



UNIVERSIDADE DE BRASÍLIA-UNB
INSTITUTO DE BIOLOGIA
DEPARTAMENTO DE BIOLOGIA CELULAR
VALQUÍRIA ALICE MICHALCZECHEN LACERDA

**PRODUÇÃO DE PROTEÍNAS MaSp DA ARANHA *Parawixia bistriata* POR
BACTÉRIAS MODIFICADAS METABOLICAMENTE, E CARACTERIZAÇÃO
MOLECULAR DAS FIBRAS SINTÉTICAS.**

BRASÍLIA - DISTRITO FEDERAL – BRASIL

ABRIL de 2015

VALQUÍRIA ALICE MICHALCZECHEN LACERDA

**PRODUÇÃO DE PROTEÍNAS MaSp DA ARANHA *Parawixia bistriata* POR
BACTÉRIAS MODIFICADAS METABOLICAMENTE, E CARACTERIZAÇÃO
MOLECULAR DAS FIBRAS SINTÉTICAS.**

Tese de doutorado apresentada à banca examinadora do Programa de Pós-graduação em Biologia Molecular do Departamento de Biologia Celular da Universidade de Brasília, como parte dos requisitos necessários à obtenção do título de Doutor em Biologia Molecular.

Orientador: Prof. Dr. Elíbio Leopoldo Rech Filho.
Co-orientador: Giovanni Rodrigues Vianna

BRASÍLIA - DISTRITO FEDERAL – BRASIL

ABRIL de 2015

VALQUÍRIA ALICE MICHALCZECHEN LACERDA

**PRODUÇÃO DE PROTEÍNAS MaSp DA ARANHA *Parawixia bistriata* POR
BACTÉRIAS MODIFICADAS METABOLICAMENTE, E CARACTERIZAÇÃO
MOLECULAR DAS FIBRAS SINTÉTICAS.**

Tese de doutorado apresentada à banca examinadora do Programa de Pós-graduação em Biologia Molecular do Departamento de Biologia Celular da Universidade de Brasília, como parte dos requisitos necessários à obtenção do título de Doutor em Biologia Molecular.

Aprovado em: 28/04/20015

Banca examinadora:

Dr. Elíbio Leopoldo Rech Filho
Embrapa Recursos Genéticos e Biotecnologia
Orientador-Membro Externo Vinculado ao Programa

Dra. Lídia Maria Pepe de Moraes
Departamento de Biologia Celular - Universidade de Brasília
Membro Interno vinculado ao Programa

Dra. Sônia Maria de Freitas
Departamento de Biologia Celular - Universidade de Brasília
Membro Interno vinculado ao Programa

Dra. Ângela Mehta dos Reis
Embrapa Recursos Genéticos e Biotecnologia
Membro Externo não vinculado ao Programa

Dra. Daniela Matias de Carvalho Bittencourt
Empresa Brasileira de Pesquisa Agropecuária
Membro Externo não vinculado ao Programa

Dr. Giovanni Rodrigues Vianna
Embrapa Recursos Genéticos e Biotecnologia
Suplente - Membro Externo não vinculado ao Programa

Dedico esse trabalho à minha mãe, Ignez.

AGRADECIMENTOS

Agradeço a minha família, Igor E. M. Lacerda e Maria Ignez Michalczechen Lacerda pelo apoio incondicional na minha caminhada em pesquisa científica. Amo vocês.

Agradeço às amigas que fiz durante o período de doutorado sanduíche em Boston, em especial Olena Tokareva e Katarzyna Kazmieska.

Agradeço também à Paula E. Farias e Oliveira pela amizade, apoio e ajuda no experimento, mesmo que à distância.

Agradeço ao pessoal do Laboratório de Biologia Sintética da Embrapa Recursos Genéticos e Biotecnologia, Marina Silva, Andressa Rezende, Juliane Viana, Jéssica Carrijo, Débora Torres, Nicolau Brito, Giovanni Vianna, Cíntia Coelho, André Murad, entre outras pessoas que passaram um período no laboratório e ajudaram na rotina e com ensinamentos.

Agradeço também a todos os meus amigos e à minha família que souberam entender a minha ausência por várias vezes, e mesmo assim, apoiavam-me sempre.

Quero agradecer ao professor David Lee Kaplan, por ter me recebido no período de doutorado sanduíche na *Tufts University*, e pelos ensinamentos. A capacidade atingida e os limites ultrapassados foram adquiridos de um modo que eu não sabia que era possível. Muito obrigada.

Gostaria de agradecer aos colaboradores Cameron Copeland e à Paula E. Faria e Oliveira da Universidade Estadual de Utah. E também ao professor Randy Lewis pela colaboração nas análises das fibras sintéticas.

Quero agradecer à analista Silviane Zanni Hubinger e aos pesquisadores Luciamara Aparecida Forato e Luiz Alberto Colnago da Embrapa Instrumentação, pelas análises de FTIR-ATR realizadas.

Agradeço à Ingrid Gracielle Martins da Silva, ao João e à professora Sonia Nair Bão, pela utilização do MEV, no laboratório de microscopia eletrônica, Instituto de Biologia, na Universidade de Brasília.

Agradeço ao pesquisador André Murad pela colaboração nas análises proteômicas e de espectrometria de massa.

Agradeço ao pesquisador Luciano P. Silva pela colaboração em várias etapas do trabalho, entre elas as análises de espectrometria *biotyper* e MFA. Muito obrigada!

Agradeço a todos os professores, mas em especial, Lídia Pepe, Ildinete Pereira Marcelo Brígido e os pós-doutorandos Rafael Burtet e Marciano Rubini pela experiência no período de estágio em docência. Foi gratificante!

Agradeço ao departamento de Pós-graduação em Biologia Molecular-UnB pelo curso e pela oportunidade em ajudar na organização de um Simpósio juntamente com outros alunos e o coordenador do curso, Robert Miller.

Agradeço à CAPES pela bolsa de estudos em território nacional, e ao CNPq pela bolsa de estudo em território internacional.

Agradeço à FAP-DF pelo apoio financeiro para a participação do congresso internacional da Sociedade Européia de Biomateriais em Liverpool, Inglaterra, 2014.

Agradeço ao meu orientador Elíbio Leopoldo Rech Filho, que sempre estava de portas abertas para novas ideias e discussões.

Agradeço à Embrapa Recursos Genéticos e Biotecnologia pelo espaço utilizado durante o trabalho e por proporcionar um ambiente científico.

RESUMO

As fibras de teia de aranha são polímeros com propriedades elásticas e de resistência, simultaneamente. Eles são compostos por alta repetição de uma sequência inicial (motivos), ricos em aminoácidos glicina, alanina e prolina e de alta massa molecular (300-500 kDa), sendo uns dos desafios em sistema heterólogos de produção. O metabolismo de bactérias pode ser modificado por biologia sintética para produzir essas espidroínas e respectivas fibras sintéticas. Além disso, os estudos concentram-se principalmente em duas espécies de aranhas norte-americanas, sendo necessário agregar valor em produtos obtidos da biodiversidade brasileira. O presente estudo concentrou-se na síntese de espidroínas da aranha *Parawixia bistriata* e em alterar o metabolismo da *Escherichia coli* aumentando a disponibilidade de glicil-tRNA. Essas novas linhagens de bactérias foram avaliadas por curva de crescimento, classificadas por perfil proteômico e avaliadas quanto ao rendimento médio das espidroínas. Posteriormente, as proteínas MaSp2 16× (54 kDa) e 32× (105 kDa) foram caracterizadas molecular e estruturalmente por MALDI-TOF, MEV, enquanto as fibras sintéticas foram caracterizadas funcional e morfológicamente por MFA, MEV e ATR-FTIR e avaliadas por testes mecânicos. As novas linhagens de bactéria apresentam uma velocidade de crescimento semelhante às linhagens comerciais, sendo comprovada a diferença dos perfis proteômico. Além disso, as espidroínas não afetaram o crescimento das bactérias, e o rendimento da MaSp2 32× com alta massa molecular foi maior para algumas bactérias modificadas metabolicamente ($p < 0,01$). As fibras sintéticas, formadas em banho de coagulação com isopropanol, apresentam a superfície lisa (MEV), semelhante à natural, mas com padrão em nanoescala heterogêneo (MFA). Essas fibras após serem hidratadas e estiradas resultam num biomaterial com tendências a melhorar suas propriedades mecanoelásticas e na presença de folha- , mostrando a importância de se processar a fibra após a sua formação. Portanto, as linhagens de bactérias modificadas por engenharia metabólica foram capazes de aumentar o rendimento da espidroína MaSp2 de alta massa molecular (105 kDa), sintetizadas pela primeira vez nesse trabalho. As fibras sintéticas formadas apresentam uma morfologia topográfica semelhante à natural, mas ainda com propriedades mecânicas inferiores. Apesar disso, o processamento das fibras sintéticas após a sua formação são etapas importantes para melhorar a qualidade desse biomaterial.

Palavras chave: Fibras, espidroínas, MaSp2, MFA, FTIR-ATR, MEV, *E. coli*.

ABSTRACT

Spider silk fibers are polymers with elastic properties and resistance simultaneously. They consist of an initial high repetition sequence (motifs), rich in amino acids glycine, alanine and proline with high molecular mass (300-500 kDa), one of the challenges in heterologous production system. The bacterial metabolism can be modified by synthetic biology to produce these spidroins and spun fibers. Also, the studies are concentrated on two species of North American spiders, being necessary to add value in products obtained from Brazilian biodiversity. These aims were spider *Parawixia bistriata* spidroins synthesis and modify the metabolism of *Escherichia coli* increasing the glycyl-tRNA. These new strains of bacteria were evaluated by growth curve, sorted by proteomic profile and evaluated as the average income of spidroins. Subsequently, the MaSp2 16 × protein (54 kDa) and 32 × (105 kDa) were characterized structurally by molecular by MALDI-TOF and SEM, while synthetic fibers are functionally and morphologically characterized by AFM, ATR-FTIR and SEM and evaluated by mechanical tests. The new bacteria strains have a growth rate similar to commercial strains, being proven the difference of proteomic profiles. Furthermore, spidroins did not affect the growth of bacteria, and the yield of MaSp2 32 × high molecular mass were higher for some metabolically modified bacteria ($p < 0.01$). Synthetic fibers, spun in coagulation bath with isopropanol, have smooth surface (SEM), similar to natural, but with standard heterogeneous nanoscale (MFA). These fibers after being hydrated and stretched results in a biomaterial with trends improving their elastic mechanic properties in the presence of β -sheet, showing the importance of processing the fiber after its formation. Therefore, the strains of bacteria modified by metabolic engineering were able to increase the yield of spidroins MaSp2 high molecular weight (105 kDa), synthesized for the first time in this study. The synthetic fibers have formed a topographical morphology similar to natural, but with lower mechanical properties. Despite this, processing synthetic fibers after spun are important steps to improve the quality of this biomaterial.

Keywords: Fibers, Spidroins, MaSp2, MFA, FTIR-ATR, MEV, *E. coli*.

LISTA DE FIGURAS

- Figura 1.** Desenho ilustrativo de uma teia orbital de aranha composta de diferentes proteínas. As molduras coloridas indicam as estruturas a serem formadas pelos motivos. A interrogação próxima à região espaçadora significa que a estrutura secundária não é conhecida. MaSp: Espidroína da glândula ampolada maior; MiSp: Espidroína da glândula ampolada menor; Flag: proteína flageliforme. Adaptada de TOKAREVA et al., 2013 e SLOTTA et al., 2012. **3**
- Figura 2.** Estrutura interna de uma seda de aranha linha de segurança. O modelo multicamadas divide cinco partes, do externo para o interno: camada de lipídeos, glicoproteínas, película, núcleo externo e núcleo interno. Modificada de SPONNER et al., 2007. **4**
- Figura 3.** Ilustração da formação natural das fibras de aranha. As espidroínas são produzidas e secretadas na região da cauda e armazenadas na ampola em altas concentrações (~50%). No ducto de formação ocorrem trocas iônicas, acidificação e alongamento. Antes de encontrar o meio externo, a fibra passa por uma válvula, que auxilia na formação de uma nova fibra, que ocorre por puxação. (Adaptado de EISOLDT et al., 2011). **6**
- Figura 4.** Aranha jovem da espécie *Parawixia bistrriata* construindo sua teia ao entardecer em ambiente periurbano. Foto adquirida por Valquíria Lacerda próximo ao Prédio de Biotecnologia, Recursos Genéticos e Biotecnologia, Brasília, DF, Brasil. **10**
- Figura 5.** Desenho esquemático resumindo o projeto desenvolvido. A- Alteração por engenharia metabólica de *E. coli* para aumentar a expressão de glicil-tRNA. pACYC184: plasmídeo de clonagem (NEB); pTetglyVXY: plasmídeo contendo uma cópia do gene *glyVXY*; pTetgly2: plasmídeo contendo duas cópias do gene *glyVXY*, pTetgly-*glyA*: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTetgly2-*glyA*: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. B- Ilustração da estratégia para obtenção de espidroínas e respectivas avaliações. A linha rosa indica as construções a serem utilizadas. MaSp: *Major Ampullate Spidroin*, PBSSK+: pBluescriptIISK+, n×: número de motivos presentes no plasmídeo. pET19b e p cong1: plasmídeos de expressão em *E. coli* e soja (*Glycine max*), respectivamente; pAC321: plasmídeo de seleção para soja. MFA: Microscopia de Força Atômica. **14**
- Figura 6.** Desenho representativo de proteínas de teia de aranha. As proteínas naturais apresentam extremidades C-terminal e N-terminal, com a parte central constituída de motivos altamente repetitivos. As proteínas sintéticas para sistemas heterólogos não precisam das sequências C-terminal e N-terminal. Adicionalmente, pode-se inserir sequências que auxiliam o processo de purificação proteica, como a sequência de histidina. **15**
- Figure 7.** Desenho representativo dos plasmídeos pJ, sintetizados pela DNA2.0 (EUA) contendo os motivos sintéticos da espécie *P. bistrriata*. e as enzimas de restrição a serem utilizadas na metodologia de produção em *E. coli*. O pJ241 contém o motivo da MaSp1(2×)+2(1×), enquanto o pJ 201 contém o motivo da MaSp2. O tamanho em pares de bases encontra-se ao centro de cada plasmídeo. **18**
- Figure 8.** Desenho representativo dos plasmídeos pJ, sintetizados pela DNA2.0 (EUA) contendo os motivos sintéticos da espécie *P. bistrriata*. e as enzimas de restrição a serem utilizadas na metodologia de produção em soja (*Glycine max*) O pJ201:27397 contém o motivo da MaSp1, enquanto o pJ 201:27396 contém o motivo da MaSp2. O tamanho em pares de bases encontra-se ao centro de cada plasmídeo. **19**

- Figure 9.** Representação da sequência peptídica e nucleotídica do peptídeo sinal da -coixina. Adaptado de Cunha (2010). **21**
- Figure 10.** Desenho esquemático dos plasmídeos construídos para aumentar a disponibilidade de glicil-tRNA e de serina hidroximetiltransferase no metabolismo de *E. coli*, utilizando o plasmídeo original pACYC184. As imagens foram montadas no programa NTIVector (versão 1 4.0, Invitrogen)..... **22**
- Figura 11.** Desenho representativo para as construções dos plasmídeos contendo os motivos das espidroínas. O motivo inicial da aranha *Parawixia bistriata* (em verde) contém sequências de enzimas de restrição *XhoI* (amarelo), *NdeI* (azul), *XmaI*(vermelho), *BspEI*(dourado) e *BamHI* (rosa). O motivo é inserido no plasmídeo de clonagem pBSSK+ pelos sítios das enzimas *XhoI* e *BamHI* (estrelas rosa e amarela no plasmídeo pBluescriptIISK+), resultando no plasmídeo contendo uma vez o motivo (pBSSK1×). Em seguida, duas alíquotas independentes desse plasmídeo sofrem digestões duplas (A): uma é digerida com as enzimas *ScaI* e *XmaI* (plasmídeo superior-estrela preta à estrela vermelha), enquanto a outra é digerida com *ScaI* e *BspEI* (plasmídeo inferior-estrela preta à estrela amarela). Apenas os fragmentos que possuem o motivo (em verde) em sua sequência são selecionados (A-fragmentos ao centro, entre os plasmídeos inferior e superior). A reação de ligação (B) reestabelece o gene de resistência à ampicilina (sítio de *ScaI*-estrelas pretas) e une as extremidades formadas for *XmaI* e *BspEI* (estrelas vermelha e amarela), mas sem reestabelecer os respectivos sítios de restrição (B-parte cinza entre duas partes verdes do plasmídeo superior). Esse plasmídeo contém o dobro de motivos do plasmídeo inicial. Esses procedimentos (A e B) podem ser repetidos inúmeras vezes até se atingir a quantidade de motivos desejada. Logo após, o fragmento contendo os n× motivos das espidroínas é retirado do plasmídeo de clonagem (B-plasmídeo superior) por digestão das enzimas *BamHI* e *NdeI*(estrelas rosa e azul) e introduzido no plasmídeo de expressão em *E.coli* pET19b (B-plasmídeo inferior), formando o plasmídeo de expressão contendo a sequência de espidroínas (C). Esse apresenta uma sequência codante de histidina na porção N-terminal. (Adaptado de TOKAREVA et al., 2013). **29**
- Figura 12.** Fluxograma ilustrativo do experimento para a curva de crescimento das bactérias. A linhagem BL21(DE03) foi transformada com os plasmídeos construídos para a engenharia metabólica de bactéria, conforme o quadrado à esquerda, sendo também co-transformadas com os vetores de expressão pET19b, pET19bMaSp2 16× e pET19bMaSp2 32×, circunferências ao centro. O crescimento com e sem indução de IPTG 0,4 mM foi acompanhado durante 10 horas, para todas as linhagens, em triplicata. A bactéria BL21(DE 03), os plasmídeos pACYC184 e pET19b foram utilizados como controles..... **33**
- Figura 13.** Desenho representativo da metodologia utilizada no cultivo das linhagens bacterianas para as análises por espectrometria de massa *biotyper*. As colônias de bactérias foram isoladas no primeiro dia em meio LB Agar a 37 °C (Placa 1), selecionadas e repicadas em 5 pontos nas mesmas condições por mais 48 h (Placas 2 e 3). Uma das colônias foi repicada em 30 pontos em outra placa por mais 24h (Placa 4), sendo que 24 foram transferidas para um tubo de prolipropileno para análise. **56**
- Figura 14.** Análise eletroforética em gel de agarose corado com brometo de etídio. A) Fragmento amplificado por PCR referente ao gene *glyVXY* (479 pb). B) Produto de digestão com enzima *EcoRI* do plasmídeo PGEMTEasy (3.000 pb) contendo o gene *glyVXY* (497 pb). Marcador: 1 Kb Ladder da Promega..... **67**
- Figura 15.** Análise eletroforética em gel de agarose corado com brometo de etídio. A) Produto das digestões das enzimas *BamHI* e *EcoRV*. 1- 1Kb Ladder Promega, 2-

- pACYC184, 3- PGEMT-Easy+ *glyVXY*. B) Confirmação do plasmídeo pTet*glyVXY* digerido com *Bam*HI e *Eco*RV (463 pb), *Hinc*II (1.066 pb) e *Nhe*I (4.520 pb)..... **68**
- Figura 16.** Análise eletroforética em gel de agarose corado com brometo de etídio. A) 1- 100 pb Invitrogen. 2-Gene *glyVXY* amplificado por PCR. B)1- 100 pb Invitrogen, 2- PGEMT-Easy+ *glyVXY* digerido com *Eco*RI, C) 1- 100 pb Invitrogen , 2 e 3- PGEMT-Easy+ *glyVXY* e pTet*glyVXY* digeridos com as enzimas *Sal*I e *Sph*I;. D) 1- 1Kb Ladder Promega, 2 e 3- plasmídeo pTet*gly2* digerido com *Sal*I e *Sph*I; *Bam*HI e *Eco*RV. **68**
- Figura 17.** Análise eletroforética em gel de agarose corado com brometo de etídio. A) 1- 1Kb Ladder Promega. 2-Produto da PCR referente ao gene *glyA*. B)1-PGEMT-Easy (3.000 pb)+ *glyA* (1.752 pb)digerido com *Eco*RI, 2- 1Kb Ladder Promega. **69**
- Figura 18.** A) Análise eletroforética em gel de agarose 1 % corado com brometo de etídio da amplificação por PCR do gene *glyA*. 1-1 Kb *Ladder Plus* Invitrogen, 2- 4 Gene *glyA* amplificado com $MgSO_4$ 1mM; 1,5; 2,0 mM, 5-branco. B).1-1Kb *Ladder Plus* Invitrogen. 2- Plasmídeo pTet*gly2* digeridos com as enzimas *Sal*I e *Hinc*II (4.025 pb), e o fragmento de 1.066 pb que foi descartado. 3-Plasmídeo pTet*gly2* digerido com as enzimas *Sal*I (5.091 pb). **70**
- Figura 19.** Análise eletroforética em gel de agarose 1% corado com brometo de etídio. A) Plasmídeos na sequência: pACYC184, pTet*glyVXY* e pTet*gly2* digeridos com: 2-4: *Bam*H1 e *Eco*RV; 5-7: *Sal*1; 8-10: *Hinc*II; 11-13: *Sal*I e *Bam*HI. B) Plasmídeo pTet*glyVXY-glyA* digerido com *Hinc*II, *Eco*RV, *Hind*III, *Sal*I, *Bam*HI e padrão não digerido. C) Plasmídeo pTet*gly2-glyA* digerido com 1: *Bam*HI e *Eco*RV, 2: *Sal*I e *Sph*I; 3- *Hinc*II; 4:1 Kb Ladder NEB. **70**
- Figura 20.** Análise eletroforética em gel de agarose 1% corado com brometo de etídio dos produtos das reações de PCR para as linhagens bacterianas. Da esquerda para a direita: BL21(DE03); BL21(DE03)pACYC184; BL21(DE03)pTet*glyVXY*; BL21(DE03)pTet*gly2*; BL21(DE03)pTet*glyVXY-glyA*; BL21(DE03)pTet*gly2-glyA*. A) Amplificação do gene *glyVXY*: 1-6 linhagens na ausência do plasmídeo pET19b; 7-12 linhagens na presença do plasmídeo pET19b sem inserto; 13-18: linhagens na presença do plasmídeo pET19b MaSp2 16×; 19-24: linhagens na presença do plasmídeo pET19bMaSp 32×. 25: controle positivo; 26-branco. B) Amplificação do gene *glyA*: 1-6 sem pET19b; 7-12 com pET19b; 13-18: com pET19b MaSp2 16×; 19-24: com pET19bMaSp 32×; 25:branco, 26:controle positivo. **71**
- Figura 21.** Alinhamento da sequência MaSp2 2× de *P. bistrata*. A sequência superior equivale às construções *in silico*, enquanto a inferior refere-se ao resultado do sequenciamento. A sequência sublinhada refere-se a ligação dos produtos das digestões das enzimas *Bsp*EI e *Xma*I. Identidade 100%. **72**
- Figura 22.** Alinhamento genético do plasmídeo contendo a MaSp1+2 de *P. bistrata* (linha inferior) com a sequência *in silico* enviada para síntese (linha superior). Identidade 100%. **73**
- Figura 23.** Análise eletroforética das construções de MaSp1. A) Plasmídeo pBSSK+ contendo 16× o motivo.1- Digestão com as enzimas *Xma*I e *Sca*I. 2- Digestão com as enzimas *Sca*I e *Bsp*EI. B) Plasmídeo pBSSK+ contendo 32× o motivo.1- Digestão com as enzimas *Sca*I e *Bsp*EI. 2- Digestão com as enzimas *Xma*I e *Sca*I. C) Plasmídeo pET19b contendo 32× o motivo digerido com *Apa*I e *Bam*HI; 1, 2 e 4 são amostras esperadas. D) Plasmídeo pET19b, contendo 32× o motivo, digerido com *Apa*I, *Bam*HI e *Sca*I. **74**
- Figura 24.** Análise eletroforética em gel de agarose corado com brometo de etídio referente às construções do módulo MaSp2 da aranha *Parawixia bistrata*.digeridos com *Bam*HI e *Xho*I.A-F: plasmídeos de clonagem contendo 1, 2, 4, 8, 16 e 32 motivo(s) de MaSp2. .. **75**

- Figura 25.** Análise eletroforética em gel de agarose corado com brometo de etídio referente às construções do módulo MaSp2 da aranha *Parawixia bistriata*. digeridos com *ScaI* e *XmaI* e também com *ScaI* e *BspEI*. A-F: plasmídeos de clonagem contendo 1, 2, 4, 8, 16 e 32 motivo(s) de MaSp2. 76
- Figura 26.** Análise eletroforética de plasmídeos contendo espidroínas MaSp2 de *Parawixia bistriata*. A) Plasmídeo pET19b MaSp2 16× digeridos com *BamHI* e *NdeI*: 1- 1Kb Ladder (Invitrogen); 2- Plasmídeo sem inserto; 3-4 Plasmídeo contendo a sequência da espidroína (1.940 pb). B) Plasmídeo pBSSK+ MaSp 32× digeridos com *BamHI* e *NdeI*: 1- 1Kb Ladder (Invitrogen); 2- Plasmídeo contendo a sequência da espidroína (3.765 pb). 76
- Figura 27.** Análise eletroforética, em gel de agarose corado com brometo de etídio referente às construções de MaSp1+2. A) Plasmídeo pJ. B-C Plasmídeo pBSSK+ digeridos com as enzimas *BamHI* e *XhoI* contendo o motivo 2× (723 pb), 4× (1.450 pb), respectivamente. D) Plasmídeo pBSSK+ digeridos com as enzimas *BamHI*, *XhoI* e *ScaI* contendo o motivo 8× (2.890 pb). 77
- Figura 28.** Análise eletroforética, em gel de agarose corado com brometo de etídio referente às construções da MaSp1+2. A) 1- 1Kb Ladder (Invitrogen); Plasmídeo pBSSK+ com 1× o motivo digerido com: 2- *XmaI* e *ScaI* (2.167pb), 3- *ScaI* e *BpsEI* (1.466 pb). B) 1- 1Kb Ladder (Invitrogen); 4- 100 pb Ladder (Invitrogen). Digestão do plasmídeo pBSSK+ 2× o motivo com as enzimas: 2- *ScaI* e *BpsEI* (1.817 pb). 3- *XmaI* e *ScaI* (2.516 pb), 5- *ScaI*, *BpsEI*, *AseI* (2.890 pb). E) 1- 1Kb Ladder (Invitrogen). Digestão do plasmídeo pBSSK+ 4× o motivo com as enzimas: 2- *ScaI* e *BpsEI* (2.519 pb). 3- *XmaI* e *ScaI* (3.220 pb). F) 1- 1Kb Ladder (Invitrogen). Digestão do plasmídeo pBSSK+ 8× o motivo com as enzimas: 2- *BspEI* e *ScaI* (3.923 pb). 78
- Figura 29.** Análise eletroforética em gel de agarose corado com brometo de etídio. Plasmídeos digeridos com *XmaI* e *BspEI*. 1- 1Kb Plus Promega, 2- pJ; 3- pET19b MaSp2 16×; 4- pET19b MaSp2 32×. 80
- Figura 30.** Análise eletroforética das digestões dos plasmídeos pJMaSp2 16× e pJMaSp2 32×. 1-3: MaSp2 16× *KpnI* e *EcoRI* (1.959 pb), 4 e 10- 1 Kb Ladder (Invitrogen), 5-6: MaSp2 32× *KpnI* e *EcoRI* (3.879 pb). 7-9: MaSp2 16× *HindIII* e *EcoRI* (1.960 pb), 11-13: MaSp2 32× *HindIII* e *EcoRI* (3.879 pb). 81
- Figura 31.** Gráficos das curvas de crescimento das linhagens de *E. coli* em meio LB pelo período de 10 horas. As induções foram realizadas com IPTG 0,4 mM com OD₆₀₀ 0,6. As bactérias BL21(DE03) e a BL21(DE03)pACYC184 foram utilizadas como controles da cepa e de plasmídeo inalterado. MaSp2: Espidroína do tipo 2 da glândula ampolada maior. pTetglyVXY: plasmídeo contendo uma cópia do gene *glyVXY*; pTetgly2: plasmídeo contendo duas cópias do gene *glyVXY*, pTetgly-glyA: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTetgly2-glyA: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. pET19b: plasmídeo de expressão em bactéria. 84
- Figura 32.** Dendograma das linhas bacterianas submetidas à análise por espectrometria de massa *biotyper*. MSP: main-spectrum. As linhagens alteradas metabolicamente foram classificadas em um grupo, enquanto as linhagens comerciais (BL21(DE03), BL21(DE03)pLysS e a linhagem utilizada como controle BL21(DE03)pACYC184 foram classificadas em outro grupo. pTetglyVXY: plasmídeo contendo uma cópia do gene *glyVXY*; pTetgly2: plasmídeo contendo duas cópias do gene *glyVXY*, pTetgly-glyA: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTetgly2-glyA: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. 86
- Figura 33.** Gráfico ilustrativo da quantificação de proteína total em cada linhagem utilizada nesse experimento. A quantificação foi realizada por absorvância 280 nm em espectrofotômetro nanodrop. pTetglyVXY: plasmídeo contendo uma cópia do gene

- glyVXY*; pTet*gly2*: plasmídeo contendo duas cópias do gene *glyVXY*, pTet*gly-glyA*: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTet*gly2-glyA*: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. pET19b: plasmídeo de expressão em bactéria. MaSp2: Espidroína do tipo 2 da glândula ampolada maior. 87
- Figura 34.** Gráfico de distribuição referente às proteínas encontradas: A) apenas em BL21(DE03). B: proteínas encontradas apenas na linhagem BL21(DE03)pACYC184. BP: processo biológico. MF: função molecular. CC: componente celular. Blast2Go®. . 88
- Figura 35.** Gráfico de distribuição referente às proteínas encontradas: A) produção aumentadas em BL21(DE03)pACYC184 vs BL21(DE03). B: produção diminuída em BL21(DE03)pACYC184 vs BL21(DE03). BP: processo biológico. MF: função molecular. CC: componente celular. Blast2Go®. 89
- Figura 36.** Análise de SDS-PAGE em gel pré-moldado 4- 12% Novex (Invitrogen) utilizando tampão MOPS para determinar a solubilidade das espidroínas durante a extração de proteínas. Os extratos das amostras contendo o precipitado (P1 e P2) e o extrato solúvel (S2) indicam que as espidroínas encontram-se na porção solúvel. Setas indicam a proteína alvo. A) MaSp2 16× (54 kDa); B) MaSp 32× (105 kDa), Marcador: *Novex sharp pre stain protein standard* (Invitrogen). 91
- Figura 37.** A) Análise por SDS-PAGE do processo de purificação da MaSp2 16×. FT: passado; L1-2: Lavado com Imidazol 40 mM; E1-2: Eluído com Imidazol 100 mM; E3-4: Imidazol 250 mM. Proteína com massa molecular de 54 kDa. B) Espidroína purificada utilizando diferentes linhagens de expressão. C) Identificação das espidroínas por meio da sequência de histidina por *Western blot*. As setas em B e C indicam a proteína alvo, e abaixo delas, são espidroínas menores que foram produzidas pela bactéria devido a erros na transcrição e tradução. 93
- Figura 38.** Gráficos representativos dos resultados de rendimento da espidroína sintética de MaSp2 16× (54kDa) de *Parawixia bistriata* em diferentes linhagens bacterianas. A) Biomassa de bactéria ao final da indução, $p < 0,01$, CV=13,68%. B) Rendimento de purificada a partir da indução em um litro, $p < 0,05$, CV= 35,87%. C) Quantidade de espidroína purificada por biomassa de bactéria, $p < 0,01$, CV=27,38%. pTet*glyVXY*: plasmídeo contendo uma cópia do gene *glyVXY*; pTet*gly2*: plasmídeo contendo duas cópias do gene *glyVXY*, pTet*gly-glyA*: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTet*gly2-glyA*: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. As linhagens BL21(DE03) e BL21(DE03)pACYC184 serviram como controle. 96
- Figura 39.** Fotografias das análises por SDS-PAGE e Western blot da MaSp2 32×. A) SDS-PAGE processo de purificação FT: passado; L1-2: Lavado com Imidazol 40 mM; E1-2: Eluído com Imidazol 100 mM; E3-4: Imidazol 250 mM. B) *Western blot* das espidroínas detectadas em extrato proteico total da bactéria. 1- *ColorBurst™ Electrophoresis Marker* (C1992-SIGMA), 6 e 7: amostras contendo a MaSp2 32× com a sequência de histidina N-terminal. C) *Western blot* para identificação da MaSp2 32× após o processo de purificação, 1-3: Linhagens utilizadas como controle negativo, 4- *Novex sharp pre stain protein standard* (Invitrogen), 5-7: Espidroína produzida em diferentes linhagens. As setas indicam a proteína com massa molecular de 105 kDa. 98
- Figura 40.** Gráficos representativos dos resultados de rendimento da espidroína sintética de MaSp2 32× (105 kDa) de *Parawixia bistriata* em diferentes linhagens bacterianas. A) Biomassa de bactéria ao final da indução, $p < 0,05$. B) Rendimento de purificada a partir da indução em um litro, $p < 0,01$, CV= 15,6%. C) Quantidade de espidroína purificada por biomassa de bactéria, $p < 0,01$, CV=13,72%. pTet*glyVXY*: plasmídeo contendo uma cópia do gene *glyVXY*; pTet*gly2*: plasmídeo contendo duas cópias do gene *glyVXY*, pTet*gly-glyA*: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTet*gly2-glyA*: plasmídeo

- contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. As linhagens BL21(DE03) e BL21(DE03)pACYC184 serviram como controle.100
- Figura 41.** Gráficos representativos das análises realizadas de MaSp2 16× (~54 kDa) por espectrometria de massa MALDI-TOF. As setas pretas indicam os picos referentes à massa molecular. A espidroína purificada foi identificada após a produção em diferentes linhagens de bactérias (A,B,C), e em dois modos de alcance: o de espectros de peptídeo/proteínas de baixa massa molecular (D) e o de detecção de proteínas maiores (E).102
- Figure 42.** Fotografias das fibras formadas a partir das espidroínas de MaSp2 16× (54kDa). A) Fragmentos da fibra sintética; B) Fibra sintética formada e depositada ao redor da área externa de um pedaço de rolo de isofilme transparente. Foto: Arquivo pessoal, Valquíria Michalczchen-Lacerda.....105
- Figura 43.** Gráficos representativos das análises de rugosidade das fibras sintéticas não tratadas (A-F) e tratadas (G-L) por MFA em modo contato intermitente (3 imagens para cada fibra)(n=10). Ra: Rugosidade média aritmética; Rq Rugosidade quadrática média; Rzjis: Rugosidade média de 10 pontos; Rp: Altura média; Rv: Profundidade média; Rz: Altura máxima. Letras diferem entre si com $p < 0,05$; * diferem entre si com $p < 0,01$. ..108
- Figura 44.** Gráfico de medidas de rugosidade das fibras sintéticas. FNT: fibras não tratadas. FT: Fibras tratadas. As letras são comparadas apenas entre os grupos das fibras, não apresentando informação para os diferentes tipos de parâmetros de análise. Ra: Rugosidade média aritmética; Rq Rugosidade quadrática média; Rzjis: Rugosidade média de 10 pontos; Rp: Altura média; Rv: Profundidade média; Rz: Altura máxima. Letras diferem entre si com $p < 0,05$109
- Figura 45.** Gráfico representativo da análise de espectroscopia de força por MFA. FNT:Fibras não tratadas; FT: Fibras tratadas. As letras são comparadas apenas entre os grupos das fibras, não apresentando informação para os diferentes tipos de parâmetros de análise. Ra: Rugosidade média aritmética; Rq Rugosidade quadrática média; Rzjis: Rugosidade média de 10 pontos; Rp: Altura média; Rv: Profundidade média; Rz: Altura máxima. Módulo de Young = módulo de elasticidade. Letras diferem entre si com $p < 0,05$110
- Figura 46.** Gráfico representativo da análise de viscoelasticidade da superfície das fibras sintéticas por MFA. FNT: Fibras não tratadas; FT: Fibras tratadas. As letras são comparadas apenas entre os grupos das fibras, não apresentando informação para os diferentes parâmetros de análise.111
- Figura 47.** Imagens adquiridas por MFA de fibras sintéticas produzidas a partir da proteína MaSp2 16× (54 kDa) de *Parawixia bistriata*. A área de varredura foi de $25 \mu\text{m}^2$. A-C: Análises de fibras sintéticas não tratadas após a extrusão; D-F: Análises de fibra sintética tratadas após a extrusão. A e D: digitalizações topográficas de altura em 3D; B e E: digitalizações de viscoelasticidade das mesmas áreas avaliadas em A e D; C e F: Gráficos representativos da curva de força.113
- Figura 48.** Fotomicrografia de espidroínas da aranha *P. bistriata* após liofilização, adquirida em microscópio eletrônico de varredura. A,B,C- Proteína MaSp2 16× (54 kDa) e D,E,F) Proteína MaSp2 32× (105 kDa) com aumento de 190×, 430× e 3.000× respectivamente. Barras em A e D: 100 μ ; B e E: 10 μm ; C e F: 1 μm . As setas indicam torções nas estruturas de microfibras proteicas.115
- Figura 49.** Análise das fibras produzidas a partir da proteína MaSp2 (54 kDa) de *Parawixia bistriata* por MEV. A) Fotomicrografia da fibra sintética após sua formação, com 43,6 μm de diâmetro; B) Fotomicrografia da fibra sintética após tratamento com isopropanol, com 31,8 μm de diâmetro; C) Gráfico representativo das médias de diâmetros dos dois

- grupos de fibras, $p < 0,05$ ($n=10$); D) Média dos diâmetros de cada fibra após a sua formação, e E) das fibras após o tratamento com isopropanol ($n=3$). **117**
- Figura 50.** Fotografia das fibras sintéticas, armazenadas em papel cartão preto em placas de petri de acrílico. A) Não tratadas. B) Hidratadas; C) Hidratadas e com óleo mineral; D) Hidratadas e esticadas; E) Hidratadas, esticadas e com óleo mineral. Fotos: Arquivo pessoal. **118**
- Figura 51.** Gráficos representativos dos diâmetros médios das fibras sintéticas de MaSp2 16× (54 kDa). C: não tratadas; H: hidratadas; HO: hidratadas e com óleo; EH: hidratadas e esticadas; EHO: hidratadas, esticadas e óleo. A) Comparação entre os cinco grupos. Letras diferem entre si em $p < 0,01$. B) Comparação entre os grupos sem tratamento e com hidratação. Letras diferem entre si em $p < 0,05$ **119**
- Figura 52.** Gráficos representativos dos testes mecânicos realizados nos cinco grupos de fibras sintéticas MaSp2 16×x (54 kDa). A) Energia de rompimento, B) Tensão máxima, C) Deformação máxima. Os grupos das fibras foram denominados como C: não tratadas; H: hidratadas; HO: hidratadas e com óleo; EH: hidratadas e esticadas; EHO: hidratadas, esticadas e óleo. **120**
- Figura 53.** Gráficos representativos de tensão máxima vs deformação máxima antes do rompimento de fibras sintéticas de espidroínas MaSp2 16× de *Parawixia bistrata*. Cada linha representa o resultado de uma fibra. Os grupos de fibras analisadas foram :A) Não tratadas ($n=7$), B) Hidratada em isopropanol 80% ($n=8$), C) Hidratada em isopropanol 80% e com uma camada de óleo mineral ($n=10$), D) Hidratada em isopropanol 80% e estirada ($n=10$), E) Hidratada em isopropanol 80%, estirada e com uma camada de óleo mineral ($n=10$). **121**
- Figura 54.** Gráfico representativo das análises por ATR-FTIR (transmitância vs n° da onda). As fibras sintéticas são de MaSp2 16× (54 kDa). FT: fibras tratadas. FNT: fibras não tratadas. A) Resultados com amplo campo das análises. B) Campo aproximado onde é possível uma melhor visualização da absorção referente a folhas- (1.670 cm^{-1}). As setas indicam padrões de amida I ($1.600-1.700 \text{ cm}^{-1}$) e amida II ($1.600-1.500 \text{ cm}^{-1}$). Linha preta: espidroína liofilizada MaSp2 16×. Linha vermelha: espidroína liofilizada MaSp2 32×. Linha azul: fibras tratadas. Linha rosa: fibras não tratadas. **123**

LISTA DE TABELAS

Tabela 1. Propriedades mecânicas das linhas de segurança obtidas de teia orbital de diferentes espécies (Adaptado de SLOTTA et al., 2012).....	5
Tabela 2. Propriedades mecânicas de alguns materiais (Adaptado de ROMER; SCHEIBEL, 2008).....	5
Tabela 3. Tamanho dos motivos sintéticos iniciais da espécie <i>P. bistriata</i> e percentual dos principais aminoácidos constituintes.....	20

LISTA DE ABREVIATURAS E SIGLAS

Amp	Ampicilina
ADF	Fibroina(s) de aranus diadematus
Ala	Alanina
Arg-C	Endopeptidase Arg-C
Asp-N	Endoproteinase Asp-N
BCIP	5-Bromo-4-Cloro-3-Indolil Fosfato
BSA	Albumina Sérica Bovina
C-terminal	Carboxi Terminal
Cl	Cloro
Comassie Blue	Reagente azul utilizado para corar géis de acrilamida
CV	Coefficiente de Variação
dNTP	Desoxirribonucleotídeos fosfatados
GPa	Gigapascal
TBE	Solução tampão contendo Tris, Ácido Bórico e EDTA
DNA	Ácido desoxirribonucleico
DTT	Ditiotreitol
EDTA	Ácido etilenodiaminotetraacético
Flag	Espidroína flageliforme
f	Femto
FH	Fibra(s) Hidratada(s)
FHE	Fibra(s) Hidratada(s) e Estirada(s)
FHO	Fibra(s) Hidratada(s) e com Camada de Óleo Mineral
FHEO	Fibra(s) Hidratada(s) Estirada(s) e com Camada de Óleo
FNT	Fibra(s) Não Tratada(s)
FT	Fibra(s) Tratada(s)
FTIR-ATR	Espectroscopia com infravermelho por transformada de fourier–reflectância total atenuada
g	Gramma
g	Força gravitacional
Gli	Glicina
Glicil-tRNA	Ácido ribonucleico transportador de glicina
Gln	Glutamina
Glu-C	Endoproteinase Glu-C
h	Hora
HCl	Ácido Clorídrico
Hz	Hertz (unidade de frequência)
HFIP	Hexafluorisopropanol
His	Histidina
IPTG	Isopropil -D-1-tiogalactopiranosídeo
Kb	Quilobase(s)
kDa	Quilodalton(s)
LB	meio Luria-Bertani
Leu	Leucina
Lys-C	Protease Lys-C
m	Massa (g)
m	Metro(s)

MALDI-TOF	Espectrometria de massa com analisador de tempo de voo e ionização a laser assistida por matriz.
MaSp	Espidroína de glândula ampolada maior
Maxi-prep	Preparação de plasmídeos em grande escala
MEV	Microscopia eletrônica de varredura
mg	Miligrama
MgCl ₂	Cloreto de magnésio
Mg ₂ SO ₄	Sulfato de magnésio
MFA	Microscopia de força atômica
min.	Minuto(s)
Misp	Glândula ampolada menor
MJ	Megajoule(s)
mL	Mililitro
mm	Milímetro
mM	Milimolar
MOPS	Solução Tampão 3-[N-morfolino]propano-ácido Sulfônico
MPa	Megapascal
MS	Espectrometria de massa
µg	Micrograma
µM	Micromolar
µm	Micrômetro
µL	Microlitro
N	Newton (unidade de força)
Na	Sódio
NaCl	Cloreto de sódio
NBT	Cloreto de tetrazólio-nitroazul
NEB	New England Biolabs
ng	Nanograma
nm	Nanômetro
nN	Nanonewton
Ni-NTA	Níquel-ácidonitrilotriacético
N-terminal	Amino terminal
OD ₆₀₀	Densidade óptica no comprimento de onda de 600 nm
°C	Graus Celsius
pb	Par(es) de base(s)
PBS	Solução Tampão contendo NaCl, KCl, Na ₂ HPO ₄ , KH ₂ PO ₄
PCR	Reação em cadeia da polimerase
pH	Potencial Hidrogeniônico
poli-Ala	Poli Alanina
Pro	Prolina
PVDF	Fluoreto de polivinilideno
Ra	Rugosidade média aritmética
RNA	Ácido ribonucleico
Rq	Rugosidade quadrática média
Rp	Altura média
rpm	Rotações por minuto
Rv	Profundidade média
Rz	Altura máxima
Rzjis	Rugosidade média de 10 pontos
s	Segundo(s)

SDS	Dodecil sulfato de sódio
SDS-PAGE	Eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio
Ser	Serina
Si ₃ N ₄	Cerâmica de nitreto de silício
SOC:	Meio de cultura enriquecido para crescimento de bactérias
TFA	Ácido trifluoracético
Tir	Tirosina
Tris	Tris(hidroximetil)aminometano
tRNA	Ácido ribonucleico transportador
U	Unidade(s)
UPLC-MS	Cromatografia líquida de ultra-performance acoplada à espectrometria de massa
V	volts
X-Gal	5-bromo-4-cloro-3-indolil- -D-galactopiranosídeo
%	porcentagem
°	grau(s)
×	multiplicação (unidade repetitiva).

SUMÁRIO

AGRADECIMENTOS	iii
RESUMO	v
ABSTRACT	vi
LISTA DE FIGURAS	vii
LISTA DE TABELAS	xiv
LISTA DE ABREVIATURAS E SIGLAS	xv
SUMÁRIO.....	xviii
INTRODUÇÃO.....	1
OBJETIVOS	11
2.1 Objetivo geral	11
2.2 Objetivos específicos	11
JUSTIFICATIVA	12
HIPÓTESE	13
MATERIAL E MÉTODOS.....	14
5.1 Elaboraões das sequências utilizadas para expressão de espidroínas em sistema heterólogos	15
5.2 Plasmídeos	17
5.2.1 Plasmídeo pACYC184	17
5.2.2 Plasmídeos pJ	17
5.2.3 Plasmídeo de clonagem pBSSK+	20
5.2.4 Plasmídeo de expressão em bactéria (<i>Escherichia coli</i>)	20
5.2.5 Plasmídeo de expressão em soja (<i>Glycine max</i>)	21
5.3 Construções dos plasmídeos para a engenharia metabólica da bactéria.....	21
5.3.1 Amplificação por PCR e confirmação dos genes	23
5.3.2 Ligação e transformação.....	24
5.3.3 Gene glyVXY	25
5.3.4 Gene glyA.....	27
5.4 Construção dos plasmídeos de expressão de espidroínas.....	28
5.4.1 Construção dos plasmídeos de expressão das MaSp em <i>E. coli</i>	28
5.4.2 Construções dos plasmídeos de expressão das MaSp em soja (<i>Glycine max</i>).....	31
5.5 Transformações	31
5.5.1 Eletroporação.....	31
5.5.2 Choque térmico	32
5.6 Curva de crescimento	33
5.7 Análise por espectrometria de massa biotyper	34
5.8 Estudo proteômico das linhagens bacterianas	57
5.9 Produção de espidroínas em <i>E.coli</i>	58
5.10 Extração proteica e determinação da solubilidade da proteína alvo.....	58
5.11 Detecção e análise das proteínas	60
5.12 Purificação das proteínas produzidas em <i>E. coli</i>	61
5.13 Espectrometria de massa MALDI-TOF	62
5.15 Análises em nanoescala por Microscopia de força atômica (MFA).....	63
5.16 Microscopia Eletrônica de Varredura (MEV)	65
5.17 Testes mecânicos das fibras.....	65
5.18 Espectroscopia por infravermelho (ATR-FTIR)	66
RESULTADOS E DISCUSSÃO	67
6.1 Plasmídeos para engenharia metabólica de bactéria.....	67

6.2 Sequenciamento dos genes utilizados para os motivos de <i>P. bistriata</i>	72
6.3 Plasmídeos de produção das espidroínas em <i>E. coli</i>	73
6.4 Plasmídeos de produção das espidroínas em soja	80
6.5 Curvas de crescimento.....	82
6.6 Análise por espectrometria de massa Biotyper.....	85
6.7 Estudo proteômico das linhagens bacterianas	87
6.8 Caracterização e purificação das espidroínas MaSp2.....	91
6.8.1.Determinação da solubilidade das espidroínas.....	91
6.8.2 Caracterização molecular e rendimento de Masp2 (16×).....	92
6.8.3 Caracterização molecular e rendimento MaSp2 (32×).....	97
6.9 Espectrometria de massa MALDI-TOF	102
6.10 Síntese e processamentos das fibras	103
6.11 Análises em nanoescala por Microscopia de Força Atômica	107
6.12 Análises por Microscopia Eletrônica de Varredura (MEV)	114
6.13 Testes mecânicos	117
6.14 Espectroscopia por infravermelho (FTIR-ATR).....	123
CONCLUSÃO.....	125
REFERÊNCIAS BIBLIOGRÁFICAS	126
ANEXO A – Teórica digestão da proteína MaSp2 16× com quimotripsina.	i
ANEXO B – Teórica digestão da proteína MaSp2 32× digerida com quimotripsina.	ii
ANEXO C- Proteínas encontradas apenas na linhagem BL21(DE03)pACYC184	iv
ANEXO D- Proteínas encontradas apenas na linhagem BL21(DE03).	vi
ANEXO E- Proteínas encontradas em maior quantidade na linhagem BL21(DE03)pACYC184 em relação à BL21(DE03).....	xii
ANEXO F- Proteínas encontradas em maior quantidade na linhagem BL21(DE03)pACYC184 em relação à BL21(DE03).....	xvii
ANEXO G –Artigos científicos publicados e participações em eventos.	xxi

INTRODUÇÃO

Durante muitos anos, as sedas de fibra de aranhas têm despertado o interesse da humanidade devido às suas propriedades mecânicas. Com os avanços da biotecnologia, muitas pesquisas têm tido como objetivo aprofundar os estudos nessas estruturas. Já se sabe que essas proteínas apresentam uma combinação única de resistência à tensão e elasticidade, sendo cinco vezes mais fortes do que o aço, e três vezes mais resistentes que o polímero químico Kevlar®*. Com isso, têm-se buscado a produção de novos materiais com base nos polímeros da seda de aranhas.

Ainda não foi possível produzir as fibras de aranhas em larga escala, por meio de cultivo dos indivíduos, pois esses animais possuem comportamento agressivo e territorialista. Sendo assim, as pesquisas biotecnológicas têm logrado considerável sucesso no desenvolvimento de reatores que permitem a produção de proteínas recombinantes e a síntese dos fios de fibras de maneira artificial. Algumas proteínas de fibras apresentaram características próximas às naturais, como as conseguidas por Oliveira (2008), e Teulé e colaboradores (2009) que têm em torno de 50 kDa, mas o maior desafio atualmente, é produzir moléculas semelhantes à encontradas na natureza (200-350 kDa) com propriedade mecânicas semelhantes.

A limitação na expressão das proteínas está relacionada principalmente à escolha do organismo a ser utilizado como bioreator. Apesar do cultivo de procariontes apresentarem melhor custo/benefício, estes organismos possuem uma limitação na síntese dessas fibras, porque essas proteínas possuem grande quantidade de resíduos de glicina e alanina (HAYASHI et al., 1999). Contudo, Xia e colaboradores (2010) fizeram alterações na rota metabólica de *E. coli* e proteínas de 150 a 289 kDa da fibra de *N. clavipes* foram sintetizadas com sucesso, após a expressão de genes que aumentaram a expressão de glicil-tRNA. Mostrando que esse é um caminho viável na produção de fibras maiores que 50 kDa.

Quanto maior a complexidade do organismo, maior é a necessidade de manutenção dos meios de cultivos, aumentando dessa forma os custos do processo. Além disso, os resultados esperados podem não ser promissores, uma vez que organismos eucariotos introduzem modificações pós-traducionais (WILDT; GERNGROOS, 2005). No entanto, as plantas têm-se mostrado um eficiente sistema de expressão para a produção de proteínas heterólogas (SCHELLER et al., 2001). A seda de aranha já foi expressa em folhas de tabaco, em *Arabidopsis* e em soja (SCHELLER et al., 2001; BARR et al., 2004).

As proteínas recombinantes de alta massa molecular são mais resistentes e apresentam melhores propriedades de fibra (XIA et al., 2010; RISING et al., 2011), e a produção em larga escala poderia ser utilizada para numerosas aplicações industriais como em cordas de pára-quedas, redes de pesca, coletes protetores (utilizados por instituições militares), produtos esportivos e matéria prima de aeronaves. Além disso, apresentam um amplo potencial para aplicação na área biomédica por causa da possibilidade de biocompatibilidade e biodegradação.

As aranhas são artrópodes classificadas na classe Aranae, dos aracnídeos, agrupando 42.751 espécies de 10 famílias, onde a Araneidae contém 169 gêneros e 3.045 espécies (PLATNICK, 2014). Essa ordem é subdividida em Mesothela e Ophisthothelae, sendo essa última constituída por aranhas que não constroem teias orbiculares (Mygalomorphae) e as que constroem teias orbiculares (Araneomorphae).

A aranha *Parawixia bistriata* (ordem: Araneae, sudordem: Araneomorphae, família: Araneidae) é uma espécie típica do cerrado, ocorrendo em outras regiões do Brasil e também na Argentina, Paraguai e Bolívia. A maioria das espécies de Araneidae apresenta um genótipo $2n=24, X_1X_2$ (WOLF; ARAÚJO, 2011). Constroem teias comunais enormes durante a noite, podendo atingir 15 m de comprimento e 3 m de altura. Diferentemente de outras espécies, vivem em colônias contendo até 200 indivíduos na mesma teia durante a juventude. Durante o dia, cada aranha ingere sua teia e se escondem em ramos de árvores ou abrigos. As aranhas jovens apresentam cor preta, com manchas vermelhas e faixas claras no abdômen, enquanto as adultas têm abdômen esverdeado, tornam-se solitárias e constroem grandes teias individuais de até um metro de diâmetro (MOTTA, 2014).

As aranhas possuem até sete diferentes tipos de glândulas, onde seis produzem diferentes tipos de sedas e uma é reposável pela produção de uma substância gelatinosa e pegajosa, podendo ocorrer variações entre as espécies. A estrutura proteica dessas fibras é formada por resíduos de aminoácidos em motivos repetitivos (HAYASHI; LEWIS, 1998; HINMAN et al., 2000). A aranha pode combinar os diferentes tipos de sedas a fim de suprir suas necessidades (SCHEIBEL, 2004; TOKAREVA et al., 2013). As diferenças entre as fibras podem ser determinadas permitindo identificar quais módulos foram utilizados (Figura 1). As primeiras fibras de aranha estudadas foram das espécies *N. clavipes* e *Araneus diadematus*, por grupos norte-americanos, por conterem tamanhos maiores, que facilita a captura e dissecação das glândulas. As glândulas presentes nas aranhas são a ampolada maior, ampolada menor, flageliforme, agregata, tubuliforme, aciniforme e a piriforme.

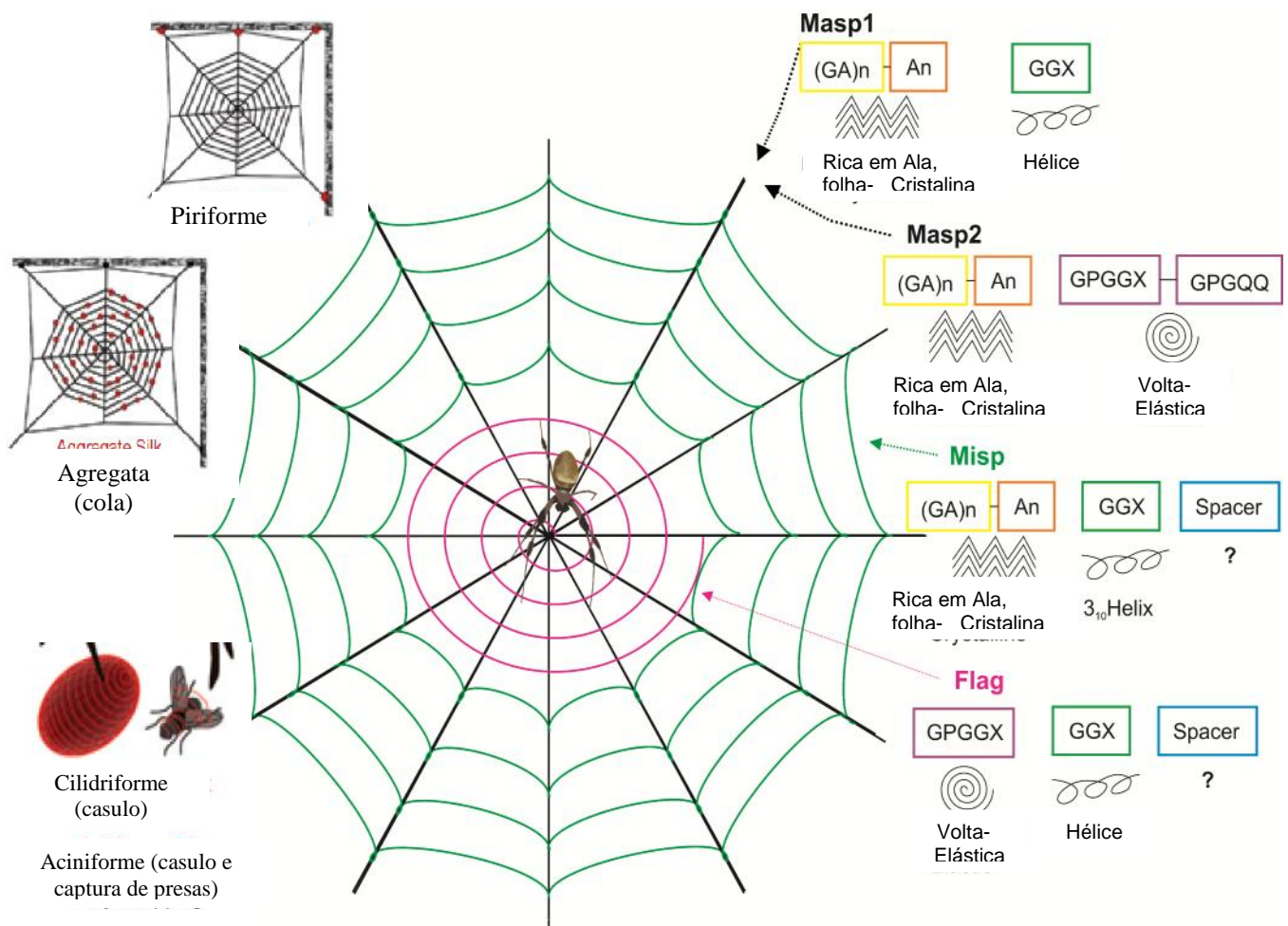


Figura 1. Desenho ilustrativo de uma teia orbital de aranha composta de diferentes proteínas. As molduras coloridas indicam as estruturas a serem formadas pelos motivos. A interrogação próxima à região espaçadora significa que a estrutura secundária não é conhecida. MaSp: Espidroína da glândula ampulada maior; MiSp: Espidroína da glândula ampulada menor; Flag: proteína flageliforme. Adaptada de TOKAREVA et al., 2013 e SLOTTA et al., 2012.

As fibras de teia de aranha naturais são constituídas de várias camadas, que podem ser agrupadas em duas partes: 1) casca e 2) núcleo. A casca possui lipídeos na camada mais externa, com morfologia semelhante às membranas celulares e formando uma camada de proteção contra microrganismos. Em seguida, a camada é composta por glicoproteínas, formando uma segunda barreira protetora e com potencial de controle hídrico na fibra; e por último, a pele ou película que atua como uma terceira camada proteção contra danos químicos e de microrganismos, além de adicionar plasticidade à fibra e servindo de suporte para a estrutura interna. A segunda parte, o núcleo, ele é dividido em partes externa e interna, onde a externa é constituído de estruturas folhas- e a interna possui estruturas de canalículos com formações de fibrilas (Figura 2) (SPONNER et al., 2007).

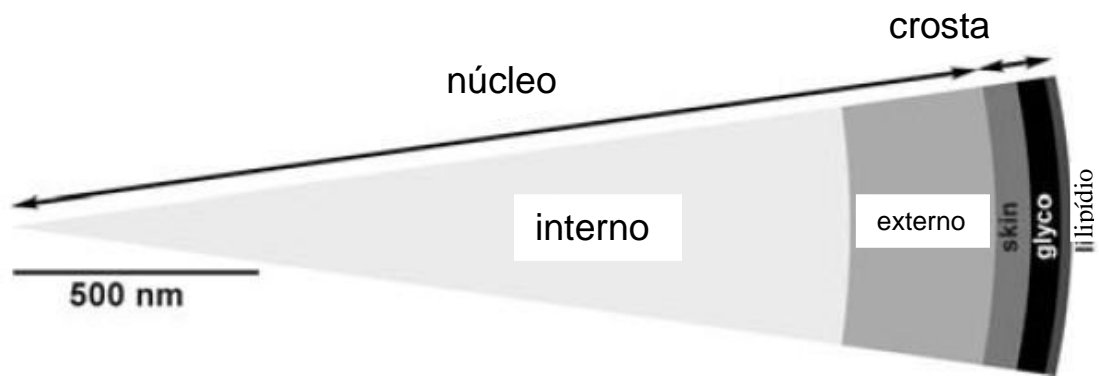


Figura 2. Estrutura interna de uma seda de aranha linha de segurança. O modelo multicamadas divide cinco partes, do externo para o interno: camada de lipídeos, glicoproteínas, película, núcleo externo e núcleo interno. Modificada de SPONNER et al., 2007.

Dentre as glândulas, a ampolada maior é a mais volumosa no abdômen da aranha, sendo responsável pela produção da fibra tipo linha de segurança, a qual serve como guia na fuga de predadores. Essa glândula apresenta o maior foco de estudos, pois produz uma seda que possui alta elasticidade e resistência à tensão (Tabela 1), sendo cinco vezes mais forte do que o aço e força tensora similar ao Kevlar (Tabela 2) (XU; LEWIS, 1990; STAUFFER et al., 1994; GOSLINE et al., 1999).

Tabela 1. Propriedades mecânicas das linhas de segurança obtidas de teia orbital de diferentes espécies (Adaptado de SLOTTA et al., 2012).

Espécie	Rigidez (GPa)	Força máxima (MPa)	Extensão (%)	Resistência MJ.m ⁻³
<i>Araneus diadematus</i>	3,6	1.599	39	193
<i>Nuctenea umbratica</i>	7,4	1.778	35	196
<i>Lairinioides cornutus</i>	7,7	1.760	31	197
<i>Nephila clavipes</i>	8,4	1.725	32	206
<i>Verrucosa arenata</i>	7,6	1.842	36	228
<i>Gasteracantha cancriformis</i>	5,2	1.929	49	254
<i>Caerostris darwini</i>	11,5	1.850	39	271

Tabela 2. Propriedades mecânicas de alguns materiais (Adaptado de ROMER; SCHEIBEL, 2008).

Material	Força Tensora _{max} .GPa ⁻¹	Elasticidade (%)	Resistência MJ.m ⁻³
Seda de <i>B. mori</i>	0,6	18	70
Linha de segurança	1,1	27	160
Seda flageliforme	0,5	270	150
Nylon	0,95	18	80
Kevlar	3,6	2,7	50
Aço de alta tensão	1,5	0,8	6

Cada glândula pode ser dividida em quatro áreas: a cauda, o lúmen, o ducto e a fiandeira. A cauda possui células epiteliais que secretam as espidroínas; o lúmen armazena a solução de fiação; o ducto orienta as proteínas e a desidratação durante a fiação e, a fiandeira é uma válvula que controla a produção da fibra final (VOLLRATH; KNIGHT, 2001). A polimerização consiste no processo de transformar proteínas solúveis em fibras insolúveis (Figura 3). À medida que a concentração da espidroína aumenta, a fibra passa da estrutura *coil* para α -hélice, permitindo que a aranha armazene grandes quantidades de proteínas sem a formação de folhas- insolúveis (LEWIS, 2006). À medida que passam pelo ducto, elas são desidratadas e perdem sódio e cloreto; são secretados íons de hidrogênio e potássio, ocorrendo uma diminuição do pH de 6.9 para 6.3 (DICKO et al., 2004) e as fibras se alinham por aproximação dos segmentos hidrofóbicos de poli-Alanina (SCHEIBEL, 2004).

Na polimerização das fibras sintéticas, uma das metodologias é a produção utilizando um banho de coagulação. A espidroína é produzida num sistema heterólogos, purificada e

solubilizada. Logo após, a solução de fiação passa por uma agulha de diâmetro pequeno, que segue seu caminho pelos adaptadores de tubos para HPLC, diminuindo o diâmetro até atingir o interior de um tubo *peek* com medidas conhecidas (de poucos micrômetros). Nessa fase as proteínas já estão sofrendo um processo inicial de alinhamento, mas a polimerização da fibra ocorre por desidratação num banho de coagulação podendo ser composto por diferentes concentrações de álcool (metanol, isopropanol) (TEULÉ et al., 2007; revisto por HEIM, M; KEERL,D. e SCHEIBEL, T.).

