Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Magnetic Resonance 178 (2006) 106–120
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

A DOQSY approach for the elucidation of torsion angle distributions in biopolymers: Application to silk

Jacco D. van Beek, Beat H. Meier *

Physical Chemistry, ETH Zurich, CH-8093 Zurich, Switzerland

Received 18 July 2005; revised 4 September 2005
Available online 21 October 2005

Abstract

Silk from the wild silkworm *Samia cynthia ricini* with a molecular mass of about 300 kDa consists of alternating repeats of nano-crystalline poly-(Ala) and non-crystalline glycine-rich domains. The backbone torsion angles between pairs of these two amino acids is determined by DOQSY solid-state NMR spectroscopy: the alanine-rich domains are predominantly in a β-sheet conformation, whereas the glycine-rich domains are found to be partially in an extended β-sheet conformation and partially in an approximately 3_1-helical conformation. In the cast film from liquid silk significantly different secondary structures were found: the alanine-rich domains are α-helical conformation, whereas the results for glycine-labelled sample are explained by a random-coil state. A detailed error analysis of the technique is presented.

Keywords: Solid-state NMR; DOQSY; Torsion-angle distribution; Regularization; Silk; *Samia cynthia ricini*

1. Introduction

Silks are fibrous proteins of considerable economic and scientific interest. The material is made by various animals for different biological purposes. Some silks exhibit interesting mechanical properties, e.g., a high toughness comparable to Kevlar combined with an extensibility of 30% or an extreme extensibility of more than 200% [1]. The primary structure of silks is typically highly repetitive and rich in glycine and alanine [2–8]. X-ray crystallography has shown the material to be partially crystalline, but much less is known about the non-crystalline parts [9].

The Eri silkworm *Samia cynthia ricini* is a wild silkworm that produces housing silk for protection during metamorphosis. Its primary structure is similar to that of spider dragline silk [10] and consists of highly conserved poly-alanine stretches of 10–14 amino acids interleaved by glycine-rich parts. The mechanical properties are quite different, though, with lower modulus and breaking energy [11]. Considering the different evolutionary pressures on the material performance, this finding is not astonishing but it would be interesting to achieve a molecular understanding of the differences. Silk is polymorphic, like many other proteins [12], and therefore the processing of the liquid material is of great importance to the final molecular structure [13].

The structure of *S. c. ricini* silk has previously been studied with one-dimensional (1D) liquid- and solid-state NMR measurements [14–24]. These 1D NMR studies mainly used the isotropic chemical-shift values to indicate the local structure of the protein backbone, by employing the empirical dependence of the chemical shift on secondary structure. These isotropic shifts can be powerful indicators for the local structure. They are, however, not unique and, in addition, difficult to use for glycine, which shows little chemical-shift variation with secondary structure. To overcome these limitations 1D solid-state NMR measurements on drawn, uniaxially oriented samples were performed as a function of the sample orientation with respect to the external magnetic field [18]. Those experiments have yielded the

* Corresponding author. Fax: +11 411 44 632 1621.
E-mail address: beme@ethz.ch (B.H. Meier).

1090-7807/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2005.09.004

orientation of the chemical-shielding tensors relative to the external magnetic field and, because the orientation of these tensors in a molecule-fixed coordinate system is approximately known, information about the secondary structure can be obtained from the chemical-shift pattern of such partially oriented samples. Similar information is obtained from 2D DECODER experiments [9,25–28]. The precision of these measurements is, however, limited by the degree of uniaxial order in the sample.

A more direct approach to the determination of the backbone conformation is the direct measurement of distances and, particularly, torsion angles. In recent years, a flurry of new solid-state NMR experiments have been proposed to this end [29–39]. Many of them are based on magic-angle spinning (MAS), which is indispensable for obtaining high-resolution data from highly ordered materials. For structurally heterogeneous materials like silk, however, the heterogeneous linewidth of MAS spectra is quite broad (6.5 ppm FWHH) and the main advantage of MAS, narrow lines and high signal-to-noise ratio, cannot be fully realized. Static experiments correlating CSA tensors [40–42] are therefore often better suited for the investigation of such materials. Even though the signal to noise ratio is lower, the sensitivity of the signal towards the structural parameters of interest is often higher. Furthermore, in contrast to fast MAS methods, the application of recoupling pulse schemes, with the associated imperfections that can lead to systematic errors in the quantitative analysis, is not necessary.

Here, we apply a variation of the 2D DOQSY experiment (correlation of the double-quantum and single-quantum spectra) [13,43–45] to determine the torsion-angle distributions in silk from *S. c. ricini*. For a two-spin system, the double-quantum dimension represents the sum chemical-shift tensor, the single-quantum dimension the chemical-shift and dipolar tensors. The 2D pattern is indicative of the relative orientation of the two involved chemical-shift tensors. The DOQSY experiment applied to ^{13}C carbonyl-labelled peptides can determine both backbone torsion angles (ϕ,ψ) in a single experiment (see Fig. 1A) without the need for oriented samples. In the experiments described here, Ala-Ala and Gly-Gly dyads were investigated.

We investigated two states of *S. c. ricini* silk, the silk fibre, being the native state of the solid material, and a cast film artificially made from the liquid silk of the gland. The experimental DOQSY spectra and corresponding best fits for both the alanine- and glycine-labelled materials are shown in Fig. 2. A short report describing the alanine-rich part of this silk has already been published [13] and similar studies have also been performed on spider dragline silk [9]. Numerical aspects of the data analysis have been described in detail in [46]. In this paper, we present a detailed analysis of the backbone conformation of silk from *S. c. ricini* in the microcrystalline alanine-rich domains, as well as in the non-crystalline glycine-rich domains. We identify possible sources of error and estimate their importance.

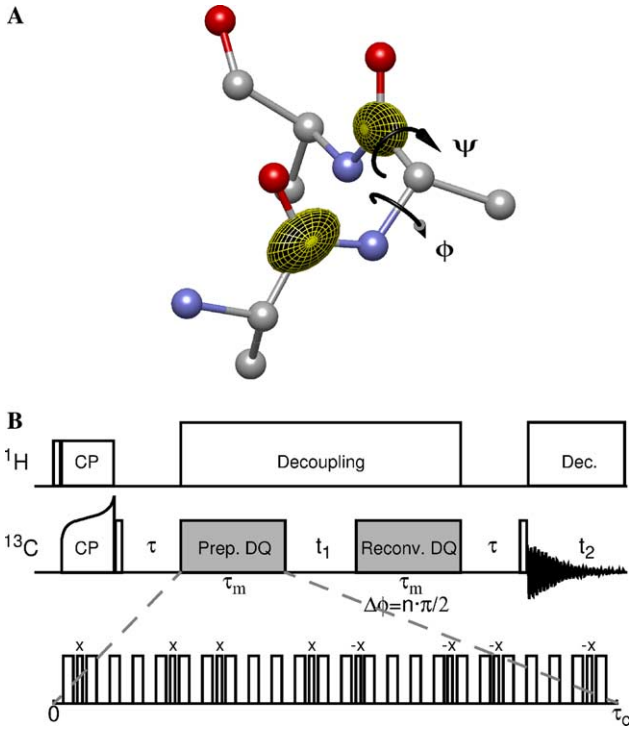

Fig. 1. (A) Goal of the experiment is to measure the backbone torsion angle pairs (ϕ,ψ) of neighbouring amino acids in a single experiment, using $^{13}C{=}O$ labelled protein samples, (B) experimental implementation. After cross-polarization and a short *z*-filter a double-quantum state is created and allowed to evolve during t_1. The same sequence was used to reconvert back into single-quantum coherence, which was followed by a short *z*-filter and detection.

2. Evaluation of the analysis

The primary structure of silks show many repetitive motifs that are rich in glycine and alanine. Cocoon silk from *S. c. ricini* contains highly conserved $(Ala)_n$ $(n = 10–12)$ and $(Gly)_2/(Gly)_3$ stretches [24], which largely correspond to the majority of the crystalline and non-crystalline domains, respectively. Thus, by applying an isotopic labelling strategy, which enriches either alanine or glycine, the backbone structure of both domains can be characterized. Assuming that two carbonyl carbons, neighbouring in the peptide chain, are isotopically labelled, the corresponding 2D DOQSY spectrum obtained is sensitive to the molecular torsion angles (ϕ,ψ) (Fig. 3). Note that an inherent symmetry in the experiment makes it impossible to distinguish the spectra $S(\phi,\psi)$ and $S(-\phi,-\psi)$.

A number of assumptions will be made to analyse these spectra in terms of the torsion angles (ϕ,ψ). These include the selectivity through labelling, the peptide geometry, CSA principal values, the orientation of the chemical-shielding tensor in the peptide, simulation of the pulse sequence, and the exclusion of multi-spin effects from the basis spectra. In the following, we show that it is acceptable to use a single set of CSA tensor values, independent of the torsion angles, and a single tensor orientation in the molecular frame to fit the data. Multi-spin effects are generally

