Chapter 5

Structure of the α-spectrin SH3 domain determined by broad-band recoupling method and spin-dilution

5.1 Introduction

Following an approach that bears many parallels to solution NMR, a solid-state MAS NMR structure of the 62-residue α-spectrin SH3 domain is achieved. The structure is calculated from a large number of inter-residues ^{13}C-^{13}C and ^{15}N-^{15}N restraints, all self-consistently obtained by solid-state MAS NMR. The main idea of our methodology for structure determination of proteins by solid-state MAS is based on the collection of a large amount of approximate distance restraints from spin-diffusion experiments, in combination with extensive but reduced ^{13}C labelling. The dilute labelling is used to suppress dipolar truncation effects. In this Chapter it is described how the long-range correlations were assigned and converted into distance restraints. The discrimination between inter- and intramolecular contacts on a experimental basis forms a part of our concept, as explained in Section 5.4.

5.2 ^{13}C-^{13}C long-range correlations from 2D ^{13}C-^{13}C PDSD spectra

The recoupling method chosen to detect long-range correlation is the proton-driven spin diffusion (PDSD) mixing scheme[1]. It was possible to use a broad-band recoupling method like PDSD since the mostly alternating spin labelling in 2- and 1,3-SH3 (see Chapter 4) leads to a reduction of dipolar truncation effects[2,3] and allows the observation of long-range interactions, while relayed polarization transfer is blocked.

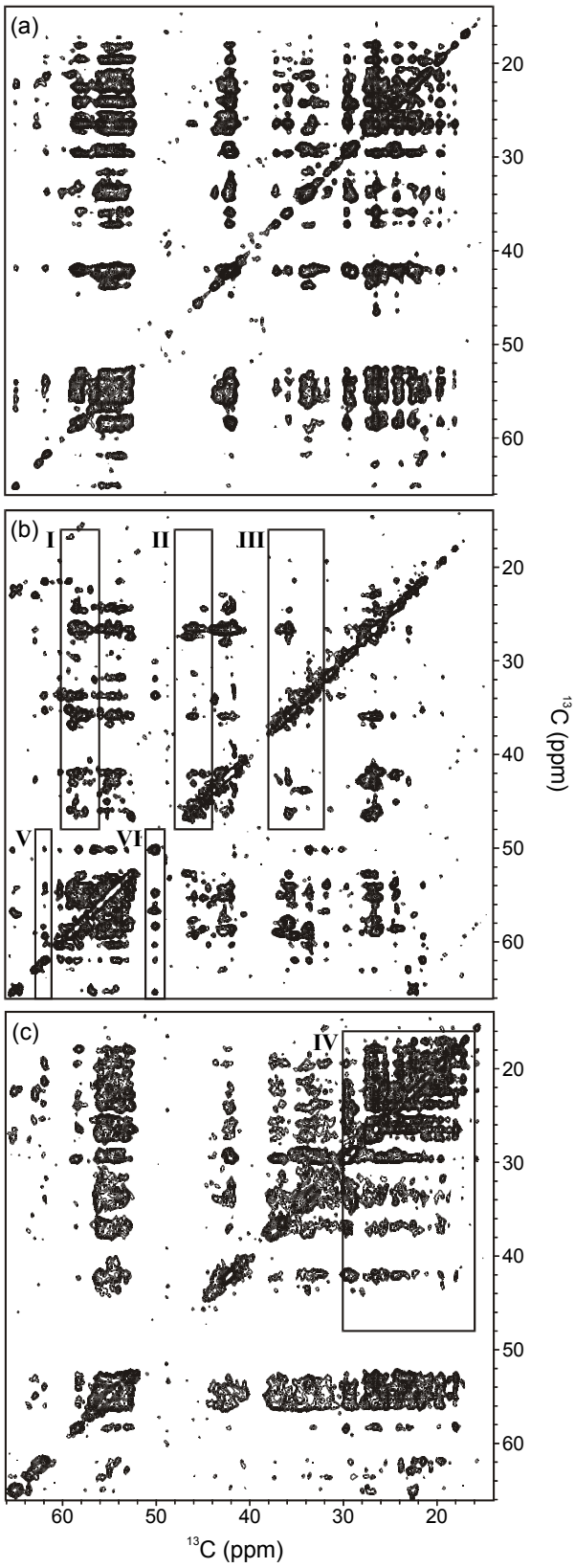

Fig. 5.1 2D ^{13}C-^{13}C PDSD spectra recorded on (U-^{13}C)-SH3 (a), 2-SH3 (b) and 1,3-SH3 (c). The spectra are acquired using a long mixing time of 500 ms. The boxed regions refer to Fig. 5.2. All the spectra are recorded at 17.6 T and using a spinning frequency of 8.0 kHz.