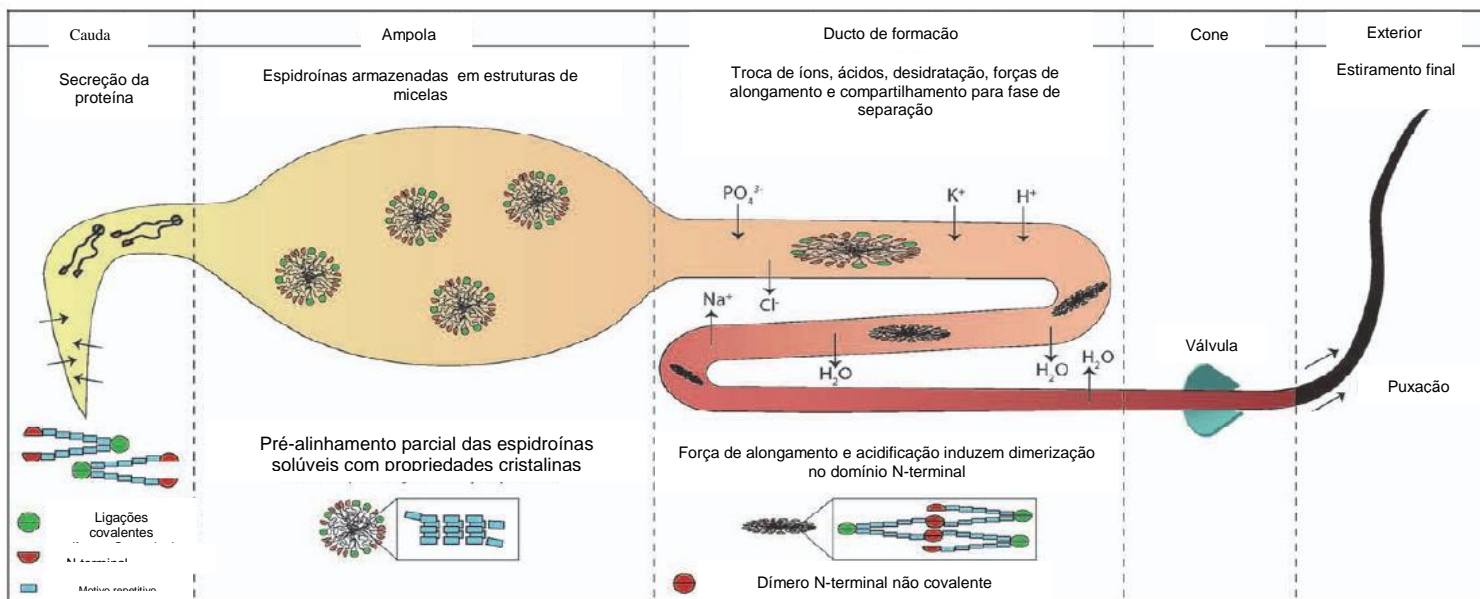


Figura 3. Ilustração da formação natural das fibras de aranha. As espidroínas são produzidas e secretadas na região da cauda e armazenadas na ampola em altas concentrações (~50%). No ducto de formação ocorrem trocas iônicas, acidificação e alongamento. Antes de encontrar o meio externo, a fibra passa por uma válvula, que auxilia na formação de uma nova fibra, que ocorre por puxação. (Adaptado de EISOLDT et al., 2011).

As sedas produzidas pelas aranhas apresentam funções distintas, como na reprodução, predação e locomoção (GUERETTE, 1996). Além disso, cada espécie é capaz de sintetizar características únicas em suas fibras (DENNY, 1980). Na espécie *N. clavipes* ela é constituída principalmente por proteínas Major Ampullate Spidroin 1 e 2 (MaSp1 e MaSp2) (HINMAN; LEWIS, 1992), enquanto em *A. diadematus* por Aranus Diademattus Fibroin 3 e 4 (ADF3 e ADF4). Segundo Lombardi e Kaplan (1990) a razão molecular entre MaSp1 e MaSp2 e entre ADF-4 e ADF-3 é de aproximadamente 3 para 2. Enquanto isso, essa proporção para aranha *P. bistriata* é de 90% para MaSp1 e 10% para MaSp2 (OLIVEIRA 2012). A diferença entre as proteínas da mesma glândula são as sequências de aminoácidos e a quantidade de motivos presentes que apresentam características específicas (Figura1). A natureza de interação entre essas proteínas não é conhecida, mas a sua associação é conservada, independentemente da espécie (TIAN et al., 2004).

Os motivos repetitivos, encontradas nas sequências de MaSp1 e MaSp2, possuem domínios ricos em alanina e glicina (poli-A e (Gli-Ala)_n) (HAYASHI et al., 1999). As sequências de alaninas formam folhas-β, que são responsáveis pelas regiões hidrofóbicas de formação de cristais (RIEKEL et al., 2000). As sequências de glicina são regiões hidrofílicas, onde GliGliX (X=Leu, Tyr, Ser, Ala) estabiliza a estrutura da fibra e são encontradas nas proteínas de MaSp1 (TATHAM; SHEWRU 2000; HOLLAND et al., 2008), enquanto a GliProGliXX (X=Gli, Gln, Tyr, Ser, Ala) é responsável pela elasticidade, sendo uma característica típica de Masp2 (GOSLINE et al., 2002). Segundo Lewis (2006), os cassetes de expressão contendo apenas os módulos repetitivos, sem regiões N-terminal e C-terminal, são de fácil expressão em micro-organismos e em plantas. Além disso, um plasmídeo de expressão que possui um promotor de alta expressão, como por exemplo, o híbrido T7, quando induzido por IPTG promove um acúmulo de 50%, aproximadamente, de proteína heteróloga a partir do extrato total da bactéria (SAMBROOK; RUSSEL, 2001).

Todas as fibras de aranha naturais são compostas por proteínas com massa molecular entre 300-500 kDa, constituídas por regiões N-terminal e C-terminal não repetitivas e altamente conservadas, e uma região central repetitiva rica em alanina, glicina e serina. A extremidade N-terminal possui códons adicionais de iniciação da transcrição sendo ainda importante no transporte das espidroínas para o lúmen glandular (SMITH, 2005). Além disso, num estudo de simulação dinâmica molecular do alinhamento, a extremidade N-terminal dos módulos de MiSp (Minor ampullate Spidroin) de *P. bistriasta* formam folhas-β. Elas aumentam as pontes de hidrogênio induzidas pela repetição de alanina e resultam numa fibra mais resistente, contribuindo significativamente na estrutura terciária da proteína que compõem a fibra (MURAD; RECH, 2010). Segundo a revisão de Omenetto e Kaplan (2010), as regiões N-terminal e C-terminal desempenham funções importantes relacionadas à concentração de proteína, controle de pH e de cátions dentro da glândula da aranha.

Um dos problemas para se obter um bom resultado é entender a relação entre a sequência da proteína e a sua função na fibra final. Primeiro é necessário o conhecimento da interação de MaSp1 e MaSp2 associadas em solução cristalina da glândula e na fibra. Em seguida, são três níveis de variabilidade que afetam a fibra nativa: 1) interespecies, as quais possuem uma grande variabilidade, 2) intraespecies, ou diferenças entre indivíduos da mesma espécie, 3) variação individual das fibras formadas pelo mesmo indivíduo (MADSEN et al., 1999). Essas variações podem ainda apresentar diferenças na sequência de aminoácidos das unidades repetitivas, na proporção de MaSp1 e MaSp2 e nas propriedades mecânicas da

fibra (MADSEN et al., 1999; BROOKS et al., 2005; BROOKS et al., 2007; BROOKS et al., 2008).

A caracterização das fibras pela diversidade de módulos iniciais é importante, pois a mesma sequência inicial, em quantidades diferentes, pode proporcionar estruturas distintas. Por exemplo, a seda das glândulas ampolada maior (MaSp) e da flageliforme (Flag) são fibras constituídas do mesmo motivo mas a diferença quantitativa faz a primeira proteína possuir elasticidade em torno de 35%, enquanto a segunda proteína apresenta mais de 200% (HINMAN et al., 2000). Essa natureza modular encontrada na seda de aranha e suas propriedades físicas são elementos que levam à pesquisa com a engenharia de módulos para expressão de proteínas análogas (HINMAN et al., 2000). Um dos problemas encontrados para a expressão dos módulos é a grande quantidade de G+C (guanina e citosina). As primeiras tentativas com procariotos não apresentaram resultados satisfatórios. Na hipótese de melhorar a expressão, diferentes sistemas, já foram testados em fungos, plantas como a soja, tabaco, tubérculo de batata, *Arabidopsis* e também em células de mamíferos (SCHELLER et al., 2001; LAZARIS et al., 2002; BARR et al., 2004, revisado por TOKAREVA et al., 2013). No entanto, quanto mais sofisticado o meio de cultivo, maior é o custo de manutenção.

A *E. coli* é o procarioto com o sistema que permite altos níveis de expressão de proteína heteróloga, por apresentar um crescimento rápido, em meios de cultura de composição simples. Além disso, já existem muitos protocolos estabelecidos e de rápida manipulação para o cultivo destes organismos. Até então, a melhor fibra com propriedades mecânicas obtidas, foi produzida em células de mamíferos com 60 kDa, sendo derivada de *Araneus diadematus* (LAZARIS et al., 2002). A partir de 2010, Xia e colaboradores anunciaram a produção de fibras de 284,9 kDa da espécie *N. clavipes* utilizando a biologia sintética para alterar o metabolismo de expressão de *E. coli*. Eles observaram um efeito sinérgico positivo na produção de proteínas de grande massa molecular ao aumentar a disponibilidade de RNA transportador de glicina e o seu respectivo aminoácido.

Sabe-se, porém que a eficiência da transformação em procariotos diminui linearmente com o aumento do tamanho do plasmídeo (HANAHAN, 1983). Bactérias gram-negativas interagem e transferem preferencialmente DNA de fita dupla, enquanto as gram-positivas interagem com DNA de fita dupla, mas só transferem fita simples (HANAHAN, 1983). As transformações por choque térmico de plasmídeos grandes e alto número de repetições são recomendadas em relação à eletroporação. Esse tipo de transformação ocorre quando as células e o DNA estão na presença de íons de cálcio (Ca^{+2}) 50 mM a 0 °C e são submetidos a um breve, mas importante, pulso termodiânico a 37-42 °C. Adicionalmente, a presença de íons

de Magnésio (Mg^{+2}) 10-20 mM em todos os meios de cultura promovem um aumento do sucesso da técnica, por ter carga positiva, ajuda na aproximação do plasmídeos e membranas celulares, que apresentam cargas negativas.

O crescimento de uma população de bactérias pode ser estudado pela análise da curva de crescimento em meio de cultura. O sistema é denominado fechado, pois os micro-organismos crescem em meio líquido num recipiente, sem ocorrer a substituição por meio fresco durante o período de incubação, ocorrendo a diminuição dos nutrientes e aumento de resíduos celulares. O crescimento por fissão binária pode ser plotado como função logarítmica do número de células viáveis pelo tempo de incubação, por exemplo (Figura 4). Esse crescimento apresenta quatro fases: lag, log, fase estacionária e morte celular. A fase lag corresponde à introdução e adaptação dos micro-organismos no meio de cultura, sem ocorrer crescimento imediato; a fase exponencial ou log ocorre quando o crescimento e divisão estão em razão máxima possível, limitado apenas ao seu potencial genético, composição do meio e condições de incubação; a fase estacionária corresponde à diminuição do crescimento da população, e a curva torna-se horizontal e o número de células viáveis permanece constante; a fase de morte celular ocorre quando o meio de cultura já está pobre em nutrientes e com excesso de resíduos tóxicos que levam à diminuição do número de células viáveis, havendo um declínio na curva (PRESCOTT, 2002). Os diferentes meios de cultura ou a toxicidade de proteínas heterólogas em plasmídeos de expressão podem descolar a curva de crescimento para a esquerda ou para a direita em relação a uma amostra controle, levando um tempo menor ou maior para alcançar o objetivo esperado (PALIY; GUNASEKERA, 2007).

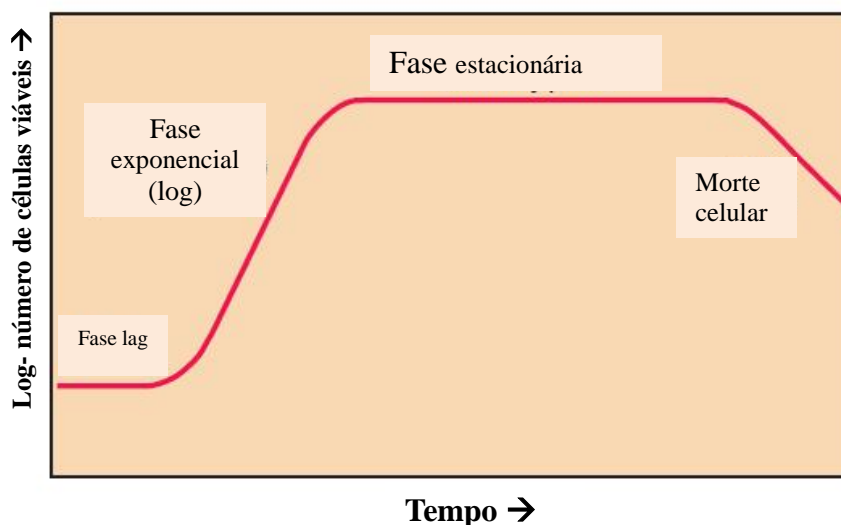


Figura 4. Desenho esquemático de uma curva de crescimento de bactéria plotando o número de células viáveis pelo tempo em sistema fechado. As fases lag, exponencial, estacionária e de morte celular estão identificadas no gráfico. Adaptado de PRESCOTT; HARLEY e KLEIN, 2005)

Os vegetais como bioreatores apresentam uma biomassa de baixo custo e um sistema econômico para a produção em larga escala de proteínas de interesse, reduzindo os custos de 10 a 50 vezes em relação ao sistema bacteriano (KUSNADI et al., 1997). A sua alta eficiência no processo de síntese, secreção e modificações pós-traducionais proporcionam uma manutenção da estrutura das proteínas recombinantes. A ausência de patógenos transmissíveis dos vegetais ao homem é outra vantagem na utilização de plantas transgênicas para a produção de proteínas heterólogas cujo objetivo final se destina ao consumo humano. Outros trabalhos com sistemas de produção de espidroínas podem ser encontrados no artigo de revisão em anexo (TOKAREVA et al., 2013).

A teia de aranha é um material interessante para estudar a relação entre a composição química, conformação estrutural e propriedades mecanoelásticas de fibras biológicas. À medida que a sequência primária é muito repetitiva torna a estrutura secundária homogênea, formando folhas, que são importantes para as propriedades mecânicas (LOMBARDI; KOPLAN, 1990). Por isso, são necessárias mais pesquisas em novas espécies de aranhas, como a *Parawixia bistriata* (Figura 4) para a aplicação de biomateriais e engenharia de tecidos, mas principalmente para agregar valor àquelas da biodiversidade brasileira.



Figura 4. Aranha jovem da espécie *Parawixia bistriata* construindo sua teia ao entardecer em ambiente periurbano. Foto adquirida por Valquíria Lacerda próximo ao Prédio de Biotecnologia, Recursos Genéticos e Biotecnologia, Brasília, DF, Brasil.

OBJETIVOS

2.1 Objetivo geral

Produzir e caracterizar diferentes linhagens de *Escherichia coli* modificadas por engenharia metabólica a fim de produzir e caracterizar proteínas e fibras sintéticas a partir de espidroínas da aranha *Parawixia bistriata*.

2.2 Objetivos específicos

- 1) Construir os plasmídeos para aumentar a disponibilidade de glicil-tRNA e serina hidroximetiltransferase no metabolismo de bactéria;
- 2) Construir os plasmídeos de expressão de espidroínas em bactéria (*E.coli*);
- 3) Estudar o perfil proteômico das linhagens bacterianas construídas;
- 4) Produzir e purificar as proteínas MaSp;
- 5) Avaliar o rendimento dos sistemas heterólogos;
- 6) Caracterizar as espidroínas por meio de espectrometria de massa, MEV, ATR-FTIR;
- 7) Caracterizar e avaliar as fibras sintéticas por meio de testes mecânicos, MFA, MEV e ATR-FTIR.

JUSTIFICATIVA

As fibras de teia de aranha apresentam um amplo potencial de utilização na indústria de um modo geral. No entanto, faz-se necessário reduzir os custos de produção e encontrar um bioreator que seja ideal para a produção de fibras com alta massa molecular e em larga escala. As fibras constituídas de proteínas de alta massa molecular (maiores que 100 kDa) apresentam melhores características de resistência e flexibilidade. A soja (*Glycine max*) e a bactéria *E. coli* modificada por engenharia metabólica (para maior produção de glicil-tRNA) são organismos candidatos a bioreatores mais eficientes para a expressão de proteínas de fibra de aranha (Figura 5).

Os estudos moleculares para a produção de biopolímeros concentraram-se, principalmente, em duas espécies de aranhas norte-americanas. A *Parawixia bistriata* é uma espécie de aranha brasileira que vive em grupo, sem canibalismo ou territorialismo. Faz-se necessário um maior investimento em estudos que envolvam produtos obtidos de espécies da biodiversidade brasileira, realizando-se a prospecção de genes das glândulas produtoras de sedas de aranhas e, a produção proteica em sistemas de expressão heterólogos, que sejam economicamente viáveis. Uma vez obtidas estas proteínas, estas poderão ser aplicadas, por exemplo, em suturas de microcirurgias, coletes militares, produtos esportivos, cordas ou redes especiais ou em condições onde se requer muita absorção de energia e alongamento, semelhante ao nylon.

HIPÓTESE

- As bactérias modificadas por engenharia metabólica apresentarão um rendimento maior para a produção das espidroínas de alta massa molecular de *Parawixia bistriata*?
- As novas sequências sintéticas de espidroínas formarão fibras?
- Se sim, essas fibras sintéticas apresentarão as propriedades mecânicas alteradas, comparando-se as fibras hidratadas e estiradas com as que não receberam o tratamento?

MATERIAL E MÉTODOS

A execução do projeto consistiu em duas grandes partes: uma com a construção das novas linhagens de bactérias (Figura 5A), e a outra com a construção da informação genética das espidroínas, por engenharia modular, e caracterização molecular e físico-química das proteínas e fibras sintéticas obtidas (Figura 5B).

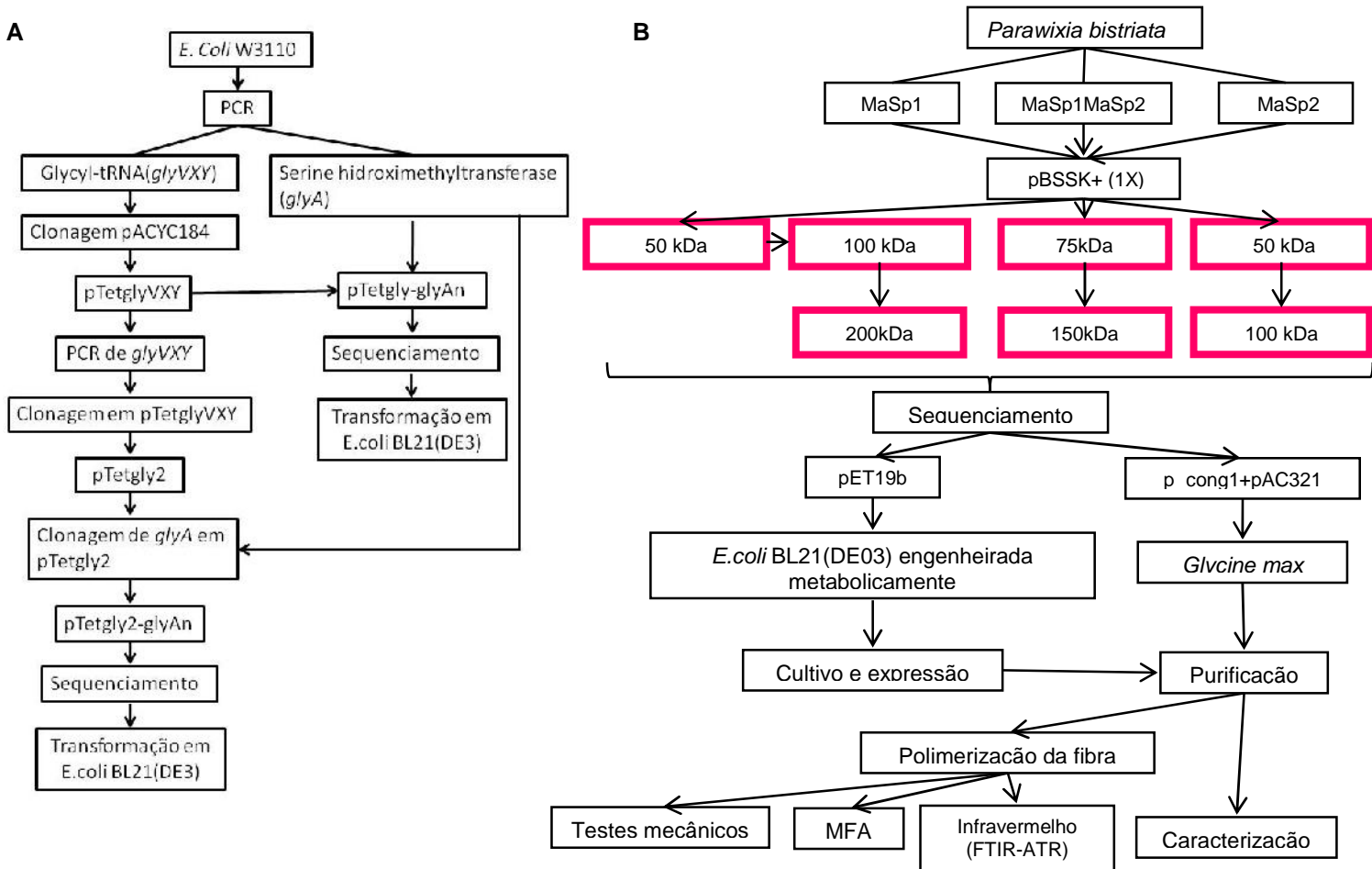


Figura 5. Desenho esquemático resumindo o projeto desenvolvido. A- Alteração por engenharia metabólica de *E. coli* para aumentar a expressão de glicil-tRNA. pACYC184: plasmídeo de clonagem (NEB); pTetglyVXY: plasmídeo contendo uma cópia do gene *glyVXY*; pTetgly2: plasmídeo contendo duas cópias do gene *glyVXY*, pTetgly-glyA: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTetgly2-glyA: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. B- Ilustração da estratégia para obtenção de espidroínas e respectivas avaliações. A linha rosa indica as construções a serem utilizadas. MaSp: Major Ampullate Spidroin, PBSSK+: pBluescriptIISK+, nx: número de motivos presentes no plasmídeo. pET19b e p cong1: plasmídeos de expressão em *E. coli* e soja (*Glycine max*), respectivamente; pAC321: plasmídeo de seleção para soja. MFA: Microscopia de Força Atômica.

5.1 Elaborações das sequências utilizadas para expressão de espidroínas em sistema heterólogos

As sequências utilizadas nas sínteses dos plasmídeos para engenharia modular e expressão foram construídas com base nas sequências originais de MaSp1 (CQ275360.1) e MaSp2 (CQ275359.1). Estas foram identificadas em uma biblioteca de cDNA, construídas previamente (OLIVEIRA, 2008), no Laboratório de Biologia Sintética, sob supervisão do pesquisador Elíbio Leopoldo Rech Filho, Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, a partir das glândulas produtoras de seda da aranha *Parawixia bistriata* (RECH et al., US 20140206851A1).

Os respectivos motivos iniciais foram construídos sem as regiões N-terminal e C-terminal dos genes das espidroínas (Figura 6) e sintetizados pela empresa DNA2.0 (Califórnia, EUA). Essas sequências foram desenhadas previamente com os códons de uso preferenciais de bactéria *E. coli*, sendo adicionados sítios de enzimas de restrição. A metodologia para as construções dos plasmídeos consiste em aumento exponencial por engenharia modular (TEULÉ et al., 2009). As sequências de DNA e as traduções proteicas sintetizadas previamente pelo grupo de pesquisa encontram-se no quadro 1 e 2.

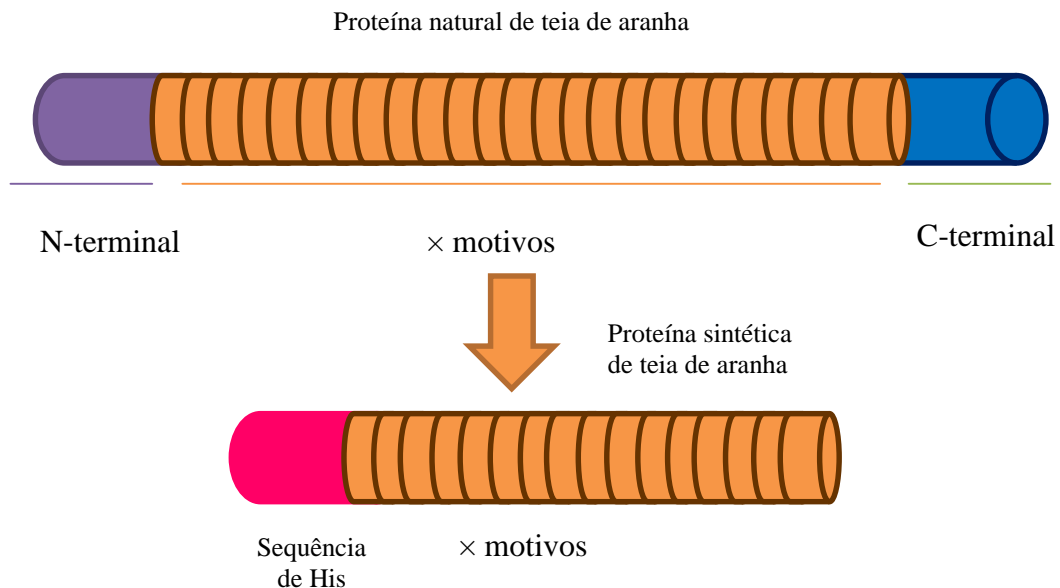


Figura 6. Desenho representativo de proteínas de teia de aranha. As proteínas naturais apresentam extremidades C-terminal e N-terminal, com a parte central constituída de motivos altamente repetitivos. As proteínas sintéticas para sistemas heterólogos não precisam das sequências C-terminal e N-terminal. Adicionalmente, pode-se inserir sequências que auxiliam o processo de purificação proteica, como a sequência de histidina.

Quadro 1. Informações das sequências genéticas da aranha *Parawixia bistrata* para as construções dos motivos sintéticos com o códon de uso adaptado para *E. coli*.

Espidroína	Sequência genética dos motivos
MaSp1	(5'GGTCTGGGTGGTCAGGGTGGTCTGGGTGGTCTGGGTCCCAGGGTGCTGTCTGGGTGGTTACGGTCAGGGTGGTGGTCTGGTCAGGGTGGTGGTCTGCTGCTGCTGCTGCTGCT3') _{n×}
MaSp2	(5'GGCGGCTATGGTCCGGGCGGCGGGCCAGCAGGGCCCCGGCGGGCGGGCAGCAGGGTCCGGGCAGCCAGGGCTCGTATGGTCCGGGCGCGGCGGGCGGGCGGGCGGGCGG3') _{n×}
MaSp1(2×)+ 2(1×)	(5'GGTGGTCTGGGTGGTCAGGGTGGTCTGGGTGGTCTGGGTCCCAGGGTGCTGGTCTGGGTGGTTACGGTCAGGGTGGTGGTCTGGTCAGGGTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGGTGGTCTGGGTGGTCAGGGTGGTCTGGTGGTCTGGGTCCCAGGGTGCTGGTCTGGGTGGTTACGGTCAGGGTGGTGGTCTGGTCAGGGTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGCGGCTATGGTCCGGGCGGCGGGCCAGCAGGGCCCCGGCGGGCGGGCCAGCAGGGTCCGGGCAGCCAGGGCTCGTATGGTCCGGGCGCGGCGGGCGGGCGGGCGGGCGGGCGGCGG3') _{n×}

MaSp: Espidroína da Glândula Ampolada Maior.

n×: número de repetições.

Quadro 2. Informações das sequências proteicas da aranha *Parawixia bistrata* para as construções dos motivos sintéticos.

Espidroína	Sequência proteica dos motivos
MaSp1	GLGGQGGLGGLGSQGAGLGGYGQGGAGQGGAAAAAAAAA
MaSp2	GGYGPGGAGQQGPAAGQQGPGSQGSYGPAAAAAAAAA
MaSp1(2×)+ 2(1×)	GGLGGQGGLGGLGSQGAGLGGYGQGGAGQGGAAAAAAAAAAGGLGGQGGLGGLGSQGAGLGGYGQGGAGQGGAAAAAAAAA ₂ GGYGPGGAGQQGPAAGQQGPGSQGSYGPAAAAAAAAA

MaSp: Espidroína da Glândula Ampolada Maior.

5.2 Plasmídeos

5.2.1 Plasmídeo pACYC184

O plasmídeo pACYC184 é utilizado em clonagens bacterianas e contém a origem de replicação do tipo p15A, ou seja, permite a co-existência de outros plasmídeos do grupo ColE1, como pBR322, pUC19, nas células. Ele possui genes de resistência à tetraciclina e cloranfenicol, é classificado como de baixa quantidade de cópias, ~15 por células, mas pode ser ampliado com antibiótico espectinomicina (manual do produto, E4152S, *New England BioLabs-NEB*).

5.2.2 Plasmídeos pJ

As sequências contendo as unidades repetitivas de espiroínas de aranha são denominado motivos. Assim, os motivos iniciais de MaSp1, MaSp2 e MaSp1+2 (Quadro 1) foram sintetizados e disponibilizados nos plasmídeos pJ, desenvolvidos pela própria empresa DNA 2.0. O motivo MaSp1+2 contém dois motivos da MaSp1 e um motivo da MaSp2 [MaSp1(2×)+2]. Eles apresentam origem de replicação pUC, genes de resistência à canamicina e alto número de cópias (500-700 por célula) e informações podem ser observados na figura 7 e 8, e na tabela 3.

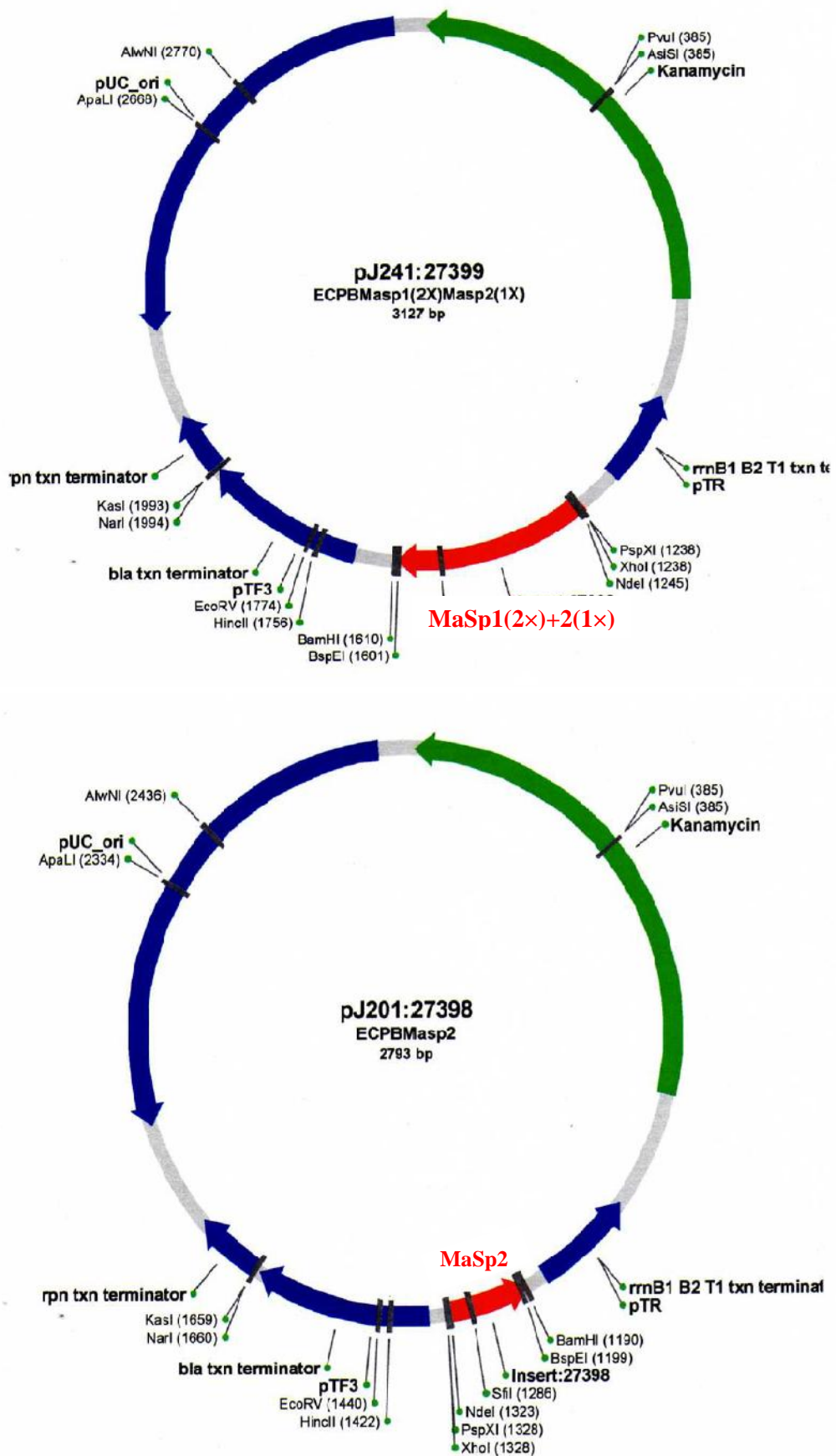


Figure 7. Desenho representativo dos plasmídeos pJ, sintetizados pela DNA2.0 (EUA) contendo os motivos sintéticos da espécie *P. bistrriata* e as enzimas de restrição a serem utilizadas na metodologia de produção em *E. coli*. O pJ241 contém o motivo da MaSp1(2x)+2(1x), enquanto o pJ 201 contém o motivo da MaSp2. O tamanho em pares de bases encontra-se ao centro de cada plasmídeo.

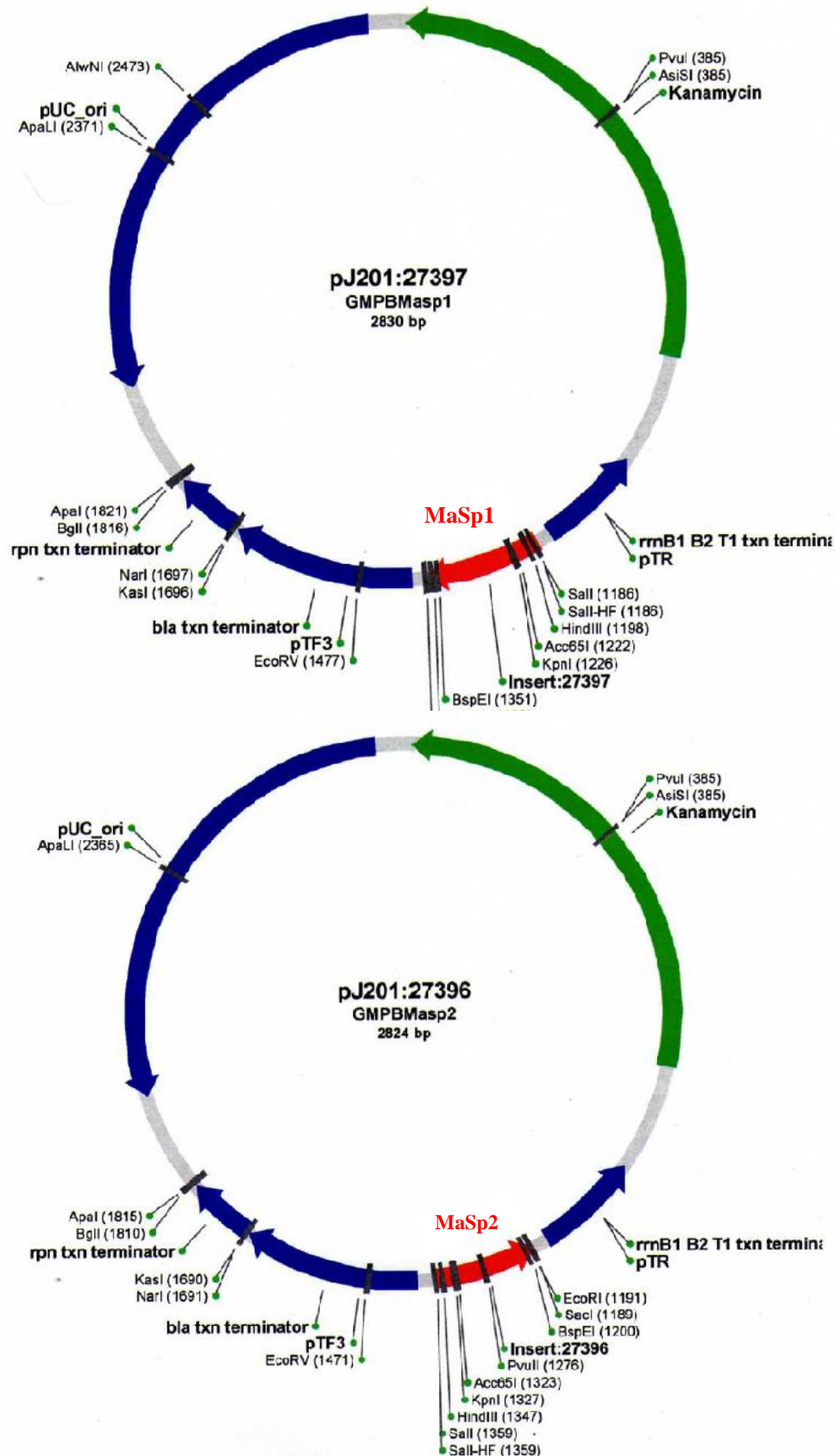


Figure 8. Desenho representativo dos plasmídeos pJ, sintetizados pela DNA2.0 (EUA) contendo os motivos sintéticos da espécie *P. bistriata* e as enzimas de restrição a serem utilizadas na metodologia de produção em soja (*Glycine max*). O pJ201:27397 contém o motivo da MaSp1, enquanto o pJ201:27396 contém o motivo da MaSp2. O tamanho em pares de bases encontra-se ao centro de cada plasmídeo.

Tabela 3. Tamanho dos motivos sintéticos iniciais da espécie *P. bistriata* e percentual dos principais aminoácidos constituintes.

Plasmídeo	Espidroína	Tamanho do motivo (DNA)	Total de aminoácidos	Glicina (%)	Alanina (%)
pJ244	MaSp1	117 pb	39	48,71	20,51
pJ201	MaSp2	111 pb	35,13	35,13	29,73
pJ241	MaSp1(2×)+(1×)	121 pb	42,15	42,15	25,62

5.2.3 Plasmídeo de clonagem pBSSK+

Todas as clonagens bacterianas a fim de multiplicar os motivos foram realizadas no plasmídeo pBlueScript SK+ (Stratagene). Ele possui resistência à ampicilina, alto número de cópias (300 – 500 por célula) e um sítio de policlonagem no gene de controle LacZ. Isso facilita a seleção de colônias após a transformação de bactérias em células competentes (XL1-Blue) na presença de X-Gal e IPTG, uma vez que as colônias contendo os plasmídeos que contêm o inserto apresentam uma coloração branca, enquanto as que contêm plasmídeos inalterados apresentam coloração azul (SAMBROOK, 2003).

5.2.4 Plasmídeo de expressão em bactéria (*Escherichia coli*).

O plasmídeo pET19b (*plasmid for expression by T7 RNA polimerase* - Novagen) utilizado para a expressão de espidroínas em *E. coli* apresenta a transcrição seletiva da T7 RNA polimerase. Essa enzima transcreve cinco vezes mais rápido do que a polimerase de *E. coli*. Além disso, esse plasmídeo apresenta um promotor T7-lac, que controla a expressão do inserto por meio do repressor lac sendo induzido por IPTG. Ele também contém um sítio de ligação ao ribossomo, terminador T7, códon ATG de início de tradução, a sequência codante para seis histidinas, repressor lacI, com o marcador de resistência a ampicilina (AmpR) e origem de replicação do plasmídeo pBR322, sendo um plasmídeo de médio número de cópias (40-100 por célula).

5.2.5 Plasmídeo de expressão em soja (*Glycine max*)

Para a síntese das proteínas em soja, um *cassete* contendo o promotor semente-específico da β -conglucina de soja foi utilizado. Ele contém o peptídeo sinal da β -Coixina de Coix (*Lacryma jobi*), sendo um endereçador subcelular N-terminal de polipeptídeos para os vacúolos de acúmulo de proteínas de sementes (Figura 9).



Figure 9. Representação da sequência peptídica e nucleotídica do peptídeo sinal da β -coixina. Adaptado de Cunha (2010).

5.3 Construções dos plasmídeos para a engenharia metabólica da bactéria

Os plasmídeos pTet*glyVXY*, pTet*gly2* e pTet*glyVXY-glyA* construídos aumentam a disponibilidade de glicil-tRNA do tipo V, X e Y (*glyVXY*) e da serina metiltransferase (*glyA*) (Figura 10). Esses genes de *E. coli* foram amplificados a partir do genoma da bactéria W3110 (ATTC Ltd, EUA) (XIA et al., 2010). A linhagem bacteriana foi cedida pelo professor Carlos Roberto Jorge Soares do Instituto de Pesquisa Energéticas e Nucleares, IPEN/ CNEN – USP, e o plasmídeo inicial pACYC184 foi adquirido da empresa NEB. O plasmídeo pTet*gly2-glyA* utilizado foi obtido em colaboração com o laboratório do professor David Lee Kaplan, *Tufts University*, Medford, MA, EUA.

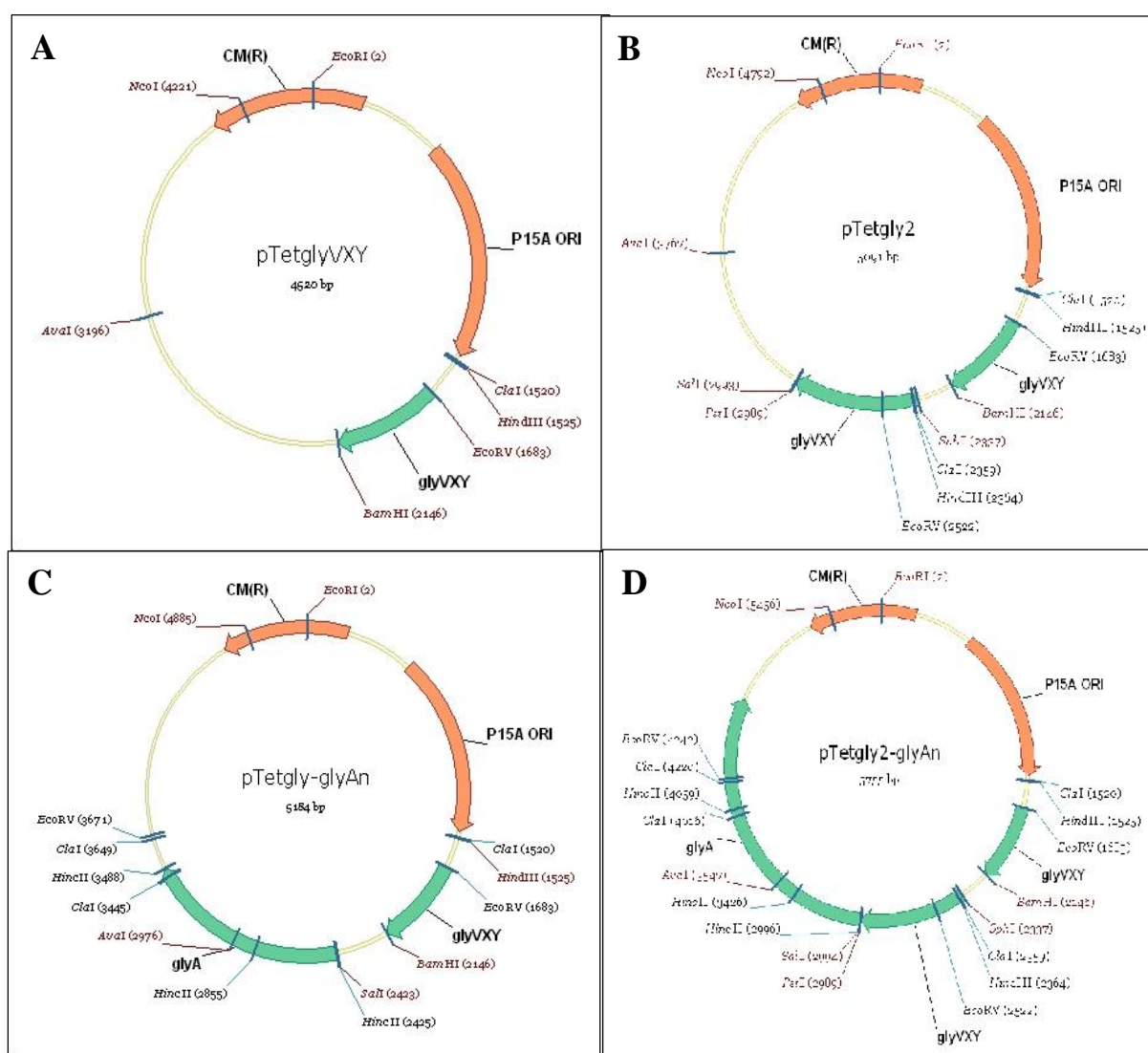


Figure 10. Desenho esquemático dos plasmídeos construídos para aumentar a disponibilidade de glicil-tRNA e de serina hidroximetiltransferase no metabolismo de *E. coli*, utilizando o plasmídeo original pACYC184. As imagens foram montadas no programa NTIVector (versão 1 4.0, Invitrogen)..

Uma alíquota da *E. coli* W3110 foi plaqueada em meio LB agar sem antibiótico. Assim que a placa chegou, duas colônias foram inoculadas em três mililitros de meio LB e cresceram sob agitação de 180 rpm à 37 °C por 16 h. Logo após, as alíquotas foram separadas em dois tubos de 1,5 mL, precipitadas por centrifugação à 500 ×g por 5 min. (mini spin, Eppendorf), o sobrenadante foi descartado, onde uma das alíquotas, de cada amostra, foi diluída em 200 µL de glicerol 10% estéril e gelado, e armazenada em congelador -80 °C. A alíquota de uma amostra foi homogeneizada em 300 µL de TE (Tris-HCl 10 mM pH 8,0; EDTA 1 mM) e a outra em 300 µL de água milli-Q. O DNA genômico foi extraído por fervura dos tubos em água durante 15 min., e armazenadas no congelador (-20 °C).

O plasmídeo pACYC184 é disponibilizado comercialmente na cepa K12 (NEB). Uma alíquota das bactérias contendo o plasmídeo foi distribuída em placa com meio LB agar, tetraciclina (30 µg/mL) e cloranfenicol (40 µg/mL) permanecendo a 37 °C por 16 h. Um pré-inóculo foi preparado em 10 mL de meio LB contendo os mesmos antibióticos da placa e incubado sob agitação de 180 rpm, a 37 °C por 16 h. Essa suspensão bacteriana foi acrescentada em 500 mL (diluição 1:50) de meio LB, em erlenmeyer de um litro, contendo os mesmos antibióticos, e incubados sob agitação de 180 rpm, a 37 °C por aproximadamente três horas. Após esse período, realizou-se a leitura da OD₆₀₀ por espectrofotômetro e, assim que a solução atingiu o valor de 0,6 adicionou-se espectinomicina (50 µg/mL) e a cultura cresceu por mais 16 h no agitador sob as mesmas condições. A extração do plasmídeo foi realizada por meio do *kit Maxi-prep* (Qiagen) utilizando *QIAfilter* (Qiagen) seguindo as recomendações do fabricante. O vetor extraído foi quantificado e uma alíquota foi aplicada em gel de agarose. Todas as construções plasmidiais realizadas a partir desse vetor inicial, também são de baixo número de cópias, sendo necessário realizar o mesmo procedimento descrito.

A fim de confirmar o perfil de restrição do plasmídeo pACYC184, quatro alíquotas foram digeridas com as enzimas: 1) *HincII*, 2) *BamHI* e *EcoR1*, 3) *NheI*, 4) *BspEI*. Os produtos dessas digestões foram submetidos à eletroforese e a amostra original armazenada em congelador (-20 °C).

5.3.1 Amplificação por PCR e confirmação dos genes

Para amplificar o gene serina metiltransferase (*glyA*) de *E. coli*, por PCR, utilizou-se os iniciadores *FglyAn 5'-AAGTCGACAGATTTGATGGCGCGATA-3' e *RglyASal 5'-CCCTGCAATGTAAATGGTTCTTTGG-3' (*GenBank* AC_000091.1) resultando num fragmento de 1734 pb (Xia et al, 2010). As condições da reação foram: solução tampão 1×; Mg₂SO₄ 1,0 mM; dNTP 200 µM (de cada); iniciador sense e antisense 0,2 µM (XIA et al., 2010); enzima *Taq Platinum High Fidelity* (Invitrogen®) 1,0 U e DNA da *E. coli* (W3110) 175 ng, num volume final de 25 µL. Os parâmetros da PCR para esse gene foi com desnaturação inicial de 94 °C por 2 min, seguidos de 40 ciclos com 94 °C por 30 s, 51 °C por 30 s, 68 °C por 1 min 45 s, acrescidos de uma extensão final de 68 °C por 5 min. Todas as PCR foram realizadas em termociclador Biorad®.

Para amplificar os genes *glyVXY* de *E. coli*, por PCR, utilizou-se os iniciadores já descritos na literatura (XIA et al., 2010), FglyEco 5'-GCTCGATATCTAACGACGCAGAAATGCGAAA-3 e RglyBam 5'-CATTGGATCCTAAGATTACAGCCTGAGGCTGTG-3 (*GenBank* AC_000091.1) resultando num fragmento de 479 pb. As condições da reação foram: solução tampão 1×; Mg₂SO₄ 2,0 mM; dNTP 200 μM (de cada); iniciadores sense e antisense 0,2 μM; enzima *Taq Polimerase High Fidelity* (Invitrogen®) 1,0 U e DNA da *E.coli* (W3110) 175 ng, num volume final de 25 μL. Os parâmetros da PCR para esse gene foi com desnaturação inicial de 94 °C por 2 min, seguidos de 40 ciclos com 94 °C por 30 s, 54 °C por 30 s, 68 °C por 1min 45 s, acrescidos de uma extensão final de 68 °C por 5min. Para confirmar os genes amplificados, o volume total dos produtos foi submetido à eletroforese, em seguida purificados, quantificados, conforme já descrito, e uma alíquota de 2 μL foi submetida à eletroforese em gel de agarose para verificar-se a qualidade do produto.

5.3.2 Ligaç o e transforma o

Os produtos purificados referentes aos genes *glyVXY* e *glyA* foram inseridos no plasm deo de clonagem pGEM®-T Easy (Promega®), de acordo com as instru es do fabricante, utilizando a raz o molar 8:1 (inserto:plasm deo), o volume de 20 μL e incubando-se   16 °C por toda a noite. A transforma o foi realizada em c lulas XL-1 Blue por eletropora o, conforme descri o posterior. Todo o volume da transforma o foi transferido para placas 90×15 mm contendo meio LB  gar com ampicilina 200 μg/mL, X-Gal 20 mg/mL (Sigma) e 6 μL IPTG 0,5 M (Sigma). As placas foram fechadas, invertidas, vedadas com parafilme e incubadas em estufa a 37 °C por 16 h. Logo ap s, as placas permaneceram a 4 °C por 2 horas. Somente as col nias com cor branca foram transferidas para crescimento em meio LB, com um palito de madeira autoclavado. O in culo foi realizado em 3 mL de meio LB com ampicilina 200 μg/mL em tubo de 15 mL, permanecendo em agitador a 180 rpm, 37 °C por 16 h.

O volume contendo c lulas em suspens o dos tubos que apresentaram turbidez foi transferido para tubo de centr fuga de 1,5 mL. O plasm deo foi extra do seguindo o protocolo de mini prepara o (Sambrook et al, 2003), quantificado e uma al quota foi submetida   eletroforese em gel de agarose. O gel foi visualizado e fotografado utilizando um

fotodocumentador (Bio Rad®). A confirmação dos insertos foi verificada por digestão enzimática com *EcoRI*, que ocorreu a 37 °C por 2 horas em banho seco (LGC Biotecnologia). Os produtos das digestões foram separados por eletroforese em gel de agarose, fotodocumentados. As amostras foram sequenciadas pela empresa MACROGEN (Coreia do Sul) com os iniciadores universais SP6 e T7. A identidade de cada gene foi analisada pelo programa DNAMAN e apenas as amostras que apresentaram identidade de 100% quando alinhadas à sequência *GeneBank* AC_000091.1 foram utilizadas no experimento.

5.3.3 Gene *glyVXY*

O gene *glyVXY* de *E.coli* foi previamente introduzido no plasmídeo PGEMT Easy (Promega) e após a confirmação por sequenciamento, uma alíquota foi digerida com as enzimas *BamHI* e *EcoRV*. O produto dessa digestão, e do plasmídeo pACYC184, digerido com as mesmas enzimas, foram submetidos à eletroforese em gel de agarose 1,0 %, corados com brometo de etídio, e os fragmentos desejados foram purificados utilizando o *kit Wizard® SV Gel and PCR Clean up System* (Promega) seguindo as recomendações do fabricante. Esses foram religados pelos sítios *BamHI* e *EcoRV* por ação da enzima T4-ligase da Promega. As células de *E. coli* (XL 1 Blue) foram transformadas por eletroporação, com o produto de ligação dos plasmídeos recém construídos, selecionadas com antibiótico adequado, e algumas colônias foram selecionadas para maxi preparação de DNA plasmidial conforme a ser descrito posteriormente. A fim de confirmar o plasmídeo pTet*glyVXY*, três alíquotas desse foram digeridas com as enzimas: 1) *HincII*, 2) *BamHI* e *EcoR1*, 3) *NheI*. Os produtos foram submetidos à eletroforese em gel de agarose a 1% e a amostra original armazenada em congelador (-20 °C).

Para amplificar o segundo fragmento do gene *glyVXY* de *E. coli*, por PCR, utilizou-se os iniciadores já descritos na literatura (Xia et al., 2010), FglySph 5'-GGCTCGCATGCTCATGTTTGACAGCTTATCATCGA-3 e RglySaPs 5' ATTGTCGACTGCTGCAGTAAGATTACAGCCTGAGGCTGTG-3 (*GeneBank* AC_000091.1). As condições da reação foram: solução tampão 1×; Mg₂SO₄ 1,5mM; dNTP 200 µM (de cada); iniciadores sense e antisense 0,2 µM (de cada) (XIA et al., 2010); enzima *Taq Polimerase High Fidelity* (Invitrogen®) 1,0 U e DNA de pTet*glyVXY* 100 ng, num

volume final de 25 μ L. Os parâmetros da PCR para esse gene utilizou uma desnaturação inicial de 94 °C por 2 min, seguidos de 40 ciclos com 94 °C por 30 s, 58 °C por 30 s, 68 °C por 1 min., acrescidos de uma extensão final de 68 °C por 3 min.

Todo o volume da reação foi submetido à eletroforese em gel de agarose 1% corado com brometo de etídio, utilizando o marcador de 1Kb Ladder (Invitrogen®) e o fragmento desejado foi purificado com *kit Wizard® SV Gel and PCR Clean-up System* (Promega®) seguindo as recomendações do fabricante. Logo após realizou-se uma quantificação em espectrofotômetro ND-1000 (NanoDrop®) e uma alíquota de 3 μ L foi submetida à eletroforese em gel de agarose 1% corado com brometo de etídio para verificação da qualidade do material genético. O produto purificado foi inseridos no vetor de clonagem pGEM®-T Easy (Promega®) conforme descrito anteriormente. Uma alíquota foi digerida com a enzima de restrição *EcoRI* a 37 °C e posteriormente submetida à eletroforese em gel de agarose a 1%, corados com brometo de etídio e fotodocumentados, para confirmação do fragmento inserido.

Uma alíquota dos vetores pTetglyVXY e pGEMT-Easy contendo o 2 ° fragmento de *glyVXY* foram digeridos com as enzimas de restrição *SalI* e *SphI* (Promega) e os produtos foram submetidos à eletroforese em gel de agarose, sendo os fragmentos de interesse purificados com *kit Wizard® SV Gel and PCR Clean-up System* (Promega®) de acordo com as recomendações do fabricante. Logo após, realizou-se uma quantificação em espectrofotômetro ND-1000 (NanoDrop®) e uma alíquota de 4 μ L foi submetida à eletroforese em gel de agarose 1% corado com brometo de etídio. O plasmídeo e o fragmento contendo as extremidades digeridas com *SalI* e *SphI* foram religados com a ação da enzima T4 DNA ligase (Promega) originando o plasmídeo pTetgly2. Logo após, os produtos das ligações foram transformados, submetidos a maxi preparação de plasmídeos, conforme a ser descrito posteriormente, e confirmados com digestões por enzimas de restrições: 1) *SalI* e *SphI*, 2) *BamHI* e *EcoRV*.

Os plasmídeos corretos foram amplificados por Maxi-prep (Qiagen) de acordo com as recomendações do fabricante, e foram inseridos na *E. coli* BL21(DE03) originando as linhagens BL21(DE03) pTetglyVXY e BL21(DE03) pTetgly2. Sequencialmente, esses plasmídeos foram utilizados nas construções de pTetglyVXY-*glyA* e pTetgly2-*glyA*.

5.3.4 Gene *glyA*

Para a amplificação do gene serina metiltransferase *glyA* de *E.coli* utilizou-se uma enzima que não adiciona uma adenina na extremidade 5' (*Pfx*, Invitrogen). Para isso, utilizou-se os mesmos iniciadores já descritos anteriormente e a amostra do plasmídeo PGEMT Easy contendo o gene inserido e que foi avaliado por sequenciamento previamente. O fragmento obtido pela PCR respeitaram as seguintes condições de reação: solução tampão 1,5×; Mg₂SO₄ 1,0 mM; dNTP 300 μM (de cada); iniciadores sense e antisense 0,3 μM (de cada) (XIA et al., 2010); enzima Taq Platinum *Pfx* (Invitrogen®) 1,0 U e DNA 300 ng, para um volume final de 25 μL. Os parâmetros da PCR utilizou uma desnaturação inicial de 94 °C por 2 min, seguidos de 40 ciclos com 94 °C por 15 s, 52 °C por 30 s, 68 °C por 2 min. 30 s, acrescidos de uma extensão final de 68 °C por 10 min.

Uma alíquota de 10 μL foi submetida à eletroforese em gel de agarose 1% corado com brometo de etídio utilizando o marcador de 1Kb Ladder (Invitrogen®). Após confirmação do fragmento amplificado, o volume de 15 μL foi digerido com enzima *SalI* (Promega®) a 37 °C por 16 h. O produto foi submetido à eletroforese e purificado com *kit Wizard® SV Gel and PCR Clean-up System* (Promega®) seguindo as recomendações do fabricante. Logo após realizou-se uma quantificação em espectrofotômetro ND-1000 (NanoDrop®) e uma alíquota de 2 μL foi submetida à eletroforese em gel de agarose 1% corado com brometo de etídio para verificação da qualidade do material genético.

Uma alíquota dos plasmídeos pTet*glyVXY* e pTet*gly2* foram digeridos com as enzimas de restrição *SalI* e *HincII* (Promega) e os produtos das digestões foram submetidos à eletroforese em gel de agarose e purificado com *kit Wizard® SV Gel and PCR Clean-up System* (Promega®) seguindo as recomendações do fabricante. Logo após realizou-se uma quantificação em espectrofotômetro ND-1000 (NanoDrop®) e uma alíquota de 4 μL foi submetida à eletroforese em gel de agarose 1% corado com brometo de etídio para verificação da qualidade do DNA. O fragmento amplificado referente ao gene *glyA* e digerido com a enzima *SalI* foi religado com a ação da enzima T4 DNA ligase (Promega) nos plasmídeos pTet*glyVXY* e pTet*gly2* digeridos com *SalI* e *HincII* na razão molar de 3:1 para inserto vetor no volume de 20 μL permanecendo a 16 °C por toda a noite. Logo após os produtos das ligações, que originaram os plasmídeos pTet*glyVXY-glyA* e pTet*gly2-glyA*, foram transformados, submetidos a maxi preparação conforme já descrito, e

confirmados com digestões por enzimas de restrições. Os plasmídeos corretos foram utilizados na (co-)transformação de *E. coli* BL21(DE03) originando as linhagens BL21(DE03) pTetglyVXY-glyA e BL21(DE03) pTetgly2-glyA.

5.4 Construção dos plasmídeos de expressão de espidroínas

5.4.1 Construção dos plasmídeos de expressão das MaSp em *E. coli*

Os plasmídeos pBSSK+ (Stratagene) e pJ244, pJ201, pJ241 contendo os motivos de MaSp1, MaSp2 e MaSp1(2×)+2(1×), respectivamente, foram digeridos com as enzimas de restrição *Xho*I (Sigma Aldrich) e *Bam*HI (Sigma Aldrich), liberando os fragmentos de 147 pb, 114 pb, 372 pb. Os produtos dessas digestões foram separados por eletroforese em gel de agarose 1% corados com brometo de etídio, e os fragmentos nucleotídicos de interesse foram purificados utilizando o *kit Wizard® SV Gel and PCR Clean up System* (Promega, USA), seguindo as recomendações do fabricante. Os fragmentos purificados foram reclonados entre os sítios *Xho*I e *Bam*HI por ação da enzima T4 DNA ligase (Promega, USA) no plasmídeo pBSSK+. Para isso, o material genético foi quantificado em NanoDrop 2000c (Thermo Scientific) e a reação de ligação utilizou a razão molar 3:1 (inserto/vetor), tampão da enzima à concentração 1×, 3U da enzima permanecendo à 4 °C por toda à noite. As células de *E. coli*. XL-1 Blue competente foram transformadas por eletroporação com o produto de ligação dos vetores pBSSK MaSp1 1×, pBSSK MaSp2 1×, pBSSK Masp(1+2) 1×, sendo adicionados o volume de um microlitro (5-10 ng) à suspensão celular e incubados no gelo por um minuto. Logo após, a solução foi transferida para um cuveta estéril (Gene pulser cuvette, 165-2089, Bio-Rad) e submetida à corrente elétrica de 1,8 kV em eletroporador (Micropulser, Bio-Rad). O volume de 800 µL de meio SOC foi adicionado, homogeneizado e a suspensão de células, transferidas para um tubo de 1,5 mL e incubadas por uma hora à 37 °C em banho seco (DB100, LGC Biotecnologia). Todo o volume foi aplicado em placas 90×15 mm contendo meio LB agar e ampicilina 200 µg/mL sendo incubadas à 37 °C por toda à noite (Incubator Shaker Series I26, New Brunswick Scientific). Algumas colônias brancas foram selecionadas e inoculadas em 3 mL de meio LB com ampicilina 200 µg/mL, para a realização de mini-preparação de DNA plasmidial (Sambrook, 2003). Esses plasmídeos iniciais possuem somente o motivo inicial. As construções da MaSp1 são uma continuação do trabalho de

Oliveira (2008), utilizando o plasmídeo pBSSKMaSp1 contendo 16 vezes o motivo inicial para atingir a multiplicação desejada.

Na natureza, as aranhas apresentam espidroínas caracterizadas por grandes repetições de motivos iniciais em um único gene. Na tentativa de mimetizar a produção proteica em sistema heterólogo em laboratório, uma metodologia foi desenvolvida para clonar esses módulos em *tandem*, descrita previamente por Lewis e colaboradores (1996), e observada na figura 11. Os plasmídeos iniciais para cada motivo foram separados em duas alíquotas. A primeira foi digerida com as enzimas de restrição *ScaI* e *BspEI*, e a outra com as enzimas *XmaI* e *ScaI*.

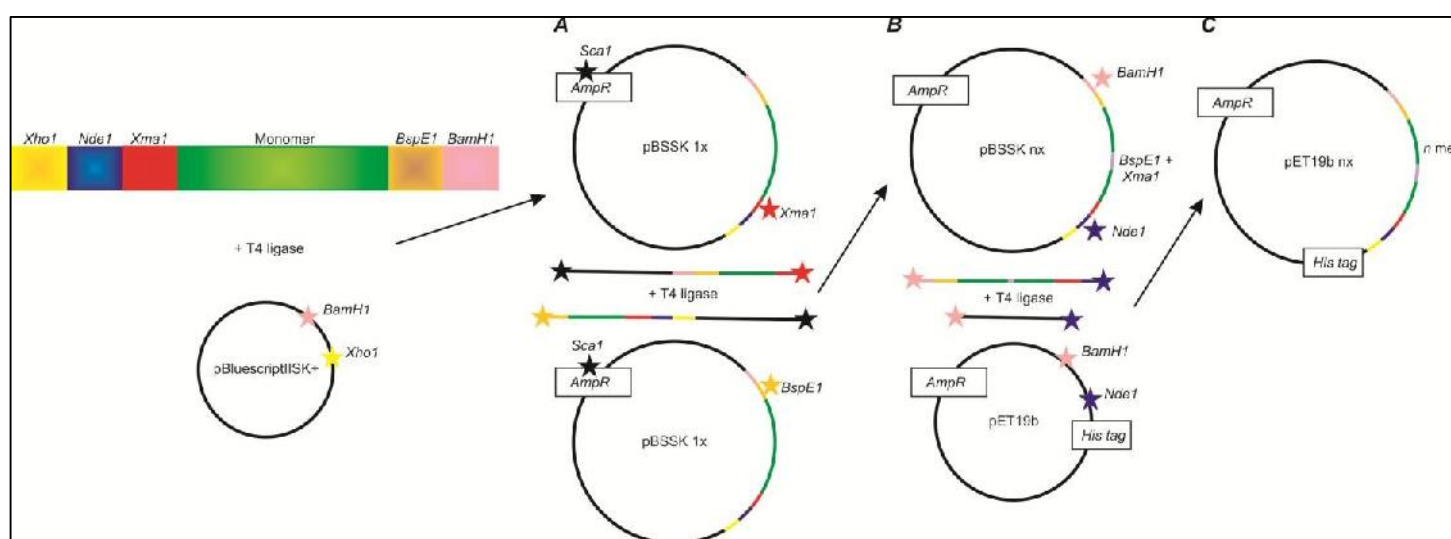


Figura 11. Desenho representativo para as construções dos plasmídeos contendo os motivos das espidroínas. O motivo inicial da aranha *Parawixia bistriata* (em verde) contém sequências de enzimas de restrição *XhoI* (amarelo), *NdeI* (azul), *XmaI* (vermelho), *BspEI* (dourado) e *BamHI* (rosa). O motivo é inserido no plasmídeo de clonagem pBSSK+ pelos sítios das enzimas *XhoI* e *BamHI* (estrelas rosa e amarela no plasmídeo pBluescriptIIISK+), resultando no plasmídeo contendo uma vez o motivo (pBSSK1x). Em seguida, duas alíquotas independentes desse plasmídeo sofrem digestões duplas (A): uma é digerida com as enzimas *ScaI* e *XmaI* (plasmídeo superior-estrela preta à estrela vermelha), enquanto a outra é digerida com *ScaI* e *BspEI* (plasmídeo inferior-estrela preta à estrela amarela). Apenas os fragmentos que possuem o motivo (em verde) em sua sequência são selecionados (A-fragmentos ao centro, entre os plasmídeos inferior e superior). A reação de ligação (B) reestabelece o gene de resistência à ampicilina (sítio de *ScaI*-estrelas pretas) e une as extremidades formadas por *XmaI* e *BspEI* (estrelas vermelha e amarela), mas sem reestabelecer os respectivos sítios de restrição (B-parte cinza entre duas partes verdes do plasmídeo superior). Esse plasmídeo contém o dobro de motivos do plasmídeo inicial. Esses procedimentos (A e B) podem ser repetidos inúmeras vezes até se atingir a quantidade de motivos desejada. Logo após, o fragmento contendo os $n \times$ motivos das espidroínas é retirado do plasmídeo de clonagem (B-plasmídeo superior) por digestão das enzimas *BamHI* e *NdeI* (estrelas rosa e azul) e introduzido no plasmídeo de expressão em *E. coli* pET19b (B-plasmídeo inferior), formando o plasmídeo de expressão contendo a sequência de espidroínas (C). Esse apresenta uma sequência codante de histidina na porção N-terminal. (Adaptado de TOKAREVA et al., 2013).

