3.2 Cross polarisation and high power proton decoupling

The most abundant magnetic nucleus in organic and bioorganic compounds is ^{1}H. Protons build a network of large dipolar couplings because of their large gyromagnetic ratio. On the one hand, the large γ results in good sensitivity due to the large population differences, but on the other hand, it leads to significant line broadening. ^{1}H-^{1}H dipolar couplings are in the range of 100 kHz, and MAS frequencies in the range of 8 - 15 kHz cannot cancel out this interaction. The larger polarisation of ^{1}H can, however, be exploited to enhance the sensitivity of the low-γ nuclei, by transfer to the low γ nuclei via cross polarisation (CP). CP uses the larger Boltzmann population of ^{1}H to enhance the population of low γ nuclei like ^{13}C and ^{15}N. In addition, the faster relaxation time of protons determines the repetition time for the experiment, which forms an additional advantage over direct carbon or nitrogen excitation. CP can be combined with high-power heteronuclear dipolar decoupling of protons during acquisition, which leads to narrow line width of the detected nuclei.

Hence, the standard 1D pulse sequence in ssNMR (figure 1-4) is:
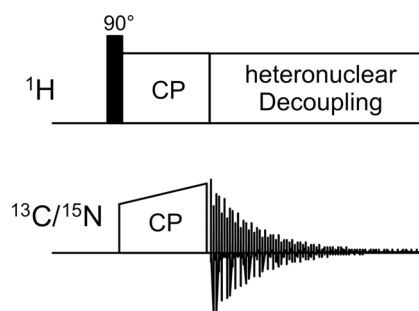
figure 1-4 – 1D CPMAS experiment. Several techniques for heteronuclear decoupling during acquisition are available, e.g. TPPM (Two Pulse Phase Modulation) or SPINAL64, for details see Reviews (10;15-17).

In this experiment, protons are first excited by a 90° pulse. In the next step, proton magnetisation is transferred to the low γ nucleus ^{13}C by the CP unit. During CP, both high- and low-γ nuclei are irradiated with rf-pulses that satisfy the Hartmann-Hahn matching condition (see below). Finally, the FID is detected under heteronuclear proton decoupling. The effects of CP and decoupling on the sensitivity and the linewidth are illustrated in figure 1-5:
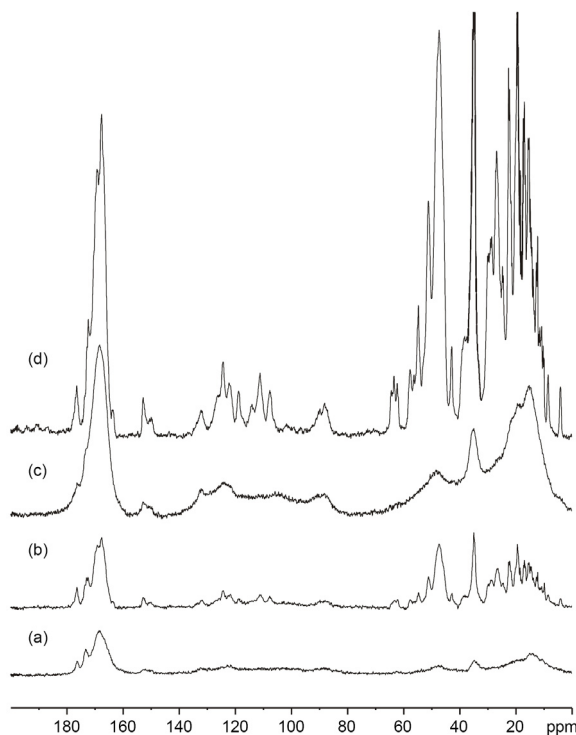
figure 1-5 – Solid-state ^{13}C MAS NMR spectra of the α-spectrin SH3 domain, acquired at a field of 9.4 T. The spinning frequency was set to ωr/2 π = 8.0 kHz and the temperature to 298 K. The acquisition time of each experiment was 10 min. (a) A single 90º pulse on the ^{13}C was applied. (b) Additional TPPM decoupling was used during acquisition. (c) The Ramp-CP method was employed but no TPPM. (d) Additional TPPM decoupling yielded the spectrum with highest intensity and highest resolution due to better signal-to-noise with respect to the spectra recorded without TPPM. Reproduced with permission from reference (4).

The spectrum in (a) was recorded applying a single 90° pulse on carbon. The spectrum shows both low resolution and intensity. The resolution enhancement observed in (b) is the result of TPPM decoupling. The strong heteronuclear dipolar couplings with the protons can be suppressed by sophisticated decoupling schemes and higher resolution can be achieved. In (c) and (d), CP is introduced. The spectrum in (d) is recorded with heteronuclear decoupling and the spectrum in (c) without. The gain in intensity is obvious in both spectra with respect to (a) and (b).

In the CP experiment, both ^{1}H and ^{13}C are "spin-locked" by two transverse rf-pulses. These spin-lock pulses result in a coherent nutation of the two spins. For energy conservation in a static sample, magnetisation will only be transferred from ^{1}H to ^{13}C if the energy difference of the two spin states on ^{1}H and ^{13}C are equal, the so-called Hartmann-Hahn matching condition (figure 1-6):

$$|\downarrow\rangle \underline{\quad{}^1\mathrm{H}\quad} \qquad |\downarrow\rangle \underline{\quad{}^{13}\mathrm{C}\quad}$$

Hartmann-Hahn Matching

$$\omega_{1H}(=\gamma_{1H}H_{1H}) \quad = \quad \omega_{13C}(=\gamma_{13C}H_{13C})$$
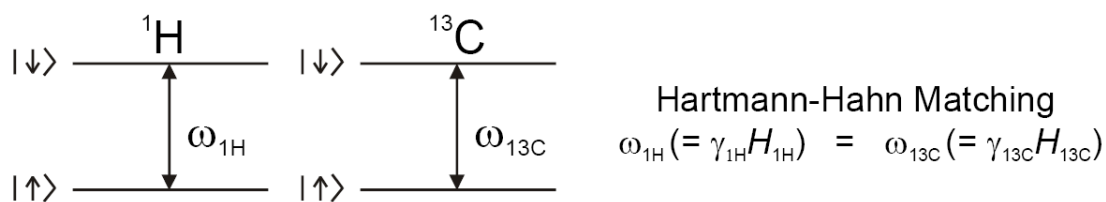
figure 1-6 – The classic Hartmann-Hahn matching condition for a static sample is shown. This matching condition allows transfer of magnetisation from ^{1}H to low γ-nuclei like ^{13}C.

Hence, Hartmann-Hahn matching is fulfilled if the energy difference between the two nuclei is matched by irradiation of two matching rf-fields on-resonance on the ^{1}H and ^{13}C channels. Under MAS, the matching condition is modified to $\omega_{1H} - \omega_{13C} = \pm n\omega_r$, n is an integer and has the value 1 or 2. In practice, a ramp on one of the channels improves signal stability and compensates for B_1 inhomogeneity.

^{1}H TPPM decoupling rotates the proton spins from one direction to the other. If the power for the decoupling is strong enough, the heteronuclear dipolar interaction is then averaged to zero. Under optimal TPPM decoupling, using well-ordered protein crystals even carbon-carbon J-couplings can be observed.

1.3.3 Cross polarisation between low-γ nuclei

The CP method is not only used to enhance sensitivity in MAS NMR. For sequential assignment purposes, this technique has been modified to direct magnetisation between nitrogen and carbon in a frequency-selective manner (18;19).

For specific cross polarisation, the low γ nuclei are spin-locked by weak rf-fields to match the specific CP condition (equation 1-12)

$$\sqrt{\Omega_C^2 + \omega_C^2} - \sqrt{\Omega_N^2 + \omega_N^2} = n\omega_r \, ,$$

equation 1-12

where Ω_C and Ω_N represent the frequency offsets of the nuclei, ω_C and ω_N represent the applied spin lock power and ω_r represents the sample spinning frequency. For matching, *n* should be ±1 or ±2. Low rf-power guarantees the specificity by band selectivity, while the offset is adjusted to fulfil the matching.

During adiabatic cross polarisation, the ^{13}C frequency is set on resonance for either Cα or C´, while the ^{15}N frequency is set on resonance for the backbone ^{15}N spins. A rectangular shape is used on the ^{15}N channel, the ^{13}C rf field amplitude is tangentially modulated from below to above the Hartmann-Hahn matching condition (19), such that magnetisation is either transferred from ^{15}N to Cα or C´ dependent on the ^{13}C offset.

By combination of N(*i*)-Cα(*i*) and N(*i*)-C´(*i*-1) experiments, backbone nitrogens can be sequentially assigned to the carbons of the neighbouring amino acid in the peptide chain (figure 1-7).
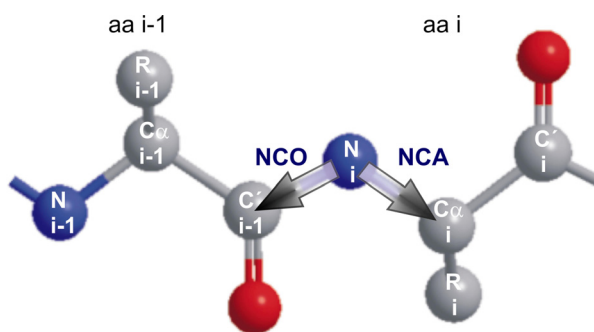
figure 1-7– backbone polarisation transfer by NCA and NCO experiments.

An extension of this approach leads to the design of NCACX and NCOCX experiments which are now common for sequential assignment (see chapter 1.4.2). Pauli and coworkers achieved the full assignment of the α-spectrin SH3 domain by this method (20).

1.3.4 Reintroduction of "lost" terms into the Hamilton operator

The effect of MAS is not only the suppression of unwanted line broadening. Dipolar interactions which contain desired structural information are partially cancelled out as well, since the term describing dipolar interactions becomes zero under fast MAS (figure 1-8a). If pulse trains are applied in a synchronized manner with the MAS, the dipolar interaction can be partly reintroduced by "recoupling" (figure 1-8b).
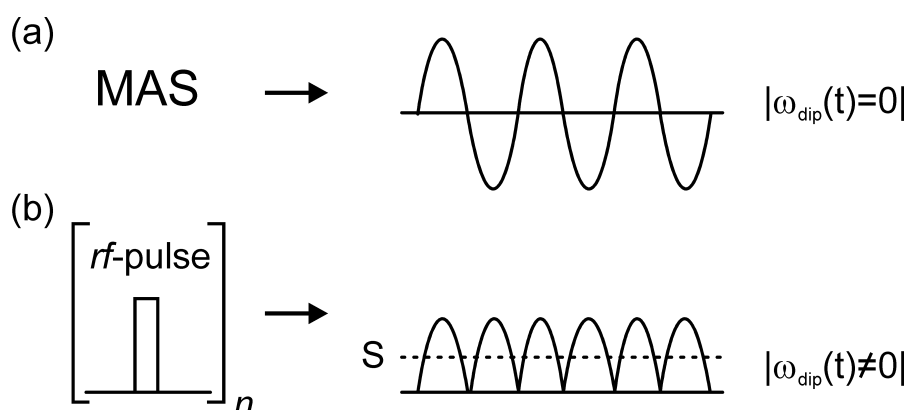
figure 1-8 – basic idea of dipolar recoupling under MAS. In (a), the dipolar coupling ω_{dip} **becomes zero, due to the oscillatory behavior of the dipolar interaction under MAS. Rotor synchronized rf-pulses yield a non-zero** ω_{dip} **in (b). figure reproduced with permission from reference (4).**

Nowadays a large number of pulse sequences that can be used are available for recoupling. For homonuclear recoupling, zero-quantum techniques like RFDR (21) and multiple quantum techniques like DREAM (22) are commonly used. In the aforementioned reviews (10;15-17) additional homonuclear recoupling schemes like DRAMA, MELODRAMA and HORROR are discussed. For heteronuclear recoupling, REDOR (23) has become a widely used technique. With methods of dipolar recoupling, spectra are recorded containing correlations through space.

1.3.5 Spin diffusion

In a magnetic field, extended systems of dipolar coupled nuclei show the effect of spin diffusion. In MAS solid-state NMR spectroscopy this effect can be exploited for homonuclear ^{13}C-^{13}C polarisation transfer in ^{13}C-^{13}C-correlation spectroscopy. In 3D triple resonance NCACX/NCOCX experiments, magnetisation from either Cα or C′ to the adjacent carbons may be transferred by spin diffusion. ^{13}C magnetisation is stored along the z-axis while heteronuclear decoupling is switched off during the spin diffusion transfer, hence long mixing times (up to seconds) may be applied without sample heating. The Hamilton Operator that describes spin diffusion is given in equation 1-13 (24):

$$H = H_S + H_{SS} + H_{SI} + H_{II}.$$

equation 1-13

The transfer is mediated by the "proton bath" represented as spin I, via which carbons (represented as spins S) interact. H_{SS}, H_{II} and H_{SI} describe the homonuclear and heteronuclear dipolar interactions of the S and I spins, respectively. H_{SI} effectively provides the energy for homonuclear dipolar coupling, in an energy-conserving manner for S spins. H_S describes the evolution of the S spins. The effect of H_{SI} was demonstrated by Ueda and coworkers (25). The distance dependence is described by the transition probability in the dipolar coupled system and is inversely proportional to r^6, as propagated by Suter and Ernst (24). The rate of spin diffusion

Γ is given by equation 1-14 (26):

$$\Gamma = \frac{1}{r_{ab}^6} f\left(\tau_c\right).$$

equation 1-14

Here $f\left(\tau_c\right)$ represents a function containing the dipolar coupling of ^{13}C and ^{1}H, while r_{ab} represents the internuclear distance between two carbon nuclei a and b. Long-range couplings are driven by protons that mediate the energy conserving transfer.

1.4 Assignment strategy

1.4.1 General Remarks

Any investigation of a protein by NMR spectroscopy starts with the resonance assignment of the individual sites of the system. Assignment is the procedure in which for each individual spin in the protein its NMR chemical shift is determined. In solution NMR, the use of scalar couplings allows polarisation transfer along the backbone and to any particular side chain nucleus. A set of triple resonance experiments connecting the ^{1}H^N and ^{15}N with the neighbouring ^{13}C′, ^{13}Cα and ^{13}Cβ of either the same or the sequential amino acid permits an unambiguous sequential assignment. In solid-state NMR, dipolar couplings are widely used to transfer magnetisation after an initial ^{1}H-^{15}N and an ensuing ^{15}N-^{13}C CP. Dipolar couplings are "through space", in contrast to *J*-couplings which are "through-bond". Hence the selective transfer to a particular atom is difficult by using dipolar couplings. Unlike in solution NMR, transfer via dipolar couplings yields a complex mixture of signals from adjacent as well as nearest-neighbour residues. Especially for large biological systems with reduced spectral resolution this deficit may hamper the assignment procedure.

The development of 3D ^{15}N-^{13}C-^{13}C experiments and new polarisation transfer schemes for solid-state NMR helped to improve and increase the reliability of the assignment procedure (20). We applied 3D NCACX, NCACB, and NCOCX triple resonance experiments and 2D ^{13}C-^{13}C correlation experiments (20) (for details see

chapter 1.4.2). To further improve the reliability of the assignment, these experiments were used to record spectra on three differently labelled types of samples. In addition to a uniformly ^{13}C-^{15}N labelled sample, we used the 2-^{13}C- and 1,3-^{13}C-glycerol derived labelling pattern (for details see chapter 1.4.3) (2;3). The samples are named accordingly U-, 2G- and 1,3G-αB, respectively.

1.4.2 Pulse sequences for the sequential assignment

For the assignment, it is advantageous to record the data under MAS conditions, since in that case, the observed chemical shifts are the isotropic shifts which may be assigned by identification of the characteristic correlation patterns. In addition, the "Biological Magnetic Resonance Data Bank" provides reference chemical shifts for amino acids in proteins (27). Pauli *et al.* have assigned the 62-residue SH3 domain of α-spectrin by identification of characteristic amino acid side chain patterns from 2D ^{13}C-^{13}C correlation spectra, which were sequentially connected by comparing NCA(CX) and NCO(CX) experiments (20). NCA spectra contain only correlations of the backbone ^{15}N of residue i with its Cα(i). NCO spectra provide the corresponding sequential correlation by correlating the ^{15}N(i) with the adjacent CO(i-1) (chapter 1.3.3). These two experiments can be extended with an additional spin diffusion mixing scheme for polarisation transfer from Cα or C´ to the side chain carbons (NCACX and NCOCX). Usually, a mixing time of 25-35 ms is sufficient for intraresidual transfer to Cδ or Cϵ, if present, which is adequate to identify the amino acid type. Sequential crosspeaks between neighbouring amino acids in the peptide chain occur at a mixing time of 50 ms and longer. 3D NCACX spectra correlate each ^{15}N(i) in the F_1 dimension with Cα(i) in F_2, and subsequently with C´(i), Cβ(i), Cγ(i), Cδ(i), Cα(i-1, i+1) or C´(i-1, i+1) in F_3 (depending on the used spin diffusion mixing time, the labelling pattern and residue type). NCOCX spectra provide complementary sequential correlations of ^{15}N(i) in F_1 with C´(i-1) in F_2 and Cα(i-1), Cβ(i-1), Cγ(i-1), Cδ(i-1), C´(i, i-2) or Cα (i) in F_3. The pulse sequences for the assignment of αB are shown in figure 1-9:
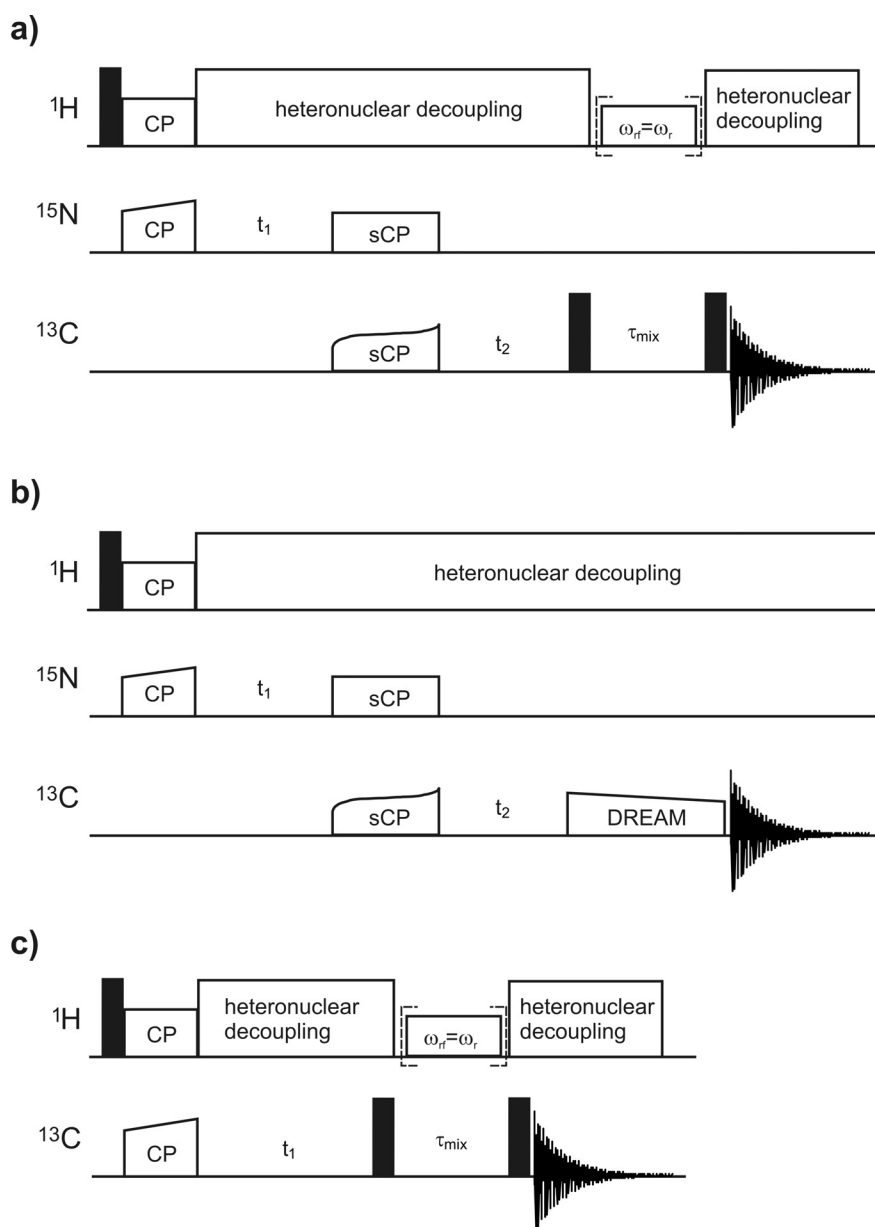
figure 1-9 In the 3D pulse sequences in figure 1-9a and b magnetisation is initially transferred from ^{1}H to ^{15}N by using a ramped CP followed by an indirect evolution time t_1 of ^{15}N. After t_1 the magnetisation is transferred by adiabatic or specific CP (sCP chapter 1.3.3) to either Cα or C´. Subsequently the polarisation is transferred from Cα or C´ to the side chain carbons by applying spin diffusion in figure 1-9a while the pulse sequence in figure 1-9b is used to transfer magnetisation from Cα to Cβ using BASE-DREAM recoupling(20). During spin diffusion mixing a field ω_{rf} equal to the MAS frequency ω_r on the ^{1}H-channel may be optionally applied to enhance the spin diffusion transfer (so called dipolar assisted rotational resonance mixing, DARR, in dashed brackets) (28). In figure 1-9c a proton driven spin diffusion pulse sequence for 2D ^{13}C-^{13}C is shown. After ^{1}H-^{13}C CP and indirect evolution of ^{13}C in t_1, spin diffusion mixing is applied for ^{13}C-^{13}C recoupling (optional DARR in dashed brackets). Filled bars represent 90° pulses. Heteronuclear ^{1}H decoupling is applied during sCP, DREAM mixing, indirect evolution and acquisition.

In the 3D pulse sequences in figure 1-9a and b magnetisation is initially transferred from ^{1}H to ^{15}N with a ramped CP. After evolution of ^{15}N signals during t_1, the magnetisation is transferred by adiabatic or specific CP (chapter 1.3.3) to either Cα or C′. Subsequently, the polarisation is transferred in figure 1-9a from Cα or C′ to the side chain carbons by allowing spin diffusion. During spin diffusion mixing, a field ω_{rf} equal to the MAS frequency ω_r on the ^{1}H-channel may be optionally applied to enhance the spin diffusion transfer (so called dipolar assisted rotational resonance mixing, DARR, in dashed brackets) (28).

In the pulse sequence in figure 1-9b the magnetisation is transferred from Cα to Cβ by using BASE-DREAM (20). During DREAM transfer a field ω_{rf} is applied on the ^{13}C channel for *frequency-selective* double-quantum transfer. The double-quantum transfer results in negative crosspeak intensities. The matching condition for the DREAM transfer (equation 1-15) contains the field ω_{rf} applied on the carbon channel and the chemical shift offset Ω_1 and Ω_2 of Cα and Cβ respectively; ω_r represents the MAS frequency and n an integer (1, 2, ...):

$$\sqrt{\omega_{rf}^2 + \Omega_1^2} + \sqrt{\omega_{rf}^2 + \Omega_2^2} = n\omega_r$$

equation 1-15

Finally a proton-driven spin diffusion (PDSD) pulse sequence for 2D ^{13}C-^{13}C is shown in figure 1-9c. After a ^{1}H-^{13}C CP and indirect evolution of ^{13}C signals during t_1, spin-diffusion mixing is applied for ^{13}C-^{13}C recoupling (or optional, dipolar-assisted rotational-resonance mixing [DARR], as indicated in dashed brackets).

1.4.3 Labelling strategy for the sequential assignment

In addition to an U-^{13}C-^{15}N labelled αB sample (U-αB), we used biosynthetically spin-diluted αB samples expressed from either 1,3-^{13}C-glycerol (1,3G-αB) or 2-^{13}C-glycerol (2G-αB) as the sole carbon source (2-4). In figure 1-10 the pattern of ^{13}C enrichment of the 20 amino acids is shown. The red colour corresponds to the degree of ^{13}C labelling obtained by growth on 2-^{13}C-glycerol (further referred to as

2G). The opposite labelling pattern which is obtained by growth on 1,3-^{13}C-glycerol is represented in green (further referred to as 1,3G).
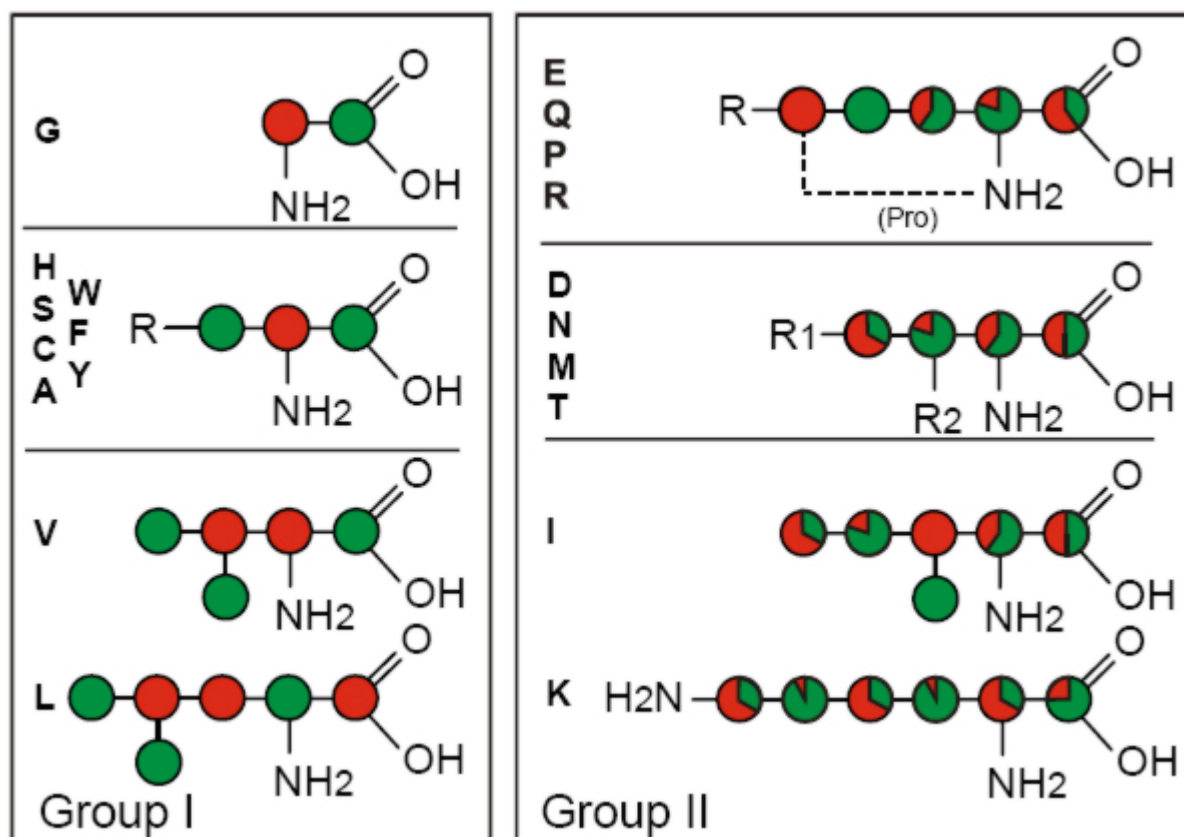
figure 1-10 Schematic representation of the effective ^{13}C enrichment of the individual amino acids, as obtained by protein expression in *E. coli*. The red colour corresponds to the degree of ^{13}C labelling obtained by growth on 2-^{13}C-glycerol while the opposite labelling pattern, obtained by growth on 1,3-^{13}C-glycerol is represented in green. Figure reproduced from reference (4).

The amino acid types are divided in two groups reflecting the biosynthetic pathway. Group 1 is synthesized from precursors of the glycolysis pathway and group 2 from precursors of the citrate cycle. In case of 2-^{13}C-glycerol labelling, only valine and leucine from group 1 have two chemically bonded ^{13}C-labelled carbons (in valine Cα-Cβ and in leucine Cβ-Cγ). Group 2 has a complex labelling scheme because the precursors pass the citrate cycle non-uniformly.

The advantage of the presented labelling pattern is a partial suppression of scalar couplings and the presence of fewer signals in the spectra resulting in a better resolution. Signals from congested regions of spectra recorded on U-^{13}C-^{15}N labelled samples can be resolved using spin dilution. Moreover, the assignment options are

reduced for the various amino acid types in the differently labelled samples. The benefits of this labelling strategy for the assignment procedure will be explained in chapter 4.1.2.1 by using examples from the αB assignment. For collecting structural distance restraints, the 1,3G- and 2G-samples help to identify interresidual correlations, since intraresidual correlations can be excluded for certain residues by alternating labelling.

Finally, an αB sample was prepared with ^{13}C and ^{15}N labelled proline, phenylalanine, tryptophane and methionine in. With this sample heterogeneity in αB was identified (see chapter 4.2).

1.5 The value of the chemical shifts

1.5.1 Chemical shift changes highlight intermolecular interaction sites

Chemical shifts are very sensitive to changes in the electronic environment of nuclei. In particular the ^{15}N resonances of amino acids in proteins exhibit site specific changes upon ligand binding. This effect is widely used in solution NMR to follow chemical shift changes in ^{1}H-^{15}N HSQC spectra when the sample is titrated with a ligand. The interaction site on the protein can then be identified. This so called "structure activity relationship" by NMR (SAR by NMR) can be used to screen a pool of ligands to study their binding to a molecule of interest or to identify the interaction sites of protein-protein, protein-DNA or protein-RNA complexes (8;29;30). As another example, crystal contact areas were determined for a microcrystalline preparation of SH3 and Metalloproteinase 12 by comparing chemical shifts from solution NMR with chemical shifts derived from MAS solid-state NMR (31;32).

1.5.2 Secondary structure from chemical shifts

The chemical shift Hamilton operator contains a term describing the shielding of a spin by its surrounding electrons. In principal the shielding of an ensemble of backbone nuclei allows a complete reconstruction of the 3D structure in which these nuclei are assembled. But since the shielding tensor is dependent on numerous

details and still not completely understood, the chemical shifts alone are presently not sufficient to calculate a structure of a protein *ab initio*. However, secondary chemical shifts, empirical methods and a combination of these are used to access structural information. The secondary chemical shift is defined as the difference between the random coil value and the measured isotropic chemical shift. As shown in figure 1-11 the secondary chemical shifts of Cα and Cβ atoms of amino acids in proteins are strongly dependent on the backbone conformation.
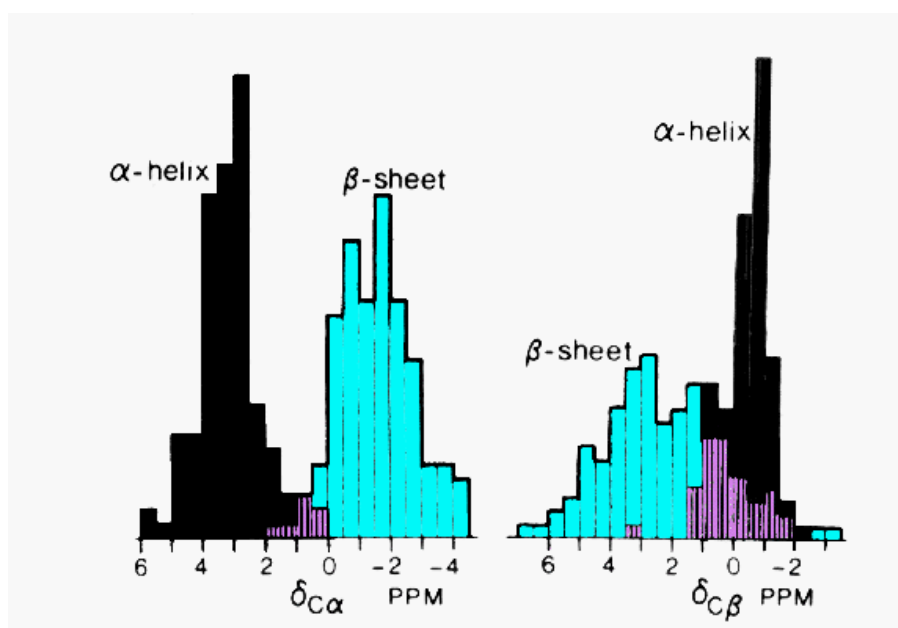
figure 1-11 Secondary shift distribution for Cα and Cβ in α-helical and β-sheet conformation are shown. Figure adapted from reference (33).

This approach has been extended by Cornilescu *et al.* by generating a database with protein structures and the associated chemical shifts (33). Their software, TALOS, searches the database for ten triplets of amino acids matching the amino acid type and secondary chemical shifts and outputs the average dihedral ϕ and ψ angles for the query sequence. Another approach ("PECAN", Protein Energetic Conformational Analysis from NMR chemical shifts) uses information from the amino acid sequence of a protein and available NMR chemical shifts for the purposes of identifying low energy conformations and secondary structure elements (34).

1.6 Structure Calculation from NMR data

The ultimate goal of structure calculation is to search for an atomic model that simultaneously fulfils measured structural restraints while maintaining a chemically allowed structure as determined from *a priori* chemical information (i.e. the force field, defining bond lengths, bond angles, improper angles and non-bonded interactions). Structural restraints that can be employed include experimentally determined distance restraints, dihedral angle restraints, hydrogen bonds and disulphide bridges. Most algorithms used for structure calculation aim to find the global minimum of a hybrid energy function (target function E), which represents *a priori* chemical knowledge and experimental data (the structural restraints). Cartesian or torsion-angle molecular dynamics may be used in a simulated annealing protocol. The global energy minimum is searched by raising and lowering the temperature of the system during a molecular dynamics simulation. The parameter "temperature" in this context has no physical meaning, but is a measure of the probability of the macromolecule to cross an energy barrier (i.e., its kinetic energy) (35). During cooling, the motion of the atoms decreases slowly unless an energy minimum is reached.

Cartesian molecular dynamics allows motions of atoms in all three spatial directions. Torsion angle molecular dynamics uses fixed bond-lengths, and only allows changes of the torsion-angles (i.e. CNS, XPLOR-NIH (36)). In all calculations presented in this thesis, torsion-angle molecular dynamics were used.

A recent approach allows the use of unassigned crosspeaks extracted as ambiguous distance restraints from solid-state NMR 2D/3D spectra (37). The software SolARIA assigns a set of crosspeaks in an automated fashion, to generate a set of distance restraints for further structure calculations. This operation is repeated iteratively to minimise the energy of a target function. In chapter 4.3, the latter strategy will be employed for structure calculation of the α-crystallin domain in αB.

1.7 The outer membrane protein G of E. coli (OmpG)

The outer membrane of Gram-negative bacteria acts as a semipermeable shell for small molecules. Escherichia coli is a Gram-negative bacterium containing several outer-membrane proteins. These so-called porins form channels that allow nutrients

to enter the cell (38;39). E. coli mutants lacking OmpF and LamB are expressing the porin OmpG (40). Thus, OmpG helps these porin-deficient mutants to grow on maltodextrin-containing media. OmpG consists of 280 amino acids and has the typical attributes of a porin: (i), CD-spectroscopy indicates a β-sheet fold (41), (ii), the last residue is a phenylalanine, which is important for the insertion of the protein into the outer membrane and (iii), OmpG serves as a nonselective channel for mono-, di-, and trisaccharides, as was shown by a proteoliposome swelling assay (41). Chemical cross-linking has shown that OmpG is monomeric (41), unlike the trimeric porins (e.g. OmpF, OmpC, PhoE and LamB). Conlan and coworkers demonstrated on in vitro refolded OmpG that it is a functional monomer in planar bilayers (42). Cryomicroscopy with 2D crystals of OmpG revealed a projection structure at 6 Å resolution showing a monomeric channel restricted by internal loops (43). Recently, the crystal structure of OmpG in 3D crystals was solved (44). Dependent on the pH, the internal loops adopt an open and close conformation. The diameter of the barrel is ~25 Å, consisting of 14 membrane-spanning β-strands. Unlike X-ray crystallography, solid-state NMR can be used to study OmpG in a natural lipid-bilayer (2D crystals)(45). OmpG is used to develop techniques to enhance the resolution of MAS solid-state NMR spectra of membrane proteins (see chapter 3).

1.8 αB-crystallin: a small heat shock protein (sHSP)

An internal control system in cells prevents missfolded proteins from aggregation and subjects them to proteolysis after ubiquitination. Failures lead to Alzheimer's, Parkinson's and related diseases. In this context, chaperones are integral to directing proteins along the proper folding pathway. They are assisted by small heat shock proteins (sHSP) which are ATP-independent molecular chaperones that act as protective reservoirs for missfolded/unfolded proteins awaiting folding by ATP-dependent chaperones (46). The level of sHSPs rises to 1% of the total cellular protein level under conditions of stress, such as changes in pH or temperature. There is no direct evidence for how sHSPs function, i.e. how they protect missfolded proteins from aggregation. It is commonly postulated that sHSPs either undergo a conformational change or dissociate into oligomers of lower molecular weight or monomers before binding denatured proteins (46).

αB-crystallin (αB) is a small heat shock protein and serves several functions in mammalian organisms. αB was first discovered in the crystalline eye-lens, where it makes up 50% of the proteins together with its homologue αA crystallin (αA) by forming polydisperse α-crystallin hetero oligomers (47). The structural integrity of α-crystallin is essential for maintaining the transparency of the eye-lens. In other tissues αB forms homo oligomers and aggregates with other sHSPs, like sHSP25 (48) and it acts as an essential cofactor for an E3-ligase (47;49;50).

For this reason, αB is also found overexpressed in cardiac, brain, and skeletal muscle cells, playing special roles, while αA-crystallin is primarily found in the lens. During cardiac ischemia, αB translocates from the cytoplasm to cardiac myofibrils where it associates with proteins such as titin, actin, and desmin. There is a pH drop of 7.5 to 6.9 during cardiac ischemia and the association of αB with titin is found to be stronger at lower pH. A mutation in αB, R120G, is found to be associated with both the occurrence of cataracts in eye lenses and desmin-related cardiomyopathy, DRM, a disease characterized by aggregates of desmin and R120G αB. In a more recent, exciting development, researchers at the University of Stanford announced that αB helped to reduce Multiple Sclerosis (MS) symptoms in mice. Furthermore, αB was found to be an essential cofactor for FBX4 required in the polyubiquitination of Cyclin D1.

Since αB forms a polydisperse oligomer of ~580 kDa built from 20 kDa (175 aa) subunits, its structure has eluded structural biologists for decades. This is a common theme for most sHSPs since they form polydisperse oligomers, and only the structures of three sHSPs which form discrete oligomers, MjHSP16.5 (from the archaea *Methanococcus jannaschii*), wHSP16.9 (from wheat), and Tsp36 (from beef tapeworm) have been solved to date. With MAS solid-state NMR it is possible to assign spectra of the heterogeneous full length αB oligomers in the solid-state and to elucidate structural details.

1.8.1 Domain architecture of αB-crystallin

A common feature of the sHSPs is the α-crystallin domain, that is flanked by N-terminal and C-terminal domains (46) (figure 1-12):

N-terminus   **α-crystallin-domain**   **C-terminus**

1   **64** ←——————→ **152** **175**
                   αB10.1

58 ←————————————————→ **157**
                αB11.4
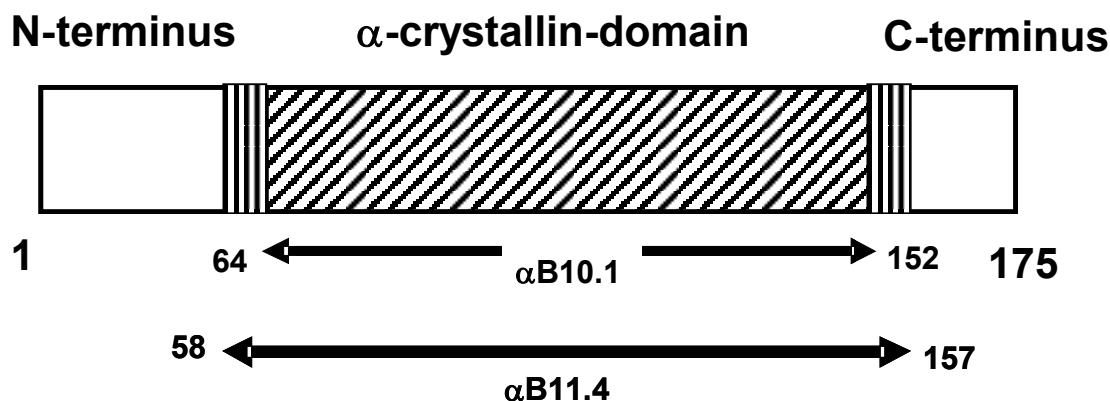
figure 1-12 The domain architecture of αB crystallin is shown. The protein consists of a α-crystallin domain, that is flanked by a N- and C-terminal domain. The protein constructs αB10.1 and αB11.4 that are lacking the N- and C-terminal domains are forming dimers but do not form higher order oligomers.

