Acknowledgements We thank L. N. Kapp for helping to prepare the manuscript; M. Jasin for providing the SCneo construct; A. Shinohara for anti-Rad51 antibody. We also thank H. Komata, C. Muranaka, T. Jo, M. Ueda, A. Kodama-Kamesako and A. Okamoto for laboratory assistance. This work was supported in part by the Ministry of Education, Science, Sports and Culture of Japan (H.T. and K.K.), and by the Nuclear Safety Research Association of Japan (H.T.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.T. (e-mail: htauchi@mx.ibaraki.ac.jp) or K.K. (e-mail: komatsu@house.rbc.kyoto-u.ac.jp).

Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy

Federica Castellani, Barth van Rossum, Annette Diehl, Mario Schubert, Kristina Rehbein & Hartmut Oschkinat

Forschungsinstitut für Molekulare Pharmakologie, Robert-Rössle-Strasse 10, 13125 Berlin, and Freie Universität Berlin, Takustrasse 3, 14195 Berlin, Germany

The determination of a representative set of protein structures is a chief aim in structural genomics. Solid-state NMR may have a crucial role in structural investigations of those proteins that do not easily form crystals or are not accessible to solution NMR, such as amyloid systems[1] or membrane proteins[2–4]. Here we present a protein structure determined by solid-state magic-angle-spinning (MAS) NMR. Almost complete ^{13}C and ^{15}N resonance assignments for a micro-crystalline preparation of the α-spectrin Src-homology 3 (SH3) domain[5] formed the basis for the extraction of a set of distance restraints. These restraints were derived from proton-driven spin diffusion (PDSD) spectra of biosynthetically site-directed, labelled samples obtained from bacteria grown using [1,3-^{13}C]glycerol or [2-^{13}C]glycerol as carbon sources. This allowed the observation of long-range distance correlations up to ~7 Å. The calculated global fold of the α-spectrin SH3 domain is based on 286 inter-residue ^{13}C–^{13}C and six ^{15}N–^{15}N restraints, all self-consistently obtained by solid-state MAS NMR. This MAS NMR procedure should be widely applicable to small membrane proteins that can be expressed in bacteria.

Solid-state NMR studies are feasible if suitably labelled samples of high structural homogeneity are available. Isotope labelling is a requirement for the applicability of NMR techniques to achieve sequence-specific resonance assignments[5–7] or to estimate distance restraints and torsion angles[8–10]. Membrane proteins can be investigated by solid-state NMR using static oriented samples, as was demonstrated with the structure of gramicidin[11]. Alternatively, MAS applied to oriented or non-oriented samples produces sufficiently resolved spectra and yields isotropic chemical shifts which give rise to amino-acid-specific correlation patterns that are important for assignment strategies, heightening expectations that structural studies of small membrane proteins or their bound ligands—such as 4–60-residue receptor agonists or antagonists—are in principle viable by solid-state NMR[12–15]. However, a suitable methodology for the collection of a large number of distance restraints in the range of 2–7 Å from a small number of samples and with minimal experimental effort is still lacking. By examining all potential short atomic distances (≤7 Å) in proteins and considering the limited resolution of the proton spectrum, it is evident that carbon–carbon restraints are of prime importance for defining a three-dimensional structure. The measurement of such restraints is hindered in fully ^{13}C-enriched samples by dipolar truncation effects[16,17]. Suppression of these effects can be achieved by applying spectroscopic techniques that contain frequency-selective dipolar recoupling schemes[9,18,19] or by combining broadband recoupling methods with dilution of ^{13}C spins[8].

We prepared a set of differently labelled samples of the α-spectrin SH3 domain[20], which enabled the semi-quantitative interpretation of cross-peak intensities. The set consisted of a uniformly ^{13}C-labelled preparation (referred to as U-SH3) and two biosynthetically site-directed ^{13}C-enriched samples, obtained by growing bacteria on [1,3-^{13}C]glycerol (1,3-SH3) and [2-^{13}C]glycerol (2-SH3)[21]. The amino-acid labelling pattern of 1,3-SH3 is shown in Fig. 1a. Amino acids synthesized from precursors made in the glycolysis pathway (group I) show a labelling pattern composed of ~100% (green) and ~0% (red) ^{13}C-enriched sites. The inverse labelling scheme is obtained with 2-SH3. A more complex labelling scheme is obtained for the amino acids synthesized from precursors related to the citric acid cycle (group II). Several isotopomers with different labelling patterns are then produced, which are shown in Fig. 1a for the case of proline. The percentage of labelling in each site (Fig. 1a), resulting from the mixture of isotopomers, was estimated from solution NMR data.

