Polymer Bulletin 45, 237–244 (2000)

Polymer Bulletin

Secondary structural analysis in the solid state for analogous sequential polypeptides of glycine-rich sequence of spider dragline silk

Yasumasa Fukushima

Department of Applied Chemistry, Faculty of Engineering, Toyo University, 2100 Kujirai, Kawagoe, Saitama 350-8585, Japan

Received: 5 August 2000/Accepted: 24 August 2000

Summary

Four sequential copolypeptides poly(X-Gly-Gly) with X being Ala, Tyr, Gln, or Leu were prepared as a model of glycine-rich sequence of dragline spider silk produced by *Nephila clavipes* and their secondary structures in the solid state were characterized by FT-IR spectroscopy. Poly(Tyr-Gly-Gly), poly(Gln-Gly-Gly), and poly(Ala-Gly-Gly) form the ß-sheet structures, whereas poly(Leu-Gly-Gly) existed as a disordered conformation as a cast film from formic acid. These results indicated that X-Gly-Gly sequences with Tyr, Gln, and Ala could contribute to the formation of the ß-sheet structure in glycine-rich sequence.

Introduction

Spider dragline silk produced by *Nephila Clavipes* has received much attention as a high performance protein fiber possessing both high tensile strength and high elasticity [1, 2]. The tensile strength is greater than that of steel and has the same order of magnitude as aramid [3]. The spider silk exhibits up to 35% elongation which is thought to be based on an entropic mechanism. Therefore, the energy absorbed to break the spider silk can exceed the values for steel and aramid. In addition, the spider silk is produced under ambient conditions from aqueous solution in nature, while steel and aramid require the high temperature in metal processing and the hazardous solvent in spinning, respectively.

The origin of these unique mechanical properties has been interested for a long time but have remained equivocal. It is essential to understand the outstanding properties of the spider silk in terms of its microscopic structure including amino acid sequence, the secondary structure, and the packing arrangement in crystalline and amorphous regions. The amino acid sequence is comprised of predominantly two kinds of repetitive amino acid sequence. One is a 13 amino acid segment containing polyalanine, and the other is a 15 amino acid highly conserved segment (glycine-rich sequence) composed of X-Gly-Gly repeat with X being alanine (Ala), tyrosine (Tyr), glitamine (Gln), or leucine (Leu) [4, 5]. Most structural studies indicate that the spider

silk is semicrystalline protein composed of an antiparallel pleated ß-sheet structure interspersed with elastic unordered segments on X-ray diffraction [6] and Fourier transform infrared (FT-IR) spectroscopy [4]. It is not clear whether the ß-sheet structure is formed by the polyalanine or the glycine-rich sequence.

Recently, genetic engineering and biotechnology provide the tools to synthesize artificial repetitive polypeptides with absolute control over molecular weight, composition, sequence, and stereochemical purity [7]. I have also prepared the repetitive polypeptide consisting of the fifteen amino acid sequence (Gly-Leu-Gly-Gly-Gln-Gly-Gly-Gly-Ala-Gly-Gln-Gly-Gly-Tyr-Gly) in glycine-rich sequence by recombinant DNA technique [8,9]. The biosynthetic polypeptide in the solid state existed as disordered structure in the lyophilized form, whereas it formed ß-sheet structure as a cast film from formic acid on FT-IR spectroscopy. These results could suggest that the ß-sheet structure in the crystalline region was composed of the glycine-rich sequence.

In this paper I intend to investigate the effect of the amino acid residue in the glycine-rich sequence on the ß-sheet structural formation. I synthesize four kinds of sequential copolypeptides, poly(X-Gly-Gly) where X is Ala, Leu, Tyr, or Gln by a conventional chemical method and analyze secondary structure of the copolypeptide as a cast film from formic acid by FT-IR measurement.

