Journal of Peptide Science
J. Peptide Sci. **10**: 381–392 (2004)
Published online 26 January 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.540

JOURNAL OF PEPTIDE SCIENCE

Conformational Analysis of Peptide Analogues of Silkmoth Chorion Protein Segments Using CD, NMR and Molecular Modelling

DIMITRA C. BENAKI,[a] EMMANUEL MIKROS[a]* and STAVROS J. HAMODRAKAS[b]

[a] Department of Pharmaceutical Chemistry, School of Pharmacy, University of Athens, GR 157 71, Athens, Greece
[b] Division of Cell Biology and Biophysics, Department of Biology, University of Athens, GR 157 71, Athens, Greece

Received 28 July 2003
Accepted 12 September 2003

Abstract: Silkmoth proteins secreted from the follicular cells that surround the oocyte form a large extracellular assembly which is important for protecting and sustaining the structure of the oocyte and the developing embryo. These proteins have been classified into two major families (A and B). Sequence analysis showed conservation of a central domain containing long stretches of six amino acid residue repeats in both families, which have been suggested to be organized in β-sheet structures. In this work NMR and CD spectra, as well as molecular calculations, have been used to investigate the conformational properties of two synthetic peptides (A and B), analogues of parts of the central domain of silkmoth chorion proteins of the A and B families, respectively. These peptides consist of three tandem repeats of the six-residue basic motif. Analysis of CD spectra of the two peptides in aqueous solutions and mixtures with organic solvents revealed β-sheet and turn structural elements with a percentage higher than 40%. NOESY spectra at low temperatures (263–273 K) show sequential nOe connectivities (i, i + 1), indicative of a relative flexibility. The presence of H_{N_i}-$H_{N_{i+1}}$ cross-peaks and medium $H\alpha_i$-$H_{N_{i+1}}$ connectivities, chemical shift deviations and temperature coefficient data provide, for the first time, experimental evidence that local folded structures around Gly residues occur in peptide segments of chorion proteins in solution. Simulated annealing calculations were used to examine the conformational space of the peptides and to probe the initial steps of amyloid fibril formation in the case of chorion proteins.

Keywords: 3D-structure; CD; NMR; molecular modelling; peptide; silkmoth chorion proteins

INTRODUCTION

Chorion, the major part of insect eggshell, is a complex extracellular proteinaceous structure surrounding the oocyte, which performs important physiological functions: it allows sperm entry–fertilization and exchange of the respiratory gases, and provides for thermal and mechanical insulation, waterproofing, exclusion of microorganisms and hatching (for reviews see [1–3]).

Silkmoth chorion is largely proteinaceous (over 95% of its dry weight) and is the part of the eggshell that exhibits a helicoidal architecture. About 200 proteins have been detected [4] and have been classified into two major classes, A and B [1]. These are products of the chorion gene superfamily, which has two branches: the α-branch

*Correspondence to: Dr Emmanuel Mikros, Department of Pharmaceutical Chemistry, School of Pharmacy, University of Athens, GR 157 71, Athens, Greece; e-mail: mikros@pharm.uoa.gr
Contract/grant sponsor: EMBO.
Contract/grant sponsor: SRMB.

and the β-branch [5]. The conservation and length invariance of the central domain of chorion proteins of the two branches [6] suggest that it adopts a precise, functionally important, three-dimensional entity [3].

A structural model has been proposed for chorion proteins combining data from amino acid sequence comparisons, secondary structure prediction, analysis of amino acid periodicities and modelling [7,8]. According to this model an antiparallel β-pleated sheet of consecutive two-residue turns, alternating with short four-residue β-strands is formed. Three repetitions of this hexapeptide periodicity (Figure 1) provide an adequate length for a study of the conformational character of the consecutive hexapeptide repeats as a minimal structural unit and to investigate its functional role in the chorion protein folding process.

Thus, our interest was focused on synthesizing [9] and studying two peptides, parts of the A and B class central domains (Figure 1). The amino acid sequences of these peptides are as follows: A peptide, GELPVAGKTAVAGRVPII, and B peptide, GNLPFLGTAGVAGEFPTA. Peptide sequences were selected on the basis of the following criteria:

(a) The A and B peptides are analogues of the segments of the central domain of proteins belonging to the α- and β-branches of the silkmoth chorion protein superfamily, respectively. (b) The hexapeptide periodicity is repeated three times. (c) A and B peptide sequences contain Pro residues at positions 4 and 16, which are highly conserved in the vast majority of silkmoth chorion protein primary structure.

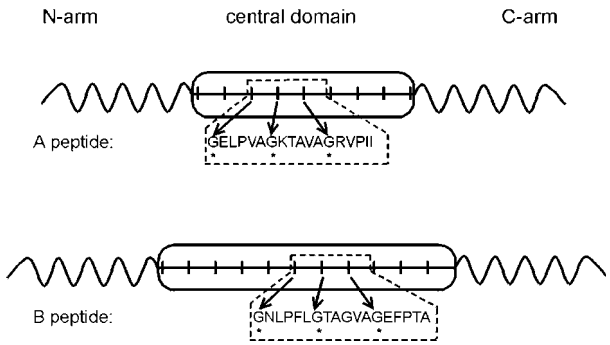

Figure 1 Schematic representation of proteins belonging to the A (a) and B (b) classes of chorion proteins showing the characteristic tripartite structure: two flanking arms (*N*- and *C*-terminal) and the central domain (box) conserved in both sequence and length. The separated areas of the central domain represent the hexapeptide periodicity and the perpendicular lines the position of the Gly residues conserved among protein sequences of the same class. A and B peptide sequences are provided (one letter code) along with their relative position in the central domain. The conserved Gly residues in the peptide sequences are marked with an asterisk.

Studies of these peptides with laser-Raman and IR absorption spectroscopies suggested a preponderance of the antiparallel β-pleated sheet structure, both in solution and in the solid state [9]. Recently, a β-sheet helix model (composed of a basic structural unit of three parallel β-strands folded into a coil [10]) has been proposed as an alternative plausible structure for silkmoth chorion proteins. The proposed structure was based on x-ray diffraction, spectroscopic data (FT-Raman and ATR-FTIR) and theoretical studies for the B peptide [11] and for the 51-residue cA peptide which comprises the entire central domain of the A family of silkmoth chorion proteins. It was also shown that these peptides fold and self-assemble forming amyloid-like fibrils under a variety of conditions [11,12], suggesting that silkmoth chorion is a natural, protective amyloid. Understanding the mechanism of amyloid formation is of great importance since the deposition of insoluble protein fibrils is associated with numerous neurodegenerative diseases such as Alzheimer's, Parkinson's and spongiform encephalopathies and other types of dementia [13,14].

