The spectrin repeat folds into a three-helix bundle in solution

Jaime Pascual[a], Mark Pfuhl[a], German Rivas[b], Annalisa Pastore[a], Matti Saraste[a,*]

[a]European Molecular Biology Laboratory, Meyerhofstr. 1, Postfach 102209, D-69012 Heidelberg, Germany
[b]Centro de Investigaciones Biologicas, CSIC, Velazquez 144, 28006 Madrid, Spain

Received 16 February 1996

Abstract Spectrin, a major component of the membrane skeleton, is mainly composed of tandemly repeated segments of approx. 106 amino acids. We have undertaken the determination of the three-dimensional structure of a chicken brain α-spectrin repeat by heteronuclear multidimensional NMR. Sedimentation equilibrium demonstrates that this repeat is monomeric at the concentration used for NMR (1 mM). Its secondary structure was identified using a collection of sequential and medium range NOEs, chemical shifts, HN-Hα coupling constants, and relaxation measurements. These data unequivocally demonstrate the presence of three long helices connected by two loops. A set of interhelical NOEs indicates that the helices assemble into a triple helical structure. Our results provide experimental evidence supporting the triple-helical bundle proposed by modelling.

Key words: Spectrin repeat; Three-helix bundle; Heteronuclear NMR; Membrane skeleton

1. Introduction

Spectrin is a major component of the membrane-associated cortical cytoskeleton. As a tetramer, it cross-links filamentous actin forming a meshwork that was originally described in erythrocytes [1]. However, spectrin (also called fodrin) is a ubiquitous structural protein which is present in most vertebrate tissues as well as in non-vertebrates such as *Drosophila*, *Dictyostelium*, *C. elegans* and possibly in higher plants [2]. Spectrin is made of two antiparallel chains. The α/β-heterodimer is an elongated molecule, 100 nm in length and 3–5 nm in width. The β-chain contains an actin-binding site that is present in the proteins of the spectrin/filamin/fimbrin family [3,4], and is made of two calponin-homology or CH domains [5]. These N-terminal CH domains of the β-chain and the C-terminal domain of the α-chain, which is structurally similar to calmodulin and contains EF-hand motifs [6,7], appear to form the contact between the two spectrin chains in the dimer [8]. Two dimers interact to form a tetramer through the partial repeats at the C-terminus of the β-chain and at the N-terminus of the α-chain [9]. Spectrin also shares domains such as SH3 [10] and PH [11] with signaling proteins.

All members of the spectrin subfamily (spectrin, dystrophin and α-actinin) contain sequence repeats of approx. 106 amino acids showing a low sequence identity (~20%) with a predicted helical structure [12]. Experimental studies using spectroscopy and protease resistance have led to the determination of the domain boundaries (phasing) and have shown that the repeat is an independent folding unit with a high helical content [13–15]. Structural models for the repeat predict a left-handed coiled-coil made by three α-helices (A–C) separated by two short loops (AB, BC) [16], although the probability for a coiled-coil fold according to the method of [17] is low. However, the X-ray structure of the 14th repeat from the *Drosophila* α-spectrin [18] shows a homodimer which has a different secondary and tertiary structure compared to that of the proposed model [16]. At the secondary structure level, helices B and C are continuous and the proposed BC loop is in fact helical. As this arrangement may be a consequence of the dimerization of the *Drosophila* repeat in solution, the authors interpret their X-ray data with a model for the monomer in which the BC loop is present. With respect to the tertiary structure, the crystal dimer is not a helical coiled-coil but rather a bundle in which the packing of helices A with B is typical of a coiled-coil (ridge to ridge) whereas the packing of C with A or B is similar to what is commonly observed in globular proteins (ridge into groove) [18]. The observed packing could be influenced by the inter-monomeric contacts as helices A and B belong to one monomer and helix C to the other in the crystal structure. This may resemble the helical packing found between the partial repeats in the α/β tetramerization site [4].

We have started a series of experiments to determine the solution structure of an isolated spectrin repeat by nuclear magnetic resonance (NMR) methods. Previous studies [15] have shown that the recombinant 16th repeat (R16) of chicken brain α-spectrin (from residues 1763 to 1872) [19] folds into a stable and monomeric helical structure. Here we present the backbone assignment, the secondary structure of R16 derived from NOE (nuclear Overhauser effect) data, coupling constants, relaxation measurements and chemical shifts as well as the packing between the helices. Sedimentation equilibrium experiments show that R16 is a stable monomer even at 1 mM concentration used in the NMR experiments. This paper presents experimental evidence that the spectrin repeat does indeed have a fold with three helices separated by two loops as predicted by theoretical models.

2. Materials and methods

2.1. Protein expression and purification

The expression construct was generated from the cDNA for the α-subunit of chicken brain spectrin [19] by standard methods using polymerase chain reaction (PCR) [20]. The PCR product was purified and ligated to a pET3d vector [21] for transformation of the *E. coli* BL21(DE3) strain. DNA sequencing confirmed the identity of the insert. The clones were grown overnight at 37°C in LB medium containing 100 μg/ml ampicillin and diluted 1:100 into 1 l volume of LB + ampicillin. 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG)

*Corresponding author. Fax: (49) (6221) 387306.

Abbreviations: R16, 16-th repeat of chicken brain α-spectrin; NOE, nuclear Overhauser effect; 2D and 3D, two-and three-dimensional, respectively; NOESY, NOE spectroscopy; TOCSY, total correlation spectroscopy; TOWNY, TOCSY without NOESY; TSP, 3-(trimethylsilyl)propionate; T_1, longitudinal relaxation time; T_2, transverse relaxation time.