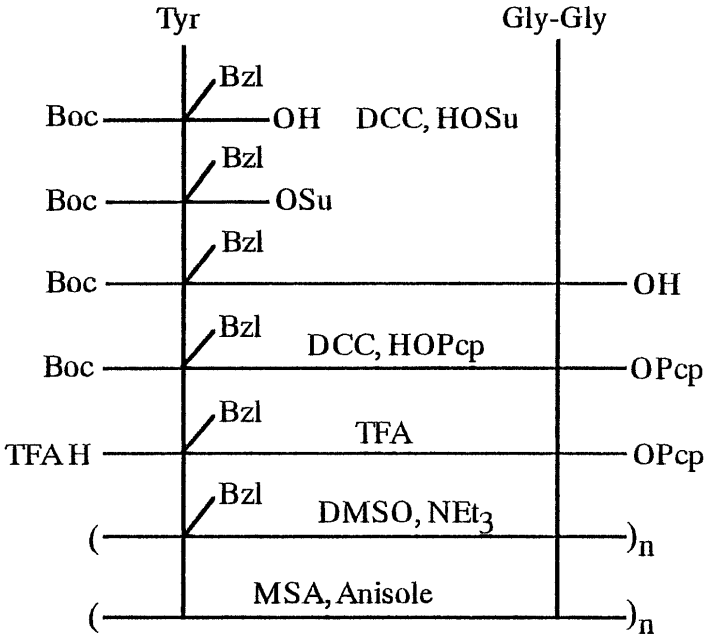
Experimental

Materials and methods

Four kinds of Boc (*tert*-butyloxycarbonyl)-protected amino acids and Gly-Gly dipeptide were purchased from Peptide Institute, Inc. and Bachem, respectively. All other reagents and solvents of high purity commercially available were used without further purification. ^{1}H NMR spectra were recorded for dimethyl sulfoxide-d_6 (DMSO-d_6) solutions at 300 MHz with a Varian VXR300. Melting points were determined on a Buchi 535 melting point apparatus without any correction. FT-IR spectra were obtained for a cast film from formic acid on JEOL JIS-5500 spectrometer. The amino acid composition of the polypeptides was determined with a Hitachi 835 amino acid analyzer. Intrinsic viscosity was measured in dichloroacetic acid with a Ubbelohde capillary viscometer at 25 °C. I estimate the order of the molecular weights of polypeptides from the intrinsic viscosity using the viscosity-molecular weight relationship for poly-benzyl-L-glutamate in dichloroacetic acid [10].

Peptide Synthesis

The synthetic routes of poly(Tyr-Gly-Gly) are illustrated in Scheme 1 and the synthetic procedure is described below as a representative of the polypeptide syntheses.



Scheme 1 Representative scheme for the synthesis of Poly(Tyr-Gly-Gly).

Boc-Tyr(Bzl)-Gly-Gly-OH (I)

N-Hydroxysuccinimide ester of Boc-O-benzyl-L-tyrosine (2.85 g, 6.08 mmol) in 70 ml of tetrahydrofuran (THF) was added to glycyl-glycine (0.80 g, 6.08 mmol) dissolved in THF (50 ml) and water (50 ml) containing sodium bicarbonate (1.02 g, 12.2 mmol) [11]. After the solution was stirred at room temperature for 18 h, the reaction mixture was concentrated, acidified with 0.5 N hydrochloric acid, and extracted with ethyl acetate. The organic layer was dried over anhydrous sodium sulfate, filtered and concentrated to give an oily product. The residual oil was triturated with ether and recrystallization from ethyl acetate and ether gave 2.01 g (68.2 %) of **I** as a white crystal.

Boc-Tyr(Bzl)-Gly-Gly-OPcp (II)

A solution of compound **I** (1.50 g, 3.09 mmol) and pentachlorophenol (Pcp) (0.91 g, 3.40 mmol) in THF (80 ml) was treated with *N, N'*-dicyclohexylcarbodiimide (DCC) (0.71 g, 3.40 mmol) at 0 °C for 15 h with stirring [12]. The produced dicyclohexylurea was filtered off and the filtrate was evaporated in vacuo and recrystallization from ethyl acetate and n-hexane gave 1.89 g (84.3 %) of **II** as a white crystal.

Poly[Tyr(Bzl)-Gly-Gly] (III)

To compound **II** (1.50 g, 2.00 mmol) was added 5 mL of trifluoroacetic acid (TFA) and the mixture was stirred at room temperature for 1h. After evaporation of TFA in vacuo, trituration with ether gave a white precipitate. To the solution of the product in 1.2 mL of DMSO was added triethylamine (0.50 mL, 3.58 mmol). The reaction mixture was stirred at room temperature to become very viscous within 2 days. After standing the mixture at room temperature for more than a week, the mixture was triturated with ether. The precipitate was washed with water, methanol and ether, the yield being 0.29 g (47.2%) of **III**.