This work presents structural studies based on CD and NMR techniques, together with restrained simulated annealing calculations, of the two 18-residue synthetic peptides A and B to investigate further the propensity of these molecules to form functional structures, aimed at visualizing their conformational preferences and at investigating the initial steps of the chorion protein self-assembly process.

MATERIALS AND METHODS

Peptide Synthesis

Peptides were synthesized by the solid-phase method using Boc chemistry [9].

Circular Dichroism Spectroscopy

CD spectra were recorded on a Jasco J-710 instrument, calibrated using the ammonium salt of 10-(+)-camphorsulfonic acid. Quartz cells of various pathlengths were employed and maintained at 278 K

using a Neslab RTE-100 water bath. Spectra were recorded in the range 250–190 nm as an average of 10 scans, with a bandwidth of 2 nm, 0.2 nm step size and a time constant of 0.2 s. The spectra were normalized for concentration and pathlength to obtain the mean residue ellipticity after subtraction of the buffer contribution. Where it was necessary, the spectra were smoothed using a third-order polynomial function.

CONTIN [15], Lincomb [16] and Chang and coworkers' [17] program analyses were used for estimation of secondary structure percentages.

Nuclear Magnetic Resonance

^{1}H NMR experiments were performed on a 400 MHz Bruker DRX-Avance spectrometer, with the probe temperature maintained using a BVT-3000 Bruker control unit.

The A peptide was studied in three different solutions: (a) in H_2O/D_2O (9 : 1) with 10 mM sample concentration at 280 and 293 K, (b) in CD_3OD/H_2O (2 : 1), and (c) in DMSO/H_2O (1 : 1) solution. The peptide concentration in organic solvents ranged from 2 to 4 mM and the studies were performed at temperatures in the range 263–300 K. As the B peptide exhibited lower solubility and formed a gel in water and methanol solutions, its NMR study was carried out at 2 mM in DMSO/H_2O (1 : 1) at temperatures of 273–300 K.

1D spectra were acquired using 32K data points and zero-filled to 64K data points before Fourier transformation. All 2D spectra were recorded in the phase-sensitive mode using time proportional phase incrementation (TPPI) [18]. NOESY spectra [19] were recorded with mixing times of 80, 100, 120, 150, 200, 250, 300, 400 and 600 ms. In the ROESY [20] experiments the spin lock time was 100, 200 and 250 ms at 268 K. TOCSY spectra [21] were recorded using the MLEV-17 spin lock sequence with mixing times of 70–80 ms. Typically, the sweep width was 12 ppm and the spectra were collected into 2048 points in the t_2 dimension (32 transients were co-added for each of 512 t_1 points). The double quantum filtered correlation experiment (DQF-COSY) [22] was recorded into 8192 points in t_2 and 2048 increments in t_1, with 64 scans per t_1 experiment. The relaxation delay in all the experiments was fixed to 2 s. Solvent suppression was achieved, either by presaturation during the relaxation and mixing period (NOESY and ROESY experiments), or by the WATERGATE gradient module [23,24].

2D spectra were Fourier transformed after applying phase-shifted squared sine-bell functions in both dimensions (the optimal phase shift was 60°). t_1 data were zero filled to 1024 points (4096 points for the DQF-COSY experiment). The polynomial baseline correction was applied in selected regions of the spectra. The chemical shifts were referenced to internal DSS at 0.0 ppm. Data were processed using the standard XWIN-NMR Bruker program on a SGI Indy. The XWIN-PLOT 1.04 program, provided in the XWIN-NMR package, was used for the presentation of selected NMR spectral regions. For the determination of the amide proton temperature coefficients, the chemical shifts of the H_N frequencies were obtained from TOCSY spectra acquired at 263, 268, 272, 276, 280, 284, 288, 295 and 300 K.

Extraction of distance constraints from 2D NOESY spectra. Interproton distances were calculated from nOe cross-peak intensities using the program MARDIGRAS [25] which incorporates all the effects of network relaxation and multiple spin effects. MARDIGRAS determines a set of distances utilizing an initial guess of molecular structure; it recalculates iteratively the relaxation rate matrix R starting with the hybrid intensity matrix and iteratively inserting experimental intensities. The process is repeated until the error between the calculated and the observed intensities reaches a minimum value. From the final relaxation matrix a set of distances corresponding to measured cross peaks is obtained. An RMS factor for the comparison between the initial model structure interproton distances and the distances proposed by MARDIGRAS is also provided. An energy minimized extended-chain structure was used as initial model generated by 'the best geometry extended structures protocol' implemented in the Crystallographic & NMR System (CNS) software [26]. An isotropic correlation time of 2 ns for all protons was used [27,28].

Simulated Annealing Calculations

Restrained simulated annealing calculations were performed using the protocol implemented in the CNS 1.0 software [26]. The energy minimized extended structure used in MARDIGRAS was further subjected to 2000 steps of energy minimization. In the first stage of the simulated annealing protocol, the system was subjected to 15 ps (1000 steps) of torsion-angle molecular dynamics [29] at 50 000 K. The scale factor used for the nOe energy term during the high temperature stage was 150, the van

der Waals energy term 0.1, to facilitate rotational barrier crossings, and the dihedral angle term 100. The system was then subjected to a slow-cooling torsion-angle molecular dynamics stage in which the temperature was reduced from 50 000 to 2000 K (temperature step: 250 K) over a period of 15 ps, while the van der Waals energy term was linearly increased from 0.1 to 1.0. The scale factor of the dihedral angle energy term was 200. The third stage of the protocol consisted of a slow-cooling stage from 2000 to 300 K for 15 ps of Cartesian molecular dynamics. During this stage of calculation the van der Waals energy term was linearly increased from 1.0 to 4.0. Finally, the structure was subjected to 10 cycles of 200 step of restraint Powell energy minimization, with a scale factor for the nOe term of 75 and for the dihedral angle of 400.

The estimated by MARDIGRAS distance restraints, (22 sequential, $H\alpha_i$-HN_{i+1}, $H\beta_i$-HN_{i+1}, HN_i-HN_{i+1} for the A and 38 for the B peptide, 15 intraresidue for the A and 25 for the B peptide) were introduced to the CNS protocol and a group of 30 possible conformations was generated. Several of these structures were used as starting models for a second cycle through MARDIGRAS. The structures that gave a minimum RMSD (~0.6) were selected as starting structures for a new simulated annealing calculation cycle. This procedure was followed until 200 structures were generated for each peptide.

The XCLUSTER 1.2 program [30], as implemented in the MACROMODEL 6.5 software [31], was used for the superposition of the generated structures in order to determine families of conformers. XCLUSTER inputs the series of conformations and computes the RMS difference between all possible pairs of conformations. Structures of the input sequence are then reordered on the basis of increasing RMS deviation. Structures were displayed using Maestro 4.1 Schrödinger Inc. Software.