Os fragmentos foram separados por eletroforese em gel de agarose a 2%, corado com brometo de etídio, e purificados com o kit Wizard® SV Gel and PCR Clean up System (Promega) seguindo as recomendações do fabricante. A enzima *ScaI* corta o gene de resistência à ampicilina, um fator importante, pois somente os plasmídeos que se reestabeleceram durante a ligação dos fragmentos proporcionaram o crescimento bacteriano. As extremidades coesivas das digestões com as enzimas *XmaI* e *BspEI* são compatíveis, mas

ao serem religados, os sítios dessas enzimas não se reestabelecem, outro fator importante para a estratégia. Assim, as amostras são submetidas novamente aos cortes das enzimas descritas para a construção dos plasmídeos com 2×, 4×, 8×, 16×, 32× motivos. Uma vantagem dessa metodologia é o controle exato da quantidade de motivos a cada clonagem. A confirmação dos motivos era realizada a cada plasmídeo de clonagem construído, por meio de digestão de uma alíquota do DNA com enzimas as *Bam*HI e *Xho*I. Além disso, todas as amostras foram enviadas para sequenciamento à empresa MacroGen® (Coréia do Sul).

Os plasmídeos de clonagem construídos contendo a sequência de DNA equivalente à expressão de uma proteína 50 kDa, e o plasmídeo de expressão pET19b foram digeridos com as enzimas *Nde*I e *Bam*HI. Os produtos das digestões foram separados por eletroforese em gel de agarose 1% corados com brometo de etídio, e os fragmentos de interesse foram purificados utilizando o kit *Wizard® SV Gel and PCR Clean up System* (Promega) seguindo as recomendações do fabricante. Os fragmentos purificados foram reclonados entre os sítios *Nde*I e *Bam*HI por ação da enzima T4 DNA ligase (Promega). Para isso, o material genético foi quantificado em espectrofotômetro NanoDrop 2000c (Thermo Scientific) e a reação de ligação utilizou a razão molar 3:1 (inserto/vetor), tampão da enzima à concentração 1×, 3U da enzima permanecendo por toda a noite à 4 °C. As *E. coli.* do tipo XL 10 Gold competentes foram transformadas por choque térmico, com o produto de ligação dos vetores pET19b MaSp1(16×); pEt19bMaSp1(32×); pET19bMaSp2 (16×), pET19bMaSp2 (32×), pET19b[MaSp1(2×)+Masp2(1×)]8×. Para isso, dois microlitros de produto de ligação (10-20 ng) foram adicionados ao volume contendo suspensão celular e, os tubos foram incubados no gelo por 30 min. Logo após, as amostras foram inseridas em banho maria (PolyScience, USA) à 42 °C durante um minuto e, retornadas ao gelo por mais 5 min. O volume de 800 µL de meio SOC (S1797, Sigma Aldrich, USA) foi adicionado, homogeneizado e permaneceram por uma hora à 37 °C em banho seco. Todo o volume foi aplicado em placas 90×15 mm contendo meio LB agar e ampicilina 200 µg/mL sendo incubadas a 37 °C por toda a noite.

Seis colônias de cada transformação foram inoculadas em meio LB, contendo antibiótico, e realizou-se a mini-preparação de DNA plasmidial segundo Sambrook (2003). Esses produtos foram quantificados por espectrofotômetro NanoDrop 2000c (Thermo Scientific), uma alíquota foi submetida à eletroforese em gel de agarose corado com brometo de etídio, para se verificar a qualidade do plasmídeo e, em seguida, outra alíquota foi digerida com as enzimas *Apa*I e *Bam*HI, a fim de confirmar o fragmento clonado. Os produtos das digestões foram submetidos à eletroforese em gel de agarose a 1% corado com brometo de etídio.

5.4.2 Construções dos plasmídeos de expressão das MaSp em soja (*Glycine max*).

O plasmídeo pJ201 sintetizado pela empresa DNA 2.0 contém o motivo inicial da espidroína MaSp2 de *P. bistriata* e sítios de endonuclease para a construção do plasmídeo de expressão em soja (*Glycine max*), apresentando ainda uma sequência referente à cauda de histidina na extremidade N-terminal. No entanto, o motivo da espidroína foi retirado desse plasmídeo por digestão das enzimas *XmaI* (Promega, Madison, WI, USA) e *BspEI* (NEB, Ipswich, MA, USA), e os motivos *in tandem* de MaSp2 16× e MaSp2 32×, foram inseridos nessa região por ação da enzima T4 DNA ligase da Promega (M1804, Madison, WI, USA), sendo denominados pJMaSp2 16× e pJMaSp2 32×. Para isso, os produtos das digestões foram separados, purificados, quantificados e submetidos à reação de ligação, conforme metodologia já descrita. Os plasmídeos recém-construídos foram inseridos em bactéria DH5 por eletroporação, essas foram cultivadas e os plasmídeos extraídos por mini preparação (Sambrook, 2003). As sequências dos genes das espidroínas MaSp2 16× e MaSp2 32× foram digeridas com as enzimas *HindIII* (Promega, Madison, WI, USA) e *EcoRI* (Promega) e inseridos no plasmídeo de expressão em soja p cong pelos mesmos sítios de restrição.

5.5 Transformações

5.5.1 Eletroporação

Para realizar as transformações bacterianas, as células *E. coli*. XL10-GOLD foram tratadas para serem competentes por eletroporação. As construções nos plasmídeos pJ, (pGEMT-Easy e pBSSK+ foram realizadas por essa metodologia. As alíquotas de células competentes foram descongeladas em gelo e 1-2 µl de DNA (1-20 ng) era(m) adicionado(s) às bactérias, e incubadas no gelo por 1 minuto. Após este período, a solução bacteriana/DNA foi transferida para uma cubeta de eletroporação previamente resfriada, com um espaço de 0,1 cm entre os eletrodos (BioRad). A corrente elétrica foi transmitida por um eletroporador (BioRad) ajustado para as seguintes condições: resistência de 200 Ω, capacitância de 25 µF e uma

voltagem de 1,8 kV. Foi aplicado apenas um pulso, onde o tempo ideal de descarga variou entre 4-5 ms para se obter uma boa eficiência. Imediatamente após a eletroporação foram adicionados 600 µL de meio SOC, e as células foram suspensas cuidadosamente e transferidas para um tubo de polipropileno de 1,5 mL. Estas foram incubadas em banho-maria a 37 °C durante 1 hora. Após esta etapa, foram plaqueados diferentes volumes da cultura bacteriana dependendo do material a ser transformado. No caso de produto de ligação de DNA todo o volume de 600 µL de solução contendo suspensão celular era transferido para 15 mL de meio LB agar, em placa de 90 mm, contendo o antibiótico adequado. Caso a transformação fosse realizada a partir de plasmídeo pronto, o volume de 50 µL de solução contendo suspensão celular era transferido para 15 mL de meio LB agar, em placa de 90 mm, contendo o antibiótico adequado. As placas eram finalmente vedadas e incubadas a 37 °C durante uma noite para o crescimento de colônias.

5.5.2 Choque térmico

Para realizar as transformações bacterianas e as produções proteicas, as células *E. coli*. BL21(DE03) eram co-transformadas com plasmídeos para alterar o metabolismo da bactéria e para produção de espidroínas. Todas as seis linhagens de bactérias [BL21(DE03), BL21(DE03)pACYC184, BL21(DE03)pTetglyVXY, BL21(DE03)pTetgly2, BL21(DE03)pTetglyVXY-glyA e BL21(DE03)pTetgly2-glyA] foram transformadas com os plasmídeos pET19b sem inserto, pET19b MaSp2 16×, pET19b MaSp2 32×, pET19b MaSp1 16×, pET19b MaSp1 32×, pET19b MaSp1+2 8×.

As alíquota de células competentes (100 µL) foram descongelada em gelo (~4 °C) e misturadas com 1-5 µL de DNA plasmidial (0,2-100 ng). Os tubos permaneceram no gelo por 30 minutos, e após este período foram transferidos para um banho maria a 42°C por 90 segundos. As amostras foram transferidas imediatamente para o gelo e incubadas por 2 minutos. Um volume de 500 µL de meio LB foi acrescentado em cada amostra, para serem incubados em banho seco à 37 °C durante 1 hora. Após esta etapa, 50 µL de suspensão celular foi plaqueada em meio LB agar, em placa de 90 mm, e incubadas a 37 °C por toda a noite para o crescimento das colônias..

5.6 Curva de crescimento

As linhagens de bactérias construídas foram plaqueadas em meio LB Ágar e, depois pré-inoculadas em três mililitros de meio LB a 37 °C, 200 rpm por 16 horas (New Brunswick). No dia seguinte, o volume de 200 µL foi inoculado em 25 mL de meio LB com uma ponteira de filtro, em erlenmyer de 125 mL, sendo mantidos sob as mesmas condições já descritas. O acompanhamento de crescimento celular foi realizado por meio de leitura em espectrofotômetro no comprimento de onda de 600 nm (OD_{600}) a cada hora, iniciando-se no momento do inóculo inicial e finalizando após 10 horas de incubação, ocorrendo em triplicata sem e com indução a IPTG 0,4 mM. O volume de um mililitro foi utilizado na leitura da OD_{600} 0,6, depois foram realizadas diluições de 1:10 (v:v). A figura 12 ilustra essa parte do experimento.

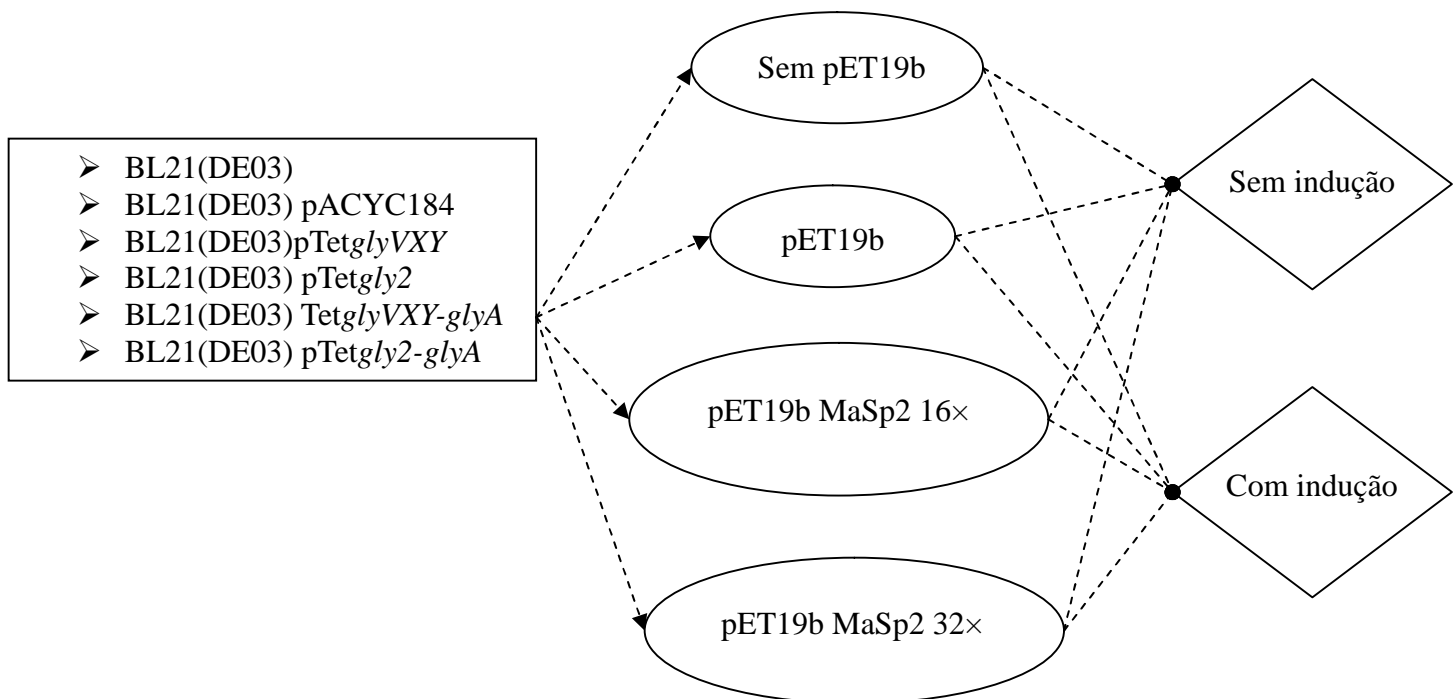


Figura 12. Fluxograma ilustrativo do experimento para a curva de crescimento das bactérias. A linhagem BL21(DE03) foi transformada com os plasmídeos construídos para a engenharia metabólica de bactéria, conforme o quadrado à esquerda, sendo também co-transformadas com os vetores de expressão pET19b, pET19bMaSp2 16× e pET19bMaSp2 32×, circunferências ao centro. O crescimento com e sem indução de IPTG 0,4 mM foi acompanhado durante 10 horas, para todas as linhagens, em triplicata. A bactéria BL21(DE 03), os plasmídeos pACYC184 e pET19b foram utilizados como controles.

Ao final do crescimento bacteriano, a solução contendo suspensão celular foi centrifugada em tubos de 50 mL à $10.000 \times g$, a 4°C por 7 min. (Eppendorf), resuspensas em 1 mL de tampão de lise $1\times$, conforme já descrito, e mantidas em congelador -80°C . Esse material foi utilizado nas análises de proteômica de segunda geração, sendo descritas posteriormente.

5.7 Análise por espectrometria de massa biotyper

A partir das linhagens de bactérias *E. coli* estoque, armazenadas em glicerol 20% e congelador -80°C , diluiu-se um microlitro em 1 mL de meio LB. Logo após, o volume de 50 μL foi plaqueado em meio LB agar, em placas de petri descartável estéril 90 mm (JProlab cod. 263, São José dos Pinhais, PR, Brasil), sem antibiótico para a linhagem BL21(DE03), e contendo cloranfenicol 25 $\mu\text{g}/\text{mL}$ para as linhagens BL21(DE03): pLys S, pACYC184, pTetglyVXY, pTetgly2, pTetglyVXY-glyA, pTetgly2-glyA, sendo incubadas em estufa seca (WTC Binder Série KB 53(E31), Tuttlingen, Germany) à temperatura de 37°C por toda à noite. Para o monitoramento de contaminação a cada período de incubação foram utilizadas placas sem e com o antibiótico contendo apenas o meio de cultura. No dia seguinte, novas placas foram divididas em 8 quadrantes iguais, com auxílio de uma caneta de marcação permanente, na superfície da placa contendo o meio LB agar. As placas foram retiradas da estufa e oito colônias de cada linhagem foram selecionadas, para serem repicadas de modo que, ao encostar a ponta da ponteira estéril em uma colônia, a mesma era encostada em cinco pontos diferentes dentro do mesmo quadrante, como mostra a figura 13.

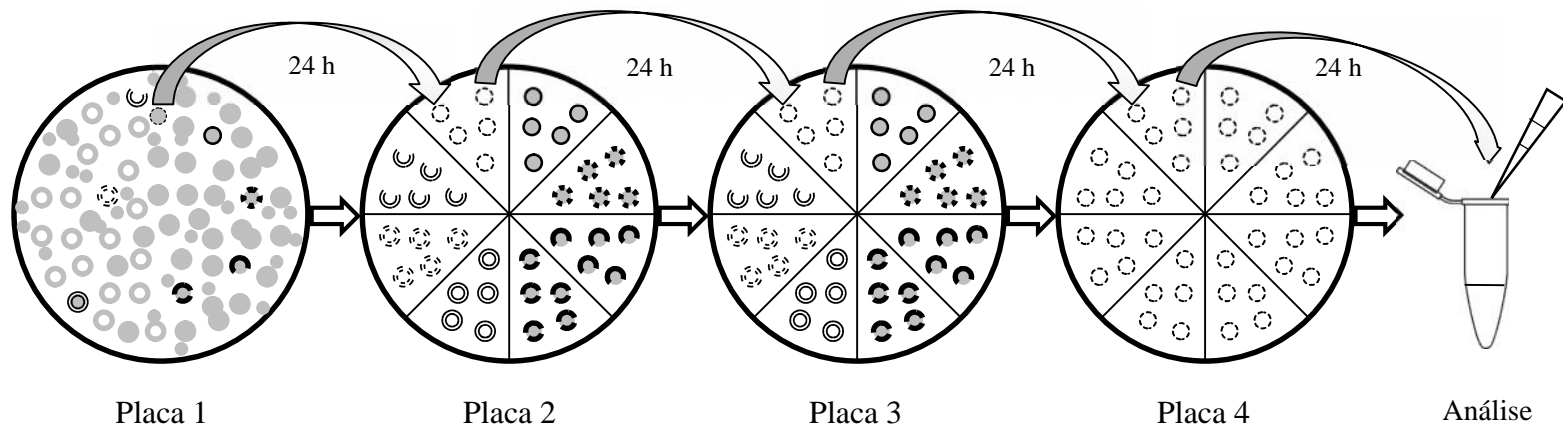


Figura 13. Desenho representativo da metodologia utilizada no cultivo das linhagens bacterianas para as análises por espectrometria de massa *biotyper*. As colônias de bactérias foram isoladas no primeiro dia em meio LB Agar a 37 °C (Placa 1), selecionadas e repicadas em 5 pontos nas mesmas condições por mais 48 h (Placas 2 e 3). Uma das colônias foi repicada em 30 pontos em outra placa por mais 24h (Placa 4), sendo que 24 foram transferidas para um tubo de prolipropileno para análise.

Após o período de 24 h, os cinco pontos que receberam as bactérias deram origem a novas colônias, idênticas à colônia de origem. Novamente, uma das colônias foi escolhida, para ser repicada em cinco pontos diferentes numa nova placa com meio de cultura (Figura 13). Durante o período de incubação as placas foram mantidas no mesmo nível, uma ao lado da outra, a fim de minimizar diferenças de temperatura. Após o crescimento da segunda repicagem (Placa 3, Figura 13), uma colônia foi selecionada para ser repicada em 30 pontos (Placa 4, Figura 13). Após o período de 24 h, a placa foi retirada da estufa e 24 colônias foram transferidas para 300 µL de água ultra pura em tubos de 2 mL (Axygen, USA) com auxílio de ponteiras de 10 µL estéreis. A solução contendo suspensão celular foi homogeneizada em vórtex por 10 s, adicionou-se o volume de 900 µL de etanol 100% e as amostras foram precipitadas por 2 minutos à 12.100 ×g. Após descartar o sobrenadante e secar o precipitado, por 30 minutos em temperatura ambiente, ele foi dissolvido em 300 µL de ácido fórmico 70 % (288, Vetec) com auxílio da ponteira. Logo após, o volume de 30 µL acetonitrila foi adicionado, as amostras foram homogeneizadas no vórtex por 10 s e centrifugadas à 12.100 ×g por 2 min. O sobrenadante foi transferido para um tubo de 0,2 mL (Axygen, USA), e utilizou-se o volume de 1 µL das amostras para realizar a co-cristalização com 1 µL de matriz α -amilase, aplicando-se os volumes diretamente na placa MSP96 polisher (part n° 224990, Bruker). A placa foi secada em fluxo laminar vertical e a análise ocorreu em espectrômetro de massa Microflex LRF (Bruker) integrado ao sistema de MALDI *Biotyper* para identificar os íons positivos e negativos das massas proteicas de cada linhagem.

5.8 Estudo proteômico das linhagens bacterianas

A partir da linhagem bacteriana BL21(DE03) foram realizadas transformações com os plasmídeos pTetglyVXY, pTetgly2, pTetglyVXY-glyA, pTetgly2-glyA, utilizando pACYC184 como controle, e co-transformações desses plasmídeos com pET19b, pET19bMaSp2(16×) e pET19bMaSp2(32×). O inóculo foi realizado em 10 mL de meio LB com antibiótico adequado sob agitação de 180 rpm, à 37 °C por 16 h (I26 New Brunswick). Logo após, o volume de 1 mL de cada amostra foi transferido para um erlenmeyer de 125 mL contendo 25 mL de meio LB, nas mesmas condições do pré-inóculo. A indução foi realizada, em triplicata, com IPTG 0,4 mM em OD₆₀₀ 0,6 durante quatro horas. As amostras foram precipitadas à 10.000 ×g, 10 min., 4 °C em tubos de 50 mL e mantidos a -80 °C por toda a noite. O extrato proteico foi adquirido conforme descrito anteriormente.

Um total de 50 µg de extrato proteico de cada amostra foi precipitada com acetona, conforme já descrito, sendo o precipitado submetido ao protocolo descrito por Murad e colaboradores (2011). As amostras foram solubilizadas em 60 µL de bicarbonato de amônio 50 mM, adicionando-se 25 µL de solução *RapiGest SF* 0,2 % (Waters). Essas foram homogeneizadas sob agitação (vórtex), aquecidas a 80 °C por 15 min. e precipitadas por centrifugação máxima durante 10s. Logo após, adicionou-se 2,5 µL de DTT 100 mM ao volume, sendo mantido por 30 min a 60 °C. Os tubos foram resfriados em temperatura ambiente, e o volume foi precipitado por centrifugação máxima durante 10 s. Logo após, as amostras receberam 2,5 µL de iodoacetamida 300 mM, homogeneizadas sob agitação (vórtex), e permaneceram por 30 min. num ambiente escuro. Foi adicionado o volume de 10 µL de solução de tripsina (Promega), as amostras foram homogeneizadas e permaneceram à 37 °C por toda a noite. Logo após, adicionou-se 10 µL de TFA 5%, homogeneizou-se as amostras no vórtex e foram incubadas a 37 °C por 90 min. Finalmente, os tubos foram centrifugados à 20.817 ×g por 30 min a 6 °C, o sobrenadante (~110 µL) foi transferido para um tubo *Waters Total Recovery* e o volume de 10 µL de fosforilase B (7 fmol/µL, 85 µL de acetonitrila 3%, ácido fórmico 0,1%) foi adicionado à amostra. Ao final, a concentração da proteína é de 250 ng/µL e a do controle interno padrão 50 fmol, num volume 200 µL, sendo analisadas em equipamento nano UPLC-MS (Waters) conforme descrito por Murad e Rech (2012).

Os resultados foram analisados e processados no programa *proteinLynx global Server* e identificados quanto a função celular no *Blast2go*®.

5.9 Produção de espidroínas em E.coli

As produções das espidroínas MaSp2 16× e MaSp2 32× foram realizadas em linhagens bacterianas comerciais e construídas nesse experimento. Uma colônia recombinante contendo os plasmídeos de expressão de cada linhagem bacteriana (BL21(DE03), BL21(DE03)pACYC184, BL21(DE03)pTetglyVXY, BL21(DE03)pTetgly2, BL21(DE03)pTetglyVXY-glyA, BL21(DE03)pTetgly2-glyA), foi inoculada em 20 mL de meio LB contendo o antibiótico adequado e cultivada à 37 °C sob agitação de 200 rpm por 16 h em (I26 New Brunswick). O volume de 10 mL dessa cultura foram inoculados em 1.000 mL (diluição 1:100) de meio LB com antibiótico adequado e cultivadas à 37 °C com agitação de 220 rpm até atingir a OD₆₀₀ 0,6. Então adicionou-se IPTG 1 mM para induzir a expressão durante 4 horas. Após esse período, a solução foi precipitada por 15 min. a 8.000 ×g a 4 °C. O precipitado foi diluído em tampão de lise (Tris50 mM - pH 8,0; NaCl 10 mM; MgCl₂10 mM) contendo lizozima 1 mg/mL (SIGMA) na proporção de 3:1 [m(g)/v(mL)], sendo mantidos por toda a noite a -80 °C em tubos de 50 mL (TPP).

5.10 Extração proteica e determinação da solubilidade da proteína alvo

O protocolo foi adaptado do manual “*A handbook for high-level expression and purification of 6x-His-tagged proteins*” (Qiagen). O volume de 1 µL de suspensão celular bacteriana, armazenada em glicerol 15% a -80 °C, foi inoculado em 10 mL de meio LB, permanecendo a 37 °C sob agitação por toda a noite (I26 New Brunswick). Esse inóculo foi transferido para 500 mL de meio LB com antibiótico e permaneceu sob agitação de 300 rpm a 37 °C até atingir a OD₆₀₀ 0,6. Uma alíquota de um mililitro dessa amostra foi precipitada, por centrifugação máxima por 2 minutos, e diluídas em 50 µL de tampão de corrida 1× (*NUPAGE LDS Sample Buffer 4×*, Invitrogen) com -mercaptoetanol. Essa amostra serviu como controle negativo do processo a fim de verificar a ausência da proteína alvo nessa etapa. Após adicionar IPTG 1 mM, as amostras permaneceram por mais quatro horas nas mesmas condições. Logo após, uma outra alíquota de um mililitro foi separada, precipitada e diluída em tampão de corrida, servindo como controle positivo da proteína alvo. O restante do volume foi precipitado a 8.000 ×g por 20 min. a 4 °C, as amostras foram diluídas em tampão

de lise, conforme já descrito, e congeladas a $-80\text{ }^{\circ}\text{C}$ por toda a noite. As amostras descongelaram em temperatura ambiente, e a extração ocorreu conforme descrito por Teulé e colaboradores (2009), com modificações. Adicionou-se lisozima 1 mg/mL (Steinheim, Alemanha, Sigma) nas amostras, permanecendo no gelo por 30 min. Em seguida, a solução recebeu ácido deoxicólico 40 mg/mL (Itália, Sigma), e sendo mantidas por 30 min. a $37\text{ }^{\circ}\text{C}$, 200 rpm (I26 New Brunswick). Logo após, foi acrescentado DNase1 $10\text{ }\mu\text{g/mL}$ (Steinheim, Alemanha, Sigma), permanecendo no gelo por 30 min e posteriormente centrifugados à $20.000\times g$, $20\text{ }^{\circ}\text{C}$ por 30 min. O sobrenadante (S1) foi transferido para outro tubo e o precipitado (P1) foi solubilizado em 5 mL de tampão de lise. Uma alíquota de $20\text{ }\mu\text{L}$ de solução proteica, do sobrenadante e do precipitado, foram guardadas para eletroforese em gel de poliacrilamida (SDS-PAGE), com tampão de corrida a $1\times$ (*NUPAGE LDS Sample Buffer 4\times*) contendo β -mercaptoetanol. Sequencialmente, o sobrenadante (S1) foi aquecido a $80\text{ }^{\circ}\text{C}$ por 20 min. e esse volume foi precipitado novamente, nas mesmas condições do primeiro, resultando em um novo precipitado (P2) e sobrenadante (S2). Esse segundo sobrenadante, teoricamente contendo a proteína alvo, foi armazenado a $-80\text{ }^{\circ}\text{C}$ até o momento da purificação proteica. Uma alíquota de $20\text{ }\mu\text{L}$ do S2 e do P2 foram retiradas para serem submetidas à SDS-PAGE. Todas as alíquotas de S1, P1, S2 e P2 foram aquecidas a $95\text{ }^{\circ}\text{C}$ por 10 min, e posteriormente centrifugadas à $15.000\times g$ por 1 min. A SDS-PAGE foi realizada utilizando gel $1,0\text{ mm}$, $4\text{-}12\%$, pré-moldado Novex (Invitrogen), com o volume $15\text{ }\mu\text{L}$ de cada amostra foi, utilizando $5\text{ }\mu\text{L}$ de padrão de corrida (*Novex® Sharp Pre-stained Protein Standard*, Estados Unidos, Invitrogen), tampão de corrida MOPS (Invitrogen) e mantendo a voltagem constante de 160 V , seguindo as condições do fabricante. O gel foi corado com azul de Coomassie utilizando o *Colloidal Blue Staining Kit* (Invitrogen), permanecendo sob agitação por 10 min. em 50 mL de solução de fixação (metanol 50%, água destilada 45%, ácido acético 5%). Em seguida, o volume foi descartado e o gel, transferido para solução de coloração (metanol 20%, solução de coloração A 20%) num volume de 50 mL e permanecendo por 10 min. sob agitação, para posterior adição de 2,5 mL de corante B do *kit*. O gel permaneceu nessa solução durante uma hora e, em seguida, foi descorado em 200 mL de água destilada e permanecendo sob agitação por toda a noite .

Para a secagem do gel , o mesmo permaneceu durante 5 min. em solução de secagem *Gel-Dry Drying solution* (Invitrogen), conforme as recomendações do fabricante, e ao mesmo tempo 2 folhas de papel celofane foram completamente embebidas na mesma solução por 20 s. O suporte de secagem (*drying frame*) foi aberto e uma das membranas foi

aderida. Em seguida, retirou-se o “pé” e as finalizações superiores do gel com uma faca, e o gel foi transferido para o papel celofane, sem formar bolhas entre as camadas. Acrescentou-se um pouco mais da solução de secagem sobre o gel e, só então a outra folha de papel celofane foi aderida sobre o gel, prevenindo-se para não formar bolhas. A outra parte do suporte de secagem foi colocada e a estrutura adquiriu suporte ao se encaixar 3 grampos de segurança. O lado que não foi firmado, serviu para abrir um pouco o suporte, e retirar-se o excesso de solução presente entre as membranas e o gel. Logo após, o último grampo foi encaixado, e o sistema foi deixado na vertical para secar durante 24 horas.

5.11 Detecção e análise das proteínas

Uma quantidade de 100 µg do extrato proteico total foi submetida à *SDS-PAGE* conforme já descrito. Logo após, os géis de poliacrilamida foram mergulhadas em solução de transferência, permanecendo sob agitação por 15 min., embebidas em metanol e em seguida as proteínas foram transferidas para uma membrana de PVP (Milipore-Immobilon P 0,45 µm) por meio de um *Trans-Blot® SD Semi-Dry Transfer Cell* (BioRad). Após desmontar o sistema, a membrana foi molhada com água destilada, embebida novamente em metanol e permaneceu sob bloqueio em BSA 2,5% em PBS 1% por 30 min sob agitação. Após o bloqueio a membrana foi incubada por 30 min. em solução de BSA 2,5% em PBS 1× com anticorpo monoclonal antipoli-histidina conjugado à fosfatase alcalina produzida em camundongo (SIGMA) diluído 5.000 vezes. Em seguida, ela foi lavada três vezes em solução tampão PBS 1× com tween 0,05 % e mais três vezes com PBS 1×. Enfim, a membrana foi imersa em 5 mL solução de revelação (Tris-HCl 0,1 M pH 9,5; NaCl 0,1 M; MgCl₂ 0,05 M) contendo substratos cromogênicos para fosfatase alcalina 50 µL de NBT 50 mg/mL e BCIP 25 mg/mL. A solução foi homogeneizada com a membrana e, assim que as bandas das proteínas foram visualizadas, acrescentou-se água destilada a fim de parar a reação. A membrana foi colocada sobre papel filme para secar e depois as mesmas foram armazenadas.

Para estimar-se a massa molecular, o coeficiente de extensão e o ponto isoelétrico das proteínas MaSp2 16× e MaSp2 32×, a sequência de aminoácidos foi informada na página da rede mundial de computadores www.expasy.org/protparam/. A massa molecular da proteína total foi confirmada por espectrometria de massa MALDI-TOF utilizando o equipamento Microflex™ (Bruker®) no modo linear positivo, com alcance entre 26 e 226 kDa e

intensidade do laser em 90 %. Para isso, uma alíquota das amostras foi separada após o procedimento de diálise e liofilização (~1 mg). As proteínas solúveis foram previamente quantificadas utilizando o espectrofotômetro Nanodrop 2.000C (Thermo Scientific), por meio de absorvância à 280 nm, informando a massa molecular e coeficiente de extensão. A quantidade de 1 mg ácido sinapínico (S8313, Sigma) foi diluído em Acetonitrila 50% (A-8596, Sigma) e TFA 0,05% (T3443, Sigma) num volume final de 100 µL em um pedaço de parafilme (Laboratory film, BEMIS).

5.12 Purificação das proteínas produzidas em *E. coli*

As biomassas de *E.coli* após a indução de espiroínas foram descongeladas em temperatura ambiente e a solução foi ajustada para concentração final NaCl 50 mM e Imidazol 5 mM pH 8. Logo após, o volume foi filtrado com membrana de polietersulfona de 0,45 µm (83.1826-SARSTEDT) e posteriormente de PVDF de 0,22 µm (SLGV033RS-Millex-GV). A purificação das proteínas recombinantes de MaSp 2 16× e 32 × ocorreu por cromatografia de afinidade por íons metálicos (Níquel), adicionando-se 3 mL de Ni-NTA Agarose (Qiagen, Hiklen) na solução proteica, mantendo-se a resina junto com a amostra sob agitação por 1 hora à 4 °C. A suspensão passou pela coluna seguida de duas lavagens com solução tampão contendo imidazol 40 mM, e logo após, a proteína foi eluída por duas vezes em solução tampão contendo de imidazol 100 mM e mais duas vezes com imidazol 250 mM. As frações eluídas foram dialisadas contra solução tampão Tris 10 mM pH 8 por 6 horas e contra água destilada à 4 °C por 48 horas sob agitação, e depois liofilizadas. As massas dos tubos de 50 mL foram medidas antes e depois da liofilização, a fim de se obter a quantidade de proteína recombinante produzida.

Para verificar a qualidade do extrato proteico total, a quantidade de 100 µg de cada amostra foi separado para a *SDS-PAGE*. Para isso, as amostras foram quantificadas pelo espectrofotômetro NanoDrop 2000C (Thermo Scientific) e precipitadas com acetona 100% gelada (4 °C) na proporção 4:1 (acetona: extrato proteico), mantendo-as por 1 h no congelador (-20 °C), sendo que a cada 15 min. as amostras eram homogeneizadas gentilmente. Ao final, os tubos foram centrifugados à 20.000 ×g, 4 °C durante 10 min., o sobrenadante descartado, deixados para secar durante 20 min., solubilizados em 20 µL tampão de corrida 1× (Tris-HCl 41,5 mM pH 8; Glicina 320 mM; SDS 0,16 %) e

submetidos à eletroforese de poliacrilamida 12,5%. O gel foi lavado 2 vezes com água destilada, fixado com solução metanol 50% + ácido acético 10% por 10 min., corado com *Blue silver* (CANDIANO et al., 2004) e descorado com água destilada.

5.13 Espectrometria de massa MALDI-TOF

As espidroínas recombinantes MaSp2 16× e MaSp2 32× foram submetidas à técnica de *Matrix-assisted laser desorption ionization time of flight mass spectrometry* (MALDI TOF MS) para confirmar as suas respectivas massas aproximadas. Para isso, foram realizadas duas metodologias: 1) O volume de um microlitro da proteína diluída em água, adquirida após o processo de diálise foi utilizada; 2) Uma alíquota da proteína liofilizada foi transferida para um tubo de 1,5 mL com ajuda de uma pinça com ponta fina (sem estática). A matriz utilizada foi de ácido sinapínico (cat n° S813 – Sigma) (HUMENIK et al., 2014) sendo diluída em acetonitrila 0,5 mg e TFA 0,05 %. A espidroína MaSp2 16× foi diluída em 5 µL de matriz e a espidroína MaSp2 32× em 10 µL de matriz. As amostras foram homogeneizadas em vórtex por 10 s e aquecidas à 80 °C por 5 min. Foi realizada uma diluição seriada do volume de 1 microlitro da amostra em três gotas de três microlitros, onde o volume inicial foi homogeneizado cinco vezes, para depois ser diluído à gota seguinte. O volume de um microlitro referente à cada diluição foi adicionado na placa do respectivo espectrômetro de massa utilizado. A placa foi colocada em fluxo vertical para secar e a análise foi realizada em equipamento Microflex (Bruker) em modo linear positivo, com extensão de massas de 26 kDa à 242 kDa, com a intensidade do laser à 90%.

5.14 Síntese das fibras

Posteriormente, as fibras foram sintetizadas artificialmente utilizando uma bomba de seringa *Pico Plus Harvard Apparatus Holliston* (Massachusetts, EUA) e banho de coagulação. A proteína MaSp2 (16x) foi diluída na concentração 20% (m/v). em de 1,1,1,3,3,3-Hexafluoro-2- propanol 95% (HFIP)(52517, Fluka, Suíça) e ácido fórmico 5% (288, Vetec, Rio de Janeiro, RJ, Brasil) A solução proteica foi aquecida à 80 °C por 10 min e transferida para uma seringa de 250 µL Hamilton Gastight® (1725, Hamilton Company, Reno, Nevada, EUA). Na ponta da agulha foram acoplados conectores permitindo o perfeito encaixe do tubo *PEEK* de 10 cm de extensão e 125 µm de diâmetro interno (AD-29003040, Zeus) e a extrusão foi realizada em banho de isopropanol 100 % numa velocidade de 415 µL/h. As fibras foram coletadas, secas, cortadas em pedaços de 3 cm e mantidas em dissecador com sílica gel azul (155, Vetec Sigma-Aldrich, Rio de Janeiro, RJ, Brasília) até serem analisadas ou processadas. Essas fibras foram separadas em dois grupos n=10, totalizando 20 fragmentos de fibras obtidas de duas extrusões. O primeiro grupo foi analisado sem tratamento pós-extrusão, sendo denominado grupo não tratado, e o segundo grupo foi imerso em isopropanol 80% por um minuto, estiradas em 1× o seu tamanho (6 cm ao final) e secas, sendo denominado grupo tratado. As fibras foram mantidas em uma placa de petri de plástico (90×15 mm) (J.Prolab, São José dos Pinhais, PR, Brasil) em dissecador contendo sílica gel (Vetec).

5.15 Análises em nanoescala por Microscopia de força atômica (MFA)

As análises topográficas e de espectroscopia de força das fibras sintéticas foram realizadas por microscopia de força atômica (MFA), em equipamento SPM-9600 (Shimadzu, Japão) em condição ambiente (média de 19.59 °C e umidade relativa do ar 21,35%). Para esse procedimento, um suporte de metal foi recoberto com fita adesiva dupla face e, em seguida as fibras de um cm de comprimento (n=10) foram depositados paralelamente entre si, para cada grupo, com ajuda de uma pinça anti-estática.

As imagens para análises de superfície foram adquiridas conforme já descrito (MENEZES et al., 2013; OLIVEIRA 2012) em modo contato intermitente – fase (viscoelasticidade), utilizando um *cantilever* de Si_3N_4 de 125 μm de comprimento, com constante da mola $\sim 42 \text{ N.m}^{-1}$, frequência de ressonância $\sim 250 \text{ kHz}$, e utilizando uma ponteira cônica com raio de curvatura de 10 nm, com resolução de aquisição 512×512 linhas e frequência de 1 Hz. A configuração do processamento ocorreu com nivelamento automático global e a remoção de linhas com ruídos. A área padrão de varredura foi de 25 μm^2 , sendo escolhidas três diferentes regiões para cada uma das 10 fibras sintéticas, totalizando 30 imagens. Os parâmetros de altura analisados foram: (1) a rugosidade média aritmética (R_a), onde os valores foram medidos a partir de uma linha média ao longo do comprimento de amostragem; (2) a rugosidade quadrática média (R_q), desvio padrão, onde os valores também são absolutos; (3) a altura máxima (R_z), onde é calculada a distância dos picos e dos vales longitudinalmente da curva de rugosidade; $R_z=R_p+R_v$; (4) a rugosidade média de 10 pontos (R_{zj10}), onde se calcula a média dos 5 vales mais profundos e dos 5 picos mais altos; (5) a altura média (R_p), onde é calculada a média da seção da curva e a distância do pico mais alto; (6) a profundidade média (R_v), calculada a partir da média da seção da curva e o vale mais profundo (Manual de instruções, SPM-9600, Shimadzu). Esses parâmetros foram selecionados com o objetivo de descrever e entender as superfícies topográficas das fibras analisadas após a sua síntese e tratamento pós-síntese, conforme previamente descrito. As imagens adquiridas no retraço das modalidades altura e fase foram analisadas e processadas em modo *off-line* (SMP- Shimadzu), corrigindo-se as médias do eixo das abscissas e posteriormente o plano de correção automática. As imagens para altura foram transformadas em 3D no modo *3D Height*, enquanto as de viscoelasticidade estão em 2D.

As avaliações por espectroscopia de força das fibras sintéticas foram adquiridas pelo mesmo equipamento e condição ambiente ($\sim 19,43 \text{ }^\circ\text{C}$ e umidade relativa do ar 22,31%). O mesmo modelo de *cantilever* foi utilizado com uma ponteira piramidal e constante de mola de $0,57 \text{ nN.nm}^{-1}$. O procedimento foi executado no modo contato com uma *loading rate* (razão de aplicação) constante de 1 nm.s^{-1} e com intervalo de ajuste de força em modo automático. Os valores de curvas força-distância adquiridos equivalem à interação da amostra com a ponteira em relação à distância entre a ponteira e a amostra (MENEZES et al., 2013). Essas medidas proporcionam valores como o módulo de elasticidade (1), a força de atuação na amostra, como as atrativas (2) e as adesivas (3), a força máxima aplicada (4), e as medidas de separação da ponteira e da amostra, como profundidade de penetração e deformação (5) a

fim de caracterizar as propriedades mecânicas. Ao menos cinco diferentes locais de cada fibra foram analisados (n=10) de cada grupo, e os resultados foram analisados em um *software* comercial (SPIP™ 5.1.11, Image Metrology, Kongens Lyngby, Denmark). Os valores do módulo de elasticidade e de curvas força-distância foram calculados por meio do *software* SPIP em modo analítico *Sneddon*, que consiste na indentação de cone (Sneddon, 1965).

5.16 Microscopia Eletrônica de Varredura (MEV)

A morfologia das proteínas liofilizadas MaSp2 16× (54 kDa) e MaSp2 32× (105 kDa), e a topografia das fibras sintéticas, as mesmas utilizadas em MFA, foram avaliadas por microscopia eletrônica de varredura (MEV). Para isso, as amostras foram depositadas sobre um suporte com fita de carbono, e metalizadas com uma camada ultrafina de paládio no vácuo (LEICA EM SCD500), sendo observadas em equipamento JEOL JSM 7001F à voltagem de 15 kV e *software* PC-SEM (ver 2.0.0.8), equipamentos localizados no laboratório de Microscopia Eletrônica e Virologia, sob a supervisão da professora Sônia Nair Bão, Instituto de Biologia, *Campus Darcy Ribeiro*, Universidade de Brasília. As imagens das proteínas liofilizadas foram adquiridas com ampliação de 190×, 430× e 3.000×. Já as fibras foram fotografadas com ampliação de 1.400× (Menezes et al 2013), para avaliação superficial, e adquirindo-se a medida do diâmetro em três diferentes áreas (n=10) tanto para o grupo controle e quanto para o grupo tratado com isopropanol.

5.17 Testes mecânicos das fibras

Os testes mecânicos das fibras sintéticas foram realizados no laboratório do professor Randy Lewis, Centro de Bioprodutos Sintéticos, Universidade Estadual de Utah, localizado em 1780N, *Research Park Way, Suite, 108 North Logan, UT 84381, USA*. As fibras possuíam cinco centímetros de comprimento e foram separadas em cinco grupos: (1) não processadas, (2) hidratadas, (3) hidratadas e com óleo mineral, (4) hidratadas e estiradas, (5) hidratadas,

estiradas e com óleo mineral, e foram enviadas em placas de petri forradas internamente com papel cartão preto.

As fibras foram cortadas em pedaços de 1,91 cm (3/4 inches) e coladas em uma moldura retangular de papel cartão. Cada fibra foi fotografada em três áreas diferentes e diâmetro foi medido três vezes em cada área, totalizando 9 medidas para cada fibra. Para isso, utilizou-se um microscópio Nikon Eclipse E200 e o programa ImageJ 1.42q (National Institute of Health, USA). Após calcular o diâmetro médio, as fibras foram avaliadas mecanicamente em condições ambiente 21-22 °C, com umidade relativa do ar entre 18-20% no equipamento MTS Synergie 100 (MTS Corporation, Eden Prairie, MN), utilizando uma carga de 10 g (transducer Techniques, Temecula, CA). Os parâmetros para energia de rompimento, módulo de elasticidade, extensão e estresse máximo foram calculados a partir dos gráficos de força por tensão na velocidade de 250 mm/min. Foram analisadas (1) 7 fibras após formação, (2) 8 fibras hidratadas, (3) 10 fibras hidratadas e com camada de óleo mineral, (4) 10 fibras hidratadas e estiradas, e (5) 10 fibras hidratadas, eticadas e contendo uma camada de óleo mineral.

5.18 Espectroscopia por infravermelho (ATR-FTIR)

As análises das proteínas liofilizadas MaSp2 16× e MaSp2 32× e dos dois grupos de fibras sintéticas, após a formação e as hidratadas e estiradas, foram realizadas em colaboração com o pesquisador Luiz Alberto Colnago no Laboratório de Caracterização Físico-Química na Embrapa Instrumentação localizada na Rua XV de Novembro, 1452 Centro, 13560970 - São Carlos, SP, Brasil. As amostras foram analisadas por espectrômetro de infravermelho médio com transformada de *Fourier* (FTIR) modelo Vertex 70 (Bruker). Utilizou-se um acessório para medida de Reflectância Total Atenuada (ATR) com cristal de diamante, onde foram adquiridos 32 digitalizações (scans) com a resolução de 4 cm⁻¹.

RESULTADOS E DISCUSSÃO

6.1 Plasmídeos para engenharia metabólica de bactéria

Após a amplificação por PCR para o primeiro fragmento do gene bacteriano *glyVXY*, o produto amplificado foi submetido a eletroforese e fotodocumentado, conforme observados na figura 14A. Logo após, esse fragmento foi inserido no plasmídeo pGEMT- Easy (Promega), e, os plasmídeos obtidos foram digeridos com a enzima de restrição *EcoRI* para confirmar a inserção (Figura 14.B).

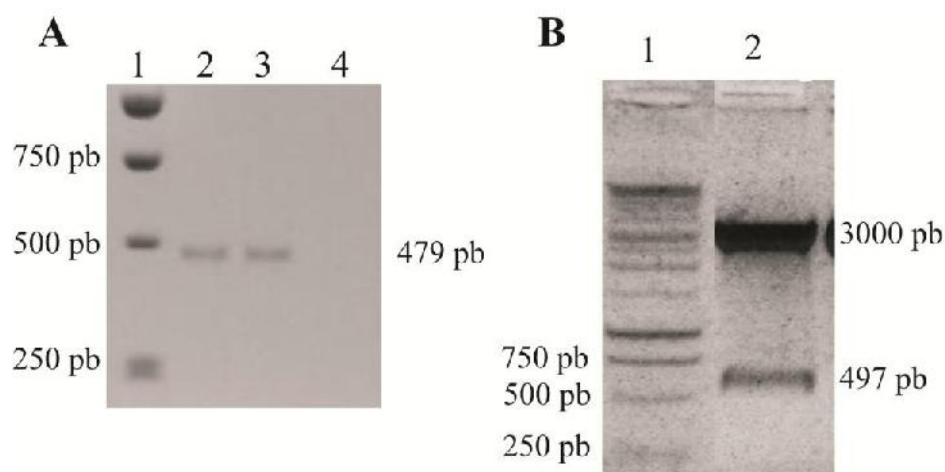


Figura 14. Análise eletroforética em gel de agarose corado com brometo de etídio. A) Fragmento amplificado por PCR referente ao gene *glyVXY* (479 pb). B) Produto de digestão com enzima *EcoRI* do plasmídeo PGEMTEasy (3.000 pb) contendo o gene *glyVXY* (497 pb). Marcador: 1 Kb Ladder da Promega.

Após a confirmação do plasmídeo pGEMT-Easy com o gene *glyVXY*, esse foi introduzido no plasmídeo pACYC184 nos sítios para as enzimas *BamHI* e *EcoRV*. O fragmento e o plasmídeo digeridos possuem 463 pb e 4.055 pb, respectivamente, e juntos originaram o plasmídeo pTet*glyVXY* (figura 15). Todos os plasmídeos derivados de pACYC184 apresentam baixo número de cópias (~15 por célula), sendo necessário a maxi preparação de plasmídeo com ação do antibiótico espectinomicina 35 µg/mL durante 12 h após a OD₆₀₀ 0,6 para amplificar esse material genético em quantidade suficiente (recomendações do fabricante, pACYC184, NEB). A resistência desse plasmídeo ao antibiótico tetraciclina deixou de existir devido as construções realizadas.

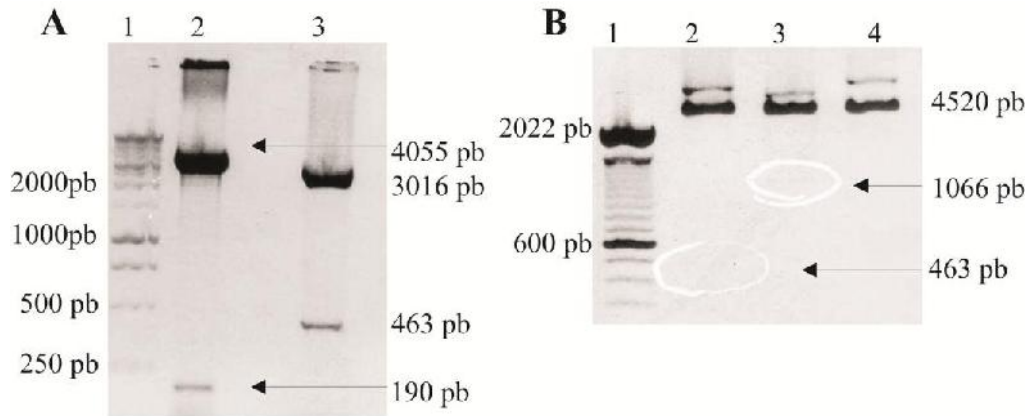


Figura 15. Análise eletroforética em gel de agarose corado com brometo de etídio. A) Produto das digestões das enzimas *Bam*HI e *Eco*RV. 1- 1Kb Ladder Promega, 2- pACYC184, 3- PGEMT-Easy+ *glyVXY*. B) Confirmação do plasmídeo pTet*glyVXY* digerido com *Bam*HI e *Eco*RV (463 pb), *Hinc*II (1.066 pb) e *Nhe*I (4.520 pb).

Após a amplificação por PCR, utilizando o plasmídeo pTet*glyVXY* (4.520 pb) como molde, o gene *glyVXY* (674 pb) foi submetido à eletroforese (Figura 16.A), purificado, e inserido no plasmídeo PGEMT-Easy (Figura 16.B). O fragmento de 674 pb foi inserido no plasmídeo pTet*glyVXY* nos sítios de restrições das enzimas *Sal*I e *Sph*I (Figura 16.C) formando o plasmídeo pTet*gly2* (5.091 pb) (Figura 16.D). O fragmento *glyVXY* e o plasmídeo pTet*glyVXY* digeridos contém 674 pb e 4.431 pb, respectivamente.

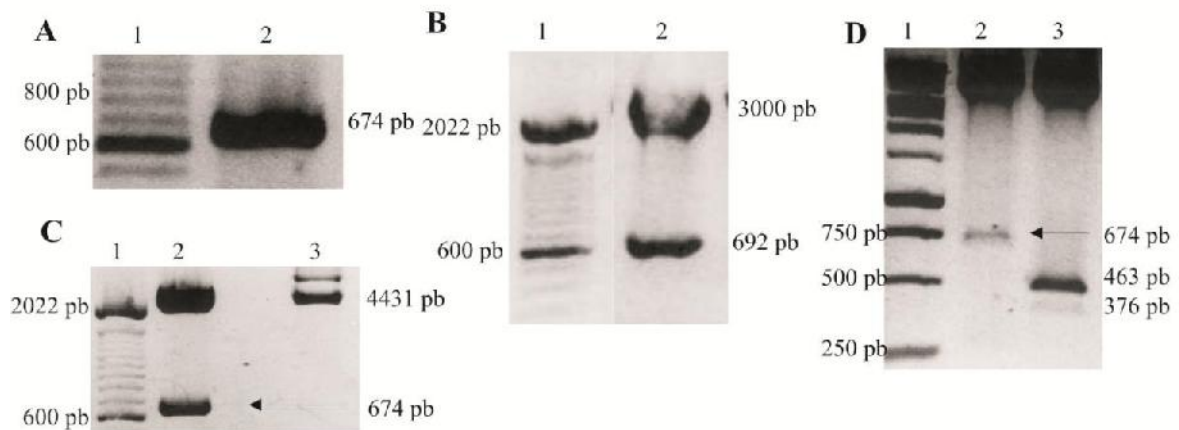


Figura 16. Análise eletroforética em gel de agarose corado com brometo de etídio. A) 1- 100 pb Invitrogen. 2- Gene *glyVXY* amplificado por PCR. B) 1- 100 pb Invitrogen, 2- PGEMT-Easy+ *glyVXY* digerido com *Eco*RI, C) 1- 100 pb Invitrogen, 2 e 3- PGEMT-Easy+ *glyVXY* e pTet*glyVXY* digeridos com as enzimas *Sal*I e *Sph*I. D) 1- 1Kb Ladder Promega, 2 e 3- plasmídeo pTet*gly2* digerido com *Sal*I e *Sph*I; *Bam*HI e *Eco*RV.

Após a amplificação por PCR, o gene *glyA* (1.734 pb) foi submetido a eletroforese e fotodocumentado, conforme observado na figura 17.A. Logo após, o fragmento foi inserido no plasmídeo pGEMT-Easy (Promega), sendo esse novo plasmídeo confirmado por meio de digestão da enzima *EcoRI* (3.000 pb e 1.752 pb)(Figura 17.B).

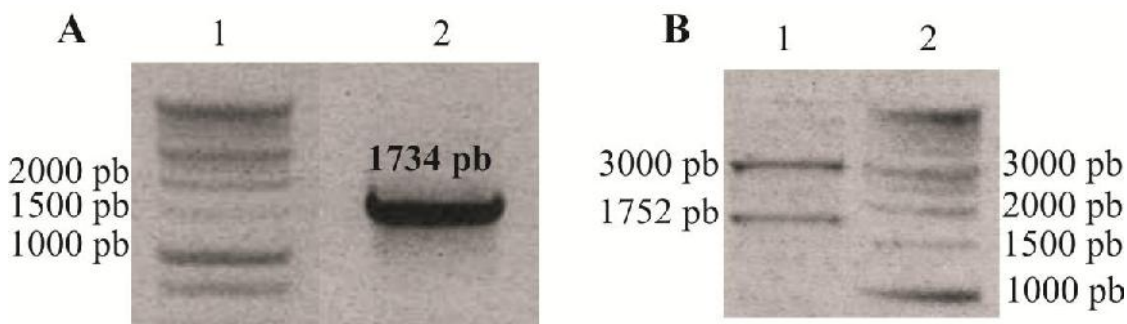


Figura 17. Análise eletroforética em gel de agarose corado com brometo de etídio. A) 1- 1Kb Ladder Promega. 2-Produto da PCR referente ao gene *glyA*. B)1-PGEMT-Easy (3.000 pb)+ *glyA* (1.752 pb)digerido com *EcoRI*, 2- 1Kb Ladder Promega.

O plasmídeo PGEMT-Easy+*glyA* (4.752 pb) foi utilizado como molde para a amplificação do gene *glyA* (1.734 pb) por PCR utilizando a enzima Pfx (Invitrogen), sendo posteriormente digerido com a enzima *SalI* (Figura 18.A). Esse gene foi introduzido no plasmídeo pTetgly2 digerido com as enzimas *SalI* e *HincII* (4.025 pb)(Figura 18.B), para originar o plasmídeo pTetgly2-*glyA* (5.755 pb). Esse foi amplificado por maxi-preparação, e sua confirmação ocorreu por meio das digestões das enzimas *BamHI* e *EcoRV* (376 pb+463 pb+1.720 pb+3.196 pb), *SalI* e *SphI* (5.098 pb + 657 pb) e *HincII* (430 pb+633 pb+4.692 pb) (Figura 19.C). A forma supercontrátil dos plasmídeos dificultou a ação da endonuclease *SalI*, necessária para a inserção do gene *glyA*, mesmo seguindo as sugestões do fabricante. Para superar essa dificuldade, fez-se necessário a utilização de 10 U da enzima para um micrograma de DNA, perdurando a incubação por toda a noite. A digestão incompleta ou parcial de *SalI* permitia a ação da enzima *HincII*, com sítio de ação na mesma região, liberando um fragmento com o tamanho aproximado, mas com duas extremidades abruptas. Por não conter a extremidade coesiva da enzima *SalI*, o gene *glyA* não era inserido nos plasmídeos pTetgly2, afim de formar o plasmídeo pTetgly2-*glyA*.

A confirmação do tamanho dos plasmídeos pACYC184, pTetglyVXY, pTetgly2 ocorreu por enzimas de digestão (Figura 19.A): *BamHI* e *EcoRI* (Fig. 19.A.2-4); *SalI*, *HincII* (Fig.19.A.5-7), *SalI* e *BamHI* (Fig.19.A.10-13). Devido à dificuldade encontrada na digestão dos plasmídeo pela enzima *SalI*, a construção do plasmídeo pTetglyVXY-*glyA* ocorreu a partir do plasmídeo pTetgly2-*glyA*. Um fragmento de DNA (839 pb) contendo uma das cópias do

gene *glyVXY* foi retirada por digestão da a enzima de restrição *HindIII* (Thermo Scientific). O fragmento com 4.916 pb foi purificado e a religação originou o plasmídeo pTet*glyVXY-glyA*, sendo amplificado por maxi-preparação. Este obteve seu tamanho confirmado por digestões das enzimas *HincII* (430 pb+633 pb+3.853 pb), *EcoRV* (1.720 pb+3.196 pb), *HindII* (4.916 pb), *SalI* (4.916 pb), *BamHI* (não digere) (Figura 19.B).

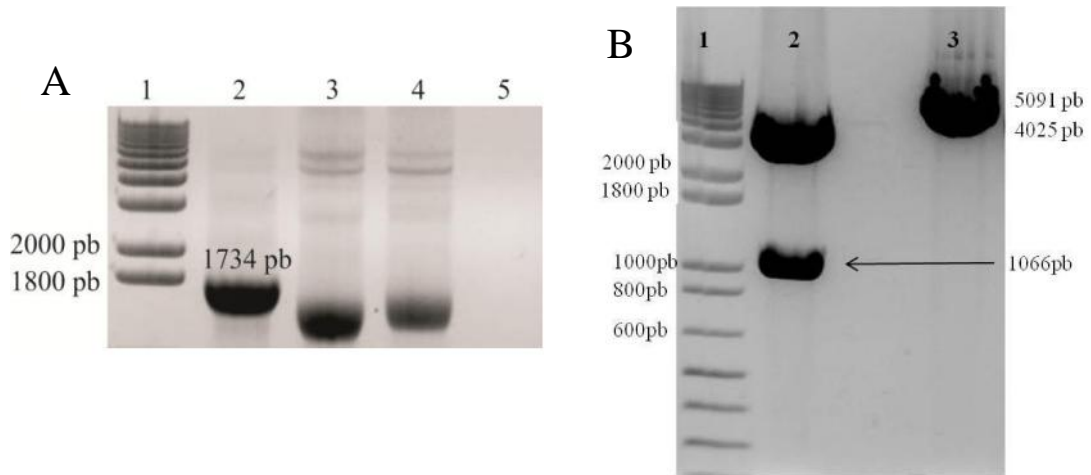


Figura 18. A) Análise eletroforética em gel de agarose 1 % corado com brometo de etídio da amplificação por PCR do gene *glyA*. 1-1 Kb *Ladder Plus* Invitrogen, 2- 4 Gene *glyA* amplificado com MgSO₄ 1 mM; 1,5; 2,0 mM, 5-branco. B) 1-1 Kb *Ladder Plus* Invitrogen. 2- Plasmídeo pTet*gly2* digeridos com as enzimas *SalI* e *HincII* (4.025 pb), e o fragmento de 1.066 pb que foi descartado. 3-Plasmídeo pTet*gly2* digerido com as enzimas *SalI* (5.091 pb).

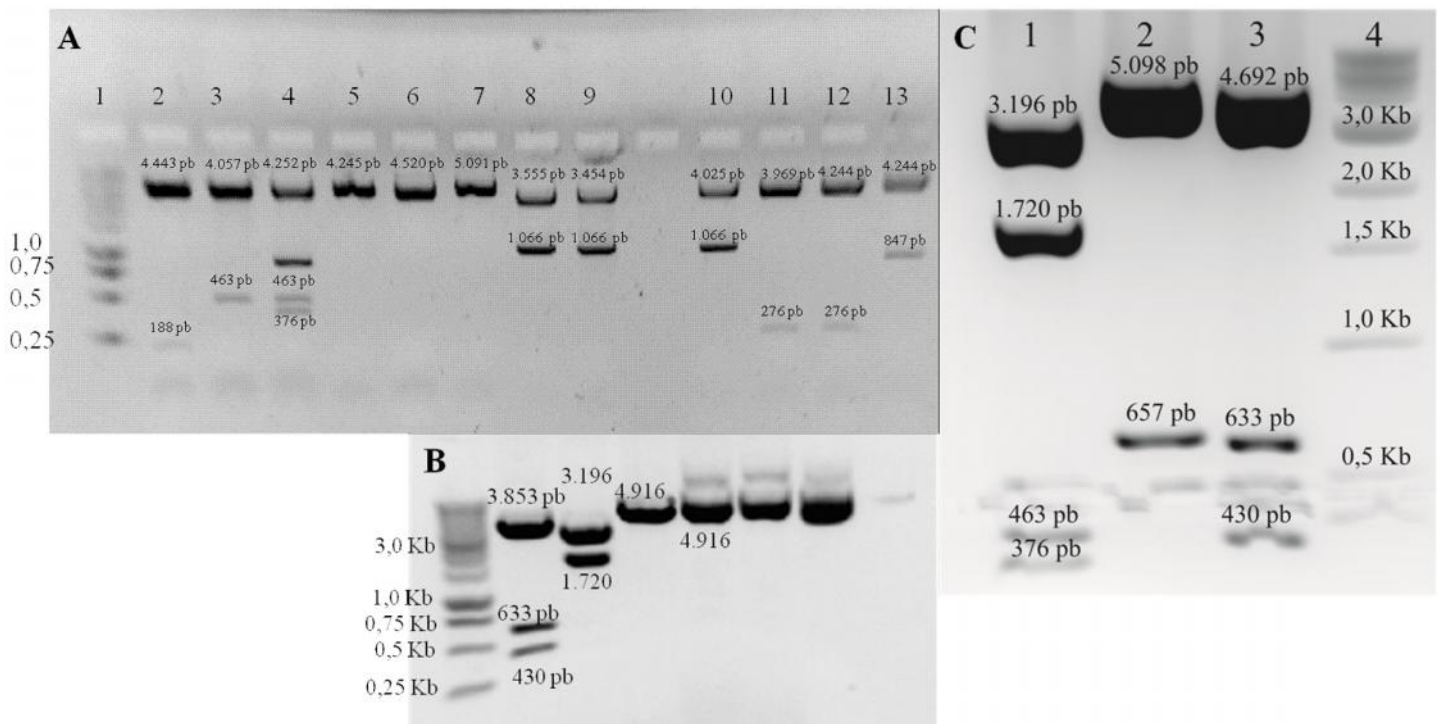


Figura 19. Análise eletroforética em gel de agarose 1% corado com brometo de etídio. A) Plasmídeos na sequência: pACYC184, pTet*glyVXY* e pTet*gly2* digeridos com: 2-4: *BamHI* e *EcoRV*; 5-7: *SalI*; 8-10: *HincII*; 11-13: *SalI* e *BamHI*. B) Plasmídeo pTet*glyVXY-glyA* digerido com *HincII*, *EcoRV*, *HindII*, *SalI*, *BamHI* e padrão não digerido. C) Plasmídeo pTet*gly2-glyA* digerido com 1: *BamHI* e *EcoRV*, 2: *SalI* e *SphI*; 3- *HincII*; 4: 1 Kb Ladder NEB.

Todas as linhagens bacterianas construídas foram confirmadas quanto à presença dos plasmídeos derivados de pACYC184. Para isso, o material genético das mini preparações de plasmídeos foram submetidas a ampliações por PCR dos genes *glyVXY* (674 pb) e *glyA* (1.734 pb) com os iniciadores e condições descritos. É importante observar que esses genes são encontrados no genoma da bactéria BL21(DE03) e podem ser amplificados durante a reação (figura 20.B). No entanto, foi possível verificar a diferença entre a amplificação do gene alvo a partir do material genético do cromossomo da bactéria (figura 20.B colunas 1 e 2) em relação ao plasmídeo (Figura 20.B, colunas 5 e 6) por meio do aumento da intensidade da banda (Figura 20).

