
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS
BIOLOGIA CELULAR E MOLECULAR

Análise Estrutural das Proteínas da Seda da Teia da Aranha
Nephila clavipes **por uma Abordagem Proteômica**

JOSÉ ROBERTO APARECIDO DOS SANTOS PINTO

Tese apresentada ao Instituto de Biociências do Câmpus de Rio Claro, Universidade Estadual Paulista, como parte dos requisitos para obtenção do título de Doutor em Biologia Celular e Molecular.

Rio Claro
São Paulo - Brasil
Março-2014

JOSÉ ROBERTO APARECIDO DOS SANTOS PINTO

**ANÁLISE ESTRUTURAL DAS PROTEÍNAS DA SEDA DA
TEIA DA ARANHA *Nephila clavipes* POR UMA
ABORDAGEM PROTEÔMICA**

Orientador: Prof. Dr. Mario Sergio Palma

Tese apresentada ao Instituto de Biociências
do Câmpus de Rio Claro, Universidade
Estadual Paulista “Júlio de Mesquita Filho”,
como parte dos requisitos para obtenção do
título de Doutor em Biologia Celular e
Molecular.

**RIO CLARO
2014**

CERTIFICADO DE APROVAÇÃO

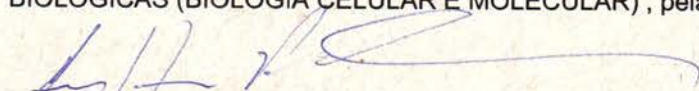
TÍTULO: Análise Estrutural das Proteínas da Seda da Teia da Aranha *Nephila clavipes* por uma Abordagem Proteômica

AUTOR: JOSÉ ROBERTO APARECIDO DOS SANTOS PINTO

ORIENTADOR: Prof. Dr. MARIO SERGIO PALMA

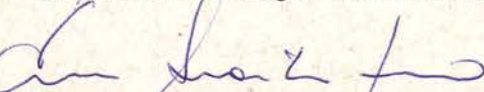
CO-ORIENTADORA: Profa. Dra. LUCILENE DELAZARI DOS SANTOS

Aprovado como parte das exigências para obtenção do Título de DOUTOR EM CIÊNCIAS BIOLÓGICAS (BIOLOGIA CELULAR E MOLECULAR), pela Comissão Examinadora:



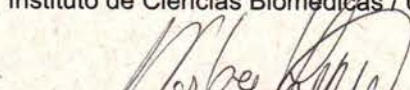
Prof. Dr. MARIO SERGIO PALMA

Departamento de Biologia / Instituto de Biociências de Rio Claro



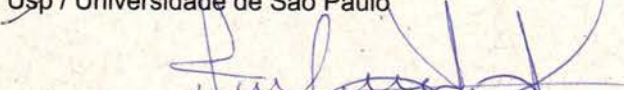
Prof. Dr. EMER SUAVINHO FERRO

Instituto de Ciências Biomédicas / Universidade de São Paulo



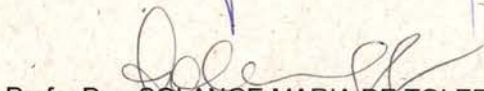
Prof. Dr. NORBERTO PEÇORINE LOPES

Departamento de Física e Química Da Faculdade de Ciências Farmacêuticas de Ribeirão Preto-
Usp / Universidade de São Paulo



Prof. Dr. JOSÉ CÉSAR ROSA

Departamento de Biologia Celular Molecular e Bioagentes Patogênicos / USP, Faculdade de
Medicina de Ribeirão Preto



Profa. Dra. SOLANGE MARIA DE TOLEDO SERRANO

Laboratório Especial de Toxinologia Aplicada / Instituto Butantan

Data da realização: 25 de março de 2014.

*À minha família e aos meus
amigos, dedico este trabalho.*

Decidi que...
NÃO VOU MAIS:

Censurar os erros do passado...
Passarei minha vida a limpo e farei dos erros meu **aprendizado**.
Paciência... Continuarei errando, porque só erra quem faz.

Criticar ou tentar modificar alguém...
Mudarei apenas minhas **atitudes**. Vou mudar para melhor.
Sabe por quê? Eu ganho, e quem me quer bem também.

Abrir mão de meus sonhos...
Lutarei sempre para que eles aconteçam. Mesmo que eu tropece.
Vencerei meu **medo** e continuarei correndo atrás do melhor para mim.

Colocar minha felicidade nas mãos de alguém...
Vou ser **feliz** de qualquer jeito. De preferência, amando quem me ama,
têm preocupações comigo e torce para me ver feliz.

Ficar com pena de mim pelos meus problemas e tropeços...
Agradecerei a Deus por me fazer **forte** e por ter sobrevivido a eles.
Lembrarei sempre que dificuldade, em vez de castigo, é o maior estímulo à minha **criatividade**.

Ser pessimista e pensar no pior...
Ocuparei meu tempo com algo útil, que me divirta ou que possa ajudar alguém.
Vou esperar pelo melhor, valorizar as **amizades** e cada momento em minha vida.

Querer ser exemplo de perfeição ou copiar o dos outros...
Vou me aceitar como sou, e dar o melhor de mim em tudo que fizer.
Acima de tudo **amar** e ser amado.

Decidi que...
A partir de hoje:

Vou viver plenamente...
Poucos são os que se dispõem a abrir mão de uma vida segura, para entrar no desconhecido caminho da **autodescoberta**, aceitando ser fiéis acima de tudo à sua própria verdade.

Mas os que ousaram arriscar e se dispuseram a ir além, apesar do medo e dos obstáculos que tiveram de atravessar, certamente têm hoje a certeza de que **valeu a pena...**

Autor desconhecido (Modificado por J.R.)

“Feeling my way through the darkness
Guided by a beating heart
I can't tell where the journey will end
But I know where to start

They tell me I'm too young to understand
They say I'm caught up in a dream
Well, life will pass me by if I don't open up my eyes
Well, that's fine by me

So wake me up when it's all over
When I'm wiser and I'm older
All this time I was finding myself
And I didn't know I was lost

So wake me up when it's all over
When I'm wiser and I'm older
All this time I was finding myself
And I didn't know I was lost

I tried carrying the weight of the world
But I only have two hands
I hope I get the chance to travel the world
But I don't have any plans
I wish that I could stay forever this young
Not afraid to close my eyes
Life's a game made for everyone
And love is a prize

So wake me up when it's all over
When I'm wiser and I'm older
All this time I was finding myself
And I didn't know I was lost

So wake me up when it's all over
When I'm wiser and I'm older
All this time I was finding myself
And I didn't know I was lost

I didn't know I was lost
I didn't know I was lost”

Wake Me Up (feat. Aloe Blacc) Avicii

“Pedras no caminho?
Guardo todas, um dia vou
construir um castelo ...”

Fernando Pessoa

“Se o experimento planejado não
funcionar, não se preocupe,
comece a planejar o próximo.”

Frederick Sanger

AGRADECIMENTOS

À vida, à ciência e a toda e qualquer força superior que nos orienta e fortalece no dia a dia na jornada que escolhemos seguir e/ou para qual somos guiados...

Ao meu orientador, Prof. Dr. Mario Sergio Palma, por todo o conhecimento científico, pela confiança, pela paciência em compreender que cada um de nós, mesmo profissionalmente, possuímos um tempo diferente de amadurecimento e também por compreender que é justamente isso que caracteriza e diferencia o trabalho em equipe em um grupo de pesquisa; pelas oportunidades proporcionadas, pelo exemplo de postura profissional e dedicação à pesquisa... pelos ensinamentos “não científicos”, pelos desabafos, pela amizade e por se permitir me conhecer um pouco mais e por permitir-se conhecer um pouco mais, pelas histórias... que não são poucas, e também pelas conversas sérias e “brincas” quando necessárias... e por aquela “pressão básica”.

Ao Prof. Dr. Gert Lubec do Laboratório de Proteômica Gert Lubec da Universidade de Medicina de Viena, Viena, Áustria, por toda a colaboração, auxílio e dedicação na realização deste estudo... pela atenção e por ter me recebido tão bem não só em seu laboratório, mas também no seu convívio compartilhando das suas experiências de vida e científica, da sua amizade e das suas ideias. Posso dizer que foi sem dúvida uma incrível e agradável experiência profissional e pessoal ter trabalhado em seu laboratório, e ter convivido com o seu grupo de pesquisa.

À Dra. Helen Andrade Arcuri, a Heleeeennnn!!! por toda a colaboração e dedicação para a realização deste estudo... pelas aventuras na bioinformática estrutural, pela paciência em entender os nossos dados experimentais e complementa-los com os dados oriundos da modelagem e dinâmica molecular, engrandecendo os nossos resultados... pela grande amizade, que mesmo quando distantes se manteve, por acreditar em mim e me apoiar, pelas conversas, por gostar de conversar... e por continuar sendo essa pessoa espontânea e transparente... SPIDROIN!!!

À minha mãe Benedita, por ser o meu exemplo de caráter, perseverança, força, dignidade, fé... por tudo o que já passamos e superamos, por me apoiar e compreender a necessidade de me manter distante, por me ouvir, pelos afagos e por sempre demonstrar que mesmo diante das dificuldades, sejam elas de qualquer natureza, devemos ser persistentes e não desistirmos.

Ao meu pai José Vitor, às minhas queridas irmãs Rosimeire e Fernanda, meus cunhados José Roberto e Dirceu, e meus sobrinhos Henrique, Matheus, Bianca e Jhenifer, por todo o carinho, apoio, por torcerem por mim e por compreenderem a minha ausência e essa necessidade de busca por conhecimentos e algo melhor.

A todos do Laboratório de Proteômica Gert Lubec da Universidade de Medicina de Viena, Viena, Áustria, por terem me recebido tão bem, por toda a troca de experiência profissional e pessoal... por toda a ajuda e amizade que me proporcionaram um bem estar durante a realização deste trabalho, por todas as conversas científicas e não científicas, pelas trocas culturais e de costumes com diferentes pessoas, de diferentes locais, proporcionadas em uma única cidade, Viena!... Sonia, Fernando, Minu, Sunetra, Ajinkya, Cláudia, Ana, Maryna, Vitali, Siva, Saras, Priya, Andras, Edit, Robert, Lin, Sophie (Wan Jia), Wei-Qiang, Seok, Gangsoo, Narkhyun, Eun-Jung, Fredi, Maryam, Soheil, Bharani, e Jana... obrigado a todos pelo convívio e pela amizade; foi um prazer conhece-los e compartilhar de um tempo e espaço com todos vocês... espero um dia poder revê-los novamente. See you guys!!!

A todos do Laboratório de Biologia Estrutural e Zooquímica, pelo trabalho em grupo, pelas trocas de experiências profissionais e pessoais... à Nathália, Marcel, Anally, Paulo, Bibiana, Kenny, Eduardo, Marcela, Fernando, Diego... à Ana Maria por toda a ajuda nas coletas e preparação de amostras, pela amizade, pelo apoio e companheirismo... à Franciele pela grande ajuda na correção das referências bibliográficas e listas de tabelas, figuras e abreviações... pelo apoio, atenção e amizade.

Ao Prof. Dr. Günther Lamprecht do Instituto de Química Analítica da Universidade de Medicina de Viena, Viena, Áustria pelas análises de ditirosina nos hidrolisados da seda; e também à Profa. Dra. Helga Priewalder do Departamento de Paleontologia, Pesquisa Geológica de Viena, Áustria pelas análises de microscopia eletrônica de varredura dos fios da seda, contribuindo com os resultados deste trabalho.

A todos os professores, pelo conhecimento e atenção... aos técnicos e funcionários pela simpatia, atenção, e dedicação... à Suzi, Gislaine e Rejane, por toda a ajuda na biblioteca, pelo carinho e por torcerem por mim... à Rose, Necis, Ciça, Felipe por toda a ajuda nas questões burocráticas e administrativas... e todos os demais que não citei.

A todos da turma de ciências biológicas noturno de 2004 por manter a sua singularidade e controvérsias, tão característicos dessa turma. Foi muito bom revê-los depois de tanto tempo e sentir que apesar dos 10 anos que se passaram, ainda continuaremos a manter contato e trocando experiências.

Aos meus amigos, próximos e distantes... pela amizade, por torcerem por mim, por vibrarem comigo pelas minhas conquistas e superação... à Lucilene, Daniel e Nicoli, Patrícia, Meiri, Grazi, Nathália, Cindi, Carlos, Daniel, Vanessa, Luciana, Ives, Sadala, Iata, Eliana... à Joice, pela amizade incondicional e paciência em me ouvir... por saber o que me dizer, e por toda a sua simplicidade, sonhos e vontade de vencer... com parcimônia! Aos meus velhos amigos Thiago Gazoni, Amá e Sérgio... e também aos novos amigos Paixão, Octávio, Fabrício e Carolina por toda amizade, discussões científicas e aleatórias, churrascos, diversões, almoços e jantares demorados... por simplesmente compartilharmos de um convívio praticamente diário, o que me ajudou muito na minha readaptação... obrigado a todos pelo companheirismo, amizade e consideração, preocupação, carinho e respeito. Cada um, à sua maneira, me proporciona e ensina muitas coisas.

Às aranhas *Nephila clavipes* que contribuíram com a seda para a realização deste estudo, em prol do desenvolvimento científico.

Ao Programa BIOprospecTA/FAPESP (Proc. 2011/51684-1 e Proc. 2010/19051-6) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq pelo suporte financeiro para a realização deste trabalho.

A todos que de alguma forma contribuíram para o meu crescimento pessoal e intelectual, que confiaram em mim e torceram por mim... Nunca ninguém sabe se estou louco para rir ou para chorar; não sou melhor, do mesmo modo que nunca serei menos do que ninguém. Só tenho que estar consciente das minhas escolhas e ser responsável pelos meus atos. Toda ação tem seu preço e seu prazer, reconhecer ambos os lados é ser realista. Não sei começar e parar, apenas faço. Essa é a única coragem de verdade. E então, aqui vou eu... ***muito obrigado!***

Tabela de aminoácidos

Aminoácido	Código de 1 letra	Código de 3 letras	Massa Média	Massa Monoisotópica
Glicina	G	Gly	57.0520	57.02146
Alanina	A	Ala	71.0788	71.03711
Serina	S	Ser	87.0782	87.03203
Prolina	P	Pro	97.1167	97.05276
Valina	V	Val	99.1326	99.06841
Treonina	T	Thr	101.1051	101.04768
Cisteína	C	Cys	103.1448	103.00919
Isoleucina	I	Ile	113.1595	113.08406
Leucina	L	Leu	113.1595	113.08406
Asparagina	N	Asn	114.1039	114.04293
Ácido Aspártico	D	Asp	115.0886	115.02694
Lisina	K	Lys	128.1742	128.09496
Glutamina	Q	Gln	128.1308	128.05858
Ácido Glutâmico	E	Glu	129.1155	129.04259
Metionina	M	Met	131.1986	131.04049
Histidina	H	His	137.1412	137.05891
Fenilalanina	F	Phe	147.1766	147.06841
Arginina	R	Arg	156.1876	156.10111
Tirosina	Y	Tyr	163.1760	163.06333
Triptofano	W	Trp	186.2133	186.07931

LISTA DE ABREVIÇÕES E SÍMBOLOS

% - por cento	mg - Miligrama
Δ - Delta	MS - Espectrometria de massas
°C - Graus Celsius	MI - Glândula ampulada menor
μg - Micrograma	MiSp1- Minor ampullate Spidroin-1
μL - Microlitro	MiSp2 - Minor ampullate Spidroin-2
2-DE - Gel de eletroforese bidimensional	MS/MS - Espectrometria de massas sequencial
3D - Estrutura tridimensional	MW - Peso molecular
Å - Ângstron	ECD - Dissociação por captura de elétrons
CCP4 - Collaborative Computational Project No 4	NanoLC-MS/MS - Cromatografia líquida em nanoescala acoplada a espectrometria de massas
CID - Dissociação induzida por colisão	ng - Nanograma
cm - Centímetro	nm - Nanômetro
Da - Dalton	ns - Nanosegundo
E - Energia	TFA - Ácido trifluoracético
ESI - Ionização por eletrospray	ng/ μL - Nanograma por microlitro
ETD - Dissociação por transferência de elétrons	pI - Ponto isoelétrico
FA - Ácido fórmico	ODA - Ácido graxo octadecanóico
g - Grama	PTMs - Modificações pós- traducionais
GC-MS - Cromatografia gasosa acoplada a espectrometria de massas	RMN - Ressonância magnética nuclear
HCL - Ácido clorídrico	RMSD - Raiz do desvio médio quadrático
HPLC - Cromatografia líquida de alta pressão	rpm - Rotações por minuto
J/Kg - Jaules por quilograma	SPC - Simple Point Charge
K - Kelvin	TDA - Ácido graxo tridecanóico
kDa - Kilodalton	V - Volts
L/min - Litros por minuto	VMD - Visual Molecular Dynamics

LC-MS/MS - Cromatografia líquida acoplada a espectrometria de massas

LiSCN - Tiocianato de lítio

M - Molar

m/z - Massa por carga

MA - Glândula ampulada maior

MALDI-MS - Matrix-assisted laser desorption/ionization MS

MaSp1 - Major ampullate Spidroin-1

PHD - Domínio prolil hidroxilase

Rg - Raio de giro

α - Alfa

β - Beta

φ - phi

ψ - psi

MaSp2 - Major ampullate Spidroin-2

SDS - Dodecil sulfato de sódio

RP-HPLC - Cromatografia líquida de alto desempenho de fase reversa

N - Normal

MET - Microscopia eletrônica de transmissão

MEV - Microscopia eletrônica de varredura

m/v - Massa por volume

PDB - Protein Data Bank

V/min - Volts por minuto

mL - Mililitros

mL/min - Mililitros por minuto

mm - Milímetros

mM - Milimolar

LISTA DE FIGURAS

	Página
Figura 1. <i>Nephila clavipes</i> - aranha construtora de teia aérea orbital coletada no campus da Universidade Estadual Paulista “Júlio de Mesquita Filho” localizado na cidade de Rio Claro, São Paulo.....	33
Figura 2. Esquema representativo das glândulas produtoras de seda das aranhas tecedoras de teias orbitais (adaptado de Vollrath, 2000). Tipos de fios de seda e suas respectivas glândulas produtoras.....	34
Figura 3. Microscopia eletrônica de varredura - MEV da glândula ampulada maior da aranha <i>N. clavipes</i> . A - cauda, B - ampola e C - ducto de fiação, indicado por uma seta (Imagem de Salles, 2004).....	34
Figura 4. Esquema representando a histologia da glândula ampulada maior (A - G) destacando o processo de fiação (L - lúmen). A, B e C - produção das espidroínas e secreção de solução viscosa por células colunares e armazenamento; D, E, F e G - histologia do ducto de fiação. A região do ducto em F parece ser altamente especializada para bombeamento de água, contendo células altas com numerosas mitocôndrias, microvilosidades apicais e uma ampla superfície formada pelas vilosidades internas da membrana plasmática. V - Z , demonstram a natureza afunilada do extenso ducto de fiação. As funções do estreitamento da glândula são reter água, alinhar as proteínas, polimerização e produção de uma fibra mecanicamente equilibrada com custos metabólicos mínimos (Adaptado de Vollrath e Knight, 2001).....	36
Figura 5. Tipos de fios de seda secretados pelas aranhas tecedoras de teias orbitais. A - glândula ampulada maior secreta os fios de seda dos raios da teia orbital, do frame e da linha de segurança “escape”; B - glândula flageliforme secreta os fios de seda da espiral de captura; C - glândula ampulada menor secreta os fios de seda que auxiliam a espiral de captura; D - glândula piriforme secreta os fios de seda das conexões entre as sedas e a fixação sobre um substrato; E - glândula agregada secreta um revestimento adesivo aquoso sobre a espiral; F - glândula tubuliforme secreta os fios de seda que cobre/protege os ovos externamente (bolsa de seda) após a postura; G - glândula aciniforme secreta os fios de seda que protege os ovos internamente (adaptado de Römer et al., 2008).	38
Figura 6. <i>Biosteel Spidersilk Fibers</i> produzida pela AMSilk® <i>high performance materials</i>	42
Figura 7. Cosméticos produzidos pela AMSilk® <i>high performance materials</i> contendo as propriedades da seda de aranhas comercialmente disponíveis.(http://www.amsilk.com/fileadmin/content/TechInsight_Cosmetics.pdf).	43
Figura 8. Esquema representativo da estrutura secundária das proteínas constituintes das fibras da seda da aranha <i>N. clavipes</i> , secretadas pelas glândulas ampulada maior e flageliforme. Esse esquema mostra a conformação estrutural das proteínas na glândula e a conformação estrutural obtida após serem secretadas durante o processo de fiação (modificado de Lefèvre et al., 2011).....	45
Figura 9. Esquema representando a teoria micelar sobre o processo de formação das fibras de seda, que ocorre ao longo dos aparatos do mecanismo de fiação	

(adaptado de Eisoldt et al., 2011). **(A)** Cauda - baixa concentração de proteínas; **(B)** Ampola - as proteínas são parcialmente pré-alinhadas com a formação de pequenas micelas apresentando propriedades de líquido-cristalino; **(C)** Ducto de fiação - alongamento e formação das fibras, acidificação do meio induzindo a dimerização do domínio N-terminal, troca iônica e remoção de água, alinhamento e formação das folhas- β , desdobramento do domínio C-terminal expondo as regiões hidrofóbicas e facilitando o dobramento das fibras; **(D)** Exterior - fibra secretada, folhas- β são alinhadas ao longo do eixo da fibra..... 46

Figura 10. A - Esquema representativo da regulação da fosforilação de uma proteína. A enzima quinase é responsável pela adição do grupo fosfato à proteína, enquanto que a enzima fosfatase é responsável pela remoção do grupo. **B** - fosfoserina, fosfotreonina e fosfotirosina..... 50

Figura 11. Esquema representativo da estrutura química geral de um peptídeo, apresentando as possibilidades de fragmentações que surgem em espectros de massas obtidos por análise *in tandem*. Os íons formados são enumerados a partir do lado N-terminal, sendo assim, os íons são denominados por -a, -b ou -c se a carga for mantida no lado N-terminal e, denominados por -x, -y ou -z se a carga for mantida no lado C-terminal. O índice indica o número do resíduo de aminoácido do fragmento em questão 57

Figura 12. Dissecção da aranha *N. clavipes* com a identificação e coleta das glândulas ampulada maior, ampulada menor, flageliforme, agregada e tubuliforme. **A** - *N. clavipes*; **B-E** - Dissecção e identificação das glândulas; **F** - Coleta da glândula ampulada maior..... 77

Figura 13. Perfil eletroforético da 2-DE da glândula ampulada maior da aranha *N. clavipes*..... 79

Figura 14. Espectro CID da sequência peptídica DFEDFINLAK, selecionando-se o íon de m/z 606.28 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína hemocianina (*spot* 2), presente na glândula ampulada maior e identificada na 2-DE da seda..... 85

Figura 15. Espectro CID da sequência peptídica AQLDFPGVQVVNVTNAK, selecionando-se o íon de m/z 950.08 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína hemocianina (*spot* 2), presente na glândula ampulada maior e identificada na 2-DE da seda..... 86

Figura 16. Espectro CID da sequência peptídica AYLPVNESFGFTADLR, selecionando-se o íon de m/z 900.40 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína fator de alongamento de tradução 2 (*spot* 44), presente na glândula ampulada maior e identificada na 2-DE da seda..... 86

Figura 17. Espectro CID da sequência peptídica GGLGGQGAGAAAAGGAGQGGYGGVGSASAASAAASR, selecionando-se o íon de m/z 954.72 na forma de $[M + 3H]^{3+}$, como precursor observado na proteína espidroína-1 (*spot* 56), presente na glândula ampulada maior e identificada na 2-DE da seda..... 87

Figura 18. Espectro CID da sequência peptídica LVQAFQFTDK, selecionando-se o íon de m/z 598.76 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína tiorredoxina peroxidase (*spot* 61), presente na glândula ampulada maior e identificada na 2-DE da seda..... 87

Figura 19. Espectro CID da sequência peptídica TITLEVEPSDTIENVK, selecionando-se o íon de m/z 894.39 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína ubiquitina (<i>spot</i> 70), presente na glândula ampulada maior e identificada na 2-DE da seda.....	88
Figura 20. Perfil eletroforético da 2-DE da glândula flageliforme da aranha <i>N. clavipes</i>	89
Figura 21. Espectro CID da sequência peptídica AILLDLEPGTMDSVR, selecionando-se o íon de m/z 815.38 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína tubulina (<i>spot</i> 23), presente na glândula flageliforme e identificada na 2-DE da seda.....	93
Figura 22. Espectro CID da sequência peptídica GHYTEGAELVDTVLDVVR, selecionando-se o íon de m/z 986.94 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína tubulina (<i>spot</i> 23), presente na glândula flageliforme e identificada na 2-DE da seda.....	94
Figura 23. Espectro CID da sequência peptídica QISLSGPLSILGR, selecionando-se o íon de m/z 670.83 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína superóxido dismutase (<i>spot</i> 47), presente na glândula flageliforme e identificada na 2-DE da seda.....	94
Figura 24. Espectro CID da sequência peptídica TNVISNALR, selecionando-se o íon de m/z 494.26 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína da seda flageliforme (<i>spot</i> 49), presente na glândula flageliforme e identificada na 2-DE da seda.....	95
Figura 25. Espectro CID da sequência peptídica AMQVALASSIAELVIAESSGGDVQR, selecionando-se o íon de m/z 834.78 na forma de $[M + 3H]^{3+}$, como precursor observado na proteína da seda flageliforme (<i>spot</i> 49), presente na glândula flageliforme e identificada na 2-DE da seda.....	95
Figura 26. Perfil eletroforético da 2-DE da glândula tubuliforme da aranha <i>N. clavipes</i>	97
Figura 27. Espectro CID da sequência peptídica IINEPTAAAIAYGLDK, selecionando-se o íon de m/z 830.43 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína de choque térmico (<i>spot</i> 1), presente na glândula tubuliforme e identificada na 2-DE da seda.....	101
Figura 28. Espectro CID da sequência peptídica SLGIADAAGLAGVLAR, selecionando-se o íon de m/z 727.97 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína da seda tubuliforme (<i>spot</i> 54), presente na glândula tubuliforme e identificada na 2-DE da seda.....	102
Figura 29. Espectro CID da sequência peptídica SLGIADAAGLAGALAR, selecionando-se o íon de m/z 713.94 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína da seda tubuliforme (<i>spot</i> 63), presente na glândula tubuliforme e identificada na 2-DE da seda.....	102
Figura 30. Perfil eletroforético da 2-DE da glândula ampulada menor da aranha <i>N. clavipes</i>	104
Figura 31. Perfil eletroforético da 2-DE da glândula agregada da aranha <i>N. clavipes</i>	109

Figura 32. Representação do mecanismo geral de ação das proteínas identificadas nas glândulas produtoras de seda da (A) aranha <i>N. clavipes</i> . (B) - glândula produtora de seda; (C) - ducto de fiação; (D) - fibras da seda, constituídas pelas espidroínas, após a secreção pelo ducto de fiação.....	118
Figura 33. SEM – Microscopia eletrônica de varredura da seda do “frame” da teia da aranha <i>N. clavipes</i> . (A) seda coletada constituída pelos fios dos raios do “frame” enrolados; (B-D) ultraestrutura dos fios da seda em diferentes pontos da teia.....	120
Figura 34. SEM – Microscopia eletrônica de varredura da seda da teia orbital da aranha <i>N. clavipes</i> . (A) seda coletada constituída pelos fios dos raios e da espiral de captura da teia orbital; (B-D) ultraestrutura dos fios da seda em diferentes pontos, destacando a presença de macroporos.....	121
Figura 35. SEM - Microscopia eletrônica de varredura da seda da aranha <i>N. clavipes</i> . (A-D) mostra o padrão de dissolução da seda, em diferentes pontos, na presença de solução saturada de tiocianato de lítio (LiSCN hidratado 38M).....	122
Figura 36. Perfil eletroforético da 2-DE da seda do “frame” da aranha <i>N. clavipes</i>	123
Figura 37. Perfil eletroforético da 2-DE da seda da teia orbital da aranha <i>N. clavipes</i>	124
Figura 38: Alinhamento múltiplo entre as sequências primárias das proteínas espidroína-1A, espidroína-1B, espidroína-2 e proteína da seda flageliforme (Flag. silk protein) identificadas na seda da aranha <i>N. clavipes</i> . As regiões de identidade mais importantes estão assinaladas em azul escuro, enfatizando os resíduos altamente conservados nesse tipo de alinhamento.....	128
Figura 39. Representação da sequência primária e caracterização da estrutura secundária (JNETPRED <i>Secondary Structure Prediction</i>) da proteína espidroína-1 que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda da aranha <i>N. clavipes</i> . A sequência foi obtida através da digestão <i>in-gel</i> utilizando diferentes enzimas proteolíticas, e mostra a localização das PTMs mais abundantes observadas na proteína.....	135
Figura 40. Representação da sequência primária e caracterização da estrutura secundária (JNETPRED <i>Secondary Structure Prediction</i>) da proteína espidroína-2 que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda da aranha <i>N. clavipes</i> . A sequência foi obtida através da digestão <i>in-gel</i> utilizando diferentes enzimas proteolíticas, e mostra a localização das PTMs mais abundantes observadas na proteína.....	136
Figura 41. Representação da sequência primária e caracterização da estrutura secundária (JNETPRED <i>Secondary Structure Prediction</i>) da proteína da seda flageliforme que foi secretada pela glândula flageliforme e identificada na 2-DE da seda da aranha <i>N. clavipes</i> . A sequência foi obtida através da digestão <i>in-gel</i> utilizando diferentes enzimas proteolíticas, e mostra a localização das PTMs mais abundantes observadas na proteína.....	137
Figura 42. (A) - Espectro CID da sequência peptídica R.TGAFTADQLD D *M S T*IGDTLK.T (49-68), selecionando-se o íon de m/z 713.34	

na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação T62 e o aminoácido substituto Asp->Gly (D59) observados na proteína espidroína-1 (*spot* 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica R.TGAFTADQLDDMSTIGDTLK.T (49-68), selecionando-se o íon de m/z 686.67 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a desfosforilação do sítio T62, que foi observado na proteína espidroína-1 (*spot* 6), após o tratamento com fosfatase alcalina..... 139

Figura 43. (A) - Espectro CID da sequência peptídica R.T*GAFTAD*QLDDMSTIGDTLK.T (49-68), selecionando-se o íon de m/z 1061.53 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação T49 e o aminoácido substituto Asp->Gly (D55) observados na proteína espidroína-1 (*spots* 3 e 5), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica R.TGAFTADQLDDMSTIGDTLK.T (49-68), selecionando-se o íon de m/z 1021.53 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando a desfosforilação do sítio T49, que foi observado na proteína espidroína-1 (*spots* 3 e 5), após o tratamento com fosfatase alcalina..... 140

Figura 44. (A) - Espectro CID da sequência peptídica K.LQALNMAFASS*MAEIAAVEQGGLSVAEK.T (85-112), selecionando-se o íon de m/z 982.50 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação S95 observado na proteína espidroína-1 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica K.LQALNMAFASSMAEIAAVEQGGLSVAEK.T (85-112), selecionando-se o íon de m/z 955.83 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a desfosforilação do sítio S95, que foi observado na proteína espidroína-1 (*spot* 4), após o tratamento com fosfatase alcalina..... 141

Figura 45. (A) - Espectro CID da sequência peptídica Y.GGLGS*QGAGRGGQGAGAAAAAAGGAGQGGY.G (60-89, 576-605), selecionando-se o íon de m/z 821.027 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando os sítios de fosforilação S64 e S580 observados na proteína espidroína-1 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY.G (60-89, 576-605), selecionando-se o íon de m/z 794.36 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a desfosforilação dos sítios S64 e S580, que foram observados na proteína espidroína-1 (*spot* 4), após o tratamento com fosfatase alcalina..... 142

Figura 46. (A) - Espectro CID da sequência peptídica GGQGAAAAAAGGAGQGGY*GGLGSQGAGR (134-161), selecionando-se o íon de m/z 811.080 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação Y151 observado na proteína espidroína-1 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica GGQGAAAAAAGGAGQGGYGGGLGSQGAGR (134-161), selecionando-se o íon de m/z 784.42 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a desfosforilação do sítio Y151, que foi observado na proteína espidroína-1 (*spot* 3), após o tratamento com fosfatase alcalina..... 143

Figura 47. (A) - Espectro CID da sequência peptídica L.GGQGAGAAAAAAGGAGQGGGLGGQGAGQGAGAS*AAAAGGAGQGGY.G

(211-255), selecionando-se o íon de m/z 863.34 na forma de $[M + 4H]^{4+}$, como precursor; e mostrando o sítio de fosforilação S243 observado na proteína espidroína-1 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica L.GGQGAGAAAAAAGGAGQGGLGGQGAGQGAGASAAAAGGAGQGGY.G (211-255), selecionando-se o íon de m/z 843.34 na forma de $[M + 4H]^{4+}$, como precursor; e mostrando a desfosforilação do sítio S243, que foi observado na proteína espidroína-1 (*spot* 3), após o tratamento com fosfatase alcalina.....

144

Figura 48. Ensaio de imunodeteção representativo da seda (20-50 μ g de proteínas) da teia total da aranha *N. clavipes* mostrando a imunoreatividade dos sítios de fosforilação observados nas proteínas da seda. **A** - Imunodeteção utilizando o anticorpo primário anti-fosfotirosina. Linhas 1 e 3 mostram a imunoreatividade dos sítios de fosfotirosina (setas); linhas 2 e 4 mostram a não imunoreatividade após o tratamento da amostra com fosfatase alcalina. **B** - Imunodeteção utilizando o anticorpo primário anti-fosfoserina. Linhas 1 e 3 mostram a imunoreatividade dos sítios de fosfoserina (setas); linhas 2 e 4 mostram a não imunoreatividade após o tratamento da amostra com fosfatase alcalina.....

145

Figura 49. (A) - Espectro CID da sequência peptídica L.N*EAGRTGAF.T (44-52), selecionando-se o íon de m/z 535.22 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de glicosilação N44 (hexosamine) observado na proteína espidroína-1 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica L.NEAGRTGAF.T (44-52), selecionando-se o íon de m/z 454.72 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando a deglicosilação do sítio N44, que foi observado na proteína espidroína-1 (*spot* 3), após o tratamento com PNGase.....

148

Figura 50. (A) - Espectro CID da sequência peptídica Y.N*GGQGSPSAAAAAASGAGQGGY.G (237-260), selecionando-se o íon de m/z 877.88 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de glicosilação N237 (HexNAc2DHex2) observado na proteína espidroína-1 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica Y.NGGQGSPSAAAAAASGAGQGGY.G (237-260), selecionando-se o íon de m/z 645.21 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a deglicosilação do sítio N237, que foi observado na proteína espidroína-1 (*spot* 4), após o tratamento com PNGase.....

149

Figura 51. Ensaio de imunodeteção representativo da seda (20-50 μ g de proteínas) da teia total da aranha *N. clavipes* mostrando a imunoreatividade do sítio de nitrotirosina observado nas proteínas da seda. Imunodeteção utilizando o anticorpo primário anti-nitrotirosina. Linha 1 mostra a amostra controle positivo (Nitro-BSA - *Nitrated bovine serum albumin*); linhas 2 e 3 mostram a imunoreatividade do sítio de nitrotirosina.....

150

Figura 52. Análises de ditirosina. **Linha (a)** - padrão de ditirosina, **Linha (b)** - hidrolisado da amostra da seda da aranha *N. Clavipes*.....

158

Figura 53. Modelos moleculares das estruturas 3D das espidroínas constituintes das fibras da seda da aranha *N. clavipes*. **A** - espidroína-1A; **B** - espidroína-1B; **C** - espidroína-2.

160

Figura 54. Potencial eletrostático para os modelos das espidroínas gerados a partir do programa PyMol. A - espidroína-1A; B - espidroína-1B; C - espidroína-2. Imagens localizadas na parte superior da imagem estão em 0° (frente) e as imagens localizadas na parte inferior sofreram uma rotação de 180° (verso).....	161
Figura 55. Avaliação da qualidade das estruturas 3D dos modelos das espidroínas constituintes das fibras da seda da aranha. Diagrama de Ramachandran para as proteínas da seda da teia de <i>N. clavipes</i> . A - espidroína-1A; B - espidroína-1B; C - espidroína-2.....	164
Figura 56. Microscopia eletrônica de transmissão (MET) com técnica de citoquímica para detecção de lipídeos. Fibras da seda mostrando a deposição de material lipídico formando uma camada lipídica sobre a fibra (setas).....	166
Figura 57. Estrutura 3D dos ácidos graxos (A) tridecanóico e (B) octadecanóico contruídos a partir do programa PyMol.....	167
Figura 58. Esquema representativo das caixas que foram utilizadas durante a simulação de dinâmica molecular. Espidroína-1 , representada em <i>ribbons</i> ; água , representada em <i>sticks</i> vermelhos; e ácidos graxos , representados em <i>sticks</i> verdes. As duas figuras representam as caixas na presença de espidroína-1, ácidos graxos e água sob dois diferentes ângulos.....	168
Figura 59. Modelos moleculares das estruturas 3D das espidroínas constituintes das fibras da seda da aranha <i>N. clavipes</i> após a dinâmica molecular de 20ns em água. A - espidroína-1A; B - espidroína-1A após a dinâmica em água; C - espidroína-1A com fosforilação, após a dinâmica em água. D - espidroína-1B; E - espidroína-1B após a dinâmica em água; F - espidroína-1B com fosforilação, após a dinâmica em água.....	168
Figura 60. (A) - RMSD da espidroína-1A (azul) em relação a espidroína-1A com fosforilação (vermelho). (B) - RMSD da espidroína-1B (azul) em relação a espidroína-1B com fosforilação (vermelho).....	170
Figura 61. (A) - Gráfico do Raio de Giro para a espidroína-1A (azul) em relação a espidroína-1A com fosforilação (vermelho). (B) - Gráfico do Raio de Giro para a espidroína-1B (azul) em relação a espidroína-1B com fosforilação (vermelho).....	172
Figura 62. (A) - Gráfico do Energia total para a espidroína-1A (azul) em relação a espidroína-1A com fosforilação (vermelho). (B) - Gráfico do Energia total para a espidroína-1B (azul) em relação a espidroína-1B com fosforilação (vermelho).....	174
Figura 63. Avaliação da qualidade das estruturas 3D dos modelos das espidroínas constituintes das fibras da seda da aranha após a dinâmica molecular de 20ns em água. Diagrama de Ramachandran para as proteínas da seda da teia de <i>N. clavipes</i> . A - espidroína-1A após a dinâmica em água; B - espidroína-1A com fosforilação, após a dinâmica em água; C - espidroína-1B após a dinâmica em água; D - espidroína-1B com fosforilação, após a dinâmica em água.....	177
Figura 64. Esquema representativo dos modelos moleculares das estruturas 3D das espidroínas em caixas durante a dinâmica de restrição para o sistema em água com ácidos graxos. A - espidroína-1A durante a dinâmica de restrição em água com ácidos graxos; B - espidroína-1A com fosforilação, durante a dinâmica de restrição em água com ácidos graxos; C - espidroína-1B durante a dinâmica	

de restrição em água com ácidos graxos; D - espidroína-1B com fosforilação, durante a dinâmica de restrição em água com ácidos graxos. Ácidos graxos , representados em <i>sticks</i> verdes.....	179
Figura 65. Avaliação da qualidade das estruturas 3D dos modelos das espidroínas constituintes das fibras da seda da aranha após a dinâmica molecular de restrição em água com ácidos graxos. Diagrama de Ramachandran para as proteínas da seda da teia de <i>N. clavipes</i> . A - espidroína-1A após a dinâmica em água com ácidos graxos; B - espidroína-1A com fosforilação, após a dinâmica em água com ácidos graxos; C - espidroína-1B após a dinâmica em água com ácidos graxos; D - espidroína-1B com fosforilação, após a dinâmica em água com ácidos graxos.....	181
Figura 66. Modelos moleculares das estruturas 3D das espidroínas constituintes das fibras da seda da aranha <i>N. clavipes</i> após a dinâmica molecular final parcial de 10ns em água com ácidos graxos. A - espidroína-1A após a dinâmica em água com ácidos graxos; B - espidroína-1A com fosforilação, após a dinâmica em água com ácidos graxos; C - espidroína-1B após a dinâmica em água com ácidos graxos; D - espidroína-1B com fosforilação, após a dinâmica em água com ácidos graxos.....	183
Figura 67. Modelos moleculares das estruturas 3D das espidroínas constituintes das fibras da seda da aranha <i>N. clavipes</i> após a dinâmica molecular final parcial de 10ns em água com ácidos graxos. A - espidroína-1B após a dinâmica em água com ácidos graxos; B - espidroína-1B com fosforilação, após a dinâmica em água com ácidos graxos. Em destaque a interação das moléculas de ácidos graxos com as estruturas da espidroína.....	184
Figura 68. Representação geral esquematizada para o mecanismo natural de fiação das fibras da seda da teia da aranha <i>N. clavipes</i> , destacando a presença de ditirosina, uma camada de ácidos graxos sobre as fibras e as principais PTMs observadas sobre as espidroínas constituintes das fibras que formam a espiral e os raios/frame da teia. (A) - aranha <i>N. clavipes</i> ; (B) - glândula produtora de seda; (C) - ducto de fiação; (D) - fibras da seda, constituídas pelas espidroínas, após a secreção pelo ducto de fiação; (E) - espidroínas-1A, espidroínas-1B e espidroínas-2 representadas em suas estruturas 3D, constituindo as fibras da seda.....	186

LISTA DE TABELAS

	Página
Tabela 1. Comparação entre as diferentes propriedades mecânicas de sedas de aranhas, seda do bicho da seda, fibras naturais e polímeros sintéticos (adaptado de Hinman et al., 2000; Lewis, 2006).....	39
Tabela 2. Algumas das mais comuns e importantes PTMs (adaptado de Mann e Jensen, 2003).....	49
Tabela 3. Moldes utilizados para a construção dos modelos 3D das espidroínas, utilizando a metodologia do alinhamento sequencial de pares.....	71
Tabela 4. Moldes utilizados para a construção dos modelos 3D das espidroínas, utilizando a metodologia de compatibilidade sequência-estrutura.....	71
Tabela 5. Simulação por dinâmica molecular para cada modelo 3D utilizando diferentes sistemas estabelecidos para as proteínas.....	73
Tabela 6. Proteínas identificadas na glândula ampulada maior da aranha <i>N. clavipes</i> através da digestão <i>in-gel</i> dos spots oriundos da 2-DE.....	80
Tabela 7. Proteínas identificadas na glândula flageliforme da aranha <i>N. clavipes</i> através da digestão <i>in-gel</i> dos spots oriundos da 2-DE.....	90
Tabela 8. Proteínas identificadas na glândula tubuliforme da aranha <i>N. clavipes</i> através da digestão <i>in-gel</i> dos spots oriundos da 2-DE.....	98
Tabela 9. Proteínas identificadas na glândula ampulada menor da aranha <i>N. clavipes</i> através da digestão <i>in-gel</i> dos spots oriundos da 2-DE.....	105
Tabela 10. Proteínas identificadas na glândula agregada da aranha <i>N. clavipes</i> através da digestão <i>in-gel</i> dos spots oriundos da 2-DE.....	110
Tabela 11. Proteínas identificadas na seda da aranha <i>N. clavipes</i> através da digestão <i>in-gel</i> dos spots oriundos da 2-DE.....	125
Tabela 12. Porcentagem de cobertura da sequência das proteínas identificadas na seda da aranha <i>N. clavipes</i> através da digestão <i>in-gel</i> utilizando diferentes enzimas proteolíticas.....	126
Tabela 13. Modificações pós-traducionais (PTMs) observadas nas proteínas da seda da aranha <i>N. clavipes</i> mostradas através do MASCOT e Modiro®.....	132
Tabela 14. Mapeamento dos sítios de fosforilação, glicosilação e nitrotirosina observados na sequência da proteína espidroína-1, a qual foi secretada pela glândula ampulada maior e identificada na seda produzida pela aranha <i>N. clavipes</i>	152
Tabela 15. Mapeamento dos sítios de fosforilação observados na sequência da proteína espidroína-2, a qual foi secretada pela glândula ampulada maior e identificada na seda produzida pela aranha <i>N. clavipes</i>	153
Tabela 16. Mapeamento dos sítios de hidroxilação, fosforilação e nitrotirosina observados na sequência da proteína da seda flageliforme, a qual foi secretada pela glândula flageliforme e identificada na seda produzida pela aranha <i>N.</i>	

<i>clavipes</i>	153
Tabela 17. Análises para os modelos moleculares das espidroínas utilizando o programa PROCHECK.....	164
Tabela 18. Análises para os modelos moleculares das espidroínas após a dinâmica molecular de 20ns em água , utilizando o programa PROCHECK.....	177
Tabela 19. Análises para os modelos moleculares das espidroínas após a dinâmica molecular de restrição em água com ácidos graxos, utilizando o programa PROCHECK.....	181
Tabela 20. Análise das estruturas secundárias, para os modelos moleculares das espidroínas, após a dinâmica molecular de 20 ns em água, e após a dinâmica molecular parcial de 10 ns em presença de ácidos graxos. Foi utilizado o programa STRIDE (Frishman e Argos, 1995) e o DSSP (Joosten et al., 2010) através do GROMACS para calcular a porcentagem e o número de estruturas secundárias.....	184

A lenda de Aracne

Segundo a mitologia grega

Numa antiga região da Ásia Menor chamada de Lídia, vivia uma jovem de nome Aracne, que tinha uma extraordinária habilidade e grande reputação na arte de tecer e bordar. As tapeçarias que desenhava eram tão belas e perfeitas que pessoas vinham de terras distantes só para contemplá-las. Devido a tanta admiração, Aracne começou a comparar-se à Atenas, deusa das fiandeiras, e que seria capaz de derrotá-la na arte da tecelagem. Quando a notícia chegou ao Olimpo, Atenas ficou furiosa com a petulância da mortal. Sentiu-se desafiada e resolveu aceitar a competição com Aracne, para ver quem merecia de fato ser considerada a melhor na arte de bordar. Antes, porém, a deusa disfarçou-se de uma humilde velhinha e foi conversar com Aracne. Pediu-lhe que a escutasse, devido à experiência de sua idade avançada: "Busque entre os mortais toda fama que desejar, mas reconheça a posição da deusa". Aracne, entretanto, não aceitou os conselhos da deusa, e, mais uma vez, a desafiou, dizendo: "Por que motivo sua deusa está evitando competir comigo"? Nesse momento, Atenas tirou o disfarce e todos, ao redor, ficaram surpresos, exceto Aracne, que permaneceu impassível. As duas, então, deram início à competição.

Ambas trabalharam com rapidez e habilidade. Atenas representou sobre a tapeçaria os doze deuses do Olimpo em toda a sua majestade e para aviso da sua rival acrescentou nos quatro cantos a representação de quatro episódios mostrando a derrota dos mortais que tinham ousado desafiar os deuses. Aracne ousou ilustrar sobre o seu trabalho as conquistas amorosas de Zeus. Sob a forma de touro, arrebatando Europa; sob a forma de águia, abordando Astéria; sob a forma de cisne, conquistando Leda; sob a forma de sátiro, fazendo amor com Antíope; Zeus fazendo-se passar por Anfitrión para seduzir Alcmena, mãe de Heraclés (Hércules); Zeus, o pastor que fez amor com Mnemosine, mulher-titã; e, ainda, Zeus conquistando Egina, Deméter e Danae, disfarçado de chama, serpente e chuva de ouro, respectivamente. Isso deixou Atenas tão enraivecida que rasgou em pedaços o trabalho e golpeou, com o bastão de tecer, a cabeça Aracne. Ultrajada e desesperada, Aracne enforca-se.

Atenas, ao constatar o que a sua cólera havia provocado, compadeceu-se de Aracne e transformou a corda que ela usara para enforcar-se numa teia. Em seguida, derramou sobre Aracne fluidos retirados das ervas da deusa Hecate e transformou-a em uma aranha. Dessa forma, Aracne foi salva da morte, embora condenada a ficar pendurada em sua teia a fiar e a tecer para sempre. Hoje, a beleza de sua arte é apreciada em todo o mundo através dos fios delicados e altamente resistentes tecidos pelas aranhas.

Para maiores informações sobre o assunto verificar: Hacquard, Georges. **Guide Mythologique de la Grèce et de Rome**. 1990. Trad. Maria Helena, Adapt. Ricardo Sérgio.
<http://www.recantodasletras.com.br/contosdefantasia/3025143>

Resumo

As aranhas construtoras de teia aérea orbital tornaram-se eficientes produtoras de diferentes tipos de sedas. Essas sedas são caracterizadas pela notável diversidade em sua composição química, estrutura e função, que vai desde a construção da teia orbital até o casulo. A seda produzida pelas aranhas são filamentos proteicos, constituindo uma interessante relação de estrutura-função e propriedades mecânicas, sendo produzida por um conjunto de glândulas abdominais. As fibras produzidas pela glândula ampulada maior são um dos tipos mais importantes de fibras, sendo um material nanoestruturado predominantemente composto por duas proteínas estruturais, espidroína-1 e espidroína-2. Enquanto que as fibras produzidas pela glândula flageliforme são compostas somente por uma proteína, a proteína da seda flageliforme. Apesar do grande interesse pelas propriedades mecânicas da seda das aranhas, visando o seu uso em aplicações biomédicas e biotecnológicas devido as suas propriedades de resistência, elasticidade e biocompatibilidade, pouco se conhece sobre os detalhes das glândulas produtoras de seda e o processo de fiação que ocorre para a produção das fibras. A seda da teia de aranha tem sido extensivamente estudada através de dados oriundos da engenharia genética e da tecnologia do DNA recombinante. No entanto, informações químicas como as modificações pós-traducionais (PTMs) podem ser perdidas, o que pode inviabilizar a manutenção das propriedades mecânicas e físico-químicas tão características da seda de aranha. Sendo assim, tanto as espidroínas naturais quanto as suas formas recombinantes foram até então, bioquímica e fisicamente, caracterizados como se apresentassem as mesmas (ou similares) propriedades físico-químicas, sem qualquer consideração sobre a existência de PTMs em suas sequências. Por tanto, no presente estudo adotamos a abordagem proteômica *bottom-up* com a utilização e combinação da 2-DE, cromatografia líquida e espectrometria de massas (NanoLC-ESI-CID/ETD-MSⁿ) para não só identificar o perfil proteico das glândulas que secretam a seda da aranha *Nephila clavipes*, como também as espidroínas constituintes dessa seda e a detecção de suas PTMs. Os resultados obtidos com a análise proteômica do conjunto de glândulas produtoras de seda da aranha *N. clavipes*, mostrou a identificação de seis diferentes grupos de proteínas, que foram classificadas de acordo com a sua função geral, sendo proposto pela primeira vez um mecanismo geral de ação dessas proteínas: (i) proteínas constituintes das fibras da seda, as espidroínas, (ii) proteínas relacionadas com o dobramento/conformação e modificação das espidroínas, (iii) proteínas relacionadas com a preservação das espidroínas (*self-protection*) contra o estresse oxidativo, (iv) proteínas relacionadas com a

preservação fibrilar das espidroínas no meio ambiente, (v) proteínas de transporte de íons e moléculas relacionadas com a estabilidade das espidroínas, e (vi) proteínas *housekeeping*. Utilizando uma abordagem proteômica *bottom-up*, mostramos pela primeira vez a análise estrutural e a presença de PTMs observadas sobre as espidroínas constituintes das fibras da seda de aranhas; com uma elevada cobertura das sequências primárias, utilizando diferentes enzimas proteolíticas, foram observados 12 sítios de fosforilação na região N-terminal e 16 sítios de fosforilação nas regiões central e C-terminal na espidroína-1; 11 sítios de fosforilação na região central e C-terminal na espidroína-2, e 3 sítios de fosforilação na proteína da seda flageliforme; também foram observados 2 sítios de glicosilação e 1 de nitrotirosina na espidroína-1, e 18 sítios de hidroxilação e 1 de nitrotirosina na proteína da seda flageliforme. A grande quantidade de sítios de fosforilação (localizados dentro dos motivos estruturais que seriam responsáveis pelas propriedades mecânicas da seda), juntamente com a provável fotoindução de hidroxilação, pode ser relevante para diversos estudos em ciências de materiais, biologia, biofísica e bioquímica. A caracterização estrutural 3D das espidroínas através da bioinformática estrutural mostrou que essas proteínas apresentam uma estrutura de grampo de folhas- β , que está relacionada com a estabilidade e dobramentos de proteínas. A modelagem e a dinâmica molecular mostraram que ocorre uma alteração da conformação estrutural nas espidroínas quando expostas, por exemplo, em um sistema aquoso; e quando em presença de ácidos graxos, a dinâmica molecular sugere que ocorre uma menor perda na quantidade de folhas- β . Os resultados obtidos para as espidroínas poderá proporcionar uma melhor compreensão das propriedades da seda visando à produção sintética/recombinante de novos materiais à base da seda de aranhas.

Palavras chave: seda de aranha; espidroínas; análise proteômica *bottom-up*; espectrometria de massas; sequenciamento; modificação pós-traducional; fosforilação; estrutura 3D; dinâmica molecular.

Abstract

The orb-web spiders became efficient producers of different types of silks. These silks are characterized by diversity in their chemical composition, structure and function, ranging from the construction of the orb-web to the cocoon. Spider web silk proteins (spidroins) have interesting relationships between their 3-D structures and mechanical properties of these protein fibers, which are secreted by specialized abdominal glands. The major ampullate silk fibers, produced by major ampullate gland, are one of the most important types of fibers spun produced by the orb-web spiders of genus *Nephila*, and are nanostructured composite materials predominantly composed of two structural proteins, designated spidroin-1 and -2. While the fibers produced by the flagelliform gland consist of only a single protein, the flagelliform silk protein. Despite the great interest in the spider silk mechanical properties, targeting its use in biomedical and biotechnological applications due to its properties of strength, elasticity and biocompatibility, the knowledge about the details of the silk-producing glands and the spinning process that occurs for fiber production are limited. The spider silk has been extensively studied using data from genetic engineering and recombinant DNA technology. However, chemical information such as post-translational modifications (PTMs) may be lost, making difficult to explain the mechanical and physico-chemical properties of the spider silks. Thus, both the natural protein and the recombinant spidroins have been biochemically and physically characterized as they would have the same (or similar) physico-chemical properties, without any consideration about the existence of PTMs in their sequences. Therefore, in this study we used a bottom-up proteomic approach combining 2-DE with *in-gel* protein digestion by different proteolytic enzymes, followed by mass spectrometry analysis (NanoLC-ESI-CID/ETD-MSⁿ) to identify the protein profile of the glands that secrete the spidroins and for protein sequencing and post-translational assignments of spidroins. The results obtained with proteomics analysis of the silk-producing glands from *Nephila clavipes* spider revealed the identification of six different groups of proteins, which have been classified according to their general function, being proposed for the first time a general mechanism of action of these proteins: (i) constituent proteins of the silk fibers, the spidroins, (ii) proteins related to folding/conformation and modification of spidroins, (iii) proteins related to spidroin preservation (self-protection) against oxidative stress, (iv) proteins related to spidroin fibrillar preservation at the environment, (v) proteins related to the transport of ions and molecules involved with spidroin stability, and (vi) housekeeping proteins.

Using a bottom-up proteomic approach, we showed for the first time the structural analysis and the presence of PTMs on spidroins constituent of spider silk fibers, with a high sequence coverage using different proteolytic enzymes, a series of phosphorylation sites were observed: 12 phosphorylation sites were observed at the N-terminal and 16 phosphorylation sites at the central part and C-terminal on spidroin-1, 11 phosphorylation sites at the central part and C-terminal on spidroin-2 and 3 phosphorylation sites on flagelliform silk protein; also were observed 2 glycosylation sites and 1 nitrotyrosine site on spidroin-1, and 18 hydroxylation sites and 1 nitrotyrosine site on flagelliform silk protein. The many site specific phosphorylations (localized within of the structural motifs, that would be responsible for the silk mechanical properties) along with the probably photoinduction of hydroxylations may be relevant for scientists in material science, biology, biochemistry and environmental scientists. The 3D structural characterization of spidroins through structural bioinformatics showed that these proteins have a β -hairpin structure, which is related to the folding and stability of proteins. The modeling and molecular dynamics showed that occurs the structural conformation change in the spidroins when they are exposed, for example in an aqueous system; and when in the presence of fatty acids, molecular dynamics suggests that there is a minor loss in the number of β -sheets. Thus, these findings may be valuable for understanding the physicochemical properties of the silk proteins and moreover, future designs of recombinantly produced spider silk proteins.

Key words: spider silk; spidroins; *bottom-up* proteomic approach; mass spectrometry; sequencing; post-translational modification; phosphorylation; 3D structure; molecular dynamic.

SUMÁRIO

	Página
1. INTRODUÇÃO	30
2. REVISÃO BIBLIOGRÁFICA	32
2.1. Aranhas construtoras de teias	32
2.2. Glândulas e tipos de sedas de aranhas	33
2.3. Aplicações biotecnológicas da seda	38
2.4. Características moleculares da seda	43
2.5. As modificações pós-traducionais dos fios da teia	48
2.6. Análise proteômica	51
2.7. Abordagens proteômicas	53
2.8. Espectrometria de massas	55
2.9. Bioinformática estrutural	58
3. OBJETIVOS	61
4. JUSTIFICATIVA	63
5. MATERIAL E MÉTODOS	64
5.1. Obtenção das glândulas e da seda da teia da aranha <i>Nephila clavipes</i>	64
5.2. Extração das proteínas das glândulas e da seda	64
5.3. Quantificação das amostras	64
5.4. 2-DE	65
5.5. Digestão <i>in gel</i>	65
5.6. Análises de espectrometria de massas	66
5.7. Identificação das proteínas	66
5.8. Identificação das modificações pós-traducionais (PTMs)	67
5.9. Tratamento com fosfatase alcalina	68
5.10. Tratamento com PNGase	68
5.11. Imunodeteção	68
5.12. Microscopia eletrônica de varredura (SEM)	69
5.13. Microscopia eletrônica de transmissão (MET)	69
5.14. Análises de ditirosina em hidrolisados da seda	70
5.15. Alinhamento múltiplo entre as sequências primárias	70
5.16. Predição da estrutura secundária	70
5.17. Modelagem molecular por homologia	70
5.18. Dinâmica molecular	73

5.19. Inserção e visualização dos sítios de fosforilação.....	74
5.20. Métodos de análise do modelo molecular.....	75
6. RESULTADOS E DISCUSSÃO.....	77
6.1. Parte 1.....	77
Análise proteômica das glândulas produtoras de seda da aranha <i>N. clavipes</i>	77
6.2. Parte 2.....	119
Análise estrutural e modificações pós-traducionais das espidroínas da seda da teia produzida pela aranha <i>N. clavipes</i>	119
6.2.1. Identificação e sequenciamento das espidroínas.....	119
6.2.2. Identificação das PTMs e aminoácidos substitutos.....	131
6.2.3. Caracterização de PTMs por NanoLC-MS/MS - abordagem proteômica.....	151
6.2.4. Caracterização estrutural 3D das espidroínas.....	159
7. RESUMO DOS RESULTADOS.....	187
8. CONCLUSÃO.....	189
9. REFERÊNCIAS BIBLIOGRÁFICAS.....	190
APÊNDICES - MATERIAL SUPLEMENTAR.....	208
Apêndices A - Tabelas S1, S2, S3 e S4.....	208
Apêndices B - Figuras S1-S63.....	248

1. Introdução

As espidroínas, proteínas da seda de aranhas, são secretadas por glândulas abdominais especializadas conectadas as fiandeiras que produzem fibras com aplicações em tarefas específicas como captura de presas, reprodução, e como linha de segurança “escape” para fugir de predadores (Vollrath, 2000). O frame e os raios da teia orbital são formados por fibras de seda predominantemente constituídas pelas espidroínas produzidas na glândula ampulada maior (Rising et al., 2007); enquanto que a espiral de captura é formada por fibras constituídas pelas espidroínas produzidas na glândula flageliforme (Römer et al., 2008).

As fibras produzidas pela glândula ampulada maior são um dos tipos mais importantes de fibras produzidas pelas aranhas do gênero *Nephila*, sendo um material nanoestruturado (Hardy et al. 2010) predominantemente composto por duas proteínas estruturais, espidroína-1 e espidroína-2 (Xu et al., 1990; Römer et al., 2008) as quais podem ser codificadas por diferentes genes (Rising et al., 2007; Ayoub et al., 2008). Enquanto que as fibras produzidas pela glândula flageliforme são compostas somente por uma proteína, a proteína da seda flageliforme. Essa seda é altamente elástica e em conjunto com as propriedades da seda do frame e do raio da teia orbital, dissipa perfeitamente o impacto da energia das presas. A enorme resistência dessas fibras é importante para a captura e detenção de presas que às vezes são maiores do que as aranhas (Hayashi et al., 2001; Römer et al., 2008; Lefèvre et al., 2011). A análise da sequência primária das espidroínas mostrou que essas proteínas são constituídas em três partes: um domínio N-terminal não repetitivo, uma região central altamente repetitiva composta por aproximadamente 100 segmentos de sequências ricas em polialanina/glicina, e um domínio C-terminal não repetitivo (Motriuk-Smith et al., 2005; Rising et al., 2011). As regiões não repetitivas, N- e C-terminal são conhecidas por desempenharem um papel importante na estabilidade das proteínas que são armazenadas no lúmen das glândulas até o momento de serem secretadas durante o processo de fiação (Askarieh et al., 2010; Hagn et al., 2011).

Devido a suas propriedades mecânicas e a combinação de resistência e extensibilidade, as fibras da seda de aranhas tem atraído cada vez mais o interesse em compreender as propriedades moleculares desse material, já que a exploração das aplicações da seda produzida pelas aranhas tem se mostrado potencialmente equivalente àquelas já conhecidas para a seda produzida pelo “bicho da seda”, a qual é utilizada em larga escala pela indústria têxtil, demonstrando também aplicações biomédicas e desenvolvimento de novos biomateriais (Kluge et al., 2008; Chen et al., 2010). Essas aplicações também podem ser estabelecidas para a seda de aranhas, oferecendo um grande potencial em aplicações biomédicas devido a uma combinação das propriedades mecânicas, biocompatibilidade e lenta biodegradabilidade (Altman

et al., 2003). O campo de aplicações para o uso da seda tem se tornado cada vez mais amplo, porém, as áreas mais promissoras têm sido as áreas médica e de engenharia de tecidos. Em estudos de Allmeling et al. (2006), fibras de seda de aranha foram coletadas e semeadas com células de *Schwann* humanas, para se investigar uma possível biocompatibilidade, sugerindo uma promissora estratégia para um futuro tratamento das lesões nervosas periféricas. Utilizados em conjunto, os biomateriais a base de seda tem se mostrado promissores na engenharia de tecidos esqueléticos como ossos, ligamentos e cartilagens, tecidos conjuntivos como a pele, atuando como fios de sutura mais finos e mais resistentes, curativos e suportes para a recuperação ou a substituição de tendões e ligamentos (Scheller et al., 2004). Estudos de biocompatibilidade *in vivo* e *in vitro* demonstraram que um recombinante, rS1/9 *scaffolds*, análogo à proteína espidroína-1 presente na seda pode ser potencialmente aplicável na regeneração de tecidos *in vivo* (Moisenovich et al., 2011; Moisenovich et al., 2012).

A seda da teia de aranha tem sido extensivamente estudada através de dados oriundos da engenharia genética e da tecnologia do DNA recombinante (Lewis, 2006; Rising et al., 2007; Lefèvre et al., 2011; Rising et al., 2011). Os dados sobre aminoácidos substitutos e modificações pós-traducionais (PTMs) são limitados. Até então, as propriedades mecânicas das espidroínas foram caracterizadas sem qualquer consideração sobre a existência de PTMs em suas sequências. Sendo assim, tanto as espidroínas naturais quanto as suas formas recombinantes (Liu et al., 2005; Koski et al., 2013) foram, até então, bioquímica e fisicamente caracterizados como se apresentassem as mesmas (ou similares) propriedades físico-químicas. Portanto, o presente trabalho representa o primeiro estudo com abordagem proteômica para um melhor conhecimento sobre as características estruturais e químicas das espidroínas ao determinar a sequência e presença/localização de PTMs nas espidroínas da seda da teia da aranha *N. clavipes*.

2. Revisão bibliográfica

2.1. Aranhas construtoras de teias

As aranhas formam um grupo monofilético, cuja monofilia é apoiada pela presença de apêndices abdominais modificados, como fiandeiras, glândulas sericígenas e fúsculas. A ordem Araneae, pertencente ao filo Arthropoda, contem aproximadamente 40 mil espécies de aranhas descritas incluídas em 108 famílias (Platnick, 2008). Dessas 108 famílias encontradas no mundo, 67 são encontradas no Brasil (Brescovit et al., 2004). Dentre a diversidade de aranhas existem as errantes (territoriais e caçadoras) e as construtoras de teias (ou tecedoras). Estas últimas desenvolveram teias como estratégia para sobrevivência e captura de alimentos (Blackledge et al., 2003).

Dentro da ordem Araneae, pertencente à família Nephilidae está a aranha *Nephila clavipes*, também conhecida como aranha construtora de teia aérea orbital, sendo encontrada nas três Américas. A família Nephilidae é composta por 58 espécies, cuja distribuição geográfica é predominantemente subtropical (Kuntner et al., 2008; Platnick, 2010). Esta família inclui quatro gêneros: *Clitaetra* (seis espécies que ocorrem na África e Índia), *Herennia* (11 espécies distribuídas na Austrália, Filipinas, Indonésia, Índia e China), *Nephila* (37 espécies procedentes da Austrália, Nova Zelândia, Filipinas, Indonésia, Paquistão, Índia, China, Japão, Estados Unidos da América e América do Sul), *Nephilengys* (quatro espécies registradas para Austrália, África, Índia, China e América do Sul). Até os dias atuais, existem apenas três espécies de Nephilidae taxonomicamente registradas que ocorrem no Brasil (Platnick, 2010), *Nephila clavipes* (Linnaeus, 1767), *Nephila sexpunctata* (Giebel, 1867) e *Nephilengys cruentata* (Fabricius, 1775). As *Nephila clavipes* são aranhas tecedoras (Figura 1) que constroem teias grandes circulares, às vezes chegando a 1 metro de diâmetro, com coloração amarelada e fios altamente resistentes, capazes de capturar pequenas aves como o beija-flor. São carnívoras, localizando-se nos espaços entre a vegetação e em áreas de florestas nas margens de rios. As teias são construídas estrategicamente entre os galhos das árvores, em uma altura em que costumam voar insetos, os quais ficam presos na teia servindo de alimento para as aranhas. Também são encontradas em áreas externas às moradias, construindo refúgio junto às paredes próximo de luminárias, o que facilita a captura de presas (Comstock, 1975).

No Brasil, estas aranhas são encontradas por toda a Mata Atlântica (Brescovit et al., 2004). Por ser um animal peçonhento, estas aranhas, naturalmente, podem picar o homem. Entretanto, não são animais agressivos, dificilmente abandonam suas teias e a ação de seu veneno não tem importância médica, causando no homem apenas vermelhidão e leve dor no local da picada (Allmeling et al., 2006).



Figura 1. *Nephila clavipes* - aranha construtora de teia aérea orbital coletada no campus da Universidade Estadual Paulista “Júlio de Mesquita Filho” localizado na cidade de Rio Claro, São Paulo.

2.2. Glândulas e tipos de sedas de aranhas

Durante milhões de anos de evolução, as aranhas tornaram-se eficientes produtoras de seda, atraindo o interesse humano para esse material que apresenta uma elevada estabilidade mecânica, biocompatibilidade, superfície lisa e fina e também por não causar inflamações e reações alérgicas em humanos (Gerritsen, 2000; Altman et al., 2003). Essas sedas são caracterizadas pela notável diversidade em sua composição química, estrutura e função, que vai desde a construção da teia orbital até o casulo. A seda produzida pelas aranhas são filamentos proteicos, constituindo uma interessante relação de estrutura-função e propriedades mecânicas, sendo produzida por um conjunto de glândulas abdominais. Existem vários tipos de glândulas e cada uma é responsável pela produção de um tipo específico de seda (Kovoor, 1987). Até então, tem sido descrito sete tipos diferentes de glândulas produtoras de seda; o número depende do gênero e espécie. As aranhas tecedoras de teias orbitais, como por exemplo, as aranhas do gênero *Nephila*, possuem um conjunto de glândulas abdominais composto por até sete tipos diferentes de glândulas (Figura 2): glândulas ampuladas maiores, ampuladas menores, flageliformes, cilíndricas ou tubuliformes, aciniformes, piriformes e agregadas (Rousseau et al., 2009). Acredita-se que essas diferentes glândulas evoluíram a partir de um único tipo e se divergiram na anatomia, no conteúdo do lúmen e na morfologia (Vollrath, 1992).

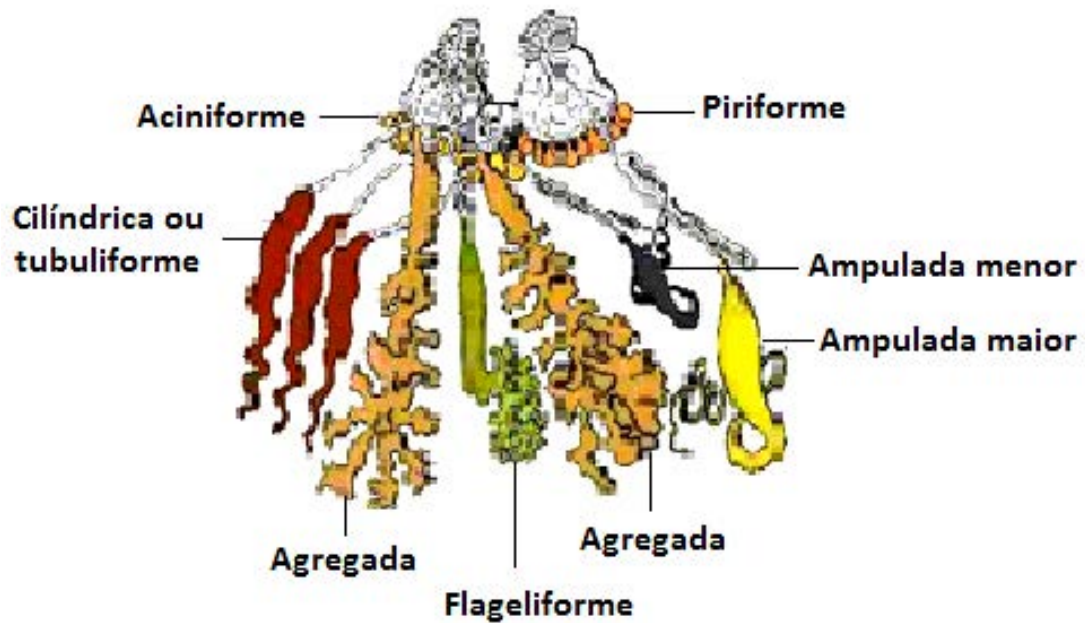


Figura 2. Esquema representativo das glândulas produtoras de seda das aranhas tecedoras de teias orbitais (adaptado de Vollrath, 2000). Tipos de fios de seda e suas respectivas glândulas produtoras.

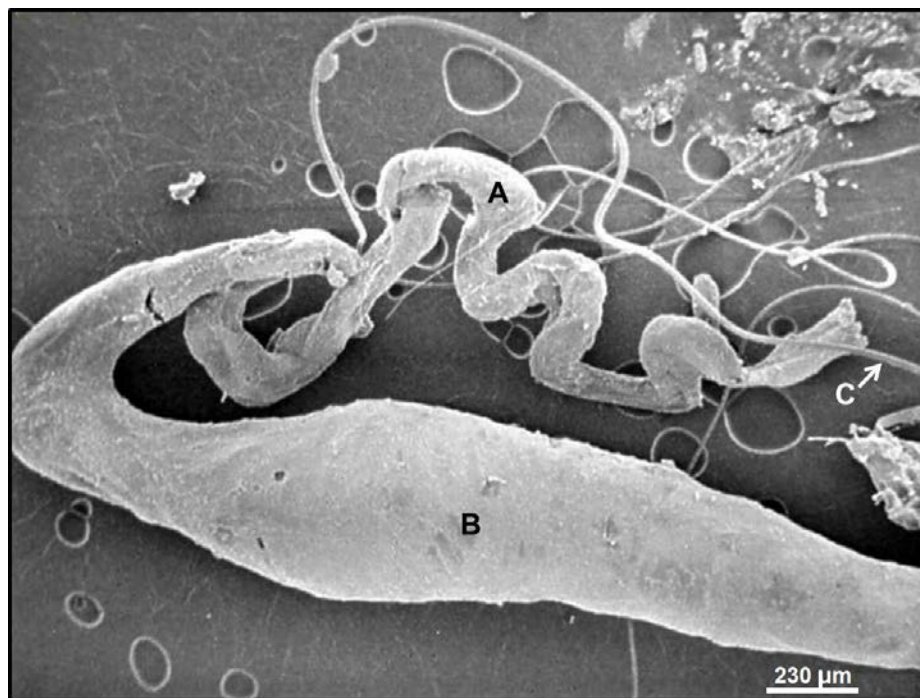


Figura 3. Microscopia eletrônica de varredura - MEV da glândula ampulada maior da aranha *N. clavipes*. A - cauda, B - ampola e C - ducto de fiação, indicado por uma seta (Imagem de Salles, 2003).

A glândula ampulada maior é a mais volumosa e de fácil acesso de todas as glândulas produtoras de seda, sendo, portanto a mais utilizada como modelo nos estudos para se compreender e decifrar as características morfológicas e funcionais do mecanismo de fiação. Estudos demonstraram que essa glândula é composta basicamente de três partes (Figura 3): a cauda, na qual as proteínas, chamadas espidroínas, são sintetizadas; a ampola, onde as proteínas são armazenadas em uma solução aquosa; e o ducto de fiação onde ocorre a formação das fibras (Knight et al., 2000; Lefèvre et al., 2011).

Estudos têm demonstrado que a cauda e a ampola da glândula ampulada maior apresentam três tipos diferentes de células epiteliais distribuídas em distintas regiões A, B e C (Figura 4). Essas células possuem diferentes tamanhos, posições específicas do núcleo, com grânulos de diferentes tamanhos, conteúdos e distribuições (Andersson et al., 2013). A maior parte do material localizado no lúmen da cauda, da ampola e do ducto de fiação provém da região A, de acordo com estudos que mostram que a seda é sintetizada na cauda da glândula (Plazaola et al., 1991; Andersson et al., 2013). A região A é composta por células colunares altas que secretam mucopolissacarídeos (Rising et al., 2005) na matriz aquosa de espidroínas, que nesta fase possui uma concentração de 25-30% m/v (Chen et al., 2002). Na região B, a matéria-prima da seda é revestida por uma solução viscosa secretada por células colunares que diferem ligeiramente em sua morfologia quando comparadas com as células colunares da região A (Vollrath e Knight, 2001). A secreção da região B é eletrodensa e aparece como uma camada contínua e distinta através do ducto de fiação até o momento da produção final da fibra, onde forma uma camada exterior que circunda a secreção oriunda da região A (Peakall, 1969; Sheibel, 2004; Andersson et al., 2013). Os grânulos da região C são mais eletrodensos do que qualquer outro componente das regiões A e B, sendo pouco provável que a secreção oriunda dessa região seja adicionada às secreções oriundas das regiões A e B, sendo retida nas microvilosidades das células (Andersson et al., 2013). A fase líquida cristalina, que é fundamental para a velocidade com a qual a aranha pode tecer a fibra de seda, possibilita que a suspensão de espidroínas possa fluir pela glândula ao mesmo tempo em que mantém uma orientação com as moléculas sendo organizadas em paralelo ao longo do eixo (Vollrath e Knight, 2001).

A habilidade de tecer uma fibra sólida a partir de uma solução altamente viscosa, porém aquosa, é única e incomparável na química dos polímeros. A aranha utiliza de um mecanismo que evita a auto-organização e a interação das proteínas antes de serem liberadas pelo ducto de fiação para a formação das fibras sólidas. Para compreender esse mecanismo, e dessa forma, reproduzir como a aranha produz a seda é necessário conhecimento não só sobre as

propriedades mecânicas das fibras da seda, mas também como as espidroínas que constituem as fibras são secretadas e armazenadas nas glândulas, e ainda como essa seda é expelida em uma fibra sólida a partir de uma solução líquida cristalina de fiação, sem usar solventes extremos ou outras condições ambientais.

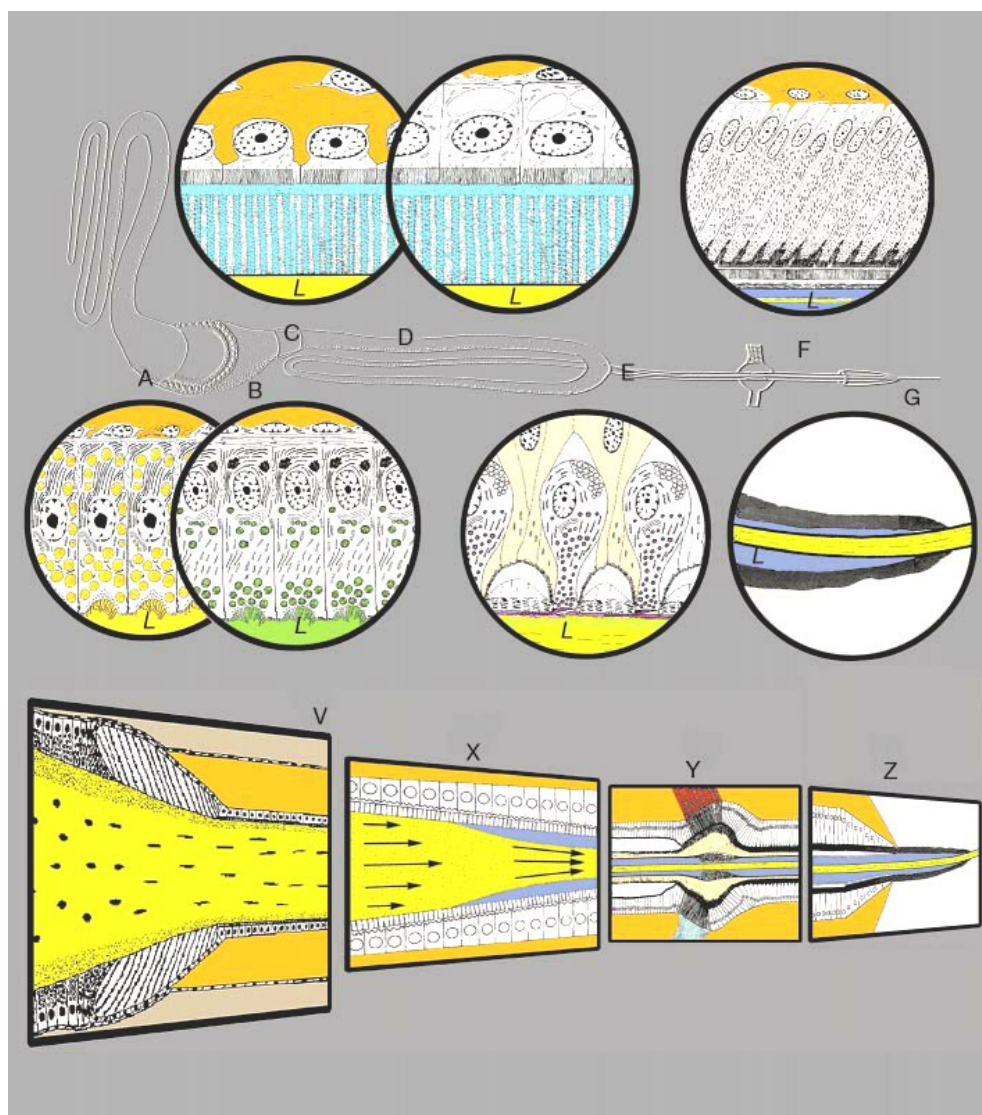


Figura 4. Esquema representando a histologia da glândula ampulada maior (A - G) destacando o processo de fiação (L - lúmen). A, B e C - produção das espidroínas e secreção de solução viscosa por células colunares e armazenamento; D, E, F e G - histologia do ducto de fiação. A região do ducto em F parece ser altamente especializada para bombeamento de água, contendo células altas com numerosas mitocôndrias, microvilosidades apicais e uma ampla superfície formada pelas vilosidades internas da membrana plasmática. V - Z, demonstram a natureza afunilada do extenso ducto de fiação. As funções do estreitamento da glândula são reter água, alinhar as proteínas, polimerização e produção de uma fibra mecanicamente equilibrada com custos metabólicos mínimos (Adaptado de Vollrath e Knight, 2001).

Possivelmente, as “sequências motivos” das espidroínas possam auxiliar na compreensão de todo o processo utilizado pelas aranhas durante o mecanismo de fiação. Essas sequências são divididas em domínios hidrofílicos e domínios hidrofóbicos, e o arranjo desses domínios permite que uma solução líquida e cristalina possa ser processada em uma fibra sólida, ao mesmo tempo em que se impede as interações prematuras que fariam com que as proteínas da seda se precipitassem. Os aminoácidos hidrofóbicos são agrupados para eliminar a água à medida que a solução começa a fluir pelo ducto de fiação, com isso cristais de folhas- β são formados para se produzir fibras fortes e insolúveis (Knight e Vollrath, 2002). Para manter a solubilidade da solução de fiação, domínios hidrofílicos circundam uma região central principalmente hidrofóbica (Bini et al., 2004). Além da interação entre os domínios hidrofílicos, domínios hidrofóbicos e o ambiente aquoso, as regiões N- e C-terminal não repetitivas e altamente conservadas das proteínas também contribuem para a formação final da fibra (Sponner et al., 2005; Eisoldt et al., 2011; Andersson et al., 2013; Schwarze et al., 2013).

Os diferentes tipos de fios de seda são comumente nomeados e diferenciados após serem secretados pelas respectivas glândulas (Figura 5); e possuem aplicações em tarefas específicas como captura de presas, reprodução, e como linha de segurança “escape” para fugir de predadores (Vollrath, 2000). Essa organização em um complexo de glândulas para produção de seda apresentado pelas aranhas, não apenas permite que a seda seja produzida para usos específicos como também, proporciona a produção de múltiplas fibras simultaneamente. Mais de 130 formas diferentes de teias de aranha são conhecidas. Os frames e os raios da teia orbital são construídos a partir de uma seda forte e bastante rígida, apresentando uma combinação excepcional de resistência e extensibilidade (Vollrath, 2000). As proteínas dessa seda são produzidas nas glândulas ampuladas maiores, sendo as fibras de seda denominadas de MA. A espiral de captura de uma teia orbital compreende fibras de seda de apenas um tipo de proteína produzida na glândula flageliforme. Essa seda é altamente elástica e dissipa perfeitamente a energia produzida pelo impacto das presas se colidindo contra teia. A enorme resistência dessas fibras é importante para a captura e detenção de presas, que às vezes são maiores do que as próprias aranhas (Römer et al., 2008).

Além das fibras de seda MA e Flageliformes, as aranhas utilizam ainda as fibras de seda produzidas de uma glândula ampulada menor (MI) para auxiliar a construção da espiral de captura. As propriedades mecânicas da seda MI são semelhantes aos da seda MA, apesar de sua resistência ser menor e a elasticidade maior do que a da seda MA (Blackledge et al., 2006). As conexões entre as sedas e a fixação da teia sobre um substrato (árvores, arbustos,

alvenaria, etc.), são realizadas por uma fibra de seda sofisticada, composta por proteínas produzidas pela glândula piriforme (Gosline, 1986; Römer et al., 2008).

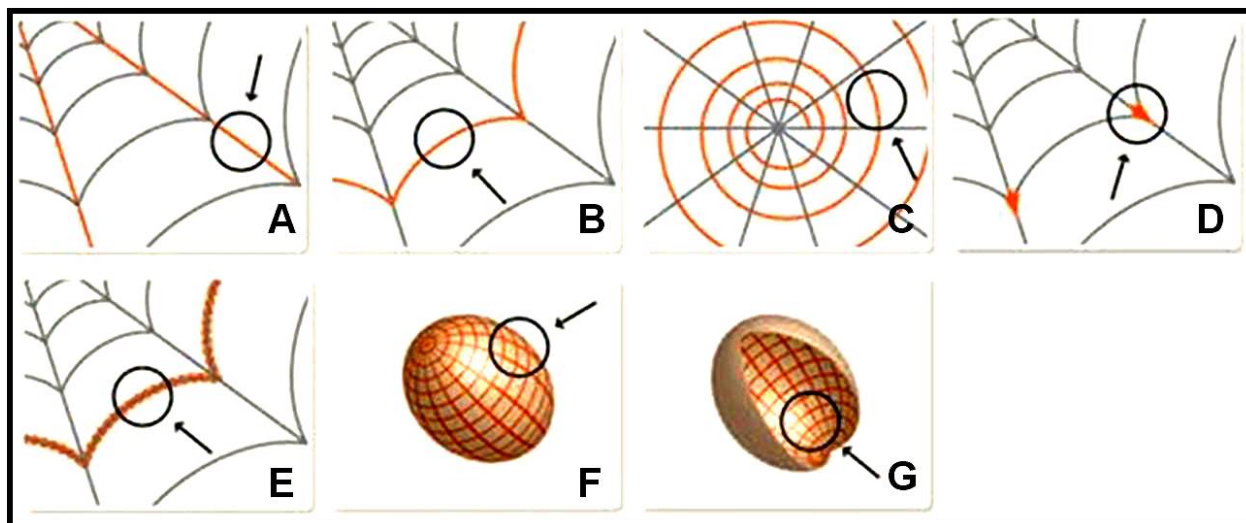


Figura 5. Tipos de fios de seda secretados pelas aranhas tecedoras de teias orbitais. **A** - glândula ampulada maior secreta os fios de seda dos raios da teia orbital, do frame e da linha de segurança "escape"; **B** - glândula flageliforme secreta os fios de seda da espiral de captura; **C** - glândula ampulada menor secreta os fios de seda que auxiliam a espiral de captura; **D** - glândula piriforme secreta os fios de seda das conexões entre as sedas e a fixação sobre um substrato; **E** - glândula agregada secreta um revestimento adesivo aquoso sobre a espiral; **F** - glândula tubuliforme secreta os fios de seda que cobre/protege os ovos externamente (bolsa de seda) após a postura; **G** - glândula aciniforme secreta os fios de seda que protege os ovos internamente (adaptado de Römer et al., 2008).

Apesar do grande interesse pelas propriedades mecânicas da seda das aranhas, visando o seu uso em aplicações biomédicas e biotecnológicas devido as suas propriedades de resistência, elasticidade e biocompatibilidade (Vollrath et al., 2002), surpreendentemente pouco se conhece sobre os detalhes das glândulas produtoras de seda e o processo de fiação que ocorre para a produção das fibras. Recentemente, alguns estudos têm abordado não só as propriedades mecânicas da seda, mas também como se dá a produção natural das espidroínas constituintes da seda, o processo de fiação e ainda a morfologia e composição das glândulas produtoras de seda (Eisoldt et al., 2011; Slotta et al., 2012; Andersson et al., 2013).

2.3. Aplicações biotecnológicas da seda

A exploração das aplicações da seda produzida pelas aranhas tem se mostrado potencialmente equivalente àquelas já conhecidas para a seda produzida pelo "bicho da seda", que tem sido domesticado a mais de cinco mil anos (Kluge et al., 2008). Atualmente a seda do "bicho da seda" é utilizada em larga escala pela indústria têxtil, demonstrando também aplicações biomédicas e desenvolvimento de novos biomateriais (Chen et al., 2010). Essas

aplicações também podem ser estabelecidas para a seda de aranhas, oferecendo um grande potencial em aplicações biomédicas devido à combinação de propriedades mecânicas, biocompatibilidade e lenta biodegradabilidade (Altman et al., 2003).

Diferente do “bicho da seda”, as aranhas não foram domesticadas para produzirem fibras em escala industrial. Essa diferença deve-se à dificuldade de se obter grandes populações de aranhas, devido a sua natureza solitária e predadora. Além disso, as sedas de aranhas são produzidas em pequena quantidade e não podem ser enoveladas como uma fibra simples, de forma semelhante ao que ocorre com o casulo do “bicho da seda” (Altman et al., 2003). Entretanto, as características físicas das sedas de aranhas são muito superiores às encontradas nas sedas de *Bombix mori*. As sedas de aranhas são biomateriais leves, extremamente fortes e elásticas, e apresentam características mecânicas comparáveis às melhores fibras sintéticas já produzidas pela alta tecnologia (Hinman et al., 2000). A tabela 1 mostra a comparação entre as diferentes propriedades mecânicas de alguns tipos de sedas produzidas pelas aranhas do gênero *Nephila* em relação a alguns materiais naturais ou sintéticos.

Tabela 1. Comparação entre as diferentes propriedades mecânicas de sedas de aranhas, seda do bicho da seda, fibras naturais e polímeros sintéticos (adaptado de Hinman et al., 2000; Lewis, 2006).

Tipos de materiais	Tensão (Pa)	Extensibilidade (%)	Energia para romper (J/Kg)
Seda produzida pela glândula ampulada maior	4×10^9	35	1×10^5
Seda produzida pela glândula ampulada menor	1×10^9	5	3×10^4
Seda produzida pela glândula flageliforme	1×10^9	>200	1×10^5
Seda produzida pelo “bicho da seda”	1.3×10^9	-----	-----
Kevlar	4×10^9	5	3×10^4
Borracha	1×10^6	600	8×10^4
Nylon tipo 6	7×10^7	200	6×10^4

Diversos estudos publicados comparam a seda de aranhas com materiais sintéticos mais resistentes, como por exemplo, o kevlar, uma fibra de poliamida aromática produzida pela empresa DuPont. O kevlar é uma inovação tecnológica, e assim como as fibras da seda de aranhas, é extremamente resistente e apresenta baixo peso molecular por isso é de grande interesse industrial, podendo ser produzida de diferentes formas dependendo de sua aplicação.

Sendo assim, é possível que alguns materiais à base de biopolímeros de seda de aranha possam ser desenvolvidos com uma grande aceitação e aplicabilidade (Hinman et al., 2000; Lewis, 2006; Slotta et al., 2012).

Considerando as suas propriedades mecânicas e físico-químicas, a seda de aranha é perfeitamente adequada para inúmeras aplicações industriais. No entanto, nenhum produto à base de seda de aranha está disponível comercialmente. A produção em larga escala das fibras da seda de aranha tornaria possível a elaboração de uma nova geração de biomateriais com elevado grau de biodegradabilidade, envolvendo aplicações práticas em diferentes segmentos do setor industrial. Alternativamente a inabilidade de domesticar as aranhas, de forma a produzirem quantidades necessárias e suficientes de proteínas para propiciar estudos e utilização comercial, a seda de aranha tem sido produzida como proteínas recombinantes utilizando a tecnologia do DNA recombinante e a engenharia de organismos hospedeiros (Hardy et al., 2008; Rising et al., 2011; Eisoldt et al., 2011; Cranford et al., 2012; Slotta et al., 2012). No entanto, informações químicas como as modificações pós-traducionais podem ser perdidas, o que pode inviabilizar a manutenção das propriedades mecânicas e físico-químicas tão características da seda de aranha.

O campo de aplicações para o uso da seda tem se tornado cada vez mais amplo, porém as áreas mais promissoras têm sido as áreas médica e de engenharia de tecidos. Em estudos de Allmeling et al. (2006), fibras de seda de aranha foram coletadas e semeadas com células de Schwann humanas, para se investigar uma possível biocompatibilidade, sugerindo uma promissora estratégia para um futuro tratamento das lesões nervosas periféricas. Utilizados em conjunto, os biomateriais a base de seda tem se mostrado promissores na engenharia de tecidos esqueléticos como ossos, ligamentos e cartilagens, tecidos conjuntivos como a pele, atuando como fios de sutura mais finos e mais resistentes, curativos e suportes para a recuperação ou a substituição de tendões e ligamentos (Scheller et al., 2004). Estudos de biocompatibilidade *in vivo* e *in vitro* demonstraram que um recombinante, rS1/9 *scaffolds*, análogo à proteína espidroína-1 presente na seda pode ser potencialmente aplicável na regeneração de tecidos *in vivo* (Moisenovich et al., 2011; Moisenovich et al., 2012).

As proteínas da seda têm se mostrado totalmente solúveis em água, solventes orgânicos e líquidos iônicos, indicando uma versatilidade de opções disponíveis na qual podem ser transformadas em novos biomateriais como fibras, filmes, géis, esponjas porosas e outros sistemas de aplicabilidade biomédica (Vepari et al., 2007). E finalmente, a seda de aranha pode ser aplicada também na indústria têxtil. Segundo pesquisadores da indústria têxtil DuPont, o fio de seda da aranha é tão resistente que seria capaz de “parar um Boing 747 em pleno vôo”, e ao

mesmo tempo, ser tão maleável que poderia ser utilizado como tecido para roupas (Fiber engineers, meet thy master, 1996). Estudos têm demonstrado a eficácia da utilização dessa seda para coletes à prova de balas, tornando-os mais leves, flexíveis e resistentes; assim como componentes para aviões e pára-choques de automóveis (Teia de aranha pode ser matéria prima para a indústria, 2002). A utilidade para os fios da seda são os mais variáveis possíveis, desde a confecção de casco de navios, airbags para automóveis, fibras ópticas até lingerie confortáveis (Helmer, 2001). Outras aplicações se referem a produtos cosméticos como xampu, sabonete e creme, realçando o brilho, maciez e resistência do produto (Yang et al., 2005; Slotta et al., 2007).

A AMSilk® *high performance materials* é uma empresa alemã, fundada em 2008, que desenvolveu um processo patenteado para a produção de seda de aranhas em escala industrial, com o desenvolvimento de novos produtos com qualidade superior e características anteriormente inatingíveis, com base nas propriedades mecânicas e físico-químicas da seda de aranhas. As proteínas da seda AMSilk mimetizam as espidroínas que constituem as fibras da seda de aranhas, com diversas propriedades desejáveis que não são encontradas em comum com o “bicho da seda” *Bombyx mori*. Essas proteínas são não imunogênicas, biocompatíveis e possuem propriedades únicas de revestimento; e podem ser produzidas com um elevado grau de pureza e qualidade. A tecnologia AMSilk® abrange diferentes setores e aplicações, incluindo aplicações biomédicas, farmacêuticas e desenvolvimento de fibras.

Recentemente a AMSilk® desenvolveu uma fibra única produzida através de proteínas recombinantes da seda de aranhas. Essa fibra foi apresentada com a marca de *Biosteel* (Figura 6), que como outras fibras sintéticas pode ser fabricada como um monofilamento ou multifilamento. Trata-se de uma fibra branca com um brilho sedoso e permeável à água; mostra-se suave sobre a pele e, como a seda normal, pode ser processada e corada. Utilizando métodos biotecnológicos, a matéria-prima - inicialmente na forma de pó - é produzida em grande quantidade e alta qualidade do que jamais poderia ser produzida pelas aranhas. O processo é baseado nas descobertas do Professor Dr. Thomas Scheibel da Universidade de Bayreuth, Alemanha, que foram desenvolvidas tecnologicamente pela AMSilk®.

The Biosteel Spidersilk Fibers permite que as propriedades versáteis da seda de aranhas possam ser utilizadas comercialmente. As características únicas das proteínas da seda podem ser adaptadas e orientadas para a fabricação de produtos sob medida. Estudos mostraram que o material é adequado para aplicações biomédicas, como por exemplo, material de sutura, curativos e outras aplicações cirúrgicas.



Figura 6. *Biosteel Spidersilk Fibers* produzida pela AMSilk® *high performance materials*.

Em paralelo com o desenvolvimento da fibra *Biosteel*, a AMSilk® está expandindo a produção de matérias-primas para apoiar o desenvolvimento de aplicações em diversas áreas, buscando uma ampla gama de produtos baseados na plataforma *Spidersilk*, incluindo revestimentos de implantes mamários de silicone, filmes, tecidos, grânulos, fibras e cosméticos a base de seda. Estudos pré-clínicos mostraram que os implantes mamários de silicone são mais bem aceitos pelo corpo quando revestido com *BioShield-S1*, um filme especial de seda de aranha não imunogênica. Especificamente, a fibrose capsular e a inflamação são significativamente reduzidas. A empresa está desenvolvendo também um novo conceito de tratamento de ferimentos baseado nas proteínas da seda de aranhas. Um produto denominado de *SanaSilk* está sendo desenvolvido para efetivamente hidratar a pele e criar um filme de seda com a função de proteger a pele. Um outro produto é o *TruSilk® Deluxe Line* um cosmético inovador para a pele que contém verdadeiras proteínas oriundas da seda de aranhas que transfere os seus benefícios para a pele. Misturado com ingredientes naturais, o creme protege a pele deixando uma camada lisa e macia de seda sobre a pele com efeitos de equilíbrio da umidade; dessa forma a pele mostra-se fresca, macia e com brilho. Recentemente, em novembro de 2013 a AMSilk® começou a disponibilizar a seda de aranhas (Figura 7) como ingrediente para cosméticos sendo a primeira seda funcional no mercado testada e aprovada para formulações de cosméticos.

A AMSilk® é uma marca registrada da AMSilk GmbH, Alemanha, União Europeia e os EUA que vem crescendo muito desde a sua fundação e atualmente possui mais de 18 aplicações de patentes, das quais diversas já foram concedidas. Apresenta diversas publicações sobre as mais diversas aplicabilidades para a seda de aranhas (Hardy et al., 2009; Lammel et al., 2011; Spiess et al., 2010; Hagn et al., 2011; Humenik et al., 2011; Eisoldt et al.,

2011; Leimer et al., 2012) e para maiores informações sobre a AMSilk® pode ser consultado o site: <http://www.amsilk.com>



Figura 7. Cosméticos produzidos pela AMSilk® *high performance materials* contendo as propriedades da seda de aranhas comercialmente disponíveis. (http://www.amsilk.com/fileadmin/content/TechInsight_Cosmetics.pdf).

Como pode ser notado, a seda da teia de aranhas se encaixa perfeitamente dentro da prática bioprospectiva do processo de transformação de recursos biológicos para a obtenção de produtos com uma aplicabilidade biotecnológica possibilitando o desenvolvimento de diversos produtos baseados nas propriedades mecânicas e físico-químicas da seda.

2.4. Características moleculares da seda

As proteínas que constituem a seda de aranha são compostas por uma grande quantidade de resíduos de aminoácidos apolares e hidrofóbicos como glicina e alanina, apresentando muito pouco ou nenhum resíduo de triptofano (Rising et al., 2005). Em comparação com as enzimas celulares comuns, as proteínas da seda possuem uma composição diferenciada de aminoácidos, apresentando sequências altamente repetitivas. Essas sequências representam mais de 90% da proteína e são compostas por fragmentos polipeptídicos com cerca de 10 a 50 resíduos de aminoácidos. Cada polipeptídeo possui características funcionais distintas resultando no entendimento das propriedades mecânicas da seda (Lewis, 2006).

A estrutura molecular da seda é composta por regiões organizadas de proteínas, separadas por regiões proteicas menos organizadas. As estruturas primárias originam diversas

estruturas secundárias que, por sua vez, direcionam as funções dos diferentes tipos de seda. A sequência primária favorece a organização da estrutura secundária da proteína em folhas- β cristalinas, que contribuem significativamente para a alta resistência das fibras da seda (Kluge et al., 2008). A formação de folhas- β ocorre através de uma reticulação física natural das sequências repetidas de aminoácidos, principalmente Ala, Ala-Gly ou Ala-Gly-Ser. As regiões menos organizadas da seda, que favorecem a sua elasticidade, são comumente formadas por espiral- β , similar a β -turn, compostas por sequências repetidas de sequência GPGXX, sendo X principalmente um resíduo de glutamina; e por estruturas helicoidais compostas de sequência GGX (Kaplan et al., 1997). Cada repetição de polipeptídeo apresenta distintas características funcionais resultantes das excelentes propriedades mecânicas da seda. As sedas produzidas pelas glândulas MA e Flageliformes, por exemplo, são formadas por até quatro sequências, que se repetem: $(GA)_n/(A)_n$, GPGGX/GPGQQ; GGX (X = A, S ou Y) e uma sequência “espaçadora” composta por resíduos de aminoácidos carregados. Análises estruturais da sequência $(GA)_n/(A)_n$ mostrou que essa sequência tende a formar estruturas hélices- α em solução correspondente ao meio aquoso presente no interior das glândulas e folhas- β nas fibras sólidas (Figura 8). Pouco se sabe sobre as estruturas constituídas pelas sequências GPGGX/GPGQQ e GGX, porém, alguns estudos descrevem essas regiões como estruturas amorfas, enquanto outros estudos sugerem a formação de uma estrutura 3_1 -helicoidal ou estruturas espiral- β , no caso da seda produzida pela flageliforme (Hu et al., 2006; Römer et al., 2008).

Estudos utilizando técnicas de RMN e difração de raios-X demonstraram que os motivos de polialanina são definidos como cristais nanométricos (2x5x7 nm) com base no empacotamento antiparalelo das folhas- β (Creager et al., 2010), no qual é muito provável que as extensões de polialanina a partir de diferentes moléculas da seda constituem ligações cruzadas não covalentes muito fortes. E estudos de RMN forneceram evidências de que a estrutura constituída pelo motivo GGX pode formar folhas- β bem como estruturas helicoidais menos organizadas (Lefevre et al., 2007; Jenkins et al., 2010), enquanto que a estrutura constituída pelo motivo GPGXX, devido ao resíduo de prolina, forma claramente voltas- β que quando repetidos produzem uma estrutura em espiral como sugerido para a elastina (Hayashi et al., 1999; Eisoldt et al., 2011).

Evidências indicam que dentro da glândula as sequências de polialanina formam estruturas hélice- α , enquanto que as regiões ricas em glicina adquirem uma conformação β -turns ou randômica (Vollrath et al., 2001). Essas sequências passam por mudanças bioquímicas e físicas em suas estruturas secundária e terciária durante a formação das fibras da seda. A estrutura final da fibra é alcançada após a remoção dos íons sódio e cloreto e de

água durante a passagem da seda (em estado líquido) pelo ducto de fiação. Ocorre também a secreção de íons potássio e íons hidrogênio, ocasionando uma mudança do pH do meio, e dessa forma ocorre uma alteração da conformação da molécula durante a passagem da solução de seda da glândula, até a saída do fio pelas fiandeiras (Knight et al., 2001).

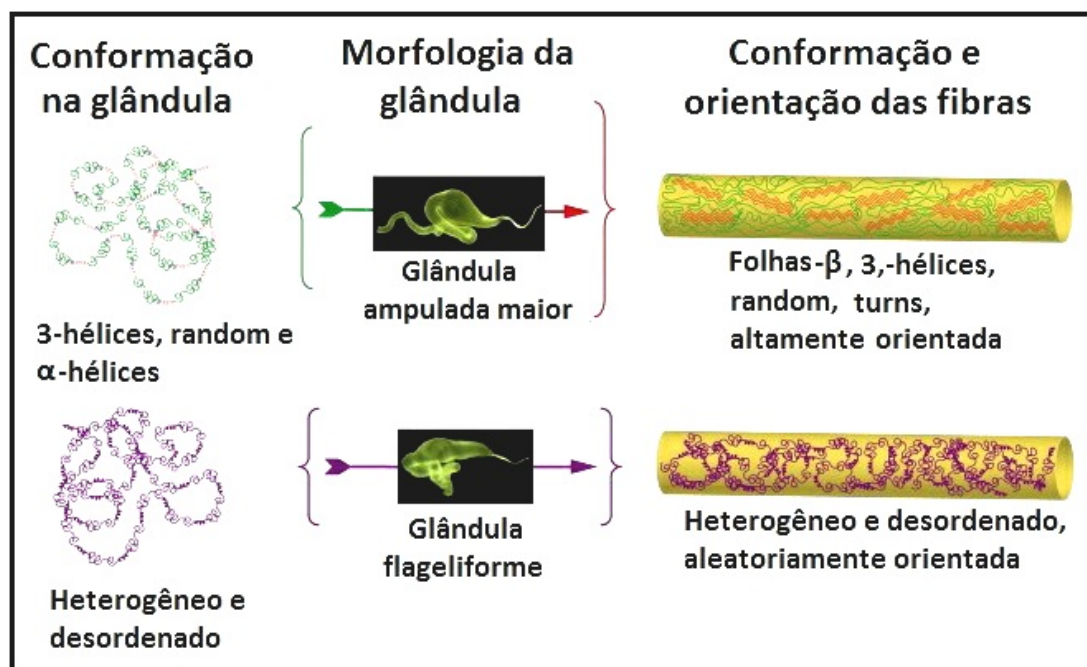


Figura 8. Esquema representativo da estrutura secundária das proteínas constituintes das fibras da seda da aranha *N. clavipes*, secretadas pelas glândulas ampulada maior e flageliforme. Esse esquema mostra a conformação estrutural das proteínas na glândula e a conformação estrutural obtida após serem secretadas durante o processo de fiação (modificado de Lefèvre et al., 2011).

O processo de fiação, que é controlado pelo sistema de glândulas e o comportamento do animal, é um dos principais determinantes sobre a estrutura das fibras. Duas teorias sobre o processo de fiação na formação das fibras de seda têm sido propostas. A primeira teoria é baseada no alinhamento cristalino subjacente das proteínas no fluxo laminar dentro do ducto de fiação. Esse alinhamento direcionado e a elevada concentração de proteínas resultam em um líquido cristalino, e essa proposta é a base para a formação de interações moleculares como força de *van der Waals* e pontes de hidrogênio entre as moléculas. Após a perda de solvente, a conversão conformacional é finalizada com a fibra sendo liberada pelo ducto de fiação. Na segunda teoria (Figura 9), as proteínas estariam em baixa concentração, ocorrendo uma organização de pequenas micelas antes do alongamento e formação das fibras ao passarem pelo ducto de fiação (Vollrath et al., 2001; Ko et al., 2004; Römer et al., 2008; Sheibel et al., 2009; Eisoltd et al., 2011).

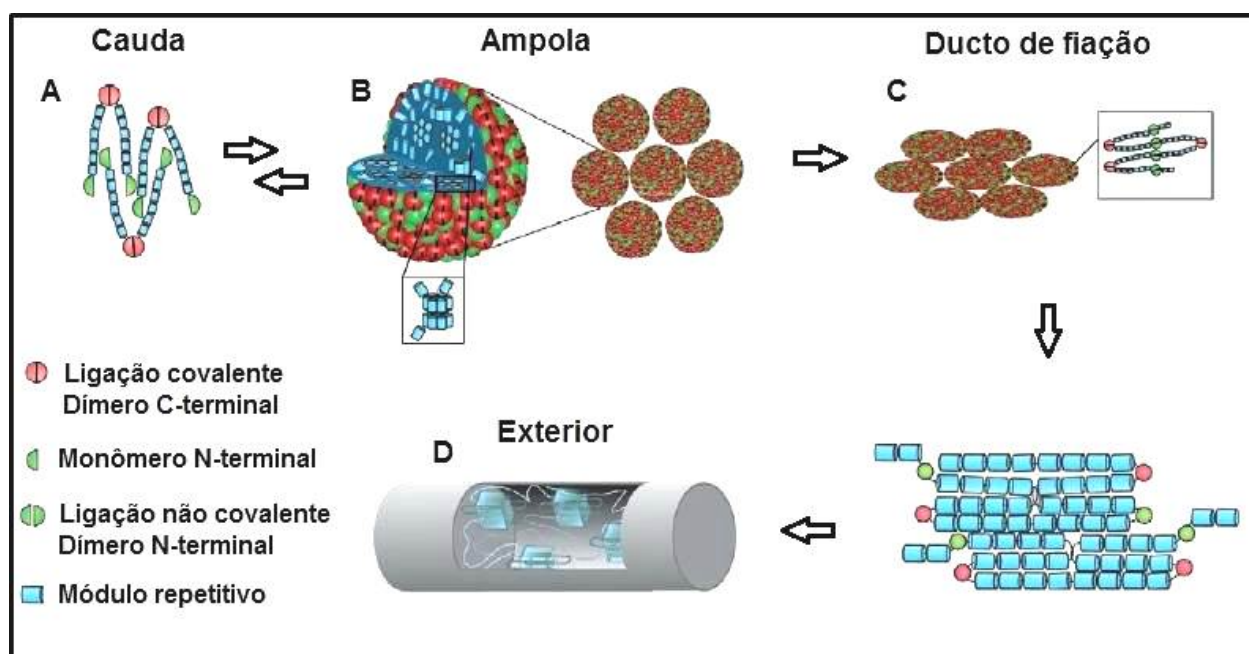


Figura 9. Esquema representando a teoria micelar sobre o processo de formação das fibras de seda, que ocorre ao longo dos aparatos do mecanismo de fiação (adaptado de Eisoldt et al., 2011). **(A)** Cauda - baixa concentração de proteínas; **(B)** Ampola - as proteínas são parcialmente pré-alinhadas com a formação de pequenas micelas apresentando propriedades de líquido-cristalino; **(C)** Ducto de fiação - alongamento e formação das fibras, acidificação do meio induzindo a dimerização do domínio N-terminal, troca iônica e remoção de água, alinhamento e formação das folhas-β, desdobramento do domínio C-terminal expondo as regiões hidrofóbicas e facilitando o dobramento das fibras; **(D)** Exterior - fibra secretada, folhas-β são alinhadas ao longo do eixo da fibra.

A concentração das proteínas durante o armazenamento na ampola, em estudos realizados com a glândula ampulada maior, pode alcançar até 50% (m/v) (Vollrath e Knight, 2001). Esta é notavelmente uma concentração elevada, uma vez que em tais condições a maioria das proteínas possui a tendência de se agregar irreversivelmente. Para evitar agregações inespecíficas, as espidroínas formam estruturas micelares controladas por sua característica anfílica (Exler et al., 2007; Hagn et al., 2010) e por suas regiões N- e C-terminal não repetitivas, as quais induzem ou suportam a formação de agrupamentos supramoleculares (Askarieh et al., 2010; Hagn et al., 2010). O líquido cristalino é exposto a um fluxo constante de alongamento permitindo que as moléculas possam fluir de uma maneira pré-alinhada ao longo do eixo do fluxo do ducto de fiação levando a formação dos cristais de folhas-β (Exler et al., 2007; Hagn et al., 2010; Eisoldt et al., 2010). No final do ducto de fiação, as forças de tração aumentam ainda mais ao puxar as fibras para o exterior, levando a uma mudança estrutural da espidroína, principalmente nas regiões terminais, acompanhado de uma exposição das superfícies hidrofóbicas e uma separação de fases entre o solvente (água que é ativamente

removida pelas células epiteliais) e as espidroínas (Knight et al., 2000; Rammensee et al., 2008). As mudanças estruturais nas regiões terminais reorganizam a posição das regiões centrais dentro das estruturas micelares, e isso, juntamente com as forças mecânicas suportam a separação de fases e o dobramento das espidroínas (Eisoldt et al., 2011).

Embora a composição proteica para todos os tipos de seda de aranha não tenha sido determinada, estudos tem demonstrado que os fios de sedas de espécies diferentes de aranhas parecem ser constituídos por diferentes composições proteicas (Renault et al., 2009). Cada organismo produtor de seda apresenta uma rica diversidade de sequências primárias e estruturas secundárias. Por exemplo, os aminoácidos constituintes das fibras da seda sintetizada pela aranha *Araneomorphae* podem ser agrupadas em quatro categorias: poli-Ala, poli-Ala-Gly, GPGXX, GGX e um espaçador de sequência (Hayashi et al., 1999). Garb et al. (2007) caracterizaram seis novas proteínas constituintes da seda da tarântula *Mygalomorphae*, que não contêm essas quatro categorias encontradas em *Araneomorphae*. Essas proteínas caracterizadas a partir da seda da tarântula não apresentam uma elevada resistência e elasticidade, o que reforça a hipótese da importância dos fragmentos poli-Ala e GPGXX na formação das sedas MA e flageliforme.

As sedas MI e MA que formam os raios da teia orbital e os frames, são exemplos bem caracterizados demonstrando serem constituídas por duas proteínas, MiSp1 e MiSp2 (*Minor ampullate Spidroin-1 e -2*); e MaSp1 e MaSp2 (*Major ampullate Spidroin-1 e -2*) respectivamente, cujas sequências foram determinadas principalmente para a aranha *Nephila clavipes* (Rousseau et al., 2009). Essas proteínas são ricas em resíduos de alanina (polialanina) e glicina, compartilhando o mesmo padrão de repetição de sequência. Devido a isto, essas proteínas formam um biomaterial semicristalino, constituído de cadeias amorfas e flexíveis, reforçadas por pequenos cristais sólidos (Simmons et al., 1996). As regiões cristalinas formadas pelos fragmentos de alanina estão dispostas em folhas- β antiparalelas sendo responsáveis pela elevada elasticidade e resistência das sedas MA e MI. A abundância dos resíduos de glicina é para fornecer extensibilidade ao segmento (Sponner et al., 2005).

Além da constituição proteica, a teia da aranha *N. clavipes* apresenta gotículas viscosas, que desempenham um papel importante na captura de presas, constituídas por diferentes compostos químicos em uma solução aquosa e uma camada lipídica que apresentam uma suspensão de vesículas contendo polipeptídeos e/ou lipídios (Salles et al., 2006). Volsi et al. (2006) relatam a identificação e a caracterização bioquímica de três peptídeos relacionados à bradicinina (nephilacininas I a III), a partir do extrato aquoso de lavagem da teia da aranha *N. clavipes*, aparentemente utilizados como toxinas para captura de presas.

A glândula MA produz fibras de sedas para diversas funções, incluindo os raios da teia orbital, os frames e o fio de segurança para fuga contra predadores. Considerando que cada fibra de seda possui uma função específica e que a aranha pode controlar sua velocidade de produção, e variar o diâmetro dos fios (Vollrath et al., 2001), as fibras reproduzidas pelas mesmas glândulas tendem a apresentar algumas variações na orientação e estrutura molecular (Rousseau et al., 2009) caracterizadas pelas modificações pós-traducionais (PTMs).

2.5. As modificações pós-traducionais dos fios da teia

As modificações pós-traducionais das proteínas constituintes da seda oriunda das glândulas da aranha podem aumentar a solubilidade das proteínas produzidas (Vend्रेly et al., 2007). As proteínas secretadas demonstram de fato uma solubilidade diferente, porém pouco ou nenhum estudo de análises detalhadas da presença e/ou influência das modificações pós-traducionais tem sido realizados. Modificações como fosforilação (Michal et al., 1996) e glicosilação (Weiskopf et al., 1996; Sponner et al., 2007) têm sido relatadas na seda de aranhas, no entanto, até então não se conhecia a exata localização/posição dessas e de outras modificações na sequência das espidroínas.

As PTMs consistem em modificações químicas de proteínas após a tradução, uma ampla estratégia utilizada por células eucarióticas para modificar rapidamente, a atividade local e específica de fatores essenciais em resposta às mudanças ambientais. Estas modificações permitem uma diversificação das atividades das proteínas codificada por todos os organismos aumentando significativamente a diversidade e a complexidade do proteoma (Walsh et al., 2005; Ribet et al., 2010). A determinação de PTMs é fundamental para a elucidação de processos que controlam os eventos celulares, como crescimento, divisão e diferenciação celular, incluindo a interação entre proteínas (Seo et al., 2008). Essas PTMs frequentemente ocorrem em múltiplos sítios, podendo ocorrer em muitos resíduos de aminoácidos do mesmo grupo; ou em uma série de cascatas que são cruciais para a função da proteína (Walsh et al., 2005). Apesar da grande importância da função biológica das modificações pós-traducionais, o seu estudo em larga escala é dificultado pela ausência de métodos apropriados, e muitas das modificações só foram descobertas mais tarde com a elucidação de vários processos biológicos (Jensen, 2004). Mais de 300 PTMs são conhecidas atualmente. Elas incluem a adição de grupos químicos como, por exemplo, o fosfato ou acetato ou moléculas mais complexas, como carboidratos ou lipídios, a ligação covalente de pequenas proteínas, como a ubiquitina e a modificação de resíduos da cadeia lateral de aminoácidos específicos (Deribe et al., 2010).

As PTMs mais comumente estudadas são a fosforilação, a ubiquitinação, a glicosilação e a hidroxilação (Tabela 2). A fosforilação é um das PTMs mais comuns em proteínas e consiste na fixação reversível de um grupo fosfato em um resíduo específico de uma proteína-alvo, atuando na sinalização e na regulação da atividade proteica. A participação das proteínas quinase e fosfatase (Figura 10) é crucial e altamente dinâmica, regulando intensivamente a adição do grupo fosfato, dada pela quinase, ou remoção do mesmo, dada pela fosfatase, em ciclos alternados que podem ocorrer em intervalos de tempo muito curto. Vários tipos de fosforilações têm sido relatados, sendo os mais freqüentes a fosforilação do grupo hidroxila dos resíduos de treonina, serina e tirosina (Johnson e Eyers et al., 2010; Ribet et al., 2010).

Tabela 2. Algumas das mais comuns e importantes PTMs (adaptado de Mann et al., 2003).

PTMs	Δ Massa (Da)*	Estabilidade durante a fragmentação por MS/MS**
Fosforilação		
Tyr	+80	+++
Ser/Thr	+80	+ / ++
Acetilação	+42	+++
Metilação	+14	+++
Glicosilação		
N-	>800	+ / ++
O-	203, >800	+ / ++
Hidroxiprolina	+16	+++
Pontes dissulfeto	-2	++
Deamidação	+1	+++
Ubiquitinação	>1,000	+ / ++
Nitrotirosina	+45	+ / ++

* Valores de Δ Massa (Da) de PTMs podem ser encontrados em: <http://www.unimod.org/>

**Estabilidade: + instável; ++ moderadamente estável; +++ estável.

As proteínas são modificadas não só pela adição de pequenas modificações químicas, mas também pela ligação covalente de pequenas proteínas como a ubiquitina em uma proteína alvo (Ikeda et al., 2008). A ubiquitina é uma proteína de aproximadamente 9 kDa, presente em todos os eucariotos, correspondente à sinalização de proteínas para a degradação em períodos definidos no ciclo celular (Ribet et al., 2010).

A hidroxilação é uma PTM predominante em proteínas secretadas, que consiste na adição de grupos hidroxilas aos resíduos de prolina. Essa modificação é catalisada pela prolil hidroxilase (PHD), garantindo uma estabilidade proteica (Kaelin, 2005). A glicosilação é também uma das PTMs mais comuns em certas proteínas secretadas de eucariotos contribuindo para atividades biológicas, solubilidade, estabilidade, interação célula-célula e resistência às proteases (Steinberg et al., 2001). Trata-se de uma modificação mais complexa com dois principais mecanismos de ligações covalentes: N-glicosilação, no qual a adição de um grupo

carboidrato ocorre em um resíduo de asparagina; e O-glicosilação, no qual a adição de um carboidrato ocorre em um resíduo de serina ou treonina (Helenius et al., 2004; Morelle et al., 2007).

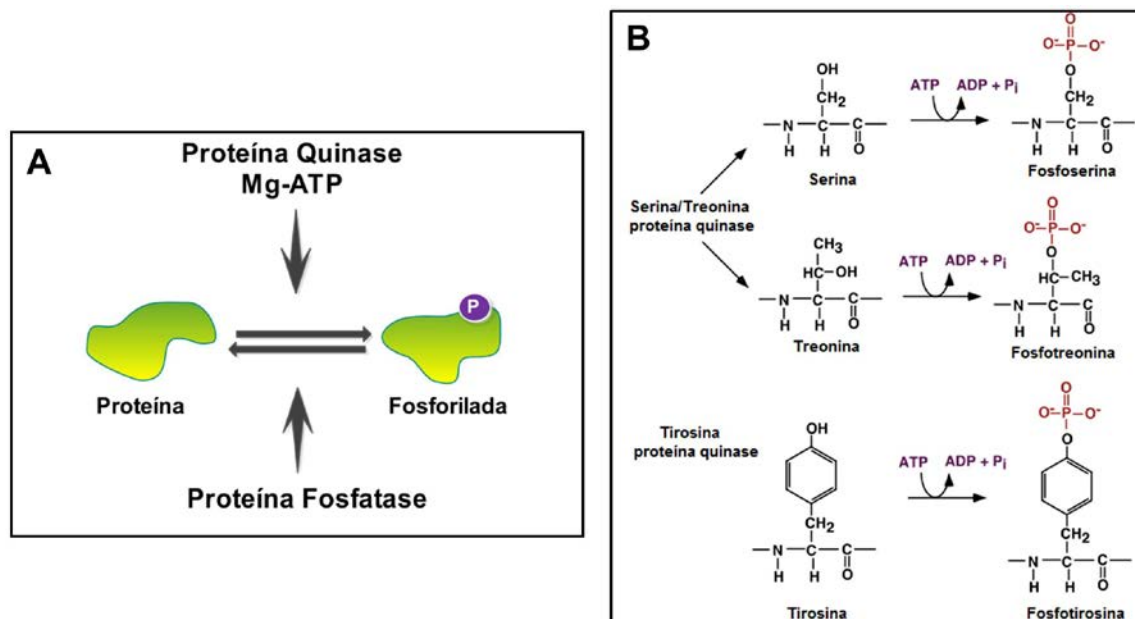


Figura 10. A - Esquema representativo da regulação da fosforilação de uma proteína. A enzima quinase é responsável pela adição do grupo fosfato à proteína, enquanto que a enzima fosfatase é responsável pela remoção do grupo. **B** - fosfoserina, fosfotreonina e fosfotirosina.

A ocorrência de diversas PTMs, atuando em seqüência e/ou em conjunto, parece ser um mecanismo conservado evolutivamente para iniciar, terminar ou refinar os resultados das vias de sinalização. Este conjunto mosaico de PTMs mostra como um complexo circuito de sinalização pode ser estabelecido pela simples adição de modificadores de um grupo constante de produtos gênicos (Deribe et al., 2010). No entanto, a identificação das PTMs é dificultada devido a sua diversidade, complexidade, baixa abundância e heterogenicidade (Shu et al., 2004; Cantin et al., 2004; Johnson e Eysers et al., 2010).

Como estão envolvidas na maioria dos processos celulares, incluindo a manutenção da estrutura proteica, as análises das PTMs têm se tornado um grande desafio às pesquisas proteômicas sendo requeridos métodos sensíveis e eficientes para a detecção das PTMs. Porém, a espectrometria de massas (MS) tem demonstrado ser extremamente útil na identificação das PTMs; e tem sido amplamente utilizada na identificação e caracterização de proteínas (Mann et al., 2003; Witze et al., 2007; Johnson e Eysers et al., 2010). A presença de modificações covalentes em proteínas altera o peso molecular dos aminoácidos modificados, e

esse aumento ou perda de massa pode ser detectada pelas análises de MS, as quais apresentam inúmeras vantagens como: alta sensibilidade, capacidade de identificação do local da PTM, possibilita a detecção de novas PTMs, capacidade de identificar PTMs em uma mistura complexa de proteínas, e condições de quantificar as mudanças relativas das PTMs em locais distintos. Nenhuma outra técnica apresenta todas essas vantagens características (Larsen et al., 2006).

A caracterização de PTMs por LC-MS/MS é atualmente viável utilizando uma série de abordagens complementares como: (i) espectrometria de massas sequencial (*tandem MS/MS*) utilizando um ou mais de um método de fragmentação disponíveis como CID (dissociação induzida por colisão), ECD (dissociação por captura de elétrons), e/ou ETD (dissociação por transferência de elétrons) tanto para a determinação da sequência de aminoácidos quanto para a determinação do tipo e localização da PTM; (ii) remoção da PTM entre análises consecutivas de espectrometria de massas; (iii) enriquecimento seletivo de proteínas ou peptídeos baseados no grupo funcional modificado antes das análises de MS/MS; e (iv) utilização de várias estratégias de MS para a determinação de PTMs específicas (Johnson e Eyers et al., 2010).

Portanto, não existe uma estratégia universal para identificar as PTMs e todas as abordagens utilizadas apresentam as suas limitações. Técnicas de separação de proteínas como eletroforese bidimensional (2-DE) e métodos cromatográficos realizam uma interface de forma on-line ou off-line com a MS. As proteínas são então convertidas em peptídeos através de tratamentos proteolíticos com enzimas específicas propiciando uma análise específica de MS e MS/MS (Lubec & Afjehi-Sadat, 2007). Métodos de ionização como MALDI-MS (*Matrix-assisted laser desorption/ionization MS*) e ESI-MS (*Electrospray ionization MS*) têm sido amplamente utilizados em análises proteômicas e na identificação de PTMs. Além disso, o sequenciamento de aminoácidos por MS/MS possibilita a identificação de PTMs em resíduos individuais de aminoácidos na proteína (Jensen, 2004). Sendo assim, análises de PTMs por uma abordagem proteômica como a que se foi proposta no presente estudo, poderá ser realizada utilizando tecnologias sofisticadas e sensíveis como a 2-DE, cromatografia líquida e espectrometria de massas (MS).

2.6. Análise proteômica

O termo “proteoma” foi introduzido e descrito como “todas as proteínas expressas em um genoma ou tecido” (Wilkins et al., 1997). No entanto, sabe-se que essa definição não é totalmente correta, pois o proteoma é o perfil proteico de todas as proteínas expressas em uma célula, tecido ou organismo em um determinado momento do sistema biológico, uma vez que o

sistema está expressando um complemento proteico que reflete todas suas interações com o meio num momento específico (Pennington et al., 1997).

A proteômica é a interação entre genes, proteínas e uma determinada condição fisiológica, farmacológica e/ou patológica. Qualquer proteína pode existir em múltiplas formas em uma célula ou em células diferentes. Modificações ocorridas podem ser oriundas de processos traducionais e/ou pós-traducionais, processos regulatórios e de degradações que afetam a estrutura da proteína, destino celular (localização) e função (Liebler, 2002). Trata-se, portanto, de uma técnica complementar a genômica, pois ela se concentra nos produtos gênicos, os quais são agentes ativos nas células. Por essa razão, a proteômica contribui diretamente para a identificação de alvos proteicos, contra os quais podem ser desenvolvidas drogas de uso terapêutico (Pandey et al., 2000). Utiliza uma combinação de métodos sofisticados como eletroforese bidimensional, técnicas de cromatografia, análises de imagens e espectrometria de massas, que possibilitam uma oportunidade maior de exibir e identificar proteínas (Xiang et al., 2004; Canelle et al., 2005).

A proteômica diferencia-se da química de proteínas tradicional por ser o estudo de sistemas de múltiplas proteínas. As análises são direcionadas às misturas complexas e suas identificações não são realizadas pela análise completa da sequência da proteína, mas por análises de sequências parciais, com o auxílio de ferramentas de bioinformática. O contexto do proteoma é o estudo das proteínas de um sistema biológico, associado à caracterização das respostas moleculares desse sistema num determinado momento (Plebani, 2005).

Dentre as principais aplicações da análise proteômica estão: i) Proteoma de expressão de proteínas, que estuda a expressão de proteínas quando o sistema biológico é exposto a drogas ou estímulos físicos, além de identificar novas proteínas como marcadores específicos de doenças; ii) Proteoma estrutural, no qual o principal objetivo é mapear a estrutura de complexos proteicos e interação entre proteínas; iii) Proteoma funcional, o qual estuda as interações entre proteínas e entre proteínas e/ou outras moléculas tendo como função determinar quais proteínas interagem em sistemas organizados e qual a função das mesmas; iv) Proteoma de mapeamento das modificações de proteínas, o qual envolve a identificação de como e onde as proteínas são modificadas, estabelecendo a natureza das modificações pós-traducionais; e v) Proteoma qualitativo, o qual identifica todas as proteínas do sistema biológico em estudo, a fim de compreender os mecanismos de ação dessas proteínas no organismo em estudo (Plebani, 2005). Essas aplicações podem ser utilizadas de forma isoladas ou em conjunto, dessa forma, o presente estudo se enquadra em diferentes aplicações da análise proteômica como o proteoma qualitativo e o proteoma de mapeamento das modificações pós-

tradicionais, com a identificação e análise estrutural das proteínas das glândulas que secretam a seda da aranha *N. clavipes*, e também das proteínas que constituem essa seda.

Nos últimos vinte anos, a engenharia proteômica tem passado por uma grande expansão e desenvolvido técnicas que possibilitam a caracterização molecular e estrutural de componentes bioativos, presentes em diferentes tipos de células, tecidos, extratos e/ou fluidos corporais. Portanto, a proteômica é uma ferramenta que começou a ser utilizada na identificação e sequenciamento molecular de proteínas em geral. Essa expansão é devida não só ao desenvolvimento de instrumentação e técnicas cada vez mais sofisticadas como a espectrometria de massas, como também ao desenvolvimento de novos métodos de preparação de amostras, sequenciamento genômico de organismos modelo de estudos, capacidades melhoradas de quantificação e disponibilidade de ferramentas de bioinformática cada vez mais confiáveis. Além disso, os avanços no estudo de PTMs e a capacidade de quantificação aumentam ainda mais a utilidade da espectrometria de massas e seu uso em análises proteômica e peptidômica (Cunningham et al., 2012).

2.7. Abordagens proteômicas

A melhor forma de se realizar um estudo de análise proteômica é determinar e/ou conhecer bem a pergunta biológica do estudo que se pretende realizar. Sendo assim, podemos definir o estudo de uma análise proteômica com uma palavra-chave: estratégia. Uma estratégia experimental bem definida é necessária para que uma pergunta biológica possa ser bem resolvida. Sem contar o fato de que ao planejar uma estratégia experimental será possível otimizar o tempo gasto para a execução dos experimentos e o elevado custo dos equipamentos/técnicas e reagentes utilizados.

Atualmente são descritas três estratégias de abordagens proteômicas: *bottom-up*, *top-down* e *middle-down* (Meyer et al., 2011). Na abordagem *bottom-up*, as proteínas são submetidas a uma digestão enzimática com uma protease, como por exemplo, tripsina, e os fragmentos peptídicos são posteriormente submetidos à análise de MS (Woods et al., 2012; Sokolowska et al., 2012). Uma desvantagem dessa abordagem é devido ao tempo para a realização de todo o processo da digestão enzimática para a obtenção dos fragmentos peptídicos antes das análises de MS. Além disso, a utilização de enzimas proteolíticas apresenta algumas limitações, como baixa termoestabilidade, possibilidade de autólise devido a valores de pH básicos e uma cobertura de sequência parcial da proteína (Baud et al., 1999; Thelen et al., 2012). Para a obtenção de uma melhor cobertura de sequência (próximo de 100%) é necessário utilizar diferentes enzimas proteolíticas além da tripsina, como por exemplo,

quimiotripsina, pepsina, Asp-N e outras proteases (Fischer et al., 2006; Swuaney et al., 2010). Dessa forma, a abordagem *bottom-up* possibilita também a identificação e localização de PTMs na sequência da proteína devido à obtenção de uma elevada cobertura da sequência. Entretanto, o aumento da complexidade introduzida por essa abordagem requer outro nível de separação, independente de um fracionamento a nível proteico previamente aplicado, antes que uma análise aprofundada possa ser realizada através da espectrometria de massas, os fragmentos peptídicos são tipicamente separados por cromatografia líquida de alto desempenho de fase reversa (RP-HPLC).

Já na abordagem *top-down*, a proteína intacta, ou seja, sem ser submetida a uma digestão enzimática, também pode ser submetida a uma análise de MS. Trata-se de um processo rápido já que não há a necessidade de digestão enzimática para as análises de MS, no entanto, requer tanto o uso de instrumentos de alta precisão de massa, bem como a utilização de algoritmos de deconvolução (Garcia et al., 2010; Tipton et al., 2011). Essa abordagem também possibilita a obtenção de uma elevada cobertura da sequência da proteína e estudos de PTMs. Apesar de ser uma estratégia mais recente para as aplicações proteômicas, a utilização da abordagem *top-down* tem aumentado cada vez mais devido aos grandes avanços na espectrometria de massas.

A combinação de ambas as abordagens tem sido considerada a melhor estratégia a ser utilizada em uma análise proteômica devido à complementação dos resultados obtidos de cada abordagem. Sendo assim, recentemente surgiu a abordagem *middle-down*, uma forma híbrida oriunda das abordagens *bottom-up* e *top-down*. Nesta abordagem, as proteínas intactas e de elevado peso molecular são submetidas a uma fragmentação limitada, dando origem aos grandes fragmentos peptídicos (3-20 kDa), os quais são sequenciados utilizando-se a abordagem *top-down*, mantendo dessa forma as vantagens de uma elevada porcentagem de cobertura da sequência e retenção das informações de PTMs. (Cannon et al., 2010; Meyer et al., 2011).

Todas essas abordagens mencionadas anteriormente podem ser realizadas em associação com tecnologias sofisticadas e sensíveis de fracionamento proteico, como a 2-DE, e cromatografia líquida de alto desempenho de fase reversa (RP-HPLC). A 2-DE é o método mais eficiente de separação simultânea de centenas ou milhares de proteínas, as quais são separadas com base em duas das suas propriedades: numa primeira dimensão de acordo com o seu ponto isoelétrico (pI) e, numa segunda dimensão, em gel de SDS-PAGE, de acordo com o seu peso molecular (MW). Também, em muitos casos, permite a separação de diferentes isoformas das proteínas. Porém, a faixa dinâmica é limitada e a quantificação das proteínas é

normalmente baseada na intensidade dos *spots*, por meio da análise das imagens geradas (Minden et al., 2009; Lewandowska-Gnatowska et al., 2011). A quantificação pode também ser comprometida se um *spot* do gel contém mais do que uma proteína, apesar da alta resolução isso pode muitas vezes ocorrer (Roepstorff, 2012). Outra razão é a dificuldade em se estudar proteínas de membrana ou proteínas que apresentam alto/baixo peso molecular ou alto/baixo pI (Brewis et al., 2010). A estratégia frequentemente aplicada, conhecida como Gel-LC-MS/MS, consiste na separação de proteínas em 2-DE seguida de digestão das proteínas utilizando uma enzima proteolítica, extração dos fragmentos peptídicos e análise dos mesmos através de LC-MS/MS (cromatografia líquida acoplada à espectrometria de massas) (Wang et al., 2010).

Outra estratégia utilizada é denominada como análise proteômica de *shotgun*. Trata-se de um método utilizado para a análise de uma amostra complexa de proteínas sem uma separação prévia. Para amostras contendo nenhum detergente e de baixa a média complexidade, a digestão em solução e injeção da amostra em cromatógrafo acoplado diretamente ao espectrômetro de massas pode ser a melhor escolha (Haas et al., 2006). Os diversos fragmentos peptídicos gerados são separados e já analisados, podendo então ser contrastados com os bancos de dados. Neste caso, isoformas de proteínas não são muitas vezes diferenciadas e informações referentes às PTMs podem ser perdidas, assim como também informações referentes às proteínas pouco abundantes na amostra. Isto é devido ao fato de que os espectrômetros de massas conseguem visualizar apenas a sequência de um número limitado de peptídeos num determinado “espaço de tempo”, e estes serão normalmente visualizados como os picos de maior intensidade, enquanto que os picos pouco abundantes na amostra não serão sequenciados. No entanto, é possível obter um maior número de identificações através da introdução de listas de exclusão, de modo que os picos que foram sequenciados numa primeira análise sejam excluídos em uma segunda análise (Roepstorff, 2012).

2.8. Espectrometria de massas

Na pesquisa proteômica é muito importante obter uma melhor compreensão sobre a estrutura e função das proteínas. Uma das principais abordagens proteômicas envolve as análises de proteínas e peptídeos baseadas na espectrometria de massas (MS) com a finalidade de solucionar sequências de aminoácidos, estruturas proteicas, assim como também a identificação e localização de PTMs. Sendo assim, a espectrometria de massas emerge como uma tecnologia indispensável para a interpretação da informação codificada pelos genes, ou seja, o proteoma. Uma das forças que impulsiona a proteômica é a habilidade de usar dados de

espectrometria de massas inerentes a peptídeos para identificar proteínas em bancos de dados.

A espectrometria de massas não é uma técnica recente, ela teve seu início em 1886 com a descoberta do íon positivo por Goldstein. O primeiro espectro de massas foi obtido por Thomson, em 1912, e o primeiro espectrômetro foi desenvolvido por Dempster em 1918, a partir desta data vários tipos de espectrômetros foram desenvolvidos, mas o grande avanço da espectrometria de massas no campo biológico ocorreu a partir da década de 80 com o desenvolvimento das técnicas MALDI (*matrix-assisted laser desorption ionization*) e ESI (*electrospray ionization*) por Hillenkamp e Fenn respectivamente (Karas et al., 1988; Fenn et al., 1989; Johnstone, et al., 1996). O advento de MALDI e ESI, assim como o avanço nos equipamentos em termos de sensibilidade e resolução, têm permitido que a verdadeira complexidade das amostras biológicas possa ser estudada e desvendada. Portanto, ocorreu um avanço paralelo entre as ciências “*omicas*” e os equipamentos/técnicas.

As análises de MS são realizadas através da integração de três componentes instrumentais: (i) uma fonte de íons, que gera íons na fase gasosa; (ii) um analisador de massas, que separa os íons na fase gasosa através da sua relação m/z ; e (iii) um detector de íons, que registra os valores de m/z de cada íon (Vestal, 2011; Lothrop et al., 2013). A espectrometria de massas é incomparável como uma ferramenta bioanalítica - permitindo ser possível medir a massa de qualquer molécula com uma exatidão extraordinária, resolução, sensibilidade, velocidade e reprodutibilidade (Watson et al., 2007; Thelen et al., 2012); além de possibilitar a fragmentação dos íons através da espectrometria de massas in *tandem*, ou também chamada de MS/MS. Sendo assim, a espectrometria de massas sequencial (MS/MS) de proteínas e peptídeos envolve a fragmentação dessas moléculas e a posterior análise de suas sequências de aminoácidos, fornecendo dessa forma informações importantes sobre a sequência e as possíveis modificações das proteínas através dos fragmentos peptídicos gerados (Breuker et al., 2008; Han et al., 2008; Tipton et al., 2011).

As análises de MS/MS desempenham um papel particularmente importante na identificação e localização de PTMs. Os mecanismos de fragmentação peptídica, comumente utilizados para as análises de PTMs, incluem a dissociação induzida por colisão (CID) (Wells et al., 2005), a dissociação por captura de elétrons (ECD) (Meng et al., 2005), e a dissociação por transferência de elétrons (ETD) (Udeshi et al., 2007; Wiesner et al., 2008). Independente do mecanismo utilizado, o desafio em todos os casos é conseguir gerar uma fragmentação que seja suficiente para identificar com precisão a sequência de aminoácidos do íon precursor e localizar a PTM nessa sequência. Uma boa fragmentação possibilita que a PTM seja retida e

sua localização pode ser facilmente mapeada, comparando-se os valores de massas de fragmentos teóricos e observados. No entanto, PTMs instáveis podem ser perdidas dificultando a localização do seu sítio específico na sequência (Palumbo et al., 2011).

O CID ainda continua a ser o método de fragmentação escolhido para a maioria das aplicações proteômicas. No entanto, o CID resulta na fragmentação preferencial das ligações com menor energia, e no caso dos peptídeos modificados, essa é muitas vezes a ligação covalente que liga a modificação com a estrutura do peptídeo. Além disso, a fragmentação pode ser afetada devido ao comprimento do peptídeo e do estado de carga do mesmo. Sendo assim, outros métodos de fragmentação MS/MS como ECD e ETD são muitas vezes necessários para melhorar a detecção de PTMs (Wiesner et al., 2008).

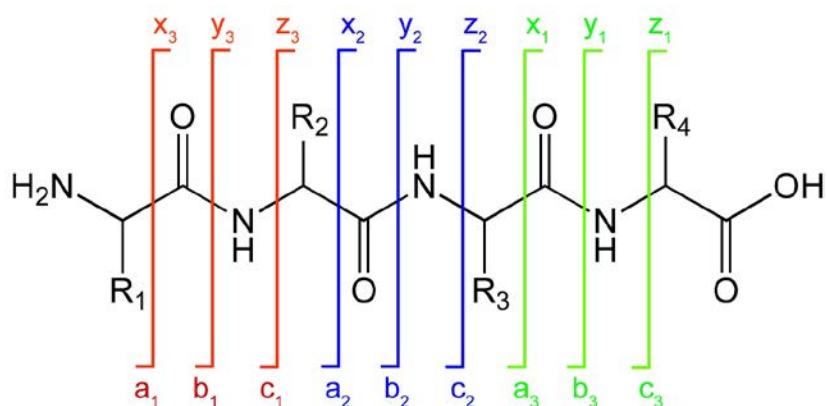


Figura 11. Esquema representativo da estrutura química geral de um peptídeo, apresentando as possibilidades de fragmentações que surgem em espectros de massas obtidos por análise *in tandem*. Os íons formados são enumerados a partir do lado N-terminal, sendo assim, os íons são denominados por -a, -b ou -c se a carga for mantida no lado N-terminal e, denominados por -x, -y ou -z se a carga for mantida no lado C-terminal. O índice indica o número do resíduo de aminoácido do fragmento em questão.

O método de fragmentação por ETD predominantemente gera íons do tipo -c e -z, enquanto que o CID gera íons do tipo -b e -y (Figura 11). Devido a natureza desses fragmentos, os dados oriundos da fragmentação por ETD complementam os dados oriundos do CID. No entanto, os dois métodos são tendenciosos no que diz respeito aos valores de *m/z* e aminoácidos. Por exemplo, o CID funciona melhor para peptídeos menores carregados duplamente, enquanto que o ETD funciona melhor para peptídeos com elevado estado de carga; e ainda o CID funciona melhor quando a ligação peptídica se encontra na extremidade amino-terminal de um resíduo de prolina, sendo esse local de fácil dissociação sob CID ao

contrário do ETD. A vantagem de se alternar o uso dos dois métodos torna-se interessante, uma vez que ocorre uma complementariedade dos dados gerados por ambos. Isso se torna claro, quando, por exemplo, o sítio de fosforilação está localizado próximo a um resíduo de prolina. Sendo o ETD incapaz de quebrar essa ligação, o CID pode ser utilizado para produzir íons de elevada intensidade no respectivo sítio de clivagem (Wiesner et al., 2008; Xia et al., 2011; Kim & Pandey, 2012).

Todos os aspectos das análises por MS têm sido beneficiados devido ao acoplamento direto entre o espectrômetro de massas e os instrumentos de cromatografia líquida (LC). De fato, a sensibilidade de detecção, a fragmentação MS/MS e a interpretação dos dados melhoram muitas vezes quando as proteínas e peptídeos são devidamente separados antes das análises por MS (Ewles et al., 2011). Atualmente, até mesmo os experimentos mais básicos por LC-MS/MS podem resultar na detecção e quantificação de dezenas de milhares de íons distintos em um intervalo curto de um gradiente padrão de LC (Di Palma et al., 2012). O uso de um sistema LC-MS/MS tem se tornado o principal método para detecção e quantificação de PTMs como tem sido demonstrado em diversos estudos (Chen et al., 2010; Redeker, 2010; Britton et al., 2011; Resende et al., 2013). No entanto, a cromatografia líquida em nanoescala acoplada à espectrometria de massas *tandem* (NanoLC-MS/MS), também tem se tornado uma ferramenta essencial em aplicações proteômicas. A sua sensibilidade apresenta vantagens sobre a LC-MS/MS convencional ao permitir a análise de misturas complexas de peptídeos em situações limitadas de amostras, como por exemplo, as proteínas proteoliticamente digeridas oriundas da eletroforese bidimensional. Esse sistema apresenta as seguintes vantagens: (i) uma eficiente capacidade de concentrar os peptídeos antes da detecção por análises de MS, (ii) a identificação das proteínas baseia-se no sequenciamento independente dos peptídeos com uma elevada cobertura de sequências e confiabilidade nas identificações, e (iii) toda a amostra é utilizada durante a análise evitando perdas. Dessa forma, o uso do sistema NanoLC-MS/MS tem se tornado eficiente devido a sua sensibilidade e confiabilidade (Chen et al., 2010; Gaspari, 2011).

