

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
CENTRO DE CIÊNCIAS MATEMÁTICAS E DA NATUREZA  
INSTITUTO DE QUÍMICA  
DEPARTAMENTO DE BIOQUÍMICA

**VIVIANE DE ALMEIDA BASTOS**

MAPEAMENTO MOLECULAR DO COMPLEXO BJ46a - JARARAGINA POR  
ESPECTROMETRIA DE MASSAS

Rio de Janeiro

2019

VIVIANE DE ALMEIDA BASTOS

**MAPEAMENTO MOLECULAR DO COMPLEXO BJ46a - JARARAGINA POR  
ESPECTROMETRIA DE MASSAS**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Bioquímica do Instituto de Química da Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Doutor em Ciências.

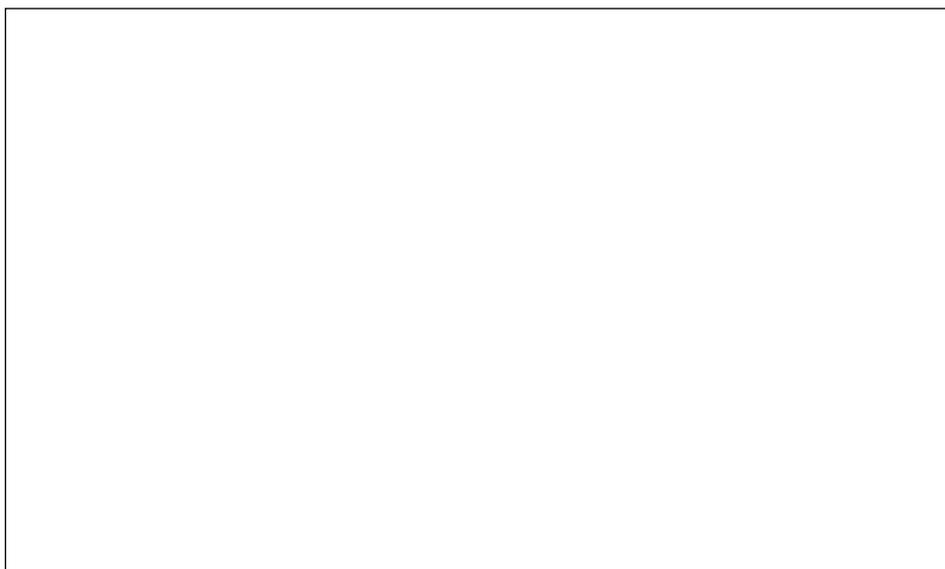
Orientadores: Dr. Richard Hemmi Valente

Dra. Ana Gisele da Costa Neves Ferreira

Rio de Janeiro

2019

Espaço reservado à ficha catalográfica



**VIVIANE DE ALMEIDA BASTOS**

**MAPEAMENTO MOLECULAR DO COMPLEXO BJ46a - JARARAGINA POR  
ESPECTROMETRIA DE MASSAS**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Bioquímica da Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de Doutor em Ciências.

Rio de Janeiro, 14 de fevereiro de 2019.

---

Prof. Dr. Richard Hemmi Valente  
Presidente – IOC / FIOCRUZ

---

Prof. Dra. Yraima Moura Lopes Cordeiro  
CCS / UFRJ

---

Prof. Dr. Ernesto Raúl Caffarena  
IOC / FIOCRUZ

---

Prof. Dr. Gilberto Barbosa Domont  
IQ/UFRJ

---

Prof. Dr. Anderson de Sá Pinheiro  
IQ / UFRJ

Ao meu avô, Octavio Lopes de Almeida.

“De todo o amor que eu tenho, metade foi tu que me deu

Salvando minh'alma da vida

Sorrindo e fazendo o meu eu...”

## **AGRADECIMENTOS**

*“A gratidão não é somente a maior das virtudes, mas a origem de todas as outras.”*

Cícero

Sou profundamente grata ao meu orientador, Dr. Richard Hemmi Valente, por toda esta caminhada, desde meus anos de graduação como aluna de iniciação científica até agora. Neste final de jornada, compreendi que é extremamente raro termos um mentor e não apenas um orientador acadêmico, e eu tive o privilégio de ter você como meu mentor. São poucas as palavras que realmente refletem o quanto sou agradecida por esta oportunidade, que foi transformadora para mim em muitos sentidos.