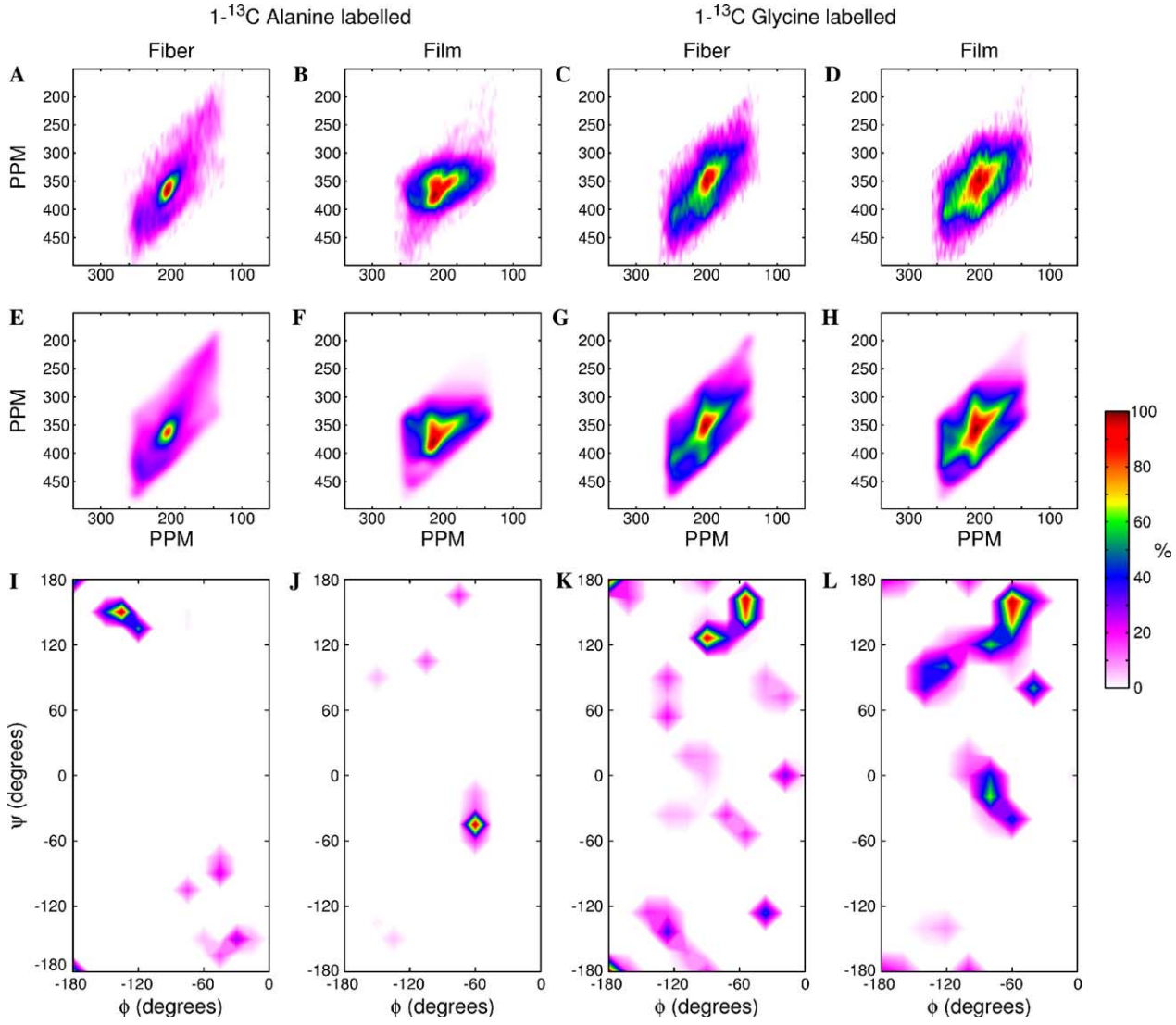

Fig. 2. Experimental 2D DOQSY spectra of (A–D) four different silk samples from the silkworm *S. c. ricini* with (E–H) corresponding best fits and (I–L) distribution functions. The samples were labelled either with [1-^{13}C]alanine or [1-^{13}C]glycine and prepared either as a fibre or a cast film. The alanine-labelled samples were measured at a ^{13}C resonance frequency of 100 MHz, whereas the glycine-labelled samples at 75 MHz. The S/N ratios are approximately 47, 60, 35, and 26 for the four spectra, respectively. The grid resolution was 15°, 15°, 18°, and 20° for the four PDFs, respectively. All PDFs, except (I) show the symmetry $S(\phi,\psi) = S(-\phi,-\psi)$. For the fibrous alanine-labelled sample a bias function was used to break the symmetry using the C_α chemical shift, as is described in the text.

small but must be accounted for on the basis of the primary structure. Furthermore, it is shown that one can simulate the spectra in frequency domain and scale them a posteriori to include effects due to finite rf field strengths. Finally, the limiting case of complex distribution functions in proteins, a random coil structure, is investigated.

2.1. Description of the DOQSY experiment

To zeroth-order average Hamiltonian theory, the idealized multiple-pulse sequence of Fig. 1B produces a pure double-quantum (DQ) Hamiltonian of the form [47–49]

$$\mathcal{H}_d = \sum_{i>j} d_{ij}(S_{ix}S_{jx} - S_{iy}S_{jy}). \tag{1}$$

The dipolar coupling constant is defined as

$$d_{ij} = -\frac{1}{2}\cdot\frac{\gamma_i\gamma_j\hbar\mu_0}{4\pi}\cdot\frac{1}{r_{ij}^3}(3\cos^2\theta_{ij} - 1) \tag{2}$$

with θ_{ij} being the angle between the static magnetic field and the internuclear vector of spins i and j. For a two-spin system, the density operator σ, after cross-polarization and a $\pi/2$ pulse, is proportional to $S_{1z} + S_{2z}$. Evolving the system for a period τ_m under the DQ Hamiltonian yields

$$\sigma(\tau_m) = \cos(d_{12}\tau_m)(S_{1z} + S_{2z}) + 2\sin(d_{12}\tau_m) \times (S_{1x}S_{2y} + S_{1y}S_{2x}), \tag{3}$$

which after the t_1 evolution, becomes

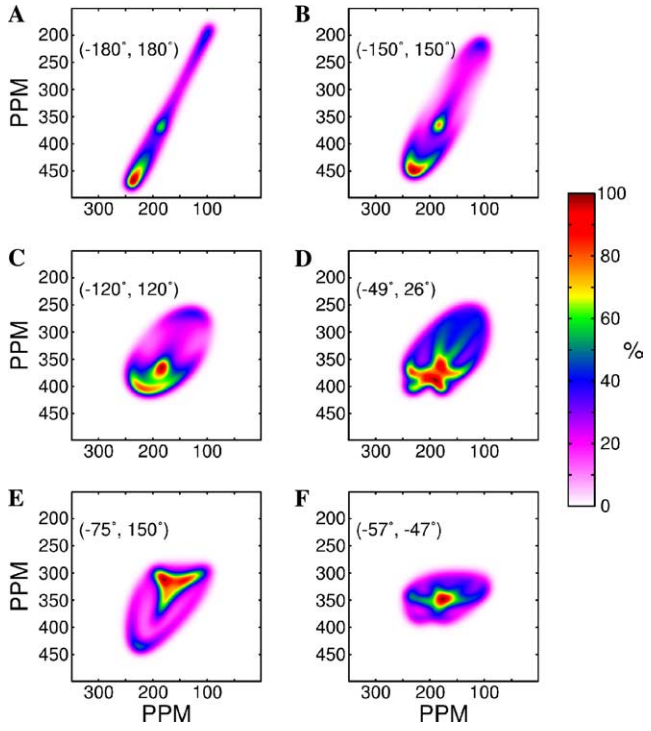

Fig. 3. Simulated 2D DOQSY spectra for different conformations and an excitation time of $\tau_m = 1.5$ ms. (A) Extended conformation with $(\phi,\psi) = (-180°, 180°)$, (B) β-sheet with $(\phi,\psi) = (-150°, 150°)$, (C) β-sheet with $(\phi,\psi) = (-120°, 120°)$, (D) 3_{10}-helix with $(\phi,\psi) = (-49°, -26°)$, (E) 3_1-helix with $(\phi,\psi) = (-75°, 150°)$, and (F) α-helix with $(\phi,\psi) = (-57°, -47°)$. A plot of basis spectra over the full (ϕ,ψ) range can be found in [13].

$$\begin{aligned}\sigma(t_1;\tau_m) = {} & \cos(d_{12}\tau_m)(S_{1z}+S_{2z}) + 2\cos(\Omega_1 t_1 + \Omega_2 t_1) \\ & \times \sin(d_{12}\tau_m)(S_{1x}S_{2y} + S_{1y}S_{2x}) \\ & - 2\sin(d_{12}\tau_m)\sin(\Omega_1 t_1 + \Omega_2 t_1)(S_{1x}S_{2x} - S_{1y}S_{2y}).\end{aligned} \tag{4}$$

Here, Ω_1 and Ω_2 are the chemical shifts of the two spins involved, which depend on the orientation of the molecule with respect to the magnetic field. Reconverting back into single-quantum coherence, and adding the signal using an appropriate phase cycle, yields (only showing terms proportional to S_z)

$$\sigma(\mathrm{t}_1;\tau_m) = \cos(\Omega_1 t_1 + \Omega_2 t_1)\sin^2(d_{12}\tau_m)(S_{1z}+S_{2z}). \tag{5}$$

For a phase-sensitive 2D signal the corresponding sine modulation in t_1 is obtained by phase-shifting the reconversion part by π/4 after each t_1 increment (TPPI). A final π/2 pulse creates transverse magnetization and the detected signal is effectively given by

$$\begin{aligned}S(t_1,t_2;\tau_m) = {} & \sin^2(d_{12}\tau_m)\exp(-\mathrm{i}\Omega_1 t_1 - \mathrm{i}\Omega_2 t_1) \\ & \times (\exp(-\mathrm{i}\Omega_1 t_2) + \exp(-\mathrm{i}\Omega_2 t_2)),\end{aligned} \tag{6}$$

if we neglect the dipolar interaction in the direct dimension. Assuming two ^{13}C-labelled carbonyl groups, neighbouring in the primary sequence (Fig. 1A), the torsion angle information (ϕ,ψ) is encoded directly in the spectrum by the correlation of the sum tensor, $(\Omega_1 + \Omega_2)$, with the individual chemical-shift tensors of the spins, Ω_1 and Ω_2. The orientation of the dipolar tensor, with respect to the two chemical-shielding tensors, is dependent on (ϕ,ψ) and results in an additional (ϕ,ψ)-specific modulation of the intensities. Also, the dipolar frequency d_{12} contains the internuclear distance r_{12}, which is modulated by the torsion angle ϕ.

2.2. III-posedness

The static lineshape of a single conformation (ϕ,ψ) in the DOQSY experiment, as it is done here, has a highly distinctive spectrum, which should allow for an accurate determination of the torsion angles. As is shown in Fig. 3, a wide range in spectral patterns is observed as a function of the torsion angles. Typically, heterogeneous materials, like silk, do not exhibit sharply defined structural parameters. Instead, a (possibly complex) distribution function of the structural parameters of interest, in this case the torsion angles, is often observed. It is well known that this complicates the analysis of experimental data considerably and a careful analysis strategy must be chosen to avoid miss- and overinterpretation of the data, whilst still using all the information contained in the experimental data. For a detailed description on the regulatory approach that was used in this work, we refer to [46]. Here, we only give the main conclusions from that work.

From a numerical point of view, the DOQSY experiment, as proposed here, is a difficult experiment to analyse when structural distributions are present. Therefore, the grid resolution of the PDF was taken to be at least 15° and in the experimental spectra of the glycine-labelled samples grid resolutions of 18° and 20° were used in order to stabilize the solution. The main reason for having to use this rather low resolution is the significant degree of correlation between the basis spectra for different conformations across the Ramachandran map. This is illustrated in Fig. 4, which shows squared-difference plots of six simulated 2D DOQSY spectra that are compared to all basis spectra in the kernel. The differences strongly depend on the torsion angles, e.g. the experiment is very distinctive around the α-helical region and around the extended conformation (±180°, ±180°). Other conformations, e.g., (∓120°, ±120°) or (±75°,±15°) show much more correlation with other basis spectra. Such correlations clearly affect the reliability of the analysis. The noise level was shown to be an important factor as well [46]. Since typically S/N ratios of only 25–35 were achieved experimentally, small but highly correlated intensities are expected to appear in the PDFs, as are indeed observed in the experimental results as shown below (Fig. 2). From a large number of tests of simulated data the positions of strong peaks in the PDF have been shown to be reliable, though [46]. Significant structural information is thus obtained, albeit of relatively low resolution.