RESULTS AND DISCUSSION

Circular Dichroism

A peptide. The CD spectra of the A peptide in aqueous solutions at pH 4–9 (Figure 2a) were characterized by negative ellipticity values, with a negative maximum at 198 nm. Analysis of the CD spectra provided an insight of the structural elements present in aqueous solutions (α-helix 0%, β-sheet 20%–35%, turns 25% and remainder

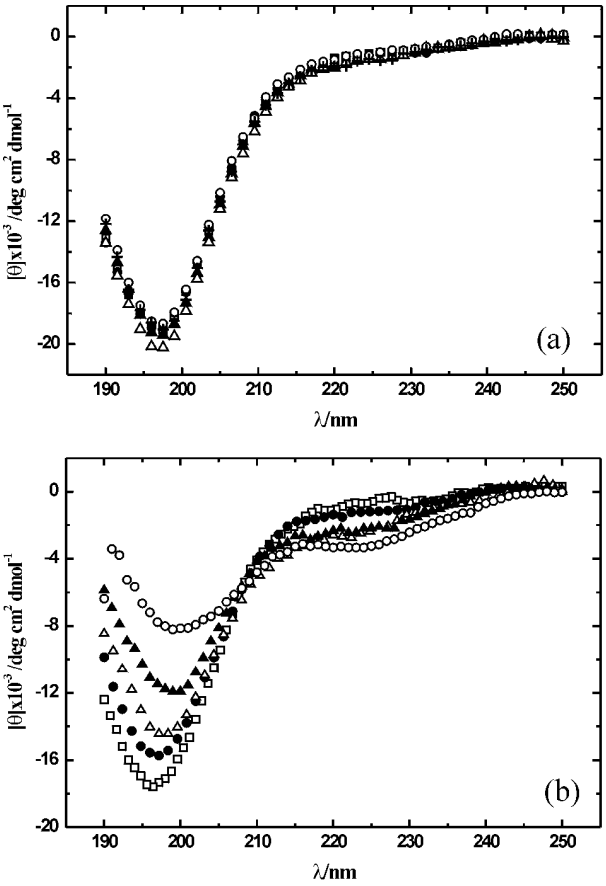

Figure 2 CD spectra of A peptide at 278 K in aqueous solutions at different pH (a): sodium acetate 50 mM, pH 4 (□) and pH 5 (●), MOPS 20 mM, pH 6 (▲) and pH 7 (Δ) and CAPS 20 mM, pH 8 (○) and pH 9 (♦) and in solutions with different concentration of methanol (b): 0% (□), 30% (●), 50% (Δ), 60% (▲) and 80% (○). Sample concentration was 30 μM.

structures 55%–60%). Under strong denaturing conditions (6 M guanidinium chloride) a decrease of the absolute value of the ellipticity (in the region 213–250 nm) was detected, suggesting a shift of the conformational equilibrium in water to less structured states.

The CD spectra of A peptide in different TFE solutions showed that the helical secondary structure component increased upon increasing the TFE percentage (α-helix 15% at 50% TFE and α-helix 27% at 100% TFE solution). It should be noted that peptides with pronounced helical propensity often reach the maximum helical content at lower TFE concentrations (30%–50%) [32]. Consequently, the A peptide should not have a marked tendency to adopt a helical structure.

In Figure 2b the CD spectra of the A peptide solutions are shown in different methanol solutions (0, 30%, 50%, 60% and 80% CH_3OH). As TFE, methanol promotes the structural organization of the specific peptide sequence: by increasing the methanol content a gradual shift of the negative maximum to higher wavelengths was detected, consistent with an increase in regular structure elements. The α-helical content estimated by the analysis programs [15–17] varied from 0 to 5% for 0–60% methanol solutions (at 80% methanol solution it exhibited a higher percentage: 27% α-helix).

The β-sheet percentage varied from 15% to 30% for 0–60% methanol solutions, and increased to 60% for the higher percentage methanol solutions.

B peptide. The B peptide exhibited similar behaviour to the A peptide in aqueous solutions of different pH values, with a negative maximum found at 198 nm. Secondary structure analysis of the CD spectra revealed 25%–35% turns and 65%–75% of the remainder structures. The Jasco program (SSEAX Jasco Corp) estimated a β-sheet percentage of the order of 20%. Addition of a strong denaturing agent (6 M guanidinium chloride) resulted in a pronounced shift of the negative maximum to lower values, a trend similar to that of the A peptide.

Nuclear Magnetic Resonance

A peptide. The assignment of proton resonances was achieved through the combined analysis of 2D TOCSY experiments, in the range 263–300 K, and 2D NOESY experiments with different mixing times (80–400 ms), using the standard protocol for sequential assignment [33]. The NMR spectra of the A peptide in aqueous solution (H_2O/D_2O, 9 : 1), although fully assigned, were characterized by small chemical shift dispersion and peak overlap (data not shown).

The A peptide was further studied in CD_3OD/H_2O and $DMSO/H_2O$ solutions providing an environment that helps the structural organization of the peptide. Additionally, these solutions allowed us to work at subzero temperatures and to 'freeze' the most stable conformations present in solution.

In both solutions the A peptide exhibited similar chemical shifts and nOe patterns (strong $H\alpha_i$-$H_{N_{i+1}}$ connectivities and H_{N_i}-$H_{N_{i+1}}$ cross-peaks). However, the chemical shift dispersion in methanol solution was relatively higher with less nOe overlapping. Thus, in the rest of the present work the CD_3OD/H_2O (2 : 1) solution is referred to.

Chemical shift values of identified protons of A peptide at 273 K are listed in Table 1. The fingerprint region and the assignment of a NOESY spectrum (mixing time 200 ms), is shown in Figure 3a.

