

Untangled: species identification of silks by protein mass spectrometry



Boyoung Lee

Linacre College

University of Oxford

Supervised by A.M. Pollard

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2018

Table of contents

Table of contents	1
Abstract	5
 Acknowledgments.....	6
List of Tables	8
List of Figures.....	11
Note for citations in foreign languages	21
 I. Introduction.....	1
 I.1 Background	2
 I.1.1 Archaeological silk controversies	3
I.1.2 Conventional methods to characterise silks and their limitations	9
I.1.3 Protein mass spectrometry as an alternative option	13
 I.2 Research objectives, questions and structure	24
 II. Silk	28
 II.1 Entomology	29
 II.1.1 Mulberry silk.....	32
II.1.2 Non-mulberry silk.....	35
II.1.3 Sericulture.....	51
 II.2 History and archaeology	56
 II.2.1 Silk in China	56
II.2.2 Silk in India	63
II.2.3 Silk on the Road.....	65
II.2.4 Silk in the West.....	68

II.3	Physicochemical properties	87
II.3.1	Mechanical and Thermal properties	87
II.3.2	Microstructure.....	88
II.3.3	Crystal structures	92
II.3.4	Protein structures	95
III.	Materials, methods and instrumentation	102
III.1	Experimental timeline	103
III.2	Preparation of silks	104
III.2.1	Silk samples	104
III.2.2	Degumming.....	105
III.3	Preparation of fibroin solution.....	107
III.3.1	Solubilisation	107
III.3.2	Desalting	109
III.4	SDS-PAGE.....	110
III.4.1	Materials	110
III.4.2	Method.....	111
III.5	Enzymatic digestion.....	112
III.5.1	Materials	112
III.5.2	Methods.....	113
III.6	Peptide analysis.....	118
III.6.1	Sample submission	118
III.6.2	LC-MS/MS data acquisition	118
III.7	Data processing tools and analytical software parameters	121
III.7.1	Sequence data	121
III.7.2	Database query and De novo analysis.....	121

IV.	Solubilisation of silk fibroins	123
IV.1	Introduction.....	124
IV.2	Solubilisation properties	127
IV.2.1	Acids.....	127
IV.2.2	Calcium chloride ternary solution (Ajisawa's reagent).....	129
IV.2.3	Chaotropic salt solutions	130
IV.2.4	Effect of the solvents on other fibres.....	138
IV.3	Desalting and concentration	139
IV.4	Protein gel electrophoresis.....	142
V.	Identification of silk fibroins by peptide analysis.....	150
V.1	Introduction.....	151
V.2	Protein sequence analysis	152
V.2.1	Sequence alignment.....	152
V.2.2	Unique sequence features	155
V.2.3	Theoretical digestion with different enzymes and their effect on unique sequence features.....	168
V.3	Database query vs. <i>de novo</i> analysis.....	171
V.4	Effect of solvents.....	176
V.4.1	<i>Bombyx mori</i>	176
V.4.2	<i>Antheraea</i> and <i>Samia</i> silks	181
V.5	Effect of enzymes	207
V.5.1	<i>Bombyx</i> silks	207
V.5.2	<i>Antheraea pernyi</i>	228
V.5.3	<i>Antheraea mylitta</i>	244
V.5.4	<i>Antheraea yamamai</i>	254
V.5.5	<i>Antheraea assamensis</i>	268
V.5.6	<i>Samia ricini</i>	280

VI.	Case studies: implication to archaeological questions.....	293
VI.1	The textiles from Naintr�.....	294
VI.1.1	Introduction	294
VI.1.2	Experimental.....	298
VI.1.3	Results and discussion.....	299
VI.2	The textiles from Palmyra	305
VI.2.1	Introduction	305
VI.2.2	Experimental.....	308
VI.2.3	Results and discussion.....	310
VII.	Summary and conclusions.....	320
VII.1	Solubilisation of silk fibroins.....	321
VII.2	Identification of silks by protein mass spectrometry	325
VII.3	Consolidating an understanding of silk proteins for archaeological textile studies	329
VII.4	A model protocol for the analysis of archaeological silks.....	332
VII.5	Further works	336
Appendix A: Light Stable Isotope analysis of preserved mulberry leaves		
.....		339
	Background.....	340
	Experimental.....	341
	Result and discussion.....	344
	Conclusion	350
Appendix B: Sequence Alignments.....		351
Appendix C: Data Tables		366
Bibliography		390

Abstract

Historical records and archaeological findings suggest the commercial use of wild silks in the ancient Middle East and Europe before the dissemination of Chinese silk. However, scientific analysis of archaeological textiles is often hindered by their physical durability and the analytical capacity of conventional microscopic methods, and thus limitations of accurate identification have resulted in many controversies. This research aims to promote the comprehensive understanding of silk, by providing literature reviews in the entomological, historical and physico-chemical aspects of the material. On this foundation, the experimental works were designed to establish the potential methods of protein mass spectrometry of silk fibroins. The objectives of the experiments were to distinguish and to identify the different species of silk. Firstly, a study on the solubilisation of silk fibroins using domesticated and wild silk species was implemented. Then the development of analytical protocols for the protein mass spectrometry of silk fibroins followed, employing peptide sequencing using tandem mass spectrometry as the primary tool of analysis. The resulting analytical protocol enabled the identification and distinction between seven different types of fresh silk samples of *Bombyx*, *Antheraea* and *Samia* silks. Furthermore, it was also effective in the identification of silks in severely degraded archaeological textiles. Although in its emerging stages, this research could offer a new, accurate and reliable approach to the identification of silks, and a way to resolve past and future disputes of archaeological textile attributions.

Acknowledgments

I wish to express my warmest gratitude to all those people whose comments, questions, criticism, support and encouragement—personal and academic—have left a mark on this work. While I alone am responsible for this thesis, it is nonetheless at least as much a product of years of interaction with, and inspiration by mentors and friends. I also wish to thank those institutions, which have supported me during the work on this thesis. Regrettably, but inevitably, the following list of names will be incomplete, and I hope that those who are missing will forgive me and will still accept my sincere appreciation of their influence on my work.

I am forever indebted to the Mass Spectrometry Research Facility (James McCullagh Group) at Chemistry Research Laboratory at Oxford University, where I have spent two valuable academic years, for providing supportive and stimulating working environments. Professor James McCullagh has always been extremely insightful and supportive, and particularly understanding my complicated academic situation. Ms. Elisabete Pires has also been an indefatigable advisor and above all a wonderful friend—I would have not been able to pull things through without you. You are a truly beautiful human being. Dr James Wickens also has provided useful advice when I'm lost with stuff and regulations. I want to express my deepest gratitude to the two wonderful Linacretes: Dr Holger Kramer of Imperial College for letting me into his precious lab at Oxford DPAG and trained me the ABC's of proteomics; and Ms Thamali Ayagama for helping me in many occasions during the training period. Moreover, I would like to thank the members of Research Laboratory for Archaeology and History of Art: my supervisor Prof. Alan Mark Pollard for his vision and guidance in doing interdisciplinary research of Archaeology and Science and always trying his best to find a solution for me. I also thank Dr Thibaut Deviese for being a good friend and mentor, as well as much useful collaboration in my own areas of endeavour; Luciana da Costa Carvalho for being my partner

in crime in the department and college.

Linacre College, Oxford, has been my home for the last five years and it will remain as is for life. The endless support and encouragement from this wonderful group of people have sustained me through the years. I thank Principal Dr Nick Brown for shaping Linacre as loving, caring and conscious community. Senior tutor and my college advisor, Dr Jane Hoverd has been my lighthouse through countless occasions and exhibited an example of an amazing human being. Linacre College has allowed me wonderful years of joy and laughter, but most importantly precious friendships. I cannot thank enough to the wonderful Guillaume Poncelet, Konstance Li, Dina Hestad, Jaeyoung Lee and Boipelo Tshwene-Mauchaza, Dragos Mosneagu, James Bae, Marios Kapsis and Dr. Amir Atepf Ibrahim for being great friends to me and Olive the golden retriever. I would particularly like to thank marvellous Carryn Yong for the long hours of endeavour for proofreading my thesis. Nyree Manoukian and Dr Kiron Neal also allowed their time for reviewing my humble work.

I must thank Prof. Tom Higham of RLAHA, University of Oxford and Prof. Matthew Collins of School of Archaeology, University of Cambridge for their invaluable insights and guidance to make this work a much better one.

Finally, this thesis is dedicated to my wonderful parents, Dr Sang-rin Lee and Mrs Geum-ja Lee, without whom it would have been impossible. My gratitude to them is beyond words, thank you for nurturing me into the person I am today.

Washington D.C, February 2019

List of Tables

Table 1. Comparison of temperate and tropical Tasar	37
Table 2. Anaphe moths used for silk production in Africa.	48
Table 3. Comparison of <i>P. otus</i> and <i>S. pyri</i>	50
Table 4. Different terms for sea silk used in the Western literatures from the 15th century.	81
Table 5. Amino acid compositions of <i>B. mori</i> proteins (excluding signal peptides).....	97
Table 6. Amino acid compositions of <i>Antheraea</i> and <i>Samia</i> silk fibroins (exclude signal peptides).....	101
Table 7. List of the silk samples selected for experiments	104
Table 8. Tryptophan and tyrosine contents in different silk fibroins.....	136
Table 9. The weights of cotton, wool and linen fibres before and after the 2 hours of solvent treatment at 125°C.....	138
Table 10. PEAKS analysis result of the SDS-PAGE processed fibroin solutions (Tryptic digestion: heavy chain only for the <i>Bombyx</i> silks; * <i>B. mandarina</i> corresponded to <i>B.</i> <i>mori</i> HC)	144
Table 11. Protein names and codes (UniProtKB) for each type of silk samples (*: Reviewed)	152
Table 12. Similarity (%) between the silk fibroin sequences through Pairwise alignment	153
Table 13. Signal peptides of each silk fibroin (highlighted in green).....	155
Table 14. Subsequence of GAGAG(X) unique to <i>B. mori</i>	156
Table 15. Sequence difference of <i>B. mori</i> and <i>B. mandarina</i> through alignment	158
Table 16. Unique peptides and residues found in <i>Antheraea</i> and <i>Samia</i> silk fibroins	160
Table 17. Cleavage rules of Trypsin and Chymotrypsin.	170
Table 18. Comparison of each theoretical digestion result of each silk fibroin. The top mass indicates the biggest peptide (MW) resulted from the enzymatic digestion, and the coverage shows the sequence coverage of the peptides in the mass range between	

500 to 3000 Da. (T: trypsin, C: chymotrypsin, T/C: trypsin and chymotrypsin)	170
Table 19. MASCOT MS/MS ion search results of silk samples (tryptic peptides, no result for <i>A. assamensis</i>); the residue marked in red indicates the unique amino acid that differentiate <i>B. mori</i> and <i>B. mandarina</i>	174
Table 20. PEAKS/ <i>De novo</i> analysis result of silk samples (tryptic digestion).....	175
Table 21. Summary of PEAKS analysis results of <i>Bombyx mori</i> silk solubilised with different reagents, in-solution tryptic digestion	177
Table 22. Summary of PEAKS analysis results of <i>Antheraea</i> and <i>Samia</i> silks solubilised with different reagents, in-solution tryptic digestion.	181
Table 23. Protein identification result of <i>A. mylitta</i> sample	187
Table 24. Theoretical tryptic cleavages of E1CGA3, residues 933-2009 (missed cleavage:0).....	197
Table 25. PEAKS analysis results of <i>Bombyx</i> silks treated with different enzymes	208
Table 26. PEAKS analysis results of <i>A. pernyi</i> fibroin treated with different enzymes....	228
Table 27. PEAKS analysis results of <i>A. mylitta</i> fibroin treated with different enzymes...	244
Table 28. PEAKS analysis results of <i>A. yamamai</i> fibroin treated with different enzymes	254
Table 29. PEAKS analysis results of <i>A. assamensis</i> fibroin treated with different enzymes	268
Table 30. PEAKS analysis results of <i>S. ricini</i> fibroin treated with different enzymes.....	280
Table 31. Naintr� site textile samples and their solubilisation behaviour	299
Table 32. Naintr� site samples: a summary of PEAKS results (Peptides -10lgP�20).....	300
Table 33. Comparison of protein identification results of N 2747 (<i>B.mori</i>) by PEAKS using <i>De novo</i> vs. UniProtKB (peptides -10lgP�15).....	303
Table 34. PEAKS result summary of the Palmyra samples (FDR=0%)	311
Table 35. Descriptions of the preserved leaf samples.....	343
Table 36. Carbon and Nitrogen distribution of the samples	345
Table 37. Summary of MASCOT search results (Database: NCBIInr).....	367

Table 38. Unique peptides to <i>B. mori</i> heavy chain.....	371
Table 39. Peptide distribution of CN and CN-FA samples of <i>A. pernyi</i> silk.....	376
Table 40. Peptide distribution of CN and CN-FA samples of <i>A. mylitta</i> silk.....	377
Table 41. Peptide distribution of CN and CN-FA samples of <i>A. yamamai</i> silk	378
Table 42. Peptide distribution of CN and CN-FA samples of <i>A. assamensis</i> silk.....	379
Table 43. Peptide distribution of CN and CN-FA samples of <i>S. ricini</i> silk.....	384

List of Figures

Figure 1. The proposed findings of early silks outside China, pre-Han to early Han times.	3
Figure 2. Cross-section of three threads from a textile recovered from a ca. 7th century BCE grave at Kerameikos, Greece. <i>Photomicrograph by I. Good</i>	5
Figure 3. A picture of Xiyin-cun cocoon (above) and <i>Rodontia mencia</i> , <i>Theophilina mandarina</i> and <i>Bombyx mori</i> cocoons in comparison (below from the left).	7
Figure 4. A general flowchart for a conventional fibre analysis.....	10
Figure 5. ATR-FT-IR spectrum of different fibres	11
Figure 6. Basics of mass spectrometry instrumentation	15
Figure 7. Two common approaches for mass spectrometry based proteomic analysis: peptide mass fingerprinting (A) and peptide sequencing (B).	17
Figure 8. Distribution of non-mulberry silk fauna	31
Figure 9. The adults (above) and cocoons (below) of <i>Bombyx mandarina</i> (left) and <i>Bombyx mori</i> (right).	34
Figure 10. Origin and dispersal of <i>A. pernyi</i> cultivation and current level of production ..	37
Figure 11. <i>Antheraea</i> moths and their cocoons: from the top <i>A. pernyi</i> , <i>A. mylitta</i> and <i>A. yamamai</i>	39
Figure 12. <i>A. assamensis</i> moth and cocoon.	41
Figure 13. <i>S. ricini</i> moth and cocoons of different colourmorphs	43
Figure 14. <i>A. atlas</i> moth and cocoons	45
Figure 15. <i>A. panda</i> and <i>A. reticulata</i> (below) moths and diagram of <i>Anaphe</i> nest cocoon	47
Figure 16. Votive panel illustration (木版畫) depicting the legend of the introduction of sericulture to Khotan. The central figure is a Chinese princess who smuggled mulberry seeds and eggs of the silk moth to Khotan, hiding them in her head. (British Museum 1907, 1111 .731907, 1111. 73)	67

Figure 17. Coa Vestis (Mus.Borbon.VIII.5.)	79
Figure 18. SEM images of raw <i>B. mori</i> silk cross-section (left) and longitudinal section (right)	88
Figure 19. Structural hierarchy of a silk fibre.....	90
Figure 20. Morphologies of the different silk species: cocoon, cross section and longitudinal section images.	91
Figure 21. Schematic crystal structure of cocoon silk fibroin	92
Figure 22. Mechanisms of silk molecular orientation and formation of hydrogen bonding.....	94
Figure 23. Structure of proteins	95
Figure 24. Schematic primary structure of <i>B. mori</i> heavy chain: the amino acid sequence for one of the crystalline domains is given to highlight the repetitive nature of the protein	99
Figure 25. Flowchart of experimental works.....	103
Figure 26. Re-aggregated proteins after sitting the samples for 24 hours at 4°C. left) fibre weight to solvent ratio at 1:1 right) at 1:1.5, with formic acid (1%)	133
Figure 27. zwitterion in simple amino acid solutions and the acid reaction.....	134
Figure 28. Yellowing of silk fibroin (<i>S. ricini</i>) solution by formic acid: 0.5mg fibre in each tubes with A) x1 7M Ca(NO ₃) ₂ , B) x1 7M Ca(NO ₃) ₂ +0.1(o.w.f.)% FA, C) x2 7M Ca(NO ₃) ₂ , D) x2 7M Ca(NO ₃) ₂ +0.1(o.w.f.)% FA. First, the solubilisation was implemented with the salt solution at 125°C for 10 minutes, and then FA was added off the heat.	135
Figure 29. Xanthoprotic reaction of tyrosine and the formation of 3-nitro-tyrosine.....	136
Figure 30. Separation of proteins in a polyacrylamide gel	142
Figure 31. SDS-PAGE result of fibroin solutions (3-8% Tris-Acetate gel, 40-400kDa): only the dye front reacted to the staining.....	144
Figure 32. SDS-PAGE result of fibroin solutions (TGX 10% gel, 15-200kD).....	146
Figure 33. SDS-PAGE result of standard proteins and fibroin solutions: 3-8% Tris-Acetate gel (40-400kDa), Coomassie blue stained.....	149

Figure 34. SDS-PAGE result of standard proteins and fibroin solutions: 3-8% Tris-Acetate gel (40-400kDa), partially silver stained over Coomassie blue.	149
Figure 35. Unique sequence features of <i>B. mori</i> fibroin heavy chain (P05790): unique residues other than the GA(X) sequences are not marked here.	157
Figure 36. Alignment of <i>Antheraea</i> and <i>Samia</i> silk fibroins: unique peptides to each protein are marked in colour: <i>A. pernyi</i> (purple); <i>A. mylitta</i> (green); <i>A. yamamai</i> (yellow); <i>A. assamensis</i> (red) and <i>S. ricini</i> (blue). The main colour scheme is by Blosum62.	167
Figure 37. Protein coverage of the heavy chain (P05790) by <i>B. mori</i> fibroin-HCl solution.	178
Figure 38. Protein coverage of the heavy chain (P05790) by <i>B. mori</i> fibroin-calcium nitrate solution.	178
Figure 39. Protein coverage of the heavy chain (P05790) by <i>B. mori</i> fibroin-calcium nitrate/formic acid (10%) solution.	180
Figure 40. Spectra of the 3 unique peptides to <i>A. pernyi</i> fibroin detected in both CN (A) and CN-FA (B) samples: 1) AAGSA ₁₂ SGAGR ($m/z = 793.41$) 2) AAGSA ₁₃ SGAGGAGGG ($m/z = 900$) and 3) SAGSA ₁₃ SGAGG ($m/z = 787.38$), $z=2$	183
Figure 41. Protein coverage of <i>A. pernyi</i> fibroin (O76786) - calcium nitrate solution.....	184
Figure 42. Protein coverage of <i>A. pernyi</i> fibroin (O76786) - calcium nitrate/ formic acid (10%) solution.....	186
Figure 43. Spectra of the 4 unique peptides to <i>A. mylitta</i> fibroin detected in both CN (A) and CN-FA (B) samples: 1) GSDSGA ₁₃ SGAGGR ($m/z = 915.94$) 2) SGA ₁₃ SGAGGRGD ($m/z = 872.42$) 3) GSGSSA ₁₁ SAAR ($m/z = 780.89$), and 4) VDSHGQLVER ($m/z = 570.29$), $z=2$	189
Figure 44. Protein coverage of <i>A. mylitta</i> fibroin (Q8ISB3)- calcium nitrate solution.	190
Figure 45. Protein coverage of <i>A. mylitta</i> fibroin (Q8ISB3) - calcium nitrate/ formic acid (10%) solution.....	191
Figure 46. Spectra of the 4 unique peptides detected in both CN (A) and CN-FA (B) samples: 1) AAGSA ₁₃ SGAGGSGR ($m/z = 929.46$), 2) AAGSA ₁₃ SGAGR	

($m/z=828.92$), 3) EEDVVINR ($m/z=487.25$) and 4) EDVVINR ($m/z=422.73$), $z=2$.	194
Figure 47. Protein coverage of <i>A. yamamai</i> fibroin (E1CGA3) - calcium nitrate solution.....	195
Figure 48. Protein coverage of <i>A. yamamai</i> fibroin (E1CGA3) - calcium nitrate/ formic acid (10%) solution.	196
Figure 49. Spectra of the unique peptides detected in both CN (A) and CN-FA (B) samples: 1) VPGASSSAAAASSASAGGR ($m/z=781.38$, $z=2$), 2) SA ₁₃ GSGGR ($m/z=722.37$, $z=2$), 3) AAGSA ₁₂ GGSGAGG ($m/z=800.88$, $z=2$) and 4) AAGSA ₁₂ GGSGAGGHGGG ($m/z=636.97$, $z=3$).	200
Figure 50. Protein coverage of <i>A. assamensis</i> fibroin (A0A0K0KR73) - calcium nitrate solution.....	200
Figure 51. Protein coverage of <i>A. assamensis</i> fibroin (A0A0K0KR73) - calcium nitrate /formic acid (10%) solution.	201
Figure 52. Spectra of the unique peptides detected in CN (A) and CN-FA (B) samples: 1) QVPQGGAASSAASSASAGSGSGAPTIIVER ($m/z=891.11$), 2) VHETGDGFLLR ($m/z=622.32$), 3) SSAASSASAGSGSGAPTIIVER ($m/z=981.99$), $z=2$.	203
Figure 53. Protein coverage of <i>S. ricini</i> fibroin (A0A0D5ZYI3) - calcium nitrate solution.	204
Figure 54. Protein coverage of <i>S. ricini</i> fibroin (A0A0D5ZYI3) - calcium nitrate/formic acid (10%) solution.	206
Figure 55. Spectra of the unique peptides to <i>B. mori</i> fibroin heavy chain (P05790): A) tryptic in-gel digestion, B) in-solution digestion; 1) GYGQGAGSAASSVSSASSR ($m/z= 843.89$), 2) DASGAVIEEQITTK ($m/z= 731.38$), and 3) SYDYSR ($m/z= 395.67$), $z=2$.	210
Figure 56. Protein coverage of <i>B. mori</i> heavy chain (P05790): tryptic in-gel digestion...	210
Figure 57. Protein coverage of <i>B. mori</i> heavy chain (P05790): tryptic in-solution digestion.	211
Figure 58. Spectra of the unique peptides to <i>B. mori</i> fibroin heavy chain (P05790): A) chymotryptic in-gel digestion, B) in-solution digestion; 1) DASGAVIEEQITTK ($m/z= 731.38$), 2) GAGAGAGAGTGSSGFGPY ($m/z= 721.32$), 3)	

AASGAGAGAGAGAGTGSSGFGPY (m/z= 899.90), and 4) SGAGAGAGAGAGAGTGSSGFGPY (m/z= 828.87), z=2	213
Figure 59. Protein coverage of <i>B. mori</i> heavy chain (P05790): chymotryptic in-gel digestion.....	215
Figure 60. Protein coverage of <i>B. mori</i> heavy chain (P05790): chymotryptic in-solution digestion.....	216
Figure 61. Spectra of the unique peptides to <i>B. mori</i> fibroin heavy chain (P05790): A) tryptic -chymotryptic in-gel digestion, B) in-solution digestion; 1) GAGAGAGAGTGSSGFGPY (m/z= 721.32), 2) GAGAGAGAGTGSSGFGPY (m/z= 721.32), 3) AGAGAGAGTGSSGFGPY (m/z= 692.8), and 4) VIEEQITTK (m/z= 530.8), z=2.....	218
Figure 62. Protein coverage of <i>B. mori</i> heavy chain (P05790): tryptic-chymotryptic in-gel digestion.....	222
Figure 63. Protein coverage of <i>B. mori</i> heavy chain (P05790): tryptic-chymotryptic in-solution digestion.....	224
Figure 64. Unique marker peptide for Q9BLL8: GQIPSQYEIPVFR (m/z=767.40, z=2) in tryptic in-gel (A) and in-solution (B) digestion.	225
Figure 65. Unique marker peptides for Q9BLL8 in chymotryptic in-gel (A) and in-solution (B) digestion: 1) EIPVFR (m/z=380.72) 2) GQIPSQYEIPVFR (m/z=767.40) and 3) TLCDFGCHH (m/z=573.73), z=2.	227
Figure 66. Unique marker peptide for Q9BLL9: DAILASSDSGFLNEEY (m/z=866.38, z=2) in tryptic-chymotryptic in-gel (A) and in-solution (B) digestion.	227
Figure 67. Spectra of tryptic unique peptides of <i>A. pernyi</i> fibroin: 1) AAGSA ₁₂ SGAGR (m/z=793.40), 2) AAGSA ₁₃ SGAGGAGGG (m/z=900.43), 3) GSGSAAAAAGSGAGGSGGG (m/z=653.78), z=2	230
Figure 68. Protein coverage of <i>A. pernyi</i> fibroin (O76786): tryptic in-gel digestion.	231
Figure 69. Protein coverage of <i>A. pernyi</i> fibroin (O76786): tryptic in-solution digestion.	232
Figure 70. Spectra of chymotryptic unique peptides of <i>A. pernyi</i> fibroin: 1) GSDSA ₁₃ GSGAGGSGGY (m/z=680.30), 2) GSDSA ₁₃ SGAGGSGGY (m/z=661.30), 3) GSDSSA ₁₀ SGAGGSGGGY (m/z=638.23), z=3	234

Figure 71. Protein coverage of <i>A. pernyi</i> fibroin (O76786): chymotryptic in-gel digestion.	236
Figure 72. Protein coverage of <i>A. pernyi</i> fibroin (O76786): chymotryptic in-solution digestion.....	238
Figure 73. Spectra of tryptic-chymotryptic unique peptides of <i>A. pernyi</i> fibroin: 1) SDSA14SGAGGSGGYGG($m/z=1055.48$), 2) GSDSSA10SGAGGSGGGY ($m/z=956.92$), 3) SA12SGAGRS($m/z=737.37$), 3) SYEASRISVH ($m/z=574.79$), $z=2$	240
Figure 74. Protein coverage of <i>A. pernyi</i> fibroin (O76786): tryptic-chymotryptic in-gel digestion.....	241
Figure 75. Protein coverage of <i>A. pernyi</i> fibroin (O76786): tryptic-chymotryptic in-solution digestion.....	243
Figure 76. Spectra of tryptic unique peptides of <i>A. pernyi</i> fibroin: 1) GAGGAAGAAAGA AASSSVR ($m/z=628.31$, $z=3$), 2) QASHGAGGAAGAAAGAAASSSVR ($m/z=730.37$, $z=2$)	245
Figure 77. Protein coverage of <i>A. mylitta</i> fibroin (Q8ISB3): tryptic in-gel digestion.....	246
Figure 78. Protein coverage of <i>A. mylitta</i> fibroin (Q8ISB3): tryptic in-solution digestion.....	246
Figure 79. Spectra of chymotryptic unique peptides of <i>A. mylitta</i> fibroin: 1) VDSHGQLVERF ($m/z=643.83$, $z=2$), 2) DEYVDSHGQLVERF ($m/z=565.27$, $z=3$), 3) GRGDGGYGSDSA ₁₂ GSGAG GRGDGGY ($m/z=957.76$, $z=2$) , 4) GGYGSDSGA ₁₃ SGAGGRGD ($m/z=1140.52$, $z=2$)	248
Figure 80. Protein coverage of <i>A. mylitta</i> fibroin (Q8ISB3): chymotryptic in-gel digestion.	249
Figure 81. Protein coverage of <i>A. mylitta</i> fibroin (Q8ISB3): chymotryptic in-solution digestion.....	250
Figure 82. Spectra of tryptic-chymotryptic peptides of <i>A. mylitta</i> fibroin: 1) GSDSGA ₁₃ SGAGGR ($m/z=915.94$), 2) RVPGASSSAAAASSASAGSGQ ($m/z=888.93$), 3) GSDSA ₉ SGAGGR ($m/z=745.35$), 4) SGA ₁₃ SGAGGRGD ($m/z=872.42$), $z=2$	252

Figure 83. Protein coverage of <i>A. mylitta</i> fibroin (Q8ISB3): tryptic-chymotryptic in-gel digestion.....	253
Figure 84. Protein coverage of <i>A. mylitta</i> fibroin (Q8ISB3): tryptic-chymotryptic in-solution digestion.....	253
Figure 85. Spectra of tryptic unique peptides of <i>A. yamamai</i> fibroin: 1) AAGSA ₈ SGAGR ($m/z=828.92$), 2) AAGSA ₁₃ SGAGGSGR ($m/z=929.46$), 3) SAAGA ₁₀ SGAGGRGD ($m/z=723.36$), 4) SAAGA ₁₀ SGAGGRGD ($m/z=836.9$), $z=2$	256
Figure 86. Protein coverage of <i>A. yamamai</i> fibroin (E1CGA3): tryptic in-gel digestion..	257
Figure 87. Protein coverage of <i>A. yamamai</i> fibroin (E1CGA3): tryptic in-solution digestion.....	258
Figure 88. Spectra of chymotryptic unique peptides of <i>A. yamamai</i> fibroin: 1) GSESA ₁₀ GSGAGGVGGGY ($m/z=953.94$), 2) VVINRVPGASSSA ₄ SSASAG ($m/z=965.50$), 3) GGYGSDSA ₁₃ SGAGGRGD ($m/z=1112.0$), $z=2$; 4) SA ₁₁ GSGAGGVGGGYGRGD ($m/z=698$), $z=3$	260
Figure 89. Protein coverage of <i>A. yamamai</i> fibroin (E1CGA3): chymotryptic in-gel digestion.....	261
Figure 90. Protein coverage of <i>A. yamamai</i> fibroin (E1CGA3): chymotryptic in-solution digestion.....	263
Figure 91. Spectra of unique tryptic-chymotryptic peptides of <i>A. yamamai</i> fibroin: 1) GSDSA ₁₃ GSGAGR ($m/z=887.43$, $z=2$), 2) AAGSA ₁₃ SGAGGSGR ($m/z=619.98$, $z=3$), 3) GGYGSDSA ₁₃ SGAGGRGD ($m/z=1112$, $z=2$), 4) GSA ₁₀ GAGASRPVGIY ($m/z=634.67$, $z=3$).....	265
Figure 92. Protein coverage of <i>A. yamamai</i> fibroin (E1CGA3): tryptic-chymotryptic in-gel digestion.....	266
Figure 93. Protein coverage of <i>A. yamamai</i> fibroin (E1CGA3): tryptic-chymotryptic in-solution digestion.....	267
Figure 94. Spectra of tryptic unique peptides of <i>A. assamensis</i> fibroin: 1) RVPGASSSA ₄ SSASAGGR ($m/z=573.29$, $z=3$), 2) VPGASSSA ₄ SSASAGGR ($m/z=781.38$, $z=2$), 3) SFVVDSS (740.35 , $z=1$), 4) A ₉ GSGGR ($m/z=536.78$, $z=2$) .	270
Figure 95. Protein coverage of <i>A. assamensis</i> fibroin (A0A0K0KR73): tryptic in-gel	

digestion.....	271
Figure 96. Protein coverage of <i>A. assamensis</i> fibroin (A0A0K0KR73): tryptic in-solution digestion.....	272
Figure 97. Spectra of unique peptides for <i>A. assamensis</i> fibroin: 1) GSDSA ₁₂ GSGGRGSGDGY ($m/z=1084.5$), 2) GSGSA ₁₃ GSGAGGAGD ($m/z=930.43$), 3) VPGASSSA ₄ SSASAGGR ($m/z=781.38$), $z=2$	274
Figure 98. Protein coverage of <i>A. assamensis</i> fibroin (A0A0K0KR73): chymotryptic in-gel digestion.....	275
Figure 99. Protein coverage of <i>A. assamensis</i> fibroin (A0A0K0KR73): chymotryptic in-solution digestion.....	276
Figure 100. Spectra of unique tryptic-chymotryptic peptides of <i>A. assamensis</i> fibroin: 1) GDGGYGSDSA ₁₂ GSGAGGS ($m/z=1070.47$), 2) VPGASSSA ₄ SSASAGGR ($m/z=781.38$), 3) DSA ₁₃ GSGGRGSG ($m/z=880.42$), 4) DSA ₁₂ GSGGRGSGD ($m/z=902.41$), $z=2$	278
Figure 101. Protein coverage of <i>A. assamensis</i> fibroin (A0A0K0KR73): tryptic-chymotryptic in-gel digestion.....	279
Figure 102. Protein coverage of <i>A. assamensis</i> fibroin (A0A0K0KR73): tryptic-chymotryptic in-solution digestion.....	279
Figure 103. Spectra of tryptic unique peptides of <i>S. ricini</i> fibroin: 1) QVPQGGAASSAASSASAGSGSGAPTIIVER ($m/z=891.11$, $z=3$), 2) QVPQGGAASSAASSASAGSGSGAPTIIVER ($m/z=1336.66$), 3) PQGGAASSAASSA ($m/z=750.86$), 4) DSA ₁₃ GGAGGG ($m/z=531.25$), 5) GGA ₁₃ GGAGSG 6) SGA ₁₃ GGAGD ($m/z=722.34$), $z=2$	283
Figure 104. Protein coverage of <i>S. ricini</i> fibroin (A0A0D5ZYI3): tryptic in-gel digestion.....	283
Figure 105. Protein coverage of <i>S. ricini</i> fibroin (A0A0D5ZYI3): tryptic in-solution digestion.....	284
Figure 106. Spectra of unique chymotryptic peptides of <i>S. ricini</i> fibroin: 1) GGDGGA ₁₃ GGSGSGYG ($m/z=954.43$), 2) GAGSA ₁₂ GGAGGG ($m/z=750.36$), 3) GGA ₁₃ GSGGGYGGD ($m/z=882.4$), 4) ASSASAGSGSGAPTIIVER ($m/z=859.44$)	

.....	286
Figure 107. Protein coverage of <i>S. ricini</i> fibroin (A0A0D5ZYI3): chymotryptic in-gel digestion.....	287
Figure 108. Protein coverage of <i>S. ricini</i> fibroin (A0A0D5ZYI3): chymotryptic in-solution digestion.....	288
Figure 109. Spectra of tryptic-chymotryptic unique peptides of <i>S. ricini</i> fibroin: 1) DGGA ₁₃ GGAGSGYGGGAR (m/z=1056), 2) DSGA ₁₂ SGGAGSGYGG(m/z=940.92), 3) GGYGHGYGSDGGAAA (m/z=648.77), 4) GGYGHGYGSDGG(m/z=542.21), 5) GGA ₁₃ GSGGGYGGD (m/z=882.40).....	290
Figure 110. Protein coverage of <i>S. ricini</i> fibroin (A0A0D5ZYI3): tryptic-chymotryptic in-gel digestion.....	291
Figure 111. Protein coverage of <i>S. ricini</i> fibroin (A0A0D5ZYI3): tryptic-chymotryptic in-solution digestion.....	292
Figure 112. Drawing of the child's grave. A) textile around the head B, C) cloth along each legs D) outer wrapping in silk damask; and underneath this, E) large bands in gold tapestry placed on each side underneath the damask along the entire length of the body F) a narrower band with gold tapestry.	295
Figure 113. Different types of textiles found in Naintr� samples.....	297
Figure 114. Microscopic images of the Palmyra samples: top) Sample S 8; Schmidt-Colinet/Stauffer Katalog Nr. 305, middle) Sample S48; Schmidt-Colinet/Stauffer Katalog Nr. 312, bottom) Sample S49; Schmidt-Colinet/Stauffer Katalog Nr. 313.307	
Figure 115. Before (left) and after (right) the solubilisation in 7M Ca(NO ₃) ₂ solution: sample S8-weft, taken after 20mins of boiling	310
Figure 116. Spectra of the highest scoring unique peptides for Q8ISB3 in S8 warp: A) SA ₁₂ GSGAGGRGD (m/z=780.4), B) SSA ₁₁ SAAR(m/z=680.35), z=2.	312
Figure 117. Spectrum of the highest scoring unique peptide for Q8ISB3 in S8 weft: GGYGSGSSA ₁₁ SAAR (m/z=919.44, z=2).....	313
Figure 118. Spectra of the highest scoring unique peptides to Q8ISB3 in S48: A) SGSSA ₁₁ SAAR (m/z=752.38), B) SGA ₁₂ SGAGGRGD (m/z=872.42), z=2	315

Figure 119. Spectrum of the highest scoring unique peptide for Q8ISB3 in S49-T1: GSGSSA ₁₁ SAAR ($m/z=780.89$, $z=2$).....	316
Figure 120. Spectrum of the highest scoring unique peptide for Q8ISB3 in S49-T2: SA ₁₁ SAAR ($m/z=636.84$, $z=2$).....	317
Figure 121. Spectrum of GSDSA ₁₃ GSGAGGR (a unique peptide to O76786) in all samples: mass = 1829.86; $m/z=610.96(z=3)$ or $915.93(z=2)$	318
Figure 122. Flowchart for initial silk screening based on the solubilisation behaviour. ...	335
Figure 123. Geographical distribution of the samples and their climatic data	342
Figure 124. Distribution of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ value of the mulberry samples.....	346
Figure 125. Distribution of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ value of <i>Morus A.</i> samples from various locations (Left).....	347
Figure 126. Distribution of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ value of <i>Morus Kagayama K.</i> samples from various locations (Right).....	347
Figure 127. Distribution of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ value of Mulberry varieties from UK.....	348
Figure 128. Distribution of $\delta^{15}\text{N}$ value of samples by yearly mean temperature ($^{\circ}\text{C}$).....	349
Figure 129. Sequence alignment of the <i>Bombyx</i> , <i>Antheraea</i> and <i>Samia</i> silks by MUSCLE algorithm (v.3.8.31) using Jalview (2.10.3b1). The colour scheme is after Blosum62: if a residue matches the consensus sequence residue at that position it is coloured dark blue. If it does not match the consensus residue but the 2 residues have a positive Blosum62 score, it is coloured light blue.....	365

Note for citations in foreign languages

Chinese Classics

All Chinese texts from ancient literatures are sourced from the Chinese Text Project Digital Library (<http://ctext.org/>). Thus, an individual in-text citation is omitted but the database is listed in the Bibliography. In text, the book titles are listed in their English translation with their original titles in Chinese are given in (). In the footnote, their *pinyin* and the name of the author and the time of publication are also given following the same format. If the title does not have an appropriate English equivalent, its better-known name (often its *pinyin*) was used.

Latin and Greek Classics

All ancient Greek and Latin texts cited here are sourced from the Loeb Classical Library (<https://www.loebclassics.com/>, Harvard University) and Perseus Digital Library (<http://www.perseus.tufts.edu/hopper/>, Tuft University), unless separately mentioned in the footnote. In the text, name of the writer, title of the book, volume and chapter/paragraph numbers are given at the end of the citation. The volume number is shown in Roman numeral, and the chapter/paragraph information is in Arabic numeral. The original text is given in the footnote. Both references are listed in the bibliography.

I. Introduction

I.1 Background

Silk fibre is commonly produced by the mulberry silkworm, *Bombyx mori*, which accounts for 95% of the modern silk production.¹ Although many other silk varieties were used for clothing throughout history, *B. mori* is the most favoured species due to its ability to produce a uniform, white fibre that can easily be reeled out to a continuous filament. However, such properties of *B. mori* only materialise through a strictly controlled growth environment: *B. mori* is the only completely domesticated species of silkworm reared by human hands, whereas other species occur in wild or semi-domesticated cultivation.

Thus, in an archaeological context, the proof of sericulture could be an index of cultural and technological development. The establishment of sericulture indicates that a community had access to resources and technology, developing economies of scale to support the domesticated silk production. Nonetheless, it must be noted that not every finding of *B. mori* silk in archaeological sites is linked to the cultivation of silk: the influx of silk products most likely preceded the introduction of sericulture. While wild silks could have been collected and used much earlier, the technology to breed and use silkworms from a domesticated stock did not exist before the late Chinese Neolithic.² Thus, these arguments must be carefully made, supported by reliable scientific analysis, a comprehensive study of the literary sources on the local use of wild silks and the introduction of Chinese sericulture.

¹ Panda, 2010

² Ahmad, *et al.*, 2016

I.1.1 Archaeological silk controversies

The identification of silk from archaeological sources has been a controversial subject, largely due to the limitations of conventional methods of analysis. Jorgensen has discussed the series of disputes concerning the identification of silk in prehistoric findings in Europe in the Antiquity, particularly over those cited in the series of studies by Good.³ Good published a study on pre-Han findings of silks in Antiquity in 1995 and listed some of the European sites (Figure 1).⁴ However, these textiles have either never been identified as silk, or confirmed as being other fibres later. In personal communications with Jorgensen, she indicated her position that these ‘silks’ have found their way into several standard overviews of the European Iron Age, causing a misconception of trade connection. She intended to start a debate with the help of Antiquity’s editor to invite Good and the German team, but it did not happen as Good passed away in 2013.

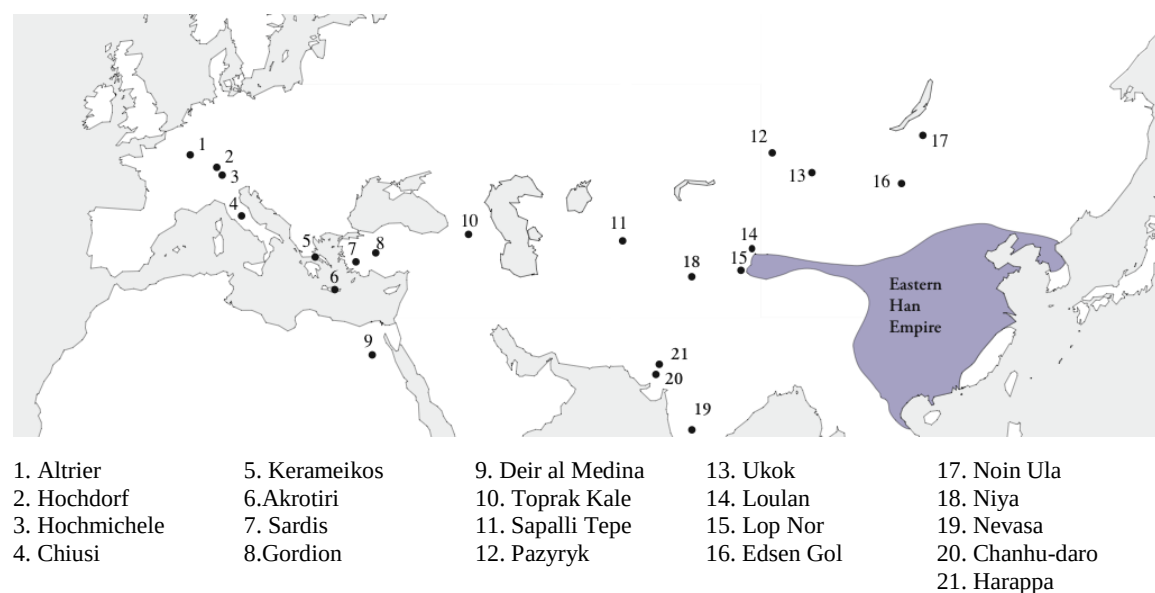


Figure 1. The proposed findings of early silks outside China, pre-Han to early Han times.⁵

³ L.B. Jorgensen, 2013

⁴ Good, 1995

⁵ Good, 2002

The textiles from Altrier (Fig1. 1) and Chuisi (Fig1. 4) were never identified as silk.⁶ The Altrier textiles were identified as wool in several studies.⁷ However, Hundt's unpublished personal thoughts about the possibility of these being silk was cited in Wild, but with caution,⁸ and this was adopted by Barber and Good.⁹ The Chiusi textile was examined by John Hayes at the Royal Ontario Museum, with the comment "a very fine long-stapled wool or, in view of the filament-like appearance of the weft, more probably silk".¹⁰ Good took his view without further advance.¹¹

The possibility of the Hochdorf (Fig1. 2), Hochmichele (Fig1. 3) and Kerameikos (Fig1. 5) textiles being silk was initially suggested by Hundt.¹² Good examined these textiles samples using amino acid analysis, and concluded that these samples were silk, and claimed that the Hochmichele and Kerameikos textiles are a type of wild silk—possibly the *Pachypasa otus*.¹³ Good provided the results of SEM (scanning electron microscope) analysis to support this argument, but as a single cross-section image of the sample fibre in a poor resolution, which was compared to a longitudinal reference image of *P. otus* (Figure 2).¹⁴ Later, Hochdorf and Hochmichele textiles were refuted by Banck-Burgess and Mann, who confirmed that they are not silk.¹⁵ The re-investigation of the alleged silks using amino acid analysis showed that both the Hochdorf and Hochmichele samples are cellulosic bast fibres.¹⁶ Although Good's later publication includes this update, she held her position about the identity of the fibres.¹⁷ The Kerameikos textiles were also re-

⁶ L.B. Jorgensen, 2013

⁷ Heyart, 1972; Hundt, 1974; Thill, 1987

⁸ Wild, 1984

⁹ Barber, 1991; Good, 1995

¹⁰ Hayes, 1977

¹¹ Good, 1995

¹² Hundt, 1985; Hundt, 1969

¹³ Good, 1995, Good, 2010, Good, 2011. See II.1.4.2 for further information.

¹⁴ Good, 2011

¹⁵ Banck-Burgess, 2012

¹⁶ Mann, 1999

¹⁷ Good, 2011

examined by Margariti *et al.* of the Ministry of Culture, Greece: the results from ESEM (environmental scanning electron microscope) and FTIR (Fourier-transform infrared spectroscopy) analysis indicated the samples are cellulosic bast fibres.¹⁸ The artefact from Akrotiri (Fig1. 6) is a fossil cocoon dated to the middle of the second millennium BCE. Panagiotakopulu *et al.* discussed this as presumably coming from either *Saturnia pyri* or *Pachypasa otus* based on the size and morphology; if true, this is entomological evidence that these native silk moths inhabited Akrotiri and elsewhere on the island of Santorini during the Bronze Age.¹⁹ However, no further examination of this cocoon has been carried out, and its identity remains unknown.

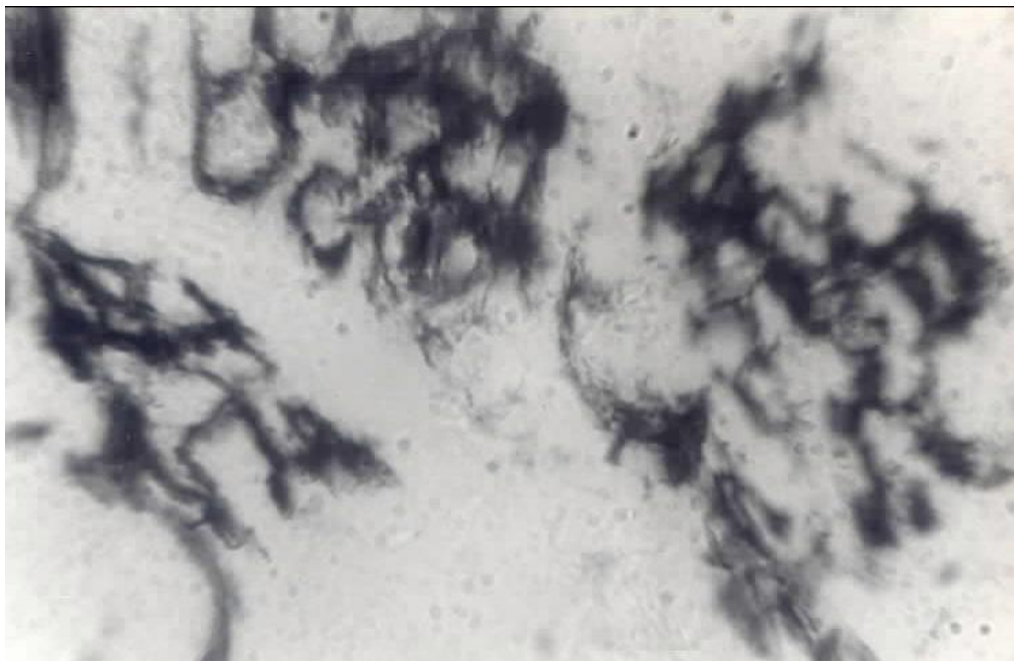


Figure 2. Cross-section of three threads from a textile recovered from a ca. 7th century BCE grave at Kerameikos, Greece. Photomicrograph by I. Good²⁰

¹⁸ Margariti, *et al.*, 2011

¹⁹ Panagiotakopulu, *et al.*, 1997

²⁰ Good, 2011 #504

In addition to the mystery of early European silks, there are also some unresolved cases among the early silk findings from China and India. One of them is the Xīyīncūn cocoon (in Taipei National Museum), known as the earliest direct evidence of cultivated silk cocoon in use. It is an artificially cut cocoon excavated by Li Chi (李濟) in 1926 at the Xīyīncūn site of the Yangshao culture, in Shanxi, China. It was dated to c. 2200-1700 B.C.E. It was identified as a *Bombyx mori* cocoon by Liu Chongle (刘崇乐), a professor of biology at Tsinghua University(清华大学),²¹ and confirmed by the Smithsonian Institution in 1928. However, there is no record left, and the means of analysis is unknown.²² The identity of this cocoon remains as a source of controversy for scholars: some have argued that it might have been left at the site by later people, or that a single finding is not enough to confirm the establishment of sericulture.²³ In 1967, Nunome published a study on this cocoon, with a reconstructed a model based on the morphological analysis from a photograph.²⁴ He argued that the cocoon is only cut to 17% of its size, and is of *Rondotia menciiana* Moore, a wild species native to China not presently existing in the north of China.²⁵ Based on Li's note on the size of the cocoon, "considerably smaller than even the smallest of silkworms raised at Xiyong village today", and Nunome's drawing, the identity of this cocoon fragment is more likely to be that of a wild species (Figure 3). It is fascinating, yet there is no means to clarify the identity of this cocoon – given the lack of scientific data to support either claim. The size of the cocoon must have evolved after

²¹ Li, 1927

²² Nunome, 1968.

“中国中央研究院歴史語言研究所の石璋如氏からの書簡によると(...)記録は残っていないようである。恐らく詳細な復元を経ることなく外部観察のみによって証言せられたものと思われる。”

²³ Kuhn, *et al.*, 2012

²⁴ According to Nunome, it was the first time that cocoon fragment was ever photographed at the Palace Museum in Taipei: “...保管者側の言によれば、曾てこの繭の写真を撮らせたことがなかったと...”

²⁵ Nunome, 1968

thousand years of domestication; based on the photograph the cocoon looks quite thin, thus, it might have been unreeled already; furthermore, it could have been cut because the person wanted to eat the chrysalis.

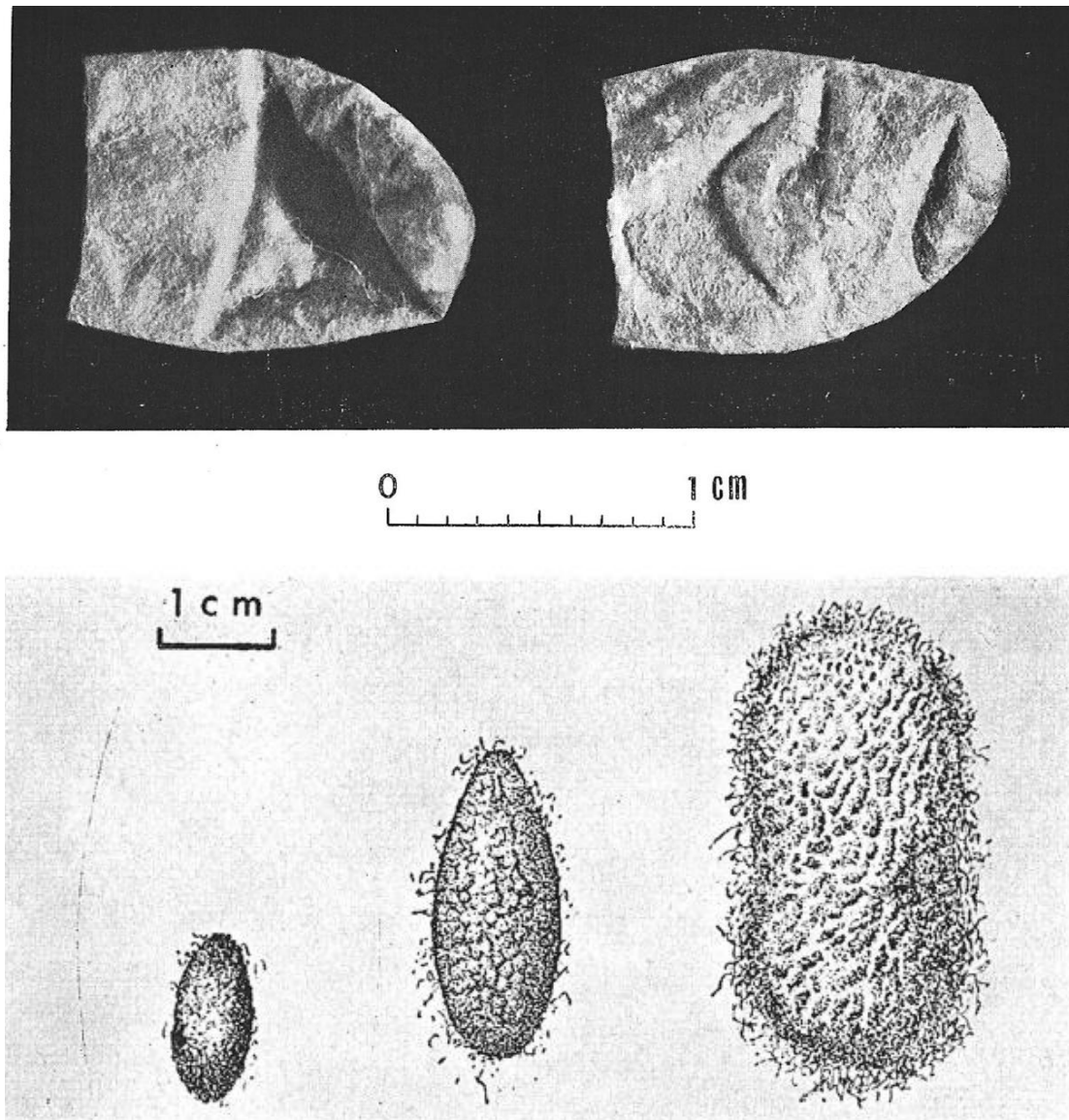


Figure 3. A picture of Xiyin-cun cocoon (above) and *Rodontia menciana*, *Theophilia mandarina* and *Bombyx mori* cocoons in comparison (below from the left).²⁶

²⁶ Ibid. The drawing is the reproduction from Pariset, E (1800).

From India, there are potential wild silk discoveries that demonstrate the use of indigenous silks in South Asia for more than 2000 years before the introduction of domesticated silk from China. The Nevasa finding is the earliest direct archaeological evidence for silk in South Asian prehistory (1500-1050 BCE).²⁷ The microscopic examination of the string threading of copper beads was reported as ‘natural’ silk, but it is noted that it may not be from Chinese silk because of the short fibre lengths.²⁸ However, the identity of the silk fibre is unknown. In 2009, Good *et al.* suggested that some of the threads from the collection of Harappa and the Chanhudaro site excavations,²⁹ belonging to the Indus Civilisation (2800-1900 BCE), were of wild silk species. Based on the morphological analysis by SEM imaging, it was proposed that the Harappa fibres are of *Antheraea mylitta* and *assamensis*, and Chanhudaro fibres are of *Samia ricini*.³⁰ However, this examination is solely based on SEM imaging, and the evidence presented is a limited quality image that is sufficient to obscure the argument.

²⁷ Allchin and Allchin, 1982

²⁸ Gulati, 1961; Janaway and Coningham, 1995; Alam, 1999

²⁹ A collection of Museum of Fine Arts, Boston.

³⁰ Good, *et al.*, 2009

I.1.2 Conventional methods to characterise silks and their limitations

Archaeological textiles represent a variety of structures and materials, sometimes complex and often distorted.³¹ If the sample is still intact in textile format and stable enough to endure the contact with analytical instruments, a series of analyses can be applied as in Figure 4. First, the investigation of thread making techniques must take place. If the textile is made of threads formed from a bundle of long filament fibres, it is highly likely to be the silk of *Bombyx mori*. In most cases, natural fibres occur in short lengths about 5 to 15 centimetres, and these fibres are gathered and spun into yarns. However, *B. mori* silkworm can produce a continuous filament of 300 m to 1500 m.³² Nonetheless, this analysis cannot be applied to very short pieces of fibre or highly deformed samples.

ATR (attenuated total reflection)-FTIR analysis can tell whether the fibre is proteinaceous or cellulosic by the presence of N-H stretching and amide bands (Figure 5). Although this method gained favour for being fast, easy and non-destructive, there are two significant drawbacks: it only gives information on the compounds; also, it cannot be applied to fragile samples, since the sample needs to be tightly pressed against the ATR crystal. Once the compound characteristic of the sample is known, optical microscopy such as SEM can be useful. Wool fibre has scales on the surface, and medulla in the middle of the cross section; cotton fibre has a ribbon-like lengthwise striation, and a kidney-shaped cross section with the lumen in the middle; linen fibre has bamboo-like nodes and a polygonal cross-section with a large lumen.

³¹ Cybulska and Maik, 2007

³² Panda, 2010

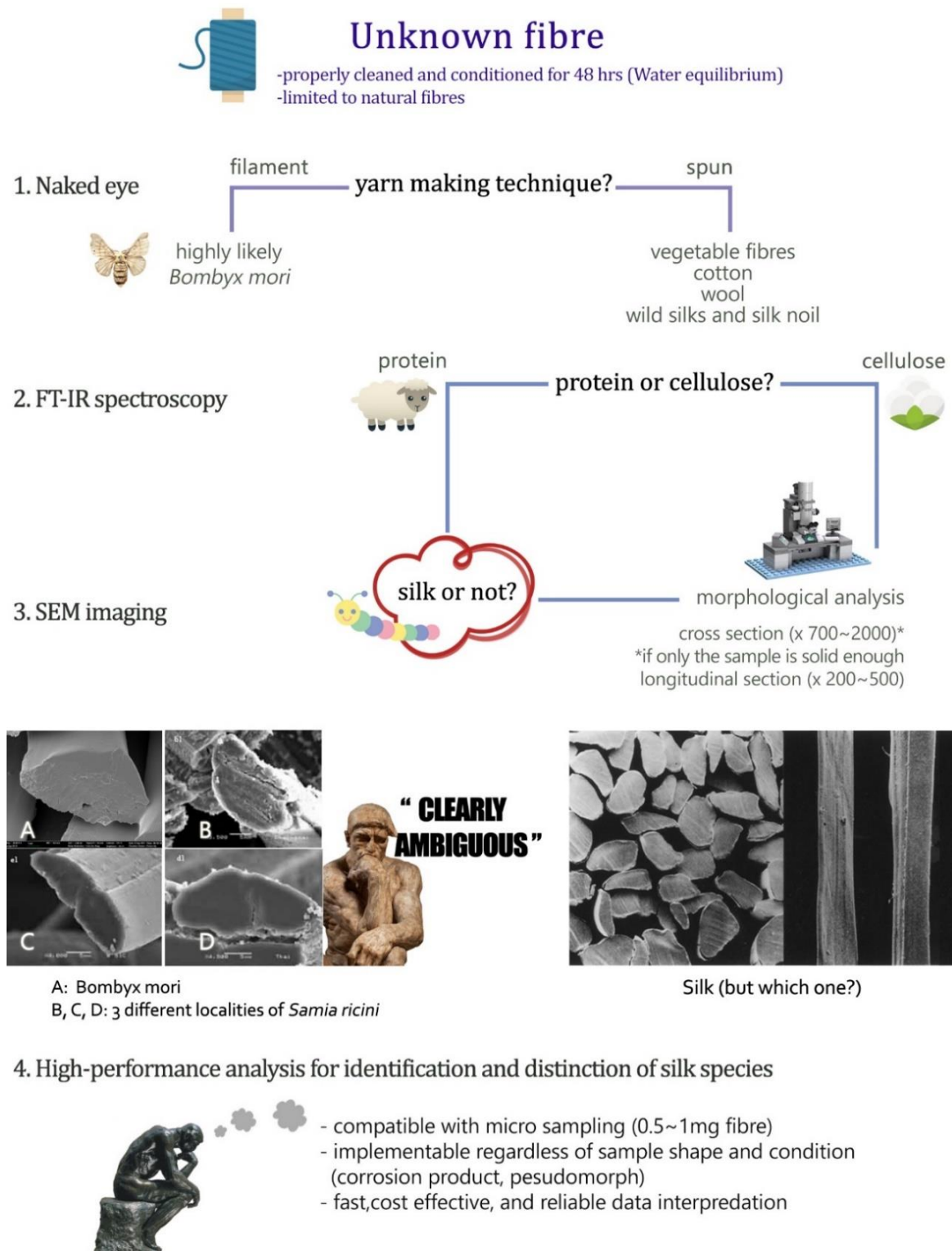


Figure 4. A general flowchart for a conventional fibre analysis

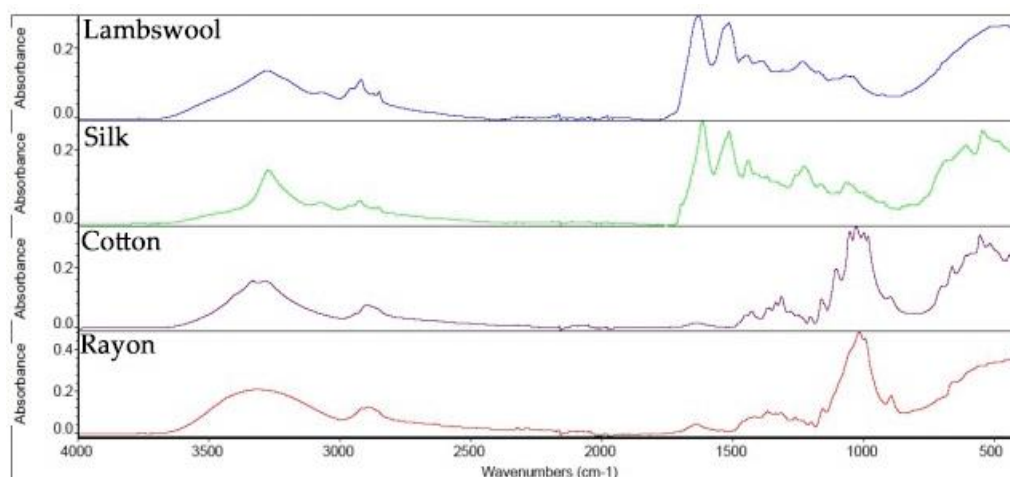


Figure 5. ATR-FT-IR spectrum of different fibres

So far, the identification of fibres in archaeological textile findings is heavily dependent on the morphological analysis by microscopic observation. However, as pointed out by Jorgensen, the documentation of archaeological textiles made with these methods is often incomplete and imprecise.³³ In theory, natural fibres can be characterised by their original microstructure, and so can the silks. The biggest obstacle in applying this method is that textiles intact in shape are a rare finding in the archaeological contexts: archaeological materials are often fragmented and fragile, severely contaminated with human remains and minerals from the previous use and the burial environment, alongside its own natural degradation. Moreover, in the worst cases, they come in the form of corrosion products or pseudomorphs.³⁴ Even if the textile is physically intact, the presence of the protective gum on the surface can also increase the ambiguity of the morphological identification; the gum easily draws more contaminants than the filament. Silk threads and

³³ L.B. Jorgensen, 2013

³⁴ When metallic substances surround fibres, it could inhibit microbiological attack, or replace the organic material with minerals: the mineralisation can result in fibres wholly or partially replaced by the mineral, or a hollow cast of the fibre can form as the fibre decays (Biek, 1963). These are sometimes termed positive and negative casts respectively, and the fully mineralised fibres are called pseudomorphs. (Janaway, 1983, Janaway, 1989)

fabrics were not always degummed: for the ease of weaving and to create specific textile effects. Not only that, but it also must be considered that in the ancient West, Chinese silk was often used together with other fibres because of its scarcity and high price. Therefore, it is important to carefully investigate such aspects of textile samples before selecting the methods for further analysis, and ultimately, to find effective and accurate analytical methods of identification of textile fibres regardless of the textile making techniques. In my opinion, morphological observation alone is not a credible method for identification of silk fibres.

Aside from the deteriorated condition of historical and archaeological silks, it must also be considered that there are many varieties of silk, and their morphological characteristics can vary by genetic, nutritional, and environmental factors. Ramos and Piegler pointed out the irregular nature of fibre formation in 29 different types of silks, and thus these descriptions should be considered as approximations of the total variation to be found within species.³⁵ Consequently, the level of ambiguity increases for much more degraded archaeological samples. Furthermore, how degradation under the edaphic condition affects the irregular nature of the fibre is barely studied.

³⁵ Ramos and Piegler, 1999

I.1.3 Protein mass spectrometry as an alternative option

Whether it is a problem of the interpreters not well versed with the scientific means of analysis, or the unfavourable condition of the samples, it seems clear that the techniques of fibre analysis in archaeology need improvement. Archaeological textiles are subject to the spontaneous degradation and damage from burial conditions such as moisture, acidity and microbial activity: this is the single most important reason to find an alternative method to morphological analysis. Even though its microstructure is damaged, the surviving proteins and peptides molecules can give us clues to the identity of fibres. What is left to do is to figure out how to read these untranslated messages, in a reliable, fast and preferably low-cost way.

Other analytical approaches have been explored in order to find a technique that is compatible with micro sampling, and applicable to the physically altered samples. Isotopic methods, although straightforward, are rather impractical to archaeological protein fibres: they are applicable only for those samples selectively grown in a controlled environment because the isotopic distribution may be affected by husbandry practices.³⁶ In contrast, proteomic approaches are gaining importance as a non-morphological approach.³⁷ One of the recent techniques is the immunoassay, a biochemical procedure for detecting or measuring specific proteins through the antigen-antibody reaction. A few studies have employed IFM (immunofluorescence microscopy) and ELISA (enzyme-linked immunosorbent assay) to localise and detect the target proteins from the painting binders³⁸ and ancient silk.³⁹ Although ELISA is specific and informative for the detection

³⁶ Pearson, *et al.*, 2007

³⁷ Proteomics stand for the global analysis of all aspects of proteins (Matthiesen, *et al.*, 2013)

³⁸ Hu, *et al.*, 2015, Palmieri, *et al.*, 2013, Zheng, *et al.*, 2015

³⁹ Liu, *et al.*, 2015, Liu, *et al.*, 2015

of silk fibroin, it is based on polyclonal antibodies that have the potential to indicate false-positives due to cross-reaction, thus not ideal for the identification of silk species.⁴⁰ Moreover, the process to develop the antibody for each protein can be costly.⁴¹

Another important technique to detect, identify and quantify protein molecules is protein mass spectrometry. A variety of methods and instruments are available for its ionisation methods, analysers, and detectors (Figure 6). In general, the proteins are analysed either intact (top-down) or as digested into fragments (bottom-up) using site-specific enzymes. Once a sample is submitted to the instrument, the ion source generates analyte ions by transferring the molecules from the condensed phase to the gas phase. The mass analyser stores and separates ions based on their mass-to-charge ratios (m/z) and the detector records a signal that can be electronically amplified and stored.⁴² The structural and sequential information gathered from mass spectrometer are submitted to databases to query protein identification.⁴³ The goal of this process is to identify large numbers of proteins fast and accurately. The key factors for a successful analysis are the quality of the MS data, the quality of the database searched, and the method used to search the database.⁴⁴

⁴⁰ Liu, *et al.*, 2015; Zheng, *et al.*, 2015

⁴¹ Scott, *et al.*, 2009

⁴² Matthiesen, *et al.*, 2013

⁴³ Graves and Haystead, 2002

⁴⁴ Ibid.

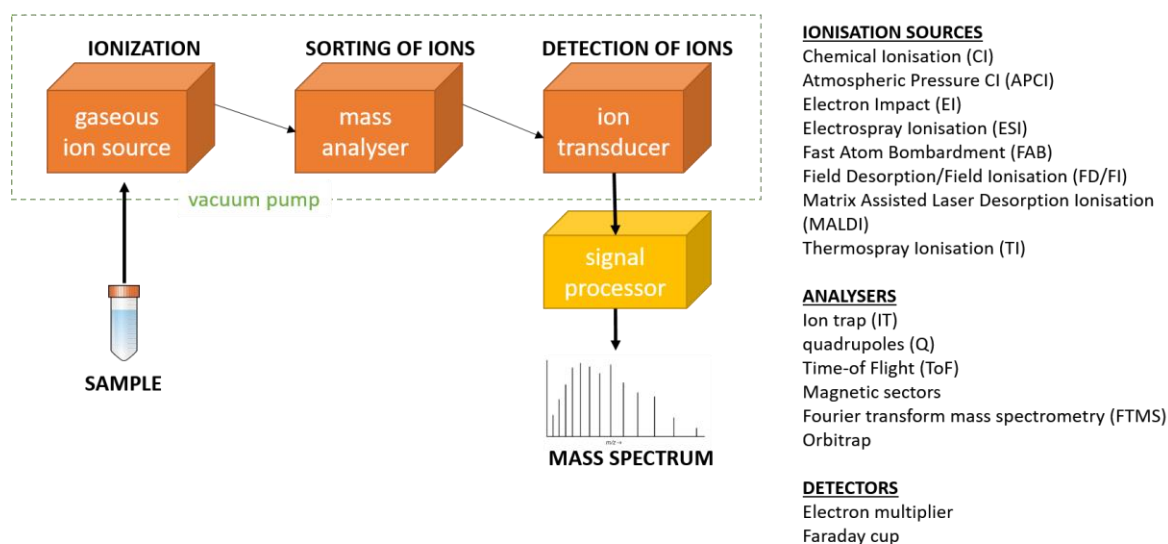


Figure 6. Basics of mass spectrometry instrumentation

The amino acid analysis is used for the determination of the amino acid composition of proteins, peptides and other amino acid containing compounds. First, the peptide bonds connecting each amino acid are hydrolysed, and then the liberated amino acids are separated, detected, and quantified.⁴⁵ Vanden Berghe and Wouters engaged liquid chromatography based amino acid analysis to the identification of silk and wool in pseudomorph samples from the first to second-century Mongolian tombs.⁴⁶ However, when Good employed amino acid analysis to distinguish different silk species, it resulted in further confusion.⁴⁷ It has been argued that amino acid analysis is not suitable to distinguish the differences between silk species because first, the level of amino acid composition in each silk fibre can vary based upon many factors; and second, the degradation-driven differences would be more significant than the inherent differences between proteins.⁴⁸

⁴⁵ Rutherford and Gilani, 2009

⁴⁶ Vanden Berghe and Wouters, 2004

⁴⁷ Good, 2011

⁴⁸ Solazzo, *et al.*, 2014

On the other hand, in peptide analysis, the proteins of interest are proteolytically digested using a sequence-specific protease.⁴⁹ It can be used to determine amino acid sequences of peptides, and to characterise a wide variety of post-translational modifications, as well as to determine absolute and relative protein quantities of thousands of proteins from complex samples.⁵⁰ This enables better detection efficiencies than analysis of intact proteins because the mass spectrometers are most efficient at obtaining sequence information from peptides that are up to approximately 20 residues long, rather than from whole proteins.⁵¹ Such characteristics are advantageous for archaeological fibre samples: first, it does not require the protein to be intact in theory; second, degradation would make the preparation of samples easier than the fresh fibres samples (silk has a very stable molecular structure); and third, the proteinous contamination from human and animal remains could interfere less. Moreover, peptide analysis using tandem mass spectrometry can provide proteomic profiling of the degradation mechanism.⁵² Solazzo *et al.* applied this technique to study the degradation pattern of *B. mori* silk by photo-oxidation of tyrosine site and the effect of historical textile finishings on the degradation.

In general, protein identification using mass spectrometry employs either of two established methodologies, peptide mass fingerprinting and peptide sequencing. Both techniques use a proteolytic enzyme to cleave the proteins into peptide fragments of an appropriate length (mass).⁵³ The basics of these two methods are provided in the flowchart (Figure 7). The advantages and limitations of these are summarised in the following section, together with the review of relevant studies.

⁴⁹ Kicman, *et al.*, 2007

⁵⁰ Aebersold and Mann, 2003; Zhang, *et al.*, 2010

⁵¹ Steen and Mann, 2004

⁵² Zhu, *et al.*, 2014; Li, *et al.*, 2015

⁵³ Kicman, *et al.*, 2007

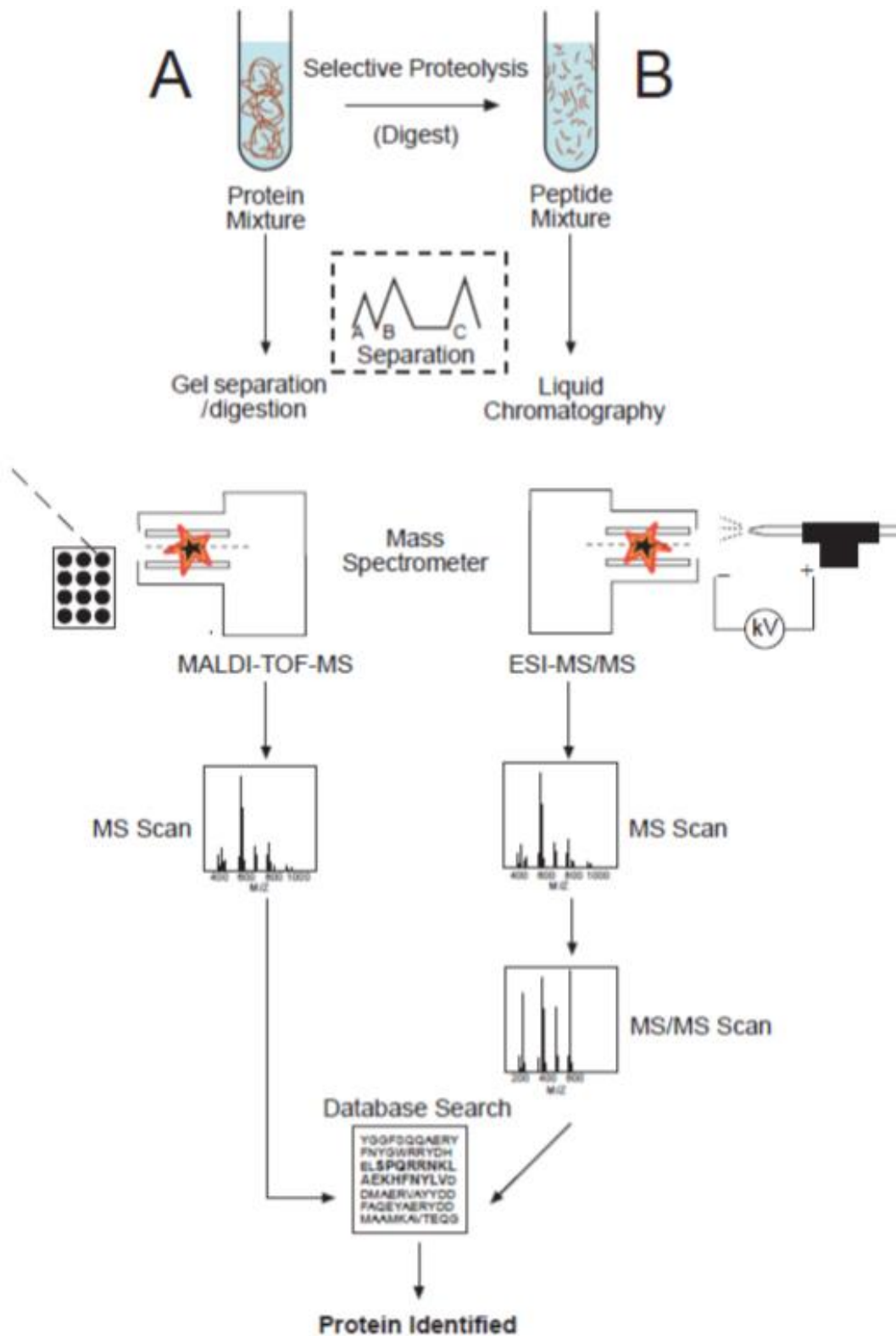


Figure 7. Two common approaches for mass spectrometry based proteomic analysis: peptide mass fingerprinting (A) and peptide sequencing (B).⁵⁴

⁵⁴ Ibid.

1.1.3.1 Peptide Mass Fingerprinting

Peptide mass fingerprinting (PMF) is a method used widely for quick and cost-effective protein identification,⁵⁵ often achieved with MALDI-TOF analysis with a database search. In MALDI, samples are co-crystallised with a UV-absorbing organic matrix to promote the co-desorption of laser irradiation and ionisation of the analyte.⁵⁶ In TOF analyser, the mass per charge ratio (m/z) of a peptide ion can be determined by measuring its drift time of flight (TOF). The resulted fragment mixture is called the peptide fingerprint, and it is very specific to a given protein: which makes the method useful to distinguish between the different subjects. The protein of interest can be identified unambiguously by comparing the spectrum data to a database. However, if it is entirely unknown or poorly databased, *de novo* sequencing of the amino acid chain becomes useful.⁵⁷

In peptide mass fingerprinting, the distribution and intensity of peptides based on the pattern of mass to charge (m/z) ratios from the unknown protein—so called fingerprint—are compared to that of theoretical digestion of protein in the database.⁵⁸ The results are compared to the database and suggested possible matches enable protein identification. The advantage of this method is that only the masses of the peptides must be known. However, the limitation is that the identification of the subject is possible only if the spectral database contains peptide mass fingerprints of the type strains, and there may be some ambiguity with protein mixtures as most algorithms assume the peptide comes from a single protein.⁵⁹

⁵⁵ Gonnet, *et al.*, 2003

⁵⁶ Matthiesen, *et al.*, 2013

⁵⁷ Buckley, *et al.*, 2009

⁵⁸ Graves and Haystead, 2002

⁵⁹ Cottrell, 2011, Singhal, *et al.*, 2015

To be successful in this method, generally five or more accurate masses are required to identify a protein with enough confidence and three or four for a more ambiguous match.⁶⁰

One of the well-known applications of MALDI-TOF/MS for the species identification of animal proteins from archaeological samples is SIAM (Species Identification of Animals using MALDI-TOF mass spectrometry of animal). This keratin-based method was developed for the identification of native feathers, down and hair: and has been applied to examine the animal hairs from the remains of Neolithic Tyrolean Iceman, the Ötzi.⁶¹ Peptide mass fingerprints from the ancient samples were compared to the reference species, which showed species-specific differences in their peptide patterns. A multidimensional scaling and binary hierarchical cluster tree analysis were applied to calculate their spectral similarities as a measure for their zoological relationship, to a family or subfamily level. The results are mainly consistent with a favoured hypothesis of the iceman's occupation as a herdsman; Red deer was found in his shoe vamp, goat in the leggings, cattle in his shoe sole and at his quiver's closing flap as well as sheep and chamois in his coat. Using sheep, cattle and perhaps goat hairs to fabricate clothing agrees with the importance of these domesticated animals as components of the early agro-pastoral economics during the Neolithic-Copper Age of central Europe.⁶²

Buckley *et al.* also proposed a similar means of rapid identification of archaeological remains using PMF termed Zooarchaeology by Mass Spectrometry (ZooMS).⁶³ This method specifically aims to distinguish similar animal species when their respective sequences are identical, using statistical approaches to make comparisons based

⁶⁰ Kicman, *et al.*, 2007

⁶¹ Hollemeyer, *et al.*, 2002, Hollemeyer, *et al.*, 2008

⁶² Sherratt, 2001

⁶³ Buckley, *et al.*, 2009

entirely upon peptide masses.⁶⁴ ZooMS uses species-specific collagen-peptide markers from mammal bones obtained from MALDI-TOF/MS and *de novo* sequencing of the peptides. This approach was successfully applied in several studies of archaeological animal fibres: Buckley *et al.* have distinguished sheep and goat in all 26 archaeological fibre samples despite the majority of samples having low collagen yields; Solazzo *et al.* have identified dog, goat and sheep wool fibres from ancient blankets confirming the existence of woolly dog in pacific northwest of North America that is now extinct; relatively well preserved intact protein markers were found from mineralised archaeological fibre samples; and some important markers of mammalian keratinaceous materials were successfully sequenced.⁶⁵ It is argued that the high protein contents in extensively mineralised samples indicate that chemical degradation was retarded by the mineralisation of the fibres, as the corroding metal acts primarily as a biocide.

⁶⁴ Richter, *et al.*, 2011

⁶⁵ Solazzo, *et al.*, 2011

1.1.3.2 Peptide sequencing

The most specific type of protein identification is peptide sequence search with database or *de novo* sequencing with MS/MS data.⁶⁶ The first approach to this method was the sequence tag, originally described by Mann and Wilm: the partial information about the amino acid sequence (sequence tag) are used as a unique identifier of a specific peptide and searched in the database containing all possible peptide sequences.⁶⁷ Later, an improvement in technology made the uninterpreted MS/MS peak lists are directly searchable against the database.⁶⁸ Many programs are now available for the identification of uninterpreted MS/MS data such as MASCOT and SEQUEST.⁶⁹

Proteins are identified based on matches to their sequences resulting from comparing observed peptide MS/MS spectra to theoretical spectra. It is the peptides that are being identified from this process, and the protein can be inferred from the identified peptides.⁷⁰ Among the distinct peptides identified in the protein group, there is a unique peptide: it is defined as a peptide, regardless of its length, which only exists in one specific protein; it may appear more than once in the same protein sequence.⁷¹ If the protein of interest is unknown or its sequence is not already in the database, a *de novo* sequencing method is used. *De novo* sequencing attempts to reconstruct individual peptide sequences based upon the MS/MS fragmentation. The ability of software to derive a sequence from and MS/MS spectrum is dependent on many factors such as the instrument mass accuracy,

⁶⁶ Cottrell, 2011

⁶⁷ Mann and Wilm, 1994

⁶⁸ Cottrell, 2011

⁶⁹ MASCOT from Matrix Science (Mowse algorithm), Crux from University of Washington (Sequest algorithm) are the web-based search engine: there are also commercially available soft wares such as Protein Pilot from AB Sciex (Mowse and Paragon algorithms), Bioworks from Thermo Electron (Sequest algorithm) and Peaks from Bioinformatic Solutions.

⁷⁰ Cottrell, 2011

⁷¹ Higdon and Kolker, 2007; Zhao and Lin, 2010

resolution, and spectral quality. Other difficulties are caused by poor peptide fragmentation, peptide ion series directionality, and cleavage abnormalities in spectra. These can be addressed by *de novo* algorithms,⁷² but even now these require independent scrutiny.

High-resolution LC-MS/MS has become an essential tool for peptide sequencing, and has significant analytical advantages, especially when dealing with sample-limited situations.⁷³ In this regard, it holds potential for archaeological studies. In the example of Xie *et al.*, protein analysis with nanoLC-MS/MS successfully identified dairy products from the residues inside of a grass bag from Early Bronze Age communities of Gumugou, China.⁷⁴ The finding of kettle milk at this archaeological site of China illustrated that milking was practised in eastern Eurasia as early as the 19th century BCE and suggested that its eastward spread was influenced by European settlers. Such a result suggests a promising extension of this approach for the analysis of archaeological textile findings, as they are often found in conditions that are not compatible with conventional methods of analysis.

Some studies of Chinese silks have already adopted this method: Zhu *et al.* have analysed silk pall and pall imprint on soil samples from Peng-state cemetery (Grave M1, Shanxi, China) by nanoLC-MS/MS.⁷⁵ The residual proteins extracted from the soils underneath the pall imprint were compared with a relatively well-preserved pall from the Warring States Period (475-221 BCE). The results matched that of *B. mori* fibroin heavy chain from the NCBI database, and thus confirmed that the pall was made of silk, offering evidence of the high social status of the deceased in the cemetery. Similarly, Li *et al.* used the same technique to examine several ancient silk remains: various types of silk samples,

⁷² Hughes, *et al.*, 2010

⁷³ Gaspari and Cuda, 2011

⁷⁴ Xie, *et al.*, 2016

⁷⁵ Zhu, *et al.*, 2014

ranging from fresh ones to extremely degraded pseudomorphs were analysed, and the peptide sequences matching to that of *B. mori* fibroin heavy and light chains were detected from all samples.⁷⁶ In the recent publication of Gong *et al.*, the analysis of ancient soil samples showed the presence of matching peptides to *B. mori* fibroin heavy chain.⁷⁷ In these studies, it is also argued that the presence of light chains in the archaeological sample could serve as the marker of degradation, because there was no sign of the light chain among the physically damaged samples: molecular weight is one of the crucial factors affecting the degradation profiles of a polymer,⁷⁸ and this is consistent with the structural characteristics of *B. mori* fibroin (See II.3 Physicochemical properties).

⁷⁶ Li, *et al.*, 2015

⁷⁷ Gong, *et al.*, 2016

⁷⁸ Ko, *et al.*, 2013

I.2 Research objectives, questions and structure

It is hoped that the application of protein mass spectrometry will allow more accurate and reliable analysis than the conventional methods of fibre examination. However, there has been only a limited amount of research concerning the proteomic analysis of silk, largely owing to the difficulty of manipulating the material and the lack of techniques. Furthermore, it seems that no prior study has explored both the historical and scientific aspects of silk beyond the introductory domain, despite the known significance of interdisciplinary approach is. This work is split into two distinct sections aiming at two objectives: one dealing with the review of silk (Chapter I and II) and one dealing with the experimental works to develop the solubilisation method and to assess the application of protein mass spectrometry to silk fibroin solutions (Chapter III and IV). The first part provides a comprehensive study of the silk as material on both the archaeological and the scientific ground, which is essential to understand what makes it difficult to elucidate the archaeological controversies around the early silk findings around the world. The second part presents the results from experimental studies. The aim has been to develop a method for the protein mass spectrometry of silks, to distinguish one from another and to identify each species. To achieve the analysis of silk proteins using high-resolution LC-MS/MS for peptide sequencing, two steps of preparation were necessary: first, the solubilisation of silk fibres; second, the enzymatic digestion to prepare the proteins into the mass spectrometry-compatible size peptides. A series of experiments were implemented in each step with the following research questions in mind: what solvents apply to the solubilisation of silk regardless of the species? which proteolytic enzyme accommodates the purpose best? is the method established through the above processes applicable to the identification of fresh,

modern silk samples as well as to the examination of archaeological silk samples?

The following is a structural framework that outlines the objectives of each chapter that comprises this research.

The first chapter is introductory that reviews the current issues in the study of archaeological silk that addresses in the limitations of conventional methods, highlights key questions and suggests protein mass spectrometry as an alternative approach to resolve these problems.

The next chapter offers a comprehensive summary of silk. The first section reviews the socio-cultural aspect of silk, based on the literary sources and archaeological evidence, summarising the history of silk and focussing on how the culture of silk has risen from the East and spread to the West. However, it is not confined to the use of domesticated silk cultivation, but also includes the existence and utilisation of wild silk species. The second section explains how the varieties of silks are different from each other by discussing their entomology and geographical habitat and gives a digest of silk cultivation and thread production. The third section describes the physical and chemical properties of domesticated and non-domesticated silks from a polymer to a molecular level.

The third chapter contains all information regarding the materials, methods and instrumentation used in the experimental studies.

The fourth chapter comprises a series of experimental works designed to develop a method to solubilise silk proteins and to prepare them suitable for a mass spectrometry analysis. The process involves solubilisation tests with various solvents, and the study of the characteristics of solubilised protein molecules using protein separation techniques. Moreover, a series of methods to desalt, concentrate, and digest the protein solutions were

tested to develop an appropriate protocol.

The fifth chapter discusses the protein mass spectrometry results of the silks. The first section describes the groundwork of silk identification using protein mass spectrometry: the consolidation of a silk protein sequence database, the analysis of unique sequence features of each silk fibroin using the sequence alignment, and the comparison of theoretical models of each silk fibroins. In the following sections, the comparison between the database query and the *de novo* analysis results for the identification of silk fibroin is presented. Furthermore, how the variables such as enzymes and solvents affected the result of the analysis are examined based on the experimental data. The raw data from mass spectrometry (*.raw) will be uploaded to an appropriate online database for open source studies. Some spectra of characteristic peptides and protein sequence coverage maps are given in the text.

In the sixth chapter, the implication of this research for the investigation of archaeological textiles through two case studies scoping out the advantages and the limitation of the current silk solubilisation and protein mass spectrometry technique. The first section studies a variety of textiles from a Gallo-Roman site in Naintr , France; and the second section examines the alleged tussah silks from the ancient Palmyra, Syria.

The seventh chapter summarises the results of experimental studies including the silk solubilisation, the protein mass spectrometry analysis, and the case studies of Naintr  and Palmyra silks.

Additional studies and data are listed in the appendices. Appendix A contains a preliminary light stable isotopic study of mulberry leaves stored in Oxford Plant Science Herbarium that was a part of my transfer submission. This study was done in 2014-2015

and submitted for the transfer process towards the D.Phil. degree. This study was implemented as a trial, however only proved that the stable isotopic analysis is neither appropriate to differentiate various types of silks nor to provenance them. Appendix B contains the sequence alignment result of all silk protein sequences used in the experimental study. Appendix C contains the data tables corresponding to some of the peptide analysis results from Chapter V.

II. Silk

II.1 Entomology

Silk is a protein spun into fibres by arthropods, including spiders, lepidopteran larvae, hymenopterans, scorpions, and mites.⁷⁹ The ability to produce silk has evolved in at least 23 groups of insects, but in general, only the silk of moth caterpillars has been used for textile manufacturing.⁸⁰ Spider silks are durable and resistant to extreme environments, but the commercialisation is difficult due to high production costs.⁸¹

The lepidopteran insects that are commercially used to produce silk mainly belong to two families: *Bombycidae* and *Saturniidae*.⁸² *Bombyx mori*, the mulberry silkworm is the most common and preferred kind for textile manufacturing. There are 350 species in *Bombycidae*, distributed worldwide except Europe. The silk moths of the family *Saturniidae*, often called non-mulberry silks, are polyphagous and occur in the wild. Among the 1,480 large species distributed worldwide belonging to *Saturniidae*, about 500 species are used in the production of non-mulberry silk, but only about 80 of those have been commercially exploited in Asia and Africa.⁸³ This includes the semi-domesticated species used for silk production spread over India, China and Japan.⁸⁴ Some of these silk producers have only ancient historical significance, but others have been used well into this century and provide silk of high economic and commercial values.⁸⁵ The major varieties of non-mulberry silk and the distribution are presented below (Figure 8).

⁷⁹ Capinera, 2008 ; Asakura and Miller, 2014

⁸⁰ Sutherland, *et al.*, 2010

⁸¹ Jolly, *et al.*, 1979 ; Babu, 2012

⁸² Arunkumar, *et al.*, 2006 ; Tikader, *et al.*, 2013

⁸³ Jolly, *et al.*, 1979

⁸⁴ Arunkumar, *et al.*, 2006

⁸⁵ Capinera, 2008 ; Panda, 2010

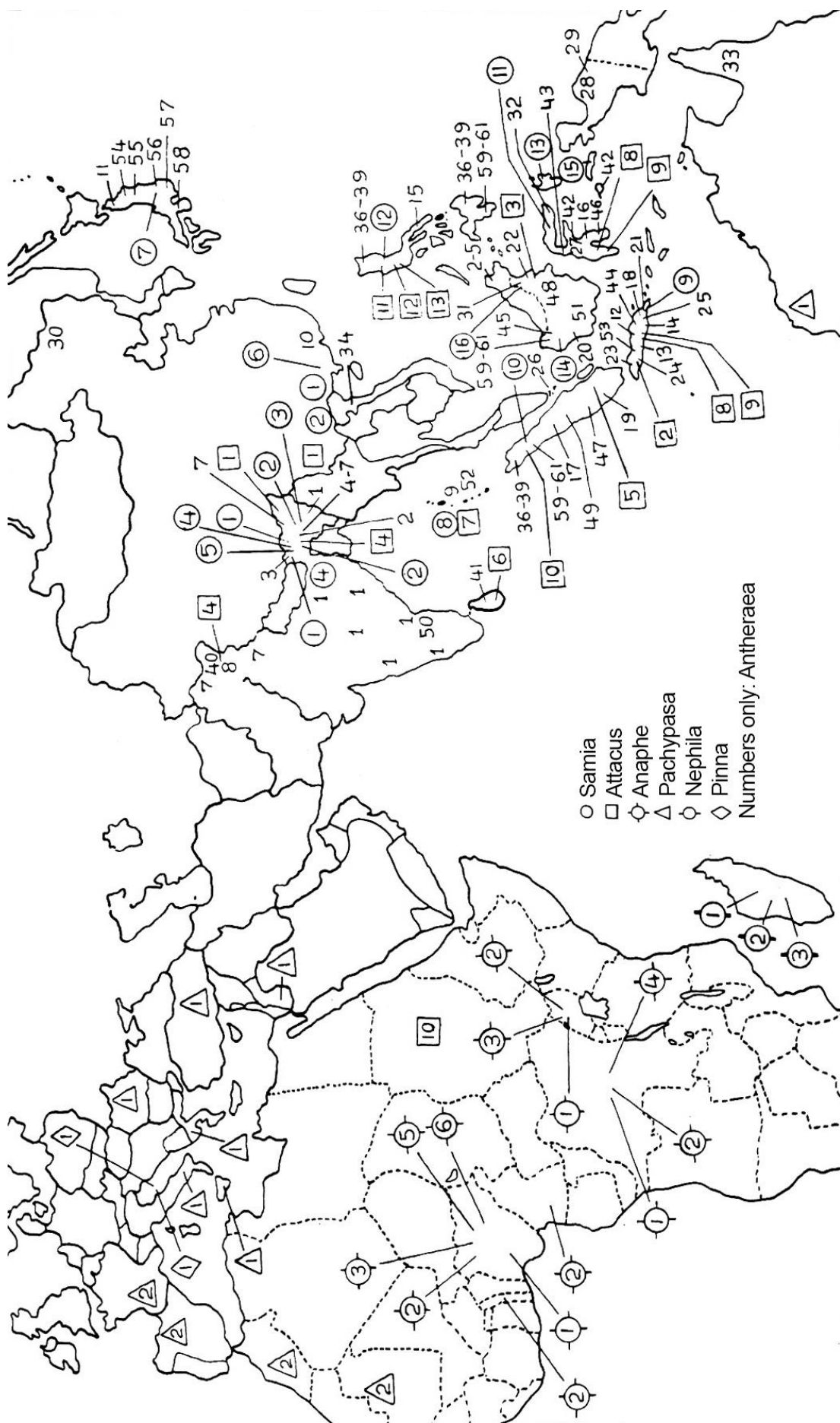


Figure 8. Distribution of non-mulberry silk fauna⁸⁶

Antheraea

1	<i>mylitta</i>	D.	21	<i>larissa</i>	Ww.	41	<i>cingalesa</i>	Mr.
2	<i>assamensis</i>	Ww.	22	<i>ridleyi</i>	Mr.	42	<i>celebensis</i>	W.&S.
3	<i>knyvetti</i>	Hmps.	23	<i>prelarissa</i>	Bouv.	43	<i>buruensis</i>	Bouv.
4	<i>compta</i>	R.&J.	24	<i>surakarta</i>	Mr.	44	<i>subcaeca</i>	Aurvil.
5	<i>frithiii</i>	Mr.	25	<i>mylittoides</i>	Bouv.	45	<i>fusca</i>	Roth.
6	<i>helpferi</i>	Mr.	26	<i>delegate</i>	Swh.	46	<i>minahassae</i>	Niep.
7	<i>roylei</i>	Mr.	27	<i>fickei</i>	Weym.	47	<i>sumatrana</i>	Niep.
8	<i>sivalica</i>	Mr.	28	<i>pristina</i>	Wkr.	48	<i>borneensis</i>	Mr.
9	<i>andamana</i>	Mr.	29	<i>sciron</i>	Ww.	49	<i>korintijina</i>	Bouv.
10	<i>pernyi</i>	G.M.	30	<i>harti</i>	Mr.	50	<i>perrotteti</i>	Guer.
11	<i>yamamai</i>	G.M.	31	<i>gephyra</i>	Niep.	51	<i>yongei</i>	Wts.
12	<i>pasteuri</i>	Bouv.	32	<i>rumphi</i>	Fldr.	52	<i>insularis</i>	Wts.
13	<i>raffrayi</i>	Bouv.	33	<i>eucalypti</i>	Sott.	53	<i>javanensis</i>	Bouv.
14	<i>jana</i>	Stoll.	34	<i>larissoides</i>	Bouv.	54	<i>hazina</i>	Butt.
15	<i>semperi</i>	Fldr.	35	<i>polyphemus</i>	Cram.*	55	<i>calida</i>	Butl.
16	<i>cordifolia</i>	Weym.	36	<i>fasciata</i>	Mr.	56	<i>morosa</i>	Butl.
17	<i>pratti</i>	Bouv.	37	<i>versicolor</i>	Mr.	57	<i>fentoni</i>	Butl.
18	<i>imperator</i>	Wts.	38	<i>pulchra</i>	Mr.	58	<i>sergustus</i>	Ww.
19	<i>brunnea</i>	Eecke	39	<i>ochripicta</i>	Mr.	59	<i>nebulosa</i>	Hutt.
20	<i>billitonensis</i>	Mr.	40	<i>fraternal</i>	Mr.	60	<i>olivescens</i>	Mr.
						61	<i>platessa</i>	Roth.

Samia

1	<i>cynthia</i>	D.	7	<i>pryeri</i>	Jord.	13	<i>tetrica</i>	Rbl.
2	<i>ricini</i>	Boisd.	8	<i>fluva</i>	Voll.	14	<i>borneensis</i>	Rbl.
3	<i>lunuloides</i>	Reb.	9	<i>insularis</i>	Wts.	15	<i>ceramensis</i>	Bouv.
4	<i>obscura</i>	Btl.	10	<i>vaneecke</i>	Wts.	16	<i>mindanaensis</i>	Rbl.
5	<i>canningii</i>	Hutt.	11	<i>vaderberghi</i>	Wts.	17	<i>advena</i>	Pack.*
6	<i>walkeri</i>	Fldr.	12	<i>luzonica</i>	Wts.			

Attacus

1	<i>atlas</i>	L.	5	<i>dohertyi</i>	Roth.	9	<i>erebus</i>	Fruhst.
2	<i>standingeri</i>	Roth.	6	<i>taprobanis</i>	Mr.	10	<i>gladiator</i>	Fruhst.
3	<i>crameri</i>	Fldr.	7	<i>macmulleri</i>	Wts.	11	<i>lorquinii</i>	Fldr.
4	<i>edwardsi</i>	White	8	<i>simalurana</i>	Wts.	12	<i>caesar</i>	M.&W.
						13	<i>imperator</i>	Ky.

Anaphe

1	<i>infracta</i>	Wals.	3	<i>monoley</i>	Druce	5	<i>reticulata</i>	Walk.
2	<i>venata</i>	Butl.	4	<i>panda</i>	Boisd.	6	<i>carteri</i>	Wals.

Spider silks

1	<i>Nephila madagascarensis</i>		2	<i>Miranda aurentia</i>		3	<i>Epeira</i> spp.	
---	--------------------------------	--	---	-------------------------	--	---	--------------------	--

Pachypasa

1	<i>otus</i>	D.	2	<i>lineosa</i>	Vill.
---	-------------	----	---	----------------	-------

Pinna

1	<i>nobilis</i>	Linn.
---	----------------	-------

* *A. polyphemus* and *P. advena* are endemic in the U.S.A.

⁸⁶ Jolly, et al., 1979

II.1.1 Mulberry silk

Bombyx mori (Linnaeus) from *Bombycidae* family, is the most dominant and productive silkworm. It is often referred to as cultivated silk, domesticated silk, or mulberry silk; because in the domestic cultivation, the silkworms are only fed the mulberry leaves in order to produce white and uniform silk. Yoshitake proposed that *B. mori* is probably originated from Chinese *B. mandarina* through selective breeding circa 4600 years ago, and subsequently introduced to Japan via Korea.⁸⁷ It has been singled out for domestication for its docile temperament and the relative easiness to contrive the growth environment.

Bombyx mandarina (Moore) of the *Bombycidae* family is a wild silkmoth occurring from inland China to Korea and Japan. It is genetically similar to *B. mori* and can cross breed to yield fertile hybrid offspring. Compared to *mori*, *mandarina* moths are smaller, slender, and brown; they form a soft, ovoid, lemon yellow coloured cocoons that can be easily reeled out (Figure 9). However, *B. mandarina* includes significant variation within species.⁸⁸ For example, *B. mandarina* inhabiting China has 28 pairs of chromosomes, similar to that of *B. mori*, whereas *B. mandarina* residing in Korea and Japan has 27 pairs of chromosomes.⁸⁹ The sequence divergence between *mori* and *mandarina* suggests significant genetic isolation of the current Korean and Japanese population of *B. mandarina* from the population or populations from which *B. mori* was derived.⁹⁰ Arunkumar *et al.* analysed the sequence of four mitochondrial loci, and the populations from Korea and Japan showed different karyotype from those of inland China; furthermore, the domestic silkmoths lack some genetic apomorphies of the north-eastern *B. mandarina*. Thus, *B. mori*

⁸⁷ Yoshitake, 1968

⁸⁸ Hwang, *et al.*, 1999, Yukuhiro, *et al.*, 2002; Banno, *et al.*, 2004;

⁸⁹ Banno, *et al.*, 2004

⁹⁰ Ibid., Yukuhiro, *et al.*, 2002

is closer to the Chinese *B. mandarina* regarding mtDNA sequence data.⁹¹

B. mori is one of the best-studied lepidopteran model systems for insect physiology, immunity, and genetics with adequately characterised mutations.⁹² During the long history of domestication, many mutations have occurred: it became a large silkworm race as a result of crossbreeding between the mutants.⁹³ More than 400 mutations have been described in *B. mori* silkworm, and about 200 mutants have been placed on linkage maps.⁹⁴ At present, *B. mori* has an enormous geographic, and genetic diversity exists all over the world, comprising many geographical breeds and commercial hybrids.⁹⁵

Their place of origin can also categorise *B. mori* silks. Chinese races are aboriginal to China.⁹⁶ The eggs are light greenish-grey and are usually lighter in colour than the Japanese variety. The larvae are plain, grows relatively faster and are bluish-white. They are resistant to disease and environmental stress. The cocoons are oval shaped and produce thin, fine threads. Japanese races have longer larvae lifespans than the Chinese variety and produce peanut-shaped cocoons with thicker and shorter threads. The eggs are greyish brown.

Tropical races are found in India and South East Asia, and their eggs are small and lustrous. The larvae are sturdy and make small and spindle-shaped cocoons that are yellow or green. European races are aboriginal in Europe and Central Asia.⁹⁷ The larvae are big and have long bodies, thus heavier in weight than other regional varieties. The cocoon is long and oval-shaped, and easy to reel out. However, it is weak against diseases and

⁹¹ Arunkumar, *et al.*, 2006

⁹² Willis, *et al.*, 1995; Ramesha, *et al.*, 2010

⁹³ Panda, 2010

⁹⁴ Doira, *et al.*, 1992

⁹⁵ Vijayan, *et al.*, 2010; Furdui, *et al.*, 2014

⁹⁶ Panda, 2010

⁹⁷ Ibid.

environmental stress, thus rather hard to breed. In the late 19th century, Europe and the Middle East experienced a drastic fall in silk production due to the silkworm epidemic called *Pébrine* disease.⁹⁸ The Introduction of Japanese and Turkish eggs immune to this disease permitted only a partial recovery: thus, the primary resource of the industry has shifted to the imported raw fibres.⁹⁹

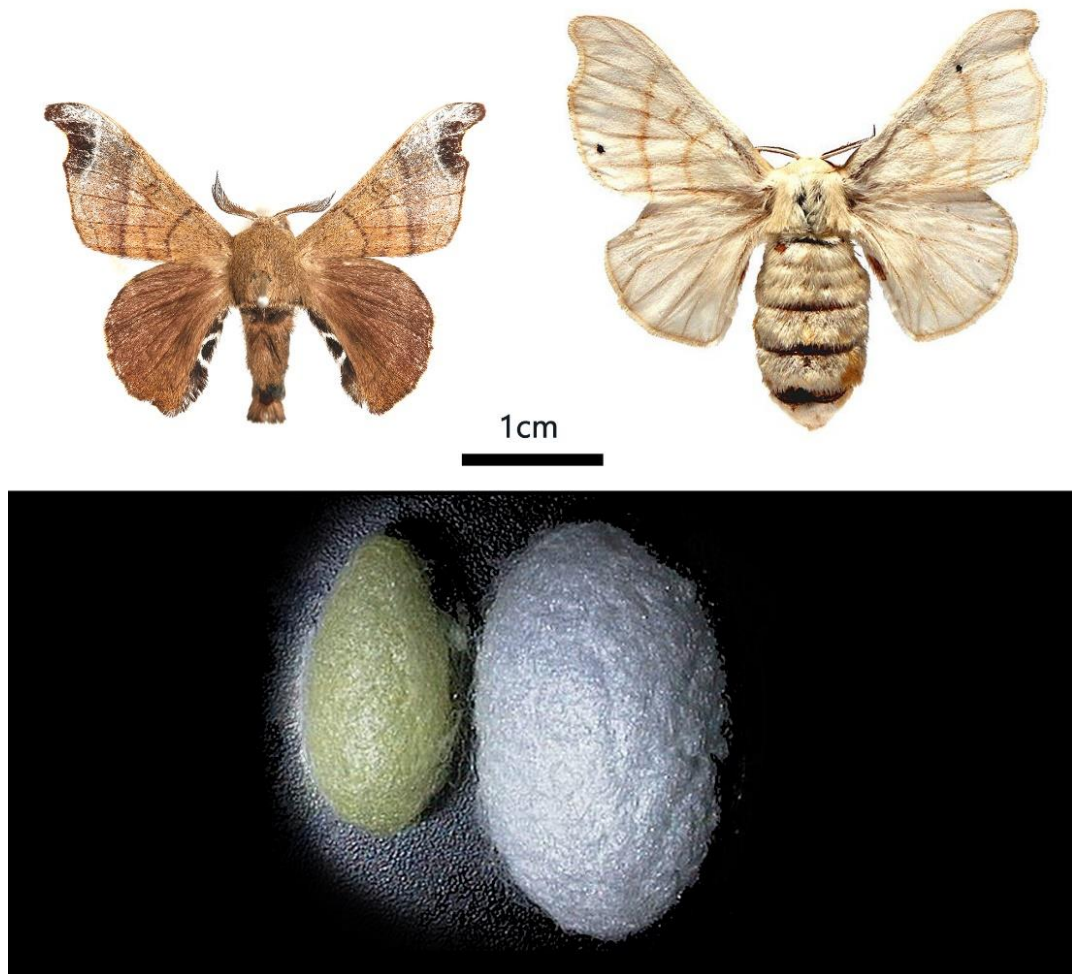


Figure 9. The adults (above) and cocoons (below) of *Bombyx mandarina* (left) and *Bombyx mori* (right).

⁹⁸ Datta and Nanavaty, 2005

⁹⁹ Issawi, 1982

II.1.2 Non-mulberry silk

II.1.2.1 *Tasar silk*

Tasar is a variety of silk moths of *Antheraea* species adapted to different climates, from tropical to temperate over China to India. Tasar silk is also called as tassar, tusser, tussar or tussah silk. The word tasar is derived from the Sanskrit word *trasara* (त्रसर) that means shuttle. The earliest record of tasar silk in literature dates to 1590 BCE: an ancient Indian epic “Ramayana” depicts the giving tasar silk as a gift.¹⁰⁰ At present, four species of *Antheraea* (Hübner) are in commercial use: *Antheraea mylitta* (Drury), *Antheraea proylei* (Jolly), *Antheraea pernyi* (Guérin-Méneville) and *Antheraea yamamai* (Guérin-Méneville). Although it belongs to the same genus, the silk produced by *Antheraea assamensis* (Helfer) is called Muga. Among the Tasar silks, *A. pernyi* of China and *A. mylitta* of India are used to produce tasar silk, but usually the Chinese variety produces better quality yarns. Indian tasar is also called as tropical tasar to distinguish it from Chinese tasar, which is also called temperate tasar (Table 1). *A. proylei* (n, chromosome number=49) is the only known fertile hybrid of *A. Pernyi* and *A. roylei* (n=30, 31, 32), distribute along the sub-Himalayan belt of India.¹⁰¹ Although recognised a century ago, the commercial exploitation of this hybrid was not attempted.¹⁰²

The Chinese tasar, *A. pernyi* (n=39) is a native of China, which feeds on Oak (*Quercus serrata* or *Quercus dealbara*), and thus is also referred to as Oak tasar. *A. pernyi* has been the monopoly of China until recently: the raw silk production is nearly 4000 metric tons per annum and accounts for 90% of tasar production in the world. The original abode

¹⁰⁰ Ojha and Panday, 2004

¹⁰¹ Jolly, *et al.*, 1979

¹⁰² Ghosh and Ghosh, 1995

is the north-eastern provinces of Liaoning(遼寧) and Shandong(山東), and its commercial exploitation dates back to Han and Wei dynasties (Figure 10). The famous Chinese Shantung and Pongee fabrics are made of this silk.¹⁰³ The cocoons are single -shelled, cream-yellow coloured, oval shaped with a small peduncle at the end (Figure 11). The cocoons of this variety are larger (than that of *B. mori*), with a high yield of superior yarn, which is difficult to reel due to a high gum or sericin content.¹⁰⁴

The Indian tasar, *A. mylitta* (n=31) is a natural fauna of tropical India. *A. mylitta* is endemic and distributed in different geographical regions of India in the form of different ecological races.¹⁰⁵ The wide distribution and polyphagia of this species had resulted in extensive variation in the population. It is produced in North Eastern India, West Bengal, Bihar, Orissa, Madhya Pradesh and Maharashtra, and mostly reared by tribal people. *A. mylitta* is known to represent about 40 eco-races, which differ from each other in qualitative and quantitative traits, such as fecundity, voltinism, cocoon weight, and in its host plant preference.¹⁰⁶ The cocoon is single-shelled, yellow-grey in colour, and oval shaped. The surface is non-flossy, and there is a well-formed dark brown peduncle with a ring at the end (Figure 11). It is known that the females spin larger cocoons than the males.¹⁰⁷ Tasar fibre has distinctive characters: it has a yellow-brown colour and coarse hand, but also has higher tensile strength, elongation, and stress-relaxation values than the mulberry silk fibre from *B. mori*.¹⁰⁸ *A. roylei* species are also used for reeling due to high yield, but they give inferior yarn.¹⁰⁹

¹⁰³ Goel and Krishna Rao, 2004

¹⁰⁴ Ghosh and Ghosh

¹⁰⁵ Mahendran, *et al.*, 2006

¹⁰⁶ Arunkumar *et al.* ; Sinha, *et al.*, 1994

¹⁰⁷ Jolly, *et al.*, 1979

¹⁰⁸ Iizuka, 2000, Rajkhowaibid.

¹⁰⁹ Ghosh and Ghosh, 1995

Table 1. Comparison of temperate and tropical Tasar

Sci. Name	<i>Antheraea pernyi</i> G.M.	<i>Antheraea mylitta</i> D.
Origin	N.E China (Shandong, Liaoning)	All over India
Size	< (both moth and cocoon)	
Feed	Oak, Acorn (silk worm oak), Chestnut	Castor
Cocoon	Yellow to brown	Yellow
Fibre length	Approx. 500m	Approx. 550-1800m
Fibre thickness	4D	
Lustre	0	X
Strength	<	
Elongation	<	
Note	More resistant to chemical solutions (vs. <i>B. mori</i>), thus less dyeable	
Similar variety	<i>Caligula Japonica</i> (Kuriwata)	

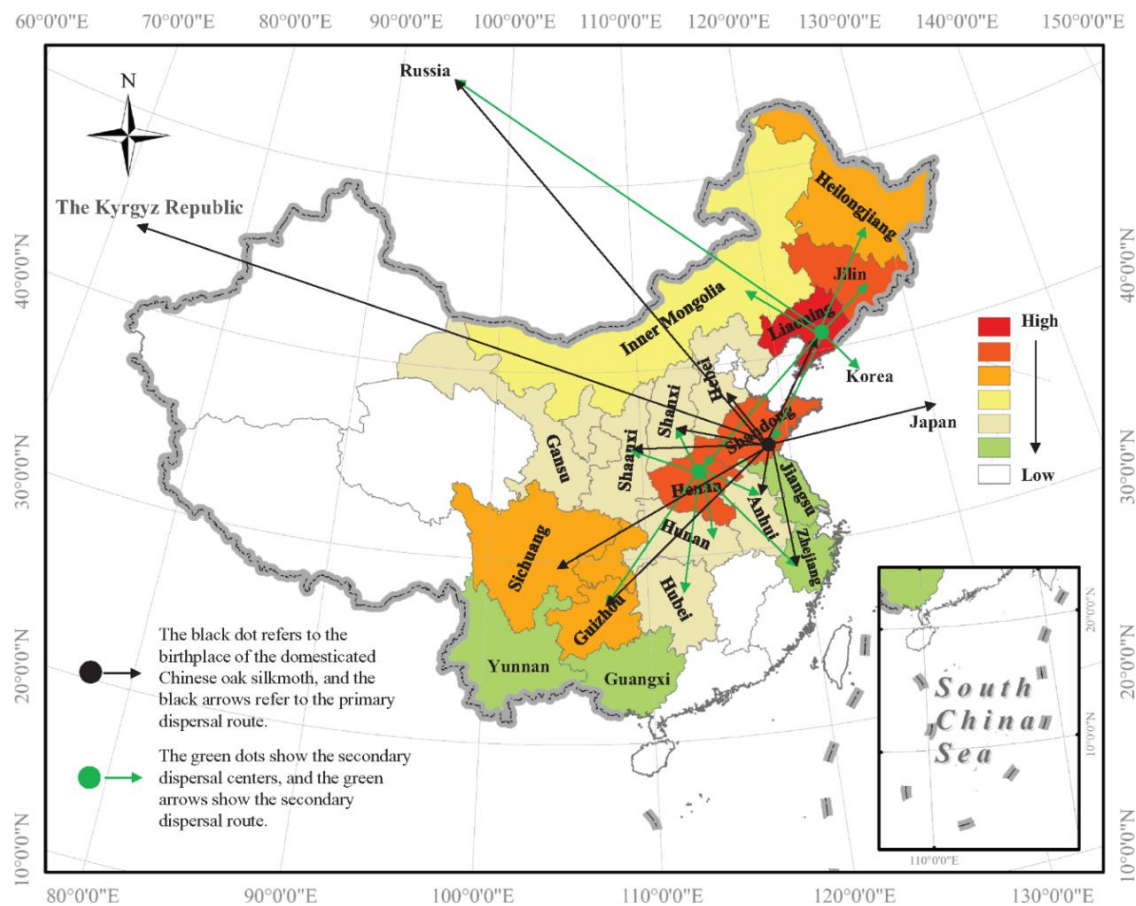


Figure 10. Origin and dispersal of *A. pernyi* cultivation and current level of production ¹¹⁰

¹¹⁰ Li, et al., 2017

The Japanese oak silkworm, *Antheraea yamamai* (Guérin-Méneville) (n=31) is a native of Japan that contributes to the production of a highly priced silk. The name Yamamai, derived from the Japanese word “*yama-mayu* (山蠶)”, literally meaning the mountain silkworm. It is distributed throughout Japan, Korea, China and Far Eastern Russia. There are at least 11 varieties in Japan, but Tensan (天蚕) is the most well-known kind. There are four sub-species, namely *bergmani* Bryk (1949), *superba* Inoue (1964), *titan* Mell (1958) and *ussuriensis* Schachbazov (1953). These are distributed from the Russian Far East south through Korea, China (to Zhejiang and Sichuan Province) and Taiwan.

A. yamamai feeds on many species of Oak (*Quercus acutissima* C., *Quercus serrata* T., *Quercus aliena* B., *Quercus dentata* T.).¹¹¹ However, it can also be found on beech (*Fagus sylvatica*), sweet chestnut (*Castanea sativa*), hornbeams (*Carpinus*), wild roses (*Rosa*) and hawthorns (*Crataegus*).¹¹² The cocoon is single-shelled, oval-shaped, and yellow to bright green, due to the existence of blue biliverdin and flavonoid (Figure 11).¹¹³ The outer layer is wrapped with loose, white floss that forms a small peduncle at the end. The thick (about five deniers), greenish fibre has a beautiful lustre, yet good strength (17.4 g/D) and elasticity (22.8%); thus, it is often called as “the diamond of silk”. The cross-section of *A. yamamai* shows characteristic voids on the cross section. Raw fibres contain a relatively high ratio of calcium carbonate, and thus pre-treatment with hydrochloric acid (HCl) is necessary before reeling; it tends to crinkle and lose lustre when immersed in water.¹¹⁴

A. yamamai was introduced to Austria from Japan in the 1860s, and it subsequently

¹¹¹ Yoon, *et al.*, 2006

¹¹² Probably resulted from the adaptation to different regions.

¹¹³ Yamada and Kato, 2004

¹¹⁴ Jolly, *et al.*, 1979

spread to Hungary, northern Italy, Switzerland, Slovenia, and Croatia.¹¹⁵ According to Wardle's record from 1881, it has been acclimatised in Europe and crossed with *Bombyx attacus Pernyi* (now *A. Pernyi*), and is successfully reared in France—their eggs can hatch at the almost freezing point.¹¹⁶

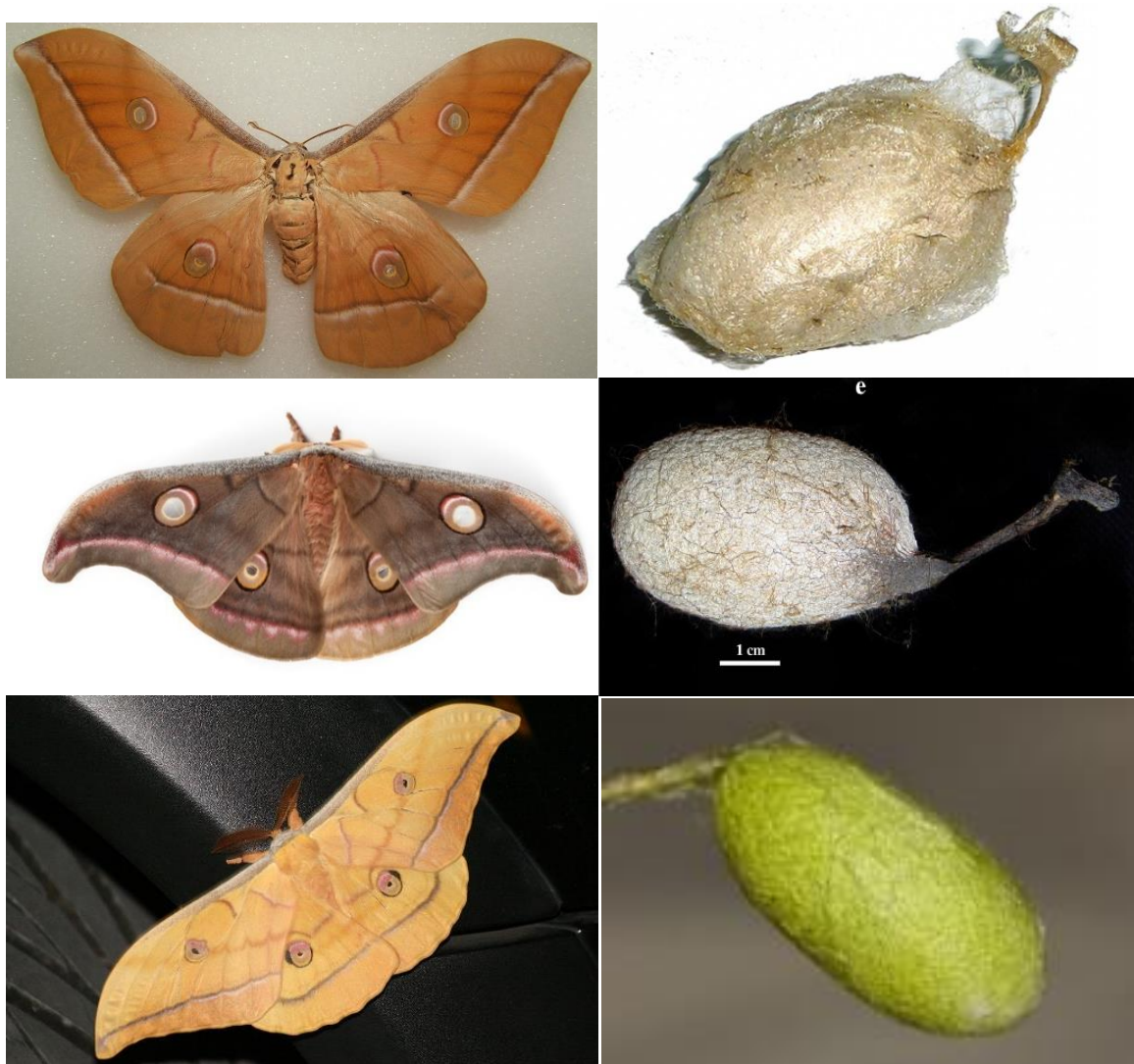


Figure 11. *Antheraea* moths and their cocoons: from the top *A. pernyi*, *A. mylitta*¹¹⁷ and *A. yamamai*.