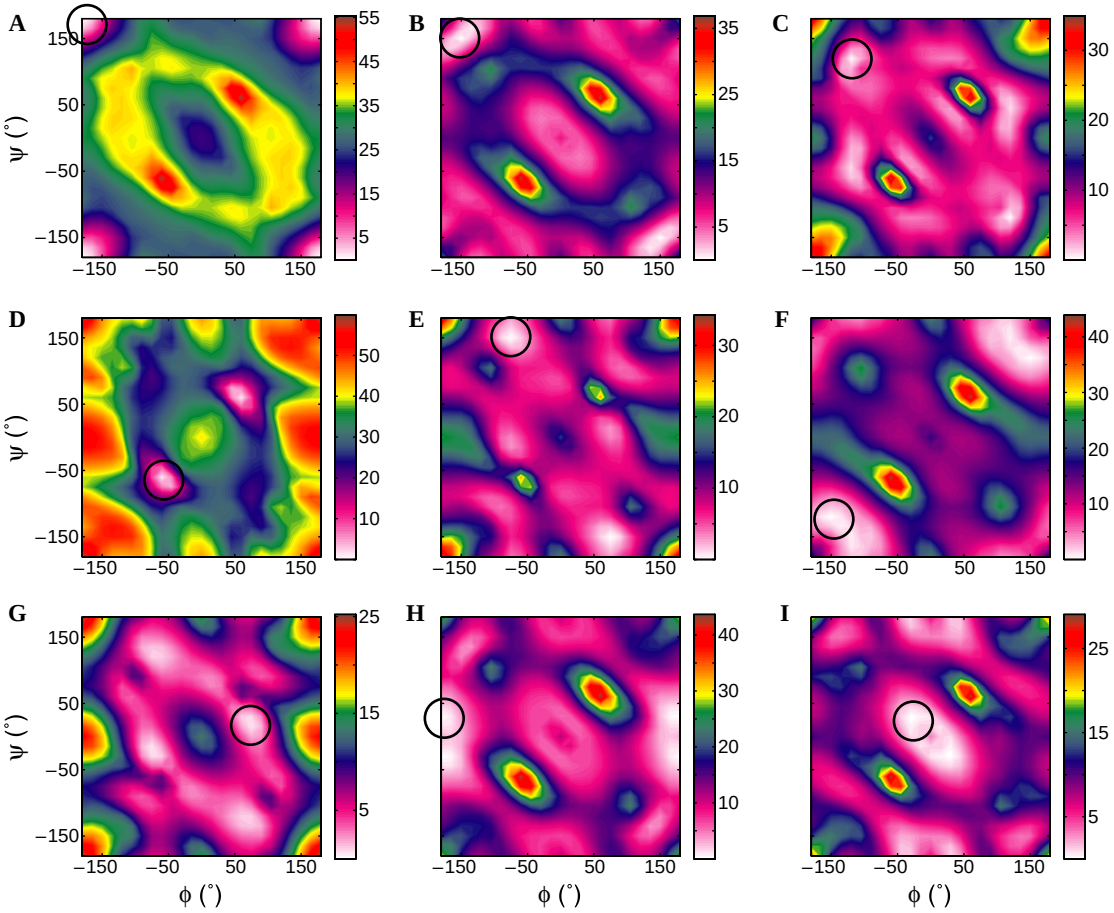

Fig. 4. Squared-difference plots of nine simulated DOQSY spectra, compared to a full set of basis spectra with a 15° grid resolution. The conformations for the single spectra are (ϕ,ψ) = (A) (−180°, −180°) (extended β-sheet), (B) (−150°, 150°) (β-sheet), (C) β-sheet (−120°, −120°), (D) helical (−60°, −60°), (E) 3_1-helical/β-turn (−75°, 150°), (F) (−150°, −120°), (G) β-turn (75°, 15°), (H) (−180°, 30°), and (I) (−30°, 30°). The black circles mark the simulated conformations. The intensity of the plots is in arbitrary units, but is absolute between the plots and hence allows direct comparison.

2.3. Selectivity

Through the double application of the excitation/reconversion DQ cycle, the intensity at a given point in the spectrum is approximately proportional to $\sin^2(d_{12}\tau_m)$, where d_{12} is the dipolar coupling constant and τ_m the excitation/reconversion time. In the initial rate regime this implies a $1/r^6$ dependence, where r is the internuclear distance between two spins. The high degree of spatial selectivity is used to ensure sufficient distinction between intra-chain and inter-chain contacts, since only next-neighbour intra-chain contacts reflect the backbone torsion angles directly.

The success of this approach strongly depends on the assumption that the shortest distance is intra-chain and care must be taken since there may also be comparably small inter-chain distances. In principle, this requires a case-to-case consideration of the system of interest, although some general remarks can be made. Fig. 5 shows the results from a search of distances between carbonyl groups, from the Protein Data Bank (release #89, 5165 protein and enzyme structures included, determined by NMR and X-ray). The $(i, i+1)$ distance ranges from 2.8 to 3.7 Å and there is a clear distinction between helical and sheet-like structures. The next smallest distances, for a helical structure, are the $(i, i+2)$ and $(i, i+3)$ contacts with distances of typically 4.5 and 4.8 Å, respectively. For sheet-like structures the smallest $(i, i+2)$ distance is typically found to be 6.5 Å, which is similar to typical inter-chain distances. Much smaller distances may be found in special cases, however, as is exemplified by the poly-glycine β-sheet structure, which has an inter-sheet distance of approximately 3.5 Å [50]. Turn-like structures can have rather small $(i, i+4)$ and $(i, i+5)$ distances of around 5–6 Å. Except for special cases, the intramolecular neighbour distance can be expected to be the shortest and this allows the determination of the backbone torsion angles.

Eq. (6) shows that the DOQSY spectrum is proportional to $\sin^2 d_{12}\tau_m$, which, in the initial rate, is proportional to $1/r^6$. Note that depending on the mixing time (i.e., how far off initial rate regime) and the dipolar coupling, the dependence may change considerably across the spectrum, and a general expression to denote the separating power between two distances cannot be given. In most cases this separation is good, but not perfect, and, depending on the system of interest, additional rationale may be needed

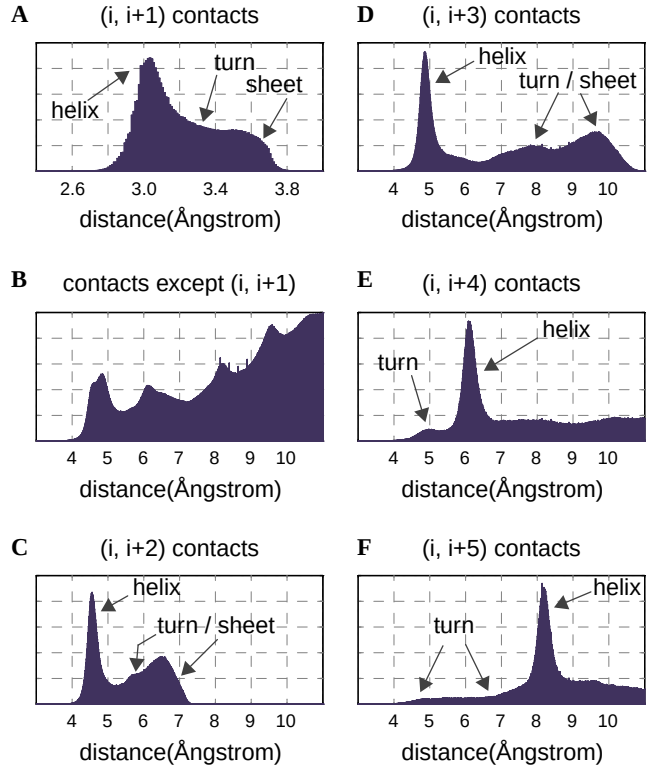

Fig. 5. Various intra-chain distances $[i, i+n]$ in proteins between C=O carbons were determined empirically by doing a search through 5165 protein and enzyme structures in the PDB database (release #89, determined by NMR and X-ray). (A) $[i, i+1]$, (B) all inter- and intra-chain distances except $[i, i+1]$, (C) $[i, i+2]$, (D) $[i, i+3]$, (E) $[i, i+4]$, and (F) $[i, i+5]$. Tentative assignments for commonly observed structural motifs are given. Approximately 2.5×10^6 distances were sampled for each plot, except for (B) where 67×10^6 distances were sampled.

a posteriori to account for certain intensities in the PDF. These can then be based on the primary structure and the observed secondary structure elements. For example, in the distribution function obtained for the fibrous alanine-labelled silks [9,13], a strong intensity was found at $(\pm 180°, \pm 180°)$. This implies a fully extended β-sheet structure for the silk, which in itself is not as energetically favourable as a less-extended sheet. By assigning this intensity to an inter-chain contact a more likely explanation was found: in a repetitive β-sheet all non-next-neighbour contacts will either be from C=O's parallel to the first spin of a pair, or parallel to the second one. This will give a sharp diagonal contribution to the spectrum. By examining the basis spectra, it was found that such a spectrum is best represented by the basis spectrum of the conformation $(-180°, 180°)$. Clearly, some inter-chain contacts may give spectra that the kernel cannot reproduce. This will lead to similar artefacts as observed for errors in the chemical-shift tensor, as shown below.

2.4. Multi-spin effects

The experimental spectra will be analysed using two-residue fragments, Eq. (6), and not using multi-spin systems. The justification of this approach depends on the isotopic labelling degree and the secondary structure. A higher labelling degree increases the probability of having three labelled neighbours within the same peptide chain, although it depends on the primary structure whether this can occur at all, and the chance of measuring signal from two spins in different peptide chains also increases. In what fashion the analysis is finally affected depends on the differences in spectrum for the two- and three-spin contributions. In the following the disturbances due to three-spin contributions, both intra- and inter-chain, are discussed.

The labelling degree of the *S. c. ricini* silk was analysed using mass spectroscopy [51] and showed only low amounts of cross labelling (Table 1). Contributions, other than from Ala-Ala or Gly-Gly, are thus small (<3%). The alanine-rich domains $(Ala)_n$ consist of 10–14 alanine residues, neighbouring in the primary sequence [24]. Assuming an average of $n = 12$, the ratio of three-spin/two-spin contributions to the DOQSY spectrum is approximately ((10/22) · labelling degree). This means that three-spin contributions are smaller than two-spin contributions by a factor of 1/23 and 1/14 for the fibre and film samples, respectively. From the primary structure of *S. c. ricini* [24] it follows that approximately 60% of the primary sequence consists of glycine, in the glycine-rich domains. The ratio of $(Gly)_1$:$(Gly)_2$:$(Gly)_3$ is approximately 5:4:2. At a labelling degree of approximately 30% this means that three-spin contributions are only attenuated by a factor 1/6, compared to two-spin contributions.