Table 1 Chemical Shifts of A Peptide in CD_3OD/H_2O (2 : 1) Solution at 273 K

Residue	HN	Hα	Hβ	Other protons	Temperature coefficient
Gly 1		3.67			
Glu 2	8.66	4.28	1.94/1.76	γ2.17	4.0
Leu 3	8.54	4.51	1.53/1.46	γ1.45 δ 0.85/0.81	7.1
Pro 4	—	4.38	2.1/	γ1.92/1.83 δ 3.75/3.52	
Val 5	8.17	3.95	1.93	γ0.85	8.6
Ala 6	8.52	4.12	1.28		9.0
Gly 7	8.45	3.85/3.74			7.0
Lys 8	8.11	4.30	1.77/1.67	γ1.36 δ1.57ε2.84	3.7
Thr 9	8.00	4.19	4.09	γ1.01	5.8
Ala 10	8.27	4.26	1.27		6.2
Val 11	8.02	3.95	1.95	γ0.83	7.6
Ala 12	8.35	4.13	1.29		7.6
Gly 13	8.37	3.82/3.75			7.5
Arg 14	7.99	4.35	1.74/1.64	γ1.49 δ 3.07ε7.15/6.65	3.3
Val 15	8.13	4.30	1.99	γ0.87	7.5
Pro 16	—	4.36	2.14/2.11	γ1.95/1.85 δ 3.78/3.58	
Ile 17	8.16	4.08	1.75	γ1.44/1.08 δ 0.83	8.6
Ile 18	7.50	4.03	1.70	γ1.33 δ 0.77	4.4

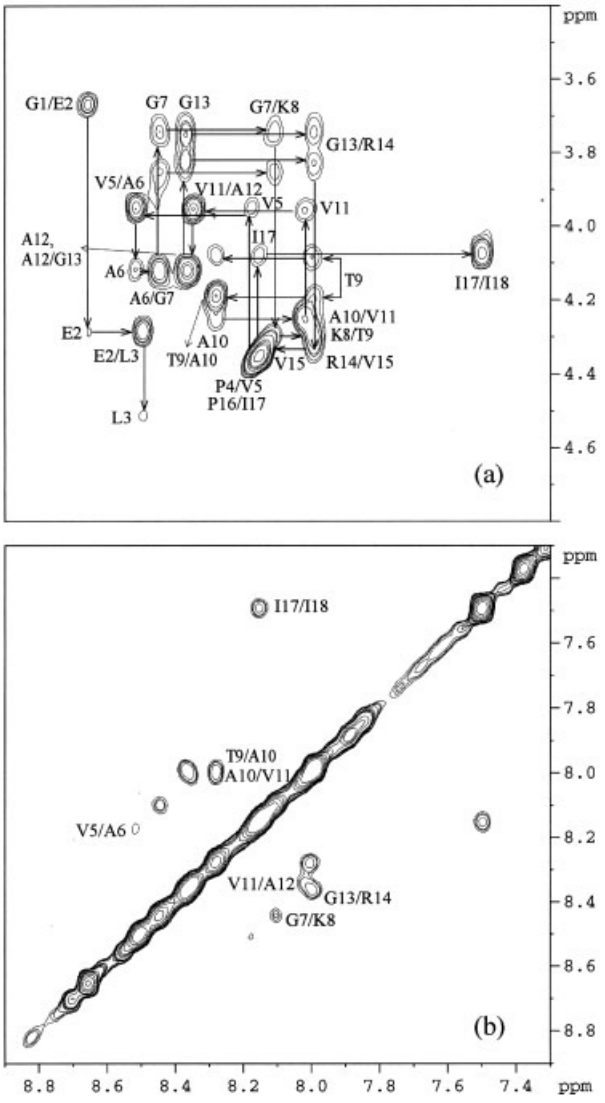

Figure 3 Fingerprint (a) and amide-amide proton (b) regions (Hα-HN and HN-HN) of NOESY spectrum of the A peptide recorded at 273 K in CD_3OD/H_2O (2 : 1) solution. The mixing time was 200 ms.

The nOes detected concern protons of successive residues in the peptide sequence (i, i + 1). For the whole temperature range studied (263–300 K) the nOe cross-peaks were negative.

Sequential $H\alpha_i$-HN_{i+1} nOes were strong along the peptide chain and intraresidue $H\alpha_i$-HN_i nOe connectivities were of medium intensity. The $d\alpha N_{(i,i+1)}/d\alpha N_{(i,i)}$ ratio was reversed in the case of Gly residues (positions 7 and 13 of the sequence) where a decrease of the interresidue connectivity was observed. Two weak HN_i-HN_{i+1} nOes for the residues Val5-Ala6 and Gly7-Lys8 and an almost complete series of weak cross-peaks between successive residues for the rest of the sequence (Thr9-Ile18) were detected (Figure 3b). No *cis*/*trans* isomerism around the Xaa-Pro bonds was observed, as suggested by the absence of $H\alpha_{i-1}$-$H\alpha_i$ nOes and the presence of the $H\alpha_{i-1}$-$H\delta_i$ nOe peaks, even at NOESY experiments with a mixing time of 600 ms. Additionally, the lack of any 'exchange' peak in the ROESY spectra, that could be attributed to different conformers, slowly interconverting in the NMR time scale, is indicative that the two Pro residues are in the *trans* conformation [34].

The relative nOe intensities, schematically represented in Figure 4a, suggest that the A peptide has a tendency for extended conformations with a preference to form loop structures in the vicinity of the two central Gly amino acids. No medium-(i,i + 2; i,i + 3; and/or i,i + 4) or long-range connectivities were detected.

Chemical shift deviations have been widely used in the course of the structural elucidation of peptides in solution. Chemical shift values provided by Wishart *et al.* [35] represent reference random coil NMR parameters for studies of peptide conformations even though the solvent used is not aqueous [36–38].

Chemical shift deviations for the Hα ($\Delta\delta$Hα) protons of the A peptide at 273 K, displayed in

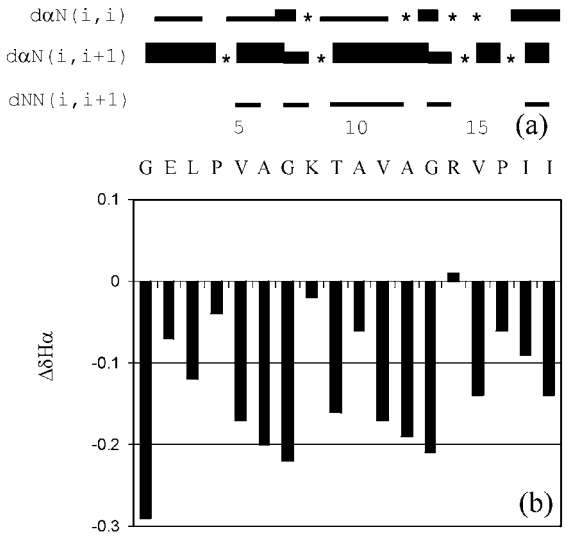

Figure 4 Schematic representation (a) of the observed nOes for the A peptide and plot of difference ($\Delta\delta$) of Hα chemical shifts between observed and random coil values (b). Line thickness for the intraresidual and sequential nOes reflects their intensity. NOe cross peaks for which we could not unambiguously estimate their relative intensity due to overlap are indicated by an asterisk (*).

Figure 4b, had mainly negative values. Interestingly, a common pattern was observed in the vicinity of the two central Gly residues: considering Gly at the i position in the sequence, upfield shifts bigger than 0.1 ppm were observed, with increasing values, from the i-2 to the i position. However, at the $i+1$ amino acid (positively charged Lys8 in the first segment, and Arg14 in the second) a clear inversion of the chemical shift deviation was observed in both segments. This pattern of the Hα chemical shift deviations was indicative of a turn and confirms that Gly residues play important role in the conformational behaviour of the peptide.

