Polymer Bulletin 9, 284–291 (1983)

Polymer Bulletin

Secondary Structure of Peptides
9. ^{13}C NMR CP/MAS Investigation of Silk Proteins

Hans R. Kricheldorf[1], Detlef Müller[2] and Klaus Ziegler[3]

[1] Institut für Angewandte Chemie der Universität, Martin-Luther-King-Platz 6, D-2000 Hamburg 13, Federal Republic of Germany

[2] Bruker Analytische Meßtechnik, D-7512 Rheinstetten, Federal Republic of Germany

[3] Deutsches Wollforschungsinstitut, Veltmanplatz 8, D-5100 Aachen, Federal Republic of Germany

Dedicated to Prof. Dr. H.-J. Cantow on the occasion of his 60th birthday

SUMMARY

50.3 MHz ^{13}C NMR cross-polymerization/magic angle spinning spectra of the raw silk of the mulberry silk moth (bombyx mori) and of crystalline silk fibroin were compared with those of synthetic copolypeptides of alanine and glycine. In crude silk, but not in crystalline fibroin , tyrosine, valine, aspartic acid and glutamic acid were detected in addition to alanine, glycine and serine. Furthermore, crude silk obviously contains ca. 10 % of an α-helix fraction which mainly consists of Ala-units. A small α-helix fraction (ca. 5 %) was also detected in the raw silk of the tussah silk moth (antheraea mylitta).

INTRODUCTION

In several previous papers we have demonstrated [1-5] that ^{13}C NMR CP/MAS spectra are a useful tool for the qualitative and quantitative analysis of the secondary structure of both homo- and copolypeptides. This work had the purpose to extend our spectroscopic investigations to naturally occuring proteins. However, because only proteins with a relative simple primary structure promise to give ^{13}C NMR CP/MAS spectra which are easily interpreted, we have begun our studies with silk fibroins.

RESULTS and DISCUSSION

The silk protein of bombyx mori consists of two components which differ largely in amino acid composition (Table I) and properties. Sericin is a water soluble, amorphous protein with a high serine content in which the fibrous fibroin is embedded. Fibroin mainly consists of glycine and alanine and contains crystalline areas with regular sequences of the

Table I Amino acid composition of silk proteins [7,8]

Amino acid	Sericin [a] (bombyx mori)	Fibroin [a] (bombyx mori)	Fibroin [a] (antheraea mylitta)
Glycine	14.7	44.5	23.5
Alanine	4.3	29.3	37.0
Serine	37.3	12.1	9.8
Threonine	8.6	1.0	1.0
Tyrosine	2.5	5.2	4.8
Aspartic acid	14.8	1.3	5.7
Glutamic acid	3.4	1.0	1.0
Arginine	3.5	0.5	13.3
Valine	-.-	2.2	0.8
Tryptophane		0.1	3.0

a) in mol % ; the less abundant amino acids are not listed

Table II ^{13}C NMR chemical shifts δ (ppm, relative to TMS) of solid raw silk (bombyx mori)

Amino acid	Secondary Structure	CO	α-C	ß-C	further carbons
Gly	ß-sheet	168.8	43.7	—	
L-Ala	ß-sheet	171.1 172.3	48.0 49.3	19.5	
	α-helix	176.8	52.9	15.6	
L-Ser	ß-sheet	(170.0)	53.0	62.8	
L-Tyr	ß-sheet	(169.6)	(52.1)	(39.3)	154.5, 128.5, 118.1
L-Val	ß-sheet	(171.5)	(58.2)	31.8	(18.5 γ-CH_3)
L-Asp	ß-sheet	(173.8)	(51.0)	37.6	(176.8)
L-Glu	ß-sheet	(172.6)	(51.3)	29.6	29.6 (γ-C)

a) The chemical shifts in brackets are those of the homopolypeptides. In the spectrum of silk fibroin they are obscured by the stronger signals of Gly and Ala residues.

type: $[-(Gly-Ala)_2-Gly-Ser-]_n$. In the raw silk we have measured (Fig. 1 A) the fibroin is about four times more abundant than the sericin. The so-called crystalline fibroin was obtained from the raw silk by dissolution in water containing the Cu^{2+}/ethylene-diamine complex and enzymatic degradation by means of chymotrypsin. The white crystalline powder which precipitated from this reaction mixture was used for the ^{13}C NMR measurements (Fig. 1 B).