If a labelled spin has two labelled neighbours, intra- or inter-chain, the signal function for the first spin is

$$S_1(t_1, t_2) = (F_{12}\cos(\Omega_1 t_1 + \Omega_2 t_1) + F_{13}\cos(\Omega_1 t_1 + \Omega_3 t_1) - F_{23}\cos(\Omega_2 t_1 + \Omega_3 t_1)) \times \exp(-i\Omega_1 t_2) \quad (7)$$

Table 1
Isotopic labelling degree of the silk samples in percentage with standard deviations given in brackets

Sample	[1-^{13}C]Ala	[1-^{13}C]Gly	[1-^{13}C]Pro	[1-^{13}C]Glu
[1-^{13}C]Ala, fibre	9.5 (6)	1.3 (5)	1.1 (5)	2.3 (39)
[1-^{13}C]Ala, film	15.5 (7)	1.2 (5)	1.5 (12)	3.0 (5)
[1-^{13}C]Gly, fibre	2.4 (4)	34.2 (19)	1.8 (4)	4.3 (21)
[1-^{13}C]Gly, film	1.5 (6)	26.2 (13)	2.2 (13)	—[a]

[a] Could not be determined.

with

$$F_{12} = \frac{16d_{12}^2\cos(D\tau_m)(d_{23}^2 + (d_{12}^2 + d_{13}^2)\cos(D\tau_m))\sin^2(D\tau_m)}{D^4},$$
$$F_{13} = \frac{16d_{13}^2\cos(D\tau_m)(d_{23}^2 + (d_{12}^2 + d_{13}^2)\cos(D\tau_m))\sin^2(D\tau_m)}{D^4},$$
$$F_{23} = \frac{128(d_{12}^2 + d_{13}^2)d_{23}^2\cos^2(\tfrac{1}{2}D\tau_m)\cos(D\tau_m)\sin^4(\tfrac{1}{2}D\tau_m)}{D^4}, \quad (8)$$

and $D = \frac{1}{2}\sqrt{d_{12}^2 + d_{13}^2 + d_{23}^2}$. The first two terms originate from directly coupled spins, but there is also a relayed contribution, which is also present if either d_{12} or d_{13} is zero. This gives increasingly negative signal intensities with longer excitation times τ_m. In the initial rate regime, though, the spectrum of a 3-spin system can be approximated by a sum of spectra from 2-spin systems. Fig. 6 shows simulated DOQSY spectra for the 2-spin and 3-spin situations, and also the absolute differences between these, using an excitation time of 1.0 ms. Differences may amount to some 10% of the maximum intensity for individual points, but in general the spectra are similar. At longer excitation times the differences are more pronounced. In practice τ_m was determined by the S/N ratio obtained for the sample of interest. For the alanine-labelled samples, an excitation time of 1.5 ms was used, which resulted in measuring times of 7 days, an excitation time τ_m of 1.0 ms would have resulted in unreasonably long measurement times in our case. For the glycine-labelled samples it was possible to use $\tau_m = 1.0$ ms. A small amount of multi-spin contributions will thus be present in all experimental spectra. Fits of simulated spectra, with a two-dimensional Gaussian distribution function with FWHH = 10°, a 1:1 mixture of 2 and 3-spin contributions, $\tau_m = 1.5$ ms excitation time and similar torsion angles as found in Fig. 2, did not show any significant effects due to these small spectral distortions, though (data not shown).

2.5. Peptide geometry and symmetries

Ideally, the DOQSY spectrum is a sum of signals arising from two coupled carbonyl tensors neighbouring in the primary structure of the protein, thus reflecting the relative orientation of the two tensors. This can be expressed by three Euler angles (α, β, γ). Assuming one knows the molecular geometry of the peptide unit, this can also be expressed as a function of the torsion angles (ϕ,ψ, ω). The angle ω is the torsion angle around the C–N bond in the peptide plane

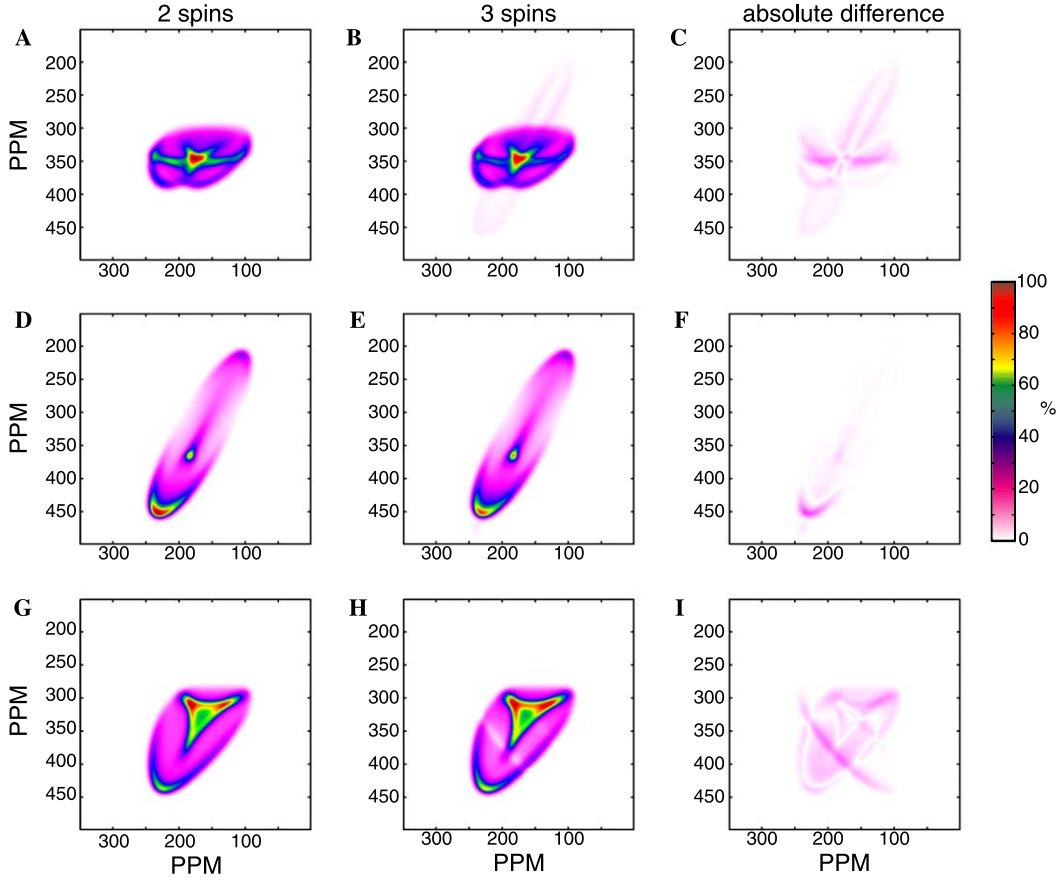

Fig. 6. Simulated 2D DOQSY spectra showing the influence of multiple-spin contacts for an excitation time of 1.0 ms. For three conformations (ϕ,ψ) = (A–C) (−57°, −47°), (D–F) (−150°, 150°), and (G–I) (−75°, 150°) the 2-spin, 3-spin and absolute difference spectra between these two are shown. A three-spin system was made using a fictitious tri-peptide with twice the same dihedral angles. For each conformation all three spectra are plotted using the same colormap allowing a direct comparison. Only positive intensities are shown (for the 3-spin case small negative intensities arise with longer excitation times).

and is usually assumed to be quite close to 180° and constant over the (ϕ,ψ) space. A search through 11876 structures of proteins and enzymes, determined by X-ray (release #95 of the Protein Data Bank), showed that the distribution in ω can be approximately described by a Lorentzian, centred around 180.1° and with a full width at half height (FWHH) of approximately 2.8° (not shown). In this work, a standard peptide geometry as defined by Pauling [52] was used and ω was fixed at 180°. The error due to this will be negligible, within the 15° grid resolution of the kernel.

It is known that a very small percentage of the peptide bonds in nature is found in the *cis*-conformation ($\omega = 0°$) [53,54]. For X–Y torsion angles (X and Y are not Pro), the occurrence is <0.05%, whereas 5% of all X-Pro fragments are found in the *cis*-conformation [54]. Furthermore, in most cases a *cis*-peptide is found close to active sites in proteins. In the following only torsion angles of Ala-Ala and Gly-Gly pairs are measured and the proteins are all assumed to be in *trans* configuration.

Analysis of the DOQSY spectra in terms of (ϕ,ψ), instead of (α, β, γ), has the advantage of considerably reducing the conformational space. In effect, the number of symmetry-related conformations is reduced from 16 [55] to 2, $S(\phi,\psi) = S(-\phi,-\psi)$. It must be noted, though, that in some regions of the Ramachandran plot $S(\psi,\phi)$ can be very similar to $S(\phi,\psi)$, as may also be deduced from Fig. 4.

Using additional information, the symmetry $S(\phi,\psi) = S(-\phi,-\psi)$ may be broken, in some cases. For [1-^{13}C] alanine-labelled samples it is possible to include the C_α chemical shift into the kernel. The chemical-shift tensors of C_α carbons of many amino acids have been calculated, as a function of the torsion angles (ϕ,ψ) [56–59]. Those values can be implemented directly into the fit by adding the penalty $\lambda\cdot\|\delta_{\phi,\psi} - \delta_{\mathrm{Exp}}\|^n$ to the first point of each basis spectrum. From the natural-abundance C_α peaks in the MAS spectrum the chemical shifts were determined to be 49.2 and 53.1 ppm for the fibre and film samples. Fig. 7 shows 1D spectra with and without application of a DREAM spin pair filter experiment [60] on the alanine-labelled fibrous sample. Considering the low labelling degree of the sample (Table 1), this can be interpreted as indicative for a single resonance in the C_α range for alanine, justifying the use of a single value δ_{Exp} for the bias function. For our purposes n was set to 1, but could in principle be varied to allow optimal distinction between conformations. For selected cases the chemical-shift bias can thus simplify the distribution considerably. In other cases though, e.g., in specific regions for certain amino acids or for the a chiral glycine residue, this bias cannot work and other experimental data would have to be included into the fit instead.

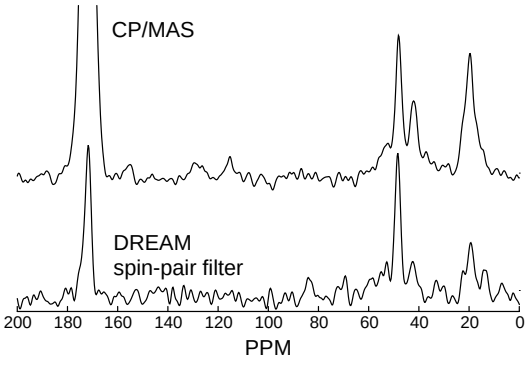

Fig. 7. 1D CP/MAS spectra of the [1-^{13}C]alanine-labelled fibrous sample (upper trace) and the corresponding spectrum when a DREAM spin-pair filter [60] was applied (lower trace). A single alanine C_α resonance is found at 49.2 ppm. The (attenuated) resonances in the C_α region are attributed to an artefact of the low labelling degree of only 8%, allowing these natural abundance contacts to show.

2.6. CSA tensor (principal values and orientation of the principal axis)

An accurate determination of the relative orientation of two carbonyl tensors, in terms of the backbone torsion angles (ϕ,ψ), requires the knowledge of both the magnitude and orientation of the carbonyl chemical-shift tensors in the molecule. From model peptide studies the orientation is rather well known. The δ_{33} component is perpendicular to the peptide plane and the δ_{22} component is almost along the C=O bond, with a small deviation away from the C–N bond. This deviation was, in model compounds, determined to be between 0° and 10° [61–67]. In absence of reliable methods to determine the tensor orientations in the silk, an average orientation of the tensors was assumed, with a 5° deviation of the δ_{22} component from the C=O bond.