¹¹⁵ Leraut and 2006

¹¹⁶ Wardle, 1881

¹¹⁷ Bhatia and Yousuf, 2015

II.1.2.2 *Muga silk*

Antheraea assamensis (n=15) is widely distributed and cultured in North-Eastern India, particularly in the State of Assam (and is mostly consumed in Assam). It is named after the Assamese word "*muga*", which means amber. It is also called *moonga* or *mooga* of the Assamese. It is a semi-domestic multivoltine species that yields fine, glossy and naturally golden fibre of high durability.¹¹⁸ It belongs to the *Antheraea* family same as Tasar, and thus the biology of the *A. assamensis* is similar to that of *A. mylitta*.¹¹⁹ The cocoon is single-shelled, golden brown, and oblong shaped. The surface is slightly flossy and has a weak peduncle at the distal end (Figure 12).

A. assamensis is polyphagous. There are about 15 plant species reported to be its host plants: the primary source is Lauraceae such as Som (*Persea bombycina* K.), Soalu (*Litseapolyantha* Juss), Mejankari (*L. citrata* B.), Dighloti (*L. salicifolia*, R.).¹²⁰ However, they also feed on certain Magnoliaceae (e.g., *Magnolia pterocarpa* R. = *M. sphenocarpa*), and this seems to be a more traditional way of rearing the worms: according to Wardle (1881), these silkworms feed upon the trees known in Assam as Champa (*Michelia* sp.), Soom, Kontoolva, Digluttee (*Tetranthera diglottica*), Pattee shoonda (*Lurus obtusifolia*) and Sonhalloo (*Tet. Macrophylla*).¹²¹ These are the varieties of muga silk called “white chapa” or “champa”, worn by Ahom nobles and kings of Assam in earlier times. It was obtained by rearing the larvae on *Michelia champaca* L. (Magnoliaceae) but has not been produced for more than a century.¹²²

¹¹⁸ International Sericulture Commission, 2013

¹¹⁹ Singh and Sachan, 2007

¹²⁰ Reddy and Prasad, 2010; Bora, *et al.*, 2013

¹²¹ Wardle, 1881

¹²² Peigler, 1993



Figure 12. *A. assamensis* moth and cocoon.

II.1.2.3 *Eri silk*

Samia cynthia (Drury), once called *Attacus cynthia*, is a multivoltine silkworm native to China, mainly domesticated in the provinces of Shantung(山東/Shandong) and Honan(河南/Henan). It is also known as *Ailanthus* silkmoth, which refers to the host plant. The closely related subspecies *Samia ricini* (n=13) is often called Eri silkmoth in India, which derives from the Assamese word ‘era’ for castor plant.¹²³ These are domesticated for silk production as well as for consumption in the larval and pupa stages by humans and as feed for domesticated animals and are the only completely domesticated kind that is not dependent on mulberry for food. Eri silk is usually raised on the big palmate leaves of plants in the family *Euphorbiaceae*, such as castor (*Ricinus communis* L.), but also feeds on a wide range of plants such as Kessuru (*Heteropnanax fragrans* S.), Payam (*Evodia flaxinifolia*) and tapioca (*Manihot utilissima*), as well as Tree of heaven (*Ailanthus glandulosa* = *A. altissima*).¹²⁴ Other wild types are distributed in natural state mainly from Japan to SE Asia, but also in Southern Europe and North America (Figure 8).¹²⁵

The cocoons can be easily distinguished from those of *Antheraea* for their elongated oval shape and soft, woolly feature. They exhibit colour polymorphism, being brick red and cream white.¹²⁶ Red cocoons are usually seen in the Indian variety; white cocoons are more valuable than the red ones.¹²⁷ The silk is extremely durable, but cannot be easily reeled off as the cocoons are gathered open-mouthed, and thus spun into yarns as in cotton or wool production.¹²⁸ There are wide morphological and quantitative variations

¹²³ Wardle, 1881

¹²⁴ Capinera, 2008

¹²⁵ Ghosh and Ghosh, op.cit.

¹²⁶ Jolly, *et al.*, 1979

¹²⁷ Ghosh and Ghosh, op.cit.

¹²⁸ U.S. Geological Survey, 2005-2011

in characters such as silk content, larval weight, cocoon weight, cocoon shell weight and silk ratio.¹²⁹ Eri silkworms were successfully acclimatized in America and Europe once: Guérin-Méneville developed a fertile hybrid between the Chinese and the Bengal species and distributed in Europe, however it did not take a firm hold.¹³⁰



Figure 13. *S. ricini* moth and cocoons of different colourmorphs

¹²⁹ Arunkumar, *et al.*, 2008.

¹³⁰ International Sericulture Commission, 2013; Wardle, 1881

II.1.2.4 *Atlas silk*

Attacus atlas (Linnaeus), or atlas moth, is one of the largest moths in the world that widely inhabits the Indo-Australian bio-geographic region, China and Sudan. *A. atlas* spins light-brown cocoons nearly 6 cm long with peduncles of varying lengths (2-10 cm), somewhat similar to that of *Eri* occurs in nature (Figure 14).¹³¹ Little is known about the biology of this silkworm. The worm is polyphagous and it has been reported to feed on a variety of leaves including *Muntingia calabura*, *Annona muricata*, *Cinnamomum verum*, *Nephelium lappaceum*, *Psidium guajava*, *Sandoricum indicum*, *Citrus* sp., Cinchona tree (*Cinchona officinalis*), camphor laurel (*Cinnamomum camphora*), Arabica coffee (*Coffea arabica*), turmeric (*Curcuma longa*), *Elettaria cardamomum*, avocado (*Persea americana*), Soalu (*Litsea polyantha*) and other evergreen trees.¹³²

Its cocoons are collected as open-mouthed and thus spun into yarns.¹³³ The brown, wool-like silk of this variety is thought to have greater durability and is known as Fagara, or Tagore, or Kotkari muga (in Assam). In India, Atlas moths are cultivated for their silk in a non-commercial capacity;¹³⁴ there is no mention of this species in a United Nations report on sericulture in China in 1980. It is likely that *A. atlas* has had occasional significance in certain southeastern Asian countries through the centuries, but the minimal usage of silk from this species occurs today.¹³⁵

¹³¹ Babu, 2012

¹³² Ahmed, 2013

¹³³ Ghosh and Ghosh, 1995 See II.2.2.2 for further explanation.

¹³⁴ Jolly, *et al.*, 1979

¹³⁵ Peigler, 1993



Figure 14. *A. atlas* moth and cocoons

II.1.2.5 *Anaphe* silk

Anaphe silks are produced by silkworms of the genus *Anaphe* of family *Notodontidae*, widely distributed in the tropical regions of southern and central Africa. The species used in silk production are *panda*, *reticulata*, *venata*, *carteri* and *moloneyi* (Table 2). The worms are polyphagous and feed on various local plants including wild coffee (*Bridelia micrantha*), tamarind (*Tamarindus indica*), and Obeche (*Triplochiton scleroxylon*).¹³⁶ The silk from *A. panda* (*infracta*) is known locally as "boko", and those from *A. moloneyi* as "Tissniam-tsamia" and "koko".¹³⁷ In Nigeria, *A. panda* and *A. venata* are used to produce "Sanyan" cloth, which is woven with local cotton yarns.¹³⁸ "Gambari" or "Hausa" silk of Nigeria is also made of *Anaphe* silk; however, the source species is unknown.

While most silk moths spin a single cocoon, *Anaphe* moths aggregate to spin a communal nest and build a huge, yellow-brown cocoon (Figure 15). Akai reported that four hundred or more larvae spin individual cells into a soft, connective silk network that is bound to an outer, papery shell.¹³⁹ The fibre of this silk is yellow-brown, very fine, and particularly strong and more elastic than that of mulberry silk.¹⁴⁰ It is reported that *Anaphe* silk can be spun to a finer denier than *B. mori* spun yarn, however with less lustre, and often used in velvet and plush making in Europe.¹⁴¹ Its fibre structure is heterogeneous but mainly made up of β -sheets, and the mechanical property of the silk can be characterized by a higher breaking point than silk spun by either the *Bombycidae* or *Saturniidae*.¹⁴² *Anaphe* silk has a crescent-shaped cross-section, which is characteristic to

¹³⁶ Zethner, *et al.*, 2008

¹³⁷ Akerele, 1970

¹³⁸ Dudgeon, 1908

¹³⁹ Akai, 2002

¹⁴⁰ Jolly, *et al.*, 1979

¹⁴¹ Saurenhaus, 1929

¹⁴² Tanaka and Asakura, 2009

Thaumetopoeidae silks; the thread has a bamboo-like appearance created by the joints at intervals of 150–170 μm .¹⁴³



Figure 15. *A. panda* and *A. reticulata* (below) moths and diagram of *Anaphe* nest cocoon

¹⁴³ Akai and Nagashima, 1999

Table 2. Anaphe moths used for silk production in Africa.

Species	Synonyms	Distribution	Food plants
<i>A. carteri</i> (Walsingham, 1855)	<i>Epanaphe carteri</i>	Cameroon, DR Congo, Gabon, Equatorial Guinea, Gambia, Ghana, Ivory Coast, Nigeria	<i>Bridelia micrantha</i> , <i>Macaranga laurentii</i>
<i>A. moloneyi</i> (Druce, 1887)	<i>Epanaphe moloneyi</i>	Cameroon, DR Congo, Gambia, Ghana, Liberia, Nigeria, Rwanda, Uganda	<i>Cassia</i> , <i>Albiza</i> , <i>Macrolobium</i> , <i>Triplochiton scleroxylon</i> , <i>Rourea coccinea</i> , <i>Bridelia micrantha</i> , <i>Ficus platyphylla</i> , <i>Isobertinia doka</i> , <i>Pericopsis laxiflora</i> , <i>Senna siamea</i>
<i>A. panda</i> (Boisduval, 1847)	<i>Bombyx panda</i> <i>Anaphe infracta</i> <i>Anaphe leplaei</i>	Angola, Cameroon, DR Congo, Ethiopia, Equatorial Guinea, Gambia, Kenya, Malawi, CAR, Mozambique, Nigeria, Rwanda, South Africa, Tanzania, Uganda, Zambia, Zimbabwe,	<i>Coffea</i> , <i>Diplorhynchus</i> , <i>Bridelia micrantha</i> , <i>Chaetocarpus africanus</i> , <i>Cynometra alexandri</i> , <i>Triumfetta macrophylla</i>
<i>A. reticulata</i> (Walker, 1855)	<i>Anaphe ambrizia</i> <i>Anaphe euprepiaeformis</i>	Angola, DR Congo, Eritrea, Ghana, Ivory Coast, Kenya, Malawi, Mozambique, S. Africa, Tanzania, Togo, Zambia, Zimbabwe	<i>Dombeya</i> , <i>Grewia</i> , <i>Bridelia micrantha</i> , <i>Cola</i> , <i>Cynometra alexandri</i> , <i>Dipiorhynchus</i> , <i>Rhus chirindensis</i> , <i>Triumfetta macrophylla</i>
<i>A. venata</i> (Butler, 1878)	<i>Anaphe nyansae</i> <i>Anaphe sericea</i>	Angola, Cameroon, CAR, Congo, DR Congo, Equatorial Guinea, Gabon, Ghana, Ivory Coast, Nigeria, Tanzania, Togo	<i>Albiza</i> , <i>Bridelia micrantha</i> , <i>Triplochiton scleroxylon</i> , <i>Syzygium owariense</i>

II.1.2.6 *Coan silk? Pachypasa otus and Saturnia pyri*

Pachypasa otus (Drury) is the largest of the European Lasiocampidae. The name *otus* means owl. The female can reach a wingspan up to 12 cm, the males are smaller to around 9 cm. It is native of the Mediterranean biogeographic region and is considered the source of the Coan Silk of Antiquity.¹⁴⁴ Although the accounts of Aristotle and Pliny indicate that this silk came from a moth native to the island of Cos, some have advocated that the Coan silk industry was based instead on the much larger and more common *Saturnia pyri* (Denis & Schiffermiiller). *S. pyri* is a nocturnal species, also very large (9-17cm). It occurs from sea level in Western Europe to over 2000m altitude in the Levant and Iran.¹⁴⁵ The moth inhabits warm slopes with open scrub and other wood-rich locations, and the caterpillar lives on deciduous trees and shrubs. It is uncertain which moth was used, as no ancient fabrics of Coan silk have survived. Moreover, both species are no longer used in silk production for centuries, and there is no study available to provide enough data to compare their physical and mechanical properties with those from other silks. Their general characteristics are compared in Table 3.

The larvae of *P. otus* and *S. pyri* are both polyphagous, but they have different preferences: *P. otus* primarily feeds on the leaves of the cypress, juniper, pine, ash and various types of oaks; *S. pyri* prefers the leaves of the walnut, pear and apple trees etc. These big moths produce large cocoons, loosely formed with soft and woolly fibres.¹⁴⁶ *P. otus* forms oval white cocoon; *S. pyri* forms pear-shaped unsealed cocoon of greyish-brown colour. Wild and Babu noted that *S. pyri* spins the finer thread than *P. otus*.¹⁴⁷

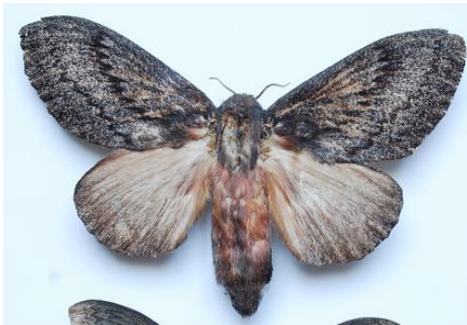
¹⁴⁴ Good, 2002, Leggett, 1949, Richter, 1929

¹⁴⁵ Pittaway, 2018

¹⁴⁶ Wild, 1970; Peigler, 1993; Babu, 2012

¹⁴⁷ Wild, 1970; Babu, 2012

Table 3. Comparison of *P. otus* and *S. pyri*

	Pachypasa otus	Saturnia pyri
Order	Lepidoptera	Lepidoptera
Family	Lasiocampidae	Saturniidae
Author	Drury, 1773	Denis & Schiffermüller, 1775
Moth		
Synonyms	Sphinx otus (Drury, 1770/1773) agrius (Olivier, 1790) dryophaga (Geyer, 1828) fulminea (Goeze, 1781) fulvescens (Kotzsch, 1932)	Bombyx pyri (Denis & Schiffermüller, 1775) Giant peacock moth Great peacock moth Viennese emperor
Distribution	Near and Middle East to Iran, Turkey Balkan Peninsula, Aegean Islands, southern Italy with Sicily	Near East to warmer areas of Europe (Iberian Peninsula to Turkey) Corsica and Sardinia
Host plants	Cypressus (cypress), Quercus (oak), Pinus (pine), Juniperus (juniper), Pistacia (pistacia), Thuja occidentalis (white cedar), Fraxinus ornus (ash), Pseudotsuga menziesii (Douglas fir). ¹⁴⁸	Acer (maple), Alnus (alder), Betula (birch), Cydonia (quince), Fagus (beech), Fraxinus (ash), Juglans (walnut), Malus (apple), Platanus (platanus), Populus (poplar), Prunus (prunus), Pyrus (pear), Rubus (blackberry, raspberry), Salix (willow), Syringa (lilac), Tilia (lime), Ulmus (elm), Corylus avellana (hazel), Aesculus hippocastanum (horse chestnut), Olea europaea (European olive). ¹⁴⁹
Cocoons	 	 

¹⁴⁸ Zolotuhin and Gurkovich, 2009

¹⁴⁹ HOSTS - a Database of the World's Lepidopteran Hostplants

II.1.3 Sericulture

Sericulture is the cultivation of silkworms to produce cocoons that yield silk filament to be made into thread. Silkworm rearing and weaving are environmentally sustainable agro-industrial activities practised globally, with the highest production in Asian and the Pacific area. Sericulture contributes substantially to many national economies. It is beneficial to the rural population in climatically suitable agricultural sectors because it provides either a primary or a secondary source of income for many workers, regardless of age and gender.¹⁵⁰ In the agrarian society, sericulture was a good source of household income in the winter season.

II.1.3.1 *Silkworm rearing*

Silkworms can breed once to multiple times a year depending on its voltine. The voltine of silkworm indicates their growing cycle, thus the number of broods per year. The major groups of silkworms are from East Asia and fall under the univoltine and bivoltine categories. Univoltine variety generates silk once a year. They are adapted to cold-warm weather zones such as Europe and hibernate during winter.¹⁵¹ The bivoltine variety is adapted to warm-hot weather zones, generating silk twice a year. These are generally found in China, Japan, and Korea.¹⁵² Polyvoltine (or multivoltine) variety generates silks up to 8 times a year. They are originally found in the tropical weather zone.¹⁵³ The quality of the filaments is inversely related to its productivity.

Female moths lay eggs then die soon afterwards. When the eggs hatch, the larvae

¹⁵⁰ Capinera, 2008

¹⁵¹ Trevisan, *et al.*

¹⁵² Korea's native variety was univoltine, and bivoltine varieties were introduced from 19th C.

¹⁵³ Trevisan, *et al.*

are carefully nurtured for the next 20-30 days and fed five times a day on chopped mulberry leaves. The larvae moult several times until they are ready to spin cocoons. The caterpillars have small openings called spinnerets under the jaw, which throws a proteinous substance that solidifies on contact with air to form the filament. Once the cocoon is completed over few days, the worm's life is ended by a process called 'stoving', or 'stifling', ¹⁵⁴ where the cocoons are heated to kill the pupae to prevent them hatching out and thus tearing the cocoon.

The growth of silkworms is largely affected by temperature and humidity. As silkworms are cold-blooded animals, the temperature will have a direct effect on various physiological activities. The seasonal differences in the environmental components considerably affect the genotypic expression in the form of phenotypic output such as cocoon weight, shell weight, and cocoon/shell ratio. The adaptability to environmental conditions in domesticated silkworms is different from those of wild silkworms because they are delicate and highly sensitive to environmental fluctuations after 5000 years of domestication. ¹⁵⁵

The appropriate temperature and humidity vary depending on the type of silkworm, age, voltine, climatic condition and nutritional condition. However, the optimum temperature range for normal growth of silkworms is between 20°C and 28°C with maximum productivity achieved in the 23°C to 28°C range. ¹⁵⁶ Temperatures above 30°C directly affect the health of the worm. If the temperature goes below 20°C, all the physiological activities are retarded, especially in early instars; with worms becoming too

¹⁵⁴ Babu, 2012

¹⁵⁵ Rahmathulla, 2012

¹⁵⁶ Benchamin and Jolly, 1984, Rahmathulla, 2012

weak and susceptible to various diseases.¹⁵⁷ Therefore, most countries engaged in sericulture are in temperate regions, within the latitude of 50° (N/S). The aptest location is the latitude range of 25~40° (N), and the leading countries of sericulture are all located within this. The current share of raw silk production is distributed between China (78%), India (19%), and then Japan, Thailand, Uzbekistan, Vietnam, Brazil and other countries with a temperate climate.

Humidity levels between 50%~90% are required for the adequate development of silkworms. The crossbreeds are more tolerant of changing humidity than pure breeds. Within the pure breeds, the East Asian breeds have superior value than European breeds. As per voltine, bivoltine are more durable than univoltine. The level of humidity requirement would decrease as the silkworm grows older. The standard humidity level for age 1 (number of moulting) to 2 is 85 ~ 90%, age 3 is 80%, and age 4 to 5 is 70-75%.¹⁵⁸

¹⁵⁷ Jabde, 2005

¹⁵⁸ Aruga, 1994

II.1.3.2 Silk reeling

Silk reeling refers to the unwinding of filaments from cocoons. Once cropped cocoons are sorted by their colour, size, shape and texture, and degummed. The degumming process is to remove the sericin gum that covers the surface of raw silks fibre: this reveals the fibroin, the core that holds the lustrous and soft characteristics of silk as a clothing material. The removal of sericin is necessary to reel out the fibres from the cocoon as well as to promote the better dyeability: the gummy surface often binds impurities that hinder the dye adsorption. A traditional method involves boiling the filaments in the light soap (alkali) solution; but at present, enzymatic solutions are often used for the silk production in industrial scale.

The content of sericin varies with the type of silk, but in general, a silk filament has about 25% of gelatinous gum (sericin) encasing the core (fibroin). However, sericin is not always to be completely removed—the degree of degumming varies depending on the desired type of threads. Some fabrics need threads with more strength or stiffness for the ease of weaving or to create specific effects. The ‘Crude’ or ‘*Ecrú*’, often referred to as ‘raw’ silk is lightly degummed fibres with maximum gum loss of 4%. The ‘Half-boiled’ or ‘*Souple*’ is partially degummed silk with 6-12% of the loss. ‘Lustre’ or ‘*Cuite*’ silk is of 18-30% gum loss thus is soft hand and high lustre.¹⁵⁹

In the cases of wild silk cocoons, it is impossible to reel out the filament for a couple of reasons. Wild silk species are difficult to cultivate, and thus most of the crops are gathered after the moths left the cocoon. They also differ in colour and texture and are often more difficult to dye than silk from the cultivated silkworm due to the strong gum and

¹⁵⁹ Karmakar, 1999

mineral contents of its surface, thus require harsh degumming treatment. Consequently, their fibres are scoured as short fibres (staples), gathered, combed and spun into yarns. However, *B. mori* silk is also cut and spun into yarns in case of double cocoons, those made from more than one or two larvae, happens to 2-3% of the population (per batch). This type of cocoon is called 玉繭 (*Tama-mayu*/ Jade cocoon)” or "同功繭 (Doupioni/ Double cocoon)”. It is tough to unreel because there are more than two strands of filaments tangled together; thus, it is mostly used as spun yarn—resulting in the characteristic irregular joints of the yarns. However, silk from this type of cocoon is famous for its superior elasticity and flexibility. *Shantung* and *Linshang* (silk dupioni) fabrics are made with such type of spun yarns.

Silk thread is measured in terms of denier (D).¹⁶⁰ Silk filaments are usually bundled together to make threads up to the desired denier scale and are sometimes given twists. This process is called Throwing. A denier number indicates the weight in grams per 9,000 meters of filament thread, thus the smaller the number, the finer the thread. If 9,000m of filament thread weighs 20g, it is called 20denier. The denier system applies both to silk and manmade filament threads but does not include waste; silk spun fibres and staple fibres—these are measured by the Tex system.^{161,162}

¹⁶⁰ A unit for the fineness of filament fibres.

$$\text{Denier (D)} = \frac{\text{Fibre weight (g)}}{\text{Fibre length (m)}} \times 9,000$$

A single strand of *B. mori* filament is about 3denier, and raw silk yarns are often made to 14 or 21denier.

¹⁶¹ A unit for the fineness of staple yarns.

$$\text{Tex (T)} = \frac{\text{Fibre weight (g)}}{\text{Fibre length (m)}} \times 1,000$$

The yarn count of 40 Tex means 1000 m of that yarn weighs 40 g.

¹⁶² Datta and Nanavaty, 2005

II.2 History and archaeology

II.2.1 Silk in China

The origin of silk cultivation is known to be China. The earliest written record of sericulture dates to 2640 BCE. According to the legend, the Empress Xi Ling Shi(西陵氏), also known as Leizu (嫫祖) was the first to raise silkworms.¹⁶³ A famous Tang dynasty scholar Zhao Rui(趙蕤) wrote on the epitaph of the Shrine of Leizu (嫫祖聖地):

嫫祖首創種桑養蠶之法, 抽絲編絹之術

“Leizu started the sericulture and the techniques to reel out the threads and to weave silk”

However, the archaeological records point to a much earlier date: the Neolithic Yangshao culture of China in the fourth millennium BCE.¹⁶⁴ Many carvings of a silkworm, chrysalis or moth have been found in various Neolithic sites in China, implying the importance of silkworm and silkmoth. The oldest artefact related to sericulture found to date is an ivory carving with four pairs of silkworms from Hemudu (ca.5000-4000 BCE) in Zhejiang.¹⁶⁵ Another carving of silkworm was found on the shard of black pottery from Meiyuan (ca.3000-2500 BCE) in Jiangsu.¹⁶⁶ Some pottery carved with a chrysalis pattern was found in Xizhuang village (ca.3500-3000 BCE), Nanyangzhuang in Hebei.¹⁶⁷ The

¹⁶³ A myth says Xi Ling Shi, the empress of Hungdi (黃帝), discovered silkworms while having a midday tea, and a cocoon fell in the tea. A fine thread started to release from the cocoon and that is said to be the first found of silk reeling (Kuhn, 1984).

¹⁶⁴ 仰韶文化

¹⁶⁵ 浙江省 河姆渡遺址

¹⁶⁶ 江蘇吳江梅堰新石器時代遺址

¹⁶⁷ 南楊庄西庄村

stone carving from Hongshan culture site (ca.3600-2000 BCE) in Houwa, Liaoning depicts a chrysalis with a pair of small wings.¹⁶⁸ Some similar designs are seen in jade carvings from various sites of Liangzhu culture (ca.3600-2000 BCE) in southern China.¹⁶⁹

The earliest direct evidence of sericulture is a small, artificially cut silk cocoon excavated from a Yangshao culture site in Xiyincun, Shanxi.^{170 171} The earliest findings of woven silk known to date are the silk tabby fabric and twined-warped fabric from a child's burial from Yangshao Culture (ca.3650 BCE) in Qingtai village in Yingyang, Henan.¹⁷² There are some fabric fragments including *Juan* (絹)¹⁷³ and a woven ribbon, dating from 2750 BCE found in Liangzhu Culture site in Qianshanyang, Huzhou, Zhejiang.¹⁷⁴ Many bronze pieces are dating from the Bronze Age Shang Dynasty left with the casts of what assumed to be from pattern-woven fabrics and embroideries.¹⁷⁵ Sylwan identified that the silk threads used in these textiles were filaments reeled from a cocoon, and not spun from short fibres.¹⁷⁶

Besides the development of domestic silk cultivation, China also has a long history of wild silk. The earliest written record is found in the *Tribute of Yu* in the *Book of Documents*, first written in the Zhou period (ca. 1046-256 BCE).¹⁷⁷

¹⁶⁸ 紅山文化后洼遺址, 辽宁省东港市

¹⁶⁹ 良渚文化

¹⁷⁰ 仰韶文化, 山西省 西陰村

¹⁷¹ Kuhn, 1984 says c. 2200-1700 BCE, while Zhao, 1998 and 1999 states it is from c. 3500-3000 BCE.

¹⁷² 河南省荊陽城東青台村的仰韶文化遺址

¹⁷³ A plain silk fabric made with multi-filament threads.

¹⁷⁴ 浙江省湖州錢山漾遺址

¹⁷⁵ Barber, 1991

¹⁷⁶ Sylwan, 1937, Sylwan, 1949

¹⁷⁷ 尚書(Shang shu) is one of the Five Classics of ancient Chinese literature, was written by 禹貢篇 (Yugong) in Zhou dynasty (ca. 1046-771 BCE). It is one of the earliest texts on geographic and ethnographic knowledge of China.

萊夷作牧. 厥筐壓絲

“Lai Yi people do pasturage, brought in their baskets the mountain mulberry silk. ”

This gives the idea of the utilisation of wild silk, although limited to the collection of the wild resource in the first stage of wild silk production: some accounts from the Han period gives more weight on wild silk production and its commercial value. In *Qifa*, written in the West Han dynasty (202 BCE- 9 CE):¹⁷⁸

野繭之絲以爲絃. 殆所謂蠟繭之韌堅者也. 野繭雖未若錦緞之美. 其來則久矣. 予嘗聞北關會寧, 楸城兩府開市. 彼人賣野繭紬. 而比我海西所產. 幅闊滋潤云. 凡山野之蠶. 取諸說考之. 則其所食楸, 檟, 櫟, 柞. 無所擇焉. 則非獨樗與壓也. 毋論山峽, 原野. 隨處廣種收繭. 則衣被之資. 可助蠶紬, 木綿之所不足矣.

“The thread of wild silk is used to make strings for the instrument, because its cocoon is quite tough. Although the beauty of wild silk does not come up to that of silk brocade, it has been in use for a long time. As I see and hear from Huinig of the North Gate, the market stands on both sides of the township, and they sell the spun fabric of wild silk. This is comparable to those from the west of our sea. It is said to be wide and lustrous. In general, according to what others say, the silkworms in the countryside eat from Catalpa, Amer cork tree, Oak, Hamamelis, and settle everywhere; which means not only fed on Ailanthus and mountain mulberry trees. They dwell in the mountain valley and the wild plain; the crops are fruitful regardless of the place. When made into garment, it is expensive as much as the cultivated silk fabric, and more valuable than cotton.”

According to the *Discourses on Salt and Iron* (81 BCE), ¹⁷⁹ the use of wild silk fabric was confined to the wedding garments for its high economic value:

繭網縑練者, 婚姻之嘉飾也

“The white *jian* silk, is used for the wedding garments.”

Jianchou (繭網), is a silk fabric made of 柞蠶, *Antheraea pernyi* silk, often called

¹⁷⁸ 七發 written by Mei cheng (枚乘) (?- 140 BCE)

¹⁷⁹ 鹽鐵論(*Yan Tie Lun*), written by Huan kuan(桓寬).

pongee or tussah. In the *Book of the Later Han*, attributed to the later Han period (25-220):¹⁸⁰

麻忒尤盛, 野蠶成繭, 被於山阜, 人收其利焉

“Hemp and beans grow abundantly, wild silkworms form cocoons, and their shells are around the mountains. People gather it to make a profit.”

In *Erya*; ¹⁸¹

屨: 山桑之絲, 其韌. 中琴瑟之絃.

蘇氏曰, 惟東萊爲此絲以之爲繒.

堅韌異常. 萊人謂之山繭

“*Yan* is a very strong yarn from Mountain Mulberry (*Morus mongolica*). It is used to make the strings of Chinese instrument *qinsè*. According to Su shi, in Donglai it is made into yarn and woven into exceptionally strong fabric. The people there call it mountain silk.”

A similar account appears in the Jin dynasty record the *Notes to things old and new* (265-420):¹⁸²

漢元帝永光四年, 東萊郡東牟山有野蠶爲繭.

繭生蛾, 蛾生卵, 卵著石, 收得萬餘石, 民以爲絮.

“In the fourth year of *Yongguang* period of the Emperor *Yuandi* of Han (40 BCE), in Donglai county’s East Mou mountain (now Muping County in Shandong), there is wild silkworm that forms cocoon. The cocoon becomes the moth, and the moth lays eggs on more than ten thousand stones. People make *xu* (floss, wool) with it.”

Also, in *Guangzhi* (265–316):¹⁸³

¹⁸⁰ 後漢書(Hou Hanshu). Written by Fan Ye (范曄, c. 398-445). 卷一上·光武帝紀第一上, 十二月戊午

¹⁸¹ 爾雅 (author unknown) is the oldest surviving Chinese dictionary or Chinese encyclopaedia known, contains the glosses from 3rd century BCE and so on. The title means roughly ‘approaching what is correct, proper, refined’ (Coblin, 1993)

¹⁸² 古今注(*Gujinzhu*) is an encyclopaedia written by Cui Bao (崔豹, active 290-306)

¹⁸³ 廣志. Written by Guo Yigong(郭義恭).

野蠶食柞葉, 民以作綿

"Wild silkworms eat oak leaves; people make it into *mian* (floss)."

It appears that the community recognised the economic value of wild silk and the production of wild silk textiles were well established. However, it was still limited to the simple collection and processing of spontaneous resource. These hand workshops are only concentrated in the Shandong Peninsula, Penglai, Yexian area.¹⁸⁴

In the late Ming Dynasty, the wild silkworm industry started to bring a human intervention in the breeding process of the wild silkworm. The recovery rate of cocoons significantly increased with the man-made control, as well as the quality of cocoons; eventually helping to promote the production of wild silk textiles.¹⁸⁵ In the late Ming to early Qing dynasty, the pasturing of wild silk (semi-cultivation) and production technology were spread outside the Shandong area to Henan, Liaoning, Shaanxi, Guizhou, Sichuan, Anhui and other places. Many texts on wild silk production were compiled in this period, including the *Mountain planting of Silkworms* (養山蠶成法), the *List of the Ailanthus silk* (樗繭譜), and the *records on wild silkworms* (野蠶錄).¹⁸⁶

The *Records on Wild silkworms* lists 15 types of wild silks in China: 柞蠶, 桑蠶, 柘蠶, 樗蠶, 椒蠶, 荊蠶, 柰蠶, 柳蠶, 娑羅蠶, 榆蠶, 松蠶, 橘蠶, 棘蠶, 欒蠶 and 蕭蠶.¹⁸⁷ 柞蠶 (*zha can*) is the *Antheraea pernyi* silkworm, the most popular breed amongst Chinese wild

¹⁸⁴ Han and Tang Silk Network, 2015

¹⁸⁵ Ibid.

¹⁸⁶ Kang, 2004

¹⁸⁷ 野蠶錄(*Ye can lu*), written by Wang Yuanting (王元綏) in 1898

silks. 桑蠶 (*sang can*), *Bombyx mandarina*, is similar to *B. mori* but much smaller in size.

柘蠶 (*zhe can*) feeds on *Cudrania tricuspidata*, and 檮蠶 (*Chu can*) is a mountain mulberry silkworms that feeds on Ailanthus oak (*Ailanthus altissima*), but according to the *List of the Ailanthus silk*, they can also feed on other types of oak trees.¹⁸⁸ 椒蠶 (*jiao can*) grows on the prickly ash (*Zanthoxylum schinifolium* S&Z), and the woven fabric would still hold the scent of the prickly ash. 荊蠶 (*jing can*) is a green silkworm similar to *A. pernyi* that grows on the Chaste tree (*Vitex agnus-castus*). 柰蠶 (*nai can*) grows on the Crabapple tree (*Malus asiatica* N.); 柳蠶 (*liu can*) grows on the Salt cedar (*Tamarix chinensis*); 娑羅蠶 (*suoluo can*) grows on Sal tree (*Shorea robusta*) in the tropical south; 榆蠶 (*yu can*) grows on the Elm tree (*Ulmus davidiana* var. *japonica*); 松蠶 (*song can*) grows on the Pine tree; 橘蠶 (*ju can*) grows on the tangerine tree (*Citrus tachibana*) and makes a very light, white coloured cocoon unsuitable to threading; 棘蠶 (*ji can*) is red-green coloured and grows on the Chinese evergreen oak (*Quercus myrsinaefolia* B.); 樂蠶 (*luan can*) grows on the *Dianthus sinensis* L., only in the south; and 蕭蠶 (*xiao can*) grows on the mugwort. Except for the first four, most of these silkworms are barely known outside China, making the matching of modern scientific names to them particularly hard. As these species were named after their habitat (i.e. food source), they testify to the wide variety of wild silk

¹⁸⁸ Zheng and Mo, 2002

species distributed in China.

As the technology matured and spread, wild silk production became an important secondary economic activity for mountain farmers.¹⁸⁹ However, the opening of the ports in the late Qing dynasty largely affected major Chinese exports, such as tea and domesticated silk; causing the decline of the local silk industry with the arrival of imported versions. Therefore, wild silk production was a good resolution to meet domestic demand, due to its low labour cost compared to silk cultivation and readily available local tree species.¹⁹⁰

¹⁸⁹ Han and Tang Silk Network, 2015

¹⁹⁰ Wang, 2002; Kang, 2010

II.2.2 Silk in India

The origin of the utilisation of silk in the Asian subcontinent is unclear. Although a number of wild silk species feature in traditional Indian textile industry, its inception is unknown.¹⁹¹ Royale pointed out that it must often have been confounded with Chinese silks, as the sources for both indigenous Indian silk and Chinese silk were so little known that the accounts of silk are frequently confounded with those of cotton.¹⁹²

In the Hindu epic *Mahabharat* (thought to have been written in BCE), the silkworms called “*Patta-keetas*” were brought from China as a gift to Yudhishtira,¹⁹³ during his reign at Indraprastha (Endraprastom).¹⁹⁴ In *Amarakosha* (ca. 4th century), silk was called “*Cheenam Sokam*”, literally meaning ‘the cloth of China’. According to a record, Chinese silk was imported into India during Mauryan (322-183 BCE) or early Han era.¹⁹⁵ In *Arthashastra* of Kautilya (ca. 5-4th BCE), there are three types of silk among the king’s treasury: *patrorna*, *kauseya*, and *cinapatta*. *Patrorna* is known to be a type of wild silk, spun from the cocoons collected from various trees.¹⁹⁶ *Kauseya* is mentioned in the *Ashtadhyayi* of Panini (ca. 6-4th BCE), as the silk made from locally cultivated silkworms, *koseyya*.¹⁹⁷ In the *Great Tang Records on the Western Regions* (646)¹⁹⁸ written by Xuanzang (ca. 602–664),¹⁹⁹ Indian silk is described by its indigenous name, *kauseya*. As seen in *Amarakosha*, *cinapatta*, or *cinansuka*, was the silk imported from China. The examples of Sanskrit vocabulary and literary references indicate that indigenous wild silks

¹⁹¹ Jolly, *et al.*, 1979

¹⁹² Royale, 1986

¹⁹³ The king of Indraprastha and later of Hastinapura (kuru). Sanskrit: युधिष्ठिर.

¹⁹⁴ Royle, 1840; Rani, 2006

¹⁹⁵ Dale, 2008

¹⁹⁶ Scharfe, 1993; Ray, 2004

¹⁹⁷ Ray, 2004

¹⁹⁸ 大唐西域記(Da Tang Xiyu Ji)

¹⁹⁹ 玄奘. A Chinese Buddhist monk, scholar and traveller.

and Chinese silk were clearly distinguished from each other from the earlier period, and the use of indigenous silks was well established.

The direct archaeological evidence for silk in South Asian prehistory is limited to a single find from Nevasa, a Chalcolithic site in Maharashtra in central India, dating to c. 1500-1050 BCE.²⁰⁰ It is thought to be silk other than Chinese silk.²⁰¹ Some of the threads from the collection of Harappa and the Chanhudaro site excavations are reported to be of wild silk species.²⁰² If true, this indicates that the utilisation of wild silk was practised in more than one Indus settlement during the height of Indus urbanism, nearly a millennium before the Nevasa finding.²⁰³ Based on ancient vocabulary and records, it is highly probable that people in the subcontinent utilised the local wild silks well before the introduction of Chinese silk. However, unfortunately, the exact identity of these alleged wild silks is yet unknown.

²⁰⁰ Allchin and Allchin, 1982

²⁰¹ Gulati, 1961; Janaway and Coningham, 1995; Alam, 1999

²⁰² Curated at the Museum of Fine Arts, Boston.

²⁰³ Good, *et al.*, 2009

II.2.3 Silk on the Road

As is well known from the history of The Silk Road, although silk products found its westward way, the skills of silk cultivation were diligently kept secret for millennia. Silk was a highly sought-after luxury traded goods, due to its scarcity of material and the labour and resource intensive nature of its production. The early evidence indicates that the techniques of degumming and reeling have also been considered exclusive to the Chinese silk industry.²⁰⁴ It is evident that the silk producers have hugely benefitted from protecting the means of production—silkworms, cocoons and even the seeds of mulberry were banned from leaving China.²⁰⁵

However, the secret of sericulture began to reach close neighbours. According to the *Book of Han, Treatise on Geography* by Ban Gu,²⁰⁶ sericulture was introduced to Korea around the twelfth century BCE. According to a legend recorded by Xuanzang and in the *Prophecy of the Li country (Li-yul lun-bstan-pa)*,²⁰⁷ the sericulture was brought to Khotan in the first few centuries by the hands of a Chinese princess who married the king of Khotan.²⁰⁸ Stein has found several wooden plaque paintings describing this legend in Dandān-Uiliq (Figure 16).²⁰⁹ Xuanzang also wrote about a thriving economy in Khotan's sericulture that he witnessed on the way from India (ca. 640-645).²¹⁰ There were remains of silk cocoons, dating from the 3rd century, alongside the seeds and desiccated trunks of

²⁰⁴ Ibid.

²⁰⁵ Zethner and Koustrup, 2007

²⁰⁶ 漢書, 地理志 (volume 28), written by 班固

²⁰⁷ A Tibetan text concerning the Buddhist history in Chinese Turkestan contains a list of the Buddhist kings of Khotan. It was compiled about 746 CE based on the older sources.

²⁰⁸ It is unclear when this event took place, and is disputed by a number of scholars, ranges from the first century to fourth or fifth century CE. See Hill, 2009.

²⁰⁹ 丹丹烏里克

²¹⁰ Stein, 1907; Hill, 2009s

mulberry trees found in Niya.²¹¹ These findings suggest that the sericulture took place here, and some simple varieties of silk fabrics may have been produced.²¹²

The silk production became a major industry and one of China's chief exports in the Han dynasty (260 BCE-220).²¹³ It seems that the economic values of silk are established in this period, when the government had an active interest in the commodification of it. For example, silk bolt was used as salary in lieu of cash to pay for Han garrison in Central Asia.²¹⁴ At the beginning of the second century BCE, silk was a popular peace offering to the Xiongnu tribes.²¹⁵ The first official exportation of silk outside of China was made when Wu Di (140-87 BCE),²¹⁶ the sixth emperor of Han, tried to establish a formal trade with the Da Xia (Bactria).²¹⁷ Until then, sericulture and silk weaving were confined to China and its close neighbours. However, with the beginning of the Silk Road trades, the knowledge of silk dyeing and weaving spread to Central Asia, including Sogdian Transoxiana and the Persian Empire of the Sassanid (224-651).²¹⁸

By the second century, India was also exporting its raw silk and silk cloth to Persia under the Parthian empire (247 BCE - 224). According to the *Periplus Maris Erythraei* (ca. 1st century), silk in the forms of floss, threads and fabrics were shipped from the inland city of Thina to the great market of Barygaza through Bactria, and to the coast of Limyrica (or Damirica) via the river Ganga.²¹⁹ Barygaza was the port for the sea trade between India and the West during the Indo-Greek period: The Red Sea for Egypt and the Persian Gulf

²¹¹ 尼雅遗址

²¹² Hansen, 2015

²¹³ The Editors of Encyclopædia Britannica

²¹⁴ Hansen and Rong, 2013

²¹⁵ 匈奴

²¹⁶ 汉武帝

²¹⁷ 大夏

²¹⁸ Wendelken, 2014

²¹⁹ Gilroy, 1853; Ray, 2004

for Babylonia. From Babylonia, the silks from the East were spun and woven into clothing for the Roman market in Palestine and Syria.²²⁰ The history of silk in the Near East begins in Syria during the Roman Empire, when the caravans brought the silks from China and India. The silks were brought to the Syrian cities of Antioch, Berytos and Tyre, and local traders and Christian pilgrims took them to Italy, France, and Spain. Although the monopoly of sericulture was over, and the Mediterranean cities started to produce their own silk, China continued to be the major supplier of high-quality silk until the large-scale industrialisation of silk production during the Song dynasty.²²¹



Figure 16. Votive panel illustration (木版畫) depicting the legend of the introduction of sericulture to Khotan. The central figure is a Chinese princess who smuggled mulberry seeds and eggs of the silk moth to Khotan, hiding them in her head. (British Museum 1907, 1111.731907, 1111.73)

²²⁰ Yarshater, 1983; Tarn, 2010

²²¹ Feltham, 2009

II.2.4 Silk in the West

The development in silk weaving and production in the West owes a great deal to the material and technology from China. However, a range of literary evidence implies that some native silks were already being produced in the Mediterranean and the Middle East, before the introduction of Chinese silk and sericulture in the region. Barber suggests that wild silk manufacture in Europe started in the Classical period to imitate Chinese imports.²²² Aristotle notices the production of indigenous silk as early as 4th century BCE in association with the Aegean island of Cos. These native species were likely to be the wild species that differs from the Chinese *B. mori* in many aspects: due to their undomesticated nature, wild silk fibres are spun into yarns, thus bulkier, and less lustrous than that of domesticated silk.²²³ Nevertheless, evidence of indigenous silk production in the West must be carefully reviewed, as they can suggest many clues to the material culture and the agriculture of the period.

II.2.4.1 *Introduction of Chinese silk and sericulture*

The silk from the East began to reach the Hellenistic world via the Persian merchants who would exchange it for gold, ivory, horses or precious stones. The Roman Empire value silk as high as gold and Hellenistic Greece appreciated the high quality of the Chinese silk and made efforts to learn to adopt its production secrets.

At first, Chinese merchandises consisted mainly of textiles than raw silks.²²⁴ The thick white Chinese fabrics were taken to the Mediterranean cities of Sidon and Tyre to be

²²² Barber, 1991

²²³ See II.3 for detailed information.

²²⁴ Marquardt , *et al.*, 1879; Hirth, 1885.

dyed, chiefly in purple, then unravelled and rewoven into light gauzes in Egypt. In Lucan's description of Cleopatra:

...candida Sidonio perlucent pectora filo, Quod Nilotis acus conpressum pectine Serum Solvit et extenso laxavit stamina velo.

“Her breasts were revealed by the fabric of Sidon, which, close-woven by the shuttle of the Seres, the Egyptian needle-worker pulls out, and loosens the thread by stretching the stuff”.²²⁵

The process of reforming Chinese silk is corroborated in Roman and Chinese sources.²²⁶ Pliny (23-79) wrote:

Seres, lanicio silvarum nobiles, perfusam aqua depectentes frondium canitiem, unde geminus feminis nostris labos redordiendi fila rursusque texendi: tam multiplici opere, tam longinquo orbe petitur ut in publico matrona traluceat.

“The Seres, so famous for the wool that is found in their forests. After steeping it in water, they comb off a soft down that adheres to the leaves; and then to the females of our part of the world, they give the twofold task of unravelling their textures, and of weaving the threads afresh. So, the manifold is this labour, and so distant are the regions which are thus ransacked to supply a dress through which our ladies may in public display their charms”.²²⁷

Theodoretus of Cyrrhus (393-ca.457) also recorded that Roman women and children unravelled the imported Chinese and Persian silk textiles so that the threads could be rewoven in a style that was more appealing to the Roman market.²²⁸ The account on Daqin²²⁹ in the *Comprehensive investigations based on literary and documentary sources*²³⁰ records:

又常利得中國縑素 解以為胡綾紺紋 數與安息諸胡交市於海中

²²⁵ Pharsalia, X.140

²²⁶ Schoff, 1912

²²⁷ N. H. VI, 20

²²⁸ Herrin, 2009

²²⁹ 大秦, the Roman Empire

²³⁰ 文獻通考(Wenxian Tongkao), compiled by Ma Duanlin (馬端臨) in 1317.

“They also make profits by *jiānsù*²³¹ from China, which they unravel and make into the foreign style purple silk damask; this is why all the foreign states trade with Parthia”

Physical evidence of re-woven silks has been recovered from Palmyra, Dura-Europos, and Coptic sites in Egypt. The damasks and embroidered silks dating from the first to the third centuries are made of tightly spun yarns; a practice not found in China.²³² The traditional way of Chinese silk thread making does not involve twisting and use of a bundle of filaments so that they yield fabrics and embroideries with a more lustrous and smoother surface. Hence, these findings are unlikely to have been made in China but were instead woven locally or at least made somewhere outside China with imported threads.²³³

The practice of spun yarns could have been developed for a number of reasons: first, for blending with other staple fibres, such as wool and linen; second, to weave a thin gauze-like fabric (it is necessary to give strength to threads to endure the tensions); third, to use inferior quality silk that required the threads to be spun.²³⁴ It is not hard to understand that Romans have done so to save cost: the practice of blended fabric²³⁵ or union fabric²³⁶ is the most common way to use a precious material at an affordable price, or to supplement some negative properties of the primary material. This sort of fabric based on Chinese silk was generally called *sericae vestes* but received a distinct name when the Romans began to import the pure woven silk goods.²³⁷ Elagabalus (reign 218-222) is said to be the first Roman who wore the pure silk garment (Lamprid. *Elag.* 26).²³⁸ *Sericam* is now divided to *holosericam* and *subsericam*: *holosericam* is a garment made of silk, and *subsericam* is

²³¹ *Jiansu* is a thick white silk made with multiple threads yarn, often refers to the fine silk used for books and paintings.

²³² Usher, 2013

²³³ Day, 1950

²³⁴ Wendelken, 2014

²³⁵ A fabric woven with yarns made of two or more different types of fibres. Ex. cotton and silk blend.

²³⁶ A fabric woven with different fibre yarns in warp and weft. Ex. Linen warp and silk weft union weave.

²³⁷ Smith, *et al.*, 1890

²³⁸ “Primus Romanorum holoserica veste usus fertur, cum iam subsericae in usu essent”

made of mixed material including silk; *tramoseria* is explained as a union fabric of linen warp and silk weft (Isid. *Orig.* XIX. 22).²³⁹ Many literary sources indicated that *holoserica* was subject to a sumptuous regulation:

vestem holosericam neque ipse in vestiario suo habuit 5neque alteri utendam dedit. et cum ab eo uxor sua peteret, ut unico pallio blatteo serico uteretur...

“A virtuous Aurelian would neither keep *holoserica* in his own wardrobe nor present to anyone else for his use; when his wife besought him to keep a single robe of purple silk”.²⁴⁰

Tacitus even enacted a decree to forbid all man to wear a garment made *holoserica* (Vospisc. *Tacitus* X).²⁴¹ The extravagant clothing habit of Elagabalus was not taken over by Severus Alexander (reign 222-235); he never wore *holoserica* or gave away even *subserica* (Lamprid, *Sev. Alex.* XL).²⁴²

The Western Empire imported large amounts of raw silks to supply its private workshops. According to the *Edict of Diocletian* (301), a wealthy Roman could easily procure a damask loom, raw silk and a skilled weaver to produce silk textiles for family use.²⁴³ Procopius (ca. 500-554) informs us about the staple industry of silk garment manufacture that existed in the ancient times in Beirut and Tyre, two cities of Phoenicia (*Anecdota*, 25).²⁴⁴ Despite the thriving silk trades in these periods, it has long been suspected that the consumption of silk in the West exceeded the amount that China could supply.²⁴⁵ Maybe it was the taste for different styles,²⁴⁶ or the desire for pure silk garments.

²³⁹ “*Holoserica tota serica; ὅλον enim totum. Tramoserica stamine lineo, trama ex serico. Holoporphrya tota ex purpura.*”

²⁴⁰ Vopisc. *Aurelian.* XLV. 4-5.

²⁴¹ “*holosericam vestem viris omnibus interdixit*”

²⁴² “*Vestes sericas ipse raras habuit; holosericam numquam induit, subsericam numquam donavit*”

²⁴³ Wild, 1984

²⁴⁴ “ἱμάτια τὰ ἐκ μετάξης ἐν Βηρυτῷ μὲν καὶ Τύρῳ πόλεσι ταῖς ἐπὶ Φοινίκης ἐργάζεσθαι ἐκ παλαιοῦ εἰώθει. οἳ τε τούτων ἔμποροί τε καὶ ἐπιδημιουργοὶ καὶ τεχνῖται ἐνταῦθα τὸ ἀνέκαθεν ᾔκουν, ἐνθὲνδε τε ἐς γῆν ἅπασαν φέρεσθαι τὸ ἐμπόλημα τοῦτο ζυνέβαινε.”

²⁴⁵ Usher, 2013

²⁴⁶ Hill, 2009. This includes the preference for lighter weave, and also for the colour: purple was not one of

Anyhow, it seems that the precarious supply was also affected by the monopoly of the Byzantine Empire: when the imperial office controlled all aspects of silk production to give priority and to protect the imperial silk workshops, driving many independent silk workers away out of Byzantine. Nonetheless, Justinian I (482-565) set a low maximum price for imported raw silk to further protect the imperial silk business, which resulted in foreign merchants' refusal to do business with the Empire.²⁴⁷

With the emerging needs for the local supply of raw silks, Justinian I persuaded two Nestorian monks to smuggle Chinese silkworms to Constantinople in about 552.²⁴⁸ It is known that the knowledge of sericulture using both indigenous black mulberry and the Chinese white mulberry was confined within Byzantine until the 12th century. However, it has been suggested that the loss of eastern provinces to Islamic conquests of 7th and 8th centuries forced the Byzantine sericulture to move farther west and north.²⁴⁹ Under Muslim rules, the liberal regulations on silk business allowed the knowledge of the sericulture and silk weaving to be spread to Lebanon, Egypt, northern Africa, Spain and Mediterranean islands of Crete and Sicily.²⁵⁰

The techniques of silk production began to spread across Western Europe with the Crusades. The Roger II of Sicily (1095–1154) attacked Corinth and Thebes (1148) during the Second Crusade and deported the silk workers and their equipment to Palermo and Calabria. By the Fourth Crusade at the beginning of the 13th century, the Byzantine silk industry was largely diminished. Italy, on the other hand, developed a domestic silk industry with the immigrants from Constantinople; who also settled in Avignon to furnish

the major shades of the East.

²⁴⁷ Lopez, 1945

²⁴⁸ There many different opinions on the year of the event. See Appendix 1 for further readings, Procopius and Theophanes recorded the anecdote.

²⁴⁹ Jacoby, 1992; Muthesius, 1993; Muthesius, 2008; Wendelken, 2014

²⁵⁰ Wendelken, 2014

the popes.²⁵¹ Sicily, Venice, Florence and Lucca became the leaders of European silk production.²⁵² In the 15th century, Louis XI founded a national silk industry in Lyon to reduce the trade deficit with Italian states. With the growing popularity of original designs, Lyon rose as the centre of the European silk market in the following century.²⁵³ In the 1680s, the revocation of the Edict of Nantes made French Huguenots to immigrate to England; many of whom were skilled silk workers, and they brought their business with them. However, it did not flourish due to the unsuitable climate.²⁵⁴

By the end of the 19th century, the European silk industry began to decline, due to epidemic silkworm diseases: the pébrine, caused by the microsporidia *Nosema bombycis*; the grasserie, caused by a virus; the flacherie, caused by eating infected mulberry leaves; and the white muscardine disease, caused by the fungus *Beauveria bassiana*. The epidemics grew to a massive scale, and after having attacked the silkworms, other viruses began to infect the mulberry trees. Although the pébrine disease started to decline from 1870 owing to the effort of Louis Pasteur (1822-1895),²⁵⁵ the European silk production never recovered. The market was already replaced with the imports from Asia, particularly from China and Japan, accelerated by the opening of the Suez Canal in 1869:²⁵⁶ and once again, the centre of the silk production returned to Asia.

²⁵¹ Gille, 1986

²⁵² Muthesius, 2007

²⁵³ Duby and Mandrou, 1964; Feltwell, 1991; Chauvy, 2000

²⁵⁴ Langer, 2001

²⁵⁵ Berche, 2012

²⁵⁶ Latham and Kawakatsu, 1994

II.2.4.2 Ancient accounts of silk

II.2.4.2.1 Bombycinae and Sericae

In the ancient Greek and Latin literature, silk was called by many names: *bombycinae*, *sericae*, and *metaxa*.²⁵⁷ *Metaxa* indicates the raw silk while the others are used for a certain type of silk textiles.²⁵⁸ The most frequently used terms in the literature are *bombycinae* and *sericae*. *Bombycinae* was first used to indicate silk, and the first instance of *sericae* (σερικόν) is from the second half of the 1st century BCE; *metaxa* (μέταξα) appears in the 2nd century.²⁵⁹ While Pliny located the *bombycum* in Assyria (a silkworm occurs in Assyria),²⁶⁰ Propertius located *bombyce* in Arabia (nor when my girl shimmers in silks of Araby).²⁶¹ Hildebrandt considers such indications of provenance are references to the ports of trade rather than places of production.²⁶² On the other hand, Isidorus and Seneca located *sericae* in Seres, the first place to provide it and famed for silk.²⁶³ ²⁶⁴ A good comparison between *bombycinae* and *sericae* is given in the *Etymologiae*:

Bombycina e bombyce vermiculo qui longissima ex se fila generat, quorum textura bombycinum dicitur; conficiturque in insula Coe. Apocalama. Serica a serico dicta, vel quod eam Seres primi miserunt.

“Silk (*bombycina*) cloth is named from the silkworm (*bombyce*), which produces extremely long threads from itself; something woven of these threads is called ‘silken’ (*bombycinus*). It is manufactured on the island of Cos. Silk (*serica*) is named from the Seres, first made it available”.²⁶⁵

²⁵⁷ Richter, 1929. The modern word for silk in Greek is μετάξι (metáxi).

²⁵⁸ Smith, *et al.*, 1890

²⁵⁹ Hildebrandt, 2017

²⁶⁰ Pliny, N.H. XI. 25: “*bombycum, in Assyria proveniens*”

²⁶¹ Propertius L.II.3,15: “*nec si quando Arabo lucet bombyce puella*”

²⁶² Marquardt, *et al.*, 1879

²⁶³ Isid. Orig. XIX. 27: “*Sericum dictum quia id Seres primi miserunt;*”

²⁶⁴ Seneca, Thy. 379: “*Seres vellere nobiles*”

²⁶⁵ Isid. Orig. XIX. 22

Also, in Pausanias's Description of Greece:

ἔστιν ἐν τῇ γῇ ζωύφιόν σφισιν, ὃν σῆρα καλοῦσιν Ἕλληνες, ὑπὸ δὲ αὐτῶν Σηρῶν ἄλλο πού τι καὶ οὐ σῆρ ὀνομάζεται.... ταῦτα τὰ ζῷα τρέφουσιν οἱ Σῆρες οἴκους κατασκευασάμενοι χειμῶνός τε καὶ θέρους ὥρα ἐπιτηδεύουσιν

“There is in the land of the Seres (Σῆρες) an insect which the Greeks call *ser* (σῆρ), though the Seres themselves give it another name. These creatures are reared by the Seres, who build them houses adapted for winter and summer”.²⁶⁶

In summary, *bombycinae* indicates silk garments derived from silkworm (*bombycae*) that comes from Assyria or Arabia, which has been locally produced; *sericae* specifically address the cultivated, imported silk of China (*Seres*). *Sericae* was imported from China, usually overland through Samarkand to the Persian Gulf, thence to Phoenicia or Egypt, and finally to Rome.²⁶⁷ For not only its high quality but also the long process of trade would have added more value to it, the Chinese silk (*serici*) was said to be worth its weight in gold in Rome.²⁶⁸ According to the *Edict of Diocletian* (301), a pound of *blattae serica* (μεταξάβλάττη), purple dyed silk fabric, was worth 150,000 denarii.²⁶⁹ Thus, the distinction between *sericae* and *bombycinae*, besides their identity, was also made in the price and social recognition as described in *Metamorphoses*:

carbasinis et bombycinis iniecti... serico contecta amiculo.

“The priests wear robes made of linen and bombycinae, where the goddess is wrapped in a silken mantle”²⁷⁰

²⁶⁶ (Paus.VI.26)

²⁶⁷ Procop. *Anecd.* 25; Arrian. *Per. M. Eryth.* 56.

²⁶⁸ Vopisc. Aurelian. XLV. 5. ““Absit ut auro fila pensentur”, “libra enim auri tunc libra serici fuit”.

²⁶⁹ According to his system of Coinage, 1 lb. of gold = 50,000 denarii. Thus approximately 4000USD.

²⁷⁰ Apul. *Met.* VIII. 27

II.2.4.2.2 Coa Vestis

Coa vestis, also *Coa*, is the light gauze fabric produced in the island of Cos. It was first mentioned in the oldest account of silk in the West, written in the 4th century BCE by Aristotle:

Ἐκ δέ τινος σκόληκος μεγάλου, ὃς ἔχει οἶον κέρατα καὶ διαφέρει τῶν ἄλλων, γίγνεται πρῶτον μὲν μεταβάλλοντος τοῦ σκόληκος κάμπη, ἔπειτα βομβυλῖς, ἐκ δὲ τούτου νεκύδαλος· ἐν ἑξὶ δὲ μηνὶ μεταβάλλει ταύτας τὰς μορφὰς πάσας. ἐκ δὲ τούτου τοῦ ζώου καὶ τὰ βομβύκια ἀναλύουσι τῶν γυναικῶν τινες ἀναπηνιζόμεναι κάπειτα ὑφαίνουσιν· πρώτη δὲ λέγεται ὑφῆναι ἐν Κῶ Παμφίλῃ Πλάτew θυγάτηρ.

“Out of a certain large larva, which has as it might be horns; and is different from the rest; there arises first a caterpillar, then a cocoon, then a moth: the transformation takes six months. Some women unwind these cocoons, by reeling the thread off, and then weave fabric from it; the first to do this weaving is said to have been a woman of Cos named Pamphile, daughter of Plates”.²⁷¹

A similar, but more elaborated description about this silk is given by the Roman naturalist and encyclopaedist Pliny the Elder (23-79 CE):

quantum inter haec genus est bombycum, in assyria proveniens, maius quam supra dicta (...) ex grandiore vermiculo gemina protendens sui generis cornuum urica fit, dein quod vocatur bombylis, ex ea necydallus, ex hoc in sex mensibus bombyx. (...) prima eas redordiri rursusque texere invenit in coo mulier pamphile, plateae filia...ut denudet feminas vestis. bombycas et in coo insula nasci tradunt, cupressi, terebinthi, fraxini, quercus florem imbribus decussum terrae halitu animante (...) atque ita subnasci sui generis plumas, quibus vestitos ad alia pensa dimitti. quae vero capta sint lanifica, umore lentescere, mox in fila tenuari iunceo fuso. nec puduit has vestes usurpare etiam viros levitatem propter aestivam(...) assyria tamen bombyce adhuc feminis cedimus.

“Another species of insect is the silk moth, occurs in Assyria. (...) A very big larva ...turns into a moth in six months. (...) The technique of unravelling the cocoons and spinning the thread was first invented on Cos by Pamphile, the daughter of Plateas (...) to reduce women’s clothing to nakedness. Silkmoths are also reported to be born in the island of Cos, where cypress, terebinth, ash and oak are. (...) the tufts of wool plucked off from cocoon are thinned out into threads with a rush spindle. Nor have even men been ashamed to make use of these dresses, because of their lightness in summer (...) so far leave the Assyrian silkmoth to women”.²⁷²

²⁷¹ H. A. V.19

²⁷² N.H. XI. 25-27

Putting aside the accountability of these records—the silkmoth of both Aristotle and Pliny, is the Assyrian silkmoth that Pamphile of Cos Island used for weaving. If it existed, this Assyrian silkmoth must be a wild type given the descriptions of worms and the fibres being spun into yarns, and the terms used for them—*βομβύκια* and *bombyce*. Forbes argues that Aristotle’s silkworm was something apparently familiar to him and his readers of Western Asia Minor, Greece, and the islands between, thus not the one from China.²⁷³

The production and export of this precious textile were centred on the Greek island of Cos, and many Romans were attracted to the island because of this industry.²⁷⁴ Coan silk is frequently mentioned by the Latin writers of the Augustan age.²⁷⁵ From their expressions, we learn that it was a fine and transparent textile, worn by women of doubtful reputation (Figure 17). It was sometimes dyed purple, or even adorned with gold weave, which fits the criteria of a favoured fashion of the contemporary Greco-Roman world. Horace said “in Coan silk you may see her almost as if naked (*Satires*.I.2.101)”²⁷⁶ and Martial said “a woman’s body shines through silk (*Epigrams*, 68)”²⁷⁷ From its revealing nature, the garments made of Coan silk was certainly not taken as virtuous. As Propertius criticised the seductive fashion of young women “to step out with adorned hair and flutter the sheer folds of a Coan dress (*L*, I. 2)”²⁷⁸ and younger Seneca congratulated his mother Helvia for “never have fancied the kind of dress that exposed no greater nakedness by being removed (*Helv*.XVI.4)”.²⁷⁹ The clue for purple Coan dress can be found in Horace accusing Lyce of dressing too young for her age, but “neither dresses of Coan purple nor

²⁷³ Forbes, 1930

²⁷⁴ Peigler, 1993

²⁷⁵ Schoff, 1912 There is no mention of Chinese silk in the Graeco-Roman world before the time of Augustus.

²⁷⁶ “Cois tibi paene videre estut nudam”

²⁷⁷ “femineum lucet sic per bombycina corpus”

²⁷⁸ “Quid iuvat ornato procedure, vita, capillo, Et tenues Coa veste movere sinus?”

²⁷⁹ “Numquam tibi placuit vestis, quae nihil amplius nudaret, cum poneretur”

can expensive jewellery bring back the days (*Odes*.IV.13)".²⁸⁰ Tibullus described that this seductive dress was adorned with golden stripes: "the gossamer robe which some woman of Cos has woven and laid it out in golden tracks (*L*.II.3, 53-54)".²⁸¹

Richter suggested that the 'Amorigine chitons' from the 5th century Greek literature could have been made of a sort of wild silk similar to *Coa vestis*; in the *Lysistrata* of Aristophanes (411 BCE), women are told to be seductive by being "naked in their Amorgian chitons", which is "diaphanous garments".²⁸² In *Cratinus*, there is a note on spinning *βρύτινον* Amorgos, which is explained to be *βομβύκινον* in the Lexikon of Hesychius.²⁸³ Richter argues that *Amorgis* was named after its origin of manufacture, Amorgos, which was next to Cos, and thus these two could have been a similar kind of clothing made of indigenous silk.²⁸⁴

However, the identities of these silks cannot be ascertained due to the lack of physical evidence: besides the petrified cocoon of Akrotiri, there have been no findings leading to either *Coa vestis* or *Armorgis*.²⁸⁵ The *Edict of Diocletian* suggests the commercial production of wild silk textiles in Cos have declined after the first century CE due to the limited output and the incompetence to *serica*.²⁸⁶ Based on the description of Aristotle and Pliny, scholars have suggested that the Coan silk is a type of wild silk from two moths: the *Saturnia pyri* and *Pachypasa otus*, which are indigenous to the central and eastern Mediterranean.²⁸⁷ A comparative study of these two silks would be helpful (see

²⁸⁰ "Nec Coae referent iam tibi purpurae, nec cari lapides tempora..."

²⁸¹ "illa gerat vestes tenues, quas femina Coa texuit, auratas disposuitque vias"

²⁸² "κὰν τοῖς χιτωνίοισι τοῖς ἀμοργίνοις γυμναὶ παρίοιμεν (I.150)" and "τὰ διαφανῆ χιτῶνια (I.45)".

²⁸³ "ἀμοργιν ἔνδον βρυτίνην νήθειν τινά." (*Fragmenta*, 103)

²⁸⁴ Richter, 1929

²⁸⁵ Panagiotakopulu, *et al.*, 1997

²⁸⁶ Wild, 1970; Babu, 2012

²⁸⁷ Richter, 1929; Leggett, *op.cit*; Good, 2002; Leggett, 1949; Richter, 1929; Olson, 2008.

II.1.2.6).



Coa Vestis. (*Mus. Borbon. viii. 5.*)

Figure 17. Coa Vestis (Mus.Borbon. VIII.5.)²⁸⁸

²⁸⁸ Smith, *et al.*, 1890

II.2.4.2.3 Sea silk: *Pinna* the Roman “water sheep”

Sea silk, also called mussel silk or byssus silk, is obtained from the largest endemic Mediterranean bivalve, *Pinna nobilis* Linnaeus, 1758 (*Pinna squamosa* Gmelin, 1791). It is a shallow water mussel, which lives along the Italian and Dalmatian shore of the Adriatic Sea. It grows up to one meter in height stands upright in the *Posidonia* beds, strong fibrous tufts attach the body to a solid surface to withstand the flow.²⁸⁹ The tufts, zoologically called byssus, are a bundle of filaments length up to 20 cm. These fibres are cut, washed several times, dried, combed and spun into yarns.²⁹⁰ A recent study of its chemical composition revealed that it lacks collagen, but its exact protein composition is yet unknown.²⁹¹

Sea silk has been made into textiles since the Greek and Roman periods.²⁹² As it was difficult to procure good quality raw material, sea silk production was very limited. It was only made into small articles such as stockings and gloves of high prices.²⁹³ A Persian geographer Istakhri said a robe made of *şūf al-baḥr* (sea wool) costs more than 1000 gold pieces (*Il Mediterraneo* V. 1).²⁹⁴ They are noted for its glossy golden yellow to rich cinnamon hue, softness, lightness,²⁹⁵ and excellent protection against cold and damp weather.²⁹⁶

Nonetheless, sea silk is not a well-known material, for both its physical rarity and historical ambiguity. Although it was known from the antiquity, confusion was created by

²⁸⁹ Šiletić, 2004

²⁹⁰ Maeder, 2008; Maeder, 2015

²⁹¹ Diana, *et al.*, 2017

²⁹² Centoducati, *et al.*, 2007

²⁹³ Yates, 1843; Gilroy, 1853;

²⁹⁴ Abu Ishaq Ibrahim ibn Muhammad al-Farisi al Istakhri, also known as Estakhri. استخرى. The record was written ca. 951, however the book was compiled ca. 1193.

²⁹⁵ Beck, 1882 said “a thin pair of stockings may be put into a regular size snuffbox.”

²⁹⁶ Riedesel, 1773; Craven, 1821

mistaking the term byssus for sea silk. The ancient term *byssus* in Latin, *byssos* in Greek, was primarily used for fine textiles, but not for sea silk; the scientific term byssus was given to the fibre beard of mussels only after the 15th century (Table 4).²⁹⁷ According to Pliny, *byssus* was a fine tissue of the ancients, produced near Elis in Achaia (*Nat. hist.* XIX, 4). Although its identity has yet to be clarified, it is now generally accepted that *byssus* of antiquity and early medieval times refers to the linen or cotton cloth. Meanwhile, sea silk was called *Pinna* (πιννα). The pedigree of *Pinna* goes back to Aristotle. In *Historia Animalia*, *Pinna* is described as a mussel that grows upright from its root anchoring in sandy and muddy places (*His. Anim.* 547 b.15; 588b.10, 15). Pliny in *Naturalis Historia* (IX, 42), and Aelian (ca. 175-235) in *De Natura Animalium* (III.29) also describe the *Pinna* in the same manner.²⁹⁸ Also, scholars have argued that a word πινικόν in the *Periplus Maris Erythraei* (ca. 80-89) indicates the textile obtained from the *Pinna*.²⁹⁹

Table 4. Different terms for sea silk used in the Western literatures from the 15th century.³⁰⁰

English	Italian	French	German
Sea silk	Bisso marino	Soie marine	Muschelseide
Marine byssus	Bisso	Soie de mer	Byssus
		Soie de pinne	
		Soie de byssus	Byssusseide
	Seta di mare		Seeseide
Pinna silk	Seta marina	Laine de mer	Fischeseide
		Laine marine	Steckmuschelseide
Pinna wool	Lana marina	Laine de pinne	
Marine wool	Lanapinna		Meeresseide
Sea wool	Lanapene	Bysse	Pinnamarina-seide
Fish wool	Lanapsece	Byssus de pinne marine	
	Lana di nacchera		Seewolle
Silkworm of the sea		Poil de nacre	Fischwolle
	Pelo d'astura		Meerwolle

²⁹⁷ Bock, 1895; Laufer, 1915; Mckinley, 1998; Maeder, 2008

²⁹⁸ Laufer, op.cit.

²⁹⁹ Yates, 1843; Gilroy, 1853; Schrader, 1883; Nagy, 1935; Maeder, 2008

³⁰⁰ Maeder, 2015

In ancient records, the product of *Pinna* was often referred to as ‘wool’, rather than called silk. Alciphron first recorded this ‘sea wool’ in his *Galenaes to Cyrton* letter (c. 2nd century CE) as:

ὅτε ἐκ Μουνυχίας ἐπέμψαμεν αὐτῷ κομιοῦντα τὸ ψώνιον Ἑρμῶνα τουτονὶ τὸν μειρακίσκον, σπόγγους ἡμῖν ἐπέταττε καὶ τὰ ἐκ τῆς ἀλὸς ἔρια, ἃ φύεται ἐπικικῶς ἐναβρυνομένης εἰς λήδιον.

“...he ordered us to bring some sponges and sea wool, which grows in fairly large quantities in the pool of Eurynome”.³⁰¹

The term sea wool is either the index of a belief that existed at that time in a marine sheep that produced the wool of *Pinna*, or directly responsible for the formation of such a notion. The idea of marine fleece and the comparison of it with sheep’s wool were constantly awake in the minds of the Greek and Latin writers.³⁰² Procopius recorded of the Chlamys made from *Pinna* wool called *πιννοῦς* was bestowed to Armenina by Roman Emperor (ca. 350).³⁰³ Tertullian (ca. 155-220) wrote about wearing pallium instead of a toga:

Nec fuit satis tunicam pangere et serere, ni etiam piscari vestitum contigisset: nam et de mari vellera, quo mucosae lanusitatis plautiores conchae comant

“...not just to comb and sow, but it was also necessary to fish for one's dress; for fleeces are obtained from the sea, where enormous shells are furnished with tufts of mossy hair.”.³⁰⁴

Pinna silk was also highly appreciated for its colour: Basileus of Caesarea (Basilios Megas, ca.330-379) praised the golden colour of the *Pinna* wool, which none of the dyers could imitate (*Hexaem.* vii).³⁰⁵ Some have suggested that perhaps this is related to the

³⁰¹ Alciphron, 1896 (E. I. 2.)

³⁰² Laufer, 1915

³⁰³ *De Aedificiis* L.III. c.1

³⁰⁴ *De Pallio* III, 6. After Yates, 1843

³⁰⁵ “Πόθεν τὸ χρυσοῦν ἔριον αἱ πίνναι τρέφουσιν, ὅπερ οὐδεὶς τῶν ἀνθοβάφων ἐμμήσατο”

Golden Fleece that Jason and his Argonauts sought in Greek mythology.³⁰⁶ Barber and McKinley have pointed out that this theory is dubious and with geohistorical errors.³⁰⁷ On the other hand, Jenkins and Silver suggested that the Golden Fleece (*χρυσείον κῶας*) signifies the trading fleece dyed with murex-purple for gold.³⁰⁸ To know more about the interesting theories on this subject, including the Colchis from India, see Lordkipanidze (2001).³⁰⁹

The traces of existence and production of sea silk are also found in some early Chinese written records on foreign culture and trade. The first known reference in Chinese literature is the *Western Regions* chapter of the *Book of Later Han*. In the description of the products of Daqin (Rome):³¹⁰

又有細布，或言水羊毳，野蠶繭所作也。

“They also have a fine cloth which some people say is made from the down of water sheep, but in fact, it is made from the wild silk cocoons.”

Bretschneider suggested that the ‘down of the water sheep’ was perhaps the *Byssus*, referring to the Chinese account: but he made a mistake to confuse byssus with *Pinna*.³¹¹ The story of the Roman ‘water-sheep’ is also found in *Weilüe* (239-265):³¹²

有織成細布，言用水羊毳，名曰海西布。³¹³

³⁰⁶ Cole, 2005; Maeder, 2015

³⁰⁷ Barber, 1991; Mckinley, 1998

³⁰⁸ Jenkins, 1985; Silver, 1992

³⁰⁹ Lordkipanidze, 2001

³¹⁰ 後漢書·列傳第七十八 西域傳 <大秦>. The book was compiled later, but the original sources were collected by the Gan Ying, during his journey to Parthia. He was sent to collect information on the Roman Empire in 97- 99.

³¹¹ Bretschneider, 1871; Hirth, 1885

³¹² 魏略(literally means A Brief history of Wei), written by Yu Huan (魚豢).

³¹³ 卷二十二 西戎, <大秦國>

“They have fine brocaded cloth that is said to be made from the down of water-sheep.
The name is *Haisipu*.”

Haisipu literally means a fabric from the west of the sea. The same name and description appears again in the *New Book of Tang*,³¹⁴ on *Folin* (拂菻, Byzantine):³¹⁵

織水羊毛爲布, 曰海西布

“They weave fabrics from the hair of water sheep, it is called *Haisipu*”

Although it does not directly address water sheep, the *Records of Strange Things* from the early sixth century also gives interesting information about a sea silk:³¹⁶

南海出蛟綃紗, 泉先潜织, 一名龙纱。其价百餘金, 以為服, 入水不濡。

“In the South Sea, there are jiao that makes raw silk. They make threads under the shallow water, which called the *longsha* (dragon silk). It is worth as hundred golds, and when made into garments, it does not get wet.”

Jiao, or *Jiaolong* in the dictionary is described as a polysemous aquatic dragon from Chinese mythology; Schafer equated this *jiao* silk with sea silk.³¹⁷ In summary, it seems that the Chinese knew about the water sheep of Daqin and its textiles by the end of the first century, and the same information appears in the later records without being much revised. Although the writer of the *Book of Later Han* corrected the myth of water sheep to wild silk, the distinction between the water sheep and the Roman wild silk was clearly made in

³¹⁴ 新唐書(*Xīn Tángshū*), compiled in 1044-1060. 卷 221 列傳 第 146 上·下 西域, 拂菻

³¹⁵ Byzantine once was called “*hē Polis* (ή Πόλις)”

³¹⁶ 通異記(*Shuyiji*), 卷上

³¹⁷ Schafer, 1967

other Chinese records on Daqin.

It is hard to assume since when the *Pinna* silk was used for weaving before the time of the authors cited above. Yates pointed out that the descriptions of its textile uses only appears from the 2nd century onwards in contrast to the abundant references to *Pinna* in general, including Aristotle's.³¹⁸ According to a legend, the skill of *Pinna* silk making was brought to Sant'Antioco by Berenice (23-ca.81) in the late first century.³¹⁹ Although there is no evidence that Berenice really spent her time there,³²⁰ at least it implies that the skills were brought from somewhere outside Sardinia. In modern references, Taranto is often regarded as the origin of sea silk weaving, but only because it was the centre of the business in the eighteenth century.³²¹ On the assumption that the water sheep of Roman empires in the Chinese records really are the *Pinna*, it is possible to infer that the weaving of *Pinna* silk was set up earlier than the first century, sometime between the time of Aristotle and the time of Gan Ying's travel to Parthia – as the industry was mature enough to be on the lips of Persian merchants.³²²

Unfortunately, there are only a few examples of early sea silk articles left to prove the theories, mainly because the *Pinna* silk is readily attacked by moths.³²³ The oldest example of sea silk was the fragments dated to the 4th century excavated in Budapest in 1912. The remains of a brownish fabric were found between the legs of a female skeleton from a grave of Roman Aquincum. It was analysed and identified as sea silk by

³¹⁸ Yates, 1843; Mckinley, 1998

³¹⁹ Paradiso, 2015

³²⁰ According to Seneca, Berenice returned to Palestine after the breakup with Titus. Racine's play does not say about her destination.

³²¹ Yates, 1843

³²² Gan Ying was sent to inspect on Rome, however only reached the shores of the Persian Gulf. But he managed to gather much information on Daqin that was new to the Chinese – presumably from seamen and other travellers he met in Parthia. The travel started in 97 and ended in 99.

³²³ Turner and Rosewater, 1958

Hollendonner, based on the diameter of the fibre and the cross-sectional characteristic that resembles a compressed ellipse.³²⁴ However, as the artefact and the original analysis were lost during the World War II, so there is no means to verify his conjecture.³²⁵ The oldest surviving object is a knitted cap excavated from St. Denis in 1978, dating from the late middle ages. In the 18th century, *Pinna* was easily found in the shores of Sicily, the Bay of Naples, the Gulf of Taranto, and particularly beyond the Cape of Posilipo,³²⁶ and the manufacture is chiefly carried on at Taranto.³²⁷ In the 19th century, the production of byssus silk was an important business for Sicilians, with a considerable manufactory established at Palermo; and the fabrics made are described as extremely elegant and vie with the finest silk.³²⁸ However, at present, the population of *P. nobiils* have greatly reduced as a result of reckless fishing and mishandling: its production is now largely confined to Mar Grande of Taranto in Italy (Ionian Sea).³²⁹

³²⁴ Hollendonner, 1917

³²⁵ Maeder, 2008 Maeder, 2015

³²⁶ Poli and Delle Chiaie, 1792

³²⁷ Riedesel, 1773; Craven, 1821

³²⁸ Simmonds, 1883

³²⁹ Centoducati, *et al.*, 2007

II.3 Physicochemical properties

II.3.1 Mechanical and Thermal properties

Silk is one of the strongest natural fibres: it is known to have higher tensile strength than glass fibre or synthetic fibres.³³⁰ Silk's elasticity is moderate to poor and would remain stretched if elongated due to the already highly extended β -conformation.³³¹ The mechanical properties of silk fibres are a direct result of the size and orientation of the crystalline domains, and their connectivity and transitions to less organised domains.³³²

Silk is hygroscopic and has good moisture regain: 11% is the accepted moisture regain coefficient for silk, and the mercantile weight of silk is derived based on this factor.³³³ Silk is also very heat stable: it is virtually unaffected by temperature up to 140°C, and stable up to about 200°C. The temperature of transiting random-coil to β -structure conformation is 212°C, and it will start to cleave from 280°C.³³⁴ Thermal analysis of *B. mori* silkworm cocoon silk examined by thermal gravimetric analysis (TGA) indicates a glass transition temperature (Tg) of 175°C and stability at ~250°C.³³⁵ Differential scanning calorimetry (DSC) and temperature modulated DSC(TMDSC) analysis showed that *A. mylitta*, *A. assama* and *S. ricini* silks have higher thermal stabilities and better water exchange mobility than *B. mori* silk.³³⁶ However, a differential thermal analysis (DTA) result shows that *B. mori* silk has higher endothermic temperature than wild silks.³³⁷

³³⁰ Babu, 2012

³³¹ Nelson, *et al.*, 2008

³³² Matsumoto, *et al.*, 2006

³³³ Lee, 1999

³³⁴ Yang, *et al.*, 2011

³³⁵ Magoshi, *et al.*, 1985. Spider dragline silk (*N. clavipes*) is thermally stable at about 230°C.

³³⁶ Mazzi, *et al.*

³³⁷ Tanaka, *et al.*

II.3.2 Microstructure

Silks are hierarchically arranged from the amino acid level up to micro and macroscopic structures.³³⁸ Silk fibre is comprised of two major proteins called fibroin and sericin, which account for about 75 wt% and 25 wt%, respectively.³³⁹ There are also minor amounts (less than 1.5 wt%) of various impurities such as fats, waxes, colourants, and mineral salts that are bound to the sericin gum.³⁴⁰ The ratio of contents is variable not only by the species but also by cocoon strains. In cross section, raw silk fibre would show two fibroin filaments (brin) encapsulated by sericin shell (bave). Fibroin is a highly hydrophobic protein placed in the core, naturally extruded from two silkworm glands as a pair. Sericin is a protective casing that binds two strands of fibroins together (Figure 18).

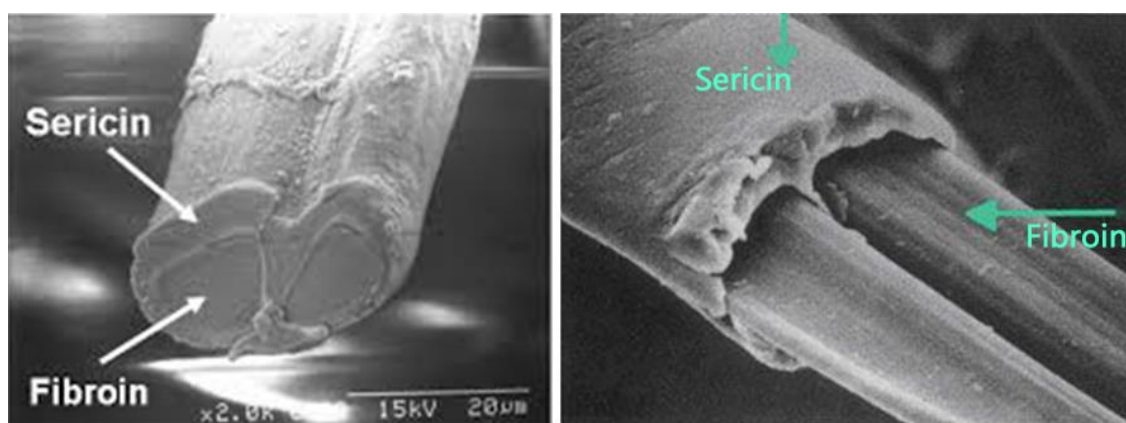


Figure 18. SEM images of raw *B. mori* silk cross-section (left) and longitudinal section (right)

The morphology of silk fibres shows many variations depending upon the silkworm species and among individual cocoons of the same species. In general, the cross section of cultivated silk fibres (*B. mori* and Eri) is often irregular, ranging from the triangular shape

³³⁸ Hakimi, *et al.*, 2006

³³⁹ Panda, 2010; Freddi, *et al.*, 2003

³⁴⁰ Sashina, *et al.*, 2006

to circular shape.³⁴¹ Wild varieties, in general, exhibit an elongated wedge shape that is large in diameter compared to *B. mori*. The irregularity of shape occurs even within the same filament, depending upon the level of the cocoon layer.³⁴² It has been observed that Tasar, Muga, and Eri silks show the presence of voids in cross section and thus are low in density, whereas *B. mori* have a uniform, compact structure.³⁴³ The degree of porousness is also variable upon the species and the location of the fibre in its cocoon layers.³⁴⁴ Another morphological property that distinguishes cultivated and wild silk is the presence of calcium oxalate crystals on the outer surfaces of the filament.³⁴⁵ Wild silkworms are reared in the open environment and conceivably need much greater protection from hazard.³⁴⁶ The salt crystals are bound to sericin, thus can be removed during the degumming process (Figure 20).

Silk fibroin belongs to fibrillar proteins. When a single fibroin is enlarged from its inner structure, it would show a bundle of fibrils, which in turn, consist of helically packed microfibrils (Figure 19). Microfibrils may play an important role in imparting enhanced strength to silks: the control of structural hierarchy is initiated from the control of chain chemistry, which provides a layout of organisation required to tune mechanical capacities.³⁴⁷

³⁴¹ Babu, 2012

³⁴² Chen, *et al.*, 2012

³⁴³ Sen and Babu, 2004

³⁴⁴ Narumi, 2004.; Prasong, *et al.*, 2011

³⁴⁵ Chen, *et al.*, 2012

³⁴⁶ Zhang, *et al.*, 2013

³⁴⁷ Sashina, *et al.*, 2006; Matsumoto, *et al.*, 2006

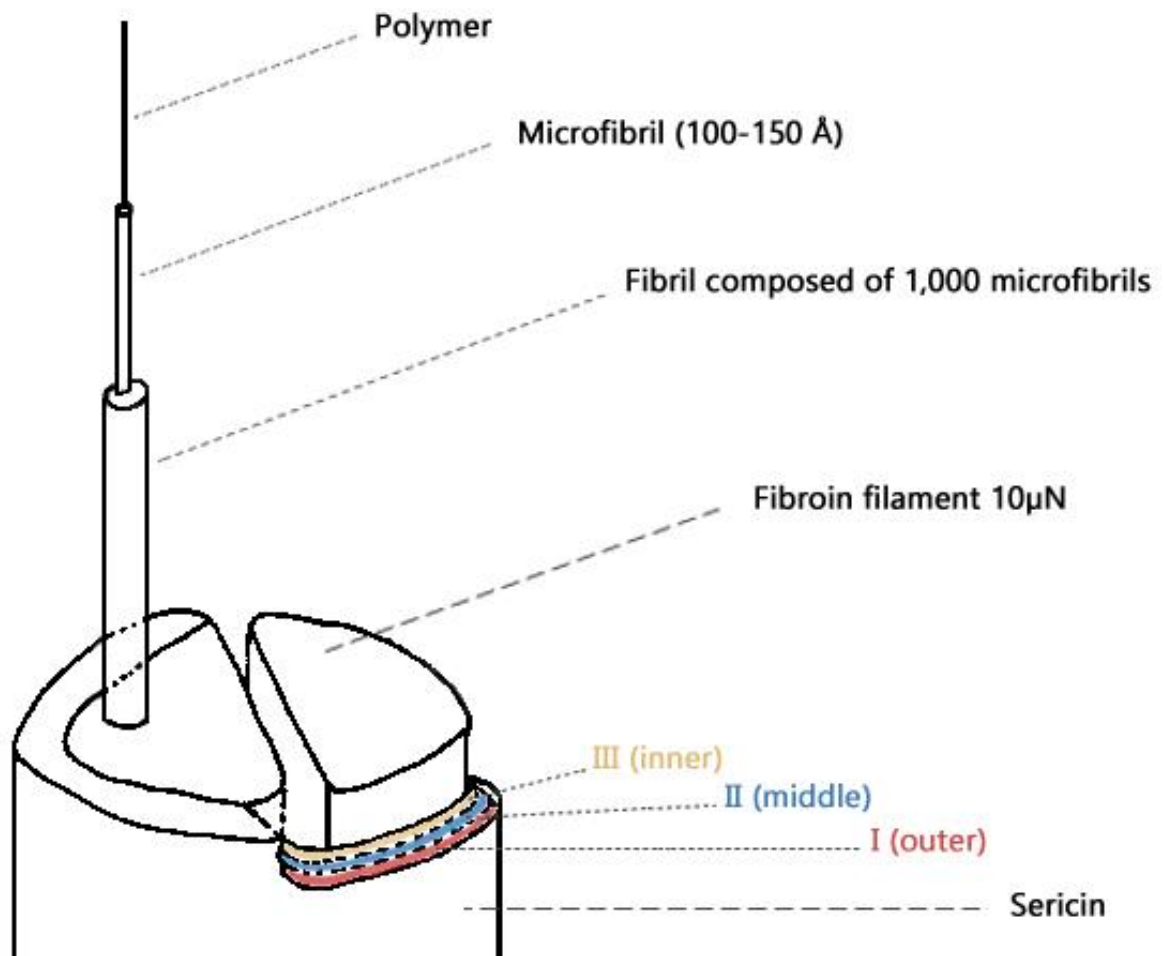


Figure 19. Structural hierarchy of a silk fibre

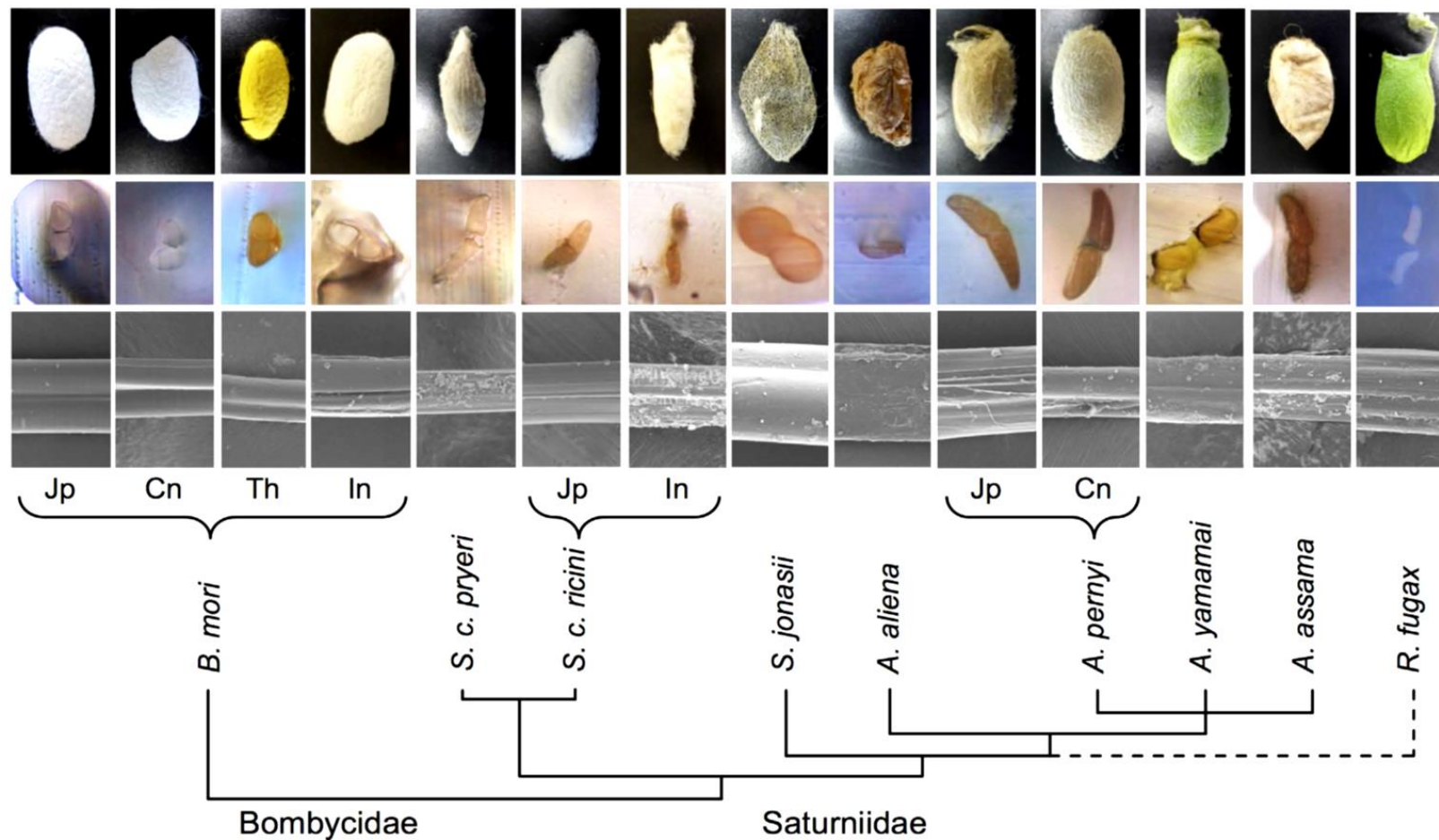


Figure 20. Morphologies of the different silk species: cocoon, cross section and longitudinal section images. ³⁴⁸

II.3.3 Crystal structures

Silk fibroin is a semi-crystalline biopolymer with highly organised nanocrystals surrounded by an amorphous matrix.³⁴⁹ The degree of crystallinity is 62–65% in *B. mori* silk, 50–63% in wild cocoon silks, and 30–50% in spider silks. The secondary structure of fibroin is in a β -sheet conformation with macro fibrils consisting of helically packed nano-fibrils (Figure 21).³⁵⁰ Most silkworm cocoon and spider dragline silk fibres have anti-parallel β -pleated sheet crystalline structures, in which the polymer chain axis is parallel to the fibre axis.³⁵¹ Other silks are known to have helical (bees, wasps, ants) or cross- β -sheet structures.

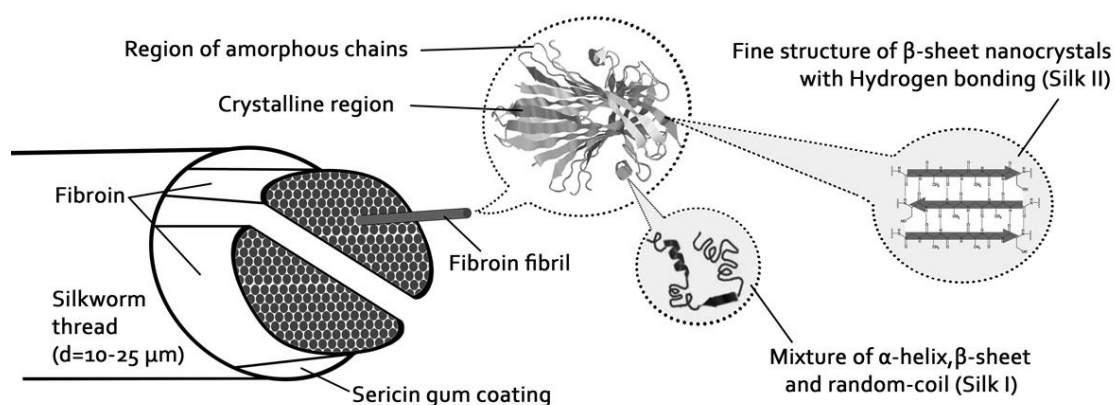


Figure 21. Schematic crystal structure of cocoon silk fibroin

The cocoon silk has two different structures, Silk II and I. Silk I is the initial form of silk in the insect's silk gland, which is liquid and water-soluble state. The crystalline regions of Silk I have α -helical structures. Once the fibre is spun out and secreted, the unstable structure converts to β -sheet of Silk II. The loss of water makes the Silk II

³⁴⁹ Fibres are consisting of high molecular polymers of crystalline and amorphous, mixed together but with orientation.

³⁵⁰ Putthanarat, *et al.*, 2000

³⁵¹ Marsh, *et al.*, 1955 ; Lucas, *et al.*, 1960 ; Lotz and Cesari, 1979 ; Kaplan and Mcgrath, 2012

insoluble, and the decrease in the distance (18.3%) between overlying sheets makes Silk II energetically stable.³⁵²

In *B. mori* silk, the β -sheets is comprised of the glycine-alanine crystalline repeats that are asymmetric; with one surface primarily projecting alanyl-methyl groups, and the other surface containing hydrogen atoms from glycine residues.³⁵³ The inter-sheet interactions are stabilised by Van der Waals force between the segments of the peptide; dipole and hydrogen bonds between carbonyl and amine groups of an adjacent peptide; and disulphide cross-links (Figure 22).³⁵⁴ The crystalline regions consist of sections of polypeptide chains of glycine, alanine and serine. These sections have simple branching; thus, the chains are closely and compactly arranged. In contrast, sections containing residues of tyrosine, proline, diamine and dicarboxylic acids have bulky residues and less oriented in packing; this forms the amorphous regions of the fibre. The groups found in the amorphous region are more accessible to the action of chemical reagents.³⁵⁵

The structural studies of wild cocoon silks are not well established as that of *B. mori*; however, it is known that they are considerably different from the domesticated silk. For example, the stress-strain curve *B. mori* silk fibroin exhibits a monotonic rapid increase in fibre, while *A. pernyi*, *A. yamamai* and *S. ricini* silk fibroin show a slow increase with the flat region.³⁵⁶ In addition, the crystallisation of the liquid silk from wild silkworms readily occurs due to heat, whereas the crystallization of *B. mori* liquid silk results from shear stress and dehydration.³⁵⁷ The phase transition from α -helix conformation to β -form is an exothermic process, resulting in peaks in the temperature range of 0 –120°C. While

³⁵² Kaplan, *et al.*, 1992

³⁵³ Matsumoto, *et al.*, 2006

³⁵⁴ Sashina, *et al.*, 2006 ; Kaplan and Mcgrath, 2012

³⁵⁵ Babu, 2012

³⁵⁶ Shimura and Hidaka, 1980 ; Iizuka, 1980 ; Vollrath and Knight, 2001

³⁵⁷ Tanaka, *et al.*, 2002

wild silks (*A. mylitta*, *A. assamensis*, and *S. ricini*) showed clear exothermic peaks corresponding to a phase transition from α -helix to β -sheet, *B. mori* showed weak endothermic peaks corresponding to the breakage of hydrogen bonds in the α -form or to the untangling of the physical network.³⁵⁸

Studies using NMR have shown that the structural transition from α -helix to β -sheet in wild silk fibroins occur in the polyalanine region and glycine-rich region.³⁵⁹ The silk fibroins of *A. pernyi* and *S. c. ricini* mainly consist of the repeated similar sequences by about 80 times where there are alternative appearances of the polyalanine region and the glycine-rich region as in spider (major ampullate) silk.³⁶⁰ It is predicted that the structural transition from silk I to II of *A. pernyi* silk fibroin induces alanine and serine residues to take the β -sheet form, leaving a small amount of α -helix and random coil form; while most of the tyrosine residue remains as a random coil.³⁶¹

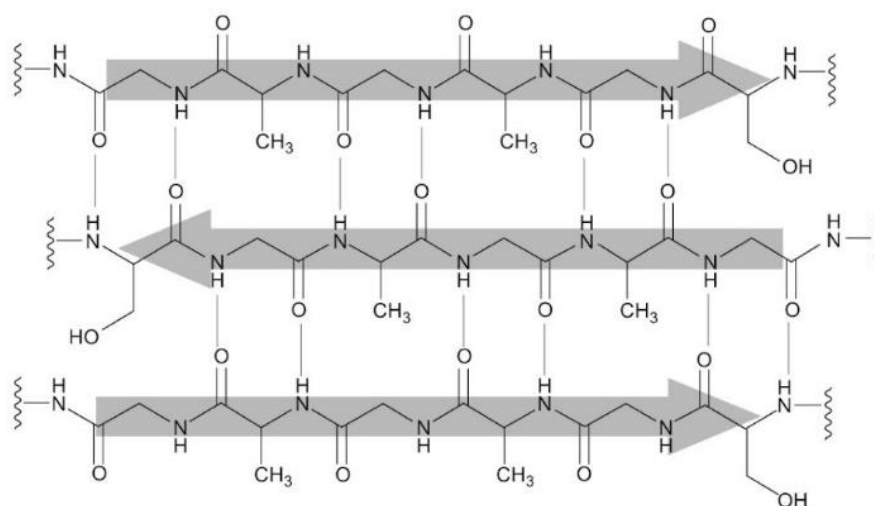


Figure 22. Mechanisms of silk molecular orientation and formation of hydrogen bonding³⁶²

³⁵⁸ Ibid.

³⁵⁹ Asakura and Murakami, 1985; Asakura, *et al.*, 1988 ; Nakazawa, *et al.*, 1999 ; Asakura, *et al.*, 1999, Asakura, *et al.*, 1988, Asakura, *et al.*, 2017, Asakura, *et al.*, 2002, Asakura, *et al.*, 1988

³⁶⁰ Yukuhiro, *et al.*, 1997; Simmons, *et al.*, 1996; Nakazawa and Asakura, 2002

³⁶¹ Nakazawa and Asakura, 2002

³⁶² Zhang, 2016

II.3.4 Protein structures

Proteins consist of polypeptide chains that are amino acids joined peptide bonds (Figure 23). The sequence of amino acids that make up a protein or polypeptide chain is called the primary structure. Secondary structure refers to the conformational arrangement occurring from the interactions of the polypeptide backbone, such as α -helix and β -sheet.³⁶³ Tertiary structure refers to a three-dimensional structure of a protein, which pertains to the assembly of the secondary structure. The three-dimensional structure is joined by the interactions between the side chains along the polypeptide chain.³⁶⁴ Most proteins contain the primary, the secondary and the tertiary structures, however, only the proteins composed of two or more polypeptide chains have the quaternary structures.

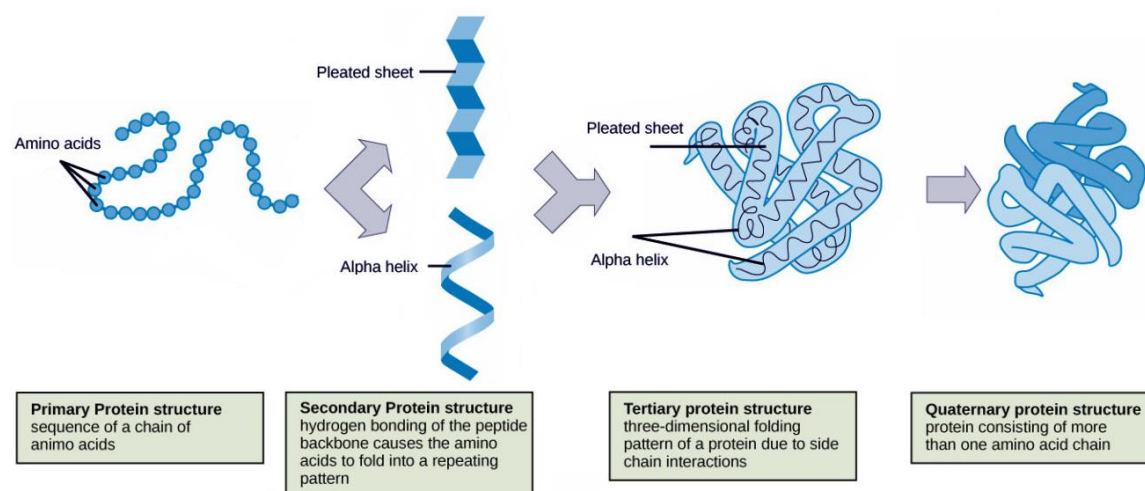


Figure 23. Structure of proteins

The most common types of proteins are globular, fibrous, membrane and disordered. Of the two proteins comprise silk, fibroin is a fibrous protein and sericin is a

³⁶³ Anfinsen, 1972; Iupac, 1997

³⁶⁴ Campbell, 2008

globular protein. Fibrous proteins—also known as Scleroprotein—such as keratin, collagen and fibroin only contain the primary and the secondary structures, while globular proteins have the tertiary structure that results from turns or bends between regions of the secondary structure.³⁶⁵ Fibrous proteins form long filament shaped like wires. However, globular proteins fold into glob-like structures. While globular protein is often water-soluble, fibrous proteins are typically inert and water insoluble. Furthermore, fibrous proteins do not denature as easily and contain many repeats of secondary structures. The structures often feature cross-links between chains such as disulfide bonds.³⁶⁶

Naturally, the primary structures of the fibroin and the sericin corresponds to the sequence of amino acids show many differences (Table 5). Sericin that wraps up the surface is a complex mixture of 5-6 polypeptides differing widely in size (40- 400 kDa),³⁶⁷ which is characterised by an unusually high content of serine ranges from 31-40 % (76-285 kDa) and other hydrophilic amino acids.³⁶⁸ As sericin consists of 76% of hydrophilic chains, it can be easily dissolved in water.³⁶⁹ However, the primary structure of fibroin, the core, holds the hydrophobic protein structure of the natural block copolymer.³⁷⁰ Fibroin is very rich in glycine and alanine, which allow a close packing of β -sheets and an interlocking arrangement of R groups. This property makes the fibres strong and resistant to breakage.³⁷¹ The amino acid composition of the silk proteins is variable depending upon the silk species.

³⁶⁵ Benjamin Cummings-Pearson Education, 2008

³⁶⁶ Campbell, 2008; Benjamin Cummings-Pearson Education, 2008

³⁶⁷ Freddi, *et al.*, 2003

³⁶⁸ Schoeser, *et al.*, 2007; Garel, *et al.*, 1997

³⁶⁹ Bae. 2002. PhD, Keimyung University.

³⁷⁰ Kim and Park, 2016

³⁷¹ Nelson, *et al.*, 2008

Table 5. Amino acid compositions of *B. mori* proteins (excluding signal peptides)

	Sericin 22-1186 residue MW 117275.95 Theoretical pl: 5.79		Heavy chain 22-5263 residue MW 389203.41 Theoretical pl: 4.32		Light chain 17-262 residue MW 25949.76 Theoretical pl: 4.91		P25 17-220 residue, MW23625.8 Theoretical pl: 6.46	
AA	# residues	Mol(%)	# residues	Mol(%)	# residues	Mol(%)	# residues	Mol(%)
Ala (A)	81	7.0	1590	30.3	34	13.8	10	4.9
Arg (R)	55	4.7	13	0.3	10	4.1	12	5.9
Asn (N)	64	5.5	19	0.4	18	7.3	13	6.4
Asp (D)	88	7.66	25	0.5	17	6.9	14	6.9
Cys (C)	3	0.3	3	0.1	3	1.2	8	3.9
Gln (Q)	29	2.5	9	0.2	15	6.1	8	3.9
Glu (E)	24	2.1	30	0.6	5	2.0	7	3.4
Gly (G)	128	11.0	2415	45.9	22	8.9	9	4.4
His (H)	17	1.5	5	0.1	5	2.0	8	3.9
Ile (I)	13	1.1	12	0.2	20	8.1	14	6.9
Leu (L)	18	1.5	5	0.1	17	6.9	19	9.3
Lys (K)	44	3.8	11	0.2	4	1.6	7	3.4
Met (M)	-	-	3	0.1	1	0.4	1	0.5
Phe (F)	6	0.5	28	0.6	7	2.8	15	7.4
Pro (P)	8	0.7	14	0.3	8	3.3	12	5.9
Ser (S)	383	32.9	635	12.1	24	9.8	13	6.4
Thr (T)	103	8.8	45	0.9	7	2.8	12	5.9
Trp (W)	10	0.9	11	0.2	2	0.8	3	1.5
Tyr (Y)	51	4.4	275	5.3	10	4.1	10	4.9
Val (V)	40	3.4	94	1.8	17	6.9	9	4.4
Instability index		55.79		-20.90		35.73		37.36
GRAVY		-1.068		0.213		-0.054		-0.276

II.3.4.1 *Bombyx silk fibroin*

B. mori fibroin consists of three elementary polypeptides: a heavy chain, a light chain, and P25. The heavy chain and the light chain are synthesised by two different mRNAs; a disulfide bond at the C-terminus forming heterodimers links the two. A group of six heterodimers is associated with a glycoprotein called P25 by hydrophobic interaction, and this forms 2.3 MDa tertiary elementary structural units of fibroin.³⁷² A single molecule of P25 is internally located and plays an essential role in maintaining the integrity of the complex. The molecular mass of silk fibroin of heavy chain, light chain, and P25 is a 6:6:1 molar ratio, in both Silk I and II.³⁷³

The heavy chain is a fibrous protein composed of 5362 amino acids, of molecular weight of about 350~390 kDa.³⁷⁴ The heavy chain consists of 12 clusters of crystalline domains varying in length and sequence that are connected by 11 amorphous domains of regular composition (Figure 24).³⁷⁵ The crystalline regions of the heavy chain are mainly formed with the repetitive arrays of (GA)_nGX (where X= S, Y, or V); glycine, alanine, and serine take about 86% of the primary structure. The light chain and P25 are smaller and have more standard amino acid compositions than the heavy chain. The primary structure of P25 was identified to have 220 amino acids and about 25 kDa in molecular weight.³⁷⁶ The light chain has 244 amino acids and molecular weight about 25-30kDa. As in the heavy chain, alanine, glycine and serine are the most abundant amino acids, but only up to 30% of the total composition. Furthermore, P25 and the light chain do have repetitive

³⁷² Shimura, 1983; Inoue, *et al.*, 2000

³⁷³ Tanaka, *et al.*, 1999, Inoue, *et al.*, 2000

³⁷⁴ Zhou, *et al.*, 2000

³⁷⁵ Ibid.; Asakura, *et al.*, 2002

³⁷⁶ Chevillard, *et al.*, 1986

sequences.³⁷⁷ The GRAVY value³⁷⁸ showed that the heavy chain is very hydrophobic, while the light chain and P25 are hydrophilic. The instability index³⁷⁹ indicates that the fibroin proteins are all stable, especially the heavy chain.

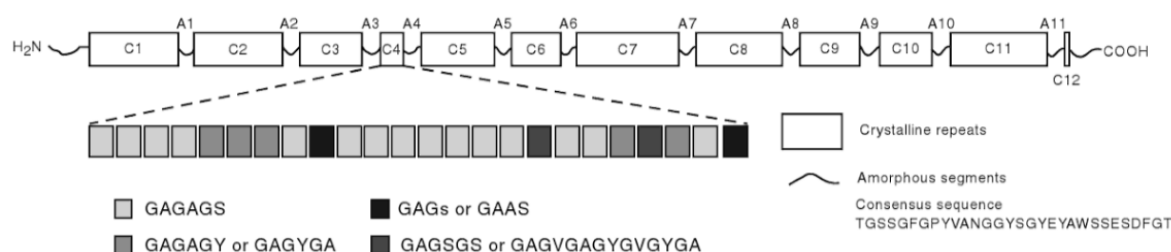


Figure 24. Schematic primary structure of *B. mori* heavy chain: the amino acid sequence for one of the crystalline domains is given to highlight the repetitive nature of the protein³⁸⁰

B. mandarina fibroin is also comprised of a heavy chain, a light chain, and P25. While the deduced sequences of the light chain and P25 are almost identical to those of *B. mori*,³⁸¹ the heavy chain sequence differs by 3.4% in the first 178 amino acids.³⁸² Unfortunately, the rest of the heavy chain sequence is not available on the database. See section IV.1.1 for further information on the difference in protein sequence between *B. mori* and *B. mandarina*.

³⁷⁷ Murphy and Romero, 2014

³⁷⁸ Grand average of hydropathicity. The increasing positive score indicates a greater hydrophobicity: the most hydrophobic amino acids are isoleucine (4.5) and valine (4.2). The most hydrophilic ones are arginine (-4.5) and lysine (-3.9). Kyte and Doolittle, 1982

³⁷⁹ The instability index provides an estimate of the stability of the protein in a test tube. A value smaller than 40 is predicted as stable, and a value above 40 predicts that the protein may be unstable.

³⁸⁰ Asakura, *et al.*, 2002; Murphy and Romero, 2014

³⁸¹ Tanaka and Mizuno, 2001

³⁸² Kusuda, *et al.*, 1986

II.3.4.2 Wild silk fibroins

In the *Saturniidae* family silk moths, fibroin is composed of a heavy chain only, with repetitive polyalanine stretches that make homodimers.³⁸³ It is reported that the presence of a light chain and P25 as in *Bombyx* silks were not detected in *A. yamamai*.³⁸⁴ The molecular weight of *A. pernyi* silk fibroin, extracted from the posterior gland, is known to be 400–450 kDa (non-reduced) and 220 kDa (reduced).³⁸⁵ *A. mylitta* fibroin, also from the gland, was calculated to be 395 kDa (non-reduced) and 197 kDa (reduced).³⁸⁶ These results suggest that *A. pernyi* and *A. mylitta* fibroins are probably a homodimer composed of two similar sized polypeptides linked by a disulfide bond. The estimated molecular weights of *A. yamamai*, *A. pernyi* and *S. cynthia ricini*, calculated from the purified RNA, are 250kDa, 220 kDa and 230 kDa, respectively.³⁸⁷ However, the exact molecular weights remain undecided because the separation of these proteins has not been studied well.

Saturniidae fibroins are predominantly of alanine, glycine and serine in their amino acid composition, however with more alanine than glycine unlike in *B. mori* silk (Gly 45.9%, Ala 30.3%); the level of serine is slightly lower than in *B. mori* (Table 6). They are also abundant in Aspartic acid and Arginine that form a characteristic Arg–Gly–Asp (RGD) sequence.³⁸⁸ The polyalanine blocks contain 12-13 contiguous alanine residues, and the rest is rich in glycine, featuring various combinations of GX and GGX (X=A, S, T, D, L or R).³⁸⁹ *S. ricini* fibroin is more glycine-rich than alanine, similar to *Bombyx* silk. However, it is reported that the primary structure of *S. c. ricini* silk fibroin consists of 100-fold

³⁸³ Lucas and Rudall, 1968; Yukuhiro, *et al.*, 1997; Gupta K, *et al.*, 2015

³⁸⁴ Tanaka and Mizuno, 2001

³⁸⁵ Tamura and Kubota, 1988

³⁸⁶ Datta, *et al.*, 2001

³⁸⁷ Tamura and Sakate, 1988

³⁸⁸ Zhao, *et al.*, 2011

³⁸⁹ Malay, *et al.*, 2016

repeated sequences of alternating polyalanine regions and glycine-rich regions as in spider silk but with more extended polyalanine region than in the spider silk.³⁹⁰ A recent study reported that *S. c. ricini* fibroin has the highest level of repetitive sequences among the *Saturniids*, at approximately 44.5% of the total length of each tandem repeat; being 43% for *A. yamamai*, 42.2% for *A. pernyi*, 39.6% for *A. assamensis*, 37% for *A. selene*, and so on.³⁹¹ The GRAVY and the instability index values confirm that these fibroins are all stable and hydrophobic except *A. mylitta*, which has an incomplete sequence.

Table 6. Amino acid compositions of *Antheraea* and *Samia* silk fibroins (exclude signal peptides)

	A. pernyi #19-2639 MW 214091.01 Theo. pl: 4.52		A. mylitta #19-507 MW 43433.37 Theo. pl: 4.98		A. yamamai #19-2856 MW 232619.64 Theo. pl: 4.48		A. assamensis #19-2809 MW 227498.35 Theo. pl: 4.50		S. ricini #19-2880 MW 225349.42 Theo. pl: 4.82	
Amino Acid	# residues	Mol (%)	# residues	Mol (%)	# residues	Mol (%)	# residues	Mol (%)	# residues	Mol (%)
Ala (A)	1133	43.2	171	35.0	1221	43.0	1190	42.6	1305	45.6
Arg (R)	70	2.7	22	4.5	80	2.8	73	2.6	47	1.6
Asn (N)	5	0.2	3	0.6	6	0.2	6	0.2	4	0.1
Asp (D)	118	4.5	28	5.7	141	5.0	127	4.6	93	3.2
Cys (C)	3	0.1	-	-	3	0.1	3	0.1	3	0.1
Gln (Q)	4	0.2	7	1.4	4	0.1	1	0.0	4	0.1
Glu (E)	17	0.6	13	2.7	18	0.6	27	1.0	17	0.6
Gly (G)	720	27.5	111	22.7	778	27.4	811	29.1	914	31.9
His (H)	25	1.0	11	2.2	32	1.1	32	1.1	39	1.4
Ile (I)	9	0.3	5	1.0	8	0.3	6	0.2	10	0.3
Leu (L)	8	0.3	6	1.2	8	0.3	8	0.3	8	0.3
Lys (K)	3	0.1	3	0.6	1	0.0	4	0.1	8	0.3
Met (M)	-	-	3	0.6	-	-	-	-	-	-
Phe (F)	5	0.2	2	0.4	3	0.1	7	0.3	2	0.1
Pro (P)	5	0.2	4	0.8	6	0.2	7	0.3	7	0.2
Ser (S)	297	11.3	51	10.4	315	11.1	287	10.3	194	6.8
Thr (T)	8	0.3	9	1.8	9	0.3	5	0.2	12	0.4
Trp (W)	34	1.3	4	0.8	44	1.6	46	1.6	16	0.6
Tyr (Y)	138	5.3	23	4.7	141	5.0	137	4.9	167	5.8
Val (V)	19	0.7	13	2.7	20	0.7	14	0.5	12	0.4
Instability index		19.69		51.35		20.15		16.51		16.01
GRAVY		0.210		-0.075		0.178		0.175		0.327

³⁹⁰ Nakazawa and Asakura, 2003

³⁹¹ Malay, *et al.*, 2016

III. Materials, methods and instrumentation

III.1 Experimental timeline

The timeline of experimental works were as presented in Figure 25.

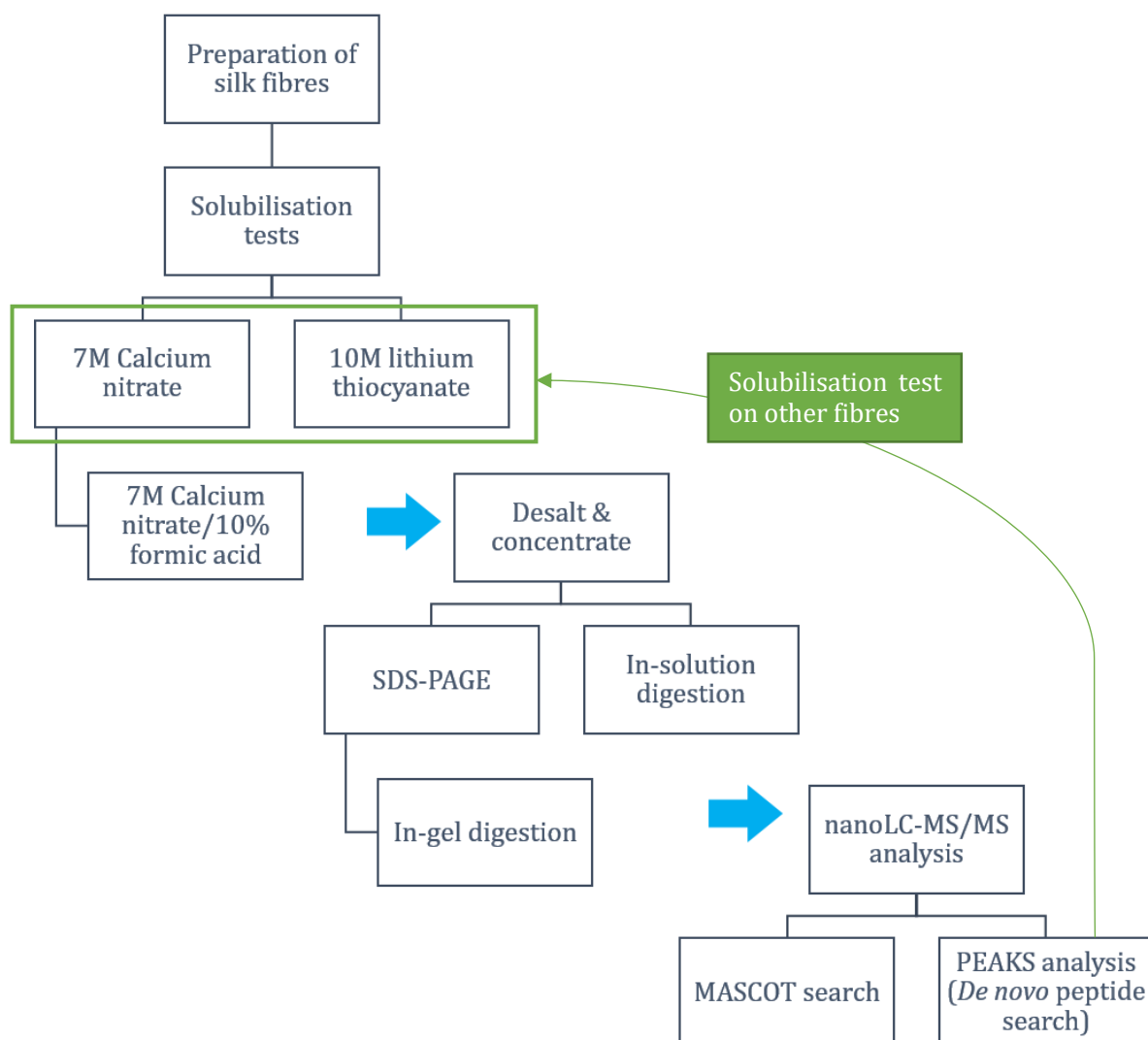


Figure 25. Flowchart of experimental works

III.2 Preparation of silks

III.2.1 Silk samples

Seven different types of silks were selected for the experiments. Besides the domesticated silkworm *B. mori* and its precursor *B. mandarina*, five other Asian varieties of wild silks that have long history of clothing application and still are commercially available were selected.³⁹² The cocoons were collected with the help of Dr. Fritz Vollarth (Oxford Zoology Department) and the late Dr. Irene Good (Harvard University). Their common names, scientific names and the place of origin are provided in Table 7 .