À Dra. Ana Gisele da Costa Neves Ferreira, por abraçar o projeto aceitando me orientar no Doutorado. Obrigada por todo o apoio, as discussões e as ideias que foram extremamente valiosas para minha formação e para o meu trabalho. Você é um exemplo de profissional para mim!

Ao Dr. Jonas Perales, por ter me aceitado na equipe do Laboratório de Toxinologia, e por todo o apoio, carinho e conselhos valiosos dados a mim por todos estes anos.

Ao Dr. Francisco Gomes Neto, pela colaboração fundamental na área de biologia estrutural, pelas discussões, por sempre me mostrar um ponto de vista que eu não tinha percebido antes, e a animação contagiosa toda vez que conseguíamos fazer algum avanço! Aprendi muito com você nestes anos e por isso sou grata por tudo.

Ao Diogo Lima Borges, obrigada pela ajuda inestimável durante a análise de XL-MS, e por sempre estar disponível para me ensinar a trabalhar e analisar os dados no SIM-XL.

À Surza Lucia Rocha, por toda a amizade por todos estes anos no laboratório, por ter sempre a mão estendida para ajudar quando precisei e sempre me

incentivando e encorajando a prosseguir, mesmo quando tudo parecia tão difícil. Obrigada por todos os momentos muito legais que passamos juntos, os congressos, os almoços no Nova América! Ganhei uma amiga para a vida, gosto muito de você!

À minha amiga Luciana Girão, dona do melhor abraço do mundo! Muito obrigada pelas conversas, as caronas, as ajudas e os momentos muito engraçados que passamos juntas. Conviver com você realmente fez meus dias mais alegres e leves! Aprendi muito com você a ver sempre um lado positivo das coisas e isso não tem preço.

À minha amiga Barbara Soares, obrigada por sempre estar ali do meu lado me apoiando, discutindo experimentos, ideias, escritas e slides, analisando mil espectros na mão, não deixando desanimar... foi muito bom contar com uma “parceira no crime” em nossos projetos tão desafiadores!

À Priscila Brandão, “General Potinho”! Obrigada por todos os momentos muito engraçados que passamos juntas, e todas as ajudas para não deixar esmorecer, incluindo áudios engraçados no whatsapp nas madrugadas de escrita!

À toda equipe do Laboratório de Toxinologia, vocês foram genuinamente minha segunda família por todos estes anos. Uma andorinha só não faz verão, e da mesma forma, uma tese não se faz apenas por esforço solitário de uma pessoa. Cada um contribuiu de maneira muito especial para que essa tese surgisse. Cresci muito, profissionalmente e pessoalmente, por estar junto dessa equipe, e esta tese é prova disso.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, via Edital Toxinologia 063/2010.

“A viagem não acaba nunca. Só os viajantes acabam. E mesmo estes podem prolongar-se em memória, em lembrança, em narrativa. Quando o viajante se sentou na areia da praia e disse: “Não há mais o que ver”, sabia que não era assim. O fim duma viagem é apenas o começo doutra. É preciso ver o que não foi visto, ver outra vez o que se viu já, ver na Primavera o que se vira no Verão, ver de dia o que se viu de noite, com sol onde primeiramente a chuva caía, ver a seara verde, o fruto maduro, a pedra que mudou de lugar, a sombra que aqui não estava. É preciso voltar aos passos que foram dados, para os repetir, e para traçar caminhos novos ao lado deles. É preciso recomeçar a viagem. Sempre. O viajante volta já.”

JOSÉ SARAMAGO, *Viagem a Portugal*

## RESUMO

BASTOS, Viviane de Almeida. **Mapeamento molecular do complexo BJ46a – jararagina por espectrometria de massas.** Tese (Doutorado em Bioquímica) – Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019.