The alternating labelling pattern in 1,3-SH3 and 2-SH3 leads to a reduction of dipolar truncation effects and allows the observation of long-range interactions, while relayed polarization transfer is blocked. These effects are demonstrated in Fig. 1b–d, which

Table 1 **^{13}C–^{13}C and ^{15}N–^{15}N distance data and statistics for α-spectrin SH3**

(**a**) Average inter-strand and sequential ^{13}C–^{13}C and ^{15}N–^{15}N distances*

Inter-strand	Distance (Å)	Sequential	Distance (Å)
$C\alpha_i$–$C\alpha_j$†	4.6 ± 0.3	$C\alpha_i$–$C\alpha_{i-1}$	3.8 ± 0.1
$C\alpha_{i-1}$–$C\alpha_{j+1}$	5.4 ± 0.3	$C\alpha_i$–$C\beta_{i-1}$	4.6 ± 0.3
		$C\beta_i$–$C\beta_{i-1}$	5.8 ± 0.4
CO_{i-1}–CO_j	4.9 ± 0.3	CO_i–CO_{i-1}	3.5 ± 0.3
CO_{i-1}–CO_{j+1}	5.0 ± 0.4		
N_{i+1}–N_j	5.1 ± 0.2	N_i–N_{i-1}	3.5 ± 0.2
N_i–N_j	5.8 ± 0.3		
N_{i-1}–N_{j+1}	4.5 ± 0.2		

(**b**) Restraints and structural statistics for the α-spectrin SH3 ensemble

Restraints	
Total experimental restraints	292
Total inter-residue C–C restraints	286
Sequential ($\|i-j\|=1$)	122
Medium range ($1<\|i-j\|\le 4$)	15
Long range ($\|i-j\|>4$)	149
Restraints in the class 2.5–4.5 Å	10
Restraints in the class 2.5–5.5 Å	21
Restraints in the class 2.5–6.5 Å	42
Restraints in the class 2.5–7.5 Å	213
Total inter-residue N–N restraints	6
Restraints in the class $3.0<r<6.0$ Å	6
Distance constraint violations > 0.3 Å	0
Dihedral angle violations > 20°	0
Energies	
Final energies (kcal mol^{-1})	⟨15⟩‡
E_{global}	53 ± 6
E_{bonds}	1.7 ± 0.4
E_{angles}	31 ± 1
$E_{impropers}$#	1.0 ± 0.2
$E_{van\ der\ Waals}$	16 ± 4
$E_{dipolar\ couplings}$	3 ± 1
r.m.s. deviation§	
Deviation from ideal values	⟨15⟩
Bonds (Å)	0.0013 ± 0.0001
Angles (deg)	0.329 ± 0.008
Impropers (deg)	0.11 ± 0.01
Backbone of β-sheet (Å)‖	1.6 ± 0.3
⟨15⟩ versus X-ray for β-sheet elements (Å)¶	2.6 ± 0.2

*Distances shown are for anti-parallel β-sheets.
†The indices *i*, *j* refer to the top left panel of Fig. 2.
‡⟨15⟩ represents the average for the 15 energy-minimized conformers.
§Evaluated by CNS[25].
‖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[26].
#Impropers are angles for maintaining geometries[25].

shows contour plots of two-dimensional ^{13}C–^{13}C PDSD[22] spectra of U-SH3 (Fig. 1b) and of 2-SH3 and 1,3-SH3 (Fig. 1c, d, respectively). All spectra were recorded with a long mixing time of 500 ms to establish long-range correlations. An advantage of the PDSD technique[22] is that longer mixing times can be applied without limitations on balancing radio-frequency-power input and sample heating. In Fig. 1b, mainly intra-residue cross-peaks are observed[5] with additional correlations, attributed to relayed polarization transfers within an amino acid or to transfers between sequentially connected amino acids. The assignment of the correlations in Fig. 1b is, however, difficult owing to overlap and ambiguity. In contrast, the 2-SH3 and 1,3-SH3 spectra (Fig. 1c, d) are simplified owing to the less extensive labelling. The spectrum of the 2-SH3 preparation (Fig. 1c) is virtually empty in the methyl region between 16 and 28 p.p.m. On the other hand, additional peaks are observed particularly in the region from 48–66 p.p.m., which can be assigned to long-range polarization transfer events. The spectrum of 1,3-SH3 (Fig. 1d) shows a large number of new peaks in the methyl region, which can also be attributed to long-range interactions. Positive side-effects of this type of labelling are the simplification of the assignment of the new correlations and the overall improved resolution as a result of suppressing homonuclear one-bond *J*-couplings[8].