Table 7. List of the silk samples selected for experiments

No.	Type	Origin	Common Name	Other names
1	<i>B.mori</i>	China	家蠶 Mulberry silk	Domesticated/ mulberry silk
2	<i>B.mandarina</i>	China	野蠶 Wild silk	<i>Bombyx mori mandarina</i> <i>Theophila mandarina</i>
3	<i>A. pernyi</i>	China	柞蠶 Chinese Tasar	Temperate Tasar
4	<i>A. mylitta</i>	India	Indian Tasar	Tropical Tasar
5	<i>A. yamamai</i>	Japan	天蠶 Japanese Oak silk	<i>Tensan</i> Japanese Tasar
6	<i>A.assamensis</i>	India	Muga silk	<i>Antheraea assamensis</i> <i>Saturnia assama</i>
7	<i>S.ricini</i>	Thailand	Eri silk	<i>Attacus ricini</i> <i>Philosamia cynthia ricini</i>

³⁹² Based on Capinera, 2008; Panda, 2010

III.2.2 Degumming

Preparation of the silk fibres from the cocoons consisted of cooking and degumming. Cooking is a pre-treatment to soften the gum material and dissolve the wax by simple boiling. Although this step does some degree of removal of sericin, it is more to free the fibres from the cocoon form, and the degumming process is necessary for the effective removal of sericin and to leave the fibroin alone. There are few different methods of degumming: traditional method with Marseilles soap, synthetic detergents methods, enzymatic degumming and acid degumming. For *B. mori* and *mandarina*, the traditional Marseilles soap method is enough. However, for other silks, the degumming process is more complicated due to the presence of metallic salts and high gum contents. The enzymatic degumming with soap became the choice of method, as it is mild and uniformly removes the surface materials.³⁹³

III.2.2.1 Marseilles soap method for *Bombyx* silks

The cut cocoon pieces were cooked in milli-Q water at 95°C for 30 minutes at 1:200 of liquor ratio.³⁹⁴ The freed fibres were collected from the cooking bath and treated in the degumming bath at 95°C for an hour, at liquor ratio 1:50. The bath was prepared with 0.5 % Marseilles soap and 0.3 % sodium bicarbonate solution (percentage on the weight of the fabric). Marseilles soap is made of olive oil containing 60-63% fatty anhydrides, 7% combined alkali, 26-30% water and up to 2.3 % free sodium hydroxide.³⁹⁵

³⁹³ Choudhury, 2006

³⁹⁴ The ratio of the weight of liquor used to the weight of material being treated (mass: liquor).

³⁹⁵ Choudhury, 2006

III.2.2.2 Marseilles soap and enzyme method for other silks

The samples were first boiled in a 95°C bath formulated with 1g/L of sodium carbonate and 1g/L of sodium hydrosulphite. The treatment took place for 30 minutes at liquor ratio 1:40. The boiled (cooked) fibres were treated in degumming bath prepared with 1g/L Alcalase[®] 2.4 (*Bacillus licheniformis*), 5g/L Sodium bicarbonate and 1g/L Marseille soap in milli-Q water.³⁹⁶ The fibres were treated for an hour at 55-60°C with constant agitation, at liquor ratio 1:50. The treated fibres were rinsed with 0.2 wt% soap solution.

All samples were rinsed thoroughly with 2 wt% sodium hydrosulphite solution to remove the excess soap and alkali, and then thoroughly rinsed with lukewarm Milli-Q water five times. The collected fibres were dried at a room temperature, avoiding direct sunlight to prevent the yellowing. After dried for at least 48 hours, silk fibres were weighted into 1 mg and distributed in Eppendorf tubes (1.5ml).

³⁹⁶ Replaceable with non-ionic surfactant, such as Trion X-100 or Tween 80.

III.3 Preparation of fibroin solution

III.3.1 Solubilisation

The degree of solubilisation of different silk fibres to the solvents was tested using various conditions. All reagents were purchased from Sigma. Silk samples were weighted to 1 mg and distributed in class vials with lids. A heating block with digital thermometer was operated inside of a fume hood cupboard. Resulted solutions were stored at -20°C. Milli-Q water was used in all process. Each set of experiments were repeated 3 times. The protein assay by UV-VIS (280nm) of the yielded solutions was tried, however it was impossible due to the high salt contents; thus, the solubilisation in this section refers to the visual disappearance of the fibres without leaving the trace of solid matter. The recovery rate of fibroin in solution was assumed to 85%, considering the possibility of incomplete solubilisation.

III.3.1.1 Acid

Hydrochloric acid

Boiling of silk fibres in 6N HCl is a well-adopted method in amino acid analysis.³⁹⁷ 1mg of fibre samples were distributed in glass tubes and immersed in 6N HCl at liquor ratio 1:1; and heated to 80°C, 105°C and 120°C.

Formic acid

As a similar approach to the HCl method, samples were first immersed in formic

³⁹⁷ Vanden Berghe and Wouters, 2004

acid (98%) at liquor ratio 1:1 and heated at 80°C, 105°C and 120°C.

III.3.1.2 Calcium Chloride ternary solution

This method uses a ternary solution of CaCl_2 : H_2O : $\text{C}_2\text{H}_5\text{OH}$ in molar ratio 1:8:2, contrived by Ajisawa.³⁹⁸ Based on the previous studies, high temperature would decrease the amount of time for dissolution; thus, the samples were treated at 105°C.³⁹⁹

III.3.1.3 Chaotropic salt solution

Lithium Bromide and Lithium Thiocyanate

It is known that 9-10M lithium bromide and thiocyanate solution can dissolve *B. mori* silk fibroin.⁴⁰⁰ Samples were immersed in 1ml solvent and heated at 60°C, 80°C and 105°C and left up to 48 hours. 10M lithium thiocyanate solution was tested at 125°C, to compare with 7M calcium nitrate solution.

Calcium Nitrate ($\text{Ca}(\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$)

7M calcium nitrate solution was prepared in 70°C water bath. As reported in Lee *et al.*, it was predicted that the temperature lower than 100°C would not affect the non-mulberry silks.⁴⁰¹ *Bombyx mori* silk dissolved in this solvent even at the room temperature during the preparation; thus, Bombyx silks were tested at 60°C, 80°C, 105°C and 125°C; and other wild silks were tested at 105°C and 125°C.

³⁹⁸ Ajisawa, 1998

³⁹⁹ Matsumoto, *et al.*, 2008, Chen, *et al.*, 2001, Solazzo, *et al.*, 2012, Li, *et al.*, 2015, Abdel-Fattah, *et al.*, 2015

⁴⁰⁰ Chen, *et al.*, 2001, Rockwood, *et al.*, 2011, Abdel-Fattah, *et al.*, 2015, Srisuwan and Srihanam, 2009 and Good, 2011

⁴⁰¹ Lee, *et al.*, 2013

III.3.2 Desalting

III.3.2.1 Dilution and centrifugation

Fibroin solutions were diluted with water to a double volume, to reduce the viscosity and to check the stability of solution once the salt concentration decreased. If necessary, pH of the solution was adjusted to neutral using Tris-HCl buffer (7.5 and 8.3). Then solutions were centrifuged for 10 minutes at 14000 rpm; supernatants and pellets were separated and stored at -20°C if necessary. The desalting of fibroin solutions was implemented using the following systems:

- Slide-A-Lyzer™ dialysis cassette/tubes MWCO 3.5kDa
- Amicon Ultra centrifugal filter units (1.5, 15, 50 ml) MWCO 3 kDa

III.3.2.2 Ultrafiltration

Fibroin solutions were mixed with one volume of buffer, and spinned at 4500g for 30 minutes. Calcium nitrate-fibroin solutions were diluted to three volumes and then mixed with buffer, due to the viscosity. Lithium thiocyanate-fibroin solutions were directly mixed with buffer without dilution. For a gradual buffer exchange, urea buffer was formulated with 8M urea in 50mM of Ammonium bicarbonate solution and the concentration of urea was reduced to 4M, 2M and 1M; a same volume fresh buffer was updated on each turn. On the 5th run, a same volume of 100mM ammonium bicarbonate solution and spinned for 30 minutes. The filtrate was further concentrated in vacuum centrifuge if needed, and the final volume was adjusted to 10µg/µl concentration in 100mM ammonium bicarbonate solution. The recovery rate from the ultrafiltration was assumed to 80% (i.e. 1mg of fibre to 0.85mg of fibroin in solution, and 0.64mg recovered after desalting).

III.4 SDS-PAGE

III.4.1 Materials

Precast SDS-PAGE gels

- Mini-PROTEAN® TGX™, 10% (1.0mm, 10 well, sample load 30µl)
- Biorad Criterion XT Tris-Acetate Gels, 3-8% (1.0mm, 12+2 well, sample load 45µl with two 15µl reference wells)

Sample buffer

- Laemmli x2 (Sigma S3401-1VL)
- Tricine x2 contains 200 mM Tris-HCl, pH 6.8, 2% SDS, 40% glycerol, 0.04% Coomassie Brilliant Blue G-250, and 2% β-mercaptoethanol (added fresh)

Running buffer

- 25 mM Tris, 1.92M glycine, 1% SDS (pH 8.3)
- 1 M Tris, 1 M Tricine, 1% SDS (pH 8.3)

Protein Standards

- Precision Plus Protein™ Dual Colour Standards (10~250kDa) (Biorad)
- HiMark™ Pre-Stained Protein Standard (31~ 460kDa) (Sigma)

Protein detection (Stains)

- Coomassie Biosafe (Biorad)
- ProteoSilver™ Silver Stain Kit (Sigma)

III.4.2 Method

10-20 μ l of fibroin solution was mixed with a same volume of running buffer and heated for 5 minutes at 95°C for denaturation. Samples were briefly centrifuged, and the supernatants were loaded to each well. Protein standard marker was loaded at the very end of the gel plate. Mini size gel was run for 45 minutes at 200V, and midi size gel was run for 65 minutes at 150V. Once removed from the cast, gels were washed with milliQ water and stained for an hour. The staining solution was removed, and the gels were detained overnight. The detailed protocols are as below.

Staining process for each type of stains were followed by the product manual.

III.5 Enzymatic digestion

III.5.1 Materials

AmBic: Ammonium bicarbonate (Sigma)

TCEP: (Sigma)

DTT: Dithiothreitol (Sigma)

IAA: Iodoacetamide (Sigma)

FA: Formic acid (HPLC grade, Sigma)

TFA: Trifluoroacetic acid (HPLC grade, Sigma)

ACN: Acetonitrile (HPLC grade, Sigma)

Potassium ferrocyanide (Sigma)

Sodium thiosulfate (Sigma)

Trypsin and Chymotrypsin (Sequencing grade, Promega)

Ultrapure water (milliQ)

Buffer A: 2% ACN, 0.1% FA in H₂O

Buffer B: 80% ACN, 0.1% FA in H₂O

III.5.2 Methods

III.5.2.1 In-gel digestion

Washing

1. Excise the protein bands and chop into smaller pieces (1-2 mm³).
2. In eppendorf tubes, wash the gel pieces in H₂O (10 min, shaking)

Distaining from Coomassie blue stain

1. Add 100µl of 100mM AmBic and shake for 10 min. Remove and discard the buffer.
2. Add 100µl of 50% ACN in 100mM AmBic and shake for 10 min. Centrifuge shortly and remove the wash solution. Repeat this step 3-4 times if necessary.
3. Add 100% ACN, enough to cover the surface and let sit for 5 min. Gel pieces should appear shrunken and opaque. Remove the liquid and let gels air dry.

Distaining from Silver stain

1. Add 200µl of 1:1 mixture of 30mM potassium ferrocyanide and 100mM sodium thiosulfate solutions and shake 20 min at RT. Gel pieces should turn yellow. Remove the liquid.
2. Add 200µl of H₂O and shake 20 min at RT, remove the liquid. Repeat 1-2 if necessary, until the gels look transparent.
3. Add 200mM AmBic enough to cover the gel and shake 20 min at RT, remove the liquid.
4. Add 100% ACN, enough to cover the surface and let sit for 10 min. Remove liquid and let gels air dry.

Reduction and Alkylation

1. Rehydrate the gel pieces in 30µl of 10mM DTT in 100mM AmBic, reduce for 30 minutes at 56°C. Remove excess liquid.
2. Add 30µl 55mM IAA in 100mM AmBic, alkylate for 30 minutes at RT. Remove the liquid.
3. Add 100µl of 100mM AmBic and shake for 5 minutes. Remove the solution.
4. Add 100µl of 100% ACN and let sit for 10 min., remove the liquid and let the gels dry.

Digestion

1. Prepare the digestion buffer. Prepare the trypsin reagent, containing 20ng/µl of trypsin or 25ng/µl chymotrypsin in 50mM AmBic. * The buffer must be kept on ice all the time.
2. Rehydrate the gel pieces 30µl (add enough to cover the surface) of digestion buffer and incubate for 15-30 min at 4°C.
3. If the initial volume has been absorbed by the gel, add more digestion buffer.
4. Centrifuge the gel pieces shortly and remove the excess digestion buffer.
5. Add sufficient amount of 50mM AmBic to cover the surface and incubate overnight at 37 °C.

* Add 1ml of ice-cold 50mM AmBic to 20µg of Trypsin to make x10 stock (20µg/ml). Always keep on ice. Right before the use, dilute trypsin stock 1:10 in 50mM AmBic to final concentration 20ng/µl. Prepare the Chymotrypsin buffer in the same manner.

Peptide extraction (optimised for the extraction of big peptides)

1. Stop the enzymatic reaction by adding FA in 50mM AmBic to 5% of the final concentration. Shake 10 min and collect the supernatant to a new tube.
2. Add 100µl of 100% ACN, shake 15 min at RT and collect the supernatant.
3. Add 100µl of 50% ACN in H₂O, shake 5 min and collect the supernatant.
4. Add 100µl of 1% FA in ACN, shake for 5min and collect the supernatant.

5. Dry the sample down in a vacuum centrifuge.
6. Resuspend in 20µl elution buffer (0.05% TFA, 2% ACN in H₂O).

Peptide extraction (standard)

1. Stop the enzymatic reaction by adding FA in 50mM AmBic to 1% of the final concentration. Shake 10 min and collect the supernatant to a new tube.
2. Add 50µl of Extraction buffer I (50% ACN, 5% FA in H₂O). Shake 10 min (and optional 5 min sonication) and collect the supernatant.
3. Add 50µl of Extraction buffer II (85% ACN, 5% FA in H₂O). Shake 10min and collect the supernatant.

III.5.2.2 In-solution digestion

Volumes are approximate, based on the quantity of protein [1mg/ml].

Reduction and alkylation

1. Prepare reducing agent: 200 mM TCEP or DTT, in 25mM AmBic
2. Take 10µl of protein sample, add 8 µl 25mM AmBic and 2 µl of reducing agent (=20mM)
3. Reduce the sample at 56°C for 30 min. Cool down to RT and spin down the samples.
4. Prepare alkylating agent: 200mM IAA solution in 25mM AmBic
5. Add 5µl alkylating agent and incubate at RT in the dark (=40mM).

Digestion

1. Add 200µl of ice-cold 50mM AmBic to 20µg of Trypsin (final concentration 100ng/µl)
2. Perform the first digestion with trypsin or chymotrypsin at 1:20 to 1:100 ratio:
 - a. Ratio of trypsin: sample = 1:20, thus 0.5µg (500ng)
 - b. Add 5µl to the sample and incubate overnight at 37°C
2. Perform a second in-solution trypsin digestion in 80% ACN environment:
 - a. Ratio of trypsin: sample = 1:100, thus 0.1µg (100ng)
 - b. Add 1µl of trypsin solution to the sample.
 - c. Add 100% ACN (LC-MS grade) up to 4/5 of the final sample volume (ex. If the sample volume is 31 µl, add 124µl). Incubate at 37°C for 3 hours.
3. Spin down the sample and stop the digestion by adding FA to 5% of the final volume (add 8µl of FA).
4. Dry the sample completely using vacuum centrifuge

III.5.2.3 Peptide cleaning (ZipTip)

Buffer A

98% Milli-Q water
2% ACN
0.1% TFA*

Buffer B

80% ACN
20% Milli-Q water
0.1% FA*

* HPLC-grade reagents and MilliQ-H₂O were used

1. Acidify sample (20-100µl) by adding FA to 0.1-1% final concentration.
2. Wet ziptip: aspirate Buffer B and dispense the solution. Repeat once.
3. Equilibrate: aspirate Buffer A and discard the solution. Repeat once.
4. Bind: aspirate and dispense the sample, repeat 10 times.
5. Wash with Buffer A and discard the waste x 4 times.
6. Elute with Buffer B to a new tube and dry using vacuum centrifuge.
7. Resuspend the sample for LC-MS/MS analysis using 0.05% TFA, 2% ACN in H₂O

III.6 Peptide analysis

III.6.1 Sample submission

10µl of each sample, enzymatically digested and zip-tip (C18) cleaned, were loaded to a separate Waters total recovery glass vial (12 x 32mm).

III.6.2 LC-MS/MS data acquisition

III.6.2.1 OXION lab (Department of Physiology, anatomy & genetics)

Material was subjected to nano-ultra performance liquid chromatography tandem mass spectrometry analysis (nano-UPLC-MS/MS). An Agilent 1100-capillary LC system (Palo Alto, CA) was used to inject 8 µL of sample onto a 5-µm Phenomenex Jupiter™ capillary LC column (150 × 0.5 mm). The fragment peptides were resolved with a linear gradient of 5–95% solvent B for 50 min, followed by a second gradient of 95–5% for 40 min at a flow rate of 15 µL/min. Mobile phase solvent A was 0.1% aqueous formic acid, and solvent B was 0.1% aqueous formic acid in 90% acetonitrile. The LC column eluate was subsequently analysed from 0 to 35 min by positive ion electrospray MS using a Waters Q-TOF micromass spectrometer (Milford, MA). The following conditions were used: capillary voltage, 3000 V; cone voltage, 35 V; extraction cone voltage, 4 V; desolvation and nebulising nitrogen gas flow, 300 L/h; desolvation temperature, 300 °C; source temperature, 80 °C; MCP detector voltage, 2400 V; TOF flight tube voltage, 5630 V.

For MS/MS analyses, the fragment peptides (prepared as above) were resolved by

nano-ultra performance LC tandem MS 60 using a 75- μ m inner diameter, 25-cm length C¹⁸ nano-Acquity™ UPLC™ column (1.7 μ m particle size; Waters) and a 90-min gradient of 2–45% solvent B (solvent A, 99.9% H₂O and 0.1% formic acid; solvent B, 99.9% acetonitrile and 0.1% formic acid) on a Waters nano-Acquity UPLC system (final flow rate, 250 nL/min; 6000–7000 psi) coupled to a Q-TOF Premier tandem mass spectrometer (Waters). Data acquisition was performed in data-directed analysis mode (MS to MS/MS switching at precursor ion counts greater than 10 and MS/MS collision energy dependent on precursor ion mass and charge state). The mass accuracy of the raw data was calibrated using [Glu1] fibrinopeptide peptide (200 fmol/ μ L in 50% acetonitrile and 0.1% formic acid; 700 nL/min flow rate), which was infused into the mass spectrometer during sample analysis. MS and MS/MS data were calibrated at intervals of 30s.

III.6.2.2 Chemistry Research Laboratory (Department of Chemistry)

Orbitrap Elite

The peptides were analysed on a NanoAcquity-UPLC system (Waters) connected to an Orbitrap Elite mass spectrometer (Thermo Fischer Scientific) possessing an EASY-Spray nano-electrospray ion source (Thermo Fischer Scientific). The peptides were trapped on an in-house packed guard column (75 μ m i.d. x 20 mm, Acclaim PepMap C18, 3 μ m, 100 Å) using solvent A (0.1% Formic Acid in water) at a pressure of 140 bar. The peptides were separated on an EASY-spray Acclaim PepMap® analytical column (75 μ m i.d. x 50 mm, RSLC C18, 3 μ m, 100 Å) using a linear gradient (length: 100 minutes, 3 % to 60 % solvent B (0.1% formic acid in acetonitrile), flow rate: 300 nL/min). The separated peptides were electro-sprayed directly into the mass spectrometer operating in a data-dependent

mode using a CID based method. Full scan MS spectra (scan range 350-1500 m/z, resolution 120000, AGC target 1e6, maximum injection time 250 ms) and subsequent CID MS/MS spectra (AGC target 5e4, maximum injection time 100 ms) of 10 most intense peaks were acquired in the Ion Trap. CID fragmentation was performed at 35 % of normalized collision energy and the signal intensity threshold was kept at 500 counts. The CID method used performs beam-type CID fragmentation of the peptides.

III.7 Data processing tools and analytical software parameters

III.7.1 Sequence data

Protein sequences of each sample silks were brought from UniProtKB,⁴⁰² and consolidated to a FASTA (text-based format of representing peptide sequences using single-letter codes).

III.7.2 Database query and De novo analysis

MS/MS database search

MS/MS ion searches were carried out using the MASCOT search engine⁴⁰³ on an in-house server (Computational Biology Research Group, University of Oxford). UniProtKB/Swiss-Prot⁴⁰⁴ and NCBI nr⁴⁰⁵ databases were used, allowing for carbamidomethyl oxidation as modifications.

De novo analysis

De novo analyses were performed with PEAKS 8.5 (Bioinformatics, Canada). The raw MS file was searched against the manually organised silk database. Trypsin,

⁴⁰² Universal protein resource/knowledgebase. (<https://www.uniprot.org/>)

⁴⁰³ Mascot is a software search engine that uses mass spectrometry data to identify proteins from peptide sequence databases. Mascot uses a probabilistic scoring algorithm for protein identification adapted from the MOWSE algorithm (<http://www.matrixscience.com/>).

⁴⁰⁴ SwissProt is a high quality, curated protein database; sequences are non-redundant. It is ideal for peptide mass fingerprint searches and MS/MS searches of well-characterised organisms where it isn't essential to match every single spectrum (<https://www.uniprot.org/>).

⁴⁰⁵ National Center for Biotechnology Information, non-redundant protein database. Part of the U.S. National Library of Medicine database (<https://www.ncbi.nlm.nih.gov/refseq/about/nonredundantproteins/>).

Chymotrypsin, and Trypsin/Chymotrypsin with a maximum number of 2 missed cleavages from both nonspecific ends was selected as the protease.⁴⁰⁶ Carbamidomethylation (cysteine), Oxidation (methionine), Deamination (asparagine, glutamine) was set as variable modifications. Precursor mass tolerance was set as 15 ppm and fragment mass tolerances for CID were set to 0.5 Da respectively. The *de novo* peptides were filtered to have ALC score $\geq 70\%$. All presented peptides have -10lgP score of > 20 , or to have false discovery rate (FDR) close to 1%. The protein identification result was filtered to have -10lgP score of > 30 . The spectra were manually checked and validated or disqualified.

⁴⁰⁶ Trypsin and chymotrypsin both cleave the C-terminal of the peptides. However, the nonspecific cleavages from both ends were allowed in consideration of unexpected proteolytic activities in the sample. It must be noted that the fibroins were denatured and degraded into polypeptides during the solubilisation process as seen in SDS-PAGE results in IV.4.

IV. Solubilisation of silk fibroins

IV.1 Introduction

Silk fibroins are large proteins of molecular sizes from 200 to 400 kDa, which is above the detection limit of the mass spectrometers in general. In other words, fibroins cannot be submitted to the mass spectrometers in their native form. Thus, sequence-specific proteases are utilised to cleave proteins into smaller fragments or peptides. However, the robust nature of the fibroins does not accommodate direct proteolysis: fibrous proteins do not denature easily and contain many repeats of secondary structures.⁴⁰⁷ A study examined the proteolysis of *B. mori* silk in fibre form and solution form with ten different proteases: the results showed only about 5-10% weight loss due to the hydrolysis in fibres, whereas the fibroin solution in Ajisawa's reagent showed 65-100% weight loss.⁴⁰⁸ Thus, solubilisation is imperative to be able to process silk fibroins for protein mass spectrometry.

However, natural silk fibroins are not soluble in water or in most of organic solvents. Once the fibre is spun and secreted, fibroin behaves like a thermoset polymer.⁴⁰⁹ The presence of a large amount of intra- and inter-molecular hydrogen bonds, its high crystallinity as well as specific physicochemical properties, all play a role in the prevention of the separation of the molecules, and as a result, resist their solubilisation into various chemicals.⁴¹⁰ Moreover, owing to the excellent heat stability of silk fibroin, melting is not easy either as heat cleaving will only happen from 280°C.⁴¹¹

There numerous studies that frequently report the ability to dissolve and reshape silk proteins for biomedical applications. As *B. Mori* silk fibroin is soluble in some solvents,

⁴⁰⁷ Campbell, 2008

⁴⁰⁸ Ha, *et al.*, 2006

⁴⁰⁹ Jo, *et al.*, 2011

⁴¹⁰ Furuhashi, *et al.*, 1994; Kweon and Park, 2001; Cai, *et al.*, 2002

⁴¹¹ Tsukada, *et al.*, 1992

mostly strong salt solutions, Ajisawa (1998)'s invention of a ternary solution—made up of calcium chloride, ethanol and water system in 1:2:8 in molar ratio—as a solvent for *B. mori* silk fibroin been widely adopted.⁴¹² *B. mori* silk is also soluble in aqueous solutions of chaotropic salt such as 9M lithium bromide solution,⁴¹³ and 10M lithium bromide solution.⁴¹⁴ *B. mori* silk can also be solubilised in non-aqueous solvents, such as N-methyl morpholine N-oxide (NMNO)⁴¹⁵ and MMNO with the addition of aprotic solvents such as Dimethyl sulfoxide.⁴¹⁶

However, other than *B. mori*, only a few other types of wild silk fibroins have been studied in this way. This is because wild silk fibres are more durable, and thus, less likely to be affected by the salts and organic solvents that readily solubilises the *B. mori* fibroin.⁴¹⁷ For example, *A. pernyi* silk fibroin showed a strong resistance to Ajisawa's reagent. It was eventually solubilised in a concentrated chaotropic salts, of which include calcium nitrate (tetrahydrate) solution.⁴¹⁸ This is because chaotropic salts denature the higher order structure of fibroin, weaken hydrophobic interactions, and thus induce the solubilisation in wild silk fibroin. The process is facilitated by the role of ions, which is induced by the disassociation of salts that inhibits the decrease of entropy in water when it come into contact with the fibroin.⁴¹⁹ Besides calcium nitrate, lithium thiocyanate,⁴²⁰ zinc chloride and zinc nitrate hexahydrate salts have also been shown to solubilise *A. pernyi* fibroin.⁴²¹ Recently, it has also been reported that the *A. yamamai* silk fibre shows good solubility in concentrated

⁴¹² Ajisawa, 1998 See Matsumoto, *et al.*, 2008; Chen, *et al.*, 1996; Min, *et al.*, 1997; Ha, *et al.*, 2006 and Jeong and Hur, 2010; Li, *et al.*, 2015

⁴¹³ Furuhashi, *et al.*, 1994; Sofia, *et al.*, 2001; Brown, *et al.*, 2015 and Patra, *et al.*, 2012

⁴¹⁴ Tse and Dupont, 2001; Good, 2011

⁴¹⁵ Freddi, *et al.*, 1999

⁴¹⁶ Sashina, *et al.*, 2003

⁴¹⁷ Tsukada, 1986

⁴¹⁸ Kweon and Park, 2001; Kweon, *et al.*, 2003; Ha, *et al.*, 2003; Lee, *et al.*, 2013; Jo, *et al.*, 2017

⁴¹⁹ Kweon, *et al.*, 2003

⁴²⁰ Tsukada, *et al.*, 1994; Ha, *et al.*, 2006 and Chen, *et al.*, 2001

⁴²¹ Kweon and Park, 2001

lithium thiocyanate and calcium nitrate (tetrahydrate) solution.⁴²² The *Samia ricini* silk fibre is also reported to be soluble in 85% phosphate, 9M lithium thiocyanate and 9M calcium nitrate solution at 100°C. However, the *S. ricini* silk fibre is not soluble in lithium bromide, zinc chloride, as well as the Ajisawa's reagent.⁴²³

Hence, in this chapter, a series of experiments based on the previous studies were designed to test the solubility of different types of silk fibres, including *B. mori*, *B. mandarina* (*B. mori* precursor) and some popular wild silk varieties: *A. pernyi*, *A. mylitta*, *A. assamensis* and *S. ricini*. The aim was to find a versatile reagent that is compatible with the variety of silks fibres so that it can be applied to the investigation of archaeological samples of unknown identity. The ideal reagent would effectively denature the protein (disrupt tertiary and secondary structure) in order to make them soluble in water, but at the same time not affect the primary structure of the protein. This is necessary for a successful peptide analysis, because of the ideal to prevent non-enzyme-specific cleavages in order to promote the confidence of analysis results.

⁴²² Jo, *et al.*, 2017

⁴²³ Wang, 2000; Srisuwan and Srihanam, 2009

IV.2 Solubilisation properties

IV.2.1 Acids

IV.2.1.1 *Hydrochloric Acid*

Hydrochloric acid (HCl) is a strong corrosive inorganic acid, commonly used for the hydrolysis of samples before amino acid analysis. However, it often involves a constant boiling in an airtight environment for more than 24 hours. As such, it was hypothesised that some part of fibroin might not be hydrolysed and could remain intact or in large peptide bonds at milder conditions.

Bombyx silks were easily solubilised in 3N to 6N HCl solution at all temperatures between 80°C and 105°C (the boiling point is 110 °C). However, *Antheraea* and *Samia* silks did not solubilise. Instead, they are partially hydrolysed and remained dispersed in the form of short fragments after 1 hour of treatment. The hydrolysis occurred faster at higher concentration of HCl and higher temperatures, but it did not facilitate solubilisation. This result is consistent with the hydrolysis behaviour of *A. pernyi* silks in HCl solution as reviewed in Kweon *et al.*,⁴²⁴ the product of acid hydrolysis was a precipitation of fine white powder, not a solution. However, it should be noted that optical observation cannot determine if any degree of solubilisation took place. Thus, a batch of samples was made with 6N HCl solution, boiled at 105°C for 1 hour, the resulting solutions were separated using 3-8% Tris-Acetate gel, and the tryptic digest was analysed by nanoLC-MS/MS. The LC-MS/MS data showed that the *Bombyx* silks exhibited the presence of small peptides just above the dye front. However, the wild silks showed very low or no trace of relevant fibroin

⁴²⁴ Kweon, *et al.*, 1998

peptides (see V.4.1, Table 21)

In another attempt, two of the wild silk variety— *A. assamensis* and *S. c. ricini*—were treated in a 6N HCl solution at 105°C with the addition of 7M calcium nitrate to 10% of the solvent volume. The addition of salt seemed to facilitate the solubilisation process. However, the resulting solutions did not precipitate any solid matter in acetone, which implies that most of the protein was hydrolysed to amino acids. Following that, the nanoLC-MS/MS analysis showed that there are only small numbers of fibroin peptides can be found at low concentrations, in contrast to the samples solubilised with only calcium nitrate (see V.4.1 and Table 21 for the result and discussion). These results suggest that HCl is not a suitable reagent to prepare silk solution for protein mass spectrometry.

IV.2.1.2 Formic acid

Formic acid (CH_2O_2 , pH 3.8-4.3) is the simplest form of carboxylic acid. Formic acid can effectively depolymerise polypeptides.⁴²⁵ It is reported that formic acid solubilised regenerated *B. mori* silk filaments.⁴²⁶ However, formic acid did not initiate the solubilisation of natural silk fibre of *Bombyx*, *Antheraea* and *Samia* silks. The samples were immersed in formic acid (98%) and heated at 80°C, but no significant changes were observed. However at 105°C (the boiling point of formic acid is 100.8 °C), the deformation of fibres to a soft, gel-like status was observed in all samples but at different phases: *B. mori* took 15 minutes, and *B. mandarina* took 30 minutes, while the rest of the samples took more than an hour. Nonetheless, there was still no sign of solubilisation in all samples.

⁴²⁵ Demassa and Boudreaux, 2013

⁴²⁶ Um, *et al.*, 2004

IV.2.2 Calcium chloride ternary solution (Ajisawa's reagent)

The Ajisawa's reagent worked well with *Bombyx* silks without boiling. However, it has been shown that it was not suitable for the solubilisation of wild silks at any temperature. At 105°C, *B. mori* and *B. mandarina* silks were solubilised in less than 30 minutes, resulting in a semi-transparent and viscose solution. However, no visible changes were shown observed after 4 hours; *Antheraea* and *Samia* silks became dispersed after 6 hours, but no significant changes or any sign of solubilisation appeared after 24 hours. In conclusion, Ajisawa's reagent only worked with the *Bombyx* silks.

IV.2.3 Chaotropic salt solutions

IV.2.3.1 *Lithium Bromide (LiBr)*

The lithium bromide aqueous solution is a widely adopted solvent for the solubilisation of *B. mori* silk together with Ajisawa's reagent, but the rate of molecular degradation of fibroin is much slower in the lithium bromide solution.⁴²⁷ In a 9M lithium bromide solution, at 60°C, *Bombyx* silks started to dissolve after 30 minutes. They were fully solubilised in one-hour yielding transparent and light amber liquids. However, there were no significant changes observed in *Antheraea* and *Samia* silks even after 48 hours. At 80°C and at 105°C, the time to solubilise *B. mori* and *B. mandarina* were halved. The other silk samples were dispersed after 1 hour, but none of them was solubilised after 24 hours. This solvent worked well with *Bombyx* silks without boiling but as shown, it was not suitable to solubilise *Antheraea* and *Samia* silks at any temperature. There is a recent publication reporting a consistent result, that lithium bromide cannot solubilise *A. yamamai* silk.⁴²⁸

⁴²⁷ Um and Park, 1998

⁴²⁸ Jo, *et al.*, 2017

IV.2.3.2 Lithium Thiocyanate (LiSCN)

The silk samples in 9M lithium thiocyanate solution exhibited similar behaviours as observed in the case of using the 9M lithium bromide solution in the previous section. At 60°C, the *B. mori* and the *B. mandarina* silks started to solubilise after 30 minutes and they were fully solubilised in 50 minutes, yielding a transparent liquid. This result is consistent with the previous study stating that the *B. mori* silk fibroin showed a better solubility in lithium thiocyanate solution than in lithium bromide solution.⁴²⁹ However, as per the results in the previous section, there were no significant changes observed in wild silk samples even after 48 hours. At 80°C and 105°C, *Bombyx* silks were solubilised in less than 15 minutes. However, the other silks did not show any sign of solubilisation.

Based on a previous study of solubilising *A. yamamai* silk in 10M lithium thiocyanate solution,⁴³⁰ the wild silk samples were treated in a 10M lithium thiocyanate solution. All samples were solubilised in an hour at 105°C, and in 15 minutes at 125°C. However, the *A. yamamai* and *S. ricini* samples took slightly more time than other samples. 1 ml of solvent was sufficient for a 1mg fibre. The results were transparent liquids.

⁴²⁹ Goto, *et al.*, 1990; Pawcenis, *et al.*, 2013

⁴³⁰ Jo, *et al.*, 2017

IV.2.3.3 Calcium Nitrate

It has been reported that the tetrahydrate form of calcium nitrate is needed to achieve the complete solubilisation of *B. mori* fibroin.⁴³¹ The *Bombyx* silks were easily solubilised in a 5M to 7M calcium nitrate (tetrahydrate) solution. At 5M concentration, the *B. mori* silk dissolved in 30 minutes and the *B. mandarina* silk took about 1 hour to be completely solubilised at 60°C. At 7M concentration, *B. mori* silk was partially solubilised at room temperature and completely solubilised at 60°C in 20 minutes. It took about 45 minutes for the *B. mandarina* silk. The process was more facilitated at 80°C; it took 15 minutes for the *B. mori* silk and less than 30 minutes for the *B. mandarina* silk. However, the *Antheraea* and *Samia* silks did not react to the 7M solvent at either of these temperatures. It has been reported that the *A. pernyi* silk fibre is not soluble in a 5M calcium nitrate solution and is only partially soluble at about a 20 wt % in a 6M solution and is only soluble in a 7M solution.⁴³²

Thus, the concentration of the salt was kept at 7M, and the temperature was raised: when heated at 105°C, all the samples were almost completely or completely solubilised in 1 hour. When the temperature was increased to 125°C, the time to achieve a same level of solubilisation decreased by large margin: it took less than 5 minutes to solubilise *Bombyx* silks completely. Wild silks also showed good solubility in this solvent. The *A. pernyi* silk reacted the fastest and solubilised in about 20 minutes. The *A. mylitta*, *A. yamamai*, *A. assamensis* and *S. ricini* silks took up to 30 minutes. All the samples resulted in a transparent viscose liquid. At high temperatures, the addition of methanol (1:1 volume to the solvent) helped the mobility of the fibre in the viscose salt solution. This is due to methanol at these

⁴³¹ Ha, *et al.*, 2003

⁴³² Kweon and Park, 2001

temperatures is readily boiling, and thus created a vortexing effect.

When cooled and centrifuged, it was observed that small amounts of sedimentation formed in the *Antheraea* and the *Samia* silk samples: The *S. ricini* showed the most significant amount of precipitation, followed by the *A. yamamai* silk, and then *A. mylitta* and *A. assamensis* silks, and finally *A. pernyi* silk, which showed the smallest amount of precipitation. The supernatants and sediments were subsequently separated, and 1ml of the fresh salt solution was added to the sediments and heated again. The solid matters were solubilised again, indicating that the precipitations are the part of the proteins re-aggregated from the solution. The results showed that the *Antheraea* and *Samia* silks were partially solubilised in 7M calcium nitrate solution at 1:1 ratio (sample weight to solvent volume). The optimal liquor ratio to solubilise 1mg of fibre were appeared to be 1:1.5 (Figure 26). However, once the solutions were cooled down to a room temperature, minor amounts of precipitations were observed again in the *A. yamamai* and *S. ricini* samples. This result implies that the temperature played an important role for the solubilisation process. Also, it should be noted that the variable nature of the silk fibres could also have affected this ratio.



Figure 26. Re-aggregated proteins after sitting the samples for 24 hours at 4°C. left) fibre weight to solvent ratio at 1:1 right) at 1:1.5, with formic acid (1%)

IV.2.3.4 Calcium nitrate with formic acid

As an alternative approach, it can be hypothesised that the 7M calcium nitrate solution and the formic acid can be used together. In theory, the addition of formic acid can affect the solubility of the protein by changing the pH and also by giving charge to the amino acids or peptides.⁴³³ Amino acids are ampholytes, and free amino acids exist as neutral zwitterions that contain both positively and negatively charged groups (Figure 27). Moreover, unlike HCl, formic acid does not destroy the primary structure of fibroin and induces hydrolysis.⁴³⁴ The difference in their behaviour has been attributed to the relative acidity of the absorbed acid concerning the zwitterion.⁴³⁵

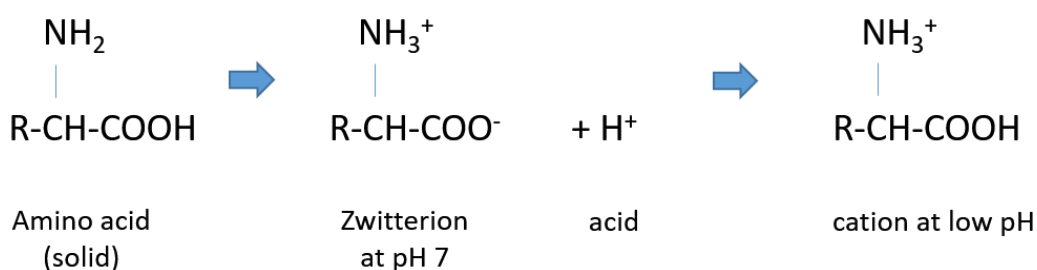


Figure 27. zwitterion in simple amino acid solutions and the acid reaction

At first, the calcium nitrate solution was added to the samples after it has been heated in formic acid. A rapid solubilisation together with the change of colour from transparent to bright yellow was observed. However, the control sample (salt solution boiled alone) did not show any change. The ratio of the salt solution was varied from 10% to 50%, and it was observed that the higher concentration of the salt solution, the faster solubilisation was achieved. However, it should be noted that most of the samples required more than 30%

⁴³³ Steiger, 1930

⁴³⁴ Zhang, *et al.*, 2012

⁴³⁵ Braga, *et al.*, 2004

salt solution to be solubilised.

In another set of experiments, the samples were first boiled 125°C in the salt solution and the formic acid was added in by varying amounts. The result showed that only 0.1% of formic acid was enough to induce the yellowing of the solution (Figure 28). Once the formic acid was added, the yellowing reaction occurred instantly followed by heavy bubbling. This heavy bubbling is thought to be from the boiling of formic acid as the boiling point of formic acid is 100.8°C, while calcium nitrate (tetrahydrate) is 132°C. The reaction that took place was much faster in sample B than in sample D. The saturation of colour was also deeper in sample B than in sample D. This development of colour was in positive relationship with the increase in formic acid content up to 10%. However, it should be noted that no further change was observed above that percentage. This result implies that the yellowing of the fibroin solution in calcium nitrate-formic acid system is the reaction between the protein contents and the formic acid.

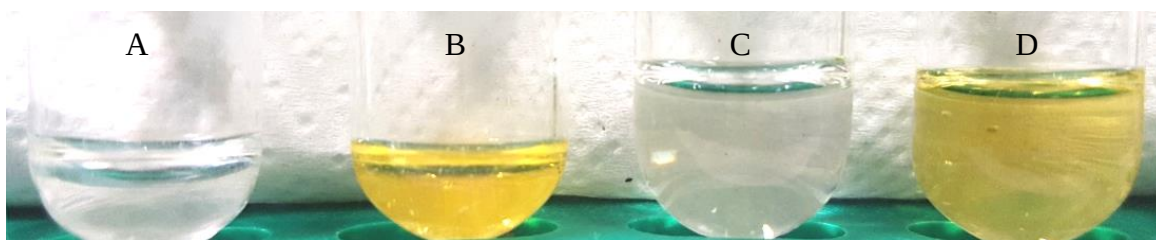


Figure 28. Yellowing of silk fibroin (*S. ricini*) solution by formic acid: 0.5mg fibre in each tubes with A) x1 7M $\text{Ca}(\text{NO}_3)_2$, B) x1 7M $\text{Ca}(\text{NO}_3)_2$ +0.1(o.w.f.)% FA, C) x2 7M $\text{Ca}(\text{NO}_3)_2$, D) x2 7M $\text{Ca}(\text{NO}_3)_2$ +0.1(o.w.f.)% FA. First, the solubilisation was implemented with the salt solution at 125°C for 10 minutes, and then FA was added off the heat.

The yellowing was related to the type of salts, but not by the concentration of salts. Four different solvents were mixed with a same volume of formic acid. Except for the slight yellowing of lithium thiocyanate-formic acid system, the Ajisawa's reagent and the calcium nitrate solvents did not show any other observable changes. In addition, the *Bombyx* silk

solutions made with the Ajisawa's reagent did not show any changes by the addition of formic acid. Thus, the most probable explanation to account for the yellowing of the calcium nitrate-fibroin solution by the addition of formic acid would be the Xanthoproteric reaction. This reaction involves the nitration of the phenyl group ($-C_6H_5$) in aromatic amino acids—tryptophan (Trp) and tyrosine (Tyr)—forming yellow nitro-substitution products, which changes to orange at alkaline pH due to the ionisation of the phenolic group.⁴³⁶ The most significant modification are 6-nitro-Trp with two λ_{max} (absorption maxima) at 326 and 380 nm,⁴³⁷ and 3-nitro-Tyr with λ_{max} at 360 nm (pH < 6).⁴³⁸ As the silk fibroins contain only minor amounts of tryptophan, the reaction is more likely to be induced the tyrosine residues (Table 8). The chromophore of tyrosine is a phenol (C_6H_5OH) bound to the $C\alpha$ carbon at the para position relative to the hydroxyl group (Figure 29).

Table 8. Tryptophan and tyrosine contents in different silk fibroins

	Tryptophan (W)		Tyrosine (Y)	
	#residues	Mol %	#residues	Mol %
<i>B. mori</i> (HC)	11	0.2	275	5.3
<i>A. pernyi</i>	34	1.3	138	5.3
<i>A. mylitta</i>	4	0.8	28	4.7
<i>A. yamamai</i>	44	1.6	141	5.0
<i>A. assamensis</i>	46	1.6	137	4.9
<i>S. ricini</i>	16	0.6	167	5.8

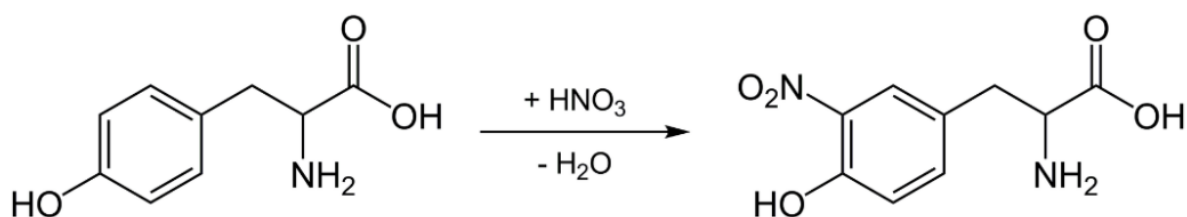


Figure 29. Xanthoproteric reaction of tyrosine and the formation of 3-nitro-tyrosine

⁴³⁶ Whitmore, 2012

⁴³⁷ De Fazi, *et al.*, 1959; Berti, 1960

⁴³⁸ Yang, *et al.*, 2010

The solubilisation result was similar to the sole use of the calcium nitrate solvent: 1mg fibre treated in up to total 1ml solvent left a small amount of sedimentation after the centrifugation. However, it should be noted that the exception to this was the *Bombyx* silks. As for the *Antheraea* and *Samia* silks, the *A. pernyi*, *A. mylitta* and *A. assamensis* exhibited a similar appearance leaving small amounts of sedimentation. However, the *A. yamamai* and *S. ricini* showed more sedimentation, indicating that these two fibroins are more rigid against chemical and thermal stimulation. From these results, it can be inferred that the solubilisation of silk fibroins is mainly affected by the number of salt molecules in the solvent, rather than by the acidity. Moreover, the addition of formic acid can protonate amino acids into zwitterion, and thus increasing the solubility by inducing stronger interactions between the salt molecules and the zwitterionic functional groups.

IV.2.4 Effect of the solvents on other fibres

From the series of experiments, it was confirmed that a 7M calcium nitrate and a 10M lithium thiocyanate solutions could successfully solubilise *Bombyx*, *Antheraea* and *Samia* silks (see Figure 25). A test was implemented to examine the effect of these two solvents on cotton, linen and wool fibres: often the archaeological or historical samples of unknown identity are mixed with other materials, which is challenging to separate. Thus, it would be useful to know if these solvents are fibroin-specific or compatible with other fibrous materials.

Boiled at 125°C, fibres in 7M calcium nitrate solution showed no significant changes and some swelling in cellulose fibres was observed after 2 hours. However, fibres boiled in 10M lithium thiocyanate solution showed active extraction of dyes from wool fibres and some degree of swelling in 30 minutes. The swelling was more severe in cellulosic fibres, and after 2 hours of treatment, a considerable loss of weight was observed. This result is consistent with the previous research that swelling of cellulose in aqueous calcium/lithium thiocyanate leads to mercerisation.^{439 440} The comparison between the two solvents in the process of further protein analysis will follow. However, this must be put into consideration in the selection of solvent when investigating a mixed material: the changes in weight of all samples are presented in Table 9.

Table 9. The weights of cotton, wool and linen fibres before and after the 2 hours of solvent treatment at 125°C

Solvent	Cotton	Wool	Linen
7M Ca(NO ₃) ₂	1.00 mg	0.92 mg	0.97 mg
10M LiSCN	0.69 mg	0.72 mg	0.53 mg

⁴³⁹ fibre-swelling and structural relaxation of cellulosic fibres often done by caustic soda solution (22-27%)

⁴⁴⁰ Katz and Derksen, 1931

IV.3 Desalting and concentration

Silk fibroin solutions were made with the saturated solution of chaotropic salts. However, salts, chaotropic agents, and detergents used in the solubilisation of protein must be removed to enable enzymatic digestion to take place, because their presence is not compatible with digestion. The most common method of desalting is the dialysis against H₂O. The two types of dialysis membrane tested in this experiment, both cassette and tube, could reduce the effort and amount of wastewater; but a large amount of precipitation was observed once the salt was removed. This may have resulted from the spontaneous nucleation of the β -form,⁴⁴¹ and the loss of salt ions seems to affect the aggregation of the disoriented protein (See II.3.3).⁴⁴² As an alternative, methanol-chloroform precipitation was applied. The proteins were successfully precipitated from the solutions; however, the formed pellets did not re-solubilised in any buffer, but in the original solvent.

Thus, to maintain the fibroin in solution, a centrifugal filter was used for buffer exchange and concentration. In the beginning, urea-thiourea buffer was used. It was suitable for in-solution digestion, but the exchange process was slow due to the high concentration of urea in the buffer, and to concentrate the solution the buffer-exchanged solution had to be diluted with a same volume of water and filtered again. This prolonged the process time much more than expected, and furthermore, overnight acetone precipitation during the in-solution digestion was necessary to remove the urea and detergent used in the buffer prior to the digestion. However, the nanoLC-MS/MS result showed that the method is compatible with the detection of silk fibroins. In another attempt, 50mM ammonium bicarbonate solution was adopted to replace the urea and detergent buffer. It is simple to prepare, and

⁴⁴¹ Li, *et al.*, 2001

⁴⁴² Tashiro, *et al.*, 1972; Magoshi, 1977; Ha, *et al.*, 2003

the buffer-exchanged solutions can be further concentrated using a vacuum centrifuge. This is useful for the preparation of SDS-PAGE concerning controlling the sample quantity, as well as the amount of additive, urea for instance. Furthermore, in this case, the acetone precipitation was no longer needed during the in-solution digestion.⁴⁴³ However, there was still some degree of aggregation occurring after the buffer exchange and concentration. The exact amount was not measured due to the lack of equipment. The degree of aggregation was different in each silk fibroin solution, and it was affected by the solvents (Figure 26). Furthermore, although the ammonium bicarbonate solution can facilitate the process, it reacts with calcium nitrate and yields calcium precipitation as the result of the double replacement (salt metathesis).

It is reported that *B. mori* silk solution in lithium bromide showed no aggregation during the dialysis and showed better stability than the silk solution in Ajisawa's reagent.⁴⁴⁴ In this study, *B. mori* and *B. mandarina* fibroin solutions in Ajisawa's reagent and 7M calcium nitrate solution showed more aggregation compared to the fibroin solutions in lithium salts. In the cases of *Antheraea* and *Samia* silks, they showed much more aggregation compared to *B. mori*, especially in *S. ricini* and *A. yamamai*. Moreover, the fibroin solution in calcium nitrate showed more aggregation than the fibroin solution in lithium thiocyanate. The re-assembly of fibroins is related to their structural characteristics and the way these solvents affect the organisation of fibroin molecules.⁴⁴⁵ Studies have shown that the silk fibroin solubilises in different manners by different reagents: calcium chloride disassembles the fibroin into nano-fibrils by breaking the chain bonds in the hydrophilic and hydrophobic blocks, while lithium bromide solubilises fibroin into smaller-

⁴⁴³ Kim, *et al.*, 2006

⁴⁴⁴ Jin and Kaplan, 2003

⁴⁴⁵ Matsumoto, *et al.*, 2008

sized molecules still preserving those blocks.⁴⁴⁶ Lithium thiocyanate is known to dismantle the molecular orientation back to random coil formation;⁴⁴⁷ calcium nitrate is reported to extensively depolymerise the fibroin and result in a mixture composed of heterogeneous molecular weights.⁴⁴⁸

In summary, if the fibroin polymer is solubilised into unfolded molecules of primary structure, it is more prone to form β -sheet structures (secondary structure) and aggregate during dialysis, due to the more exposed hydrophobic residue. However, it is yet unclear why the severe aggregation behaviour of wild silk fibroin solutions is also affected by the loss of heat, which is related to the crystallisation of the liquid silks in wild silk species.⁴⁴⁹ Furthermore, it is reported that the addition of formic acid (carboxylic acid) induces gelation of fibroin solution, assembling the small molecules into micelles and micrometre-sized globules;⁴⁵⁰ however, it did not affect the mass spectrometry analysis result (see V.4). The aggregation induced by the molecular folding is unavoidable at this point but leaves room for further researches.

⁴⁴⁶ Zhang, *et al.*, 2015

⁴⁴⁷ Goto, *et al.*, 1990

⁴⁴⁸ Bhat and Ahirrao, 1983; Jin and Kaplan, 2003; Hu, *et al.*, 2006; Zhang, *et al.*, 2015; Zheng, *et al.*, 2016

⁴⁴⁹ Tanaka, *et al.*, 2002 See II.3.3 Secondary structure.

⁴⁵⁰ Kweon and Park, 2001

IV.4 Protein gel electrophoresis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) provides a low-resolution separation of solubilised proteins based on molecular mass. This technique enables to estimate the relative molecular mass and the relative abundance. Furthermore, if the mixture of peptides is highly complex, it is advantageous to use a separation step prior to tandem mass spectrometry to limit the number of peptides the mass spectrometer sees over the time of the analysis.⁴⁵¹ In this study, it was also used to confirm how the solubilisation has affected the properties of the protein. The separation of proteins followed by excision of the individual spots from the gel and in-gel digestion allows focussing on the subject proteins (Figure 30).

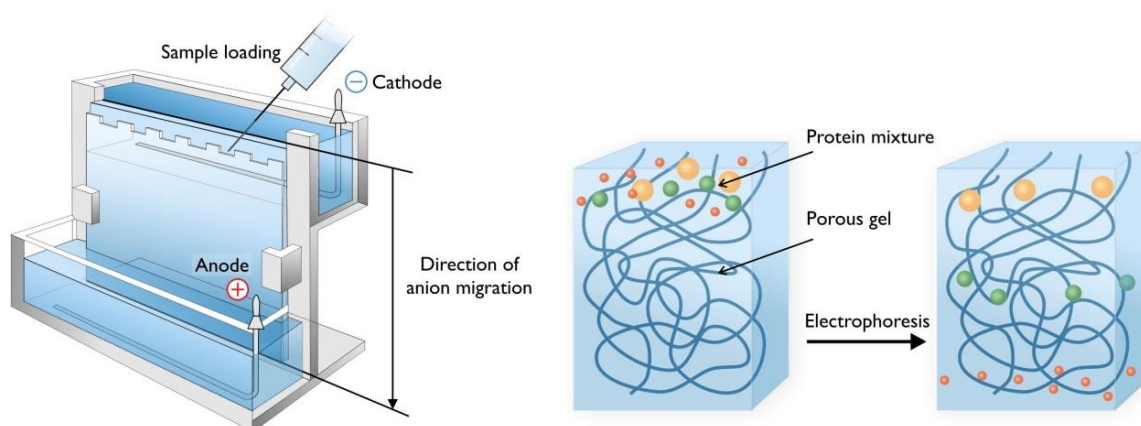


Figure 30. Separation of proteins in a polyacrylamide gel⁴⁵²

In the initial trial, there were no prominent protein bands observed within the area of coverage of protein separation; only the dye front, where unresolved proteins held up, were reacted to the Coomassie blue staining (Figure 31). The SDS-PAGE result showed that

⁴⁵¹ Delahunty and Yates, 2005

⁴⁵² Hegyi, *et al.*, 2013

subsided appearance was similar in all samples regardless of the reagent used for solubilisation. However, fibroin solutions made with 7M calcium nitrate showed more prominent staining than those made with 6N HCl. The visible bands from the dye front were cut and digested with trypsin. The nanoLC-MS/MS analysis result of 7M calcium nitrate-fibroin solutions showed the presence of proteins in the dye front, but at low protein sequence coverage (Table 10).

When the protein separation was resumed, desalting and concentration methods of fibroin solutions using buffer exchange were improved. Thus, the new set of samples made with revised protocols was tested with SDS-PAGE. As a result, lightly stained, continuous bars appeared after the Coomassie blue staining. In *Bombyx* silks, it was possible to spot the bands at 25-30kDa region corresponding to the light chain and P25. However, the *Bombyx* heavy chain and the other wild fibroins appeared as a tailing stream instead of a clean band; *S. ricini* did not show the streamline, and only the dye front was stained. The appearance of tailing stream is reported in the previous SDS-PAGE studies of solvent-dissolved *A. Pernyi* and *A. yamamai* fibroins.⁴⁵³ The most probable explanation to this would be that the solubilisation process denatured the proteins, and further depolymerised into a pool of polypeptides of various MW during the process.

⁴⁵³ Kweon and Park, 2001; Tao, *et al.*, 2007; Jo, *et al.*, 2017

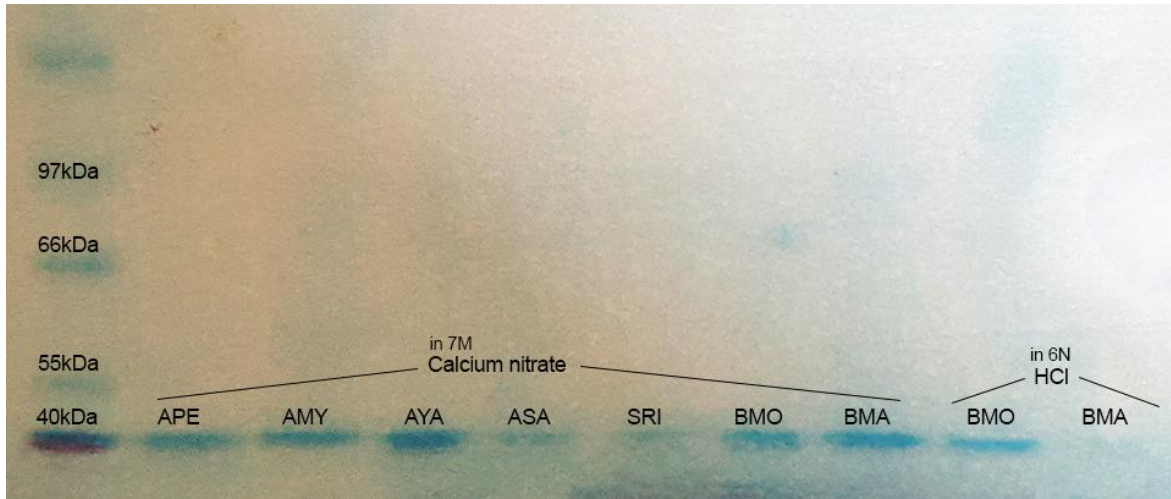


Figure 31. SDS-PAGE result of fibroin solutions (3-8% Tris-Acetate gel, 40-400kDa): only the dye front reacted to the staining

Table 10. PEAKS analysis result of the SDS-PAGE processed fibroin solutions (Tryptic digestion: heavy chain only for the *Bombyx* silks; **B. mandarina* corresponded to *B. mori* HC)

	BMO	BMA*	APE	AMY	AYA	AAS	SRI
-10lgP (confidence)	132.2	91.85	235.43	167.99	67.09	272.74	61.41
Coverage (%)	2	1	6	17	1	6	3
(expectation)	(2.3)	(58.6)	(22.8)	(71.6)	(21.2)	(10)	(10.3)
#Peptides	11	2	40	19	3	43	5
#Unique peptides	11	2	3	2	1	26	5
Peptide-Spectrum Matches	225	74	129	77	44	119	17
Peptide sequences	119	50	64	39	22	69	10
De novo Only spectra	812	1115	737	903	1030	1103	1027

The PEAKS peptide score (-10lgP)⁴⁵⁴ was the highest in *B. mori*, and then followed by *A. pernyi* and *A. assamensis*. The low values of *B. mandarina* and *A. mylitta* are thought to be affected by the partial sequences: if the detected spectra were from the unsequenced

⁴⁵⁴ -10lgP score: the ions score for an MS/MS match is based on the calculated probability, P, that the observed match between the experimental data and the database sequence is a random event. The -10lgP score is a conversion of a p-value on an increasing scale. For a given score x, its equivalent P-value is the probability that a false identification has a score greater than or equal to x. A P-value of 1% is equivalent to -10lgP of 20; a greater than 20 score is of relatively high confidence. The higher means the more significant peptide-spectrum match (<http://www.bioinform.com/dbscoring-tutorial/>).

areas of the protein, the confidence level of peptide-spectrum matches would decrease. On the other hand, the low $-10\lg P$ scores of *A. yamamai* and *S. ricini* are thought to have resulted from the poor resolution of proteins in the gel: this is by what was observed in the solubilisation tests, that *A. yamamai* and *S. ricini* showed the less effective reaction to the reagents compared to other silk samples. Such behaviour is thought to be related to their structural characteristics. In theory, *S. ricini* fibroin is the most hydrophobic among all samples. However, although the high contents of hydrophobic amino acids and polyalanine sequence must have done their parts, *A. yamamai* fibroin is less hydrophobic than *A. pernyi* or *A. assamensis* (Table 6). Thus, it is more likely to be affected by the repetitive block structure of the fibroins: the level of repetitive sequences (tandem repeat) is the highest in *S. ricini* (44.5%), then in *A. yamamai* (43%)—which is higher than in *A. pernyi* (42.2%) and *A. assamensis* (39.6%).⁴⁵⁵

The results may have been improved by adjusting the protein concentration of the samples or using another staining method of higher sensitivity. However, unfortunately, a further study using protein gel separation could not be implemented for a year due to the change of a research facility and lack of equipment. Thus, during the period, the method of digestion was switched to the in-solution process. When the protein separation by SDS-PAGE was resumed, desalting and concentration methods of fibroin solutions using buffer exchange were improved. Thus, a new set of samples made with revised protocols was tested with SDS-PAGE. As a result, lightly stained, continuous bars appeared after Coomassie blue staining (Figure 32). In *Bombyx* silks, it was possible to spot the bands at 25-30kDa region corresponding to the light chain and P25. However, the *Bombyx* heavy chain and the other wild fibroins appeared as a tailing stream instead of a clean band; *S.*

⁴⁵⁵ Malay, *et al.*, 2016

ricini did not show the streamline, and only the dye front was stained. The appearance of tailing stream is reported in the previous SDS-PAGE studies of solvent-dissolved *A. Pernyi* and *A. yamamai* fibroins.⁴⁵⁶ An explanation could be that the solubilisation process denatures the proteins, and they are further depolymerised into a pool of polypeptides of various MW during the process.

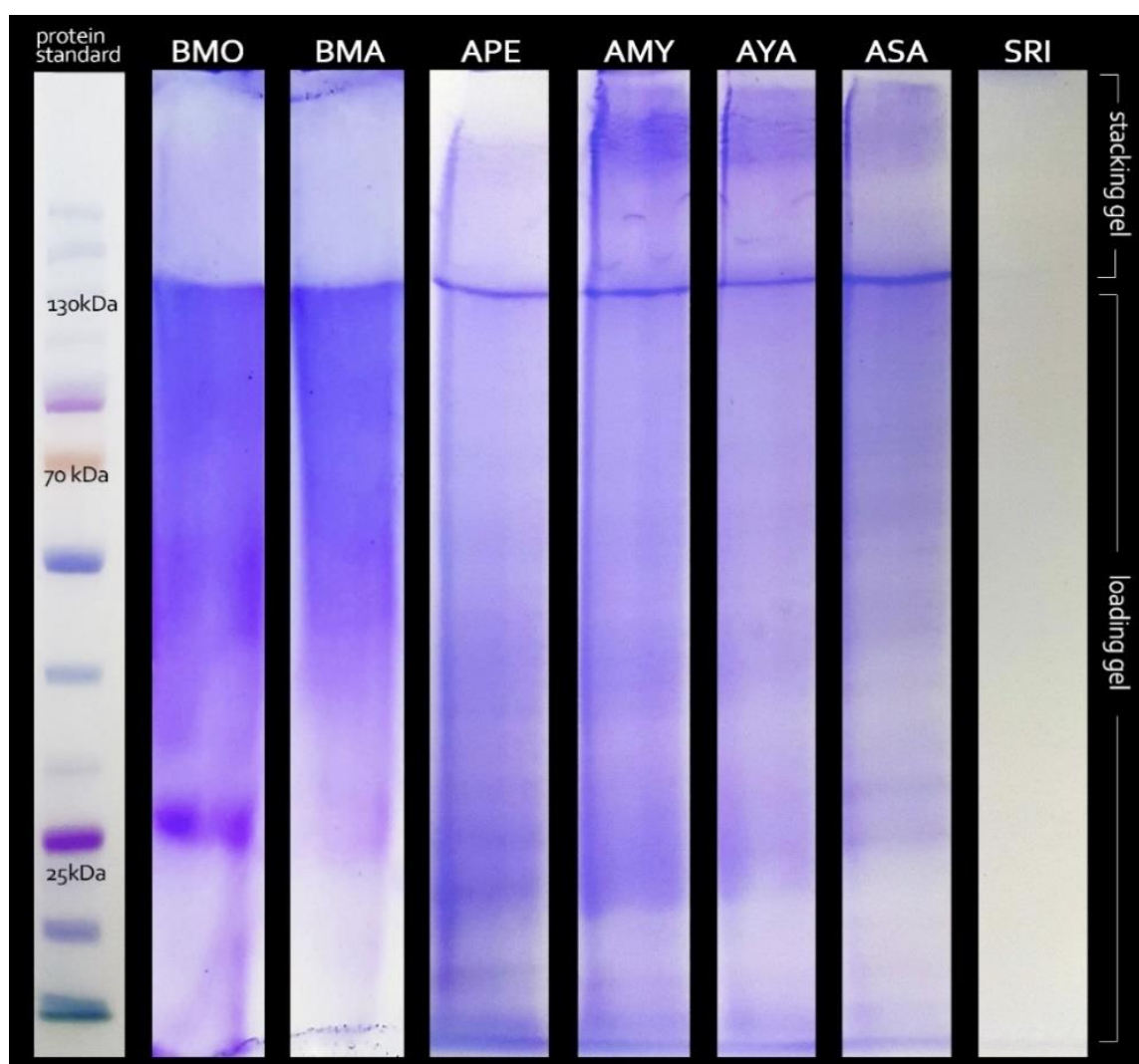


Figure 32. SDS-PAGE result of fibroin solutions (TGX 10% gel, 15-200kD)

⁴⁵⁶ Kweon and Park, 2001; Jo, *et al.*, 2017

In another attempt, the samples were run in a midi-sized gel of a bigger separation capacity with a few standard proteins. After the Coomassie blue staining, the standard proteins (2µg/µl) showed clean results (Figure 33): a broad, strong band for BSA was observed at 66.5kDa; both collagen I from human and rat tail both showed multiple bands corresponding to β region (200-300kDa) and α region (approx. 100kDa); and cytochrome-c (12 kDa) was observed near the dye front. However, the silk fibroin samples did not show a visible result except the band corresponding to the light chain and P25 in *B. mori*. Based on the previous mass spectrometry result, the proteins must have been resolved, but at a low ratio.

A double-staining method using silver stain over the Coomassie stain was applied to confirm this. The detection sensitivity of Coomassie stain (G250) is 30-40ng, whereas silver stain can detect as little as 5-10ng of proteins. The result was significantly improved: the same tailing stream, however much stronger and visible, were observed in all samples (Figure 34). This result indicates the only low amount of proteins were resolved into the gel, and this also explains why in some samples only the dye front was reacted to the staining—where unresolved proteins stock up at the end of the run. The concentration of fibroin samples was at least 5µg/µl concentration (from fully solubilised 1mg/1ml solution, suppose 50% were retrieved after ultracentrifugation, and then the final volume was concentrated to 100µl): thus, compared to the behaviour of other standard proteins, silk fibroins were not readily resolved into the gel.

After the silver stain, it was possible to see the difference in molecular weight distribution between *Bombyx* silks and *Antheraea* and *Samia* silks. The former was distributed from around 400kDa, which is correspondent to the *B. mori* heavy chain (~390kDa); the latter are distributed from around 230kDa, which is correspondent to the

approximate MW of each fibroin based on the sequence data.⁴⁵⁷ Furthermore, the SDS-PAGE analysis of silk fibroins aided the fibre identification. The unknown fibre sample on the right end (N2747, debris from a Gallo-Roman tomb) appeared to be the *B. mori* silk: the sample showed a very similar distribution to *B. mori*, indicating the presence of the heavy chain (~390kDa) molecule. The gel separation could be a useful and straightforward identification method to distinguish *Bombyx* silks from others; however, it is difficult to distinguish one species from another, as the *Antheraea* and *Samia* silk fibroins have similar molecular weight to each other.

Unfortunately, as is well known, silver stained gel samples did not show good compatibility with the mass spectrometry. Thus, to be able to apply the gel separation method, the resolution problem must be tackled. However, to promote the ease of the digestion, the gel or filter-aided digestion methods can be applied.

⁴⁵⁷ The MW of *B. mandarina* and *A. mylitta* fibroins are unavailable due to the limited sequence information.

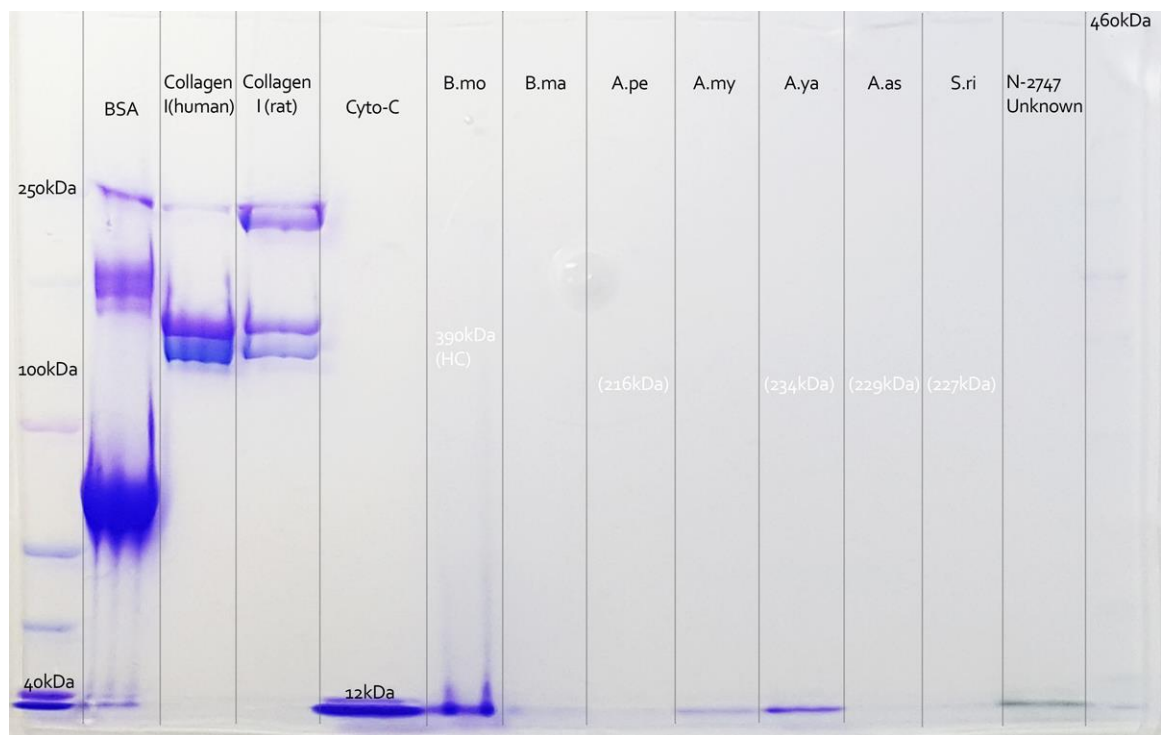


Figure 33. SDS-PAGE result of standard proteins and fibroin solutions: 3-8% Tris-Acetate gel (40-400kDa), Coomassie blue stained.

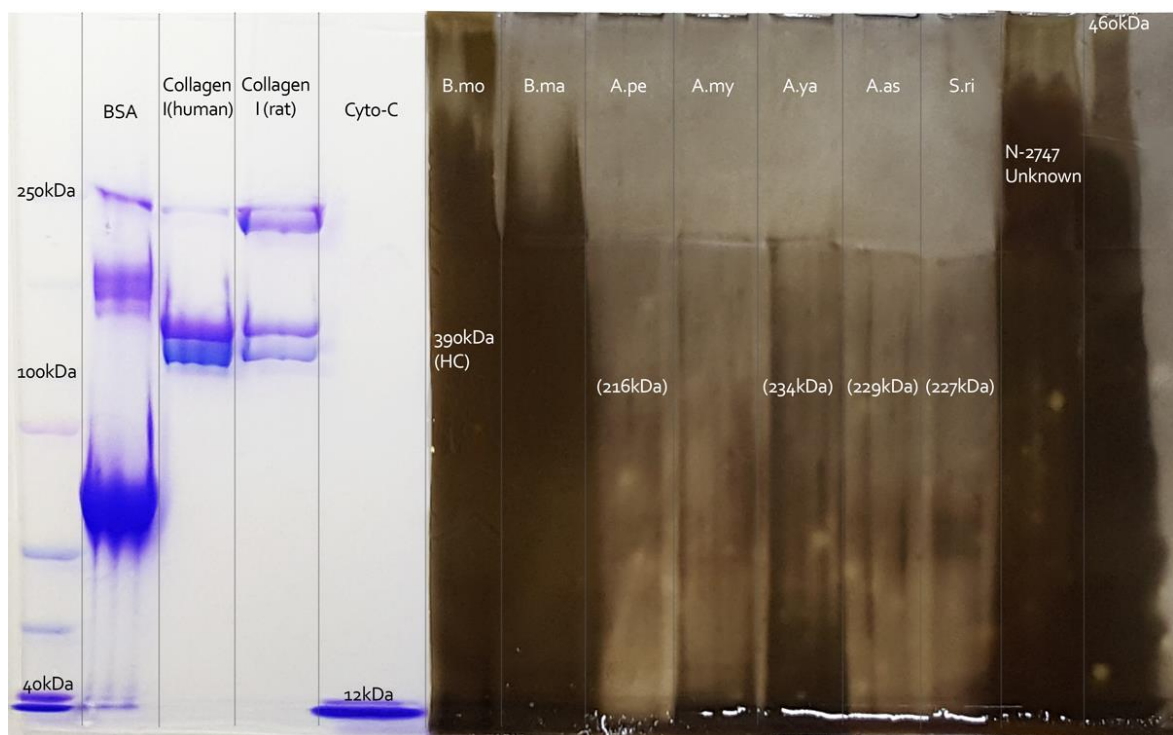


Figure 34. SDS-PAGE result of standard proteins and fibroin solutions: 3-8% Tris-Acetate gel (40-400kDa), partially silver stained over Coomassie blue.

V. Identification of silk fibroins by peptide analysis

V.1 Introduction

Although *B. mori* silk was studied with high-resolution LC-MS/MS in a few previous studies,⁴⁵⁸ *Antheraea* and *Samia* silks have never been examined using the same technique. Thus, there is no established protocol to prepare fibroin solutions for protein mass spectrometry. Hence, in this chapter, a series of studies were implemented in order to consolidate a workflow for the successful protein mass spectrometry of various silk fibroins. First off, protein sequences of *Bombyx*, *Antheraea* and *Samia* silk fibroins were analysed to define criteria for analysis. The aim was to find characteristic features of each protein sequence that distinguish one genus or species from another. Secondly, silk fibroin solutions resulted from the previous chapter were examined by nanoLC-MS/MS, and the resulted spectra were analysed by both database query and *De novo* methods. The solvents used to process the fibroin into solution, the enzymes used to cleave the protein into peptides, and the methods of digestion were varied to study their effect on the protein identification results. At last, the analytical protocol developed through the above processes was applied to a case study of severely degraded archaeological textiles.

⁴⁵⁸ Solazzo, *et al.*, 2012, Li, *et al.*, 2015, Gong, *et al.*, 2016

V.2 Protein sequence analysis

V.2.1 Sequence alignment

It is essential to recognise the characteristic protein sequence features to distinguish a silk fibroin from another using peptide analysis. Sequence alignment was used to compare the seven different fibroins: MSA (multiple sequence alignment) aligns three or more biological sequences (protein or nucleic acid) of similar length. The homology between the sequences and the unique features of each protein were studied based on the output. The protein sequences of each silk fibroin were retrieved from UniProtKB (Table 11).

Among the 11 protein sequences of seven different species of silks, only the *Bombyx* silks sequences are manually annotated through experiments and reviewed (Swiss-Prot), and the others are automatically annotated from their genomic DNA or mRNA and not reviewed (TrEMBL). Thus, sometimes a silk fibroin has multiple sequences that differ from each other. In such cases, the most recent sequence was selected, and the partial sequences were omitted if it is a part of a more extended sequence.

Table 11. Protein names and codes (UniProtKB) for each type of silk samples (*: Reviewed)

Species	Protein name	Acquisition code	Length
<i>B. mori</i>	Fibroin heavy chain	P05790 (FIBH_BOMMO)*	5263
	Fibroin light chain	P21828 (FIBL_BOMMO)*	262
	P25(Fibrohexamerin)	P04148 (SI25_BOMMO)*	220
<i>B. mandarina</i>	Fibroin heavy chain	Q99050 (FIBH_BOMMA)*	178
	Fibroin light chain	Q9BLL9 (Q9BLL9_BOMMA)	262
	P25(Fibrohexamerin)	Q9BLL8 (Q9BLL8_BOMMA)	220
<i>A. pernyi</i>	Fibroin	O76786 (O76786_ANTPE)	2639
<i>A. mylitta</i>	Fibroin	Q8ISB3 (Q8ISB3_ANTMY)	507
<i>A. yamamai</i>	Fibroin	E1CGA3 (E1CGA3_ANTYA)	2856
<i>A. assamensis</i>	Fibroin heavy chain	A0A0K0KR73 (A0A0K0KR73_9NEOP)	2809
<i>S. ricini</i>	Fibroin	A0A0D5ZYI3 (A0A0D5ZYI3_SAMRI)	2880

Unfortunately, *B. mandarina* heavy chain (Q99050) and *A. mylitta* fibroin (Q8ISB3) were only available in partial sequences of only 178 and 507 residues, respectively. *B. mori* fibroin comprises three proteins, and all three sequences are reviewed on UniProtKB (P05790, P12828 and P04148). *B. mandarina* has a reviewed heavy chain sequence, but it is a partial sequence of 178 lengths (Q99050); the light chain and p25 are translated from their mRNA (Q9BLL9 and Q9BLL8, respectively). The protein sequences of *Antheraea* and *Samia* silk fibroin are translated from their genomic DNA. These fibroins have 30-40% similarity to *Bombyx* silks showing 70-80% similarity between the *Antheraea* species; *S. ricini* and *Antheraea* silk fibroins are about 60% similar to each other (Table 12). Thus, mapping out the differences among these sequences is the first step to establish the criteria to identify each species through mass spectrometry, and to avoid the ambiguity of analysis.

Table 12. Similarity (%) between the silk fibroin sequences through Pairwise alignment

	<i>B. ma (HC)</i>	<i>A. pernyi</i>	<i>A. mylitta</i> *	<i>A. yamamai</i>	<i>A. assamensis</i>	<i>S. ricini</i>
<i>B. mo (HC)</i>	97.77%	40.91%	38.96%	41.66%	42.89%	42.54%
<i>B. ma (HC)*</i>		33.15%	32.02%	34.07%	32.96%	36.46%
<i>A. pernyi</i>			75%	77.79%	73.43%	60.77%
<i>A. mylitta</i>				68.94%	72.67%	57.44%
<i>A. yamamai</i>					77.33%	61.85%
<i>A. assamensis</i>						60.77%

*partial sequence

The selected protein sequences were compiled to a silk fibroin sequence database. The sequence alignments were made using Jalview (2.10.3b1) and MUSCLE algorithm (v.3.8.31). MUSCLE (multiple sequence comparison by log-expectation) is an accurate MSA tool especially good with proteins, claimed to achieve both better average accuracy and better speed.⁴⁵⁹ Each fibroin sequences and the complete alignment result is presented in the Appendix A (Figure 129): the alignment result illustrates the similarity between the sequences as well as their unique features that either only belongs to one protein or only shared among the same genus. The unique features are either a single amino acid in a specific location that differs from other sequences in the alignment, or a peptide that is often repeated multiple times throughout the sequence.

⁴⁵⁹ Edgar, 2004

V.2.2 Unique sequence features

V.2.2.1 Genus level

The similarity and difference between the genera of silks are well exhibited in their signal peptide. Table 13 shows the signal peptide at N-terminus of each fibroin, the putative signal peptidase recognition site of *Antheraea* species are characterised by AXA,⁴⁶⁰ while Bombyx species have VAY and *Samia ricini* has VTA in the same location.⁴⁶¹ However, it is unlikely to capture the signal peptides in MS/MS result. Thus, an analysis of fibroin sequences to examine the unique features was manually implemented.

Table 13. Signal peptides of each silk fibroin (highlighted in green)

	1	5	10	15	20	25
<i>B. mori</i>	M	R	V	K	T	F
<i>B. mandarina</i>	M	R	V	K	T	F
<i>A. pernyi</i>	M	R	V	I	A	F
<i>A. mylitta</i>	M	R	V	I	A	F
<i>A. yamamai</i>	M	R	V	T	A	F
<i>A. assamensis</i>	M	R	V	I	A	F
<i>S. ricini</i>	M	R	V	T	A	F

The most distinctive feature of Bombyx silks are the repetitive glycine-alanine hexapeptides of GAGAG(X), where X is S (432), A (151), Y (120), T (14), V (11) and G (1).⁴⁶² This particular arrangement takes up more than 80% of the entire sequence of *B. mori* yet is not found in *Antheraea* and *Samia*. As GAGAG is unique to *B. mori*, there is a variety of unique subsequence (Table 14). For example, GAGAGAG(X) occurs 150 times throughout the sequence, where X is Y (63) > A (60) > S (15) > T (12). The same rule applies to *B. mandarina*; however, it is not covered here due to the short length of the available

⁴⁶⁰ Gierasch, 1989

⁴⁶¹ There is no other *Samia* species fibroin sequence available yet.

⁴⁶² Zhou, *et al.*, 2001 listed that GAGAGS (433 times) and GAGAGY (120 times) are the most repeated throughout the amino acid sequence of *B. mori*.

sequence. There are some unique sequences other than the glycine-alanine variation: there is 34 occurrences of GVGAGY, 11 of GTGSSGFG and GYEVAW, 7 of PYVANGGYS and 3 of PYVAHGGYS, which are all unique to *B. mori* (Figure 35).⁴⁶³

On the other hand, one of the most distinctive characteristics of *Antheraea* and *Samia* silks is the consecutive alanine sequence of 4 or more. For example, (X)A₁₂(X), where X is often G, S, and R, is not present in *Bombyx* but abundant in *Antheraea* and *Samia*. The longest repetitive alanine sequence found in *Bombyx* silks is only one triple alanine sequence (AAA) at 159-161. SGAGG and RRAGHDR are also unique to *Antheraea* species and not present in *Bombyx* silks and *S. ricini*.

Table 14. Subsequence of GAGAG(X) unique to *B. mori*

GAGAG S	(432)	GAGAGS G	(431)	GAGAGSG A	(423)
				GAGAGSG T	(4)
				GAGAGSG S	(3)
				GAGAGSG V	(1)
		GAGAGS A	(1)		
GAGAG A	(151)	GAGAGA G	(150)	GAGAGAG Y	(63)
				GAGAGAG A	(60)
				GAGAGAG S	(15)
				GAGAGAG T	(12)
		GAGAGA A	(1)		
GAGAG Y	(120)	GAGAGY G	(120)	GAGAGYG A	(113)
				GAGAGYG V	(3)
				GAGAGYG T	(2)
				GAGAGYG G	(1)
				GAGAGYG Q	(1)
GAGAG T	(14)	GAGAGT G	(14)	GAGAGTG S	(11)
				GAGAGTG A	(3)
GAGAG V	(11)	GAGAGV G	(11)	GAGAGVG YGA	(11)
GAGAG G	(1)				

⁴⁶³ Residues marked in red are the unique residues and those marked in black do not affect the number of occurrences, however always presented together with the unique residues.

V.2.2.2 Species level

Although it seems straightforward to distinguish *Bombyx* from *Antheraea* and *Samia*, differentiating between *B. mori* and *B. mandarina* would be more complicated. As the two are closely related thus sharing many sequence characteristics. Their heavy chains have a similarity of 97.77% (1-178); the light chains 99.62% and the P25 99.09%. There are only minor differences between the *B. mori* and *B. mandarina* sequences that can be used to distinguish one from another in the mass spectrometry analysis (Table 15).

Table 15. Sequence difference of *B. mori* and *B. mandarina* through alignment

	Heavy chain (length=5263/178)	Light chain (length=262)	P25 (length=220)
	58 141 145	183	62 219
<i>B. mori</i>	...E Q ITT.....Y T SDF ST S...	...SDS S FL...	...VF Q FE.....FGC Q H
<i>B. mandarina</i>	...E E ITT.....Y R SDF A -S...	...SDS G FL...	...VF R FE.....FGC H H

Based on their chemical structure, in theory, the ageing and deterioration of the fibre would affect the light chain and P25 before the heavy chain. Thus, in the analysis of deteriorated archaeological samples, the characteristics of the heavy chain would be the essential criteria of identification. Once the complete sequence of *B. mandarina* heavy chain becomes available, it would give us a better idea of the unique features of each protein, as well as how the protein has evolved through thousands of years of domestication.

Compared to *Bombyx* silks, the *Antheraea* species have a relatively large number of unique sequence features that differentiate one species from another. Although there were many unique residues exclusive to one fibroin at multiple locations, these are only counted to the 507th residue, as *A. mylitta* sequence is incomplete. Also, it must be noted that the unique sequence features found in *Samia ricini* fibroin may not be exclusive to the species but shared among genus *Samia*. However, this can only be verified once another species of

Samia silk fibroin sequence becomes available. The list of some unique sequence features is presented in Table 16.

In *A. pernyi* fibroin sequence, YGGYGGYG is the most abundant unique peptide occurring 6 times throughout the protein sequence. There is also one occurrence of YGGYGGYGGY. AASGARGSG and SAARRAGHDS occur 3 times each. Among the unique residues, serine (S) was the most abundant amino acid.

In *A. mylitta* fibroin sequence of 507 residues, only RGHGG occurred more than once and other unique features are of a single occurrence: SYEWD, QATVVMDGAM, MVLTRAQQQLAAAA. Threonine (T) at the 107th residue, serine (S) at the 124th residue, and valine (V) at the 132nd residue are unique to the position that is not shared with other *Antheraea* and *Samia* species. In *A. yamamai* fibroin sequence, GDYGW is the most abundantly found unique peptide that occurs 6 times; GSGAGGAGGD occurs 5 times; VGGGYGG, ASSGAGGAGGGYGW and HAAGSSGGGYSWDYSSYGSES occur twice, respectively. QAGHE, SGRG and RRV occur once. Among the unique residues, the asparagine(N) at the 84th residue is characteristic; serine (S) was found the most frequently. In *A. assamensis* fibroin, more numbers of unique peptides are spotted than in other *Antheraea* species: featuring GSSGA –or in form of a longer peptide GGYGSGSSGA₁₂ – occurs 15 times, GSGAGGAGD occurs 13 times, EGYGSG occurs 10 times, GGSGAGG occurs 7 times, GDGYGW occurs 6 times, and GGSGAGGH occurs twice throughout the sequence. Among the unique residues, the asparagine(N) at the 75th residue and the phenylalanine (F) at the 114th residue is characteristic.

S. ricini fibroin shows the greatest number of characteristic features, mainly because no other *Samia* species were compared in this alignment. Consequently, whether these peptides are unique to a species *Samia ricini* or a genus *Samia* is unclear at this moment.

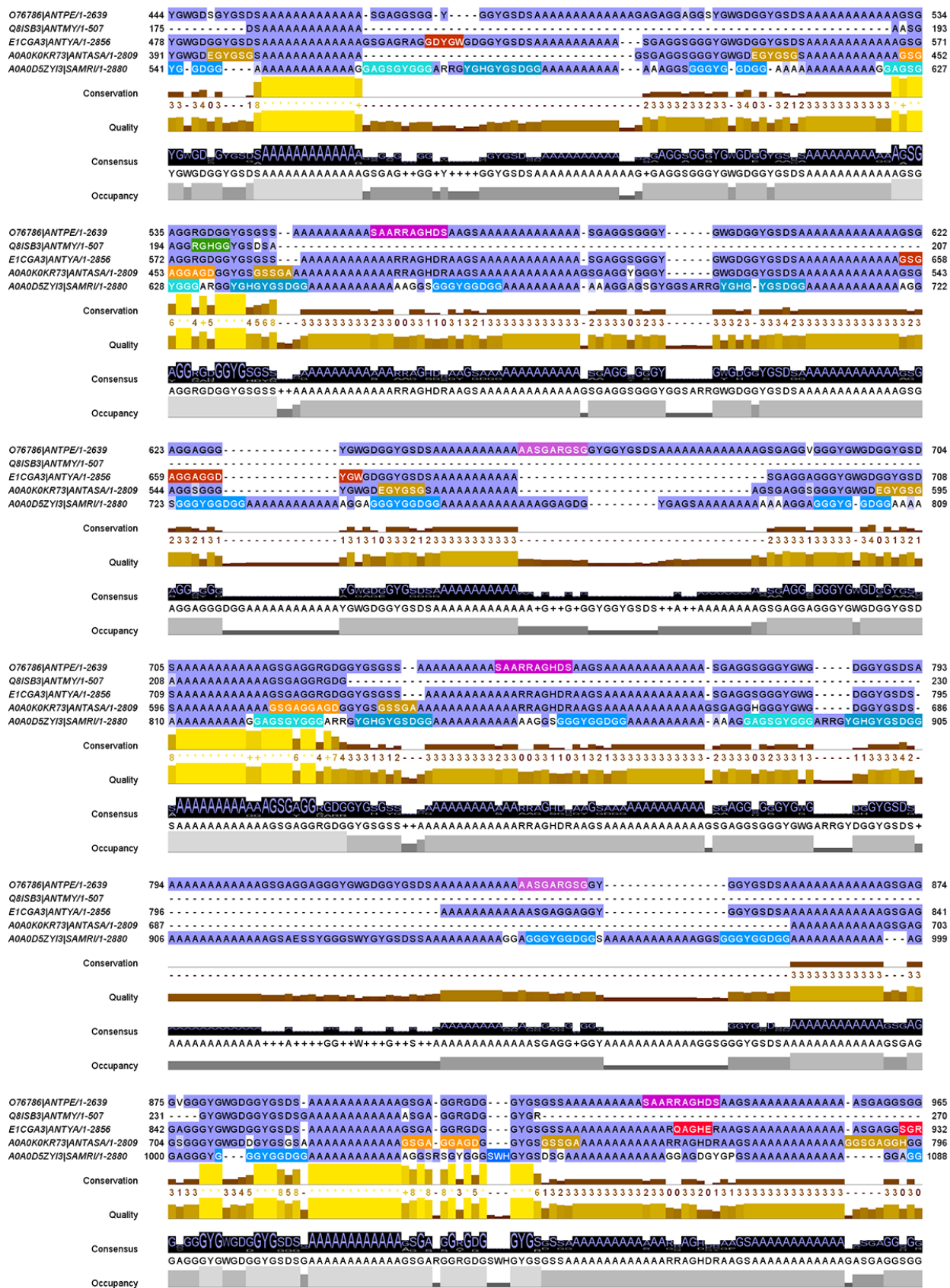
The occurrence of GGGYGGDGG is 35 times, GAGSGYGGG is 31 times, YGHGYGSD(GG) is 21(20) times, SWH is 12 times, SYGSDSGA is 8 times, and GTA is twice. These peptides are often adjacent to each other and make up a repetitive block. For example, GGGYGGDGG is coupled with polyalanine peptide A₁₁> for 34 times; GAGSGYGGG and YGHGYGSD(GG) is either coupled together (4 times) or connected by non-unique peptide ARGG (11 times) and ARRG (5 times). Also, GAGSGYGGG, SWH and SYGSDSGA are present as connected sequence (7 times).

Table 16. Unique peptides and residues found in *Antheraea* and *Samia* silk fibroins

Protein	sequence	location
<i>A. pernyi</i> fibroin (O76786 ANTPE)	YGGYGGYG	223-230, 1250-1257, 2083-2090, 2116-2123, 2269-2276, 2300-2307
	YGGYGGYGGY	1803-1809
	AASGARGSG	653-661, 839-847, 1021-1033
	SAARRAGHDS	559-568, 745-754, 931-940
	R	135
	N	144
	V	165
	G	214
	A	253, 504
	S	185, 190, 233, 249, 314, 346, 449, 507
<i>A. mylitta</i> fibroin (Q8ISB3 ANTMY)	RGHGG	167-171, 197-201
	SYEWD	337-341
	QATVVMDGAM	298-307
	MVLTRAQQQLAAAA	310-323
	T	107
	S	129
	V	132
<i>A. yamamai</i> fibroin (E1CGA3 ANTYA)	GDYGW	511-515, 664-668, 967-971, 1120- 1124, 1422-1426, 1575-1579
	GSGAGGAGGD	656-665, 959-968, 1112-21, 1414-23, 1567-76
	ASSGAGGAGGGYGW	159-172, 1783-1796
	HAAGSSGGGYSWDYSSYGSES	379-399, 2693-2713
	VGGGYGG	416-422, 2730-2736
	QAGHE	902-906
	SGRG	930-933
	RRV	223-225
	Y	44
	N	84
	S	137, 160, 455, 474
	G	430

<i>A. assamensis</i> fibroin (Q8ISB3 ANTMY)	GSSGA (GGYGSGSSGA ₁₂)	182-186, 304-308, 464-468, 623-627, 748-752, 908-912, 1067-71, 1224-28, 1348-52, 1544-48, 1735-39, 1896-1900, 2021-25, 2145-49, 2338-42
	GSGAGGAGD	168-176, 290-298, 450-458, 609-617, 734-742, 1210-18, 1334-42, 1530-38, 1721-29, 1882-90, 2007-15, 2131-39, 2324-32
	EGYGSG	271-276, 396-401, 431-436, 556-561, 590-595, 1509-14, 1827-32, 1863-68, 1987-92, 2303-08
	GGSGAGG	1209-1215, 1264-1270, 1586-1592, 2185-2191, 2380-2386
	GDGYGW	862-872, 1025-30, 1467-72, 2227-32, 2618-23, 2651-56
	GGSGAGGH	787-794, 823-830
	NGQLF	31-35
	ETRPNL	49-54
	RGGFYES	104-110
	N	75
	F	114
	S	123
	R	139
<i>Samia ricini</i> (A0A0D5ZYI3 SAMRI)	GGGYGGDGG	255-263, 336-344, 406-414, 465-473, 490-498, 538-546, 600-608, 662-670, 724-732, 749-757, 797-805, 860-868, 952-960, 977-985, 1016-24, 1087-95, 1149-57, 1229-37, 1258-66, 1339-47, 1401-09, 1481-89, 1562-70, 1620-28, 1682-90, 1707-15, 1765-73, 1883-91, 1945-53, 2026-34, 2144-52, 2168-76, 2311-19, 2532-40, 2631-39
	GAGSGYGGG	278-286, 359-367, 429-437, 561-569, 623-631, 821-829, 883-891, 1110-18, 1171-79, 1362-70, 1423-31, 1504-12, 1585-93, 1643-51, 1730-38, 1788-96, 1906-14, 1968-76, 2049-57, 2086-94, 2191-99, 2253-61, 2334-42, 2371-79, 2404-12, 2485-93, 2555-63, 2588-96, 2687-95, 2718-26
	YGHGYGSD(GG)	287-296, 372-381, 574-583, 636-645, 698-707, 834-843, 896-905, 1123-32, 1375-84, 1513-22, 1594-1603, 1656-65, 1739-48, 1801-10, 1919-28, 2062-2071, 2204-13, 2347-56, 2498-07, 2601-10, 2731-38 (YGHGYGSD only)
	SWH	438-440, 1038-40, 1180-82, 1290-92, 1432-34, 1977-79, 2095-97, 2262-64, 2380-82, 2413-15, 2564-66, 2663-65
	SYGSDSGA	441-448, 1837-1944, 1980-87, 2098-2105, 2265-72, 2333-40, 2416-23, 2557-64
	GTA	175-177, 248-250







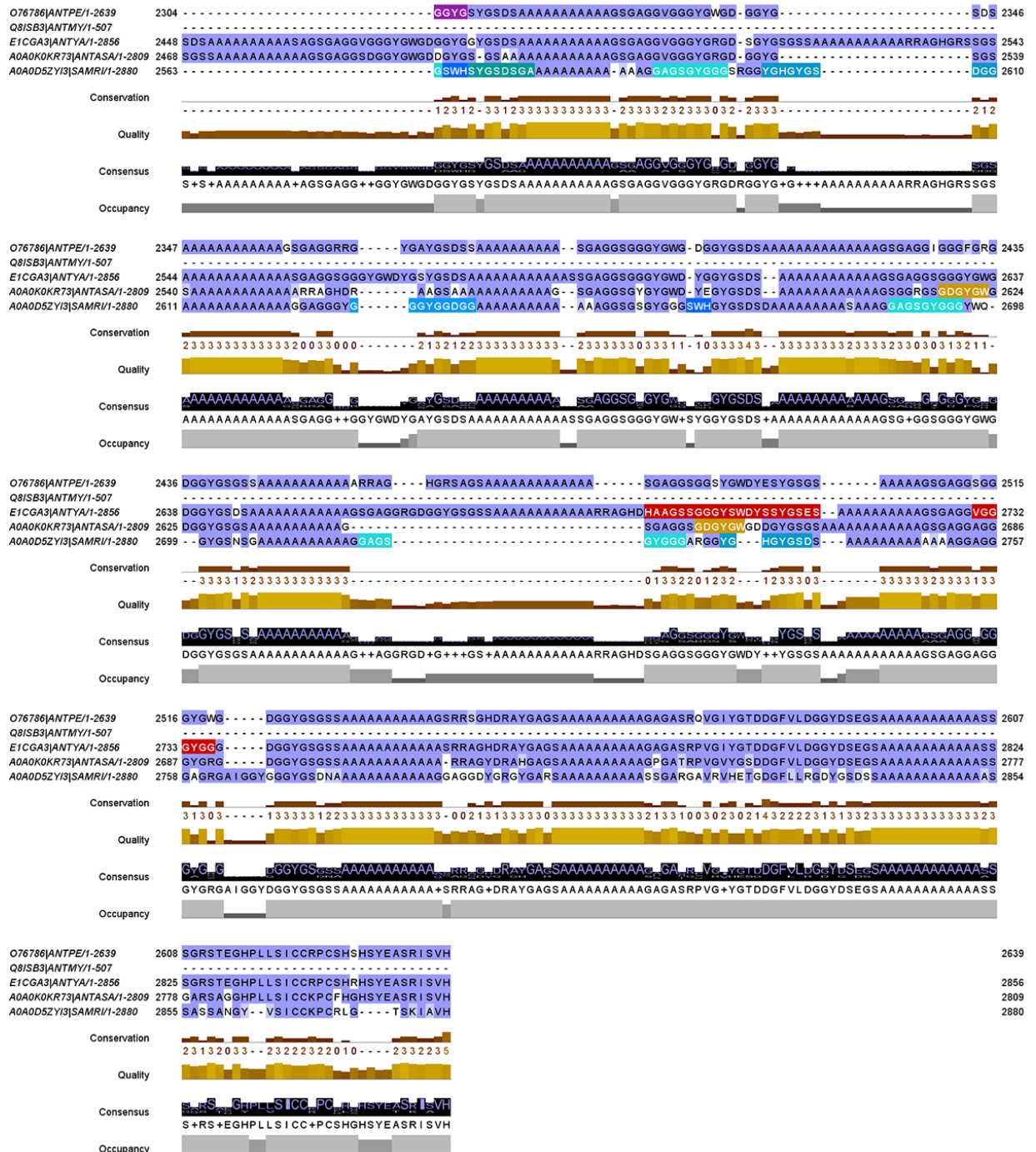


Figure 36. Alignment of *Antheraea* and *Samia* silk fibroins: unique peptides to each protein are marked in colour: *A. pernyi* (purple); *A. mylitta* (green); *A. yamamai* (yellow); *A. assamensis* (red) and *S. ricini* (blue). The main colour scheme is by Blosum62.⁴⁶⁴

⁴⁶⁴ If a residue matches the consensus sequence residue at that position it is coloured dark blue. If it does not match the consensus residue but the 2 residues have a positive Blosum62 score, it is coloured light blue.

V.2.3 Theoretical digestion with different enzymes and their effect on unique sequence features

Silk fibroins are polymers of high molecular weight about 300-400kDa, which is not the range that can be adequately analysed by mass spectrometers. Thus in protein mass spectrometry, samples are often cleaved into smaller peptides using proteolytic enzymes. Trypsin is the most commonly adopted enzyme for this purpose. Furthermore, it is reported to have a similar specificity characteristic to Cocoonase, a proteolytic enzyme that silkworms produce and spit out to soften the cocoon to exit at the end of metamorphosis.⁴⁶⁵ However, recent studies of *Bombyx* silk using nanoLC-MS/MS sequencing showed favour for chymotrypsin.⁴⁶⁶ These enzymes catalyse the hydrolysis of the esters of different amino acid groups, thus yielding peptides of different sizes and lengths. Trypsin cleaves peptides on the C-terminal side of basic amino acid, lysine (K) and arginine (R) residues; while chymotrypsin cleaves the C-terminal side of hydrophobic aromatic amino acids phenylalanine (F), tryptophan (W) and tyrosine (Y) and at a lower rate leucine (L). See Table 17 for their cleave rules and exceptions.

As chymotrypsin has more cleavage sites, it would create more peptides of smaller sizes. However, it appears that many of the unique peptides from *Antheraea* and *Samia* silk contain tyrosine, so this might affect the result of an analysis. Theoretical digestions using PeptideMass were performed to predict the behaviour of the silk fibroins against these two enzymes.⁴⁶⁷ Table 18 shows the size of the biggest resulting peptide and the protein coverage between 500 and 3000 Da of each silk fibroins, digested with trypsin and chymotrypsin, with up to 2 missed cleavages allowed. As a result, tryptic digestion of silk

⁴⁶⁵ Kafatos, *et al.*, 1967; Kafatos, *et al.* Ibid.

⁴⁶⁶ Solazzo, *et al.*, 2012, Zhu, *et al.*, 2014, Li, *et al.*, 2015, Gong, *et al.*, 2016

⁴⁶⁷ https://web.expasy.org/peptide_mass/ Wilkins, *et al.*, 1997; Gasteiger, *et al.*, 2005

fibroins yields more of bigger sized peptides than chymotryptic digestion; the protein sequence coverage of the silk fibroins in the range between 500 and 3000Da was almost three times higher in chymotryptic digestion than in tryptic digestion. This result was consistent in all samples except *B. mandarina* and *A. mylitta*; those have partial sequences: if their sequences were complete, they would show similar results to other silks.

The protein sequence coverage of *Bombyx* silk fibroin heavy chains in chymotryptic digestion is less than 50%, while *Antheraea* silks fibroins are about 70% and *Samia ricini* fibroin is 96.2%. Theoretically, if trypsin and chymotrypsin are used together, the size of the peptides decrease, and the coverage goes up to 100% in *Antheraea* and *Samia* silks, and to 47.5% in *B. mori* (100% in *B. mandarina*). However, it is yet to be tested: it is assumed that the harsh condition of solubilisation might induce a degree of degradation of the protein molecules and cause non-enzymatic cleavages. Moreover, in the case of severely degraded archaeological samples, the protein molecule might not be present in the physically intact condition; this would create more numbers of smaller peptides than predicted. The experimental results would verify these arguments.

Table 17. Cleavage rules of Trypsin and Chymotrypsin.⁴⁶⁸

	P2	P1	P1'	P2'
Trypsin	W	K	P	-
	M	R	P	-
(Exceptions)	C or D	K	D	-
	C	K	H or Y	-
	C	R	K	-
	R	R	H or R	-
Chymotrypsin-high specificity (C-term to [FYW], not before P)	-	F or Y	not P	-
	-	W	not M or P	-
Chymotrypsin-low specificity (C-term to [FYWML], not before P)	-	F, L or Y	not P	-
	-	W	not M or P	-
	-	M	not P or Y	-
	-	H	not D, M, P or W	-

Table 18. Comparison of each theoretical digestion result of each silk fibroin. The top mass indicates the biggest peptide (MW) resulted from the enzymatic digestion, and the coverage shows the sequence coverage of the peptides in the mass range between 500 to 3000 Da. (T: trypsin, C: chymotrypsin, T/C: trypsin and chymotrypsin)

	<i>B. mori</i> (HC)			<i>B. mandarina</i> (HC)		
	T	C	T/C	T	C	T/C
Top mass (MW)	37184	11592	11592	9032	9906	4035
Coverage (%)	21.2	45.9	47.5	58.6	46.5	100
	<i>A. pernyi</i>			<i>A. mylitta</i>		
	T	C	T/C	T	C	T/C
Top mass (MW)	19415	10155	4720	11195	10118	5802
Coverage (%)	22.8	71.3	100	71.6	67.7	92
	<i>A. yamamai</i>			<i>A. assamensis</i>		
	T	C	T/C	T	C	T/C
Top mass (MW)	26621	10088	5402	15535	7400	4708
Coverage (%)	21.2	70.0	100	10.1	67.7	100
	<i>S. ricini</i>					
	T	C	T/C			
Top mass (MW)	41178	12479	5893			
Coverage (%)	10.3	96.2	100			

⁴⁶⁸ https://web.expasy.org/peptide_cutter/peptidecutter_enzymes.html#exceptions

V.3 Database query vs. *de novo* analysis

The MS/MS results of calcium nitrate-fibroin solutions from IV.4 (Figure 31) were queried against the MASCOT search engine using UniProtKB and NCBIInr databases. *Bombyx* silks were successfully identified from UniProtKB; however, *Antheraea* and *Samia* silks were only identified from NCBIInr (Table 19).

In *Bombyx* silks, the protein score ($-10\lg P$) of the light chain was higher than that of the heavy chain and P25. However, given the molar ratio of the heavy chain, the light chain and P25 is 6:6:1, the result implies that only small amounts of the heavy chain protein were resolved into the gel; this was later confirmed by silver staining of the SDS-PAGE gel (Figure 34). In *B. mandarina* sample, the heavy and the light chains were identified as both *B. mori* and *B. mandarina* fibroin by the same probability. However, one unique peptide GQIPSQYEIPV**FR**(F) that distinguishes *mandarina* from *mori* was detected from P25, thus P25 is identified as that of *B. mandarina* with a higher probability. The heavy chain peptides matched to P05790 (*B. mori* HC) indicates that the unsequenced part of the *B. mandarina* heavy chain protein is highly similar to that of *B. mori*.

The overall result of *Antheraea* and *Samia* silk identification was not successful except in *A. pernyi* sample. *A. mylitta* and *S. ricini* silks were identified as *A. pernyi*, and *A. yamamai* sample retrieved the *A. yamamai* and *A. pernyi* by the same $-10\lg P$ score; *A. assamensis* sample did not show any adequate protein match. This is largely due to the similarity between the sequences. For example, NAATRPHLSGNERL and GDGGYGS₁₂GSSA₁₂R and are shared among *A. pernyi*, *A. mylitta* and *A. yamamai*; VPGASSSA₄SSASAGSGQTIIVER presents in *A. pernyi* and *A. yamamai*; other detected peptides are also shared between at least two *Antheraea* species. The result suggested that *de novo* analysis might be necessary to distinguish the *Antheraea* and *Samia* silk fibroins.

Moreover, the protein sequence coverage must be improved to be able to trace unique sequence features to distinguish each species. A list of the peptides resulted from this experiment is presented in Appendix B (Table 37).

The same set of samples were re-examined by another nanoLC-MS/MS using a different method with longer running time, which would allow peptides to stretch better.⁴⁶⁹ The MS/MS data were analysed by PEAKS software using *De novo* method: the experimental data were matched against the FASTA database of silk fibroin sequences compiled in V.2.1. The PEAKS analysis result displays some useful features for grasping the peptide identification result: the per cent coverage of a protein sequence is calculated by dividing the number of amino acids in all found peptides by the total number of amino acids in the entire protein sequence(the higher the better); the number of peptides is the number of detected peptides matching the target protein sequence; and the number of unique peptides indicates the number of peptides unique to the target protein within the database used for *de novo*.⁴⁷⁰

The new MS/MS analysis detected a higher number of peptides than the first run; the PEAKS analysis showed better identification results than the MASCOT analysis although there is still room for improving the protein sequence coverage (Table 20). In *B. mori* sample (BMO), the light chain and P25 were identified as *B. mandarina*; while the heavy chain was identified as *B. mori*. In *B. mandarina* sample (BMA), the most confident suggestion was the light chain that identified as both *B. mori* and *B. mandarina*; however, the -10lgP score of P25 was higher than that of the heavy chain. However, it is not clear from where the confusions between the light chain and P25 of *B. mori* and *B. mandarina*

⁴⁶⁹ I had to move from lab to lab twice during this period due to the facility closure and moving.

⁴⁷⁰ A unique peptide is a peptide with a -10lgP score above the peptide filtering threshold that can be mapped to only one protein group: ex) unique peptide to a protein within the FASTA database used for *De novo*. The resulted unique peptide sequences can be different from the ones listed in IV.1.2. i.e.) expanded form or only containing fragments

have resulted: this could be due to the sequence homology, or a sample run over.

All *Antheraea* and *Samia* silk samples showed the best match to their sequences apart from *A. mylitta* sample. In *A. pernyi* sample (APE), the result suggested the protein to be O76786 with -10lgP score of 296.36; there were 36 peptide hits with 1 unique peptide KHFER (42-46). APE sample also showed a high probability to be *A. yamamai* and *A. mylitta*. *A. mylitta* sample (AMY) was identified as *A. yamamai* with 29 peptides correspond to E1CGA3 with -10lgP score of 327.18; and identified as *A. pernyi* with 25 peptides correspond to O76786 with -10lgP score of 326.95. *A. mylitta* came up as the 3rd possibility with 19 peptides correspond to Q8ISB3 with -10lgP score of 247.88. While there were 2 unique peptides for *A. mylitta* were found, QASHGAGGAAGAAAGAAASSSVR (111-133) and SHGQLVER (29-36), there was no unique peptide for O76786 and only 1 unique peptide for E1CGA3 was found— GGSGFYETHD (135-144). The result suggests that the unsequenced part of *A. mylitta* fibroin is similar to that of *A. yamamai* and *A. pernyi*. However, it is not clear yet whether GGSGFYETHD also exists in the unsequenced part of *A. mylitta* fibroin or is it caused by sample run over during the gel process. In *A. yamamai* sample (AYA), there were 5 peptides correspond to E1CGA3 with -10lgP score of 143.2: NAATRPHLSGNER (47-59) and VPGASSSAAAASSASAG SG (86-104) are unique to E1CGA3, and the other 3 are shared with *A. assamensis*. *A. assamensis* sample (ASA) showed the largest number of unique peptides in this sample set. There were 25 unique peptides among the 39 peptides matched to A0A0K0KR73, with -10lgP score of 306.58. *S. ricini* sample (SRI) showed 12 peptides matched to A0A0D5ZYI3 with -10lgP score of 285.54 and all of them were unique.

Table 19. MASCOT MS/MS ion search results of silk samples (tryptic peptides, no result for *A. assamensis*); the residue marked in red indicates the unique amino acid that differentiate *B. mori* and *B. mandarina*.

	Accession	Description	Score	Peptide
<i>Bombyx mori</i>	P21828	Fibroin light chain/ OS=Bombyx mori	348	SGNFAGFR YIAQAASQVHV AWDYVDDTDK YSVGPALGCAGGGR SIAILNVQEILK GVGNNGNDATGLVANAQR
	P04148	Fibrohexamerin/ OS=Bombyx mori	157	EHIFGK LTTVFDK TAQWLSK TLAQHMSFK VSEFYDNVR DLTLLHIYER
	P05790	Fibroin heavy chain/ OS=Bombyx mori	114	NEKMIK NHGILGKNEK DASGAVIEEQITTK GYGQGAGSAASSVSSASSR
<i>Bombyx mandarina</i>	QBL9 (P21828)	Fibroin light hain/ OS=Bombyx mandarina (OS=Bombyx mori)	411	DIDDGKASSVISR GVGNNGNDATGLVANAQR SGNFAGFR YIAQAASQVHV YSVGPALGCAGGGR SIAILNVQEILK
	P05790	Fibroin heavy chain/ OS=Bombyx mori	213	DASGAVIEEQITTK DASGAVIEEQITTK GYGQGAGSAASSVSSASSR
	QBL8	Fibrohexamerin/ OS=Bombyx mandarina	84	TFVITDSDGNESIVEEDVLMK DLTLLHIYER GQIPSQYEIPVFR(F) LTTVFDK VSEFYDNVR
<i>Antheraea pernyi</i>	gi 3450883	fibroin [Antheraea pernyi]	256	NAATRPHLSGNERL QASHGAGGAAGAAAGAAAGSSAR HHDEYVDNHGQLVER VPGASSSAAAAASSASAGSGQTIIVER RVPGASSSAAAAASSASAGSGQTIIVER
	gi 15077406 gi 301017132	fibroin [Antheraea yamamai]	124	NAATRPHLSGNER QASHGAGGAAGAAAGAAAGSSAR VPGASSSAAAAASSASAGSGQTIIVER
	gi 23476394	fibroin [Antheraea mylitta]	54	NAATRPHLSGNER
<i>Antheraea mylitta</i>	gi 3450883	fibroin [Antheraea pernyi]	82	VPGASSSAAAAASSASAGSGQTIIVER RVPGASSSAAAAASSASAGSGQTIIVER
	gi 15077406 gi 301017132	fibroin [Antheraea yamamai]	72	VPGASSSAAAAASSASAGSGQTIIVER
<i>Antheraea yamamai</i>	gi 15077406 gi 301017132	fibroin [Antheraea yamamai]	103	NAATRPHLSGNER QASHGAGGAAGAAAGAAAGSSAR VPGASSSAAAAASSASAGSGQTIIVER
	gi 3450883	fibroin [Antheraea pernyi]	103	NAATRPHLSGNER QASHGAGGAAGAAAGAAAGSSAR VPGASSSAAAAASSASAGSGQTIIVER
	gi 23476394	fibroin [Antheraea mylitta]	46	NAATRPHLSGNER
<i>Samia ricini</i>	gi 3450883	fibroin [Antheraea pernyi]	104	VPGASSSAAAAASSASAGSGQTIIVER
	gi 15077406 gi 301017132	fibroin [Antheraea yamamai]	104	GDGGYGSGSSAAAAAAAAAAAAAAR VPGASSSAAAAASSASAGSGQTIIVER GDGGYGSGSSAAAAAAAAAAAAAAR
	gi 13991888 gi 1507659	fibroin [Antheraea pernyi]	61	GDGGYGSGSSAAAAAAAAAAAAAAR

Table 20. PEAKS/*De novo* analysis result of silk samples (tryptic digestion)

Sample	accession	-10lgP	Protein coverage (%)	#Peptides	#Unique peptides
BMO	Q9BLL9 FIBL_BOMMA	227.58	85	60	60
	Q9BLL8 P25_BOMMA	166.62	60	23	23
	P05790 FIBH_BOMMO	141.37	2	11	11
BMA	Q9BLL9 FIBL_BOMMA	274.35	38	10	10
	/P21828 FIBL_BOMMO				
	Q9BLL8 P25_BOMMA	192.24	22	5	5
	P05790 FIBH_BOMMO	121.86	0	1	1
APE	O76786 ANTPE	296.36	5	36	1
	E1CGA3 ANTYA	294.87	4	35	1
	Q8ISB3 ANTMY	259.43	23	24	2
AMY	E1CGA3 ANTYA	327.18	5	27	1
	Q8ISB3 ANTMY	247.88	17	18	2
	A0A0K0KR73 ANTASA	195.05	3	10	1
AYA	E1CGA3 ANTYA	143.2	2	5	2
	A0A0K0KR73 ANTASA	77.00	1	3	0
AAS	A0A0K0KR73 ANTASA	203.56	4	12	8
	O76786 ANTPE	164.70	4	11	7
SRI	A0A0D5ZYI3 SAMRI	284.54	4	12	12
	E1CGA3 ANTYA	85.84	1	3	2

Overall, the detection of *Bombyx* silks was successful in both methods. However, it appeared that *de novo* analysis using PEAKS is more efficient for the detection of *Antheraea* and *Samia* silk, as this method supports the screening of unique peptides. Nevertheless, it must be noted that the presence of other fibroins from BMA to SRI in the PEAK analysis result could have been affected by the sample run over during the gel process; this would be the best explanation to the presence of unique peptides from similar fibroins other than the subject fibroin. This can be avoided by inserting blanks in between the samples in the gel or switching to in-solution digestion and minimising contamination during the handling process.

V.4 Effect of solvents

V.4.1 *Bombyx mori*

The peptide analysis result of the same samples but solubilised with different reagents were compared. Trypsin was used for the proteolysis as it creates fewer cleavages than Chymotrypsin and yields bigger sized peptides—thus appropriate to assess the effect of the solvent. The preparation process for the MS/MS analysis and the analytical parameters was identical. The comparison was made based on PEAKS analysis results of each sample (Table 21). The examination of *B. mandarina* has been omitted as the result can be inferred from that of *B. mori*, and the available protein sequence is incomplete.

For the *B. mori* silk, calcium nitrate-formic acid (10%) system (CN-FA) appeared to be a better solvent for the heavy chain than calcium nitrate used alone (CN) or hydrochloric acid used alone (HCl). The HCl-fibroin solution showed high coverages for the light chain (84%) and P25 (39%), but only 1% for the heavy chain. Contrarily, CN and CN-FA samples showed 51% and 42% coverages in the light chain, but 4% and 8% coverages in the heavy chain, respectively: given that the length of the heavy chain is 20 times longer than that of the light chain, the increase in the heavy chain sequence coverage of 3-7% is a big improvement. The light chain and the P25 are much smaller proteins (30 and 25 kDa, respectively) than the heavy chain (350kDa) and have higher contents of hydrophilic amino acids. Thus, it is expected that these are readily solubilised in HCl solution and digested into the mass-spectrometry appropriate-sized peptides, although this solvent was not effective to solubilise the heavy chain: Figure 37 shows that the heavy chain peptides in this sample were only occurred near the N- and C-terminus. On the other hand, CN based solvents were effective for the heavy chain but probably too harsh for the light chain and

P25—thus the loss of these two proteins. However, the identification of the *Bombyx* silks in archaeological samples would be more dependent on the contents of the heavy chain than the light chain or P25, as these two are prone to degrade than the heavy chain. The results imply that the higher protein coverage for the heavy chain would be necessary to achieve a reliable protein identification of *Bombyx* silks by peptide analysis.

The comparison between the CN and CN-FA samples shows the effect of solvent on the heavy chain (P05790). There were 42 peptides for CN sample (all unique peptides), and 73 peptides (53 unique peptides) for CN-FA sample that correspond to P05790, including 28 common peptides. Both samples show the distribution of the peptides mostly occurring from the residues 1-1330 (Figure 38). However, CN-FA sample shows peptides occurring from 1750-5050 including 1 *de novo* tag CNGSGSGAGAGATPR at 1758-1766 (Figure 39). This implies that the addition of carboxylic acid aided in disturbing the structural orientation of the heavy chain, especially the predominant GAGAGX blocks, resulting in silk solutions prone to proteolysis. In contrast, the addition of inorganic acid or the strong acidity did not affect the solubility of silk. See Appendix B for the list of unique peptides (Table 38).