SSDI S0014-5793(96)00251-7

was added to induce expression when the absorbance at 600 nm was 0.5. After shaking for 3 h at 37°C, the cells were harvested by centrifugation and resuspended in 20 mM Tris-HCl (pH 8) containing 1 mM EDTA and 1 mM dithiothreitol. After passage through a French press and centrifugation to pellet unbroken cells and debris, the supernatant was passed through a 0.22 μm filter and loaded onto a cation exchange S-Sepharose column (Pharmacia) equilibrated with 20 mM MES (pH 6) buffer. R16-containing fractions were concentrated and further purified by gel filtration on a Superdex 75 column (Pharmacia) using 10 mM potassium phosphate buffer at pH 6. A fast performance liquid chromatography (FPLC) instrument (Pharmacia) was used in protein purification. Purity of the protein was monitored by SDS-PAGE [22] and electro-spray mass spectroscopy. N-terminal sequencing of the protein showed that the first residue (Met) was missing. The protein concentration was determined according to [23], obtaining a molar extinction coefficient in the buffer used of 11 415 M^{-1} cm^{-1}. All NMR samples were prepared in 10 mM potassium phosphate buffer (pH 6) in 90% H_2O/10% D_2O or 100% D_2O at 1 mM protein concentration. Integrity of the protein was checked after acquisition of the NMR spectra by SDS-PAGE and mass spectroscopy.

2.2. Protein labelling

For preparation of the ^{15}N-labelled sample, transformed cells were grown overnight at 37°C in M9 minimal medium [24] containing 5 g/l of $^{15}NH_4Cl$ and 4 g/l of glucose, diluted 1:4 into 1 l volumes of ^{15}N-M9 minimal medium and treated as the unlabelled sample. The percentage of labelling in ^{15}N checked by mass spectroscopy was 100%. In order to prepare a $^{15}N/^{13}C$ double-labelled protein transformed clones were grown overnight at 37°C in M9 minimal medium containing $^{15}NH_4Cl$ (10 g/l) and glycerol (5 g/l), diluted 4:100 into 1 l volumes of M9 minimal medium with [^{13}C]glycerol (0.3 g/l) and incubated further. When A_{600} reached 0.3, ^{13}C-hydrolysate from *Chaenopodium rubrum* cell wall was added at a final concentration of 3 g/l (Ashurst et al., manuscript in preparation) and the expression was induced as above. The protein was purified as mentioned before and found to be 100% labelled in ^{15}N and 89% in ^{13}C.

2.3. Sedimentation equilibrium

Sedimentation equilibrium experiments were performed in a Beckman Optima XL-A analytical ultracentrifuge using a Ti60 rotor and double sector 12 mm centerpieces of epon-charcoal (at the highest concentration used, some experiments were performed using 4 mm centerpieces). Samples of the protein (70 μl) in 10 mM KH_2PO_4, 100 mM NaCl, pH 6, ranging in protein concentration from 0.02 to 0.8 mM, were centrifuged at 30 000 rpm and 25°C. Radial scans at different wavelengths (from 280 to 300 nm) were taken after equilibrium was reached. To obtain the weight-average molecular weight of the repeat (M_w), the equation which defines the radial distribution of an ideal solute at sedimentation equilibrium [25] was fitted to the experimental data using the programs XLAEQ and EQWASSOC (supplied by Beckman [26]). The partial specific volume of the repeat was 0.734 ml/g, calculated from its amino acid composition [27].

2.4. NMR spectroscopy

All NMR spectra were acquired at 27 or 35°C on Bruker AMX-500/600 spectrometers. Water suppression was accomplished using WATERGATE [28]. Typically, ^{15}N-HSQC (heteronuclear single quantum coherence) [29] and ^{13}C-HSQC [30] were collected with 256 points in the indirect dimension and 1024 points in the acquisition dimension. The spectral widths were 13.5 ppm for ^1H, 30 ppm for ^{15}N and 70 ppm for ^{13}C. For the ^{3}JHN-Hα coupling constant measurement, a HMQC-J (heteronuclear multiple-quantum coherence) spectrum [31] was acquired with 512 points in the indirect dimension providing a resolution of 3.0 Hz. After transformation and linear prediction up to 1024 points an improved resolution of 1.5 Hz was achieved. Standard pulse sequences were used to monitor the relaxation process, ^{15}N-T_1, T_2, and ^{15}N-^1H heteronuclear NOE [32–34]. These experiments were all recorded with 224 points in the ^{15}N dimension, and 32 scans/increment. A series of 9 experiments with relaxation delays from 128 to 1218 ms were carried out for T_1 measurement. T_2 was measured with a series of 10 experiments with relaxation delays between 16 and 206 ms. The relaxation times T_1 and T_2 were obtained by exponential fitting of the peak intensities. Two HSQC experiments were recorded to measure the heteronuclear NOE, one with saturation of protons during the relaxation delay of 2.5 s and one without saturation. Proton saturation was accomplished by a series of 120° pulses spaced by delays of 10 ms. The NOE effect was calculated as the ratio of peak intensities in spectra collected with and without saturation.

^{15}N 3D TOWNY-HSQC [35] with 70 ms mixing time, ^{15}N 3D NOESY-HSQC [29] with 100 ms mixing time, ^{13}C resolved 3D NOESY [36] using 80 ms as a mixing time, and the triple resonance backbone experiments HNCA [37], HN(CO)CA [38], HNCO [37], and HN(CA)CO [39] were recorded for backbone assignment and identification of secondary structure elements. Commonly 64, 88, 128, 184 and 1024 complex points were collected for ^{15}N, ^{13}C, ^{13}CO, indirectly detected ^1H, and observed ^1H dimensions, respectively. NOESY mixing times were chosen after recording several 2D-spectra with mixing times ranging from 60 to 180 ms. The 3D spectra were recorded in the phase-sensitive mode using the States-TPPI [40] method and linear prediction was employed in both indirect dimensions. All spectra were processed using UXNMR for the 2D spectra and AZARA [41] for the 3D spectra. The plotting and analysis of the data was performed with AURELIA.