All basis spectra have been calculated using the principal values shown in Table 2, which were obtained from fits of static 1D patterns of the respective unoriented silk samples. The tensor was thus assumed independent of the secondary structure. It is known, however, that the δ_{22} component may shift by more than 10 ppm as a function of hydrogen bonding strength [68,69]. Even though this is not very large compared to the inhomogeneous linewidth, in silk a 6.5 ppm wide line (FWHH) is typically observed for the ^{13}C=O isotropic chemical shift under MAS, a single set of tensor values may give rise to artefacts in the PDF. Fig. 8 shows the effect of fitting experimental spectra with incorrect tensor values. The experimentally determined difference in the δ_{22} component for the fibre and film samples is almost 15 ppm and these two sets of tensor values are assumed to reflect the maximum difference in ^{13}C=O chem-

Table 2
Principal values for the chemical 1-^{13}C shielding tensors as obtained from non-linear fits of the 1D patterns of unoriented samples, fitting a single set of tensor values only. The accuracy is estimated to be ±2 ppm

Sample	δ_{11} (ppm)	δ_{22} (ppm)	δ_{33} (ppm)
[1-^{13}C]Ala, fibre	242.0	181.3	92.0
[1-^{13}C]Ala, film	244.8	194.7	91.6
[1-^{13}C]Gly, fibre	242.0	178.3	94.3
[1-^{13}C]Gly, film	244.3	182.8	92.1

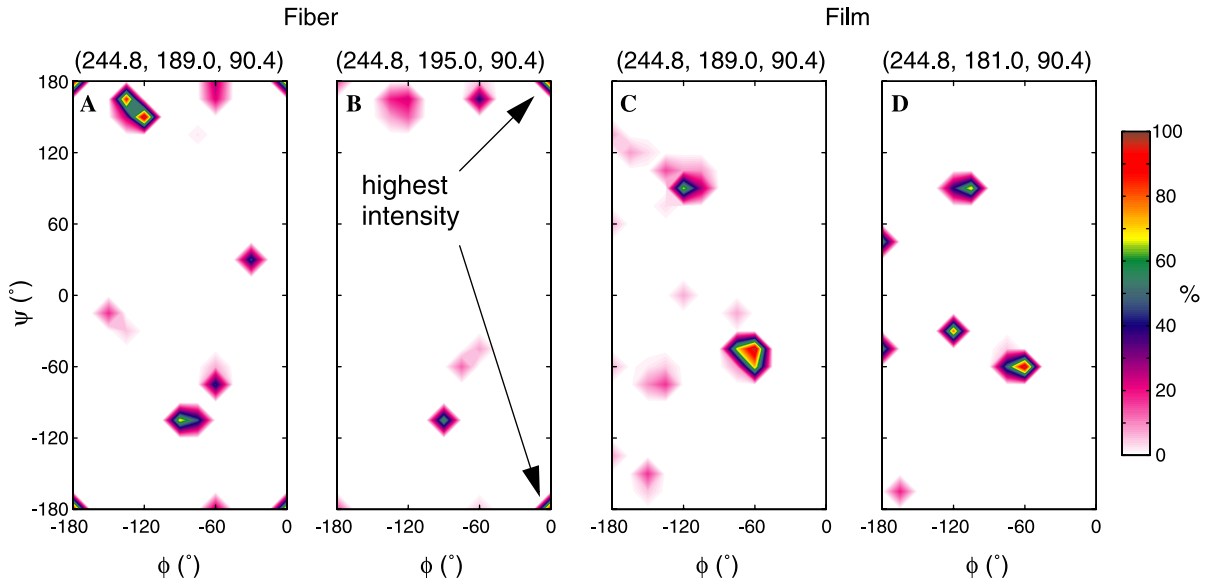

Fig. 8. Results of fitting the experimental DOQSY spectra of Figs. 2A and B with different chemical shift values. The values are plotted above each figure. Mainly the δ_{22} component was changed by (A) +7.7 ppm and (B) +13.7 ppm for the fibrous sample and (C) −5.7 ppm and (D) −13.7 ppm for the film. The DOQSY spectrum for α-helical conformations is very distinct, and therefore the PDFs of the film, which is mostly α-helical, are less affected by variations of the principal values than are those of the fibre.

ical-shift tensor in proteins. The δ_{11} and δ_{33} components were taken similar to the values determined from experimental 1D spectra, and the δ_{22} component was changed, for each sample, in direction of the value for the other sample. By comparing with Fig. 2, it shows that with increasing deviation of the δ_{22} value the discrepancy between spectrum and fit increased and a multitude of sharp highly correlated intensities appeared in the PDF, often in energetically unfavourable regions of the Ramachandran map. The spectrum of the film sample could be fitted using the intermediate value for δ_{22}, whereas the fit for the fibre already showed significant deviations (Figs. 8A and B). At the highest deviation of the δ_{22} tensor value neither spectrum could be reasonably explained by the kernel and the PDFs no longer reflected the underlying torsion angles.

In particular for broad or multi-peaked distributions of torsion angles, this approximation can be a significant source of systematic errors. The situation could potentially be improved by including the (ϕ,ψ)-dependence of the chemical-shift tensor into the analysis, but this requires better knowledge of the exact effects of conformation and hydrogen bonding. Still, the CSA values of Table 2 were used without a posteriori optimization to fit the DOQSY spectra. From the fact that both PDFs for the [1-^{13}C]alanine-labelled samples show sharp distribution functions, without dominating intensities in forbidden regions, it is concluded that the single-tensor approach is acceptable. The glycine-labelled samples show more complex distribution functions, with many small peaks that shift in position for slightly different fitting conditions. Consequently, the grid resolution of the analysis had to be lowered. The position of the highest intensities were stable, though, and the interpretation of the PDFs was based on those intensities.

2.7. Rf performance

For an ideal spectrum (e.g., using δ pulses) Eq. (6) can be used to directly calculate the basis spectra in the frequency-domain, as this is much faster than the corresponding time-domain calculation of the 2D spectrum. It was found that a frequency-domain calculation is also possible, in good approximation, for finite pulses if the rf excitation profile for the pulse sequence is calculated separately and multiplied a posteriori into the "ideal" basis spectra. Fig. 9 shows the offset-dependencies and their influence on the DQ excitation efficiency across the 2D DOQSY pattern for pulses with 50 kHz rf field amplitude and $\tau_m = 1.0$ ms or $\tau_m = 1.5$ ms. Despite the fact that the excitation sequence is well compensated and broadbanded [49], comparable to equivalent sequences under MAS [70], the excitation profiles indicate that the effects on the spectrum can be significant. In fact, for the glycine-labelled samples it was preferable to measure the spectra at a magnetic field of 7.0 T because the analysis for the spectra taken at 9.4 T, with a correspondingly broader CSA tensor, was not stable. Figs. 9C and E show the marked improvement of the excitation profile upon lowering the magnetic field. This instability was not observed for sharply defined distribution functions, i.e., the alanine-labelled samples, but only for the glycine-labelled samples. Fig. 10 shows the effect of including the finite rf power effects in the kernel on the resulting PDF for the alanine-labelled samples. Within experimental error the differences in the PDFs are negligible, although a significant decrease in the discrepancy between spectrum and fit was observed, upon inclusion of the excitation profile in the frequency-domain simulated kernel. No further decrease was observed for the kernel cal-

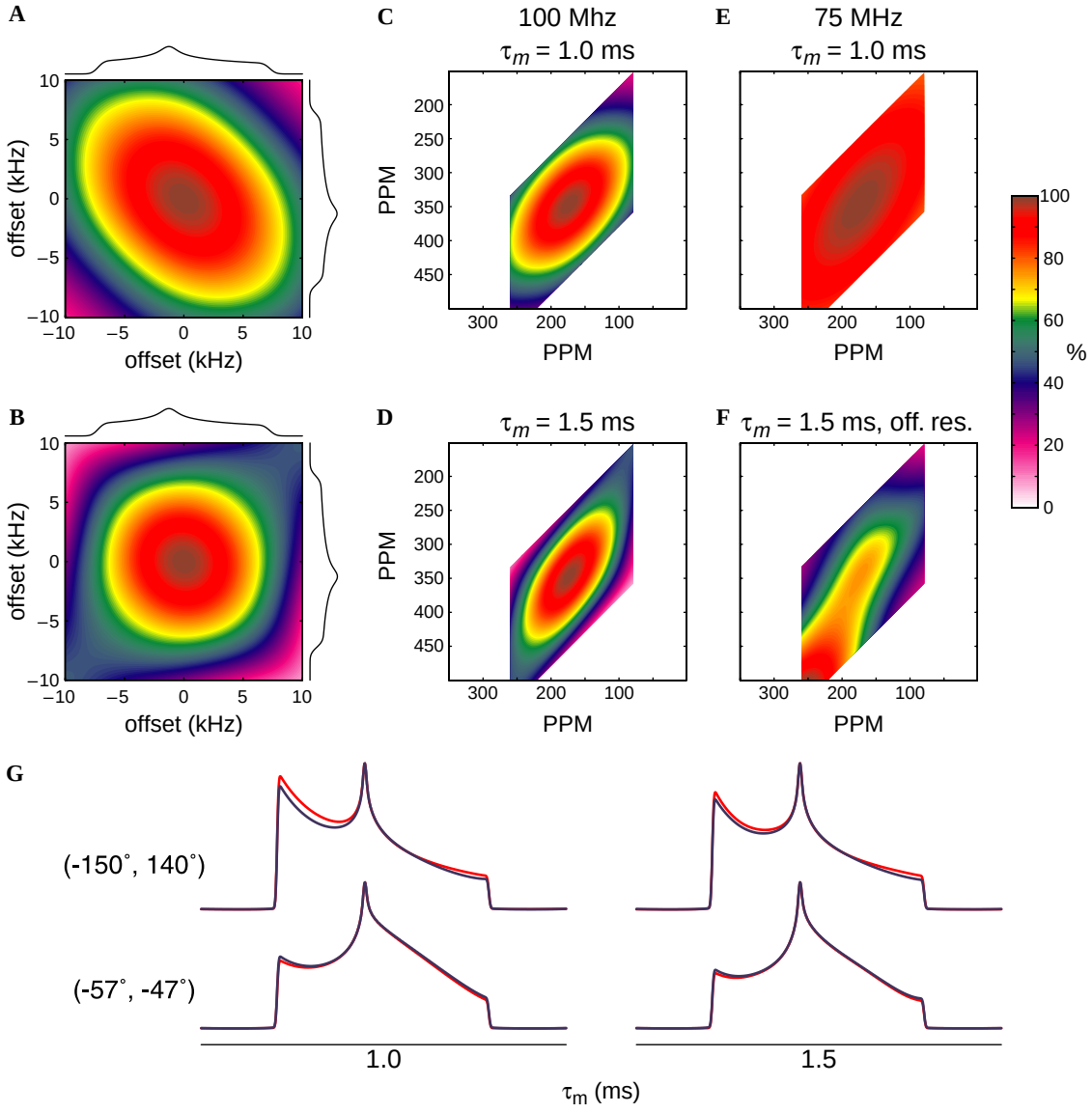

Fig. 9. Simulated DQ efficiency for the pulse sequence of Fig. 1B for an rf field strength of 50 kHz and an excitation cycle time of (A) 1.0 ms and (B) 1.5 ms and a dipolar coupling of 150 Hz. The plots are normalized to their respective maxima. Actual scaling of the theoretical DOQSY spectrum measured at a ^{13}C resonance frequency of 100 MHz, when calculated in frequency domain using Eq. (6), for (C) 1.0 ms and (D) 1.5 ms excitation time. (E) The excitation profile of the sequence for a ^{13}C resonance frequency of 75 MHz and 1.0 ms DQ excitation time. (F) The spectrum in Fig. 2B was measured with off-resonance irradiation and the excitation profile is attenuated accordingly. (G) Simulated one-dimensional SQ → DQ → SQ spectra for $(\phi,\psi) = (-150°, 140°)$ (top curves) and $(-57°, -47°)$ (bottom curves) for rf field strengths of 250 kHz (red curves) and 50 kHz (blue curves) to illustrate the effect of the scaling. All blue curves are normalized to the maximum of the corresponding red curves. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

culated in the time domain. In selected cases, even off-resonance irradiation need not be fatal even though the excitation profile is very pronounced indeed (Fig. 9F). The spectrum shown in Fig. 2B was recorded in this fashion. Due to the fact that the experiment is very distinctive around the α-helical conformation, in the sense that the basis spectra of the α-helix show very little correlation with other basis spectra, the analysis was stable and there was no reason to remeasure the spectrum on-resonance.