Table 21. Summary of PEAKS analysis results of *Bombyx mori* silk solubilised with different reagents, in-solution tryptic digestion

		7M CN	7M CN-FA (10%)	6N HCl
P05790 FIBH_BOMMO	-10lgP	242.58	210.85	175.37
	Coverage (%)	4	8	1
	#Peptides/#unique	42/42	73/53	7/7
P12828 FIBL_BOMMO	-10lgP	242.58	184.43	330.96
	Coverage (%)	51	42	84
	#Peptides/#unique	56/56	51/51	74/74
P01418 P25_BOMMO	-10lgP	95.69*	84.34	157.80*
	Coverage (%)	13*	8	39*
	#Peptides/#unique	7*/7*	4/4	9*/9*

* Detected as Q9BLL8|P25_BOMMA

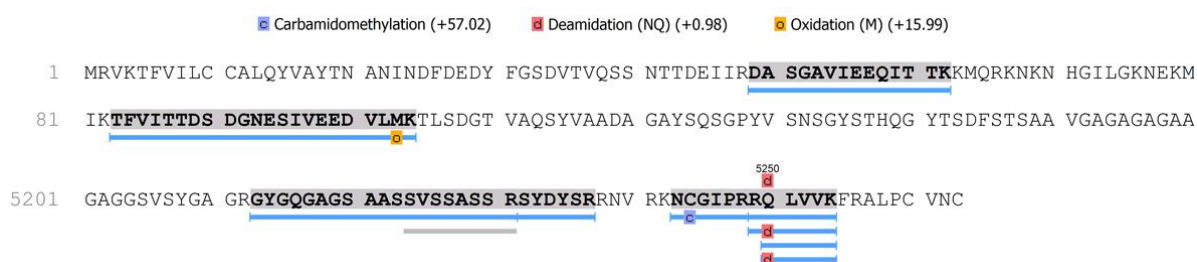


Figure 37. Protein coverage of the heavy chain (P05790) by *B. mori* fibroin-HCl solution.

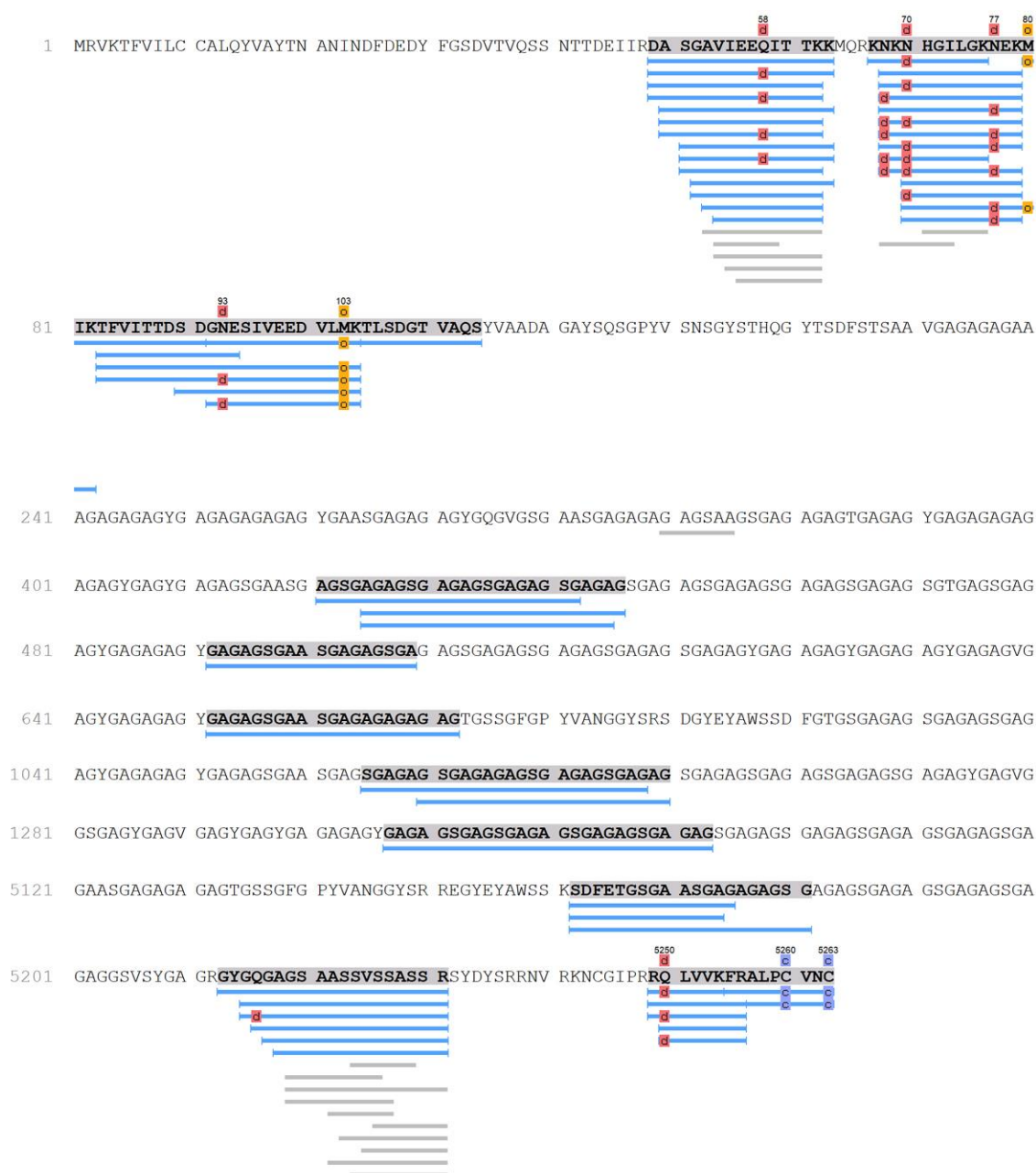
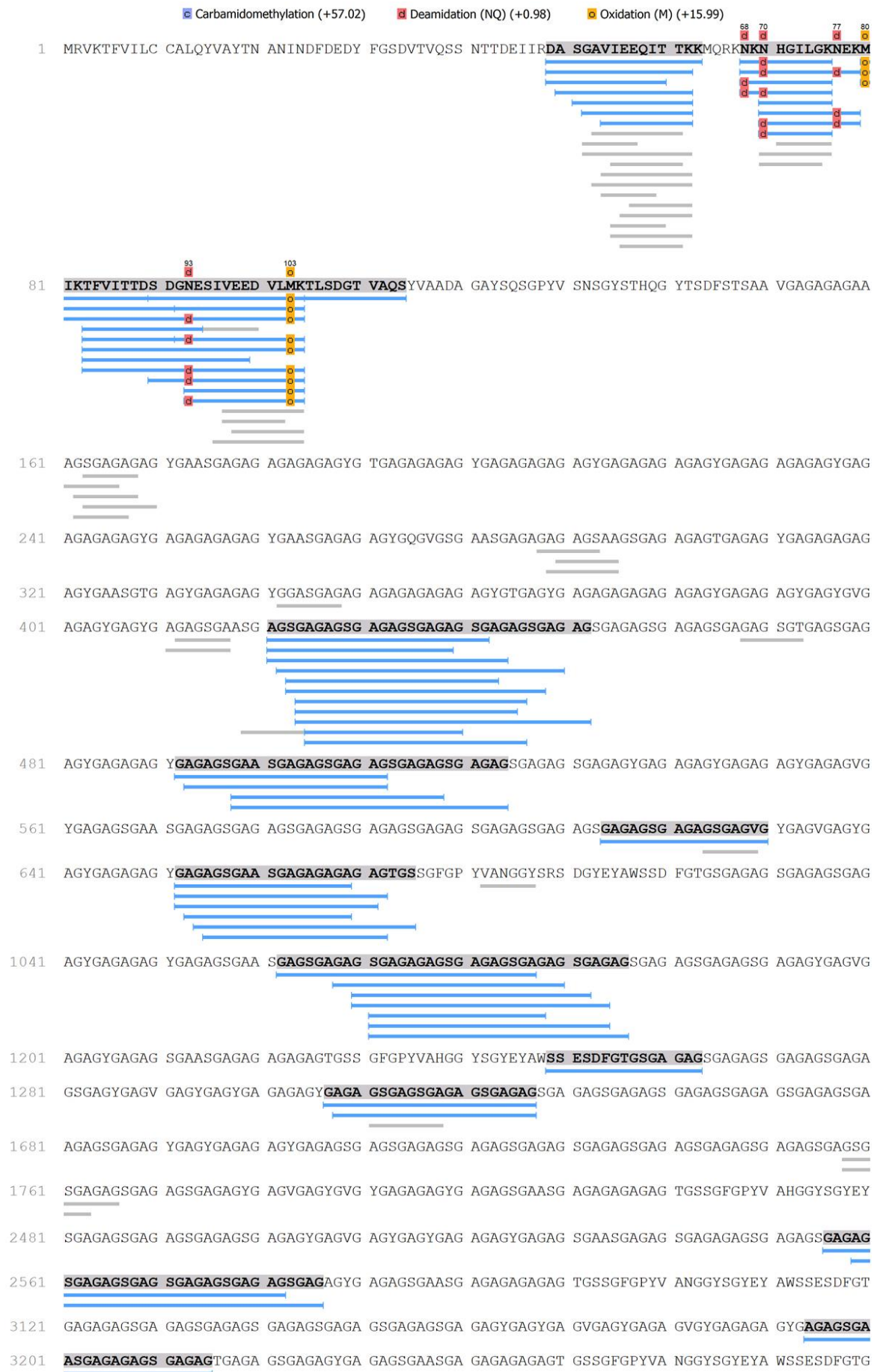


Figure 38. Protein coverage of the heavy chain (P05790) by *B. mori* fibroin-calcium nitrate solution.



V.4.2 Antheraea and Samia silks

Antheraea and *Samia* silks showed similar results to that of *B. mori*. In general, CN-FA samples showed better protein coverages than CN samples, especially in *A. assamensis* and *S. ricini* – which have shown the most resistance to solubilisation in the previous chapter (see IV.2.3). As also observed during the solubilisation, the addition of formic acid induces Xanthoprotic reaction, which is related to the nitration of tyrosine (Figure 29). It is reported that the degradation of tyrosine is related to the ageing of the silk.⁴⁷¹ Thus, the results from CN-FA samples might give an idea of how the aged and yellowed silks would react to the calcium nitrate solution. HCl alone did not solubilise these silk fibres thus excluded. The list of supporting peptides was grouped by the peptides that are common in both CN and CN-FA samples and the ones exclusive to each sample. A summary of PEAKS analysis results is presented in Table 22.

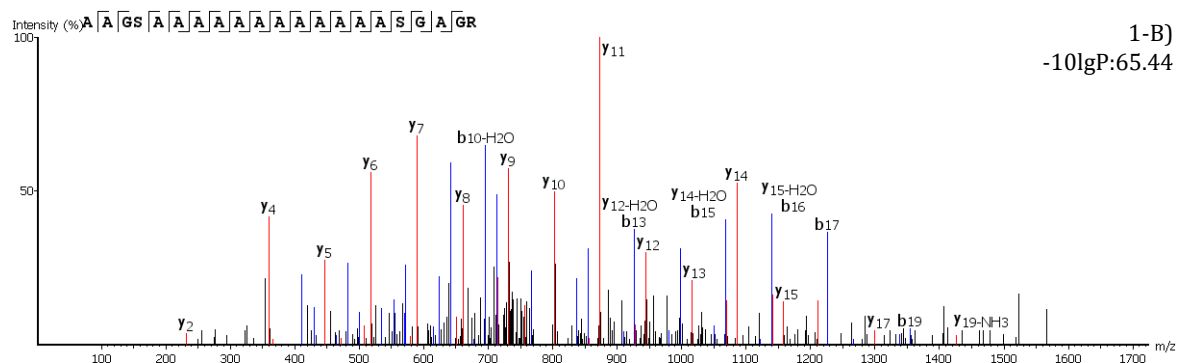
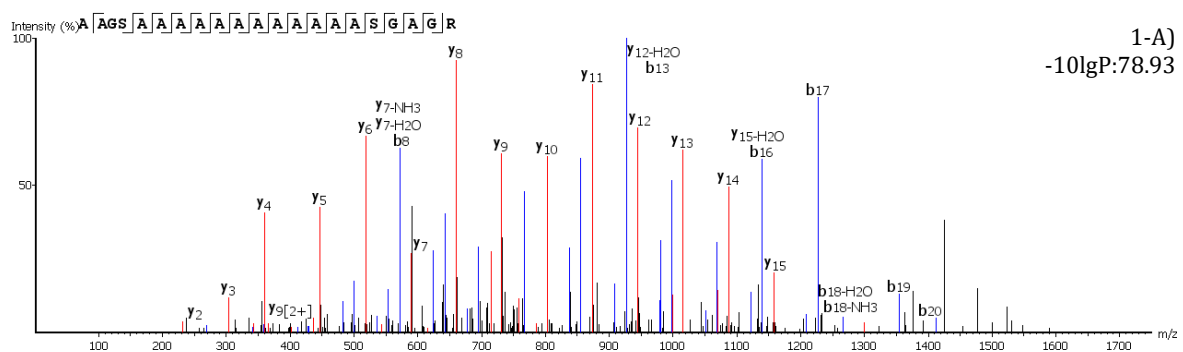
Table 22. Summary of PEAKS analysis results of *Antheraea* and *Samia* silks solubilised with different reagents, in-solution tryptic digestion.

		7M CN	7M CN – FA (10%)
A. pernyi O76786 ANTPE	-10lgP	247.04	252.98
	Coverage (%)	10	16
	#Peptides / #unique	38/8	56/18
A. mylitta Q8ISB3 ANTMY	-10lgP	207.92	189.17
	Coverage (%)	28	35
	#Peptides / #unique	16/5	29/12
A. yamamai E1CGA3 ANTYA	-10lgP	257.44	255.63
	Coverage (%)	9	13
	#Peptides / #unique	44/5	65/13
A. assamensis A0A0K0KR73 ANTASA	-10lgP	307.11	289.38
	Coverage (%)	5	17
	#Peptides / #unique	40/26	81/55
S. ricini A0A0D5ZY13 SAMRI	-10lgP	285.19	366.80
	Coverage (%)	8	22
	#Peptides / #unique	77/76	148/141

⁴⁷¹ Solazzo, *et al.*, 2012

V.4.2.1 *Antheraea pernyi*

In *A. pernyi* silk, CN-FA sample showed 6% higher protein coverage, along with more peptides and unique peptides. There were 38 peptides detected from CN sample and 8 of them were unique to O76786; there were 56 peptides detected from CN-FA sample and 18 of them were unique to O76786. The list of unique peptides is presented in Appendix C (Table 39). The two samples shared 31 peptides and 3 of them were unique to O76786: AAGSA₁₂SGAGR, AAGSA₁₃SGAGGAGGG and SAGSA₁₃SGAGG (Figure 40). The coverage map shows that the latter yielded more peptides from the residues 383-2620, including two *de novo* only tags FHSGYAAAGSGAGKDR (156-162) and M(+15.99)RC(+57.02)YSAAAAAAAA AAC(+57.02)MHR(1530-1541) and 6 unique peptides: GSGSSA₁₀SAAR (544-562), GGRGDGGYGSAGSSA₁₀SA (536-560), GSDSA₁₃SGAR (638-658), SA₁₃GSGAGGRGD (2161-2184), SSA₁₀SGAGGSGGGYGW (2375-2398), and SA₁₄GSGAGGIGGGFGRGD (2407-2436) (Figure 41, Figure 42).



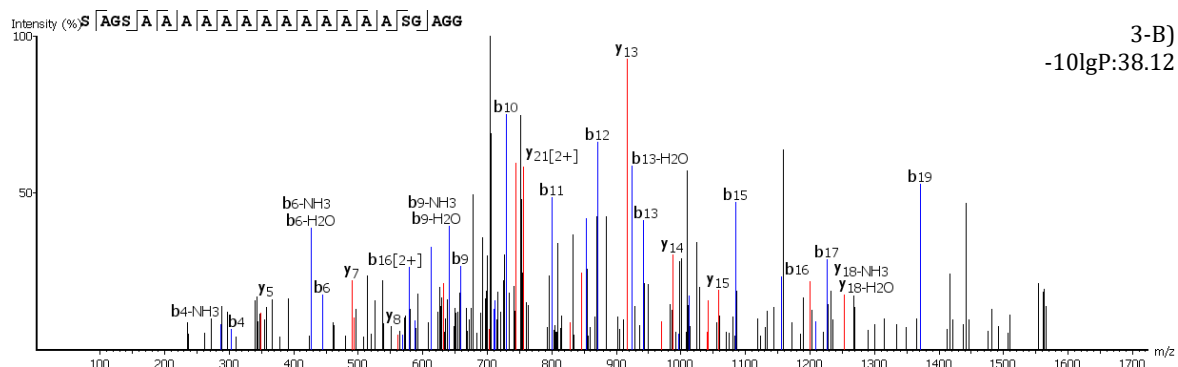
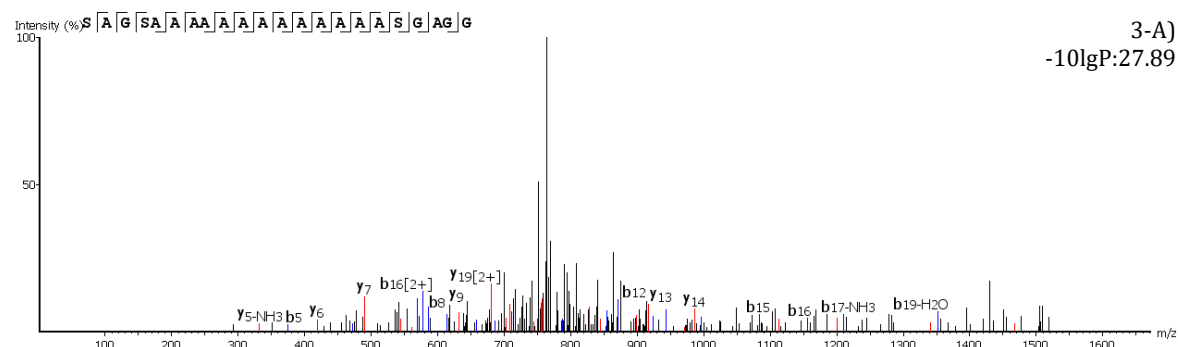
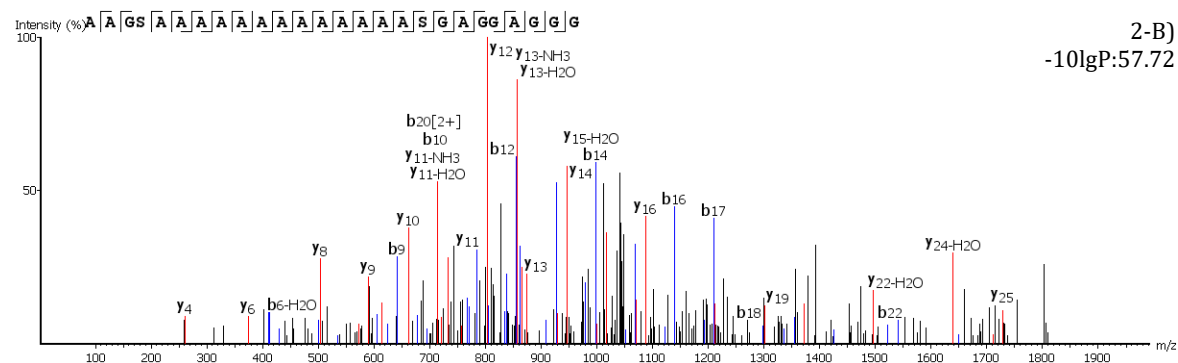
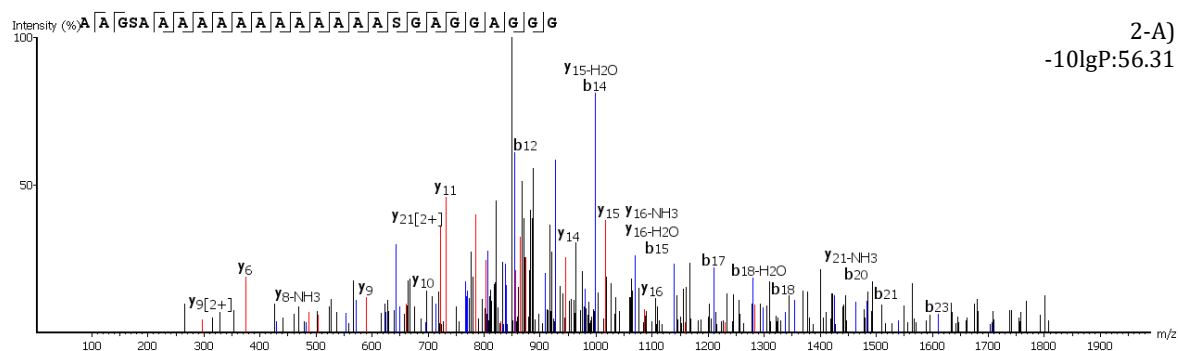


Figure 40. Spectra of the 3 unique peptides to *A. pernyi* fibroin detected in both CN (A) and CN-FA (B) samples: 1) AAGSA₁₂SGAGR (m/z =793.41) 2) AAGSA₁₃SGAGGAGGG (m/z =900) and 3) SAGSA₁₃ SGAGG (m/z = 787.38), $z=2$

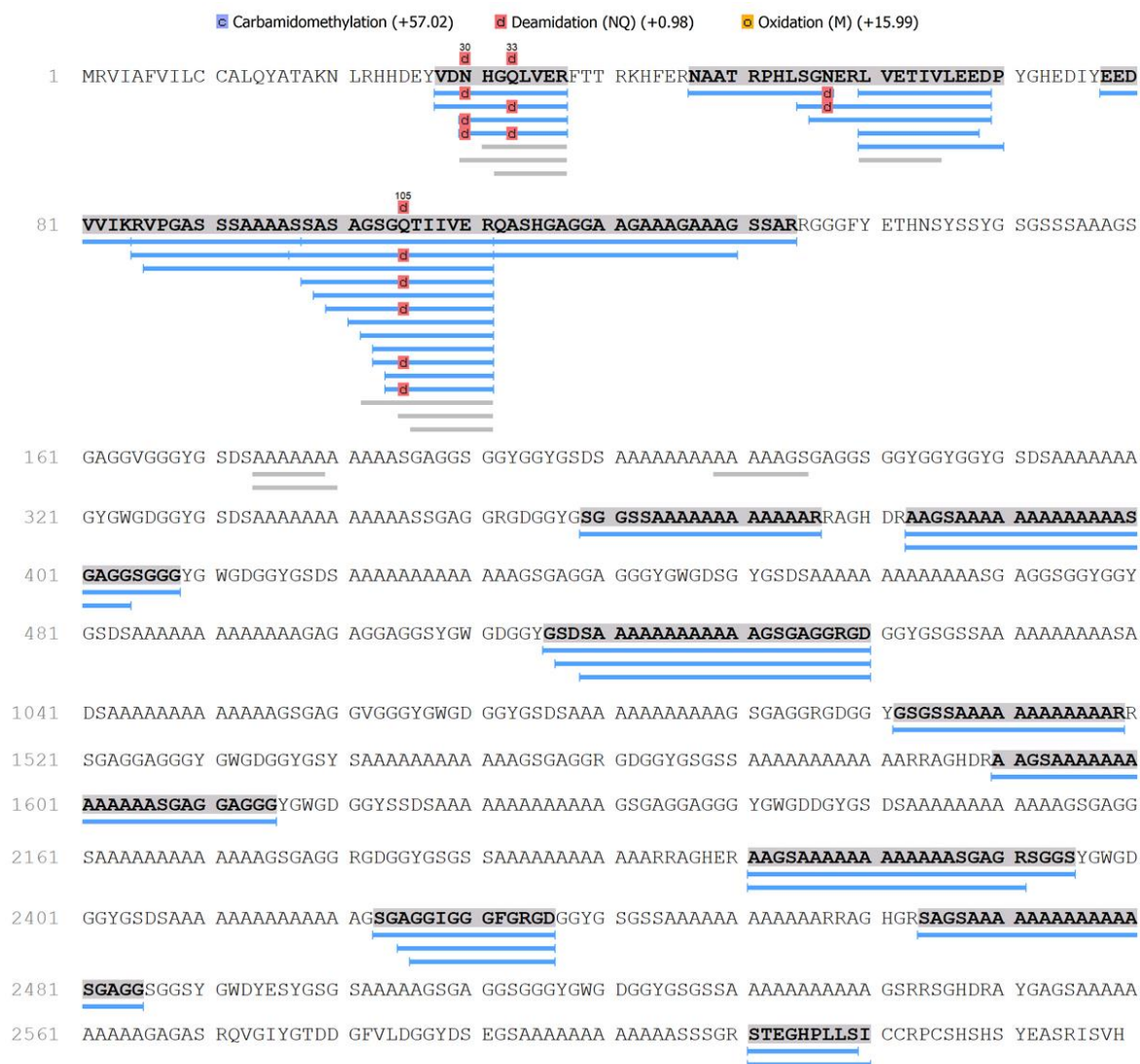


Figure 41. Protein coverage of *A. pernyi* fibroin (076786) - calcium nitrate solution.

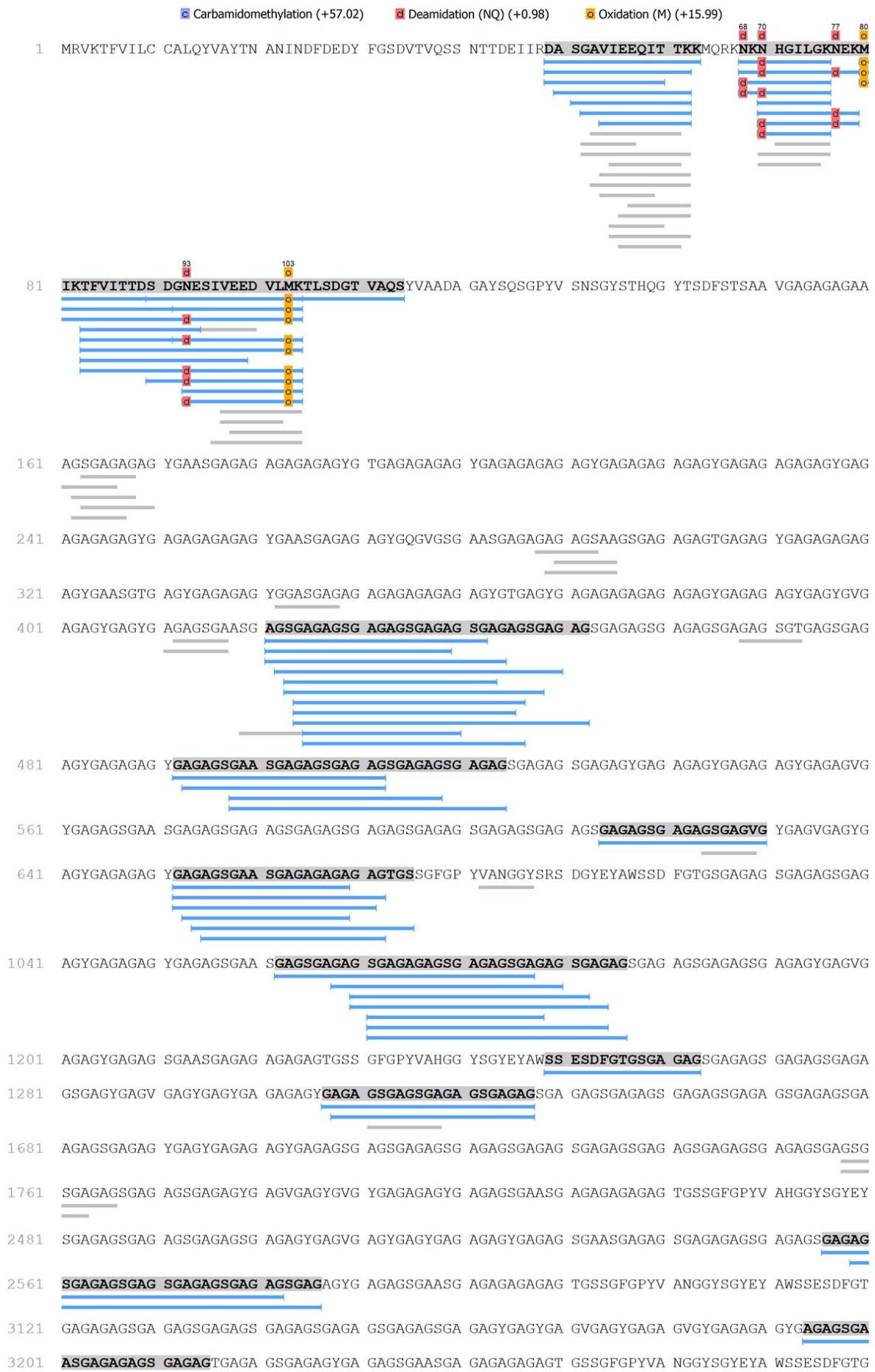




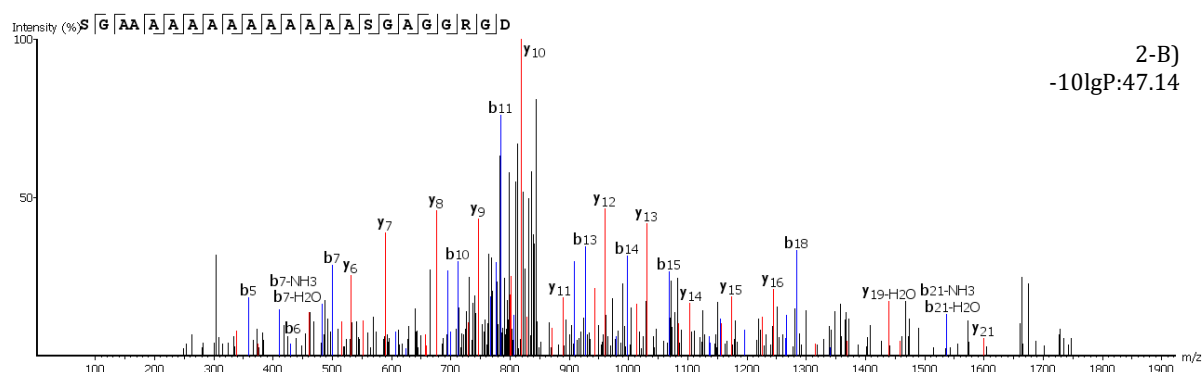
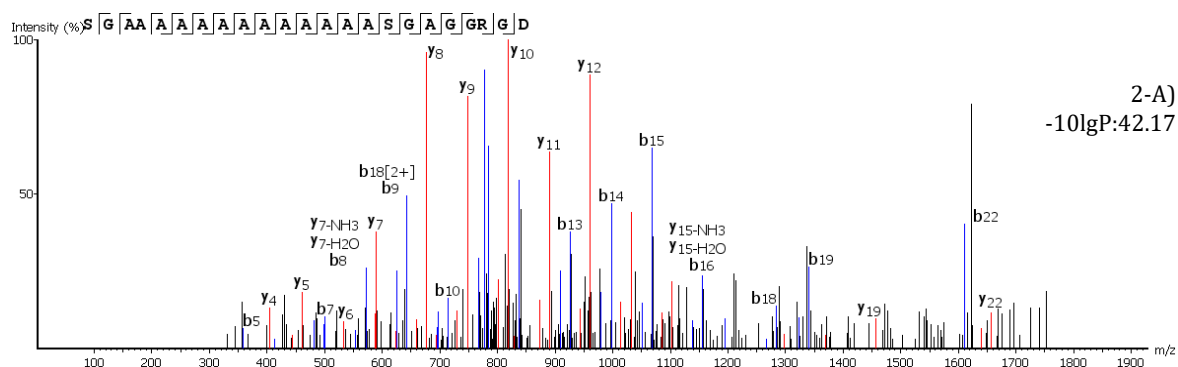
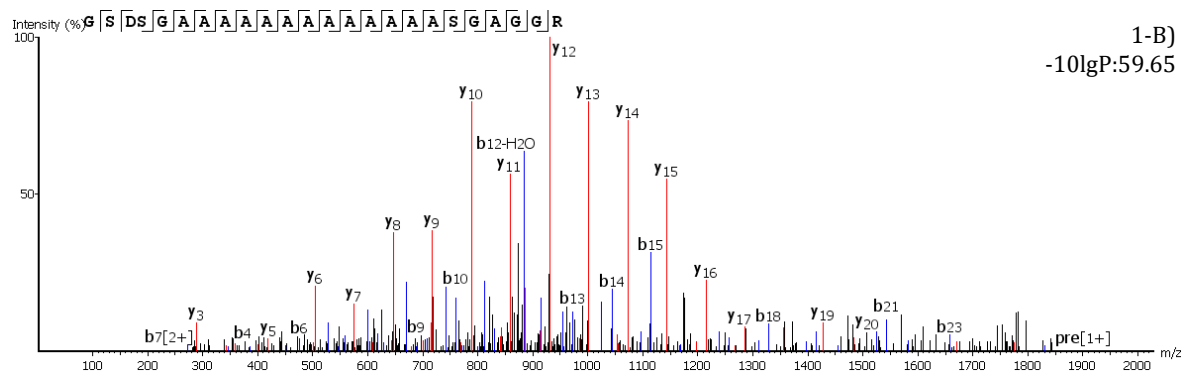
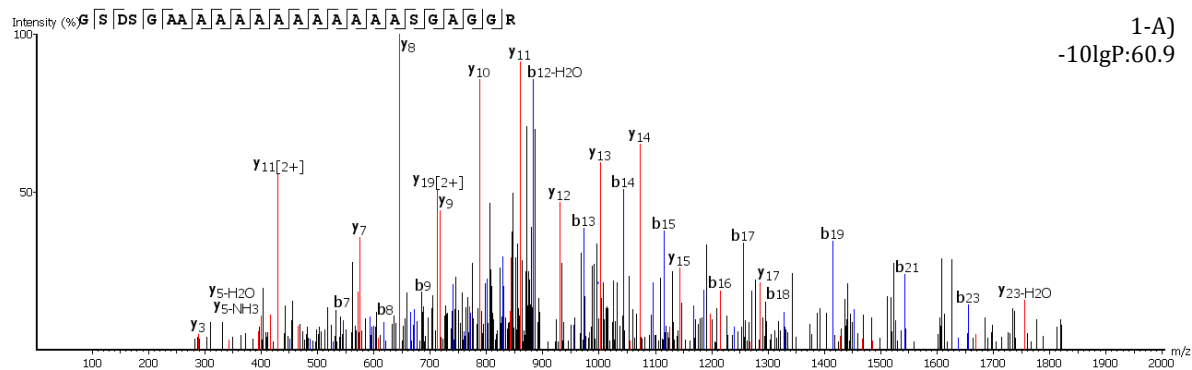
Figure 42. Protein coverage of *A. pernyi* fibroin (076786) - calcium nitrate/ formic acid (10%) solution.

V.4.2.2 *Antheraea mylitta*

In *A. mylitta* silk, CN-FA sample showed higher protein coverage than CN sample and yielded more peptides. However, CN sample showed a slightly higher -10lgP score. The protein identification result lists *A. mylitta* as the first choice in CN sample, the probability scores of *A. yamamai* and *A. pernyi* exceeded that of *A. mylitta* in CN-FA sample (Table 23). This implies that in CN-FA sample, more peptides occurred from the unsequenced part of the protein that is similar to *A. yamamai* and *A. pernyi* fibroin. There were 16 peptides correspond to Q8ISB3 detected in CN sample and 5 of them were unique to the protein. On the other hand, there were 29 peptides correspond to Q8ISB3 detected in CN-FA sample and 12 of them were unique to the protein. Among the unique peptides, 4 peptides GSDSGA₁₃SGAGGR, SGA₁₃SGAGGRGD, GSGSSA₁₁SAAR and VDSHGQLVER were found in both samples (Figure 43). The protein coverage map shows that there were more random cleavages in CN sample, and more *de novo* only tags found in the residues 81-230 in CN-FA sample (Figure 44, Figure 45). The list of unique peptides is presented in Appendix C (Table 40).

Table 23. Protein identification result of *A. mylitta* sample

	Accession	-10lgP	Coverage (%)	#Peptides	#Unique	Avg. Mass
CN	Q8ISB3(A.M)	207.92	28	16	5	45402
	A0A0K0KR73(A.A)	146.5	2	6	2	229493
	A0A0D5ZYI3(S.R)	97.94	0	1	1	227360
CN-FA	E1CGA3(A.Y)	210.51	8	39	3	234574
	O76786(A.P)	205.73	6	36	2	216058
	Q8ISB3(A.M)	189.17	35	29	12	45402
	A0A0K0KR73(A.A)	121.7	2	10	2	229493
	A0A0D5ZYI3(S.R)	86.65	1	2	2	227360



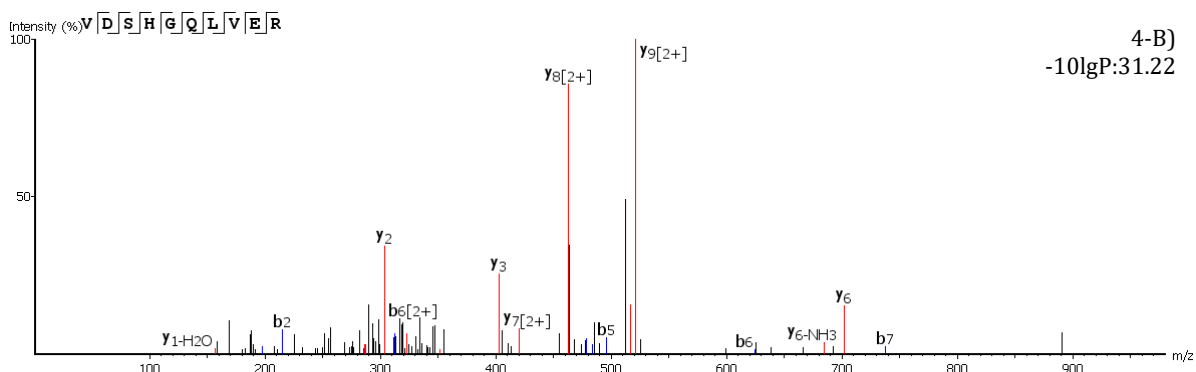
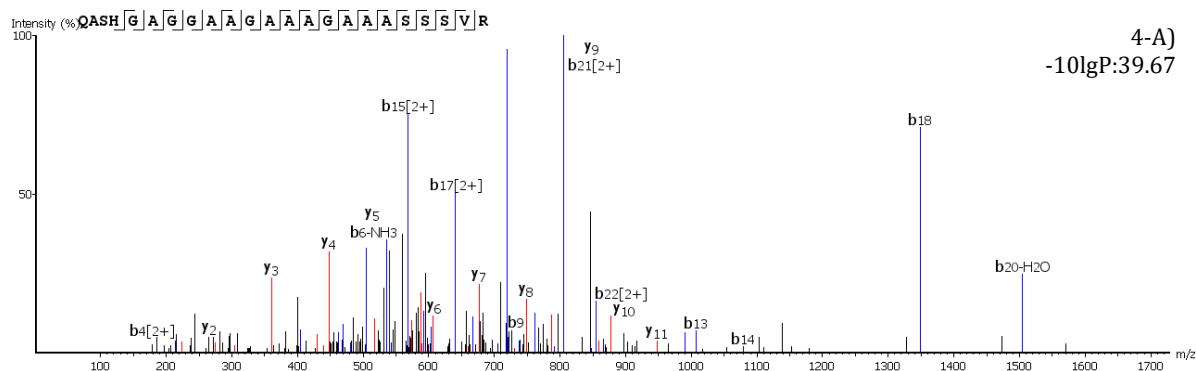
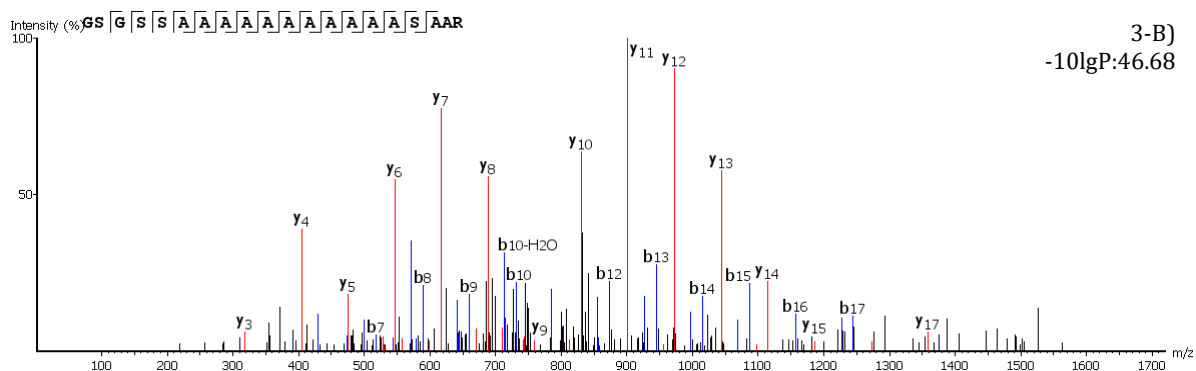
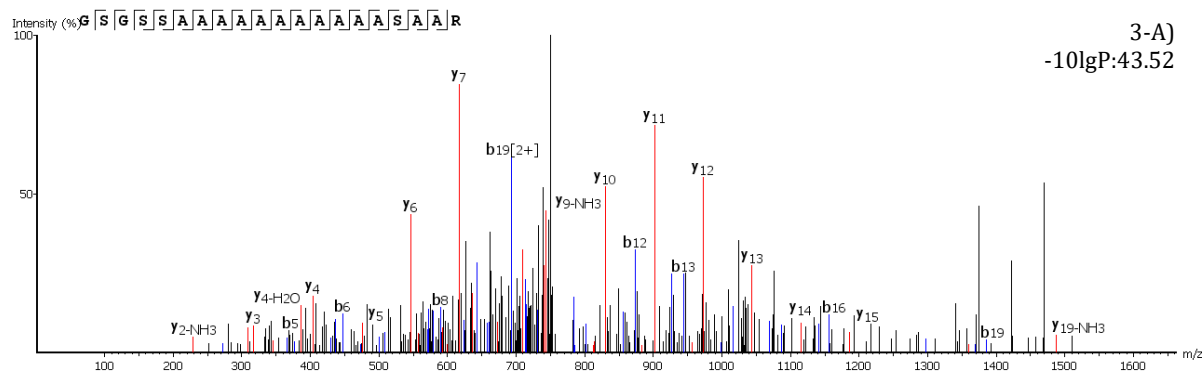


Figure 43. Spectra of the 4 unique peptides to *A. mylitta* fibroin detected in both CN (A) and CN-FA (B) samples: 1) GSDSGA₁₃SGAGGR (m/z = 915.94) 2) SGA₁₃SGAGGRGD (m/z = 872.42) 3) GSGSSA₁₁SAAR (m/z = 780.89), and 4) VDSHGQLVER (m/z = 570.29), $z=2$.

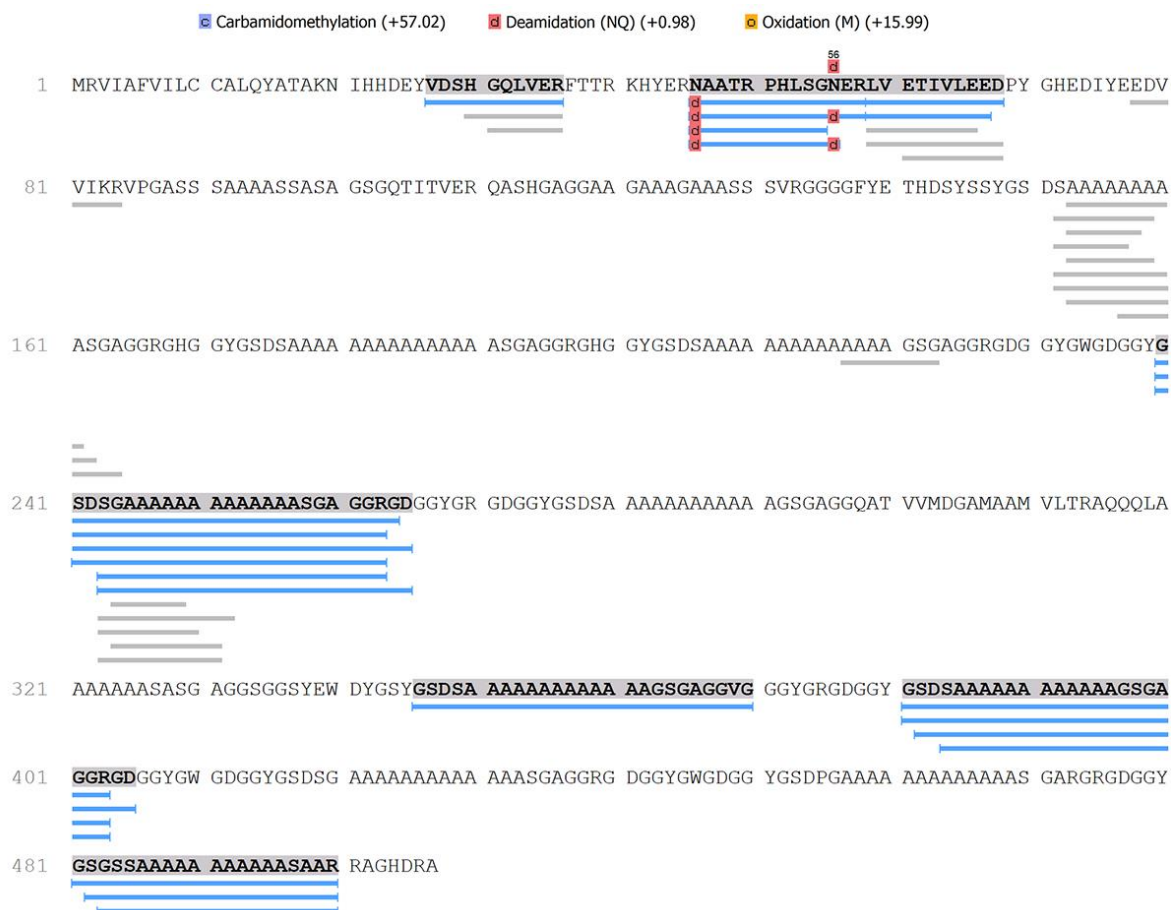


Figure 44. Protein coverage of *A. mylitta* fibroin (Q8ISB3)- calcium nitrate solution.

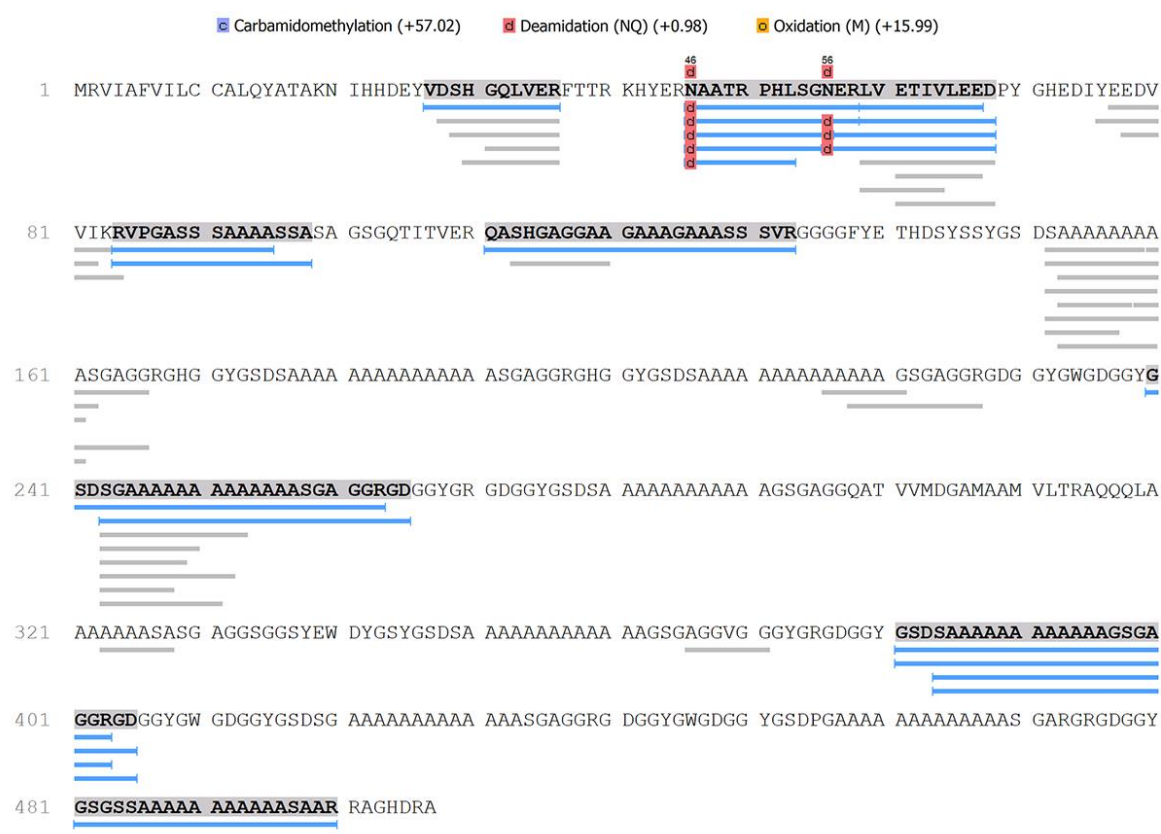
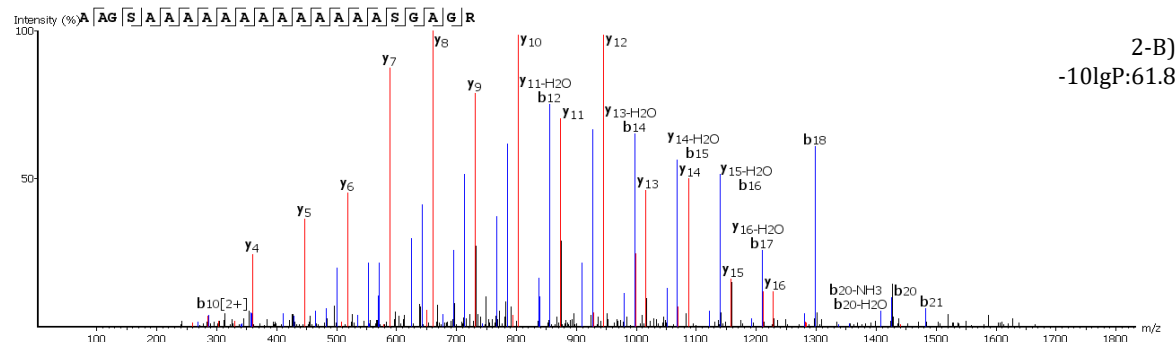
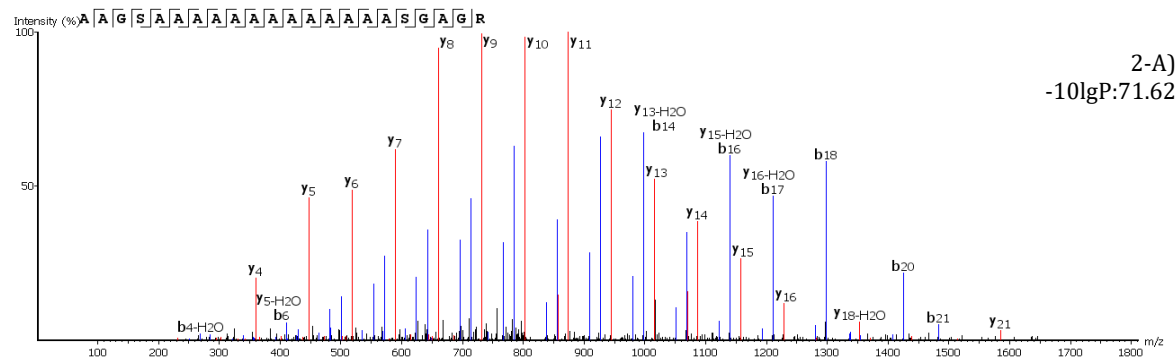
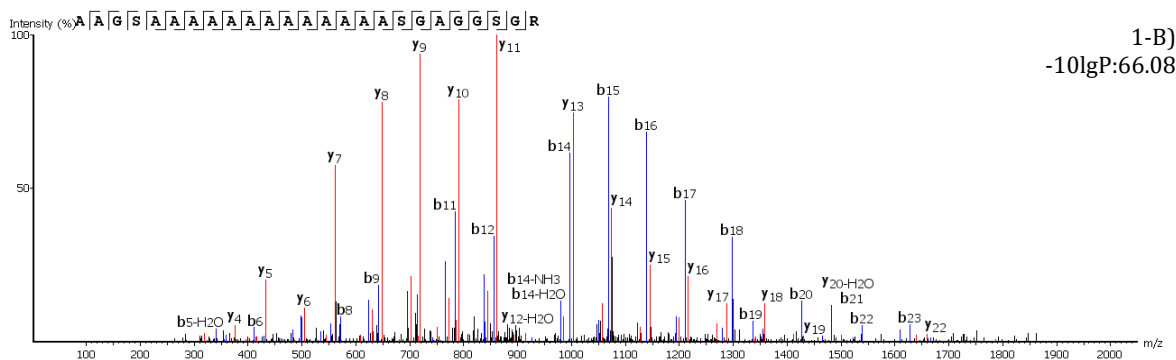
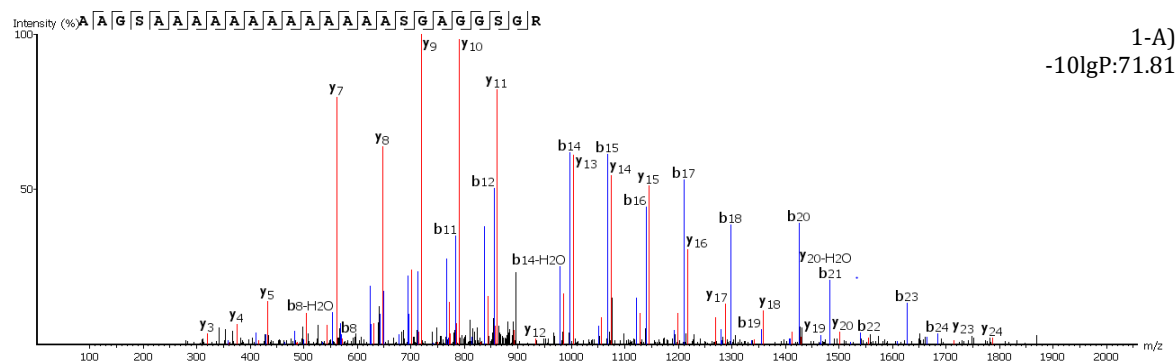


Figure 45. Protein coverage of *A. mylitta* fibroin (Q8ISB3) - calcium nitrate/ formic acid (10%) solution.

V.4.2.3 *Antheraea yamamai*

In *A. yamamai* silk, CN-FA sample showed 4% higher protein coverage, along with more peptides and unique peptides. There were 44 peptides detected from CN sample and 5 of them were unique to E1CGA3; there were 65 peptides detected from CN-FA sample and 13 of them were unique to E1CGA3. There were 30 peptides shared among the *Antheraea* fibroins and 4 unique peptides to E1CGA3 detected from both CN and CN-FA samples: EDVVINR, EEDVVINR, AAGSA₁₃SGAGR and AAGSA₁₃SGAGGSGR (Figure 46). The coverage map shows that CN-FA sample yielded more peptides from the residues 400-900 region, including two unique peptides GSDSA₁₃GSGAGR (486-508) and A_{7/8}GSGAGGAGGD (648/9-665) (Figure 47, Figure 48). However, the repetitive polyalanine block region between the residues 933-1880 did not yield any peptides in either sample. This is thought to be from the limitation of Trypsin. Table 24 shows the theoretical tryptic peptides of E1CGA3 between the residues 933-2009. There is a pattern of a set comprised of a very big peptide with 4-5 polyalanine blocks followed by GDGGYGS₁₂SSA₁₂R, R, and AGHDR. GDGGYGS₁₂SSA₁₂R and AGHDR are common among the *Antheraea* silk fibroins and not exclusive to one fibroin. The big block peptides are too large to be analysed by the mass spectrometer. Thus, unless the protease can effectively cleave the big blocks and yield some unique peptide to the subject protein, it would not affect the result of the current protein identification result. However, there was one unique peptide to E1CGA3 occurred from this region, GSDSAAGA₁₀SGAGGGRG (1987-2010) in both CN and CN-FA samples: this peptide was probably occurred by a random cleavage happened during the solubilisation, and one missed tryptic cleavage at the N-terminus. The list of unique peptides is presented in Appendix C (Table 41).



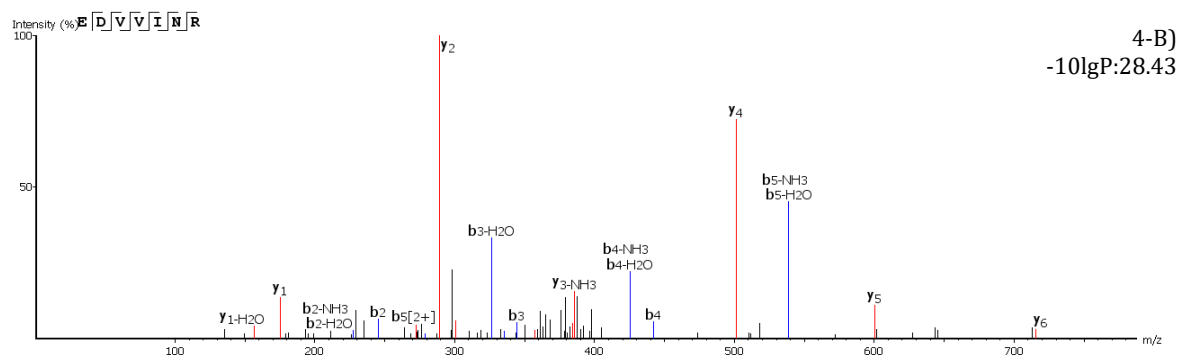
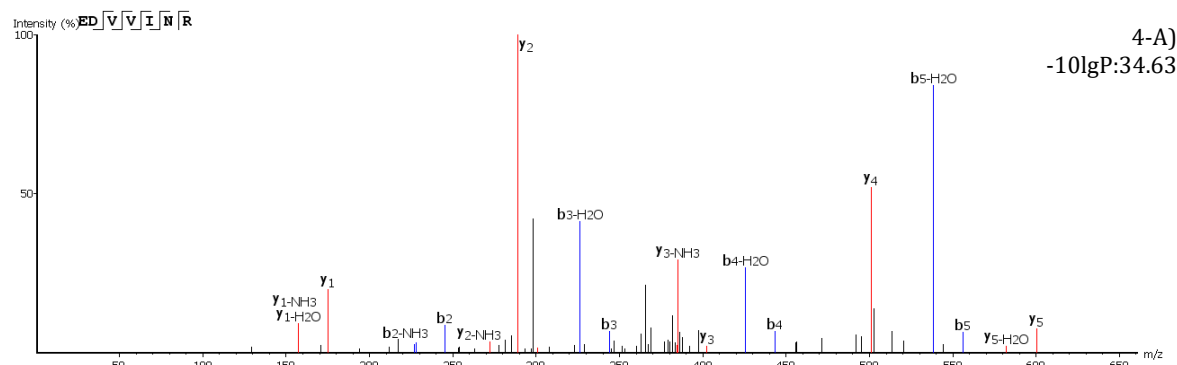
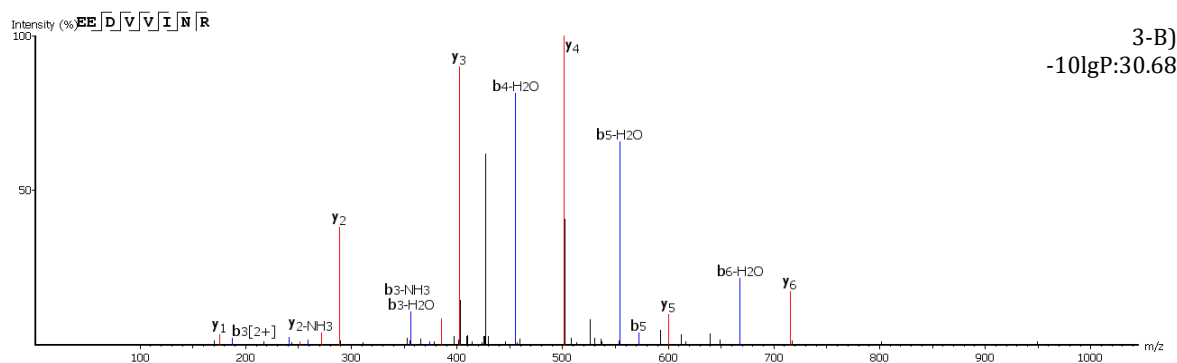
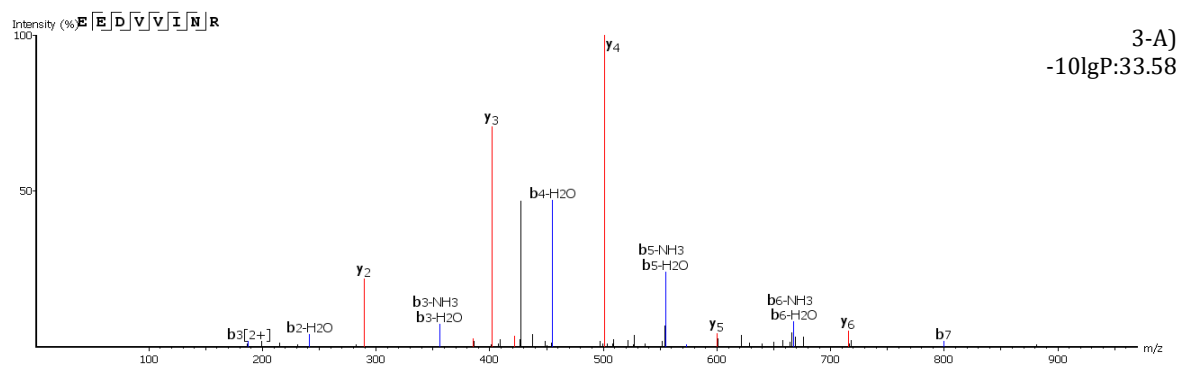


Figure 46. Spectra of the 4 unique peptides detected in both CN (A) and CN-FA (B) samples:
1) AAGSA13SGAGGSGR ($m/z=929.46$), 2) AAGSA13SGAGR ($m/z=828.92$), 3) EEDVVINR ($m/z=487.25$) and 4) EDVVINR ($m/z=422.73$), $z=2$.



Figure 47. Protein coverage of *A. yamamai* fibroin (E1CGA3) - calcium nitrate solution.

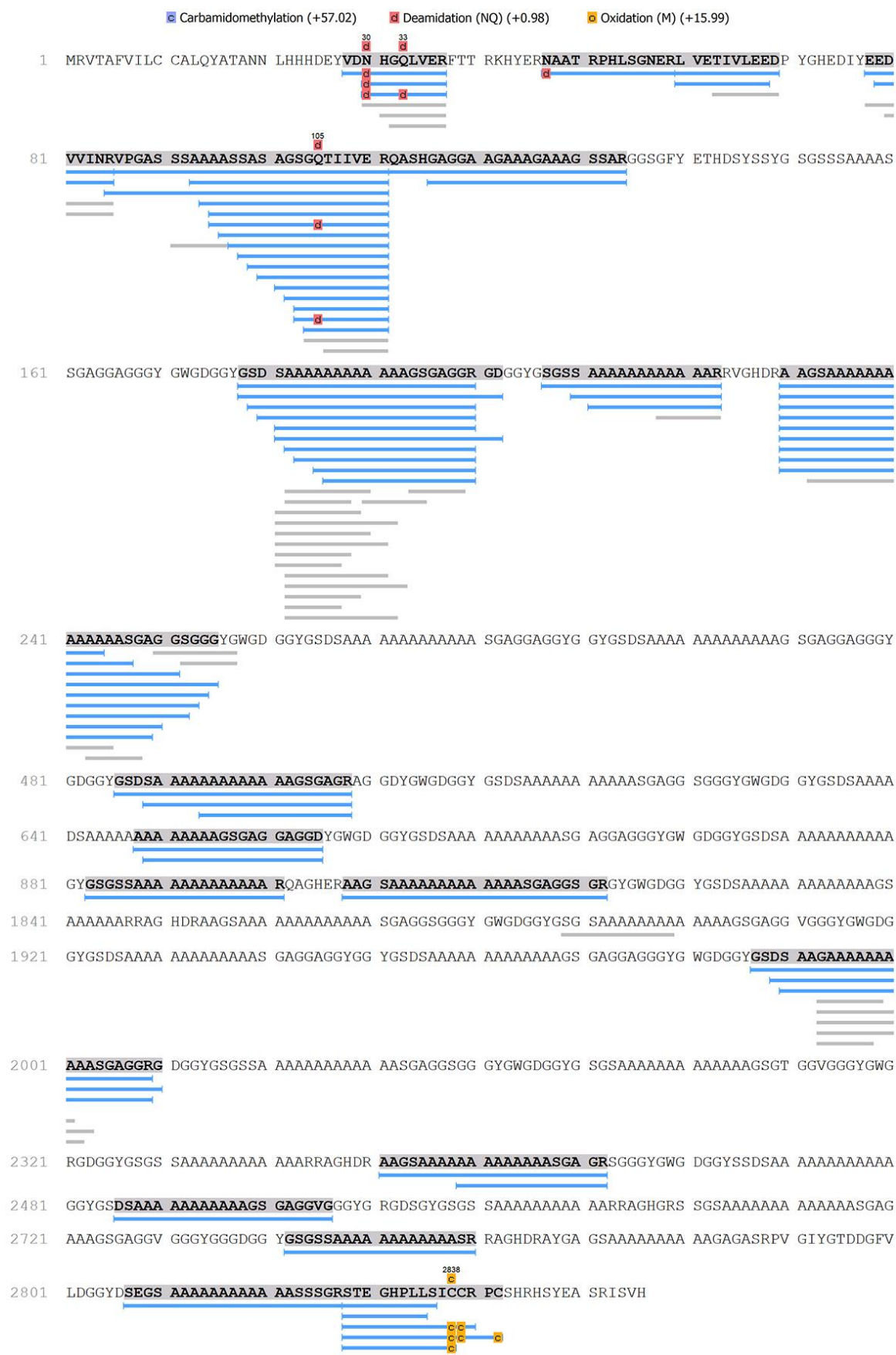


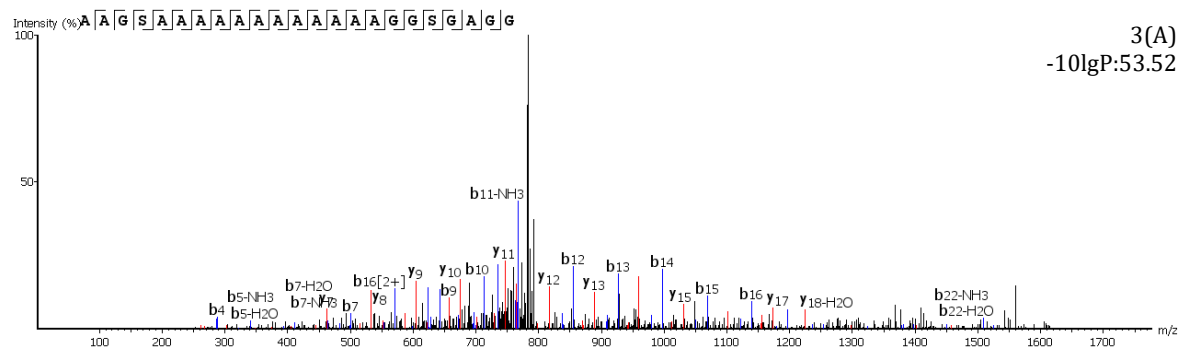
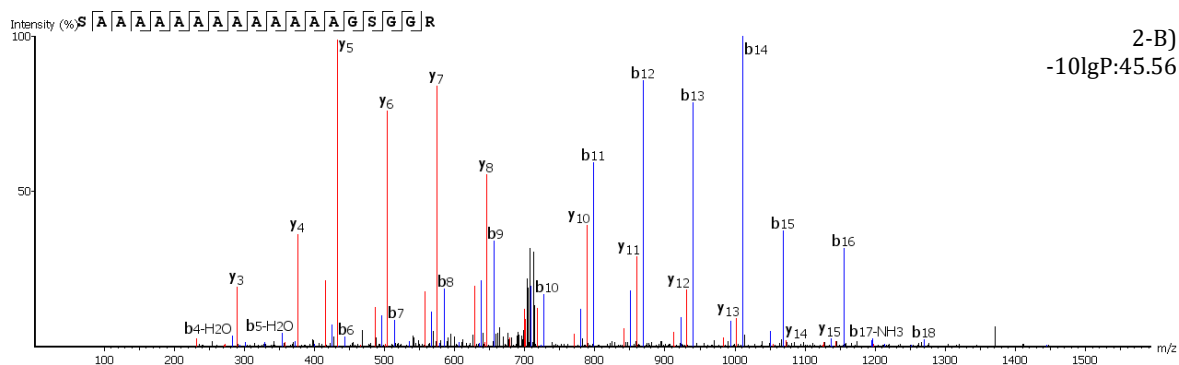
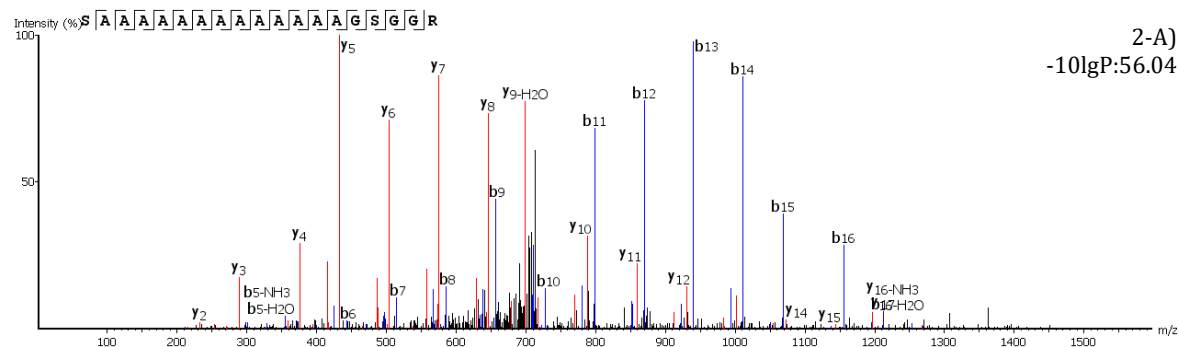
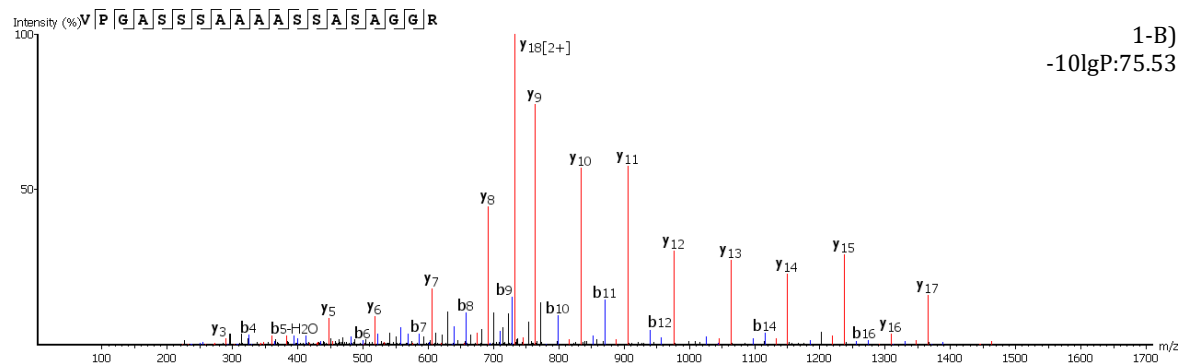
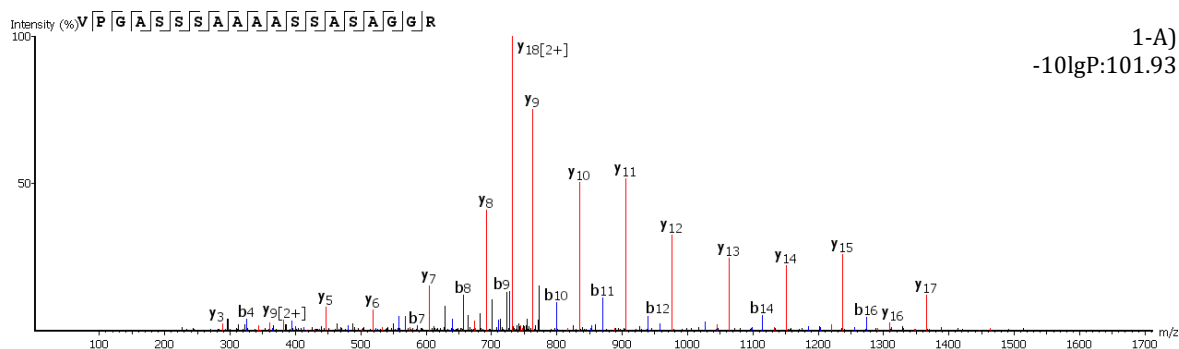
Figure 48. Protein coverage of *A. yamamai* fibroin (E1CGA3) - calcium nitrate/ formic acid (10%) solution.

Table 24. Theoretical tryptic cleavages of E1CGA3, residues 933-2009 (missed cleavage:0)

location	sequence	mass
933-1031	GYGWGDGGYGSDSA13GSGAGGAGGDYGWGDGGYGSDSA11SGAGSGGGYGWGDGGYGSDSA12GSGAGGR	8053.42
1032-1054	GDGGYSGSSA12R	1851.86
1055-1055	R	175.11
1056-1060	AGHDR	555.26
1061-1184	AAGSA13SGAGSGGGYGWGDGGYGSDSA13GSGAGGAGGDYGWGDGGYGSDSA11SGAGGAGGGYGWGDGGYGS SDSA12GSG AGGR	9807.25
1185-1207	GDGGYSGSSA12R	1851.86
1208-1208	R	175.12
1209-1213	AGHDR	555.26
1214-1333	AAGSA13SGAGSGGGYGWGDGGYGSDSA12SGAGGAGGGYGSDSA13GSGAGGAGGGYGWGDGGYGSD SA12GSGAGGR	9318.10
1334-1356	GDGGYSGSSA12R	1851.86
1357-1357	R	175.12
1358-1362	AGHDR	555.26
1363-1486	AAGSA13SGAGSGGGYGWGDGGYGSDSA13GSGAGGAGGDYGWGDGGYGSDSA11SGAGGSGGGYGWGDGG YGS SDSA12GSG AGGR	9793.24
1487-1509	GDGGYSGSSA12R	1851.86
1510-1510	R	175.12
1511-1515	AGHDR	555.26
1516-1639	AAGSA13SGAGSGGGYGWGDGGYGSDSA13GSGAGGAGGDYGWGDGGYGSDSA11SGAGGAGGGYGWGDGG YGS SDSA12GSG AGGR	9777.24
1640-1662	GDGGYSGSSA12R	1851.86
1663-1663	R	175.12
1664-1668	AGHDR	555.26
1669-1824	AAGSA13SGAGSGGGYGWGDGGYGSDSA13GSGAGGVGGGYGWGDGGYGSDSA13SGAGGAGGGYGSDSA13SSGAGGAGGGYGWGDGGYGSDSA12GSGAGGR	12200.35
1825-1847	GDGGYSGSSA12R	1851.86
1848-1848	R	175.12
1849-1853	AGHDR	555.26
1854-2009	AAGSA13SGAGSGGGYGWGDGGYGSDSA13GSGAGGVGGGYGWGDGGYGSDSA13SGAGGAGGGYGSDS A13GSGAGGAGGGYGWGDGGYGSDSAAGA10SGAGGR	12112.34

V.4.2.4 *Antheraea assamensis*

In *A. assamensis* silk, both CN and CN-FA sample were identified as *A. assamensis* fibroin with high probability score. CN sample showed only 5% sequence coverage with 26 unique peptides to A0A0K0KR73; however, CN-FA sample showed 17% coverage with 55 unique peptides to A0A0K0KR73. The high number of unique peptides compare to other *Antheraea* and *Samia* silk samples is in accordance with the sequence feature analysis result reviewed in V.2.2.2. CN sample showed 14 peptides shared with *A. pernyi* fibroin and 26 unique peptides to A0A0K0KR73, and 20 of them were also detected in CN-FA sample. Most of these 20 unique peptides occurred from near C-and N-terminal, except the two polyalanine peptides AAGSA₁₂GGSGAGG (771-788) (AAGSA₁₂GGSGAGGHGGG at 771-797 in CN-FA) and SA₁₃GSGGR (1128-1146). Among the unique peptides, VPGASSSAAAASSASAGGR (86-104) showed the highest probability score: -10lgP score of 101.93 in CN sample and -10lgP score of 75.53 in CN-FA sample (Figure 49). The coverage map shows that CN-FA sample yielded more peptides from the residues 300-2700 region, and most of them were unique to A0A0K0KR73 (Figure 50, Figure 51). The list of unique peptides is presented in Appendix C (Table 42).



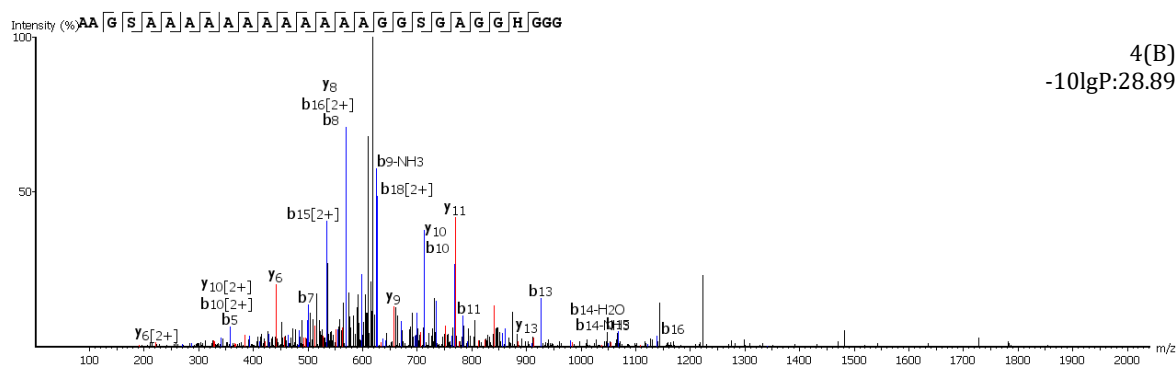


Figure 49. Spectra of the unique peptides detected in both CN (A) and CN-FA (B) samples: 1) VPGASSSAAAASSASAGGR ($m/z=781.38$, $z=2$), 2) SA₁₃GSGGR ($m/z=722.37$, $z=2$), 3) AAGSA₁₂GGSGAGG ($m/z=800.88$, $z=2$) and 4) AAGSA₁₂GGSGAGGHGGG ($m/z=636.97$, $z=3$).

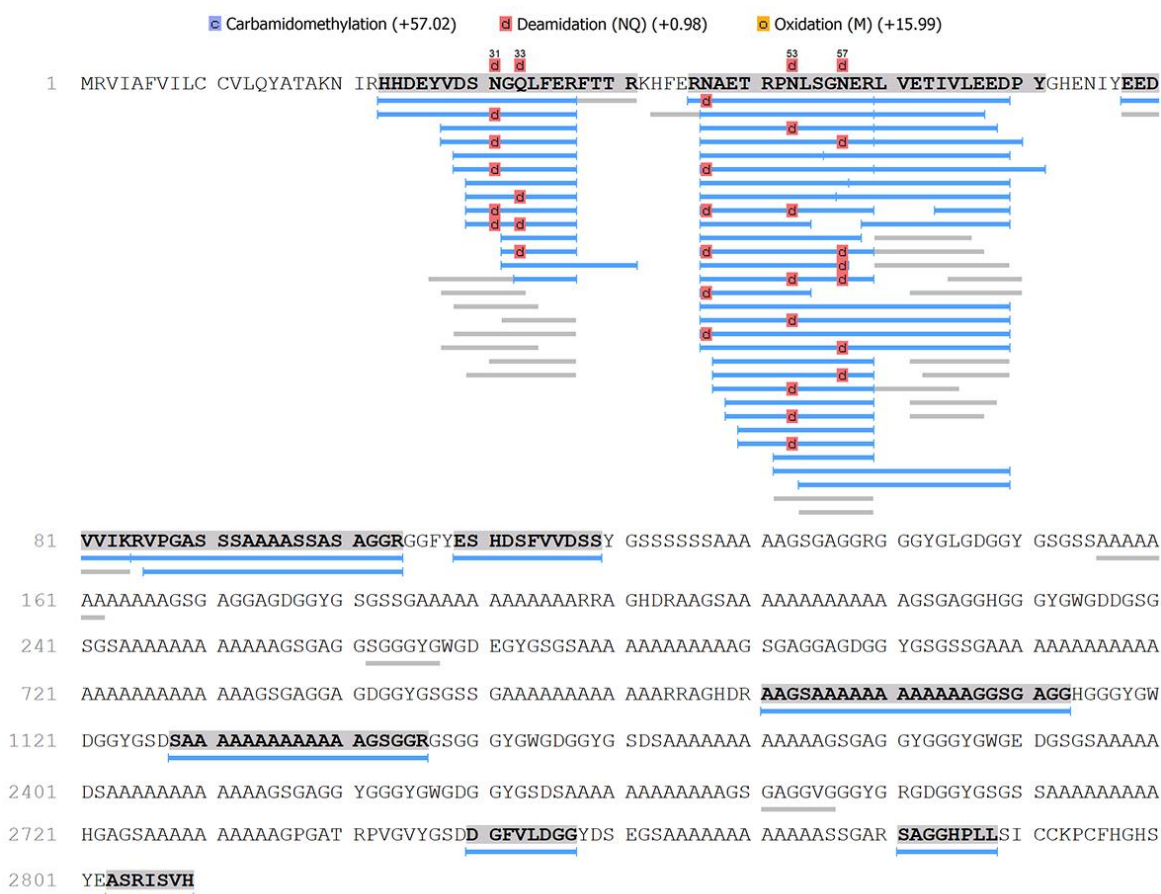


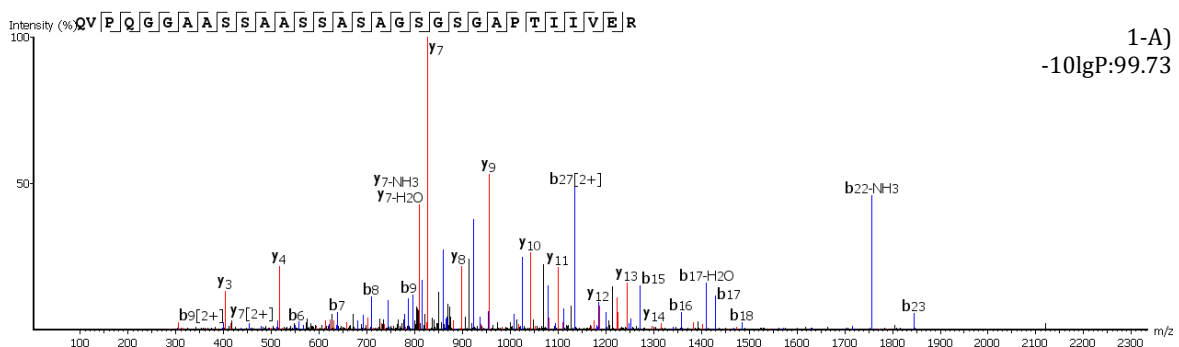
Figure 50. Protein coverage of *A. assamensis* fibroin (A0A0K0KR73) - calcium nitrate solution.



Figure 51. Protein coverage of *A. assamensis* fibroin (A0A0K0KR73) - calcium nitrate /formic acid (10%) solution.

V.4.2.5 *Samia ricini*

In *S. ricini* silk, CN sample yielded 4% sequence coverage with 56 peptides including 53 unique peptides to A0A0D5ZYI3. CN-FA sample yielded 5 times more sequence coverage (20%) with 111 peptides including 105 unique peptides to A0A0D5ZYI3. Among the unique peptides to the subject protein, 19 unique peptides were detected in both samples, mostly occurred from the C-and N-terminus of the protein sequence (Figure 53, Figure 54). In CN sample, QVPQGGAASSAASSASAGSGSGAPTIIVER (101-130) was the highest scoring peptide followed by VHETGDGFLLR (2822-2832). In CN-FA sample, peptides occurred all over the protein sequence except the residues 541-800 and 1220-2385. In CN-FA sample, SSAASSASAGSGSGAPTIIVER (109-130) was the highest scoring peptide, followed by QVPQGGAASSAASSASAGSGSGAPTIIVER (101-130); VHETGDGFLLR (2822-2832) was also abundant but there were other peptides with higher $-10\lg P$ scores occurred close to the C-and N-terminus of the protein sequence (Figure 52). The list of unique peptides is presented in Appendix B (Table 43).



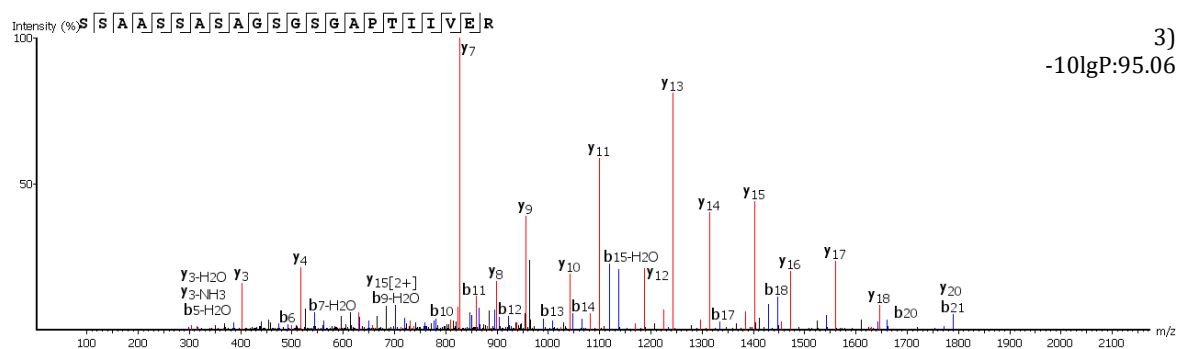
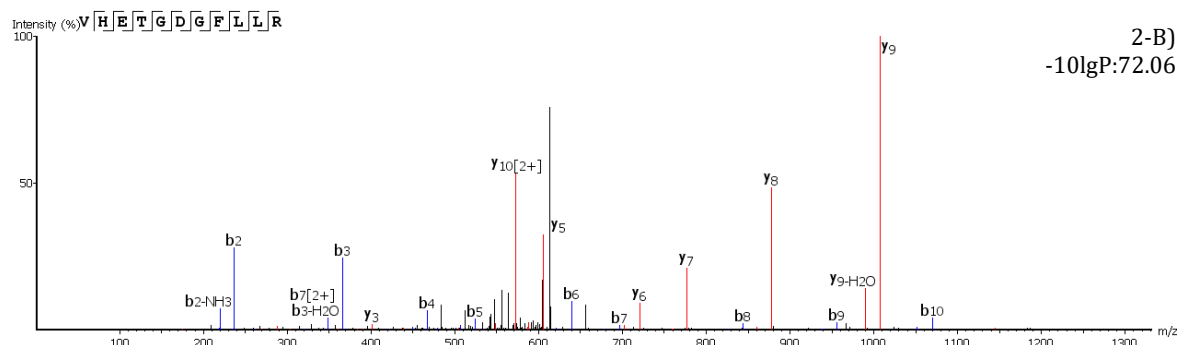
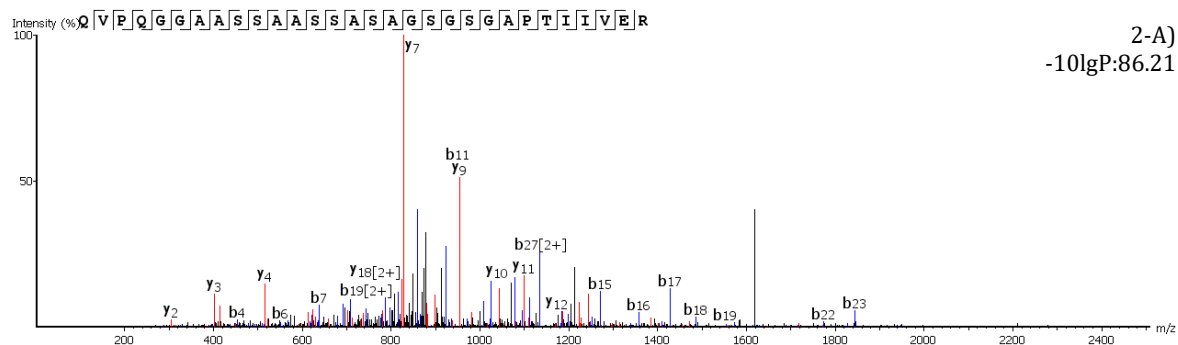
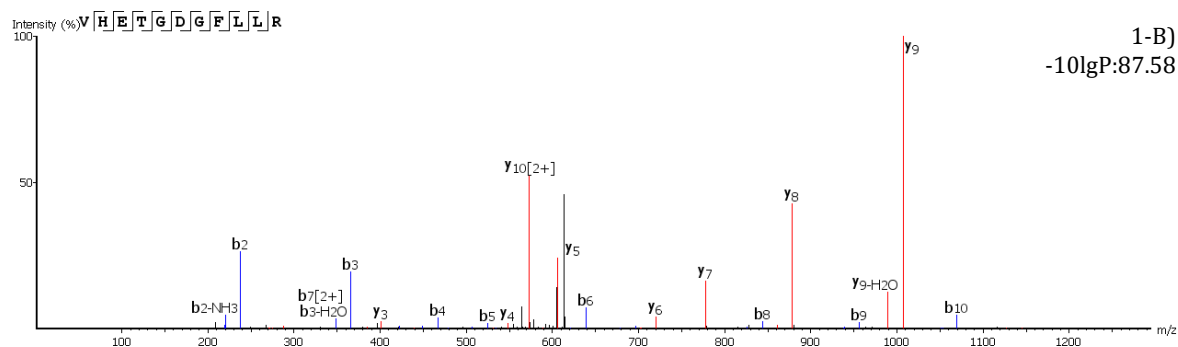


Figure 52. Spectra of the unique peptides detected in CN (A) and CN-FA (B) samples: 1) QVPQGGAASSAASSASAGSGSGAPTIIVER (m/z=891.11), 2) VHETGDGFLLR (m/z=622.32), 3) SSAASSASAGSGSGAPTIIVER (m/z=981.99), z=2.

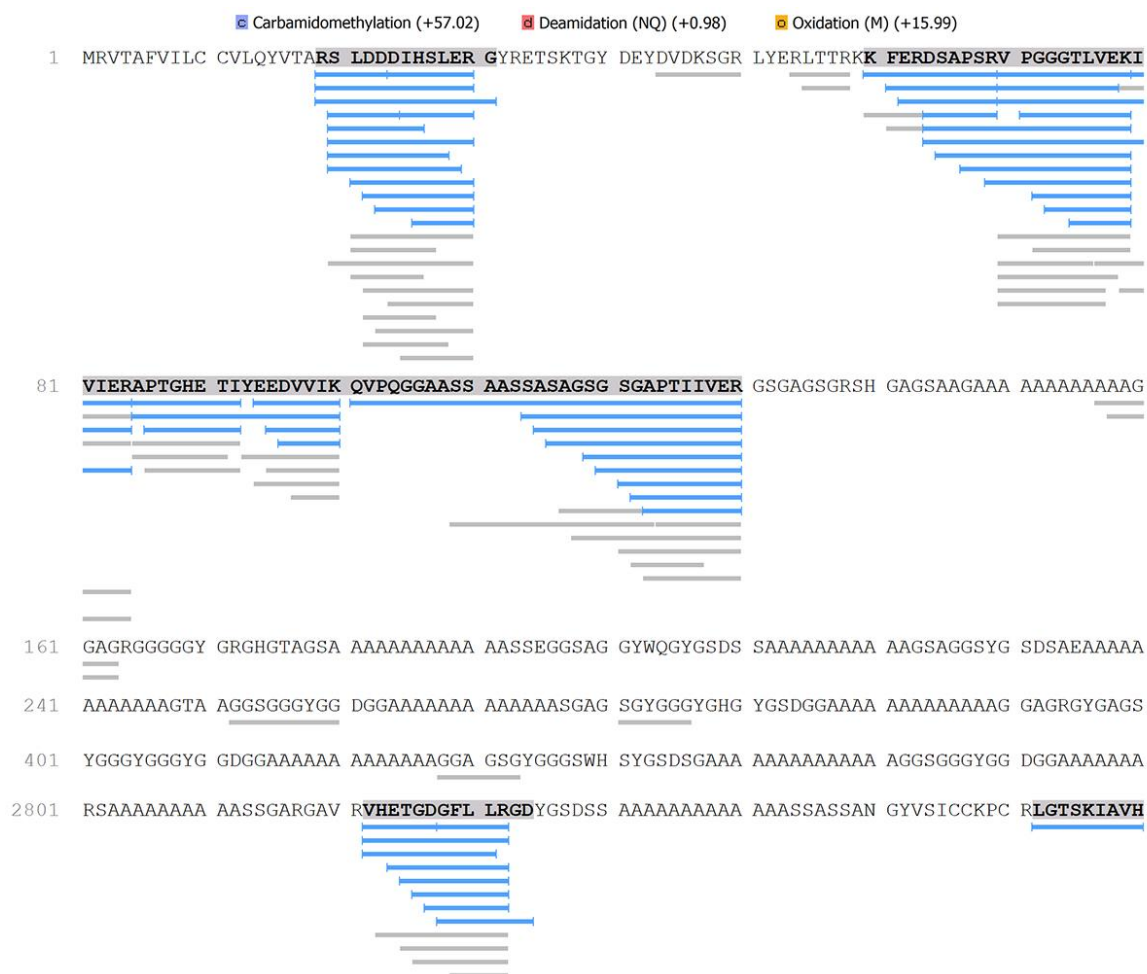


Figure 53. Protein coverage of *S. ricini* fibroin (A0A0D5ZYI3) - calcium nitrate solution.





Figure 54. Protein coverage of *S. ricini* fibroin (A0A0D5ZYI3) - calcium nitrate/formic acid (10%) solution.

V.5 Effect of enzymes

The theoretical projection of enzymatic digestion suggests that trypsin and chymotrypsin used together would be the most appropriate choice of enzyme for the mass spectrometry of silk fibroins. However, the SDS-PAGE result showed that the proteins were degraded into polypeptides by the solubilisation process. Thus to understand their actual behaviour to each enzyme, fibroin solutions were digested with trypsin, chymotrypsin, and 1:1 mixture of the two; and their MS/MS results analysed by PEAKS is compared. For convenience, enzymes are abbreviated to T (trypsin), C (chymotrypsin) and TC (trypsin-chymotrypsin); in-gel digestion is abbreviated to G and in-solution digestion is abbreviated to S.

V.5.1 Bombyx silks

A summary of PEAKS analysis results is presented in Table 25. As predicted in V.2.3, the experimental results also showed that the mixture of trypsin-chymotrypsin yielded better protein coverages than chymotrypsin or trypsin used alone; chymotrypsin was more effective than trypsin to prepare the *Bombyx* fibroins for the mass spectrometry purpose.

In the *Bombyx mori* sample, the best protein identification result for the heavy chain was found in TCG. Tryptic-chymotryptic digestion of *B. mori* silk fibroin yielded 356 peptides exclusive to P05790 that achieved 25% sequence coverage. Although there is room for improvement, considering the large molecular size of the *B. mori* fibroin's heavy chain (5263 residues), 25% protein coverage can be considered good. However, higher the coverage of the heavy chain achieved, lower the coverage of the light chain was achieved. The probability scores of P25 protein in tryptic-chymotryptic and CS were too low thus

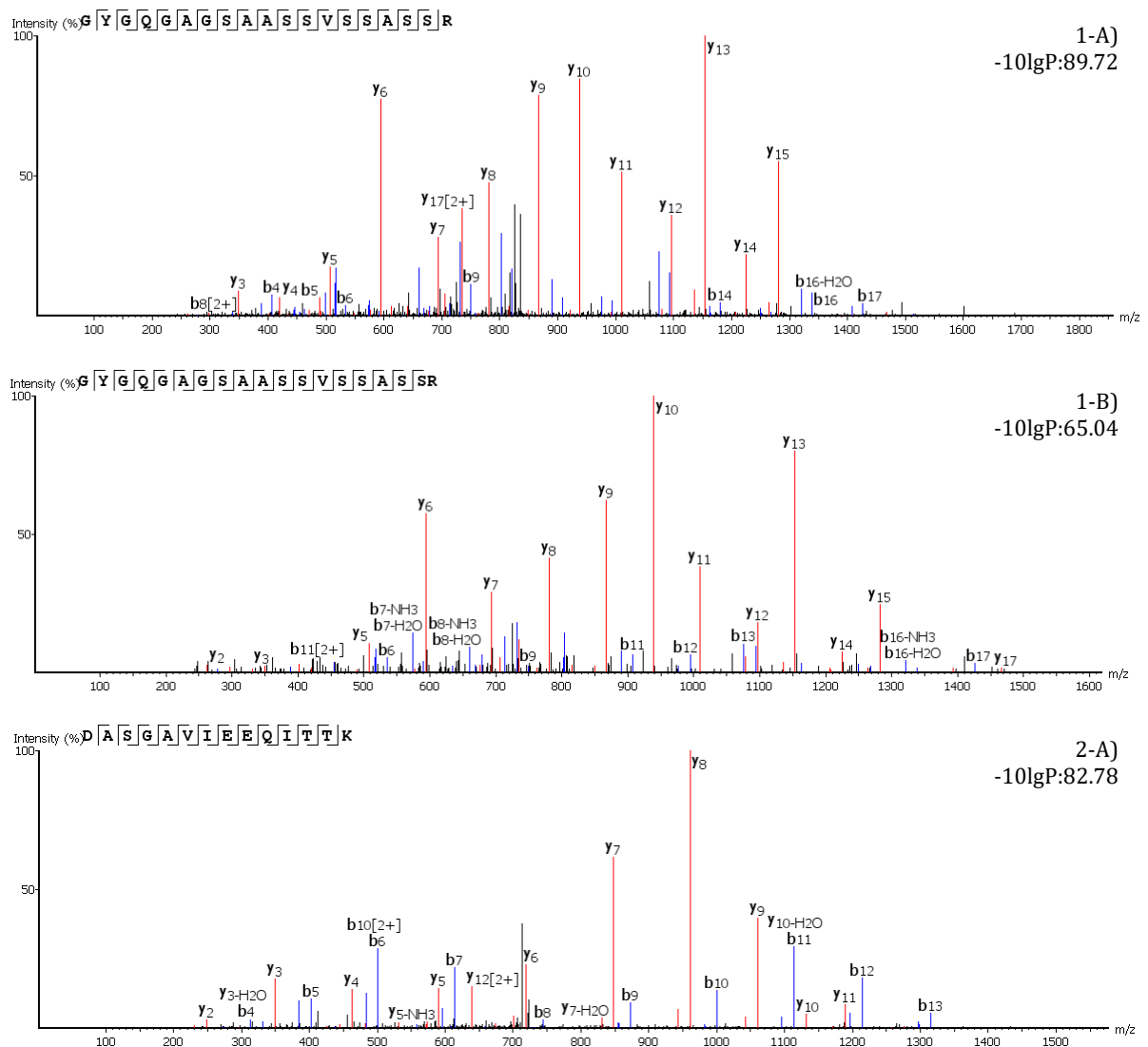
screened out (-10lgP >30). The decrease in the light chain and P25 protein coverages according to the increase in the heavy chain coverage indicates that these smaller proteins (262, 220 residues respectively) are more likely to be damaged by the condition that effectively denatures the heavy chain protein.

Table 25. PEAKS analysis results of *Bombyx* silks treated with different enzymes

<i>Bombyx mori</i>	Trypsin		Chymotrypsin		Try-Chymo	
P05790 FIBH_BOMMO	In-gel	In-sol	In-gel	In-sol	In-gel	In-sol
-10lgP	189.77	280.11	288.58	270.53	216.28	192.48
Coverage (%)	2	5	11	7	25	21
#Peptides/#Unique	10/8	81/59	219/165	62/46	356/324	153/131
P21828 FIBL_BOMMO (Q9BLL9 FIBL_BOMMA)						
-10lgP	300.43	305.82	219.55	184.41	124.78	39.86
Coverage (%)	85	93	82	70	59	8
#Peptides/#Unique	61/61	147/147	118/118	33/33	48/48	2/2
P04148 P25_BOMMO						
-10lgP	211.39	167.86	75.87	*	-	-
Coverage (%)	60	68	16	-	-	-
#Peptides/#Unique	23/23	28/28	7/7	-	-	-
<i>Bombyx mandarina</i>						
P05790 FIBH_BOMMO	In-gel	In-sol	In-gel	In-sol	In-gel	In-sol
-10lgP	121.86	241.86	284.02	224.01	197.62	144.25
Coverage (%)	0	1	6	5	7	10
#Peptides/#Unique	1/1	11/11	103/103	27/18	55/34	39/39
Q9BLL9 FIBL_BOMMA						
-10lgP	274.35	198.78	236.89	197.78	140.90	96.16
Coverage (%)	38	38	65	64	53	44
#Peptides/#Unique	9/0	8/8	85/85	41/41	27/27	11/11
Q9BLL8 P25_BOMMA						
-10lgP	192.24	101.31	97.17	53.49	90.03	55.3
Coverage (%)	22	10	25	10	29	7
#Peptides/#Unique	5/0	2/2	8/8	2/2	7/1	1/1

V.5.1.1 *B. mori* - Tryptic digestion

In *B. mori*, The -10lgP score for protein identification for P05790 was much higher in TS than in TG. Most of the detected peptides were occurring close from the C- and N-terminus. There were 7 unique peptides found in both samples: GYGQGAGSAASSVSS ASSR (5213-5231) showed the highest probability score in both samples, followed by DASGAVIEEQITTK (49-62) and SYDYSR (5232-5237) (Figure 55). The exclusive features of P05790 as reviewed in V.2.2, GTGSSGFG, (PY)VANGG and GYEYAW were found in IS; but GVGAGY and (PY)VHAGG were not found in either sample (Figure 56, Figure 57).



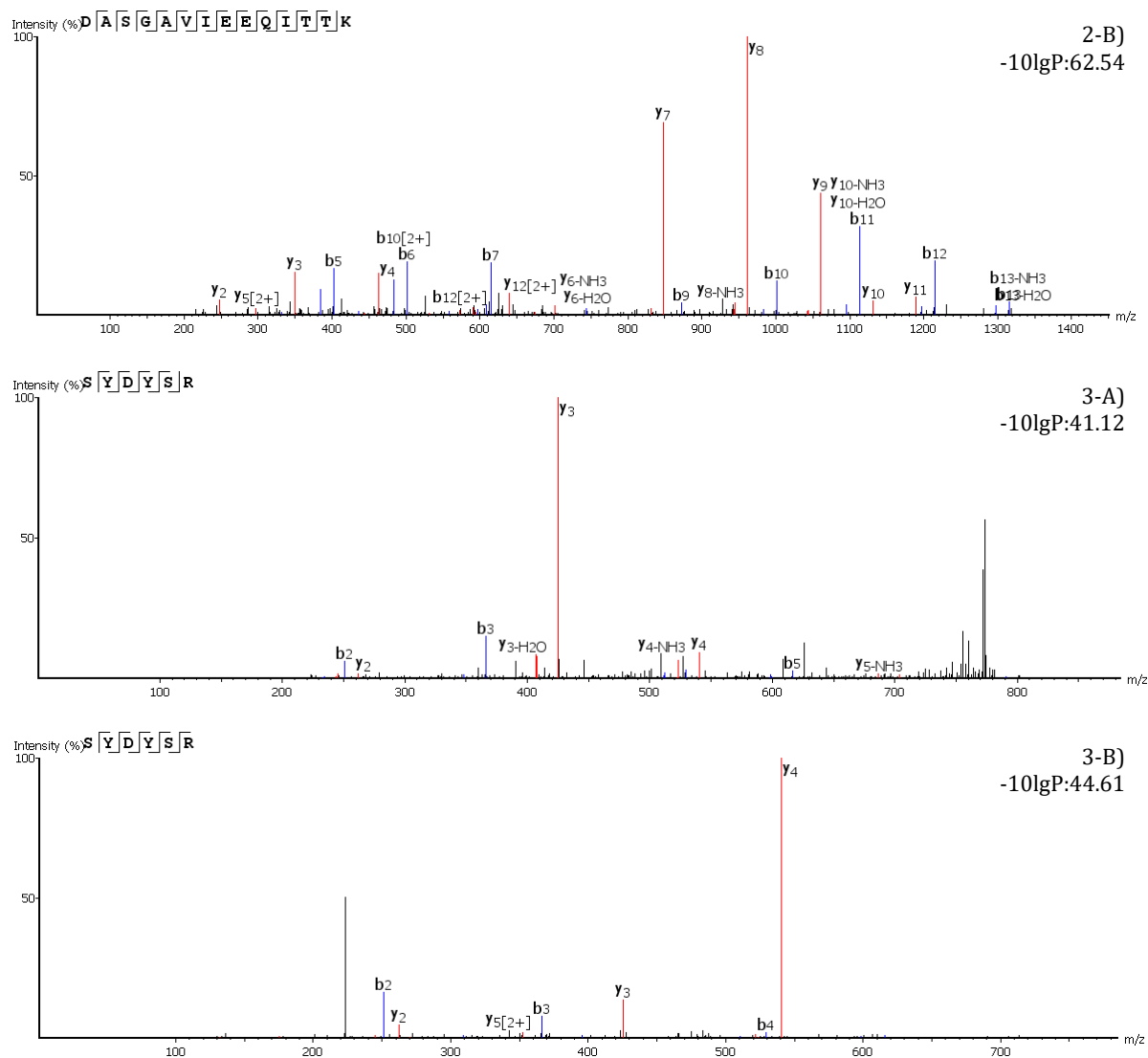


Figure 55. Spectra of the unique peptides to *B. mori* fibroin heavy chain (P05790): A) tryptic in-gel digestion, B) in-solution digestion; 1) GYGQGAGSAASSVSSASSR ($m/z = 843.89$), 2) DASGAVIEEQITTK ($m/z = 731.38$), and 3) SYDYSR ($m/z = 395.67$), $z = 2$

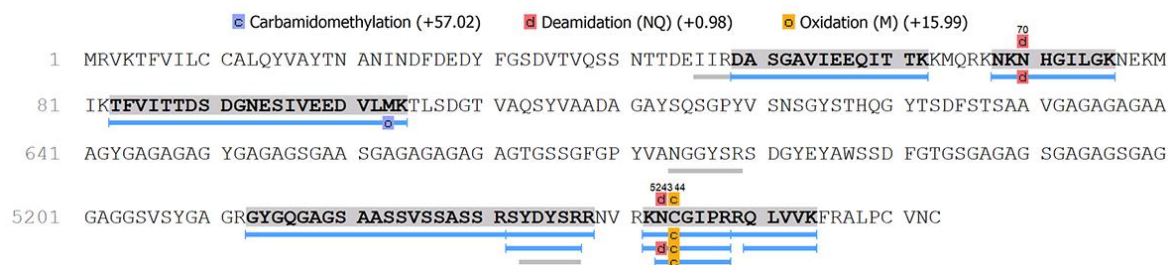


Figure 56. Protein coverage of *B. mori* heavy chain (P05790): tryptic in-gel digestion.

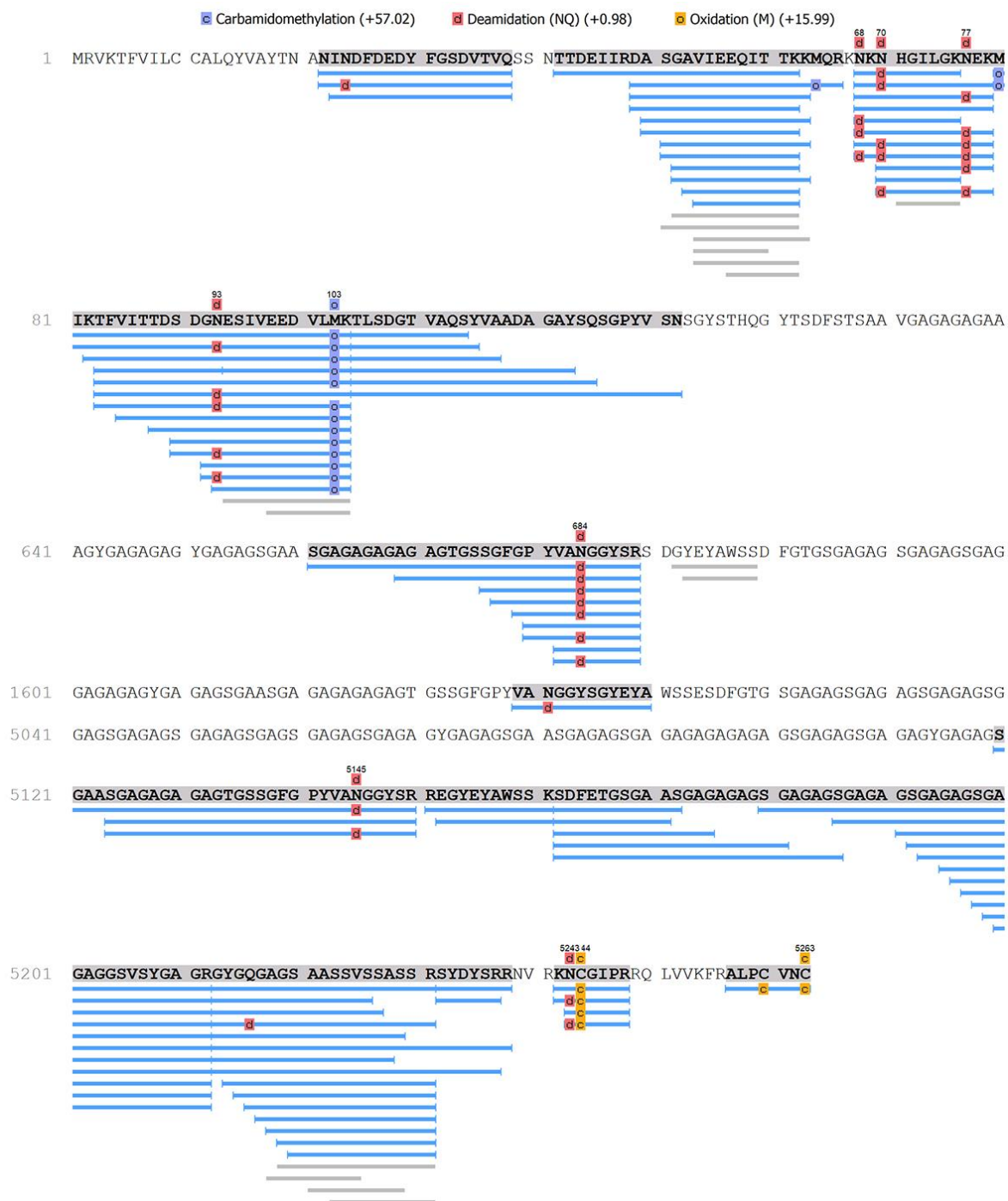
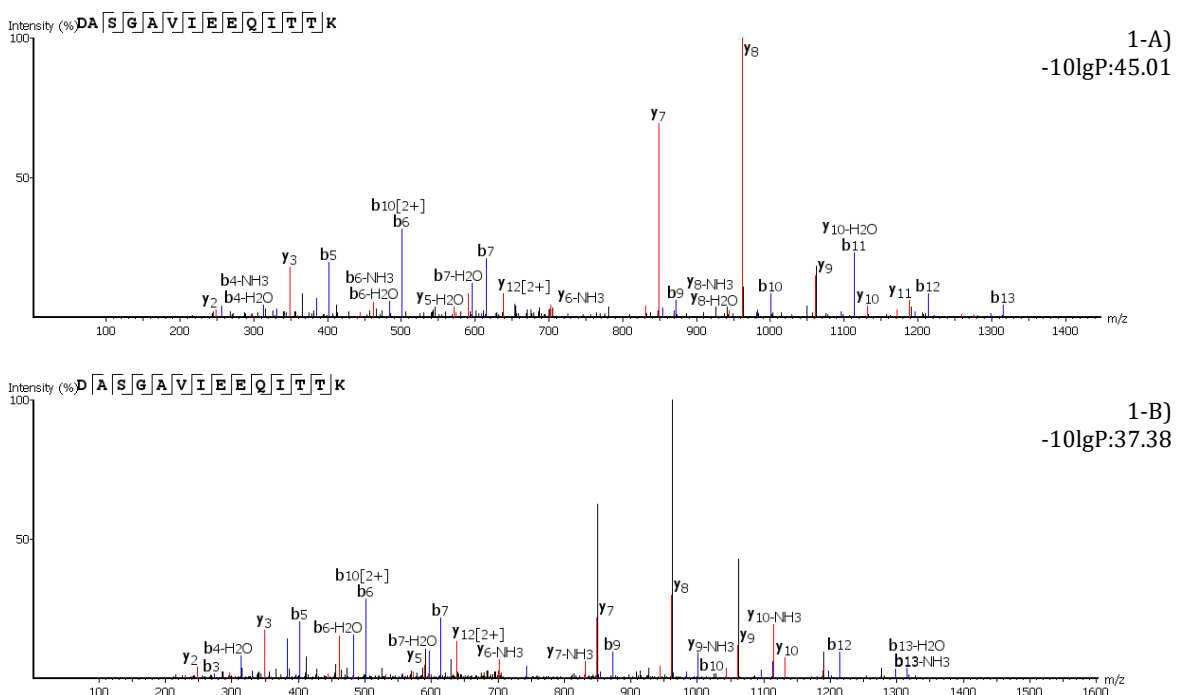


Figure 57. Protein coverage of *B. mori* heavy chain (P05790): tryptic in-solution digestion.

V.5.1.2 *B. mori* - Chymotryptic digestion

The protein identification score for P05790 was similar in the two samples. However, CS showed no peptides from the residues 1700-5000 while in-gel digestion yielded peptides throughout the protein sequence (Figure 59, Figure 60). CG yielded 219 peptides correspond to P05790 and 165 of them were unique; IS yielded 62 peptides and 46 of them were unique. There were 18 unique peptides found in both samples, including DASGAVIEE**Q**ITTK (49-62) that carries the marker sequence for *B. mori* that differentiate it from *B. mandarina* heavy chain. The unique sequence features such as **GTGSS**GFG, (PY)**VANGG**, GYE**YAW** and (PY)**VHAGG** were all found in both samples; however, GVGAGY was only found in CG. Among the unique peptides, the ones from residues 661-681 showed the highest probability scores: GAGAGAGAGTGSSGFGPY (664-681) showed the highest -10lgP score among the unique peptides found in both samples. In CG, AASGAGAGAGAGAG TGSSGFGPY (659-681) showed -10lgP score of 53.32. In IS, SGAGAGAGAGAGTGSSG FGPY (661-681) showed -10lgP score of 75.52 (Figure 58).



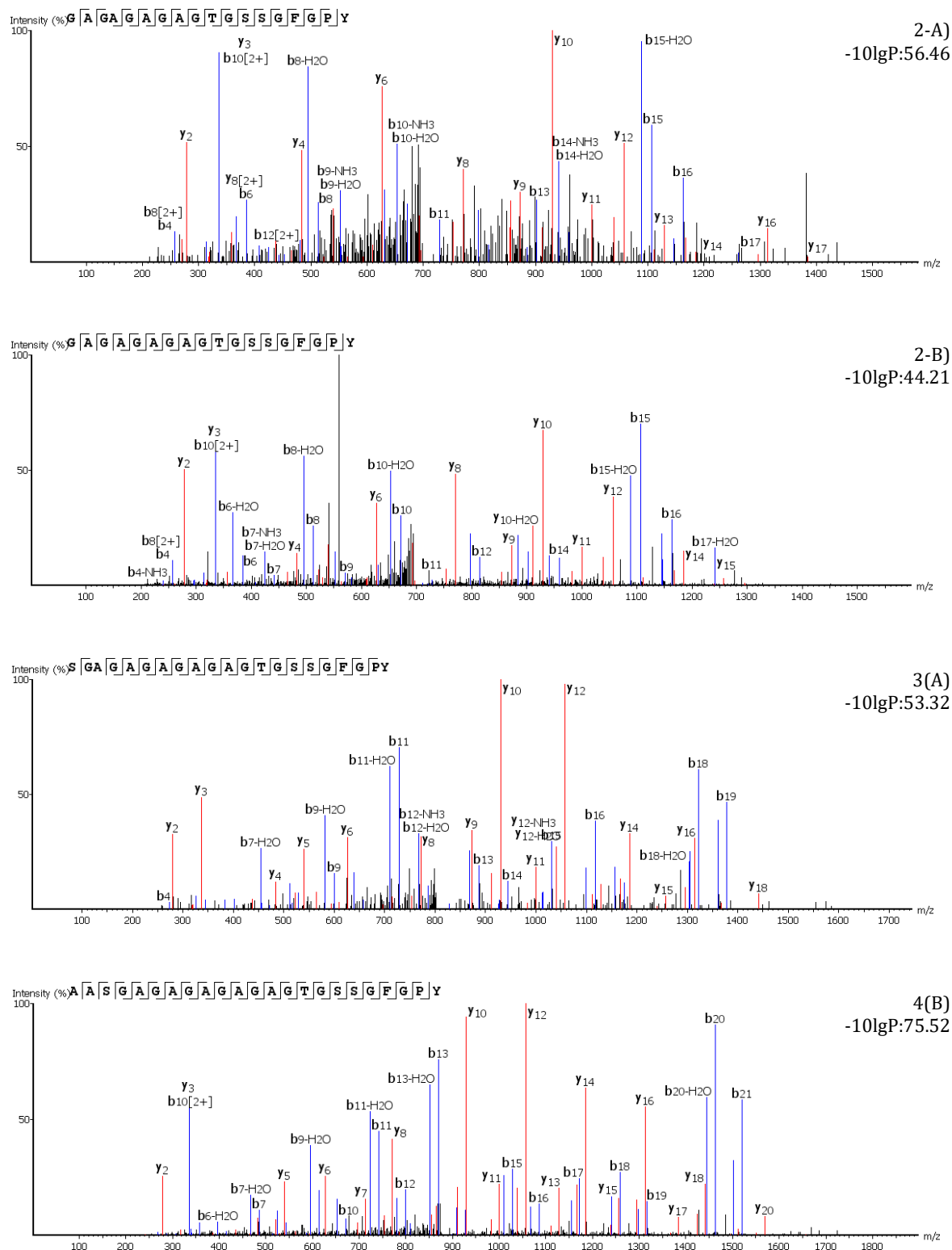


Figure 58. Spectra of the unique peptides to *B. mori* fibroin heavy chain (P05790): A) chymotryptic in-gel digestion, B) in-solution digestion; 1) DASGAVIEEQITTK (m/z = 731.38), 2) GAGAGAGAGTGSSGFGPY (m/z = 721.32), 3) AASGAGAGAGAGAGTGSSGFGPY (m/z = 899.90), and 4) SGAGAGAGAGAGTGSSGFGPY (m/z = 828.87), $z=2$



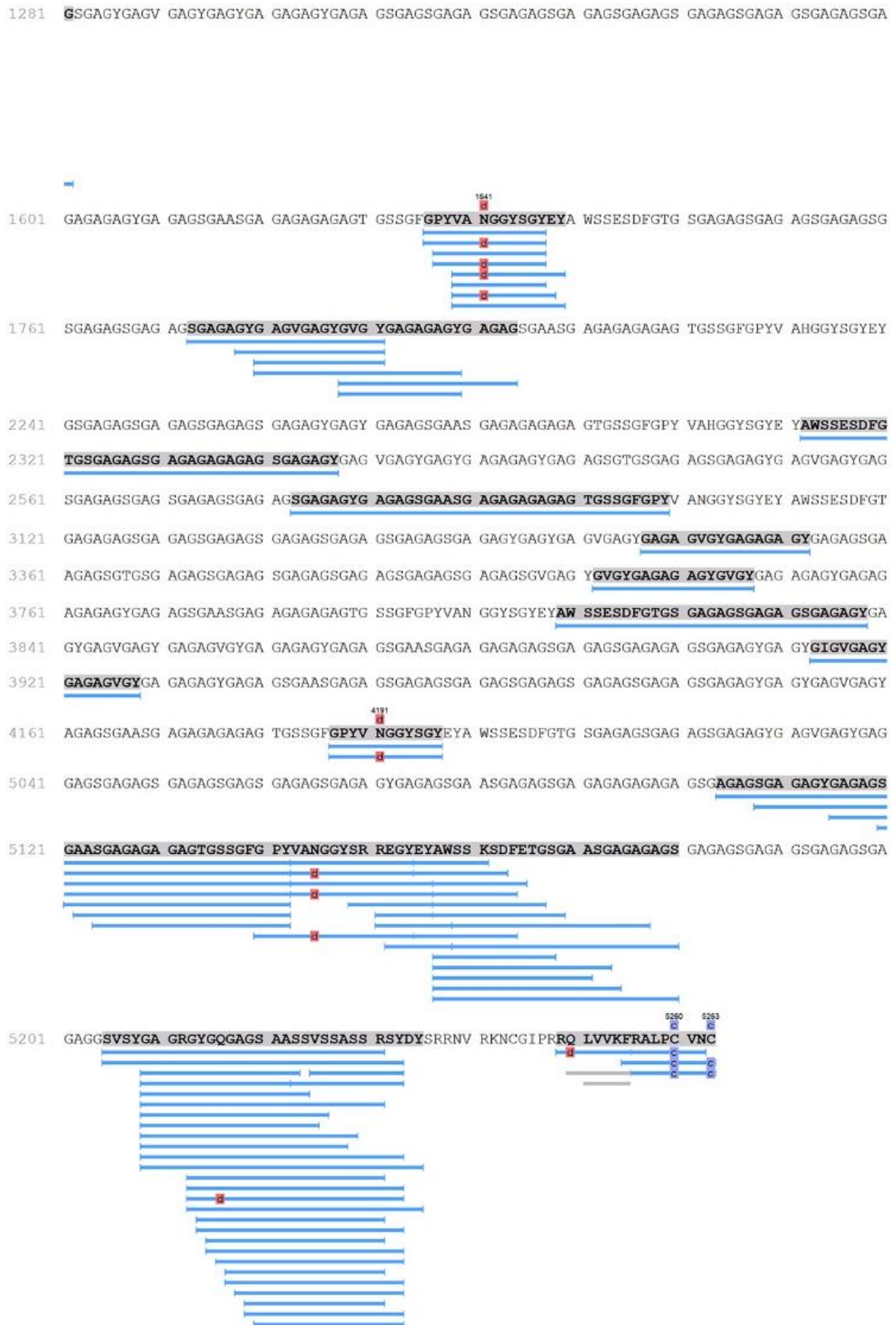


Figure 59. Protein coverage of *B. mori* heavy chain (P05790): chymotryptic in-gel digestion.