3. Results

3.1. Weight-average molecular weight

The results shown in Fig. 1 indicate that R16 behaves as a monomer at a millimolar concentration. Analysis of the data taking into account nonideality and hypothetical monomer-dimer equilibrium indicates that the repeat is a monomer. The molecular weight of R16 measured by sedimentation equilibrium is very close to the values obtained by mass spectroscopy (12 804 Da) or calculated from the sequence (12 803.3 Da). The anhydrous molecular weight is independent of protein

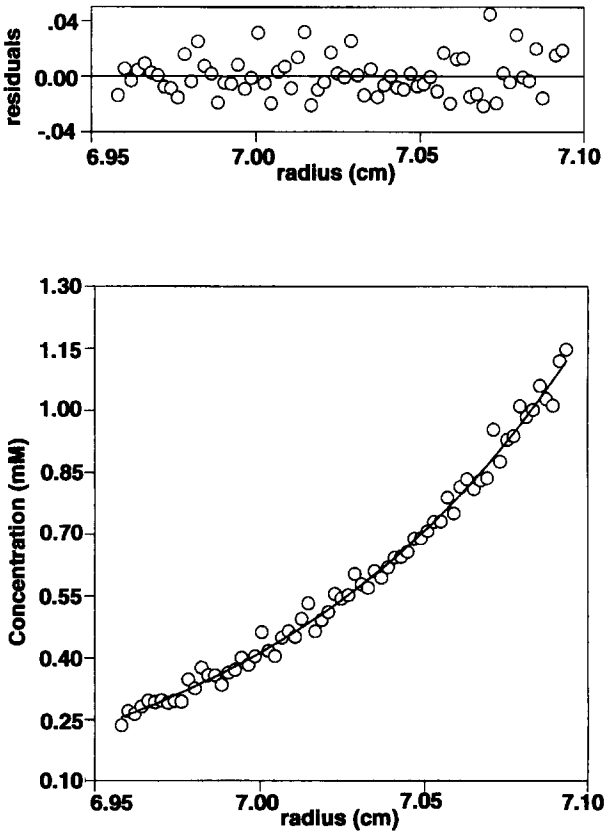

Fig. 1. Sedimentation equilibrium distribution of R16. Protein concentration is shown versus the distance from the center of rotation at sedimentation equilibrium. The sample is 0.8 mM R16 pH 6, 10 mM KH_2PO_4, 100 mM NaCl. Circles are experimental values and the curve is a calculated best fit to the radial distribution of an ideal solute at sedimentation equilibrium. The residuals show a random distribution indicative of an accurate fitting.

Table 1
Molecular weights (M_w)

Protein (mM)[a]	M_w (Da) ± 2 S.D.
0.02	13 100 ± 500
0.10	12 700 ± 300
0.25	13 200 ± 400
0.50	12 100 ± 600

[a]Loading concentration.

concentration (Table 1) indicating that R16 has no tendency to oligomerize.

3.2. Backbone and aromatic side chains assignment

NMR studies were conducted on uniformly ^{15}N or $^{15}N/^{13}C$ labelled R16. Sequential assignments of ^1H, ^{15}N, and ^{13}C resonances of the protein backbone were obtained by identifying the resonances of contiguous residues using two sets of heteronuclear 3D NMR experiments. HNCA and HN(CO)CA experiments were used to correlate the amide proton and nitrogen resonances of residue i with the $^{13}C\alpha$ resonance of residues i and $i-1$. The second set of experiments, HNCO and HN(CA)CO, take advantage of the dispersion of the carbonyl resonances to solve ambiguities caused by $^{13}C\alpha$ degeneracy. These experiments correlate the amide proton and nitrogen frequencies of residue i with the ^{13}CO frequencies of residues $i-1$ and i. The contiguous backbone resonances identified from these two sets of experiments were further checked by NOEs from the ^{15}N NOESY-HSQC and ^{13}C HSQC-NOESY. In order to obtain sequential assignment an approach combining information from both ^{15}N TOCSY-and NOESY-HSQC spectra was used. Analysis of the TOCSY gave the $^1H\alpha$ resonance of almost all HN groups and allowed the identification of the spin systems that were sequentially connected in the ^{15}N-resolved NOESY. Confirmation of the $^1H\beta$ resonances was obtained from the ^{13}C resolved NOESY. Fig. 2 presents strips at selected ^{15}N frequencies of the ^{15}N NOESY-HSQC and HNCA illustrating the sequential assignment procedure for the region of the BC loop (residues 76–80). From these experiments, the backbone assignment was obtained for all the residues except Ala-1 as shown in Table 2, where all the chemical shifts are referenced according to [42]. For Pro-61, the observation of a NOE between $^1H\alpha$ of Glu-60 and $^1H\delta$ of Pro-61 indicates that the 60–61 peptide bond is in the *trans* conformation.

A set of 2D spectra recorded with the unlabelled sample in 100% D_2O was used for assignment of the aromatic side chains and for identification of the contacts that these residues make.

3.3. Secondary structure and packing of the helices

A summary of the NMR data that define the secondary structure of R16 is shown in Fig. 3. The secondary structure elements were identified from the patterns of sequential and medium range NOEs [43,44], deviation of ^1H, $^{13}C\alpha$, and ^{13}CO chemical shifts from their random-coil values [45–47], and the value of the $^3J_{HN-H\alpha}$ coupling constant [44]. As depicted in Fig. 3, the secondary structure of the R16 consists of three α-helices (helix A from residue 9 to 32, helix B from residue 41 to 75 and helix C from residue 81 to 107) separated by two non-helical segments (from residues 33 to 40 and from residues 76 to 80). Helix B shows a distortion between positions 58 and 62 caused by the Pro-61 as demonstrated by the $^3J_{HN-H\alpha}$ values and chemical shift index for those residues. The first eight residues at the N-terminus and the last three at the C-terminus are not in a helical conformation. Interhelical NOEs, summarized in Fig. 4, allowed the identification of contacts between the three helices. The assigned NOEs can be clustered into three areas of the molecule. The first one is defined by interactions between the side chains of Phe-12 (A7, according to the nomenclature of [18]), Gly-70 (B29) and Ile-84 (C4). The second is represented by Glu-20 (A15), Val-66 (B25) and Phe-91 (C11). Trp-22 (A17), His-59 (B18) and Trp-95 (C15) define the last region.