Furthermore, positive side-effects of the reduced labelling are a simplification of the assignment of the long-range correlations and an overall improved resolution as a result of suppressing homonuclear one-bond *J*-couplings, as discussed in Chapter 4. The pulse-program of a standard PDSD experiment is shown in Fig. 3.3. An advantage of the PDSD technique is that longer mixing times can be applied without limitations on balancing *rf* power input and sample heating. In Fig. 5.1, contour plots of 2D ^{13}C-^{13}C PDSD spectra of (U-^{13}C)-SH3 (a), 2-SH3 (b) and of 1,3-SH3 (c) are depicted. All spectra are recorded with long mixing times of 500 ms, to establish long-range correlations. In the PDSD spectrum of (U-^{13}C)-SH3 (a), mainly intraresidue cross-peaks are observed with some additional correlations that are attributed to relayed polarization transfers within an amino-acid or to transfers between sequentially-connected amino-acids. The assignment of the correlations in (a) is difficult due to overlap and ambiguity. In contrast, the 2-SH3 and 1,3-SH3 PDSD spectra (Fig. 5.1b and c) are simplified due to the less extensive labelling and the higher resolution. The spectrum of the 2-SH3 preparation (Fig. 5.1b) is virtually empty in the methyl region between 16 and 28 ppm, because of absence of labelling of side-chain methyl groups. On the other hand, in the region from 48–66 ppm, correlations are observed that can be assigned to long-range polarization transfer events, mainly between C^{α}. The spectrum of 1,3-SH3 in Fig. 5.1c shows a large number of new peaks in the methyl region, which are attributed to long-range interactions.

We assigned the cross-peaks by applying the criterion that each peak must be part of a correlation pattern involving several nuclei of two interacting amino-acids. As an example, the assignment of long-range cross-peaks due to transfer between L33 and V44 is shown in Fig. 5.2 (panels I-IV), which contains the regions indicated in Fig. 5.1b and c. In the panels I, II and III correlations between C^{α} and C^{β} of V44 and C^{β} and C^{γ} of L33 are detected for 2-SH3, whereas correlations between the methyl groups appear in the spectrum of 1,3-SH3 (panel IV).

Of particular interest is the region around 50 ppm (panel VI), where for 2-SH3 a large number of cross-peaks due to the proline-δ signals (P20 and P54) are observed, whereas in the corresponding area of (U-^{13}C)-SH3 (see Fig. 5.1a), no correlations are detected. Additionally, around 62 ppm (Fig. 5.2, panel V), where the C^{α} of T24 and P54 resonate, all cross-peaks involving the two sequential α-carbons are observed, as well as a correlation between the α-carbon signals of P54 and W41. The inter-strand C^{α}–C^{α} correlation between

T24 and Q16 is also observed in the same region, but only in the 1,3-SH3 spectrum due to a low percentage of Q16 C^{α} labelling in the 2-SH3 preparation. These observations reflect systematically appearing correlations between α-carbons in adjacent strands, over distances which range from 4.6–5.4 Å (see Fig. 5.2, top left, and Tab. 5.1), and define the β-sheet topology.