From these results, we conclude that the scaling of the basis spectra in the frequency domain may be used in excellent approximation to calculate DOQSY spectra in the finite rf-amplitude case.

2.8. Broad distributions

Naturally, the resolution of the regulatory approach used in this work is expected to decrease, if the complexity of the distribution function underlying a spectrum increases. The random coil structure of a protein is an example of a complex distribution. We define a random-coil structure by a PDF, which is determined solely by the energy surface

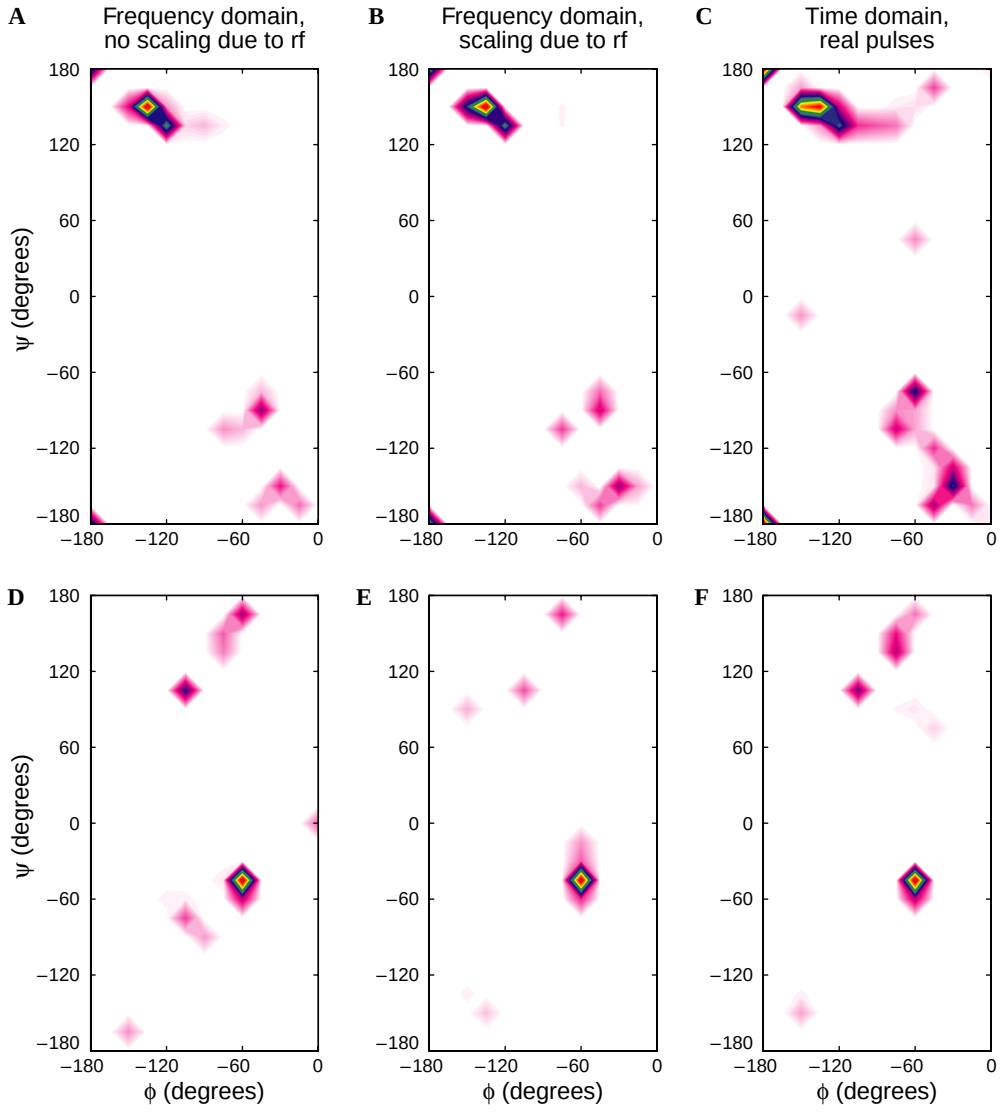

Fig. 10. Contrary to the glycine-labelled samples, the analysis of the spectra for the alanine-labelled samples was very stable towards the rf performance of the pulse sequence. (A–C) The distributions obtained for the fibrous sample when not correcting the basis spectra for the excitation profile, including the correction in the frequency domain or calculating the kernel in time domain with realistic pulses. (D–F) The respective results obtained for the film sample.

$E(\phi,\psi)$ of the amino acid residue via the Boltzmann distribution [71]. Based on the energy surface for glycine given in [71], an approximate random coil distribution function was created and is shown in Fig. 11A. The corresponding DOQSY spectrum was simulated and line broadening, similar to that observed for experimental spectra for different silks (15 ppm), and noise were added. The algorithm is clearly able to reconstruct the PDF. Furthermore, the reconstruction worked for a wide range of line broadening (not shown), when neglecting the (ϕ,ψ)-dependence of the CSA.

2.9. Relaxation

Relaxation could in principle distort the experimental DOQSY spectrum. From the experimental data presented here and in previous work [9,13], it seems unlikely though that relaxation plays an important role. Certainly, fast relaxing components, giving rise to significant distortions from the simulations, are not present.

3. Experimental section

3.1. Silk

Isotope-labelled silk fibroin was obtained from Prof. Tetsuo Asakura and was prepared according to the procedure described in [13]. The $^{13}C{=}O$ enrichment of the samples was checked using mass spectroscopy [51].

3.2. NMR spectra

NMR spectra were obtained at room temperature on a Bruker DMX 400 and on a Varian Infinity+ 300 spectrometer. The pulse sequence used (Fig. 1B) was comparable to the one used by Schmidt-Rohr [43], except for the DQ exci-

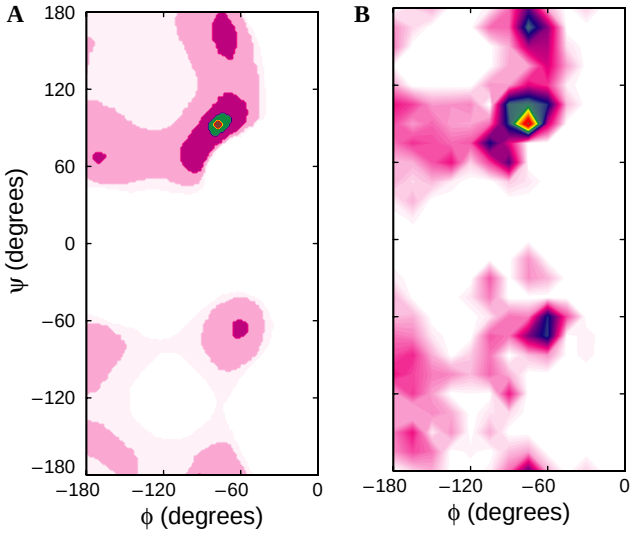

Fig. 11. The limits of the algorithm, with respect to complex distribution function, were explored using (A) an approximate random-coil distribution, generated from the energy surface of glycine, given in [71]. The resulting fit (B) is in excellent agreement with the input distribution.

tation for which the modification by Antzukin et al. [49] of the method by Yen and Pines [48] was used. Radio-frequency field strengths of 50 kHz were applied, except for decoupling at 100 kHz. The cross-polarization contact time was 1.7 ms, the z-filter time $\tau = 1$ ms, the recycle delay 4 s. A single excitation cycle of 1.5 or 1.0 ms was used to generate DQ coherence. Phase sensitivity in the indirect dimension was achieved by TPPI in all spectra, except the alanine-labelled film sample (Fig. 2B) where the rf carrier was positioned outside the spectrum (32 ppm or 3.2 kHz from the δ_{11} tensor component). The cable length between probe and pre-amplifier was carefully selected to minimize phase transients and achieve constant performance over a wide range of carbon field strengths (30–70 kHz). The total measuring time per DOQSY spectrum was approximately 7 days for samples of approximately 30 mg. The ppm scale was calibrated using external adamantane as a reference [72]. All spectra were processed using the matNMR processing package [73].

The $^{13}C_\alpha$ chemical shifts were determined for the [1-^{13}C]alanine-labelled samples by applying a DREAM spin pair filter experiment [60] at a MAS frequency of 25 kHz.

3.3. Analysis

NMR spectra were calculated using the GAMMA C++ simulation environment [74]. For frequency space simulations the frequencies for a given crystallite orientation were determined by rotation of the spatial tensors, extracting the frequencies, and combining them into the spectral matrix. Spectra calculated in the frequency domain had a frequency resolution of 1 ppm/point and were corrected a posteriori, corresponding to the excitation profile of the DOQSY experiment, described in detail below, unless explicitly mentioned otherwise. Time-domain calculations included all delays and pulses as executed in the actual experiment. A full set of basis spectra used in the analysis is referred to as "kernel."

The grid resolution in (ϕ,ψ) was 15°, 15°, 18°, and 20° for the PDFs of [1-^{13}C]alanine alanine fibre, [1-^{13}C]alanine-labelled film, [1-^{13}C]glycine-labelled fibre, [1-^{13}C]glycine-labelled film sample, respectively. Each interval, defined by four corner points on the parameter grid, was linearly interpolated using splines, with the points of this subgrid defined by a two-point Gaussian quadrature integration scheme. For the PDFs of the four samples, each interval was divided in 3, 3, 4, and 4 sub intervals (of two points each), respectively. Powder averaging was done following the Cheng method [75] and 28,657, 28,657, 75,025, and 75,025 powder points were used for the four kernels. For the fits of the experimental spectra, linebroadening was essentially a free parameter and was optimized in coarse steps. For the four experiments, the Gaussian linebroadening corresponded to 16, 17, 14, and 14 ppm, respectively. The differences in linebroadening are mainly caused by the different number of powder points used for simulation of the basis spectra.

4. Results and discussion

Fig. 2 shows experimental 2D DOQSY spectra for the fibrous and the film samples, labelled either with [1-^{13}C]alanine or [1-^{13}C]glycine, and the corresponding fits and probability distribution functions (PDFs). Significant differences are observed between the spectra of the fibrous and film samples, and for the two domains. The alanine spectra are the same as described before [13] and the small differences between the alanine PDFs presented here and the ones published before are due to minor changes in the fitting procedure. Table 2 shows the principal values for all four silk samples as used in the fits, which were determined from static 1D NMR spectra of unoriented samples. The values for the fibrous samples are very similar to those observed for fibrous spider dragline silk [9], suggesting similar secondary structures. The values for the [1-^{13}C]alanine-labelled film sample are clearly different from those in the fibrous state, though.