2.9. Bioinformática estrutural

Podemos considerar a bioinformática como uma linha de pesquisa que envolve aspectos multidisciplinares (ciência da computação, matemática/estatística, biologia, bioquímica, física, medicina, entre outros) e que surgiu a partir do momento em que se iniciou a utilização de metodologias computacionais para a análise de dados biológicos, químicos, médicos, entre outros (Bayat, 2002). Dentro desse contexto, e da impossibilidade da análise manual desses

dados, se torna necessário o emprego de computadores e abordagens computacionais como ferramentas para análise, captura e interpretação de dados biológicos.

A bioinformática utiliza sistemas de informações para entender processos biológicos; portanto, possui um papel essencial, seja decifrando dados gerados por tecnologias de *high-throughput* relacionadas aos projetos genoma, proteoma e transcriptoma, seja organizando, coordenando e viabilizando o acesso racional aos dados gerados a partir da biologia tradicional. Mais do que metodologias e protocolos de análises de sequências de nucleotídeos e proteínas, a bioinformática tem fornecido a base conceitual e os métodos práticos, para a detecção de comportamentos sistemáticos funcionais de células e organismos e para a elucidação de processos somente detectáveis quando se estuda o organismo como um todo.

O aumento no número de estruturas tridimensionais (3D) disponíveis em bancos de dados como o PDB (*Protein Data Bank*) (Berman et al., 2000), levou ao crescimento da bioinformática estrutural. Assim, a predição de estruturas 3D de proteínas permanece uma área de grande interesse, sendo que a principal categoria de predições de estruturas de proteínas tem sido a modelagem molecular por homologia, baseada nos conhecimentos da estereoquímica dos aminoácidos e na alta homologia de uma sequência por uma estrutura conhecida (Sanchez et al., 1997; De Lima, 2006). E a metodologia de Dinâmica Molecular, atualmente aumenta a possibilidade no avanço da compreensão de processos biológicos em escala atômica e molecular, sendo assim, tornou-se uma ferramenta essencial para a investigação de moléculas (Namba et al., 2008). Através da técnica de simulação do movimento de partículas em um sistema, obtêm-se as flutuações das posições relativas dos átomos de uma determinada molécula em solução, em função do tempo. Utilizando esse conhecimento é possível obter várias informações sobre a molécula como, por exemplo, as possíveis mudanças na estrutura 3D da molécula que esteja associada a sua atividade biológica, estudar interações da molécula com possíveis ligantes ou a importância dessas interações para a estabilidade e/ou atividade biológica (Karplus & Petsko, 1990). Com este método podemos estudar tanto estruturas obtidas experimentalmente como aquelas obtidas por modelagem molecular.

Entre as ferramentas mais utilizadas na bioinformática estrutural encontram-se aquelas relacionadas à comparação de sequência de aminoácidos através do alinhamento de sequências. O alinhamento de pares é uma comparação aproximada entre duas sequências, a qual tem como objetivo evidenciar padrões comuns, identidade e similaridade, que podem ocorrer globalmente ou localmente entre duas sequências comparadas e o alinhamento múltiplo é uma comparação aproximada com mais de duas sequências que reflete a existência de regiões comuns compartilhadas entre as sequências (Gibas & Jambeck, 2001).

A comparação de pares de sequências é fundamental para a realização da modelagem molecular por homologia. Se o grau de identidade entre as sequências primárias das proteínas, que podem ser utilizadas como moldes, e da proteína alvo for igual ou superior a cerca de 30%, existe uma grande probabilidade de que estas proteínas apresentem estruturas 3D semelhantes e pode-se construir um modelo 3D para a proteína alvo utilizando metodologias de modelagem molecular por homologia (Sander & Schneider, 1991). Existem vários programas que são utilizados no processo de alinhamento como o BLAST (Altschul et al., 1997), o FASTA (Lipman & Pearson, 1985), o CLUSTAL (Thompson et al., 1994), entre outros.

Nos casos em que a identidade entre as sequências é inferior a cerca de 30%, a melhor maneira de solucionar o problema é utilizar métodos de compatibilidade sequência-estrutura (*threading* ou *3D template matching methods*) (Jones et al., 1992). Estes métodos baseiam-se na avaliação da compatibilidade entre a predição da possível estrutura 3D da proteína alvo e de potenciais proteínas molde determinados experimentalmente. Ou seja, é escolhida a proteína molde, cuja estrutura 3D, apresenta uma maior compatibilidade em relação à sequência da proteína alvo. A avaliação sequência/estrutura é realizada por meio de um potencial empírico derivado de uma tabela de contatos de resíduos observados em proteínas de estrutura 3D determinadas experimentalmente. Esse método de seleção de proteínas molde é menos preciso que o método de alinhamento de pares, porém, trata-se de uma ferramenta muito útil quando não são encontradas proteínas moldes com identidade superior a 30% (Rost & Sander, 1996).

3. Objetivos

No presente estudo adotamos a abordagem *bottom-up* com a utilização e combinação de tecnologias como a 2-DE, cromatografia líquida e espectrometria de massas (NanoLC-ESI-CID/ETD-MSⁿ) para não só identificar o perfil proteico das glândulas que secretam a seda da aranha *N. clavipes*, como também as proteínas constituintes dessa seda com a obtenção de uma elevada cobertura da sequência e detecção de suas PTMs. Sendo assim, os objetivos do presente estudo foram:

- Identificar as proteínas presentes no conjunto de glândulas distintas especializadas na secreção de diferentes tipos de seda que constituem a teia, o frame, a linha de segurança “escape” e o casulo para a proteção dos ovos;
- Isolar e identificar as proteínas constituintes das fibras da seda que formam os fios da teia orbital e do frame utilizando uma abordagem proteômica;
- Realizar a digestão *in gel* dos *spots* proteicos individualmente com diferentes enzimas proteolíticas para obter uma maior cobertura da sequência das proteínas;
- Analisar os fragmentos peptídicos por espectrometria de massas (MS e MSⁿ) para a identificação das proteínas;
- Analisar e caracterizar as modificações pós-traducionais (PTMs) das proteínas identificadas;
- Verificar a presença de modificações pós-traducionais como fosforilação, glicosilação e nitrotirosina através de tratamentos específicos com fosfatase alcalina, PNGase e ensaios de imunodeteção;
- Realizar a modelagem molecular por homologia e caracterizar as modificações pós-traducionais na estrutura tridimensional da proteína;

- Propor um mecanismo geral de atuação das proteínas identificadas nas glândulas produtoras de seda, as quais podem estar envolvidas com a produção, secreção, armazenamento, transporte, proteção e mudanças conformacionais das espidroínas constituintes da seda durante todo o processo de fiação;
- Estabelecer um esquema representativo para o mecanismo natural de fiação das fibras da seda da teia da aranha *N. clavipes*.

4. Justificativa

A seda da teia de aranha tem sido extensivamente estudada através de dados oriundos da engenharia genética e da tecnologia do DNA recombinante (Lazaris et al., 2002; Lewis, 2006; Rising et al., 2007; Lefèvre et al., 2011; Rising et al., 2011). As surpreendentes propriedades mecânicas da seda têm motivado muitas pesquisas para o estabelecimento de técnicas de produção biotecnológica, as quais são necessárias para a obtenção de quantidades suficientes de proteínas da seda para as aplicações industriais (Vendrely et al., 2007). Porém, a análise proteômica prospectiva apresenta vantagens com relação às análises genéticas, já que normalmente, genes mutados podem não ser transcritos por inúmeras razões epigenéticas. Estudos proteômicos com análises de PTMs são necessários e importantes para que possamos compreender a relação estrutura-função e a interação entre as proteínas constituintes da seda após serem secretadas pelas glândulas. Tem-se tornado muito grande o interesse pelos fios da seda da teia de aranhas, e estudos tem demonstrado a eficácia do uso dessa seda principalmente com um grande potencial em aplicações biomédicas.

No presente estudo foi possível obtermos informações químicas e estruturais das características das proteínas da seda da teia da aranha *Nephila clavipes*, possibilitando o seu desenvolvimento em aplicações biotecnológicas. Dessa forma, esse estudo procurou alcançar alguns dos objetivos principais do projeto BIOprospecTA da FAPESP que surgiu como um projeto piloto para a bioprospecção da biodiversidade do Estado de São Paulo. O grupo de pesquisas em Biologia Estrutural e Zooquímica do CEIS/IBRC-UNESP possui experiência na área de caracterização de compostos orgânicos, proteínas e peptídeos em secreções de artrópodes e tem realizado a prospecção química de maneira sistemática, em insetos e aranhas da fauna do Estado de São Paulo, junto ao programa BIOprospecTA/FAPESP (Proc. 2011/51684-1).

5. Material e métodos

5.1. Obtenção das glândulas e da seda da teia da aranha *Nephila clavipes*

As aranhas *Nephila clavipes* foram coletadas no campus da Universidade Estadual Paulista “Júlio de Mesquita Filho” localizado na cidade de Rio Claro, São Paulo, de maneira georeferenciada (S22°23'42.9'', W047°32'32.5''), para a dissecação de suas glândulas produtoras de seda. Também foram coletados alguns dos diferentes tipos de fios de seda, tais como a “teia orbital”, e o “frame” (que suporta toda a estrutura da teia).

5.2. Extração das proteínas das glândulas e da seda

As glândulas foram dissecadas e armazenadas individualmente em frascos distintos. Logo em seguida, foram maceradas individualmente em presença de tampão acetato de amônio 10 mM, pH 6.8 contendo coquetel de inibidores de proteases (Sigma). As amostras foram então centrifugadas à 8000 rpm por 10 min, com o sobrenadante sendo coletado em um tubo. As amostras foram liofilizadas e armazenadas à -80 °C até o momento de serem utilizadas. Os fios de sedas da teia orbital e do frame, coletados, foram limpos com o auxílio de uma pinça para a retirada de partículas, restos de insetos, folhas e galhos aderidos à seda, evitando dessa forma uma contaminação da amostra. Após a limpeza, foram pesados 30 mg dos fios de seda da teia orbital e do frame individualmente em frascos distintos, nos quais foram adicionados 2 mL de solução saturada de tiocianato de lítio (25g de LiSCN hidratado em 10 mL de água, 38 M). As amostras foram incubadas à 25°C sob agitação por 2h. Após esse período, as amostras foram centrifugadas à 14000 rpm por 15 min, com o sobrenadante sendo transferido para um tubo adequado de uma unidade filtrante da centrífuga Ultrafree-4 (Millipore) com adição de 2 mL de solução de uréia recém preparado (Tris 20mM, uréia 7M, tiouréia 2M, CHAPS 4% m/v, 1,4-Ditioeritritol 10 mM, EDTA 1 mM, PMSF 1mM, coquetel de inibidores de proteases 0.2% v/v, coquetel de inibidores de fosfatases 0.2% v/v). As amostras juntamente com a solução de uréia foram centrifugadas a 4500 rpm até o volume ser reduzido a 0.5 mL. Esse procedimento foi realizado por cinco vezes. Dessa forma a solução saturada de LiSCN foi removida da amostra sendo trocada pela solução de uréia, estando em condições favoráveis para a realização da 2-DE. As amostras das glândulas também foram solubilizadas com a solução de uréia para a realização da 2-DE.

5.3. Quantificação das amostras

A quantificação de proteínas das amostras foi determinada pelo método de Bradford (BioRad), utilizando-se albumina bovina (BSA) como padrão (Bradford, 1976).

5.4. 2-DE

As amostras (250 μ g) foram aplicadas em fitas *Immobilized pH 3-10* e *pH 7-10*, 18 cm gradiente não linear. A focalização isoeletrica foi iniciada em 200 V com um aumento gradual até 8000 V (4V/min) se mantendo constante por um período de 3h (150000 Vh). Para a segunda dimensão, as fitas foram equilibradas por 15 min sob agitação em solução de equilíbrio SDS (Tris-HCl 50mM, pH 8.8, uréia 6M, glicerol 30% v/v, SDS 2% m/v, traço de azul de bromofenol), contendo DTT (1% m/v) e depois por mais 15 min na mesma solução de equilíbrio SDS contendo iodoacetamida (4% m/v). A segunda dimensão foi realizada em um gel SDS-PAGE (7-10% e 10-16%), 50V constante “overnight”, 200V constante por 5h. Após a fixação das proteínas por 4h em solução de fixação (metanol 50% v/v e ácido acético 10% v/v), o gel foi corado com *Colloidal Coomassie Blue* por 8h sendo o excesso de corante retirado com água destilada. Em seguida foi escaneado para a aquisição das imagens. As imagens foram analisadas utilizando o software Master Platinum v.7 (GE Healthcare).

5.5. Digestão in gel

Os *spots* proteicos foram excisados do gel e individualmente foram lavados com solução de bicarbonato de amônio 10 mM e ACN 50% v/v até serem completamente descorados, sendo em seguida secos em *Speedvac*. Logo depois, foram incubados com tripsina (Promega), 40 ng/ μ L em tampão de digestão (octil- β -D-glicopiranosídeo - OGP 5mM e bicarbonato de amônio 10mM, pH 7.8) à 37°C por 16h; quimiotripsina (Roche Diagnostics), 40 ng/ μ L em tampão de digestão (octil- β -D-glicopiranosídeo – OGP 5mM e bicarbonato de amônio 25mM, pH 7.8) incubado à 30°C por 3h; pepsina (Sigma), 40 ng/ μ L em tampão de digestão (HCl 0.1 M, pH 1.0) incubado a 37°C por 4h; Glu-C/V8 protease (Sigma), 40 ng/ μ L em bicarbonato de amônio 50mM, pH 7.8) incubado à 37°C por 16h; proteinaseK (Sigma), 40 ng/ μ L em tampão de digestão (bicarbonato de amônio 50mM, pH 7.8) incubado à 37°C por 2h; subtilisina (Sigma), 40 ng/ μ L em tampão de digestão (uréia 6 M e Tris-HCl 1 M, pH 8.8) incubado à 37°C por 2h; proteinase 10 (sin.: termolisina), 40 ng/ μ L em tampão de digestão (Tris-HCl 25mM, pH 7.5) incubado a 50°C por 2h; e Asp-N protease (Promega, 40 ng/ μ L em bicarbonato de amônio 25mM, pH 7.8) incubado à 37°C por 16h. A extração dos peptídeos de cada *spot* foi realizada com 15 μ L de ácido fórmico 0.1% em octil- β -D-glicopiranosídeo – OGP 5mM sob agitação por 30 min. e 15 μ L de ácido fórmico 0.1% em ACN 15% sob agitação por 30 min. Os peptídeos foram secos em *Speedvac* e armazenados a -20°C até o momento de serem analisados.

5.6. Análises de espectrometria de massas

Para as análises de espectrometria de massas (NanoLC-ESI-CID/ETD-MSⁿ) foi utilizado um espectrômetro do tipo HCT ETD II e um Amazon ETD (Bruker Daltonics, Bremen, Alemanha) acoplado a um HPLC Ultimate 3000 system (Dionex, Sunnyvale, CA, EUA), equipado com uma “*trap column*” PepMap100 C-18 (75 mm x 150 mm) e uma coluna analítica PepMap100 C-18 (300 mm x 5 mm). Foi utilizado um gradiente: A - 0.1% de ácido fórmico em água, B - 0.08% de ácido fórmico em ACN (4-30% de B (0-105 min), 80% de B (105-110 min), 4% de B (110-125 min). Repetidamente, os espectros de MS foram registrados seguidos por três espectros CID MSⁿ e três espectros ETD MSⁿ gerados a partir dos cinco íons precursores de maior intensidade. A voltagem entre “ion spray tip” e “spray shield” foi fixada em 1500 V. O gás nitrogênio foi aquecido á 150°C com um fluxo de 10L/min. A energia de colisão foi definida automaticamente de acordo com a massa e estado de carga dos peptídeos selecionados para a fragmentação. Os espectros MSⁿ foram interpretados e uma lista de picos foi gerada pelo software DataAnalysis 4.0 (Bruker Daltonics) como descrito anteriormente por Bae et al. (2012).

5.7. Identificação das proteínas

Para a identificação das proteínas foi utilizado o servidor MASCOT 2.2.06 (Matrix Science, London, UK) realizando buscas contra bancos de dados como o NCBI e Swiss-Prot. Os parâmetros de busca foram utilizados de acordo com a taxonomia do organismo de estudo (*Metazoa / Other Metazoa*), assim como também, conforme dados oriundos durante a manipulação experimental das amostras como, o tipo de enzima proteolítica utilizada e possíveis modificações fixas e variáveis que possam ter ocorridas. Dessa forma foram utilizados os seguintes parâmetros:

- a) Enzima: Tripsina (ou enzima correspondente)
- b) Carbamidometilação (C) como modificação fixa
- c) Oxidação (M), Fosforilação (Y) e Fosforilação (ST) como modificação variável
- d) N° de clivagens perdidas pela enzima: 1 (2 ou 3 de acordo com a enzima utilizada)
- e) Massa molecular do tipo: Monoisotópica
- f) Massa molecular: sem restrições
- g) Erro de tolerância de peptídeos $\pm 0.2\text{Da}$ e erro de tolerância de MSⁿ $\pm 0.5\text{ Da}$
- h) Protonação: +2, +3, +4 para o estado da carga dos peptídeos
- i) Tipo de instrumento: ESI-Trap
- j) Protein Scores maiores que 200 foram significativos ($p < 0.05$) e admitidos para a identificação das proteínas.

5.8. Identificação das modificações pós-traducionais (PTMs)

Para as análises de PTMs foi utilizada a espectrometria de massas MS^n combinando dois princípios de dissociação de íons que se complementam: ETD (Dissociação por Transferência de Elétrons) e CID (Dissociação Induzida por Colisão). A ETD fragmenta os peptídeos utilizando a técnica química de reações do tipo íon a íon. Os fragmentos são gerados pela transferência de um elétron de um radical aniônico para o peptídeo protonado. Essa transferência libera energia que induz a fragmentação peptídica causando clivagem das ligações C α -N. Esse tipo de fragmentação gera uma complementariedade ao formar os íons dos tipos c- e z- em vez dos íons típicos dos tipos b- e y- observados por CID. Os espectros MS^n foram interpretados e uma lista de picos foi gerada utilizando o software DataAnalysis 4.0. Após a identificação das proteínas foi realizada uma pesquisa de erro de tolerância para detectar clivagens inespecíficas não atribuídas às modificações. Pesquisas sobre as PTMs foram realizadas utilizando o software MODIROTM - The PTM Explorer 1.1 (Protagen AG, Dortmund, Germany). MODIRO é um software específico e extremamente eficiente para a identificação rápida e automatizada de PTMs em um conjunto de dados MS^n . Esse software é projetado para trabalhar com uma lista de picos do qual ele extrai todas as informações dos dados estatísticos de MS^n , proporcionando ótimos resultados independentemente da origem dos dados. O software é compatível com os principais fabricantes de instrumentos como Bruker, Applied Biosystems e Shimadzu. Foram utilizados os seguintes parâmetros:

- a) Enzima: Tripsina (ou enzima correspondente)
- b) Carbamidometilação (C) como modificação 1
- c) Oxidação (M) como modificação 2
- d) N° de clivagens perdidas pela enzima: 1 (2 ou 3 de acordo com a enzima utilizada)
- e) Massa molecular do tipo: Monoisotópica
- f) Massa molecular: sem restrições
- g) Erro de tolerância de peptídeos ± 0.2 Da e erro de tolerância de $MS^n \pm 0.5$ Da
- h) Tipo de instrumento: ESI-Trap
- i) Formato dos dados de MS/MS: MGF.

Uma lista com 172 modificações incluindo fosforilação (Y, S, T) e glicosilação (O-glicosilação e N-glicosilação) foram selecionadas e adicionados para a busca de PTMs nos fragmentos peptídicos dos espectros de MS^n obtidos experimentalmente. Após a busca foi gerada uma lista de fragmentos peptídicos para visualização nos espectros. Cada fragmento

peptídico identificado foi considerado significativo com base na qualidade da fragmentação dos íons, sendo considerado score significativo (> 80) e íon score (> 200). Para a busca de alterações de massas desconhecidas, substituição de resíduos de aminoácidos e cálculos de significância foi selecionado o modo de estratégia de busca avançada de PTMs dentro do software que emite uma lista de picos do qual é extraído as informações estatísticas dos dados de MSⁿ. As informações obtidas sobre as identificações e modificações das proteínas foram analisadas manualmente de forma minuciosa para a confirmação dos dados obtidos.

5.9. Tratamento com fosfatase alcalina

Para verificar os sítios de fosforilação observados através dos dados de MSⁿ, os *spots* foram excisados do gel, descorados, e em seguida secos em *Speedvac*. Subsequentemente foram incubados por 1h a 37°C em uma solução de bicarbonato de amônio 100 mM, contendo 0.5 mL de fosfatase alcalina (*calf intestine alkaline phosphatase*, New England Biolabs). Logo após, os *spots* foram lavados em uma solução de bicarbonato de amônio 100 mM, desidratados em ACN e secos em *SpeedVac*, prontos para serem submetidos à digestão enzimática.

5.10. Tratamento com PNGase

Para verificar os sítios de glicosilação observados através dos dados de MSⁿ, os *spots* foram excisados do gel, descorados, e em seguida secos em *Speedvac*. Subsequentemente foram incubados por 16h a 37°C em uma solução de bicarbonato de amônio 25 mM, contendo PNGase F 10% v/v (*Peptide-N-glycosidase F*, Sigma). Logo após, os *spots* foram lavados em uma solução de bicarbonato de amônio 25 mM, desidratados em ACN e secos em *SpeedVac*, prontos para serem submetidos à digestão enzimática.

5.11. Imunodeteção

Para verificar a ocorrência das modificações nitrotirosina, fosfotirosina e fosfoserina nas sequências das proteínas da seda, amostras contendo 20-50 µg de proteínas foram submetidas ao ensaio de imunodeteção. Após a realização da SDS-PAGE, as proteínas foram transferidas para uma membrana de PVDF (Millipore) usando um sistema de eletrotransferência semi-seco (Bio-Rad). A membrana foi bloqueada por 1h em solução T-TBS (0.1% (v/v) Twen 20, 5% caseína para a imunodeteção de nitrotirosina). Após seis lavagens com T-TBS (0.1% (v/v) Twen 20) por 10 min, a membrana foi incubada com anticorpo primário anti-nitrotirosina (1:1000, Millipore) à 4°C por 16h. Em seguida, foi lavada seis vezes com solução T-TBS (0.1% (v/v) Twen 20) por 10 min. e incubada com anticorpo secundário HRP-coupled anti-rabbit IgG

(1:3000, Abcam). Como controle positivo da imunodeteção foi utilizado Nitro-BSA (*Nitrated bovine serum albumin*, Sigma) que apresenta MW 68 kDa. O mesmo procedimento descrito acima foi realizado para a imunodeteção de fosfotirosina e fosfoserina. Nesse caso, a membrana foi bloqueada com 5% BSA – *bovine serine albumin* e incubada com anticorpo primário anti-fosfotirosina ou anti-fosfoserina (1:2000, Abcam); em seguida foi incubada com anticorpo secundário HRP-coupled anti-mouse IgG (1:10000, Abcam). E ainda, como controle do experimento, a amostra foi tratada com fosfatase alcalina (*calf intestine alkaline phosphatase*, New England Biolabs), conforme descrito anteriormente, antes de ser submetida à SDS-PAGE. As membranas foram submetidas à revelação fotográfica para detecção com o Kit de imunodeteção *Enhanced Chemiluminescence* (ECL plus, GE Healthcare).

5.12. Microscopia eletrônica de varredura (SEM)

Para demonstrar a solubilidade das fibras da seda, uma pequena porção (5x5mm) foi adicionada em uma solução saturada de tiocianato de lítio (LiSCN hidratado). Em seguida, as fibras da seda foram retiradas da solução e fixadas sobre uma placa de aquecimento do microscópio eletrônico, sendo submetida a um aquecimento por várias horas a 70°C. A análise foi realizada em um Vega II - microscópio eletrônico de varredura (Tascan, Brno, República Checa), em 7 kV e uma distância cerca de 8mm. Essa análise foi realizada em colaboração com a Dra. Helga Priewalder do Departamento de Paleontologia, Pesquisa Geológica de Viena, Áustria.

5.13. Microscopia eletrônica de transmissão (MET)

Para a detecção específica de lipídeos, fibras da seda livremente suspensas foram expostas ao vapor de ácido ósmico 4% (m/v) em água durante 60 min. Em seguida foram fixadas em solução de formaldeído 37% (v/v) em etanol por 60 min. As fibras da seda foram então, incubadas em tampão imidazol 0.1 M, pH 7.5 por 30 min. à temperatura ambiente no escuro. Posteriormente, as fibras foram lavadas duas vezes em tampão fosfato salino (PBS) e desidratada em soluções crescentes de etanol (álcool 90%, 95% e 100%, três vezes por 5 min. cada). As fibras foram incluídas em resina pura (Epon Araldite) homogeneizada com quatro gotas do catalisador DMP (2,4,6- dimetil amino fenol) e polimerizada em estufa à 60°C por 24h. Esse material foi então coletado em telas de cobre e contrastado com acetato de uranila e citrato de chumbo, sendo examinado e fotografado com Microscópio Eletrônico de Transmissão (Philips – Modelo CM100).

5.14. Análises de ditirosina em hidrolisados da seda da aranha *N. clavipes*

Amostra da seda (1.0-1.4 mg) foi hidrolisada em ácido clorídrico 6N à 90°C por 18 h. Os compostos voláteis foram removidos utilizando um evaporador rotativo à 30 mbar e 60°C e os resíduos foram reconstituídos em 400 µL de acetato de amônio 0.1M, pH 4.8. Para a análise, um sistema de cromatografia compreendendo 2 bombas (Merck/Hitachi L-6300, Darmstadt, Alemanha), injetor automático (Merck/Hitachi AS-2000A), coluna termostática (Thermotech, Langenzersdorf, Áustria), controle de troca de válvula (Merck ELV 7000-2) e um detector (Merck/detector de fluorescência Hitachi F-1080) foi utilizado. Como padrão, o composto ditirosina foi sintetizado como descrito anteriormente por Dong-Ik et al. (2008). Para a análise de ditirosina foi utilizado o seguinte procedimento: Primeira dimensão - coluna Purospher Star RP18, 5 µm, 250 x 4 mm; eluente A - 0.1% de HCOOH + NH₃, pH 3.0, eluente B - metanol, gradiente linear de 0 a 20 min, compreendido entre 95% A + 5% de B até 75% de A + 25% de B; fluxo de 0.8 mL/min, volume de injeção de 40 µL; intervalo de troca 7.7-9.1 min, tempo de retenção da ditirosina 8.4 min. Segunda dimensão - "coluna Nucleosil strong acid cation exchanger SA", 125 x 4 mm, 10 µm tamanho de partícula; eluente A - 0.1% de HCOOH, eluente B - 0.1 M de HCOOH + NH₃, pH 3.7; A + B = 60+40 (v/v), fluxo 1.0 mL/min, 285/410 nm. Essa análise foi realizada em colaboração com o Professor Dr. Günther Lamprecht do Instituto de Química Analítica da Universidade de Viena, Áustria.

5.15. Alinhamento múltiplo entre as sequências primárias das espidroínas

As sequências primárias foram alinhadas utilizando o programa CLUSTAL W (Thompson et al., 1994) e editadas no programa JALVIEW (Clamp et al., 2004). Nesse tipo de alinhamento as regiões de identidade mais importantes são assinaladas em azul escuro, enfatizando os resíduos altamente conservados.

5.16. Predição da estrutura secundária

A predição da estrutura secundária foi realizada através do programa JNETPRED *Secondary Structure Prediction* (Cuff et al., 2000) e editada no programa JALVIEW (Clamp et al., 2004).

5.17. Modelagem molecular por homologia

Todas as ferramentas de bioinformática estrutural descritas abaixo estão integradas para executarem as simulações de modelagem molecular por homologia e dinâmica molecular em larga escala em um *cluster Beowulf* de 15 nós sob o sistema operacional Linux Fedora 19

(BioComp - <http://izalco.incor.usp.br/ganglia>). Exceto para as análises e visualização dos resultados que foram executadas usando somente o servidor do *cluster Beowulf*. Esse estudo das estruturas 3D das proteínas da seda foi realizado em colaboração com a Dra. Helen Andrade Arcuri do CEIS/IBRC-UNESP e Faculdade de Medicina de São Paulo - Incor/HC.

A modelagem molecular por homologia é o principal método computacional utilizado para prever estruturas terciárias de grandes ou pequenas moléculas. A realização da modelagem molecular por homologia das espidroínas consistiu de quatro etapas:

1) Busca e seleção do molde: foi realizada uma busca no banco de dados de estruturas 3D de proteínas, PDB (<http://www.rcsb.org/pdb/>) por estruturas 3D determinadas experimentalmente que fossem compatíveis com as espidroínas. Durante essa busca, foram identificados somente moldes que cobrem parcialmente as espidroínas-1A, -1B e -2 (proteínas alvo), não sendo encontrados moldes compatíveis com a proteína da seda flageliforme (Tabela 3).

Tabela 3. Moldes utilizados para a construção dos modelos 3D das espidroínas, utilizando a metodologia do alinhamento sequencial de pares.

Proteína alvo	Região de cobertura	Identidade / Similaridade (%)	Código PDB_cadeia	Referências
Espidroína-1A	24-155	51 / 70	3LR2_A	Askarieh et al., 2010
Espidroína-1A	934-1008	80 / 92	2KHM_A	Hang et al., 2010
Espidroína-1B	24-155	52 / 72	3LR2_A	Askarieh et al., 2010
Espidroína-1B	934-1008	80 / 92	2KHM_A	Hang et al., 2010
Espidroína-2	28-94	54 / 75	3LR2_A	Askarieh et al., 2010
Espidroína-2	707-777	73 / 90	2KHM_A	Hang et al., 2010

Sendo assim, baseado nesses resultados do alinhamento sequencial de pares, foi utilizada então, a metodologia de compatibilidade sequência-estrutura do programa THREADER com o objetivo de identificar possíveis moldes para aquelas regiões das espidroínas que não foram identificadas com o alinhamento sequencial de pares (Tabela 4).

Tabela 4. Moldes utilizados para a construção dos modelos 3D das espidroínas, utilizando a metodologia de compatibilidade sequência-estrutura.

Proteína alvo	Região de cobertura	Grau de confiança	p-value	Código PDB_cadeia	Referências
Espidroína-1A	30-960	certo	<0.0001	3SYJ_A	Meng et al., 2011
Espidroína-1B	30-960	Certo	<0.0001	3SYJ_A	Meng et al., 2011
Espidroína-2	2-780	Médio	<0.01	3SYJ_A	Meng et al., 2011

A porcentagem de identidade da sequência primária entre as regiões das espidroínas e dos moldes identificados indicam que essas estruturas 3D são bons moldes para serem utilizadas na modelagem das regiões das espidroínas. Para a escolha dos moldes foi utilizado o parâmetro “*Confidence level*” que informa a estatística da qualidade do alinhamento que avalia o nível de confiança em conjunto com o valor do “*p-value*” que indica a probabilidade de falso positivo.

2) Alinhamento: As sequências primárias dos moldes foram alinhadas com as sequências primárias das espidroínas, para verificar a porcentagem de identidade entre eles. Para isto foi utilizado o programa de alinhamento BLAST para realizar a identificação dos moldes para as espidroínas; e também o programa THREADER 3.5 (Jones & Hadley, 2000) para encontrar moldes que tivessem compatibilidade sequência-estrutura com as espidroínas. Os moldes escolhidos foram aqueles que apresentaram uma identidade elevada não sendo inferior a 30% com as espidroínas. Os valores dos parâmetros de cortes mostrados nas tabelas 3 e 4 demonstram que os moldes escolhidos podem ser perfeitamente utilizados para a realização da modelagem das espidroínas.

3) Modelagem: o programa MODELLER 9v11 (Šali & Bludell, 1993) foi utilizado recebendo os arquivos contendo as sequências alinhadas e as coordenadas tridimensionais dos moldes. Dessa forma, o programa construiu as estruturas 3D das espidroínas baseado nas coordenadas tridimensionais dos moldes. O MODELLER utiliza as coordenadas atômicas dos moldes para gerar um conjunto de restrições espaciais que são aplicadas à sequência alvo. As restrições limitam, por exemplo, a distância entre dois resíduos no modelo que está sendo construído, baseadas na distância entre dois resíduos homólogos do molde. As restrições também podem ser aplicadas aos ângulos de ligação, ângulos diédricos e pares de diedros. Ao aplicar um bom número dessas restrições espaciais, o MODELLER limita efetivamente o número de configurações que o modelo pode assumir (Šali et al., 2001). As coordenadas atômicas 3D das águas e ligantes foram removidas de suas respectivas estruturas durante o processo de modelagem. As imagens das estruturas 3D das espidroínas foram geradas a partir do programa PyMol (Delano, 2002).

4) Análise: foram gerados 1000 modelos para cada espidroína e a escolha do melhor modelo foi baseada na qualidade estereoquímica utilizando o programa PROCHECK (Laskowski et al., 1993) que pertence ao pacote CCP4 (*Collaborative Computational Project No 4*) e a função objetiva do MODELLER (Shen & Šali, 2006); os modelos das espidroínas foram avaliados como um todo, bem como em regiões individuais. Essa análise consistiu em estudar

as posições atômicas dos moldes e dos modelos, de modo que foi possível determinar a qualidade estereoquímica dos modelos 3D das espidroínas.

5.18. Dinâmica molecular

A metodologia de Dinâmica Molecular, atualmente aumenta a possibilidade no avanço da compreensão de processos biológicos em escala atômica e molecular, sendo possível estudar tanto as estruturas obtidas experimentalmente como aquelas obtidas por modelagem molecular.

A simulação por dinâmica molecular para cada modelo 3D construído para as proteínas da seda da aranha *N. clavipes* foi executada pelo programa GROMACS 4.5 (Van Der Spoel et al., 2005). Diferentes sistemas (Tabela 5) foram estabelecidos para as proteínas espidroína-1A e -1B, com e sem a fosforilação (modificação mais abundante observada na sequência das espidroínas), com a finalidade de obter o maior número de informações diferentes sobre essas proteínas. Todos os sistemas foram imersos em uma caixa cúbica preenchida com moléculas de água do tipo SPC - *Simple Point Charge*, os contra íons foram inseridos nas posições eletrostaticamente mais favoráveis para neutralizar o sistema, e a distância entre cada átomo e a parede da caixa foi de 1,0 nm.

Tabela 5. Simulação por dinâmica molecular para cada modelo 3D utilizando diferentes sistemas estabelecidos para as proteínas.

Proteínas	Número de Moléculas			Contra íons	Campo de força
	H ₂ O	ODA*	TDA**		
Espidroína-1A sem PTMs	183333	-	-	Cl (15)	Gromos43a1
Espidroína-1B sem PTMs	153512	-	-	Cl (13)	Gromos43a1
Espidroína-1A com PTMs	183338	-	-	Na (2)	Gromos45a3
Espidroína-1B com PTMs	153487	-	-	Na (11)	Gromos45a3
Espidroína-1A sem PTMs	176863	176	176	Na (337)	Gromos43a1
Espidroína-1B sem PTMs	147001	176	176	Na (339)	Gromos43a1
Espidroína-1A com PTMs	176854	176	176	Na (354)	Gromos45a3
Espidroína-1B com PTMs	146953	176	176	Na (363)	Gromos45a3

H₂O - água; *ODA - ácido graxo octadecanóico (com 18 C); **TDA - ácido graxo tridecanóico (com 13 C).

As estruturas 3D dos ácidos graxos tridecanóico (TDA) e octadecanóico (ODA) foram construídos utilizando o programa PyMol. A informação sobre a topologia dos ácidos graxos não estava disponível nos campos de forças utilizados neste estudo, sendo assim, para a construção desta topologia utilizamos o servidor PRODRG (<http://davapc1.bioch.dundee.ac.uk/programs/prodrgr/> - Schuettelkopf & Van Aalten, 2004) o qual gera parâmetros para as ligações, ângulos e cargas. O arquivo no formato pdb gerado pelo PyMol foi inserido como entrada no servidor PRODRG; no arquivo de saída gerado pelo

servidor PRODRG, as cargas das moléculas foram alteradas pelas cargas que foram calculadas por meio do cálculo quântico e foram adotadas as nomenclaturas dos átomos aceitas pelos campos de forças utilizados neste estudo. Antes de iniciar a simulação da dinâmica molecular final, uma simulação de dinâmica molecular de minimização de energia foi realizada com o objetivo de eliminar indevidos contatos entre os átomos da molécula utilizando o algoritmo *steepest descent*. E outra simulação de dinâmica molecular com restrição de posições foi realizada, nessa simulação a molécula é mantida estável enquanto as moléculas do solvente são relaxadas. As simulações de dinâmica molecular de restrição e final foram realizadas conforme amostra isotérmica isobárica em temperatura de 300 K e, no tempo total de simulação de 1ns e 20ns respectivamente. O controle da temperatura foi obtido através de Berendsen (Berendsen et al., 1984), onde o sistema é lentamente aquecido até que se atinja a temperatura fixada e pressão constante de 1 bar. Os cálculos foram realizados utilizando-se os campos de força Gromos 43A1 (Daura et al., 1998) e Gromos45A3 (Schuler et al., 2001) como mostrados na tabela 3. O sistema foi simulado utilizando-se condições periódicas de contorno com um raio de corte de 1,0 nm para interações de curta distância, atualizando-se a lista de vizinhos a cada 10 passos. Interações eletrostáticas de longa distância foram tratadas com o método *Particle Mesh Ewald* -PME (Darden et al., 1993). O algoritmo LINCS foi usado para restringir ligações covalentes das proteínas e o algoritmo SETTLE (Miyamoto et al., 1992) foi usado para restringir a geometria das moléculas de água; e as equações de movimento nas simulações da dinâmica molecular final foram integradas com o algoritmo *leapfrog* (Hockney et al., 1974).

Para avaliar o comportamento global das estruturas 3D das proteínas foi utilizado a Raiz do Desvio Médio Quadrático (RMSD) que calcula os desvios entre as conformações das proteínas no decorrer da simulação de dinâmica molecular em relação a uma estrutura de referência; o Raio de Giro (Rg) para observar a compactação e expansão da estrutura 3D das proteínas; e a variação da energia (E) em relação ao tempo, tomando como referência as estruturas 3D de cada uma das proteínas após a simulação de dinâmica molecular de restrição.

5.19. Inserção e visualização dos sítios de fosforilação

A inserção e visualização das modificações de fosforilação na estrutura 3D das espidroínas foram realizadas no ambiente gráfico do PyMol (Delano, 2002). E a análise e visualização dos arquivos produzidos pela simulação da dinâmica molecular foram editadas utilizando o ambiente gráfico VMD - *Visual Molecular Dynamics* (Humphrey et al., 1996). Os gráficos foram construídos no programa Excel 2010.

5.20. Métodos de análise do modelo molecular

Para as análises dos modelos das espidroínas foram utilizados a função objetiva do MODELLER (Shen & Šali, 2006) e o programa PROCHECK (Laskowski et al., 1993) que pertence ao pacote CCP4 (*Collaborative Computational Project No 4*) utilizado para verificar a qualidade estereoquímica de moléculas determinadas experimentalmente. O programa PROCHECK analisa a geometria global da estrutura ou cada resíduo individualmente, utilizando parâmetros estereoquímicos derivados de estruturas de alta resolução ou bem refinados que constituem sua base de dados. Os parâmetros estereoquímicos utilizados são aqueles descritos por Morris et al. (1992): ligações covalentes, planaridade de grupos planares (aromáticos, ligações peptídicas, etc.) ângulos diédricos, quiralidade, interações não covalentes, ligações de hidrogênio da cadeia principal e pontes de dissulfeto. Os resultados são apresentados em vários arquivos diferentes, entre estes arquivos foram escolhidos o diagrama de Ramachandran e o G-factor.

- O diagrama de Ramachandran analisa a qualidade estereoquímica dos modelos construídos, pois representa através de um plano, a posição dos aminoácidos de acordo com os ângulos de torção ϕ e ψ formados entre o carbono- α e o nitrogênio (C α -N) e os ângulos formados entre o carbono e o carbono- α (C-C α) respectivamente. Em princípio, ϕ e ψ podem ter qualquer valores entre -180° e $+180^\circ$, porém, muitos valores dos ângulos ϕ e ψ são proibidos pelas interferências estéricas entre átomos pertencentes ao mesmo esqueleto polipeptídico e pelas cadeias laterais dos resíduos de aminoácidos (Gibas et al., 2001).
- O G-factor verifica os comprimentos de ligação da cadeia principal, ângulos da cadeia principal, distribuição dos ângulos Φ/ψ , distribuição dos ângulos χ_1/χ_2 e ω . Fornece uma medida do desvio de uma dada propriedade estereoquímica (ângulos de torção e geometria covalente), baseado nas distribuições observadas nos parâmetros estereoquímicos em sua base de dados e uma medida da normalidade da estrutura como um todo. O cálculo do G-factor global envolve os diversos G-factors para toda a estrutura, num único número. O G-factor global é a melhor medida da qualidade estereoquímica total de uma molécula. Por exemplo, os resíduos que caem nas regiões não permitidas do diagrama de Ramachandran terão um G-factor baixo. Se uma proteína tiver muitos resíduos com G-factor baixo sugere que algo pode estar errado com a geometria total da proteína avaliada.

Devido a grande quantidade de dados que foram gerados e, para que a interpretação e leitura desses dados no decorrer da tese possam ocorrer de forma compreensível, optou-se por apresentar os resultados e a discussão divididos em duas partes independentes, mas que se complementam:

Parte 1 - Análise proteômica das glândulas produtoras de seda da aranha *Nephila clavipes*

Parte 2 - Análise estrutural e modificações pós-traducionais das espidroínas da seda da teia produzida pela aranha *Nephila clavipes*

6. Resultados e discussão

6.1. Parte 1 - Análise proteômica das glândulas produtoras de seda da aranha *Nephila clavipes*

As proteínas constituintes das glândulas produtoras de seda da aranha *N. clavipes* foram submetidas a uma análise proteômica comparativa. As aranhas foram dissecadas para a retirada das glândulas ampulada maior, ampulada menor e flageliforme localizadas no abdômen. Essas glândulas foram selecionadas por sintetizarem fios de sedas de alta resistência, extensibilidade e elasticidade que constituem os fios de seda da teia orbital e do frame. A glândula ampulada maior é a mais volumosa e sua forma distinta em forma de ampola facilitou a identificação no momento da dissecação (Figura 12). Além desses três tipos de glândulas selecionadas inicialmente, também foi possível identificar outros dois tipos de glândulas: a agregada e a tubuliforme. Sendo assim, essas glândulas também foram coletadas para a extração de suas proteínas e submetidas à análise proteômica. Dos sete tipos glandulares presentes na *N. clavipes*, cinco foram identificados e coletados em quantidades suficientes para a realização do estudo; somente não foi possível identificar e coletar as glândulas aciniforme e piriforme por apresentarem uma consistência homogênea o que dificultou a separação das mesmas.

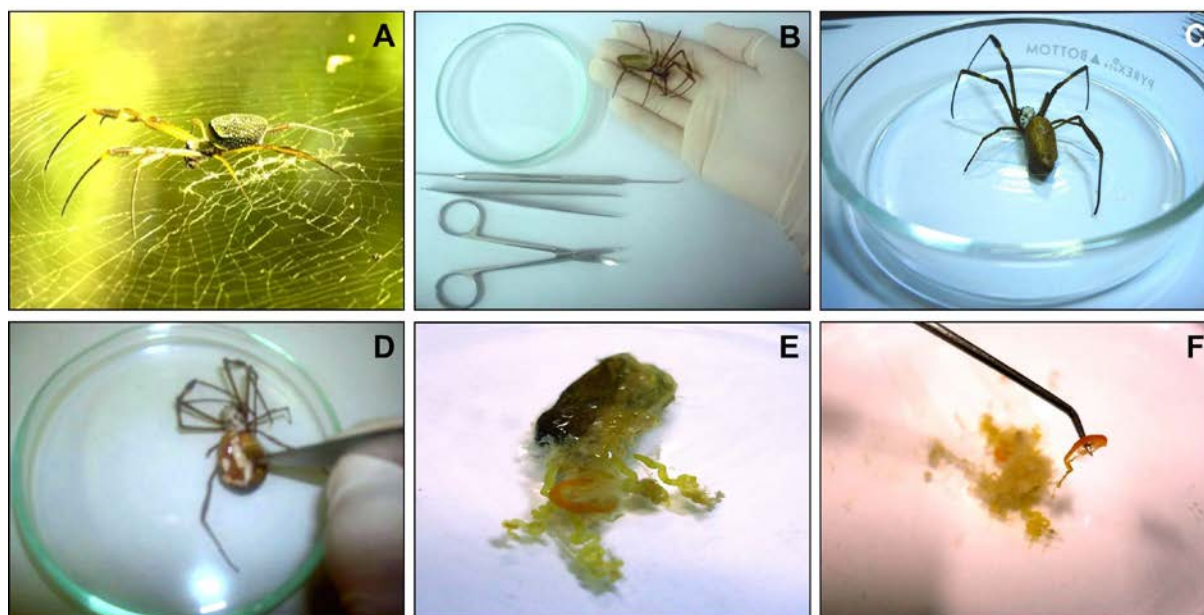


Figura 12. Dissecação da aranha *N. clavipes* com a identificação e coleta das glândulas ampulada maior, ampulada menor, flageliforme, agregada e tubuliforme. **A** - *N. clavipes*; **B-E** - Dissecação e identificação das glândulas; **F** - Coleta da glândula ampulada maior.

Após a extração das proteínas, todas as amostras foram quantificadas pelo método de Bradford e submetidas à eletroforese bidimensional - 2-DE (n=3), apresentando um alto grau de identidade entre os três géis analisados de cada amostra individualmente. A figura 13 mostra o perfil da eletroforese bidimensional da glândula ampulada maior, sendo possível visualizar um total de 70 *spots* com MW entre 10 e 80 kDa e pI de 3 a 10 (gradiente não linear). Para a identificação por espectrometria de massas, os *spots* mais dominantes foram selecionados, excisados dos géis e submetidos à digestão enzimática com tripsina. A tabela 6 mostra a identificação de 63 proteínas. Uma série de proteínas como hemocianina (*spots* 1 a 24; 26 a 31), enolase (*spots* 32, 33 e 35), 2-fosfo-D-glicerato hidrolase (*spots* 34, 36 a 38), actina (*spots* 39 a 42), transaldolase (*spot* 43), fator de alongamento de tradução 2 (*spot* 44), gliceraldeído 3-fosfato desidrogenase (*spots* 45 a 49; 57), proteína de choque térmico (*spots* 50 a 51; 53 a 54), dihidropteridina redutase (*spot* 52), proteína sarcoplasmática ligante de cálcio (*spot* 55), espidroína-1 (*spot* 56), tioredoxina peroxidase (*spots* 60 e 61), superóxido dismutase (*spot* 62), peptidil-prolil cis-trans isomerase (*spots* 63 e 64), nucleosídeo difosfato quinase (*spot* 66) e ubiquitina (*spot* 70) foram identificadas.

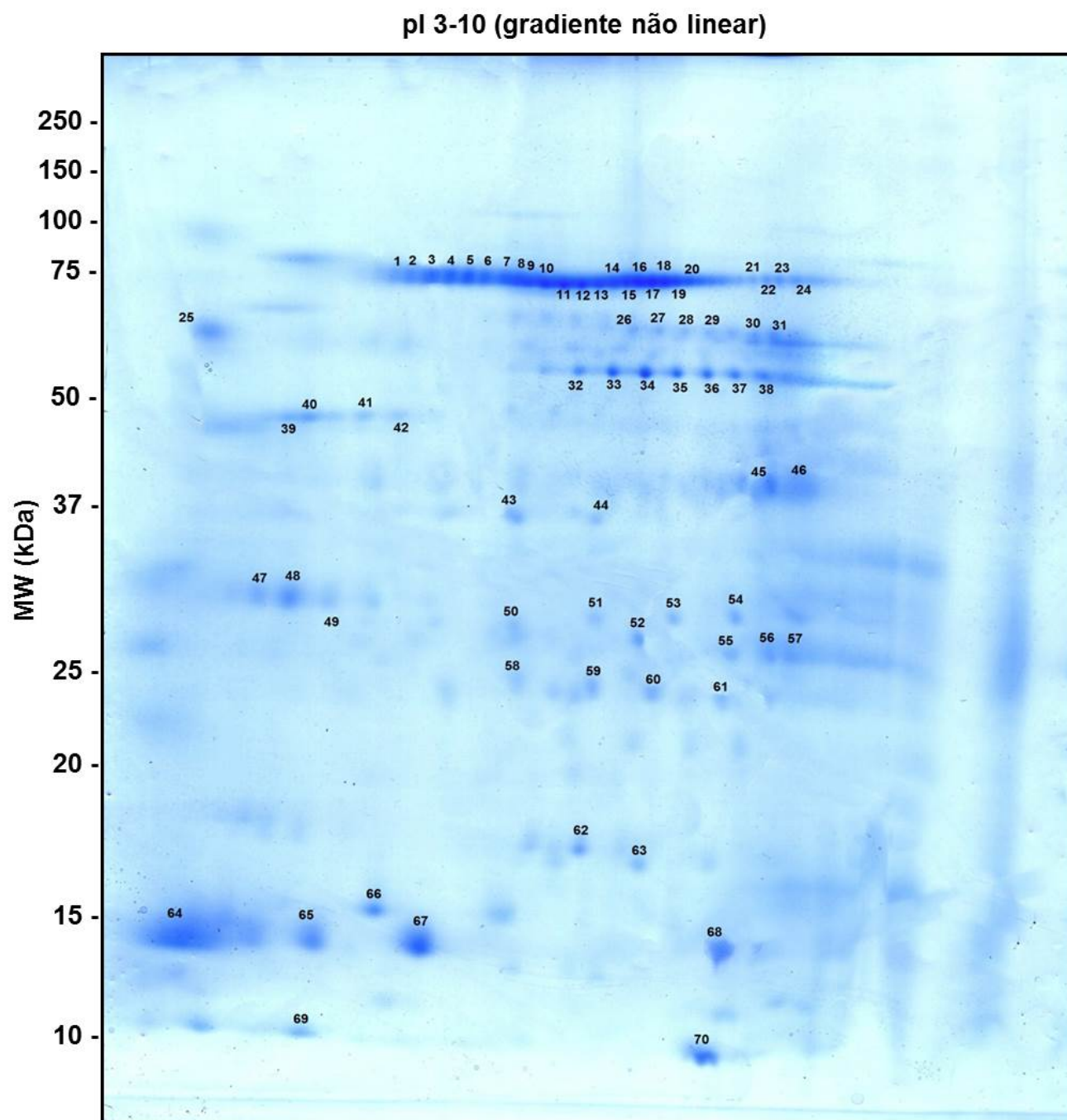


Figura 13. Perfil eletroforético da 2-DE da glândula ampulada maior da aranha *N. clavipes*.

Tabela 6. Proteínas identificadas na glândula ampulada maior da aranha *N. clavipes* através da digestão *in-gel* dos spots oriundos da 2-DE.

Spot	Código de acesso	Taxonomia	Proteína	Mascot Score	% Cobertura	Exp. / Calc. MW (kDa)	Sequência peptídica MS/MS (Ion Score)
1	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	940	41	78623 / 71709	VCSLFTHLTSISR (78); LAQGELFSCFHEK (51); DFEDFINLAK (88); SFANEGLFVYAASVAILHR (54); GVTVPPIQEIFPDR (66); FVPSETISLALK (55); EVTNHPDKDIEVEIESTGNILDPEYK (49); LTEDNGIDILGSIVESSYESK (56); LFYGSLHNWGHVMMANITDPDGR (44); FNENPGVMSDTSTSLRDPFYR (60); AQLDFPGVQVVNVTNNAK (124); VPNLVTTFMK (44); TLCIELDKFQVELAAGK (77); QLLAGEGVSENTTEFCSCGWPEHMLIPR (95)
2	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	1201	42	77482 / 71709	VCSLFTHLTSISR (69); LAQGELFSCFHEK (49); DFEDFINLAK (88); GVTVPPIQEIFPDR (55); FVPSETISLALK (64); EVTNHPDKDIEVEIESTGNILDPEYK (53); KGELFYMHQQMCAR (75); LTEDNGIDILGSIVESSYESK (64); LFYGSLHNWGHVMMANITDPDGR (67); FNENPGVMSDTSTSLR (126); DPIFYR (26); AQLDFPGVQVVNVTNNAK (142); VPNLVTTFMK (45); TLCIELDKFQVELAAGK (75); LSSVTVSEHTFK (98); QLLAGEGVSENTTEFCSCGWPEHMLIPR (81); KPMGFPPDR (31)
3	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	1093	37	77341 / 71709	VCSLFTHLTSISR (73); LAQGELFSCFHEK (50); DFEDFINLAK (88); GVTVPPIQEIFPDR (62); FVPSETISLALK (57); EVTNHPDKDIEVEIESTGNILDPEYK (42); DIEVEIESTGNILDPEYK (40); KGELFYMHQQMCAR (56); LTEDNGIDILGSIVESSYESK (55); FNENPGVMSDTSTSLR (116); FNENPGVMSDTSTSLRDPFYR (27); FIDNIFQDYK (17); AQLDFPGVQVVNVTNNAK (151); VPNLVTTFMK (37); TLCIELDK (18); TLCIELDKFQVELAAGK (70); LSSVTVSEHTFK (81); YPDKKPMGFPPDR (23); KPMGFPPDR (29)
4	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	1123	41	77341 / 71709	LAQGELFSCFHEK (57); DFEDFINLAK (88); SFANEGLFVYAASVAILHR (108); GVTVPPIQEIFPDR (81); FVPSETISLALK (42); EVTNHPDKDIEVEIESTGNILDPEYK (65); MSYFR (20); KGELFYMHQQMCAR (38); LTEDNGIDILGSIVESSYESK (23); FNENPGVMSDTSTSLR (98); AQLDFPGVQVVNVTNNAK (129); VPNLVTTFMK (38); YDELGNR (30); TLCIELDKFQVELAAGK (70); FQVELAAGK (46); LSSVTVSEHTFK (92); QLLAGEGVSENTTEFCSCGWPEHMLIPR (75); DQKYPDKKPMGFPPDR (33)
5	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	880	27	76781 / 71709	VCSLFTHLTSISR (83); LHGIGK (19); DFEDFINLAK (88); GVTVPPIQEIFPDR (56); FVPSETISLALK (44); EVTNHPDKDIEVEIESTGNILDPEYK (85); MSYFR (24); LFYGSLHNWGHVMMANITDPDGR (71); FNENPGVMSDTSTSLR (97); FNENPGVMSDTSTSLRDPFYR (61); DPIFYR (27); AQLDFPGVQVVNVTNNAK (85); YDELGNR (40); TLCIELDKFQVELAAGK (75); FQVELAAGK (24)
6	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	1113	41	76503 / 71709	VCSLFTHLTSISR (92); DFEDFINLAK (88); SFANEGLFVYAASVAILHR (31); GVTVPPIQEIFPDR (53); FVPSETISLALK (46); EVTNHPDKDIEVEIESTGNILDPEYK (51); DIEVEIESTGNILDPEYK (16); KGELFYMHQQMCAR (54); LTEDNGIDILGSIVESSYESK (51); LFYGSLHNWGHVMMANITDPDGR (64); FNENPGVMSDTSTSLR (95); FNENPGVMSDTSTSLRDPFYR (51); DPIFYR (20); AQLDFPGVQVVNVTNNAK (123); VPNLVTTFMK (46); TLCIELDKFQVELAAGK (97); FQVELAAGK (40); QLLAGEGVSENTTEFCSCGWPEHMLIPR (61); KPMGFPPDR (40)
7	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	963	34	76642 / 71709	VCSLFTHLTSISR (79); LHGIGK (9); LAQGELFSCFHEK (50); DFEDFINLAK (88); SFANEGLFVYAASVAILHR (50); GVTVPPIQEIFPDR (60); FVPSETISLALK (38); DIEVEIESTGNILDPEYK (79); LTEDNGIDILGSIVESSYESK (61); LFYGSLHNWGHVMMANITDPDGR (69); AQLDFPGVQVVNVTNNAK (98); VPNLVTTFMK (31); TLCIELDKFQVELAAGK (75); FQVELAAGK (41); LSSVTVSEHTFK (93); KPMGFPPDR (46)
8	Q86N91	<i>Nephila inaurata</i>	Hemocyanin subunit E	1013	38	75815 / 71709	VCSLFTHLTSISR (79); LAQGELFSCFHEK (51); DFEDFINLAK (88);

		<i>madagascariensis</i>					SFANEGLFVYAASVAILHR (44); GVTVPPIQEIFPDR (52); FVPSETISLALK (50); EVTNHPDKDIEVEIESTGNILDPEYK (60); DIEVEIESTGNILDPEYK (34); DRKGELFYMHQQMCAR (36); KGELFYMHQQMCAR (77); LTEDNGIDILGSIVESSYESK (52); FNENPGVMSDTSTSLRDPIFYR (56); AQLDFPGVQVNVTVNAK (148); VPNLVTTFMK (45); TLCIELDKFQVELAAGK (72); QLLAGEGVSENTTEFCSCGWPEHMLIPR (79)
9	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	939	27	75952 / 71786	QDQILPLFEK (50); LIGVGVLPDR (60); AADFADFLNLAK (114); DVVNEGLFVYALSAVAVVHR (49); GVTLPPIQEIVFPDR (71); KGELFYMHQQMCAR (80); MVAFHNFEEK (62); DLTEVDVQDMER (72); ILEAIDLK (49); FQENPGVMSDTSTSLR (95); FQENPGVMSDTSTSLRDPIFYR (47); FVDNIFQQYK (61); YHHLDEHPFSITVNAQNSSGAPK (78); QLAPGNNVISR (22)
10	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	1002	32	75542 / 71786	QDQILPLFEK (49); LIGVGVLPDR (60); AADFADFLNLAK (114); DVVNEGLFVYALSAVAVVHR (105); GVTLPPIQEIVFPDR (64); KGELFYMHQQMCAR (86); LEGYAPHLTSLVSGLHYASR (62); DLTEVDVQDMER (57); GFYGLHNWGHVMMAR (43); FQENPGVMSDTSTSLR (104); FVDNIFQQYK (65); YHHLDEHPFSITVNAQNSSGAPK (69); TICNDAVSYCGAK (65); LPTENTCVTDIK (59)
11	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	1096	31	74865 / 71786	QDQILPLFEK (52); LIGVGVLPDR (71); AADFADFLNLAK (114); DVVNEGLFVYALSAVAVVHR (44); GVTLPPIQEIVFPDR (65); FIPAETINLASK (40); KGELFYMHQQMCAR (57); DLTEVDVQDMER (89); ILEAIDLK (55); GFYGLHNWGHVMMAR (43); FQENPGVMSDTSTSLR (130); FQENPGVMSDTSTSLRDPIFYR (63); FVDNIFQQYK (67); LSPDEQRPLFIELDK (62); TICNDAVSYCGAK (73); LPTENTCVTDIK (73)
12	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	903	33	75542 / 71786	LIGVGVLPDR (65); AADFADFLNLAK (113); DVVNEGLFVYALSAVAVVHR (22); GVTLPPIQEIVFPDR (49); FIPAETINLASK (42); MVAFHNFEEK (52); LEGYAPHLTSLVSGLHYASR (42); DLTEVDVQDMER (89); GNEVPLDETNGANILGTIEASSDSPNK (39); GFYGLHNWGHVMMAR (46); FQENPGVMSDTSTSLR (125); FQENPGVMSDTSTSLRDPIFYR (44); FVDNIFQQYK (63); QLAPGNNVISR (46); LPTENTCVTDIK (69)
13	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	860	30	75270 / 71786	LIGVGVLPDR (68); AADFADFLNLAK (113); DVVNEGLFVYALSAVAVVHR (43); GVTLPPIQEIVFPDR (66); FIPAETINLASK (46); KGELFYMHQQMCAR (35); GELFYMHQQMCAR (33); MVAFHNFEEK (29); DLTEVDVQDMER (92); FQENPGVMSDTSTSLRDPIFYR (63); FVDNIFQQYK (61); YHHLDEHPFSITVNAQNSSGAPK (55); QLAPGNNVISR (30); TICNDAVSYCGAK (51); LPTENTCVTDIK (75)
14	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	993	35	75406 / 72326	RILPLFEHLASLTR (75); ILPLFEHLASLTR (68); DIVNEGFMVYSTSVALLHR (36); GVTVPPIQEIFPDR (37); FIPSETINQAIK (30); KGELFYMHQQMCAR (61); LEGYSPHLTSLVSGHHYANR (10); PAGLSLHDLNELVDVQDMAR (57); YKENPGVMSDTSTSLRDPIFYR (19); ENPGVMSDTSTSLRDPIFYR (19); FIDNIFQEYK (75); DLDFFPGSVVNVTVNAK (123); LPNIVNTYK (66); RFFIELDKFHAELAPGK (42); FFIELDKFHAELAPGK (32); SICDAVSYCGAK (90); SMGFPPDR (28); TLAEFSSANMK (89)
15	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	1101	32	75678 / 72326	RILPLFEHLASLTR (103); DIVNEGFMVYSTSVALLHR (44); GVTVPPIQEIFPDR (47); FIPSETINQAIK (45); KGELFYMHQQMCAR (23); GELFYMHQQMCAR (24); MIPYHNFNEK (23); EYYGNLHNSGHVMMARI (40); ENPGVMSDTSTSLRDPIFYR (30); FIDNIFQEYK (73); DLDFFPGSVVNVTVNAK (130); LPNIVNTYK (69); RFFIELDKFHAELAPGK (71); FFIELDKFHAELAPGK (99); SICDAVSYCGAK (101); SMGFPPDR (33); TLAEFSSANMK (84)
16	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	1095	33	75678 / 72326	RILPLFEHLASLTR (58); ILPLFEHLASLTR (75); DFEDFMR (40); DIVNEGFMVYSTSVALLHR (37); GVTVPPIQEIFPDR (67); FIPSETINQAIK (41); KGELFYMHQQMCAR (72); PAGLSLHDLNELVDVQDMAR (91);

							YKENPGVMSDTSTSLRDPIFYR (39); ENPGVMSDTSTSLRDPIFYR (18); FIDNIFQEYK (74); DLDFPGVSVVNVTVNAK (127); LPNIVNTYLK (63); FFIELDKFHAELAPGK (52); SICADAVSYCGAK (86); SMGFPPDR (18); TLAEFSSANMK (80)
17	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	1152	32	75542 / 72326	RILPLFEHLASLTR (63); ILPLFEHLASLTR (91); DIVNEGMFVYSTSVALLHR (38); GVTVPPIQEIPDR (59); FIPSETINQAIK (56); KGELFYMHQQMCAR (69); GELFYMHQQMCAR (53); MIPYHNFNEK (32); PAGLSLHDLNELVDVQDMAR (77); YKENPGVMSDTSTSLR (41); YKENPGVMSDTSTSLRDPIFYR (17); ENPGVMSDTSTSLRDPIFYR (61); FIDNIFQEYK (73); KDLDFPGVSVVNVTVNAK (35); DLDFPGVSVVNVTVNAK (90); LPNIVNTYLK (63); EDQLEVSHGVSLK (62); RFFIELDKFHAELAPGK (74); FFIELDKFHAELAPGK (76); SMGFPPDR (33)
18	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	1392	38	75406 / 72326	RILPLFEHLASLTR (76); ILPLFEHLASLTR (73); DFEDFMR (39); DIVNEGMFVYSTSVALLHR (43); GVTVPPIQEIPDR (51); FIPSETINQAIK (60); KGELFYMHQQMCAR (43); GELFYMHQQMCAR (32); MIPYHNFNEK (30); PAGLSLHDLNELVDVQDMAR (80); YKENPGVMSDTSTSLR (30); YKENPGVMSDTSTSLRDPIFYR (48); ENPGVMSDTSTSLRDPIFYR (40); FIDNIFQEYK (74); DLDFPGVSVVNVTVNAK (122); LPNIVNTYLK (57); LPNIVNTYLKEDQLEVSHGVSLK (67); EDQLEVSHGVSLK (95); LSLNEAR (30); RFFIELDKFHAELAPGK (57); FFIELDKFHAELAPGK (72); SICADAVSYCGAK (81); SMGFPPDR (22); TLAEFSSANMK (81)
19	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	1384	38	75815 / 72326	RILPLFEHLASLTR (62); ILPLFEHLASLTR (73); DFEDFMR (38); DIVNEGMFVYSTSVALLHR (53); GVTVPPIQEIPDR (57); FIPSETINQAIK (48); KGELFYMHQQMCAR (72); GELFYMHQQMCAR (37); MIPYHNFNEK (30); PAGLSLHDLNELVDVQDMAR (74); YKENPGVMSDTSTSLRDPIFYR (39); ENPGVMSDTSTSLRDPIFYR (48); FIDNIFQEYK (70); DLDFPGVSVVNVTVNAK (137); LPNIVNTYLK (67); EDQLEVSHGVSLK (76); LSLNEAR (38); RFFIELDKFHAELAPGK (61); FFIELDKFHAELAPGK (66); SICADAVSYCGAK (98); SMGFPPDR (35); TLAEFSSANMK (84)
20	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	870	30	75952 / 72326	RILPLFEHLASLTR (72); ILPLFEHLASLTR (64); DFEDFMR (32); DIVNEGMFVYSTSVALLHR (36); GVTVPPIQEIPDR (71); FIPSETINQAIK (47); PAGLSLHDLNELVDVQDMAR (59); ENPGVMSDTSTSLRDPIFYR (32); FIDNIFQEYK (61); DLDFPGVSVVNVTVNAK (99); LPNIVNTYLK (62); PNIVNTYLKEDQLEVSHGVSLK (59); FFIELDKFHAELAPGK (57); SICADAVSYCGAK (77); SMGFPPDR (22)
21	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	670	30	76089 / 72483	DFGDFINLAK (74); GVILPPIQEVFPDK (64); GVILPPIQEVFPDKFIPAETINR (16); TGNIIDPEYNLAYFR (81); KGELFYMHQQMCAR (54); LEGYSAHLISLISGLNYASR (54); ILSAIHTGQVIDSNGQEVPLDLER (35); GLDILGALIESSQESLNK (72); ENPGVMDDTSTALRDPIFYR (25); WMDNIFQEYK (51); SVDSSVTLSHVPTFEELEK (69); ETEYCSCGWPEHMLVPR (58)
22	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	935	32	76089 / 72483	DFGDFINLAK (74); GVILPPIQEVFPDK (57); FIPAETINR (43); TGNIIDPEYNLAYFR (112); KGELFYMHQQMCAR (73); LEGYSAHLISLISGLNYASR (81); ILSAIHTGQVIDSNGQEVPLDLER (50); GLDILGALIESSQESLNK (118); FKENPGVMDDTSTALRDPIFYR (41); ENPGVMDDTSTALRDPIFYR (13); WMDNIFQEYK (71); ETEYCSCGWPEHMLVPR (47); VLDDSATVVCSDAVSYCGAK (77); AMGYPPDRPIK (38)
23	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	619	24	76089 / 72483	LFSCFHEDHLGEAQALYEVLYEAK (43); DFGDFINLAK (74); FIPAETINR (37); TGNIIDPEYNLAYFR (91); KGELFYMHQQMCAR (65); GLDILGALIESSQESLNK (73); FKENPGVMDDTSTALRDPIFYR (24); ENPGVMDDTSTALRDPIFYR (23); WMDNIFQEYK (75); ETEYCSCGWPEHMLVPR (56)
24	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	736	28	76089 / 72483	DFGDFINLAK (71); GVILPPIQEVFPDK (53); TGNIIDPEYNLAYFR (70); KGELFYMHQQMCAR (68); ILSAIHTGQVIDSNGQEVPLDLER (26); GLDILGALIESSQESLNK (120); FKENPGVMDDTSTALRDPIFYR (20);

							ENPGVMDDTSTALRDPIFYR (22); WMDNIFQEYK (69); SVDSSVTLSHVPTFEELEK (87); ETEYCSCGWPEHMLVPR (66); ETEYCSCGWPEHMLVPR (42); AMGYPFDRPIK (26)
26	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	431	14	62537 / 72326	ILPLFEHLASLTR (64); GVTVPPIQEIPDR (34); KGELFYMHQMCAR (77); YKENPGVMSDTSTSLRDPIFYR (87); FIDNIFQEYK (66); DLDFPGVSVNVTVNAK (107)
27	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	298	11	62433 / 72326	ILPLFEHLASLTR (68); GVTVPPIQEIPDR (52); YKENPGVMSDTSTSLR (9); FIDNIFQEYK (70); DLDFPGVSVNVTVNAK (103)
28	Q9NFL5	<i>Aphonopelma californicum</i>	Hemocyanin F chain	130	6	61920 / 72674	TVQDK (4); GVTVPPIQEIPDR (48); FVPAETVNQAVK (12); FIDNIFQEYK (66)
29	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	244	9	61920 / 72326	DIVNEGFMVYSTSVALLHR (25); GVTVPPIQEIPDR (45); FIDNIFQEYK (74); DLDFPGVSVNVTVNAK (101)
30	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	124	8	60418 / 72483	TGNIIDPEYNLAYFR (53); KGELFYMHQMCAR (32); ILSAHTGQVIDSNGQEVPLDLER (39)
31	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	181	8	60124 / 72483	DFGDFINLAK (60); ILSAHTGQVIDSNGQEVPLDLER (76); FKENPGVMDDTSTALRDPIFYR (47)
32	A1YQ87	<i>Bombyx mori</i>	Enolase	180	5	53642 / 47164	LGANAILGVSLAVAK (107); DGKYDLDFK (47); YDLDFK (29)
33	G1ERJ1	<i>Bucculatrix staintonella</i>	Enolase	267	10	53804 / 40877	SKLGANAILGVSLAVAK (57); LGANAILGVSLAVAK (94); GVPLYR (27); HLADLAGVK (66); YDLDFK (23)
34	Q6QWQ0	<i>Phormictopus sp. SBH266263</i>	2-phospho-D-glycerate hydrolase	204	10	53401 / 40026	KFGLAATGVGDEGGFAPDIQENK (18); FGLAATGVGDEGGFAPDIQENK (113); YDLDFK (23); SNGWGTMVSHR (52)
35	G1ERJ1	<i>Bucculatrix staintonella</i>	Enolase	257	8	53321 / 40877	SKLGANAILGVSLAVAK (48); LGANAILGVSLAVAK (113); HLADLAGVK (65)
36	Q6QWQ0	<i>Phormictopus sp.</i>	2-phospho-D-glycerate hydrolase	178	7	53321 / 40026	FGLAATGVGDEGGFAPDIQENK (142); YDLDFK (36)
37	Q6QWP9	<i>Limulus polyphemus</i>	2-phospho-D-glycerate hydrolase	177	8	53242 / 39993	IEIGMDVAASEFHK (79); YDLDFK (22); VNQIGSVTEAIK (76)
38	Q6QWQ0	<i>Phormictopus sp.</i>	2-phospho-D-glycerate hydrolase	131	6	52689 / 40026	KFGLAATGVGDEGGFAPDIQENK (38); FGLAATGVGDEGGFAPDIQENK (95)
39	B2XY36	<i>Latrodectus hesperus</i>	Actin	510	29	46588 / 34974	AVFPSIVGR (54); IWHHTFYNELR (65); DLTDYLMK (39); LCYVALDFEQEMATAASSSSLEKS (73); SYELPDGQVITIGNER (63); KDLYANTVLSGGTTMYPGIADR (72); DLYANTVLSGGTTMYPGIADR (111)
40	B3VN78	<i>Bombyx mandarina</i>	Actin	674	36	46338 / 42174	AVFPSIVGR (57); AVFPSIVGRPR (41); VAPEEHPVLLTEAPLNPK (79); TTGIVLDSGDGVSHTVPIYEGYALPHAILR (42); DLTDYLMK (42); GYSFTTTAER (52); SYELPDGQVITIGNER (89); DLYANTVLSGGTTMYPGIADR (83); EITALAPSTMK (41); QEYDESGPSIVHR (66)
41	B2XY36	<i>Latrodectus hesperus</i>	Actin	739	56	46588 / 34974	AVFPSIVGR (48); IWHHTFYNELR (75); VAPEEHPVLLTEAPLNPK (55); TTGIVLDSGDGVSHTVPIYEGYALPHAILR (72); DLTDYLMK (34); GYSFTTTAER (58); EKLCYVALDFEQEMATAASSSSLEK (59); LCYVALDFEQEMATAASSSSLEK (54); SYELPDGQVITIGNER (91); CPEALFQPSFLGMESCGIHETTYNSIMK (51); KDLYANTVLSGGTTMYPGIADR (51); DLYANTVLSGGTTMYPGIADR (92)
42	B2XY36	<i>Latrodectus hesperus</i>	Actin	476	28	46588 / 34974	IWHHTFYNELR (45); DLTDYLMK (56); GYSFTTTAER (48); LCYVALDFEQEMATAASSSSLEK (58); SYELPDGQVITIGNER (83); KDLYANTVLSGGTTMYPGIADR (78); DLYANTVLSGGTTMYPGIADR (108)
43	E7D1B4	<i>Latrodectus hesperus</i>	Transaldolase	114	8	35881 / 32016	LFVLFGCEILK (64); ILDWYVANTDQK (50)
44	Q1HPK6	<i>Bombyx mori</i>	Translation elongation factor 2	103	5	35489 / 95750	DEQDRCITIKSTAISMFFEELEK (12); STAISMFFEELEKDLVFITNPDQR (5); AYLPVNESFGFTADLR (90)
45	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate	495	36	38323 / 33251	YMVYMFK (27); IHVFQEMKPSSIPWGK (68); VGAEYVVESTGVFTTLEK (88); VINDNFGIVEGLMTTVHAITATQK (41); GAAQNIIPASTGAAK (81); LTGMAFR (32);

			dehydrogenase				VPTPDVSVVDTLCR (87); AGISLNNNFVK (72)
46	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	551	40	38406 / 33251	IHFVQEMKPSSIPWGK (60); VGAEYVVESTGVFTTLEK (93); VINDNFGIVEGLMTTVHAITATQK (79); GAAQNIIPASTGAAK (80); VPTPDVSVVDTLCR (90); GILGYTEDEVSSDFIGDSHSSIFDAK (56); AGISLNNNFVK (93)
47	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	224	22	30283 / 33251	VGAEYVVESTGVFTTLEK (102); VINDNFGIVEGLMTTVHAITATQK (39); GILGYTEDEVSSDFIGDSHSSIFDAK (83)
48	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	98	13	30067 / 33251	VGAEYVVESTGVFTTLEK (68); VINDNFGIVEGLMTTVHAITATQK (30);
49	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	71	14	29986 / 33251	VGAEYVVESTGVFTTLEK (31); GILGYTEDEVSSDFIGDSHSSIFDAK (40)
50	B0WP93	<i>Culex quinquefasciatus</i>	Heat shock 70 kDa protein cognate 4	260	7	28306 / 71787	SFFPEEISSMVLTK (61); TVTNAVVTVPAYFNDSQR (94); IINEPTAAAIAYGLDK (105)
51	B7PAR6	<i>Ixodes scapularis</i>	Heat shock protein	230	7	28757 / 71292	TVTNAVVTVPAYFNDSQR (59); DAGTIAGLNVLR (72); IINEPTAAAIAYGLDK (100)
52	E2A057	<i>Camponotus floridanus</i>	Dihydropteridine reductase	114	11	27666 / 25941	AALLETGPMIGYGMK (114); GALGSACVSQFK
53	Q2PPI9	<i>Tetranychus urticae</i>	Heat shock protein	242	5	28783 / 71516	IINEPTAAAIAYGLDK (118); STAGDTHLGGEDFDNR (62)
54	F0J8P3	<i>Amblyomma variegatum</i>	Heat shock protein	578	15	28810 / 56269	HWPFEVISDGGKPK (82); EIAEAYLGK (37); DAGTIAGLNVLR (83); IINEPTAAAIAYGLDK (117); STAGDTHLGGEDFDNR (70); MVNHFVQEFK (48)
55	P86909	<i>Chionoecetes opilio</i>	Sarcoplasmic calcium-binding protein	55	100	26611 / 842	VATVSLPR (55)
56	P19837	<i>Nephila clavipes</i>	Spidroin-1	96	9	26693 / 60607	GGLGGQGAGAAAAAGGAGQGGYGGVSGASAAASAAASR (52); GGLGGQGAGAVAAAAAGGAGQGGYGGVSGAGR (44)
57	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	65	16	26693 / 33251	VINDNFGIVEGLMTTVHAITATQK (17); GILGYTEDEVSSDFIGDSHSSIFDAK (47)
60	A6N9S1	<i>Ornithodoros parkeri</i>	Thioredoxin peroxidase	115	10	24107 / 22143	QITVNDLPVGR (47); LVQAFQFTDK (69)
61	A6N9S1	<i>Ornithodoros parkeri</i>	Thioredoxin peroxidase	114	10	23683 / 22143	QITVNDLPVGR (50); LVQAFQFTDK (64)
62	B0X9L3	<i>Culex quinquefasciatus</i>	Superoxide dismutase	75	19	16785 / 15662	QISLSGPLSILGR (45); HAGDLGNVVADAGGVAK (30)
63	A1KYY3	<i>Suidasia medanensis</i>	Peptidyl-prolyl cis-trans isomerase	199	14	16436 / 17893	TSWLDGK (29); HVVFGQVVEGMDVVK (106)
64	A1KYY3	<i>Suidasia medanensis</i>	Peptidyl-prolyl cis-trans isomerase	89	14	14651 / 17893	TSWLDGK (35); HVVFGQVVEGMDVVK (54)
66	Q201X7	<i>Acyrtosiphon pisum</i>	Nucleoside diphosphate kinase	124	22	15417 / 17231	TFIMIKPDGVQR (65); GDLCIQVGR (46); NIIHGSDAVESANK (5)
70	Q4LAY8	<i>Eucinetus sp.</i>	Ubiquitin	352	28	9700 / 18107	MKIFVK (16); TLTGKTITLEVEVSDTIENVK (20); TITLEVEPSDTIENVK (96); TLDYNIQK (65); ESTLHLVLR (65)

Considerando a grande quantidade de espectros que foram gerados na identificação das sequências peptídicas de todas as proteínas oriundas da 2-DE da glândula ampulada maior, como exemplos serão mostrados somente alguns espectros CID de algumas sequências peptídicas que foram identificadas nas proteínas hemocianina (*spot* 2; Figuras 14 e 15), fator de alongamento de tradução 2 (*spot* 44; Figura 16), espidroína-1 (*spot* 56; Figura 17), tiorredoxina peroxidase (*spot* 61; Figura 18) e ubiquitina (*spot* 70; Figura 19).

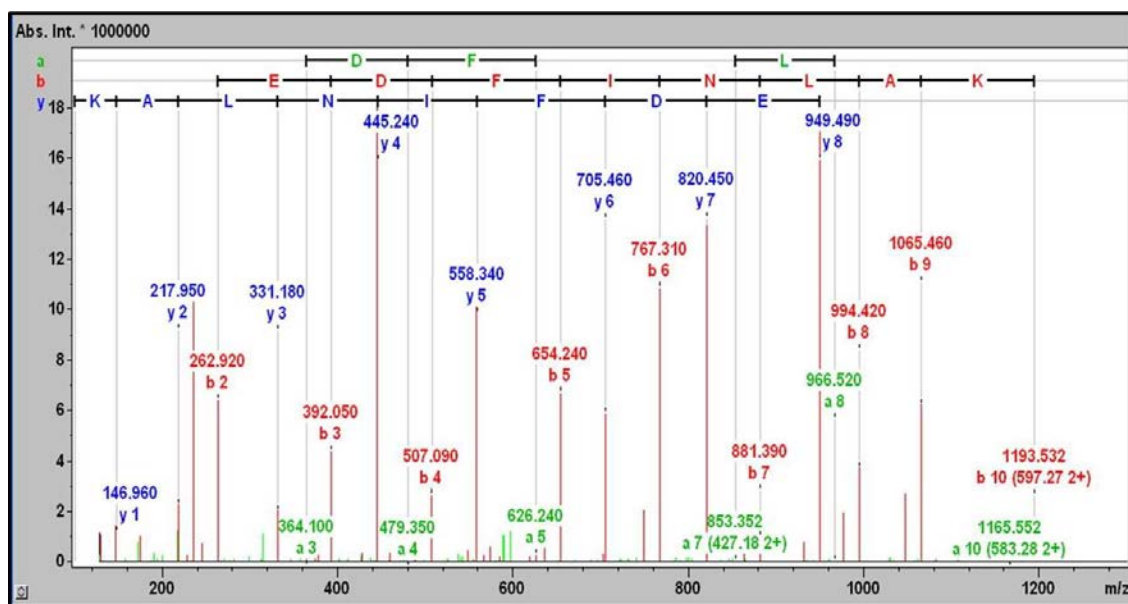


Figura 14. Espectro CID da sequência peptídica DFEDFINLAK, selecionando-se o íon de m/z 606.28 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína hemocianina (*spot* 2), presente na glândula ampulada maior e identificada na 2-DE da seda.

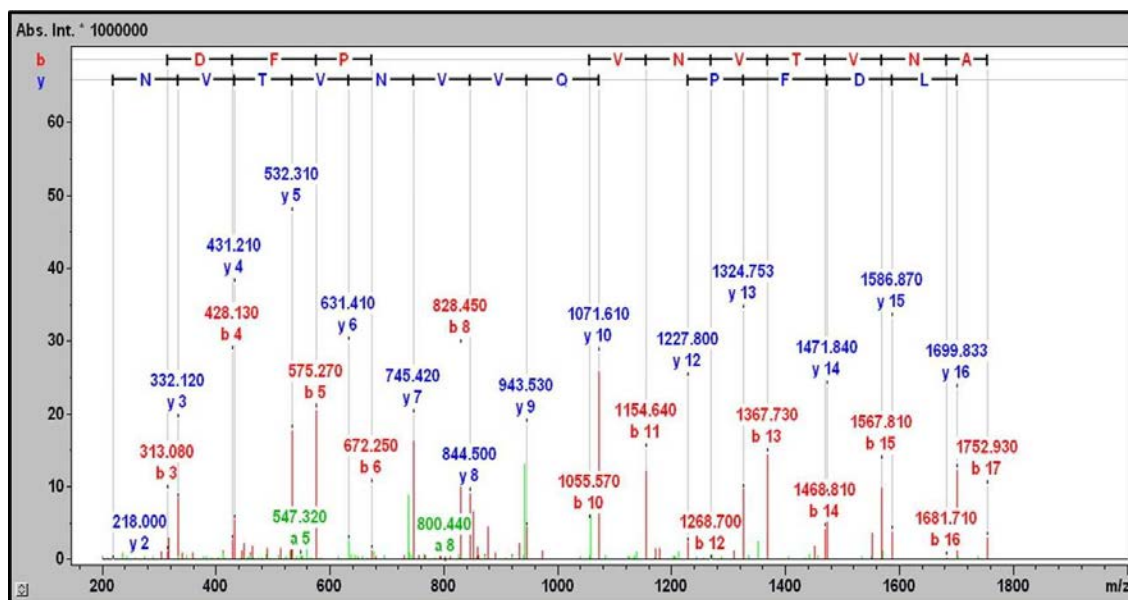


Figura 15. Espectro CID da sequência peptídica AQLDFPGVQVVNVTVNAK, selecionando-se o íon de m/z 950.08 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína hemocianina (spot 2), presente na glândula ampulada maior e identificada na 2-DE da seda.

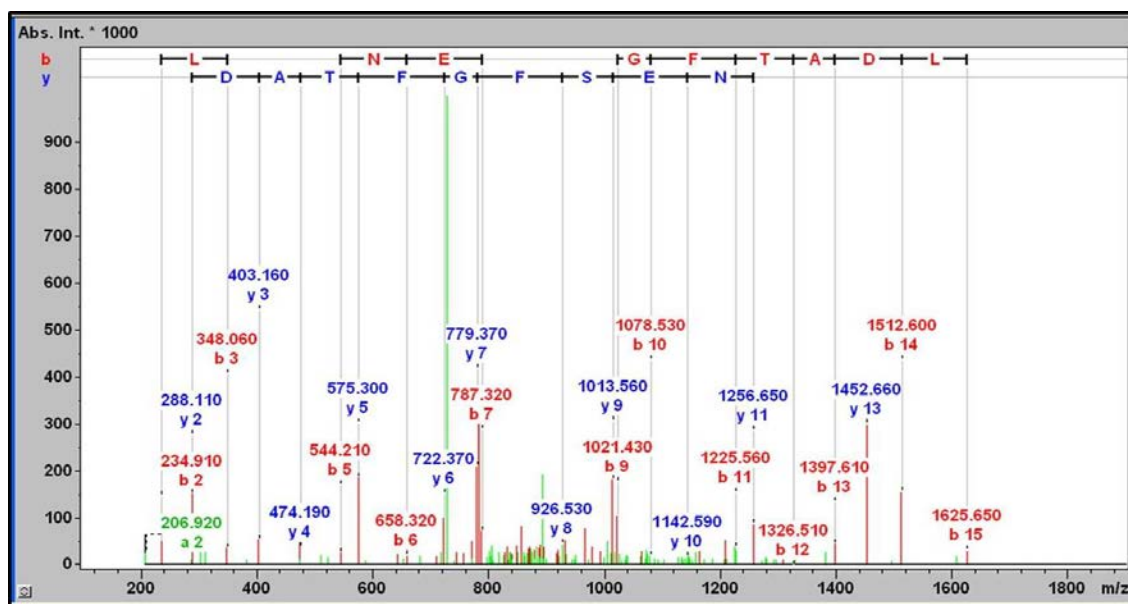


Figura 16. Espectro CID da sequência peptídica AYLPVNESFGFTADLR, selecionando-se o íon de m/z 900.40 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína fator de alongamento de tradução 2 (spot 44), presente na glândula ampulada maior e identificada na 2-DE da seda.

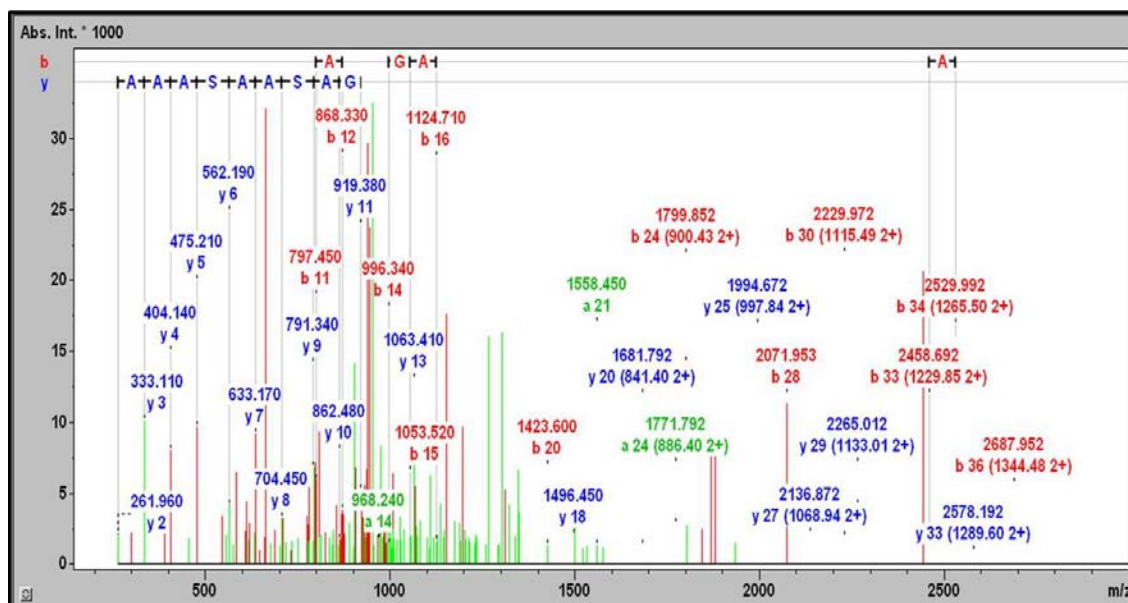


Figura 17. Espectro CID da sequência peptídica GGLGGQGAGAAAAGGAGQGGYGGVGSGASAASAAASR, selecionando-se o íon de m/z 954.72 na forma de $[M + 3H]^3+$, como precursor observado na proteína espidroína-1 (*spot* 56), presente na glândula ampulada maior e identificada na 2-DE da seda.

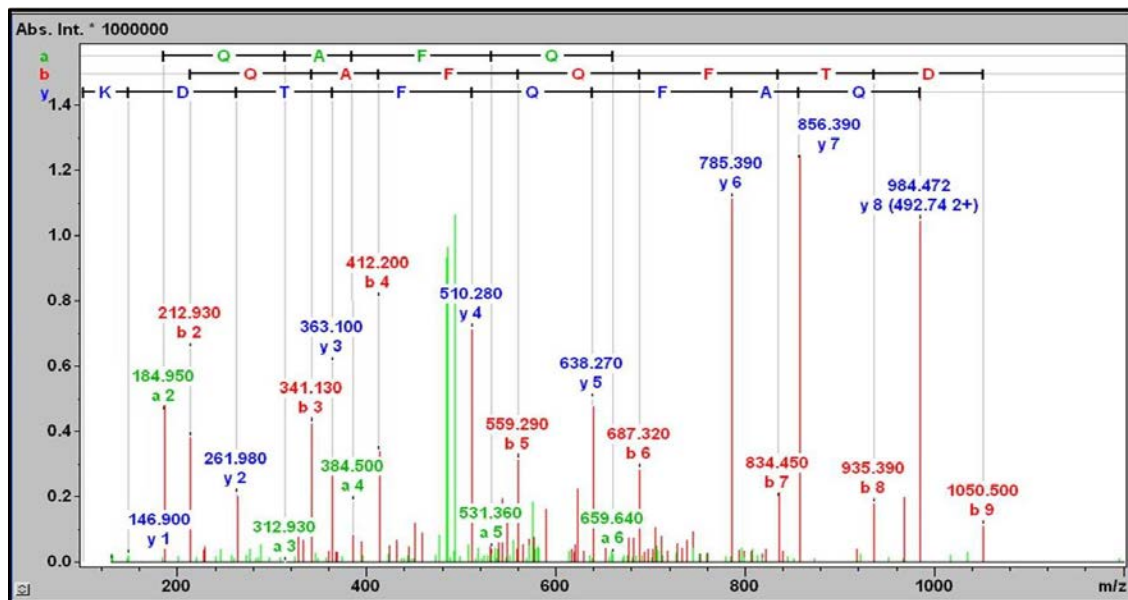


Figura 18. Espectro CID da sequência peptídica LVQAFQFTDK, selecionando-se o íon de m/z 598.76 na forma de $[M + 2H]^2+$, como precursor observado na proteína tioredoxina peroxidase (*spot* 61), presente na glândula ampulada maior e identificada na 2-DE da seda.

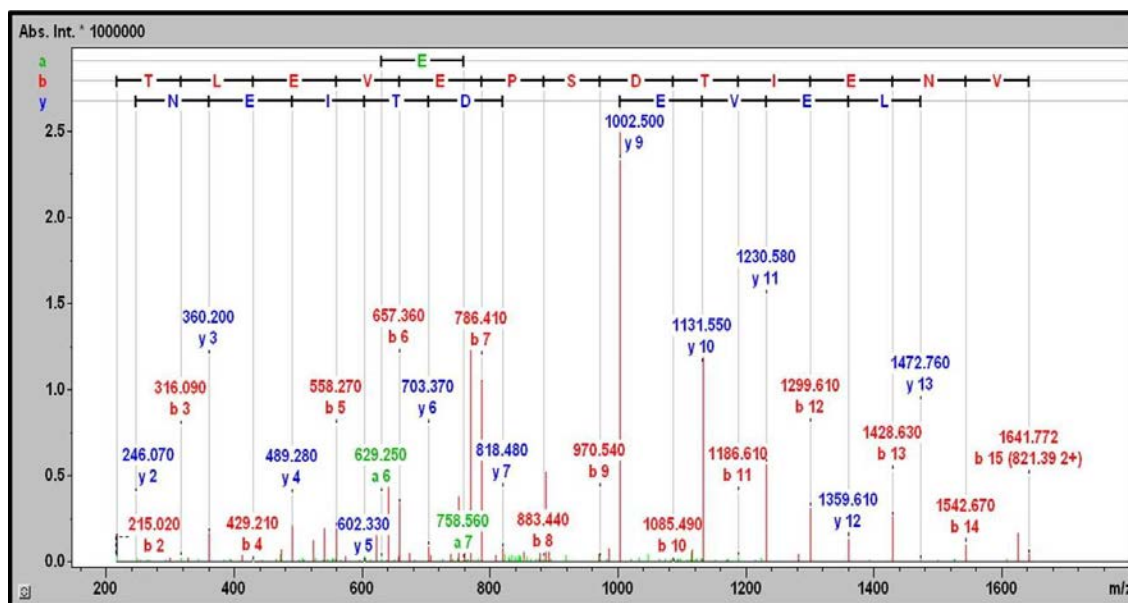


Figura 19. Espectro CID da sequência peptídica TITLVEPSDTEIENVK, selecionando-se o íon de m/z 894.39 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína ubiquitina (*spot* 70), presente na glândula ampulada maior e identificada na 2-DE da seda.

A figura 20 mostra o perfil da eletroforese bidimensional da glândula flageliforme sendo possível visualizar um total de 54 *spots* com MW entre 10 e 80 kDa e *pI* de 3 a 10 (gradiente não linear). Para a identificação por espectrometria de massas, os *spots* mais dominantes foram selecionados, excisados dos géis e submetidos à digestão enzimática com tripsina. A tabela 7 mostra a identificação de 43 proteínas. Uma série de proteínas como proteína de choque térmico (*spots* 1 a 4), hemocianina (*spots* 5 a 16; 19), tubulina (*spot* 17; 23 e 24), calreticulina (*spot* 22), enolase (*spots* 25 a 28), 2-fosfo-D-glicerato hidrolase (*spots* 29 e 30), actina (*spots* 31 a 36), peptidil-proil cis-trans isomerase (*spots* 42 e 48), dihidropteridina redutase (*spot* 43), triosefosfato isomerase (*spot* 45), proteína sarcoplasmática ligante de cálcio (*spot* 46), superóxido dismutase (*spot* 47), proteína da seda flageliforme (*spot* 49), nucleosídeo difosfato quinase (*spots* 50 e 52) e ubiquitina (*spot* 54) foram identificadas.

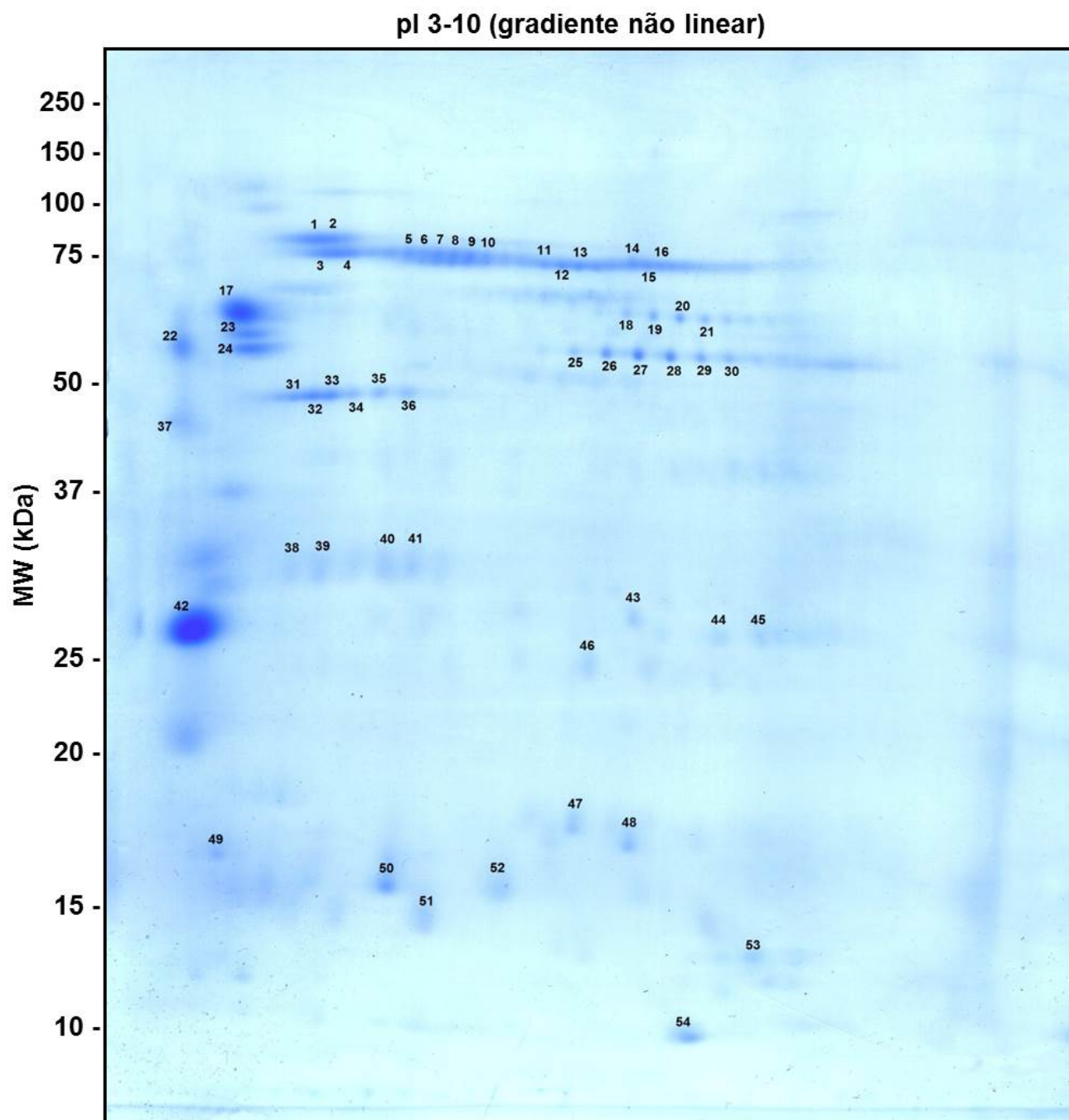


Figura 20. Perfil eletroforético da 2-DE da glândula flageliforme da aranha *N. clavipes*.

Tabela 7. Proteínas identificadas na glândula flageliforme da aranha *N. clavipes* através da digestão *in-gel* dos spots oriundos da 2-DE.

Spot	Código de acesso	Taxonomia	Proteína	Mascot Score	% Cobertura	Exp. / Calc. MW (kDa)	Sequência peptídica MS/MS (Ion Score)
1	E4W3Z2	<i>Haemaphysalis longicornis</i>	Heat shock protein	522	8	89460 / 72544	ITPSYVAFTEAGER (84); NQLTSNPENTVFDK (104); IINEPTAAAIAYGLDK (94); AKFEELNMDLFR (86); FEELNMDLFR (80)
2	F0J9B6	<i>Amblyomma variegatum</i>	Heat shock protein	197	13	89103 / 23953	ITPSYVAFTEAGER (81); NQLTSNPENTVFDK (116)
3	B7PAR6	<i>Ixodes scapularis</i>	Heat shock protein	422	8	83329 / 71292	TVTNAVVTVPAYFNDSQR (91); DAGTIAGLNVLR (90); IINEPTAAAIAYGLDK (98); ARFEELNADLFR (78); FEELNADLFR (65)
4	F0J8P3	<i>Amblyomma variegatum</i>	Heat shock protein	472	12	83329 / 56269	HWPFEVISDGGKPK (62); DAGTIAGLNVLR (71); IINEPTAAAIAYGLDK (122); MVNHFVQEFK (50); ARFEELNADLFR (46); FEELNADLFR (65)
5	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	592	21	83176 / 71709	VCSLFTHLTSISR (102); LAQGELFSCFHEK (45); DFEDFINLAK (69); GVTVPPIQEIFPDR (61); FVPSETISLALK (44); KGELFYMHQQMCAR (47); FNENPGVMSDTSLSRDPFYR (50); AQLDFPGVQVVNVTVNAK (111); TLCIELDKFQVELAAGK (64)
6	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	477	23	82569 / 71709	LAQGELFSCFHEK (46); DFEDFINLAK (77); SFANEGLFVYAASVAILHR (32); GVTVPPIQEIFPDR (50); FVPSETISLALK (45); KGELFYMHQQMCAR (33); AQLDFPGVQVVNVTVNAK (63); TLCIELDKFQVELAAGK (76); QLLAGEGVSENTTEFCSCGWPEHMLIPR (58)
7	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	575	25	82569 / 71709	LHGIGK (27); LAQGELFSCFHEK (79); DFEDFINLAK (77); SFANEGLFVYAASVAILHR (30); GVTVPPIQEIFPDR (52); FVPSETISLALK (48); EVTNHPDKDIEVEIESTGNILDPEYK (36); FNENPGVMSDTSLSRDPFYR (41); AQLDFPGVQVVNVTVNAK (112); TLCIELDKFQVELAAGK (80)
8	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	555	21	81529 / 71709	VCSLFTHLTSISR (64); DFEDFINLAK (81); SFANEGLFVYAASVAILHR (27); GVTVPPIQEIFPDR (57); FVPSETISLALK (39); FNENPGVMSDTSLSRDPFYR (45); AQLDFPGVQVVNVTVNAK (116); VPNLVTTFMK (25); TLCIELDKFQVELAAGK (106)
9	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	626	23	82120 / 71709	VCSLFTHLTSISR (80); LAQGELFSCFHEK (56); DFEDFINLAK (81); SFANEGLFVYAASVAILHR (37); GVTVPPIQEIFPDR (68); FVPSETISLALK (32); FNENPGVMSDTSLSRDPFYR (49); AQLDFPGVQVVNVTVNAK (110); VPNLVTTFMK (54); TLCIELDKFQVELAAGK (63)
10	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	744	31	81824 / 71709	VCSLFTHLTSISR (89); LAQGELFSCFHEK (44); DFEDFINLAK (74); SFANEGLFVYAASVAILHR (44); GVTVPPIQEIFPDR (44); FVPSETISLALK (31); KGELFYMHQQMCAR (44); LTEDNGIDILGSIVESSYESK (56); FNENPGVMSDTSLSR (66); AQLDFPGVQVVNVTVNAK (125); TLCIELDKFQVELAAGK (66); QLLAGEGVSENTTEFCSCGWPEHMLIPR (63)
11	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	377	17	79666 / 71709	DFEDFINLAK (62); SFANEGLFVYAASVAILHR (38); GVTVPPIQEIFPDR (22); FVPSETISLALK (42); FNENPGVMSDTSLSRDPFYR (64); AQLDFPGVQVVNVTVNAK (84); TLCIELDKFQVELAAGK (68)
12	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	672	27	79247 / 71786	QDQILPLFEK (46); LIGVGVLP (55); AADFADFLNLAK (113); DVVNEGLFVYALSVAVVHR (15); GVTLPPIQEVFPDR (75); FIPAETINLASK (14); KGELFYMHQQMCAR (68); GFYGSLSHNWGHVMMAR (56); FQENPGVMSDTSLSR (93); FVDNIFQQYK (61); YHLDHEPFSITVNAQNSSGAPK (30); LSPDEQRPLFIELDK (46)
13	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	672	26	79109 / 71786	LIGVGVLP (55); AADFADFLNLAK (95); DVVNEGLFVYALSVAVVHR (39); GVTLPPIQEVFPDR (61); FIPAETINLASK (31); LEGYAPHLTSLVSGLHYASR (47); GFYGSLSHNWGHVMMAR (46); FQENPGVMSDTSLSR (88); FVDNIFQQYK (73); YHLDHEPFSITVNAQNSSGAPK (81); SPDEQRPLFIELDK (55)
14	Q86N90	<i>Nephila inaurata</i>	Hemocyanin subunit F	497	19	80373 / 72326	ILPLFEHLASLTR (71); DIVNEGMFVYSTSVALLHR (31); GVTVPPIQEIFPDR (50);

		<i>madagascariensis</i>					KGELFYMHQQMCAR (76); ENPGVMSDSTSLRDPIFYR (17); FIDNIFQEYK (74); DLDFFGVSVVNVTVNAK (124); FFIELDKFHAEAPGK (59)
15	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	493	20	79387 / 72326	ILPLFEHLASLTR (72); DIVNEGFMFVYSTSVALLHR (38); GVTVPPIQEIPDR (40); KGELFYMHQQMCAR (66); PAGLSLHDLNELVDVQDMAR (60); YKENPGVMSDSTSLRDPIFYR (18); ENPGVMSDSTSLRDPIFYR (35); FIDNIFQEYK (70); DLDFFGVSVVNVTVNAK (96)
16	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	657	21	79247 / 72326	RILPLFEHLASLTR (85); ILPLFEHLASLTR (76); DIVNEGFMFVYSTSVALLHR (36); GVTVPPIQEIPDR (53); KGELFYMHQQMCAR (74); ENPGVMSDSTSLRDPIFYR (51); FIDNIFQEYK (70); DLDFFGVSVVNVTVNAK (120); LPNIVNTYLK (56); LPNIVNTYLKEDQLEVSHGVSLK (44)
17	P52273	<i>Bombyx mori</i>	Tubulin alpha chain	548	31	66262 / 50558	TIGGGDDSFNTFFSETGAGK (103); LIGQIVSSITASLR (122); IHFPLVTYAPVISAER (72); AYHEQLSVAEITNACFEPANQMVK (62); TIQFVDWCPTGFK (68); VGINYQPPTVVPGLDLAK (41); AVCMLSNTTAAEAWAR (32); AFVHWYVGEEMEEGEFSEAR (50)
19	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	69	4	65506 / 72326	ILPLFEHLASLTR (52); GVTVPPIQEIPDR (19)
22	Q64K80	<i>Ixodes ovatus</i>	Calreticulin	136	11	58571 / 48044	FYGLSTK (30); TMVVQFTVK (26); KVHVIFNYK (42); CKDDVYTHLYTLVVKPDNTYVVK (39)
23	E7D198	<i>Latrodectus hesperus</i>	Beta-tubulin	674	43	61407 / 33681	INVYYNEATGGK (24); AILLDLEPGTMDSVR (97); GHYTEGAELVDVTVR (120); IMNTFSVMPSPK (48); LTTPTYGDLNLHLSVTMSGVTTCLR (42); FPGQLNADLR (38); LAVNMVPPFR (76); LHFFMPGFAPLTSR (75); ALTPVELTQQMFDK (104)
24	Q8T8B2	<i>Bombyx mori</i>	Beta-tubulin	251	15	58118 / 50638	SGPFGQIFRPDNFVFGQSGAGNNWAK (62); LHFFMPGFAPLTSR (54); NSSYFVEWIPNNVK (51); MAATFIGNSTAIQELFK (86)
25	D6X017	<i>Tribolium castaneum</i>	Enolase	194	11	57316 / 48448	LGANAILGVSLAVAK (82); DIILPVPAFNVINGGSHAGNK (36); LAMQEFMILPTGAK (78)
26	D6X017	<i>Tribolium castaneum</i>	Enolase	259	13	57053 / 48448	LGANAILGVSLAVAK (88); DIILPVPAFNVINGGSHAGNK (58); LAMQEFMILPTGAK (78); DGQYDLDFK (37)
27	D6X009	<i>Tribolium castaneum</i>	Enolase	203	8	56878 / 47124	LGANAILGVSLAVAK (87); LAMQEFMILPTGAK (74); DGQYDLDFK (42)
28	D6X009	<i>Tribolium castaneum</i>	Enolase	229	8	56965 / 47124	LGANAILGVSLAVAK (96); LAMQEFMILPTGAK (93); DGQYDLDFK (39)
29	Q6QWQ0	<i>Phormictopus sp.</i>	2-phospho-D-glycerate hydrolase	190	7	56273 / 40026	KFGLAATGVGDEGGFAPDIQENK (35); FGLAATGVGDEGGFAPDIQENK (132); YDLDFK (24)
30	Q6QWQ0	<i>Phormictopus sp.</i>	2-phospho-D-glycerate hydrolase	164	9	56359 / 40026	FGLAATGVGDEGGFAPDIQENK (138); AGYTGK (5); YDLDFK (22)
31	B2XY36	<i>Latrodectus hesperus</i>	Actin	564	38	49517 / 34974	AGFAGDDAPR (72); AVFPSIVGR (49); IWHHTFYNELR (70); VAPEEHPVLLTEAPLNPK (27); GYSFTTTAER (56); LCYVALDFEQEMATAASSSSLEK (55); SYELPDGQVITIGNER (72); DLYANTVLSGGTTMYPGIADR (85)
32	B3VN78	<i>Bombyx mandarina</i>	Actin	614	31	49517 / 42174	AVFPSIVGR (52); IWHHTFYNELR (78); TTGIVLDSGDGVSHTVPIYEGYALPHAILR (73); DLTDLMLK (43); GYSFTTTAER (61); SYELPDGQVITIGNER (84); KDLYANTVLSGGTTMYPGIADR (63); DLYANTVLSGGTTMYPGIADR (96); EITALAPSTMK (45)
33	P04829	<i>Bombyx mori</i>	Actin	580	22	49517 / 42207	AVFPSIVGR (57); IWHHTFYNELR (75); GYSFTTTAER (59); SYELPDGQVITIGNER (98); KDLYANTVLSGGTTMYPGIADR (69); DLYANTVLSGGTTMYPGIADR (119); QEYDESGPSIVHR (60)
34	B3VN78	<i>Bombyx mandarina</i>	Actin	539	32	49861 / 42174	AVFPSIVGR (42); AVFPSIVGRPR (34); HQGVMVGMGQKDLVYGDEAQSK (18); IWHHTFYNELR (76); VAPEEHPVLLTEAPLNPK (30); GYSFTTTAER (55); SYELPDGQVITIGNER (73); KDLYANTVLSGGTTMYPGIADR (67); DLYANTVLSGGTTMYPGIADR (97); EITALAPSTMK (47)
35	A9QQ62	<i>Lycosa</i>	Actin	487	37	50140 / 30525	IWHHTFYNELR (83); VAPEEHPVLLTEAPLNPK (51); DLTDLMLK (45);

		<i>singoriensis</i>					EKLCYVALDFEQEMATAASSSSLEK (37); LCYVALDFEQEMATAASSSSLEK (50); SYELPDGQVITIGNER (95); KDLYANTVLSGGTTMYPGIADR (58); DLYANTVLSGGTTMYPGIADR (71)
36	P84183	<i>Bombyx mori</i>	Actin	644	47	50140 / 42194	AVFPSIVGR (46); IWHHTFYNELR (48); VAPEEHPVLLTEAPLNPK (54); EKLCYVALDFEQEMATAASSSSLEK (38); LCYVALDFEQEMATAASSSSLEK (59); SYELPDGQVITIGNER (88); CPEALFQPSFLGMEACGIHETTYNSIMK (22); KDLYANTVLSGGTTMYPGIADR (67); DLYANTVLSGGTTMYPGIADR (88); EITALAPSTMK (47); YSVWIGGSILASLSTFQQMWISK (3); QEYDESGPSIVHR (66)
42	D3TSE2	<i>Glossina morsitans morsitans</i>	FKBP-type peptidyl-prolyl cis-trans isomerase	65	12	28023 / 23528	NGFISHDEFSGPK (35); VDVISVPEVCEQK (30)
43	E2A057	<i>Camponotus floridanus</i>	Dihydropteridine reductase	71	10	28933 / 25941	AALEETPGMIGYGMK (31); ILAQIENTLK (40)
45	F0J8Y1	<i>Amblyomma variegatum</i>	Triosephosphate isomerase	74	13	27556 / 24625	VVIAYEPVWAIGTGK (30); GAFTGEISPAMIK (44)
46	P86909	<i>Chionoecetes opilio</i>	Sarcoplasmic calcium-binding protein	56	100	25652 / 842	VATVSLPR (56)
47	B0X9L3	<i>Culex quinquefasciatus</i>	Superoxide dismutase	92	15	17402 / 15662	QISLSGPLSILGR (42); EHGAPDASVR (50)
48	A1KYY3	<i>Suidasia medanensis</i>	Peptidyl-prolyl cis-trans isomerase	194	14	17031 / 17893	TSWLDGK (28); HVVFGQVVEGMDVVK (109); HVVFGQVVEGMDVVK (57)
49	O44358	<i>Nephila clavipes</i>	Flagelliform silk protein	150	3	16720 / 70997	AMQVALASSIAELVIAESSGGDVQR (102); TNVISNALR (47)
50	Q0PXX9	<i>Diaphorina citri</i>	Nucleoside diphosphate kinase	126	13	15988 / 17151	TFLMIKPDGVQR (66); GDLCIQVGR (60)
52	E7D144	<i>Latrodectus hesperus</i>	Nucleoside diphosphate kinase	136	21	15897 / 17258	TFIMIKPDGVQR (68); NIIHGSDSVPSADKEIALWFK (53); EIALWFK (15)
54	P68197	<i>Ceratitis capitata</i>	Ubiquitin	239	67	10000 / 8560	MQIFVK (27); TLTGKTITLEVEPSDTIENVKAKIQDK (3); TITLEVEPSDTIENVK (102); TLDYNIQK (58); ESTLHLVLR (49)

Considerando a grande quantidade de espectros que foram gerados na identificação das sequências peptídicas de todas as proteínas oriundas da 2-DE da glândula flageliforme, como exemplos serão mostrados somente alguns espectros CID de algumas sequências peptídicas que foram identificadas nas proteínas tubulina (*spot* 23; Figuras 21 e 22), superóxido dismutase (*spot* 47; Figura 23) e proteína da seda flageliforme (*spot* 49; Figuras 24 e 25).

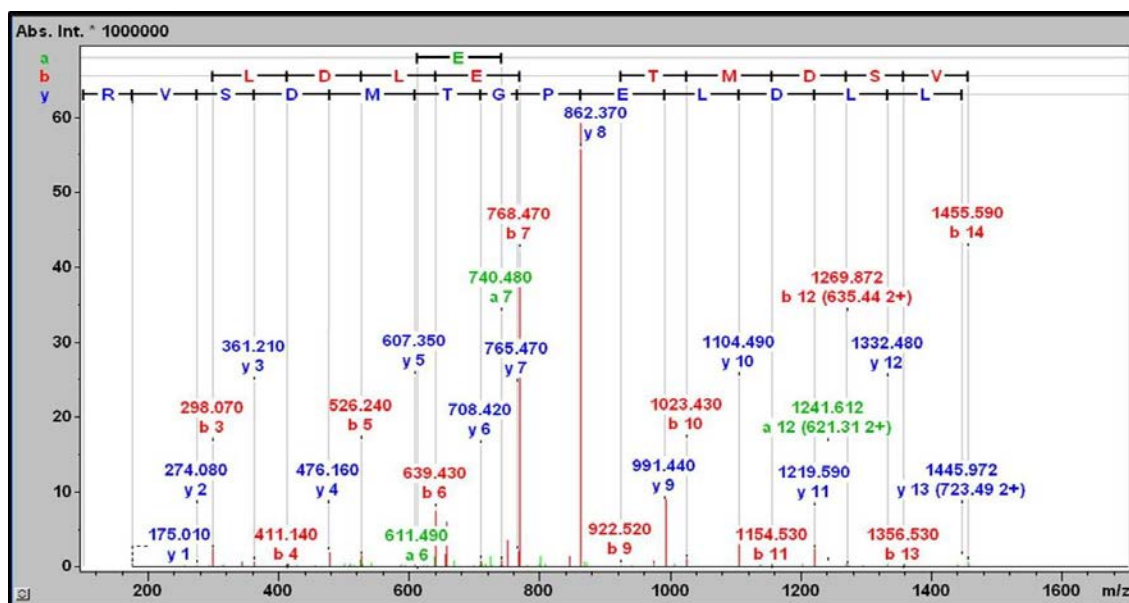


Figura 21. Espectro CID da sequência peptídica AILLDLEPGTMDSVR, selecionando-se o íon de m/z 815.38 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína tubulina (*spot* 23), presente na glândula flageliforme e identificada na 2-DE da seda.

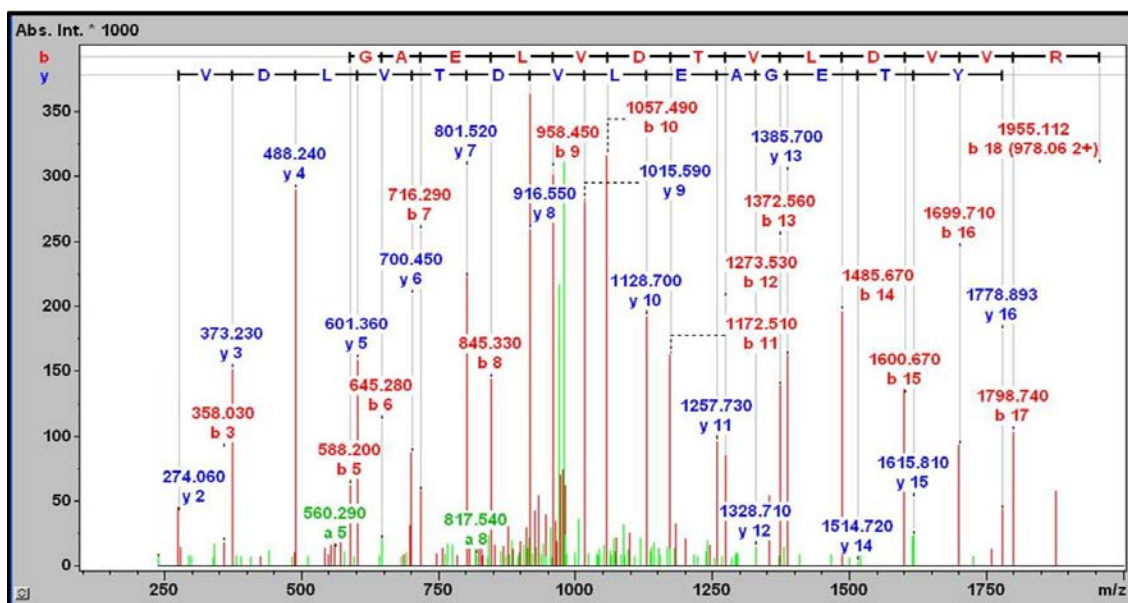


Figura 22. Espectro CID da sequência peptídica GHYTEGAELVDVTVR, selecionando-se o íon de m/z 986.94 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína tubulina (*spot* 23), presente na glândula flageliforme e identificada na 2-DE da seda.

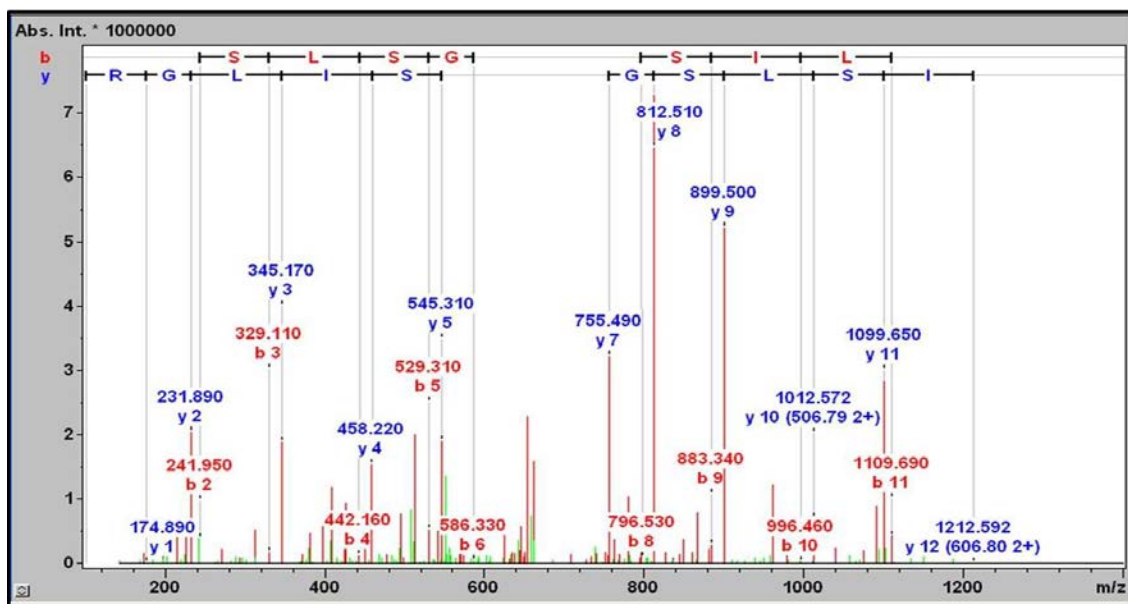


Figura 23. Espectro CID da sequência peptídica QISLSGPLSILGR, selecionando-se o íon de m/z 670.83 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína superóxido dismutase (*spot* 47), presente na glândula flageliforme e identificada na 2-DE da seda.

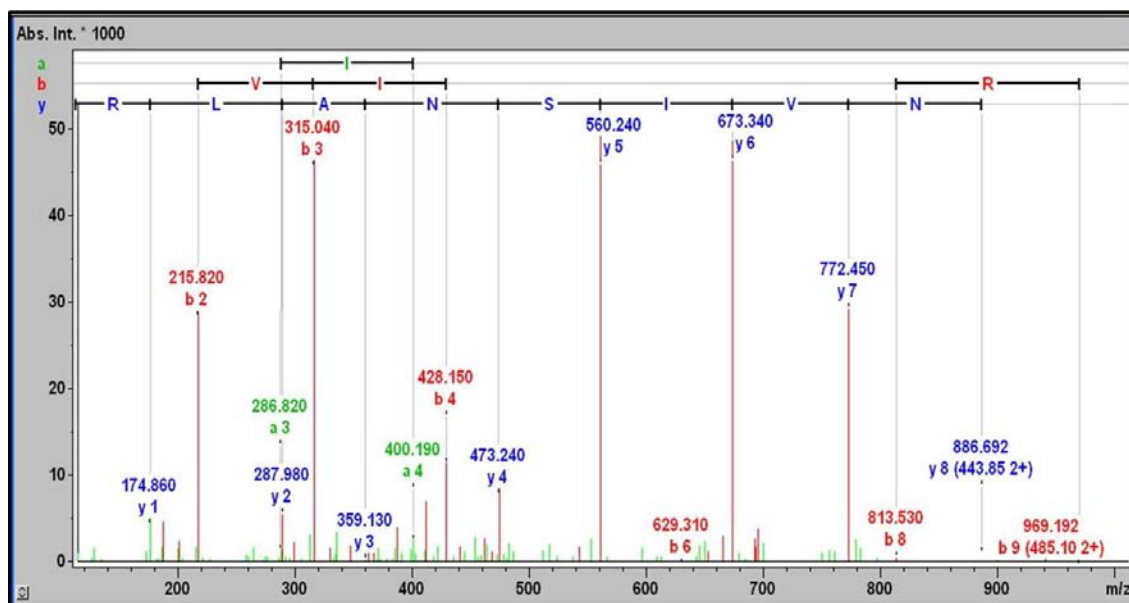


Figura 24. Espectro CID da sequência peptídica TNVISNALR, selecionando-se o íon de m/z 494.26 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína da seda flageliforme (spot 49), presente na glândula flageliforme e identificada na 2-DE da seda.

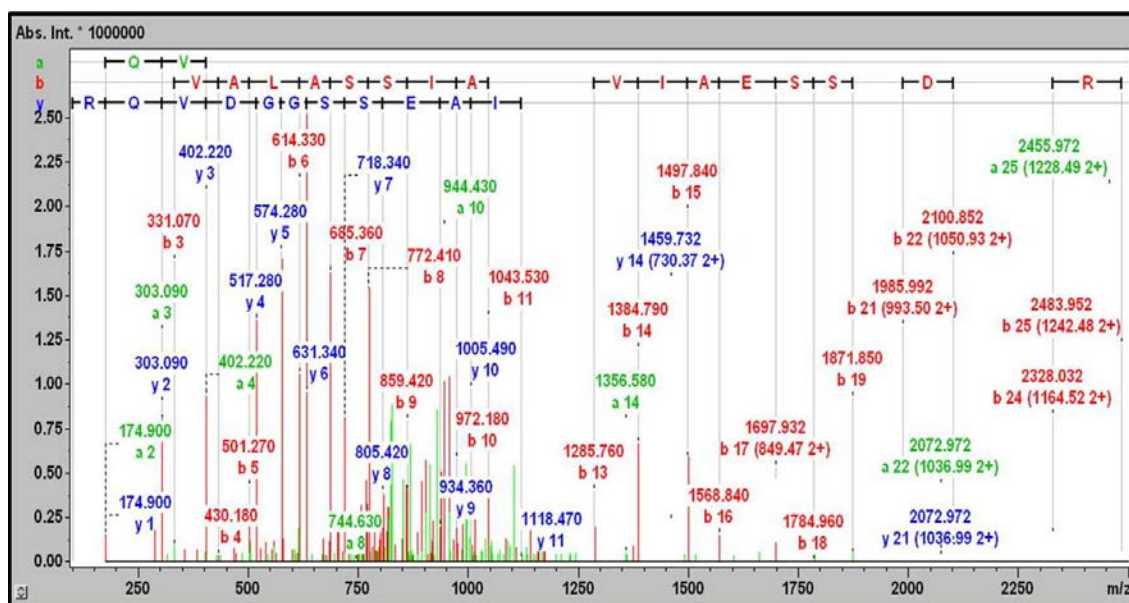


Figura 25. Espectro CID da sequência peptídica AMQVALASSIAELVIAESSGGDVQR, selecionando-se o íon de m/z 834.78 na forma de $[M + 3H]^{3+}$, como precursor observado na proteína da seda flageliforme (spot 49), presente na glândula flageliforme e identificada na 2-DE da seda.

A figura 26 mostra o perfil da eletroforese bidimensional da glândula tubuliforme sendo possível visualizar um total de 66 *spots* com MW entre 10 e 80 kDa e pI de 3 a 10 (gradiente não linear). Para a identificação por espectrometria de massas, os *spots* mais dominantes foram selecionados, excisados dos géis e submetidos à digestão enzimática com tripsina. A tabela 8 mostra a identificação de 47 proteínas. Uma série de proteínas como proteína de choque térmico (*spots* 1 a 7; 21 a 22), hemocianina (*spots* 8 a 16; 18), proteína dissulfeto isomerase (23), aminopeptidase (*spots* 25; 27 a 30), 2-fosfo-D-glicerato hidrolase (*spots* 32, 33, 35, 37 enolase (*spots* 34, 36 e 38), gliceraldeído 3-fosfato desidrogenase (*spots* 39 a 42), tiol peroxiredoxina (*spot* 49), troponina T (*spot* 50), superóxido dismutase (*spot* 52), proteína da seda tubuliforme (*spots* 54, 62 e 63), nucleosídeo difosfato quinase (*spots* 55 e 56), chaperonina (*spot* 61) e ubiquitina (*spots* 65 e 66) foram identificadas.

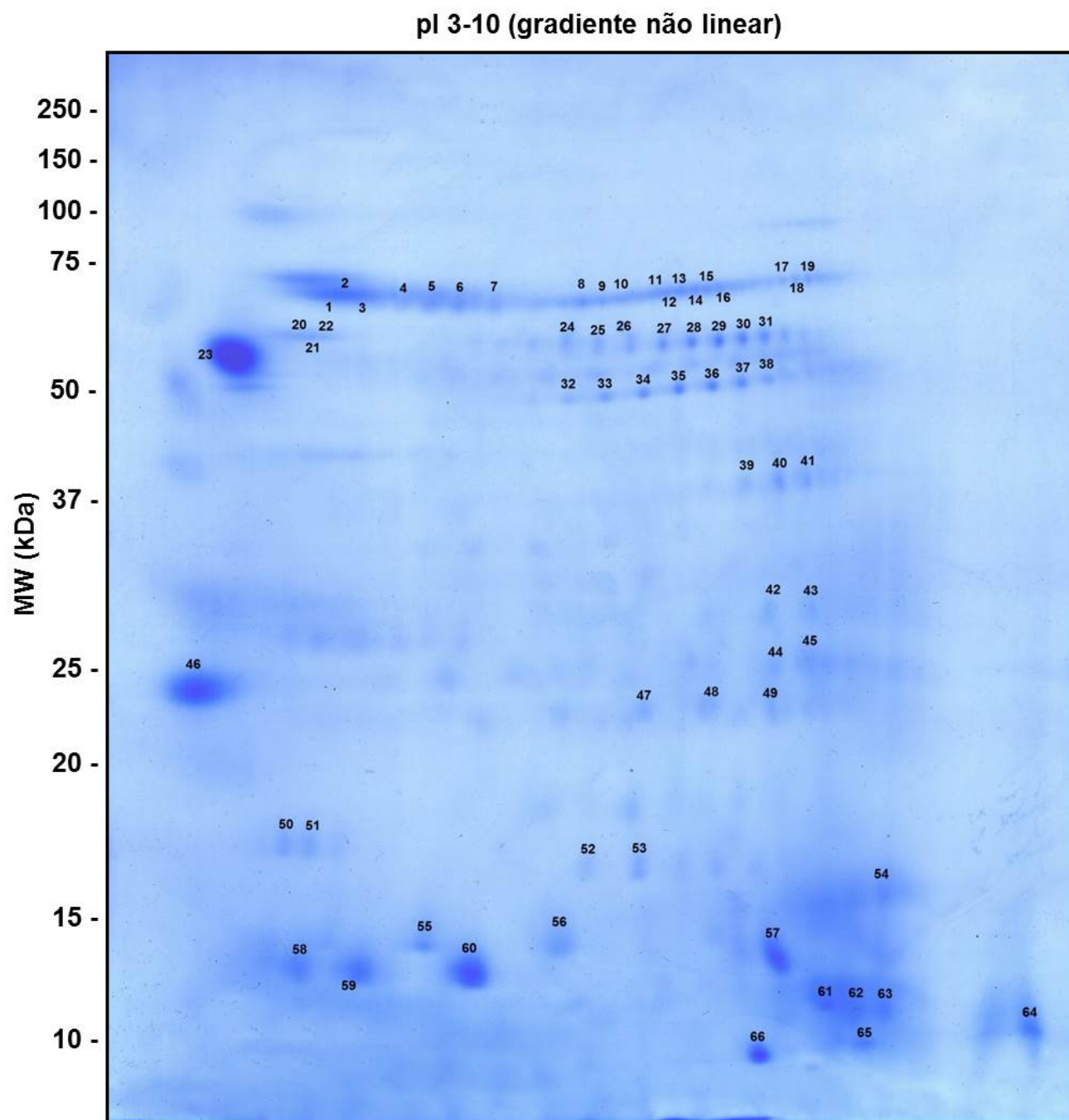


Figura 26. Perfil eletroforético da 2-DE da glândula tubuliforme da aranha *N. clavipes*.

Tabela 8. Proteínas identificadas na glândula tubuliforme da aranha *N. clavipes* através da digestão *in-gel* dos spots oriundos da 2-DE.

Spot	Código de acesso	Taxonomia	Proteína	Mascot Score	% Cobertura	Exp. / Cal. MW (kDa)	Sequência peptídica MS/MS (Ion Score)
1	B7PAR6	<i>Ixodes scapularis</i>	Heat shock protein	456	16	67446 / 71292	VEIANDQGNR (54); EIAEAYLGK (39); TVTNAVVTVPAYFNDSQR (39); IINEPTAAAIAYGLDK (103); STAGDTHLGGEDFDNR (64); ARFEELNADLFR (56); FEELNADLFR (73); SINPDEAVAYGAAVQAAILIGDK (30)
2	Q17310	<i>Ceratitis capitata</i>	Heat shock protein	560	15	67446 / 71435	VEIANDQGNR (48); TTPSYVAFETER (95); HWPFEVVSADGKPK (22); DAGTIAGLNVLR (80); IINEPTAAAIAYGLDK (106); IINEPTAAAIAYGLDKK (81); FEELNADLFR (67); QTQTFTTYSNQPGLVLIQVYEGE (69)
3	B7PAR6	<i>Ixodes scapularis</i>	Heat shock protein	536	12	67791 / 71292	MKEIAEAYLGK (63); EIAEAYLGK (35); TVTNAVVTVPAYFNDSQR (71); DAGTIAGLNVLR (95); IINEPTAAAIAYGLDK (118); MVNHVQEFK (58); ARFEELNADLFR (38); FEELNADLFR (61)
4	Q17310	<i>Ceratitis capitata</i>	Heat shock protein	495	11	66322 / 71435	VEIANDQGNR (69); TTPSYVAFETER (102); IINEPTAAAIAYGLDK (102); IINEPTAAAIAYGLDKK (79); ARFEELNADLFR (73); QTQTFTTYSNQPGLVLIQVYEGE (78)
5	B7PAR6	<i>Ixodes scapularis</i>	Heat shock protein	600	14	66544 / 71292	MKEIAEAYLGK (32); EIAEAYLGK (39); TVTNAVVTVPAYFNDSQR (107); DAGTIAGLNVLR (72); IINEPTAAAIAYGLDK (118); STAGDTHLGGEDFDNR (50); MVNHVQEFK (77); ARFEELNADLFR (57); FEELNADLFR (52)
6	B7PAR6	<i>Ixodes scapularis</i>	Heat shock protein	584	14	65993 / 71292	MKEIAEAYLGK (57); EIAEAYLGK (26); TVTNAVVTVPAYFNDSQR (101); DAGTIAGLNVLR (94); IINEPTAAAIAYGLDK (110); STAGDTHLGGEDFDNR (82); MVNHVQEFK (50); ARFEELNADLFR (67)
7	Q9U639	<i>Manduca sexta</i>	Heat shock protein	411	13	65884 / 71672	VEIANDQGNR (60); TFPPEEVSSMVLTK (32); IINEPTAAAIAYGLDK (128); MVNHVQEFK (43); ARFEELNADLFR (73); FEELNADLFR (70); STMEDEKLK (1); CNDTIK (5)
8	P80476	<i>Androctonus australis</i>	Hemocyanin AA6 chain	86	5	65884 / 72197	GITVPPIQEVFPDR (33); FQENPGVMSDSTSLRDPFYR (55)
9	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	701	25	65993 / 71786	AADFADFLNLAK (113); GVTLPPIQEVFPDR (65); FIPAETINLASK (40); KGELFYMHQQMCAR (24); MVAFHNFEK (55); DLTEVDVQDMER (102); ILEAIDLK (54); GFYGSLSHNWGHVMMAR (53); FQENPGVMSDSTSLRDPFYR (43); FVDNIFQQYK (61); LSPDEQRPLFIELDK (33); TICNDAVSYYCGAK (48)
10	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	532	19	65993 / 71786	LIGVGVLPDR (55); AADFADFLNLAK (109); GVTLPPIQEVFPDR (50); FIPAETINLASK (42); KGELFYMHQQMCAR (78); FVDNIFQQYK (61); YHLDHEPFSITVNAQNSSGAPK (57); LSPDEQRPLFIELDK (61); QLAPGNVISR (20)
11	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	436	18	67791 / 71786	LIGVGVLPDR (55); AADFADFLNLAK (114); GVTLPPIQEVFPDR (53); FIPAETINLASK (36); GFYGSLSHNWGHVMMAR (13); FQENPGVMSDSTSLRDPFYR (34); FVDNIFQQYK (79); LSPDEQRPLFIELDKFKH (55)
12	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	438	19	68024 / 72326	RILPLFEHLASLTR (63); DIVNEGMFVYSTSVALLHR (29); GVTVPPIQEIFPDR (42); YKENPGVMSDSTSLRDPFYR (19); FIDNIFQEYK (70); DLDFPGSVSVNVTNNAK (101); EDQLEVSHGVSLK (53); SICADAVSYCGAK (66)
13	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	743	24	68731 / 72326	RILPLFEHLASLTR (65); ILPLFEHLASLTR (79); DIVNEGMFVYSTSVALLHR (23); GVTVPPIQEIFPDR (52); KGELFYMHQQMCAR (83); YKENPGVMSDSTSLR (14); YKENPGVMSDSTSLRDPFYR (18); FIDNIFQEYK (74); DLDFPGSVSVNVTNNAK (122); LPNIVNTYLKEDQLEVSHGVSLK (80); LSLNEAR (35); TLAEFSSANMK (80)
14	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	617	25	68731 / 72326	LPLFEHLASLTR (73); DIVNEGMFVYSTSVALLHR (33); GVTVPPIQEIFPDR (31); FIPSETINQAIK (29); KGELFYMHQQMCAR (69); EYYGNLHNSGHVMMAR (23); YKENPGVMSDSTSLRDPFYR (21); ENPGVMSDSTSLRDPFYR (47); FIDNIFQEYK

							(78); DLDFFPGVSVVNVTVNAK (77); SICADAVSYCGAK (61); TLAEFSSANMK (81)
15	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	755	24	68731 / 72326	RILPLFEHLASLTR (66); DIVNEGMFVYSTSVALLHR (40); GVTVPPIQEIPDR (49); KGEIFYMHQMCAR (73); YKENPGVMSDTSTSLRDPFYR (30); ENPGVMSDTSTSLRDPFYR (66); FIDNIFQEYK (74); DLDFFPGVSVVNVTVNAK (127); LPNIVNTYLK (59); SICADAVSYCGAK (81); SMGFPPDR (30); TLAEFSSANMK (64)
16	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	441	18	69577 / 72326	DIVNEGMFVYSTSVALLHR (39); GVTVPPIQEIPDR (41); PAGLSLHDLNELVDVQDMAR (50); FIDNIFQEYK (70); SMGFPPDR (18); DLDFFPGVSVVNVTVNAK (118); EDQLEVSHGVSLK (41); SICADAVSYCGAK (71)
18	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	194	11	71596 / 72483	FIPAETINR (14); EVDVQDMER (30); VPNIHTYSK (33); LVLEDQR (24); ETEYCSCGWPEHMLVPR (49); AMGYPPDRPIK (16); TQNSFTEVR (27)
21	A3QME2	<i>Lucilia cuprina</i>	Heat shock protein	174	4	58415 / 61575	NVIEQSWGSPK (87); TLDELEVIEGMK (88)
22	A3QME2	<i>Lucilia cuprina</i>	Heat shock protein	173	9	58415 / 61575	NVIEQSWGSPK (69); LVQDVANNTNEEAGDGTATTATVLAR (79); TLDELEVIEGMKFDR (28)
23	F1CJ13	<i>Hottentotta judaica</i>	Protein disulfide isomerase	111	7	54892 / 37275	IFGGDIK (26); NFDEVAFDK (50); INSFTPIK (38)
25	Q17GL0	<i>Aedes aegypti</i>	Leucyl aminopeptidase	60	4	57369 / 56798	GVTYDTGGADIK (55); KEDFDMHSGK (5)
27	Q17GL0	<i>Glossina m. morsitans</i>	Leucyl aminopeptidase	70	4	57448 / 56798	GVTYDTGGADIK (25); KEDFDMHSGK (45)
28	B0WS64	<i>Culex quinquefasciatus</i>	Leucyl aminopeptidase	84	5	57686 / 56101	GVTYDTGGADIK (44); VEQYLNDGVFDK (40)
29	Q17GL0	<i>Aedes aegypti</i>	Leucyl aminopeptidase	88	5	57846 / 56798	GVTYDTGGADIK (46); NSVGEECYVSEMITSR (42)
30	B0WS64	<i>Culex quinquefasciatus</i>	Leucyl aminopeptidase	86	5	58333 / 56101	GVTYDTGGADIK (38); VEQYLNDGVFDK (48)
32	Q6QWP9	<i>Limulus polyphemus</i>	2-phospho-D-glycerate hydrolase	156	7	48892 / 39993	IEIGMDVAASEFHK (85); VNQIGSVTEAIK (71)
33	Q6QWP9	<i>Limulus polyphemus</i>	2-phospho-D-glycerate hydrolase	157	7	49198 / 39993	IEIGMDVAASEFHK (78); VNQIGSVTEAIK (79)
34	D2WJ58	<i>Atteva punctella</i>	Enolase	239	11	49814 / 40720	LGANAILGVSLAVAK (97); DIILPVPAPFNVINGGSHAGNK (68); DGKYDLDFK (54); YDLDFK (22)
35	Q6QWQ0	<i>Phormictopus sp.</i>	2-phospho-D-glycerate hydrolase	249	7	50312 / 40026	KFGLAATGVGDEGGFAPDIQENK (73); FGLAATGVGDEGGFAPDIQENK (154); YDLDFK (23)
36	A1YQ87	<i>Bombyx mori</i>	Enolase	194	6	50940 / 47164	LGANAILGVSLAVAK (99); DGKYDLDFK (52); YDLDFK (23); YNQILR (23)
37	Q6QWQ0	<i>Phormictopus sp.</i>	2-phospho-D-glycerate hydrolase	205	7	51322 / 40026	KFGLAATGVGDEGGFAPDIQENK (60); FGLAATGVGDEGGFAPDIQENK (120); YDLDFK (28)
38	Q967P1	<i>Cryphalus abietis</i>	Enolase	180	19	51643 / 15461	LGANAILGVSLAVAK (97); LAMQEFMILPTGAK (83)
39	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	297	24	38906 / 33251	VGINGFGR (23); IHVFQEMKPSSIPWGK (76); VGEAYVVESTGVFTTLEK (100); VINDNFGIVEGLMTTVHAITATQK (24); AGISLNNNFVK (76)
40	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	369	29	38997 / 33251	YMVYMFK (31); IHVFQEMKPSSIPWGK (38); VGEAYVVESTGVFTTLEK (81); VPTPDVSVWDLTCR (52); GILGYTEDEVSSDFIGDSHSSIFDAK (97); AGISLNNNFVK (72)
41	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	205	14	39366 / 33251	IHFVFQEMKPSSIPWGK (39); VGEAYVVESTGVFTTLEK (79); AGISLNNNFVK (87)
42	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	158	22	29442 / 33251	VGEAYVVESTGVFTTLEK (85); VINDNFGIVEGLMTTVHAITATQK (18); GILGYTEDEVSSDFIGDSHSSIFDAK (55)
49	Q6T3A7	<i>Bombyx mori</i>	Thiol peroxiredoxin	72	10	23435 / 22073	LVQAFQFTDK (37); ATAVVNGEFK (35)
50	A8E4K0	<i>Haemaphysalis</i>	Troponin T	60	6	16755 / 46442	EGEEMGGGSALGSSGFDK (46); TFTDKR (17)

		<i>qinghaiensis</i>					
52	B0X9L3	<i>Culex quinquefasciatus</i>	Superoxide dismutase	93	18	16272 / 15662	QISLSGPLSILGR (42); GTIYFEQNADSDAVK (51)
54	Q3BCG2	<i>Nephila clavipes</i>	Tubuliform spidroin	135	5	15576 / 56145	SLGIADAAGLAGVLAR (60); VSSLAQSFASALSASR (75)
55	E7D144	<i>Latrodectus hesperus</i>	Nucleoside diphosphate kinase	180	21	14009 / 17258	TFIMIKPDGVQR (51); NIIHGSDSVPSADK (53); NIIHGSDSVPSADKEIALWFK (67)
56	E7D144	<i>Latrodectus hesperus</i>	Nucleoside diphosphate kinase	127	21	14042 / 17258	TFIMIKPDGVQR (60); NIIHGSDSVPSADKEIALWFK (67)
61	Q685Z5	<i>Mesobuthus gibbosus</i>	Chaperonin	176	50	11595 / 6762	GGIMIEPK (47); VQSATVVAVGPGGR (87); VLLPEYGGTK (43)
62	Q3BCG2	<i>Nephila clavipes</i>	Tubuliform spidroin	131	5	11098 / 56145	SLGIADAAGLAGVLAR (59); VSSLAQSFASALSASR (72)
63	Q3BCG2	<i>Nephila clavipes</i>	Tubuliform spidroin	201	5	11098 / 56145	SLGIADAAGLAGALAR (63); SLGIADAAGLAGVLAR (138)
65	Q4PM40	<i>Ixodes scapularis</i>	Ubiquitin	135	21	9416 / 18209	TITLEVEPSDTIENVK (81); TLSDYNIQK (28); ESTLHLVLR (25)
66	Q4LAY8	<i>Eucinetus sp.</i>	Ubiquitin	390	27	8949 / 18107	TITLEVEASDTIENVK (103); EGIPPDQQR (17); TLSDYNIQK (57); TLSDYNIQKESTLHLVLR (44); ESTLHLVLR (65)

Considerando a grande quantidade de espectros que foram gerados na identificação das sequências peptídicas de todas as proteínas oriundas da 2-DE da glândula tubuliforme, como exemplos serão mostrados somente alguns espectros CID de algumas sequências peptídicas que foram identificadas nas proteínas de choque térmico (*spot* 1; Figura 27) e proteína da seda tubuliforme (*spot* 54; Figura 28/ *spot* 63; figura 29).