3.4. Relaxation data

The experimental T_1, T_2, and NOE values are plotted against the amino acid sequence in panels A–C of Fig. 5, respectively. The average T_1 is 800 ms, the average T_2 is 90 ms, and the overall tumbling correlation time (τ_m) is 9.4 ns. For the majority of residues, a positive ^{15}N-^1H heteronuclear NOE is found. However, residues at both ends of the molecule show a significant decrease, in particular residues Lys-2 and Leu-3 where that NOE was negative indicative of high mobility. The relaxation data have been fitted to the simple two-parameter model-free approach [48,49] and, when necessary, to an extension of this model that takes into account the

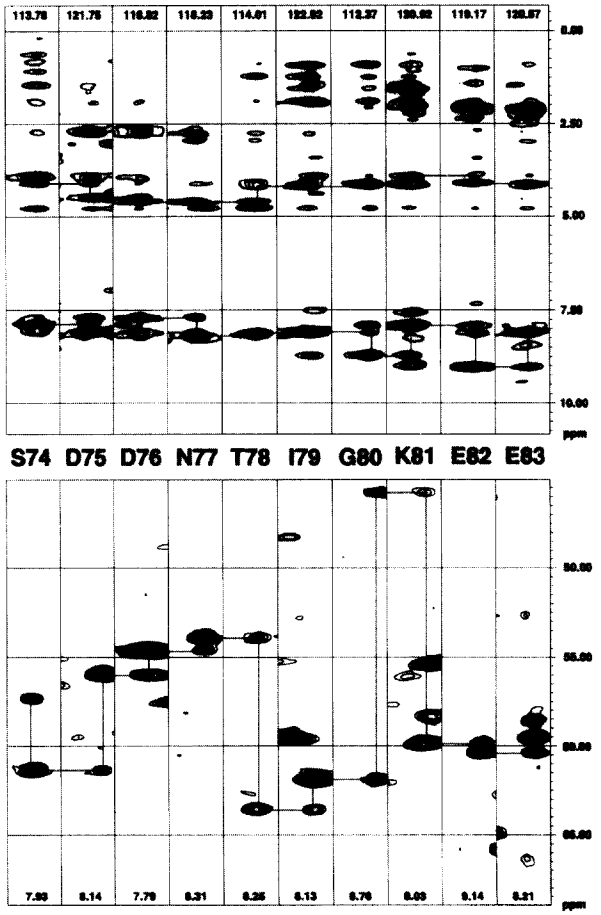

Fig. 2. ^1H-^1H and ^1H-^{13}C planes extracted from the 3D ^{15}N NOESY-HSQC (upper panel) and HNCA (lower panel), respectively. The ^{15}N resonance values are given at the top of each strip and the ^1H ones at the bottom (both in ppm). Shown are the ^1HN-^1HN and ^1HN-$^1H\alpha$ correlations for the ^{15}N edited NOESY, and the $^{13}C\alpha$-^1HN correlations for the HNCA corresponding to the region around the BC loop (from Asp-76 to Gly-80).