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. For example, the assignment of long-range cross-peaks involving L33 and V44 is shown in Fig. 2 (panels I–IV), which contains the regions indicated in Fig. 1c, d. Furthermore, we could identify the secondary structure and topology of the SH3 domain from correlation patterns involving residues 14 and 26, 15 and 25, 16 and 24, or between 34 and 43, 33 and 44, 31 and 46, which indicate the presence of an antiparallel β-sheet. Additionally, at around 62 p.p.m. (Fig. 2, panel V), where the Cα of T24 and P54 resonate, all cross-peaks involving the two sequential α-carbons are observed, as is a correlation between the α-carbon signals of P54 and W41. The inter-strand Cα–Cα correlation between T24 and Q16 is also observed in the same region but only in the 1,3-SH3 spectrum, owing to a low percentage of Q16 Cα-labelling in the

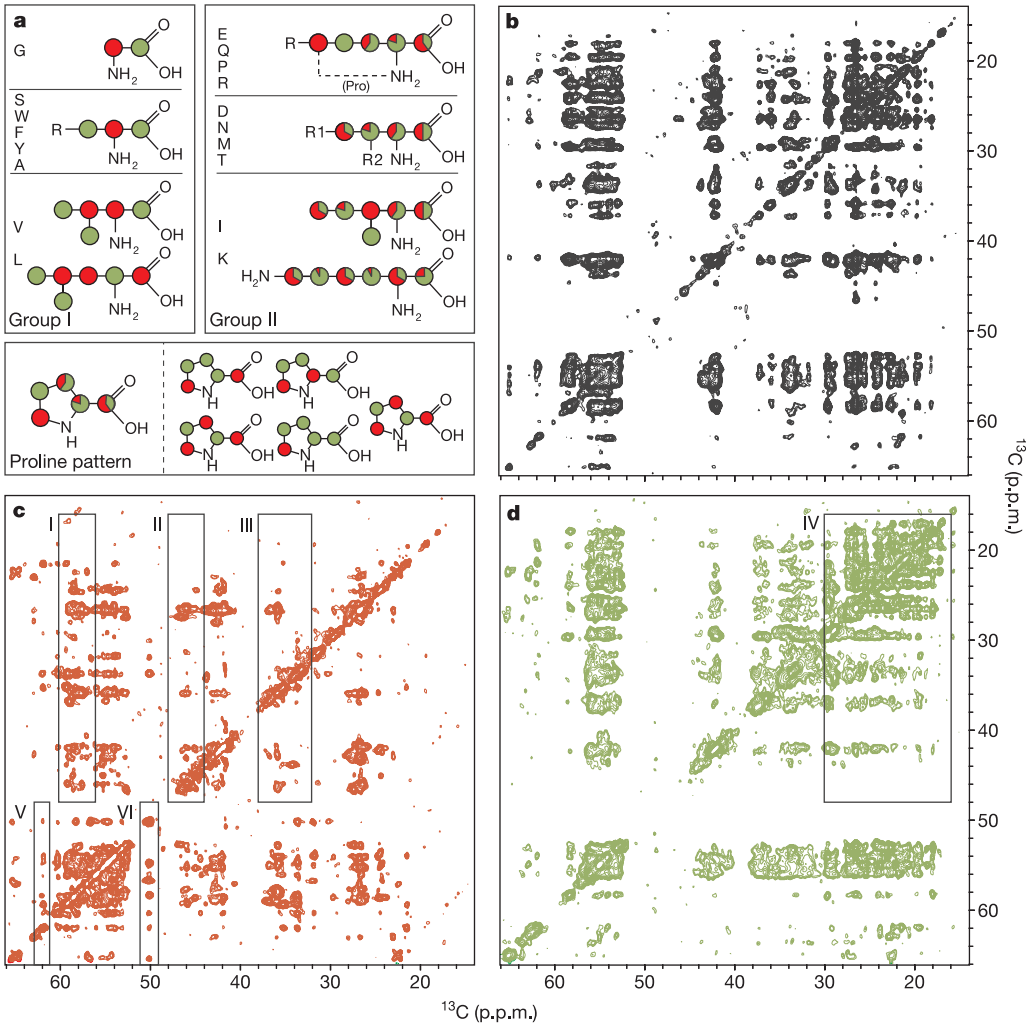

Figure 1 Labelling patterns and NMR spectra for the different α-spectrin SH3 domain preparations. **a**, Schematic representation of the effective ^{13}C enrichment for the indicated residues, as obtained by protein expression in *E. coli* BL21 (DE3). The green colour corresponds to the degree of ^{13}C labelling pattern obtained by growth on [1,3-^{13}C]glycerol; the opposite labelling pattern, obtained by growth on [2-^{13}C]glycerol, is represented in red. In cases with mixed labelling, the percentage label from [2-^{13}C]glycerol and [1,3-^{13}C]glycerol for a particular atom is represented using relative red/green colouring. The percentages of labelling for each site were estimated by solution NMR from satellite signals. The fractional labelling for residues of group II is explained by the production of different labelling patterns, as illustrated for proline at the bottom. **b**–**d**, Contour plots of two-dimensional ^{13}C–^{13}C PDSD spectra recorded with a mixing time of 500 ms at 17.6 T and a spinning frequency of $\omega_R/2\pi = 8.0$ kHz on U-SH3 (**b**), on 2-SH3 (**c**) and on 1,3-SH3 (**d**). The colours of the spectra (red for 2-SH3 and green for the 1,3-SH3) correspond to the labelling pattern in **a**. The boxed regions (I–VI) in **c** and **d** relate to Fig. 2.

2-SH3 preparation. These observations reflect systematically appearing distances between α-carbons in adjacent strands that range from 4.6–5.4 Å (Fig. 2 top left and Table 1a). Similarly, inter-strand distances between nitrogens, CO and Cα are of that order (Table 1a). These characteristic β-sheet cross-peak patterns prove the reliability of the assignment. Likewise, characteristic signal patterns (for example, between residues i and $i+3$ or $i+4$) are expected in spectra of α-helical proteins.