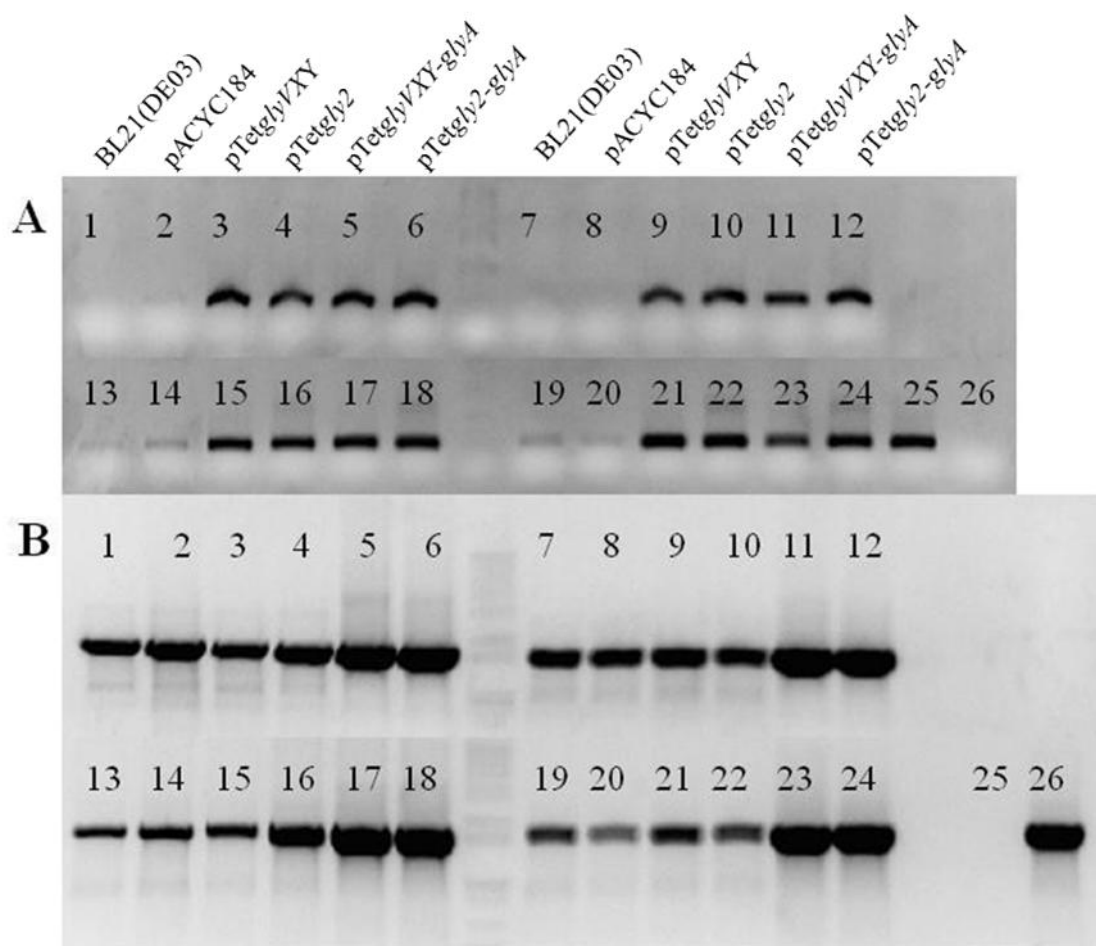


Figura 20. Análise eletroforética em gel de agarose 1% corado com brometo de etídio dos produtos das reações de PCR para as linhagens bacterianas. Da esquerda para a direita: BL21(DE03); BL21(DE03)pACYC184; BL21(DE03)pTetglyVXY; BL21(DE03)pTetgly2; BL21(DE03)pTetglyVXY-glyA; BL21(DE03)pTetgly2-glyA. A) Amplificação do gene *glyVXY*: 1-6 linhagens na ausência do plasmídeo pET19b; 7-12 linhagens na presença do plasmídeo pET19b sem inserto; 13-18: linhagens na presença do plasmídeo pET19b MaSp2 16×; 19-24: linhagens na presença do plasmídeo pET19bMaSp 32×. 25: controle positivo; 26-branco. B) Amplificação do gene *glyA*: 1-6 sem pET19b; 7-12 com pET19b; 13-18: com pET19b MaSp2 16×; 19-24: com pET19bMaSp 32×; 25:branco, 26:controle positivo.

6.2 Sequenciamento dos genes utilizados para os motivos de *P. bistrata*

O resultado do alinhamento entre o motivo inicial MaSp2 (linha inferior) e o fragmento presente no plasmídeo pBSSK+ (linha superior) encontra-se na da figura 21. Todas as construções para essa espidroína nos plasmídeos de clonagem e de expressão em bactéria apresentaram as sequências e o códon de iniciação sem alterações.

1	GCATGCTGGTTCACTCGAGCATATGCCCGGGGGCGGCTATGGTCCGGGCGGCGCGGGCCA
1GGCGGCTATGGTCCGGGCGGCGCGGGCCA
61	GCAGGGCCCGGCGGCGGGCCAGCAGGGTCCGGGCAGCCAGGGCTCGTATGGTCCGGGCGC
30	GCAGGGCCCGGCGGCGGGCCAGCAGGGTCCGGGCAGCCAGGGCTCGTATGGTCCGGGCGC
121	GGCGGCGGCGGCGGCGGGCGGCGGCGG <u>TCCGGG</u> GGCGGCTATGGTCCGGGCGGCGCGGGCCAGCA
90	GGCGGCGGCGGCGGCGGGCGGCGGCGG <u>TCCGGG</u> GGCGGCTATGGTCCGGGCGGCGCGGGCCAGCA
181	GGGCCCGGCGGCGGGCCAGCAGGGTCCGGGCAGCCAGGGCTCGTATGGTCCGGGCGGGC
150	GGGCCCGGCGGCGGGCCAGCAGGGTCCGGGCAGCCAGGGCTCGTATGGTCCGGGCGGGC
241	GGCGGCGGCGGCGGGCGGCGGCGGCGGATAAGGATCCACTAGTTCTAGAGCGGCCGCCACCGC
210	GGCGGCGGCGGCGGGCGGCG

Figura 21. Alinhamento da sequência MaSp2 2× de *P. bistrata*. A sequência superior equivale às construções *in silico*, enquanto a inferior refere-se ao resultado do sequenciamento. A sequência sublinhada refere-se a ligação dos produtos das digestões das enzimas *BspEI* e *XmaI*. Identidade 100%.

O resultado do alinhamento entre o motivo inicial MaSp1(2×)+2(1×) de *Parawixia bistrata* (linha inferior) e o fragmento presente no plasmídeo pBSSK+ (linha superior) encontram-se na figura 22.


```

1   TCGTACTGGGTCACCTCGAGCATATGCCCGGGGTGGTCTGGGTGGTCAGGGTGGTCTGG
      |||
1   .....GGTGGTCTGGGTGGTCAGGGTGGTCTGG
61  GTGGTCTGGGTTCCCAGGGTGCTGGTCTGGGTGGTTACGGTCAGGGTGGTCTGGTCAGG
      |||
29  GTGGTCTGGGTTCCCAGGGTGCTGGTCTGGGTGGTTACGGTCAGGGTGGTCTGGTCAGG
      |||
121 GTGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGTGGTCTGGGTGGTCAGGGTGGTCTGGGTG
      |||
89  GTGGTGCTGCTGCTGCTGCTGCTGCTGCTGGTGGTCTGGGTGGTCAGGGTGGTCTGGGTG
      |||
181 GTCTGGGTTCCCAGGGTGCTGGTCTGGGTGGTTACGGTCAGGGTGGTCTGGTCAGGGTG
      |||
149 GTCTGGGTTCCCAGGGTGCTGGTCTGGGTGGTTACGGTCAGGGTGGTCTGGTCAGGGTG
      |||
241 GTGCTGCTGCTGCTGCTGCTGCTGCTGGCGGCTATGGTCCGGGCGGCAGCAGG
      |||
209 GTGCTGCTGCTGCTGCTGCTGCTGCTGGCGGCTATGGTCCGGGCGGCAGCAGG
      |||
301 GCCCGGCGGGCCAGCAGGGTCCGGGCGCCAGGGCTCGTATGGTCCGGGCGGCAGG
      |||
269 GCCCGGCGGGCCAGCAGGGTCCGGGCGCCAGGGCTCGTATGGTCCGGGCGGCAGG
      |||
361 CGGCGGCGGGCGGGCTCCGGATAAGGATCCACTAGTTCTAGAGCGGCCACCAGG
      |||
329 CGGCGGCGGGCGGGCGG

```

Figura 22. Alinhamento genético do plasmídeo contendo a MaSp1+2 de *P. bistrata* (linha inferior) com a sequência *in silico* enviada para síntese (linha superior). Identidade 100%.

6.3 Plasmídeos de produção das espidroínas em *E. coli*

Os fragmentos referentes ao gene de MaSp1 das digestões realizadas com as enzimas *ScaI* e *XmaI* com 16× motivos (3.688 pb), 32× motivos (5.560 pb) e as digestões com *ScaI/BspE1* com 16× motivos (2.984 pb), 32× motivos (4.856 pb) podem ser observada na figura 23 (A e B). O plasmídeo pET19b contendo 32× motivos foi digerido com as enzimas *ApaI* e *BamHI* (4.871 pb), e com *NdeI*, *BamHI* e *ScaI* (3.765 pb) sendo os seus fragmentos observados na figura 23 (C e D).

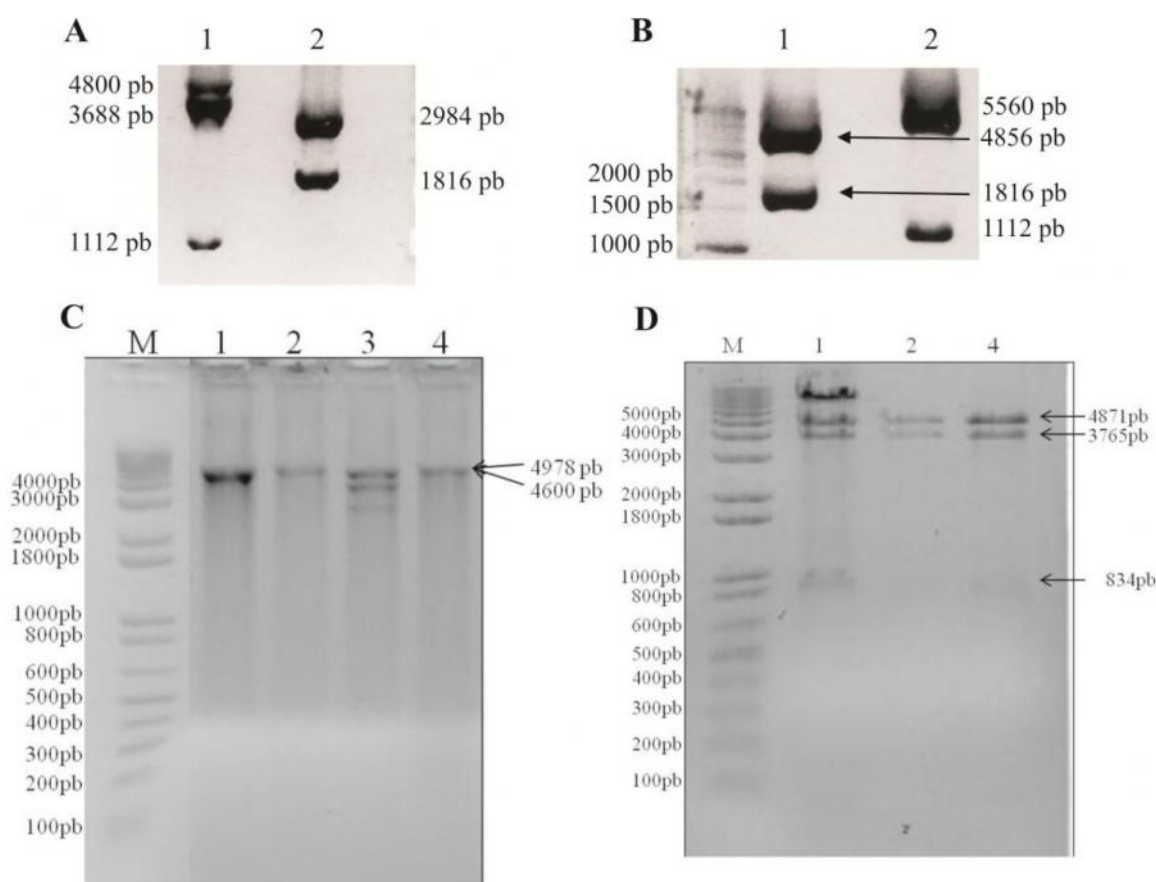


Figura 23. Análise eletroforética das construções de MaSp1. A) Plasmídeo pBSSK+ contendo 16× o motivo. 1- Digestão com as enzimas *XmaI* e *ScaI*. 2- Digestão com as enzimas *ScaI* e *BspEI*. B) Plasmídeo pBSSK+ contendo 32× o motivo. 1- Digestão com as enzimas *ScaI* e *BspEI*. 2- Digestão com as enzimas *XmaI* e *ScaI*. C) Plasmídeo pET19b contendo 32× o motivo digerido com *ApaI* e *BamHI*; 1, 2 e 4 são amostras esperadas. D) Plasmídeo pET19b, contendo 32× o motivo, digerido com *ApaI*, *BamHI* e *ScaI*.

Apesar dos plasmídeos de clonagem (pBSSK+) e de expressão apresentarem tamanhos corretos, não houve a produção das espiroínas MaSp1 16× e da inédita MaSp1 32× nas linhagens de bactérias. As análises de sequenciamento confirmaram o material genético, não sendo encontrada uma justificativa para a não produção dessas espiroínas. Em trabalhos anteriores realizados no grupo do laboratório, a mesma espiroína MaSp1 16× foi produzida, caracterizada molecularmente e formou fibras sintéticas (Oliveria, 2008; Oliveira, 2012).

Os fragmentos referentes à digestão de sequências genéticas dos motivos da MaSp2 podem ser observados na figura 24. Digestão com as enzimas para *Bam*HI e *Xho*I, com quantidade de motivos 1× (114 pb)(24.A), 2× (261 pb)(24.B), 4× (501 pb)(24.C), 8× (981 pb)(24.D), 16× (1.941 pb)(24.E), 32× (3.861 pb)(24.F).

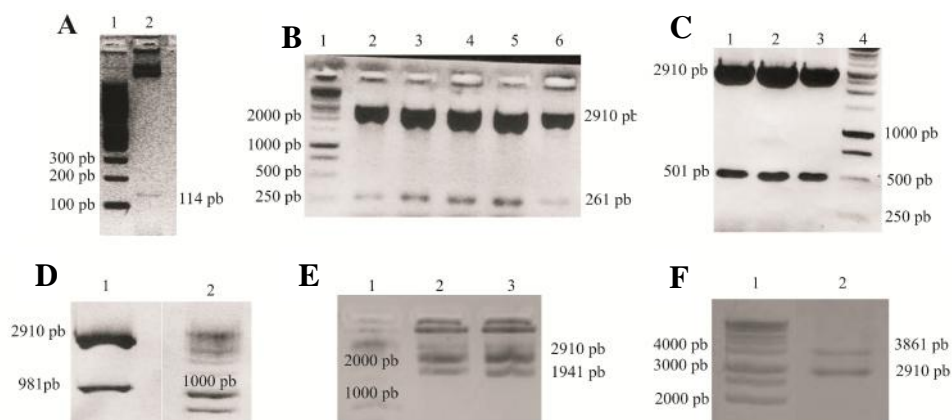


Figura 24. Análise eletroforética em gel de agarose corado com brometo de etídio referente às construções do módulo MaSp2 da aranha *Parawixia bistriata*. digeridos com *Bam*HI e *Xho*I. A-F: plasmídeos de clonagem contendo 1, 2, 4, 8, 16 e 32 motivo(s) de MaSp2.

Os resultados das duplas digestões com as enzimas *ScaI* e *XmaI* nos plasmídeos de clonagem contendo genes de espidroínas encontram-se na figura 25; com a quantidade de motivo(s) 1× (1.936 pb)(25.A), 2× (2.056 pb)(25.B), 4× (2.296 pb)(25.C), 8× (2.776 pb)(25.D), 16× (3.736 pb)(25.E), 32× (5.656 pb)(25.F). Nessa mesma figura, as digestões com *ScaI* e *BspEI* proporcionam motivo(s) 1× (1.235 pb)(25.A), 2× (1.355 pb)(25.B), 4× (1.595 pb)(25.C), 8× (2.776 pb)(25.D), 16× (3.035 pb)(25.E), 32× (4.955 pb)(25.F).

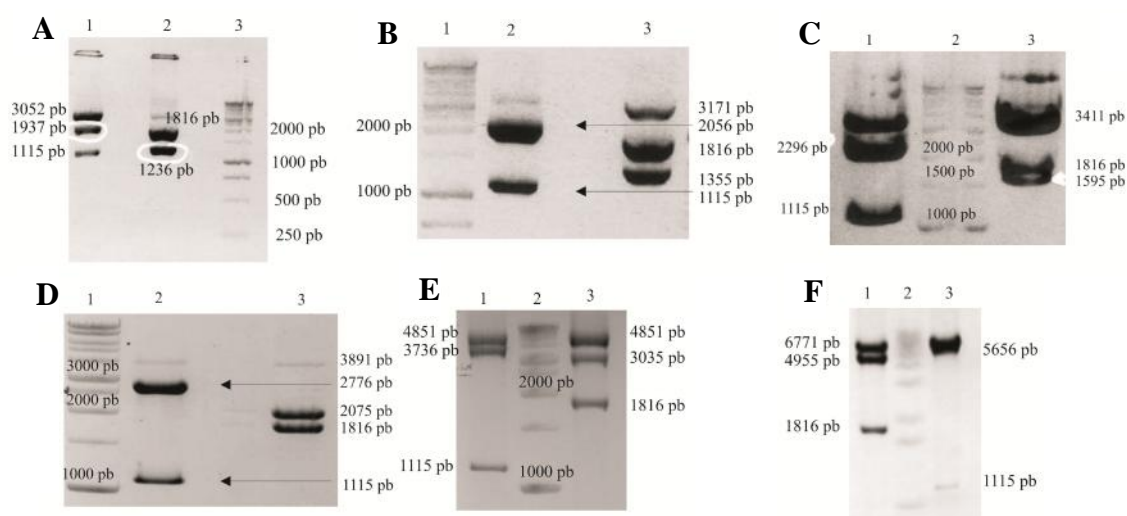


Figura 25. Análise eletroforética em gel de agarose corado com brometo de etídio referente às construções do módulo MaSp2 da aranha *Parawixia bistrata*. digeridos com *ScaI* e *XmaI* e também com *ScaI* e *BspEI*. A-F: plasmídeos de clonagem contendo 1, 2, 4, 8, 16 e 32 motivo(s) de MaSp2.

Os fragmentos de material genético das espidroínas MaSp 16× (1.940 pb) e MaSp e 32 × (3.765 pb) obtidos a partir da digestão das enzimas *Bam* HI e *NdeI* encontram-se na figura 26.

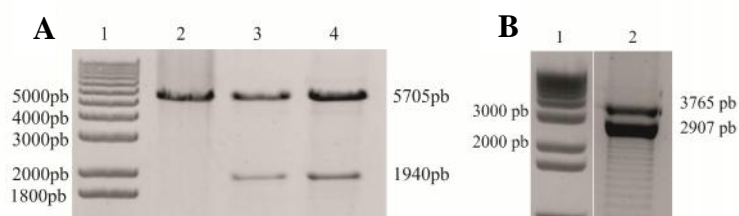


Figura 26. Análise eletroforética de plasmídeos contendo espidroínas MaSp2 de *Parawixia bistrata*. A) Plasmídeo pET19b MaSp2 16× digeridos com *Bam* HI e *NdeI*: 1- 1Kb Ladder (Invitrogen); 2- Plasmídeo sem inserto; 3-4 Plasmídeo contendo a sequência da espidroína (1.940 pb). B) Plasmídeo pBSSK+ MaSp 32× digeridos com *Bam* HI e *NdeI*: 1- 1Kb Ladder (Invitrogen); 2- Plasmídeo contendo a sequência da espidroína (3.765 pb).

Os plasmídeos construídos contendo a sequência genética da MaSp1+2, foram digeridos com as enzimas *Bam*HI e *Xho*I, contendo o(s) motivo(s) em: 1× (372 pb)(27.A), 2× (723 pb)(27.B), 4× (1.450 pb)(27.C), 8× (2.890 pb) (27.D).

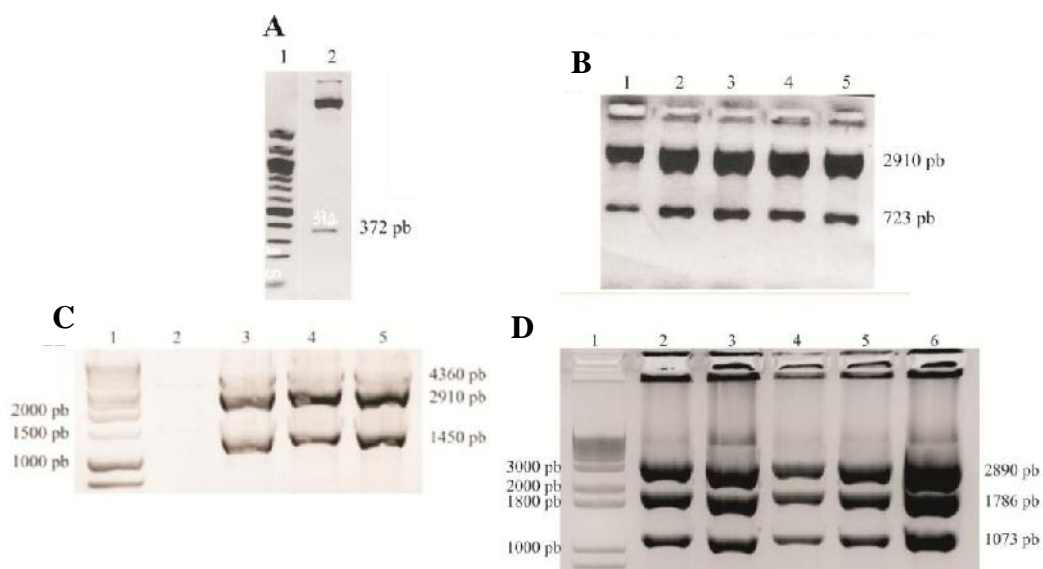


Figura 27. Análise eletroforética, em gel de agarose corado com brometo de etídio referente às construções de MaSp1+2. A) Plasmídeo pJ. B-C Plasmídeo pBSSK+ digeridos com as enzimas *Bam*HI e *Xho*I contendo o motivo 2× (723 pb), 4× (1.450 pb), respectivamente. D) Plasmídeo pBSSK+ digeridos com as enzimas *Bam*HI, *Xho*I e *Sca*I contendo o motivo 8× (2.890 pb).

As digestões realizadas com as enzimas *Sca*I e *Xma*I com a quantidade de módulos 1× (2.167 pb)(28.A), 2× (2.516 pb)(28.B), 4× (3.220 pb)(28.C), e as digestões com *Sca*I e *Bsp*E1 com a quantidade de módulos 1×(1.466 pb)(28.A), 2×(1.817 pb)(28.B), 4×(2.519 pb)(28.C), 8×(3.923 pb)(28.D).

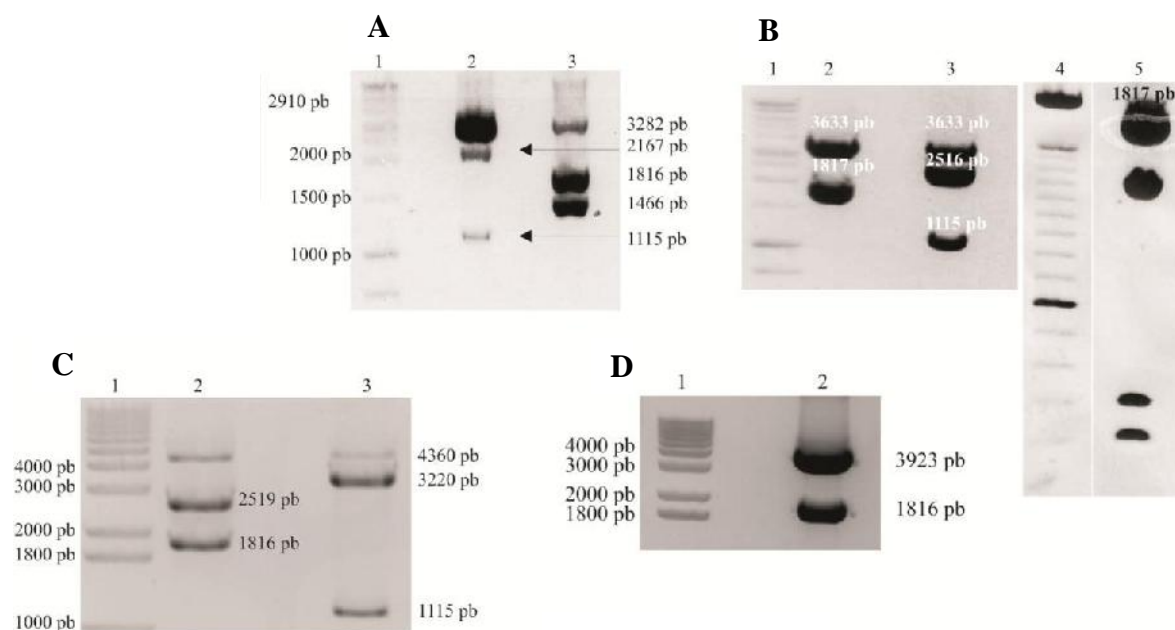


Figura 28. Análise eletroforética, em gel de agarose corado com brometo de etídio referente às construções da MaSp1+2. A) 1- 1Kb Ladder (Invitrogen); Plasmídeo pBSSK+ com 1× o motivo digerido com: 2- *XmaI* e *ScaI* (2.167pb) , 3- *ScaI* e *BpsEI* (1.466 pb). B) 1- 1Kb Ladder (Invitrogen); 4- 100 pb Ladder (Invitrogen). Digestão do plasmídeo pBSSK+ 2× o motivo com as enzimas: 2- *ScaI* e *BpsEI* (1.817 pb). 3- *XmaI* e *ScaI* (2.516 pb), 5- *ScaI*, *BpsEI*, *AseI* (2.890 pb). E) 1- 1Kb Ladder (Invitrogen). Digestão do plasmídeo pBSSK+ 4× o motivo com as enzimas: 2- *ScaI* e *BpsEI* (2.519 pb). 3- *XmaI* e *ScaI* (3.220 pb). F) 1- 1Kb Ladder (Invitrogen). Digestão do plasmídeo pBSSK+ 8× o motivo com as enzimas: 2- *BspEI* e *ScaI* (3.923 pb).

As construções de plasmídeos com uma grande quantidade de motivos como MaSp1 48× ou MaSp1 64× resultaram em fragmentos menores ao tamanho esperado (dados não apresentados). Sabe-se que quanto maior o tamanho do inserto e a alta repetição de motivos causam instabilidade nos plasmídeos durante a transformação e duplicação de material genético (TEULÉ et al., 2009), mesmo ao serem utilizadas as linhagens bacterianas com deleções ou mutações para genes de recombinação como XL10GOLD, XL1Blue e DH5 .

Na tentativa de se obter o plasmídeo contendo MaSp1 64× utilizou-se a linhagem STBL2 (Invitrogen), desenvolvida comercialmente para a clonagem de insertos instáveis, adquirindo com sucesso o tamanho de plasmídeo e insertos esperados (dados não apresentados). Xia e colaboradores (2010) construíram plasmídeos com 96 e até 128 motivos, não sendo informado detalhes da metodologia. Recentemente, outro grupo de pesquisa conseguiu produzir uma espidroína de 190 kDa em bactéria, que ao ser dimerizada, alcança a massa molecular de 378 kDa (LIN, et al., 2013). Porém, a produção dessas proteínas requer a utilização de linhagens de expressão com características semelhantes, BLR por exemplo, mas que não estavam disponibilizadas na época do experimento no laboratório. Xia e colaboradores (2010) construíram plasmídeos com MaSp1 96× e 128×, mas o trabalho não informa detalhes dos protocolos utilizados. Sabe-se que a porção C-terminal das espidroínas, que não foi incluída na sequência genética neste experimento, está relacionada à

formação e à estabilidade das proteínas de teia de aranha (Ittah, et al., 2006) (Hang, et al., 2010) (Hedhammar, et al., 2008). Além disso, a introdução de motivos diferentes na construção genética final parece estabilizar os plasmídeos e/ou as proteínas (Hauptmann, et al., 2013).

6.4 Plasmídeos de produção das espidroínas em soja

Os fragmentos de DNA de MaSp2 16× (1.920 pb) e MaSp2 32× (3.840 pb) foram retirados dos plasmídeos pET19bMaSp2 16× (4.851 pb) e pET19bMaSp2 32× (6.771 pb), desfosforilados e inseridos no plasmídeo pJ201:27396 GMPB (2.713 pb) nos sítios de restrição das enzimas *Xma*I e *Bsp*E1, a fim inserir uma sequência de histina na extremidade N-terminal e obter os sítios de restrições necessários para a construção do plasmídeo de transformação em soja (*Glycine max*) (Figura 29).

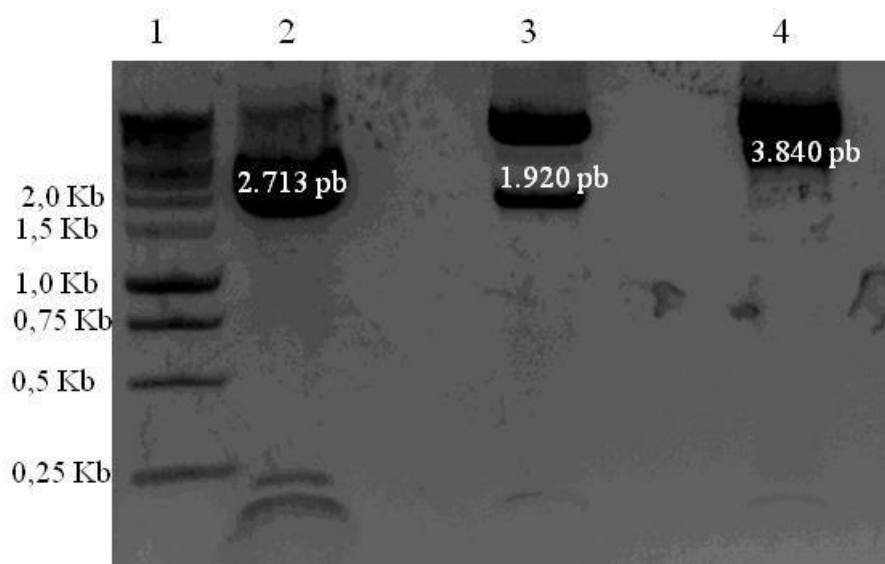


Figura 29. Análise eletroforética em gel de agarose corado com brometo de etídio. Plasmídeos digeridos com *Xma*I e *Bsp*E1. 1- 1Kb Plus Promega, 2- pJ; 3- pET19b MaSp2 16×; 4- pET19b MaSp2 32×.

Os plasmídeos pJ MaSp2 16× (4.613 pb) e pJ MaSp2 32× (6.533 pb) foram confirmados por tamanho com duas digestões enzimáticas, *Kpn*I e *Eco*RI, *Hind*III e *Eco*RI (Figura 30). Os fragmentos MaSp2 16× (1.959 pb) e MaSp2 32× (3.879 pb) digeridos com *Eco*RI e *Hind*III foram inseridos no plasmídeo p cong (~6.000 pb) entre esses sítios de restrições (Figura 30).

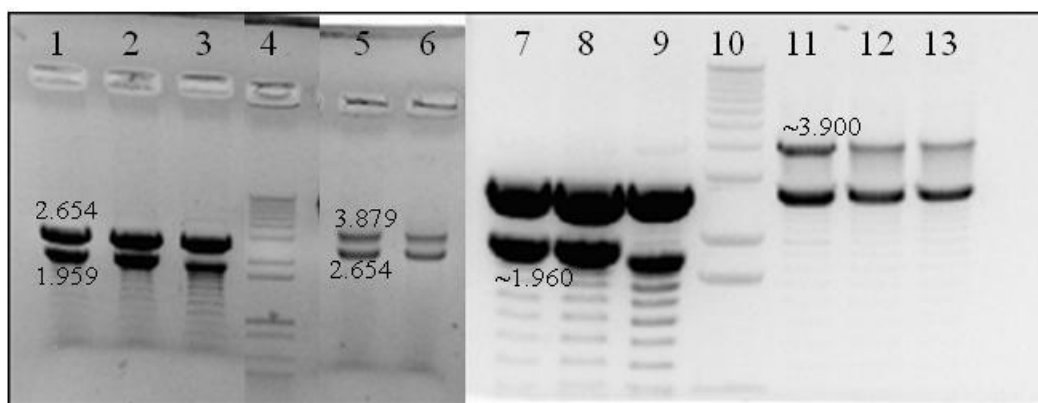


Figura 30. Análise eletroforética das digestões dos plasmídeos pJMaSp2 16× e pJMaSp2 32×. 1-3: MaSp2 16× *KpnI* e *EcoRI* (1.959 pb), 4 e 10- 1 Kb Ladder (Invitrogen), 5-6: MaSp2 32× *KpnI* e *EcoRI* (3.879 pb). 7-9: MaSp2 16× *HindIII* e *EcoRI* (1.960 pb), 11-13: MaSp2 32× *HindIII* e *EcoRI* (3.879 pb).

Os plasmídeos contendo as sequências de espidroínas MaSp2 16× e MaSp2 32× para a produção em soja foram construídos, mas não foram utilizados. A metodologia apresentada (RECH, VIANNA e ARAGÃO, 2008) foi bastante utilizada pelo grupo de pesquisa, alcançando resultados importantes na utilização da soja como biofábricas de moléculas como fator IX de coagulação sanguínea, hormônio do crescimento humano e cianovirina-N, que impede a entrada do vírus da AIDS nas células (CUNHA et al., 2011a; CUNHA, et al., 2011b, VIANNA, ARAGÃO e RECH, 2011; VIANNA, CUNHA e RECH, 2011).

Com o surgimento de novas tecnologias como TALEN e CRISPR é possível editar o genoma dos organismos (com adição, alteração ou destruição) em sequência específica (WOOD, et al., 2011) (CONG, et al., 2013) a fim de aumentar o rendimento de proteínas desejadas. Esses protocolos foram adotados pelo grupo nas práticas do laboratório, levando a substituições das outras metodologias (CHEN et al., 1986; CUNHA, et al., 2011a).

As espidroínas MaSp e Flag de outras espécies de aranha e de massas moleculares variando de 12,9 à 127 kDa já foram produzidas em plantas como tabaco (*Nicotiana tobaccum*, *Nicotiana benthamiana*), batata (*Solanum tuberosum*), *Arabidopsis thaliana* e soja (*Glycine max*) (Revisto por TOKAREVA et al, 2013), mas ainda encontram-se longe de um biorreator ideal. Barr e colaboradores (2004), utilizando uma metodologia semelhante à proposta neste trabalho, conseguiram produzir espidroínas de 64 kDa, acumulando de 0,54% à 1,64 % do total de proteína estocada em sementes de soja, mas distante dos 8,5 % acumulados em *Arabidopsis thaliana* (YANG, et al., 2005). Recentemente, a espidroína Flag de *Neplila clavipes* contendo múltiplos motivos repetitivos e inteínas nas extremidades (proteínas *splicing*) foram produzidas, com massas moleculares de 47 kDa à 250 kDa em *Nicotiana benthamiana* (HAUPTMANN, et al., 2013).

6.5 Curvas de crescimento

Com o intuito de verificar o tempo de crescimento das quatro novas linhagens de bactérias (1-BL21(DE03)pTetglyVXY, 2-BL21(DE03)pTetgly2, 3-BL21(DE03)pTetglyVXY-glyA, 4- BL21(DE03)pTetgly2-glyA) as curvas de crescimento foram estabelecidas em meio de cultura LB, à 37 °C durante um período de 10 horas na ausência ou na presença de um indutor (IPTG 0,4 mM). As linhagens BL21(DE03) e BL21(DE03)pACYC184 foram utilizadas como controles (Figura 31). Assim como a *E. coli* BL21(DE03), todas as linhagens avaliadas cresceram satisfatoriamente no meio de cultura utilizado, nas condições submetidas (ausência ou presença de IPTG), apresentando curvas e velocidades de crescimento semelhantes aos controles utilizados (Figura 31.A e B) e ao encontrado na literatura (PRESCOTT; HARLEY e KLEIN, 2005).

Após confirmar que a velocidade de crescimento dessas novas linhagens é semelhante à BL21(DE03), o mesmo procedimento de avaliação foi realizado na produção das espiroínas MaSp2 16× e MaSp2 32×, utilizando o plasmídeo pET19b como controle nas diferentes linhagens. A linhagem BL21(DE03)pTetglyVXY contendo o plasmídeo para MaSp2 16× e na ausência do indutor, demorou mais para crescer em relação às outras linhagens, chegando a entrar na fase estacionária da curva de crescimento (linha verde deslocada para a direita na figura 31.C). Fato esse que não foi observado na presença do indutor, onde as linhagens apresentam um crescimento muito semelhante (Figura 31.D). Como essas linhagens são novas, uma análise proteômica quantitativa e qualitativa é capaz de justificar essa diferença.

Já os resultados para a produção da MaSp2 32× inédita, as linhagens apresentaram curvas e velocidades de crescimento semelhantes aos controles utilizados, tanto na ausência quanto na presença do indutor (Figura 31.D e H). Esses resultados diferem do trabalho de Xia e colaboradores (2010), onde a presença do plasmídeo pTetglyVXY aumentou a produção da biomassa durante a indução das espiroínas MaSp1 32× e MaSp1 96×, mas corroboram com os resultados encontrados por Silva (2014), onde não houve diferença na curva de crescimento entre as linhagens BL21(DE03) e BL21(DE03)pTeglyVXY durante a indução de MaSp2 16×.

Após o estabelecimento das curvas de crescimentos de todas as linhagens, foi possível observar também que a indução das espiroínas MaSp2 16× e MaSp2 32× não alteraram o seus crescimentos. Isso significa que as proteínas produzidas não apresentam propriedades tóxicas

ou afetaram o metabolismo e o crescimento das bactérias de maneira negativa. Esses resultados são semelhantes ao processo de produção de outras espiroínas utilizando bactéria como biofábricas (FAHNESTOCK; IRWIN, 1997; BITTENCOURT, 2007; BROOKS et al., 2008; SOUTO, 2008; Xia et al., 2010; AN, et al., 2011; DAMS-KOZLOWSKA et al, 2012; OLIVEIRA, 2012; SILVA, 2014).

Diferentemente dos outros resultados, todas as linhagens que apresentavam o plasmídeo pET19b como controle, atingiram a fase estacionária após a adição do indutor e permaneceram assim durante todo o período de avaliação (Figura 31.F). No entanto, essas mesmas linhagens apresentaram o mesmo resultado de outras condições, onde as curvas são muito próximas entre si (Figura 31.B). Uma das justificativas para isso seria o funcionamento do operon *lac*, já que a bactéria BL21(DE03) apresenta no genoma um DE03 lisogênio, contendo o gene da T7 RNA polimerase sob o controle do promotor *lacUV5* (*lacI*), tendo a expressão inibida pelo repressor *lac* da própria bactéria. Esse repressor influencia também no promotor T7lac do plasmídeo de expressão do gene de interesse. Ao adicionar o IPTG no volume contendo suspensão a celular, ele liga-se ao repressor *lac*, liberando as regiões operadoras dos promotores *lacUV5*(genoma) e T7lac (plasmídeo). A T7 RNA polimerase produzida atua no promotor T7lac do plasmídeo (*pET system manual*, Novagen). Uma vez que o plasmídeo não contém um inserto, as células demandam uma grande quantidade de energia para produzir essa polimerase, que transcreve fragmentos da região promotora à terminadora, mas que não apresentam função alguma.

Contabilizando essa fase, todas as 24 linhagens foram colocadas na presença ou não de IPTG 0,4 mM, em três réplicas, totalizando 144 curvas de crescimento estabelecidas.

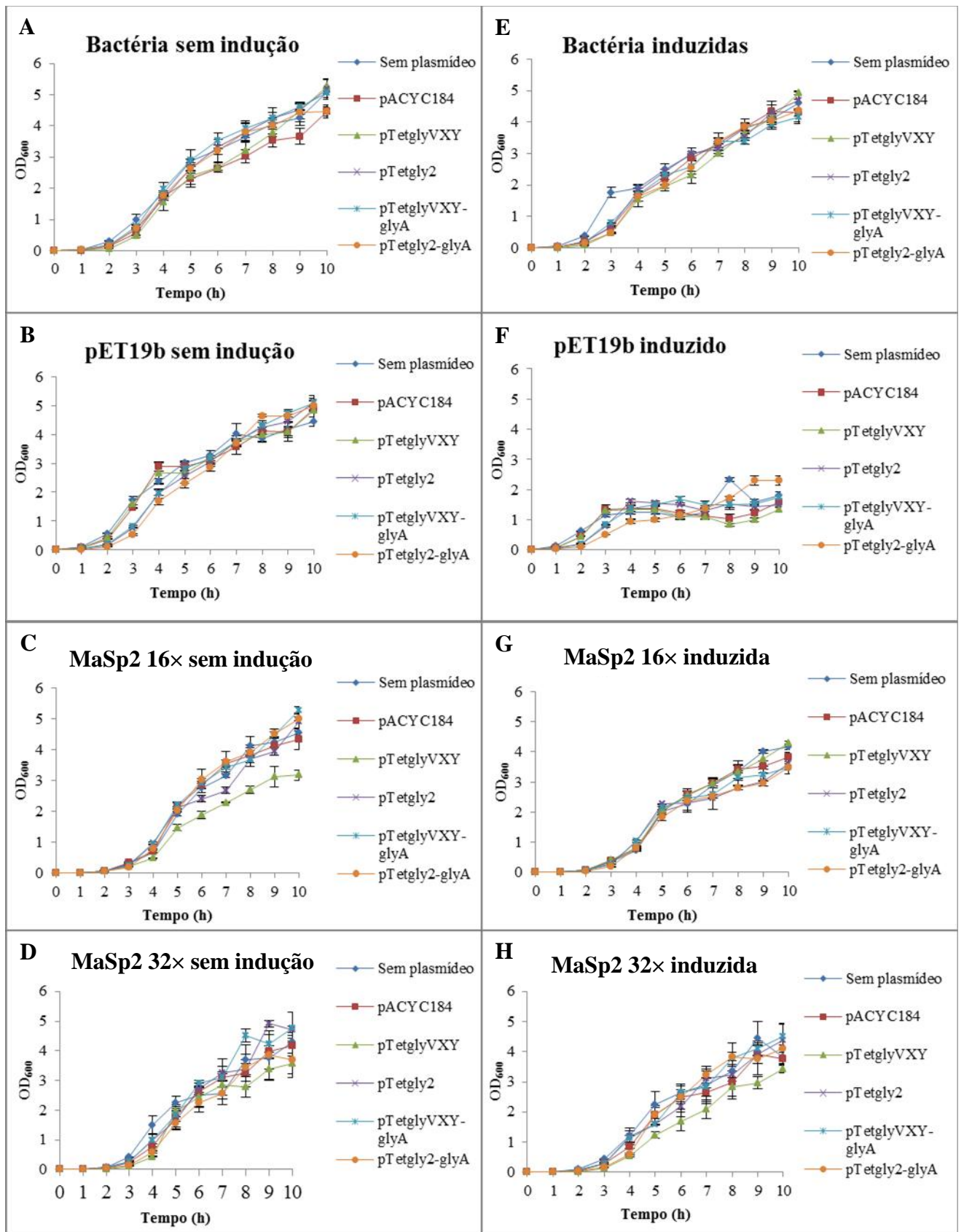


Figura 31. Gráficos das curvas de crescimento das linhagens de *E. coli* em meio LB pelo período de 10 horas. As induções foram realizadas com IPTG 0,4 mM com OD₆₀₀ 0,6. As bactérias BL21(DE03) e a BL21(DE03)pACYC184 foram utilizadas como controles da cepa e de plasmídeo inalterado. MaSp2: Espidroína do tipo 2 da glândula amplosada maior. pTetglyVXY: plasmídeo contendo uma cópia do gene *glyVXY*; pTetgly2: plasmídeo contendo duas cópias do gene *glyVXY*, pTetgly-glyA: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTetgly2-glyA: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. pET19b: plasmídeo de expressão em bactéria.

6.6 Análise por espectrometria de massa Biotyper

As 24 colônias das sete linhagens de bactérias *E. coli* 1-BL21(DE03), 2-BL21(DE03)pLysS, 3- BL21(DE03)pACYC184, 4-BL21(DE03)pTetglyVXY, 5-BL21(DE03)pTetgly2, 6- BL21(DE03)pTetglyVXY-glyA, 7-BL21(DE03)pTetgly2-glyA foram isoladas e cultivadas por 72 h em placas de meio LB Agar a 37 °C, totalizando 168 colônias analisadas. Logo após, o extrato proteico total foi analisado pelo espectrômetro de massa Microflex LRF integrado ao sistema de MALDI *Biotyper*.

As análises por espectrometria de massa MALDI-TOF por meio da utilização da série *microflex* (Bruker), são capazes de identificar perfil proteômicos clínicos, genômica funcional, de pequenas moléculas, analisar polímeros e de micro-organismos em alta performance em questão de poucos minutos. O sistema LRF proporciona precisão das massas e sensibilidade para íons negativos e positivos. Com esse sistema, é possível criar um biblioteca com perfil metabólico de bactérias para futuras identificações dos micro-organismos (*Microflex series*. Disponível em: <<https://www.bruker.com/products/mass-spectrometry-and-separations/maldi-toftof/microflex/technical-details.html>> Acesso em: 05 maio 2015).

As proteínas e os peptídeos das sete linhagens de bactérias analisadas foram organizados em espectros de massa crescente, revelando o padrão característico de cada uma, sendo classificadas em dois grandes grupos. Esses resultados foram armazenados no banco de dados do equipamento e o dendrograma pode ser observado na figura 32. As linhagens comerciais BL21(DE03) e BL21(DE03)pLysS e a linhagem controle BL21(DE03)pACYC184 apresentam um perfil semelhante, sendo inclusas no mesmo grupo (linha azul, Figura 32). Já as linhagens BL21(DE03)pTetglyVXY, BL21(DE03)pTetgly2, BL21(DE03)pTetglyVXY-glyA, BL21(DE03)pTetgly2-glyA para aumentar a disponibilidade de glicil-tRNA e serina hidroximetiltransferase, apresentaram um perfil proteico mais próximos entre si e foram classificadas em outro grupo (linhas vermelhas, Figura 32), confirmando uma alteração metabólica desses novos organismos em relação às linhagens comerciais. Nessa metodologia utilizada, os espectros são detectados quanto ao tamanho, mas não ocorre uma identificação dos peptídeos e/ou proteínas quanto à função e à quantificação das amostras.

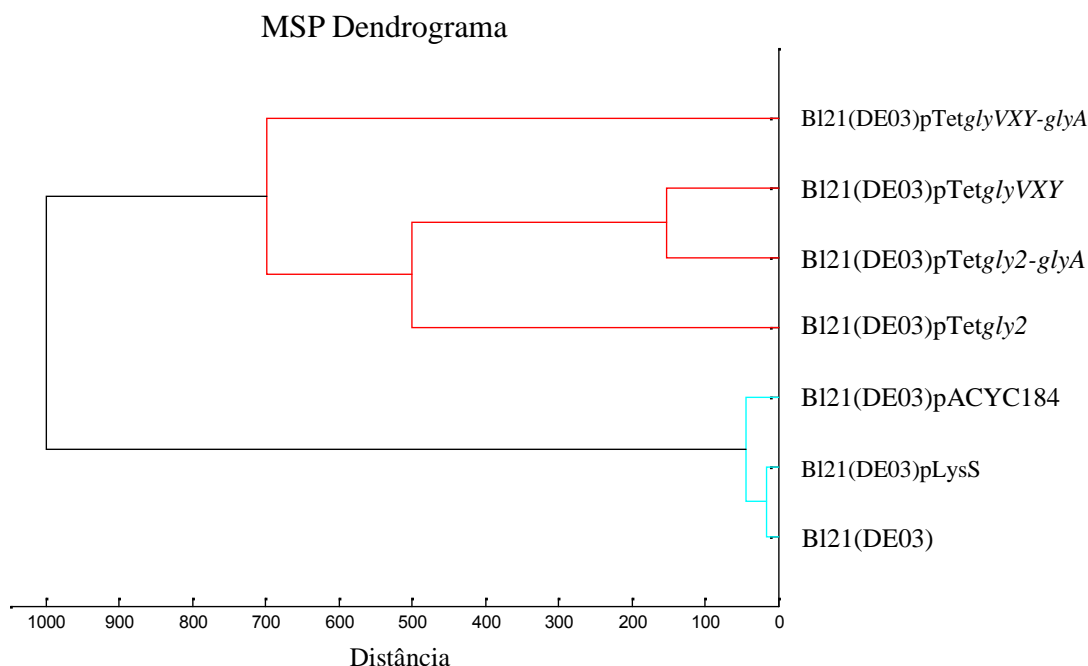


Figura 32. Dendrograma das linhas bacterianas submetidas à análise por espectrometria de massa *biotyper*. MSP: main-spectrum. As linhagens alteradas metabolicamente foram classificadas em um grupo, enquanto as linhagens comerciais (BL21(DE03), BL21(DE03)pLysS e a linhagem utilizada como controle BL21(DE03)pACYC184 foram classificadas em outro grupo. . pTetglyVXY: plasmídeo contendo uma cópia do gene *glyVXY*; pTetgly2: plasmídeo contendo duas cópias do gene *glyVXY*, pTetgly-glyA: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTetgly2-glyA: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*.

As linhagens de bactérias construídas por meio de engenharia metabólica aumentam a quantidade de RNA transportadores (tRNA) de glicina (*glyVXY*) e da enzima serina hidroximetiltransferase (*glyA*). Os diversos tipos de tRNAs apresentam sequências genéticas conservadas e de tamanho pequeno, o que dificulta a sua identificação específica. As técnicas de PCR quantitativo não permitem a identificação e quantificação dessas moléculas e a técnica de RNA seq ainda é muito onerosa para responder a questão. Como a alteração metabólica é para um aminoácido específico, convém avaliar enzimas e proteínas que estejam relacionadas a essa rota. A avaliação por espectrometria de massa por *biotyper* foi utilizada para comprovar que os plasmídeos construídos funcionaram e foram capazes de modificar o metabolismo em relação a linhagens convencionais e ao controle. Para uma melhor compreensão dos resultados, análises quantitativas e qualitativas que identifiquem essas proteínas precisam ser realizadas.

6.7 Estudo proteômico das linhagens bacterianas

Após realizar as curvas de crescimento ($n=3$), os volumes contendo suspensões celulares de todas as amostras foram centrifugadas, armazenadas e as proteínas foram extraídas conforme descrito previamente. O extrato proteico total de todas as linhagens (induzidas ou não) foi quantificado em espectrofotômetro (Figura 33) a fim de verificar a semelhança entre as amostras. Logo após, uma alíquota foi submetida à digestão trípica e analisada qualitativamente e quantitativamente em um espectrômetro de massa Synapt G2 HDMSTM (Waters). Até o momento da defesa desta Tese, os resultados permitiram comparar a linhagem BL21(DE03)pACYC184 em relação à BL21(DE03), sendo identificadas e quantificadas 587 proteínas ao total. Destas proteínas, 55 (9,4 %) foram de ocorrência únicas em BL21(DE03)pACYC184, 146 (24,87 %) foram únicas em BL21(DE03), 106 (18,05 %) foram de um aumento produtivo e 98 proteínas (16,7 %) de diminuição produtiva na linhagem BL21(DE03)pACYC184, e as outras 182 (31,0 %) não apresentaram diferença entre as linhagens. As tabelas com as informações do estudo proteômico encontram-se nos anexos.

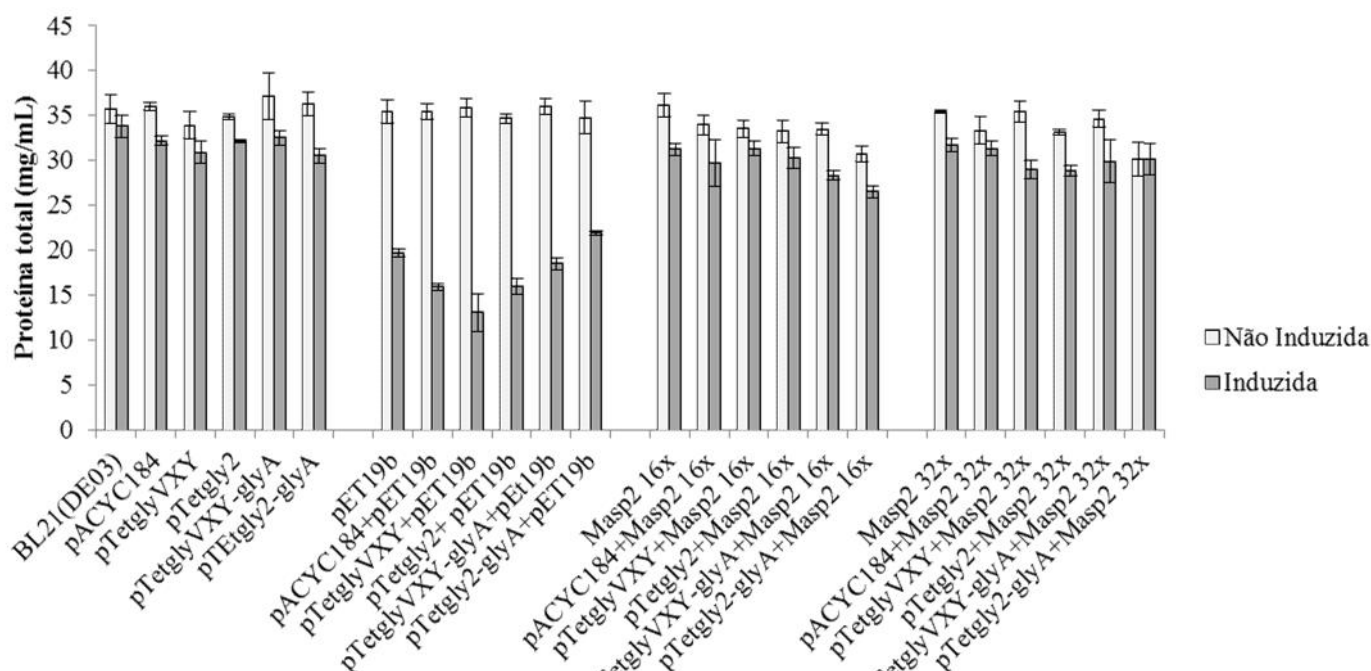


Figura 33. Gráfico ilustrativo da quantificação de proteína total em cada linhagem utilizada nesse experimento. A quantificação foi realizada por absorbância 280 nm em espectrofotômetro nanodrop. pTetglyVXY: plasmídeo contendo uma cópia do gene *glyVXY*; pTetgly2: plasmídeo contendo duas cópias do gene *glyVXY*; pTetgly-glyA: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTetgly2-glyA: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. pET19b: plasmídeo de expressão em bactéria. MaSp2: Espidrofina do tipo 2 da glândula ampolada maior.

Em seguida, os dados foram analisados, classificados e agrupados por função celular pelo programa Blast2Go®. Os grupos de peptídeos e das proteínas, das amostras com ocorrência únicas podem ser observados nas figuras 34.A e 34.B. Enquanto os grupos de proteínas que apresentaram aumento e diminuição de produção entre as linhagens BL21(DE)3pACYC184 e BL21(DE)03 encontram-se nas figuras 35.A e 35.B.

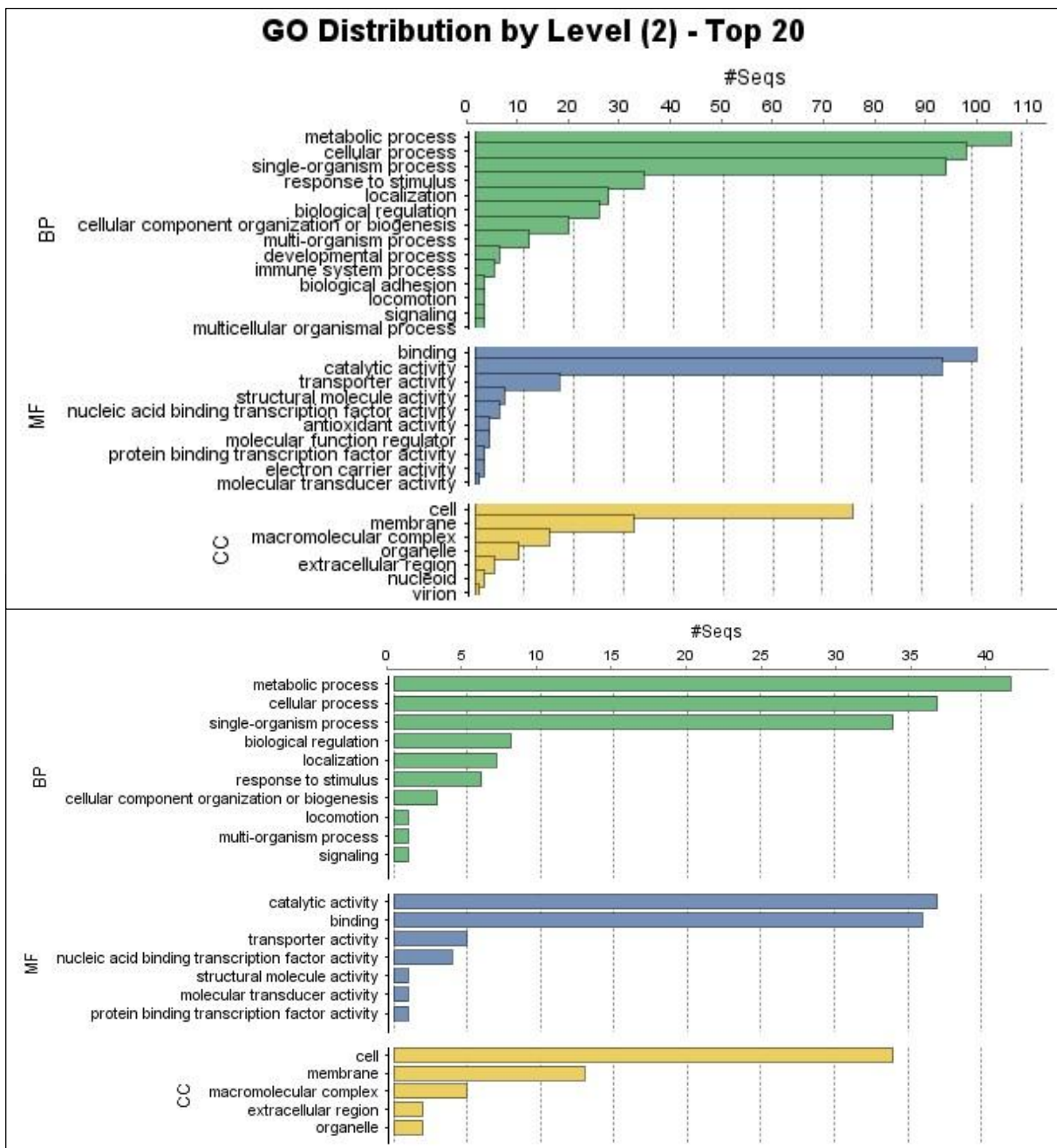


Figura 34. Gráfico de distribuição referente às proteínas encontradas: A) apenas em BL21(DE)03. B: proteínas encontradas apenas na linhagem BL21(DE)03pACYC184. BP: processo biológico. MF: função molecular. CC: componente celular. Blast2Go®.

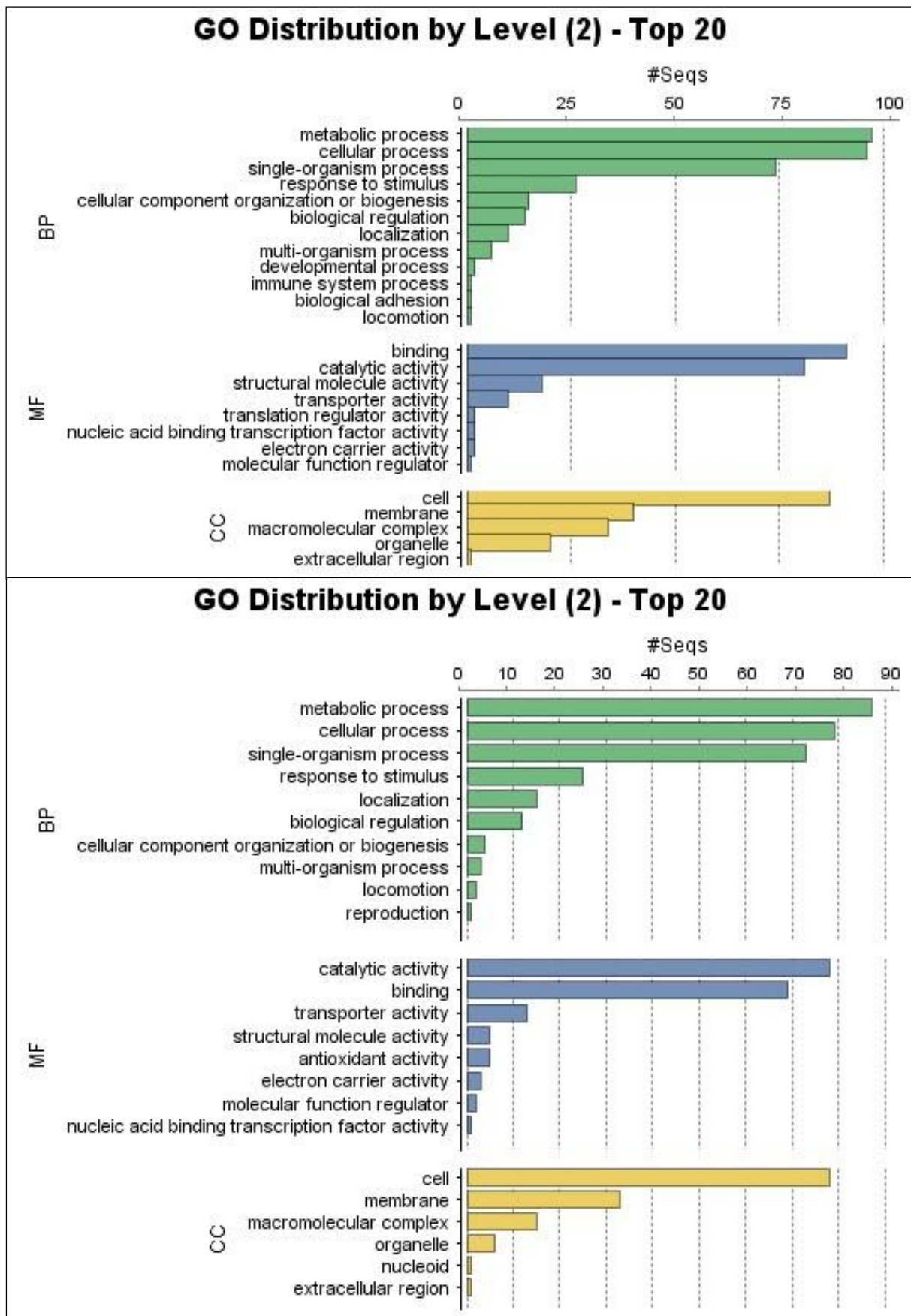


Figura 35. Gráfico de distribuição referente às proteínas encontradas: A) produção aumentadas em BL21(DE03)pACYC184 vs BL21(DE03). B: produção diminuída em BL21(DE03)pACYC184 vs BL21(DE03). BP: processo biológico. MF: função molecular. CC: componente celular. Blast2Go®.

A caracterização do perfil proteômico das linhagens induzidas BL21(DE03) e BL21(DE03)pACYC184, realizados por espectrometria de massa, apresentaram diferenças para o perfil metabólico de glicina. Previamente, um estudo comparativo entre as técnicas de quantificação proteica por Bradford e espectrofotometria foi realizado, e neste caso, a quantificação por espectrofotometria assemelhou-se aos resultados quantitativos do espectrômetro de massa (dados não mostrados). Nota-se uma quantidade maior de proteínas únicas com função molecular, de processo biológico e como componente celular no metabolismo da linhagem BL21(DE03) em relação à mesma linhagem contendo o plasmídeo pACYC184 (Figura 27). A presença desse plasmídeo aumentou quantitativamente a produção de 106 proteínas e a diminuição de 98 proteínas. Dentre as 106 proteínas, foram classificadas 12 categorias para o processo biológico, 8 categorias para a função molecular e 5 categorias para componente celular. Já o grupo com 98 proteínas, a classificação identificou 10 categorias para processos biológicos, 8 categorias para função molecular e 6 categorias para componente celular.

Xia e colaboradores (2010) realizaram um estudo comparativo de perfil proteômico de bactérias BL21(DE03) produzindo espidroínas MaSp1 de *Nephila clavipes* de 100,7 kDa e 284,9 kDa. Após a eletroforese 2D, 31 proteínas da linhagem contendo a espidroína de 289,4 kDa encontravam-se com maior produção e foram identificadas por espectrometria de massa. Ao correlacionar essa lista de proteínas com o presente estudo, verificou-se que a presença do plasmídeo pACYC184 foi capaz de aumentar, em relação à BL21(DE03), a produção de 8 dessas proteínas relacionadas ao estresse metabólico para síntese de proteínas de teia de aranha em bactéria (transcetolase1, subunidade alfa da ATP sintase, cisteína A sintase, aspartato amônia-liase, aspartato aminotransferase, proteína ribossomal L9 50S, chaperonina 60 kDa, chaperonina 10 kDa). Os futuros resultados precisam ser analisados com atenção, a fim de indicar outras alterações metabólicas que podem ser realizadas para melhorar o sistema de produção das espidroínas em *E. coli*.

6.8 Caracterização e purificação das espidroínas MaSp2

6.8.1. Determinação da solubilidade das espidroínas

A solubilidade da proteína alvo foi determinada conforme as recomendações do manual *The QIAexpressionist™* 5ª edição (Qiagen) com as modificações descritas previamente. As alíquotas do sobrenadante (S2) e dos precipitados (P1 e P2) foram separadas por SDS-PAGE e corados com azul de *Comassie* (Figura 36). Todas as linhagens avaliadas apresentaram a produção das espidroínas MaSp2, não havendo diferenças nos processos de extração e purificação. Uma etapa importante durante essa extração é o aquecimento a 80 °C, por auxiliar na precipitação proteica de *E. coli* (AN, et al., 2012; OLIVEIRA, 2012, TEULÉ et al., 2012.), e manter as espidroínas na porção solúvel. Durante a precipitação ocorreu a mudança da cor da solução, de amarela translúcida para branca opaca.