In αB, the N-terminal domain contains an internal duplication with residues 1-34 homologous to residues 35-63 (48). αB protein constructs lacking the first 34 residues still form polydisperse oligomers of the typical size of ~580 kDa. However, upon removal of the complete N-terminus, the α-crystallin domain forms dimers and a small amount of tetramers, whereas the N-terminus alone lacking the α-crystallin domain and the C-terminus aggregates unspecifically (48). Two constructs lacking both N- and C-termini, αB10.1 and αB11.4 are known to form exclusively dimers in solution under given conditions (51;52). Consequently, the C-terminal domain interacts with the α-crystallin domain (53) and may be involved in the formation of tetramers (48) and the N-terminus is involved in the formation of higher order oligomers. Carver *et al.* could assign sequentially ^{1}H resonances by solution NMR of the residues 1-5 and 166-175 from the N- and C-terminus of αB oligomers. The oligomer size of ~580 kDa makes the use of solution NMR impractical to study the protein core. Obviously, the first (1-5) and last (166-175) residues must be flexible and have a high degree of mobility that is essentially independent of the protein core spanning residues 5-166 (54).

Full length αB binds with unfolded proteins to form larger oligomers than the original ones (46;55). However, the underlying mechanism is not completely understood. αB10.1 and αB11.4 show substantial chaperone activity without forming

oligomers (51;52). Hence, the chaperone activity of αB is independent of the termini. Recently, a mini-αB was found comprising the residues 73-92 preventing unfolded proteins from aggregation and precipitation (56). Whether dissociation is necessary for the binding of denatured proteins or not is unknown (46). However, cross linked αB oligomers incapable of dissociation also show the capability to prevent aggregation of missfolded proteins (57).

The αB10.1 construct is stable and shows a favourable resolution in solution NMR spectra (52). In **chapter 4.2.4** chemical shifts derived from the αB oligomer will be compared with shifts derived from solution NMR of the αB10.1 (kindly provided by Ponni Rajagopal (52)). Differences are expected for residues that are involved in intermolecular contacts (32).

1.8.2 Related structures of sHSPs and the EM structure of αB

1.8.2.1 Related structures of sHSPs

The organization as large oligomers and the polydispersity of sHSPs has impeded a structure determination by solution NMR and X-ray crystallography to date. However three structures of sHSPs forming discrete oligomers have been solved.

The eukaryotic wheat sHSP16.9 (wHSP16.9) forms a 12-mer hexameric double disc (58) and the prokaryotic sHSP16.5 of *Methanococcus jannashii* (MjHSP16.5) forms a 24-mer hollow sphere (59). Both oligomers are build from tetramers of dimers and the α-crystallin domain is folded as a β-sheet sandwich which is involved in the dimerization by domain swapping. In both wHSP16.9 and MjHSP16.5, β6 of the interdomain loop and β2 are involved in intermolecular strand sharing and are a part of the dimer interface (figure 1-13). In MjHSP16.5, the β6 strand is additionally stabilized by a β1 strand.
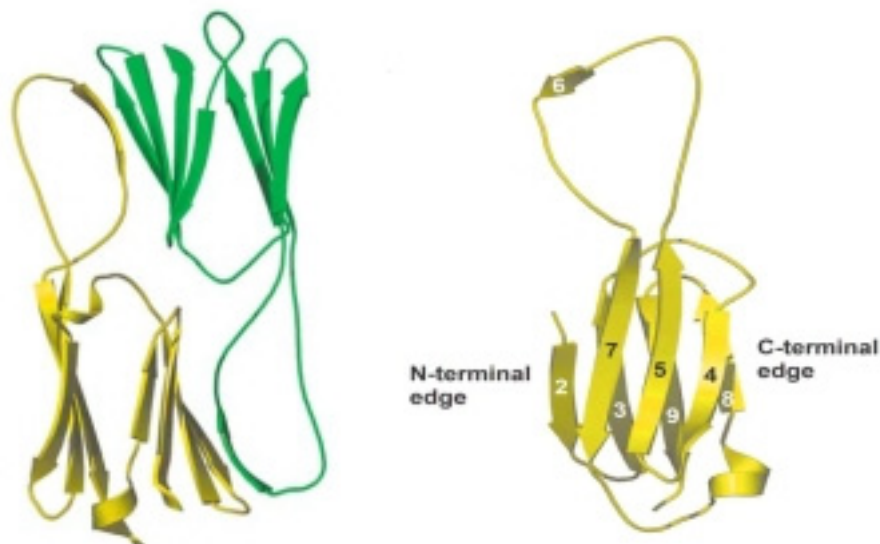
figure 1-13 The left panel shows the α-crystallin domains of the wHSP16.9 dimer with the intersubunit composite β-sheet. On the right a ribbon diagram of the α-crystallin domain of the wHSP16.9 monomer is shown with the conventional strand numbering (58). The figure is adapted from reference (58).

The C-terminal domains of wHSP16.9 and MjHSP16.5 form an intermolecular β-strand (β10) for the assembly of tetramers. The N-terminal domain of wHSP16.9 of each second monomer is folded and involved in intermolecular contacts to other N-terminal domains and the α-crystallin domain. On the other hand, the N-terminal domain of MjHSP16.5 is disordered (58).

In figure 1-14 the sequences of αB, αA, TSP36 (HSP36 of the beef-tapeworm), wHSP16.9 and MjHSP16.5 are aligned. The α-crystallin domain shares ~25% sequence homology between MjHSP16.5, wHSP16.9 and eukaryotic sHSPs while the similarity raises up to 60% among eukaryotic sHSPs. However, the N- and C-termini are poorly conserved throughout the species. The α-crystallin domain reveals multiple conserved residues among the species. However, the interdomain loop (blue coloured in figure 1-14) is much shorter in αB, αA and TSP36, compared with wHSP16.9 and MjHSP16.5. Consequently, the dimer assembly of the α-crystallins must be different.

```
                             20                  40                  60
human αB crystallin MDIAIHHPWIRR-----PFFPFHSPSRLFDQFFGEHLLESDLFP-TSTSLSPFYLRPPSF
human αA crystallin MDVTIQHPWFKR--TLGPFY----PSRLFDQFFGEGLFEYDLLPFLSSTISPYY-R-QSL
Tsp36 (104-186)     ----------------------------------------------------------
wHSP16.9            ------MSIVRRTNVFDPFAD------LWADPFDTFRSIVPAISGGGSE-----------
MJHSP16.5           --------MFGR----DPFDS------LFERMFKEFFATPM----TGTTMIQSSTG----


                             80                 100                 120
human αB crystallin LRAPS-WFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGD-VIEVHGKHEERQDEHG
human αA crystallin FRT---VLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQDD-FVEIHGKHNERQDDHG
Tsp36 (104-186)     ---------YEV-GKDGRLHFKVYFNVKNFKA-EEITIKA-DKNKLVVRAQKSVACGDAA
wHSP16.9            ---TAAFANARMDWKE-TPEAHVFKADLPGVKKEEVKVEVEDGNVLVVSGERTKEKEDKN
MJHSP16.5           -IQISGKGFMPISIIEGDQHIKVIAWLPGVNKEDIILNAVGD--TLEIRAKRSPLMITES
                      β1             β2           β3             β4         β5
                             140                160                 180
human αB crystallin -------F-ISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQ----VSGPERTIP
human αA crystallin -------Y-ISREFHRRYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHAERAIP
Tsp36 (104-186)     -------MSES--VGRSIPLPPSVDRNHIQATITTDDVLVIEAP----------------
wHSP16.9            DKWHRVERSSG-KFVRRFRLLEDAKVEEVKAGLEN-GVLTVTVPKAE---VKKPEVKAIQ
MJHSP16.5           ERIIYSEIPEEEEIYRTIKLPATVKEENASAKFEN-GVLSVILPKAE-----SSIKKGIN
                      β6             β7               β8        β9              β10
                             195
human αB crystallin ITREEKPAVTAAPKK
human αA crystallin VSREEKP--TSAPSS
Tsp36 (104-186)     ---------------
wHSP16.9            ISG------------
MJHSP16.5           IE-------------
```

figure 1-14 Alignment of human αB and αA-crystallin amino acid sequences with those of beef tapeworm HSP36 (TSP36), wheat HSP16.9 and *Methanococcus jannashii* HSP16.5. Conserved residues are printed in bold. The β-strands of the α-crystallin domain are shaded and numbered in the sequences of known structure. The numbering includes the gaps. From TSP36 only the first α-crystallin domain in the sequence is aligned spanning from residue 104-186.

Since TSP36 has a short interdomain loop alike the α-crystallins, this may suggest a similar assembly. However, TSP36 contains two α-crystallin domains in one peptide

sequence. The first ranges from residue 104 to 186, the second from 236 to 311. Furthermore, the N-terminal domain consists of more than 100 amino acids and the linker between the two α-crystallin domains contains additional ~50 residues. Hence, except for the α-crystallin domain, the homology between the two α-crystallins and TSP36 is weak. The C-terminal domain contains a conserved IXI/V motif throughout the aligned species, except for TSP36. In wHSP16.9 and MjHSP16.5 this motif is involved in the assembly of tetramers and binds into the groove between β4 and β8. However, the orientation varies among the species. In conclusion, the sHSPs show a reasonable sequence homology for the α-crystallin domain and a conserved IXI/V motif, in the C-terminal domain. Low homology is observed for the N-terminal domain and large parts of the C-termini.

1.8.2.2 EM structure of αB

Haley *et al.* (1998) obtained the EM structure of human αB at 3.6 nm resolution (60). The structure indicates a hollow sphere with variable quaternary structure for the αB oligomer (figure 1-15):
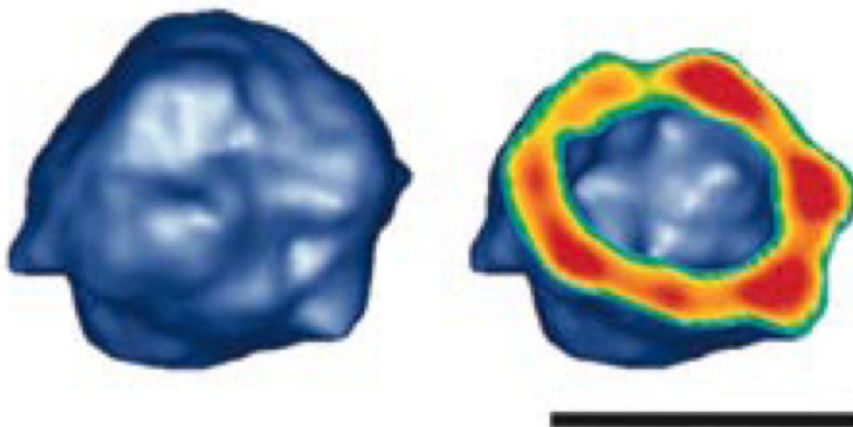
figure 1-15 Structure of αB oligomers at 3.6 nm resolution (approximately 32 subunit species) solved by cryo-electronmicroscopy (60). The structure indicates a hollow sphere with variable quaternary structure for the αB oligomer. The scale bar represents 10 nm. The figure is adapted from reference (60).

The study revealed that human αB oligomers exhibit a range of diameters, 8 to 18

nm (60). Gel-filtration chromatography confirmed a polydispersity in molecular mass (510-790 kDa). To compensate for heterogeneity, they selected image data of the subspecies assembled from 32 monomers for 3D image reconstruction. But still the 3D variance and structural divergence is significant within the protein shell. To understand the assembly and function of αB oligomers, structural data at an improved resolution are needed. However, crystallization attempts were unsuccessful so far, and solution NMR methods allowed to observe only the first five and the last ten residues. MAS solid-state NMR can help collecting structural restraints of αB, because neither high quality single-crystals nor rapid molecular tumbling is a prerequisite for this technique. In chapter 4.1, the sample preparation and resonance assignment for αB will be described.

1.9 Overview of the thesis

In the main part of this thesis, structural features of αB-crystallin oligomers are elucidated using solid-state NMR methodology. The combination of solid-state NMR and solution NMR delivers insight into the interaction of monomers in the oligomer. For large systems, the high number of signals leads to overlap and hampers the assignment process. To overcome this hurdle, methods for resolution enhancement are presented in this thesis.

Each structure determination process using NMR spectroscopy starts with the resonance assignment. Spectral editing is one way to reduce the number of signals, to ease the assignment process. The selection of specific moieties may be achieved by virtue of their unique chemical, spectral and/or conformational properties. A filtering block that selects methyl resonances will be described in **chapter 3.1**. Methyl filtering is achieved by utilising different CP-dynamics of CH-, CH_2- and CH_3-groups. The filter is implemented in 2D pulse schemes which are applied to the SH3-domain as a model system. The resolution enhancement coming along with this filter is further demonstrated on a preparation of OmpG. Another possibility to enhance the resolution in solid-state NMR spectra is to remove J-couplings. In **chapter 3.2**, the use of J-decoupling is demonstrated using a microcrystalline SH3-domain and 2D crystals of OmpG in lipid bilayers. A combination of refocusing of J-couplings during the indirect evolution t_1 and post-acquisition data processing in t_2 is carried out to

remove *J*-couplings in 2D ^{13}C-^{13}C correlation spectra.

      In **chapter 4.1**, the MAS solid-state NMR assignment of αB-crystallin is described. We use for this purpose multidimensional pulse schemes, methods for resolution enhancement, 2-^{13}C-glycerol/1,3-^{13}C-glycerol labelled samples (2) and site-specific amino acid labelling. The obtained chemical shifts are further analysed in **chapter 4.2**. The Information contained in the chemical shifts are used via the software TALOS (33) to obtain dihedral angles. We compare the dihedral angles of αB-crystallin with the secondary structure from related proteins. Moreover, a method is presented to define the topology of strands and to distinguish between intra- and intermolecular strands in β-sheets. In addition, solution NMR shifts of αB10.1 are compared with the solid-state shifts of the oligomer to identify systematic differences which are an evidence for intermolecular interactions. In **chapter 4.3**, structural features of αB-crystallin oligomers at atomic resolution are shown that are based on experimental restraints obtained on the full-length protein. Distance restraints are collected from a set of 3D triple resonance spectra and used in structure calculations, yielding a β-sandwich fold. In this context, recently developed structure calculation techniques (37) are for the first time applied on a protein of which the structure is unknown. Finally, first results from the R120G mutant are discussed with the structure at hand.

2. Materials and Methods

2.1.1 Sample preparation

αB was expressed and purified as described elsewhere (Muchowski et al., 1999) with some modifications to the growth protocol (51;61). BL21 Star (DE3) cells (Novagen, Milwaukee, WI) were used instead of BL21 (DE3) cells. They were grown in M9 minimal media supplemented with ^{15}NH4Cl (1 g/L) and $^{13}C_6$-glucose (2 g/L). Dilutely labelled ^{13}C samples were overexpressed in bacteria grown in M9 media supplemented with either 2-^{13}C-glycerol or 1,3-^{13}C-glycerol (2 g/L). The samples are named accordingly U-, 2G- and 1,3G-αB, respectively. Cells were incubated at 37 °C until a cell density of 0.4 at 600 nm was reached. The cultures were then transferred to a 22 °C shaker and protein expression was induced with 1 mM IPTG. Cells were harvested after 12-16 hrs. The PFMW-αB sample has proline, phenylalanine, methionine and tryptophane ^{13}C, ^{15}N labelled. The sample was overexpressed in bacteria growing on M9 minimal medium containing 100 mg/L of each labelled and unlabelled amino acid, 2 g of glucose and 0.5 g of NH4Cl per litre of culture, according to the protocol for amino acid selective labelling of Hiller *et al*. (62). This sample is further referred to as PFMW-αB. The αB sample consisting of 1,3-^{13}C-glycerol/^{14}N labelled monomers and ^{15}N labelled monomers was prepared by mixing the differently labelled samples 1:1 and subsequent denaturing with 6M guanidinium chloride. The unfolded monomers were refolded in a desalting column to yield differently labelled αB oligomers.

Lyophilised full length αB was dissolved in 50mM sodium phosphate buffer (pH 7.6) containing 50 mM NaCl and 2 mM NaN$_3$ to a final concentration of 40mg/ml. 0.5 ml (20 mg) of the αB solution was mixed with ~80-100 μl of a 50% PEG8000 solution until the start of precipitation. The precipitation was complete after 3 days in a vapour diffusion chamber over a reservoir solution with 50% PEG8000. A ZrO$_2$ rotor was filled with the precipitant and used for solid-state NMR experiments.

All column chromatography was performed on the AKTA system. ^{15}N NH$_4$Cl and ^{13}C-glucose were obtained from CIL (Cambridge Isotope Labs). 1,3-^{13}C-glycerol and 2-^{13}C-glycerol were purchased from Cambridge Isotope Laboratories. Protease

inhibitor cocktail for non His-tagged proteins and protease inhibitors such as Pefabloc and PMSF were purchased from SIGMA.

The plasmid for the R120G mutant was kindly provided by Ponni Rajagopal in the group of Rachel Klevit. The expression and purification for the R120G mutant is equivalent to the wild type.

The 1,3G-OmpG sample was expressed using 1,3-^{13}C-glycerol (2 g/L culture) as the carbon source and ^{15}NH$_4$Cl (0.5 g/L culture) as the nitrogen source. The 2G-OmpG and the 2G-SH3 were prepared from minimal medium containing 2-^{13}C-glycerol (2 g/L culture) as the sole carbon source as described in the literature (3;45). According to this conditions, only Valine and Leucine show a pattern with directly bounded ^{13}C-atoms, e.g. Cα-Cβ for Valine and Cβ-Cγ for Leucine. The uniformly ^{13}C and ^{15}N labelled α-spectrin SH3 domain is referred to as U-SH3 and was prepared as described elsewhere (63). For the solid-state cross-polarisation (CP) MAS NMR correlation experiments, samples containing typically ~ 1.4 µmol (10 mg) U-SH3 domain and ~ 0.6 µmol (20 mg) 1,3G-OmpG were used.

2.1.2 Solid-state NMR spectra recorded of αB

All solid-state NMR spectra on αB were recorded at 270 K and a MAS frequency of $\omega_R/2\pi$ = 10.0 kHz. The 2D experiments were acquired at a field of 16.4 T on a wide-bore Bruker Avance 700 spectrometer. The 3D experiments on 2G- and 1,3G-αB were acquired at a field of 14.1 T on a wide-bore Bruker Avance 600 spectrometer. The 3D spectra on U-αB were recorded on a wide-bore Bruker Avance 700 spectrometer operating at 16.4 T. The wide-bore 600 and 700 spectrometers were equipped with a 4 mm triple-resonance CP/MAS probe (Bruker, Karlsruhe, Germany).

To establish sequential assignments and collect structural restraints, triple resonance 3D NCACX, 3D NCOCX (3) and 3D NCACB (20) spectra were recorded on U-, 2G-, and 1,3G-αB. NCACX and NCACB spectra were acquired with 48 scans as a data matrix of 1024 x 64 x 40 with acquisition times of 12 ms, 5.6 ms and 6 ms for ^{13}C, ^{13}Cα/^{13}C´ and ^{15}N, respectively. NCOCX spectra were acquired with 48 scans as data matrix of 1024 x 32 x 40 with acquisition times of 12 ms, 5.6 ms and 6 ms, respectively. The smaller spectral width of C´ enabled acquisition with fewer

increments without a loss in resolution. For U-αB, an NCOCX spectrum with 35 ms DARR (28) mixing and an NCACX with 25 ms PDSD (24) mixing was recorded. On 2G- and 1,3G-αB, NCOCX spectra with 75 ms DARR- and 150, 200 and 500 ms PDSD mixing were recorded. NCACX with 60 ms DARR- and 200 and 500 ms PDSD mixing were recorded. For U-αB, an NCACB spectrum with a 2 ms DREAM mixing (20) time was recorded. DREAM mixing was established with a weak ^{13}C B$_1$ field that was linearly amplitude-modulated from 6 to 3 kHz. The carrier frequency was centred at 45 ppm between Cα and Cβ. For DARR mixing (28), a B$_1$ field on protons equal to ω_R was used. To establish ^{1}H-^{15}N CP, a 100-75% ramp on ^{1}H was used. The B$_1$ fields during CP were ~50 kHz and ~60 kHz for ^{1}H and ^{15}N, respectively. Magnetisation transfer between the low-γ nuclei ^{15}N and ^{13}C was established with a 3.5 ms adiabatic CP using a field of 25 kHz and 35 kHz respectively. 2D PDSD spectra were recorded with a 15 ms mixing time for U-αB and 50, 75 and 500 ms for 2-G-αB and 1,3G-αB.

The 2D CHHC spectrum was recorded with an initial CP contact time of 1.5 ms. The second and third CP were each 50 µs long to guarantee for one bond ^{1}H-^{13}C transfer. 40 µs ^{1}H spin diffusion mixing was used to avoid recoupling of more distant protons than these attached to Cα of residues i and j. For CP, a field of 60 kHz and 70 kHz was used on the ^{1}H and ^{13}C channel, respectively. 384 transients were recorded in t$_1$ with a total indirect acquisition time of ~8 ms. The direct acquisition time was 15 ms.

For detection of intermolecular restraints a 2D ^{15}N-^{13}C correlation spectrum was recorded from an αB sample consisting of 1,3-^{13}C-glycerol/^{14}N labelled monomers and ^{15}N labelled monomers. For ^{13}C-^{15}N exchange, a TEDOR transfer scheme was used. REDOR (23) mixing times of 3 and 6 ms were applied in these studies, to achieve transfer over distances from ~ 4 Å (3ms) up to about 5-6 Å for the longest mixing times[1]. Typical ^{15}N π–pulse lengths during the REDOR mixing were 12 µs. After t$_1$-evolution of ^{15}N, a z-filter was used to retain rotor synchronisation by compensating for pulse durations and to remove transverse spin components (64). During t$_1$-evolution, acquisition and REDOR-mixing, heteronuclear decoupling at ~ 87 kHz was performed. The t$_1$-increment was chosen to be 1/ω_R = 100 µs. The total indirect acquisition time amounted to 4 ms.

---

[1] The acronym "REDOR" is used to refer to the mixing blocks and the mixing times, whilst the acronym "TEDOR" is used for the 2D-experiment where two REDOR mixing-blocks are sandwiched by an INEPT step.

Typical π/2-pulses on ^{13}C and ^{1}H were 3.5-4.3 µs and 3.8 µs, respectively. SPINAL64 decoupling of ^{1}H spins with 87 kHz was applied in all 2D and 3D experiments.

The 2D INEPT experiment was optimised for [1]H-^{13}C transfer in methyl-groups. Hence, a delay of 3.125 ms, corresponding to J_{CH3} = 60 Hz was chosen.

Typical π/2-pulse length on ^{13}C and ^{1}H were 3.5-4.3 µs and 3.8 µs, respectively. The J-decoupled PDSD was recorded using States-TPPI for phase sensitive detection in the indirect dimension; all other spectra were recorded using TPPI. SPINAL64 (65) decoupling of ^{1}H spins with 87 kHz was applied in all 2D and 3D experiments. The 2D spectra were zero filled to 2048 points in each dimension.

2.1.3 Methyl-filtered spectra

All methyl-filtered solid-state NMR spectra of OmpG and SH3 were recorded at a MAS frequency $\omega_R/2\pi$ = 8.0 kHz. The 2D [13]C-^{13}C and [13]C-^{15}N dipolar correlation experiments of OmpG and SH3 were acquired at 280 K (αB 270K), at a field of 9.4 T on a wide-bore Avance-400 spectrometer (Bruker, Karlsruhe, Germany). The spectrometer was equipped with a 4 mm triple-resonance CP/MAS probe (Bruker, Karlsruhe, Germany). The methyl-filtered DARR experiment of αB was recorded at $\omega_R/2\pi$ = 12.5 kHz at a field of ~21 T and the methyl filtered REDOR on αB at $\omega_R/2\pi$ = 10 kHz at a field of 16.4 T (Bruker Avance 600)..

A ramped CP (75% - 100%) from ^{1}H to ^{13}C was applied, followed by a squared phase-inversion CP at the n = 1 Hartmann-Hahn condition. This polarisation inversion (PI) period cancelled the CH-signals, whereas CH_2 groups were inverted to approximately $-1/3$ of their initial intensity. The PI time was determined experimentally. It was found that 50 µs lead to the desired signal inversion. Following PI, the CH_2 and residual CH signals were zeroed by applying a short (60 µs) dipolar-dephasing period. The methyl (and non-protonated) carbons were weaker coupled to ^{1}H and were not dephased. The carbon chemical shift evolution was refocused by a π–pulse in the middle of the dephasing period.

For the filtered [13]C-^{13}C homonuclear correlation experiment on SH3 and αB , a 70 ms dipolar-assisted rotational resonance (DARR) (28) mixing period was used. For the [13]C-^{15}N exchange, a TEDOR transfer scheme was introduced (66;67) after selection of methyl carbons by CPPI and dipolar dephasing. REDOR (23) mixing times of 2, 3,

6 or 8 ms were applied in these studies, to achieve transfer over distances from ~ 2.5 Å (short mixing time) up to about 5 Å for the longest mixing times[1]. Typical ^{15}N π–pulse lengths during the REDOR mixing were 12 µs for the SH3 domain and 13.3 µs for OmpG and αB. After t_1-evolution of ^{15}N, a z-filter was used to retain rotor synchronization by compensating for pulse durations and to remove transverse spin components (64). During t_1-evolution, acquisition and REDOR-mixing, two-pulse phase-modulation (TPPM) decoupling at ~ 75 kHz (SH3 domain) or ~ 87 kHz (OmpG, αB) were used for heteronuclear decoupling (68). The t_1-increment was chosen to $1/\omega_R$ = 125 µs (αB 100 µs). The total indirect acquisition time amounted to 6 ms.

2.1.4 *J*-decoupled spectra

All experiments on the 2G-SH3 sample were recorded on a Bruker Avance 400 spectrometer equipped with a 4mm triple resonance MAS probe (Bruker, Karlsruhe). The MAS frequency was set to 8 kHz (10 kHz for OmpG and αB) and heteronuclear decoupling was performed using TPPM, with a nutation frequency of 70 kHz (86 kHz for OmpG and αB). The reference spectrum of the 2G-SH3 domain was recorded using the same conditions as described above but without refocusing pulses in t_1. The spectra of OmpG and αB were recorded on a Bruker Avance 750 MHz supplied with a 4mm double resonance probe using similar conditions.

The B_1 fields during ^1H-^{13}C CP (1.5ms) in 2D experiments were ~55 kHz and ~65 kHz for ^1H and ^{13}C, respectively. A 50-100% ramp on the ^{13}C channel was used for CP. 2D ^{13}C-^{13}C spectra were recorded as a data matrix of 1024 x 768 points with acquisition times of 15 ms and 10 ms in the direct and indirect dimension, respectively.

For homonuclear *J*-decoupling during t_1, a selective Gaussian-shaped π-pulse was used with a truncation level of 3%. The duration of the selective pulse was set to 350 µs (17.6 T) or 1 ms (9.4 T) to obtain a flip angle of 180° on resonance. Phase-sensitive data acquisition in t_1 was performed using States-TPPI. The following calibration scheme was used: CP-hard($\pi/2$)-soft(π)-hard($\pi/2$) with simultaneously heteronuclear decoupling of the proton spins using TPPM.

2.1.5 Data processing and analysis

The *J*-deconvolution by MaxEnt in the direct dimension was performed, using the RowlandNMRToolkit version 3.0 (http://structbio.uchc.edu/hochlab_files/rnmrtk.html). Chemical shifts of ^{15}N, Cα, Cβ and C′ resonances were used as input for PECAN (34) for secondary structure determination. The solution-state NMR chemical shifts of Cα and Cβ were corrected for ^{2}H isotope shift before being used as input. Backbone dihedral angles were obtained from TALOS (33) using ^{15}N, Cα, Cβ and C′ resonances.

The *J*-coupling constant for MaxEnt reconstruction was set to 35Hz. The 2D spectra were zero filled to 2048 or 4096 points in each dimension. All 3D spectra were zero filled to a data matrix of 2048 x 256 x 256 points. For all spectra, linear prediction in the indirect dimensions was used to increase the number of points by a factor of 1.3. Gaussian and squared cosine filters were used in the direct and indirect dimensions, respectively. Data processing was performed using XWINNMR or Topspin 1.3/2.0 (Bruker, Karlsruhe, Germany) and further analysed using Sparky3.110 (T.D. Goddard & D.G. Kneller, University of California, San Francisco).

2.1.6 Structure calculation

Structures were calculated with the protocol SolARIA using CNS, version 1.0. Calculations were performed using the simulated annealing protocol with torsion-angle dynamics, starting with 200 randomised conformers. The ^{13}C-^{13}C restraints obtained from 2D and 3D spectroscopy, were automatically assigned by the programme SolARIA and further used iteratively for structure calculations in CNS. All restraints were classified as distances between 2.5 and 6.5 Å. Prediction of torsion angles was performed using the TALOS software. In total, 55 predictions for the backbone torsion angles were obtained and included in the restraint list for the SolARIA calculation. The error margins for the angular restraints were set to a minimum value of 25 °. The 10 lowest-energy structures calculated using both distance and angular restraints showed no distance violations greater than 0.5 Å and no angular restraints violations larger than 25 °. The αB dimer was calculated with the programme XPLOR-NIH using the restraints from the monomer and the intermolecular ^{15}N-^{13}C distances restraints from the mixed αB.

Results and Discussion

3. Methods for Resolution Enhancement

3.1 Spectral editing: selection of methyl groups in multidimensional solid-state magic-angle spinning NMR

3.1.1 Introduction

Membrane proteins and heterogeneous oligomers are mostly of high molecular weight and show very crowded NMR spectra, which can make it difficult to achieve complete resonance assignments (45). Moreover, these proteins may suffer from inherent structural inhomogeneity, resulting in reduced resolution. Therefore, it is desirable to develop methods that allow for simplification of NMR spectra. One way to alleviate spectral crowding and to achieve additional resolution enhancement is spectral editing. Resonances of specific moieties may be selected by virtue of their unique chemical, spectral and/or conformational properties. Selective detection of methyl groups, for instance, may ease assignments and provide useful structural information, since contacts involving methyl groups form major parameters for structure calculations. Furthermore, methyl resonances are good monitors for ligand binding as demonstrated by Zech *et al.* (2004), who investigated changes in the chemical shifts of methyl resonances of Bcl-xL induced by ligand binding. Opella and Frey (1979) demonstrated that spectral editing for solid-state NMR can be achieved using dipolar dephasing, in order to distinguish chemically-different carbons on the basis of the number of attached protons. Likewise, it was shown that cross-polarisation with polarisation inversion (CPPI) can be applied to select between carbons with different multiplicities (69-73). Other techniques using ^{1}H double-quantum filtering (74) and chemical shift anisotropy dephasing (75) have been suggested. A different approach for spectral editing has been proposed by Emsley and coworkers, that relies on scalar couplings and multiple-quantum filtering to

distinguish carbons with different proton multiplicities (76;77). The advantage of the latter approach is that the filtering is independent of molecular dynamics. On the other hand, long delays for evolution of *J*-couplings and the use of unlabelled material may limit the sensitivity of the experiment.

In this chapter, a simple but robust method is presented that combines CPPI and dipolar dephasing in a double filtering step, to allow filtering of methyl resonances. It will be demonstrated that this methyl-filtering technique can be applied to remove the overlap between Cδ1-Cδ2 and Cγ-Cδ1/2 crosspeaks of leucine residues. Moreover, it will be shown that the double filter implemented into a ^{15}N-^{13}C TEDOR (66;67) correlation experiment enables straightforward detection of correlations between backbone nitrogens and CH$_3$ groups. The pulse sequences have been applied to the α-spectrin SH3 domain and outer-membrane protein G (OmpG) to achieve assignments of leucine resonances and sequential connectivities via methyl filtering. In **chapter 4.1.2.2**, methyl filtering will be applied to aid the assignment of αB-crystallin.

3.1.2 CPPI and dipolar dephasing pulse sequences

Pulse sequences were designed for the following two experiments: a CH$_3$-double-filtered DARR (figure 3-1a) and a CH$_3$-double-filtered TEDOR (figure 3-1b). For both experiments, a short 50 μs polarisation inversion was implemented after a long initial CP of 2 ms. The polarisation-inversion time was chosen as the time needed for signals from rigid CH groups to cross zero (i.e., going from positive to negative intensity). Under these conditions, strongly-coupled, rigid CH$_2$ moieties attained -1/3 of their initial intensity (72). For methyl groups, fast rotation lead to self-decoupling, reducing the effective ^{1}H-^{13}C dipolar couplings. Hence, CH$_3$ carbon spins exchanged more slowly during CP and retained a positive intensity. Likewise, non-protonated carbon spins experienced weaker heteronuclear dipolar interactions and remained positive.
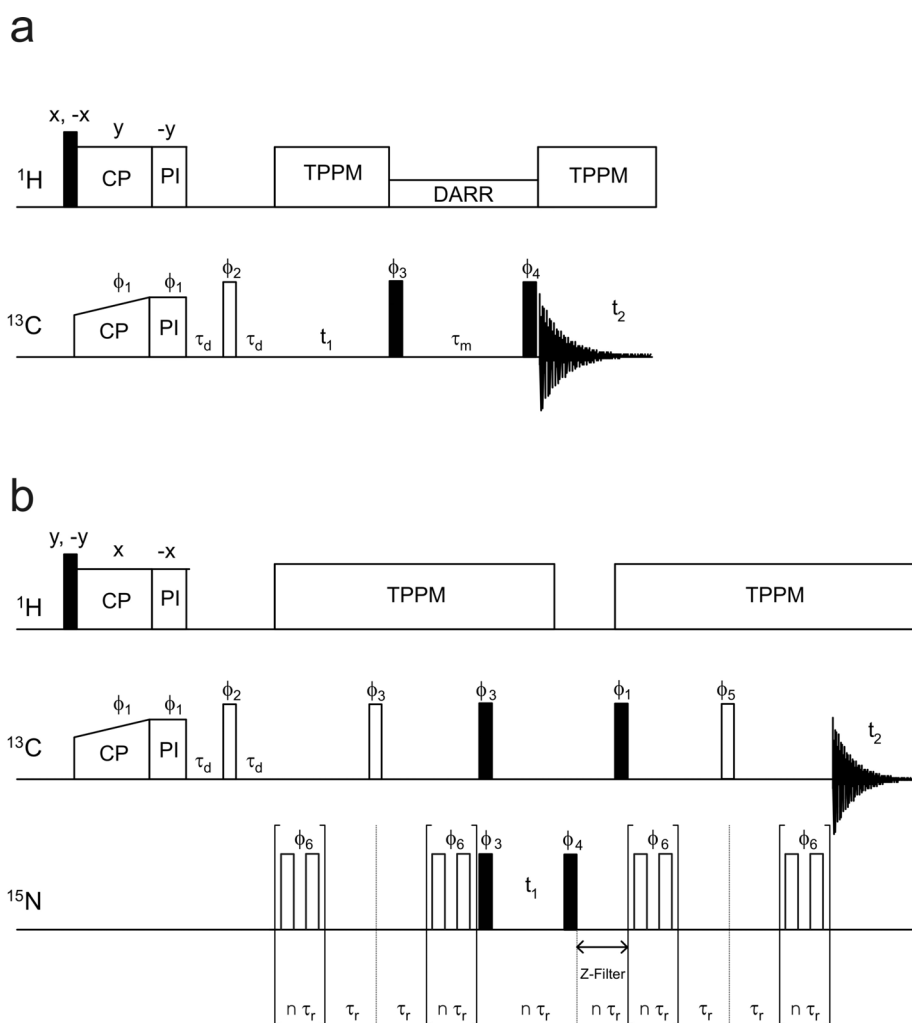
figure 3-1 Pulse sequences for ^{13}C-homonuclear (a) and ^{15}N-^{13}C heteronuclear (b) dipolar correlation experiments used in this work. In both sequences, polarisation was first transferred from protons to carbons using a ramped cross-polarisation (CP). In a next step, cross-polarisation with polarisation inversion (CPPI) (71), in combination with dipolar dephasing (78;79) was used to select methyl groups. In (a), the filter was followed by ^{13}C evolution during t_1 and a dipolar-assisted rotational resonance (DARR) mixing scheme (28). In (b), the filtering block was followed by a rotational-echo double resonance (REDOR) mixing scheme that sandwiched the evolution of ^{15}N in t_1. During t_1-evolution, REDOR mixing and data acquisition, a two-pulse phase modulation (TPPM) scheme was applied to decouple the protons (68). The following phase cycle was applied (a): φ1 = x, -x, y, -y, -x, x, -y, y, φ2 = y, y, -x, -x, -y, -y, x, x, φ3 = y, y, x, x, -y, -y, -x, -x, φ4 = -y, -y, -x, -x, y, y, x, x, φrec = x, x, y, y, -x, -x, -y, -y, (b): φ1 = -x, -x, -y, -y, x, x, y, y, φ2 = y, -y, x, -x, φ3 = x, x, y, y, -x, -x, -y, -y, φ4 = -x, -x, -y, -y, x, x, y, y, x, x, y, y, -x, -x, -y, -y, φ5 = y, -y, -x, x, -y, y, x, -x, φrec = x, -x, y, -y, -x, x, -y, y, -x, x, -y, y, x, -x, y, -y, φ6 was cycled according to the XY16 sheme (80): x, y, x, y, y, x, y, x, −x, −y, −x, −y, −y, −x, −y, -x.

The CPPI filter block was combined with a dipolar-dephasing period in the sequences shown in figure 3-1a and b. During dipolar dephasing, proton decoupling was not applied. This lead to rapid transverse relaxation of methylene (CH_2) and methyne (CH) carbon signals due to strong heteronuclear interactions with their directly bonded protons. This efficiently dephased the negative CH_2 magnetisation and any residual CH signals, whilst ~35 % of the CH_3 signals remained.

This two-step filtering provided a better sensitivity for selection of methyl groups than a dipolar dephasing period alone. For CH groups, a relatively long dephasing period of ~200 μs was required to zero the signals. On the other hand, to cancel out the CH signals using CPPI, a short 50 μs PI-time was sufficient. This CPPI period reduces the CH_2 signals to approximately (minus) 30% of their initial intensity; for dephasing of the remaining CH_2 signal after PI a short dephasing period of 60 μs was adequate. In contrast, a straightforward dipolar-dephasing period to zero both CH and CH_2 coherences needed to be much longer than the combined filtering step and resulted in a sensitivity of only ~15 %. Hence, the most effective filter combination was to initially zero out the CH signals using CPPI, and subsequently, to zero out the residual CH and negative CH_2 signals with dipolar dephasing.