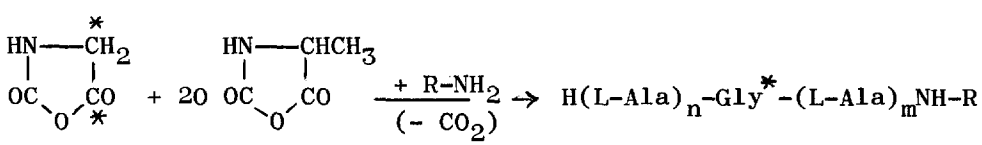
Both 50.3 and 75.4 MHz spectra of raw silk (bombyx mori) were measured. Because the resolution of the 75.4 MHz spectra was not better, whereas the spinning side bands were considerably more intensive, only 50.3 MHz spectra were measured in all other cases. The original spectrum and several resolution enhanced versions of this spectrum (Fig. 1 A) yielded the following information. The presence of Gly, Ala, Ser and Tyr units is evident from their CO-, α-C, ß-C or phenyl-C signals. Their chemical shifts agree well with those found for the ß-sheet structure of the corresponding homopolypeptides [2]. The line widths of 200 - 400 Hz prevent a distinction between different kinds of ß-sheets if such were present [6]. Of particular interest are the signals at 177, 53 and 15 ppm, which are characteristic of Ala units in an α-helix structure. On the basis of the chemical shifts alone these three signals can also be attributed to other amino acids having a ß-structure:

1) The 17.7 ppm signal might stem from the ß-carboxyl group of aspartic acid.
2) The 53 ppm signal stems at least partially from the α-C of serine.
3) The 15 ppm signal might stem from the δ-CH_3 group of isoleucine [2].

However, all three signals are significantly more intensive than expected for those alternative assignments. For instance, the weak signal at 38.5 ppm represents the ß-CH_2 group of Asp. The signal at 66 ppm stems from the CH_2 group of Ser; its intensity is only 1/3 that of the 53 ppm signal, and the content of Ile is <0.8 %. Hence, the three signals at 177, 53 and

15 ppm suggest that crude silk contains short helical blocks mainly consisting of Ala units.

Since the mulberry silk possesses a high content of glycine, the question arises whether the helical parts may contain Gly units and whether glycine in helical environments is detectable in the ^{13}C NMR spectra. Because poly(glycine) itself does not form an α-helix we have synthesized a so-called guest-host polypeptide from glycine-N-carboxyanhydride (NCA) containing 20 % ^{13}C (denoted $*$) and from a 20 fold molar excess of L-Ala-NCA. (Eq. 1). The resulting polypeptide was mainly helical, and thus, most Gly units were forced into an α-helix structure (Fig. 2 A). The CO-signal of the helical Gly units absorbs ca. 3 ppm downfield of its ß-sheet position [2] whereas the CH_2 signal is insensitive to a change of the secondary structure[2]. Because the CO-signal of helical Gly overlaps with the CO-signal of the Ala units in the ß-sheet structure, the ^{13}C NMR CP/MAS spectra of silk fibroin do not allow the detection of helical Gly units. We have also synthesized a copolypeptide with a random sequence of Gly and Ala units in a mole ratio of 1:4. Figure 2 B shows that this copolypeptide contains ca. 60 % α-helix structure. Hence, we may conclude that Ala sequences containing up to ca. 10 % Gly units may form α-helices. Furthermore, it is to be noted that helical poly(alanine) may contain up to 40 mol % Val units [9]. Thus the helical structures in raw silk may contain a variety of amino acids including glycine and valine, even though only the signals of the helical Ala units were detectable in our spectra.



$n + m = 20$; $* \approx 20$ % ^{13}C enrichment.

In the spectrum of crystalline fibroin the "α-helix signals" are not detectable, in agreement with the pleated sheet structure reported in literature[10]. Furthermore, the signals of tyrosine are absent along with the signals u, v and w. The

absence of these signals in Fig. 1B supports their assignments to Asp (ß-CH_2), Val (ß-CH) and Glu (γ-CH_2) (Table I). Furthermore, it is noteworthy that the intensity ratios of Gly, Ala and Ser signals have changed. Evidently, the crystalline fibroin contains more alanine and less serine compared to raw silk,in good agreement with the data of Table I.