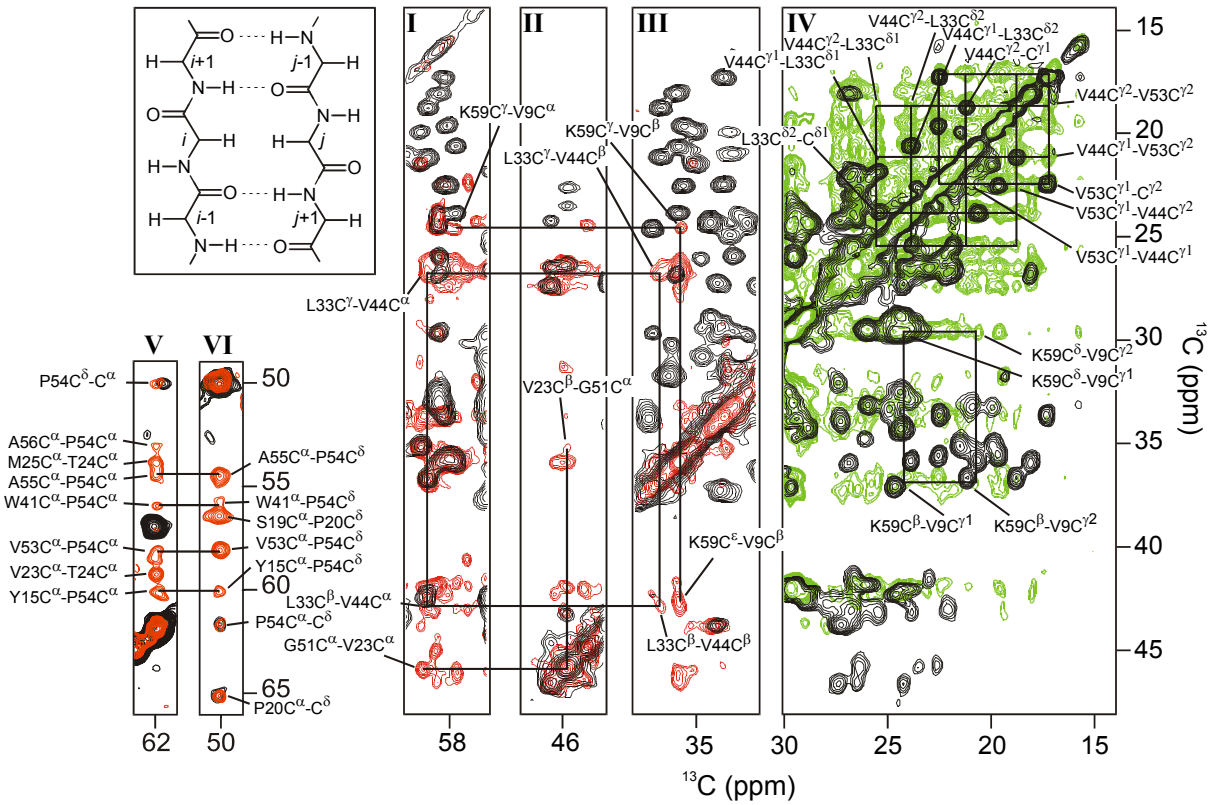

Fig. 5.2 Regions extracted from the spectra of Fig. 5.1, superimposed on a PDSD spectrum of uniformly labelled SH3 domain (black), the latter recorded with a short mixing time of 15 ms (Fig. 4.14a). Part of the assignment of the long-range correlations is reported in the figure, and the lines help as guide for the different correlation patterns. As an example, correlations between residues L33 and V44 are shown. In the upper left corner, a schematic representation of an antiparallel β-sheet is shown.

Furthermore, the secondary β-sheet structure and topology of the SH3 domain could be identified from correlation patterns involving residues 14/26, 15/25 and 16/24, and between 34/43, 33/44 and 31/46. These correlation patterns are characteristic for an antiparallel β-sheet, which in turn proves the reliability of the assignment.

5.3 ^{15}N-^{15}N long-range correlations from 2D ^{15}N-^{15}N PDSD spectra

Similarly as done for the ^{13}C-^{13}C restraints, ^{15}N-^{15}N distances can be extracted from ^{15}N-^{15}N PDSD spectra recorded on ^{15}N-labelled samples. In Fig. 5.3, a 2D ^{15}N correlation spectrum is shown, that was recorded at 17.6 T on a ^{15}N-labelled α-spectrin SH3 domain. A long PDSD mixing time of 4.0 s was applied to exchange magnetization between the weakly coupled ^{15}N spins. In an anti-parallel β-sheet topology, distances between sequential backbone nitrogens are about 3.5 ± 0.2 Å, while the closest distance between nitrogens in different strands of the β-sheet is about 4.5 ± 0.2 Å.