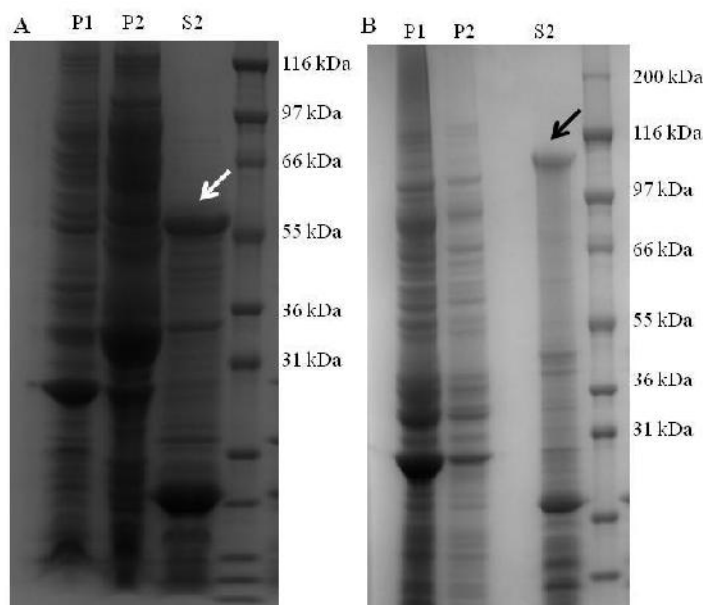


Figura 36. Análise de SDS-PAGE em gel pré-moldado 4- 12% Novex (Invitrogen) utilizando tampão MOPS para determinar a solubilidade das espidroínas durante a extração de proteínas. Os extratos das amostras contendo o precipitado (P1 e P2) e o extrato solúvel (S2) indicam que as espidroínas encontram-se na porção solúvel. Setas indicam a proteína alvo. A) MaSp2 16× (54 kDa); B) MaSp 32× (105 kDa), Marcador: *Novex sharp pre stain protein standard* (Invitrogen).

As espidroínas MaSp2 16× (54 kDa) e MaSp2 32× (105 kDa) foram detectadas na porção solúvel durante a extração proteica para todas as linhagens avaliadas. Esses resultados corroboram com a extração de outras espidroínas sintéticas, com ou sem sequências de histidina, produzidas em bactérias e plantas (SCHELLER et al., 2001; DAMS-KOZLOWSKA et al., 2012; OLIVEIRA, 2012; TEULÉ, et al., 2012). Após a purificação por cromatografia de afinidade, essas proteínas que apresentam o ponto isoelétrico próximo ao neutro (6,68 e 6,67, respectivamente) eram desalinizadas em solução tampão alcalino (Tris 10 mM pH 8,0) pelas primeiras 24 horas e posteriormente em água a 4 °C para evitar e/ou diminuir suas precipitações. Mesmo seguindo essas recomendações, notavam-se microfilamentos na solução, que eram retirados por centrifugação. Não é aconselhada a dessalinização ou concentração dessas proteínas por meio de tubos contendo filtros, pois as espidroínas aderem ao material utilizado na fabricação do filtro. As proteínas precipitadas não devem ser utilizadas porque a mudança de conformação do estado solúvel para insolúvel dificulta a resolubilização dessas proteínas para o processo de formação de fibras sintéticas. Esses resultados diferem aos estudos bioquímicos realizados com fragmentos de ADF3 e ADF4, onde a ausência do domínio C-terminal da proteína não permite a sua agregação e formação de nanofibras (HUEMMERICH, et al., 2004).

6.8.2 Caracterização molecular e rendimento de Masp2 (16×)

A proteína MaSp2 16× é composta de 650 resíduos de aminoácidos, sendo os mais abundantes a glicina (227), alanina (176), glutamina (80) e prolina (65), representando 34,9%, 27,1%, 12,3% e 10% da composição final. Ela apresenta uma massa molecular média estimada de 54,227 kDa, massa monoisotópica 54,195 kDa o ponto isoelétrico é de 6,68, com coeficiente de extensão molar (280 nm) de $40,960 \text{ cm}^{-1}\text{M}^{-1}$, valor de absorvância molar de $0,76 \text{ cm}^{-1}$ e índice de hidrofobicidade de -0,44. Além disso, não apresenta sequência para N-glicosilação, não apresenta resíduos de cisteína e a sequência da espidroína não apresenta sítios para ação de proteases como tripsina, Asp-N, Arg-C, Lys-C, Glu-C para. No entanto, a quimotripsina é capaz de hidrolisar ligações peptídicas na poção carboxi terminal de aminoácidos hidrofóbicos como tirosina, triptofano, fenilalanina e leucina, resultando uma fragmentação, *in silico*, da proteína MaSp2 16× em 33 partes. A meia-vida *in vivo* estimada em *Escherichia coli* e fungos é maior que 10 h e 20 h, respectivamente, e em

mamíferos é de aproximadamente 30 h. Os parâmetros físico químicos da proteína foram calculados na página www.expasy.org, informando a sequência de aminoácidos. Os resultados por SDS-PAGE e *Western blot* do processo de purificação e a confirmação da espidroína purificada encontram-se na figura 37.

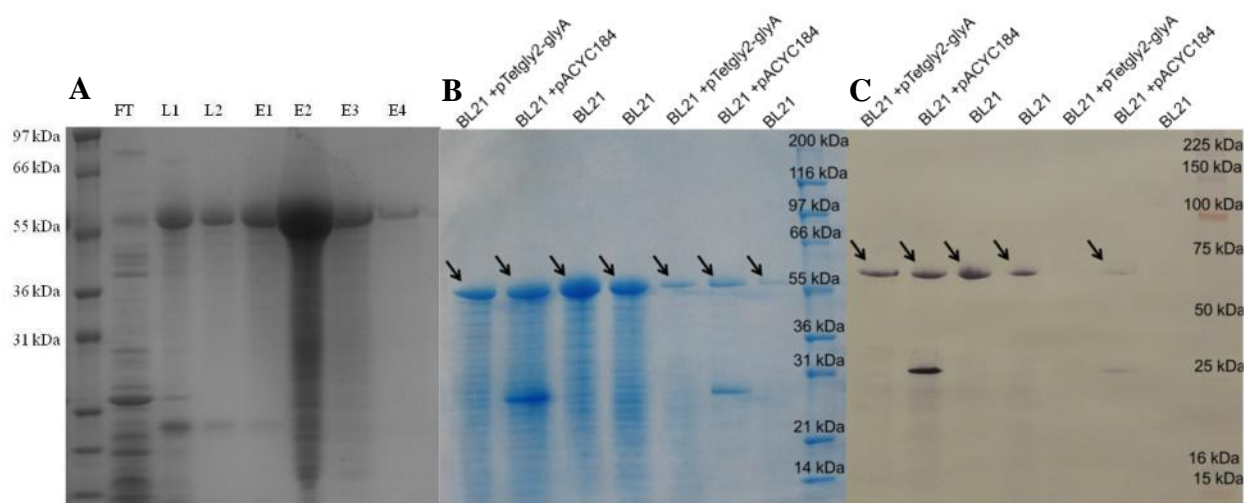


Figura 37. A) Análise por SDS-PAGE do processo de purificação da MaSp2 16 \times . FT: passado; L1-2: Lavado com Imidazol 40 mM; E1-2: Eluído com Imidazol 100 mM; E3-4: Imidazol 250 mM. Proteína com massa molecular de 54 kDa. B) Espidroína purificada utilizando diferentes linhagens de expressão. C) Identificação das espidroínas por meio da sequência de histidina por *Western blot*. As setas em B e C indicam a proteína alvo, e abaixo delas, são espidroínas menores que foram produzidas pela bactéria devido a erros na transcrição e tradução.

Após o processo de purificação, as soluções contendo as espidroínas foram submetidas à SDS-PAGE e *Western blot*, sendo detectadas com o tamanho esperado, mas outras proteínas menores sempre apareciam nos géis. Esse fato já foi observado em outros trabalhos (FAHNESTOCK; BEDZYK, 1997; BARR et al., 2004; HAUPTMANN et al., 2013). Por serem espidroínas também, não impedem a formação da fibra, podendo auxiliar no processo de polimerização. Esses fragmentos menores ocorrem por problemas durante a transcrição e tradução das sequências repetitivas (HINMAN; JONES e LEWIS, 2000; revisto por RISING et al., 2011).

A biomassa de cada linhagem foi adquirida por balança de precisão, após realizar a indução proteica e precipitá-la por meio de centrifugação. A linhagem BL21(DE03)pTetglyVXY-glyA, contendo o plasmídeo de produção para a MaSp2 16 \times , obteve a menor quantidade de biomassa, média de 4,23 g, em relação às linhagens de BL21(DE03), média de 6,68 g, BL21(DE03)pACYC184, média de 6,84 g, BL21(DE03)pTetgly2-glyA, média de 6,73 g ($p < 0,01$ e $CV = 13,68\%$). No entanto, a linhagem BL21(DE03) não diferiu

das linhagens BL21(DE03)pTetglyVXY, média de 5,25 g, e BL21(DE03)pTetgly2, média de 5,45 g (Figura 38.A).

O rendimento total da proteína MaSp2 16× (54 kDa) após a purificação também foi medido por balança de precisão. A linhagem BL21(DE03)pTetgly2 obteve um rendimento maior na produção da espidroína, com média de 16,03 mg/L, em relação às linhagens BL21(DE03)pACYC184 (6,94 mg/L), BL21(DE03)pTetglyVXY (9,3 mg/L), BL21(DE03)pTetglyVXY-glyA (7,3 mg/L) e BL21(DE03)pTetgly2-glyA (7,26 mg/L) ($p < 0,05$, $CV = 35,87\%$). Contudo, não houve diferença no rendimento da MaSp2 16× entre a linhagem BL21(DE03)pTetgly2 (16,03 mg/L) e BL21(DE03) (15,42 mg/L) (Figura 38.B).

Por existirem processos fermentativos sem e com parâmetros controlados, como os fermentadores comerciais, tem-se modificado o modo de apresentar os resultados para essas metodologias. Sugere-se que eles sejam evidenciados pela razão da quantidade de proteína purificada com a biomassa adquirida. Sendo assim, a linhagem BL21(DE03)pTetgly2 obteve uma média de rendimento maior para a produção da espidroína MaSp2 16× por biomassa (2,94 mg/g) em relação às linhagens BL21(DE03)pACYC184 (0,89 mg/g) e BL21(DE03)pTetgly2-glyA (1,07 mg/g) ($p < 0,01$, $CV = 27,38\%$). A linhagem BL21(DE03) apresentou uma média de produção maior (2,36 mg/g), em relação à BL21(DE03)pACYC184 (0,89 mg/g), mas com rendimento semelhante à BL21(DE03)pTetgly2 (2,94 mg/g) ($p < 0,01$, $CV = 27,38\%$). Por conseguinte, as linhagens BL21(DE03)pTetglyVXY (1,76 mg/g) e BL21(DE03)pTetglyVXY-glyA (1,72 mg/g) apresentaram resultados semelhantes às demais linhagens (Figura 38.C).

Apesar da espidroína MaSp2 16× com 54 kDa de massa molecular apresentar uma grande quantidade de resíduos de aminoácidos específicos (glicina, alanina, glutamina e prolina), a bactéria parece não ter esgotado seus recursos metabólicos para a produção dessa proteína. Além disso, as novas linhagens de bactéria apresentaram um rendimento semelhante à linhagem comercial BL21(DE03). Esperava-se que as bactérias alteradas metabolicamente aumentassem o rendimento para as espidroínas testadas, uma vez que o sistema de produção por plasmídeos pET19b são capazes de acumular o produto desejado em até 50% de proteína total das células (*manual pET system*, Novagen). Esses resultados corroboram com os resultados de Xia e colaboradores (2010), uma vez que não foram encontradas diferenças para o rendimento das espidroínas MaSp1 16× e 32× utilizando as mesmas linhagens de bactérias descritas por este trabalho. No entanto, divergem de outro trabalho, onde a produção desta mesma espidroína purificada apresentou uma média aritmética com rendimento menor para as linhagens BL21(DE03) (6,35 mg), BL21(DE03)pTetglyVXY (10,51 mg) e

BL21(DE03)pTet*gly*2 (4,93 mg) (SILVA, 2014). Isso se deve ao fato do cultivo dos pré-inóculos ser menor que 24 horas. Para um rendimento maior de proteínas, as bactérias devem ser cultivadas por 48 horas em meio de cultura, antes de serem induzidas (Sambrook, 2003).

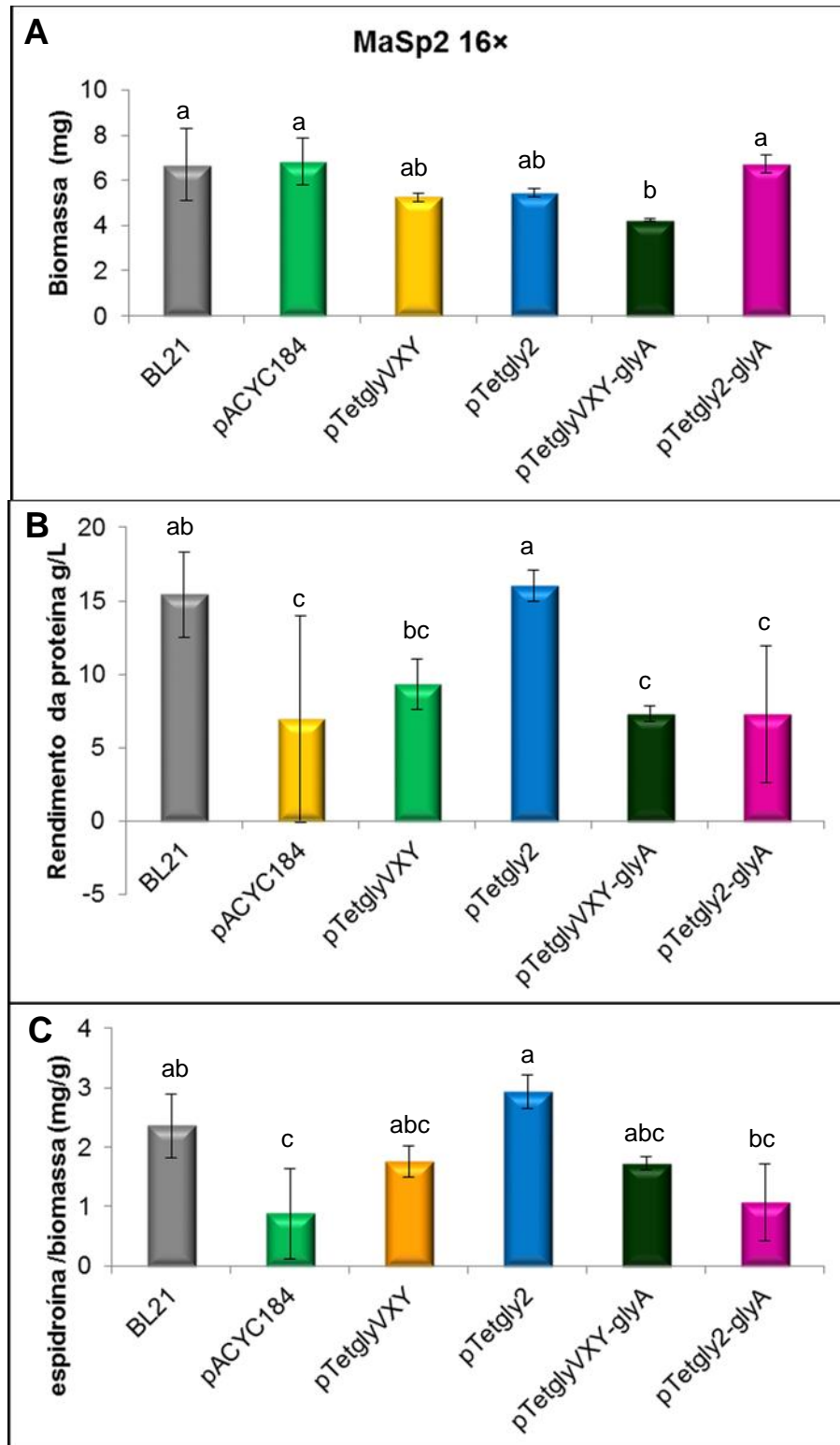


Figura 38. Gráficos representativos dos resultados de rendimento da epidrioina sintética de MaSp2 16× (54kDa) de *Parawixia bistrata* em diferentes linhagens bacterianas. A) Biomassa de bactéria ao final da indução, $p < 0,01$, CV=13,68%. B) Rendimento de purificada a partir da indução em um litro, $p < 0,05$, CV= 35,87%. C) Quantidade de epidrioina purificada por biomassa de bactéria, $p < 0,01$, CV=27,38%. pTetglyVXY: plasmídeo contendo uma cópia do gene *glyVXY*; pTetgly2: plasmídeo contendo duas cópias do gene *glyVXY*, pTetgly-glyA: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTetgly2-glyA: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. As linhagens BL21(DE03) e BL21(DE03)pACYC184 serviram como controle.

6.8.3 Caracterização molecular e rendimento MaSp2 (32×)

A proteína MaSp2 32× com a sequência amino terminal de histidina é composta de 1.274 resíduos de aminoácidos, sendo os mais abundantes a glicina (451), alanina (352), glutamina (160) e prolina (129), representando 35,4 %, 27,6%, 12,6% e 10,1% da estrutura final. Ela apresenta uma suposta massa molecular de 105,383 kDa, massa monoisotópica de 105,322 kDa, o ponto isoelétrico é 6,67, com coeficiente de extensão molar (280 nm) de $81.920 \text{ cm}^{-1}\text{M}^{-1}$, valor de absorvância molar à $0,78 \text{ cm}^{-1}$ e índice de hidrofobicidade de -0,41. As descrições de modificações pós-traducionais e das próteses são as mesmas da MaSp2 16×. Sendo assim, a quimotripsina é capaz de fragmentar, *in silico*, a proteína MaSp2 32× em 65 partes. A meia-vida *in vivo* estimada em *Escherichia coli* e fungos é maior que 10 h e 20 h, respectivamente, e em mamíferos é de aproximadamente 30 h. Os parâmetros da proteína foram calculados na página www.expasy.org, informando a sequência de aminoácidos.

Após a extração de proteína das linhagens BL21(DE03), BL21(DE03)pACYC184, BL21(DE03)pTetglyVXY, BL21(DE03)pTetgly2, BL21(DE03)pTetglyVXY-glyA, BL21(DE03)pTetgly2-glyA, as amostras foram purificadas por coluna de afinidade carregada com níquel *Histrap* (Invitrogen) de 5 mL, sendo submetidas a um processo de gradiente de eluição com imidazol no sistema Akta (GE). O início da eluição da proteína de 105 kDa ocorreu com solução tampão contendo imidazol 100 mM e finalizou com 135 mM. Esse resultado foi confirmado por ELISA e SDS- PAGE (dados não mostrados). Além disso, esse processo também confirmou que a metodologia utilizada para a lavagem da coluna com solução tampão contendo imidazol 40 mM não fez a proteína se desligar da coluna, e as eluições iniciando com imidazol 100 mM e finalizando com 250 mM são ideais para a recuperação da proteína alvo. A detecção da proteína MaSp2 32×, por esta metodologia de purificação, ocorreu por meio de SDS-PAGE e *Western blot* utilizando um anticorpo contra sequência de histidina (Figura 39).

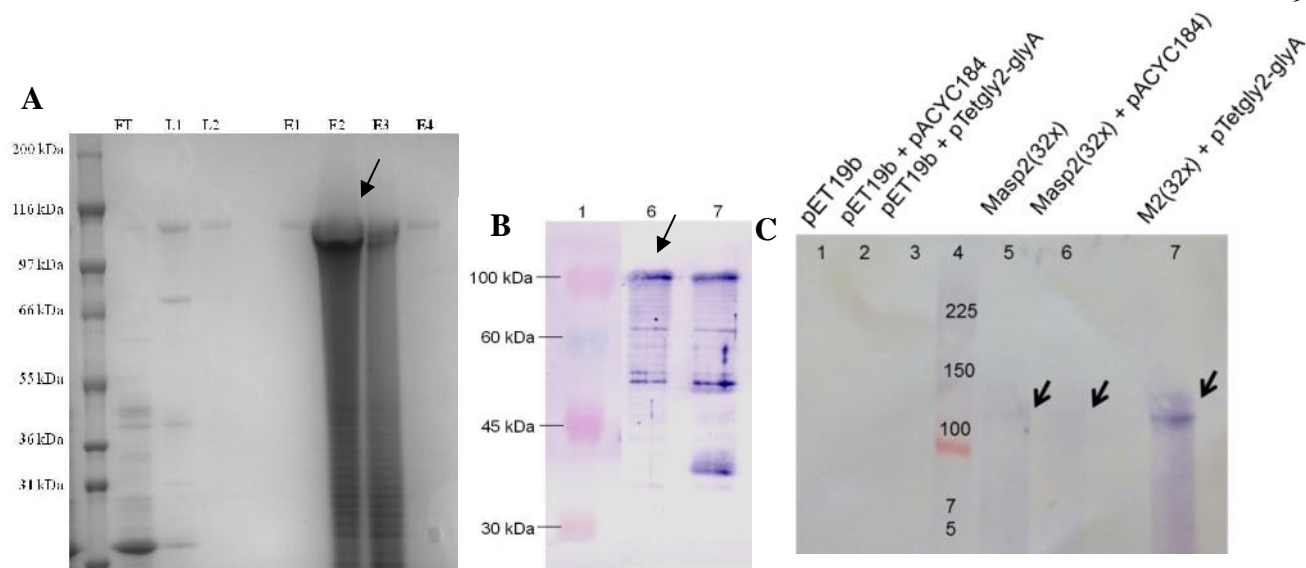


Figura 39. Fotografias das análises por SDS-PAGE e Western blot da MaSp2 32 \times . A) SDS-PAGE processo de purificação FT: passado; L1-2: Lavado com Imidazol 40 mM; E1-2: Eluído com Imidazol 100 mM; E3-4: Imidazol 250 mM. B) *Western blot* das espidroínas detectadas em extrato proteico total da bactéria. 1- *ColorBurst™ Electrophoresis Marker* (C1992-SIGMA), 6 e 7: amostras contendo a MaSp2 32 \times com a sequência de histidina N-terminal. C) *Western blot* para identificação da MaSp2 32 \times após o processo de purificação, 1-3: Linhagens utilizadas como controle negativo, 4- *Novex sharp pre stain protein standard* (Invitrogen), 5-7: Espidroína produzida em diferentes linhagens. As setas indicam a proteína com massa molecular de 105 kDa.

Após as purificações, as soluções proteicas foram desalinizadas por diálise durante 48 horas, congeladas, liofilizadas, quantificadas em balança de precisão e os rendimentos encontram-se na figura 40. A biomassa de cada linhagem também foi adquirida por balança de precisão, após realizar a indução proteica e precipitá-la por centrifugação. As linhagens BL21(DE03)pACYC184 e BL21(DE03)pTetglyVXY, contendo o plasmídeo de produção de MaSp2 32 \times , obtiveram a menor quantidade de biomassa, média de 4,71 g e 4,62 g, em relação às linhagens de BL21(DE03) (6,32 g), BL21(DE03)pTetgly2 (5,05 g), BL21(DE03)pTetglyVXY-glyA (4,95 g) e BL21(DE03)pTetgly2-glyA (5,23 g) ($p < 0,05$) (Figura 40.A).

O rendimento total da proteína MaSp2 32 \times (105 kDa) após a purificação também foi medido por balança de precisão. A linhagem BL21(DE03)pTetgly2 apresentou um rendimento maior (12,33 mg/L - média aritmética) para a produção da espidroína do que as linhagens BL21(DE03) (4,93 mg/L), BL21(DE03)pACYC184(4,66 mg/L), BL21(DE03)pTetglyVXY (4,3 mg/L). A linhagem BL21(DE03)pTetglyVXY-glyA apresentou um rendimento similar (14,04 mg/L) em relação à linhagem BL21(DE03)pTetgly2(12,33 mg/L) e à BL21(DE03)pTetgly2-glyA(12,33 mg/L). Já a linhagem BL21(DE03)pTetgly2-glyA apresentou o maior rendimento para a produção dessa espidroína (17,75 mg/L)($p < 0,01$; CV= 15,6 %)(Figura 40.B).

Ao relacionar os dados obtidos pela razão da quantidade de proteína purificada com a biomassa adquirida, a linhagem BL21(DE03)pTet*gly*2 o foi maior (média de 2,45 mg/g) em relação às linhagens BL21(DE03) (0,79 mg/g), BL21(DE03)pACYC184(1,34 mg/g), BL21(DE03)pTet*gly*VXY (1,36 mg/g). A linhagem BL21(DE03)pTet*gly*VXY-*gly*A apresentou um rendimento similar (2,85 mg/g) comparada à linhagem BL21(DE03)pTet*gly*2 (2,45 mg/g) e à BL21(DE03)pTet*gly*2-*gly*A(3,39 mg/g.). Já a linhagem BL21(DE03)pTet*gly*2-*gly*A apresentou o maior rendimento (3,39 mg/g) ($p < 0,01$; CV= 13,72 % (Figura 40.C).

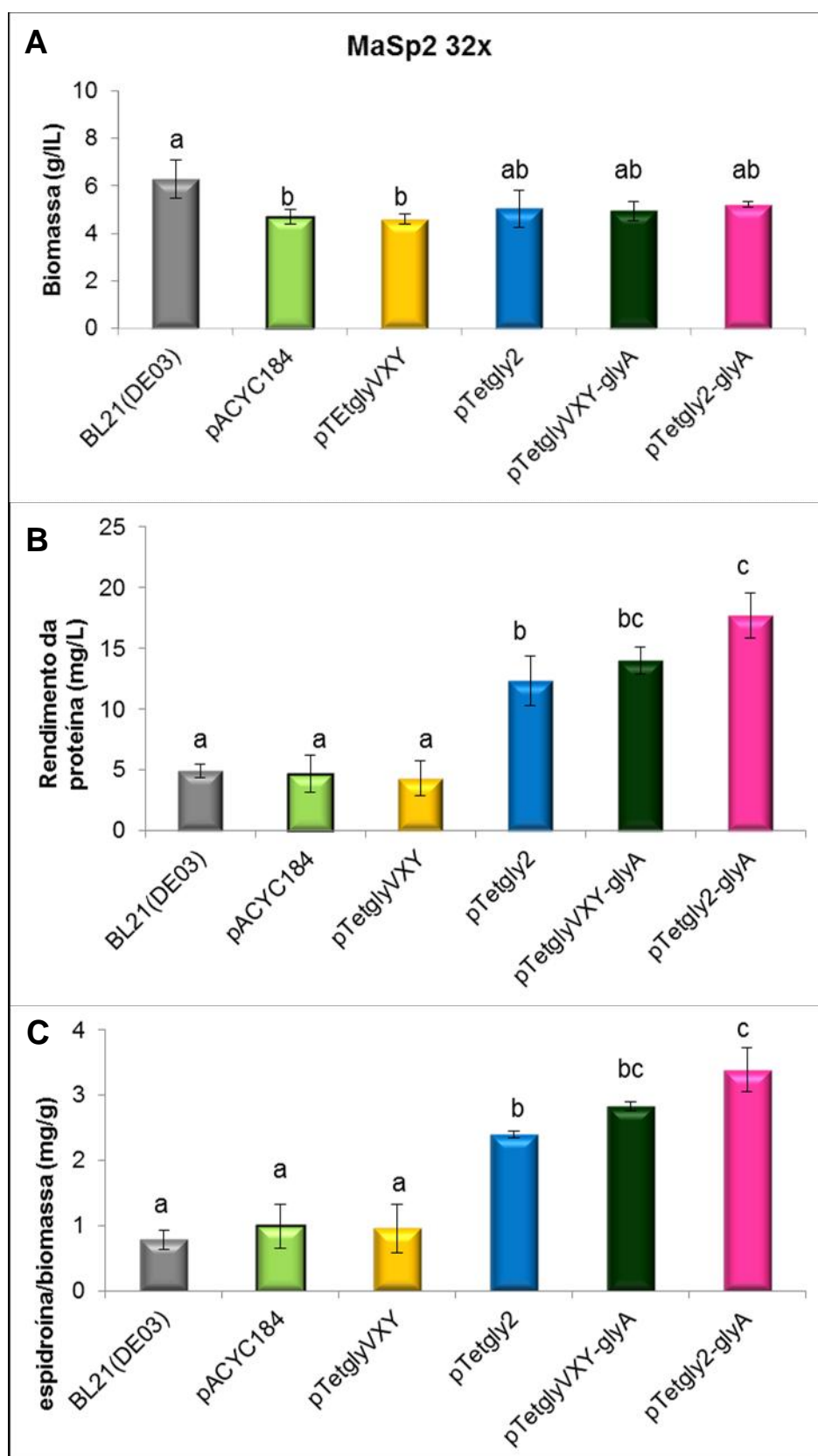


Figura 40. Gráficos representativos dos resultados de rendimento da epidroína sintética de MaSp2 32× (105 kDa) de *Parawixia bistrriata* em diferentes linhagens bacterianas. A) Biomassa de bactéria ao final da indução, $p < 0,05$. B) Rendimento de purificada a partir da indução em um litro, $p < 0,01$, CV= 15,6%. C) Quantidade de epidroína purificada por biomassa de bactéria, $p < 0,01$, CV=13,72%. pTetglyVXY: plasmídeo contendo uma cópia do gene *glyVXY*; pTetgly2: plasmídeo contendo duas cópias do gene *glyVXY*, pTetgly-glyA: plasmídeo contendo uma cópia do gene *glyVXY* e *glyA*; pTetgly2-glyA: plasmídeo contendo duas cópias do gene *glyVXY* e uma cópia do gene *glyA*. As linhagens BL21(DE03) e BL21(DE03)pACYC184 serviram como controle.

As proteínas MaSp2 32 × de *Parawixia bistriata* ao serem induzidas nas linhagens contendo os plasmídeos pTetgly2, pTetglyVXY-glyA e pTetgly2-glyA apresentaram um rendimento maior comparado à linhagem convencional BL21(DE03), como informado na figura 40. Apesar da bactéria convencional BL21(DE03) ser capaz de produzir a espidroína de 105 kDa, claramente existe um limite de recursos metabólicos devido à composição da proteína de interesse (32% de glicina, 35,4% de alanina, 12,6 % de glutamina e 10% de prolina). Contudo, esse limite parece ser superado pelas linhagens contendo uma quantidade maior de tRNA de glicina, construídas neste trabalho. A linhagem BL21(DE03)pTetglyVXY foi a única a não ser capaz de aumentar o rendimento de MaSp2 32×, apesar do seu metabolismo diferenciado. De acordo com a literatura, a capacidade dessas mesmas linhagens de *E.coli* em aumentar a produção de espidroínas de alta massa molecular só foi detectada a partir de MaSp1 48× (XIA et al., 2010).

6.9 Espectrometria de massa MALDI-TOF

Após o processo de liofilização, as proteínas MaSp2 16× e MaSp2 32× foram submetidas à análise por espectrometria de massa MALDI-TOF pela equipe de suporte da *Tufts University*, Boston, MA, EUA (Figura 41.A-C), e no Laboratório de Biologia Sintética- Embrapa Recursos Genéticos e Biotecnologia (Figura 41.D-E), utilizando ácido sinapínico como matriz. A massa molecular da espidroína MaSp2 16× de 54 kDa foi identificada com sucesso (Figura 41). Todavia, a espidroína de 105 kDa não foi detectada com as metodologias utilizadas.

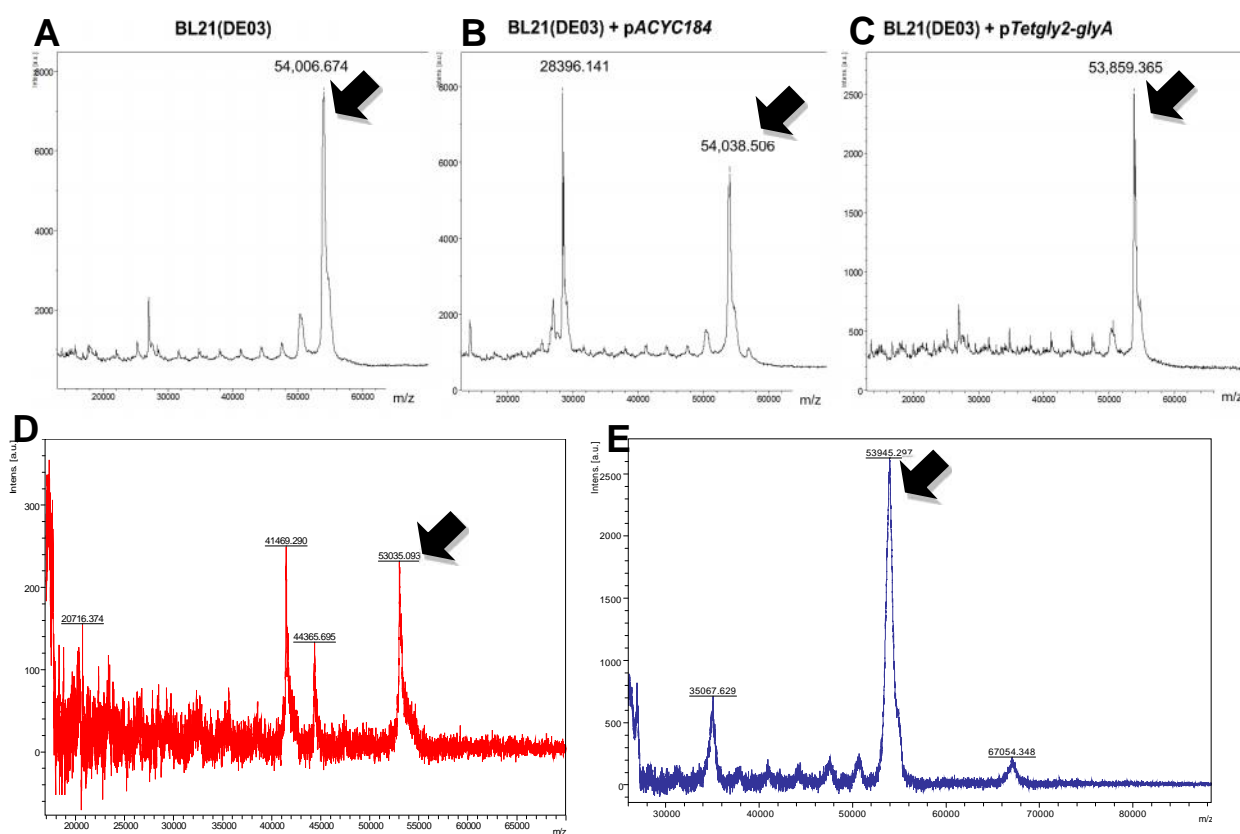


Figura 41. Gráficos representativos das análises realizadas de MaSp2 16× (~54 kDa) por espectrometria de massa MALDI-TOF. As setas pretas indicam os picos referentes à massa molecular. A espidroína purificada foi identificada após a produção em diferentes linhagens de bactérias (A,B,C), e em dois modos de alcance: o de espectros de peptídeo/proteínas de baixa massa molecular (D) e o de detecção de proteínas maiores (E).

Essas espidroínas apresentam características hidrofóbicas, e não são facilmente ionizadas. A matriz de ácido sinapínico foi escolhida por ter sido capaz de ionizar e possibilitar a detecção de outras proteínas de teia de aranha (HUMENIK et al., 2014). A espidroína com 54 kDa (MaSp2 16×) foi detectada com sucesso (Figura 38). Os espectros menores e com pouca intensidade são fragmentos de espidroínas, provenientes de tradução bloqueada ou interrompida durante a síntese proteica na bactéria, identificados previamente nas análises de SDS-PAGE e *Western blot* (Figura 34). Além disso, a amostra diluída em água apresentou um sinal melhor do que a proteína liofilizada. No entanto, a MaSp2 32× (105 kDa) não foi detectada por ser de alta massa molecular e de difícil ionização. Não foram encontrados relatos de detecção de proteínas de teia de aranha identificadas por espectrometria de massa com esse tamanho, sem serem digeridas com enzimas.

Conforme apresentado, essas proteínas não sofrem ação da maioria das enzimas utilizadas para análises de espectrometria de massa. Apenas a enzima quimotripsina fragmenta essas espidroínas, resultando em quatro sequências diferentes, onde duas se repetem várias vezes por terem a mesma sequência e o mesmo tamanho (anexo), impossibilitando a confirmação da massa total (www.alphalyse.com/gpmaw_lite.html).

6.10 Síntese e processamentos das fibras

Após o processo de purificação e diálise, as espidroínas MaSp2 16× e MaSp2 32× foram liofilizadas e mantidas em condições de baixa umidade (dessecador com sílica). A espidroína MaSp2 16× foi submetida a diferentes solventes e concentrações para determinar a melhor condição de solubilização. O solvente 1,1,1,3,3,3- Hexafluoroisopropanol (HFIP) absoluto foi testado para as concentrações proteicas (m/V) 15%, 20% e 25%, não se apresentando solúvel, mesmo após homogeneização no vórtex e aquecimento à 80 °C. A solução adquirida apresentava uma cor opaca e branca, sendo o mesmo problema encontrado em outro trabalho (TUCKER et al., 2014). Logo após, uma solução de HFIP 95% e ácido fórmico 5% foi preparada e adicionada às concentrações de proteína 15% e 20%. Após aquecer a solução a 80 °C por 5 min, as duas amostras apresentavam uma coloração amarela translúcida, indicando que a proteína havia solubilizado.

Utilizando a concentração de 20% (15 mg de proteína em 75 μ L de HFIP 95%+ ácido fórmico 5%), a solução foi inserida numa seringa 250 μ L (Hamilton), e adaptada numa bomba de seringa com conectores de tubo *peek* (cor vermelha), com diâmetro interno de 125 μ m à uma velocidade de 415 μ L/h. A fibra foi extruída com sucesso em banho de coagulação com isopropanol 100%, sendo depositada na área externamente um pedaço de rolo de isofilme transparente (Figura 42) e mantidas em condições de baixa umidade. O isopropanol foi escolhido como o solvente de coagulação, pois as fibras sintéticas produzidas nessas condições apresentaram formações estruturais de folhas- (CHEN et al., 2001), além de propriedades físico-químicas e moleculares superiores aos banhos contendo diferentes concentrações de metanol, isopropanol e água (TEULÉ, et al., 2012; MENEZES et al 2013) .

As proteínas MaSp2 16 \times foram testadas em vários solventes antes de encontrar a solução de solubilização ideal (YANG; ASAKURA, 2005). A presença de grânulos e da cor branca e opaca indica que a proteína não foi solubilizada e podem atrapalhar a síntese da fibra ou mesmo obstruir o equipamento utilizado. Durante a extrusão das fibras em banho de coagulação com isopropanol absoluto e com uma bomba de seringa, a velocidade constante da bomba possibilitou a formação de um fio longo e contínuo. O volume de 100 μ L rendeu vários metros de material, mas que não foram calculados. Uma alternativa para a extrusão de fibras é a utilização de uma câmara, onde é possível controlar as condições da formação de fio da seda sintética de *Bombyx mori* (KINAHAN, et al., 2011), mas que não formaram a fibra de teia de aranha, pois o pH das soluções são muito ácidos (pH 0,3).

As espidroínas MaSp2 32 \times também foi submetida a diferentes concentrações e solventes a fim de determinar as condições de solubilização (YANG; ASAKURA, 2005). As concentrações de 15% da proteína (m/V) em HFIP 75% + ácido fórmico 5%, ou em solução TFA 50%+HFIP 50%, não solubilizaram a proteína, mesmo aquecendo a 80 $^{\circ}$ C por 5 min. Porém, em soluções de HFIP 80 % + ácido fórmico 20% , em ácido clorídrico absoluto ou TFA 50%+ ácido fórmico 50%, a contração de 20% da MaSp2 32 \times foi solubilizada. Na tentativa de formar a fibra, o volume de cinco microlitros, de cada solução, foi pipetado em isopropanol. As proteínas em solução de ácido clorídrico e de HFIP + ácido fórmico sedimentaram formando aglomerados no fundo do béquer, sem formar uma fibra. Todavia, ao realizar esse mesmo procedimento com a solução TFA e ácido fórmico, houve a formação de uma fibra. Sendo assim, a espidroína de 105 kDa foi submetida a essa condição e ao mesmo processo de extrusão de fibras que a espidroína de 54 kDa. Um fio contínuo e com cor branca foi formado no banho de coagulação de isopropanol, mas era frágil demais para ser recuperada, pois ao encostar-se à fibra com uma pinça, ela rompia-se.

Na tentativa de encontrar uma justificativa para a não formação das fibras, o pH dos diferentes solvente utilizados foram mensurados. O reagente HFIP absoluto apresenta um pH 3,0, enquanto as outras soluções testadas (HFIP 90%+Ácido fórmico10%; HFIP 95%+ TFA 5%; HFIP 95%+ TFA 10%) apresentaram ~ pH 1-2, resultando no mesmo problema da câmara com parâmetros controlados descrito. Scheller e colaboradores (2001) afirmam que as espidroínas são solúveis em solução ácida, mas os valores precisam ser determinados, pois soluções muito ácidas impedem a interação molecular das espidroínas, impedindo a formação das fibras. Como não foi encontrado um solvente ideal para solubilizar a espidroína de 105 kDa e formar fibras estáveis, não foi possível a síntese desse material. Recentemente, foi evidenciada um aumento nas propriedade mecânicas de biomateriais compostos de MaSp1 e MaSp2 de *Nephila clavipes* ao adicionar glutaraldeído nas espidroínas solúveis em água ou HFIP (TUCKER, et al., 2014) Mais recente ainda, uma nova técnica denominada como método de solvatação, permitiu a produção de biomateriais a partir de espidroínas dissolvida em água, ajustando as condições de temperatura e pressão (JONES, et al., 2015), mas que ainda não foram testadas nas espidroínas de *Parawixia bistrata*.

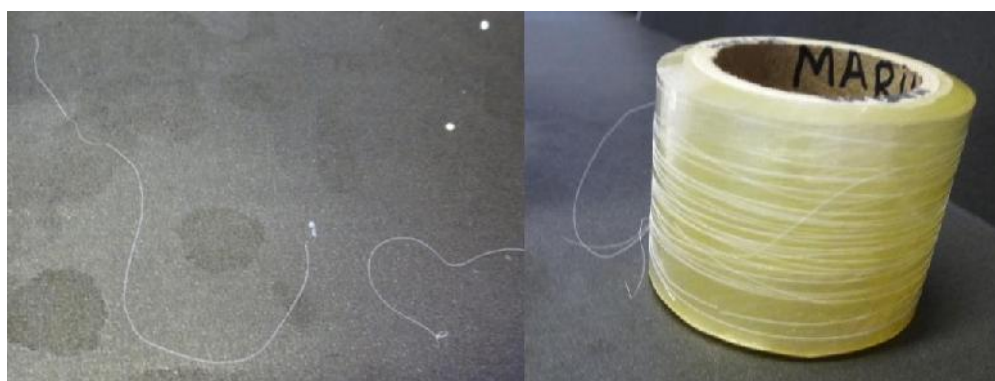


Figure 42. Fotografias das fibras formadas a partir das espidroínas de MaSp2 16× (54kDa). A) Fragmentos da fibra sintética; B) Fibra sintética formada e depositada ao redor da área externa de um pedaço de rolo de isofilme transparente. Foto: Arquivo pessoal, Valquíria Michalczechen-Lacerda.

As fibras formadas durante as duas extrusões foram fragmentadas em tamanhos diferentes, para serem caracterizadas molecularmente e mecanicamente. Para isso, fez-se necessário a utilização de um espaço com iluminação, onde a bancada foi forrada com papel cartão preto e na ausência de correntes de ar, pois tanto a proteína liofilizada, quanto as fibras sofrem influência de estática do ambiente. As fibras foram cortadas em pedaços com 2 cm de comprimento e separados para análises topográficas por MFA e MEV, para análises estruturais por espectroscopia por infravermelho (FTIR-ATR), enquanto os fragmentos com 5

cm foram separados para testes mecânicos. Durante o processamento de hidratação das fibras em isopropanol 80%, elas liberaram pequenas bolhas no meio. Esse tratamento foi escolhido por ter aumentado as propriedades mecânicas de fibras sintéticas da MaSp1 *Parawixia bistriata* (Oliveria, 2012). Ademais, as fibras que receberam uma camada de óleo mineral sofriram menos ação da estática local. Todos os grupos de fibras foram armazenados em placas de petri, revestidas internamente com papel cartão preto, as quais permaneceram em condições de baixa umidade até serem analisadas mecanicamente.

6.11 Análises em nanoescala por Microscopia de Força Atômica

As fibras sintéticas foram separadas em dois grupos para serem analisadas por MFA: 1) Grupo não tratado, cujas fibras não sofreram processamentos (n=10); 2) Grupo tratado, cujas fibras foram processadas após a formação, sendo hidratadas e estiradas ao dobro do tamanho inicial (n=10). As imagens com características morfológicas de superfície foram adquiridas no modo contato intermitente utilizando uma ponteira cônica, e as análises de espectroscopia de força foram adquiridas em modo contato com uma ponteira piramidal.

Ao longo de cada fibra, foram adquiridas três aquisições de imagens para proporcionar os valores médio de Ra (rugosidade média aritmética), Rz (altura máxima), Rzjis (rugosidade média de 10 pontos), Rq (rugosidade quadrática média), Rp (altura média) e Rv (profundidade média). Analisando os resultados entre fibras sintéticas do mesmo grupo, conclui-se que esse material apresenta uma superfície heterogenia (Figura 43), existindo diferença ao longo desse biomaterial. No entanto, a variação dos valores médios e dos desvios padrões diminuíram no grupo das fibras processadas (Figura 43.G-L) em relação ao grupo das fibras não processadas (Figura 43.A-F)

Os resultados foram analisados por teste de Kruskal-Wallis e, os valores foram comparados dentro do mesmo grupo. As fibras não tratadas apresentaram diferenças para as análises de Rz e Rp ($p < 0,05$), não sendo evidenciadas diferenças para Ra, Rzjis, Rq e Rv. Já as fibras processadas foram diferentes entre si em Ra, Rz, Rzjis com $p < 0,01$, os valores de Rq e Rp diferiram com $p < 0,05$, não sendo evidenciadas diferenças para os resultados em Rv.

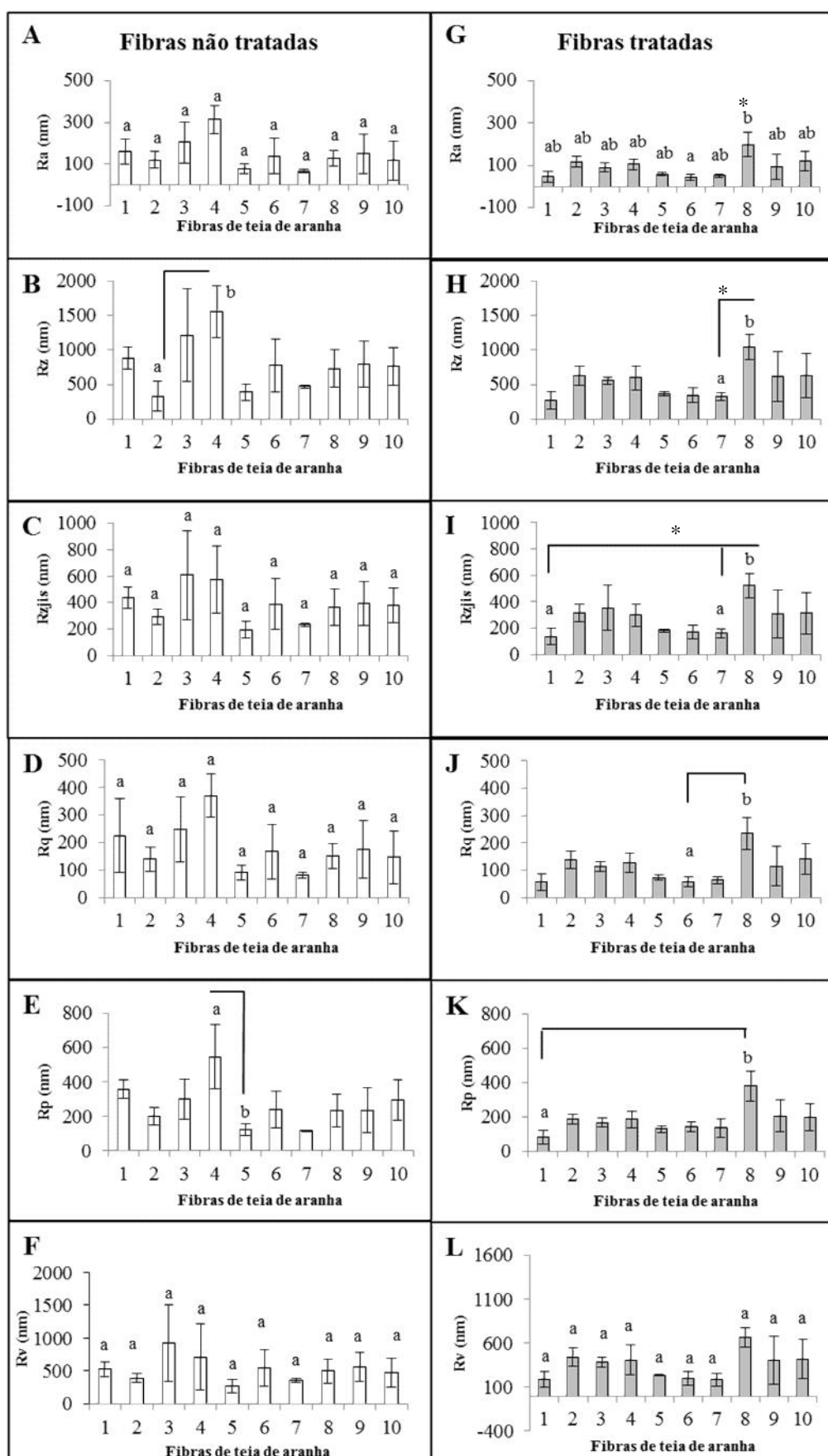


Figura 43. Gráficos representativos das análises de rugosidade das fibras sintéticas não tratadas (A-F) e tratadas (G-L) por MFA em modo contato intermitente (3 imagens para cada fibra)(n=10). Ra: Rugosidade média aritmética; Rq Rugosidade quadrática média; Rzjis: Rugosidade média de 10 pontos; Rp: Altura média; Rv: Profundidade média; Rz: Altura máxima. Letras diferem entre si com $p < 0,05$; * diferem entre si com $p < 0,01$.

Comparando os resultados entre tratamentos, as medidas de rugosidade de superfície para Ra (rugosidade média aritmética), Rz (altura máxima), Rzjis (rugosidade média de 10 pontos), Rq (rugosidade quadrática média) e Rv (profundidade média) apresentam distribuição normal para ANOVA e os teste t resultaram em valores de p iguais a 0,0584; 0,0822; 0,0592; 0,0502 e 0,0374 respectivamente (Figura 44). Os valores para análises Rv apresentaram diferença entre os grupos, ($p < 0,05$) com coeficiente de variação em 38,55 %, ou seja, houve uma diminuição do tamanho dos vales na superfície da fibra. Os resultados de Rp (a altura média) apresentaram uma distribuição estatística não-paramétrica, e foram analisados por Mann-Whitney, onde não foi constatada uma diferença entre os grupos ($P = 0,0892$) (Figura 44).

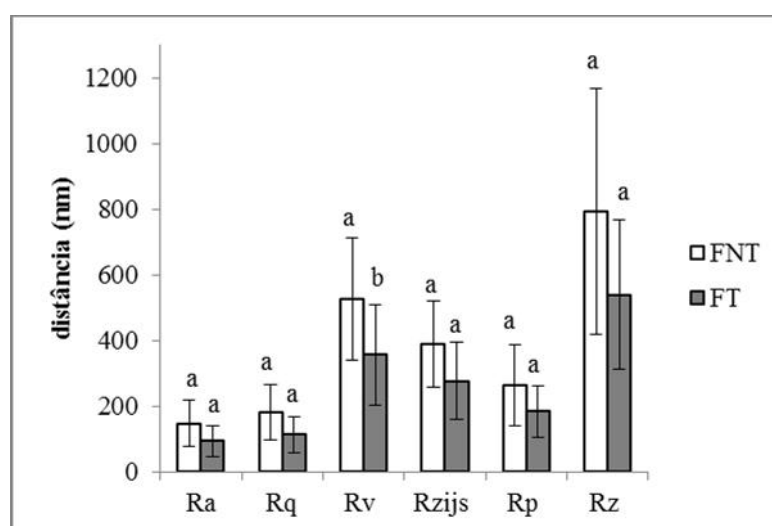


Figura 44. Gráfico de medidas de rugosidade das fibras sintéticas. FNT: fibras não tratadas. FT: Fibras tratadas. As letras são comparadas apenas entre os grupos das fibras, não apresentando informação para os diferentes tipos de parâmetros de análise. Ra: Rugosidade média aritmética; Rq Rugosidade quadrática média; Rzjis: Rugosidade média de 10 pontos; Rp: Altura média; Rv: Profundidade média; Rz: Altura máxima. Letras diferem entre si com $p < 0,05$.

Os resultados de espectroscopia de força para força máxima aplicada e módulo de elasticidade (módulo de Young) foram submetidos ao teste t, onde as médias entre os dois grupos foram consideradas iguais, com valores de P iguais a 0,9727 e 0,0866, respectivamente. Os resultados de módulo de força de atração, força de separação e energia aplicada foram avaliados por Mann-Whitney e os valores de P foram iguais a 0,00288; 0,529; 0,579; sendo a média de força de atração considerada diferente entre os grupos ($p < 0,05$)(Figura 45).

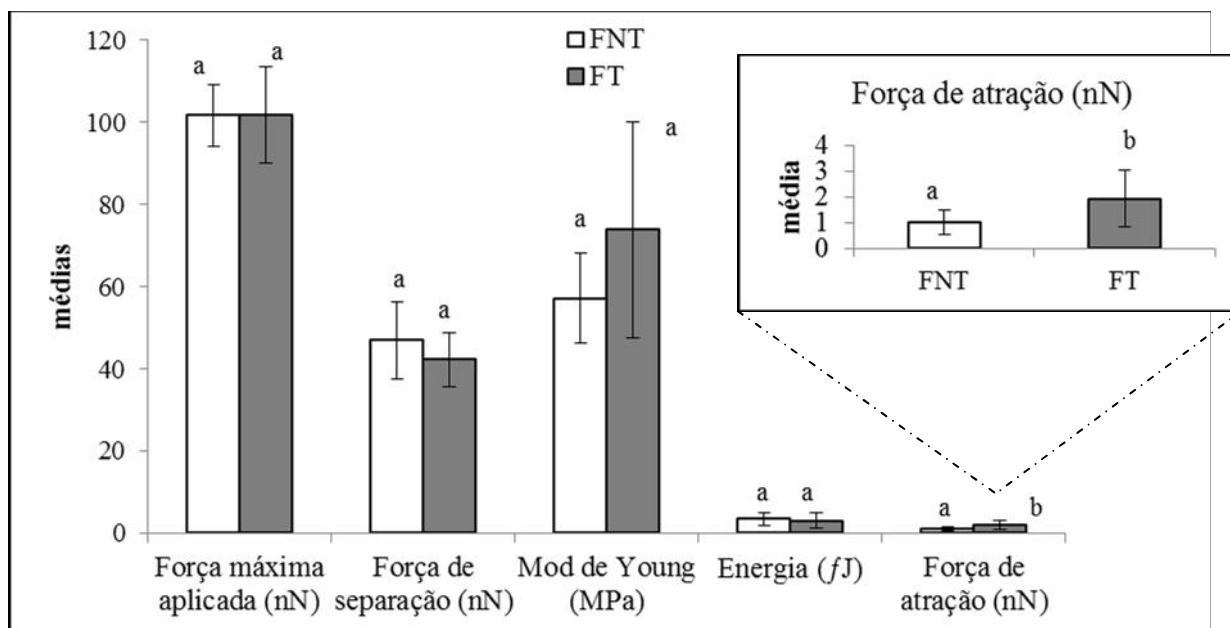


Figura 45. Gráfico representativo da análise de espectroscopia de força por MFA. FNT: Fibras não tratadas; FT: Fibras tratadas. As letras são comparadas apenas entre os grupos das fibras, não apresentando informação para os diferentes tipos de parâmetros de análise. Ra: Rugosidade média aritmética; Rq: Rugosidade quadrática média; Rzj10: Rugosidade média de 10 pontos; Rp: Altura média; Rv: Profundidade média; Rz: Altura máxima. Módulo de Young = módulo de elasticidade. Letras diferem entre si com $p < 0,05$.

Os resultados topográficos de viscoelasticidade foram adquiridos no momento da análise de rugosidade das fibras sintéticas. Da mesma maneira, foram realizadas três medições nas 10 fibras de cada grupo. Os valores de Rzj10, Rq e Rv foram submetidos ao teste t, enquanto Ra, Rz e Rp foram submetidos ao teste de Mann-Whitney. Todas as médias foram consideradas iguais entre os grupos sendo encontrados valores de p iguais a 0,278; 0,0977; 0,4729 para as análises Rzj10, Rq e Rv, e valores de p iguais a 0,0892; 0,912; 0,296 para as análises Ra, Rz e Rp, respectivamente (Figura 46).

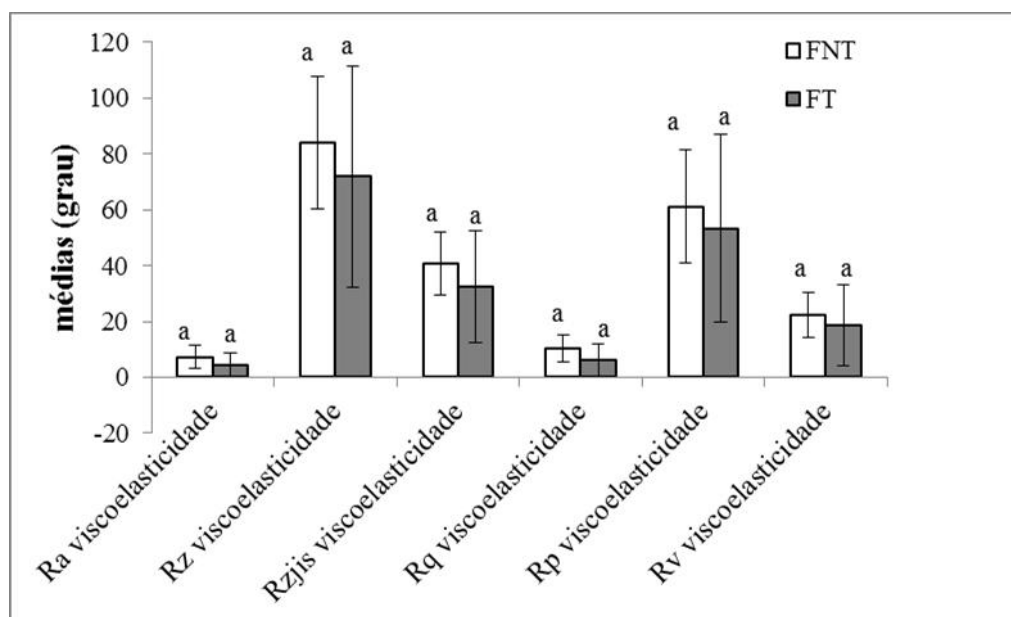


Figura 46. Gráfico representativo da análise de viscoelasticidade da superfície das fibras sintéticas por MFA. FNT: Fibras não tratadas; FT: Fibras tratadas. As letras são comparadas apenas entre os grupos das fibras, não apresentando informação para os diferentes parâmetros de análise.

Os dois grupos de fibras apresentam uma heterogenicidade quanto a rugosidade, viscoelasticidade e espectroscopia de força. Conforme informado, as fibras processadas apresentaram uma diminuição da altura máxima (Rz) em relação às fibras não processadas. A morfologia topográfica média tridimensional (3D) desses dois grupos encontra-se na figura 47.A e D. As proporções de viscoelasticidade encontram-se como imagens de fase, e são das mesmas aquisições das imagens em 3D (Fig. 47 B e E), onde as áreas escuras equivalem a porções viscosas e as áreas claras equivalem a porções elásticas.

As superfícies das fibras sintéticas não processadas e processadas apresentaram uma superfície lisa nas análises por MFA e MEV. Elas assemelham-se às fibras naturais das espécies *Nephila cruentata*, *Nephila clavipes* as quais apresentam essa morfologia (MADSEN; VOLLRAH, 2000; MENEZES, et al., 2013; SWADSON, et al., 2009; SILVA, et al., 2013). As análises de rugosidade por MFA realizadas nas fibras sintéticas MaSp2 16× processadas (hidratadas e estiradas) em isopropanol 80% apresentaram uma tendência em tornarem-se mais lisas em relação às fibras não processadas, para os parâmetros Ra, Rz, Rzjis, Rp e Rq, já que os valores estatísticos estão muito próximos aos considerados significativos. Esse processamento pode ter auxiliado na reestruturação intermolecular, formando ou estabilizando as folhas. Ao estirar as fibras ocorreu uma compressão das moléculas, e resíduos de aminoácidos hidrofílicos podem ter ficado expostos à porção externa

criando forças momentâneas de van der Waals. Isso explicaria o aumento da força de atração ocorrida entre a amostra e a ponteira do microscópio durante o experimento de indentação. Sabe-se também que a água pode interagir com estruturas amorfas na superfície da fibra e diminuir a sua rugosidade (OLIVEIRA, 2012, TEULÉ, et al., 2012). Entretanto, os resultados diferem de outro trabalho, onde as fibras sintéticas contendo motivos de Flag e MaSp, após serem estiradas em isopropanol absoluto, reportaram a diminuição da força de atração amostra-ponteira em relação ao grupo controle e ao processamento com metanol durante análises de espectroscopia de força por MFA (MENEZES, et al., 2013). Os resultados adquiridos por espectroscopia de força das fibras sintéticas por MFA (Figura 42) apresentaram uma tendência no aumento do módulo de elasticidade para o grupo de fibras processadas em isopropanol (hidratadas e estiradas) em relação às fibras não processadas. Sabe-se que o tratamento com álcool proporciona uma reorganização molecular das proteínas, onde as fibras deixam o seu estado supercontrátil por terem sido expostas a esse solvente. Além disso, as forças aplicadas durante o estiramento podem ser capazes de alterar a elasticidade das fibras, devido a mudanças na microestrutura e do rearranjo das microfibras de material compactado natural ou sintético (SHAO; VOLLRATH, 1999; GOULD et al., 1999). Em outro estudo, as fibras sintéticas de Flag com sequência de polialanina e massa molecular de 57,6 kDa quando processadas em isopropanol obtiveram um aumento no módulo de elasticidade topográfico, o que não foi observado no processamento em metanol. Apesar das diferenças encontradas, o estudo caracterizou esse biomaterial como heterogêneo, onde as diferenças eram encontradas ao longo da mesma fibra e entre as fibras do mesmo grupo (MENEZES et al., 2013).

Ainda se tratando de análises de espectroscopia de força por MFA, a força de atração foi maior para as fibras processadas em relação a fibras não processadas ($p < 0,05$) (figura 42). Essa força é calculada durante a aproximação da ponteira à superfície da amostra antes do contato, quando os valores superam a constante elástica do cantilever, devido à forças eletrostáticas ou de van der Waals. Vale ressaltar que tanto as proteínas liofilizadas quanto as fibras sofriam interações eletrostáticas do ambiente durante a manipulação desse biomaterial. O tratamento dessas fibras em solução de isopropanol 80% podem ter reestruturado os resíduos de aminoácidos deixando os hidrofílicos expostos .

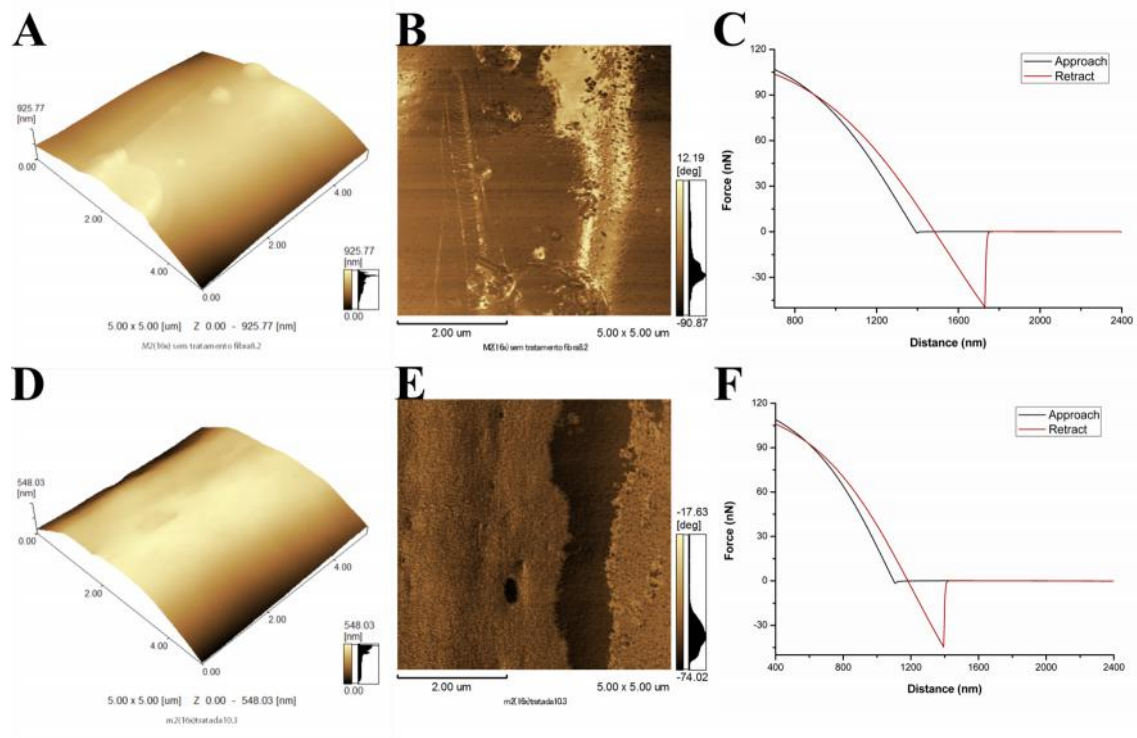


Figura 47. Imagens adquiridas por MFA de fibras sintéticas produzidas a partir da proteína MaSp2 16× (54 kDa) de *Parawixia bistrigata*. A área de varredura foi de $25 \mu\text{m}^2$. A-C: Análises de fibras sintéticas não tratadas após a extrusão; D-F: Análises de fibra sintética tratadas a pós a extrusão. A e D: digitalizações topográficas de altura em 3D; B e E: digitalizações de viscoelasticidade das mesmas áreas avaliadas em A e D; C e F: Gráficos representativos da curva de força.

6.12 Análises por Microscopia Eletrônica de Varredura (MEV)

Após o processo de purificação e diálise, as espidroínas MaSp2 16× e MaSp2 32× foram liofilizadas, a fim de verificar as suas respectivas morfologias. Para isso, uma parte das amostras foi submetida à MEV, e as imagens podem ser observadas na figura 48.