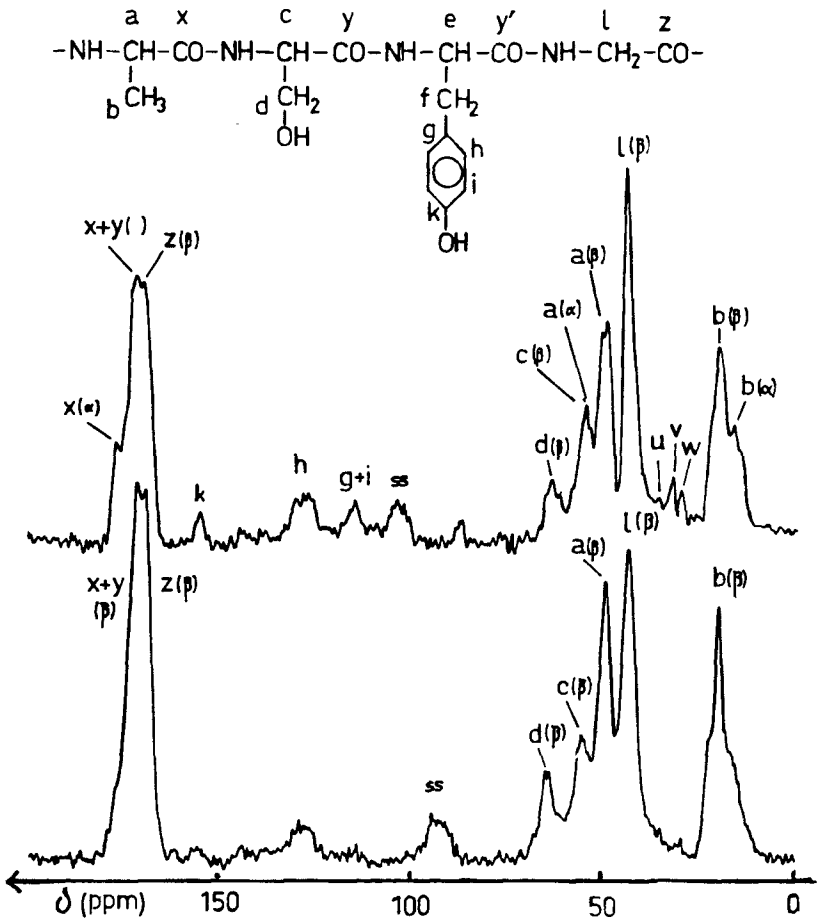


Fig.1 50.3 MHz ^{13}C NMR CP/MAS spectra (slightly resolution enhanced) of A) raw mulberry silk, and B) crystalline silk fibroin.

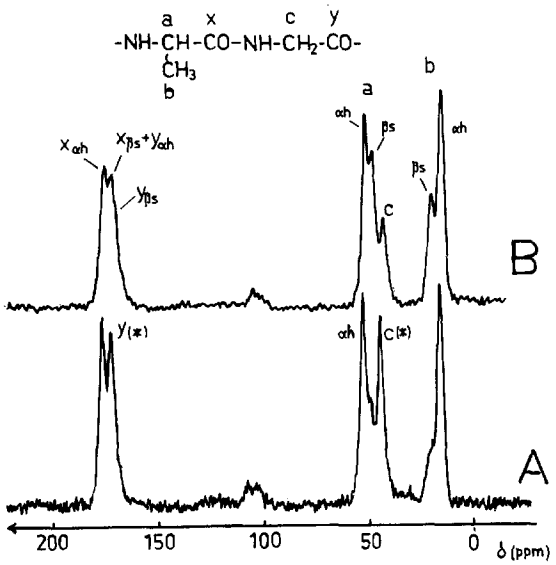
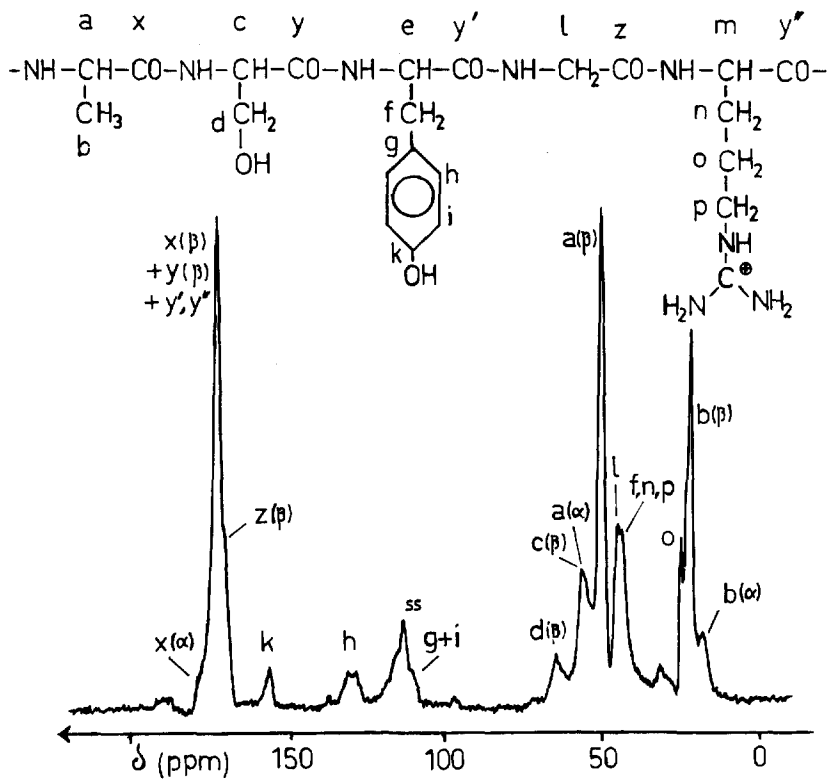


Fig.2 50.3 MHz ^{13}C NMR CP/MAS spectra of random copolypeptides: A) (Gly*/Ala)$_n$ mole ratio 1:20; B) (Gly/Ala)$_n$ mole ratio 1:4.