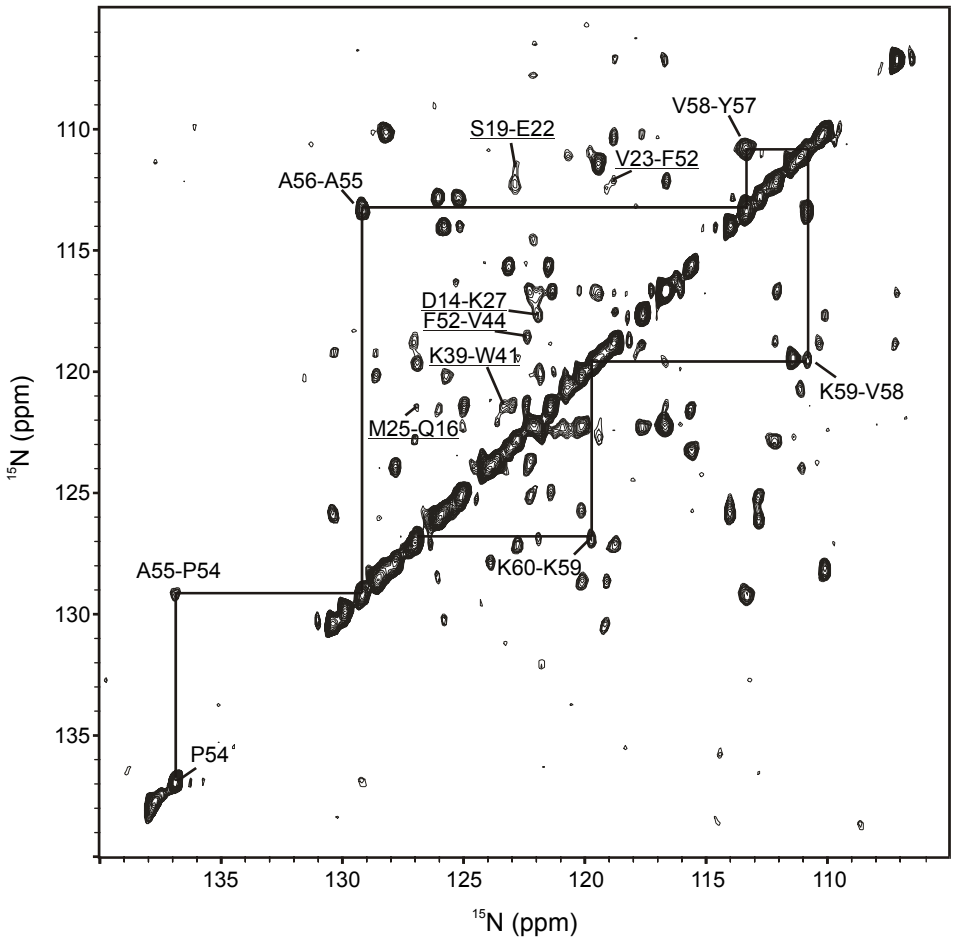

Fig. 5.3 Contour plot of a 2D homonuclear dipolar correlation spectrum of precipitated (U-^{15}N) α-spectrin SH3 domain, recorded at a field of 17.6 T, with a spinning frequency $\omega_R/2\pi$ = 8.0 kHz and at a temperature of 298 K. The data were obtained using a PDSD mixing time of 4.0 s. The dashed line indicates the correlation walk from P54 to K60. The assignment of six ^{15}N-^{15}N long-range correlations is also reported in the figure.

Analysis of the ^{15}N-^{15}N PDSD experiment revealed that most of the observed cross-peaks are related to transfers between the amide ^{15}N spins of sequential residues. As an example, the correlations in the subsequence P54 to K60 are depicted in Fig. 5.3. In a similar way, other

sequential correlations can be found and assigned, accounting for most of the cross-peaks in Fig. 5.3.

Six correlations are detected in the 2D ^{15}N-^{15}N spectrum that can not be assigned to transfers between amides of sequential residues. These correlations likely involve long-range transfers and provide restraints for the calculation of the fold of the SH3 domain from the solid-state NMR data. For instance, a correlation is observed that can be identified as V23-F52 and/or V23-Y15. According to the x-ray structure of the α-spectrin SH3 domain[4], the distance between the amides is 4.0 Å for V23 and F52, and 9.0 Å for V23 and Y15. The observed correlation therefore most likely involves transfer over the shortest distance, between V23 and F52. Likewise, it was found that S19 correlates with E17 (5.8 Å) and/or E22 (4.9 Å). Since these long-range correlations could not be assigned unambiguously from the current solid-state data, they are included as 'ambiguous' restraints in the structure calculations. In the figure, the six inter-residue correlations are indicated, together with the most likely assignment due to distances in the range 3.8–5.0 Å, measured in the x-ray structure[4]. Few long-range ^{15}N-^{15}N correlations were detected because the couplings between ^{15}N spins are approximately six fold weaker than couplings between carbon spins at comparable distances, and nitrogen atoms form a less dense network relative to carbon atoms.

5.4 Intermolecular correlations

In solid samples of proteins, intermolecular interactions are expected to cause additional cross-peaks in PDSD spectra. Initial structure calculations on our SH3 data, for example, showed the necessity of discriminating between inter- and intramolecular contacts on an experimental basis. Since the SH3 domain is a small protein, the surface-to-core ratio is too unfavourable to achieve this by analysing the restraint violations. For this reason and as a part of the general concept, we prepared samples containing 80% unlabelled material and 20% 1,3-SH3 or 2-SH3. A set of ^{13}C-^{13}C PDSD spectra with 500 ms mixing time was recorded on these diluted samples. Analysis showed that several correlations were reduced approximately 5-fold in intensity and are most likely due to intermolecular contacts. In Fig. 5.4 aromatic regions of the spectrum of 1,3-SH3 (in green), overlaid with spectrum of the diluted 1,3-SH3 (in red) sample, are shown. On the right side of the figure, the corresponding slices of the 2D are depicted, which include also the carbonyl regions. The spectra are adjusted in such a way

that the carbonyl signals have the same intensities. Signals assigned to L8C$^{\delta2}$-R21C$^{\zeta}$ and S19C$^{\beta}$-Y57C$^{\zeta}$ correlations are almost not detectable in the diluted sample and we attributed them to intermolecular correlations. In this way, contacts between residues Y57-S19, Y57-P20, R21-L8, and P20-Y13 were identified as "intermolecular", and were removed from the restraint list. The proximity of the aromatic residue Y13 of one molecule to P20 of a neighbouring molecule was also deduced from the change in chemical shift of P20C$^{\delta}$, due to ring-current effect, as discussed in more details in Chapter 7.