Tanto as espidroínas MaSp2 16× quanto MaSp2 32× apresentam estruturas globulares e de microfibras. Foi possível observar uma microfibra de MaSp2 16× com 2,95 μm de diâmetro (Fig. 48.C) e de MaSp2 32× com 1,58 μm de diâmetro (Fig. 48.F). Além disso, algumas torções também foram observadas nas microfibras supercontráteis (setas figura 48.C e F), ocorrendo com mais frequência nas espidroínas liofilizadas MaSp2 32×.Essas torções podem ser organizações precoce das espidroínas para a formação da fibra, e que foram formadas durante o processo de liofilização.

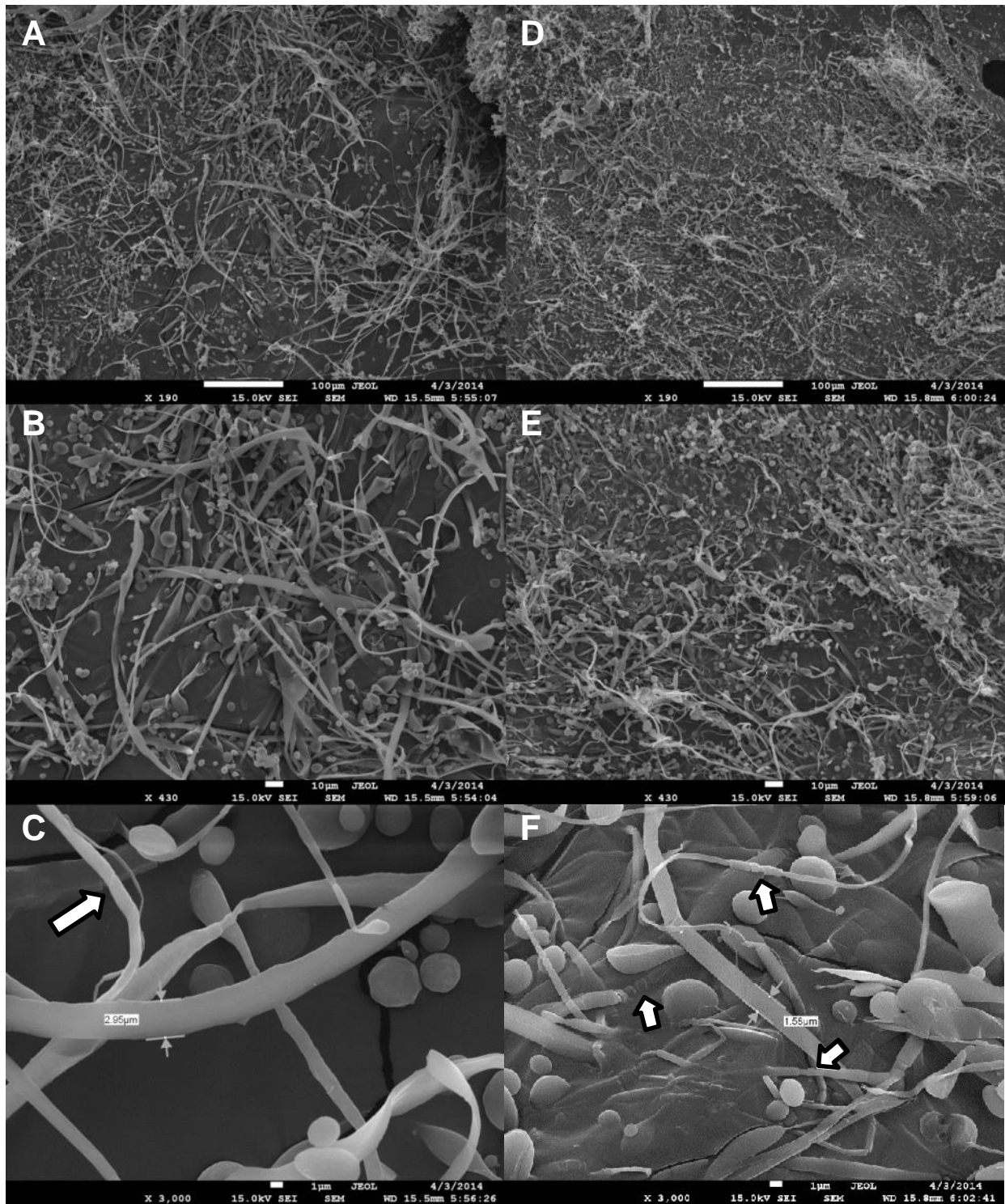


Figura 48. Fotomicrografia de espiroínas da aranha *P. bistriata* após liofilização, adquirida em microscópio eletrônico de varredura. A,B,C- Proteína MaSp2 16× (54 kDa) e D,E,F) Proteína MaSp2 32× (105 kDa) com aumento de 190×, 430× e 3.000× respectivamente. Barras em A e D: 100 μ; B e E: 10 μm; C e F: 1 μm. As setas indicam torções nas estruturas de microfibras proteicas.

As espidroínas MaSp2 16× e 32× liofilizadas foram caracterizadas com estruturas globulares e com formação de nano e microfibras com regiões supercontráteis (Figura 48). Esses resultados confirmam os dois modelos propostos para a formação de fibras: 1) teoria do cristal líquido, onde as estruturas alongadas organizam-se paralelamente formando filamentos maiores por atração física de folhas- ; 2) teoria de micela, onde os filamentos de moléculas anfífilicas organizam-se formando estruturas globulares similares a micelas, as quais unem-se posteriormente para formar uma estrutura maior, como as fibras, por exemplo (revisto por HEIM et al., 2009). Além disso, por meio de MEV e uma extrusão “clássica” é possível verificar as duas ocorrências para a organização e formação das fibras constituídas pela espidroína de 54 kDa (SILVA, 2014).

As fibras sintéticas formadas, a partir da MaSp2 16× (54 kDa), foram avaliadas morfológicamente quanto à estrutura de superfície e os diâmetros calculados. As fibras após a extrusão (Fig. 49.A) e as que foram processadas (Fig. 49.B) apresentaram a superfície lisa. Aumentos de campo em 30.000× no MEV realizados confirmaram a superfície lisa das fibras (dados não mostrados). Os valores dos diâmetros calculados para cada fibra (n=3) e para cada grupo (n=10) encontram-se na figura 49 (Fig. 49.D, E e C), existindo uma diminuição para a média dos diâmetros das fibras processadas.

Nas fotomicrografia de MEV, observou-se uma quantidade maior de filamentos e filamentos supercontráteis na amostra de MaSp2 32×, podendo ser um dos motivos da dificuldade em solubilizar essa proteína. Sabe-se que o processo de liofilização influencia nas formações de estruturas secundárias de proteínas (FORATO et al., 1998), sendo um dos pontos críticos da metodologia para a produção de fibras sintéticas.

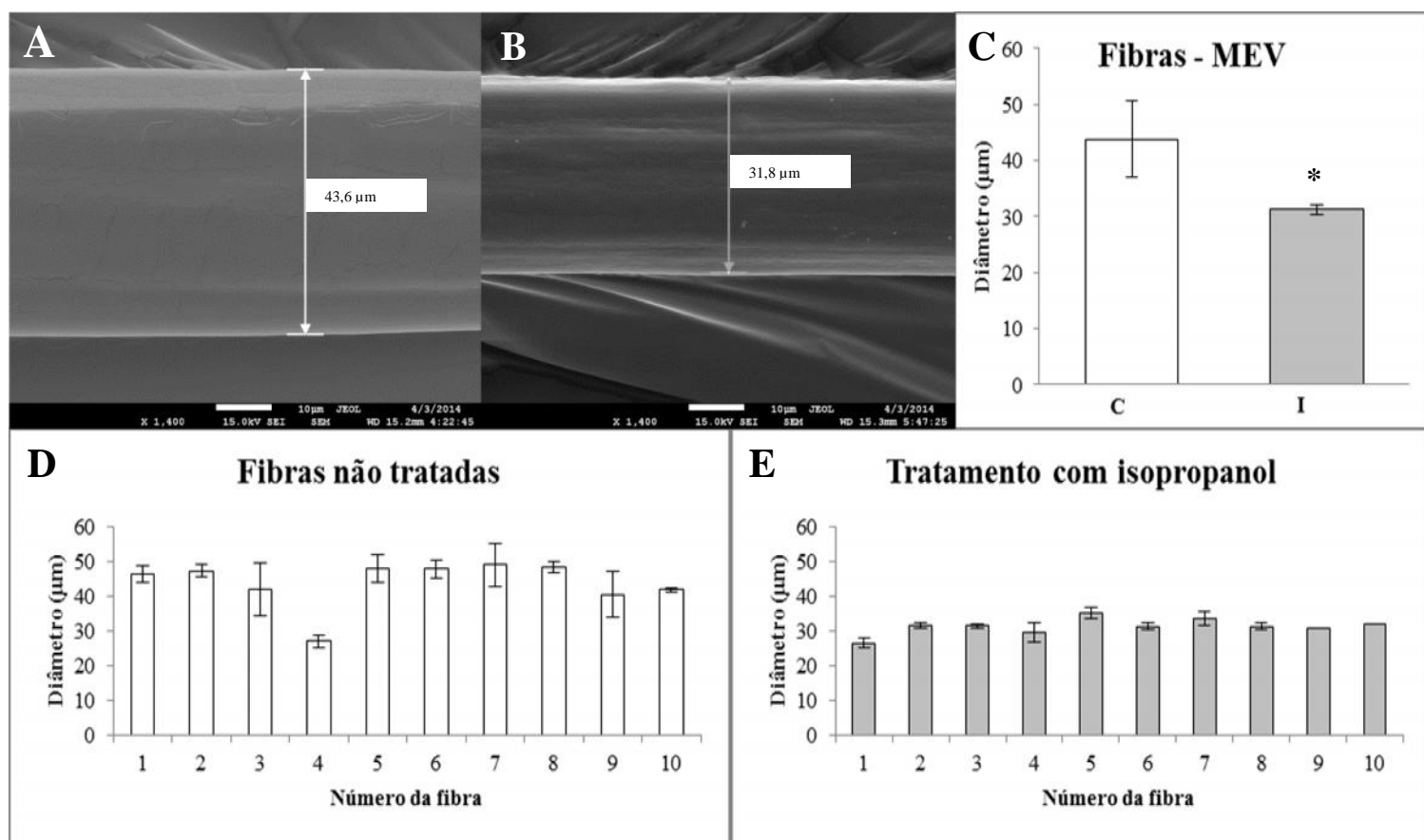


Figura 49. Análise das fibras produzidas a partir da proteína MaSp2 (54 kDa) de *Parawixia bistriata* por MEV. A) Fotomicrografia da fibra sintética após sua formação, com 43,6 μm de diâmetro; B) Fotomicrografia da fibra sintética após tratamento com isopropanol, com 31,8 μm de diâmetro; C) Gráfico representativo das médias de diâmetros dos dois grupos de fibras, $p < 0,05$ ($n=10$); D) Média dos diâmetros de cada fibra após a sua formação, e E) das fibras após o tratamento com isopropanol ($n=3$).

6.13 Testes mecânicos

As fibras sintéticas foram cortadas em fragmentos de 5 cm e separadas em cinco grupos: 1) Grupo não tratado, cujas fibras não sofreram processamentos ($n=12$); 2) Grupo hidratado, cujas fibras foram hidratadas em isopropanol ($n=20$); 3) Grupo hidratado e com camada de óleo mineral ($n=15$); 4) Grupo hidratado e esticado, cujas fibras foram estiradas ao dobro do seu tamanho inicial ($n=20$) e 5) Grupo hidratado, esticado e com camada de óleo

mineral (n=15). Esse material foi enviado para o colaborador Randy Lewis, Universidade de Utah, EUA, em placas de petri revestidas internamente com papel cartão preto (Figura 50).

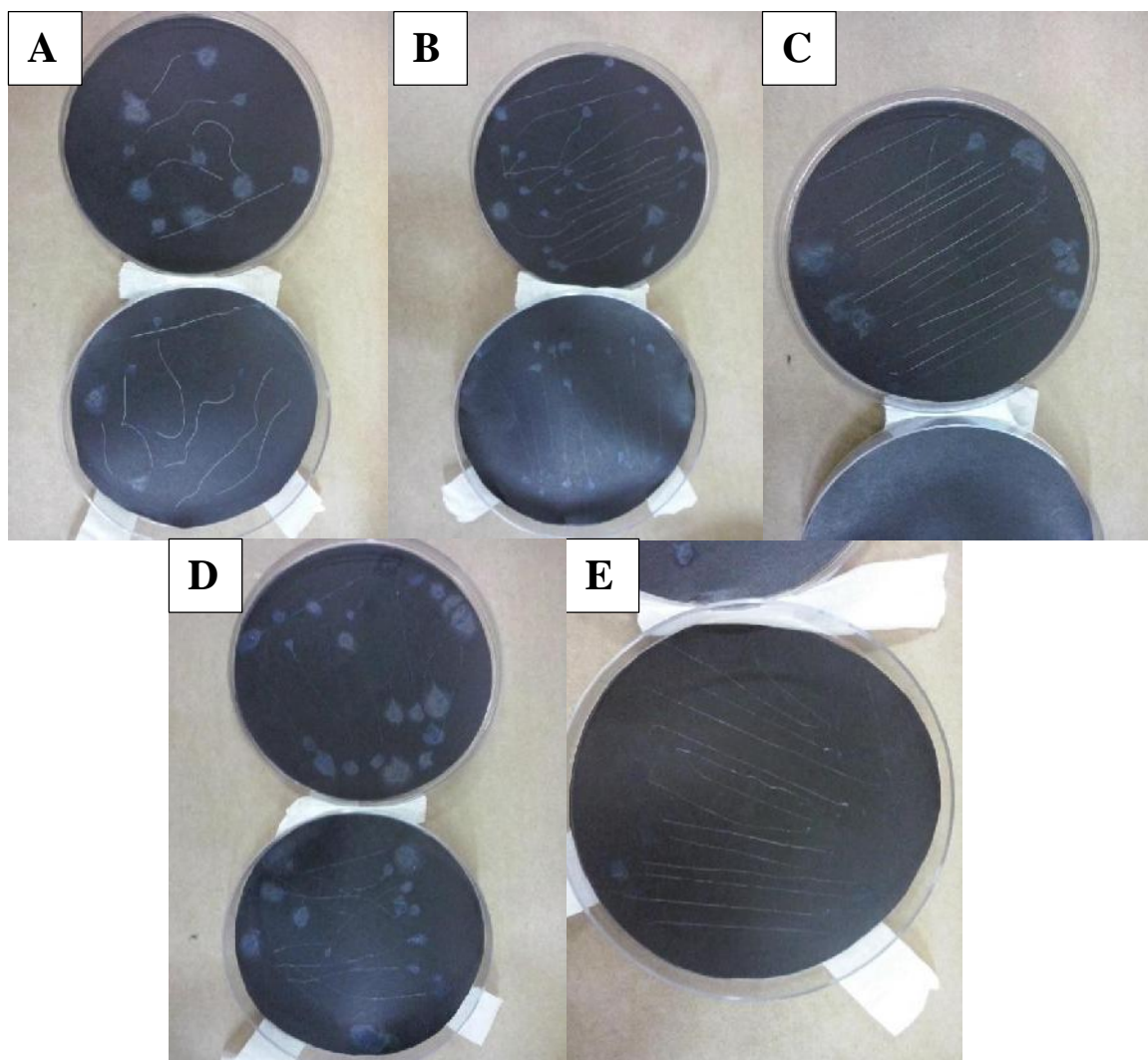


Figura 50. Fotografia das fibras sintéticas, armazenadas em papel cartão preto em placas de petri de acrílico. A) Não tratadas. B) Hidratadas; C) Hidratadas e com óleo mineral; D) Hidratadas e esticadas; E) Hidratadas, esticadas e com óleo mineral. Fotos: Arquivo pessoal.

As fibras sintéticas foram avaliadas por testes mecânicos. Para isso, os diâmetros foram mensurados por microscopia de luz e *software*, conforme já descrito, e os resultados encontram-se na figura 51. Os diâmetros dos dois grupos das fibras hidratadas e estiradas diminuíram em relação aos outros grupos ($p < 0,01$) (Figura 51.A). Comparando os grupos das fibras não tratadas com as hidratadas, uma diminuição do diâmetro médio das fibras também foi identificado ($p < 0,05$) (Figura 51.B).

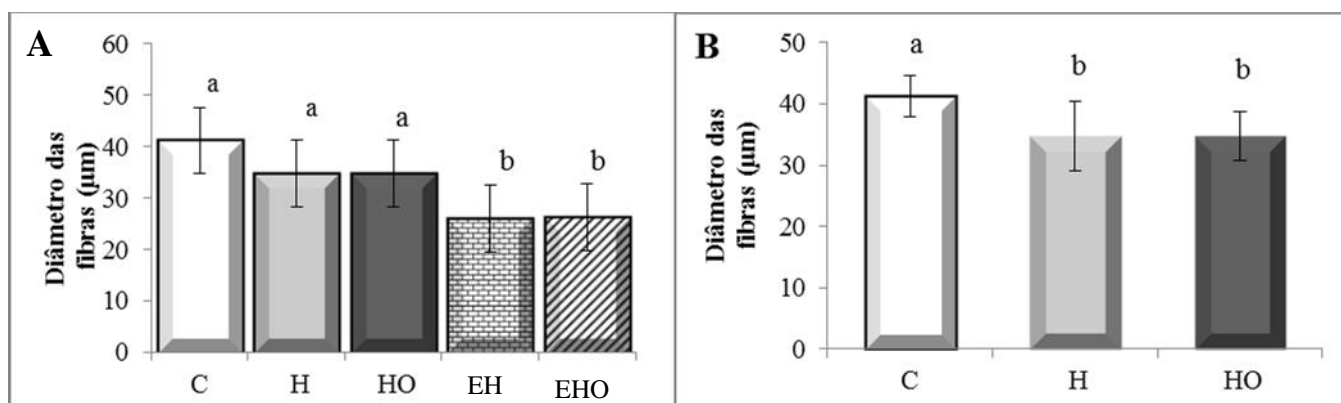


Figura 51. Gráficos representativos dos diâmetros médios das fibras sintéticas de MaSp2 16× (54 kDa). C: não tratadas; H: hidratadas; HO: hidratadas e com óleo; EH: hidratadas e esticadas; EHO: hidratadas, esticadas e óleo. A) Comparação entre os cinco grupos. Letras diferem entre si em $p < 0,01$). B) Comparação entre os grupos sem tratamento e com hidratação. Letras diferem entre si em $p < 0,05$.

Os testes mecânicos avaliados foram em energia de rompimento, na tensão máxima, na deformação máxima das fibras e módulo de elasticidade (módulo de Young). Os resultados foram analisados pelos testes de Tukey ou Kruskal-Wallis e os processamentos de hidratação, hidratação e estiramento e/ou adição de uma camada de óleo mineral na fibra sintética não diferiram da fibra sem tratamento (Figura 52). Os gráficos contendo os resultados do módulo de elasticidade de cada grupo encontram-se na figura 51. Os processamentos das fibras parecem não alterar a elasticidade do biomaterial (Figura 52). Esperava-se que as fibras MaSp2 fossem mais estáveis e apresentassem uma estrutura mais elástica do que resistente, já que essas espidroínas apresentam essa propriedade na natureza (Revisto por TOKAREVA et al., 2013).

O módulo de elasticidade é um parâmetro físico que proporciona a medida de rigidez de um material sólido e depende da energia de ligação entre os átomos do material, classificando-os como flexíveis e rígidos. Quanto maior o valor do módulo de elasticidade de um material, menor são as chances dele sofrer alteração e deformação em resposta a uma força aplicada, sendo considerado um material rígido (HECK, disponível em: <http://www.ct.ufrgs.br/ntcm/graduacao/ENG06638/IEM-Texto-4.pdf>).

Após a formação das fibras, essas foram analisadas morfológicamente por MEV após as análises em MFA. O diâmetro das fibras variou ao longo da mesma fibra e entre as fibras do mesmo grupo. Apesar disso, a variação desses diâmetros foi menor no grupo de fibras processadas, o qual foi também caracterizado por apresentar diâmetro médio significativamente menor em relação às fibras não processadas ($p < 0,01$)(Figura 46).

Estima-se que a hidratação e o estiramento realmente permitam uma finalização das interações moleculares resultando numa fibra mais homogeneia (AN et al., 2011; OLIVEIRA, 2012; TEULÉ et al., 2012) apesar dos dois grupos serem caracterizados morfologicamente como estruturas lisas.

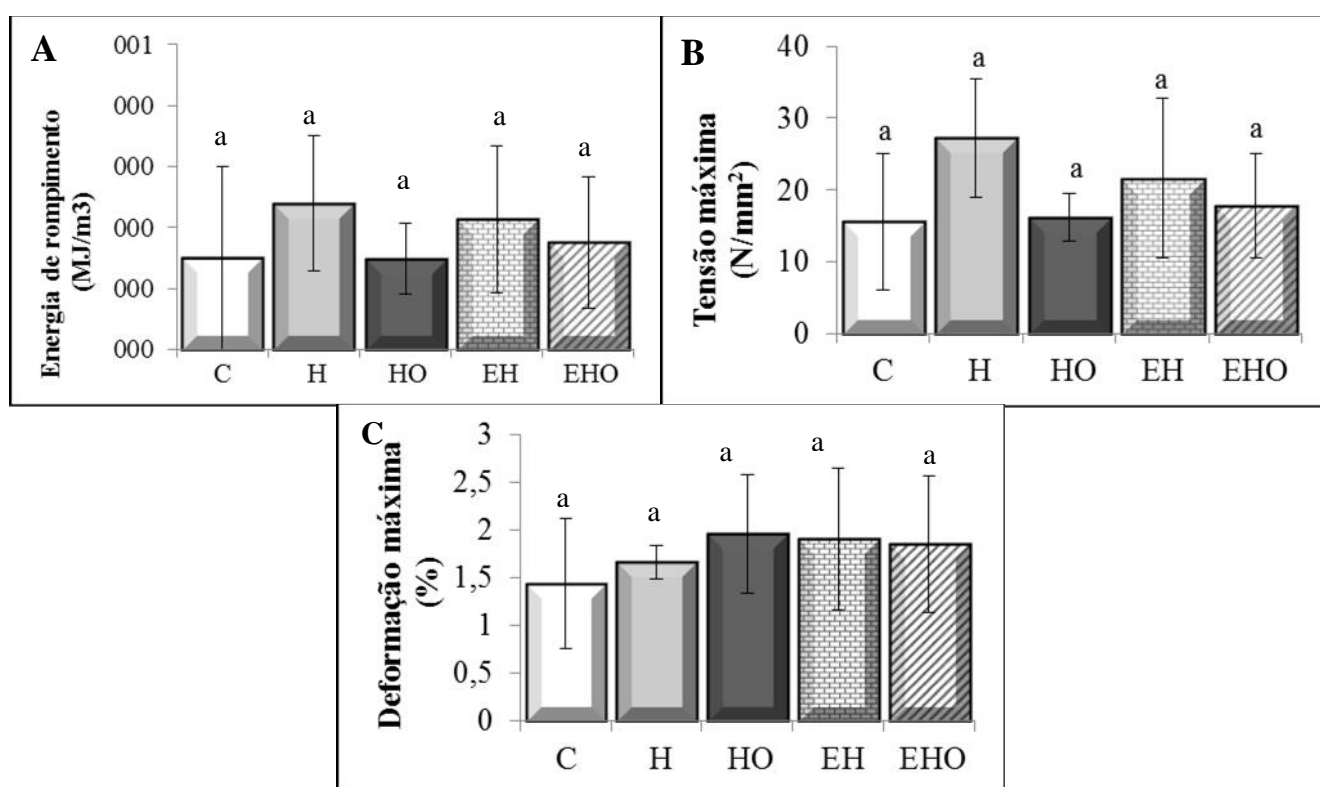


Figura 52. Gráficos representativos dos testes mecânicos realizados nos cinco grupos de fibras sintéticas MaSp2 16x (54 kDa). A) Energia de rompimento, B) Tensão máxima, C) Deformação máxima. Os grupos das fibras foram denominados como C: não tratadas; H: hidratadas; HO: hidratadas e com óleo; EH: hidratadas e esticadas; EHO: hidratadas, esticadas e óleo.

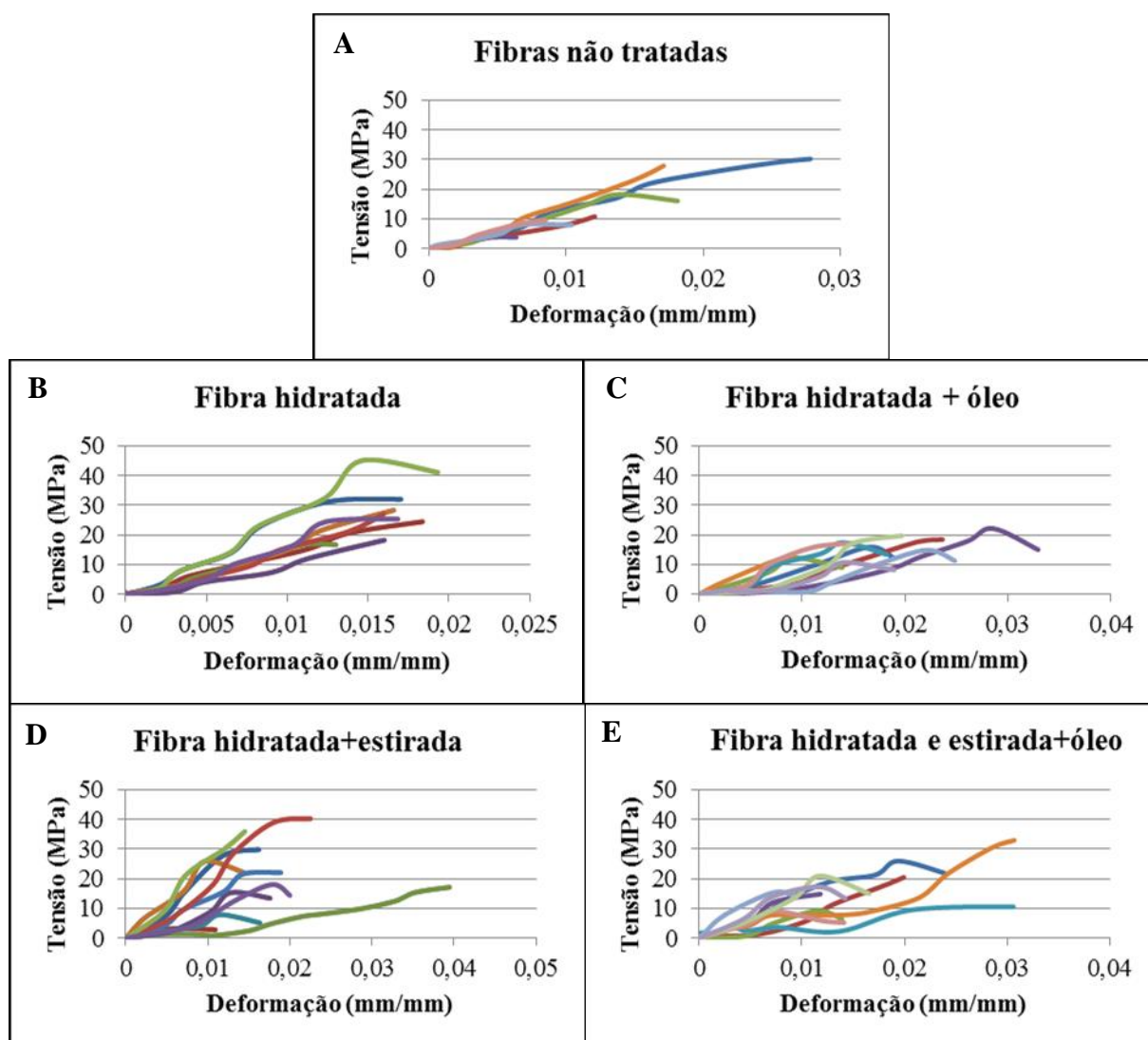


Figura 53. Gráficos representativos de tensão máxima vs deformação máxima antes do rompimento de fibras sintéticas de espidroínas MaSp2 16× de *Parawixia bistrriata*. Cada linha representa o resultado de uma fibra. Os grupos de fibras analisadas foram :A) Não tratadas (n=7), B) Hidratada em isopropanol 80% (n=8), C) Hidratada em isopropanol 80% e com uma camada de óleo mineral (n=10), D) Hidratada em isopropanol 80% e estirada (n=10), E) Hidratada em isopropanol 80%, estirada e com uma camada de óleo mineral (n=10).

De acordo com Rech e colaboradores (2010), as fibras naturais de *Parawixia bistrriata* foram caracterizadas com diâmetros de $7,56 \pm 1,21 \mu\text{m}$, extensibilidade de $21,54 \pm 5,01 \%$, força tensora $0,76 \pm 0,14 \text{ GPa}$ e rigidez de $0,069 \pm 0,024 \text{ GPa}$. Desse mesmo trabalho, sequências semelhantes às proteínas aciniformes, Flag, MiSp, MaSp1, MaSp2 entre outras espidroínas de glândulas foram identificadas e as sequências depositadas no GenBank (BENSON et al, 1999). As fibras sintéticas possuem diâmetros maiores, mas forças tensoras

menores, fazendo-se necessário uma melhoria no processo de extrusão das fibras e processamentos pós-formação, a fim de melhorar a qualidade das fibras sintéticas. Resultados semelhantes foram encontrados na literatura (GEURTS et al., 2010; XIA, et al., 2010; GNESA et al., 2012; AN, et al., 2011; TEULÉ, et al., 2012;).

Na natureza as fibras de linha de segurança são compostas por MaSp1 e MaSp2, ocorrendo variações no percentual dessas proteínas. Essa fibra da espécie *Argiope aurantia* foi caracterizada com 41% de MaSp1 e 59% de MaSp2, a de *Nephila clavipes* com 81% de MaSp1 e 19% de MaSp2, enquanto a *Parawixia bistrata* contém 90% de MaSp1 e 10% de MaSp2 (Revisto por HU et al., 2006; AN et al., 2011; OLIVEIRA, 2012). Sabe-se que as sedas da glândula ampolada maior são formadas por proteínas com 250-320 kDa (SPONNER et al., 2005; AYOUB; HAYASHI, 2007) e que fibras sintéticas constituídas de proteínas de alta massa molecular possuem uma qualidade maior (XIA et al., 2010, AN et al., 2011).

6.14 Espectroscopia por infravermelho (FTIR-ATR)

As espidroínas MaSp2 16× (54 kDa) e MaSp2 32× (105 kDa), e os grupos das fibras não tratadas e processadas (hidratadas e estiradas) foram analisadas por espectroscopia de infravermelho com reflexão total atenuada (ATR-FTIR). As identificações de folhas- e volta- são visualizadas pela absorção do sinal no comprimento de onda em torno de 1670 cm^{-1} e 1685 cm^{-1} conforme as identificadas no gráfico da figura 54. As espidroínas liofilizadas MaSp2 16× e MaSp2 32× e as fibras processadas apresentaram absorções semelhantes, enquanto as fibras não tratadas apresentaram um perfil de transmitância diferente das fibras tratadas e das proteínas liofilizadas (Figura 54).

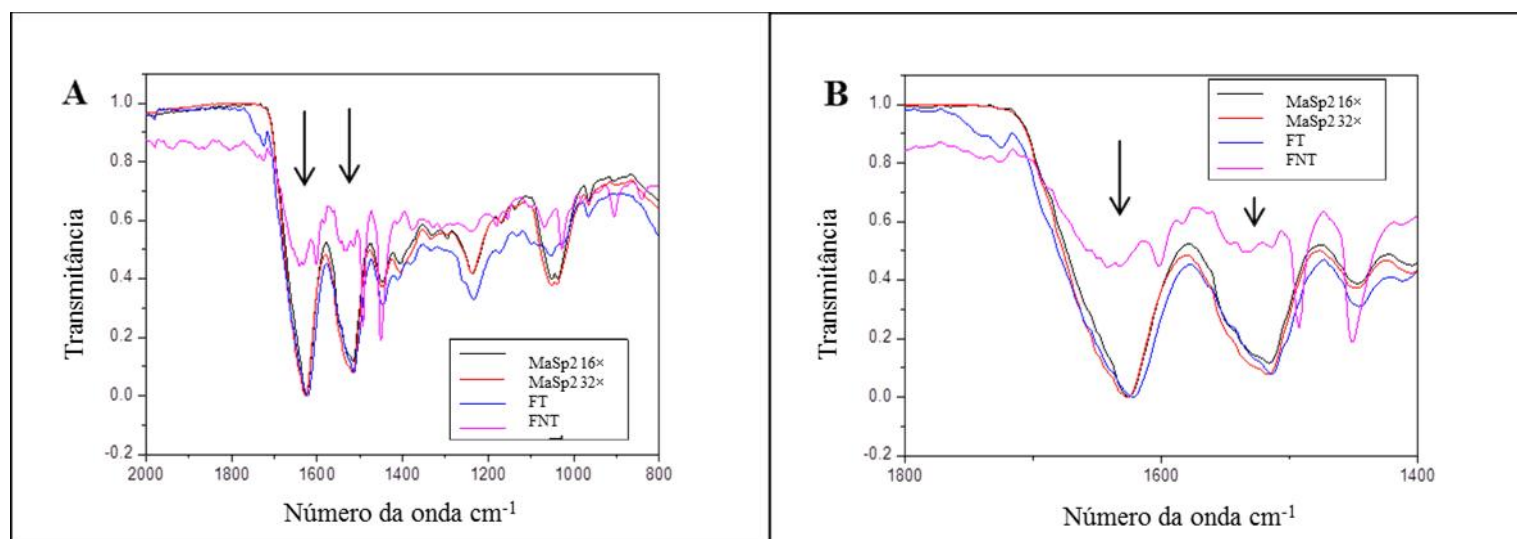


Figura 54. Gráfico representativo das análises por ATR-FTIR (transmitância vs nº da onda). As fibras sintéticas são de MaSp2 16× (54 kDa). FT: fibras tratadas. FNT: fibras não tratadas. A) Resultados com amplo campo das análises. B) Campo aproximado onde é possível uma melhor visualização da absorção referente a folhas- (1670 cm^{-1}). As setas indicam padrões de amida I ($1600\text{-}1700\text{ cm}^{-1}$) e amida II ($1600\text{-}1500\text{ cm}^{-1}$). Linha preta: espidroína liofilizada MaSp2 16×. Linha vermelha: espidroína liofilizada MaSp2 32×. Linha azul: fibras tratadas. Linha rosa: fibras não tratadas.

A FTIR é uma técnica recomendada para estudar estruturas secundárias de proteínas insolúveis em água (FORATO et al., 1998). As análises realizadas identificaram espectros para amida I ($1600\text{ - }1700\text{ cm}^{-1}$) e amida II ($1600\text{-}1500\text{ cm}^{-1}$), que se referem às ligações duplas C=O e simples N-H, corroborando com resultados de sedas de aranha (OLIVEIRA, 2012; SILVA, 2014; JIANG et al., 2014). As estruturas de folha beta (1506 e 1517 cm^{-1}) são encontradas na decomposição da amida II de fibras de teia de aranha (PAQUET-MERCIER et al., 2013). A literatura atribui à banda de amida I sendo compostas pela absorção em α -hélice (1656 cm^{-1}), folhas- (1624 , 1627 , 1632 , 1638 e 1642 cm^{-1}) e volta- (1666 , 1672 , 1680

e 1.688 cm^{-1}) (BYLER; SUSI, 1986; FORATO, 1998). Já as análises em fibras naturais de *Parawixia bistrinata* quantificaram 17% α -hélice, 47% de folhas- e 32% estruturas em espiral (RECH et al., 2010). Infelizmente até o momento da defesa, esses resultados não haviam sido quantificados. No entanto, a presença de estruturas de folhas- nas fibras foi confirmada, componente importante para a estabilidade da fibra (LAZARIS, et al., 2002).

A fim de obter fibras sintéticas com propriedades semelhantes às naturais, faz-se necessário a obtenção de proteínas de massa molecular maiores, inserindo ou alternando as sequências dos motivos na mesma proporção que elas ocorrem na natureza.

7. CONCLUSÃO

Conclui-se que as linhagens de bactéria modificadas por engenharia metabólica aumentaram o rendimento na produção das espidroínas sintéticas de alta massa molecular de *Parawixia bistrriata*. Isso indica que o limite de recursos para a produção de proteínas heterólogas pelo metabolismo das bactérias convencionais pode ser superado. Adicionalmente, as sequências das espidroínas MaSp2 de *Parawixia bistrriata* foram produzidas, caracterizadas e purificadas com sucesso. A espidroína MaSp2 16×(54 kDa) formou longas fibras sintéticas contínuas e com morfologia lisa nas condições apresentadas, sendo identificadas estruturas de folhas- importantes na sua constituição. A inédita espidroína MaSp2 32× (105 kDa) demanda a aplicação de novas metodologias para torna-la solúvel e novas tentativas de produzir fibras e outros biomateriais. Além disso, o processamento de hidratação e estiramento, realizado nas fibras sintéticas após a extrusão, são etapas importantes com tendência em melhorar qualitativamente as propriedades mecânicas desse biomaterial.

8. REFERÊNCIAS BIBLIOGRÁFICAS

AN, B.; HINMAN, M.B.; HOLLAND, G.P.; YARGER, J.L.; LEWIS, R.V. Inducing β -sheets formation in synthetic spider silk fibers by aqueous post-spin stretching. **Biomacromolecules**, v. 12(6), p. 2375-2382, 2011.

AN, B., JENKINS, J.E.; SAMPATH, S.; HOLLAND, G.P.; HINMAN, M.; YARGER, J.L.; LEWIS, R. Reproducing Natural Spider Silks' Copolymer Behavior in Synthetic Silk Mimics. **Biomacromolecules**, v. 13(12), p. 3938-3948, 2012.

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**, v. 2(6), e514, 2007, doi:10.1371/journal.pone.0000514.

BARR, L.A.; FAHNESTOCK, S.R.; YANG, J. Production and purification of recombinant DP1B silk-like protein in plants. **Molecular Breeding**, v. 13, p. 345-356, 2004.

BLACKLEDGE, T.A.; SWINDEMAN, J.E.; HAYASHI, C.Y. Quasistatic and continuous dynamics characterization of the mechanical properties of silk from the cobweb of the black widow spider *Latrodectus Hesperus*. **Journal of Experimental Biology**, v. 208(10), p.1937-1949, 2005.

BITTENCOURT, D.M. de C. **Caracterização molecular de proteínas de sedas de aranhas da biodiversidade brasileira**. 2007. 138 f. Tese (Doutorado em Biologia Molecular). Programa de Pós-Graduação "Stricto Sensu" em Biologia Molecular, Universidade de Brasília-UnB, Brasília, 138 p. 2007.

BROOK, A.E.; STEINKRAUS, H.B.; LEWIS, R.V. An investigation of the divergence of major ampullate silk fibers from *Nephila clavipes* and *Argiope aurantia*. **Biomacromolecules**, v. 6(6), p. 3095-3099, 2005.

BROOK, A.E.; BROTHERS, T.J.; CREAGER, M.S.; LEWIS, R.V. A novel methodology to explore the viscoelasticity of spider of spider major ampullate silk. **Journal of Applied Biomaterials & Biomechanics**, v. 5(3), p. 158-165, 2007.

BROOK, A.E.; NELSON, S.R.; JONES, J.A.; KOENIG, C.; HINMAN, M.; STRICKER, S.; LEWIS, R.V. Distinct contribution of MaSp1 and MaSp2 to mechanical properties of major ampullate silk fiber. **Nanotechnology, Science and Applications**, v. 1, p. 09-16, 2008.

CANDIANO, G.; BRUSCHI, M.; MUSANTE, L.; SANTUCCI, L.; GHIGGERI, G.M.; CARNEMOLLA, B.; ORECCHIA, P.; ZARDI, L.; RIGHETTI, P.G. Blue silver: A very

sensite coloidal Coomassie G-250 staining for proteome analysis. **Electrophoresis**, v. 25, p. 1327-1333, 2004.

CHEN Z.L.; SCHULER M.A.; BEACHY R.N. Functional analysis of regulatory elements in a plant embryo specific gene. **Proceeding of the National Academy of Science**, v. 83, p. 8560–8564, 1986.

CONG, L.; et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. **Science**, v. 339(6121), p. 819-82, 2013.

CUNHA, N.B.; MURAD, A.M.; RAMOS, G.L.; MARANHÃO, A.Q.; BRÍGIDO, M.M.; ARAÚJO, A.C.G.; LACORTE, C.; ARAGÃO, F.J.L.; COVAS, D.T.; FONTES, A.M.; SOUZA, G.H.M.F.; VIANNA, G.R.; RECH, E.L. Accumulation of functional recombinant human coagulation factor IX in transgenic soybean seeds. **Transgenic Research**, v. 20, p. 841-855, 2011a.

CUNHA, N.B.; MURAD, A.M. ; CIPRIANO, T.M. ; ARAÚJO, A.C.G.; ARAGÃO, F.J.L.; LEITE, A.; VIANNA, G.R.; MCPHEE, T.R.; SOUZA, G.H.M.F.; WATERS, M.J.; RECH, E.L. Expression of functional recombinant human growth hormone in transgenic soybean seeds. **Transgenic Research**, v. 20, p. 811-826, 2011b.

DAMS-KOZLOWSKA, H.; MAJER, A.; TOMASIEWICK, P.; LORINSKA, J. KAPLAN, D.L.; MACKIEWICZ, A. Purification and cytotoxicity of tag-free bioengineered spider silk proteins. **Journal of Biomedical Materials Research**, v. 101A, p. 456-464, 2012.

DENNY, M.W. Silks - their properties and functions. **In: Vincent. JFV, Currey, JD (Eds). Mechanical properties of Biological Materials.** Cambridge University Press, Cambridge, p. 247-272, 1980.

DICKO, C.; et al. Secondary structures and conformational changes in flagelliform, cylindrical, major, and minor ampullate silk proteins. Temperature and concentration effects. **Biomacromolecules**, v. 5, p. 2105-2115, 2004.

EISOLDT, L.; THAMM, C.; SCHEIBEL, T. The role of terminal domains during storage and assembly of spider silk proteins. Review. **Biopolymers**, v. 97(6), p. 355-361, 2011.

FAHNESTOCK, S.R.; BEDZYK, L.A. Synthetic spider dragline silk proteins and their production in *Escherichia coli*. **Applied microbiology and biotechnology**, v. 47, p. 23-32, 1997.

FORATO, L.A.; FILHO, R.B.; COLNAGO, L.A. Estudo de métodos de aumento de resolução de espectros de FTIR para análise de estruturas secundárias de proteínas. **Química Nova**, v. 20(5), p. 1-5, 1997.

GEURTS, P.; ZHAO, L. HSIA, Y.; GNESE, E.; TANG, S.; JEFFERY, F.; MATTINA, C.L.; FRANZ, A.; LARKIN, L.; VIERRA, C. Synthetic Spider Silk Fibers Spun from Pyriform Spidroin 2, A Glue Silk Protein Discovered in Orb-Weaving Spider Attachment Discs. *Biomacromolecules*, v. 11(12), p. 3495-3503, 2012.

GNESE, E.; HSIA, Y. YAGER, J.L.; WEBER, W.; LIN-CEREGHINO, J.; LINCEREGHINO, G. TANG, S.; AGARI, K.; VIERRA, C. Conserved CTerminalDomain of Spider Tubuliform Spidroin 1 Contributes to Extensibility inSynthetic Fibers. *Biomacromolecules*, v.13(2), p. 304-312, 2012.

GOSLINE, J.M.; GUERETTE, P.A.; ORTLEPP, C.S.; SAVAGE, K.N. The mechanical design of spider silk: From fibroin sequence to mechanical function. **The Journal of Experimental Biology**, v. 202, p. 3295-3303, 1999.

GOSLINE, J.M.; et al. Elastic proteins: biological roles and mechanical properties. **Philosophical Transactions of Royal Society Lond. B Biological Sciences**, v. 357, p. 121-132, 2002.

GUERETTE, P.A., *et al...* Silk properties determined by gland-specific expression of a spider fibroin gene family. **Science**, v. 272, p. 112-115, 1996.

GOUL, S.A.C.; TRAN, K.T.; SPANGA, J.C.; MOORE, A.M.F.; SHULMAN, J.B. Short and long range order of the morphology of silk from *Latrodectus Hesperus* (Black Widow) as characterized by atomic force microscopy. **International Journal of Biological Macromolecules**, v. 24(2-3), p. 151-157, 1999.

HANG, F.; et al. A conserved spider silk domain acts as a molecular switch that controls fibre assembly. **Nature**. v. 465, p. 239-242, 2010.

HANAHAN, D.; JESSE, J.; BLOOM, F.R. Plasmid transformation of Escherichia coli and other bacteria. **Methods in Enzymology**, v. 204, p. 63-113, 1983.

HAUPTMANN, V.; et al. Native-sized spider silk proteins synthesized in planta intein-based multimerization. **Transgenic Research**. v. 22(2), p. 369-377, 2013.

HAYASHI, C.Y.; LEWIS, R.V. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. **Journal of Molecular Biology**, v. 275, p. 773-784, 1998.

HAYASHI, C.Y.; et al. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. **International Journal of Biological Macromolecules**, v. 24, p. 271-275, 1999.

HECK, N.C. Introdução à engenharia metalúrgica, UFRGS. Disponível em: <http://www.ct.ufrgs.br/ntcm/graduacao/ENG06638/IEM-Texto-4.pdf>. Acesso em: 04 maio 2015.

HEDHAMMAR, M.; et al. Structural properties of recombinant nonrepetitive and repetitive parts of major ampullate spidroin 1 from *Euprosthops australis*: implications for fiber formation. **Biochemistry**, v. 18(47), p. 3407-17, 2008.

HINMAN, M.B.; JONES, J.A.; LEWIS, R.V. Synthetic spider silk: a modular fiber. **Trends in Biotechnology**, v. 18(9), p. 374-379, 2000.

HOLLAND, G.P.; et al. Solid-state NMR investigation of major and minor ampullate spider silk in the native and hydrated states. **Biomacromolecules**, v. 9, p. 651-657, 2008.

HUEMMERICH, D., et al. Primary structure elements of spider dragline silk and their contribution to protein solubility. **Biochemistry**, v. 43, p. 395-411, 2004.

ITTAH, S.; et al. An essential role for the C-terminal domain of a dragline spider silk protein in directing fiber formation. **Biomacromolecules**, v. 7, p. 1790-1795, 2006.

JIANG, P.; MARÍ-BUYÉ, N.; MADURGA, R.; ARROYO-HERNÁNDEZ, M.; SOLANAS, C.; GAÑÁN, A.; DAZA, R.; PLAZA, G.R.; GUINEA, G.V.; ELICES, M.; CENIS, J.L.; PÉREZ-RIGUEIRO. Spider silk gut: Development and characterization of a novel strong spider silk fiber. **Scientific Reports**, 2014, doi: 10.1038/srep07326.

KINAHAN, M.E.; et al. Tunable silk: using microfluidics to fabricate silk fibers with controllable properties. **Biomacromolecules**, v. 5(12), p. 1504-1511, 2011.

KUSNADI, A.R.; et al. Production of Recombinant proteins in transgenic plants: practical considerations. **Biotechnology and Bioengineering**, v. 56(5), p. 473-484, 1997.

LAZARIS, A.; et al. Spider silk fiber spun from soluble recombinant silk produced in mammalian cells. **Science**, v. 295, p. 472-476, 2002.

LEWIS, R.V. Spider silk: ancient ideas for new biomaterials. **Chemical Reviews**, v. 106, p. 3762-3774, 2006.

LOMBARDI, S.; KAPLAN, D. The amino acid composition of major ampullate gland silk (dragline) of *Nephila clavipes* (araneae, tetragnathidae). **Journal of Arachnology**, v. 18, p. 297-306, 1990.

MADSEN, B.; SHAO, Z.Z.; VOLLRATH, F. Variability in the mechanical properties of spider silk on three levels: interspecific, intraspecific and intraindividual. **International Journal of Biological Macromolecules**, v. 24 (2-3), p. 301-306, 1999.

MADSEN, B.; VOLLRATH, F. Mechanics and morphology of silk drawn from anesthetized spiders. **Naturwissenschaften**. v. 87(3), p. 148-153, 2000.

MARHABAIE, M.; LEEPER, T.C.; BLACKLEDGE, T.A. Protein composition correlates with the mechanical properties of spider (*Argiope trifasciata*) dragline silk. **Biomacromolecules**, v. 15(1), p. 20-29, 2014.

MENEZES, G.M.; TEULÉ, F.; LEWIS, R.V.; SILVA, L.P.; RECH, E.L. Nanoscale investigations of synthetic spider silk fibers modified by physical and chemical processes. **Polymer Journal**, v. 45, p. 997-1006, 2013.

MOTTA, P.C. **Aracnídeos do cerrado**. 1 ed., Rio de Janeiro Technical Books Editora, 2014. 209 p.

MURAD, A.M.; RECH, E.L. Molecular dynamics simulations of the minor ampullate spidroin modular amino acid sequence from *Parawixia bistrriata*: insights into silk tertiary structure and fibre formation. **Journal of Molecular Modeling**, doi:10.1007/s00894-010-0823-4.

MURAD, A.M. RECH, E.L. NanoUPLC-MS^E proteomic data assessment of soybean seeds using the Uniprot database. **BioMed Central Biotechnology**, v. 12(82), p.1-17, 2012.

OLIVEIRA, P.E.F. **Expressão e purificação MaSp2 isolada da glândula produtora de teia da aranha *Parawixia bistrriata***. 2008. 70 f. Dissertação (Mestrado em Ciências Genômicas e Biotecnologia). Programa de Pós-Graduação “Stricto Sensu” em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília-UCB, Brasília, 70p. 2008.

OLIVEIRA, P.E.F. **Produção e caracterização molecular de fibras sintéticas das aranhas *Parawixia bistrriata* e *Nephilengys cruentata***. 2012. 83 f. Tese (Doutorado em Biologia Molecular). Programa de Pós-Graduação “Stricto Sensu” em Biologia Molecular, Universidade de Brasília-UnB, Brasília, 83 p. 2012.

OMENETTO, F.G.; KAPLAN, D.L. New opportunities for an ancient material. Review, **Science**, v. 329, p. 528-531, 2010.

PALIY, O; GUNASEKERA, T.S. Growth of *E. coli* BL21 in minimal with different gluconeogenic carbon sources and salt contents. **Applied Microbial and Cell Physiology**, v. 73, p.1169-1172, 2007.

PAQUET-MERCIER, F.; LEFÈVRE, T.; AUGER, M.; PÉZOLT M. Evidence by infrared spectroscopy of the presence of two types of β -sheets in major ampullate spider silk and silkworm silk. **Journal of Materials Chemistry (Soft Matter)**, v. 9, p. 208-215, 2013.

PLATNICK, N.I. 2014. The world spider catalog, version 15. **American Museum of Natural History**, disponível on-line: http://research.amnh.org/iz/spiders/catalog_15.0/index.html.

PRESCOTT, Lansing M.; HARLEY, John, P.; KLEIN, Donald A. **Microbiology**. 6.ed. Boston, MA, Estados Unidos: McGraw-Hill, 992 p, 2005.

RIEKEL, C.; et al. X-ray diffraction on spider silk during controlled extrusion under a synchrotron radiation X-ray beam. **Biomacromolecules**, v. 1, p. 622-626, 2000.

RECH, E.L., et al. High-efficiency transformation by biobalistics of soybean, common bean and cotton transgenic plants. **Nature Protocols**, v. 3(3), p. 410-418, 2008.

RECH FILHO, E.L.; et al. 2010. Proteins from the webs of *Nephilengys cruentata*, *Avicularia juruensis* and *Parawixia bistrriata* spiders isolated from brazilian biodiversity. **US20100311645 A1** Estados Unidos, 2010.

RISING, A.; WIDHE, M.; JOHANSSON, J.; HEDHAMMAR, M. Spider silk proteins: recent advances in recombinant production, structure-function relationships and biomedical applications. **Review, Cellular and Molecular Life Science**, v. 68, p. 169-184, 2011.

ROMER, L.; SCHEIBEL, T. The elaborate structure of spider silk. Review. **Prion**, v. 2(4), p. 154-161, 2008.

SAMBROOK, J.; RUSSELL, D.W. **Molecular Cloning: A Laboratory Manual**. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2003.

SCHEIBEL, T. Spider silks: recombinant synthesis, assembly, spinning, and engineering of synthetic proteins. **Microbial Cell Factories**, v. 3(14), 2004, doi:[10.1186/1475-2859-3-14](https://doi.org/10.1186/1475-2859-3-14).

SCHELLER, J.; et al. Production of spider silk proteins in tobacco and potato. **Nature Biotechnology**, v. 19, p. 573-577, 2001.

SILVA, O.B. **Expressão e purificação de proteína de glândula sericígena da aranha *Parawixia bistrriata* em bactérias geneticamente modificadas**. 2014. 98p. Dissertação (Mestrado em Biologia Molecular). Programa de pós-graduação "Stricto Sensu" em Biologia Molecular. Universidade de Brasília- UnB, Brasília, 98p. 2014.

SILVA, L.P.; RECH, E.L. Unravelling the biodiversity of nanoscale signatures of spider silk fibres. **Nature communications**. v. 4(18), p. 1-9, 2013.

SLOTTA, U.; MOUGIN, N.; RÖMER, L.; LEIMER, A.H. Synthetic spider silk proteins and threads. **Society for Biological Engineering**, 43-49, 2012, disponível em: http://www.amsilk.com/fileadmin/user_upload/SyntheticSpiderSilk_May2012_CEP.pdf.

SMITH, D.M., et al. Analysis of the conserved N-terminal domains in major ampollate spider silk proteins. **Biomacromolecules**, v. 6, p. 3152-3159, 2005.

SNEDDON, I.N. the relation between load and penetration in the axisymmetric boussinesq problem for a punch of arbitrary profile. **International Journal of engineering Science**. v. 3(1), p. 47-57, 1965.

SOUTO, B. de MORAIS. **Expressão e purificação de proteínas de glândulas produtoras de seda das aranhas *Nephilengys cruentata* e *Avicularia juruensis***. 2008. 102p. Dissertação (Mestrado em Biologia Molecular). Programa de pós-graduação “Stricto Sensu” em Biologia Molecular. Universidade de Brasília- UnB, Brasília, 102p. 2008.

SPONNER, A.; VATER, W.; MONAJEMBASHI, S.; UNGER, E.; GROSSE, F.; WEISSHART. Composition and hierarchical organization of a spider silk. **Plos One**, v. 2 (10), e998, 2007.

STAUFFER, S.; et al. Comparison of physical properties of three silks from *Nephila clavipes* and *Araneus gemmoides*. **Journal of Arachnology**, v. 22, p. 5-11, 1994.

STEPHENSON, FRANK H. Calculations for molecular biology and biotechnology. A Guide to mathematics in the laboratory. **Academic Press**, v. 1, p. 302, 2003.

SWADSON, B.; et al. The evolution of complex biomaterial performance the case of spider silk. **Integrative and Comparative Biology**, v. 49, p. 21-31, 2009.

TATHAM, A.S.; SHEWRY, P.R. Elastomeric Proteins: Biological roles, Structures and Mechanisms. **Trends Biochemical Sciences**, v. 25, p. 567-571, 2000.

TEULÉ, F., et al. Modifications of spider silk sequences in an attempt to control the mechanical properties of the synthetic fibers. **Journal of Material Science Materials in Medicine**, v. 42, p. 8974-8985, 2007.

TEULÉ, F.; et al. A protocol for the production of recombinant spider silk-like proteins for artificial fiber spinning. **Nature Protocols**, v. 4(3), p. 341-355, 2009.

TEULÉ, F.; ADDISON, B.; COOPER, A.R., AVON, J.; HENNING, R.W.; BENMORE, C.J.; HOLLAND, G.P.; YARGER, J.L.; LEWIS, R.V. Combining flagelliform and dragline spider silk motifs to produce tunable synthetic biopolymer fibers. **Biopolymers**, v. 97(6), p.418-431, 2012.

TIAN, M., et al. R. Analysis of major ampullate silk cDNAs from two non-orb-weaving spiders. **Biomacromolecules**, v. 5, p. 657-660, 2004.

TOKAREVA, O.; MICHALCZECHEN-LACERDA, V.A.; RECH, E.L.; KAPLAN, D.L. Recombinant DNA production of spider silk proteins. **Microbial biotechnology**, v. 6(6), p. 651-663, 2013.

TUCKER, C.L.; et al. Mechanical and physical properties of recombinant spider silk films using organic and aqueous solvents. **Biomacromolecules**. v. 8(15), p. 3158-70, 2014.

VIANNA, G.R.; ARAGÃO, F.J.L.; RECH, E.L. A minimal DNA cassette as a vector for genetic transformation of soybean (*Glycine max*). **Genetics and Molecular Research**, v. 10, p. 382-390, 2011.

VIANNA, G.R.; CUNHA, N.B.; RECH, E.L. Expression and accumulation of heterologous molecules in the protein storage vacuoles of soybean seeds. **Nature Protocol Exchange**. 2011, doi:10.1038/protex.2011.206.

VOLLRATH, F.; KNIGHT, D.P. Liquid crystalline spinning of spider silk. **Nature**, v. 410, p. 541-548, 2001.

WOOD, A.; et al. Targeted Genome Editing Across Species Using ZFNs and TALENs. **Science**, v. 333(6040), p. 307, 2011.

WOLF, M.J.; ARAÚJO, D. Análise cromossômica de aranhas de teias orbiculares (Araneae, Orbiculariae) do parque nacional de ilha grande. **Anais do encontro de iniciação científica-ENIC**, Universidade Estadual de Mato Grosso do Sul, ISSN 2179-7447, n3, 2011.

YANG, J.; BARR, L.A.; FAHNESTOCK, S.R.; LIU, Z.B. High yield recombinant silk-like protein production in transgenic plants through protein targeting. **Transgenic Research**, v. 14(3), p. 313-324, 2005.

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. **Journal of Biochemistry**, v. 137, p. 721-729, 2005.

XIA, X.X.; et al. Native-sized recombinant spider silk protein produced in metabolically engineered *Escherichia coli* results in a strong fiber. **PNAS**, v. 107(32), p. 14059-14063, 2010.

XU, M.; Lewis, R. Structure of a protein superfiber: Spider dragline silk. **Biochemistry**, v. 87, p. 7120-7124, 1990.

ANEXO A – Teórica digestão da proteína MaSp2 16× com quimotripsina.

Nº	de>até	MH+	Tempo HPLC	pI	Sequência
1	1-29	3.346,38	22,29	6,68	MGHHHHHHHHHHSSGHIDDDDKHMPGGGY
2	30-52	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
3	53-68	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
4	69-91	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
5	92-107	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
6	108-130	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
7	131-146	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
8	147-169	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
9	170-185	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
10	186-208	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
11	209-224	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
12	225-247	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
13	248-263	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
14	264-286	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
15	287-302	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
16	303-325	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
17	326-341	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
18	342-364	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
19	365-380	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
20	381-403	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
21	404-419	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
22	420-442	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
23	443-458	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
24	459-481	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
25	482-497	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
26	498-520	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
27	521-536	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
28	537-559	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
29	560-575	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
30	576-598	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
31	599-614	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
32	615-637	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
33	638-650	942,46	9,18	5,70	GPGAAAAAAAAASG

ANEXO B – Teórica digestão da proteína MaSp2 32× digerida com quimotripsina.

Nº	de>até	MH+	Tempo HPLC	pI	Sequência
1	1-29	3.346,38	22,29	6,68	MGHHHHHHHHHHSSGHIDDDDKHMPGGGY
2	30-52	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
3	53-68	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
4	69-91	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
5	92-107	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
6	108-130	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
7	131-146	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
8	147-169	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
9	170-185	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
10	186-208	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
11	209-224	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
12	225-247	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
13	248-263	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
14	264-286	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
15	287-302	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
16	303-325	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
17	326-341	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
18	342-364	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
19	365-380	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
20	381-403	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
21	404-419	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
22	420-442	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
23	443-458	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
24	459-481	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
25	482-497	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
26	498-520	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
27	521-536	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
28	537-559	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
29	560-575	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
30	576-598	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
31	599-614	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
32	615-637	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
33	638-653	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
34	654-676	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
35	677-692	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
36	693-715	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
37	716-731	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
38	732-754	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
39	755-770	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
40	771-793	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
41	794-809	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
42	810-832	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
43	833-848	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
44	849-871	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
45	872-887	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
46	888-910	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
47	911-926	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
48	927-949	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
49	950-965	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
50	966-988	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
51	989-1004	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
52	1005-1027	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY

53	1028-1043	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
54	1044-1066	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
Nº	de>até	MH+	Tempo HPLC	pI	Sequência
55	1067-1082	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
56	1083-1105	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
57	1106-1121	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
58	1122-1144	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
59	1145-1160	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
60	1161-1183	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
61	1184-1199	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
62	1200-1222	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
63	1223-1238	1.219,57	13,57	5,69	GPGAAAAAAAAASGGGY
64	1239-1261	2.013,90	18,75	5,69	GPGGAGQQGPAAGQQGPGSQGSY
65	1262-1274	942,46	9,18	5,70	GPGAAAAAAAAASG

ANEXO C- Proteínas encontradas apenas na linhagem BL21(DE03)pACYC184

	Acesso	Descrição	Escore
1	YLCE_ECOLI	Putative uncharacterized protein ylcE OS=Escherichia coli (strain K12) GN=ylcE PE=5 SV=1	651.29
2	SOPB_ECOLI	Protein sopB OS=Escherichia coli (strain K12) GN=sopB PE=1 SV=1	223.41
3	ASTB_ECOLI	N-succinylarginine dihydrolase OS=Escherichia coli (strain K12) GN=astB PE=1 SV=1	146.49
4	DKGB_ECOLI	2,5-diketo-D-gluconic acid reductase B OS=Escherichia coli (strain K12) GN=dkgB PE=1 SV=2	189.4
5	FCL_ECOLI	GDP-L-fucose synthase OS=Escherichia coli (strain K12) GN=fcl PE=1 SV=2	227
6	HEMX_ECOLI	Putative uroporphyrinogen-III C-methyltransferase OS=Escherichia coli (strain K12) GN=hemX PE=1 SV=1	142.36
7	PYRH_ECOLI	Uridylate kinase OS=Escherichia coli (strain K12) GN=pyrH PE=1 SV=2	214.83
8	SPEA_ECOLI	Biosynthetic arginine decarboxylase OS=Escherichia coli (strain K12) GN=speA PE=1 SV=2	173.88
9	YFFB_ECOLI	Protein Yffb OS=Escherichia coli (strain K12) GN=yffB PE=1 SV=1	387.67
10	YIAF_ECOLI	Uncharacterized protein yiaF OS=Escherichia coli (strain K12) GN=yiaF PE=4 SV=2	371.37
11	KCY_ECOLI	Cytidylate kinase OS=Escherichia coli (strain K12) GN=cmk PE=1 SV=1	436.35
12	AROC_ECOLI	Chorismate synthase OS=Escherichia coli (strain K12) GN=aroC PE=1 SV=4	178.6
13	RTCA_ECOLI	RNA 3'-terminal phosphate cyclase OS=Escherichia coli (strain K12) GN=rtcA PE=1 SV=3	262.88
14	MIAB_ECOLI	(Dimethylallyl)adenosine tRNA methylthiotransferase miaB OS=Escherichia coli (strain K12) GN=miaB PE=1 SV=1	172.81
15	DCEB_ECOLI	Glutamate decarboxylase beta OS=Escherichia coli (strain K12) GN=gadB PE=1 SV=1	133.08
16	UBID_ECOLI	3-octaprenyl-4-hydroxybenzoate carboxy-lyase OS=Escherichia coli (strain K12) GN=ubiD PE=1 SV=1	153.02
17	B8YM76_ECOLI	Alkaline phosphatase (Fragment) OS=Escherichia coli (strain K12) PE=3 SV=1	221.92
18	CYSH_ECOLI	Phosphoadenosine phosphosulfate reductase OS=Escherichia coli (strain K12) GN=cysH PE=1 SV=3	168.08
19	RIMM_ECOLI	Ribosome maturation factor rimM OS=Escherichia coli (strain K12) GN=rimM PE=1 SV=1	176.99
20	Y3808_ECOLI	Putative uncharacterized protein b3808 OS=Escherichia coli (strain K12) GN=b3808 PE=5 SV=2	287.1
21	TOLC_ECOLI	Outer membrane protein tolC OS=Escherichia coli (strain K12) GN=tolC PE=1 SV=3	233.34
22	PDXH_ECOLI	Pyridoxine/pyridoxamine 5'-phosphate oxidase OS=Escherichia coli (strain K12) GN=pdxH PE=1 SV=2	309.21
23	YDHB_ECOLI	Uncharacterized HTH-type transcriptional regulator ydhB OS=Escherichia coli (strain K12) GN=ydhB PE=4 SV=1	204.77
24	YHDH_ECOLI	Putative quinone oxidoreductase YhdH OS=Escherichia coli (strain K12) GN=yhdH PE=1 SV=1	163.89
25	RLPA_ECOLI	Rare lipoprotein A OS=Escherichia coli (strain K12) GN=rlpA PE=1 SV=1	139.46
26	HIS1_ECOLI	ATP phosphoribosyltransferase OS=Escherichia coli (strain K12) GN=hisG PE=1 SV=1	217.41
27	PIL1_ECOLI	Pilin OS=Escherichia coli (strain K12) GN=traA PE=1 SV=1	190.46

28	IFI_ECOLI	Translation initiation factor IF-1 OS=Escherichia coli (strain K12) GN=infA PE=1 SV=2	295.45
29	GPDA_ECOLI	Glycerol-3-phosphate dehydrogenase [NAD(P)+] OS=Escherichia coli (strain K12) GN=gpsA PE=3 SV=1	179.04
30	YHAJ_ECOLI	Uncharacterized HTH-type transcriptional regulator yhaJ OS=Escherichia coli (strain K12) GN=yhaJ PE=4 SV=1	163.21
31	YKFA_ECOLI	Uncharacterized protein ykfA OS=Escherichia coli (strain K12) GN=ykfA PE=4 SV=2	207.88
32	MPRA_ECOLI	Transcriptional repressor mprA OS=Escherichia coli (strain K12) GN=mprA PE=2 SV=1	332.79
33	YCIO_ECOLI	Uncharacterized protein yciO OS=Escherichia coli (strain K12) GN=yciO PE=1 SV=1	284.08
34	DCEA_ECOLI	Glutamate decarboxylase alpha OS=Escherichia coli (strain K12) GN=gadA PE=1 SV=1	133.08
35	RMLB1_ECOLI	dTDP-glucose 4,6-dehydratase 1 OS=Escherichia coli (strain K12) GN=rfbB PE=3 SV=2	160.04
36	FUMC_ECOLI	Fumarate hydratase class II OS=Escherichia coli (strain K12) GN=fumC PE=1 SV=1	148.13
37	FLIC_ECOLI	Flagellin OS=Escherichia coli (strain K12) GN=fliC PE=1 SV=2	203.62
38	GABD_ECOLI	Succinate-semialdehyde dehydrogenase [NADP+] OS=Escherichia coli (strain K12) GN=gabD PE=1 SV=1	144.78
39	YIBP_ECOLI	Uncharacterized protein yibP OS=Escherichia coli (strain K12) GN=yibP PE=4 SV=2	172.66
40	YDAT_ECOLI	Uncharacterized protein ydaT OS=Escherichia coli (strain K12) GN=ydaT PE=4 SV=1	215.68
41	HYPE_ECOLI	Hydrogenase isoenzymes formation protein hypE OS=Escherichia coli (strain K12) GN=hypE PE=1 SV=2	159.01
42	CAPP_ECOLI	Phosphoenolpyruvate carboxylase OS=Escherichia coli (strain K12) GN=ppc PE=1 SV=1	133.61
43	YJIL_ECOLI	Uncharacterized protein yjiL OS=Escherichia coli (strain K12) GN=yjiL PE=4 SV=2	199.55
44	OMPR_ECOLI	Transcriptional regulatory protein ompR OS=Escherichia coli (strain K12) GN=ompR PE=1 SV=1	329.22
45	MBHM_ECOLI	Hydrogenase-2 large chain OS=Escherichia coli (strain K12) GN=hybC PE=1 SV=2	203.68
46	XGPT_ECOLI	Xanthine phosphoribosyltransferase OS=Escherichia coli (strain K12) GN=gpt PE=1 SV=1	203.36
47	YJGR_ECOLI	Uncharacterized protein yjgR OS=Escherichia coli (strain K12) GN=yjgR PE=1 SV=1	136.81
48	SECD_ECOLI	Protein-export membrane protein secD OS=Escherichia coli (strain K12) GN=secD PE=3 SV=1	170.17
49	TAUB_ECOLI	Taurine import ATP-binding protein TauB OS=Escherichia coli (strain K12) GN=tauB PE=2 SV=1	150.74
50	ARAH_ECOLI	L-arabinose transport system permease protein AraH OS=Escherichia coli (strain K12) GN=araH PE=3 SV=2	437.91
51	RSUA_ECOLI	Ribosomal small subunit pseudouridine synthase A OS=Escherichia coli (strain K12) GN=rsuA PE=1 SV=1	511.28
52	CYSN_ECOLI	Sulfate adenyltransferase subunit 1 OS=Escherichia coli (strain K12) GN=cysN PE=1 SV=1	173.86
53	DPPD_ECOLI	Dipeptide transport ATP-binding protein DppD OS=Escherichia coli (strain K12) GN=dppD PE=3 SV=1	348.51
54	GLGC_ECOLI	Glucose-1-phosphate adenyltransferase OS=Escherichia coli (strain K12) GN=glgC PE=1 SV=2	175.19
55	AMID_ECOLI	N-acetylmuramoyl-L-alanine amidase AmiD OS=Escherichia coli (strain K12) GN=amiD PE=1 SV=1	170.27

ANEXO D- Proteínas encontradas apenas na linhagem BL21(DE03).