The temperature dependence of the amide proton chemical shifts was determined over the range 263 to 300 K (Table 1). The changes of the amide proton chemical shifts varied linearly with temperature (data not shown), suggesting that no major conformational rearrangements occur in the temperature range studied. The coefficients for most of the residues were similar, varying from 6.5 to 8.0 ppb/K. In the case of Glu2, Lys8 and Arg14 (i.e. the charged amino acids following the Gly residues in the sequence) the coefficients were found to be considerably lower (3.9, 3.7 and 3.3 ppb/K, respectively), suggesting a relatively higher protection from the solvent and indicating that in a significant percentage of the structures existing in solution these amide protons might participate in intramolecular hydrogen bonds.

No additional information on the structural characterization of the peptide was provided by the $^3J_{HN\alpha}$ coupling constants. Only for residues Leu3, Val15, Ile17 and Ile18 the coupling constant values were greater than 8 Hz, but their non sequential position in the peptide sequence did not allow any conformational conclusion to be drawn.

In summary, the NMR characteristics of most of the residues point to extended structures. However, a common pattern was observed around the Gly residues: low intensity of the $d\alpha N_{(i,i+1)}$ parameters, low temperature coefficients of the vicinal polar amino acids, and upfield shifts for the Hα nuclei. The segments of the peptide marked with these characteristics should adopt backbone arrangements with dihedral angles in the α-helical region of the (ϕ, ψ) conformational space [39].

B peptide. The B peptide, although exhibiting homology with the A peptide sequence, was only sparingly soluble in water. At concentrations of 1–3 mM it formed transparent solutions, which a few hours later (10–12 h) were characterized by high viscosity. It has been argued that insolubility of short peptides in water is an indicator of β-sheet structure formation [40]. The B peptide bears only one charged residue (Glu), whereas the A peptide has three (Glu, Lys and Arg) and one polar (Thr). In the B peptide the residues corresponding to Lys and Arg of the A peptide sequence are Thr and Glu. Moreover two Phe residues are also present in the B peptide. It is thus expected that the B peptide would be less water soluble compared with the A peptide. A similar behaviour was observed in CD_3OD/H_2O solutions at the concentrations required for the NMR experiments [11].

B peptide was thus studied in a $DMSO/H_2O$ (1 : 1) solution at a concentration of 2 mM and in the temperature range from 273 to 300 K. The chemical shift values are listed in Table 2.

The amino acid sequence was easily identified by a series of TOCSY and NOESY experiments performed at different temperatures. Part of the NOESY spectrum, illustrating the $H\alpha_i$-HN_i, $H\alpha_i$-HN_{i+1} connectivities, is provided in Figure 5a. An almost full series of medium to weak HN_i-HN_{i+1} nOe cross-correlations were detected (Figure 5b). The relative nOe intensities of the sequential backbone protons are depicted in Figure 6a. As in the case of the A peptide, a decrease of the $H\alpha_i$-HN_{i+1} intensity was observed at the Gly residues.

The two Pro residues (Pro4, Pro16) are both in the *trans* conformation, as in the case of the A peptide.

Hα chemical shift deviations from the random coil ($\Delta\delta$Hα) were negative along the whole peptide sequence (Figure 6b), with differences suggesting a relatively higher structural stability and organization. No medium- (i,i + 2; i,i + 3; and/or i,i + 4) or long-range nOe connectivities were observed. However, the rest of the structural characteristics referred to above provide evidence for the existence of structured populations in solution (probably 'open' loops).

Temperature coefficient data (Table 2) were more informative than those of the A peptide, as two residues (Gly7 and Glu14) exhibited considerably low values (1 ppb/K and 2.5 ppb/K, respectively).

$^3J_{HN\alpha}$ values (Phe5, Thr8, Ala9, Val11, Glu14, Thr17), measured at 278 K from a DQF-COSY spectrum, were indicative of extended conformations as they varied from 8.3 to 9.3 Hz. These values were larger compared with those of the A peptide, implying a higher structural stability. For the Asn2, Gly7, Gly10, Gly13 and Phe15 residues it was not possible to measure the $^3J_{HN\alpha}$ values due to signal overlapping. Leu3, Leu6 and Ala18 have

Table 2 Chemical Shifts of B Peptide in DMSO/H_2O (1 : 1) Solution at 278 K

Residue	HN	Hα	Hβ	Other protons	Temperature coefficient
Gly 1		3.62			
Asn 2	8.66	4.63	2.62/2.49	γNH_2 7.60/6.98	4.9
Leu 3	8.28	4.45	1.56/1.46	γ1.38 δ0.82	7.4
Pro 4	—	4.22	2.00/1.63	γ1.81δ3.61/3.47	
Phe 5	8.15	4.42	2.97/2.90	7.17/7.15	9.2
Leu 6	8.20	4.17	1.45/1.41	γ1.41 δ0.80/0.73	7.8
Gly 7	7.81	3.76/3.70			1.0
Thr 8	7.93	4.17	4.07	γ1.07	6.5
Ala 9	8.37	4.18	1.25		7.0
Gly 10	8.22	3.75			4.4
Val 11	7.85	4.00	1.95	γ0.78	4.9
Ala 12	8.43	4.11	1.25		7.1
Gly 13	8.24	3.77/3.67			6.3
Glu 14	7.90	4.11	2.13/2.03	γ1.76/1.70	2.5
Phe 15	8.24	4.68	3.05/2.75	7.21	7.4
Pro 16	—	4.40	2.16/1.89	γ1.89 δ3.64/3.53	
Thr 17	8.08	4.15	4.12	γ1.10	7.8
Ala 18	7.90	4.01	1.23		6.0

intermediate $^3J_{HN\alpha}$ values (7.3 Hz), typical of a flexible molecule in solution.

In conclusion, the B peptide is characterized by high chemical shift deviations, very low temperature coefficients for Gly7 and Glu14, low $H\alpha_i$-HN_{i+1} nOe intensities and an almost full range of sequential HN-HN. These parameters are indicative of turn and/or α-helical structures, but the absence of any medium range nOe connectivity (i,i + 2; i,i + 3; and/or i,i + 4) suggests the formation of 'open' loop structures.

Theoretical Calculations

Due to the presence of conformational averaging for these linear peptides in solution, the use of intraresidue and sequential nOes did not lead to single structures and were used only for the determination of structural families.

Introducing the nOe and chemical shift restraints in the simulated annealing calculations, all the generated structures are related through two common structural elements: the existence of 'open' loops located around the Gly residues alternating with short extended segments of three to four amino acids. Conformers were clustered in a small number of families representative of the conformational preferences of the A and B peptides.