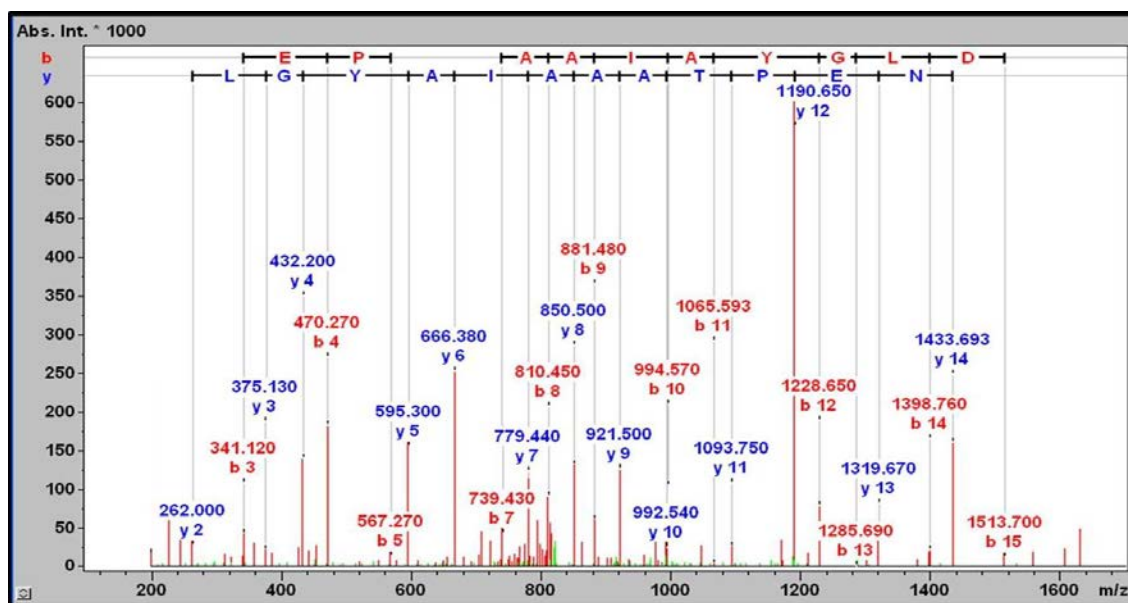


Figura 27. Espectro CID da sequência peptídica IINEPTAAAIYGLDK, selecionando-se o íon de m/z 830.43 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína de choque térmico (*spot* 1), presente na glândula tubuliforme e identificada na 2-DE da seda.

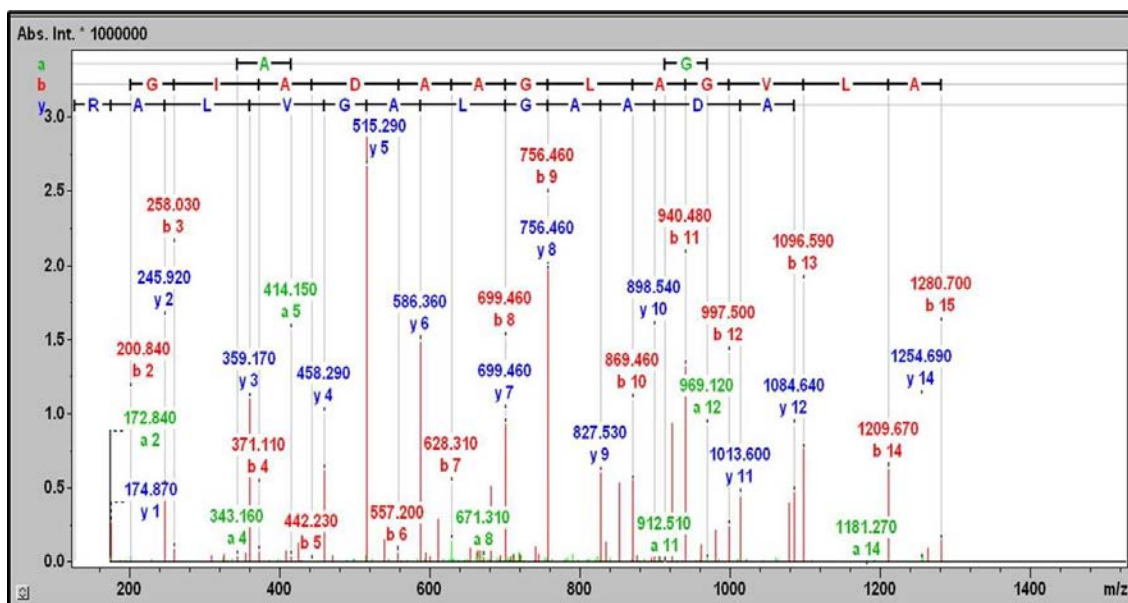


Figura 28. Espectro CID da sequência peptídica SLGIADAAGLAGVLAR, selecionando-se o íon de m/z 727.97 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína da seda tubuliforme (*spot* 54), presente na glândula tubuliforme e identificada na 2-DE da seda.

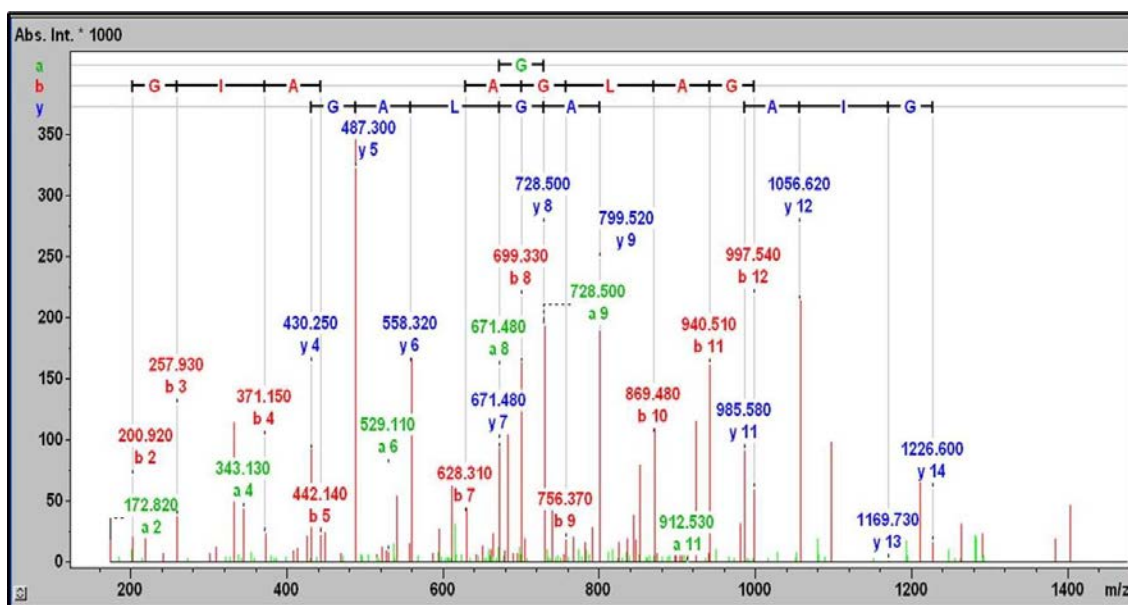


Figura 29. Espectro CID da sequência peptídica SLGIADAAGLAGALAR, selecionando-se o íon de m/z 713.94 na forma de $[M + 2H]^{2+}$, como precursor observado na proteína da seda tubuliforme (*spot* 63), presente na glândula tubuliforme e identificada na 2-DE da seda.

A figura 30 mostra o perfil da eletroforese bidimensional da glândula ampulada menor sendo possível visualizar um total de 80 *spots* com MW entre 10 e 80 kDa e pI de 3 a 10 (gradiente não linear). Para a identificação por espectrometria de massas, os *spots* mais dominantes foram selecionados, excisados dos géis e submetidos à digestão enzimática com tripsina. A tabela 9 mostra a identificação de 52 proteínas. Uma série de proteínas como hemocianina (*spots* 4 a 11; 14; 17 a 30; 32 a 35), aldeído desidrogenase (*spots* 37 e 38), enolase (*spots* 39, 40, 42 a 45), 2-fosfo-D-glicerato hidrolase (*spot* 41), actina (*spots* 49 a 53), isocitrato desidrogenase (*spots* 54 a 56), gliceraldeído 3-fosfato desidrogenase (*spots* 57 a 59; 61), dihidropteridina redutase (*spot* 65), superóxido dismutase (*spot* 71), peptidil-prolil cis-trans isomerase (*spot* 72) e ubiquitina (*spot* 80) foram identificadas.

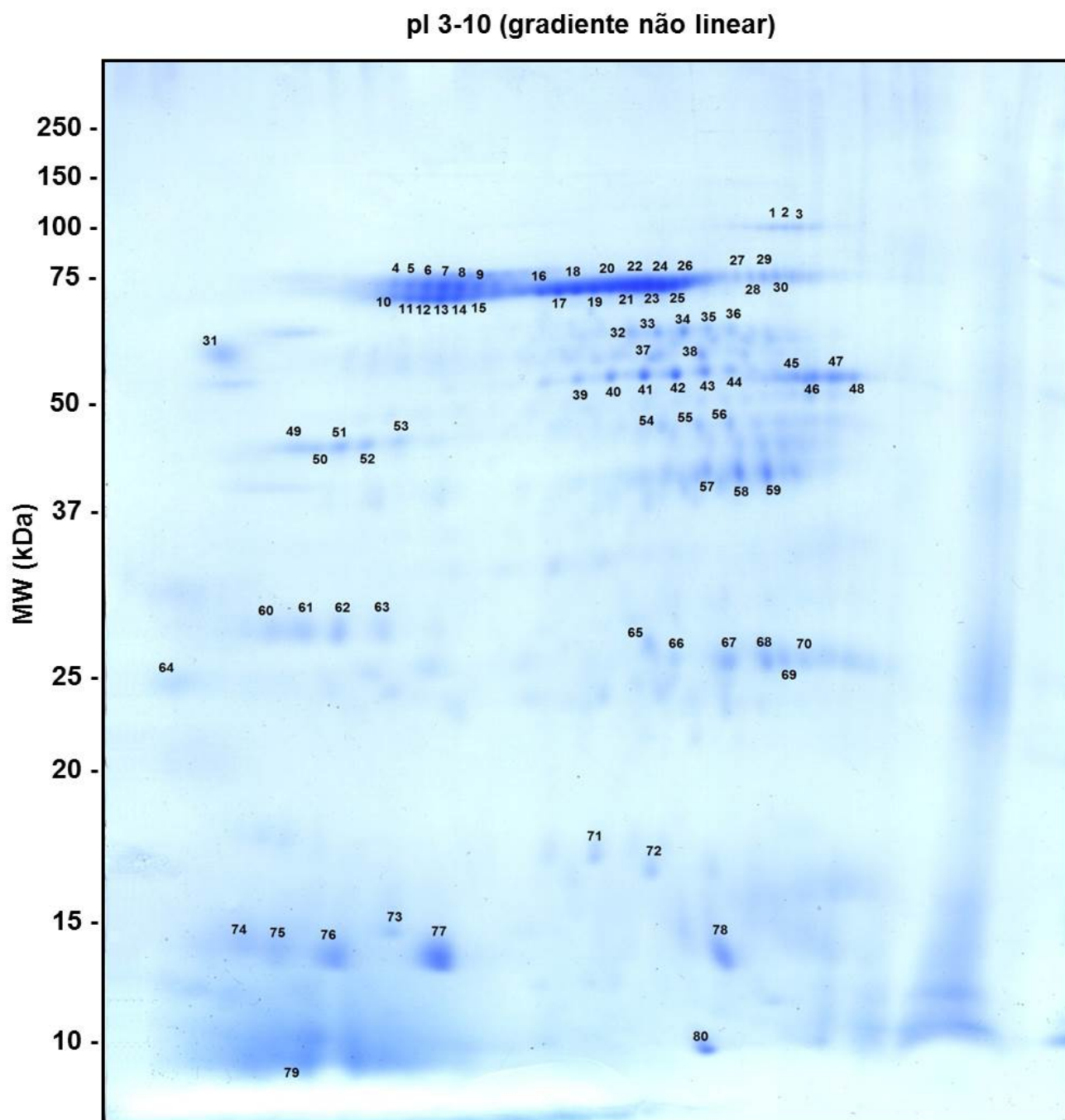


Figura 30. Perfil eletroforético da 2-DE da glândula ampulada menor da aranha *N. clavipes*.

Tabela 9. Proteínas identificadas na glândula ampulada menor da aranha *N. clavipes* através da digestão *in-gel* dos spots oriundos da 2-DE.

Spot	Código de acesso	Taxonomia	Proteína	Mascot Score	% Cobertura	Exp. / Calc. MW (kDa)	Sequência peptídica MS/MS (Ion Score)
4	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	809	28	80248 / 71709	VCSLFTHLTSISR (65); LYETLYAAK (58); DFEDFINLAK (83); GVTVPPIQEIPDR (60); KGELFYMHQQMCAR (30); LFYGSLSHNWGHVMMANITDPDGR (46); FNENPGVMSDTSTSLR (73); AQLDFPGVQVNVTVNAK (151); TLCIELDKFQVELAAGK (81); FQVELAAGK (55); LSSVTVSEHTFK (27); QLLAGEGVSENTTEFCSCGWPEHMLIPR (82)
5	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	493	15	79925 / 71709	LYETLYAAK (78); DFEDFINLAK (83); GVTVPPIQEIPDR (63); MSYFR (11); KGELFYMHQQMCAR (46); FNENPGVMSDTSTSLR (115); YDELGNR (27); LSSVTVSEHTFK (30); KPMGFPPDR (45)
6	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	706	20	79285 / 71709	VCSLFTHLTSISR (84); LYETLYAAK (70); DFEDFINLAK (83); GVTVPPIQEIPDR (69); EVTNHPDKDIEVEIESTGNILDPEYK (35); FNENPGVMSDTSTSLR (116); YDELGNR (29); TLCIELDKFQVELAAGK (119); FQVELAAGK (45); LSSVTVSEHTFK (62)
7	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	770	22	79285 / 71709	VCSLFTHLTSISR (83); LHGIK (28); LYETLYAAK (67); DFEDFINLAK (83); SFANGLFVYAASVAILHR (45); GVTVPPIQEIPDR (67); FVPSETISLAK (37); KGELFYMHQQMCAR (60); FNENPGVMSDTSTSLR (122); VPNLVTTFMK (34); TLCIELDKFQVELAAGK (102); FQVELAAGK (50)
8	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	946	34	79925 / 71709	VCSLFTHLTSISR (65); LAQGEFSCFHEK (46); LYETLYAAK (76); DFEDFINLAK (83); GVTVPPIQEIPDR (39); FVPSETISLAK (20); DIEVEIESTGNILDPEYK (13); KGELFYMHQQMCAR (47); LFYGSLSHNWGHVMMANITDPDGR (40); FNENPGVMSDTSTSLR (103); AQLDFPGVQVNVTVNAK (143); VPNLVTTFMK (45); TLCIELDKFQVELAAGK (111); FQVELAAGK (46); QLLAGEGVSENTTEFCSCGWPEHMLIPR (73)
9	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	842	29	79604 / 71709	VCSLFTHLTSISR (86); LAQGEFSCFHEK (68); LYETLYAAK (69); DFEDFINLAK (88); FVPSETISLAK (23); EVTNHPDKDIEVEIESTGNILDPEYK (60); MSYFR (25); KGELFYMHQQMCAR (48); LTEDNGIDILGSIVESSYESK (19); FNENPGVMSDTSTSLR (129); DPIFYR (26); AQLDFPGVQVNVTVNAK (95); FQVELAAGK (56); LSSVTVSEHTFK (59)
10	Q86N94	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit A	446	17	75887 / 72628	GIIVPPIQEIFADR (54); PAGDESDVIIDVQETGNILDPEYK (36); TGELFYMHQQMCAR (43); HYAPRPDGLAMTDLR (23); LVDVQDMQR (63); FIDNVFQEYK (72); AMGYPPDR (23); TITAQSHEEFITPNMK (79)
11	Q86N94	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit A	337	13	75291 / 72628	EDIGVNAHWHVHVPVSVYDSK (45); TGELFYMHQQMCAR (54); HYAPRPDGLAMTDLR (43); LVDVQDMQR (69); DPIFYR (24); FIDNVFQEYK (75)
14	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	52	4	75589 / 71709	DFEDFINLAK (23); SFANGLFVYAASVAILHR (32)
17	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	302	18	78180 / 71786	AADFADFLNLAK (56); MVAFHNFEEK (32); DLTEVDVQDMER (41); ILEAIDLK (34); FQENPGVMSDTSTSLR (38); FQENPGVMSDTSTSLRDPIFYR (31); FVDNIFQQYK (42); YHHLDEPFISITVNAQNSSGAPK (30)
18	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	220	12	78180 / 71786	AADFADFLNLAK (65); FQENPGVMSDTSTSLRDPIFYR (35); FVDNIFQQYK (45); YHHLDEPFISITVNAQNSSGAPK (50); QLAPGNVISR (28)
19	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	341	13	78967 / 71786	AADFADFLNLAK (113); DVVNEGLFVYALSVAVVHR (30); GVTLPPIQEVFPDR (55); FIPAETINLASK (32); KGELFYMHQQMCAR (36); DLTEVDVQDMER (76);
20	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	258	9	79285 / 72326	DIVNEGFMFVYSTSVALLHR (52); GVTVPPIQEIPDR (26); FIDNIFQEYK (74); DLDFPGSVVNVTVNAK (108)
21	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	589	20	79444 / 72326	RILPLFEHLASLTR (52); ILPLFEHLASLTR (41); GVTVPPIQEIPDR (35); KGELFYMHQQMCAR (61); ENPGVMSDTSTSLRDPIFYR (82); FIDNIFQEYK (70);

							DLDFPGVSVVNVTVNAK (109); LPNIVNTYLK (42); RFFIELDKFHAELAPGK (96)
22	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	555	21	79444 / 72326	ILPLFEHLASLTR (43); DIVNEGFMVYSTSVALLHR (36); GVTVPPIQEIPDR (47); FIPSETINQAIK (36); ENPGVMSDTSTSLRDPFYR (39); FIDNIFQEYK (74); DLDFPGVSVVNVTVNAK (134); LPNIVNTYLK (68); LPNIVNTYLKEDQLEVSHGVSLK (53); SMGFPPDR (32)
23	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	558	21	79765 / 72326	RILPLFEHLASLTR (66); ILPLFEHLASLTR (79); GVTVPPIQEIPDR (45); FIPSETINQAIK (49); GELFYMHQQMCAR (24); ENPGVMSDTSTSLRDPFYR (14); FIDNIFQEYK (74); DLDFPGVSVVNVTVNAK (129); LPNIVNTYLK (52); EDQLEVSHGVSLK (18); SMGFPPDR (21)
24	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	814	28	79444 / 72326	RILPLFEHLASLTR (66); ILPLFEHLASLTR (76); DFEDFMR (31); DIVNEGFMVYSTSVALLHR (16); GVTVPPIQEIPDR (47); KGELFYMHQQMCAR (29); YKENPGVMSDTSTSLR (11); YKENPGVMSDTSTSLRDPFYR (24); ENPGVMSDTSTSLRDPFYR (49); FIDNIFQEYK (70); DLDFPGVSVVNVTVNAK (130); LPNIVNTYLK (61); LPNIVNTYLKEDQLEVSHGVSLK (63); FFIELDKFHAELAPGK (39); SMGFPPDR (27); TLAEFSSANMK (89)
25	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	581	24	79765 / 72326	DIVNEGFMVYSTSVALLHR (37); GVTVPPIQEIPDR (59); FIPSETINQAIK (37); GELFYMHQQMCAR (26); YKENPGVMSDTSTSLRDPFYR (21); FIDNIFQEYK (70); DLDFPGVSVVNVTVNAK (151); LPNIVNTYLK (65); LPNIVNTYLKEDQLEVSHGVSLK (50); FFIELDKFHAELAPGK (46); SMGFPPDR (25)
26	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	689	23	80410 / 72326	RILPLFEHLASLTR (53); DFEDFMR (45); GVTVPPIQEIPDR (49); FIPSETINQAIK (46); KGELFYMHQQMCAR (70); GELFYMHQQMCAR (23); YKENPGVMSDTSTSLRDPFYR (29); ENPGVMSDTSTSLRDPFYR (35); FIDNIFQEYK (74); DLDFPGVSVVNVTVNAK (144); LPNIVNTYLK (61); FFIELDKFHAELAPGK (51); SMGFPPDR (16)
27	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	311	14	83052 / 72483	DFGDFINLAK (57); GVILPPIQEVFPDKFIPAETINR (18); TGNIDPEYNLAYFR (74); KGELFYMHQQMCAR (50); WMDNIFQEYK (54); ETEYCSCGWPEHMLVPR (58)
28	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	474	22	82884 / 72483	DFGDFINLAK (74); GVILPPIQEVFPDK (51); GVILPPIQEVFPDKFIPAETINR (40); TGNIDPEYNLAYFR (65); ILSAIHTGQVIDSNGQEVPLDLER (75); ENPGVMDTSTALRDPFYR (17); WMDNIFQEYK (69); SVDSSVTLSHVPTFEELEK (40); ETEYCSCGWPEHMLVPR (45)
29	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	293	13	83052 / 72483	DFGDFINLAK (74); TGNIDPEYNLAYFR (73); LEGYSAHLISLISGLNYASR (46); ENPGVMDTSTALRDPFYR (29); ETEYCSCGWPEHMLVPR (72)
30	Q86N92	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit D	276	15	83388 / 72483	DFGDFINLAK (64); GVILPPIQEVFPDKFIPAETINR (22); TGNIDPEYNLAYFR (49); KGELFYMHQQMCAR (22); WMDNIFQEYK (70); SVDSSVTLSHVPTFEELEK (34); SVDSSVTLSHVPTFEELEKGENIPNR (16)
32	Q9NFL5	<i>Aphonopelma californicum</i>	Hemocyanin F chain	63	7	65817 / 72674	DIVNEGLFVYSVSVAILHR (9); GVTVPPIQEIPDR (43); FVPAETVNQAVK (13)
33	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	252	14	65703 / 72326	ILPLFEHLASLTR (20); DIVNEGFMVYSTSVALLHR (33); GVTVPPIQEIPDR (44); KGELFYMHQQMCAR (30); FIDNIFQEYK (73); DLDFPGVSVVNVTVNAK (55)
34	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	141	10	66275 / 72326	DIVNEGFMVYSTSVALLHR (23); GVTVPPIQEIPDR (33); KGELFYMHQQMCAR (32); DLDFPGVSVVNVTVNAK (53)
35	Q86N90	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit F	136	9	66740 / 72326	DIVNEGFMVYSTSVALLHR (40); GVTVPPIQEIPDR (22); YKENPGVMSDTSTSLR (12); FIDNIFQEYK (66)
37	B7QAL5	<i>Ixodes scapularis</i>	Aldehyde dehydrogenase	122	3	60501 / 54721	IAKEEIFGPVQQIMK (49); EEIFGPVQQIMK (73)
38	B7QAL5	<i>Ixodes scapularis</i>	Aldehyde dehydrogenase	117	3	60885 / 54721	IAKEEIFGPVQQIMK (58); EEIFGPVQQIMK (59)
39	Q967P1	<i>Cryphalus abietis</i>	Enolase	183	19	55611 / 15461	LGANAILGVSLAVAK (92); LAMQEFMILPTGAK (91)
40	A1YQ87	<i>Bombyx mori</i>	Enolase	210	7	56096 / 47164	DNIKSEYHGK (63); SKLGANAILGVSLAVAK (50); LGANAILGVSLAVAK (116)
41	Q6QWQ0	<i>Phormictopus sp.</i>	2-phospho-D-glycerate	174	7	56424 / 40026	FGLAATGVGDEGGFAPDIQENK (143); YDLDFK (31)

			hydrolase				
42	Q967P1	<i>Cryphalus abietis</i>	Enolase	189	19	56755 / 15461	LGANAILGVSLAVAK (102); LAMQEFMILPTGAK (87)
43	Q967P1	<i>Cryphalus abietis</i>	Enolase	198	22	57431 / 15461	LGANAILGVSLAVAK (108); AGAAK (36); LAMQEFMILPTGAK (84)
44	Q967P1	<i>Cryphalus abietis</i>	Enolase	173	19	57346 / 15461	LGANAILGVSLAVAK (95); LAMQEFMILPTGAK (78)
45	Q967P1	<i>Cryphalus abietis</i>	Enolase	96	22	55933 / 15461	LGANAILGVSLAVAK (38); AGAAK (26); LAMQEFMILPTGAK (32)
49	P84183	<i>Bombyx mori</i>	Actin	488	31	45548 / 42194	AVFPSIVGR (50); AVFPSIVGRPR (17); DLTDYLMK (42); LCYVALDFEQEMATAASSSSLEK (79); SYELPDGQVITIGNER (92); CPEALFQPSFLGMEACGIHETTYNSIMK (27); KDLYANTVLSGGTTMYPGIADR (48); DLYANTVLSGGTTMYPGIADR (84); EITALAPSTMK (51)
50	B3VN78	<i>Bombyx mori</i>	Actin	414	26	45190 / 42174	AVFPSIVGR (57); HQGVMVGMGQKDLVVGDEAQS (15); DLTDYLMK (37); LCYVALDFEQEMATAASSSSLEK (79); SYELPDGQVITIGNER (82); KDLYANTVLSGGTTMYPGIADR (57); DLYANTVLSGGTTMYPGIADR (90)
51	Q5D579	<i>Ixodes ricinus</i>	Actin	436	18	45369 / 41821	AVFPSIVGR (57); GYSFTTTAER (43); SYELPDGQVITIGNER (100); KDLYANTVLSGGTTMYPGIADR (70); DLYANTVLSGGTTMYPGIADR (119); EITALAPSTMK (47)
52	B7PBG2	<i>Ixodes scapularis</i>	Actin	426	33	45608 / 37854	IWHHTFYNELR (46); LCYVALDFEQEMATAASSSSLEK (81); SYELPDGQVITIGNER (80); CPEALFQPSFLGMEACGIHETTYNSIMK (24); KDLYANTVLSGGTTMYPGIADR (84); DLYANTVLSGGTTMYPGIADR (80); EITALAPSTMK (33)
53	B2XY36	<i>Latrodectus hesperus</i>	Actin	479	29	46153 / 34974	AVFPSIVGR (42); IWHHTFYNELR (40); DLTDYLMK (30); EKLCYVALDFEQEMATAASSSSLEK (24); LCYVALDFEQEMATAASSSSLEK (99); SYELPDGQVITIGNER (89); KDLYANTVLSGGTTMYPGIADR (77); DLYANTVLSGGTTMYPGIADR (79)
54	Q2F681	<i>Bombyx mori</i>	Isocitrate dehydrogenase	85	2	48416 / 46546	FKDIFQDIYDR (39); DIFQDIYDR (45)
55	C1BS27	<i>Lepeophtheirus salmonis</i>	Isocitrate dehydrogenase cytoplasmic	74	6	48872 / 46885	IIWDLIK (45); NILGGTVFR (19); LIDDMVAQAMK (10)
56	Q2F681	<i>Bombyx mori</i>	Isocitrate dehydrogenase	135	4	48741 / 46546	IIWDLIK (49); FKDIFQDIYDR (39); DIFQDIYDR (48)
57	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	322	21	42410 / 33251	YMVYMFK (17); IHVFQEMKPSSIPWGK (28); VGAEYVVESTGVFTTLEK (119); VPTPDVSVVDLTCR (83); AGISLNNNFVK (76)
58	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	394	28	42410 / 33251	YMVYMFK (28); IHVFQEMKPSSIPWGK (33); VGAEYVVESTGVFTTLEK (74); VINDNFGIVEGLMTTVHAITATQK (94); VPTPDVSVVDLTCR (89); AGISLNNNFVK (76)
59	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	428	37	42302 / 33251	YMVYMFK (31); IHVFQEMKPSSIPWGK (40); VGAEYVVESTGVFTTLEK (71); VINDNFGIVEGLMTTVHAITATQK (52); VPTPDVSVVDLTCR (91); GILGYTEDEVSSDFIGDSHSSIFDAK (56); AGISLNNNFVK (90)
61	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	78	11	29044 / 33251	VGAEYVVESTGVFTTLEK (43); VSNASCTTNCLAPLAK (35)
65	E2A057	<i>Camponotus floridanus</i>	Dihydropteridine reductase	73	14	28029 / 25941	AALLETGPMIGYGMK (41); ADTSTWTPDLFVSNLFWK (32)
71	B0X9L3	<i>Culex quinquefasciatus</i>	Superoxide dismutase	81	15	16781 / 15662	QISLSGPLSILGR (40); EHGAPDASVR (41)
72	A1KYY3	<i>Suidasia medanensis</i>	Peptidyl-prolyl cis-trans isomerase	74	9	16481 / 17893	HVVFQGVVEGMDVVK (51); HVVFQGVVEGMDVVK (23)
80	Q4LAY8	<i>Eucinetus sp.</i>	Ubiquitin	194	10	9635 / 18107	TITLEVEASDTIENVK (89); TITLEVEPSDTIENVK (104)

Considerando a grande quantidade de espectros que foram gerados na identificação das sequências peptídicas de todas as proteínas oriundas da 2-DE das glândulas, e que a grande maioria das proteínas identificadas está presente em todas as glândulas, como exemplos somente foram mostrados anteriormente alguns espectros CID de algumas sequências peptídicas que foram identificadas nas glândulas ampulada maior, flageliforme e tubuliforme.

A figura 31 mostra o perfil da eletroforese bidimensional da glândula agregada sendo possível visualizar um total de 87 *spots* com MW entre 10 e 100 kDa e pI de 3 a 10 (gradiente não linear). Para a identificação por espectrometria de massas, os *spots* mais dominantes foram selecionados, excisados dos géis e submetidos à digestão enzimática com tripsina. A tabela 10 mostra a identificação de 48 proteínas. Uma série de proteínas como hemocianina (*spots* 6 a 16), proteína sarcoplasmática ligante de cálcio (*spots* 24; 26; 28; 30; 36 e 38), proteína de choque térmico (*spots* 25 e 69), proteína de dissulfeto isomerase-2 (*spot* 33), enolase (*spots* 42 a 44), actina (*spots* 46; 48 a 50), isocitrato desidrogenase (*spot* 51), acil-CoA desidrogenase (52), acetil-CoA acetiltransferase (*spot* 53), malato desidrogenase citosólica (*spots* 55; 56 e 59), gliceraldeído 3-fosfato desidrogenase (*spots* 57 e 58), chaperona multifuncional (*spot* 63), triosefosfato isomerase (*spots* 67 e 68), tiorredoxina peroxidase (*spot* 72), fator liberador de histamina (*spot* 74), fator de iniciação da tradução 5A (*spot* 75), superóxido dismutase (*spot* 77), peptidil-prolil cis-trans isomerase (*spot* 78), nucleosídeo difosfato quinase (*spots* 79 e 80), proteína receptora de odor-gustativo (*spot* 82), paramiosina (*spot* 85) e ubiquitina (*spot* 87) foram identificadas.

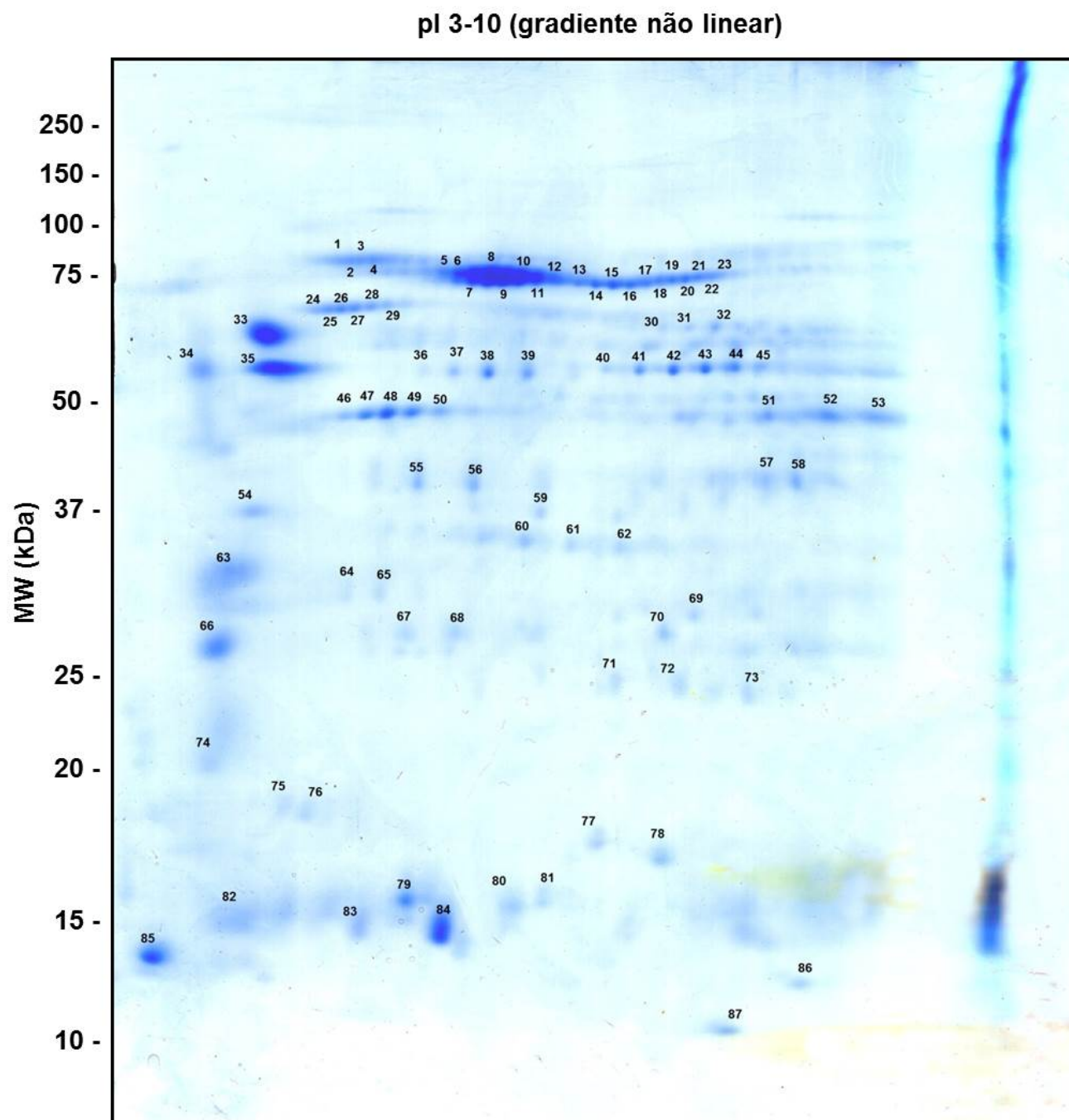


Figura 31. Perfil eletroforético da 2-DE da glândula agregada da aranha *N. clavipes*.

Tabela 10. Proteínas identificadas na glândula agregada da aranha *N. clavipes* através da digestão *in-gel* dos spots oriundos da 2-DE.

Spot	Código de acesso	Taxonomia	Proteína	Mascot Score	% Cobertura	Exp. / Calc. MW (kDa)	Sequência peptídica MS/MS (Ion Score)
6	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	391	14	85106 / 71709	GVTVPPIQEIPDR (65); FVPSETISLALK (55); FNENPGVMSDTSTSLRDPFYR (67); AQLDFPGVQVVNVTNNAK (114); VPNLVTTFMK (30); TLCIELDKFQVELAAGK (62)
7	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	133	5	84566 / 71709	FVPSETISLALK (49); VPNLVTTFMK (32); LSSVTVSEHTFK (54)
8	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	398	18	84208 / 71709	VCSLFTHLTSISR (58); LAQGELFSCFHEK (38); GVTVPPIQEIPDR (55); FVPSETISLALK (52); KGELFYMHQQMCAR (32); FNENPGVMSDTSTSLRDPFYR (44); AQLDFPGVQVVNVTNNAK (76); VPNLVTTFMK (45)
9	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	536	22	84029 / 71709	LAQGELFSCFHEK (45); LYETLYAAK (42); GVTVPPIQEIPDR (69); FVPSETISLALK (51); KGELFYMHQQMCAR (58); FNENPGVMSDTSTSLRDPFYR (39); AQLDFPGVQVVNVTNNAK (120); VPNLVTTFMK (31); YDELGNRLDPEHQR (21); LSSVTVSEHTFK (65)
10	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	754	31	84566 / 71709	VCSLFTHLTSISR (55); DFEDFINLAK (64); SFANEGLFVYAASVAILHR (59); GVTVPPIQEIPDR (59); FVPSETISLALK (57); EVTNHPDKDIEVEIESTGNILDPEYK (45); KGELFYMHQQMCAR (72); LTEDNGIDILGSIVESSYESK (71); FNENPGVMSDTSTSLRDPFYR (66); AQLDFPGVQVVNVTNNAK (100); VPNLVTTFMK (48); TLCIELDKFQVELAAGK (62)
11	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	606	23	84208 / 71709	VCSLFTHLTSISR (66); LAQGELFSCFHEK (61); LYETLYAAK (45); SFANEGLFVYAASVAILHR (61); GVTVPPIQEIPDR (52); FVPSETISLALK (48); KGELFYMHQQMCAR (41); FNENPGVMSDTSTSLRDPFYR (75); AQLDFPGVQVVNVTNNAK (128); VPNLVTTFMK (31)
12	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	433	17	83851 / 71709	VCSLFTHLTSISR (81); SFANEGLFVYAASVAILHR (65); GVTVPPIQEIPDR (25); FVPSETISLALK (45); FNENPGVMSDTSTSLRDPFYR (56); AQLDFPGVQVVNVTNNAK (125); VPNLVTTFMK (38)
13	Q86N91	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit E	279	13	82966 / 71709	VCSLFTHLTSISR (33); GVTVPPIQEIPDR (39); FVPSETISLALK (52); AQLDFPGVQVVNVTNNAK (50); VPNLVTTFMK (34); TLCIELDKFQVELAAGK (72)
14	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	118	7	81568 / 71786	AADFADFLNLAK (24); FQENPGVMSDTSTSLRDPFYR (29); FVDNIFQQYK (65)
15	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	85	3	81050 / 71786	LIGVGVLP (43); FIPAETINLASK (42)
16	Q86N89	<i>Nephila inaurata madagascariensis</i>	Hemocyanin subunit G	71	3	80707 / 71786	GVTLPPIQEVFPDR (45); FVDNIFQQYK (26)
24	P86909	<i>Chionoecetes opilio</i>	Sarcoplasmic calcium-binding protein	89	100	71884 / 842	VATVSLPR (89)
25	A3QME2	<i>Lucilia cuprina</i>	Heat shock protein	139	6	72027 / 61575	LVQDVANNTNEEAGDGTATVTLAR (52); TLTDELEVIEGMK (87)
26	P86909	<i>Chionoecetes opilio</i>	Sarcoplasmic calcium-binding protein	86	100	72171 / 842	VATVSLPR (86)
28	P86909	<i>Chionoecetes opilio</i>	Sarcoplasmic calcium-binding protein	72	100	72896 / 842	VATVSLPR (72)
30	P86909	<i>Chionoecetes opilio</i>	Sarcoplasmic calcium-binding	62	100	66822 / 842	VATVSLPR (62)

			protein				
33	A5LHV9	<i>Haemaphysalis longicornis</i>	Protein disulfide isomerase-2	62	6	64777 / 57270	QLAPIYDELAEK (32); EQVPAMRIIQLEGDMTR (30)
34	P86909	<i>Chionoecetes opilio</i>	Sarcoplasmic calcium-binding protein	50	100	56861 / 842	VATVSLPR (50)
36	P86909	<i>Chionoecetes opilio</i>	Sarcoplasmic calcium-binding protein	55	100	56773 / 842	VATVSLPR (55)
38	P86909	<i>Chionoecetes opilio</i>	Sarcoplasmic calcium-binding protein	54	100	56165 / 842	VATVSLPR (54)
42	O02654	<i>Loligo pealeii</i>	Enolase	106	8	56685 / 47738	EVILPVPAPFNVINGGSHAGNK (11); LAMQEFMILPTGAK (95)
43	Q967P1	<i>Cryphalus abietis</i>	Enolase	133	19	56949 / 15461	LGANAILGVSLAVAK (67); LAMQEFMILPTGAK (67)
44	O02654	<i>Loligo pealeii</i>	Enolase	87	8	57127 / 47738	LGANAILGVSLAVAK (32); LAMQEFMILPTGAK (55)
46	P04829	<i>Bombyx mori</i>	Actin	76	6	49059 / 42207	VAPEEHPVLLTEAPLNPK (51); IIPPER (25)
48	Q6X4V4	<i>Rhipicephalus microplus</i>	Actin	588	39	49325 / 42196	AVFPSIVGR (47); AVFPSIVGRPR (33); IWHHTFYNELR (39); VAPEEHPVLLTEAPLNPK (68); DLTDLMLK (49); LCYVALDFEQEMATAASSSSLEK (63); SYELPDGQVITIGNER (80); CPEALFQPSFLGMESCGIHETTYNSIMK (39); KDLYANTVLSGGTTMYPGIADR (52); DLYANTVLSGGTTMYPGIADR (81); EITALAPSTMK (45)
49	B2XY30	<i>Paraphidippus aurantius</i>	Actin	486	35	49526 / 40121	AVFPSIVGR (45); IWHHTFYNELR (24); VAPEEHPVLLTEAPLNPK (62); DLTDLMLK (32); GYSFTTTAER (53); LCYVALDFEQEMATAASSSSLEK (73); SYELPDGQVITIGNER (78); DLYANTVLSGGTTMYPGIADR (74); EITALAPSTMK (47)
50	Q6X4V4	<i>Rhipicephalus microplus</i>	Actin	596	37	49796 / 42196	AVFPSIVGR (50); IWHHTFYNELR (70); DLTDLMLK (35); GYSFTTTAER (43); EKLCYVALDFEQEMATAASSSSLEK (67); LCYVALDFEQEMATAASSSSLEK (60); SYELPDGQVITIGNER (11); CPEALFQPSFLGMESCGIHETTYNSIMK (40); DLYANTVLSGGTTMYPGIADR (78); EITALAPSTMK (50)
51	Q2F681	<i>Bombyx mori</i>	Isocitrate dehydrogenase	148	4	48861 / 46546	NILGGTVFR (35); FKDIFQDIYDR (67); DIFQDIYDR (46)
52	E2BX37	<i>Harpegnathos saltator</i>	Probable medium-chain specific acyl-CoA dehydrogenase	105	7	48993 / 42519	IFQIYEGTAQIR (55); EINMGQRASDTR (50)
53	B7PDP6	<i>Ixodes scapularis</i>	Acetyl-CoA acetyltransferase	80	8	48861 / 43608	ILANLTYELR (45); VQDLDLVDVNEAFASQFLAVER (35)
55	Q2F5P8	<i>Bombyx mori</i>	Cytosolic malate dehydrogenase	141	10	40555 / 35612	VLVVGNPANTNALICSK (86); WVSMGVSDGSYGTPR (55)
56	E2EJM5	<i>Epicopeia mencia</i>	Cytosolic malate dehydrogenase	121	17	40363 / 14598	VLVVGNPANVNALICSK (30); VLVVGNPANTNALICSK (70); VLVVGNPANCNALICSKYAPSIPK (27)
57	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	322	18	40848 / 33251	VGINGFGR (43); YMVYMFK (24); IHVFQEMKPSSIPWGK (83); VPTPDVSVVDLTCR (111); AGISLNNNFVK (62)
58	E7D199	<i>Latrodectus hesperus</i>	Glyceraldehyde 3-phosphate dehydrogenase	497	40	40701 / 33251	VGINGFGR (44); YMVYMFK (40); IHVFQEMKPSSIPWGK (78); VGAEYVVESTGVFTTLEK (83); VINDNFGIVEGLMTTVHAITATQK (35); VPTPDVSVVDLTCR (92); GILGYTEDEVSSDFIGDSSHIFDAK (54); AGISLNNNFVK (73)
59	E2EIY6	<i>Glyphipterix haworthana</i>	Cytosolic malate dehydrogenase	104	21	37724 / 14747	EQGQAMDRVARK (26); VLVVGNPANTNALICSK (78)
63	B7Q5I4	<i>Ixodes scapularis</i>	Multifunctional chaperone	235	22	32399 / 29599	NLLSVAYK (33); EICQDILDVLDK (36); YLAEFATGNDR (52); AASDIAMTELPPHPIR (57); DSTLMQLLR (59)
67	Q8T5G9	<i>Archaeopotamobius sibiriensis</i>	Triosephosphate isomerase	81	13	28037 / 24754	VVIAYEPVWAIGTGK (47); DCGCEWVILGHSE (34)
68	Q8T5G9	<i>Archaeopotamobius sibiriensis</i>	Triosephosphate isomerase	85	13	27975 / 24754	VVIAYEPVWAIGTGK (55); NVFNEPDTLISEK (30)

69	B0WP93	<i>Culex quinquefasciatus</i>	Heat shock protein	311	9	29293 / 71787	SFFPEEISSMVLTK (85); TVTNAVVTVPAYFNDSQR (70); DAGTIAGLNVLR (90); IINEPTAAAIAYGLDK (70)
72	A6N9S1	<i>Ornithodoros parkeri</i>	Thioredoxin peroxidase	112	10	24339 / 22143	QITVNDLPVGR (48); LVQAFQFTDK (64)
74	E7D139	<i>Latrodectus hesperus</i>	Histamine release factor	101	17	19649 / 19801	DLITGDEMFTDSSK (61); WKEVGMNESEIADAK (40)
75	B5M6E4	<i>Haplopelma schmidtii</i>	Translation initiation factor 5A	129	20	17940 / 17154	VHLVGLDIFTNK (24); KYEDLCPSTHNIDVPNVNR (58); YEDLCPSTHNIDVPNVNR (47)
77	B0X9L3	<i>Culex quinquefasciatus</i>	Superoxide dismutase	68	15	16869 / 15662	QISLSGPLSILGR (38); EHGAPDASVR (30)
78	E0VI73	<i>Pediculus humanus corporis</i>	Peptidyl-prolyl cis-trans isomerase	197	7	16372 / 23488	HVVFVGKVVVEGMDVVK (126); HVVFVGKVVVEGMDVVKK (72)
79	Q0PXX9	<i>Diaphorina citri</i>	Nucleoside diphosphate kinase	131	13	15118 / 17151	TFLMIKPDGVQR (76); GDLCIQVGR (55)
80	Q0PXX9	<i>Diaphorina citri</i>	Nucleoside diphosphate kinase	109	13	15029 / 17151	TFLMIKPDGVQR (53); GDLCIQVGR (55)
82	A7UV31	<i>Anopheles gambiae</i>	Gustatory receptor	89	6	14391 / 46678	QLSSSLNALLK (59); HTFISVPYAPLK (30)
85	Q967Z0	<i>Dermatophagoides farina</i>	Paramyosin	76	3	13057 / 81437	RIHEQEIEIK (44); QYQIEIEQLNMR (32)
87	Q4LAY8	<i>Eucinetus sp.</i>	Ubiquitin	270	19	9982 / 18107	MKIFVK (26); TITLEVEASDTIENVK (86); TITLEVEPSDTIENVK (88); ESTLHLVLR (73)

As aranhas tecedoras de teias orbitais desenvolveram um sistema complexo de glândulas distintas especializadas na secreção de diferentes tipos de fibras de seda que constituem a teia orbital, o frame, a linha de segurança “escape” e o casulo para a proteção dos ovos. Acredita-se que essas glândulas evoluíram de um único tipo de glândula, e subsequentemente divergiram em sua anatomia, conteúdo luminal e morfologia (Vollrath et al., 2001), mantendo reservatórios separados de proteínas de seda individuais conectados às fiandeiras. Esta organização permite com que a aranha utilize cada tipo de seda para uma função específica e também proporciona a produção de múltiplas fibras de seda simultaneamente. Para compreender esse processo é necessário um conhecimento sobre como as proteínas que constituem a seda são secretadas e armazenadas nas glândulas, e toda a mudança física e bioquímica que ocorre até o momento de serem expelidas na forma de uma fibra sólida.

O processo de fiação é, portanto, uma das principais atividades desenvolvidas pela aranha e está envolvido em cada etapa de sua vida. Juntamente com a estrutura primária da proteína, o processo de fiação é um dos determinantes pelas características mecânicas e organização estrutural da fibra, como a cristalinidade, a orientação molecular, estrutura secundária e microestruturas (Lefèvre et al., 2011).

As espidroínas constituintes das fibras da seda são armazenadas nas glândulas até serem processadas em fibras. Durante o processo de fiação, essas proteínas se movem através da glândula encontrando mudanças físicas e bioquímicas em seu ambiente. Essas alterações são acompanhadas por uma separação da fase líquido-líquido, para uma fase líquido-sólido, resultando em uma fibra preliminar. A estrutura final da fibra é atingida após a perda de solvente pelo ar, ao ser liberada pelo ducto de fiação (Sheibel et al., 2009). Portanto, as proteínas identificadas nas glândulas ampulada maior, ampulada menor, flageliforme, tubuliforme e agregada do presente estudo possivelmente estão envolvidas na secreção das proteínas da seda, transporte, regulação de atividades proteolíticas e preservação das proteínas da seda contra o estresse oxidativo e degradação, durante todo o processo de fiação.

Considerando-se que as composições proteômicas de cada glândula estudada são muito semelhantes entre si, sob o ponto de vista de funções proteicas identificadas, neste trabalho estamos propondo um mecanismo geral de ação destas proteínas para a formação das fibras da seda da teia da aranha *N. clavipes*. Para isso, as proteínas identificadas em todos os tipos de glândulas foram organizadas e classificadas em seis diferentes grupos de acordo com a sua função geral: (i) proteínas constituintes das fibras da seda, as espidroínas, (ii) proteínas relacionadas com o dobramento/conformação e modificação das espidroínas, (iii) proteínas

relacionadas com a preservação das espidroínas (*self-protection*) contra o estresse oxidativo, (iv) proteínas relacionadas com a preservação fibrilar das espidroínas no meio ambiente, (v) proteínas de transporte de íons e moléculas relacionadas com a estabilidade das espidroínas, e (vi) proteínas *housekeeping*.

O primeiro grupo é composto basicamente pelas proteínas constituintes das fibras da seda, as espidroínas. Nesse grupo foram identificadas as proteínas espidroína-1 na glândula ampulada maior, proteína da seda flageliforme na glândula flageliforme, e proteína da seda tubuliforme na glândula tubuliforme. Somente não foram identificadas as espidroínas da glândula ampulada menor, possivelmente porque no momento em que essa glândula foi dissecada não estava ocorrendo a síntese das espidroínas e/ou ainda essas proteínas não estavam em quantidades suficientes para serem detectadas na 2-DE. Esse grupo será melhor discutido adiante neste texto.

O segundo grupo incluiu as proteínas relacionadas com o dobramento/conformação e modificação das espidroínas. Nesse grupo as proteínas identificadas foram peptidil-prolil cis-trans isomerase e dissulfeto isomerase, que podem atuar na complementação do dobramento das espidroínas; e também foi identificada uma série de proteínas de choque térmico como proteína de choque térmico 60, chaperonina e calreticulina. Sob condições normais, essas proteínas facilitam o transporte de proteínas entre compartimentos celulares, o dobramento de proteínas recém-sintetizadas e a degradação de proteínas instáveis (Parcellier et al, 2003). As proteínas de choque térmico também foram identificadas nos estudos de Hou et al. (2007) com a glândula de seda de *Bombix mori*. Embora a função das proteínas de choque térmico seja incerta, no presente estudo, é sugerido que desempenham a função no dobramento e transporte das espidroínas durante o processo de formação das fibras no ducto de fiação. A calreticulina é uma proteína ligante de cálcio, com função chaperonina, que pode estar envolvida no processo de maturação das espidroínas através do dobramento dessas proteínas para atingir a sua correta conformação estrutural (Villagomez et al., 2009; Uniprot/GO: Gene Ontology, código de acesso: Q64K80). Outra proteína incluída nesse grupo é a aminopeptidase, enzima que atua como protease catalisando a remoção de aminoácidos da região N-terminal de peptídeos e proteínas (Taylor, 1993; Uniprot/GO: Gene Ontology, código de acesso: D3TLD9).

O terceiro grupo, as proteínas relacionadas com a proteção das espidroínas, incluiu as proteínas tioredoxina peroxidase, superóxido dismutase e tiol peroxiredoxina. São proteínas envolvidas na proteção das espidroínas contra o estresse oxidativo (Kanzok et al., 2001; Apel et al., 2004; Dayer et al., 2008). Provavelmente estão relacionadas com a proteção das espidroínas contra o estresse oxidativo dentro das glândulas, no sentido de preservar intactas

as estruturas das proteínas que constituem as fibras dos diferentes tipos de sedas das teias. Tem sido sugerido que as proteínas peroxidases podem estar envolvidas também na formação de ligações cruzadas covalentes de ditirosina, para fortificar as fibras da seda. A atividade de peroxidases já havia sido detectada na glândula ampulada maior das aranhas do gênero *Nephila*, e foi proposto que essa atividade poderia catalisar a formação de ligações cruzadas de ditirosina, e dessa forma ajudar na estabilização das proteínas das fibras na seda. No entanto, até então essa formação de ligações cruzadas nas fibras da seda não tem sido relatada, permanecendo por ser elucidada (Vollrath e Knight, 1999; Pouchkina et al., 2003). As peroxidases também foram identificadas nas glândulas de seda de *Bombix mori* e descritas como possivelmente envolvida em *cross-linking* da matrix extracelular da glândula de seda (Suderman et al., 2006; Thein et al., 2009; Dong et al., 2013).

O quarto grupo, proteínas relacionadas com a preservação fibrilar das espidroínas no meio ambiente, incluiu as proteínas acil-CoA desidrogenase e acetil-CoA acetiltransferase. Essas proteínas podem estar envolvidas em processos metabólicos, como catálise e oxidação de ácidos graxos (Thorpe et al., 1995). Estudos relatam que as fibras da seda de aranhas são revestidas com uma camada lipídica que é aderida sobre a fibra, após a sua formação pela aranha (Vollrath et al., 2001; Salles et al., 2006; Sponner et al., 2007). Assim como outras substâncias, os lipídeos encontram-se na solução viscosa, a qual é secretada pela glândula agregada e aderida sobre a fibra (Vollrath & Knight, 2001; Schulz, 2001; Romer et al., 2008). E análises de lipídeos sobre essas fibras tem indicado a presença de uma grande variedade de ácidos graxos. Estudos de análises de GC-MS realizados por Salles et al. (2006) com as fibras da teia da aranha *N. clavipes* demonstraram a presença de uma diversidade de ácidos graxos saturados e insaturados sobre as fibras. No presente estudo essas proteínas somente foram identificadas na glândula agregada e podem estar envolvidas com o metabolismo dos ácidos graxos que são depositados sobre as fibras após a liberação da seda; e essa camada de ácidos graxos pode estar envolvida com a proteção das fibras contra a degradação (Schulz, 2001), e/ou até mesmo como um repelente para alguns insetos e também apresentar uma ação antimicrobiana (Avis et al., 2000; Dani et al., 2003; Salles et al., 2006); e ainda podemos sugerir que essa camada de ácidos graxos pode estar envolvida com a lubrificação da seda, protegendo contra as ações naturais do meio e dessa forma manter a estrutura conformacional das espidroínas e consequentemente as propriedades mecânicas das fibras. Isso será melhor discutido adiante neste texto.

O quinto grupo, proteínas de transporte de íons e moléculas relacionadas com a estabilidade das espidroínas, é constituído por uma série de isoformas de hemocianinas D, E, F

e G. Essas proteínas estão relacionadas principalmente com o transporte de oxigênio e também são capazes de desempenharem atividade de fenoloxidase nos cheliceratas (Rehm et al., 2012). A fenoloxidase ocorre em quase todos os organismos, sendo essencialmente envolvida em vários processos como a resposta imune, cicatrização de feridas, pigmentação e esclerotização nos artrópodes (Averdam et al., 2003; Cunningham et al., 2007; Jaenicke et al., 2008). Além disso, as hemocianinas foram descritas relacionadas, por exemplo, com o transporte de moléculas de oxigênio e íons cloreto, e processos de oxirredução (Rehm et al., 2012; Uniprot/GO: Gene Ontology, código de acesso: Q86N91). Durante o processo de fiação ocorre uma notável mudança no meio, com o aumento na concentração dos íons potássio e fosfato, e uma diminuição na concentração dos íons cloreto e sódio, remoção de água e uma leve acidificação do meio (Knight et al., 2001; Sheibel et al., 2009). A estabilidade das proteínas em uma solução aquosa, geralmente é afetada pelos íons circundantes. De acordo com estudos realizados, íons como sódio e cloreto, presentes na glândula inibem a agregação das proteínas da seda, e conseqüentemente a formação das fibras impedindo a separação da fase líquido-líquido (Sheibel et al., 2009). Enquanto que íons como potássio e fosfato promovem a agregação das proteínas e formação da estrutura. Essa mudança de íons, agindo sinergicamente com a diminuição do pH no meio favoreceria o desdobramento e β -cristalização das proteínas da seda (Knight et al., 2000; Vollrath et al., 2001). Estudos de Knight et al. (2001), sugerem que a capacidade de transportar prótons e íons para dentro do lúmen da glândula, enquanto transporta água para fora pode ser fundamental para a formação das fibras da seda, visto que a aranha ativa e controla cuidadosamente o ambiente iônico durante o processo de fiação, o que proporciona conseqüências importantes durante o dobramento das espidroínas. Sendo assim, os muitos tipos de hemocianinas identificadas em todos os tipos glandulares podem estar relacionados com o transporte/remoção de íons cloreto, sódio, potássio e fosfato; e moléculas de água que atuam no meio influenciando/favorecendo a estabilidade e dobramento das espidroínas.

O sexto e último grupo de proteínas identificado, as denominadas proteínas *housekeeping*, é provavelmente oriundo das células/tecidos glandulares; são proteínas que provavelmente extravasaram durante o método de preparação do extrato proteico das glândulas. As proteínas incluídas nesse grupo foram enolase, 2-fosfo-D-glicerato hidrolase, transaldolase, gliceraldeído 3-fosfato desidrogenase, isocitrato desidrogenase, tubulina, malato desidrogenase citosólica, aldeído desidrogenase e triosefosfato isomerase relacionadas com as funções metabólicas; a maioria associada ao processo de produção de energia, possivelmente para suportar a grande síntese de proteínas nas glândulas. Outras proteínas incluídas nesse

grupo foram actina, paramiosina, fator liberador de histamina, proteína sarcoplasmática ligante de cálcio, nucleosídeo difosfato quinase, dihidropteridina redutase, troponina T, proteína receptora de odor (gustatório) e ubiquitina; essas proteínas podem ser contaminantes do conteúdo glandular oriundas do tecido/musculatura glândular e estão relacionadas com a estrutura, atividade catalítica, organização e regulação do ciclo celular. E ainda as proteínas fator de alongamento da tradução 2 e fator de iniciação da tradução 5A, que associadas aos ribossomos, podem estar relacionadas com a biossíntese de proteínas ao acoplarem a ligação entre os aminoácidos na cadeia peptídica (Kamiie et al., 1993).

A maioria das proteínas identificadas em todas as glândulas produtoras de seda da aranha *N. clavipes*, também já foram identificadas de forma semelhante e descritas nos mais diversos estudos com a glândula de seda de *Bombix mori* (Zhang et al., 2006; Hou et al., 2007; Jia et al., 2007; Li et al., 2012; Dong et al., 2013; Yi et al., 2013). Porém, até então, não havia sido proposto nenhuma relação desses grupos de proteínas identificados com um mecanismo geral de ação destas proteínas para a formação das fibras da seda. Sendo assim, com os resultados obtidos no presente estudo, estamos propondo pela primeira vez, um mecanismo geral de ação das proteínas identificadas nas glândulas produtoras de seda da aranha *N. clavipes* (Figura 32) que possam estar envolvidas com a produção, secreção, armazenamento, transporte, proteção e mudanças conformacionais das espidroínas constituintes da seda durante todo o processo de fiação até a seda ser secretada e também após a sua secreção, já formando as fibras sólidas.

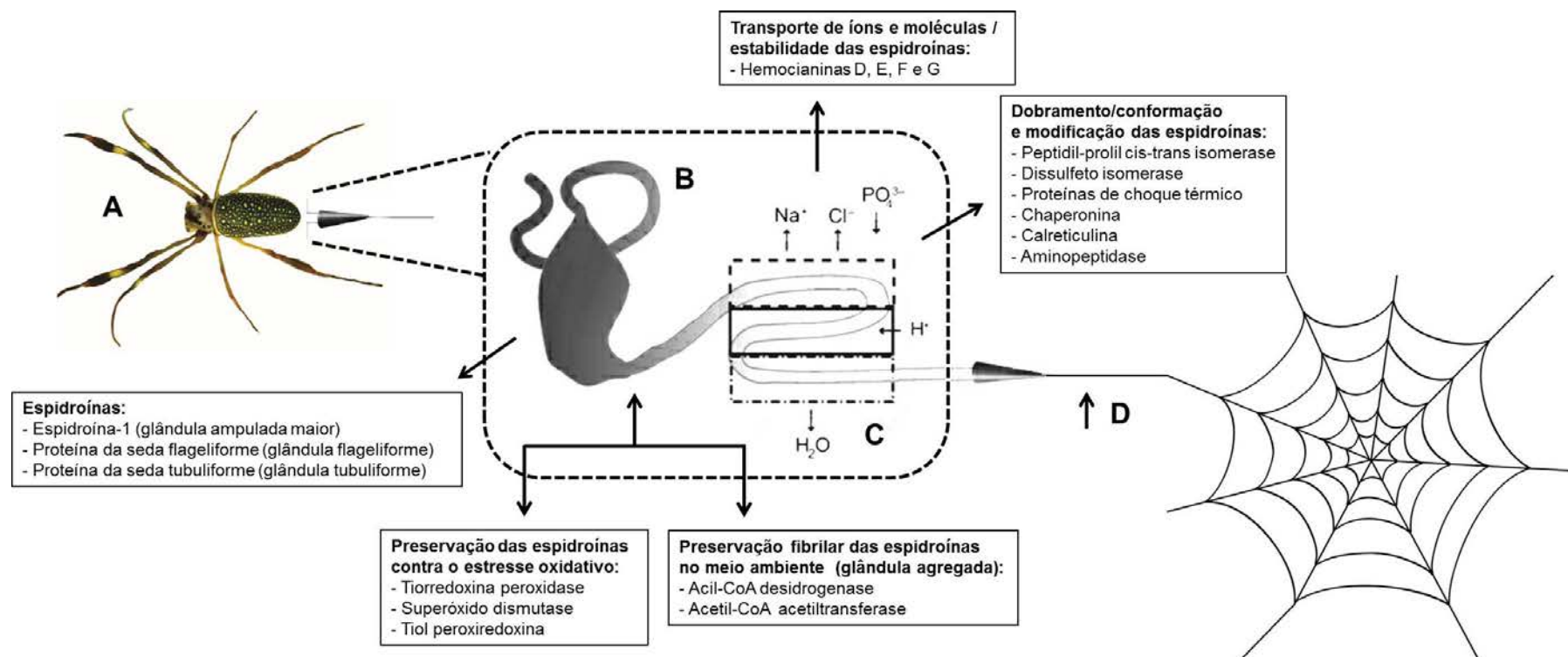


Figura 32. Representação do mecanismo geral de ação das proteínas identificadas nas glândulas produtoras de seda da **(A)** aranha *N. clavipes*. **(B)** - glândula produtora de seda; **(C)** - ducto de fiação; **(D)** - fibras da seda, constituídas pelas espidroínas, após a secreção pelo ducto de fiação.

6.2. Parte 2 - Análise estrutural e modificações pós-traducionais das espidroínas da seda da teia produzida pela aranha *Nephila clavipes*

6.2.1. Identificação e sequenciamento das espidroínas

As proteínas constituintes da seda da teia da aranha *N. clavipes* foram submetidas à análise proteômica com estudos de modificações pós-traducionais. Essa seda apresenta elevada resistência, extensibilidade e elasticidade que caracterizam a teia orbital e o frame. Os fios de seda que constituem a teia orbital e o frame foram coletados e armazenados em frascos individualmente. A teia orbital foi coletada com o auxílio de uma folha de papel mais resistente “tipo cartolina”. Os fios de seda coletados foram limpos com o auxílio de uma pinça para a retirada de partículas, restos de insetos, folhas e galhos aderidos na seda evitando dessa forma uma contaminação da amostra. E com essa limpeza acabou ocorrendo uma perda parcial de fibras. Após a limpeza, tanto os fios de seda da teia orbital quanto os do frame foram pesados em frascos individuais sendo obtidos 300 mg de fios de seda da teia orbital, e 450 mg de fios de seda do frame.

Com a finalidade de se obter informações sobre as características estruturais da superfície dos fios da seda, análises de microscopia eletrônica de varredura (SEM) foram realizadas mostrando a ultraestrutura dos fios da seda constituintes do frame e da teia orbital (Figuras 33 e 34). As fibras parecem ser constituídas de vários fios entrelaçados entre si, que formam estruturas em camadas sobrepostas; também é possível notarmos, o que pode ser uma emenda dos fios (Figura 34B), realizado pela aranha; e ainda podemos visualizar a presença de estruturas macroporosas conectadas entre si (Figura 34D).

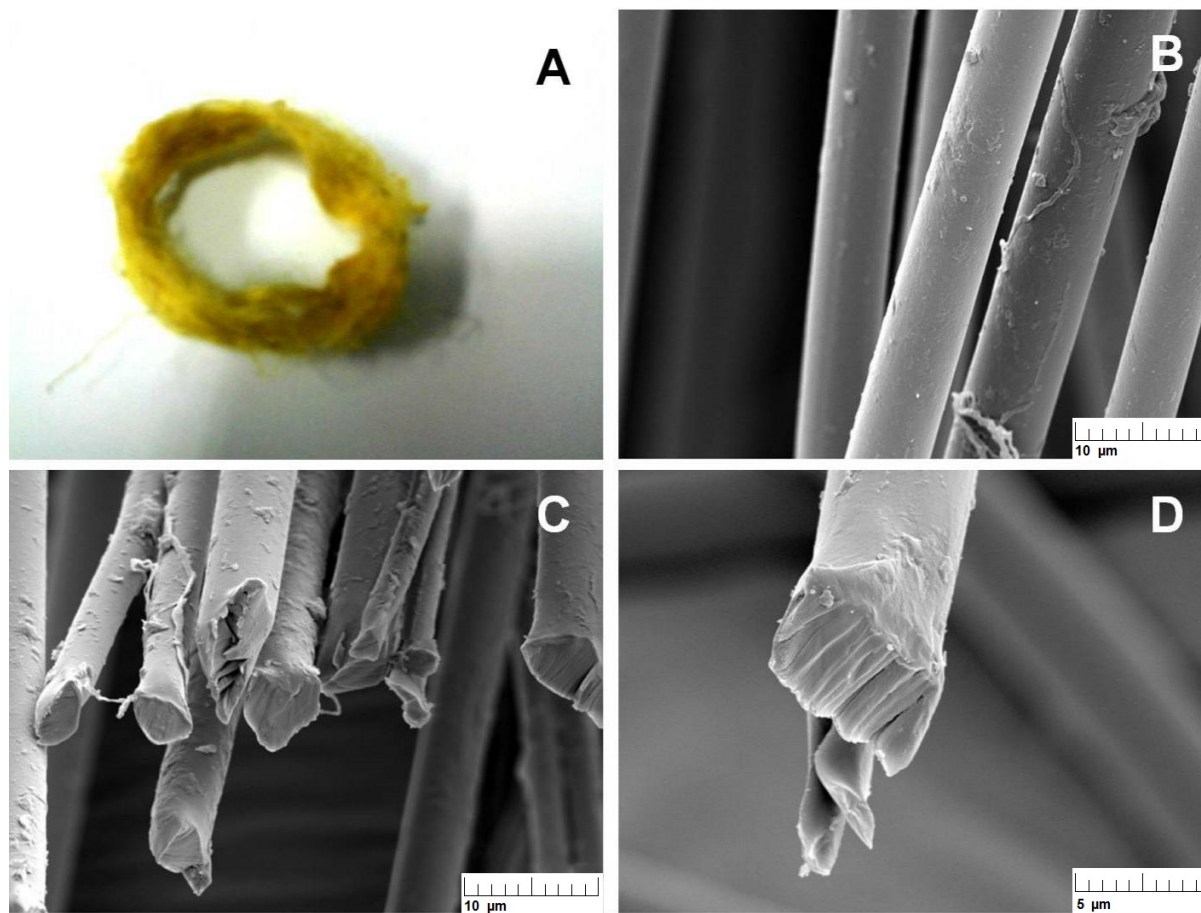


Figura 33. SEM – Microscopia eletrônica de varredura da seda do “frame” da teia da aranha *N. clavipes*. **(A)** seda coletada constituída pelos fios dos raios do “frame” enrolados; **(B-D)** ultraestrutura dos fios da seda em diferentes pontos da teia.

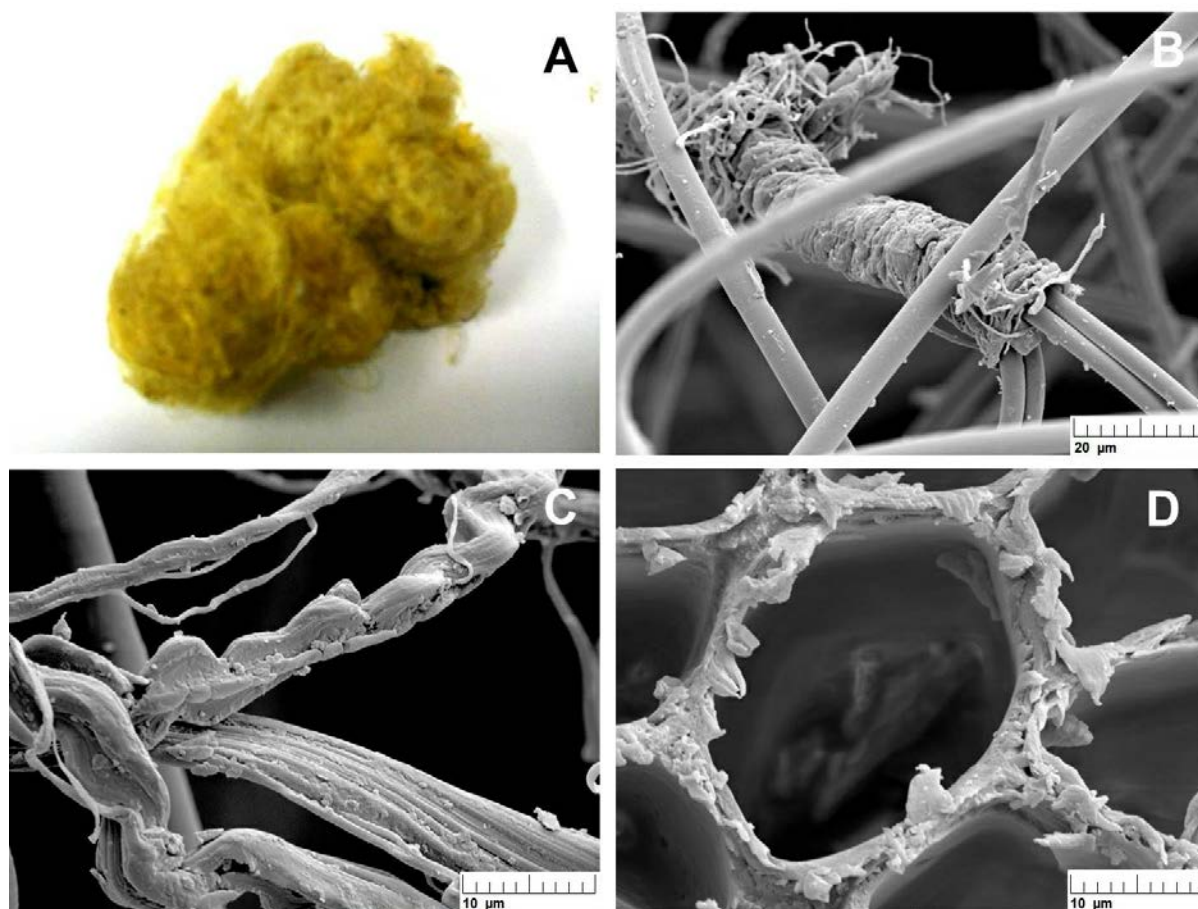


Figura 34. SEM – Microscopia eletrônica de varredura da seda da teia orbital da aranha *N. clavipes*. **(A)** seda coletada constituída pelos fios dos raios e da espiral de captura da teia orbital; **(B-D)** ultraestrutura dos fios da seda em diferentes pontos, destacando a presença de macroporos.

Um dos aspectos técnicos que contribuiu para o sucesso na análise dos fios de seda já pronta (madura), a partir do material obtido diretamente da teia, foi a solubilização dos fios de seda em tiocianato de lítio. E para demonstrar a solubilidade da seda durante a preparação da amostra, resultados de microscopia eletrônica de varredura revelaram a ultraestrutura da seda durante o processo de dissolução dos fios em solução saturada de tiocianato de lítio (Figura 35). A dissolução completa da seda após 60 minutos de contato com essa solução mostrou a homogeneidade da amostra que foi preparada para a análise proteômica.

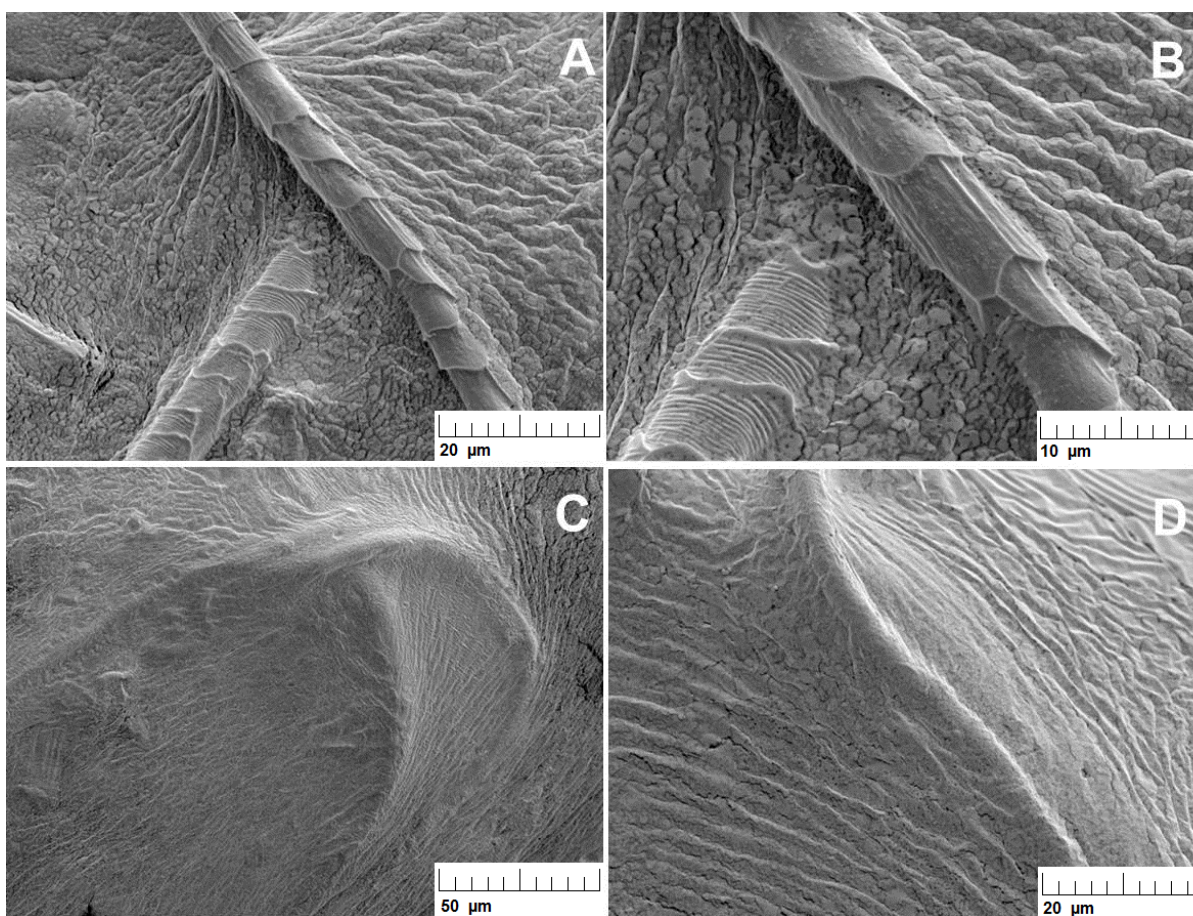


Figura 35. SEM - Microscopia eletrônica de varredura da seda da teia da aranha *N. clavipes*. (A-D) mostra o padrão de dissolução da seda, em diferentes pontos, na presença de solução saturada de tiocianato de lítio (LiSCN hidratado 38M).

Após a extração das proteínas, as mesmas foram quantificadas pelo método de Bradford, e submetidas à eletroforese bidimensional - 2-DE (n=8), apresentando um alto grau de identidade entre os géis analisados de cada amostra individualmente. Como as proteínas constituintes da seda de aranhas são descritas na literatura com elevado peso molecular (acima de 200 kDa), optou-se por realizar a 2-DE em um gel SDS-PAGE com gradiente de 7-10%, e também utilizamos fitas *Immobilized* pH 7-10 gradiente não linear para uma melhor resolução do perfil proteico. A figura 36 mostra o perfil da eletroforese bidimensional da seda do frame, sendo possível visualizar um total de 4 *spots* com MW acima de 250 kDa, e pI no intervalo de pH 9 a 10 (gradiente não linear); enquanto a figura 37 mostra o perfil da eletroforese bidimensional da seda da teia orbital, sendo possível visualizar um total de 6 *spots* com MW acima de 250 kDa, e pI no intervalo de pH 9 a 10 (gradiente não linear).

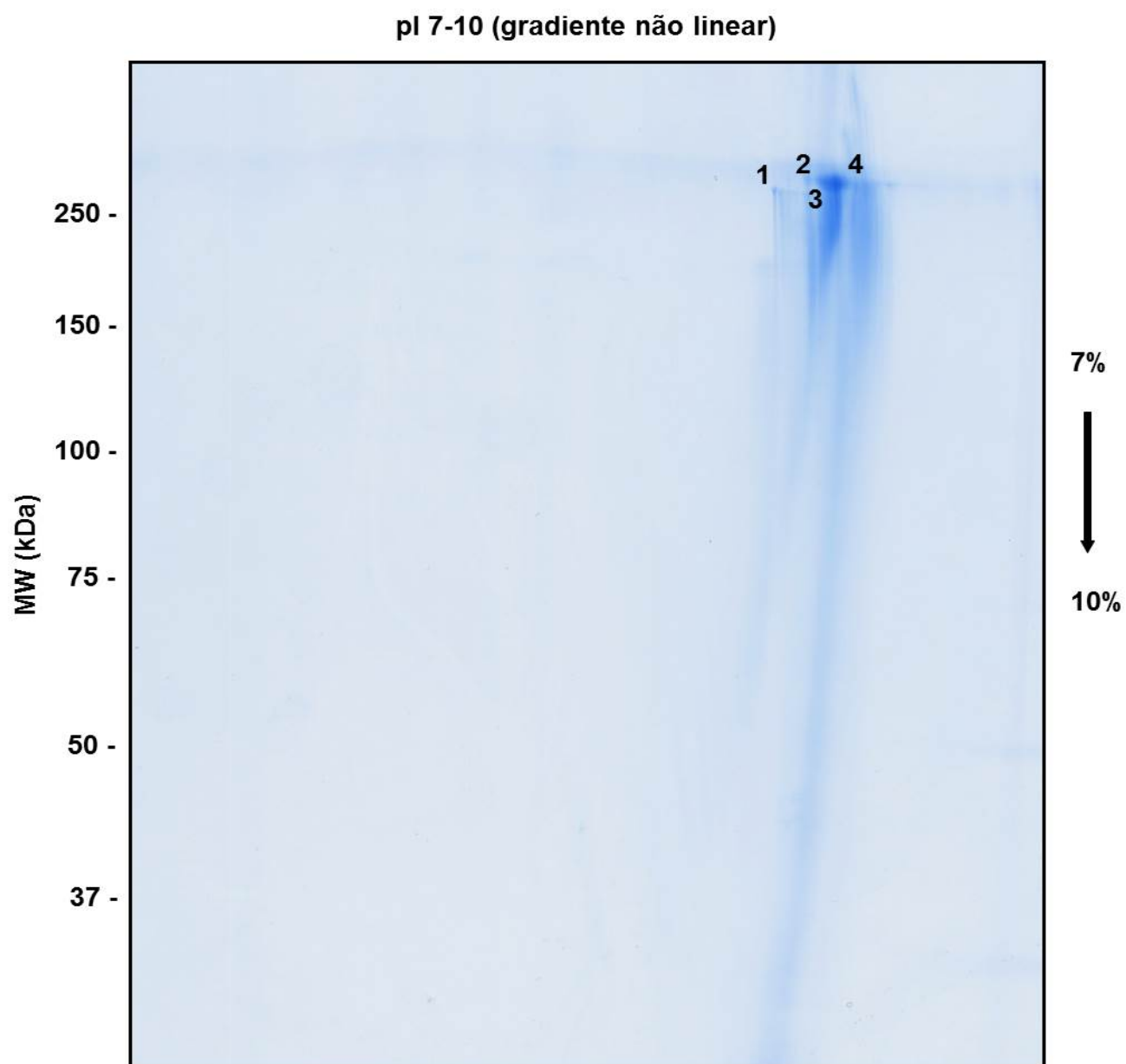


Figura 36. Perfil eletroforético da 2-DE da seda do “frame” da aranha *N. clavipes*.

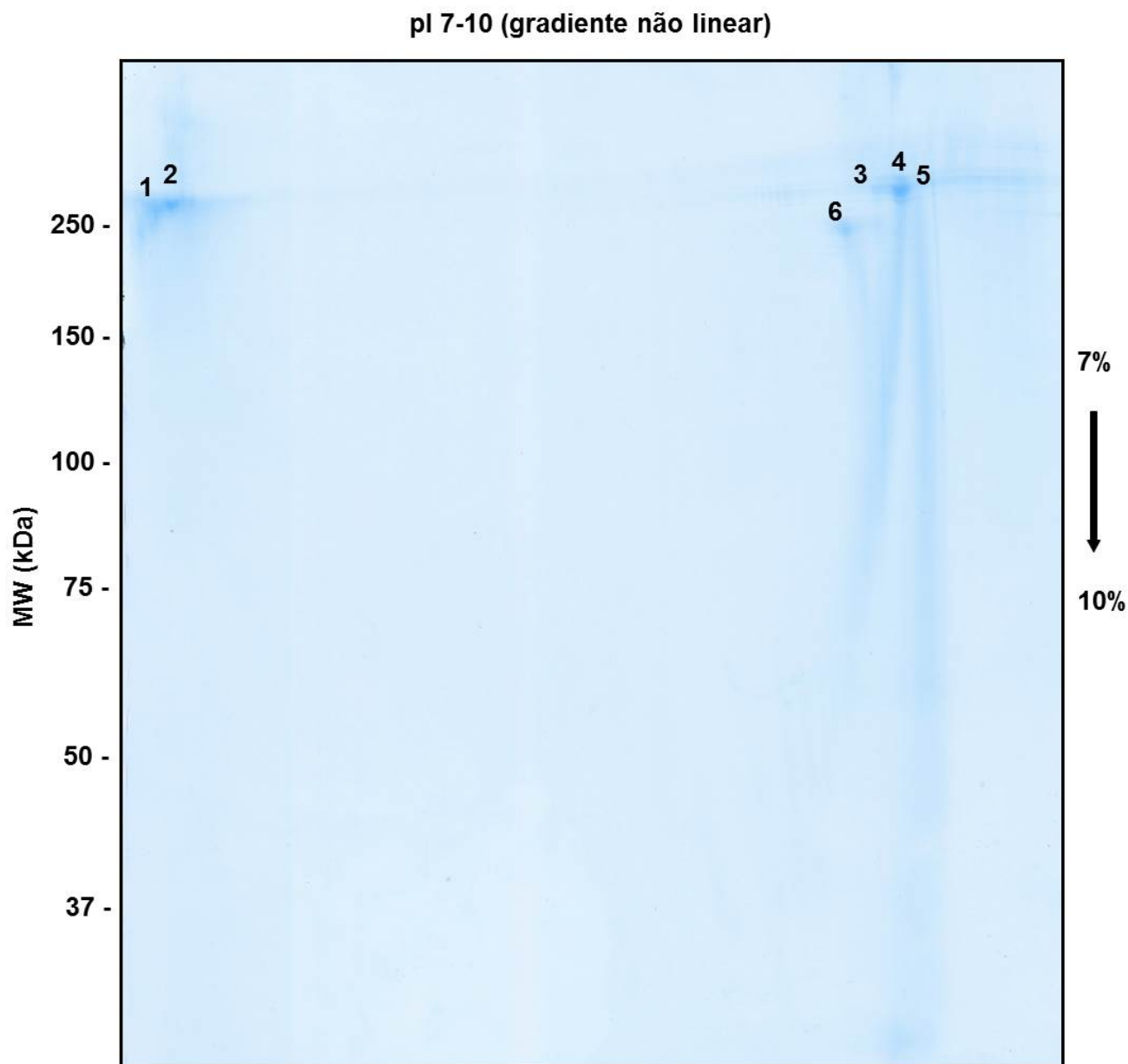


Figura 37. Perfil eletroforético da 2-DE da seda da teia orbital da aranha *N. clavipes*.

Todos os *spots* foram excisados dos géis e submetidos à digestão enzimática utilizando-se 8 diferentes enzimas proteolíticas como tripsina, quimiotripsina, pepsina, Glu-C/V8 protease, proteinase 10, proteinase K, Asp-N protease e subtilisina. Após a digestão enzimática, as amostras foram analisadas por espectrometria de massas. A tabela 11 mostra a identificação de todos os *spots* analisados, sendo identificadas as proteínas espidroína-1 (*spots* 1, 2, 3 e 4) e espidroína-2 (*spots* 3 e 4) na seda do frame; e as proteínas da seda flageliforme (*spots* 1 e 2) e espidroína-1 (*spots* 3, 4, 5 e 6) na seda da teia orbital. Na seda do frame as espidroínas-1 e -2

foram identificadas nos mesmos *spots* (3 e 4), sugerindo uma coexistência das duas formas nas fibras, associadas talvez por *cross-links*.

Tabela 11. Proteínas identificadas na seda da aranha *N. clavipes* através da digestão *in-gel* dos *spots* oriundos da 2-DE.

Spots	Código de Acesso	Proteína	Exp. MW (kDa)	Calc. MW (kDa)	% cobertura da sequência *
Proteínas identificadas na 2-DE da seda do frame					
1, 2, 3 e 4	P19837	Espidroína-1	> 250	61.21	98.12
3 e 4	B5SYS7 P46804	Espidroína-2	> 250	69.03	91.02
Proteínas identificadas na 2-DE da seda da teia orbital					
1 e 2	O44358 O44359	Proteína da seda flageliforme	> 250	74.33	96.14
3, 4, 5 e 6	B5SYS5 B5SYS6 P19837	Espidroína-1	> 250	86.77	98.94

* MS/MS: Todas as sequências peptídicas são mostradas no material suplementar, apêndices A - Tabelas S1, S2, S3 e S4.

A proteína da seda flageliforme foi identificada com os códigos de acesso GenBank ID O44358 (N-terminal) e O44359 (região central e C-terminal) apresentando uma cobertura da sequência de 96.14%; a espidroína-1 foi identificada com os códigos de acesso GenBank ID B5SYS5 (N-terminal - espidroína-1A), B5SYS6 (N-terminal - espidroína-1B) e P19837 (região central e C-terminal) apresentando uma cobertura da sequência de 98.94%; nas aranhas do gênero *Nephila* existem evidências da ocorrência de dois genes que codificam a proteína espidroína-1, sendo denominado então de espidroína-1A e espidroína-1B (Rising et al., 2007; Gaines et al., 2008); e a espidroína-2 foi identificada com os códigos de acesso B5SYS7 (N-terminal) e P46804 (região central e C-terminal) apresentando uma cobertura da sequência de 91.02%. Todas as sequências foram identificadas a partir de dados de sequências parciais depositados nos bancos de dados (Xu et al., 1990; Hinman et al., 1992; Beckwitt et al., 1994; Hayashi et al., 1998; Gaines et al., 2008). Valores de massas moleculares teóricos foram determinados para espidroína-1 de 86.77 kDa, para espidroína-2 de 69.03 kDa e para proteína da seda flageliforme de 74.33 kDa. Estudos demonstram que as espidroínas constituintes das fibras da seda apresentam MW em torno de 250-300 kDa após serem expelidas pelo ducto de fiação formando as fibras sólidas (Sponner et al., 2005; Ayoub et al., 2007; Hagn et al., 2010).

No entanto, os elevados valores observados para as massas moleculares mostradas na 2-DE (Figuras 36 e 37) e tabela 9 podem ser devido à existência de *cross-links* ou glicosilação, modificação conhecida por reduzir a mobilidade das proteínas no gel; ou/e ainda por outras modificações pós-traducionais que também possam influenciar o peso molecular ou o comportamento das espiroínas no gel (Thiel et al., 1994; Michal et al., 1996; Augsten et al., 2000). Isso será melhor discutido adiante neste texto.

Considerando-se a grande quantidade de fragmentos peptídicos que foram gerados na identificação e sequenciamento das proteínas oriundas da 2-DE da seda, e para que a redação da tese ocorresse de forma contínua de modo que a sua leitura não fosse interrompida, a relação completa de todos os fragmentos peptídicos identificados está demonstrada no material suplementar, apêndices A - tabelas S1, S2, S3 e S4 constando todas as informações dos dados oriundos da análise proteômica.

Resultados da digestão *in gel* utilizando diferentes enzimas proteolíticas mostrou uma elevada cobertura das sequências. A tabela 12 mostra a eficiência de cada enzima utilizada para todas as proteínas identificadas.

Tabela 12. Porcentagem de cobertura da sequência das proteínas identificadas na seda da aranha *N. clavipes* através da digestão *in-gel* utilizando diferentes enzimas proteolíticas.

Enzima	Espiroína-1	Espiroína-2	Proteína da seda flageliforme
Quimiotripsina (%)	89.37	72.16	86.76
Tripsina (%)	46.39	25.19	10.80
Pepsina (%)	31.91	25.83	Não identificado
Glu-C/V8 protease (%)	23.96	Não identificado	26.85
Proteinase 10 (%)	16.48	45.82	55.45
Asp-N protease (%)	13.87	Não identificado	Não identificado
Proteinase K (%)	12.03	Não identificado	Não identificado
Subtilisina (%)	9.61	Não identificado	19.29
Total (%)	98.94	91.02	96.14

Eficiência enzimática foi avaliada pela combinação da cobertura da sequência obtida de todos os *spots* correspondentes a cada proteína individualmente.

A figura 38 mostra o alinhamento das sequências primárias das proteínas espidroína-1A, espidroína-1B, espidroína-2 e proteína da seda flageliforme identificadas na seda da aranha *N. clavipes*. Os valores de identidade e similaridade entre essas sequências foram 90 e 94.4%, respectivamente, entre a espidroína-1A e a espidroína-1B; 48.5 e 60.5%, respectivamente, entre a espidroína-1A e a espidroína-2; 46.4 e 58.7%, respectivamente, entre a espidroína-1B e a espidroína-2; 39.3 e 59.8%, respectivamente, entre a espidroína-1A e a proteína da seda flageliforme; 38.5 e 61.2%, respectivamente, entre a espidroína-1B e a proteína da seda flageliforme; e 45.5 e 68.9%, respectivamente, entre a espidroína-2 e a proteína da seda flageliforme. Podemos observar que os resíduos assinalados em azul escuro são aqueles que se conservam em todas as sequências primárias que foram alinhadas, ou seja, são os resíduos de aminoácidos idênticos em todas as sequências obtidas; em sua grande maioria os trechos de sequência conservados são aqueles ricos em resíduos de glicina. Os resíduos assinalados em azul claro são os resíduos de aminoácidos que apresentam propriedades físico-químicas similares.

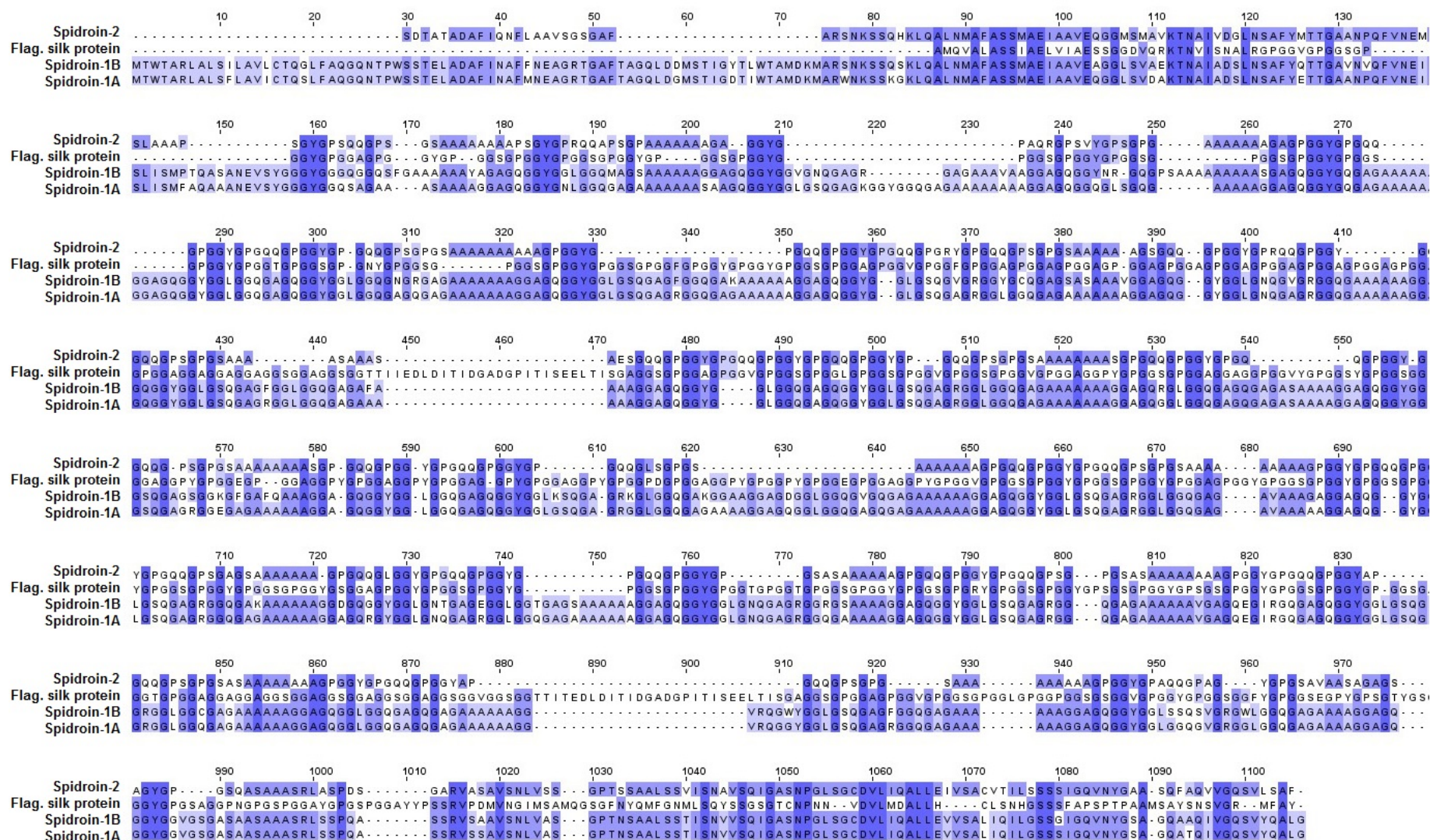


Figura 38: Alinhamento múltiplo entre as sequências primárias das proteínas espidroína-1A, espidroína-1B, espidroína-2 e proteína da seda flageliforme (Flag. silk protein) identificadas na seda da aranha *N. clavipes*. As regiões de identidade mais importantes estão assinaladas em azul escuro, enfatizando os resíduos altamente conservados nesse tipo de alinhamento.

As proteínas espidroína-1 e espidroína-2 são secretadas pela glândula ampulada maior, e constituem as fibras da seda dos raios da teia orbital, do frame e da linha de segurança “escape” (Vollrath, 2000; Römer et al., 2008); e podem ser codificadas por alguns genes diferentes (Beckwitt & Arcidiacono 1994; Römer et al., 2008; Rising et al., 2007; Ayoub et al., 2008). A espidroína-1 está presente em maior quantidade (aprox. 80%) na seda do que espidroína-2, sendo encontrada uniformemente distribuída na fibra. Já a espidroína-2 está ausente nas regiões periféricas formando aglomerados em determinadas áreas centrais da fibra (Sponner et al., 2005; Yazawa et al., 2011). As propriedades mecânicas da seda produzida pela glândula ampulada maior dependem da região central rica em unidades altamente repetitivas de glicina seguida por alanina (Rising et al., 2007). A espidroína-1 possui a repetição de sequências GGX, (GA)_n, e (A)_n, enquanto a espidroína-2 possui a repetição de sequências (A)_n e GPGXX. As propriedades de força e resistência da fibra são devido à presença das unidades altamente repetitivas de (A)_n e/ou (GA)_n, que formam estruturas folhas-β cristalinas na fibra, enquanto as propriedades de elasticidade são devido à presença das unidades altamente repetitivas de GGX e GPGXX (Simmons et al., 1996; Gatesy et al., 2001; Rising et al. 2011). Os segmentos ricos em glicina são postulados para formar diferentes estruturas como as espiral-β e estruturas aleatórias (*coil*) (Hayashi et al., 1999; Van Beek et al., 2002). A espidroína-1 apresenta uma quantidade baixa de resíduos de prolina, enquanto espidroína-2 apresenta uma grande quantidade desse aminoácido em sua sequência (Savage et al., 2008). Rauscher et al. (2006) identificaram o resíduo de prolina como sendo o determinante primário das propriedades elásticas na espidroína-2, enquanto que a espidroína-1 apresenta maiores propriedades de força e resistência. Sendo assim, a interação dessas proteínas na fibra proporciona uma combinação excepcional de resistência e extensibilidade na seda.

A proteína da seda flageliforme é secretada pela glândula flageliforme, e constitui as fibras da seda da espiral de captura da teia orbital. Essa seda é altamente elástica e em conjunto com as propriedades da seda do frame e do raio da teia orbital, dissipa perfeitamente o impacto da energia das presas. A enorme resistência dessas fibras é importante para a captura e detenção de presas que às vezes são maiores do que as aranhas (Römer et al., 2008; Lefèvre et al., 2011). A proteína da seda flageliforme possui a repetição de sequências GPGGX e GGX, para as quais tem sido proposta a formação de estruturas espiral-β que podem atuar como “*nanosprings*” moleculares proporcionando elasticidade às fibras (Hu et al., 2006). Ao contrário das outras proteínas constituintes da seda, a proteína da seda flageliforme não apresenta mudança conformacional durante o processo de fiação (Lefèvre et al., 2011). Estudos indicam a ausência de estruturas cristalinas na fibra (Craig et al., 2002) e sua estrutura

secundária tem sido descrita como apresentando uma distribuição de conformações, sem uma orientação molecular preferencial. A grande quantidade de resíduos de prolina distribuída regularmente ao longo da repetição de sequência GPGGX, evita a formação de estruturas folhas- β cristalinas na fibra (Ohgo et al., 2006; Rousseau et al., 2009).

As espidroínas identificadas são sintetizadas pelas células epiteliais das glândulas e secretadas no lúmen glandular, onde são armazenadas como uma solução de fição líquida e cristalina altamente concentrada (Scheibel, 2004) até serem processadas em fibras. Essas proteínas são compostas por um domínio N-terminal não repetitivo, uma região central altamente repetitiva composta por aproximadamente 100 segmentos de sequências ricas em polialanina/glicina, e um domínio C-terminal não repetitivo (Rising et al., 2011). Estudos têm demonstrado que as três partes que compõem a espidroína são estruturalmente e funcionalmente distintas, e até o desenvolvimento do presente estudo ainda não havia sido detectada experimentalmente a proteína completa, i.e., com os domínios N-terminal – região central – e domínio C-terminal numa única molécula. O conhecimento que se tem sobre isso decorre da sequência completa de DNA dos genes das espidroínas, e da detecção isolada de cada um destes domínios na forma proteica. As propriedades mecânicas das fibras da seda têm sido consideradas tradicionalmente como o mais importante parâmetro funcional. No entanto, a produção, secreção, armazenamento, e em particular a conversão controlada das espidroínas do estado solúvel para insolúvel apresenta desafios significativos, sendo provável que os diferentes domínios sejam especializados para atuar em diferentes etapas do processo de fição (Rising et al., 2011). A região N-terminal é a parte mais conservada das proteínas da seda e, embora a sua função esteja relacionada ao transporte das espidroínas para o lúmen glandular devido à presença de um peptídeo sinal (Moutriuk-Smith et al., 2005), a sua função na estrutura proteica ainda permanece desconhecida. Recentemente, estudos estruturais através de raios-X de alta resolução realizados com o N-terminal recombinante da espidroína de *E. australis* têm sugerido que o domínio N-terminal controla a formação das fibras, inibindo a agregação precoce das espidroínas durante o armazenamento, e também acelerando e direcionando a formação das fibras durante o processo de fição conforme o pH vai sendo reduzido ao longo do ducto de fição (Askarieh et al., 2010). A região C-terminal, constituída por aproximadamente 100 resíduos de aminoácidos, também se apresenta altamente conservada entre as espidroínas. Devido a sua alta conservação, estudos tem sugerido uma importante atuação na polimerização das fibras (Sponner et al., 2005). Acredita-se que, a medida que o pH e a composição de sais são alteradas ao longo do ducto de fição, os domínios N e C-terminal possam desencadear a polimerização das fibras (Hagn et al., 2010).

6.2.2. Identificação das PTMs e aminoácidos substitutos

Com uma elevada cobertura das sequências, uma série de modificações pós-traducionais e aminoácidos substitutos foram observados (Tabelas S1, S2, S3 e S4 - material suplementar, apêndices A). Uma série de aminoácidos substitutos (polimorfismos) foi observada nas proteínas espidroína-1 e proteína da seda flageliforme, baseados nas sequências depositadas nos bancos de dados:

- **Espidroína-1, código de acesso P19837:** G574W; G617W; A39N; Q41R; Q464R; R69F; R161F; R585F; G75K; G302K; G309K; G317K; G403K; A97V; A115V; A131V; A334V; A109S; A111S; A437S; A466S; L102Y; G104C; A172F; A270F; A273F; A174Q; G229R; R265S; E268K; A318G; A319G; A378G; Q326D; A412D; R415G; Q423T; Q433T; R427E; Q534C; G609S; G610S; G612S; S664A; S723G; T736A.
- **Espidroína-1, código de acesso B5SYS5:** D55G; D59G; K68W; S77W; Q126E; N144S; S149A; S150A; G177A; E202G; e R216K.
- **Espidroína-1, código de acesso B5SYS6:** L43F; D55G; D65Y; K68W; Q104A; F146P; A147T; A169F; S177Y; G192M; S202G; A226V; e G238R.
- **Proteína da seda flageliforme:** Y75S; G104N; S135Y; V287L; A335V; E394A; S403P; Y404D; G567R; S568Y; V710L; G794S; Y799N; e N864D.

Essas substituições de aminoácidos podem ser devido às mutações, polimorfismo ou simplesmente refletem a alteração na composição de aminoácidos dessas proteínas em resposta à pressão ambiental (como p. Ex.: clima, variação do tipo de presas, etc). Estudos demonstram que a expressão de proteínas da seda de aranhas é alterada devido à interação entre a vibração, manipulação e tipo de presas (Tso et al., 2005; Blamires et al., 2010; Blamires et al., 2012). Essas substituições podem alterar a conformação proteica, e até mesmo as propriedades mecânicas e elásticas da seda durante a formação das fibras. A tabela 13 mostra uma série de modificações pós-traducionais como fosforilação, glicosilação, hidroxilação, nitrotirosina, dihidroxilação, deamidação, metilação e oxidação que foram observadas nas sequências das proteínas espidroína-1, espidroína-2 e proteína da seda flageliforme; e detectadas com o uso dos algoritmos MASCOT e do Modiro[®] na análise dos dados experimentais. Ao final do presente estudo foi possível apresentar a sequência primária obtida

das proteínas identificadas, mostrando a localização das PTMs observadas. A relação completa de todos os fragmentos peptídicos identificados mostrando a localização das PTMs observadas nas sequências, de acordo com o código de acesso de cada proteína, está demonstrada no material suplementar, apêndices A - Tabelas S1, S2, S3 e S4.

Tabela 13. Modificações pós-traducionais (PTMs) observadas nas proteínas da seda da aranha *N. clavipes* mostradas através do MASCOT e Modiro®.

Spot	Código de acesso	Proteína	PTMs observadas através do MASCOT	PTMs observadas através do Modiro®
PTMs observadas nas proteínas identificadas na 2-DE da seda do Frame				
1	P19837	Espidroína-1	Phosphorylation (Y59, Y89, Y151, S260, Y353, Y387, Y478); Hydroxymethyl (N128); Deamidation (R161); Methylation + Deamidation (Q314, Q326, Q332, Q336)	Methylation (Q37, Q41, Q282, Q291, Q295, S742); Deamidation (Q181, R585); Dimethylation (R308)
2	P19837	Espidroína-1	Methylation (Q37, Q41, Q129, Q136); Phosphorylation (Y59, Y151, Y353, Y387, Y478)	Methylation (Q37, Q41, R69, Q72, Q282, Q291, Q295); Deamidation (Q181, Q252, Q611, Q621, Q633); Methylation + Deamidation (Q504, Q514)
3	P19837	Espidroína-1	Methylation (Q37, Q41); Phosphorylation (S64, Y151, S260)	Deamidation (Q30, Q136, Q148, R161, Q181, Q336, Q350); Methylation (Q95, Q282, Q291, Q295)
3	B5SYS7	Espidroína-2	Oxidation (M90, M96, M107, M109, M126, M139); Methylation + Deamidation (N121, N137)	Dihydroxylation (S95); Hydroxylation (N89); deamidation (R140); Aminotirosina (Y125)
	P46804		Phosphorylation (S116, S245, S383, T555, S601, S602, S603, Y609, S613) Deamidation (Q237)	Dihydroxylation (P189, P230)
4	P19837	Espidroína-1	Deamidation (Q56, Q129, Q148, Q336, Q350, Q464); Phosphorylation (Y151, S156, Y184, S462); Dimethylation (R161)	Methylation (Q37, Q41, Q332, Q336, S742); Deamidation (R69, Q72, R161, Q181, R585, Q588)
4	B5SYS7	Espidroína-2	Oxidation (M90, M107, M109, M126, M139); Deamidation (N132)	Oxidation (M90, M107, M109, M126, M139); Hydroxylation (N89, N132); Deamidation (R140)
	P46804		Phosphorylation (S22, Y61, Y609, S613)	Methylation (Q621)
PTMs observadas nas proteínas identificadas na 2-DE da seda da teia orbital				
1	O44359	Proteína da seda flageliforme	Dihydroxylation (P379); Phosphorylation (Y380)	Nitrotirosina (Y336); Hydroxylation (P382)
2	O44358	Proteína da seda flageliforme		Oxidation (M52); Deamidation (R38)
	O44359		Oxidation (M52, M843); Phosphorylation (T95, Y389)	Hydroxylation (P2, P17, P27, P37, P47, P57, P67, P382, P406, P415, P435, P460, P470, P480, P490, P500, P520, P530); Methylation (E362); Acetylation (S453); Deamidation (N828); Oxidation (M826, M831, M834)
3	B5SYS6	Espidroína-1	Phosphorylation (T49, S95, S144); Sulphone (M60); Oxidation (M90, M96, M145)	Oxidation (M1, M71, M74, M90, M96, M145); Hexose (N44); Hydroxylation (D58); Methylation + Deamidation (Q86)
	P19837		Phosphorylation (Y31, S64, S156, S260, S526, Y575, S580)	Nitrotirosina (Y18); Phosphorylation (S243); Deamidation (R261, R308); Methylation (S742)
4	B5SYS6	Espidroína-1	Oxidation (M71, M96, M145); Phosphorylation	Sulphone (M60); Oxidation (M74, M90, M96, M145);

			(S95, S144, S195); Deamidation (Q191)	Methylation (E103); Phosphorylation (T127); HexNAc2DHex2 (N237)
	P19837		Phosphorylation (S64, S156, Y184, S260, Y575, S580); Methylation (Q37, Q41)	Nitrotyrosina (Y18); Phosphorylation (Y151, S243); Deamidation (R652)
5	B5SYS6	Espidroína-1	Phosphorylation (T49, S119, S195); Oxidation (M74, M96, M145); Methylation (Q215)	Hexose (N44); Sulphone (M60); Oxidation (M1, M74, M90, M96, M145); Deamidation (R140)
	P19837		Phosphorylation (S64, S156, S260, Y387, Y575, S580) Methylation + Deamidation (Q129, Q136)	Deamidation (R615)
6	B5SYS5	Espidroína-1	Sulphone (M60); Oxidation (M43, M71, M90, M96, M145); Methylation (T113); Phosphorylation (S164, S170, S211, S244)	Phosphorylation (T53, T62, S164, S170, S211, S244); Sulphone (M60); Oxidation (M1, M43, M90, M96); Methylation (E103, S108, Q241)
	P19837		Phosphorylation (Y31, Y59, Y151, S156, S260, Y353, Y387, Y478)	Deamidation (Q326, Q433); Methylation + Deamidation (Q572, Q588)

Muitas dessas modificações incluindo deamidação, metilação e oxidação da metionina podem ser consideradas artefatos experimentais oriundos da manipulação de amostras e procedimentos analíticos realizados durante os experimentos. No entanto, estudos relatam a presença de deamidases e deiminases que conduzem a deamidação como uma modificação (Chikuma et al., 1996; Kubota et al., 2005; Gyorgy et al., 2006; Smith et al., 2009). No presente estudo foram observados sítios de deamidação em alguns resíduos de arginina nas sequências das proteínas espidroína-1, espidroína-2 e proteína da seda flageliforme. A citrulinização, produto da deamidação da arginina, também foi identificada na fibroína de cadeia leve da seda de *Bombyx mori* (Chen et al., 2010). E embora os mecanismos e as consequências funcionais da deamidação nas proteínas estruturais da seda sejam indeterminados, possivelmente este tipo de PTM pode estar envolvido na construção da conformação estrutural da seda.

Na sequência obtida da espidroína-1 foi observada uma grande quantidade de sítios de fosforilação: 12 sítios de fosforilação em T49, T53, T62, S95, S119, T127, S144, S164, S170, S195, S211 e S244 foram observados no domínio N-terminal; e 16 sítios de fosforilação em Y31, Y59, S64, Y89, Y151, S156, Y184, S243, S260, Y353, Y387, Y462, Y478, S523, Y575 e S580 foram observados na região central e no domínio C-terminal. Também foram observados dois sítios de glicosilação em N44 (Hexose) e N237 (HexNAc2DHex2), um sítio de dihidroxilação (Sulphone) em M60, e um sítio de nitrotyrosina em Y18. A figura 39 mostra a representação da sequência primária da espidroína-1 destacando a localização das PTMs mais abundantes observadas na proteína. A localização das PTMs observadas, ou seja, a posição em que elas ocorrem na sequência foi estabelecida de acordo com a sequência primária das proteínas obtidas no presente estudo, após o alinhamento e sobreposição de todos os fragmentos peptídicos identificados, conforme mostrado nas figuras de representação das sequências primárias de cada proteína. Sendo assim, a numeração dessa localização na

sequência obtida não corresponde à posição das sequências presentes nos códigos de acesso pelos quais as proteínas foram identificadas nos bancos de dados, já que as regiões N-terminal, central e C-terminal de uma mesma proteína foram identificadas com diferentes códigos de acesso.

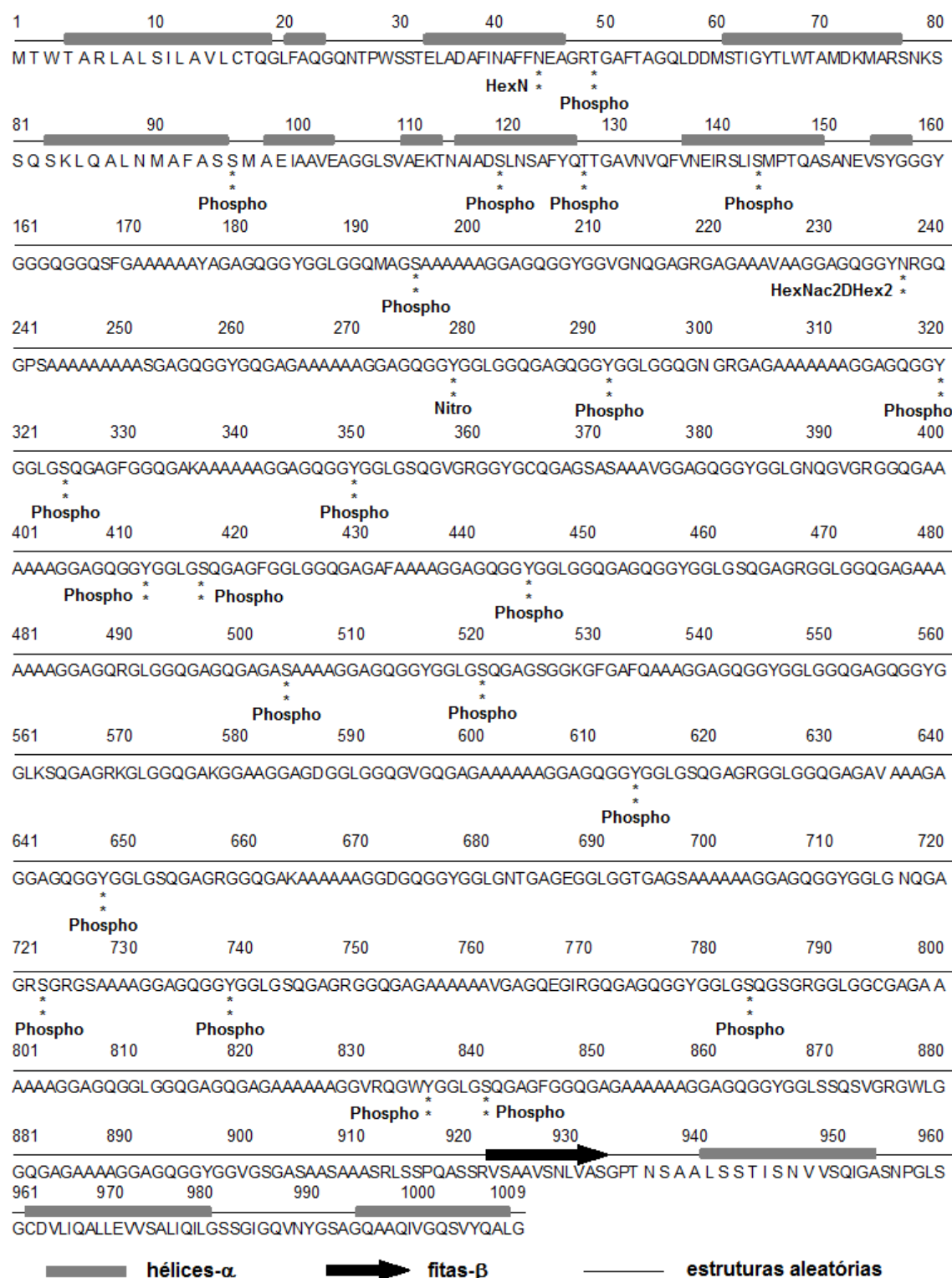


Figura 39. Representação da sequência primária e caracterização da estrutura secundária (JNETPRED *Secondary Structure Prediction*) da proteína espidroína-1 que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda da aranha *N. clavipes*. A sequência foi obtida através da digestão *in-gel* utilizando diferentes enzimas proteolíticas, e mostra a localização das PTMs mais abundantes observadas na proteína.

Na sequência da espidroína-2 foram observados 11 sítios de fosforilação em S22, Y61, S116, S245, S383, T555, S601, S602, S603, Y609 e S613. A figura 40 mostra a representação da sequência primária da espidroína-2 destacando a localização das PTMs mais abundantes observadas na proteína.

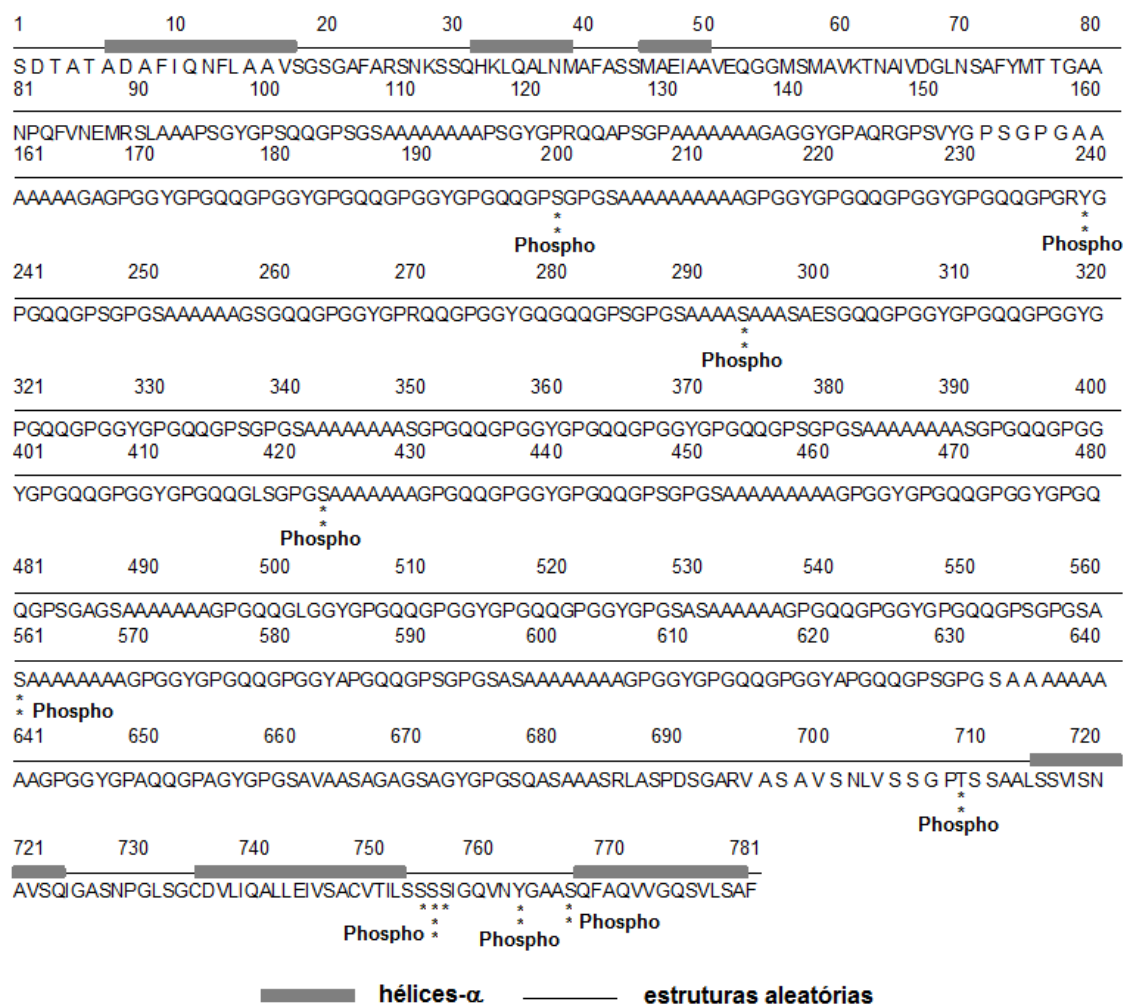


Figura 40. Representação da sequência primária e caracterização da estrutura secundária (JNETPRED *Secondary Structure Prediction*) da proteína espidroína-2 que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda da aranha *N. clavipes*. A sequência foi obtida através da digestão *in-gel* utilizando diferentes enzimas proteolíticas, e mostra a localização das PTMs mais abundantes observadas na proteína.

Na sequência obtida da proteína da seda flageliforme foram observados 18 sítios de hidroxilação no resíduo de prolina, sendo caracterizados como hidroxiprolina: P2, P17, P27, P37, P47, P57, P67, P382, P406, P415, P435, P460, P470, P480, P490, P500, P520 e P530;

também foram observados três sítios de fosforilação em T95, Y380 e Y389; e um sítio de nitrotirosina no resíduo Y336. A figura 41 mostra a representação da sequência primária da proteína da seda flageliforme destacando a localização das PTMs mais abundantes observadas na proteína.

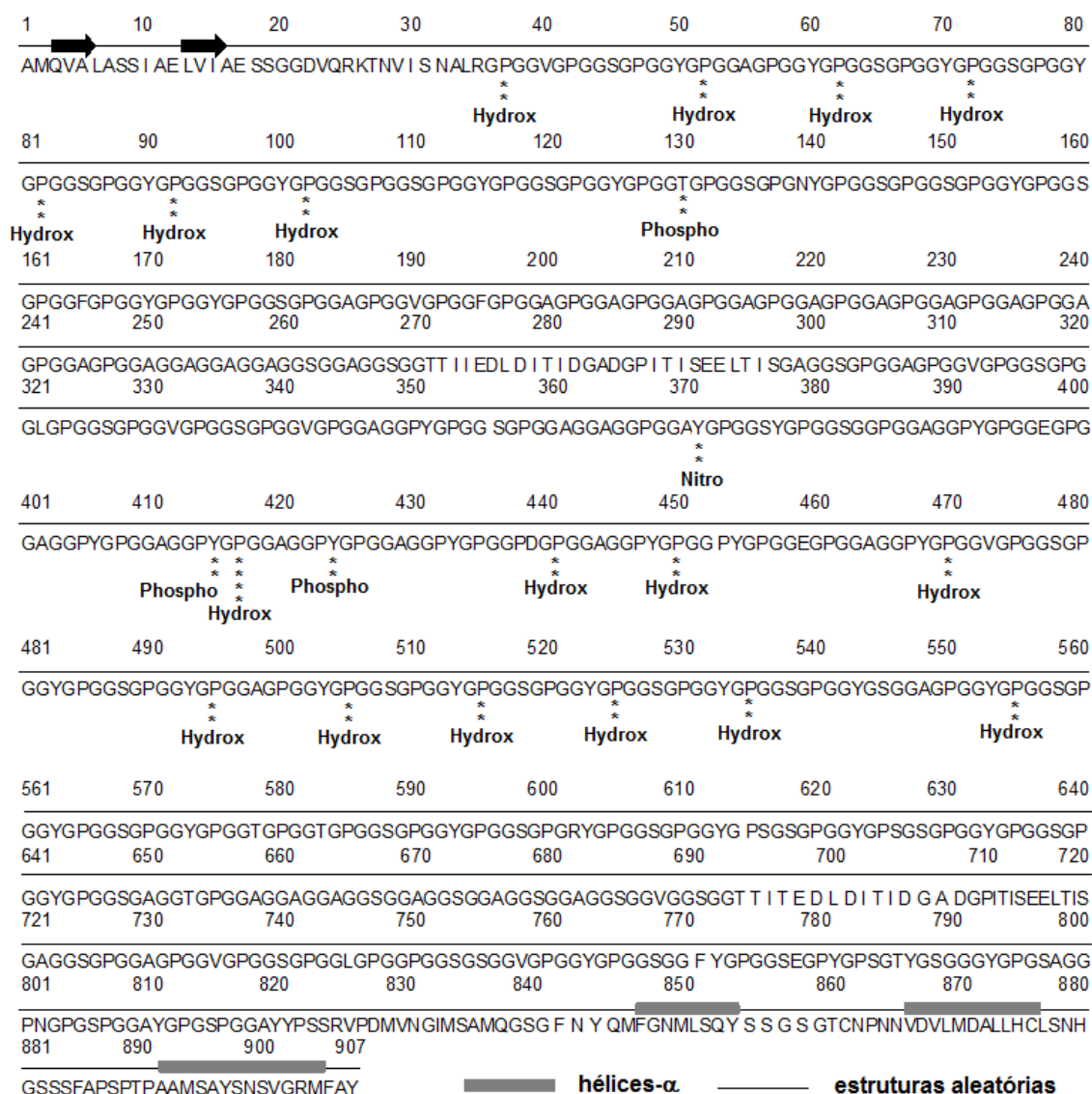


Figura 41. Representação da sequência primária e caracterização da estrutura secundária (JNETPRED *Secondary Structure Prediction*) da proteína da seda flageliforme que foi secretada pela glândula flageliforme e identificada na 2-DE da seda da aranha *N. clavipes*. A sequência foi obtida através da digestão *in-gel* utilizando diferentes enzimas proteolíticas, e mostra a localização das PTMs mais abundantes observadas na proteína.

Para verificar a ocorrência da modificação de fosforilação, os *spots* oriundos da 2-DE foram submetidos ao tratamento com fosfatase alcalina e analisados novamente por MSⁿ. Os espectros mostraram uma diminuição de 80 unidades de massa (Da), confirmando dessa forma a perda do grupo fosfato em alguns fragmentos peptídicos onde a PTM estava presente. Esses resultados estão demonstrados nas figuras 42, 43, 44, 45, 46 e 47.

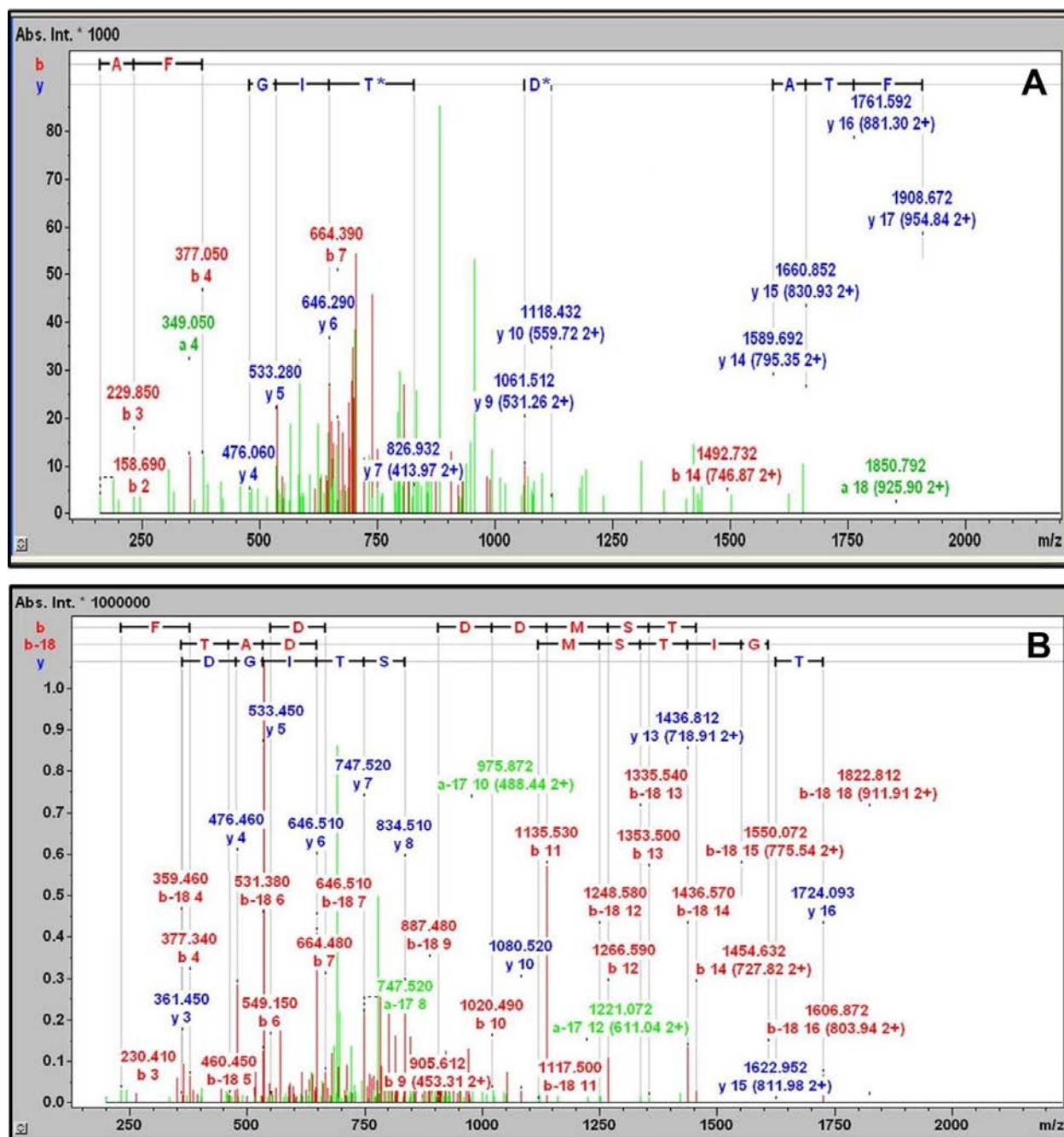


Figura 42. (A) - Espectro CID da sequência peptídica R.TGAFTADQLD**M***ST*IGDTLK.T (49-68), selecionando-se o íon de m/z 713.34 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação T62 e o aminoácido substituído Asp->Gly (D59) observados na proteína espidroína-1 (spot 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica R.TGAFTADQLDD**M**STIGDTLK.T (49-68), selecionando-se o íon de m/z 686.67 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a desfosforilação do sítio T62, que foi observado na proteína espidroína-1 (spot 6), após o tratamento com fosfatase alcalina.

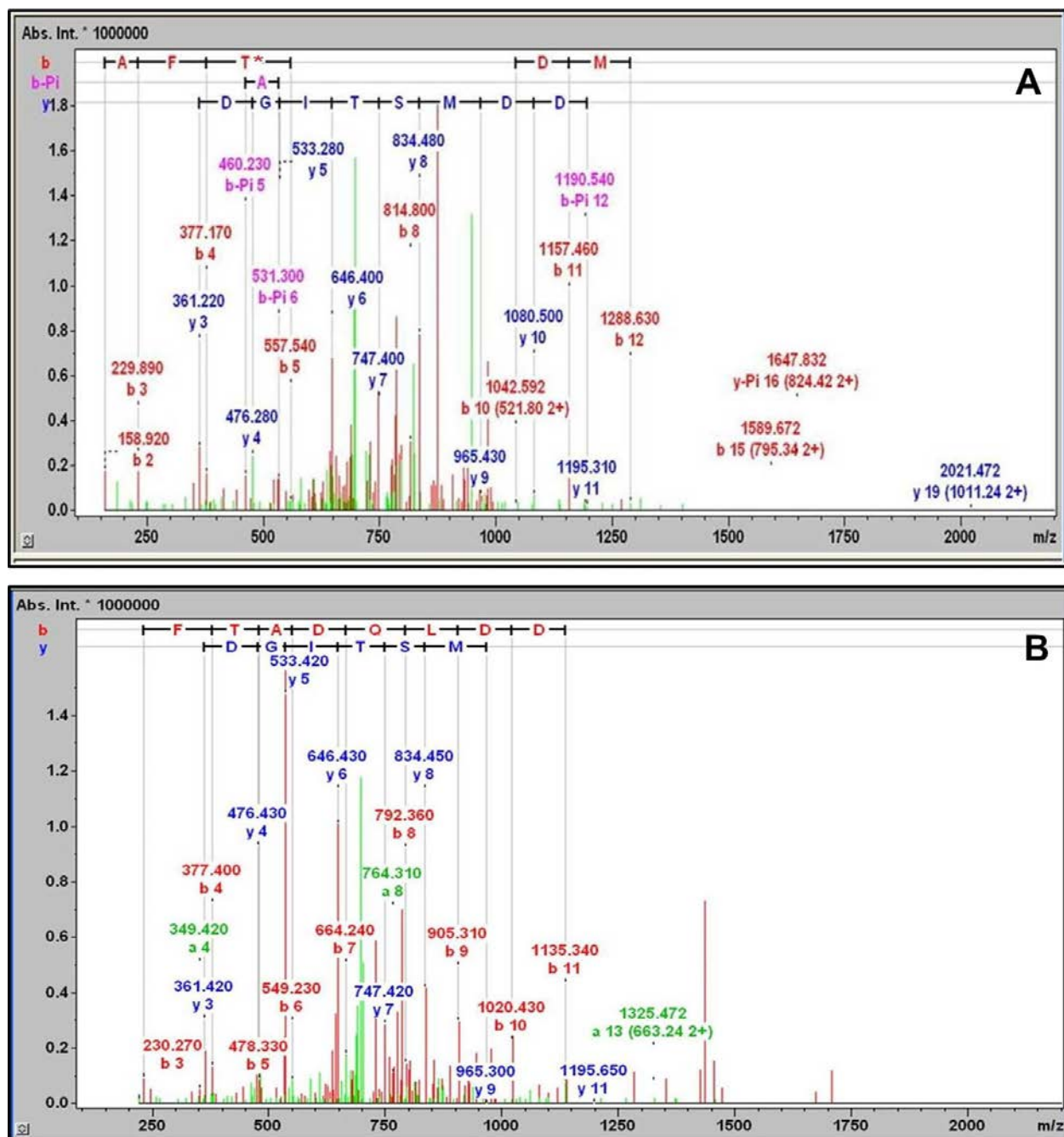


Figura 43. (A) - Espectro CID da sequência peptídica R.T*GAFTAD*QLDDMSTIGDTLK.T (49-68), selecionando-se o íon de m/z 1061.53 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação T49 e o aminoácido substituído Asp→Gly (D55) observados na proteína espiroína-1 (spots 3 e 5), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica R.TGAFTADQLDDMSTIGDTLK.T (49-68), selecionando-se o íon de m/z 1021.53 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando a desfosforilação do sítio T49, que foi observado na proteína espiroína-1 (spots 3 e 5), após o tratamento com fosfatase alcalina.

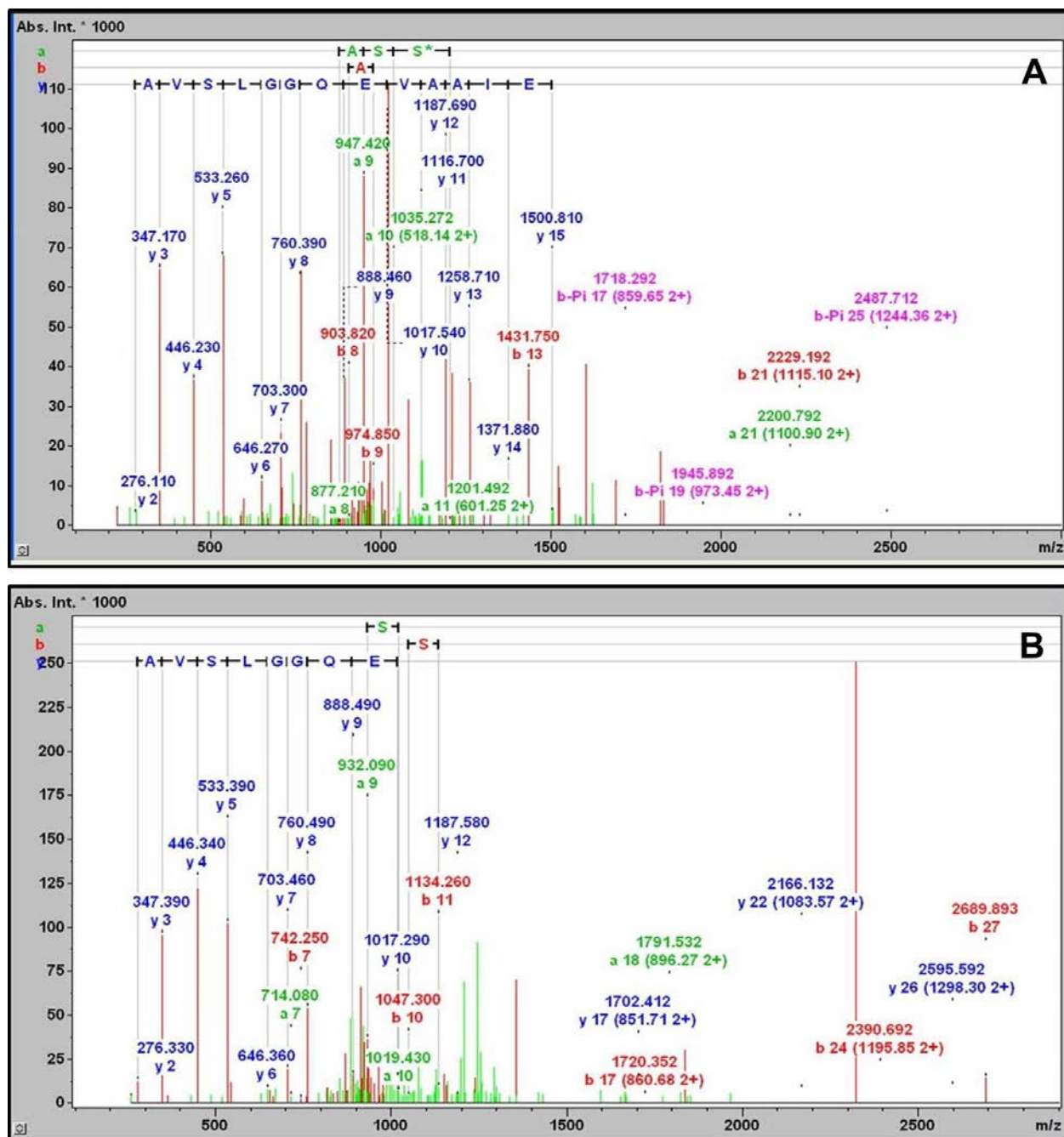


Figura 44. (A) - Espectro CID da sequência peptídica K.LQALNMAFAS^{S*}MAEIAAVEQGGLSVAEK.T (85-112), selecionando-se o íon de m/z 982.50 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação S95 observado na proteína espidroína-1 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica K.LQALNMAFASSMAEIAAVEQGGLSVAEK.T (85-112), selecionando-se o íon de m/z 955.83 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a desfosforilação do sítio S95, que foi observado na proteína espidroína-1 (*spot* 4), após o tratamento com fosfatase alcalina.

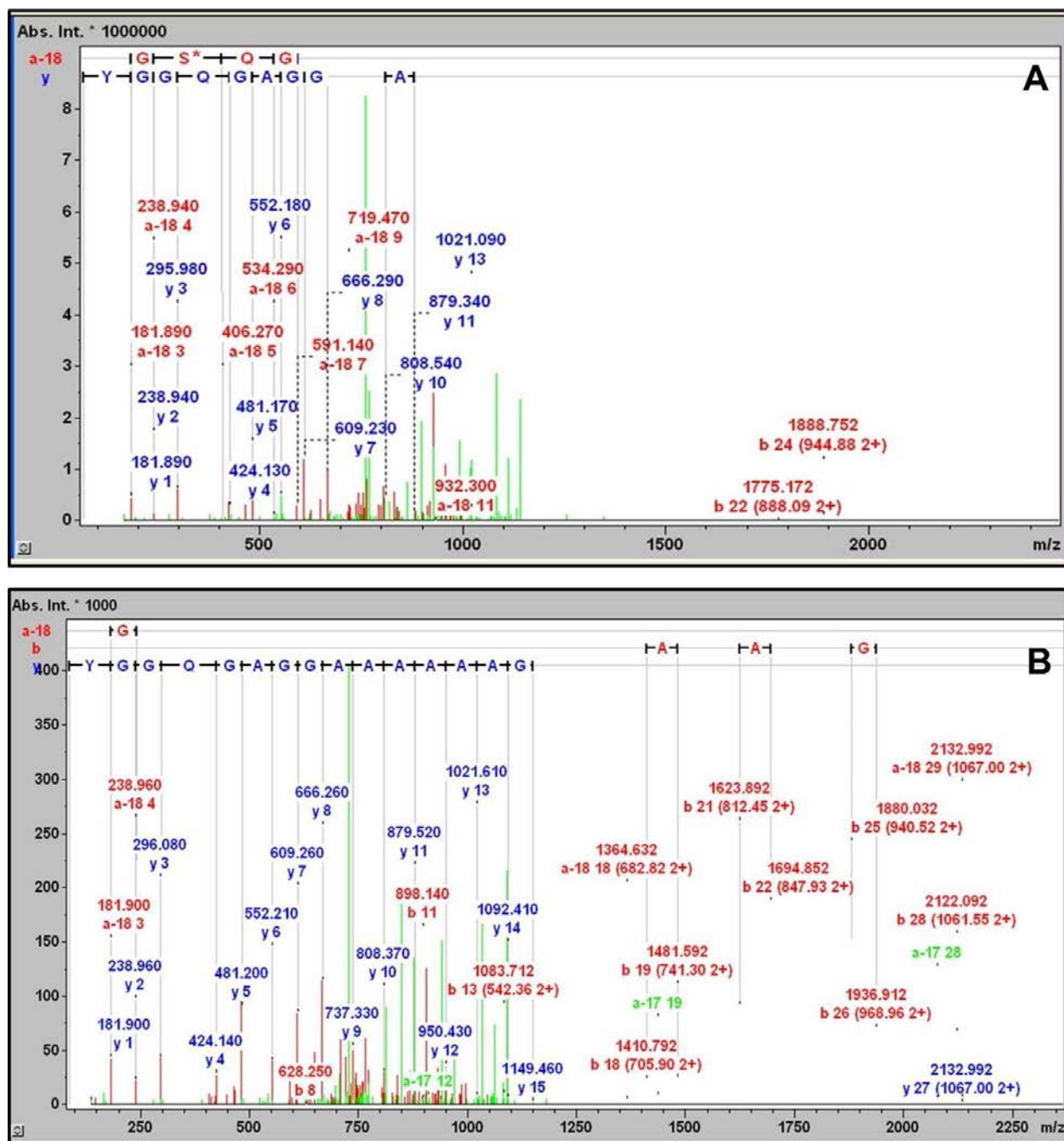


Figura 45. (A) - Espectro CID da sequência peptídica Y.GGLGS*QGAGRGGQGAGAAAAAGGAGQGGY.G (60-89, 576-605), selecionando-se o íon de m/z 821.027 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando os sítios de fosforilação S64 e S580 observados na proteína espidroína-1 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. (B) - Espectro CID adquirido para a mesma proteína, sequência peptídica Y.GGLGSQGAGRGGQGAGAAAAAGGAGQGGY.G (60-89, 576-605), selecionando-se o íon de m/z 794.36 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a desfosforilação dos sítios S64 e S580, que foram observados na proteína espidroína-1 (*spot* 4), após o tratamento com fosfatase alcalina.