	Acesso	Descrição	Escore
1	UPP_ECOLI	Uracil phosphoribosyltransferase OS=Escherichia coli (strain K12) GN=upp PE=1 SV=1	390.39
2	Q04QB4_LEPBJ	Putative uncharacterized protein OS=Leptospira borgpetersenii serovar Hardjo-bovis (strain JB197) GN=LBJ_2455 PE=4 SV=1	288.04
3	NUSG_ECOLI	Transcription antitermination protein nusG OS=Escherichia coli (strain K12) GN=nusG PE=1 SV=2	168.55
4	HFLK_ECOLI	Modulator of FtsH protease HflK OS=Escherichia coli (strain K12) GN=hflK PE=1 SV=1	230.51
5	ISPH_ECOLI	4-hydroxy-3-methylbut-2-enyl diphosphate reductase OS=Escherichia coli (strain K12) GN=ispH PE=1 SV=1	157.92
6	YEDJ_ECOLI	Uncharacterized protein yedJ OS=Escherichia coli (strain K12) GN=yedJ PE=1 SV=2	330.33
7	RODZ_ECOLI	Cytoskeleton protein rodZ OS=Escherichia coli (strain K12) GN=rodZ PE=1 SV=2	218.64
8	YAET_ECOLI	Outer membrane protein assembly factor yaeT OS=Escherichia coli (strain K12) GN=yaeT PE=1 SV=1	174.59
9	RNR_ECOLI	Ribonuclease R OS=Escherichia coli (strain K12) GN=rnr PE=1 SV=2	338.77
10	YCBB_ECOLI	Probable L,D-transpeptidase YcbB OS=Escherichia coli (strain K12) GN=ycbB PE=1 SV=2	268.86
11	K6PF1_ECOLI	6-phosphofructokinase isozyme 1 OS=Escherichia coli (strain K12) GN=pfkA PE=1 SV=1	184.73
12	PTKA_ECOLI	Galactitol-specific phosphotransferase enzyme IIA component OS=Escherichia coli (strain K12) GN=gatA PE=4 SV=1	307.89
13	PPID_ECOLI	Peptidyl-prolyl cis-trans isomerase D OS=Escherichia coli (strain K12) GN=ppiD PE=1 SV=1	185.64
14	YQCA_ECOLI	Uncharacterized protein yqcA OS=Escherichia coli (strain K12) GN=yqcA PE=3 SV=1	792.65
15	RS20_ECOLI	30S ribosomal protein S20 OS=Escherichia coli (strain K12) GN=rpsT PE=1 SV=2	733.63
16	UMUC_ECOLI	Protein umuC OS=Escherichia coli (strain K12) GN=umuC PE=1 SV=2	308.36
17	MAO2_ECOLI	NADP-dependent malic enzyme OS=Escherichia coli (strain K12) GN=maeB PE=1 SV=1	318.19
18	YFJP_ECOLI	Uncharacterized protein yfjP OS=Escherichia coli (strain K12) GN=yfjP PE=4 SV=4	244.55
19	FOLX_ECOLI	D-erythro-7,8-dihydroneopterin triphosphate epimerase OS=Escherichia coli (strain K12) GN=folX PE=1 SV=2	459.35
20	SPEE_ECOLI	Spermidine synthase OS=Escherichia coli (strain K12) GN=speE PE=1 SV=2	227.63
21	RL31_ECOLI	50S ribosomal protein L31 OS=Escherichia coli (strain K12) GN=rpmE PE=1 SV=1	581.06
22	GYRB_ECOLI	DNA gyrase subunit B OS=Escherichia coli (strain K12) GN=gyrB PE=1 SV=2	175.03
23	GATY_ECOLI	D-tagatose-1,6-bisphosphate aldolase subunit GatY OS=Escherichia coli (strain K12) GN=gatY PE=1 SV=1	428.96
24	E4SYX3_LACDN	CRISPR-associated protein OS=Lactobacillus delbrueckii subsp. bulgaricus (strain ND02) GN=LDBND_0803 PE=4 SV=1	196.22
25	YEEZ_ECOLI	Protein yeeZ OS=Escherichia coli (strain K12) GN=yeeZ PE=1 SV=1	503.35
26	YJEN_ECOLI	Uncharacterized protein yjeN OS=Escherichia coli (strain K12) GN=yjeN PE=4 SV=1	967.6

27	YHBO_ECOLI	Protein yhbO OS=Escherichia coli (strain K12) GN=yhbO PE=1 SV=2	1069.89
28	GLO2_ECOLI	Hydroxyacylglutathione hydrolase OS=Escherichia coli (strain K12) GN=gloB PE=1 SV=1	391.18
29	FRMA_ECOLI	S-(hydroxymethyl)glutathione dehydrogenase OS=Escherichia coli (strain K12) GN=frmA PE=1 SV=3	380.81
30	MDAB_ECOLI	Modulator of drug activity B OS=Escherichia coli (strain K12) GN=mdaB PE=1 SV=1	302.87
31	POTD_ECOLI	Spermidine/putrescine-binding periplasmic protein OS=Escherichia coli (strain K12) GN=potD PE=1 SV=1	175.43
32	TORT_ECOLI	Periplasmic protein torT OS=Escherichia coli (strain K12) GN=torT PE=1 SV=3	245.41
33	MIOC_ECOLI	Protein mioC OS=Escherichia coli (strain K12) GN=mioC PE=1 SV=4	216.78
34	GLCE_ECOLI	Glycolate oxidase subunit glcE OS=Escherichia coli (strain K12) GN=glcE PE=4 SV=1	342.75
35	SLYB_ECOLI	Outer membrane lipoprotein slyB OS=Escherichia coli (strain K12) GN=slyB PE=2 SV=1	296.87
36	PSPA_ECOLI	Phage shock protein A OS=Escherichia coli (strain K12) GN=pspA PE=1 SV=2	220.06
37	YCDY_ECOLI	Uncharacterized protein ycdY OS=Escherichia coli (strain K12) GN=ycdY PE=3 SV=1	230.06
38	YGIW_ECOLI	Protein ygiW OS=Escherichia coli (strain K12) GN=ygiW PE=1 SV=1	342.99
39	GUAA_ECOLI	GMP synthase [glutamine-hydrolyzing] OS=Escherichia coli (strain K12) GN=guaA PE=1 SV=1	205.94
40	YQHD_ECOLI	Alcohol dehydrogenase YqhD OS=Escherichia coli (strain K12) GN=yqhD PE=1 SV=1	200.01
41	YBED_ECOLI	UPF0250 protein ybeD OS=Escherichia coli (strain K12) GN=ybeD PE=1 SV=1	491.19
42	AMN_ECOLI	AMP nucleosidase OS=Escherichia coli (strain K12) GN=amn PE=1 SV=1	273.38
43	CSPG_ECOLI	Cold shock-like protein CspG OS=Escherichia coli (strain K12) GN=cspG PE=1 SV=1	333.61
44	KRT86_HUMAN	Keratin type II cuticular Hb6 OS=Homo sapiens GN=KRT86 PE=1 SV=1	196.52
45	STPA_ECOLI	DNA-binding protein stpA OS=Escherichia coli (strain K12) GN=stpA PE=3 SV=1	260.81
46	PGM_ECOLI	Phosphoglucomutase OS=Escherichia coli (strain K12) GN=pgm PE=1 SV=1	326.05
47	HYCA_ECOLI	Formate hydrogenlyase regulatory protein hycA OS=Escherichia coli (strain K12) GN=hycA PE=1 SV=1	250.9
48	B8LFD7_ECOLI	Galactoside permease OS=Escherichia coli (strain K12) PE=4 SV=1	200.41
49	GHRB_ECOLI	Glyoxylate/hydroxypyruvate reductase B OS=Escherichia coli (strain K12) GN=ghrB PE=1 SV=3	762.8
50	D8FLJ6_LACDE	CRISPR-associated endonuclease Cas1 OS=Lactobacillus delbrueckii subsp. bulgaricus PB2003/044-T3-4 GN=cas1 PE=4 SV=1	196.22
51	YEBF_ECOLI	Protein yebF OS=Escherichia coli (strain K12) GN=yebF PE=1 SV=2	202.72
52	PAAC_ECOLI	Phenylacetic acid degradation protein paaC OS=Escherichia coli (strain K12) GN=paaC PE=1 SV=1	293.03
53	PHNB_ECOLI	Protein phnB OS=Escherichia coli (strain K12) GN=phnB PE=1 SV=1	269.2

54	YEHR_ECOLI	Uncharacterized lipoprotein yehR OS=Escherichia coli (strain K12) GN=yehR PE=1 SV=2	345.02
55	YCIT_ECOLI	Uncharacterized HTH-type transcriptional regulator yciT OS=Escherichia coli (strain K12) GN=yctT PE=4 SV=1	204.98
56	DSBA_ECOLI	Thiol:disulfide interchange protein DsbA OS=Escherichia coli (strain K12) GN=dsbA PE=1 SV=1	288.19
57	Q054L7_LEPBL	Putative uncharacterized protein OS=Leptospira borgpetersenii serovar Hardjo-bovis (strain L550) GN=LBL_0657 PE=4 SV=1	288.04
58	ISCR_ECOLI	HTH-type transcriptional regulator iscR OS=Escherichia coli (strain K12) GN=iscR PE=1 SV=1	249.45
59	A5PFJ8_ECOLI	Disrupted ArcA (Fragment) OS=Escherichia coli (strain K12) GN=arcA2 PE=4 SV=1	806.31
60	YHIH_ECOLI	Uncharacterized ABC transporter ATP-binding protein YhiH OS=Escherichia coli (strain K12) GN=yhiH PE=1 SV=3	168.5
61	UGPB_ECOLI	sn-glycerol-3-phosphate-binding periplasmic protein ugpB OS=Escherichia coli (strain K12) GN=ugpB PE=1 SV=1	235.14
62	ALF1_ECOLI	Fructose-bisphosphate aldolase class 1 OS=Escherichia coli (strain K12) GN=fbaB PE=1 SV=2	278.52
63	PTNAB_ECOLI	PTS system mannose-specific EIIAB component OS=Escherichia coli (strain K12) GN=manX PE=1 SV=2	229.44
64	RS6_ECOLI	30S ribosomal protein S6 OS=Escherichia coli (strain K12) GN=rpsF PE=1 SV=1	372.02
65	RPIA_ECOLI	Ribose-5-phosphate isomerase A OS=Escherichia coli (strain K12) GN=rpiA PE=1 SV=1	232.85
66	YIIF_ECOLI	Uncharacterized protein yiiF OS=Escherichia coli (strain K12) GN=yiiF PE=4 SV=1	320.13
67	MALM_ECOLI	Maltose operon periplasmic protein OS=Escherichia coli (strain K12) GN=malM PE=4 SV=1	299.08
68	PSS_ECOLI	CDP-diacylglycerol--serine O-phosphatidyltransferase OS=Escherichia coli (strain K12) GN=pssA PE=1 SV=2	180.7
69	RPIR_ECOLI	HTH-type transcriptional regulator rpiR OS=Escherichia coli (strain K12) GN=rpiR PE=1 SV=1	332.53
70	RPOD_ECOLI	RNA polymerase sigma factor rpoD OS=Escherichia coli (strain K12) GN=rpoD PE=1 SV=2	220.36
71	YRDA_ECOLI	Protein YrdA OS=Escherichia coli (strain K12) GN=yrdA PE=1 SV=1	462.89
72	KRT81_HUMAN	Keratin type II cuticular Hb1 OS=Homo sapiens GN=KRT81 PE=1 SV=3	187.33
73	CYOB_ECOLI	Ubiquinol oxidase subunit 1 OS=Escherichia coli (strain K12) GN=cyoB PE=1 SV=1	231.71
74	NUDE_ECOLI	ADP compounds hydrolase nudE OS=Escherichia coli (strain K12) GN=nudE PE=1 SV=1	185.01
75	CATE_ECOLI	Catalase HP11 OS=Escherichia coli (strain K12) GN=katE PE=1 SV=1	250.75
76	GLTS_ECOLI	Sodium/glutamate symport carrier protein OS=Escherichia coli (strain K12) GN=gltS PE=1 SV=1	326.47
77	UDP_ECOLI	Uridine phosphorylase OS=Escherichia coli (strain K12) GN=udp PE=1 SV=3	808.95
78	PROX_ECOLI	Glycine betaine-binding periplasmic protein OS=Escherichia coli (strain K12) GN=proX PE=1 SV=1	239.05
79	KGUA_ECOLI	Guanylate kinase OS=Escherichia coli (strain K12) GN=gmk PE=1 SV=1	279.55
80	ASTD_ECOLI	N-succinylglutamate 5-semialdehyde dehydrogenase OS=Escherichia coli (strain K12) GN=astD PE=1 SV=1	324.91

81	YCAR_ECOLI	UPF0434 protein ycaR OS=Escherichia coli (strain K12) GN=ycaR PE=1 SV=1	956.85
82	YDIA_ECOLI	Putative phosphotransferase ydiA OS=Escherichia coli (strain K12) GN=ydiA PE=1 SV=1	220.62
83	FKBX_ECOLI	FKBP-type 16 kDa peptidyl-prolyl cis-trans isomerase OS=Escherichia coli (strain K12) GN=fkpB PE=1 SV=2	349.4
84	YFIO_ECOLI	UPF0169 lipoprotein yfiO OS=Escherichia coli (strain K12) GN=yfiO PE=1 SV=1	191.55
85	COBT_ECOLI	Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase OS=Escherichia coli (strain K12) GN=cobT PE=3 SV=3	553.6
86	FDHD_ECOLI	Protein FdhD OS=Escherichia coli (strain K12) GN=fdhD PE=3 SV=1	236.32
87	YLIJ_ECOLI	Uncharacterized GST-like protein yliJ OS=Escherichia coli (strain K12) GN=yliJ PE=1 SV=1	665.69
88	IRAP_ECOLI	Anti-adapter protein iraP OS=Escherichia coli (strain K12) GN=iraP PE=1 SV=1	960.2
89	BTUR_ECOLI	Cob(I)yrinic acid a,c-diamide adenosyltransferase OS=Escherichia coli (strain K12) GN=btuR PE=3 SV=1	241.52
90	SLP_ECOLI	Outer membrane protein slp OS=Escherichia coli (strain K12) GN=slp PE=1 SV=1	319.82
91	TORA_ECOLI	Trimethylamine-N-oxide reductase 1 OS=Escherichia coli (strain K12) GN=torA PE=1 SV=2	196.57
92	LACY_ECOLI	Lactose permease OS=Escherichia coli (strain K12) GN=lacY PE=1 SV=1	200.41
93	G6PI_ECOLI	Glucose-6-phosphate isomerase OS=Escherichia coli (strain K12) GN=pgi PE=1 SV=1	302.2
94	YCDO_ECOLI	UPF0409 protein ycdO OS=Escherichia coli (strain K12) GN=ycdO PE=1 SV=1	253.64
95	YKGD_ECOLI	Uncharacterized HTH-type transcriptional regulator ykgD OS=Escherichia coli (strain K12) GN=ykgD PE=4 SV=1	138.51
96	PABA_ECOLI	Para-aminobenzoate synthase glutamine amidotransferase component II OS=Escherichia coli (strain K12) GN=pabA PE=1 SV=1	197.24
97	ADHP_ECOLI	Alcohol dehydrogenase, propanol-preferring OS=Escherichia coli (strain K12) GN=adhP PE=1 SV=1	188.67
98	CSGE_ECOLI	Curli production assembly/transport component CsgE OS=Escherichia coli (strain K12) GN=csgE PE=4 SV=1	521.26
99	PROB_ECOLI	Glutamate 5-kinase OS=Escherichia coli (strain K12) GN=proB PE=1 SV=1	250.21
100	BTUE_ECOLI	Vitamin B12 transport periplasmic protein BtuE OS=Escherichia coli (strain K12) GN=btuE PE=3 SV=1	337.29
101	NIFU_ECOLI	NifU-like protein OS=Escherichia coli (strain K12) GN=nifU PE=1 SV=1	1170.19
102	YAIE_ECOLI	UPF0345 protein yaiE OS=Escherichia coli (strain K12) GN=yaiE PE=3 SV=1	210.63
103	ERPA_ECOLI	Iron-sulfur cluster insertion protein erpA OS=Escherichia coli (strain K12) GN=erpA PE=1 SV=1	1084.54
104	YBFF_ECOLI	Esterase ybfF OS=Escherichia coli (strain K12) GN=ybfF PE=1 SV=1	215.71
105	AK1H_ECOLI	Bifunctional aspartokinase/homoserine dehydrogenase 1 OS=Escherichia coli (strain K12) GN=thrA PE=1 SV=2	132.24
106	DLD_ECOLI	D-lactate dehydrogenase OS=Escherichia coli (strain K12) GN=dld PE=1 SV=3	168.85
107	THRC_ECOLI	Threonine synthase OS=Escherichia coli (strain K12) GN=thrC PE=1 SV=1	219.68

108	GLNB_ECOLI	Nitrogen regulatory protein P-II 1 OS=Escherichia coli (strain K12) GN=glnB PE=1 SV=1	677.87
109	GSHR_ECOLI	Glutathione reductase OS=Escherichia coli (strain K12) GN=gor PE=1 SV=1	260.15
110	YEBV_ECOLI	Uncharacterized protein yebV OS=Escherichia coli (strain K12) GN=yebV PE=4 SV=1	583.16
111	RIHC_ECOLI	Non-specific ribonucleoside hydrolase rihC OS=Escherichia coli (strain K12) GN=rihC PE=1 SV=1	419.53
112	MSYB_ECOLI	Acidic protein msyB OS=Escherichia coli (strain K12) GN=msyB PE=4 SV=1	209.56
113	YCFD_ECOLI	Uncharacterized protein ycfD OS=Escherichia coli (strain K12) GN=ycfD PE=4 SV=2	189.45
114	ISPZ_ECOLI	Probable intracellular septation protein OS=Escherichia coli (strain K12) GN=ispZ PE=1 SV=1	257.72
115	YJGA_ECOLI	UPF0307 protein yjgA OS=Escherichia coli (strain K12) GN=yjgA PE=1 SV=1	627.48
116	OXAA_ECOLI	Inner membrane protein oxaA OS=Escherichia coli (strain K12) GN=oxaA PE=1 SV=2	198.17
117	KDSB_ECOLI	3-deoxy-manno-octulosonate cytidylyltransferase OS=Escherichia coli (strain K12) GN=kdsB PE=1 SV=2	174.41
118	FENR_ECOLI	Ferredoxin--NADP reductase OS=Escherichia coli (strain K12) GN=fpr PE=1 SV=4	334.61
119	TOP1_ECOLI	DNA topoisomerase 1 OS=Escherichia coli (strain K12) GN=topA PE=1 SV=2	126.28
120	NADE_ECOLI	NH(3)-dependent NAD(+) synthetase OS=Escherichia coli (strain K12) GN=nadE PE=1 SV=2	158.12
121	MURE_ECOLI	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase OS=Escherichia coli (strain K12) GN=murE PE=1 SV=2	113.37
122	G6PD_ECOLI	Glucose-6-phosphate 1-dehydrogenase OS=Escherichia coli (strain K12) GN=zwf PE=1 SV=1	204.43
123	YBIT_ECOLI	Uncharacterized ABC transporter ATP-binding protein YbiT OS=Escherichia coli (strain K12) GN=ybiT PE=1 SV=1	182.35
124	MURA_ECOLI	UDP-N-acetylglucosamine 1-carboxyvinyltransferase OS=Escherichia coli (strain K12) GN=murA PE=1 SV=1	213.06
125	DAPB_ECOLI	Dihydrodipicolinate reductase OS=Escherichia coli (strain K12) GN=dapB PE=1 SV=1	230.05
126	TOLA_ECOLI	Protein tolA OS=Escherichia coli (strain K12) GN=tolA PE=1 SV=1	265.7
127	ENGB_ECOLI	Probable GTP-binding protein EngB OS=Escherichia coli (strain K12) GN=engB PE=1 SV=1	208.73
128	YJBR_ECOLI	Uncharacterized protein yjbr OS=Escherichia coli (strain K12) GN=yjbr PE=1 SV=1	555.34
129	P5CR_ECOLI	Pyrroline-5-carboxylate reductase OS=Escherichia coli (strain K12) GN=proC PE=1 SV=1	266.35
130	PYRI_ECOLI	Aspartate carbamoyltransferase regulatory chain OS=Escherichia coli (strain K12) GN=pyrI PE=1 SV=2	454.69
131	OPGH_ECOLI	Glucans biosynthesis glucosyltransferase H OS=Escherichia coli (strain K12) GN=mdoH PE=1 SV=1	127.36
132	YTFP_ECOLI	Gamma-glutamylcyclotransferase family protein ytfP OS=Escherichia coli (strain K12) GN=ytfP PE=1 SV=1	244.52
133	HLDE_ECOLI	Bifunctional protein hldE OS=Escherichia coli (strain K12) GN=hldE PE=1 SV=1	214.38
134	SERC_ECOLI	Phosphoserine aminotransferase OS=Escherichia coli (strain K12) GN=serC PE=1 SV=4	281.12

135	YJIP_ECOLI	Putative uncharacterized protein yjiP OS=Escherichia coli (strain K12) GN=yjiP PE=5 SV=2	199.38
136	GLXK2_ECOLI	Glycerate kinase 2 OS=Escherichia coli (strain K12) GN=garK PE=1 SV=2	241.98
137	RIR1_ECOLI	Ribonucleoside-diphosphate reductase 1 subunit alpha OS=Escherichia coli (strain K12) GN=nrdA PE=1 SV=2	210.21
138	YHHX_ECOLI	Uncharacterized oxidoreductase yhhX OS=Escherichia coli (strain K12) GN=yhhX PE=1 SV=1	334.72
139	QOR1_ECOLI	Quinone oxidoreductase 1 OS=Escherichia coli (strain K12) GN=qorA PE=1 SV=1	292.89
140	OSMF_ECOLI	Putative osmoprotectant uptake system substrate-binding protein osmF OS=Escherichia coli (strain K12) GN=osmF PE=2 SV=1	273.27
141	FECE_ECOLI	Fe(3+) dicitrate transport ATP-binding protein FecE OS=Escherichia coli (strain K12) GN=fecE PE=1 SV=1	232.57
142	PANB_ECOLI	3-methyl-2-oxobutanoate hydroxymethyltransferase OS=Escherichia coli (strain K12) GN=panB PE=1 SV=1	222.89
143	DDPA_ECOLI	Probable D,D-dipeptide-binding periplasmic protein ddpA OS=Escherichia coli (strain K12) GN=ddpA PE=2 SV=1	291.95
144	SELD_ECOLI	Selenide, water dikinase OS=Escherichia coli (strain K12) GN=selD PE=1 SV=1	356.91
145	YPDF_ECOLI	Aminopeptidase ypdF OS=Escherichia coli (strain K12) GN=ypdF PE=1 SV=1	189.1
146	YUAN_ECOLI	Uncharacterized protein YuaN OS=Escherichia coli (strain K12) GN=yuaN PE=4 SV=1	647.3

ANEXO E- Proteínas encontradas em maior quantidade na linhagem BL21(DE03)pACYC184 em relação à BL21(DE03).

	Acesso	descrição	Escore	Razão pACYC184:BL21(DE03)	razão log(e)	log(e) variância
1	MTNN_ECOLI	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase OS=Escherichia coli (strain K12) GN=mtnN PE=1 SV=1	438.08	70.10541235	4.25	0.38
2	RHO_ECOLI	Transcription termination factor Rho OS=Escherichia coli (strain K12) GN=rho PE=1 SV=1	519.16	1.433329435	0.36	0.1
3	NUOF_ECOLI	NADH-quinone oxidoreductase subunit F OS=Escherichia coli (strain K12) GN=nuoF PE=1 SV=3	359.88	1.221402762	0.2	0.13
4	EFTU2_ECOLI	Elongation factor Tu 2 OS=Escherichia coli (strain K12) GN=tufB PE=1 SV=1	38326.05	1.051271097	0.05	0.03
5	RS4_ECOLI	30S ribosomal protein S4 OS=Escherichia coli (strain K12) GN=rpsD PE=1 SV=2	6084.9	1.197217372	0.18	0.07
6	FUMA_ECOLI	Fumarate hydratase class I, aerobic OS=Escherichia coli (strain K12) GN=fumA PE=1 SV=2	545.41	1.1502738	0.14	0.09
7	CLPB_ECOLI	Chaperone protein ClpB OS=Escherichia coli (strain K12) GN=clpB PE=1 SV=1	1804.89	1.083287066	0.08	0.06
8	CH10_ECOLI	10 kDa chaperonin OS=Escherichia coli (strain K12) GN=groS PE=1 SV=1	6121.47	1.16183425	0.15	0.07
9	RS15_ECOLI	30S ribosomal protein S15 OS=Escherichia coli (strain K12) GN=rpsO PE=1 SV=2	1723.44	1.127496849	0.12	0.08
10	HTPG_ECOLI	Chaperone protein htpG OS=Escherichia coli (strain K12) GN=htpG PE=1 SV=1	772.47	1.296930074	0.26	0.08
11	ISCS_ECOLI	Cysteine desulfurase OS=Escherichia coli (strain K12) GN=iscS PE=1 SV=1	186.05	1.233678052	0.21	0.16
12	EFTU1_ECOLI	Elongation factor Tu 1 OS=Escherichia coli (strain K12) GN=tufA PE=1 SV=1	38326.05	1.051271097	0.05	0.02
13	PCKA_ECOLI	Phosphoenolpyruvate carboxykinase [ATP] OS=Escherichia coli (strain K12) GN=pckA PE=1 SV=2	1626.47	1.11627807	0.11	0.06
14	DNAK_ECOLI	Chaperone protein DnaK OS=Escherichia coli (strain K12) GN=dnaK PE=1 SV=2	7978.2	1.127496849	0.12	0.03
15	CSPC_ECOLI	Cold shock-like protein CspC OS=Escherichia coli (strain K12) GN=cspC PE=1 SV=2	4703.18	1.258600015	0.23	0.06
16	USPF_ECOLI	Universal stress protein F OS=Escherichia coli (strain K12) GN=uspF PE=1 SV=2	1798.32	1.616074385	0.48	0.17
17	MIND_ECOLI	Septum site-determining protein minD OS=Escherichia coli (strain K12) GN=minD PE=1 SV=2	846.16	1.296930074	0.26	0.15
18	FTNA_ECOLI	Ferritin-1 OS=Escherichia coli (strain K12) GN=ftnA PE=1 SV=1	3217.81	1.491824707	0.4	0.1
19	RS10_ECOLI	30S ribosomal protein S10 OS=Escherichia coli (strain K12) GN=rpsJ PE=1 SV=1	1846.44	1.233678052	0.21	0.12
20	DBHA_ECOLI	DNA-binding protein HU-alpha OS=Escherichia coli (strain K12) GN=hupA PE=1 SV=1	30092.34	1.083287066	0.08	0.04
21	TKT1_ECOLI	Transketolase 1 OS=Escherichia coli (strain K12) GN=tktA PE=1 SV=5	497.48	1.16183425	0.15	0.07
22	ISPG_ECOLI	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase OS=Escherichia coli (strain K12) GN=ispG PE=1 SV=1	256.14	1.632316236	0.49	0.33

23	RS18_ECOLI	30S ribosomal protein S18 OS=Escherichia coli (strain K12) GN=rpsR PE=1 SV=2	3278.06	1.377127754	0.32	0.14
24	FABB_ECOLI	3-oxoacyl-[acyl-carrier-protein] synthase 1 OS=Escherichia coli (strain K12) GN=fabB PE=1 SV=1	815.1	1.209249595	0.19	0.11
25	IF2_ECOLI	Translation initiation factor IF-2 OS=Escherichia coli (strain K12) GN=infB PE=1 SV=1	296.65	1.197217372	0.18	0.1
26	PTA_ECOLI	Phosphate acetyltransferase OS=Escherichia coli (strain K12) GN=pta PE=1 SV=2	201.02	1.246076729	0.22	0.14
27	AAT_ECOLI	Aspartate aminotransferase OS=Escherichia coli (strain K12) GN=aspC PE=1 SV=1	2629.43	1.1502738	0.14	0.07
28	TNAA_ECOLI	Tryptophanase OS=Escherichia coli (strain K12) GN=tnaA PE=1 SV=1	14890.54	1.197217372	0.18	0.02
29	ATPB_ECOLI	ATP synthase subunit beta OS=Escherichia coli (strain K12) GN=atpD PE=1 SV=2	1614.4	1.138828378	0.13	0.06
30	SYGA_ECOLI	Glycyl-tRNA synthetase alpha subunit OS=Escherichia coli (strain K12) GN=glyQ PE=1 SV=2	232.71	1.323129814	0.28	0.19
31	RS13_ECOLI	30S ribosomal protein S13 OS=Escherichia coli (strain K12) GN=rpsM PE=1 SV=2	514.5	1.462284582	0.38	0.18
32	DADA_ECOLI	D-amino acid dehydrogenase small subunit OS=Escherichia coli (strain K12) GN=dadA PE=2 SV=1	813.6	1.271249144	0.24	0.11
33	OMPF_ECOLI	Outer membrane protein F OS=Escherichia coli (strain K12) GN=ompF PE=1 SV=1	734.47	1.41906754	0.35	0.13
34	ARNB_ECOLI	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase OS=Escherichia coli (strain K12) GN=arnB PE=1 SV=2	248.17	1.271249144	0.24	0.16
35	ODO2_ECOLI	Dihydropyridoxylsuccinyltransferase component of 2-oxoglutarate dehydrogenase complex OS=Escherichia coli (strain K12) GN=sucB PE=1 SV=2	4634.85	1.094174288	0.09	0.05
36	TALB_ECOLI	Transaldolase B OS=Escherichia coli (strain K12) GN=talB PE=1 SV=2	2293.4	1.10517092	0.1	0.06
37	RL20_ECOLI	50S ribosomal protein L20 OS=Escherichia coli (strain K12) GN=rplT PE=1 SV=2	1500.89	1.185304853	0.17	0.11
38	CYSK_ECOLI	Cysteine synthase A OS=Escherichia coli (strain K12) GN=cysK PE=1 SV=2	4203.58	1.363425117	0.31	0.04
39	RL15_ECOLI	50S ribosomal protein L15 OS=Escherichia coli (strain K12) GN=rplO PE=1 SV=1	5534.56	1.16183425	0.15	0.1
40	KPYK2_ECOLI	Pyruvate kinase II OS=Escherichia coli (strain K12) GN=pykA PE=1 SV=3	1323.09	1.138828378	0.13	0.1
41	SYD_ECOLI	Aspartyl-tRNA synthetase OS=Escherichia coli (strain K12) GN=aspS PE=1 SV=1	226.38	1.197217372	0.18	0.15
42	LIPA_ECOLI	Lipoyl synthase OS=Escherichia coli (strain K12) GN=lipA PE=1 SV=1	316.62	1.296930074	0.26	0.18
43	B8LFD6_ECOLI	Beta-galactosidase 2 OS=Escherichia coli (strain K12) GN=lacZ2 PE=1 SV=1	6338.66	1.040810773	0.04	0.02
44	RL7_ECOLI	50S ribosomal protein L7/L12 OS=Escherichia coli (strain K12) GN=rplL PE=1 SV=2	3925.98	1.127496849	0.12	0.08
45	RL28_ECOLI	50S ribosomal protein L28 OS=Escherichia coli (strain K12) GN=rpmB PE=1 SV=2	5227.78	1.296930074	0.26	0.1
46	OMPA_ECOLI	Outer membrane protein A OS=Escherichia coli (strain K12) GN=ompA PE=1 SV=1	16385.71	1.584073998	0.46	0.03
47	BGAL_ECOLI	Beta-galactosidase OS=Escherichia coli (strain K12) GN=lacZ PE=1 SV=2	6338.66	1.051271097	0.05	0.03
48	ATPA_ECOLI	ATP synthase subunit alpha OS=Escherichia coli (strain K12) GN=atpA PE=1 SV=1	2778.14	1.10517092	0.1	0.06
49	NFSA_ECOLI	Oxygen-insensitive NADPH nitroreductase OS=Escherichia coli (strain K12) GN=nfsA PE=1	271.72	1.349858824	0.3	0.19

		SV=2				
50	RL10_ECOLI	50S ribosomal protein L10 OS=Escherichia coli (strain K12) GN=rplJ PE=1 SV=2	11694.35	1.185304853	0.17	0.06
51	RPOC_ECOLI	DNA-directed RNA polymerase subunit beta' OS=Escherichia coli (strain K12) GN=rpoC PE=1 SV=1	634.18	1.138828378	0.13	0.05
52	ARNA_ECOLI	Bifunctional polymyxin resistance protein ArnA OS=Escherichia coli (strain K12) GN=arnA PE=1 SV=1	373.76	1.233678052	0.21	0.1
53	ODO1_ECOLI	2-oxoglutarate dehydrogenase E1 component OS=Escherichia coli (strain K12) GN=sucA PE=1 SV=1	865.29	1.10517092	0.1	0.06
54	FUMB_ECOLI	Fumarate hydratase class I, anaerobic OS=Escherichia coli (strain K12) GN=fumB PE=1 SV=2	452.7	1.197217372	0.18	0.1
55	RL14_ECOLI	50S ribosomal protein L14 OS=Escherichia coli (strain K12) GN=rplN PE=1 SV=1	1288.13	1.258600015	0.23	0.11
56	IPYR_ECOLI	Inorganic pyrophosphatase OS=Escherichia coli (strain K12) GN=ppa PE=1 SV=2	335.52	1.377127754	0.32	0.15
57	DEGP_ECOLI	Protease do OS=Escherichia coli (strain K12) GN=degP PE=1 SV=1	256.73	1.296930074	0.26	0.14
58	EFG_ECOLI	Elongation factor G OS=Escherichia coli (strain K12) GN=fusA PE=1 SV=2	4027.5	1.072508182	0.07	0.04
59	RPOB_ECOLI	DNA-directed RNA polymerase subunit beta OS=Escherichia coli (strain K12) GN=rpoB PE=1 SV=1	305.81	1.10517092	0.1	0.07
60	HSLU_ECOLI	ATP-dependent protease ATPase subunit HslU OS=Escherichia coli (strain K12) GN=hslU PE=1 SV=1	521.74	1.404947596	0.34	0.14
61	RNE_ECOLI	Ribonuclease E OS=Escherichia coli (strain K12) GN=rne PE=1 SV=6	212.82	1.284025417	0.25	0.22
62	RS3_ECOLI	30S ribosomal protein S3 OS=Escherichia coli (strain K12) GN=rpsC PE=1 SV=2	4876.25	1.083287066	0.08	0.07
63	BOLA_ECOLI	Protein BofA OS=Escherichia coli (strain K12) GN=bofA PE=1 SV=1	554.01	1.377127754	0.32	0.29
64	FABF_ECOLI	3-oxoacyl-[acyl-carrier-protein] synthase 2 OS=Escherichia coli (strain K12) GN=fabF PE=1 SV=2	497.65	1.173510867	0.16	0.12
65	ACRA_ECOLI	Acriflavine resistance protein A OS=Escherichia coli (strain K12) GN=acrA PE=1 SV=1	225.13	1.246076729	0.22	0.14
66	GHRA_ECOLI	Glyoxylate/hydroxypyruvate reductase A OS=Escherichia coli (strain K12) GN=ghrA PE=1 SV=2	290.98	3.095656485	1.13	0.57
67	ACCA_ECOLI	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha OS=Escherichia coli (strain K12) GN=accA PE=1 SV=2	1363.55	1.185304853	0.17	0.15
68	YNCB_ECOLI	Putative NADP-dependent oxidoreductase yncB OS=Escherichia coli (strain K12) GN=yncB PE=3 SV=3	514.56	1.185304853	0.17	0.13
69	CPTA_ECOLI	Phosphoethanolamine transferase CptA OS=Escherichia coli (strain K12) GN=cptA PE=1 SV=1	206.82	1.632316236	0.49	0.35
70	ACP_ECOLI	Acyl carrier protein OS=Escherichia coli (strain K12) GN=acpP PE=1 SV=2	2217.57	1.094174288	0.09	0.07
71	CH60_ECOLI	60 kDa chaperonin OS=Escherichia coli (strain K12) GN=groL PE=1 SV=2	11061.78	1.040810773	0.04	0.03

72	ASPA_ECOLI	Aspartate ammonia-lyase OS=Escherichia coli (strain K12) GN=aspA PE=1 SV=1	1587.5	1.083287066	0.08	0.08
73	PROQ_ECOLI	ProP effector OS=Escherichia coli (strain K12) GN=proQ PE=1 SV=2	182.3	1.246076729	0.22	0.17
74	RL17_ECOLI	50S ribosomal protein L17 OS=Escherichia coli (strain K12) GN=rplQ PE=1 SV=1	2458.14	1.11627807	0.11	0.12
75	FABG_ECOLI	3-oxoacyl-[acyl-carrier-protein] reductase FabG OS=Escherichia coli (strain K12) GN=fabG PE=1 SV=1	689.36	1.246076729	0.22	0.2
76	RPOA_ECOLI	DNA-directed RNA polymerase subunit alpha OS=Escherichia coli (strain K12) GN=rpoA PE=1 SV=1	2883.62	1.072508182	0.07	0.07
77	SUCC_ECOLI	Succinyl-CoA ligase [ADP-forming] subunit beta OS=Escherichia coli (strain K12) GN=sucC PE=1 SV=1	8003.11	1.061836545	0.06	0.05
78	PYRC_ECOLI	Dihydroorotase OS=Escherichia coli (strain K12) GN=pyrC PE=1 SV=2	411.61	1.309964465	0.27	0.23
79	HINT_ECOLI	HIT-like protein hinT OS=Escherichia coli (strain K12) GN=hinT PE=1 SV=1	1456.92	1.1502738	0.14	0.13
80	PPIB_ECOLI	Peptidyl-prolyl cis-trans isomerase B OS=Escherichia coli (strain K12) GN=ppiB PE=1 SV=2	481.51	1.173510867	0.16	0.18
81	DACA_ECOLI	D-alanyl-D-alanine carboxypeptidase dacA OS=Escherichia coli (strain K12) GN=dacA PE=1 SV=1	266.5	1.934792385	0.66	0.56
82	YGDH_ECOLI	LOG family protein ygdH OS=Escherichia coli (strain K12) GN=ygdH PE=1 SV=1	238.18	2.159766213	0.77	0.9
83	GLRX4_ECOLI	Glutaredoxin-4 OS=Escherichia coli (strain K12) GN=grxD PE=1 SV=1	1041.32	1.127496849	0.12	0.09
84	SUCD_ECOLI	Succinyl-CoA ligase [ADP-forming] subunit alpha OS=Escherichia coli (strain K12) GN=sucD PE=1 SV=2	7140.11	1.051271097	0.05	0.06
85	KDSA_ECOLI	2-dehydro-3-deoxyphosphooctonate aldolase OS=Escherichia coli (strain K12) GN=kdsA PE=1 SV=1	760.72	1.11627807	0.11	0.12
86	CYDA_ECOLI	Cytochrome d ubiquinol oxidase subunit 1 OS=Escherichia coli (strain K12) GN=cydA PE=1 SV=1	1318.34	1.094174288	0.09	0.09
87	CRP_ECOLI	Catabolite gene activator OS=Escherichia coli (strain K12) GN=crp PE=1 SV=1	321.29	1.197217372	0.18	0.19
88	PURA_ECOLI	Adenylosuccinate synthetase OS=Escherichia coli (strain K12) GN=purA PE=1 SV=2	943.4	1.10517092	0.1	0.1
89	HLDD_ECOLI	ADP-L-glycero-D-manno-heptose-6-epimerase OS=Escherichia coli (strain K12) GN=hldD PE=1 SV=1	953.89	1.11627807	0.11	0.12
90	ATPG_ECOLI	ATP synthase gamma chain OS=Escherichia coli (strain K12) GN=atpG PE=1 SV=1	266.99	1.233678052	0.21	0.19
91	LOLA_ECOLI	Outer-membrane lipoprotein carrier protein OS=Escherichia coli (strain K12) GN=lolA PE=1 SV=1	280.67	1.69893226	0.53	0.63
92	PTKC_ECOLI	Galactitol permease IIC component OS=Escherichia coli (strain K12) GN=gatC PE=1 SV=1	304.36	1.233678052	0.21	0.21
93	RS2_ECOLI	30S ribosomal protein S2 OS=Escherichia coli (strain K12) GN=rpsB PE=1 SV=2	5629.37	1.061836545	0.06	0.05
94	YBJP_ECOLI	Uncharacterized lipoprotein ybjP OS=Escherichia coli (strain K12) GN=ybjP PE=1 SV=1	696.27	1.173510867	0.16	0.16
95	GALU_ECOLI	UTP--glucose-1-phosphate uridylyltransferase OS=Escherichia coli (strain K12) GN=galU PE=1 SV=2	269.31	1.447734622	0.37	0.37

96	C7MTL6_SACVD	CRISPR-associated protein, Cas1 family OS=Saccharomonospora viridis (strain ATCC 15386 / DSM 43017 / JCM 3036 / NBRC 12207 / P101) GN=Svir_17190 PE=4 SV=1	204.22	2.1382762	0.76	0.59
97	ODP1_ECOLI	Pyruvate dehydrogenase E1 component OS=Escherichia coli (strain K12) GN=aceE PE=1 SV=2	1090.65	1.061836545	0.06	0.06
98	PHS2_RABIT	GLYCOGEN PHOSPHORYLASE, MUSCLE FORM (EC	1453.15	1.040810773	0.04	0.04
99	NDK_ECOLI	Nucleoside diphosphate kinase OS=Escherichia coli (strain K12) GN=ndk PE=1 SV=2	1035.56	1.127496849	0.12	0.13
100	RL5_ECOLI	50S ribosomal protein L5 OS=Escherichia coli (strain K12) GN=rplE PE=1 SV=2	1096.79	1.061836545	0.06	0.08
101	SYN_ECOLI	Asparaginyl-tRNA synthetase OS=Escherichia coli (strain K12) GN=asnS PE=1 SV=2	566.99	1.11627807	0.11	0.12
102	SYFB_ECOLI	Phenylalanyl-tRNA synthetase beta chain OS=Escherichia coli (strain K12) GN=pheT PE=1 SV=2	284.93	1.197217372	0.18	0.18
103	YEBE_ECOLI	Inner membrane protein yebE OS=Escherichia coli (strain K12) GN=yebE PE=1 SV=2	249.45	1.390968147	0.33	0.34
104	CYSI_ECOLI	Sulfite reductase [NADPH] hemoprotein beta-component OS=Escherichia coli (strain K12) GN=cysI PE=1 SV=4	186.66	1.768267039	0.57	0.6
105	GCST_ECOLI	Aminomethyltransferase OS=Escherichia coli (strain K12) GN=gcvT PE=1 SV=3	421.89	1.10517092	0.1	0.12
106	YAJC_ECOLI	UPF0092 membrane protein yajC OS=Escherichia coli (strain K12) GN=yajC PE=1 SV=1	354.33	1.173510867	0.16	0.17

ANEXO F- Proteínas encontradas em maior quantidade na linhagem BL21(DE03)pACYC184 em relação à BL21(DE03).

	Acesso	Descrição	Escore	Razão pACYC184: BL21(DE03)	razão log(e)	log(e) variância
1	SYK1_ECOLI	Lysyl-tRNA synthetase OS=Escherichia coli (strain K12) GN=lysS PE=1 SV=2	699.91	0.923116348	-0.08	0.09
2	RRAB_ECOLI	Regulator of ribonuclease activity B OS=Escherichia coli (strain K12) GN=rraB PE=1 SV=1	1629.17	0.802518799	-0.22	0.22
3	DHAS_ECOLI	Aspartate-semialdehyde dehydrogenase OS=Escherichia coli (strain K12) GN=asd PE=1 SV=1	828.97	0.860707971	-0.15	0.12
4	ALDA_ECOLI	Lactaldehyde dehydrogenase OS=Escherichia coli (strain K12) GN=aldA PE=1 SV=2	1830.47	0.923116348	-0.08	0.08
5	RS19_ECOLI	30S ribosomal protein S19 OS=Escherichia coli (strain K12) GN=rpsS PE=1 SV=2	1179.4	0.771051593	-0.26	0.26
6	ODP2_ECOLI	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex OS=Escherichia coli (strain K12) GN=aceF PE=1 SV=3	969.89	0.904837417	-0.1	0.1
7	AMPP_ECOLI	Xaa-Pro aminopeptidase OS=Escherichia coli (strain K12) GN=pepP PE=1 SV=2	242.68	0.755783741	-0.28	0.36
8	GATZ_ECOLI	D-tagatose-1,6-bisphosphate aldolase subunit gatZ OS=Escherichia coli (strain K12) GN=gatZ PE=1 SV=1	7016.36	0.951229424	-0.05	0.05
9	RL22_ECOLI	50S ribosomal protein L22 OS=Escherichia coli (strain K12) GN=rplV PE=1 SV=1	750.21	0.886920439	-0.12	0.12
10	YCGM_ECOLI	Uncharacterized protein ycgM OS=Escherichia coli (strain K12) GN=ycgM PE=1 SV=1	469.19	0.543350861	-0.61	0.61
11	OPPA_ECOLI	Periplasmic oligopeptide-binding protein OS=Escherichia coli (strain K12) GN=oppA PE=1 SV=2	1768.09	0.951229424	-0.05	0.04
12	SURA_ECOLI	Chaperone surA OS=Escherichia coli (strain K12) GN=surA PE=1 SV=1	617.5	0.869358235	-0.14	0.12
13	SSPA_ECOLI	Stringent starvation protein A OS=Escherichia coli (strain K12) GN=sspA PE=1 SV=2	2559.05	0.913931182	-0.09	0.1
14	LTAE_ECOLI	Low specificity L-threonine aldolase OS=Escherichia coli (strain K12) GN=ltaE PE=1 SV=1	220.31	0.554327299	-0.59	0.43
15	ASTC_ECOLI	Succinylornithine transaminase OS=Escherichia coli (strain K12) GN=astC PE=1 SV=1	918.65	0.878095435	-0.13	0.1
16	GLNA_ECOLI	Glutamine synthetase OS=Escherichia coli (strain K12) GN=glnA PE=1 SV=2	619.69	0.860707971	-0.15	0.13
17	OPDA_ECOLI	Oligopeptidase A OS=Escherichia coli (strain K12) GN=prlC PE=3 SV=3	585.44	0.843664815	-0.17	0.14
18	LPTB_ECOLI	Lipopolysaccharide export system ATP-binding protein LptB OS=Escherichia coli (strain K12) GN=lptB PE=1 SV=2	375.41	0.444858065	-0.81	0.6
19	YGAM_ECOLI	Uncharacterized protein ygaM OS=Escherichia coli (strain K12) GN=ygaM PE=4 SV=1	813.65	0.794533599	-0.23	0.23
20	PT1_ECOLI	Phosphoenolpyruvate-protein phosphotransferase OS=Escherichia coli (strain K12) GN=ptsI PE=1 SV=1	491.23	0.878095435	-0.13	0.11
21	FTSZ_ECOLI	Cell division protein ftsZ OS=Escherichia coli (strain K12) GN=ftsZ PE=1 SV=1	318.18	0.826959136	-0.19	0.18
22	RL24_ECOLI	50S ribosomal protein L24 OS=Escherichia coli (strain K12) GN=rplX PE=1 SV=2	808.21	0.810584251	-0.21	0.16

23	PURU_ECOLI	Formyltetrahydrofolate deformylase OS=Escherichia coli (strain K12) GN=purU PE=1 SV=1	646.08	0.625002269	-0.47	0.42
24	IDH_ECOLI	Isocitrate dehydrogenase [NADP] OS=Escherichia coli (strain K12) GN=icd PE=1 SV=1	17204.54	0.96078944	-0.04	0.03
25	THIO_ECOLI	Thioredoxin-1 OS=Escherichia coli (strain K12) GN=trxA PE=1 SV=2	2332.41	0.904837417	-0.1	0.07
26	DKSA_ECOLI	DnaK suppressor protein OS=Escherichia coli (strain K12) GN=dksA PE=1 SV=1	269.15	0.565525443	-0.57	0.48
27	ASSY_ECOLI	Argininosuccinate synthase OS=Escherichia coli (strain K12) GN=argG PE=1 SV=2	317.81	0.748263574	-0.29	0.26
28	RS16_ECOLI	30S ribosomal protein S16 OS=Escherichia coli (strain K12) GN=rpsP PE=1 SV=1	1044.21	0.886920439	-0.12	0.11
29	FKBA_ECOLI	FKBP-type peptidyl-prolyl cis-trans isomerase fkpA OS=Escherichia coli (strain K12) GN=fkpA PE=1 SV=1	2026.15	0.878095435	-0.13	0.08
30	Q1PI59_ECOLI	Malate dehydrogenase (Fragment) OS=Escherichia coli (strain K12) GN=mdh PE=4 SV=1	8011.39	0.852143792	-0.16	0.08
31	ACON1_ECOLI	Aconitate hydratase 1 OS=Escherichia coli (strain K12) GN=acnA PE=1 SV=3	696.51	0.618783398	-0.48	0.07
32	YEAD_ECOLI	Putative glucose-6-phosphate 1-epimerase OS=Escherichia coli (strain K12) GN=yeaD PE=1 SV=2	186.35	0.548811623	-0.6	0.39
33	ATPD_ECOLI	ATP synthase subunit delta OS=Escherichia coli (strain K12) GN=atpH PE=1 SV=1	802.63	0.740818212	-0.3	0.17
34	ABDH_ECOLI	Gamma-aminobutyraldehyde dehydrogenase OS=Escherichia coli (strain K12) GN=prp PE=1 SV=1	284.09	0.49658531	-0.7	0.22
35	EFTS_ECOLI	Elongation factor Ts OS=Escherichia coli (strain K12) GN=tsf PE=1 SV=2	1730.04	0.895834136	-0.11	0.07
36	YCAC_ECOLI	Uncharacterized protein ycaC OS=Escherichia coli (strain K12) GN=ycaC PE=1 SV=1	502.66	0.571209062	-0.56	0.27
37	TRXB_ECOLI	Thioredoxin reductase OS=Escherichia coli (strain K12) GN=trxB PE=1 SV=2	533.73	0.818730751	-0.2	0.07
38	B8LFD8_ECOLI	Galactoside O-acetyltransferase OS=Escherichia coli (strain K12) PE=4 SV=1	964.17	0.657046828	-0.42	0.26
39	ACEA_ECOLI	Isocitrate lyase OS=Escherichia coli (strain K12) GN=aceA PE=1 SV=1	5960.83	0.886920439	-0.12	0.05
40	CISY_ECOLI	Citrate synthase OS=Escherichia coli (strain K12) GN=gltA PE=1 SV=1	5894.92	0.886920439	-0.12	0.05
41	AGP_ECOLI	Glucose-1-phosphatase OS=Escherichia coli (strain K12) GN=agp PE=1 SV=1	1155.5	0.733446954	-0.31	0.12
42	SYS_ECOLI	Seryl-tRNA synthetase OS=Escherichia coli (strain K12) GN=serS PE=1 SV=1	1612.59	0.843664815	-0.17	0.11
43	SYK2_ECOLI	Lysyl-tRNA synthetase, heat inducible OS=Escherichia coli (strain K12) GN=lysU PE=1 SV=2	1702.94	0.895834136	-0.11	0.1
44	TKT2_ECOLI	Transketolase 2 OS=Escherichia coli (strain K12) GN=tktB PE=1 SV=1	246.74	0.802518799	-0.22	0.16
45	GLRX3_ECOLI	Glutaredoxin-3 OS=Escherichia coli (strain K12) GN=grxC PE=1 SV=2	3126.09	0.895834136	-0.11	0.08
46	WRBA_ECOLI	Flavoprotein wrbA OS=Escherichia coli (strain K12) GN=wrbA PE=1 SV=2	4187.83	0.810584251	-0.21	0.07
47	MDH_ECOLI	Malate dehydrogenase OS=Escherichia coli (strain K12) GN=mdh PE=1 SV=1	18302.48	0.852143792	-0.16	0.04
48	MALE_ECOLI	Maltose-binding periplasmic protein OS=Escherichia coli (strain K12) GN=maleE PE=1 SV=1	3327.67	0.740818212	-0.3	0.06
49	YIHD_ECOLI	Protein yihD OS=Escherichia coli (strain K12) GN=yihD PE=1 SV=1	5496.99	0.677056884	-0.39	0.11
50	LPP_ECOLI	Major outer membrane lipoprotein Lpp OS=Escherichia coli (strain K12) GN=lpp PE=1 SV=1	27208.61	0.733446954	-0.31	0.11
51	ADHE_ECOLI	Aldehyde-alcohol dehydrogenase OS=Escherichia coli (strain K12) GN=adhE PE=1 SV=2	546.98	0.794533599	-0.23	0.1

52	GLRX2_ECOLI	Glutaredoxin-2 OS=Escherichia coli (strain K12) GN=grxB PE=1 SV=1	2369.54	0.71177032	-0.34	0.07
53	AMPA_ECOLI	Cytosol aminopeptidase OS=Escherichia coli (strain K12) GN=pepA PE=1 SV=1	178.2	0.755783741	-0.28	0.19
54	PTM3C_ECOLI	PTS system mannitol-specific EIICBA component OS=Escherichia coli (strain K12) GN=mtlA PE=1 SV=1	518.79	0.697676316	-0.36	0.14
55	DHAK_ECOLI	PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit dhaK OS=Escherichia coli (strain K12) GN=dhaK PE=1 SV=2	342.1	0.506616989	-0.68	0.35
56	G3P1_ECOLI	Glyceraldehyde-3-phosphate dehydrogenase A OS=Escherichia coli (strain K12) GN=gapA PE=1 SV=2	8032.08	0.869358235	-0.14	0.04
57	YBAY_ECOLI	Uncharacterized lipoprotein ybaY OS=Escherichia coli (strain K12) GN=ybaY PE=1 SV=1	845.26	0.786627865	-0.24	0.13
58	THGA_ECOLI	Galactoside O-acetyltransferase OS=Escherichia coli (strain K12) GN=lacA PE=1 SV=1	964.17	0.65050909	-0.43	0.25
59	USPA_ECOLI	Universal stress protein A OS=Escherichia coli (strain K12) GN=uspA PE=1 SV=2	9967.71	0.543350861	-0.61	0.1
60	YAHK_ECOLI	Uncharacterized zinc-type alcohol dehydrogenase-like protein YahK OS=Escherichia coli (strain K12) GN=yahK PE=1 SV=1	513.43	0.336216482	-1.09	0.4
61	ALF_ECOLI	Fructose-bisphosphate aldolase class 2 OS=Escherichia coli (strain K12) GN=fbaA PE=1 SV=2	2783.87	0.843664815	-0.17	0.05
62	RL29_ECOLI	50S ribosomal protein L29 OS=Escherichia coli (strain K12) GN=rpmC PE=1 SV=1	11781.31	0.522045789	-0.65	0.07
63	HSLO_ECOLI	33 kDa chaperonin OS=Escherichia coli (strain K12) GN=hslo PE=1 SV=1	204.48	0.625002269	-0.47	0.24
64	YGAU_ECOLI	Uncharacterized protein ygaU OS=Escherichia coli (strain K12) GN=ygaU PE=1 SV=2	3276.12	0.690734327	-0.37	0.1
65	BFR_ECOLI	Bacterioferritin OS=Escherichia coli (strain K12) GN=bfr PE=1 SV=1	2027.3	0.63128364	-0.46	0.08
66	KBL_ECOLI	2-amino-3-ketobutyrate coenzyme A ligase OS=Escherichia coli (strain K12) GN=kbl PE=1 SV=1	1386.23	0.810584251	-0.21	0.11
67	ENO_ECOLI	Enolase OS=Escherichia coli (strain K12) GN=eno PE=1 SV=2	12638.86	0.794533599	-0.23	0.05
68	DPS_ECOLI	DNA protection during starvation protein OS=Escherichia coli (strain K12) GN=dps PE=1 SV=2	24402.57	0.690734327	-0.37	0.04
69	SYE_ECOLI	Glutamyl-tRNA synthetase OS=Escherichia coli (strain K12) GN=gltX PE=1 SV=1	198.95	0.65050909	-0.43	0.22
70	YAHO_ECOLI	Uncharacterized protein YahO OS=Escherichia coli (strain K12) GN=yahO PE=3 SV=1	4590.46	0.690734327	-0.37	0.16
71	OSMY_ECOLI	Osmotically-inducible protein Y OS=Escherichia coli (strain K12) GN=osmY PE=1 SV=1	1015.85	0.802518799	-0.22	0.1
72	KATG_ECOLI	Catalase-peroxidase OS=Escherichia coli (strain K12) GN=katG PE=1 SV=2	1761.56	0.60653066	-0.5	0.07
73	YTFQ_ECOLI	ABC transporter periplasmic-binding protein ytfQ OS=Escherichia coli (strain K12) GN=ytfQ PE=1 SV=1	4477.53	0.786627865	-0.24	0.09
74	YJGF_ECOLI	RutC family protein yjgF OS=Escherichia coli (strain K12) GN=yjgF PE=1 SV=2	3367.23	0.65050909	-0.43	0.09
75	UXUA_ECOLI	Mannonate dehydratase OS=Escherichia coli (strain K12) GN=uxuA PE=1 SV=2	389.4	0.543350861	-0.61	0.45
76	DEOB_ECOLI	Phosphopentomutase OS=Escherichia coli (strain K12) GN=deoB PE=1 SV=1	1191.85	0.826959136	-0.19	0.07
77	PEPD_ECOLI	Aminoacyl-histidine dipeptidase OS=Escherichia coli (strain K12) GN=pepD PE=1 SV=3	1568.68	0.786627865	-0.24	0.08
78	DPPA_ECOLI	Periplasmic dipeptide transport protein OS=Escherichia coli (strain K12) GN=dppA PE=1 SV=1	458.07	0.594520559	-0.52	0.17

79	RISB_ECOLI	6,7-dimethyl-8-ribityllumazine synthase OS=Escherichia coli (strain K12) GN=ribH PE=1 SV=1	371.52	0.576949804	-0.55	0.3
80	AHPC_ECOLI	Alkyl hydroperoxide reductase subunit C OS=Escherichia coli (strain K12) GN=ahpC PE=1 SV=2	6635.46	0.878095435	-0.13	0.05
81	PFLB_ECOLI	Formate acetyltransferase 1 OS=Escherichia coli (strain K12) GN=pflB PE=1 SV=2	679.36	0.810584251	-0.21	0.09
82	OSMC_ECOLI	Peroxiredoxin osmC OS=Escherichia coli (strain K12) GN=osmC PE=1 SV=2	3194.39	0.618783398	-0.48	0.12
83	YJBJ_ECOLI	UPF0337 protein yjbj OS=Escherichia coli (strain K12) GN=yjbj PE=1 SV=1	6280.05	0.532591804	-0.63	0.1
84	PTKB_ECOLI	Galactitol-specific phosphotransferase enzyme IIB component OS=Escherichia coli (strain K12) GN=gatB PE=1 SV=2	4559.65	0.794533599	-0.23	0.1
85	PTGA_ECOLI	Glucose-specific phosphotransferase enzyme IIA component OS=Escherichia coli (strain K12) GN=crr PE=1 SV=2	5049.42	0.826959136	-0.19	0.07
86	PGK_ECOLI	Phosphoglycerate kinase OS=Escherichia coli (strain K12) GN=pgk PE=1 SV=2	17083.44	0.771051593	-0.26	0.04
87	ALDB_ECOLI	Aldehyde dehydrogenase B OS=Escherichia coli (strain K12) GN=aldB PE=1 SV=2	778.92	0.431710535	-0.84	0.18
88	CLPP_ECOLI	ATP-dependent Clp protease proteolytic subunit OS=Escherichia coli (strain K12) GN=clpP PE=1 SV=1	636.54	0.748263574	-0.29	0.13
89	GMHA_ECOLI	Phosphoheptose isomerase OS=Escherichia coli (strain K12) GN=gmhA PE=1 SV=1	299.2	0.394553708	-0.93	0.12
90	TPIS_ECOLI	Triosephosphate isomerase OS=Escherichia coli (strain K12) GN=tpiA PE=1 SV=1	449.51	0.748263574	-0.29	0.12
91	ACON2_ECOLI	Aconitate hydratase 2 OS=Escherichia coli (strain K12) GN=acnB PE=1 SV=3	5886.92	0.860707971	-0.15	0.03
92	ARGT_ECOLI	Lysine-arginine-ornithine-binding periplasmic protein OS=Escherichia coli (strain K12) GN=argT PE=1 SV=3	966.2	0.755783741	-0.28	0.13
93	GLTI_ECOLI	Glutamate/aspartate periplasmic-binding protein OS=Escherichia coli (strain K12) GN=gltI PE=1 SV=2	1015.13	0.860707971	-0.15	0.08
94	SODM_ECOLI	Superoxide dismutase [Mn] OS=Escherichia coli (strain K12) GN=sodA PE=1 SV=2	1518.27	0.444858065	-0.81	0.14
95	FABH_ECOLI	3-oxoacyl-[acyl-carrier-protein] synthase 3 OS=Escherichia coli (strain K12) GN=fabH PE=1 SV=1	433.73	0.571209062	-0.56	0.3
96	KPYK1_ECOLI	Pyruvate kinase I OS=Escherichia coli (strain K12) GN=pykF PE=1 SV=1	758.96	0.740818212	-0.3	0.1
97	AGAL_ECOLI	Alpha-galactosidase OS=Escherichia coli (strain K12) GN=meIA PE=1 SV=1	3288.49	0.543350861	-0.61	0.08
98	YQJD_ECOLI	Uncharacterized protein yqjD OS=Escherichia coli (strain K12) GN=yqjD PE=4 SV=1	24982.8	0.367879441	-1	0.14

ANEXO G – Artigos científicos publicados e participações em eventos.

microbial biotechnology

Open Access

Minireview

Recombinant DNA production of spider silk proteins

Olena Tokareva,^{1†} Valquíria A. Michalczechen-Lacerda,² Elíbio L. Rech³ and David L. Kaplan^{1*}¹Department of Biomedical Engineering, Tufts University, Medford, MA 02155, USA.²Department of Cell Biology, Campus Universitario Darcy Ribeiro, Institute of Biology, University of Brasilia, Brasilia, DF 70910-900, Brazil.³Embrapa Genetics Resources and Biotechnology, Biotechnology Unit, Parque Estação Biológica PqEB W5 Norte, Brasília 70770-900, DF Brazil.

Abstract

Spider dragline silk is considered to be the toughest biopolymer on Earth due to an extraordinary combination of strength and elasticity. Moreover, silks are biocompatible and biodegradable protein-based materials. Recent advances in genetic engineering make it possible to produce recombinant silks in heterologous hosts, opening up opportunities for large-scale production of recombinant silks for various biomedical and material science applications. We review the current strategies to produce recombinant spider silks.