Table 2
Backbone chemical shifts of R16

AA	^{15}N	NH	^{13}Cα	Hα	^{13}CO	AA	^{15}N	NH	^{13}Cα	Hα	^{13}CO
1 A					177.7	56 L	119.8	8.80	59.4	4.51	179.3
2 K	130.5	8.01	56.9	4.21	176.1	57 A	120.5	7.97	55.2	4.38	176.2
3 L	124.1	8.49	55.6	4.35		58 A	119.1	8.18	54.0	4.27	178.4
4 N	119.8	8.64	53.6	4.66	175.1	59 H	116.1	7.95	56.1	4.28	174.9
5 E	121.7	8.50	56.8	4.22	176.8	60 E	120.3	7.64	54.7	4.38	
6 S	116.4	8.37	59.6	4.25		61 P	–	–	66.1	4.31	179.1
7 H	119.3	8.18	59.0	4.24	179.5	62 A	121.6	7.23	54.5	4.00	180.3
8 R	120.8	8.55	55.5	4.09		63 I	119.8	7.28	63.3	3.93	177.9
9 L	120.5	7.92	54.6	4.27		64 Q	118.0	8.36	58.2	3.79	177.8
10 H	116.6	8.13		4.38	178.7	65 G	106.0	8.15	47.4	3.96/3.96	176.4
11 Q[a]	118.9	8.33	59.6	4.15	177.0	66 V	124.0	7.49	66.9	3.87	179.0
12 F	116.8	8.03	60.3	3.94	176.3	67 L	119.9	7.90	58.8	4.04	179.7
13 F	116.8	8.03	60.3	3.94	176.4	68 D	120.6	9.07	57.6	4.39	179.2
14 R	120.3	7.76	59.4	4.06	177.9	69 T	119.4	8.39	67.0	3.82	175.8
15 D	121.3	8.45	57.7	4.29		70 G	109.2	8.77	48.3	3.40/2.77	174.4
16 M	119.6	8.35	59.9	3.36	177.3	71 K	123.0	8.08	59.6	3.92	178.0
17 D	120.5	8.62	57.9	4.38	179.4	72 K	119.3	7.52	59.4	4.04	180.1
18 D	122.2	8.48	57.9	4.55	180.0	73 L	117.6	8.01	57.4	3.98	179.7
19 E	120.9	8.15	58.6	4.15	179.3	74 S	113.8	7.93	61.4	4.12	175.2
20 E	120.6	9.48	60.5	4.36	180.3	75 D	121.7	8.14	56.0	4.47	176.6
21 S	116.9	8.50	62.3	4.25	175.8	76 D	118.8	7.79	54.7	4.60	175.9
22 W	125.8	7.93	62.7	4.14	177.9	77 N	118.2	8.31	54.0	4.61	174.6
23 I	119.2	9.02	65.8	4.11	180.8	78 T	114.0	8.25	63.6	4.21	174.7
24 K	119.3	8.22	60.1	3.89	178.5	79 I	122.9	8.13	61.9	4.20	176.0
25 E	119.1	7.89	59.5	3.92	179.3	80 G	112.4	8.78	45.7	4.15/4.15	174.9
26 K	118.2	7.82	57.0	3.84	178.0	81 K	120.9	8.03	59.9	3.91	177.3
27 K	121.2	8.90	60.1	3.64	178.9	82 E	119.2	9.14	60.4	4.11	180.0
28 L	120.3	7.83	58.0	3.94	179.1	83 E	120.6	8.21	59.6	4.18	178.9
29 L	119.5	7.45	58.0	4.10	179.4	84 I	118.7	8.08	66.2	3.92	177.2
30 V	114.8	8.11	63.9	3.65	176.1	85 Q	117.3	8.42	59.8	3.95	178.7
31 S	113.9	7.74	59.6	4.38	174.0	86 Q	119.8	8.29	59.1	4.19	176.9
32 S	116.2	7.43	58.9	4.38	174.9	87 R	118.9	8.34	58.6	4.05	179.7
33 E	124.3	8.59	57.5	4.25	179.0	88 L	120.7	8.98	58.4	4.18	177.9
34 D	119.3	8.19	54.6	4.54	175.9	89 A	120.9	7.92	55.4	4.15	180.0
35 Y	121.6	8.05	59.3	4.35	176.0	90 Q	116.5	7.57	58.4	3.98	176.8
36 G	107.8	8.22	45.5	3.88/3.88	173.6	91 F	121.6	8.24	61.0	4.55	176.9
37 R	119.6	8.11	56.4	4.35	175.2	92 V	118.1	8.56	67.0	3.54	178.1
38 D	119.8	8.27	53.7	4.62	175.9	93 D	120.5	7.72	58.0	4.49	179.1
39 L	122.9	8.47	56.6	4.27	177.9	94 H	119.3	8.52	57.9	3.74	177.6
40 T	113.0	8.33	63.7	4.25	175.5	95 W	123.4	8.60	59.4	3.56	176.4
41 G	111.3	8.33	46.0	4.22/3.96	175.1	96 K	118.2	8.87	60.4	3.58	179.2
42 V	120.9	8.29	65.0	3.91	176.9	97 E	119.3	8.43	59.3	4.01	177.5
43 Q	120.5	8.64	58.6	4.12	177.3	98 L	120.3	7.73	58.5	3.98	177.3
44 N	118.5	8.20	55.1	4.60	176.9	99 K	117.1	7.91	60.5	3.55	179.5
45 L	122.2	8.21	57.9	4.05		100 Q	119.9	8.26	59.0	4.04	178.6
46 R	120.0	8.64	59.9	4.05	178.3	101 L	122.3	8.51	57.8	4.08	178.4
47 K	119.4	7.82	59.7	4.05	179.0	102 A	121.0	8.14	54.7	4.12	
48 K	120.2	8.01	59.6	4.02	178.8	103 A	121.0	8.14	54.7	4.12	
49 H	121.3	8.64	60.2	4.39	177.8	104 A	121.0	8.14	54.7	4.12	178.9
50 K	118.9	8.33	58.6	4.15	179.5	105 R	117.9	8.02	57.7	4.11	177.8
51 R	119.8	7.75	59.1	4.18	178.2	106 G	105.8	8.01	46.6	3.77/3.32	175.3
52 L	122.8	7.72	58.1	4.29	178.6	107 Q	119.2	7.71	57.6	4.15	176.8
53 E	118.5	8.64	60.6	4.14	179.4	108 R	118.2	7.63	56.6	4.27	176.0
54 A	122.1	7.93	55.0	4.32	180.5	109 L	120.0	7.67	55.0	4.32	175.7
55 E	122.0	8.33	60.0	4.31	179.7	110 E	124.9	7.48	58.4	4.01	

All chemical shifts are referenced according to [42], using TSP as internal reference at 0.0 ppm.
[a]Tentative assignments.

presence of fast and intermediate internal motions [50]. This extension of the model includes two rapid motions. A very fast motion on a picosecond time scale characterized by τ_e, and an intermediate motion on a nanosecond time scale characterized by τ_i. The order parameters, S^2 and S_2^f, derived from analysis of the relaxation data are plotted against amino acid sequence in panels E and F of Fig. 5, respectively. The average order parameter (S^2) is 0.8 indicative of a compact structure. All residues within the helices have an order parameter (S^2) of 0.72 or greater. However, both loop regions have the smallest values of S^2 found in the protein with the exception of the N-terminus, 0.52 for the AB loop and 0.55 for the BC loop on average. In panels G and H of Fig. 5, the correlation time for fast (τ_e) and intermediate (τ_i) internal motions is represented versus sequence. Interestingly, both loops show motions in the nanosecond time scale. Concretely, the AB and BC loops give τ_i values in the range of 2–6 ns and 2–3 ns and S_2^f values of 0.17–0.39 and 0.18–0.62, respectively. The exchange contribution (R_{ex}) to the ^{15}N line width (Fig. 5D) is always less than 3 Hz.

4. Discussion

Sedimentation equilibrium experiments indicate that R16 does not self-associate even at a millimolar protein concentration (Table 1). Complete backbone assignment allowed the determination of the secondary structure of this monomeric spectrin repeat using a combination of short and medium range NOEs, the chemical shift index and *J*-coupling data (Fig. 3). α-Helices have been characterized by strong HN-HN($i,i+1$) NOEs, weak Hα-HN($i,i+1$) NOEs and specifically by the presence of Hα-HN($i,i+3$) and Hα-Hβ($i,i+3$) NOEs, a consensus chemical shift index of −1 and HN-Hα coupling constant <6.0 Hz. Using these criteria three helices (A–C) were defined separated by two non-helical segments (AB and BC). Additionally, unambiguously assigned side chain NOEs indicate a tight packing between the three helices (Fig. 4). In the alignment of the spectrin repeats (Pascual et al., manuscript in preparation), most of these NOEs belong to conserved residues that occupy the positions a or d within the heptad pattern which are commonly populated by hydrophobic amino acids forming the core of the protein.