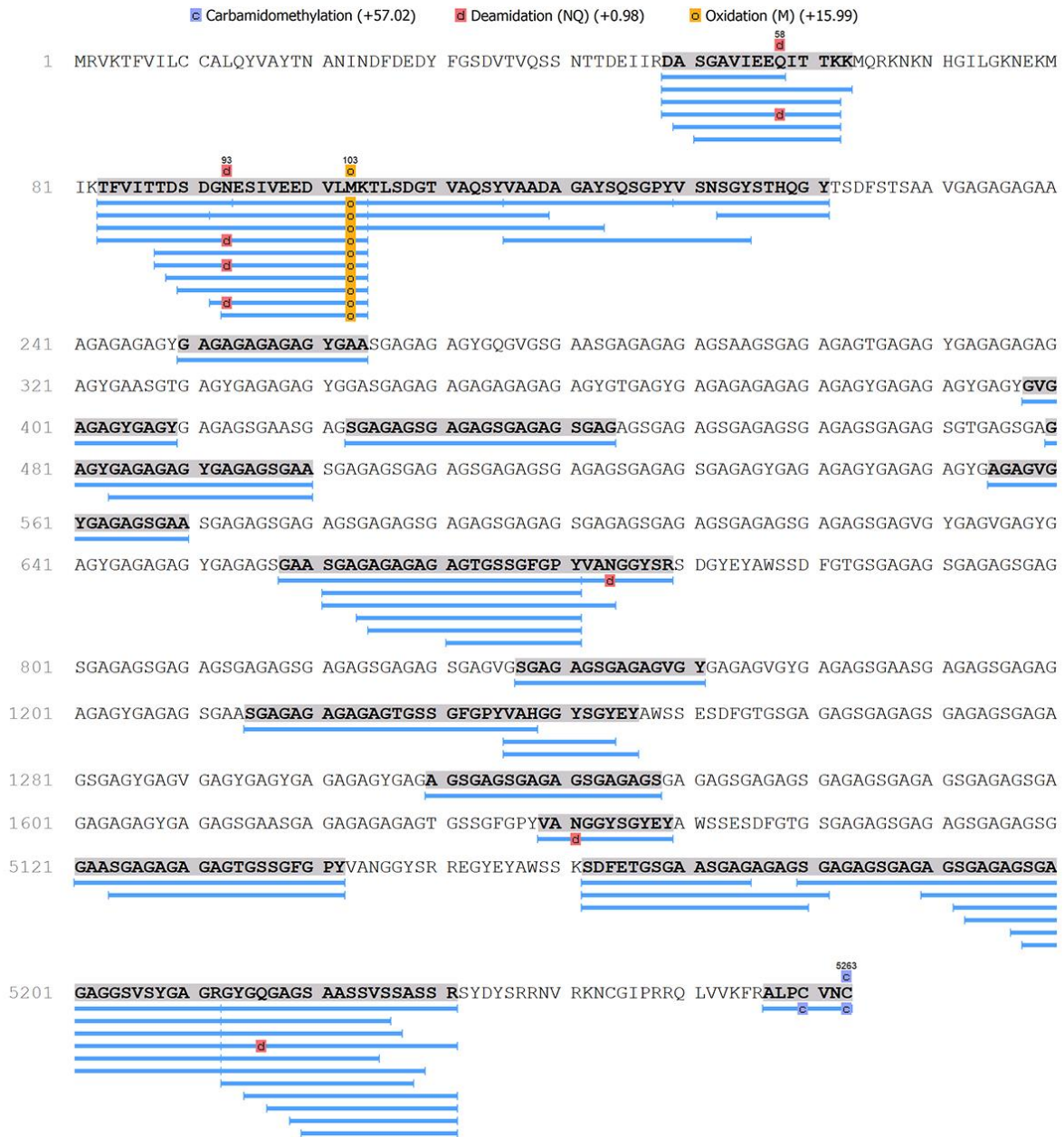
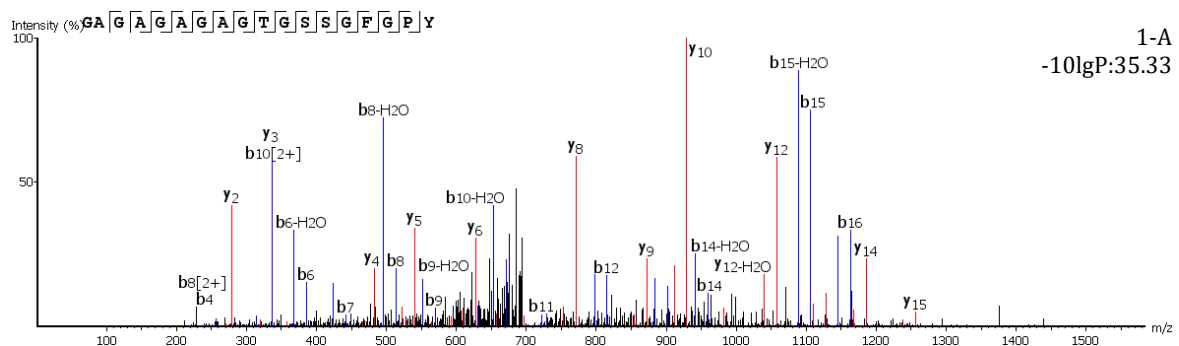


Figure 60. Protein coverage of *B. mori* heavy chain (P05790): chymotryptic in-solution digestion.

V.5.1.3 *B. mori* - Tryptic-chymotryptic digestion

The PEAKS analysis results were similar to the chymotryptic digestion. However, the protein coverages were better in tryptic-chymotryptic digestion. Between TCG and TCS, TCG sample showed better protein coverages and higher probability scores for both the heavy and light chain; P25 was not detected. The better the protein coverage, the more the peptides occurrence from the highly repetitive block chain regions was observed (see Figure 24) (Figure 62, Figure 63). TCG yielded 356 peptides correspond to P05790 and 324 of them were unique; IS yielded 153 peptides correspond to P05790 and 131 of them were unique. There were 37 unique peptides found in both samples. Among them, GAGAGAGAGTGSSGFGPY (664-681) showed the highest probability scores, then GAGAGAGYGAGAG (194-206) and GAGYGAGVGAGY (1112-1123) followed. The most abundant unique peptide in TCG was ALPCVNC (5267-5263), which is located at N-terminus of P05790, then followed by AGAGAGAGTGSSGFGPY (665-681). On the other hand, GAGAGAGAGAGAGAGY (348-363) showed the highest probability score in IS (Figure 61). The marker sequences for *B. mori* were only detected in TCG, as VIEEQITTK (54-62) and SGYSTHQGYTSDFFTS (133-148)—the peptides occurred in TCS had a low probability score thus screened out. The unique sequence features GVGAGY, GTGSSGFG and were all found in both samples; however, GYEYAW, (PY)VANGG were only found in the in-gel digestion sample; (PY)VHAGG were not found in either sample.



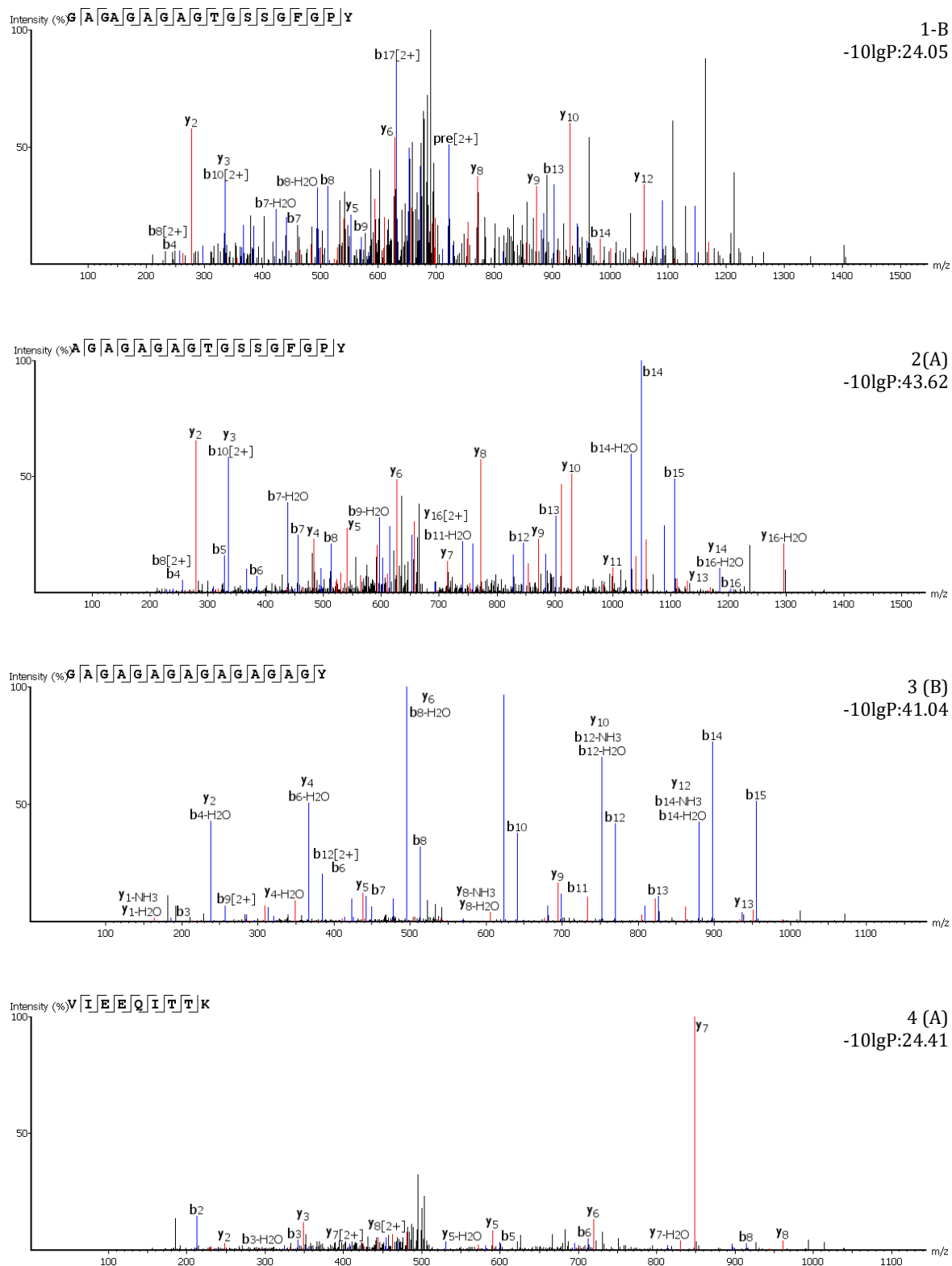
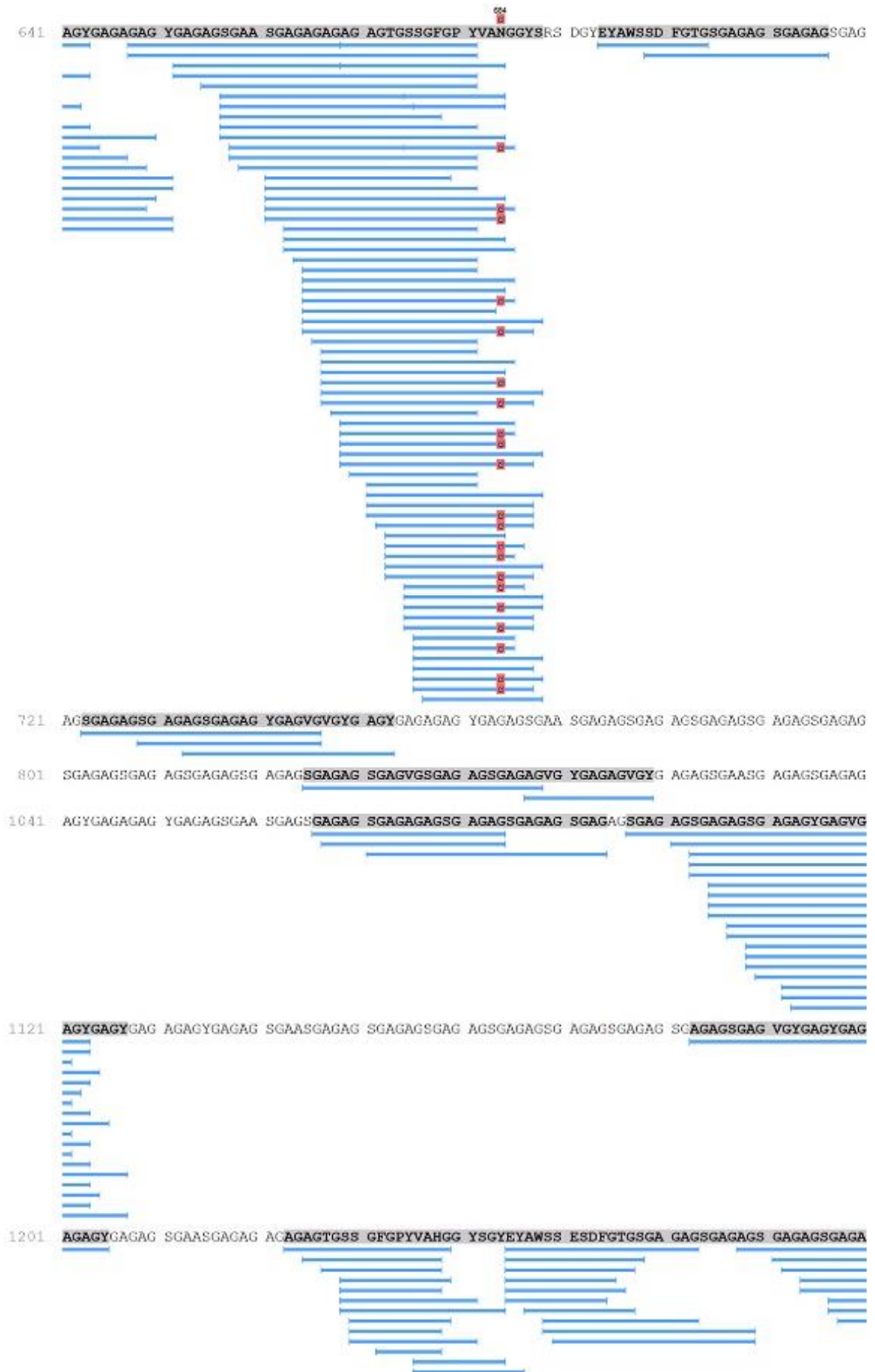


Figure 61. Spectra of the unique peptides to *B. mori* fibroin heavy chain (P05790): A) tryptic - chymotryptic in-gel digestion, B) in-solution digestion; 1) GAGAGAGAGTGS S GFGPY ($m/z=721.32$), 2) GAGAGAGAGTGS S GFGPY ($m/z=721.32$), 3) AGAGAGAGTGS S GFGPY ($m/z=692.8$), and 4) VIEEQITTK ($m/z=530.8$), $z=2$





1281 GSGAGYGAGV GAGYGAGYGA GAGAGYGAGA GSGAGSGAGA GSGAGAGSGA GAGSGAGAGS GAGAGSGAGA GSGAGAGSGA

1361 GAGSGAGAGY GAGYGAGAGA GYGAGAGSGA GSGAGAGSGA GAGSGAGAGS GAGAGSGAGA GSGAGAGSGA GAGSGAGAGY

1601 GAGAGAGYGA GAGSGAASGA GAGAGAGAGT GSSGFGPYVA NGGYSGYEY WSESDFGTG SGAGAGSGAG AGSGAGAGSG

1681 AGAGSGAGAG YGAGYGAGAG AGYGAGAGSG AGSGAGAGSG AGAGSGAGAG SGAGAGSGAG AGSGAGAGSG AGAGSGAGSG

1761 SGAGAGSGAG AGSGAGAGY AGVGAGYGVG YGAGAGAGY AGAGSGAASG AGAGAGAGAG TGSSGFGPYV AHGGYSGYEY

2081 GYGAGAGSGA GSGAGAGSGA GAGYGAGAGA GYGAGYGAGA GAGYGAGAGT GAGSGAGAGS GAGAGSGAGA GSGAGAGSGA

2321 TGSGAGAGSG AGAGAGAGAG SGAGAGYAG VAGYGAGYAG AGAGAGYAG AGSGTGSGAG AGSGAGAGY AGVGAGYAG

2401 AGSGAAGAG AGAGAGSGAG AGSGAGAGSG AGAGSGAGAG SGAGAGYAG YGAGVGAGY AGAGSGAASG AGAGSGAGAG

2481 SGAGAGSGAG AGSGAGAGSG AGAGYGAGVG AGYGAGYAG AGAGYGAGAG SGAASGAGAG SGAGAGAGSG AGAGSGAGAG

2561 SGAGAGSGAG SGAGAGSGAG AGSGAGAGY AGAGSGAASG AGAGAGAGAG TGSSGFGPYV AHGGYSGYEY WSESDFGT

2801 GAGYGVYGA GAGAGYGAGA GSGAGSGAGA GSGAGAGSGA GAGSGAGAGS GAGSGAGAGS GAGAGSGAGA GSGAGSGAGA

2881 GSGAGAGYGV GYGAGAGAGY GAGAGSGAGS GAGAGSGAGA GSGAGAGSGA GSGAGAGSGA GAGSGAGAGS GAGAGYAGAGV

3121 GAGAGAGSGA GAGSGAGAGS GAGAGSGAGA GSGAGAGSGA GAGYGAGYGA GVGAGYAGA GVGYAGAGA GYGAGAGSGA

3361 AGAGSGTGSG AGAGSGAGAG SGAGAGSGAG AGSGAGAGSG AGAGSGAGAG YGVGYGAG AGYGVGAGAG AGAGYAGAG



Figure 62. Protein coverage of *B. mori* heavy chain (P05790): tryptic-chymotryptic in-gel digestion.

1 MRVKTFVILC CALQYVAYTN ANINDFDEDY FGSDVTVQSS NTTDEIIRDA SGAVIEEQIT TKKMQRKNKN HGILGKNEKM
81 IKTFVITTDS DGNESIVEED VLMKTLSDGT VAQSYVAADA GAYSQSGPYV SNSGYSTHQG YTSDFS TSAA VGAGAGAGAA

161 AGSGAGAGAG YGAASGAGAG AGAGAGAGYG TGAGAGAGAG YGAGAGAGAG AGYGAGAGAG AGAGYGAGAG AGAGAGYGAG

241 AGAGAGAGYG AGAGAGAGAG YGAASGAGAG AGYGQGVGSG AASGAGAGAG AGSAAGSGAG AGAGTGAGAG YGAGAGAGAG

321 AGYGAASGTG AGYGAGAGAG YGGASGAGAG AGAGAGAGAG AGYGTGAGYG AGAGAGAGAG AGAGYGAGAG AGYGAGYGVG

401 AGAGYGAGYG AGAGSGAASG AGSGAGAGSG AGAGSGAGAG SGAGAGSGAG AGSGAGAGSG AGAGSGAGAG SGTGAGSGAG

481 AGYGAGAGAG YGAGAGSGAA SGAGAGSGAG AGSGAGAGSG AGAGSGAGAG SGAGAGYGAG AGAGYGAGAG AGYGAGAGVG

561 YGAGAGSGAA SGAGAGSGAG AGSGAGAGSG AGAGSGAGAG SGAGAGSGAG AGSGAGAGSG AGAGSGAGVG YGAGVGAGYG

641 AGYGAGAGAG YGAGAGSGAA SGAGAGAGAG AGTGSSGFGP YVANGGY SRS DGYEYAWSS FGTGSGAGAG SGAGAGSGAG

801 SGAGAGSGAG AGSGAGAGSG AGAGSGAGAG SGAGVSGAG AGSGAGAGVG YGAGAGVGYG AGAGSGAASG AGAGSGAGAG

961 AGAGSGAGAG SGAGAGSGAG AGSGAGAGSG AGAGSGAGAG SGAGAGSGAG AGSGAGAGSG AGAGSGAGAG YGAGAGAGYG

1041 AGYGAGAGAG YGAGAGSGAA SGAGSGAGAG SGAGAGAGSG AGAGSGAGAG SGAGAGSGAG AGSGAGAGSG AGAGYGAGVG

1121 AGYGAGYGAG AGAGYGAGAG SGAASGAGAG SGAGAGSGAG AGSGAGAGSG AGAGSGAGAG SGAGAGSGAG VYGAGYGAG

1201 AGAGYGAGAG SGAASGAGAG AGAGAGTGSS GFGPYVAHGG YSGYEYAWSS ESDFGTGSGA GAGSGAGAGS GAGAGSGAGA

1281 GSGAGYGAGV GAGYGAGYGA GAGAGYGAGA GSGAGSGAGA GSGAGAGSGA GAGSGAGAGS GAGAGSGAGA GSGAGAGSGA

1681 AGAGSGAGAG YGAGYGAGAG AGYGAGAGSG AGSGAGAGSG AGAGSGAGAG SGAGAGSGAG AGSGAGAGSG AGAGSGAGSG

1761 SGAGAGSGAG AGSGAGAGYG AGVGAGYGVG YGAGAGAGYG AGAGSGAASG AGAGAGAGAG TGSSSGFPGYV AHGGYSGY EY

1841 AWSSESDFGT GSGAGAGSGA GAGSGAGAGS GAGAGSGAGA GSGAGAGSGA GAGY GAGVGA GYGAAYCAGA GAGYGAGAGS

2081 GY GAGAGSGA GSGAGAGSGA GAGYGAGAGA GYGAGYGAGA GAGY GAGAGT GAGSGAGAGS GAGAGSGAGA GSGAGAGSGA

2161 GAGSGAGAGS GAGSGSGAGA GSGAGAGSGA GAGSGAGAGS GAGAGSGAGA GYGAGAGAGY GAGYGAGAGA GY GAGAGSGA

2241 **GSGAGAGSGA GAGSGAGAGS GAGAGY**GAGY GAGAGSGAAS GAGAGAGAGA GTGSSGFGPY VAHGGYSGYE YAWSSSEDFG
 2321 TGSGAGAG**SG AGAGAGAGAG SGAGAGY**GAG VGAGYGAGYG AGAGAGYGAG AGSGTGSGAG AGSGAGAGY**G AGVGAGYGAG**
 2401 **AGSGAAF**GAG AGAGAGSGAG AGSGAGAGSG AGAGSGAGAG SGAGAGYGAG YGAGVGAGYG AGAGSGAASG AGAGSGAGAG
 2481 SGAGAGSGAG AGSGAGAGSG AGAGYGAGVG AGYGAGYGAG AGAGY**GAGAG SGAASGAGAG SGAGAGAGSG AGAGSGAGAG**
 2801 GAGYGVGYGA GAGAGYGAGA GSGAGSGAGA GSGAGAGSGA GAGSGAGAGS GAGSGAG**AGS GAGAGSGAGA GSGAGSGAGA**
 2881 **GSGAGAGY**GV GYGAGAGAGY GAGAGSGAGS GAGAGSGAGA GSGAGAGSGA GSGAGAGSGA GAGSGAGAGS GAGAGYGAGV
 3121 GAGAGAGSGA GAGSGAGAGS GAGAGSGAGA GSGAGAGSGA GAGYGAGY**GA GVGAGYGAGA GVGYGAGAGA GYGAGAGSGA**
 3201 **ASGAGAGAGS GAGAGTGAGA GSGAGAGY**GA GAGSGAASGA GAGAGAGAGT GSSGFGPYVA NGGYSGYEYA WSSSEDFGTG
 3361 AGAGSGTGSG AGAGSGAGAG SGAGAGSGAG AG**SGAGAGSG AGAGSGVGAG Y**GVGYGAGAG AGYGVGYGAG AGAGYGAGAG
 3841 GYGAGVGAGY GAGAGVGAGA GAGAGYGAGA GSGAASGAGA GAG**AGAGSGA GAGSGAGAGA GSGAGAGY**GA GY**GIGVGAGY**
 3921 **GAGAGVGY**GA GAGAGYGAGA GSGAASGAGA GSGAGAGSGA GAGSGAGAGS GAGAGSGAGA GSGAGAGYGA GYGAGVGAGY
 4081 GSGAGAGSGA GAGYGAG**GVGA GYGAGYGGAG AGY**GAGAGSG AASGAGAGSG AGAGSGAGAG SGAGAGSGAG AGSGAGAGY
 4321 AGYGAGVGAG YGAGYGAGAG AGYGAGAGSG AASGAGAGSG AGAGAGSGAG AGSGAGAGSG AGAGSGAG**AG SGAGAGSGAG**
 4401 **SGAGAGSGAG AGYGAGY**GAG VGAGYGAGAG VYGAGAGAG YGAGAGSGAA SGAGAGSGSG AGSGAGAGSG AGAGSGAGAG
 4881 GSGAGAGSGA GAGSGAGAGS GAGAGSGAGA GSGAGSGAGA GSGAGAGYGA **GYGAGVGAGY GAGAGY**GAGY **GVGAGAGYGA**
 4961 **GAGSGAGSGA GAGSGAGAGS GAGAGSGAGA GSGAGAGSGA GSGAGAGY**GA GAGAGYGAGA GAGY**GAGAGS GAASGAGAGA**
 5041 **GAGSGAGAGS GAGAGSGAGS GAGAGSGAGA GY**GAGAGSGA ASGAGAGSGA GAGAGAGAGA GSGAGAGSGA GAGYGAGAGS
 5121 GAASGAGAGA GAGTGSSGFG PYVANGGYSR REGYEWSS KSDFETGSGA ASGAGAGAGS GAGAGSGAGA GSGAGAG**SGA**
 5201 **GAGGSVSYGA GRGY**GQAGS AASSVSSASS RSYDYSRRNV RKNCGIPRRQ LVVKFRALPC VNC

Figure 63. Protein coverage of *B. mori* heavy chain (P05790): tryptic-chymotryptic in-solution digestion.

V.5.1.4 *B. mandarina* (partial)

As the protein sequence of *B. mandarina* heavy chain(Q99050) is incomplete, samples corresponded to *B. mori* heavy chain (P05790). The identification of *B. mandarina* at this point is completely relying on a small number of unique residues: E (58) and R (141) in the heavy chain, G (183) in the light chain, and R (62) and H (219) in the P25. All samples, regardless of enzymes and digestion methods, corresponded to P05790 for the heavy chain by the marker peptide DASGAVIEEQITTK (49-62). This might have resulted from the sequence variability. However, the light chain and P25 were often identified as that of *B. mandarina*.

In tryptic digestion, TG corresponded to Q9BLL9(subset P21818) for the light chain and Q9BLL8 for P25. In the light chain, none of the detected peptides was unique to either protein. In P25, a unique marker for Q9BLL8 was found as GQIPSQYEIPVFR (50-62). IS showed the same result as TG for the light chain and P25 (Figure 64).

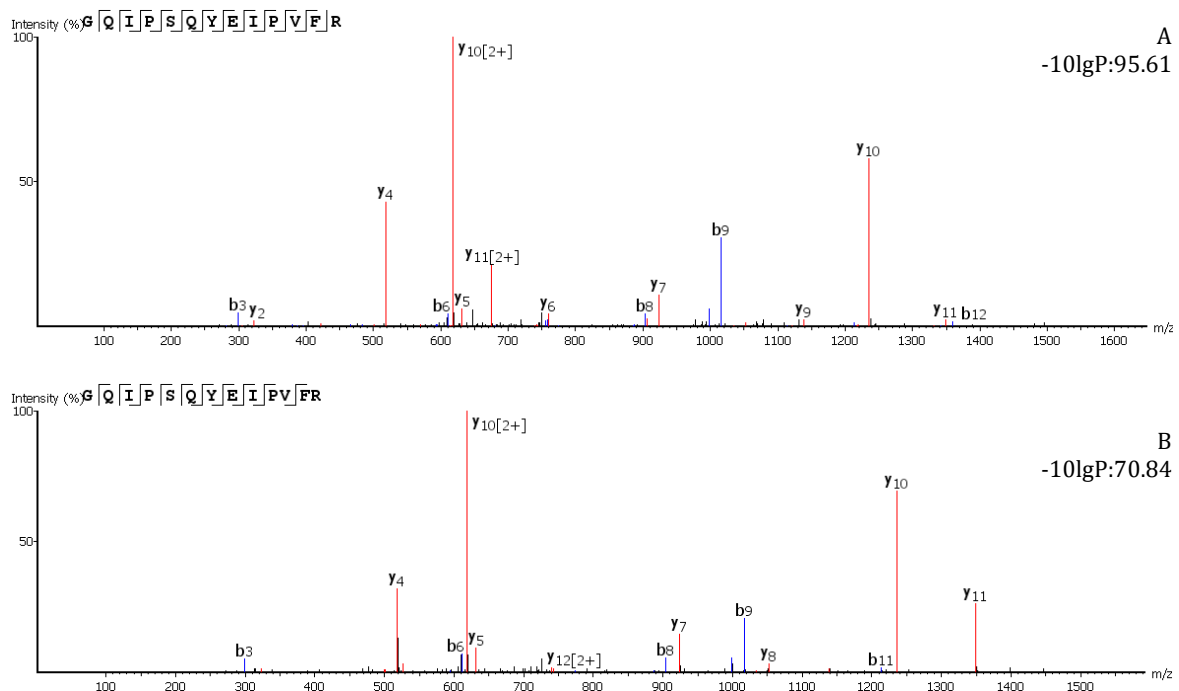
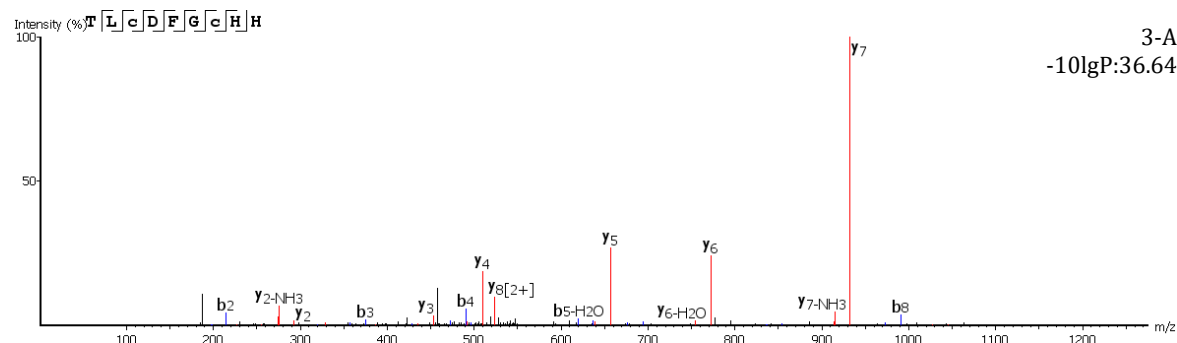
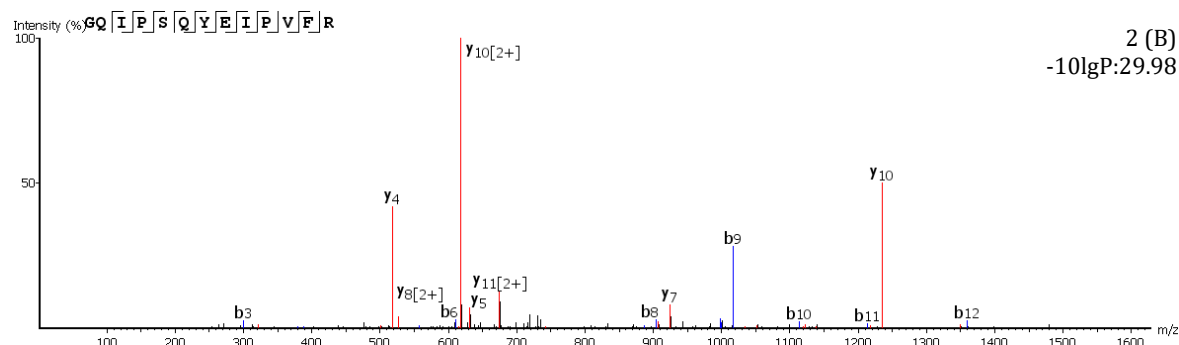
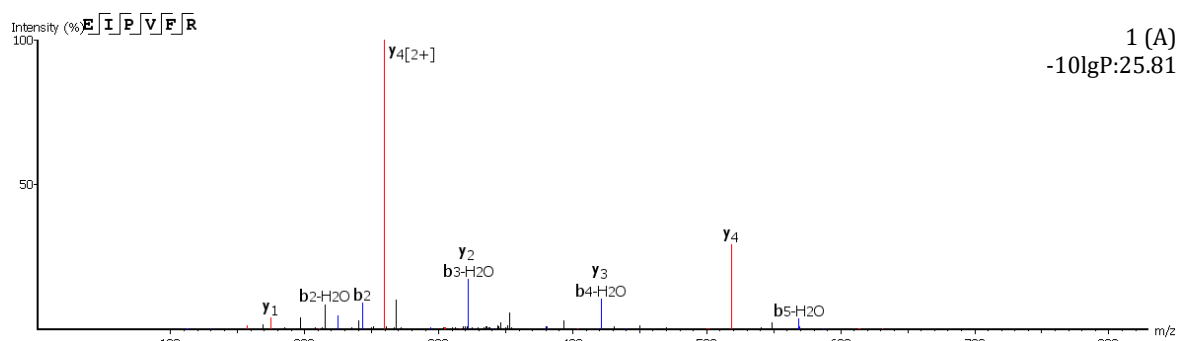


Figure 64. Unique marker peptide for Q9BLL8: GQIPSQYEIPVFR (m/z=767.40, z=2) in tryptic in-gel (A) and in-solution (B) digestion.

The protein identification results in chymotryptic digestion were similar to that of tryptic digestion: the heavy chain corresponded to P05790, the light chain corresponded to both Q9BLL9 and P21818, and P25 corresponded to Q9BLL8 by the unique residue **R** (58). In TG, both the unique residues in Q9BLL8 were detected as EIPVFR (58-62) and TLCDFG**HH** (212-220). In TS, GQIPSQYEIPVFR (50-62) and TLCDFG**HH** (212-220) were both detected, but the latter was screened out for the PEAKS analysis result due to the low probability score($-10\lg P=14.61$) (Figure 65).



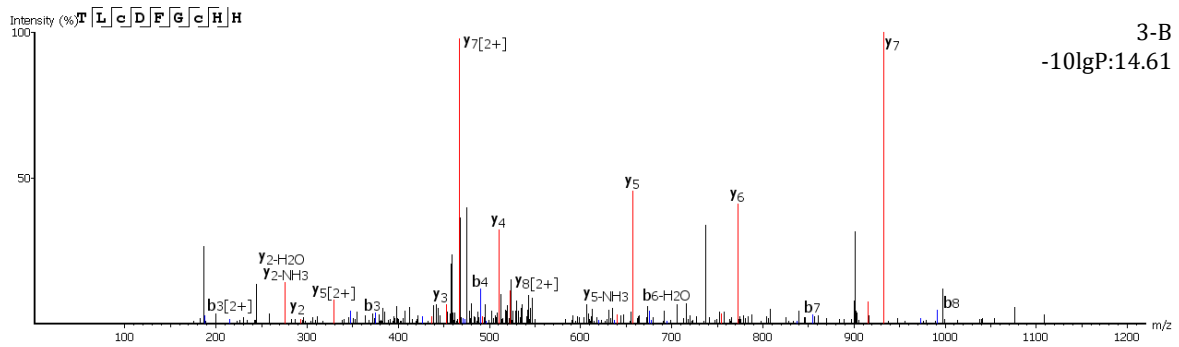


Figure 65. Unique marker peptides for Q9BLL8 in chymotryptic in-gel (A) and in-solution (B) digestion: 1) EIPVFR ($m/z=380.72$) 2) GQIPSQYEIPVFR ($m/z=767.40$) and 3) TLCDFGCHH ($m/z=573.73$), $z=2$.

On the other hand, in tryptic-chymotryptic digestion, P25 corresponded to both Q9BLL8 and P04148 (*B. mori*), but the light chain corresponded to Q9BLL9 in both TCG and TCS. A unique marker peptide for Q9BLL9 was detected in both samples as DAILASSDSGFLNEEY (174-189), however screened out for PEAKS analysis result due to the low probability scores (Figure 66).

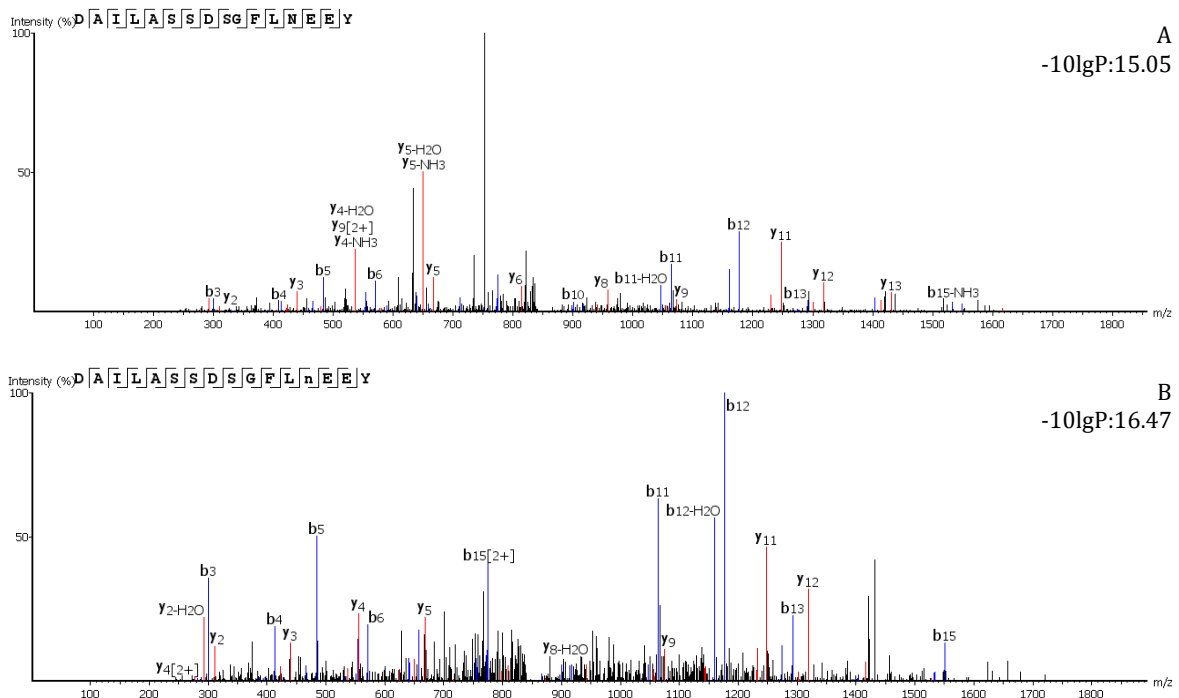


Figure 66. Unique marker peptide for Q9BLL9: DAILASSDSGFLNEEY ($m/z=866.38$, $z=2$) in tryptic-chymotryptic in-gel (A) and in-solution (B) digestion.

V.5.2 *Antheraea pernyi*

In *A. pernyi* silk, the chymotryptic in-gel digested sample showed the highest protein coverage (45%) with protein -10lgP score (330.49). There were 194 peptides matching to O76786 and 96 of them were unique to the protein. Tryptic peptides showed low protein coverages and the number of peptides despite the high -10lgP score. This result corresponds to the theoretical projection in V.2.2; supporting that the use of chymotrypsin promoted more cleavages, thus resulted in more MS-appropriate-sized peptides than the tryptic digestion. Although in theory the tryptic-chymotryptic digestion would be the most effective method, chymotryptic digestion showed the slightly better result in this set of experiment. This must have been affected by the incomplete protein solubilisation and digestion, thus further examinations to define the chymotrypsin alone or in combination with trypsin is more effective for the proteolysis of silk fibroins for mass spectrometry purpose is necessary. However, this was not covered in this research due to the limited time and resource.

Table 26. PEAKS analysis results of *A. pernyi* fibroin treated with different enzymes

	Trypsin		Chymotrypsin		Try-Chymo	
	In-gel (TG)	In-sol (TS)	In-gel (CG)	In-sol (CS)	In-gel (TCG)	In-sol (TCS)
<i>Antheraea pernyi</i>	(O76786 ANTPE)					
-10lgP	229.53	278.71	336.90	228.85	192.70	285.65
Coverage (%)	8	12	48	39	26	38
#Peptides	29	54	233	236	64	158
#Unique	3	11	113	104	25	68

V.5.2.1 Tryptic digestion

In tryptic digestion, both tryptic peptides and peptides with non-enzymatic cleavages were found. In Figure 68 and Figure 69, the protein sequence of O76786 is marked in black and green in turn by the tryptic cleavages. Although most of the resulted peptides are either tryptic or has non-specific cleavage at one end, a few numbers of peptides cleaved by no specific rule were detected in TS sample. But their peptide probability scores were not as high as the tryptic peptides. For example, a unique tryptic peptide of O76786 (R)AAGSA₁₂SGAGR(S) at the residues 2211-2231 was the highest scoring peptide in TS (-10lgP=68.96) and the second highest scoring peptide in TG (-10lgP=62.22). However, the highest scoring peptide in TG (-10lgP=62.45) (R)AAGSA₁₃SGAGGAGGG(Y) at the residues 1590-1615 has one non-specific cleavage at N-terminus. On the other hand, the unique peptides found in TS, (Y)GSGSA₅GSGAGGSGGG(Y) at 2498-2516 (-10lgP=48.06) and (A)A₉GSGAGGAGGG(Y) at 425-443 (-10lgP=37.76) were occurred by the alanine and tyrosine cleavages at both ends (Figure 67).

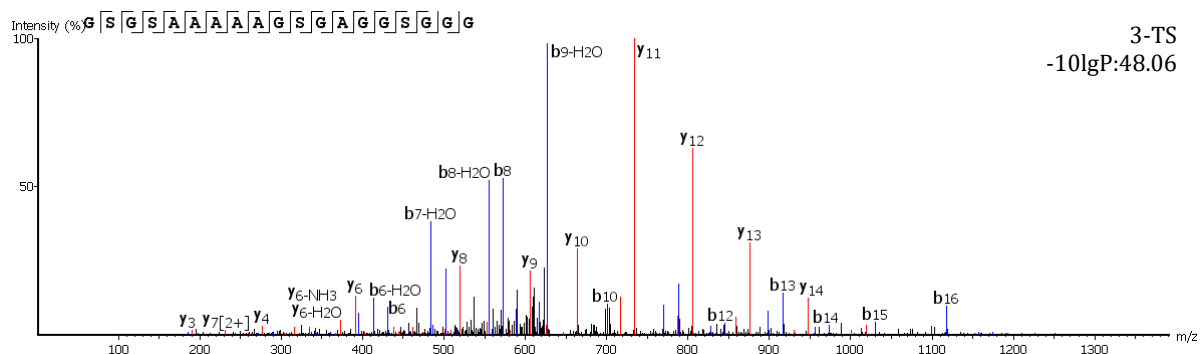
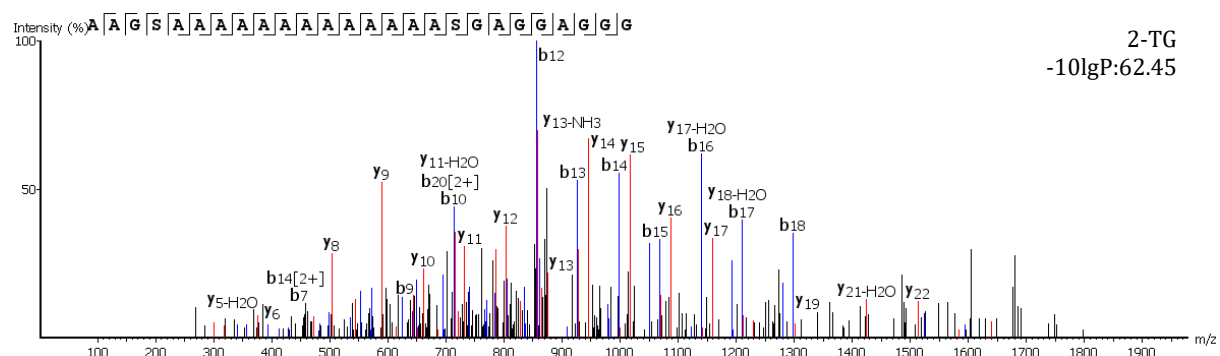
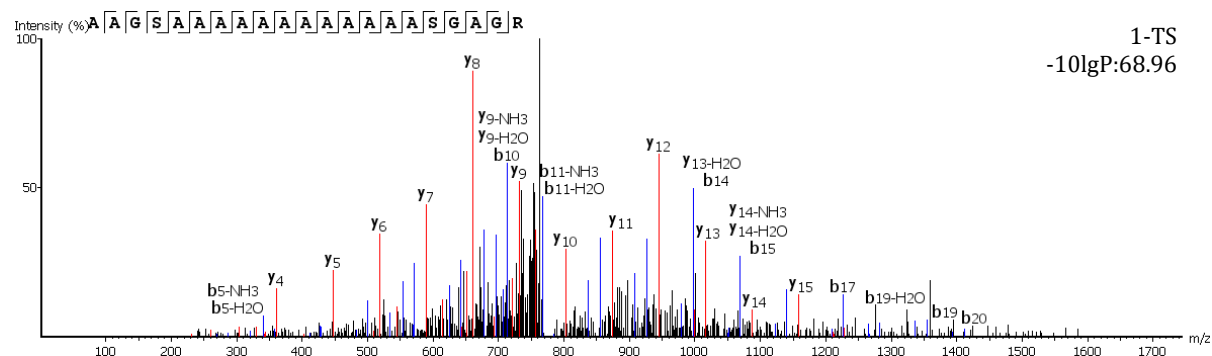
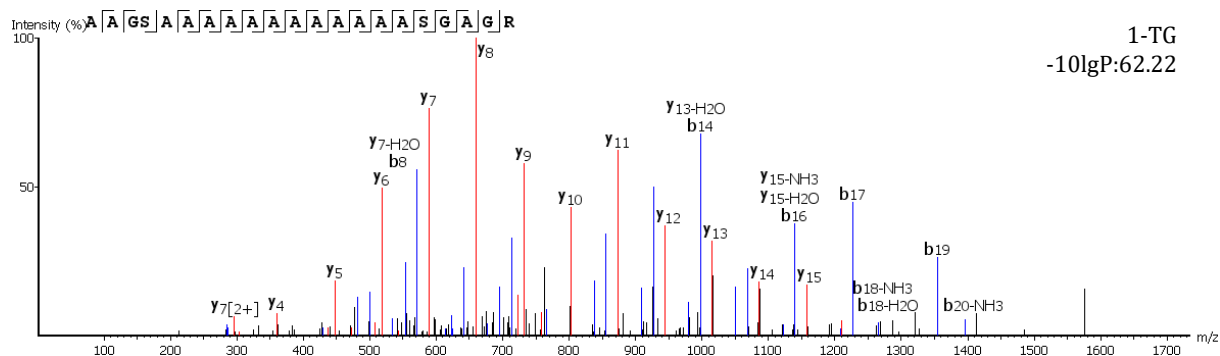


Figure 67. Spectra of tryptic unique peptides of *A. pernyi* fibroin: 1) AAGSA₁₂SGAGR ($m/z=793.40$), 2) AAGSA₁₃SGAGGAGGG ($m/z=900.43$), 3) GSGSAAAAAGSGAGGSGGG ($m/z=653.78$), $z=2$

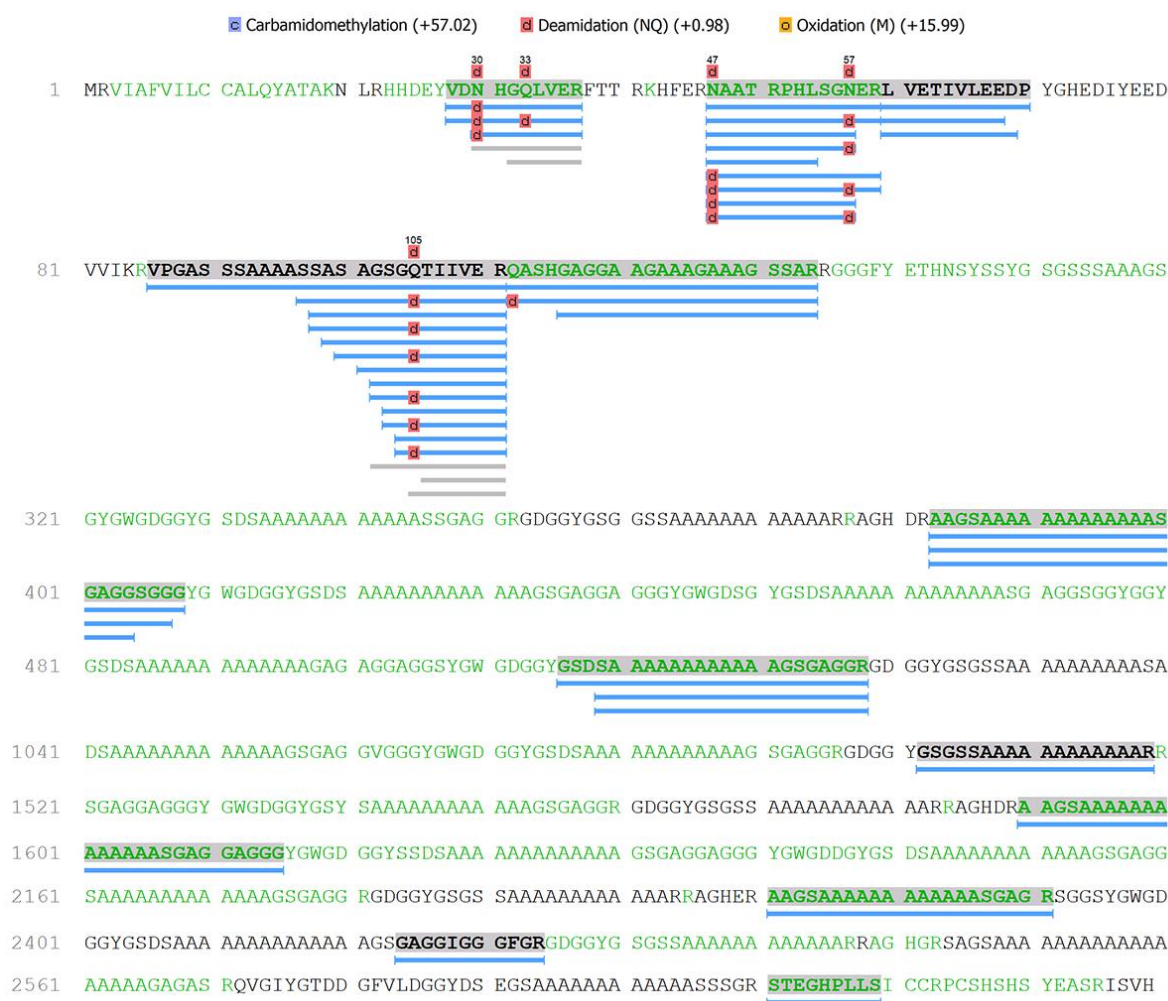


Figure 68. Protein coverage of *A. pernyi* fibroin (076786): tryptic in-gel digestion.



Figure 69. Protein coverage of *A. pernyi* fibroin (076786): tryptic in-solution digestion.

V.5.2.2 Chymotryptic digestion

The chymotryptic digestion was successful for *A. pernyi* silk, yielding high protein probability score and good sequence coverage in both CG and CS: the peptide distribution shows that if properly worked, in-gel digestion can be beneficial to achieve the high protein coverage. However, there were no peptides detected corresponding to the residues 700-1100 in both samples. In Figure 71 and Figure 81Figure 72, the protein sequence marked in black and green in turn indicates the chymotryptic cleavages. Similar peptides were detected in both samples; however, CG showed the regular distribution of peptides corresponding to O76786 and less non-enzyme specific peptides. As in the tryptic digestion, peptides resulted from non-specific cleavages in both ends were only found in CS: for example, unique peptides (D)SA₁₄GSGAGGIGGGFGRGD(G) (2407-2436) with -10lgP score of 46.96 and (D)SA₁₃GSGAGGRGD(G) (2161-2183) with -10lgP score of 39.29 are only occurred in CS.

In CG, a chymotryptic unique peptide GSDSA₁₃GSGAGGSGGY (197-223) yielded the highest probability score (-10lgP= 68.74), followed by unique chymotryptic peptides GSDSA₁₃SGAGGSGGY (452-477) with -10lgP score of 67.76. In CS, the highest scoring peptide was GSDSA₁₃GSGAGGVGGGY(-10lgP=67.22) that shared between O76786 and E1CGA3 (*A. yamamai*). Then a chymotryptic unique peptide GSDSSA₁₀SGAGGSGGGY (2372-2396) followed (-10lgP=63.23). This peptide was also found in CG with -10lgP score of 64.72 (Figure 70).

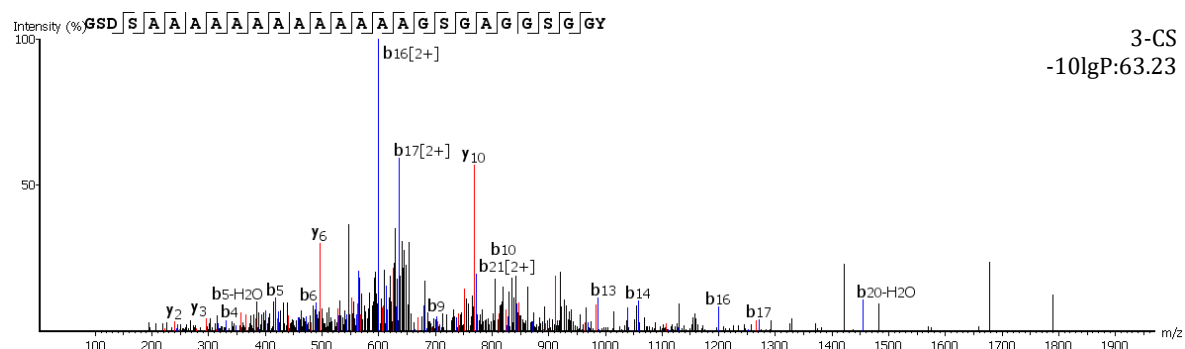
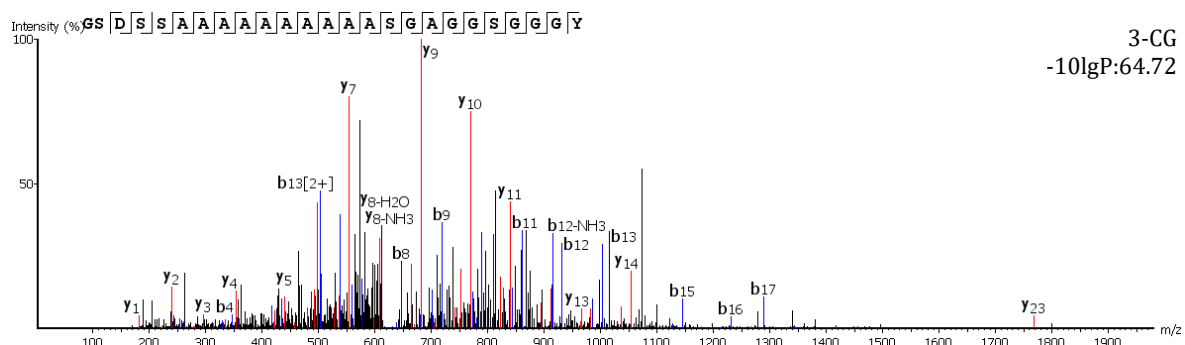
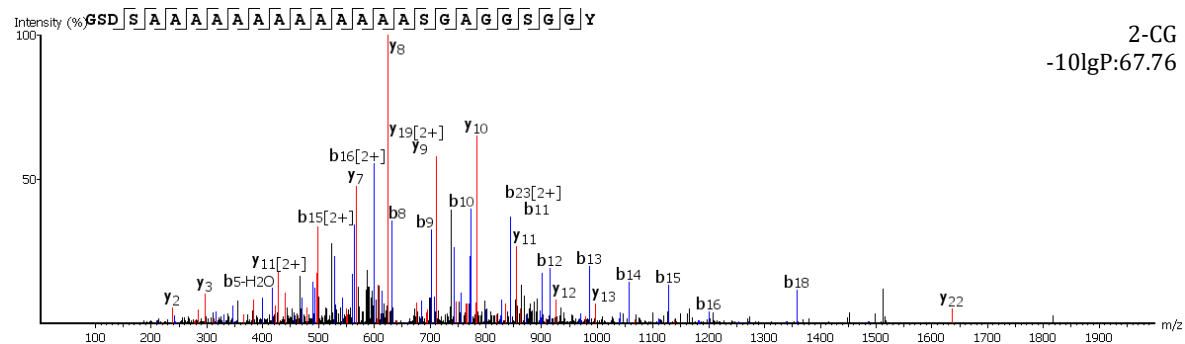
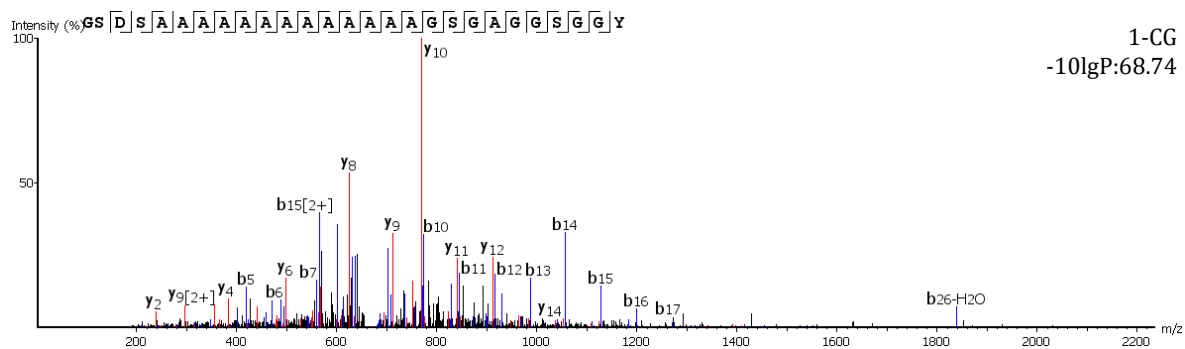
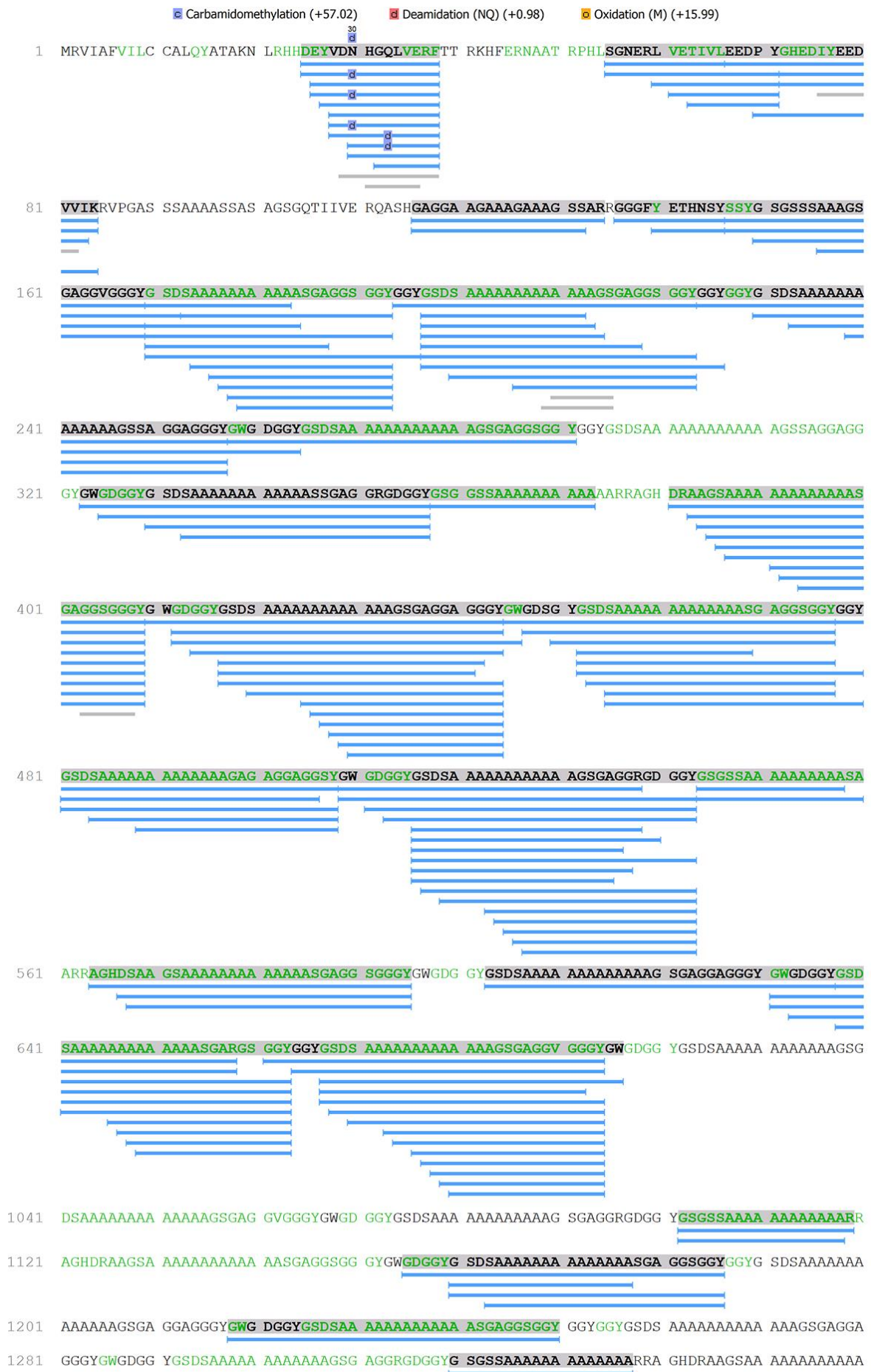


Figure 70. Spectra of chymotryptic unique peptides of *A. pernyi* fibroin: 1) GSDSA₁₃GSGAGSGGGY (m/z =680.30), 2) GSDSA₁₃SGAGSGGGY (m/z =661.30), 3) GSDSSA₁₀SGAGSGGGY (m/z =638.23), z =3



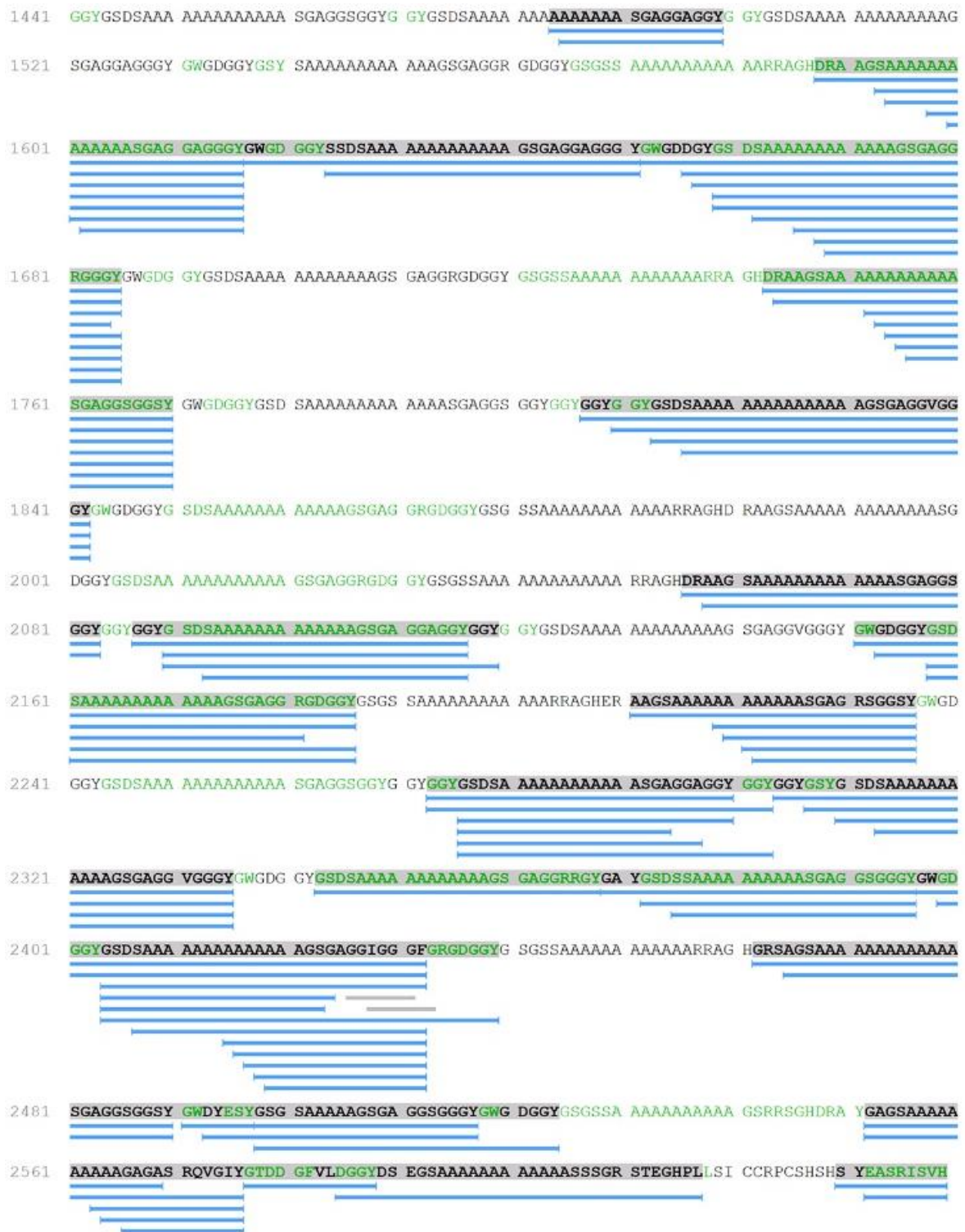


Figure 71. Protein coverage of *A. pernyi* fibroin (076786): chymotryptic in-gel digestion.



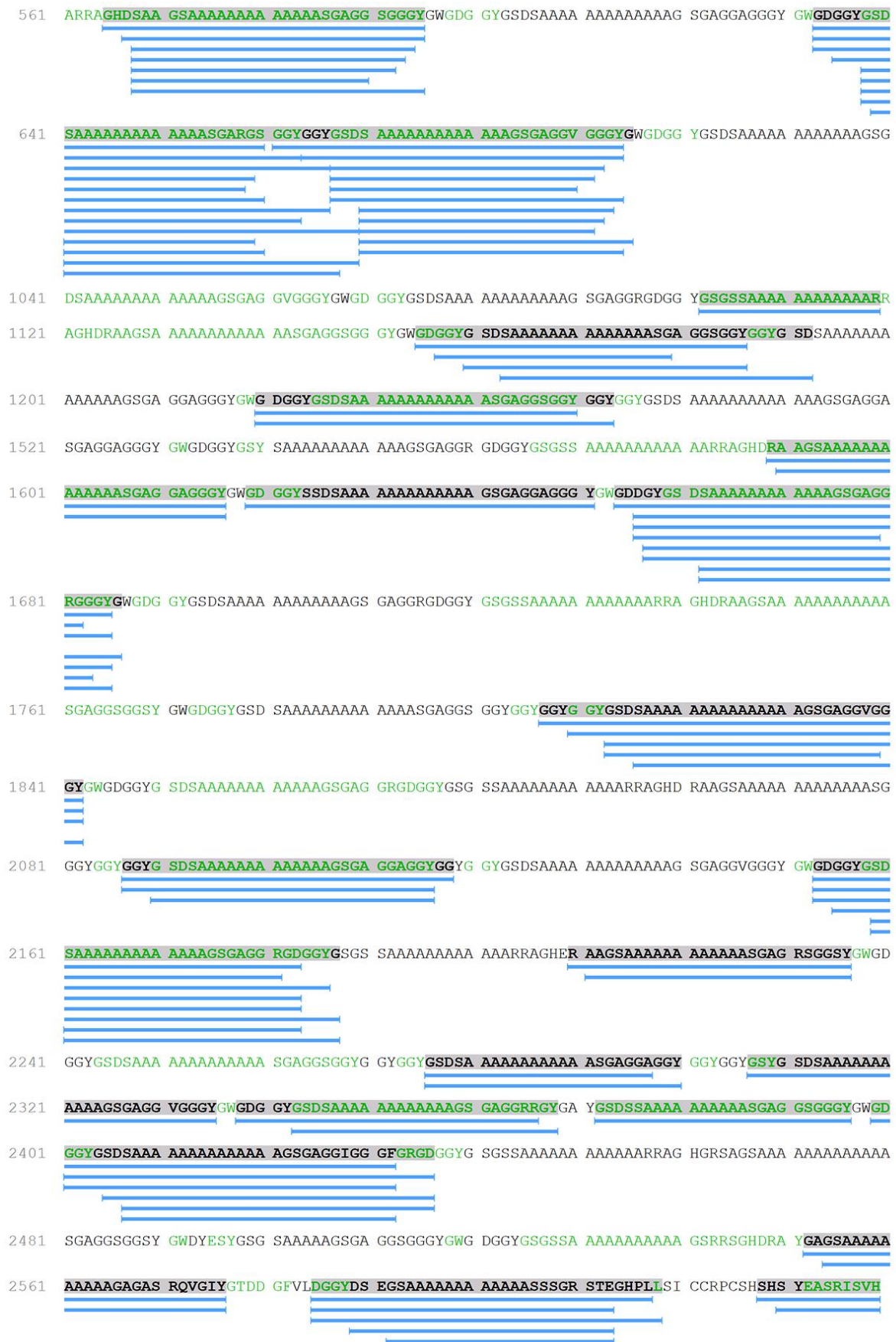


Figure 72. Protein coverage of *A. pernyi* fibroin (076786): chymotryptic in-solution digestion.

V.5.2.3 *Tryptic-chymotryptic digestion*

In tryptic-chymotryptic digestion, TCS showed better protein coverage than TCG with much higher probability score; the most probable cause of this result is thought to be the poor resolution of protein into the gel. As CS and TCS show very similar results, it is assumed that if properly resolved, TCG would have shown a similar result to that of CG. As in the chymotryptic digestion, the peptides occurring from the residues 700-1100 were not detected in tryptic-chymotryptic digestion. In Figure 74 and Figure 75, the protein sequence marked in black and green in turn indicates the tryptic-chymotryptic cleavages.

The highest scoring peptides in both samples were tryptic-chymotryptic peptides, but not unique to O76786 and shared with E1CGA3 (*A.yamamai*): (Y)DSEGSA₁₂SSSGR(S) with -10lgP score of 53.78 and (Y)GSDSA₁₃GSGAGGAGGG(Y) with -10lgP score of 47.37 were found in TCG; (Y)GSDSA₁₃GSGAGGVGGGY(G) with -10lgP score of 74.06 and (K)RVPGASSS AAAASSASAGSGQ(T) with -10lgP score of 70.17 were found in TCS. The highest scoring unique peptides in both samples had non-enzyme specific cleavage on glycine and tyrosine, but on different ends: (G)SDSA₁₄SGAGGSGGYGG(Y) with -10lgP score of 44.90 was found in TCG and (Y)GSDSSA₁₀SGAGGSGGGY(G) with -10lgP score of 68.92 was found in TCS. Some unique peptides resulted by non-enzyme specific cleavages were found in both samples. In TCG, (G)SDSA₁₄SGAGGSG(G) (1161-1184) with -10lgP score of 28.64 and (G)SA₁₂SGAGRS(G) (2214-2232) with -10lgP score of 34.64 were found. In TCS, (D)SA₁₂SSGAGGRGD(G) (333-354) with -10lgP score of 37.02, (D)SA₁₄GSGAGG IGGGFGRGD(G) (2407-2436) -10lgP score of 40.47 and (H)SYEASRISVH (2630-2639) with -10lgP score of 54.96 were found (Figure 73).

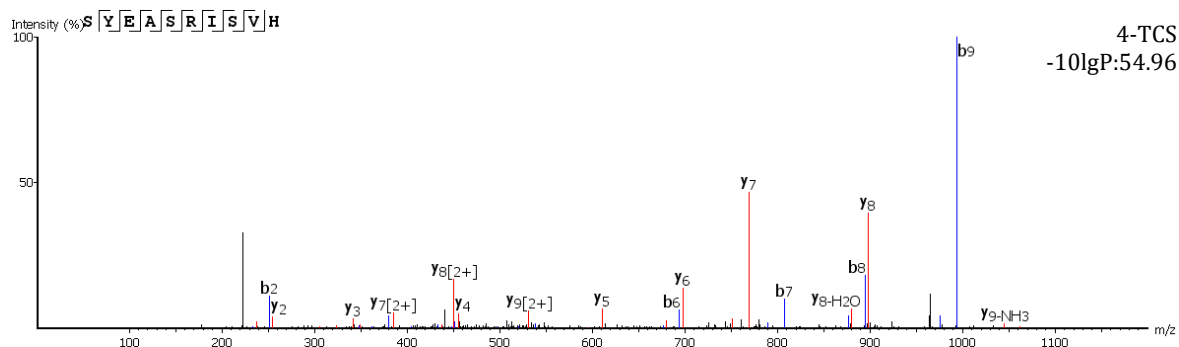
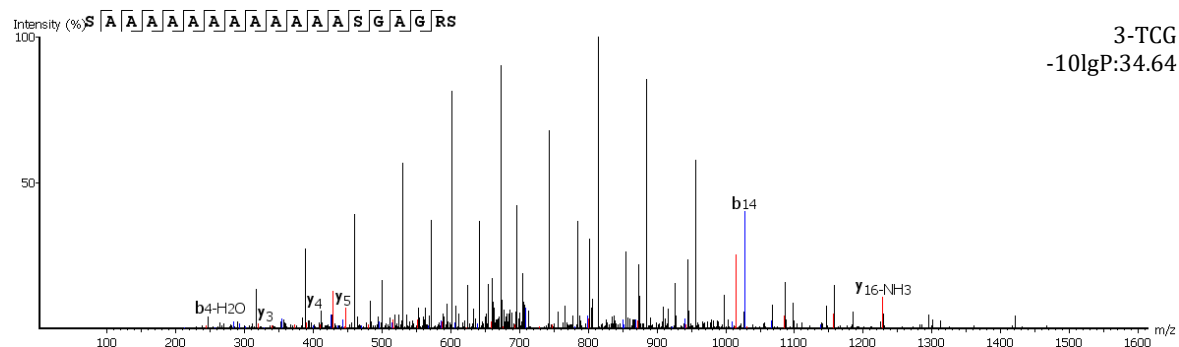
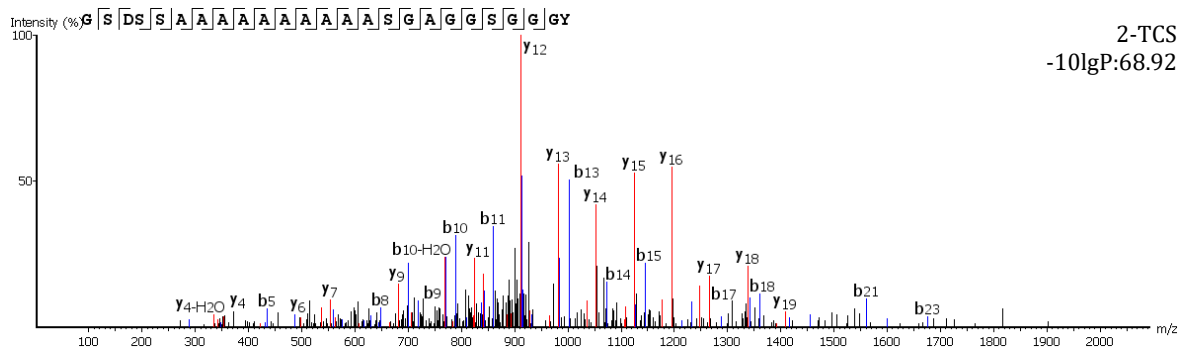
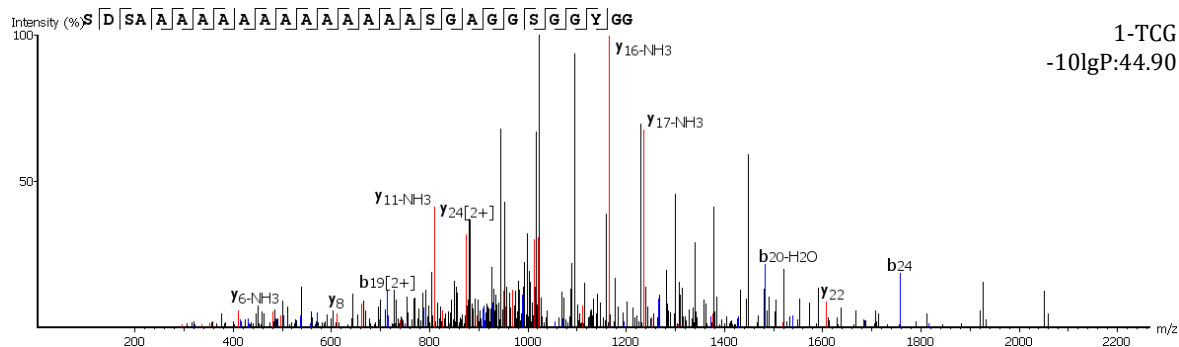


Figure 73. Spectra of tryptic-chymotryptic unique peptides of *A. pernyi* fibroin: 1) SDSA14SGAGGSGGYGG($m/z=1055.48$), 2) GSDSSA10SGAGGSGGGY ($m/z=956.92$), 3) SA12SGAGRS($m/z=737.37$), 3) SYEASRISVH ($m/z=574.79$), $z=2$

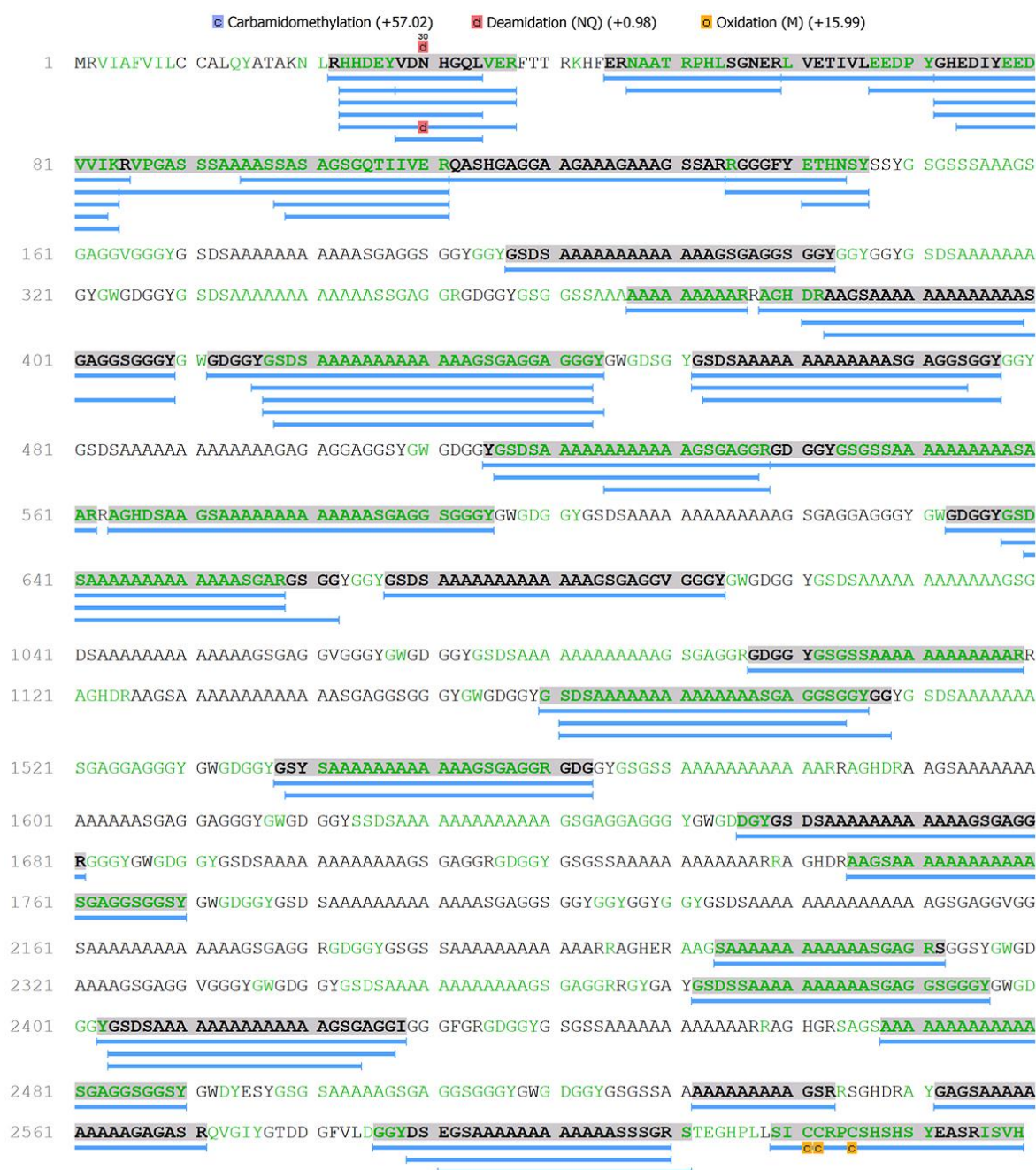


Figure 74. Protein coverage of *A. pernyi* fibroin (076786): tryptic-chymotryptic in-gel digestion.

1 MRVIAFVILC CALQYATAKN LRHHDEYVDN HGQLVERFTT RKFHERNAAT RPHLSGNERL VETIVLEEDP YGHEDIYEED

81 VVIKRVPGAS SSAAAASSAS AGSGQTIIVE RQASHGAGGA AGAAAGAAAG SSARRGGGFY ETHNSYSSYG SGSSSAAAGS

161 CAGGVGGGYG SDSAAAAAAA AAAASGAGGS GGYGGYGSDS AAAAAAAAAA AAAGSGAGGS GGYGGYGGYG SDSAAAAAAA

241 AAAAAAGSSA GGAGGGYGWG DGGYGSDSAA AAAAAAAAAA AGSGAGGSGG YGGYGSDSAA AAAAAAAAAA AGSSAGGAGG

321 GYGWGDGGYG SDSAAAAAAA AAAAASSGAG GRGDGGYGSG GSSAAAAAAA AAAAARRAGH DRAAGSAAAA AAAAAAAS

401 GAGSGGGYG WGDGGYGSDS AAAAAAAAAA AAAGSGAGGA GGGYGWGDSC YGSDSAAAAA AAAAAAASG AGSGGGYGGY

481 GSDSAAAAAA AAAAAAGAG AGGAGGSYGW GDGGYGSDSA AAAAAAAAAA AGSGAGGRGD GYGSGSSAA AAAAAAAS

561 ARRAGHDSAA GSAAAAAAA AAAAASGAGG SGGGYGWGDG YGSDSAAAA AAAAAAAG SGAGGAGGGY WGDGGYGSD

641 SAAAAAAA AAAASGARGS GGYGGYGSDS AAAAAAAAAA AAAGSGAGGV GGGYGWGDG YGSDSAAAA AAAAAAGSG

1041 DSAAAAAAA AAAAGSGAG GVGGYGWG DGGYGSDSAA AAAAAAAG SGAGGRGDG YGSGSSAAA AAAAAAARR

1121 AGHDRAGSA AAAAAAAA AASGAGGSGG YGWGDGGYG SDSAAAAAAA AAAAAASGA GSGGGYGGYG SDSAAAAAAA

1201 AAAAAAGSGA GGAGGGYGWG DGGYGSDSAA AAAAAAAAAA ASGAGGSGGY GGYGGYGSDS AAAAAAAA AAAGSGAGGA

1521 SGAGGAGGGY WGDGGYGSY SAAAAAAA AAAGSGAGGR GDGGYGSSS AAAAAAAA AARRAGHRA AGSAAAAAAA

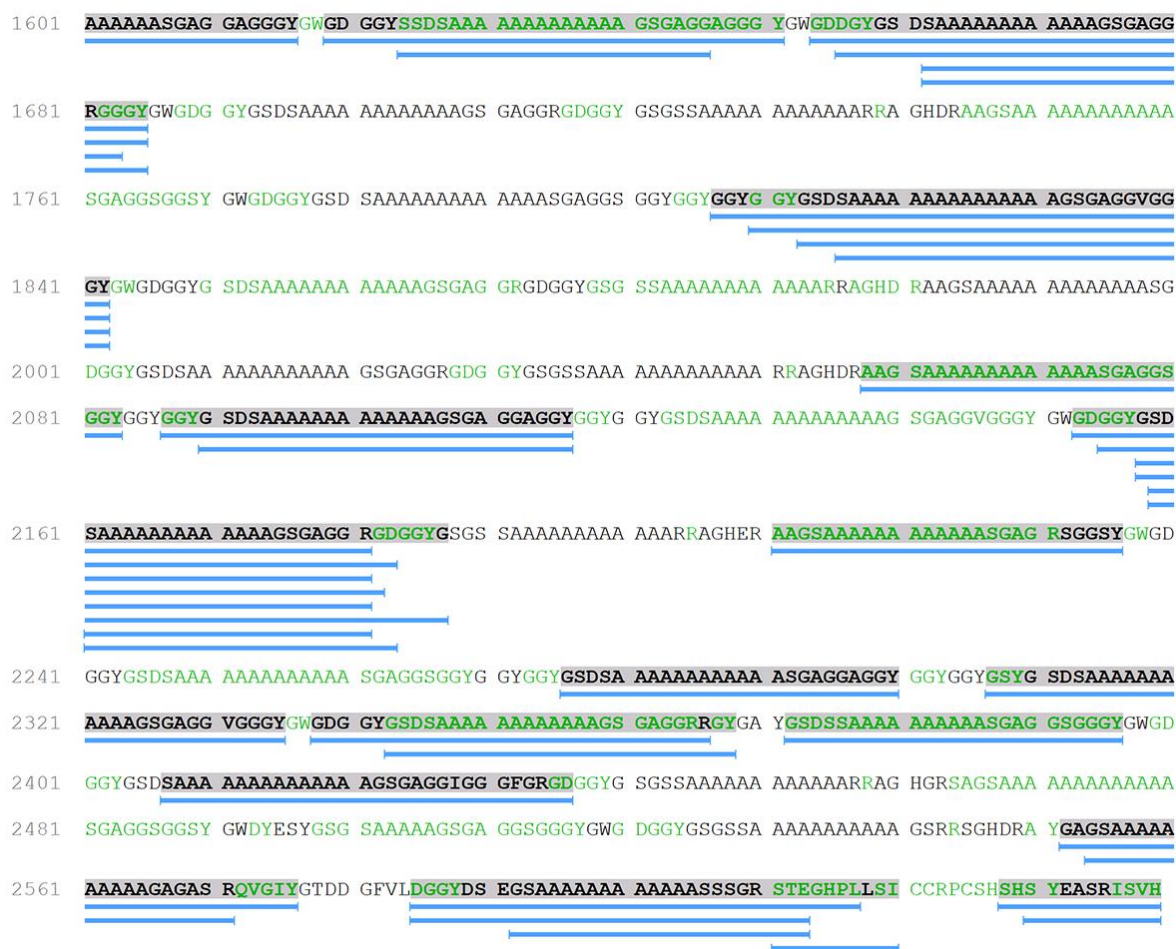


Figure 75. Protein coverage of *A. pernyi* fibroin (076786): tryptic-chymotryptic in-solution digestion.

V.5.3 *Antheraea mylitta*

In *A. mylitta* silk, samples showed better protein coverages compare to other silk fibroins, given that the sequence is only partial of 507 residues. Although the highest protein probability score resulted in tryptic digestion, the highest protein coverage was achieved by chymotryptic digestion. However, the tryptic-chymotryptic samples showed poor results than projected: this might have resulted from the poor resolution of the protein in SDS-PAGE process and less amount of protein to begin with (Table 27).

In some of the samples, the highest protein matches were not *A. mylitta*: in TG and TCS it was *A. yamamai*, and in TCG it was *A. assamensis*; *A. yamamai* was the second choice of protein match in TS, CG and CS samples. However, most of the unique peptides corresponding to these other *Antheraea* fibroins were occurring after the 507th residue or occurring multiple times throughout the sequence. Thus, in the identification of *A. mylitta* silk from an unknown source, it would be beneficial to manually review the peptides and trace the unique markers that differentiate this subject protein from other *Antheraea* fibroins.

Table 27. PEAKS analysis results of *A. mylitta* fibroin treated with different enzymes

	Trypsin		Chymotrypsin		Try-Chymo	
	In-gel	In-sol	In-gel	In-sol	In-gel	In-sol
<i>Antheraea mylitta</i>	(Q8ISB3 ANTMY)					
-10lgP	247.88	230.28	237.67	213.84	145.71	197.90
Coverage (%)	17	27	72	46	29	33
#Peptides	17	24	145	47	17	18
#Unique	2	6	67	20	8	5

V.5.3.1 Tryptic digestion

TG showed tryptic peptides near C-terminus and TS showed more protein coverage and more peptides with non-enzyme specific cleave on one end. The highest scoring peptide in TG was a tryptic unique peptide QASHGAGGAAGAAAGAAASSSVR (111-113) with -10lgP score of 73.56. The highest scoring peptide was also a unique tryptic peptide GAGGAAGAAAGAAASSSVR, occurred from a similar location (115-133) with -10lgP score of 65.52, followed by QASHGAGGAAGAAAGAAASSSVR (111-113) with and without deamidation of glutamine (-10lgP score= 56.41/ 54.78, respectively). In TS, a peptide with random cleavages on both ends (A)₈SGAGGRGD(G) (251-265) with -10lgP score of 36.37 was found; however, this was not unique to Q8ISB3 (Figure 76). In Figure 77 and Figure 78, the protein sequence of Q8ISB3 is marked in black and green in turn by the tryptic cleavages.

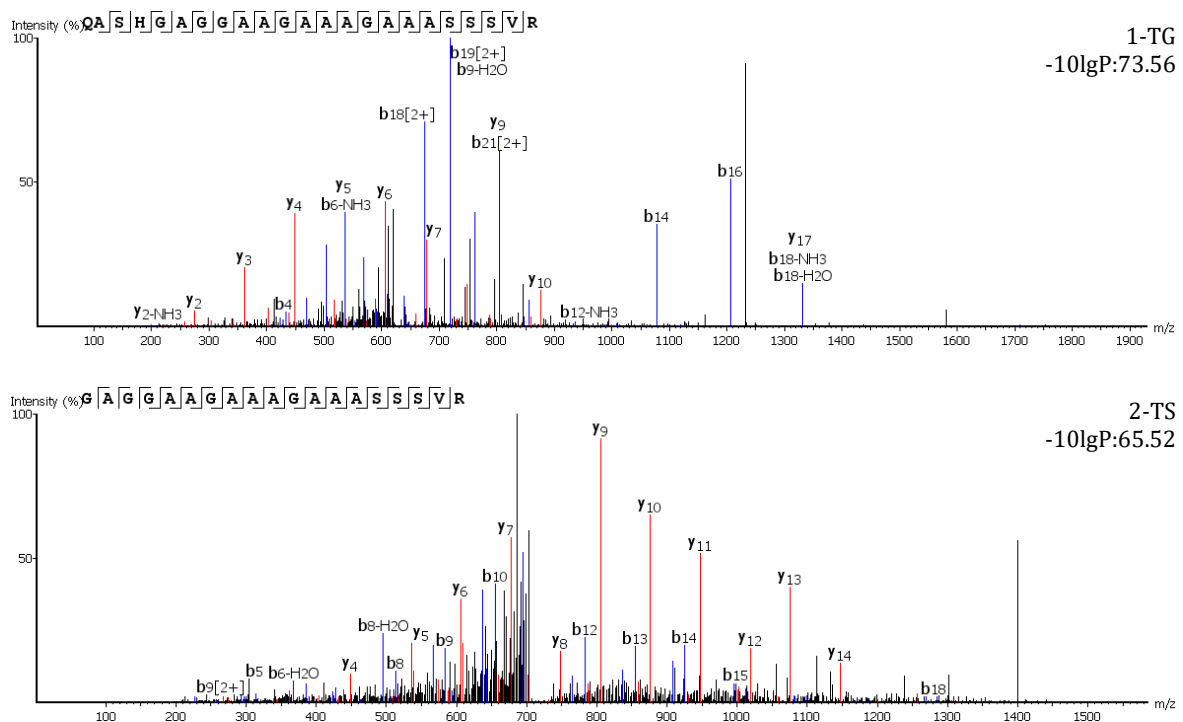


Figure 76. Spectra of tryptic unique peptides of *A. pernyi* fibroin: 1) GAGGAAGAAAGAAASSSVR ($m/z=628.31$, $z=3$), 2) QASHGAGGAAGAAAGAAASSSVR ($m/z=730.37$, $z=2$)

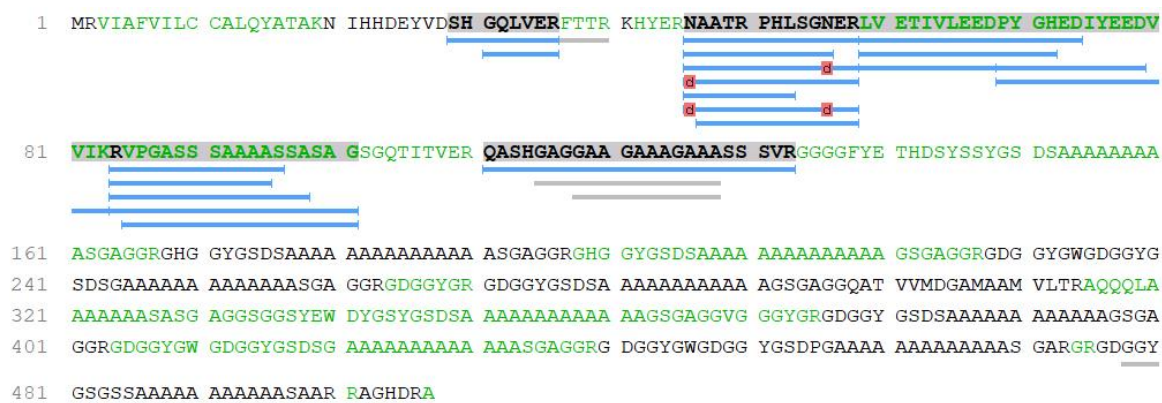


Figure 77. Protein coverage of *A. mylitta* fibroin (Q8ISB3): tryptic in-gel digestion.

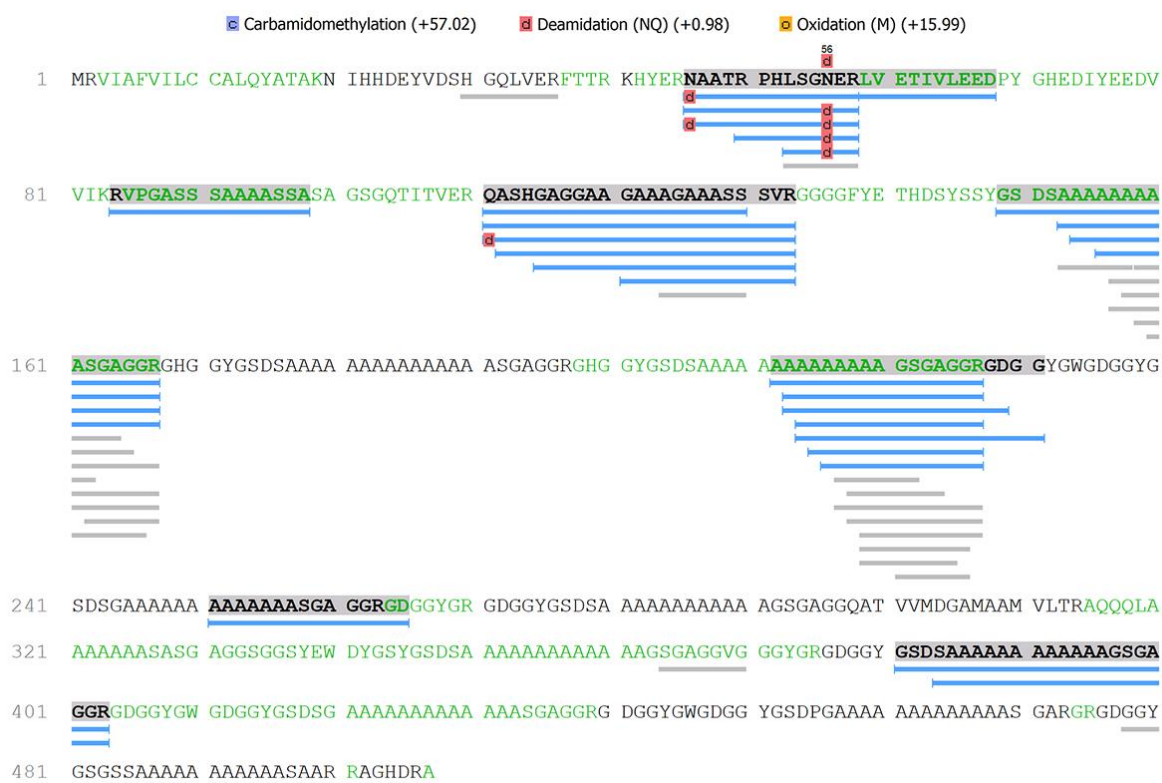


Figure 78. Protein coverage of *A. mylitta* fibroin (Q8ISB3): tryptic in-solution digestion.

V.5.3.2 *Chymotryptic digestion*

CG showed a desirable result with 72% protein coverage with -10lgP score of 261.09. Q8ISB3 was the first choice of identification, and one of the unique peptides with one specific cleavage on C-terminus carried the last residue of the signal peptide: **DEYVDSHGQLVERF** (24-37) with -10lgP score of 34.35. The highest scoring peptide in this sample was a chymotryptic unique peptide VDSHGQLVERF (27-37), yielded -10lgP score of 52.01. Most of the peptides were cleaved by the chymotrypsin rule; however, some peptides were cleaved after alanine (A) and aspartic acid(D) in C-terminus. In Figure 80 and Figure 81, the protein sequence of Q8ISB3 is marked in black and green in turn by the chymotryptic cleavages.

In CS, the highest scoring unique peptide was a chymotryptic peptide GRGDGGYGSDSA₁₂GSGAGGRGDGGY (374-468) with -10lgP score of 46.01. There were numbers of peptides yielded by random cleavages at both ends and some of them were unique to Q8ISB3. For example, (D)GGYGS(D)SGA₁₃SGAGGRGD(G)GYG(R) (237-269) yielded two peptides both by non-enzyme specific cleavages on both ends; (A)(G)GVGGGYGRGDGGYGSD(S)(367-383) also yielded two peptides by non-enzyme specific cleavages on both ends. Most of the non-enzyme specific cleavages were occurred after aspartic acid (D) and before glutamine (G) and arginine (R). Among them, GGYGSDSGA₁₃SGAGGRGD showed the highest probability score; -10lgP score of 42.08 (Figure 79).

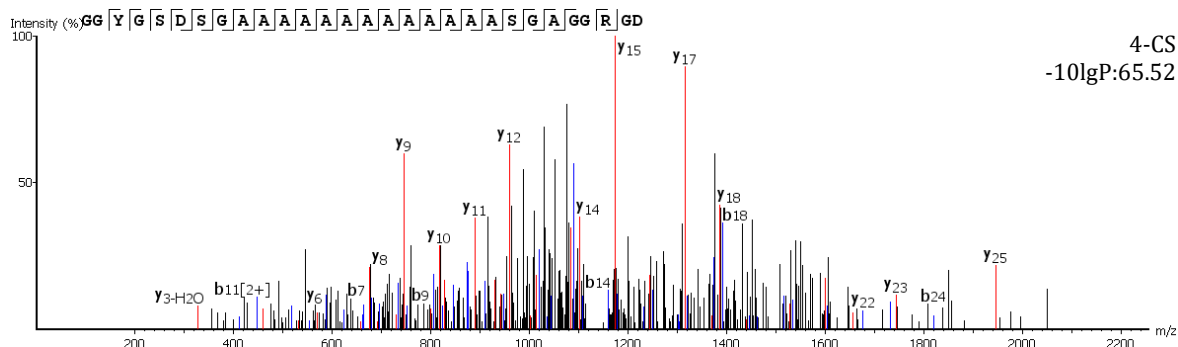
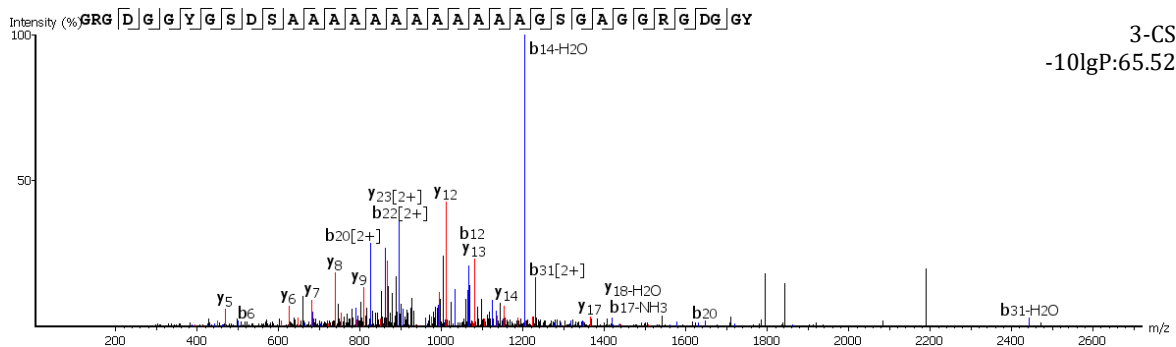
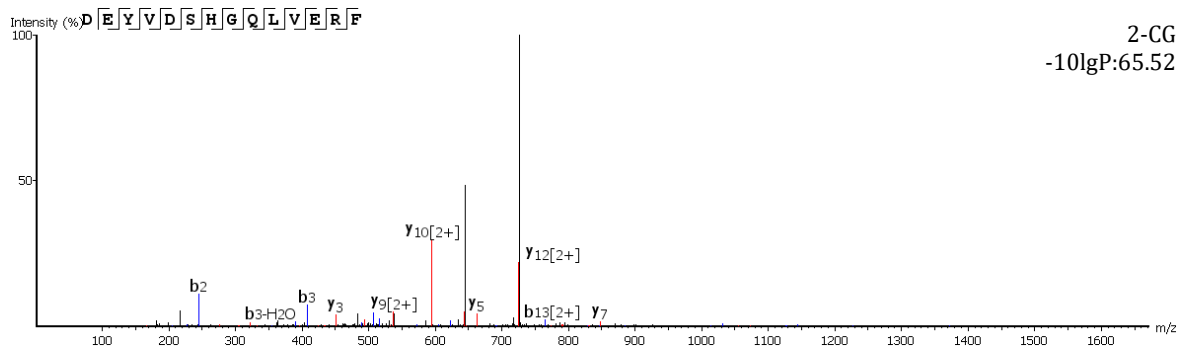
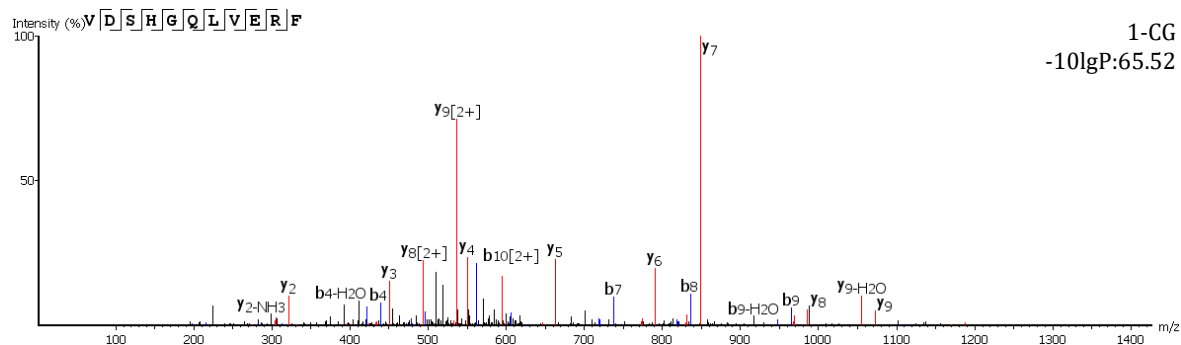


Figure 79. Spectra of chymotryptic unique peptides of *A. mylitta* fibroin: 1) VDSHGQLVERF ($m/z=643.83$, $z=2$), 2) DEYVDSHGQLVERF ($m/z=565.27$, $z=3$), 3) GRGDGGYGSDSA₁₂GSGAGRGDGGY ($m/z=957.76$, $z=2$), 4) GGYGSDSGA₁₃SGAGGRGD ($m/z=1140.52$, $z=2$)



Figure 80. Protein coverage of *A. mylitta* fibroin (Q8ISB3): chymotryptic in-gel digestion.



Figure 81. Protein coverage of *A. mylitta* fibroin (Q8ISB3): chymotryptic in-solution digestion.

V.5.3.3 Tryptic-chymotryptic digestion

Despite the better theoretical digestion result, tryptic-chymotryptic digestion did not yield the best protein coverages of *A. mylitta* than chymotryptic digestion. Also, in both samples the best protein identification match was not *A. mylitta*—while both tryptic and chymotryptic digestion results showed *A. mylitta* as the most confident match. TCG showed slightly higher probability scores for *A. assamensis* (-10lgP=162.7) and *A. pernyi* (-10lgP=145.89) than *A. mylitta*; which was the 3rd match for protein identification with -10lgP score of 145.71 and 29% protein sequence coverage. TCS also showed a slightly higher probability score for *A. yamamai* (-10lgP=205.12) than *A. mylitta*, which showed -10lgP score of 197.9 with 36% protein coverage. The protein matches to other *Antheraea* fibroins with higher probability scores implies that the combination of trypsin and chymotrypsin attacks more efficiently the unsequenced part of the protein.

In TCG, the highest scoring peptide was (Y)GSDSGA₁₃SGAGGR(G) (240-263) with -10lgP score of 38.27, which is unique to Q8ISB3. However in TCS, the highest scoring peptide was (K)RVPGASSSAAAASSASAGSGQ(T) with -10lgP score of 68.65, which is shared among *A. pernyi*, *A. mylitta* and *A. yamamai*: the highest scoring unique peptide to Q8ISB3 was (Y)GSDSA₉SGAGGR(G) (149-167) with -10lgP score of 48.96. In both samples, the highest scoring unique peptides were cleaved by the combination of chymotrypsin and trypsin. The peptides with both non-enzyme specific ends were only found in TCS: (D)SA₁₂GSGAGGRGD(G) (384-405) and (D)SGA₁₃SGAGGRGD(G) (243-265) were both cleaved by aspartic acid; however, only the latter is unique to Q8ISB3 (Figure 82). In Figure 83 and Figure 84, the protein sequence of Q8ISB3 is marked in black and green in turn by the tryptic-chymotryptic cleavages.

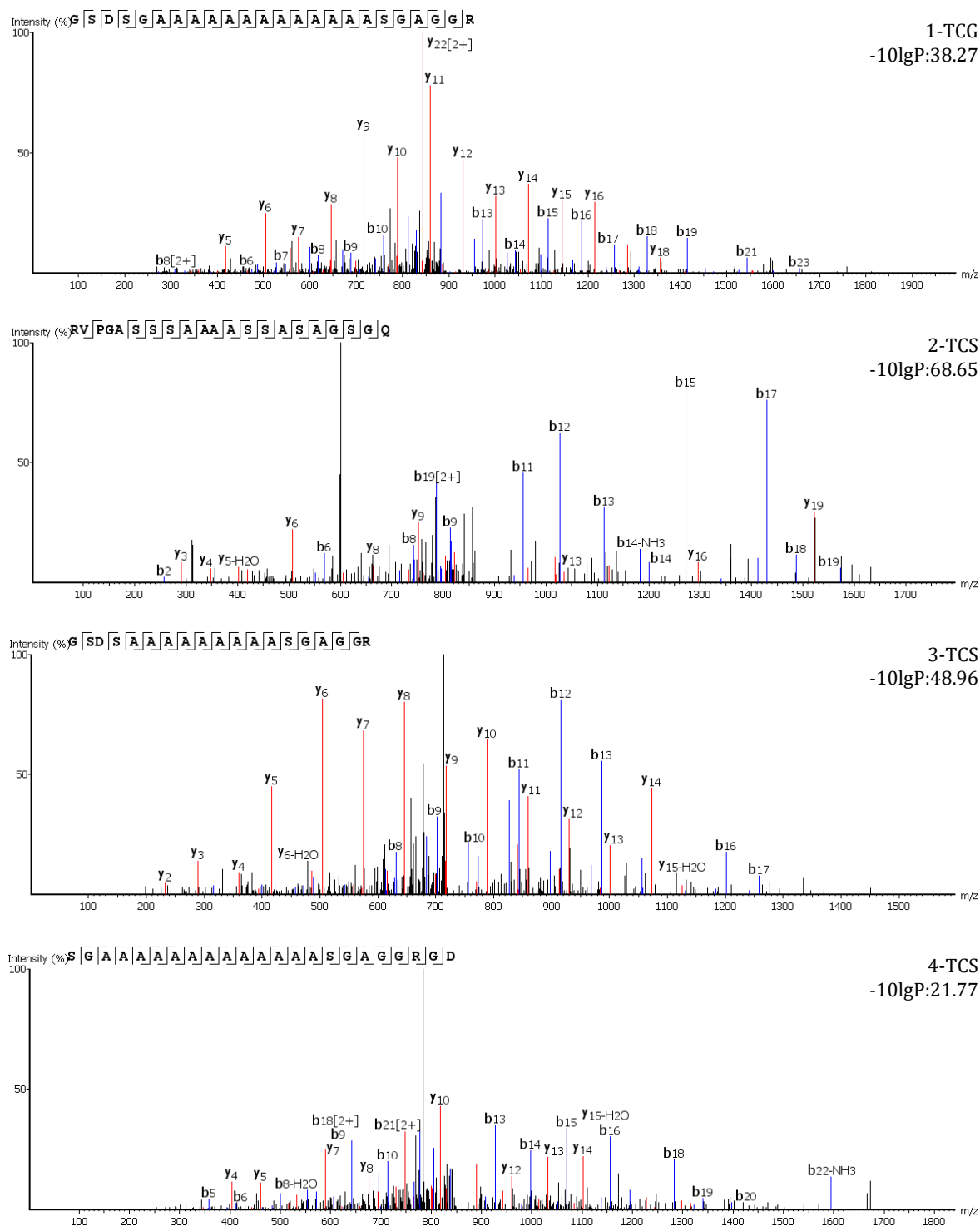


Figure 82. Spectra of tryptic-chymotryptic peptides of *A. mylitta* fibroin: 1) GSDSGA₁₃SGAGGR ($m/z=915.94$), 2) RVPGASSSAAAASSASAGSGQ ($m/z=888.93$), 3) GSDSA₉SGAGGR ($m/z=745.35$), 4) SGA₁₃SGAGGRGD ($m/z=872.42$), $z=2$

1 MRVIAFVILC CALQYATAKN IHHDEYVDSH GQLVERFTTR KHYERNAATR PHLSGNERLV ETIVLEEDPY GHEDIYEEDV
81 VIKRVPGASS SAAAASSASA GSGQTITVER QASHGAGGAA GAAAGAAASS SVRGGGGFYE THDSYSSYGS DSAAAAAAAAA
161 ASGAGGRGHG GYGSDSAAAA AAAAAAAAAA ASGAGGRGHG GYGSDSAAAA AAAAAAAAAA GSGAGGRGDG GYGWGDGGYG
241 SDSGAAAAAA AAAAAASGA GGRGDGGYGR GDGGYGSDSA AAAAAAAAAA AGSGAGGQAT VMDGAMAAM VLTRAQQQLA
321 AAAAAASAG AGSGSGSYEW DYGSYGSDSA AAAAAA AAAAAGSGAGGVG GGYGRGDGGY GSDSAAAAAA AAAAAAGSGA
401 GGRGDGGYGW GDGGYGSDSG AAAAAAAAAA AAASGAGGRG DGGYGWGDGG YGSDPGAAAA AAAAAAAS GARGRGDGGY
481 GSGSSAAAAA AAAAAASAAR RAGHDRA

Figure 83. Protein coverage of *A. mylitta* fibroin (Q8ISB3): tryptic-chymotryptic in-gel digestion.

1 MRVIAFVILC CALQYATAKN IHHDEYVDSH GQLVERFTTR KHYERNAATR PHLSGNERLV ETIVLEEDPY GHEDIYEEDV
81 VIKRVPGASS SAAAASSASA GSGQTITVER QASHGAGGAA GAAAGAAASS SVRGGGGFYE THDSYSSYGS DSAAAAAAAAA
161 ASGAGGRGHG GYGSDSAAAA AAAAAAAAAA ASGAGGRGHG GYGSDSAAAA AAAAAAAAAA GSGAGGRGDG GYGWGDGGYG
241 SDSGAAAAAA AAAAAASGA GGRGDGGYGR GDGGYGSDSA AAAAAAAAAA AGSGAGGQAT VMDGAMAAM VLTRAQQQLA
321 AAAAAASAG AGSGSGSYEW DYGSYGSDSA AAAAAA AAAAAGSGAGGVG GGYGRGDGGY GSDSAAAAAA AAAAAAGSGA
401 GGRGDGGYGW GDGGYGSDSG AAAAAAAAAA AAASGAGGRG DGGYGWGDGG YGSDPGAAAA AAAAAAAS GARGRGDGGY
481 GSGSSAAAAA AAAAAASAAR RAGHDRA

Figure 84. Protein coverage of *A. mylitta* fibroin (Q8ISB3): tryptic-chymotryptic in-solution digestion.

V.5.4 *Antheraea yamamai*

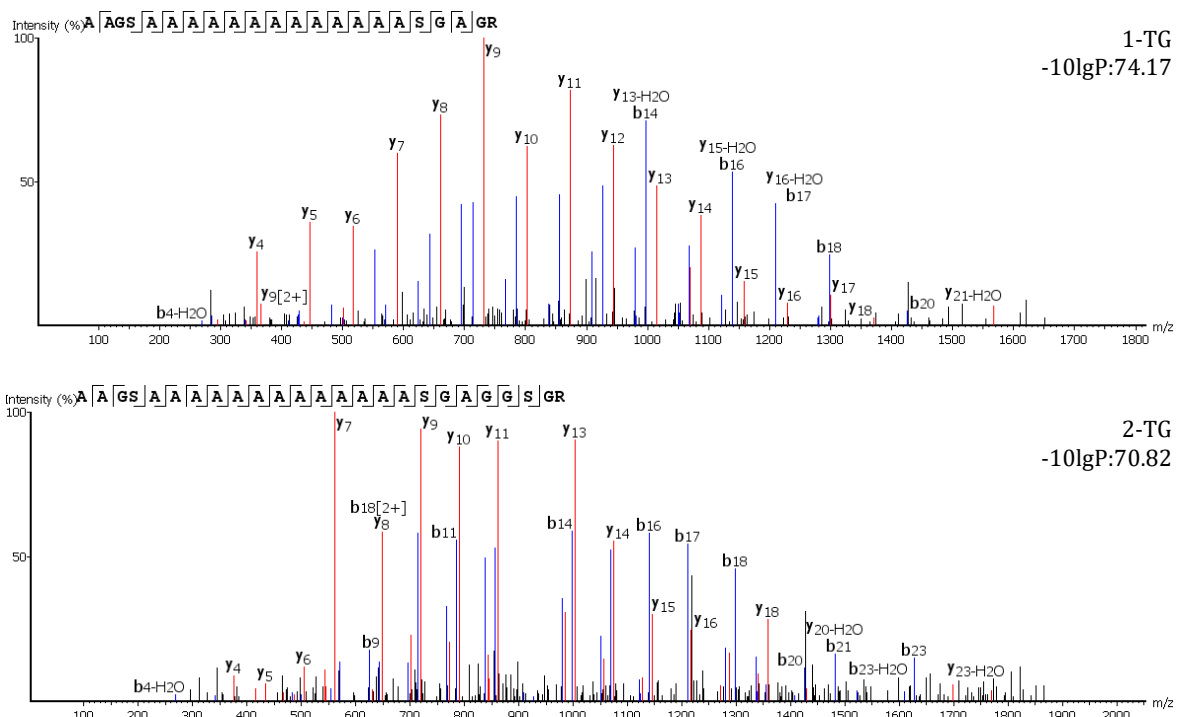
In *A. yamamai* silk, tryptic digestion showed relatively good result compare to other tasar silks, *A. pernyi* and *A. mylitta*: TG showed 9% protein coverage with -10lgP score of 263.54, and TS showed 16% protein coverage with -10lgP score of 322.75. The chymotryptic digestion showed the highest protein coverages: CG showed 24% protein coverage with -10lgP score of 279.90 and CS showed 26% protein coverage with -10lgP score of 244.46. The tryptic-chymotryptic digestion showed better protein coverages than the tryptic digestion, however no better than the chymotryptic digestion: TCG showed relatively low probability score (-10lgP=182.54) compare to other samples regardless of 16% protein coverage— this is thought to have resulted from the poor protein resolution into the gel. TCS showed 23% protein coverage with -10lgP score of 245.64. Thus, in projection, TCG would have performed similarly to CG if appropriately prepared (Table 28).

Table 28. PEAKS analysis results of *A. yamamai* fibroin treated with different enzymes

	Trypsin		Chymotrypsin		Try-Chymo	
	In-gel	In-sol	In-gel	In-sol	In-gel	In-sol
<i>Antheraea yamamai</i>	(E1CGA3 ANTYA)					
-10lgP	263.54	322.75	279.90	244.46	182.54	245.63
Coverage (%)	9	14	24	26	16	23
#Peptides	33	63	63	135	26	92
#Unique	4	18	28	40	13	26

V.5.4.1 Tryptic digestion

In both samples, two unique tryptic peptides (R)AAGSA₈SGAGR(S) at the residues 908-932 and (R)AAGSA₁₃SGAGGSGR(G) at the residues 2351-2372 yielded the highest probability score (Figure 85). In TG, most of the peptides were tryptic; however, one unique peptide peptides cleaved by non-enzyme specific ends was found: (G)SSA₁₃ASGAGG(S) at the residues 2018-2037 with -10lgP score of 36.53. In TS, more numbers of unique peptides were created by non-enzyme specific cleavages: for instance, (D)SAAGA₁₀SGAGGRGD(G) (1990-2011) with -10lgP score of 51.01 was cleaved by aspartic acid and glycine, (G)SGQTIIVERQASHGAGGAAGAAAGAAAGSSARGG(S) (103-136) with -10lgP score of 41.06 was cleaved by glycine and serine, and (D)SA₁₃GSGAGGAGGD(Y) (642-665) with -10lgP score of 32.87 was cleaved by aspartic acid and tyrosine. Although TS yielded more numbers of peptides from the residues 260-880 than TG, the block between the residues 933-1899 remained unaffected in both samples (Figure 86, Figure 87).