As described above, the alanine-C_α isotropic chemical shift of the carbon was used to break the $S(\phi,\psi) = S(-\phi, -\psi)$ symmetry. An appropriate value for the isotropic-shift penalty parameter λ was determined by evaluating the effects of increasing its amplitude. For the fibrous sample a very small λ, causing an increase in squared deviation in the order of 0.1% only, was enough to break the symmetry. In the case of the film, the bias had no effect, since the isotropic shift values cannot distinguish between the left- and right handed α-helical conformations [56], and was therefore not included in the fit, nor were the corresponding values for glycine-labelled samples.

The alanine-rich domains are predominantly found in β-sheet conformation in the fibre and in an α-helical confor-

mation in the film sample. The PDFs of Fig. 2 show narrow distributions with integrated intensities, in the designated regions, of 70 and 60% for the fibrous and film sample, respectively. The highest probability for the fibrous sample was found for (−135°, 150°), corresponding to the β-sheet region. This is in accordance with previous 1D NMR measurements, which were evaluated in terms of a single structure and where both the information on the chemical shift and spectra from oriented silk were best explained by a structure in the β-sheet range of the Ramachandran plot [17,18]. The film sample shows a sharp maximum at (±60°, ±45°), which is close to idealized values for α-helices (±57°, ±47°), and is in accordance with previous findings [13,17,22]. Similar values have also been reported recently for model peptides that were based on the *S. c. ricini* silk [76]. Also, liquid-state NMR experiments of dissolved silk have indicated that 70% of all alanine-rich domains reside in an α-helical state in aqueous solution [14,15].

In contrast to the narrow distributions observed for the alanine-labelled samples, the PDFs from the glycine-labelled samples exhibit several maxima across the Ramachandran map. The glycine-rich domains in the fibrous sample nevertheless seem to have a defined local structure, resembling that observed for spider dragline silk [9]. From static 1D NMR spectra of oriented silk fibre Asakura et al. [18] have argued that the glycine-rich domains are predominantly in a β-sheet conformation. There it was predicted that the torsion angles are most likely between −140° and −125° for ϕ and between 131° and 148° for ψ, at variance with our data. In contrast, for spider dragline silk a partial β-sheet and partial 3_1-helical arrangement was proposed [9,77]. We interpret our data for *S. c. ricini* as indicative of a comparable structure as observed for spider silk, which is also reflected by the chemical-shift tensor values. Even though the differences between the PDFs of these two types of silk are apparent, they are within the experimental error of the analysis, considering the high complexity of the distribution functions, the relatively low resolution that could be used in the analysis and the systematic errors associated with using a single set of chemical-shift tensor values. The glycine-rich domains in the film state show a PDF corresponding to a significantly disordered system and can be explained, within experimental error, by a random-coil structure. The alanine as well as the glycine torsion-angle distributions are reminiscent of the structure found for dissolved silk [18].

5. Conclusions

A detailed analysis of the DOQSY approach for determining torsion-angle distributions in silks was presented. Using this approach, we have investigated the backbone structure of the silk from the silkworm *S. c. ricini*, in its fibrous state as well as in a cast film. Previous results from indirect 1D NMR data had indicated β-sheet and α-helical conformations for the alanine-rich domains for the fibrous and film sample, respectively. The glycine-rich domains in the fibre were described to be in a β-sheet conformation. The results presented here have extended these results by explicitly measuring the distribution of the backbone torsion angles. The previously published results for the alanine-rich domains were confirmed for both the fibre and the film state. The glycine-rich domains in fibrous silk were found to be in a multipeaked distribution function, partially in β-sheet and partially in 3_1-helical conformation. In the cast film they were found in a broad distribution, within experimental error describing a random coil.

We believe that the method has a considerable potential to be applied to other fibrous proteins and polymers. Although the resolution of this approach is not very high, the method is robust towards various potential sources of error, and first in its kind to provide detailed structural information in non-crystalline heterogeneous proteins, like silk, which are generally difficult to address. The method works best for relatively narrow distributions. Broad distributions are clearly recognized as such but the detailed characterization is less unambiguous. We expect that taking into account the torsion-angle dependence of the CSA principle values using quantum-chemical calculations would improve the performance for broad distributions. The method described here is not limited to pairs of the same amino acid, though, and should work with a modified labelling strategy for any pair of residues, provided the principal values of both chemical shielding tensors are known.

Acknowledgments

This research has been supported by the Swiss National Science Foundation and by the European Science Foundation through the Network: Silk, Properties and Production. T. Asakura is gratefully acknowledged for providing the silk samples.

References

[1] M.B. Hinman, J.A. Jones, R.V. Lewis, Synthetic spider silk: a modular fiber, TIBTECH 18 (2000) 374–379.
[2] K. Mita, S. Ichimura, T.C. James, Highly repetitive structure and its organization of the silk fibroin gene, J. Mol. Evol. 38 (1994) 583–592.
[3] C.Z. Zhou, F. Confalonieri, N. Medina, Y. Zivanovic, C. Esnault, T. Yang, M. Jacquet, J. Janin, M. Duguet, R. Perasso, Z.G. Li, Fine organization of *Bombyx mori* fibroin heavy chain gene, Nucleic Acids Res. 28 (2000) 2413–2419.
[4] K. Yukuhiro, T. Kanda, T. Tamura, Preferential codon usage and two types of repetitive motifs in the fibroin gene of the chinese oak silkworm, *Antheraea pernyi*, Insect Mol. Biol. 6 (1997) 89–95.
[5] M. Xu, R.V. Lewis, Structure of a protein superfiber: spider dragline silk, Proc. Natl. Acad. Sci. USA 87 (1990) 7120–7124.
[6] M.B. Hinman, R.V. Lewis, Isolation of a clone encoding a second dragline silk fibroin, J. Biol. Chem. 267 (1992) 19320–19324.
[7] P.A. Guerette, D.G. Ginzinger, B.H.F. Weber, J.M. Gosline, Silk properties determined by gland-specific expression of a spider fibroin gene family, Science 272 (1996) 112–114.
[8] C.Y. Hayashi, R.V. Lewis, Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks, J. Mol. Biol. 275 (1998) 773–784.

[9] J.D. van Beek, S. Hess, F. Vollrath, B.H. Meier, The molecular structure of spider dragline silk, Proc. Natl. Acad. Sci. USA 99 (2002) 10266–10271.
[10] C. Zhao, T. Asakura, Structure of silk studied with NMR, Progr. Nucleic Magn. Res. Spectrosc. 39 (2001) 301–352.
[11] T. Asakura, personal communication.
[12] D. Wilson, R. Valuzzi, D. Kaplan, Conformational transitions in model silk peptides, Biophys. J. 78 (2000) 2690–2701.
[13] J.D. van Beek, L. Beaulieu, H. Schäfer, M. Demura, T. Asakura, B.H. Meier, Solid-state NMR determination of the secondary structure of *Samia cynthia ricini* silk, Nature 405 (2000) 1077–1079.
[14] T. Asakura, T. Murakami, NMR of silk fibroin. 4. Temperature- and urea-induced helix-coil transitions of the -(ala)$_n$- sequence in *Philosamia cynthia ricini* silk fibroin protein monitored by carbon-13 NMR spectroscopy, Macromolecules 18 (1985) 2614–2619.
[15] T. Asakura, H. Kashiba, H. Yoshimizu, NMR of silk fibroin. 8. Carbon-13 NMR analysis of the conformation and the conformational transition of *Philosamia cynthia ricini* silk fibroin protein on the basis of bixon-scheraga-lifson theory, Macromolecules 21 (1988) 644–648.
[16] H. Saito, M. Ishida, M. Yokoi, T. Asakura, Dynamic features of side chains in tyrosine and serine residues of some polypeptides and fibroins in the solid as studied by high-resolution solid-state carbon-13 NMR spectroscopy, Macromolecules 23 (1990) 83–88.
[17] M. Ishida, T. Asakura, M. Yokoi, H. Saito, Solvent- and mechanical-treatment-induced conformational transition of silk fibroins studies by high-resolution solid-state carbon-13 NMR spectroscopy, Macromolecules 23 (1990) 88–94.
[18] T. Asakura, T. Ito, M. Okudaira, T. Kameda, Structure of alanine and glycine residues of *Samia cynthia ricini* silk fibers studied with solid-state ^{15}N and ^{13}C NMR, Macromolecules 32 (1999) 4940–4946.
[19] T. Asakura, M. Iwadate, M. Demura, M.P. Williamson, Structural analysis of silk with ^{13}C NMR chemical shift contour plots, Int. J. Biol. Macromol. 24 (1999) 167–171.
[20] T. Kameda, Y. Ohkawa, K. Yoshizawa, J. Naito, A.S. Ulrich, T. Asakura, Hydrogen-bonding structure of serine side chains in *Bombyx mori* and *Samia cynthia ricini* silk fibroin determined by solid-state 2H NMR, Macromolecules 32 (1999) 7166–7171.
[21] T. Kameda, Y. Ohkawa, K. Yoshizawa, E. Nakano, T. Hiraoki, A.S. Ulrich, T. Asakura, Dynamics of the tyrosine side chain in *Bombyx mori* and *Samia cynthia ricini* silk fibroin studied by solid state 2H NMR, Macromolecules 32 (1999) 8491–8495.
[22] Y. Nakazawa, T. Nakai, T. Kameda, T. Asakura, A ^{13}C NMR study on the structural change of silk fibroin from *Samia cynthia ricini*, Chem. Phys. Lett. 311 (1999) 362–366.
[23] Y. Nakazawa, T. Asakura, Heterogeneous exchange behaviour of *Samia cynthia ricini* silk fibroin during helix-coil transition studied with ^{13}C NMR, FEBS Lett. 529 (2002) 188–192.
[24] Y. Nakazawa, M. Bamba, S. Nishio, T. Asakura, Tightly winding structure of sequential model peptide for repeated helical region in *Samia cynthia ricini* silk fibroin studied with solid-state NMR, Prot. Sci. 12 (2003) 666–671.
[25] P.M. Henrichs, Molecular-orientation and structure in solid polymers with ^{13}C NMR—a study of biaxial films of poly(ethylene-terephthalate), Macromolecules 20 (1987) 2099–2112.
[26] K. Schmidt-Rohr, M. Hehn, D. Schäfer, H.W. Spiess, Two-dimensional nuclear magnetic resonance with sample flip for characterizing orientation distributions, and its analogy to X-ray scattering, J. Chem. Phys. 97 (1992) 2247–2262.
[27] B.H. Chmelka, K. Schmidt-Rohr, H.W. Spiess, Molecular orientation distributions in PET thin films and fibers from multidimensional DECODER NMR spectroscopy, Macromolecules 26 (1993) 2282–2296.
[28] P.T. Eles, C.A. Michal, A decoder NMR study of backbone orientation in *Nephila clavipes* dragline silk under varying strain and draw rate, Biomacromolecules 5 (2004) 661–665.
[29] D.P. Weliky, R. Tycko, Determination of peptide conformations by two-dimensional magic angle spinning NMR exchange spectroscopy with rotor synchronization, J. Am. Chem. Soc. 118 (1996) 8487–8488.
[30] X. Feng, Y.K. Lee, D. Sandstrom, M. Edén, H. Maisel, A. Sebald, M.H. Levitt, Direct determination of a molecular torsional angle by solid-state NMR, Chem. Phys. Lett. 257 (1996) 314–320.
[31] P.R. Costa, J.D. Gross, M. Hong, R.G. Griffin, Solid-state NMR measurement of psi in peptides: A NCCN 2Q-heteronuclear local field experiment, Chem. Phys. Lett. 280 (1997) 95–103.
[32] M. Hong, J.D. Gross, R.G. Griffin, Site-resolved determination of peptide torsion angle phi from the relative orientation of backbone N–H and C–H bonds by solid-state NMR, J. Phys. Chem. B 101 (1997) 5869–5874.
[33] M. Hong, J.D. Gross, C.M. Rienstra, R.G. Griffin, K.K. Kumashiro, K. Schmidt-Rohr, Coupling amplification in 2D MAS NMR and its application to torsion angle determination in peptides, J. Magn. Res. 129 (1997) 85–92.
[34] M. Hong, J.D. Gross, W. Hu, R.G. Griffin, Determination of the peptide torsion angle phi by ^{15}N chemical shift and ^{13}C–^{1}H dipolar tensor correlation in solid-state mas NMR, J. Magn. Res. 135 (1998) 169–177.
[35] M. Edén, A. Brinkmann, H. Luthman, L. Eriksson, M.H. Levitt, Determination of molecular geometry by high-order multiple-quantum evolution in solid-state NMR, J. Magn. Res. 144 (2000) 266–279.
[36] D. Huster, S. Yamaguchi, M. Hong, Efficient beta-sheet identification in proteins by solid-state NMR spectroscopy, J. Am. Chem. Soc. 122 (2000) 11320–11327.
[37] I. Sack, Y.S. Balazs, S. Rahimipour, S. Vega, Solid-state NMR determination of peptide torsion angles: applications of 2H-dephased REDOR, J. Am. Chem. Soc. 122 (2000) 12263–12269.
[38] I. Sack, Y.S. Balazs, S. Rahimipour, S. Vega, Peptide torsion angle measurements: effects of nondilute spin pairs on carbon-observed, deuterium-dephased PM5-REDOR, J. Magn. Res. 148 (2001) 104–114.
[39] V. Ladizhansky, M. Veshtort, R.G. Griffin, NMR determination of the torsion angle psi in alpha-helical peptides and proteins: the HCCN dipolar correlation experiment, J. Magn. Res. 154 (2002) 317–324.
[40] P.M. Henrichs, M. Linder, Carbon-13 spin diffusion in the determination on intermolecular structure in solids, J. Magn. Res. 58 (1984) 458–461.
[41] R. Tycko, G. Dabbagh, A simple theory of C-13 nuclear spin diffusion in organic solids, Isr. J. Chem. 32 (1992) 179–184.
[42] P. Robyr, M. Tomaselli, J. Straka, C. Grob-Pisano, U.W. Suter, B.H. Meier, R.R. Ernst, Rf-driven and proton-driven NMR polarization transfer for investigating local order: an application to solid polymers, Mol. Phys. 84 (1995) 995–1020.
[43] K. Schmidt-Rohr, A double-quantum solid-state NMR technique for determining torsion angles in polymers, Macromolecules 29 (1996) 3975–3981.
[44] K. Schmidt-Rohr, W. Hu, N. Zumbulyadis, Elucidation of the chain conformation in a glassy polyester, pet, by two-dimensional NMR, Science 280 (1998) 714–717.
[45] M. Utz, Measurement of structural distribution functions in disordered systems: a general approach for sensitivity estimation, J. Chem. Phys. 109 (1998) 6110–6124.
[46] J.D. van Beek, B.H. Meier, H. Schäfer, Inverse methods in two-dimensional NMR spectral analysis, J. Magn. Res. 162 (2003) 141–157.
[47] W.S. Warren, D.P. Weitekamp, A. Pines, Theory of selective excitation of multiple-quantum transitions, J. Chem. Phys. 73 (1980) 2084–2099.
[48] Y.S. Yen, A. Pines, Multiple-quantum NMR in solids, J. Chem. Phys. 78 (1983) 3579–3582.
[49] O.N. Antzutkin, R. Tycko, High-order multiple quantum excitation in ^{13}C nuclear magnetic resonance spectroscopy of organic solids, J. Chem. Phys. 110 (1999) 2749–2752.
[50] B. Lotz, Beta structure of periodic copolypeptides of L-alanine and glycine, J. Mol. Biol. 87 (1974) 193–203.
[51] S. Hess, J.D. van Beek, L.K. Pannell, Acid hydrolysis of silk fibroins and determination of the enrichment of isotopically labeled amino