To extract distance restraints from the cross-peak intensities, we recorded and analysed a set of spectra with mixing times of 50, 100, 200 and 500 ms. 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 Cβ and Cγ 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) increased in the first 500 ms. The time dependence of the signal intensities involving interactions in the 4–6 Å range, such as inter-strand Cα–Cα interactions in β-sheets, showed lag phases of varying length similar to effects in NOESY spectra of proteins in solution[23]. As peak intensities are influenced by several effects, a rigorous quantitative evaluation was not feasible. We therefore categorized the carbon–carbon distance restraints empirically in four restraint classes (Table 1b). We determined the upper distance boundaries of these classes according to the first appearance of the reference peaks at the various mixing times. The reference distance for the first class was defined by cross-peaks between sequential Cα atoms (~3.8 Å) that appeared within 50 ms, whereas interactions between sequential Cα and Cβ (~4.6 Å) and inter-strand Cα–Cα (4.6–5.4 Å) appeared first at 100 ms, defining the second class. The third class contained sequential Cβ–Cβ 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 six nitrogen–nitrogen restraints, primarily defining a β-sheet architecture. They were obtained from a single ^{15}N–^{15}N PDSD spectrum with a mixing time of 4.0 s, and classified in the long-distance range 3–6 Å. Few long-range ^{15}N–^{15}N correlations were detected because the couplings between ^{15}N spins are approximately sixfold weaker than couplings between carbon spins at comparable distances, and nitrogen atoms form a less dense network relative to carbon atoms.

In solid samples of proteins, intermolecular interactions are expected to cause additional cross-peaks in the 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. As 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 PDSD spectra with 500 ms mixing time was recorded on these diluted samples in which correlations attributable to intermolecular contacts were reduced approximately fivefold relative to the intramolecular contacts. In this way, contacts between residues Y57–S19, Y57–P20, R21–L8 and P20–Y13 were identified as intermolecular and removed from the restraint list, in line with earlier observations[24].

A list of 292 inter-residue distance restraints for residues 7–61 was generated and subjected to a conventional structure calculation