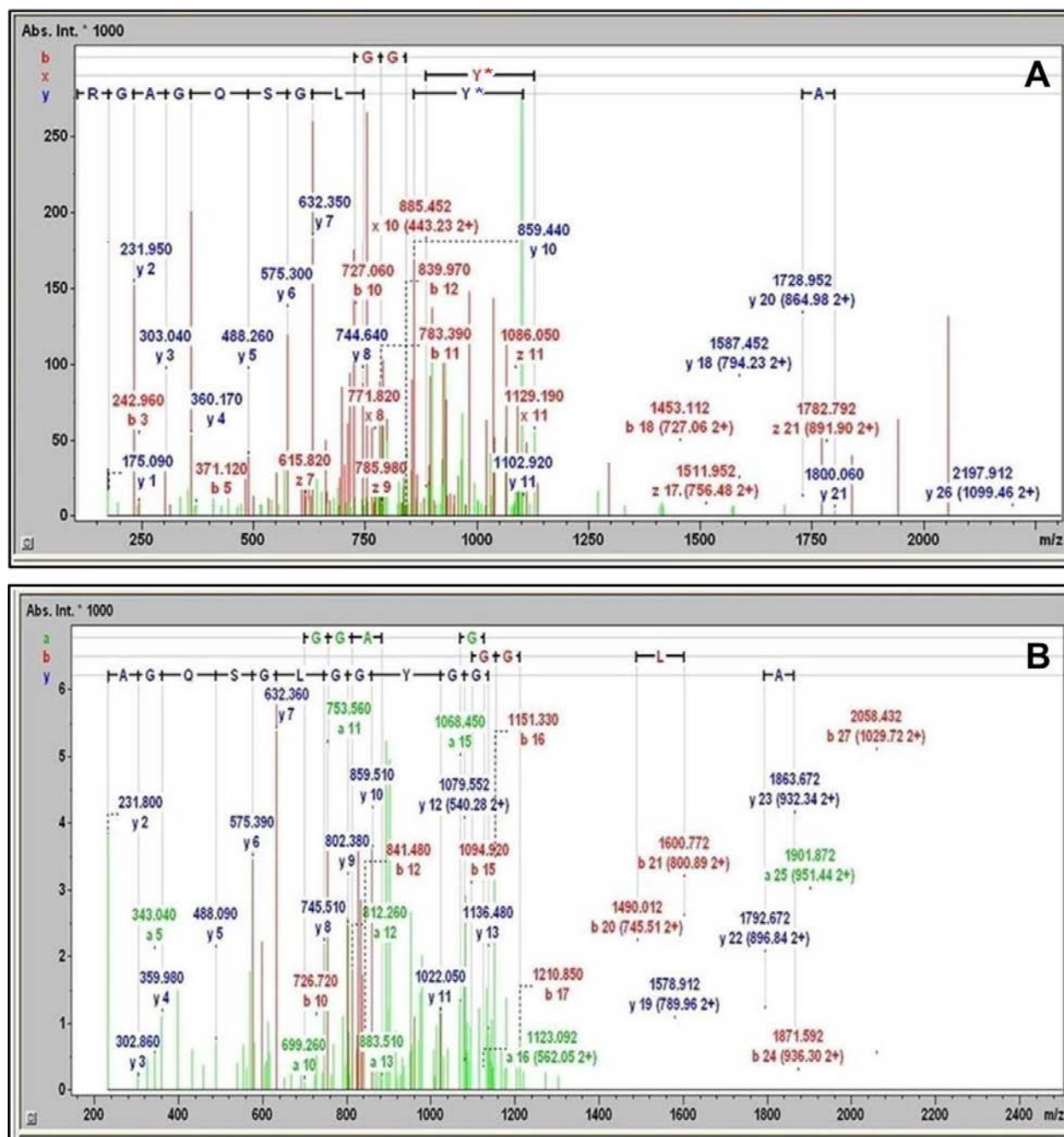


Figura 46. (A) - Espectro CID da sequência peptídica GGQGAAAAAGGAGQGGY^{*}GGLGSQGAGR (134-161), selecionando-se o íon de m/z 811.080 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação Y151 observado na proteína espidroína-1 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica GGQGAAAAAGGAGQGGY^{*}GGLGSQGAGR (134-161), selecionando-se o íon de m/z 784.42 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a desfosforilação do sítio Y151, que foi observado na proteína espidroína-1 (*spot* 3), após o tratamento com fosfatase alcalina.

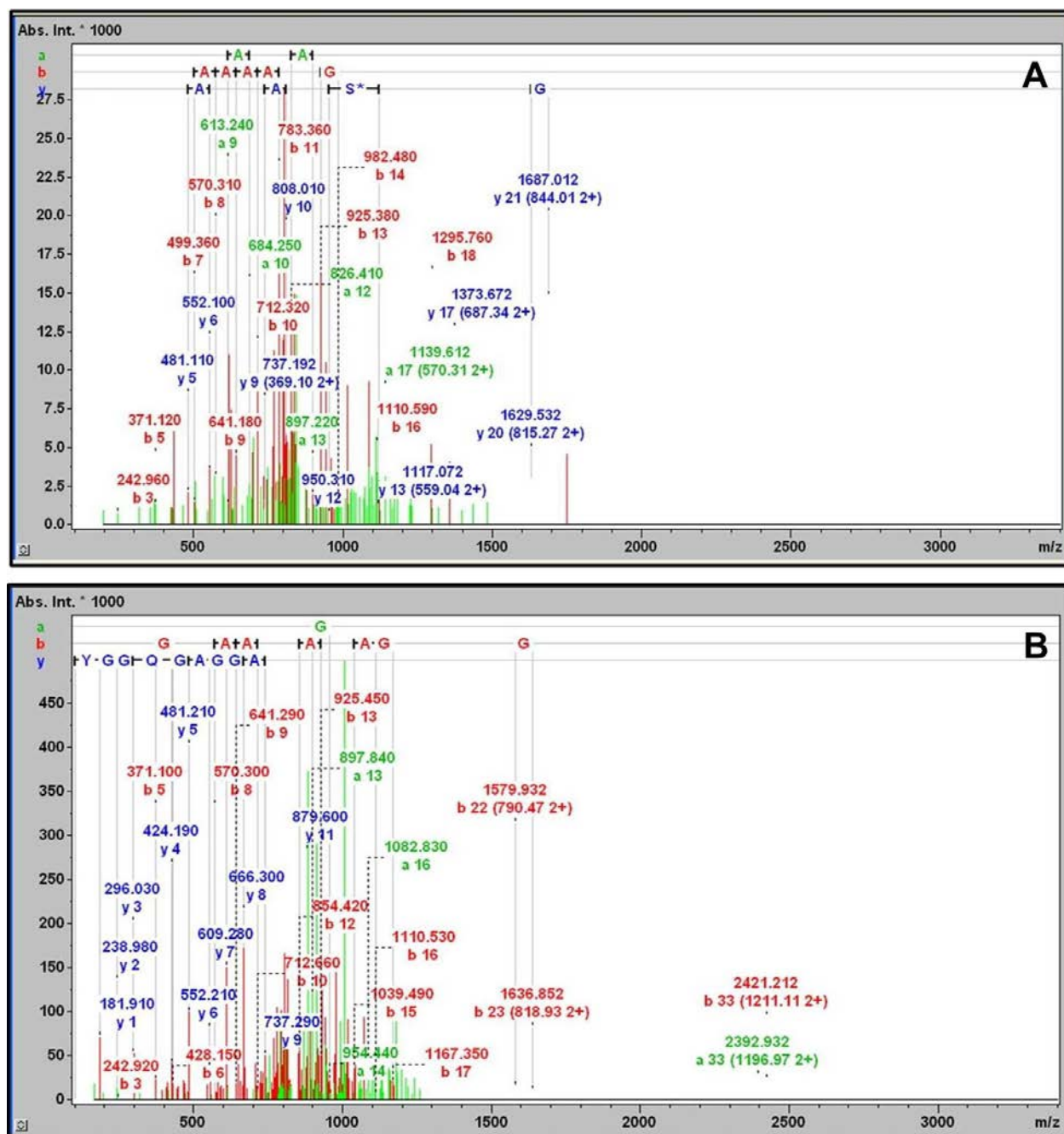


Figura 47. (A) - Espectro CID da sequência peptídica L.GGQGAGAAAAAGGAGQGGLGGQGAGQGAGAS*AAAAGGAGQGGY.G (211-255), selecionando-se o íon de m/z 863.34 na forma de $[M + 4H]^{4+}$, como precursor; e mostrando o sítio de fosforilação S243 observado na proteína espiroína-1 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. (B) - Espectro CID adquirido para a mesma proteína, sequência peptídica L.GGQGAGAAAAAGGAGQGGLGGQGAGQGAGASAAAAGGAGQGGY.G (211-255), selecionando-se o íon de m/z 843.34 na forma de $[M + 4H]^{4+}$, como precursor; e mostrando a desfosforilação do sítio S243, que foi observado na proteína espiroína-1 (spot 3), após o tratamento com fosfatase alcalina.

E ainda para confirmar a ocorrência da fosforilação nas proteínas da seda da teia total, amostras contendo 20-50 μg de proteínas foram submetidas ao ensaio de imunodeteção com anticorpos primários anti-fosfotirosina e anti-fosfoserina. Esse resultado está demonstrado na figura 48.

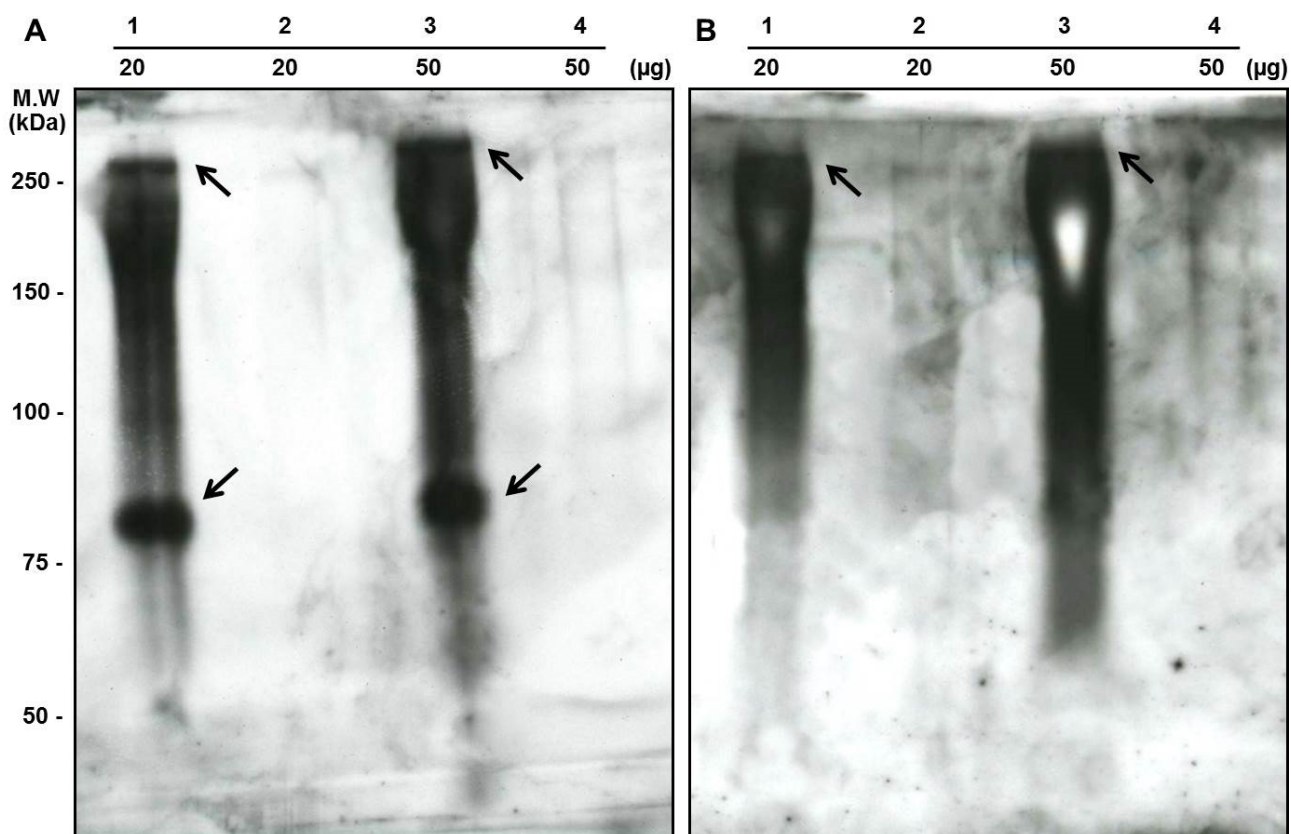


Figura 48. Ensaio de imunodeteção representativo da seda (20-50 μg de proteínas) da teia total da aranha *N. clavipes* mostrando a imunoreatividade dos sítios de fosforilação observados nas proteínas da seda. **A** - Imunodeteção utilizando o anticorpo primário anti-fosfotirosina. Linhas 1 e 3 mostram a imunoreatividade dos sítios de fosfotirosina (setas); linhas 2 e 4 mostram a não imunoreatividade após o tratamento da amostra com fosfatase alcalina. **B** - Imunodeteção utilizando o anticorpo primário anti-fosfoserina. Linhas 1 e 3 mostram a imunoreatividade dos sítios de fosfoserina (setas); linhas 2 e 4 mostram a não imunoreatividade após o tratamento da amostra com fosfatase alcalina.

O resultado obtido para o ensaio de imunodeteção para a fosfotirosina demonstrou a presença de duas bandas, uma banda acima de 250 kDa e outra em torno de 85 kDa, que apresentaram imunoreatividade dos sítios de fosfotirosina presente nas espidroínas constituintes das fibras da seda. Conforme resultados do perfil da eletroforese bidimensional, apresentados anteriormente, podemos verificar que todos os *spots* visualizados apresentam MW acima de 250 kDa, não sendo possível visualizar nenhum *spot* em torno de 85 kDa. A

imunodeteção é um método extremamente sensível que possibilita a detecção da interação proteína-anticorpo, mesmo quando em quantidades mínimas de amostras. Ao contrário da eletroforese bidimensional corada com *Coomassie*, que é um método de detecção de proteínas, menos sensível, sendo necessárias quantidades de amostra acima de 100 ng para que ocorra a detecção da mesma no gel de eletroforese. Sendo assim, possivelmente a banda em torno de 85 kDa visualizada na imunodeteção não estava em quantidades mínima necessária para ser detectada pelo *Coomassie* no gel.

Estudos demonstram que as espidroínas constituintes das fibras da seda apresentam MW em torno de 250-300 kDa após ser expelida pelo ducto de fiação formando as fibras sólidas (Sponner et al., 2005; Ayoub et al., 2007; Hagn et al., 2010); porém, quando armazenadas dentro das glândulas, as espidroínas apresentam MW em torno de 30 kDa (Braunitzer & Wolff, 1955; Bell et al., 1969). No presente estudo as espidroínas foram detectadas com massas moleculares de 26.693 Da para espidroína-1, 16.720 Da para proteína da seda flageliforme e com 15.576 Da e 11.098 Da para a proteína da seda tubuliforme, dentro das respectivas glândulas produtoras de seda. Considerando-se que foi utilizado um coquetel de inibidores de proteases durante toda a preparação das amostras, e que proteases não foram identificadas na análise proteômica das glândulas, pode-se considerar que as espidroínas identificadas nas glândulas são produtos de uma síntese interrompida no momento em que as aranhas foram coletadas; ou ainda que, dentro da glândula as espidroínas ocorreriam na forma de monômeros que formariam ligações covalentes através da região C-terminal formando dímeros (Eisoldt et al., 2011). Sendo assim, ainda permanece indefinido qual seria o real peso molecular das espidroínas dentro das glândulas, e como ocorreria a ligação entre elas formando as espidroínas com elevado pesos moleculares que constituem as fibras da seda. Outra hipótese a ser considerada é de que a síntese de cada domínio pode ocorrer separadamente dentro das glândulas, e que estes domínios são ligados (condensados) uns aos outros ao término de suas sínteses individuais.

Na análise proteômica da seda, as espidroínas constituintes das fibras apresentaram valores de massas moleculares teóricos de 86.77 kDa para a espidroína-1, 69.03 kDa para a espidroína-2 e de 74.33 kDa para a proteína da seda flageliforme. Estudos demonstram que as espidroínas constituintes das fibras da seda apresentam MW em torno de 250-300 kDa após serem expelidas pelo ducto de fiação formando as fibras sólidas (Sponner et al., 2005; Ayoub et al., 2007; Hagn et al., 2010). No entanto, os elevados valores observados para as massas moleculares mostradas na 2-DE (Figuras 36 e 37) e tabela 9 podem ser devido à existência de *cross-links*, ou seja, as espidroínas ocorreriam na sua forma monomérica, constituindo ligações

covalentes que levam à formação de dímeros, trímeros e tetrâmeros para compor as fibras. Considerando o resultado do ensaio de imunodeteção para a fosfotirosina, que demonstrou a presença de uma banda em torno de 85 kDa, que não foi detectada na 2-DE da seda, possivelmente por estar presente em quantidades mínimas nas fibras da seda, já que, nas fibras ocorre a predominância de espidroínas com valores de massas acima de 250 kDa; e considerando que o valor de massa teórico obtido para a espidroína-1 foi de 86.77 kDa, poderia-se sugerir que as espidroínas detectadas por uma abordagem proteômica com valores de massa experimental acima de 250 kDa na seda da aranha *N. clavipes* estariam ocorrendo na forma de trímeros.

E ainda deve-se levar em consideração a ocorrência das PTMs que foram observadas sobre as espidroínas, como por exemplo, a glicosilação, modificação conhecida por reduzir a mobilidade das proteínas no gel; e/ou ainda outras PTMs, como a grande quantidade de fosforilação que foi observada, e que também podem influenciar o peso molecular ou o comportamento das espidroínas no gel.

Para verificar a ocorrência de glicosilação, os *spots* oriundos da 2-DE foram submetidos ao tratamento com PNGase e analisados novamente por MSⁿ. Os espectros mostraram uma diminuição de 162 Da de massa correspondente a PTM do tipo N-glicosilação (N44) hexose, e uma diminuição de 698 Da de massa correspondente a PTM do tipo N-glicosilação (N237) HexNAc2dHex2, confirmando dessa forma a perda dessas moléculas nos fragmentos peptídicos onde a PTM estava presente. Esses resultados estão demonstrados nas figuras 49 e 50.

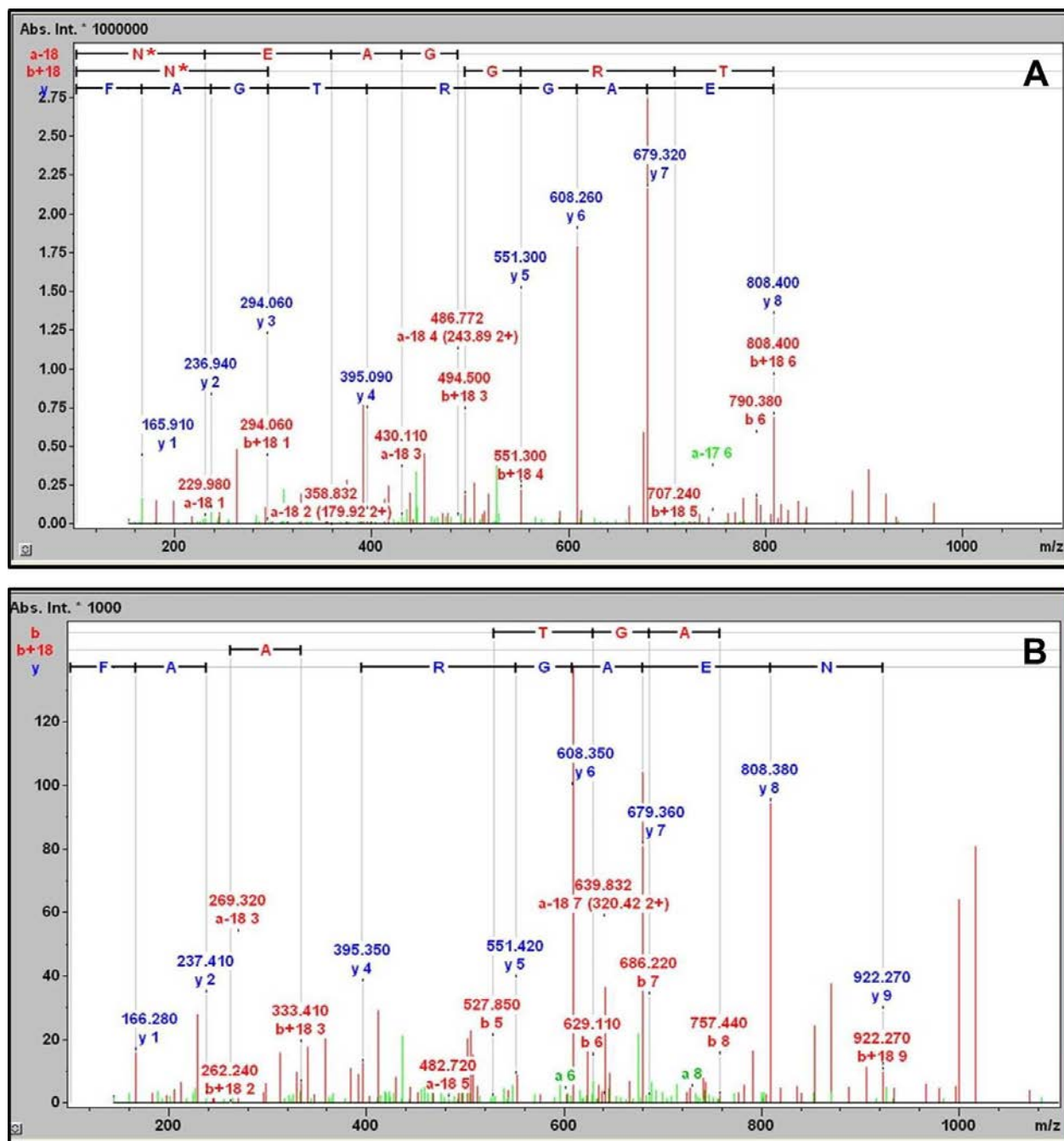


Figura 49. (A) - Espectro CID da sequência peptídica L.N*EAGRTGAF.T (44-52), selecionando-se o íon de m/z 535.22 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de glicosilação N44 (hexose) observado na proteína espidroína-1 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica L.NEAGRTGAF.T (44-52), selecionando-se o íon de m/z 454.72 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando a deglicosilação do sítio N44, que foi observado na proteína espidroína-1 (spot 3), após o tratamento com PNGase.

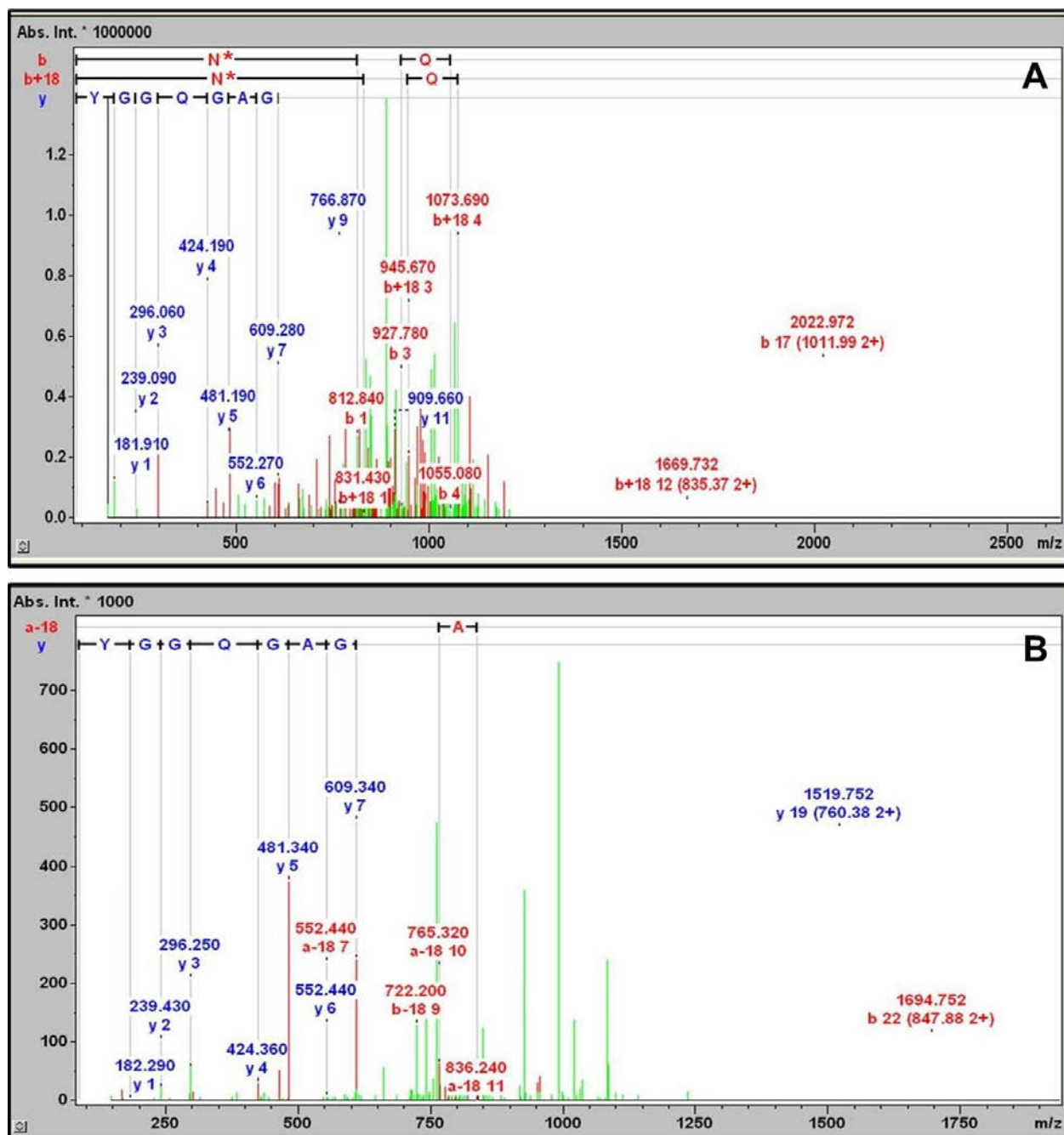


Figura 50. (A) - Espectro CID da sequência peptídica Y.N*GGQGPSAAAAAASGAGQGGY.G (237-260), selecionando-se o íon de m/z 877.88 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de glicosilação N237 (HexNac2DHex2) observado na proteína espidroína-1 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro CID adquirido para a mesma proteína, sequência peptídica Y.NGGQGPSAAAAAASGAGQGGY.G (237-260), selecionando-se o íon de m/z 645.21 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando a deglicosilação do sítio N237, que foi observado na proteína espidroína-1 (*spot* 4), após o tratamento com PNGase.

Para verificar a ocorrência da modificação nitrotirosina nas proteínas da seda da teia total, amostras contendo 20-50 μg de proteínas foram submetidas ao ensaio de imunodeteção com anticorpo primário anti-nitrotirosina. Esse resultado está demonstrado na figura 51.

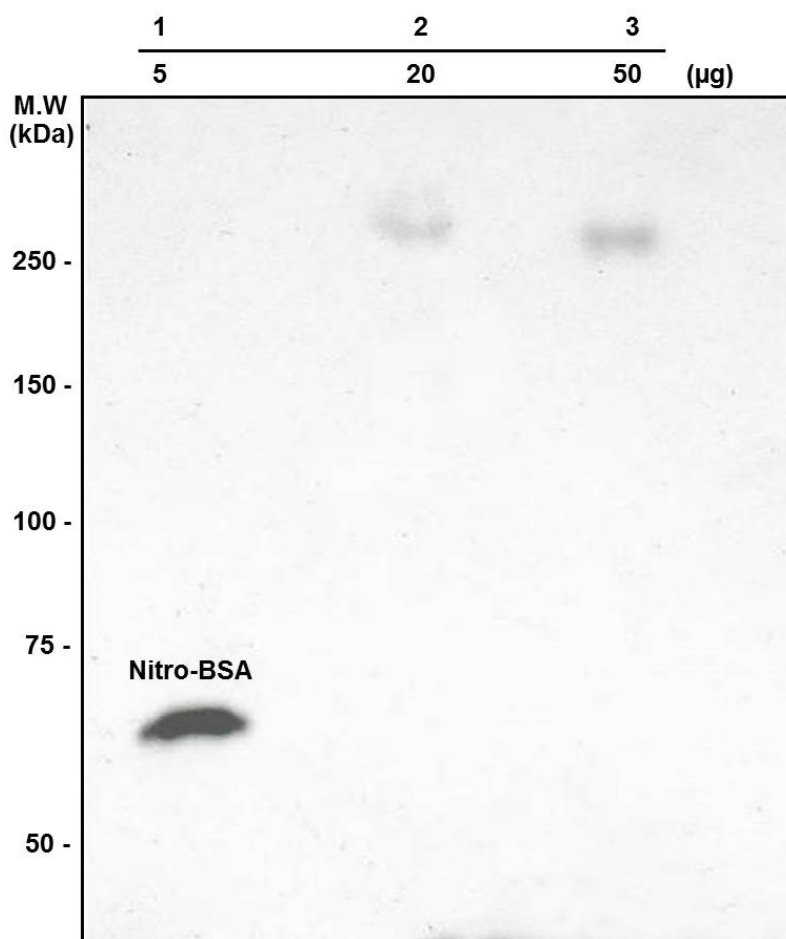


Figura 51. Ensaio de imunodeteção representativo da seda (20-50 μg de proteínas) da teia total da aranha *N. clavipes* mostrando a imunoreatividade do sítio de nitrotirosina observado nas proteínas da seda. Imunodeteção utilizando o anticorpo primário anti-nitrotirosina. Linha 1 mostra a amostra controle positivo (Nitro-BSA - *Nitrated bovine serum albumin*); linhas 2 e 3 mostram a imunoreatividade do sítio de nitrotirosina.

As modificações de fosforilação, glicosilação, nitrotirosina e hidroxilação serão melhores discutidas, quanto as suas presenças e possíveis funções sobre as espidroínas constituintes das fibras da seda da aranha *N. clavipes*, mais adiante neste texto.

6.2.3. Caracterização de PTMs por NanoLC-MSⁿ – abordagem proteômica

Como estão envolvidas na maioria dos processos celulares, incluindo a manutenção da estrutura proteica, as análises das PTMs têm se tornado um grande desafio às pesquisas proteômicas sendo requeridos métodos sensíveis e eficientes para a detecção das PTMs. Porém, a espectrometria de massas (MSⁿ) tem demonstrado ser extremamente útil na identificação das PTMs; e tem sido amplamente utilizada na identificação e caracterização de proteínas (Mann et al., 2003; Witze et al., 2007; Johnson e Evers et al., 2010). Não existe uma estratégia universal para identificar as PTMs, e todas as abordagens utilizadas apresentam aspectos positivos, bem como algumas limitações. No entanto, a caracterização de PTMs por NanoLC-MSⁿ é atualmente viável através da abordagem de espectrometria de massas (MSⁿ), utilizando um ou mais de um método de fragmentação como o CID (dissociação induzida por colisão) e o ETD (dissociação por transferência de elétrons), tanto para a determinação da sequência de aminoácidos, quanto para a determinação do tipo e localização de PTMs. A vantagem de se alternar o uso dos dois métodos de fragmentação torna-se interessante, uma vez que ocorre uma complementariedade dos dados gerados por ambos.

Atualmente estudos têm utilizado a MSⁿ em todas as suas abordagens e/ou a combinação da MSⁿ com outros métodos para a identificação e localização de PTMs nos mais diversos estudos com as mais diversas amostras biológicas (Chen et al., 2010; Redeker, 2010; Britton et al., 2011; Bae et al., 2012; Heo et al., 2012; Resende et al., 2013; Dos Santos-Pinto et al., 2014). No presente estudo adotamos uma estratégia combinando a 2-DE com NanoLC-MSⁿ utilizando-se dois diferentes métodos de fragmentação, CID e ETD, para identificar e sequenciar as proteínas constituintes das fibras da seda da aranha *N. clavipes*; assim como também, identificar e localizar as PTMs presentes nessas proteínas.

As tabelas 14, 15 e 16 mostram o mapeamento dos sítios de fosforilação, hidroxilação, glicosilação e nitrotirosina observados na sequência das proteínas espidroína-1, espidroína-2 e proteína da seda flageliforme, mostrando a posição dessas modificações na sequência e o método de fragmentação pelo qual foram observadas. Podemos notar que a maioria das PTMs foram observadas tanto por CID quanto por ETD, presentes nos fragmentos oriundos da digestão enzimática com tripsina e quimiotripsina, enzimas que apresentaram uma melhor fragmentação das espidroínas possibilitando uma elevada cobertura das sequências obtidas no presente estudo. Além disso, muitas dessas PTMs se repetem em fragmentos peptídicos de diferentes *spots*.

Tabela 14. Mapeamento dos sítios de fosforilação, glicosilação e nitrotirosina observados na sequência da proteína espidroína-1, a qual foi secretada pela glândula ampulada maior e identificada na seda produzida pela aranha *N. clavipes*.

PTMs	Enzima	Método de fragmentação	Anexo dados referentes
Fosforilação - código de acesso B5SYS5 / B5SYS6			
T49	Tripsina/Quimiotripsina	CID/ETD	Tabela S3 (spots 3 e 5)
T53	Tripsina	CID	Tabela S4 (spot 6)
T62	Tripsina	CID	Tabela S4 (spot 6)
S95	Tripsina	CID	Tabela S3 (spots 3 e 4)
S119	Glu-C	CID/ETD	Tabela S3 (spot 5)
T127	Quimiotripsina	CID/ETD	Tabela S4 (spot 4)
S144	Glu-C	CID	Tabela S3 (spots 3 e 4)
S164	Quimiotripsina	CID	Tabelas S3 (spot 6); S4 (spot 6)
S170	Quimiotripsina	CID	Tabela S4 (spot 6)
S195	Quimiotripsina	CID	Tabela S3 (spot 4 e 5)
S211	Quimiotripsina	CID	Tabela S3 (spot 6)
S244	Quimiotripsina	CID	Tabela S3 (spot 6)
Glicosilação - código de acesso B5SYS6			
N44	Quimiotripsina	CID	Tabela S4 (spots 3 e 5)
N237	Quimiotripsina	CID	Tabela S4 (spot 4)
Fosforilação - código de acesso P19837			
Y31	Quimiotripsina	CID	Tabela S3 (spots 3 e 6)
Y59	Tripsina/Quimiotripsina	CID/ETD	Tabelas S1 (spot 2); S3 (spot 6)
S64	Tripsina/Quimiotripsina	CID	Tabelas S1 (spot 3); S3 (spots 3, 4 e 5)
Y89	Tripsina	CID/ETD	Tabela S1 (spot 1)
Y151	Tripsina	CID/ETD	Tabelas S1 (spots 2, 3 e 4); S3 (spot 6); S4 (spot 4)
S156	Quimiotripsina	CID	Tabelas S1 (spot 4); S3 (spots 3, 4, 5 e 6)
Y184	Tripsina/Quimiotripsina	CID	Tabelas S1 (spot 4); S3 (spot 4)
S243	Quimiotripsina	CID/ETD	Tabela S4 (spots 3 e 4)
S260	Quimiotripsina	CID/ETD	Tabelas S1 (spots 1, 3); S3 (spots 4, 5 e 6)
Y353	Tripsina	CID/ETD	Tabelas S1 (spot 2); S3 (spot 6)
Y387	Tripsina	CID/ETD	Tabelas S1 (spot 2); S3 (spot 6)
S462	Quimiotripsina	CID/ETD	Tabela S1 (spot 4)
Y478	Tripsina	CID/ETD	Tabelas S1 (spot 4); S3 (spot 6)
S523	Tripsina	CID	Tabela S3 (spot 4)
Y575	Tripsina	CID	Tabela S3 (spots 4 e 6)
S580	Tripsina/Quimiotripsina	CID	Tabela S3 (spots 3, 4 e 5)
Nitrotirosina - código de acesso P19837			
Y18	Quimiotripsina	CID	Tabela S4 (spots 3 e 4)

Tabela 15. Mapeamento dos sítios de fosforilação observados na sequência da proteína espidroína-2, a qual foi secretada pela glândula ampulada maior e identificada na seda produzida pela aranha *N. clavipes*.

PTMs	Enzima	Método de fragmentação	Anexo dados referentes
Fosforilação - código de acesso P46804			
S22	Quimiotripsina	CID	Tabela S1 (<i>spot</i> 4)
Y61	Proteinase 10	CID	Tabela S1 (<i>spot</i> 4)
S116	Proteinase 10	CID	Tabela S1 (<i>spot</i> 3)
S245	Quimiotripsina	CID	Tabela S1 (<i>spot</i> 3)
S383	Proteinase 10	CID	Tabela S1 (<i>spot</i> 3)
T555	Proteinase 10	CID	Tabela S1 (<i>spot</i> 3)
S601	Proteinase 10	CID/ETD	Tabela S1 (<i>spot</i> 3)
S602	Proteinase 10	CID/ETD	Tabela S1 (<i>spot</i> 3)
S603	Proteinase 10	CID/ETD	Tabela S1 (<i>spot</i> 3)
Y609	Proteinase 10	CID	Tabela S1 (<i>spots</i> 3 e 4)
S613	Proteinase 10	CID/ETD	Tabela S1 (<i>spots</i> 3 e 4)

Tabela 16. Mapeamento dos sítios de hidroxilação, fosforilação e nitrotirosina observados na sequência da proteína da seda flageliforme, a qual foi secretada pela glândula flageliforme e identificada na seda produzida pela aranha *N. clavipes*.

PTMs	Enzima	Método de fragmentação	Anexo dados referentes
Hidroxilação - código de acesso O44359			
P2	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P17	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P27	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P37	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P47	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P57	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P67	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P382	Quimiotripsina	CID	Tabela S4 (<i>spots</i> 1 e 2)
P406	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P415	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P435	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P460	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P470	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P480	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P490	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P500	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P520	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
P530	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 2)
Fosforilação - código de acesso O44359			
T95	Proteinase 10	CID	Tabela S3 (<i>spot</i> 2)
Y380	Quimiotripsina	CID	Tabela S3 (<i>spot</i> 1)
Y389	Quimiotripsina	CID	Tabela S3 (<i>spot</i> 2)
Nitrotirosina - código de acesso O44359			
Y336	Quimiotripsina	CID	Tabela S4 (<i>spot</i> 1)

O método de fragmentação CID mostrou uma melhor fragmentação dos íons de carga 1^+ e 2^+ , enquanto que o ETD mostrou uma melhor fragmentação dos íons de carga 3^+ . No entanto, os dados gerados foram utilizados de forma complementar, sendo possível detectar as PTMs nos fragmentos gerados por ambos, tanto CID quanto ETD. Apesar de o método ETD ser considerado eficiente para a detecção das PTMs, que são lábeis, como a fosforilação e glicosilação pela CID, no presente estudo essas modificações também foram observadas pelo método CID. A fosforilação nos resíduos de tirosina é considerada estável durante a fragmentação por MSⁿ, enquanto que nos resíduos de serina e treonina apresenta uma estabilidade moderadamente estável. Já a glicosilação e a nitrotirosina são consideradas de instáveis a moderadamente estáveis; e a hidroxiprolina é considerada estável (Mann et al., 2003; Wiesner et al., 2008; Johnson e Evers et al., 2010; Drabik et al., 2012). Portanto, todas essas modificações são possíveis de serem detectadas por ambos CID e ETD.

Considerando-se a grande quantidade de espectros que foram gerados na identificação das sequências peptídicas de todas as proteínas oriundas da 2-DE da seda, e para que a redação da tese ocorresse de forma contínua, de modo que a sua leitura não fosse interrompida, todos os resultados de CID/ETD-MSⁿ destacando as PTMs observadas em todas as proteínas do presente estudo estão demonstrados no material suplementar, apêndices B - figuras S1-S63.

Até então, nenhum estudo utilizando a abordagem proteômica sobre as proteínas constituintes da seda de aranha havia sido publicado. Os estudos publicados sobre essas proteínas são principalmente oriundos de dados da engenharia genética e da tecnologia do DNA recombinante (Lewis, 2006; Rising et al., 2007; Lefèvre et al., 2011; Rising et al., 2011); dados sobre a variação polimórfica de resíduos de aminoácidos e PTMs, por exemplo, são limitados. Modificações como fosforilação (Michal et al., 1996) e glicosilação (Weiskopf et al., 1996; Sponner et al., 2007) têm sido relatadas na seda de aranhas, no entanto, até então não se conhecia a exata localização/posição dessas e de outras modificações na sequência das espidroínas. Recentemente publicamos parte dos resultados obtidos com a proteína espidroína-1 da seda (somente o frame sem a teia orbital) da aranha *N. clavipes* do presente estudo, juntamente com resultados obtidos com a seda (somente o frame sem a teia orbital) de outras duas espécies de aranhas, *N. edulis* e *N. madagascariensis*. Nesse estudo descrevemos oito sítios de fosforilação considerados confiáveis sobre a espidroína-1 produzida pela *N. clavipes*,

quatro sobre a espidroína-1 produzida pela *N. madagascariensis*, e dois sobre a espidroína-1 produzida pela *N. edulis* (Dos Santos-Pinto et al., 2014).

No presente estudo, as PTMs mais comumente estudadas, como fosforilação, glicosilação e hidroxilação, foram observadas nas sequências das proteínas constituintes da seda da aranha *N. clavipes*. A fosforilação é uma das PTMs mais comuns em proteínas e consiste na fixação reversível de um grupo fosfato em um resíduo específico de uma proteína-alvo atuando na sinalização e na regulação da atividade proteica (Ribet et al., 2010). Em geral, a introdução de um ou mais grupos fosfato nos resíduos de serina, treonina e tirosina é conhecido por induzir mudanças conformacionais significativas e consequentemente ocasionar efeitos sobre a atividade e interação das proteínas (Eyers et al., 2008), e têm sido relatados em diferentes tipos de seda. Zhang et al. (2006) relataram a presença de sítios de fosforilação na fibroína de cadeia leve e P25, um dos três polipeptídeos componentes da fibroína da seda de *Bombyx Mori*; e em estudos de Chen et al. (2010), também com a seda do casulo de *Bombyx Mori*, foram identificados 12 sítios de fosforilação presentes na fibroína de cadeia pesada. Estudos demonstram a presença de uma grande quantidade de resíduos de fosfato no ducto de fiação sugerindo que a fosforilação ocorre durante o processamento da fibra para ser liberada pela glândula ampulada maior (Michal et al., 1996; Heim et al., 2009). Além disso, foi demonstrado que o alinhamento/dobramento das proteínas é regulado por alterações físicas e bioquímicas que ocorrem ao longo do ducto de fiação, com uma diminuição de pH e concentração de íons Na^+ e aumento na concentração dos íons K^+ e fosfato (Knight et al., 2001; Dicko et al., 2004; Askarieh et al., 2010). Winkler et al. (2000), demonstraram que o alinhamento/dobramento de folhas- β em proteínas semelhantes às proteínas da seda foi controlado pelas reações enzimáticas de fosforilação/desfosforilação. Foi proposto que os resíduos de serina, treonina e tirosina fosforilados desencadeiam a interação entre os motivos hidrofóbicos de polialanina, causando a agregação das espidroínas e outras proteínas da seda (Arakawa et al., 1985; Huemmerich et al., 2004).

Uma observação cuidadosa sobre a localização dos sítios de fosforilação observados na espidroína-1 e espidroína-2 do presente estudo, nos mostra que, a maioria dessas PTMs estão localizadas exatamente dentro dos motivos estruturais que seriam responsáveis pelas propriedades mecânicas da seda; na espidroína-1 a fosforilação foi observada principalmente no motivo GGX, o qual pode ser responsável pelas propriedades elásticas. Já na espidroína-2, a fosforilação foi observada nos motivos GPGXX e $(\text{A})_n$, os quais podem ser responsáveis pelas propriedades elásticas e de resistência da seda, respectivamente. Estudos evidenciam que a estrutura constituída pelo motivo GGX pode formar folhas- β bem como estruturas helicoidais

menos organizadas (Lefèvre et al., 2007; Jenkins et al., 2010; Rising et al., 2011); enquanto que a estrutura constituída pelo motivo GPGXX, devido ao resíduo de prolina, forma claramente voltas- β que quando repetidos produzem uma estrutura em espiral (Hayashi et al., 1999; Eisoldt et al., 2011). O significado biológico da grande quantidade de fosforilação observada nas proteínas da seda da aranha *N. clavipes* do presente estudo, ainda é desconhecido. Porém, corroborando com os dados já descritos na literatura, podemos dizer que a sua presença pode levar a uma determinação da conformação estrutural da proteína; assim como também atuar na ligação e interação entre as proteínas da seda ao influenciar as propriedades físico-químicas da fibra.

A glicosilação é também uma das PTMs mais comuns em certas proteínas secretadas de eucariotos contribuindo para atividades biológicas, solubilidade, estabilidade, interação célula-célula e resistência às proteases (Steinberg et al., 2001). Estudos relatam evidências de glicosilação na proteína fibrohexamerina da seda de *Bombix mori* (Chen et al., 2010); e também há evidências de proteínas glicosiladas presentes no revestimento viscoso da seda da teia das aranhas *Argiope aurantia*, *Argiope trifasciata* e *Araneus diadematus* (Tillinghast, 1981; Vollrath et al., 1990; Vollrath et al., 1991). Análises de microscopia de luz e microscopia eletrônica de varredura (SEM) usando lectinas também mostraram evidências de glicosilação em proteínas da seda da aranha *N. clavipes* (Augsten et al., 2000). No presente estudo foram observados dois sítios de N-glicosilação N44 e N237 na sequência da espidroína-1 da aranha *N. clavipes*. No resíduo Asn44 foi observada uma hexose que acrescenta um valor de 162 unidades de massa (Da); e no resíduo Asn237 foi observada uma HexNAc2dHex2 que acrescenta um valor de 698 unidades de massa (Da) na massa molecular da proteína. No presente estudo, o papel da glicosilação observada nas proteínas da seda da aranha *N. clavipes*, pode estar relacionado com a estrutura conformacional, contribuindo para a resistência às proteases e estabilidade das proteínas.

A hidroxilação, uma PTM predominante em proteínas secretadas, também foi observada no presente estudo. Essa modificação que consiste na adição de grupos hidroxilas aos resíduos de prolina é catalisada pela prolil hidroxilase (PHD), garantindo uma estabilidade proteica (Kaelin, 2005). No presente estudo foi observada uma grande quantidade de hidroxiprolina ao longo da sequência da proteína da seda flageliforme, que constitui a seda da espiral de captura da teia da aranha *N. clavipes*. Essa seda é conhecida por ser altamente elástica, e a observação de uma grande quantidade de resíduos de prolina hidroxilados pode perfeitamente caracterizar essa propriedade elástica apresentada pelas fibras, assim como ocorre no colágeno encontrado no mexilhão *Mytilus edulis* (Qin et al., 1995; Qin et al., 1997). A

hidroxiprolina é um resíduo de aminoácido essencial presente no colágeno, e estudos indicam a presença de domínios estruturais no colágeno de *M. edulis*, semelhantes aos domínios encontrados na seda. E esses domínios podem aumentar a força e extensibilidade do colágeno (Qin et al., 1995; Qin et al., 1997; Waite et al., 1998; Baoyong et al., 2010). No caso da proteína da seda flageliforme os sítios de hidroxiprolina estão localizados no motivo GPGGX, para o qual tem sido proposta a formação de estruturas espiral- β que podem atuar como “nanosprings” moleculares proporcionando elasticidade às fibras (Hu et al., 2006). Estudos evidenciam que a proteína da seda flageliforme não apresenta mudança conformacional durante o processo de fiação; e a grande quantidade de resíduos de prolina distribuídas regularmente ao longo da repetição de sequência GPGGX, evita a formação de estruturas folhas- β cristalinas na fibra (Ohgo et al., 2006; Rousseau et al., 2009).

Um dos principais mecanismos para a hidroxilação pode ser a ação de hidroxilases/oxidases ou radical-hidroxila, devido à exposição ambiental da seda da teia, como por exemplo, à radiação UV. Estudos tem demonstrado que a oxidação de proteínas pode conduzir a fragmentação da cadeia polipeptídica, assim como a formação de *cross-linkages* entre as proteínas (Lubec, 1996; Lubec et al., 1998; Stadtman & Levine, 2003; Stadtman, 2006). No presente estudo foram observadas algumas modificações como dihidroxilação da metionina (sulphone), oxidação da tirosina (aminotirosina), oxidação da tirosina (nitrotirosina), dihidroxilação da prolina e hidroxilação do ácido aspártico. Chen et al. (2010) também relatou a presença da hidroxilação do ácido aspártico na fibroína de cadeia leve, presente na seda do casulo de *Bombyx Mori*, porém a função dessa modificação permanece indefinida.

A oxidação da tirosina, por exemplo, pode favorecer a formação de *cross-linkages* pela combinação dos radicais tirosina, tyr-tyr, entre duas proteínas diferentes (Pryor et al., 1993; Stadtman, 1993; Stadtman, 2006). Estudos ainda indicam que a nitração do resíduo de tirosina pode estar relacionado com o mecanismo de fosforilação/desfosforilação do grupo hidroxila deste resíduo, em algumas proteínas ao inibir a fosforilação (Kong et al., 1996; Stadtman & Levine, 2003), modificação já discutida anteriormente. A relevância funcional da aminotirosina e nitrotirosina observadas nas espidroínas do presente estudo permanece indefinida, porém estudos sobre as consequências biológicas de proteínas nitradas revelaram efeitos da oxidação da tirosina sobre as propriedades mecânico-elásticas, catabolismo (Parastatidis et al., 2008), e função das proteínas (Peluffo et al., 2007). Estudos de Bae et al. (2012) com a proteína *bilin-binding* presente na asa da borboleta *Hebomoia glaucippe* revelou a presença de nitrotirosina, sendo proposta uma atuação dessa modificação nas propriedades mecânico-elásticas e interação proteína-proteína.

O resultado das análises nos hidrolisados da seda da aranha *N. clavipes* demonstraram a presença de ditirosina (Figura 52). A ditirosina observada na amostra da seda pode representar um processo de oxidação, embora permaneça incerto se a oxidação foi introduzida por peroxidases em algum momento durante a produção/armazenamento das espidroínas na glândula; ou após a formação e liberação das fibras e exposição à radiação UV (Pouchkina et al., 2003).

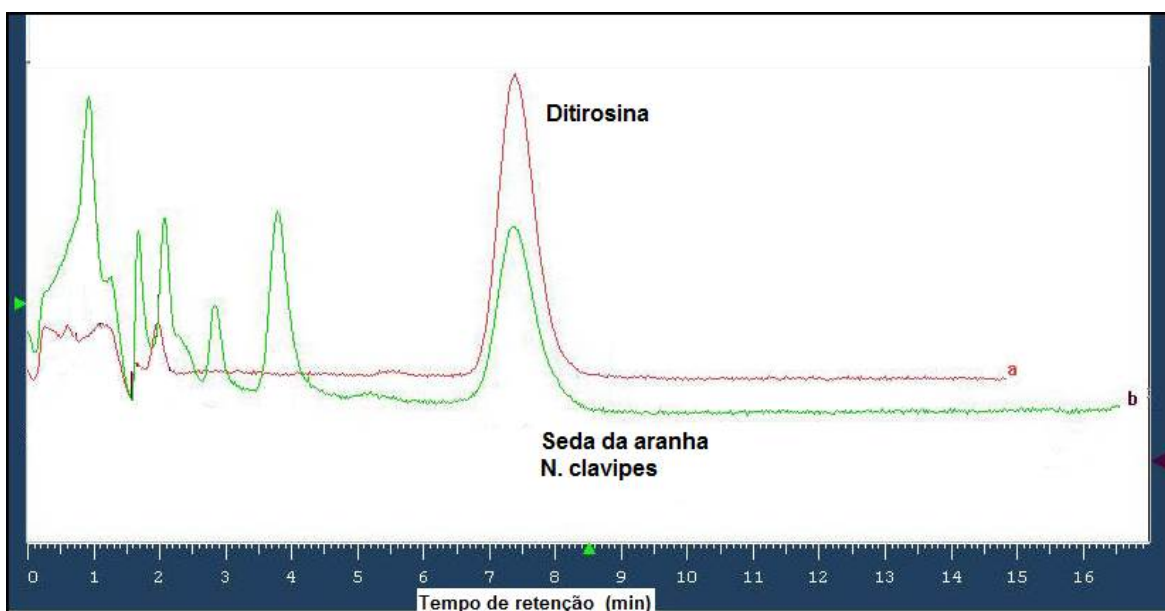


Figura 52. Análises de ditirosina. **Linha (a)** - padrão de ditirosina, **Linha (b)** - hidrolisado da amostra da seda da aranha *N. clavipes*. Procedimento realizado: Primeira dimensão - coluna Purospher Star RP18, 5 μ m, 250 x 4 mm; eluente A - 0.1% de HCOOH + NH₃, pH 3.0, eluente B - metanol, gradiente linear de 0 a 20 min, compreendido entre 95% A + 5% de B até 75% de A + 25% de B; fluxo de 0.8 mL/min, volume de injeção de 40 μ L; intervalo de troca 7.7-9.1 min, tempo de retenção da ditirosina 8.4 min. Segunda dimensão - "coluna Nucleosil strong acid cation exchanger SA", 125 x 4 mm, 10 μ m tamanho de partícula; eluente A - 0.1% de HCOOH, eluente B - 0.1 M de HCOOH + NH₃, pH 3.7; A + B = 60+40 (v/v), fluxo 1.0 mL/min, 285/410 nm.

As sequências das espidroínas apresentam cerca de 3% de resíduos de tirosina (Dicko et al., 2004); e foi sugerido que esses resíduos podem atuar na formação de ligações cruzadas covalentes de ditirosina, para fortificar as fibras da seda. Quando a atividade de peroxidases foi detectada na glândula ampulada maior das aranhas do gênero *Nephila*, como já foi discutido anteriormente neste texto, foi proposto que a atividade das peroxidases poderia catalisar a formação de ligações cruzadas de ditirosina, e dessa forma ajudar na estabilização das

proteínas das fibras da seda. No entanto, até então essa formação nas fibras da seda não tem sido relatado, permanecendo por ser elucidada (Vollrath e Knight, 1999; Pouchkina et al., 2003).

6.2.4. Caracterização estrutural 3D das espidroínas

Devido aos resultados obtidos no presente estudo com a identificação, sequenciamento e localização das PTMs sobre as sequências primárias das espidroínas constituintes das fibras da seda, decidiu-se realizar a caracterização da estrutura 3D das espidroínas, através da bioinformática estrutural, para se analisar a conformação das estruturas dos motivos responsáveis pelas propriedades mecânicas das fibras com a presença das PTMs. Sendo assim, as sequências das espidroínas foram submetidas a uma busca contra o banco de dados de estruturas 3D de proteínas determinadas experimentalmente (*Protein Data Bank* - <http://www.rcsb.org/pdb/>), para a busca e seleção de moldes (*templates*) que fossem compatíveis com as espidroínas; só não foi possível encontrar moldes compatíveis com a proteína da seda flageliforme. Os moldes escolhidos foram aqueles que apresentaram uma identidade elevada, não sendo inferior a 30% com as espidroínas. Os códigos PDB dos moldes escolhidos estão mostrados na metodologia nas tabelas 3 e 4. Os valores dos parâmetros de cortes mostrados, conforme metodologia descrita anteriormente demonstram que os moldes escolhidos podem ser perfeitamente utilizados para a realização da modelagem das espidroínas. A confiabilidade da modelagem molecular por homologia de proteínas está relacionada à porcentagem de identidade na qual o modelo é baseado, estabelecendo-se uma correlação entre a similaridade sequencial e estrutural das duas proteínas (Koehl, 1999; Marti-Renom et al., 2000).

As imagens das estruturas 3D dos modelos moleculares das espidroínas constituintes das fibras da seda da aranha *N. clavipes* foram geradas a partir do programa PyMol (Delano, 2002) e estão ilustradas na figura 53. Posteriormente, as modificações pós-traducionais de fosforilação foram inseridas nos modelos das proteínas espidroínas-1A e -1B através do programa PyMol; e todos os modelos foram submetidos à simulação de dinâmica molecular.

Pela primeira vez, na literatura, se apresenta um modelo 3D para as espidroínas constituintes das fibras da seda de aranhas, constando das regiões N- e C-terminal e da região central repetitiva das sequências primárias dessas proteínas da seda da teia da aranha *N. clavipes*. Até então, somente foram descritas partes de estruturas 3D das regiões N- e C-terminal da espidroína-1 (Askarieh et al., 2010; Hagn et al., 2010; Hagn, 2012).

Os resultados obtidos para os modelos 3D das espidroínas constituintes das fibras da seda da aranha *N. clavipes* corroboram com outros estudos, que indicam que as espidroínas

produzidas pela glândula ampulada maior, espidroínas-1 e -2 apresentam uma conformação estrutural altamente orientada constituída pelas estruturas de folhas- β , 3_{10} -helicoidal, hélice- α e estruturas aleatórias (randômicas) nas fibras sólidas da seda, após serem secretadas durante o processo de fiação (Römer et al., 2008; Lefèvre et al., 2011; Eisoldt et al., 2011).

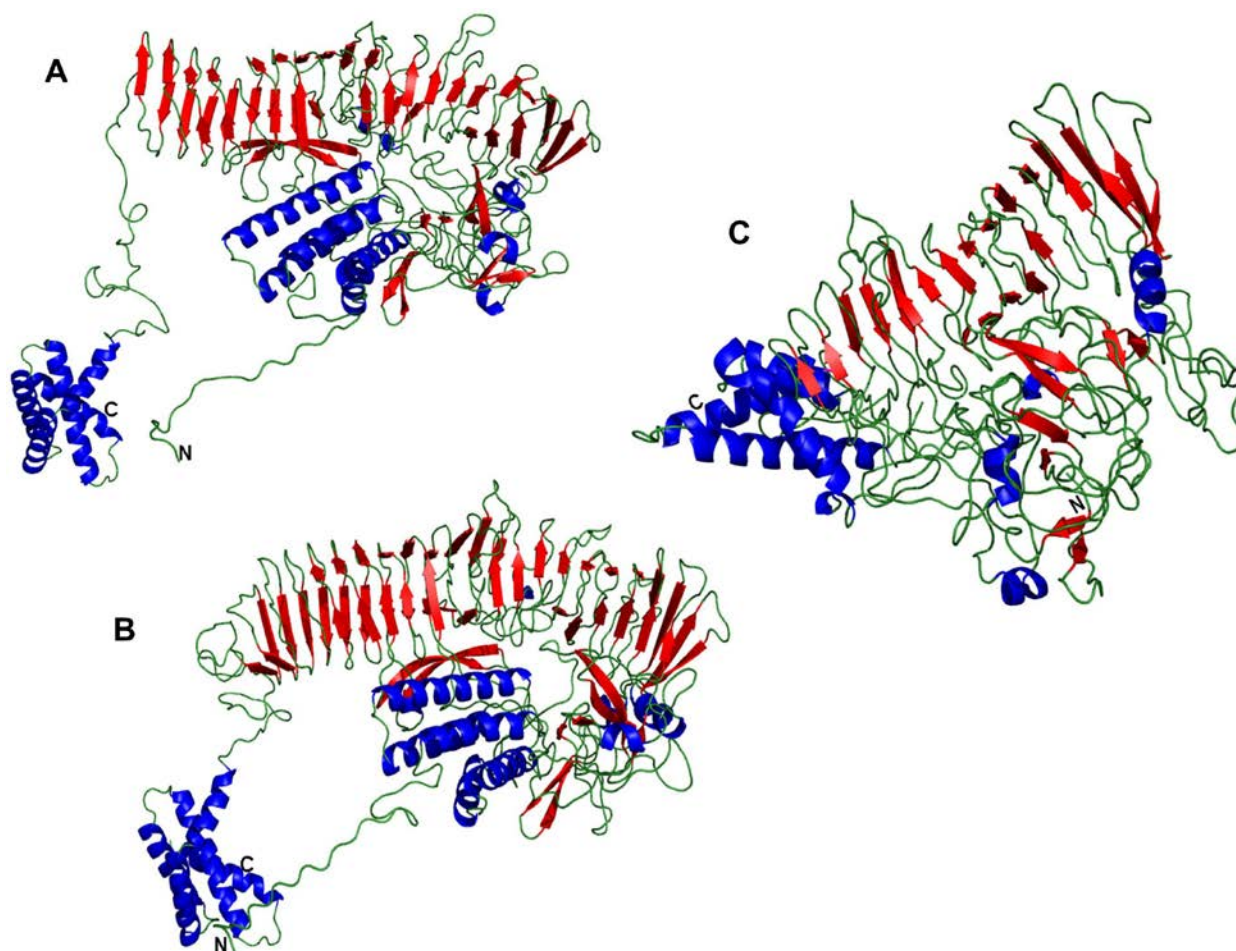


Figura 53. Modelos moleculares das estruturas 3D das espidroínas constituintes das fibras da seda da aranha *N. clavipes*. **A** - espidroína-1A; **B** - espidroína-1B; **C** - espidroína-2.

Nos modelos obtidos no presente estudo pode-se constatar a presença de superestruturas denominadas de grampos de folhas- β . O grampo de folhas- β é caracterizado pela presença de fitas- β anti-paralelas próximas, formando um trecho de folhas- β , o qual tem sido relacionado como uma estrutura importante para a estabilidade e nos dobramentos de proteínas (Ramirez-Alvarado et al., 1997; Simpson et al., 2005; Trevino et al., 2007; Madan et al., 2014). Representa um excelente sistema modelo para estudar a relação entre a sequência

e a estrutura secundária, permitindo uma abordagem hierárquica para um melhor entendimento sobre os fatores que influenciam o dobramento e a estabilidade de proteínas (Kiehna et al., 2003). Entre esses fatores estão as “voltas” localizadas entre as estruturas, as interações cruzadas entre as fitas, através das cadeias laterais, e a presença de regiões hidrofóbicas; assim como ocorre com a região central repetitiva das espidroínas, essa região é tipicamente necessária para a estabilidade das proteínas que apresentam grampos de folhas- β (Syud et al., 1999; Kobayashi et al., 2000; Espinosa & Gellman, 2000; Tatko & Waters, 2002; Syud et al., 2001).

O potencial eletrostático para cada modelo foi gerado a partir do programa PyMol e está demonstrado na figura 54. Quando se compara a superfície eletrostática dos modelos, pode-se verificar que a espidroína-1A apresenta uma superfície um pouco mais eletropositiva (regiões em azul) que a espidroína-1B; enquanto que a espidroína-2 apresenta uma distribuição de cargas totalmente diferente das outras duas espidroínas.

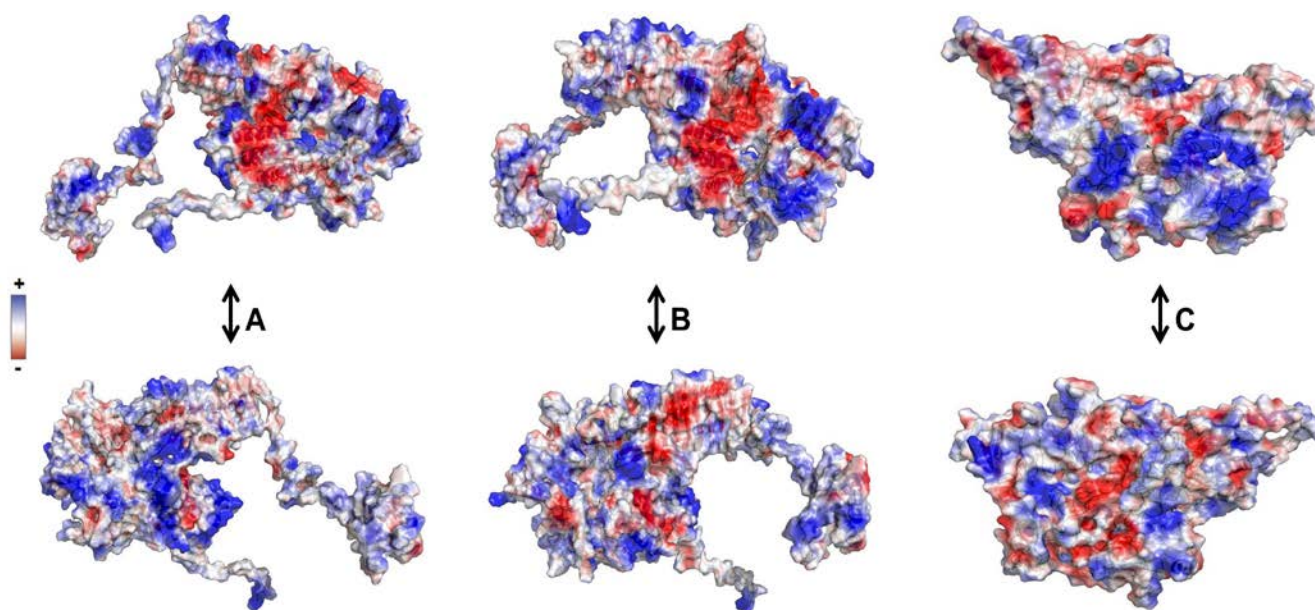


Figura 54. Potencial eletrostático para os modelos das espidroínas gerados a partir do programa PyMol. **A** - espidroína-1A; **B** - espidroína-1B; **C** - espidroína-2. Imagens localizadas na parte superior da imagem estão em 0° (frente) e as imagens localizadas na parte inferior sofreram uma rotação de 180° (verso).

Avaliar a qualidade do modelo é um quesito fundamental, pois ela será proporcional à qualidade da estrutura tridimensional utilizada como molde. Por isso, deve-se avaliar o modelo, como um todo, e, por região. Diversas metodologias existem para a avaliação da qualidade de uma estrutura, seja ela determinada experimentalmente ou por modelagem por homologia. Após a escolha do melhor modelo, análises mais criteriosas e detalhadas foram efetuadas com a utilização do PROCHECK (diagrama de Ramachandran e G-factor), para verificar a qualidade do modelo construído. O programa PROCHECK foi utilizado para analisar a geometria global dos modelos através do diagrama de Ramachandran e o G-factor. Segundo Ramachandran, uma estrutura de boa qualidade deve ter 90% ou mais de seus resíduos nas regiões mais favoráveis. As áreas em branco correspondem a regiões onde existem os choques estereoquímicos na proteína. Estas regiões não permitidas estereoquimicamente valem para todos os aminoácidos, exceto a glicina (Gly) que é a única que não tem cadeia lateral e pode adotar os ângulos de torção ϕ e ψ em todos os quadrantes do diagrama de Ramachandran, são representadas por triângulos no gráfico. As áreas vermelhas correspondem à conformação onde não há nenhum choque estereoquímico, ou seja, são as regiões permitidas e aonde se encontram as hélices- α e folhas- β . As áreas em amarelo escuro e claro mostram regiões limitantes, porém ainda ângulos permitidos ou generosamente permitidos, respectivamente.

No presente estudo, a análise dos diagramas de Ramachandran dos modelos das espidroínas, mostrados na figura 55 e tabela 17, demonstra que mais de 90% dos resíduos de aminoácidos estão localizados sobre as regiões favoráveis do gráfico. Portanto, os modelos gerados para as espidroínas apresentam uma boa qualidade estereoquímica, levando em consideração que essas proteínas apresentam uma grande quantidade de resíduos de aminoácidos, e ainda composta em sua maior proporção por estruturas aleatórias. Os resíduos que estão na região não favorecida estão localizados em sua maior parte, na região de estruturas aleatórias e alguns poucos em região de superfície da proteína em contato com o meio (solvente), por isso não comprometem a qualidade do modelo, pois resíduos de superfície (expostos ao solvente) e em regiões de estruturas aleatórias geralmente apresentam uma maior flexibilidade de rotação do que os resíduos localizados mais internamente na proteína.

O G-factor médio dos ângulos de torção e geometria covalente representa uma medida do desvio de uma dada propriedade estereoquímica para cada resíduo de aminoácido da estrutura. O valor ideal deste fator deve ser acima de -0.5; valores no intervalo de -0.5 a -1.0 significam que existem resíduos na proteína que devem ser investigados e, valores abaixo de -1.0 indicam que os modelos apresentam qualidade estereoquímica ruim e devem ser investigados. O G-factor global é a melhor medida da qualidade estereoquímica total de uma

molécula, pois representa a média de todos os G-factor obtidos para cada aminoácido da estrutura. A análise do G-factor mostrou que os modelos estão dentro dos intervalos de confiança que validam a estrutura 3D; sendo assim, pode-se afirmar que os modelos apresentam uma boa qualidade quanto ao comprimento de ângulos e sua geometria. As análises utilizando o programa PROCHECK (diagrama de Ramachandran e G-factor), estão demonstradas na tabela 17.

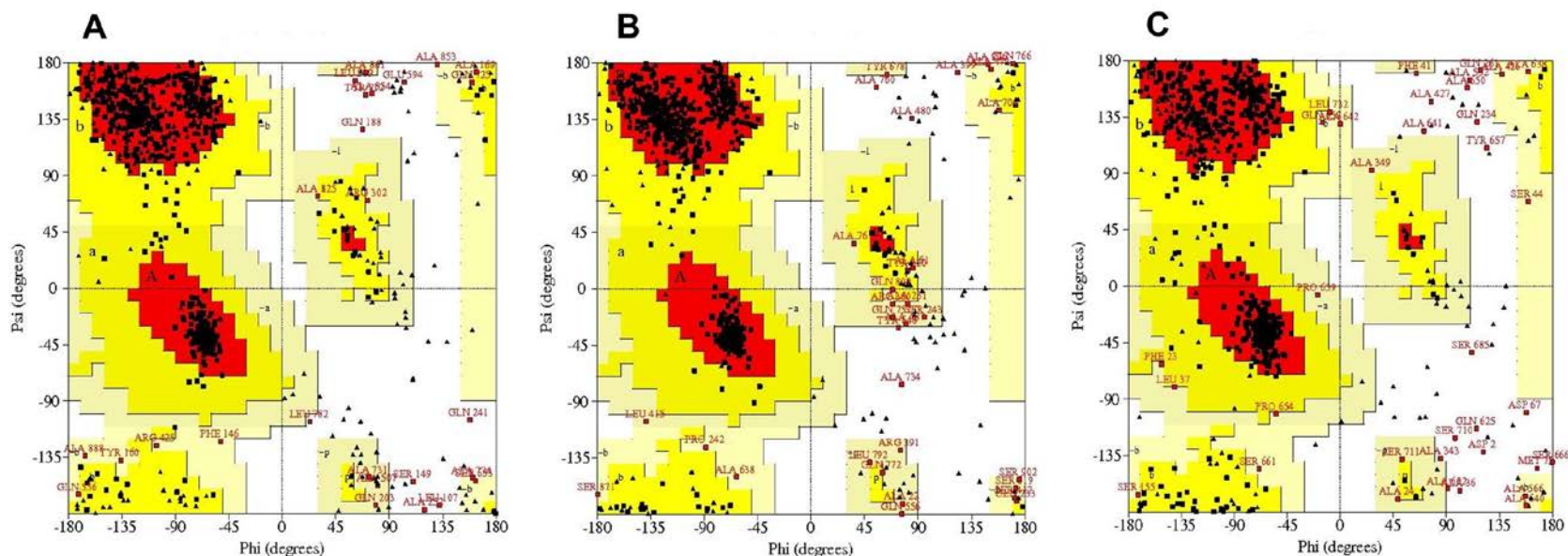


Figura 55. Avaliação da qualidade das estruturas 3D dos modelos das espidroínas constituintes das fibras da seda da aranha. Diagrama de Ramachandran para as proteínas da seda da teia de *N. clavipes*. **A** - espidroína-1A; **B** - espidroína-1B; **C** - espidroína-2.

Tabela 17. Análises para os modelos moleculares das espidroínas utilizando o programa PROCHECK.

Modelos 3D	Regiões do diagrama de Ramachandran ϕ - ψ (%)				G-factor		
	Mais Favorecida	Adicionalmente Favorecida	Generosamente Favorecida	Não Favorecida	Ângulos de torção	Geometria covalente	G-factor Global
Espidroína-1A	85.9	10.0	2.5	1.6	0.30	-0.49	-0.18
Espidroína-1B	87.5	7.5	4.1	0.8	-0.21	-0.56	-0.34
Espidroína-2	77.8	15.1	3.2	3.8	-0.38	-0.52	-0.48

Estudos relatam que as fibras da seda de aranhas são revestidas com uma camada lipídica que é aderida sobre a fibra, após a sua formação, pela aranha (Vollrath et al., 2001; Salles et al., 2006; Sponner et al., 2007). Os lipídeos que recobrem os fios das teias, principalmente os ácidos graxos, constituem um fator importante para consolidar a estratégia de captura utilizada pelas aranhas construtoras de teias aéreas como a *N. clavipes* (Schulz, 2001). Assim como outras substâncias, os lipídeos encontram-se na solução viscosa a qual é secretada pela glândula agregada e aderida sobre a fibra (Vollrath & Knight, 2001; Schulz, 2001; Romer et al., 2008). E análises de lipídeos sobre essas fibras tem indicado a presença de uma grande variedade de lipídeos. Estudos realizados com o uso de análises GC-MS realizados por Salles et al. (2006) com as fibras da teia da aranha *N. clavipes* demonstraram a presença de uma diversidade de ácidos graxos saturados como o ácido decanóico, ácido dodecanóico, ácido tridecanóico, ácido tetradecanóico, ácido octadecanóico e ácido icosanóico; e também de ácidos graxos insaturados como o ácido (Z)-tetradec-9-enoic. Alguns desses lipídeos também já foram relatados previamente em outros estudos sobre os lipídeos recobrindo as fibras da teia (Prouvost et al., 1999; Schulz, 2001). Nas teias de *N. clavipes* a composição de lipídeos revelou que 85% dos mesmos eram compostos de uma mistura de ácido octadecanóico e ácido tridecanóico na proporção 1:1 (Salles et al., 2006). A função dessa camada de ácidos graxos sobre as fibras da seda ainda não está bem esclarecida, porém muitos dos ácidos graxos relatados podem estar envolvidos com a estrutura central da fibra, proporcionar uma superfície suave, regulação do conteúdo de água, repelindo-a da seda, e proteção da seda contra a degradação (Schulz, 2001; Salles et al., 2006). E estudos de Microscopia Eletrônica de Transmissão (MET) com técnica de citoquímica para detecção de lipídeos demonstraram que tanto na superfície quanto no interior das fibras se encontra uma camada de ácidos graxos (Salles, 2003).

Devido aos resultados já descritos na literatura sobre essa camada de ácidos graxos presente nas fibras da seda, e devido aos resultados obtidos no presente estudo com a identificação das proteínas da glândula agregada, a qual secreta uma solução viscosa que é depositada sobre as fibras; a determinação da sequência primária, identificação e localização de PTMs, e determinação da estrutura 3D das espidroínas, decidiu-se então investigar, através do uso de ferramentas de bioinformática estrutural, se essa camada de ácidos graxos contribui de alguma forma para a manutenção das propriedades mecânicas da seda ao entrar em contato com as espidroínas. Ou seja, se essa camada de ácidos graxos atuaria como uma camada protetora ao manter a conformação estrutural das espidroínas, e consequentemente as propriedades mecânicas das fibras da seda quando em contato com o meio ambiente.

Sendo assim, utilizando-se a microscopia eletrônica de transmissão (MET) com técnica de citoquímica para detecção de lipídeos, foi possível observar uma intensa marcação de ácidos graxos (aparece como precipitado negro) sobre a superfície das fibras da seda da aranha *N. clavipes*. Esse resultado está mostrado na figura 56. Abaixo dessa espessa camada lipídica encontra-se uma faixa menos contrastada, constituída pelas espidroínas.



Figura 56. Microscopia eletrônica de transmissão (MET) com técnica de citoquímica para detecção de lipídeos. Fibras da seda mostrando a deposição de material lipídico formando uma camada lipídica sobre a fibra (setas).

Com a confirmação da presença de ácidos graxos nas fibras da seda, e com a determinação da estrutura 3D das espidroínas, decidiu-se então utilizar a proteína espidroína-1 como modelo, para se verificar através da bioinformática estrutural, como se dá essa interação da camada de ácidos graxos com as espidroínas constituintes das fibras da seda; desta maneira, construiu-se caixas virtuais compostas por água, e compostas por ácidos octadecanóico, ácidos tridecanóico e água, conforme discutido adiante.

A metodologia de Dinâmica Molecular, atualmente aumenta a possibilidade no avanço da compreensão de processos biológicos em escala atômica e molecular, sendo possível estudar tanto as estruturas obtidas experimentalmente como aquelas obtidas por modelagem molecular. Utilizando a metodologia de dinâmica molecular para cada modelo 3D contruído para as proteínas da seda da aranha *N. clavipes*, diferentes sistemas (Tabela 5) foram

estabelecidos para as proteínas espidroína-1A e -1B, com e sem as fosforilações (modificação mais abundante observada na sequência das espidroínas), com a finalidade de obter o maior número de informações diferentes sobre as estruturas 3D dessas proteínas.

Para a representação do sistema com ácidos graxos, utilizaram-se as formas mais abundantes desses ácidos que foram identificados sobre as fibras da seda da aranha *N. clavipes*, conforme descrito por Salles et al. (2006). As estruturas 3D dos ácidos graxos tridecanóico (TDA) e octadecanóico (ODA) foram construídos utilizando o programa PyMol (Figura 57). A informação sobre a topologia dos ácidos graxos não estava disponível nos campos de forças utilizados neste estudo, sendo assim, para a construção desta topologia utilizamos o servidor PRODRG (<http://davapc1.bioch.dundee.ac.uk/programs/prodrg/> - Schuettelkopf & Van Aalten, 2004) o qual gera parâmetros para as ligações, ângulos e cargas.



Figura 57. Estrutura 3D dos ácidos graxos (A) tridecanóico e (B) octadecanóico contruídos a partir do programa PyMol.

Para a construção das caixas para a realização da simulação da dinâmica molecular, foram utilizados os modelos 3D oriundos da modelagem molecular, as moléculas de ácidos graxos oriundas do programa PyMol e as moléculas de água padrões do GROMACS, conforme detalhado na metodologia, tabela 5. Todas as caixas foram submetidas aos parâmetros já discutidos anteriormente na metodologia. A figura 58 apresenta uma visão geral das caixas utilizadas durante a simulação da dinâmica molecular.

As simulações de dinâmica molecular foram realizadas com a finalidade de simular as condições naturais, as quais as espidroínas são expostas constituindo as fibras sólidas na construção da teia. Dessa forma foi possível verificar como se dá a interação das moléculas de água e a camada de ácidos graxos com as espidroínas, e também se a camada de ácidos graxos contribui de alguma forma para a manutenção da conformação estrutural das espidroínas na seda quando em contato com o meio ambiente.

Utilizando a metodologia de dinâmica molecular para cada modelo 3D, primeiramente realizamos a simulação de dinâmica das espidroínas em água, com e sem a fosforilação, sendo então apresentados aqui os resultados da dinâmica final de 20 ns. As imagens das estruturas

3D dos modelos moleculares das espidroínas após essa dinâmica foram geradas a partir do programa PyMol (Delano, 2002), e estão ilustradas na figura 59.

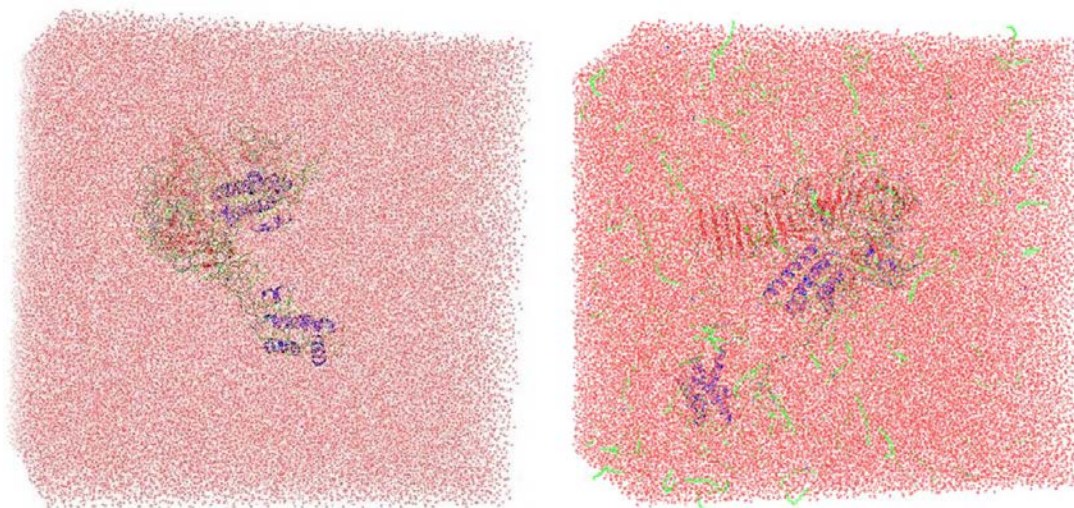


Figura 58. Esquema representativo das caixas que foram utilizadas durante a simulação de dinâmica molecular. **Espidroína-1**, representada em *ribbons*; **água**, representada em *sticks* vermelhos; e **ácidos graxos**, representados em *sticks* verdes. As duas figuras representam as caixas na presença de espidroína-1, ácidos graxos e água sob dois diferentes ângulos.

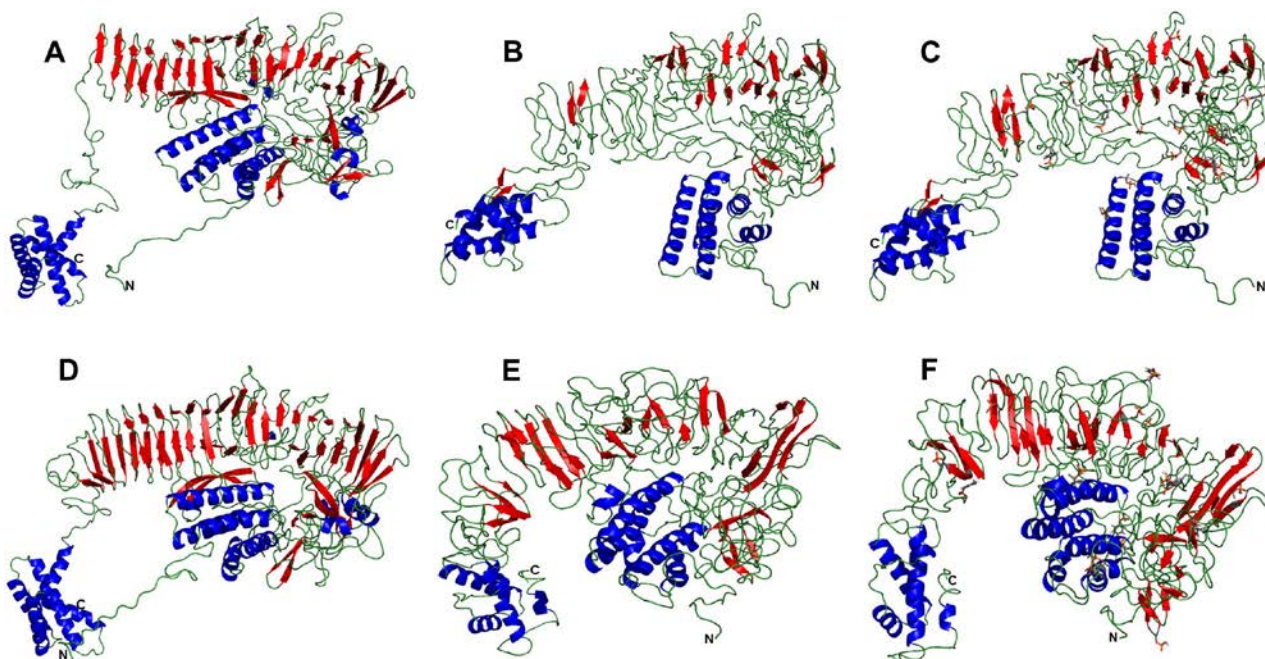


Figura 59. Modelos moleculares das estruturas 3D das espidroínas constituintes das fibras da seda da aranha *N. clavipes* após a dinâmica molecular de 20ns em água. **A** - espidroína-1A; **B** - espidroína-1A após a dinâmica em água; **C** - espidroína-1A com fosforilação, após a dinâmica em água. **D** - espidroína-1B; **E** - espidroína-1B após a dinâmica em água; **F** - espidroína-1B com fosforilação, após a dinâmica em água.

Para avaliar o comportamento global das estruturas 3D das espidroínas foi utilizado a Raiz do Desvio Médio Quadrático (RMSD) que calcula os desvios entre as conformações das proteínas no decorrer da simulação de dinâmica molecular em relação a uma estrutura de referência; o Raio de Giro (Rg) para observar a compactação e expansão da estrutura 3D das proteínas; e a variação da energia (E) em relação ao tempo, tomando como referência as estruturas 3D de cada uma das proteínas após a simulação de dinâmica molecular de restrição.

Nas simulações de dinâmica molecular, mesmo após a etapa prévia de equilíbrio e restrição, quando os átomos do sistema atingem uma distribuição de velocidade compatível com a temperatura, o relaxamento das estruturas ainda continua por tempos que vão além de centenas de picos segundos (ps). Para avaliar este comportamento global das estruturas foi calculado o RMSD, e esse resultado está demonstrado na figura 60.

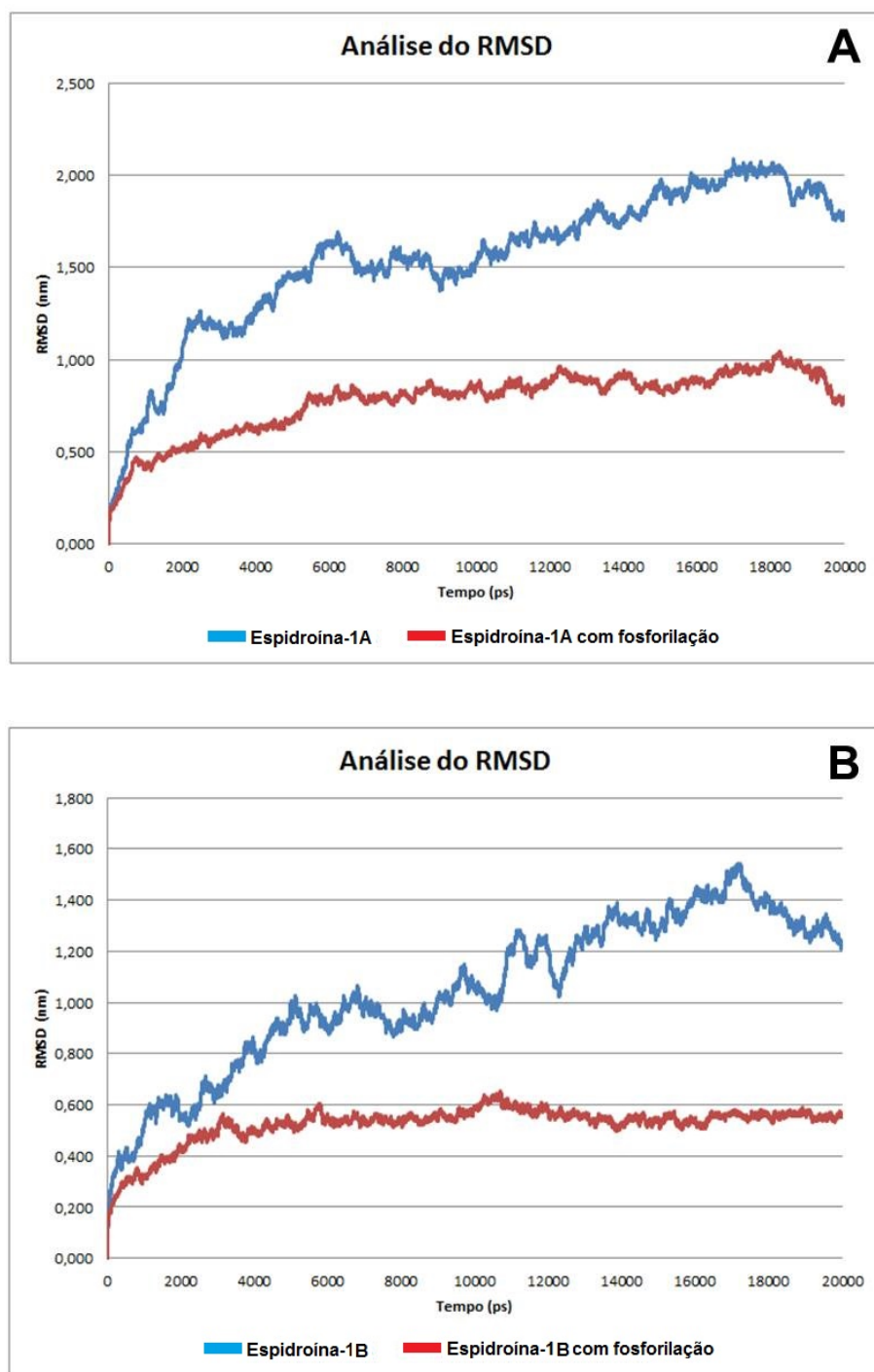


Figura 60. (A) - RMSD da espidroína-1A (azul) em relação à espidroína-1A com fosforilação (vermelho). **(B)** - RMSD da espidroína-1B (azul) em relação à espidroína-1B com fosforilação (vermelho).

Na figura 60 pode-se observar que as espidroínas sem a fosforilação apresentam uma maior variação em sua estrutura 3D em relação às espidroínas com a fosforilação, ou seja, podemos sugerir que a presença da fosforilação representa um fator importante na estabilidade das espidroínas. De um modo geral, o desvio estrutural no início da dinâmica e a estabilização do RMSD podem ser interpretados como variações estruturais devido à mudança de ambiente, o qual as proteínas foram submetidas, e às características particulares de cada modelo 3D.

A análise das simulações pelo Raio de Giro nos fornece informações com relação à evolução da geometria das proteínas, se estão realizando movimentos periódicos de contração e expansão ou se estão girando. Esse resultado está demonstrado na figura 61.

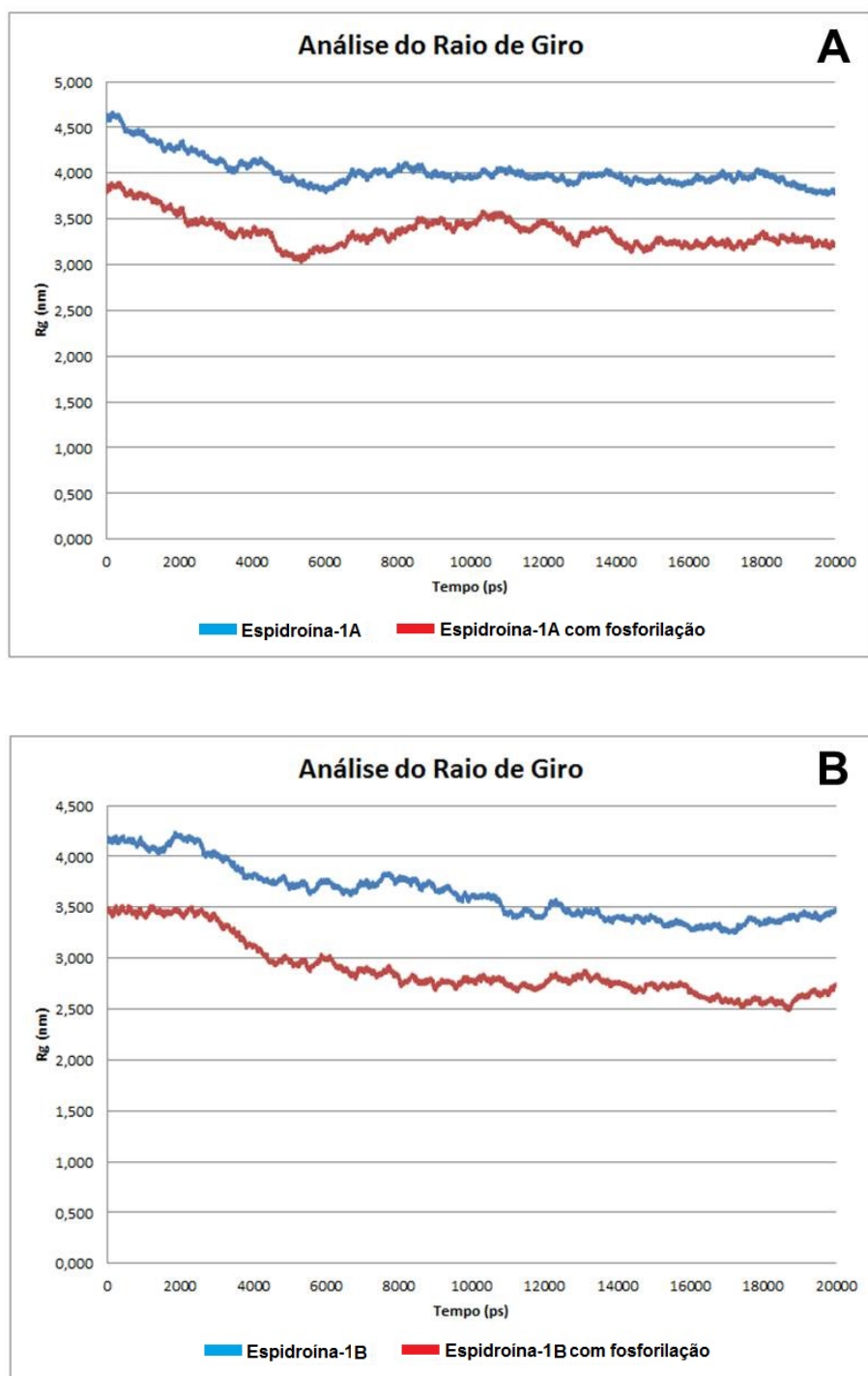


Figura 61. (A) - Gráfico do Raio de Giro para a espidroína-1A (azul) em relação à espidroína-1A com fosforilação (vermelho). **(B)** - Gráfico do Raio de Giro para a espidroína-1B (azul) em relação à espidroína-1B com fosforilação (vermelho).

Na figura 61 pode-se observar que as espidroínas fosforiladas apresentam-se um pouco mais compactas em relação às espidroínas sem qualquer PTM.

Para verificar se o sistema, no qual as espidroínas foram submetidas, está estável energeticamente, realizou-se a análise das simulações por Energia construindo o gráfico da evolução da energia total do sistema em função do tempo. Na figura 62 pode-se observar que o sistema é estável durante toda a simulação de dinâmica molecular, e que a introdução dos grupos fosfato nas espidroínas elevou o nível de energia destas moléculas, em relação às moléculas sem PTMs, sem perder a estabilidade.

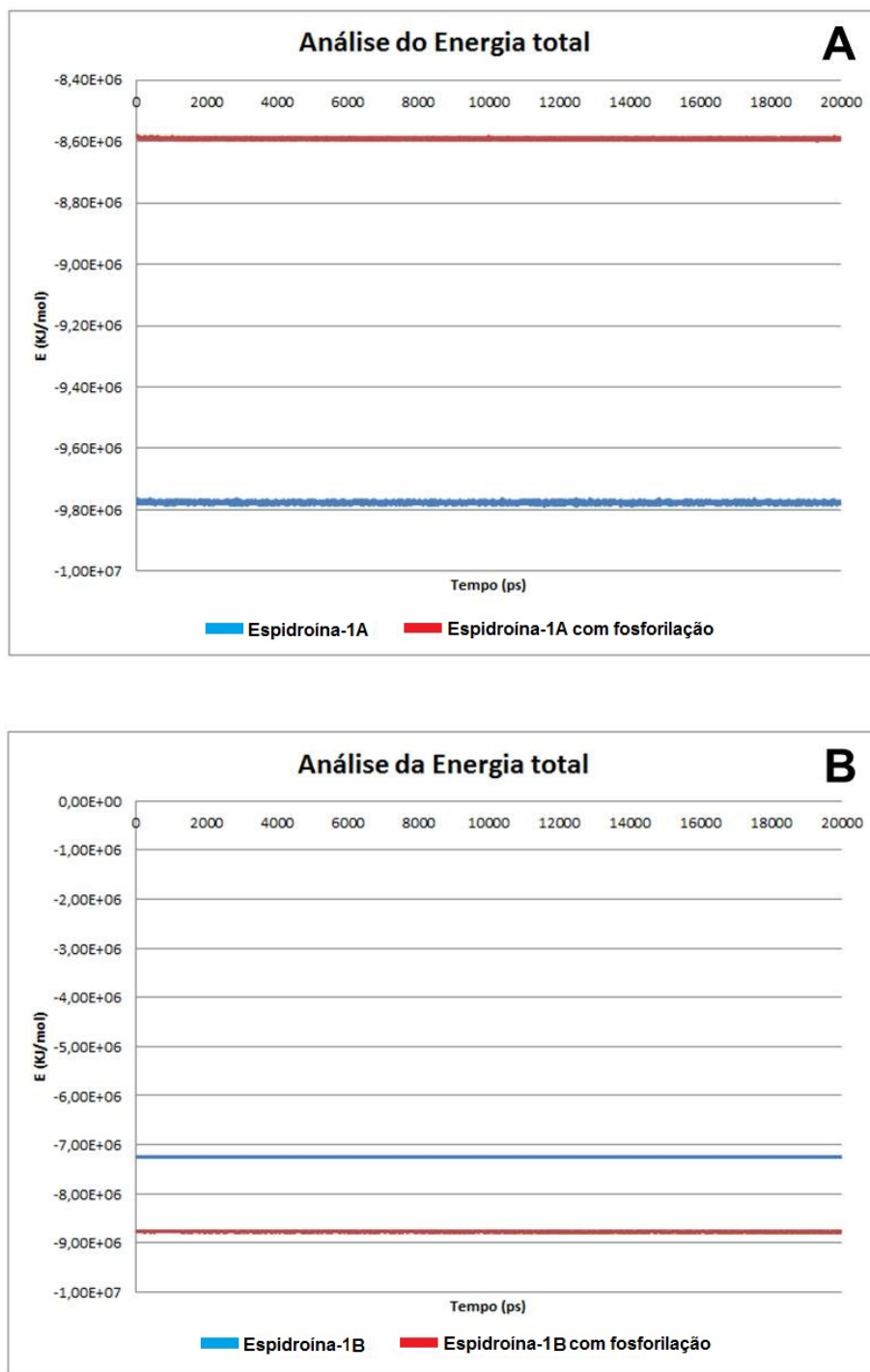


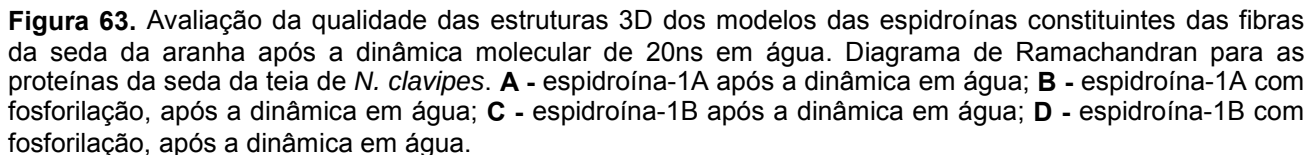
Figura 62. (A) - Gráfico do Energia total para a espidroína-1A (azul) em relação à espidroína-1A com fosforilação (vermelho). **(B)** - Gráfico do Energia total para a espidroína-1B (azul) em relação à espidroína-1B com fosforilação (vermelho).

O programa PROCHECK foi utilizado para analisar a geometria global dos modelos, após a dinâmica molecular final em água, através do diagrama de Ramachandran e o G-factor. A análise dos diagramas de Ramachandran dos modelos das espidroínas, mostrados na figura 63 e tabela 18, demonstra que mais de 90% dos resíduos de aminoácidos estão localizados sobre as regiões favoráveis do gráfico. Portanto, os modelos das espidroínas apresentam uma boa qualidade estereoquímica, levando em consideração que essas proteínas apresentam uma grande quantidade de resíduos de aminoácidos, e ainda composta em sua maior proporção por estruturas aleatórias. Os resíduos que estão na região não favorecida, estão localizados em sua maior parte na região de estruturas aleatórias e alguns poucos em região de superfície da proteína em contato com o meio (solvente), por isso não comprometem a qualidade do modelo, pois resíduos de superfície (expostos ao solvente) e em regiões de estruturas aleatórias geralmente apresentam uma maior flexibilidade de rotação do que os resíduos localizados mais internamente na proteína. A análise do G-factor mostrou que os modelos estão dentro dos intervalos de confiança que validam a estrutura 3D; sendo assim, pode-se afirmar que os modelos apresentam uma boa qualidade quanto ao comprimento de ângulos e sua geometria. As análises utilizando o programa PROCHECK (diagrama de Ramachandran e G-factor), estão demonstradas na tabela 18.

Após a dinâmica molecular de 20ns em água pode-se notar que ocorre alguma perda da conformação estrutural nas espidroínas, sendo bem visível a perda da superestrutura de grampo de folhas- β (Figura 59), que apresenta uma extensa região hidrofóbica. Conforme já discutido anteriormente essa estrutura está relacionada com a estabilidade e dobramentos de proteínas (Trevino et al., 2007; Madan et al., 2014). Sendo assim, pode-se sugerir que quando as fibras estão expostas a um meio aquoso, como por exemplo, nas condições ambientais, as fibras podem perder as suas propriedades mecânicas, já que essas propriedades dependem da região altamente repetitiva que formam as folhas- β (Rising et al., 2007; Rising et al., 2011).

A espidroína-1 possui a repetição de sequências GGX, $(GA)_n$, e $(A)_n$, enquanto a espidroína-2 possui a repetição de sequências $(A)_n$ e GPGXX. As propriedades de força e resistência da fibra são devido à presença das unidades altamente repetitivas de $(A)_n$ e/ou $(GA)_n$ as quais formam estruturas folhas- β cristalinas na fibra, enquanto as propriedades de elasticidade são devido à presença das unidades altamente repetitivas de GGX e GPGXX (Simmons et al., 1996; Gatesy et al., 2001; Rising et al. 2011). Os sítios de fosforilação observados nas espidroínas estão localizados dentro desses motivos estruturais que seriam responsáveis pelas propriedades mecânicas da seda. Apesar de ainda não sabermos qual o

significado biológico da fosforilação observada sobre as espidroínas da seda da aranha *N. clavipes*, pode-se constatar que somente a sua presença não é suficiente para manter a conformação estrutural das proteínas em um meio aquoso. Em alguns estudos foi proposto que os resíduos de serina, treonina e tirosina fosforilados desencadeiam a interação entre os motivos hidrofóbicos de polialanina, causando a agregação das espidroínas e outras proteínas da seda (Arakawa et al., 1985; Huemmerich et al., 2004). Levando esses estudos em consideração e os nossos resultados com a dinâmica molecular final em água, podemos sugerir/corroborar que a fosforilação, presente nas espidroínas diminui a hidrofobicidade da região do grampo de folhas- β , tornando tais regiões um pouco mais expostas ao meio aquoso, o que desorganiza um pouco a conformação desta região, porém mantém as moléculas de espidroínas ainda estáveis e um pouco mais compactas, ocasionando a agregação das espidroínas para a formação das fibras. A partir do momento que as espidroínas perdem essa conformação estrutural constituída pelo grampo de folhas- β , ocorre a perda dessa agregação entre as espidroínas e conseqüentemente a perda das propriedades mecânicas da seda.



Regiões do diagrama de Ramachandran ϕ - ψ (%)					G-factor		
Modelos 3D	Mais Favorecida	Adicionalmente Favorecida	Generosamente Favorecida	Não Favorecida	Ângulos de torção	Geometria covalente	G-factor Global
Espidroína-1A	73.3	24.1	0.8	1.9	-0.41	-0.51	-0.46
Espidroína-1A*	74.1	20.5	3.3	2.2	-0.22	-0.39	-0.31
Espidroína-1B	69.2	27.7	1.9	1.2	-0.20	-0.43	-0.32
Espidroína-1B*	73.1	23.0	2.0	1.9	-0.22	-0.34	-0.28

Utilizando a metodologia de dinâmica molecular para cada modelo 3D, realizamos a simulação de dinâmica em água com ácidos graxos, com e sem a fosforilação. Antes de iniciar a dinâmica molecular final, uma simulação de dinâmica molecular de restrição de posições é realizada para verificar se está ocorrendo algum choque estereoquímico entre os átomos das moléculas. Somente após o resultado dessa dinâmica é realizada a dinâmica molecular final. Sendo assim, para o momento, serão apresentados os resultados da dinâmica de restrição para o sistema em água com ácidos graxos.

As imagens das estruturas 3D dos modelos moleculares das espidroínas, durante a dinâmica de restrição para o sistema em água com ácidos graxos, foram geradas a partir do programa PyMol (Delano, 2002), e estão ilustradas na figura 64. Durante a análise da dinâmica de restrição foi possível notar que as moléculas de ácidos graxos apresentam uma aproximação com as espidroínas, demonstrando uma interação entre as moléculas. Possivelmente, após a dinâmica molecular final, essa interação se apresentará de forma mais visível, o que possibilitará verificarmos, se ocorrerá alguma alteração conformacional nas estruturas das espidroínas em um sistema com água e ácidos graxos com e sem a fosforilação.

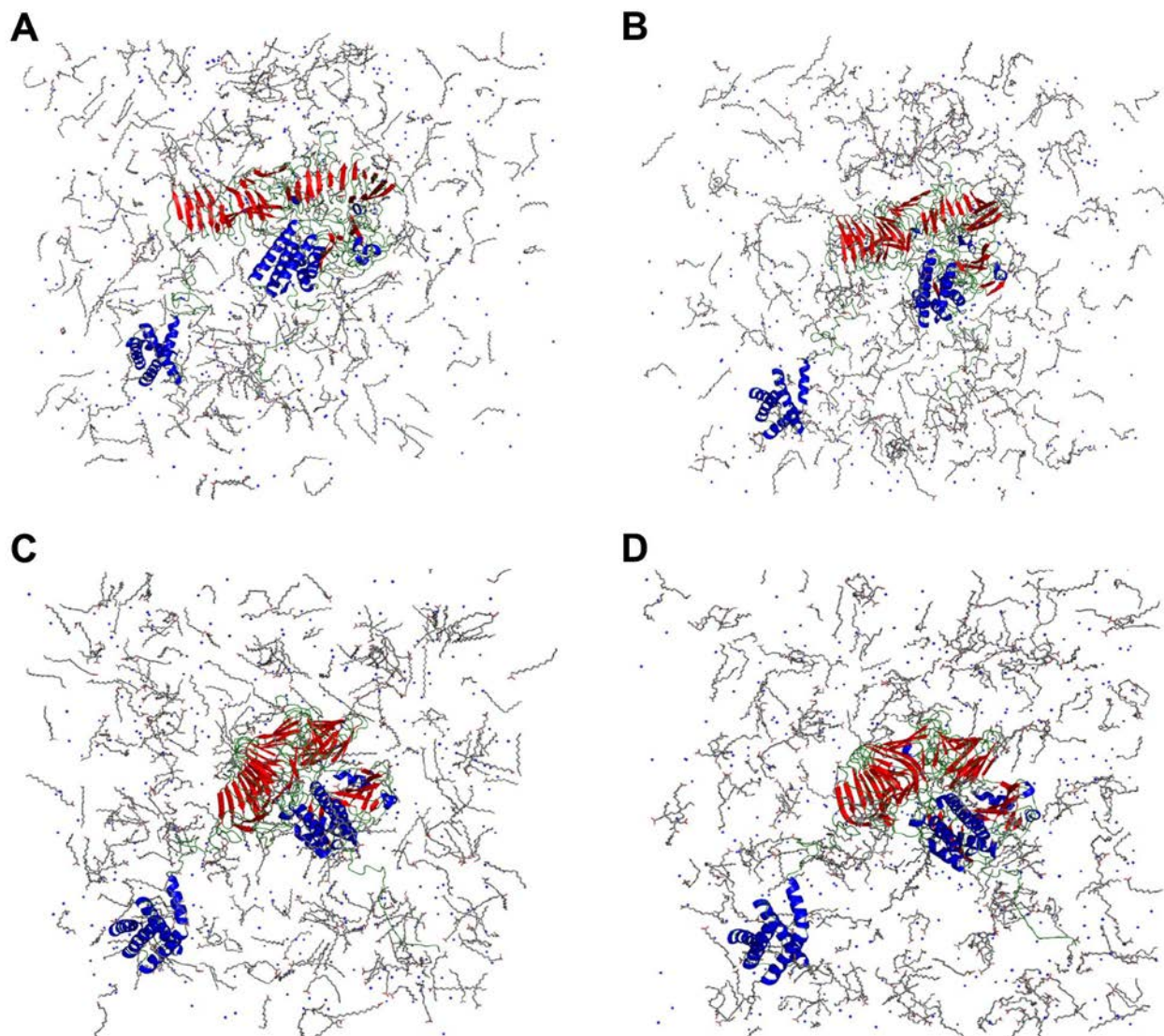


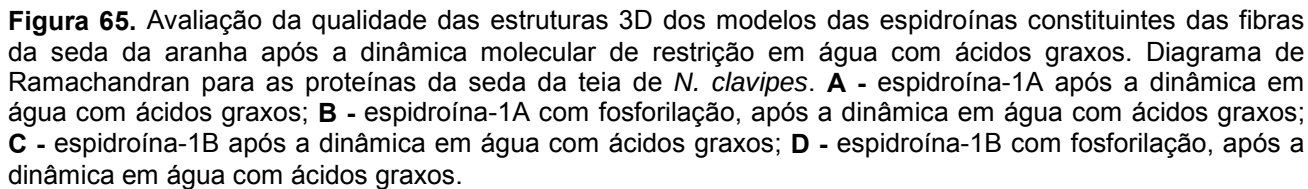
Figura 64. Esquema representativo dos modelos moleculares das estruturas 3D das epidoroínas em caixas durante a dinâmica de restrição para o sistema em água com ácidos graxos. **A** - epidoroína-1A durante a dinâmica de restrição em água com ácidos graxos; **B** - epidoroína-1A com fosforilação, durante a dinâmica de restrição em água com ácidos graxos; **C** - epidoroína-1B durante a dinâmica de restrição em água com ácidos graxos; **D** - epidoroína-1B com fosforilação, durante a dinâmica de restrição em água com ácidos graxos. **Ácidos graxos**, representados em *sticks* verdes.

O programa PROCHECK foi utilizado para analisar a geometria global dos modelos, após a dinâmica molecular de restrição em água com ácidos graxos, através do diagrama de Ramachandran e o G-factor. A análise dos diagramas de Ramachandran dos modelos das epidoroínas, mostrados na figura 65 e tabela 19, demonstra que mais de 90% dos resíduos de aminoácidos estão localizados sobre as regiões favoráveis do gráfico. Portanto, os modelos das

espidroínas apresentam uma boa qualidade estereoquímica, levando em consideração que essas proteínas apresentam uma grande quantidade de resíduos de aminoácidos, e ainda composta em sua maior proporção por estruturas aleatórias.

Analisando os resultados do gráfico de Ramachandran obtidos para os modelos 3D com a dinâmica final em água, e os resultados obtidos até então com a dinâmica de restrição em água com ácidos graxos, podemos constatar que ocorre uma mudança na quantidade de resíduos entre as regiões favorecidas, porém, ao mesmo tempo ocorre um leve aumento no número de resíduos na região não favorecida após a simulação da dinâmica molecular. Isto pode sugerir que as espidroínas sofreram um rearranjo estrutural, e que o meio no qual estão inseridas induziu este rearranjo. Conforme já discutido anteriormente, os resíduos que estão na região não favorecida estão localizados em sua maior parte na região de estruturas aleatórias e alguns poucos estão localizados em região de superfície da proteína em contato com o meio (solvente), por isso não comprometem a qualidade do modelo, pois resíduos de superfície (expostos ao solvente) e em regiões de estruturas aleatórias geralmente apresentam uma maior flexibilidade de rotação do que os resíduos localizados mais internamente na proteína.

A análise do G-factor mostrou que os modelos estão dentro dos intervalos de confiança que validam a estrutura 3D; sendo assim, pode-se afirmar que os modelos apresentam uma boa qualidade quanto ao comprimento de ângulos e sua geometria. As análises utilizando o programa PROCHECK (diagrama de Ramachandran e G-factor), estão demonstradas na tabela 19.



Regiões do diagrama de Ramachandran ϕ - ϕ (%)					G-factor		
Modelos 3D	Mais Favorecida	Adicionalmente Favorecida	Generosamente Favorecida	Não Favorecida	Ângulos de torção	Geometria covalente	G-factor Global
Espidroína-1A	73.1	21.6	3.0	2.3	-0.21	-0.49	-0.38
Espidroína-1A*	73.0	22.7	2.3	2.0	-0.20	-0.43	-0.30
Espidroína-1B	74.5	20.2	3.6	1.7	-0.25	-0.55	-0.40
Espidroína-1B*	71.9	24.1	2.5	1.6	-0.22	-0.51	-0.37

181

Utilizando a metodologia de dinâmica molecular para cada modelo 3D, realizamos a simulação de dinâmica em água com ácidos graxos, com e sem a fosforilação. Estão sendo apresentados aqui os resultados da dinâmica final parcial de 10 ns, o que corresponde 50% da dinâmica molecular necessária para se analisar qualquer mudança conformacional na molécula. E a partir deste ponto a proteína já se encontra estável no meio ao qual foi exposta.

As imagens das estruturas 3D dos modelos moleculares das espidroínas após a dinâmica final parcial de 10 ns para o sistema em água com ácidos graxos, foram geradas a partir do programa PyMol (Delano, 2002), e estão ilustradas na figura 66. Com a análise da dinâmica final parcial de 10 ns foi possível verificar que ocorre uma interação das moléculas de ácidos graxos com as espidroínas, demonstrando uma maior preservação das estruturas de folhas- β , onde essa interação é mais evidente. Esse resultado está demonstrado na figura 67, e comparando com a análise da dinâmica molecular em água, essa dinâmica sugere que ocorre uma propensão a uma menor perda na quantidade de folhas- β e um empacotamento maior destas proteínas (Tabela 20) em um sistema com água e ácidos graxos com e sem a fosforilação.

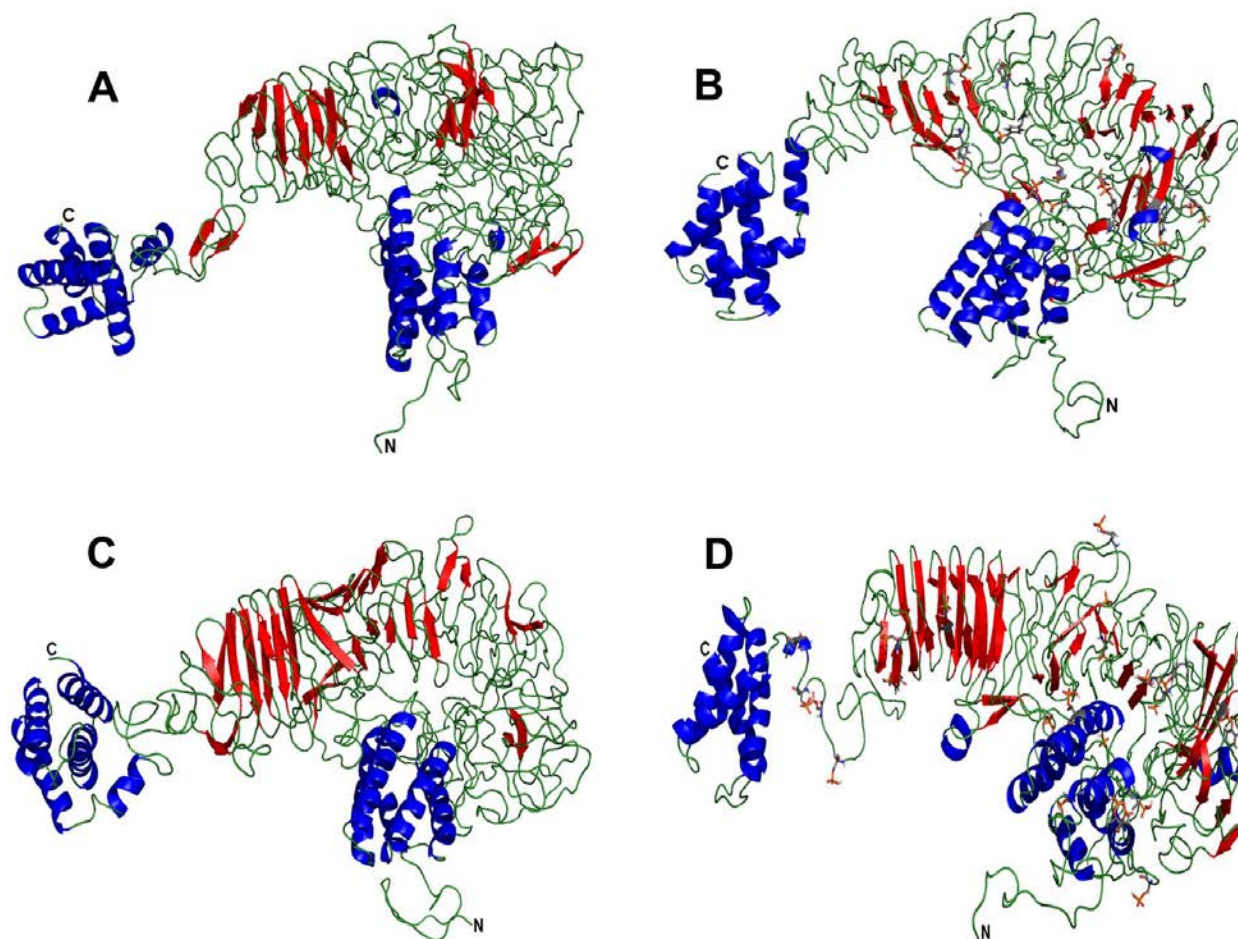


Figura 66. Modelos moleculares das estruturas 3D das espidroínas constituintes das fibras da seda da aranha *N. clavipes* após a dinâmica molecular final parcial de 10ns em água com ácidos graxos. **A** - espidroína-1A após a dinâmica em água com ácidos graxos; **B** - espidroína-1A com fosforilação, após a dinâmica em água com ácidos graxos; **C** - espidroína-1B após a dinâmica em água com ácidos graxos; **D** - espidroína-1B com fosforilação, após a dinâmica em água com ácidos graxos.

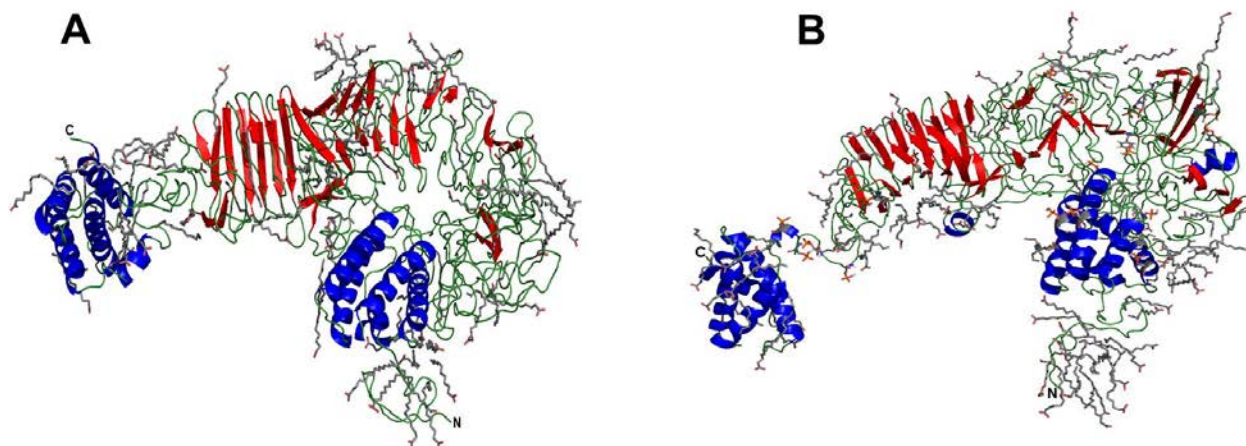


Figura 67. Modelos moleculares das estruturas 3D das espidroínas constituintes das fibras da seda da aranha *N. clavipes* após a dinâmica molecular final parcial de 10ns em água com ácidos graxos. **A** - espidroína-1B após a dinâmica em água com ácidos graxos; **B** - espidroína-1B com fosforilação, após a dinâmica em água com ácidos graxos. Em destaque a interação das moléculas de ácidos graxos com as estruturas da espidroína.

Tabela 20. Análise das estruturas secundárias, para os modelos moleculares das espidroínas, após a dinâmica molecular de 20 ns em água, e após a dinâmica molecular parcial de 10 ns em presença de ácidos graxos. Foi utilizado o programa STRIDE (Frishman e Argos, 1995) e o DSSP (Joosten et al., 2010) através do GROMACS para calcular a porcentagem e o número de estruturas secundárias.

Modelos 3D	Condição/sistema estabelecido para a proteína	Estruturas secundárias (%)				Área de Superfície (Å ²)
		α-hélice	hélice 3 ₁₀ e hélice π	fita	coil	
Espidroína-1A (sem PTM)	Água	13.9	0.9	16.3	68.9	32622
Espidroína-1A (sem PTM)	Água + ácidos graxos	16.5	0.4	15.3	67.9	31676
Espidroína-1A (fosforilada)	Água	14.3	0.9	15.4	69.4	43108
Espidroína-1A (fosforilada)	Água + ácidos graxos	17.3	0.3	14.9	67.6	30911
Espidroína-1B (sem PTM)	Água	13.5	1.2	27.8	57.6	33989
Espidroína-1B (sem PTM)	Água + ácidos graxos	15.8	0.11	20.9	62.2	31630
Espidroína-1B (fosforilada)	Água	14.7	0.6	17.7	67.1	42640
Espidroína-1B (fosforilada)	Água + ácidos graxos	14.5	0	19.6	65.9	32594

A Dinâmica Molecular total do sistema em água com ácidos graxos, ainda está em execução. Ainda não foi finalizada devido a questões de ordens técnicas, já que todas essas análises dependem de programas, cálculos, softwares e computadores para que sejam realizadas; e essas análises demandam de tempo suficiente para que os resultados obtidos possam ser validados, sendo assim, foram apresentados os resultados da dinâmica molecular parcial de 10 ns para o sistema em água com ácidos graxos.

Diante os resultados obtidos com a identificação e sequenciamento das espidroínas, e ainda a caracterização estrutural 3D dessas proteínas através da bioinformática estrutural, estamos propondo pela primeira vez, um esquema representativo (Figura 68) para o mecanismo natural de fiação das fibras da seda da teia da aranha *N. clavipes*, destacando a presença de ditirosina, uma camada de ácidos graxos sobre as fibras e as principais PTMs observadas sobre as espidroínas constituintes das fibras que formam a espiral e os raios/frame da teia.

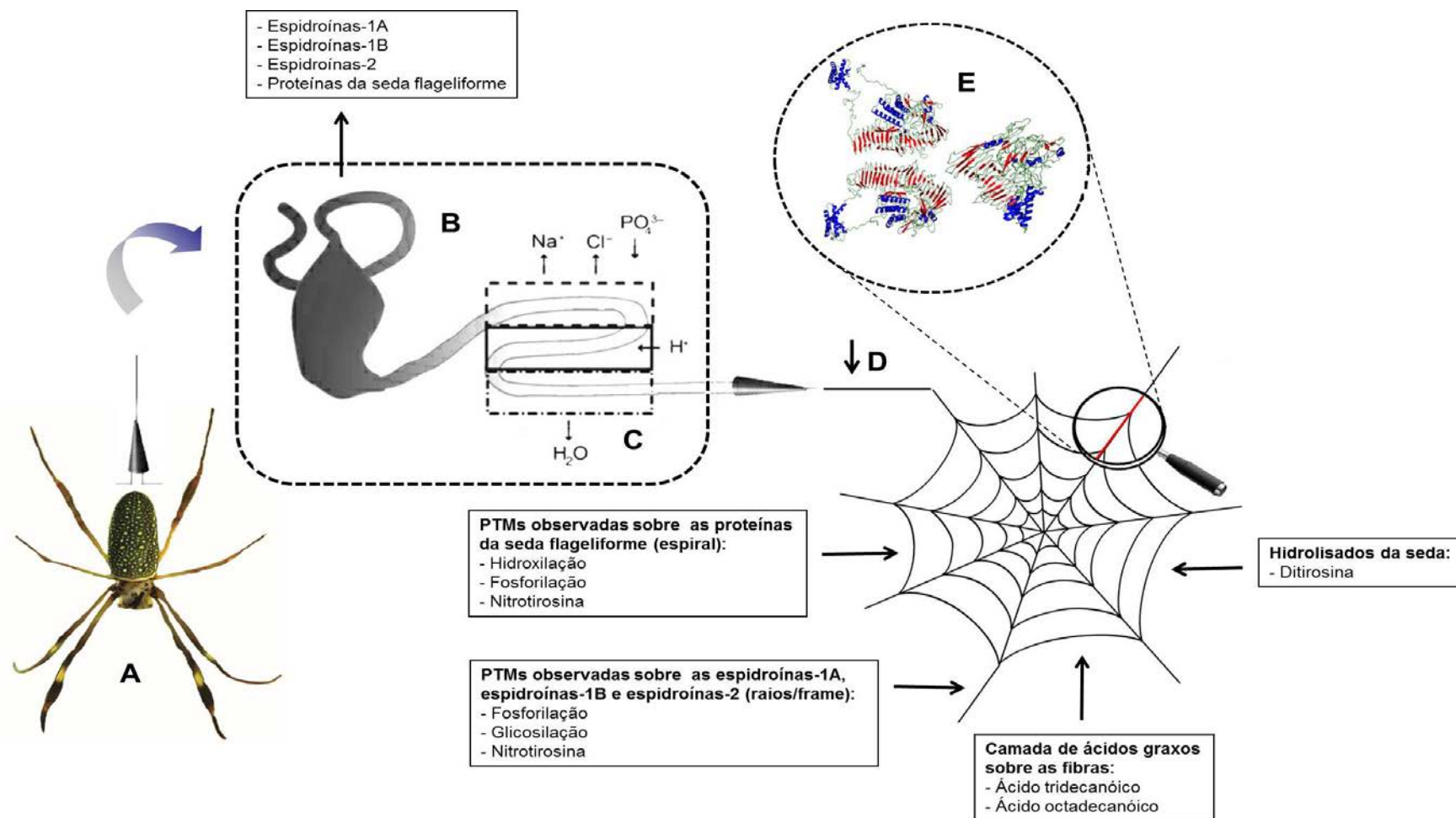


Figura 68. Representação geral esquematizada para o mecanismo natural de fiação das fibras da seda da teia da aranha *N. clavipes*, destacando a presença de ditirosina, uma camada de ácidos graxos sobre as fibras e as principais PTMs observadas sobre as espidroínas constituintes das fibras que formam a espiral e os raios/frame da teia. **(A)** - aranha *N. clavipes*; **(B)** - glândula produtora de seda; **(C)** - ducto de fiação; **(D)** - fibras da seda, constituídas pelas espidroínas, após a secreção pelo ducto de fiação; **(E)** - espidroínas-1A, espidroínas-1B e espidroínas-2 representadas em suas estruturas 3D, constituindo as fibras da seda.

7. Resumo dos resultados

- No presente estudo, a análise proteômica do conjunto de glândulas produtoras de seda da aranha *N. clavipes* mostrou a identificação de diversas proteínas presentes em praticamente todas as glândulas;
- Foram identificados seis diferentes grupos de proteínas, de acordo com a sua função geral: (i) proteínas constituintes das fibras da seda, as espidroínas, (ii) proteínas relacionadas com o dobramento/conformação e modificação das espidroínas, (iii) proteínas relacionadas com a preservação das espidroínas (*self-protection*) contra o estresse oxidativo, (iv) proteínas relacionadas com a preservação fibrilar das espidroínas no meio ambiente, (v) proteínas de transporte de íons e moléculas relacionadas com a estabilidade das espidroínas, e (vi) proteínas *housekeeping*;
- O mecanismo geral de atuação das proteínas identificadas nas glândulas demonstra que essas proteínas podem estar envolvidas com a produção, secreção, armazenamento, transporte, proteção e mudanças conformacionais das espidroínas constituintes da seda durante todo o processo de fiação;
- No presente estudo, utilizando a abordagem proteômica *bottom-up*, mostramos pela primeira vez a análise estrutural e a presença de PTMs observadas sobre as espidroínas constituintes das fibras da seda de aranhas;
- Obtivemos uma elevada cobertura das sequências primárias oriundas da análise proteômica utilizando diferentes enzimas proteolíticas; até então informações sobre as sequências primárias só foram obtidas através de estudos de cDNA;
- Foram observados diversos sítios de fosforilação: 12 na região N-terminal e 16 na região central e C-terminal da espidroína-1, 11 na espidroína-2, e 3 na proteína da seda flageliforme;
- Também foram observados 2 sítios de glicosilação e 1 de nitrotirosina na espidroína-1. E ainda 18 sítios de hidroxilação e 1 de nitrotirosina na proteína da seda flageliforme;
- A grande quantidade de sítios de fosforilação (localizados dentro dos motivos estruturais que seriam responsáveis pelas propriedades mecânicas da seda), juntamente com a

provável fotoindução de hidroxilação pode ser relevante para diversos estudos em ciências de materiais, biologia, biofísica e bioquímica;

- Até então, todas as propriedades mecânicas das fibras da seda têm sido caracterizadas sem considerar as PTMs nas sequências das espidroínas. Tanto as espidroínas naturais quanto as recombinantes têm sido bioquímica e fisicamente caracterizadas como se apresentassem as mesmas (ou similares) propriedades físico-químicas;
- A presença de ditirosina na amostra do hidrolisado da seda pode indicar a presença de ligações cruzadas entre as proteínas, e isso ajudaria na estabilidade das fibras da seda;
- Os resultados obtidos para as espidroínas poderá proporcionar uma melhor compreensão das propriedades da seda visando à produção sintética/recombinante de novos materiais à base da seda de aranhas;
- As sequências das espidroínas também podem ser utilizadas para as análises filogenéticas e para a determinação das relações evolutivas entre as aranhas;
- Devido aos resultados obtidos com a identificação, sequenciamento e localização das PTMs sobre as sequências primárias das espidroínas constituintes das fibras da seda, foi possível realizar a caracterização estrutural 3D das espidroínas através da bioinformática estrutural;
- Utilizando a Microscopia eletrônica de transmissão (MET) com técnica de citoquímica para detecção de lipídeos foi possível observar uma intensa marcação de ácidos graxos sobre a superfície das fibras da seda da aranha *N. clavipes*;
- Com a dinâmica molecular de 20ns em água podemos notar que ocorre uma perda da conformação estrutural nas espidroínas, sendo bem visível a perda da superestrutura de grampo de folhas- β , que está relacionada com a estabilidade e dobramentos de proteínas;
- Na dinâmica molecular, para o sistema em água com ácidos graxos, foi possível notar que as moléculas de ácidos graxos apresentam uma aproximação com as espidroínas, demonstrando uma interação entre as moléculas; e a dinâmica molecular sugere que ocorre uma menor perda na quantidade de folhas- β .

8. Conclusão

Com os resultados obtidos no presente estudo, estamos propondo pela primeira vez, um mecanismo geral de ação das proteínas identificadas nas glândulas produtoras de seda da aranha *N. clavipes* que possam estar envolvidas com a produção, secreção, armazenamento, transporte, proteção e mudanças conformacionais das espidroínas constituintes da seda durante todo o processo de fiação até a seda ser secretada e também após a sua secreção, já formando as fibras sólidas. Até então, não havia sido proposto nenhuma relação dessas proteínas com um mecanismo geral de ação para a formação das fibras da seda.

E ainda, no presente estudo foi observada uma diversidade de PTMs nas espidroínas constituintes das fibras da seda da aranha *N. clavipes*. Cada uma dessas PTMs pode contribuir para a determinação do dobramento da estrutura conformacional, estabilidade, resistência às proteases, ligação e interação entre as proteínas, caracterizando as propriedades da seda. A seda da aranha é um material natural que apresenta propriedades mecânicas superiores ou iguais aos materiais sintéticos existentes, devido a uma combinação de suas propriedades elásticas e de resistência. E, possivelmente, podemos dizer que essas PTMs podem estar atuando de forma coordenada, sinérgica, na caracterização dessas propriedades. Sendo assim, esses resultados obtidos para a seda da aranha *N. clavipes* fornece a base para os estudos das propriedades mecânico-elásticas e conformacional da seda visando possíveis aplicações biotecnológicas e biomédicas.

9. Referências Bibliográficas

- ALLMELING, C. et al. Use of spider silk fibres as an innovative material in a biocompatible artificial nerve conduit. **J. Cell. Mol. Med.** 10, 770-777, 2006.
- ALTSCHUL, S.F.; MADDEN, T.L.; SCHAFER, A.A. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. **Nucleic Acids Res.** 25, 3389-3402, 1997.
- ALTMAN, G.H. et al. Silk-based biomaterials. **Biomaterials** 24, 401-416, 2003.
- ANDERSSON, M.; HOLM, L.; RIDDERSTRÅLE, Y.; JAN JOHANSSON, J.; RISING, A. Morphology and Composition of the Spider Major Ampullate Gland and Dragline Silk. **Biomacromolecules** 14, 2945-2952, 2013.
- APEL, K.; HIRT, H. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. **Annu. Rev. Plant Biol.** 55, 373-99, 2004.
- ARAKAWA, T.; TIMASHEFF, S.N. Theory and protein solubility. **Methods Enzymol** 114,49-77, 1985.
- ASKARIEH, G.; HEDHAMMAR, M.; NORDLING, K.; SAENZ, A.; CASALS, C.; RISING, A.; JOHANSSON, J.; KNIGHT, S.D. Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. **Nature** 465,236-238, 2010.
- AUGSTEN, K.; MUEHLIG, P.; HERMMAN, C. Glycoproteins and skin-core structure in *Nephila clavipes* spider silk observed by light and electron microscopy. **Scanning** 22,12-15, 2000.
- AVERDAM, A.; MARKL, J.; BURMESTER, T. Subunit sequences of the 4 x 6-mer hemocyanin from the golden orb-web spider, *Nephila inaurata*. **Eur J Biochem.** 270(16), 3432-9, 2003.
- AVIS, T.J.; BOULANGER, R.R.; BÉLANGER, R.R. Synthesis and biological characterization of (Z)-9-heptadecenoic and (Z)-6-methyl-9-heptadecenoic acids: fatty acids with antibiotic activity produced by *Pseudozyma flocculosa*. **J. Chem. Ecol.** 26, 987-1000, 2000.
- AYOUB, N.A.; GARB, J.E.; TINGHITELLA, R.M.; COLLIN, M.A.; HAYASHI, C.Y. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. **PLoS ONE** 2: e514, 2007.
- AYOUB, N. A. & HAYASHI, C. Y. Multiple recombining loci encode MaSp1, the primary constituent of dragline silk, in widow spiders (Latrodectus: Theridiidae). **Mol. Biol. Evol.** 25, 277-286, 2008.
- BAE, N.; LI, L.; LÖDL, M.; LUBEC, G. Peptide toxin glacontryphan-M is present in the wings of the butterfly *Hebomoia glaucippe* (Linnaeus, 1758) (Lepidoptera: Pieridae). **Proc Natl Acad Sci U S A.** 109(44), 17920-4, 2012.
- BAOYONG, L.; JIAN, Z.; DENG LONG, C.; MIN, L. Evaluation of a new type of wound dressing made from recombinant spider silk protein using rat models. **Burns** 36(6),891-6, 2010.
- BAYAT, A. Science, medicine, and the future: Bioinformatics. **BMJ**, v. 324, p. 1018-1022, 2002.

- BAUD, F.; KARLIN, S. Measures of residue density in protein structures. **Proc Natl Acad Sci USA** 96(22),12494-12499, 1999.
- BECKWITT, R.; ARCIDIACONO, S. Sequence conservation in the C-terminal region of spider silk proteins (spidroin) from *Nephila clavipes* (Tetragnathidae) and *Araneus bicentenarius* (Araneidae). **J Biol Chem** 9,6661-6663, 1994.
- BELL, A.L.; PEAXALL, D.B. Changes in Fine Structure of Spider. **The Journal of cell biology** 42, 284-295, 1969.
- BERENDSEN, H.J.C.; POSTMA, J.P.M.; DINOLA, A.; et al., Molecular dynamics with coupling to an external bath. **J. Chem. Phys.** 81, 3684-3690, 1984.
- BERMAN, H. M.; WESTBROOK, J.; FENG, Z.; GILLILAND, G.; BHAT, T. N.; WEISSIG, H.; SHINDYALOV, I. N.; BOURNE, P. E. The Protein Data Bank, **Nucleic Acids Research** 28, 235-242, 2000.
- BINI, E.; KNIGHT, D.P.; KAPLAN, D.L. Mapping domain structures in silks from insects and spiders related to protein assembly. **J. Mol. Biol.** 335,27-40, 2004.
- BLACKLEDGEL, T. A.; CODDINGTON, J. A.; GILLESPIE, R. G. Are three-dimensional spider webs defensive adaptations? **Ecology Letters**, 6 (1): 13-18, 2003.
- BLACKLEDGE, T. A.; HAYASHI, C. Y. Unraveling the mechanical properties of composite silk threads spun by cribellate orb-weaving spiders. **J. Exp. Biol.** 209, 2452–2461, 2006.
- BLAMIRE, S.J.; CHAO, I.C.; TSO, I.M. Prey type, vibrations and handling interactively influence spider silk expression. **J. Exp. Biol.** 213, 3906-3910, 2010.
- BLAMIRE, S.J.; WU, C.L.; TSO, I.M. Variation in protein intake induces variation in spider silk expression. **PLoS ONE** 7(2) e31626, 2012.
- BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Anal. Biochem.** 72, 248-254, 1976.
- BRAUNITZER, G.; WOLFF, D. Vergleichende chemische Untersuchungen fiber die Fibroine von *Bombyx mori* und *Nephila madagascariensis*. **Z. Naturforsch** . 10b, 404, 1955.
- BRESCOVIT, A.D.; BERTANI, R.; PINTO-DA-ROCHA, R.; RHEIMS, C.A. **Aracnídeos da Estação Ecológica Juréia-Itatins: Inventário preliminar e história natural**. In Marques, O. A. V. e Duleba, W. (Editores). Estação Ecológica Juréia-Itatins: Ambiente físico, flora e fauna. Ed. Holos. Ribeirão Preto. .198-221, 2004.
- BREUKER, K.; JIN, M.; HAN, X.; JIANG, H. & MCLAFFERTY, F.W. Top-down identification and characterization of biomolecules by mass spectrometry. **J. Am. Soc. Mass Spectrom.** 19, 1045-1053, 2008.
- BREWIS, I. A.; BRENNAN P. Proteomics technologies for the global identification and quantification of proteins. **Adv. Protein Chem. Struct. Biol.** 80,1-44, 2010.

- BRITTON, L.M.; GONZALES-COPE, M.; ZEE, B.M.; GARCIA, B.A. Breaking the histone code with quantitative mass spectrometry. **Expert Rev. Proteomics** 8, 631-643, 2011.
- CANELLE, L.; BOUSQUET, J.; PIONNEAU, C.; DENEUX, L.; IMAMSGHIOUAR, N.; CARON, M.; JOUBERT-CARON, R. Na efficient proteomics-based approach for the screening of autoantibodies. **Journal of Immunological Methods** 299, 77-89, 2005.
- CANNON J.; LOHNES K.; WYNNE C.; WANG Y.; EDWARDS N.; FENSELAU C. High-throughput middle-down analysis using an orbitrap. **J. Proteome Res.** 9, 3886-3890, 2010.
- CANTIN, G. T.; YATES, J. R. Strategies for shotgun identification of post-translational modifications by mass spectrometry. **J. Chromatogr.** 1053, 7-14, 2004.
- CHEN, X.; KNIGHT, D.P.; SHAO, Z.; VOLLRATH, F. Conformation transition in silk protein films monitored by time-resolved Fourier transform infrared spectroscopy: effect of potassium ions on Nephila spidroin films. **Biochemistry** 41,14944-14950, 2002.
- CHEN, W.Q.; PRIEWALDER, H.; JOHN, J. P.; LUBEC, G. Silk cocoon of Bombyx mori: proteins and posttranslational modifications-heavy phosphorylation and evidence for lysine-mediated cross links. **Proteomics.** (3),369-79, 2010.
- CHIKUMA, T.; et al. A fluorometric assay for measuring deamidase (lysosomal protective protein) using high-performance liquid chromatography. **Analytical Biochemistry** 233,36-41, 1996.
- CLAMP, M.; CUFF, J.; SEARLE, S. M.; BARTON, G. J. The Jalview Java Alignment Editor, **Bioinformatics**, v. 20, p. 426-427, 2004.
- Collaborative Computational Project N° 4., The CCP4 suite: programs for proteins crystallography. *Acta Crystallogr.* D50, 760-763, 1994.
- COMSTOCK, J. H. **The spider book**. 4^a ed., London, Cornell University Press, 1975.
- CRAIG, C. L. & RIEKEL, C. Comparative architecture of silks, fibrous proteins and their encoding genes in insects and spiders. **Comp. Biochem. Physiol.** Part B, 133, 493-507, 2002.
- CRANFORD, S.W.; et al. Nonlinear Material Behavior of Spider Silk Yields Robust Webs. **Nature**, 482, 72-76, 2012.
- CREAGER, M.S.; JENKINS, J.E.; THAGARD-YEAMAN, L.A.; BROOKS, A.E.; JONES, J.A.; LEWIS, R.V.; HOLLAND, G.P.; YARGER, J.L. Solid-state NMR comparison of various spiders' dragline silk fiber. **Biomacromolecules** 11(8), 2039-43, 2010
- CUFF, J. A.; BARTON, G.J. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. **Proteins** 40,502-511, 2000.
- CUNNINGHAM, M.G.F.; POLLERO, R.J. Arachnid lipoproteins: comparative aspects. **Comp Biochem Physiol** 146C:79-87, 2007.
- CUNNINGHAM, R.; DI, M.A.; LINGJUN, L.I. Mass spectrometry-based proteomics and peptidomics for systems biology and biomarker discovery. **Front. Biol.** 7, n.4, 313-335, 2012.

- DANI, F.R.; JONES, G.R.; MORGAN, E.D.; TURILAZZI, S. Reevaluation of the chemical secretion of the sternal glands of *Polistes social* wasps (Hymenoptera, Vespidae). **Ethol. Ecol. Evol.** 15, 73, 2003.
- DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. **J. Chem. Phys.** 98, 10089-10092, 1993.
- DAURA, X.; MARK, A.E.; VAN GUNSTEREN, W.F. Parametrization of aliphatic CH_n united atoms of GROMOS96 force field. **J. Comp. Chem.** 19(5), 535-547, 1998.
- DAYER, R.; FISCHER, B. B.; EGGEN, R. I.; LEMAIRE, S. D. The peroxiredoxin and glutathione peroxidase families in *Chlamydomonas reinhardtii*. **Genetics** 179(1), 41-57, 2008.
- DELANO, W. L. **The PyMOL Molecular Graphics System DeLano Scientific**, Palo Alto, CA, USA., 2002.
- DE LIMA, T. **Algoritmos evolutivos para predição de estruturas de proteínas**. São Carlos, 2006. 139 f. Dissertação (Mestrado em Ciências de computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo.
- DERIBE, Y. L.; PAWSON, T.; DIKIC, I. Post-translational modifications in signal integration. **Nat Struct Mol Biol.** 17(6):666-72, 2010.
- DICKO, C.; VOLLRATH, F.; KENNEY, J.M. Spider silk protein refolding is controlled by changing pH. **Biomacromolecules** 5:704-710, 2004.
- DI PALMA, S.; HENNRICH, M.L.; HECK, A.J.; MOHAMMED, S. Recent advances in peptide separation by multidimensional liquid chromatography for proteome analysis. **J. Proteomics** 75, 3791-3813, 2012.
- DONG, Z.; ZHAO, P.; WANG, C.; ZHANG, Y.; CHEN, J.; WANG, X.; LIN, Y.; XIA, Q. Comparative Proteomics Reveal Diverse Functions and Dynamic Changes of *Bombyx mori* Silk Proteins Spun from Different Development Stages. **J. Proteome Res.** 12, 5213-5222, 2013.
- DONG-IK, L.; SANGPILL, H.; YUN, C.J.; IK-SUNG, A.; CHANG-HA, L. A convenient preparation of dityrosine via Mn(III)-mediated oxidation of tyrosine. **Process Biochemistry** 43,999-1003, 2008.
- DOS SANTOS-PINTO, J.R.A.; DOS SANTOS, L.D.; ANDRADE ARCURI, H.; CASTRO, F.M.; KALIL, J.E.; PALMA, M.S. Using Proteomic Strategies for Sequencing and Post-Translational Modifications Assignment of Antigen-5, a Major Allergen from the Venom of the Social Wasp *Polybia paulista*. **J. Proteome Res.** 13(2), 855-865, 2014.
- DOS SANTOS-PINTO, J.R.A.; LAMPRECHT, G.; CHEN, W.Q.; HEO, S.; HARDY, J.G.; PRIEWALDER, H.; SCHEIBEL, T.R.; PALMA, M.S.; LUBEC, G. Structure and post-translational modifications of the web silk protein spidroin-1 from Nephila spiders. **J Proteomics**. pii: S1874-3919(14)00008-6. doi: 10.1016/j.jprot., 2014.
- DRABIK, A.; BODZON-KULAKOWSKA, A.; SUDER, P. Application of the ETD/PTR reactions in top-down proteomics as a faster alternative to bottom-up nanoLC-MS/MS protein identification. **J Mass Spectrom.** 47(10),1347-52, 2012.