A peptide. Thirty nine lower energy structures were selected and classified into the three conformational families (A-I, -II, -III) shown in Figure 7. Bends around the Gly residues were evident, while the other parts of the sequence showed extended conformations. The side chains of the charged amino acids following the Gly residues always point outward of the loops. The difference between families A-I and -II concerns the relative orientation of the *N*- and *C*-terminal strands (amino acids 1–6 and 15–18): in family A-I the two strands are almost parallel to each other and they are located on the same plane. In contrast, in family A-II the *N*- and *C*-terminal extended parts of the backbone are almost perpendicular to each other.

The common characteristics, which led to the clustering of the A-III family of conformers, were the presence of an additional bend localized at Val11 of the sequence. This arrangement is supported by the experimentally observed HN_i-HN_{i+1} weak correlations between Ala10-Val11 and Val11-Ala12.

B peptide. Twenty one structures with the lowest energy generated from simulated annealing calculations (B-I and B-II in Figure 7) share many similarities with the A peptide, as a consequence of both their sequence homology and the nOe restraints, which concerned only sequential connectivities, introduced in the calculations. The B-I family is related to the A-I family: bends of the sequence are localized at charged or polar

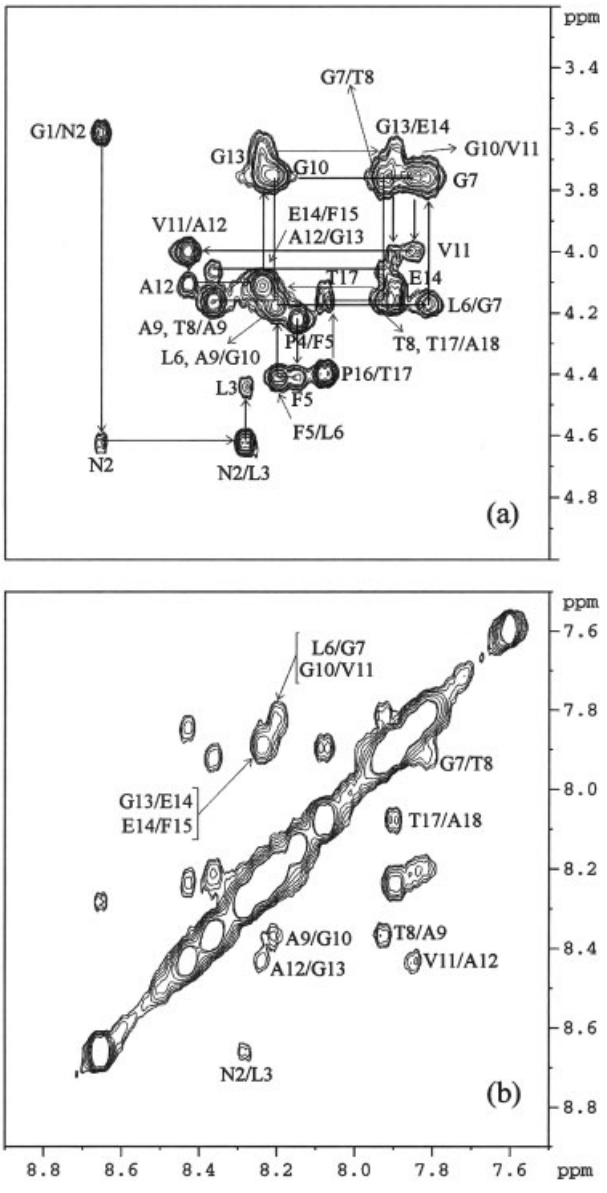

Figure 5 Fingerprint (a) and amide-amide proton (b) regions (Hα-HN and HN-HN) of NOESY spectrum ($t_m = 200$ ms) of the B peptide recorded at 278 K in DMSO/H_2O (1 : 1) solution.

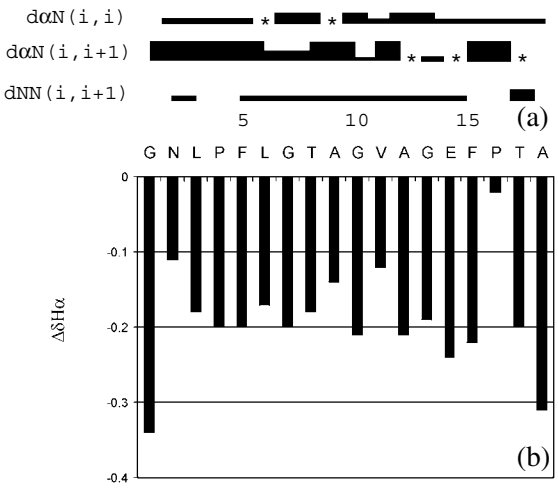

Figure 6 Schematic representation (a) of the observed nOes detected for the B peptide and plot of difference ($\Delta\delta$) of Hα chemical shifts between observed and random coil values (b).

residues following the Gly residues at positions 7 and 13.

Family B-II is similar to family A-III, but the bend localized at Val11 is manifested more clearly in the case of the B peptide. Probably, the flexibility of the sequence is affected by the presence of Gly10.

The above presented experimental and theoretical data allow the visualization of the conformational preferences of chorion peptides A and B in solution. Even though NMR data have only limited capacity to distinguish between a single folded peptide conformation and various mixtures of folded and unfolded conformations [41] the spectroscopic results described here highlight the tendency of the A and B peptides to form local structural elements: extended segments of the sequence and localized loops. The Gly residues (positions 7 and 13), followed by charged or polar amino acid, break the continuity of the extended segments in both the peptide sequences. The small temperature coefficient values of charged amino acids, the ratios of NOESY cross-peak intensities, the $d\alpha N_{(i,i+1)}/d\alpha N_{(i,i)}$ ratio of each Gly with the following residue, and the Hα chemical shift deviations support the formation of two bends around residues 7–8 and 13–14 for both peptides. Although these bends change the backbone direction, allowing the 3–6 and 15–18 extended parts of the sequence to freely move relatively to the central segment (residues 9–12), they do not result in a backbone reversal. Consequently, no interactions are established among side-chain protons of residues located on successive strands (extended part of the sequence), neither intramolecular hydrogen bonds are formed. Though charged and polar residues exist in the peptide sequences, backbone arrangements do not provide the necessary spatial proximity for coulombic interactions to be developed. In agreement with NMR data, analysis of CD spectra revealed β-sheet and turn structural elements with a percentage higher than 40%.