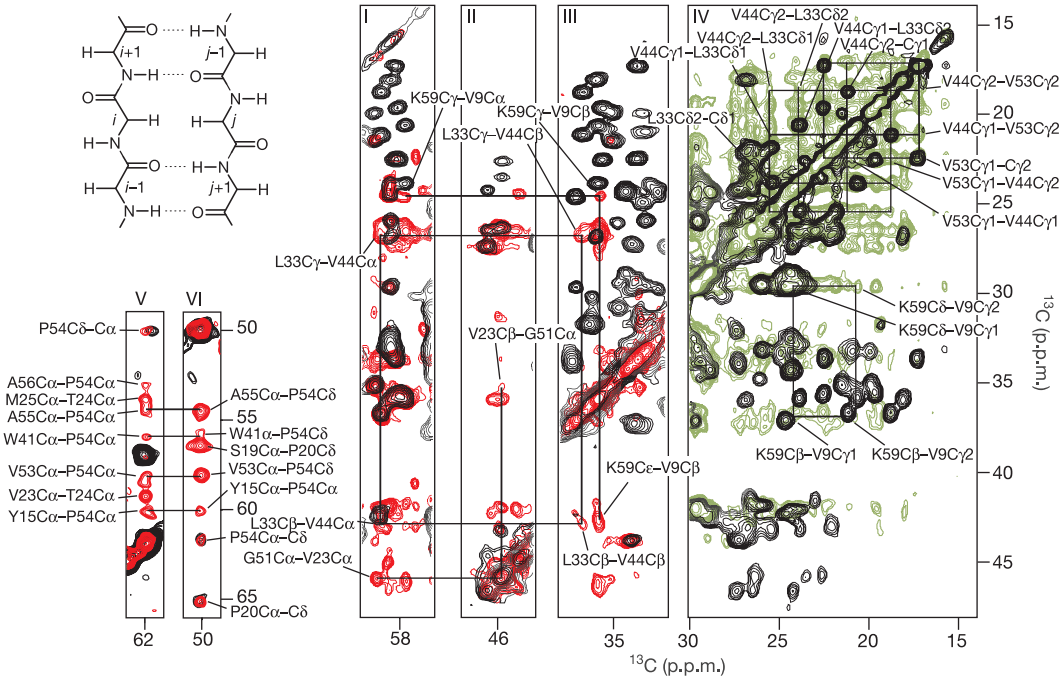

Figure 2 Assignment strategy. Regions extracted from the spectra of Fig. 1c and d (I–VI) superimposed on a PDSD spectrum of uniformly labelled SH3 domain (black), the latter recorded with a short mixing time of 15 ms (ref. 5. Part of the assignment of the long-range correlations is reported in the figure, and the lines define the different correlation patterns. As an example, correlations between residues L33 and V44 are shown. The correlations between the Cα and Cβ signals of V44 and Cβ and Cγ signals of L33 are observed in panels I–III 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 p.p.m. (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 the U-SH3 sample (see Fig. 1b), no correlations are detected. In the upper left corner, a schematic representation of an antiparallel β-sheet is shown. The numbering of the residues (i and j) corresponds to the numbering in Table 1.

