Components of Silk Proteins and Their Gene Loci in the Silkworm

By TAKUMA GAMO

Department of Molecular Biology, National Institute of Agrobiological Resources
(Yatabe, Ibaraki, 305 Japan)

Introduction

Silk fiber produced by the reeling of thread from cocoons of the silkworm, *Bombyx mori*, is composed of two classes of proteins, fibroin and sericin. These silk proteins are synthesized in the silk gland existing in the inside of the larvae. Fibroin is synthesized specifically in the posterior silk gland, while sericin is synthesized in the middle silk gland. Fibroin protein secreted into the gland lumen moves from the posterior to the middle section of the gland for storage, where it is coated by sericin protein, which acts to cement the two silk threads which are spun out from both sides of spinneret[1,10].

As fibroin and sericin can be obtained easily, a number of investigations have been conducted to elucidate their molecular structure. These efforts have not been successful, however, due to difficulties in solubilizing the proteins. Recent development in the solubilization and separation of proteins has enabled observation of genetic variation in polypeptides useful for the linkage analysis.

Solubilization of silk proteins

For the linkage analysis of genes encoding fibroin and sericin proteins, these proteins are dissolved in mild condition without cleavage of peptide bonds[2,4].

1) Liquid silk from silk gland

Usually, liquid silk existing in the lumen of silk gland is used as a material of genetic analysis. In this case, liquid silk taken out from the different sections of silk gland was dissolved in 8 M urea (Fig. 1). For the analysis of polypeptides in sericin proteins existing in the anterior section of the middle silk gland, protein solution dissolved in 8 M urea

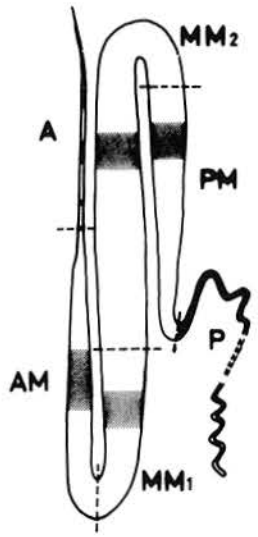

Fig. 1. Diagrammatic representation of silk gland in the mature larvae of the silkworm
Shadowed parts in the figure indicate the section used for the extraction of silk proteins, fibroin and sericin.
A: the anterior gland, AM: the anterior section in the middle gland, MM: the middle section in the middle gland, PM: the posterior section in the middle gland, P: the posterior gland

can be used directly for gel electrophoresis without disulfide-cleavage. SDS should be added into the protein solution at the concentration of 10%, if polyacrylamide gel electrophoresis (PAGE) containing SDS will be used for the analysis of sericin proteins. Fibroin and one component of sericin synthesized in the posterior section of the middle silk gland are composing disulfide bond. Therefore, disulfide bond of these proteins should be reduced

with 2-mercaptoethanol at pH 8.6 and the resulting residues are masked by the alkylation with iodoacetoamide.

2) Silk proteins in cocoon thread

Sericin proteins in cocoon thread can be dissolved by the incubation with 8 M urea solution containing 2-mercaptoethanol[2]. After alkylation of the resulting residues and removing the extra reagents by dialysis, protein solution was applied to gel electrophoresis. For the analysis of sericin proteins composing no disulfide bond, these sericins can be extracted from cocoon thread by the incubation with 8 M urea without 2-mercaptoethanol and applied to PAGE directly. Fibroin in cocoon thread can not be dissolved by this method. Silk fibroin can, however, be solubilized with 60% LiSCN and resulting polypeptides show the same electrophoretic pattern as that extracted from the silk gland[6].

Analysis of genetic variation

PAGE containing 4 M urea at acid pH or SDS-PAGE can be used for the detection of genetic variation in components of fibroin and sericin proteins. Laemmlie's SDS-PAGE[9] system gives the highest resolution of fibroin and sericin proteins[12].

Component of polypeptides in fibroin and sericin

Polypeptides extracted from the lumen of the silk gland show the existence of different components among the sections of the silk gland (Plate 1)[4]. Fibroin is synthesized in the posterior silk gland and secreted into the gland lumen. Secreted liquid fibroin moves to the middle gland for storage. By the reduction of disulfide bond, fibroin is cleaved into two polypeptides of high (f-1) and low (f-2) molecular weight. These two polypeptides show different molecular weight and amino acid composition (Tables 1 and 2)[4].