3.1.3 CPPI and dipolar dephasing in 2D ^{13}C-^{13}C dipolar correlation spectroscopy of proteins

In the case of the 2D ^{13}C-^{13}C correlation experiment, the filter block was implemented prior to evolution of ^{13}C signals during t_1 (figure 3-1a), that was followed by a homonuclear ^{13}C recoupling period using a DARR (28) mixing scheme. During the mixing, CH and CH_2 resonances received magnetisation from nearby methyl carbons and could be observed along F_2. Hence, crosspeaks involving CH and CH_2 signals appeared only at one side of the diagonal since they did not evolve during t_1, whilst correlations involving methyl carbons appeared at both sides. figure 3-1 shows an example of the double filter performed on uniformly ^{13}C, ^{15}N labelled α-spectrin SH3 domain (U-SH3). This spectrum shows a superposition of leucine $C\delta1$-$C\gamma$, $C\delta2$-$C\gamma$ and $C\delta1$-$C\delta2$ crosspeaks on the left side of the diagonal. On the right side of the diagonal, isolated $C\delta2$-$C\delta1$ crosspeaks allows the extraction of the respective

chemical shifts. The efficiency of the filtering scheme can be appreciated by disappearance of the strong isoleucine $C\delta$-$C\gamma 1$ crosspeak at 11.2/26.6 ppm, respectively, on the right-hand side of the diagonal.
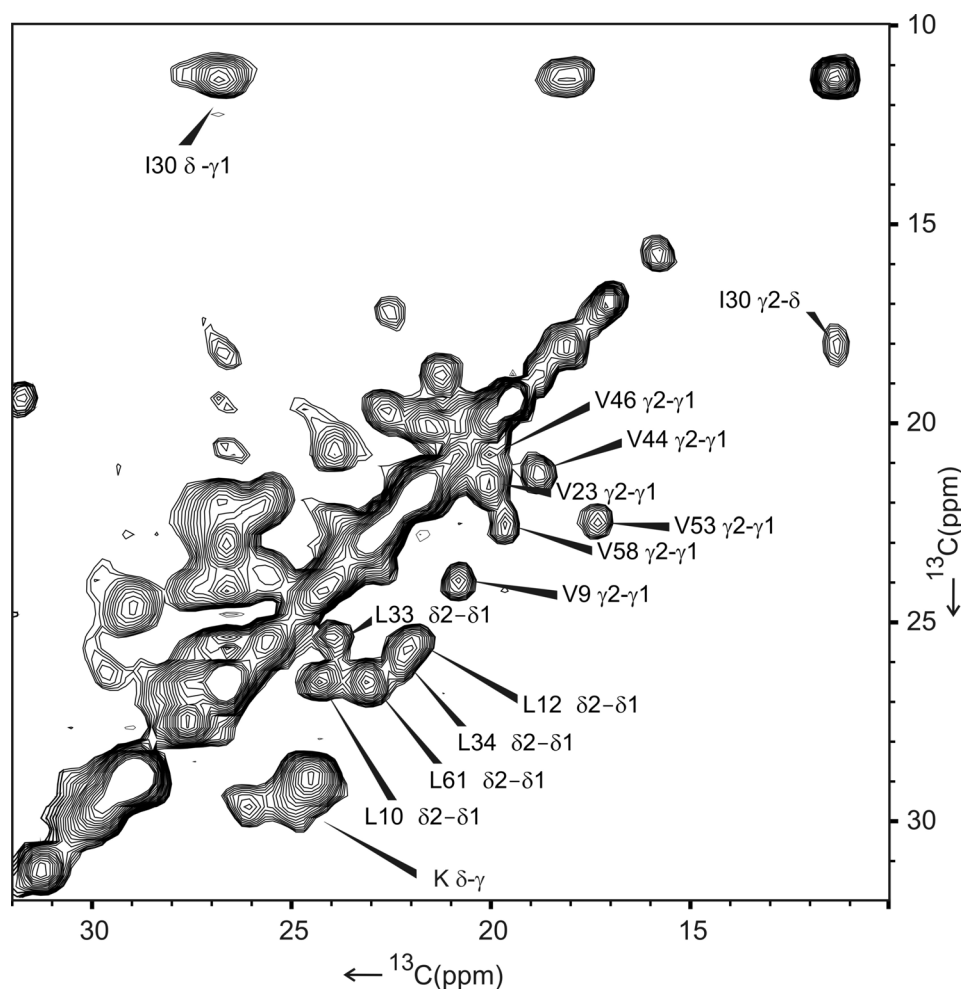
figure 3-2 Contour plot of a two-dimensional ^{13}C-^{13}C homonuclear correlation experiment of uniformly-^{13}C, ^{15}N labelled α-spectrin SH3 domain. The spectrum was recorded using cross-polarisation with polarisation inversion (CPPI) filtering and dipolar dephasing. It demonstrates the resolution enhancement in the methyl region. The spectrum was obtained with 50 μs phase inversion after CP and a 70 ms dipolar-assisted rotational resonance mixing time. Assignments of methyl-methyl crosspeaks are indicated for the amino acids that contain two methyl groups, like isoleucines, leucines and valines.

The effects of the double filter and of side chain mobility are illustrated in figure 3-3. There, spectra of the U-SH3 domain recorded with a CPPI filter only (a), and with CPPI in combination with an additional 60 μs dipolar dephasing step (b), are compared. The spectral region around 42 ppm in F_1 and 25-30 ppm in F_2 is shown. It

is dominated by Cβ-Cγ crosspeaks of leucine residues and Cε-Cδ/Cε-Cγ crosspeaks of lysine residues. In the CPPI-DARR spectrum (figure 3-3a), the CH$_2$ signals from non-flexible parts of the protein were not zeroed out due to the absence of dipolar dephasing, and instead evolved during t$_1$ with a negative intensity. Consequently, the rigid Cβ spins of leucines produced crosspeaks with negative intensity (indicated by dotted contour lines in figure 3-3a). Since the selection of signals with the CPPI filter relied on the relative strengths of heteronuclear dipolar interactions, carbons that were weakly coupled to protons remained positive during CPPI. This was not only the case for methyl groups and non-protonated carbons, but also for CH and CH$_2$ signals of mobile side chains, i.e. the Cε signals of mobile lysine side chains remained positive (figure 3-3a).
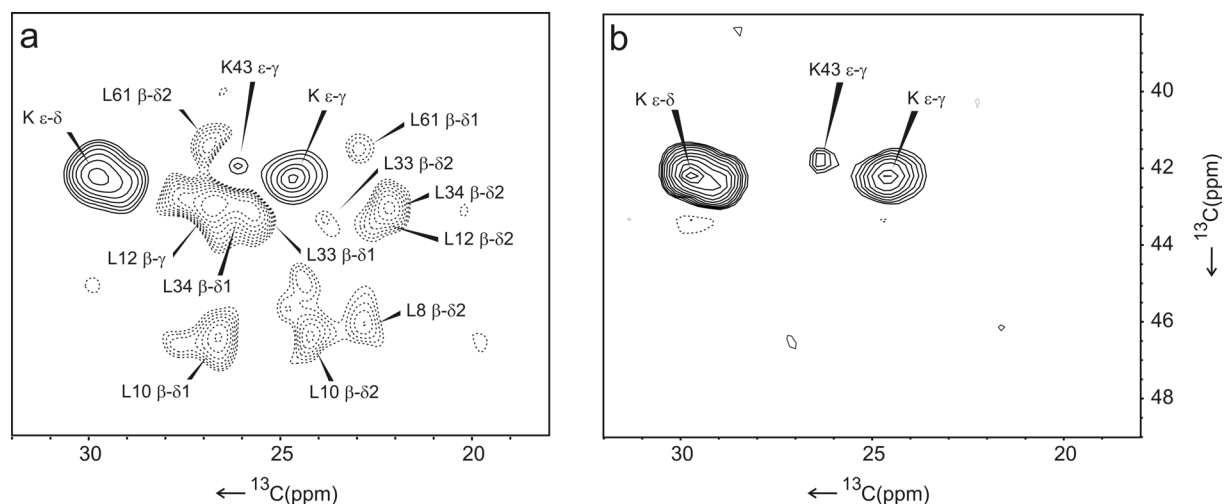
figure 3-3 Contour plots of two-dimensional ^{13}C-^{13}C homonuclear correlation experiments recorded on uniformly-^{13}C, ^{15}N labelled α-spectrin SH3 domain. The spectral region dominated by Cβ-Cγ crosspeaks of leucines and Cε-Cδ/Cε-Cγ crosspeaks of lysines is shown. The ^{13}C-^{13}C correlation spectrum shown in (a) was recorded without dipolar dephasing. The spectrum shown in (b) was recorded using dipolar dephasing after polarisation inversion. Positive contour levels were drawn with solid lines, while negative contour levels were indicated by dotted lines.

Likewise, due to the weaker couplings, mobile side chains were also marginally affected by the dipolar dephasing. Hence, the combination of CPPI and dipolar dephasing only zeroed out negative signals arising from rigid Cβ of leucine residues due to their multiplicity and strong carbon-proton dipolar couplings, whilst mobile C$_ε$ resonances of lysine side chains were hardly affected (figure 3-3b). Side chain

mobility may reduce filter selectivity, and signals from non-methyl groups may "leak through," also remaining positive. Hence, care must be taken when analysing the spectra. In the spectra of the α-spectrin SH3 domain, such leaking appeared rather systematically, in that only CH_2 signals from the end of long hydrophilic side chains (and among those, mainly lysines) were not suppressed by the double filter.

3.1.4 CPPI and dipolar dephasing in 2D ^{15}N-^{13}C correlation spectroscopy of proteins

CPPI in combination with dipolar dephasing can also be used in connection with a TEDOR transfer scheme to recouple methyl carbon and nitrogen spins, using the pulse sequence depicted in figure 3-1b.

A spectrum of uniformly ^{13}C-^{15}N labelled α-spectrin SH3 domain, recorded with a 2 ms REDOR mixing time is shown in figure 3-4a. The spectrum shows pure methyl-carbon to backbone nitrogen correlations. At this short REDOR mixing time, crosspeaks were visible for $^{13}CH_3$ and ^{15}N nuclei that showed distances in the range of ~ 2.5 – 3.0 Å. These were mostly intraresidual ^{15}N-^{13}C pairs of alanines. In addition to all three alanines, Intraresidual correlations for one of the six valines (V53) was observed and for one of the four threonines (T37). At a longer REDOR mixing time of 6 ms, more signals involving methyl carbons and nitrogens in their environment were observed (figure 3-4b). For comparison, a TEDOR experiment without the double filter step is shown in figure 3-4c. This spectrum is more crowded, due to overlap with non-methyl signals. In particular, arginine Cγ signals are found in the chemical shift (CS) range from 12 to 38 ppm (centred around 27 ppm), leucine Cγ signals in the range 15 - 42 ppm (centred around 26 ppm) and isoleucine Cγ1 signals in the range 9 - 38 ppm (centred around 27 ppm) (27). Additional overlap may occur between methyl and lysine C$_\gamma$ signals in the chemical shift range from 17 to 40 ppm (centred at 25 ppm). In the methyl filtered experiment (figure 3-4b), crosspeaks of none of these resonances interfere with crosspeaks from methyl carbons.
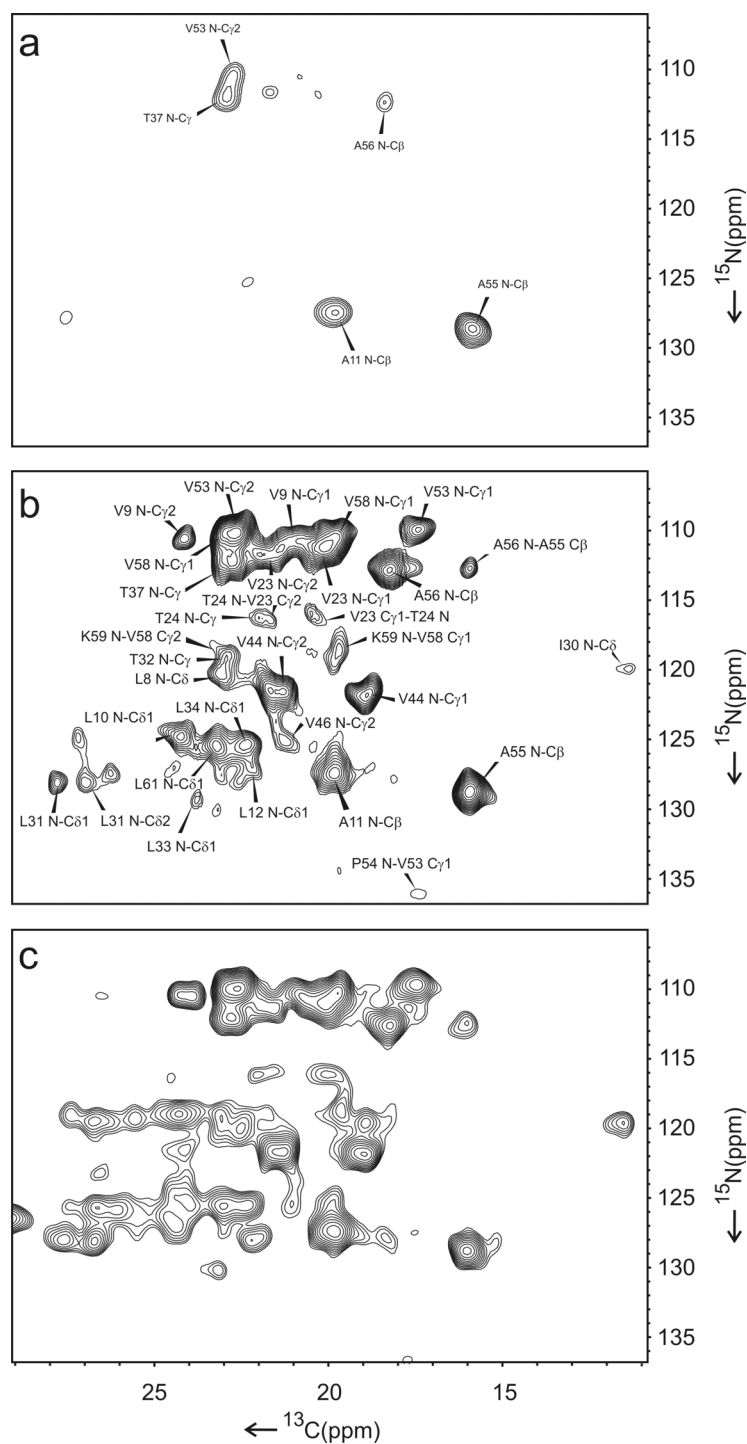
figure 3-4 Contour plot of ^{15}N-^{13}C methyl-selective transferred-echo double resonance (TEDOR) correlation spectra of uniformly-^{13}C, ^{15}N labelled *α*-spectrin SH3 domain, using the pulse sequence depicted in figure 3-1b. For the spectrum shown in (a), a short rotational-echo double resonance mixing of 2 ms was applied. The contour plot in (b) shows a methyl-selective TEDOR spectrum recorded with a mixing time of 6 ms. For comparison, a standard TEDOR experiment (i.e., without the double filter), obtained using a mixing time of 6 ms, is shown in (c).

Most of the correlations in figure 3-4b were again intraresidual. However, some correlations involved sequential residues. Ala 55 and 56, for example, showed a sequential crosspeak at 113 ppm and 15.7 ppm in F_1 and F_2, respectively, that was weaker than the intraresidual crosspeaks of Ala 55 (F_1=129.1 ppm, F_2=15.7 ppm) and Ala 56 (F_1=113 ppm, F_2=18.1 ppm). Likewise, a weak correlation of a methyl carbon signal with a proline backbone nitrogen signal at 136.9 ppm (F_1) and 17.2 ppm (F_2) was observed. This crosspeak involved transfer between the amide nitrogen of Pro 54 and the Cγ1 of Val 53.

Additionally, structural information could be extracted from the cross-peak pattern of branched amino acids. The distances between the valine methyl groups, for example, and the nitrogens in their neighbourhood are conformation-dependent. Only one of the methyl groups of Val 53 showed a crosspeak to the sequential proline. Similarly, a weak crosspeak at 11.2 ppm in F_2 and 119.8 ppm in F_1 was observed, involving Cδ of Ile 30 and its backbone nitrogen. The crosspeak involving Cγ2 was weaker and could only be observed at a deeper contour level (not shown). Indeed, according to the recent X-ray structure (PDB code 1U06)(81), Cγ2 of Ile 30 is more distant from the backbone nitrogen than the Cδ. As a final example, the two crosspeaks involving the methyl groups of Leu 31 and the backbone amide differed slightly in their intensity. Again, this was consistent with the crystal structure, which showed that the individual Cδ are ~ 3.5 and ~ 4.3 Å away from the intraresidual nitrogen.

In a previous paragraph, it is discussed that signals from mobile parts of the protein may leak through the filter. The CH_2 coherences from mobile side chains that may pass the filter, however, will not show crosspeaks involving nitrogens, since the mobility reduces the effective heteronuclear dipolar coupling and interferes with the REDOR transfer.

The methyl-filtered TEDOR experiment was applied to OmpG. OmpG is a membrane-integrated pore for non-selective transport of mono-, di- and trisaccharides (41). The protein consists of 281 amino acids and forms 2D crystals that show well resolved MAS NMR spectra (45). Despite the favourable line width, resonance overlap complicates the sequence-specific assignment of resonances in crowded regions.
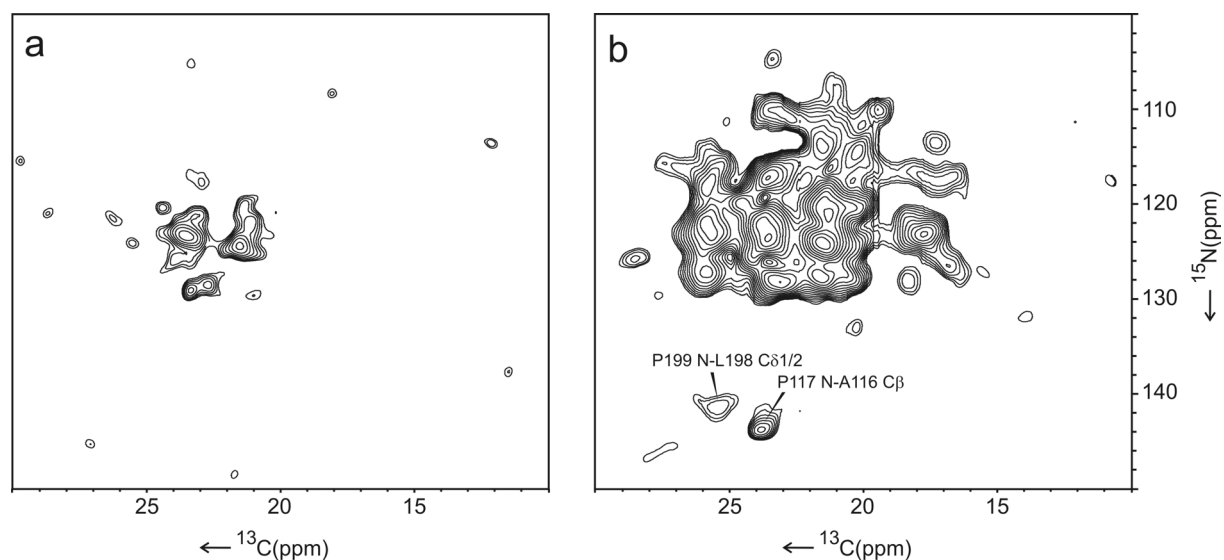
figure 3-5 Contour plots of methyl-selective transferred-echo double resonance correlation spectra of 1,3-^{13}C glycerol-grown outer-membrane protein G, recorded using rotational-echo double resonance mixing times of 3 ms (a) and 8 ms (b). The data were collected using the pulse sequence of figure 3-1b. The sequential assignments are indicated.

By using the CH_3-double-filtered TEDOR sequence at short mixing times (3 ms), many correlations between the backbone nitrogen and the methyl carbons of alanines are observed using a 1,3-^{13}C-glycerol labelled OmpG sample (figure 3-5a). In this spectrum, intraresidual N-Cβ correlations of alanines were expected to appear with the strongest signal intensity; valine, threonine and leucine methyl groups could also give rise to intraresidual cross-peaks. They, however, tended to be weaker since these methyls were more distant from the backbone nitrogens.

In the spectrum recorded with an 8 ms mixing time (figure 3-5b), many signals that were not present in figure 3-5a were found. These signals predominantly arose due to transfer from methyl groups of isoleucines (Cγ2/Cδ), valines (Cγ1/2), leucines (Cδ1/2) and threonines (Cγ) to backbone nitrogens. Furthermore, a strong signal at 23.8 ppm in F_2 and 143.7 ppm in F_1 was observed. The downfield shift in the ^{15}N dimension for this signal ensured that it could be attributed to a correlation involving a backbone amide of a proline. The methyl filtering excluded many assignment options for the correlated carbon. Hence, this correlation could be attributed to sequential transfer from a methyl group at residue i-1 to the nitrogen of a proline at position i. According to the protein sequence, there were only three potential residue pairs that could account for this transfer: A116-P117, L198-P199, T208-P209. Transfer from the Cγ of T208 could, however, be excluded since the chemical shifts detected for the

threonine Cγ's did not match (45). The correlation at 23.8 ppm/143.7 ppm (F$_2$/F$_1$) was most likely due to the A116-P117 pair. A second, weaker, sequential cross-peak was observed at 25.4 ppm in F$_2$ and 141.5 ppm in F$_1$. This peak was tentatively assigned to the other sequential candidate, involving a methyl group of Leu 198 and the backbone nitrogen of Pro 199, although we should keep in mind that a long range correlation could not be excluded. The weaker intensity was consistent with the larger distance between the methyl groups of leucine and the backbone nitrogen of the next amino acid. However, sequential correlations of this type can provide a good starting point for the sequential assignment.

3.2 Resolution enhancement by removal of J-couplings from 2D ^{13}C-^{13}C-correlation experiments

3.2.1 Introduction

MAS NMR still relies strongly on ^{13}C detection. Hence, ^{13}C-^{13}C scalar couplings can impair both resolution and sensitivity in ^{13}C-labelled preparations. *J*-couplings lead to a splitting of the carbon lines into multiplets, or, when not resolved, become manifest in an effective line broadening. Cα signals, for example, contain *J*-couplings to the C' and – if present – the Cβ spins, with coupling constants of $J_{C\alpha C'} \sim 55$ Hz and $J_{C\alpha C\beta} \sim 35$ Hz, respectively. Hence it would be advantageous to remove J-couplings from 2D-^{13}C-^{13}C spectra to improve the resolution.

Various experimental approaches exist for *J*-decoupling. The downside of many experimental methods is that they can bring about a penalty in signal-to-noise ratio (S/N) that is often intolerable for solid-state NMR of 'difficult' biological samples, like membrane proteins. *J*-couplings in the indirect dimension (t$_1$), for example, can be eliminated by using so-called "constant-time" delays (82). However, this method may result in limited sensitivity due to relaxation that takes place during the constant-time interval. In solids, relaxation processes are generally fast, which renders this method less practical for sensitivity reasons. Another approach relies on refocusing of *J*-couplings by combining band-selective and hard 180°-pulses to selectively suppress couplings between certain groups of spins. This method was first established for

liquid-state NMR (83) and later introduced into solid-state NMR (84). A drawback is that it can only be applied in t_1. In a different strategy, homonuclear J-couplings in the direct dimension (t_2) may be removed by using an adiabatic pulse train for selective irradiation during acquisition (85). This method, however, requires dedicated hardware for simultaneous irradiation and detection on the same NMR nucleus. An innovative method that allows J-decoupling in both dimensions is based on combination of an IPAP spin-state selection filter within a constant-time interval in t_1 with a zero-quantum PDSD (proton-driven spin diffusion) sequence, resulting in a spin-state selective polarisation transfer (86;87). The constant time interval and the duration of the selective pulses to assure spin-state selection, however, may cause a substantial loss in sensitivity due to relaxation.

An alternative strategy to remove J-couplings is by post-acquisition data processing. Deconvolution of the raw time-domain data using maximum-entropy (MaxEnt) reconstruction to remove J-couplings has been demonstrated for liquid-state NMR data (88-91). In this chapter it is demonstrated, that deconvolution of MAS NMR data with MaxEnt reconstruction allows removal of splittings due to J-couplings without expenses in sensitivity. We show that a combination of MaxEnt reconstruction in t_2 with a selective pulse in t_1, produces fully J-resolved data in both dimensions. The method was applied to preparations of proteins expressed on media containing 2-^{13}C-glycerol as the only carbon source, leading to labelling schemes that result in a restricted set of J-couplings so that the effect of decoupling is readily observed. When using 2-^{13}C-glycerol as carbon source, valines contain $^{13}C\alpha$-$^{13}C\beta$ spin-pairs that can be suitably used for investigation of the performance of the decoupling technique without being obscured by overlap from signals from other residues. As a test system we used a 2-^{13}C-glycerol preparation of the α-spectrin SH3 domain, which is denoted as 2G-SH3 (2). As an example of a membrane protein with higher molecular weight (281 residues), we used a preparation of the outer-membrane protein G obtained using 2-^{13}C-glycerol as carbon source, denoted as 2G-OmpG (45).

3.2.2 *J*-deconvolution using Maximum Entropy (MaxEnt) reconstruction

Maximum Entropy (MaxEnt) reconstruction is an alternative for Discrete Fourier

Transformation (89). Unlike DFT, MaxEnt handles the conversion of time domain data to frequency domain data as an inverse problem. During MaxEnt processing, spectra are iteratively generated until one spectrum f fits the experimental data. In this context, the generated spectrum f is inverse Fourier transformed (IDFT) and multiplied with a convolution kernel ker [2] to generate an FID m which can be compared with the experimental data d.

$$m_k = ker * \mathrm{IDFT}(f(k\Delta t))$$

equation 3-1

To compare m_k with the experimental data we assume that the noise and the experimental error is normally distributed. The experimentally measured resonance signal d_k is composed of the sum of the pure signal $\bar{d}_k$ and the noise ε_k at the point in time $k\Delta t$.

$$d_k = \bar{d}_k + \varepsilon_k$$

equation 3-2

Hence, the difference of the experimental d_k and the generated m_k can be used for a least square fit χ^2 (the noise is considered normally distributed):

$$C(\mathbf{f}) = \frac{\chi^2(\mathbf{f})}{2} = \sum_{k=0}^{M-1} |m_k - d_k|^2$$

equation 3-3

To make sure that the generated spectrum matches the experimental data, we can set an upper allowed margin of the error C_0:

$$C(\mathbf{f}) \leq C_0$$

equation 3-4

An additional restraint for the MaxEnt reconstruction is the so called principle of maximum entropy (89). The entropy of a probability distribution is a measure for its information density. During MaxEnt processing, the number of points of the reconstruction is (1) restricted to M data points and (2) may not contain additional information than the experimental data provide. Thus, the entropy $S(f)$ of the reconstruction reaches a maximum if the information density of the experimental data is equal to the information density of the reconstruction (89).

---

[2] The convolution kernel ker represents the folding core, that contains for instance the term $(cos \pi Jt)$ for J-deconvolution or operations to compensate for non-linear sampled data.

The reconstruction of a spectrum from NMR time domain data by using MaxEnt comprises the maximization of the entropy $S(f)$ and the fit with the experimental data according to equation 3-4. The Rowland NMR Toolkit offers a software platform that solves the problem described above with a numerical algorithm developed by Hoch and coworkers (89).

J-couplings appear in the FID as a cosine modulation *(cos πJt)* and result in a peak splitting after DFT whilst the separation distance corresponds to the J-coupling constant in Hertz. To eliminate J-couplings, a division of the FID by *(cos πJt)* would be sufficient, but divisions by values near zero are numerical inconsistent and thus result in spikes and high noise levels after DFT. For J-deconvolution by using MaxEnt, the convolution kernel *ker* from equation 3-1 contains the term "*(cos πJt)*" and is multiplied with IDFT(f), whereas the spectrum f does not contain J-couplings. Hence, the generated FID m_k contains J-couplings and is compared with the experimental data. We applied MaxEnt processing to remove Cα-Cβ J-couplings of 35 Hz from 2D ^{13}C-^{13}C correlation spectra.

3.2.3 J-decoupling in t_1 by selective refocusing

J-couplings may be removed from the indirect dimension by selective refocusing in t_1 (84). A PDSD pulse sequence with J-decoupling is shown in figure 3-6.
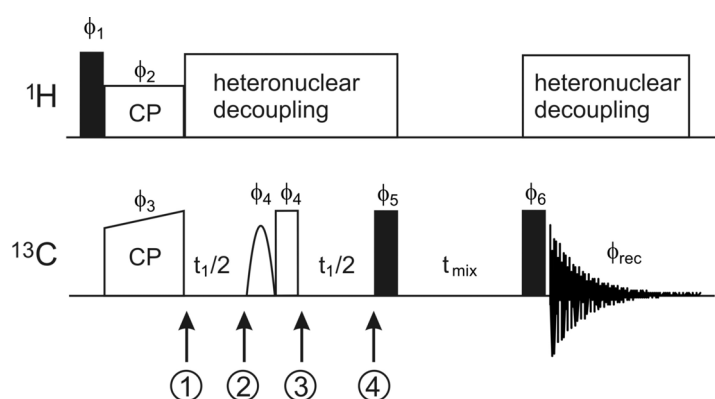
figure 3-6 PDSD pulse sequence with the suppression of *J*-couplings in t_1. Filled bars represent hard 90° pulses. Magnetisation is first transferred from ^{1}H to ^{13}C using a ramped cross polarisation (CP), followed by evolution in t_1. In the middle of t_1, a selective and a hard refocusing 180° pulse are applied. The phase listing is as following: ϕ_1 = y, -y; ϕ_2 = x; ϕ_3 = x, x, y, y; ϕ_4 = y, -x, -x, -y, -y, x, x, y; ϕ_5 = y, -y, x, -x; ϕ_6 = -y, y, -x, x and ϕ_{rec} = -x, x, -y, y.

The experiment uses a band-selective 180° pulse followed by a hard 180° pulse at the frequency of the Cα during t_1. ^{13}C-^{13}C mixing is achieved by a spin-diffusion mixing after t_1.

To describe the function of the experiment for a two spin ½ system, we ignore dipolar couplings which are eliminated by MAS and heteronuclear decoupling. Furthermore, off-resonance effects of the pulses are disregarded. Hence we can describe the spin system with the weak coupling Hamilton operator in the rotating frame (equation 3-5):

$$\mathbf{H} = \sum_i \Omega_i S_{iz} + 2\pi \sum_{j<k} J_{ik} S_{jz} S_{kz}$$

equation 3-5

Ω_i represents the isotropic chemical shift and J_{ik} represents the scalar coupling constant between S_i and S_k. We can follow the time dependent evolution of the magnetisation indicated at the numbered time steps from figure 3-6 by using the product operator formalism. Spin S_1 is selected by the selective pulse after step 3 which is responsible for decoupling in ω_1:

S_{1z}+S_{2z}

$\xrightarrow{\quad CP \quad}$

①
S_{1x}+ S_{2x}

$\xrightarrow{\Omega_1 t_1/2 S_{1z} \quad \Omega_2 t_1/2 S_{2z} \quad 2\pi J_{12} t_1/2 S_{1z} S_{2z}}$

②

$S_{1x}\cos\pi Jt\cos\Omega_1 t_1/2 + 2S_{1y}S_{2z} \sin\pi Jt\cos\Omega_1 t_1/2$
$+S_{1y}\cos\pi Jt\sin\Omega_1 t_1/2 - 2S_{1x}S_{2z} \sin\pi Jt\sin\Omega_1 t_1/2$
$+S_{2x}\cos\pi Jt\cos\Omega_2 t_1/2 + 2S_{1z}S_{2y}\sin\pi Jt\cos\Omega_2 t_1/2$
$+S_{2y}\cos\pi Jt\sin\Omega_2 t_1/2 - 2S_{1z}S_{2x} \sin\pi Jt\sin\Omega_2 t_1/2$

$\xrightarrow{\pi S_{1x} \quad \pi(S_{1x}+S_{2x})}$

③

$S_{1x}\cos\pi Jt\cos\Omega_1 t_1/2$ - $2S_{1y}S_{2z}\sin\pi Jt\cos\Omega_1 t_1/2$

$+S_{1y}\cos\pi Jt\sin\Omega_1 t_1/2+ 2S_{1x}S_{2z}\sin\pi Jt\sin\Omega_1 t_1/2$
$+S_{2x}\cos\pi Jt\cos\Omega_2 t_1/2$ - $2S_{1z}S_{2y}\sin\pi Jt\cos\Omega_2 t_1/2$
$-S_{2y}\cos\pi Jt\sin\Omega_2 t_1/2$ - $2S_{1z}S_{2x}\sin\pi Jt\sin\Omega_2 t_1/2$

$$\xrightarrow{\Omega_1 t_1/2 S_{1z}} \xrightarrow{\Omega_2 t_1/2 S_{2z}} \xrightarrow{2\,\pi J_{12} t_1/2 S_{1z}S_{2z}}$$

④

$S_{1x}\cos\Omega_1 t_1+S_{1y}\sin\Omega_1 t_1$ + S_{2x}

The spin S_2 is not affected by the band-selective pulse applied in step 3, and is refocused during t_1. In a 2D correlation experiment its resonance appears at the carrier frequency parallel to the F_2 axis. The magnetisation of the selected spin S_1 evolves during t_1 under its chemical shift. The crosspeak of the S_1 spin shows in the indirect dimension a resolution enhancement equal to the scalar coupling constant of spin S_1 and the coupled spin. To remove the signal of the unselected spin S_2, the phase of the selective refocusing pulse is cycled x, y (0°, 90°) while the receiver phase is cycled simultaneously x, -x (0°, 180°). To illustrate the phase cycling, we follow the magnetisation again using the product operator formalism, beginning from time step 3 in figure 3-6:

$$\xrightarrow{\pi S_{1y}} \xrightarrow{\pi(S_{1x}+S_{2x})}$$

③.1

$-S_{1x}\cos\pi Jt\cos\Omega_1 t_1/2$ + $2S_{1y}S_{2z}\sin\pi Jt\cos\Omega_1 t_1/2$
$-S_{1y}\cos\pi Jt\sin\Omega_1 t_1/2$ - $2S_{1x}S_{2z}\sin\pi Jt\sin\Omega_1 t_1/2$
$+S_{2x}\cos\pi Jt\cos\Omega_2 t_1/2$ - $2S_{1z}S_{2y}\sin\pi Jt\cos\Omega_2 t_1/2$
$-S_{2y}\cos\pi Jt\sin\Omega_2 t_1/2$ - $2S_{1z}S_{2x}\sin\pi Jt\sin\Omega_2 t_1/2$

$$\xrightarrow{\Omega_1 t_1/2 S_{1z}} \xrightarrow{\Omega_2 t_1/2 S_{2z}} \xrightarrow{2\,\pi J_{12} t_1/2 S_{1z}S_{2z}}$$

④.1

$$-(S_{1x}\cos\Omega_1 t_1 + S_{1y}\sin\Omega_1 t_1) + S_{2x}$$

The selective pulse has now phase y, and the selected spin S_1 has opposite sign, while spin S_2 did not change its sign. If we cycle the receiver phase by 180°, the signal of spin S_1 adds up while other signals are subtracted. We can conclude that the minimum phase cycle for the experiment is φ_4 = x, y, φ_{rec} = x, -x, keeping the other phases constant.

3.2.4 Results and discussion

3.2.5 Generating fully *J*-decoupled spectra in t_1 and t_2

figure 3-7a–c shows the valine Cα-Cβ crosspeaks of 2G-SH3, recorded or processed under different conditions.
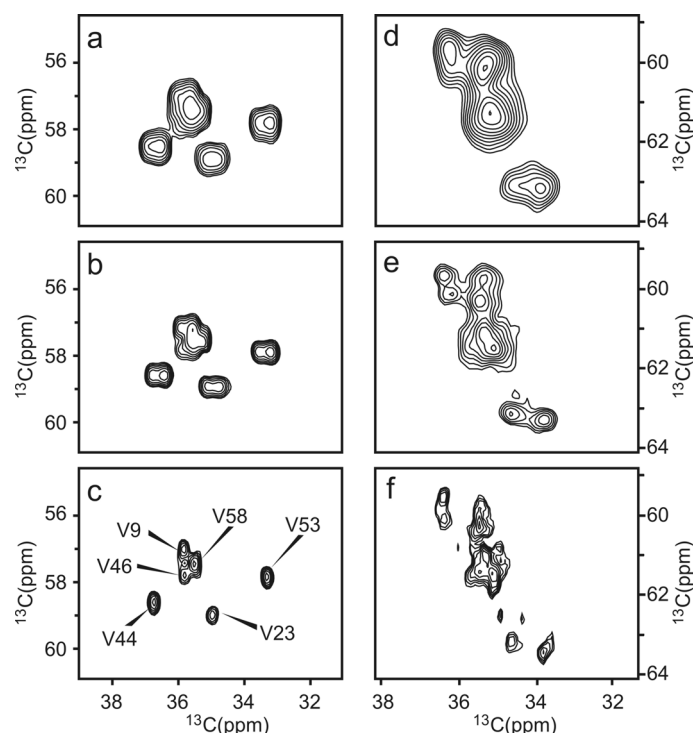
figure 3-7 Valine Cα-Cβ crosspeaks of 2G-SH3 (a–c) and 2G-OmpG (d–f). The data shown in (a) and (d) are 2D PDSD reference spectra (25ms mixing) (24). The spectra in (b) and (e) additionally contained a hard 180° pulse and a band-selective refocusing pulse on Cα during t_1. Figures (c) and (f) show the same data as in (b) and (e), respectively, deconvolved with maximum-entropy reconstruction instead of FFT in t_2, using the Rowland NMR Toolkit

version3.0. The *J*-coupling constant for MaxEnt reconstruction was set to 35Hz. The proteins were expressed in bacteria on media with 2-^{13}C-glycerol (92) as the sole carbon source.

The data in figure 3-7a is the reference spectrum, which is a 2D PDSD (24) correlation experiment (25ms mixing, pulse sequence of figure 3-6 without refocusing pulses in t_1). The spectrum in figure 3-7b was obtained with similar experimental settings, but the pulse sequence additionally contained a hard 180°-pulse and a band-selective 180°-pulse at the frequency of the Cα during t_1 (figure 3-6). Figure 3-7c shows the same experiment as in figure 3-7b, but processed with MaxEnt deconvolution in t_2. In figure 3-8, we compare various columns and rows extracted from the data in figure 3-7a–c. However, the reduction in line width under free precession and under homonuclear decoupling amounts to about 20-25 Hz, which is somewhat less than the expected narrowing upon removal of the scalar coupling constant $J_{C^\alpha C^\beta}$ of ~35 Hz (table 3-1). This difference is attributed to the limited FID resolution in the indirect dimension (the *J*-splittings of the Cα-Cβ crosspeaks are resolved in t_2 but not in t_1). Hence, even though the *J*-couplings are not resolved in the indirect dimension of the experiment, the line widths are reduced by the selective pulses (figure 3-8, rows I and II). In the direct dimension, the splittings due to the *J*-coupling between the valine Cα and Cβ spins are partially resolved and collapse into single peaks after performing *J*-deconvolution with MaxEnt (figure 3-8, rows III and IV, respectively). The deconvolution results in reduction of 36-40 Hz of the line width, which is in good agreement with the expected line narrowing (table 3-1). The S/N value for the various crosspeaks increases from figure 3-7a to b.
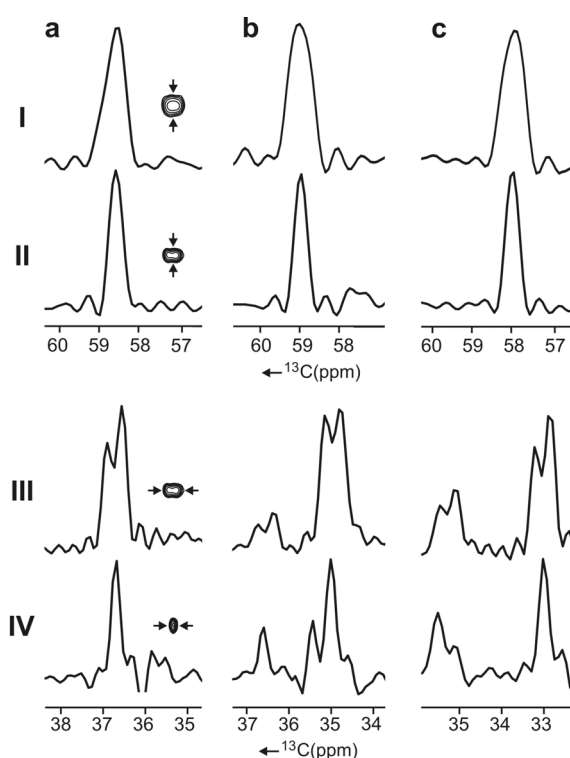
figure 3-8 Slices extracted from figure 3-7 for the SH3 Cα-Cβ crosspeaks of V44 (a), V23 (b) and V53 (c). The small icons schematically depict the crosspeaks as observed in the 2D data and show the dimension of the extracted slices. Columns extracted from figure 3-7a are displayed in Panel I; columns and rows taken from figure 3-7b are shown in Panels II and III, respectively; rows taken from figure 3-7c are shown in Panel IV.