In general, secondary structure elements are characterized by high order parameters (S^2) which are indicative of low mobility, while low S^2 values, indicative of high mobility, are only found at the termini and in loops [51,52]. The relaxation data depicted in Fig. 5 corroborate the secondary structure previously identified and show different relaxation properties between the helical and non-helical residues. Significantly, there is a clear indication that both non-helical regions are flexible in solution supporting the proposal that the regions connecting the helices are loops. The differences in the values of the S^2 between the helical and non-helical regions indicate that these non-helical residues undergo motions of significant amplitude on a time scale faster than the overall tumbling correlation time (τ_m). The intermediate time scale motions defined by S_f^2 in the extended model are found at the N-and C-termini and in both loops. In fact, nearly all of the residues whose relaxation data were analyzed taking into account intermediate internal motions, are located either in the loop regions or at the ends of the protein. Moreover, for the loop regions HN-HN($i,i+1$) NOEs have been observed corroborating the intermediate time scale of the motions on those regions.

Comparison of the secondary structure of the 16th chicken repeat as determined above with that of the crystal dimer of the 14th *Drosophila* repeat obtained using the dssp programme [53] shows several differences. The most important is the absence in the crystal structure of the BC loop. The equivalent residues in the crystal structure (from Gln-71 to Ala-75) are in a helical conformation ($\phi = -64.24$ and $\psi = -43.86$ on average). Regarding the tertiary structure, the interhelical NOEs indicate that the helix packing is similar

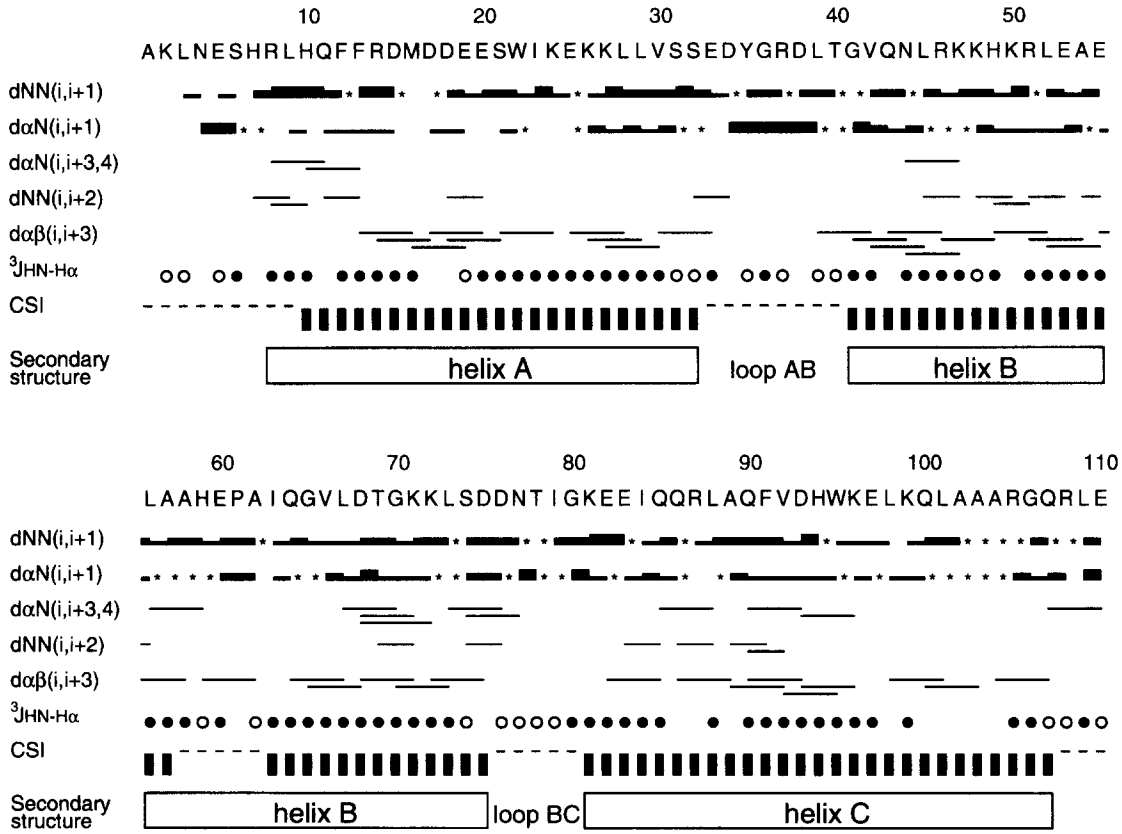

Fig. 3. Summary of the NMR data used to identify the secondary structure of R16. The amino acid sequence and residue number are given at the top. Sequential NOE connectivities are indicated by bars, where the thickness represents the relative NOE intensities. Asterisks indicate NOEs that were not calibrated due to overlap. For Pro-61 the ^{1}Hδ has been used as ^{1}HN. Medium-range NOEs are indicated by lines covering the length of the connected residues. In row $^3J_{HN-H\alpha}$, the open circles indicate residues with a coupling constant greater than 6.0 Hz, whereas the filled circles correspond to those with smaller coupling constants. The consensus chemical shift index (CSI) shown in the figure has been calculated combining chemical shifts from ^{1}Hα, ^{13}Cα, and ^{13}CO. Thin rectangles represent consensus chemical shift values equal to 0 (coil), and thick rectangles represent those values equal to −1 (helix). The sequence locations of the three helices identified from the data shown in this figure are indicated at the bottom.