A polypeptide of sericin named s-4 is synthesized in the posterior section of the middle gland, and secreted into the gland lumen. This polypeptide exists in the lumen of the middle gland except for the anterior section (Plate 1). This component shows higher molecular weight when disulfide bond was not cleaved. In the rear part of the middle section, two polypeptides of sericin named s-1 and s-3 are synthesized, and these two polypeptides exist in the lumen of the anterior section of the middle gland. Furthermore, two other polypeptides named s-2 and s-5 are synthesized in the anterior section of the middle gland. Molecular weight of these polypeptides shows different values (Table 1). Amino acid composition is closely similar to each other (Table 2). However, different composition was observed in some amino acids. Polypeptides in sericin also show

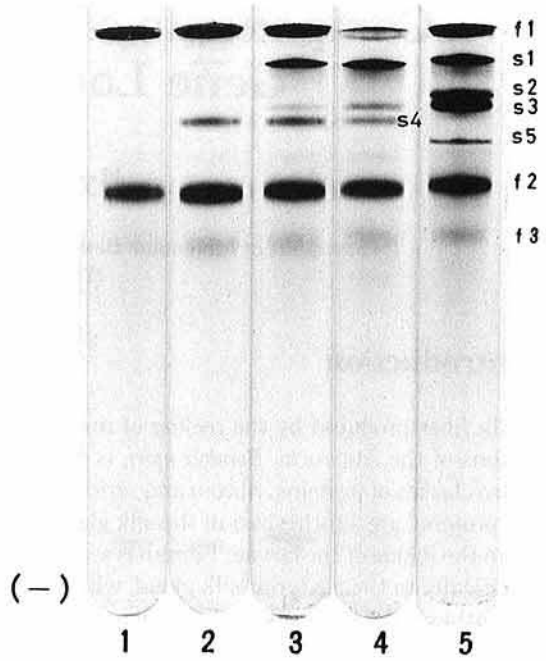

Plate 1. Electrophoresis of polypeptides in fibroin and sericin extracted by the reductive cleavage of disulphide bonds from different sections of the silk gland, on 5% polyacrylamide gel at pH 2.9 containing 4 M urea
Fraction numbers were given for the major bands decided as fibroin (f) or sericin (s).
Number on the abscissa.
1: the posterior gland, 2: the posterior section in the middle gland, 3: the rear part of the middle section in the middle gland, 4: the fore part of the middle section in the middle gland, 5: the anterior section in the middle gland

Table 1. Molecular weight of polypeptides in fibroin and sericin proteins

Polypeptide	Molecular weight*
Fibroin	
f-1 (H chain)	356 Kd
f-2 (L chain)	27
Sericin	
s-1	309
s-2	177
s-3	145
s-4	80
s-5	134

* Molecular weight was determined from the mobility in different gel percentages of acid PAGE at pH 2.9 containing 4 M urea.

the same pattern in SDS-PAGE. However, some polypeptide bands separate into more fractions by Laemmlie's PAGE system[12].

Linkage analysis of polypeptides in fibroin and sericin

1) Sericin

Variant components in sericin have first been observed in mutant strains, *Nd* and *Nd-s*, of the silkworm defecting the synthesis of fibroin[3]. A variant sericin found in the *Nd-s* mutant strain was clarified to be encoded by the allelic gene to the one of the s-2 sericin (Plate 2) and the gene loci was

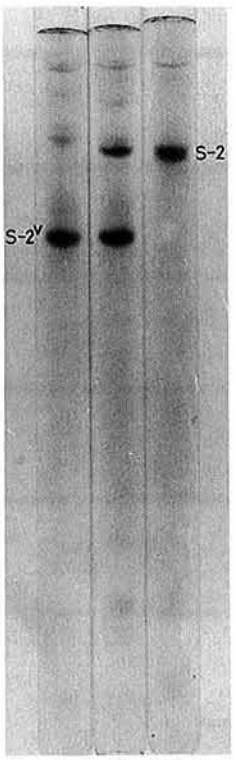

Plate 2. SDS-PAGE of sericin proteins extracted from cocoon shell by the reductive cleavage of disulphide bonds in 8 M urea
1: s-2^{v}/s-2^{v}, 2: s-2^{v}/+, 3: +/+.