Poly(Tyr-Gly-Gly) (IV)

Compound **III** (0.11 g, 0.30 mmol) was treated with 1 mL of methanesulfonic acid (MSA) in the presence of anisole (0.1 mL) at room temperature for 1h and the mixture was triturated with ether to give a white precipitate [13]. The solid was washed with methanol and ether and dried to obtain 0.046 g (55.5%) of **IV**.

RESULTS and DISCUSSION

Synthesis and Polymerization of Tripeptides

Each of the four tripeptide monomers was synthesized similarly in THF/water with $NaHCO_3$ by the conventional active ester method, adding Boc amino acid N-hydroxysuccinimide ester to a growing Gly-Gly dipeptide with a free α-carboxylic acid. In order to avoid racemization, the syntheses were conducted by positioning glycine at the C-terminal residue. DCC has been used to synthesize N-hydroxysuccinimide active esters, which are colorless crystalline derivatives with good stability [14]. The hydroxyl groups of Tyr was protected with the benzyl (Bzl) group. The Boc group was selectively removed by anhydrous TFA, while the Bzl group was eliminated after completion of polymerization of the polypeptide by a treatment with MSA in the presence of anisole [13].

The polymerization of sequential polypeptides having a defined repeating unit of amino acids is usually achieved by a self-condensation of monomer peptide active esters. For polymerizing the tripeptide monomers, the free acid was activated by conversion to its pentachlorophenyl ester. Compound **II** was synthesized by coupling compound **I** with pentachlorophenol in the presence of DCC [12]. The physical properties of the N-protected peptide active esters as a monomer are summarized in Table 1. All of the monomers were generally obtained in good yield and found to be homogeneous by melting point and chromatography. In addition, they gave elemental analyses that agreed well with those expected for unhydrated. These results indicate that all of the monomers are sufficiently pure to be used for the polymerization reaction in the next step.

Table 1 Physical properties and elemental analyses of monomers

Monomer	mp (℃)	R_f^*	Elemental analysis (%) Calc. (Found) C	H	N
Boc-Ala-Gly-Gly-OPcp	181-183	0.56	39.19 (38.47)	3.65 (3.42)	7.62 (7.39)
Boc-Gln-Gly-Gly-OPcp	139-141	0.05	39.47 (40.03)	3.81 (3.98)	9.20 (8.87)
Boc-Leu-Gly-Gly-OPcp	124-127	0.63	42.48 (43.11)	4.41 (4.62)	7.08 (6.85)
Boc-Tyr(Bzl)-Gly-Gly-OPcp	108-112	0.47	50.74 (49.69)	4.12 (3.78)	5.73 (5.65)

* Chloroform-methanol-acetic acid (95 : 5 : 3, v/v)

The polymerization was carried out in DMSO in the presence of triethylamine after removing the Boc group of compound **II** by TFA. After the polycondensation, the resulting polypeptides were isolated by precipitation from the reaction system by addition of diethyl ether. All signals of the ^{1}H NMR spectrum of compound **III** broadened due to polymerization.

The Bzl groups of two protected polypeptides were removed by MSA in the presence of anisole. The ^{1}H NMR spectra of compound **IV** shows a disappearance of the signals assigned to the all protecting groups after the cleavage reaction.

Polymerization results of the sequential polypeptides are summarized in Table 2. The polypeptides have low solubility in general organic solvents and water. I can estimate the molecular weights of the polypeptides by the intrinsic viscosity measurements in dichloroacetic acid which is a random-coil forming solvent for polypeptides. Here, if I employ Doty's equation (1) which was obtained in the case of poly-benzyl-L-glutamate in dichloroacetic acid [10].

$$[\eta] = 2.78 \times 10^{-5} M^{0.87} \quad \text{------} \quad (1)$$

It is apparent from the intrinsic viscosity values in Table 2 that each polypeptide has a molecular weight greater than 20,000 which corresponds to a intrinsic viscosity value of 0.15 dL/g. Thus the polypeptides are found to be produced in a sufficiently

high degree of polymerization for the characterization of their secondary conformation. In addition, each polypeptide has similar amino acid ratio to that expected.