Signal-to-noise and line widths are summarized in table 3-1. The experimental *J*-decoupling (hard 180° pulse and selective 180° pulse on the valine α carbons in t_1) results in a 25-40% gain in sensitivity; this gain is less than the expected gain upon collapsing a doublet into a singlet, which we ascribe to the negative influence of the selective pulse on the sensitivity. It is difficult to compare this to the S/N in the data reconstructed with MaxEnt (figure 3-7c), since the apparent noise can be tuned by the choice of parameters used for the reconstruction. Our goal was to remove the *J*-coupling and not the noise, hence we *chose* the parameters in such a way that the experimental data and the reconstructed data have about the same S/N. It is however noted that S/N is an experimental constant, whilst the apparent noise in the reconstructed data very much depends on the exact choice of parameters used for the MaxEnt reconstruction; hence, by using different parameter settings, the S/N can be tuned. For the sake of completeness, we have added the S/N values for the reconstructed data in table 3-1 as well.

The results for 2G-OmpG are shown in figure 3-7d–f. The data in figure 3-7d contains the reference spectrum whilst the spectrum in figure 3-7e was recorded using a band-selective refocusing pulse at the Cα frequency during t_1. Figure 3-7f shows the same experimental data as in figure 3-7e, but processed with MaxEnt deconvolution in t_2. Without *J*-decoupling (figure 3-7d), the 2D ^{13}C-^{13}C PDSD spectrum is very congested in the valine Cα-Cβ region, which, for OmpG, comprises the response from 14 valine Cα-Cβ crosspeaks. With homonuclear *J*-decoupling applied in t_1, the region of the Cα-Cβ crosspeaks becomes better resolved (figure 3-7e). The best resolution is obtained if the data are additionally processed with MaxEnt in t_2 (figure 3-7f).

table 3-1 . Signal-to-noise (S/N) and line widths (full-width at half height, FWHH) of three resolved 2G-SH3 valine Cα-Cβ crosspeaks (V23, V44 and V53) determined from the spectra shown in figure 3-7. The S/N-values observed in spectra processed with MaxEnt are in quotation marks, see main text for details.

	spin pair	S/N	FWHH ω_1 (Hz)	FWHH ω_2 (Hz)
figure 3-7a	V44 Cα-Cβ	40	61	63
	V23 Cα-Cβ	33	65	64
	V53 Cα-Cβ	31	66	64
figure 3-7b	V44 Cα-Cβ	56	41	63
	V23 Cα-Cβ	42	42	64
	V53 Cα-Cβ	44	44	64
figure 3-7c	V44 Cα-Cβ	"51"	41	22
	V23 Cα-Cβ	"32"	42	28
	V53 Cα-Cβ	"48"	44	25

MaxEnt deconvolution needs to be performed carefully in order to guarantee that the 'mock' FID (the product of the inverse Fourier transform of the reconstructed spectrum with the convolution kernel) closely resembles the experimental FID and

differences should be within the error tolerance set for the calculations (89). For the MaxEnt reconstruction, the error tolerance (the '*aim*' parameter in NMR Toolkit language) was set to a value just above the experimental noise, whilst the scale factor (the *def* parameter) was set to a fraction of about 0.02 of the aim parameter. Rules of thumb for the choice of these parameters and several illustrative demonstrations of their effect on the calculations can be found in the book from Hoch and Stern (89). In the calculations on the OmpG data (figure 3-7e and f), some small additional signals appear in the reconstructed data, hence it is recommended to compare MaxEnt processed data with the FT-transformed data. As these signals do not correspond with crosspeaks in the experimental data, they can be recognized as artefacts. Possible artefacts that may arise in the reconstructed spectra are ghost signals or phase distortions.

3.2.6 Performance of Maximum Entropy Reconstruction

In order to obtain a qualitative estimate about the performance of the method, we compared spectra obtained from 2G-OmpG deconvolved in t_1 using the MaxEnt reconstruction method, with data recorded with experimental *J*-decoupling (figure 3-9):
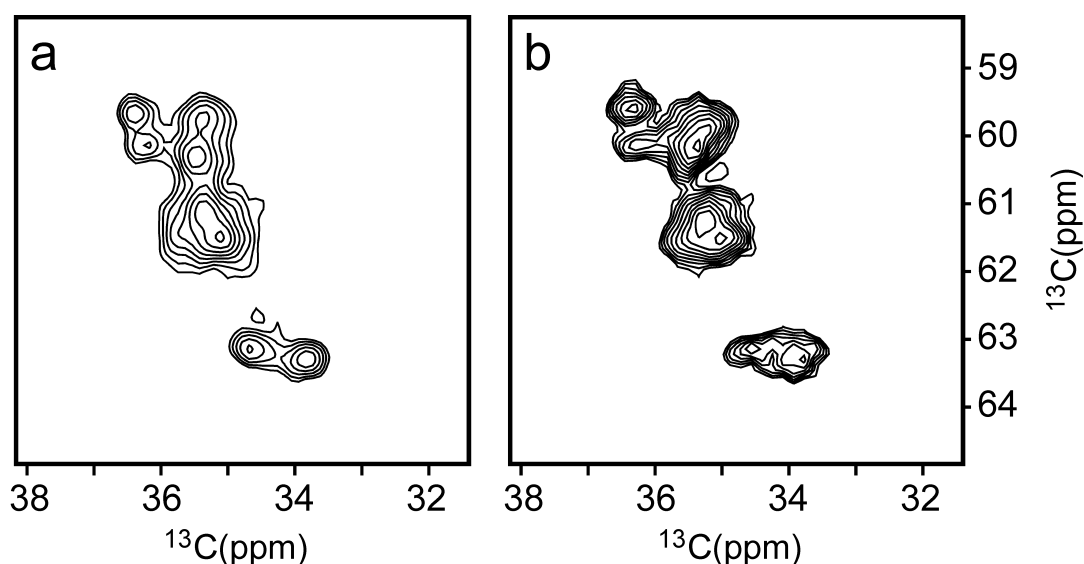
figure 3-9 Comparison of experimental J-decoupling using a selective pulse on the Cα spins during t_1 (a) with deconvolution using maximum-entropy reconstruction in t_1 (b). The data were recorded at 17.6 T on a preparation of 2G-OmpG.

In this figure, experimental removal of the Cα-Cβ *J*-couplings of valines was achieved by application of a hard 180°-pulse followed by a selective 180°-pulse on the valine Cα spins (a), using a preparation of 2G-OmpG. Deconvolution with MaxEnt reconstruction was used to remove *J*-couplings in the indirect dimension from a standard PDSD experiment, recorded on 2G-OmpG (b). Both techniques provide *J*-decoupled data in the indirect dimension that compare well. There are, however, some small differences in the cross-peak positions ($\leq$0.2 ppm) in the indirect dimension and in the line widths observed in the direct dimension, most likely resulting from slightly different measurement conditions.

3.3 Conclusions

In conclusion, we have demonstrated that *J*-decoupling using refocusing pulses in t_1 and MaxEnt reconstruction in t_2 leads to well-resolved spectra without sacrificing the S/N to an intolerable amount. Since MaxEnt is a post-acquisition processing method, existing data can be reprocessed and analysed in greater detail. MaxEnt reconstruction can be seen as a powerful and viable alternative to experimental approaches for *J*-decoupling. However, MaxEnt processing of 2D data in both dimensions is difficult, since it can not be applied sequential, but both dimensions must be processed at once. The performance of such processing with regard to artefacts makes it advisable to restrict MaxEnt processing to one dimension.

Another approach to enhance the resolution of 2D ^{13}C-^{13}C solid-state NMR spectra has proved to be Methyl Filtering by CPPI and dipolar-dephasing.

The combination of CPPI and dipolar dephasing provides a simple way to obtain efficient filtering of methyl groups. This double filter can be easily implemented as a building block into multidimensional MAS NMR correlation experiments, for instance by using it as preparative filtering step prior to homonuclear (^{13}C-^{13}C) or heteronuclear (^{15}N-^{13}C) transfer schemes. The filtering relies on the relative strengths of the ^{1}H-^{13}C dipolar interactions. As a consequence, coherences from dynamic side chains may behave differently to those from rigid side chains, and some non-methyl signals can escape from the filter. On the other hand, this 'leaking-through' is limited to a small number of signals from mobile side chains, mainly lysines. In particular, in

the heteronuclear variant of the experiment, mobile fractions do not pass the REDOR transfer step and do not interfere with the filtered signals. Suppression of CH and CH_2 signals can simplify significantly spectral regions containing methyl groups. The assignment of leucine signals, for example, may profit from the reduction of overlap between $C\delta2$-$C\delta1$ and $C\gamma$-$C\delta$ crosspeaks. Moreover, it is shown that a methyl-selective ^{15}N-^{13}C correlation experiment can be particularly helpful for achieving starting points for the assignment. Using a short REDOR mixing time of 2 - 3 ms, it is possible to select intraresidual correlations between amide nitrogens and methyl groups that are separated by distances of ~ 3 Å. Hence, when combined with a short REDOR transfer, this experiment selects predominantly methyl-nitrogen correlations for short amino acid side chains, like alanines (methyl group at $C\beta$), and to a lesser extent residues with a methyl group at $C\gamma$ (*e.g.*, valines, threonines). For longer REDOR mixing times, additional sequential and conformational information can be extracted.

4. Structure Determination of αB-crystallin

4.1 ^{13}C and ^{15}N assignment of αB-crystallin

4.1.1 Preparation of αB oligomers for solid-state NMR measurements

αB is highly soluble, up to several hundreds of milligrams in a millilitre of water. It does not crystallise even at high concentrations as in the eye lens, which is an essential attribute to maintain the lens transparent. For our structural studies of full length αB by solid-state NMR, an immobilisation of the protein is required. In 1981, Lawson *et al.* showed by circular dichroism spectroscopy that the structure of αB does not change upon addition of PEG (93). We assumed that an increasing concentration of PEG will decrease the solvation of the αB oligomer resulting in precipitation of oligomers. The structure of the αB oligomer is pH-dependent and seems to be more stable at pH 7.3 (94). We used the following conditions for precipitation: 20-30 mg/ml αB dissolved in a 50 mM sodium phosphate buffer containing 50 mM NaCl and 0.002 M NaN$_3$ at pH 7.6-7.7. A preparative vapour diffusion chamber (figure 4-1) was used, applying the sitting "drop" method. The precipitation starts after addition of PEG8000 solution (50% w/v) to a final concentration of ~8.5-10% (w/v) and is complete after 2-3 days. The reservoir solution contained 50% (w/v) PEG8000. A ZrO$_2$ rotor was filled with the precipitant and further used for solid-state NMR experiments.
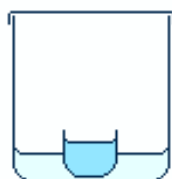
figure 4-1 The figure shows a vapour diffusion chamber with a little vessel containing the protein mix inside the reservoir chamber surrounded by the reservoir solution.

To probe the sample quality, it is common in solid-state NMR to record 1D ^{13}C-spectra which can be analysed in terms of linewidth. The spectrum in figure 4-2a was recorded from dry lyophilised αB. The resonance lines are broad and do not show typical sharp features as the spectrum in figure 4-2b. The latter was recorded from

αB precipitated from a solution with PEG8000. The αB preparation used to record spectrum b shows resolved Cβ- and methyl-resonances of threonines and isoleucines, respectively. This suggests folded protomers in the precipitate. We assume that the precipitate consists of intact oligomers consisting of protomers, which are stabilised in the molecular network of the oligomer.
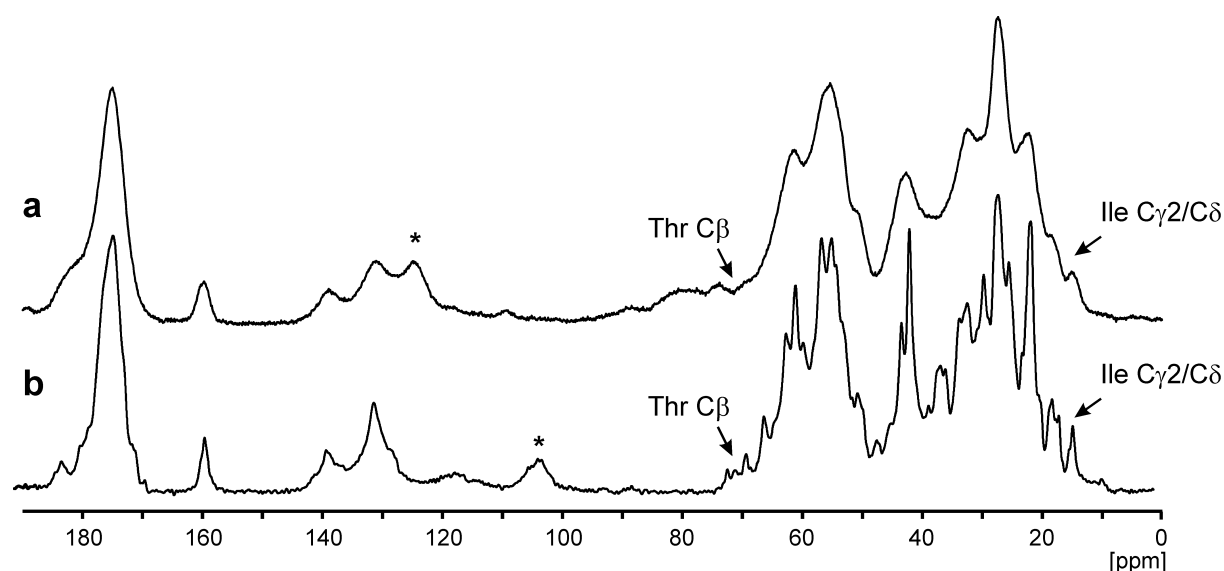
figure 4-2 The figure shows two 1D ^{13}C CPMAS spectra of different αB-crystallin preparations. The spectrum in a was recorded from lyophilised αB while the spectrum in b was recorded from αB precipitated from a solution with PEG8000. Both spectra were recorded with ~85 kHz ^{1}H decoupling and 15 ms acquisition at 285 K. During CP a field of ~60 kHz and ~50 kHz was applied on ^{13}C and ^{1}H respectively. The spectrum in a was recorded at a field of 18.8 Tesla and a MAS frequency of ω_r=8 kHz. The spectrum in b was recorded at a field of 16.4 Tesla and a MAS frequency of ω_r=10 kHz. Spinning sidebands are marked with an asterisk.

Our initial experiments showed that the sample quality decreases quickly (within one week of measurement time), when using moderately-high ^{1}H-decoupling (~85 kHz) at T=285K. To prolong the lifetime of our samples, subsequent experiments were done at a temperature of 270 K, at which the sample is not frozen and survives several months of measurement.

In figure 4-3, a 2D ^{13}C-^{13}C-correlation spectrum with 15 ms spin-diffusion mixing time is shown. As expected from the 1D spectrum in figure 4-2, the chemical shift distribution of the threonine resonances allows the assignment of secondary structure elements to individual residues. At least four out of nine signal sets for threonines are observed, while only three threonines are in the α-crystallin domain.

Likewise, we observe at least six out of nine Cβ resonances of isoleucines, implicating that the four isoleucines of the α-crystallin domain and at least two of the isoleucines located in the C- and N-termini (two and three respectively) are observed. In particular those in the C-terminal IXI/V motif are of interest, since this motif is supposed to cover a chaperone binding site and is involved in intermolecular interactions (47;48). Hence, solid-state NMR experiments on full-length αB may reveal structural details elucidating the roles of the termini.
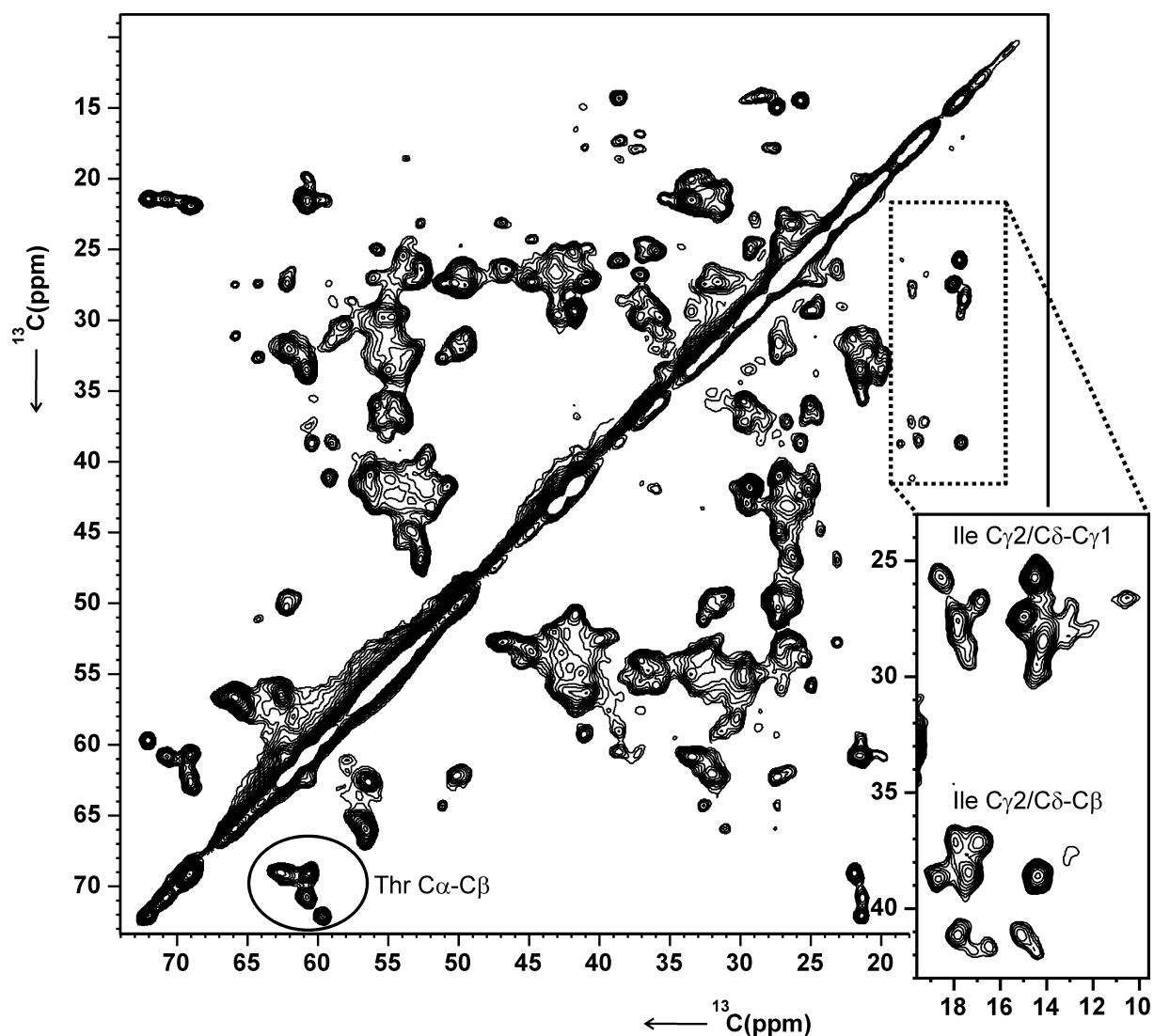
figure 4-3 2D ^{13}C-^{13}C-PDSD spectrum recorded of αB precipitated with PEG using a mixing time of 15 ms. The spectrum was recorded at a field of 16.4 Tesla, at 270 K, a MAS frequency of ω_r=10 kHz and a ~87 kHz ^{1}H decoupling field. The threonine Cα-Cβ crosspeaks are marked with a circle and the isoleucine Cγ2/Cδ-Cγ1 and -Cβ correlations are highlighted.

The linewidth observed for methyl resonances is around 80 Hz which is more than usually observed for microcrystalline samples (63). The heterogeneity of the molecules in the oligomer may be the cause for that observation.

4.1.2 Assignment of full-length αB-crystallin oligomers

Amino acid types from αB are identified and sequentially connected by using various 3D ^{15}N-^{13}C-^{13}C correlation spectra. In addition, J-decoupling in t_1 and filtering of methyl resonances helped to resolve peaks that are in congested regions of 2D ^{13}C-^{13}C spectra. A Methyl REDOR spectrum (chapter 3.1) was used to corroborate the assignment of an alanine and for the sequential assignment of I124 and P125. The assignments for M68 in a putative β2-strand were obtained from a PFMW αB sample, that contains selectively ^{15}N/^{13}C labelled proline-, phenylalanine-, methionine- and tryptophane-residues.

4.1.2.1 Assignment of strips extracted from 3D spectra

The sequential assignment is illustrated for the segment L131-T132-I133-T134-S135-S136-L137 in figure 4-4. The sequential walk starts at the ^{15}N frequency of L137 in an NCACX spectrum of U-αB (strip **a**). The amino acid type is identified from characteristic chemical shifts, which is confirmed by the following assignment procedure. In the NCOCX experiment of 1,3G-αB (strip **b**) at the same ^{15}N frequency, we find a crosspeak at 66 ppm (labelled **1**), characteristic for serine Cβ, and a signal at 52.9 ppm (labelled **2**). Since only the carbons indicated by green letters in figure 4-4 are labelled in 1,3G-αB, the signal at 52.9 ppm is attributed to a sequential transfer from S136 C´ to L137 Cα. The Cα frequency of S136 is found at 56.7 ppm in the NCOCX spectrum of U-αB at the ^{15}N frequency of L137 (strip **c**). These assignments are confirmed by the NCACB spectrum of U-αB. At the ^{15}N frequency of S136 (strip **d**), correlations involving its Cα and Cβ are observed. In the NCOCX spectrum of U-αB at the S136 ^{15}N frequency (strip **e**), the C´ (173.3 ppm), Cα (56.9 ppm) and Cβ (66.9 ppm) resonances of S135 are identified.

$\delta^{15}N(L137)=122.9$ ppm

U - NCACX
$\tau_{mix} = 25$ ms

L137N-Cα-Cβ   L137N-Cα-Cγ
L137N-Cα-Cδ'

a)

$\delta^{15}N(L137)=122.9$ ppm

1,3G - NCOCX
$\tau_{mix} = 75$ ms

L137N-S136C-Cβ   L137N-S136C-L137Cα

b)

$\delta^{15}N(L137)=122.9$ ppm

U - NCOCX
$\tau_{mix} = 35$ ms

L137N-S136C-Cβ   L137N-S136C-Cα   K103N-G102C-Cα

c)

$\delta^{15}N(S136)=118.3$ ppm

U - NCACB
$\tau_{mix} = 2$ ms

S136N-Cα-Cβ

d)

$\delta^{15}N(S136)=118.3$ ppm

U - NCOCX
$\tau_{mix} = 35$ ms

S136N-S135C-Cβ   S136N-S135C-Cα   S76N-F75C-Cβ

e)

$\delta^{15}N(S135)=112.1$ ppm

U - NCACB
$\tau_{mix} = 2$ ms

S135N-Cα-Cβ   E87N-Cα-Cβ

f)

$\delta^{15}N(S135)=112.1$ ppm

2G - NCACX
$\tau_{mix} = 200$ ms

S135N-Cα-T134C   S135N-Cα-T134Cα

g)

$\delta^{15}N(S135)=112.1$ ppm

U-NCOCX
$\tau_{mix} = 35$ ms

S135N-T134C-Cα   S135N-T134C-Cβ   G102N-H101C-Cα   S135N-T134C-Cγ2

h)

$\delta^{15}N(T134)=117.3$ ppm

U-NCACX
$\tau_{mix} = 25$ ms

T134N-Cα-C   T134N-Cα-Cβ   S139N-Cα-Cβ   T134N-Cα-Cγ2

i)

$\delta^{15}N(T134)=117.3$ ppm

2G - NCACX
$\tau_{mix} = 200$ ms

T134N-Cα-I133C

j)

$\delta^{15}N(T134)=117.3$ ppm

U-NCOCX
$\tau_{mix} = 35$ ms

T134N-I133C-Cα   T134N-I133C-Cβ   T134N-I133C-Cγ1   T134N-I133C-Cγ2

k)

$\delta^{15}N(I133)=121.9$ ppm

U-NCACB
$\tau_{mix} = 2$ ms

I133N-Cα-Cβ

l)

$\delta^{15}N(I133)=121.9$ ppm

U-NCOCX
$\tau_{mix} = 35$ ms

I133N-T132C-Cβ   I133N-T132C-Cα   I133N-T132C-Cγ2

m)

$\delta^{15}N(T132)=106.4$ ppm

U-NCACX
$\tau_{mix} = 25$ ms

T132N-Cα-C   T132N-Cα-Cβ   T132N-Cα-Cγ2

n)

$\delta^{15}N(T132)=106.4$ ppm

U-NCOCX
$\tau_{mix} = 35$ ms

T132N-L131C-Cα   T132N-L131C-Cβ   T132N-L131C-Cγ2

o)

$\delta^{15}N(L131)=115.5$ ppm

U-NCACX
$\tau_{mix} = 25$ ms

L131N-Cα-C   L131N-Cα-Cβ   L131N-Cα-Cγ

p)

$\delta^{15}N(L131)=115.5$ ppm

1,3G-NCACX
$\tau_{mix} = 500$ ms

L131N-Cα-T132Cβ   L131N-Cα-T132Cβ   L131N-Cα-Cδ
(L131N-Cα-P130Cα)

q)

^{13}C(ppm)

figure 4-4 Strips extracted from various 3D triple resonance spectra are shown. The assignment of the fragment L131 to L137 is illustrated using strips from 3D NCACX, NCACB and NCOCX spectra recorded on uniformly, 1,3-^{13}C-glycerol and 2-^{13}C-glycerol labelled samples. On the top right of each strip the ^{15}N frequency in F$_1$ is given at which the strip is extracted. The F$_2$ dimension is an indirect ^{13}C dimension containing either Cα or C′ frequencies. The F$_3$ dimension is the direct ^{13}C dimension containing intraresidual and sequential correlations, depending on the mixing time and the labelling pattern. The labelling scheme, experiment type and mixing time are indicated at the right hand side. "U" denotes a uniformly ^{13}C, ^{15}N, "1,3G" a 1,3-^{13}C-glycerol, ^{15}N and "2G" a 2-^{13}C-glycerol, ^{15}N labelled sample. The experiment type is indicated with NCOCX, NCACX or NCACB. In NCACX spectra, the ^{15}N and Cα of residue i is correlated with the side chain of residue i, in NCACB, the ^{15}N and Cα of residue i is correlated with the Cβ of residue i. NCOCX spectra provide corresponding sequential information and correlate ^{15}N of residue i and C′ of residue i-1 with the side chain of residue i-1 (see main text for details).

The Cα and Cβ shifts of S135 are confirmed in the NCACB spectrum at the ^{15}N frequency of S135 (strip **f**). In the NCACX spectrum of 2G-αB at the S135 ^{15}N frequency (strip **g**), S135 Cα correlates with a C′ at 173 ppm (labelled **3**) and a Cα at 60.1 ppm (labelled **4**). In the NCOCX spectrum of U-αB at the S135 ^{15}N frequency (strip **h**), we observe a crosspeak involving a Cα resonance at 60.1 ppm, correlated with a C′ at 173 ppm. In addition, we observe a crosspeak involving a Cβ resonance at 72.5 ppm, characteristic for a threonine. The sequential Cα-Cα (labelled **4**) and Cα-C′ (labelled **3**) crosspeaks in strip **g** connect S135 and T134 unambiguously. The assignments for T134 Cα, Cβ and C′ are confirmed by the NCACX spectrum of U-αB (strip **i**) at the ^{15}N frequency of T134. A sequential Cα-C′ crosspeak (labelled **5**) is observed in an NCACX spectrum recorded from 2G-αB at the ^{15}N frequency of T134, which was assigned to C′ of I133 (strip **j**). This assignment is confirmed by the NCOCX spectrum of U-αB, where at the T134 ^{15}N frequency (strip **k**), correlations involving the C′ of I133 and the Cα (60.5 ppm), Cβ (37.7 ppm), Cγ1 (27.5 ppm) and Cγ2 (18.2 ppm) resonances are observed, as characteristic for isoleucine. The ^{15}N frequency of I133 can be found by searching for matching Cα and Cβ frequencies in an NCACB spectrum of U-αB (strip **l**). The ^{13}C resonances of T132 correlate with the backbone ^{15}N signals of T132 itself (in the NCACX spectrum, strip **n**) and of I133 (in the NCOCX spectrum, strip **m**), both spectra are recorded on U-αB. In strip **n** correlations involving T132 Cα (61.0 ppm) and T132 Cβ (69.4 ppm), Cγ (22.2 ppm)

and C' (174.1 ppm) are observed. In the NCOCX spectrum (strip **m**), Correlations involving the Cα, Cβ, Cγ and C' resonances of T132 are again observed correlated with the backbone amide of I133. Following a similar procedure, L131 can be sequentially correlated with T132. Strip **o** shows correlations involving the ^{15}N frequency of T132 with the Cα (56.7 ppm), Cβ (41.0 ppm), Cγ (27.6 ppm) and C' (178.6 ppm) resonances of L131. The same L131 Cα, Cβ, Cγ and C' signals are observed in the NCACX spectrum recorded on U-αB at the ^{15}N frequency of L131 (strip **p**). The sequential assignment of L131 is corroborated by sequential correlations between L131 Cα and T132 Cα (labelled **6**) and T132 Cβ (labelled **7**) in an NCACX spectrum of 1,3G-αB at ^{15}N frequency of L131 (strip **q**).

As illustrated above, the labelling patterns from 1,3G-αB and 2G-αB samples remove certain assignment options in the spectra and hence reduce ambiguity.

Mc Haourab and coworkers found a β-strand in αA involving the residues E113, R112, S111 and I110. These residues were found to be involved in intermolecular contacts with the same β-strand of a second molecule (95). The homologous sequence in αB (E117, R116, S115 and I114) may be important for a better understanding of intermolecular interactions in αB. Hence an unambiguous assignment is inalienable and will be demonstrated in figure 4-5 and figure 4-6.
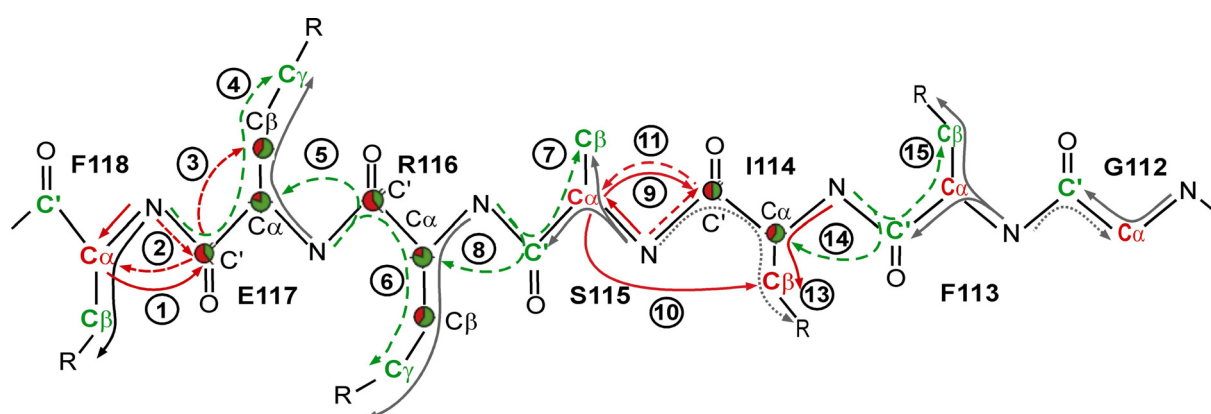
figure 4-5 The sequence of the α-crystallin domain spanning residues G112-F118 is shown. Dotted and dashed arrows represent magnetisation transfers in NCOCX experiments and solid arrows in NCACX experiments. The colours green, red and black refer to the three differently labelled samples, 1,3G-, 2G- and U-αB. The arrows are coloured according to the sample of which the experiment was recorded. The colour of the nuclei represent the labelling pattern observed in the 1,3G-αB and 2G-αB sample. Some nuclei are labelled in both samples, in this case the amount of labelling in each sample is indicated by coloured circles. The numbering

refers to observed crosspeaks in figure 4-6.

In figure 4-5 the polarisation transfer schemes are indicated on a molecular model of the sequence spanning residues G112-F118. Dashed and dotted arrows represent NCOCX and solid arrows NCACX transfers, respectively. Magnetisation transfer in three differently labelled samples, 1,3G-αB, 2G-αB, and U-αB is shown in three different colours, green, red, and black, respectively. The colour of the nuclei corresponds to the labelling in either the 1,3G-αB (green) or 2G-αB sample (red). A filled circle with different percentages of red and green reflects nuclei which are labelled with different percentages in both samples, this is the case for E117, R116 and I114. The spin pattern of F118 is identified in the NCACX spectrum of the PFMW-labelled αB sample at the corresponding ^{15}N plane, which is not shown in figure 4-6. Phenylalanine can be identified from characteristic Cβ shifts around 38-43 ppm and typical shifts for aromatic Cγ resonances at ~140 ppm. Strip **a** from figure 4-6 shows the ^{15}N plane of F118 in an NCACX spectrum of 2G-αB. Dashed lines guide the eye through the strips, following the sequential connections of the amino acids. The C$'$ frequency of E117 is found at 174.7 ppm in F$_3$, correlated with the Cα frequency of F118 in F$_2$ (numbered with **1** in strip **a**). The NCOCX strip **b** of 2G-αB is extracted at the ^{15}N frequency of F118 (121.5 ppm). The correlation at 174.7 ppm in F$_2$ and 56.6 ppm in F$_3$ is attributed to a sequential transfer from E117 C$'$ to F118 Cα, indicated by **2** in strip **b**. Crosspeak **3** in strip **b** at 33.8 ppm in F$_3$ can be assigned to a Cβ resonance. Given that the spectrum is recorded from 2G-αB, and the chemical shift of the Cβ is 33.8 ppm, the amino acid type can be constrained to a glutamate, glutamine, arginine, or proline, because these residues have simultaneously C$'$ and Cβ labelled in 2G-αB and have average Cβ shifts that match 33.8 ppm. The Cβ resonance at 33.8 ppm is found again in the NCOCX spectrum of 1,3G-αB extracted at the ^{15}N frequency of F118 (strip **c**) together with a crosspeak at 36.0 ppm in F$_3$ (labelled **4**), which is attributed to the Cγ resonance of E117. Consequently, we can identify the amino acid type and connect it sequentially as E117 to F118, which is the sole glutamate (i-1) phenylalanine (i) pair in αB. The assignments for E117 are confirmed by the NCACX spectrum of U-αB (strip **d**) at the ^{15}N frequency of E117 (119.7 ppm).

F118N-Cα-E117C

δ¹⁵N(F118)=121.5 ppm — 2G-NCACX, τ_mix = 60 ms

a)

F118 Cα

E117 C′

δ¹⁵N(F118)=121.5 ppm — 2G-NCOCX, τ_mix = 75 ms

F118N-E117C-F118Cα   F118N-E117C-Cβ

b)

E117 Cβ

δ¹⁵N(F118)=121.5 ppm — 13G-NCOCX, τ_mix = 75 ms

F118N-E117C-Cγ   F118N-E117C-Cβ

c)

E117 Cγ

δ¹⁵N(E117)=119.7 ppm — U-NCACX, τ_mix = 25 ms

E117N-Cα-Cβ   E117N-Cα-Cγ

d)

δ¹⁵N(E117)=119.7 ppm — U-NCACB, τ_mix = 2 ms

E117N-Cα-Cβ

e)

E117 Cα

δ¹⁵N(E117)=119.7 ppm — 13G-NCOCX, τ_mix = 75 ms

E117N-R116C-E117Cα   E117N-R116C-Cβ   E117N-R116C-Cγ

C′

f)

R116 Cβ   R116 Cγ

δ¹⁵N(R116)=115.0 ppm — U-NCACX, τ_mix = 25 ms

R116N-Cα-C   R116N-Cα-Cδ   R116N-Cα-Cβ   R116N-Cα-Cγ

g)

R116 Cα

δ¹⁵N(R116)=115.0 ppm — U-NCACB, τ_mix = 2 ms

R116N-Cα-Cβ

h)

δ¹⁵N(R116)=115.0 ppm — 13G-NCOCX, τ_mix = 75 ms

S115 C′   R116N-S115C-Cβ   R116N-S115C-R116Cα

i)

S115N-Cα-C   S115 Cβ   S115N-Cα-Cβ

δ¹⁵N(S115)=121.8 ppm — U-NCACX, τ_mix = 25 ms

j)

S115N-Cα-I114C   S115 Cα

δ¹⁵N(S115)=121.8 ppm — 2G-NCACX, τ_mix = 200 ms

S115N-Cα-I114Cβ

k)

I114 C′   S115N-I114C-S115Cα   S115N-I114C-Cβ

δ¹⁵N(S115)=121.8 ppm — 2G-NCOCX, τ_mix = 75 ms

l)

I114 Cβ

S115N-I114C-Cα   S115N-I114C-Cβ   S115N-I114C-Cδ

δ¹⁵N(S115)=121.8 ppm — U-NCOCX, τ_mix = 35 ms

m)

I114 Cα

δ¹⁵N(I114)=119.8 ppm — 2G-NCACX, τ_mix = 60 ms

I114N-Cα-I114Cβ

n)

F113 C′   I114N-F113C-Cβ   I114N-F113C-I114Cα

δ¹⁵N(I114)=119.8 ppm — 13G-NCOCX, τ_mix = 75 ms

o)

F113 Cβ

F113N-Cα-C   F113N-Cα-Cβ

δ¹⁵N(F113)=120.9 ppm — PFMW-NCACX, τ_mix = 25 ms

p)

G112 C′   F113N-G112C-Cα

δ¹⁵N(F113)=120.9 ppm — U-NCOCX, τ_mix = 35 ms

q)

G112N-Cα-C   G112 Cα   G112N-Cα-Cα

δ¹⁵N(G112)=109.5 ppm — U-NCACX, τ_mix = 25 ms

r)

¹³C(ppm)

figure 4-6 Strips extracted from various 3D triple resonance spectra are shown. The assignment of the fragment G112 to F118 is illustrated using strips from 3D NCACX, NCACB and NCOCX spectra recorded on uniformly, 1,3-^{13}C-glycerol and 2-^{13}C-glycerol labelled samples to demonstrate the sequential assignment along a β-strand that may be involved in the dimerisation of αB-crystallin. On the top right of each strip the ^{15}N frequency in F$_1$ is given at which the strip is extracted. The F$_2$ dimension is an indirect ^{13}C dimension containing either Cα or Cʹ frequencies. The F$_3$ dimension is the direct ^{13}C dimension containing intraresidual and sequential correlations, depending on the mixing time and the labelling pattern. The labelling scheme, experiment type and mixing time are indicated at the right hand side. "U" denotes a uniformly ^{13}C, ^{15}N, "1,3G" a 1,3-^{13}C-glycerol, ^{15}N and "2G" a 2-^{13}C-glycerol, ^{15}N labelled sample. In addition a sample was prepared which has proline, phenylalanine, methionine and tryptophane selectively ^{13}C, ^{15}N labelled (blue contours, labelled "PFMW"). The experiment type is indicated with NCOCX, NCACX or NCACB. In NCACX spectra, the ^{15}N and Cα of residue i is correlated with the side chain of residue i, in NCACB, the ^{15}N and Cα of residue i is correlated with the Cβ of residue i. NCOCX spectra provide corresponding sequential information and correlate ^{15}N of residue i and Cʹ of residue i-1 with the side chain of residue i-1 (see main text for details).