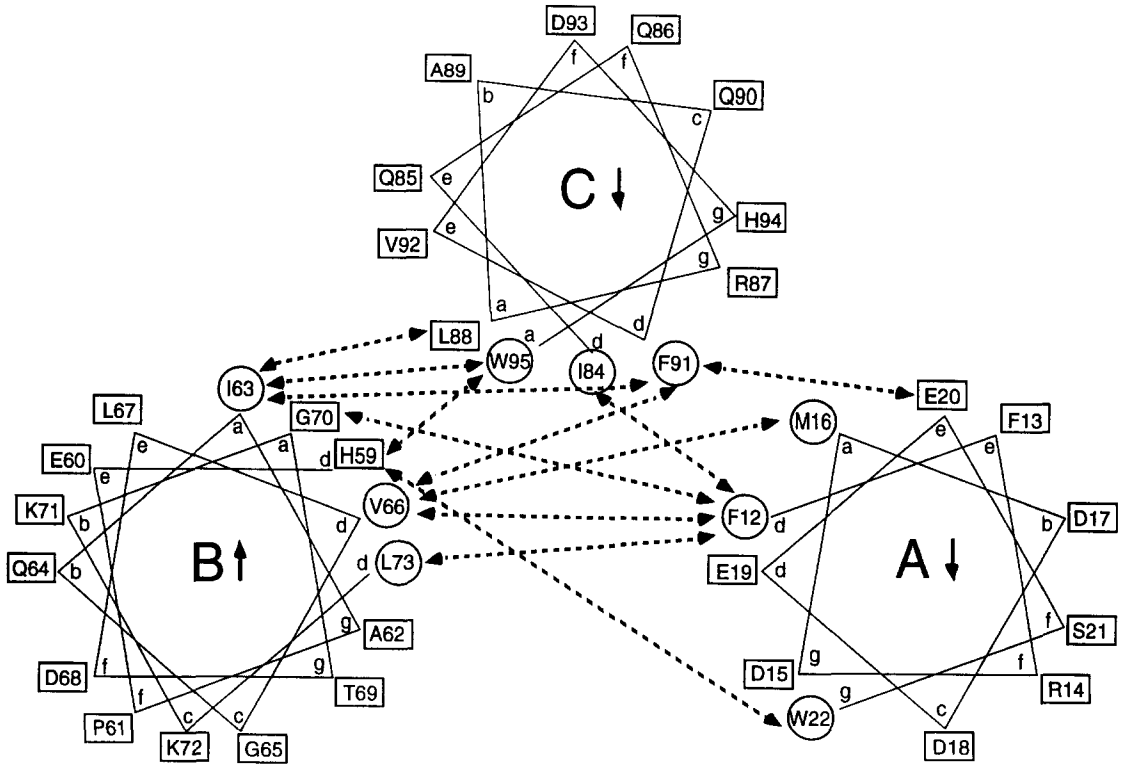

Fig. 4. Helical wheel representation of the hydrophobic core of R16. Interhelical NOEs are indicated by arrows, and the position in the heptad pattern for each residue is depicted. Residues surrounded by a circle indicate conserved hydrophobic positions in the spectrin repeat alignment. The arrow inside each wheel points towards the C-terminus of the helix, above or below the plane of the drawing.

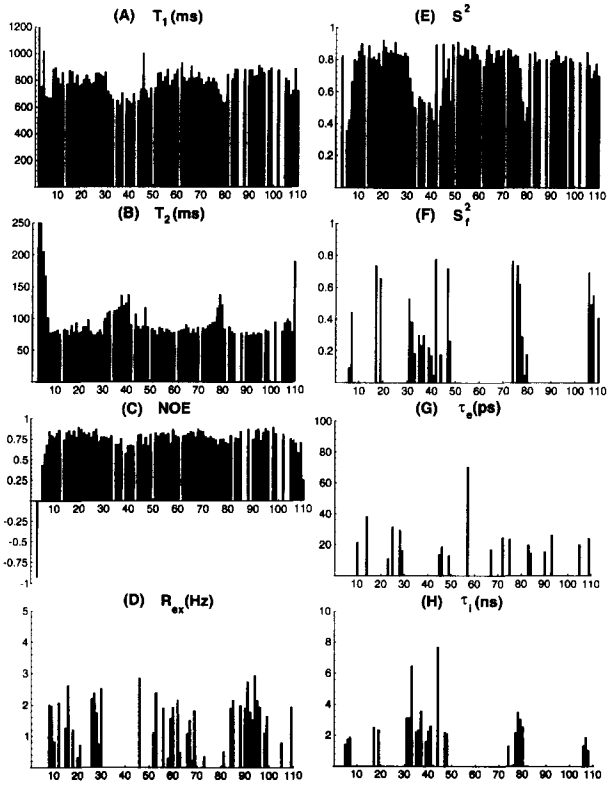

Fig. 5. Summary of the relaxation data for R16, residue by residue. (A) ^{15}N T_1 values, (B) ^{15}N T_2 values and (C) ^{15}N-^1H heteronuclear NOE values. (E,F) Plot of the calculated values for S^2 and S_f^2, respectively, versus sequence. (G,H) Plot of the calculated correlation time for fast (τ_e) and intermediate (τ_i) motions, respectively. (D) Plot showing the exchange contribution (R_{ex}).

to that observed in the crystal structure of the dimer, although the ongoing structure determination still has to distinguish properly between the tertiary structure of the two different models, either a left-handed coiled-coil [16] or a left-handed three-helix bundle [18].

The elucidation of the three-dimensional structure of the spectrin repeat is a step towards a fuller understanding of the structure and dynamics of the membrane skeleton. The elastic properties of the spectrin repeat as well as pairs of repeats need to be further investigated in the light of an accurate structure of the repeat. Molecular explanations for mutations such as those causing human elliptocytosis related to defective tetramerization of spectrin in red cells [9] may also derive from the refined structure of the repeat.

Acknowledgements: We thank Dr. A. Musacchio for assisting in protein purification, and Dr. M. Kelly for the protocols for protein labelling. We also thank Dr. F. Blanco for critically reading the manuscript and very helpful discussions. G.R. was supported by grant PB92007 from the Spanish DGICYT.

References

[1] Elgsaeter, A., Stokke, B.T., Mikkelsen, A. and Branton, D. (1986) Science 234, 1217–1222.
[2] Bennett, V. and Gilligan, D.M. (1993) Annu. Rev. Cell Biol. 9, 27–66.
[3] Sheterline, P., Clayton, J. and Sparrow, J. (1995) Protein Profile 2, 1–68.
[4] Hartwig, J.H. (1994) Protein Profile 1, 711–762.
[5] Castresana, J. and Saraste, M. (1995) FEBS Lett. 374, 149–151.
[6] Trave, G., Pastore, A., Hyvonen, M. and Saraste, M. (1995) Eur. J. Biochem. 227, 35–42.
[7] Trave, G., Lacombe, P.J., Pfuhl, M., Saraste, M. and Pastore, A. (1995) EMBO J. 14, 4922–4931.