Table 2 Synthetic results of the sequential polypeptides

Polypeptide	[η](dL/g)[a]	Molecular weight [b]	Amino acid composition (in %)	
			calcd (found)	calcd (found)
poly(Ala-Gly-Gly)	0.33	48000	Gly : 2 (2.07)	Ala : 1 (0.93)
poly(Gln-Gly-Gly)	0.17	23000	Gly : 2 (2.04)	Glx : 1 (0.96)[c]
poly(Leu-Gly-Gly)	0.24	33000	Gly : 2 (1.97)	Leu : 1 (1.03)
poly(Tyr-Gly-Gly)	0.17	23000	Gly : 2 (1.92)	Tyr : 1 (1.08)

a) Measured at 25 °C in dichloroacetic acid.

b) Calculated by $[\eta] = 2.78 \times 10^{-5} M^{0.87}$

c) Glx = Gln + Glu

Characterization of Secondary Structure of the Polypeptides in the Solid State by FT-IR

Figure 1 shows the IR spectra of the four polypeptides as cast film from formic acid in amide I and II regions, which have been used for characterization of the secondary structure of many proteins [15-17]. The IR spectrum of poly(Leu-Gly-Gly) exhibits amide I and II vibrational modes at 1650 and 1540 cm^{-1}, respectively, which can be characteristic of a disordered conformation, and a shoulder around 1694 cm^{-1}, which can be characteristic of the regular alternation of a ß-sheet chain direction. Because the shoulder is a small peak and the other bands assigned to a ß-sheet structure are not observed, the spectrum indicates that poly(Leu-Gly-Gly) exists as predominantly a disordered conformation.

The IR spectra of poly(Ala-Gly-Gly), poly(Gln-Gly-Gly), and poly(Tyr-Gly-Gly) are very similar one another, which show three bands around 1694, 1650, and 1626 cm^{-1} in amide I region as well as two bands around 1540 and 1520 cm^{-1} in amide II region. The bands at 1626 and 1520 cm^{-1} can be assigned to a ß-sheet conformation, and the bands around 1694, 1650, and 1694 cm^{-1} are characteristic of the conformation described above. Therefore, these polypeptides form the ß-sheet conformations. Because the ß-sheet conformational contents increase with increasing the signal strengths at 1626 and 1520 cm^{-1}, the ß-sheet content of poly(Tyr-Gly-Gly) is as almost high as that of poly(Gln-Gly-Gly), and poly(Ala-Gly-Gly) has slightly lower ß-sheet content than poly(Tyr-Gly-Gly) and poly(Gln-Gly-Gly). It is found that the ß-sheet content is

decreased in following order; poly(Tyr-Gly-Gly) = poly(Gln-Gly-Gly) > poly(Ala-Gly-Gly) > poly(Leu-Gly-Gly).