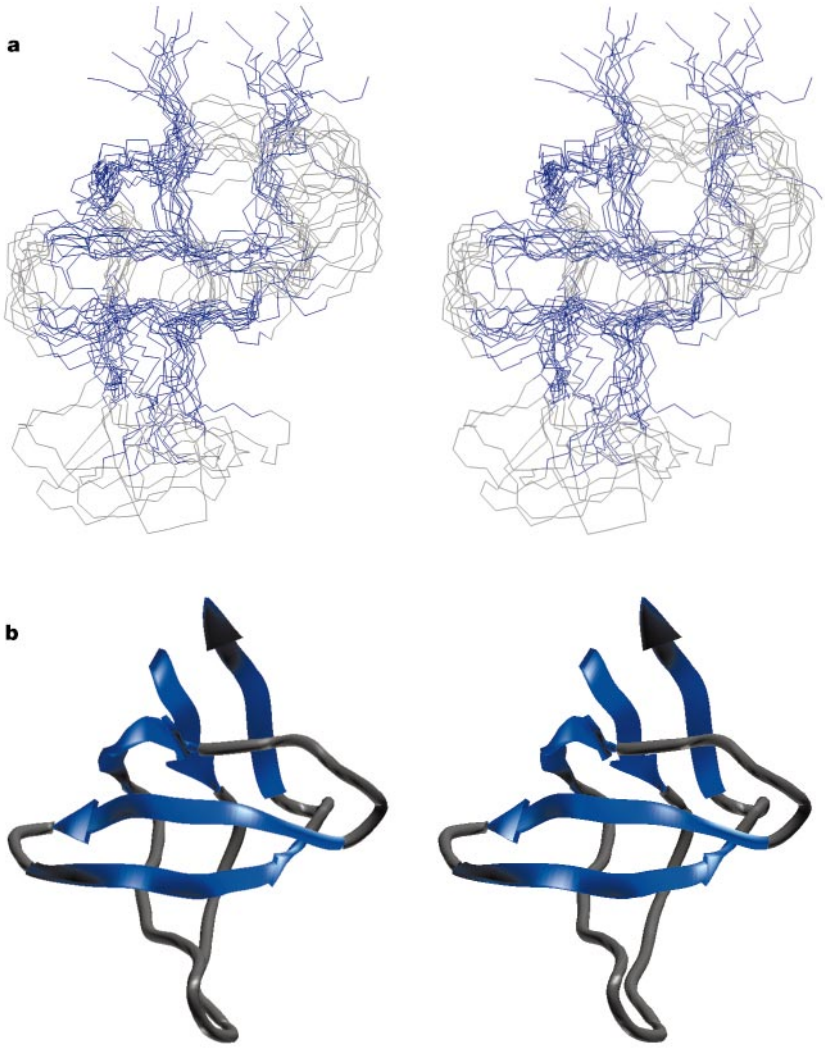

Figure 3 Solid-state structure of the α-spectrin SH3 domain. **a**, Stereo view of twelve of the fifteen lowest-energy structures, representing the fold of the SH3 domain. The three structures with the largest r.m.s. deviation to the average structure are not displayed. The β-strand regions are shown in blue. **b**, The X-ray structure[26] 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 is not indicated in blue.

protocol[25]. The 15 lowest-energy structures were selected out of 200, which represent the fold of the SH3 domain (Fig. 3a). The X-ray structure is shown below (Fig. 3b) for comparison[26]. The Cα coordinates of the regular structure elements show a root mean square (r.m.s.) deviation of 1.6 Å to the average structure, and of 2.6 Å to the X-ray structure (Table 1b). 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.

We present a solid-state MAS NMR structure that satisfactorily describes the β-sandwich fold of the 62-residue SH3 domain. This result paves the way for the structure determination of amyloids, small membrane proteins or receptor–ligand complexes, provided that the proteins can be expressed in bacteria and suitably labelled[4]. Although the current work is performed on micro-crystalline material, its periodicity is not a requirement for the applicability of the presented methodology, but aids in providing a structurally homogeneous environment. For membrane proteins, structural or conformational homogeneity is supported by either embedding the protein into the membrane or embedding an agonist (or antagonist) into the receptor-binding site, which serves as a matrix and forces the ligand into a unique environment ensuring sufficiently narrow lines. Our method provides a sufficiently large number of long-range carbon–carbon distance restraints with reasonable precision from dipolar carbon–carbon correlation spectroscopy. The data set may be supplemented by ^{1}H–^{1}H restraints obtained from proton–proton correlation spectroscopy[10,27] using samples deuterated at non-exchangeable sites, and by ^{13}C–^{15}N restraints from C–N dipolar correlation spectroscopy using doubly-labelled samples. In the foreseeable future, techniques will be developed to include ^{1}H–^{13}C and/or ^{1}H–^{15}N restraints. In general, the application of three-dimensional techniques to resolve, for example, carbon–carbon correlation spectra by means of ^{15}N chemical shifts would be necessary to study larger proteins. However, for the small SH3 domain presented here, two-dimensional spectroscopy was sufficient. The study of bacterial membrane proteins within structural genomics projects is an obvious next application of the new technique described herein, together with the study of receptor-bound agonists and antagonists. □

Methods

Sample preparation

Plasmid pET3d coding for α-spectrin SH3 protein from chicken brain was a gift of

M. Saraste. The SH3 protein was expressed in *Escherichia coli* BL21 (DE3), using M9 minimal medium. A total of 0.5 g $^{15}NH_4Cl$, 2.0 g $NaH^{12}CO_3$ and 2.0 g [1,3-^{13}C]glycerol per litre of medium was added in the case of ([1,3-^{13}C]glycerol, ^{15}N) SH3. For the preparation of the ([2-^{13}C]glycerol, ^{15}N) SH3 sample, 0.5 g $^{15}NH_4Cl$, 2.0 g $NaH^{13}CO_3$ and 2.0 g [2-^{13}C]glycerol were used[21]. The proteins were purified as previously described[20]. The final yield was approximately 20 mg of protein per litre of culture. A 200 mM $(NH_4)_2SO_4$ solution (pH 3.5, 0.04% NaN_3) was added to a 1.15 mM SH3 solution (pH 3.5) at a volume ratio of 1:1. The protein samples for MAS measurement were precipitated as previously described[20].