acids using precolumn derivatization and high-performance liquid chromatography-electrospray ionization-mass spectroscopy, Anal. Biochem. 311 (2002) 19–26.
[52] L. Pauling, R.B. Corey, H.R. Branson, The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain, Proc. Natl. Acad. Sci. USA 37 (1951) 205–211.
[53] M.S. Weiss, R. Hilgenfeld, A method to detect nonproline cis peptide bonds in proteins, Biopolymers 50 (1999) 536–544.
[54] A. Jabs, M.S. Weiss, R. Hilgenfeld, Non-proline cis peptide bonds in proteins, J. Mol. Biol. 286 (1999) 291–304.
[55] K. Schmidt-Rohr, H.W. Spiess, Multidimensional Solid-State NMR and Polymers, Academic Press, London, 1994.
[56] J. Heller, D.D. Laws, M. Tomaselli, D.S. King, D.E. Wemmer, A. Pines, R.H. Havlin, E. Oldfield, Determination of dihedral angles in peptides through experimental and theoretical studies of alpha-carbon chemical shielding tensors, J. Am. Chem. Soc. 119 (1997) 7827–7831.
[57] R.H. Havlin, H. Le, D.D. Laws, A.C. deDios, E. Oldfield, An ab initio quantum chemical investigation of carbon-13 NMR shielding tensors in glycine, alanine, valine, isoleucine, serine and threonine: comparisons between helical and sheet tensors, and the effects of chi1 on shielding, J. Am. Chem. Soc. 119 (1997) 11951–11958.
[58] R.H. Havlin, D.D. Laws, H.M.L. Bitter, L.K. Sanders, H. Sun, J.S. Grimley, D.E. Wemmer, A. Pines, E. Oldfield, An experimental and theoretical investigation of the chemical shielding tensors of $^{13}C_{\alpha}$ of alanine, valine and leucine residues in solid peptides and in proteins in solution, J. Am. Chem. Soc. 123 (2001) 10362–10369.
[59] H. Sun, L.K. Sanders, E. Oldfield, Carbon-13 NMR shielding in the twenty common amino acids: comparisons with experimental results in proteins, J. Am. Chem. Soc. 124 (2002) 5486–5495.
[60] R. Verel, M. Baldus, M. Ernst, B.H. Meier, A homonuclear spin-pair filter for solid-state NMR based on adiabatic-passage techniques, Chem. Phys. Lett. 287 (1998) 421–428.
[61] A. Naito, S. Ganapathy, K. Akasaka, C.A. McDowell, Chemical shielding tensor and ^{13}C–^{14}N dipolar splitting in single crystals of L-alanine, J. Chem. Phys. 74 (1981) 3190–3197.
[62] R.A. Haberkorn, R.E. Stark, H. van Willigen, R.G. Griffin, Determination of bond distances and bond angles by solid-state nuclear magnetic resonance. ^{13}C and ^{14}N NMR study of glycine, J. Am. Chem. Soc. 103 (1981) 2534–2539.
[63] R.E. Stark, L.W. Jelinski, D.J. Ruben, D.A. Torchia, R.G. Griffin, Carbon-13 chemical shift and carbon-13/nitrogen-15 dipolar tensors for the peptide bond: 1-^{13}C-glycyl-^{15}N-glycine hydrochloride hydrate, J. Magn. Res. 55 (1983) 266–273.
[64] T.G. Oas, C.J. Hartzell, F.W. Dahlquist, G.P. Drobny, The amide nitrogen-15 chemical shift tensors of four peptides determined from carbon-13 dipole-coupled chemical shift powder patterns, J. Am. Chem. Soc. 109 (1987) 5962–5966.
[65] T.G. Oas, C.J. Hartzell, T.J. McMahon, G.P. Drobny, F.W. Dahlquist, The carbonyl carbon-13 chemical shift tensors of five peptides determined from nitrogen-15 dipole-coupled chemical shift powder patterns., J. Am. Chem. Soc. 109 (1987) 5956–5962.
[66] Q. Teng, M. Iqbal, T.A. Cross, Determination of the carbon-13 chemical shift and nitrogen-14 electric field gradient tensor orientations with respect to the molecular frame in a polypeptide., J. Am. Chem. Soc. 114 (1992) 5312–5321.
[67] N. Takeda, S. Kuroki, H. Kurosu, I. Ando, ^{13}C NMR chemical shift tensor and hydrogen-bonded structure of glycine-containing peptides in a single crystal, Biopolymers 50 (1999) 61–69.
[68] T. Kameda, I. Ando, The relationship between the helical conformation and ^{13}C NMR chemical shift of amino acid residue carbonyl carbons of polypeptides in the solid state, J. Mol. Struct. 412 (1997) 197–203.
[69] G. Zheng, L. Wang, J. Hu, X. Zhang, L. Shen, C. Ye, G.A. Webb, Hydrogen bonding effects on the ^{13}C NMR chemical shift tensors of some amino acids in the solid state, Magn. Res. Chem. 35 (1997) 606–608.
[70] M. Baldus, D.G. Geurts, B.H. Meier, Broadband dipolar recoupling in rotating solids: a numerical comparison of some pulse schemes, Solid State NMR 11 (1998) 157–168.
[71] P.J. Flory, Statistical Mechanics of Chain Molecules, John Wiley, New York, 1969.
[72] W.L. Earl, D.L. VanDerHart, Measurement of ^{13}C chemical shifts in solids, J. Magn. Res. 48 (1982) 35–54.
[73] matNMR is a toolbox for processing NMR/EPR data in Matlab and can be downloaded freely at <http://matNMR.sourceforge.net>.
[74] S. Smith, T. Levante, B.H. Meier, R.R. Ernst, Computer simulations in magnetic resonance: an object oriented programming approach, J. Magn. Res. A 106 (1994) 75–105.
[75] V.B. Cheng, J.H.H. Suzakawa, M. Wolfsberg, Investigations of a nonrandom numerical method for multidimensional integration, J. Chem. Phys. 59 (1973) 3992–3999.
[76] Y. Nakazawa, T. Asakura, Structure determination of a peptide model of the repeated helical domain in *Samia cynthia ricini* silk fibroin before spinning by a combination of advanced solid-state NMR methods, J. Am. Chem. Soc. 125 (2003) 7230–7237.
[77] J. Kümmerlen, J.D. van Beek, F. Vollrath, B.H. Meier, Local structure in spider dragline silk investigated by two-dimensional spin-diffusion nuclear magnetic resonance, Macromolecules 29 (1996) 2920–2928.