Table 2. Amino acid composition of polypeptides in fibroin and sericin fractionated by gel electrophoresis at acid pH in 4 M urea

Amino acid	Composition (mole %)							
	f-1	f-2	f-3*	s-1	s-2	s-3	s-4*	s-5*
Glycine	45.5	11.4	15.6	18.0	12.5	14.2	11.4	17.5
Alanine	29.5	13.7	8.5	6.2	6.9	6.2	4.2	7.2
Valine	1.8	5.5	3.1	3.7	2.8	1.0	3.8	0.3
Leucine	0.2	7.0	6.1	2.1	1.5	0.6	2.8	2.7
Isoleucine	0.3	6.4	4.5	1.6	1.3	0.4	1.8	2.2
Serine	10.8	10.9	9.1	29.1	30.2	38.1	32.5	16.4
Threonine	1.5	3.4	4.8	8.5	8.0	4.1	11.1	4.0
Aspartic acid	1.0	14.7	11.7	13.0	12.9	12.6	11.6	13.9
Glutamic acid	1.0	8.9	6.4	5.4	5.2	10.7	6.1	10.3
Lysine	0.4	2.2	4.1	2.2	4.8	5.3	2.2	6.3
Arginine	0.3	4.1	3.8	4.0	3.9	3.0	3.6	3.5
Histidine	0.7	2.2	3.3	1.6	1.5	0.8	1.9	3.1
Tyrosine	6.2	3.1	9.5	3.3	6.8	2.1	2.8	9.0
Phenylalanine	0.7	1.7	4.1	1.5	0.7	0.4	3.4	1.2
Proline	0.4	3.4	2.9	Trace	Trace	Trace	Trace	Trace
Methionine	0.2	0.3	Trace	Trace	Trace	Trace	Trace	Trace
Cystine (half)**	0.2	1.3	2.7	Trace	1.0	0.5	0.7	2.5