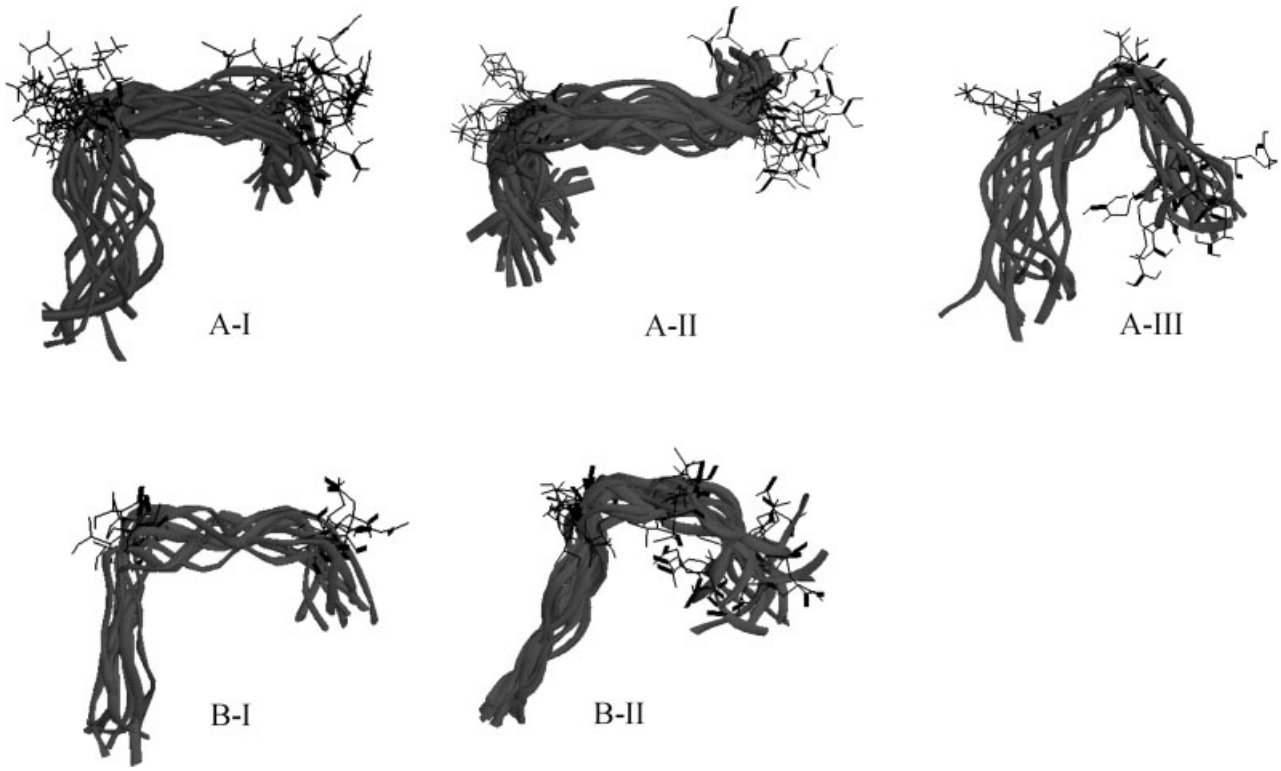

Figure 7 Superposition of selected conformers of the A and B peptides (clustered into families), as a result of our theoretical calculations.

Experimental data from laser-Raman and FT-IR absorption studies of the A and B peptides both in solution and in the solid state [9], as well as x-ray diffraction, FT-Raman and ATR-FTIR data of amyloid-like fibrils formed from chorion peptides [11,12] suggest the preponderance of the antiparallel β-pleated sheet structure. Additionally, the left-handed parallel β-sheet helix [10] model structure has recently been proposed by Iconomidou *et al.* [11,12], as an attractive alternative to describe the folding and self-assembly of chorion peptides. According to this model, there is a central hydrophobic core region forming a triangular prism-like helix; the edges of the prism are occupied by Gly and charged or polar residues, while four-residue extended segments lie on the prism planes.

It has been anticipated that the hexapeptide periodicity and the presence of Gly residues dictate the secondary structure of the fibrils in chorion proteins and peptide analogues [3]. The data reported in this work are of considerable importance since they clearly establish, for the first time, experimental evidence that local folded structures around the Gly residues do occur in segments of the chorion proteins in solution. The structural characteristics of both peptides in solution, emerging from the experimental data, provide a better insight for the initiation of the chorion protein folding process. The structures described in this work may represent early folding intermediates capable of directing subsequent folding events and the nucleation of the final functionally important conformation, the amyloid-like fibrils. However, due to the short length of the peptide sequences, plausible 'triangular' [10] and/or antiparallel β-sheet structures [9] are probably not stable enough. In any case, it remains to be seen which of the two models (left-handed parallel β-helix or 'cross-β' antiparallel β-sheet structure) actually represents the basic folding motif of the amyloid-like fibrils formed by chorion peptides.

Acknowledgements

D.B. was the recipient of short term EMBO and SRMB fellowships. Thanks are due to Dr S.T. Case for the synthesis of peptides A and B. The authors gratefully acknowledge Dr A. Pastore for continuing support during this work.