In the NCACB spectrum of U-αB at the ^{15}N frequency of E117 (strip **e**), a Cα-Cβ correlation is observed, while the absence of a Cα-Cγ crosspeak of E117 further confirms the previous assignment of the amino acid type. The NCOCX spectrum of 1,3G-αB (strip **f**) is extracted at the ^{15}N frequency of E117 (119.7 ppm). Crosspeak **5** indicates the correlation N(E117)-Cʹ(R116)-Cα(E117) which connects E117 and R116. The crosspeaks **6** in strip **f** at 37.3 ppm and 30.0 ppm in F$_3$ and 175.0 ppm in F$_2$ are attributed to the Cʹ-Cβ and Cʹ-Cγ correlations of R116, respectively. The NCACX spectrum of U-αB in strip **g** at the ^{15}N frequency of R116 (115.0 ppm) confirms the Cʹ, Cα, Cβ, Cγ and Cδ resonances of R116. The presence of the Cα-Cβ and the absence of the Cα-Cγ correlation of R116 in the NCACB strip **h** of U-αB at the ^{15}N frequency of R116 corroborate the previous assignments. In the NCOCX experiment of 1,3G-αB (strip **i**) at the ^{15}N frequency of R116, a crosspeak at 65.8 ppm (labelled **7**) is found, characteristic for serine Cβ (S115 Cβ) and a signal at 55.2 ppm (labelled **8**). The signal at 55.2 ppm is attributed to a sequential transfer from S115 Cʹ to R116 Cα. The Cα of S115 at 57.1 ppm is extracted from the NCOCX spectrum of U-αB at the ^{15}N frequency of R116 (data not shown). The NCACX spectrum of U-αB at the ^{15}N frequency of S115 (strip **j**) matches the Cʹ, Cα and Cβ

resonances of S115 obtained from the NCOCX spectra. The NCACX strip **k** of 2G-αB connects the Cα of S115 with the C′ (173.0 ppm) and Cβ (42.1 ppm) of I114 (labelled **9** and **10**). An NCOCX spectrum of 2G-αB contains C′-Cβ crosspeaks for isoleucines, since C′ and Cβ of isoleucines are simultaneously labelled in 2G-αB. The NCOCX of 2G-αB (strip **l**) extracted at the ^{15}N frequency of S115 correlates the C′ frequency of I114 in F$_2$ with the Cβ frequency of I114 in F$_3$ (labelled **12**). Crosspeak **11** indicates the sequential correlation N(S115)-C′(I114)-Cα(S115). The remaining resonances for Cα and Cδ are observed in the NCOCX strip **m**. In the 2G-αB sample only valine and isoleucines are simultaneously Cα and Cβ ^{13}C labelled. Hence, characteristic Cα-Cβ crosspeaks can be identified in NCACX spectra of 2G-αB, as demonstrated for I114 in strip **n**, which is extracted at the ^{15}N frequency of I114; the Cα-Cβ crosspeak at 57.0 ppm in F$_2$ and 42.1 ppm in F$_3$ is indicated with **13**. Again the sequential connection to F113 is established from an NCOCX strip recorded from 1,3G-αB (strip **o**). Crosspeak **14** indicates the sequential correlation N(I114)-C′(F113)-Cα(I114). Crosspeak **15** indicates the correlation N(I114)-C′(F113)-Cβ(F113) at 175.3 ppm in F$_2$ and 42.75 ppm in F$_3$. The ^{15}N frequency of F113 is found by searching an NCACX spectrum recorded from PFMW-αB for a ^{15}N plane matching the frequencies found for F113 (strip **p**). The NCOCX strip **q** delivers the C′ and Cα chemical shifts of G112 at 171.5 ppm in F$_2$ and 44.8 ppm in F$_3$, respectively. Again, the ^{15}N (109.5 ppm) resonance of G112 is found by searching an NCACX spectrum recorded from U-αB for a slice matching the Cα and C′ chemical shifts of G112 (strip **r**).

As demonstrated in the strip plots from figure 4-4 and figure 4-6, the ambiguity of the sequential assignment decreases with the number of found sequential crosspeaks in the spectra recorded from 1,3G-αB and 2G-αB. This labelling strategy was indispensable to compensate for the lack of experiment types during the assignment process of full-length αB crystallin from solid-state NMR data. In addition, the 1,3-^{13}C-glycerol and 2-^{13}C-glycerol labelling pattern removes certain assignment options in the spectra and hence reduces ambiguity. With this procedure we could assign the α-crystallin domain and the C-terminus from full-length αB almost completely from solid-state NMR data. The amino acids from the N-terminal domain (~50% assigned) yield crosspeaks with less intensity than those from the α-crystallin domain which could be due to heterogeneity or mobility.

4.1.2.2 Assignments from methyl filtered spectra

Filtering of methyl resonances (chapter 3.1) was used during the assignment process of αB and helped to resolve leucine spin patterns and to assign the segment A126-P125-I124. Correlations involving Cδ1 and Cδ2 of leucine and Cγ1 and Cγ2 of valine are indicated as Cδ* and Cγ*, respectively. The spectrum in figure 4-7 on the right shows a section from a methyl-filtered DARR experiment containing leucine Cδ*-Cα and Cδ*-Cβ crosspeaks. For comparison, the spectrum on the left shows the same region of an unedited 2D ^{13}C-^{13}C DARR experiment. Both spectra were recorded at a field of 21.1 T with 70 ms spin diffusion mixing. In the methyl filtered experiment, the polarisation is transferred exclusively from methyl groups to the other carbons in the side chain by spin diffusion. In essence, only signals of those amino acids appear in the spectra that have a methyl group.
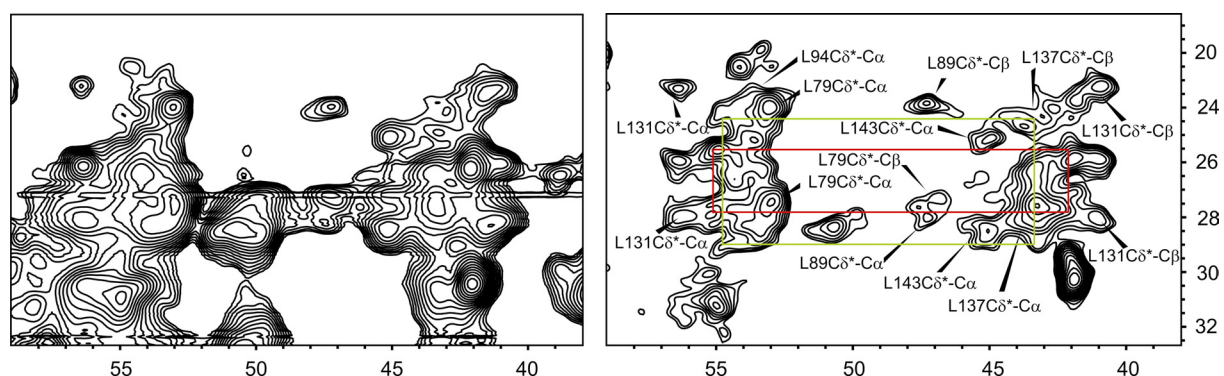
figure 4-7 A methyl filtered 2D ^{13}C-^{13}C DARR experiment on the right hand side is compared with the corresponding unedited 2D ^{13}C-^{13}C DARR experiment. The methyl filtered spectrum contains Cδ*-Cα and Cδ*-Cβ correlations of leucine residues from αB crystallin. The unedited spectrum shows the same spectral region. Both spectra were recorded at a field of 21.1 T with 70 ms spin diffusion mixing time and 90 kHz ^{1}H decoupling. The methyl filtered experiment was recorded with cross polarisation and 50 µs phase inversion followed by a 70 µs dipolar dephasing period for filtering of methyl resonances.

The methyl-filtered spectrum shows improved resolution for methyl-bearing amino acids compared with the unfiltered spectrum. In particular, the Cδ*-Cα crosspeak from L131 at 27.6 ppm in F$_1$ and 56.7 ppm in F$_2$ is now partly resolved in the edited spectrum. Moreover, the methyl filtered DARR experiment reveals further leucine

signal patterns which are hidden in the spectrum on the left. The patterns of yet unassigned leucines from the N-terminal domain, are indicated by green and red coloured rectangles.

Methyl-filtered REDOR spectra can aid the sequential assignment for amino acids containing a methyl-group (chapter 3.1). In the case of αB, methyl-filtered ^{15}N-^{13}C spectra were analysed to assign the triplet I124-P125-A126. In 2D ^{13}C-^{13}C correlation spectra of αB, two out of six alanines are observed. In 3D spectra, the ^{15}N-Cα-Cβ correlations of the alanines are not observed or weak and hence the ^{15}N frequencies could not be assigned unambiguously. However, the methyl-filtered REDOR spectrum with 3 ms mixing time, as shown on the right in figure 4-8, contains the ^{15}N-Cβ correlation of an alanine at 18.8 ppm (Cβ) in F_2 and 120.6 ppm (^{15}N) in F_1. The Cβ resonance matches the alanine Cβ resonances observed in 2D ^{13}C-^{13}C correlation spectra of αB. Further analysis of 3D NCOCX spectra confirmed that the ^{15}N frequency found in the spectrum in figure 4-8 belongs to A126 and was further used to assign the ^{13}C-resonances of P125 (spectra not shown).
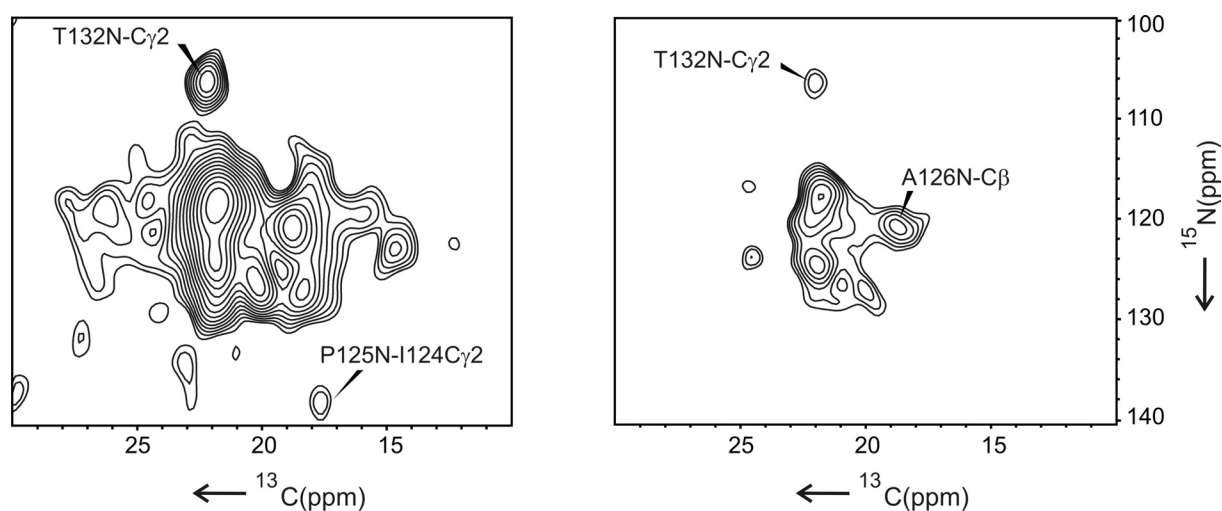
figure 4-8 Contour plots of methyl-filtered transferred-echo double resonance correlation spectra of U-αB, recorded using rotational-echo double resonance mixing times of 3 ms for the spectrum to the right and 6 ms for the spectrum to the left. The data were collected using the pulse sequence from chapter 3.1. The assignments are indicated.

Subsequently, the sequential crosspeak in the spectrum on the left in figure 4-8, recorded with 6 ms REDOR mixing time, connects the Cγ2 from I124 in F_2 at 17.7 ppm with the ^{15}N of P125 in F_1 at 138.4 ppm. The ^{15}N-Cγ2 correlation of T132 is

indicated in both spectra and was used to calibrate the spectra.

4.1.2.3 The assignment of mobile residues in the N- and C-terminus of αB

A 2D ^{1}H-^{13}C INEPT spectrum was recorded to assign residues in the mobile termini in αB. ^{1}H-^{13}C J-couplings are used in INEPT spectra to transfer magnetisation and ^{1}H decoupling is not applied. Hence, mobile residues with vanishing dipolar couplings are detected in that type of experiments. In figure 4-9, a contour-plot of a 2D ^{1}H-^{13}C INEPT-spectrum, recorded from U-αB is shown.
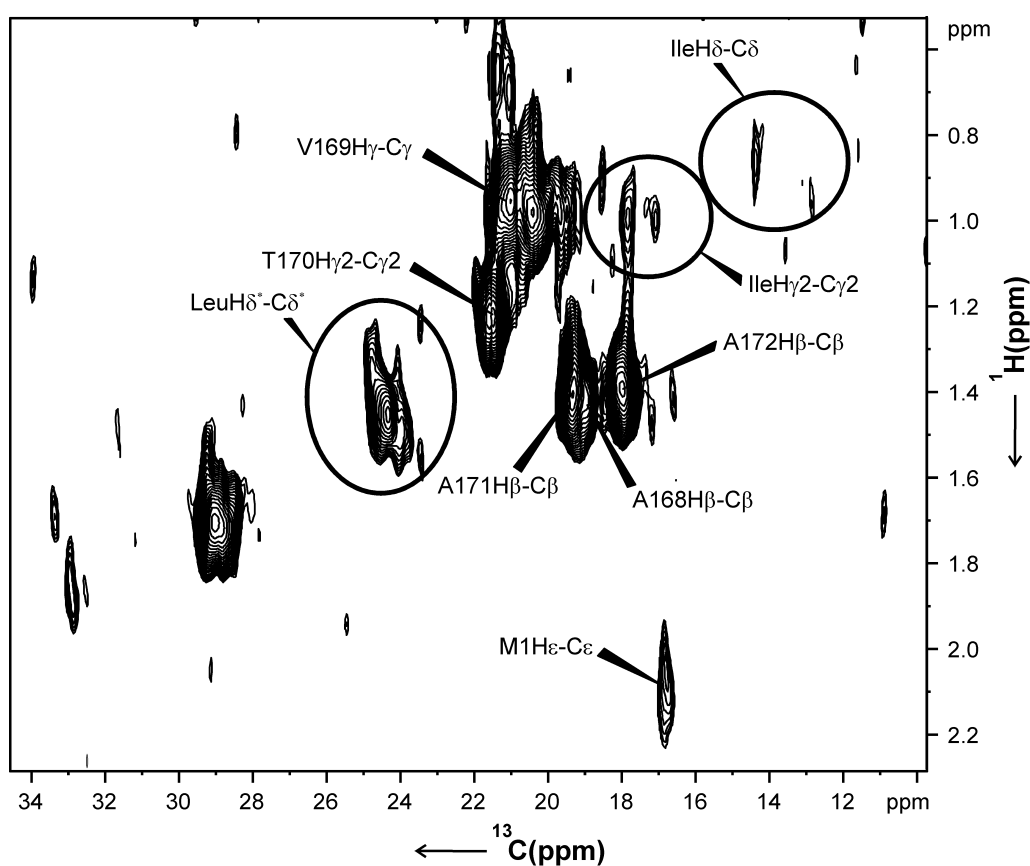
figure 4-9 Contour plot of a 2D ^{1}H-^{13}C INEPT spectrum recorded from U-αB. The methyl-resonances of mobile residues in the termini are observed.

The methyl-resonances of A168, V169, T170, A171 and A172 are observed (the assignments were adapted from Carver *et al.* (54)). Additional Hδ*-Cδ* crosspeaks of leucines appear. From the N-terminus, a crosspeak for Hε-Cε of M1 is observed, whereas only weak Hδ-Cδ and Hγ2-Cγ2 crosspeaks of isoleucines appear.

In summary, the results from the assignment of mobile residues in the termini of αB reveal a mobile C-terminal tail of precipitated full length αB crystallin oligomers and flexibility of M1. The first five residues in the N-terminus could not be assigned in the INEPT spectrum, except M1. The INEPT spectrum shows correlations for methyl-groups of isoleucines, that may indicate flexibility for I3, I5 or I10.

4.2 The topology and chemical shift analysis of the α-crystallin domain and the C-terminal domain of αB-crystallin

4.2.1 Introduction

In the previous chapter, the resonance assignment of full-length αB using 2D and 3D solid-state NMR experiments was demonstrated. In particular, backbone chemical shifts form a rich source of information: secondary structure can be predicted with high confidence from backbone ^{15}N and ^{13}C chemical shifts and ^{15}N chemical shift changes are routinely used to detect sites of intermolecular contacts. In this chapter the chemical shifts of full length αB are used to predict the secondary structure of the α-crystallin domain and the C-terminal domain. In addition, the chemical shifts of the isolated α-crystallin domain of αB10.1 dimers studied by solution NMR in the laboratory of R. Klevit, Seattle, are compared with the chemical shifts of the α-crystallin domain of αB-oligomers derived from solid-state NMR. The observed differences provide information about interaction sites in the oligomer.

4.2.2 Secondary structure of the α-crystallin domain

The secondary structure of the α-crystallin domain in αB oligomers was predicted from the chemical shifts of the Cα, Cβ, C′ and ^{15}N resonances, using the secondary structure prediction programme PECAN(34). Dihedral angles were determined using the software TALOS (33) using the same constraints. The dihedral angles were further used to calculate a structure of the α-crystallin domain together with distance restraints derived from solid-state NMR (see chapter 4.3). The predicted secondary

structure elements are indicated by arrows above the sequence of αB in figure 4-10. The α-crystallin domain in αB has six β-strands numbered β3-β9: the strand numbering follows the convention in the sHSP field (46;47). For the purpose of comparison, the sequences of the homologue αA-crystallin, wHSP16.9 and MjHSP16.5 are also shown in figure 4-10. The secondary structure elements of wHSP16.9 and MjHSP16.5 are indicated by shading the respective sequence elements.

```
                           HR1    80                100              120
human αB crystallin  LRAPS-WFDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGD-VIEVHGKHEERQDEHG
human αA crystallin  FRT---VLDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQDD-FVEIHGKHNERQDDHG
wHSP16.9                ---TAAFANARMDWKE-TPEAHVFKADLPGVKKEEVKVEVEDGNVLVVSGERTKEKEDKN
MJHSP16.5              -IQISGKGFMPISIIEGDQHIKVIAWLPGVNKEDIILNAVGD-TLE-IRAKRSPLMITES
                         β1          β2          β3              β4        β5

                                        140              160              180
human αB crystallin  F--------ISREFHRKYRIPADVDPLTITSSLSSDGVLTVNGPRKQ----VSGPERTIP
human αA crystallin  Y--------ISREFHRRYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHAERAIP
wHSP16.9             DKWHRVERSSG-KFVRRFRLLEDAKVEEVKAGLEN-GVLTVTVPKAE---VKKPEVKAIQ
MJHSP16.5            ERIIYSEIPEEEEIYRTIKLPATVKEENASAKFEN-GVLSVILPKAE-----SSIKKGIN
                         β6          β7          β8     β9                β10

                         195
human αB crystallin  ITREEKPAVTAAPKK
human αA crystallin  VSREEKP--TSAPSS
wHSP16.9             ISG------------
MJHSP16.5            IE-------------
```

figure 4-10 Amino acid sequences of the α-crystallin domains and the C-termini of αB, αA, wHSP16.9 and MjHSP16.5. Secondary structure elements of the α-crystallin domains from αB (arrows), wHSP16.9 (grey) and MjHSP16.5 (grey) are indicated. The secondary structure elements obtained by solid-state NMR data are shown as black arrows above the sequence of αB. The region corresponding to β2 in wHSP16.9 and MjHSP16.5 is labelled Heterogeneous Region 1 (HR1) in αB because these residues exhibit heterogeneity. The β-strands of wHSP16.9 and MjHSP16.5 are indicated by shading the respective sequence elements.

There are two notable differences in αB compared to wHSP16.9 and MjHSP16.5: (i) a contiguous 11-residue β-strand is predicted between β5 and β8 instead of the two

shorter β6 and β7 strands connected by a loop and (ii) heterogeneity of M68 is observed, which is in the middle of the sequence comprising a β2-strand. The 11-residue β-strand is referred to as β6/7.

As analysis of 3D spectra did not yield unambiguous assignments for the HR1 region, we prepared a PFMW-labelled αB sample to assign M68 as well as to provide a starting point for sequential assignment of the adjacent residues. A contour plot of a 2D ^{13}C-^{13}C correlation spectrum with 15 ms spin diffusion mixing time collected on the PFMW-αB is shown in figure 4-11. The average chemical shifts for the Cα, Cβ, and Cγ of methionine are 56.1 ppm, 33.0 ppm and 32.0 ppm, respectively (BMRB (27)). More than five Cα-Cβ correlations are observed for methionine residues in αB, although only two, M1 and M68, are present in the sequence. In chapter 4.1.2.3 the flexible residues in αB were identified by solid-state NMR. These include the N-terminal M1 and the last 8-10 residues in the C-terminus (54). Normally it is difficult to observe flexible residues like M1 by solid-state NMR due to inefficient cross polarisation transfer, so it is likely that all observed crosspeaks in figure 4-11 arise from M68. The Cα shifts vary from 54.5 to 56 ppm and the Cβ shifts vary from 34 to 37.5 ppm. The spread and multiplicity of chemical shifts in the solid-state spectra indicates that M68 is found in multiple distinct environments within the oligomer. The HR1 region is assumed to be involved in building interfaces within oligomers. We expect several types of interfaces exist, consistent with the polydisperse nature of αB. Hence this region in αB is referred to as HR1: Heterogeneous Region 1.
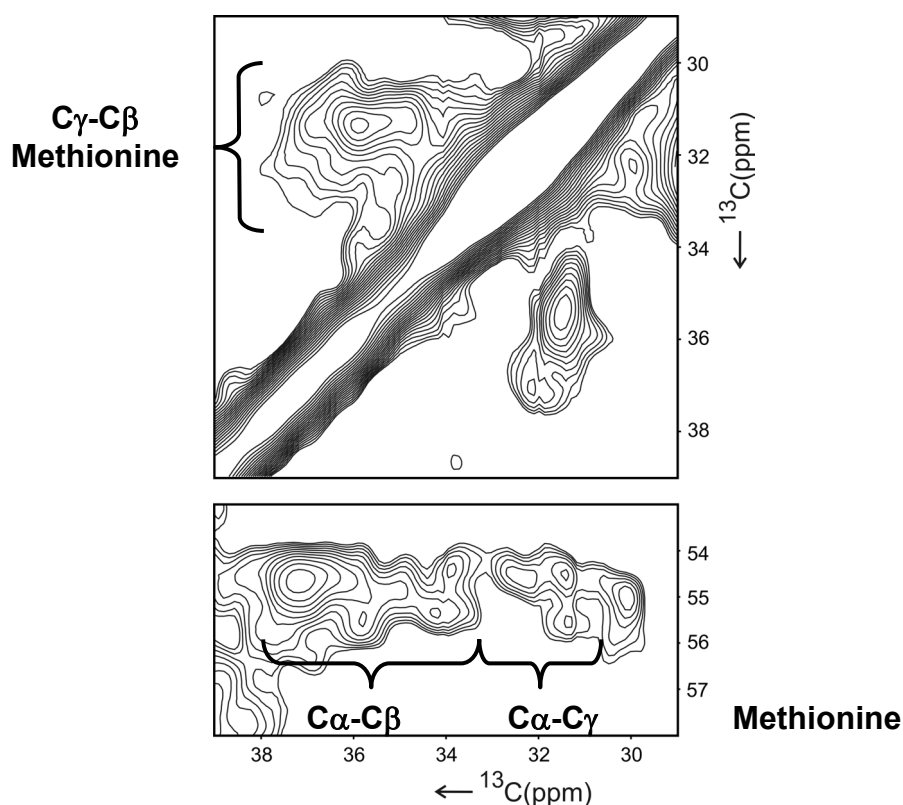
figure 4-11 A contour plot of a 2D ^{13}C-^{13}C PDSD experiment recorded with 15 ms mixing time of PFMW-αB is shown. The plot shows the Cγ-Cβ, Cα-Cβ and Cα-Cγ crosspeaks from methionines. The spectrum was recorded at a field of 14.1 T and with 86 kHz ^{1}H decoupling.

Additional heterogeneity is observed for V77 which is in β3. Two signal sets for V77 are found at δ^{15}N=117.5 ppm and δ^{15}N=119.5 ppm. In contrast to M68, which shows multiple Cα and Cβ shifts, the Cα and Cβ shifts of V77 vary by a little around 59.5 ppm and 34.5 ppm, respectively. Hence, V77 is thought to have a discrete backbone conformation, since the Cα and Cβ resonances of amino acids are sensitive to changes of the backbone dihedral angles.

The lack of a β2-strand in αB10.1 corroborates the observed heterogeneity in HR1. ^{15}N backbone dynamics measurements were performed on αB10.1 with solution-state NMR by Ponni Rajagopal at the University of Washington, Seattle (USA). Negative to low heteronuclear-NOE values (-1.0 to 0.6) characterize the HR1 and neighbouring residues. Model free analysis yields low order parameters (S^2 = 0.4-0.6) for this region indicating increased flexibility of these residues in αB10.1 (52).

4.2.3 The topology of the β-sheets in the α-crystallin domain of αB

To determine the topology of β-strands in αB, a combined approach was used: (i) specific recoupling of Cα carbons in β-sheets and (ii) the evaluation of dihedral angles. From dihedral angles and using the achieved assignments, six β-strands were identified in αB. A common structural feature of β-sheets is the parallel or antiparallel packing of strands. The antiparallel packing is more stable than parallel packing due to the planar hydrogen bonds between the strands. In figure 4-12a, a section of a β-sheet is shown, exhibiting the hydrogen bonding patterns of amino acid i+1 with j-1 and i-1 with j+1. The amino acids i and j can form hydrogen bonds to another strand, thus extending the sheet (not shown). The side-chains of opposite residues in the sheet point into the same direction, i.e. the side chains of residue i-1, j+1 and of residues i, j face up or down, respectively (see figure 4-12). The side chains and Hα atoms of amino acids involved in hydrogen bonding with each other point away from each other (i.e. j+1 and i-1), whereas side-chains and Hα atoms of amino acids not involved in hydrogen bonding with each other point towards each other (i.e. i and j). This periodicity is characteristic for β-sheets and is used to identify the residues involved using typical ^{13}C-^{13}C crosspeak patterns in solid-state NMR spectra.

The proximity of Cα spins allows the identification of typical Cα-Cα crosspeaks from 2D ^{13}C-^{13}C correlation experiments. Unfortunately, the Cα-Cα crosspeak region in 2D ^{13}C-^{13}C spectra recorded on U-αB using ^{13}C-^{13}C spin diffusion is highly congested. Due to the spectral crowding, sequential crosspeaks from the Cα atoms of residue j with j+1 and j-1 and residue i with i-1 and i+1 are hard to distinguish from inter-strand crosspeaks (Cα of j with i etc.).

In contrast to the PDSD which utilises direct ^{13}C-^{13}C spin-diffusion, the CHHC experiment employs indirect ^{13}C-^{13}C recoupling via the bound ^{1}H spins. Deuteration of the exchangeable protons in the sample (by placing the protein in D_2O) and fine-tuning of the experiment (see below) mainly yields interstrand Cα-Cα correlations between residues i and j in the Cα-region of a 2D ^{13}C-^{13}C CHHC spectrum because only their Hα atoms are in close proximity. Subsequently, these correlations can be used to determine the topology of β-strands in αB.

a)

b)
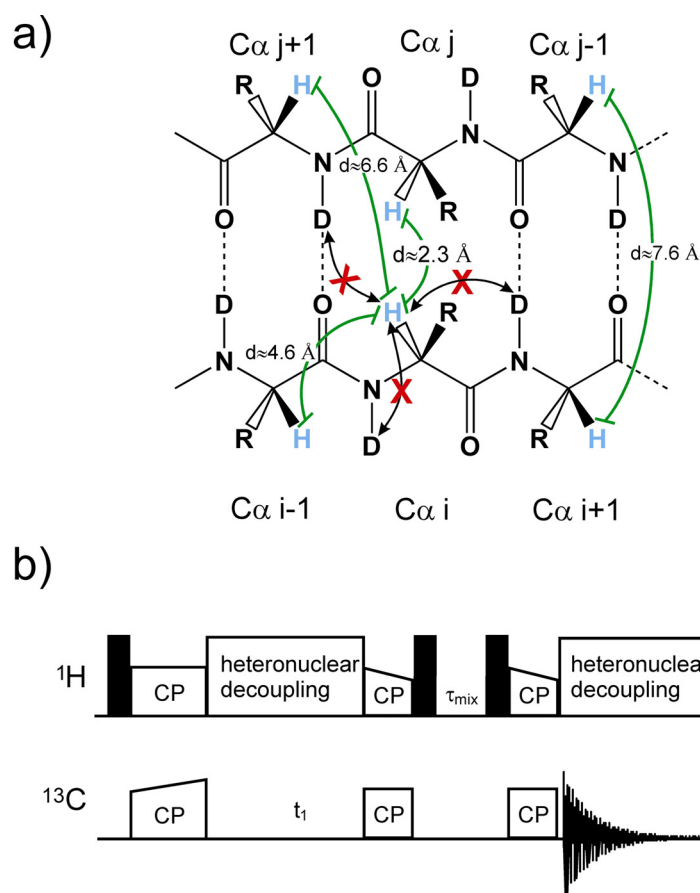
figure 4-12 In panel a), two strands in a β-sheet are shown together with typical ^{1}H-^{1}H distances detected in CHHC transfer experiments (green lines). The red crosses indicate ^{1}H-^{1}H-transfer that is blocked when protons are exchanged with deuterons at exchangeable sites. In panel b) the CHHC pulse sequence is shown that is used to record 2D ^{13}C-^{13}C correlation spectra involving proton mixing(96). Dependent on the used mixing time, the experiment only correlates ^{13}C spins for which the attached ^{1}H spins are in close proximity.

In figure 4-12b, the pulse sequence of a CHHC experiment is given (96). The experiment starts with a ^{1}H-^{13}C CP followed by an indirect evolution time of ^{13}C spins. The magnetisation is then transferred back to the protons. The ^{1}H magnetisation generated is stored along the z-axis for a ^{1}H-^{1}H spin diffusion mixing time, τ_{mix}. After polarisation exchange of protons, another CP step transfers the magnetisation to the ^{13}C spins for detection. The CHHC experiment starts with a long CP (i.e. 1.5 ms) and generates magnetisation on all ^{13}C spins of the protein. After the t_1 evolution period, a contact time of 50 μs for the second CP guarantees mostly one-bond transfer from Cα to Hα and from Cβ to Hβ etc. Carbonyl carbons do not have directly bonded protons and are therefore not considered. The ^{1}H-^{1}H distances given in figure 4-12a give an indication of the spin-diffusion mixing time which should be used. Sequential

Cα-Cα distances are usually ~3.9 Å and the corresponding Hα-Hα distances are 4.6 Å. The interstrand distances between Cα nuclei from residues i+1 and j-1 and i and j are ~5.6 and ~4.2 Å respectively. For these ^{13}C spins, the Hα distances differ substantially: 2.3 Å for i and j but 7.6 Å for i+1 and j-1. Transfer between Hα atoms of residues i and j (d=2.3 Å) is established by using a short mixing-time of 40-50 µs. Transfer over a distance of more than 4 Å is rarely observed at a 40 µs mixing time. Consequently ^{1}H-^{1}H transfer is avoided between sequential and interstrand i+1/j-1 amino acids at short mixing times. To prevent coupling to H^N and relayed transfer via H^N (red crosses in figure 4-12a) to sequential and/or interstrand ^{13}C spins, the H^N protons were exchanged with deuterons by purifying and precipitating the sample from D$_2$O. The third CP contact time is again chosen to be 50 µs (see above) for one bond transfer (i.e. from Hα to Cα). In summary, the CHHC experiment provides Cα-Cα crosspeaks exclusively between ^{13}C spins of which the bound ^{1}H spins are close to one another, with the choice of experimental parameters as written above. Specifically, for β-sheets these crosspeaks predominantly arise between Cα atoms in close contact across the strands (residues i and j in figure 4-12a).

The description of the αB protomers within oligomers requires the determination of whether a β-strand is part of an inter- or intramolecular β-sheet. Therefore, uniformly ^{13}C, ^{15}N-labelled αB oligomers were disassembled into unfolded monomers using a chaotropic reagent and mixed under denaturing conditions with 80% unlabelled αB. The mixture containing 20% U-^{13}C, ^{15}N-labelled and 80% unlabelled αB was finally refolded in a desalting column, yielding oligomers formed from 20% labelled- and 80% unlabelled αB. αB oligomers are known to form stable aggregates, even in solutions containing 1-5 M of chaotropic agents. However, in the presence of 6M Guanidinium Chloride, αB oligomers were found to dissociate and αB was exclusively detected in the monomeric form. Analytical gel filtration chromatography was performed under denaturing conditions (6M Guanidinium Chloride) to assure the complete disassembly of αB oligomers. The 20 kDa αB monomers were detected together with the 25 kDa marker protein. The 20% U-^{13}C, ^{15}N-labelled monomers in the oligomer less likely have a labelled neighbour molecule and hence intermolecular crosspeaks are not observed. The αB oligomers prepared from 20% U-^{13}C, ^{15}N-labelled monomers and 80% unlabelled monomers will be referred to as dilute-^{13}C, ^{15}N αB.
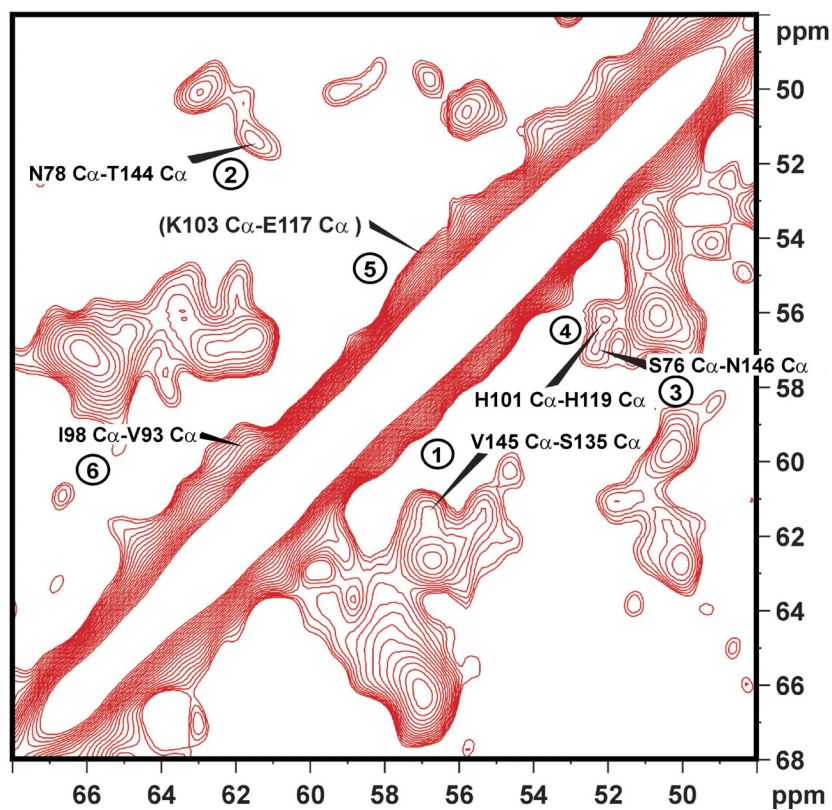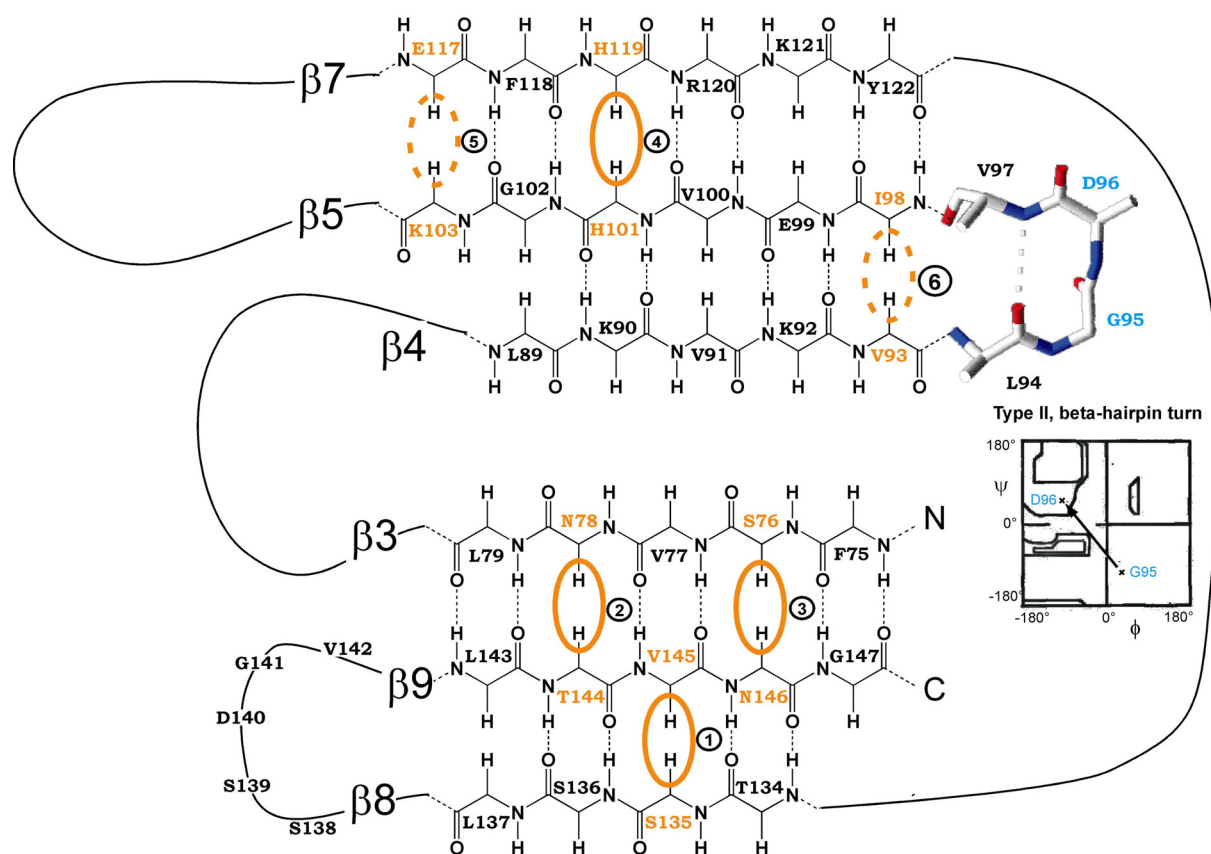
figure 4-13 Top: the topology of β-strands 3-9 of the α-crystallin in αB-crystallin. The dihedral angles for G95 and D96, derived from TALOS, are plotted in the Ramachandran Plot. Bottom:

^{13}C-^{13}C-CHHC spectrum recorded from dilute-^{13}C, ^{15}N αB-crystallin. Cα-Cα crosspeaks between residues i and j are indicated by numbers.