Introduction

Spider silks have been a focus of research for almost two decades due to their outstanding mechanical and biophysical properties. Spider silks are remarkable natural polymers that consist of three domains: a repetitive middle core domain that dominates the protein chain, and non-repetitive N-terminal and C-terminal domains. The large core domain is organized in a block copolymer-like arrangement, in which two basic sequences, crystalline [poly(A) or poly(GA)] and less crystalline (GGX or GPGXX) polypeptides alternate. At least seven different types of silk proteins are known for one orb-weaver species of spider

(Lewis, 2006a). Silks differ in primary sequence, physical properties and functions (Hu *et al.*, 2006). For example, dragline silks used to build frames, radii and lifelines are known for outstanding mechanical properties including strength, toughness and elasticity (Gosline *et al.*, 1984). On an equal weight basis, spider silk has a higher toughness than steel and Kevlar (Vepari and Kaplan, 2007; Heim *et al.*, 2009). Flagelliform silk found in capture spirals has extensibility of up to 500%. Minor ampullate silk, which is found in auxiliary spirals of the orb-web and in prey wrapping, possesses high toughness and strength almost similar to major ampullate silks, but does not supercontract in water. Figure 1 depicts the location and structural elements of MaSp, MiSp and Flag silks.

Finally, there are other silk types such as aciniform, pyriform, aggregate and tubuliform (egg case) with unusual primary structure, composition and properties. Diverse and unique biomechanical properties together with biocompatibility and a slow rate of degradation make spider silks excellent candidates as biomaterials for tissue engineering, guided tissue repair and drug delivery, for cosmetic products (e.g. nail and hair strengthener, skin care products), and industrial materials (e.g. nanowires, nanofibres, surface coatings).

Recent advances in genetic engineering have provided a route to produce various types of recombinant spider silks (Prince *et al.*, 1995; Fahnestock and Bedzyk, 1997; Rabotyagova *et al.*, 2009; Xu *et al.*, 2012). However, production of spider silk proteins at a larger scale remains challenging. Moreover, recombinant silk threads do not recapitulate the full potential of native fibres in terms of mechanical properties. Different heterologous host systems have been investigated to develop suitable production systems. In this review, we discuss recent advances in the production of recombinant spider silks in heterologous host systems with the main focus on microbial production. In particular, we focus on dragline silks. Current cloning strategies, expression systems and purification strategies will be discussed to help researchers to engineer customized synthetic spider silk-like proteins for various needs, including biomaterials and material science applications.

Structure of silk proteins

Spider silks are fascinating polymers, as is the spinning process that members of *Araneidae* family use to make

Received 15 June, 2013; revised 31 July, 2013; accepted 1 August, 2013. *For correspondence. E-mail david.kaplan@tufts.edu; Tel. (+617) 626 3251; Fax (+617) 627 3231. †Published previously as O. S. Rabotyagova.

Microbial Biotechnology (2013) 6(6), 651–663
doi:10.1111/1751-7915.12081

Funding Information The authors thank the NIH (EB014976, EB002520; EY020856; DE017207; EB014283), the National Council for Scientific and Technological Development, Embrapa Genetics Resources and Biotechnology, and Foundation for Research Support, Brasilia, DF, Brazil for support of the studies reviewed here.

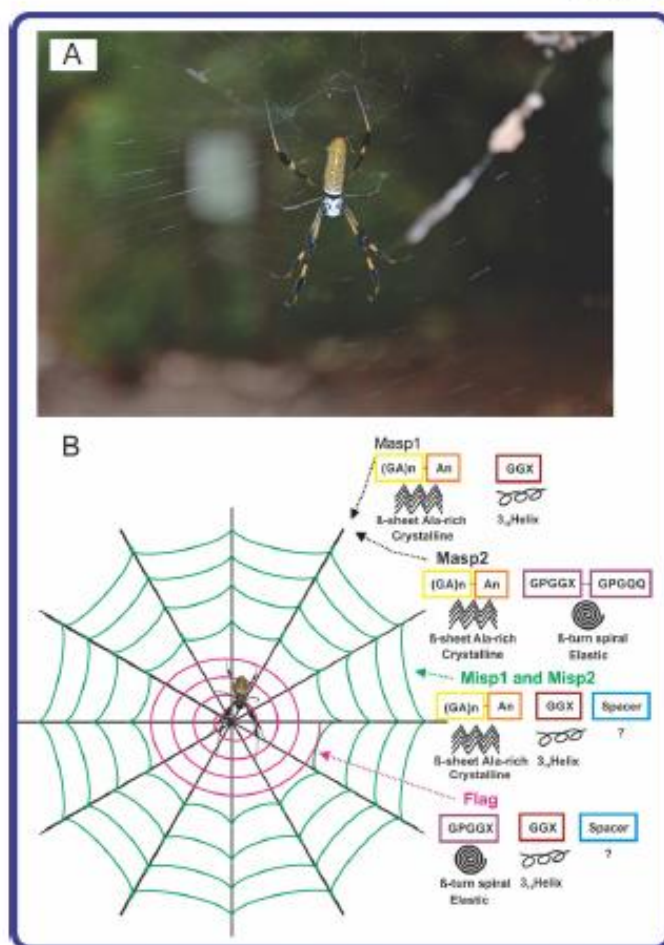


Fig. 1. A. An adult female orb weaver spider *Nephila clavipes* and her web. B. Schematic overview of *N. clavipes* web composed of three different spider silk proteins and their structures. The coloured boxes indicate the structural motifs in silk proteins. An empty box marked '?' indicates that the secondary structure of the 'spacer' region is unknown. Note: MaSp1 or MaSp2: major ampullate spidroin 1 or 2; MiSp1 and 2: minor ampullate spidroin 1 and 2; Flag: flagelliform protein. The photo was taken by Olena and Artem Tokarev in the Florida Keys.

these exceptional materials. Spiders use complex spinning to rapidly transform water soluble, high molecular weight, silk proteins into solid fibres at ambient temperature and pressure, giving rise to an environmentally safe, biodegradable and high performance material (Asakura *et al.*, 2007; Lewicka *et al.*, 2012; Teulé *et al.*, 2012a). The details on anatomy and physiology of the spider spinning apparatus (*N. clavipes*) can be found elsewhere (Knight and Vollrath, 2001; 2002; Eisoldt *et al.*, 2011; Rising *et al.*, 2011).

In order to understand the challenges and needs associated with biotechnological production of recombinant spider silks, primary protein motifs, composition and secondary structural elements must be discussed. As mentioned earlier, one spider is capable of producing up to seven different types of silks with varying mechanical properties. In spite of different mechanical and physiological properties, the majority of spider silks share a common primary structural pattern comprised of a large central core of repetitive protein domains flanked by non-repetitive N-

and C-terminal domains. The most investigated silk is dragline silk, which shows a remarkable combination of strength and elasticity. The golden orb-weaver spider, *N. clavipes*, produces dragline silk in the major ampullate gland (Knight and Vollrath, 2001). Dragline silk is the protein complex composed of major ampullate dragline silk protein 1 (MaSp1) and major ampullate dragline silk protein 2 (MaSp2). Both silks are approximately 3500 amino acid long. MaSp1 can be found in the fibre core and the periphery, whereas MaSp2 forms clusters in certain core areas. The large central domains of MaSp1 and MaSp2 are organized in block copolymer-like arrangements, in which two basic sequences, crystalline [poly(A) or poly(GA)] and less crystalline (GGX or GPGXX) polypeptides alternate in core domain. The main difference between MaSp1 and MaSp2 is the presence of proline (P) residues accounting for 15% of the total amino acid content in MaSp2 (Hu *et al.*, 2006), whereas MaSp1 is proline-free. By calculating the number of proline residues in *N. clavipes* dragline silk, it is possible to estimate the presence of the two proteins in fibres; 81% MaSp1 and 19% MaSp2 (Brooks *et al.*, 2005). Different spiders have different ratios of MaSp1 and MaSp2. For example, a dragline silk fibre from the orb weaver *Argiope aurantia* contains 41% MaSp1 and 59% MaSp2 (Huemmerich *et al.*, 2004). Such changes in the ratios of major ampullate silks can dictate the performance of the silk fibre (Vollrath and Knight, 1999). Specific secondary structures have been assigned to poly(A)/(GA), GGX and GPGXX motifs including β -sheet, 3_0 -helix and β -spiral respectively (Humenik *et al.*, 2011). The primary sequence, composition and secondary structural elements of the repetitive core domain are responsible for mechanical properties of spider silks; whereas, non-repetitive N- and C-terminal domains are essential for the storage of liquid silk dope in a lumen and fibre formation in a spinning duct (Itah *et al.*, 2006). The primary amino acid sequence, composition and secondary structural elements of other silk types are reviewed elsewhere (Lewis, 2006b; Humenik *et al.*, 2011).

Production of recombinant silk proteins

Spiders cannot be farmed, in contrast to silkworms, due to their aggressive behaviour and territorial nature (Kluge *et al.*, 2008). Collecting silk from webs is a time-consuming task. It took 8 years to make a golden spider silk cape from 1.2 million golden orb webs (Chung *et al.*, 2012). Therefore, biotechnological production of recombinant spider silks is the only practicable solution to harvest silks on a larger scale and to meet growing needs of medicine and biotechnology. A variety of heterologous host systems have been explored to produce different types of recombinant silks (Table 1 and Table 2). Recombinant partial spidroins as well as engineered silks have been cloned and

expressed in bacteria (*Escherichia coli*), yeast (*Pichia pastoris*), insects (silkworm larvae), plants (tobacco, soybean, potato, *Arabidopsis*), mammalian cell lines (BHT/hamster) and transgenic animals (mice, goats).

Unicellular organisms as heterologous host systems.

Unicellular organisms, such as bacteria and yeast, have been investigated as host systems for recombinant silks. A gram-negative, rod-shaped bacterium *E. coli* is a well-established host for industrial scale production of proteins. Therefore, the majority of recombinant spider silks have been produced in *E. coli* (Lewis *et al.*, 1996; Fahnestock and Irwin, 1997; Wang *et al.*, 2006; Rabotyagova *et al.*, 2009; Rabotyagova *et al.*, 2010; An *et al.*, 2011; An *et al.*, 2012; Teulé *et al.*, 2012a). *E. coli* is easy to manipulate, has a short generation time, is relatively low cost and can be scaled up for larger amounts protein production. The recombinant DNA approach enables the production of recombinant spider silks with programmed sequences, secondary structures, architectures and precise molecular weight (Rabotyagova *et al.*, 2011). There are four main steps in the process: (i) design and assembly of synthetic silk-like genes into genetic 'cassettes', (ii) insertion of this segment into a DNA vector, (iii) transformation of this recombinant DNA molecule into a host cell and (iv) expression and purification of the selected clones. Figure 2 summarizes the recombinant DNA approach used to prepare silk-like proteins.

The monomeric silk-like gene sequences can be synthesized as short single-stranded oligonucleotides (up to 100 bp) by commercial oligonucleotide synthesis or used directly as polymerase chain reaction products from cDNA libraries. Large repetitive sequences can be constructed by using concatemerization, step-by-step directional approach and recursive ligation (Fig. 3). Concatemerization is a useful method when a library of genes of different sizes is desired but has limitations in the preparation of genes with specific sizes (Meyer and Chilkoti, 2002). To overcome limitations of concatemerization, recursive directional ligation or a step-by-step ligation is employed (Meyer and Chilkoti, 2002; Wright and Conticello, 2002). Recursive directional ligation allows for facile modularity, where control over the size of the genetic cassettes is achieved. Moreover, recursive directional ligation eliminates the restriction sites at the junctions between monomeric genetic cassettes without interrupting key gene sequences with additional base pairs that makes it different from the step-by-step ligation approach (Higashiya *et al.*, 2007).

For example, we have employed step-by-step directional ligation to produce various partial recombinant spider silks as well as engineered silk-like proteins based on the sequences of dragline silk originated from *N. clavipes* (Prince *et al.*, 1995; Wang *et al.*, 2006; Huang

Table 1. Summary of recombinantly expressed spider silks in *Escherichia coli*. Spider silk origin, number of monomers, molecular weight, cloning and expression plasmids as well as restriction enzymes and purification strategies used to produce recombinant silks are shown.

Type	Host	Origin	Protein	Number of monomers	MW (kDa)	Cloning plasmid	RE	Expression plasmid	Purification Strategy	Yield (mg L ⁻¹)	References
Bacteria	<i>E. coli</i>	<i>Nephila clavipes</i>		15, 32, 64, 96	85, 100, 193, 285	pET30a(+)	<i>Nhe I</i> / <i>Spe I</i>	pET30a(+)	Ammonium sulphate		Xia et al., 2010
			MaSp 1	6; 15	16, 36.5			pETNX	PA	96.8 mg L ⁻¹ ; 200 mg L ⁻¹ ; 300 mg L ⁻¹	Dams-Kozbawska et al., 2012 Fahnestock and Iwih, 1997 Fahnestock and Iwih, 1997
				8; 16	65–163	pBR322 derived	<i>Pst I</i>	pFP202 (pET19a + pET11a) pFP202, pFP204, IMAC	IMAC	N.A. ¹ 25 mg ml ⁻¹	An et al., 2011 Rabotyagova et al., 2009 Fahnestock and Iwih, 1997
			poly(A) and GGX	16, 24 1, 2, 3, 6	46, 70 10, 18	pBSSKII+ pET30a(+)	<i>AvrII</i> , <i>NheI</i> / <i>Spe I</i> / <i>NotI</i>	pET19k pET30a(+)		300 mg L ⁻¹ 10 mg g ⁻¹	Lewis et al., 1996 Brooks et al., 2008
				8; 16	65–163	pBR322 derived	<i>Pst I</i>	pFP202, pFP204, or pFP207	IMAC		Brooks et al., 2008
				8, 16, 32 16	31, 58, 112 63	pBBSK	<i>ScaI</i> / <i>XmaI</i> / <i>BspEI</i>	pET19k pET30a(+)		N.A. ¹	Brooks et al., 2008
				12	71	pET30a(+)	N.A. ¹	pET30a(+)			Brooks et al., 2008
				8	67						Brooks et al., 2008
				24/16	62/47	pBSSKII+	<i>XmaI</i> / <i>ScaI</i> / <i>BspEI</i>	pET19k	IMAC	120 mg L ⁻¹	An et al., 2011
				1x-18x	15, 23, 32, 41	pUC18	<i>Spe I</i> / <i>NotI</i>	pQE-9		15, 7, 3, 2 mg L ⁻¹	Price et al., 1995
				2, 3, 4	19, 38, 51.7, 76.1	pET32		pET32	IMAC	80 mg L ⁻¹ ; 22 mg L ⁻¹ ; 40 mg L ⁻¹	Xu et al., 2012
				11	190			<i>Bam</i> HI/ <i>Bgl</i> II/ <i>Bsp</i> RI/ <i>Xma</i> I/ <i>Pvu</i> II/ <i>Bsp</i> E1			Widmayer et al., 2009
				1x-3x	30–56	pL2	<i>Hind</i> III/ <i>Xba</i> I	pTRONa_Cm	SEC	N.A. ¹	Widmayer et al., 2009
				1x-3x							Widmayer et al., 2009
				1x-3x							Widmayer et al., 2009
				8, 16	65	pBR322 derived	<i>Pst I</i>	pFP684	Ammonium sulphate	665 mg L ⁻¹	Fahnestock and Bedzyk, 1997

Table 2. Summary of recombinantly expressed spider silks in insects, plants and mammals. Spider silk origin, number of monomers, molecular weight, cloning and expression plasmid as well as restriction enzymes and purification strategies used to produce recombinant silks are shown.

Type	Host	Origin	Protein	Number of monomers	MW (kDa)	Cloning plasmid	RE	Expression plasmid	Purification strategy	Yield (mg L ⁻¹)	References
Insects	<i>B. mori</i>	<i>Nephila clavipes</i>	MspI	2	85	pSL1180a	SpeI/NdeI	pBa13xP3-DeRedII	IMAC	N.A. ¹	Wen et al., 2010
			MspI-Flag	4	70	pSL1180	SpeI/XbaI/	pFastBacHTC		6 mg/larva	Zhong et al., 2009
Plants	Nicotiana glaucum Solanum tuberosum Nicotiana benthamiana	<i>Nephila clavipes</i>	MspI-Flag	multiple	75–130	pBSSKII+ and pSL1180a	SpeI/XbaI/ BspEI	pBaC(3xP3-DeRedII)		N.A. ¹	Tadé et al., 2012b
			MspI	multiple	12.9–68.8	pUC19	AgeI/NotI/ HindIII	pRTR473	(NH ₄) ₂ SO ₄ , 10–60% saturation	0.5–2 % total protein	Scheller et al., 2001
			Flag (Intein)	4; 10	47, 72, 100, 250	pRTR15	splicing events	pCB301-Kan	IMAC	1.9 mg/50 g leaf material; 0.34%; 0.03% in leaves; 1.2% in seeds; 0.78% in 1% in somatic embryos	Hauptmann et al., 2013
Mammals	Transgenic mice COS-1 cells	<i>Nephila clavipes</i> <i>Euglossinops sp.</i>	MspI	6, 16	64, 127	pBSSK+	AgeI/BamHI	Cong ¹ + Pha3 ²	ammonium sulphate		Barr et al., 2004
			MspI	6	31–66	pGEM-5zf	BamHI/KpnI	pBC1	centrifugation	11.7 mg L ⁻¹	Xu et al., 2007
Mammals	Baby hamster kidney	<i>Nephila clavipes</i> <i>Araneus diadematus</i>	MspI	N.A. ¹	28, 22	pER-14	BamHI/ EcoRV	pSecTag2/ Hygro A	N.A. ¹	N.A. ³	Grip et al., 2006
			MspI	N.A. ¹	59, 109/ 59	pBSSK+	ApaI/SapI	CMV promoter	ammonium sulfate		
			MspI	N.A. ¹	63, 60, 110, 140	pSecTag-C	MscI/PvuI			25–50 mg L ⁻¹	Lazaris et al., 2002

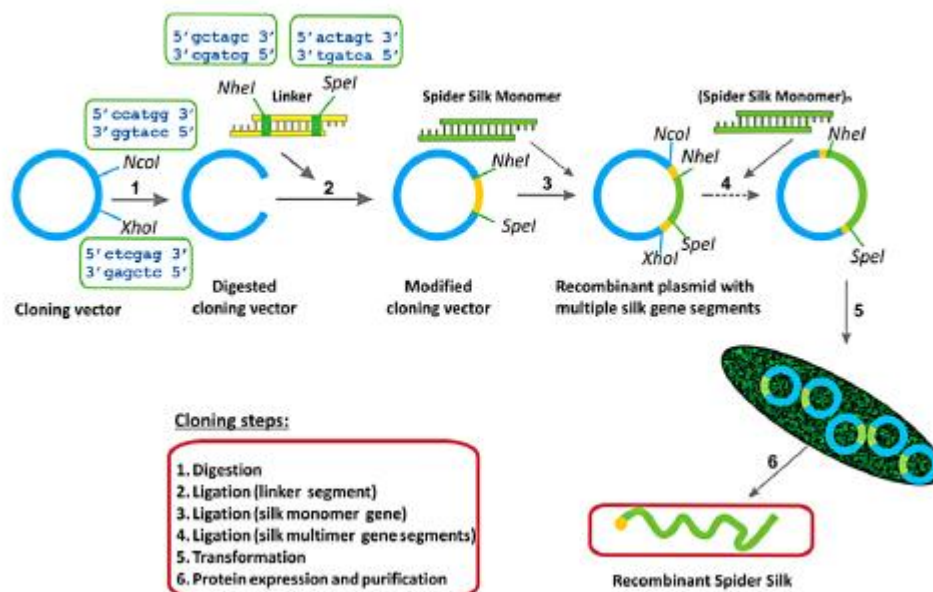


Fig. 2. Recombinant DNA approach used to prepare silk-like proteins.

et al., 2007; Rabotyagova *et al.*, 2009; 2010; Mieszawska *et al.*, 2010; Gomes *et al.*, 2011; Numata *et al.*, 2012). As one example, spider silk block copolymers were generated in *E. coli* (Rabotyagova *et al.*, 2009; 2010). In the first cloning step, a commercially available pET30a(+) vector (Novagen, San Diego, CA, USA) was modified with an adaptor sequence, carrying *NheI* and *SpeI* restriction sites. The adaptor was inserted into *XhoI* and *NcoI* sites of a pET30a(+) to generate pET30L. The coding sequences of two spider silk-like monomers A (hydrophobic block) and B (hydrophilic) were designed to carry *SpeI* and *NheI* restriction sites at the ends of the sequences. This allowed ligation of the domains into a pET30L vector. By using a step-by-step directional ligation approach, direct control over the assembly of monomeric genes into complex sequences was achieved. Six different constructs were cloned and transformed into the bacterial host for expression. An N-terminal His-tag was used for protein purification by immobilized metal affinity chromatography (Rabotyagova *et al.*, 2009).

Another genetic engineered strategy has been proposed by Lewis Laboratory to assemble long repetitive spider silk genes (Teule *et al.*, 2009). This cloning strategy employs a one-step head-to-tail ligation that can produce large inserts in precise manner (Lewis *et al.*, 1996; Brooks

et al., 2008; Teule *et al.*, 2009; Teulé *et al.*, 2012a). The spider silk synthetic genes were optimized for codon usage in *E. coli* and were cloned into a plasmid vector pBluescriptII SK(+) (Stratagene). Each silk module was carrying compatible *XmaI* and *BspEI* restriction sites at the ends on the coding sequences. The vector also contained a unique restriction site (*ScaI*) in the ampicillin resistance gene. By simultaneously performing two double digestion reactions *ScaI* – *XmaI* and *ScaI* – *BspEI* two fragments each containing a copy of a silk monomer gene were obtained. The fragments were ligated together using T4 ligase resulting in the doubling of the size of silk genes and restoring the ampicillin resistance of the plasmid (Fig. 4). Several round of cloning were performed to obtain repetitive sequences of a desired size. Next, the multimeric synthetic genes were subcloned into an expression pET19b vector using *NdeI* and *BamHI* restriction sites. Since the expression vector was carrying *NdeI* and *BamHI* sites, the liberated inserts were cloned in-frame with pET19b. Similar to pET30L, silk genes in pET19b are under control of the T7 promoter and require the addition of isopropyl- β -D-1-thiogalactopyranoside to initiate protein expression. The expressed proteins can be purified by immobilized metal affinity chromatography (IMAC) due to the presence of an N-terminal His-tag.

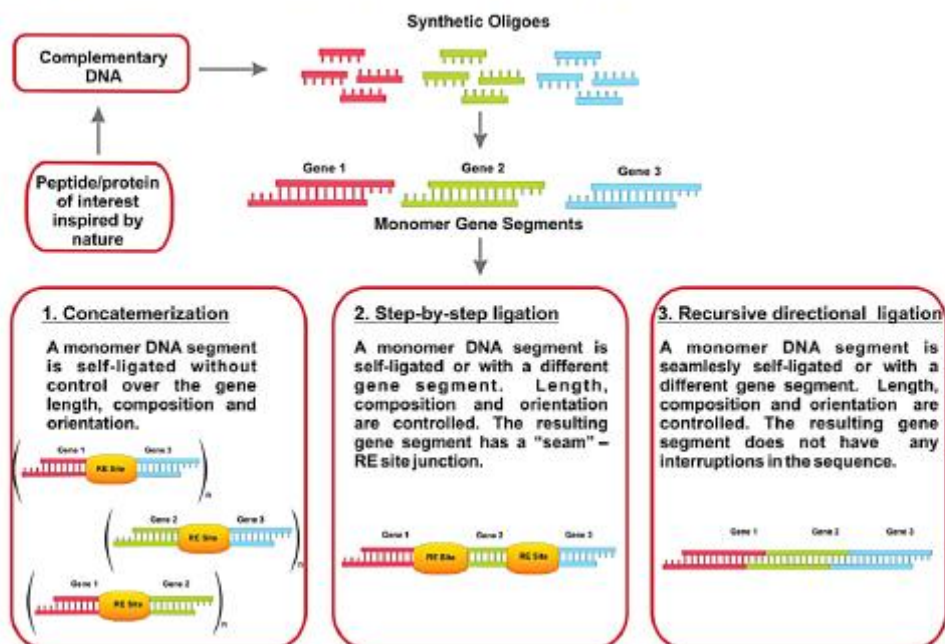


Fig. 3. Gene multimerization approaches. Note: RE Site stands for a restriction enzyme site.

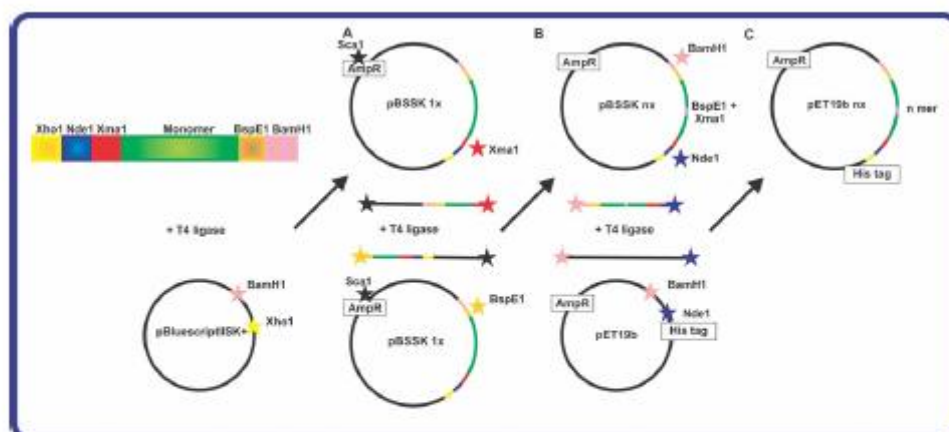


Fig. 4. Cloning strategy used by the Lewis group to engineer long repetitive spider silk sequences (n mer).

A. Cloning of a silk monomer into the vector pBluescript II SK⁺.

B. The resulting plasmid is double digested and fragments containing silk monomers are ligated again to produce longer sequences.

C. The synthetic spider silk multimer is ligated into pET19b expression vector. Note: Restriction digestion sites are indicated by star. Adapted from reference (Toule et al., 2009).

Several recombinant spider silk proteins from different species were produced using this genetic engineering strategy including silks from *N. clavipes* (Teule *et al.*, 2009) *Argiope aurantia* (Brooks *et al.*, 2008). Recombinant spider silk proteins from *Nephylengys cruentata*, *Parawixia bistriata* and *Avicularia juruensis* were produced employing this cloning strategy (Leopoldo *et al.*, 2007) (US patent 20 100 311 645). Figure 4 summarizes the strategy.

A three module cloning strategy based on the sequences of ADF-3 and ADF-4 was developed by Scheibel research group (Huemmerich *et al.*, 2004), designed so that multiple modules can be combined. Moreover, additional coding sequences such as N- or C-terminal domains can be added if needed. The purification protocol is based on heat resistance of silk proteins followed by an ammonium sulphate precipitation that is different from Ni-NTA IMAC.

Different purification strategies have been employed recently to optimize small and large-scale production of recombinant silks. Most of the spider silk proteins are produced with an N- or C-terminal His-tags to make purification simple and produce enough amounts of the protein. However, the presence of this tag can affect protein secondary structure and interfere with the process of spider silk fibre formation. Dams-Kozłowska *et al.* (2012) proposed two strategies to purify spider silks from lysates without the use of a His-tag. These protocols are based on thermal treatment and organic acid resistance of silk proteins and do not require the presence of the His-tag. After purification, silk proteins based on MaSp1 gene sequence were formed into films that subsequently were used to grow murine fibroblast cell culture. The results demonstrated that silk films were non-toxic to the cells (Dams-Kozłowska *et al.*, 2012).

Because of the highly repetitive core sequence of spider silk genes, frequent homologous recombination, deletions, transcription errors, translation pauses, accumulation in inclusion bodies and low yields were observed during the production of recombinant silks in *E. coli*. Moreover, when the protein size was increased from 43 kDa to higher (the size of native spidroins is between 300 and 350 kDa), protein yields decreased dramatically. Codon optimization for the specific host expression system helped maximize the translation of the foreign gene transcripts and thus, improved protein yields (Fahnestock and Bedzyk, 1997; Lewis, 2006b). It was also suggested that depletion of tRNA pools upon protein expression resulted in transcription and translation errors (Rosenberg *et al.*, 1993). Recently, Xia *et al.* (2010) employed a metabolic engineering strategy to enhance the production of recombinant spider silks. The authors reported production of full length (284.9 kDa) recombinant *N. clavipes* dragline silk proteins that were rich in glycine (43–45%). Production of these silk

proteins was enhanced by the use of the metabolically engineered expression host within which the glycyl-tRNA pool was elevated. The fibres spun with the native-sized recombinant spider silk protein showed tenacity, elongation and Young's modulus of 508 MPa, 15% and 21 GPa, respectively, comparable to those of native spider dragline silk (Xia *et al.*, 2010). Through extensive proteomic analysis, serine hydroxymethyltransferase (GlyA) and β -subunit of glycyl-tRNA synthetase (GlyS) were found to be upregulated to meet the high cellular demand for glycyl-tRNA when expressing glycine-rich silk proteins. Increased glycine biosynthetic flux by overexpressing glycyl-tRNA synthetase elevated the total tRNA^{Gly} pool and resulted in enhanced production of high molecular weight recombinant spider silks.

Recently, large spider recombinant egg case silk protein from *Nephila antipodiana*, 378 kDa, was engineered using *E. coli*, where gene multimers were chemically linked by cysteine disulfide bonds. The recombinant silk sequence consisted of two silk proteins: tubuliform spidroin 1 (TuSp1) and C-terminal domain of MisP1. Non-repetitive C-terminal domain of MisP1 was chosen due to its higher water solubility and stability compared with the C-terminal domain of TuSp1. A disulfide linkage between two C-terminal domains was formed by introducing a point mutation (S76 to S76C). This link allowed the formation of a hybrid DNA construct that was expressed in *E. coli* (DE3). The recombinant protein was expressed in *E. coli*. Moreover, the artificial fibres spun from this protein showed higher tensile strength and Young's modulus than natural egg case protein (Lin *et al.*, 2013).

The highly repetitive silk gene arrangement and the unusual mRNA secondary structure result in inefficient translation that limits the size of the silks produced in *E. coli*. To minimize the presence of truncated silk proteins and allow the extracellular secretion of silks, the myxotrophic yeast *P. pastoris* has been used. Fahnestock and Bedzyk (1997) produced *N. clavipes* spider dragline silks in yeast *P. pastoris*. Synthetic genes were expressed at high levels under control of the methanol-inducible *AOX1* promoter. Transformants containing multiple gene copies produced elevated levels of silk protein. Results demonstrated that *P. pastoris* can be used to successfully produce long repetitive proteins (Fahnestock and Bedzyk, 1997).

Spider silks from *Araneus diadematus* (ADF-1, 2 and 3) have also been expressed using the type III secretion system of a gram-negative, non-spore-forming, enterobacterium *Salmonella*. The authors reported yield values range from 90 to 410 nmol L⁻¹h⁻¹ that is similar to 10 mg L⁻¹h⁻¹ for a protein the size of ADF-2. The results demonstrated the feasibility to use *Salmonella* for the large-scale spider silk production (Widmaier *et al.*, 2009). Mammalian cell lines, such as bovine mammary epithelial

alveolar and baby hamster kidney cells, were used to express MaSp1 and MaSp2 (Lazaris *et al.*, 2002). The cells expressed recombinant proteins; however, as size of silk gene increased, the yield decreased dramatically due to inability of mammalian cells to cope with large repetitive sequences. Several factors have attributed to the decreased yields including, but not limited to, inefficient transcription, insufficient secretion, low copy numbers and translational limitations. The produced silk proteins were spun into fibres, and their mechanical properties were tested. It was noted that those recombinant silks that were produced without a His-tag demonstrated better mechanical properties compared with fibres made of silk proteins with a His-tag (i.e. fibres were brittle). Similar problems (i.e. transcription and translation limitations) have been reported when green monkey kidney fibroblast-like cell lines (COS-1) were used to express a 636-base pair gene fragment of MaSp1 from the African spider *Euprosthenois* sp. (Grip *et al.*, 2006). Table 1 summarizes genetic engineering approaches, cloning strategies, and production yields of recombinant silk proteins produced in unicellular heterologous host systems.

Multicellular organisms as heterologous host systems. Due to the low production rate and instability (i.e. frequent homologous recombination, deletions, transcription errors, translation pauses) of spider silk repetitive genes in unicellular organisms, multicellular organisms such as insects, plants and mammals have been studied for production of recombinant spider silk proteins.

Silkworms (*B. mori*) can be farmed and produce cocoons containing large quantities of silkworm silk known as fibroin (Vepari and Kaplan, 2007; Hu and Kaplan, 2011). Moreover, to produce a solid thread, silkworms employ a spinning process that is similar to that used by spiders to make dragline silk. The presence of a natural silk production system in silkworms makes them excellent candidates to investigate as heterologous hosts for spider silk production. There have been several reports of the transfer of silk genes from spiders to silkworms (Motohashi *et al.*, 2005; Zhang *et al.*, 2011; Teulé *et al.*, 2012b).

Baculovirus-based expression systems have been used to introduce silk genes into a heterologous host. Baculovirus infects silkworms and allows for production of large quantities of heterologous proteins in a short period of time (Motohashi *et al.*, 2005). Using this expression system, MaSp1 from *N. clavipes* linked with an enhanced green fluorescent protein (EGFP) fusion protein was cloned and expressed in the *B. mori* cell line (BmN) and larvae (Zhang *et al.*, 2008). The authors reported successful production of a recombinant EGFP-MaSp1 fusion protein in both systems. In the silkworm larvae, a total of 6 mg of fusion protein was expressed, whereas in the BmN cells, 5% of the cell total protein was occupied by this

recombinant silk. The major limitations of this expression system were low solubility of silk proteins and inability to assemble spider silk fibres. It was shown that more than 60% of the fusion proteins formed aggregates via self-assembly. To overcome solubility issues, MaSp1 C-terminal domain is to be incorporated due to its role to prevent aggregate formation. To produce fibres, germline-transgenic silkworms (*B. mori*) were produced by injecting silkworm eggs with a *piggyBac* transformation vector carrying MaSp1 sequence (Wen *et al.*, 2010). The insects were capable of spinning fibres and forming cocoons containing recombinant spider silk. However, the mechanical properties of the fibres were lower than dragline MaSp1 silk due to the low ratio of MaSp1 in the total silk protein.

In a recent effort to develop tough fibres, transgenic silkworms encoding chimeric silkworm/spider silk proteins were produced using *piggyBac* vectors (Teulé *et al.*, 2012b). The vector, used previously by the Tamada group (Kojima *et al.*, 2007) included the *B. mori* fibroin heavy chain promoter and enhancer, a genetic sequencing encoding a 78 kDa synthetic spider silk protein, and an EGFP tag. Strong EGFP signals were observed by fluorescence (Fig. 5). The composite fibres were tougher than the parental silkworm silk fibres and as tough as native dragline spider silk fibres.

These results demonstrate that silkworms can be engineered to generate composite silk fibres containing stably integrated spider silk protein sequences, which significantly improved overall mechanical properties.

Transgenic plants have also been investigated as heterologous host systems to produce recombinant spider silks. Advances in genetic engineering technology and transformation methods make it possible to produce non-plant proteins in plants (Yang *et al.*, 2005; Rech *et al.*, 2008). Moreover, one plant offers several different expression systems, such as seeds, leaves, tubers and roots with potential for organelle-specific accumulation of recombinant proteins (Scheller and Conrad, 2005).

Stable transgenic tobacco and potato lines were engineered to express MaSp1 genes from *N. clavipes* ranging from 420 to 3600 bp (Scheller *et al.*, 2001). Recombinant spider silk proteins were found in the endoplasmic reticulum (ER) of tobacco and potato leaves at the accumulation of 2% of total soluble protein. Moreover, the production levels were independent of the size of silk genes. Purification was performed using high temperature treatment followed by acidification and ammonium sulphate precipitation. Additionally, recombinant MaSp1-like proteins were also produced in the leaves and seeds of *Arabidopsis* (small flowering plants related to cabbage) as well as in somatic soybean embryos (Barr *et al.*, 2004). The expression of recombinant silks was driven by the 35S promoter in leaves and the β -conglycinin α' subunit promoter in seeds and somatic soybean embryos. The

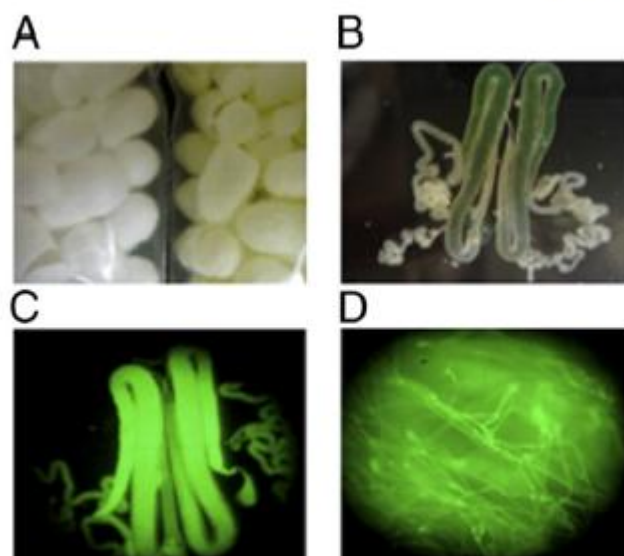


Fig. 5. Expression of the chimeric silkworm/spider silk/EGFP protein in (A) cocoons, (B and C) silk glands and (D) silk fibres from spider 6-GFP silkworms. Reproduced with permission from (Taulé *et al.*, 2012b).

results demonstrated that recombinant spider silk proteins had higher accumulation levels in seeds than in the leaves. Recently, a native-sized FLAG protein from *N. clavipes* was cloned and expressed in the ER of tobacco plant (*Nicotiana benthamiana*) leaf cells using an intein-based posttranslational protein fusion technology (Hauptmann *et al.*, 2013). This method avoids the need for highly repetitive transgenes resulting in a higher genetic and transcriptional stability. Additional details on production of fibrous proteins in plants can be found elsewhere (Scheller and Conrad, 2005).

Transgenic production of recombinant silk proteins in mammary glands and secretion of them into milk has been investigated in mice and goats (Williams, 2003; Xu *et al.*, 2007). In case of transgenic mice production, MaSp1 and MaSp2 synthetic genes (40 and 55 kDa) were synthesized and cloned into the pBC1 expression vector (Invitrogen, Carlsbad, CA, USA) together with a goat β -casein signal sequence. The chimeric gene construct was microinjected into pronuclei of fertilized eggs of Kunming white mice (Xu *et al.*, 2007). Southern blot analysis was used to identify mice containing transgene construct as well as a copy number of transgene. The expression of dragline silk in milk was confirmed by Northern blot followed by Western blot analysis. The results revealed that transgenic mice were capable of expressing recombinant silk proteins in their milk. Geneti-

cally engineered (transgenic) goats capable of expressing spider silk proteins based on the sequences of MaSp1 and MaSp 2 were produced by Nexia Biotechnologies, and later by the Lewis group (Lazaris *et al.*, 2002; Service, 2002). Silk protein expression was controlled by the β -casein promoter and was expressed in the milk of transgenic goats. Silk proteins were observed only in mammary tissues as confirmed by Western blot (Steinkraus *et al.*, 2012). Maximum yields observed for the recombinant silk production in transgenic animals were low (11.7 mg l^{-1}) when compared with bacterial expression (Table 1 and Table 2). Today, the large-scale production of recombinant silk proteins from transgenic animals is relatively expensive and challenging in terms of animal breeding.

Future outlook

Over the last decade there has been considerable progress in understanding the genetic organization encoding spider silks. Cloning, expression and purification of spider silks has improved, and the self-assembly and processing of spider silk into many material formats is now better understood. Recently a native-sized (285 kDa) recombinant protein of the spider *N. clavipes* was produced and spun into a fibre displaying mechanical properties comparable to those of the native silk, indicating a breakthrough in

standard recombinant production of spider silks. Moreover, a variety of heterologous host systems have been explored to produce different types of recombinant silks. For example, transgenic silkworm/spider silk production systems have been developed to produce tough fibres. It is possible to mix and match key modules via recombinant approaches, providing additional insights into the role of individual modules and effects of neighbouring elements on properties. This approach should lead to the development of custom structures built from specific silk elements. Future challenges will include development of tailor-made production systems for recombinant silks keeping in mind differences in chemical and physical properties of individual silk modules, scaling up silk production, prevention of the formation of aggregates and matches to the mechanical properties of silk fibres.

Acknowledgements

The authors thank the NIH (EB014976, EB002520; EY020856; DE017207; EB014283) and the BSF for support of the studies reviewed here. This study was supported in part by the National Council for Scientific and Technological Development (CNPq), Brazil, Embrapa Genetics Resources and Biotechnology, Brasilia, DF, Brazil and Foundation for Research Support (FAP-DF), Brazil.

References

- An, B., Hinman, M.B., Holland, G.P., Yarger, J.L., and Lewis, R.V. (2011) Inducing β -sheets formation in synthetic spider silk fibers by aqueous post-spin stretching. *Biomacromolecules* **12**: 2375–2381.
- An, B., Jenkins, J.E., Sampath, S., Holland, G.P., Hinman, M., Yarger, J.L., and Lewis, R. (2012) Reproducing natural spider silks' copolymer behavior in synthetic silk mimics. *Biomacromolecules* **13**: 3938–3948.
- Asakura, T., Umemura, K., Nakazawa, Y., Hirose, H., Higham, J., and Knight, D. (2007) Some observations on the structure and function of the spinning apparatus in the silkworm *Bombyx mori*. *Biomacromolecules* **8**: 175–181.
- Barr, L.A., Fahnestock, S.R., and Yang, J. (2004) Production and purification of recombinant DP1B silk-like protein in plants. *Mol Breed* **13**: 345–356.
- Brooks, A.E., Steinkraus, H.B., Nelson, S.R., and Lewis, R.V. (2006) An investigation of the divergence of major ampullate silk fibers from *Nephila clavipes* and *Argiope aurantia*. *Biomacromolecules* **6**: 3095–3099.
- Brooks, A.E., Stricker, S.M., Joshi, S.B., Kamerzell, T.J., Middaugh, C.R., and Lewis, R.V. (2008) Properties of synthetic spider silk fibers based on *Argiope aurantia* MaSp2. *Biomacromolecules* **9**: 1506–1510.
- Chung, H., Kim, T.Y., and Lee, S.Y. (2012) Recent advances in production of recombinant spider silk proteins. *Curr Opin Biotechnol* **23**: 957–964.
- Dams-Kozłowska, H., Majer, A., Tomasiewicz, P., Łozinska, J., Kaplan, D.L., and Mackiewicz, A. (2012) Purification and cytotoxicity of tag-free bioengineered spider silk proteins. *J Biomed Mater Res A* **101A**: 456–464.
- Eisoldt, L., Smith, A., and Scheibel, T. (2011) Decoding the secrets of spider silk. *Mater Today* **14**: 80–86.
- Fahnestock, S.R., and Bedzyk, L.A. (1997) Production of synthetic spider dragline silk protein in *Pichia pastoris*. *Appl Microbiol Biotechnol* **47**: 33–39.
- Fahnestock, S.R., and Irwin, S.L. (1997) Synthetic spider dragline silk proteins and their production in *Escherichia coli*. *Appl Microbiol Biotechnol* **47**: 23–32.
- Gomes, S.C., Leonor, I.B., Mano, J.F., Reis, R.L., and Kaplan, D.L. (2011) Antimicrobial functionalized genetically engineered spider silk. *Biomaterials* **32**: 4255–4266.
- Gosline, J.M., Denny, M.W., and DeMont, M.E. (1984) Spider silk as rubber. *Nature* **309**: 551–552.
- Grip, S., Rising, A., Nimmervoll, H., Storckenfeldt, E., McQueen-Mason, S., Pouchkina-Stantcheva, N., et al. (2006) Transient expression of a major ampullate spidroin 1 gene fragment from *Euprosthenois* sp. in mammalian cells. *Cancer Genomics and Proteomics* **3**: 83–87.
- Hauptmann, V., Weichert, N., Menzel, M., Knoch, D., Paegel, N., Scheller, J., et al. (2013) Native-sized spider silk proteins synthesized in planta via intein-based multimerization. *Transgenic Res* **22**: 369–377.
- Heim, M., Keerl, D., and Scheibel, T. (2009) Spider silk: from soluble protein to extraordinary fiber. *Angew Chem Int Ed* **48**: 3584–3596.
- Higashiya, S., Topilina, N.I., Ngo, S.C., Zagorevskii, D., and Welch, J.T. (2007) Design and Preparation of b-Sheet Forming Repetitive and Block-Copolymerized Polypeptides. *Biomacromolecules* **8**: 1487–1497.
- Hu, X., and Kaplan, D.L. (2011) Silk Biomaterials. In *Comprehensive Biomaterials*. Ducheyne, P. (ed.). Oxford, UK: Elsevier, pp. 207–219.
- Hu, X., Vasanthavada, K., Kohler, K., McNary, S., Moore, A., and Vierra, C. (2006) Molecular mechanisms of spider silk. *Cell Mol Life Sci* **63**: 1986–1999.
- Huang, J., Wong, C., George, A., and Kaplan, D.L. (2007) The effect of genetically engineered spider silk-dentin matrix protein 1 chimeric protein on hydroxyapatite nucleation. *Biomaterials* **28**: 2358–2367.
- Huemmerich, D., Helsen, C.W., Quedzuweit, S., Oschmann, J., Rudolph, R., and Scheibel, T. (2004) Primary structure elements of spider dragline silks and their contribution to protein solubility. *Biochemistry* **43**: 13604–13612.
- Humenik, M., Scheibel, T., and Smith, A. (2011) Spider silk understanding the structure-function relationship of a natural fiber. *Prog Mol Biol Transl Sci* **103**: 131–185.
- Ittiah, S., Cohen, S., Garty, S., Cohn, D., and Gat, U. (2006) An essential role for the C-terminal domain of a dragline spider silk protein in directing fiber formation. *Biomacromolecules* **7**: 1790–1796.
- Kluge, J.A., Rabotyagova, O., Leisk, G.G., and Kaplan, D.L. (2008) Spider silks and their applications. *Trends Biotechnol* **26**: 244–251.
- Knight, D.P., and Vollrath, F. (2001) Changes in element composition along the spinning duct in a *Nephila* spider. *Naturwissenschaften* **88**: 179–182.

- Knight, D.P., and Vollrath, F. (2002) Spinning an elastic ribbon of spider silk. *Philosophical Transactions of the Royal Society B. Biol Sci* **357**: 219–227.
- Kojima, K., Kuwana, Y., Sezutsu, H., Kobayashi, I., Uchino, K., Tamura, T., and Tamada, Y. (2007) A new method for the modification of fibroin heavy chain protein in the transgenic silkworm. *Biosci Biotechnol Biochem* **71**: 2943–2951.
- Lazaris, A., Arcidiacono, S., Huang, Y., Zhou, J.F., Duguay, F., Chretien, N., *et al.* (2002) Spider silk fibers spun from soluble recombinant silk produced in mammalian cells. *Science* **295**: 472–476.
- Leopoldo, R.F.E., Cristina, V.F.N., Rodrigues, V.G., Rodrigues, D.S.F., Jose, L.A.F., Alberto, C.L., *et al.* (2007) US20100311645 – Proteins from the webs of *Nephilengys cruentata*, *Avicularia juruensis* and *Psecowia bistriata* spiders isolated from Brazilian biodiversity. In: *US20100311645 A1 20101209*: Empresa Brasileira de Pesquisa Agropecuaria- EMBRAPA (Brasilia – DF, BR); Fundacao Universidade de Brasilia (Brasilia – DF, BR).
- Lewicka, M., Hermanson, O., and Rising, A.U. (2012) Recombinant spider silk matrices for neural stem cell cultures. *Biomaterials* **33**: 7712–7717.
- Lewis, R.V. (2006a) Spider silk: ancient ideas for new biomaterials. *Chem Rev* **106**: 3762–3774.
- Lewis, R.V. (2006b) Spider silk production. In *Bionanotechnology. Proteins to Nanodevices*. Renugopalakrishnan, V. and Lewis, R.V. (eds). Dordrecht, the Netherlands: Springer, pp. 61–78.
- Lewis, R.V., Hinman, M., Kothakota, S., and Fournier, M.J. (1996) Expression and purification of a spider silk protein: a new strategy for producing repetitive proteins. *Protein Expr Purif* **7**: 400–406.
- Lin, Z., Deng, Q., Liu, X.Y., and Yang, D. (2013) Engineered large spider eggcase silk protein for strong artificial fibers. *Adv Mater* **25**: 1216–1220.
- Meyer, D.E., and Chilkoti, A. (2002) Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. *Biomacromolecules* **3**: 357–367.
- Mieszawska, A.J., Fourligas, N., Georgakoudi, I., Ouhib, N.M., Belton, D.J., Perry, C.C., and Kaplan, D.L. (2010) Osteoinductive silk-silica composite biomaterials for bone regeneration. *Biomaterials* **31**: 8902–8910.
- Motohashi, T., Shimojima, T., Fukagawa, T., Maenaka, K., and Park, E.Y. (2005) Efficient large-scale protein production of larvae and pupae of silkworm by *Bombyx mori* nuclear polyhedrosis virus bacmid system. *Biochem Biophys Res Commun* **326**: 564–569.
- Numata, K., Mieszawska-Czajkowska, A.J., Kvenvold, L.A., and Kaplan, D.L. (2012) Silk-based nanocomplexes with tumor-homing peptides for tumor-specific gene delivery. *Macromol Biosci* **12**: 75–82.
- Prince, J.T., McGrath, K.P., DiGirolamo, C.M., and Kaplan, D.L. (1995) Construction, cloning, and expression of synthetic genes encoding spider dragline silk. *Biochemistry* **34**: 10879–10885.
- Rabotyagova, O.S., Cebe, P., and Kaplan, D.L. (2009) Self-assembly of genetically engineered spider silk block copolymers. *Biomacromolecules* **10**: 229–236.
- Rabotyagova, O.S., Cebe, P., and Kaplan, D.L. (2010) Role of polyalanine domains in β -sheet formation in spider silk block copolymers. *Macromol Biosci* **10**: 49–59.
- Rabotyagova, O.S., Cebe, P., and Kaplan, D.L. (2011) Protein-based block copolymers. *Biomacromolecules* **12**: 269–289.
- Rech, E.L., Vianna, G.R., and Aragao, F.J.L. (2006) High-efficiency transformation by biolistics of soybean, common bean and cotton transgenic plants. *Nat Protocols* **3**: 410–418.
- Rising, A., Widhe, M., Johansson, J., and Hedhammar, M. (2011) Spider silk proteins: recent advances in recombinant production, structure–function relationships and biomedical applications. *Cell Mol Life Sci* **68**: 169–184.
- Rosenberg, A.H., Goldman, E., Dunn, J.J., Studier, F.W., and Zubay, G. (1993) Effects of consecutive AGG codons on translation in *Escherichia coli*, demonstrated with a versatile codon test system. *J Bacteriol* **175**: 716–722.
- Scheller, J., and Conrad, U. (2005) Plant-based material, protein and biodegradable plastic. *Curr Opin Plant Biol* **8**: 188–196.
- Scheller, J., Guhrs, K.H., Grosse, F., and Conrad, U. (2001) Production of spider silk proteins in tobacco and potato. *Nat Biotechnol* **19**: 573–577.
- Service, R.F. (2002) Materials science. Mammalian cells spin a spidery new yarn. *Science* **295**: 419–421.
- Steinkraus, H.B., Rothfuss, H., Jones, J.A., Dissen, E., Shefferly, E., and Lewis, R.V. (2012) The absence of detectable fetal microchimerism in nontransgenic goats (*Capra aegagrus hircus*) bearing transgenic offspring. *J Anim Sci* **90**: 481–488.
- Teule, F., Cooper, A.R., Furin, W.A., Bittencourt, D., Rech, E.L., Brooks, A., and Lewis, R.V. (2009) A protocol for the production of recombinant spider silk-like proteins for artificial fiber spinning. *Nat Protoc* **4**: 341–355.
- Teulé, F., Addison, B., Cooper, A.R., Ayon, J., Henning, R.W., Benmore, C.J., *et al.* (2012a) Combining flagelliform and dragline spider silk motifs to produce tunable synthetic biopolymer fibers. *Biopolymers* **97**: 418–431.
- Teulé, F., Miao, Y.-G., Sohn, B.-H., Kim, Y.-S., Hull, J.J., Fraser, M.J., *et al.* (2012b) Silkworms transformed with chimeric silkworm/spider silk genes spin composite silk fibers with improved mechanical properties. *Proc Natl Acad Sci U S A* **109**: 923–928.
- Vepari, C., and Kaplan, D.L. (2007) Silk as a biomaterial. *Prog Polym Sci* **32**: 991–1007.
- Vollrath, F., and Knight, D.P. (1999) Structure and function of the silk production pathway in the spider *Nephila edulis*. *Int J Biol Macromol* **24**: 243–249.
- Wang, C., Patwardhan, S.V., Belton, D.J., Kitchel, B., Anastasiades, D., Huang, J., *et al.* (2006) Novel nanocomposites from spider silk-silica fusion (chimeric) proteins. *Proc Natl Acad Sci U S A* **103**: 9428–9433.
- Wen, H., Lan, X., Zhang, Y., Zhao, T., Wang, Y., Kajjura, Z., and Nakagaki, M. (2010) Transgenic silkworms (*Bombyx mori*) produce recombinant spider dragline silk in cocoons. *Mol Biol Rep* **37**: 1815–1821.
- Widmaier, D.M., Tullman-Ercok, D., Mirsky, E.A., Hill, R., Govindarajan, S., Minshull, J., and Voigt, C.A. (2009) Engineering the *Salmonella* type III secretion system to export spider silk monomers. *Mol Syst Biol* **5**: 1–9.

663 O. Tokareva, V. A. Michalczychen-Lacerda, E. L. Rech and D. L. Kaplan

- Williams, D. (2003) Sows' ears, silk purses and goats' milk: new production methods and medical applications for silk. *Med Device Technol* **14**: 9–11.
- Wright, E.R., and Conticello, V.P. (2002) Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. *Adv Drug Delivery Rev* **54**: 1057–1073.
- Xia, X.X., Qian, Z.G., Ki, C.S., Park, Y.H., Kaplan, D.L., and Lee, S.Y. (2010) Native-sized recombinant spider silk protein produced in metabolically engineered *Escherichia coli* results in a strong fiber. *Proc Natl Acad Sci U S A* **107**: 14059–14063.
- Xu, H.T., Fan, B.L., Yu, S.Y., Huang, Y.H., Zhao, Z.H., Lian, Z.X., *et al.* (2007) Construct synthetic gene encoding artificial spider dragline silk protein and its expression in milk of transgenic mice. *Anim Biotechnol* **18**: 1–12.
- Xu, L., Rainey, J.K., Meng, Q., and Liu, X.-Q. (2012) Recombinant minimalist spider wrapping silk proteins capable of native-like fiber formation. *PLoS ONE* **7**: e50227.
- Yang, J., Barr, L.A., Fahnestock, S.R., and Liu, Z.B. (2005) High yield recombinant silk-like protein production in transgenic plants through protein targeting. *Transgenic Res* **14**: 313–324.
- Zhang, Y., Hu, J., Miao, Y., Zhao, A., Zhao, T., Wu, D., *et al.* (2008) Expression of EGFP-spider dragline silk fusion protein in BmN cells and larvae of silkworm showed the solubility is primary limit for dragline proteins yield. *Mol Biol Rep* **35**: 329–335.
- Zhang, Y., Zhao, T., Zhao, A., and Nakagaki, M. (2011) Stably express spider flagelliform silk protein in *Bombyx mori* cell line by piggyBac transposon-derived vector. *Tianjin*. 779–782.

POSTER PRESENTATION

Open Access

Synthetic biology increases efficiency of *Escherichia coli* to produce *Parawixia bistrriata* spider silk protein

Valquíria Alice Michalczechen-Lacerda^{1,2,3*}, Olena Tokareva³, Andressa de Rezende Bastos^{1,2}, Marina Silveira da Silva^{1,2}, Giovanni Rodrigues Vianna¹, André Melro Murad¹, David Lee Kaplan³, Elfbio Leopoldo Rech¹

From 5th Congress of the Brazilian Biotechnology Society (SBBIOTEC)
 Florianópolis, Brazil. 10-14 November 2013

Background

Spider dragline silk is considered to be the toughest biopolymer on Earth due to an extraordinary combination of strength and elasticity. With synthetic biology it is possible to express recombinant spider silk proteins, which are characterized by a highly repetitive rich glycine and alanine sequence [1]. However, production of high molecular weight spider silk protein can be difficult due to DNA instability, transcription and translation errors. Here we show, for the first time, Masp2 (105 kDa) spider silk protein production from the Brazilian spider *Parawixia bistrriata* in different metabolically engineered *E.coli* strains.

Methods

A *Masp2* monomer gene from *P. bistrriata* was designed (DNA2.0), and a 32 mer plasmid was constructed [2]. To increase the glycyl-tRNA pool, tRNA^{Gly} and glycyl-tRNA synthetase genes were cloned in pACYC184 plasmid [3]. All vectors were confirmed by DNA sequencing. The bacterium BL21(DE03) was co-transformed with the 32 mer and one metabolic plasmid. pACYC184 and pET19b vectors was used as controls. Cells were grown in a 2L flask culture with 1L of LB medium with antibiotics, at 37°C and 200 rpm. Silk proteins were induced with 1 mM IPTG, for four hours, and growth curves were established. Protein histag N-terminal extraction was performed under native conditions and purified by IMAC. All samples were analyzed by SDS-PAGE gels, staining with Colloidal Blue

and by Western blot. Dialysis was at 4°C against 10 mM Tris-HCl pH 8.0 for the first 24 hours, and water for an additional 24 hours. Samples were lyophilized and weighed. Statistical analyses were determined by ANOVA and unpaired Tukey test (ASSISTAT 7.7 BETA 2013), being statistically significant at $P < 0.01$.

Results and conclusions

All *E. coli* that received Masp2 grew, expressing spider protein, and data yields are shown in table 1 (N = 3). There were no expression differences between BL21 (DE03) and pACYC184, nor in the strain containing one gene copy to increase tRNA^{Gly} pool. This shows no interference from the initial plasmid used for bacterial metabolic engineering. The amount of tRNA was probably not enough to supply the metabolic stress. When the tRNA^{Gly} and glycine pool were overexpressed by pTetgly2, pTetglyVXY-glyA and pTetgly2-glyA, an improvement was noted in the production of 105 kDa Masp2 ($P < 0.01$). In similar studies, the same synergistic effect was reported for high molecular weight protein from *Nephila clavipes*

Table 1 Spider silk Masp2 protein yield in different engineered *E.coli*

Bacteria	Means	*
pTetgly2-glyA	3.39	a
pTetglyVXY-glyA	2.83	ab
pTetgly2	2.40	b
pTetglyVXY	0.96	c
pACYC184	0.99	c
BL21 (DE03)	0.79	c

Means reference: mg protein/g pellet. $P < 0.01$; CV = 13.72%

¹Embrapa Genetic Resources and Biotechnology-Parque Estação Biológica, Brasília, DF, 70770-917, Brazil

Full list of author information is available at the end of the article

[3]. The same authors also reported that protein sizes are directly associated with fiber quality properties. The 105 kDa Masp2 produced was able to be spun in fibers, which will be characterized in future analyses. Currently, there is no ideal heterologous organism to produce spider silk proteins, and metabolic engineering together with synthetic biology can optimize spider silk protein production. Different spider silk proteins may result in new types of protein-based biomaterials with wide applications in medicine and industry.

Acknowledgements

Embrapa Genetics Resources and Biotechnology, University of Brasilia, CNPq, CAPES, FAPDF, the NIH P41 Resource Center (EB002520), Tufts University.

Authors' details

¹Embrapa Genetic Resources and Biotechnology-Parque Estação Biológica, Brasilia, DF, 70770-917, Brazil. ²University of Brasilia, Campus Darcy Ribeiro, Brasilia, DF, 70910-900, Brazil. ³Tufts University, Medford, MA, 02155, USA.

Published: 1 October 2014

References

1. Hinman MB, Jones JA, Lewis RV: **Synthetic spider silk: a modular fiber.** *Trends Biotechnology* 2000, **18**(9):374-379.
2. Teulé F, Cooper AR, Furin WA, Bittencourt D, Rech EL, Brooks A, Lewis RL: **A protocol for the production of recombinant spider silk-like proteins for artificial fiber spinning.** *Nature Protocols* 2009, **4**(3):341-355.
3. Xia XX, Qian ZG, Ki CS, Park YH, Kaplan DL, Lee SY: **Native-sized recombinant spider silk protein produced in metabolically engineered Escherichia coli results in a strong fiber.** *PNAS* 2010, **107**(32):14059-14063.

doi:10.1186/1753-6561-8-S4-P231

Cite this article as: Michalczewen-Lacerda et al: Synthetic biology increases efficiency of *Escherichia coli* to produce *Parawixia bistriata* spider silk protein. *BMC Proceedings* 2014 **8**(Suppl 4):P231.

ESB 2014

it's liverpool

31st August – 3rd September 2014UNIVERSITY OF
LIVERPOOL

Musculoskeletal Biology
Institute of Ageing & Chronic Disease
UK Centre for Tissue Engineering
Faculty of Health & Life Sciences
Ground Floor
Duncan Building
Daulby Street
Liverpool
L69 3GA

T +44-151 706 4913
E esb2014@liv.ac.uk

16th September 2014

CERTIFICATE OF POSTER PRESENTATION

I certify that :

Valquíria Michalczechen-Lacerda

of

University of Brasilia

Gave a posterrpresentation entitled:

“Nanoscale Analyses of Parawixia bistriata Synthetic Spider Silk Fibres”

at the 26th European Conference on Biomaterials, at the Liverpool Arena and Convention Centre, 31st August – 3rd September 2014.

UNIVERSITY OF
LIVERPOOL
The Institute of Ageing and Chronic Disease
4th Floor UCD
Duncan Building
Daulby Street
Liverpool L69 3GA

Dr Nicholas Rhodes
Technical & Financial Director, ESB Liverpool 2014

on behalf of Professor John Hunt
Chair, ESB Liverpool 2014

UNIVERSITY OF
LIVERPOOL

UK Centre for Tissue Engineering



Nanoscale Analyses of *Parawixia bistriata* Synthetic Spider Silk Fibres

Valquíria A Michalczechen-Lacerda^{1,2*}, Giovanni R Vianna², André M Murad², Luciano P Silva², Elibio L Rech²

¹Department of Cell Biology/ Institute of Biology, University of Brasília, Brazil, * kira21bio@gmail.com

²Synthetic Biology and Nanobiotechnology Group/Embrapa Genetic Resources and Biotechnology, Brazil,

INTRODUCTION

Spider silks are remarkable natural polymers and, with advances in genetic engineering, it is possible to produce synthetic silks based on recombinant proteins obtained from heterologous hosts even at large-scale. The core repetitive domain is a block copolymer-like arrangement, where amino acids sequence can result in a protein with unique biomechanical properties. Different protein sequences may result in new types of spider silks-based biomaterials. They also display biocompatibility and a slow rate of degradation, making synthetic spider silks promising candidates as biomaterials for tissue engineering, guided tissue repair, cosmetic products and industrial materials. The aim of the present study was to characterize synthetic *P. bistriata* MaSp2 fibres, for surface topography, viscoelasticity, roughness, Young's Modulus, separation and interacting forces using atomic force microscopy (AFM) image analyses and force spectroscopy experiments.

EXPERIMENTAL METHODS

A silk monomer gene from *P. bistriata* major ampullate 2 (MaSp2) was designed (DNA2.0). The 16 mcr plasmid was constructed using a "head to tail" cloning strategy¹. All plasmids were confirmed by DNA sequencing. The 54 kDa protein containing a histag was induced at *Escherichia coli* BL21(DE03) with 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) for 4 hours, at 37°C, 200 rpm, in LB medium. After immobilized metal affinity chromatography (IMAC) purification, protein was dialyzed against distilled water, lyophilized and dissolved into 95% hexafluoroisopropanol (HFIP) + 5% formic acid. The fibers were wet spun by a pump syringe and extruded in 100% isopropanol coagulation bath. A total of 20 fibers with 1 cm length were separated in two groups: control (10 fibres) and isopropanol 80% treatment with 1 \times stretch (10 fibres). Surface and force spectroscopy analyses were performed by AFM under ambient conditions on a SFM-9600 instrument (Shimadzu, Kyoto, Japan) as previously described². Statistical significance was determined by student t test using the software Assistat 7.7 Beta (Assistat, Campina Grande, Brazil).

RESULTS AND DISCUSSION

The fibres presented inhomogeneous topographical characteristics. At least 2-4 different areas of each fibre were analysed. All groups presented a similar general pattern for Rz (maximum height), Rzjis (10 points mean roughness based on the five highest peaks and lowest valleys in the entire sampling length), Rp (maximum peaks and lowest valleys) presenting no difference between control and isopropanol groups. However, Ra

(arithmetic average roughness), Rq (root-mean-square roughness), Rv (maximum valley depth) analysis showed a decreased for isopropanol treatment compared to control group ($P < 0.05$).

For the fibre force spectroscopy measurements there were no difference between control group and isopropanol treatment for detachment force (comprised adhesive forces), Young's Modulus (local elasticity), maximum load force and energy dissipated during the analyses. Only the snap-in, which illustrate the attractive force resulting from the primary surface interactions of the fibre with the tip, showed an increase for the isopropanol treatment group ($P < 0.01$).

The isopropanol-treated group did not differ significantly from the control group as expected. It is known that alcohols might be responsible for induction β -sheet formation. In a previous work², isopropanol seems to result in a better fibre formation than using methanol for post-spinning treatments, and they also found a decrease for Rv parameter². Other studies have shown inferior mechanical properties for synthetic fibres from recombinant protein with low molecular mass compared to natural fibres. However, synthetic fibres can increase its mechanical quality as a natural fibres as long as the monomers and protein molecular mass are increased³. Synthetic spider silk fibre with high molecular mass as native one (300 kDa) needs to be measured with this technology.

CONCLUSION

The synthetic fibres presented irregular surface, elasticity and strength. This biomaterial depends on their molecular organizations which are not completed even after extrusion. Different bath coagulation and a treatment post-spinning still needs to be optimized for the production of synthetic fibres with quality similar to natural dragline silk fibre. Analyses by AFM are a reliable method to qualitatively describe the mechanical properties of silk fibres.

REFERENCES

1. Teulé F *et al.*, Nature Protocols. 4:341-355, 2009
2. Menezes GM *et al.*, Polymer Journal. 45:997-1006, 2013
3. Xia XX *et al.*, PNAS.107:14059-14063, 2010

ACKNOWLEDGMENTS

The authors would like to thank National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), Fundação de Apoio a Pesquisa (FAP-DF), Embrapa Genetic Resources and Biotechnology and University of Brasília (UnB) for providing financial support to this project.