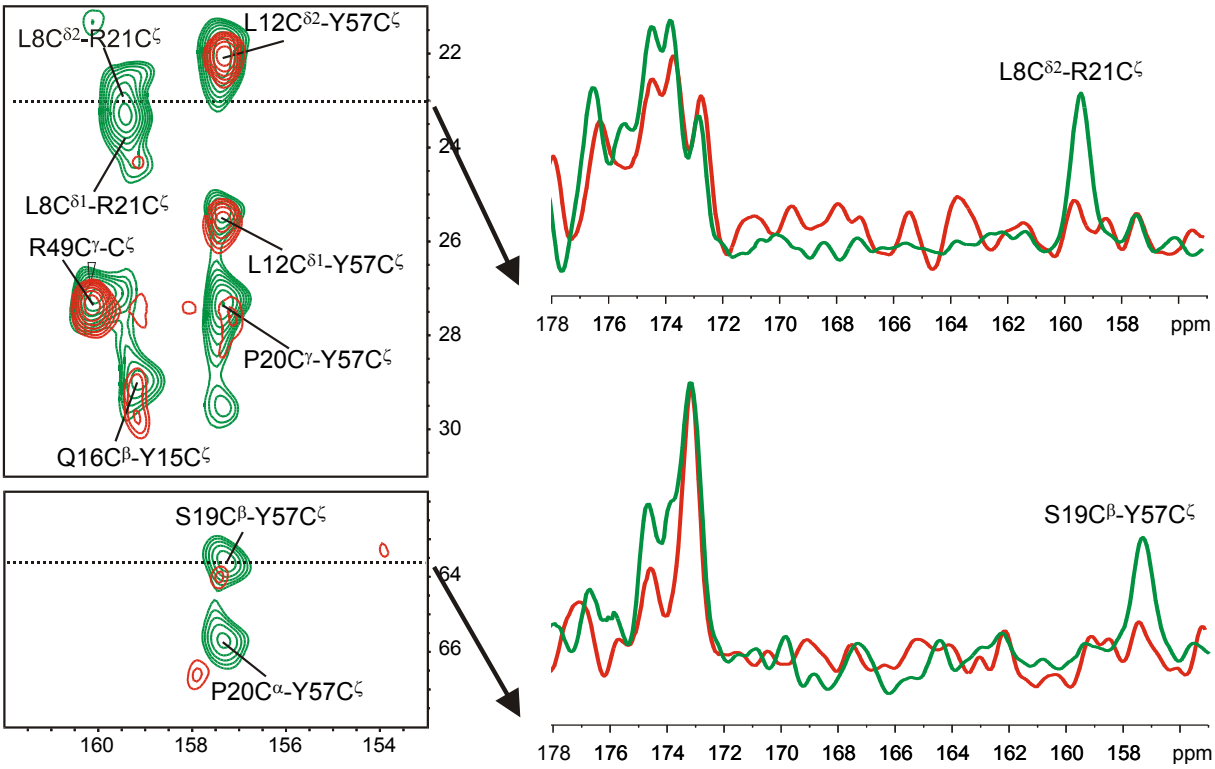

Fig. 5.4 Regions of the 2D ^{13}C-^{13}C PDSD spectrum of 1,3-SH3 (in green), overlaid with the spectrum of a sample containing 80% unlabelled material and 20% 1,3-SH3 (in red). The spectra are recorded using a mixing time of 500 ms. On the right side, the slices are shown, indicated with dotted lines in the 2D spectra.

5.5 Distance classes

To extract distance restraints from the cross-peak intensities, we recorded and analysed a set of ^{13}C-^{13}C PDSD spectra with mixing times of 50, 100, 200 and 500 ms. In Fig. 5.5 the spectra recorded on 2-SH3 are shown. The build-up of the integrated signal intensities was evaluated for a number of reference peaks. Conformation-independent distances between sequential α-carbons, methyl groups within leucine and valine and α- and γ-carbons within one amino-acid and between C$^{\beta}$ and C$^{\gamma}$ of the leucines were used as references. In general, the

intensities of peaks arising from shorter distances of 2.5 Å built up relatively fast reaching a maximum at around 100 ms, and then decayed slowly. The build-up curves corresponding to transfers between sequential α-carbons (~3.8 Å) progressively increase during the 500 ms mixing. The time dependence of the signal intensities involving interactions in the 4–6 Å range, like inter-strand C^{α}-C^{α} interactions in β-sheets, showed lag phases of varying length similar to effects in NOESY spectra of proteins in solution.