NMR spectroscopy

All solid-state carbon–carbon NMR 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). The protein was confined to the centre of the rotor with the use of spacers. Approximately 6–7 mg of protein was used in all experiments, except for the fivefold diluted samples, where 9–10 mg of protein was used. All two-dimensional correlation spectra were acquired at MAS frequencies of 8.0 kHz or 13.0 kHz using time-proportional phase incrementation 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[22]. Spin diffusion periods of 50–500 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[28] during ^{13}C acquisition and evolution. The two-dimensional ^{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 16–32 h. 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).

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

Structure calculations

Structures were calculated with the program CNS version 1.0 (ref. 25). Calculations were performed using the simulating annealing protocol with torsion-angle dynamics, starting with 200 randomized conformers. The 286 ^{13}C–^{13}C restraints were categorized into 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) classes. The six ^{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 three-dimensional fold of the α-spectrin SH3 domain. The atomic coordinates and NMR restraints were deposited in the RCSB Protein DataBank under entry number 1M8M.

Received 30 January; accepted 6 August 2002; doi:10.1038/nature01070.

1. Tycko, R. Biomolecular solid state NMR: Advances in structural methodology and applications to peptide and protein fibrils. *Annu. Rev. Phys. Chem.* **52,** 575–606 (2001).
2. Opella, S. J. NMR and membrane proteins. *Nature Struct. Biol.* **4** suppl., 845–848 (1997).
3. Griffin, R. G. Dipolar recoupling in MAS spectra of biological solids. *Nature Struct. Biol.* **5** NMR suppl., 508–512 (1998).
4. De Groot, H. J. M. Solid-state NMR spectroscopy applied to membrane proteins. *Curr. Opin. Struct Biol.* **10,** 593–600 (2000).
5. Pauli, J., Baldus, M., van Rossum, B., de Groot, H. & Oschkinat, H. Backbone and side-chain ^{13}C and ^{15}N signal assignments of the α-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 tesla. *ChemBioChem* **2,** 272–281 (2001).
6. Straus, S. K., Bremi, T. & Ernst, R. R. Experiments and strategies for the assignment of fully ^{13}C/^{15}N-labelled polypeptides by solid-state NMR. *J. Biomol. NMR* **12,** 39–50 (1998).
7. Hong, M. Resonance assignment of ^{13}C/^{15}N labelled solid proteins by two-and three-dimensional magic-angle-spinning NMR. *J. Biomol. NMR.* **15,** 1–14 (1999).
8. Hong, M. Determination of multiple φ-torsion angles in proteins by selective and extensive ^{13}C labeling and two-dimensional solid-state NMR. *J. Magn. Reson.* **139,** 389–401 (1999).
9. Jaroniec, C. P., Tounge, B. A., Herzfeld, J. & Griffin, R. G. Frequency selective heteronuclear dipolar recoupling in rotating solids: accurate ^{13}C-^{15}N distance measurements in uniformly ^{13}C,^{15}N-labeled peptides. *J. Am. Chem. Soc.* **123,** 3507–3519 (2001).
10. Brown, S. P. & Spiess, H. W. Advanced solid-state NMR methods for the elucidation of structure and dynamics of molecular, macromolecular, and supramolecular systems. *Chem. Rev.* **101,** 4125–4156 (2001).
11. Ketchem, R. R., Lee, K.-C., Huo, S. & Cross, T. A. Macromolecular structural elucidation with solid-state NMR-derived orientational constraints. *J. Biomol. NMR* **8,** 1–14 (1996).
12. Shochat, S. *et al.* ^{13}C MAS NMR evidence for a homogeneously ordered environment of tyrosine M210 in reaction centres of *Rhodobacter sphaeroides. Spectrochim. Acta* **51A,** 135–144 (1995).
13. Verhoeven, M. A. *et al.* Ultra-high-field MAS NMR assay of a multispin labeled ligand bound to its G-protein receptor target in the natural membrane environment: electronic structure of the retilylidene chromophore in rhodopsin. *Biochemistry* **40,** 3282–3288 (2001).
14. Egorova-Zachernyuk, T. A. *et al.* Characterization of pheophytin ground states in *Rhodobacter sphaeroides* R26 photosynthetic reaction centers from multispin pheophytin enrichment and 2-D ^{13}C MAS NMR dipolar correlation spectroscopy. *Biochemistry* **36,** 7513–7519 (1997).
15. Griffiths, J. M. *et al.* Dipolar correlation NMR spectroscopy of a membrane protein. *J. Am. Chem. Soc.* **116,** 10178–10181 (1994).
16. Hodgkinson, P. & Emsley, L. The accuracy of distance measurements in solid-state NMR. *J. Magn. Reson.* **139,** 46–59 (1999).
17. Kiihne, S. *et al.* Distance measurements by dipolar recoupling two-dimensional solid-state NMR. *J. Phys. Chem. A* **102,** 2274–2282 (1998).
18. Nielsen, N. C., Bildsoe, H., Jakobsen, H. J. & Levitt, M. H. Double-quantum homonuclear rotary resonance: efficient dipolar recovery in magic-angle spinning nuclear magnetic resonance. *J. Chem. Phys.* **101,** 1805–1812 (1994).
19. Costa, P. R., Sun, B. Q. & Griffin, R. G. Rotational resonance tickling: accurate internuclear distance measurement in solids. *J. Am. Chem. Soc.* **119,** 10821–10830 (1997).
20. Pauli, J., van Rossum, B., Förster, H., De Groot, H. J. M. & Oschkinat, H. Sample optimization and identification of signal patterns of amino acid side chains in 2D RFDR spectra of the α-spectrin SH3 domain. *J. Magn. Reson.* **143,** 411–416 (2000).
21. LeMaster, D. M. & Kushlan, D. M. Dynamical mapping of *E. coli* thioredoxin via ^{13}C NMR relaxation analysis. *J. Am. Chem. Soc.* **118,** 9255–9264 (1996).
22. Szeverenyi, N. M., Sullivan, M. J. & Maciel, G. E. ^{13}C Spin exchange by 2D FT ^{13}C CP/MAS. *J. Magn. Reson.* **47,** 462–475 (1982).
23. Wüthrich, K. *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
24. van Rossum, B. J., Castellani, F., Rehbein, K., Pauli, J. & Oschkinat, H. Assignment of the non exchanging protons of the α-spectrin SH3 domain by two- and three-dimensional ^{1}H-^{13}C solid-state magic-angle-spinning NMR and comparison of solution and solid-state proton chemical shifts. *ChemBioChem* **2,** 906–914 (2001).
25. Brünger, A. T. *et al.* Crystallography and NMR system (CNS): a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
26. Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. & Saraste, M. Crystal structure of a Src-homology 3 (SH3) domain. *Nature* **359,** 851–855 (1992).
27. Reif, B., Jaroniec, C. P., Rienstra, C. M., Hohwy, M. & Griffin, R. G. ^{1}H-^{1}H MAS correlation spectroscopy and distance measurements in a deuterated peptide. *J. Magn. Reson.* **151,** 320–327 (2001).
28. Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V. & Griffin, R. G. Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **103,** 6951–6958 (1995).