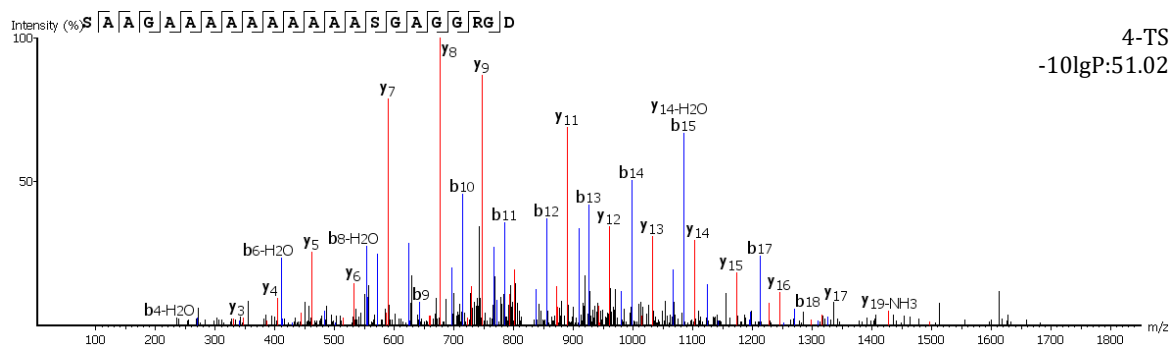
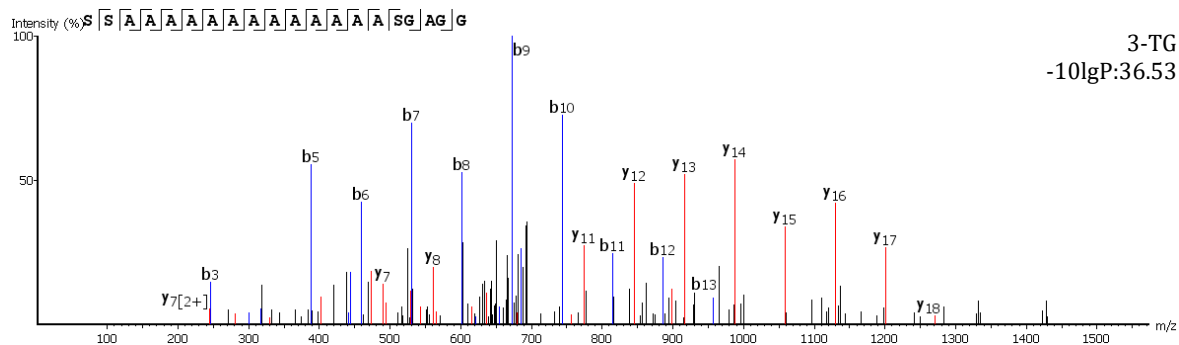
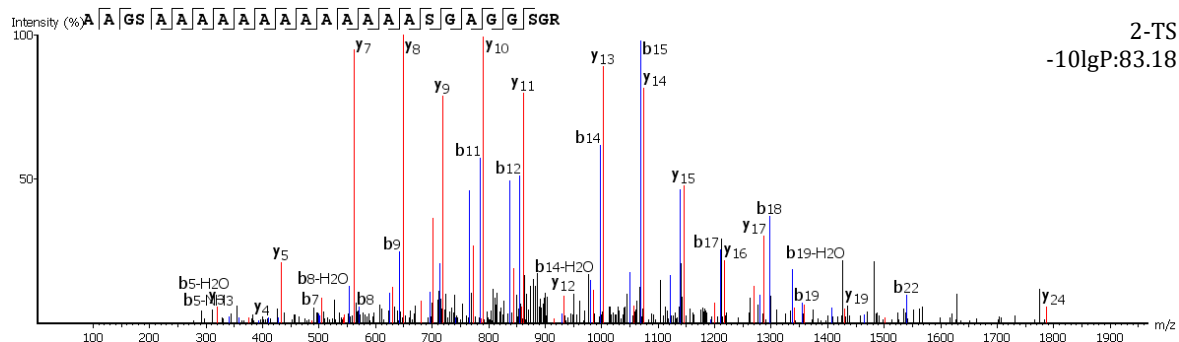
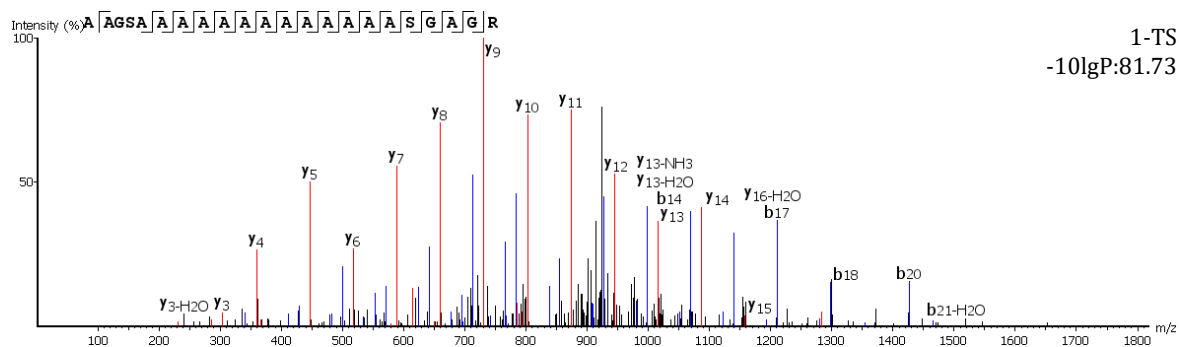


Figure 85. Spectra of tryptic unique peptides of *A. yamamai* fibroin: 1) AAGSA₈SGAGR ($m/z=828.92$), 2) AAGSA₁₃SGAGGSGR ($m/z=929.46$), 3) SAAGA₁₀SGAGGRGD ($m/z=723.36$), 4) SAAGA₁₀SGAGGRGD ($m/z=836.9$), $z=2$

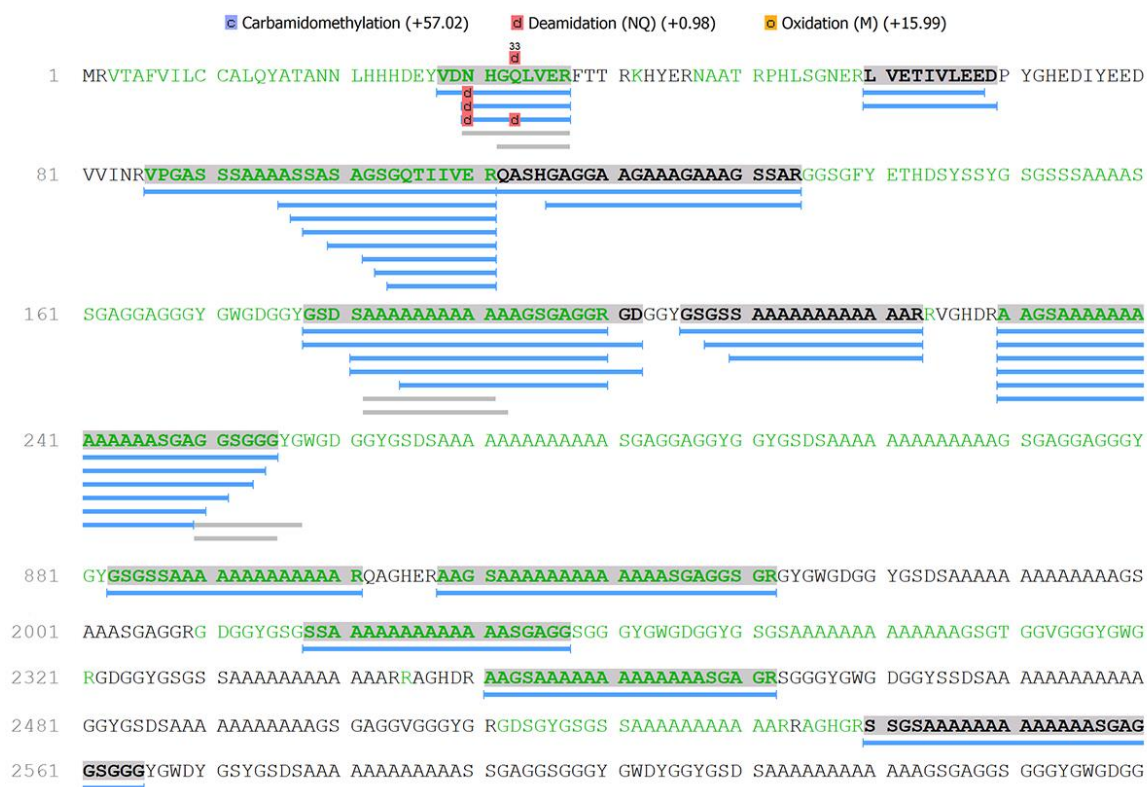


Figure 86. Protein coverage of *A. yamamai* fibroin (E1CGA3): tryptic in-gel digestion.

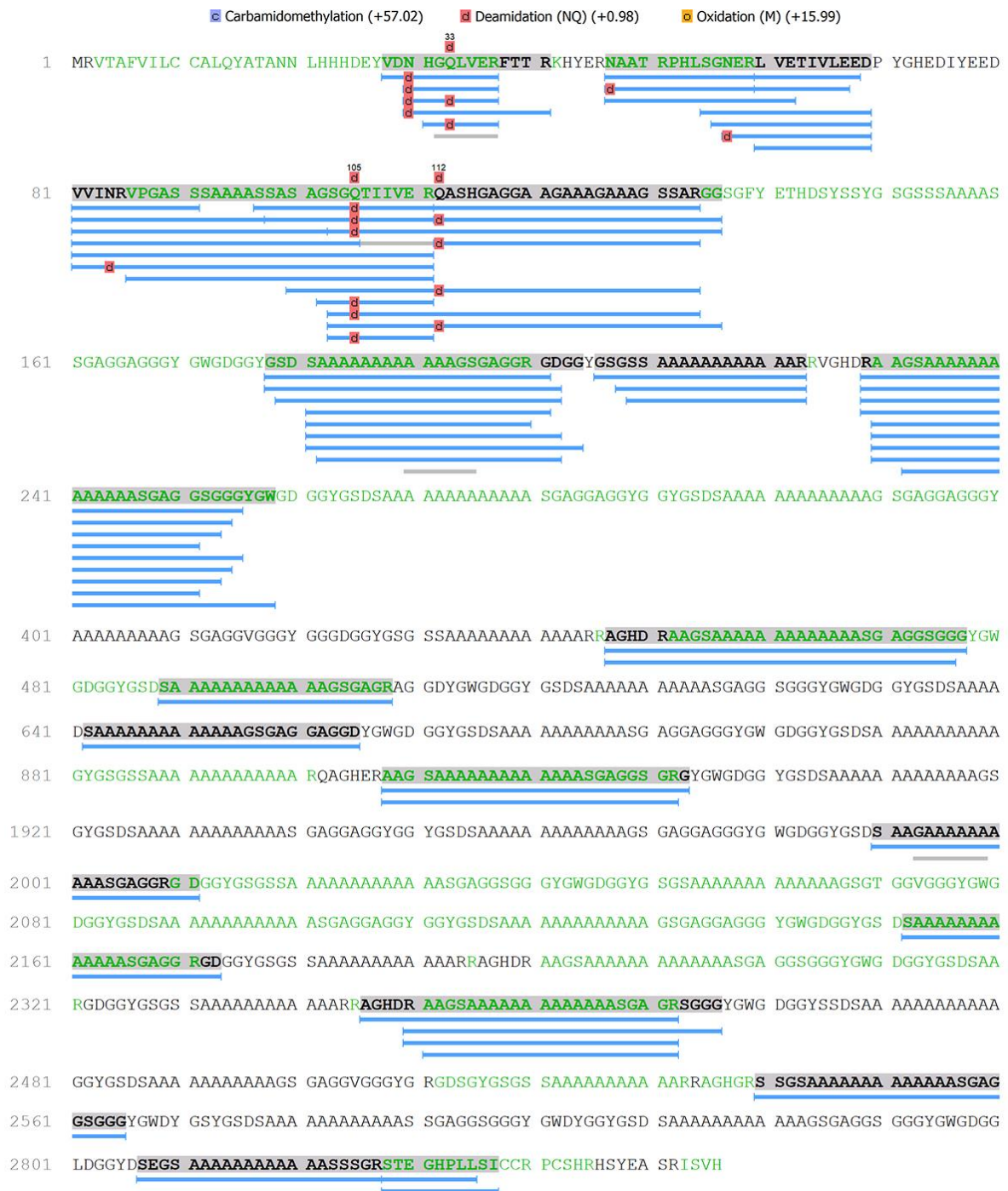
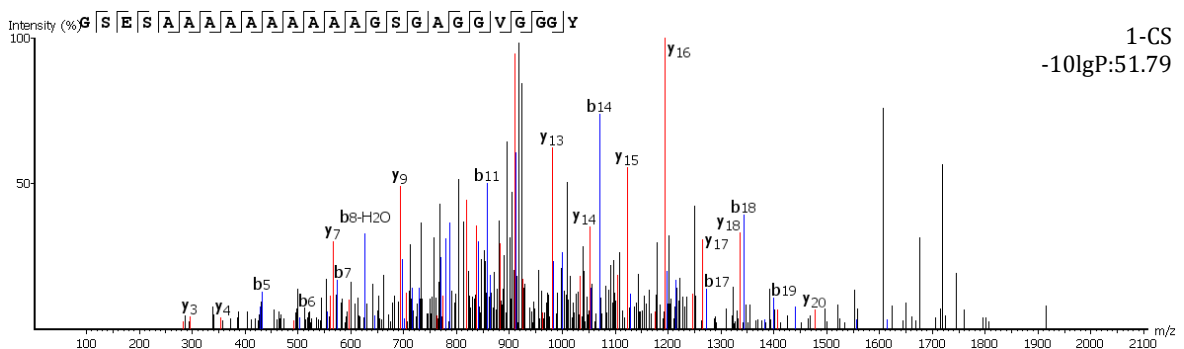
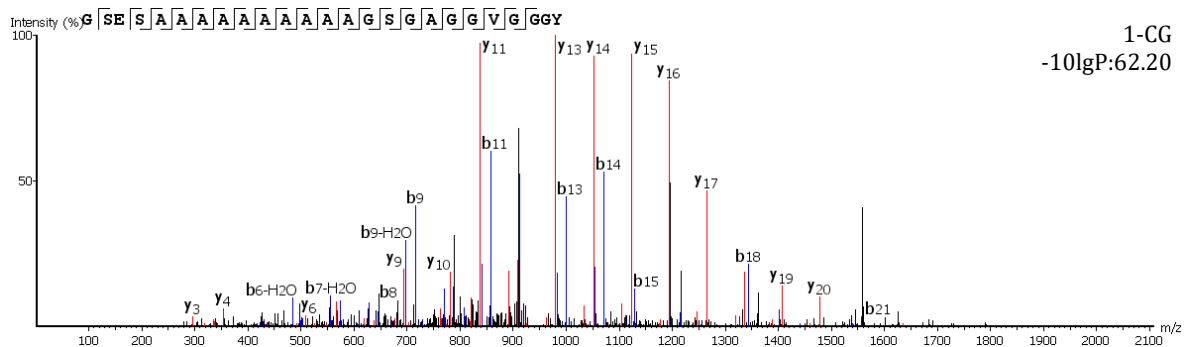


Figure 87. Protein coverage of *A. yamamai* fibroin (E1CGA3): tryptic in-solution digestion.

V.5.4.2 Chymotryptic digestion

In CG, all detected peptides were a chymotryptic peptide. The highest scoring peptides in both samples were not unique to E1CGA3. The highest scoring unique peptide in CG was (Y)GSESA₁₀GSGAGGVGGGY(G) (396-420) with -10lgP score of 62.20. In CS, the same peptide was showed the highest probability score (-10lgP=51.79) among the unique peptide to E1CGA3. CS showed some peptides with non-enzyme specific cleavages in both ends: (D)VVINRVPGASSSA₄SSASAG(S) at the residues 81-102, (D)GGY(G)S(D)SA₁₃ SGAGGRGD(G) at the residues 2146-2173 and (D)SA₁₁GSGAGGVGGGYGRGD(S) at the residues 2487-2513 were cleaved after aspartic acid and glycine (Figure 88). The distribution of the peptides was similar in both samples; however, CS showed more missed cleavages from the same sequences (Figure 89, Figure 90).



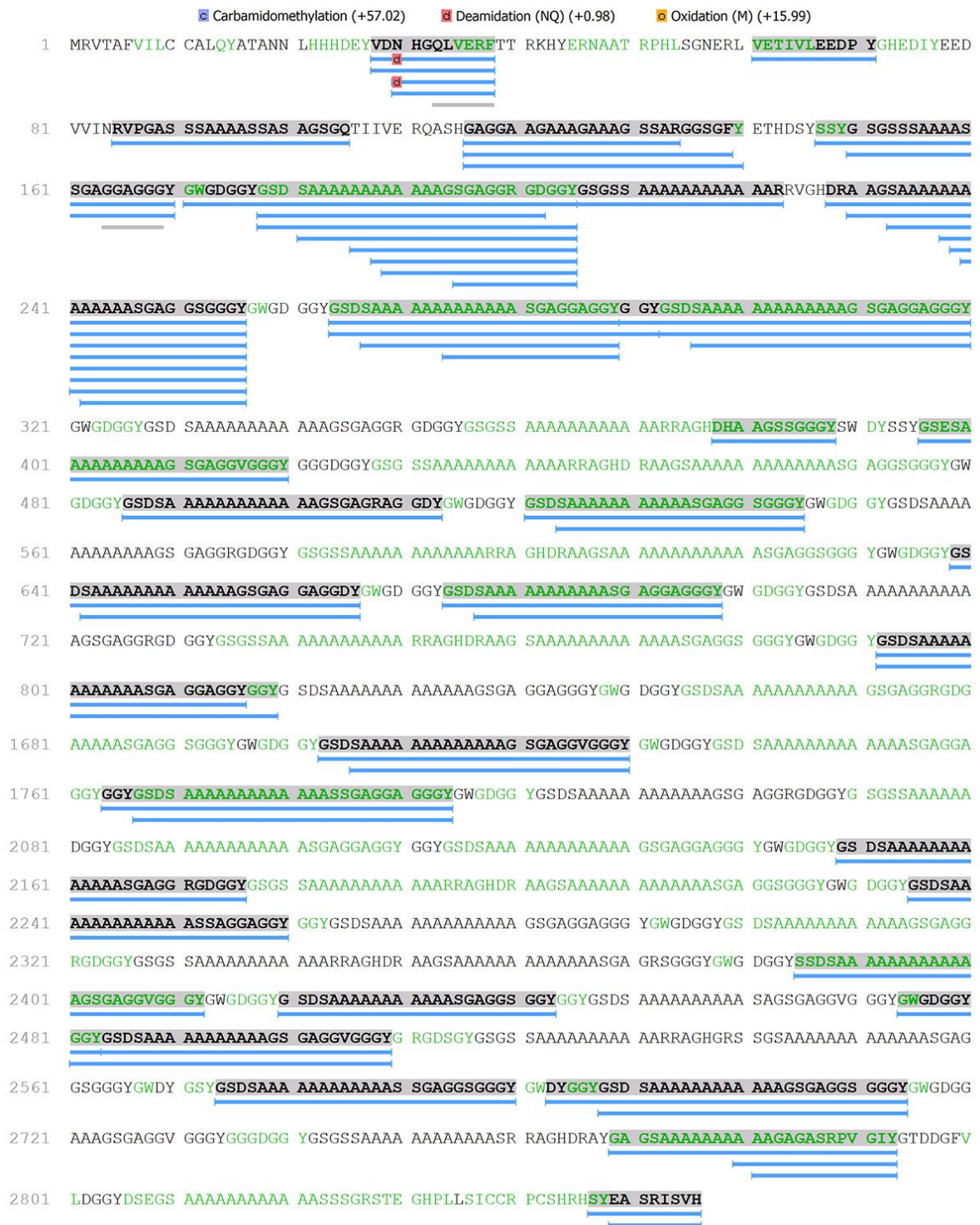


Figure 89. Protein coverage of *A. yamamai* fibroin (E1CGA3): chymotryptic in-gel digestion.





Figure 90. Protein coverage of *A. yamamai* fibroin (E1CGA3): chymotryptic in-solution digestion.

V.5.4.3 Tryptic-chymotryptic digestion

In both TCG and TCS, the first and second highest scoring peptides were not unique to E1CGA3. In TCG, the highest scoring unique peptide was (Y)GSDSA₁₃GSGAGR(A) at the residues 486-508 with -10lgP score of 37.15; in TCS, the highest scoring unique peptide was (R)AAGSA₁₃SGAGGSGR(G) at the residues 908-932 with -10lgP of 55.23. In TCS, a few numbers of non-enzyme specific cleavages were observed. The majority was cleaved after the aspartic acid: for example, a unique peptide (D)GGYGSDSA₁₃SGAGGRGD(G) was cleaved after the aspartic acid in both ends; another unique peptide (A)GSA₁₀GAGASRPVGIY(G) was randomly cleaved after alanine on the C-terminus but shows a chymotryptic cleavage on the N-terminus (Figure 91). In regards of peptide distribution, TCS showed a TCS showed more numbers of missed cleavages from a same sequence between the residues 170-320 and yielded more peptides between the residues 700-1800 and 2300-2700 (Figure 92, Figure 93).

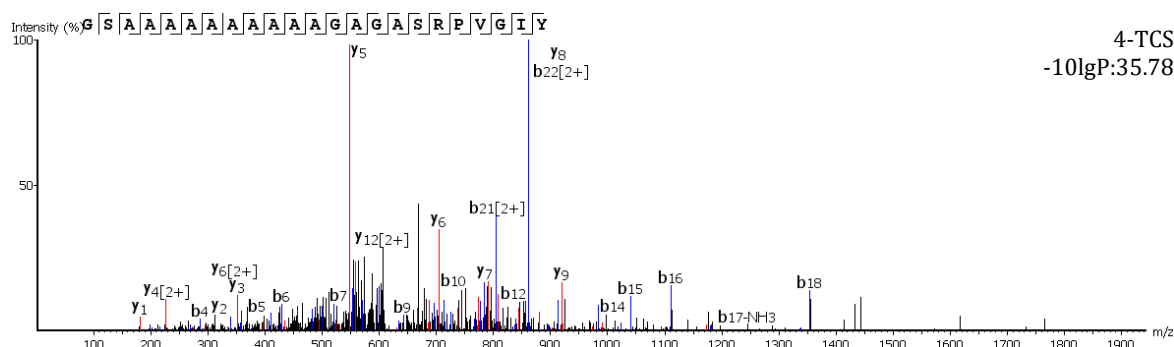
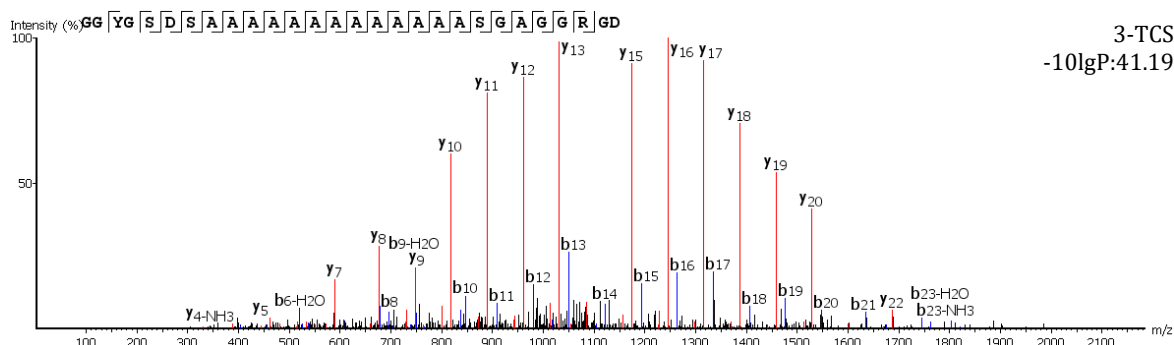
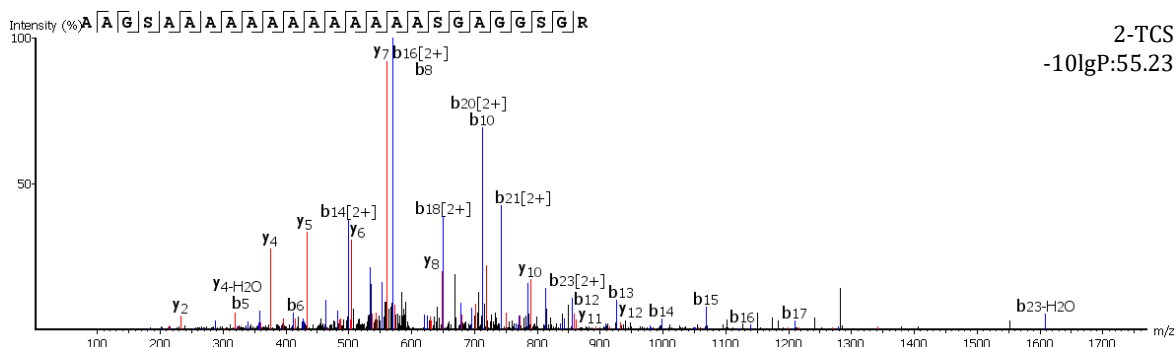
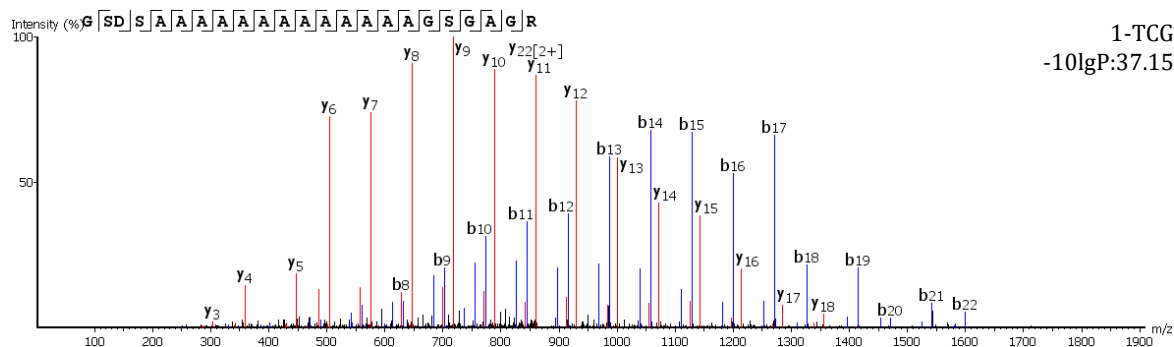


Figure 91. Spectra of unique tryptic-chymotryptic peptides of *A. yamamai* fibroin: 1) GSDSA₁₃GSGAGR ($m/z=887.43$, $z=2$), 2) AAGSA₁₃SGAGGSGR ($m/z=619.98$, $z=3$), 3) GGYGSDSA₁₃SGAGGRGD ($m/z=1112$, $z=2$), 4) GSA₁₀GAGASRPVGIIY ($m/z=634.67$, $z=3$)



Figure 92. Protein coverage of *A. yamamai* fibroin (E1CGA3): tryptic-chymotryptic in-gel digestion.

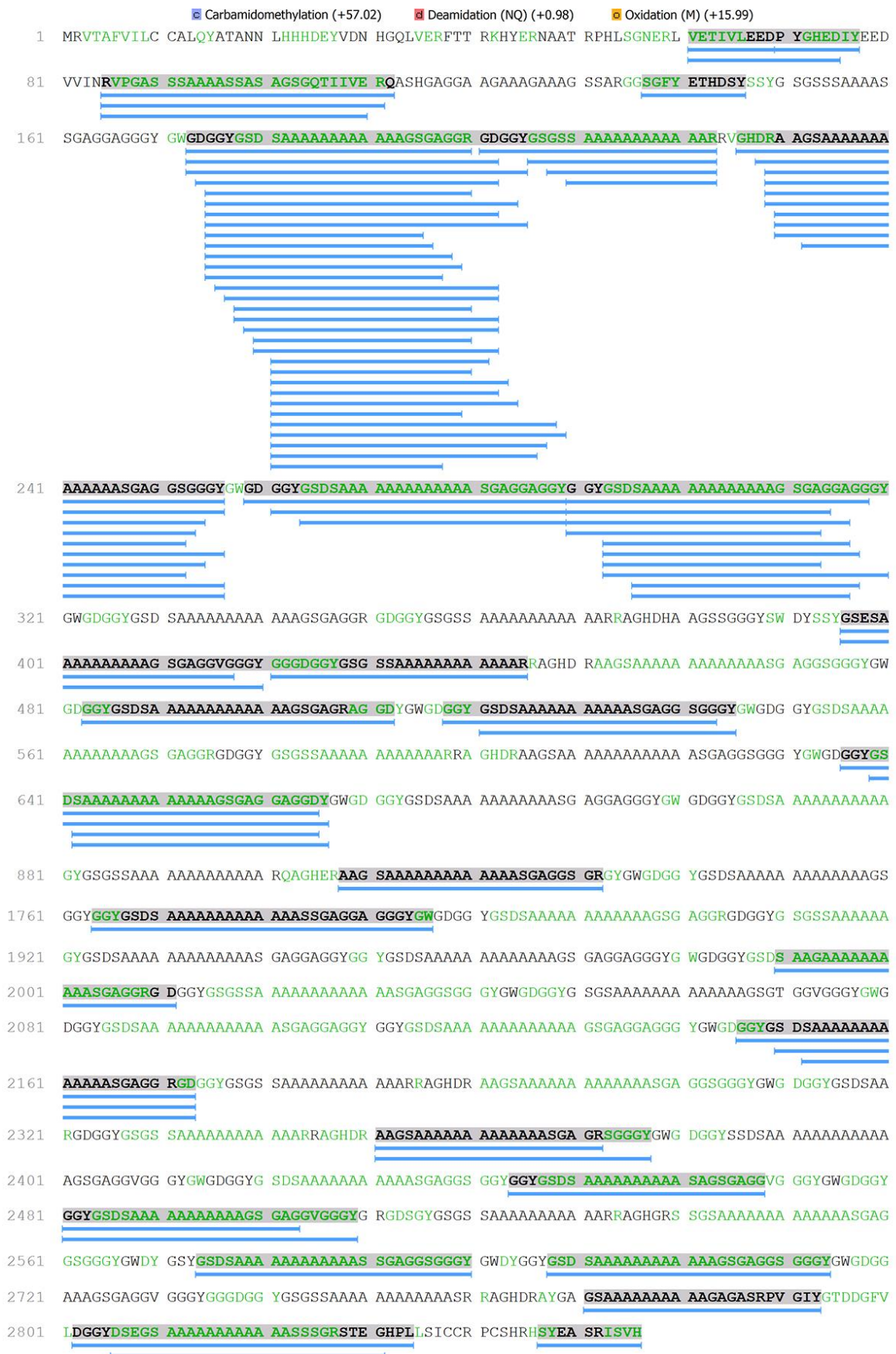


Figure 93. Protein coverage of *A. yamamai* fibroin (E1CGA3): tryptic-chymotryptic in-solution digestion.

V.5.5 *Antheraea assamensis*

In case of *A. assamensis* silk, chymotryptic peptides showed the highest protein coverage (16 and 21%), however the ratio of unique peptides was higher in tryptic digestion. The protein probability score was higher in tryptic digestion than in chymotryptic digestion. Moreover, CG showed much lower protein probability score than CS, thus the concentration of the sample was improved, it is expected to show a better protein coverage. However, the overall result was poor compare to the other silk the samples. Despite the 100% protein sequence coverage result of theoretical digestion (see Table 18), tryptic-chymotryptic digestion of *A. assamensis* did not show a good result. Unfortunately, time was not permitting to do more experiments to find out the reason, however the peptide distribution of the tryptic-chymotryptic digestion results suggest that the proteolysis was not optimised: TCG showed 10% protein coverage despite the relatively low protein probability score (161.71), whereas TCS showed good protein probability score (239.88) but low protein coverage (Table 29).

Table 29. PEAKS analysis results of *A. assamensis* fibroin treated with different enzymes

	Trypsin		Chymotrypsin		Try-Chymo	
	In-gel	In-sol	In-gel	In-sol	In-gel	In-sol
<i>Antheraea assamensis</i> (A0A0K0KR73 ANTASA)						
-10lgP	278.06	308.97	136.05	212.76	161.71	239.88
Coverage (%)	7	12	16	21	10	5
#Peptides	58	101	29	83	16	16
#Unique	41	73	21	69	15	10

V.5.5.1 Tryptic digestion

In TG, the highest scoring peptide was a unique peptide to A0A0K0KR73, (K)RVPGASSSA₄SSASAGGR(G) at the residues 85-104 that showed -10lgP score of 70.89. In TS, another unique peptide derived from the same sequence was the highest scoring peptide: (R)VPGASSSA₄SSASAGGR(G) at the residues 86-104 showed -10lgP score of 76.72. In both samples, some peptides have non-enzyme specific ends were found. In TG, a unique peptide (D)SFVVDSS(Y) (113-119) with -10lgP score of 23.55 was resulted by the cleavages after aspartic acid (D), serine (S) and glycine (G). In TS, (G)GSGGGYGWGD(E) was the only peptide with two non-enzyme specific ends. However, it was not unique to A0A0K0KR73. However, unique peptides with random C-terminus and tryptic N-terminus were found; for example, unique peptides (A)A₉GSGGR(G) and (A)AAGSGAGGAGGGYGR(G) were cleaved after alanine(A) on C-terminus (Figure 94). Between the two samples, TS showed a greater number of peptides between the residues 800-2550 (Figure 95, Figure 96).

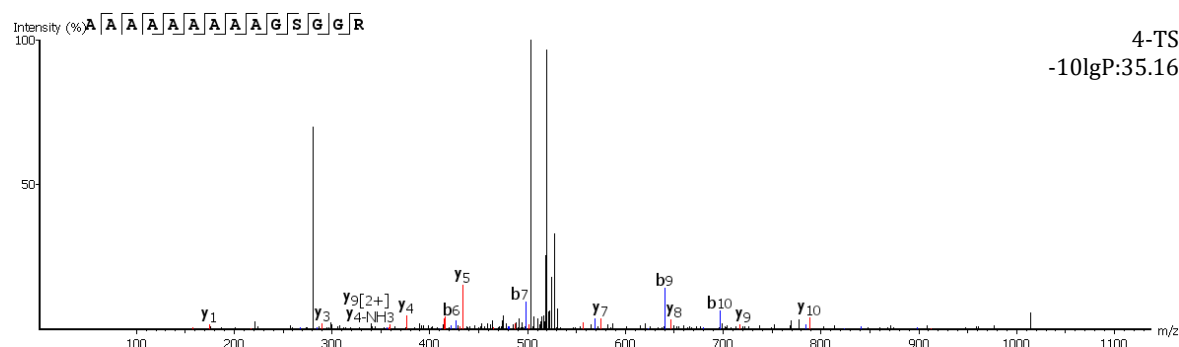
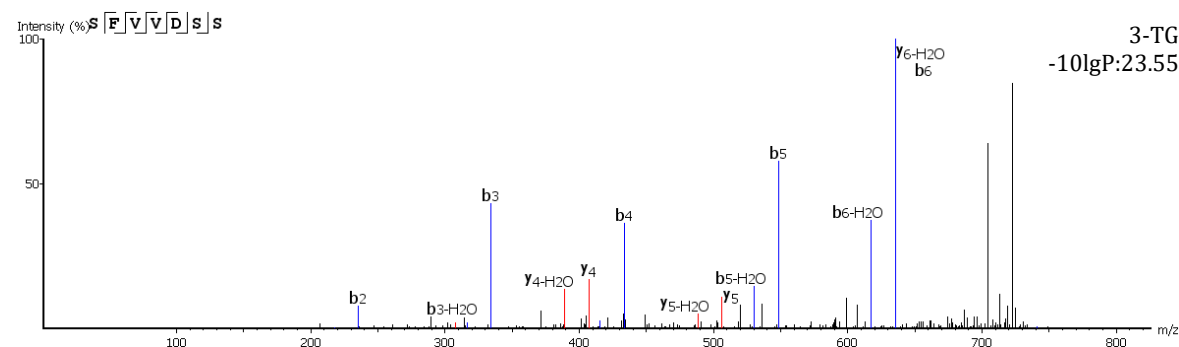
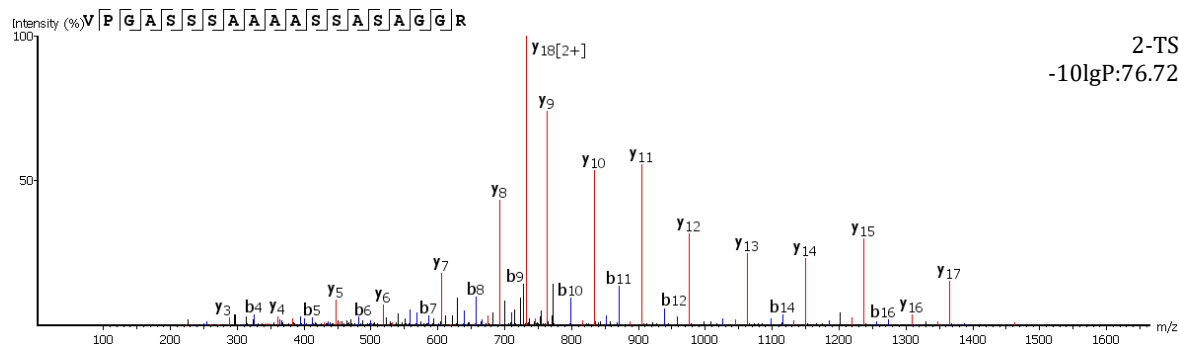
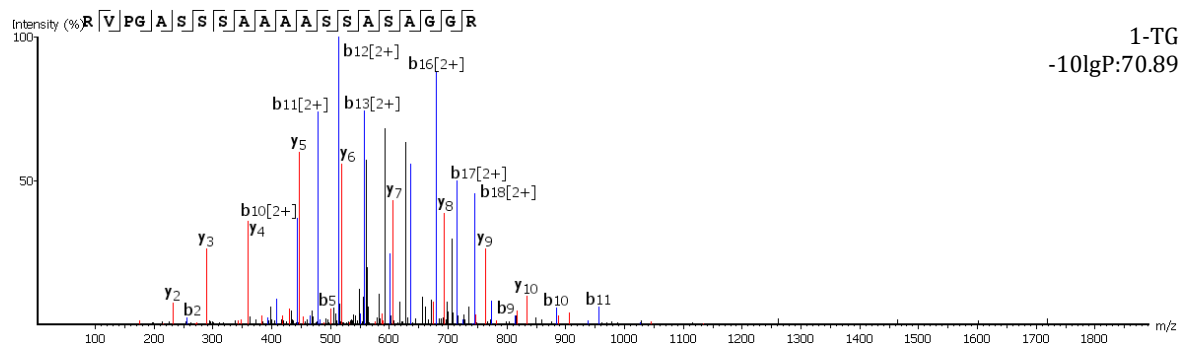


Figure 94. Spectra of tryptic unique peptides of *A. assamensis* fibroin: 1) RVPGASSA₄SSASAGGR ($m/z=573.29$, $z=3$), 2) VPGASSA₄SSASAGGR ($m/z=781.38$, $z=2$), 3) SFVVDSS (740.35 , $z=1$), 4) A₉GSGGR ($m/z=536.78$, $z=2$)



Figure 95. Protein coverage of *A. assamensis* fibroin (A0A0K0KR73): tryptic in-gel digestion.

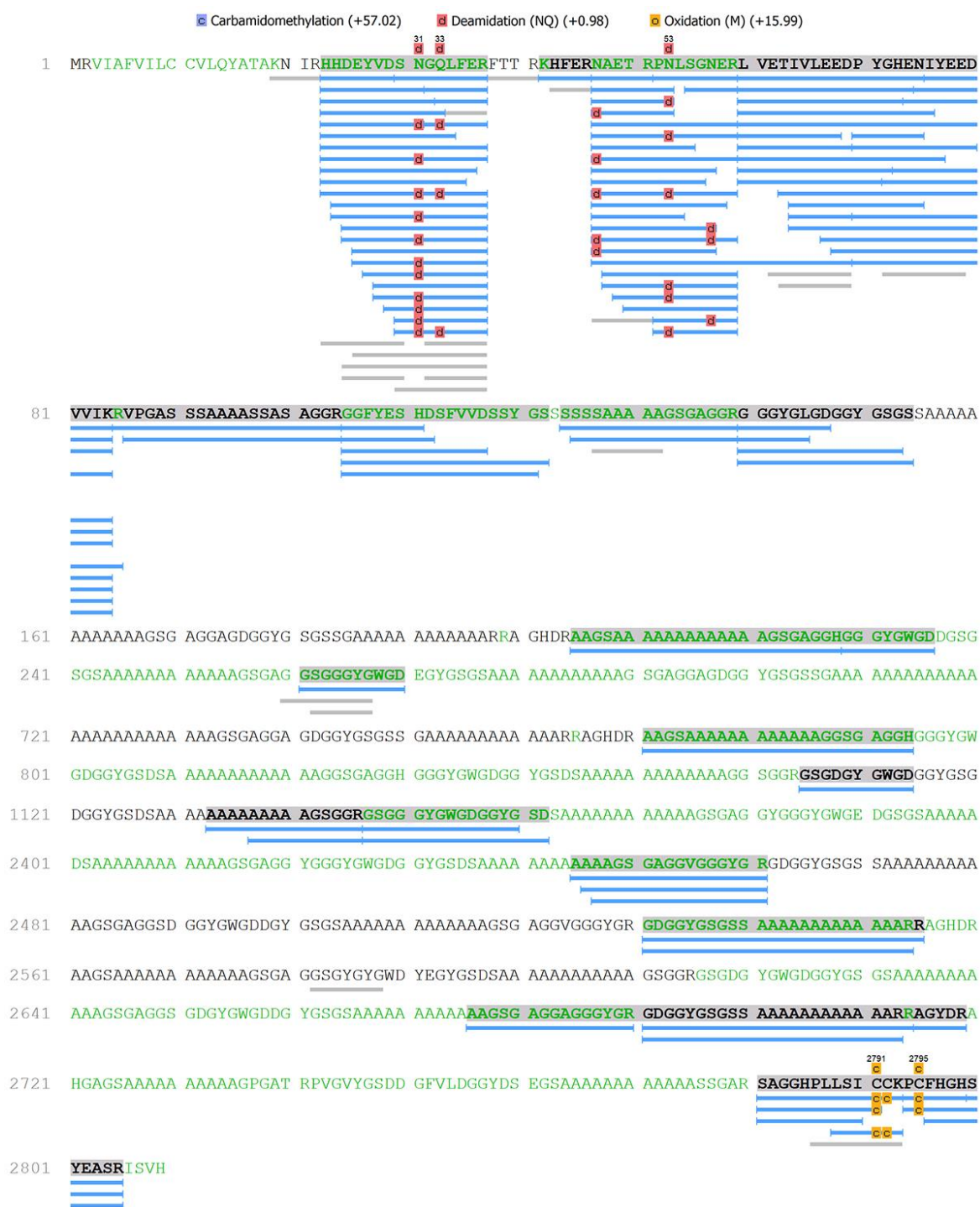


Figure 96. Protein coverage of *A. assamensis* fibroin (A0A0K0KR73): tryptic in-solution digestion.

V.5.5.2 Chymotryptic digestion

In CG, all peptides were chymotryptic and the highest scoring peptide was unique to A0A0K0KR73: (Y)GSDSA₁₂GSGGRGSGDGY(G) at the residues 2204-2230 with 10lgP score of 40.72. In CS, some unique peptides with non-enzyme specific ends were found: (R)(V) PGASSSA₄SSASAGGR(G) (86-104) were both unique to A0A0K0KR73, and showed -10lgP score of 48.44 and 41.32, respectively; (D)RAAGSA₁₂GGSGAGG(H), and (R)AAGSA₁₃GSGAGG(H), (G)ASSSA₄SSASAGGR(G) are unique peptides also resulted from non-enzyme specific cleavages after aspartic acid (D), arginine (R), and glycine (G). The highest scoring peptide in CS was unique to A0A0K0KR73: (Y)GSGSA₁₃GSGAGGAGD(G) at the residues 717-742 with 10lgP score of 52.10 (Figure 97). CS showed relatively even distribution of peptides throughout the protein sequence than CG, however, this seems to be affected by the low protein score and the coverage due to the poor gel resolution (Figure 98, Figure 99).

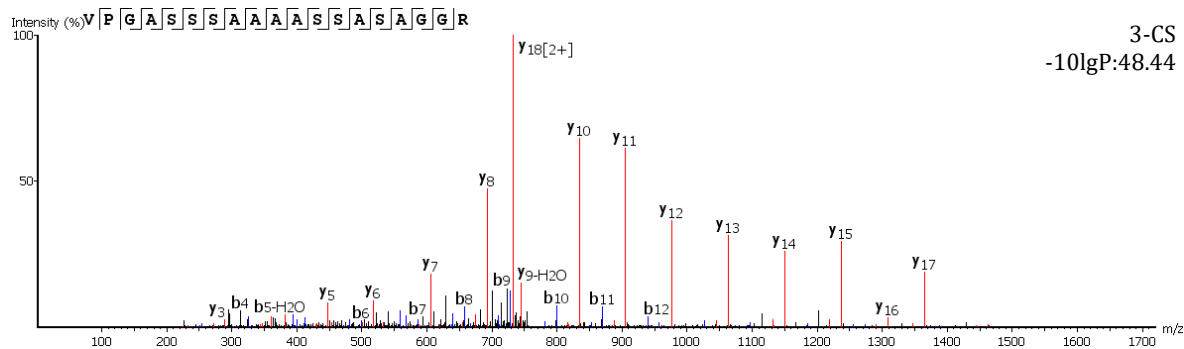
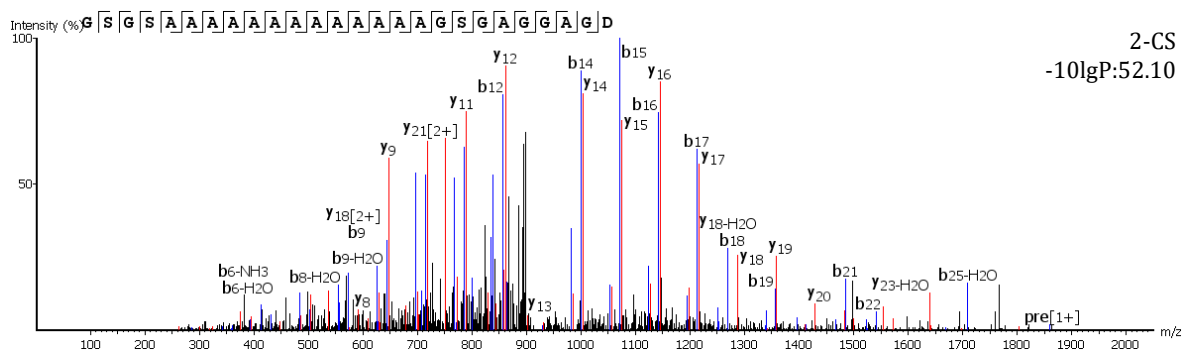
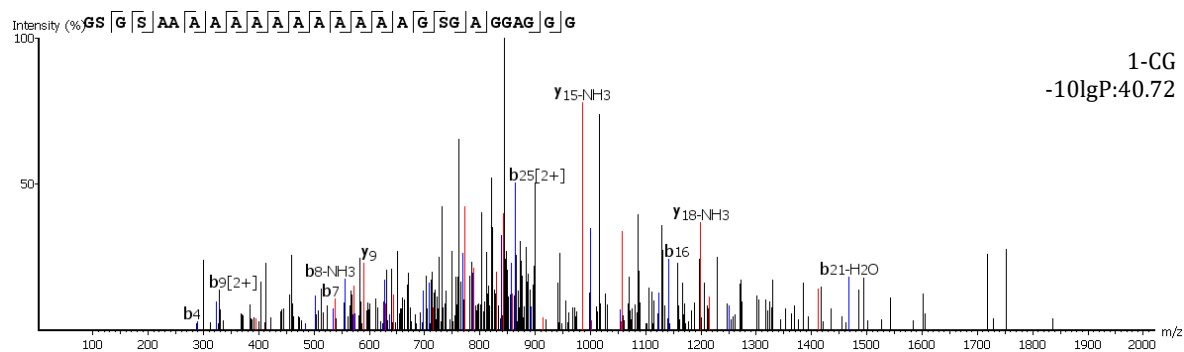


Figure 97. Spectra of unique peptides for *A. assamensis* fibroin: 1) GSDSA₁₂GSGGRGSGDGY ($m/z=1084.5$), 2) GSGSA₁₃GSGAGGAGD ($m/z=930.43$), 3) VPGASSA₄SSASAGGR ($m/z=781.38$), $z=2$

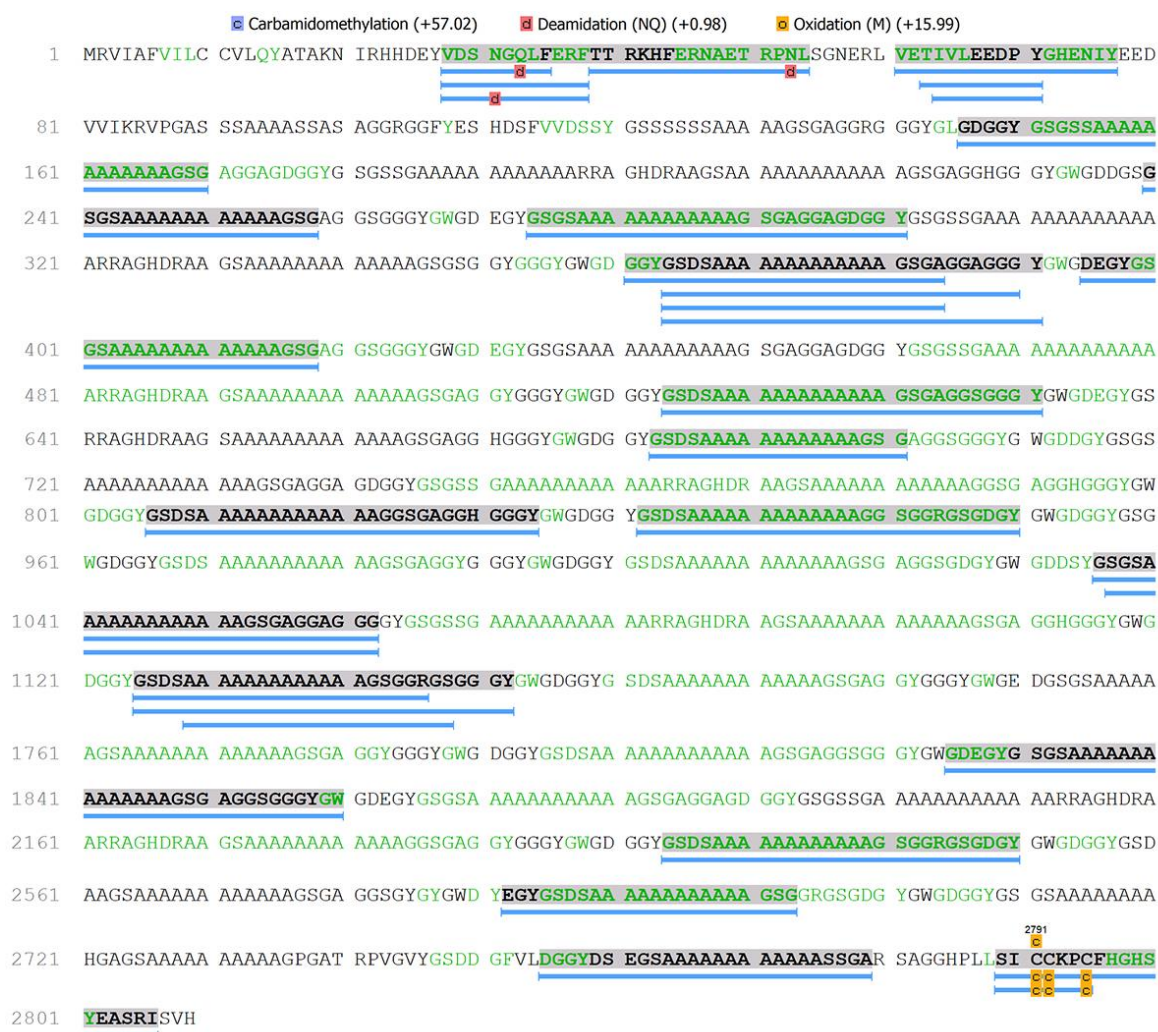


Figure 98. Protein coverage of *A. assamensis* fibroin (A0A0K0KR73): chymotryptic in-gel digestion.



V.5.5.3 Tryptic-chymotryptic digestion

In TCG, the highest scoring peptide was a chymotryptic peptide (W)GDGGYGSDSA₁₂GSGAGGS(G) at the residues 678-705 with -10lgP score of 63.03, which is unique to A0A0K0KR73. In TCS, a tryptic unique peptide (R)VPGASSSA₄SSASAGGR(G) at the residues 86-104 showed -10lgP score of 90.23. In TCG, several unique peptides were yielded by non-enzyme specific cleavages after serine(S), glycine(G), and aspartic acid (D): for example, (S)DSA₁₃GSGGRGSG(G) (1127-1149) showed -10lgP score of 43.86, and (S)DSA₁₂GSGGRGSGD(G)(2206-2228) showed -10lgP score of 26.18. On the other hand, in TCS, (E)TIVLEEDPYGHE(N) was the only peptide with non-enzyme specific ends, but this was not unique to A0A0K0KR73(Figure 100). In addition, TCS showed low protein sequence coverage and only yielded peptides near C- and N-terminus; thus, the poor result is more likely due to the poor proteolysis (Figure 101, Figure 102).

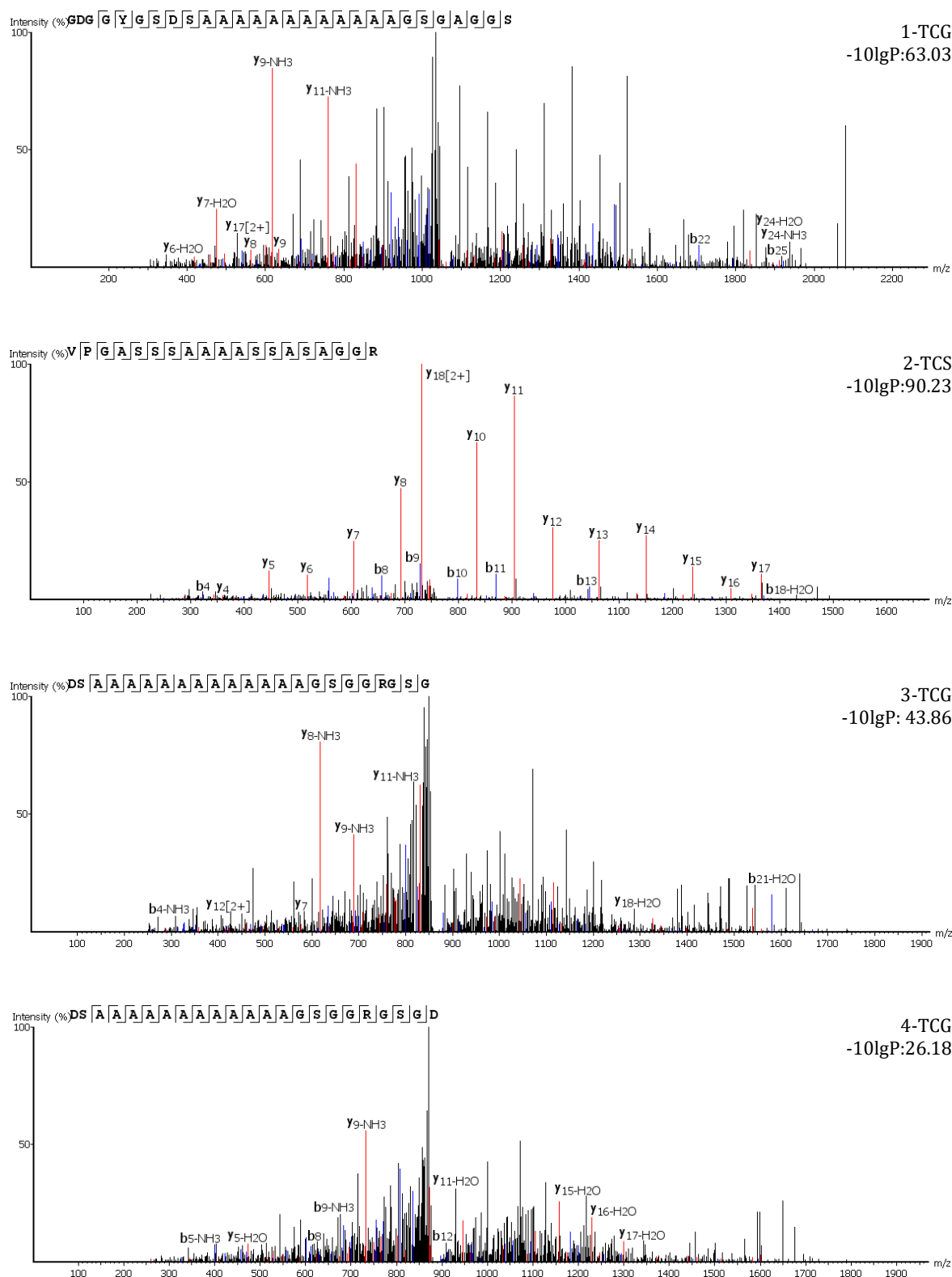


Figure 100. Spectra of unique tryptic-chymotryptic peptides of *A. assamensis* fibroin: 1) GDGGYGSDSA₁₂GSGAGGS (m/z =1070.47), 2) VPGASSSA₄SSASAGGR (m/z =781.38), 3) DSA₁₃GSGGRGSG(m/z =880.42), 4) DSA₁₂GSGGRGSGD (m/z =902.41), z =2.



Figure 101. Protein coverage of *A. assamensis* fibroin (A0A0K0KR73): tryptic-chymotryptic in-gel digestion.

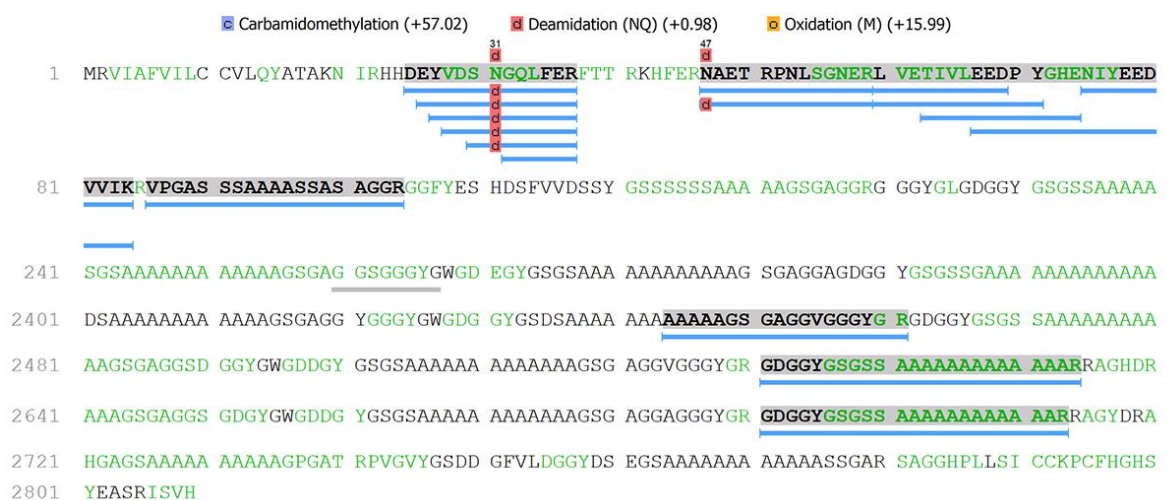


Figure 102. Protein coverage of *A. assamensis* fibroin (A0A0K0KR73): tryptic-chymotryptic in-solution digestion.

V.5.6 *Samia ricini*

In *S. ricini* silk, tryptic-chymotryptic digestion showed relatively better identification result than tryptic or chymotryptic digestion. TCG showed 12% protein coverage with probability score of 259.01, and TCS showed 17% protein coverage with -10lgP score of 175.89. However, the greatest number of peptides and the highest probability score was yielded by TS, with 12% protein coverage and -10lgP score of 327.23. Considering the lower probability score but the higher protein coverages of tryptic-chymotryptic digestion samples compare to TS, results can be improved if the protein concentration and digestion condition are optimised (Table 30). However, the overall results were the poorest among the 7 different types of silks, suggesting that the physical characteristics of the fibroin are the cause. *S. ricini* silk showed the least solubility in both calcium nitrate and lithium thiocyanate solvents; which thought to be affected by the rigid polyalanine structure of the protein that resembles the spider silk.⁴⁷²

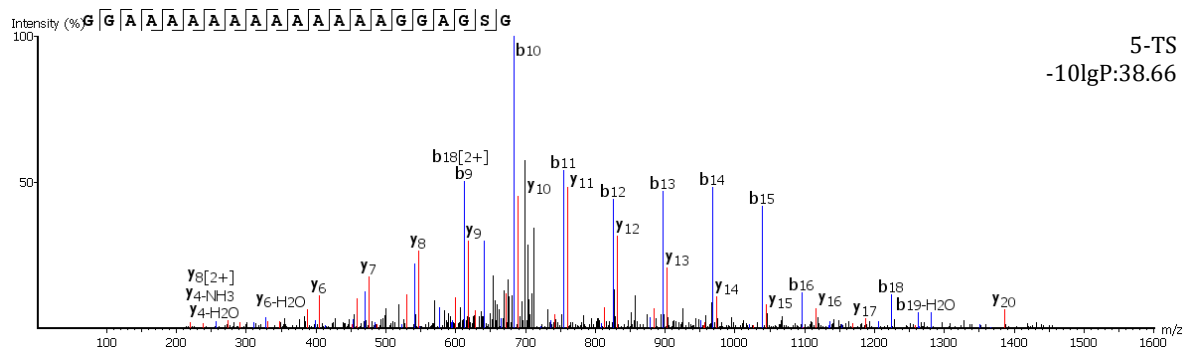
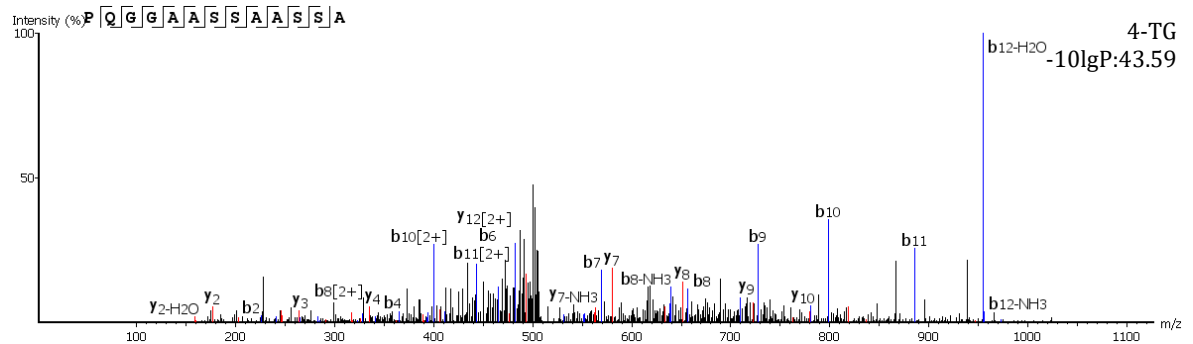
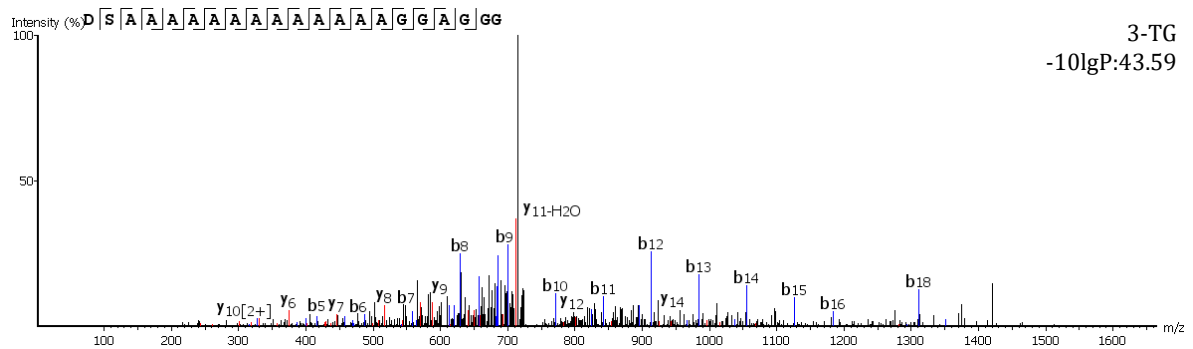
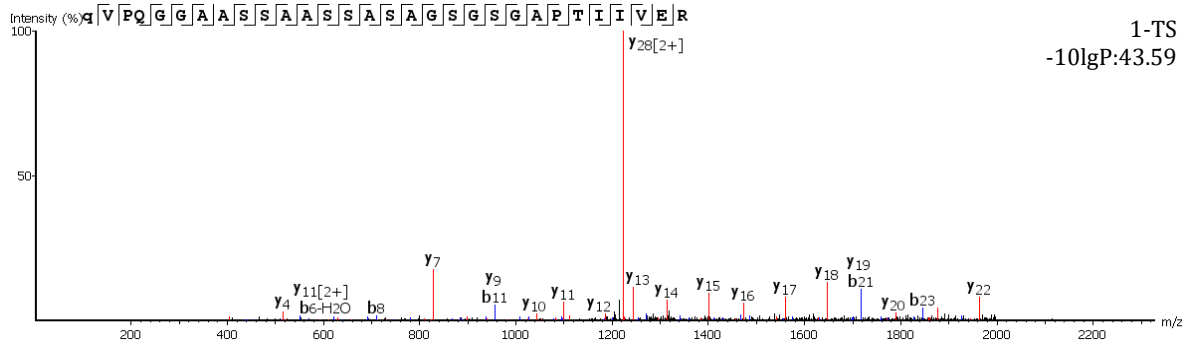
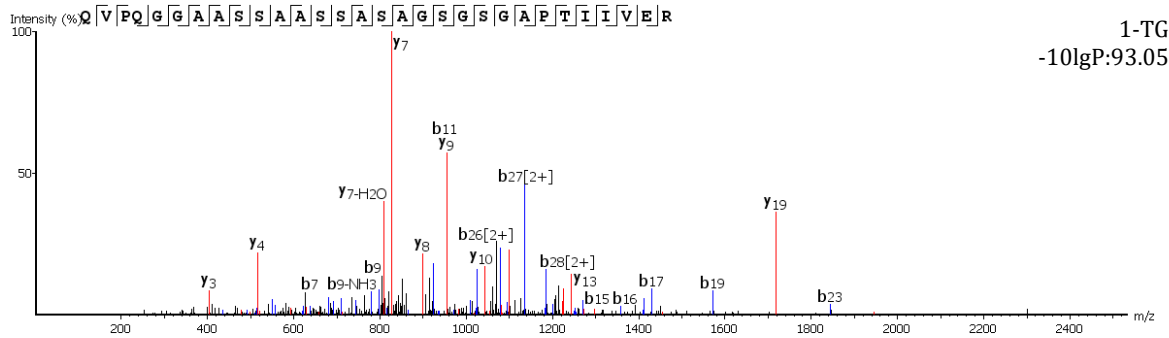
Table 30. PEAKS analysis results of *S. ricini* fibroin treated with different enzymes

	Trypsin		Chymotrypsin		Try-Chymo	
	In-gel	In-sol	In-gel	In-sol	In-gel	In-sol
<i>Samia ricini</i> (A0A0D5ZYI3 SAMRI)						
-10lgP	353.64	327.23	281.59	192.36	259.01	175.89
Coverage (%)	5	12	14	7	12	17
#Peptides	46	70	23	22	49	34
#Unique	45	67	20	22	49	33

⁴⁷² Nakazawa and Asakura, 2003

V.5.6.1 Tryptic digestion

The highest scoring peptide in both samples was a unique peptide to A0A0D5ZYI3, (K)QVPQGGAASSAASSASAGSGSGAPTIIVER(G) at the residues 101-130: it showed -10lgP score of 93.05 in TG and 89.89 with deamidation on Q (+.98) or 89.31 without deamidation in TS. In TG, most of the peptides were tryptic peptides and were unique to A0A0D5ZYI3; however, there were a few peptides cleaved by non-enzyme specific rules, including (V)PQGGAASSAASSA(S) (103-115) and DSA₁₃GGAGGG (2738-2758). In TS, the majority of the peptides were tryptic or tryptic with one non-enzyme specific end after aspartic acid (D), alanine (A), serine (S) and glycine (G). Some of them had two non-enzyme specific ends: for example, (D)GGA₁₃GGAGSG(Y) (413-433) with -10lgP score of 38.66, (D)SGA₁₃GGAGD(G) (1046-1065) with -10lgP score of 36.13 are cleaved after aspartic acid and glycine (Figure 103). While TG only yielded peptides near the C-and N- terminus, TS showed a more diverse distribution of peptide throughout the sequence (Figure 104, Figure 105).



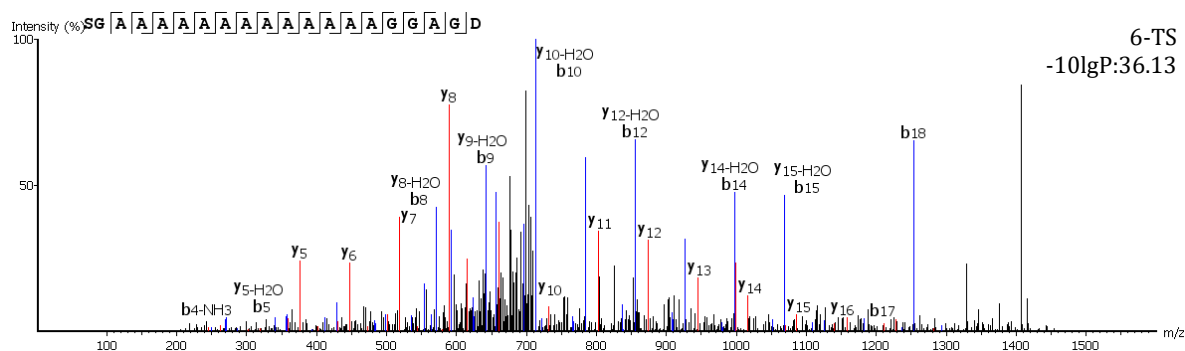


Figure 103. Spectra of tryptic unique peptides of *S. ricini* fibroin: 1) QVPQGGAASSAASSASAGSGSGAPTIIVER ($m/z=891.11$, $z=3$), 2) QVPQGGAASSAASSASAGSGSGAPTIIVER ($m/z=1336.66$), 3) PQGGAASSAASSA ($m/z=750.86$), 4) DSA₁₃GGAGGG ($m/z=531.25$), 5) GGA₁₃GGAGSG 6) SGA₁₃GGAGD ($m/z=722.34$), $z=2$



Figure 104. Protein coverage of *S. ricini* fibroin (A0A0D5ZYI3): tryptic in-gel digestion.



Figure 105. Protein coverage of *S. ricini* fibroin (A0A0D5ZYI3): tryptic in-solution digestion.

V.5.6.2 Chymotryptic digestion

In both CG and CS, all detected peptides were unique to A0A0D5ZYI3. The highest scoring peptide in CG was (Y)GGDGGA₁₃GGSGSGYG(G) at the residues 2635-2660 with -10lgP score of 53.26; the highest scoring peptide in CS was (Y)GAGSA₁₂GGAGGG(Y) at the residues 519-540 with -10lgP score of 66.66. In CG, many of the resulted peptides had one non-chymotryptic end: for example, (G)SDGGA₁₃GGAGSGYG(G) (1807-1831) and (G)SDGGA₁₃GGSEGGYG(G) (2210-2234) are both cleaved after glycine on both ends, and (D)GGA₁₃GSGGGYGGD(G) (2151-2174) was cleaved after aspartic acid and glycine. A similar result was observed in CS: there were peptides with non-chymotryptic ends such as (D)GGA₁₃GGSGGG(Y) (472-492) and (A)ASSASAGSGSGAPTIIVER(G) (112-130), which are cleaved after aspartic acid, glycine and arginine (Figure 106.). Although CS showed almost the same number of unique peptides to CG, there were no peptides detected between the residues 860-2710 while CG showed relatively even distribution of peptides throughout the protein sequence (Figure 107, Figure 108).

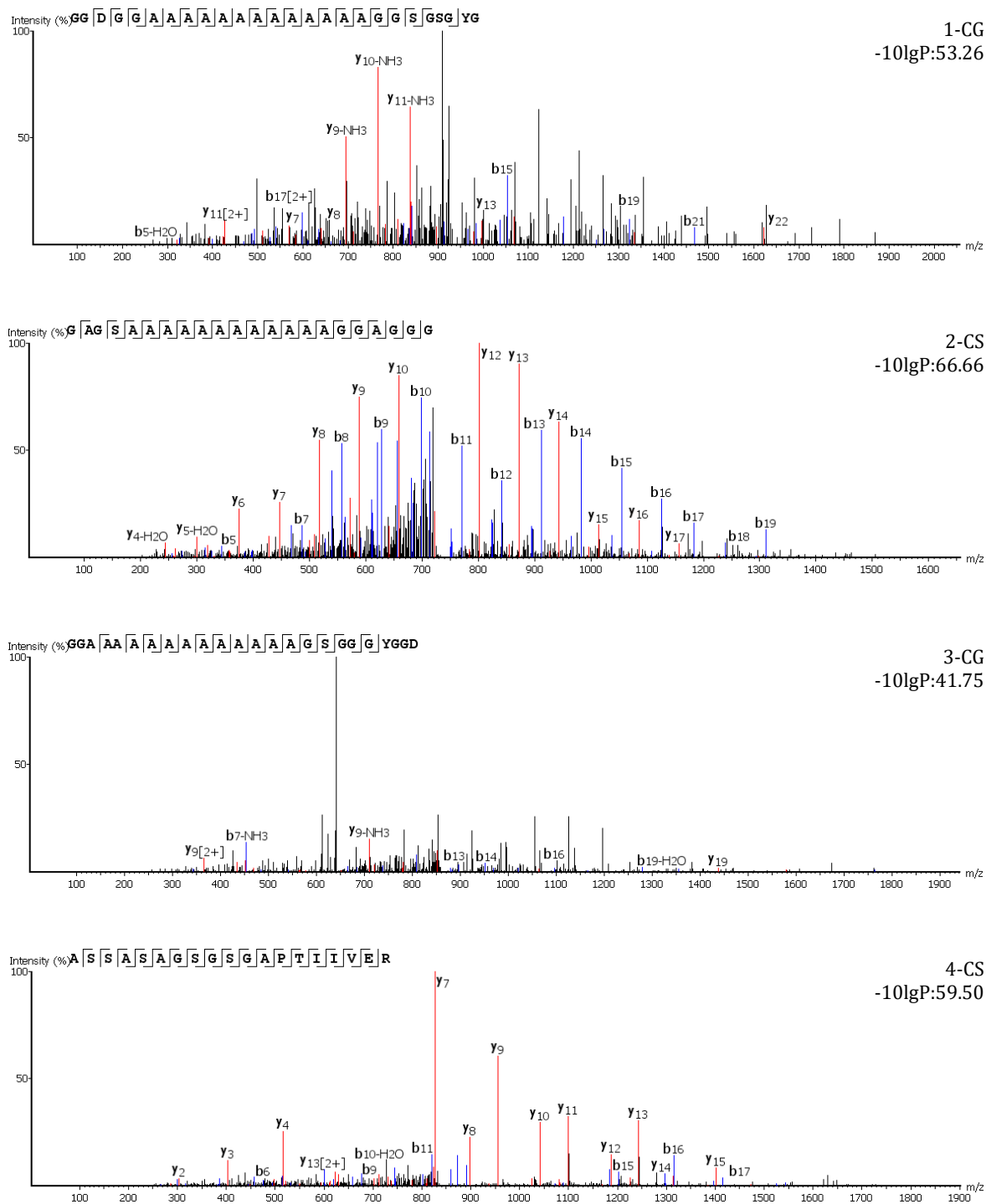


Figure 106. Spectra of unique chymotryptic peptides of *S. ricini* fibroin: 1) $\text{GGDGGA}_{13}\text{GGSGSGYG}$ ($m/z=954.43$), 2) $\text{GAGSA}_{12}\text{GGAGGG}$ ($m/z=750.36$), 3) $\text{GGA}_{13}\text{GSGGGYG}$ ($m/z=882.4$), 4) $\text{ASSASAGSGSGAPTIIVER}$ ($m/z=859.44$)

1 MRVTAFVILC CVLQYVTARS LDDDIHSLER GY**RETSKTGY DEYDVKSGR LY**ERLTTRKK FERDSAPSRV PGGGTILVEKI
 161 GAGRGGGGY GRHGHTAGSA AAAAAAAAAA AASSEGGGAG GYWQGY**GSDS SAAAAAAAAA AAGSAGGSYG** SDSAEAAAAA
 241 AAAAAAGTA AGGSGGGY**GG DGGAAAAAAAA AAAAAASGAG SGYGGYGHG YGSDGGA AAAAAGAG GAGRGYGAGS**
 321 AAAAAAAAAA AAGSGGGYG GDGGA AAAAASGA GSGYGGGARG GYGHGY**GSDG GAAAAAAAAA AAAAGGAGGG**
 401 YGGGYGGYG GDGGA AAAAAGGA GSGYGGGWH SY**GSDSGAAA AAAAAAAAAA AGGSGGGYGG DGGAAAAA**
 481 AAAAAAGGSG GYGDDGGA AAAAAAAAAA AGGAGDGYGA GSAAAAAA AAAAGGAGG YGGDGA AAAAAGG
 721 GSGGGY**GGD GAAAAAAAAA AAAAGGAGG GY**GDDGGA AAAAAGGAGG YGGDGA AAAAAGGAGG
 801 GGDGGA AAAAAGGAG GAGSGYGGGA RRGYGHGYGS DGGAAAAA AAAAAGGSG GYGDDGGA AAAAAAAAAA
 881 AGGAGSGYG GARRGYGHGY GSDGGA AAAAAGG AESSYGGGSW YGY**GSDSSAA AAAAAAGG AGGGY**GDDG
 961 SAAAAAAAAA AAGSGGGY GDDGGA AAAAAGG AGGGYGGYG **GDGGA AAAAAGG RSGY**GGGWH
 1121 RGYGHGYSD GAAAAAAAAA AAAAGGSGG Y**GDDGGA AAAAAGG GAGSGY**GGGS WHGYSDSA AAAAAAAAAA
 1761 AGGAGGGYG DGGAAAAA AAAAGGAG SGYGGGARG YGHGYSDG **AAAAAGGAGG GGGWNSYGS**
 2081 AAAAGGAGSG YGGGWSHYG SDGAAAAA AAAAAGGA EGGYAGSAA AAAAAAAAAA GGAGGGYGD **GGAAAAAA**
 2161 AAAAGSGGG **YGGD**GGA AAAAAGGAG GAGSGYGGGA RRGYGHGYGS **DGGAAAAA AAAAAGGSE GYGDDGGA**
 2321 AAAAAAAAAA AAGGAGSGY GGARGYGHG YGSDGGA AAAAAGG GAGSGYGGGS WHSYSDSA **AAAAAGGAGG**
 2401 **ASGAGSGY GGS**WSHYSD SGAAAAAA AAGAGGAG GYGAGSAAA AAAAAAGG AGGGYGRDG **AAAAAGGAGG**
 2561 GGSWSHYGS **DSGAAAAA AAAAAGGAG SGYGGGSRGG** YGHGYSDG **AAAAAGGAGG GGGYGGDGA**
 2641 AAAAAAAAAA AAGSGSGYG GGSWSHYSD SDAAAAAA ASAAAGGAGS YGGGYWQGY **GSNSGAAAAA AAAAAGGAGG**
 2721 **SGYGGGARGG YGHGYSDSA AAAAAAGGAG RGAIGGYGG YGSDNAAAAA AAAAAGGAG GGDYGRGYGA**
 2801 **RSAAAAAA AASSGARGAV RVHETGDGFL** LRGDYSDSS AAAAAAGG AASSASSAN GYVSICCKPC RLGTSKIIVH

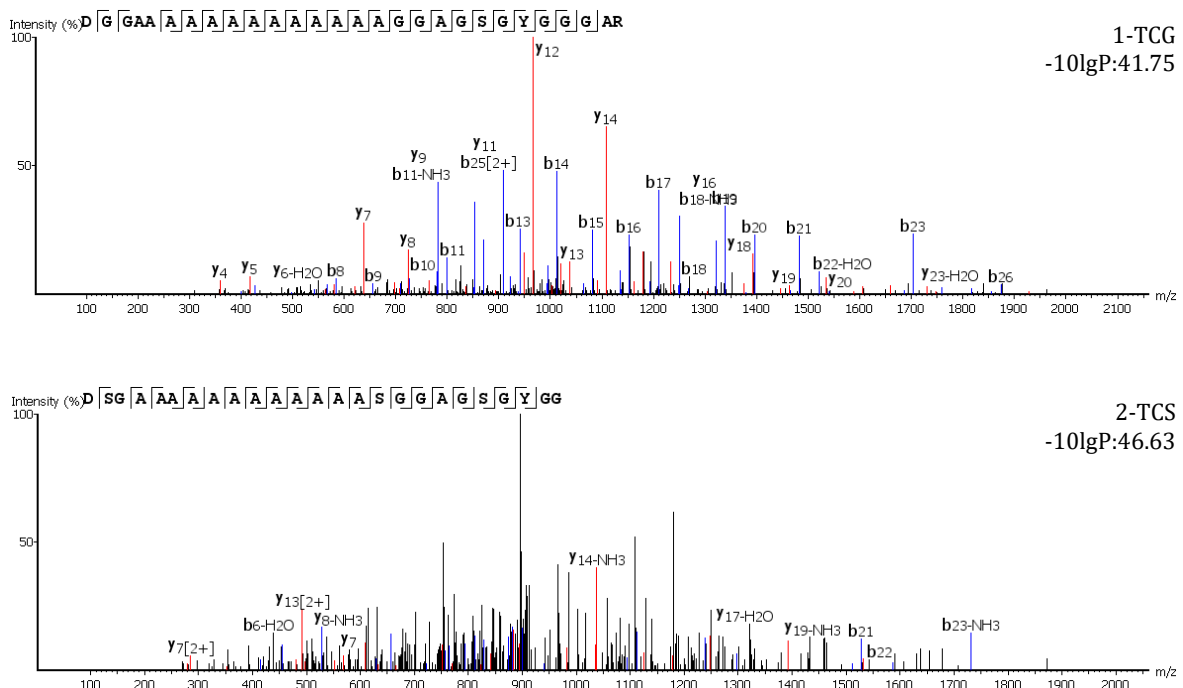
Figure 107. Protein coverage of *S. ricini* fibroin (A0A0D5ZYI3): chymotryptic in-gel digestion.

1 MRVTAFVILC CVLQYVTARS **LDDDIHSL**ER GYRETSKTGY DEYDVKSGR LYERLTTRKK FERDSAPS**RV PGGGTIVEKI**
 81 VIER**APTGHE TTYEEDVVIK QVPQGAASS AASSASAGSG SGAPTIIVER** GSGAGSGR**SH GAGSAAGAAA AAAAAAAAAAG**
 161 **GAGR**GGGGGY GRGHGTAGSA AAAAAAAAAA AASSEGGSSAG GYWQYGS**DS SAAAAAAAAA AAGSAGGSYG** SDSAEAAAAA
 241 AAAAAAAGTA AGGSGGGYGG **DGGAAAAAAAA AAAAAASGAG** SGYGGGYGHG YGSDGAAAA AAAAAAAG GAGRGY**GAGS**
 321 AAAAAAAAAA AAGSGGGYG GDGGAATAA AAAAAASGA GSGYGGGARG GYGHGYGS**DG GAAAAAAAAA AAAAGGAGGG**
 401 YGGGYGGGYG GD**GGAAAAA AAAAAAGGA GSGYGGGSWH** SYGSDSGAAA AAAAAAAAAA AGGSGGGYGG **DGGAATAA**
 481 **AAAAAAGGSG GGYGGDGGAA AAAAAAAAAA AGGAGDGYGA GSAAAAAAAA AAAAGGAGGG** YGGDGGAAAA AAAAAAAG
 881 AGGAGSGYGG **GARRGYGHGY GSDGGAAAA AAAAAAAGS AESSYGGGSW** YGYGSDSSAA AAAAAAAGG AGGGYGGDGG
 2721 SGYGGGARGG YGHGYGS**DSA AAAAAAAAAA AAGGAGGGAG RGAIGGYGGG** YGSDNAAAAA AAAAAAGGA GGDYGRGYGA
 2801 **RSAAAAAAAA AASSGARGAV RVHETGDGFL** LRGDYGS**DSS AAAAAAAAAA AAASSASSAN** GYVSICCKPC RLGTSKIAPH

Figure 108. Protein coverage of *S. ricini* fibroin (A0A0D5ZYI3): chymotryptic in-solution digestion.

V.5.6.3 Tryptic-chymotryptic digestion

In TCG, the highest scoring peptide was (G)DGGA₁₃GGAGSGYGGGAR(R) (544-571) with -10lgP score of 85.58; in TCS, (S)DSGA₁₂SGGAGSGYGG(G) (2387-2411) with -10lgP score of 46.63 was the highest scoring peptide; both peptides are unique to A0A0D5ZYI3. All detected peptides were unique to A0A0D5ZYI3 in TCG and TCS except (Y)EEDVVIK(Q). Both samples showed numbers of peptides with non-enzyme specific cleavages, mostly after aspartic acid(D) and alanine (A): in TCG, (G)GGYGHGYGSDGGA₃(A) at the residues 285-299 and (G)GGYGHGYG(S)DGG(A) at the residues 285-292/296 were cleaved after glycine, alanine and serine on both ends; in TCS, the highest scoring peptide DSGA₁₂SGGAGSGYGG was cleaved after serine and glycine, and (D)GGA₁₃GSGGGYGGD(G) at the residues 2151-2174 was cleaved after aspartic acid on both ends (Figure 109). TCG showed more peptides occurring between the residues 1-700 (Figure 110, Figure 111).



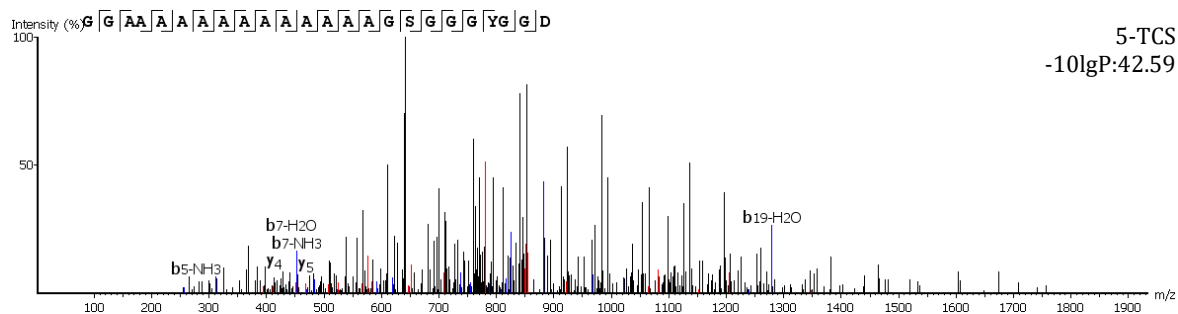
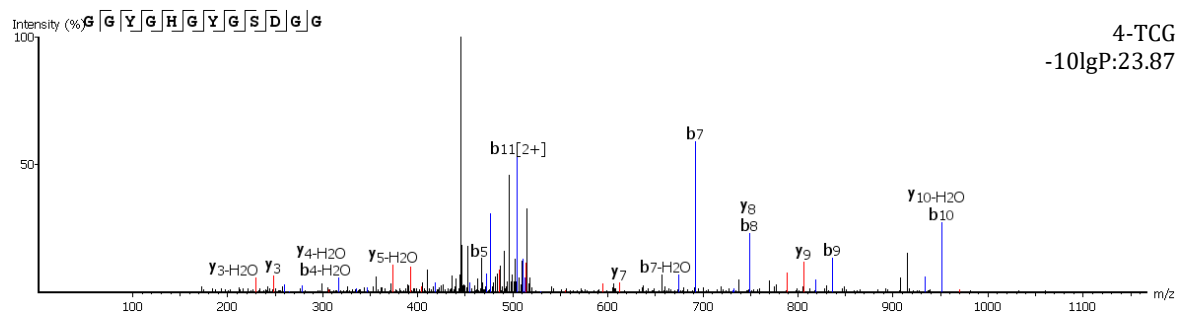
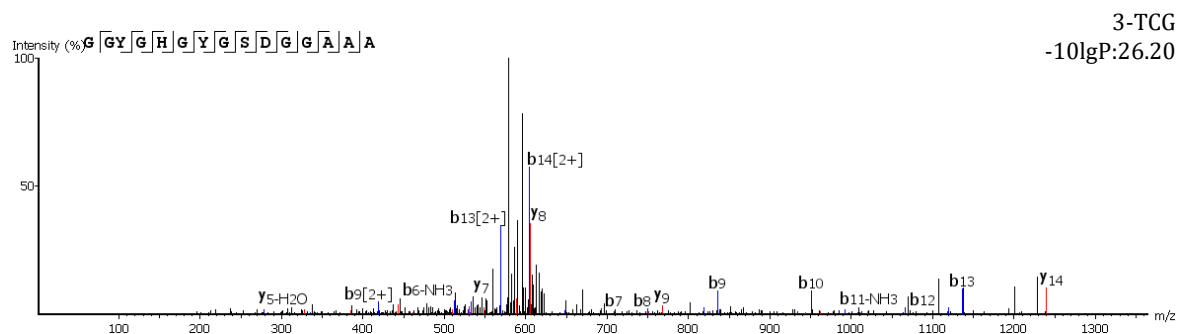


Figure 109. Spectra of tryptic-chymotryptic unique peptides of *S. ricini* fibroin: 1) DGA₁₃GGAGSGYGGAR (m/z=1056), 2) DSGA₁₂SGGAGSGYGG(m/z=940.92), 3) GGYGHGYGSDGGAAA (m/z=648.77), 4) GGYGHGYGSDGG(m/z=542.21), 5) GGA₁₃GSGGGYGGD (m/z=882.40)

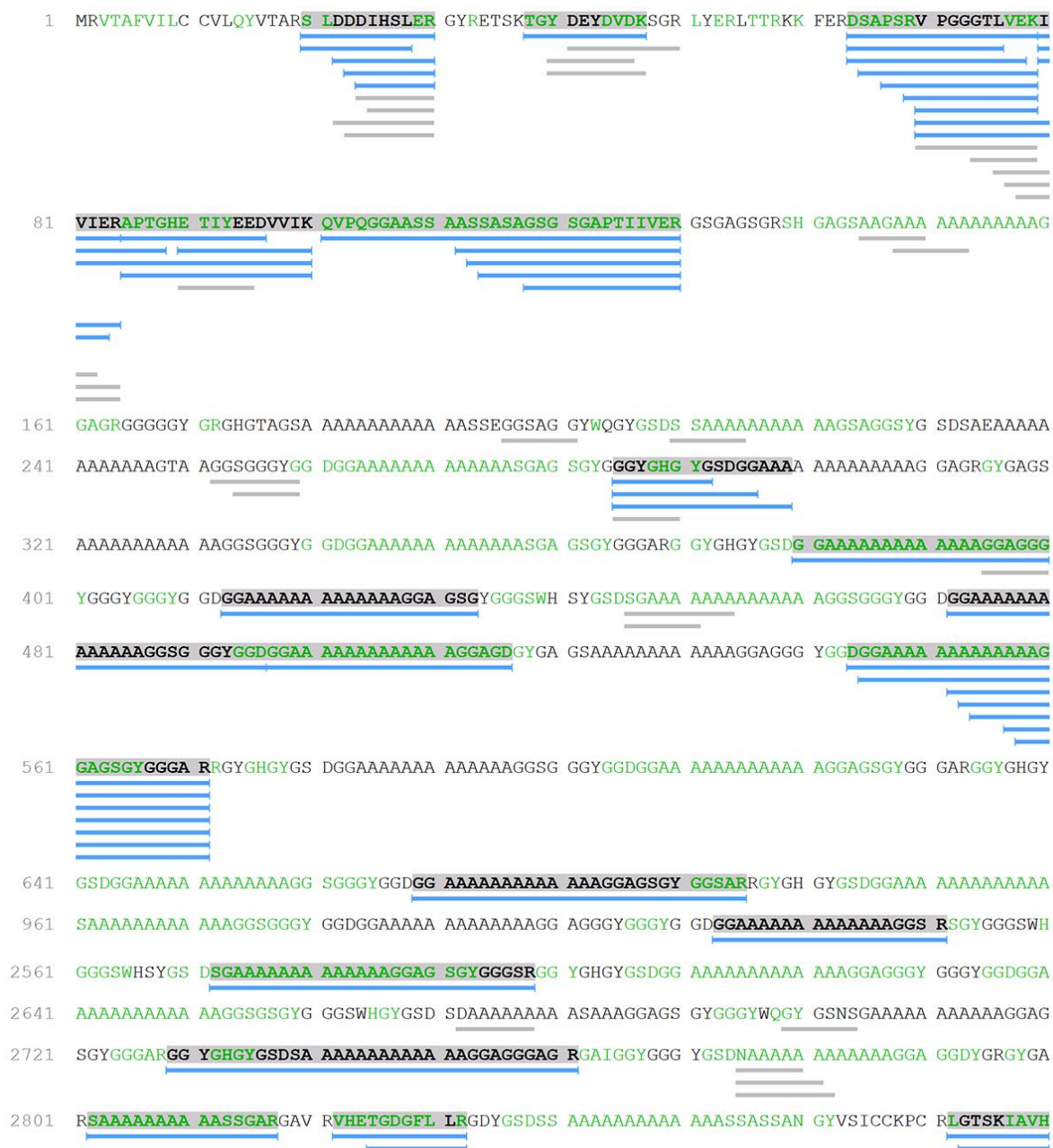


Figure 110. Protein coverage of *S. ricini* fibroin (A0A0D5ZYI3): tryptic-chymotryptic in-gel digestion.

VI. Case studies: implication to archaeological questions

VI.1 The textiles from Naintré

VI.1.1 Introduction

A series of textiles were found in two Gallo-Roman tombs excavated in Naintré, France, thought to be from 2-4th CE. There were layers of burial textiles found with a child's body. The initial status of preservation was noted as exceptionally good; however, since then most of the textiles have become unstable given the lack of appropriate environmental storage conditions. An initial study of the textile materials recorded the pattern and structure of the decorative fabric, which was woven with gold filé.⁴⁷³ The study of the textile was not expanded further, and the fibres were too brittle for SEM examination. Considering its poor condition, the textile fragments now piles of debris were thought to be a good candidate for protein mass spectrometry analysis.

Five samples were selected from a range of textile debris collected from the various locations within the child's coffin (Figure 112). They comprise at least four different types of textiles. In sample N2702, there are four different types of fabrics. There is a dark and thin gauze weave with a different type of thread used as weft. Layer 1 seems to be a part of the damask (D) as there are both warp-faced twill and weft-faced twill structures are seen. Layer 2 is a thick plain (tabby) fabric woven from coarser spun yarn, possibly made of wool. Then there is a layer made of dark resin with some imprints of the textiles on the surface indicates that it was placed between the dark gauze and the Layer 1 (Figure 113). In sample N2747, two more types of textiles are found in addition to those found in N2702. One is a fine gold-ornamented fabric, and the other is a thin, plain-woven fabric that is possibly made of silk dyed in a dark shade (purple). The gold threads are

⁴⁷³ Bedat, *et al.*, 1999

thought to belong to ornamental bands (E, F). Sample TD 152b is a layer of textile that only partially preserves its textile format. One side of it is completely covered with resin, and the other surviving surface is covered with calcified residue. It may be that this side was in contact with the body and the resinous material was applied upon it. Sample TD168 must have been a part of E or F: the gold threads on the surface attached to the resin might have been from other tapestry or the folded side of the same textile; on the side of the chunk, some of the textiles are exposed between the cracks in the resin-covered surface.

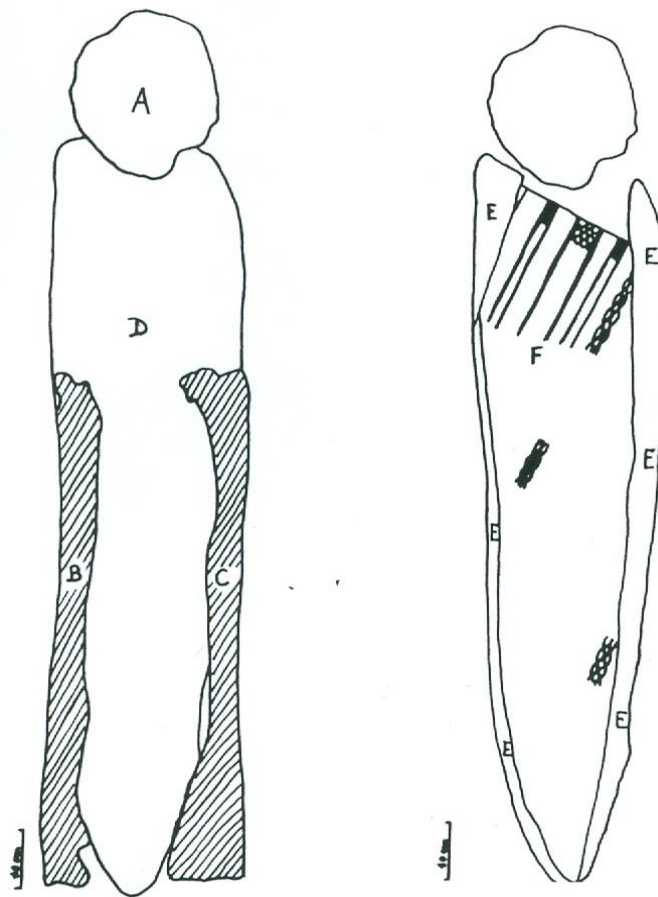
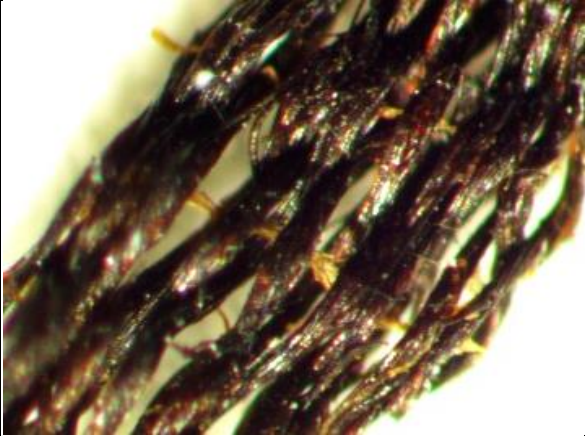
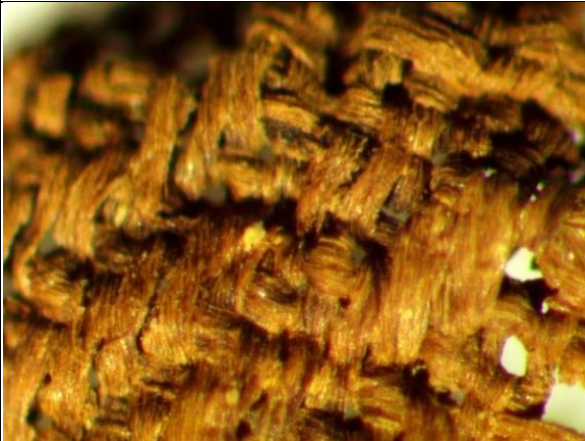
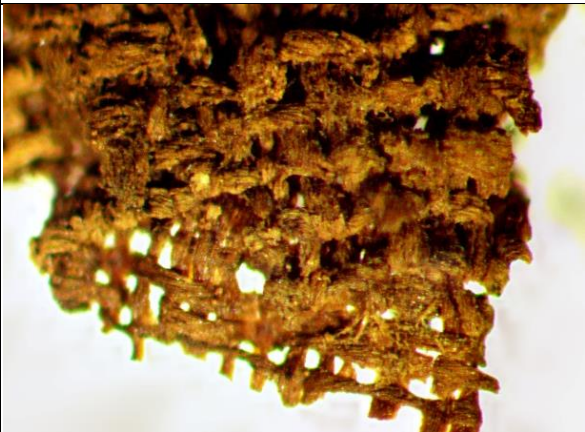
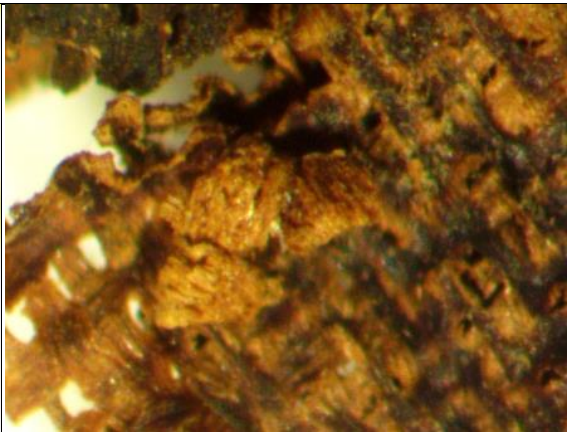


Figure 112. Drawing of the child's grave. A) textile around the head B, C) cloth along each legs D) outer wrapping in silk damask; and underneath this, E) large bands in gold tapestry placed on each side underneath the damask along the entire length of the body F) a narrower band with gold tapestry.⁴⁷⁴

⁴⁷⁴ Ibid.

N2702	
Dark gauze	
	
Layer 1	
	
Layer 2	
	
Resin	

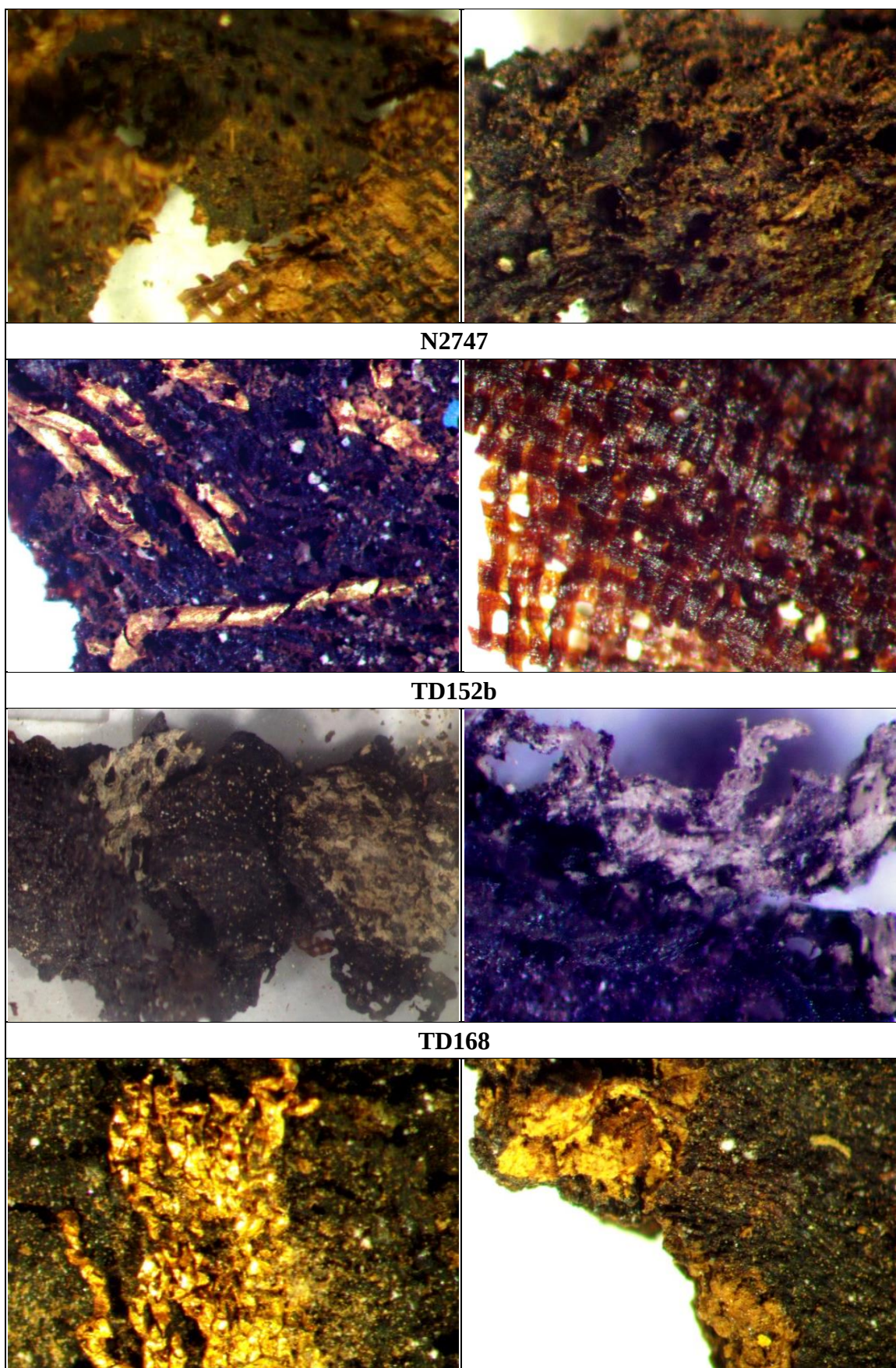


Figure 113. Different types of textiles found in Naintré samples.

VI.1.2 Experimental

Samples were taken from the debris of the selected specimens. 10 mgs of pulverised fibres were taken into glass tubes and treated with 2ml methanol at 60°C for 10 minutes. As most of the samples were mixed with soil and other contaminants, samples were not weighted. Some of the samples induced a change in colour following the methanol treatment, probably from the dyes used for each textile. The coloured liquids were removed to clean tubes, and 1ml of fresh methanol was added with 1ml of 7M calcium nitrate solution; the mixture was heated at 125°C for 20 minutes. Once cooled, solutions were centrifuged for 10 minutes at 14000 rpm, and the supernatants were transferred to clean tubes and the leftover (soil and organic matter such as resin) was discarded.

Fibre solutions were desalted using NWCO 3K ultrafiltration tubes. Solutions were mixed with 1:1 volume of 50mM ammonium bicarbonate buffer and spun for 30 minutes; then another volume of buffer was added to the tube and spun for 45 minutes. The concentrates were collected into new tubes and dried in a vacuum centrifuge. The final volume of the samples was adjusted to 50µl using the same buffer.

10µl of each sample was digested with trypsin and chymotrypsin (1:1). The resulted peptides were cleaned using ZipTip and submitted to nanoLC-MS/MS for peptide sequencing; injection blanks were placed before and after each sample.

The silk peptides were analysed by the high-resolution LC-MS/MS, with BSA as a standard and blanks (elution buffer only) before and after each sample. Resulted MS/MS spectra were analysed by PEAKS and Proteome Discover software using *De novo* method and the silk FASTA database (see III.7.1).

VI.1.3 Results and discussion

VI.1.3.1 Solubilisation behaviour

During the swelling, N2702-O (outer layer, dark gauze) made the methanol turn light blue, and N2702-I (inner layer, fine plain weave) made the liquid turn yellow, showing that trace of dyes has been dissolved into the methanol. The former could be the shellfish purple; the latter is red polychromatic dyestuff: the methanol extractions shifted colours when a small amount of acetic acid was added to adjust the pH to 3; the blue liquid turned yellow, and the yellow liquid turned red. The solubilisation in 7M $\text{Ca}(\text{NO}_3)_2$ solution was almost instant. However, N2702-O (outer layer) did not react to the solvent for silk; whereas other fibre samples were well-solubilised yielding reddish-brown liquid. Thus, this fibre is not silk; however, it is too fine to be of vegetable or animal origin. TD152 solubilised and resulted in a reddish brown to amber colour. TD168 yielded a lighter brown shade solution from the thread used as a core for the gold thread (Table 31). The methanol extracts can be separately examined with other instruments (i.e. HPLC) for dye analysis; however, it was not implemented in this research.

Table 31. Naintr  site textile samples and their solubilisation behaviour

Sample	Description	Solubilisation behaviour	
		In MeOH	In $\text{Ca}(\text{NO}_3)_2$
N2702-O	Dark gauze (outer layer)	Liquid turned blue	No change
N2702-I	Layer 1 damask (inner layer)	Liquid turned yellow	Fibres solubilised to reddish brown liquid, leaving soil etc.
N2747	Tabby weave	Liquid turned yellow	Fibres solubilised to reddish brown liquid, leaving soil etc.
TD152b	Textile debris (chunk)	Liquid turned yellow	Fibres solubilised to amber liquid, leaving dark matters
TD168	Gold threads	No visible change	Fibres solubilised to reddish brown liquid, leaving gold and soils

VI.1.3.2 Peptide analysis

The PEAKS analysis results suggested that N2702-I, N2747, TD152b and TD168 are highly likely to be the Chinese silk, as the identification result showed the highest probability for *B. mori* heavy chain. N2747 was also processed with Protein gel, showing the distribution that is highly likely to be of *B. mori* silk (Figure 34). The light chain peptides were only observed in N2702-I and N2747, indicating that these textiles are experiencing the different phases of degradation (Table 32). All peptides corresponding to P05790 in all samples were unique to *Bombyx* silks. However, as the limited presence of the light chain implies, each sample showed a different distribution of heavy chain peptides. In N2702-I, VAADAGAYSQSGPYVSNSGY (116-135) was the highest scoring peptide with -10lgP score of 49.05, followed by GQGAGSAASSVSSASSR (5215-5231) but with 10lgP score of 30.63. In N2747, GAASGAGAGAGAGAGAGY (172-189) was the highest scoring peptide with -10lgP score of 40.85; GGASGAGAGAGAGAGAGAGAGY (342-363) was the next showing -10lgP score 39.73. On the other hand, in TD152b and TD168, GAGAGAGAGSGAGAGSGAGAGSGAGSG (5035-5061) was the highest scoring peptide with -10lgP score of 43.75 and 49.60 respectively. The samples still showed a trace of the light chain showed more heavy chain peptides from the non-repetitive region near C-terminus.

Table 32. Naintr  site samples: a summary of PEAKS results (Peptides -10lgp 20)

Sample	Accession	-10lgP	Coverage (%)	#Peptides	#Unique	Avg. Mass
N2702-I	P05790 FIBH_BOMMO	106.48	2	9	9	391610
	P12828 FIBL_BOMMA	67.14	4	2	2	27639
N2747	P05790 FIBH_BOMMO	187.99	18	119	119	391610
	P12828 FIBL_BOMMO	96.92	41	10	10	27639
TD152b	P05790 FIBH_BOMMO	166.39	8	42	42	391610
TD168	P05790 FIBH_BOMMO	176.84	12	77	77	391610

Besides the *B. mori* proteins, there were traces of other peptides suggested to be that of wild silks. However, their presence and probability scores were very minor, appearing to be the peptide carryover from the previous samples. This was confirmed by checking the injection blanks before and after each sample.⁴⁷⁵ As an example, the protein analysis result of N2747 using *de novo* analysis based on the silk database and UniProtKB database search were compared (Table 33). Where there were three peptides matches to *A. assamensis* in the *de novo* analysis using silk database, there was no identification match for non-mulberry silks in the *de novo* analysis using UniProtKB; however, other proteins including human, mouse and spider web were suggested. The majority of human contamination is assumed to be from the contamination during the sample handling, but the child's remain that was in contact with the textiles might have affected this too. Traces of mouse keratin and *Nephila* silk protein (spidroin-1) are probably originated from the burial environment. Remains of trypsin and chymotrypsin were also detected. This example suggests that for protein analysis of an unknown subject, such as highly contaminated archaeological samples—especially when too fragile to withstand cleaning—it may be useful to compile multiple analysis using different databases, or software using different algorithms.

The peptide distribution of the archaeological silk was compared the fresh modern silk, using two samples of similar protein coverage for *B. mori* heavy chain (P05790): TD168 showed 26% protein coverage and the tryptic-chymotryptic in-gel digested fresh *B. mori* sample showed 25% protein coverage. Although the initial mass of the sample was 2-3 times bigger in TD 168(considering the weight of soil and contamination), the -10lgP score was much higher in the fresh sample, indicating that the protein content in the archaeological sample is highly degraded. The two shared 30 supporting peptides for the *B.*

⁴⁷⁵ Hendy, *et al.*, 2018

mori heavy chain, out of 187 peptides from TD168 and 378 peptides from the modern silk. Among the peptides unique to each sample, 148 peptides from TD 168 (79%) and 292 peptides from the fresh silk (77%) occurred in the highly repetitive region (149-5206).

However, the distribution of the peptides across the entire sequence in the two samples showed a different aspect. In fresh silk, 53% of the peptides occurred in the first 1000 residues of the sequences and 24% occurred in the residues 1001-2000. Although the number of peptides decreased along the sequence, in TD168, 28% of the peptides occur in the first 1000 residues with the remaining relatively evenly distributed. Moreover, despite the lower -10lgP score and a smaller number of peptides detected, the archaeological sample showed more peptides occurred between the residues 3000-5000 than the fresh silk; implying that the degradation of the protein made it more susceptible to solubilisation.

Table 33. Comparison of protein identification results of N 2747 (*B.mori*) by PEAKS using *De novo* vs. UniProtKB (peptides -10lgP≥15)

Accession	-10lgP	Coverage (%)	#Peptide	#Uniq	PTM	Avg. Mass	Description
<i>De novo</i> (4)							
P05790 FIBH_BOMMO	187.99	29	238	238	Y	391610	Fibroin heavy chain OS=Bombyx mori OX=7091 GN=FIBH PE=1 SV=4
Q9BLL9 FIBL_BOMMA	96.92	46	22	22	Y	27639	Fibroin light chain OS=Bombyx mandarina OX=7092 GN=Fib-L PE=2 SV=1
P21828 FIBL_BOMMO	96.92	46	22	22	Y	27669	Fibroin light chain OS=Bombyx mori OX=7091 GN=FIBL PE=1 SV=1
A0A0K0KR73 ANTASA	33.95	3	3	3	N	229493	Fibroin heavy chain OS=Antheraea assamensis OX=91021 GN=Fhc PE=4 SV=1
<i>UniProtKB</i> (40)							
P05790 FIBH_BOMMO	207.18	25	217	217	Y	391610	Fibroin heavy chain OS=Bombyx mori OX=7091 GN=FIBH PE=1 SV=4
P21828 FIBL_BOMMO	119.42	46	26	26	Y	27669	Fibroin light chain OS=Bombyx mori OX=7091 GN=FIBL PE=1 SV=1
P04264 K2C1_HUMAN	108.8	20	17	10	N	66039	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6
P48668 K2C6C_HUMAN	88.97	12	8	1	N	60025	Keratin, type II cytoskeletal 6C OS=Homo sapiens GN=KRT6C PE=1 SV=3
P02538 K2C6A_HUMAN	88.97	12	8	1	N	60045	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3
P35527 K1C9_HUMAN	85.89	21	11	10	N	62064	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3
P04259 K2C6B_HUMAN	85.43	9	7	0	N	60067	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5
P35908 K22E_HUMAN	82.3	8	7	1	N	65433	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2
P50446 K2C6A_MOUSE	78.82	6	5	0	N	59335	Keratin, type II cytoskeletal 6A OS=Mus musculus GN=Krt6a PE=1 SV=3
Q9Z331 K2C6B_MOUSE	78.82	6	5	0	N	60322	Keratin, type II cytoskeletal 6B OS=Mus musculus GN=Krt6b PE=1 SV=3
Q6IG05 K2C75_RAT	78.82	6	5	0	N	59027	Keratin, type II cytoskeletal 75 OS=Rattus norvegicus GN=Krt75 PE=3 SV=2
Q8BGZ7 K2C75_MOUSE	78.82	6	5	0	N	59741	Keratin, type II cytoskeletal 75 OS=Mus musculus GN=Krt75 PE=1 SV=1
P08779 K1C16_HUMAN	70.85	17	7	2	Y	51268	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4
Q8VED5 K2C79_MOUSE	69.42	6	5	0	N	57552	Keratin, type II cytoskeletal 79 OS=Mus musculus GN=Krt79 PE=1 SV=2
P02533 K1C14_HUMAN	67.15	15	7	2	Y	51562	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4
P13645 K1C10_HUMAN	60.41	13	7	5	Y	58827	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6
P00766 CTRA_BOVIN	60.02	30	8	8	Y	25666	Chymotrypsinogen A OS=Bos taurus PE=1 SV=1
P00761 TRYP_PIG	47.34	16	5	5	N	24409	Trypsin OS=Sus scrofa PE=1 SV=1
P81605 DCD_HUMAN	39.31	19	2	2	N	11284	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2
A5A6M0 K1C17_PANTR	38.58	7	3	1	N	48121	Keratin, type I cytoskeletal 17 OS=Pan troglodytes GN=KRT17 PE=2 SV=1
Q04695 K1C17_HUMAN	38.58	7	3	1	N	48106	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=1 SV=2
Q9QWL7 K1C17_MOUSE	38.58	7	3	1	N	48162	Keratin, type I cytoskeletal 17 OS=Mus musculus GN=Krt17 PE=1 SV=3
Q6IFU8 K1C17_RAT	38.58	7	3	1	N	48123	Keratin, type I cytoskeletal 17 OS=Rattus norvegicus GN=Krt17 PE=1 SV=1
A1L595 K1C17_BOVIN	38.58	7	3	1	N	48712	Keratin, type I cytoskeletal 17 OS=Bos taurus GN=KRT17 PE=2 SV=1
Q6X2U2 POLN_RUBVC	29.16	0	1	1	N	230397	Non-structural polypeptide p200 OS=Rubella virus (strain BRDII) PE=3 SV=1
Q88DD1 MUTL_PSEPK	28.57	3	3	3	Y	69745	DNA mismatch repair protein MutL OS=Pseudomonas putida (strain ATCC 47054 / DSM 6125 / NCIMB 11950 / KT2440) GN=mutL PE=3 SV=1

<u>B0KKZ7 MUTL_PSEPG</u>	28.57	3	3	3	Y	69686	DNA mismatch repair protein MutL OS=Pseudomonas putida (strain GB-1) GN=mutL PE=3 SV=1
<u>P19837 SPD1_NEPCL</u>	26.79	4	2	2	Y	60585	Spidroin-1 (Fragment) OS=Nephila clavipes PE=1 SV=3
<u>P05886 ENV_SIVVT</u>	26.55	1	1	1	Y	99033	Envelope glycoprotein gp160 OS=Simian immunodeficiency virus agm.vervet (isolate AGM TYO-1) GN=env PE=3 SV=1
<u>O00570 SOX1_HUMAN</u>	25.82	8	2	1	N	39023	Transcription factor SOX-1 OS=Homo sapiens GN=SOX1 PE=1 SV=2
<u>Q2YAV0 HIS6_NITMU</u>	25.09	3	1	1	N	26975	Imidazole glycerol phosphate synthase subunit HisF OS=Nitrosospira multiformis (strain ATCC 25196 / NCIMB 11849 / C 71) GN=hisF PE=3 SV=1
<u>Q38897 BEL1_ARATH</u>	24.68	1	1	1	N	69493	Homeobox protein BEL1 homolog OS=Arabidopsis thaliana GN=BEL1 PE=1 SV=2
<u>C4LIW5 SYH_CORK4</u>	22.53	3	1	1	N	46770	Histidine--tRNA ligase OS=Corynebacterium kroppenstedtii (strain DSM 44385 / JCM 11950 / CIP 105744 / CCUG 35717) GN=hisS PE=3 SV=1
<u>P67777 PP2AA_RABIT</u>	22.37	4	1	1	N	35594	serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform OS=Oryctolagus cuniculus GN=PPP2CA PE=2 SV=1
<u>Q18740 K1C9_CANLF</u>	21.21	1	1	0	N	76354	Keratin, type I cytoskeletal 9 OS=Canis lupus familiaris GN=KRT9 PE=3 SV=1
<u>A1VF83 PYRF_DESVV</u>	20.88	4	1	1	Y	24025	Orotidine 5'-phosphate decarboxylase OS=Desulfovibrio vulgaris subsp. vulgaris (strain DP4) GN=pyrF PE=3 SV=1
<u>Q1GSD9 SYA_SPHAL</u>	20.81	2	1	1	N	94347	alanine--tRNA ligase OS=Sphingopyxis alaskensis (strain DSM 13593 / LMG 18877 / RB2256) GN=alaS PE=3 SV=1
<u>A1SLW5 LEUC_NOCSI</u>	20.72	3	1	1	N	50124	3-isopropylmalate dehydratase large subunit OS=Nocardioides sp. (strain ATCC BAA-499 / JS614) GN=leuC PE=3 SV=1
<u>Q9HGI8 ERF3_KLULA</u>	20.52	1	1	1	Y	77639	Eukaryotic peptide chain release factor GTP-binding subunit OS=Kluyveromyces lactis (strain ATCC 8585 / CBS 2359 / DSM 70799 / NBRC 1267 / NRRL Y-1140 / WM37) GN=SUP35 PE=3 SV=1
<u>Q59ZV5 EIF3G_CANAL</u>	20.48	5	1	1	N	30686	Eukaryotic translation initiation factor 3 subunit G OS=Candida albicans (strain SC5314 / ATCC MYA-2876) GN=TIF35 PE=3 SV=1

VI.2 The textiles from Palmyra

VI.2.1 Introduction

Palmyra is an ancient oasis city in the present-day Homs province of Syria, north-east of Damascus. It was located halfway between the Euphrates and the Eastern Mediterranean coast.⁴⁷⁶ Archaeological finds date back the original foundation of Palmyra to the Neolithic period. The Mari Archives first mentioned the city in the document in the early 2nd millennium BCE. Palmyra came under the Roman control in the mid-1st century CE, as a part of the Roman province of Syria.⁴⁷⁷ It became an established caravan oasis during the Roman imperial times, interlinking city of routes to Persia, India and China with the Roman Empire. In consequence, the art and architecture of Palmyra adopted Graeco-Roman techniques and blended with local traditions and Persian influences.⁴⁷⁸ This is reflected in the way Palmyrene people dressed: some preferred the Parthian style of trousers and caftan, others the Roman tunic and the Greek himation.⁴⁷⁹

The textiles from Palmyra give us a clear and unique insight into the economy, trade and textile production of antiquity. So far, there are more than 2,000 textile fragments of over 500 different types were found from the 9 different tower tombs, comprising one of the largest groups of dated antique textiles with a proven origin.⁴⁸⁰ They are dated to 1st century BCE to the 2nd century CE, as the tower tombs are dated to this period by their foundation inscriptions. Parts of them were found by a French mission and studied by Rudolf Pfister, exhibited and deposited in the National Museum of Damascus.⁴⁸¹ Another series of textiles

⁴⁷⁶ Stoneman, 1994

⁴⁷⁷ Gawlikowski, 2014

⁴⁷⁸ Ibid.; Aruz, *et al.*, 2018

⁴⁷⁹ Schmidt-Colinet, *et al.*, 1992; Schmidt-Colinet and Gaballa, 1995

⁴⁸⁰ Schmidt-Colinet, *et al.*, 2000

⁴⁸¹ Pfister, 1934, Pfister, 1937, Pfister, 1940

from different tower tombs have been recovered by Syrian, Polish, and German archaeologists. An international project led by Khaled al-As'ad, the director of Antiquities of Palmyra, Andreas Schmidt-Colinet and Annemarie Stauffer of Berne University.⁴⁸²

The textiles from Palmyra can be grouped into two: the textiles produced locally, and the woven silks imported from China. This reflects the wealth and the taste of aristocratic elites who could afford and had connections to acquire the luxurious silks, but not the general fashion of Palmyrene people unlike the cases of contemporary finds from Mons Claudianus⁴⁸³ and Masada.⁴⁸⁴ Palmyrene people cut precious garments into strips and used in burial as a decoration over the body wrapped in several layers of plain clothes. This custom explains why many of the textiles found at Palmyra are in the form of bands in various sizes.⁴⁸⁵ It has been reported that there are some specimens of wild silk found among the Palmyra textiles: “tussah” silks were used for plain-woven, monochrome fabrics of fine quality.⁴⁸⁶ And there is a union fabric made of wild silk warp and cotton weft, which probably was made to imitate the more expensive combination of *Bombyx* silk and cashmere hair. However, the origin of the material or the place of fabrication has not been identified.⁴⁸⁷

I inquired Professor Annamarie Stauffer of *Technische Hochschule Köln* if she still has access to any of the materials. She kindly offered 5 different samples of alleged “Tussha” silk. Sample S8-warp and weft from the Schmidt-Colinet/Stauffer Katalog Nr. 305, Sample S48 from the Schmidt-Colinet/Stauffer Katalog Nr. 312, and S49-system 1 and 2 from the Schmidt-Colinet/Stauffer Katalog Nr.313. Although the amount of the samples was minimum, as about 0.1mg each, it was enough for a single test. S8 is made of very shiny,

⁴⁸² Schmidt-Colinet, *et al.*, 2000

⁴⁸³ Jørgensen, 2018

⁴⁸⁴ Sheffer and Granger-Taylor, 1994

⁴⁸⁵ Schmidt-Colinet, *et al.*, 2000; Stauffer, 2000

⁴⁸⁶ Schmidt-Colinet, *et al.*, 2000

⁴⁸⁷ Stauffer, 2000

flat fibres: only the weft has slight torsion in the Z direction, thus less shiny than the warp.