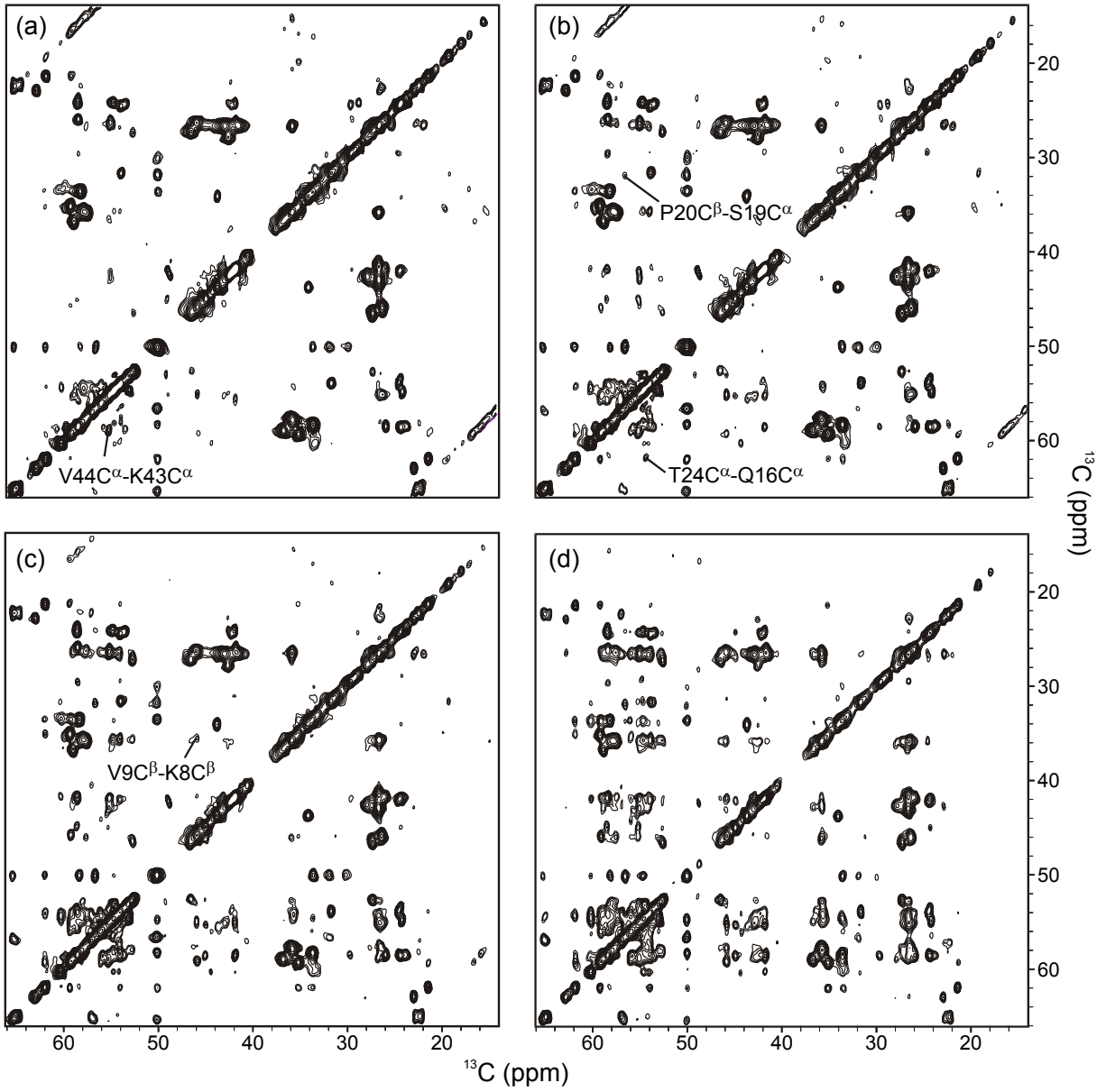

Fig. 5.5 2D ^{13}C-^{13}C PDSD spectra of 2-SH3, recorded using mixing times of 50 ms (a), 100 ms (b), 200 ms (c) and 500 ms (d). Some of the peaks used to define the four distance classes are indicated in the spectra with the relative assignment.

Since peak intensities are influenced by several effects, a rigorous quantitative evaluation was not feasible. We therefore categorised the carbon-carbon distance restraints

empirically in four restraint classes (Table 5.1). We determined the upper distance-boundaries of these classes according to the first appearance of the reference peaks at the various mixing times. Some examples are shown in Fig. 5.5. The reference distance for the first class was defined by cross-peaks between sequential C^{α} atoms (~3.8 Å) that appeared within 50 ms, whereas interactions between sequential C^{α} and C^{β} (~4.6 Å) and inter-strand C^{α}-C^{α} (4.6–5.4 Å) appeared first at 100 ms, defining the second class. The third class contained sequential C^{β}-C^{β} interactions (~5.8 Å). We assigned all other interactions to the fourth class, with distances in the range of 2.5–7.5 Å. The lower bound was always kept at 2.5 Å, to account for an apparent lower signal intensity due to incomplete suppression of dipolar truncation effects and/or fractional labelling. The carbon-carbon restraints were supplemented by the six nitrogen-nitrogen restraints, as discussed in Section 5.3, and classified in the long distance range 3–6 Å.