- EISOLDT, L.; HARDY, J.G.; HEIM, M.; SCHEIBEL, T.R. The role of salt and shear on the storage and assembly of spider silk proteins. **J Struct Biology** 170,413-419, 2010.
- EISOLDT, L.; SMITH, A.M.; SCHEIBEL, T. Decoding the secrets of spider silk. **Materials Today** 14, 80-86, 2011.
- ESPINOSA, J.F.; GELLMAN, S.H. A designed β -hairpin containing a natural hydrophobic cluster. **Angew Chem. Int. Edit.** 39, 2330-2333, 2000.
- EWLES, M. & GOODWIN, L. Bioanalytical approaches to analyzing peptides and proteins by LC-MS/MS. **Bioanalysis** 3, 1379-1397, 2011.
- EXLER, J.H.; HÜMMERICH, D.; SCHEIBEL, T. The amphiphilic properties of spider silks are important for spinning. **Angew Chem Int Ed Engl.** 46(19),3559-62, 2007.
- EYERS, C.E.; GASKELL, S.J. Mass spectrometry to identify post-translational modifications. Wiley Encyclopedia of Chemical Biology. doi:10.1002/9780470048672.webc469, 2008.
- FENN, J.B.; MANN, M.; MENG, C.K.; WONG, S.F.; WHITEHOUSE, C.M. Electrospray ionization for the mass spectrometry of large biomolecules. **Science** 246, 64-71, 1989.
- FIBER ENGINEERS, MEET THY MASTER. On-line, disponível em http://www.floridanature.org/articles/fiber_enginners.asp, 1996.
- FISCHER, F.; POETSCH, A. Protein cleavage strategies for an improved analysis of the membrane proteome. **Proteome Sci** 4, 2, 2006.
- FRISHMAN, D.; ARGOS, P. Knowledge-based protein secondary structure assignment. *Proteins: Structure, function, and genetics*, 23, 566-579, 1995.
- GAINES, W.A. & MARCOTTE Jr, W.R. Identification and characterization of multiple Spidroin 1 genes encoding major ampullate silk proteins in *Nephila clavipes*. **Insect Molecular Biology** 17(5), 465-474, 2008.
- GARB, J.E. et al. Expansion and intragenic homogenization of spider silk genes since the triassic: evidence from mygalomorphae spidroins. **Mol. Biol. Evol.** 24, 2454-2464, 2007.
- GARCIA, B.A. What does the future hold for top down mass spectrometry? **J. Am.Soc. Mass Spectrom.** 21, 193-202, 2010.
- GASPARI, M.; CUDA, G. Nano LC-MS/MS: a robust setup for proteomic analysis. **Methods Mol Biol.** 790,115-26, 2011.
- GATESY, J.; HAYASHI, C.; MOTRIUK, D.; WOODS, J.; LEWIS, L. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. **Science** 291,2603-2605, 2001.
- GERRITSEN, V. B. An airbus could tiptoe on spider silk. **Protein Spotlight**, 24,1-2, 2000.
- GIBAS, C.; JAMBECK, P. **Desenvolvendo Bioinformática**. Rio de Janeiro: Campus, 179-180, 2001.

GOSLINE, J.M.; DEMONT, M.E.; DENNY, M.W. The structure and properties of spider silk. **Endeavour N.S.** 10, 31-43, 1986.

GYORGY, B.; TOTH, E.; TARCSA, E.; FALUS, A.; BUZAS, E.I. Citrullination: a posttranslational modification in health and disease. **Int J Biochem Cell Biol** 38:1662-1677, 2006.

HAAS, W.; FAHERTY, B.K.; GERBER, S.A.; ELIAS, J.E.; BEAUSOLEIL, S.A.; BAKALARSKI, C.E. Optimization and use of peptide mass measurement accuracy in shotgun proteomics. **Mol Cell Proteomics** 5(7),1326-37, 2006.

HAGN, F.; EISOLDT, L.; HARDY, J.; VENDRELY, C.; COLES, M.; SCHEIBEL, T.; KESSLER, H. A conserved spider silk domain acts as a molecular switch that controls fibre assembly. **Nature** 465:239-242, 2010.

HAGN, F.; THAMM, C.; SCHEIBEL, T.; KESSLER, H. pH-dependent dimerization and salt-dependent stabilization of the N-terminal domain of spider dragline silk - Implications for fiber formation. **Angew. Chem. Int. Ed.** 50(1), 310-3, 2011.

HAGN, F. A structural view on spider silk proteins and their role in fiber assembly. **J pept Sci.** 18(6):357-65, 2012.

HAN, X.; ASLANIAN, A.; YATES, J.R. Mass spectrometry for proteomics. **Curr. Opin. Chem. Biol.** 12, 483-490, 2008.

HARDY, J.G.; SCHEIBEL, T. Silk-inspired polymers and proteins. **Biochem. Soc. Trans.** 37,677-681, 2009.

HARDY, J.G.; et al. Polymeric Materials Based on Silk Proteins. **Polymer** 49(20), 4309-4327, 2008.

HARDY, J.G.; SCHEIBEL, T.R. Composite materials based on silk proteins. **Progress in Polymer Science** 35,1093-1115, 2010.

HAYASHI, C.Y.; LEWIS, R.V. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. **J Mol Biol.** 275(5),773-84, 1998.

HAYASHI, C.Y.; SHIPLEY, N.H.; LEWIS, R.V. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. **Int J Biol Macromol** 24:271-275, 1999.

HAYASHI, C.Y.; LEWIS, R.V. Spider flagelliform silk: Lessons in protein design, gene structure and molecular evolution. **Bioessays**, 23:750-6, 2001.

HEIM, M.; KEERL, D.; SCHEIBEL, T. Spider silk: From soluble protein to extraordinary fiber. **Angew Chem Int Ed** 48:3584-3596, 2009.

HELENIUS, A.; AEBI, M. Roles of N-linked glycans in the endoplasmic reticulum. **Annu. Rev. Biochem.** 73, 1019-1049, 2004.

HELMER, M. Web of intrigue. **Nature**, 2001.

HEO, S.; CSASZAR, E.; JUNG, G.; BEUK, T.; HÖGER, H.; LUBEC, G. Hippocampal levels and activity of the sodium/potassium transporting ATPase subunit alpha-3 (AT1A3) are paralleling memory training in the multiple T-Maze in the C57BL/6J mouse. **Neurochem Int.** 61(5), 702-12, 2012.

HINMAN, M.B.; JONES, J.A.; LEWIS, R.V. Synthetic spider silk: a modular fiber. **Trends Biotechnol.** 18,374-379, 2000.

HOCKNEY, R.W.; GOEL, S.P.; EASTWOOD, J. Quiet high resolution computer models of plasma. **J. Comp. Phys.** 14, 148-158, 1974.

HOU, Y.; XIA, Q.; ZHAO, P.; ZOU, Y.; LIU, H.; GUAN, J.; GONG, J.; XIANG, Z. Studies on middle and posterior silk glands of silkworm (*Bombyx mori*) using two-dimensional electrophoresis and mass spectrometry. **Insect Biochemistry and Molecular Biology** 37, 486-496, 2007.

HUEMMERICH, D.; et al. Primary structure elements of spider dragline silks and their contribution to protein solubility. **Biochemistry** 43:13604-13612, 2004.

HU, X.; VASANTHAVADA, K.; KOHLER, K.; MCNARY, S.; MOORE, A.M.F.; VIERRA, C.A. Molecular mechanisms of spider silk. **Cell. Mol. Life Sci.** 63, 1986-1999, 2006.

HUMENIK, M.; SMITH, A.M.; SCHEIBEL, T. Recombinant spider silks - biopolymers with potential for future applications. **Polymers** 3, 640-661, 2011.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD - Visual Molecular Dynamics. **J. Molec. Graphics** 14(1), 33-38, 1996.

IKEDA, F.; DIKIC, I. Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series. **EMBO Rep.** (6),536-42, 2008.

JAENICKE, E.; DECKER, H. Kinetic properties of catecholoxidase activity of tarantula hemocyanin. **FEBS Journal** 275, 1518-1528, 2008.

JENKINS, J.E.; CREAGER, M.S.; BUTLER, E.B.; LEWIS, R.V.; YARGER, J.L.; HOLLAND, G.P. Solid-state NMR evidence for elastin-like beta-turn structure in spider dragline silk. **Chem Commun (Camb)**. 46(36),6714-6, 2010.

JENSEN, O.N. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. **Curr. Opin. Chem. Biol.** 8:33-41, 2004.

JIA, S.-H.; LI, M.-W.; ZHOU, B.; LIU, W.-B.; ZHANG, Y.; MIAO, X.-X.; ZENG, R.; HUANG, Y.-P. Proteomic Analysis of Silk Gland Programmed Cell Death during Metamorphosis of the Silkworm *Bombyx mori*. **Journal of Proteome Research** 6, 3003-3010, 2007.

JOHNSON, H.; EYERS, C.E. Analysis of post-translational modifications by LC-MS/MS. **Methods Mol Biol.** 658,93-108, 2010.

JOHNSTONE, R.A.W.; ROSE, M.E. **Mass Spectrometry for Chemists and Biochemists**. New York: Cambridge University Press. 2.Ed. 501p., 1996.

JONES, D.T.; TAYLOR, W.R.; THORNTON, J.M. A new approach to protein fold recognition. **Nature** 358, 86-89, 1992.

JONES, D.T.; HADLEY, C. **Threading methods for protein structure prediction.** (In) Bioinformatics: Sequence, structure and databanks. Higgins, D. & Taylor, W.R. Eds., pp1-13, Springer-Verlag, Heidelberg, 2000.

JOOSTEN, R.P.; TE BEEK, T.A.H.; KRIEGER, E.; HEKKELMAN, M.L.; HOOFT, R.W.W.; SCHNEIDER, R.; SANDER, C.; VRIEND G. A series of PDB related databases for everyday needs. *Nucleic Acids Research* 1-9, 2010.

KAMIIE, K.; TAIRA, H.; OURA, H.; KAKUTA, A. Nucleotide sequence of the cDNA encoding silk gland elongation factor 1 alpha **Nucleic Acids Res.** 21(3), 742, 1993.

KAMIIE, K.; YAMASHITA, T.; TAIRA, H.; KIDOU, S.; EJIRI, S. Interaction between elongation factors 1beta and 1gamma from *Bombyx mori* silk gland. **Biosci. Biotechnol. Biochem.** 67, 1522-1529, 2003.

KANZOK, S.M.; FECHNER, A.; BAUER, H.; ULSCHMID, J.K.; MULLER, H.M.; BOTELLA-MUNOZ, J.; SCHNEUWLY, S.; SCHIRMER, H.; BECKER, K. Substitution of the thioredoxin system for the glutathione reductase in *Drosophila melanogaster*. **Science** 291, 643-646, 2001.

KAPLAN, D.L. et al. Silk. In **Protein-Based Materials**, pp. 103-131, Birkhauser, 1997.

KAELIN, W. G. Proline hydroxylation and gene expression. **Annu Rev Biochem.** 74:115-28, 2005.

KARAS, M.; HILLENKAMP, F. Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. **Anal. Chem.** 60, 1299-2301, 1988.

KARPLUS, M.; PETSKO, G.A. Molecular dynamics simulations in biology. **Nature** 347(6294), 631-639, 1990.

KIEHNA, S.E.; WATERS, M.L. Sequence dependence of beta-hairpin structure: comparison of a salt bridge and an aromatic interaction. **Protein Sci.** 12(12), 2657-67, 2003.

KIM, M.-S. & PANDEY, A. Electron Transfer Dissociation Mass Spectrometry in Proteomics. **Proteomics** 12(0): 530-542, 2012.

KLUGE, J.A.; RABOTYAGOVA, O.; LEISK, G.G.; KAPLAN, D. L. Spider silks and their applications. **Trends Biotechnol.** 26(5):244-51, 2008.

KNIGHT, D. P.; KNIGHT, M. M.; VOLLRATH, F. Beta transition and stress-induced phase separation in the spinning of spider dragline silk. **Int. J. Biol. Macromol.** 27, 205-210, 2000.

KNIGHT, D.P.; VOLLRATH, F. Changes in element composition along the spinning duct in a *Nephila* spider. **Naturwissenschaften**, 88, 179-182, 2001.

KNIGHT, D. P.; VOLLRATH, F. Comparison of the spinning of selachian egg case ply sheets and orb web spider dragline filaments. **Biomacromolecules** 2, 323-334, 2001.

- KNIGHT, D.P.; VOLLRATH, F. Spinning an elastic ribbon of spider silk. **Philos Trans R Soc Lond B Biol Sci.** 357, 219-227, 2002.
- KO, F. K.; JOVICIC, J. Modeling of mechanical properties and structural design of spider web. *Biomacromolecules* 5, 780-785, 2004.
- KOBAYASHI, N.; HONDA, S.; YOSHII, H.; MUNEKATA, E. Role of side chains in the cooperative β -hairpin folding of the short C-terminal fragment derived from streptococcal protein G. **Biochemistry** 39, 6564-6571, 2000.
- KOEHL, P.; LEVITT, M. A brighter future for protein structure prediction. **Nature Structural Biology**, v. 6, p. 180-111, 1999.
- KONG, S-K.; YIM, M.B.; STADTMAN, E.R.; CHOCK, P.B. Peroxynitrite disables tyrosine phosphorylation regulatory mechanism: lymphocyte-specific tyrosine kinase fails to phosphorylate nitrated cdc16(6-20)NH₂ peptide. **Proc Natl Acad Sci USA** 93, 3377-3382, 1996.
- KOSKI, K.J.; AKHENBLIT, P.; MCKIERNAN, K.; YARGER, J.L. Non-invasive determination of the complete elastic moduli of spider silks. **Nature Materials** 12, 262-267, 2013.
- KOVOOR, J. **Comparative structure and histochemistry of silkproducing organs in arachnids**. In *Ecophysiology of Spiders*; Nentwig, W., Ed.; Springer: Berlin, 160-186, 1987.
- KUBOTA, K.; YONEYAMA-TAKAZAWA, T.; ICHIKAWA, K. Determination of sites citrullinated by peptidylarginine deiminase using ¹⁸O stable isotope labeling and mass spectrometry. **Rapid Communications in Mass Spectrometry** 19:683-688, 2005.
- KUNTNER, M.; CODDINGTON, J.A.; HORMIGA, G. Phylogeny of extant nephilid orbweaving spiders (Araneae, Nephilidae): testing morphological and ethological homologies. **Cladistics** 24(2), 147-217, 2008.
- LAMMEL, A.; SCHWAB, M.; HOFER, M.; WINTER, G.; SCHEIBEL, T. Recombinant spider silk particles as drug delivery vehicles. **Biomaterials** 32(8), 2233-40, 2011.
- LARSEN, M. R.; TRELLE, M. B.; THINGHOLM, T. E. Analysis of posttranslational modifications of proteins by tandem mass spectrometry. **BioTechniques**, 40:790-798, 2006.
- LASKOWSKI, R.A.; MACARTHUR, M.W.; MOSS, D.S.; et al., PROCHECK - A program to check the stereochemical quality of protein structures. **J. Appl. Cryst.** 26, 283-291, 1993.
- LASKOWSKI, R. A.; MACARTHUR, M. W.; SMITH, D. K.; JONES, D. T.; HUTCHINSON, E. G.; MORRIS, A. L.; NAYLOR, D.; MOSS, D. S.; THORNTON, J. M. PROCHECK v. 3.0: **Program to check the stereochemistry quality of protein structures** – Operating instructions. [S.l.:s.n.], 1994.
- LAZARIS, A. et al. Spider silk fibers spun from soluble recombinant silk produced in mammalian cells. **Science** 295, 472–476, 2002.
- LEFÈVRE, T.; ROUSSEAU, M.E.; PÉZOLET, M. Protein secondary structure and orientation in silk as revealed by Raman spectromicroscopy. **Biophys J.** 92(8), 2885-95, 2007.

- LEFÈVRE, T.; BOUDREAU, S.; CLOUTIER, C.; PÉZOLET, M. Diversity of molecular transformations involved in the formation of spider silks. **J. Mol. Biol.** 405, 238-253, 2011.
- LEIMER, A.; RÖMER, L.; MOUGIN, N.; SLOTTA, U. Synthetic Spider Silk Proteins and Threads. **Society for Biological Engineering** 43-49, 2012.
- LEWANDOWSKA-GNATOWSKA, E.; JOHNSTON, M.L.; ANTOINE, W.; SZCZEGIELNIAK, J.; MUSZYŃSKA, G.; MIERNYK, J.A. Using multiplex-staining to study changes in the maize leaf phosphoproteome in response to mechanical wounding. **Phytochemistry** 72, 1285-1292, 2011.
- LEWIS, R. V. Spider silk: Ancient ideas for new biomaterials. **Chem Rev**, 106:3762-74, 2006.
- LI, J.; YE, L.; LAN, T.; YU, M.; LIANG, J.; ZHONG, B. Comparative proteomic and phosphoproteomic analysis of the silkworm (*Bombyx mori*) posterior silk gland under high temperature treatment. **Mol Biol Rep** 39, 8447-8456, 2012.
- LIEBLER, D. C. The Proteome. In: Liebler, D. C., editor. **Introduction to proteomics: tools for the new biology**. Human Press, Totowa, p. 1-13, 2002.
- LIPMAN, D.J.; PEARSON, W.R. Rapid and sensitive protein similarity searches. **Science** 1985, 227, 1435-1441.
- LIU, Y.; SHAO, Z.; VOLLRATH, F. Relationships between supercontraction and mechanical properties of spider silk. **Nature Materials** 4,901-905, 2005.
- LOTHROP, A.P.; TORRES, M.P.; FUCHS, S.M. Deciphering post-translational modification codes. **FEBS Letters** 587(8), 1247-1257, 2013.
- LUBEC, G. The hydroxyl radical: from chemistry to human disease. **J Investig Med** 44(6),324-346, 1996.
- LUBEC, B.; HERMON, M.; HOEGER, H.; LUBEC, G. Aromatic hydroxylation in animal models of diabetes mellitus. **FASEB J** 12(14),1581-1587, 1998.
- LUBEC, G.; AFJEHI-SADAT, L. Limitations and pitfalls in protein identification by mass spectrometry. **Chem Rev.** 107(8),3568-84, 2007.
- MADAN, B.; SEO, S.; LEE, S.G. Structural and sequence features of two residue turns in beta-hairpins. **Proteins**, doi: 10.1002/prot.24526, 2014.
- MANN, M.; JENSEN, O. N. Proteomic analysis of post-translational modifications. **Nat. Biotechnol.** 21,255-261, 2003.
- MARTI-RENOM, M.A.; STUART, A.C.; FISER, A.; SÁNCHEZ, R.; MELO, F.; ŠALI, A. Comparative protein structure modeling of genes and genomes. **Annual Review of Biophysics & Biomolecular Structure**, v. 29, p. 291-325, 2000.
- MENG, F.; FORBES, A.J.; MILLER, L.M.; KELLEHER, N.L. Detection and localization of protein modifications by high resolution tandem mass spectrometry. **Mass Spectrom. Rev.** 24, 126-134, 2005.

- MENG, G.; SPAHICH, N.; KENJALE, R.; WAKSMAN, G.; Crystal structure of the *Haemophilus influenzae* Hap adhesin reveals an intercellular oligomerization mechanism for bacterial aggregation. **EMBO J.** 30(18), 3864-3874, 2011.
- MEYER, B.; PAPASOTIRIOU, D.G.; KARAS, M. 100% protein sequence coverage: a modern form of surrealism in proteomics. **Amino Acids** 41, 291-310, 2011.
- MICHAL, C.A.; SIMMONS, A.H.; CHEW, B.G.; ZAX, D.B.; JELINSKI, L.W. Presence of phosphorus in *Nephila clavipes* dragline silk. **Biophys J** 70,489-493, 1996.
- MINDEN, J.S.; DOWD, S.R.; MEYER, H.E., STUHLER, K. Difference gel electrophoresis. **Electrophoresis**, Suppl. 1, S156-S161, 2009.
- MIYAMOTO, S.; KOLLMAN, P.A.; SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. **J. Comput. Chem.** 13, 952-962, 1992.
- MOISENOVICH, M. M.; et al. In vitro and in vivo biocompatibility studies of a recombinant analogue of spidroin 1 scaffolds. **Journal of Biomedical Materials Research A** 96, 1 127, 125-131, 2011.
- MOISENOVICH, M. M.; et al. Tissue regeneration in vivo within recombinant spidroin 1 scaffolds. **Biomaterials** 33, 3887e3898, 2012.
- MORELLE, W.; MICHALSKI, J.C. Analysis of protein glycosylation by mass spectrometry. **Nat. protoc.** 2, 1585-1602, 2007.
- MORRIS, A. L.; MACARTHUR, M. W.; HUTCHINSON, E. G.; et al. Stereochemical quality of protein structure coordinates. **Proteins**, v.12, 345-364, 1992.
- MOTRIUK-SMITH, D.; SMITH, A.; HAYASHI, C.Y.; LEWIS, R.V. Analysis of conserved N-terminal domains in major ampullate spider silk proteins. **Biomacromolecules** 6:3152-3259, 2005.
- NAMBA, A.M.; et al. Dinâmica molecular: teoria e aplicações em planejamento de fármacos. **Eclética Química** 33, 13-24, 2008.
- OHGO, K.; KAWASE, T.; ASHIDA, J. & ASAKURA, T. Solid-state NMR analysis of a peptide (Gly-Pro-Gly-Gly-Ala) 6-Gly derived from a flagelliform silk sequence of *Nephila clavipes*. **Biomacromolecules**, 7,1210-1214, 2006.
- PALUMBO, A.M.; SMITH, S.A.; KALCIC, C.L.; DANTUS, M.; STEMMER, P.M.; REID, G.E. Tandem mass spectrometry strategies for phosphoproteome analysis. **Mass Spectrom. Rev.** 30, 600-625, 2011.
- PANDEY, A.; MANN, M. Proteomics to study genes and genomes. **Nature**, New York, 405, 837-846, 2000.
- PARASTATIDIS, I.; THOMSON, L.; BURKE, A.; CHERNYSH, I.; NAGASWAMI, C.; VISSER, J.; STAMER, S.; LIEBLER, D.C.; KOLIAKOS, G.; HEIJNEN, H.F.; FITZGERALD, G.A.; WEISEL, J.W.; ISCHIROPOULOS, H. Fibrinogen beta-chain tyrosine nitration is a prothrombotic risk factor. **J Biol Chem** 283(49), 33846-53, 2008.

- PARCELLIER, A.; GURBUXANI, S.; SCHMITT, E.; SOLARY, E.; GARRIDO, C. Heat shock proteins cellular chaperones that modulate mitochondrial cell death pathways. **Biochem. Biophys. Res. Commun.** 304, 505-512, 2003.
- PEAKALL, D. B. Synthesis of silk, mechanism and location. **Am. Zool.** 9, 71-79, 1969.
- PELUFFO, G.; RADI R. Biochemistry of protein tyrosine nitration in cardiovascular pathology. **Cardiovasc Res.** 75(2),291-302, 2007.
- PENNINGTON, S.R.; WILKINS, M.R.; HOCHSTRASSER, D.F.; DUNN, M.J. Proteome analysis: from protein characterization to biological function. **Trends Cell. Biol.**, Cambridge, 7, 168-173, 1997.
- PLATNICK, N.I. 2008: **The world spider catalog**, version 8.5 American Museum of Natural History. Disponível em: <<http://research.amnh.org/entomology/spiders/catalog>>.
- PLATNICK N.I. 2010: **The world spider catalog**, version 11.0. American Museum of Natural History. Disponível em: <<http://research.amnh.org/iz/spiders/catalog>>.
- PLAZAOLA, A.; CANDELAS, G. C. Stimulation of fibroin synthesis elicits ultrastructural modifications in spider silk secretory cells. **Tissue Cell** 23(2), 277-84, 1991.
- POUCHKINA, N.N.; STANCHEV, B.S.; MCQUEEN-MANSON, S.J. From EST sequence to spider silk spinning: identification and molecular characterisation of Nephila senegalensis major ampullate gland peroxidase NsPox. **Insect Biochemistry and Molecular Biology** 33,229-238, 2003.
- PROUVOST, O.; TRABALON, M.; PAPKE, M.; SCHULZ, S. Contact sex signals on web and cuticle of Tegenaria atrica (Araneae, Agelenidae). **Arch. Insect Biochem. Physiol.** 40, 194, 1999.
- PRYOR, W.A.; UPPU, R.M. A kinetic model for the competitive reactions of ozone with amino acid residues in proteins in reverse micelles. **J Biol Chem** 268:3120-3126, 1993.
- QIN, X.; WAITE, J.H. Exotic collagen gradients in the byssus of the mussel Mytilus edulis. **J Exp Biol.** 198(Pt 3):633-44, 1995.
- QIN, X.X.; COYNE, K.J.; WAITE, J.H. Tough tendons. Mussel byssus has collagen with silk-like domains. **J Biol Chem.** 272(51):32623-7, 1997.
- RAMMENSEE, S.; SLOTTA, U.; SCHEIBEL, T.; BAUSCH, A.R. Assembly mechanism of recombinant spider silk proteins. **Proc Natl Acad Sci U S A.** 105(18),6590-5, 2008.
- RAMIREZ-ALVARADO, M.; BLANCO, F.J.; NIEMANN, H.; SERRANO, L. Role of beta-turn residues in beta hairpin formation and stability in designed peptides. **J Mol Biol** 273, 898-912, 1997.
- RAUSCHER, S.; BAUD, S.; MIAO, M.; KEELEY, F.W.; POMES, R. Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. **Structure** 14:1667-1676, 2006.

- REHM, P.; PICK, C.; BORNER, J.; MARKL, J.; BURMESTER, T. The diversity and evolution of chelicerate hemocyanins. **BMC Evol Biol.** 12,19, 1-13, 2012.
- RENAULT, A.; RIOUX-DUBÉ, J.F.; LEFÈVRE, T. Surface properties and conformation of *Nephila clavipes* spider recombinant silk proteins at the air-water interface. **Langmuir.** 25(14):8170-80, 2009.
- REDEKER, V. Mass spectrometry analysis of C-terminal posttranslational modifications of tubulins. **Methods Cell Biol.** 95, 77-103, 2010.
- RESENDE, V.M.; VASILJ, A.; SANTOS, K.S.; PALMA, M.S.; SHEVCHENKO, A. Proteome and phosphoproteome of Africanized and European honeybee venoms. **Proteomics** 13(17),2638-48, 2013.
- RIBET, D.; COSSART, P. Post-translational modifications in host cells during bacterial infection. **FEBS Letters** 584, 2748–2758, 2010.
- RISING, A.; NIMMERVOLL, H.; GRIP, S.; FERNANDEZ-ARIAS, A.; et al. Spider silk proteins-mechanical property and gene sequence. **Zoolog Sci.** 22:273-81, 2005.
- RISING, A.; et al. Major ampullate spidroins from *euprosthernops australis*: multiplicity at protein, mRNA and gene levels. **Insect Mol Biol** 16:551-561, 2007.
- RISING, A.; WIDHE, M.; JOHANSSON, J.; HEDHAMMAR, M. Spider silk proteins: recent advances in recombinant production, structure–function relationships and biomedical applications. **Cell. Mol. Life Sci.** 68:169-184, 2011.
- ROEPSTORFF, P. Mass spectrometry based proteomics, background, status and future needs. **Protein Cell.** 3(9), 641-7, 2012.
- RÖMER, L.; SCHEIBEL, T. The elaborate structure of spider silk: structure and function of a natural high performance fiber. **Prion.** 2(4):154-61, 2008.
- ROST, R.; SANDER, C. Bridging the Protein Sequence-Structure Gap by Structure Predictions. **Ann. Rev. Bioph. Biomol. Struc.** 25, 113-136, 1996.
- ROUSSEAU, M.E.; LEFÈVRE, T.; PÉZOLET, M. Conformation and orientation of proteins in various types of silk fibers produced by *Nephila clavipes* spiders. **Biomacromolecules.** 10(10):2945-53, 2009.
- ŠALI, A.; BLUNDEL, T.L. Comparative protein modelling by satisfaction of spatial restraints. **J. Mol. Biol.** 234, 779-815, 1993.
- ŠALI, A.; SANCHEZ, R.; MARTI-RENOM, M.A.; et al., **Manual do Modeller: A Program for Protein Structure Modeling**, 2001.
- SALLES, H.C. **Estratégias químicas da teia da aranha *Nephila clavipes* (Linnaeus, 1767) usadas para captura de presas.** Universidade Estadual Paulista, Instituto de Biociências, Rio Claro [s.n.], 2003.

- SALLES, H.C.; VOLSI, E.C.F.R.; MARQUES, M.R.; SANTOS, L.D.; PALMA, M.S. The Venomous Secrets of the Web Droplets from the Viscid Spiral of the Orb-Weaver Spider *Nephila clavipes* (Araneae, Tetragnatidae). **Chemistry & Biodiversity**, vol 3, 2006.
- SANCHEZ, R.; ŠALI, A. Advances in comparative protein-structure modeling. **Current Opinion in Structural Biology**, v. 7, p. 206-214, 1997.
- SANDER, C.; SCHNEIDER, R. Database of homology-derived protein structures and the structural meaning of sequence alignment. **Proteins** 9(1), 56-68, 1991.
- SAVAGE, K.N.; GOSLINE, J.M. The effect of proline on the network structure of major ampullate silks as inferred from their mechanical and optical properties. **J Exp Biology** 211:1937-1947, 2008.
- SCHEIBEL, T.; HEIM, M.; KEERL, D. Spider silk: From soluble protein to extraordinary fiber. **Angew. Chem. Int.** 48, 3584-3596, 2009.
- SCHEIBEL, T. Spider silks: recombinants synthesis, assembly, spinning, and engineering of synthetic proteins. **Microb. Cell Fact.** 3, 14, 2004.
- SCHELLER, J.; HENGGELE, D.; VIVIANI, A. Purification of spider silk-elastin from transgenic plants and application for human chondrocyte proliferation. **Transgenic Res.** 13(1),51-7, 2004.
- SCHWARZE, S.; ZWETTLER, F.U.; JOHNSON, C.M.; NEUWEILER, H. The N-terminal domains of spider silk proteins assemble ultrafast and protected from charge screening. **Nat Commun.** 4,2815, 2013.
- SEO, J.; JEONG, J.; KIM, Y. M.; HWANG, N. Strategy for comprehensive identification of post-translational modifications in cellular proteins, including low abundant modifications: application to glyceraldehyde-3-phosphate dehydrogenase. **J Proteome Res.** 7(2),587-602, 2008.
- SHEN, M.-Y.; ŠALI, A. Statistical potential for assessment and prediction of protein structures. **Protein Science** 15(11), 2507-2524, 2006.
- SCHULER, L.D.; DAURA, X.; VAN GUNSTEREN, W.F. An Improved GROMOS96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase. **J Comp. Chem.** 22(11), 1205-1218, 2001.
- SCHÜTTELKOPF, A.W.; VAN AALTEN, D.M.F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. **Acta Crystallogr.** D60, 1355-1363, 2004.
- SCHULZ, S. Composition of the silk lipids of the spider *Nephila clavipes*. **Lipids** 36(6), 637-647, 2001.
- SHU, H.; CHEN, S.; BI, Q.; MUMBY, M.; BREKKEN, D. L. Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line. **Mol. Cell. Proteomics** 3, 279-286, 2004.
- SIMMONS, A.; MICHAL, C.; JELINSKI, L. Molecular-orientation and 2-component nature of the crystalline fraction of spider dragline silk. **Science** 271, 84-87, 1996.

- SIMPSON, E.R.; MELDRUM, J.K.; BOFILL, R.; CRESPO, M.D.; HOLMES, E.; SEARLE, M.S. Engineering enhanced protein stability through beta-turn optimization: insights for the design of stable peptide beta-hairpin systems. **Angew Chem Int Ed Engl** 44, 4939-4944, 2005.
- SLOTTA, U.; HESS, S.; SPIESS, K.; STROMER, T.; SERPELL, L.; SCHEIBEL, T. Spider silk and amyloid fibrils: A structural comparison. **Macromol Biosci**, 7,183-8, 2007.
- SLOTTA, U.; MOUGIN, N.; RÖMER, L.; LEIMER, A.H. Synthetic spider silk proteins and threads. **Society for Biological Engineering** 43-49, 2012.
- SMITH, B.C.; DENU, J.M. Chemical mechanisms of histone lysine and arginine modifications. **Biophys Acta** 1789, 45-57, 2009.
- SOKOLOWSKA, I. Proteomic analysis of plasma membranes isolated from undifferentiated and differentiated HepaRG cells. **Proteome Sci** 10(1), 47, 2012.
- SPIESS, K.; LAMMEL, A.; SCHEIBEL, T. Recombinant spider silk proteins for applications in biomaterials. **Macromolecular Biosciences** 10(9), 998-1007, 2010.
- SPONNER, A.; SCHLOTT, B.; VOLLRATH, F.; UNGER, E. Characterization of the protein components of *Nephila clavipes* dragline silk. **Biochemistry** 44(12), 4727-4736, 2005.
- SPONNER, A.; VATER, W.; MONAJEMBASHI, S.; UNGER, E.; GROSSE, F.; WEISSHART, K. Composition and hierarchical organisation of a spider silk. **PLoS ONE** 2(10),e998, 2007.
- STADTMAN, E. R. Oxidation of free amino acids and amino acid residues in proteins by radiolysis and by metal-catalyzed reactions. **Annu Rev Biochem** 62, 797-821, 1993.
- STADTMAN, E.R.; LEVINE, R.L. Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. **Amino acids** 25:207-218, 2003.
- STADTMAN, E. R. Protein oxidation and aging. **Free Radical Research** 40(12),1250-1258, 2006.
- STEINBERG, T.H.; TOP, K.; BERGGREN, K.N. Rapid and simple single nanogram detection of glycoproteins in polyacrylamide gels and on electroblots. **Proteomics**, 841-855, 2001.
- SUDERMAN, R. J.; DITTMER, N. T.; KANOST, M. R.; KRAMER, K. J. Model reactions for insect cuticle sclerotization: cross-linking of recombinant cuticular proteins upon their laccase-catalyzed oxidative conjugation with catechols. **Insect Biochem. Mol. Biol.** 36(4), 353-365, 2006.
- SWANEY, D.L.; WENGER, C.D.; COON, J.J. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. **J Proteome Res** 9(3), 1323-1329, 2010.
- SYUD, F.A.; ESPINOSA, J.F.; GELLMAN, S.H. NMR-based quantification of β -sheet populations in aqueous solution through use of reference peptides for the folded and unfolded states. **J. Am. Chem. Soc.** 121, 11577-11578, 1999.
- SYUD, F.A.; STANGER, H.E.; GELLMAN, S.H. Interstrand side chain-side chain interactions in a designed β -hairpin: Significance of both lateral and diagonal pairings. **J. Am. Chem. Soc.** 123, 8667-8677, 2001.

- TATKO, C.D.; WATERS, M.L. Selective aromatic interactions in β -hairpin peptides. *J. Am. Chem. Soc.* 124, 9372-9373, 2002.
- TAYLOR, A. Aminopeptidases: structure and function. *FASEB J.* 1;7(2),290-8, 1993.
- TEIA DE ARANHA PODE SER MATÉRIA PRIMA PARA A INDÚSTRIA. Revista do Ministério da Ciência e Tecnologia. On-line, disponível em www.mct.gov.br/comunicação/textos/default.asp, 2002.
- THEIN, M. C.; WINTER, A. D.; STEPEK, G.; MCCORMACK, G.; STAPLETON, G.; JOHNSTONE, I. L.; PAGE, A. P. Combined extracellular matrix cross-linking activity of the peroxidase MLT-7 and the dual oxidase BLI-3 is critical for post-embryonic viability in *Caenorhabditis elegans*. *J. Biol. Chem.* 284(26), 17549-63, 2009.
- THELEN, J.J.; MIERNYK, J.A. The proteomic future: where mass spectrometry should be taking us. *Biochem J* 444(2),169-181, 2012.
- THIEL, B.L.; KUNKEL, D.D.; VINEY, C. Physical and Chemical Microstructure of Spider Dragline: A Study by Analytical Transmission Electron Microscopy. *Biopolymers* 34, 1089-1097, 1994.
- THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673-4680, 1994.
- THORPE, C.; KIM, J.J. Structure and mechanism of action of the acyl-CoA dehydrogenases. *FASEB J.* 9(9),718-25, 1995.
- TILLINGHAST, E.K. Selective removal of glycoproteins from the adhesive spiral of the spiders orb web. *Naturwissenschaften* 68:526-527, 1981.
- TIPTON, J.D.; TRAN, J.C.; CATHERMAN, A.D.; AHLF, D.R.; DURBIN, K.R.; KELLEHER, N.L. Analysis of intact protein isoforms by mass spectrometry. *J. Biol. Chem.* 286, 25451-25458, 2011.
- TREVINO, S.R.; SCHAEFER, S.; SCHOLTZ, J.M.; PACE, C.N. Increasing protein conformational stability by optimizing beta-turn sequence. *J Mol Biol* 373, 211-218, 2007.
- TSO, I.M.; WU, H.C.; HWANG, I.R. Giant wood spider *Nephila clavipes* alters silk protein in response to prey variation. *J. Exp. Biol.* 208, 1053-1061, 2005.
- UDESHI, N.D.; SHABANOWITZ, J.; HUNT, D.F.; ROSE, K.L. Analysis of proteins and peptides on a chromatographic timescale by electron-transfer dissociation MS. *FEBS J.* 274, 6269-6276, 2007.
- VAN BEEK, J.D.; HESS, S.; VOLLRATH, F.; MEIER, B.H. The molecular structure of spider dragline silk: folding and orientation of the protein backbone. *Proc Natl Acad Sci USA* 99,10266-10271, 2002.
- VAN DER SPOEL, D.; LINDAHL, E.; HESS, B.; et al., GROMACS: fast, flexible, and free. *J Comp. Chem.* 26(16), 1701-1718, 2005.

- VENDRELY, C.; SCHEIBEL, T. Biotechnological Production of Spider-Silk Proteins Enables New Applications. **Macromol. Biosci.** 7, 401-409, 2007.
- VEPARI, C. & KAPLAN, D.L. Silk as a Biomaterial. **Prog Polym Sci.** 32(8-9),991-1007, 2007.
- VESTAL, M.L. The future of biological mass spectrometry. **J. Am. Soc. Mass Spectrom.** 22, 953-959, 2011.
- VOLLRATH, F.; et al. Compounds in the droplets of the orb spider's viscid spiral. **Nature** 345:526-527, 1990.
- VOLLRATH, F.; TILLINGAST, E.K. Glicoprotein glue beneath a spider web's aqueous coat. **Naturwissenschaften** 78, 557-559, 1991.
- VOLLRATH, F. Spider webs and silk. **Sci. Am.** 266, 70-76, 1992.
- VOLLRATH, F.; KNIGHT, D.P. Structure and function of the silk production pathway in the spider *Nephila edulis*. **Int J Biol Macromol** 24,243-249, 1999.
- VOLLRATH, F. Strength and structure of spiders' silks. **J. Biotechnol.**, 74, 67, 2000.
- VOLLRATH, F.; KNIGHT, D. P. Liquid crystalline spinning of spider silk. **Nature**, 410, 541-548, 2001.
- VOLLRATH, F.; MADSEN, B.; SHAO, Z. Z. The effect of spinning conditions on the mechanics of a spider's dragline silk. **Proc Biol Sci.** 22;268(1483):2339-46, 2001.
- VOLLRATH, F.; BARTH, P.; BASEDOW, A.; ENGSTROM, W.; LIST, H. Local tolerance to spider silks and protein polymers in vivo. **In Vivo** 16(4), 229-34, 2002.
- VOLSI, E. C.F.R.; MENDES, M. A.; MARQUES, M. R.; SANTOS, L. D.; PALMA, M. S. Multiple bradykinin-related peptides from the capture web of the spider *Nephila clavipes* (Araneae, Tetragnatidae). **Peptides**, 690-697, 2006.
- VILLAGOMEZ, M.; SZABO, E.; PODCHEKO, A.; FENG, T.; PAPP, S.; OPAS, M. Calreticulin and focal-contact-dependent adhesion. **Biochem Cell Biol.** 87(4),545-56, 2009.
- WAITE, J.H.; QIN, X.X.; COYNE, K. The peculiar collagens of mussel byssus. **Matrix Biol.** 17(2):93-106, 1998.
- WALSH, C.T.; GARNEAU-TSODIKOVA, S.; GATTO, G.J. Protein posttranslational modifications: the chemistry of proteome diversifications. **Angew Chem Int Ed Engl.** 44(45),7342-72, 2005.
- WANG, H.; CHANG-WONG, T.; TANG, H.Y.; SPEICHER, D.W. Comparison of extensive protein fractionation and repetitive LC-MS/MS analyses on depth of analysis for complex proteomes. **J Proteome Res.** (2),1032-40, 2010.
- WATSON, J.T. & SPARKMAN, O.D. **Introduction to Mass Spectrometry: Instrumentation, Applications, and Strategies for Data Interpretation**, 4^a ed, John Wiley & Sons, Chichester, 2007.

- WEISKOPF, A.; SENEAL, K.; VOUIROS, P.; KAPLAN, D.; MELLO, C.M. The carbohydrate composition of a spider silk: *Nephila clavipes* Dragline. **Glycobiology** 6,1703-1703, 1996.
- WELLS, J.M. & MCLUCKEY, S.A. Collision-induced dissociation (CID) of peptides and proteins. **Methods Enzymol.** 402, 148-185, 2005.
- WIESNER, J.; PREMSLER, T.; SICKMANN, A. Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications. **Proteomics** 8, 4466-4483, 2008.
- WILKINS, M.R.; WILLIAMS, K.L.; APPEL, R.D.; HOCHSTRASSER, D.F. **Proteome research: new frontiers in functional genomics**. Springer, Berlin, 243, 1997.
- WINKLER, S.; WILSON, D.; KAPLAN, D.L. Controlling beta-sheet assembly in genetically engineered silk by enzymatic phosphorylation/dephosphorylation. **Biochemistry** 39,12739-12746, 2000.
- WITZE, E.S.; OLD, W.M.; RESING, K.A.; AHN, N.G. Mapping protein post-translational modifications with mass spectrometry. **Nat Methods**. 4(10),798-806, 2007.
- WOODS, A.G.; SOKOLOWSKA, I.; DARIE, C.C. Identification of consistent alkylation of cysteine-less peptides in a proteomics experiment. **Biochem Biophys Res Commun** 419(2),305-308, 2012.
- XIA, Q.; LEE, M.V.; ROSE, C.M.; MARSH, A.J. Characterization and diagnostic value of amino acid side chain neutral losses following electron-transfer dissociation. **J Am Soc Mass Spectrom.** 22:255-264, 2011.
- XIANG, Y.; SEKINE, T.; NAKAMURA, H.; IMAJOH-OHMI, S.; FUKUDA, H.; NISHIOKA, K.; KATO, T. Proteomic surveillance of autoimmunity in osteoarthritis: identification of triosephosphate isomerase as a na autoantigen in patients with osteoarthritis. **Arthritis Rheum.** 50, 1511-1521, 2004.
- XU, M.; LEWIS, R.V. Structure of a protein superfiber: spider dragline silk. **Proc Natl Acad Sci USA** 87,7120-7124, 1990.
- YANG, M.; ASAKURA, T. Design, expression and solid-state NMR characterization of silk-like materials constructed from sequences of spider silk, *Samia cynthia ricini* and *Bombyx mori* silk fibroins. **J. Biochem.** 137, 721, 2005.
- YAZAWA, K.; YAMAGUCHI, E.; KNIGHT, D.; ASAKURA, T. ^{13}C solid state NMR study of the ^{13}C -labeled peptide, $(\text{E})_8\text{GGLGGQGAG}(\text{A})_6\text{GGAGQGGYGG}$ as a model for the local structure of *Nephila clavipes* dragline silk (MaSp1) before and after spinning. **Biopolymers** 97(6),347-54, 2011.
- YI, Q.; ZHAO, P.; WANG, X.; ZOU, Y.; ZHONG, X.; WANG, C.; XIANG, Z.; XIA, Q.Y. Shotgun proteomic analysis of the *Bombyx mori* anterior silk gland: An insight into the biosynthetic fiber spinning process. **Proteomics**. 13(17),2657-63, 2013.
- ZHANG, P.B.; ASO, Y.; YAMAMOTO, K.; BANNO, Y.; WANG, Y.Q.; TSUCHIDA, K.; KAWAGUCHI, Y.; FUJII, H., Proteome analysis of silk gland proteins from the silkworm, *Bombyx mori*. **Proteomics** 6, 2586-2599, 2006.

APÊNDICES A

Material suplementar Tabelas S1, S2, S3 e S4

Tabela S1. Sequenciamento peptídico das proteínas spidroin-1 e spidroin-2 identificadas na 2-DE da seda do frame da aranha *N. clavipes* através de digestão *in-gel* utilizando diferentes enzimas proteolíticas. Número do spot, nome da proteína, enzima utilizada, posição na sequência, *m/z* observado, massa experimental, massa teórica, valor delta entre a massa experimental e a massa teórica, número de clivagens perdidas, sequência peptídica, PTMs, aminoácidos substitutos, método de fragmentação e ion score foram listados para todos os peptídeos.

Spot	Código de Acesso	Proteína	Enzima	Posição da sequência peptídica na proteína	Valor <i>m/z</i> observado (carga)	Valor massa experimental (Da)	Valor massa teórica (Da)	Delta	Clivagens perdidas	Sequência peptídica	PTMs a aminoácidos substitutos / Método de fragmentação	Ion score
1	P19837	Spidroin-1	Trypsin	30 - 69	1091.1425 (+3)	3270.4057	3270.4552	-0.0495	0	YGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	59
1	P19837	Spidroin-1	Trypsin	70 - 99	861.0592 (+3)	2580.1558	2580.1660	-0.0102	0	R.GGQGAGAAAAAAGGAGQGGY*GGLGSQGAGR.G	Phospho (Y89) / CID + ETD	42
1	P19837	Spidroin-1	Trypsin	134 - 161	844.1465 (+3)	2529.4177	2529.2361	0.1816	0	R.GGQGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	83
1	P19837	Spidroin-1	Trypsin	266 - 308	1164.1327 (+3)	3489.3763	3489.4965	-0.1202	0	R.GGEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	CID + ETD	74
1	P19837	Spidroin-1	Trypsin	398 - 427	856.1444 (+3)	2565.4114	2565.1582	0.2532	1	R.GGQGAGAAAAAAGGAGQRGYGGGLGNQGAGR.G	CID + ETD	76
1	P19837	Spidroin-1	Trypsin	462 - 488	730.3604 (+3)	2188.0594	2188.0006	0.0588	0	R.GGQGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	63
1	P19837	Spidroin-1	Trypsin	572 - 585	632.8975 (+2)	1263.7804	1263.5956	0.1848	0	R.QGGYGGLGSQGAGR.G	CID	67
1	P19837	Spidroin-1	Trypsin	616 - 652	954.7554 (+3)	2861.2444	2861.3401	-0.0957	0	R.GGLGGQGAGAAAAGGAGQGGYGGVGSASAASAAASR.L	CID + ETD	65
1	P19837	Spidroin-1	Chymotrypsin	1 - 21	885.0600 (+2)	1768.1054	1767.9214	0.1840	1	QGAGAAAAAAGGAGQGGYGGL.G	CID	80
1	P19837	Spidroin-1	Chymotrypsin	19 - 34	653.3969 (+2)	1304.7792	1304.6110	0.1683	2	Y.GGLGGQGAGQGGYGGL.G	CID	85
1	P19837	Spidroin-1	Chymotrypsin	22 - 31	975.2430 (+1)	974.2357	974.4255	-0.1897	0	L.GGQGAGQGGY.G	CID	70
1	P19837	Spidroin-1	Chymotrypsin	32 - 59	710.9552 (+3)	2129.8438	2129.9839	-0.1401	1	Y.GGLGGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	73
1	P19837	Spidroin-1	Chymotrypsin	35 - 59	696.6243 (+3)	2086.8511	2086.8399	0.0112	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	91
1	P19837	Spidroin-1	Chymotrypsin	60 - 89	787.2978 (+3)	2358.8716	2359.1014	-0.2298	1	Y.GGLGSQAGRGGQGAGAAAAAAGGAGQGGY.G	CID	47
1	P19837	Spidroin-1	Chymotrypsin	60 - 92	863.1161 (+3)	2586.3265	2586.2284	0.0981	2	Y.GGLGSQAGRGGQGAGAAAAAAGGAGQGGYGGL.G	CID + ETD	56
1	P19837	Spidroin-1	Chymotrypsin	63 - 89	787.4326 (+3)	2359.2760	2359.1993	0.0767	0	L.GSQAGRGGQGAGAAAAAAGGAGQGGY.G	ETD	69
1	P19837	Spidroin-1	Chymotrypsin	93 - 123	1216.1916 (+2)	2430.3686	2430.1749	0.1938	1	L.GSQAGRGGLGGQGAGAAAAAAGGAGQ*GGY.G	Gln->Lys (Q120) / CID	98
1	P19837	Spidroin-1	Chymotrypsin	124 - 151	1144.9919 (+2)	2287.9692	2288.0642	-0.0950	1	Y.GGLGN*QGAGRGGQGAAAAAAGGAGQGGY.G	Hydroxymethyl (N128) / CID	65
1	P19837	Spidroin-1	Chymotrypsin	127 - 151	710.9552 (+3)	2129.8438	2130.0063	-0.1626	0	L.GNQG*AGRGGQGAAAAAAGGAGQGGY.G	Gly->Arg (G130) / CID + ETD	83
1	P19837	Spidroin-1	Chymotrypsin	152 - 184	853.0140 (+3)	2556.0202	2556.2178	-0.1976	2	Y.GGLGSQAGRGGLGGQGAGAAAAAAGGAGQGGY.G	ETD	53
1	P19837	Spidroin-1	Chymotrypsin	155 - 184	787.6600 (+3)	2359.9582	2360.0854	-0.1272	1	L.GSQAGR*GGLGGQGAGAAAAAAGGAGQGGY.G	Deamidated (R161) / CID + ETD	61
1	P19837	Spidroin-1	Chymotrypsin	165 - 184	917.5339 (+2)	1833.0532	1832.8693	0.1839	0	L.GGQGAGAAAAAAGGAGQGGY.G	CID	59
1	P19837	Spidroin-1	Chymotrypsin	232 - 255	611.6497 (+3)	1831.9273	1831.8198	0.1075	0	L.GGQGAGQGAGASAAAAGGAGQGGY.G	ETD	70
1	P19837	Spidroin-1	Chymotrypsin	256 - 298	1140.5500 (+3)	3418.6282	3418.5748	0.0534	3	Y.GGLGSQAGRGGEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	ETD	112
1	P19837	Spidroin-1	Chymotrypsin	259 - 285	1072.9525 (+2)	2143.8904	2143.8556	0.0348	0	L.GS*QGAGR*GGEGAGAAAAAAGGAGQGGY.G	Phospho (S260); Arg->Ser (R265) / CID	31
1	P19837	Spidroin-1	Chymotrypsin	259 - 288	863.1161 (+3)	2586.3265	2586.1630	0.1635	1	L.GSQAGRGGEGAGAAAAAAGGAGQGGYGGL.G	CID + ETD	59
1	P19837	Spidroin-1	Chymotrypsin	312 - 353	1067.8700 (+3)	3200.5882	3200.4355	0.1527	1	L.GGQ*GAGAAAAGGAGQ*GGLGGQ*GAGQ*GAGAAAAAAGGAGQGGY.G	Methyl+Deamidated (Q314; Q326; Q332; Q336) / CID + ETD	67
1	P19837	Spidroin-1	Chymotrypsin	327 - 353	696.6243 (+3)	2086.8511	2086.9893	-0.1382	1	Y.GGLGGQGAGQ*GAGAAAAAAGGAGQGGY.G	Gln->Arg (Q336) / CID + ETD	68
1	P19837	Spidroin-1	Chymotrypsin	330 - 353	916.9295 (+2)	1831.8444	1831.8198	0.0247	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	93
1	P19837	Spidroin-1	Chymotrypsin	421 - 451	1215.9900 (+2)	2429.9654	2430.1385	-0.1730	1	L.GN*QGAGRGGLGGQGAGAAAAAAGGAGQGGY.G	Asn->Ser (N422) / CID	64
1	P19837	Spidroin-1	Chymotrypsin	455 - 478	1049.6085 (+2)	2097.2024	2097.0536	0.1488	0	L.GNQAGRGGQGAAAAAAGGAGQGGY.G	CID	96
1	P19837	Spidroin-1	Chymotrypsin	637 - 653	703.3476 (+2)	1404.6806	1404.6957	-0.0151	0	Y.GGVGSASAASAA*ASRL.S	Ala->Ser (A649) / CID	56
1	P19837	Spidroin-1	Chymotrypsin	654 - 680	854.7849 (+3)	2561.3329	2561.2569	0.0760	1	L.SSPEASSRVSSAVSNLVSSGPTNSAAL.S	ETD	58
1	P19837	Spidroin-1	Chymotrypsin	654 - 680	843.8329 (+3)	2528.4769	2528.2831	0.1938	1	L.SSPQASSRVSS*AVSNLVASGPTNSAAL.S	Ser->Ala (S664) / ETD	71
1	P19837	Spidroin-1	Chymotrypsin	699 - 709	666.7959 (+2)	1331.5772	1331.7241	-0.1468	2	L.SGCDVLIQALL.E	CID	88
1	P19837	Spidroin-1	Chymotrypsin	720 - 729	424.4347 (+3)	1270.2823	1270.4156	-0.1333	0	L.GSSSIGQVNY.G	CID + ETD	65
1	P19837	Spidroin-1	Chymotrypsin	730 - 744	718.3276 (+2)	1434.6406	1434.7103	-0.0697	0	Y.GSAGQAT*QIVGQSVY.Q	Thr->Ala (T736) / CID + ETD	56
1	P19837	Spidroin-1	Pespin	709 - 715	730.5411 (+1)	729.5338	729.4272	0.1066	1	L.LEVVSAL.I	CID	57

2	P19837	Spidroin-1	Trypsin	47 - 69; 139-161; 341-363; 375-397; 466-488	1067.4369 (+2)	2132.8592	2132.9124	-0.0532	0	AAAAAGGAGQGGY*GGLGSQGAGR.G	Phospho (Y59; Y151; Y353; Y387; Y478) / CID + ETD	31
2	P19837	Spidroin-1	Trypsin	70 – 99	903.4444 (+3)	2707.3114	2707.3418	-0.0304	0	R.GGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	ETD	85
2	P19837	Spidroin-1	Trypsin	134 - 161	739.6358 (+3)	2215.8856	2216.0683	-0.1827	0	R.GGQ*GAAAAAAGGAGQGGYGGLGSQGAGR.G	Gln->Leu (Q136) / ETD	72
2	P19837	Spidroin-1	Trypsin	266 - 308	1140.5853 (+3)	3418.7341	3418.6112	0.1229	0	R.GGE*GAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	Glu->Lys (E268) / CID + ETD	106
2	P19837	Spidroin-1	Trypsin	398 - 427	777.5744 (+4)	3106.2685	3106.2193	0.0492	1	R.GGQGAGAAAAAAGGAGQRGYGGGLGNQGAGR.G	ETD	53
2	P19837	Spidroin-1	Trypsin	462 -488	739.6358 (+3)	2215.8856	2216.0319	-0.1463	0	R.GGQGAAAAAAGGAGQGGYGGLGSQGAGR.G	ETD	68
2	P19837	Spidroin-1	Trypsin	616 - 652	954.7546 (+3)	2861.2420	2861.3401	-0.0981	0	R.GGLGGQGAGAAAAGGAGQGGYGGVGSASAASAAASR.L	CID + ETD	101
2	P19837	Spidroin-1	Trypsin	653 - 661	366.2066 (+3)	1095.5980	1095.4134	0.1845	0	R.LSSPQASSR.V	CID + ETD	60
2	P19837	Spidroin-1	Chymotrypsin	19 -31	539.8101 (+2)	1077.6056	1077.4840	0.1217	1	Y.GGLGGQGAGQGGY.G	CID	88
2	P19837	Spidroin-1	Chymotrypsin	32 – 59	720.4140 (+3)	2158.2202	2158.0264	0.1938	1	Y.GGLGGQGAGQ*GAGAAAAAAGGAGQGGY.G	Gln->Arg (Q41) / CID + ETD	76
2	P19837	Spidroin-1	Chymotrypsin	32 – 59	720.3643 (+3)	2158.0711	2158.0152	0.0559	1	Y.GGLGGQ*GAGQ*GAGAAAAAAGGAGQGGY.G	Methyl (Q37; Q41) / CID + ETD	65
2	P19837	Spidroin-1	Chymotrypsin	35 – 59	711.0872 (+3)	2130.2398	2130.0818	0.1580	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	73
2	P19837	Spidroin-1	Chymotrypsin	60 – 89	793.0484 (+3)	2376.1234	2376.1358	-0.0125	1	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY.G	ETD	79
2	P19837	Spidroin-1	Chymotrypsin	63 – 89	811.0855 (+3)	2430.2347	2430.1676	0.0670	0	L.GSQGAGRGGQGAGAAAAAAGGAGQGGY.G	ETD	56
2	P19837	Spidroin-1	Chymotrypsin	63 – 92	863.1161 (+3)	2586.3265	2586.3263	0.0002	1	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGL.G	CID + ETD	59
2	P19837	Spidroin-1	Chymotrypsin	93 - 123	1216.1916 (+2)	2430.3686	2430.1749	0.1938	1	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQ*GGY.G	Gln->Lys (Q120) / CID + ETD	98
2	P19837	Spidroin-1	Chymotrypsin	127 - 151	687.2925 (+3)	2058.8557	2058.9580	-0.1023	0	L.GNQ*GAGRGGQ*GAAAAAAGGAGQGGY.G	Methyl (Q129; Q136) / CID + ETD	70
2	P19837	Spidroin-1	Chymotrypsin	127 - 154	796.0260 (+3)	2385.0562	2385.1170	-0.0608	1	L.GNQGAGRGGQGAAAAAAGGAGQGGYGGL.G	ETD	97
2	P19837	Spidroin-1	Chymotrypsin	152 - 184	904.0123 (+3)	2709.0151	2709.2117	-0.1967	2	Y.GGLGSQGAGRGGL*GGQGAGAAAAAAGGAGQGGY.G	Leu->Arg (L164) / CID + ETD	82
2	P19837	Spidroin-1	Chymotrypsin	155 - 197	1164.4833 (+3)	3490.4281	3490.5959	-0.1678	3	L.GSQGAGRGGGLGGQGA*G*AAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Gly->Glu (G170) / CID + ETD	59
2	P19837	Spidroin-1	Chymotrypsin	162 - 184	873.8873 (+2)	1745.7600	1745.8081	-0.0481	1	Y.GGLGGQGAGAAAAAAGGAGQGGY.G	CID	120
2	P19837	Spidroin-1	Chymotrypsin	232 - 255	916.9248 (+2)	1831.8350	1831.8198	0.0153	0	L.GGQGAGQGAGASAAAAGGAGQGGY.G	CID	113
2	P19837	Spidroin-1	Chymotrypsin	256 - 298	1140.5000 (+3)	3418.4782	3418.6112	-0.1330	3	Y.GGLGSQGAGRGGE*GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Glu->Lys (E268) / CID + ETD	147
2	P19837	Spidroin-1	Chymotrypsin	259 - 285	787.2978 (+3)	2358.8716	2359.0360	-0.1644	0	L.GSQGAGRGGEGAGAAAAAAGGAGQGGY.G	ETD	71
2	P19837	Spidroin-1	Chymotrypsin	324 - 356	886.7305 (+3)	2657.1697	2657.2576	-0.0879	2	AGQGGGLGGQGAGQGAGAAAAAAGGAGQGGYGGL.G	CID + ETD	50
2	P19837	Spidroin-1	Chymotrypsin	330 - 353	916.9295 (+2)	1831.8444	1831.8198	0.0247	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID	93
2	P19837	Spidroin-1	Chymotrypsin	455 - 478	711.0872 (+3)	2130.2398	2130.0679	0.1719	0	L.GNQGAGRGGQGAAAAAGGAGQGGY.G	ETD	54
2	P19837	Spidroin-1	Chymotrypsin	606 - 636	825.0261 (+3)	2472.0565	2472.1855	-0.1290	2	Y.GGLGG*QGVGRGGLGGQGAGAAAAGGAGQGGY.G	Gly->Ser (G610) / CID + ETD	93
2	P19837	Spidroin-1	Chymotrypsin	654 - 680	843.7867 (+3)	2528.3383	2528.2831	0.0552	1	L.SSPQASSRV*S*SAVSNLVASGPTNSAAL.S	Ser->Ala (S663) / CID + ETD	61
2	P19837	Spidroin-1	Chymotrypsin	670 - 698	1033.6792 (+3)	3098.0158	3098.2090	-0.1932	1	L.VASGPTNSAALSSTISNVVSIQASNPGL.S	CID + ETD	63
2	P19837	Spidroin-1	Chymotrypsin	699 - 709	666.7959 (+2)	1331.5772	1331.7241	-0.1468	2	L.SGCDVLIQALL.E	CID	84
2	P19837	Spidroin-1	Chymotrypsin	716 - 729	739.8900 (+2)	1477.7672	1477.7776	-0.0104	1	L.IQILGSSSIGQVNY.G	CID	73
2	P19837	Spidroin-1	Chymotrypsin	730 - 744	740.4366 (+2)	1478.8586	1478.7365	0.1221	0	Y.GSAGQATQIVGQSV*Y.Q	Val->Ile (V743) / CID + ETD	84
3	P19837	Spidroin-1	Trypsin	70 - 99	777.4058 (+3)	2329.1956	2329.0908	0.1048	0	R.GGQGAGAAAAAAGGAGQGGYGGLGS*QGAGR.G	Ser->Gly (S94) / CID + ETD	61
3	P19837	Spidroin-1	Trypsin	134 - 161	811.0803 (+3)	2430.2191	2430.0462	0.1729	0	R.GGQGAAAAAAGGAGQGGY*GGLGSQGAGR.G	Phospho (Y151) / CID + ETD	42
3	P19837	Spidroin-1	Trypsin	167 - 207	1164.5409 (+3)	3490.6009	3490.5134	0.0875	0	QGTDAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQSGR.G	ETD	74
3	P19837	Spidroin-1	Trypsin	462 - 488	730.6673 (+3)	2188.9801	2189.0448	-0.0647	0	R.GGQGAAAAAAGGAGQGGYGGLGSQGAGR.G	ETD	76
3	P19837	Spidroin-1	Trypsin	489 - 509	582.6598 (+3)	1744.9576	1744.8912	0.0664	0	R.GGQGAGAAAAAAGVAGQEGIR.G	CID + ETD	85
3	P19837	Spidroin-1	Trypsin	586 - 615	777.4058 (+3)	2329.1956	2329.0908	0.1047	0	R.GGQGAGAAAAAAGGAGQGGYGGLGGQG*V*GR.G	Val->Ala (V613) / CID + ETD	99
3	P19837	Spidroin-1	Trypsin	616 - 652	954.8393 (+3)	2861.4961	2861.3401	0.1560	0	R.GGLGGQGAGAAAAGGAGQGGYGGVGSASAASAAASR.L	CID + ETD	75
3	P19837	Spidroin-1	Chymotrypsin	01 - 21	885.0556 (+2)	1768.0966	1767.9214	0.1752	1	QGAGAAAAAAGGAGQGGYGGL.G	CID	94
3	P19837	Spidroin-1	Chymotrypsin	11 - 18	769.1602 (+1)	768.1529	768.2861	-0.1332	0	Y.*GQGQGGY.G	Gly->Cys (G11) / CID	74

3	P19837	Spidroin-1	Chymotrypsin	22 - 31	938.3310 (+1)	937.3237	937.3712	-0.0475	0	L.GGQGAGQGGY.G	CID	95
3	P19837	Spidroin-1	Chymotrypsin	22 - 34	1106.6645 (+1)	1105.6572	1105.4789	0.1783	1	L.GGQGAGQGGYGGL.G	CID	75
3	P19837	Spidroin-1	Chymotrypsin	32 - 62	796.0260 (+3)	2385.0562	2385.1422	-0.0860	2	Y.GGLGGQ*GAGQ*GAGAAAAAAGGAGQGGYGGL.G	Methyl (Q37; Q41) / CID + ETD	79
3	P19837	Spidroin-1	Chymotrypsin	32 - 59	1080.4482 (+2)	2158.8818	2159.0230	-0.1411	1	Y.GGLGGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	76
3	P19837	Spidroin-1	Chymotrypsin	35 - 59	696.5679 (+3)	2086.6819	2086.8399	-0.1580	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	63
3	P19837	Spidroin-1	Chymotrypsin	60 - 89	787.2978 (+3)	2358.8716	2359.1014	-0.2298	1	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY.G	CID	97
3	P19837	Spidroin-1	Chymotrypsin	63-89	738.0579 (+3)	2213.1739	2213.1319	0.0420	0	L.GS*QGAGRGGQGAGAAAAAAGGAGQGGY.G	Phospho (S64) / CID	41
3	P19837	Spidroin-1	Chymotrypsin	90 - 123	957.0334 (+3)	2868.0784	2868.2520	-0.1736	2	Y.GGLGSQGAGRGGGLGGQGAGAAAAAAGGAGQGGY.G	CID + ETD	81
3	P19837	Spidroin-1	Chymotrypsin	103 - 123	810.9276 (+2)	1619.8406	1619.7288	0.1118	0	L.GGQGAGAA*AAAAAGGAGQGGY.G	Ala->Thr (A110) / CID	76
3	P19837	Spidroin-1	Chymotrypsin	127 - 154	796.0260 (+3)	2385.0562	2385.1170	-0.0608	1	L.GNQGAGRGGQGAAAAAAGGAGQGGYGGL.G	CID + ETD	97
3	P19837	Spidroin-1	Chymotrypsin	152 - 184	904.0123 (+3)	2709.0151	2709.2117	-0.1967	2	Y.GGLGSQGAGRGGL*GGQGAGAAAAAAGGAGQGGY.G	Leu->Arg (L164) / CID + ETD	62
3	P19837	Spidroin-1	Chymotrypsin	155 - 197	1164.4833 (+3)	3490.4281	3490.5959	-0.1678	3	L.GSQGAGRGGGLGGQGA*AAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Gly->Glu (G170) / ETD	59
3	P19837	Spidroin-1	Chymotrypsin	165 - 184	873.9798 (+2)	1745.9450	1745.9061	0.0390	0	L.GGQGAGAAAAAAGGAGQGGY.G	CID + ETD	116
3	P19837	Spidroin-1	Chymotrypsin	256 - 298	1140.7931 (+3)	3419.3575	3419.5588	-0.2013	3	Y.GGLGSQGAGRGGEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	CID + ETD	53
3	P19837	Spidroin-1	Chymotrypsin	259 - 285	715.6513 (+3)	2143.9321	2143.8556	0.0764	0	L.GS*QGAGR*GEGAGAAAAAAGGAGQGGY.G	Phospho (S260); Arg->Ser (R265) / CID	31
3	P19837	Spidroin-1	Chymotrypsin	312 - 353	1049.8049 (+3)	3146.3929	3146.4570	-0.0641	1	L.GGQGAGAAAAGGAGQGGGLGGQGAGQAGAAAAAAGGAGQGGY.G	CID + ETD	85
3	P19837	Spidroin-1	Chymotrypsin	327 - 353	1044.3793 (+2)	2086.7440	2086.9893	-0.2452	1	Y.GGLGGQGAGRGAAGAAAAAAGGAGQGGY.G	CID	71
3	P19837	Spidroin-1	Chymotrypsin	354 - 387	904.0445 (+3)	2709.1117	2709.2716	-0.1600	2	Y.GGLGSQGAGRGGL*GGQGAGAVAAAAAGGAGQGGY.G	Leu->His (L366) / CID + ETD	81
3	P19837	Spidroin-1	Chymotrypsin	388 -417	787.2978 (+3)	2358.8716	2359.1014	-0.2298	1	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY.G	ETD	67
3	P19837	Spidroin-1	Chymotrypsin	388 - 420	1036.5219 (+3)	3106.5439	3106.3812	0.1626	2	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQRGYGGL.G	CID + ETD	61
3	P19837	Spidroin-1	Chymotrypsin	421 - 451	839.4113 (+3)	2515.2121	2515.1548	0.0572	1	L.GNQGAGRGGGLGGQGAGAA*AAAAGGAGQGGY.G	Ala->Glu (A439) / CID + ETD	73
3	P19837	Spidroin-1	Chymotrypsin	455 - 478	1009.3915 (+2)	2016.7684	2016.9110	-0.1426	0	L.GNQGAGRGGQGAAAAAGGAGQGGY.G	CID	77
3	P19837	Spidroin-1	Chymotrypsin	529 - 550	873.8873 (+2)	1745.7600	1745.8081	-0.0481	1	Y.GGLGGQGAGAAAAAAGGAGQGGY.G	CID	120
3	P19837	Spidroin-1	Chymotrypsin	549 - 575	1044.3793 (+2)	2086.7440	2086.9893	-0.2452	1	Y.GGLGGQGAGRGAAGAAAAAAGGAGQGGY.G	CID + ETD	71
3	P19837	Spidroin-1	Chymotrypsin	576 -605	787.2978 (+3)	2358.8716	2359.1014	-0.2298	1	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY.G	CID + ETD	87
3	P19837	Spidroin-1	Chymotrypsin	619 - 636	794.4931 (+2)	1586.9716	1586.8053	0.1664	0	L.GGQGAGAAAAGGAGQGGY.G	CID	99
3	P19837	Spidroin-1	Chymotrypsin	654 - 680	844.1017 (+3)	2529.2833	2529.3035	-0.0202	1	L.SSPQ*ASSRVSSAVSNLVASGPTNSAAL.S	Gln->Leu (Q657) / CID + ETD	66
3	P19837	Spidroin-1	Chymotrypsin	670 - 698	1033.6792 (+3)	3098.0158	3098.2090	-0.1932	1	L.VASGPTNSAALSSTISNVVSIQIGASNPL.S	CID + ETD	63
3	P19837	Spidroin-1	Chymotrypsin	716 - 729	739.8909 (+2)	1477.7672	1477.7776	-0.0104	1	L.IQILGSSSIGQVNY.G	CID	73
3	P19837	Spidroin-1	Chymotrypsin	730 - 744	740.4366 (+2)	1478.8586	1478.7365	0.1221	0	Y.GSAGQATQIVGQSV*Y.Q	Val->Ile (V743) / CID + ETD	84
3	B5SYS7	Spidroin-2	Trypsin	85 – 112	974.0588 (+3)	2919.1546	2919.3486	-0.1940	0	K.LQALNM*AFASSM*AEIAAVEQGM*SM*AVK.T	Oxidation (M90, M96, M107, M109) / CID	46
3	B5SYS7	Spidroin-2	Trypsin	113 - 140	766.5621 (+4)	3062.2193	3062.4114	-0.1921	0	K.TNAIVDGLNSAFYM*TTGAANPQFVN*EMR.S	Oxidation (M126); Methyl + Deamidated (N137) / CID + ETD	44
3	B5SYS7	Spidroin-2	Trypsin	113 - 140	766.5621 (+4)	3062.2193	3062.4114	-0.1921	0	K.TNAIVDGLN*SAFYMTTGAANPQFVNEM*R.S	Oxidation (M139); Methyl + Deamidated (N121) / CID + ETD	35
3	B5SYS7	Spidroin-2	Trypsin	113 - 140	1022.0964 (+3)	3063.2674	3063.4066	-0.1392	0	K.TNAIVDGLNSAFYM*TTGAANPQFVNEM*R.S	Oxidation (M126, M139) / CID + ETD	91
3	B5SYS7	Spidroin-2	Trypsin	221 - 244	708.0834 (+3)	2121.2284	2121.0351	0.1932	0	R.QQAPSGPAAAAAAGAGGYGPAQR.G	CID + ETD	70
3	B5SYS7	Spidroin-2	Pepsin	44 – 52	766.3304 (+1)	765.3231	765.3657	-0.0426	0	L.AAVSGSGAF.T	CID	31
3	P46804	Spidroin 2	Chymotrypsin	13-35, 261-283	983.8948 (+2)	1965.775	1965.8105	-0.0355	0	G.GYGPGQQGPSGPGSAAAAAAAAA.A	CID + ETD	68
3	P46804	Spidroin 2	Chymotrypsin	229-252	1025.8929 (+2)	2049.7712	2049.8793	-0.1081	0	Q.GPGGYGPGQ*QGLSGPGS*AAAAAAA.G	Deamidation (Q237); Phospho (S245) / CID	49
3	P46804	Spidroin 2	Chymotrypsin	270-306	1054.5132 (+3)	3160.5178	3160.3636	0.1542	0	P.SGPGSAAAAAAAAAGPGGYGPGQQGPGGYGPGQQGPS.G	CID + ETD	46
3	P46804	Spidroin 2	Chymotrypsin	322-356	827.0589 (+4)	3304.2065	3304.1949	0.0116	0	Q.QGLGGYGPGQQGPGGYGPGQQGPGGYGPGSASAAA.A	ETD	43

3	P46804	Spidroin 2	Chymotrypsin	350-390	1130.7759 (+3)	3389.3059	3389.4433	-0.1374	0	P.GSASAAAAAAGPGQQGPGGYGPGQQGPSGPGSASAAAAAAA.A	CID	59
3	P46804	Spidroin 2	Chymotrypsin	365-386	636.9284 (+3)	1907.7634	1907.7687	-0.0053	0	Q.GPGGYGPGQQGPSGPGSASAAA.A	CID + ETD	42
3	P46804	Spidroin 2	Chymotrypsin	371-389, 408-426	816.8076 (+2)	1631.6006	1631.6941	-0.0934	0	G.PGQQGPSGPGSASAAAAA.A	CID	36
3	P46804	Spidroin 2	Chymotrypsin	378-396, 415-433	777.9294 (+2)	1553.8442	1553.6511	0.1931	0	S.GPGSASAAAAAAGPGGY.G	CID	60
3	P46804	Spidroin 2	Chymotrypsin	438-464	773.9981 (+3)	2318.9725	2318.9805	-0.008	0	Q.QGPGGYAPGQQGPSGPGSAAAAAAA.G	CID + ETD	39
3	P46804	Spidroin-2	Chymotrypsin	534 - 560	820.4852 (+3)	2458.3938	2458.2199	0.1738	1	L.ASPDSGARVASAVSNLVSSGPTSSAAL.S	CID + ETD	65
3	P46804	Spidroin-2	Chymotrypsin	542-560	859.3886 (+2)	1716.7626	1716.8894	-0.1267	0	R. VASAVSNLVSSGPTSSAAL.S	CID + ETD	85
3	P46804	Spidroin-2	Chymotrypsin	616 - 624	450.8070 (+2)	899.5994	899.5076	0.0918	0	F.AQVVGQSVL.S	CID	44
3	P46804	Spidroin-2	Pepsin	168-189, 209-230	955.3342 (+2)	1908.6538	1908.8715	-0.1176	0	A. AAAASGPGQQGPGGYGPGQQGP.G	CID	58
3	P46804	Spidroin 2	Pepsin	568-584	837.3651 (+2)	1672.7156	1672.8091	-0.0934	0	A.VSQIGASNPGLSGCDVL.I	CID	56
3	P46804	Spidroin-2	Pepsin	600 - 615	801.9045 (+2)	1601.7944	1601.7322	0.0623	1	L.SSSSIGQVNYGAASQF.A	CID	73
3	P46804	Spidroin-2	Pepsin	616 - 624	450.8063 (+2)	899.5980	899.5076	0.0904	0	F.AQVVGQSVL.S	CID	49
3	P46804	Spidroin-2	Proteinase 10	3-26	1072.410 (+2)	2142.805	2142.935	-0.1301	0	G.GYGPQQGPGGYGPGQQGPSGPGS.A	CID	60
3	P46804	Spidroin-2	Proteinase 10	4-26	1043.960 (+2)	2085.905	2085.914	-0.0086	0	G.YGPGQQGPGGYGPGQQGPSGPGS.A	CID + ETD	77
3	P46804	Spidroin-2	Proteinase 10	4-27	1079.490 (+2)	2156.965	2156.951	0.0142	0	G.YGPGQQGPGGYGPGQQGPSGPGSA.A	CID	37
3	P46804	Spidroin-2	Proteinase 10	8-26	856.900 (+2)	1711.785	1711.755	0.0304	0	G.QQGPGGYGPGQQGPSGPGS.A	CID + ETD	96
3	P46804	Spidroin-2	Proteinase 10	14-25; 261-272; 369-380	1101.430 (+1)	1100.422	1100.488	-0.0660	0	G.YGPGQQGPSGPG.S	CID	23
3	P46804	Spidroin-2	Proteinase 10	50-72	1158.450 (+2)	2314.885	2314.915	-0.0304	0	G.GYGPQQGPGGRYGPGQQGPSGPG.S	CID + ETD	33
3	P46804	Spidroin-2	Proteinase 10	56-77	709.960 (+3)	2126.858	2126.857	0.0010	0	Q.QGPGRYGPGQQGPSGPGSAAAA.A	CID + ETD	21
3	P46804	Spidroin-2	Proteinase 10	61-89	888.350 (+3)	2662.028	2662.048	-0.0205	0	R.YGPGQQGPSGPGSAAAAAAGSQQGPGGY.G	CID + ETD	29
3	P46804	Spidroin-2	Proteinase 10	74-86	568.750 (+2)	1135.485	1135.465	0.0196	0	S.AAAAAAGSQQGP.G	CID	22
3	P46804	Spidroin-2	Proteinase 10	81-110	960.070 (+3)	2877.188	2877.221	-0.0335	0	G.SGQQGPGGYGPRQQGPGGYGQQQQGPSGPG.S	CID + ETD	34
3	P46804	Spidroin-2	Proteinase 10	109-119	924.410 (+1)	923.402	923.374	0.0278	0	G.PGSAAAA S *AAA.S	Phospho (S116) / CID	26
3	P46804	Spidroin-2	Proteinase 10	122-140	897.880 (+2)	1793.745	1793.700	0.0448	0	A.ESGQQGPGGYGPGQQGPGG.Y	CID	38
3	P46804	Spidroin-2	Proteinase 10	136-161	780.390 (+3)	2338.148	2338.036	0.1118	0	Q.QGPGGYGPGQQGPGGYGPGQQGPSGP.G	CID + ETD	24
3	P46804	Spidroin-2	Proteinase 10	171-190	877.350 (+2)	1752.685	1752.781	-0.0962	0	A.ASGPGQQGPGGYGPGQQGPG.G	CID	60
3	P46804	Spidroin-2	Proteinase 10	171-203	965.130 (+3)	2892.368	2892.281	0.0869	0	A.ASGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPG.S	CID + ETD	45
3	P46804	Spidroin-2	Proteinase 10	171-205	1017.780 (+3)	3050.318	3050.350	-0.0322	0	A.ASGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSA.A	CID + ETD	51
3	P46804	Spidroin-2	Proteinase 10	172-190	841.820 (+2)	1681.625	1681.744	-0.1191	0	A.SGPGQQGPGGYGPGQQGPG.G	CID + ETD	42
3	P46804	Spidroin-2	Proteinase 10	172-203	941.440 (+3)	2821.298	2821.244	0.0540	0	A.SGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPG.S	CID + ETD	57
3	P46804	Spidroin-2	Proteinase 10	172-205	994.050 (+3)	2979.128	2979.313	-0.1851	0	A.SGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPGSA.A	CID + ETD	16
3	P46804	Spidroin-2	Proteinase 10	251-273	989.920 (+2)	1977.825	1977.893	-0.0675	0	A.AAGPGQQGPGGYGPGQQGPSGPG.S	CID	36
3	P46804	Spidroin-2	Proteinase 10	251-274	1033.380 (+2)	2064.745	2064.925	-0.1795	0	A.AAGPGQQGPGGYGPGQQGPSGPGS.A	CID	20
3	P46804	Spidroin-2	Proteinase 10	368-391	986.380 (+2)	1970.745	1970.908	-0.1628	0	G.GYGPQQGPSGPGSASAAAAAAA.G	CID	25
3	P46804	Spidroin-2	Proteinase 10	383-392	811.390 (+1)	810.382	810.327	0.0555	0	A. S *AAAAAAG.P	Phospho (S383) / CID	35
3	P46804	Spidroin-2	Proteinase 10	401-417	756.280 (+2)	1510.545	1510.680	-0.1347	0	Q.QGPGGYAPGQQGPSGPG.S	CID	28
3	P46804	Spidroin-2	Proteinase 10	401-419	835.340 (+2)	1668.665	1668.749	-0.0838	0	Q.QGPGGYAPGQQGPSGPGSA	CID	19
3	P46804	Spidroin-2	Proteinase 10	492-513	1007.530 (+2)	2013.045	2012.862	0.1825	0	G.GYGPAQQGPAGYGPGSAVAASA.G	CID	27
3	P46804	Spidroin-2	Proteinase 10	532-556	902.770 (+3)	2705.288	2705.090	0.1981	0	S.RLASPSGARVASAVSNLVSSGPTS.S	CID + ETD	29
3	P46804	Spidroin-2	Proteinase 10	537-549	663.900 (+2)	1325.785	1325.597	0.1878	0	P.DSGARVASAVSNL.V	CID	33
3	P46804	Spidroin-2	Proteinase 10	552-560	870.360 (+1)	869.352	869.353	-0.0004	0	S.SGPT T *SSAAL.S	Phospho (T555) / CID	24

3	P46804	Spidroin-2	Proteinase 10	593-624	1183.640 (+3)	3547.538	3547.478	0.0596	0	V.SACVTILS S*S*S *IGQVNYGAAS S *QFAQVVGQSVL.S	Phospho (S601, S602, S603, S613) / CID + ETD	27
3	P46804	Spidroin-2	Proteinase 10	602-621	1086.050 (+2)	2170.085	2169.913	0.1721	0	S.SSIGQVNY*GAAS S *QFAQVVGQ.S	Phospho (Y609, S613) / CID	21
3	P46804	Spidroin-2	Proteinase 10	606-626	762.060 (+3)	2283.158	2282.997	0.1608	0	G.QVNYGAASQFAQVVGQSVLSA.F	CID + ETD	32
4	P19837	Spidroin-1	Trypsin	70 - 99	844.0748 (+3)	2529.2026	2529.1494	0.0532	0	R.GGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	ETD	91
4	P19837	Spidroin-1	Trypsin	134 - 161	811.0614 (+3)	2430.1624	2430.0462	0.1162	0	R.GGQGAAAAAAGGAGQGGY Y *GGLGSQGAGR.G	Phospho (Y151) / CID + ETD	34
4	P19837	Spidroin-1	Trypsin	462 -488	739.6358 (+3)	2215.8856	2216.0319	-0.1463	0	R.GGQGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	68
4	P19837	Spidroin-1	Trypsin	616 - 652	954.7713 (+3)	2861.2921	2861.3401	-0.0480	0	R.GGLGGQGAGAAAAAGGAGQGGYGGVGSGASAASAAASR.L	CID + ETD	68
4	P19837	Spidroin-1	Chymotrypsin	19 – 31	539.8064 (+2)	1077.5982	1077.4840	0.1143	1	Y.GGLGGQGAGQGGY.G	CID	64
4	P19837	Spidroin-1	Chymotrypsin	19 – 34	653.3912 (+2)	1304.7678	1304.6110	0.1569	2	Y.GGLGGQGAGQGGYGGL.G	CID	81
4	P19837	Spidroin-1	Chymotrypsin	32 – 59	711.3251 (+3)	2130.9535	2130.9679	-0.0144	1	Y.GGLGGQGAGQGAGAAAAAAGGAG Q *GGY.G	Deamidated (Q56) / ETD	57
4	P19837	Spidroin-1	Chymotrypsin	60 – 89	787.4159 (+3)	2359.2259	2359.1014	0.1245	1	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY.G	CID + ETD	74
4	P19837	Spidroin-1	Chymotrypsin	63 – 89	787.4159 (+3)	2359.2259	2359.1993	0.0266	0	L.GSQGAGRGGQGAGAAAAAAGGAGQGGY.G	ETD	77
4	P19837	Spidroin-1	Chymotrypsin	70 – 89	817.8326 (+2)	1633.6506	1633.7478	-0.0972	0	L.GGQGAGAAAAAAGGAGQGGY.G	CID	60
4	P19837	Spidroin-1	Chymotrypsin	90 - 102	543.8757 (+2)	1085.7368	1085.5578	0.1791	1	Y.GGLGSQGAGRSSL.G	CID	54
4	P19837	Spidroin-1	Chymotrypsin	127 - 154	754.4114 (+3)	2260.2124	2260.0217	0.1907	1	L.GN Q *GAGRGGQAAAAAAGGAG Q *GGYGGL.G	Deamidated (Q129; Q148) / ETD	68
4	P19837	Spidroin-1	Chymotrypsin	155-184	826.0866 (+3)	2475.2380	2475.1434	0.0946	1	L.G S *QGAG R *GGLGGQGAGAAAAAAGGAGQGGY.G	Phospho (S156) / CID	48
4	P19837	Spidroin-1	Chymotrypsin	162 - 184	873.8873 (+2)	1745.7600	1745.8081	-0.0481	1	Y.GGLGGQGAGAAAAAAGGAGQGGY.G	CID	120
4	P19837	Spidroin-1	Chymotrypsin	165-187	604.3130 (+3)	1809.9172	1809.7432	0.1740	1	L.GGQGAGAAAAAAGGAGQGGY Y *GGL*.G	Phospho (Y184); Leu->Pro (L187) / CID	47
4	P19837	Spidroin-1	Chymotrypsin	232 - 255	711.5478 (+3)	2131.6216	2131.7793	-0.1577	0	L.GGQGAG Q *GAGASAAAGGAGQGGY.G	Gln->Leu (Q238) / CID + ETD	61
4	P19837	Spidroin-1	Chymotrypsin	256 - 298	1140.5642 (+3)	3418.6708	3418.6112	0.0596	3	Y.GGLGSQGAGRGG E *GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Glu->Lys (E268) / CID + ETD	98
4	P19837	Spidroin-1	Chymotrypsin	259 - 285	1067.0010 (+2)	2131.9874	2132.0108	-0.0233	0	L.GSQGAGRGGEGAGAAAAAAGGAGQGGY.G	CID	103
4	P19837	Spidroin-1	Chymotrypsin	327 - 353	711.3251 (3)	2130.9535	2130.9679	-0.0144	1	Y.GGLGG Q G *AGQGAGAAAAAAGGAGQGGY.G	Gly->Glu (G333) / CID + ETD	53
4	P19837	Spidroin-1	Chymotrypsin	330 – 353	612.2660 (+2)	1833.7762	1833.7878	-0.0116	0	L.GGQGAG Q *GAGAAAAAAGGAG Q *GGY.G	Deamidated (Q336; Q350) / CID	74
4	P19837	Spidroin-1	Chymotrypsin	330 – 353	917.3265 (+2)	1832.6384	1832.8038	-0.1653	0	L.GGQGAG Q *GAGAAAAAAGGAGQGGY.G	Deamidated (Q336) / CID	65
4	P19837	Spidroin-1	Chymotrypsin	367 – 387	817.8326 (+2)	1633.6506	1633.7445	-0.0938	0	L.GGQGAG A *VAAAAAGGAGQGGY.G	Ala->Ser (A373) / CID	66
4	P19837	Spidroin-1	Chymotrypsin	455 – 478	678.3532 (+3)	2032.0378	2031.9107	0.1271	0	L.GNQGAGRGG Q G *AAAAAGGAGQGGY.G	Gly->Glu (G465) / CID + ETD	61
4	P19837	Spidroin-1	Chymotrypsin	462 – 478	817.8326 (+2)	1633.6506	1633.5606	0.0900	0	L. S * G Q *GAAAAAGGAGQGGY.G	Phospho (S462); Deamidation (Q464) / CID + ETD	54
4	P19837	Spidroin-1	Pepsin	716 - 737	730.0071 (+3)	2186.9995	2187.0920	-0.0925	1	L.IHILGSSSIGQVNYGSAGQATQ	CID + ETD	71
4	B5SYS7	Spidroin-2	Trypsin	77 - 84	421.7651 (+3)	1262.2735	1262.4226	-0.1491	1	R.SNKSSQHK.L	/ CID + ETD	33
4	B5SYS7	Spidroin-2	Trypsin	85 -112	968.7642 (+3)	2903.2708	2903.3537	-0.0829	0	K.LQALN M *AFASSMAEIAAVEQGG M * SM *AVK.T	Oxidation (M90, M107, M109) / CID + ETD	56
4	B5SYS7	Spidroin-2	Trypsin	113 - 140	1022.0810 (+3)	3063.2212	3063.4066	-0.1933	0	K.TNAIVDGLNSAFY M *TTGAAN R *PQFVN M *R.S	Oxidation (M126, M139); Deamidated (N132) / CID + ETD	44
4	B5SYS7	Spidroin-2	Trypsin	221 - 244	708.0703 (+3)	2121.1891	2121.0351	0.1539	0	R.QQAPSGPAAAAAAGAGGYGPAQR.G	CID + ETD	72
4	B5SYS7	Spidroin-2	Chymotrypsin	30 - 43	757.3006 (+2)	1512.5866	1512.7096	-0.1230	2	W.SDTATADAFIQNFL.A	CID	56
4	B5SYS7	Spidroin-2	Chymotrypsin	126 - 142	933.8878 (+2)	1865.7610	1865.8764	-0.1153	1	Y.MTTGAANPQFVNEMRSL.I	CID	58
4	B5SYS7	Spidroin-2	Chymotrypsin	218 - 249	979.1485 (+3)	2934.4237	2934.4485	-0.0248	1	Y.GPRQQAPSGPAAAAAAGAGGYGPAQRGPSVY.G	CID + ETD	52
4	B5SYS7	Spidroin-2	Pepsin	86 - 92	405.6899 (+2)	809.3652	809.3742	-0.0089	1	L.QALN M *AF.A	Oxidation (M90) / CID	51
4	B5SYS7	Spidroin-2	Pepsin	125 - 135	608.7779 (+2)	1215.5412	1215.5230	0.0182	0	F.Y M *TTGAANPQF.V	Oxidation (M126) / CID	45
4	P46804	Spidroin 2	Chymotrypsin	22-41	813.4248 (+2)	1624.835	1624.6882	0.1468	0	P. S *GPGSAAAAAAAAAAGPGGY.G	Phospho (S22) / CID	62
4	P46804	Spidroin-2	Chymotrypsin	52 - 99	1499.0216 (+3)	4494.0430	4493.9313	0.1117	2	Y.GPGQQGPGRYGPGQQGPGSPGSAAAAAAGSGQQGPGGYGPRQQGPGGY.G	CID + ETD	34
4	P46804	Spidroin 2	Chymotrypsin	378-396, 415-433	777.9027 (+2)	1553.7908	1553.6511	0.1397	0	S.GPGSASAAAAAAGPGGY.G	CID	77

4	P46804	Spidroin-2	Chymotrypsin	534 - 560	820.4442 (+3)	2458.3108	2458.2299	0.0808	1	L.ASPDSGARVASAVSNLVSSGPTSSAAL.S	CID + ETD	94
4	P46804	Spidroin 2	Pepsin	34-46, 281-293, 389-401, 426-438	565.7719 (+2)	1129.5292	1129.5152	0.014	0	A.AAAGPGGYGPGQQ.G	CID	40
4	P46804	Spidroin 2	pepsin	95-107	635.755 (+2)	1269.4954	1269.4663	0.0291	0	Q.GPGGYGQGQQGPS.G	CID	40
4	P46804	Spidroin-2	Pepsin	550 - 560	606.7945 (+2)	1211.5744	1211.4005	0.1740	0	L.VSSGPTSSAAL.S	CID	31
4	P46804	Spidroin 2	Pepsin	570-579	943.385 (+1)	942.3777	942.4771	-0.0993	0	S.QIGASNPGLS.G	CID	38
4	P46804	Spidroin-2	Proteinase 10	3-22; 181-200	923.370 (+2)	1844.725	1844.807	-0.0824	0	G.GYGPQQGPGGYGPGQQGPS.G	CID + ETD	48
4	P46804	Spidroin-2	Proteinase 10	3-25	1028.920 (+2)	2055.825	2055.903	-0.0781	0	G.GYGPQQGPGGYGPGQQGPSGPG.S	CID + ETD	83
4	P46804	Spidroin-2	Proteinase 10	3-26	1072.420 (+2)	2142.825	2142.935	-0.1101	0	G.GYGPQQGPGGYGPGQQGPSGPGS.A	CID	35
4	P46804	Spidroin-2	Proteinase 10	4-22; 141-159; 182-200	894.880 (+2)	1787.745	1787.786	-0.0409	0	G.YGPQQGPGGYGPGQQGPS.G	CID	35
4	P46804	Spidroin-2	Proteinase 10	4-27	1079.440 (+2)	2156.865	2156.951	-0.0858	0	G.YGPQQGPGGYGPGQQGPSGPGSA.A	CID	48
4	P46804	Spidroin-2	Proteinase 10	8-25; 363-380	813.310 (+2)	1624.605	1624.723	-0.1176	0	G.QQGPGGYGPGQQGPSGPG.S	CID + ETD	52
4	P46804	Spidroin-2	Proteinase 10	9-24; 364-379	720.800 (+2)	1439.585	1439.643	-0.0575	0	Q.QGPGGYGPGQQPSGP.G	CID + ETD	23
4	P46804	Spidroin-2	Proteinase 10	13-26; 261-274	623.270 (+2)	1244.525	1244.542	-0.0168	0	G.GYGPQQGPSGPGS.A	CID	20
4	P46804	Spidroin-2	Proteinase 10	14-25; 261-272; 369-380	1101.460 (+1)	1100.452	1100.488	-0.0360	0	G.YGPQQGPSGPG.S	CID	26
4	P46804	Spidroin-2	Proteinase 10	14-26; 261-273	1188.540 (+1)	1187.532	1187.520	0.0120	0	G.YGPQQGPSGPGS.A	CID	24
4	P46804	Spidroin-2	Proteinase 10	56-63	911.360 (+1)	910.352	910.369	-0.0171	0	Q.QGPGRY*GP.G	Phospho (Y61) / CID	36
4	P46804	Spidroin-2	Proteinase 10	171-191	905.860 (+2)	1809.705	1809.803	-0.0976	0	A.ASGPGQQGPGGYGPGQQPGG.Y	CID	56
4	P46804	Spidroin-2	Proteinase 10	171-203	965.070 (+3)	2892.188	2892.281	-0.0931	0	A.ASGPGQQGPGGYGPGQQPGGYGPGQQGPSGPG.S	ETD	68
4	P46804	Spidroin-2	Proteinase 10	172-190	841.840 (+2)	1681.665	1681.744	-0.0791	0	A.SGPGQQGPGGYGPGQQGPG.G	CID + ETD	45
4	P46804	Spidroin-2	Proteinase 10	172-203	941.410 (+3)	2821.208	2821.244	-0.0360	0	A.SGPGQQGPGGYGPGQQGPGGYGPGQQGPSGPG.S	CID + ETD	84
4	P46804	Spidroin-2	Proteinase 10	251-274	1033.410 (+2)	2064.805	2064.925	-0.1195	0	A.AAGPGQQGPGGYGPGQQGPSGPGS.A	CID	52
4	P46804	Spidroin-2	Proteinase 10	381-414	988.730 (+3)	2963.168	2963.283	-0.1154	0	G.SASAAAAAAGPGGYGPGQQGPGGYAPGQQGPS.G	CID + ETD	23
4	P46804	Spidroin-2	Proteinase 10	401-418	799.840 (+2)	1597.665	1597.712	-0.0467	0	Q.QGPGGYAPGQQGPSGPGS.A	CID + ETD	22
4	P46804	Spidroin-2	Proteinase 10	513-545	739.850 (+4)	2955.370	2955.347	0.0237	0	S.AGAGSAGYGPSQASAAASRLASPDSGARVASA.V	CID + ETD	30
4	P46804	Spidroin-2	Proteinase 10	602-621	1086.040 (+2)	2170.065	2169.913	0.1521	0	S.SSIGQVNY*GAAS*QFAQVVGQ.S	Phospho (Y609, S613) / CID	29

Tabela S2. Caracterização MS/MS através do Modiro®. Sequenciamento das proteínas spidroin-1 e spidroin-2 identificadas na 2-DE da seda do frame da aranha *N. clavipes* através de digestão *in-gel* utilizando diferentes enzimas proteolíticas. Número do spot, nome da proteína, enzima utilizada, posição na sequência, *m/z* observado, *m/z* teórico, valor delta entre o *m/z* observado e o *m/z* teórico, valor da carga do íon precursor, sequência peptídica, PTMs, aminoácidos substitutos, método de fragmentação, score significativo e ion score foram listados para todos os peptídeos.

Spot	Código de acesso	Proteína	Enzima	Posição da sequência peptídica na proteína	Valor <i>m/z</i> observado	Valor <i>m/z</i> teórico	Delta	Z	Sequência peptídica	PTMs e aminoácidos substitutos / Método de fragmentação	Ion Score	Score Sig.
1	P19837	Spidroin-1	Trypsin	266 – 308	1140.73	1140.53	0.1978	3	R.GGE ^{-Q-} GAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	Substitution Glu->Gln (E268) / CID + ETD	446	100
1	P19837	Spidroin-1	Trypsin	266 – 308	862.82	862.90	-0.0848	4	R.GGEGAGAAAAAAGGAGQGGYGGLGG ^{Me} GAG ^{Me} GGYGGLGSQGAGR.G	Methylation (Q291; Q295) / ETD	451	100
1	P19837	Spidroin-1	Trypsin	266 – 308	862.82	862.90	-0.0848	4	R.GGEGAGAAAAAAGGAG ^{Me} GGYGGLGG ^{Me} GAGQGGYGGLGSQGAGR.G	Methylation (Q282; Q291) / ETD	409	99.1
1	P19837	Spidroin-1	Trypsin	266 – 308	862.82	862.90	-0.0848	4	R.GGEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR ^{Me2} .G	Dimethylation (R 308) / ETD	368	91.9
1	P19837	Spidroin-1	Chymotrypsin	32 – 59	720.17	720.34	-0.1757	3	Y.GGLGG ^{Me} GAG ^{Me} GAGAAAAAAGGAGQGGY.G	Methylation (Q37; Q41) / CID + ETD	339	100
1	P19837	Spidroin-1	Chymotrypsin	155 – 197	1140.77	1140.86	-0.0902	3	L.GSQGAGRGGLGGQGAGAAAAAAGGAG ^{Deamid} GGYGGLGGQGAGQGGY.G	Deamidation (Q181) / CID + ETD	374	100
1	P19837	Spidroin-1	Chymotrypsin	579 – 605	712.14	711.99	0.1466	3	L.GSQGAGR ^{Deamid} GGQGAGAAAAAAGGAGQGGY.G	Deamidation (R585) / CID + ETD	292	99.9
1	P19837	Spidroin-1	Chymotrypsin	579 – 608	795.94	795.94	-0.1063	3	L.GSQGA ^{-P-} GRGGQGAGAAAAAAGGAGQGGYGGL.G	Substitution Ala->Pro (A583) / CID + ETD	318	99.9
1	P19837	Spidroin-1	Chymotrypsin	730 – 744	740.44	740.37	0.0645	2	Y.GSAGQATQIVGQS ^{Me} VY.Q	Methylation (S742) / CID	382	100
1	P19837	Spidroin-1	Pepsin	155 – 187	1017.43	1017.4167	0.0133	3	L.GSQGAGRGGLGGQGAGAAAAAAGGAGQGGYGG ^{463.04} L.G	Unknown shift / CID + ETD	353	100
1	P19837	Spidroin-1	Pepsin	211 – 258	890.66	890.6748	-0.0148	4	L.GGQGAGAAAAAAGGAGQGGGLGGQGAG ^{-38.05} QGASAAAAGGAGQGGYGGL.G	Unknown shift / ETD	345	99
1	P19837	Spidroin-1	Pepsin	259 – 288	768.31	768.3691	-0.0591	3	L.GSQGAGRGGEGAGAAAAAAGGAGQGGYGG ^{-58.17} L.G	Unknown shift / ETD	283	95.4
1	P19837	Spidroin-1	Pepsin	421 – 454	881.73	881.7661	-0.0361	3	L.GNQGAGRGGLGGQGAGAAAAAAGGAGQGG ^{-42.1} YGGL.G	Unknown shift / CID + ETD	353	99.9
1	P19837	Spidroin-1	Pepsin	609 – 653	954.34	954.1918	0.1482	4	L.GGQGVGRGGLGGQGAGAAAAAGGAGQGGYGGVGSASASAAASR ^{227.59} L.S	Unknown shift / ETD	292	91.3

1	P19837	Spidroin-1	Pepsin	709 – 715	730.51	730.4345	0.0755	1	L.LEVVSAL.I	- CID	255	99.9
2	P19837	Spidroin-1	Trypsin	70 – 99	795.33	795.36	-0.0360	3	R.GGQGAGAAAAAGGAGQGGYGGL ^{~H} GSQGAGR.G	Substitution Leu->His (L92) / CID + ETD	240	91.2
2	P19837	Spidroin-1	Chymotrypsin	32 – 59	720.17	720.34	-0.1757	3	Y.GGLGGQ ^{Me} GAGQ ^{Me} GAGAAAAAAGGAGQGGY.G	Methylation (Q37; Q41) / CID + ETD	339	100
2	P19837	Spidroin-1	Chymotrypsin	60 – 89	796.55	796.71	-0.1682	3	Y.GGLGSQGAGR ^{Me} GGQ ^{Me} GAGAAAAAAGGAGQGGY.G	Methylation (R69; Q72) / CID + ETD	441	99.9
2	P19837	Spidroin-1	Chymotrypsin	63 – 92	795.94	796.04	-0.1063	3	L.GSQGA ^{~P} GRGGQGAGAAAAAAGGAGQGGYGGL.G	Substitution Ala->Pro (A67) / CID	318	99.9
2	P19837	Spidroin-1	Chymotrypsin	155 – 197	1140.77	1140.86	-0.0902	3	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQ ^{Deamid} GGYGGGQGAGQGGY.G	Deamidation (Q181) / CID + ETD	374	100
2	P19837	Spidroin-1	Chymotrypsin	165 – 184	540.13	540.26	-0.1309	3	L.G ^{~R} GQGAGAAAAAAGGAGQGGY.G	Substitution Gly->Arg (G165) / CID + ETD	320	94.9
2	P19837	Spidroin-1	Chymotrypsin	232 – 258	692.92	692.98	-0.0625	3	L.GGQGAGQGAGASAAAAGGAGQ ^{Deamid} GGYGGL.G	Deamidation (Q252) / ETD	268	87.7
2	P19837	Spidroin-1	Chymotrypsin	522 – 551	772.87	773.04	-0.1763	3	L.GSQSGRGGGLGGQGAGAAAA ^{-9.82} AAGGAGQGGL.G	Unknown shift / ETD	203	87.5
2	P19837	Spidroin-1	Chymotrypsin	606 – 636	816.22	816.04	0.1704	3	Y.GGLGGQ ^{Deamid} GVGRGGLGGQ ^{Deamid} GAGAAAAGGAGQ ^{Deamid} GGY.G	Deamidation (Q611; Q621; Q633) / CID + ETD	220	86.7
2	P19837	Spidroin-1	Chymotrypsin	730 – 747	739.86	739.64	0.2192	3	Y.GSAGQATQIVGQSVYQ ^{439.66} AL.G	Unknown shift / ETD	349	80.8
2	P19837	Spidroin-1	Pepsin	22 - 34; 188 - 200; 289 - 301	581.8	581.7492	0.0508	2	L.GGQGAGQG ^{84.1} GYGGL.G	Unknown shift / CID	200	80.3
2	P19837	Spidroin-1	Pepsin	63 – 102	1073.23	1073.1811	0.0489	3	L.GSQGAGRGGQGAGAAAA ^{17.15} AAGGAGQGGYGGLGSQGAGRGG.L	Unknown shift / ETD	289	99
2	P19837	Spidroin-1	Pepsin	201 – 231	811	810.827	0.173	3	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQGGL ^{49.82} .G	Unknown shift / CID + ETD	467	99.9
2	P19837	Spidroin-1	Pepsin	259 – 288	798.04	798.0357	0.0043	3	L.GSQGAGRGGEGAGAAAAAGG ^{31.01} AGQGGYGGL.G	Unknown shift / ETD	219	82.5
2	P19837	Spidroin-1	Pepsin	302 – 329	1075.93	1076.0312	-0.1012	2	L.G ^{-17.19} SQGAGRGGGLGGQGAGAAAAGGAGQGGL.G	Unknown shift / CID	226	85.1
2	P19837	Spidroin-1	Pepsin	391 – 420	837.63	837.4009	0.2291	3	L.GSQGAGRGGQGAGAAAAAAGGAGQRGYG ^{51.69} GL.G	Unknown shift / CID + ETD	297	87.8
2	P19837	Spidroin-1	Pepsin	421 – 454	698.36	698.3264	0.0336	4	L.GNQGAGRGGGLGGQGAGAAAAAAGGAGQGGYGG ^{105.13} L.G	Unknown shift / ETD	259	84.2
3	P19837	Spidroin-1	Trypsin	266 – 308	873.87	873.90	-0.0322	4	R.GGEGAGAAAAAG ^{~E} GAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	Substitution Gly->Glu (G278) / ETD	428	99.9
3	P19837	Spidroin-1	Trypsin	266 – 308	862.82	862.90	-0.0848	4	R.GGEGAGAAAAAAGGAGQGGYGGLGGQ ^{Me} GAGQ ^{Me} GGYGGLGSQGAGR.G	Methylation (Q291; Q295) / ETD	451	100
3	P19837	Spidroin-1	Trypsin	266 – 308	862.82	862.90	-0.0848	4	R.GGEGAGAAAAAAGGAGQ ^{Me} GGYGGLGGQ ^{Me} GAGQGGYGGLGSQGAGR.G	Methylation (Q282; Q291) / ETD	409	99.1
3	P19837	Spidroin-1	Chymotrypsin	21 – 59	1064.42	1064.48	-0.0677	3	Y.GGLGGQGAGQ ^{Deamid} GGYGGLGGQGAGQGAGAAAAAAGGAGQGGY.G	Deamidation (Q30) / CID + ETD	210	87
3	P19837	Spidroin-1	Chymotrypsin	93 – 123	612.02	612.04	-0.0258	4	L.GSQ ^{Me} GAGRGGGLGGQGAGAAAAAAGGAGQGGY.G	Methylation (Q95) / ETD	304	97.4
3	P19837	Spidroin-1	Chymotrypsin	127 – 154	754.41	754.34	0.0622	3	L.GNQGAGRGGQ ^{Deamid} GAAAAAAGGAGQ ^{Deamid} GGYGGL.G	Deamidation (Q136; Q148) / CID	286	99.9
3	P19837	Spidroin-1	Chymotrypsin	155 – 197	856.04	855.89	0.1430	4	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQ ^{Deamid} GGYGGLGGQGAGQGGY.G	Deamidation (Q181) / ETD	381	100
3	P19837	Spidroin-1	Chymotrypsin	155 – 197	856.04	855.89	0.1430	4	L.GSQGAGR ^{Deamid} GGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Deamidation (R161) / ETD	379	100
3	P19837	Spidroin-1	Chymotrypsin	330 – 353	917.33	917.40	-0.0792	2	L.GGQGAGQ ^{Deamid} GAGAAAAAAGGAGQGGY.G	Deamidation (Q336) / CID	438	100
3	P19837	Spidroin-1	Chymotrypsin	330 – 353	612.27	612.26	0.0001	3	L.GGQGAGQ ^{Deamid} GAGAAAAAAGGAGQ ^{Deamid} GGY.G	Deamidation (Q336; Q350) / CID + ETD	251	92
3	P19837	Spidroin-1	Chymotrypsin	482 – 518	1026.93	1027.17	-0.2403	3	L.GSQGAGRGGQGAGAAAAAAGGAG ^{-49.71} QEGIRQQAGQGGY.G	Unknown shift / CID + ETD	256	92.5
3	P19837	Spidroin-1	Pepsin	1-34	871.3939	871.4201	-0.0263	3	>.QGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY ^{-80.07} GGL.G	Unknown shift / CID + ETD	273	97.6
3	P19837	Spidroin-1	Pepsin	22 – 62	836.1741	836.1216	0.0525	4	L.GGQGAGQGGYGGLGGQGAGQGAGAAAAAAGGAG ^{151.21} QGGYGGL.G	Unknown shift / ETD	217	95.6
3	P19837	Spidroin-1	Pepsin	63 – 102	814.6628	814.6377	0.0251	4	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGLG ^{55.1} SQGAGRGL.G	Unknown shift / ETD	234	91.9
3	P19837	Spidroin-1	Pepsin	165 – 210	895.4569	895.6827	-0.2258	4	L.GGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGL ^{-67.89} GSQGAGRGL.G	Unknown shift / ETD	245	97.3
3	P19837	Spidroin-1	Pepsin	188 - 210; 289 - 311	938.8642	938.9594	-0.0952	2	L.GGQGAGQGGYGGLGSQG ^{-42.18} AGRGL.G	Unknown shift / CID	218	98.8
3	P19837	Spidroin-1	Pepsin	259 – 288	843.7384	843.7024	0.036	3	L.G ^{168.11} SQAGRGGEAGAAAAAAGGAGQGGYGGL.G	Unknown shift / CID + ETD	210	93.6
3	P19837	Spidroin-1	Pepsin	431 – 481	995.4478	995.4701	-0.0223	4	L.GGQGA ^{-8.08} GAAAAAAGGAGQGGYGGLGNQGAGRGGQGAAAAAGGAGQGGYGGL.G	Unknown shift / ETD	265	94
3	P19837	Spidroin-1	Pepsin	482 – 521	695.2188	695.1305	0.0882	5	L.GSQGAG ^{116.44} RGQQGAGAAAAAAGGAGQEGIRQQAGQGGYGGL.G	Unknown shift / ETD	326	95.6
3	P19837	Spidroin-1	Pepsin	609 – 653	927.7086	927.6918	0.0168	4	L.GGQGVGRGGLGGQGAGAAAAGGAGQGGYGGVSGASASAAASRL ^{121.07} .S	Unknown shift / ETD	241	94.1
3	P19837	Spidroin-1	Pepsin	699 – 708	460.7017	460.7655	-0.0638	2	L.SGCDVLI ^{-98.12} QAL.L	Unknown shift / CID	274	95
3	B5SYS7	Spidroin 2	Trypsin	77 - 84	478.65	478.73	-0.08	2	R.SNKSSQH ^{39.83} K.L	Unknown shift / CID	218	89.2
3	B5SYS7	Spidroin-2	Trypsin	85 - 112	974.06	974.12	-0.0635	3	K.LQALNMAFAS ^S ^{Dihydroxy} MAEIAAVEQGGM ^{Ox} SM ^{Ox} AVK.T	Dihydroxy (S95) / CID	269	99.8
3	B5SYS7	Spidroin-2	Trypsin	85 - 112	974.04	974.12	-0.08	3	K.LQALN ^{Ox} MAFASSM ^{Ox} AEIAAVEQGGM ^{Ox} SM ^{Ox} AVK.T	Hydroxylation (N89) / CID + ETD	402	100
3	B5SYS7	Spidroin-2	Trypsin	113 - 140	1017.09	1017.13	-0.0492	3	K.TNAIVDGLNSAFYM ^{Ox} TTGAANPQFVNEM ^R ^{Deamid} .S	Deamidation (R140) / CID	397	100.0
3	B5SYS7	Spidroin-2	Trypsin	113 - 140	766.56	766.61	-0.0529	4	K.TNAIVDGLNSAFY ^{Amino} MTTGAANPQFVNEM ^{Ox} R.S	Amino (Y125); Oxidation (M139) / ETD	337	92.7
3	B5SYS7	Spidroin 2	Trypsin	191 - 232	1227.33	1227.25	0.0757	3	A.APSGYGPSQQGPSGSAAAAAAPSgyGPRQQAPSGPAAAAA.A	No enzyme (two ends) / CID	250	86.7
3	B5SYS7	Spidroin 2	Trypsin	221 - 244	708.06	708.01	0.041	3	R.QQAPSGPAAAAAAGAGGYGPAQR.G	- / CID + ETD	518	100
3	B5SYS7	Spidroin 2	Trypsin	239 - 268	847.52	847.41	0.1059	3	G.YGPAQRGPSVYGPSGPGAAAAAAGAGPGG.Y	No enzyme (two ends) / CID + ETD	251	81.6
3	B5SYS7	Spidroin 2	Trypsin	245 - 276	1001.16	1001.08	0.0739	3	R.GPSVYGPSGPGAAAAAAGAGPGGYGPGQQ ^{349.22} GP.-	Unknown shift / CID + ETD	258	86.3

3	P46804	Spidroin 2	Chymotrypsin	5-41	1088.2	1088.14	0.0586	3	Y.GPGQQGPGGYGPGQQGPSGPSASAAAAAAA ^{197,18} AAGPGGY.G	Unknown shift / CID + ETD	281	82.4
3	P46804	Spidroin 2	Chymotrypsin	5 - 14; 42 - 51; 132 - 141; 142 - 151; 183 - 192; 224 - 233; 289 - 298; 328 - 337; 338 - 347; 397 - 406; 434 - 443	819.96	819.98	-0.02	1	Y.G ^{-97,44} PGQQGPGGY.G	Unknown shift / CID	211	81.5
3	P46804	Spidroin-2	Chymotrypsin	42 - 61	737.64	737.65	-0.0177	3	Y.GPGQQGPGGYGPGQQGPGRY.G	CID	241	87.9
3	P46804	Spidroin 2	Chymotrypsin	62 - 89	836.57	836.35	0.2118	3	Y.G ^{167,63} PGQQGPSGPSASAAAAAGSGQQGPGGY.G	Unknown shift / CID + ETD	270	91
3	P46804	Spidroin 2	Chymotrypsin	62 - 99	918.04	917.88	0.1513	4	Y.GPGQQGP ^{331,6} SGPGSAAAAAAGSGQQGPGGYGPRQQGPGGY.G	Unknown shift / ETD	289	85.3
3	P46804	Spidroin 2	Chymotrypsin	124 - 141; 134 - 151; 175 - 192; 216 - 233; 330 - 347	554.61	554.59	0.02	3	S.GQQGPGGYGPGQQGPGGY.G	- CID + ETD	284	82.2
3	P46804	Spidroin 2	Chymotrypsin	152 - 192; 193 - 233	878.21	878.148	0.062	4	Y.GPGQQGPSGPSASAAAAAASGPGQQGPGGYGPGQQG ^{pDihydroxy} GGY.G	Dihydroxy (P189, P230) / ETD	282	84.3
3	P46804	Spidroin 2	Chymotrypsin	234 - 262	841.77	841.72	0.0488	3	Y.G ^{86,15} PGQQGLSGPGSAAAAAAGPGQQGPGGY.G	Unknown shift / CID + ETD	210	88.3
3	P46804	Spidroin 2	Chymotrypsin	241 - 262	869.35	869.41	-0.06	2	L.SGPSASAAAAAGP ^{62,73} GQQGPGGY.G	Unknown shift / CID	268	98.7
3	P46804	Spidroin 2	Chymotrypsin	263 - 288	1000.26	1000.49	-0.2332	2	Y.G ^{-96,46} PGQQGPSGPSASAAAAAAGPGGY.G	Unknown shift / CID	210	82.7
3	P46804	Spidroin 2	Chymotrypsin	299 - 324	554.03	554.01	0.0177	4	Y.GPGQQ ^{79,07} GPSGASAAAAAAGPGQQGL.G	Unknown shift / ETD	332	87.2
3	P46804	Spidroin 2	Chymotrypsin	299 - 327	1479.18	1479.07	0.1097	2	Y.GPGQQGPSGASAAAAAAGPGQ ^{546,22} QGLGGY.G	Unknown shift / CID	240	85
3	P46804	Spidroin 2	Chymotrypsin	370 - 406	806.97	806.85	0.1134	4	Y.GPGQQGPSGPSASAAAAAAGP ^{143,45} GGYGPGQQGPGGY.A	Unknown shift / ETD	268	82.5
3	P46804	Spidroin 2	Chymotrypsin	407 - 443	814.96	814.86	0.0995	4	Y.APGQ ^{161,4} QGSPGPSASAAAAAAGPGGYGPGQQGPGGY.A	Unknown shift / ETD	311	87.2
3	P46804	Spidroin 2	Chymotrypsin	494 - 520	838.98	838.88	0.10	3	Y.G ^{237,87} PAQQGPAGYGPGSAVAASAGAGSAGY.G	Unknown shift / CID + ETD	327	85.1
3	P46804	Spidroin 2	Chymotrypsin	504 - 520	451.87	451.84	0.03	3	Y.GPGSAV ^{2,97} AASAGAGSAGY.G	Unknown shift / CID + ETD	243	97.4
3	P46804	Spidroin 2	Chymotrypsin	521 - 533	583.65	583.80	-0.1546	2	Y.G ^{-6,3} PGSQASAAASRL.A	Unknown shift / CID	242	89.8
3	P46804	Spidroin 2	Chymotrypsin	534 - 560	820.48	820.41	0.0627	3	L.ASPDSGARVASAVSNLVSSGPTSSAAL.S	- / CID + ETD	412	100
3	P46804	Spidroin 2	Chymotrypsin	534 - 549	940.34	940.32	0.02	2	L.A ^{377,91} SPDSGARVASAVSNL.V	Unknown shift / CID	226	80.1
3	P46804	Spidroin 2	Chymotrypsin	550 - 560	672.16	672.15	0.01	2	L.V ^{366,82} SSGPTSSAAL.S	Unknown shift / CID	201	83.5
3	P46804	Spidroin 2	Chymotrypsin	561 - 588	926.59	926.56	0.03	3	L.SSVISNAVSIQIGASNPGLSGCDV ^{77,37} LIQAL.L	Unknown shift / CID + ETD	383	85.6
3	P46804	Spidroin 2	Chymotrypsin	610 - 624	743.31	743.26	0.05	2	Y.GAASQF ^{23,84} AQVVGQSVL.S	Unknown shift / CID	226	87.5
4	P19837	Spidroin-1	Chymotrypsin	32 – 62	795.94	796.05	-0.1147	3	Y.GGLGQ ^{Me} GAG ^{QMe} GAGAAAAAAGGAGQGGYGGL.G	Methylation (Q37; Q41) / CID + ETD	362	100
4	P19837	Spidroin-1	Chymotrypsin	63 – 89	712.14	711.99	0.1466	3	L.GSQGAGRGG ^{QDeamid} GAGAAAAAAGGAGQGGY.G	Deamidation (Q72) / CID + ETD	300	100
4	P19837	Spidroin-1	Chymotrypsin	63 – 89	712.14	711.99	0.1466	3	L.GSQGAG ^{RDeamid} GGQGAGAAAAAAGGAGQGGY.G	Deamidation (R69) / CID + ETD	292	99.9
4	P19837	Spidroin-1	Chymotrypsin	63 – 92	795.94	796.04	-0.1063	3	L.GSQGA ^{-P-} GRGGQGAGAAAAAAGGAGQGGYGGL.G	Substitution Ala->Pro (A67) / CID + ETD	318	99.9
4	P19837	Spidroin-1	Chymotrypsin	155 – 197	856.04	855.89	0.1430	4	L.GSQGAGRGGLGGQGAGAAAAAAGGAG ^{QDeamid} GGYGGLGGQGAGQGGY.G	Deamidation (Q181) / ETD	381	100
4	P19837	Spidroin-1	Chymotrypsin	155 – 197	856.04	855.89	0.1430	4	L.GSQGAG ^{RDeamid} GGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Deamidation (R161) / ETD	379	100
4	P19837	Spidroin-1	Chymotrypsin	312 – 353	1052.72	1052.81	-0.0910	3	L.GGQGAGAAAAGGAGQGGGL ^{-Q-} GGQGAGQGAGAAAAAAGGAGQGGY.G	Substitution Leu->Glu (L329) / CID + ETD	261	99.9
4	P19837	Spidroin-1	Chymotrypsin	312 – 353	793.28	793.12	0.1557	4	L.GGQGAGAAAAGGAGQGGGLGG ^{QMe} GAG ^{QMe} GAGAAAAAAGGAGQGGY.G	Methylation (Q332; Q336) / ETD	438	100
4	P19837	Spidroin-1	Chymotrypsin	579 – 605	712.14	711.99	0.1466	3	L.GSQGAGRGG ^{QDeamid} GAGAAAAAAGGAGQGGY.G	Deamidation (Q588) / CID + ETD	300	100
4	P19837	Spidroin-1	Chymotrypsin	579 – 605	712.14	711.99	0.1466	3	L.GSQGAG ^{RDeamid} GGQGAGAAAAAAGGAGQGGY.G	Deamidation (R585) / CID + ETD	292	99.9
4	P19837	Spidroin-1	Chymotrypsin	579 – 608	795.94	795.94	-0.1063	3	L.GSQGA ^{-P-} GRGGQGAGAAAAAAGGAGQGGYGGL.G	Substitution Ala->Pro (A583) / CID + ETD	318	99.9
4	P19837	Spidroin-1	Chymotrypsin	730 – 744	740.44	740.37	0.0645	2	Y.GSAGQATQIVGQ ^{SMe} VY.Q	Methylation (S742) / CID	382	100
4	P19837	Spidroin-1	Pepsin	63 - 92; 579 - 608	600.2853	600.2826	0.0027	4	L.GSQGAGRGGQGAGAAA ^{38,01} AAAGGAGQGGYGGL.G	Unknown shift / ETD	227	92.7
4	P19837	Spidroin-1	Pepsin	155 – 200	1232.2649	1232.2412	0.0237	3	L.GS ^{48,07} QGAGRGGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGL.G	Unknown shift / CID + ETD	299	96.8
4	P19837	Spidroin-1	Pepsin	165 – 187	827.3855	827.4113	-0.0258	2	L.GGQGAGAAAAAAGG ^{-93,04} AGQGGYGGL.G	Unknown shift / CID	229	98.7
4	P19837	Spidroin-1	Pepsin	188 – 231	845.4148	845.4154	-0.0006	4	L.GGQGAGQGGYGGLGSQGA ^{-61,99} GRGGLGGQGAGAAAAAAGGAGQGGGL.G	Unknown shift / ETD	295	92.9
4	P19837	Spidroin-1	Pepsin	211 – 231	801.3928	801.3768	0.0161	2	L.GGQGAGAAAAAAGGA ^{61,03} GQGGL.G	Unknown shift / CID	203	99.7
4	P19837	Spidroin-1	Pepsin	232 – 258	1019.4247	1019.4781	-0.0534	2	L.GGQGAG ^{-38,1} QGAGASAAAAGGAGQGGYGGL.G	Unknown shift / CID	219	100
4	P19837	Spidroin-1	Pepsin	259 – 301	868.17	868.147	0.023	4	L.GSQGAGRGGEAGAGAAAAAAGGAGQGGYGGLGGQ ^{49,08} GAGQGGYGGL.G	Unknown shift / CID	277	94
4	P19837	Spidroin-1	Pepsin	312 – 356	1099.8506	1099.8619	-0.0113	3	L.GGQGAGAAAAGGAGQGGGLGGQGAGQGAGAAAAAAGGAGQ ^{-71,02} GGYGGL.G	Unknown shift / CID + ETD	201	93.3
4	P19837	Spidroin-1	Pepsin	357 – 390	884.7179	884.7729	-0.055	3	L.GS ^{-34,16} QGAGRGGLGGQGAGAVAAAAAGGAGQGGYGGL.G	Unknown shift / CID + ETD	254	94.9
4	P19837	Spidroin-1	Pepsin	579 – 618	855.9287	855.8929	0.0359	4	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGGQGVGRG ^{222,14} L.G	Unknown shift / ETD	244	93.7
4	B5SYS7	Spidroin 2	Trypsin	75 - 86	692.35	692.37	-0.0262	2	M.ARSNKSSQHLQ.A	No enzyme (two ends) / CID	297	89.6
4	B5SYS7	Spidroin 2	Trypsin	85 - 112	974.04	974.12	-0.0835	3	K.LQALN ^{Ox} AFASSM ^{Ox} AEIAAVEQGGM ^{Ox} SM ^{Ox} AVK.T	Oxidation (M90, M96, M107, M109) / CID + ETD	423	100
5	B5SYS7	Spidroin-2	Trypsin	85 - 112	974.05	974.12	-0.0735	3	K.LQALN ^{Ox} MAFASSM ^{Ox} AEIAAVEQGGM ^{Ox} SM ^{Ox} AVK.T	Hydroxylation (N89) / CID + ETD	428	100

4	B5SYS7	Spidroin-2	Trypsin	113 - 140	762.84	762.86	-0.0202	4	K.TNAIVDGLNSAFYMTTGAAN ^{OX} PQFVNEMR.S	Hydroxylation (N132) / ETD	390	99.8
4	B5SYS7	Spidroin 2	Trypsin	113 - 140	1022.41	1022.47	-0.0608	3	K.TNAIVDGLNSAFY ^{OX} TTGAANPQFVNE ^{OX} R ^{Deamid} .S	Oxidation (M126, M139); Deamidation (R140) / CID + ETD	282	99.4
4	B5SYS7	Spidroin 2	Trypsin	189 - 220	934.46	934.44	0.0139	3	A.AAAPSGYGPSQQGPSGSAAAAAAAAPSGYGPR.Q	No enzyme (one end) / CID + ETD	248	88
4	B5SYS7	Spidroin 2	Trypsin	221 - 244	750.26	750.01	0.2462	3	R.QQAPSGPAAAAA ^{152.74} AAGAGGYGPAQR.G	Unknown shift / CID + ETD	274	84
4	B5SYS7	Spidroin 2	Trypsin	245 - 276	1001.31	1001.08	0.2239	3	R.GPSVYGPSGPGAAAAAAGAPGGYGPGQQGP ^{349.67} .-	Unknown shift / CID + ETD	203	85.8
4	B5SYS7	Spidroin-2	Chymotrypsin	126 - 142	628.36	628.29	0.0623	3	Y. ^{OX} TTGAANPQFVNEMRSL.I	Oxidation (M126) / CID + ETD	397	100
4	P46804	Spidroin 2	Chymotrypsin	1-14	844.23	844.16	0.07	2	>.P ^{395.88} GGYGPGQQGPGGY.G	Unknown shift / CID	202	82.5
4	P46804	Spidroin 2	Chymotrypsin	5 - 14; 42 - 51; 132 - 141; 142 - 151; 183 - 192; 224 - 233; 289 - 298; 328 - 337; 338 - 347; 397 - 406; 434 - 443	876.94	876.98	-0.04	1	Y.GPGQQGP ^{-40.46} GGY.G	Unknown shift / CID	202	99.2
4	P46804	Spidroin 2	Chymotrypsin	15 - 41	846.7	846.67	0.0231	3	Y.G ^{371.07} PGQQGPSGPGSAAAAAAAAGPGGY.G	Unknown shift / CID + ETD	271	86.2
4	P46804	Spidroin 2	Chymotrypsin	15 - 61	999.51	999.72	-0.2161	4	Y.GPGQQGPSGP ^{-67.86} GSAAAAAAAAGPGGYGPGQQGPGGYGPGQQGPGRY.G	Unknown shift / ETD	249	82.4
4	P46804	Spidroin 2	Chymotrypsin	42 - 61	999.51	999.44	0.0643	2	Y.GPGQ ^{83.13} QGPGGYGPGQQGPGRY.G	Unknown shift / CID	214	89.9
4	P46804	Spidroin 2	Chymotrypsin	62 - 89	566.87	567.02	-0.1504	4	Y.GPGQQ ^{-75.59} GPSGPGSAAAAAAGSGQQGPGGY.G	Unknown shift / ETD	331	86.5
4	P46804	Spidroin 2	Chymotrypsin	100 - 131	768.52	768.31	0.2083	4	Y.GQQQQGPSGPGSAA ^{311.83} AASAAASAESGQQGPGGY.G	Unknown shift / ETD	285	81.3
4	P46804	Spidroin 2	Chymotrypsin	132 - 151; 328 - 347	598.66	598.93	-0.2797	3	Y.G ^{-21.83} PGQQGPGGYGPGQQGPGGY.G	Unknown shift / CID + ETD	291	95.6
4	P46804	Spidroin 2	Chymotrypsin	152 - 182; 193 - 223	873.48	873.40	0.0795	3	Y.GPGQQGPSGPGSAAA ^{39.24} AAAAAGPGQQGPGGY.G	Unknown shift / CID + ETD	291	88.1
4	P46804	Spidroin 2	Chymotrypsin	263 - 298	780.94	780.84	0.0914	4	Y.GPGQQGPSGPGSAAAAAAAAGPGGYGPGQQ ^{126.36} GPGGY.G	Unknown shift / ETD	273	87.1
4	P46804	Spidroin 2	Chymotrypsin	585 - 599	749.91	749.97	-0.06	2	L.IQALLEI ^{-144.52} VSAC ^{CAMe} VTIL.S	Unknown shift / CID	276	87.2
4	P46804	Spidroin 2	Chymotrypsin	289 - 324	1017.91	1017.81	0.0983	3	Y.GPGQQGPGGYG ^{19.29} PGQQGPSGAGSAAAAAAGPGQQGL.G	Unknown shift / CID + ETD	203	85.4
4	P46804	Spidroin 2	Chymotrypsin	289 - 327	841.27	841.13	0.1328	4	Y.GPGQQGPGGYGP ^{52.53} GQQGPSGAGSAAAAAAGPGQQGLGGY.G	Unknown shift / ETD	292	84.5
4	P46804	Spidroin 2	Chymotrypsin	299 - 327	899.86	899.71	0.144	3	Y.GPGQQ ^{286.43} GPSGAGSAAAAAAGPGQQGLGGY.G	Unknown shift / CID + ETD	224	92.1
4	P46804	Spidroin 2	Chymotrypsin	325 - 347	668.06	668.30	-0.2484	3	L.GGYGPGQQGPGGYG ^{-90.74} PGQQGPGGY.G	Unknown shift / CID + ETD	313	81.9
4	P46804	Spidroin 2	Chynotrypsin	338 - 369	749.23	749.33	-0.1001	4	Y.GPGQQGPGGYGPGSASAAAAAAGPGQQGPGGY.G	ETD	274	81.1
4	P46804	Spidroin 2	Chymotrypsin	348 - 369	874.8	874.91	-0.119	2	Y.GPGSASAAAAAAG ^{-51.23} PGQQGPGGY.G	Unknown shift / CID	228	89.8
4	P46804	Spidroin 2	Chymotrypsin	370 - 396	827.23	827.00	0.2214	3	Y.GPGQQGPSGPGSASAAAA ^{296.66} AAAAGPGGY.G	Unknown shift / CID + ETD	233	88.1
4	P46804	Spidroin 2	Chymotrypsin	444 - 469	548.76	548.75	0.0058	4	Y.A ^{82.02} PGQQGPSGPGSAAAAAAAAGPGGY.G	Unknown shift / ETD	301	80.9
4	P46804	Spidroin 2	Chymotrypsin	504 - 520	495.9	495.88	0.0190	3	Y.GPGSAV ^{135.06} AASAGAGSAGY.G	Unknown shift / CID + ETD	203	82.7
4	P46804	Spidroin 2	Chymotrypsin	550 - 578	1313.73	1313.78	-0.05	2	L.VSSGPTSSAA ^{-31.9} LSSVISNAVSIQASNPGL.S	Unknown shift / CID	222	80.3
4	P46804	Spidroin 2	Chymotrypsin	590 - 609	1036.3	1036.43	-0.1300	2	L.EIVSAC ^{CAMe} VTILSSSSIGQV ^{-55.46} NY.G	Unknown shift / CID	211	88
4	P46804	Spidroin 2	Chymotrypsin	600 - 609	540.42	540.24	0.1740	2	L.SSSS ^{38.35} IGQVNY.G	Unknown shift / CID	228	86.3
4	P46804	Spidroin 2	Chymotrypsin	610 - 624	831.39	831.38	0.0016	2	Y.GAASQF ²⁰⁰ AQVVGQSVL.S	Unknown shift / CID	221	88.3
4	P46804	Spidroin 2	Chymotrypsin	616 - 624	450.69	450.76	-0.0711	2	F.AQVVGQSVL.S	CID	527	100
4	P46804	Spidroin 2	Chymotrypsin	616 - 627	407.21	407.22	-0.0176	3	F.AQVVG ^{Me} SVLSAF.-	Methylation (Q621) / CID + ETD	217	86.8

Tabela S3. Sequenciamento peptídico das proteínas flageliform silk protein e spidroin-1 identificadas na 2-DE da seda da teia orbital da aranha *N. clavipes* através de digestão *in-gel* utilizando diferentes enzimas proteolíticas. Número do spot, nome da proteína, enzima utilizada, posição na sequência, *m/z* observado, massa experimental, massa teórica, valor delta entre a massa experimental e a massa teórica, número de clivagens perdidas, sequência peptídica, PTMs, aminoácidos substitutos, método de fragmentação e íon score foram listados para todos os peptídeos.