S48 and S49 have similar looking flat fibres for both warp and weft (Figure 114).

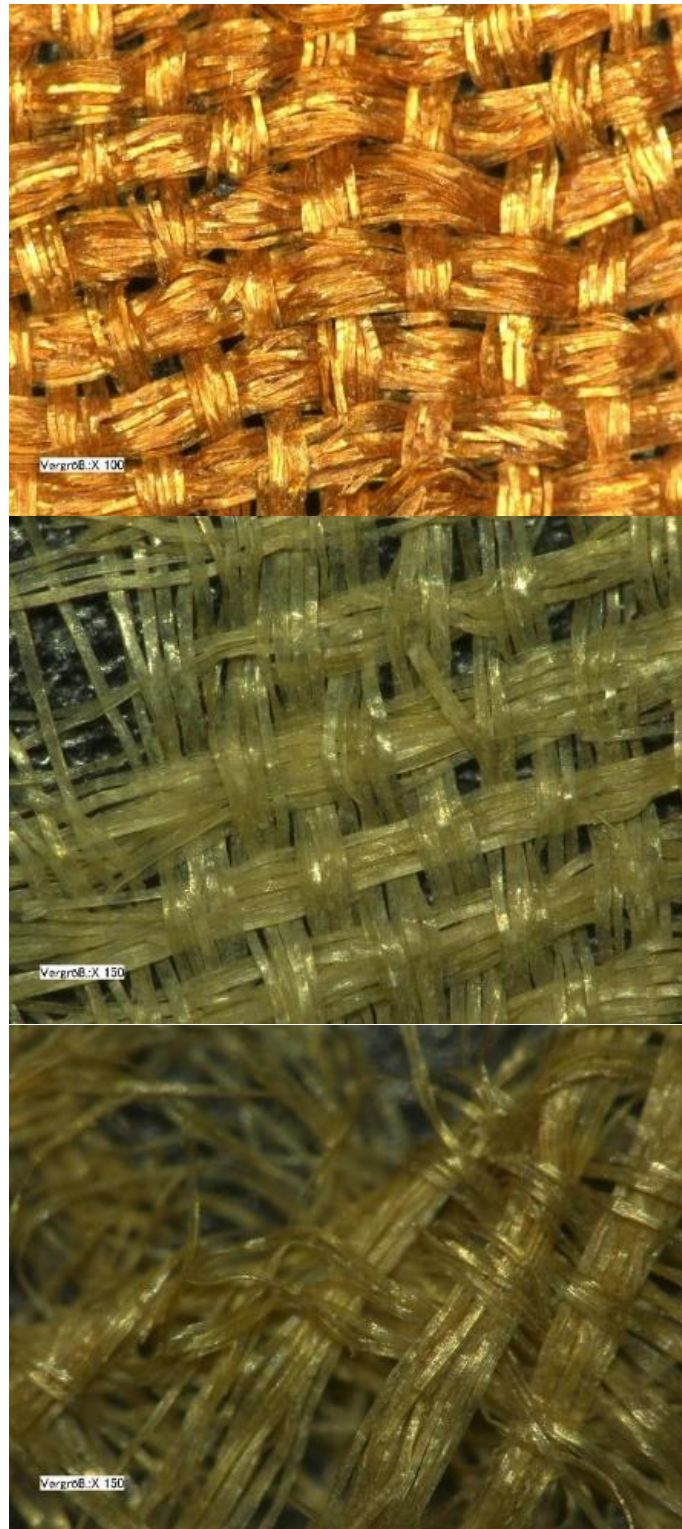


Figure 114. Microscopic images of the Palmyra samples: top) Sample S 8; Schmidt-Colinet/Stauffer Katalog Nr. 305, middle) Sample S48; Schmidt-Colinet/Stauffer Katalog Nr. 312, bottom) Sample S49; Schmidt-Colinet/Stauffer Katalog Nr. 313.

VI.2.2 Experimental

Other than S8, samples only contained a few strands of fibres thus the weights were not measured to prevent any loss. However, the samples were clean and thus directly moved to clean glass tubes and heated at 125°C for 30-40 minutes in 300µl of 7M calcium nitrate solution. Some of the samples showed a colour shift following the methanol treatment, probably caused by the extraction of the dyes. Due to the low volume of the protein concentration, desalting and digestion steps were modified based on the Filter Aided Sample Prep (FASP) method.⁴⁸⁸ The silk solution must be diluted prior to the use of a filter unit due to the high concentration of the salt.

The fibroin solutions were diluted to 1ml volume with 20mM concentration TCEP in 100mM ammonium bicarbonate solution, sonicated for 30 seconds and incubated for 30 minutes at room temperature. Fibroin solutions were moved to NWCO 3K ultrafiltration tubes (~4ml) with two volume of 8M urea in 100mM ammonium bicarbonate solution, vortexed and spun at 4500g for 60 minutes. This process may take longer than the suggested time, as the concentration of the calcium salt slows the buffer exchange process. Once the volume was decreased to near 1ml, one volume of 4M urea in 100mM ammonium bicarbonate solution was added, vortexed and spinned for 30 minutes. The same step using 2M urea buffer was implemented. For the alkylation, 1ml of 10mM IAA solution in 1M urea and 100mM ammonium bicarbonate solution concentration was added, vortexed and incubated in dark at room temperature for 20 minutes; then spun for 30 minutes. Lastly, one volume of 100mM ammonium bicarbonate solution was added to the filters, vortexed

⁴⁸⁸ Wiśniewski, *et al.*, 2009

and spinned for 30 minutes: this step was repeated once more. The concentrates were then mixed with the trypsin/chymotrypsin mixture and incubated overnight at 37 °C (see III.5.2.2 In-solution digestion protocol). The filtrates were collected by spinning for 30 minutes: first it was spinned as it is, then 500 µl of MilliQ water was added to the filters, vortexed and spun again for 30 minutes. The collected peptides were completely dried in vacuum centrifuge and eluted to 100µl with 0.1% formic acid. The samples were ZipTip cleaned (see III.5.2.3. ZipTip process). 10µl of each sample was submitted to nanoLC-MS/MS for peptide sequencing; injection blanks were placed before and after each sample. Resulted MS/MS spectra were analysed by PEAKS and Proteome Discover software using *De novo* method and the silk FASTA database.

VI.2.3 Results and discussion

VI.2.3.1 *Solubilisation behaviour*

All five samples from 3 textiles were successfully solubilised in the 7M calcium nitrate solution under the heat. However, the process took much longer than the Naintr  samples, indicating that these are non-Bombyx silks. The fibres started to melt within the first 10 minutes; at 20 minutes it was observed that the large part of the fibres was dissolved, however, the mass had no fluidity; it took more than 30 minutes of heating to solubilise the fibres until the solutions exhibit no solid mass (Figure 115). The yielded solutions of S8 (warp and weft both) and S48 showed pale pinkish-yellow colour, exhibiting the presence of dyes. S49 could also have been dyed but it was hard to notice any change with naked eyes because the quantity of the sample was too small compared to the other two.



Figure 115. Before (left) and after (right) the solubilisation in 7M $\text{Ca}(\text{NO}_3)_2$ solution: sample S8-weft, taken after 20mins of boiling

VI.2.3.2 Peptide analysis

The PEAKS analysis result of the samples showed relevant results to the speculations of the previous studies that these are the “tussah” silk (Table 34). In the database search using NCBI nr, none of the samples showed any match to silk proteins. However, in *de novo* analysis, all samples were suggested to be *Antheraea* silks. The unique sequence features of *Antheraea* silk fibroins, SGAGG and A_{n≥8} were found frequently in all samples, supporting that the samples are of *Antheraea* origin. There were also a few numbers of peptides corresponding to *Samia ricini* fibroin; in most cases, these were confirmed to be the contaminations from the blank samples ran before and after these samples. The relatively low -10lgP score in S 48 and S49-T2 in comparison with the other three samples seem to reflect the very small sample quantity.

Table 34. PEAKS result summary of the Palmyra samples (FDR=0%)

Sample	Accession	-10lgP	Coverage (%)	#Peptides	#Uniq.
S8 warp	Q8ISB3 ANTMY	163.51	27	16	8
	O76786 ANTPE	160.3	5	18	4
	E1CGA3 ANTYA	151.28	4	14	1
	A0A0K0KR73 ANTASA	121.62	4	7	2
S8 weft	Q8ISB3 ANTMY	238.35	29	20	13
	O76786 ANTPE	197.23	5	11	2
	E1CGA3 ANTYA	195.64	6	13	3
S48	O76786 ANTPE	94.79	6	9	5
	Q8ISB3 ANTMY	90.8	22	7	4
	A0A0D5ZYI3 SAMRI	70.79	3	4	4
S49 T1	Q8ISB3 ANTMY	178.49	23	10	9
	E1CGA3 ANTYA	172.99	4	8	3
	O76786 ANTPE	170.24	3	7	2
S49 T2	O76786 ANTPE	79.96	4	7	1
	Q8ISB3 ANTMY	79.49	23	8	3
	A0A0K0KR73 ANTASA	70.32	4	6	1

In S8, the warp was identified as *A. mylitta* with -10lgP score of 163.51. It showed 29% sequence coverage for Q8ISB3 (partial, 507 residues) with 8 unique peptides to the protein. The highest scoring peptide among the unique peptides was SA₁₂GSGAGGRGD (-10lgP=47.33), which is shared among *Antheraea* species. The highest scoring unique peptide for Q8ISB3 was SGA₁₃SGAGGR (-10lgP=46.77) that appears twice in the protein sequence at residues 243-263 and 419-439, followed by SSA₁₁SAAR (-10lgP=44.12) that carries an exclusive marker SSA₁₁S at the residues 484-499 (Figure 116). The result also showed less than 5% coverage for other *Antheraea* and *Samia* fibroins with small numbers of unique peptides.

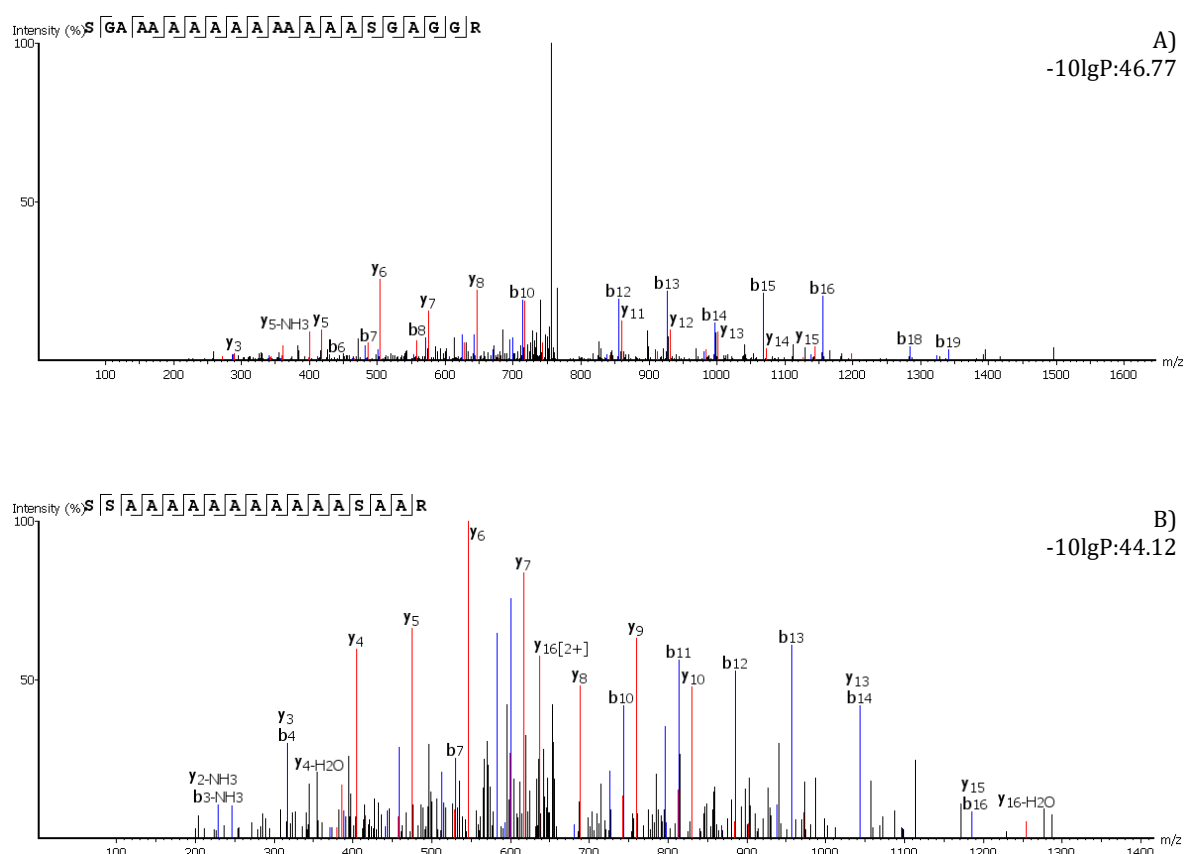


Figure 116. Spectra of the highest scoring unique peptides for Q8ISB3 in S8 warp: A) SA₁₂GSGAGGRGD (m/z=780.4), B) SSA₁₁SAAR(m/z=680.35), z=2.

The weft was also identified as *A. mylitta* with -10lgP score of 238.35. It showed 29% sequence coverage for Q8ISB3, with 13 unique peptides to Q8ISB3. Although the result also showed 5% coverage for other *A. pernyi* fibroin and 6% coverage for *A. yamamai* fibroin with small numbers of unique peptides, the protein probability score for Q8ISB3 was dominant. The highest scoring peptide among the unique peptides was GGYGSGSSA₁₁SAAR at residues at 478-500 (-10lgP=70.75): the exclusive marker SSA₁₁S was also found in this peptide (Figure 117).

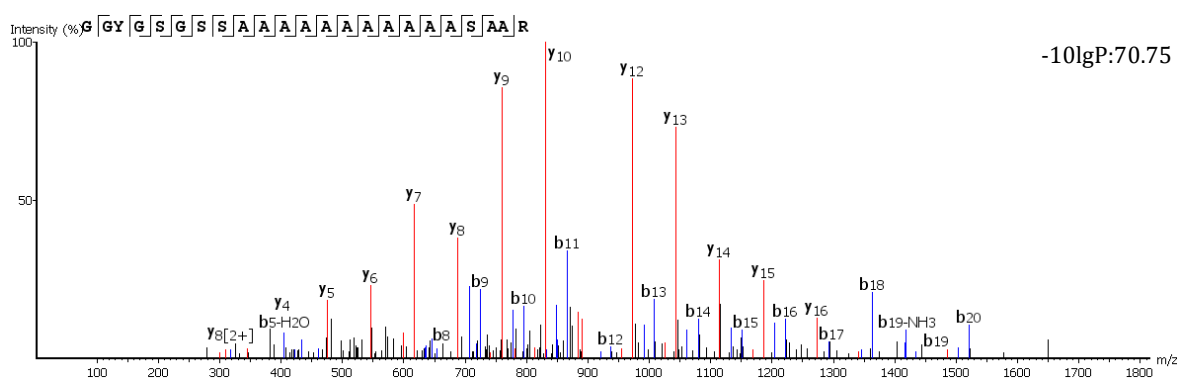


Figure 117. Spectrum of the highest scoring unique peptide for Q8ISB3 in S8 weft: GGYGSGSSA₁₁SAAR ($m/z=919.44$, $z=2$).

Overall, the peptide analysis result strongly suggests that both the warp and weft of S8 are *A. mylitta* silk. In both samples, the peptides matched to *A. mylitta* were dominant up to the 507th residues, and the rest of the peptides matched to other *Antheraea* fibroins mainly occurred after the 507th residue. Considering the sequence homology between the *Antheraea* silk fibroins, it is highly likely that these peptides are also present in unsequenced part of *A. mylitta* fibroin. For example, the unique sequence features of *A. pernyi* fibroin found in S8 warp and weft were containing either of the two unique markers of O76786—SA₁₃GSGAGGR at the residues 2162-2181 and SSA₁₀S that presents in O76786 total 4 times at the residues 547-559, 733-745, 919-931 and 2375-2387. In S8 warp, there was one

unique peptide corresponding to *A. assamensis* fibroin A0A0K0KR73, AAGSA₁₂GGSG (-10lgP=28.27). This peptide contains a unique marker AAGSA₁₂G, which presents in the protein sequence 6 times.

S48 showed the highest match to *A. pernyi* fibroin O76786 with 6% coverage and 5 unique peptides at protein probability score of 94.79. It also showed 22% coverage for Q8ISB3 with 4 unique peptides at protein probability score of 90.80, and 3% coverage for A0A0D5ZYI3 with 4 unique peptides at probability score of 70.79. Although *A. pernyi* fibroin was of the highest probability, it is early to conclude that S48 is of *pernyi* silk. Interestingly, the highest scoring peptide in this sample was GDGGA₁₃GGSRSGY (-10lgP=36.04) that presents in A0A0D5ZYI3 at the residues 1011-1034. It was followed by SGSSA₁₁SAAR (-10lgP=35.72) that is unique to Q8ISB3, and then SA₁₃GSGAGGR (-10lgP=35.48) that is unique to O76786. The distribution of the unique peptides reveals that the peptides corresponding to *A. mylitta* is more dominant until the 507th residue, and the peptides corresponding to *A. pernyi*, *A. assamensis* *S. ricini* were distributed after the 507th residue. All unique peptides to Q8ISB3 were containing the unique marker SSA₁₁S (484-499). Furthermore, the second highest scoring unique peptide to Q8ISB3 SGA₁₂SGAGGRGD (-10lgP=32.86) carries one of the peptide markers to distinguish *A. mylitta* and *A. pernyi* —GA₁₂—that is present in *A. mylitta*, *A. assamensis* and *S. ricini*, but not in *A. pernyi* (Figure 118). The unique peptides to O76786 were including the unique marker SA₁₃GSGAGGR (2162-2181), A₁₄GSGAGGV (1818-1838) and SRQV (2570-2573); however, none of them was corresponding to the first 507 residues. Thus, a hypothesis of S48 also being *A. mylitta* silk can be established based on these factors.

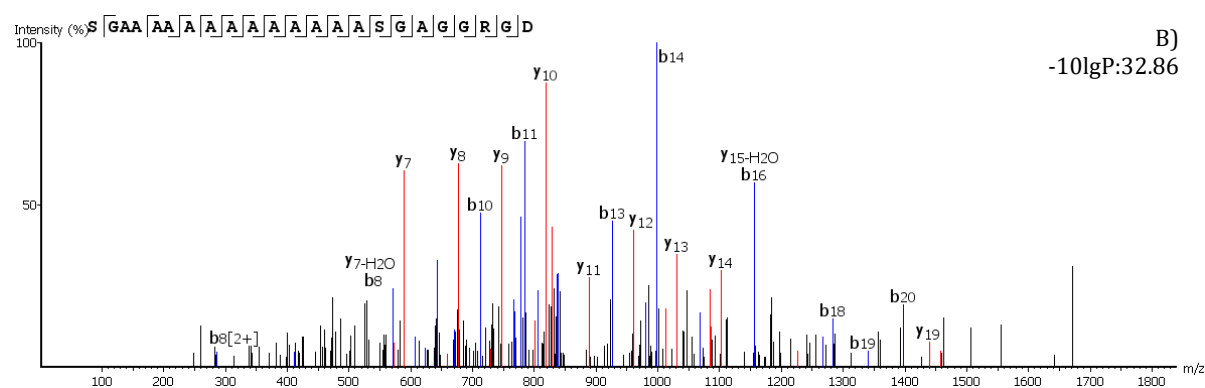
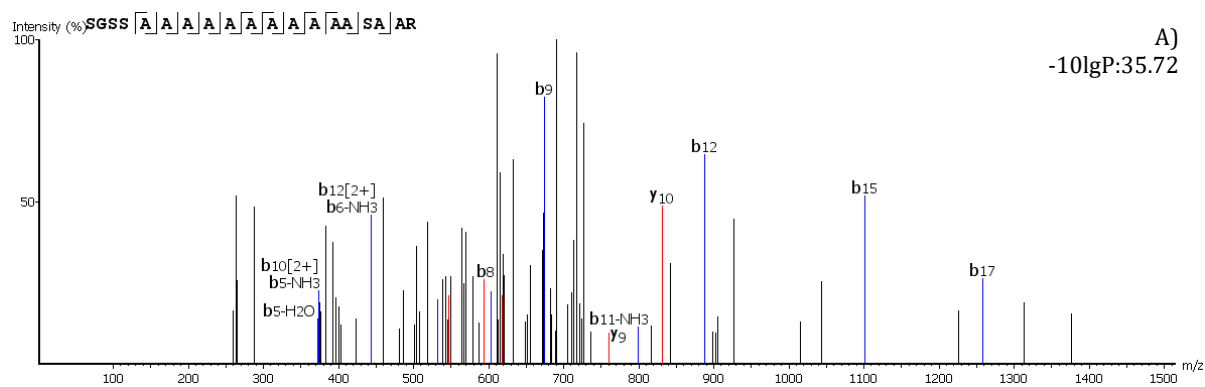


Figure 118. Spectra of the highest scoring unique peptides to Q8ISB3 in S48: A) **SGSSA₁₁SAAR** ($m/z=752.38$), B) **SGA₁₂SGAGGRGD** ($m/z=872.42$), $z=2$

In S49, T1 showed the highest match to *A. mylitta* fibroin, followed by *A. yamamai* and *A. pernyi*. The highest scoring peptide was AAGSA₁₃SGAGGSGGGY (-10lgP=69.63), that present in both E1CGA3 (230-256) and O76786(383-409); it is likely that this peptide also presents in *A. mylitta* fibroin, but after the 507th residue. The next highest scoring peptide was GSGSSA₁₁SAAR (-10lgP=64.48), which is exclusive to Q8ISB3 (Figure 119).

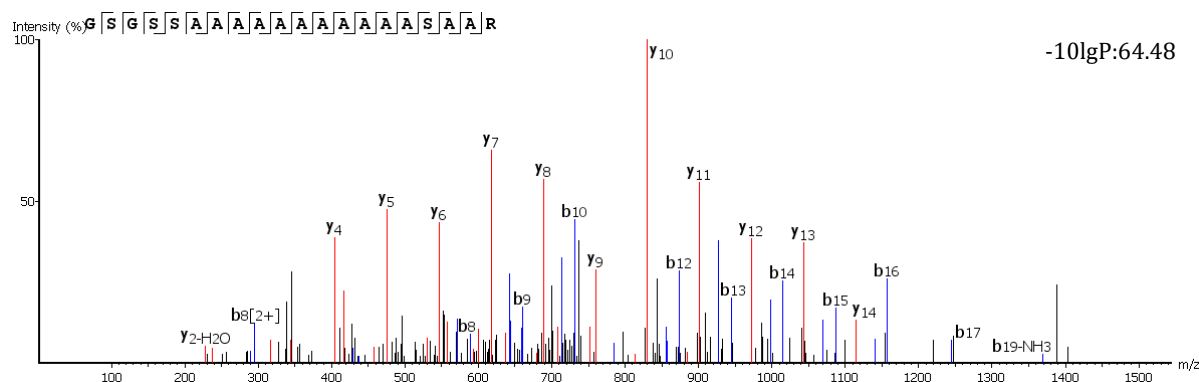


Figure 119. Spectrum of the highest scoring unique peptide for Q8ISB3 in S49-T1: GSGSSA₁₁SAAR ($m/z=780.89$, $z=2$).

T2 showed the highest match to *A. pernyi* fibroin, followed by *A. mylitta* and *A. yamamai*. However, the difference in the protein probability score between O76786 and Q8ISB3 was only 0.47. There was only one unique peptide to O76786, while there were 3 unique peptides corresponding to Q8ISB3. The highest scoring peptide was SA₁₂GSGAGGR (-10lgP=29.96), which is shared between O76786 and Q8ISB3; the second highest was SA₁₃GSGA, which is shared among O76786, Q8ISB3 and A0A0K0KR73. It was followed by the only unique peptide to *A. assamensis* fibroin A0A0K0KR73, AAGSA₁₂GGSG (-10lgP=25.04); this peptide was also found in S8 warp (-10lgP=17.29). The 4th highest was SA₁₁SAAR, a unique peptide to Q8ISB3 (-10lgP=23.30) (Figure 120). The only unique peptide corresponding to O76786 was GSDSA₁₃GSGAGGR (2158-2181): however, the identical peptide was found in all other samples; including S49-T1, S8 warp and weft that were identified to be *A. mylitta* fibroin (Figure 121).

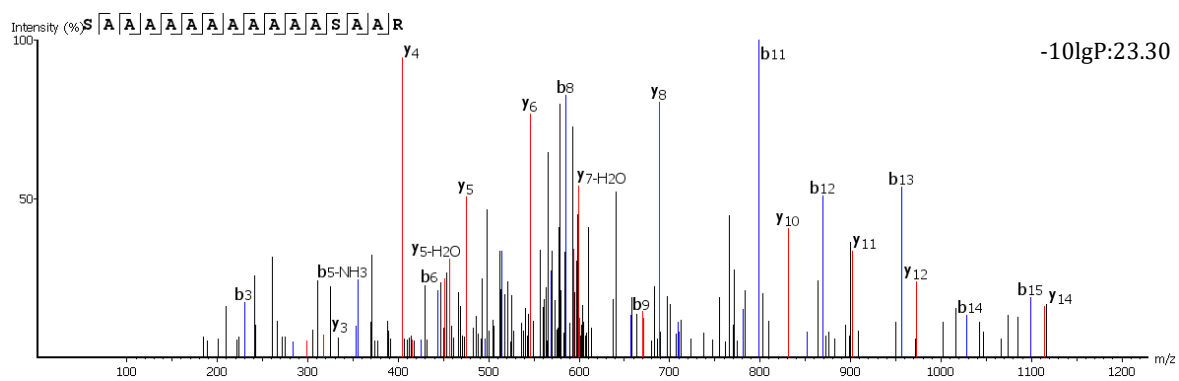
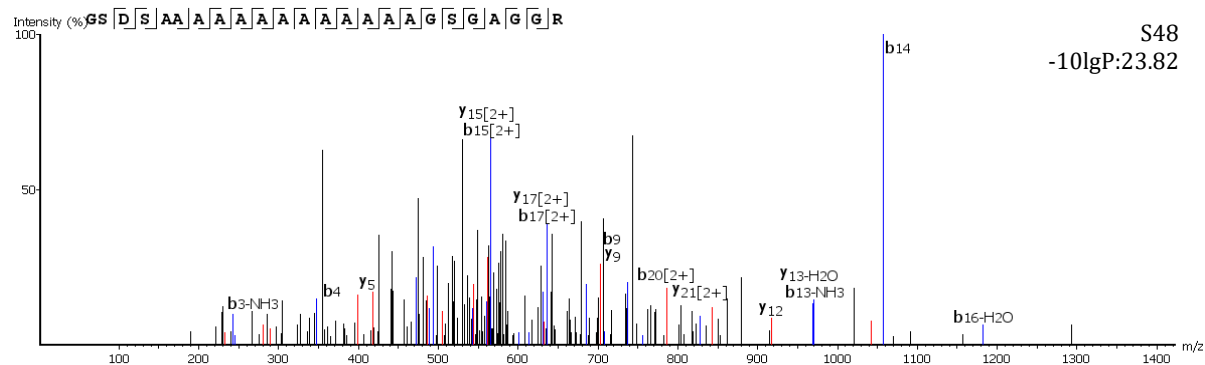
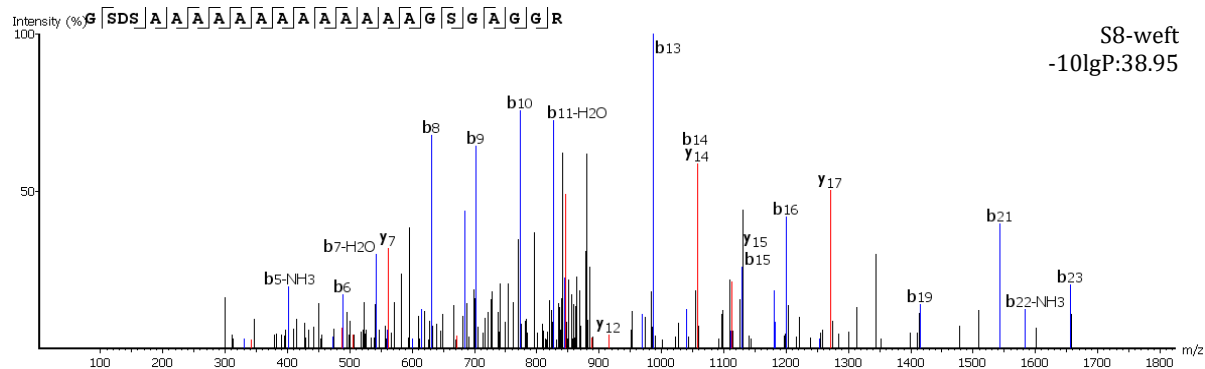
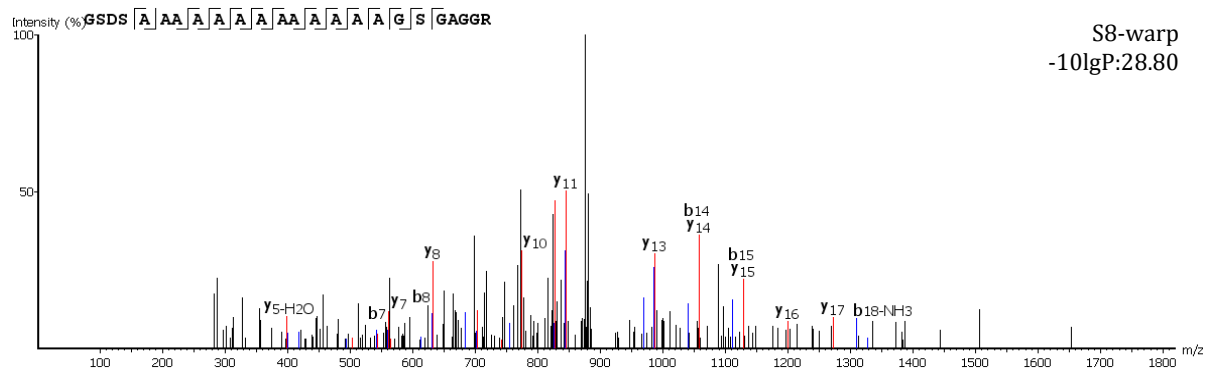


Figure 120. Spectrum of the highest scoring unique peptide for Q8ISB3 in S49-T2: SA₁₁SAAR ($m/z=636.84$, $z=2$).



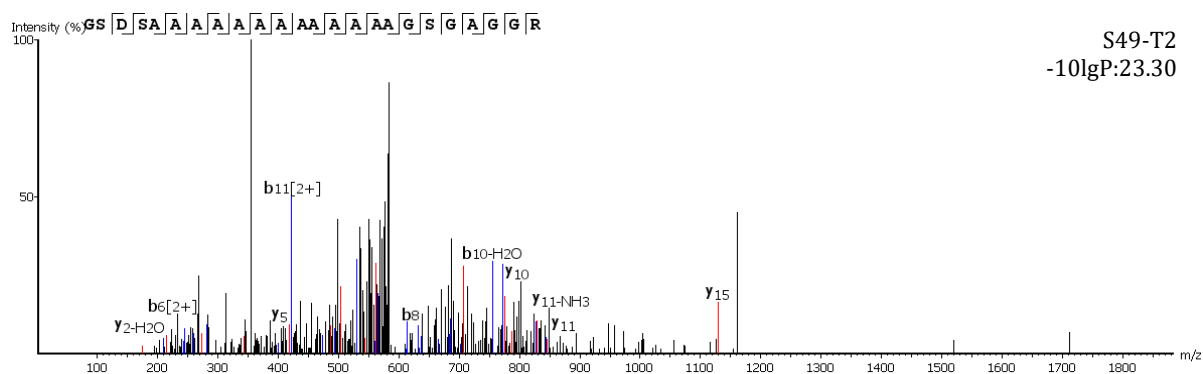
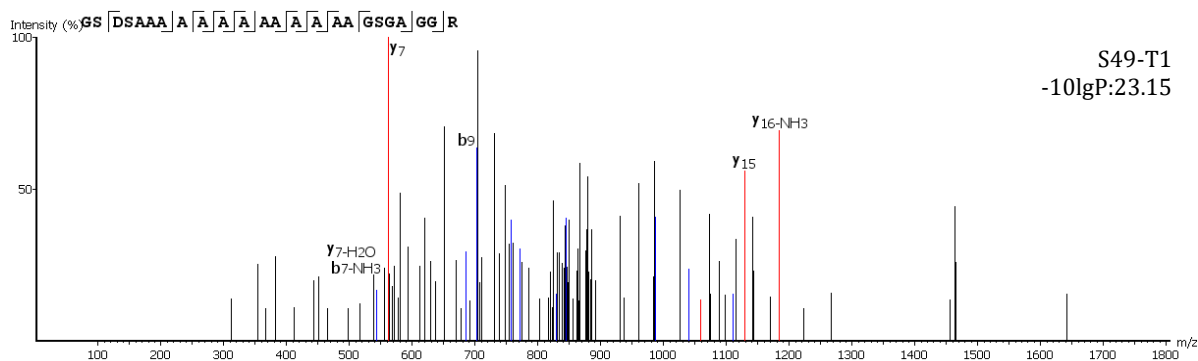


Figure 121. Spectrum of GSDSA₁₃GSGAGGR (a unique peptide to 076786) in all samples:
mass = 1829.86; m/z =610.96(z =3) or 915.93(z =2)

In summary, the Palmyra samples were identified to be either *A. mylitta* or *A. pernyi* silks by PEAKS analysis. The result confirms that S8 (both warp and weft) and S49-T1 is made of *A. mylitta* silk. And although S48 and S49-T2 were identified to be *A. pernyi* silk by PEAKS, the manually reviewed peptide analysis results suggest that these are also of *A. mylitta* silk: the peptides corresponding to *A. pernyi* fibroin O76786 in both samples were also found in the other three confirmed to be *A. mylitta* silk. However, it would be necessary to have the full protein sequence of *A. mylitta* to elucidate this matter. The minor portion of peptides corresponding to *S. ricini* in S8 warp and S48 is mainly from the contamination (checked with injection blanks in between each sample), but it is also possible that unsequenced part of the *A. mylitta* fibroin shares these features with *S. ricini* fibroin.

A. mylitta and *A. pernyi* are both called Tasar silk (tussah) but *pernyi* is of Chinese origin, and *mylitta* is of Indian origin (see II.1.2.1). The finding of Indian silks from Palmyra is a good example of their prospered international trade. And this is the first scientifically proven evidence of the exportation of raw silks and silk textiles from India as recorded in *Periplus Maris Erythraei*, which is about a century older than the height of the Palmyrene trade.

VII. Summary and conclusions

VII.1 Solubilisation of silk fibroins

A series of experiments to solubilise the silk fibroins demonstrated that a saturated solution of chaotropic salts, calcium nitrate (tetrahydrate) and lithium thiocyanate, are effective for *Bombyx*, *Antheraea*, and *Samia* silk fibroins. *B. mori* silk was soluble in calcium chloride-ethanol ternary solution and saturated solutions of chaotropic salts, lithium bromide, lithium thiocyanate and calcium nitrate (tetrahydrate). *B. mandarina* showed a similar behaviour to that of *B. mori*: however, in general, the reaction took place slightly slower. On the contrary, *A. pernyi*, *A. mylitta*, *A. assamensis*, *A. yamamai* and *S. ricini* silks did not react to calcium chloride-ethanol ternary solution and lithium bromide solution. These were solubilised in lithium thiocyanate solution and calcium nitrate solution, but only under high heat.

The advantage of lithium thiocyanate over calcium nitrate is the less viscosity that facilitates the desalting and concentration process. On the other hand, the advantage of calcium nitrate is that it does not interfere with dyes and coatings, while lithium thiocyanate can affect other organic materials and cellulosic fibres. In addition, calcium nitrate is much more economical than lithium thiocyanate. The addition of formic acid in the calcium nitrate solution facilitated the solubilisation process, as well as induced a shift in colour from transparent to yellow. This is thought to be the product of a Xanthoprotic reaction, due to the nitration of tyrosine chromophore. The yellowing of solution by addition of formic acid can confirm the existence protein in the solution, thus can be useful as an initial screening method for silk and non-silk; however, lithium thiocyanate yellows by the addition of formic acid without protein, while calcium nitrate remains unaffected without protein.

The two most important factors in the solubilisation of silk fibroins were the concentration of salts and the high temperature. The dissolution process is a physical

phenomenon that requires proper ratios of salt ions and ligands, which is dependent on the nature of the salt counterion.⁴⁸⁹ The concentration of salt solutions required for the solubilisation varied by the type of silk: *Bombyx* silks require a lower amount of salts to be solubilised than *Antheraea* and *Samia* silks. However, *Antheraea* and *Samia* silks required 10M concentration for lithium thiocyanate and 7M for calcium nitrate (tetrahydrate): when using 5M calcium nitrate solution, the solubilisation of *Antheraea* and *Samia* silks were not as effective as using 7M concentration. Thus to examine an archaeological sample of unknown identity, a higher concentration of salts would be necessary. Moreover, when using 7M calcium nitrate solution at 1:1 ratio (liquor ratio to the weight of fibre), *Antheraea* and *Samia* silks showed a small amount of precipitation after centrifugation, especially *A. yamamai* and *S. ricini*. The amount of precipitation decreased at 1:1.5 ratio (silk/salt solution w/v) indicates the solubilisation of fibroins was affected by the salt contents in the solution.

The role of heat in the solubilisation of silk fibroins was also significant, especially in *Antheraea* and *Samia* silks. Firstly, the increase in temperature decreased the time to solubilise the fibroins in all samples. Although *Bombyx* silks were soluble in 7M Ca (NO₃)₂ solution at room temperature, the reaction took place much faster when the temperature went above 60°C. Secondly, the transition from the water-soluble Silk I to insoluble Silk II in *Antheraea* and *Samia* silks are known to be exothermic (see II.3.3).⁴⁹⁰ Thus, the reverse-engineering of Silk II to Silk I would require heat: this was well demonstrated by the fact that solubilisation of *Antheraea* and *Samia* silks barely took place below 100°C. However, none of the previous studies on solubilisation of silk fibroin employed higher temperature above 100°C, and thus the degree of solubilisation was low: thus in this study, higher

⁴⁸⁹ Gordon, 1975

⁴⁹⁰ Tanaka, *et al.*, 2002

temperatures were adopted to accommodate the mass spectrometry of fibroin, which needs the entire protein structure to be denatured. *Antheraea* and *Samia* silks were solubilised in 2-4 hours at 105°C; however, the process took place in less than an hour at 125°C. The optimum condition tested in this study was 125-130°C, which successfully dissolved the fibroins while not boiling off the salt solutions.

Once solubilised, fibroin solutions were desalted by buffer exchange and further concentrated to 100µl (estimatedly 5 µg/µl concentration). There was some degree of aggregation observed from all samples once the heat and salts were removed. It was much less in *Bombyx* silks than in *Antheraea* and *Samia* silks: *A. yamamai* and *S. ricini* showed the most amount of aggregation in accordance with their solubilisation behaviour. The solubilisation and aggregation behaviour of the silk fibroins is related to their structural characteristics: *S. ricini* and *A. yamamai* are reported to have the highest level of repetitive polypeptide sequences, and the stability of their structures makes them less affected by the solvent.⁴⁹¹

The concentrated silk fibroin solutions were examined using SDS-PAGE. Each sample showed corresponding molecular weights to their known experimental or theoretical molecular weights. However, the proteins were not separated to bands but as tailing streams, indicating the proteins now exists in the form of mixtures of polypeptides; this is thought to be the result of denaturation of primary structures by salts and heat. The effect of heat-induced denaturation seems unavoidable at this moment, as the solubilisation of *Antheraea* and *Samia* silks does not take place without the heat. However, several studies are reporting that the heat or microwaved induced denaturation of proteins prior to the in-solution digestion could help to avoid the use of urea or SDS and increasing the digestion

⁴⁹¹ Malay, *et al.*, 2016

efficiency.⁴⁹²

In conclusion, the solubilisation of silk fibroin is possible by using a high concentration of chaotropic salts and heat. The molecular structure of fibroin is denatured into multiple polypeptides after the solubilisation, thus become susceptible to chemicals than in its fibre form. To examine an unknown fibre by protein mass spectrometry, it is recommended to use 7M calcium nitrate as a solvent, and test the sample starting from 60°C: if the solubilisation happens easily, it is highly likely to be the *B. mori* silk; if no reaction is observed, increase the temperature to 125°C and treat for 30 minutes to one hour. In other words, if the sample is silk, the solubilisation should take place under this condition. For a further application of solubilisation of fibroins on the identification of silk, see VII.4.A model protocol for the analysis of archaeological silk.

⁴⁹² Park and Russell, 2001; Russell, *et al.* Ibid.; Pramanik, *et al.*, 2002; Kim, *et al.*, 2006

VII.2 Identification of silks by protein mass spectrometry

The protein sequences of seven different species of modern silk samples were downloaded from the UniProtKB database and analysed by the multiple sequence alignment. The alignment result showed the similarity and uniqueness of each protein sequence. The presence of unique sequence features can be used for a target mass spectrometry, or to develop an antibody for a faster identification: however, it is unknown whether these peptides survive after the years of degradation or harsh solubilisation. The protein sequences were *in-silico* digested using trypsin, chymotrypsin, and 1:1 use of both. The result suggested that the tryptic-chymotryptic digestion would result in the highest protein coverage in all samples.

Silk fibroin solutions and SDS-PAGE gels prepared in Chapter IV were analysed by nanoLC-MS/MS. Although the protein analysis result showed that the various silk proteins were successfully identified, the protein sequence coverages were significantly lower than the theoretical projection. The resulted spectra were interpreted using two different peptide search methods: database query and *de novo* analysis using PEAKS software. The database query was submitted to the MASCOT search engine using MS/MS ion search with UniProtKB and NCBIInr databases. As a result, *Bombyx* silks were identified from both databases, but the *Antheraea* and *Samia* silks were identified only from NCBIInr. This is because most of the silk protein sequences are not experimental, thus not listed in the Swiss-Prot section that MASCOT server uses from UniProtKB. Moreover, there were many cross-identifications due to the multiple sequences that are only slightly different from each other. On the other hand, there was less confusion in *de novo* analysis results, because the manually assembled FASTA database only contained up-to-date protein sequences of

each fibroin. Moreover, the *de novo* method was helpful to exclude disturbances from non-silk protein contaminations. In conclusion, *de novo* analysis using relevant software (PEAKS) was better suited for the identification of silk proteins.

The peptide analysis by PEAKS showed that the reagents used to solubilise silks affected the protein coverage of the samples. The addition of formic acid to 7M calcium nitrate solution promoted the solubilisation process, highly likely by destabilising of the glycine-alanine and polyalanine structure. It is also known that formic acid at high temperature induces proteolysis of aspartic acid in proteins.⁴⁹³ The PEAKS analysis result showed that the distribution of peptides in 7M calcium nitrate-fibroin solution mainly occurred from the less-repetitive region of the protein sequences, which are more susceptible to chemical and mechanical stimulations. However, when formic acid was added, more peptides were occurring from the repetitive region of the proteins. For example, *A. mylitta* fibroin yielded higher protein sequence coverage than other *Antheraea* fibroin samples. This sequence only has the first 507 residues, which is quite similar to other *Antheraea* species. The GRAVY and the instability value indicate that this part of the protein is hydrophilic and unstable, while the complete sequence is supposed to be stable and hydrophobic (Table 6). The result gives an idea of how degraded and aged silks would react to the calcium nitrate solution. In the future, it would be interesting to compare the peptide distributions resulting from the archaeological silk samples and the formic acid-aided modern silk samples.

The effect of various digestion methods on peptide analysis results was investigated. The comparison of PEAKS analysis results from in-gel and in-solution digested samples showed that the difference between the two is not significant; however, if the protein was

⁴⁹³ Crimmins, *et al.*, 2005, Zhang and Basile, 2007

properly resolved into the gels, in-gel digestion showed relatively more even distribution of peptides and higher protein probability than in-solution digestion. The sign of poor digestion process was recognisable by the distribution of peptides concentrated on each terminus of the protein sequence, which leaves repetitive regions unaffected. On the other hand, the use of different protease significantly affected the protein coverage and the distribution of the peptides. In terms of digestion efficiency, chymotryptic peptides yielded more numbers of unique peptides and higher protein coverages than tryptic peptides, except in *S. ricini*. However, chymotryptic peptides and tryptic-chymotryptic peptides did not show much difference despite the projection that the latter would yield higher protein coverage (Table 18). Moreover, tryptic-chymotryptic peptides showed a smaller number of unique peptides than chymotryptic peptide; the yield of unique peptides is important as it is a key feature to avoid the ambiguity of the identification results. Although more experiments must follow until a clear conclusion can be drawn, in this research, the use of chymotrypsin accommodated the peptide analysis of silk fibroins the best. In some samples the results were poor regardless of the enzyme: this must have resulted from either the poor gel resolution or incomplete digestion. This must be further examined once the protein coverages are improved. Also, it must be noted that the protein coverages were no higher than 50% (except in *A. mylitta*), which is much lower than their theoretical projection. This also must be improved before we can conclude which enzyme copes better with the silk fibroins.

In *Antheraea* and *Samia* silks, it was observed that chymotryptic and tryptic-chymotryptic digestions yielded more numbers of peptides resulted by random cleavages, while tryptic peptides mostly yielded tryptic peptides. The non-enzyme specific cleavages were observed in one or both ends of the peptides, after alanine (A), glycine (G), serine(S) and aspartic acid (D): alanine, glycine and serine are the three most abundant amino acids

present in all silk fibroin, thus these cleavages are thought to be the product of solubilisation and denaturation process; while the cleavage of aspartic acid is thought to be initiated by the use of formic acid.⁴⁹⁴ If only time and resource were permitting, it would have been beneficial to further investigate the relationship between the effect of solvent and temperature on the yield of non-enzymatic cleavages. However, hitherto in this research, these non-enzyme specific cleavages did not affect the result of the protein identification, often yielding unique peptides to subject proteins.

⁴⁹⁴ Enzymes: supported reagents (http://www.matrixscience.com/help/enzyme_help.html)

VII.3 Consolidating an understanding of silk proteins for archaeological textile studies

Several interesting new results have been found from the two case studies of archaeological textiles using the fibroin solubilisation technique and protein mass spectrometry. The first case manifested that these techniques are a useful and reliable application to examine the samples cannot be dealt with microscopic methods. The second case has proved that the fibroin solubilisation and protein mass spectrometry can provide significant evidence to long-standing archaeological questions.

The first case studied the textiles from a Gallo-Roman grave of Naintré, dating to the 2-4th century CE. The samples were in very bad condition and mostly disintegrated and lost the physical conformation as fibres. The experimental results proved that the protein mass spectrometry could be a useful tool to identify silk from the samples that cannot be examined by microscopic analysis due to their physical stability. The solubilisation process could screen out the non-silk fibres: out of six samples, one did not react to the calcium nitrate solvent; and all five samples that successfully solubilised were identified as *B. mori* silk by the protein analysis using PEAKS. In other words, if the unknown sample did not react to the solubilisation process, it is unlikely to be silk (Figure 122). The multiple analysis using both *de novo* method and UniProtKB database search was beneficial: the former was apt for focusing on the identification of the silk proteins, and the latter offered the background information that corresponds to the environment that the sample came from. In the comparison of the archaeological and fresh silk samples prepared with an identical method, the archaeological sample showed that more peptide occurred between the residues 3000-5000 of the sequence, whereas fresh silk samples, in general, yielded more peptides

near the C- and N-terminus. It is conjectured that the less hydrophobic region of the protein was already affected by the years of degradation, the surviving—hydrophobic region of the sequence also has become more susceptible to change. In consequence, the distribution of peptides in the archaeological sample slightly differs from that of fresh silk. The conclusion can be drawn once the protein analysis of more archaeological samples of *Antheraea* and *Samia* silks becomes available. In summary, the protein mass spectrometry results of the samples suggest that these textiles from the Gallo-Roman grave were made with silk threads imported from China. Based on the previous study on their weaving structures, these silk textiles could have been something similar to the reproduced silk gauzes made in the Middle East.

The second case studied the textiles from Palmyra, dating to the 1st century BCE to 2nd century CE. It has been reported that there are some specimens of “tussah silks” were found; however, these were never examined due to the limitation of scientific means as the researchers stated.⁴⁹⁵ Five different samples taken from 3 textiles—that were speculated to be tussah silks—were investigated using protein mass spectrometry. All of them were solubilised in 7M Ca(NO₃)₂ solution at a high temperature but the process took about an hour: the solubilisation behaviour implied that these are non-*Bombyx* silks. The peptide analysis results strongly suggested that S8 warp and weft, S 49-T1 are *A. mylitta* silk. S48 and S49 T-2 showed more probability to be *A. pernyi* silk; however, the distribution of peptides suggests that these two are also likely to be *A. mylitta*. The reasons for such counter-identification are 1) incomplete sequence of *A. mylitta* and 2) high homologue between the two species. In my opinion, these are all *A. mylitta* silk; however, for better credibility, S48 and S49 T-2 must be re-evaluated once a full protein sequence of *A. mylitta*

⁴⁹⁵ Schmidt-Colinet, *et al.*, 2000

becomes available. *A. mylitta* and *A. pernyi* are both called Tasar silk, but *pernyi* is of Chinese origin, and *mylitta* is of Indian origin (see II.1.2.1). These two species contain many similar peptides in their protein sequences, and S48 and S49 T-2 seemed to have yield more peptides from the unsequenced part of the protein, which are in a great similarity to *A. pernyi* fibroin.

The finding of this case study is meaningful as it provides the first direct evidence of Palmyrene trade with India, especially the exportation of raw silks and silk textiles from India as recorded in *Periplus Maris Erythraei*. The text discusses a well-established trade route involving the ports of Apologos and Omana in Arabia and Barygaza in India.⁴⁹⁶ Barygaza was the port for the sea trade between India and the West during the Indo-Greek period, to the Red Sea for Egypt and the Persian Gulf for Babylonia. Scholars have been speculated that the silks in Palmyra were imported from the Indian port of Barygaza, via Charax spasinou that was connected to Apologos by Gulf sea lane.⁴⁹⁷ Charax is considered as the most important and principal destination of Palmyrean caravans: Palmyrene inscriptions of the 2nd century also mention the trade caravans to Forat and Charax.⁴⁹⁸ It looks that the cargo imported via Charax to Palmyra was of rather high value: Chinese silk, pearls from the Gulf, agate and carnelian beads from India and incense from India and the Himalayan region.⁴⁹⁹ Thus, either being *A. mylitta* from India or *A. pernyi* from China, these non-domesticated silks manifest that they have travelled through the trade route and was being used for commercial productions.

⁴⁹⁶ Salles, 1993

⁴⁹⁷ Żuchowska, 2013

⁴⁹⁸ Benjamin, 2018

⁴⁹⁹ Żuchowska, 2013

VII.4 A model protocol for the analysis of archaeological silks

Based on the experimental results of this research, here I propose a protocol for the proteomic analysis of archaeological silks. See Figure 122 for a summary of the processes.

Distinguishing silks from non-silk fibres using chaotropic salt solution

1. Clean samples using milliQ water, methanol or ethanol at room temperature if necessary.
2. If the fibre sample is dyed, methanol extraction can be helpful.
 - a. Immerse samples in 1:10 v/v methanol in glass vials or Eppendorf tubes and vortex for 30 seconds.
 - b. Gradually heat to 80°C with the caps opened.
 - c. If necessary, update fresh methanol when the volume decreases, or remove the extract and repeat the process with a fresh load of methanol.
 - d. Extracts can be filtered and vacuum dried for a further dye analysis.
3. If dye extraction is unnecessary, treat samples in methanol at 60°C for 10 min for cleaning, then add solvent (1:1 w/v) and heat to 80°C. Use clean glass vials.
4. If the solubilisation occurs instantly, the fibre is *B. mori* silk.
5. If not, raise the temperature above 125°C, but be careful not to boil off the salts – this will lead to Xanthoprotic reaction and result in protein solution turning yellow.
6. If the subject is silk but of non-bombyx origin, solubilisation of fibres should take place in an hour at an appropriate temperature.
7. Addition of methanol can create vortex thus can aid the solubilisation.
8. If calcium nitrate was used, dilute the fibroin solution with milliQ water to 2-4

volumes of the original solvent volume to facilitate the buffer exchange process. Vortex and sonicate for 30 seconds.

9. If necessary, spin fibroin solutions at maximum speed to remove impurity. Samples such as soil, pulverised fibres and metal threads required this additional step.
10. Fibroin solutions can be stored at 4°C, up to 2 weeks. If set more than 48 hours, reaggregation of proteins will be observed.

Preparation of silk solution for protein mass spectrometry

This process involves desalting, digestion of protein and purification of fibroin solution for any proteomic analysis by mass spectrometry that requires the subject protein to be proteolyzed (MALDI, LC-MS). The biggest obstacles of preparing silk fibroins for mass spectrometry is that silk fibroins do not react to the common buffers used for protein resuspension. Once the salt is removed, the protein starts to reaggregate and once precipitate, it will not be solubilised in urea or ammonium bicarbonate buffer. The reaggregation can be minimised using gradual buffer exchange.

1. Mix 1 volume of fibroin solution and 1 volume of 6M urea-tris/HCl buffer (in MilliQ water) with 10mM TCEP or DTT. Vortex and incubate for 30 minutes at room temperature. If $\text{Ca}(\text{CO}_3)_2$ solution was used as solvent, dilute the sample at least to 5 volumes of the original solvent volume. LiSCN or Ajisawa's solution did not show such problem.
2. Spin the mixture at advised speed/g using MWCO 3K filter for 30 minutes. Update 1 volume of 4M urea buffer and spin for 30 minutes.
3. Update 1 volume of 2M urea buffer and spin for 30 minutes.
4. Update 1 volume of 1M urea buffer with 10mM IAA and incubate for 20 minutes in dark. Spin for 30minutes.

5. Add 1 volume of 100mM Ammonium bicarbonate and spin for 30 minutes.
6. Collect the concentrate and dry in vacuum centrifuge if necessary.
7. Digest the protein regarding the common protocol.
8. Use Empore disk or Ziptip (C18) to clean the peptides.
9. Completely dry the peptides in vacuum centrifuge and elute in appropriate buffer.

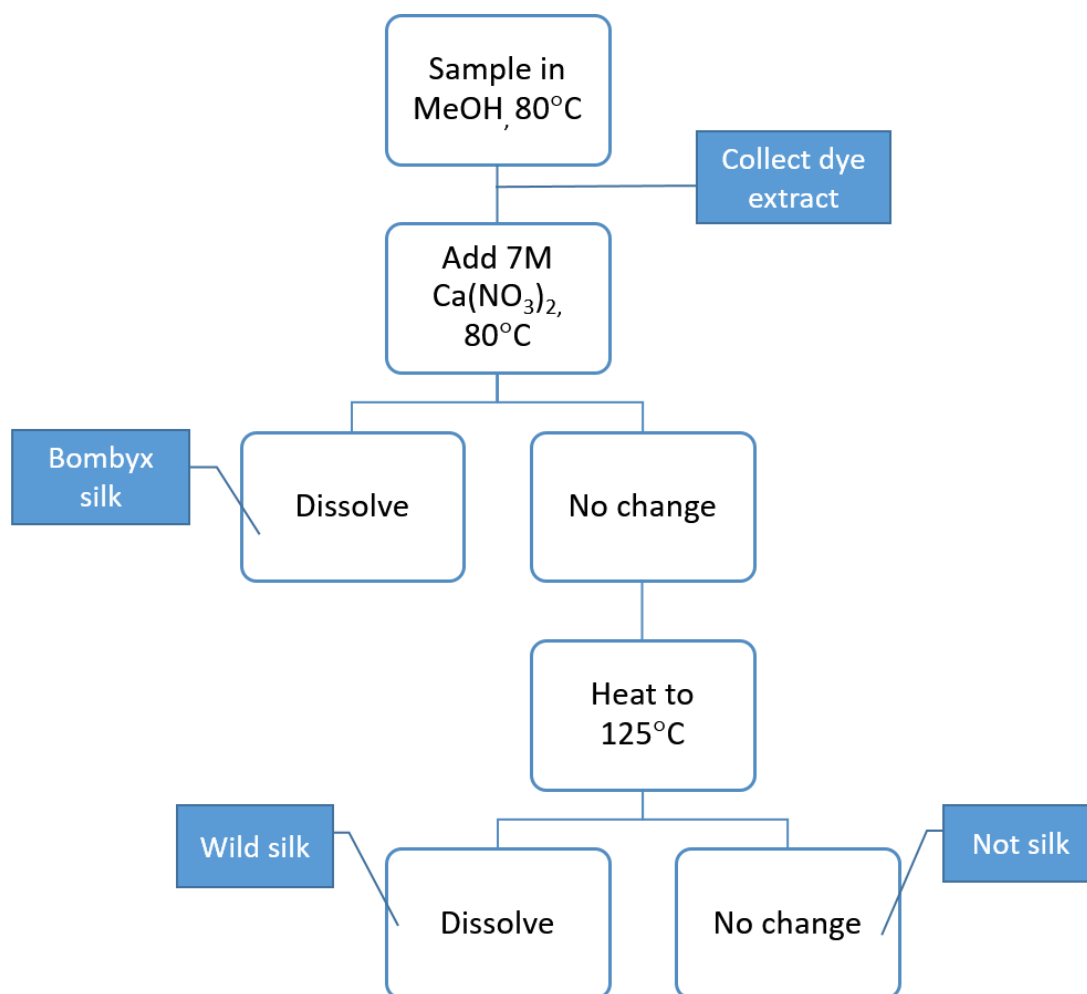


Figure 122. Flowchart for initial silk screening based on the solubilisation behaviour.

VII.5 Further works

Many different adaptations, tests, and experiments have been left for the future due to the time-consuming nature of the experimental research. Also, the limited resources and hardship to access the experiment facility in the first couple of years of the study have limited the time frame of the research. The preliminary results of the experiments were satisfactory to prove the compatibility of the protein mass spectrometry for the analysis of silk. However, there are several lines of studies arising from this research, which should be pursued in order to refine and improve the proposed methods. Further work concerns the additional analysis of particular species, proposals to try different methods and curiosity.

Following from the Section IV.2, the mass spectrometry results in this research were based on the use of calcium nitrate solvent as it affects other fibres less: however, a further investigation of how lithium thiocyanate solvent affects the mass spectrometry result of silk proteins would clarify the argument.

Following from the Section IV.4, the protein gel separation was not as successful as expected; neither were the mass spectrometry results from the in-gel digested samples, despite the method generally leading to a high protein coverage. The main reason was the poor resolution of the subjects due to the complexity of the protein. The adoption of the gel or filter-aided digestion method, instead of using electrophoresis could improve the results. Alternatively, 2-dimensional separation could be used to achieve a higher level of resolution and higher loading capacity: in this technique, proteins are first separated by their isoelectric point (pI), and then by the molecular mass.

Following from Section V.4, the experimental protein coverages of each sample from the tryptic-chymotryptic digestion were not as high as their theoretical projection. The

possible explanations are: 1) some part of the proteins were lost during the solubilisation, due to the low molecular size below the cut-off limit of the desalting membrane, and thus the use of two enzymes have created too many small cleavages, 2) even after the solubilisation, the resulted molecules were still too stable to react with the protease, and could not be digested, thus were not able to be analysed by the mass spectrometry. However, the contemplation about this matter should be followed by the improvement of the protein coverage. Moreover, it would be beneficial to investigate the effect of enzymes, by comparing when they are used as a mixture and used separately, as well as in different orders. It would be necessary to refine the current digestion protocol.

In addition, to expand the scope of this research, the amino acid sequencing of various species to broaden the silk protein database should follow. Firstly, it is necessary to improve the reliability of the analysis, especially for those that have partial sequences. *B. mandarina* always suggested *B. mori* heavy chain instead of *B. mandarina* heavy chain, implying that the unsequenced area of *B. mandarina* heavy chain from the 178th residue is almost identical to that of *B. mori*. In the same manner, the protein analysis result of *A. mylitta* fibroin almost always showed a high probability for *A. yamamai* and *A. pernyi*, suggesting that the unsequenced part of this protein from the 507th residue would be similar to that of *A. yamamai* and *A. pernyi*; such ambiguity can be avoided by acknowledging the analysis of the unique features and the similarity between the sequences. Secondly, the availability of more protein sequences will enable the analysis of the controversial silk findings such as Coan silk: neither the protein nor the gene sequence data of *Pachypasa otus* and *Saturnia pyri* fibroin, the two candidates of Coan silk moth, are yet available. Now that I have a method to solubilise their silks, the amino acid sequencing can be done easily.

Once the full sequence of the silk proteins becomes available, their isoelectric

values (pI) will be known. This will enable the identification of the proteins after the 2-dimensional gel separation, even before the mass spectrometry analysis. However, despite their benefits, the gel-based protein separations are labour intensive and have a limited dynamic range.⁵⁰⁰ Based on both the theoretical foundation and the experimental results, the 2-dimensional LC-MS/MS technique seems to be more suitable for the protein analysis of silk: in this method, a sample is first separated sequentially using two different liquid chromatography columns and then analysed by two mass spectrometers in series. The combination of ion-exchange chromatography followed by reversed-phase chromatography is the most common combination. However, in the case of silk proteins, the combination of size-exclusion chromatography and reversed phase chromatography capillary electrophoresis could be useful. Unfortunately, the 2D LC-MS/MS setting was not available in our facility due to the compatibility issue, but I strongly hope to be able to use this method in the future.

Concerning the analysis of archaeological textiles, I have a line of research planned to collaborate with Dr Christina Margariti from the Hellenic Ministry of Culture, to analyse the ancient Greek textile samples from Kerameikos and Akrotiri that are thought to be the remnants of the Coan silk, in comparison with modern day *Saturnia pyri* and *Pachypasa otus* cocoon silks. I hope to have more chances to examine samples such as Xiyincun cocoon and be able to help to trace the development of material culture through the movement of textile materials in the ancient world.

⁵⁰⁰ Nagele, *et al.*, 2004; Delahunty and Yates, 2005

Appendix A: Light Stable Isotope analysis of preserved mulberry leaves

Background

In theory, it is likely to be able to provenance *B. Mori* silks with light stable isotopes. Knaller and Ströbele argued that the result of their stable isotope analysis indicate that it is possible to provenance the regions where the silk fibres were produced.⁵⁰¹ However, no specific data has been published as of 2015. The quality of mulberry leaves is associated with the health of silkworm larvae and the weight of the cocoons. The quality of mulberry leaves, in other words the protein and amino acids contents, often directly influences the quality of cocoons.⁵⁰² It also varies with the type of leaves used for rearing the spring or the summer-autumn silkworms. In any case, the mulberry leaves used in spring rearing yield better cocoon quality than those used in summer-autumn rearing, which implies the effect of temperature on the nutritional quality of the leaves.⁵⁰³ The mulberry plants are cultivated both in tropical and temperate regions, and there are three important factors that determine the quality of leaves, viz. good variety of mulberry, appropriate agronomic practices and plant protection measures.⁵⁰⁴

To sum up, there are many variables for the nutritional composition of mulberry leaves that eventually influence on the chemical composition of silk. Thus keeping in view of the above mentioned factors, an experiment was conducted to investigate the carbon and nitrogen stable isotope distribution of the mulberry leaves from various locations. Although the exact rate of fractionation is yet unknown, it is a good place to start with to examine the applicability of the stable isotope technique for the provenance of silks.

⁵⁰¹ Knaller and Ströbele, 2014, *ibid*.

⁵⁰² Aruga, 1994, Mechii and Katagiri, 1991 and Ravikumar, 1988

⁵⁰³ Aruga, 1994,

⁵⁰⁴ Qader, *et al.*, 1992, cited by Kumar, *et al.*, 2013

Experimental

Material

Mulberry leaf samples were secured from historical plant samples preserved in collections in the University of Oxford Herbaria (Plant Science Department) with the help of the curator, Dr. Stephen Harris. The criteria for sample selections were 1) location 2) date 3) method of preservation. A total of 18 samples collected from Asia (8), UK (5) and rest of Europe (5) with date record, and preferably preserved without glue (stitched or taped to the background) were chosen as the samples (Table 35). Known origins of the samples and their geographical and climate data are presented in Figure 123. However, the climatic information of Japan is that of Shizuoka prefecture in general—there was no specific data available for Mt. Amagi (天城山, Izumo, Japan). The climatic classification in colours follows the Köppen–Geiger climate classification system.

Sample preparation

The botanical specimens were preserved in the late 19th to early 20th century, thus in many cases mercuric chloride (HgCl₂) had been applied as a preservative. The samples were chosen with care to avoid heavily preserved ones. However, organic additives likely to remain were targeted by a solvent wash, comprising the following steps: acetone (45 min, 45 °C), methanol (45 min, 45 °C) and chloroform (45 min, 20 °C). This procedure was applied to every sample whether contamination was suspected or not.⁵⁰⁵

Solvent cleaned samples were kept in freeze dryer overnight, then weighed into tin capsules contains approximately 1 mg, and 6 capsules were prepared for each sample. Samples were combusted in an elemental analyser coupled to a mass spectrometer (Carlo Erba EA1108),

⁵⁰⁵ Dee, *et al.*, 2010.

which measured the $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ value. The obtained data, 6 sets of each, were consistent per category and did not show much variance, thus the average values were engaged for the analysis.



Location	Detail	Köppen-Geiger classification	Latitude(N)	Longitude(E)	Mean Ann. Temp(°C)	Mean Precipitation /Month (mm)
Burma	Toungoo District	Tropical wet and dry savanna: AW	18.56	96.26	27	197.5
India	Jaunsar-New Delhi	Humid Continental: Csa	30.75	77.83	25.1	65.83
India	Dehradun		30.33	78.06	21.25	169
Japan	Mt. Amagi, Shizuoka	Humid subtropical : Cfa	34.51	139	16.08	193.92
Italy	Modena	Mediterranean: Csa	44.39	10.56	13.39	67.42
France	Canet Plage		42.42	3	17.8	38.9
France	Paris	Marine west coast : Cfb	48.5	2.2	20	54.3
Poland	Breslau	Cold and temperate : Dfb	51.6	17.2	9.8	47.2
UK	Greater London	Maritime temperate: Cfb	51.3	-0.2	12.5	49

Figure 123. Geographical distribution of the samples and their climatic data

Table 35. Descriptions of the preserved leaf samples

No.	Collection	Col. no.	Name	Location	Year	Date	Method	Note
A1	19771	94	Morus A.	Burma			stitched	
A2		343	Morus Indica Linn.	Burma	1914	April	taped	Toungoo district. Cultivated
A3		248	Morus Laevigata W.	Burma	1914	Mar 31	taped	Toungoo district.
A4	19852	96	Morus A.	India			stitched/taped	
A5		234	Morus Serrata R.	India	1891	May	glued/taped	Flora of Jaunsar, N.-W. Provinces/New Delhi
A6			Morus A. L.	India	1893	Mar	stitched	Dehradun Herbarium
A7			Morus. A. L.	India	1895	Feb	stitched	Dehradun Herbarium
A8	133597	TOFO 28	Morus bombycis K.	Japan	1974	Jun 14	stitched/taped	Honshu/Shizuoka: Mt. Amagi
E1			Morus A.L. (ramus)	Italy	1916	Jan	glued	Modena: Casinalbo, bulta. 2 pgs
E2	151930	A21/79	Morus Kagayamae K.	France	1978	Sep	stitched/taped	Canet Plage, Dept. Roussillon
E3	153445	A22/79	Morus Kagayamae K.	France	1979	Jun 12	stitched/taped	Paris, Jardin du Luxemburg, tree about 50 yrs old
E4	16041	1394	Morus A.L.	Poland	1905	Jun 2/Jul 22	stitched	"v. constantinopolitana Loud. Heimisch in China"
E5	16645		Morus A.L.	Poland	1905	Jun 2/Jul 22	stitched	Same as above. Marked as from Breslau
U1	12629	16	Morus A.	UK	1893		taped	Trees of GB and Ireland
U2	12486	5.85	Morus A. Tokwa	UK	1911	Jun 21	stitched	
U3	29094		Morus A. Latifolia	UK	1911	Jun 21	taped	
U4	12628		Morus. A. Pumila	UK	1911	Jun 21	taped	
U5	29092	129-90	Morus. A. Pendula	UK	1911	Jun 21	taped	Royal Botanic Gardens, Kew Richmond Surrey

Result and discussion

¹³C and ¹⁵N distribution

Carbon and nitrogen distribution of the samples are presented in Table 36 and Figure 124; each number of samples shows distinctive carbon and nitrogen values. Regardless of the location and cultivar, $\delta^{13}\text{C}$ values were ranged from -27.18 to -23. 55, while $\delta^{15}\text{N}$ values were largely dispersed: this implies that $\delta^{15}\text{N}$ values of each sample could be a marker to distinguish each other. First off, the difference of cultivar seems to be one of the factors that affect this variance. For example, while *Morus A.* from Toungoo, Burma (A1 and A4) showed 8.71 and 10.19 ‰ of $\delta^{15}\text{N}$ values, respectively, *Morus Indica L.* (A2) showed -0.28 ‰ of $\delta^{15}\text{N}$ value.

The difference in locational origins also seems to affect the result. Figure 125 and Figure 126 are the result of *Morus A.* and *Morus Kagayamae K.* samples only, respectively. The samples showed different $\delta^{15}\text{N}$ values among the same species group. In Figure 127, a group of samples from the locations within the U.K. of all different cultivars,⁵⁰⁶ also showed $\delta^{15}\text{N}$ values different from each other's. However, the same cultivar samples from close locations—A6 and A7 (*Morus A.* from Dehradun, India), E2 and E3 (*Morus Kagayama K.*, France), and E4 and E5 (*Morus A.* from Breslau, Poland) — showed a little difference in values but stayed close compared to the same cultivar samples from remote locations.

To better observe the relationship between $\delta^{15}\text{N}$ values of samples and environmental factors, the data were matched to the climatic data from each location in Figure 128. The yearly mean temperature (°C) considered to be the most reasonable choice to compare, where the samples are not from a controlled environment thus lacks information such as Growing-degree

⁵⁰⁶ Out of 5 UK samples, U1-U4 lacks information of origin. However, all these belonged to a same collection and thus assumed to be sourced from a same botanical garden with U5.

days(GDD) and the time of crop. Although the result did not exactly match with the Köppen–Geiger’s classification in Figure 123, there were some trends observed: the samples from similar climatic categories stayed together within a range. The blue zone stands for the locations of yearly mean temperature under 15°C, the green zone stands for the a little more temperate, but mountainous environment (Mt. Amagi, Elevation 1,460m), the orange zone stands for the much warmer environment (~22°C), and the red zone stands for the tropical environment (25°C~).

Table 36. Carbon and Nitrogen distribution of the samples

No.	Name	Location	W/V	N	¹⁵ N	C	¹³ C	C/N Molar(%)
A1	Morus A.	Burma	1.06	46.58	10.19	456.80	-24.18	11.49
A2	Morus Indica Linn.	Burma	1.05	52.30	-0.28	398.54	-25.77	8.91
A3	Morus Laevigata W.	Burma	1.09	36.56	-1.11	443.33	-27.18	14.16
A4	Morus A.	India	1.07	38.11	8.71	450.70	-25.50	14.19
A5	Morus Serrata R.	India	1.04	42.99	1.59	415.03	-25.74	11.36
A6	Morus A. L.	India	1.06	37.96	3.65	432.47	-24.19	13.46
A7	Morus. A. L.	India	1.07	61.69	6.73	456.28	-23.57	8.64
A8	Morus bombycis K.	Japan	1.09	34.55	-2.77	433.58	-25.33	14.71
E1	Morus A.L.(ramus)	Italy	1.08	41.78	2.66	448.06	-25.53	12.62
E2	Morus Kagayamae K.	France	1.17	58.64	5.37	504.23	-24.34	10.11
E3	Morus Kagayamae K.	France	1.10	35.40	4.74	467.93	-24.74	15.50
E4	Morus A.L.	Poland	1.11	27.17	2.09	429.21	-23.55	19.55
E5	Morus A.L.	Poland	1.05	53.95	4.35	385.55	-22.21	8.61
U1	Morus A.	UK	1.08	37.46	2.02	442.51	-24.84	13.94
U2	Morus A. Tokwa	UK	1.10	42.23	5.06	433.92	-25.00	12.02
U3	Morus A. Latifolia	UK	1.09	34.21	3.15	444.98	-24.14	15.19
U4	Morus. A. Pumila	UK	1.10	46.07	5.97	455.76	-25.42	11.62
U5	Morus. A. Pendula	UK	1.07	43.79	1.16	439.17	-24.43	11.94

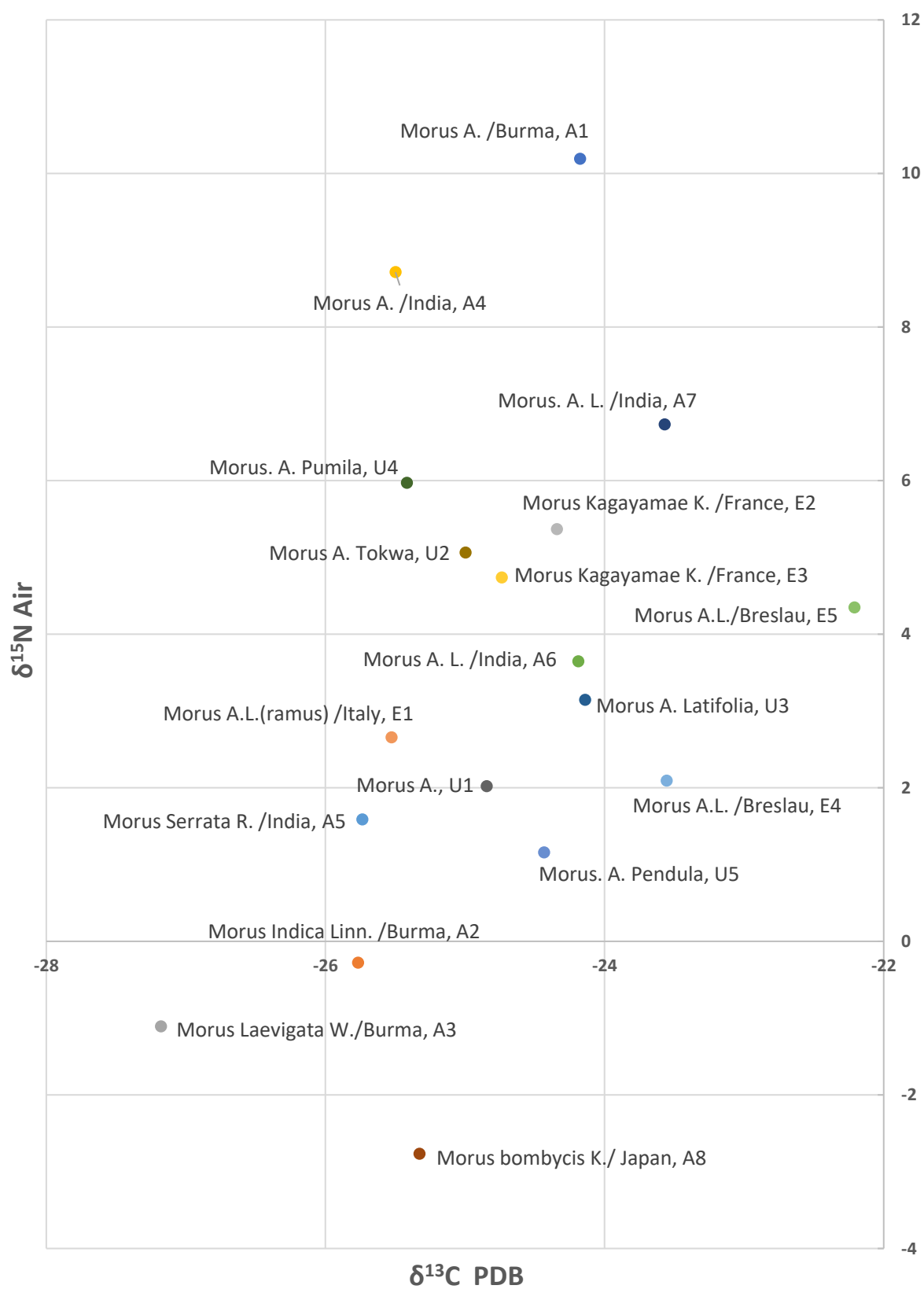


Figure 124. Distribution of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ value of the mulberry samples

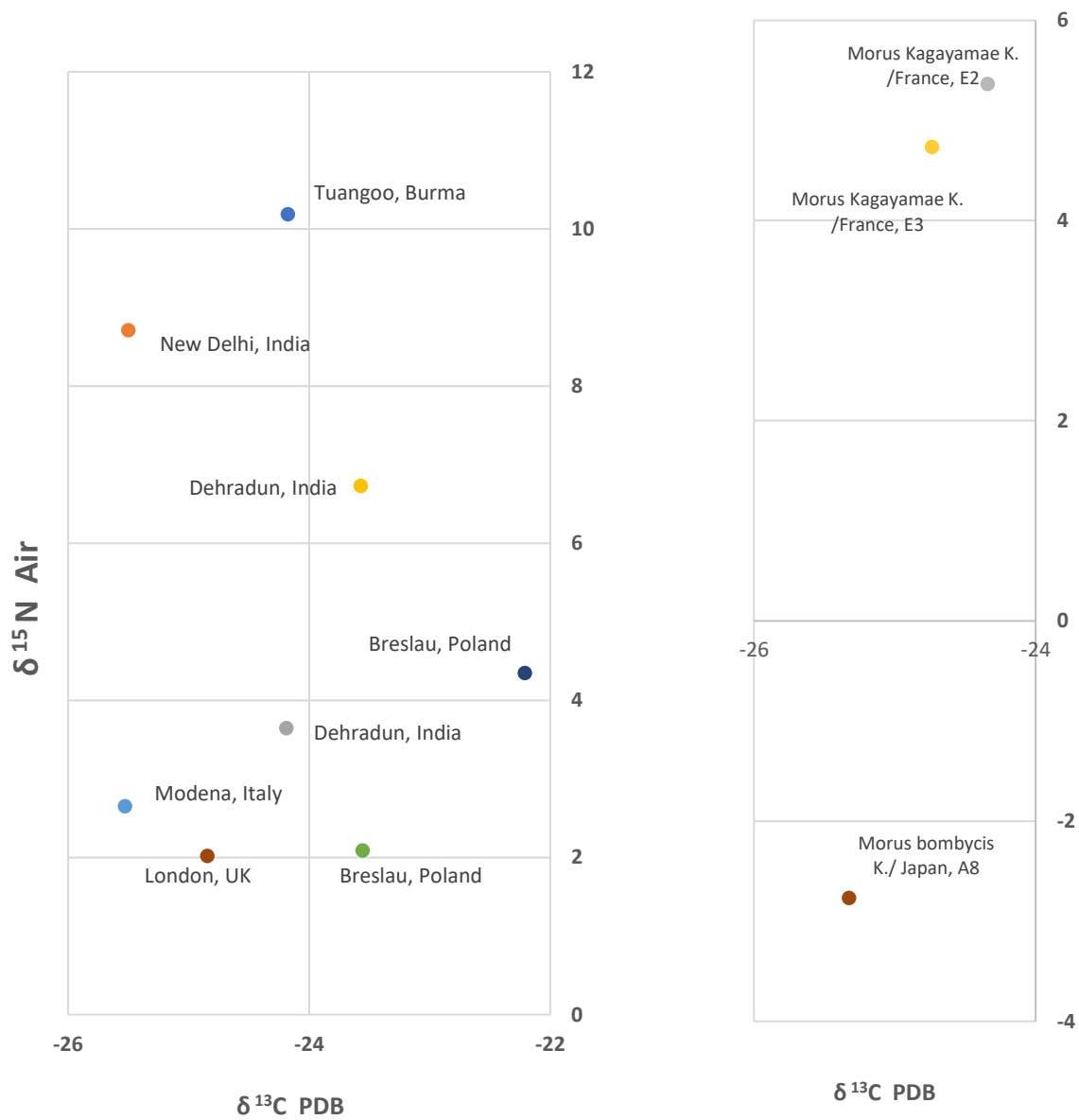


Figure 125. Distribution of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ value of *Morus A.* samples from various locations (Left)

Figure 126. Distribution of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ value of *Morus Kagayama K.* samples from various locations (Right)



Figure 127. Distribution of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ value of Mulberry varieties from UK

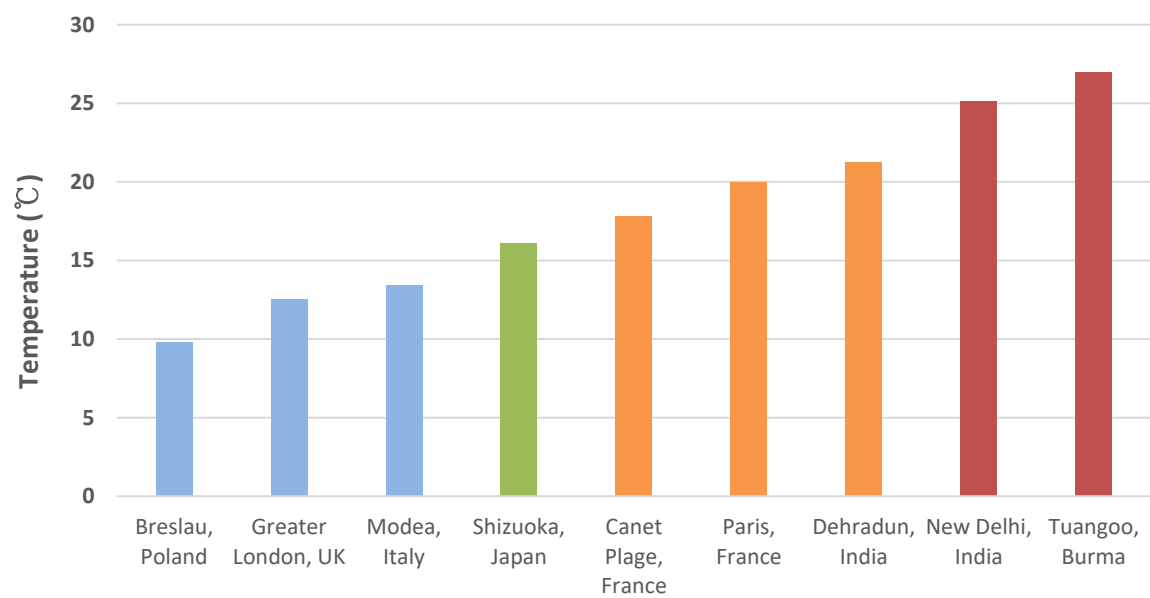
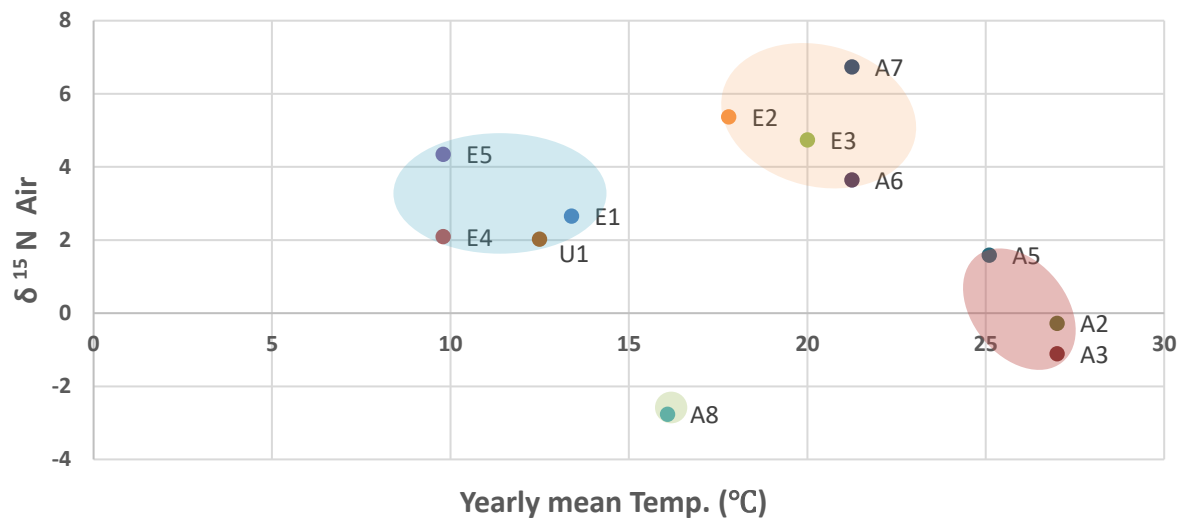


Figure 128. Distribution of $\delta^{15}\text{N}$ value of samples by yearly mean temperature (°C)

Conclusion

This experiment did not yield a useful dataset to study the relationship between the mulberry leaves and the distribution of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values. There are several factors that affect the $\delta^{15}\text{N}$ values: location (latitude), cultivar and temperature. Moreover, mulberry trees are highly likely to be fertilized for a high yield, thus the quality of soil also would affect the divergence of data. However, in comparison with other factors, the temperature of its grown origin is more closely related to the distribution of $\delta^{15}\text{N}$ values of the samples.

The performance of any mulberry variety with respect to leaf quality and cocoon production varies with agro-climatic conditions and cultivation practices. Physical and chemical properties of leaves directly influence the amount of leaf consumption and digestion.⁵⁰⁷ The larval growth of silkworm is influenced by the nutritional level of different cultivars of mulberry which ultimately reflect the economic traits namely larval weight, cocoon and shell weight, effective rate of rearing and so on.⁵⁰⁸ The quality of mulberry leaves and environmental conditions influence the dietary efficiency and growth of silkworm.⁵⁰⁹

Thus, it seems that the study of hydrogen and oxygen would be more appropriate for provenance mulberry leaves or mulberry silks from the leaves. First the water from the region must be collected, then mulberry leaves grew on that water, and *Bombyx mori* silkworm grew on those leaves must be studied: also, the fractionation ratio between the gland silk and the cocoon silk must be investigated.

⁵⁰⁷ Ghosh, *et al.*, 2000

⁵⁰⁸ Ibid., Zannoon, *et al.*, 2012

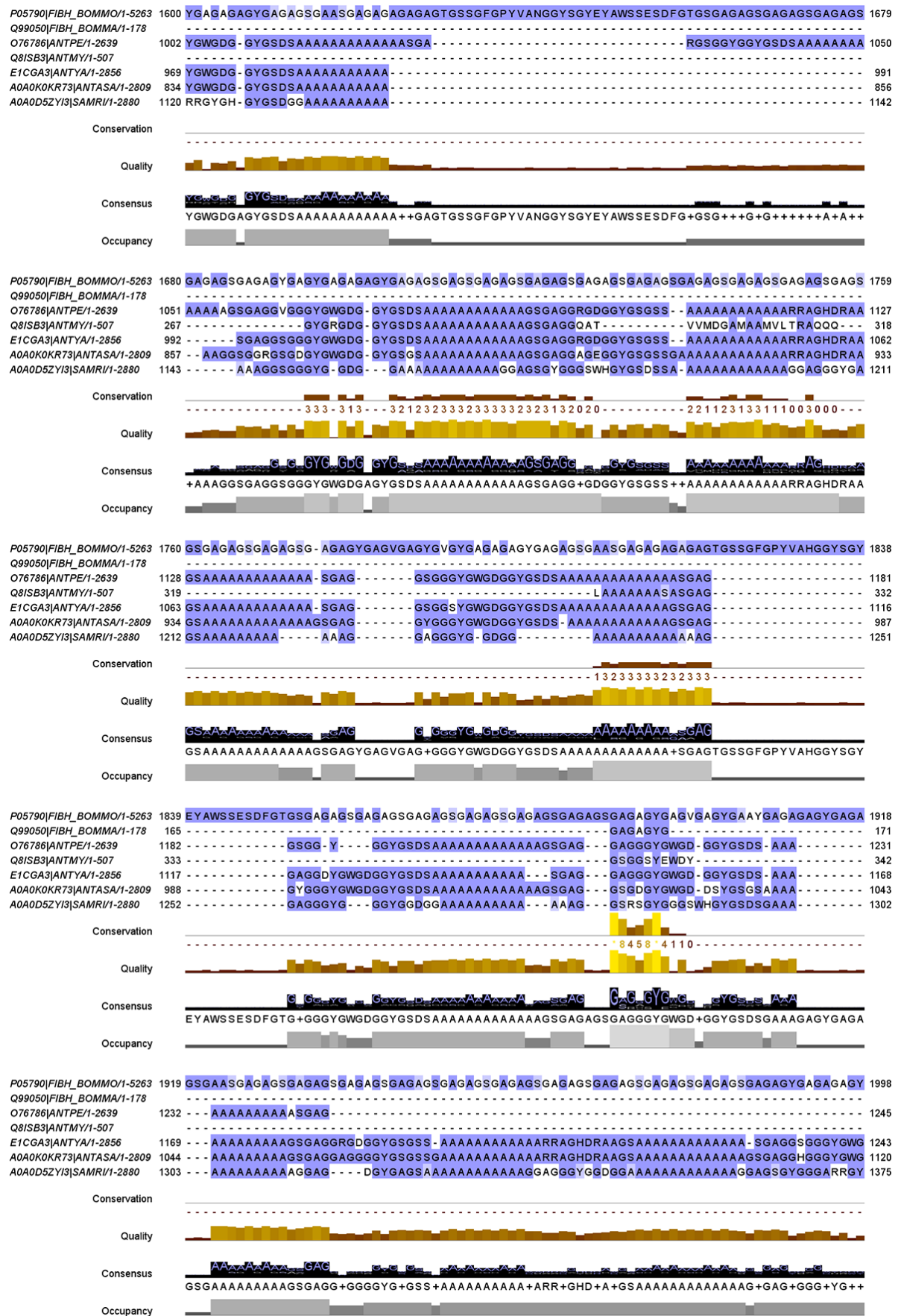
⁵⁰⁹ Rahmathulla, 2012

Appendix B:

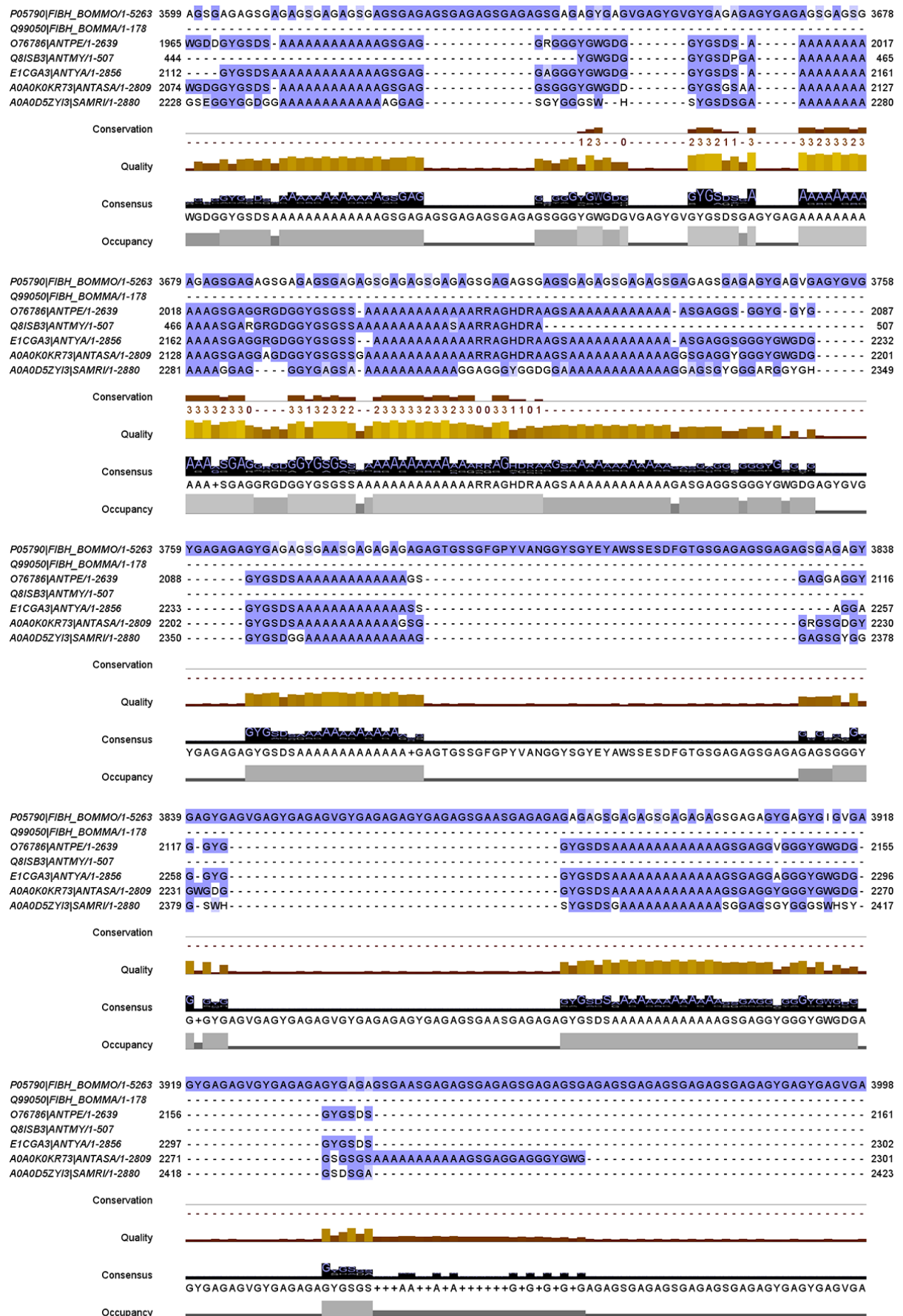
Sequence Alignments











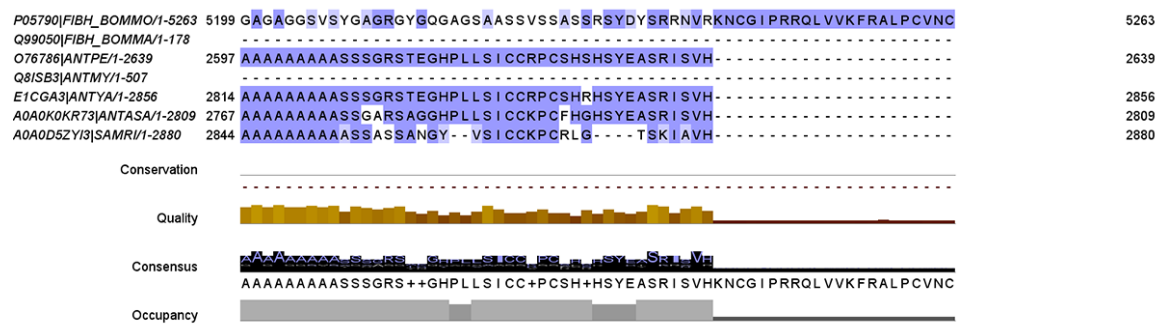


Figure 129. Sequence alignment of the *Bombyx*, *Antheraea* and *Samia* silks by MUSCLE algorithm (v.3.8.31) using Jalview (2.10.3b1). The colour scheme is after Blosum62: if a residue matches the consensus sequence residue at that position it is coloured dark blue. If it does not match the consensus residue but the 2 residues have a positive Blosum62 score, it is coloured light blue.

Appendix C:

Data Tables

Table 37. Summary of MASCOT search results (Database: NCBIInr)

Accession	Mem.	Protein description	Score	Mass	Matches (sig.)		Sequences (sig.)	
B. mori (7M Calcium Nitrate)								
gi 119850	1	RecName: Full=Fibroin light chain; Precursor [Bombyx mori]	255	27823	11	6	6	5
gi 13383193	1	fibroin L-chain [Bombyx mandarina]	255	27793	11	6	6	5
gi 112984494	1	fibroin light chain precursor [Bombyx mori]	255	27908	10	6	5	5
gi 24637964		fibroin light chain [Bombyx mori]	224	27806	8	5	5	4
gi 304443558		FIB-L protein [Bombyx mori]	220	27824	9	5	5	4
gi 13195043		fibroin light chain [Bombyx mori]	142	24832	8	3	4	3
gi 126024018		fibroin light chain [Bombyx mori]	78	5196	2	1	1	1
gi 452378		Nd-s mutant fibroin light chain [Bombyx mori]	78	30889	3	1	2	1
gi 452392		Nd-sD mutant fibroin light chain [Bombyx mori]	78	30863	3	1	2	1
gi 449142588		fibroin, partial [Bombyx mandarina]	78	6802	3	1	2	1
gi 58416325		fibroin light chain [Bombyx mori]	78	5434	2	1	1	1
gi 161898273		fibroin light chain [Bombyx mori]	74	2314	3	1	1	1
gi 13383195	1	fibroin P25 [Bombyx mandarina]	93	25701	10	3	7	2
gi 225543283	1	fibrohexamerin precursor [Bombyx mori]	93	25664	9	3	6	2
gi 164448672	1	fibroin heavy chain precursor [Bombyx mori]	91	391653	2	2	2	2
gi 3599678		fibroin H-chain [Bombyx mori]	77	8141	1	1	1	1
gi 765323		silk fibroin heavy chain, partial [Bombyx mori]	77	49526	1	1	1	1
gi 164506963		fibroin heavy chain Fib-h [Bombyx mori]	54	15908	1	1	1	1
gi 5869		putative reading frame (fibroin) [Bombyx mori]	54	17595	1	1	1	1
gi 381352997		Chain A, N-Terminal Domain Of Bombyx mori Fibroin Mediates The Assembly Of Silk In Response To Ph Decrease	54	14891	1	1	1	1
B. mori (6N HCl)								
gi 119850	1	RecName: Full=Fibroin light chain; Precursor [Bombyx mori]	394	27823	24	15	7	7
gi 13383193	1	fibroin L-chain [Bombyx mandarina]	394	27793	24	15	7	7
gi 24637964		fibroin light chain [Bombyx mori]	352	27806	20	12	6	6
gi 304443558		FIB-L protein [Bombyx mori]	318	27824	22	13	6	6
gi 13195043		fibroin light chain [Bombyx mori]	222	24832	13	9	5	5
gi 452378		Nd-s mutant fibroin light chain [Bombyx mori]	141	30889	7	4	3	3
gi 449142588		fibroin, partial [Bombyx mandarina]	136	6802	4	3	2	2

gi 58416325		fibroin light chain [<i>Bombyx mori</i>]	125	5434	2	2	1	1
gi 161898273		fibroin light chain [<i>Bombyx mori</i>]	88	2314	4	3	1	1
gi 164448672	1	fibroin heavy chain precursor [<i>Bombyx mori</i>]	162	391653	4	4	2	2
gi 3599678		fibroin H-chain [<i>Bombyx mori</i>]	111	8141	3	3	1	1
gi 765323		silk fibroin heavy chain, partial [<i>Bombyx mori</i>]	111	49526	3	3	1	1
gi 164506963		fibroin heavy chain Fib-h [<i>Bombyx mori</i>]	93	15908	1	1	1	1
gi 5869		putative reading frame (fibroin) [<i>Bombyx mori</i>]	93	17595	1	1	1	1
gi 381352997		Chain A, N-Terminal Domain Of <i>Bombyx mori</i> Fibroin Mediates The93	93	14891	1	1	1	1
		Assembly Of Silk In Response To Ph Decrease						
gi 13383195	1	fibroin P25 [<i>Bombyx mandarina</i>]	76	25701	3	3	2	2
gi 225543283		fibrohexamerin precursor [<i>Bombyx mori</i>]	66	25664	2	2	1	1
gi 112983118	1	seroin 1 precursor [<i>Bombyx mori</i>]	53	11875	3	1	2	1
B. mandarina								
gi 119850	1	RecName: Full=Fibroin light chain; Precursor [<i>Bombyx mori</i>]	411	27823	22	15	7	7
gi 13383193	1	fibroin L-chain [<i>Bombyx mandarina</i>]	411	27793	22	15	7	7
gi 112984494		fibroin light chain precursor [<i>Bombyx mori</i>]	400	27908	20	14	6	6
gi 24637964		fibroin light chain [<i>Bombyx mori</i>]	348	27806	19	12	6	6
gi 304443558		FIB-L protein [<i>Bombyx mori</i>]	338	27824	20	13	6	6
gi 13195043		fibroin light chain [<i>Bombyx mori</i>]	285	24832	11	9	5	5
gi 452378		Nd-s mutant fibroin light chain [<i>Bombyx mori</i>]	135	30889	6	4	3	3
gi 452392		Nd-sD mutant fibroin light chain [<i>Bombyx mori</i>]	135	30863	6	4	3	3
gi 449142588		fibroin, partial [<i>Bombyx mandarina</i>]	126	6802	4	3	2	2
gi 126024018		fibroin light chain [<i>Bombyx mori</i>]	118	5196	2	2	1	1
gi 58416325		fibroin light chain [<i>Bombyx mori</i>]	118	5434	2	2	1	1
gi 161898273		fibroin light chain [<i>Bombyx mori</i>]	108	2314	3	3	1	1
gi 164448672	1	fibroin heavy chain precursor [<i>Bombyx mori</i>]	213	391653	5	5	4	4
gi 3599678		fibroin H-chain [<i>Bombyx mori</i>]	145	8141	1	1	1	1
gi 765323		silk fibroin heavy chain, partial [<i>Bombyx mori</i>]	145	49526	1	1	1	1
gi 381352997		Chain A, N-Terminal Domain Of <i>Bombyx mori</i> Fibroin Mediates The114	114	14891	4	4	3	3
		Assembly Of Silk In Response To Ph Decrease						
gi 5869		putative reading frame (fibroin) [<i>Bombyx mori</i>]	114	17595	4	4	3	3
gi 164506963		fibroin heavy chain Fib-h [<i>Bombyx mori</i>]	101	15908	3	3	2	2
gi 5867		fibroin [<i>Bombyx mandarina</i>]	50	18368	1	1	1	1
gi 9087164		RecName: Full=Fibroin heavy chain; Short=Fib-H; AltName: Full=H-50 fibroin; Flags: Precursor, partial [<i>Bombyx mandarina</i>]	50	18429	1	1	1	1

gi 13383195	1	fibroin P25 [<i>Bombyx mandarina</i>]	84	25701	6	4	4	3
gi 225543283		fibrohexamerin precursor [<i>Bombyx mori</i>]	74	25664	4	3	3	2
A. pernyi								
gi 3450883	1	fibroin [<i>Antheraea pernyi</i>]	256	216219	20	16	5	5
gi 15077406		fibroin [<i>Antheraea yamamai</i>]	124	218889	8	7	3	3
gi 301017132		fibroin [<i>Antheraea yamamai</i>]	124	234724	8	7	3	3
gi 23476394		fibroin [<i>Antheraea mylitta</i>]	54	45489	6	5	1	1
gi 110347853	1	serpin-like protein [<i>Antheraea mylitta</i>]	142	16373	2	2	1	1
gi 119850	1	RecName: Full=Fibroin light chain; Precursor [<i>Bombyx mori</i>]	108	27823	5	3	5	3
gi 13383193	1	fibroin L-chain [<i>Bombyx mandarina</i>]	108	27793	5	3	5	3
gi 112984494	1	fibroin light chain precursor [<i>Bombyx mori</i>]	108	27908	5	3	5	3
gi 24637964		fibroin light chain [<i>Bombyx mori</i>]	99	27806	4	2	4	2
gi 304443558		FIB-L protein [<i>Bombyx mori</i>]	82	27824	4	2	4	2
gi 13195043		fibroin light chain [<i>Bombyx mori</i>]	72	24832	3	2	3	2
gi 126024018		fibroin light chain [<i>Bombyx mori</i>]	62	5196	1	1	1	1
gi 452378		Nd-s mutant fibroin light chain [<i>Bombyx mori</i>]	62	30889	1	1	1	1
gi 452392		Nd-sD mutant fibroin light chain [<i>Bombyx mori</i>]	62	30863	1	1	1	1
gi 58416325		fibroin light chain [<i>Bombyx mori</i>]	62	5434	1	1	1	1
gi 449142588		fibroin, partial [<i>Bombyx mandarina</i>]	62	6802	1	1	1	1
gi 161898273		fibroin light chain [<i>Bombyx mori</i>]	49	2314	1	1	1	1
A. mylitta								
gi 3450883	1	fibroin [<i>Antheraea pernyi</i>]	82	216219	2	2	2	2
gi 15077406		fibroin [<i>Antheraea yamamai</i>]	72	218889	1	1	1	1
gi 301017132		fibroin [<i>Antheraea yamamai</i>]	72	234724	1	1	1	1
gi 110347853	1	serpin-like protein [<i>Antheraea mylitta</i>]	58	16373	1	1	1	1
gi 512931106	1	PREDICTED: myb-like protein X-like [<i>Bombyx mori</i>]	56	169144	1	1	1	1
A. yamamai								
gi 3450883	1	fibroin [<i>Antheraea pernyi</i>]	103	216219	7	4	3	3
gi 15077406	1	fibroin [<i>Antheraea yamamai</i>]	103	218889	7	4	3	3
gi 301017132	1	fibroin [<i>Antheraea yamamai</i>]	103	234724	7	4	3	3
gi 23476394		fibroin [<i>Antheraea mylitta</i>]	46	45489	5	2	1	1
A. assamensis (no result)								
S. ricini								

gi 3450883	1	fibroin [<i>Antheraea pernyi</i>]	104	216219	2	2	2	2
gi 15077406	1	fibroin [<i>Antheraea yamamai</i>]	104	218889	2	2	2	2
gi 15077406	1	fibroin [<i>Antheraea yamamai</i>]	104	218889	2	2	2	2
gi 301017132	1	fibroin [<i>Antheraea yamamai</i>]	104	234724	2	2	2	2
gi 301017132	1	fibroin [<i>Antheraea yamamai</i>]	104	234724	2	2	2	2
gi 13991888		fibroin [<i>Antheraea pernyi</i>]	61	37177	1	1	1	1
gi 1507659		<i>Antheraea pernyi</i> fibroin	61	36007	1	1	1	1

Table 38. Unique peptides to B. mori heavy chain

CN	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
R.DASGAVIEEQITTK.K	52.67	1460.736	14	731.3748	2	29.52	49	62	
R.DASGAVIEEQITTKK.M	32.33	1588.831	15	795.4235	2	25.65	49	63	
R.DASGAVIEEQIT.T	22.24	1231.593	12	616.8041	2	35.32	49	60	
D.ASGAVIEEQITTK.K	35.45	1345.709	13	673.8619	2	25.38	50	62	
A.SGAVIEEQITTK.K	33.92	1274.672	12	638.3426	2	25.46	51	62	
S.GAVIEEQITTK.K	22.8	1187.64	11	594.8218	2	41.94	52	62	
A.VIEEQITTK.K	23.16	1059.581	9	530.798	2	20.39	54	62	
G.AGSGAGAGSGAGAGSGAGAGSG.A	24.32	1490.635	22	746.3241	2	16.12	421	442	
G.AGSGAGAGSGAGAGSGAGAGSGAG.A	23.59	1618.693	24	810.3533	2	16.68	421	444	
G.AGSGAGAGSGAGAGSGAGAGSGAGA.G	23.06	1289.56	19	645.7871	2	16.18	421	439	
A.GSGAGAGSGAGAGSGAGAGSGAGAGSGAG.A	24.76	1947.827	29	974.9213	2	17.6	422	450	
G.SGAGAGSGAGAGSGAGAGSGAGAGSGAGSG.A	28.08	1762.747	26	882.3817	2	17.04	423	448	
G.SGAGAGSGAGAGSGAGAGSGAGAGSGA.G	24.46	1433.613	21	717.8138	2	16.71	423	443	
S.GAGAGSGAGAGSGAGAGSGAGAGSGAGAG.S	28.48	1531.661	23	766.8379	2	16.81	424	446	
S.GAGAGSGAGAGSGAGAGSGAGAGSGAGAG.S	24.31	1474.64	22	738.3276	2	17.12	424	445	
S.GAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGAG.S	22.88	1931.832	29	966.9222	2	18.01	424	452	
G.AGAGSGAGAGSGAGAGSGAGAGSGAGAG.S	25.25	1474.64	22	738.3275	2	16.61	425	446	
G.AGAGSGAGAGSGAGAGSGAGAG.S	22.71	1074.469	16	538.2416	2	15.2	425	440	
Y.GAGAGSGAASGAGAGSGAGAGSGAGAG.S	26.22	1417.618	21	709.8167	2	17.12	492	512	
G.AGAGSGAASGAGAGSGAGAGSGAGAG.S	24.26	1360.597	20	681.3059	2	16.66	493	512	
S.GAASGAGAGSGAGAGSGAGAGSGAGAGSGAGAG.S	26.89	1817.789	27	909.9017	2	18.11	498	524	
S.GAASGAGAGSGAGAGSGAGAGSGAGAG.S	25.68	1417.618	21	709.8165	2	16.61	498	518	
S.GAGAGSGAGAGSGAGAGSGAGAGSGAGAGSGAGV.G.Y	22.92	1159.522	17	580.7685	2	17.57	614	630	
Y.GAGAGSGAASGAGAGAGAGAGAGAG.T	26.14	1401.623	21	701.8187	2	17.52	652	672	
Y.GAGAGSGAASGAGAGAGAGAGAGAG.G	22.51	1216.543	18	609.2797	2	17.03	652	669	
Y.GAGAGSGAASGAGAGAGAGAGAGAGAG.G	22.27	1344.602	20	673.308	2	17.83	652	671	
G.AGAGSGAASGAGAGAGAGAGAGAG.G	25.37	1159.522	17	580.7681	2	16.82	653	669	
A.GAGSGAASGAGAGAGAGAGAGAGTGS.S	24.6	1518.666	22	760.3409	2	17.68	654	675	

G.AGSGAASGAGAGAGAGAG.T	24.8	1216.543	18	609.2791	2	16.46	655	672	
S.GAGSGAGAGSGAGAGAGSGAGAGSGA.G	25.92	1746.752	26	874.3838	2	17.72	1062	1087	
A.GAGSGAGAGAGSGAGAGSGAGAG.S	22.45	1531.661	23	766.8389	2	22.74	1068	1090	
A.GSGAGAGAGSGAGAGSGAGAGSGAG.A	27.36	1675.715	25	838.8656	2	16.92	1070	1094	
A.GSGAGAGAGSGAGAGSGAGAGSGA.A	23.33	1547.656	23	774.8354	2	16.4	1070	1092	
G.SGAGAGAGSGAGAGSGAGAGSGAGAG.S	26.62	1746.752	26	874.3834	2	17.2	1071	1096	
G.SGAGAGAGSGAGAGSGAGAGSGAGAG.A	25.64	1618.693	24	810.3542	2	17.17	1071	1094	
G.SGAGAGAGSGAGAGSGAG.A	22.92	1218.523	18	610.2684	2	15.4	1071	1088	
W.SSESDFGTGSGAGAG.S	24.2	1285.506	15	643.76	2	22.42	1249	1263	
Y.GAGAGSGAGSGAGAGSGAGAG.S	25.75	1403.603	21	702.8086	2	16.25	1307	1327	
G.AGAGSGAGSGAGAGSGAGAG.S	22	1346.581	20	674.2975	2	16.04	1308	1327	
S.GAGAGSGAGAGSGAGSGAGAGSGAGAG.S	27.16	1803.773	27	902.8934	2	17.42	2556	2582	
G.AGSGAGAGSGAGSGAGAGSGAGAGSGAG.A	22.02	1890.805	28	946.4095	2	17.64	2559	2586	
G.AGAGSGAASGAGAGAGSGAGAG.T	22.71	1488.655	22	745.3354	2	17.29	3194	3215	
A.GAGAGSGAGAGSGAGAGAGSGAGAG.Y	26.53	1659.72	25	830.8679	2	17.42	3883	3907	
A.GAGAGSGAGAGSGAGAGAGS.G	24.21	1346.581	20	674.2982	2	16.56	3883	3902	
S.GAASGAGAGAGAGSGAGAG.S	22.34	1273.565	19	637.7897	2	16.62	5031	5049	
K.SDFETGSGAASGAGAGAGSGAGAGSG.A	34.5	2012.831	26	1007.423	2	24.43	5162	5187	
K.SDFETGSGAASGAG.A	33.18	1212.489	14	607.2525	2	22.67	5162	5175	
K.SDFETGSGAASG.A	26.68	1084.431	12	543.2228	2	21.19	5162	5173	
G.SGAASGAGAGAGSGAGAGSGAG.A	23.45	1504.65	22	753.3325	2	16.81	5168	5189	
S.GAGAGSGAGAGSGAGAGSGAGAGG.S	24.49	1588.683	24	795.3486	2	16.93	5181	5204	
A.GAGSGAGAGSGAGAGSGAGAGG.S	24.19	1460.624	22	731.3192	2	16.41	5183	5204	
R.GYGQGAGSAASSVSSASSR.S	41.11	1685.761	19	843.8885	2	20.55	5213	5231	
Y.GQ(+.98)GAGSAASSVSSASSR.S	41.4	1466.66	17	734.3373	2	21.19	5215	5231	Deamidation (NQ)
Y.GQGAGSAASSVSSASSR.S	41.3	1465.676	17	733.8455	2	16.96	5215	5231	
G.QGAGSAASSVSSASSR.S	26.88	1408.654	16	705.335	2	16.83	5216	5231	
Q.GAGSAASSVSSASSR.S	37.28	1280.596	15	641.3055	2	16.33	5217	5231	
G.AGSAASSVSSASSR.S	32.81	1223.574	14	612.7944	2	15.73	5218	5231	
G.SAASSVSSASSR.S	22.2	1095.516	12	548.765	2	15.49	5220	5231	
R.RQ(+.98)LVVKFR.A	24.17	1045.64	8	523.8271	2	31.8	5249	5256	Deamidation (NQ)

CN-FA	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
R.DASGAVIEEQITTK.K	71.38	1460.736	14	731.3757	2	29.87	49	62	
R.DASGAVIEEQITTKK.M	58.32	1588.831	15	795.4226	2	25.86	49	63	
R.DASGAVIEEQ(+.98)ITTK.K	53.4	1461.72	14	731.8669	2	32.4	49	62	Deamidation (NQ)
R.DASGAVIEEQ(+.98)ITTKK.M	41.48	1589.815	15	795.9152	2	27.73	49	63	Deamidation (NQ)
D.ASGAVIEEQITTK.K	55.23	1345.709	13	673.8615	2	25.31	50	62	
D.ASGAVIEEQITTKK.M	35.07	1473.804	14	737.9092	2	22.04	50	63	
D.ASGAVIEEQ(+.98)ITTK.K	31.06	1346.693	13	674.3539	2	35.96	50	62	Deamidation (NQ)
A.SGAVIEEQITTK.K	51.47	1274.672	12	638.3427	2	24.79	51	62	
A.SGAVIEEQ(+.98)ITTKK.M	29.02	1403.751	13	702.8813	2	22.32	51	63	Deamidation (NQ)
A.SGAVIEEQITTKK.M	24.66	1402.767	13	702.3903	2	20.49	51	63	
S.GAVIEEQITTK.K	32.56	1187.64	11	594.8268	2	23.7	52	62	
S.GAVIEEQITTKK.M	30.05	1315.735	12	658.8745	2	19.97	52	63	
G.AVIEEQITTK.K	36.55	1130.618	10	566.3167	2	22.94	53	62	
A.VIEEQITTK.K	28.61	1059.581	9	530.7975	2	20.44	54	62	
R.KNKN(+.98)HGILGK.N	35.17	1108.635	10	555.3248	2	0.54	67	76	Deamidation (NQ)
K.N(+.98)KNHGILGKNEK.M	49.05	1351.721	12	451.5807	3	3.7	68	79	Deamidation (NQ)
K.NKN(+.98)HGILGKNEK.M	38.2	1351.721	12	451.5807	3	4.22	68	79	Deamidation (NQ)
K.NKN(+.98)HGILGKN(+.98)EK.M	34.49	1352.705	12	677.3591	2	16.3	68	79	Deamidation (NQ)
K.NKNHGILGKN(+.98)EK.M	34.45	1351.721	12	676.8668	2	15.46	68	79	Deamidation (NQ)
K.N(+.98)KNHGILGKN(+.98)EK.M	32.86	1352.705	12	451.9091	3	15.34	68	79	Deamidation (NQ)
K.NKNHGILGKNEK.M	31.94	1350.737	12	451.2525	3	1.06	68	79	
K.N(+.98)KN(+.98)HGILGKNEK.M	30.96	1352.705	12	451.9088	3	15.86	68	79	Deamidation (NQ)
K.N(+.98)KN(+.98)HGILGK.N	26.8	981.5243	9	491.7695	2	17.12	68	76	Deamidation (NQ)
K.N(+.98)KN(+.98)HGILGKN(+.98)EK.M	25.57	1353.689	12	452.2368	3	17.23	68	79	Deamidation (NQ)
K.N(+.98)HGILGKNEK.M	41.06	1109.583	10	555.7982	2	15.97	70	79	Deamidation (NQ)
K.NHGILGKN(+.98)EK.M	34.86	1109.583	10	555.7983	2	16.84	70	79	Deamidation (NQ)
K.NHGILGKN(+.98)EKM(+15.99)IK.T	29.53	1497.797	13	500.2731	3	16.52	70	82	Deamidation (NQ); Oxidation (M)
K.NHGILGKNEK.M	28.68	1108.599	10	555.3066	2	4.34	70	79	
K.M(+15.99)IKTFVITDSD.G	29.87	1385.675	12	693.8449	2	31.24	80	91	Oxidation (M)