Os acidentes ofídicos são um problema crescente de saúde pública, afetando milhões de pessoas em todo o mundo. A administração intravenosa do soro antiofídico é o único tratamento disponível, e os danos locais são apenas parcialmente neutralizados. Uma vertente da pesquisa de novas terapias antiofídicas dedica-se ao estudo de inibidores naturais e sintéticos que possam ser utilizados logo após o acidente, para minimizar os extensos danos locais que são frequentemente observados nas vítimas. O inibidor BJ46a é uma glicoproteína isolada do plasma da serpente *Bothrops jararaca* que inibe a ação catalítica de metaloendopeptidases de venenos de serpentes por meio da formação de um complexo estável, de alta afinidade, e mantido por forças de natureza não-covalente. Entretanto, os requisitos estruturais que regem esta interação não são conhecidos. Neste trabalho, aplicamos as técnicas de *cross-linking* (XL-MS) e troca isotópica hidrogênio-deutério monitoradas por espectrometria de massas (HDX-MS) com o objetivo de analisar as regiões de interação entre BJ46a e a metaloendopeptidase de veneno de serpente jararagina. As restrições espaciais geradas pelos experimentos de XL-MS foram exploradas para guiar a modelagem molecular de BJ46a e jararagina, bem como as simulações de *docking* proteína-proteína. Os dados de HDX-MS indicaram regiões de proteção/desproteção na interface do complexo entre BJ46a e jararagina. Em jararagina, regiões importantes para a coordenação do átomo de zinco catalítico tiveram seus perfis de incorporação isotópica alterados quando em complexo; para BJ46a, a região N-terminal é significativamente protegida da troca isotópica. O entendimento do mecanismo de inibição de metaloendopeptidases por BJ46a pode levar ao desenvolvimento de novos fármacos efetivos para a prevenção do dano tecidual local nos acidentes ofídicos, bem como para outras doenças associadas a um desequilíbrio na atividade catalítica e/ou expressão de metaloendopeptidases.

**Palavras-chave:** resistência natural; inibidor de metaloendopeptidases; modelagem molecular.

**ABSTRACT**

BASTOS, Viviane de Almeida. **Mapeamento molecular do complexo BJ46a – jararagina por espectrometria de massas.** Tese (Doutorado em Bioquímica) – Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019.

Snakebite envenoming is a growing public health problem, affecting millions of people worldwide. The intravenous administration of antivenom is the only available treatment and several studies have shown that local effects are only partially neutralized. A current trend in antivenom research is the study of natural and synthetic inhibitors that could be used soon after the envenomation event to restrain local tissue degradation. The inhibitor BJ46a is a glycoprotein isolated from *Bothrops jararaca* serum that inhibits the catalytic activity of snake venom metalloendopeptidases (SVMPs) through the formation of a stable high-affinity non-covalent complex with the toxin. However, the structural features that govern this interaction are largely unknown. In this work, the regions of interaction between BJ46a and the SVMP jararhagin have been analyzed using the techniques of cross-linking-mass spectrometry (XL-MS) and hydrogen-deuterium exchange-MS (HDX-MS). We explored the distance restraints generated from XL-MS experiments to guide the modeling of BJ46a and jararhagin, as well as the protein-protein docking simulations. HDX-MS data pinpointed regions of protection/deprotection in the interface of BJ46a-jararhagin complex. In jararhagin, key structural features in the coordination of the catalytic zinc ion had their deuteration profiles altered when in complex, whereas for BJ46a the N-terminal region is significantly protected from the isotopic exchange. The understanding of BJ46a inhibitory mechanism may lead to the development of new effective targeted therapeutics to prevent local tissue damage in ophidic accidents, and for other diseases associated with an imbalance in the catalytic activity and/or expression of metalloendopeptidases.