Spot	Código de acesso	Proteína	Enzima	Posição da sequência peptídica na proteína	Valor <i>m/z</i> observado (carga)	Valor massa experimental (Da)	Valor massa teórica (Da)	Delta	Clivagens perdidas	Sequência peptídica	PTMs e aminoácidos substitutos / Método de fragmentação	Íon score
1	O44358	Flageliform silk protein	Trypsin	14-38	834.7257 (+3)	2501.1553	2501.2795	-0.1243	0	K.AMQVALASSIAELVIAESSGGDVQR.K	CID + ETD	97
1	O44358	Flageliform silk protein	Trypsin	39-48	558.3191 (+2)	1114.6236	1114.6458	-0.0222	1	R.KTNVISNALR.N	CID + ETD	24
1	O44358	Flageliform silk protein	Trypsin	40-48	494.2738 (+2)	986.5330	986.5509	-0.0178	0	K.TNVISNALR.N	CID	43
1	O44359	Flageliform silk protein	Chymotrypsin	01-25	971.9553 (+2)	1941.8960	1941.8606	0.0354	1	Y.GPGGVGPGGSGPGGYGPGGAGPGGY.G	CID	97
1	O44359	Flageliform silk protein	Chymotrypsin	26 - 45; 36 - 55; 46 - 65; 56 - 75; 469 - 488; 479 - 498; 489 - 508; 519 - 538	804.9257 (+2)	1607.8368	1607.7044	0.1324	1	Y.GPGGSGPGGYGPGGSGPGGY.G	CID + ETD	38
1	O44359	Flageliform silk protein	Chymotrypsin	121-140	796.3983 (+2)	1590.7820	1590.6700	0.1121	1	Y.GPGGSGPGGYGPGGSGPGGY.G	CID	108

1	O44359	Flageliform silk protein	Chymotrypsin	317-336	743.3921 (+2)	1484.7696	1484.6645	0.1052	0	Y.GPGGSGPGGAGGAGGPGGA*Y.G	Ala->Val (A335) / CID	106
1	O44359	Flageliform silk protein	Chymotrypsin	317-336	751.8770 (+2)	1501.7394	1501.6182	0.1212	0	Y.GPGGSGPGGAGGAGGPGGAY.G	CID	51
1	O44359	Flageliform silk protein	Chymotrypsin	337-357	831.9373 (+2)	1661.8600	1661.7071	0.1530	1	Y.GPGGSYGPGGSGPGGAGGPY.G	CID	63
1	O44359	Flageliform silk protein	Chymotrypsin	343-371	770.3529 (+3)	2308.0369	2307.8869	0.1500	1	Y.GPGGSGGPGGAGGPYGPGGEGPGGAGGPY.G	CID + ETD	53
1	O44359	Flageliform silk protein	Chymotrypsin	358 - 371; 420 - 433	565.3246 (+2)	1128.6346	1128.4836	0.1510	0	Y.GPGGEGPGGAGGPY.G	CID + ETD	63
1	O44359	Flageliform silk protein	Chymotrypsin	358-380	930.4636 (+2)	1858.9126	1858.8314	0.0812	1	Y.GPGGEGPGGAGGPYGPGGAGGPY.G	CID	81
1	O44359	Flageliform silk protein	Chymotrypsin	372-389	779.3906 (+2)	1556.7666	1556.5933	0.1733	1	Y.GPGGAGGP*Y*GPGGAGGPY.G	Dihydroxylation (P379); Phospho (Y380) / CID	30
1	O44359	Flageliform silk protein	Chymotrypsin	381-398	752.4004 (+2)	1502.7862	1502.6427	0.1436	1	Y.GPGGAGGPYGPGGEGGPY.G	CID	82
1	O44359	Flageliform silk protein	Chymotrypsin	390-398	732.3884 (+1)	731.3811	731.3239	0.0572	0	Y.GPGGE*GGPY.G	Glu->Ala (E394) / CID	46
1	O44359	Flageliform silk protein	Chymotrypsin	399-419	886.9786 (+2)	1771.9426	1771.7802	0.1624	2	Y.GPGVS*Y*GPGGAGGPYGPGGPY.G	Tyr->Asp (Y404) / CID	69
1	O44359	Flageliform silk protein	Chymotrypsin	414-433	829.4463 (+2)	1656.8780	1656.7169	0.1612	1	Y.GPGGPYGPGGEGPGGAGGPY.G	CID + ETD	119
1	O44358	Flageliform silk protein	Chymotrypsin	509-528	804.9257 (+2)	1607.8368	1607.6601	0.1767	1	Y.GSGGAGPGGYGPGGSGPGGY.G	CID	43
1	O44358	Flageliform silk protein	Glu-C	26-30	544.3100 (+1)	543.3027	543.3268	-0.0241	0	E.LVIAE.S	CID	29
1	O44358	Flageliform silk protein	Glu-C	31-61	1078.7969 (+3)	3233.3689	3233.5582	-0.1893	1	E.SSGGDVQRKTNVISNALRNALMSTTGSPNEE.F	CID + ETD	39
1	O44358	Flageliform silk protein	Glu-C	66-80	894.5055 (+2)	1786.9964	1786.8771	0.1194	1	E.VQDLIQMLSSEQINE.V	CID	87
1	O44358	Flageliform silk protein	Glu-C	81-109	900.7024 (+3)	2699.0854	2699.2060	-0.1206	0	E.VDTSGPGQYYRSSSSGGGGGGGG*PVITE.T	Gly->Asn (G104) / CID + ETD	40
1	O44359	Flageliform silk protein	Proteinase 10	25-44; 478-497; 518-537;	796.300 (+2)	1.590.585	1.590.670	-0.0845	0	G.YGPGGSGPGGYGPGGSGPGG.Y	CID + ETD	48
1	O44359	Flageliform silk protein	Proteinase 10	109-129	786.810 (+2)	1.571.605	1.571.660	-0.0547	0	G.GSGPGGSGPGGYGPGGSGPGG.F	CID + ETD	70
1	O44359	Flageliform silk protein	Proteinase 10	352-374	921.920 (+2)	1.841.825	1.841.796	0.0285	0	G.GAGGPYGPGGEGPGGAGGPYGP.G	CID	62
1	O44359	Flageliform silk protein	Proteinase 10	408-427	829.370 (+2)	1.656.725	1.656.716	0.0086	0	G.GAGGPYGPGGPYGPGGEGPG.G	CID + ETD	64
1	O44359	Flageliform silk protein	Proteinase 10	410-427	765.320 (+2)	1.528.625	1.528.658	-0.0329	0	A.GGPYGPGGPYGPGGEGPG.G	CID + ETD	68
1	O44359	Flageliform silk protein	Subtilisin	6-15; 26-35; 46-55; 66-75; 81-90; 111-120; 131-140; 449-458; 469-478; 489-498, 529-538; 549-558; 569-578	805.3023 (+1)	804.2950	804.3403	-0.0452	0	V.GPGGSGPGGY.G	CID	39
1	O44359	Flageliform silk protein	Subtilisin	234-242	918.3370 (+1)	917.3297	917.4706	-0.1409	0	S.GGTTHIEDL.D	CID	35
1	O44359	Flageliform silk protein	Subtilisin	281-290; 291-300; 301-310; 704-713	741.3011 (+1)	740.2938	740.3454	-0.0515	0	G.GSGPGGVPGG.G	CID	51
2	O44358	Flageliform silk protein	Trypsin	39-48	372.5829 (+3)	1114.7269	1114.6458	0.0810	1	R.KTNVISNALR.N	CID + ETD	32
2	O44358	Flageliform silk protein	Trypsin	40-48	494.2347 (+2)	986.4548	986.5509	-0.0960	0	K.TNVISNALR.N	CID + ETD	35
2	O44359	Flageliform silk protein	Chymotrypsin	1 - 15; 434 - 448	586.7690 (+2)	1171.5222	1171.5211	-0.0011	0	GPGGVGPGGSGPGGY.G	CID	58
2	O44359	Flageliform silk protein	Chymotrypsin	1-25	648.2941 (+3)	1941.8646	1941.8623	0.0023	1	GPGGVGPGGSGPGGYGPGGAGPGGY.G	CID + ETD	36
2	O44359	Flageliform silk protein	Chymotrypsin	1-45	1172.5101 (+3)	3514.5294	3514.5247	0.0047	3	GPGGVGPGGSGPGGYGPGGAGPGGYGPGGSGPGGYGPGGSGPGGY.G	CID	33
2	O44359	Flageliform silk protein	Chymotrypsin	16-35; 459-478	788.3410 (+2)	1574.6808	1574.6754	0.0054	1	Y.GPGGAGPGGYGPGGSGPGGY.G	CID	36
2	O44359	Flageliform silk protein	Chymotrypsin	26 - 35; 36 - 45; 46 - 55; 56 - 65; 66 - 75; 81 - 90; 131 - 140; 449 - 458; 469 - 478; 479 - 488; 489 - 498; 499 - 508; 519 - 528; 529 - 538; 599 - 608	403.1701 (+2)	804.3419	804.3414	0.0005	0	Y.GPGGSGPGGY.G	CID	40

2	O44359	Flageliform silk protein	Chymotrypsin	26 - 45; 36 - 55; 46 - 65; 56 - 75; 469 - 488; 479 - 498; 489 - 508; 519 - 538	796.3401 (+2)	1590.6707	1590.6706	0.0001	1	Y.GPGGSGPGGYGPGGSGPGGY.G	CID	53
2	O44359	Flageliform silk protein	Chymotrypsin	26 - 55; 36 - 65; 46 - 75; 469 - 498; 479 - 508	1189.5001 (+2)	2377.0024	2376.9917	0.0107	2	Y.GPGGSGPGGYGPGGSGPGGYGPGGSGPGGY.G	CID + ETD	46
2	O44359	Flageliform silk protein	Chymotrypsin	26 - 65; 36 - 75; 469 - 508	1055.4512 (+3)	3163.3300	3163.3280	0.0020	3	Y.GPGGSGPGGYGPGGSGPGGYGPGGSGPGGYGPGGSGPGGY.G	CID + ETD	31
2	O44359	Flageliform silk protein	Chymotrypsin	76-90	618.7610 (+2)	1235.5162	1235.5229	-0.0067	1	Y.GPGGYGPGGSGPGGY.G	CID	31
2	O44359	Flageliform silk protein	Chymotrypsin	106-120	580.7501 (+2)	1159.4890	1159.4880	-0.0010	0	Y.GPGGSGPGGSGPGGY.G	CID	34
2	O44359	Flageliform silk protein	Chymotrypsin	106-130	644.2820 (+3)	1929.8153	1929.8207	-0.0054	1	Y.GPGGSGPGGSGPGGYGPGGSGPGGF.G	CID	38
2	O44359	Flageliform silk protein	Chymotrypsin	260-316	1059.7501 (+4)	4234.9727	4234.9614	0.0113	0	L.TISGAGGSGPGGAGPGGVGPGGSGPGGVGPGGSGPGGVGPGGSGPGGVGPGGAGGPY.G	ETD	40
2	O44359	Flageliform silk protein	Chymotrypsin	317-336	729.3212 (+2)	1456.6380	1456.6331	0.0049	0	Y.GPGGSGPGGAGGAGGPGGAY.G	CID	32
2	O44359	Flageliform silk protein	Chymotrypsin	337-357	831.8520 (+2)	1661.7033	1661.7059	-0.0026	1	Y.GPGGSYGPGGSGGPGGAGGPY.G	CID + ETD	64
2	O44359	Flageliform silk protein	Chymotrypsin	343-357	572.7510 (+2)	1143.4902	1143.4906	-0.0004	0	Y.GPGGSGGPGGAGGPY.G	CID	36
2	O44359	Flageliform silk protein	Chymotrypsin	343-371	752.3200 (+3)	2253.9517	2253.9644	-0.0127	1	Y.GPGGSGGPGGAGGPYGPGEPPGGAGGPY.G	CID + ETD	30
2	O44359	Flageliform silk protein	Chymotrypsin	343-380	990.0901 (+3)	2967.2791	2967.2830	-0.0039	2	Y.GPGGSGGPGGAGGPYGPGEPPGGAGGPYGPGGAGGPY.G	CID + ETD	42
2	O44359	Flageliform silk protein	Chymotrypsin	358 - 371; 420 - 433	565.2411 (+2)	1128.4804	1128.4808	-0.0004	0	Y.GPGGEGPGGAGGPY.G	CID	68
2	O44359	Flageliform silk protein	Chymotrypsin	358-380	921.9002 (+2)	1841.8024	1841.7982	0.0042	1	Y.GPGGEGPGGAGGPYGPGGAGGPY.G	CID	48
2	O44359	Flageliform silk protein	Chymotrypsin	372 - 380; 381 - 389; 405 - 413; 791 - 799	366.6620 (+2)	731.3209	731.3201	0.0008	0	Y.GPGGAGGPY.G	CID	44
2	O44359	Flageliform silk protein	Chymotrypsin	372-389	723.3211 (+2)	1444.6301	1444.6310	-0.0009	1	Y.GPGGAGGPYGPGGAGGPY*.G	Phospho (Y389) / CID	66
2	O44359	Flageliform silk protein	Chymotrypsin	372-398	1108.9822 (+2)	2215.9609	2215.9541	0.0068	2	Y.GPGGAGGPYGPGGAGGPYGPGEPPY.G	CID	36
2	O44359	Flageliform silk protein	Chymotrypsin	381-398	752.3210 (+2)	1502.6351	1502.6408	-0.0057	1	Y.GPGGAGGPYGPGEPPY.G	CID + ETD	50
2	O44359	Flageliform silk protein	Chymotrypsin	405-419	630.7800 (+2)	1259.5503	1259.5509	-0.0006	1	Y.GPGGAGGPYGPGGPY.G	CID	50
2	O44359	Flageliform silk protein	Chymotrypsin	405-433	1186.0210 (+2)	2370.0330	2370.0301	0.0029	2	Y.GPGGAGGPYGPGGPYGPGEPPGGAGGPY.G	CID	31
2	O44359	Flageliform silk protein	Chymotrypsin	414-433	829.3613 (+2)	1656.7101	1656.7107	-0.0006	1	Y.GPGGPYGPGEPPGGAGGPY.G	CID + ETD	52
2	O44359	Flageliform silk protein	Chymotrypsin	420-448	1142.0112 (+2)	2282.0198	2281.9923	0.0175	1	Y.GPGGEGPGGAGGPYGPGGVPGGSGPGGY.G	CID	42
2	O44359	Flageliform silk protein	Chymotrypsin	434-458	979.9421 (+2)	1957.8707	1957.8552	0.0155	1	Y.GPGGVPGGSGPGGYGPGGSGPGGY.G	CID + ETD	40
2	O44359	Flageliform silk protein	Chymotrypsin	609-665	1042.7031 (+4)	4166.8009	4166.7361	0.0648	0	Y.GPGGSGAGGTGPGGAGGAGGAGSGGAGGSGGAGGSGGAGGSGGVGGSGGTTITEDL.D	ETD	38
2	O44359	Flageliform silk protein	Chymotrypsin	746-760	621.7623 (+2)	1241.5111	1241.4700	0.0411	0	Y.GPGGSGSGGVGPGGY.G	CID	33
2	O44359	Flageliform silk protein	Chymotrypsin	746-768	889.8900 (+2)	1777.7646	1777.7624	0.0022	1	Y.GPGGSGSGGVGPGGYGPGGSGGF.Y	CID + ETD	40
2	O44359	Flageliform silk protein	Chymotrypsin	746-769	971.4201 (+2)	1940.8260	1940.8293	-0.0033	2	Y.GPGGSGSGGVGPGGYGPGGSGGFY.G	CID	35
2	O44359	Flageliform silk protein	Chymotrypsin	770-784	691.7921 (+2)	1381.5714	1381.5756	-0.0012	1	Y.GPGGSEGPYGPSGTY.G	CID + ETD	32
2	O44359	Flageliform silk protein	Chymotrypsin	809-844	1024.1731 (+4)	4092.6605	4092.5010	0.1595	3	Y.GPGSPGGAYYPSSRVPDMVNGIMSAMQGS GFNYQM*.F.G	Oxidation (M843) / ETD	38
2	O44359	Flageliform silk protein	Glu-C	26-30	544.3056 (+1)	543.2983	543.3268	-0.0285	0	E.LVIAE.S	CID	19
2	O44359	Flageliform silk protein	Glu-C	31-60	1078.4963 (+3)	3232.4671	3232.5742	-0.1071	0	E.SSGGDVQRKTNVISNALRNALMSTTGSPNE.E	CID + ETD	39
2	O44359	Flageliform silk protein	Glu-C	31-65	941.4182 (+4)	3761.6437	3761.8045	-0.1608	2	E.SSGGDVQRKTNVISNALRNALM*.STTGSPNEEFVHE.V	Oxidation (M52) / ETD	31
2	O44359	Flageliform silk protein	Glu-C	66-80	894.4152 (+2)	1786.8158	1786.8771	-0.0612	1	E.VQDLIQMLSQEIQINE.V	CID	87
2	O44359	Flageliform silk protein	Glu-C	81-109	900.6979 (+3)	2699.0719	2699.2060	-0.1341	0	E.VDTSGPGQYYRSSSSGGGGGGQGGPVVTE.T	CID + ETD	87

2	O44359	Flageliform silk protein	Proteinase 10	4-14	798.340 (+1)	797.332	797.366	-0.0341	0	G.GVPGGSGPGG.Y	CID	31
2	O44359	Flageliform silk protein	Proteinase 10	5-14	741.340 (+1)	740.332	740.345	-0.0126	0	G.VGPGGSGPGG.Y	CID	36
2	O44359	Flageliform silk protein	Proteinase 10	5-18	1.115.410 (+1)	1.114.402	1.114.504	-0.1017	0	G.VGPGGSGPGGYGPG.G	CID	39
2	O44359	Flageliform silk protein	Proteinase 10	5-19	586.760 (+2)	1.171.505	1.171.525	-0.0204	0	G.VGPGGSGPGGYGPGG.A	CID + ETD	35
2	O44359	Flageliform silk protein	Proteinase 10	5-24	756.330 (+2)	1.510.645	1.510.680	-0.0347	0	G.VGPGGSGPGGYGPGGAGPGG.Y	CID	70
2	O44359	Flageliform silk protein	Proteinase 10	5-28	943.410 (+2)	1.884.805	1.884.839	-0.0337	0	G.VGPGGSGPGGYGPGGAGPGGYGPG.G	CID + ETD	56
2	O44359	Flageliform silk protein	Proteinase 10	15-34; 458-477	788.320 (+2)	1.574.625	1.574.675	-0.0496	0	G.YGPGGAGPGGYGPGGSGPGG.Y	CID + ETD	34
2	O44359	Flageliform silk protein	Proteinase 10	20-38	759.790 (+2)	1.517.565	1.517.653	-0.0881	0	G.AGPGGYGPGGSGPGGYGPG.G	CID	45
2	O44359	Flageliform silk protein	Proteinase 10	20-54	906.360 (+3)	2.716.058	2.716.153	-0.0958	0	G.AGPGGYGPGGSGPGGYGPGGSGPGGYGPGGSGPGG.Y	CID + ETD	31
2	O44359	Flageliform silk protein	Proteinase 10	24-44; 477-497; 517-537	824.820 (+2)	1.647.625	1.647.691	-0.0660	0	G.GYGPGGSGPGGYGPGGSGPGG.Y	CID	71
2	O44359	Flageliform silk protein	Proteinase 10	24-48	1.011.880 (+2)	2.021.745	2.021.850	-0.1050	0	G.GYGPGGSGPGGYGPGGSGPGGYGPG.G	CID	97
2	O44359	Flageliform silk protein	Proteinase 10	24-54	812.320 (+3)	2.433.938	2.434.021	-0.0829	0	G.GYGPGGSGPGGYGPGGSGPGGYGPGGSGPGG.Y	ETD	29
2	O44359	Flageliform silk protein	Proteinase 10	24-58	937.080 (+3)	2.808.218	2.808.180	0.0380	0	G.GYGPGGSGPGGYGPGGSGPGGYGPGGSGPGGYGPG.G	CID + ETD	47
2	O44359	Flageliform silk protein	Proteinase 10	25-38	590.220 (+2)	1.178.425	1.178.499	-0.0738	0	G.YGPGGSGPGGYGPG.G	CID	39
2	O44359	Flageliform silk protein	Proteinase 10	25-43	767.810 (+2)	1.533.605	1.533.648	-0.0430	0	G.YGPGGSGPGGYGPGGSGPG.G	CID + ETD	84
2	O44359	Flageliform silk protein	Proteinase 10	25-48	983.380 (+2)	1.964.745	1.964.829	-0.0835	0	G.YGPGGSGPGGYGPGGSGPGGYGPG.G	CID + ETD	106
2	O44359	Flageliform silk protein	Proteinase 10	25-49	1.011.890 (+2)	2.021.765	2.021.850	-0.0850	0	G.YGPGGSGPGGYGPGGSGPGGYGPGG.S	CID + ETD	46
2	O44359	Flageliform silk protein	Proteinase 10	25-53	774.280 (+3)	2.319.818	2.319.978	-0.1600	0	G.YGPGGSGPGGYGPGGSGPGGYGPGGSGPG.G	ETD	66
2	O44359	Flageliform silk protein	Proteinase 10	25-54	793.310 (+3)	2.376.908	2.376.999	-0.0915	0	G.YGPGGSGPGGYGPGGSGPGGYGPGGSGPGG.Y	CID + ETD	55
2	O44359	Flageliform silk protein	Proteinase 10	25-63	1.036.380 (+3)	3.106.118	3.106.307	-0.1897	0	G.YGPGGSGPGGYGPGGSGPGGYGPGGSGPGGYGPGGSGPG.G	CID	74
2	O44359	Flageliform silk protein	Proteinase 10	29-44	609.230 (+2)	1.216.445	1.216.510	-0.0655	0	G.GSGPGGYGPGGSGPGG.Y	CID	35
2	O44359	Flageliform silk protein	Proteinase 10	29-53	973.870 (+2)	1.945.725	1.945.819	-0.0937	0	G.GSGPGGYGPGGSGPGGYGPGGSGPG.G	CID	109
2	O44359	Flageliform silk protein	Proteinase 10	29-58	793.340 (+3)	2.376.998	2.376.999	-0.0015	0	G.GSGPGGYGPGGSGPGGYGPGGSGPGGYGPG.G	CID + ETD	37
2	O44359	Flageliform silk protein	Proteinase 10	29-63	911.690 (+3)	2.732.048	2.732.148	-0.1007	0	G.GSGPGGYGPGGSGPGGYGPGGSGPGGYGPGGSGPG.G	CID + ETD	71
2	O44359	Flageliform silk protein	Proteinase 10	29-64	930.690 (+3)	2.789.048	2.789.170	-0.1221	0	G.GSGPGGYGPGGSGPGGYGPGGSGPGGYGPGGSGPGG.Y	ETD	52
2	O44359	Flageliform silk protein	Proteinase 10	30-64	911.690 (+3)	2.732.048	2.732.148	-0.1007	0	G.SGPGGYGPGGSGPGGYGPGGSGPGGYGPGGSGPGG.Y	CID + ETD	48
2	O44359	Flageliform silk protein	Proteinase 10	55-88	944.700 (+3)	2.831.078	2.831.125	-0.0468	0	G.YGPGGSGPGGYGPGGSGPGGYGPGGYGPGGSGPG.G	CID + ETD	50
2	O44359	Flageliform silk protein	Proteinase 10	92-114	910.890 (+2)	1.819.765	1.819.716	0.0491	0	G.PGG ^T *GPGGSGPGGYGPGGSGPGG.S	Phospho (T95) / CID	26
2	O44359	Flageliform silk protein	Proteinase 10	99-123	973.880 (+2)	1.945.745	1.945.819	-0.0737	0	G.GSGPGGYGPGGSGPGGSGPGGYGPG.G	CID	45
2	O44359	Flageliform silk protein	Proteinase 10	99-124	1.002.370 (+2)	2.002.725	2.002.840	-0.1152	0	G.GSGPGGYGPGGSGPGGSGPGGYGPGG.S	CID	65
2	O44359	Flageliform silk protein	Proteinase 10	99-129	786.980 (+3)	2.357.918	2.357.989	-0.0717	0	G.GSGPGGYGPGGSGPGGSGPGGYGPGGSGPGG.F	CID + ETD	39
2	O44359	Flageliform silk protein	Proteinase 10	100-119	758.290 (+2)	1.514.565	1.514.638	-0.0732	0	G.SGPGGYGPGGSGPGGSGPGG.Y	CID	55
2	O44359	Flageliform silk protein	Proteinase 10	104-123	796.300 (+2)	1.590.585	1.590.670	-0.0845	0	G.GYGPGGSGPGGSGPGGYGPG.G	CID	74
2	O44359	Flageliform silk protein	Proteinase 10	104-129	1.002.360 (+2)	2.002.705	2.002.840	-0.1352	0	G.GYGPGGSGPGGSGPGGYGPGGSGPGG.F	CID	102
2	O44359	Flageliform silk protein	Proteinase 10	105-129	973.870 (+2)	1.945.725	1.945.819	-0.0937	0	G.YGPGGSGPGGSGPGGYGPGGSGPGG.F	CID + ETD	82
2	O44359	Flageliform silk protein	Proteinase 10	109-129	786.840 (+2)	1.571.665	1.571.660	0.0053	0	G.GSGPGGSGPGGYGPGGSGPGG.F	CID	54

				549-558; 569-578								
2	O44359	Flageliform silk protein	Subtilisin	273-290; 283-300	660.2815 (+2)	1318.5484	1318.6266	-0.0782	0	A.GPGGVGPGGSGPGGVGP.G	CID + ETD	84
2	O44359	Flageliform silk protein	Subtilisin	276-290; 286-300	1108.3436 (+1)	1107.3363	1107.5309	-0.1946	0	G.GVGPGGSGPGGVGP.G	CID	54
2	O44359	Flageliform silk protein	Subtilisin	281-290; 291-300; 301-310; 704-713	741.3361 (+1)	740.3288	740.3454	-0.0165	0	G.GSGPGGVGP.G	CID	48
2	O44359	Flageliform silk protein	Subtilisin	281-300	732.2972 (+2)	1462.5798	1462.6801	-0.1003	0	G.GSGPGGVGPGGSGPGGVGP.G	CID + ETD	90
2	O44359	Flageliform silk protein	Subtilisin	390-398	790.2277 (+1)	789.2204	789.3293	-0.1089	0	Y.GPGGEGGPY.G	CID	42
3	B5SYS6	Spidroin-1	Trypsin	49-68	1061.5312 (+2)	2121.0478	2120.9337	0.1141	0	R.T*GAFTAD*QLDDMSTIGDTL.K	Phospho (T49); Asp->Gly (D55) / CID	40
3	B5SYS6	Spidroin-1	Trypsin	49-68	711.3253 (+3)	2130.9541	2130.9627	-0.0086	0	R.TGAFTADQLDDM*STIGDTL.K	Sulphone (M60)	48
3	B5SYS6	Spidroin-1	Trypsin	80-112	1129.2910 (+3)	3384.8512	3384.6541	0.1971	1	K.SSQSKLQALNM*AFASSM*AEIAAVEQ*GGLSVAEK.T	Oxidation (M90, M96) / CID + ETD	26
3	B5SYS6	Spidroin-1	Trypsin	85-112	964.4119 (+3)	2890.2139	2890.3493	-0.1354	0	K.LQALNM*AFASS*M*AEIAAVEQ*GGLSVAEK.T	Oxidation (M90, M96); Phospho (S95); Gln->Ala (Q104) / CID	36
3	B5SYS6	Spidroin-1	Trypsin	113-140	1015.1800 (+3)	3042.5182	3042.5047	0.0135	0	K.TNAIADSLNSAFYQTTGAVNVQFVNEIR.S	CID + ETD	79
3	P19837	Spidroin-1	Trypsin	70-99	811.0642 (+3)	2430.1708	2430.1749	-0.0041	0	R.GGQGA*G*AAAAAGGAGQGGYGGLGSQGAGR.G	Gly->Lys (G75) / CID + ETD	79
3	P19837	Spidroin-1	Trypsin	134-161	1138.0829 (+2)	2274.1512	2274.0486	0.1026	0	R.GGQGA*AAAAAGGAGQGGYGGLGSQGAGR.G	CID	34
3	P19837	Spidroin-1	Trypsin	266-308	861.3936 (+3)	3441.5453	3441.5196	0.0257	0	R.GGEGAGAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	CID + ETD	35
3	P19837	Spidroin-1	Trypsin	309-363	1049.4651 (+4)	4193.8313	4193.9684	-0.1371	0	R.GGLGGQGAGA*AAAGGAGQGGGAGQGAGAAAAAGGAGQGGYGGLGSQGAGR.G	Ala->Gly (A318) / ETD	81
3	P19837	Spidroin-1	Trypsin	364-397	907.4190 (+3)	2719.2352	2719.2481	-0.0130	0	R.GGLGGQGAGAVAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	32
3	P19837	Spidroin-1	Trypsin	398-415	735.3429 (+2)	1468.6712	1468.7495	-0.0782	0	R.GGQGA*G*AAAAAGGAGQR.G	Gly->Lys (G403) / CID	35
3	P19837	Spidroin-1	Trypsin	398-427	1036.4752 (+3)	3106.4038	3106.2193	0.1844	1	R.GGQGAGAAAAAGGAGQRGYGGLGNQGAGR.G	CID	43
3	P19837	Spidroin-1	Trypsin	462-488	1095.0335 (+2)	2188.0524	2188.0482	0.0042	0	R.GGQ*GAAAAAGGAGQGGYGGLGSQGAGR.G	Gln->Arg (Q464) / CID + ETD	159
3	P19837	Spidroin-1	Trypsin	572-585	777.3693 (+2)	1552.7240	1552.5861	0.1379	0	R.QG*Y*GGLGS*QGAGR.G	Gly->Trp (G574); Phospho (Y575, S580) / CID	25
3	P19837	Spidroin-1	Trypsin	616-652	954.8108 (+3)	2861.4106	2861.3401	0.0705	0	R.GGLGGQGAGAAAGGAGQGGYGGVGSGASAASAAASR.L	CID + ETD	100
3	B5SYS6	Spidroin-1	Chymotrypsin	30-42	693.3436 (+2)	1384.6726	1384.6510	0.0216	2	W.SSTELADAFINAF.L	CID	54
3	B5SYS6	Spidroin-1	Chymotrypsin	43-52	518.2402 (+2)	1034.4658	1034.5145	-0.0486	1	F.LNEAGRTGAF.T	CID	61
3	B5SYS6	Spidroin-1	Chymotrypsin	43-67	871.3867 (+3)	2611.1383	2611.2072	-0.0689	3	F.LNEAGRTGAFTADQLDDMSTIGDTL.K	CID + ETD	123
3	B5SYS6	Spidroin-1	Chymotrypsin	44-67	876.7756 (+3)	2627.3050	2627.1098	0.1951	2	L.NEAGRT*GAFTADQLDDMSTIGDTL.K	Phospho (T49) / CID + ETD	64
3	B5SYS6	Spidroin-1	Chymotrypsin	53-67	798.3548 (+2)	1594.6950	1594.7032	-0.0082	1	F.TADQLDDMSTIGDTL.K	CID + ETD	79
3	B5SYS6	Spidroin-1	Chymotrypsin	108-125	950.9991 (+2)	1899.9836	1899.9214	0.0623	2	L.SVAEKTNAIADSLNSAFY.Q	CID	83
3	B5SYS6	Spidroin-1	Chymotrypsin	136-146	662.8503 (+2)	1323.6860	1323.6857	0.0004	1	F.VNEIRSLISM*F.A	Oxidation (M145) / CID	59
3	B5SYS6	Spidroin-1	Chymotrypsin	147-156	520.1704 (+2)	1038.3262	1038.4618	-0.1355	0	F.AQASANEVSY.G	CID + ETD	68
3	B5SYS6	Spidroin-1	Chymotrypsin	157-185	787.3754 (+3)	2359.1044	2359.0214	0.0830	1	Y.GGGYGGGQGGQS*A*GAAAAASAGAGQGGY.G	Ala->Phe (A169) / CID + ETD	36
3	B5SYS6	Spidroin-1	Chymotrypsin	186-209	917.4282 (+2)	1832.8418	1832.8402	0.0017	1	Y.GGLGGQGAGSAAAAAAS*GAGQGGY.G	Ser->Gly (S202) / CID	121
3	B5SYS6	Spidroin-1	Chymotrypsin	210-236	787.3356 (+3)	2358.9850	2359.1580	-0.1731	0	Y.GGVGNQGAGRGAGAAAAAGGAGQGGY.N	CID + ETD	64
3	P19837	Spidroin-1	Chymotrypsin	1-31	873.6586 (+3)	2617.9540	2618.0970	-0.1430	2	QGAGAAAAAGGAGQGGYGGLGGQGAGQGGY.G	CID + ETD	38
3	P19837	Spidroin-1	Chymotrypsin	19-31; 185-197; 286-298;	539.7255 (+2)	1077.4364	1077.4840	-0.0475	1	Y.GGLGGQGAGQGGY.G	CID	55
3	P19837	Spidroin-1	Chymotrypsin	19-34; 185-200; 286-301	653.2820 (+2)	1304.5494	1304.6110	-0.0615	2	Y.GGLGGQGAGQGGYGGL.G	CID	64
3	P19837	Spidroin-1	Chymotrypsin	22-34	539.6789 (+2)	1077.3432	1077.4840	-0.1407	1	L.GGQGAGQGGYGGL.G	CID + ETD	44
3	P19837	Spidroin-1	Chymotrypsin	30-59	825.3603 (+3)	2473.0591	2473.0408	0.0183	2	G*Y*GGLGGQGA*GQGAGAAAAAGGAGQGGY.G	Phospho (Y31); Ala->Asn (A39) / CID	60
3	P19837	Spidroin-1	Chymotrypsin	32-59	720.3233 (+3)	2157.9481	2158.0264	-0.0783	1	Y.GGLGGQGAGQ*GAGAAAAAGGAGQGGY.G	Gln->Arg (Q41) / CID + ETD	95
3	P19837	Spidroin-1	Chymotrypsin	35-59	1016.9622 (+2)	2031.9098	2031.9147	-0.0049	0	L.GGQGAGQGAGAAAAAGGAGQGGY.G	CID	96
3	P19837	Spidroin-1	Chymotrypsin	60-92; 576-608	886.7379 (+3)	2657.1919	2657.1620	0.0299	2	Y.GGLGS*QGAGR*GGQGAGAAAAAGGAGQGGYGGL.G	Phospho (S64, S580); Arg->Phe (R69, 585) / CID	103
3	P19837	Spidroin-1	Chymotrypsin	63-89; 579-605	735.3099 (+3)	2202.9079	2202.9080	-0.0001	0	L.GS*QGAGR*GGQGAGAAAAAGGAGQGGY.G	Phospho (S64, 580); Arg->Phe (R69, 585) / CID	46

3	P19837	Spidroin-1	Chymotrypsin	63-92; 579-608	863.1109 (+3)	2586.3109	2586.3263	-0.0154	1	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGL.G	CID + ETD	70
3	P19837	Spidroin-1	Chymotrypsin	90-102; 152-164; 198-210; 299-311; 354-366	568.7285 (+2)	1135.4424	1135.5370	-0.0946	1	Y.GGLGSQGAGRGG.L	CID	70
3	P19837	Spidroin-1	Chymotrypsin	93-123	816.3677 (+3)	2446.0813	2446.1334	-0.0521	1	L.GSQGAGRGGLGQQGAGAA*AAAAGGAGQGGY.G	Ala->Ser (A111) / CID + ETD	76
3	P19837	Spidroin-1	Chymotrypsin	127-151	696.6301 (+3)	2086.8685	2086.9893	-0.1208	0	L.GNQGAGRGGQAAAAAAGGAGQGGY.G	CID + ETD	80
3	P19837	Spidroin-1	Chymotrypsin	152-184	853.0706 (+3)	2556.1900	2556.2178	-0.0278	2	Y.GGLGSQGAGRGGLGQQGAGAAAAAAGGAGQGGY.G	CID + ETD	119
3	P19837	Spidroin-1	Chymotrypsin	155-184	811.0091 (+3)	2430.0055	2430.0350	-0.0295	1	L.GS*QGAGR*GGLGGQGAGAAAAAAGGAGQGGY.G	Phospho (S156); Arg->Phe (R161) / CID	61
3	P19837	Spidroin-1	Chymotrypsin	155-197	1164.1392 (+3)	3489.3958	3489.5084	-0.1126	3	L.GS*QGAGR*GGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Phospho (S156); Arg->Phe (R161) / CID	50
3	P19837	Spidroin-1	Chymotrypsin	165-184	788.8373 (+2)	1575.6600	1575.7026	-0.0426	0	L.GGQGAGAAA*AAGGAGQGGY.G	Ala->Gln (A174) / CID	46
3	P19837	Spidroin-1	Chymotrypsin	201-231	1216.0100 (+2)	2430.0054	2430.1385	-0.1330	1	L.GSQGAGRGGLGQQGAGAAAAAAGGAGQGG.L	CID	97
3	P19837	Spidroin-1	Chymotrypsin	232-255	916.8785 (+2)	1831.7424	1831.8198	-0.0773	0	L.GGQGAGQGAGASAAAAGGAGQGGY.G	CID + ETD	163
3	P19837	Spidroin-1	Chymotrypsin	256-285	787.3356 (+3)	2358.9850	2359.1378	-0.1528	1	Y.GGLGSQGAGRGGE*GAGAAAAAAGGAGQGGY.G	Glu->Lys (E268) / CID + ETD	59
3	P19837	Spidroin-1	Chymotrypsin	259-285	1066.9708 (+2)	2131.9270	2132.0108	-0.0837	0	L.GSQGAGRGGE*GAGAAAAAAGGAGQGGY.G	Glu->Lys (E268) / CID	149
3	P19837	Spidroin-1	Chymotrypsin	259-298	1064.7983 (+3)	3191.3731	3191.4842	-0.1111	2	L.GSQGAGRGGE*GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Glu->Lys (E268) / CID	138
3	P19837	Spidroin-1	Chymotrypsin	302-329	746.9998 (+3)	2237.9776	2238.1214	-0.1438	1	L.GSQGAGRGGLGQQGAG*AAAAGGAGQGG.L	Gly->Lys (G317) / CID + ETD	58
3	P19837	Spidroin-1	Chymotrypsin	330-353	917.4282 (+2)	1832.8418	1832.8168	0.0250	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID	79
3	P19837	Spidroin-1	Chymotrypsin	421-451	811.0216 (+3)	2430.0430	2430.1385	-0.0955	1	L.GNQGAGRGGLGQ*GAGAAAAAAGGAGQGGY.G	Gln->Thr (Q433) / CID + ETD	77
3	P19837	Spidroin-1	Chymotrypsin	452-478	735.3099 (+3)	2202.9079	2203.0115	-0.1036	1	Y.GGLGNQGAGRGGQGA*AAAAGGAGQGGY.G	Ala->Ser (A466) / CID + ETD	49
3	P19837	Spidroin-1	Chymotrypsin	455-478	1049.6106 (+2)	2097.2066	2097.0536	0.1530	0	L.GNQGAGRGGQAAAAAAGGAGQGGY.G	CID	26
3	P19837	Spidroin-1	Chymotrypsin	519-531	543.7438 (+2)	1085.4730	1085.5578	-0.0848	1	Y.GGLGSQSGRGG.L	CID + ETD	69
3	P19837	Spidroin-1	Chymotrypsin	522-551	794.4000 (+3)	2380.1782	2380.0340	0.1442	1	L.GSQSGRGGLGQ*GAGAAAAAAGGAGQGG.L	Gln->Cys (Q534) CID	56
3	P19837	Spidroin-1	Chymotrypsin	606-636	825.0649 (+3)	2472.1729	2472.1855	-0.0126	2	Y.GGLG*QGVGRGGLGQQGAGAAAAGGAGQGGY.G	Gly->Ser (G610) / CID + ETD	100
3	P19837	Spidroin-1	Chymotrypsin	637-653	695.3213 (+2)	1388.6280	1388.7008	-0.0727	0	Y.GGVGSGASAASAAASRL.S	CID	110
3	P19837	Spidroin-1	Chymotrypsin	654-669	780.9046 (+2)	1559.7946	1559.7903	0.0043	0	L.SSPQASSRVSS*AVSNL.V	Ser->Ala (S664) / CID	50
3	P19837	Spidroin-1	Chymotrypsin	654-680	843.8081 (+3)	2528.4025	2528.2831	0.1194	1	L.SSPQASSRVSS*AVSNLVASGPTNSAAL.S	Ser->Ala (S664) / CID + ETD	112
3	P19837	Spidroin-1	Chymotrypsin	730-744	718.3326 (+2)	1434.6506	1434.7103	-0.0597	0	Y.GSAGQAT*QIVGQSVY.Q	Thr->Ala (T736) / CID	68
3	B5SYS6	Spidroin-1	Glu-C	34-45	686.3299 (+2)	1370.6452	1370.6176	0.0276	0	E.LADAFINAF.LNE.A	CID + ETD	69
3	B5SYS6	Spidroin-1	Glu-C	99-103	502.2771 (+1)	501.2698	501.2798	-0.0100	0	E.IAAVE.Q	CID	28
3	B5SYS6	Spidroin-1	Glu-C	104-111	760.3629 (+1)	759.3556	759.3763	-0.0207	0	E.QGGLSVAE.K	CID	25
3	B5SYS6	Spidroin-1	Glu-C	112-138	968.1615 (+3)	2901.4627	2901.4145	0.0482	0	E.KTNAIADSLNSAFYQTTGAVNVQFVNE.I	CID + ETD	105
3	B5SYS6	Spidroin-1	Glu-C	139-153	842.4478 (+2)	1682.8810	1682.8062	0.0748	0	E.IRSLIS*MFAQASANE.V	Phospho (S144) / CID	31
3	B5SYS6	Spidroin-1	AspN	45-54	490.7567 (+2)	979.4988	979.4723	0.0265	0	N.EAGRTGAFTA.D	CID + ETD	27
3	B5SYS6	Spidroin-1	AspN	55-64	1110.2678 (+1)	1109.2605	1109.4547	-0.1942	2	A.DQLDDMSTIG.D	CID	46
3	B5SYS6	Spidroin-1	AspN	58-64	738.2764 (+1)	737.2691	737.2902	-0.0211	1	L.DDMSTIG.D	CID	27
3	B5SYS6	Spidroin-1	AspN	65-71	779.4400 (+1)	778.4327	778.3895	0.0432	0	G.DTLKTAM.D	CID	25
3	B5SYS6	Spidroin-1	AspN	98-110	622.3300 (+2)	1242.6454	1242.6456	-0.0001	1	A.EIAAVEQGGLSVA.E	CID + ETD	72
3	B5SYS6	Spidroin-1	AspN	98-117	986.0500 (+2)	1970.0854	1970.0320	0.0534	2	A.EIAAVEQGGLSVAEKTNAIA.D	CID	46
3	B5SYS6	Spidroin-1	Proteinase 10	52-66	823.3500 (+2)	1644.6854	1644.6825	0.0029	0	A.FTADQLDDMSTIGDT.L	CID	58
3	B5SYS6	Spidroin-1	Proteinase 10	199-211	993.4500 (+1)	992.4427	992.4312	0.0115	0	A.AAAS*GAGQGGYGG.V	Ser->Gly (S202) / CID	54
4	B5SYS6	Spidroin-1	Trypsin	49-68	1050.4813 (+2)	2098.9480	2098.9729	-0.0248	0	R.TGAFTADQLDDMSTIGDTLK.T	CID	127
4	B5SYS6	Spidroin-1	Trypsin	49-73	912.7489 (+3)	2735.2249	2735.1942	0.0306	1	R.TGAFTADQLDDMSTIGDTLK*TAM*DK.M	Oxidation (M71); Lys->Trp (K68) / CID + ETD	43
4	B5SYS6	Spidroin-1	Trypsin	85-112	982.5089 (+3)	2944.5049	2944.3381	0.1668	0	K.LQALNMAFAS*S*AEIAAVEQGGLSVAEK.T	Phospho (S95); Oxidation (M96) / CID	72
4	B5SYS6	Spidroin-1	Trypsin	113-140	1015.1816 (+3)	3042.5230	3042.5047	0.0183	0	K.TNAIADSLNSAFYQTTGAVNVQFVNEIR.S	CID + ETD	125

4	P19837	Spidroin-1	Trypsin	70-99	855.9769 (+3)	2564.9089	2564.9643	-0.0555	0	R.GGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	30
4	P19837	Spidroin-1	Trypsin	134-161	927.0870 (+3)	2778.2392	2778.2589	-0.0197	0	R.GGQGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	76
4	P19837	Spidroin-1	Trypsin	162-207	1328.0965 (+3)	3981.1297	3980.9355	0.1942	0	R.GGLGGQ*GAGAAAAAAGGAGQ*GGY*GGLGGQ*GAGQGGYGGLGSQGAGR.G	Phospho (Y184) / CID	38
4	P19837	Spidroin-1	Trypsin	266-308	878.9195 (+4)	3511.6489	3511.5654	0.0835	0	R.GGEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	ETD	32
4	P19837	Spidroin-1	Trypsin	309-363	1049.7062 (+4)	4194.7957	4194.9525	-0.1568	0	R.GGLGGQGAGAAAAGGAGQ*GGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	Gln->Asp (Q326) / ETD	63
4	P19837	Spidroin-1	Trypsin	364-397	950.8087 (+3)	2849.4043	2849.3570	0.0473	0	R.GGLGGQGAGAVAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	51
4	P19837	Spidroin-1	Trypsin	398-415	721.8601 (+2)	1441.7056	1441.6658	0.0398	0	R.GGQGAGAAAAAAGGAGQR.G	Ala->Asp (A412) / CID	96
4	P19837	Spidroin-1	Trypsin	462-488	1095.0163 (+2)	2188.0180	2188.0370	-0.0189	0	R.GGQGAA*AAAAGGAGQGGYGGLGSQGAGR.G	CID	195
4	P19837	Spidroin-1	Trypsin	489-509	873.4306 (+2)	1744.8466	1744.8912	-0.0445	0	R.GGQGAGAAAAAAGVAGQEGIR.G	CID	30
4	P19837	Spidroin-1	Trypsin	510-528	858.8629 (+2)	1715.7112	1715.7013	0.0099	0	R.GQGAGQGGYGGLGS*QSGSR.G	Phospho (S523) / CID	20
4	P19837	Spidroin-1	Trypsin	572-585	824.8953 (+2)	1647.7760	1647.7610	0.0151	0	R.QGGY*GGLGSQGAGR.G	Phospho (Y575) / CID	57
4	P19837	Spidroin-1	Trypsin	586-615	786.6874 (+3)	2357.0404	2357.1221	-0.0817	0	R.GGQGAGAAAAAAGGAGQGGYGGLGGQGVGR.G	CID + ETD	32
4	P19837	Spidroin-1	Trypsin	616-652	954.7892 (+3)	2861.3458	2861.3401	0.0057	0	R.GGLGGQGAGAAAAGGAGQGGYGGVGSASAASAAASR.L	CID + ETD	114
4	B5SYS6	Spidroin-1	Chymotrypsin	30-38	470.7156 (+2)	939.4166	939.4185	-0.0019	1	W.SSTELADAF.I	CID	40
4	B5SYS6	Spidroin-1	Chymotrypsin	30-42	693.3751 (+2)	1384.7356	1384.6510	0.0846	2	W.SSTELADAFINAF.L	CID	65
4	B5SYS6	Spidroin-1	Chymotrypsin	39-52	740.9023 (+2)	1479.7900	1479.7470	0.0430	2	F.INAFLNEAGRTGAF.T	CID + ETD	82
4	B5SYS6	Spidroin-1	Chymotrypsin	43-52	518.2603 (+2)	1034.5060	1034.5145	-0.0084	1	F.LNEAGRTGAF.T	CID	55
4	B5SYS6	Spidroin-1	Chymotrypsin	43-67	914.1401 (+3)	2739.3985	2739.2099	0.1886	3	F.LNEAGRTGAFTADQLDDMSTIGD*TL.K	Asp->Tyr (D65) / CID + ETD	47
4	B5SYS6	Spidroin-1	Chymotrypsin	53-67	798.3585 (+2)	1594.7024	1594.7032	-0.0008	1	F.TADQLDDMSTIGDTL.K	CID	82
4	B5SYS6	Spidroin-1	Chymotrypsin	93-107	717.3886 (+2)	1432.7626	1432.6868	0.0759	0	F.ASSMAEIAAVEQGGL.S	CID	80
4	B5SYS6	Spidroin-1	Chymotrypsin	108-124	868.9095 (+2)	1735.8044	1735.8740	-0.0696	1	L.SVAEKTNAIADSLNSAF.Y	CID + ETD	37
4	B5SYS6	Spidroin-1	Chymotrypsin	108-125	950.9848 (+2)	1899.9550	1899.9214	0.0337	2	L.SVAEKTNAIADSLNSAFY.Q	CID	86
4	B5SYS6	Spidroin-1	Chymotrypsin	126-135	532.6987 (+2)	1063.3828	1063.5298	-0.1470	0	Y.QTTGAVNVQF.V	CID + ETD	49
4	B5SYS6	Spidroin-1	Chymotrypsin	126-142	938.5004 (+2)	1874.9862	1874.9850	0.0012	1	Y.QTTGAVNVQFVNEIRSL.I	CID	60
4	B5SYS6	Spidroin-1	Chymotrypsin	136-146	662.8455 (+2)	1323.6764	1323.6857	-0.0092	1	F.VNEIRSLISM*F.A	Oxidation (M145) / CID	53
4	B5SYS6	Spidroin-1	Chymotrypsin	147-160	687.2927 (+2)	1372.5708	1372.5895	-0.0187	1	F.AQASANEVSYGGGY.G	CID	98
4	B5SYS6	Spidroin-1	Chymotrypsin	161-185	975.3881 (+2)	1948.7616	1948.8624	-0.1007	0	Y.GGGQGGQSAGAAAAAASAGAGQGGY.G	CID + ETD	133
4	B5SYS6	Spidroin-1	Chymotrypsin	186-209	917.4404 (+2)	1832.8662	1832.8402	0.0261	1	Y.GGLGGQGAGSAAAAAAS*GAGQGGY.G	Ser->Gly (S202) / CID	135
4	B5SYS6	Spidroin-1	Chymotrypsin	189-209	860.3426 (+2)	1718.6706	1718.6783	-0.0077	0	L.GGQ*GAGS*AAAAAASGAGQGGY.G	Deamidated (Q191); Phospho (S195) / CID	106
4	B5SYS6	Spidroin-1	Chymotrypsin	210-236	1079.9740 (+2)	2157.9334	2158.0264	-0.0930	0	Y.GGVGNQGAGRGAGAAAAS*AAGGAGQGGY.N	Ala->Val (A226) / CID	43
4	P19837	Spidroin-1	Chymotrypsin	19-31; 185-197; 286-298	539.7410 (+2)	1077.4674	1077.4840	-0.0165	1	Y.GGLGGQGAGQGGY.G	CID + ETD	66
4	P19837	Spidroin-1	Chymotrypsin	19-34; 185-200; 286-301	653.2959 (+2)	1304.5772	1304.6110	-0.0337	2	Y.GGLGGQGAGQGGYGGL.G	CID + ETD	64
4	P19837	Spidroin-1	Chymotrypsin	22-31; 188-197; 289-298	426.2005 (+2)	850.3864	850.3570	0.0295	0	L.GGQGAGQGGY.G	CID	42
4	P19837	Spidroin-1	Chymotrypsin	22-34; 188-200; 289-301	1078.4620 (+1)	1077.4547	1077.4840	-0.0292	1	L.GGQGAGQGGYGGL.G	CID	45
4	P19837	Spidroin-1	Chymotrypsin	32-59	1079.9692 (+2)	2157.9238	2158.0264	-0.1026	1	Y.GGLGGQGAGQ*GAGAAAAAAGGAGQGGY.G	Gln->Arg (Q41) / CID + ETD	118
4	P19837	Spidroin-1	Chymotrypsin	32-62	787.2927 (+3)	2358.8563	2359.0538	-0.1975	2	Y.GGLGGQGAGQGAGAAAAAAGGAGQGGYGGL.G	CID + ETD	54
4	P19837	Spidroin-1	Chymotrypsin	35-59	966.4341 (+2)	1930.8536	1930.8882	-0.0345	0	L.GGQ*GAGQ*GAGAAAAAAGGAGQGGY.G	Methyl (Q37, Q41) / CID + ETD	117
4	P19837	Spidroin-1	Chymotrypsin	60-89; 576-605	821.0272 (+3)	2460.0598	2460.1382	-0.0784	1	Y.GGLGS*QGAGRGGQGAGAAAAAAGGAGQGGY.G	Phospho (S64, S580) / CID	67
4	P19837	Spidroin-1	Chymotrypsin	63-89; 579-605	1102.3634 (+2)	2202.7122	2202.9080	-0.1958	0	L.GS*QGAGR*GGQGAGAAAAAAGGAGQGGY.G	Phospho (S64, S580); Arg->Phe (R69, R585) / CID	54
4	P19837	Spidroin-1	Chymotrypsin	63-89; 579-605	735.2918 (+3)	2202.8536	2202.9080	-0.0544	0	L.GS*QGAGR*GGQGAGAAAAAAGGAGQGGY.G	Phospho (S64, S580); Arg->Phe (R69, R585) / CID	62
4	P19837	Spidroin-1	Chymotrypsin	63-92; 579-608	886.7416 (+3)	2657.2030	2657.2946	-0.0917	1	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGL.G	CID + ETD	80
4	P19837	Spidroin-1	Chymotrypsin	90-102; 152-164; 198-210; 299-311; 354-366	557.2950 (+2)	1112.5754	1112.6051	-0.0296	1	Y.GGLGSQGAGRGGGL.G	CID	51

4	P19837	Spidroin-1	Chymotrypsin	93-102	444.2492 (+2)	886.4838	886.4621	0.0218	0	L.GSQGA*GRGGL.G	Ala->Val (A97) / CID	32
4	P19837	Spidroin-1	Chymotrypsin	93-123	1279.0403 (+2)	2556.0660	2556.0925	-0.0265	1	L.GSQGAGRGGLG*QGAGAAAAAAGGAGQGGY.G	Gly->Cys (G104) / CID	39
4	P19837	Spidroin-1	Chymotrypsin	103-123; 431-451	903.3680 (+2)	1804.7214	1804.6893	0.0321	0	L.GGQGAGAAAAAAGGAGQGGY.G	CID	105
4	P19837	Spidroin-1	Chymotrypsin	127-151	687.3686 (+3)	2059.0840	2058.9580	0.1260	0	L.GNQGA*GRGGQGAAAAAAGGAGQGGY.G	Ala->Val (A131) / CID + ETD	69
4	P19837	Spidroin-1	Chymotrypsin	152-184	853.0738 (+3)	2556.1996	2556.2178	-0.0182	2	Y.GGLGSQAGRGGGLGGQGAGAAAAAAGGAGQGGY.G	CID + ETD	126
4	P19837	Spidroin-1	Chymotrypsin	155-184	811.0114 (+3)	2430.0124	2430.0350	-0.0226	1	L.GS*QGAGR*GGLGGQGAGAAAAAAGGAGQGGY.G	Phospho (S156); Arg->Phe (161) / CID	59
4	P19837	Spidroin-1	Chymotrypsin	155-197	1164.2290 (+3)	3489.6652	3489.6483	0.0169	3	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	CID + ETD	52
4	P19837	Spidroin-1	Chymotrypsin	201-231	794.3982 (+3)	2380.1728	2380.1592	0.0136	1	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQGGL.G	CID + ETD	66
4	P19837	Spidroin-1	Chymotrypsin	232-255	916.9128 (+2)	1831.8110	1831.8198	-0.0087	0	L.GGQGAGQGAGASAAAAGGAGQGGY.G	CID + ETD	200
4	P19837	Spidroin-1	Chymotrypsin	256-285	839.6989 (+3)	2516.0749	2516.0830	-0.0081	1	Y.GGLGS*QGAGRGGEGAGA*AAAAGGAGQGGY.G	Phospho (S260); Ala->Phe (A273) CID + ETD	78
4	P19837	Spidroin-1	Chymotrypsin	256-298	1178.1664 (+3)	3531.4774	3531.5815	-0.1041	3	Y.GGLGS*QGAGRGGEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Phospho (S260) CID + ETD	60
4	P19837	Spidroin-1	Chymotrypsin	259-285	1066.9704 (+2)	2131.9262	2132.0108	-0.0845	0	L.GSQGAGRGGE*GAGAAAAAAGGAGQGGY.G	Glu->Lys (E268) / CID	111
4	P19837	Spidroin-1	Chymotrypsin	259-298	1064.7769 (+3)	3191.3089	3191.4842	-0.1753	2	L.GSQGAGRGGE*GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Glu->Lys (E268) / CID	131
4	P19837	Spidroin-1	Chymotrypsin	299-329	815.6952 (+3)	2444.0638	2444.1541	-0.0904	2	Y.GGLGSQAGRGGGLGGQGAGAAAAGGAGQGGL.G	ETD	106
4	P19837	Spidroin-1	Chymotrypsin	302-329	747.0295 (+3)	2238.0667	2238.1214	-0.0547	1	L.GSQGAGRGGLG*GQGAGAAAAGGAGQGGL.G	Gly->Lys (G312) / CID + ETD	67
4	P19837	Spidroin-1	Chymotrypsin	312-353	1069.7750 (+3)	3206.3032	3206.4587	-0.1555	1	L.GGQGAGAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	40
4	P19837	Spidroin-1	Chymotrypsin	330-353	930.8681 (+2)	1859.7216	1859.8511	-0.1294	0	L.GGQGA*GQGAGAAAAAAGGAGQGGY.G	Ala->Val (A334) / CID	106
4	P19837	Spidroin-1	Chymotrypsin	421-451	844.0531 (+3)	2529.1375	2529.1858	-0.0483	1	L.GNQGAGRGGGLGGQGAGAAAAAAGGAGQGGY.G	ETD	147
4	P19837	Spidroin-1	Chymotrypsin	452-478	782.0197 (+3)	2343.0373	2343.1177	-0.0804	1	Y.GGLGNQAGRGGQGAAAAAGGAGQGGY.G	ETD	78
4	P19837	Spidroin-1	Chymotrypsin	455-478	1049.4724 (+2)	2096.9302	2097.0536	-0.1234	0	L.GNQGAGRGGQGAAAAAGGAGQGGY.G	CID	45
4	P19837	Spidroin-1	Chymotrypsin	519-531	543.7720 (+2)	1085.5294	1085.5578	-0.0284	1	Y.GGLGSQGSGRGGL.G	CID	41
4	P19837	Spidroin-1	Chymotrypsin	606-636	825.0544 (+3)	2472.1414	2472.1855	-0.0441	2	Y.GGLG*QGVGRGGLGGQGAGAAAAGGAGQGGY.G	Gly->Ser (G610) / CID + ETD	116
4	P19837	Spidroin-1	Chymotrypsin	609-636	749.3412 (+3)	2245.0018	2245.0585	-0.0567	1	L.*G*QGVGRGGLGGQGAGAAAAGGAGQGGY.G	Gly->Ser (G609) / CID + ETD	77
4	P19837	Spidroin-1	Chymotrypsin	619-636	811.2701 (+2)	1620.5256	1620.7084	-0.1827	0	L.GGQGAGAAAAGGAGQGGY.G	CID	30
4	P19837	Spidroin-1	Chymotrypsin	637-653	695.3213 (+2)	1388.6280	1388.7008	-0.0727	0	Y.GGVGSGASAASAAASRL.S	CID	119
4	P19837	Spidroin-1	Chymotrypsin	654-669	780.9019 (+2)	1559.7892	1559.7903	-0.0011	0	L.SSPQASSRVSS*AVSNL.V	Ser->Ala (S664) / CID	62
4	P19837	Spidroin-1	Chymotrypsin	654-680	1265.0828 (+2)	2528.1510	2528.2831	-0.1320	1	L.SSPQASSRVSS*AVSNLVASGPTNSAAL.S	Ser->Ala (S664) / CID	65
4	P19837	Spidroin-1	Chymotrypsin	670-680	494.2318 (+2)	986.4490	986.5033	-0.0542	0	L.VASGPTNSAAL.S	CID + ETD	65
4	P19837	Spidroin-1	Chymotrypsin	730-744	718.3473 (+2)	1434.6800	1434.7103	-0.0303	0	Y.GSAGQAT*QIVGQSVY.Q	Thr->Ala (T736) / CID	82
4	B5SYS6	Spidroin-1	Glu-C	34-45	669.3472 (+2)	1336.6798	1336.6663	0.0136	0	E.LADAFINAFLNE.A	CID + ETD	64
4	B5SYS6	Spidroin-1	Glu-C	99-103	502.2798 (+1)	501.2725	501.2798	-0.0073	0	E.IAAVE.Q	CID	16
4	B5SYS6	Spidroin-1	Glu-C	104-111	760.3901 (+1)	759.3828	759.3763	0.0065	0	E.QGGLSVAE.K	CID	29
4	B5SYS6	Spidroin-1	Glu-C	112-138	968.1525 (+3)	2901.4357	2901.4145	0.0212	0	E.KTNAIADSLNSAFYQTTGAVNVQFVNE.I	CID + ETD	107
4	B5SYS6	Spidroin-1	Glu-C	139-153	842.4368 (+2)	1682.8590	1682.7699	0.0892	0	E.IRSLIS*M*F*AQASANE.V	Phospho (S144); Oxidation (M145); Phe->Pro (F146) / CID	27
4	B5SYS6	Spidroin-1	AspN	45-54	490.7559 (+2)	979.4972	979.4723	0.0249	0	N.EAGRTGAFTA.D	CID	24
4	B5SYS6	Spidroin-1	AspN	58-64	738.2191 (+1)	737.2118	737.2902	-0.0784	1	L.DDMSTIG.D	CID	27
4	B5SYS6	Spidroin-1	AspN	65-71	390.2325 (+2)	778.4504	778.3895	0.0610	0	G.DTLKTAM.D	CID	19
4	B5SYS6	Spidroin-1	AspN	98-110	622.3152 (+2)	1242.6158	1242.6456	-0.0297	1	A.EIAAVEQGGLSVA.E	CID + ETD	31
4	P19837	Spidroin-1	AspN	222-255	863.4600 (+3)	2587.3582	2587.1760	0.1822	0	A.AAGGAGQGGLGGQGAGQGAGASAAAAGGAGQGGY.G	CID + ETD	54
4	P19837	Spidroin-1	AspN	327-353	1030.4900 (+2)	2058.9654	2058.9468	0.0187	0	Q.GGLGGQGAGQGAGAAAAAAGGAGQGGY.G	CID	112
4	P19837	Spidroin-1	Proteinase K	229-234; 549-554	488.3100 (+1)	487.3027	487.2391	0.0636	0	Q.GGLGGQ.G	CID	34
4	P19837	Spidroin-1	Proteinase K	96-105; 204-213; 305-314; 424-433	415.1800 (+2)	828.3454	828.4202	-0.0748	0	Q.GAGRGGGLGGQ.G	CID + ETD	41
4	P19837	Spidroin-1	Proteinase K	691-699	408.2700 (+2)	814.5254	814.4185	0.1070	0	Q.IGASNPLS.G	CID	44

4	P19837	Spidroin-1	Proteinase K	708-712	572.4500 (+1)	571.4427	571.3581	0.0846	0	A.LLEV.V.S	CID	31
4	B5SYS6	Spidroin-1	Subtilisin	58-67	542.2000 (+2)	1082.3854	1082.4438	-0.0583	0	L.DDMSTIGDTL.K	CID + ETD	55
4	B5SYS6	Spidroin-1	Subtilisin	97-110	657.8400 (+2)	1313.6654	1313.6827	-0.0172	0	M.AEIAAVEQGGLSVA.E	CID	46
4	B5SYS6	Spidroin-1	Subtilisin	98-110	622.2900 (+2)	1242.5654	1242.6456	-0.0801	0	A.EIAAVEQGGLSVA.E	CID	55
4	B5SYS6	Spidroin-1	Proteinase 10	34-38	536.2400 (+1)	535.2327	535.2642	-0.0315	0	E.LADAF.I	CID	31
4	B5SYS6	Spidroin-1	Proteinase 10	52-62	630.2300 (+2)	1258.4454	1258.5024	-0.0569	0	A.FTADQLDDMST.I	CID + ETD	50
4	B5SYS6	Spidroin-1	Proteinase 10	52-66	823.3400 (+2)	1644.6654	1644.6825	-0.0171	0	A.FTADQLDDMSTIGDT.L	CID + ETD	64
4	B5SYS6	Spidroin-1	Proteinase 10	54-66	699.2900 (+2)	1396.5654	1396.5664	-0.0010	0	T.ADQLDDMSTIGDT.L	CID	59
4	B5SYS6	Spidroin-1	Proteinase 10	154-168	665.8100 (+2)	1329.6054	1329.5586	0.0469	0	E.VSYGGGYGGGQGGQS.A	CID	41
4	B5SYS6	Spidroin-1	Proteinase 10	154-170	729.7700 (+2)	1457.5254	1457.6172	-0.0917	0	E.VSYGGGYGGGQGGQSAG.A	CID + ETD	51
4	P19837	Spidroin-1	Proteinase 10	8-20; 141-153; 245-257; 343-355; 441-453; 468-480; 626-638	993.4500 (+1)	992.4427	992.4312	0.0115	0	A.AAAGGAGQGGYGG.L	CID	40
4	P19837	Spidroin-1	Proteinase 10	10-20; 81-91; 143-153; 247-257; 345-355; 443-453; 470-480; 628-638	851.3800 (+1)	850.3727	850.3570	0.0158	0	A.AGGAGQGGYGG.L	CID	42
5	B5SYS6	Spidroin-1	Trypsin	49-68	1058.4443 (+2)	2114.8740	2114.9678	-0.0938	0	R.TGAFTADQLDDMSTIGDTLK.T	CID	123
5	B5SYS6	Spidroin-1	Trypsin	69-79	604.8166 (+2)	1207.6186	1207.5979	0.0207	2	K.TAMDKM*ARSNK.S	Oxidation (M74) / CID + ETD	24
5	B5SYS6	Spidroin-1	Trypsin	85-112	956.8128 (+3)	2867.4166	2867.3681	0.0485	0	K.LQALNMAFASSM*AEIAAVEQGGLSVAEK.T	Oxidation (M96) / CID + ETD	89
5	B5SYS6	Spidroin-1	Trypsin	113-140	1015.1518 (+3)	3042.4336	3042.5047	-0.0711	0	K.TNAIADSLNSAFYQTTGAVNVQFVNEIR.S	ETD	84
5	P19837	Spidroin-1	Trypsin	70-99	796.3443 (+3)	2386.0111	2386.1123	-0.1012	0	R.GGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	132
5	P19837	Spidroin-1	Trypsin	134-161	773.0195 (+3)	2316.0367	2316.1319	-0.0953	0	R.GGQGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	79
5	P19837	Spidroin-1	Trypsin	266-308	870.1750 (+3)	3476.6709	3476.5802	0.0907	0	R.GGEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	CID + ETD	40
5	P19837	Spidroin-1	Trypsin	309-363	1049.4836 (+4)	4193.9053	4193.9684	-0.0631	0	R.GGLGGQGAGA*AAAGGAGQGGGAGQGAGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	Ala->Gly (A318) / ETD	67
5	P19837	Spidroin-1	Trypsin	364-397	1049.4844 (+3)	3145.4314	3145.4102	0.0212	0	R.GGLGGQGAGAVAAAAAGGAGQGGY*GGLGSQGAGR.G	Phospho (Y387) / CID + ETD	33
5	P19837	Spidroin-1	Trypsin	398-415	728.3115 (+2)	1454.6084	1454.6974	-0.0890	0	R.GGQGAGAAAAAAGGAGQR.G	CID	35
5	P19837	Spidroin-1	Trypsin	416-427	540.2807 (+2)	1078.5468	1078.5156	0.0313	0	R.GYGGLGNQ*GAGR.G	Gln->Thr (Q423) / CID	37
5	P19837	Spidroin-1	Trypsin	462-488	730.3094 (+3)	2187.9064	2188.0006	-0.0942	0	R.GGQGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	118
5	P19837	Spidroin-1	Trypsin	572-585	725.3201 (+2)	1448.6256	1448.5834	0.0422	0	R.QGGY*GGLGSQGAGR.G	Phospho (Y575) / CID	54
5	P19837	Spidroin-1	Trypsin	616-652	1064.4246 (+3)	3190.2520	3190.2356	0.0164	0	R.GGLGGQGAGAAAAGGAGQGGYGGVGSASAASAAASR.L	CID	49
5	B5SYS6	Spidroin-1	Chymotrypsin	30-42	693.3403 (+2)	1384.6660	1384.6510	0.0150	2	W.SSTELADAFINAF.L	CID	52
5	B5SYS6	Spidroin-1	Chymotrypsin	39-52	740.9131 (+2)	1479.8116	1479.7470	0.0646	2	F.INAFLNEAGRTGAF.T	CID + ETD	81
5	B5SYS6	Spidroin-1	Chymotrypsin	43-52	518.2482 (+2)	1034.4818	1034.5145	-0.0326	1	F.LNEAGRTGAF.T	CID	57
5	B5SYS6	Spidroin-1	Chymotrypsin	43-67	871.4188 (+3)	2611.2346	2611.2072	0.0274	3	F.LNEAGRTGAFTADQLDDMSTIGDTL.K	CID + ETD	115
5	B5SYS6	Spidroin-1	Chymotrypsin	44-67	1306.5412 (+2)	2611.0678	2611.1149	-0.0471	2	L.NEAGRT*GAFTADQLDDMSTIGDTL.K	Phospho (T49) / CID	46
5	B5SYS6	Spidroin-1	Chymotrypsin	53-67	806.3700 (+2)	1610.7254	1610.6982	0.0273	1	F.TADQLDDMSTIGDTL.K	CID	85
5	B5SYS6	Spidroin-1	Chymotrypsin	93-107	725.3487 (+2)	1448.6828	1448.6817	0.0012	0	F.ASSM*AEIAAVEQGGL.S	Oxidation (M96) / CID	80
5	B5SYS6	Spidroin-1	Chymotrypsin	108-125	950.9865 (+2)	1899.9584	1899.9214	0.0371	2	L.SVAEKTNAIADSLNSAFY.Q	CID + ETD	71
5	B5SYS6	Spidroin-1	Chymotrypsin	126-135	1064.5102 (+1)	1063.5029	1063.5298	-0.0269	0	Y.QTTGAVNVQF.V	CID	40
5	B5SYS6	Spidroin-1	Chymotrypsin	136-146	662.8542 (+2)	1323.6938	1323.6857	0.0082	1	F.VNEIRSLISM*F.A	Oxidation (M145) / CID	69
5	B5SYS6	Spidroin-1	Chymotrypsin	147-156	520.2187 (+2)	1038.4228	1038.4618	-0.0389	0	F.AQASANEVSY.G	CID	60
5	B5SYS6	Spidroin-1	Chymotrypsin	147-160	687.2529 (+2)	1372.4912	1372.5895	-0.0983	1	F.AQASANEVSYGGGY.G	CID	87
5	B5SYS6	Spidroin-1	Chymotrypsin	161-185	975.4124 (+2)	1948.8102	1948.8624	-0.0521	0	Y.GGGQGGQSAGAAAAAASAGAGQGGY.G	CID + ETD	140
5	B5SYS6	Spidroin-1	Chymotrypsin	186-209	917.4289 (+2)	1832.8432	1832.8402	0.0031	1	Y.GGLGGQGAGSAAAAAS*GAGQGGY.G	Ser->Gly (S202) / CID	144
5	B5SYS6	Spidroin-1	Chymotrypsin	189-209	860.3455 (+2)	1718.6764	1718.7001	-0.0237	0	L.GGQGAGS*AAAAASGAGQGGY.G	Phospho (S195) / CID	86

5	B5SYS6	Spidroin-1	Chymotrypsin	189-209	895.8872 (+2)	1789.7598	1789.7091	0.0508	0	L.GGQ G *AG S *AAAAAASGAGQGGY.G	Gly->Met (G192); Phospho (S195) / CID	81
5	B5SYS6	Spidroin-1	Chymotrypsin	210-236	1072.9931 (+2)	2143.9716	2144.0108	-0.0391	0	Y.GGVGN Q *GAGRGAGAAAAAAGGAGQGGY.N	Methyl (Q215) / CID	92
5	P19837	Spidroin-1	Chymotrypsin	1-31	811.0358 (+3)	2430.0856	2430.0909	-0.0053	2	QGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	CID + ETD	63
5	P19837	Spidroin-1	Chymotrypsin	19-31; 185-197; 286-298	539.7291 (+2)	1077.4436	1077.4840	-0.0403	1	Y.GGLGGQGAGQGGY.G	CID	59
5	P19837	Spidroin-1	Chymotrypsin	19-34; 185-200; 286-301	653.3004 (+2)	1304.5862	1304.6110	-0.0247	2	Y.GGLGGQGAGQGGYGGL.G	CID + ETD	73
5	P19837	Spidroin-1	Chymotrypsin	32-59	1080.0178 (+2)	2158.0210	2158.0264	-0.0054	1	Y.GGLGGQGAG Q *GAGAAAAAAGGAGQGGY.G	Gln->Arg (Q41) / CID + ETD	123
5	P19837	Spidroin-1	Chymotrypsin	35-59	611.6100 (+3)	1831.8082	1831.8198	-0.0116	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	54
5	P19837	Spidroin-1	Chymotrypsin	60-92; 576-608	886.7281 (+3)	2657.1625	2657.1620	0.0005	2	Y.GGLGS S *QGAG R *GGQGAGAAAAAAGGAGQGGYGGL.G	Phospho (S64, S580); Arg->Phe (R69, R585) / CID	90
5	P19837	Spidroin-1	Chymotrypsin	90-102; 152-164; 198-210; 299-311; 354-366	543.7749 (+2)	1085.5352	1085.5578	-0.0225	1	Y.GGLGSQGAGR G L.G	CID	54
5	P19837	Spidroin-1	Chymotrypsin	93-123	854.4102 (+3)	2560.2088	2560.0841	0.1247	1	L.GSQGAGR G L*GGQGAGAAAAAAGGAGQGGY.G	Leu->Tyr (L102) / CID + ETD	82
5	P19837	Spidroin-1	Chymotrypsin	103-123; 431-451	796.3479 (+2)	1590.6812	1590.7153	-0.0340	0	L.GGQGAGAAAAAAGGAGQGGY.G	CID	39
5	P19837	Spidroin-1	Chymotrypsin	124-151	763.6485 (+3)	2287.9237	2288.0530	-0.1293	1	Y.GGLGN Q *GAGRGG Q *GAAAAAAGGAGQGGY.G	Methyl+Deamidated (Q129, Q136) / CID	62
5	P19837	Spidroin-1	Chymotrypsin	127-151	696.6722 (+3)	2086.9948	2086.9893	0.0055	0	L.GNQGAGRGGQGAAAAAAGGAGQGGY.G	CID + ETD	78
5	P19837	Spidroin-1	Chymotrypsin	152-184	853.0855 (+3)	2556.2347	2556.2178	0.0169	2	Y.GGLGSQGAGR G LGGQGAGAAAAAAGGAGQGGY.G	CID + ETD	131
5	P19837	Spidroin-1	Chymotrypsin	155-184	1215.9508 (+2)	2429.8870	2430.0350	-0.1480	1	L.G S *QGAG R *GGLGGQGAGAAAAAAGGAGQGGY.G	Phospho (S156); Arg->Phe (R161) / CID	46
5	P19837	Spidroin-1	Chymotrypsin	155-184	839.3114 (+3)	2514.9124	2515.0990	-0.1866	1	L.G S *QGAGR G LGGQGAG A *AAAAGGAGQGGY.G	Phospho (S156); Ala->Phe (A172) / CID	56
5	P19837	Spidroin-1	Chymotrypsin	155-197	1164.1616 (+3)	3489.4630	3489.5084	-0.0454	3	L.G S *QGAG R *GGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Phospho (S156); Arg->Phe (R161) / CID	41
5	P19837	Spidroin-1	Chymotrypsin	201-231	794.3919 (+3)	2380.1539	2380.1592	-0.0053	1	L.GSQGAGR G LGGQGAGAAAAAAGGAGQGG.L	CID + ETD	72
5	P19837	Spidroin-1	Chymotrypsin	232-255	916.9068 (+2)	1831.7990	1831.8198	-0.0207	0	L.GGQGAGQGAGASAAAAGGAGQGGY.G	CID	197
5	P19837	Spidroin-1	Chymotrypsin	256-285	787.3263 (+3)	2358.9571	2359.1378	-0.1807	1	Y.GGLGSQGAGR G E*GAGAAAAAAGGAGQGGY.G	Glu->Lys (E268) / CID + ETD	61
5	P19837	Spidroin-1	Chymotrypsin	256-298	1140.4969 (+3)	3418.4689	3418.6112	-0.1423	3	Y.GGLGSQGAGR G E*GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Glu->Lys (E268) / CID + ETD	144
5	P19837	Spidroin-1	Chymotrypsin	259-285	1066.9432 (+2)	2131.8718	2132.0108	-0.1389	0	L.GSQGAGR G E*GAGAAAAAAGGAGQGGY.G	Glu->Lys (E268) / CID	106
5	P19837	Spidroin-1	Chymotrypsin	259-298	1117.1746 (+3)	3348.5020	3348.4294	0.0726	2	L.G S *QGAGR G EG A *GAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Phospho (S260); Ala->Phe (A270) CID + ETD	88
5	P19837	Spidroin-1	Chymotrypsin	299-329	815.7249 (+3)	2444.1529	2444.1541	-0.0013	2	Y.GGLGSQGAGR G LGGQGAGAAAAGGAGQGG.L	CID + ETD	124
5	P19837	Spidroin-1	Chymotrypsin	302-329	747.0415 (+3)	2238.1027	2238.1214	-0.0187	1	L.GSQGAGR G *GLGGQGAGAAAAGGAGQGG.L	Gly->Lys (G309) / CID + ETD	55
5	P19837	Spidroin-1	Chymotrypsin	302-353	1013.9908 (+4)	4051.9341	4051.9306	0.0035	2	L. G *SQGAGR G LGGQGAGAAAAGGAGQGG.LGGQGAGQGAGAAAAAAGGAGQGGY.G	Gly->Lys (G302) / ETD	71
5	P19837	Spidroin-1	Chymotrypsin	330-353	927.9190 (+2)	1853.8234	1853.8017	0.0218	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID	53
5	P19837	Spidroin-1	Chymotrypsin	421-451	810.9968 (+3)	2429.9686	2430.0909	-0.1223	1	L.GNQGAGR R *GGLGGQGAGAAAAAAGGAGQGGY.G	Arg->Glu (R427) / CID + ETD	63
5	P19837	Spidroin-1	Chymotrypsin	452-478	720.3476 (+3)	2158.0210	2158.0264	-0.0055	1	Y.GGLGNQGAGR G Q *GAAAAAGGAGQGGY.G	Gln->Val (Q464) / CID + ETD	83
5	P19837	Spidroin-1	Chymotrypsin	606-636	825.0667 (+3)	2472.1783	2472.1855	-0.0072	2	Y.GGLG G *QGVGRGGLGGQGAGAAAAGGAGQGGY.G	Gly->Ser (G610) / CID + ETD	41
5	P19837	Spidroin-1	Chymotrypsin	609-636	749.3831 (+3)	2245.1275	2245.0585	0.0690	1	L.GGQ G *VGRGGLGGQGAGAAAAGGAGQGGY.G	Gly->Ser (G612) / ETD	51
5	P19837	Spidroin-1	Chymotrypsin	637-653	695.3277 (+2)	1388.6408	1388.7008	-0.0599	0	Y.GGVGSGASAAASAAASRL.S	CID	100
5	P19837	Spidroin-1	Chymotrypsin	654-669	780.9307 (+2)	1559.8468	1559.7903	0.0565	0	L.SSPQASSRV S *AVSNL.V	Ser->Ala (S664) / CID	62
5	P19837	Spidroin-1	Chymotrypsin	654-680	843.7766 (+3)	2528.3080	2528.2831	0.0249	1	L.SSPQASSRV S *AVSNLVASGPTNSAAL.S	Ser->Ala (S664) / CID + ETD	89
5	P19837	Spidroin-1	Chymotrypsin	730-744	740.3714 (+2)	1478.7282	1478.7365	-0.0083	0	Y.GSAGQATQIVGQSVY.Q	CID	45
5	B5SYS6	Spidroin-1	Glu-C	34-45	669.3305 (+2)	1336.6464	1336.6663	-0.0198	0	E.LADAFINAF L NE.A	CID	65
5	B5SYS6	Spidroin-1	Glu-C	104-138	1288.5461 (+3)	3865.6385	3865.8047	-0.1663	1	E.QGGLSVAEKT N AIAD S *LNSAFYQTTGAVNVQFVNE.I	Phospho (S119) / CID + ETD	42
5	B5SYS6	Spidroin-1	Glu-C	112-138	968.1548 (+3)	2901.4426	2901.4145	0.0281	0	E.KT N AIADSLNSAFYQTTGAVNVQFVNE.I	CID + ETD	97
5	B5SYS6	Spidroin-1	Glu-C	139-153	827.4434 (+2)	1652.8722	1652.8192	0.0531	0	E.IRSLIS M *FAQASANE.V	Oxidation (M145) / CID	74
5	B5SYS6	Spidroin-1	AspN	45-54	490.6700 (+2)	979.3254	979.4723	-0.1469	0	N.EAGRTGA F TA.D	CID	29
5	B5SYS6	Spidroin-1	AspN	58-64	738.2800 (+1)	737.2727	737.2902	-0.0175	1	L.DDMSTIG.D	CID	32
5	B5SYS6	Spidroin-1	AspN	59-64	639.2000 (+1)	638.1927	638.2581	-0.0654	0	D.DMSTIG.D	CID	19
5	B5SYS6	Spidroin-1	AspN	65-71	795.4000 (+1)	794.3927	794.3844	0.0083	0	G.DTLKT A M *.D	Oxidation (M71) / CID	39

5	B5SYS6	Spidroin-1	AspN	98-110	622.3000 (+2)	1242.5854	1242.6456	-0.0601	1	A.EIAAVEQGGLSVA.E	CID + ETD	60
5	B5SYS6	Spidroin-1	AspN	98-117	985.9900 (+2)	1969.9654	1970.0320	-0.0665	2	A.EIAAVEQGGLSVAEKTNAIA.D	CID	30
5	P19837	Spidroin-1	Proteinase K	19-24; 32-37; 286-291	488.3400 (+1)	487.3327	487.2391	0.0936	0	Y.GGLGGQ.G	CID	43
5	P19837	Spidroin-1	Proteinase K	66-74	365.7200 (+2)	729.4254	729.3518	0.0736	0	Q.GAGRGGQGA.G	CID + ETD	38
5	B5SYS6	Spidroin-1	Subtilisin	58-67	1083.3800 (+1)	1082.3727	1082.4438	-0.0711	0	L.DDMSTIGDTL.K	CID	38
5	B5SYS6	Spidroin-1	Proteinase 10	52-62	630.2500 (+2)	1258.4854	1258.5024	-0.0169	0	A.FTADQLDDMST.I	CID	48
5	B5SYS6	Spidroin-1	Proteinase 10	52-66	823.3400 (+2)	1644.6654	1644.6825	-0.0171	0	A.FTADQLDDMSTIGDT.L	CID + ETD	62
5	P19837	Spidroin-1	Proteinase 10	8-20; 141-153; 245-257; 343-355; 441-453; 468-480; 626-638	993.4100 (+1)	992.4027	992.4312	-0.0285	0	A.AAAGGAGQGGYGG.L	CID	35
5	P19837	Spidroin-1	Proteinase 10	20-33; 288-300	1078.4400 (+1)	1077.4327	1077.4840	-0.0512	0	G.LGGQGAGQGGYGG.L	CID	45
6	B5SYS5	Spidroin-1	Trypsin	49-68	1066.4849 (+2)	2130.9552	2130.9627	-0.0075	0	R.TGAFTADQLDDM*STIGDTIK.T	Sulphone (M60) / CID	93
6	B5SYS5	Spidroin-1	Trypsin	49-73	912.7489 (+3)	2735.2249	2735.1942	0.0306	1	R.TGAFTADQLDDMSTIGDTIK*TAM*DK.M	Oxidation (M71); Lys->Trp (K68) / CID + ETD	43
6	B5SYS5	Spidroin-1	Trypsin	85-112	952.1636 (+3)	2853.4690	2853.3888	0.0802	0	K.LQALNM*AFASSMAEIAAVEQGGLSVDAK.T	Oxidation (M90) / CID + ETD	63
6	B5SYS5	Spidroin-1	Trypsin	113-140	1005.1752 (+3)	3012.5038	3012.4577	0.0461	0	K.TNAIADSLNSAFYQTTGAANPQFVNEIR.S	ETD	92
6	P19837	Spidroin-1	Trypsin	47-69; 139-168; 341-363; 375-397; 466-488	1067.5362 (+2)	2132.0578	2131.9641	0.0938	0	AAAAAGGAGQGGY*GGLGSQGAGR.G	Phospho (Y59, Y151, Y353, Y387, Y478) / CID + ETD	72
6	P19837	Spidroin-1	Trypsin	70-99	811.0424 (+3)	2430.1054	2430.1385	-0.0331	0	R.GGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	70
6	P19837	Spidroin-1	Trypsin	266-308	1140.4915 (+3)	3418.4527	3418.6112	-0.1585	0	R.GGE*GAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	Glu->Lys (E268) / CID + ETD	141
6	P19837	Spidroin-1	Trypsin	309-363	1049.5031 (+4)	4193.9833	4193.9684	0.0149	0	R.GGLGGQGAGAA*AAGGAGQGGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	Ala->Gly (A319) / ETD	59
6	P19837	Spidroin-1	Trypsin	462-488	837.0252 (+3)	2508.0538	2508.2461	-0.1923	0	R.GGQGAAAAAGGAGQGGYGGLGSQGAGR.G	CID + ETD	87
6	P19837	Spidroin-1	Trypsin	616-652	954.7894 (+3)	2861.3464	2861.3401	0.0063	0	R.GGLGGQGAGAAAAGGAGQGGYGGVGSGASAASAAASR.L	CID + ETD	110
6	B5SYS5	Spidroin-1	Chymotrypsin	43-69	947.8109 (+3)	2840.4109	2840.2804	0.1304	2	F.MNEAGRTGAFTADQLDDMSTIGDTIKT.	CID + ETD	81
6	B5SYS5	Spidroin-1	Chymotrypsin	53-69	912.9739 (+2)	1823.9332	1823.8459	0.0874	1	F.TADQLDDMSTIGDTIKT.	CID	104
6	B5SYS5	Spidroin-1	Chymotrypsin	30-38	470.7005 (+2)	939.3864	939.4185	-0.0321	1	W.SSTELADAF.I	CID	46
6	B5SYS5	Spidroin-1	Chymotrypsin	30-42	693.3724 (+2)	1384.7302	1384.6510	0.0792	2	W.SSTELADAFINAF.M	CID	75
6	B5SYS5	Spidroin-1	Chymotrypsin	39-52	740.8777 (+2)	1479.7408	1479.7470	-0.0062	1	F.INAFMNEAGRTGAF.T	CID + ETD	50
6	B5SYS5	Spidroin-1	Chymotrypsin	43-52	535.2206 (+2)	1068.4266	1068.4658	-0.0392	0	F.M*NEAGRTGAF.T	Oxidation (M43) / CID + ETD	62
6	B5SYS5	Spidroin-1	Chymotrypsin	86-92	405.6807 (+2)	809.3468	809.3742	-0.0273	1	L.QALNM*AF.A	Oxidation (M90) / CID	29
6	B5SYS5	Spidroin-1	Chymotrypsin	93-107	725.3441 (+2)	1448.6736	1448.6817	-0.0080	0	F.ASSM*AEIAAVEQGGL.S	Oxidation (M96) / CID	81
6	B5SYS5	Spidroin-1	Chymotrypsin	93-120	917.4747 (+3)	2749.4023	2749.2964	0.1059	1	F.ASSM*AEIAAVEQGGLSVDAKTNAIADSL.N	Oxidation (M96) / CID + ETD	62
6	B5SYS5	Spidroin-1	Chymotrypsin	93-125	1111.2025 (+3)	3330.5857	3330.5925	-0.0069	3	F.ASSM*AEIAAVEQGGLSVDAKT*NAIADSLNSAFY.Q	Oxidation (M96); Methyl (T113) / ETD	51
6	B5SYS5	Spidroin-1	Chymotrypsin	108-125	943.9768 (+2)	1885.9390	1885.9057	0.0333	2	L.SVDAKTNAIADSLNSAFY.Q	CID + ETD	92
6	B5SYS5	Spidroin-1	Chymotrypsin	126-142	923.5173 (+2)	1845.0200	1844.9381	0.0820	1	Y.QTTGAANPQFVNEIRSL.I	CID + ETD	69
6	B5SYS5	Spidroin-1	Chymotrypsin	136-142	830.3934 (+1)	829.3861	829.4657	-0.0796	0	F.VNEIRSL.I	CID	36
6	B5SYS5	Spidroin-1	Chymotrypsin	136-146	662.8581 (+2)	1323.7016	1323.6857	0.0160	1	F.VNEIRSLIN*M*F.A	Asn->Ser (N144); Oxidation (M145) / CID	71
6	B5SYS5	Spidroin-1	Chymotrypsin	147-156	520.2103 (+2)	1038.4060	1038.4618	-0.0557	0	F.AQS*SANEVSY.G	Ser->Ala (S149) / CID + ETD	66
6	B5SYS5	Spidroin-1	Chymotrypsin	157-182	1014.4151 (+2)	2026.8156	2026.8729	-0.0573	1	Y.GGGYGGQSAGAAAAAAGGGGQGGY.G	CID	131
6	B5SYS5	Spidroin-1	Chymotrypsin	161-182	934.4031 (+2)	1866.7916	1866.6935	0.0982	0	Y.GGQS*AGAAA*S*AAAAGG*GQGGY.G	Phospho (S164, S170); Gly->Ala (G177) / CID	48
6	B5SYS5	Spidroin-1	Chymotrypsin	183-206	959.9995 (+2)	1917.9844	1917.8929	0.0915	1	Y.GNLGGQGAGAAAAAASAAE*QGGY.G	Glu->Gly (E202) / CID + ETD	142
6	B5SYS5	Spidroin-1	Chymotrypsin	207-219	1136.4617 (+1)	1135.4544	1135.5370	-0.0826	1	Y.GGLGSQGAGRGGY.G	CID	39
6	B5SYS5	Spidroin-1	Chymotrypsin	210-219	539.1730 (+2)	1076.3314	1076.3746	-0.0432	0	L.GS*QGAGRGGY.G	Phospho (S211) / CID	31
6	B5SYS5	Spidroin-1	Chymotrypsin	244-260	817.8701 (+2)	1633.7256	1633.5606	0.1651	0	L.*SQGAAAAAGGAGQGGY.G	Phospho (S244) / CID	66
6	P19837	Spidroin-1	Chymotrypsin	1-31	817.8242 (+2)	1633.6338	1633.8011	-0.1673	0	QGAGAAAAAAGGAGQGGY.G	CID + ETD	68

6	P19837	Spidroin-1	Chymotrypsin	19-31; 185-197; 286-298	1078.4591 (+1)	1077.4518	1077.4840	-0.0321	1	Y.GGLGGQGAGQGGY.G	CID	41
6	P19837	Spidroin-1	Chymotrypsin	19-34; 185-200; 286-301	653.2807 (+2)	1304.5468	1304.6110	-0.0641	2	Y.GGLGGQGAGQGGYGGL.G	CID + ETD	62
6	P19837	Spidroin-1	Chymotrypsin	19-59	1102.0940 (+3)	3303.2602	3303.3749	-0.1148	3	Y.GGLGGQGAGQGGY*GGL*GGQGAGQGAGAAAAAAGGAGQGGY.G	Phospho (Y31) / CID	48
6	P19837	Spidroin-1	Chymotrypsin	22-31; 188-197; 289-298	851.3018 9 (+1)	850.2945	850.3570	-0.0624	0	L.GGQGAGQGGY.G	CID	32
6	P19837	Spidroin-1	Chymotrypsin	30-59	797.3112 (+3)	2388.9118	2389.0085	-0.0967	2	G*GGLGGQGAGQGAGAAAAAAGGAGQGGY.G	Phospho (Y31) / CID	59
6	P19837	Spidroin-1	Chymotrypsin	32-59	720.3300 (+3)	2157.9682	2158.0152	-0.0470	1	Y.GGLGGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	96
6	P19837	Spidroin-1	Chymotrypsin	35-59	1016.9444 (+2)	2031.8742	2031.9147	-0.0405	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID	113
6	P19837	Spidroin-1	Chymotrypsin	60-89; 576-605	1215.9782 (+2)	2429.9418	2430.1385	-0.1966	1	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY.G	CID	63
6	P19837	Spidroin-1	Chymotrypsin	63-92; 579-608	886.7077 (+3)	2657.1013	2657.2946	-0.1934	1	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGL.G	CID + ETD	54
6	P19837	Spidroin-1	Chymotrypsin	63-102	1074.1031 (+3)	3219.2875	3219.4338	-0.1463	2	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGL.G	CID + ETD	41
6	P19837	Spidroin-1	Chymotrypsin	90-102; 152-164; 198-210; 299-311; 354-366	568.7209 (+2)	1135.4272	1135.5370	-0.1098	1	Y.GGLGSQGAGRGL.G	CID	52
6	P19837	Spidroin-1	Chymotrypsin	93-123	820.3891 (+3)	2458.1455	2458.1698	-0.0243	1	L.GSQGAGRGLGGQGAGAAAAA*GGAGQGGY.G	Ala->Val (A115) / CID + ETD	54
6	P19837	Spidroin-1	Chymotrypsin	103-123; 431-451	862.8334 (+2)	1723.6522	1723.6039	0.0484	0	L.GGQGAGAAAAAAGGAGQGGY.G	CID	53
6	P19837	Spidroin-1	Chymotrypsin	127-151	696.6449 (+3)	2086.9129	2086.9893	-0.0764	0	L.GNQGAGRGGQAAAAAAGGAGQGGY.G	ETD	83
6	P19837	Spidroin-1	Chymotrypsin	152-184	839.3921 (+4)	2515.1545	2515.1913	-0.0368	2	Y.GGLGSQGAGRGLGGQGAGAAAAAAGGAGQGGY.G	ETD	105
6	P19837	Spidroin-1	Chymotrypsin	155-184	810.9999 (+3)	2429.9779	2430.0350	-0.0571	1	L.GS*QGAGR*GGLGGQGAGAAAAAAGGAGQGGY.G	Phospho (S156); Arg->Phe (R161) / CID	69
6	P19837	Spidroin-1	Chymotrypsin	155-197	1164.1452 (+3)	3489.4138	3489.6119	-0.1981	3	L.GSQGAGRGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	CID + ETD	44
6	P19837	Spidroin-1	Chymotrypsin	165-184	788.8367 (+2)	1575.6588	1575.7390	-0.0801	0	L.GGQGAGAAAAAAGGAGQGGY.G	CID	107
6	P19837	Spidroin-1	Chymotrypsin	201-231	794.4027 (+3)	2380.1863	2380.1592	0.0271	1	L.GSQGAGRGLGGQGAGAAAAAAGGAGQGGL.G	CID + ETD	49
6	P19837	Spidroin-1	Chymotrypsin	232-255	916.8867 (+2)	1831.7588	1831.8198	-0.0609	0	L.GGQGAGQGAGASAAAAGGAGQGGY.G	CID	150
6	P19837	Spidroin-1	Chymotrypsin	256-285	1180.4838 (+2)	2358.9530	2359.1378	-0.1847	1	Y.GGLGSQGAGRGGE*GAGAAAAAAGGAGQGGY.G	Glu->Lys (E268) / CID	110
6	P19837	Spidroin-1	Chymotrypsin	256-298	1140.4726 (+3)	3418.3960	3418.5748	-0.1788	3	Y.GGLGSQGAGRGGE*GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Glu->Lys (E268) / CID	103
6	P19837	Spidroin-1	Chymotrypsin	259-285	1072.9476 (+2)	2143.8806	2143.8556	0.0250	0	L.GS*QGAGR*GGEGAGAAAAAAGGAGQGGY.G	Phospho (S260); Arg->Ser (R265) / CID + ETD	102
6	P19837	Spidroin-1	Chymotrypsin	330-353	620.9536 (+3)	1859.8390	1859.8511	-0.0121	0	L.GGQGAGQGAGAAAAAAGGAGQGGY.G	CID + ETD	61
6	P19837	Spidroin-1	Chymotrypsin	452-481	796.0186 (+2)	2385.0340	2385.1534	-0.1195	2	Y.GGLGNQGAGRGGQGAAAAAGGAGQGGYGGL.G	CID	89
6	P19837	Spidroin-1	Chymotrypsin	532-551	760.3206 (+2)	1518.6266	1518.6811	-0.0545	0	L.GGQGAGAAAAAAGGAGQGGL.G	CID	105
6	P19837	Spidroin-1	Chymotrypsin	654-680	843.7779 (+3)	2528.3119	2528.2831	0.0288	1	L.SSPQASSRVSS*AVSNLVASGPTNSAAL.S	Ser->Ala (S664) / CID + ETD	55
6	P19837	Spidroin-1	Chymotrypsin	730-744	718.3632 (+2)	1434.7118	1434.7103	0.0015	0	Y.GSAGQAT*QIVGQSVY.Q	Thr->Ala (T736) / CID	62
6	P19837	Spidroin-1	Chymotrypsin	522-551	794.3177 (+3)	2379.9313	2380.0340	-0.1027	1	L.GSQSGRGLGLGGQ*GAGAAAAAAGGAGQGGL.G	Gln->Cys (Q534) / CID	50
6	B5SYS5	Spidroin-1	Glu-C	34-45	686.3299 (+2)	1370.6452	1370.6176	0.0276	0	E.LADAFINAFM*NE.A	Oxidation (M43) / CID	69
6	B5SYS5	Spidroin-1	Glu-C	46-69	858.0842 (+3)	2571.2308	2571.2123	0.0185	0	E.AGRTGAFTADQLDDMSTIGDTIKT	CID + ETD	45
6	B5SYS5	Spidroin-1	Glu-C	99-103	502.2771 (+1)	501.2698	501.2798	-0.0100	0	E.IAAVE.Q	CID	18
6	B5SYS5	Spidroin-1	Glu-C	104-138	900.8882 (+4)	3599.5237	3599.7146	-0.1909	0	E.QGGLSVDAKTNAIADSLNSAFYQTTGAANPQFVNE.I	ETD	66
6	B5SYS5	Spidroin-1	Glu-C	139-153	842.4336 (+2)	1682.8526	1682.8297	0.0229	0	E.IRSLINM*FAQSSANE.V	Oxidation (M145) / CID	80
6	B5SYS5	Spidroin-1	AspN	36-44	1058.4300 (+1)	1057.4227	1057.4539	-0.0311	0	A.DAFINAFM*N.E	Oxidation (M43) / CID	43
6	B5SYS5	Spidroin-1	AspN	36-54	1010.4900 (+2)	2018.9654	2018.9156	0.0498	1	A.DAFINAFM*NEAGRTGAFTA.D	Oxidation (M43) / CID	55
6	B5SYS5	Spidroin-1	AspN	45-54	490.7200 (+2)	979.4254	979.4723	-0.0469	0	N.EAGRTGAFTA.D	CID	21
6	B5SYS5	Spidroin-1	AspN	58-64	738.2400 (+1)	737.2327	737.2902	-0.0575	1	L.DDMSTIG.D	CID	31
6	B5SYS5	Spidroin-1	AspN	59-64	639.2500 (+1)	638.2427	638.2581	-0.0154	0	D.DMSTIG.D	CID	28
6	B5SYS5	Spidroin-1	AspN	110-117	402.2100 (+2)	802.4054	802.4185	-0.0130	0	V.DAKTNAIA.D	CID	20
6	B5SYS5	Spidroin-1	AspN	118-137	1072.9700 (+2)	2143.9254	2143.9811	-0.0556	0	A.DSLNSAFYQTTGAANPQFVN.E	CID + ETD	71
6	B5SYS5	Spidroin-1	Proteinase K	55-67	720.3700 (+2)	1438.7254	1438.6134	0.1121	0	A.DQLDDMSTIGDTI.K	CID + ETD	80
6	B5SYS5	Spidroin-1	Proteinase K	56-67	662.8600 (+2)	1323.7054	1323.5864	0.1190	0	D.QLDDMSTIGDTI.K	CID	76

6	B5SYS5	Spidroin-1	Proteinase K	58-67	606.2900 (+2)	1210.5654	1210.5387	0.0267	0	L.DDMSTIGDTIK.T	CID	46
6	B5SYS5	Spidroin-1	Proteinase K	61-68	417.7200 (+2)	833.4254	833.4494	-0.0240	0	M.STIGDTIK.T	CID + ETD	38
6	B5SYS5	Spidroin-1	Proteinase 10	34-38	536.2600 (+1)	535.2527	535.2642	-0.0115	0	E.LADAF.I	CID	31
6	B5SYS5	Spidroin-1	Proteinase 10	48-66	705.9100 (+3)	2114.7082	2114.8847	-0.1765	0	G.RTGAFTADQLDDMSTIGDT.L	ETD	63
6	B5SYS5	Spidroin-1	Proteinase 10	52-66	823.3500 (+2)	1644.6854	1644.6825	0.0029	0	A.FTADQLDDMSTIGDT.I	CID + ETD	58
6	B5SYS5	Spidroin-1	Proteinase 10	154-164	1031.4500 (+1)	1030.4427	1030.4356	0.0071	0	E.VSYGGGYGGQS.A	CID	45

Tabela S4. Caracterização MS/MS através do Modiro®. Sequenciamento das proteínas flagelliform silk protein e spidroin-1 identificadas na 2-DE da seda da teia orbital da aranha *N. clavipes* através de digestão *in-gel* utilizando diferentes enzimas proteolíticas. Número do spot, nome da proteína, enzima utilizada, posição na sequência, *m/z* observado, *m/z* teórico, valor delta entre o *m/z* observado e o *m/z* teórico, valor da carga do íon precursor, sequência peptídica, PTMs, aminoácidos substitutos, método de fragmentação, score significativo e íon score foram listados para todos os peptídeos.

Spot	Código de acesso	Proteína	Enzima	Posição da sequência peptídica na proteína	Valor <i>m/z</i> observado	Valor <i>m/z</i> teórico	Delta	z	Sequência peptídica	PTMs e aminoácidos substitutos / Método de fragmentação	Ion score	Score Sig.
1	O44359	Flagelliform silk protein	Chymotrypsin	1-25	971.96	971.93	0.0224	2	>.GPGGVGPGGSGPGGYGPGGAGPGGY.G	CID	440	100
1	O44359	Flagelliform silk protein	Chymotrypsin	1-35	928.46	928.404	0.056	3	>.GPGGVGPGGSGPGGYGPG ^{54.17} GAGPGGYGPGGSGPGGY.G	Unknown shift / CID + ETD	346	100
1	O44359	Flagelliform silk protein	Chymotrypsin	16 - 25; 459 – 468	405.75	405.67	0.0701	2	Y.GPGGAGP ^{21.14} GGY.G	Unknown shift / CID	257	82.6
1	O44359	Flagelliform silk protein	Chymotrypsin	16 - 35; 459 – 478	754.41	754.39	0.0234	2	Y.GPGGAGPGGYGPGGS ^{-67.86} GPGGY.G	Unknown shift / CID	226	89.8
1	O44359	Flagelliform silk protein	Chymotrypsin	26 - 45; 36 - 55; 46 - 65; 56 - 75; 469 - 488; 479 - 498; 489 - 508; 519 – 538	804.93	804.84	0.0878	2	Y.GPGGSGPGGYG ^{17.17} PGGSGPGGY.G	Unknown shift / CID + ETD	240	100
1	O44359	Flagelliform silk protein	Chymotrypsin	26 - 55; 36 - 65; 46 - 75; 469 - 498; 479 - 508	804.4	804.34	0.0595	3	Y.GPGGSGPGGYGPGGSGPGGYGP ^{33.18} GGSGPGGY.G	Unknown shift / CID + ETD	213	80
1	O44359	Flagelliform silk protein	Chymotrypsin	56 – 80	667.65	667.60	0.0507	3	Y.GPGGSGPGGYGPGG ^{-21.91} SGPGGYGPGGY.G	Unknown shift / ETD	243	98.6
1	O44359	Flagelliform silk protein	Chymotrypsin	66 – 80	660.95	660.76	0.1824	2	Y.GPGGSGPG ^{84.36} GYGPGGY.G	Unknown shift / CID	287	89.5
1	O44359	Flagelliform silk protein	Chymotrypsin	76 – 105	805.76	805.679	0.081	3	Y.GPGGYGPGGSGPGGYGPGGTGPGG ^{23.24} SGPGGY.G	Unknown shift / CID + ETD	231	86.8
1	O44359	Flagelliform silk protein	Chymotrypsin	81 – 105	631.47	631.61	-0.1489	3	Y.GPGG ^{-68.44} SGPGGYGPGGTGPGGSGPGGY.G	Unknown shift / CID + ETD	257	91.6
1	O44359	Flagelliform silk protein	Chymotrypsin	81 – 120	863.91	863.83	0.0743	4	Y.GPGGSGPGGY ^{350.3} GPGGTGPGGSGPGGYGPGGSGPGGSGPGGY.G	Unknown shift / ETD	326	81.3
1	O44359	Flagelliform silk protein	Chymotrypsin	106 – 120	360.31	360.50	-0.1937	3	Y.GPGGSGP ^{-81.57} GGSGPGGY.G	Unknown shift / CID + ETD	233	86.1
1	O44359	Flagelliform silk protein	Chymotrypsin	106 – 140	1105.62	1105.39	0.2281	3	Y.GPGGSGPGGSGPGGYGPGG ^{597.68} SGPGGFPGGSGPGGY.G	Unknown shift / CID + ETD	230	80
1	O44359	Flagelliform silk protein	Chymotrypsin	141 – 160	711.48	711.46	0.0236	2	Y.GPGGSGPGG ^{-73.73} AGPGGVPGGF.G	Unknown shift / CID	228	87.5
1	O44359	Flagelliform silk protein	Chymotrypsin	260 - 287; 683 - 710	1046.54	1046.50	0.0337	2	L.TISGAGGSGPGGAGPGGVGPGGSGPGGV ^{-L} .G	Substitution Val->Leu (V287, V710) / CID	224	87.5
1	O44359	Flagelliform silk protein	Chymotrypsin	317 - 336	751.88	751.81	0.0636	2	Y.GPGGSGPGGAGGAGPGGA ^{Y^{Nitro}} .G	Nitro (Y336) / CID	305	100
1	O44359	Flagelliform silk protein	Chymotrypsin	317 – 342	758.38	758.28	0.0908	3	Y.G ^{297.27} PGGSGPGGAGGAGPGGAYGPGGSY.G	Unknown shift / CID + ETD	250	86.7
1	O44359	Flagelliform silk protein	Chymotrypsin	337 - 342; 399 - 404	817.23	817.21	0.0226	1	Y.GP ²⁸⁰ GGSY.G	Unknown shift / CID	204	83.5
1	O44359	Flagelliform silk protein	Chymotrypsin	337 – 357	831.94	831.86	0.0792	2	Y.GPGGSYGPGGSGPGGAGGPY.G	CID	399	100
1	O44359	Flagelliform silk protein	Chymotrypsin	337 – 371	937.81	937.734	0.076	3	Y.GPGGSYGP ^{38.23} GSGGPGGAGGPYPGGEGPGGAGGPY.G	Unknown shift / CID + ETD	316	100
1	O44359	Flagelliform silk protein	Chymotrypsin	343 – 357	565.32	565.30	0.0214	2	Y.GPGGS ^{-14.86} GGPGGAGGPY.G	Unknown shift / CID	476	100
1	O44359	Flagelliform silk protein	Chymotrypsin	343 – 380	791.1	791.07	0.0225	4	Y.GPGGSGGPGGAGGPYPGGEGPGGAGG ^{193.09} PYGPGGAGGPY.G	Unknown shift / ETD	308	85.7
1	O44359	Flagelliform silk protein	Chymotrypsin	358 - 371; 420 - 433	573.81	573.74	0.0609	2	Y.G ^{17.12} PGGEGPGGAGGPY.G	Unknown shift / CID	278	99.1
1	O44359	Flagelliform silk protein	Chymotrypsin	358 – 380	930.46	930.40	0.0543	2	Y.GPGGEGPGGAGGP ^{17.11} YGPGGAGGPY.G	Unknown shift / CID	318	100
1	O44359	Flagelliform silk protein	Chymotrypsin	372 - 380; 381 - 389; 405 - 413; 791 - 799	732.39	732.33	0.0589	1	Y.GPGGAGGPY.G	CID	204	100
1	O44359	Flagelliform silk protein	Chymotrypsin	372 – 389	731.45	731.32	0.1267	2	Y.GPGGAGGPYGP ^{0x} GGAGGPY.G	Hydroxylation (P382) / CID	241	96

1	O44359	Flagelliform silk protein	Chymotrypsin	381 – 398	752.4	752.32	0.0714	2	Y.GPGGAGGPYGPGGEGGPY.G	CID	402	100
1	O44359	Flagelliform silk protein	Chymotrypsin	390 – 413	1017.01	1016.93	0.0752	2	Y.GPGGEGGPYGP GSY ^{11.15} GPGGAGGPY.G	Unknown shift / CID	204	98.8
1	O44359	Flagelliform silk protein	Chymotrypsin	399 – 419	886.98	886.96	0.0212	2	Y.GPGGS ^{-5.82} YGPGGAGGPYGP GGPY.G	Unknown shift / CID	404	100
1	O44359	Flagelliform silk protein	Chymotrypsin	405 – 433	791.1	791.017	0.0827	3	Y.GPGGAGGPYGP GGPYGP GGE GPGGAGGPY.G	ETD	325	100
1	O44359	Flagelliform silk protein	Chymotrypsin	414 – 433	837.93	837.86	0.0643	2	Y.G ^{17.13} PGGPYGP GGE GPGGAGGPY.G	Unknown shift / CID	369	100
1	O44359	Flagelliform silk protein	Chymotrypsin	434 – 458	479.22	479.22	-0.0011	4	Y.GPGGVGP GSGSGPGGYGP ⁻⁴⁵ GGSGPGGY.G	Unknown shift / ETD	280	80.9
1	O44359	Flagelliform silk protein	Chymotrypsin	449 – 468	825.39	825.34	0.0452	2	Y.G ^{74.09} PGGSGPGGYGP GAGPGGY.G	Unknown shift / CID	278	85.9
1	O44359	Flagelliform silk protein	Chymotrypsin	449 – 478	868.22	868.00	0.2112	3	Y.GPGGSG ^{240.63} PGGYGP GAGPGGYGP GSGPGGY.G	Unknown shift / CID + ETD	242	83.1
1	O44359	Flagelliform silk protein	Chymotrypsin	499 – 518	741.96	741.95	0.0137	2	Y.GPGGSGPGGYSGGAG ^{-82.74} PGGY.G	Unknown shift / CID	252	91.6
1	O44359	Flagelliform silk protein	Chymotrypsin	509 – 518	913.51	913.33	0.1781	1	Y.GSGGAGPGGY ^{134.18} .G	Unknown shift / CID	224	98
1	O44359	Flagelliform silk protein	Chymotrypsin	509 – 528	523.67	523.55	0.1113	3	Y.GSG ^{3.33} GAGPGGYGP GSGPGGY.G	Unknown shift / CID + ETD	287	87.2
1	O44359	Flagelliform silk protein	Chymotrypsin	529 – 578	965.44	965.41	0.0284	4	Y.GPGGSGPGGYGP GGTGP GGTGP GSGPGGYGP GSGPGGSGPGGSGPGGY.G	CID + ETD	290	80.8
1	O44359	Flagelliform silk protein	Chymotrypsin	559 – 578	796.4	796.34	0.0578	2	Y.GPGGSGPGGS ^{-Y} GP GSGPGGY.G	Substitution Ser->Tyr (S568) / CID	425	100
1	O44359	Flagelliform silk protein	Chymotrypsin	559 – 588	608.61	608.52	0.0881	4	Y.GPGGSGPGG ^{-R} SGPGGSGPGGYGPSGSGPGGY.G	Substitution Gly->Arg (G567) / ETD	380	93.7
1	O44359	Flagelliform silk protein	Chymotrypsin	579 – 598	549.25	549.23	0.0211	3	Y.GPSGSGPGG ^{-5.95} YGPSGSGPGGY.G	Unknown shift / CID	263	96.4
1	O44359	Flagelliform silk protein	Chymotrypsin	746 – 769	816.43	816.28	0.1464	3	Y.GPGGSGSGGVPGG ^{505.44} YGP GSGGGFY.G	Unknown shift / CID + ETD	355	98.6
1	O44359	Flagelliform silk protein	Chymotrypsin	770 – 784	685.08	685.29	-0.2166	2	Y.GPGGSEGPYGP ^{-13.42} SGTY.G	Unknown shift / CID	206	80.6
1	O44359	Flagelliform silk protein	Chymotrypsin	785 – 799	584.04	584.25	-0.2198	2	Y.GSGGGYGP GAGG ^{-43.43} PY.G	Unknown shift / CID	204	83.3
1	O44359	Flagelliform silk protein	Chymotrypsin	840 – 848	617.84	617.74	0.0986	2	F.NYQ ^{101.2} MFGNML.S	Unknown shift / CID	291	84.9
1	O44359	Flagelliform silk protein	Chymotrypsin	849 – 866	610.71	610.68	0.0325	3	L.SQYSSGSGTCNP ^{-10.69} NNVNL.M	Unknown shift / CID + ETD	322	87.7
1	O44359	Flagelliform silk protein	Chymotrypsin	852 – 870	674.07	673.95	0.1159	3	Y.SSG ^{127.35} SGTCNPNNVNLMDAL.L	Unknown shift / CID	206	87.9
1	O44359	Flagelliform silk protein	Chymotrypsin	875 – 897	858.15	858.00	0.1428	3	L.HCLSNHGSSSFAPSPT ^{236.43} PAAM ^{Ox} SAY.S	Unknown shift / CID + ETD	335	100
1	O44359	Flagelliform silk protein	Chymotrypsin	878 – 885	412.75	412.67	0.0775	2	L.S ^{2.15} NHGSSSF.A	Unknown shift / CID	239	99.9
1	O44359	Flagelliform silk protein	Chymotrypsin	886 – 907	794.84	794.68	0.1561	3	F.APSP ^{74.47} TPAAM ^{Ox} SAYSNSVGRM ^{Ox} FAY.-	Unknown shift / CID + ETD	237	80.8
1	O44358	Flagelliform silk protein	Glu-C	1-25	850.37	850.46	-0.0951	3	>.MGKGRHDTKAK ^{-50.28} AKAMQVALASSIAE.L	Unknown shift / CID + ETD	310	89.6
1	O44358	Flagelliform silk protein	Glu-C	13 – 25	659.83	659.82	0.012	2	A.KAMQVALASSIAE.L	CID	220	100
1	O44358	Flagelliform silk protein	Glu-C	26 – 30	544.31	544.33	-0.0241	1	E.LVIAE.S	CID	255	99.8
1	O44358	Flagelliform silk protein	Glu-C	31 – 61	809.12	809.15	-0.0308	4	E.SSGGDVQRKTNVISNALRNALMSTTGSPNEE.F	/ ETD	396	99.8
1	O44358	Flagelliform silk protein	Glu-C	62 – 80	815.35	815.33	0.023	3	E.FVHEVQDLIQM ^{Ox} LS ^{127.92} QEQINE.V	Unknown shift / CID + ETD	273	82.6
1	O44358	Flagelliform silk protein	Glu-C	66 – 80	894.42	894.44	-0.0258	2	E.VQDLIQMLSQEQINE.V	/ CID	529	100
1	O44358	Flagelliform silk protein	Glu-C	77 – 109	906.64	906.61	0.0242	4	E.QINE ^{439.1} VDTSGPGQYYRSSSSGGGGGGQGGPVVTE.T	Unknown shift / ETD	388	81.8
1	O44358	Flagelliform silk protein	Glu-C	81 – 109	900.7	900.74	-0.0426	3	E.VDTSGPGQYYRSSSSGGGGGGQGGPVVTE.T	/ CID + ETD	488	100
1	O44358	Flagelliform silk protein	Glu-C	504 - 521	1110.47	1110.44	0.0207	2	E.DLDITIDGA ^{346.04} DGPITISEE.L	Unknown shift / CID	231	90.9
2	O44359	Flagelliform silk protein	Chymotrypsin	1 - 15; 434 - 448	594.76	594.76	0	2	>.GP ^{Ox} GGVGP GSGPGGY.G	Hydroxylation (P2; P435) / CID	277	100
2	O44359	Flagelliform silk protein	Chymotrypsin	1 – 25	979.94	979.935	0.0078	2	>.GPGGVGP GSGPGGYGP ^{Ox} GGAGPGGY.G	Hydroxylation (P17) / CID	243	100

2	O44359	Flagelliform silk protein	Chymotrypsin	1 – 35	928.71	928.70	0.0172	3	>.GPGGVGPGGSGPGGYGPGG ^{54.92} AGPGGYGPGGSGPGGY.G	Unknown shift / CID	210	100
2	O44359	Flagelliform silk protein	Chymotrypsin	1 – 45	1172.51	1172.51	0.0016	3	>.GPGGVGPGGSGPGGYGPGGAGPGGYGPGGSGPGGYGPGGSGPGGY.G	CID + ETD	251	100
2	O44359	Flagelliform silk protein	Chymotrypsin	16 - 45; 459 - 488	1189.50	1189.50	0.0009	2	Y.GP ^{ox} GGAGPGGYGPGGSGPGGYGPGGSGPGGY.G	Hydroxylation (P17, P460) / CID	290	100
2	O44359	Flagelliform silk protein	Chymotrypsin	16 - 45; 459 - 488	1197.49	1197.50	-0.0095	2	Y.GPGGAGPGGYGPGGSGPGGYGPGGSGPGGY.G	CID	207	100
2	O44359	Flagelliform silk protein	Chymotrypsin	26 - 35; 36 - 45; 46 - 55; 56 - 65; 66 - 75; 81 - 90; 131 - 140; 449 - 458; 469 - 478; 479 - 488; 489 - 498; 499 - 508; 519 - 528; 529 - 538; 599 - 608	403.17	403.17	0.0003	2	Y.GPGGSGPGGY.G	CID	254	100
2	O44359	Flagelliform silk protein	Chymotrypsin	26 - 45; 36 - 55; 46 - 65; 56 - 75; 469 - 488; 479 - 498; 489 - 508; 519 - 538	796.34	796.34	0.0001	2	Y.GPGGSGPGGYGPGGSGPGGY.G	CID	298	100
2	O44359	Flagelliform silk protein	Chymotrypsin	26 - 45; 36 - 55; 46 - 65; 56 - 75; 469 - 488; 479 - 498; 489 - 508; 519 - 538	812.33	812.33	0.001	2	Y.GP ^{ox} GGSGPGGYGP ^{ox} GGSGPGGY.G	Hydroxylation (P27, P37, P47, P57, P67, P470, P480, P490, P500, P520, P530) / CID	230	100
2	O44359	Flagelliform silk protein	Chymotrypsin	26 - 55; 36 - 65; 46 - 75; 469 - 498; 479 - 508	793.34	793.34	0.0026	3	Y.GPGGSGPGGYGPGGSGPGGYGPGGSGPGGY.G	CID + ETD	242	100
2	O44359	Flagelliform silk protein	Chymotrypsin	26 - 65; 36 - 75; 469 - 508	1055.45	1055.45	0.0007	3	Y.GPGGSGPGGYGPGGSGPGGYGPGGSGPGGYGPGGSGPGGY.G	CID + ETD	242	99.7
2	O44359	Flagelliform silk protein	Chymotrypsin	56 – 80	973.92	973.91	0.0063	2	Y.GPGGSGPGGYGPGGSGPGGY ^{-S-} GPGGY.G	Substitution Tyr->Ser (Y75) / CID	201	99.9
2	O44359	Flagelliform silk protein	Chymotrypsin	66 – 80	610.76	610.77	-0.0005	2	Y.GPGGSGPGGYGPGGY.G	CID	308	100
2	O44359	Flagelliform silk protein	Chymotrypsin	66 – 90	974.41	974.43	-0.0133	2	Y.GPGGSGPGGYGP ^{-75.02} GGYGPGGSGPGGY.G	Unknown shift / CID	210	100
2	O44359	Flagelliform silk protein	Chymotrypsin	76 – 90	586.76	586.74	0.0249	2	Y.GPGGY ^{-63.98} GPGGSGPGGY.G	Unknown shift / CID	344	100
2	O44359	Flagelliform silk protein	Chymotrypsin	81 – 105	981.91	981.89	0.0248	2	Y.GPGGSGPGGYGPG ^{1.98} GTGPGGSGPGGY.G	Unknown shift / CID + ETD	221	100
2	O44359	Flagelliform silk protein	Chymotrypsin	81 – 120	1030.77	1030.77	-0.0027	3	Y.GPGGSGPGGYGPGGT ⁻¹² GPGGSGPGGYGPGGSGPGGSGPGGY.G	Unknown shift / CID + ETD	205	84.2
2	O44359	Flagelliform silk protein	Chymotrypsin	91 – 105	618.76	618.75	0.0045	2	Y.GPGGT ^{62.01} GPGGSGPGGY.G	Unknown shift / CID	202	100
2	O44359	Flagelliform silk protein	Chymotrypsin	91 – 120	558.52	558.48	0.0423	4	Y.GPGGTGPGGSGP ^{-84.92} GGYGPGGSGPGGSGPGGY.G	Unknown shift / ETD	253	86.9
2	O44359	Flagelliform silk protein	Chymotrypsin	106 – 120	580.75	580.752	-0.0013	2	Y.GPGGSGPGGSGPGGY.G	CID	274	100
2	O44359	Flagelliform silk protein	Chymotrypsin	106 – 135	1181.50	1181.50	-0.0022	2	Y.GPGGSGPGGSGPGGYGPGGSGPGGFPGGS ^{-Y-} .G	Substitution Ser->Tyr (S135) / CID	329	100
2	O44359	Flagelliform silk protein	Chymotrypsin	106 – 140	1050.45	1050.39	0.062	3	Y.GPGGSGPGGSGPGGYGPGGSGPGGFPGGS ^{432.18} GPGGY.G	Unknown shift / CID + ETD	239	99.6
2	O44359	Flagelliform silk protein	Chymotrypsin	141 – 160	709.83	709.84	-0.019	2	Y.GPGGS ^{-77.03} GPGGAGPGGVGPGGF.G	Unknown shift / CID	218	91
2	O44359	Flagelliform silk protein	Chymotrypsin	243 - 259; 666 - 682	830.42	830.43	-0.0118	2	L.DITID ^{-99.01} GADGPITISEEL.T	Unknown shift / CID	210	100
2	O44359	Flagelliform silk protein	Chymotrypsin	317 – 336	743.33	743.33	-0.0001	2	Y.GPGGSGPGGAGGAGGPGGAY.G	/ CID + ETD	414	100
2	O44359	Flagelliform silk protein	Chymotrypsin	317 – 342	861.71	861.62	0.0906	3	Y.GPGGSGPGGAGGAGGPGGAYGPGGSY ^{607.27} .G	Unknown shift / CID + ETD	217	99.9
2	O44359	Flagelliform silk protein	Chymotrypsin	337 – 357	839.85	839.85	0.0013	2	Y.GPGGSYGPGGSGGPGGAGGPY.G	CID	201	100
2	O44359	Flagelliform silk protein	Chymotrypsin	337 – 371	920.07	920.04	0.0329	3	Y.GPGGSYGPGG ^{-14.98} SGGPGGAGGPYGPGEPPGGAGGPY.G	Unknown shift / CID + ETD	233	100
2	O44359	Flagelliform silk protein	Chymotrypsin	343 – 357	580.75	580.752	-0.0015	2	Y.GPGGSGGPGGAGGPY.G	CID	314	100
2	O44359	Flagelliform silk protein	Chymotrypsin	343 – 371	757.66	757.66	-0.0013	3	Y.GPGGSGGPGGAGGPYGPGEPPGGAGGPY.G	CID	257	100
2	O44359	Flagelliform silk protein	Chymotrypsin	343 – 380	995.43	995.43	0.0036	3	Y.GPGGSGGPGGAGGPYGPGEPPGGAGGPYGPGGAGGPY.G	CID	211	89.2
2	O44359	Flagelliform silk protein	Chymotrypsin	358 - 371; 420 – 433	566.32	566.24	0.0789	2	Y.GPGGEGPGA ^{2.16} GGPY.G	Unknown shift / CID	217	99.1
2	O44359	Flagelliform silk protein	Chymotrypsin	358 – 380	928.907	928.91	-0.0066	2	Y.GPGGE ^{Me} GPGGAGGPYGPGGAGGPY.G	Methylation (E362) / CID	231	100
2	O44359	Flagelliform silk protein	Chymotrypsin	358 – 380	929.90	929.90	0.0009	2	Y.GPGGEGPGGAGGPYGPGGAGGPY.G	CID	232	100
2	O44359	Flagelliform silk protein	Chymotrypsin	358 – 389	714.78	714.78	0.0039	4	Y.G ^{300.01} PGGEGPGGAGGPYGPGGAGGPYGPGGAGGPY.G	Unknown shift / ETD	255	95.5

2	O44359	Flagelliform silk protein	Chymotrypsin	372 - 380; 381 - 389; 405 - 413; 791 - 799	366.66	366.66	0.0004	2	Y.GPGGAGGPY.G	ETD	230	99.9
2	O44359	Flagelliform silk protein	Chymotrypsin	372 – 389	731.32	731.32	-0.0007	2	Y.GPGGAGGPYGP ^{ox} GGAGGPY.G	Hydroxylation (P382) / CID	263	100
2	O44359	Flagelliform silk protein	Chymotrypsin	372 – 389	739.32	739.32	-0.0008	2	Y.GPGGAGGPYGPGGAGGPY.G	CID	253	100
2	O44359	Flagelliform silk protein	Chymotrypsin	372 – 398	1116.97	1116.98	-0.005	2	Y.GPGGAGGPYGP ^{ox} GGAGGPYGPGE ^{ox} GPY.G	Hydroxylation (P382) / CID	225	100
2	O44359	Flagelliform silk protein	Chymotrypsin	381 – 398	760.32	760.326	-0.0025	2	Y.GP ^{ox} GGAGGPYGPGE ^{ox} GPY.G	Hydroxylation (P382) / CID	233	100
2	O44359	Flagelliform silk protein	Chymotrypsin	381 – 398	768.32	768.32	0.0054	2	Y.GP ^{ox} GGAGGPYGPGE ^{ox} GPY.G	Hydroxylation (P382) / CID	200	100
2	O44359	Flagelliform silk protein	Chymotrypsin	381 – 404	672.63	672.61	0.0332	3	Y.GPGGAGGPYGPGE ^{ox} GPYGP ^{ox} GSY ^{-5.97} .G	Unknown shift / CID + ETD	273	100
2	O44359	Flagelliform silk protein	Chymotrypsin	390 – 404	651.79	651.76	0.0328	2	Y.GPGE ^{ox} GPYGP ^{ox} GSY ^{-5.97} .G	Unknown shift / CID	299	100
2	O44359	Flagelliform silk protein	Chymotrypsin	399 – 413	631.302	631.27	0.0265	2	Y.G ^{11.05} P ^{ox} GSYGP ^{ox} GGAGGPY.G	Unknown shift / CID	203	100
2	O44359	Flagelliform silk protein	Chymotrypsin	399 – 419	894.89	894.90	-0.0029	2	Y.GP ^{ox} S ^{-P} YGP ^{ox} GGAGGPYGP ^{ox} GPY.G	Substitution Ser->Pro (S403) / CID	225	100
2	O44359	Flagelliform silk protein	Chymotrypsin	405 – 419	646.78	646.78	0.0084	2	Y.GP ^{ox} GGAGGPYGP ^{ox} GPY.G	Hydroxylation (P406, P415) / CID	201	100
2	O44359	Flagelliform silk protein	Chymotrypsin	405 – 419	638.78	638.78	0.0014	2	Y.GP ^{ox} GGAGGPYGP ^{ox} GPY.G	Hydroxylation (P406) / CID	297	100
2	O44359	Flagelliform silk protein	Chymotrypsin	405 – 419	638.80	638.78	0.0258	2	Y.GP ^{ox} GGAGGPYGP ^{ox} GPY.G	Hydroxylation (P415) / CID	203	97.1
2	O44359	Flagelliform silk protein	Chymotrypsin	405 – 433	796.34	796.349	-0.0049	3	Y.GP ^{ox} GGAGGPYGP ^{ox} GPYGPGE ^{ox} GP ^{ox} GGAGGPY.G	Hydroxylation (P406) / CID	202	100
2	O44359	Flagelliform silk protein	Chymotrypsin	414 – 433	837.36	837.36	0.0005	2	Y.GP ^{ox} GPYGPGE ^{ox} GP ^{ox} GGAGGPY.G	CID	284	100
2	O44359	Flagelliform silk protein	Chymotrypsin	414 – 433	837.36	837.36	0.0012	2	Y.GP ^{ox} GPYGPGE ^{ox} GP ^{ox} GGAGGPY.G	Hydroxylation (P415) / CID	234	100
2	O44359	Flagelliform silk protein	Chymotrypsin	414 – 448	937.75	937.75	-0.0011	3	Y.GP ^{ox} GPYGPGE ^{ox} GP ^{ox} GGAGGPYGP ^{ox} GVGP ^{ox} SGP ^{ox} GGY.G	CID + ETD	202	97.3
2	O44359	Flagelliform silk protein	Chymotrypsin	420 – 448	1142.00	1142.00	0.0088	2	Y.GPGE ^{ox} GP ^{ox} GGAGGPYGP ^{ox} GVGP ^{ox} SGP ^{ox} GGY.G	CID + ETD	214	100
2	O44359	Flagelliform silk protein	Chymotrypsin	420 – 458	1018.78	1018.76	0.0213	3	Y.GPGE ^{ox} GP ^{ox} GGAGGPYGP ^{ox} GVGP ^{ox} SGP ^{ox} GGYGP ^{ox} GG ^{-14.98} SGP ^{ox} GGY.G	Unknown shift / CID + ETD	216	99.9
2	O44359	Flagelliform silk protein	Chymotrypsin	434 – 458	667.628	667.62	-0.0013	3	Y.GP ^{ox} GVGP ^{ox} SGP ^{ox} GGYGP ^{ox} GG ^{-14.98} SGP ^{ox} GGY.G	Acetylation (S453) / CID	205	100
2	O44359	Flagelliform silk protein	Chymotrypsin	434 – 468	906.39	906.404	-0.0104	3	Y.GP ^{ox} GV ^{-12.02} GP ^{ox} SGP ^{ox} GGYGP ^{ox} GG ^{-12.02} SGP ^{ox} GGYGP ^{ox} GGAGP ^{ox} GGY.G	Unknown shift / CID + ETD	207	99.5
2	O44359	Flagelliform silk protein	Chymotrypsin	449 – 478	793.674	793.672	0.002	3	Y.GP ^{ox} SGP ^{ox} GGYGP ^{ox} GA ^{16.99} GP ^{ox} GGYGP ^{ox} SGP ^{ox} GGY.G	Unknown shift / CID + ETD	209	99.9
2	O44359	Flagelliform silk protein	Chymotrypsin	499 – 518	604.95	604.892	0.0653	3	Y.GP ^{ox} SGP ^{ox} GGYGP ^{ox} GA ^{16.99} GP ^{ox} GGYGP ^{ox} SGP ^{ox} GGY.G	Unknown shift / CID + ETD	224	96.5
2	O44359	Flagelliform silk protein	Chymotrypsin	499 – 528	811.31	811.324	-0.0132	3	Y.GP ^{ox} SGP ^{ox} GGYGP ^{ox} GA ^{16.99} GP ^{ox} GGYGP ^{ox} SGP ^{ox} GGY.G	CID	202	84.1
2	O44359	Flagelliform silk protein	Chymotrypsin	519 – 558	1055.78	1055.78	0.0054	3	Y.GP ^{ox} SGP ^{ox} GGYGP ^{ox} SGP ^{ox} GGYGP ^{ox} GGTGP ^{49.01} GTGP ^{ox} SGP ^{ox} GGY.G	Unknown shift / CID + ETD	210	98.7
2	O44359	Flagelliform silk protein	Chymotrypsin	559 – 578	532.22	532.22	0.0056	3	Y.GP ^{ox} SGP ^{ox} GGSGP ^{ox} GGSGP ^{ox} GG ^{79.02} Y.G	Unknown shift / CID + ETD	201	90.8
2	O44359	Flagelliform silk protein	Chymotrypsin	559 – 588	817.35	817.33	0.0252	3	Y.GP ^{ox} SGP ^{ox} GGSGP ^{ox} GGSGP ^{ox} GGYGP ^{ox} SGSGP ^{ox} GGY.G	Unknown shift / CID + ETD	204	98.9
2	O44359	Flagelliform silk protein	Chymotrypsin	579 - 588; 589 - 598	449.69	449.68	0.0145	2	Y.GPSGS ^{63.03} GP ^{ox} GGY.G	Unknown shift / CID	201	89.4
2	O44359	Flagelliform silk protein	Chymotrypsin	589 – 608	804.34	804.34	-0.0075	2	Y.G ^{-14.01} PSGSGP ^{ox} GGYGP ^{ox} SGSGP ^{ox} GGY.G	Unknown shift / CID	243	100
2	O44359	Flagelliform silk protein	Chymotrypsin	746 – 768	889.89	889.89	0.0012	2	Y.GP ^{ox} SGSGGVGP ^{ox} GGYGP ^{ox} SGSGGF.Y	CID	211	100
2	O44359	Flagelliform silk protein	Chymotrypsin	746 – 769	971.42	971.42	-0.0016	2	Y.GP ^{ox} SGSGGVGP ^{ox} GGYGP ^{ox} SGSGGF.Y	CID	201	100
2	O44359	Flagelliform silk protein	Chymotrypsin	769 – 778	482.74	482.72	0.0214	2	F.YGP ^{-18.93} GGSEGPY.G	Unknown shift / CID + ETD	204	86.3
2	O44359	Flagelliform silk protein	Chymotrypsin	770 – 778	533.26	533.17	0.0919	2	Y.GP ^{ox} SGSEGP ^{245.18} Y.G	Unknown shift / CID	204	96.5
2	O44359	Flagelliform silk protein	Chymotrypsin	770 – 784	691.796	691.79	-0.0006	2	Y.GP ^{ox} SGSEGPYGP ^{ox} SGTY.G	CID	208	100
2	O44359	Flagelliform silk protein	Chymotrypsin	779 – 790	642.79	642.72	0.071	2	Y.G ^{225.14} PSGTYGSGGGY.G	Unknown shift / CID	211	99.5
2	O44359	Flagelliform silk protein	Chymotrypsin	785 – 799	620.76	620.76	-0.0006	2	Y.GSGGGYGP ^{ox} G ^{-S} AGGPY.G	Substitution Gly->Ser (G794) / CID	203	99.9

2	O44359	Flagelliform silk protein	Chymotrypsin	785 – 808	635.62	635.61	0.0148	3	Y.GSGGGYGPGGAGGPY ^N GPSPGGAY.G	Substitution Tyr->Asn (Y799) / CID + ETD	212	96.5
2	O44359	Flagelliform silk protein	Chymotrypsin	800 – 817	712.80	712.83	-0.0291	2	Y.GPGSPGGAYGPGSP ^{-81.05} GGAY.Y	Unknown shift / CID	209	100
2	O44359	Flagelliform silk protein	Chymotrypsin	818 – 839	794.01	794.01	-0.0027	3	Y.YPSSRVPDM ^{Ox} VN ^{Deamid} GIM ^{Ox} SAM ^{Ox} QSGSF.N	Deamidation (N828); Oxidation (M826, M831, M834) / CID	208	80.9
2	O44359	Flagelliform silk protein	Chymotrypsin	818 – 841	860.40	860.36	0.0432	3	Y.YPSSRVPDM ^{-28.95} VNGIMSAMQSGSFNY.Q	Unknown shift / CID + ETD	211	85.2
2	O44359	Flagelliform silk protein	Chymotrypsin	842 – 848	421.77	421.69	0.0857	2	Y.QMF ^{-2.17} GNML.S	Unknown shift / CID	206	80
2	O44359	Flagelliform silk protein	Chymotrypsin	849 – 870	1113.99	1114.01	-0.0189	2	L.SQYSSG ^{-117.03} SGTC ^{CAMe} NPNNVNVLM ^{Ox} DAL.L	Unknown shift / CID	213	94.7
2	O44359	Flagelliform silk protein	Chymotrypsin	852 – 871	1032.48	1032.47	0.0137	2	Y.SSGSGTC ^{CAMe} NPNNVN ^{-D} VLMDALL.A	Substitution Asn->Asp (D864) / CID	205	82.9
2	O44359	Flagelliform silk protein	Chymotrypsin	878 – 885	412.72	412.67	0.0568	2	L.SNHGSS ^{2.11} SF.A	Unknown shift / CID	419	97.4
2	O44359	Flagelliform silk protein	Chymotrypsin	898 – 905	527.27	527.216	0.0557	2	Y.S ^{156.11} NSVGRMF.A	Unknown shift / CID	204	94.4
2	O44358	Flagelliform silk protein	Glu-C	1-25	1363.2	1363.19	0.006	2	>.MGKGRHDTKAKAKAMQV ^{126.01} ALASSIAE.L	Unknown shift / CID + ETD	259	85.2
2	O44358	Flagelliform silk protein	Glu-C	13 – 25	659.83	659.82	0.012	2	A.KAMQVALASSIAE.L	CID + ETD	228	99.8
2	O44358	Flagelliform silk protein	Glu-C	26 – 30	544.33	544.33	-0.0041	1	E.LVIAE.S	ETD	243	99.9
2	O44358	Flagelliform silk protein	Glu-C	31 – 61	1083.83	1083.86	-0.0337	3	E.SSGGDVQRKTNVISNALRNALM ^{Ox} STTGSPNEE.F	Oxidation (M52) / CID + ETD	317	100
2	O44358	Flagelliform silk protein	Glu-C	31 – 61	1078.8	1078.86	-0.06	3	E.SSGGDVQR ^{Deamid} KTNVISNALRNALMSTTGSPNEE.F	Deamidation (R38) / CID	299	99.9
2	O44358	Flagelliform silk protein	Glu-C	31 – 61	809.35	809.39	-0.0468	4	E.SSGGDVQR ^{Deamid} KTNVISNALRNALMSTTGSPNEE.F	Deamidation (R38) / ETD	452	98.8
2	O44358	Flagelliform silk protein	Glu-C	62 – 80	757.07	757.05	0.021	3	E.FVHEVQDLIQML ^{-30.92} SSEQINE.V	Unknown shift / CID + ETD	299	93.1
2	O44358	Flagelliform silk protein	Glu-C	66 – 80	894.51	894.44	0.0642	2	E.VQDLIQMLSSEQINE.V	CID	525	100
2	O44358	Flagelliform silk protein	Glu-C	77 – 109	1078.52	1078.48	0.0347	3	E.QINEVDTSGPGQYYRSSSSGGGGGGQGG ^{49.1} PVVTE.T	Unknown shift / CID + ETD	205	91.8
2	O44358	Flagelliform silk protein	Glu-C	81 – 109	900.73	900.74	-0.0126	3	E.VDTSGPGQYYRSSSSGGGGGGQGGPVVTE.T	CID + ETD	523	100
2	O44358	Flagelliform silk protein	Glu-C	504 – 520	944.93	944.928	0.002	2	E.DLDITIDGADGPIT ¹⁴⁴ SE.E	Unknown shift / CID	234	87.5
3	B5SYS6	Spidroin-1	Trypsin	49 – 68	1058.46	1058.49	-0.0312	2	R.TGAFTADQLDDMSTIGDTLK.T	CID	465	100
3	B5SYS6	Spidroin-1	Trypsin	49 – 68	706.06	705.99	0.0635	3	R.TGAFTADQLD ^{Ox} DMSTIGDTLK.T	Hydroxylation (D58) / CID	229	96.1
3	B5SYS6	Spidroin-1	Trypsin	49 – 85	1003.8	1003.72	0.0709	4	R.TGAFTADQLDDMSTIGDTLKTA ^M ^{Ox} DKM ^{Ox} ARSNKSSQSKL.Q	Oxidation (M71, M74) / ETD	235	99.9
3	B5SYS6	Spidroin-1	Trypsin	69 – 79	687.29	687.31	-0.0287	2	K.TAMDKM ^{Ox} ARSNK.S	Oxidation (M74) / CID	274	88.8
3	B5SYS6	Spidroin-1	Trypsin	74 – 84	783.88	783.80	0.0709	2	K.M ^{Ox} ARSNKSSQS ^{327.14} K.L	Unknown shift / CID	289	92.3
3	B5SYS6	Spidroin-1	Trypsin	74 – 79	414.23	414.19	0.0349	2	K.M ^{Ox} ARSNK.S	Oxidation (M74) / CID	234	83.6
3	B5SYS6	Spidroin-1	Trypsin	85 – 112	956.8	956.80	-0.0088	3	K.LQALNMAFASSMAEIAAVEQGGLSVAEK.T	CID + ETD	295	99.6
3	B5SYS6	Spidroin-1	Trypsin	113 – 140	1015.18	1015.17	0.0045	3	K.TNAIADSLNSAFYQTTGAVNVQFVNEIR.S	CID + ETD	465	100
3	B5SYS6	Spidroin-1	Trypsin	144 – 173	903.46	903.39	0.0671	3	I.SMFAQASANEVSYGGGYGGGQGGSAGAAA.A	No enzyme (two ends) / CID + ETD	203	82
3	B5SYS6	Spidroin-1	Trypsin	160 – 205	888.95	888.90	0.0425	4	G.YGGGQGGQSAGAAAAAASAGAGQGGYGGLGGQGAGSAAAAAASGAG.Q	No enzyme (two ends) / ETD	352	83.7
3	B5SYS6	Spidroin-1	Trypsin	164 – 182	751.31	751.35	-0.0431	2	G.QGGQSAGAAAAAASAGAGQ.G	No enzyme (two ends) / CID	243	91
3	B5SYS6	Spidroin-1	Trypsin	179 – 190	475.7	475.72	-0.02	2	A.GAGQGGYGGLGG.Q	No enzyme (two ends) / CID	254	88
3	B5SYS6	Spidroin-1	Trypsin	187 – 207	793.84	793.87	-0.0395	2	G.GLGGQGAGSAAAAAASGAGQG.G	No enzyme (two ends) / CID	200	92
3	B5SYS6	Spidroin-1	Trypsin	195 – 223	773.74	773.69	0.0411	3	G.SAAAAAASGAGQGGYGGVGNQAGRGAGA.A	No enzyme (two ends) / CID + ETD	258	86
3	B5SYS6	Spidroin-1	Trypsin	220 – 261	839.22	839.123	0.097	4	R.GAGAAAAAAGGAGQGGYNGGQGPSAAAAA ^{105.39} AAAASGAGQGGYG.-	Unknown shift / ETD	329	89.2
3	B5SYS6	Spidroin-1	Chymotrypsin	1-9	792.28	792.23	0.050	2	>.M ^{Ox} TWTARLA ^{504.98} L.S	Oxidation (M1) / CID	272	85.9
3	B5SYS6	Spidroin-1	Chymotrypsin	10-20	1079.16	1079.12	0.046	1	L.SIL ^{-38.46} AVLCTQGL.F	Unknown shift / CID	257	92.3
3	B5SYS6	Spidroin-1	Chymotrypsin	13 – 21	1023.5	1023.48	0.021	1	L.AVLC ^{CAMe} TQGL ^{14.98} F.A	Unknown shift / CID	206	97.9

3	B5SYS6	Spidroin-1	Chymotrypsin	16 – 29	523.56	523.58	-0.0206	3	L. C ^{CAME} TQGLF ^{-39.05} AQQQNTPW.S	Unknown shift / CID	233	81.6
3	B5SYS6	Spidroin-1	Chymotrypsin	22 – 38	908.89	908.91	-0.0258	2	F.AQQQNTPWSST ^{-6.04} ELADAF.I	Unknown shift / CID + ETD	253	81.5
3	B5SYS6	Spidroin-1	Chymotrypsin	30 – 38	940.42	940.42	-0.0058	1	W.SSTELADAF.I	CID	272	99.7
3	B5SYS6	Spidroin-1	Chymotrypsin	30 – 42	693.37	693.33	0.0372	2	W.SSTELADAFINAF.L	CID	440	100
3	B5SYS6	Spidroin-1	Chymotrypsin	39 – 52	740.88	740.88	-0.0008	2	F.INAFLNEAGRTGAF.T	CID + ETD	272	100
3	B5SYS6	Spidroin-1	Chymotrypsin	43 – 52	518.2	518.26	-0.0645	2	F.LNEAGRTGAF.T	CID	435	100
3	B5SYS6	Spidroin-1	Chymotrypsin	43 – 67	876.74	876.74	-0.0013	3	F.LNEAGRTGAFTADQLDDMSTIGDTL.K	CID + ETD	503	100
3	B5SYS6	Spidroin-1	Chymotrypsin	44 – 52	535.22	535.25	-0.0354	2	L. N ^{HexN} EAGRTGAF.T	Hexosamine (N44) / CID	366	100
3	B5SYS6	Spidroin-1	Chymotrypsin	44 – 67	876.73	876.71	0.0167	3	L.NE ^{113.05} AGRTGAFTADQLDDMSTIGDTL.K	CID + ETD	502	100
3	B5SYS6	Spidroin-1	Chymotrypsin	53 – 67	806.35	806.35	-0.0063	2	F.TADQLDDMSTIGDTL.K	CID	464	100
3	B5SYS6	Spidroin-1	Chymotrypsin	68 – 88	955.44	955.41	0.0278	3	L.KTAMDKMARS ^{541.08} NKSSQSKLQAL.N	Unknown shift / CID + ETD	218	81.4
3	B5SYS6	Spidroin-1	Chymotrypsin	86 – 92	405.68	405.69	-0.0144	2	L.QALN M ^{Ox} AF.A	Oxidation (M90) / CID	368	100
3	B5SYS6	Spidroin-1	Chymotrypsin	93 – 107	725.34	725.34	-0.0081	2	F.ASS M ^{Ox} AEIAAVEQGG.L.S	Oxidation (M96) / CID	472	100
3	B5SYS6	Spidroin-1	Chymotrypsin	93 – 124	1028.07	1028.18	-0.1154	3	F.ASSMAEIAAVEQGGLSVAEKTN ^{-70.34} AIADSLNSAF.Y	Unknown shift / CID + ETD	228	94.9
3	B5SYS6	Spidroin-1	Chymotrypsin	93 – 125	1111.2	1111.20	-0.0048	3	F.ASS M ^{Ox} AEIAAVEQGGLSVAEKTNAIADSLNSAFY.Q	Oxidation (M96) / CID + ETD	411	100
3	B5SYS6	Spidroin-1	Chymotrypsin	108 – 120	716.82	716.79	0.034	2	L.SVAEKTNAIADSL ^{113.95} .N	Unknown shift / CID	230	100
3	B5SYS6	Spidroin-1	Chymotrypsin	108 – 124	869.46	869.43	0.0237	2	L.SVAEKTNAIADSLNSAF.Y	CID	357	100
3	B5SYS6	Spidroin-1	Chymotrypsin	108 – 125	950.99	950.968	0.022	2	L.SVAEKTNAIADSLNSAFY.Q	CID	352	100
3	B5SYS6	Spidroin-1	Chymotrypsin	121 – 125	643.29	643.13	0.1555	2	L.NS ^{684.31} AFY.Q	Unknown shift / CID	259	80.2
3	B5SYS6	Spidroin-1	Chymotrypsin	121 – 135	778.9	778.89	0.011	2	L.NSAFYQTTGAVNV ^{-89.98} QF.V	Unknown shift / CID	257	84.8
3	B5SYS6	Spidroin-1	Chymotrypsin	126 – 135	532.76	532.77	-0.0122	2	Y.QTTGAVNVQF.V	CID	394	100
3	B5SYS6	Spidroin-1	Chymotrypsin	126 – 142	923.52	923.48	0.048	2	Y.QTTGAV ^{-29.95} NVQFVNEIRSL.I	Unknown shift / CID	273	99.7
3	B5SYS6	Spidroin-1	Chymotrypsin	136 – 146	662.85	662.85	-0.0001	2	F.VNEIRSLIS M ^{Ox} F.A	Oxidation (M145) / CID + ETD	405	100
3	B5SYS6	Spidroin-1	Chymotrypsin	147 – 156	520.21	520.23	-0.0282	2	F.AQASANEVSY.G	CID	523	100
3	B5SYS6	Spidroin-1	Chymotrypsin	147 – 160	642.68	642.802	-0.122	2	F.AQASANEVS ^{-89.23} YGGGY.G	Unknown shift / CID + ETD	218	98.2
3	B5SYS6	Spidroin-1	Chymotrypsin	157 – 185	820.39	820.33	0.0527	3	Y.GGGYGG ^{175.16} GQGGQSAGAAAAAASAGAGQGGY.G	Unknown shift / CID + ETD	434	99.8
3	B5SYS6	Spidroin-1	Chymotrypsin	157 – 188	886.71	886.67	0.040	3	Y.G ^{146.99} GGYGGGQGGQSAGAAAAAASAGAGQGGYGGL.G	Unknown shift / CID	309	99.4
3	B5SYS6	Spidroin-1	Chymotrypsin	161 – 185	975.43	975.43	-0.0084	2	Y.GGGQGGQSAGAAAAAASAGAGQGGY.G	CID	548	100
3	B5SYS6	Spidroin-1	Chymotrypsin	161 – 188	796.02	796.00	0.0163	3	Y.G ^{209.05} GGQGGQSAGAAAAAASAGAGQGGYGGL.G	Unknown shift / CID + ETD	381	98.4
3	B5SYS6	Spidroin-1	Chymotrypsin	186 – 209	917.42	917.42	-0.0074	2	Y.GGLGGQGAGSAAAAA S ^{-G} GAGQGGY.G	Substitution Ser->Gly (S202) / CID	538	100
3	B5SYS6	Spidroin-1	Chymotrypsin	186 – 236	974.83	974.96	-0.1361	4	Y.GGLGGQGAGSAAAAAASGAGQG ^{-79.54} GYGGVGNQGAGRGAGAAAAAAGGAGQGGY.N	Unknown shift / ETD	296	92.1
3	B5SYS6	Spidroin-1	Chymotrypsin	189 – 209	860.34	860.26	0.080	2	L.GGQGA ^{82.94} GSAAAAAASGAGQGGY.G	Unknown shift / CID	478	100
3	B5SYS6	Spidroin-1	Chymotrypsin	210 – 236	1065.99	1066.00	-0.0148	2	Y.GGVGNQGAGRGAGAAAAAAGGAGQGGY.N	CID	434	100
3	B5SYS6	Spidroin-1	Chymotrypsin	213 – 236	959.4	959.37	0.030	2	V.GNQGAGRGAGAAAAAAGGAGQGGY.N	CID	386	100
3	B5SYS6	Spidroin-1	Chymotrypsin	214 – 236	930.93	930.90	0.0301	2	G.NQGAGRGAGAAAAAAGGAGQGGY.N	CID	508	100
3	B5SYS6	Spidroin-1	Chymotrypsin	215 – 236	873.89	873.81	0.080	2	N.QGAGRGAGAAAAAAGGAGQGGY.N	CID	452	100
3	B5SYS6	Spidroin-1	Chymotrypsin	237 – 260	927.41	927.44	-0.0334	2	Y.NGGQGPSAAAAAAAAS ^{-79.06} GAGQGGY.G	Unknown shift / CID	289	99.8
3	B5SYS6	Spidroin-1	Chymotrypsin	237 – 261	1080.91	1080.88	0.035	2	Y.N ^{170.91} GGQGPSAAAAAAAASGAGQGGYG.-	Unknown shift / CID	305	100
3	B5SYS6	Spidroin-1	Glu-C	25 – 33	525.22	525.18	0.040	2	G.QNTPWSSTE.L	CID	336	100
3	B5SYS6	Spidroin-1	Glu-C	34 – 45	669.35	669.34	0.0096	2	E.LADAFINAF.LNE.A	CID	415	100
3	B5SYS6	Spidroin-1	Glu-C	83 – 98	865.89	885.86	0.031	2	Q.SKQLALN M ^{Ox} AFASS M ^{Ox} AE.I	Oxidation (M90, M96) / CID	350	99.9
3	B5SYS6	Spidroin-1	Glu-C	99 – 103	502.24	502.28	-0.0471	1	E.IAAVE.Q	CID	236	97.4
3	B5SYS6	Spidroin-1	Glu-C	99 – 111	734.32	734.27	0.051	2	E.I ^{223.98} AAVEQGGLSVAE.K	Unknown shift / CID + ETD	299	98
3	B5SYS6	Spidroin-1	Glu-C	104 – 111	380.69	380.69	-0.0054	2	E.QGGLSVAE.K	CID	286	97.5