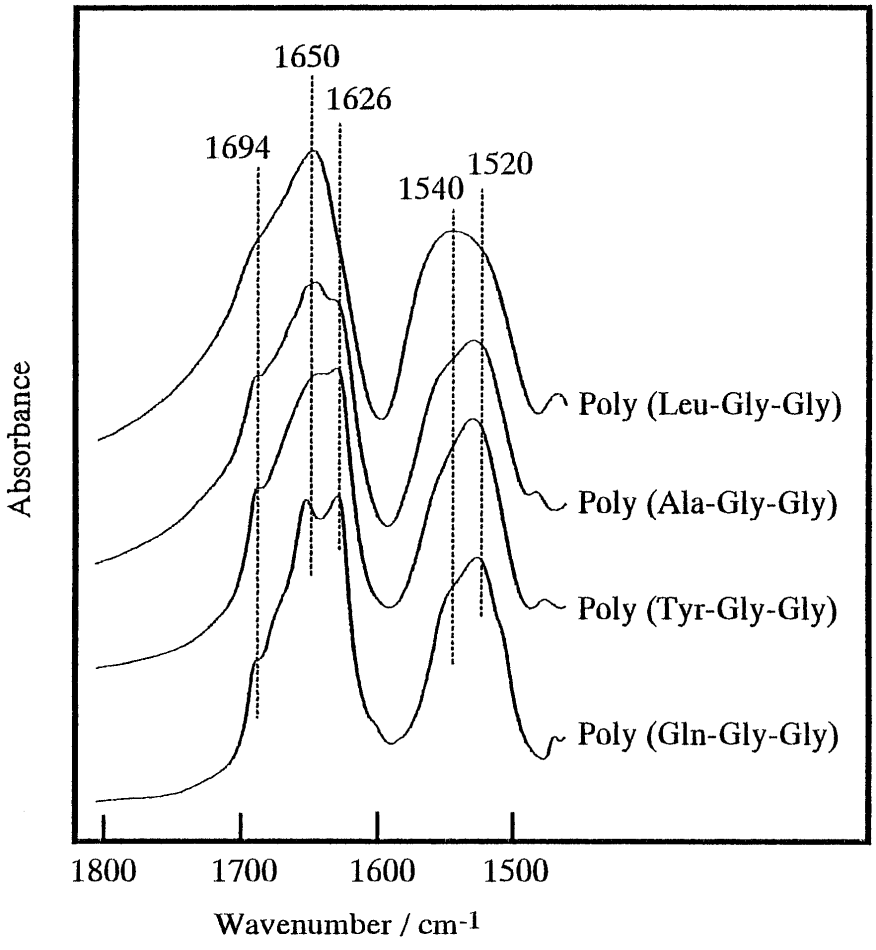


Figure 1 FT-IR spectra in amide I and amide II regions on the samples of the sequential polypeptides prepared as cast film from formic acid.

Poly(Ala-Gly-Gly) adopts the ß-sheet structure even though Gly and Ala have poor ß-sheet propensities and Gly has a very strong tendency to form ß-turn structure as proposed by Chou and Fasman [18]. Poly(Tyr-Gly-Gly) and poly(Gln-Gly-Gly) form the ß-sheet structure due to strong tendency of Tyr and Gln to adopt the ß-sheet. Although Leu has a very strong tendency for the formation of α-helix, poly(Leu-Gly-Gly) exists as a disordered conformation.

Glycine-rich sequence, which is X-Gly-Gly sequence with the bulky residues such as Ala, Tyr, Gln, or Leu, exists as the ß-sheet structure in natural dragline spider silk fiber. These results indicate that X-Gly-Gly sequences with Tyr, Gln, and Ala form the ß-sheet structure while Leu-Gly-Gly sequence cannot contribute to the ß-sheet structural formation.

Acknowledgments

I thank Dr. M. Hirami of Unitika Ltd. for many helpful discussions and valuable comments on this work.

References

1 Gosline JM, DeMont ME, Denny MW (1986) Endeavor 10: 37
2 Kaplan DL, Fossey S, Mello CM, Arcidiacono S, Senecal K, MullerW, Stockwell S, Beckwitt R, Viney C, Kerkam K (1992) MRS Bull. 17: 41
3 Tirrell DA (1996) Science 271: 39
4 Xu M, Lewis RV (1990) Proc. Natl. Acad. Sci. USA 87: 7120
5 Hinman MB, Lewis RV (1992) J. Biol. Chem. 267: 19320
6 Warwicker JO (1960) Faraday Soc. Trans. 52: 554
7 McGrath KP, Fournier MJ, Mason TL, Tirrell DA (1992) J. Am. Chem. Soc. 114: 727
8 Fukushima Y (1998) Biopolymers 45: 269
9 Fukushima Y (1998) Chem. Lett. 939
10 Doty P, Bradbury JH, Holtzer AM (1956) J. Am. Chem. Soc. 78: 947
11 Sugihara T, Imanishi Y, Higashimura T, Shimonishi Y (1976) J. Org. Chem. 41: 2584
12 Kovacs J, Giannotti R, Kapoor A (1966) J. Am. Chem. Soc. 88: 2282
13 Fujino M, Wakimasu M, Shinagawa S, Kitada C, Yajima H (1978) Chem. Pharm. Bull. 26: 539
14 Doyle BB, Traub , Lorenzi GP, Brown FR, Blout BR (1970) J. Mol. Biol. 51: 47
15 Miyazawa T, Blout ER (1961) J. Am. Chem. Soc. 83: 712
16 Ambrose EJ, Elliot A (1951) Proc. R. Soc. London A205: 47
17 Bandekar J, Krimm S (1988) Biopolymers 27: 909
18 Chou PY, Fasman GD (1978) Adv. Enzymol. 47: 45