**Keywords:** natural resistance; metalloendopeptidase inhibitor; molecular modeling.

## LISTA DE ILUSTRAÇÕES

|   |    |
|---|----|
| FIGURA 1: INCIDÊNCIA MUNDIAL DE ACIDENTES OFÍDICOS. ....  | 1  |
| FIGURA 2: REPRESENTAÇÃO ESQUEMÁTICA DOS EFEITOS LOCAIS E SISTÊMICOS<br>CARACTERÍSTICOS DO ACIDENTE BOTRÓPICO. ....  | 5  |
| FIGURA 3: ABORDAGENS BASEADAS EM ESPECTROMETRIA DE MASSAS APLICADAS À BIOLOGIA<br>ESTRUTURAL INTEGRATIVA.....   | 22 |
| FIGURA 4: FLUXOGRAMA ESQUEMÁTICO DE UMA REAÇÃO DE XL-MS.....  | 25 |
| FIGURA 5: REPRESENTAÇÃO ESQUEMÁTICA DE MODIFICAÇÕES ENCONTRADAS EM PEPTÍDEOS<br>NA TÉCNICA DE XL-MS.....  | 26 |
| FIGURA 6: LOCALIZAÇÃO DOS ÁTOMOS DE HIDROGÊNIO EM PROTEÍNAS. ....   | 29 |
| FIGURA 7: A INFLUÊNCIA DO PH NA CONSTANTE QUÍMICA DE TROCA ( $K_{CH}$ , PAINEL A) E NA MEIA-<br>VIDA DA INCORPORAÇÃO ISOTÓPICA DE UM HIDROGÊNIO AMÍDICO (PAINEL B) EM UM<br>PEPTÍDEO POLI-DL-ALANINA A 20°C. .... | 30 |
| FIGURA 8: A CONSTANTE QUÍMICA DE TROCA $K_{CH}$ EM FUNÇÃO DA TEMPERATURA.....   | 31 |
| FIGURA 9: CARACTERIZAÇÃO DOS REGIMES EX1 E EX2 EM HDX-MS PELA ANÁLISE DOS<br>ENVELOPES ISOTÓPICOS.....  | 32 |
| FIGURA 10: FLUXOGRAMA ESQUEMÁTICO DE UM EXPERIMENTO DE HDX-MS.....  | 33 |
| FIGURA 11: DESENHO ESQUEMÁTICO DO SISTEMA EXPERIMENTAL PARA HDX-MS. ADAPTADO<br>A PARTIR DE MYSLING (2013) .....  | 51 |
| FIGURA 12: <i>IMMUNOBLOTTING</i> DO PLASMA DE <i>BOTHROPS JARARACA</i> E DO INIBIDOR BJ46A.<br>.....  | 62 |
| FIGURA 13: CROMATOGRAFIA DE IMUNOAFINIDADE DO PLASMA DE <i>B. JARARACA</i> EM COLUNA<br>HiTRAP NHS-IGG ANTI-BJ46A.....  | 63 |
| FIGURA 14: CROMATOGRAFIA EM FASE REVERSA EM COLUNA C4 SEMI-PREPARATIVA (10 X<br>250 MM) DOS <i>POOLS</i> DE FRAÇÕES LIGADAS ORIUNDAS DA CROMATOGRAFIA DE<br>IMUNOAFINIDADE.....                                   | 64 |
| FIGURA 15: SDS-PAGE DAS FRAÇÕES 6 E 7 ORIUNDAS DA CROMATOGRAFIA DE FASE<br>REVERSA.....   | 66 |
| FIGURA 16: RECROMATOGRAFIA EM FASE REVERSA DE BJ46A PURIFICADA (FRAÇÃO 7) APÓS<br>OS PASSOS CROMATOGRÁFICOS DE IMUNOAFINIDADE E FASE REVERSA. ....  | 67 |
| FIGURA 17: ALINHAMENTO DA SEQUÊNCIA DE BJ46A COM HSF, HLP, CMSF E JMSF<br>REALIZADAS PELO BLAST.....  | 70 |
| FIGURA 18: ALINHAMENTO DA SEQUÊNCIA DE BJ46A COM OS INIBIDORES JMSF E CMSF..  | 71 |