*20 hr hydrolysis ** S-carboxyamidomethylated cysteine

Linkage Maps of *Bombyx mori*

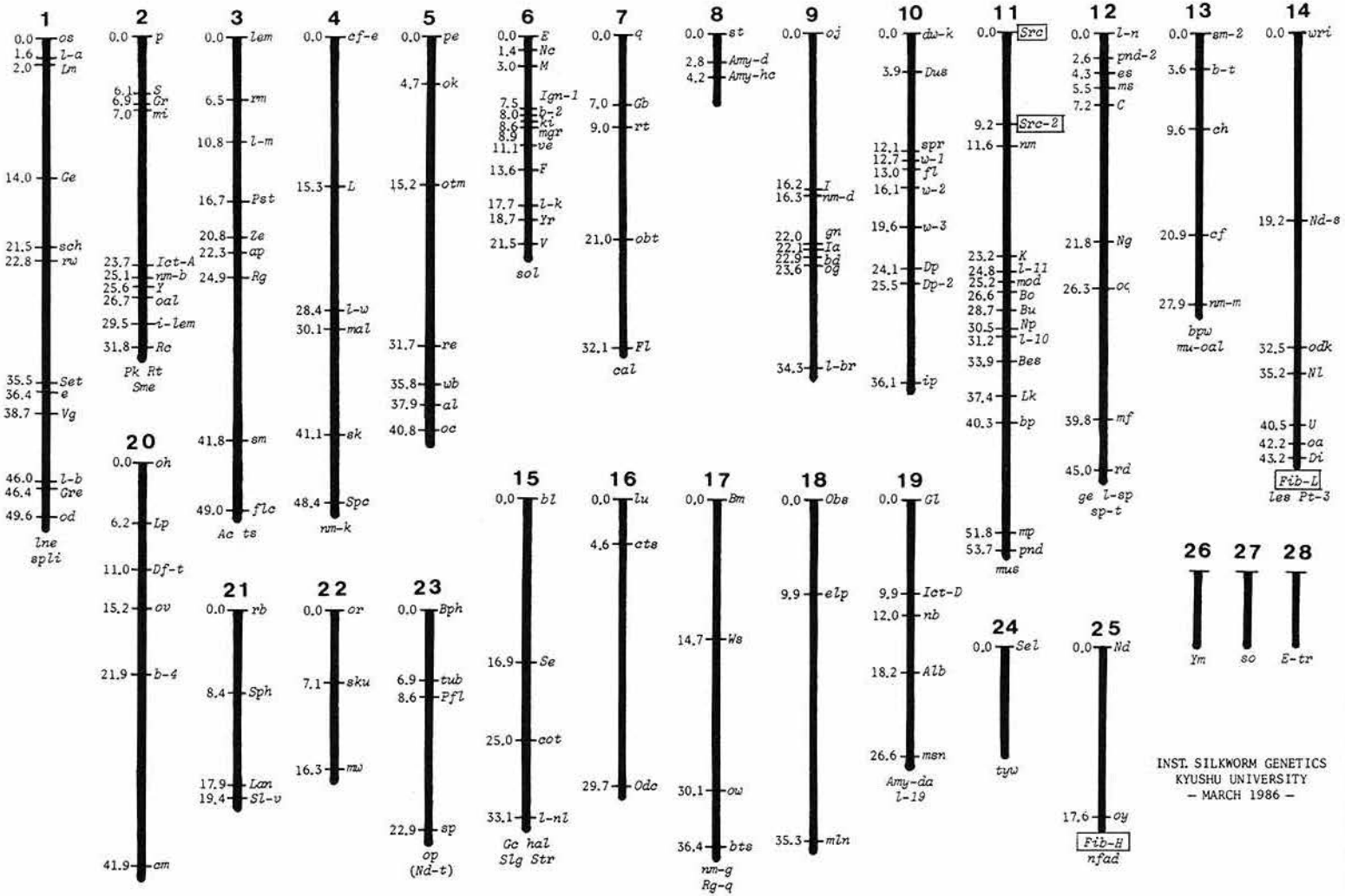

Fig. 2. Linkage maps of the silkworm, *Bombyx mori*, showing the localization of genes encoding fibroin and sericin proteins

analysed to be located at 0.0 position on the 11th linkage group[5]. The molecular weight of s-2^{v} variant polypeptide is lower by approximately 62,500 daltone than that of s-2 sericin. Therefore, it can be considered that this variant allele arose by deletion within the s-2 coding sequence in the *Src-2* gene locus.

Genetic variations were also observed in the other polypeptides of sericin and linkage analysis of genes encoding these polypeptides was carried out[12]. Three other polypeptides of sericin migrate with different mobility. However, genes encoding these three polypeptides exist in a closely adjacent region on the same chromosome 11. Since no recombination was observed among these three genes, they named tentatively *Src(s)* for these genes and the gene locus was determined at 0.0 point. Recombination value between the *Src* and *Src-2* was 9.2%. Therefore, the location of gene locus of the *Src-2* should be revised at 9.2 on chromosome 11 (Fig. 2).

Most of the sericin genes, therefore, can be considered to be located on a single chromosome. However, gene loci encoding the minor components of sericin proteins still remain obscure.

2) Fibroin

Fibroin protein is composed of two polypeptides, heavy (H) and light (L), linked with disulphide bonds. Genetic variations have also been observed by means of SDS-PAGE in heavy chain of fibroin[6,13]. Using these genetic variations, linkage analyses of the *Fib-H* gene were carried out and its gene locus was found to be located at the adjacent region of the *Nd* (naked pupa) gene defecting the fibroin synthesis, on chromosome 23[7,11].

Genetic variations were also observed in the light chain of fibroin[8]. Linkage analysis of the *Fib-L* gene showed that this gene is not located on the 25th chromosome carrying the *Fib-H* gene, but was found to be linked to the *U* gene on chromosome 14. A three point test cross was performed involving *U*, *Nd-s* and *Fib-L* genes[14]. No crossover was observed between *Fib-L* and *Nd-s*. Therefore, the *Fib-L* gene can be considered to locate at the adjacent region of the *Nd-s* on chromosome 14. The *Nd-s* gene is also defecting the fibroin synthesis.