The topology of β-strands 3-9 in the α-crystallin domain of αB was determined by assigning interstrand Cα-Cα crosspeaks of residues i and j as described above using the spectrum shown in figure 4-13 of dilute-^{13}C, ^{15}N αB: The topology consists of two β-sheets with strands β3-β9-β8 forming one sheet and strands β4-β5-β7 forming the second sheet. Crosspeaks connecting β3, β8 and β9 are S135Cα/V145Cα (labelled 1), T144Cα/N78Cα (labelled 2) and N146Cα/S76Cα (labelled 3). The intramolecular connection between β5 and β7 is supported by the H119Cα/H101Cα crosspeak at 55.6 ppm in F_1 and 52.3 ppm in F_2 (labelled 4 in figure 4-13). Crosspeaks of Cα spins differing 1 ppm or less in their chemical shift are near the diagonal and their presence only appears as a convexity of the diagonal, i.e. K103Cα/E117Cα (labelled 5) or I98Cα/V93Cα at 59.6 ppm in F_1 and 61 ppm in F_2 (labelled 6). Dihedral angles from TALOS predict a type 2 turn between β4 and β5. The typical dihedral angle distribution of G95 and D96 is indicated in the Ramachandran Plot in figure 4-13.

wHSP16.9 forms a β-sandwich with sheets β2-β3-β9-β8 and β4-β5-β7 comprising the two faces of the sandwich. The results here show that αB and wHSP16.9 have a similar topology, except for the β2- and β6-strand. In addition, our results are consistent with site directed spin labelling (SDSL) studies of the topology and arrangement of the β-strands in the homologue αA-crystallin (which shares 60% sequence homology with αB) (95;97-100). Sets of mutants were generated, in which consecutive amino acids were replaced with cysteine. After attaching a nitroxide spin label to the side-chain of the cysteine, electron spin resonance (EPR) measurements were performed on the set of proteins. From EPR characteristics, the secondary structure could be deduced, i.e. by studying the accessibility and mobility of the label, which varies in β-strands, with a periodicity of two (i.e. low, high, low, high etc.). Moreover, in double mutants, containing two simultaneously spin-labelled residues, the dipolar coupling of the two labels allowed the determination of the orientation of two strands with respect to each other. A topology of the strands in αA-crystallin similar to that of wHSP16.9 was found. The only differences are a missing β6-strand, and β7 adopts an extended β-strand conformation. With respect to the sequence corresponding to EMRLE (β2-strand) in the homologue αA-crystallin, the EPR results

studies were interpreted as a β-strand conformation (99). However, the site-directed spin labelling studies did not reveal the periodicity of the electron spin resonance usually observed for β-strands. This periodicity was found in all other strands of the α-crystallin domain of αA-crystallin, but not for the proposed β2 strand. In addition, dipolar couplings between the β2 and β3 strand were measured and it was found that these were weaker than usual for couplings between β-strands. But nevertheless a proximity of the 67-70 fragment to β3 was proved. In the light of our results, HR1 of αB may be in proximity of β3 while being in an heterogeneous environment (see paragraph 4.2.2).

4.2.4 Chemical shift differences in the α-crystallin domain of αB crystallin oligomers and αB10.1

In this section, chemical shifts of the α-crystallin domain derived from solid-state NMR of the oligomer and from solution NMR of αB10.1 are compared and analysed in terms of intermolecular interactions. Figure 4-14a, b, c, and d show histograms of chemical shift differences (δ_{solid}-$\delta_{solution}$) between solid-state shifts (δ_{solid}) and solution-state shifts ($\delta_{solution}$). ^{15}N chemical shift differences are shown in a; ^{13}Cα, ^{13}Cβ, and ^{13}C′ chemical shift differences are depicted in b, c, and d, respectively. The mean chemical shift differences are 1.3 ± 1.4, 0.6 ± 0.7, 0.5 ± 0.4, and 0.6 ± 0.6 ppm for ^{15}N, ^{13}Cα, ^{13}Cβ, and ^{13}C′, respectively. The mean values are in the range observed in other comparisons of solid versus solution-state data, e.g., SH3 domain (32), GB3 (101), Ubiquitin (102), Zinc metalloproteinase (31), to name a few. Most importantly, negligible differences, i.e. differences smaller than the mean value, are observed for a majority of residues indicating that the structure of the isolated α-crystallin domain in solution is preserved in the oligomer. Differences smaller than the mean value are clustered in certain regions. To facilitate interpretation of these observations, the sum of the absolute values of ^{15}N, ^{13}Cα, ^{13}Cβ and ^{13}C′ chemical shift differences ($|\delta_{solid}$-$\delta_{solution}|^{N,C\alpha,C\beta,C'}$) were computed and plotted versus residue number as shown in the histogram in figure 4-15. It can be discerned from the histogram that there are clusters of residues exhibiting differences smaller than one standard deviation (S.D.) (2.5 ppm) from the mean.
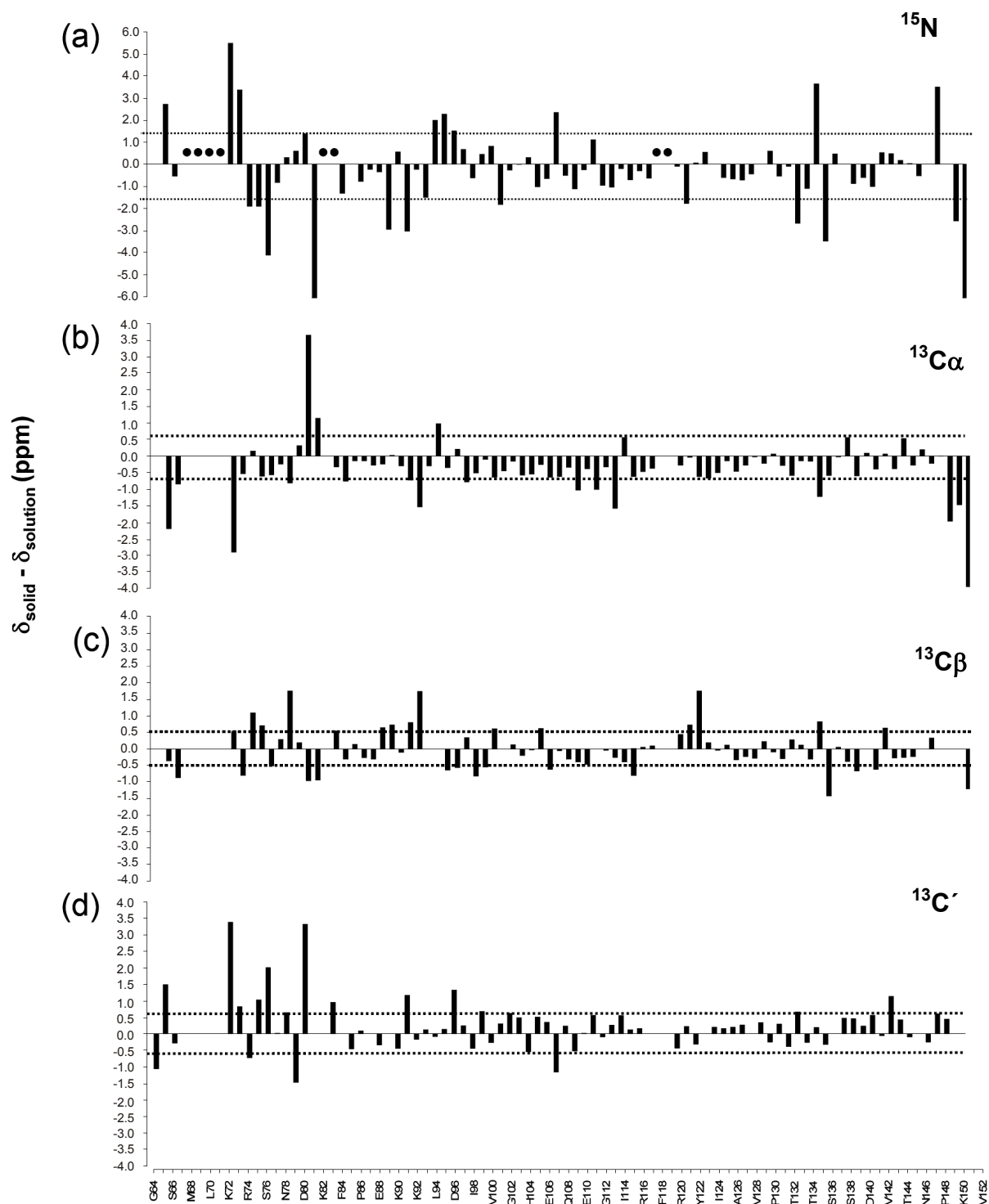
figure 4-14 Histogram of chemical shift differences (δ_{solid} - $\delta_{solution}$) plotted versus residue number. The chemical shift differences of ^{15}N, Cα, Cβ, and C$^\prime$ resonances between solid (αB) and solution-state (αB10.1) data are shown in the four panels. The dashed horizontal lines represent the mean values. • represent residues with missing assignments.
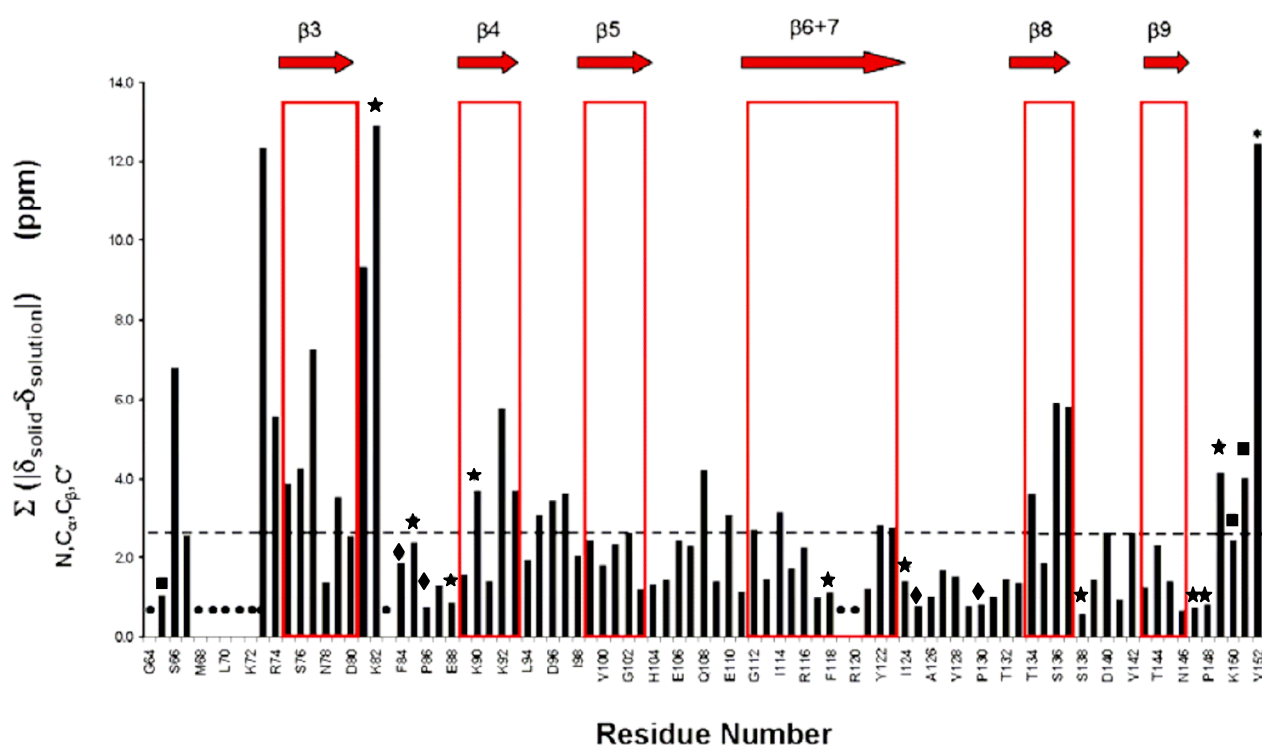
figure 4-15 Sum of the absolute values of the chemical shift differences of ^{15}N, ^{13}Cα, ^{13}Cβ, and ^{13}C′ resonances ($|\delta_{solid}-\delta_{solution}|^{N,C\alpha,C\beta,C'}$) plotted versus residue number. The secondary structure is shown at the top of the histogram. The dashed horizontal line represents the mean value. Residues that have either some or all resonance assignments missing are indicated as follows:

● : ^{15}N , ^{13}Cα, ^{13}Cβ and ^{13}C′ missing; ♦ ^{15}N missing; ★^{13}C′ missing, ■ ^{13}C′ and ^{13}Cβ missing

The location of these clusters is mapped onto the topology of the α-crystallin domain of αB in figure 4-16 and a colouring scheme ranging from yellow to red was used to depict the chemical shift changes. Yellow denotes the least and red the maximum chemical shift change. A large number of "hot spots" corresponding to maximal chemical shift differences cluster to the edge strands, β3, β4, and β8, modelling the environment of the α-crystallin domain in the oligomer. In wHSP16.9, N-terminal residues are in close proximity to the edge strand β2. The conserved IXI/V motif in the C-terminus binds at the hydrophobic groove formed by the edge strands β4 and β8 in wHSP16.9 and MjHSP16.5. The interactions of the N- and C-termini with the α-crystallin domain in wHSP16.9 are inter-subunit in nature and thus give rise to the higher order oligomeric structure (58). Analogous to wHSP16.9, the edge strand β3 in αB could be in close proximity to sites in HR1 as well as the rest of the N-terminal region. β4 and β8 could be putative sites of interactions with the C-terminal region in oligomeric αB that is absent in the isolated domain. Significant perturbations (> 4

ppm) are observed for the loop-segments between (i) HR1 and β3 (residues D73 and R74) and (ii) β3 and β4 (residues V81 and K82). Since some assignments are missing in these segments, a rigorous comparison between the solid and solution chemical shift data cannot be made. The missing resonance assignments are indicated in figure 4-15.
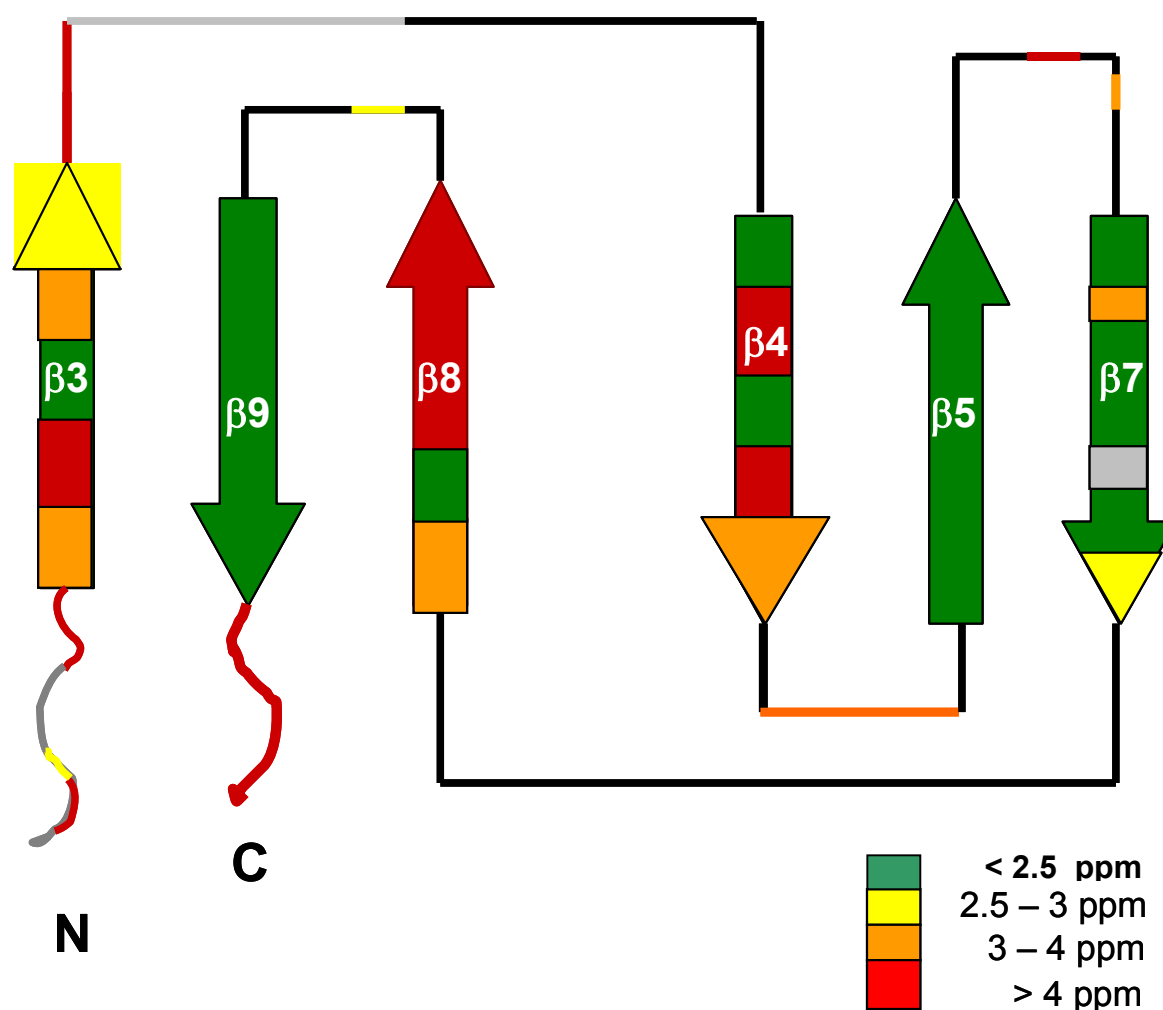
figure 4-16 Topology of the α-crystallin domain from αB illustrating regions that exhibit chemical shift perturbations in the context of the oligomer. Colour coding of the perturbed regions is as follows: 2.5-3 ppm: yellow; 3-4 ppm: orange; > 4 ppm: red. Edge strands β3, β4, and β8, some loop regions as well as the N and C-termini experience chemical shift perturbations. No information is available for regions coloured grey due to missing assignments.

The "cold spots" corresponding to chemical shift similarities cluster in β5, β7, and β9 as well as loops β5-β7, β7-β8 and β8-β9. Lack of differences in β5 and β9 is

not unexpected, since they are in the middle of the sheet (figure 4-16). There are some minor perturbations in $\beta 7$, but most remarkably, there is a concentration of cold spots as well, indicative of a similar environment in both the isolated domain and the oligomer. This can be rationalized in terms of a dimer model where $\beta 7$ from protomer A stacks in an antiparallel fashion to another $\beta 7$ from protomer B. Thus, residues in $\beta 7$ would be buried in the dimer interface and will not be exposed to a different environment in the oligomer. These results are well in line with SDSL experiments from the homologue αA-crystallin evidencing the dimer interface in the extended $\beta 6/7$-strand.

4.3 Structure of the α-crystallin domain in αB-crystallin

In this chapter, the collection of distance restraints for structure calculations of the α-crystallin domain in αB will be described. Furthermore, preliminary results from the R120G mutant will be shown and a model of the dimeric subunit will be presented based on distance restraints from solid-state NMR data.

4.3.1 Collection of distance restraints from spectra recorded from 2G- and 1,3G-αB

In this paragraph, the collection of long-range distance restraints for a structure calculation using SolARIA will be delineated. The programme SolARIA automatically assigns peak lists from 3D NCACX, 3D NCOCX and 2D ^{13}C-^{13}C spectra to generate distance restraints. In doing so, each assignment option is weighted against the corresponding distance obtained from the intermediate structure to provide a "likeliness" for this assignment option. If more assignment options are possible for a given crosspeak, the corresponding distances are combined into a single "ambiguous" distance. The 1,3-^{13}C-glycerol and 2-^{13}C-glycerol labelling pattern is implemented in the software and hence the software "knows" which assignment options are possible for the two types of labelling schemes. The separation between two restrained nuclei is preassigned to a distance range of 2.5-6.5 Å. The current version of SolARIA does not take peak volumes into account since these are difficult to interpret in spin-diffusion spectra (see chapter 1.3.5).

Not all crosspeaks contain relevant information: sequential short range Cα[i]-Cα[i-1], Cα[i]-C´[i-1], and C´[i-1]-C´[i] crosspeaks lack relevant structural information because their separation is almost invariable. These crosspeaks were assigned manually and excluded from the structure calculation, to prevent their misinterpretation as long-range transfer. The secondary structure elements are defined by characteristic long-range Cα-Cα, Cα-Cβ, Cα-C´, and C´-C´ distances derived from 3D NCACX and NCOCX spectra. The system was additionally restrained using dihedral angle restraints (chapter 4.2). The quality of the lowest-energy structures was assessed

after each calculation by analysing the backbone root mean square deviation (rmsd) in the family of structures, distance restraint violations, close contacts, dihedral-angle violations and the presence of optimally folded secondary structure elements. SolARIA-derived distance assignments are analysed interactively and verified using the experimental data. During the structure refinement, unambiguous peaks are used step-by-step as preset distances to restrain the system prior to the automatic assignment. H-bonds were carefully added for optimising the secondary structure elements in cases where typical crosspeak patterns between the involved residues are found (i.e. C´[i]-Cα[j-1] C´[i]-C´[j-1]) and typical backbone torsion-angles were present. A selection of distance restraints from the β-sheet composed of β3, β8 and β9 is shown in figure 4-17. The restraints are numbered **1**-**14**, and were automatically assigned by SolARIA. Dihedral-angle restraints further define the turns between the β-strands. A three-residue β-hairpin connects β8 with β9. The typical ψ- and ϕ-angle distribution for S138-V142 as predicted by TALOS is shown on top of figure 4-17, in the Ramachandran plot to the right. The structure of the turn in αB is shown to the left. S138 is the last residue in β8 and V142 is the first residue in β9. S139 adopts the right-handed α-helical conformation while D140 lies in the bridging region between α-helix and β-sheet. The third residue (G141) adopts dihedral angles close to the left-handed helical conformation. The hairpin conformation is further corroborated in strip **a** to the right in figure 4-17, extracted from an NCACX experiment at the Cα frequency of D140 (52.8 ppm in F_2). The spectrum was recorded with 500 ms spin-diffusion mixing from 1,3G-αB. At 116.7 ppm in F_1 (^{15}N), two crosspeaks are observed that correlate the Cα of D140 with the Cβ of S138 (labelled **1**) and the C´ (labelled **2**) of S138 at 65.5 ppm and 177.3 ppm in F_3, respectively. Strip **b** is extracted from the same spectrum at the ^{15}N frequency of T144 (124.2 ppm) and shows correlations for T144 in β9. At 61.4 ppm in F_2 and 51.6 ppm in F_3, there is the interstrand correlation connecting N78 Cα in β3 with the Cα of T144 (labelled **3**). The magenta ellipse in the sheet model indicates that this correlation was also observed in the CHHC spectrum in chapter 4.2. The crosspeak (labelled **4**) for T144 at 61.3 ppm in F_2 and ~53 ppm in F_3 was used as an ambiguous distance restraint correlating T144 with either L137 Cα or L79 Cα. The Cα resonances of both leucines is ~53 ppm and within the error range of the resonance assignment. The separation between T144 Cα and the Cα of the two leucines is almost equal (Dashed green

arrows numbered with **4** in the sheet model). Finally, strip **c**, extracted from the NCACX spectrum of 1,3G, shows the ^{15}N plane of L137. At 52.9 ppm in F_2 and 61.3 ppm in F_3 a strong correlation (labelled **5**) can be observed indicating proximity between T144 Cα and L137 Cα.
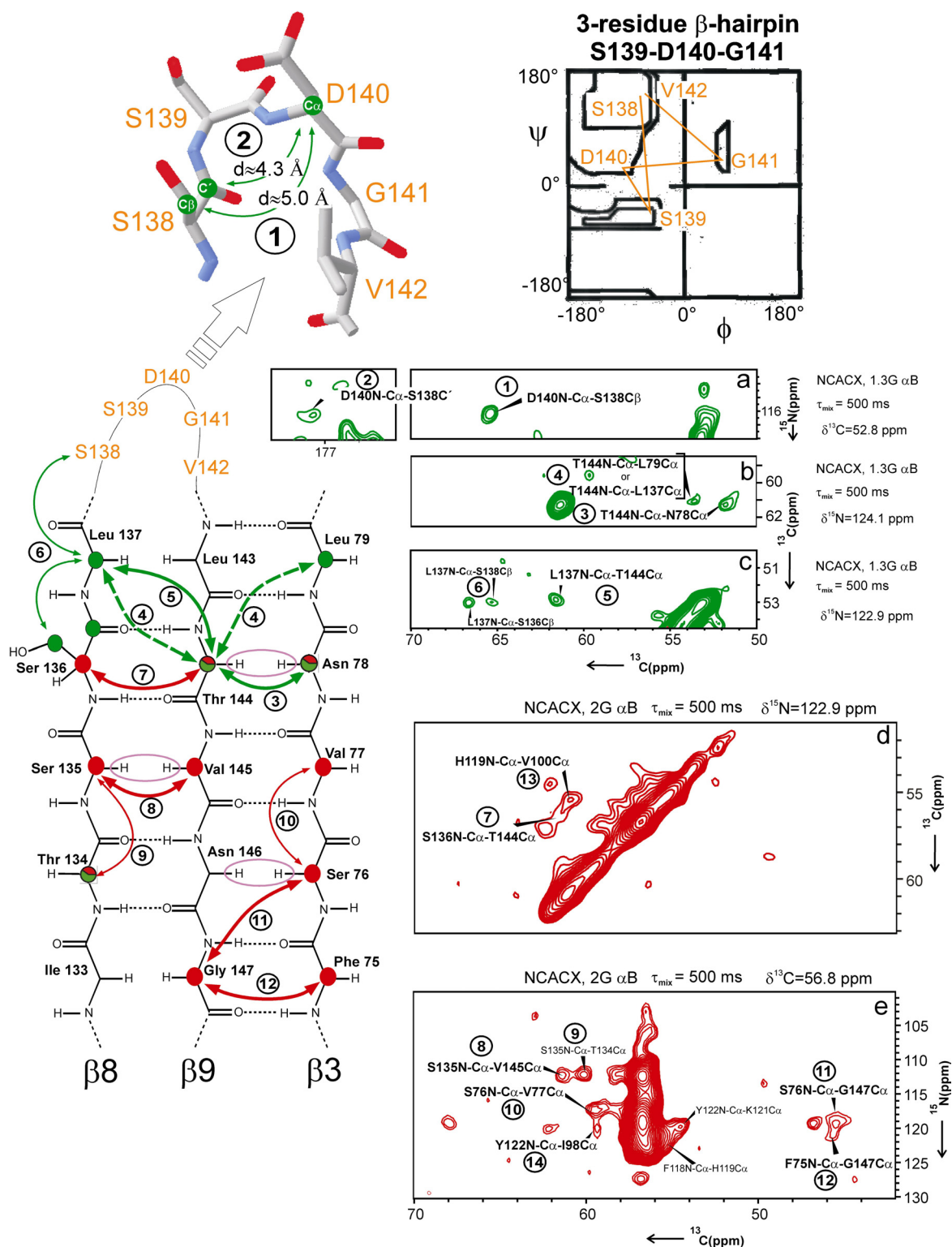
figure 4-17 Schematic representation of the β-sheet composed of β3, β8 and β9 (left). On top,

the S138-V142 turn connecting β8 with β9 is shown. The ψ and φ dihedral angle distribution in the Ramachandran plot (top right) is characteristic for a 3-residue beta hairpin. Distance restraints are indicated by arrows. Interactive validation of the distances, by matching the resonances of the experimental data, is shown in strips from 3D NCACX spectra to the right. CHHC restraints (chapter 4.2) are indicated with magenta ellipses. The colour of the contours and of the nuclei in the sheet-model indicates the labelling: Red and green represent nuclei labelled in 1,3G- and 2G-αB-crystallin, respectively. Fractionated coloured circles represent nuclei labelled in both samples. The experiment type, the mixing time and the ^{13}C or ^{15}N frequency at which the strip is extracted, is indicated on the right or on top of the strip.

The two crosspeaks numbered with **6** indicate sequential Cα-Cβ correlations of the Cα of L137 with the Cβ of S136 (66.4 ppm in F_3) and S138 (65.5 ppm in F_3). Strip **d** is extracted at the ^{15}N frequency of S136 (118.3 ppm) from an NCACX spectrum with 500 ms mixing time recorded from 2G-αB. The inter-strand correlation at 56.6 ppm in F_2 and 61.4 ppm in F_3 (labelled **7**) together with the restraints from the NCACX recorded from 1,3G-αB, provide evidence for a spatial arrangement of T144 in proximity to S136, N78, L137 and perhaps L79 (ambiguous from **b**). Strip **e** shows an NCACX plane extracted at 56.8 ppm in F_2 (^{13}Cα-plane) and is recorded from 2G-αB with a 500 ms mixing time. Interstrand correlations for S76 (118 ppm in F_1) and F75 (121.1 ppm in F_1) are observed with the Cα of G147 at 45.4 ppm in F_3 (indicated with red arrows and numbered with **11** and **12**). The sequential Cα-Cα crosspeaks (labelled **9** and **10**) connecting S76 with V77 and S135 with T134 are excluded from the peak list used for structure calculation. The interstrand crosspeak, S135Cα-V145Cα (labelled **8**) at 112.1 ppm in F_1 and 61.3 ppm in F_3, corroborates the proximity of S135 Cα and V145 Cα, as previously observed in the CHHC spectrum depicted in chapter 4.2, and connects β8 with β9.
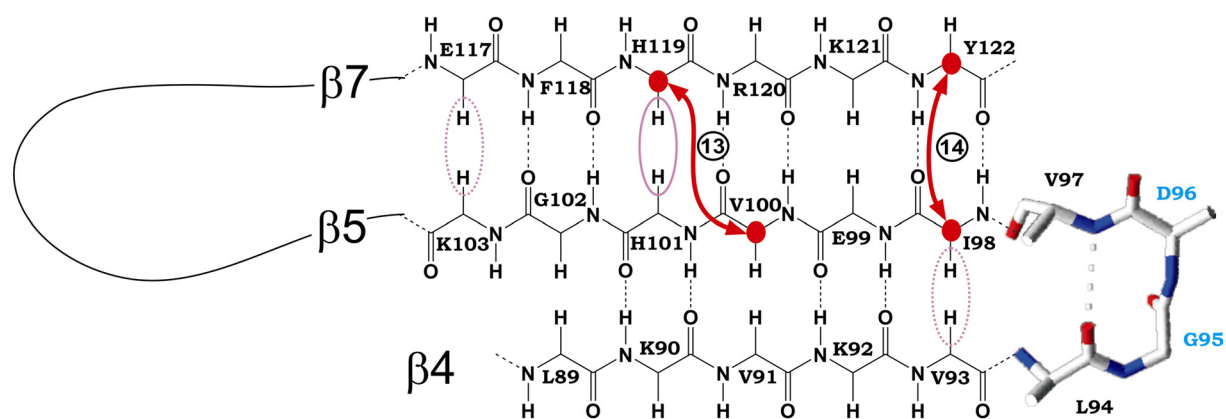
figure 4-18 Schematic representation of the β-sheet built from β3, β8 and β9. Distance restraints assigned using SolARIA are indicated with arrows and the numbers link to crosspeaks shown in figure 4-17. Magenta ellipses indicate CHHC restraints. For further details see text.

The two interstrand crosspeaks in the strips **d** and **e** derive from correlations of H119Cα/V100Cα and Y122Cα/I98Cα, numbered **13** and **14** respectively. These correlations are indicated in the schematic view in figure 4-18 with red arrows and numbered as in the strips **d** and **e**, in figure 4-17. In both figures 4-17 and 4-18, restraints derived from the CHHC experiment (chapter 4.2) are indicated, highlighted with magenta ellipses. CHHC crosspeaks that are near the diagonal are indicated with dashed ellipses.

Interstrand backbone-backbone correlations in β-sheets and typical Cα[i]-Cα[i+4] correlations in helices are indispensable for defining secondary structure elements and are essential for obtaining high-quality protein structures. 3D NCACX and NCOCX experiments are ideally suited to provide such restraints. In addition, long-range correlations between side-chain nuclei have to be collected to restrain the geometrical orientation of secondary structure elements with respect to each other. For instance, the backbone-backbone inter-sheet distance of β-sheet sandwiches is ~10 Å, which is difficult to detect using PDSD mixing. Hence, to arrange two β-sheets with respect to each other, side-chain correlations between the two sheets are required. These restraints can be obtained from 2D ^{13}C-^{13}C correlation spectra recorded from 2G-αB or 1,3G-αB employing spin-diffusion mixing times of 500 ms. Hydrophobic interactions may be studied by examination of methyl-methyl crosspeaks from 2D PDSD spectra recorded from 1,3G-αB, since all methyl groups are labelled in this sample (figure 4-19a). Correlations involving Cδ1/2 of leucine and Cγ1/2 of valine are treated as stereo-chemically equal in a floating assignment

procedure and indicated as Cδ* and Cγ*. In 2G-αB, the Cα-Cβ of valine and Cβ-Cγ of isoleucine and leucine are labelled so that especially structure relevant interresidual Cβ-Cβ, Cβ-Cγ and Cβ-Cα interactions of hydrophobic residues involved in van der Waals interactions are observed (figure 4-19b).
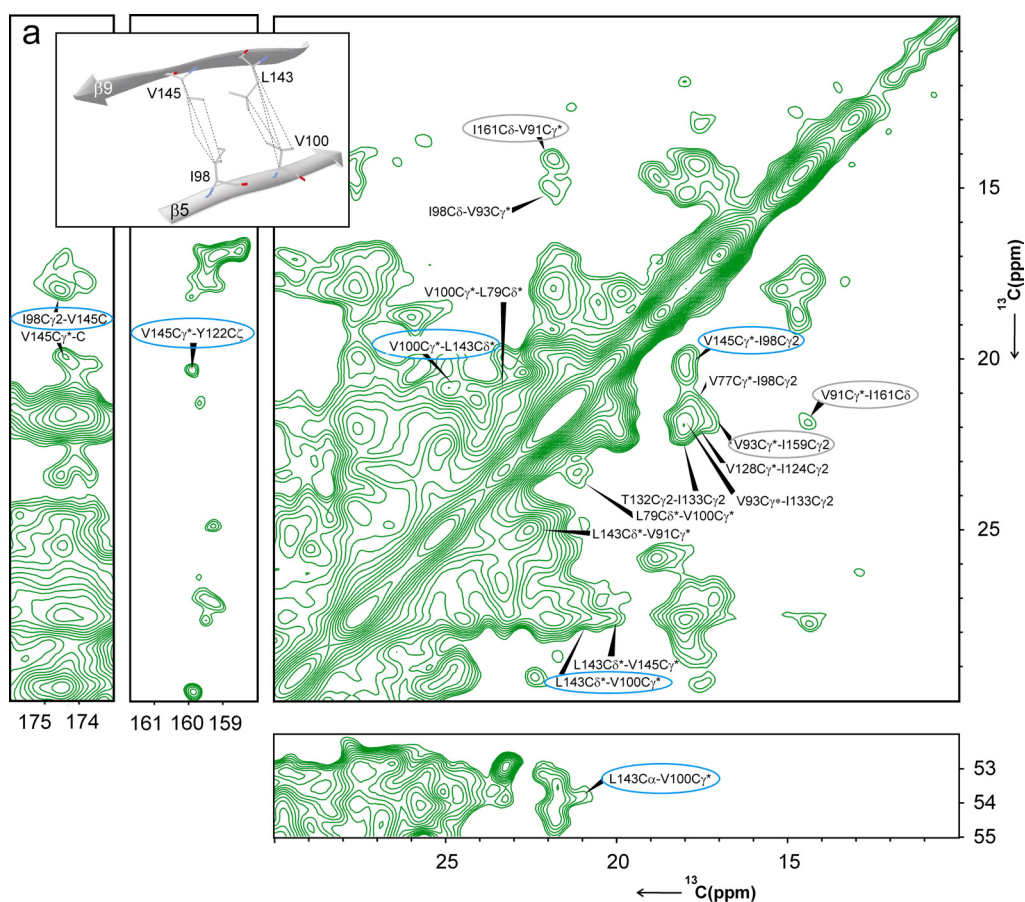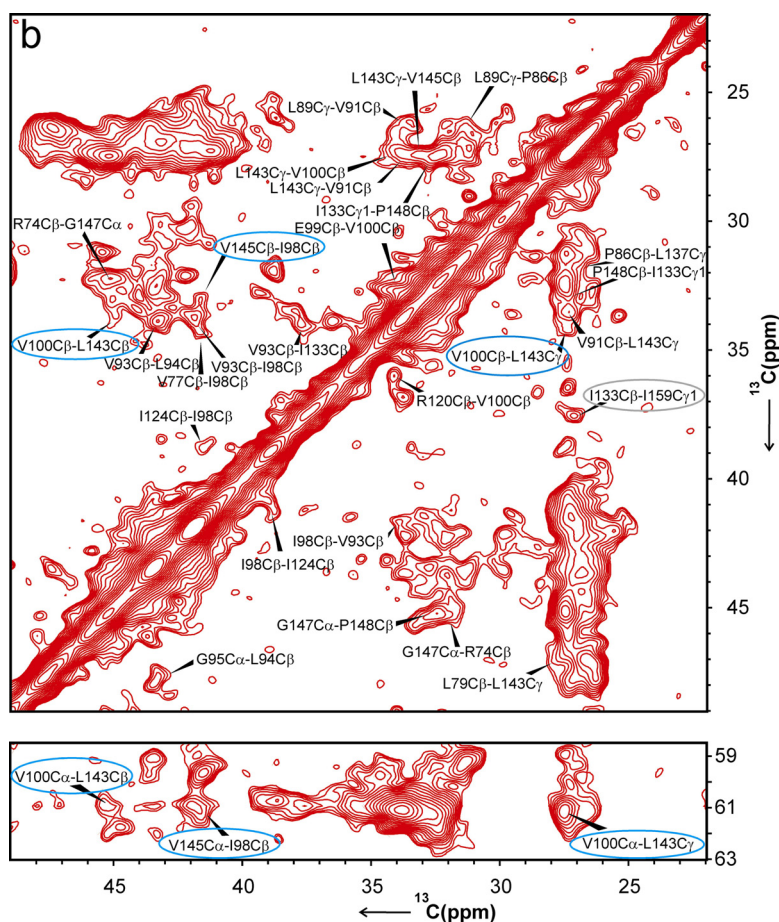
figure 4-19 Contour plots of 2D ^{13}C-^{13}C PDSD experiments using 500 ms mixing time recorded from 1,3G-αB (a) and 2G-αB (b). The orientation of β5 with respect to β9 is shown in the insert in (a). Correlations involving Cδ1/2 of leucine and Cγ1/2 of valine are treated stereo-chemically equal in a floating assignment procedure and indicated as Cδ* and Cγ*.

Crosspeaks involving hydrophobic residues in both β-sheets of αB characterise the relative geometrical orientation of the two sheets. Several of such crosspeaks have been assigned unambiguously and uniquely defined the tertiary topology; crosspeaks that could not be assigned unambiguously had at least one assignment option that supported the tertiary structure and hence formed a consistent network together with the unambiguous assignments. The Cγ2 of V145 on β9 and I98 on β5 have unique chemical shifts at 20.1 ppm and 18.0 ppm, respectively. A crosspeak between both is observed in the 2D PDSD spectrum shown in figure 4-17a recorded from 1,3G-αB. In addition crosspeaks between the Cγ2 of I98 and the C´ of V145 (174 ppm) and between the Cγ2 of V145 and the Cζ of Y122 (β7, 159.8 ppm) are observed in the same spectrum. Y122 (β7) is opposite to I98 (β5) on a different strand in the same β−sheet (c.f. figure 6.2); both side chains point out of the β−sheet plane in the same

4-101

direction and show crosspeaks to the Cγ2 of V145 in β9. In the 2D PDSD spectrum recorded from 2G-αB (figure 4-19b) additional crosspeaks between the Cβ (41.5 ppm) and the Cα (61.3 ppm) of V145 and I98 corroborate this proximity between the two side chains. The V100 (β5) is two residues upstream of I98, while L143 (β9) is two residues downstream of V145. Crosspeaks involving the Cα (61 ppm)/Cβ (34.1 ppm) of V100 and the Cβ (45.4 ppm)/Cγ (27.7 ppm) of L143 are observed in the spectrum shown in figure 4-19b recorded from 2G-αB. Additional V100Cγ*-L143Cδ*/Cα crosspeaks are found in the spectrum recorded from 1,3G-αB, shown in figure 4-19a. The restraints involving the side chains of I98 (β5), V100 (β5), Y122 (β7), L143 (β9) and V145 (β9) delineate the orientation of the two sheets on top of each other in a β-sheet sandwich.

Grey ellipses indicate correlations involving residues from the C-terminal domain and are discussed in paragraph 4.3.3.