3	B5SYS6	Spidroin-1	Glu-C	112 – 138	968.12	968.14	-0.0254	3	E.KTNAIADSLNSAFYQTTGAVNVQFVNE.I	CID + ETD	528	100
3	B5SYS6	Spidroin-1	Glu-C	139 – 153	827.42	827.41	0.0031	2	E.IRSLISM ^{0x} FAQASANE.V	Oxidation (M145) / CID	403	100
3	B5SYS6	Spidroin-1	Glu-C	139 – 153	842.42	842.42	-0.0022	2	E.IRSLISM ^{0x} FA ^{-T} QASANE.V	Substitution Ala->Thr (A147) / CID	211	99.8
3	P19837	Spidroin-1	Trypsin	68 – 100	877.36	877.42	-0.0624	3	A.GRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRG.G	No enzyme (two ends) / CID + ETD	207	98.9
3	P19837	Spidroin-1	Trypsin	70 – 99	811.06	811.04	0.0189	3	R.GGQGAGA ^{71.06} AAAAAGGAGQGGYGGLGSQGAGR.G	Unknown shift / CID + ETD	485	100
3	P19837	Spidroin-1	Trypsin	94 – 124	811.07	811.05	0.0166	3	G.SQGAGRGGLGGQGAGAAAAAAGGAGQGGYG.G	No enzyme (two ends) / CID + ETD	375	99.6
3	P19837	Spidroin-1	Trypsin	100 - 133; 428 - 461	884.85	884.82	0.031	3	R.GGLGGQGAGAAAAAA ^{-32.74} GGAGQGGYGGLGNQGAGR.G	Unknown shift / CID + ETD	236	99
3	P19837	Spidroin-1	Trypsin	128 – 153	696.98	696.99	-0.01	3	G.NQGAGRGGQGAAAAAAGGAGQGGYGG.L	No enzyme (two ends) / ETD	238	96.4
3	P19837	Spidroin-1	Trypsin	134 – 161	730.37	730.32	0.052	3	R.G ^{-42.95} GQGAAAAAGGAGQGGYGGLGSQGAGR.G	Unknown shift / ETD	618	100
3	P19837	Spidroin-1	Trypsin	156 – 198	1140.5	1140.53	-0.0322	3	G.SQGAGRGGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYG.G	No enzyme (two ends) / CID + ETD	356	100
3	P19837	Spidroin-1	Trypsin	162 – 184	873.9	873.91	-0.0113	2	R.GGLGGQGAGAAAAAAGGAGQGGY.G	Chymotryptic cleavage (one end) / CID + ETD	480	100
3	P19837	Spidroin-1	Trypsin	162 – 207	964.64	964.43	0.2073	4	R.G ^{208.83} GLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	Unknown shift / ETD	323	85
3	P19837	Spidroin-1	Trypsin	165 – 207	1140.51	1140.53	-0.0222	3	L.GGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID + ETD	517	100
3	P19837	Spidroin-1	Trypsin	169 - 208; 270 - 309	1059.8	1059.83	-0.0317	3	G.AGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRG.G	No enzyme (two ends) / CID + ETD	324	99.9
3	P19837	Spidroin-1	Trypsin	170 - 208; 271 - 309	1036.07	1036.15	-0.0826	3	A.GAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGRG.G	No enzyme (two ends) / CID + ETD	288	100
3	P19837	Spidroin-1	Trypsin	173 - 207; 274 - 308	950.76	950.78	-0.0203	3	A.AAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / ETD	459	100
3	P19837	Spidroin-1	Trypsin	177 - 207; 278 - 308	856.2	856.06	0.1359	3	A.GGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID + ETD	316	99.9
3	P19837	Spidroin-1	Trypsin	211 – 253	1050.79	1050.82	-0.0367	3	L.GGQGAGAAAAAAGGAGQGGGLGGQGAGQGAGASAAAGGAGQG.G	No enzyme (two ends) / CID + ETD	249	96.4
3	P19837	Spidroin-1	Trypsin	266 – 308	856.2	856.143	0.057	4	R.GGEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR ^{Deamid} .G	Deamidation (R308) / ETD	433	99.2
3	P19837	Spidroin-1	Trypsin	320 – 363	1138.08	1138.20	-0.1277	3	A.AAGGAGQGGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID + ETD	454	100
3	P19837	Spidroin-1	Trypsin	333 – 356	909.4	909.42	-0.0299	2	Q.GAGQGAGAAAAAAGGAGQGGYGGL.G	No enzyme (two ends) / CID	351	100
3	P19837	Spidroin-1	Trypsin	357 – 388	839.38	839.40	-0.0244	3	L.GSQGAGRGGLGGQGAGAVAAAAAGGAGQGGYG.G	No enzyme (two ends) / CID + ETD	288	89.7
3	P19837	Spidroin-1	Trypsin	364 – 397	958.46	958.43	0.0205	3	R.GGLGGQGAGAVA ^{187.06} AAAAGGAGQGGYGGLGSQGAGR.G	Unknown shift / CID + ETD	406	100
3	P19837	Spidroin-1	Trypsin	398 – 415	735.34	735.32	0.021	2	R.GGQ ^{70.99} GAGAAAAAAGGAGQR.G	Unknown shift / CID	299	99.4
3	P19837	Spidroin-1	Trypsin	398 – 427	1036.48	1036.40	0.0754	3	R.GGQGAGAAAAAAGGAG ^{621.23} QRGYGGLGNQGAGR.G	Unknown shift / CID + ETD	358	100
3	P19837	Spidroin-1	Trypsin	416 – 427	1098.52	1098.53	-0.0137	1	R.GY ⁻⁸ GGLGNQGAGR.G	Unknown shift / CID	288	99.9
3	P19837	Spidroin-1	Trypsin	462 – 488	1265.64	1265.51	0.1299	2	R.G ^{369.26} GQGAAAAAGGAGQGGYGGLGSQGAGR.G	Unknown shift / CID	350	100
3	P19837	Spidroin-1	Trypsin	495 – 518	1044.49	1044.51	-0.0245	2	G.AAAAAAVGAGQEGIRQQGAGQGGY.G	No enzyme (one end) / CID	271	95.4
3	P19837	Spidroin-1	Trypsin	529 – 571	1091.86	1091.84	0.021	3	R.GGLGGQGAGAAAAAA ^{40.97} GGAGQGGGLGGQGAGQGAGAAAAAAGGV.R	Unknown shift / CID + ETD	267	86.4
3	P19837	Spidroin-1	Trypsin	572 – 585	598.95	598.87	0.0775	3	R.QGGYGGLGSQGA ^{530.23} GR.G	Unknown shift / CID + ETD	298	92.8
3	P19837	Spidroin-1	Trypsin	572 – 615	1184.64	1184.60	0.041	3	R.QGGYGGLGSQGAGRGQGAGAAAAAAG ^{-51.8} GAGQGGYGGLGGQGVGR.G	Unknown shift / CID + ETD	322	80.1
3	P19837	Spidroin-1	Trypsin	586 – 615	826.73	826.71	0.0154	3	R.GGQGAGAAAAAAGGAGQGG ^{120.05} YGGLGGQGVGR.G	Unknown shift / CID + ETD	236	92.7
3	P19837	Spidroin-1	Trypsin	610 – 636	720.35	720.34	0.010	3	G.GQGVGRGGLGGQGAGAAAAGGAGQGGY.G	No enzyme (one end) / CID	360	100
3	P19837	Spidroin-1	Trypsin	616 – 652	962.14	962.12	0.0194	3	R.GGLGGQGAGAAAAGGAGQGG ^{22.06} YGGVGSASAASAAASR.L	Unknown shift / CID + ETD	445	100
3	P19837	Spidroin-1	Trypsin	619 – 654	945.78	945.78	-0.0037	3	L.GGQGAGAAAAGGAGQGGYGGVGSASAASAAASRLS.S	No enzyme (two ends) / CID	297	89
3	P19837	Spidroin-1	Chymotrypsin	1-18	725.86	725.81	0.0411	2	>.QGAGAAAAAAGGAGQGGY ^{Nitro} .G	Nitro (Y18) / CID	258	98.5
3	P19837	Spidroin-1	Chymotrypsin	1-31	819.79	819.75	0.040	3	>.QGAGAAAAAAGGAGQG ^{-7.75} GYGGLGGQGAGQGGY.G	Unknown shift / CID + ETD	291	99.4
3	P19837	Spidroin-1	Chymotrypsin	19 - 31; 185 - 197; 286 - 298	539.75	539.74	0.0008	2	Y.GGLGGQGAGQGGY.G	CID	409	100
3	P19837	Spidroin-1	Chymotrypsin	19 - 34; 185 - 200; 286 - 301	787.4	787.31	0.0873	2	Y.GGLGGQGAGQGG ^{268.17} YGGL.G	Unknown shift / CID	227	83.2
3	P19837	Spidroin-1	Chymotrypsin	22 - 31; 188 - 197; 289 - 298	820.14	820.36	-0.2242	1	L.GGQ ^{-31.21} GAGQGGY.G	Unknown shift / CID	253	99.9
3	P19837	Spidroin-1	Chymotrypsin	22 – 59	1013.96	1013.78	0.176	3	L.GGQGAGQGGYGGLGGQGAGQ ^{76.53} GAGAAAAAAGGAGQGGY.G	Unknown shift / CID	263	95.9
3	P19837	Spidroin-1	Chymotrypsin	32 – 59	687.31	687.33	-0.0252	3	Y.GGLGGQGAGQ ^{-71.07} GAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	445	100
3	P19837	Spidroin-1	Chymotrypsin	32 – 62	635.68	635.535	0.145	4	Y.GGLGGQGAGQGAGAAAAAA ^{181.58} GGAGQGGYGGL.G	Unknown shift / ETD	326	97.1

3	P19837	Spidroin-1	Chymotrypsin	35 – 59	916.92	916.93	-0.0157	2	L.GGQGAGQGAGA ^{-71.02} AAAAAAGGAGQGGY.G	Unknown shift / CID	537	100
3	P19837	Spidroin-1	Chymotrypsin	35 – 62	907.42	907.33	0.0848	3	L.GGQGAGQGAGAAAAAAGGAGQGGYGG ^{589.25} .L.G	Unknown shift / CID + ETD	327	96
3	P19837	Spidroin-1	Chymotrypsin	60 - 92; 576 - 608	1036.11	1036.08	0.0266	3	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGG ^{519.08} .L.G	Unknown shift / CID	475	100
3	P19837	Spidroin-1	Chymotrypsin	63 – 102	1155.43	1155.40	0.032	3	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGR ^{263.75} GGL.G	Unknown shift / CID + ETD	203	93.3
3	P19837	Spidroin-1	Chymotrypsin	63 - 89; 579 - 605	1030.47	1030.49	-0.0244	2	L.G ^{-73.04} SQGAGRGGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	492	100
3	P19837	Spidroin-1	Chymotrypsin	63 - 92; 579 – 608	960.45	960.37	0.0756	3	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGL ^{519.23} .G	Unknown shift / CID + ETD	438	100
3	P19837	Spidroin-1	Chymotrypsin	90 - 102; 152 - 164; 198 - 210; 299 - 311; 354 - 366	568.67	568.65	0.021	2	Y.GGLGSQGAGRGG ^{49.77} .G	Unknown shift / CID	324	100
3	P19837	Spidroin-1	Chymotrypsin	93 – 123	793.38	793.38	-0.0068	3	L.GSQGAGRGGGLGGQGAGAAAAAA ^{-53.01} GGAGQGGY.G	Unknown shift / CID + ETD	329	98.9
3	P19837	Spidroin-1	Chymotrypsin	93 – 126	655.8	655.82	-0.0236	4	L.G ^{-38.09} SQGAGRGGGLGGQGAGAAAAAAGGAGQGGYGGL.G	Unknown shift / ETD	284	88.7
3	P19837	Spidroin-1	Chymotrypsin	103 - 123; 431 - 451	760.35	760.36	-0.0164	2	L.GGQGAGAAAAAAG ^{-71.02} GAGQGGY.G	Unknown shift / CID	336	100
3	P19837	Spidroin-1	Chymotrypsin	103 - 126; 431 - 454	863.34	863.42	-0.0899	2	L.GGQGAGAAAAAAGGAGQGGY ^{-92.17} GGL.G	Unknown shift / CID + ETD	339	100
3	P19837	Spidroin-1	Chymotrypsin	127 – 151	1049.78	1049.74	0.043	2	L.GNQGAGRGGQGAAAA ^{66.62} AAGGAGQGGY.G	Unknown shift / CID	234	96.7
3	P19837	Spidroin-1	Chymotrypsin	127 – 154	586.32	586.27	0.0493	4	L.GNQGAGRGGQGAAAAAAG ^{83.2} GAGQGGYGGL.G	Unknown shift / ETD	377	97
3	P19837	Spidroin-1	Chymotrypsin	152 – 184	835.06	835.08	-0.0234	3	Y.GGLGSQGAGRGG ^{-84.06} GGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	294	99.3
3	P19837	Spidroin-1	Chymotrypsin	155 – 184	853.11	853.04	0.0689	3	L.G ^{197.21} SQGAGRGGGLGGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	512	100
3	P19837	Spidroin-1	Chymotrypsin	155 – 187	1041.76	1041.75	0.0099	3	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQGGYGG ^{536.03} .L.G	Unknown shift / CID + ETD	264	99.8
3	P19837	Spidroin-1	Chymotrypsin	155 – 197	1140.51	1140.53	-0.0222	3	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	CID + ETD	495	100
3	P19837	Spidroin-1	Chymotrypsin	165 – 184	873.92	873.84	0.0722	2	L.G ^{227.14} GQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	484	100
3	P19837	Spidroin-1	Chymotrypsin	165 – 197	873.51	873.39	0.1179	3	L.GGQGAGA ^{39.35} AAAAAGGAGQGGYGGLGGQGAGQGGY.G	Unknown shift / CID + ETD	333	99.9
3	P19837	Spidroin-1	Chymotrypsin	198 – 231	858.34	858.436	-0.096	3	Y.GGLGSQGAGRGG ^{-35.28} GGQGAGAAAAAAGGAGQGGL.G	Unknown shift / CID + ETD	203	92.3
3	P19837	Spidroin-1	Chymotrypsin	201 – 231	1328.14	1328.08	0.0531	2	L.GSQGAGR ^{274.11} GGLGGQGAGAAAAAAGGAGQGGL.G	Unknown shift / CID	227	86.5
3	P19837	Spidroin-1	Chymotrypsin	211 – 255	863.34	863.38	-0.0446	4	L.GGQGAGAAAAAAGGAGQGGLGGQGAGQGAGAS ^{HPO3} AAAAGGAGQGGY.G	Phosphorylation (S243) / CID + ETD	349	94.6
3	P19837	Spidroin-1	Chymotrypsin	232 – 255	1013.96	1013.91	0.0454	2	L.GGQGAGQGAGASA ^{178.09} AAAGGAGQGGY.G	Unknown shift / CID	215	100
3	P19837	Spidroin-1	Chymotrypsin	256 – 285	811.06	811.03	0.0243	3	Y.GGLGSQGAGRGGEG ^{70.07} AGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	466	100
3	P19837	Spidroin-1	Chymotrypsin	256 – 288	837.74	837.74	-0.0047	3	Y.GGLGSQGAGRGGEGAGAAAAAAGGAGQGG ⁻⁷⁷ YGGL.G	Unknown shift / CID + ETD	243	86.9
3	P19837	Spidroin-1	Chymotrypsin	259 – 285	787.39	787.32	0.0633	3	L.G ^{226.19} SQGAGRGGEGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	383	100
3	P19837	Spidroin-1	Chymotrypsin	259 – 298	1064.81	1064.83	-0.0253	3	L.GSQGAGRGGE ^{-K-} GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Substitution Glu->Lys (E268) / CID	500	100
3	P19837	Spidroin-1	Chymotrypsin	299 – 329	1152.57	1152.59	-0.0247	2	Y.GGLGSQGAGRGGGLGGQGAGAAAAGGAG ^{-91.04} QGGL.G	Unknown shift / CID	211	89.2
3	P19837	Spidroin-1	Chymotrypsin	302 – 353	971.47	971.47	-0.0015	4	L.GSQGAGRGGGLGGQGAGAAAAGGAGQGGLGGQGAGQG ⁹⁹ AGAAAAAAGGAGQGGY.G	Unknown shift / ETD	274	80.6
3	P19837	Spidroin-1	Chymotrypsin	312 – 353	844.69	844.61	0.0735	4	L.GGQGAGAAAAGGA ^{234.29} GQGGLGGQGAGQGAGAAAAAAGGAGQGGY.G	Unknown shift / ETD	299	84.1
3	P19837	Spidroin-1	Chymotrypsin	312 – 356	1146.27	1146.19	0.0748	3	L.GGQGAGAAAAGGAGQGGLGGQGAGQGAGAAAAAAGG ^{68.22} AGQGGYGGL.G	Unknown shift / CID + ETD	216	80.7
3	P19837	Spidroin-1	Chymotrypsin	330 – 353	1044.48	1044.41	0.0629	2	L.GGQGAG ^{255.13} QGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	419	100
3	P19837	Spidroin-1	Chymotrypsin	330 – 356	657.31	657.32	-0.0129	3	L.GGQGAGQGAGAAAAAA ^{-90.03} GGAGQGGYGGL.G	Unknown shift / CID + ETD	216	89.8
3	P19837	Spidroin-1	Chymotrypsin	330 – 366	723.34	723.349	-0.009	4	L.GGQGAGQGAGAAA ^{-10.03} AAAGGAGQGGYGGLGSQGAGRGG.L.G	Unknown shift / ETD	326	97.9
3	P19837	Spidroin-1	Chymotrypsin	354 – 387	886.41	886.43	-0.0295	3	Y.GGLGSQGAG ^{-29.08} RGGLGGQGAGAVAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	286	93.7
3	P19837	Spidroin-1	Chymotrypsin	357 – 390	876.76	876.77	-0.0129	3	L.G ^{-58.03} SQGAGRGGGLGGQGAGAVAAAAAGGAGQGGYGGL.G	Unknown shift / CID + ETD	213	81
3	P19837	Spidroin-1	Chymotrypsin	367 – 387	860.44	860.382	0.058	2	L.GGQGAGAVAAA ^{101.12} AAGGAGQGGY.G	Unknown shift / CID	223	97.6
3	P19837	Spidroin-1	Chymotrypsin	367 – 390	669.99	669.96	0.0239	3	L.GGQGAG ^{162.07} AVAAAAAGGAGQGGYGGL.G	Unknown shift / CID	248	85.6
3	P19837	Spidroin-1	Chymotrypsin	388 – 420	959.87	959.77	0.0934	3	Y.GGLGSQGAGRGG ^{191.28} QGAGAAAAAAGGAGQRGYGGL.G	Unknown shift / CID + ETD	204	94.4
3	P19837	Spidroin-1	Chymotrypsin	391 – 430	1299.58	1299.54	0.0354	3	L.G ^{570.11} SQGAGRGGQGAGAAAAAAGGAGQRGYGGLGNQGAGRGG.L.G	Unknown shift / CID + ETD	274	84.1
3	P19837	Spidroin-1	Chymotrypsin	418 – 451	1129.57	1129.43	0.1373	3	Y.G ^{701.41} GLGNQGAGRGGGLGGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	224	89.6
3	P19837	Spidroin-1	Chymotrypsin	421 – 430	518.23	518.22	0.0019	2	L.GNQGAGR ¹⁴⁹ GL.G	Unknown shift / CID	294	82.3
3	P19837	Spidroin-1	Chymotrypsin	452 – 478	720.4	720.37	0.032	3	Y.G ^{-28.83} GLGNQGAGRGGQGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	378	100

3	P19837	Spidroin-1	Chymotrypsin	479 – 518	892.27	892.16	0.1087	4	Y.GGLGSQGAGRGGQGAGAAAAAAVGAGQEGIRGQGAGQ ^{210,43} GGY.G	Unknown shift / ETD	355	90.1
3	P19837	Spidroin-1	Chymotrypsin	482 – 518	1102.09	1102.05	0.043	3	L.GSQGAGRGGQGAGAAAAAAVGAGQEG ^{175,76} IRGQGAGQGGY.G	Unknown shift / CID	355	99.9
3	P19837	Spidroin-1	Chymotrypsin	519 – 551	908.76	908.75	0.0047	3	Y.GGLGSQGSGRG ^{171,01} GLGGQGAGAAAAAAGGAGQGGGL.G	Unknown shift / CID + ETD	243	89.5
3	P19837	Spidroin-1	Chymotrypsin	532 – 575	1205.09	1204.88	0.2087	3	L.GGQGAGAAAAAAGGAGQGGGLGGQGA ^{202,63} GQGAGAAAAAAGGVRQGGY.G	Unknown shift / CID + ETD	299	84.9
3	P19837	Spidroin-1	Chymotrypsin	552 – 578	1374.04	1374.03	0.0039	2	L.GGQGAG ^{560,01} QGAGAAAAAAGGVRQGGYGGL.G	Unknown shift / CID	255	80.8
3	P19837	Spidroin-1	Chymotrypsin	609 – 636	825.13	825.02	0.1068	3	L.G ^{257,32} GQGVGRGGLGGQGAGAAAAGGAGQGGY.G	Unknown shift / CID + ETD	514	100
3	P19837	Spidroin-1	Chymotrypsin	619 – 636	639.27	639.31	-0.0407	2	L.GGQGAG ^{-100,07} AAAAGGAGQGGY.G	Unknown shift / CID	282	100
3	P19837	Spidroin-1	Chymotrypsin	619 – 653	956.16	956.10	0.0537	3	L.GGQGAGAAAAGGAG ^{118,16} QGGYGGVGSASAASAAASRL.S	Unknown shift / CID + ETD	232	83.9
3	P19837	Spidroin-1	Chymotrypsin	637 – 653	695.36	695.35	0.0023	2	Y.GGVGSASAASAAASRL.S	CID	249	100
3	P19837	Spidroin-1	Chymotrypsin	654 – 669	789.42	789.39	0.0281	2	L.SSPQASSR ^{Deamid} VSSAVSNL.V	Deamidation (R261) / CID	359	99.9
3	P19837	Spidroin-1	Chymotrypsin	654 – 680	854.78	854.76	0.0134	3	L.SSPQASSRVSSAVSNLVA ^{17,04} SGPTNSAAL.S	Unknown shift / CID + ETD	394	99.4
3	P19837	Spidroin-1	Chymotrypsin	681 – 708	971.47	971.43	0.043	3	L.SSTISNVVSQIGASNPGLSGCDVL ¹⁸² IQAL.L	Unknown shift / CID + ETD	269	85.7
3	P19837	Spidroin-1	Chymotrypsin	699 – 708	502.27	502.27	-0.0062	2	L.SG ^{CAME} DVLI ⁻⁷² QAL.L	Unknown shift / CID	246	87
3	P19837	Spidroin-1	Chymotrypsin	709 – 719	841.47	841.37	0.0937	2	L.LEVVSALIQ ^{484,19} IL.G	Unknown shift / CID	216	87.2
3	P19837	Spidroin-1	Chymotrypsin	710 – 715	440.26	440.17	0.0811	2	L.EVVSA ^{262,16} L.I	Unknown shift / CID + ETD	233	95.5
3	P19837	Spidroin-1	Chymotrypsin	710 – 729	1018.46	1018.56	-0.1024	2	L.EVVSA ^{-41,2} LIQILGSSSIGQVNY.G	Unknown shift / CID	205	93.8
3	P19837	Spidroin-1	Chymotrypsin	730 – 744	740.38	740.37	0.0045	2	Y.GSAGQATQIVGQS ^{Me} VY.Q	Methylation (S742) / CID	457	100
4	B5SYS6	Spidroin-1	Trypsin	1-6	802.42	802.37	0.0488	1	>.M ^{37,05} TWTAR.L	Unknown shift / CID	298	85
4	B5SYS6	Spidroin-1	Trypsin	25 – 40	897.42	897.42	-0.0007	2	G.QNTPWSSTELADAFIN.A	No enzyme (two ends) / CID	405	100
4	B5SYS6	Spidroin-1	Trypsin	49 – 57	462.23	462.22	0.0029	2	R.TGAFTADQL.D	No enzyme (one end) / CID	313	100
4	B5SYS6	Spidroin-1	Trypsin	49 – 68	1066.43	1066.48	-0.0586	2	R.TGAFTADQLDDM ^{+O2} STIGDTLK.T	Sulphone (M60) / CID	404	100
4	B5SYS6	Spidroin-1	Trypsin	69 – 79	883.9	883.81	0.0895	2	K.TAMDKMARSN ^{514,18} K.S	Unknown shift / CID + ETD	263	83.3
4	B5SYS6	Spidroin-1	Trypsin	74 – 84	783.88	783.80	0.0709	2	K.M ^{Ox} ARSNKSSQS ^{327,14} K.L	Oxidation (M74) / CID	243	84.2
4	B5SYS6	Spidroin-1	Trypsin	85 – 112	956.82	956.80	0.0112	3	K.LQALNM ^{Ox} AFASSM ^{Ox} AEIAAVEQGGLSVAEK.T	Oxidation (M90, M96) / CID + ETD	506	100
4	B5SYS6	Spidroin-1	Trypsin	101 – 110	465.74	465.74	-0.0082	2	A.AVEQGGLSVA.E	No enzyme (two ends) / CID	251	99.2
4	B5SYS6	Spidroin-1	Trypsin	113 – 120	402.71	402.70	0.0015	2	K.TNAIADSL.N	No enzyme (one end) / CID	375	100
4	B5SYS6	Spidroin-1	Trypsin	113 – 128	858.86	858.90	-0.0474	2	K.TNAIADSLNSAFYQTT.G	No enzyme (one end) / CID + ETD	331	100
4	B5SYS6	Spidroin-1	Trypsin	113 – 134	1143.09	1143.05	0.0341	2	K.TNAIADSLNSAFYQTTGAVNVQ.F	No enzyme (one end) / CID	356	100
4	B5SYS6	Spidroin-1	Trypsin	113 – 140	1015.11	1015.17	-0.0655	3	K.TNAIADSLNSAFYQTTGAVNVQFVNEIR.S	CID + ETD	499	100
4	B5SYS6	Spidroin-1	Trypsin	121 – 128	466.2	466.21	-0.0114	2	L.NSAFYQTT.G	No enzyme (two ends) / CID	236	99.8
4	B5SYS6	Spidroin-1	Trypsin	129 – 140	673.38	673.36	0.0152	2	T.GAVNVQFVNEIR.S	No enzyme (one end) / CID + ETD	374	100
4	B5SYS6	Spidroin-1	Trypsin	141 – 150	1070.42	1070.51	-0.0987	1	R.SLISM ^{Ox} FAQAS.A	Oxidation (M145) / CID	204	100
4	B5SYS6	Spidroin-1	Trypsin	151 – 172	957.87	957.91	-0.0443	2	S.ANEVSYGGGYGGGQGGQSAGAA.A	No enzyme (two ends) / CID	467	100
4	B5SYS6	Spidroin-1	Trypsin	187 – 207	793.88	793.87	0.0005	2	G.GLGGQGAGSAAAAAASGAGQG.G	No enzyme (two ends) / CID	225	81.4
4	B5SYS6	Spidroin-1	Trypsin	204 – 221	767.33	767.36	-0.0313	2	G.AGQGGYGGVGNQGAGRGA.G	No enzyme (two ends) / CID + ETD	280	88.4
4	B5SYS6	Spidroin-1	Chymotrypsin	1-7	896.59	896.45	0.1347	1	>.MTWTA ^{18,13} RL.A	Unknown shift / CID	223	82.2
4	B5SYS6	Spidroin-1	Chymotrypsin	16 – 29	779.87	779.85	0.0135	2	L.CTQGLF ^{8,03} AQGQNTPW.S	Unknown shift / CID	243	82.3
4	B5SYS6	Spidroin-1	Chymotrypsin	30 – 38	470.72	470.71	0.0035	2	W.SSTELADAF.I	CID + ETD	493	99.6
4	B5SYS6	Spidroin-1	Chymotrypsin	30 – 42	693.38	693.33	0.0472	2	W.SSTELADAFINAF.L	CID	438	100
4	B5SYS6	Spidroin-1	Chymotrypsin	39 – 52	740.9	740.88	0.0192	2	F.INAFLNEAGRTGAF.T	CID	368	100
4	B5SYS6	Spidroin-1	Chymotrypsin	43 – 52	518.26	518.26	-0.0045	2	F.LNEAGRTGAF.T	CID	382	100
4	B5SYS6	Spidroin-1	Chymotrypsin	43 – 67	876.75	876.74	0.0087	3	F.LNEAGRTGAFTADQLDDMSTIGDTL.K	CID + ETD	521	100
4	B5SYS6	Spidroin-1	Chymotrypsin	53 – 67	912.93	912.85	0.0711	2	F.TADQLDDMSTIGDTL ^{229,14} .K	Unknown shift / CID	398	100

4	B5SYS6	Spidroin-1	Chymotrypsin	86 – 107	747.03	747.02	0.0007	3	L.QALNMAFASSMAEIAAVE ^{Me} QGGL.S	Methylation (E103) / CID + ETD	392	86.8
4	B5SYS6	Spidroin-1	Chymotrypsin	89 – 120	1088.47	1088.45	0.023	3	L.NMAFASSM ^{Ox} AEIAAVEQGGLSVAEKTNAIADSL ^{50.85} .N	Oxidation (M96) / CID	236	89.6
4	B5SYS6	Spidroin-1	Chymotrypsin	93 – 107	725.42	725.34	0.0719	2	F.ASSM ^{Ox} AEIAAVEQGGL.S	Oxidation (M96) / CID	437	99.9
4	B5SYS6	Spidroin-1	Chymotrypsin	108 – 124	961.94	961.93	0.0037	2	L.SVAEKTNAIADSLNS ^{185.01} AF.Y	Unknown shift / CID	281	99.4
4	B5SYS6	Spidroin-1	Chymotrypsin	108 – 125	950.98	950.968	0.012	2	L.SVAEKTNAIADSLNSAFY.Q	CID	376	100
4	B5SYS6	Spidroin-1	Chymotrypsin	121 – 135	614.55	614.53	0.020	3	L.NSAFYQT ^{HPO3} TGAVNVQF.V	Phosphorylation (T127) / CID + ETD	236	97.6
4	B5SYS6	Spidroin-1	Chymotrypsin	126 – 135	532.7	532.77	-0.0722	2	Y.QTTGAVNVQF.V	CID	425	99.9
4	B5SYS6	Spidroin-1	Chymotrypsin	126 – 142	938.5	938.49	0.0002	2	Y.QTTGAVNVQFVNEIRSL.I	CID	247	100
4	B5SYS6	Spidroin-1	Chymotrypsin	127 – 135	614.28	614.25	0.030	2	Q. ^{T^{HPO3}} TGAVNVQF.V	Phosphorylation (T127) / CID	245	95.4
4	B5SYS6	Spidroin-1	Chymotrypsin	136 – 146	662.85	662.85	-0.0001	2	F.VNEIRSLISM ^{Ox} F.A	Oxidation (M145) / CID	330	100
4	B5SYS6	Spidroin-1	Chymotrypsin	147 – 156	1039.43	1039.46	-0.0391	1	F.AQASANEVSY.G	CID	258	98.4
4	B5SYS6	Spidroin-1	Chymotrypsin	147 – 160	687.29	687.302	-0.012	2	F.AQASANEVSYGGGY.G	CID	501	100
4	B5SYS6	Spidroin-1	Chymotrypsin	147 – 185	873.7	873.61	0.0824	4	F.AQASANEVSYGGGYGG ^{187.33} GQGGQSAGAAAAAASAGAGQGGY.G	Unknown shift / ETD	406	96.6
4	B5SYS6	Spidroin-1	Chymotrypsin	157 – 185	739.97	740.004	-0.034	3	Y.GGGYGGGQGGQSAGAAAAAASA ^{-66.09} GAGQGGY.G	Unknown shift / CID	364	99.1
4	B5SYS6	Spidroin-1	Chymotrypsin	157 – 188	886.74	886.713	0.027	3	Y.G ^{147.08} GGYGGGQGGQSAGAAAAAASAGAGQGGYGGL.G	Unknown shift / CID + ETD	411	100
4	B5SYS6	Spidroin-1	Chymotrypsin	161 – 185	975.39	975.43	-0.0484	2	Y.GGGQGGQSAGAAAAAASAGAGQGGY.G	CID	522	100
4	B5SYS6	Spidroin-1	Chymotrypsin	161 – 188	863.11	863.00	0.1063	3	Y.GGGQGGQSAGAAAAA ^{410.32} SAGAGQGGYGGL.G	Unknown shift / CID	380	98.9
4	B5SYS6	Spidroin-1	Chymotrypsin	186 – 209	902.41	902.43	-0.0226	2	Y.GGLGGQGAGSAAA ^{-60.04} AAASGAGQGGY.G	Unknown shift / CID	235	100
4	B5SYS6	Spidroin-1	Chymotrypsin	186 – 236	996.65	996.46	0.1839	4	Y.GGLGGQGAGSAAAAAASGAGQGGYGGVGNQGAG ^{7.73} RGAGAAAAAAGGAGQGGY.N	Unknown shift / ETD	358	93.6
4	B5SYS6	Spidroin-1	Chymotrypsin	189 – 209	803.8	803.86	-0.0639	2	L.GGQGAGSAAAAAS ^{-G⁻} GAGQGGY.G	Substitution Ser->Gly (S202) / CID	348	100
4	B5SYS6	Spidroin-1	Chymotrypsin	189 – 236	957.43	957.18	0.2457	4	L.GGQGAGSAAAAASGAGQGGYGGVGNQ ^{77.98} GAGRGAGAAAAAAGGAGQGGY.N	Unknown shift / ETD	461	99.9
4	B5SYS6	Spidroin-1	Chymotrypsin	210 – 236	1030.5	1030.50	-0.0048	2	Y.G ⁻⁷¹ GVGNQGAGRGAGAAAAAAGGAGQGGY.N	Unknown shift / CID	450	100
4	B5SYS6	Spidroin-1	Chymotrypsin	237 – 260	877.88	877.72	0.1571	3	Y. ^{N^{HexNAc2DHex2}} GGQGPSAAAAAASGAGQGGY.G	HexNAc2DHex2 (N237) / CID	233	94
4	B5SYS6	Spidroin-1	Glu-C	34 – 45	686.36	686.33	0.0274	2	E.LADAFINAF ^{L^{-F}} NE.A	Substitution Leu->Phe (L43) / CID	373	100
4	B5SYS6	Spidroin-1	Glu-C	99 – 111	919.97	919.83	0.1399	2	E.I ^{595.28} AAVEQGGLSVAE.K	Unknown shift / CID	293	80.6
4	B5SYS6	Spidroin-1	Glu-C	104 – 111	760.38	760.38	-0.0036	1	E.QGGLSVAE.K	CID	185	99.3
4	B5SYS6	Spidroin-1	Glu-C	112 – 138	968.15	968.14	0.0046	3	E.KTNAIADSLNSAFYQTTGAVNVQFVNE.I	CID + ETD	563	100
4	B5SYS6	Spidroin-1	Glu-C	139 – 153	827.43	827.41	0.0131	2	E.IRSLISM ^{Ox} FAQASANE.V	Oxidation (M145) / CID	408	100
4	P19837	Spidroin-1	Trypsin	3-38	936.1	936.10	-0.0011	3	G.AGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGGQG.A	No enzyme (two ends) / CID	230	97.5
4	P19837	Spidroin-1	Trypsin	22 - 34; 188 - 200; 289 - 301	539.73	539.74	-0.0192	2	L.GGQGAGQGGYGGL.G	No enzyme (two ends) / CID	242	99.7
4	P19837	Spidroin-1	Trypsin	31 – 69	1045.49	1045.49	-0.0064	3	G.YGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID + ETD	254	95.6
4	P19837	Spidroin-1	Trypsin	49 - 69; 79 - 99; 141 - 161; 245 - 265; 343 - 363; 377 - 397; 468 – 488	860.36	860.41	-0.0515	2	A.AAAGGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID	382	100
4	P19837	Spidroin-1	Trypsin	52 - 69; 82 - 99; 144 - 161; 248 - 265; 346 - 363; 380 - 397; 471 – 488	753.8	753.85	-0.0558	2	A.GGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID	224	99.4
4	P19837	Spidroin-1	Trypsin	56 - 69; 86 - 99; 148 - 161; 194 - 207; 252 - 265; 295 - 308; 350 - 363; 384 - 397; 475 - 488; 572 - 585	632.79	632.80	-0.0151	2	G.QGGYGGLGSQGAGR.G	No enzyme (one end) / CID	368	100
4	P19837	Spidroin-1	Trypsin	59 - 69; 89 - 99; 151 - 161; 197 - 207; 255 - 265; 298 - 308; 353 - 363; 387 - 397; 478 - 488; 575 - 585	511.72	511.75	-0.0343	2	G.YGGLGSQGAGR.G	No enzyme (one end) / CID	466	100
4	P19837	Spidroin-1	Trypsin	70 – 99	844.08	844.04	0.0389	3	R.G ^{170.12} GQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	Unknown shift / CID + ETD	521	100
4	P19837	Spidroin-1	Trypsin	100 - 133; 428 - 461	1098.86	1098.76	0.0939	3	R.G ^{609.28} GLGGQGAGAAAAAAGGAGQGGYGGLGNQGAGR.G	Unknown shift / CID + ETD	238	94.5
4	P19837	Spidroin-1	Trypsin	134 – 161	771.3	771.34	-0.0436	3	R.GGQGAAAAAAGGAGQGGY ^{HPO3} GGLGSQGAGR.G	Phosphorylation (Y151) / CID + ETD	244	95.7

4	P19837	Spidroin-1	Trypsin	162 – 207	907.38	907.43	-0.0527	4	R.GGLGGQGAGAAAAA ^{-20.2} AGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	Unknown shift / ETD	388	99.2
4	P19837	Spidroin-1	Trypsin	165 – 207	1140.48	1140.53	-0.0522	3	L.GGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID	532	100
4	P19837	Spidroin-1	Trypsin	168 - 207; 269 - 308	1059.82	1059.83	-0.0117	3	Q.GAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID	315	98.5
4	P19837	Spidroin-1	Trypsin	170 - 208; 271 - 309	1036.11	1036.15	-0.0426	3	A.GAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (two ends) / CID + ETD	392	100
4	P19837	Spidroin-1	Trypsin	173 - 207; 274 - 308	950.79	950.78	0.0097	3	A.AAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID + ETD	522	100
4	P19837	Spidroin-1	Trypsin	211 – 249	946.59	946.44	0.1428	3	L.GGQGAGAAAAAAGGAGQGGGLGGQGAGQGAGASAAAAGG.A	No enzyme (two ends) / CID	219	95.3
4	P19837	Spidroin-1	Trypsin	266 – 308	870.17	870.147	0.023	4	R.GGEGAGAA ^{57.09} AAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	Unknown shift / ETD	483	100
4	P19837	Spidroin-1	Trypsin	325 – 364	1048.34	1048.16	0.1766	3	A.GQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	No enzyme (two ends) / CID + ETD	219	91.3
4	P19837	Spidroin-1	Trypsin	326 – 363	1010.19	1010.14	0.0409	3	G.QGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID + ETD	448	100
4	P19837	Spidroin-1	Trypsin	364 – 397	958.46	958.43	0.0205	3	R.GGLGGQGAGAVAAAAA ^{187.06} GGAGQGGYGGLGSQGAGR.G	Unknown shift / CID + ETD	465	100
4	P19837	Spidroin-1	Trypsin	398 – 416	728.31	728.356	-0.046	2	R.GGQGAGAAAAAAGGAGQRG.Y	No enzyme (one end) / CID	373	100
4	P19837	Spidroin-1	Trypsin	416 – 427	1014.44	1014.53	-0.0937	1	R.G ^{92.08} YGGLGNQGAGR.G	Unknown shift / CID	203	85.5
4	P19837	Spidroin-1	Trypsin	462 – 488	1094.96	1095.02	-0.0657	2	R.GGQGA AAAAGGAGQGGYGGLGSQGAGR.G	CID	542	100
4	P19837	Spidroin-1	Trypsin	536 – 576	1056.7	1056.84	-0.1474	3	G.AGAAAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGVRQGGYG.G	No enzyme (two ends) / CID + ETD	230	94.1
4	P19837	Spidroin-1	Trypsin	572 – 585	917.44	917.30	0.1349	2	R.Q ^{569.27} GGYGGLGSQGAGR.G	Unknown shift / CID	331	100
4	P19837	Spidroin-1	Trypsin	586 – 615	922.43	922.38	0.0487	3	R.GGQGAGAAAAAAGGAGQGGYGGLGGQGV ^{407.15} GR.G	Unknown shift / CID + ETD	369	100
4	P19837	Spidroin-1	Trypsin	616 – 652	955.08	955.11	-0.0353	3	R.GGLGGQGAGAAAAGGAGQGGYGGVGSASAASAAASR ^{Deamid} .L	Deamidation (R652) / CID	230	94.7
4	P19837	Spidroin-1	Chymotrypsin	1-18	725.85	725.81	0.0311	2	QGAGAAAAAAGGAGQGGY ^{Nitro} .G	Nitro (Y18) / CID	210	97.4
4	P19837	Spidroin-1	Chymotrypsin	1-21	1044.48	1044.38	0.0901	2	Q ^{455.18} GAGAAAAAAGGAGQGGYGGL.G	Unknown shift / CID	345	100
4	P19837	Spidroin-1	Chymotrypsin	1-31	793	793.04	-0.0445	3	QGAGAAAAAAGGAGQGGYGGLGGQ ^{-88.12} GAGQGGY.G	Unknown shift / CID + ETD	327	98.5
4	P19837	Spidroin-1	Chymotrypsin	19 - 31; 185 - 197; 286 - 298	539.74	539.74	-0.0092	2	Y.GGLGGQGAGQGGY.G	CID	391	100
4	P19837	Spidroin-1	Chymotrypsin	19 - 34; 185 - 200; 286 - 301	653.29	653.31	-0.0227	2	Y.GGLGGQGAGQGGYGGL.G	CID	353	100
4	P19837	Spidroin-1	Chymotrypsin	22 - 31; 188 - 197; 289 - 298	860.38	860.36	0.0158	1	L.G ^{9.02} GQGAGQGGY.G	Unknown shift / CID	99	100
4	P19837	Spidroin-1	Chymotrypsin	22 – 59	1062.8	1062.78	0.016	3	L.G ^{223.05} GQGAGQGGYGGLGGQGAGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	202	92.5
4	P19837	Spidroin-1	Chymotrypsin	32 – 59	1102.49	1102.49	-0.0068	2	Y.GGLGGQGAGQGAGAAAAAAGGAGQGGY.G	CID	422	100
4	P19837	Spidroin-1	Chymotrypsin	35 – 59	916.9	916.93	-0.0357	2	L.GGQGAGQGAGA ^{-71.06} AAAAAAGGAGQGGY.G	Unknown shift / CID	494	100
4	P19837	Spidroin-1	Chymotrypsin	60 - 89; 576 - 605	1144.97	1145.05	-0.0879	2	Y.GGLGSQGAGRGGQGAGAA ^{-71.17} AAAAGGAGQGGY.G	Unknown shift / CID	414	100
4	P19837	Spidroin-1	Chymotrypsin	60 - 92; 576 - 608	851.08	851.08	-0.0034	3	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY ⁻³⁶ GGL.G	Unknown shift / CID + ETD	308	99.1
4	P19837	Spidroin-1	Chymotrypsin	63 – 102	1180.92	1180.84	0.0722	3	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGG ^{340.22} L.G	Unknown shift / CID + ETD	231	97
4	P19837	Spidroin-1	Chymotrypsin	63 - 89; 579 - 605	1180.51	1180.49	0.0156	2	L.G ^{227.03} SQGAGRGGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	351	100
4	P19837	Spidroin-1	Chymotrypsin	63 - 92; 579 - 608	960.43	960.37	0.0556	3	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGL ^{519.17} .G	Unknown shift / CID + ETD	370	100
4	P19837	Spidroin-1	Chymotrypsin	90 - 102; 152 - 164; 198 - 210; 299 - 311; 354 - 366	508.78	508.78	-0.0062	2	Y.GGL ⁻⁷⁰ GSQGAGRGL.G	Unknown shift / CID	253	91.5
4	P19837	Spidroin-1	Chymotrypsin	90 – 123	825.33	825.32	0.0064	4	Y.GGLGSQGAGRGLGGQGAGAAAAAAGGAGQGGY ^{640.02} .G	Unknown shift / ETD	355	87.7
4	P19837	Spidroin-1	Chymotrypsin	93 – 126	980.64	980.42	0.2109	3	L.GSQ ^{281.63} GAGRGLGGQGAGAAAAAAGGAGQGGYGGL.G	Unknown shift / CID + ETD	241	90
4	P19837	Spidroin-1	Chymotrypsin	103 - 123; 431 - 451	803.86	803.86	-0.0039	2	L.GGQGAGA ^{-S-} AAAAAAGGAGQGGY.G	Substitution Ala->Ser (A109, A437) / CID	457	100
4	P19837	Spidroin-1	Chymotrypsin	103 – 151	980.5	980.44	0.0524	4	L.GGQGAGAAAAAAGGAGQGGYGG ^{88.21} LGNQGAGRGGQGAAAAAAGGAGQGGY.G	Unknown shift / ETD	345	99.5
4	P19837	Spidroin-1	Chymotrypsin	124 – 151	800.02	800.00	0.021	3	Y.GGLGNQGAGRGGQGAAAAA ^{138.98} AGGAGQGGY.G	Unknown shift / CID + ETD	247	83.7
4	P19837	Spidroin-1	Chymotrypsin	127 – 151	1049.73	1049.67	0.062	2	L.GNQGAGRGGQGA ^{66.52} AAAAAGGAGQGGY.G	Unknown shift / CID	258	100
4	P19837	Spidroin-1	Chymotrypsin	152 – 184	839.37	839.41	-0.0467	3	Y.GGLGSQGAGRGLGGQGAGAAA ^{-71.13} AAAGGAGQGGY.G	Unknown shift / CID + ETD	524	100
4	P19837	Spidroin-1	Chymotrypsin	155 – 184	853.09	853.04	0.0489	3	L.G ^{197.15} SQGAGRGLGGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	561	100
4	P19837	Spidroin-1	Chymotrypsin	155 – 187	863.08	863.08	-0.0034	3	L.GSQGAGRGLGGQGAGAAAAAAGGAGQGGYGGL.G	CID + ETD	412	100
4	P19837	Spidroin-1	Chymotrypsin	165 – 184	903.41	903.34	0.0622	2	L.G ^{286.12} GQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	512	100
4	P19837	Spidroin-1	Chymotrypsin	165 – 197	640.36	640.54	-0.1859	4	L.G ^{-20.73} GQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Unknown shift / ETD	268	82.7

4	P19837	Spidroin-1	Chymotrypsin	201 – 231	785.43	785.37	0.062	3	L.GSQGAGRGGLGGQGAGA ^{-26.88} AAAAAAGGAGQGGL.G	Unknown shift / CID + ETD	251	83.1
4	P19837	Spidroin-1	Chymotrypsin	211 – 255	863.35	863.38	-0.0346	4	L.GGQGAGAAAAAAGGAGQGGLGGQGAGQGAGAS ^{HP03} AAAAGGAGQGGY.G	Phosphorylation (S243) / CID + ETD	356	84.5
4	P19837	Spidroin-1	Chymotrypsin	232 – 255	952.45	952.41	0.0354	2	L.GGQGAGQGAGAS ^{55.07} AAAAGGAGQGGY.G	Unknown shift / CID	491	100
4	P19837	Spidroin-1	Chymotrypsin	256 – 288	837.73	837.74	-0.0147	3	Y.GGLGSQGAGRGGEAGAAAAAAGG ^{-77.03} AGQGGYGGL.G	Unknown shift / CID + ETD	250	93.6
4	P19837	Spidroin-1	Chymotrypsin	259 – 285	1066.95	1067.01	-0.0626	2	L.GSQGAGRGGE ^{-K-} GAGAAAAAAGGAGQGGY.G	Substitution Glu->Lys (E268) / CID	466	100
4	P19837	Spidroin-1	Chymotrypsin	259 – 288	787.71	787.70	0.0076	3	L.GSQGAGRGGEAGAAAAAAGGAGQGGYGGL.G	/ CID + ETD	413	100
4	P19837	Spidroin-1	Chymotrypsin	259 – 298	1064.14	1064.15	-0.0112	3	L.GSQGAGRG ^{-3.02} GEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Unknown shift / CID + ETD	446	100
4	P19837	Spidroin-1	Chymotrypsin	299 – 329	1152.54	1152.59	-0.0547	2	Y.GGLGSQGAGRGGLGGQGAGAAAAG ^{-91.1} GAGQGGL.G	Unknown shift / CID	209	80.9
4	P19837	Spidroin-1	Chymotrypsin	302 – 353	993.38	993.47	-0.0915	4	L.GSQGAGRGGLGGQGAGAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGY.G	ETD	296	84.8
4	P19837	Spidroin-1	Chymotrypsin	312 – 353	1031.64	1031.81	-0.1795	3	L.GGQGAGAA ^{-48.53} AAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	293	99.1
4	P19837	Spidroin-1	Chymotrypsin	312 – 356	865.71	865.64	0.0618	4	L.GGQG ^{91.25} AGAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGL.G	Unknown shift / ETD	321	87
4	P19837	Spidroin-1	Chymotrypsin	330 – 353	1044.47	1044.41	0.0529	2	L.GGQGA ^{255.11} QGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	461	100
4	P19837	Spidroin-1	Chymotrypsin	330 – 356	1130.07	1129.98	0.0894	2	L.G ^{199.18} QQAGQGAGAAAAAAGGAGQGGYGGL.G	Unknown shift / CID	285	99.9
4	P19837	Spidroin-1	Chymotrypsin	330 – 366	1066.01	1065.79	0.2137	3	L.GGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGRGG ^{295.64} L.G	Unknown shift / CID + ETD	209	80.6
4	P19837	Spidroin-1	Chymotrypsin	354 – 387	931.34	931.24	0.133	3	Y.GGLGSQ ^{105.7} GAGRGGLGGQGAGAVAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	239	93.3
4	P19837	Spidroin-1	Chymotrypsin	367 – 387	917.43	917.382	0.048	2	L.GGQGAGAVA ^{215.1} AAAAGGAGQGGY.G	Unknown shift / CID	471	100
4	P19837	Spidroin-1	Chymotrypsin	388 – 417	883.44	883.40	0.0391	3	Y.GGLGSQGAGRGQGAGAAAAAAGGAGQR ^{189.12} GY.G	Unknown shift / CID + ETD	261	84.2
4	P19837	Spidroin-1	Chymotrypsin	391 – 417	729.6	729.69	-0.0919	3	L.GSQGAGR ^{-45.27} GGQGAGAAAAAAGGAGQRGY.G	Unknown shift / CID	407	100
4	P19837	Spidroin-1	Chymotrypsin	418 – 451	1012.86	1012.76	0.0939	3	Y.GGLNQGAGRGGLGGQG ^{351.28} AGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	219	97.6
4	P19837	Spidroin-1	Chymotrypsin	421 – 451	810.34	810.39	-0.0504	3	L.GNQGAG ^{-29.14} RGGLGGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	352	99.9
4	P19837	Spidroin-1	Chymotrypsin	431 – 478	931.34	931.43	-0.0983	4	L.GGQGAGAAAAAAGGAGQ ^{37.38} GGYGGLNQGAGRGGQGAAAAAGGAGQGGY.G	Unknown shift / ETD	310	91.8
4	P19837	Spidroin-1	Chymotrypsin	455 – 478	1009.04	1008.95	0.0879	2	L.G ^{56.18} NQGAGRGGQGAAAAAAGGAGQGGY.G	Unknown shift / CID	330	100
4	P19837	Spidroin-1	Chymotrypsin	482 – 518	862.27	862.12	0.1404	4	L.GSQGAGRGGQGAGAAAAAAGVAGQEGIRGQGA ^{317.56} QGGY.G	Unknown shift / ETD	320	98.5
4	P19837	Spidroin-1	Chymotrypsin	519 – 551	1405.64	1405.62	0.0107	2	Y.GGLGSQGSGRGGLGGQGAGAAAAA ^{257.02} GGAGQGGL.G	Unknown shift / CID	267	88.7
4	P19837	Spidroin-1	Chymotrypsin	552 – 575	971.47	971.47	-0.0026	2	L.GGQGAGQGAGA ⁻¹⁸ AAAAAGGVRQGGY.G	Unknown shift / CID	266	83.8
4	P19837	Spidroin-1	Chymotrypsin	579 – 618	1074.16	1074.12	0.043	3	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGGQGVG ^{21.92} RGGL.G	Unknown shift / CID + ETD	466	100
4	P19837	Spidroin-1	Chymotrypsin	609 – 636	720.96	721.02	-0.0632	3	L.GGQGVGRGGLGGQGAGAAAAGGA ^{-55.18} QGGY.G	Unknown shift / CID	311	100
4	P19837	Spidroin-1	Chymotrypsin	619 – 636	817.87	817.81	0.0593	2	L.G ^{257.12} QQGAGAAAAGGAGQGGY.G	Unknown shift / CID	482	100
4	P19837	Spidroin-1	Chymotrypsin	637 – 653	695.35	695.35	-0.0077	2	Y.GGVGSGASAASAAASRL.S	CID	277	100
4	P19837	Spidroin-1	Chymotrypsin	654 – 669	780.89	780.90	-0.0125	2	L.SSPQASSRVSS ^{-A-} AVSNL.V	Substitution Ser->Ala (S664) / CID	200	98.4
4	P19837	Spidroin-1	Chymotrypsin	654 – 680	843.79	843.76	0.0217	3	L.SSPQASSRVSS ^{-A-} SAVSNLVASGPTNSAAL.S	Substitution Ser->Ala (S664) / CID	415	99.1
4	P19837	Spidroin-1	Chymotrypsin	670 – 704	1108.89	1108.88	0.0096	3	L.VASGPTNSAALSSTI ^{51.03} SNVVSQIGASNPGLSGCDVL.I	Unknown shift / CID + ETD	400	80.9
4	P19837	Spidroin-1	Chymotrypsin	681 – 698	720.05	719.96	0.0812	3	L.S ^{427.24} STISNVVSQIGASNPGL.S	Unknown shift / CID + ETD	278	97.8
4	P19837	Spidroin-1	Chymotrypsin	730 – 744	718.35	718.36	-0.0124	2	Y.GSAGQAT ^{-A-} QIVGQSVY.Q	Substitution Thr->Ala (T736) / CID	327	100
4	P19837	Spidroin-1	Chymotrypsin	730 – 748	890.44	890.46	-0.0283	2	Y.GSAGQATQI ^{-55.05} VGQSVYQALG.-	Unknown shift / CID	254	80.4
5	B5SYS6	Spidroin-1	Trypsin	1-6	802.41	802.37	0.0388	1	>.M ^{37.04} TWTAR.L	Unknown shift / CID	284	89.5
5	B5SYS6	Spidroin-1	Trypsin	25 – 48	884.82	884.76	0.0589	3	G.QNTPWSSTELADAFINAFLNEAGR.T	No enzyme (one end) / CID + ETD	474	100
5	B5SYS6	Spidroin-1	Trypsin	30 – 48	1013.53	1013.49	0.0326	2	W.SSTELADAFINAFLNEAGR.T	Chymotryptic cleavage (one end) / CID	399	100
5	B5SYS6	Spidroin-1	Trypsin	49 – 68	1066.48	1066.48	-0.0086	2	R.TGAFTADQLDDM ^{+O2} STIGDTLK.T	Sulphone (M60) / CID	449	100
5	B5SYS6	Spidroin-1	Trypsin	49 – 68	711.32	711.32	-0.0082	3	R.TGAFTADQLDDM ^{+O2} STIGDTLK.T	Sulphone (M60) / CID	306	100
5	B5SYS6	Spidroin-1	Trypsin	69 – 79	883.89	883.81	0.0795	2	K.TAMDKMARSN ^{514.16} K.S	Unknown shift / CID	224	89.3
5	B5SYS6	Spidroin-1	Trypsin	74 – 84	783.85	783.80	0.0409	2	K.M ^{Ox} ARSNKSSQS ^{327.08} K.L	Oxidation (M74) / CID	250	85.4
5	B5SYS6	Spidroin-1	Trypsin	80 – 112	982.51	982.42	0.0892	4	K.SSQSKLQAL ^{541.36} NM ^{Ox} AFASSM ^{Ox} AEIAAVEQGGLSVAEK.T	Oxidation (M90, M96) / ETD	447	99.5

5	B5SYS6	Spidroin-1	Trypsin	85 – 112	956.87	956.80	0.0612	3	K.LQALN ^{M^{Ox}} AFASS ^{M^{Ox}} AEIAAVEQGGLSVAEK.T	Oxidation (M90, M96) / CID + ETD	541	100
5	B5SYS6	Spidroin-1	Trypsin	98 – 112	750.89	750.89	-0.0088	2	A.EIAAVEQGGLSVAEK.T	No enzyme (one end) / CID	384	100
5	B5SYS6	Spidroin-1	Trypsin	101 – 112	594.28	594.317	-0.037	2	A.AVEQGGLSVAEK.T	No enzyme (one end) / CID	385	99.9
5	B5SYS6	Spidroin-1	Trypsin	113 – 128	858.91	858.90	0.0026	2	K.TNAIADSLNSAFYQTT.G	No enzyme (one end) / CID	489	100
5	B5SYS6	Spidroin-1	Trypsin	113 – 140	1015.53	1015.50	0.0265	3	K.TNAIADSLNSAFYQTTGAVNVQFVNEI ^{R^{Deamid}} .S	Deamidation (R140) / CID	290	99.3
5	B5SYS6	Spidroin-1	Trypsin	129 – 140	673.36	673.36	-0.0048	2	T.GAVNVQFVNEIR.S	No enzyme (one end) / CID	432	100
5	B5SYS6	Spidroin-1	Trypsin	220 – 238	774.38	774.36	0.0109	2	R.GAGAAAAAAGGAGQGGYN ^{G^{-R⁺}} .G	Substitution Gly->Arg (G238) / CID	509	100
5	B5SYS6	Spidroin-1	Trypsin	220 – 261	1107.83	1107.82	0.0018	3	R.GAGAAAAAAGGAGQGGYNGGQGPSAAAAAAAAS ^{73.01} GAGQGGYG.-	Unknown shift / CID + ETD	240	87.1
5	B5SYS6	Spidroin-1	Chymotrypsin	1-9	792.3	792.28	0.0107	2	>. ^{M^{Ox}} TWTAR ^{505.02} LAL.S	Oxidation (M1) / CID	282	88.4
5	B5SYS6	Spidroin-1	Chymotrypsin	4-12	927.54	927.60	-0.0691	1	W.T ^{30.06} ARLALSIL.A	Unknown shift / CID	263	89.1
5	B5SYS6	Spidroin-1	Chymotrypsin	10-20	725.85	725.81	0.0321	2	L.SILAVLCTQ ^{333.06} GL.F	Unknown shift / CID	258	83.8
5	B5SYS6	Spidroin-1	Chymotrypsin	16 – 29	780.89	780.85	0.0335	2	L.CTQGLFAQG ^{10.07} QNTPW.S	Unknown shift / CID	267	81.6
5	B5SYS6	Spidroin-1	Chymotrypsin	30 – 38	470.71	470.71	-0.0065	2	W.SSTELADAF.I	CID	467	99.7
5	B5SYS6	Spidroin-1	Chymotrypsin	30 – 42	693.34	693.33	0.0072	2	W.SSTELADAFINAF.L	CID	423	100
5	B5SYS6	Spidroin-1	Chymotrypsin	39 – 52	740.91	740.88	0.0292	2	F.INAFLNEAGRTGAF.T	CID	375	100
5	B5SYS6	Spidroin-1	Chymotrypsin	43 – 52	518.25	518.26	-0.0145	2	F.LNEAGRTGAF.T	CID + ETD	335	100
5	B5SYS6	Spidroin-1	Chymotrypsin	43 – 67	871.42	871.40	0.0103	3	F.LNEAGRTGAFTADQLDDMSTIGDTL.K	CID + ETD	498	100
5	B5SYS6	Spidroin-1	Chymotrypsin	44 – 52	535.24	535.25	-0.0154	2	L. ^{N^{HexN}} EAGRTGAF.T	Hexosamine (N44) / CID	373	100
5	B5SYS6	Spidroin-1	Chymotrypsin	44 – 67	884.06	884.04	0.0134	3	L.NEAGRTGAFTADQLDDMSTIGDTL.K	CID + ETD	305	97.9
5	B5SYS6	Spidroin-1	Chymotrypsin	53 – 67	806.37	806.35	0.0136	2	F.TADQLDDMSTIGDTL.K	CID	509	100
5	B5SYS6	Spidroin-1	Chymotrypsin	53 – 77	944.15	944.10	0.0433	3	F.TADQLDDMSTIGDTLKTAMDKMARS.N	CID + ETD	221	94.1
5	B5SYS6	Spidroin-1	Chymotrypsin	68 – 85	657.59	657.68	-0.0922	3	L.KTAM ^{Ox} DK ^{M^{Ox}} ARSNKSSQSK ^{-72.27} L.Q	Unknown shift / CID + ETD	294	88.1
5	B5SYS6	Spidroin-1	Chymotrypsin	86 – 92	405.68	405.69	-0.0144	2	L.QALN ^{M^{Ox}} AF.A	Oxidation (M90) / CID	347	100
5	B5SYS6	Spidroin-1	Chymotrypsin	93 – 107	725.35	725.34	0.0019	2	F.ASS ^{M^{Ox}} AEIAAVEQGGL.S	Oxidation (M96) / CID	521	99.9
5	B5SYS6	Spidroin-1	Chymotrypsin	93 – 120	917.07	917.12	-0.0535	3	F.ASS ^{M^{Ox}} AEIAAVEQGGLSVAEKTNAIADSL.N	Oxidation (M96) / CID + ETD	451	100
5	B5SYS6	Spidroin-1	Chymotrypsin	93 – 125	1111.28	1111.20	0.0752	3	F.ASS ^{M^{Ox}} AEIAAVEQGGLSVAEKTNAIADSLNSAFY.Q	Oxidation (M96) / CID + ETD	433	100
5	B5SYS6	Spidroin-1	Chymotrypsin	108 – 125	634.31	634.31	-0.0044	3	L.SVAEKTNAIADSLNSAFY.Q	CID	298	100
5	B5SYS6	Spidroin-1	Chymotrypsin	126 – 135	532.74	532.77	-0.0322	2	Y.QTTGAVNVQF.V	CID	353	100
5	B5SYS6	Spidroin-1	Chymotrypsin	126 – 142	679.01	679.00	0.0077	3	Y.QTTGAVNVQFVNE ^{159.02} IRSL.I	Unknown shift / CID + ETD	267	92.6
5	B5SYS6	Spidroin-1	Chymotrypsin	136 – 146	662.85	662.85	-0.0001	2	F.VNEIRSLIS ^{M^{Ox}} F.A	Oxidation (M145) / CID	456	99.9
5	B5SYS6	Spidroin-1	Chymotrypsin	147 – 156	520.22	520.23	-0.0182	2	F.AQASANEVSY.G	CID	493	100
5	B5SYS6	Spidroin-1	Chymotrypsin	147 – 160	687.25	687.302	-0.052	2	F.AQASANEVSYGGGY.G	CID	448	100
5	B5SYS6	Spidroin-1	Chymotrypsin	147 – 185	855.63	855.61	0.0124	4	F.AQASANEVSYGGGYG ^{115.05} GGQGGQSAGAAAAAASAGAGQGGY.G	Unknown shift / ETD	347	95.2
5	B5SYS6	Spidroin-1	Chymotrypsin	157 – 188	863.08	863.05	0.0233	3	Y.GGGYGGGQGGQSAGAAAAAAS ^{S^{-Y}} AGAGQGGYGGL.G	Substitution Ser->Tyr (S177) / CID + ETD	301	98.2
5	B5SYS6	Spidroin-1	Chymotrypsin	161 – 185	975.41	975.43	-0.0284	2	Y.GGGQGGQSAGAAAAAASAGAGQGGY.G	CID	522	100
5	B5SYS6	Spidroin-1	Chymotrypsin	161 – 188	787.71	787.67	0.0396	3	Y.GGGQGGQSAGAAAAAASAGAG ^{184.12} QGGYGGL.G	Unknown shift / CID + ETD	353	99.1
5	B5SYS6	Spidroin-1	Chymotrypsin	186 – 209	917.43	917.42	0.0026	2	Y.GGLGGQGAGSAAAAAAS ^{S^{-G}} GAGQGGY.G	Substitution Ser->Gly (S202) / CID	505	100
5	B5SYS6	Spidroin-1	Chymotrypsin	186 – 236	993.38	993.46	-0.0861	4	Y.G ^{-5.34} GLGGQGAGSAAAAAASGAGQGGYGGVGNQGAGRGAGAAAAAAGGAGQGGY.N	Unknown shift / ETD	280	81.9
5	B5SYS6	Spidroin-1	Chymotrypsin	189 – 209	803.86	803.86	-0.0039	2	L.GGQGAGSAAAAAAS ^{S^{-G}} GAGQGGY.G	Substitution Ser->Gly (S202) / CID	457	100
5	B5SYS6	Spidroin-1	Chymotrypsin	210 – 236	1065.99	1066.00	-0.0148	2	Y.GGVGNQGAGRGAGAAAAAAGGAGQGGY.N	CID	445	100
5	B5SYS6	Spidroin-1	Chymotrypsin	237 – 260	820.38	820.298	0.082	3	Y.N ^{526.25} GGQGPSAAAAAAAASGAGQGGY.G	Unknown shift / CID + ETD	417	99.5
5	B5SYS6	Spidroin-1	Chymotrypsin	237 – 261	973.43	973.45	-0.0241	2	Y.NGGQGPSAAAAAAAAS ^{44.04} ASGAGQGGYG.-	Unknown shift / CID	213	99.6
5	B5SYS6	Spidroin-1	Glu-C	1-33	886.51	886.71	-0.2029	4	MTWTARL ^{-108.8} ALSILAVL ^{C^{CAme}} TQGLFAQGQNTPWSSTE.L	Unknown shift / ETD	258	81
5	B5SYS6	Spidroin-1	Glu-C	34 – 45	669.36	669.34	0.0196	2	E.LADAFINAFLE.A	CID	361	100

5	B5SYS6	Spidroin-1	Glu-C	99 – 111	919.88	919.83	0.0499	2	E.I ^{595.1} AAVEQGGLSVAE.K	Unknown shift / CID	272	82.9
5	B5SYS6	Spidroin-1	Glu-C	104 – 111	760.39	760.38	0.0064	1	E.QGGLSVAE.K	CID	228	88.7
5	B5SYS6	Spidroin-1	Glu-C	112 – 138	968.5	968.47	0.0266	3	E.KTNAIADSLNSAFYQTTGAVNVQFVNE.I	CID + ETD	456	100
5	B5SYS6	Spidroin-1	Glu-C	139 – 153	827.46	827.41	0.0431	2	E.IRSLISM ^{0x} FAQASANE.V	Oxidation (M145) / CID	330	100
5	P19837	Spidroin-1	Trypsin	50 - 69; 80 - 99; 142 - 161; 246 - 265; 344 - 363; 378 - 397; 469 – 488	824.9	824.89	0.0071	2	A.AAGGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID	424	100
5	P19837	Spidroin-1	Trypsin	60 - 89; 576 - 605	787.22	787.37	-0.1544	3	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY.G	Chymotryptic cleavage (two ends) / CID + ETD	454	100
5	P19837	Spidroin-1	Trypsin	70 – 99	811.08	811.04	0.0389	3	R.GGQG ^{71.12} AGAAAAAAGGAGQGGYGGLGSQGAGR.G	Unknown shift / CID + ETD	565	100
5	P19837	Spidroin-1	Trypsin	134 – 161	730.33	730.35	-0.0249	3	R.G ^{43.07} GQGAAAAAAGGAGQGGYGGLGSQGAGR.G	Unknown shift / CID + ETD	473	100
5	P19837	Spidroin-1	Trypsin	162 – 207	928.17	928.15	0.022	4	R.GGLGGQG ^{62.95} AGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	Unknown shift / ETD	391	99.8
5	P19837	Spidroin-1	Trypsin	170 - 208; 271 - 309	777.37	777.36	0.0037	4	A.GAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (two ends) / ETD	376	96.6
5	P19837	Spidroin-1	Trypsin	174 - 207; 275 - 308	927.09	927.10	-0.0112	3	A.AAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID + ETD	488	100
5	P19837	Spidroin-1	Trypsin	179 - 207; 280 - 308	817.98	818.04	-0.0698	3	G.AGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / CID + ETD	279	84.9
5	P19837	Spidroin-1	Trypsin	216 – 265	960.94	960.95	-0.0181	4	A.GAAAAAAGGAGQGGLGGQGAGQGAGASAAAGGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / ETD	489	100
5	P19837	Spidroin-1	Trypsin	266 – 308	1140.81	1140.86	-0.0502	3	R.GGEGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQGAGR.G	CID + ETD	351	100
5	P19837	Spidroin-1	Trypsin	312 – 363	996.29	996.22	0.0685	4	L.GGQGAGAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGLGSQGAGR.G	No enzyme (one end) / ETD	514	100
5	P19837	Spidroin-1	Trypsin	364 – 397	1056.16	1056.10	0.0538	3	R.GGLGGQGA ^{480.16} GAVAAAAAGGAGQGGYGGLGSQGAGR.G	Unknown shift / CID + ETD	254	95.3
5	P19837	Spidroin-1	Trypsin	398 – 415	994.42	994.34	0.0747	2	R.GGQGAGAAAAAAGGAGQR ^{589.15} .G	Unknown shift / CID	364	100
5	P19837	Spidroin-1	Trypsin	398 – 427	846.07	846.02	0.053	3	R.GGQGAGAAAAAAGGAG ⁵⁰ QRGYGLGNQGAGR.G	Unknown shift / CID + ETD	214	83.6
5	P19837	Spidroin-1	Trypsin	462 – 488	1095.02	1095.03	-0.0114	2	R.GGQ ^{R~} GAAAAAGGAGQGGYGGLGSQGAGR.G	Substitution / CID	540	100
5	P19837	Spidroin-1	Trypsin	529 – 571	1048.09	1048.20	-0.112	3	R.GGLGGQGAGAAAAAAGG ^{-90.33} AGQGGLGGQGAGQGAGAAAAAAGGVR.Q	Unknown shift / CID + ETD	278	85.2
5	P19837	Spidroin-1	Trypsin	572 – 585	917.4	917.30	0.0949	2	R.Q ^{569.19} GGYGGLGSQGAGR.G	Unknown shift / CID	382	100
5	P19837	Spidroin-1	Trypsin	572 – 615	951.21	951.18	0.0259	4	R.QG ^{198.1} YGGLGSQAGRGGQGAGAAAAAAGGAGQGGYGGLGGQGVGR.G	Unknown shift / ETD	290	83.9
5	P19837	Spidroin-1	Trypsin	586 – 615	786.69	786.71	-0.0246	3	R.GGQGAGAAAAAAGGAGQGGYGGLGGQGVGR.G	CID + ETD	328	99.9
5	P19837	Spidroin-1	Trypsin	616 – 652	954.79	954.78	0.0027	3	R.GGLGGQGAGAAAAGGAGQGGYGGVGSASAASAAASR.L	CID + ETD	539	100
5	P19837	Spidroin-1	Trypsin	616 – 661	949.72	949.70	0.0123	4	R.GGLGGQGAGAAAAGGAGQGGYGGVGS ^{20.05} GASAASAAASRLSSPQASSR.V	Unknown shift / ETD	321	87
5	P19837	Spidroin-1	Trypsin	653 – 661	523.26	523.24	0.0166	2	R.LSSPQA ^{113.03} SSR.V	Unknown shift / CID	293	90.7
5	P19837	Spidroin-1	Chymotrypsin	1-18	721.84	721.82	0.0136	2	QGA ^{37.03} GAAAAAAGGAGQGGY.G	Unknown shift / CID	258	93.4
5	P19837	Spidroin-1	Chymotrypsin	1-21	817.55	817.38	0.1601	2	QGAGAAAAA ^{1.32} GGAGQGGYGGL.G	Unknown shift / CID	208	96.5
5	P19837	Spidroin-1	Chymotrypsin	1-31	881.17	881.04	0.1255	3	QGAGAAAAAAGGAGQGG ^{176.38} YGGLGGQGAGQGGY.G	Unknown shift / CID + ETD	232	87.8
5	P19837	Spidroin-1	Chymotrypsin	19 - 34; 185 - 200; 286 - 301	653.28	653.31	-0.0327	2	Y.GGLGGQGAGQGGYGGL.G	CID	382	100
5	P19837	Spidroin-1	Chymotrypsin	22 - 34; 188 - 200; 289 – 301	539.68	539.74	-0.0692	2	L.GGQGAGQGGYGGL.G	CID	333	100
5	P19837	Spidroin-1	Chymotrypsin	22 – 59	1025.46	1025.45	0.0093	3	L.GGQGAGQGGYGG ^{111.03} LGGQGAGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	272	99.1
5	P19837	Spidroin-1	Chymotrypsin	32 – 59	687.31	687.33	-0.0252	3	Y.GGLGGQGAGQ ^{71.07} GAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	480	100
5	P19837	Spidroin-1	Chymotrypsin	35 – 59	1102.46	1102.43	0.0243	2	L.GGQGAG ^{300.05} QGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	492	100
5	P19837	Spidroin-1	Chymotrypsin	35 – 62	534.31	534.25	0.0568	4	L.GGQG ^{3.23} AGQGAGAAAAAAGGAGQGGYGGL.G	Unknown shift / ETD	332	84.1
5	P19837	Spidroin-1	Chymotrypsin	60 - 89; 576 - 605	787.34	787.37	-0.0344	3	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGY.G	CID + ETD	489	100
5	P19837	Spidroin-1	Chymotrypsin	60 - 92; 576 - 608	1036.13	1036.08	0.0466	3	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGG ^{519.14} L.G	Unknown shift / CID + ETD	499	100
5	P19837	Spidroin-1	Chymotrypsin	63 - 89; 579 - 605	1030.47	1030.49	-0.0244	2	L.GSQG ^{-73.04} AGRGGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	496	100
5	P19837	Spidroin-1	Chymotrypsin	63 - 92; 579 - 608	960.45	960.37	0.0756	3	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGL ^{519.23} .G	Unknown shift / CID + ETD	491	100
5	P19837	Spidroin-1	Chymotrypsin	90 - 102; 152 - 164; 198 - 210; 299 - 311; 354 - 366	543.74	543.78	-0.0462	2	Y.GGLGSQGAGRGG.L.G	CID	294	100
5	P19837	Spidroin-1	Chymotrypsin	93 – 123	811.02	811.05	-0.0334	3	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQGGY.G	CID + ETD	479	100
5	P19837	Spidroin-1	Chymotrypsin	93 – 126	1059.8	1059.76	0.0376	3	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQGGYGG ^{519.11} L.G	Unknown shift / CID + ETD	393	100

5	P19837	Spidroin-1	Chymotrypsin	103 – 123; 431 - 451	787.32	787.36	-0.0464	2	L.GGQGAGA ^{-17.08} AAAAAAGGAGQGGY.G	Unknown shift / CID	326	99.8
5	P19837	Spidroin-1	Chymotrypsin	124 – 154	1240.38	1240.59	-0.2176	2	Y.GGLGNQGAGRGGQGAAAAAAGGAGQGGYG ^{-6.43} GL.G	Unknown shift / CID	220	84.5
5	P19837	Spidroin-1	Chymotrypsin	127 – 151	1050.99	1050.97	0.0194	2	L.GNQGAGRGGQGAAA ^{69.04} AAAGGAGQGGY.G	Unknown shift / CID	274	100
5	P19837	Spidroin-1	Chymotrypsin	127 – 154	735.69	735.69	-0.0018	3	L.GNQGAGRGGQG ⁻⁵⁴ AAAAAAGGAGQGGYGGL.G	Unknown shift / CID + ETD	252	83.4
5	P19837	Spidroin-1	Chymotrypsin	152 – 184	839.38	839.41	-0.0367	3	Y.GGLGSQGAGRGGGLGGQGAGAAAAA ^{-71.1} AGGAGQGGY.G	Unknown shift / CID + ETD	551	100
5	P19837	Spidroin-1	Chymotrypsin	155 – 184	853.07	853.04	0.0289	3	L.G ^{197.09} SQGAGRGGGLGGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	555	100
5	P19837	Spidroin-1	Chymotrypsin	165 – 184	859.89	859.84	0.0422	2	L.G ^{199.08} GQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	216	99.9
5	P19837	Spidroin-1	Chymotrypsin	165 – 187	839.4	839.41	-0.0113	2	L.GGQGAGAAAAAA ^{-69.01} GGAGQGGYGGL.G	Unknown shift / CID	263	98.9
5	P19837	Spidroin-1	Chymotrypsin	165 – 197	1088.45	1088.39	0.0579	3	L.G ^{684.17} GQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Unknown shift / CID + ETD	497	100
5	P19837	Spidroin-1	Chymotrypsin	188 - 210; 289 - 311	674.34	674.30	0.0313	3	L.GGQGAGQGGYGGLGSQGAGR ^{102.09} GL.G	Unknown shift / CID	221	90.3
5	P19837	Spidroin-1	Chymotrypsin	201 – 231	827.4	827.42	-0.0202	3	L.GSQGAGRGGGLGGQGAGAAAAAAGGAGQ ^{G^{-R}} GL.G	Substitution Gly->Arg (G229) / CID + ETD	426	98.7
5	P19837	Spidroin-1	Chymotrypsin	211 – 255	896.21	896.143	0.067	4	L.GGQG ^{211.27} AGAAAAAAGGAGQGGLGGQGAGQGAGASAAAAGGAGQGGY.G	Unknown shift / ETD	269	69.3
5	P19837	Spidroin-1	Chymotrypsin	232 – 255	916.88	916.91	-0.0372	2	L.GGQGAGQGAGASAAAAGGAGQGGY.G	CID	526	100
5	P19837	Spidroin-1	Chymotrypsin	256 – 285	1180.52	1180.55	-0.0379	2	Y.GGLGSQGAGRGGEGAGAAAAAAGGAGQGGY.G	CID	348	100
5	P19837	Spidroin-1	Chymotrypsin	259 – 285	787.38	787.32	0.0533	3	L.GSQG ^{226.16} AGRGGEAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	388	100
5	P19837	Spidroin-1	Chymotrypsin	259 – 298	1064.8	1064.83	-0.0353	3	L.GSQGAGRGG ^{E^{-K}} GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Substitution Glu->Lys (E268) / CID + ETD	496	100
5	P19837	Spidroin-1	Chymotrypsin	302 – 329	604.62	604.51	0.1008	4	L.GSQGAGRGGGL ^{G^{247.4}} GQGAGAAAAGGAGQGGL.G	Unknown shift / ETD	305	82.9
5	P19837	Spidroin-1	Chymotrypsin	312 – 353	1014.47	1014.48	-0.0162	3	L.GGQGAGAAAAGGAGQ ^{-100.04} GGLGGQGAGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	209	95
5	P19837	Spidroin-1	Chymotrypsin	330 – 353	1044.43	1044.41	0.0129	2	L.GGQGAGQ ^{255.03} GAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	433	100
5	P19837	Spidroin-1	Chymotrypsin	330 – 356	493.23	493.244	-0.014	4	L.GGQGAGQGAGAAAAAAGGAGQ ^{-90.05} GGYGGL.G	Unknown shift / ETD	278	88.5
5	P19837	Spidroin-1	Chymotrypsin	354 – 387	995.21	995.10	0.1038	3	Y.GGLGSQGAGRGGGLGGQG ^{297.31} AGAVAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	295	84.8
5	P19837	Spidroin-1	Chymotrypsin	357 – 387	815.69	815.72	-0.0353	3	L.GSQGAGRGGGLGGQGAGAVAAA ^{A^{-G}} AGGAGQGGY.G	Substitution Ala->Gly (A378) / CID + ETD	383	100
5	P19837	Spidroin-1	Chymotrypsin	357 – 390	892.74	892.77	-0.0329	3	L.G ^{-10.09} SQGAGRGGGLGGQGAGAVAAAAAGGAGQGGYGGL.G	Unknown shift / CID + ETD	313	99.8
5	P19837	Spidroin-1	Chymotrypsin	367 – 387	902.4	902.382	0.018	2	L.GGQGAGAVA ^{185.04} AAAAGGAGQGGY.G	Unknown shift / CID	353	100
5	P19837	Spidroin-1	Chymotrypsin	388 – 417	805.25	805.40	-0.1509	3	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQ ^{-45.44} RGY.G	Unknown shift / CID + ETD	213	83.6
5	P19837	Spidroin-1	Chymotrypsin	388 – 420	1012.46	1012.44	0.0167	3	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQRGYGGL ^{349.05} .G	Unknown shift / CID + ETD	387	100
5	P19837	Spidroin-1	Chymotrypsin	418 – 430	372.76	372.63	0.132	3	Y.GGLGNQGAGR ^{G^{2.69}} GL.G	Unknown shift / CID + ETD	304	86.5
5	P19837	Spidroin-1	Chymotrypsin	421 – 451	806.33	806.39	-0.0604	3	L.GNQGAGRGGGLGGQGAGA ^{-41.17} AAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	446	100
5	P19837	Spidroin-1	Chymotrypsin	431 – 478	968.73	968.68	0.0417	4	L.GGQGAGAAAAAAGGAGQGGYGGL ^{112.17} GNQGAGRGGQGAAAAAAGGAGQGGY.G	Unknown shift / ETD	345	99.4
5	P19837	Spidroin-1	Chymotrypsin	479 – 518	1141.29	1141.21	0.0773	3	Y.GGLGSQGAGRGGQGAGAAAAAV ^{66.23} GAGQEGIRQGAGQGGY.G	Unknown shift / CID	263	98.1
5	P19837	Spidroin-1	Chymotrypsin	522 – 551	747	747.04	-0.0463	3	L.GSQSGRGGGLGGQGAG ^{-87.13} AAAAAAGGAGQGGL.G	Unknown shift / CID + ETD	466	99.9
5	P19837	Spidroin-1	Chymotrypsin	532 – 578	893.34	893.44	-0.1045	4	L.GGQGAGAAAAAAG ^{-67.41} GAGQGGLGGQGAGQGAGAAAAAAGGVRQGGYGGL.G	Unknown shift / ETD	295	93.3
5	P19837	Spidroin-1	Chymotrypsin	576 – 617	1147.78	1147.88	-0.1083	3	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQGGYGGLGGQGVGRG ^{G^{-W}} .L	Substitution Gly->Trp (G617) / CID + ETD	217	99.5
5	P19837	Spidroin-1	Chymotrypsin	606 – 636	815.22	815.39	-0.1736	3	Y.GGLGGQGVGR ^{R^{Deamid}} GGLGGQGAGAAAAGGAGQGGY.G	Deamidation (R615) / CID	414	100
5	P19837	Spidroin-1	Chymotrypsin	609 – 636	825.06	825.02	0.0368	3	L.G ^{257.11} GQGVGRGGLGGQGAGAAAAGGAGQGGY.G	Unknown shift / CID + ETD	520	100
5	P19837	Spidroin-1	Chymotrypsin	619 – 653	667.37	667.58	-0.2116	4	L.GGQGAGAAAAGGAGQ ^{-81.84} GGYGGVGSASASAAASRL.S	Unknown shift / ETD	279	80.8
5	P19837	Spidroin-1	Chymotrypsin	637 – 653	695.32	695.35	-0.0377	2	Y.GGVGSGASASAAASRL.S	CID	380	100
5	P19837	Spidroin-1	Chymotrypsin	654 – 669	780.9	780.90	-0.0025	2	L.SSPQASSRVSS ^{-A⁻} AVSNL.V	Substitution Ser->Ala (S664) / CID	367	99.8
5	P19837	Spidroin-1	Chymotrypsin	654 – 680	843.81	843.76	0.0417	3	L.SSPQASSRV ^{S^{-A⁻}} SAVSNLVASGPTNSAAL.S	Substitution Ser->Ala (S664) / CID + ETD	472	99.9
5	P19837	Spidroin-1	Chymotrypsin	681 – 708	1330.11	1330.20	-0.0926	2	L.SSTISNVV ^{-71.18} SQIGASNPLSGCDVLIQAL.L	Unknown shift / CID	210	81.9
5	P19837	Spidroin-1	Chymotrypsin	716 – 744	965.92	965.83	0.0869	3	L.IQILGSS ^{S^{-G⁻}} IGQVNYGSAGQATQIVGQSVY.Q	Substitution Ser->Gly (S723) / CID + ETD	245	87.1
5	P19837	Spidroin-1	Chymotrypsin	730 – 744	729.35	729.36	-0.0177	2	Y.GSAGQAT ^{-8.03} QIVGQSVY.Q	Unknown shift / CID	248	100
6	B5SYS5	Spidroin-1	Trypsin	1-6	401.69	401.68	0.0008	2	M ³⁷ TWTAR.L	Unknown shift / CID	311	97.1

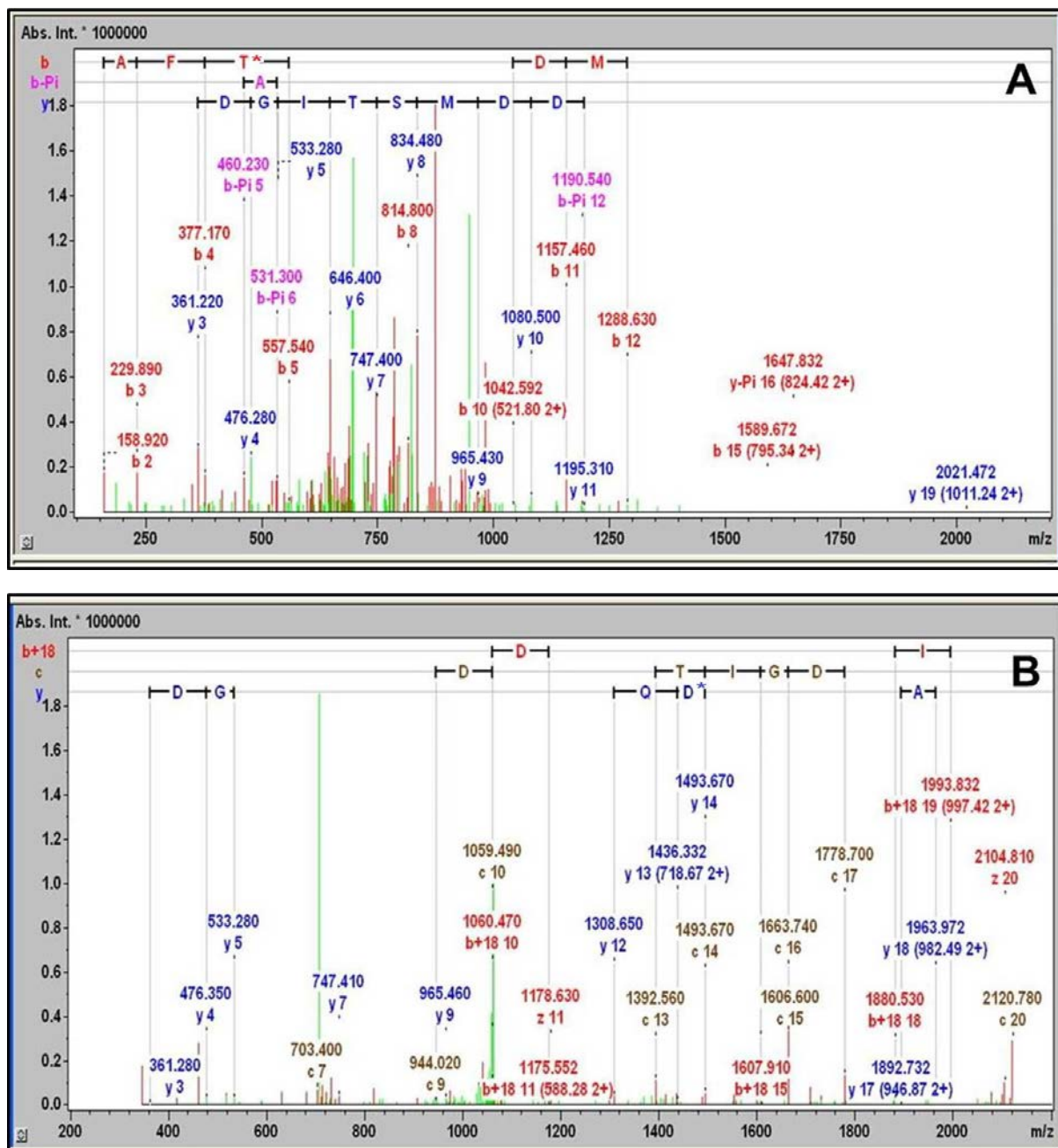
6	B5SYS5	Spidroin-1	Trypsin	49 – 68	711.33	711.32	0.0018	3	R.TGAFTADQLDDM ^{+O2} STIGDTIK.T	Sulphone (M60) / CID	375	100
6	B5SYS5	Spidroin-1	Trypsin	49 – 68	1066.43	1066.48	-0.0586	2	R.TGAFTADQLDDM ^{+O2} STIGDTIK.T	Sulphone (M60) / CID	404	100
6	B5SYS5	Spidroin-1	Trypsin	49 – 68	707.993	707.983	0.010	3	R.TGAFT ^{HPO3} AD ^{-G-} QLDDMSTIGDTIK.T	Phosphorylation (T53); Substitution Asp->Gly (D55) / CID	278	99
6	B5SYS5	Spidroin-1	Trypsin	49 – 68	713.342	713.321	0.021	3	R.TGAFTADQLDD ^{-G-} MS ^{T^{HPO3}} IGDTIK.T	Phosphorylation (T62); Substitution Asp->Gly (D59) / CID	267	98.2
6	B5SYS5	Spidroin-1	Trypsin	49 – 73	912.43	912.41	0.0194	3	R.TGAFTADQLDDM ^{Ox} STIGDTIK ^{57.06} TAM ^{Ox} DK.M	Unknown shift / CID + ETD	269	91.1
6	B5SYS5	Spidroin-1	Trypsin	52 – 77	954.79	954.78	0.0079	3	A.FTADQLDDMSTIGDTIKTAMDKMARS.N	No enzyme (two ends) / CID + ETD	243	88.4
6	B5SYS5	Spidroin-1	Trypsin	69 – 79	883.86	883.81	0.0495	2	K.TAMDKMARSN ^{514.1} K.S	Unknown shift / CID	251	82.7
6	B5SYS5	Spidroin-1	Trypsin	74 – 82	946.41	946.52	-0.1104	1	K.M ^{Ox} ARSNKSS ^{-78.1} K.G	Oxidation (M74) / CID	250	88.3
6	B5SYS5	Spidroin-1	Trypsin	80 – 112	1150.33	1150.23	0.0996	3	K.SSKGKLQ ^{123.3} ALNMAFASSM ^{Ox} AEIAAVEQGGLSVDAK.T	Unknown shift / CID + ETD	269	82.1
6	B5SYS5	Spidroin-1	Trypsin	85 – 112	952.17	952.13	0.0331	3	K.LQALNM ^{Ox} AFASSM ^{Ox} AEIAAVEQGGLSVDAK.T	Oxidation (M90, M96) / CID + ETD	492	100
6	B5SYS5	Spidroin-1	Trypsin	113 – 140	1005.16	1005.15	0.0001	3	K.TNAIADSLNSAFYQTTGAANPQFVNEIR.S	CID + ETD	506	100
6	B5SYS5	Spidroin-1	Trypsin	203 – 216	632.78	632.80	-0.0251	2	E.QGGYGGLGSQGAGR.G	No enzyme (one end) / CID	377	99.9
6	B5SYS5	Spidroin-1	Trypsin	217 – 261	1202.92	1202.86	0.0563	3	R.GGYGGQGAGAAAAAAGGAGQGGQGLSGQGAAAAA ^{144.17} GGAGQGGYG	Unknown shift / CID + ETD	248	87.2
6	B5SYS5	Spidroin-1	Chymotrypsin	1-7	819.28	819.45	-0.1702	1	M ^{Ox} TWTA ^{-75.16} RL.A	Oxidation (M1) / CID	249	92.6
6	B5SYS5	Spidroin-1	Chymotrypsin	4-11	412.7	412.75	-0.0584	2	W.TARLAL ^{-54.11} SF.L	Unknown shift / CID	324	99
6	B5SYS5	Spidroin-1	Chymotrypsin	12-21	1023.49	1023.59	-0.1014	1	F.LAVICTQS ^{-71.09} LF.A	Unknown shift / CID	214	89.3
6	B5SYS5	Spidroin-1	Chymotrypsin	13 – 29	984.47	984.45	0.0135	2	L.AVICTQS ^{104.03} LFAQQNTPW.S	Unknown shift / CID	218	82.1
6	B5SYS5	Spidroin-1	Chymotrypsin	22 – 29	808.35	808.41	-0.0663	1	F.AQQQN ^{-93.06} TPW.S	Unknown shift / CID	212	90.1
6	B5SYS5	Spidroin-1	Chymotrypsin	30 – 38	940.41	940.42	-0.0158	1	W.SSTELADAF.I	CID	290	99.2
6	B5SYS5	Spidroin-1	Chymotrypsin	30 – 42	693.31	693.33	-0.0228	2	W.SSTELADAFINAF.M	CID	379	100
6	B5SYS5	Spidroin-1	Chymotrypsin	43 – 52	518.27	518.26	0.0055	2	F.MNEAGRTGAF.T	CID	339	100
6	B5SYS5	Spidroin-1	Chymotrypsin	53 – 77	708.39	708.33	0.0581	4	F.TADQLDDMSTIGDTIKTAMDKM ^{Ox} ARS ^{-W-} .N	Substitution Ser->Trp (S77) / ETD	290	85.4
6	B5SYS5	Spidroin-1	Chymotrypsin	86 – 107	747.04	747.02	0.0107	3	L.QALNMAFASSM ^{Ox} AEIAAVE ^{Me} QGGL.S	Methylation (E103) / CID	392	97.5
6	B5SYS5	Spidroin-1	Chymotrypsin	93 – 120	879.05	879.11	-0.0683	3	F.ASSM ^{Ox} AEIAAVE ^{-100.2} QGGLSVDAKTNAIADSL.N	Unknown shift / CID	220	85.4
6	B5SYS5	Spidroin-1	Chymotrypsin	108 – 125	951	950.968	0.032	2	L.S ^{Me} VDAKTNAIADSLNSAFY.Q	Methylation (S108) / CID	432	100
6	B5SYS5	Spidroin-1	Chymotrypsin	126 – 142	923.47	923.47	-0.0063	2	Y.QTTGAANPQFVNEIRSL.I	CID	226	100
6	B5SYS5	Spidroin-1	Chymotrypsin	136 – 146	662.84	662.85	-0.0101	2	F.VNEIRSLIN ^{-S-} M ^{Ox} F.A	Substitution Asn->Ser (N144) / CID	284	99.9
6	B5SYS5	Spidroin-1	Chymotrypsin	147 – 156	520.23	520.23	-0.0082	2	F.AQSS ^{-A-} ANEVSY.G	Substitution Ser->Ala (S150) / CID	421	100
6	B5SYS5	Spidroin-1	Chymotrypsin	147 – 182	1007.04	1007.11	-0.0736	3	F.AQSSANEVSYGGGYGGQSAGAAASA ^{-45.21} AAAGGGGQGGY.G	Unknown shift / CID + ETD	220	98
6	B5SYS5	Spidroin-1	Chymotrypsin	161 – 182	934.270	934.260	0.010	2	Y.GGQS ^{HPO3} AGAAA ^{S^{HPO3}} AAAAGGG ^{-A-} GQGGY.G	Phosphorylation (S164, S170); Substitution Gly->Ala (G177) / CID	256	98.1
6	B5SYS5	Spidroin-1	Chymotrypsin	161 – 185	685.03	684.97	0.0582	3	Y.GGQSAGAAASAA ^{75.17} AAGGGGQGGYGNL.G	Unknown shift / CID	240	80.6
6	B5SYS5	Spidroin-1	Chymotrypsin	183 – 206	720.35	720.312	0.038	3	Y.GNLGGQGAGAAAAAAS ^{168.11} AAEQGGY.G	Unknown shift / CID + ETD	318	99.3
6	B5SYS5	Spidroin-1	Chymotrypsin	183 – 209	886.73	886.68	0.0424	3	Y.G ^{440.13} NLGGQGAGAAAAAASAAEQGGYGGL.G	Unknown shift / CID	361	99.2
6	B5SYS5	Spidroin-1	Chymotrypsin	207 – 219	568.74	568.77	-0.0358	2	Y.GGLGSQGAGRGGY.G	CID	352	100
6	B5SYS5	Spidroin-1	Chymotrypsin	207 – 243	855.42	855.35	0.0671	4	Y.GGLGSQGAGRGGYGGQGAGAAAAAAGGAGQG ^{504.27} GQGL.S	Unknown shift / ETD	367	94.6
6	B5SYS5	Spidroin-1	Chymotrypsin	210 – 219	961.350	961.330	0.020	1	L.GS ^{HPO3} QGAGR ^{-K-} GGY.G	Phosphorylation (S211); Substitution Arg->Lys (R216) / CID	287	96.8
6	B5SYS5	Spidroin-1	Chymotrypsin	210 – 260	991.91	991.96	-0.0575	4	L.GSQGAGRGGYGGQGAGAAAAAAGGAGQGGQGLSGQG ^{-54.22} AAAAAGGAGQGGY.G	Unknown shift / ETD	283	81.9
6	B5SYS5	Spidroin-1	Chymotrypsin	220 – 243	794.39	794.29	0.0973	3	Y.GGQGAGAA ^{584.29} AAAAAAGGAGQGGQGL.S	Unknown shift / CID + ETD	345	95
6	B5SYS5	Spidroin-1	Chymotrypsin	220 – 260	1048.27	1048.15	0.1103	3	Y.GGQGAGAAAAAAGGAGQGGQ ^{Me} GLSGQGAAAAAGGAGQGGY.G	Methylation (Q241) / CID + ETD	282	99.7
6	B5SYS5	Spidroin-1	Chymotrypsin	220 – 261	873.39	873.373	0.017	4	Y.GGQGAGAAAAAAGGAGQGGQGLSG ^{305.07} QGAAAAAGGAGQGGYG.-	Unknown shift / ETD	348	87.5
6	B5SYS5	Spidroin-1	Chymotrypsin	244 – 260	715.715	715.702	0.013	2	L.S ^{HPO3} GQGAAAAAGGAGQGGY.G	Phosphorylation (S244) / CID	289	94.2
6	B5SYS5	Spidroin-1	Chymotrypsin	244 – 261	696.67	696.54	0.1236	3	L.S ^{680.37} GQGAAAAAGGAGQGGYG.-	Unknown shift / CID	313	99.1
6	B5SYS5	Spidroin-1	Glu-C	1-33	1337.08	1336.93	0.1424	3	M ^{Ox} TWTA ^{334.43} RLALSFLAVICTQSLFAQGQNTPWSSTE.L	Oxidation (M1) / CID	270	87.8

6	B5SYS5	Spidroin-1	Glu-C	34 – 45	686.32	686.31	0.0039	2	E.LADAFINAFM ^{ox} NE.A	Oxidation (M43) / CID	328	100
6	B5SYS5	Spidroin-1	Glu-C	99 – 103	502.24	502.28	-0.0471	1	E.IAAVE.Q	CID	236	97.4
6	B5SYS5	Spidroin-1	Glu-C	104 – 138	900.89	900.93	-0.0427	4	E.QGGLSVDAKTNAIADSLNSAFYQ ^{-E-} TTGAANPQFVNE.I	Substitution Gln->Glu (Q126) / ETD	442	99.3
6	B5SYS5	Spidroin-1	Glu-C	139 – 153	834.44	834.42	0.0153	2	E.IRSLINMFAQSSANE.V	CID	424	100
6	B5SYS5	Spidroin-1	Glu-C	154 – 202	999.66	999.45	0.2089	4	E.VSYGGGYGGQSAGAAASAAAAGGGGQGGYG ^{51.84} GNLGG ^{51.84} QGAGAAAAAASAAE.Q	Unknown shift / ETD	261	91.5
6	P19837	Spidroin-1	Trypsin	70 – 99	1014.1	1014.04	0.0589	3	R.G ^{680.18} QQGAGAAAAAAGGAGQGGYGGLGSQAGR.G	Unknown shift / CID	251	97.1
6	P19837	Spidroin-1	Trypsin	134 – 161	1073.02	1073.02	-0.0087	2	R.G ^{-87.01} QQGAAAAAAGGAGQGGYGGLGSQAGR.G	Unknown shift / CID	222	89.4
6	P19837	Spidroin-1	Trypsin	165 – 207	1140.49	1140.53	-0.0422	3	L.GGQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGYGGLGSQAGR.G	No enzyme (one end) / CID + ETD	491	100
6	P19837	Spidroin-1	Trypsin	266 – 308	861.14	860.897	0.243	4	R.GGEGAGAAAAAAGGAGQGGYG ^{20.97} GLGGQGAGQGGYGGLGSQAGR.G	Unknown shift / ETD	320	92.3
6	P19837	Spidroin-1	Trypsin	333 – 363	811.04	861.11	0.031	3	Q.GAGQGAGAAAAAAGGAGQGGYGGLGSQAGR.G	No enzyme (one end) / CID	393	100
6	P19837	Spidroin-1	Trypsin	364 – 397	1055.18	1055.10	0.0738	3	R.GGLGGQGAGAVAAAAAGG ^{477.22} AGQGGYGGLGSQAGR.G	Unknown shift / CID + ETD	252	99.5
6	P19837	Spidroin-1	Trypsin	398 – 415	665.3	665.34	-0.0453	2	R.GGQGAGAAAAA ^{-69.08} AGGAGQR.G	Unknown shift / CID	290	89.3
6	P19837	Spidroin-1	Trypsin	398 – 427	1036.15	1036.07	0.0788	3	R.GGQGAGAAAAAAGGAGQGRGYGGLGN ^{620.24} QGAGR.G	Unknown shift / CID + ETD	320	99.8
6	P19837	Spidroin-1	Trypsin	416 – 461	948.67	948.55	0.124	4	R.GYGGLGNQGAGRGLGGQGAGAAAAAAGGAGQGG ^{18.86} YGGLGNQGAGR.G	Unknown shift / ETD	298	89.3
6	P19837	Spidroin-1	Trypsin	462 – 488	706.61	706.67	-0.0658	3	R.GGQGAA ^{-43.19} AAAGGAGQGGYGGLGSQAGR.G	Unknown shift / CID	287	100
6	P19837	Spidroin-1	Trypsin	489 – 528	1101.98	1101.94	0.041	3	R.GGQGAGAAAAAAGVAGQEGIRGQAGQ ^{-67.68} GGYGGLGSQSGR.G	Unknown shift / CID + ETD	299	86.5
6	P19837	Spidroin-1	Trypsin	572 – 585	598.94	598.87	0.0675	3	R.QGGYGGLGSQAG ^{530.2} R.G	Unknown shift / CID + ETD	317	85.1
6	P19837	Spidroin-1	Trypsin	572 – 615	1216.94	1216.90	0.0307	3	R.Q ^{Me+Deamid} GGYGGLGSQAGRGGQ ^{Me+Deamid} GAGAAAAAAGGAGQGGYGGLGGQ ^{Me+Deamid} GVGR.G	Methylation and Deamidation (Q572, Q588) / CID	365	97.8
6	P19837	Spidroin-1	Trypsin	586 – 615	828.42	828.38	0.0387	3	R.GGQGAGAAAAAAGGAGQGG ^{125.12} YGGLGGQGVGR.G	Unknown shift / CID	242	83.2
6	P19837	Spidroin-1	Trypsin	616 – 652	954.79	954.78	0.0027	3	R.GGLGGQGAGAAAAAGGAGQGGYGGVGSASAASAAASR.L	CID + ETD	534	100
6	P19837	Spidroin-1	Trypsin	653 – 661	523.25	523.24	0.0066	2	R.LSSPQ ^{113.01} ASSR.V	Unknown shift / CID	242	81
6	P19837	Spidroin-1	Trypsin	698 – 710	687.29	687.37	-0.0809	2	G.LSGCDVLIQALLE.V	No enzyme (two ends) / CID	213	96.9
6	P19837	Spidroin-1	Chymotrypsin	1-18	511.29	511.22	0.07	3	QGAGAA ^{126.21} AAAAGGAGQGGY.G	Unknown shift / CID + ETD	341	99.2
6	P19837	Spidroin-1	Chymotrypsin	1-31	860.38	860.37	0.0022	3	Q ^{114.01} GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Unknown shift / CID + ETD	305	100
6	P19837	Spidroin-1	Chymotrypsin	19 - 31; 185 - 197; 286 - 298	539.73	539.74	-0.0192	2	Y.GGLGGQGAGQGGY.G	CID	373	100
6	P19837	Spidroin-1	Chymotrypsin	22 - 34; 188 - 200; 289 - 301	539.67	539.74	-0.0792	2	L.GGQGAGQGGYGGL.G	CID	329	100
6	P19837	Spidroin-1	Chymotrypsin	22 – 59	1049.51	1049.45	0.0593	3	L.GGQGAGQGGY ^{183.18} GGLGGQGAGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	345	99.9
6	P19837	Spidroin-1	Chymotrypsin	32 – 59	1016.94	1016.99	-0.0592	2	Y.GGLGG ^{-98.11} QGAGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID	471	100
6	P19837	Spidroin-1	Chymotrypsin	35 – 59	916.91	916.93	-0.0257	2	L.GGQGAGQGAGA ^{-71.04} AAAAAAGGAGQGGY.G	Unknown shift / CID	577	100
6	P19837	Spidroin-1	Chymotrypsin	60 - 89; 576 - 605	768.34	768.37	-0.0344	3	Y.GGLGSQAGRGGQGAGA ^{-57.09} AAAAAGGAGQGGY.G	Unknown shift / CID + ETD	554	100
6	P19837	Spidroin-1	Chymotrypsin	60 - 92; 576 - 608	1036.14	1036.08	0.0566	3	Y.GGLGSQAGRGGQGAGAAAAAAGGAGQGGYGG ^{519.17} L.G	Unknown shift / CID + ETD	504	100
6	P19837	Spidroin-1	Chymotrypsin	63 – 102	785.88	785.88	-0.0077	4	L.GSQGAGRGGQGAGAAAAA ^{-60.02} AGGAGQGGYGGLGSQAGRGG.L	Unknown shift / ETD	333	82.3
6	P19837	Spidroin-1	Chymotrypsin	63 - 89; 579 - 605	1031.49	1031.49	-0.0044	2	L.GSQGAGRGGQGAGA ⁻⁷¹ AAAAAGGAGQGGY.G	Unknown shift / CID	416	100
6	P19837	Spidroin-1	Chymotrypsin	63 - 92; 579 - 608	960.43	960.37	0.0556	3	L.GSQGAGRGGQGAGAAAAAAGGAGQGGYGGL ^{519.17} .G	Unknown shift / CID + ETD	452	100
6	P19837	Spidroin-1	Chymotrypsin	90 – 102; 152 – 164; 198 – 210; 299 – 311; 354 - 366	543.77	543.78	-0.0162	2	Y.GGLGSQAGRGG.L	CID	331	100
6	P19837	Spidroin-1	Chymotrypsin	90 – 123	863.38	863.42	-0.0491	3	Y.GGLGSQAGRGG ^{-70.14} GGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	493	100
6	P19837	Spidroin-1	Chymotrypsin	93 – 123	811.02	811.05	-0.0334	3	L.GSQAGRGGGLGGQGAGAAAAAAGGAGQGGY.G	CID + ETD	514	100
6	P19837	Spidroin-1	Chymotrypsin	93 – 126	1059.78	1059.76	0.0176	3	L.GSQAGRGGGLGGQGAGAAAAAAGGAGQGGYGG ^{519.05} L.G	Unknown shift / CID + ETD	415	100
6	P19837	Spidroin-1	Chymotrypsin	103 - 123; 431 - 451	903.37	903.36	0.0036	2	L.G ^{215.01} QQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	421	100
6	P19837	Spidroin-1	Chymotrypsin	103 – 151	957.43	957.44	-0.0176	4	L.GGQGAGAAAAAAGG ^{-4.06} AGQGGYGGLGNQGAGRGGQGAAAAAAGGAGQGGY.G	Unknown shift / ETD	492	100
6	P19837	Spidroin-1	Chymotrypsin	124 – 154	863.11	863.06	0.0425	3	Y.GGLGNQ ^{101.13} GAGRGGQAAAAAAGGAGQGGYGGL.G	Unknown shift / CID + ETD	494	100
6	P19837	Spidroin-1	Chymotrypsin	127 – 151	820.39	820.31	0.0738	3	L.GNQGAGRGGQGAA ^{427.22} AAAAGGAGQGGY.G	Unknown shift / CID + ETD	355	99.8

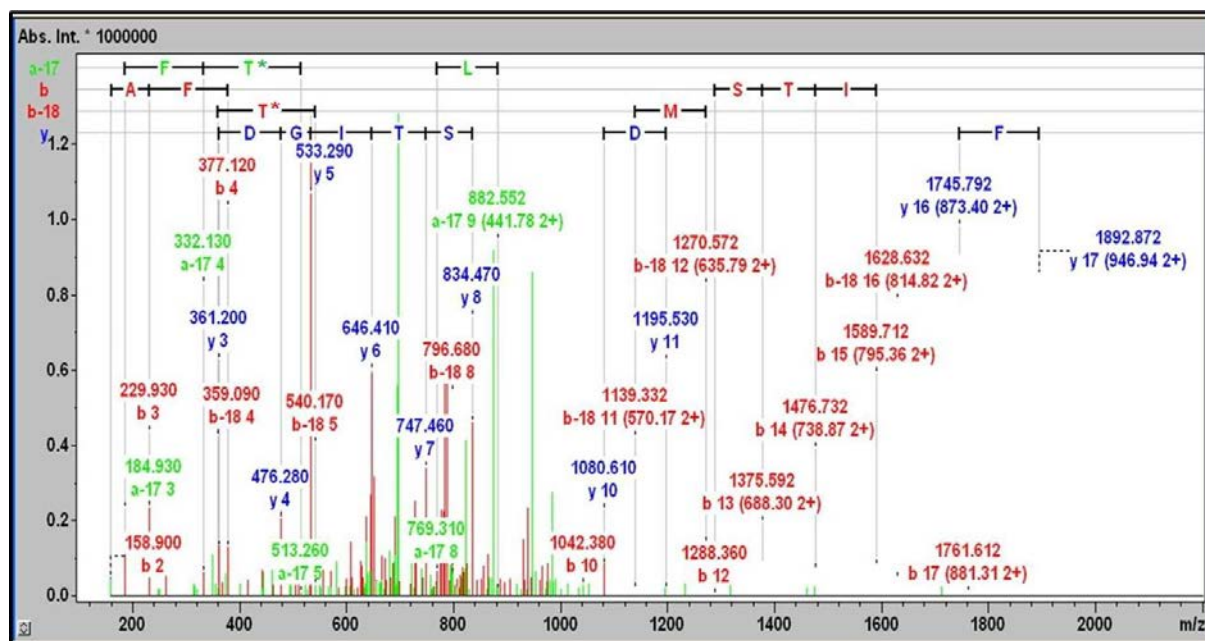
6	P19837	Spidroin-1	Chymotrypsin	152 – 184	839.7	839.75	-0.0501	3	Y.GGLGSQGAGRGGLGGQGAGAA ^{-70.14} AAAAGGAGQGGY.G	Unknown shift / CID + ETD	589	100
6	P19837	Spidroin-1	Chymotrypsin	155 – 187	1042.49	1042.41	0.0733	3	L.GSQGAGRGGLGGQGAGAAAAAAGGAGQG ^{538.22} GYGGL.G	Unknown shift / CID	292	99.6
6	P19837	Spidroin-1	Chymotrypsin	165 – 184	735.34	735.34	-0.0078	2	L.GGQGAGAAA ^{50.01} AAAGGAGQGGY.G	Unknown shift / CID	292	99
6	P19837	Spidroin-1	Chymotrypsin	165 - 197	1088.47	1088.39	0.0779	3	L.G ^{684.23} GQGAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Unknown shift / CID + ETD	489	100
6	P19837	Spidroin-1	Chymotrypsin	201 - 231	792.38	792.39	-0.0137	3	L.GSQGAG ^{-6.03} RGGLGGQGAGAAAAAAAGGAGQGGL.G	Unknown shift / CID + ETD	215	83.6
6	P19837	Spidroin-1	Chymotrypsin	211 - 255	1092.1	1092.18	-0.0883	3	L.GGQGAGAAAAAAGGAGQGGLGGQGAGQ ^{-96.26} GAGASAAAAGGAGQGGY.G	Unknown shift / CID + ETD	274	99.9
6	P19837	Spidroin-1	Chymotrypsin	232 - 255	1031.99	1031.91	0.0754	2	L.GGQGAGQG ^{214.15} AGASAAAAGGAGQGGY.G	Unknown shift / CID	202	100
6	P19837	Spidroin-1	Chymotrypsin	256 - 285	1180.5	1180.57	-0.0761	2	Y.GGLGSQGAGRGGE ^{-K} GAGAAAAAAGGAGQGGY.G	Substitution Glu->Lys (E268) / CID	379	100
6	P19837	Spidroin-1	Chymotrypsin	256 - 288	837.71	837.74	-0.0347	3	Y.GGLGSQGAGRGGEAGAAAAAAGG ^{-77.09} AGQGGYGGL.G	Unknown shift / CID + ETD	216	90.1
6	P19837	Spidroin-1	Chymotrypsin	259 - 285	1049.9	1049.98	-0.0865	2	L.GSQGAGRGGEAGA ^{-35.16} AAAAAGGAGQGGY.G	Unknown shift / CID	222	100
6	P19837	Spidroin-1	Chymotrypsin	259 - 298	1064.78	1064.83	-0.0553	3	L.GSQGAGRGGE ^{-K} GAGAAAAAAGGAGQGGYGGLGGQGAGQGGY.G	Substitution Glu->Lys (E268) / CID + ETD	504	100
6	P19837	Spidroin-1	Chymotrypsin	302 - 329	747.03	747.02	0.0068	3	L.GSQGAGRGGLGGQGAG ^{71.02} AAAAGGAGQGGL.G	Unknown shift / CID + ETD	515	99.8
6	P19837	Spidroin-1	Chymotrypsin	302 - 353	996.65	996.46	0.1825	4	L.GSQGAGRGGLGGQGAGAAAAGGAGQ ^{Deamid} GGLGGQGAGQGAGAAAAAAGGAGQGGY.G	Deamidation (Q326) / ETD	333	87.9
6	P19837	Spidroin-1	Chymotrypsin	312 - 353	1049.55	1049.48	0.0638	3	L.GGQGAGAAAAGGAGQGGL ^{5.19} GGQGAGQGAGAAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	355	99.9
6	P19837	Spidroin-1	Chymotrypsin	312 - 356	871.72	871.64	0.0718	4	L.G ^{115.29} GQGAGAAAAGGAGQGGLGGQGAGQGAGAAAAAAGGAGQGGYGGL.G	Unknown shift / ETD	312	83.3
6	P19837	Spidroin-1	Chymotrypsin	330 - 366	776.97	776.849	0.121	4	L.GGQGAGQGAGAAAAAAGGAGQGGYG ^{204.48} GLGSQGAGRGGL.G	Unknown shift / ETD	510	88.5
6	P19837	Spidroin-1	Chymotrypsin	354 - 387	1014.48	1014.43	0.0405	3	Y.G ^{355.12} GLGSQGAGRGGLGGQGAGAVAAAAAGGAGQGGY.G	Unknown shift / CID	389	99.1
6	P19837	Spidroin-1	Chymotrypsin	357 - 387	1045.46	1045.39	0.0628	3	L.GSQGAGRGGLGGQGAGAVAAAAAGGA ^{675.19} GQGGY.G	Unknown shift / CID + ETD	314	100
6	P19837	Spidroin-1	Chymotrypsin	367 - 387	1014.45	1014.38	0.068	2	L.GGQGAG ^{409.14} AVAAAAAGGAGQGGY.G	Unknown shift / CID	326	100
6	P19837	Spidroin-1	Chymotrypsin	367 - 390	657.32	657.29	0.0205	3	L.G ^{124.06} GQGAGAVAAAAAGGAGQGGYGGL.G	Unknown shift / CID + ETD	260	94.7
6	P19837	Spidroin-1	Chymotrypsin	388 - 417	793.72	793.73	-0.0143	3	Y.GGLGSQGAGRGGQGAGAAAAAA ^{-80.03} GGAGQRGY.G	Unknown shift / CID + ETD	239	85.9
6	P19837	Spidroin-1	Chymotrypsin	388 - 420	1111.8	1111.77	0.0234	3	Y.GGLGSQGAGRGGQGAGAAAAAAGGAGQR ^{647.07} GYGGL.G	Unknown shift / CID	434	100
6	P19837	Spidroin-1	Chymotrypsin	391 - 417	1066.97	1066.99	-0.0244	2	L.GSQGAGRGGQGAGAAAAAAGGAGQR ^{-G} GY.G	Substitution Arg->Gly (R415) / CID	446	100
6	P19837	Spidroin-1	Chymotrypsin	418 - 430	557.3	557.29	0.0084	2	Y.GGLGNQGAGRGGL.G	CID	323	100
6	P19837	Spidroin-1	Chymotrypsin	418 - 451	1032.18	1032.09	0.0806	3	Y.GGLGNQGAGRGGLGGQGAG ^{409.24} AAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	288	90.7
6	P19837	Spidroin-1	Chymotrypsin	421 - 430	518.26	518.22	0.0319	2	L.GNQGAGRGG ^{149.06} L.G	Unknown shift / CID	232	90.7
6	P19837	Spidroin-1	Chymotrypsin	421 - 451	1229.99	1230.07	-0.084	2	L.GNQGAGRGGLGGQ ^{Deamid} GAGAAAAAAGGAGQGGY.G	Deamidation (Q433) / CID	268	100
6	P19837	Spidroin-1	Chymotrypsin	421 - 454	962.49	962.43	0.0573	3	L.GNQGAGRGGLGGQGAGAAAA ^{200.17} AAAGGAGQGGYGGL.G	Unknown shift / CID + ETD	258	82.7
6	P19837	Spidroin-1	Chymotrypsin	452 - 478	805.71	805.67	0.0305	3	Y.GGLGNQGAG ^{227.09} RGGQAAAAAGGAGQGGY.G	Unknown shift / CID + ETD	507	100
6	P19837	Spidroin-1	Chymotrypsin	455 - 478	969.89	969.95	-0.0621	2	L.GNQGAGRGGQ ^{22.11} GAAAAAGGAGQGGY.G	Unknown shift / CID	268	100
6	P19837	Spidroin-1	Chymotrypsin	479 - 521	1220.66	1220.58	0.0717	3	Y.GGLGSQGAGRGGQGAGAAAAAAVGAGQE ^{77.21} IRGQGAGQGGYGGL.G	Unknown shift / CID + ETD	291	87.4
6	P19837	Spidroin-1	Chymotrypsin	522 - 531	444.25	444.22	0.0299	2	L.GSQGS ^{12.06} GRGGL.G	Unknown shift / CID	221	98.9
6	P19837	Spidroin-1	Chymotrypsin	522 - 551	794.4	794.37	0.0204	3	L.GSQSGRGGLGGQG ^{55.06} AGAAAAAGGAGQGGL.G	Unknown shift / CID + ETD	487	100
6	P19837	Spidroin-1	Chymotrypsin	532 - 575	877.23	877.16	0.0672	4	L.GGQGAGAAAAAAGGAGQGGLGGQGAGQGAGAAAAAA ^{95.27} GGVRQGGY.G	Unknown shift / ETD	285	99
6	P19837	Spidroin-1	Chymotrypsin	579 - 618	1142.86	1142.85	0.0053	3	L.GSQGAGRGGQG ^{228.02} AGAAAAAGGAGQGGYGGLGGQGVGRRGL.G	Unknown shift / CID + ETD	241	94.9
6	P19837	Spidroin-1	Chymotrypsin	609 - 636	1123.45	1123.53	-0.0865	2	L.GGQG ^{-S} VGRGGLGGQGAGAAAAGGAGQGGY.G	Substitution Gly->Ser (G612) / CID	380	100
6	P19837	Spidroin-1	Chymotrypsin	619 - 653	1075.9	1075.77	0.127	3	L.G ^{477.38} GQGAGAAAAGGAGQGGYGGVGSASAASAAASRL.S	Unknown shift / CID + ETD	282	93.6
6	P19837	Spidroin-1	Chymotrypsin	637 - 653	695.32	695.35	-0.0377	2	Y.GGVGSASAASAAASRL.S	CID	423	100
6	P19837	Spidroin-1	Chymotrypsin	654 - 680	1265.08	1265.14	-0.0688	2	L.SSPQASSRVSS ^{-A} AVSNLVASGPTNSAAL.S	Substitution Ser->Ala (S664) / CID	266	100
6	P19837	Spidroin-1	Chymotrypsin	670 - 680	494.23	494.25	-0.0289	2	L.VASGPTNSAAL.S	CID	343	100
6	P19837	Spidroin-1	Chymotrypsin	709 - 715	872.5	872.43	0.0655	1	L.L ^{142.07} EVVSAL.I	Unknown shift / CID	289	81.1
6	P19837	Spidroin-1	Chymotrypsin	716 - 729	843.95	843.89	0.0539	2	L.I ^{208.11} QILGSSSIGQVNY.G	Unknown shift / CID	234	83.8
6	P19837	Spidroin-1	Chymotrypsin	730 - 744	718.35	718.36	-0.0124	2	Y.GSAGQAT ^{-A} QIVGQSVY.Q	Substitution Thr->Ala (T736) / CID	348	100

APÊNDICES B
Material suplementar
Figuras S1-S63

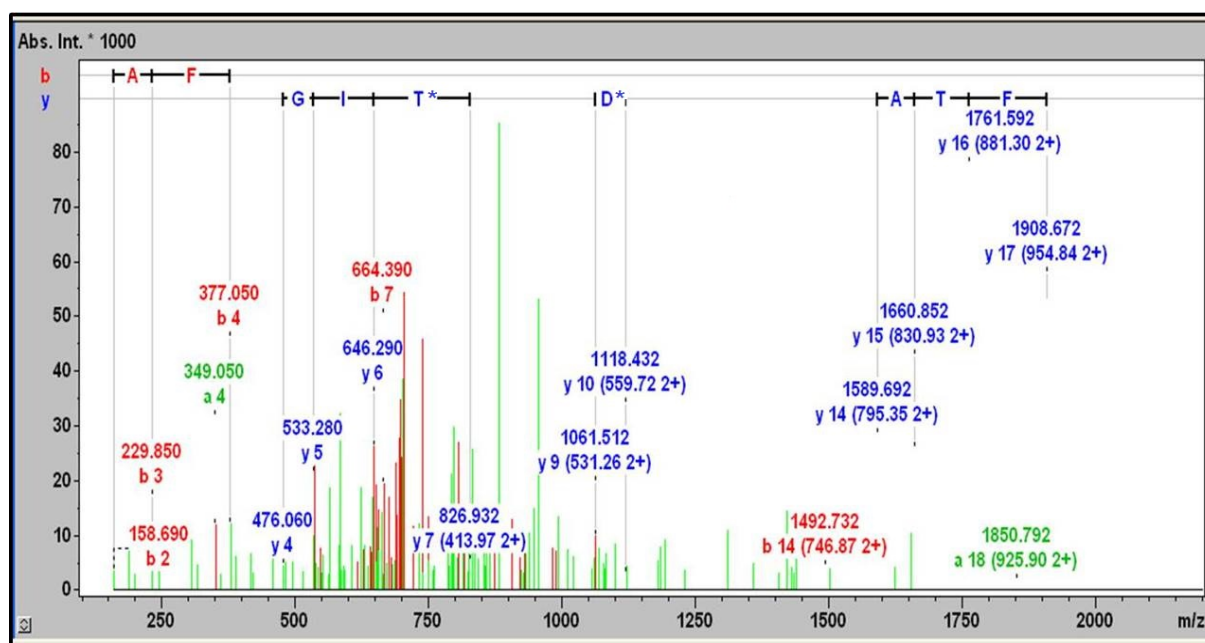
Figuras dos espectros representativos das modificações pós-traducionais observadas na proteína espidroína-1 que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda da aranha *N. clavipes*.



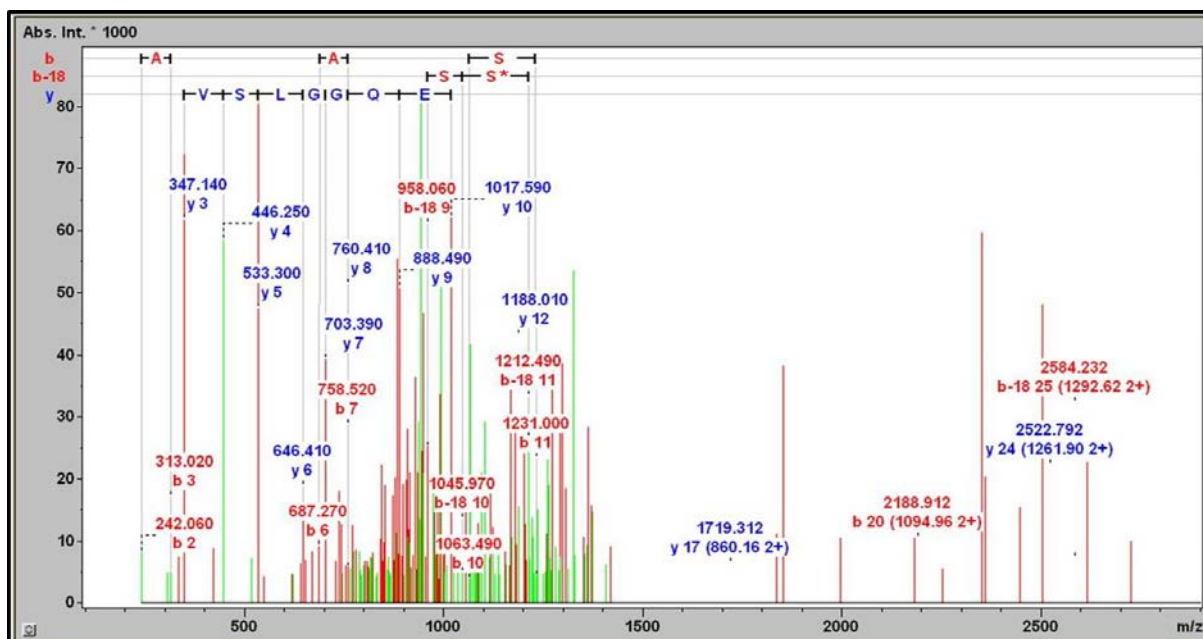
S1. (A) - Espectro CID da sequência peptídica R.T*GAFTAD*QLDDMSTIGDTLK.T (49-68), selecionando-se o íon de m/z 1061.53 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação T49 e o aminoácido substituído Asp→Gly (D55) observados na proteína espidroína-1 (*spots* 3 e 5), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro ETD adquirido para a mesma proteína.



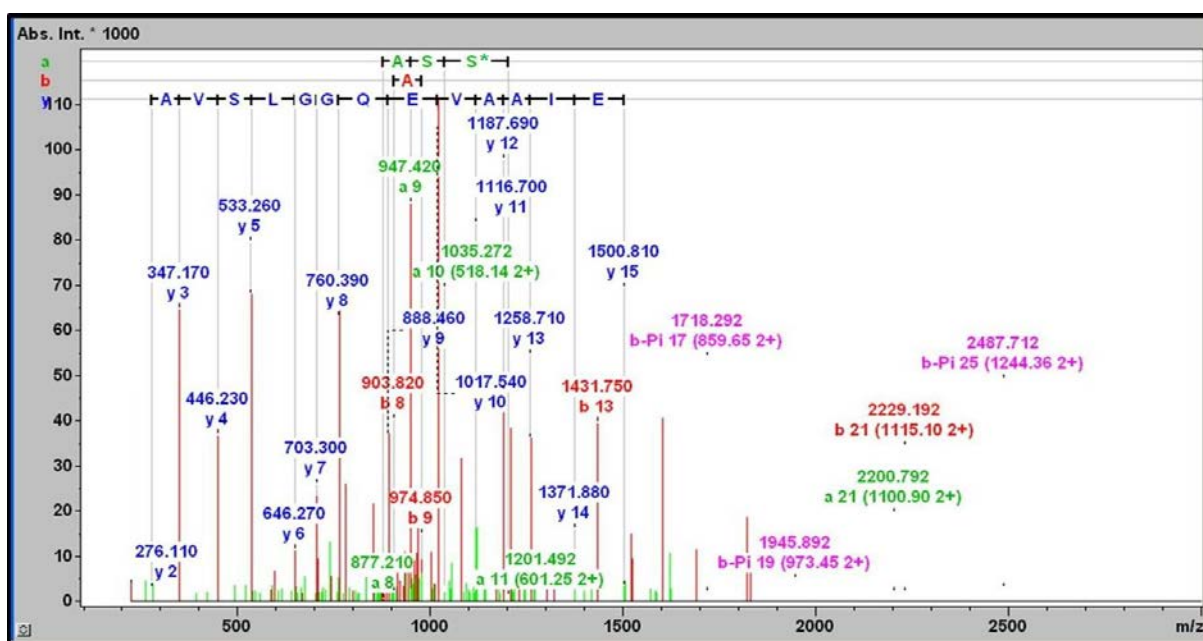
S2. Espectro CID da sequência peptídica R.TGAFT*ADQLDDMSTIGDTIK.T (49-68), selecionando-se o íon de m/z 707.99 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação T53 observado na proteína espidroína-1 (spot 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



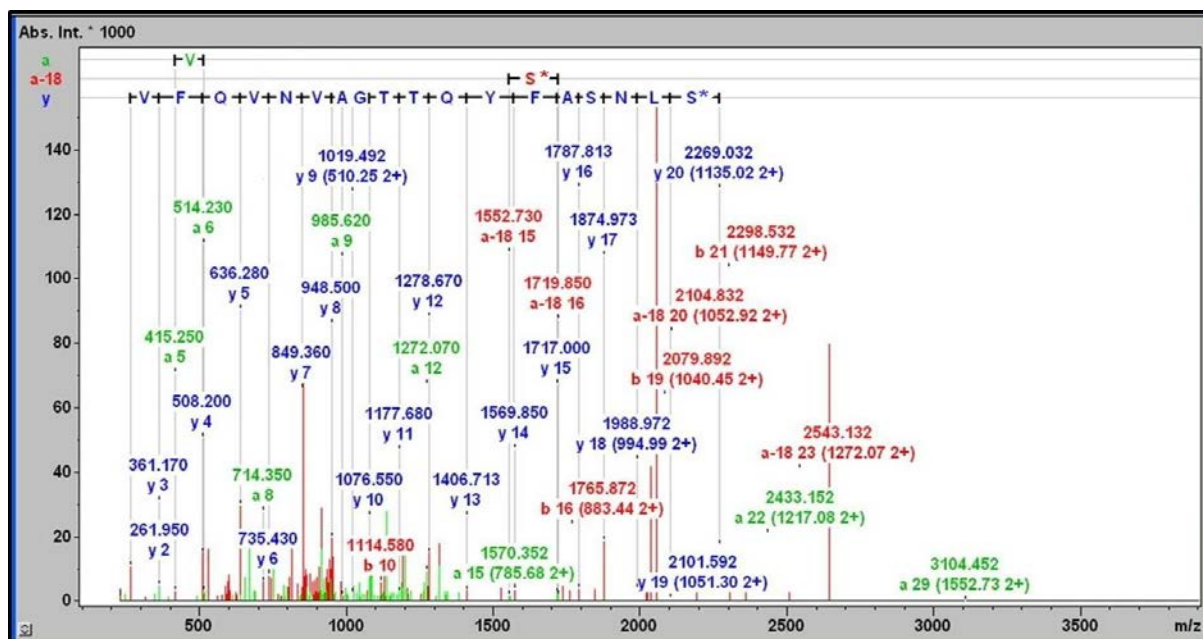
S3. Espectro CID da sequência peptídica R.TGAFTADQLDD*MST*IGDTLK.T (49-68), selecionando-se o íon de m/z 713.34 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação T62 e o aminoácido substituído Asp->Gly (D59) observados na proteína espidroína-1 (spot 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



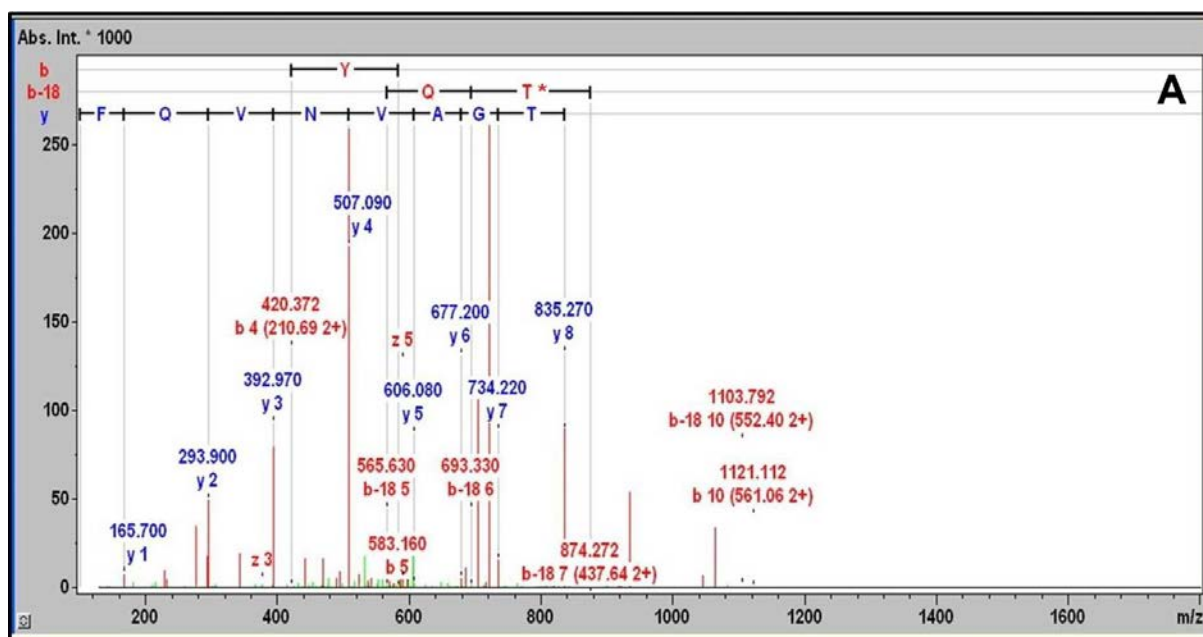
S4. Espectro CID da sequência peptídica K.LQALNMAFAS*MAEIAAVEQGGLSVAEK.T (85-112), selecionando-se o íon de m/z 964.411 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação S9 observado na proteína espiroína-1 (spots 3 e 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

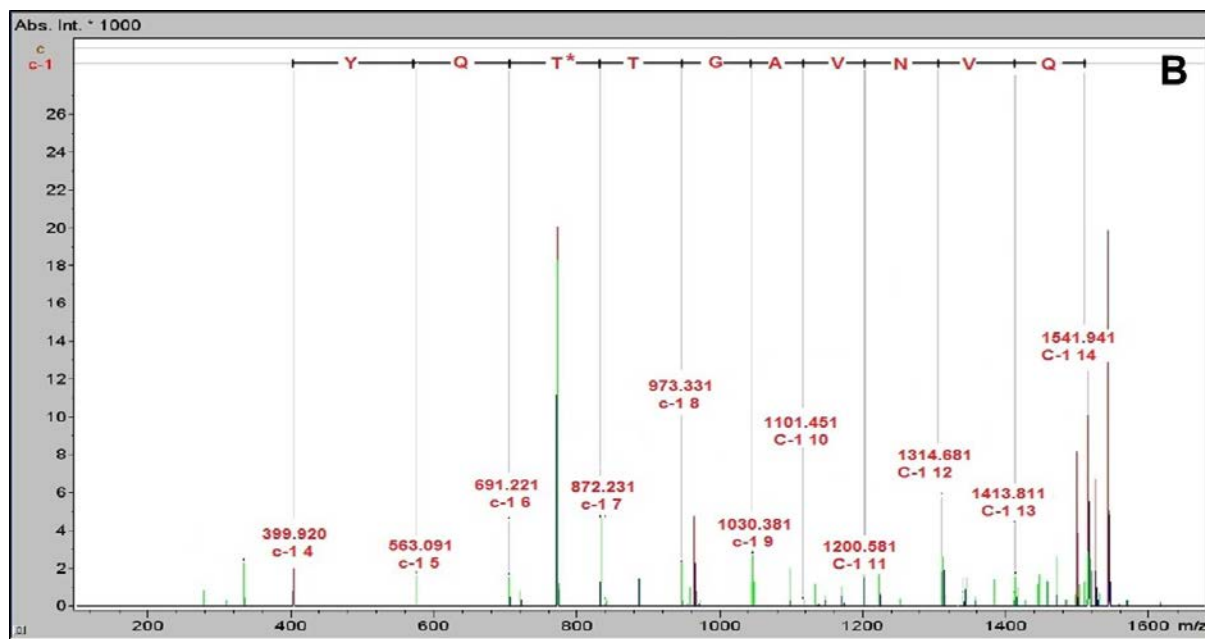


S5. Espectro CID da sequência peptídica K.LQALNMAFAS*MAEIAAVEQGGLSVAEK.T (85-112), selecionando-se o íon de m/z 982.50 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação S9 observado na proteína espiroína-1 (spot 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

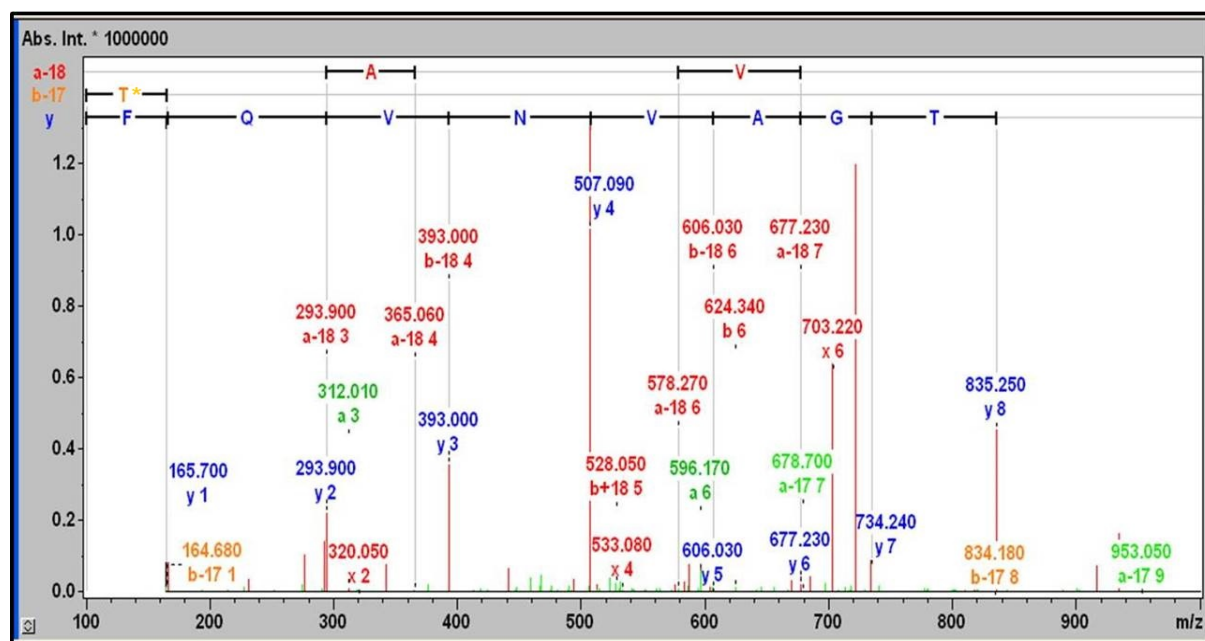


S6. Espectro CID da sequência peptídica E.QGGLSVAEKTNAIAD S* LNSAFYQTTGAVNVQFVNE.I (104-138), selecionando-se o íon de m/z 1288.546 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação S119 observado na proteína espidroína-1 (spot 5), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

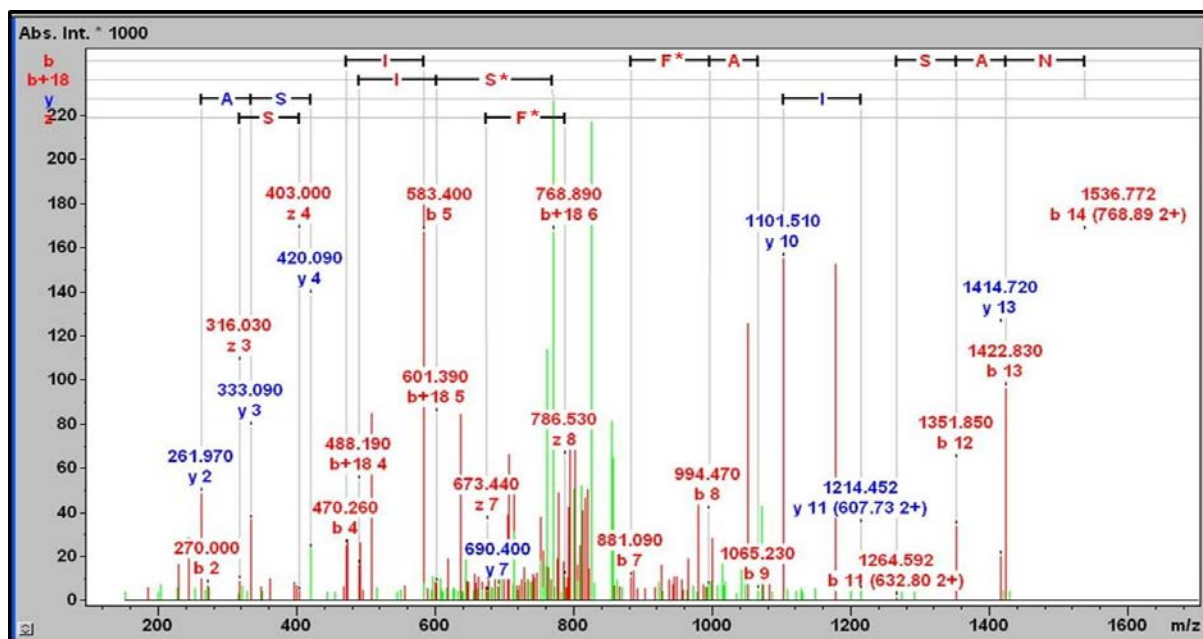




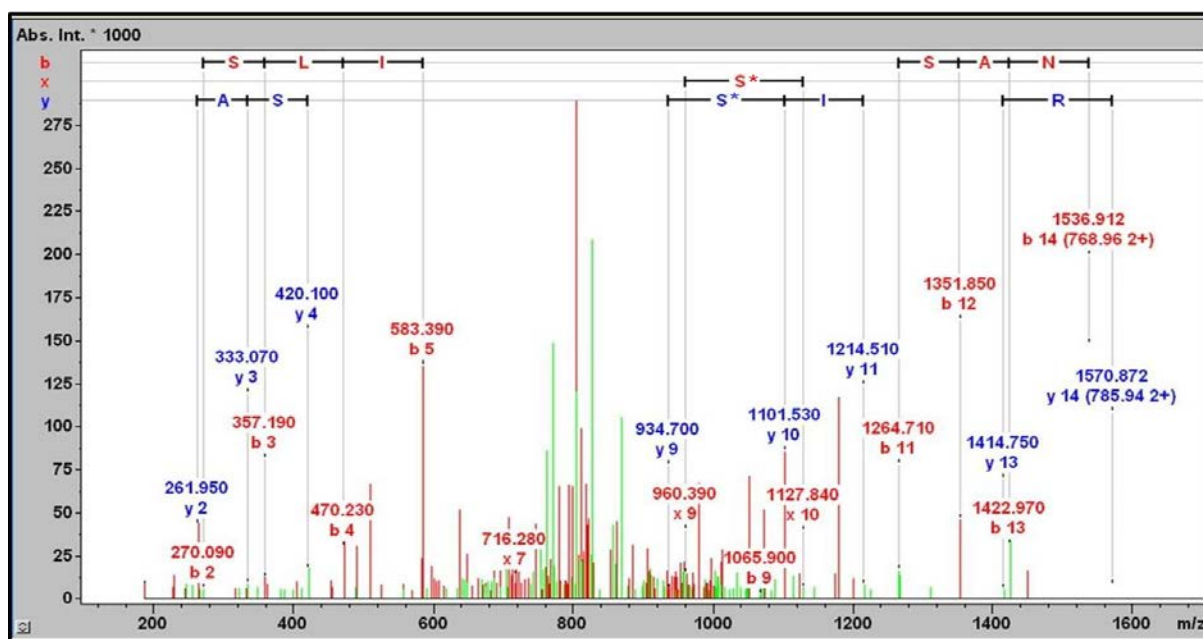
S7. (A) - Espectro CID da sequência peptídica L.NSAFYQ**T***TGAVNVQF.V (121-135), selecionando-se o íon de m/z 614.55 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação T127 observado na proteína espidroína-1 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro ETD adquirido para a mesma proteína.