Both fibroin genes encoding the heavy and light chains are locating at the adjacent region of the mutant genes defecting the fibroin synthesis. Therefore, it can be expected that the *Nd* and *Nd-s* genes are caused by the defect in the promoter region of the *Fib-H* and *Fib-L* genes, respectively, and further studies are in progress to elucidate the different DNA sequences in the 5′ flanking region of the two fibroin genes[15,16].

Acknowledgement

The study presented in this paper was carried out in the Sericultural Experiment Station where the author attended three years before. He wishes to express his sincere thanks to many co-workers. He also thanks Dr. H. Doira for permission to use the linkage map of the silkworm.

References

1) Fukuda, T. & Flokin, M.: Contributions to silkworm biochemistry. VII. Ordered progression of fibroinogen in the reservoir of the silkgland during the 5th instar. *Arch. Int. Physiol. Biochem.*, **67**, 214–222 (1959).
2) Gamo, T.: Electrophoretic analyses of the proteins extracted with disulphide cleavage from cocoons of the silkworm, *Bombyx mori* L. *J. Sericult. Sci. Jpn.*, **42**, 17–23 (1973) [In Japanese with English summary].
3) Gamo, T.: Genetically different components of fibroin and sericin in the mutants, *Nd* and *Nd-s*, of the silkworm *Bombyx mori*. *Jpn. J. Genet.*, **48**, 99–104 (1973).
4) Gamo, T., Inokuchi, T. & Laufer, H.: Polypeptides of fibroin and sericin secreted from the different sections of the silk gland in *Bombyx mori*. *Insect Biochem.*, **7**, 285–295 (1977).
5) Gamo, T.: Genetic variants of the *Bombyx mori* silkworm encoding sericin proteins of different lengths. *Biochem. Genet.*, **20**, 165–177 (1982).
6) Hyodo, A. & Shimura, K.: The occurrence of hereditary variants of fibroin in the silkworm *Bombyx mori*. *Jpn. J. Genet.*, **55**, 203–209 (1980).
7) Hyodo, A., Gamo, T. & Shimura, K.: Linkage analysis of the fibroin gene in the silkworm, *Bombyx mori*. *Jpn. J. Genet.*, **55**, 297–300 (1980).
8) Hyodo, A. et al.: Linkage analysis of the fibroin light chain gene in the silkworm, *Bombyx mori*. *Jpn. J. Genet.*, **59**, 285–296 (1984).
9) Laemmlie, U. K.: Cleavage of structural proteins during the assembly of the head of bacteriophage. *Nature*, **227**, 980–988 (1970).
10) Machida, J.: On the secretion of the silk substance in the silkworm (*Bombyx mori* L.). *J. Coll. Agr. Imp. Univ.*

Tokyo. **9**, 119–138 (1927).
11) Nakayama, M., Doira, H. & Xu, Meng-Kui: Linkage analysis of the *Fib-H* gene encoding heavy chain of fibroin. *Proc. Kyushu Sericult. Sci.* **15**, 51 (1984) [In Japanese].
12) Shonozaki, N., Doira, H. & Machida, Y.: Locus of the sericin gene *Src*. *53rd Annu, Conf. Jpn. Soc. Sericult. Sci.* 62 (1983) [In Japanese].
13) Sprague, K. U. et al.: Alleles of the fibroin gene coding the proteins of different length. *Cell*, **17**, 407–413 (1979).
14) Takei, F. et al.: Genetic analysis of the *Nd-s* mutation in the silkworm, *Bombyx mori*. *Jpn. J. Genet.*, **59**, 307–313 (1984).
15) Takei, F. et al.: Reduced level of secretion and absence of subunit combination for the fibroin synthesized by a mutant slikworm, *Nd* (2). *J. Cell Biol.*, **99**, 2005–2010 (1984).
16) Ueda, H. et al.: Sequence polymorphisms in the 5′-upstream region of the fibroin H-chain gene in the silkworm, *Bombyx mori*. *Gene*, **28**, 241–248 (1984).

(Received for publication, September 24, 1986)