5.6 Structure calculations

A list of 292 inter-residue distance restraints for residues 7–61 was generated from 2D ^{13}C-^{13}C and ^{15}N-^{15}N spectroscopy and subjected to a conventional structure calculation protocol[5]. The fifteen lowest energy structures were selected out of two hundreds which represent the fold of the SH3 domain (Fig. 5.6a). The x-ray structure is shown below (Fig. 5.6b) for comparison[4]. The C^{α} coordinates of the regular structure elements show an rmsd of 1.6 Å to the average structure, and of 2.6 Å to the x-ray structure (see Tab. 5.1). The coordinates of the grey shaded loop region at the bottom consisting of residues 18–22 show a high degree of divergence in this set of structures. However, we do not attribute this to flexibility but rather to the lack of long-range intramolecular restraints.

Statistics for the α-spectrin SH3	
Restraints	
Total experimental restraints	292
Total inter-residue C-C restraints	286
Sequential ($\|i\text{-}j\| = 1$)	122
Medium range($1 < \|i\text{-}j\| \leq 4$)	15
Long range ($\|i\text{-}j\| > 4$)	149
Restraints in the class 2.5 < r < 4.5 Å	10
Restraints in the class 2.5 < r < 5.5 Å	21
Restraints in the class 2.5 < r < 6.5 Å	42
Restraints in the class 2.5 < r < 7.5 Å	213
Total inter-residue N-N restraints	6
Restraints in the class 3.0 < r < 6.0 Å	6
Distance restraint violations > 0.3 Å	0
Dihedral angle violations > 20°	0
Energies	
Final energies, kcal/mol	<15> †
E_{global}	53 ±6
E_{bonds}	1.7 ±0.4
E_{angles}	31 ±1
$E_{impropers}$	1.0 ±0.2
E_{vdW}	16 ±4
E_{DC}	3 ±1
Rmsd ‡	
Deviation from ideal values	<15>
Bonds, Å	0.0013 ±0.0001
Angles, °	0.329 ±0.008
Impropers, °	0.11 ±0.01
Backbone of β-sheet, Å §	1.6 ±0.2
<15> vs. x-ray for β-sheet elements, Å ¶	2.6 ±0.2

Tab. 5.1 Restraints and structural statistics for the α-spectrin SH3 ensemble.† <15> represents the average for the fifteen energy-minimized conformers.‡ Evaluated by CNS.§ The β-sheet comprises residues 7–11, 14–17, 23–26, 29–34, 41–46, 49–54 and 58–61.¶ R.m.s. deviation of the mean structure from the x-ray structure.

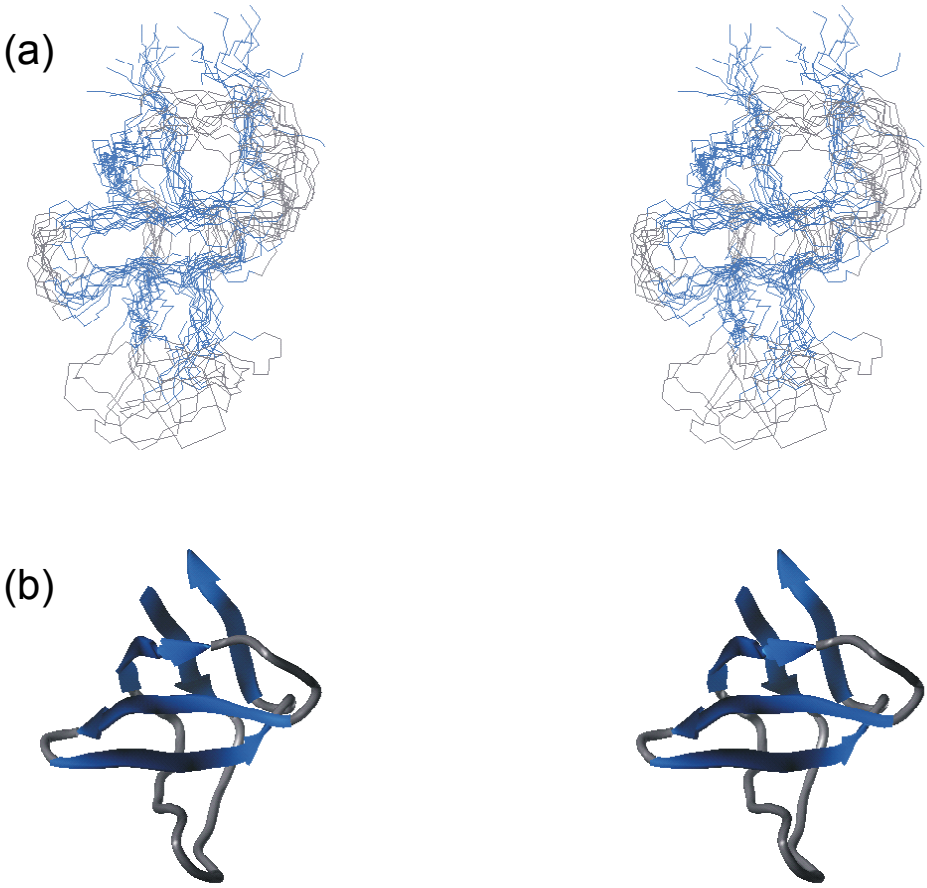

Fig. 5.6 Solid-state structure of the α-spectrin SH3 domain. (a), Stereoview of twelve of the fifteen lowest energy structures, representing the fold of the SH3 domain. The three structures with the largest rmsd to the average structure are not displayed. The β-strand regions are shown in blue. (b), The x-ray structure is shown for comparison. In this case, the part of the β-sheet in the region 14–17 and 23–26 is non-ideal and therefore not indicated in blue.