Acknowledgements H. de Groot is acknowledged for access to the high-field NMR facility in Leiden. The authors thank R. Kühne, P. Schmieder, G. Krause and C. Glaubitz for discussions, and L. Ball, K. Heuer and K. Zierler for carefully reading the manuscript.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.O. (e-mail: oschkinat@fmp-berlin.de) or B.v.R. (brossum@fmp-berlin.de).

Absolute comparison of simulated and experimental protein-folding dynamics

Christopher D. Snow*†, Houbi Nguyen†‡, Vijay S. Pande* & Martin Gruebele‡

** Biophysics Program and Department of Chemistry, Stanford University, Stanford, California 94305-5080, USA*
‡ Departments of Chemistry and Physics, and Center for Biophysics and Computational Biology, University of Illinois, Urbana, Illinois 61801, USA
† These authors contributed equally to this work

Protein folding is difficult to simulate with classical molecular dynamics. Secondary structure motifs such as α-helices and β-hairpins can form in 0.1–10 μs (ref. 1), whereas small proteins have been shown to fold completely in tens of microseconds[2]. The longest folding simulation to date is a single 1-μs simulation of the villin headpiece[3]; however, such single runs may miss many features of the folding process as it is a heterogeneous reaction involving an ensemble of transition states[4,5]. Here, we have used a distributed computing implementation to produce tens of thousands of 5–20-ns trajectories (700 μs) to simulate mutants of the designed mini-protein BBA5. The fast relaxation dynamics these predict were compared with the results of laser temperature-jump experiments. Our computational predictions are in excellent agreement with the experimentally determined mean folding times and equilibrium constants. The rapid folding of BBA5 is due to the swift formation of secondary structure. The convergence of experimentally and computationally accessible timescales will allow the comparison of absolute quantities characterizing *in vitro* and *in silico* (computed) protein folding[6].