Fig.3 50.3 MHz ^{13}C NMR CP/MAS spectrum (slightly resolution enhanced) of raw tussah silk.



Finally, we have measured the raw silk fibroin of the tussah silk moth (antheraea mylitta). The amino acid composition of this fibroin differs from that of bombyx mori silk in that the former contains more alanine than glycine and more arginine than serine or tyrosine (Table I). In agreement with this composition signals of Ala, Gly, Arg, Ser and Tyr units are detectable and the α-C signal of Ala is 1.5 times more intensive than that of glycine and ca. 4 times more intensive than that of serine (Fig. 3). Most interesting is the presence of signals (or shoulders) at 176 and 15 ppm which are characteristic for helical alanine. Furthermore, the α-C signal of serine at 53 ppm is three times more intensive than its ß-C signal at 62 ppm. Hence the 53 ppm signal might contain a contribution from the α-C signal of helical Ala-units. Because the concentration of isoleucine is below 0.3 % the signal at 15 ppm cannot result from the δ-CH_3 group of this amino acid. Hence, the assignment of the 15 ppm signal to helical alanine is still more reliable in the case of tussah silk than in the case of mulberry silk. Thus, we may conclude that tussah silk, similar to mulberry silk, contains a small fraction (ca. 5 %) of α-helices which has hitherto not yet been detected. Thus Figures 1-3 clearly demonstrate that ^{13}C NMR CP/MAS spectra allow a qualitative and semiquantitative analysis of composition and secondary structure of fibrous proteins. However, considering our studies of synthetic copolypeptides [9,11], we must emphasize that these spectra do not yield any sequence information.

EXPERIMENTAL

Copolymerizations: A) A mixture of 5 mmol ^{13}C enriched Gly-NCA and 100 mmol L-Ala-NCA were dissolved in 150 ml dry dioxane and 1 mmol benzylamine was added. The reaction mixture was stored for 4 days at 20°C, then diluted with 400 ml diethyl ether and filtered. B) A mixture of 10 mmol Gly-NCA and 40 mmol L-Ala-NCA was dissolved in 100 ml dry dioxane and polymerized with 0.5 mmol benzylamine at 100°C. After 20 h the reaction mixture was cooled, diluted with 400 ml diethyl ether and filtered. ^{13}C NMR sequence analyses demonstrate [12] that under

these conditions random Gly/Ala sequences are obtained.

NMR-Measurements: The 50.3 MHz ^{13}C NMR CP/MAS spectra were measured on a Brüker CXP-200 in deuterated PMMA rotors at a spinning rate of ca. 3 kHz. A single contact pulse sequence with alternation of the 90^{o} pulse phase, a contact time of 3 ms and a repetition time of 4 s were used. The magic angle was checked with glycine, and the chemical shifts were referenced to TMS as described previously [2]. Because the silk fibers do not allow a dense packing of the rotors, accumulation of ca. 15 000 transients was required for Figs. 1 and 3, whereas ca. 600 transients were sufficient for Figs. 2 A and B.

References

1) D.Müller and H.R.Kricheldorf; Polym.Bull 6 101 (1981)
2) H.R.Kricheldorf and D.Müller; Macromolecules in pr.(Part 3)
3) H.R.Kricheldorf, M.Mutter, F.Maser, D.Müller and H.Förster; Biopolymers in press (Part 4).
4) H.R.Kricheldorf, D.Müller and H.Förster; Polym.Bull.8 487
5) H.R.Kricheldorf and D.Müller;Polym.Bull. 8 495 (1982)
6) B.Lotz, A.Brack and G.Spach; J.Mol.Biol. 87 193 (1974)
7) S.Seiffer and P.M.Gallop in "The Proteins" (H.Neurath,Ed.) Vol.VI, Academic Press, New York 1966
8) F.Lucas, J.T.B. Shaw and S.G.Smith; J.Mol.Biol. 2 339 (1960)
9) H.R.Kricheldorf and D.Müller; Part 14 of this series (in preparation)
10) F.Lucas and K.M.Rudall in "Symposium on Fibrous Proteins" (W.G.Grewther Ed.) Plenum Press New York 1968
11) H.R.Kricheldorf unpublished results
12) H.R.Kricheldorf and G.Schilling; Makromol.Chem. 179 1175 (1978)

Acknowledgment: We thank Dr. C.Bragadin (Milano) for a gift of tussah silk and the Deutsche Forschungsgemeinschaft for financial support.

Received October 24, accepted after correction December 22, 1982 C