REFERENCES

1. Regier JC, Kafatos FC. In *Comprehensive Insect Biochemistry, Physiology and Pharmacology*, Gilbert LI, Kerkut GA (eds). Pergamon Press: Oxford, 1985; **1**: 113–151.
2. Goldsmith MR, Kafatos FC. Developmentally regulated genes in silkmoths. *Annu. Rev. Genet.* 1984; **18**: 443–487.
3. Hamodrakas SJ. In *Results and Problems in Cell Differentiation*, Case ST (ed.). Springer: Berlin, 1992; **19**: 115–186.
4. Kafatos FC, Regier JC, Mazur GD, Nadel MR, Blau HM, Petri WH, Wyman AR, Gelinas RE, Moore PB, Paul M, Efstratiadis A, Vournakis JN, Goldsmith MR, Hunsley JR, Baker B, Nardi J, Koehler M. In *Results and Problems in Cell Differentiation*, Beerman W (ed.). Springer: Berlin, 1977; **8**: 45–145.
5. Lecanidou R, Rodakis GC, Eickbush TH, Kafatos FC. Evolution of the silkmoth chorion gene superfamily: gene families CA and CB. *Proc. Natl Acad. Sci. USA* 1986; **83**: 6514–6518.
6. Hamodrakas SJ, Jones CW, Kafatos FC. Secondary structure predictions for silkmoth chorion proteins. *Biochim. Biophys. Acta* 1982; **700**: 42–51.
7. Hamodrakas SJ, Etmektzoglou T, Kafatos FC. Amino acid periodicities and their structural implications for the evolutionarily conservative central domain of some silkmoth chorion proteins. *J. Mol. Biol.* 1985; **186**: 583–589.
8. Hamodrakas SJ, Bosshard HE, Carlson CN. Structural models of the evolutionarily conservative central domain of silkmoth chorion proteins. *Protein Eng.* 1988; **2**: 201–207.
9. Benaki DC, Aggeli A, Chryssikos GD, Yiannopoulos YD, Kamitsos EI, Brumley E, Case ST, Boden N, Hamodrakas SJ. Laser-Raman and FT-IR spectroscopic studies of peptide analogues of silkmoth chorion protein segments. *Int. J. Biol. Macromol.* 1998; **23**: 49–59.
10. Yoder MD, Jurnak F. Protein motifs. 3. The parallel β-helix and other coiled folds. *FASEB J.* 1995; **9**: 335–342.
11. Iconomidou VA, Chryssikos GD, Gionis V, Vriend G, Hoenger A, Hamodrakas SJ. Amyloid-like fibrils from an 18-residue peptide analogue of a part of the central domain of the B-family of silkmoth chorion proteins. *FEBS Lett.* 2001; **499**: 268–273.
12. Iconomidou VA, Vriend G, Hamodrakas SJ. Amyloids protect the silkmoth oocyte and embryo. *FEBS Lett.* 2000; **479**: 141–145.
13. Sunde M, Blake CCF. From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quart. Rev. Biophys.* 1998; **31**: 1–39.
14. Sunde M, Serpell LC, Bartlam M, Fraser PE, Pepys MB, Blake CCF. Common core structure of amyloid fibrils by synchrotron x-ray diffraction. *J. Mol. Biol.* 1997; **273**: 729–739.
15. Provencher SW, Glockner J. Analysis of the components present in kinetics (or titration) curves. *J. Biochem. Biophys. Methods* 1983; **7**: 331–334.
16. Perczel A, Hollosi M, Tusnady G, Fasman GD. CONVEX constraint analysis: a natural deconvolution of circular dichroism curves of proteins. *Protein Eng.* 1991; **4**: 669–679.
17. Chang CT, Wu CSC, Yang JT. Circular dichroic analysis of protein conformation: inclusion of the β-turns. *Anal. Biochem.* 1978; **91**: 13–31.
18. Bodenhausen G, Kogler H, Ernst RR. Selection of coherence-transfer pathways in NMR pulse experiments. *J. Magn. Reson.* 1984; **58**: 370–388.
19. Kumar A, Ernst RR, Wüthrich K. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* 1980; **95**: 1–6.
20. Bothner-By AA, Stephens RL, Lee JM, Warren CD, Jeanloz RW. Structure determination of a tetrasaccharide: transient nuclear Overhauser effects in the rotating frame. *J. Am. Chem. Soc.* 1984; **106**: 811–813.
21. Bax A, Davis DG. MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J Magn. Reson.* 1985; **65**: 355–360.
22. Rance M, Sorensen OW, Bodenhausen G, Wagner G, Ernst RR, Wüthrich K. Improved spectral resolution in COSY ^{1}H-NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* 1983; **117**: 479–485.
23. Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* 1992; **2**: 661–665.
24. Sklenar V, Piotto M, Leppik R, Saudek V. Gradient-tailored water suppression for ^{1}H-^{15}N HSQC experiments optimized to retain full sensitivity. *J. Magn. Reson., Ser. A* 1993; **102**: 241–245.
25. Borgias BA, James TL. MARDIGRAS A procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure. *J. Magn. Reson.* 1990; **87**: 475–487.
26. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR system. A new software suite for macromolecular structure determination. *Acta Crystallogr.* 1998; **D 54**: 905–921.
27. Palmer AG, Rance M, Wright PE. Intramolecular motions of a zinc finger DNA-binding domain from XFIN characterized by proton-detected natural abundance ^{12}C heteronuclear NMR spectroscopy. *J. Am. Chem. Soc.* 1991; **113**: 4371–4380.

28. Genest D, Simorre JP. Method for evaluating the reliability of distances and rotational correlation times deduced from 2D ^{1}H-NMR NOESY experiments. *Magn. Reson. Chem.* 1990; **28**: 21–24.
29. Stein EG, Rice LM, Brunger AT. Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation. *J. Magn. Reson.* 1997; **124**: 154–164.
30. Shenkin PS, McDonald DQ. Cluster-analysis of molecular conformations. *J. Comput. Chem.* 1994; **15**: 899–916.
31. Mohamadi F, Richards NGJ, Guida WC, Liskamp R, Lipton M, Caufield C, Chang G, Hendrickson T, Still WC. MACROMODEL. An integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* 1990; **11**: 440–467.
32. Zhou NE, Mant CT, Hodges RS. Effect of preferred binding domains on peptide retention behaviour in reversed-phase chromatography: amphipathic α-helices. *Pept. Res.* 1990; **3**: 8–20.
33. Wüthrich K. In *NMR of Proteins and Nucleic Acids*. Wiley: New York, 1986.
34. Hinck AP, Eberhardt ES, Markley JL. NMR strategy for determining Xaa-Pro peptide bond configurations in proteins mutants of staphylococcal nuclease with altered configuration at proline-117. *Biochemistry* 1993; **32**: 11 810–11 818.
35. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* 1995; **5**: 67–81.
36. Gesell J, Zasloff M, Opella SJ. Two-dimensional ^{1}H-NMR experiments show that the 23-residue magainin antibiotic peptide is an α-helix in dodecylphosphocholine micelles, sodium dodecylsulfate micelles, and trifluoroethanol/water solution. *J. Biomol. NMR* 1997; **9**: 127–135.
37. Blond A, Cheminant M, Segalas-Milazzo I, Peduzzi J, Barthelemy M, Goulard C, Salomon R, Moreno F, Farias R, Rebuffat S. Solution structure of microcin J25, the single macrocyclic antimicrobial peptide from *Escherichia coli*. *Eur. J. Biochem.* 2001; **268**: 2124–2133.
38. Bak M, Sorensen MD, Sorensen ES, Rasmussen LK, Sorensen OW, Petersen TE, Nielsen NC. The structure of the membrane-binding 38C-terminal residues from bovine PP3 determined by liquid- and solid-state NMR spectroscopy. *Eur. J. Biochem.* 2000; **267**: 188–199.
39. Dyson HJ, Sayre JR, Merutka G, Shin HC, Lerner RA, Wright PE. Folding of peptide fragments comprising the complete sequence of proteins: models for initiation of protein folding. 2. Plastocyanin. *J. Mol. Biol.* 1992; **226**: 819–835.
40. Narita M, Doi M, Nakai T, Takegahara H. Synthesis and solubility properties of peptide fragments of human hemoglobulin α-chain. *Int. J. Pept. Protein Res.* 1988; **32**: 200–207.
41. Daura X, Antes I, van Gunsteren WF, Thiel W, Mark AE. The effect of motional averaging on the calculation of NMR-derived structural properties. *Proteins* 1999; **36**: 542–555.