4.3.2 The structure of the α-crystallin domain in αB-crystallin oligomers

The analysis of various spectra from αB, and the interactive restraint collection using the programme SolARIA yielded 558 ^{13}C-^{13}C distance restraints and backbone ψ- and φ-angles for 55 residues of the α-crystallin domain in αB. A structure quality report generated with the PSVS protein validation suite for the 10 lowest energy structures is given below (103;104):

Secondary Structure Elements:
beta strands (ordered residues): 75-80, 142-147, 134-136, 89-94, 97-103, 117-123

RMSD	All residues	Ordered residues
All backbone atoms	1.8 Å	1.4 Å
All heavy atoms	2.6 Å	2.2 Å

Ramachandran Plot Summary for ordered residues from Procheck

Most favoured regions	Additionally allowed regions	Generously allowed regions	Disallowed regions
89.0%	10.6%	0.3%	0.1%

Global quality scores

Programme[3]	Procheck (phi-psi)	Procheck (all)	MolProbity Clashscore
Raw score	-0.75	-0.59	31.69
Z-score	-2.64	-3.49	-3.91

Close Contacts and Deviations from Ideal Geometry (from PDB validation software)

Number of close contacts (within 1.6 Å for H atoms, 2.2 Å for heavy atoms): 12

RMS deviation for bond angles: 0.4 °

RMS deviation for bond lengths: 0.005 Å

RMS deviation for Impropers: 0.2 °

Restraints

Interresidue C-C distance restraints 3.5 Å < r > 6.5 Å   554

Interresidue C-C distance restraints 4.0 Å < r > 5.5 Å (from CHHC)   4

Total torsion angle restraints   123

Total hydrogen bond restraints (two restraints per H-bond)   38

The 10 lowest-energy structures calculated using both distance and angular restraints showed no distance violations greater than 0.5 Å and no angular restraints violations larger than 5 °. The rmsd from the ensemble is low, even for side-chain atoms and hence side-chain orientation converged in the ensemble. The Z-scores provide information from van der Waals overlap and are in the standard range for high-quality structures. A Ramachandran-plot analysis of the backbone-torsion angles reveals that 89% of the residues are in the most favoured regions and 10.6% in the additionally allowed regions from the ensemble of the 10 lowest energy structures. The superimposed 10 lowest-energy structures are shown as stereo image in figure 4-20a and as ribbon chart turned by 180° in figure 4-20b.

---

[3] The software suite PSVS performs several tests on a given ensemble of structures. For instance the *MolProbity Clashscore* provides information on close contacts of van-der-Waals radii. The Z-score should be larger than -5. For further details see references (103;104).
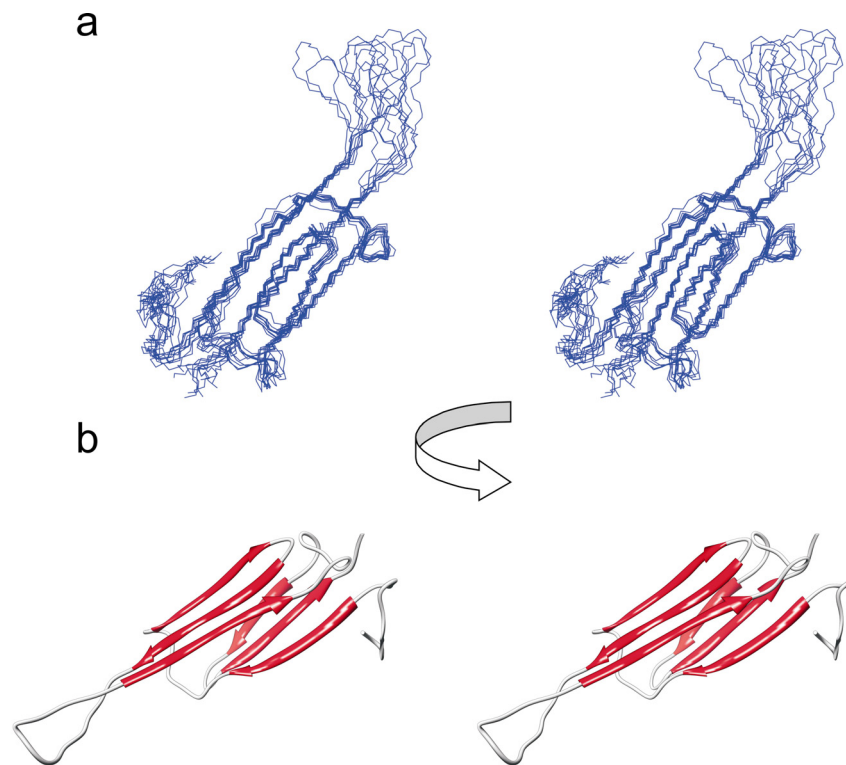
figure 4-20 Solid-state structure of the α-crystallin domain from αB-crystallin. A: Stereoview of ten lowest energy structures are shown superimposed in A. B: Stereoview of the lowest energy structure as ribbon chart.

The topology of the β-strands in the α-crystallin domain (figure 4-21) is well in line with the results from chapter 4.2. β3, β8 and β9 form one sheet on top of the sheet composed of β4, β5 and β7. Subsequent to the last residue in β4 (L89), the residues E88 and E87 form a turn.
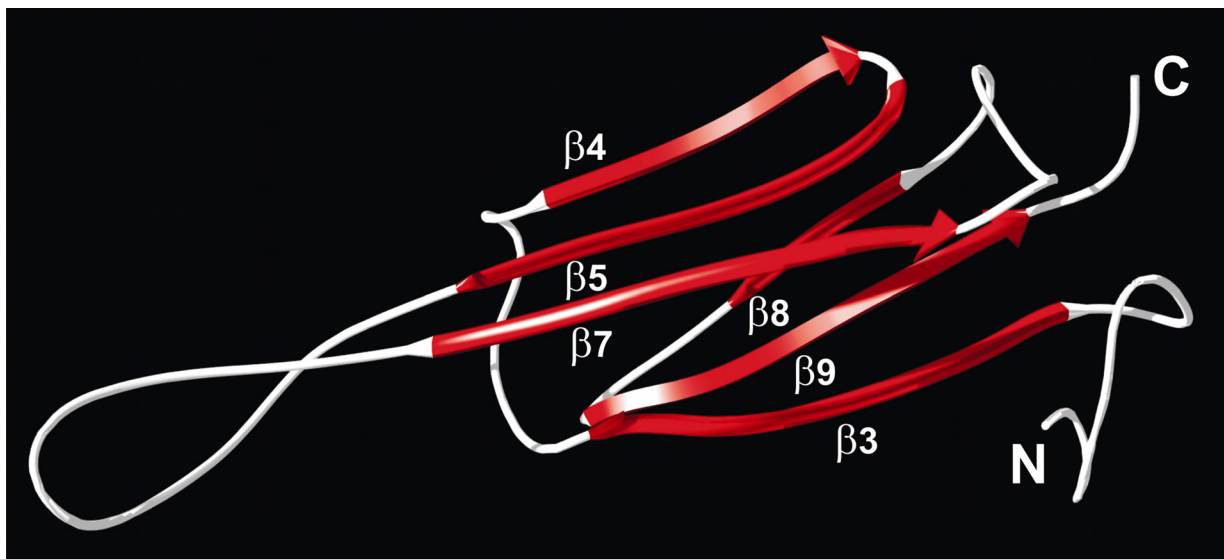
figure 4-21 topology of β-strands in the α-crystallin domain of αB-crystallin.

P86 invokes an ensuing loop connecting β4 and β3. The loop connecting the extended β6/7 strand with the β8 strand was manually refined using the restraints presented in figure 4-22. The Cδ of proline is labelled in 2G-αB, while the Cα is labelled in 1,3G-αB. In the strip extracted at the ^{15}N frequency of P125 from an NCACX spectrum, recorded from 1,3G-αB, a crosspeak at 51 ppm in F_2 and 62.4 ppm in F_3 appears. Since the Cδ of P125 is unlabelled in 1,3G-αB, the correlation must be an interresidual correlation. D129 is the sole residue with a Cα chemical shift of 51 ppm in full-length αB, and is labelled in the 1,3G-αB sample. Hence, the distance restraint was unambiguously assigned and introduced in the structure calculation. Moreover, in 2G-αB the Cδ of proline and the Cα of valine are simultaneously labelled. Interestingly, in an NCACX spectrum recorded from 2G-αB, a crosspeak involving the Cδ of P125 and the Cα of V128 is found (data not shown). P125 invokes a loop involving four residues.
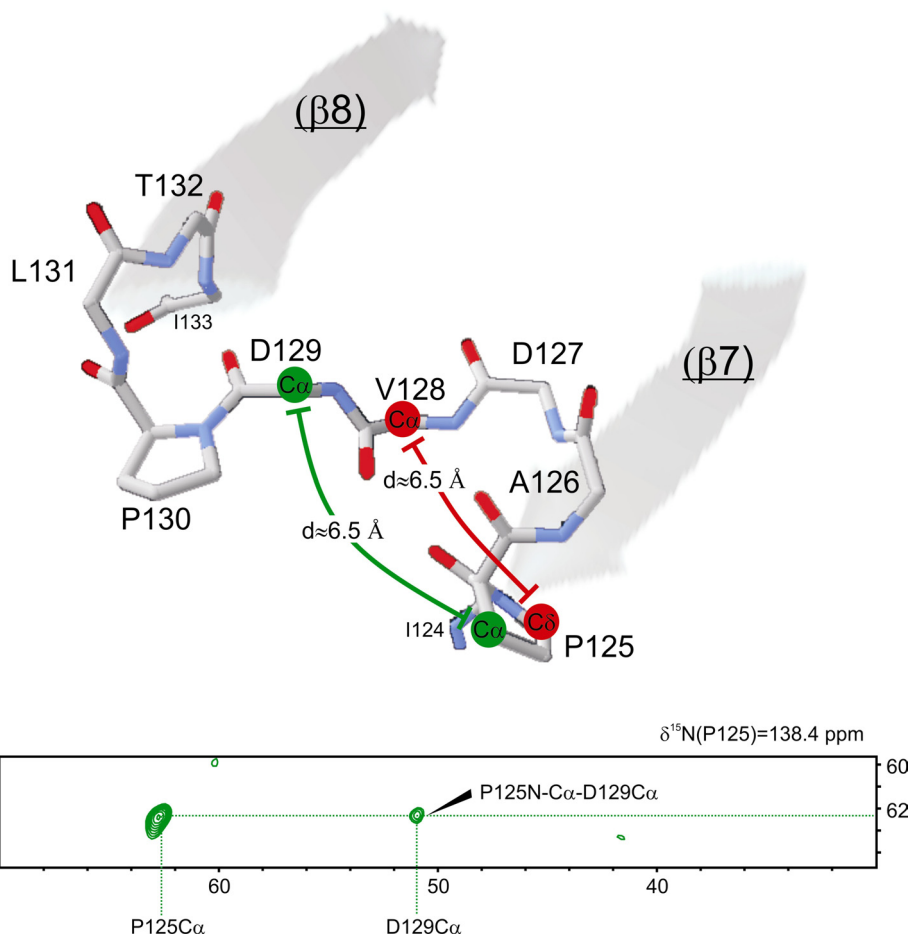
figure 4-22 Structure refinement of the loop connecting β7 with β8. The crosspeak at 51 ppm in

F$_3$ and 62.4 ppm in F$_2$, in the NCACX strip extracted at the ^{15}N frequency of P125, indicates proximity of P125 Cα and D129 Cα.

The subsequent P130 breaks the invoked loop, and forces the loop into the opposite direction forming another loop towards the β8 strand. This sequence motif is not present in sHSPs with known structures.

The hydrophobic core in the plane between the two sheets of the α-crystallin domain, and the chemical properties of the surfaces of the β-sheet-sandwich are summarized in figure 4-23. As expected from interresidual crosspeaks of methyl-bearing amino acids in figure 4-19, the two sheets are hold together by van der Waals interactions between hydrophobic residues. However, F75 (β3), V77 (β3), L79 (β3), V81 (loop between β3 and β4), L94 (β4), V97 (β5), V128 (β9), L137 (β8), V91 (β4) and V93 (β4) have exposed hydrophobic surfaces, if not covered by the N- and C- termini of αB. Isoleucine 124 in β7 and isoleucine 133 right before β8 are in close proximity and, together with V128, build a hydrophobic network at the boundary of the β-sheet-sandwich. I124 is important for the structural integrity of αA, as found by Koteiche *et al.* They prepared cysteine mutants of αA-crystallin, and found a large fraction of unfolded and insoluble protein for the I124C mutant only (95). The surface of the β-sheet composed of β3, β8 and β9 contains polar but uncharged residues. In contrast, the β-sheet composed of β4, β5 and β7 has an excess of charged residues on the surface. One arginine (R123), two histidines (H119 and H101) and four lysines (K121, K103, K92 and K90) are responsible for the positive charge excess. Two glutamates (E117 and E99) are negatively charged residues with exposed side-chains. αB has proven to be involved in the complexation of two-fold positively charged ions like Cu^{2+} and Zn^{2+}. E99, H119 and H101 are thought to be electron donators for the complex, however a fourth ligand has not yet been identified (105-107). The side-chains of R120 and R116 are pointing away from the accessible surface towards the buried sites of the sheet, a finding that will be discussed in the next paragraph.

polar, uncharged residues point
out of the plane of β3, β8, β9
2xT, 2xN, 2xS

β3
β9
F75
β8
V77
V145 L143
V128
I133
I124 V93
I98   V91
V97
L94

V142
L137
V81
L89
L79
V100
β4
β5
β7

\-   +   +   -   +
+

charged residues point out of the plane
from β4, β5, β7. 1xR, 2xH, 4xK, 2xE, 1xD

+   -   +   +
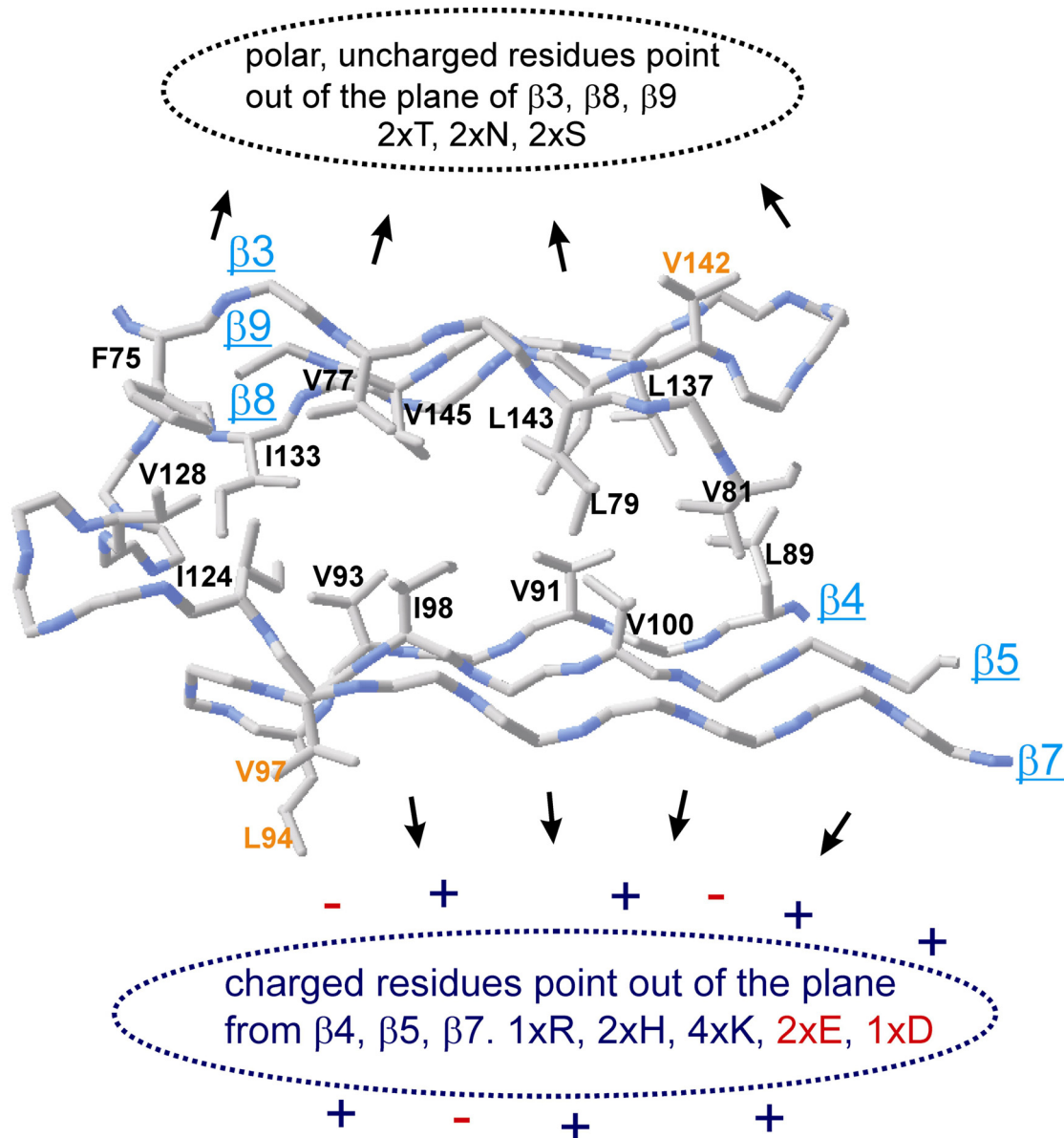
figure 4-23 The hydrophobic core in the plane between the two β-sheets of the α-crystallin domain of αB-crystallin.

The interdomain loop contains three glutamates at positions 110, 105 and 106; the latter two disrupt the β-sheet by forming a β-bulge. The remaining amino acids in the loop comprise one arginine (R107), one glutamine (Q108), one aspartate (D109) and one histidine (H111). The interdomain loop is less defined in our structure due to the lack of intermolecular crosspeaks supporting one particular conformation. In the following paragraph a model for the dimer interface will be presented that is well in line with the observed structural restraints and previous findings from site-directed spin-labelling experiments of αA.

4.3.3 Intermolecular interactions and a dimer model based on solid-state NMR restraints

In this paragraph, a model of the dimeric subunit in αB is presented which is based on experimental distance restraints from solid-state NMR and the prior knowledge of the structure of the monomeric protomer. In chapter 4.2 the disassembly of αB oligomers using 6M chaotropic guanidinium chloride was described. For the collection of intermolecular restraints, again 6M guanidinium chloride was used to disassemble αB oligomers in unfolded protomers and refold a mixture of differently labelled, monomeric protomers, to build mixed labelled oligomers. Therefore a 1,3-^{13}C-glycerol/^{14}N labelled αB preparation was mixed 1:1 with a ^{12}C/^{15}N labelled αB preparation. In ^{15}N-^{13}C correlation experiments, only intermolecular correlations are expected, since none of the protomers is both ^{13}C and ^{15}N labelled. In the following, the mixed [1,3G-^{13}C-^{14}N]/[^{12}C-^{15}N] αB oligomer will be referred to as mixed-αB. ^{15}N-^{13}C correlation experiments were recorded by using the TEDOR pulse sequence shown in figure 3-1b, but without the methyl filter. In figure 4-24, contour plots of 2D ^{15}N-^{13}C TEDOR experiments recorded from mixed-αB with 3 ms and 6 ms REDOR mixing time are shown in a and b, respectively:
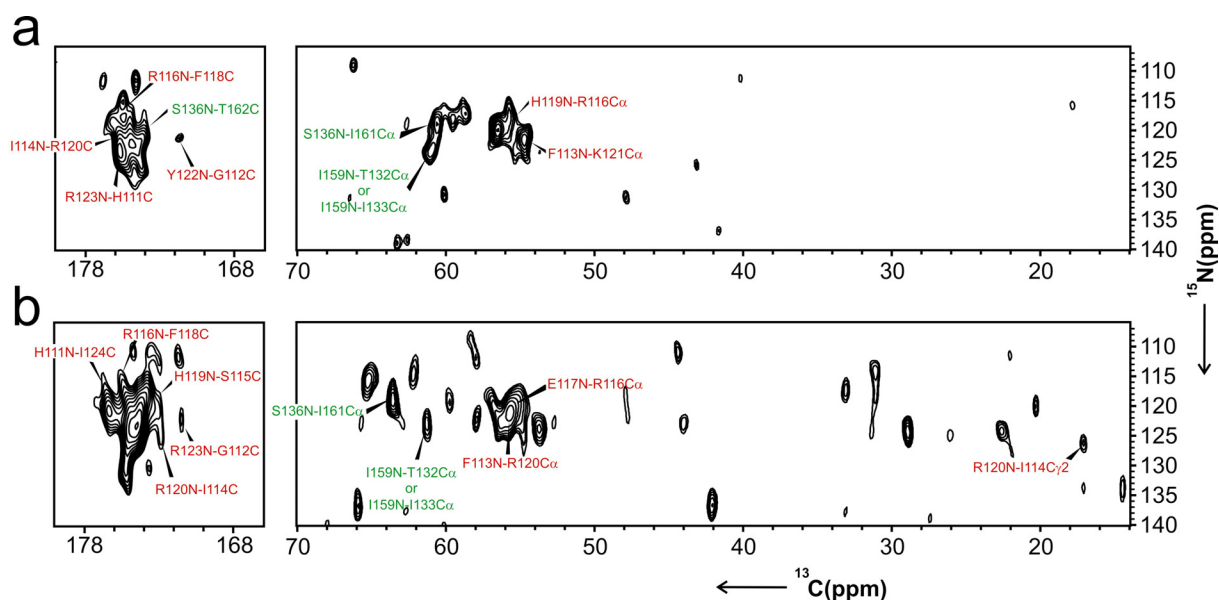
figure 4-24 Contour plots of ^{15}N-^{13}C correlation experiments recorded from αB-crystallin oligomers assembled from ^{12}C-^{15}N labelled- and 1,3-^{13}C-glycerol/^{14}N labelled monomeric protomers. The spectra in a and b were recorded with 3 ms and 6 ms REDOR mixing time, respectively. Green and red assignments indicate correlations involved in C-terminal binding

and dimerisation, respectively.

Green labelled assignments represent intermolecular correlations involving residues from the C-terminal domain.

The putative intermolecular binding site of the C-terminal IXI-motif is the hydrophobic groove on the surface of the subunit between β4 and β8 (108;109). The β4-strand has recently been identified to be involved in the binding of denatured proteins and hence is part of the chaperone binding site (56). Sharma and coworkers found that immobilisation of the C-terminal domain in αA-crystallin by cross-linking, reduces the chaperone activity (109). This suggests that denatured proteins bind in the hydrophobic groove, and displace at least a part of the C-terminal domain. As expected, the C-terminal IXI motif correlates intermolecular with residues of β8. The ^{15}N of I159 (in the C-terminal domain) correlates with T132Cα or I133Cα (both are in the loop before β8). This correlation appears already at 3 ms REDOR mixing time (figure 4-24a). S138N (β8) is in proximity to I161Cα and T162C′ (figure 4-24a). The crosspeaks highlighted with grey ellipses in the ^{13}C-^{13}C correlation experiments in figure 4-19a and b recorded from 1,3G-αB and 2G-αB, respectively, indicate proximity of the side-chains of I133, V93 and V91 in β-strands four and eight to I161 and I159 in the C-terminal IXI-motif. These findings support the assumption, that the side-chains of I159 and I161 in the C-terminal domain cover the hydrophobic groove between β4 and β8.

The crosspeaks indicated with red labels in figure 4-24 derive from residues in the extended β6/7-strand and are involved in the dimerization of αB. The even-numbered residues from β7 are hydrogen bonded to β5 and hence the odd residues may build hydrogen-bonds to a second αB protomer and form an intermolecular, antiparallel β-strand. Koteiche *et al.* found a two-fold symmetry for the dimeric subunit of αA. In SDSL experiments they found decreasing proximity from E113-R112-S111 to I110. Interestingly, the sequence in αB has an exact homology: E117-R116-S115-I114. The odd residues E117 and S115 may form intermolecular hydrogen-bonds. From the spectrum in figure 4-24a, a conspicuous intermolecular crosspeak appears at 119.9 ppm in F_1 and 171.4 ppm in F_3. The C′ frequency can be unambiguously assigned to G112C′, but in the ^{15}N dimension the assignment is less obvious since three residues in the β7 stretch resonate at $\delta^{15}N$ =119.8 ($\pm$ 0.2 ppm). However, at 6

ms REDOR mixing time, the C´ of G112 correlates with a second backbone amide at 122.7 ppm. The combination of the two backbone ^{15}N frequencies is unique in the β7 stretch and hence the two crosspeaks can be assigned to Y122N-G112C´ and the latter to R123N-G112C´. R120 is in proximity to I114, K121 to F113 and F118 to R116. These intermolecular, interstrand correlations define the dimer-interface of αB. In addition, a correlation R120N-I114Cγ2 is observed in figure 4-24b. From these restraints, a model of the αB dimer could be calculated and is shown in figure 4-25a. The extended β6/7-strand from two αB monomers forms a long intermolecular, antiparallel β-sheet. This domain swapping connects the two β-sheets composed of β4, β5 and β7 of two monomeric protomers to a large β-sheet spanning over six strands. Intermolecular cation-π interactions of the side-chains of R116 from one protomer to F118 to a second protomer and backbone hydrogen bonds stabilise the dimerisation (figure 4-25b). Cation-π interactions are frequently found in protein-protein interfaces and are responsible for strong side-chain interactions that stabilize the position of the molecules with respect to each other. However, the twist of the intermolecular β-sheet may slightly vary and hence form another source of heterogeneity in the αB oligomer. Unfortunately, in the presented model the negatively charged E117 of the two protomers face each other in an unfavourable manner (figure 4-25c). However, a salt-bridge between the side-chain carboxyl of E117 and the amine-group of K103 can compensate the charge. Additional interactions to compensate the charges could emerge between E117 and the imidazole-group of H119. Ponni Rajagopal found a dimer to monomer transition of the truncated αB11.4 if the pH is decreased from 7.4 to 6.5 (52). Since the protonation state of the imidazole group of histidines is pH sensitive, the interactions of H119, E117 and K103 may be disrupted at lower pH. This may lead to a repulsion of residues in the dimer interface (i.e. E117) and subsequently to a disruption of the dimer. Interestingly, in the case of cellular stress, i.e. cardiac ischemia, the pH also decreases in cells. The geometrical orientation of H119, E117 and K103 would be ideal for chelating ions like Zn^{2+} or Cu^{2+}. In summary, the dimer model matches the experimental restraints and is consistent with the results from SDSL experiments of the homologue αA. In the following paragraph, first results from the R120G mutant will be shown. Again these results support the proposed model for the dimer interface.
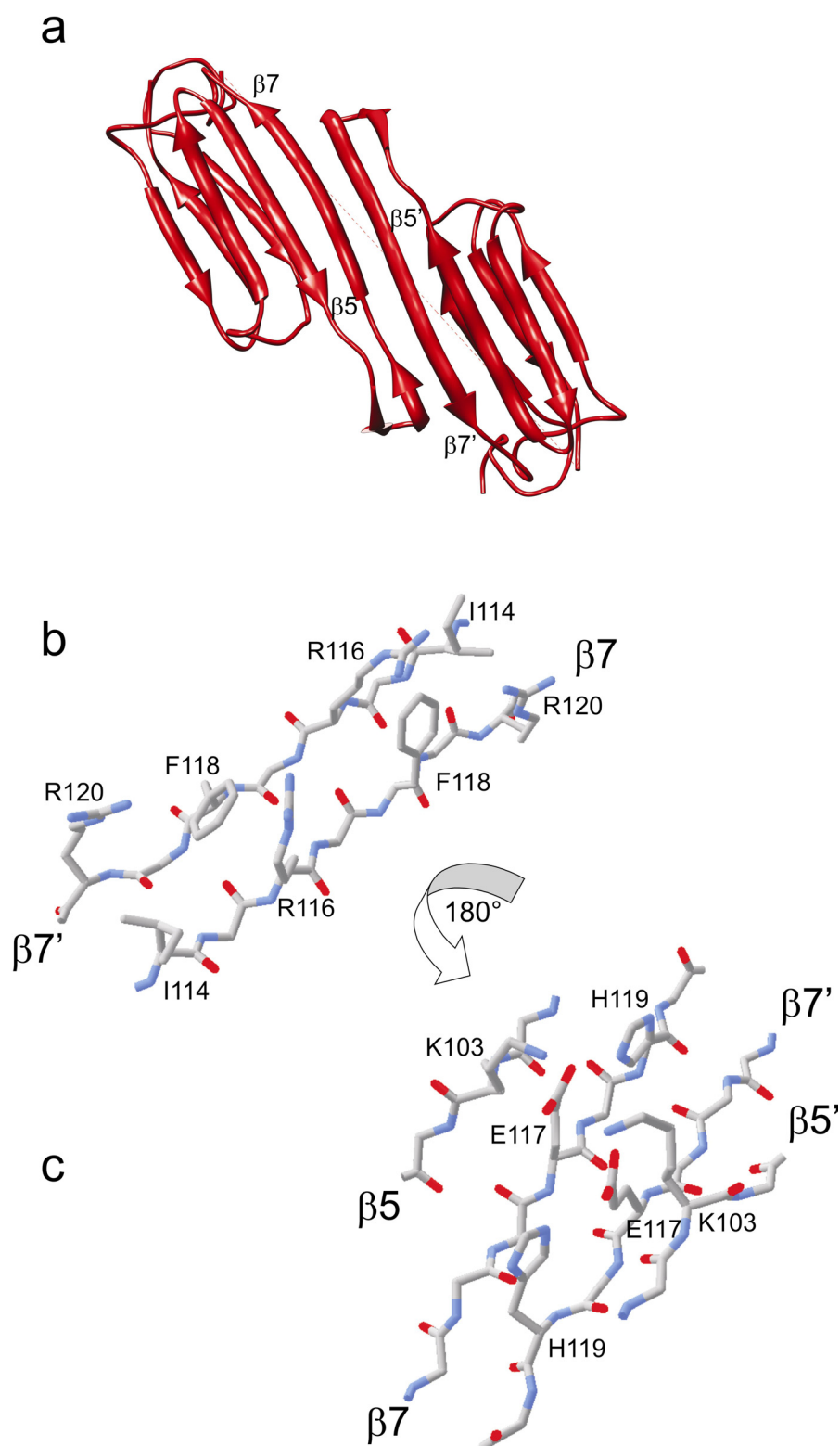
figure 4-25 a: Model of the αB dimeric subunit based on experimental restraints derived from solid-state NMR. In b and c, the side-chains that are involved in intermolecular and intramolecular interactions are shown.

4.3.4 The R120G mutant causes chemical-shift changes in β7 and β5

A well-known αB mutant is R120G, which has the R120 in the β7 strand replaced by a glycine. Although only one residue is replaced, this mutant shows completely different biological behaviour and has different structural characteristics. The oligomer size is increased and the chaperone activity diminished. In the present paragraph, first results from the R120G mutant will be presented and discussed in the light of the results from the previous paragraphs. A 2-^{13}C-glycerol/^{15}N labelled R120G αB mutant was prepared and compared with the wild type in a 2D ^{13}C-^{13}C correlation spectrum. In figure 4-26 a contour plot of a 2D ^{13}C-^{13}C PDSD spectrum recorded from 2-^{13}C-glycerol, ^{15}N labelled R120G αB (blue) is shown together with a spectrum from wild-type αB (red). The Cα and Cβ of Isoleucine is simultaneously labelled in 2-^{13}C-glycerol labelled samples and, hence, Cα-Cβ crosspeaks from isoleucines can be observed and analysed. I124, which is on the same β-strand as R120, shows little chemical shift difference compared to I98, which is on the β5-strand opposed to R120, and shifted by about two residues. I114 shows the largest chemical shift difference compared to the other isoleucines. A cartoon on the bottom of figure 4-26 illustrates the position of the isoleucines in proximity to the R120 mutation site in the dimeric subunit of αB. In the dimer structure, I114 from one αB protomer faces R120 of a second αB protomer in the dimer interface. The chemical shift changes upon mutation are larger for I114, than for I124 which is four residues away compared with I114 which is six residues away from R120. Further work to understand the structural and/or chemical differences between the R120G mutant and the wild-type is in progress.
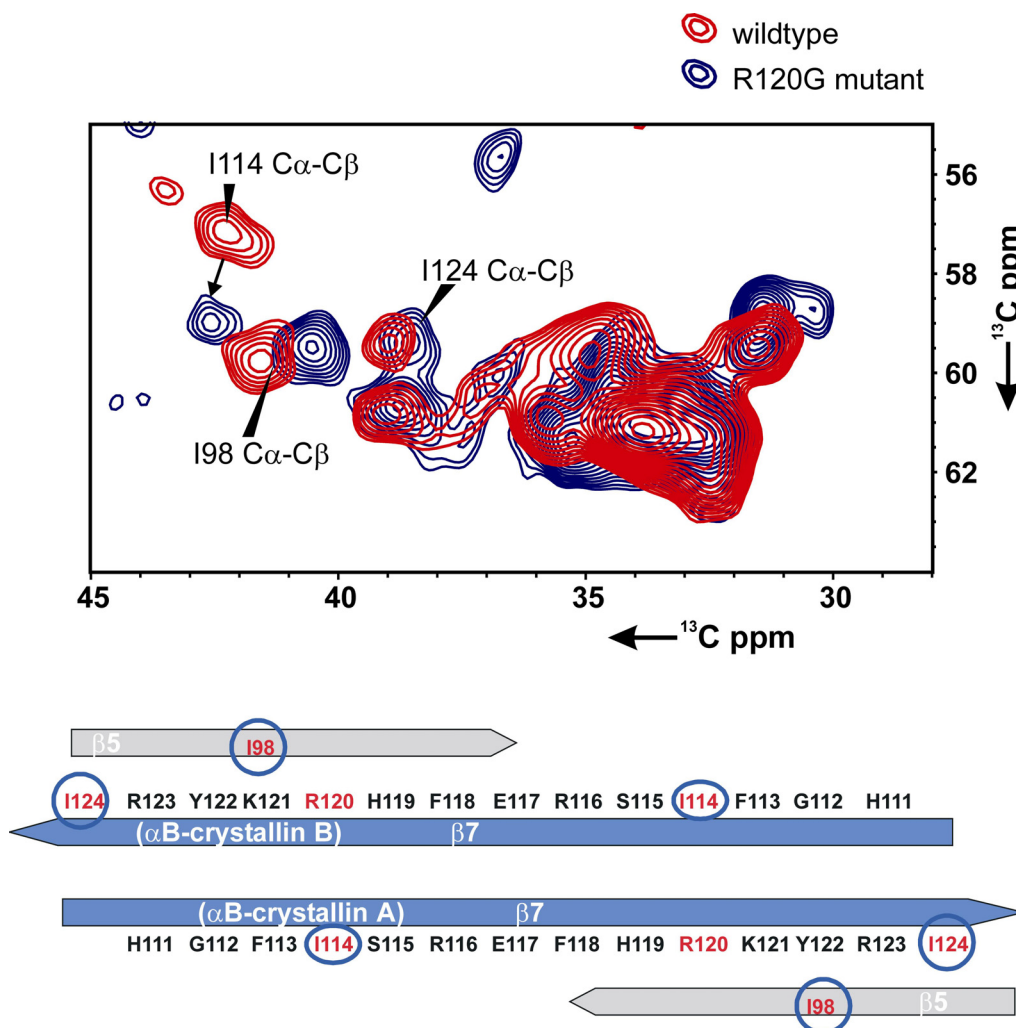
figure 4-26 On top, contour plots of 2D ^{13}C-^{13}C PDSD spectra (75 ms mixing time) from 2-^{13}C-glycerol labelled αB (red) an from 2-^{13}C-glycerol labelled R120G αB-crystallin (blue) are superimposed. The isoleucine Cα-Cβ crosspeak region is shown. On the bottom a cartoon of the dimer interface is given. Isoleucine residues surrounding the R120G mutation site are highlighted.

4.4 Conclusions

The structure of the α-crystallin domain from αB is the first example of a structure with atomic resolution obtained from human oligomeric sHSPs. The structure determination was done by using distance restraints from solid-state NMR data, dihedral-angles, derived from chemical shifts and a selective Cα-Cα correlation spectrum. This shows that solid-state NMR is a powerful tool for researching large heterogeneous systems that cannot be addressed by solution NMR or X-ray

crystallography alone. The presented data give a satisfactory description of the β-sandwich fold of the 89-residue α-crystallin domain from full-length αB-crystallin oligomers. The backbone torsion-angles were extracted from the chemical-shifts and further used as structural restraints for a structure calculation. The analysis of crosspeaks from 3D ^{15}N-^{13}C-^{13}C and 2D ^{13}C-^{13}C correlation experiments yielded sufficient structural restraints to calculate a high-quality solid-state NMR structure of the α-crystallin domain. In addition intermolecular contact areas for the dimerization and the interaction site of the C-terminal IXI-motif were pictured by using mixed-labelled oligomers. These data are a first move towards understanding the structure-function relationship of heterogeneous human sHSPs. To date, there are only structures of distantly related sHSPs in the PDB, viz., MjHSP16.5, wHSP16.9, and Tsp36. One of the differences between the structures of the α-crystallin domains in αB and these sHSPs is the region comprising β6. In full-length αB, β6 is found to be fused with β7, forming one contiguous strand. The NMR data are consistent with a dimer interface comprised mainly of the long β6/7 strand of the α-crystallin domain. Using spin-labelled forms of αA-crystallin, Mchaourab and colleagues proposed that αA residues 109-120 are at the dimer interface (95). The corresponding residues in αB are 113-123, i.e. the residues in the β6/7 strand. In another modelling study using X-ray scattering, a V-shaped dimer structure was proposed for the α-crystallin domain (51). This model, although not consistent with the NMR data in terms of secondary structure elements, predicted that part of β6/7 forms the dimer interface.

EPR studies on spin labelled αA-crystallin revealed a lack of the characteristic periodicity observed for β-strand structure in the region that would correspond to the β2 strand (99). A multitude of signals was observed for M68 in full-length αB indicating structural heterogeneity. This observation was made on a sample containing only PFMW-labelled αB. There are only two Met residues in αB, M1 and M68. M68 is in the middle of the HR1 region (the region that is analogous to β2 strand in other sHSPs), and this heterogeneity could arise from different conformations and/or chemical environments. The sequential assignment connecting M68 to its neighbours proved difficult due to multiple assignment options. The observation of heterogeneity has an important ramification: HR1 is likely to be involved in inter-subunit interactions that give rise to the known polydisperse property of αB.

The NMR data and their structural interpretation provide insight into potential functional aspects of αB. Residues in β3 and β4 in αB have been shown to be important for chaperone activity (56). Comparison of solution and solid-state NMR data provide evidence that β3 and β4 are involved in inter-subunit interactions in the αB oligomer (figure 4-16). sHSPs are thought to dissociate into smaller multimeric units or undergo a conformational change before binding denatured proteins.

The unique architecture of αB (different dimerisation motif) may make it particularly suitable for the diverse functions described in the literature (47). For instance, cardiac ischemia is accompanied by acidosis. Interestingly, αB11.4 dissociates from a dimer into a monomer under acidic conditions (52). sHSPs are thought to dissociate into smaller multimeric units before binding denatured proteins. For example, it has been shown that αB binds the cardiac muscle protein, titin, more tightly at lower pH. During cardiac ischemia, there is a pH drop from 7.5 to 6.9 and αB translocates from the cytoplasm to the myofibrils (110;111). Does αB undergo a conformational change on acidosis that is a prerequisite for more efficient binding to cardiac muscle proteins? How this change is transmitted to the full-length αB remains to be investigated. Finally, the approaches presented here will be applicable to a wide variety of dynamic complexes with similar properties, opening new avenues for structural biology.