K.TFVITTDSDGNESIVEEDVLM(+15.99)K.T	67.55	2457.147	22	1229.581	2	43.54	83	104	Oxidation (M)
K.TFVITTDSDGN(+.98)ESIVEEDVLM(+15.99)K.T	44.83	2458.131	22	1230.076	2	44.93	83	104	Deamidation (NQ); Oxidation (M)
K.TFVITTDSDGNE.S	31.27	1297.567	12	649.7911	2	34.34	83	94	
D.SDGNESIVEEDVLM(+15.99)K.T	26.41	1679.756	15	840.8853	2	32.33	90	104	Oxidation (M)
D.GN(+.98)ESIVEEDVLM(+15.99)K.T	42.62	1478.681	13	740.349	2	39.03	92	104	Deamidation (NQ); Oxidation (M)
D.GNESIVEEDVLM(+15.99)K.T	32.84	1477.697	13	739.8563	2	31.21	92	104	Oxidation (M)
K.TLSDGTVAQS.Y	31.24	977.4666	10	489.7405	2	20.23	105	114	
G.AGSGAGAGSGAGAGSGAGAGSG.A	26.42	1490.635	22	746.3249	2	16.48	421	442	
G.AGAGSGAGAGSGAGAGSGAGAG.S	33.44	1474.64	22	738.3278	2	16.67	425	446	
G.AGAGSGAGAGSGAGAGSGAGA.G	26.56	1417.618	21	709.8158	2	16.91	425	445	
Y.GAGAGSGAASGAGAGSGA.G	25	1232.538	18	617.2765	2	16.56	492	509	
Y.GAGAGSGAASGAGAGAGAGAG.T	25.79	1401.623	21	701.8189	2	17.61	652	672	
G.SGAGAGSGAGAGAGSGAGAGSGAG.A	24.78	1618.693	24	810.3536	2	16.79	1065	1088	
A.GSGAGAGAGSGAGAGSGAGAG.S	35.54	1403.603	21	702.8085	2	16.34	1070	1090	
Y.GAGAGSGAGSGAGAGSGAGAGSGAGAG.S	24.6	1803.773	27	902.8949	2	17.56	1307	1333	
K.SDFETGSGAASGAGAGAGSG.A	44.74	1612.66	20	807.3398	2	23.64	5162	5181	
K.SDFETGSGAASGAG.A	37.79	1212.489	14	607.2516	2	22.26	5162	5175	
K.SDFETGSGAASGA.G	26.35	1155.468	13	578.7413	2	22.88	5162	5174	
R.GYGQGAGSAASSVSSASSR.S	48.44	1685.761	19	843.8874	2	20.49	5213	5231	
Y.GQGAGSAASSVSSASSR.S	52.56	1465.676	17	733.8452	2	16.98	5215	5231	
Y.GQ(+.98)GAGSAASSVSSASSR.S	44.53	1466.66	17	734.3381	2	17.58	5215	5231	Deamidation (NQ)
G.QGAGSAASSVSSASSR.S	47.45	1408.654	16	705.3344	2	16.89	5216	5231	
Q.GAGSAASSVSSASSR.S	44.87	1280.596	15	641.3045	2	16.41	5217	5231	
G.AGSAASSVSSASSR.S	42.54	1223.574	14	612.7936	2	15.84	5218	5231	
R.RQLVVKFR.A	37.34	1044.656	8	523.3347	2	18.89	5249	5256	
R.RQ(+.98)LVVKFR.A	30.44	1045.64	8	523.827	2	24.5	5249	5256	Deamidation (NQ)
R.RQ(+.98)LVVK.F	26.99	742.4701	6	372.2423	2	14.99	5249	5254	Deamidation (NQ)
R.QLVVKFR.A	36.37	888.5545	7	445.2845	2	22.17	5250	5256	
R.Q(+.98)LVVKFR.A	35.34	889.5385	7	445.7767	2	23.59	5250	5256	Deamidation (NQ)

K.FRALPC(+57.02)VNC(+57.02)	41.37	1135.527	9	568.7707	2	28.23	5255	5263	Carbamidomethylation
R.ALPC(+57.02)VNC(+57.02)	59.92	832.3572	7	417.1858	2	26.3	5257	5263	Carbamidomethylation
HCI	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
R.GYGQGAGSAASSVSSASSR.S	89.43	1685.761	19	843.8879	2	22.26	5213	5231	
R.DASGAVIEEQITTK.K	75.03	1460.736	14	731.3755	2	26.75	49	62	
K.TFVITTDSDGNESIVEEDVLM(+15.99)K.T	66.28	2457.147	22	820.0574	3	33.16	83	104	Oxidation (M)
R.SYDYSR.R	39.77	789.3293	6	395.6716	2	20.94	5232	5237	
K.NC(+57.02)GIPR.R	31.45	715.3435	6	358.6787	2	20.48	5243	5248	Carbamidomethylation
R.Q(+.98)LVVK.F	26.78	586.369	5	587.3762	1	21.04	5250	5254	Deamidation (NQ)
R.QLVVK.F	24.97	585.385	5	293.6993	2	20.08	5250	5254	
R.RQLVVK.F	24.36	741.4861	6	371.7497	2	14.84	5249	5254	
R.RQ(+.98)LVVK.F	24.29	742.4701	6	372.2416	2	16.7	5249	5254	Deamidation (NQ)

Table 39. Peptide distribution of CN and CN-FA samples of *A. pernyi* silk

CN	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
R.AAGSAAAAAAAAAAAAASGAGR.S	78.93	1584.797	21	793.4056	2	46.86	2211	2231	
R.AAGSAAAAAAAAAAAAASGAGGAGGG.Y	56.31	1798.856	26	900.4359	2	55.13	1590	1615	
R.AAGSAAAAAAAAAAAAASGAGRSGGS.Y	54.16	1872.904	25	937.4615	2	46.83	2211	2235	
G.SGAGGIGGGFGRGD.G	42.28	1163.532	14	582.7733	2	38.51	2423	2436	
G.SGGSSAAAAAAAAAAAAAR.R	35.97	1401.696	18	701.8542	2	42.18	359	376	
G.AGGIGGGFGRGD.G	33.27	1019.478	12	510.7468	2	37.88	2425	2436	
A.GGIGGGFGRGD.G	29.11	948.4413	11	475.2274	2	37.51	2426	2436	
R.SAGSAAAAAAAAAAAAASGAGG.S	27.89	1572.749	22	787.381	2	56.3	2464	2485	
CN-FA	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
R.AAGSAAAAAAAAAAAAASGAGR.S	65.44	1584.797	21	793.4069	2	47.54	2211	2231	
Y.GSDSAAAAAAAAAAAAASGAR.G	61.32	1658.797	21	830.405	2	51.28	638	658	
R.AAGSAAAAAAAAAAAAASGAGGAGGG.Y	57.72	1798.856	26	900.4345	2	56.05	1590	1615	
R.SAGSAAAAAAAAAAAAASGAGGSGG.S	48.58	1773.824	25	887.9205	2	56.57	2464	2488	
A.GSGAGGIGGGFGR.G	44.23	1048.505	13	525.2596	2	38.56	2422	2434	
D.SAAAAAAAAAAAAASGAR.G	43.08	1399.717	18	700.8654	2	46.84	641	658	
R.GDGGYGSGSSAAAAAAAAAAAAAR.R	41.62	1907.872	24	954.9484	2	52.02	353	376	
D.SAAAAAAAAAAAAAGSGAGGIGGGFGRGD.G	39.04	2302.105	30	768.3777	3	56.22	2407	2436	
R.SAGSAAAAAAAAAAAAASGAGG.S	38.12	1572.749	22	787.3819	2	57.31	2464	2485	
G.SGAGGIGGGFGR.G	35.12	991.4835	12	496.7492	2	38.64	2423	2434	
Y.GSGSSAAAAAAAAAAAAAR.R	34.93	1458.717	19	730.3665	2	42.8	358	376	
A.GGRGDGGYGSGSSAAAAAAAAASA.A	32.58	1980.888	25	991.4512	2	56.68	536	560	
Y.GSGSSAAAAAAAAASAAR.R	31.71	1488.728	19	745.3712	2	43.72	544	562	
A.AGSGAGGIGGGFGR.G	30.66	1119.542	14	560.7784	2	38.71	2421	2434	
D.SSAAAAAAAAAAAAAGAGGSGGGYGW.G	28.93	1895.84	24	632.9533	3	51.19	2375	2398	
D.SAAAAAAAAAAAAAGSGAGGRGD.G	28.65	1742.83	23	872.4233	2	45.81	2161	2183	
S.GAGGIGGGFGR.G	25.71	904.4515	11	453.2331	2	38.43	2424	2434	
D.VVIKRVPGASSAAAASSASAGSGQTIIVER.Q	25.55	2926.584	31	732.6535	4	39.91	81	111	

Table 40. Peptide distribution of CN and CN-FA samples of *A. mylitta* silk

CN	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
R.QASHGAGGAAGAAAGAAASSVR.G	66.46	1881.904	23	628.3087	3	36.8	111	133	
Y.GSDSGAAAAAAAAAAAAASGAGGR.G	60.9	1829.862	24	915.9385	2	51.33	240	263	
D.SGAAAAAAAAAAAAASGAGGRGD.G	47.14	1742.83	23	872.4223	2	48.11	243	265	
Y.GSGSSAAAAAAAAAAAAASAAR.R	43.52	1559.765	20	780.8904	2	45.51	481	500	
Y.VDSHGQLVER.F	39.67	1138.573	10	570.2876	2	38.49	27	36	
CN-FA	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
Y.GSDSGAAAAAAAAAAAAASGAGGR.G	59.65	1829.862	24	915.9382	2	52.62	240	263	
D.SGAAAAAAAAAAAAASGAGGR.G	46.98	1570.781	21	786.398	2	46.95	243	263	
Y.GSGSSAAAAAAAAAAAAASAAR.R	46.68	1559.765	20	780.8909	2	44.92	481	500	
Y.GSDSGAAAAAAAAAAAAASGAGGRG.D	44.7	1886.883	25	944.4497	2	49.63	240	264	
D.SGAAAAAAAAAAAAASGAGGRGD.G	42.17	1742.83	23	872.4233	2	47.5	243	265	
G.SDSGAAAAAAAAAAAAASGAGGR.G	41.24	1772.84	23	887.428	2	50.52	241	263	
G.SGSSAAAAAAAAAAAAASAAR.R	39.72	1502.744	19	752.3802	2	44.69	482	500	
Y.GSDSGAAAAAAAAAAAAASGAGGRGD.G	39.04	2001.91	26	1001.962	2	51.07	240	265	
S.GSSAAAAAAAAAAAAASAAR.R	36.79	1415.712	18	708.8636	2	44.01	483	500	
Y.VDSHGQLVER.F	31.22	1138.573	10	380.5312	3	32.15	27	36	
D.SAAAAAAAAASGAGGR.G	29.33	1229.611	16	615.8126	2	36.11	152	167	
Y.GSDPGAAAAAAAAAAAAAS.G	28.62	1441.68	19	721.8529	2	67.68	452	470	

Table 41. Peptide distribution of CN and CN-FA samples of *A. yamamai* silk

CN	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
Y.GSDSAAGAAAAAAAAAASGAGGRG.D	41.16	1815.846	24	908.9301	2	46.93	1987	2010	
CN-FA	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
R.AAGSAAAAAAAAAAAAASGAGGSGR.G	66.08	1856.909	25	929.4623	2	49.43	908	932	
R.AAGSAAAAAAAAAAAAASGAGR.S	61.8	1655.834	22	828.9243	2	49.84	2351	2372	
Y.GSDSAAAAAAAAAAAAAAGSGAGR.A	43.99	1772.84	23	887.4284	2	49.4	486	508	
A.AAAAAAAGSGAGR.A	38.5	1142.579	15	572.2969	2	32.91	494	508	
Y.GSGSSAAAAAAAAAAAAASR.R	35.45	1488.728	19	745.3715	2	43.75	2742	2760	
D.SAAGAAAAAAAAAAAAASGAGGR.G	33.81	1499.744	20	750.8798	2	44.25	1990	2009	
A.AAAAAAAGSGAGGAGGD.Y	33.52	1201.532	17	601.7736	2	33.33	649	665	
D.SAAAAAAAAAAAAAAGSGAGR.A	32.99	1513.76	20	757.8875	2	44.9	489	508	
A.AAAAAAAGSGAGGAGGD.Y	31.85	1272.569	18	637.2923	2	33.9	648	665	
Y.EEDVVINR.V	30.68	972.4876	8	487.2511	2	36.63	78	85	
Y.GSDSAAGAAAAAAAAAAAAASGAGGR.G	29.25	1758.825	23	880.4191	2	47.58	1987	2009	
S.DSAAGAAAAAAAAAAAAASGAGGRG.D	28.6	1671.792	22	836.9041	2	48.52	1989	2010	
E.EDVVINR.V	28.43	843.445	7	422.7297	2	35.16	79	85	

Table 42. Peptide distribution of CN and CN-FA samples of *A. assamensis* silk

Peptide	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
R.VPGASSSAAAASSASAGGR.G	101.93	1560.749	19	781.3802	2	36.38	86	104	
R.HHDEYVDSN(+.98)GQLFER.F	74.84	1845.792	15	616.2698	3	44.02	23	37	Deamidation (NQ)
R.HHDEYVDSNGQLFER.F	74.28	1844.808	15	615.9415	3	42.01	23	37	
R.NAETRPNLSGNER.L	72.24	1456.702	13	729.3566	2	34.39	47	59	
D.SN(+.98)GQLFER.F	69.07	950.4457	8	476.2292	2	43.67	30	37	Deamidation (NQ)
Y.VDSN(+.98)GQLFER.F	68.6	1164.541	10	583.2765	2	48.8	28	37	Deamidation (NQ)
R.NAETRPNLSGNE.R	67.55	1300.601	12	651.306	2	36.35	47	58	
R.PNLSGNER.L	66.74	885.4304	8	443.7213	2	34.06	52	59	
Y.VDSNGQLFER.F	66.06	1163.557	10	582.7845	2	44.11	28	37	
K.RVPGASSSAAAASSASAGGR.G	63.88	1716.85	20	573.2896	3	34.74	85	104	
V.DSNGQLFER.F	62.51	1064.489	9	533.2504	2	47.78	29	37	
R.N(+.98)AETRPNLSGNER.L	62.49	1457.686	13	729.8485	2	35.4	47	59	Deamidation (NQ)
R.N(+.98)AETRPNLSGN(+.98)ER.L	60.18	1458.67	13	487.2294	3	37.31	47	59	Deamidation (NQ)
R.NAETRPN(+.98)LSGNER.L	57.7	1457.686	13	486.9009	3	35.56	47	59	Deamidation (NQ)
R.NAETRPNLSG.N	57.44	1057.515	10	529.7632	2	35.64	47	56	
V.DSN(+.98)GQLFER.F	57.4	1065.473	9	533.7422	2	50.94	29	37	Deamidation (NQ)
D.SNGQLFER.F	57.38	949.4617	8	475.7368	2	40.52	30	37	
D.SNGQ(+.98)LFER.F	56.11	950.4457	8	476.2289	2	41.96	30	37	Deamidation (NQ)
D.SAAAAAAAAAAAAAGSGGR.G	56.04	1442.723	19	722.3669	2	49.04	1128	1146	
D.SN(+.98)GQ(+.98)LFER.F	55.31	951.4297	8	476.7211	2	46.14	30	37	Deamidation (NQ)
R.NAETRPNLSGN(+.98).E	54.1	1172.542	11	587.2771	2	36.65	47	57	Deamidation (NQ)
R.AAGSAAAAAAAAAAAAAGSGGAGG.H	53.52	1599.76	23	800.8844	2	55.04	771	793	
R.NAETRPNLSGN(+.98)ER.L	52.23	1457.686	13	729.8486	2	35.07	47	59	Deamidation (NQ)
E.TRPNLSGNER.L	50.58	1142.579	10	572.2954	2	34.3	50	59	
N.AETRPNLSGNER.L	50.14	1342.659	12	672.3351	2	34.28	48	59	
R.NAETRPNLSGN.E	49.78	1171.558	11	586.7847	2	35.59	47	57	
R.NAETRPNL.S	47.78	913.4617	8	457.7371	2	36.31	47	54	
R.N(+.98)AETRPN(+.98)LSGNER.L	47.74	1458.67	13	730.3402	2	36.14	47	59	Deamidation (NQ)

N.GQLFER.F	47.12	748.3868	6	375.1995	2	40.25	32	37	
N.AETRPNLSGN(+.98)ER.L	46.12	1343.643	12	448.8871	3	34.64	48	59	Deamidation (NQ)
R.NAETRPNLS.G	45.49	1000.494	9	501.2527	2	35.35	47	55	
R.SAGGHPLL.S	45.23	750.4024	8	376.2072	2	39.49	2781	2788	
N.GQ(+.98)LFER.F	44.28	749.3708	6	375.6915	2	41.86	32	37	Deamidation (NQ)
N.AETRPN(+.98)LSGNER.L	42.97	1343.643	12	672.8274	2	37.87	48	59	Deamidation (NQ)
R.NAETRPN(+.98)LSGN(+.98)ER.L	42.31	1458.67	13	730.3398	2	36.68	47	59	Deamidation (NQ)
A.ETRPNLSGNER.L	42.24	1271.622	11	636.8167	2	35.37	49	59	
R.NAETRPNLSGN(+.98)ERLVETIVLEED.P	42.09	2598.277	23	867.0976	3	64.83	47	69	Deamidation (NQ)
R.NAETRPN(+.98)LSGNERLVETIVLEED.P	40.66	2598.277	23	867.0971	3	63.39	47	69	Deamidation (NQ)
E.TRPN(+.98)LSGNER.L	40.56	1143.563	10	572.7875	2	36.03	50	59	Deamidation (NQ)
R.PNLSGNERLVETIVLEED.P	39.79	2026.022	18	1014.016	2	63.89	52	69	
R.N(+.98)AETRPNL.S	39.07	914.4457	8	458.2291	2	39.86	47	54	Deamidation (NQ)
R.NAETRPNLSGNERLVETIVLEED.P	36.82	2597.293	23	866.7681	3	59.22	47	69	
N.GQLFERFTTR.K	36.8	1253.652	10	627.8323	2	46.62	32	41	
R.N(+.98)AETRPNLSGNERLVETIVLEED.P	35.31	2598.277	23	867.0978	3	62.39	47	69	Deamidation (NQ)
G.QLFER.F	34.66	691.3653	5	346.689	2	39.72	33	37	
Y.ESHDSFVVDSS.Y	33.99	1207.499	11	604.7559	2	44.19	109	119	
A.ETRPN(+.98)LSGNER.L	33.96	1272.606	11	637.309	2	37.54	49	59	Deamidation (NQ)
E.RN(+.98)AETRPNLSGNER.L	32.55	1613.787	14	538.9348	3	34.96	46	59	Deamidation (NQ)
CN-FA	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
R.VPGASSSAAAASSASAGGR.G	75.53	1560.749	19	781.3805	2	38.64	86	104	
E.GGYGSGSSGAAAAAAAAAAAAAAR.R	67.97	1806.861	23	904.4365	2	52.89	903	925	
H.DEYVDSN(+.98)GQLFER.F	62.96	1571.674	13	786.8433	2	55.51	25	37	Deamidation (NQ)
D.EYVDSNGQLFER.F	59.98	1455.663	12	728.838	2	49.89	26	37	
D.GGYGSGSSGAAAAAAAAAAAAAAR.R	59.49	1735.824	22	868.9189	2	50.09	177	198	
H.DEYVDSNGQLFER.F	58.39	1570.69	13	786.351	2	53.26	25	37	
D.GGYGSGSSGAAAAAAAAAAAAAAR.R	57.15	1877.898	24	939.9548	2	55.28	299	322	
W.GDGGYGSDSAAAAAAAAAAAAAAGSGGR.G	56.97	2150.958	27	1076.485	2	58.55	1120	1146	
R.HHDEYVDSN(+.98)GQLFER.F	56.84	1845.792	15	616.2693	3	44.42	23	37	Deamidation (NQ)
R.LVETIVLEEDPYGHEN.I	56.63	1855.884	16	928.948	2	59.43	60	75	

R.HHDEYVDSNGQLFER.F	55.55	1844.808	15	615.9421	3	43.39	23	37	
D.PYGHENIYEEDVVIK.R	54.99	1803.868	15	902.9392	2	47.9	70	84	
Y.VDSNGQLFER.F	53.22	1163.557	10	582.7849	2	44.78	28	37	
D.EYVDSN(+.98)GQLFER.F	52.44	1456.647	12	729.3294	2	51.64	26	37	Deamidation (NQ)
E.YVDSN(+.98)GQLFER.F	52.27	1327.604	11	664.8085	2	50.02	27	37	Deamidation (NQ)
E.YVDSNGQLFER.F	52.19	1326.62	11	664.3165	2	47.25	27	37	
R.LVETIVLEEDPYGHENI.Y	50.99	1968.968	17	985.489	2	62.08	60	76	
D.SN(+.98)GQLFER.F	50.28	950.4457	8	476.2293	2	43.44	30	37	Deamidation (NQ)
Y.VDSN(+.98)GQLFER.F	49.08	1164.541	10	583.2765	2	46.64	28	37	Deamidation (NQ)
R.NAETRPNLSGNE.R	48.6	1300.601	12	651.3064	2	38.69	47	58	
R.N(+.98)AETRPNLSGNER.L	47.13	1457.686	13	729.8489	2	38.2	47	59	Deamidation (NQ)
D.SNGQLFER.F	46.83	949.4617	8	475.7373	2	41.48	30	37	
R.LVETIVLEEDPYGHENIYEEDVVIK.R	46.33	2944.459	25	982.4924	3	62.54	60	84	
H.HDEYVDSNGQLFER.F	46.06	1707.749	14	570.256	3	46.44	24	37	
D.SAAAAAAAAAAAAAGSGGR.G	45.56	1442.723	19	722.3676	2	48.9	1128	1146	
R.AAGSAAAAAAAAAAAAAGSGAGGHGGGYG.W	43.25	2141.984	29	715.0004	3	54.94	205	233	
V.DSNGQLFER.F	42.73	1064.489	9	533.2505	2	45.31	29	37	
R.NAETRPNLSG.N	42.38	1057.515	10	529.764	2	38.36	47	56	
R.NAETRPNLSGNER.L	41.67	1456.702	13	486.5742	3	37.57	47	59	
V.DSN(+.98)GQLFER.F	41.42	1065.473	9	533.7423	2	47.69	29	37	Deamidation (NQ)
R.NAETRPNLSGN.E	41.26	1171.558	11	586.7855	2	38.32	47	57	
K.RVPGASSSAAAASSASAGGR.G	41.25	1716.85	20	573.2899	3	37.87	85	104	
R.NAETRPN(+.98)LSGNER.L	41.08	1457.686	13	729.8489	2	39.44	47	59	Deamidation (NQ)
D.PYGHENIYEEDVVIK.R	40.49	1959.969	16	654.3298	3	44.19	70	85	
D.SAAAAAAAAAAAAAGSGGR.G	39.69	1499.744	20	750.8784	2	48.68	845	864	
D.SN(+.98)GQ(+.98)LFER.F	39.41	951.4297	8	476.7212	2	46.6	30	37	Deamidation (NQ)
R.NAETRPNLS.G	38.87	1000.494	9	501.2531	2	38.19	47	55	
R.NAETRPNLSGN(+.98).E	38.18	1172.542	11	587.2775	2	38.75	47	57	Deamidation (NQ)
R.AAGSAAAAAAAAAAAAAGSGAGGHGGG.Y	38.12	1921.899	27	641.6385	3	52.66	205	231	
D.SAAAAAAAAAAAAAGSGGR.G	37.26	1371.685	18	686.8491	2	45.61	2207	2224	
R.N(+.98)AETRPN(+.98)LSGNER.L	35.89	1458.67	13	487.2294	3	38.6	47	59	Deamidation (NQ)

N.GQ(+.98)LFER.F	34.34	749.3708	6	375.6917	2	42.7	32	37	Deamidation (NQ)
R.NAETRPNL.S	33.97	913.4617	8	457.7372	2	38.72	47	54	
N.AETRPNL.SGNER.L	33.7	1342.659	12	672.3357	2	37.52	48	59	
N.GQLFER.F	32.3	748.3868	6	375.1999	2	41.3	32	37	
R.GGFYESHDSFVVD.S	31.65	1457.61	13	729.811	2	55.07	105	117	
R.GGFYESHD.S	31.01	910.3457	8	456.179	2	40.04	105	112	
D.PYGHENIYEED.V	30.63	1364.552	11	683.2817	2	42.93	70	80	
R.N(+.98)AETRPNL.S	29.92	914.4457	8	458.2291	2	40.93	47	54	Deamidation (NQ)
R.HHDEYVDSN(+.98).G	28.91	1115.416	9	558.7141	2	37.54	23	31	Deamidation (NQ)
R.AAGSAAAAAAAAAAAAAGSGAGGGHGGG.Y	28.89	1907.883	27	636.9675	3	48.8	771	797	
D.GGYGSGSSGAAAAAAAAAAAAAAR.R	28.53	1948.935	25	650.651	3	57.73	1539	1563	
D.SNGQ(+.98)LFERFTTR.K	28.52	1455.711	12	728.8608	2	52.03	30	41	Deamidation (NQ)
R.GGFYESHDSFVVDSSYGS.S	27.55	1938.791	18	970.4012	2	56.97	105	122	
N.GQLFERFTTR.K	27.46	1253.652	10	627.832	2	46.74	32	41	
R.AAGSAAAAAAAAAAAAAGSGAGGGHG.G	27.21	1807.856	25	603.6248	3	53.35	205	229	
R.GGGYGLGD.G	27.1	694.2922	8	695.2985	1	46.99	140	147	
D.EYVDSN(+.98)GQ(+.98)LFER.F	26.79	1457.631	12	729.8217	2	53.52	26	37	Deamidation (NQ)
A.ETRPNL.SGNER.L	25.85	1271.622	11	636.8169	2	37.48	49	59	
R.HHDEYVDSN(+.98)GQ(+.98)LFER.F	25.62	1846.776	15	616.603	3	62.33	23	37	Deamidation (NQ)
R.GGFYESHDSFVVD.S	24.41	1544.642	14	773.3264	2	54.27	105	118	
R.HHDEYVDSNG.Q	24.4	1171.453	10	586.733	2	37.24	23	32	
R.AAGSAAAAAAAAAAAAAGSGAGGSGYGYG.W	24.25	2126.962	28	710.0009	3	49.79	2561	2588	
R.N(+.98)AETRPNL.SG.N	24.13	1058.499	10	530.2557	2	39.89	47	56	Deamidation (NQ)
G.QLFER.F	23.34	691.3653	5	346.6891	2	40.91	33	37	
D.SN(+.98)GQLFERFTTR.K	22.67	1455.711	12	486.2433	3	51.98	30	41	Deamidation (NQ)
G.GSGGGYGWGD.EGYGSG.S	22.56	1461.543	16	731.7776	2	55.75	261	276	
R.SAGGHPLLSI.C	21.91	950.5185	10	476.2656	2	50.8	2781	2790	
R.NAETRPN(+.98)LSG.N	21.82	1058.499	10	530.2559	2	38.88	47	56	Deamidation (NQ)
R.NAETRPNL.SGN(+.98)ERLVETIVLEED.P	21.73	2598.277	23	867.0985	3	63.79	47	69	Deamidation (NQ)
G.SGSSGAAAAAAAAAAAAAARR.A	21.55	1628.834	20	543.9504	3	46.34	907	926	
D.GGYGSGSSGAAAAAAAAAAAAAARR.A	21.44	2033.999	25	679.0056	3	51.21	299	323	

R.NAETRPNLSGNERLVETIVLEED.P	21.25	2597.293	23	866.7689	3	59.37	47	69	Deamidation (NQ)
E.TRPN(+.98)LSGNER.L	20.82	1143.563	10	572.7881	2	38.58	50	59	
R.AAGSAAAAAAAAAAAAAGG.S	20.45	1270.627	18	636.3191	2	55.01	771	788	

Table 43. Peptide distribution of CN and CN-FA samples of *S. ricini* silk

CN	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
K.QVPQGGAASSAASSASAGSGSGAPTIIVER.G	99.73	2670.321	30	891.1126	3	48.89	101	130	
R.VHETGDGFLLR.G	86.21	1242.636	11	622.3241	2	44.88	2822	2832	
R.SLDDDIHSLE.R	80.53	1298.61	11	650.3113	2	48.34	20	30	
R.DSAPSRVPGGGTLVEKIVIER.A	80.06	2179.196	21	727.4041	3	51.95	64	84	
R.VPGGGTLVEK.I	77.69	955.5338	10	478.7729	2	38.25	70	79	
S.LDDDIHSLE.R	74.71	1211.578	10	606.7952	2	42.59	21	30	
R.VPGGGTLVEKIVIER.A	73.71	1565.914	15	783.9629	2	52.31	70	84	
R.DSAPSR.V	72.88	631.2925	6	316.6526	2	20.59	64	69	
R.APTGHETIYEEDVVIK.Q	72.62	1799.894	16	600.971	3	46.5	85	100	
S.SASAGSGSGAPTIIVER.G	71.43	1558.795	17	780.403	2	45.52	114	130	
A.SAGSGSGAPTIIVER.G	69.83	1400.726	15	701.369	2	45.26	116	130	
A.RSLDDDIHSLE.R	67.82	1454.711	12	728.3618	2	43.7	19	30	
R.DSAPSRVPGGGTLVEK.I	66.81	1568.816	16	785.4138	2	40.51	64	79	
L.DDDIHSLE.R	64.57	1098.494	9	550.2534	2	39.07	22	30	
S.GSGAPTIIVER.G	63.31	1098.603	11	550.308	2	44.85	120	130	
K.FERDSAPSR.V	62.52	1063.505	9	532.7585	2	33.48	61	69	
R.SLDDDIHSLE.R	61.28	1142.509	10	572.2612	2	53.85	20	29	
G.SGAPTIIVER.G	60.75	1041.582	10	521.7972	2	44.78	121	130	
E.TGDGFLLR.G	59.83	877.4658	8	439.7386	2	50.33	2825	2832	
D.DDIHSLE.R	59.63	983.4672	8	492.7397	2	37.27	23	30	
V.PGGGTVEK.I	59.06	856.4654	9	429.2392	2	51.42	71	79	
S.ASAGSGSGAPTIIVER.G	58.62	1471.763	16	736.887	2	45.45	115	130	
D.SAPSRVPGGGTLVEK.I	58.4	1453.789	15	727.9001	2	37.52	65	79	
G.SGSGAPTIIVER.G	58.28	1185.635	12	593.8242	2	44.84	119	130	
R.SLDDDIHSLE.R	57.41	1013.467	9	507.7392	2	53.02	20	28	
K.KFERDSAPSR.V	52.94	1191.6	10	596.806	2	30.5	60	69	
P.GGGTVEK.I	49.93	759.4127	8	380.7129	2	34	72	79	
R.VHETGDGFLLR	48.72	1086.535	10	544.2743	2	56.69	2822	2831	

S.GAPTIIVER.G	48.55	954.5498	9	478.2815	2	44.38	122	130
D.DIHSLER.G	47.42	868.4402	7	435.2264	2	38.41	24	30
R.APTGHETI.Y	47.04	824.4028	8	413.208	2	35.27	85	92
G.DGFLLR.G	46.95	719.3966	6	360.7053	2	56.73	2827	2832
T.GDGFLLR.G	46.84	776.418	7	389.2153	2	52.2	2826	2832
R.LGTSKIAVH	45.67	924.5392	9	463.2762	2	42.17	2872	2880
G.APTIIVER.G	44.49	897.5283	8	449.7701	2	43.05	123	130
D.IHSLER.G	41.58	753.4133	6	252.1444	3	32.76	25	30
G.GGTIVEK.I	40.07	702.3912	7	352.2022	2	33.87	73	79
R.VHETGD.G	39.71	656.2766	6	657.2819	1	19.37	2822	2827
R.VPGGGTLVE.K	39.24	827.4388	9	828.4457	1	47.84	70	78
K.IVIER.A	38.66	628.3908	5	315.2022	2	37.82	80	84
I.HSLER.G	35.08	640.3292	5	321.171	2	32.93	26	30
A.PSRVPGGGTLVEK.I	34.9	1295.72	13	648.8661	2	40.59	67	79
R.SLDDDIH.S	34.74	813.3505	7	407.6816	2	39.98	20	26
G.TLVEK.I	34.55	588.3483	5	295.1808	2	33.41	75	79
D.GFLLR.G	33.68	604.3696	5	303.1915	2	47.45	2828	2832
H.ETGDGFLLR.G	32.63	1006.508	9	504.258	2	52.66	2824	2832
S.RVPGGGTIVEK.I	32.57	1111.635	11	556.8235	2	35	69	79
R.SLDDD.I	28.52	563.2075	5	564.2138	1	34.91	20	24
A.RSLDD.D	27.84	604.2816	5	605.2872	1	21.66	19	23
A.RSLDDDIHSLERG.Y	26.78	1511.733	13	504.9174	3	40.39	19	31
D.GFLLRGD.Y	26.31	776.418	7	389.2154	2	49.33	2828	2834
F.ERDSAPSR.V	24.85	916.4362	8	459.2242	2	33.57	62	69
A.PTGHETI.Y	24.53	753.3657	7	754.3709	1	34.64	86	92

CN-FA	-10lgP	Mass	Length	m/z	z	RT	Start	End	PTM
A.SSAASSASAGSGSGAPTIIVER.G	95.06	1961.965	22	981.9882	2	51.82	109	130	Deamidation (NQ)
K.Q(+.98)VPQGGGAASSAASSASAGSGSGAPTIIVER.G	89.64	2671.305	30	1336.656	2	54.93	101	130	
K.QVPQGGGAASSAASSASAGSGSGAPTIIVER.G	87.58	2670.321	30	891.1132	3	54.28	101	130	
S.SAASSASAGSGSGAPTIIVER.G	85.47	1874.933	21	938.4727	2	51.81	110	130	
G.SDSAAAAAAAAAAAAAGGAGGGGAGR.G	82.34	1870.888	25	936.4496	2	59.64	2737	2761	

K.QVPQ(+.98)GGAASSAASSASAGSGSGAPTIIVER.G	81.28	2671.305	30	891.4409	3	54.81	101	130	Deamidation (NQ)
S.AASSASAGSGSGAPTIIVER.G	78.01	1787.901	20	894.9562	2	51.51	111	130	
D.SAAAAAAAAAAAAAGGAGGGAGR.G	76.17	1668.829	23	835.4197	2	55.62	2739	2761	
R.VHETGDGFLLR.G	72.06	1242.636	11	622.3237	2	51.74	2822	2832	
S.SASAGSGSGAPTIIVER.G	69.16	1558.795	17	780.4031	2	51.33	114	130	
R.VPGGGTLVEK.I	68.16	955.5338	10	478.7735	2	46.47	70	79	
A.ASSASAGSGSGAPTIIVER.G	67.97	1716.864	19	859.4373	2	51.37	112	130	
A.SAGSGSGAPTIIVER.G	64.23	1400.726	15	701.3693	2	51.07	116	130	
S.AGSGSGAPTIIVER.G	62.74	1313.694	14	657.853	2	50.92	117	130	
A.SSASAGSGSGAPTIIVER.G	60.99	1645.827	18	823.9193	2	51.31	113	130	
R.SAAAAAAAAAASSGAR.G	58.32	1273.638	16	637.8239	2	45.09	2802	2817	
G.SAAAAAAAAAAAAAGGAGGG.Y	57.29	1313.632	19	657.8223	2	56.85	522	540	
R.SLDDDIHSLER.G	56.8	1298.61	11	650.3096	2	2.94	20	30	
K.Q(+.98)VPQGGAASSAASSA.S	55.45	1288.59	15	645.3005	2	45.89	101	115	Deamidation (NQ)
K.QVPQGGAASSAASSASAG.S	54.91	1502.696	18	752.3537	2	45.78	101	118	
G.SAAGAAAAAAAAAAAAAGGAGR.G	54.39	1554.786	21	778.3991	2	56.94	144	164	
G.SAAAAAAAAAAAAAGGAGG.G	53.83	1256.611	18	629.3113	2	56.97	522	539	
R.GHGTAGSAAAAAAAAAAAAASSEGG.S	53.42	1925.883	25	963.9461	2	58.26	173	197	
S.ASAGSGSGAPTIIVER.G	53.3	1471.763	16	736.8871	2	51.26	115	130	
S.SAAAAAAAAAAAAAGGAGGG.Y	53.11	1384.669	20	693.3408	2	60.9	1189	1208	
S.GSGAPTIIVER.G	51.1	1098.603	11	550.3078	2	50.81	120	130	
K.Q(+.98)VPQ(+.98)GGAASSAASSA.S	50.73	1289.574	15	645.7927	2	46.89	101	115	Deamidation (NQ)
D.GGAAAAAAAAAAAAAGGSR.S	50.64	1412.712	19	707.3624	2	54.78	1013	1031	
E.TGDGFLLR.G	50.47	877.4658	8	439.7391	2	57.49	2825	2832	
G.SGSGAPTIIVER.G	49	1185.635	12	593.8237	2	50.79	119	130	
R.SHGAGSAAGAAAAAAAAAAAAAGGA.G	48.98	1750.835	24	876.4226	2	59.12	139	162	
A.AAAAAAAGGAGGGAGR.G	47.86	1297.649	18	649.8305	2	44.53	2744	2761	
V.PGGGTLVEK.I	45.78	856.4654	9	429.239	2	57.37	71	79	
R.DSAPSRVPGGGTLVEK.I	45.73	1568.816	16	785.4135	2	47.96	64	79	
S.DGGAAAAAAAAAAAAAGGAG.R	45	1412.664	20	707.3386	2	69.26	294	313	
R.SHGAGSAAGAAAAAAAAAAAAAGGAG.R	44.56	1807.856	25	904.9331	2	58.47	139	163	

D.GGAAAAAAAAAAAAAGGAGR.G	43.99	1453.739	20	727.8752	2	56.09	295	314	Deamidation (NQ)
D.SGAAAAAAAAAAAAAGGAGD.G	42.93	1442.675	20	722.3439	2	62.21	1046	1065	
R.VHETGDGFLL.R	42.17	1086.535	10	544.2736	2	62.44	2822	2831	
D.GGAAAAAAAAAAAAAGGAGGG.Y	41.06	1411.68	21	706.8454	2	61.55	380	400	
R.SHGAGSAAGAAAAAAAAAAAAAGG.A	40.95	1679.798	23	840.9039	2	57.44	139	161	
K.Q(+.98)VPQGGAASSAA.S	40.29	1043.488	12	522.7504	2	45.53	101	112	
D.GGAAAAAAAAAAAAASGAG.S	40.19	1327.648	19	664.83	2	62.14	262	280	
A.PTIIVER.G	40.15	826.4912	7	414.252	2	48.01	124	130	
S.DSAAAAAAAAAAAAAGGAG.G	40.12	1385.653	19	693.8327	2	80.93	2738	2756	
G.SAAAAAAAAAAAAAGGAG.G	40.03	1199.589	17	600.8005	2	56.99	522	538	
S.AEAAAAAAAAAAAAAGTAAG.G	39.99	1427.7	19	714.8615	2	70.97	234	252	
S.SAAAAAAAAAAAAAGGAGG.G	39.92	1327.648	19	664.8302	2	61.05	1189	1207	
G.GAAAAAAAAAAAAAGGAGGG.Y	39.72	1354.659	20	678.3347	2	60.13	381	400	
S.DGGAAAAAAAAAAAAAGGA.G	39.62	1355.643	19	678.8272	2	75.36	294	312	
D.NAAAAAAAAAAAAAGGAGG.D	39.6	1283.622	18	642.8167	2	57.72	2775	2792	
D.GGAAAAAAAAAAAAAGGAGD.G	39.11	1412.664	20	707.3381	2	62.38	497	516	
D.SAEAAAAAAAAAAAAAGTAAGG.S	38.92	1571.754	21	786.8823	2	65.97	233	253	
A.GSGSGAPTIIVER.G	38.74	1242.657	13	622.3341	2	50.8	118	130	
A.AAAAAAAAAAAGGAGGGAGR.G	38.67	1368.686	19	685.3485	2	45.12	2743	2761	
G.SGAPTIIVER.G	38.59	1041.582	10	521.7974	2	50.91	121	130	
R.GHGTAGSAAAAAAAAAAAAASSE.G	38.39	1811.84	23	906.9249	2	58.66	173	195	
S.DSAAAAAAAAAAAAAGGAGGG.A	37.61	1499.696	21	750.853	2	78	2738	2758	
D.SSAAAAAAAAAAAAAGSAGG.S	37.4	1302.616	18	652.314	2	55.14	210	227	
R.DSAPSRVPGGG.T	36.22	998.4781	11	500.2451	2	45.41	64	74	
D.NAAAAAAAAAAAAAGGAGGD.Y	36.01	1398.649	19	700.3302	2	58.48	2775	2793	
D.SGAAAAAAAAAAGAAGGAGG.G	35.42	1370.654	20	686.3325	2	55.79	2421	2440	
G.GDGGAAAAAAAAAAAAAGGAG.S	34.61	1469.686	21	735.8488	2	70.41	411	431	
R.GHGTAGSAAAAAAAAAAAAAS.S	34.53	1595.765	21	798.8883	2	56.27	173	193	
Y.GAGSAAAAAAAAAAAAAGGAGG.G	34.35	1441.691	21	721.8517	2	61.49	519	539	
Y.GGDGGAIAAAAAAAAAAAGGAG.S	34.32	1526.707	22	764.3596	2	69.83	410	431	
R.APTGHETI.Y	34.13	824.4028	8	413.2078	2	44.58	85	92	

D.GGAAAAAAAAAAAAAGGA.G	33.98	1240.616	18	621.314	2	62.5	295	312	
S.GAPTIIVER.G	33.95	954.5498	9	478.2807	2	50.46	122	130	
G.GAAAAAAAAAAAAAGGAGSG.Y	33.95	1384.669	20	693.3406	2	59.75	414	433	
S.DSAAAAAAAAAAAAAGGAGG.G	33.13	1442.675	20	722.3428	2	78.44	2738	2757	
G.SDGGAAAAAAAAAAAAAGG.A	33.1	1371.638	19	686.8253	2	65.79	293	311	
G.SDSGAAAAAAAAAAAAAGGAG.D	32.83	1529.707	21	765.8591	2	66.25	1044	1064	
D.SGAAAAAAAAAAGAAGGAGGG.Y	31.92	1427.675	21	714.8434	2	55.68	2421	2441	
D.AAAAAAAAAASAAAGGAG.S	31.7	1199.589	17	600.8005	2	50.56	2673	2689	
G.SDGGAAAAAAAAAAAAAGGAG.R	31.7	1499.696	21	750.8542	2	70.3	293	313	
H.ETGDGFLLR.G	31.5	1006.508	9	504.2604	2	58.51	2824	2832	
D.GGAAAAAAAAAAAAAGGAG.R	31.32	1297.638	19	649.8249	2	61.75	295	313	
A.AAAAAAAGGAGR.G	31.21	1055.547	14	528.78	2	44.41	151	164	
D.SGAAAAAAAAAAAAAGGAG.D	30.97	1327.648	19	664.8307	2	61.47	1046	1064	
D.GGAAAAAAAAAAAAAGGSG.G	30.93	1313.632	19	657.8221	2	60.25	472	490	
G.GAAAAAAAAAAAAAGGSG.G	30.92	1256.611	18	629.3112	2	58.76	473	490	
R.SHGAGSAAGAAAAAAAAAAG.G	30.84	1622.776	22	812.3935	2	56.39	139	160	
N.SGAAAAAAAAAAGGAG.S	30.68	1185.574	17	593.7935	2	54.04	2704	2720	
D.GGAAAAAAAAAAAAAGGSGG.G	30.6	1370.654	20	686.332	2	60.18	472	491	
D.N(+.98)AAAAAAAAAAAAAGGAGG.D	29.7	1284.606	18	643.308	2	76.35	2775	2792	Deamidation (NQ)
D.GGAAAAAAAAAAAAAGGSGGG.Y	29.55	1427.675	21	476.898	3	60.04	472	492	
D.SGAAAAAAAAAAAAASGGAG.S	29.19	1343.643	19	672.8284	2	58.88	2388	2406	
D.GGAAAAAAAAAAAAAGGS.G	28.96	1256.611	18	629.3116	2	60.53	472	489	
Y.GPGSAAAAAAAAAAAAAGGA.G	28.89	1353.664	19	677.8436	2	66.01	1068	1086	
G.GAAAAAAAAAAAAAGGAGG.G	28.56	1297.638	19	649.8246	2	60.27	381	399	
A.AGAAAAAAAAAAAAAGGA.G	28.55	1183.595	17	592.8036	2	61.03	146	162	
D.N(+.98)AAAAAAAAAAAAAGGA.G	28.43	1170.563	16	586.2877	2	78.52	2775	2790	Deamidation (NQ)
D.SSAAAAAAAAAAAAAGGAGG.G	28.39	1414.68	20	708.345	2	62.94	1188	1207	
G.GAAAAAAAAAAAAAGGAGS.G	28.13	1327.648	19	664.8303	2	60.05	414	432	
R.SHGAGSAAGAAAAAAAAA.A	27.94	1494.717	20	748.3641	2	52.37	139	158	
A.AAAAAAAGGAGGG.Y	27.83	1084.526	16	543.2693	2	46.59	385	400	
A.EAAAAAAAAAAGTAAG.G	27.72	1356.663	18	679.3433	2	67.21	235	252	

A.AAAAAAAAAAGGSR.S	27.26	1085.558	14	543.7852	2	44.62	1018	1031
D.GGAAAAAAAAAAAAAGG.A	27.05	1169.579	17	585.7957	2	60.56	295	311
G.AAAAAAAAAAAAAAGGAGD.G	26.87	1298.622	18	650.3168	2	82.99	499	516
D.GGAAAAAAAAAAAAAGGAG.S	26.61	1368.675	20	685.3435	2	64.17	804	823
A.GSAAAAAAAAAAAAAGGAGGG.Y	26.51	1370.654	20	686.3325	2	58.72	521	540
G.GGTLVEK.I	26.31	702.3912	7	352.202	2	26.73	73	79
G.AAAAAAAAAAAAAAGGAGDG.Y	26.21	1355.643	19	678.8267	2	81.64	499	517
S.AEAAAAAAAAAAAAAG.T	25.78	1127.557	15	564.785	2	83.01	234	248
R.SHGAGSAAGAAAAAAAA.A	25.76	1352.643	18	677.328	2	47.8	139	156
A.AGAAAAAAAAAAAAAGGAGR.G	25.66	1396.717	19	699.3648	2	54.98	146	164
K.IVIER.A	25.63	628.3908	5	315.2021	2	45.77	80	84

Bibliography

- Abdel-Fattah, W. I., Atwa, N. & Ali, G. W. 2015. Influence of the Protocol of Fibroin Extraction on the Antibiotic Activities of the Constructed Composites. *Prog Biomater*, 4, 77-88.
- Aebersold, R. & Mann, M. 2003. Mass Spectrometry-Based Proteomics. *Nature*, 422, 198-207.
- Ahmad, S., Tripathi, R., Singh, I., *et al.* 2016. Studies on the Haemocytes of Mulberry Silkworm Bombyx Mori L in the Region of District Amethi, Uttar Pradesh, India. *International Journal of Entomology Research*, 1, 29-32.
- Ahmed, S. 2013. Attacus Atlas L.(Lepidoptera: Saturniidae): A New Record on Ailanthus Excelsa Roxb. From Assam, India. *World Journal of Zoology*, 8, 127-130.
- Ajisawa, A. 1998. Dissolution of Silk Fibroin with Calcium Chloride/Ethanol Aqueous Solution. *J. Seric. Sci. Jpn.*, 67, 91-94.
- Akai, H. 2002. Knowledge of Wild Silk a, B, C (in Japanese). *Monthly Dying* 11, 11, 33-67.
- Akai, H. & Nagashima, T. 1999. Fine-Structural Characteristics of Anaphe Cocoon Filament. *International Journal of Wild Silkworm and Silk (Japan)*.
- Akerele, L. L. 1970. *A Review of the Technique of Silkworm Rearing and Silk Production*, Federal Ministry of Industries.
- Alam, I. 1999. History of Sericulture in India to the 17th century. *Proceedings of the Indian History Congress*, 60, 339-352.
- Alciphron 1896. *Alciphron, Literally and Completely Translated from the Greek, with Introduction and Notes*, Athens, Privately printed for the Athenian Society.
- Allchin, B. & Allchin, R. 1982. *The Rise of Civilization in India and Pakistan*, Cambridge University Press.
- Anfinsen, C. B. 1972. The Formation and Stabilization of Protein Structure. *Biochemical Journal*, 128, 737-749.
- Aruga, H. 1994. *Principles of Sericulture*, New Delhi Oxford & IBH Pub. Co.
- Arunkumar, K. P., Metta, M. & Nagaraju, J. 2006. Molecular Phylogeny of Silkworms Reveals the Origin of Domesticated Silkworm, Bombyx Mori from Chinese Bombyx Mandarina and Paternal Inheritance of Antheraea Proylei Mitochondrial DNA. *Mol Phylogenet Evol*, 40, 419-27.
- Arunkumar, K. P., Tomar, A., Daimon, T., *et al.* 2008. Wild Silk Base: An Est Database of Wild Silkworms. July 2008 ed.: BMC Genomics

- Aruz, J., Al-Asa'Ad, W. K., Cussini, E., *et al.* 2018. *Palmyra: Mirage in the Desert*, Yale University Press.
- Asakura, T., Iwadate, M., Demura, M., *et al.* 1999. Structural Analysis of Silk with ¹³C Nmr Chemical Shift Contour Plots. *International Journal of Biological Macromolecules*, 24, 167-171.
- Asakura, T., Kashiba, H. & Yoshimizu, H. 1988. 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, 644-648.
- Asakura, T. & Miller, T. (eds.) 2014. *Biotechnology of Silk*: Springer.
- Asakura, T., Nishimura, A., Kametani, S., *et al.* 2017. Refined Crystal Structure of Samia Cynthia Ricini Silk Fibroin Revealed by Solid-State Nmr Investigations. *Biomacromolecules*, 18, 1965-1974.
- Asakura, T., Yao, J., Yamane, T., *et al.* 2002. Heterogeneous Structure of Silk Fibers from Bombyx Mori Resolved by ¹³C Solid-State Nmr Spectroscopy. *Journal of the American Chemical Society*, 124, 8794-8795.
- Asakura, T., Yoshimizu, H. & Yoshizawa, F. 1988. Nmr of Silk Fibroin. 9. Sequence and Conformation Analyses of the Silk Fibroins from Bombyx Mori and Philosamia Cynthia Ricini by ¹⁵N Nmr Spectroscopy. *Macromolecules*, 21, 2038-2041.
- Babu, K. M. 2012. Silk Fibres. In: Kozlowski, R. M. (ed.) *Handbook of Natural Fibres : Types, Properties and Factors Affecting Breeding and Cultivation*. Elsevier Science.
- Bae, H. S. 2002. *Study on the Dyeability, Antibacterial and Deodorization Activities of Extracts from Coptis Chinensis Franch by Electrolytic Water*. PhD, Keimyung University.
- Banck-Burgess, J. 2012. *Case Study: The Textiles from the Princely Burial at Eberdingen-Hochdorf, Germany*, na.
- Banno, Y., Nakamura, T., Nagashima, E., *et al.* 2004. M Chromosome of the Wild Silkworm, Bombyx Mandarina (N = 27), Corresponds to Two Chromosomes in the Domesticated Silkworm, Bombyx Mori (N = 28). *Genome*, 47, 96-101.
- Barber, E. J. W. 1991. *Prehistoric Textiles : The Development of Cloth in the Neolithic and Bronze Ages with Special Reference to the Aegean*, Princeton, N.J., Princeton University Press.
- Beck, S. W. 1882. *The Draper's Dictionary. A Manual of Textile Fabrics; Their History and Applications*, London, The Warehousemen & Drapers' Journal Office.

- Bedat, I., Desrosiers, S., Moulherat, C., *et al.* Two Gallo-Roman Graves Recently Found in Naintré (Vienne, France). *In: Pritchard, F. & Wild, J. P., eds. Northern Archaeological Textiles*, 1999 Edinburgh. Oxbow Books, 5-11.
- Benchamin, K. V. & Jolly, M. S. Studies on Ingestion, Digestion and Conversion Efficiency in Polyvoltine and Bivoltine Silkworm Breeds (*Bombyx Mori* L.). *Proceeding of the 3rd Oriental Entomology Symposium*, 1984. 195-206.
- Benjamin, C. 2018. *Empires of Ancient Eurasia: The First Silk Roads Era, 100 Bce – 250 Ce*, Cambridge University Press.
- Benjamin Cummings-Pearson Education, I. P. 2008. *Biology, 8th Edition, Campbell & Reece, 2008: Biology*.
- Berche, P. 2012. Louis Pasteur, from Crystals of Life to Vaccination. *Clinical Microbiology and Infection*, 18, 1-6.
- Berti, G. 1960. Studio Spettrofotometrico Dell'acidita Di Alcuni Nitroderivati Indolici. *Gazz. Chim. Ital.*, 90, 539-558.
- Bhat, N. & Ahirrao, S. 1983. Investigation of the Structure of Silk Film Regenerated with Lithium Thiocyanate Solution. *Journal of Polymer Science Part A: Polymer Chemistry*, 21, 1273-1280.
- Bhatia, N. K. & Yousuf, M. 2015. Effects of Forestry Host Plants, Rearing Seasons and Their Interaction on Cocoon Productivity of Tropical Tasar Silkworm, *Antheraea Mylitta* in Uttarakhand. *International Journal of Industrial Entomology*, 30, 31-39.
- Biek, L. 1963. Lib Soils. *Science in Archaeology: A Comprehensive Survey of Progress and Research*, 108.
- Bock, F. 1895. *Die Textilien Byssus-Reliquien Des Christlichen Abendlandes, Aufbewahrt in D. Kirchen Zu Köln, Aachen, Cornelimünster, Mainz U. Prag*, La Ruelle'sche Accidenzdr.
- Bora, D. S., Deka, B. & Sen, A. 2013. Host Plant Selection by Larvae of the Muga Silk Moth, *Antheraea Assamensis*, and the Role of the Antenna and Maxillary Palp. *Journal of Insect Science*, 13, 52.
- Braga, D., Maini, L., Polito, M., *et al.* 2004. Hydrogen Bonding Interactions between Ions: A Powerful Tool in Molecular Crystal Engineering. *In: Mingos, D. M. P., Braga, D., Grepioni, F., et al. (eds.) Supramolecular Assembly Via Hydrogen Bonds II*. Springer.
- Bretschneider, E. 1871. *On the Knowledge Possessed by the Ancient Chinese : Of the Arabs and Arabian Colonies, and Other... Western Countries, Mentioned in Chinese*

Books, London, Trübner & co.

- Brown, J., Lu, C. L., Coburn, J., *et al.* 2015. Impact of Silk Biomaterial Structure on Proteolysis. *Acta Biomater*, 11, 212-21.
- Buckley, M., Collins, M., Thomas-Oaies, J., *et al.* 2009. Species Identification by Analysis of Bone Collagen Using Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry. *Rapid communications in mass spectrometry : RCM*, 23, 3843-3854.
- Buckley, M., Whitcher Kansa, S., Howard, S., *et al.* 2010. Distinguishing between Archaeological Sheep and Goat Bones Using a Single Collagen Peptide. *Journal of Archaeological Science*, 37, 13-20.
- Cai, K., Yao, K., Cui, Y., *et al.* 2002. Influence of Different Surface Modification Treatments on Poly (D, L-Lactic Acid) with Silk Fibroin and Their Effects on the Culture of Osteoblast in Vitro. *Biomaterials*, 23, 1603-1611.
- Campbell, N. A. 2008. *Biology*, Pearson Education Australia.
- Capinera, J. L. 2008. Encyclopedia of Entomology *In*: Capinera, J. L. & Springerlink (eds.). Dordrecht : Springer Netherlands.
- Centoducati, G., Tarsitano, E., Bottalico, A., *et al.* 2007. Monitoring of the Endangered *Pinna Nobilis* Linne, 1758 in the Mar Grande of Taranto (Ionian Sea, Italy). *Environ Monit Assess*, 131, 339-47.
- Chauvy, G. 2000. La Dure Condition Des Forçats Du Luxe. *Historia*. Tallandier.
- Chen, F., Porter, D. & Vollrath, F. 2012. Morphology and Structure of Silkworm Cocoons. *Materials Science and Engineering: C*, 32, 772-778.
- Chen, G., Arai, M. & Hirabayashi, K. 1996. Isolation of Tyrosine from Silk Fibroin by Enzyme Hydrolysis. *J. Seric. Sci. Jpn.*, 65, 182-184.
- Chen, X., Knight, D. P., Shao, Z., *et al.* 2001. Regenerated Bombyx Silk Solutions Studied with Rheometry and Ftir. *Polymer*, 42, 09969-09974.
- Chevillard, M., Couble, P. & Prudhomme, J.-C. 1986. Complete Nucleotide Sequence of the Gene Encoding the Bombyx Mori Silk Protein P25 and Predicted Amino Acid Sequence of the Protein. *Nucleic acids research*, 14, 6341.
- Choudhury, A. R. 2006. *Textile Preparation and Dyeing*, Science publishers.
- Coblin, W. S. 1993. "Erh Ya (爾雅)". *In*: Loewe, M. (ed.) *Early Chinese Texts: A Bibliographical Guide*. Berkeley: Society for the Study of Early China.

- Cole, C. 2005. The Golden Fleece. *Noesis: The Journal of the Mega Society*, 7-15.
- Cottrell, J. S. 2011. Protein Identification Using Ms/Ms Data. *J Proteomics*, 74, 1842-51.
- Craven, K. R. 1821. *A Tour through the Southern Provinces of the Kingdom of Naples*, Rodwell and Martin.
- Crimmins, D. L., Mische, S. M. & Denslow, N. D. 2005. Chemical Cleavage of Proteins in Solution. *Current Protocols in Protein Science*, 41, 11.4.1-11.4.11.
- Cybulska, M. & Maik, J. 2007. Archaeological Textiles – a Need for New methods of Analysis and Reconstruction. *Fibers and Textiles in Eastern Europe*, 15, 64-65.
- Dale, S. F. 2008. “Silk Road, Cotton Road Or . . . Indo-Chinese Trade in Pre-European Times”. *Modern Asian Studies*, 43, 79.
- Datta, A., Ghosh, A. K. & Kundu, S. 2001. Purification and Characterization of Fibroin from the Tropical Saturniid Silkworm, *Antheraea Mylitta*. *Insect Biochemistry and Molecular Biology*, 31, 1013-1018.
- Datta, R. K. & Nanavaty, M. 2005. *Global Silk Industry : A Complete Source Book*, Boca Raton, FL., Universal Publishers.
- Day, F. E. 1950. Silks of the near East. *The Metropolitan Museum of Art Bulletin*, 9, 108-117.
- De Fazi, R., Berti, G. & Passetti, A. D. S. 1959. *Il 6-Nitrotriptofano*.
- Dee, M. W., Brock, F., Harris, S. A., *et al.* 2010. Investigating the Likelihood of a Reservoir Offset in the Radiocarbon Record for Ancient Egypt. *Journal of Archaeological Science*, 37, 687-693.
- Delahunty, C. & Yates, J. R., 3rd 2005. Protein Identification Using 2d-Lc-Ms/Ms. *Methods*, 35, 248-55.
- Demassa, J. M. & Boudreaux, E. A. 2013. Formic Acid: A Significant but Ignored Product in the Miller-Urey Experiment. *Creation Research Society Quarterly*, 49, 211-217.
- Diana, A., Reguzzoni, M., Congiu, T., *et al.* 2017. The Byssus Threads of *Pinna Nobilis*: A Histochemical and Ultrastructural Study. *European Journal of Histochemistry : EJH*, 61, 2779.
- Doira, H., Fujii, H., Kawaguchi, Y., *et al.* 1992. Genetic Stocks and Mutations of *Bombyx Mori*. *Institute of Genetic Resources, Kyushu University, Japan*, 54.
- Duby, G. & Mandrou, R. 1964. *A History of French Civilization*, Random House.

- Dudgeon, G. C. 1908. Silk-Producing Insects of West Africa. *Nature*, 79, 160.
- Edgar, R. C. 2004. Muscle: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic acids research*, 32, 1792-1797.
- Feltham, H. B. 2009. Justinian and the International Silk Trade. *SINO-PLATONIC PAPERS*.
- Feltwell, J. 1991. *The Story of Silk*, St. Martin's Press.
- Forbes, W. T. M. 1930. The Silkworm of Aristotle. *Classical Philology*, 25, 22-26.
- Freddi, G., Mossotti, R. & Innocenti, R. 2003. Degumming of Silk Fabric with Several Proteases. *Journal of Biotechnology*, 106, 101-112.
- Freddi, G., Pessina, G. & Tsukada, M. 1999. Swelling and Dissolution of Silk Fibroin (Bombyx Mori) in N-Methyl Morpholine N-Oxide. *International Journal of Biological Macromolecules*, 24, 251-263.
- Furdui, E. M., Marghitas, L. A., Dezmirean, D. S., *et al.* 2014. Genetic Characterization of Bombyx Mori (Lepidoptera: Bombycidae) Breeding and Hybrid Lines with Different Geographic Origins. *J Insect Sci*, 14.
- Furuhata, K., Okada, A., Chen, Y., *et al.* 1994. Dissolution of Silk Fibroin in Lithium Halide/Organic Amide Solvent Systems. *J. Seric. Sci. Jpn.*, 63, 315-322.
- Garel, A., Deleage, G. & Prudhomme, J.-C. 1997. Structure and Organization of the Bombyx Mori Sericin 1 Gene and of the Sericins 1 Deduced from the Sequence of the Ser 1b Cdna. *Insect Biochemistry and Molecular Biology*, 27, 469-477.
- Gaspari, M. & Cuda, G. 2011. Nano LC-MS/MS: A Robust Setup for Proteomic Analysis. *Methods Mol Biol*, 790, 115-26.
- Gasteiger, E., Hoogland, C., Gattiker, A., *et al.* 2005. Protein Identification and Analysis Tools on the ExPASy Server. *The Proteomics Protocols Handbook*. Springer.
- Gawlikowski, M. 2014. Palmyra as a Trading Centre. *Iraq*, 56, 27-33.
- Ghosh, G. K. & Ghosh, S. 1995. *Indian Textiles : Past and Present*, New Delhi, APH Pub.
- Ghosh, J., Senapati, S. & Baral, K. 2000. Performance of Mulberry Varieties on Bivoltine Silkworm (Bombyx Mori L.) with Reference to Terai Ecological Conditions of West Bengal. *Indian Journal of Agricultural Research*, 34, 141-146.
- Gierasch, L. M. 1989. Signal Sequences. *Biochemistry*, 28, 923-930.
- Gille, B. 1986. *History of Techniques*, New York, Gordon and Breach Science Publishers.

- Gilroy, C. G. 1853. *The History of Silk, Cotton, Linen, Wool, and Other Fibrous Substances*, New York, New London, Conn., C. M. Saxton; E. R. Fellows.
- Goel, R. K. & Krishna Rao, J. V. 2004. *Oak Tasar Culture : Aboriginal of Himalayas*, New Delhi, A.P.H. Pub. Corp.
- Gong, Y., Li, L., Gong, D., *et al.* 2016. Biomolecular Evidence of Silk from 8,500 Years Ago. *PLoS One*, 11, e0168042.
- Gonnet, F., Gilles Lemaître, Gilles Waksman, *et al.* 2003. Maldi/Ms Peptide Mass Fingerprinting for Proteome Analysis: Identification of Hydrophobic Proteins Attached to Eucaryote Keratinocyte Cytoplasmic Membrane Using Different Matrices in Concert. *Proteome Science*, 1.
- Good, I. 1995. On the Question of Silk in Pre-Han Eurasia. *Antiquity*, 69, 959-968.
- Good, I. The Archaeology of Early Silk. Textile Society of America 8th Biennial Symposium, 2002 Northampton, MA. Textile Society of America.
- Good, I. 2010. When East Met West: Interpretive Problems in Assessing Eurasian Contact and Exchange in Antiquity. *Archäologische Mitteilungen aus Iran und Turan*, 42, 23-45.
- Good, I. 2011. Strands of Connectivity: Assessing the Evidence for Long Distance Exchange of Silk in Later Prehistoric Eurasia. In: Wilkinson, T., Sherratt, S. & Bennet, J. (eds.) *Interweaving Worlds: Systemic Interactions in Eurasia, 7th to the 1st Millennium Bc*. Oxford, U.K.: Oxbow Books.
- Good, I. L., Kenoyer, J. M. & Meadow, R. H. 2009. New Evidence for Early Silk in the Indus Civilization. *Archaeometry*, 51, 457-466.
- Gordon, J. E. 1975. *Organic Chemistry of Electrolyte Solutions*, Wiley.
- Goto, Y., Tsukada, M. & Minoura, N. 1990. Molecular Cleavage of Silk Fibroin by an Aqueous Solution of Lithium Thiocyanate
Subtitle_in_Japanese. *The Journal of Sericultural Science of Japan*, 59, 402-409.
- Graves, P. R. & Haystead, T. A. 2002. Molecular Biologist's Guide to Proteomics. *Microbiol Mol Biol Rev*, 66, 39-63; table of contents.
- Gulati, A. N. 1961. Technical Reports on Archaeological Remains. In: Clutton-Brock, J., Mittre, V. & Gulati, A. N. (eds.) *Technical Reports on Archaeological Remains*. 03/01 ed. Poona: Deccan College.
- Gupta K, A., Mita, K., Arunkumar, K. P., *et al.* 2015. Molecular Architecture of Silk Fibroin of Indian Golden Silkmoth, *Antheraea Assama*. *Scientific Reports*, 5, 12706.

- Ha, J. S., Song, J. J., Cho, H. K., *et al.* 2006. Preparation of Soluble Silk Peptides by Food-Grade Proteinases. *Korean J. Microbil. Biotechnol.*, 34, 115-120.
- Ha, S.-W., Park, Y. H. & Hudson, S. M. 2003. Dissolution of Bombyx Mori Silk Fibroin in the Calcium Nitrate Tetrahydrate-Methanol System and Aspects of Wet Spinning of Fibroin Solution. *Biomacromolecules*, 4, 488-496.
- Hakimi, O., Knight, D. P., Knight, M. M., *et al.* 2006. Ultrastructure of Insect and Spider Cocoon Silks. *Biomacromolecules*, 7, 2901-2908.
- Han and Tang Silk Network 2015. 丝绸文史:野蚕历史(Silk History: Wild Silkworm History). 都市丝绸 苏州通讯(Silk City Suzhou Communication). Suzhou, China: Suzhou Silk Industry Association.
- Hansen, V. 2015. *The Silk Road: A New History*, Oxford University Press.
- Hansen, V. & Rong, X. 2013. How the Residents of Turfan Used Textiles as Money, 273–796 Ce. *Journal of the Royal Asiatic Society of Great Britain & Ireland*, 23, 281-305.
- Hayes, J. 1977. Some Etruscan Textile Remains in the Royal Ontario Museum. *Studies in textile history*, 144-48.
- Hegyi, G., Kardos, J., Kovács, M., *et al.* 2013. 7.3. Polyacrylamide Gel Electrophoresis (Page). *Introduction to Practical Biochemistry*. Eötvös Loránd University.
- Hendy, J., Welker, F., Demarchi, B., *et al.* 2018. A Guide to Ancient Protein Studies. *Nature Ecology & Evolution*, 2, 791-799.
- Herrin, J. 2009. *Byzantium: The Surprising Life of a Medieval Empire*, Princeton University Press.
- Heyart, H. 1972. Textilreste. *Hémecht*, 24, 500-501.
- Higdon, R. & Kolker, E. 2007. A Predictive Model for Identifying Proteins by a Single Peptide Match. *Bioinformatics (Oxford, England)*, 23, 277-280.
- Hildebrandt, B. 2017. Silk Production and Trade in Roman Empire. In: Hildebrandt, B. & Gillis, C. (eds.) *Silk: Trade & Exchange Along the Silk Roads between Rome and China in Antiquity*. Oxford & Philadelphia: Oxbow Books.
- Hill, J. E. 2009. Appendix A: Introduction of Silk Cultivation to Khotan in the 1st century Ce. *Through the Jade Gate to Rome: A Study of the Silk Routes During the Later Han Dynasty 1st to 2nd Centuries Ce : An Annotated Translation of the Chronicle*

- on the 'Western Regions' in the Hou Hanshu. BookSurge Publishing.
- Hill, J. E. 2009. *Through the Jade Gate - China to Rome: A Study of the Silk Routes 1st to 2nd Centuries Ce: An Annotated Translation of the Chronicle on the 'Western Regions' in the Hou Hanshu*, BookSurge Publishing.
- Hirth, F. 1885. *China and the Roman Orient: Researches into Their Ancient and Mediaeval Relations as Represented in Old Chinese Records*, G. Hirth.
- Hollemeyer, K., Altmeyer, W. & Heinzle, E. 2002. Identification and Quantification of Feathers, Down, and Hair of Avian and Mammalian Origin Using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Anal Chem*, 74, 5960-8.
- Hollemeyer, K., Altmeyer, W., Heinzle, E., *et al.* 2008. Species Identification of Oetzi's Clothing with Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Based on Peptide Pattern Similarities of Hair Digests. *Rapid Commun Mass Spectrom*, 22, 2751-67.
- Hollemeyer, K., Altmeyer, W., Heinzle, E., *et al.* 2012. Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Combined with Multidimensional Scaling, Binary Hierarchical Cluster Tree and Selected Diagnostic Masses Improves Species Identification of Neolithic Keratin Sequences from Furs of the Tyrolean Iceman Oetzi. *Rapid Commun Mass Spectrom*, 26, 1735-45.
- Hollendonner, F. 1917. Der Stoff Eines Römischen Gewebes Von Aquincum / Az Aquincumi Ròmai Szövet Anyaga. *Botanische Mitteilungen / Botanikai Közlemények*, 16, 13-15.
- Hu, W., Zhang, H. & Zhang, B. 2015. Identification of Organic Binders in Ancient Chinese Paintings by Immunological Techniques. *Microsc Microanal*, 21, 1278-87.
- Hu, X., Kaplan, D. & Cebe, P. 2006. Determining Beta-Sheet Crystallinity in Fibrous Proteins by Thermal Analysis and Infrared Spectroscopy. *Macromolecules*, 39, 6161-6170.
- Hughes, C., Ma, B. & Lajoie, G. A. 2010. De Novo Sequencing Methods in Proteomics. *In: Hubbard, S. & Jones, A. K. G. (eds.) Proteome Bioinformatics*. Humana Press, a part of Springer Science+Business Media, LLC.
- Hundt, H. 1969. Über Vorgeschichtliche Seidenfunde. *Jahrbuch des RömischGermanischen Zentralmuseums*, 16, 65-71.
- Hundt, H. 1974. Gewebereste Aus Dem Fürstengrab Von Worms-Herrnsheim. *Jahrbuch des Römisch-Germanisches Zentralmuseum*, 18, 113-17.

- Hundt, H. 1985. Die Textilien Im Grab Von Hochdorf. *Der Keltenfürst von Hochdorf. Methoden und Ergebnisse der Landesarchäologie, Katalog zur Ausstellung Stuttgart, Kunstgebäude vom, 14*, 106-15.
- Hwang, J.-S., Lee, J.-S., Goo, T.-W., *et al.* 1999. Molecular Genetic Relationships between Bombycidae and Saturniidae Based on the Mitochondria DNA Encoding of Large and Small Rrna. *Genetic analysis: biomolecular engineering*, 15, 223-228.
- Iizuka, E. 1980. 絹糸の力学的性質 (the Mechanical Propertis of Silk Fibre). *In*: Hojo, N. (ed.) 続 絹糸の構造 (*Structure of Silk II*). Ueda:Japan: Shinshū University.
- Iizuka, E. 2000. Physical Properties of Silk Thread from Cocoons of Various Wild Silkmoths Including Domestic Silkmoth. *International Journal of Wild Silkmoth & Silk*, 5.
- Inoue, S., Tanaka, K., Arisaka, F., *et al.* 2000. Silk Fibroin of Bombyx Mori Is Secreted, Assembling a High Molecular Mass Elementary Unit Consisting of H-Chain, L-Chain, and P25, with a 6:6:1 Molar Ratio. *J Biol Chem*, 275, 40517-28.
- International Sericulture Commission. 2013. *Types of Silk* [Online]. Nagalore, India: International Sericultural Commission.
Available: http://inserco.org/en/types_of_silk [Accessed Nov. 4 2014].
- Issawi, C. P. 1982. *An Economic History of the Middle East and North Africa*, New York, Columbia University Press.
- Iupac 1997. Primary Structure, Secondary Structure, Tertiary Structure, Quaternary Structure. *In*: Mcnaught, A. D. & Wilkinson, A. (eds.) *Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*. XML on-line corrected version ed. Oxford: Blackwell Scientific Publications.
- Jabde, P. V. 2005. *Textbook of Applied Zoology*, Discovery Publishing House.
- Jacoby, D. 1992. Silk in Western Byzantium before the Fourth Crusade. *Byzantinische Zeitschrift*.
- Janaway, R. C. 1983. Textile Fibre Characteristics Preserved by Metal Corrosion: The Potential of S.E.M. Studies. *The Conservator*, 7, 48-52.
- Janaway, R. C. Corrosion Preserved Textile Evidence: Mechanism, Bias and Interpretation. *In*: Janaway, R. C. & Scott, B., eds. Evidence Preserved in Corrosion Products: New Field Studies in Artifact Studies: Proceedings of a joint conference between UKIC Archaeology Section and the Council for British Archaeology Science Committee, 1989 Leeds. United Kingdom Institute for Conservation, 21-29.

- Janaway, R. C. & Coningham, R. a. E. 1995. A Review of Archaeological Textile Evidence from South Asia. *South Asian Studies*, 11, 159-176.
- Jenkins, I. D. 1985. The Ambiguity of Greek Textiles. *Arethusa*, 18.
- Jeong, J. & Hur, W. 2010. Even-Numbered Peptides from a Papain Hydrolysate of Silk Fibroin. *J Chromatogr B Analyt Technol Biomed Life Sci*, 878, 836-40.
- Jin, H.-J. & Kaplan, D. L. 2003. Mechanism of Silk Processing in Insects and Spiders. *Nature*, 424, 1057-1061.
- Jo, Y.-Y., Bae, S. M. & Kweon, H. 2017. Dissolution of Degummed *Antheraea Yamamai* Silkworm Cocoon. *International Journal of Industrial Entomology*, 34, 6-10.
- Jo, Y.-Y., Kweon, H.-Y., Lee, K.-G., *et al.* 2011. The Promotion of Cell Attachment and Proliferation on Silk Fibroin. *Journal of Applied Biological Chemistry*, 54, 166-170.
- Jolly, M. S., Sen, S. K., Sonwalkar, T. N., *et al.* 1979. *Non-Mulberry Silks*, Rome, FAO Agricultural Service Bulletin.
- Jørgensen, L. B. 2018. Textiles from Mons Claudianus, ‘Abu Sha’ar and Other Roman Sites in the Eastern Desert. In: Brun, J., Faucher, T., Redon, B., *et al.* (eds.) *The Eastern Desert of Egypt During the Greco-Roman Period: Archaeological Reports*. Paris: Collège de France.
- Kafatos, F. C., Law, J. H. & Tartakoff, A. M. 1967. Cocoonase: II. Substrate Specificity, Inhibitors, and Classification of the Enzyme. *The Journal of Biological Chemistry*, 242, 1488-1494.
- Kafatos, F. C., Tartakoff, A. M. & Law, J. H. 1967. Cocoonase: I. Preliminary Characterization of a Proteolytic Enzyme from Silk Moths. *The Journal of Biological Chemistry*, 242, 1477-1487.
- Kang, P. K. 2004. A Study of the Silkworms and Mulberry Tree: *Nongshu*(農書) of Jiangnan in Qing Period. *Journal of Chinese Historical Researches*, 171-200.
- Kang, P. K. 2010. The Sericulture in Shandong Province in the Qing Dynasty: Focusing on *Yecanlu*. *Journal of Chinese historical researches*, 66, 25-50.
- Kaplan, D. & McGrath, K. 2012. *Protein-Based Materials*, Birkhäuser Boston.
- Kaplan, D. L., Fossey, S., Mello, C. M., *et al.* 1992. Biosynthesis and Processing of Silk Proteins. *MRS Bulletin*, 17, 41-47.
- Karmakar, S. R. 1999. Chemical Technology in the Pre-Treatment Processes of Textiles.

- Katz, J. R. & Derksen, J. C. 1931. Zur Erklärung Der Quellkraft Der Rhodanide Und Über Verbindungen Von Cellulose Und Neutralsalzen (Vorläufige Mitteilung) (Explaining the Swelling Power of Rhodanides and Compounds of Cellulose and Neutral Salts: Provisional Communication). *Recueil des Travaux Chimiques des Pays-Bas*, 50, 149-152.
- Kicman, A. T., Parkin, M. C. & Iles, R. K. 2007. An Introduction to Mass Spectrometry Based Proteomics—Detection and Characterization of Gonadotropins and Related Molecules. *Molecular and Cellular Endocrinology*, 260–262, 212-227.
- Kim, M. H. & Park, W. H. 2016. Chemically Cross-Linked Silk Fibroin Hydrogel with Enhanced Elastic Properties, Biodegradability, and Biocompatibility. *International journal of nanomedicine*, 11, 2967.
- Kim, S. C., Chen, Y., Mizra, S., *et al.* 2006. A Clean, More Efficient Method for in-Solution Digestion of Protein Mixtures without Detergent or Urea. *Journal of Proteome Research*, 5, 3446-3452.
- Knaller, R. & Ströbele, F. 2014. The Heritage of Tang Dynasty Textiles from the Famen Temple, Shaanxi, China: Technological and Stable Isotope Studies. *Studies in Conservation*, 59, S62-S65.
- Ko, J. S., Yoon, K., Ki, C. S., *et al.* 2013. Effect of Degumming Condition on the Solution Properties and Electrospinnability of Regenerated Silk Solution. *International journal of biological macromolecules*, 55, 161-168.
- Kuhn, D. 1984. Tracing a Chinese Legend: In Search of the Identity of the "First Sericulturalist". *T'oung Pao, 2nd Series*, 70, 213-245.
- Kuhn, D. 1984. Tracing a Chinese Legend: In Search of the Identity of the "First Sericulturalist". *T'oung Pao*, 70, 213-245.
- Kuhn, D., Chen, J., Huang, N., *et al.* 2012. *Chinese Silks*, Yale University Press.
- Kumar, H., Priya, Y. S., Kumar, M., *et al.* 2013. Effect of Different Mulberry Varieties and Seasons on Growth and Economic Traits of Bivoltine Silkworm (*Bombyx Mori*). *Journal of Entomology*, 10, 147-155.
- Kusuda, J., Tazima, Y., Onimaru, K., *et al.* 1986. The Sequence around the 5' End of the Fibroin Gene from the Wild Silkworm, *Bombyx Mandarina*, and Comparison with That of the Domesticated Species, *B. Mori*. *Molecular and General Genetics MGG*, 203, 359-364.
- Kweon, H., Lee, K. G., Lee, Y. W., *et al.* 1998. Hydrolysis Behaviour of *Antheraea Pernyi* Silk Fiber Treated with HCl. *Korean Journal of Sericultural Science*.

- Kweon, H., Lee, K. G. & Yeo, J. H. 2003. Dissolution of *Antheraea Pernyi* Silk Fiber and Structure of Regenerated Fibroin from Zinc Nitrate Solution. *Korean journal of sericultural science*, 45, 121-125.
- Kweon, H. & Park, Y. 2001. Dissolution and Characterization of Regenerated *Antheraea Pernyi* Silk Fibroin. *Journal of Applied Polymer Science*, 82, 750-758.
- Kyte, J. & Doolittle, R. F. 1982. A Simple Method for Displaying the Hydropathic Character of a Protein. *J Mol Biol*, 157, 105-32.
- Langer, W. L. 2001. *An Encyclopedia of World History; Ancient, Medieval, and Modern, Chronologically Arranged*, Boston, Houghton Mifflin Co.
- Latham, A. J. H. & Kawakatsu, H. 1994. *Japanese Industrialization and the Asian Economy*, Routledge.
- Laufer, B. 1915. The Story of the Pinna and the Syrian Lamb. *The Journal of American Folklore*, 28, 103-128.
- Lee, K. G., Jo, Y.-Y., Yeo, J.-H., *et al.* 2013. Dissolution of *Antheraea Pernyi* Raw Silk with Calcium Nitrate Ethanol Solution. *Korean Journal of Sericultural Science*, 51, 15-19.
- Lee, Y.-W. 1999. Silk Reeling and Testing Manual. *FAO Agricultural Service Bulletin No. 136* [Online]. Available: <http://www.fao.org/docrep/x2099e/x2099e00.htm#con> [Accessed 2015-5-27].
- Leggett, W. F. 1949. *The Story of Silk*, New York, Lifetime Editions.
- Leraut, P. & 2006. *Moths of Europe: Saturnids, Lasiocampids, Hawkmoths, Tiger Moths.*, Verrières le Buisson.
- Li, C. 1927. *西陰村史前的遺存* (Excavation report: the Xiyincun Prehistoric Site), 清華學校研究院(Chinghua University Research Institute).
- Li, G., Zhou, P., Shao, Z., *et al.* 2001. The Natural Silk Spinning Process. *The FEBS Journal*, 268, 6600-6606.
- Li, L., Gong, Y., Yin, H., *et al.* 2015. Different Types of Peptide Detected by Mass Spectrometry among Fresh Silk and Archaeological Silk Remains for Distinguishing Modern Contamination. *PLoS One*, 10, e0132827.
- Li, W., Zhang, Z., Lin, L., *et al.* 2017. *Antheraea Pernyi* (Lepidoptera: Saturniidae) and Its Importance in Sericulture, Food Consumption, and Traditional Chinese Medicine. *Journal of Economic Entomology*, 110, 1404-1411.

- Liu, M., Li, Y., Zheng, H., *et al.* 2015. Development of a Gold-Based Immunochromatographic Strip Assay for the Detection of Ancient Silk. *Analytical Methods*, 7, 7824-7830.
- Liu, M., Xie, J., Zheng, H., *et al.* 2015. Identification of Ancient Silk Using an Enzyme-Linked Immunosorbent Assay and Immuno-Fluorescence Microscopy. *Anal Sci*, 31, 1317-23.
- Lopez, R. S. 1945. Silk Industry in the Byzantine Empire. *Speculum*, 20, 1-42.
- Lordkipanidze, O. 2001. The Golden Fleece : Myth, Euhemeristic Explanation and Archaeology. *Oxford Journal of Archaeology*, 20, 1-38.
- Lotz, B. & Cesari, F. C. 1979. The Chemical Structure and the Crystalline Structures of Bombyx Mori Silk Fibroin. *Biochimie*, 61, 205-214.
- Lucas, F. & Rudall, K. M. 1968. Extracellular Fibrous Proteins: The Silks. . In: Florkin M., S. E. H., Stotz E.H (ed.) *Comprehensive Biochemistry*. Amsterdam: Elsevier.
- Lucas, F., Shaw, J. T. B. & Smith, S. G. 1960. Comparative Studies of Fibroins: I. The Amino Acid Composition of Various Fibroins and Its Significance in Relation to Their Crystal Structure and Taxonomy. *Journal of Molecular Biology*, 2, 339-349.
- Maeder, F. 2008. Sea-Silk in Aquincum: First Production Proof in Antiquity. In: Alfaro, C. & Karali, L. (eds.) *Vestidos, Textiles Y Tintes(Purpureae Vestes Ii): Estudios Sobre La Produccion De Bienes De Consumo En La Antigüedad* Valencia: UNIVERSITAT DE VALENCIA. SERVEI DE PUBLICACIONS.
- Maeder, F. Sea-Silk – the Rediscovery of the Ancient Textile Material Raises New Questions. In: Kim, A., ed. Exhibitions and Interpretation: Proceedings of the ICOM Costume Committee annual meeting, 8-13 September 2015 Toronto. ICOM.
- Magoshi, J. 1977. Physical Properties and Structure of Silk: 4. Spherulites Grown from Aqueous Solution of Silk Fibroin. *Polymer*, 18, 643-646.
- Magoshi, J., Magoshi, Y. & Nakamura, S. 1985. Crystallization, Liquid Crystal, and Fiber Formation of Silk Fibroin. *J. Appl. Polym. Sci*, 41, 187-204.
- Mahendran, B., Raman, N. & Kim, D. J. 2006. Purification and Characterization of Tannase from *Paecilomyces Variotii*: Hydrolysis of Tannic Acid Using Immobilized Tannase. *Applied Microbiology and Biotechnology*, 70, 444-450.
- Malay, A. D., Sato, R., Yazawa, K., *et al.* 2016. Relationships between Physical Properties and Sequence in Silkworm Silks. *Sci Rep*, 6, 27573.
- Mann, K. 1999. Aminosäurenanalysen. In: Banck-Burgess, J. (ed.) *Hochdotflv, Die*

- Textilfunde Aus Dem Späthallstattzeitlichen Fürstengrab Von Eberdingen-Hochdorf (Kreis Ludwigsburg) Und Weitere Grabtextilien Aus Hallstatt-Latènezeitlichen Kulturgruppen (Forschungen Und Berichte Zur Vbrund Frühgeschichte 70)*. Stuttgart: Konrad Theiss.
- Mann, M. & Wilm, M. 1994. Error-Tolerant Identification of Peptides in Sequence Databases by Peptide Sequence Tags. *Analytical Chemistry*, 66, 4390-4399.
- Margariti, C., Protopapas, S. & Orphanou, V. 2011. Recent Analyses of the Excavated Textile Find from Grave 35 Htr73, Kerameikos Cemetery, Athens, Greece. *Journal of Archaeological Science*, 38, 522-527.
- Marquardt , J., Mommsen , T., Dessau , H., *et al.* 1879. Das Privatleben Der Römer/ the Private Lif of the Romans. S. Hirzel.
- Marsh, R. E., Corey, R. B. & Pauling, L. 1955. An Investigation of the Structure of Silk Fibroin. *Biochimica et Biophysica acta*, 16, 1-34.
- Matsumoto, A., Kim, H. J., Tsai, I. Y., *et al.* 2006. Silk. *In*: Lewin, M. (ed.) *Handbook of Fiber Chemistry*. 3 ed. London: CRC Press.
- Matsumoto, A., Lindsay, A., Abedian, B., *et al.* 2008. Silk Fibroin Solution Properties Related to Assembly and Structure. *Macromol Biosci*, 8, 1006-18.
- Matthiesen, R., Rockwood, A. & Henneman, A. 2013. *Mass Spectrometry Data Analysis in Proteomics*, Humana Press.
- Mazzi, S., Zulker, E., Buchicchio, J., *et al.* 2014. Comparative Thermal Analysis of Eri, Mori, Muga, and Tussar Silk Cocoons and Fibroin Fibers. *Journal of Thermal Analysis and Calorimetry*, 116, 1337-1343.
- Mckinley, D. 1998. Pinna and Her Silken Beard: A Foray into Historical Misappropriations. *Ars Textrina*, 29, 9-223.
- Mechii, H. & Katagiri, K. 1991. Varietal Differences in Nutritive Values of Mulberry Leaves for Rearing Silkworms. *Japanese Agricultural Research*, Q, 202-208.
- Min, S., Kai, T., Teramoto, A., *et al.* 1997. Effects of Temperature and Humidity on the Solubility and Degradability of Fibroin Membrane. *J. Seric. Sci. Jpn.*, 66, 387-392.
- Murphy, A. R. & Romero, I. S. 2014. 8 - Biochemical and Biophysical Properties of Native Bombyx Mori Silk for Tissue Engineering Applications A2 - Kundu, S.C. *Silk Biomaterials for Tissue Engineering and Regenerative Medicine*. Woodhead Publishing.
- Muthesius, A. 1993. The Byzantine Silk Industry: Lopez and Beyond. *Journal of Medieval*

History, 19, 1-67.

Muthesius, A. 2007. Silk in the Medieval World. In: Jenkins, D. T. (ed.) *The Cambridge History of Western Textiles*. Cambridge, U.K.; New York: Cambridge University Press.

Muthesius, A. 2008. *Studies in Byzantine, Islamic and near Eastern Silk Weaving*, Pindar.

Nagele, E., Vollmer, M., Horth, P., *et al.* 2004. 2d-Lc/Ms Techniques for the Identification of Proteins in Highly Complex Mixtures. *Expert Rev Proteomics*, 1, 37-46.

Nagy, L. 1935. Aquincumi Mumia-Temetkezesek :Mumiengrăbnisse Aus Aquincum. In: Hungarici, M. N. (ed.) *Dissertationes Pannonicae* Budapest: Magyar Nemzeti Múzeum.

Nakazawa, Y. & Asakura, T. 2002. High-Resolution ¹³C Cp/Mas Nmr Study on Structure and Structural Transition of *Antheraea Pernyi* Silk Fibroin Containing Poly(L-Alanine) and Gly-Rich Regions. *Macromolecules*, 35, 2393-2400.

Nakazawa, Y. & Asakura, T. 2003. 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. *Journal of the American Chemical Society*, 125, 7231-7237.

Nakazawa, Y., Nakai, T., Kameda, T., *et al.* 1999. A ¹³C Nmr Study on the Structural Change of Silk Fibroin from *Samia Cynthia Ricini*. *Chemical physics letters*, 311, 362-366.

Narumi, T. 2004. ヤママユガ科の絹糸の形態的構造特性に関する研究 (Study on Morphological and Structural Properties of Silk Threads of Saturniidae). 農学博士論文(Doctoral dissertation for Agricultural Science), 東京大学 (Tokyo University).

Nelson, D. L., Nelson, D. L., Lehninger, A. L., *et al.* 2008. *Lehninger Principles of Biochemistry*, New York, W.H. Freeman.

Nunome, J. 1968. 山西省西陰村出土の仰韶期繭殻について(On the Yangshao culture cocoon found in Xiyin village, Shanxi province). *日本蚕糸学雑誌 (Journal of Sericology)*, 37, 187-194.

Ojha, N. G. & Panday, P. N. 2004. *Silk Production : Role of Feed on Tasar Silk and Egg Production*, New Delhi, APH Pub. Corp.

Olson, K. 2008. *Dress and the Roman Woman : Self-Presentation and Society*, Milton Park,

- Abingdon, Oxon; New York, Routledge.
- Palmieri, M., Vagnini, M., Pitzurra, L., *et al.* 2013. Identification of Animal Glue and Hen-Egg Yolk in Paintings by Use of Enzyme-Linked Immunosorbent Assay (Elisa). *Anal Bioanal Chem*, 405, 6365-71.
- Panagiotakopulu, E., Buckland, P. C., Day, P. M., *et al.* 1997. A Lepidopterous Cocoon from Thera and Evidence for Silk in the Aegean Bronze Age. *Antiquity*, 71, 420.
- Panda, H. 2010. *The Complete Book on Textile Processing and Silk Reeling Technology*, Delhi, Asia Pacific Business Press, Inc. : Distributed by Niir Project Consultancy Services.
- Paradiso, M. 2015. Chiara Vigo: The Last Woman Who Makes Sea Silk. *BBC News Magazine*. BBC.
- Park, Z. Y. & Russell, D. H. 2001. Identification of Individual Proteins in Complex Protein Mixtures by High-Resolution, High-Mass-Accuracy Maldi ToF-Mass Spectrometry Analysis of in-Solution Thermal Denaturation/Enzymatic Digestion. . *Analytical Chemistry*, 73, 2558-2564.
- Paszkiwicz, K. & Studholme, D. J. 2010. De Novo Assembly of Short Sequence Reads. *Brief Bioinform*, 11, 457-72.
- Patra, C., Talukdar, S., Novoyatleva, T., *et al.* 2012. Silk Protein Fibroin from *Antheraea Mylitta* for Cardiac Tissue Engineering. *Biomaterials*, 33, 2673-80.
- Pawcenis, D., Koperska, M. A., Milczarek, J. M., *et al.* 2013. Size Exclusion Chromatography for Analyses of Fibroin in Silk: Optimization of Sampling and Separation Conditions. *Applied Physics A*, 114, 301-308.
- Pearson, J. A., Buitenhuis, H., Hedges, R., *et al.* 2007. New Light on Early Caprine Herding Strategies from Isotope Analysis: A Case Study from Neolithic Anatolia. *Journal of Archaeological Science*, 34, 2170-2179.
- Peigler, R. S. 1993. Wild Silks of the World. *American Entomologist*, 39, 151-162.
- Pfister, R. 1934. *Textiles De Palmyre*, Paris, Les Éditions d'art et d'histoire.
- Pfister, R. 1937. *Nouveaux Textiles De Palmyre (Tour D'elahbel)*, Paris, Les Éditions d'art et d'histoire.
- Pfister, R. 1940. *Textiles De Palmyre*, Paris, Les Éditions d'art et d'histoire.
- Pittaway, A. R. 2018. *Saturnia Pyri (Denis & Schiffermüller, 1775) -- Great Peacock Moth* [Online]. Available: <http://tpittaway.tripod.com/silk/satlist.htm> [Accessed December 1 2016].

- Poli, G. S. & Delle Chiaie, S. 1792. *Testacea Utriusque Siciliae Eorumque Historia Et Anatome Tabulis Aeneis Illustrata*, Parmae, Ex regio typographeio.
- Pramanik, B. N., Mirza, U. A., Hain, Y., *et al.* 2002. Microwave-Enhanced Enzyme Reaction for Protein Mapping by Mass Spectrometry: A New Approach to Protein Digestion in Minutes. *Protein Science*, 11, 2676-2687.
- Prasong, S., Wilaiwan, S. & Yaowalak, S. 2011. Cross-Section Images of Eri (*Samia Ricini*)-Silk Fibers and Their Secondary Structures after Treatment with Different Organic Solvents. *Journal of Biological Sciences*, 11, 46-51.
- Putthanarat, S., Stribeck, N., Fossey, S., *et al.* 2000. Investigation of the Nanofibrils of Silk Fibers. *Polymer*, 41, 7735-7747.
- Qader, M. A., Haque, R. & Absar, N. 1992. Nutritative Effects of Different Types of Mulberry Leaves on Larval Growth and Cocoon Characteristics of *Bombyx Mori* L. . *Pakistan Journal of Zoology*, 24, 341-345.
- Rahmathulla, V. K. 2012. Management of Climatic Factors for Successful Silkworm (*Bombyx Mori* L.) Crop and Higher Silk Production: A Review. *Psyche: A Journal of Entomology*, 2012, 1-12.
- Rajkhowa, R. 2000. Structure Property Correlation of Non-Mulberry and Mulberry Silk Fibers. *International Journal of Wild Silkworm & Silk*, 5, 287-297.
- Ramesha, C., Kumari, S. S., Anuradha, C. M., *et al.* 2010. Nutrigenomic Analysis of Mulberry Silkworm (*Bombyx Mori* L.) Strains Using Polymerase Chain Reaction-Simple Sequence Repeats (Pcr-Ssr). *International Journal of Biotechnology and Molecular Biology Research*, 1, 92-100.
- Ramos, C. N. & Piegler, R. S. 1999. Comparative Ultrasturucture of Silk Fibers for Identifying Silk Textiles. *International Journal of Wild Silkworm & Silk*, 4.
- Rani, G. S. 2006. *Women in Sericulture*, Discovery Publishing House Pvt. Limited.
- Ravikumar, C. 1988. Western Ghat as a Bivoltine Region Prospects, Challenges and Strategies for Its Development. . *Indian Silk*, 26, 39-54.
- Ray, H. P. 2004. 2. Far-Flung Fabrics- Indian Textiles in Ancient Maritime Trade. In: Barnes, R. (ed.) *Textiles in Indian Ocean Societies*. Taylor & Francis.
- Reddy, K. V. & Prasad, G. R. 2010. *Non Mulberry Silkworms and Mulberry Rearing Requirements for the Course of Sericulture*, Hyderabad, India, Govt. of Andhra Pradesh.
- Richter, G. M. A. 1929. Silk in Greece. *American Journal of Archaeology*, 33, 27-33.

- Richter, K. K., Wilson, J., Jones, A. K. G., *et al.* 2011. Fish 'N Chips: Zooms Peptide Mass Fingerprinting in a 96 Well Plate Format to Identify Fish Bone Fragments. *Journal of Archaeological Science*, 38, 1502-1510.
- Riedesel, J. H. V. 1773. *Travels through Sicily and That Part of Italy Formerly Called Magna Græcia and a Tour through Egypt, with an Accurate Description of Its Cities, and the Modern State of the Country.* , London, Thomson Gale.
- Rockwood, D. N., Preda, R. C., Yucel, T., *et al.* 2011. Materials Fabrication from Bombyx Mori Silk Fibroin. *Nat. Protocols*, 6, 1612-1631.
- Royale, J. G. 1986. Silk Culture in India. *Indian Silk*, 25.
- Royle, J. F. 1840. *Essay on the Productive Resources of India*, W. H. Allen.
- Russell, W. K., Park, Z. Y. & Russell, D. H. 2001. Proteolysis in Mixed Organic-Aqueous Solvent Systems: Applications for Peptide Mass Mapping Using Mass Spectrometry. . *Analytical Chemistry*, 73, 2682-2685.
- Rutherford, S. M. & Gilani, G. S. 2009. Amino Acid Analysis. *Curr Protoc Protein Sci*, Chapter 11, Unit 11.9.
- Salles, J.-F. 1993. The Periplus of the Erythraean Sea and the Arab-Persian Gulf. *Topoi. Orient-Occident*, 493-523.
- Sashina, E., Bochek, A., Novoselov, N., *et al.* 2006. Structure and Solubility of Natural Silk Fibroin. *Russian Journal of Applied Chemistry*, 79, 869-876.
- Sashina, E., Novoselov, N. & Heinemann, K. 2003. Dissolution of Silk Fibroin in N-Methylmorpholine-N-Oxide and Its Mixtures with Organic Solvents. *Russian journal of applied chemistry*, 76, 128-131.
- Sashina, E. S., Bochek, A. M., Novoselov, N. P., *et al.* 2006. Structure and Solubility of Natural Silk Fibroin. *Russian Journal of Applied Chemistry*, 79, 869-876.
- Saurenhau, M. 1929. African Wild Silk and Its Uses. *The Melliand*, 1.
- Schafer, E. H. 1967. *The Vermilion Bird*, University of California Press.
- Scharfe, H. 1993. *Investigations in Kauṭalya's Manual of Political Science*, Harrassowitz.
- Schmidt-Colinet, A., As'ad, K. & Müting-Zimmer, C. 1992. *Das Tempelgrab Nr. 36 in Palmyra: Studien Zur Palmyrenischen Grabarchitektur Und Ihrer Ausstattung*, P. von Zabern.
- Schmidt-Colinet, A. & Gaballa, G. A. 1995. *Palmyra : Kulturbegegnung Im Grenzbereich*, Mainz am Rhein, Verlag Philipp von Zabern.

- Schmidt-Colinet, A., Stauffer, A. & As'ad, K. 2000. *Die Textilien Aus Palmyra*, Ph. von Zabern.
- Schoeser, M., Macdonald, J. & Marcandalli, B. 2007. *Silk*, New Haven, Conn. London, Yale University Press.
- Schoff, W. H. 1912. *The Periplus of the Erythræan Sea: Travel and Trade in the Indian Ocean*, Longmans, Green, and Company.
- Schrader, O. 1883. *Sprachvergleichung Und Urgeschichte: Linguistisch-Historische Beiträge Zur Erforschung Des Indogermanischen Altertums*, H. Costenoble.
- Scott, D. A., Warmlander, S., Mazurek, J., et al. 2009. Examination of Some Pigments, Grounds and Media from Egyptian Cartonnage Fragments in the Petrie Museum, University College London. *Journal of Archaeological Science*, 36, 923-932.
- Sen, K. & Babu, K. M. 2004. Studies on Indian Silk. I. Macrocharacterization and Analysis of Amino Acid Composition. *Journal of Applied Polymer Science*, 92, 1080-1097.
- Sheffer, A. & Granger-Taylor, H. 1994. Textiles from Masada. In: Barag, D. E. A. (ed.) *Masada Iv: The Yigael Yadin Excavations 1963-1965, Final Reports*. Jerusalem: Israel Exploration Society, The Hebrew University.
- Sherratt, A. G. 2001. The Transformation of Early Agrarian Europe: The Later Neolithic and Copper Ages 4500-2500 Bc. In: Cunliffe, B. W. (ed.) *The Oxford Illustrated History of Prehistoric Europe*. Oxford: Oxford University Press.
- Shimura, K. 1983. Chemical Composition and Biosynthesis of Silk Proteins. *Experientia*, 39, 455-461.
- Shimura, K. & Hidaka, Y. 1980. フィブロインの化学構造 (the Chemical Structure of Fibroin). In: Hojo, N. (ed.) 続. 絹糸の構造 (*Structure of Silk II*). Ueda:Japan: Shinshū University.
- Šiletić, T. 2004. Die Edle Steckmuschel: *Pinna Nobilis* (Linnaeus, 1758) / *La Conchiglia Nobile*:
- Pinna Nobilis* (Linnaeus, 1758). In: Maeder, F., Hänggi, A. & Wunderlin, D. (eds.) *Exhibition Catalogue: Muschelseide – Goldene Fäden Vom Meeresgrund/Bisso Marino - Fili D'oro Dal Fondo Del Mare*. Milano: 5 Continents Editions.
- Silver, M. 1992. *Taking Ancient Mythology Economically*, Brill.
- Simmonds, P. L. 1883. *The Commercial Products of the Sea;or Marine Contributions to Food,Industry, and Art*, London; New York, Griffith and Farran; E. P. Dutton & co.

- Simmons, A. H., Michal, C. A. & Jelinski, L. W. 1996. Molecular Orientation and Two-Component Nature of the Crystalline Fraction of Spider Dragline Silk. *Science*, 271, 84-87.
- Singh, R. & Sachan, G. C. 2007. Elements of Entomology.
- Singhal, N., Kumar, M., Kanaujia, P. K., *et al.* 2015. MALDI-TOF Mass Spectrometry: An Emerging Technology for Microbial Identification and Diagnosis. *Frontiers in microbiology*, 6, 791-791.
- Sinha, A. K., Sinha, R., Goel, A., *et al.* 1994. A Review on the Breeding and Genetic Aspects of Tropical Tasar Silkworm, *Antheraea Mylitta D.* *Proceedings of the Conference on Cytology and Genetics*, 4, 7-16.
- Smith, W., Sir., Wayte, W., Yates, J., *et al.* 1890. *A Dictionary of Greek and Roman Antiquities.*, Albermarle Street, London, John Murray.
- Smith, W., Wayte, W. & Marindin, G. E. 1890. Σεΐρικόν, Βομβύκια). *In: Anthon, C. (ed.) A Dictionary of Greek and Roman Antiquities.* London: John Murray.
- Sofia, S., McCarthy, M. B., Gronowicz, G., *et al.* 2001. Functionalized Silk-Based Biomaterials for Bone Formation. *Journal of Biomedical materials research*, 54, 139-148.
- Solazzo, C., Dyer, J. M., Deb-Choudhury, S., *et al.* 2012. Proteomic Profiling of the Photo-Oxidation of Silk Fibroin: Implications for Historic Tin-Weighted Silk. *Photochem Photobiol*, 88, 1217-26.
- Solazzo, C., Heald, S., Ballard, M. W., *et al.* 2011. Proteomics and Coast Salish Blankets: A Tale of Shaggy Dogs? *Antiquity*, 85, 1418-1432.
- Solazzo, C., Rogers, P. W., Weber, L., *et al.* 2014. Species Identification by Peptide Mass Fingerprinting (Pmf) in Fibre Products Preserved by Association with Copper-Alloy Artefacts. *Journal of Archaeological Science*, 49, 524-535.
- Srisuwan, Y. & Srihanam, P. 2009. Dissolution of Philosamia Ricini Silk Film: Properties and Functions in Different Solutions. *Journal of Applied Sciences*, 9, 978-982.
- Stauffer, A. 2000. The Textiles from Palmyra: Technical Analyses and Their Evidence for Archaeological Research. *In: Cardon, D. & Feugère, M. (eds.) Archéologie des textiles des origines au Ve siècle, Actes du colloque de Lattes.* Montpellier, France.
- Steen, H. & Mann, M. 2004. The Abc's (and Xyz's) of Peptide Sequencing. *Nat Rev Mol Cell Biol*, 5, 699-711.
- Steiger, R. E. 1930. The Formylation of Amino Acids. *Journal of Biological Chemistry*, 86,

695-702.

- Stein, M. A. 1907. *Ancient Khotan: Detailed Report of Archaeological Explorations in Chinese Turkestan*, Clarendon Press.
- Stoneman, R. 1994. *Palmyra and Its Empire: Zenobia's Revolt against Rome*, University of Michigan Press.
- Sutherland, T. D., Young, J. H., Weisman, S., *et al.* 2010. Insect Silk: One Name, Many Materials. *Annu Rev Entomol*, 55, 171-88.
- Sylwan, V. 1937. Silk from the Yin Dynasty. *Bulletin of the Museum of Far Eastern Antiquities*, 9, 119-126.
- Sylwan, V. Investigation of Silk from Edsen-Gol and Lop-Nor. Sino-Swedish Expedition (1927-1935). 1949.
- Tamura, T. & Kubota, T. 1988. Ultrastructure of Silk Glands and Cocoon Filaments of Wild Silkmoths, *Antheraea Yamamai* and *Antheraea Pernyi*. *Wild Silkmoths'88*, 67.
- Tamura, T. & Sakate, S. 1988. Isolation and Characterization of Fibroin Mrnas from the Saturniid Silkworms, *Antheraea Yamamai*, *Antheraea Pernyi* and *Philosamia Cynthia Ricini*. *Insect biochemistry*, 18, 169-175.
- Tanaka, C. & Asakura, T. 2009. Synthesis and Characterization of Cell-Adhesive Silk-Like Proteins Constructed from the Sequences of Anaphe Silk Fibroin and Fibronectin. *Biomacromolecules*, 10, 923-8.
- Tanaka, K., Inoue, S. & Mizuno, S. 1999. Hydrophobic Interaction of P25, Containing Asn-Linked Oligosaccharide Chains, with the H-L Complex of Silk Fibroin Produced by *Bombyx Mori*. *Insect Biochem Mol Biol*, 29, 269-76.
- Tanaka, K. & Mizuno, S. 2001. Homologues of Fibroin L-Chain and P25 of *Bombyx Mori* Are Present in *Dendrolimus Spectabilis* and *Papilio Xuthus* but Not Detectable in *Antheraea Yamamai*. *Insect Biochemistry and Molecular Biology*, 31, 665-677.
- Tanaka, T., Magoshi, J., Magoshi, Y., *et al.* 2002. Thermal Properties of *Bombyx Mori* and Several Wild Silkworm Silks: Phase Transition of Liquid Silk. *Journal of Thermal Analysis and Calorimetry*, 70, 825-832.
- Tao, W., Li, M. & Zhao, C. 2007. Structure and Properties of Regenerated *Antheraea Pernyi* Silk Fibroin in Aqueous Solution. *Int J Biol Macromol*, 40, 472-8.
- Tarn, W. W. 2010. *The Greeks in Bactria and India*, Cambridge University Press.
- Tashiro, Y., Otsuki, E. & Shimadzu, T. 1972. Sedimentation Analyses of Native Silk Fibroin in Urea and Guanidine·Hcl. *Biochimica et Biophysica Acta (BBA) - Protein*

Structure, 257, 198-209.

The Editors of Encyclopædia Britannica. Silk. *Encyclopædia Britannica* [Online]. Available: <https://www.britannica.com/topic/silk>.

Thill, G. Altrier. In: Duval, A. & Eluère, C., eds. Trésors des princes celtes, Galeries nationales du Grand Palais, 20 October, 1987 1987 Paris. Réunion des musées nationaux, 251-54.

Tikader, A., Vijayan, K. & Saratchandra, B. 2013. Muga Silkworm, *Antheraea Assamensis* (Lepidoptera: Saturniidae) - an Overview of Distribution, Biology and Breeding. *EJE*, 110, 293-300.

Trevisan, A. A., Hart, M., Yuste, V., *et al.* Cocoon Silk: A Natural Architecture. *Sense of Nature*.

Tse, S. & Dupont, A.-L. 2001. Measuring Silk Deterioration by High-Performance Size-Exclusion Chromatography, Viscometry, and Electrophoresis. ACS Publications.

Tsukada, M. 1986. Structural Changes Induced in Tussah Silk (*Antheraea Pernyi*) Fibroin Films by Immersion in Methanol. *Journal of Polymer Science Part B: Polymer Physics*, 24, 1227-1232.

Tsukada, M., Freddi, G., Gotoh, Y., *et al.* 1994. Physical and Chemical Properties of Tussah Silk Fibroin Films. *Journal of Polymer Science Part B: Polymer Physics*, 32, 1407-1412.

Tsukada, M., Freddi, G., Nagura, M., *et al.* 1992. Structural Changes of Silk Fibers Induced by Heat Treatment. *APP Journal of Applied Polymer Science*, 46, 1945-1953.

Turner, R. D. & Rosewater, J. 1958. The Family Pinnidae in the Western Atlantic. *Johnsonia*, 3, 285-326.

U.S. Geological Survey 2005-2011. Butterflies and Moths of North America. National Biological Information Infrastructure (NBII) Program and the USGS Northern Prairie Wildlife Research Center.

Um, I.-C. & Park, Y.-H. 1998. Morphology of Silk Fibroin/Poly(Vinyl Alcohol) Blend Film. *Korean Journal of Seric. Sci.*, 40, 169-175.

Um, I. C., Kweon, H., Lee, K. G., *et al.* 2004. Wet Spinning of Silk Polymer: I. Effect of Coagulation Conditions on the Morphological Feature of Filament. *International Journal of Biological Macromolecules*, 34, 89-105.

Usher, A. P. 2013. *A History of Mechanical Inventions: Revised Edition*, Dover Publications.

- Vanden Berghe, I. & Wouters, J. Identification and Condition Evaluation of Deteriorated Protein Fibres at the Sub-Microgram Level by Calibrated Amino Acid Analysis. Scientific analysis of ancient and historic textiles: informing preservation, display and interpretation: postprints, first annual conference, 13-15 July 2004, 2004. Archetype Publications, 151-158.
- Vijayan, K., Nair, C. & Urs, S. R. 2010. Assessment of Genetic Diversity in the Tropical Mulberry Silkworm (*Bombyx Mori* L.) with Mtdna-Sscp and Ssr Markers. *Emirates Journal of Food and Agriculture*, 22, 71.
- Vollrath, F. & Knight, D. P. 2001. Liquid Crystalline Spinning of Spider Silk. *Nature*, 410, 541.
- Wang, X. 2000. Production and Physical Properties of Soluble Wild Silk Powder. *Int. J. Wild Silkmoth & Silk*, 5, 57-60.
- Wang, Y. 2002. *野蠶錄*(*Ye Can Lu*), Shanghai, 上海古籍出版社 (Shanghai gu ji chu ban she).
- Wardle, T. S. 1881. Handbook of the Collection Illustrative of the Wild Silks of India, in the Indian Section of the South Kensington Museum, with a Catalogue of the Collection. London: Victoria and Albert Museum.
- Wendelken, R. W. 2014. Wefts and Worms: The Spread of Sericulture and Silk Weaving in the West before 1300. In: Netherton, R. & Owen-Crocker, G. R. (eds.) *Medieval Clothing and Textiles*. Boydell Press.
- Whitmore, F. C. 2012. *Organic Chemistry, Volume One: Part I: Aliphatic Compounds Part II: Alicyclic Compounds*, Dover Publications.
- Wild, J. P. 1970. *Textile Manufacture in the Northern Roman Provinces*, Cambridge University Press.
- Wild, J. P. 1984. Some Early Silk finds in Northwest Europe. *The Textile Museum Journal*, 23, 17-23.
- Wilkins, M. R., Lindskog, I., Gasteiger, E., *et al.* 1997. Detailed Peptide Characterization Using Peptidomass—a World-Wide-Web-Accessible Tool. *Electrophoresis*, 18, 403-408.
- Willis, J. H., Wilkins, A. S. & Goldsmith, M. R. 1995. A Brief History of Lepidoptera as Model Systems. *Molecular model systems in the Lepidoptera*, 1-20.
- Wiśniewski, J. R., Zougman, A., Nagaraj, N., *et al.* 2009. Universal Sample Preparation Method for Proteome Analysis. *Nature Methods*, 6, 359.

- Xie, M., Shevchenko, A., Wang, B., *et al.* 2016. Identification of a Dairy Product in the Grass Woven Basket from Gumugou Cemetery (3800 bp, Northwestern China). *Quaternary International*.
- Yamada, H. & Kato, Y. 2004. Green Coloration of Cocoons in *Antheraea Yamamai* (Lepidoptera : Saturniidae): Light-Induced Production of Blue Bilin in the Larval Hemolymph. *Journal of Insect Physiology*, 5, 393-401.
- Yang, H., Zhang, Y. & Pöschl, U. 2010. Quantification of Nitrotyrosine in Nitrated Proteins. *Analytical and bioanalytical chemistry*, 397, 879-886.
- Yang, J. C., Lee, S. Y., Chen, C. C., *et al.* 2011. *High Yield Dialysis-Free Process for Producing Organosoluble Regenerated Silk Fibroin*.
- Yarshater, E. 1983. *The Cambridge History of Iran: Seleucid Parthian*, Cambridge University Press.
- Yates, J. 1843. *Textrinum Antiquorum: An Account of the Art of Weaving among the Ancients : With an Appendix on the Period of the Invention of Linen Paper, on Feltin, on Netting, on the Onomasticon of Julius Pollux*, Taylor & Walton.
- Yoon, H. J., Kang, P. D., Kim, S. E., *et al.* 2006. Breeding of Major Characteristics of the Wild Silkworm, *Antheraea Yamamai* Indoor-and Outdoor-Reared. *Korea J. Seric. Sci.*, 48, 61-67.
- Yoshitake, N. 1968. Phylogenetic Aspects on the Origin of Japanese Race of the Silkworm, *Bombyx Mori* L. *J. Sericol. Sci. Jpn*, 37, 83-7.
- Yukuhiro, K., Kanda, T. & Tamura, T. 1997. Preferential Codon Usage and Two Types of Repetitive Motifs in the Fibroin Gene of the Chinese Oak Silkworm, *Antheraea Pernyi*. *Insect Molecular Biology*, 6, 89-95.
- Yukuhiro, K., Sezutsu, H., Itoh, M., *et al.* 2002. Significant Levels of Sequence Divergence and Gene Rearrangements Have Occurred between the Mitochondrial Genomes of the Wild Mulberry Silkworm, *Bombyx Mandarina*, and Its Close Relative, the Domesticated Silkworm, *Bombyx Mori* *Molecular Biology and Evolution*, 19, 1385–1389.
- Zannoon, A. H. A., E.M.M. Hassan, S.S. Ei-Akkad, *et al.* 2012. Biological and Technol. Effects of Mulberry Varieties and Nutritional Additives on Silkworm *Bombyx Mori* Development. *Egyptian Journal of Biology*, 10, 1-10.
- Zethner, O. & Koustrup, R. 2007. *Silkens Veje : Gamle Tråde Og Nye Netværk*, Farum; [Haslev], Koustrup & Co. ; Ekspedition: Nordisk Bog Center.
- Zethner, O., Koustrup, R. & Raina, S. K. 2008. *African Ways of Silk: Ancient Threads-New*

Possibilities Cape town, Centre for Advanced Studies of African Society.

- Zhang, F., Lu, Q., Yue, X., *et al.* 2015. Regeneration of High-Quality Silk Fibroin Fiber by Wet Spinning from CaCl₂–Formic Acid Solvent. *Acta Biomaterialia*, 12, 139-145.
- Zhang, F., Zuo, B., Fan, Z., *et al.* 2012. Mechanisms and Control of Silk-Based Electrospinning. *Biomacromolecules*, 13, 798-804.
- Zhang, G., Annan, R. S., Carr, S. A., *et al.* 2010. Overview of Peptide and Protein Analysis by Mass Spectrometry. *Curr Protoc Protein Sci*, Chapter 16, Unit16.1.
- Zhang, J., Kaur, J., Rajkhowa, R., *et al.* 2013. Mechanical Properties and Structure of Silkworm Cocoons: A Comparative Study of Bombyx Mori, Antheraea Assamensis, Antheraea Pernyi and Antheraea Mylitta Silkworm Cocoons. *Mater Sci Eng C Mater Biol Appl*, 33, 3206-13.
- Zhang, S. & Basile, F. 2007. Site-Specific Pyrolysis-Induced Cleavage at Aspartic Acid Residue in Peptides and Proteins. *Journal of proteome research*, 6, 1700-1704.
- Zhang, X. C. 2016. *Science and Principles of Biodegradable and Bioresorbable Medical Polymers: Materials and Properties*, Elsevier Science.
- Zhao, C., Wu, X., Zhang, Q., *et al.* 2011. Enzymatic Degradation of Antheraea Pernyi Silk Fibroin 3d Scaffolds and Fibers. *Int J Biol Macromol*, 48, 249-55.
- Zhao, F. 1998. Art of Silk and Art on Silk in China. In: Lee, S. E., Museum, S. R. G. & Bilbao, M. G. (eds.) *China, 5000 Years: Innovation and Transformation in the Arts*. Solomon R. Guggenheim Museum.
- Zhao, F. 1999. *Treasures in Silk : An Illustrated History of Chinese Textiles*, Hong Kong, ISAT/Costume Squad.
- Zhao, Y. & Lin, Y. H. 2010. Whole-Cell Protein Identification Using the Concept of Unique Peptides. *Genomics Proteomics Bioinformatics*, 8, 33-41.
- Zheng, Q., Wu, X., Zheng, H., *et al.* 2015. Development of an Enzyme-Linked-Immunosorbent-Assay Technique for Accurate Identification of Poorly Preserved Silks Unearthed in Ancient Tombs. *Anal Bioanal Chem*, 407, 3861-7.
- Zheng, Z., Guo, S., Liu, Y., *et al.* 2016. Lithium-Free Processing of Silk Fibroin. *J Biomater Appl*, 31, 450-63.
- Zheng, Z. & Mo, Y. 2002. *柶繭譜*(*Chu Jian Pu*), Shanghai, 上海古籍出版社(Shanghai gu ji publishing).
- Zhou, C.-Z., Confalonieri, F., Medina, N., *et al.* 2000. Fine Organization of Bombyx Mori

- Fibroin Heavy Chain Gene. *Nucleic acids research*, 28, 2413-2419.
- Zhou, C. Z., Confalonieri, F., Jacquet, M., *et al.* 2001. Silk Fibroin: Structural Implications of a Remarkable Amino Acid Sequence. *Proteins: Structure, Function, and Bioinformatics*, 44, 119-122.
- Zhu, Z., Chen, H.-F., Li, L., *et al.* 2014. Biomass Spectrometry Identification of the Fibre Material in the Pall Imprint Excavated from Grave M1, Peng-State Cemetery, Shanxi, China. *Archaeometry*, 56, 681-688.
- Zolotuhin, V. & Gurkovich, A. 2009. A Review of the Genus *Pachypasa* Walker, 1865, *Sensu Lato* in Africa (Lepidoptera, Lasiocampidae). *Neue Entomol. Nachr*, 63, 1-75.
- Żuchowska, M. 2013. Palmyra and the Far Eastern Trade. *Studia Palmyreńskie 12: Fifty Years of Polish Excavations in Palmyra 1959–2009*. Warsaw, Poland: Polish Centre of Mediterranean Archaeology, University of Warsaw.