[8] Viel, A. and Branton, D. (1994) Proc. Natl. Acad. Sci. USA 91, 10839–10843.
[9] Tse, W.T., Lecomte, M.C., Costa, F.F., Garbarz, M., Feo, C., Boivin, P., Dhermy, D. and Forget, B.G. (1990) J. Clin. Invest. 86, 909–916.
[10] Musacchio, A., Noble, M., Pauptit, R., Wirenga, R. and Saraste, M. (1992) Nature 359, 851–854.
[11] Macias, M.J., Musacchio, A., Ponstingl, H., Nilges, M., Saraste, M. and Oschkinat, H. (1994) Nature 369, 675–677.
[12] Speicher, D.W. and Marchesi, V.T. (1984) Nature 311, 177–180.
[13] Winograd, E., Hume, D. and Branton, D. (1991) Proc. Natl. Acad. Sci. USA 88, 10788–10791.
[14] Lusitani, D.M., Qtaishat, N., LaBrake, C.C., Yu, R.N., Davis, J., Kelley, M.R. and Fung, L.W. (1994) J. Biol. Chem. 269, 25955–25958.
[15] MacDonald, R.I., Musacchio, A., Holmgren, R.A. and Saraste, M. (1994) Proc. Natl. Acad. Sci. USA 91, 1299–1303.
[16] Parry, D.A., Dixon, T.W. and Cohen, C. (1992) Biophys. J. 61, 858–867.
[17] Lupas, A., Van Dyke, M. and Stock, J. (1991) Science 252, 1162–1164.
[18] Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S.C. and Branton, D. (1993) Science 262, 2027–2030.
[19] Wasenius, V.M., Saraste, M., Salven, P., Eramaa, M., Holm, L. and Lehto, V.P. (1989) J. Cell. Biol. 108, 79–93.
[20] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1992) Current Protocols in Molecular Biology, Wiley, New York.
[21] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Methods Enzymol. 185, 60–89.
[22] Laemmli, U. (1970) Nature 227, 680–685.
[23] Gill, S.C. and Hippel, P.H. (1989) Anal. Biochem. 182, 319–326.
[24] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Plainview, NY.
[25] Ralston, G. (1993) Introduction to Analytical Ultracentrifugation, Beckman Instruments Inc., Palo Alto, CA.
[26] Minton, A.P. (1994) Modern Analytical Ultracentrifugation, pp. 81–93, Birkhauser, Boston.
[27] Laue, T.M., Shah, B.D., Ridgeway, T.M. and Pelletier, S.L. (1992) Analytical Ultracentrifugation in Biochemistry and Polymer Science, pp. 90–125, Royal Society of Chemistry, Cambridge.
[28] Piotto, M., Saudek, V. and Sklenar, V. (1992) J. Biomol. NMR 2, 661–665.
[29] Sklenar, V., Piotto, M., Leppik, R. and Saudek, V. (1993) J. Magn. Reson. A 102, 241–245.
[30] Stonehouse, J., Shaw, G.L., Keeler, J. and Laue, E.D. (1994) J. Magn. Reson. A 107, 178–184.
[31] Kay, L.E. and Bax, A. (1990) J. Magn. Reson. 86, 110–126.
[32] Barbato, G., Ikura, M., Kay, L.E., Pastor, R.W. and Bax, A. (1992) Biochemistry 31, 5269–5278.
[33] Kay, L.E., Torchia, D.A. and Bax, A. (1989) Biochemistry 28, 8972–8979.
[34] Kay, L.E., Nicholson, L.K., Delagio, F., Bax, A. and Torchia, D.A. (1992) J. Magn. Reson. 97, 359–367.
[35] Kadkhodaei, M., Hwang, T.L., Tang, J. and Shaka, A.J. (1993) J. Magn. Reson. A 104, 105–107.
[36] Majumdar, A. and Zuiderweg, E.P. (1993) J. Magn. Reson. B 102, 242–244.
[37] Kay, L.E., Ikura, M., Tschudin, R. and Bax. A. (1990) J. Magn. Reson. 89, 496–514.
[38] Grzesiek, S. and Bax, A. (1992) J. Magn. Reson. 96, 432–440.
[39] Clubb, R.T., Thanabal, V. and Wagner, G. (1992) J. Magn. Reson. 97, 213–217.
[40] Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) J. Magn. Reson. 85, 393–397.
[41] Boucher, W. (1994) AZARA v1.0, Department of Biochemistry, University of Cambridge, UK.
[42] Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, J., Oldfield, E., Markley, J. and Sykes, B.D. (1995) J. Biomol. NMR 6, 135–140.
[43] Wüthrich, K., Billeter, M. and Braun, W. (1984) J. Mol. Biol. 180, 715–740.
[44] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
[45] Pastore, A. and Saudek, V. (1990) J. Magn. Reson. 90, 165–176.
[46] Wishart, D.S., Sykes, B.D. and Richards, F.M. (1992) Biochemistry 31, 1647–1651.
[47] Wishart, D.S. and Sykes, B.D. (1994) J. Biomol. NMR 4, 171–180.
[48] Lipari, G. and Szabo, A. (1982) J. Am. Chem. Soc. 104, 4546–4559.
[49] Lipari, G. and Szabo, A. (1982) J. Am. Chem. Soc. 104, 4559–4570.
[50] Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C. and Gronenborn, A.M. (1990) J. Am. Chem. Soc. 112, 4989–4991.
[51] Redfield, C., Boyd, J., Smith, L.J., Smith, R.A.G. and Dobson, C.M. (1992) Biochemistry 31, 10431–10437.
[52] Mandel, A.M., Akke, M. and Palmer III, A.G. (1995) J. Mol. Biol. 246, 144–163.
[53] Kabsch, W. and Sander, C. (1983) Biopolymers 22, 2577–2637.