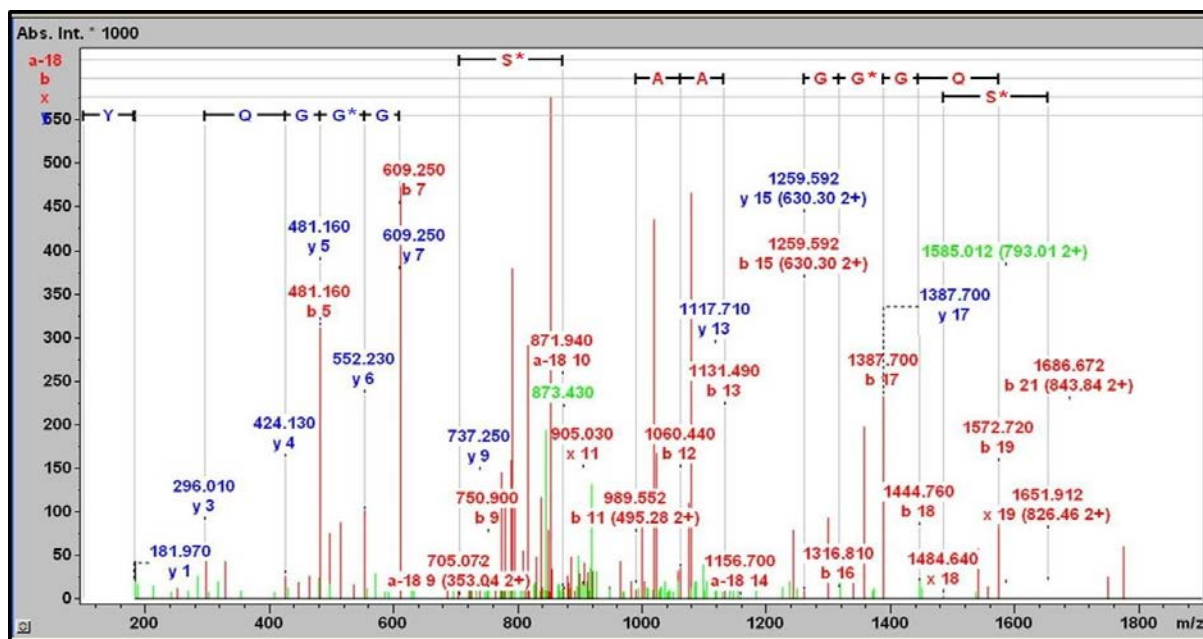
S8. Espectro CID da sequência peptídica Q.**T***TGAVNVQF.V (127-135), selecionando-se o íon de m/z 614.28 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação T127 observado na proteína espidroína-1 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



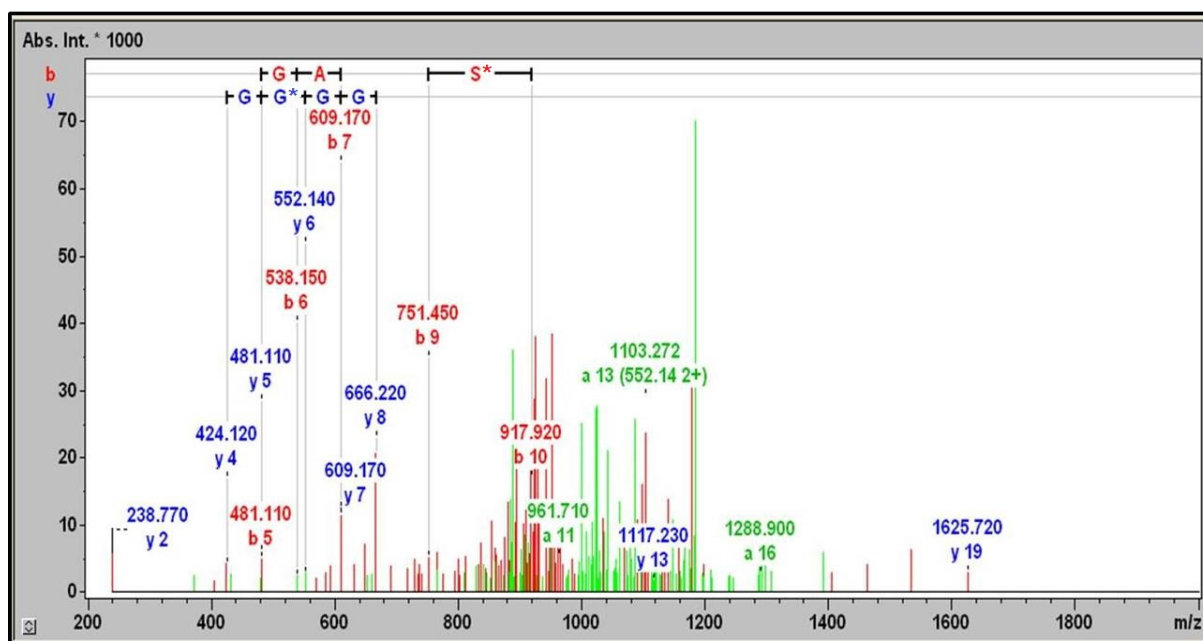
S9. Espectro CID da sequência peptídica E.IRSLIS*MF* AQASANE.V (139-153), selecionando-se o íon de m/z 842.436 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação S144 e o aminoácido substituído Phe→Pro (F146) observados na proteína espidroína-1 (spots 3 e 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



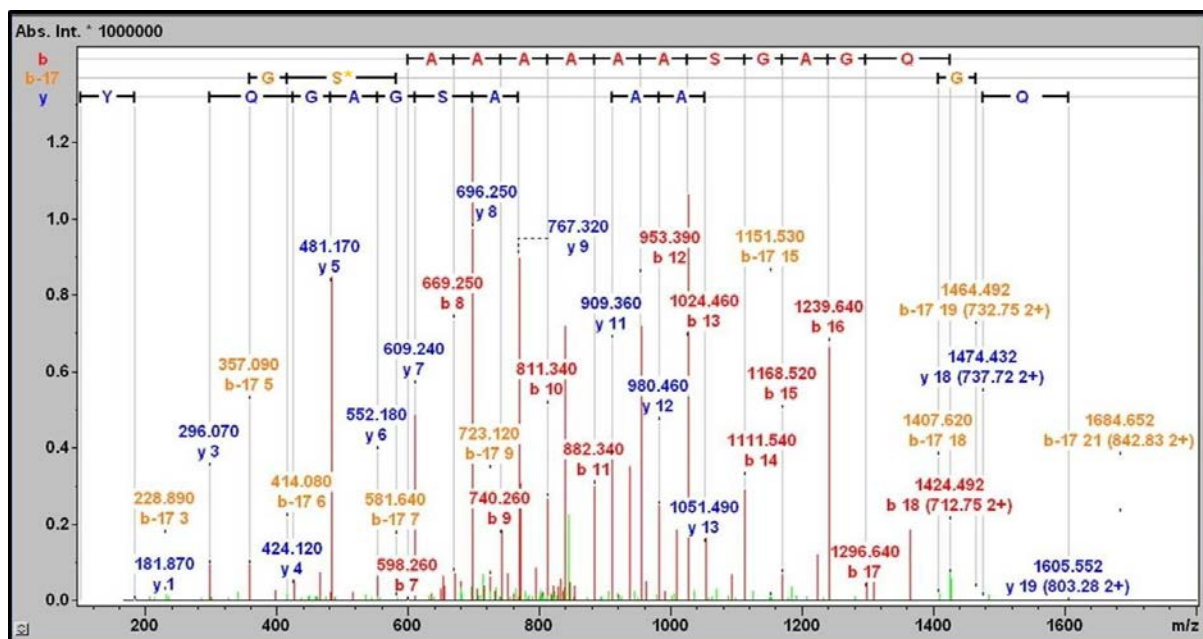
S10. Espectro CID da sequência peptídica E.IRSLIS*MFAQASANE.V (139-153), selecionando-se o íon de m/z 842.447 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação S144 observado na proteína espidroína-1 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



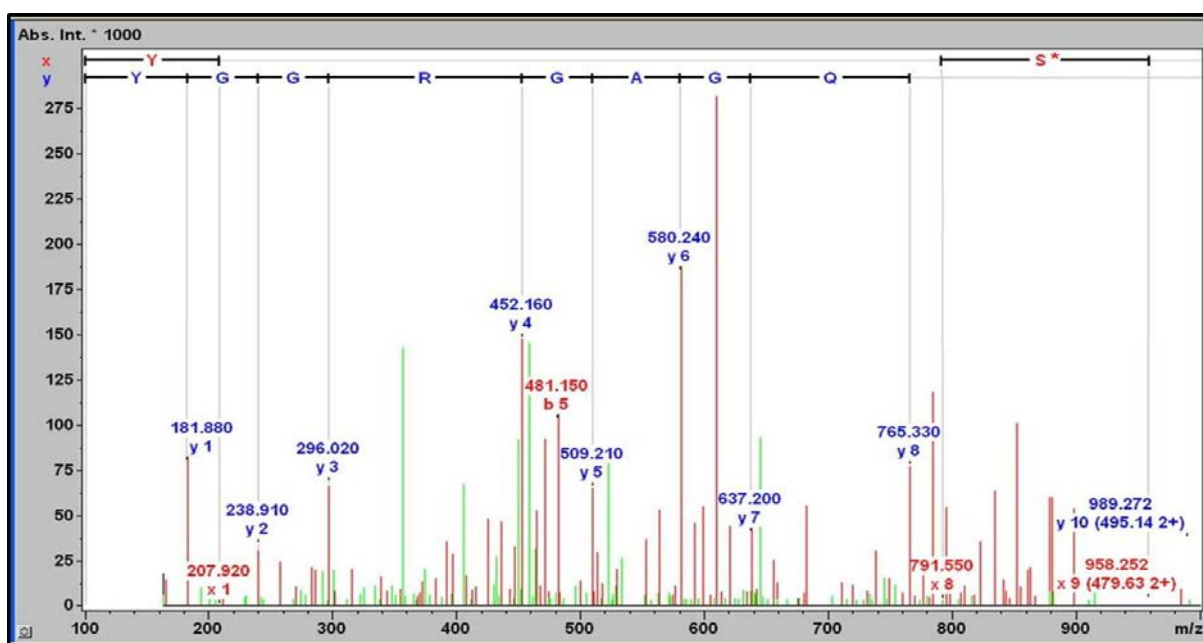
S11. Espectro CID da sequência peptídica Y.GGQS*AGAAAS*AAAAGG*GQGGY.G (161-182), selecionando-se o íon de m/z 934.403 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando os sítios de fosforilação S164 e S170 e o aminoácido substituído Gly→Ala (G177) observados na proteína espidroína-1 (spot 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



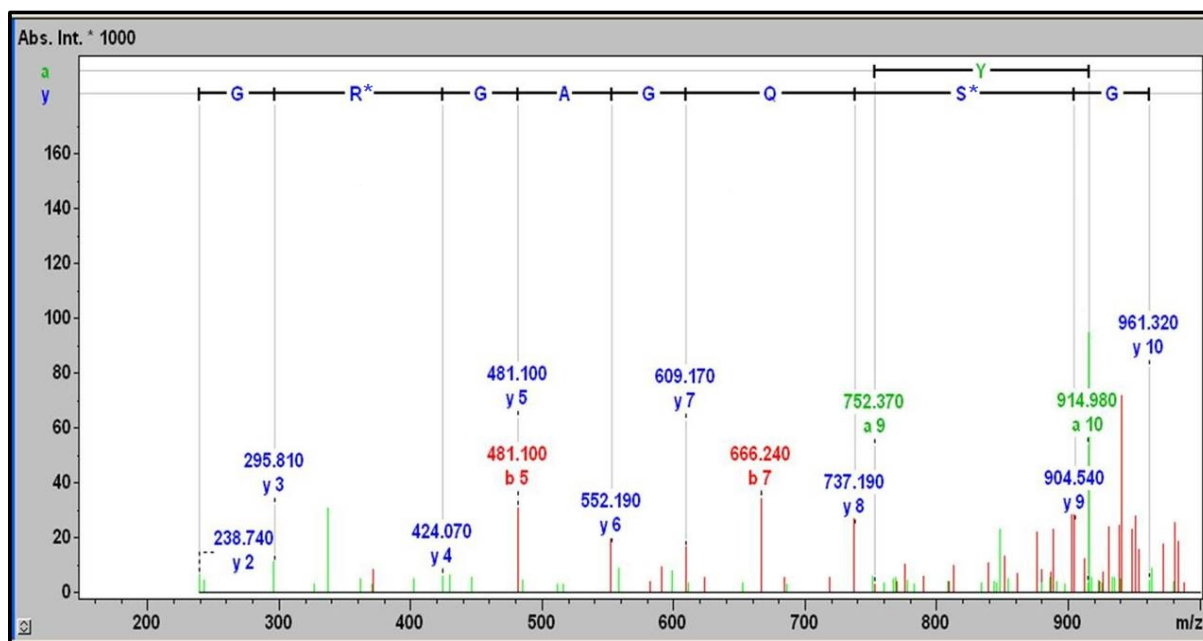
S12. Espectro CID da sequência peptídica Y.GGQS*AGAAAS*AAAAGG*GQGGY.G (161-182), selecionando-se o íon de m/z 934.27 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando os sítios de fosforilação S164 e S170 e o aminoácido substituído Gly→Ala (G177) observados na proteína espidroína-1 (spot 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



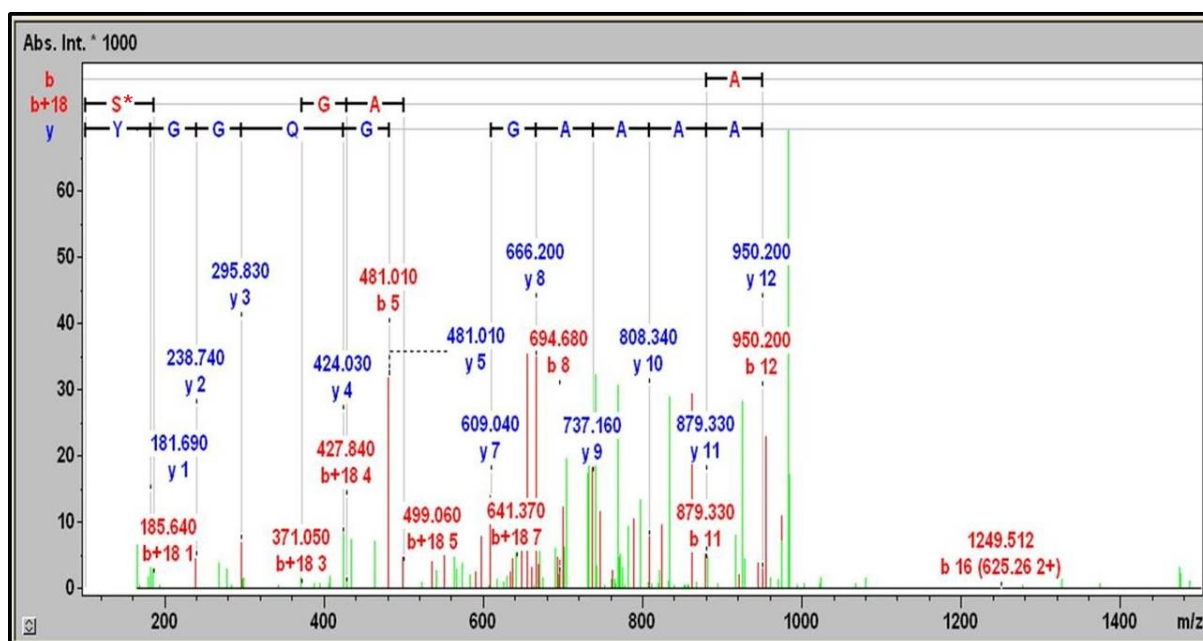
S13. Espectro CID da sequência peptídica L.GGQGAGS*AAAAAASGAGGGY.G (189-209), selecionando-se o íon de m/z 860.342 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação S195 observado na proteína espidroína-1 (spot 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



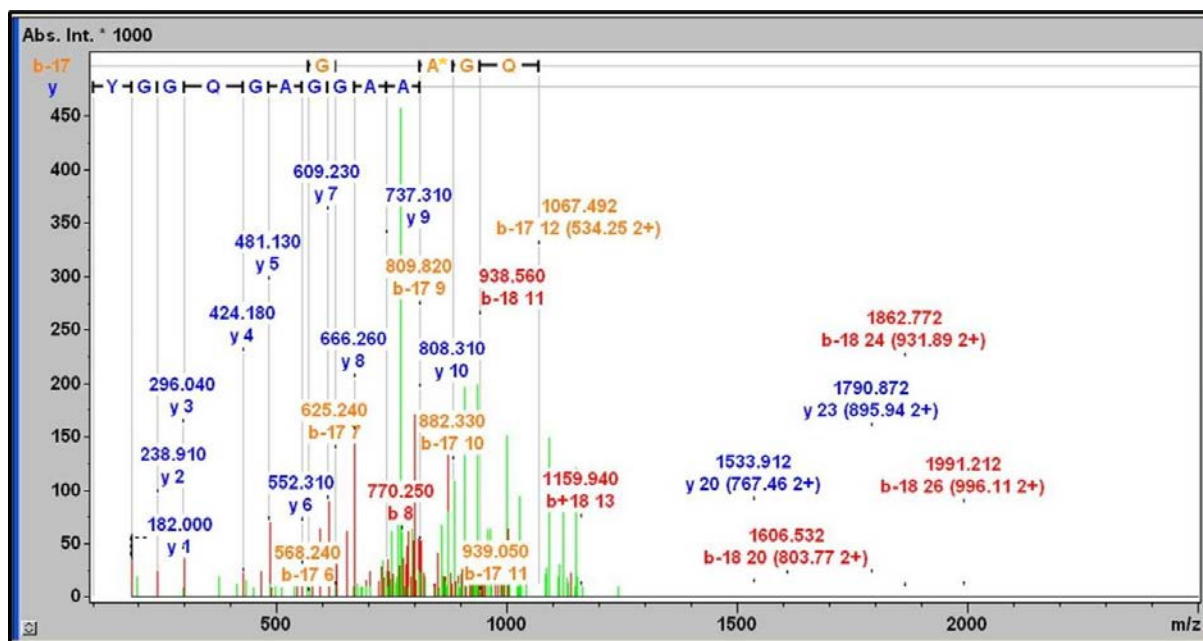
S14. Espectro CID da sequência peptídica L.GS*QGAGRGGY.G (210-219), selecionando-se o íon de m/z 539.173 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação S211 observado na proteína espidroína-1 (spot 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



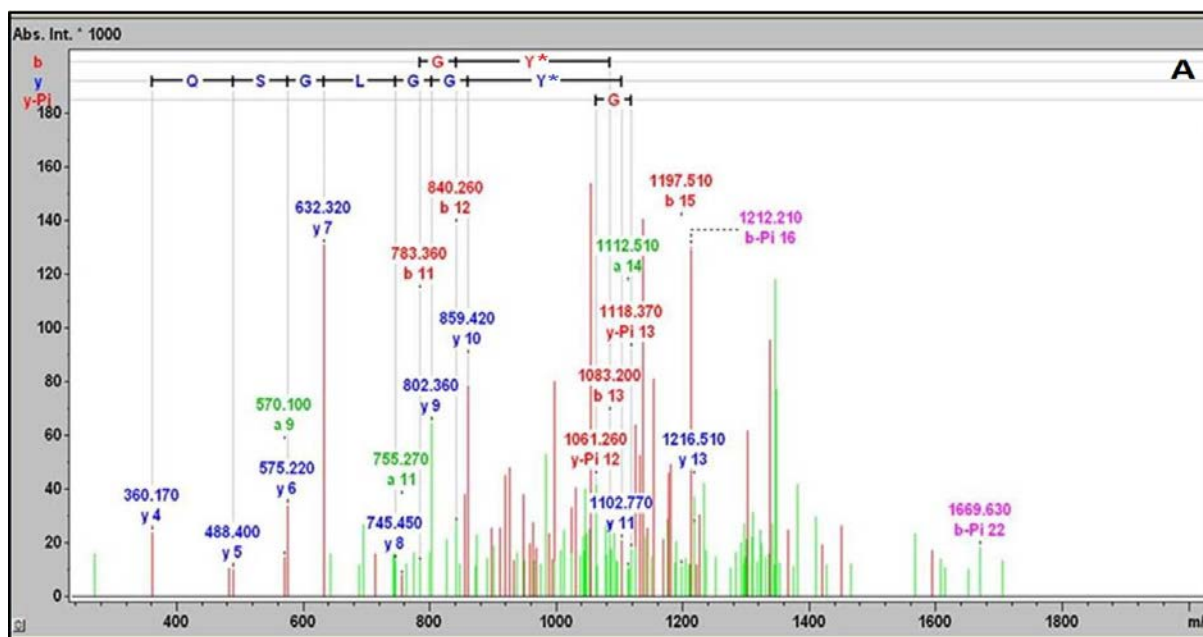
S15. Espectro CID da sequência peptídica L.GS*QGAGR*GGY.G (210-219), selecionando-se o íon de m/z 961.35 na forma de $[M + H]^+$, como precursor; e mostrando o sítio de fosforilação S211 e o aminoácido substituído Arg→Lys (R216) observado na proteína espidroína-1 (spot 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

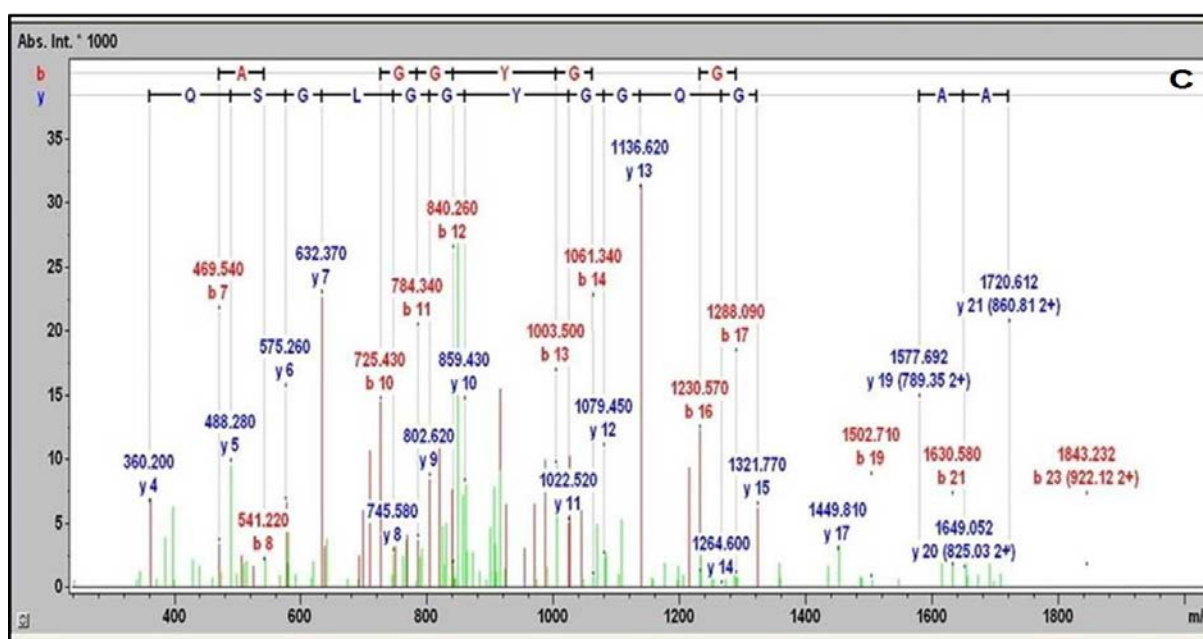
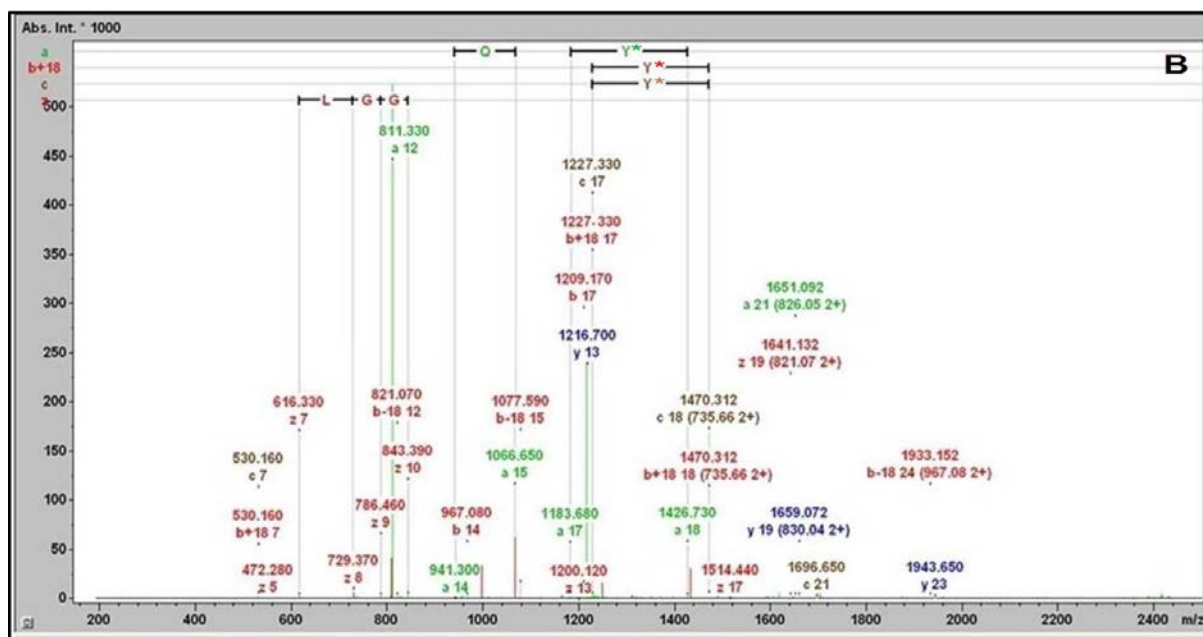


S16. Espectro CID da sequência peptídica L.S*QGAAAAAGGAGQGGY.G (244-260), selecionando-se o íon de m/z 715.710 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação S244 observado na proteína espidroína-1 (spot 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

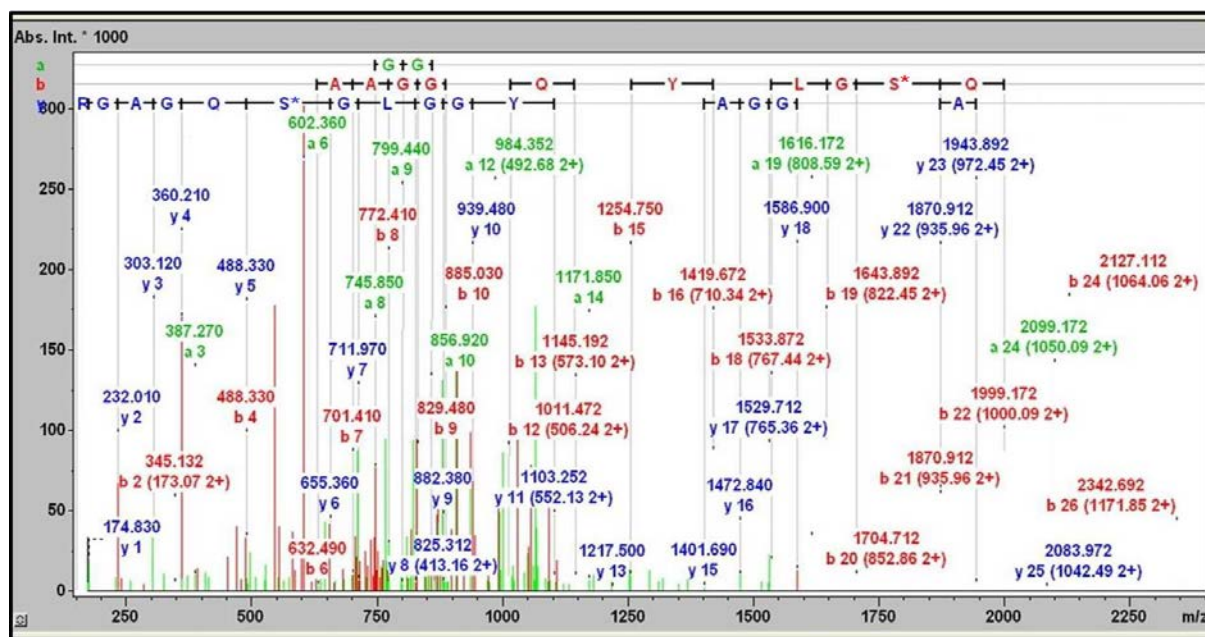


S17. Espectro CID da sequência peptídica GY*GGLGGQGA*GQGAGAAAAAAGGAGQGGY.G (30-59), selecionando-se o íon de m/z 825.360 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação Y31 e o aminoácido substituído Ala→Asn (A39) observados na proteína espidroína-1 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

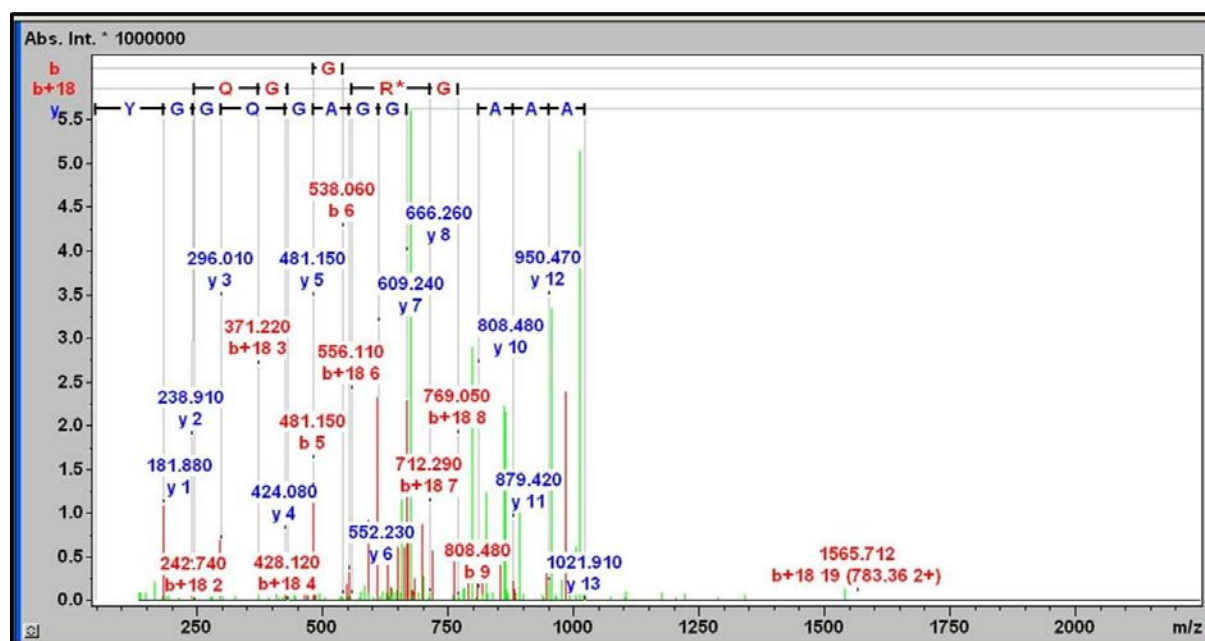




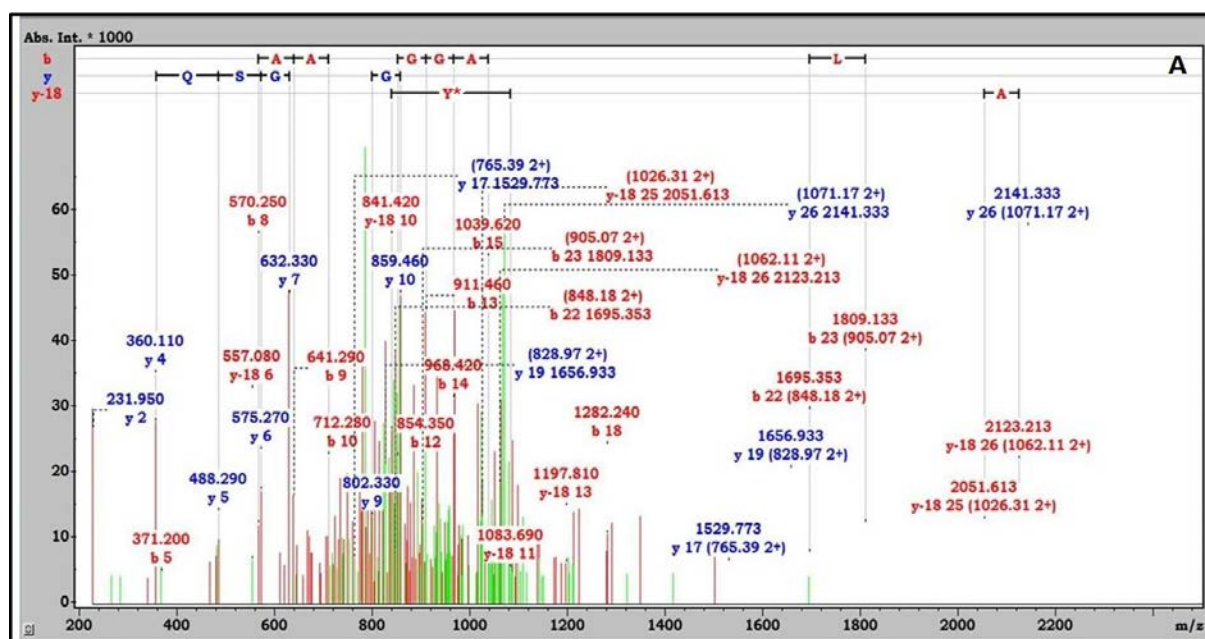
S18. (A) - Espectro CID da sequência peptídica AAAAAGGAGQGGY*GGLGSQGAGR (47-69, 139-161, 341-363, 375-397, 466-488), selecionando-se o íon de m/z 1067.439 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando os sítios de fosforilação Y59, Y151, Y353, Y387 e Y478 observados na proteína espidroína-1 (*spots* 2 e 6), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro ETD adquirido para a mesma proteína. **(C)** - Espectro CID da sequência peptídica AAAAAGGAGQGGY*GGLGSQGAGR (47-69, 139-161, 341-363, 375-397, 466-488), selecionando-se o íon de m/z 1027.490 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando a desfosforilação dos sítios Y59, Y151, Y353, Y387 e Y478, que foram observados na proteína espidroína-1 (*spots* 2 e 6), após o tratamento com fosfatase alcalina.

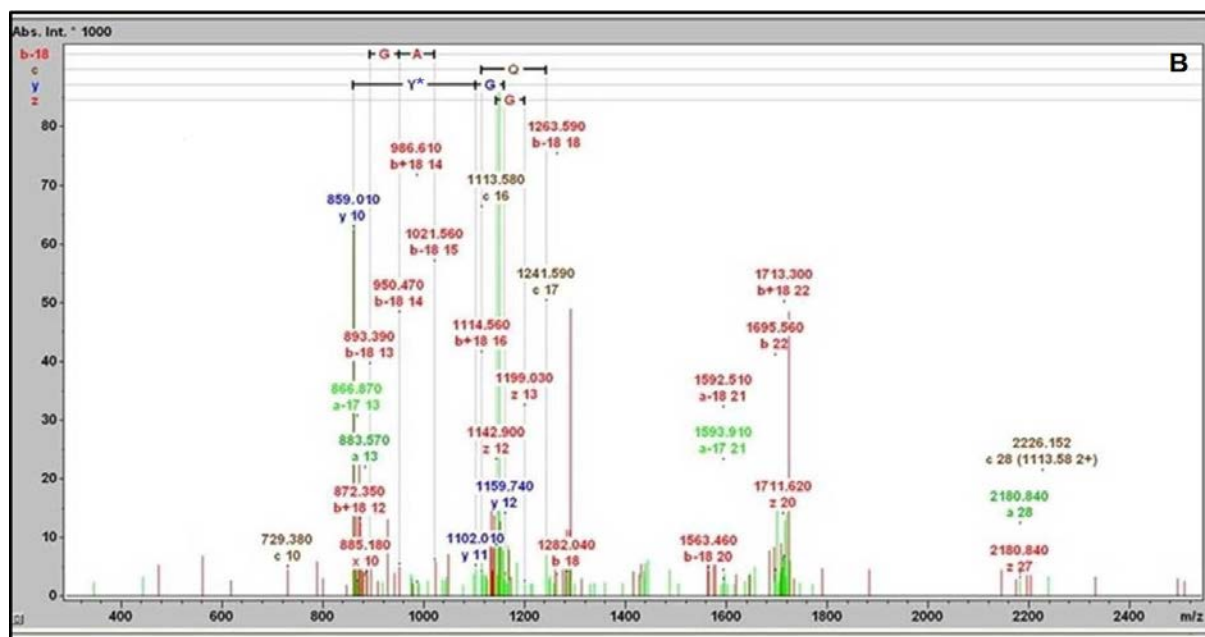


S19. Espectro CID da sequência peptídica AAAAAAGGAGQGGYGGGLGS*QGAGR.G (45-69), selecionando-se o íon de m/z 738.059 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação S64 observado na proteína espidroína-1 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

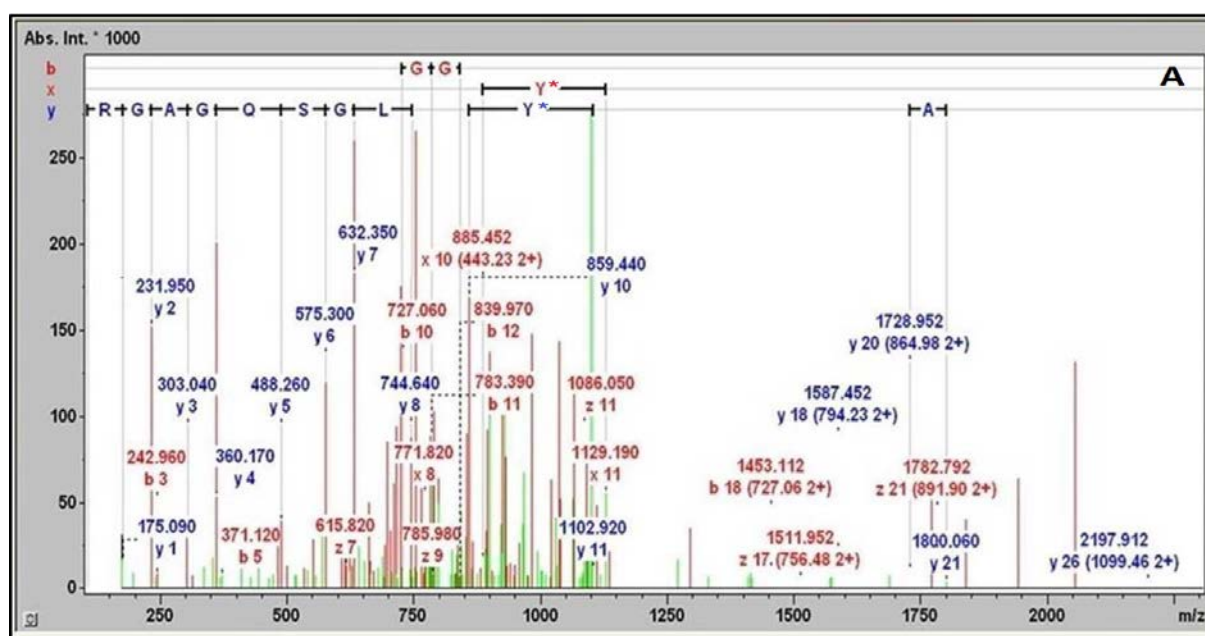


S20. Espectro CID da sequência peptídica L.GS*QGAGR*GGQGAGAAAAAGGAGQGGY.G (63-89, 579-605), selecionando-se o íon de m/z 735.309 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando os sítios de fosforilação S64 e S580 e o aminoácido substituído Arg->Phe (R69, R585) observados na proteína espidroína-1 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.





S22. (A) - Espectro CID da sequência peptídica GGQGAGAAAAAGGAGQGGY*GGLGSQGAGR (70-99), selecionando-se o íon de m/z 861.060 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação Y89 observado na proteína espidroína-1 (spot 1), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro ETD adquirido para a mesma proteína.

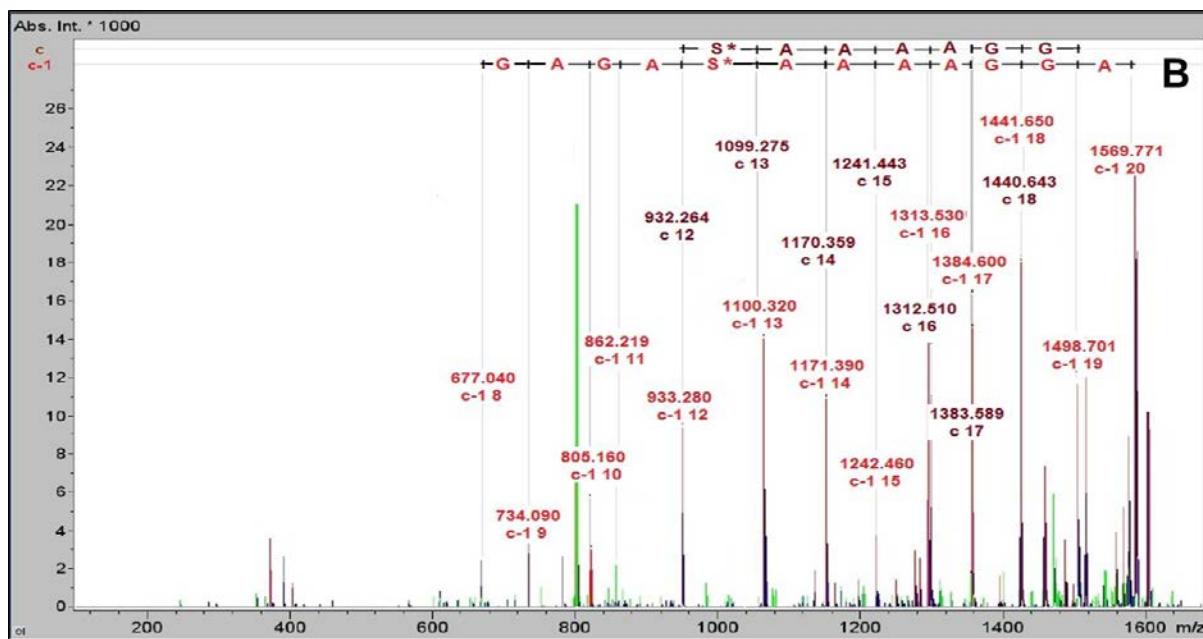


Abs. Int. * 1000

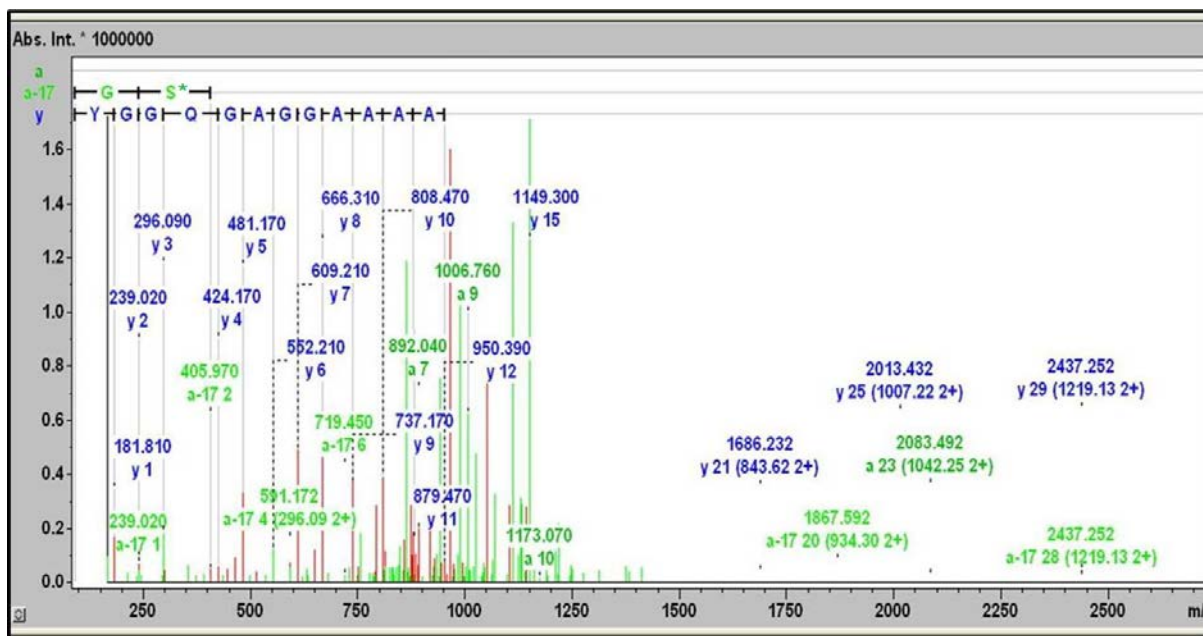
A

Mass spectrum showing relative intensity (Abs. Int. * 1000) versus m/z. The spectrum displays numerous peaks, with major peaks labeled in red (b-ions), green (a-ions), and blue (y-ions). The x-axis ranges from 500 to 3000 m/z, and the y-axis ranges from 0.0 to 27.5. A dashed box highlights the region from m/z 800 to 1100.

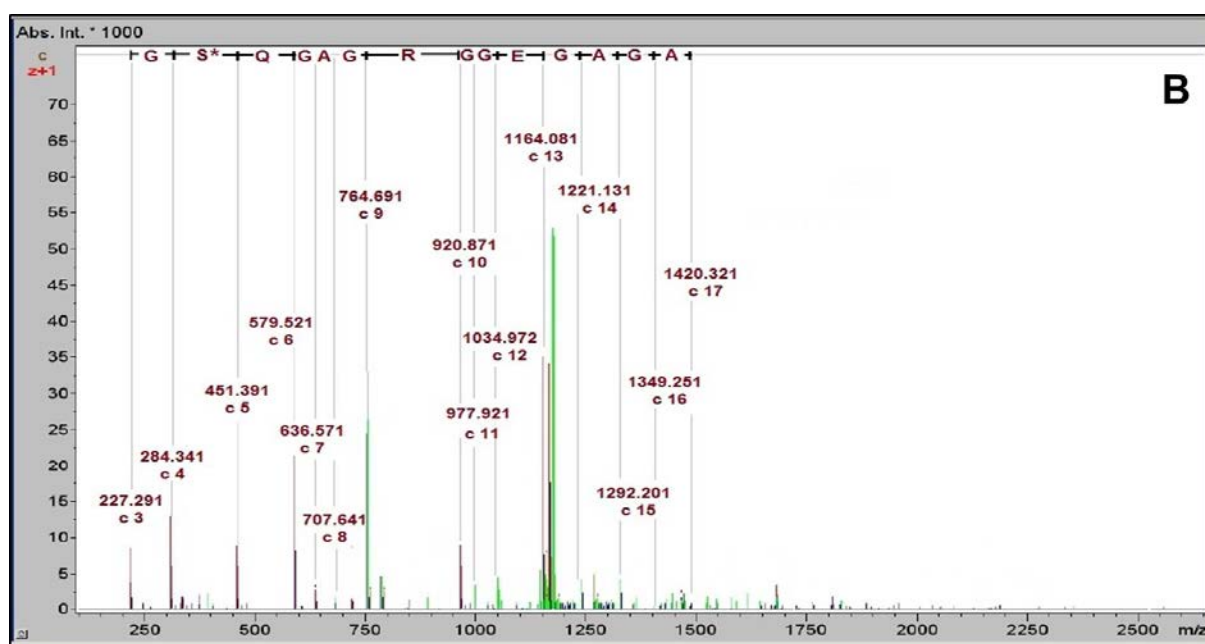
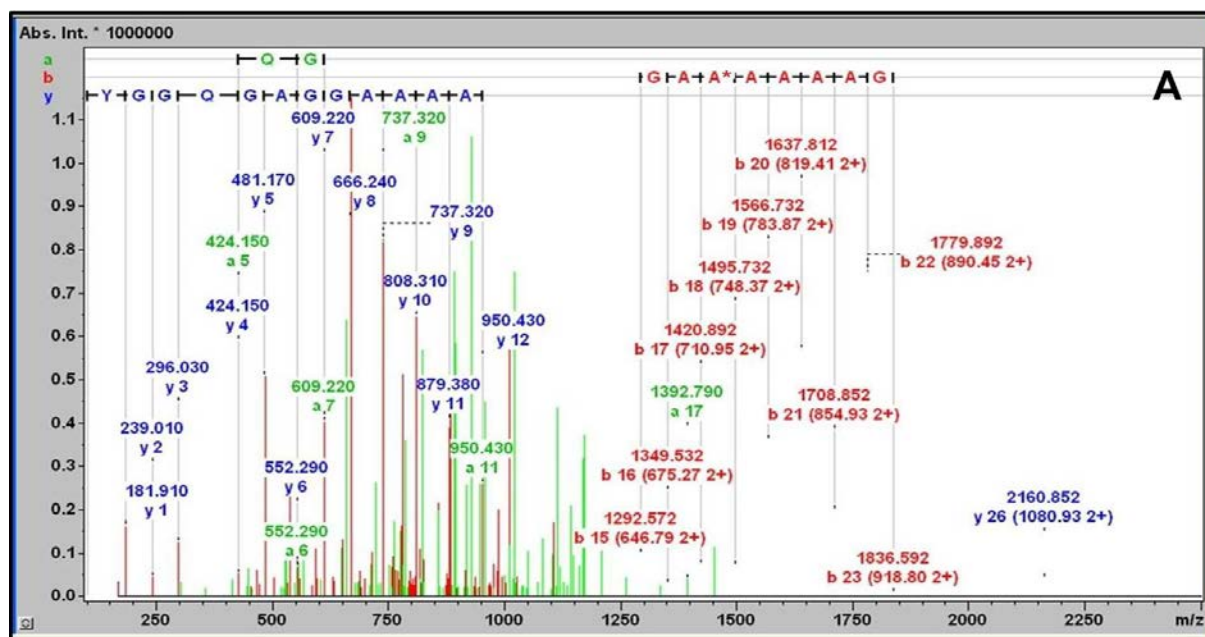
m/z	Ion Type	Relative Intensity (approx.)
242.960	b 3	2.5
371.120	b 5	5.5
481.110	y 5	10.0
499.360	b 7	17.5
552.100	y 6	12.5
570.310	b 8	20.0
613.240	a 9	25.0
641.180	b 9	5.5
684.250	a 10	17.5
712.320	b 10	12.5
737.192	y 9	10.0
773.360	b 11	25.0
783.360	b 11	25.0
808.010	y 10	20.0
826.410	a 12	15.0
897.220	a 13	5.5
925.380	b 13	20.0
950.310	y 12	2.5
982.480	b 14	22.5
1099.500	S*	25.0
1110.590	b 16	5.5
1117.072	y 13	2.5
1139.612	a 17	10.0
1295.760	b 18	17.5
1373.672	y 17	12.5
1499.500	G	25.0
1629.532	y 20	5.5
1687.012	y 21	22.5
1687.012	y 21 (844.01 2+)	22.5



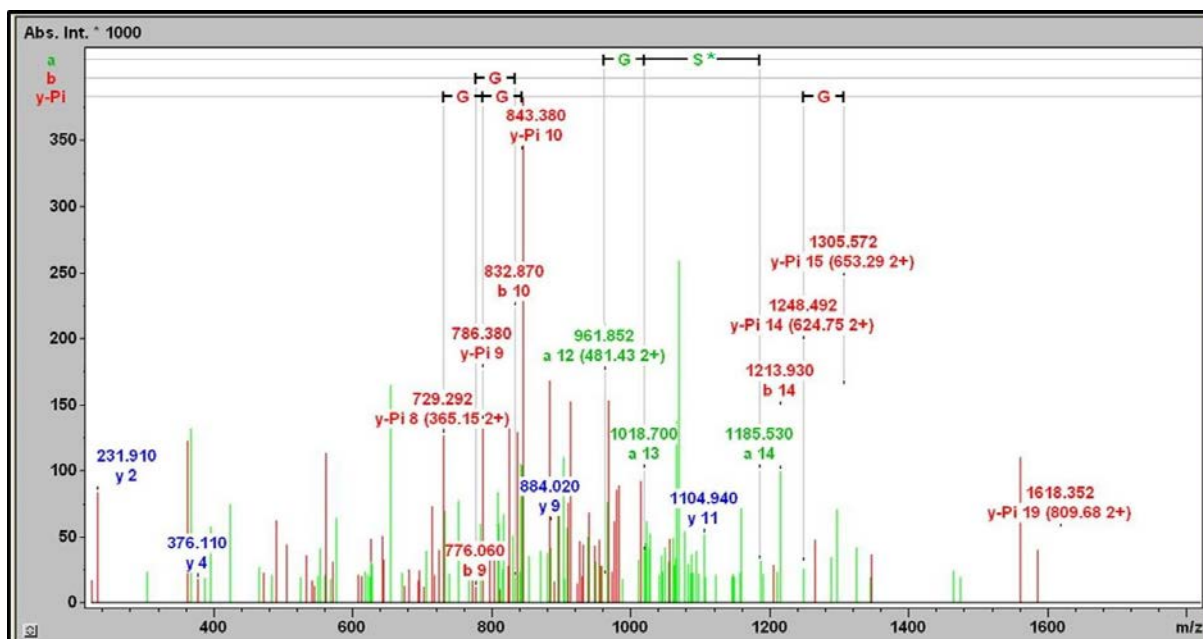
S24. (A) - Espectro CID da sequência peptídica L.GGQGAGAAAAAAGGAGQGGLGGQGAGQGAGAS*AAAAGGAGQGGY.G (211-255), selecionando-se o íon de m/z 863.34 na forma de $[M + 4H]^{4+}$, como precursor; e mostrando o sítio de fosforilação S243 observado na proteína espidroína-1 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. (B) - Espectro ETD adquirido para a mesma proteína.



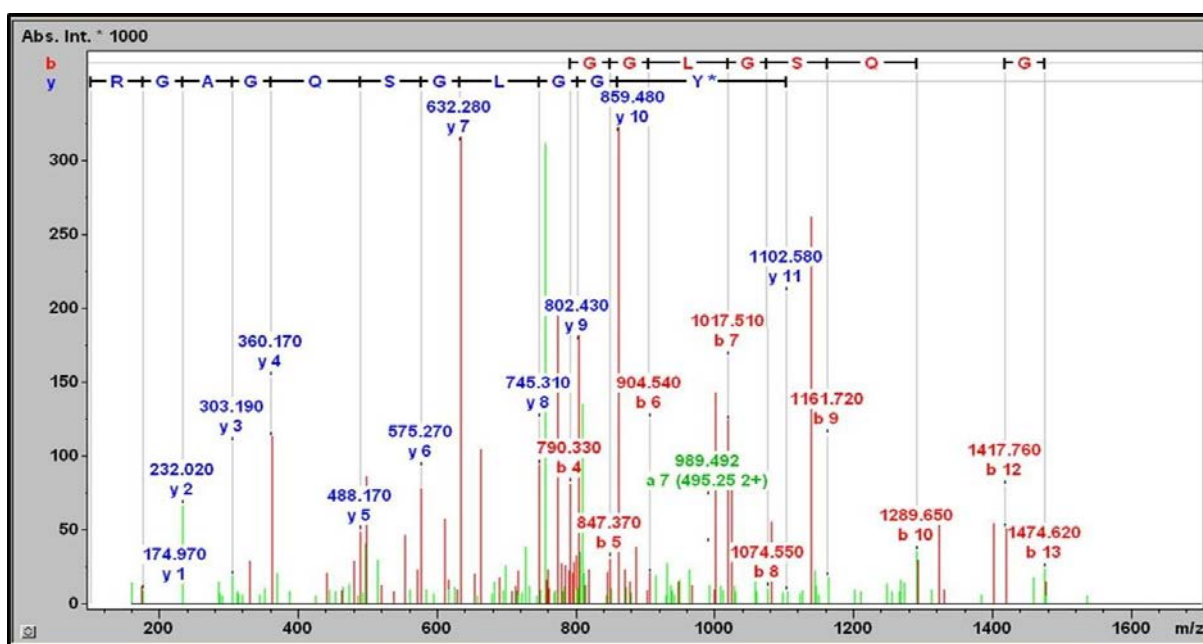
S25. Espectro CID da sequência peptídica GS*QGAGRGGEGAGAAAAAAGGAGQGGY (259-285), selecionando-se o íon de m/z 1072.952 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação S260 observado na proteína espidroína-1 (*spot* 1), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



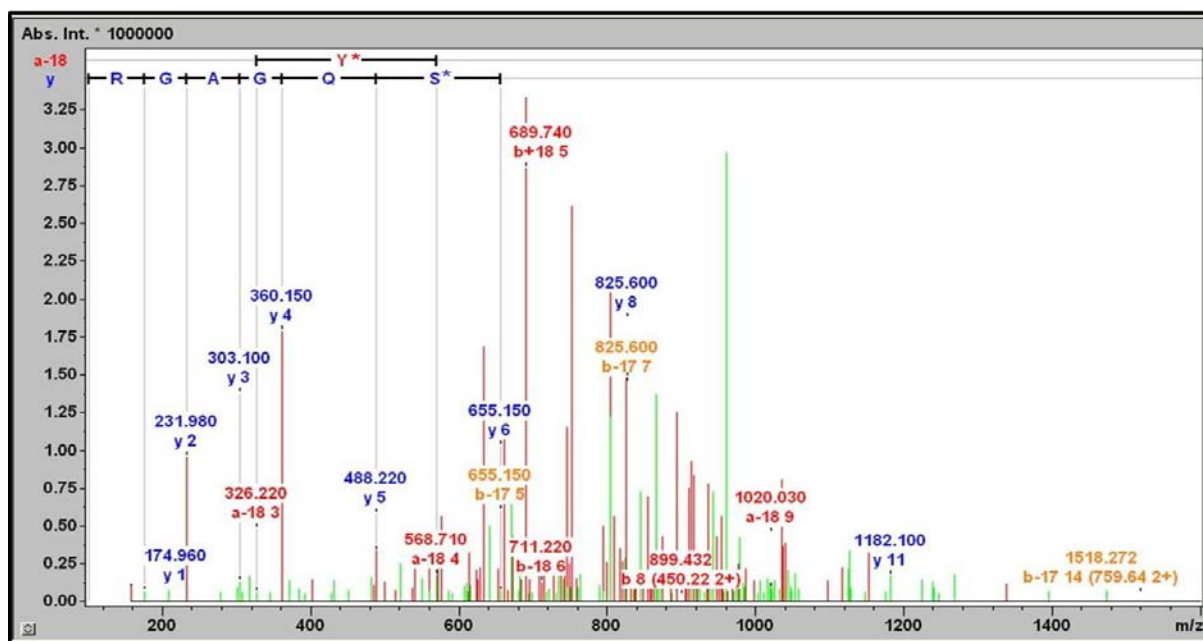
S26. (A) - Espectro CID da sequência peptídica Y.GGLGS*QGAGRGGEGAGAA*AAAAGGAGQGGY.G (256-298), selecionando-se o íon de m/z 839.698 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de fosforilação S260 e o aminoácido substituído Ala->Phe (A273) observados na proteína espidroína-1 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro ETD adquirido para a mesma proteína.



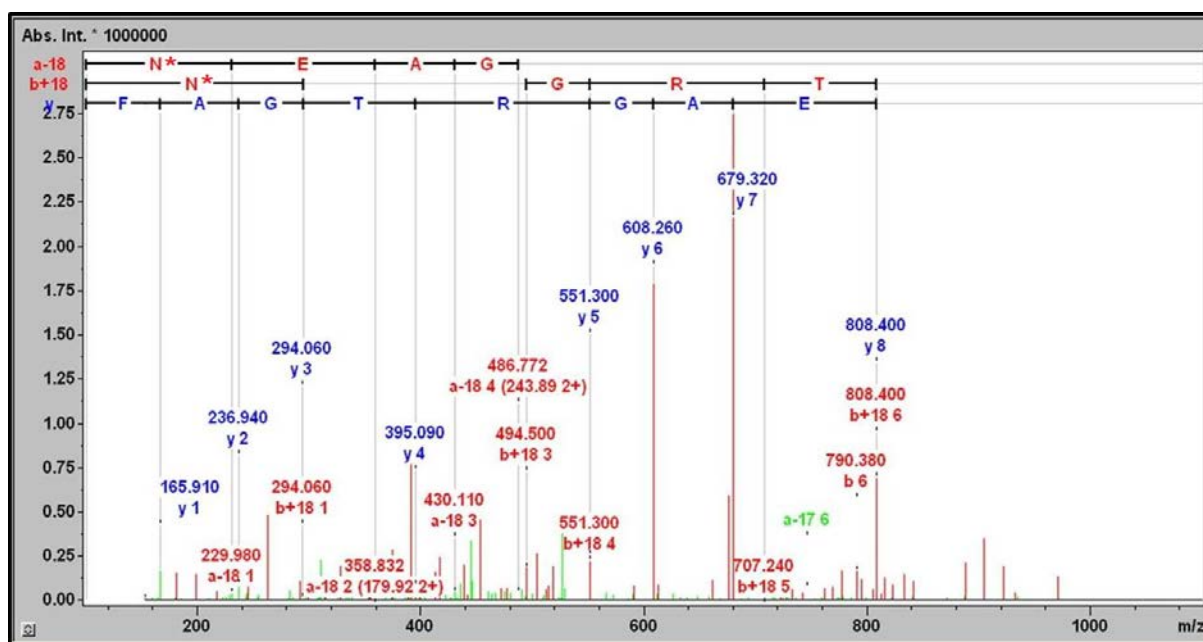
S28. Espectro CID da sequência peptídica R.GQGAGQGGYGGGLG[S*]QGSGR.G (510-528), selecionando-se o íon de m/z 858.862 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação S523 observado na proteína espidroína-1 (spot 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



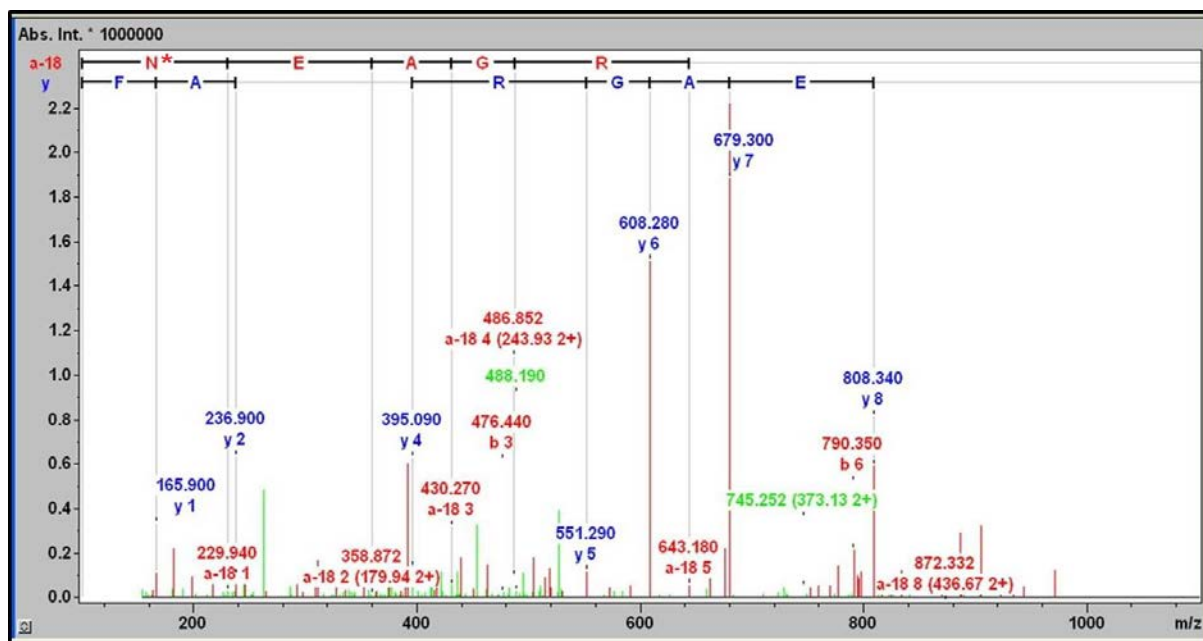
S29. Espectro CID da sequência peptídica R.QGGY[G*]GGLGSQGAGR.G (572-585), selecionando-se o íon de m/z 824.895 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação Y575 observado na proteína espidroína-1 (spot 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



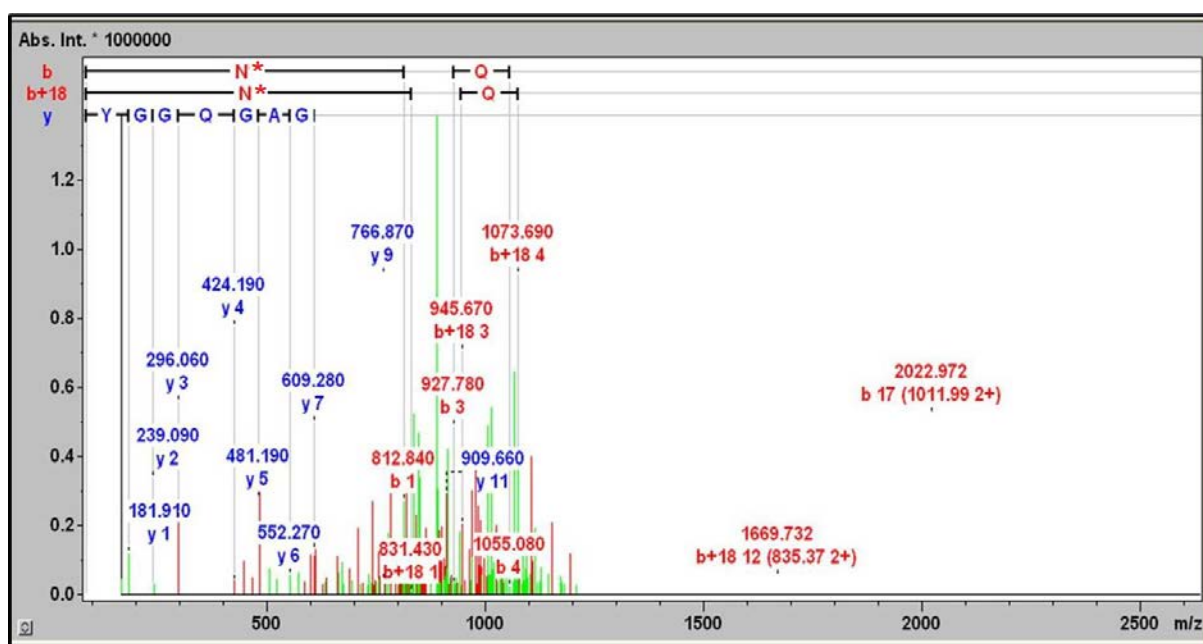
S30. Espectro CID da sequência peptídica R.QGGY*GGLGS*QGAGR.G (572-585), selecionando-se o íon de m/z 777.369 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando os sítios de fosforilação Y575 e S580 observados na proteína espidroína-1 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



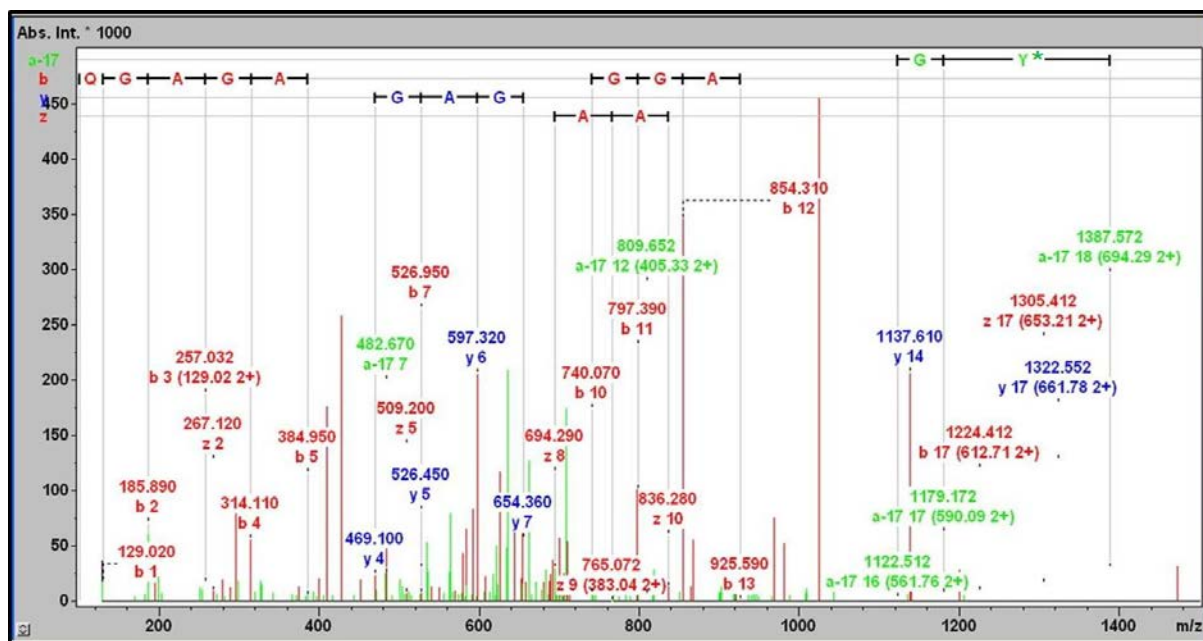
S31. Espectro CID da sequência peptídica L.N*EAGRTGAF.T (44-52), selecionando-se o íon de m/z 535.22 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de glicosilação N44 (hexosamine) observado na proteína espidroína-1 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



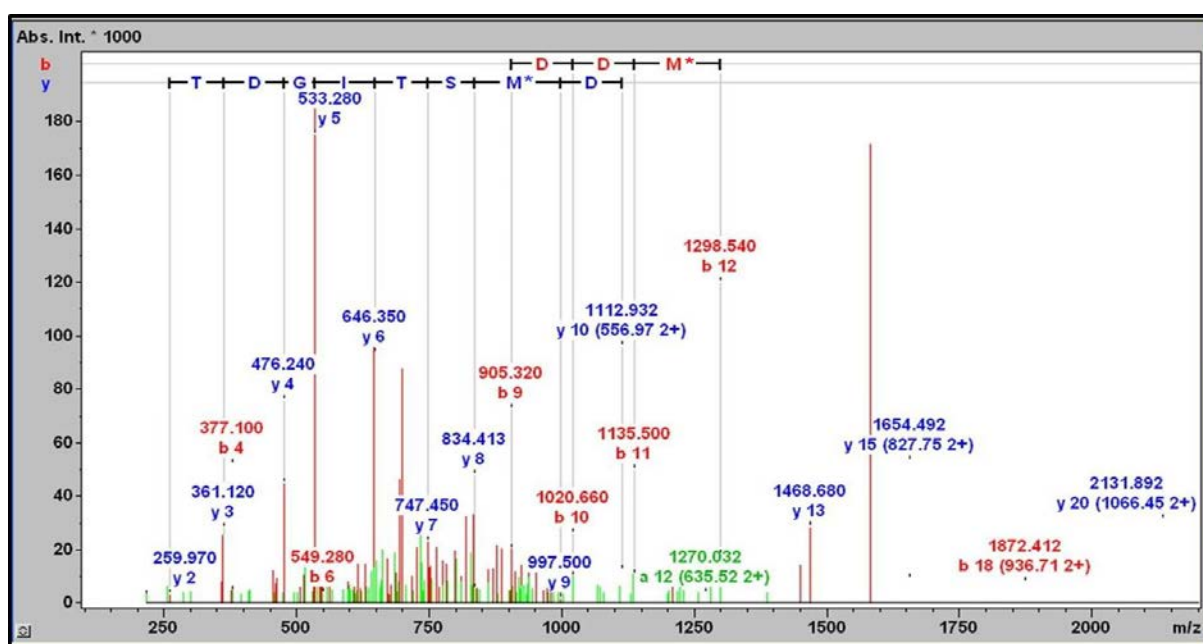
S32. Espectro CID da sequência peptídica L.N*EAGRTGAF.T (44-52), selecionando-se o íon de m/z 535.24 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de glicosilação N44 (hexosamine) observado na proteína espidroína-1 (spot 5), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



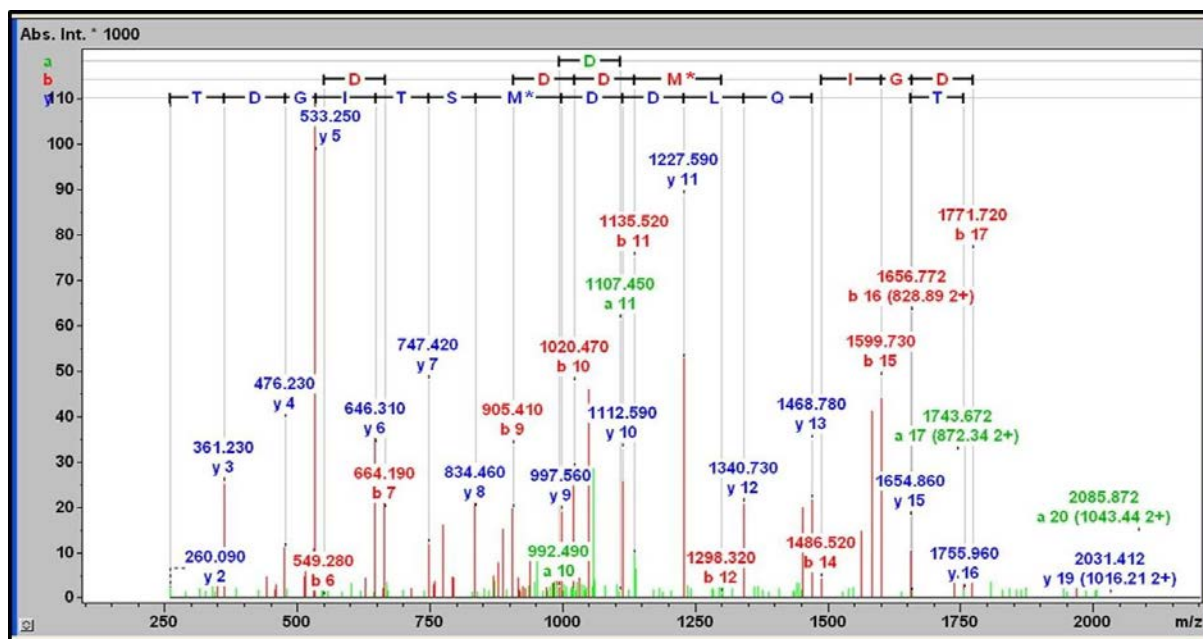
S33. Espectro CID da sequência peptídica Y.N*GGQGPSAAAAAASGAGQGGY.G (237-260), selecionando-se o íon de m/z 877.88 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de glicosilação N237 (HexNAc2DHex2) observado na proteína espidroína-1 (spot 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



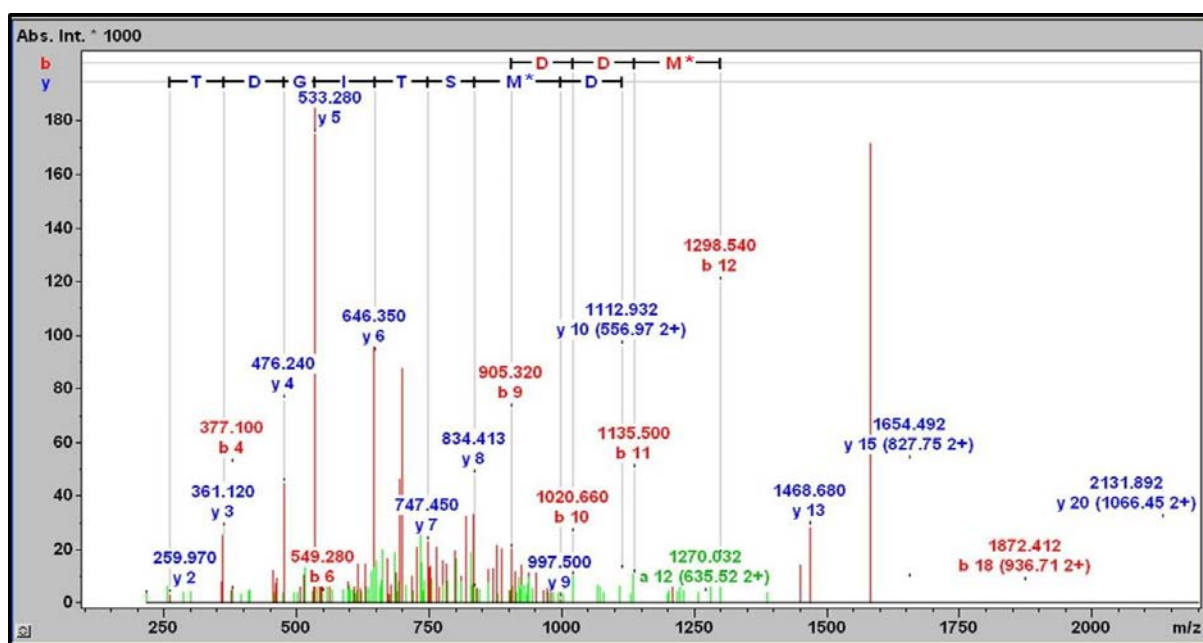
S34. Espectro CID da sequência peptídica QGAGAAAAAGGAGQGGY*.G (1-18), selecionando-se o íon de m/z 725.86 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando a nitrotirosina Y18 observado na proteína espidroína-1 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



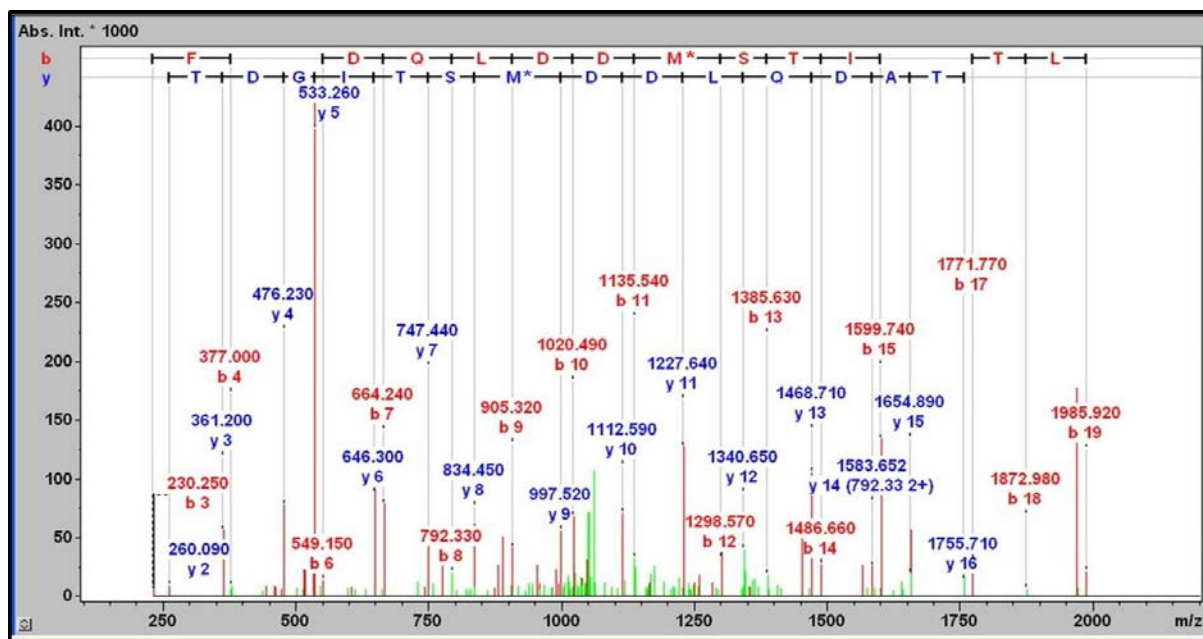
S35. Espectro CID da sequência peptídica R.TGAFTADQLDDM*.STIGDTLK.T (49-68), selecionando-se o íon de m/z 711.325 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de dihidroxilação (Sulphone) M60 observado na proteína espidroína-1 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



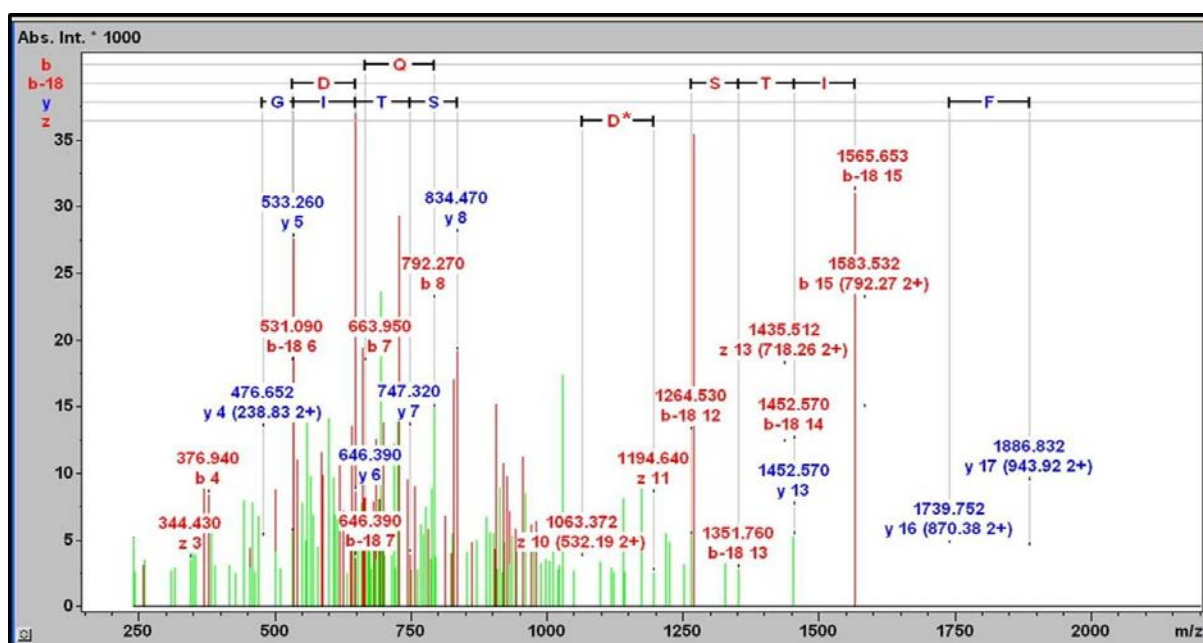
S36. Espectro CID da sequência peptídica R.TGAFTADQLDDM*STIGDTLK.T (49-68), selecionando-se o íon de m/z 1066.43 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de dihidroxilação (sulphone) M60 observado na proteína espidroína-1 (spot 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



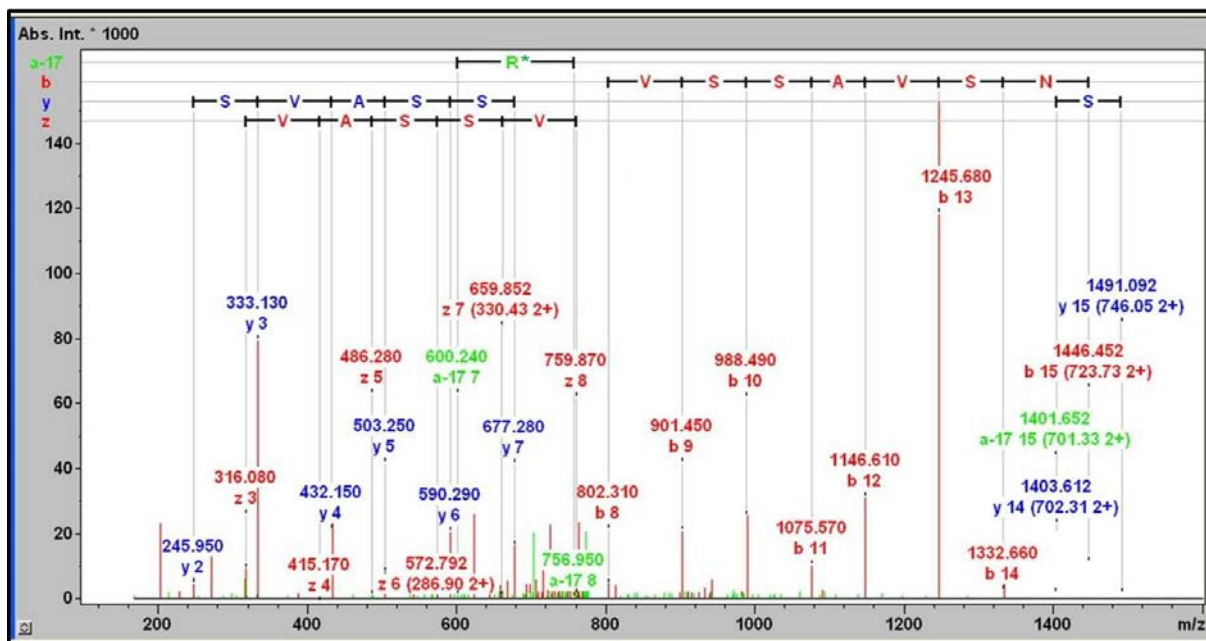
S37. Espectro CID da sequência peptídica R.TGAFTADQLDDM*STIGDTLK.T (49-68), selecionando-se o íon de m/z 711.32 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de dihidroxilação (sulphone) M60 observado na proteína espidroína-1 (spot 5), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



S38. Espectro CID da sequência peptídica R.TGAFTADQLDDM*STIGDTLK.T (49-68), selecionando-se o íon de m/z 1066.48 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de dihidroxilação (sulphone) M60 observado na proteína espidroína-1 (spot 5), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

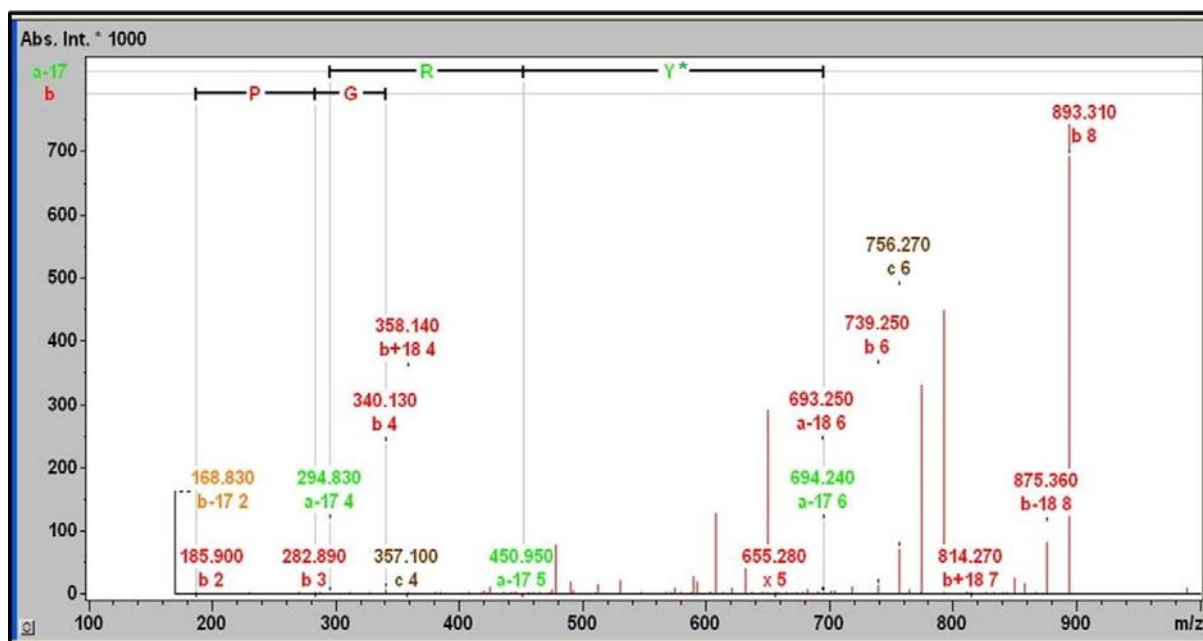


S39. Espectro CID da sequência peptídica R.TGAFTADQLD*DMSTIGDTLK.T (49-68), selecionando-se o íon de m/z 706.06 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de hidroxilação D58 observado na proteína espidroína-1 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

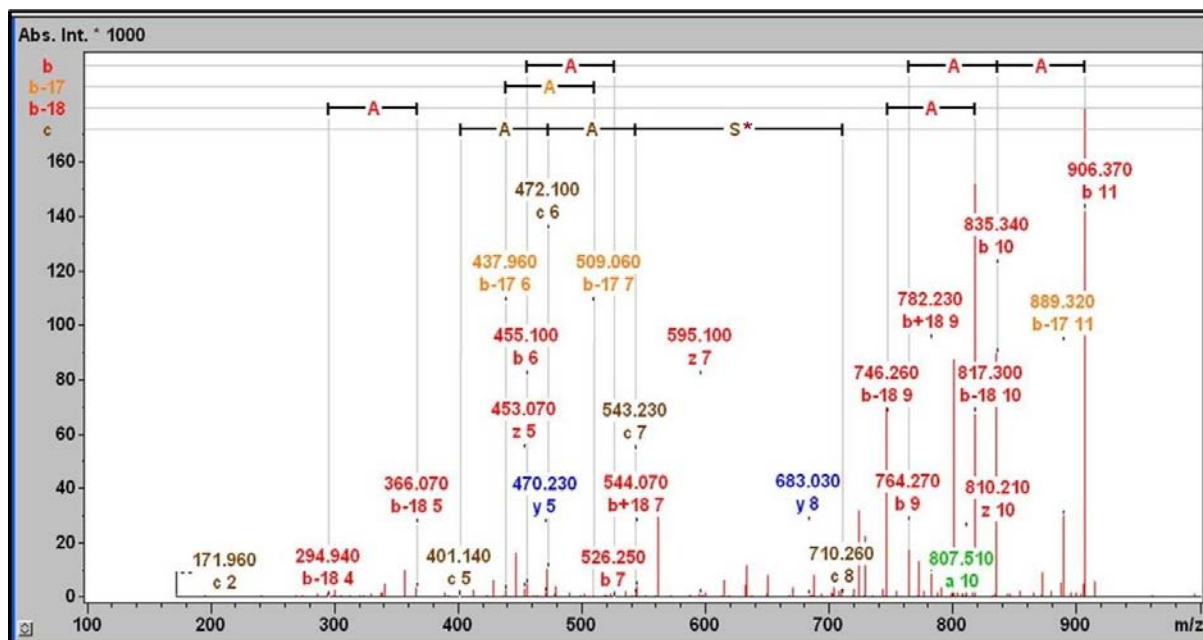


S40. Espectro CID da sequência peptídica L.SSPQASSR*VSSAVSNL.V (654-669), selecionando-se o íon de m/z 789.42 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de deaminação R261 observado na proteína espidroína-1(*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

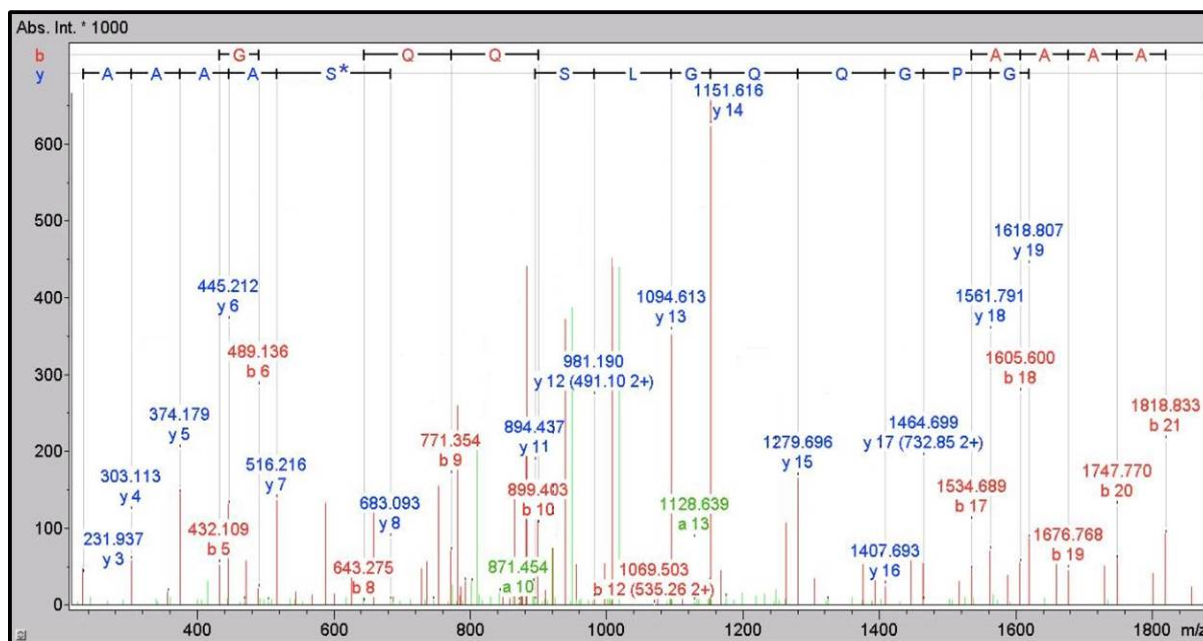
Figuras dos espectros representativos das modificações pós-traducionais observadas na proteína espidroína-2 que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda da aranha *N. clavipes*.



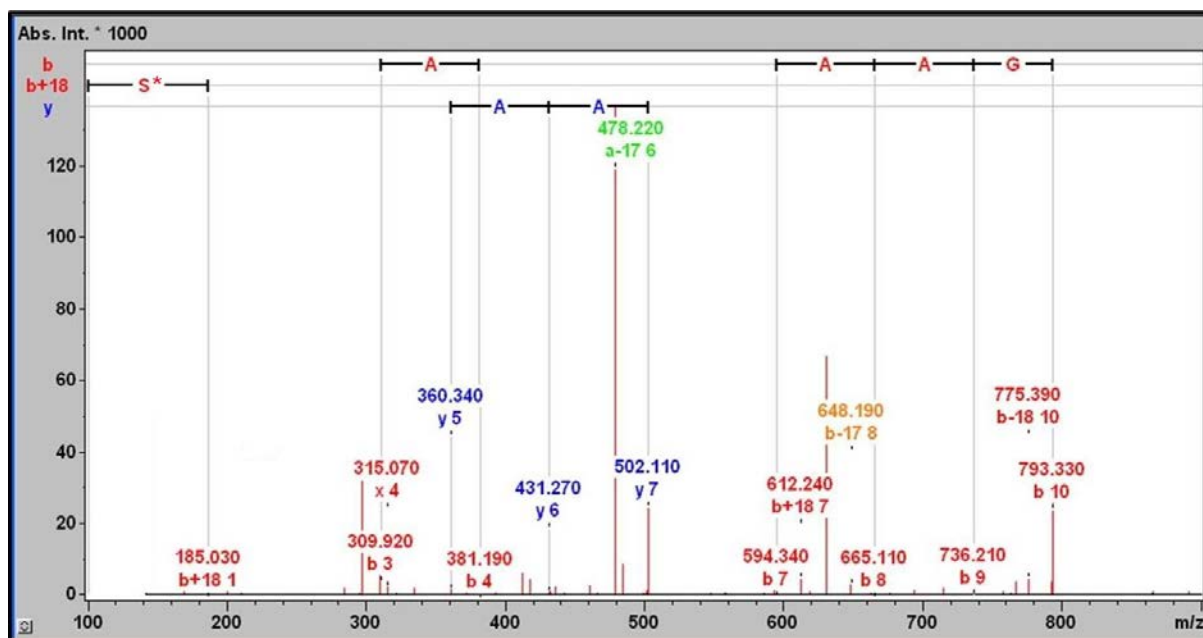
S41. Espectro CID da sequência peptídica Q.QGPGRY*GP.G (56-63), selecionando-se o íon de m/z 911.360 na forma de $[M + H]^+$, como precursor; e mostrando o sítio de fosforilação Y61 observado na proteína espidroína-2 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



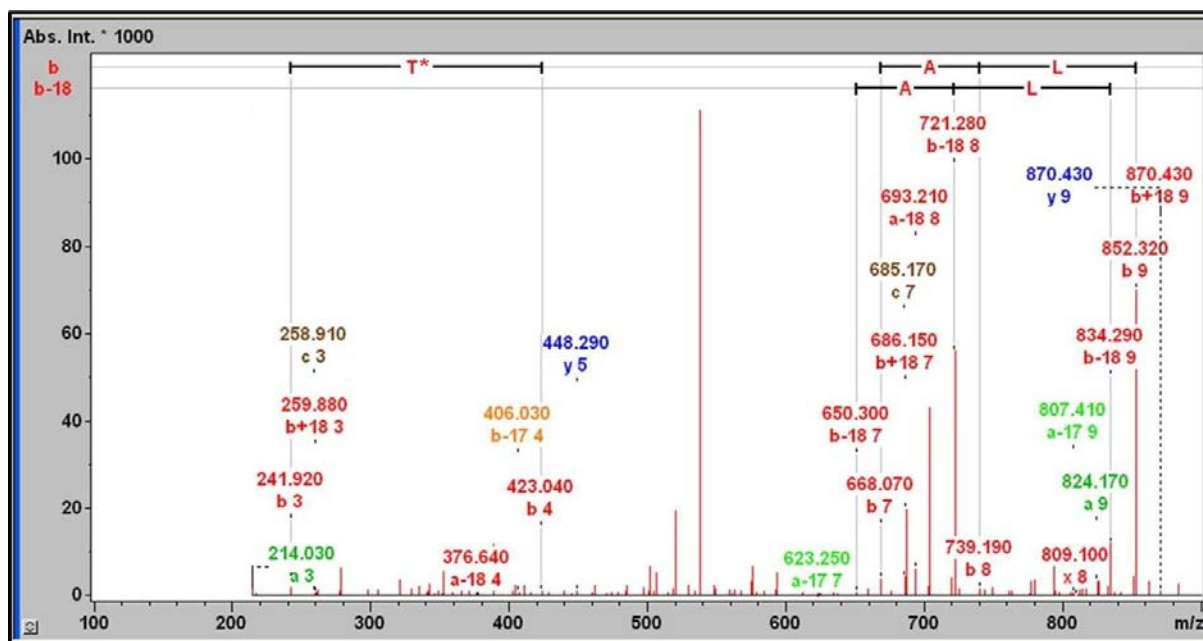
S42. Espectro CID da sequência peptídica G.PGSAAAAS*AAA.S (109-119), selecionando-se o íon de m/z 924.410 na forma de $[M + H]^+$, como precursor; e mostrando o sítio de fosforilação S116 observado na proteína espidroína-2 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



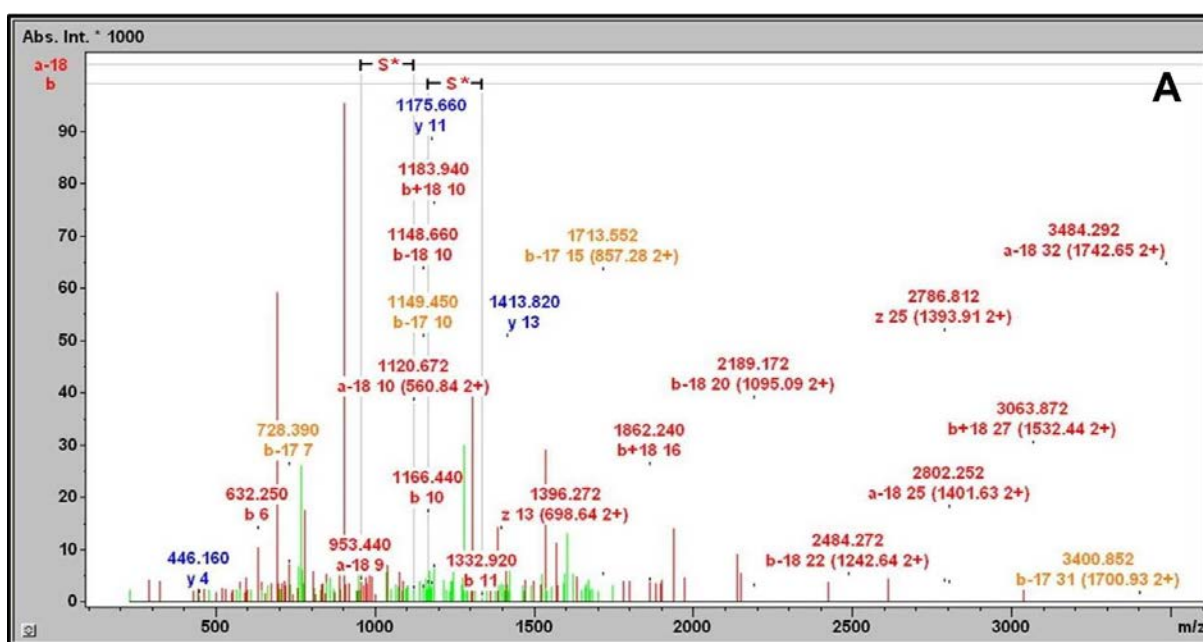
S43. Espectro CID da sequência peptídica Q.GPGGYGPGQQGLSGPG**S***AAAAAAA.G (229 -252), selecionando-se o íon de m/z 1025.890 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação S245 observado na proteína espidroína-2 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

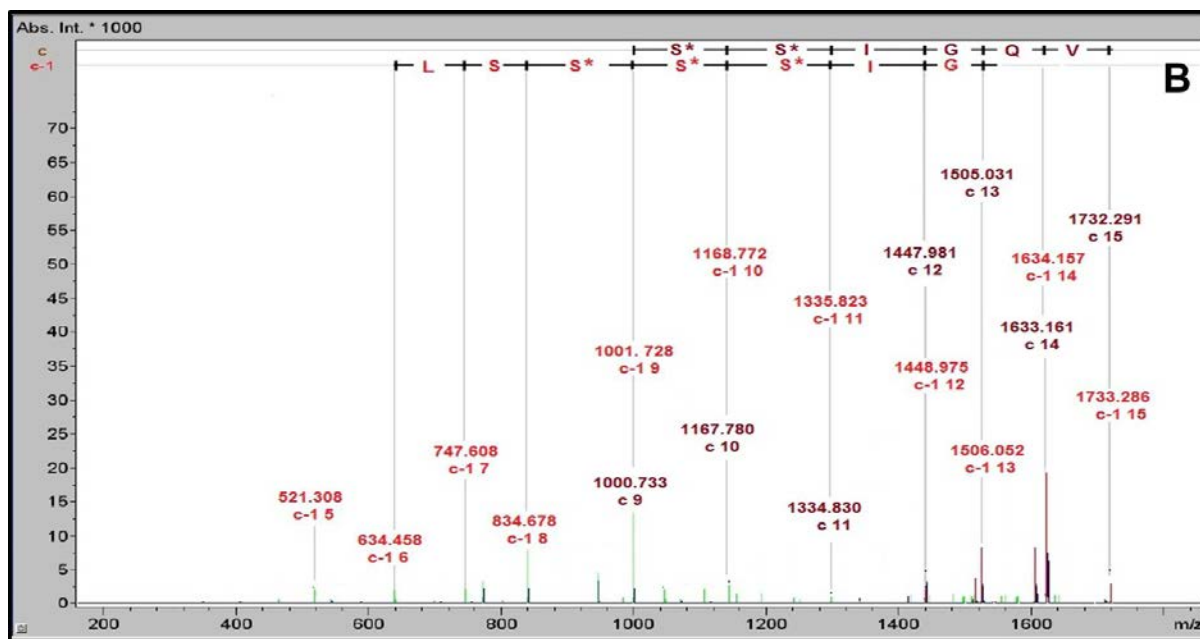


S44. Espectro CID da sequência peptídica A.**S***AAAAAAAAG.P (383-392), selecionando-se o íon de m/z 811.390 na forma de $[M + H]^+$, como precursor; e mostrando o sítio de fosforilação S383 observado na proteína espidroína-2 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.

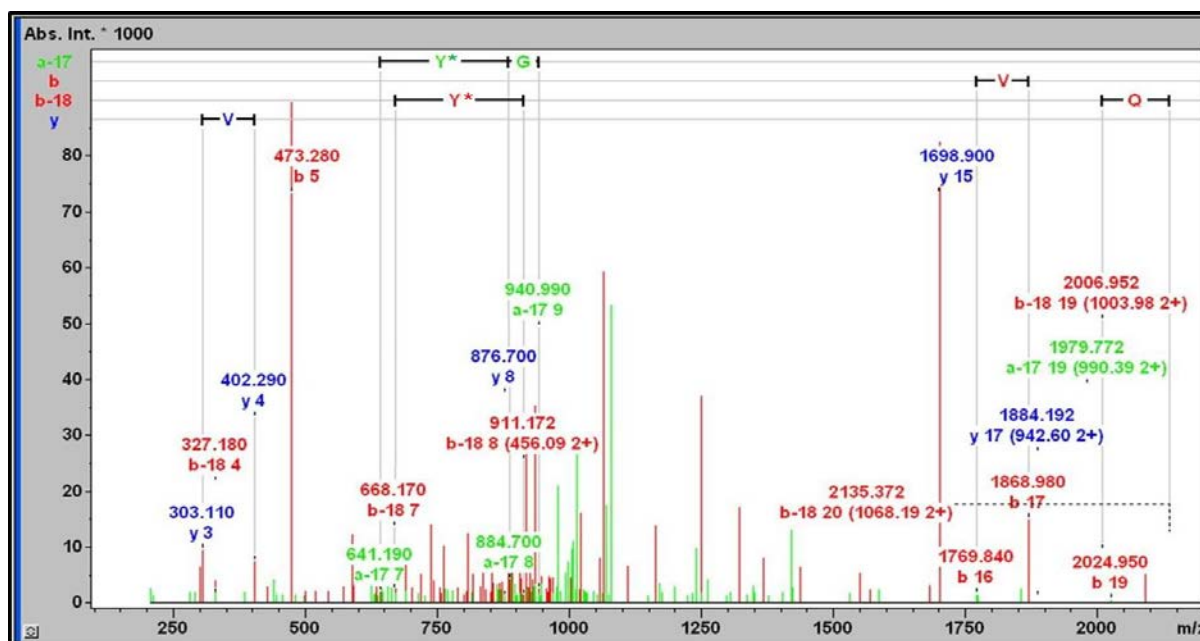


S45. Espectro CID da sequência peptídica S.SGPT*SSAAL.S (552-560), selecionando-se o íon de m/z 870.360 na forma de $[M + H]^+$, como precursor; e mostrando o sítio de fosforilação T555 observado na proteína espidroína-2 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.





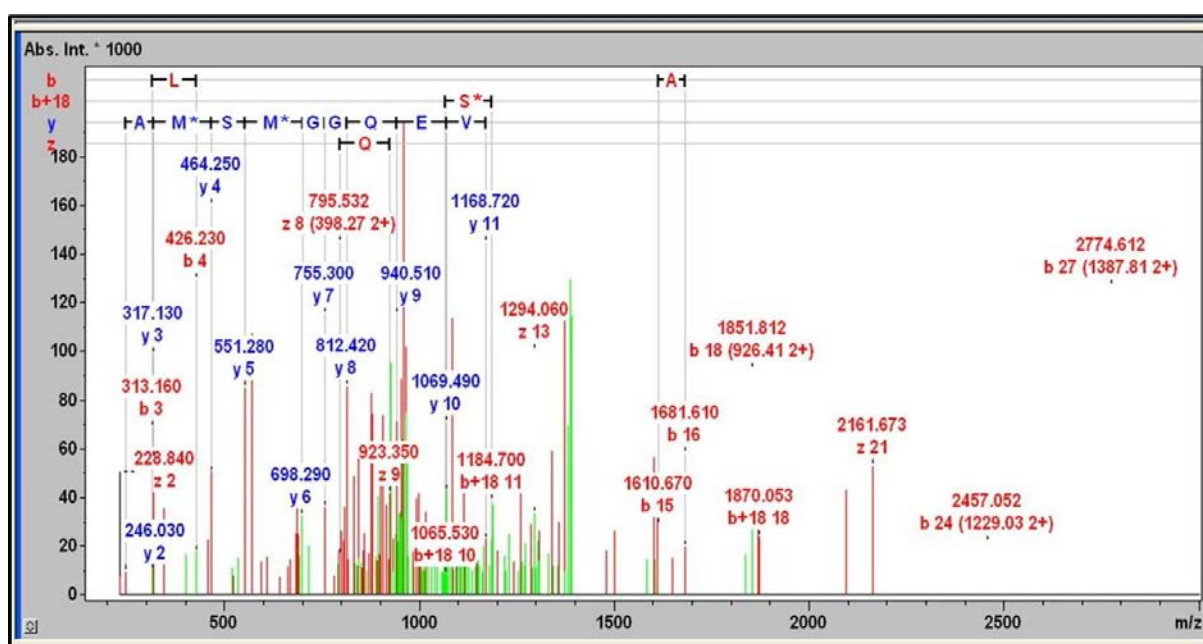
S46. (A) - Espectro CID da sequência peptídica V.SACVTILSS*S*S*IGQVNYGAAS*QFAQVVGQSVL.S (593-624), selecionando-se o íon de m/z 1183.640 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando os sítios de fosforilação S601, S602, S603 e S613 observados na proteína espidroína-2 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda. **(B)** - Espectro ETD adquirido para a mesma proteína.



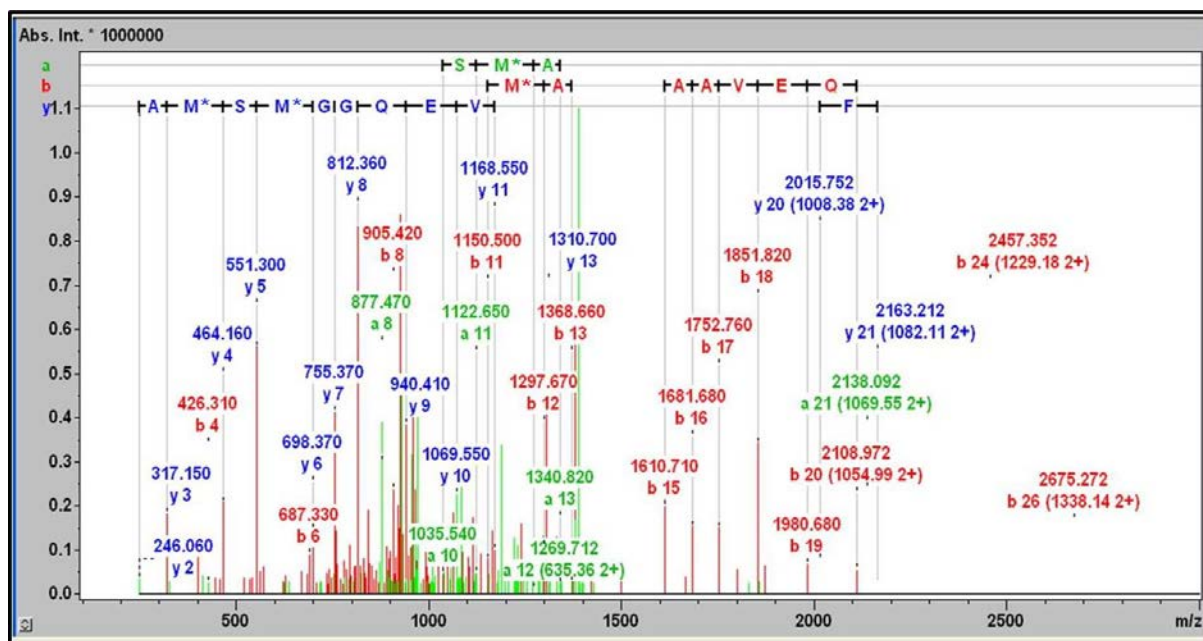
S47. Espectro CID da sequência peptídica S.SIGQVNY*GAAS*QFAQVVGQ.S (602-621), selecionando-se o íon de m/z 1086.050 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando os sítios de fosforilação Y609 e S613 observados na proteína espidroína-2 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



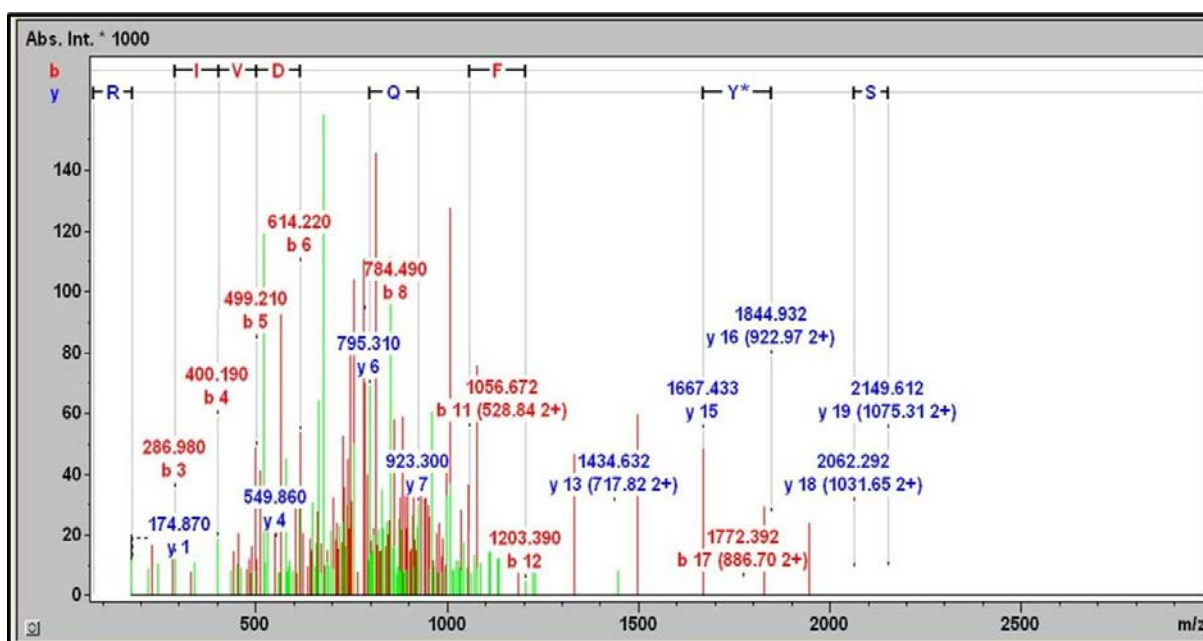
S48. Espectro CID da sequência peptídica S.SSIGQVNY*GAAS*QFAQVVGQ.S (602-621), selecionando-se o íon de m/z 1086.040 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando os sítios de fosforilação Y609 e S613 observados na proteína espidroína-2 (*spot* 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



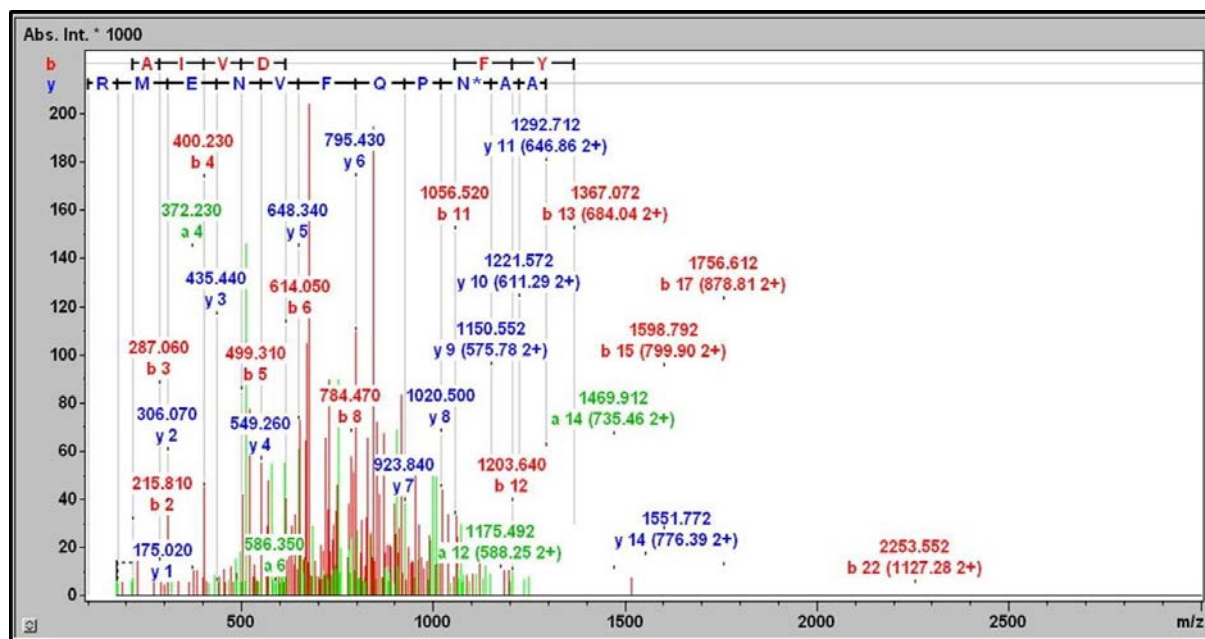
S49. Espectro CID da sequência peptídica K.LQALNMAFAS*MAEIAAVEQGGM*SM*AVK.T (85-112), selecionando-se o íon de m/z 974.06 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de dihidroxilação S95 e os sítios de oxidação M107 e M109 observados na proteína espidroína-2 (*spot* 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



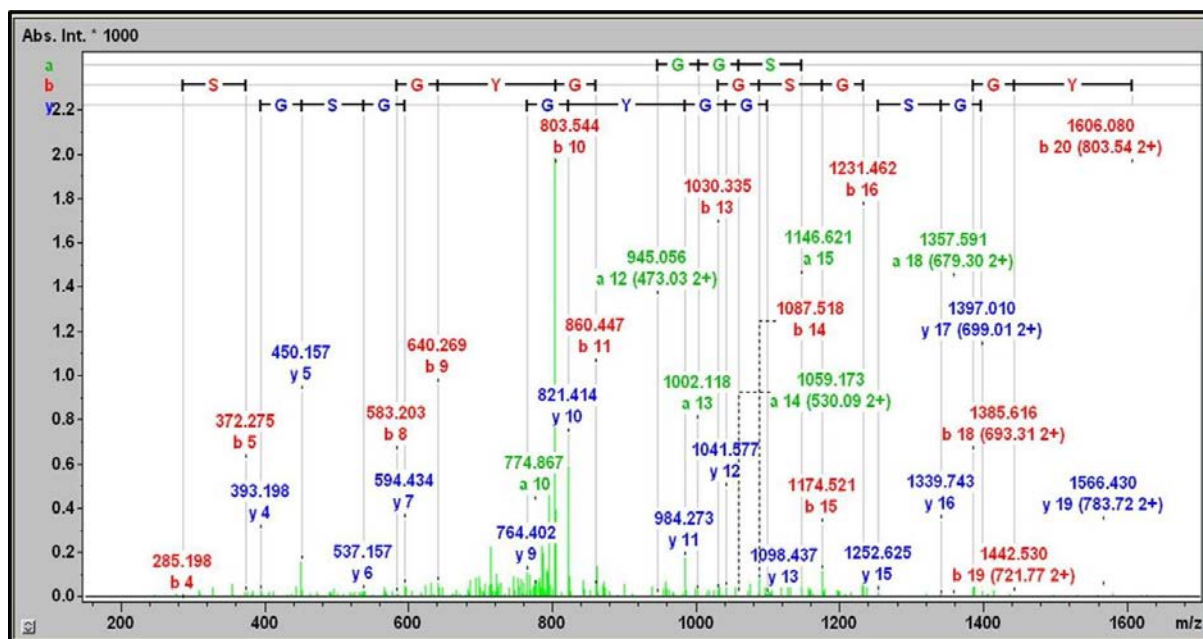
S50. Espectro CID da sequência peptídica K.LQALN*MAFASSM*AEIAAVEQGGM*SM*AVK.T (85-112), selecionando-se o íon de m/z 974.04 na forma de $[M + 3H]^{3+}$, como precursor; e mostrando o sítio de hidroxilação N89 e os sítios de oxidação M96, M107 e M109 observados na proteína espidroína-2 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



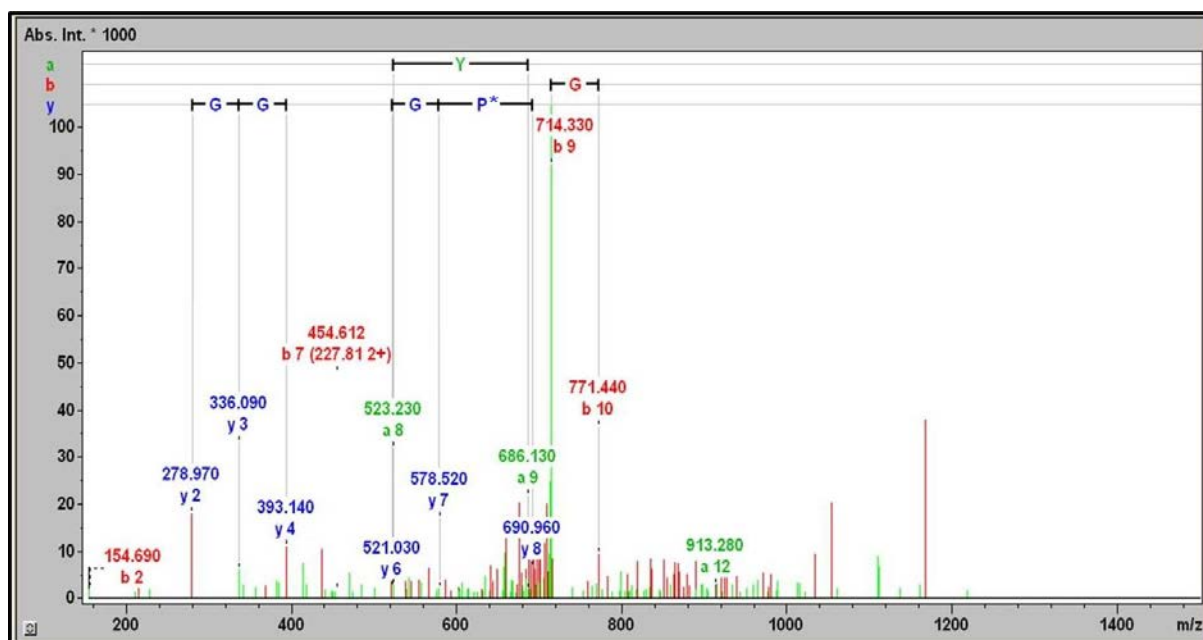
S51. Espectro CID da sequência peptídica K.TNAIVDGLNSAFY*MTTGAANPQFVNEMR.S (113-140), selecionando-se o íon de m/z 766.56 na forma de $[M + 4H]^{4+}$, como precursor; e mostrando o aminotirosina Y127 observado na proteína espidroína-2 (spot 3), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



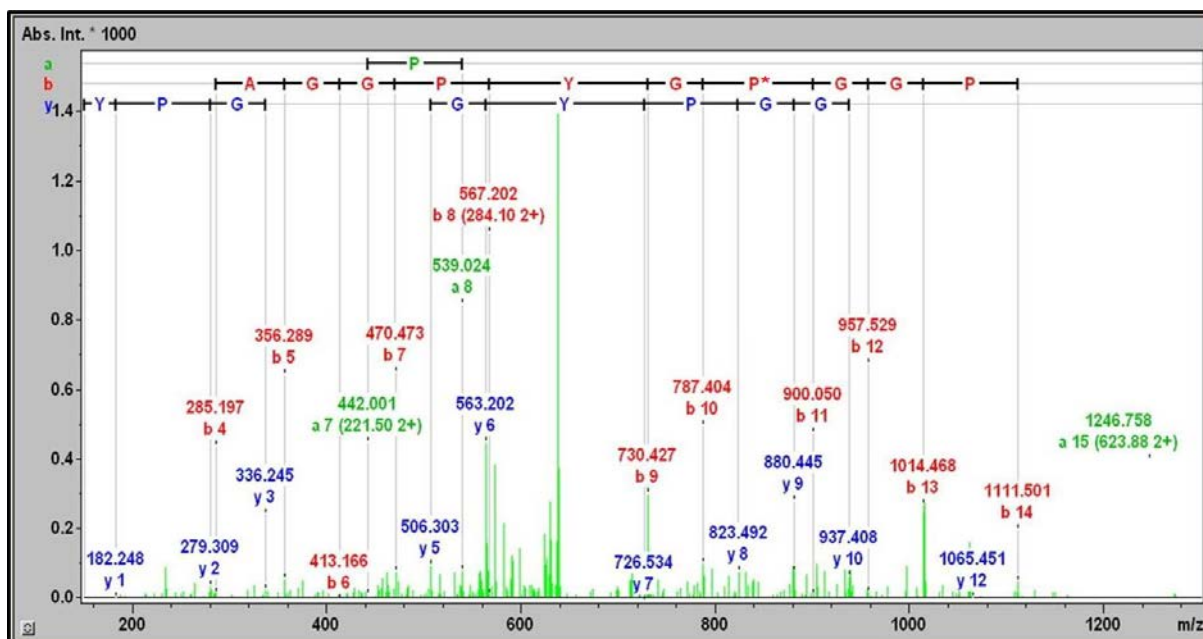
S52. Espectro CID da sequência peptídica K.TNAIVDGLNSAFYMTTGAAN*PQFVNEMR.S (113-140), selecionando-se o íon de m/z 762.84 na forma de $[M + 4H]^{4+}$, como precursor; e mostrando o sítio de hidroxilação N132 observado na proteína espidroína-2 (spot 4), que foi secretada pela glândula ampulada maior e identificada na 2-DE da seda.



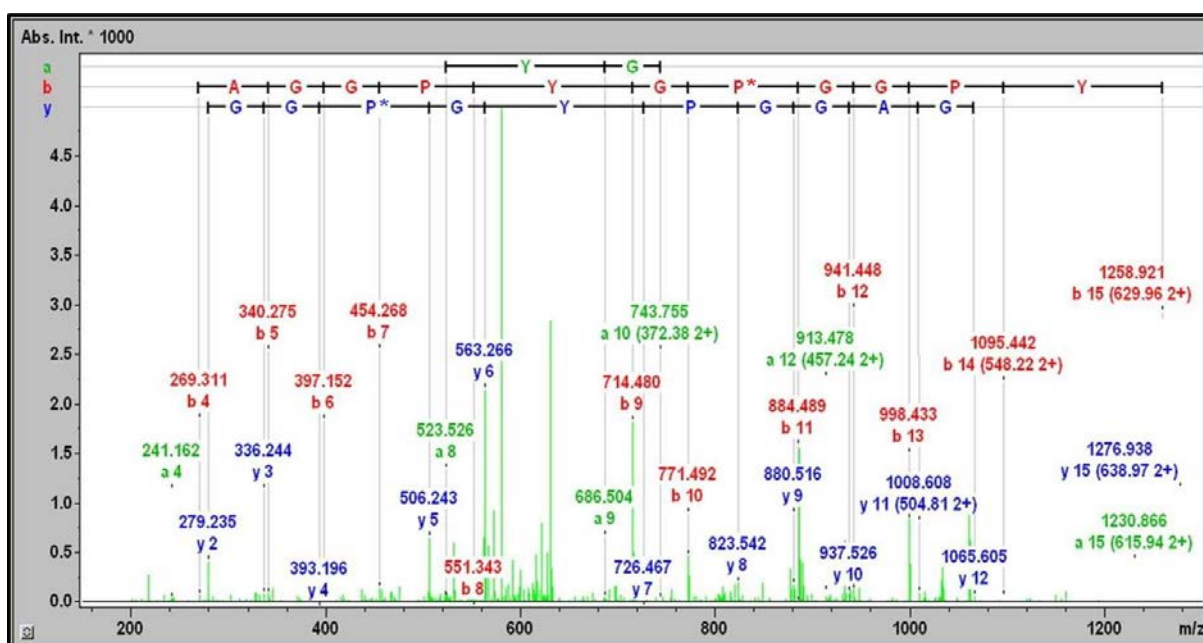
S55. Espectro CID da sequência peptídica Y.GP*GGSGPGGYG (26 - 45; 36 - 55; 46 - 65; 56 - 75; 469 - 488; 479 - 498; 489 - 508; 519 - 538), selecionando-se o íon de m/z 812.33 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando os sítios de hidroxilação P27, P37, P47, P57, P67, P470, P480, P490, P500, P520 e P530 observados na proteína da seda flageliforme (spot 2), que foi secretada pela glândula flageliforme e identificada na 2-DE da seda.



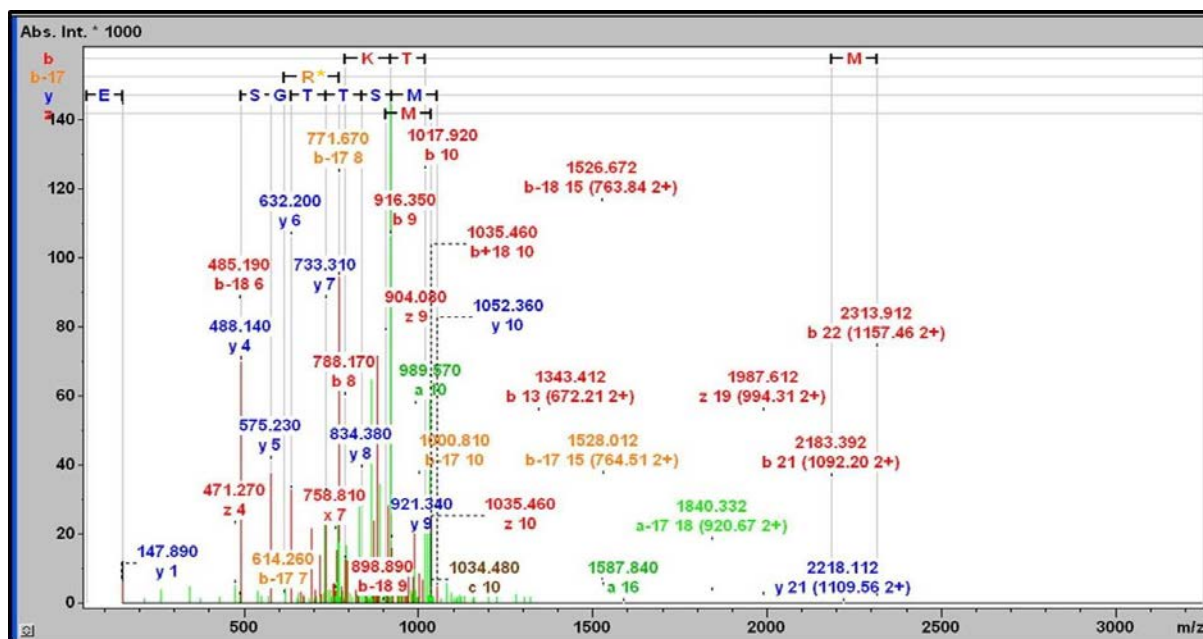
S56. Espectro CID da sequência peptídica Y.GPGGAGGPYP*GGAGGPY.G (372-389), selecionando-se o íon de m/z 731.45 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de hidroxilação P382 observado na proteína da seda flageliforme (spot 1), que foi secretada pela glândula flageliforme e identificada na 2-DE da seda.



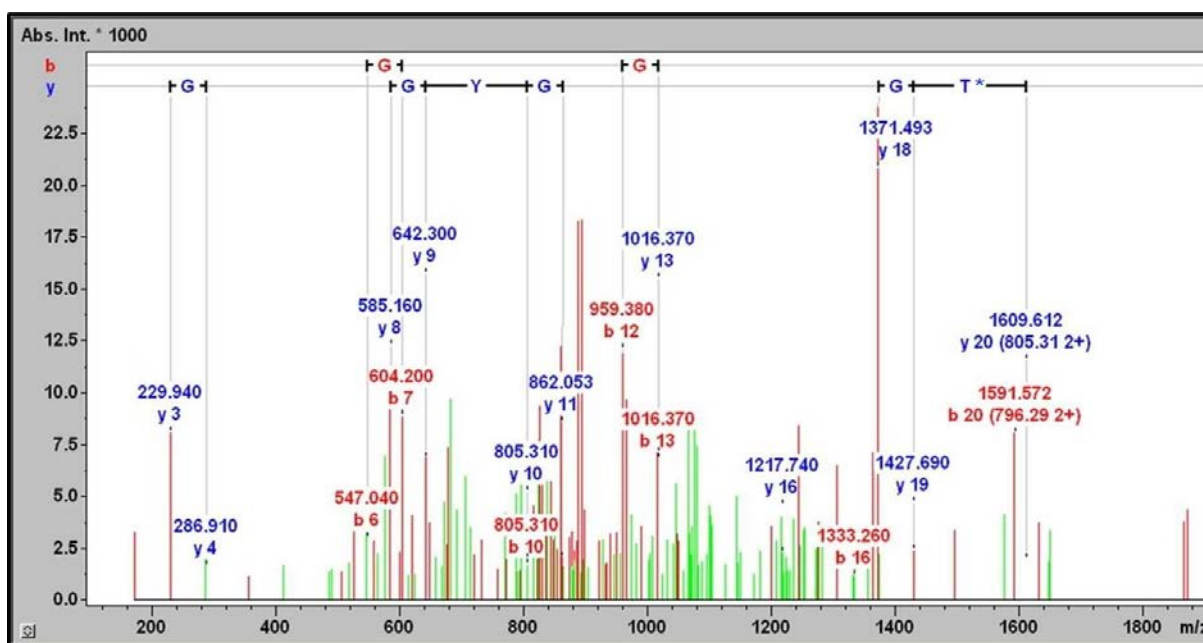
S57. Espectro CID da sequência peptídica Y.GP*GGAGGPYP*GGPY.G (405 - 419), selecionando-se o íon de m/z 646.78 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando os sítios de hidroxilação P406 e P415 observados na proteína da seda flageliforme (*spot* 2), que foi secretada pela glândula flageliforme e identificada na 2-DE da seda.



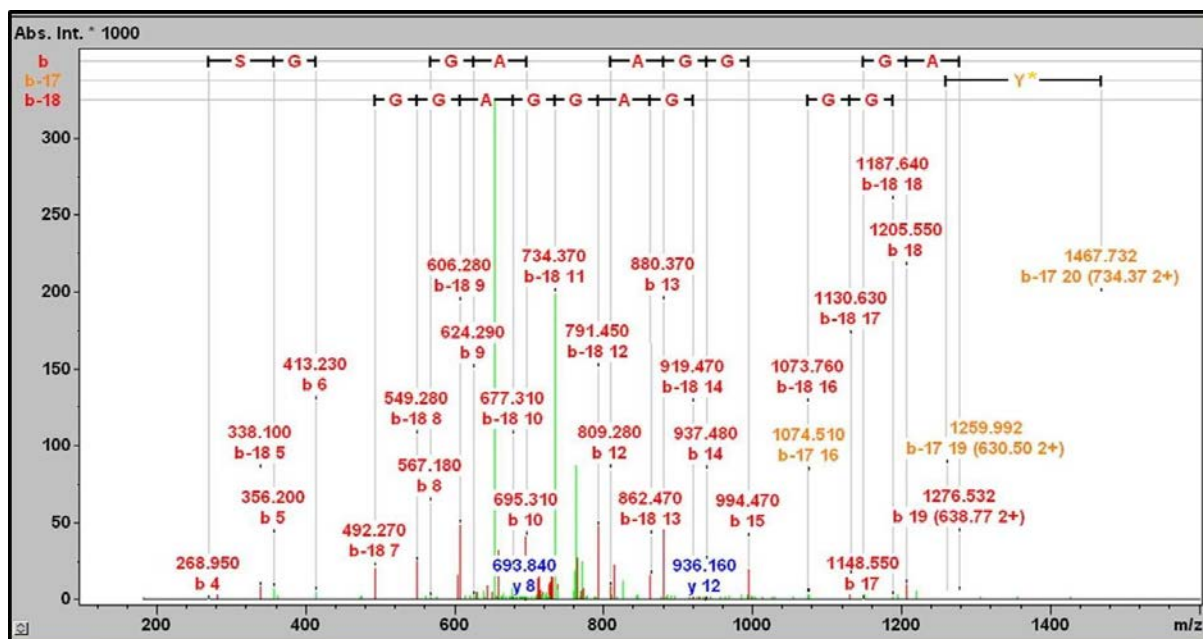
S58. Espectro CID da sequência peptídica Y.GP*GGAGGPYP*GGPY.G (405 - 419), selecionando-se o íon de m/z 638.80 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de hidroxilação P415 observado na proteína da seda flageliforme (*spot* 2), que foi secretada pela glândula flageliforme e identificada na 2-DE da seda.



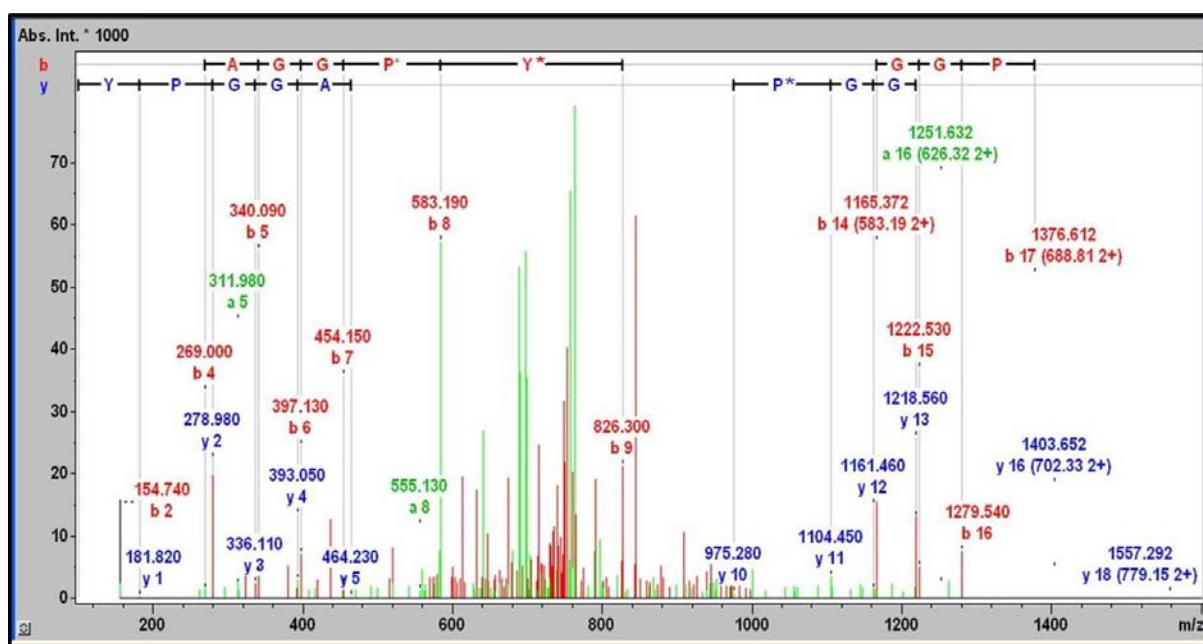
S59. Espectro CID da sequência peptídica E.SSGGDVQ**R***KTNVISNALRNALMSTTGSPNEE.F (31 - 61), selecionando-se o íon de m/z 809.35 na forma de $[M + 4H]^{4+}$, como precursor; e mostrando o sítio de deamidação R38 observado na proteína da seda flageliforme (*spot* 2), que foi secretada pela glândula flageliforme e identificada na 2-DE da seda.



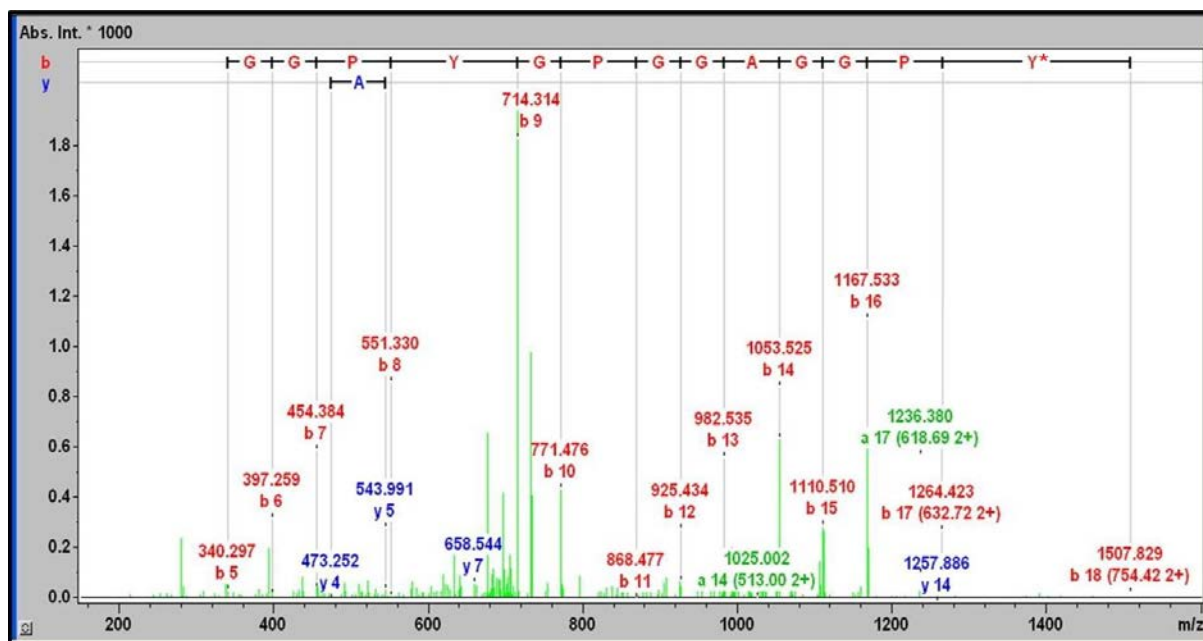
S60. Espectro CID da sequência peptídica G.PGG**T***GPGGSGPGGYGPGGSGPGG.S (92 - 114), selecionando-se o íon de m/z 910.890 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação T95 observado na proteína da seda flageliforme (*spot* 2), que foi secretada pela glândula flageliforme e identificada na 2-DE da seda.



S61. Espectro CID da sequência peptídica Y.GPGGSGPGGAGGAGGPGGA Y*.G (317-336), selecionando-se o íon de m/z 751.88 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando a nitrotirosina Y336 observado na proteína da seda flageliforme (spot 1), que foi secretada pela glândula flageliforme e identificada na 2-DE da seda.



S62. Espectro CID da sequência peptídica Y.GPGGAGGP*Y*GPGGAGGPY.G (372-389), selecionando-se o íon de m/z 779.39 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de dihidroxilação P379 e o sítio de fosforilação Y380 observados na proteína da seda flageliforme (spot 1), que foi secretada pela glândula flageliforme e identificada na 2-DE da seda.



S63. Espectro CID da sequência peptídica Y.GPGGAGGPYGPAGGPY*.G (372-389), selecionando-se o íon de m/z 723.32 na forma de $[M + 2H]^{2+}$, como precursor; e mostrando o sítio de fosforilação Y389 observado na proteína da seda flageliforme (spot 2), que foi secretada pela glândula flageliforme e identificada na 2-DE da seda.

Rio Claro, 25 de março de 2014.

Prof. Dr. Mario Sergio Palma
Orientador

José Roberto Aparecido dos Santos-Pinto
Doutorando