5. References

Reference List

1. McDermott, A. and Polenova, T. (2007) *Curr. Opin. Struct. Biol.* **17,** 617-622

2. Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002) *Nature* **420,** 98-102

3. Castellani, F., van Rossum, B. J., Diehl, A., Rehbein, K., and Oschkinat, H. (2003) *Biochemistry* **42,** 11476-83

4. Castellani, F. (2003) Structure Determination of Immobilized Proteins By Solid-State NMR Spectroscopy.

5. Ferguson, N., Becker, J., Tidow, H., Tremmel, S., Sharpe, T. D., Krause, G., Flinders, J., Petrovich, M., Berriman, J., Oschkinat, H., and Fersht, A. R. (2006) *Proc. Natl. Acad. Sci. U. S. A* **103,** 16248-16253

6. Ritter, C., Maddelein, M. L., Siemer, A. B., Luhrs, T., Ernst, M., Meier, B. H., Saupe, S. J., and Riek, R. (2005) *Nature* **435,** 844-848

7. Lange, A., Giller, K., Hornig, S., Martin-Eauclaire, M. F., Pongs, O., Becker, S., and Baldus, M. (2006) *Nature* **440,** 959-962

8. Liu, S., Li, P., Dybkov, O., Nottrott, S., Hartmuth, K., Luhrmann, R., Carlomagno, T., and Wahl, M. C. (2007) *Science* **316,** 115-120

9. Al-Amoudi, A., Diez, D. C., Betts, M. J., and Frangakis, A. S. (2007) *Nature* **450,** 832-837

10. Smith, S. O., Aschheim, K., and Groesbeek, M. (1996) *Q. Rev. Biophys.* **29,** 395-449

11. Liang, B. and Tamm, L. K. (2007) *Proc. Natl. Acad. Sci. U. S. A* **104,** 16140-16145

12. Marassi, F. M. and Opella, S. J. (2003) *Protein Science* **12,** 403-411

13. van Dijk, A. D., Boelens, R., and Bonvin, A. M. (2005) *FEBS J.* **272,** 293-312

14. James, T. L. (1998) *Fundamentals of NMR,*

15. Brown, S. P. and Emsley, L. (2001) Solid-State NMR. In Vo-Dinh and Gauglitz, editors. *NMR Spectroscopy*, Wiley,

16. Brown, S. P. and Spiess, H. W. (2001) *Chem. Rev.* **101,** 4125-56

17. Laws, D. D., Bitter, H. M., and Jerschow, A. (2002) *Angew. Chem. Int. Ed. Engl.* **41,** 3096-129

18. Baldus, M., Petkova, A. T., Herzfeld, J., and Griffin, R. G. (1998) *Mol. Phys.* **95,** 1197-1205

19. Baldus, M., Geurts, D. G., Hediger, S., and Meier, B. H. (2008) *J. Magn. Reson.* **A,** 140-144

20. Pauli, J., Baldus, M., van Rossum, B., de Groot, H., and Oschkinat, H. (2001) *Chembiochem* **2,** 272-81

21. Bennet, A. E., Rienstra, C. M., and Griffiths, J. M. (1998) *J. Chem. Phys.* **108,** 9463-9479

22. Verel, R., Baldus, M., Ernst, M., and Meier, B. H. (1998) *Chem. Phys. Letters* **287,** 421-428

23. Gullion, T. and Schaefer, J. (1989) *J. Magn. Reson.* **81,** 196-200

24. Suter, D. and Ernst, R. R. (1985) *PHYSICAL. REVIEW B. CONDENSED. MATTER* **32,** 5608-5627

25. Ueda, T., Masui, H., and Nakamura, N. (2001) *Solid State Nucl. Magn. Reson.*

20, 145-155

26. Griffiths, J. M., Bennett, A. E., Engelhard, M., Siebert, F., Raap, J., Lugtenburg, J., Herzfeld, J., and Griffin, R. G. (2000) *Biochemistry* **39,** 362-371

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

28. Takegoshi, K., Nakamura, S., and Terao, T. (2001) *Chemical Physics Letters* **344,** 631-637

29. Hajduk, P. J. (2006) *Mol. Interv.* **6,** 266-272

30. Shuker, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) *Science* **274,** 1531-1534

31. Balayssac, S., Bertini, I., Falber, K., Fragai, M., Jehle, S., Lelli, M., Luchinat, C., Oschkinat, H., and Yeo, K. J. (2007) *Chembiochem.* **8,** 486-489

32. van Rossum, B. J., Castellani, F., Rehbein, K., Pauli, J., and Oschkinat, H. (2001) *Chembiochem* **2,** 906-14

33. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289-302

34. Eghbalnia, H. R., Wang, L., Bahrami, A., Assadi, A., and Markley, J. L. (2005) *J. Biomol. NMR* **32,** 71-81

35. Fossi, M., Oschkinat, H., Nilges, M., and Ball, L. J. (2005) *J. Magn Reson.* **175,** 92-102

36. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallographica Section D-Biological Crystallography* **54,** 905-921

37. Fossi, M., Castellani, F., Nilges, M., Oschkinat, H., and van Rossum, B. J. (2005) *Angew. Chem. Int. Ed Engl.* **44,** 6151-6154

38. Nikaido, H. (1998) *Curr. Opin. Microbiol.* **1,** 516-523

39. Nikaido, H. (2003) *Microbiol. Mol. Biol. Rev.* **67,** 593-656

40. Misra, R. and Benson, S. A. (1989) *J. Bacteriol.* **171,** 4105-4111

41. Fajardo, D. A., Cheung, J., Ito, C., Sugawara, E., Nikaido, H., and Misra, R. (1998) *J. Bacteriol.* **180,** 4452-4459

42. Conlan, S., Zhang, Y., Cheley, S., and Bayley, H. (2000) *Biochemistry* **39,** 11845-11854

43. Behlau, M., Mills, D. J., Quader, H., Kuhlbrandt, W., and Vonck, J. (2001) *J. Mol. Biol.* **305,** 71-77

44. Yildiz, O., Vinothkumar, K. R., Goswami, P., and Kuhlbrandt, W. (2006) *EMBO J.* **25,** 3702-3713

45. Hiller, M., Krabben, L., Vinothkumar, K. R., Castellani, F., van Rossum, B. J., Kuhlbrandt, W., and Oschkinat, H. (2005) *Chembiochem.* **6,** 1679-1684

46. Haslbeck, M., Franzmann, T., Weinfurtner, D., and Buchner, J. (2005) *Nat. Struct. Mol. Biol.* **12,** 842-846

47. Augusteyn, R. C. (2004) *Clin. Exp. Optom.* **87,** 356-366

48. Augusteyn, R. C. (1998) *Int. J. Biol. Macromol.* **22,** 253-262

49. Barbash, O., Lin, D. I., and Diehl, J. A. (2007) *Cell Div.* **2,** 2

50. Lin, D. I., Barbash, O., Kumar, K. G., Weber, J. D., Harper, J. W., Klein-Szanto, A. J., Rustgi, A., Fuchs, S. Y., and Diehl, J. A. (2006) *Mol. Cell* **24,** 355-366

51. Feil, I. K., Malfois, M., Hendle, J., van Der, Z. H., and Svergun, D. I. (2001) *J. Biol. Chem.* **276,** 12024-12029

52. Rajagopal, P. and Klevit, R. (2007) aB10.1 forms dimers in solution.

53. Ghosh, J. G. and Clark, J. I. (2005) *Protein Sci.* **14,** 684-695

54. Carver, J. A., Aquilina, J. A., Truscott, R. J., and Ralston, G. B. (1992) *FEBS Lett.* **311,** 143-149

55. Hendrick, J. P. and Hartl, F. U. (1993) *Annu. Rev. Biochem.* **62,** 349-384

56. Bhattacharyya, J., Padmanabha Udupa, E. G., Wang, J., and Sharma, K. K. (2006) *Biochemistry* **45,** 3069-3076

57. Siezen, R. J., Bindels, J. G., and Hoenders, H. J. (1980) *Eur. J. Biochem.* **111,** 435-444

58. van Montfort, R. L., Basha, E., Friedrich, K. L., Slingsby, C., and Vierling, E. (2001) *Nat. Struct. Biol.* **8,** 1025-1030

59. Kim, K. K., Kim, R., and Kim, S. H. (1998) *Nature* **394,** 595-599

60. Haley, D. A., Horwitz, J., and Stewart, P. L. (1998) *J. Mol. Biol.* **277,** 27-35

61. Muchowski, P. J., Wu, G. J., Liang, J. J., Adman, E. T., and Clark, J. I. (1999) *J. Mol. Biol.* **289,** 397-411

62. Hiller, M., Higman, V. A., Jehle, S., van Rossum, B. J., Kuhlbrandt, W., and Oschkinat, H. (2008) *J. Am. Chem. Soc.* **130,** 408-409

63. Pauli, J., van, R. B., Forster, H., de Groot, H. J., and Oschkinat, H. (2000) *J. Magn Reson.* **143,** 411-416

64. Hong, M. and Griffin, R. G. (1998) *JACS* **120,** 7113-7114

65. Sinha, N., Grant, C. V., Wu, C. H., De Angelis, A. A., Howell, S. C., and Opella, S. J. (2005) *J. Magn Reson.* **177,** 197-202

66. Jaroniec, C. P., Tounge, B. A., Herzfeld, J., and Griffin, R. G. (2001) *JACS* **123,** 3507-3519

67. Jaroniec, C. P., Filip, C., and Griffin, R. G. (2002) *JACS* **124,** 10728-10742

68. Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V., and Griffin, R. G. (1995) *J. Chem. Phys.* **103,** 6951-6958

69. Burns, S. T., Wu, X., and Zilm, K. W. (2000) *J. Magn. Reson.* **143,** 352-359

70. Wu, X. L. and Zilm, K. W. (1993) *J. Magn. Reson.* **104,** 119-122

71. Wu, X. L., Burns, S. T., and Zilm, K. W. (1994) *J. Magn. Reson.* **111,** 29-36

72. Wu, X. L. and Zilm, K. W. (1993) *J. Magn. Reson. A* **102,** 205-213

73. Sangill, R., Rastrupandersen, N., Bildsoe, H., Jakobsen, H. J., and Nielsen, N. C. (1994) *J. Magn. Reson.* **107,** 67-78

74. Rossi, P., Subramanian, R., and Harbison, G. S. (1999) *J. Magn. Reson.* **141,** 159-63

75. Peng, J. and Frydman, L. (1995) *J. Magn. Reson.* **113,** 247-250

76. Lesage, A., Steuernagel, S., and Emsley, L. (1998) *J. Am. Chem. Soc.* **120,** 7095-7100

77. Sakellariou, D., Lesage, A., and Emsley, L. (2001) *J. Magn. Reson.* **151,** 40-7

78. Opella, S. J., Frey, M. H., and Cross, T. A. (1979) *J. Am. Chem. Soc.* **101,** 5856-5857

79. Opella, S. J. and Frey, M. H. (1979) *J. Am. Chem. Soc.* **101,** 5854-5856

80. Gullion, T., Baker, D. B., and Conradi, M. S. (1990) *J. Magn. Reson.* **89,** 479-484

81. Chevelkov, V., Faelber, K., Diehl, A., Heinemann, U., Oschkinat, H., and Reif, B. (2005) *J. Biomol. NMR* **31,** 295-310

82. Ernst R.R., Bodenhausen G., and Wokaun A. (2002) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press: Oxford.,

83. Bruschweiler, R., Griesinger, C., Sorensen O.W., and Ernst R.R. (1988) *J. Magn. Reson.* **A,** 178-185

84. Straus, S. K., Bremi, T., and Ernst, R. R. (1996) *Chemical Physics Letters* **262,** 709-715

85. Chevelkov, V., Chen, Z., Bermel, W., and Reif, B. (2005) *J. Magn Reson.* **172,** 56-62

86. Duma, L., Hediger, S., Brutscher, B., Bockmann, A., and Emsley, L. (2003) *JACS* **125,** 11816-7

87. Duma, L., Hediger, S., Lesage, A., and Emsley, L. (2003) *J. Magn. Reson.* **164,** 187-95

88. Delsuc, M. A. and Levy, G. C. (1988) *Journal of Magnetic Resonance (1969)* **76,** 306-315

89. Hoch, J. C. and Stern, A. S. (1996) *Maximum entropy reconstruction in NMR: an alternative to DFT in NMR Data Processing*, Wiley-Liss: New York,

90. Jones, J. A. and Hore, P. J. (1991) *Journal of Magnetic Resonance (1969)* **92,** 276-292

91. Shimba, N., Stern, A. S., Craik, C. S., Hoch, J. C., and Dotsch, V. (2003) *J. Am. Chem. Soc.* **125,** 2382-2383

92. LeMaster, D. M. and Kushlan, D. M. (1996) *J. Am. Chem. Soc.* **118,** 9255-9264

93. Aerts, T., Clauwaert, J., Haezebrouck, P., Peeters, E., and Van, D. H. (1997) *Eur. Biophys. J.* **25,** 445-454

94. Liu, L., Ghosh, J. G., Clark, J. I., and Jiang, S. (2006) *Anal. Biochem.* **350,** 186-195

95. Berengian, A. R., Parfenova, M., and Mchaourab, H. S. (1999) *J. Biol. Chem.* **274,** 6305-6314

96. Gardiennet, C., Loquet, A., Etzkorn, M., Heise, H., Baldus, M., and Bockmann, A. (2008) *J. Biomol. NMR* **40,** 239-250

97. Berengian, A. R., Bova, M. P., and Mchaourab, H. S. (1997) *Biochemistry* **36,** 9951-9957

98. Koteiche, H. A., Berengian, A. R., and Mchaourab, H. S. (1998) *Biochemistry* **37,** 12681-12688

99. Koteiche, H. A. and Mchaourab, H. S. (1999) *J. Mol. Biol.* **294,** 561-577

100. Mchaourab, H. S., Berengian, A. R., and Koteiche, H. A. (1997) *Biochemistry* **36,** 14627-14634

101. Franks, W. T., Zhou, D. H., Wylie, B. J., Money, B. G., Graesser, D. T., Frericks, H. L., Sahota, G., and Rienstra, C. M. (2005) *J. Am. Chem. Soc.* **127,** 12291-12305

102. Igumenova, T. I., Wand, A. J., and McDermott, A. E. (2004) *J. Am. Chem. Soc.* **126,** 5323-5331

103. Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) *Proteins* **66,** 778-795

104. Bhattacharya, A., Wunderlich, Z., Monleon, D., Tejero, R., and Montelione, G. T. (2008) *Proteins* **70,** 105-118

105. Ganadu, M. L., Aru, M., Mura, G. M., Coi, A., Mlynarz, P., and Kozlowski, H. (2004) *J. Inorg. Biochem.* **98,** 1103-1109

106. Hawse, J. R., Cumming, J. R., Oppermann, B., Sheets, N. L., Reddy, V. N., and Kantorow, M. (2003) *Invest Ophthalmol. Vis. Sci.* **44,** 672-679

107. Narayanan, S., Kamps, B., Boelens, W. C., and Reif, B. (2006) *FEBS Lett.* **580,** 5941-5946

108. Augusteyn, R. C., Hum, T. P., Putilin, T. P., and Thomson, J. A. (1987) *Biochim. Biophys. Acta* **915,** 132-139

109. Sharma, K. K., Kumar, G. S., Murphy, A. S., and Kester, K. (1998) *J. Biol. Chem.* **273,** 15474-15478

110. Golenhofen, N., Htun, P., Ness, W., Koob, R., Schaper, W., and Drenckhahn, D. (1999) *J. Mol. Cell Cardiol.* **31,** 569-580

111. Golenhofen, N., Arbeiter, A., Koob, R., and Drenckhahn, D. (2002) *J. Mol. Cell Cardiol.* **34,** 309-319

Summary

Advances in structural biology have been strongly supported by the success of three leading techniques: X-ray crystallography, solution NMR and Electron microscopy. These techniques are widely used for structure determination of proteins. Today, structure determination of non-crystalline solid protein preparations has been made possible through rapid progress of solid-state NMR methodology for biological systems. Castellani *et. al.* have solved and refined a first structure of a microcrystalline protein by using only solid-state MAS NMR spectroscopy. Its success relies on the perspective to access systems that are difficult to crystallise or that form large heterogeneous complexes and insoluble aggregates. Examples are ligands bound to membrane integrated receptors. Protein fibrils and heterogeneous proteins form another target for solid-state NMR.

In this thesis, the structure of the α-crystallin domain of human αB crystallin (αB) is solved using solid-state NMR methodology. Atomic level structural information on αB, a prominent member of the small Heat Shock Protein (sHSP) family has been a challenge to obtain due its polydisperse, oligomeric nature. It is shown that magic-angle spinning solid-state NMR can be used to obtain high-resolution information on ~ 580 kDa human αB assembled from 175-residue, 20 kDa subunits. An ~90-residue α-crystallin domain is common to all sHSPs. Chemical shifts of the backbone and the Cβ resonances have been obtained for residues 64-163 (α-crystallin domain plus part of the C-terminus) in αB and the isolated α-crystallin (αB10.1) domain by solid- and solution-state NMR, respectively. From the N-terminal domain 32 residues could be assigned by solid-state NMR. The presented data give a satisfactory description of the β-sandwich fold of the α-crystallin domain from full-length αB-crystallin oligomers. In the presented structure, the β-sandwich consists of six strands, three in each sheet. A majority of residues in the α-crystallin domain have similar chemical shifts in both solid- and solution-state indicating a similar structure for the domain in its isolated and oligomeric forms. Sites of inter-subunit interaction are identified from chemical shift differences that cluster to specific regions of the α-crystallin domain. Multiple signals are observed for the resonances of M68 in the oligomer, identifying the region containing this residue as existing in heterogeneous environments within

αB.

In addition intermolecular contact areas for the dimerisation and the interaction site of the C-terminal IXI-motif were pictured by using mixed-labelled oligomers. These data are a first move towards understanding the structure-function relationship of heterogeneous human sHSPs.

Since solid-state NMR does not require rapid tumbling of proteins, there is no inherent limitation on the particle's size. However, for large systems, the number of signals forms the bottleneck due to overlapping signals. To overcome this hurdle, methods for resolution enhancement are presented in this thesis.

A simple spectroscopic filtering technique is presented that may aid the assignment of ^{13}C and ^{15}N resonances of methyl-containing amino acids in solid-state magic-angle spinning (MAS) NMR. A filtering block that selects methyl resonances is introduced in two-dimensional (2D) ^{13}C-homonuclear and ^{15}N-^{13}C heteronuclear correlation experiments. As model systems, the α-spectrin SH3 domain and the outer membrane protein G of E. *coli* are used. In addition, these methods were shown to be valuable for the assignment of αB.

As another resolution enhancement method, *J*-decoupling using refocusing pulses in the indirect dimension and MaxEnt reconstruction in the direct dimension is shown in this thesis. These techniques yield well-resolved spectra without sacrificing the signal to noise to an intolerable amount. Since MaxEnt is a post-acquisition processing method, existing data can be reprocessed and analysed in greater detail. MaxEnt reconstruction can be seen as a powerful and viable alternative to experimental approaches for *J*-decoupling.

Finally, the approaches presented here will be applicable to a wide variety of dynamic complexes, opening new avenues for structural biology.

Zusammenfassung

Die Fortschritte in der Strukturbiologie sind im Wesentlichen auf die Anwendung dreier Techniken, der Röntgenkristallographie, der Lösungs-NMR und der Elektronenmikroskopie zurückzuführen. Diese Methoden sind weit verbreitet für die Strukturaufklärung von biologischen Systemen. Castellani *et. al.* haben als erste die

Struktur eines mikrokristallinen Proteins mit Hilfe der Festkörper NMR Spektroskopie gelöst. Diese Technik ist vor allem deshalb so erfolgreich, weil sie Zugang zu Systemen ermöglicht, die nicht kristallisieren, große heterogene Komplexe oder unlösliche Aggregate bilden. Geeignete Systeme für die Festkörper NMR Spektroskopie sind beispielsweise Binder an Rezeptoren die in Membranen eingebettet sind, Protein-Fibrillen oder heterogene Biomoleküle.

In dieser Doktorarbeit wird die Struktur der α-Kristallin Domäne von menschlichem αB-Kristallin (αB) mit Hilfe der Festkörper NMR Spektroskopie aufgeklärt. Wegen der polydispersen und oligomeren Eigenschaften von αB war es bis jetzt eine Herausforderung, Strukturinformationen auf atomarer Ebene von αB, einem bekannten Mitglied der kleinen Hitzeschock Protein-Familie (sHSP), zu erlangen. Es wird gezeigt, daß mittels Festkörper NMR Spektroskopie hoch aufgelöste Daten von αB, welches ~580 kDa große Oligomere aus 20 kDa Untereinheiten bildet, erhalten werden können. Eine ca. ~90 Reste umfassende α-Kristallin Domäne ist eine charakteristische Eigenschaft aller sHSPs. Mit Hilfe der Festkörper NMR Spektroskopie wurden die Signale der Reste 64-162 zugeordnet und von 32 Aminosäuren aus dem N-Terminus. Die chemischen Verschiebungen wurden mit denen der isolierten Domäne (Reste 64-152), welche mit Hilfe der Lösungs-NMR Spektroskopie erhalten wurden, verglichen. Die vorliegenden Daten beschreiben die β-Sandwich Faltung der α-Kristallin Domäne des vollständigen αB-Kristallins sehr gut. In der präsentierten Struktur, besteht das β-Sandwich aus zwei Faltblättern, die je aus drei Strängen zusammengesetzt sind. Die Mehrheit der Reste in der α-Kristallin Domäne zeigen eine ähnliche chemische Verschiebung bei der Messung mittels Festkörper NMR und Lösungs- NMR, was auf eine ähnlich Struktur der isolierten Domäne und der Domäne im Oligomer schließen läßt. Aus den systematischen Unterschieden der chemischen Verschiebungen in der α-Kristallin Domäne lassen sich Rückschlüsse auf die Interaktionsstellen der Untereinheiten ziehen. Für M68 werden mehrere Signalsätze beobachtet, was darauf schließen läßt, daß sich dieser Rest in einer heterogenen Umgebung im αB Oligomer befindet.

Zusätzlich wurde ein Dimerisierungs-Modell und die intermolekulare Bindungsstelle des C-terminalen IXI-Motivs beschrieben. Hierzu wurden Oligomere hergestellt die aus unterschiedlich markierten Monomeren bestehen. Diese Daten sind der erste Schritt hin zum Verständnis des Struktur-Funktion Zusammenhanges von

menschlichen heterogenen sHSPs.

Molekularbewegung ist keine Bedingung für die Festkörper NMR Spektroskopie, deshalb gibt es keine Größenbegrenzung für die zu untersuchende Probe. Große Systeme führen jedoch zu einer hohen Anzahl an Signalen, und folglich zu nicht interpretierbaren Spektren. In dieser Doktorarbeit werden deshalb Methoden zur Verbesserung der Auflösung präsentiert, um diese Hürde zu überwinden.

Es wird eine einfache Filter-Methode zur selektiven Anregung von Aminosäuren, die Methyl-Gruppen enthalten, vorgestellt. Diese Methode vereinfacht die Zuordnung von ^{13}C- und ^{15}N-Resonanzen dieser Gruppe von Aminosäuren in der Festkörper magic-angle spinning (MAS) NMR Spektroskopie. Ein für Methylgruppen selektiver Filterblock wird hierfür in 2D ^{13}C homonukleare und ^{13}C-^{15}N heteronukleare Experimente eingebaut. Zur Entwicklung der Methode wurden die SH3 Domäne des α-Spektrins und das Außenmembran-Protein G von E. *coli* als Modelsysteme verwendet. Darüber hinaus wird gezeigt, daß die Filter-Methode einen wertvollen Beitrag bei der Resonanz-Zuordnung von αB leistet.

Als weitere Methode zur Verbesserung der Auflösung wurde *J*-Entkopplung eingesetzt. Um die Auflösung in der indirekten Dimension zu verbessern wird ein Teil der Spins während der indirekten Evolutionszeit refokussiert. Um *J*-Kopplungen aus der direkten Dimension zu entfernen, werden die Daten mit der Methode der Maximalen Entropie (MaxEnt) prozessiert. Diese Methode liefert gut aufgelöste Spektren mit guten Signal zu Rausch Ausbeuten. Da die Methode der maximalen Entropie zum Prozessieren der Daten nach der Datenaufnahme eingesetzt wird, können bereits aufgenommene Daten erneut prozessiert werden und daraufhin detaillierter ausgewertet werden. Maximum Entropie Rekonstruktion kann als eine leistungsstarke und brauchbare Alternative zu experimenteller *J*-Entkopplung betrachtet werden.

Die in dieser Doktorarbeit vorgestellten Methoden können auf eine Vielzahl von dynamischen Komplexen angewendet werden und öffnen neue Wege für die Strukturbiologie.

Curriculum vitae

Aufgrund von Datenschutz wird der Lebenslauf nicht veröffentlicht.

Danksagung

Zunächst möchte ich meinen Eltern dafür danken, daß Sie an meine Zukunft geglaubt haben und Zeit ihres Lebens für mich da waren. Sie waren ganz besondere Menschen, die aufopferungsvoll für Ihre Familie gekämpft und gearbeitet haben. Ich würde diesen Moment sehr gerne mit Ihnen teilen und möchte an dieser Stelle ganz besonders Ihrer gedenken.

Prof. Hartmut Oschkinat möchte ich dafür danken, daß er mir die Möglichkeit gab an einem herausragenden Forschungsprojekt teilzuhaben. Von Anfang an hatte ich die Möglichkeit mich wissenschaftlich frei zu entfalten. Ich habe gelernt wissenschaftliche Ergebnisse kritisch zu betrachten und wasserdicht auszuarbeiten. Die Erfahrung mit internationalen Forschergruppen zusammenzuarbeiten hat mich persönlich und wissenschaftlich weiter gebracht.

Dr. Barth-Jan van Rossum möchte ich meinen besonderen Dank aussprechen. Du hast mich in den vergangenen Jahren stets motiviert und hast mit viel Begeisterung und Kreativität mit mir kooperiert. Dein Empfehlungsbrief für den Vortrag in Taiwan hat es mir ermöglicht einen sehr bedeutenden Moment zu erleben, den ich nie vergessen werde. Vielen Dank, daß Du mir mit viel Geduld wissenschaftliches Englisch und die Macht der Bilder beigebracht hast. Ich habe viel von Dir gelernt.

Meinen Kooperationspartnern in Seattle, Ponni Rajagopal und Rachel E. Klevit, möchte ich für die offene und herzliche Arbeitsatmosphäre danken. Ich bin froh und dankbar, daß ich in einem internationalen Forscherteam mit erfahrenen Wissenschaftlern arbeiten durfte. Ich hatte dabei immer das Gefühl von Akzeptanz, was mir sehr wichtig ist.

Für die Zusammenarbeit im Labor, möchte ich Kristina Rehbein, Anne Diehl und Katja Fälber danken. Auch wenn die Diskussionen manchmal hitzig waren, so war das Ergebnis meistens positiv. Vielen Dank.

Dr. Peter Schmieder möchte ich an dieser Stelle gesondert erwähnen, weil er mit

Appendix A

Table of the ^{13}C and ^{15}N chemical shifts derived from solid-state NMR of precipitated αB-crystallin at 270 K and at pH=7.6.

continued

Residue	N	C´	Cα	Cβ	Cγ	Cγ1	Cγ2	Cδ	Cδ1	Cδ2	Cε	Cε1	Cε2	Cζ	Nε2
W9	-	-	58.8	-	116.5	-	-	-	-	-	-	-	-	-	-
F14	121.7	178.2	-	-	-	-	-	-	-	-	-	-	-	-	-
F15	110.8	173.2	56.03	40.62	-	-	-	-	-	-	-	-	-	-	-
P16	135.3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
S19	-	174.6	-	64.68	-	-	-	-	-	-	-	-	-	-	-
P20	141.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Q26	117.1	-	59.61	31.75	-	-	-	179.6	-	-	-	-	-	-	118.1
F27	-	173.7	56.78	41.79	-	-	-	-	-	-	-	-	-	-	-
F28	123	174.4	56.59	42.27	-	-	-	-	-	-	-	-	-	-	-
G29	112.3	171	43.71	-	-	-	-	-	-	-	-	-	-	-	-
E30	118.5	175.5	56.97	-	34.68	-	-	184.3	-	-	-	-	-	-	-
L32	-	175.8	-	-	26.43	-	-	-	-	-	-	-	-	-	-
L33	119.6	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L37	-	-	54.32	-	-	-	-	-	-	-	-	-	-	-	-
F38	-	174.7	55.62	-	-	-	-	-	-	-	-	-	-	-	-
P39	132.1	176.6	62.34	32.8	27.27	-	-	50.2	-	-	-	-	-	-	-
T40	124	-	62.38	69.7	-	-	-	-	-	-	-	-	-	-	-
S41	111.9	171.8	57.63	66.03	-	-	-	-	-	-	-	-	-	-	-
T42	112.6	175.3	60.63	61.81	22.21	-	-	-	-	-	-	-	-	-	-
L44	-	177.4	-	44.18	27.06	-	-	-	-	-	-	-	-	-	-
S45	119.2	173.7	57.48	62.91	-	-	-	-	-	-	-	-	-	-	-
P46	141.5	-	-	-	-	-	-	50.52	-	-	-	-	-	-	-
P51	-	173.9	-	31.5	-	-	-	-	-	-	-	-	-	-	-
P52	132	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Residue	N	C'	Cα	Cβ	Cγ	Cγ1	Cγ2	Cδ	Cδ1	Cδ2	Cε	Cε1	Cε2	Cζ	Nε2
A57	-	177.8	52.31	19	-	-	-	-	-	-	-	-	-	-	-
P58	136.7	-	-	-	-	-	-	-	-	-	-	-	-	-	-
S59	120.4	-	58	64.28	-	-	-	-	-	-	-	-	-	-	-
W60	125.2	173.3	55.62	29.59	116.6	-	-	-	-	-	-	-	139.5	-	-
F61	117.5	174.1	56.88	41.07	141.2	-	-	-	-	-	-	-	-	-	-
D62	120.9	175	52.46	42.79	-	-	-	-	-	-	-	-	-	-	-
T63	116.2	-	62.5	70.73	-	-	-	-	-	-	-	-	-	-	-
G64	109.6	172	43.96	-	-	-	-	-	-	-	-	-	-	-	-
L65	129.4	176.7	54.33	44.61	26.75	-	-	-	-	-	-	-	-	-	-
S66	119.4	175.8	56.82	-	-	-	-	-	-	-	-	-	-	-	-
E67	122.4	175.7	56.27	29.95	35.92	-	-	-	-	-	-	-	-	-	-
M68	122.8	175.2	55.23	33.68	30.1	-	-	-	-	-	14.77	-	-	-	-
R69	121.9	175.9	54.08	32.78	27.33	-	-	-	-	-	-	-	-	-	-
L70	124	176.5	55.18	43.49	27.42	-	-	27.56	-	-	-	-	-	-	-
E71	121.6	175.2	55.32	29.25	36.1	-	-	-	-	-	-	-	-	-	-
K72	124.7	175.2	52.9	33.44	26.48	-	-	-	-	-	-	-	-	-	-
D73	123	175.7	54.5	42.23	180	-	-	-	-	-	-	-	-	-	-
R74	121.8	175.1	54.96	31.99	27.38	-	-	-	-	-	-	-	-	-	-
F75	121.1	172.9	56.54	42.71	139.5	-	-	131	-	-	128.7	-	-	-	-
S76	118	173.6	56.89	66.22	-	-	-	-	-	-	-	-	-	-	-
V77	117.8	175.5	59.74	34.58	-	21.26	-	-	-	-	-	-	-	-	-
N78	119.9	173.7	51.66	42.26	177.3	-	-	-	-	-	-	-	-	-	-
L79	123.5	175.2	53.25	46.78	27.33	-	-	-	26.05	23.35	-	-	-	-	-
D80	125.4	174.8	55.2	41.49	179.9	-	-	-	-	-	-	-	-	-	-
V81	120.3	177	59.43	31.32	-	21.79	-	-	-	-	-	-	-	-	-
K82	117.5	175.3	55.13	32.66	26.85	-	-	-	-	-	-	-	-	-	-
H83	123.3	-	56.38	32.35	132.7	-	-	-	-	-	-	-	-	-	-
F84	-	175.4	56.97	41.94	-	-	-	132.9	-	-	-	-	-	-	-
S85	119.6	174.2	56.57	62.78	-	-	-	-	-	-	-	-	-	-	-
P86	134.3	178.4	66.23	31.43	27.79	-	-	49.82	-	-	-	-	-	-	-
E87	112	177.7	58.04	28.92	37.61	-	-	-	-	-	-	-	-	-	-

Residue	N	C'	Cα	Cβ	Cγ	Cγ1	Cγ2	Cδ	Cδ1	Cδ2	Cε	Cε1	Cε2	Cζ	Nε2
E88	120.6	175.1	55.92	30.18	37.66	-	-	184.2	-	-	-	-	-	-	-
L89	119.8	175.4	53.16	47.34	26.37	-	-	-	27.37	23.38	-	-	-	-	-
K90	118.6	175.1	55.12	36.84	26.4	-	-	29.29	-	-	-	-	-	-	-
V91	125.7	174.7	61	33.61	-	22.03	21.44	-	-	-	-	-	-	-	-
K92	124.9	175.2	54.27	37.42	24.81	-	-	29.85	-	-	42.02	-	-	-	-
V93	124.7	175.2	61.07	33.85	-	21.75	21.58	-	-	-	-	-	-	-	-
L94	129.4	176.6	53.49	43.26	27.08	-	-	-	-	-	-	-	-	-	-
G95	115.9	173.3	47.69	-	-	-	-	-	-	-	-	-	-	-	-
D96	124.7	175	52.82	40.25	181.6	-	-	-	-	-	-	-	-	-	-
V97	118.9	175.9	60.91	35.71	-	21.68	-	-	-	-	-	-	-	-	-
I98	123.5	175.2	59.6	41.55	-	27.68	18.06	15.36	-	-	-	-	-	-	-
E99	125.9	174.9	54.18	32.77	37.24	-	-	182.1	-	-	-	-	-	-	-
V100	124	174.2	60.99	34.09	-	21.15	21.39	-	-	-	-	-	-	-	-
H101	127.9	173.7	52.45	34.5	135.6	-	-	-	-	-	-	-	-	-	-
G102	110.9	171.3	43.48	-	-	-	-	-	-	-	-	-	-	-	-
K103	123.1	175.1	56.21	36.52	25.5	-	-	29.37	-	-	-	-	-	-	-
H104	123.3	175.7	53.72	32.59	127.9	-	-	-	-	-	-	130.7	-	-	-
E105	123.1	175.8	54.85	30.1	36.04	-	-	183.3	-	-	-	-	-	-	-
E106	120.4	176.4	58.44	31.11	36.9	-	-	183.1	-	-	-	-	-	-	-
R107	124.3	174.8	53.35	33.18	27.48	-	-	-	-	-	-	-	-	-	-
Q108	124.6	175	54.97	29.74	32.81	-	-	177.5	-	-	-	-	-	-	-
D109	127.5	176.5	51.92	41.8	182.6	-	-	-	-	-	-	-	-	-	-
E110	115	176.2	58.17	28.41	37.13	-	-	-	-	-	-	-	-	-	-
H111	115.5	175.6	55.4	32.27	131.5	-	-	-	-	-	-	-	-	-	-
G112	109.2	171.3	44.92	-	-	-	-	-	-	-	-	-	-	-	-
F113	121	175.3	56.45	42.75	139.8	-	-	-	-	-	-	-	-	-	-
I114	119.8	173.1	57.02	42.14	-	28.05	17.31	14.96	-	-	-	-	-	-	-
S115	121.9	173.1	57.06	65.63	-	-	-	-	-	-	-	-	-	-	-
R116	115	175.2	55.18	37.4	30.11	-	-	43.35	-	-	-	-	-	159.9	-
E117	119.7	174.6	55.53	33.82	36.05	-	-	183.7	-	-	-	-	-	-	-
F118	121.7	174.9	56.66	42.51	139.4	-	-	-	-	-	130.5	-	-	-	-

Residue	N	C´	Cα	Cβ	Cγ	Cγ1	Cγ2	Cδ	Cδ1	Cδ2	Cε	Cε1	Cε2	Cζ	Nε2
H119	118.3	175.1	55.63	33.16	133.2	-	-	-	-	-	-	-	-	-	-
R120	125	175.9	55	36.01	27.42	-	-	-	-	-	-	-	-	-	-
K121	123.4	175.3	54.5	37.33	24.3	-	-	-	-	-	-	-	-	-	-
Y122	119.6	175.5	56.46	42.16	128.7	-	-	134.3	-	-	117.7	-	-	159.7	-
R123	122.7	175.5	55.28	32.52	27.32	-	-	43.12	-	-	-	-	-	-	-
I124	127	176.6	59.41	38.99	-	28.74	17.8	14.66	-	-	-	-	-	-	-
P125	138.4	176.7	62.39	32.25	27.7	-	-	50.1	-	-	-	-	-	-	-
A126	120.6	177.9	54.05	18.81	-	-	-	-	-	-	-	-	-	-	-
D127	112.3	174.2	52.54	39.26	-	-	-	-	-	-	-	-	-	-	-
V128	119.4	174.7	60.95	33.39	-	21.83	21.78	-	-	-	-	-	-	-	-
D129	126.8	177	51.05	41.88	179.3	-	-	-	-	-	-	-	-	-	-
P130	143	178.2	64.77	33.02	27.85	-	-	51.58	-	-	-	-	-	-	-
L131	115.5	178.6	56.66	41.1	27.3	-	-	-	27.6	22.7	-	-	-	-	-
T132	106.5	174.1	61	69.38	-	-	22.23	-	-	-	-	-	-	-	-
I133	121.9	176.2	60.55	37.74	-	27.44	18.47	13.53	-	-	-	-	-	-	-
T134	117.3	172.9	60.15	72.5	-	-	21.87	-	-	-	-	-	-	-	-
S135	112.1	173.3	56.93	66.88	-	-	-	-	-	-	-	-	-	-	-
S136	118.3	171.6	56.68	66.46	-	-	-	-	-	-	-	-	-	-	-
L137	122.9	176.2	52.92	43.57	26.98	-	-	-	23.47	27.26	-	-	-	-	-
S138	123.2	177.3	57.36	65.55	-	-	-	-	-	-	-	-	-	-	-
S139	117.4	175.1	61.02	62.69	-	-	-	-	-	-	-	-	-	-	-
D140	116.7	176.6	52.8	39.93	180	-	-	-	-	-	-	-	-	-	-
G141	107.2	172.2	47.37	-	-	-	-	-	-	-	-	-	-	-	-
V142	117.1	176.6	61.29	32.36	-	21.05	21	-	-	-	-	-	-	-	-
L143	134.8	174.9	53.86	45.43	27.51	-	-	-	27.76	24.66	-	-	-	-	-
T144	124.2	174.9	61.23	71.01	-	-	21.95	-	-	-	-	-	-	-	-
V145	127.6	174.5	61.31	32.95	-	22.09	20.13	-	-	-	-	-	-	-	-
N146	122.1	173.9	52.18	43.44	177.4	-	-	-	-	-	-	-	-	-	-
G147	105.8	169.6	45.4	-	-	-	-	-	-	-	-	-	-	-	-
P148	134.7	176.3	62.27	32.63	27.98	-	-	50.09	-	-	-	-	-	-	-
R149	123.3	177.2	54.88	-	27.51	-	-	-	-	-	-	-	-	-	-

132

Residue	N	C′	Cα	Cβ	Cγ	Cγ1	Cγ2	Cδ	Cδ1	Cδ2	Cε	Cε1	Cε2	Cζ	Nε2
K150	-	176.5	54.57	-	25.7	-	-	-	-	-	-	-	-	-	-
Q151	120.5	-	54.53	-	-	-	-	-	-	-	-	-	-	-	-
V152	118.5	-	59.74	31.99	-	-	-	-	-	-	-	-	-	-	-
S153	-	172.9	57.19	65.98	-	-	-	-	-	-	-	-	-	-	-
G154	111.7	-	43.3	-	-	-	-	-	-	-	-	-	-	-	-
P155	131.6	-	61.46	32.34	-	-	-	49.67	-	-	-	-	-	-	-
E156	117.4	-	56.79	-	35.54	-	-	-	-	-	-	-	-	-	-
T158	-	175.8	62.78	68.8	-	-	-	-	-	-	-	-	-	-	-
I159	123	176.9	57.67	37.41	-	27.16	17.32	10.84	-	-	-	-	-	-	-
P160	135.9	-	62.77	31.89	28.07	-	-	-	-	-	-	-	-	-	-
I161	123.4	177.9	60.84	38.94	-	26.09	18.99	14.68	-	-	-	-	-	-	-
T162	122.1	174.2	63.04	69.13	-	-	-	-	-	-	-	-	-	-	-
R163	129.4	-	54.38	34.23	-	-	-	-	-	-	-	-	-	-	-