5.7 Conclusions

This structure exemplifies one of the first attempts to calculate the fold of a protein system from exclusively solid-state MAS data and satisfactorily describes the β-sandwich fold of the 62-residue SH3 domain. The application of 3D-techniques to resolve, for example, carbon-carbon correlation spectra via ^{15}N chemical shifts would be necessary for investigations on larger proteins or determination of structures with higher resolution. In chapter 6, a refined structure of the SH3 domain is obtained, by means of 3D ^{15}N-^{13}C-^{13}C correlation spectroscopy and chemical shift analysis.

5.8 Materials and methods

5.8.1 Sample preparation

The (U-^{13}C)-SH3, 2-SH3 and 1,3-SH3 protein sample were prepared as described in Chapter 3.4.1. Approximately 6 –7 mg of protein was used in all experiments, except for the 5-fold diluted samples, where 9–10 mg of protein were used.

5.8.2 Solid-state NMR spectroscopy

The 2D ^{13}C-^{13}C PDSD experiments were performed at a field of 17.6 T on a DMX-750 narrow-bore spectrometer, equipped with a 4 mm double-resonance MAS probe (Bruker, Karlsruhe, Germany). All 2D correlation spectra were acquired at MAS frequencies of 8.0 kHz or 13.0 kHz using TPPI for phase-sensitive detection. Ramped cross polarization from ^{1}H to ^{13}C created the initial transverse carbon magnetization; spin-lock fields were 36 kHz for ^{1}H and 18–36 kHz for the ^{13}C ramp. After the first ^{13}C evolution period, carbon magnetization was exchanged by using a PDSD mixing scheme[4]. Spin diffusion periods of 15–500 ms were applied. Typical carbon 90° pulse lengths were 5.3 μs. A proton *rf* field of ~60 kHz was applied for the two-pulse phase modulation decoupling during ^{13}C acquisition and evolution. The 2D ^{13}C-^{13}C spectra were recorded with 32–64 scans, and with ~ 6 ms evolution in the indirect dimension, leading to experimental times of 16 to 32 hours. The spectra of the 80% unlabelled and 20% labelled protein were recorded with 128 scans in 2.5 days. The ^{15}N-^{15}N PDSD spectrum was recorded under similar experimental conditions, but with a mixing time of 4 s and on a DMX-750 wide-bore spectrometer, equipped with a 4 mm triple-resonance MAS probe (Bruker, Karlsruhe, Germany).

All the solid-state data were processed with the XWINNMR software, version 2.6 (Bruker, Karlsruhe, Germany) and subsequently analysed using the program Sparky, version 3.100 (T.D. Goddard & D.G. Kneller, University of California).

5.8.3 Structure calculations

Structures were calculated with the program CNS version 1.0[5]. Calculations were performed using the simulating annealing protocol with torsion-angle dynamics, starting with 200

randomised conformers. The 286 ^{13}C-^{13}C restraints were categorised in strong (2.5–4.5 Å), medium (2.5–5.5 Å), weak (2.5–6.5 Å) or very weak (2.5–7.5 Å). The 6 ^{15}N-^{15}N correlations were constrained to 3–6 Å. Twelve of the fifteen lowest energy structures with no violations larger than 0.3 Å were used to represent the 3D fold of the α-spectrin SH3 domain. The atomic coordinates and NMR restraints were deposited in the RCSB Protein Data Bank, entry number 1M8M.

References

1. Szeverenyi, N. M., Sullivan, M. J., & Maciel, G. E. (1982). Observation of Spin Exchange by Two-Dimensional Fourier-Transform C-13 Cross Polarization-Magic-Angle Spinning. *J. Magn. Reson.* **47**, 462-475.

2. Hodgkinson, P. & Emsley, L. (1999). The accuracy of distance measurements in solid-state NMR. *J. Magn. Reson.* **139**, 46-59.

3. Kiihne, S., Mehta, M. A., Stringer, J. A., Gregory, D. M., Shiels, J. C., & Drobny, G. P. (1998). Distance measurements by dipolar recoupling two-dimensional solid-state NMR. *Journal of Physical Chemistry A* **102**, 2274-2282.

4. Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., & Saraste, M. (1992). Crystal structure of a Src-homology 3 (SH3) domain. *Nature* **359**, 851-855.

5. 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., & Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* **54 (Pt 5)**, 905-921.