|  |    |
|--|----|
| FIGURA 19: MAPA DO ALINHAMENTO DOS PEPTÍDEOS SEQUENCIADOS <i>DE NOVO</i> DE BJ46A COM A SEQUÊNCIA DEPOSITADA NO UNIPROT.....                             | 73 |
| FIGURA 20: COBERTURA DE SEQUÊNCIA DO INIBIDOR BJ46A.....   | 73 |
| FIGURA 21: COBERTURA DE SEQUÊNCIA PARA BJ46A UTILIZANDO AS FERRAMENTAS PEAKS DB (ABORDAGEM PSM).....   | 74 |
| FIGURA 22: EXPERIMENTO DE FORMAÇÃO DE COMPLEXO ENTRE O INIBIDOR BJ46A E A METALOENDOPEPTIDASE JARARAGINA.....  | 78 |
| FIGURA 23: ENSAIO DE DILUIÇÃO SERIADA DO COMPLEXO BJ46A - JARARAGINA COMO PROVA DE CONCEITO PARA OS EXPERIMENTOS DE HDX-MS. ....                         | 79 |
| FIGURA 24: ENSAIO DE INIBIÇÃO DE ATIVIDADE AZOCASEINOLÍTICA DE JARARAGINA POR BJ46A.....   | 80 |
| FIGURA 25: ESTABILIZAÇÃO DO COMPLEXO ENTRE O INIBIDOR BJ46A E JARARAGINA USANDO O ALC BS <sup>3</sup> . ....   | 83 |
| FIGURA 26: ESTABILIZAÇÃO COVALENTE DO COMPLEXO BJ46A-JARARAGINA PELA METODOLOGIA XPLEX. ....   | 85 |
| FIGURA 27: PURIFICAÇÃO DAS AMOSTRAS ESTABILIZADAS POR BS <sup>3</sup> POR CROMATOGRAFIA DE EXCLUSÃO MOLECULAR. ....                                      | 86 |
| FIGURA 28: ESTRUTURA DOS ÍONS-REPÓRTER DETECTADOS EM ESPECTROS MS/MS DE PEPTÍDEOS COVALENTEMENTE ESTABILIZADOS COM O ALC BS <sup>3</sup> . ....          | 88 |
| FIGURA 29: <i>CROSS-LINKS</i> DETECTADOS PARA BJ46A LIVRE UTILIZANDO O ALC BS <sup>3</sup> . ....  | 91 |
| FIGURA 30: <i>CROSS-LINKS</i> DETECTADOS PARA BJ46A LIVRE UTILIZANDO A ABORDAGEM XPLEX. ....   | 92 |
| FIGURA 31: MAPAS BIDIMENSIONAIS DE RESTRIÇÕES DE DISTÂNCIA PARA BJ46A GERADAS PELAS ABORDAGENS BS <sup>3</sup> E XPLEX. ....                             | 93 |
| FIGURA 32: DISTRIBUIÇÃO DOS <i>CROSS-LINKS</i> ENCONTRADOS EM BJ46A NA ABORDAGEM XPLEX, DE ACORDO COM O TIPO DE <i>CROSS-LINK</i> FORMADO. ....          | 94 |
| FIGURA 33: <i>CROSS-LINKS</i> DETECTADOS PARA JARARAGINA LIVRE UTILIZANDO O ALC BS <sup>3</sup> . ....   | 95 |
| FIGURA 34: <i>CROSS-LINKS</i> DETECTADOS PARA JARARAGINA LIVRE UTILIZANDO A ABORDAGEM XPLEX. ....  | 96 |
| FIGURA 35: MAPAS BIDIMENSIONAIS DE RESTRIÇÕES DE DISTÂNCIA PARA JARARAGINA GERADOS A PARTIR DAS ABORDAGENS BS <sup>3</sup> E XPLEX. ....                 | 97 |
| FIGURA 36: REPRESENTAÇÃO DETALHADA DA ESTRUTURA CRISTALOGRÁFICA DE BOTROPASINA, UMA SVMP QUE APRESENTA 95% DE IDENTIDADE SEQUENCIAL COM JARARAGINA. .... | 98 |

|   |     |
|---|-----|
| FIGURA 37: <i>CROSS-LINKS</i> DETECTADOS PARA O COMPLEXO BJ46A-JARARAGINA UTILIZANDO O ALC BS <sup>3</sup> . .....  | 101 |
| FIGURA 38: <i>CROSS-LINKS</i> DETECTADOS PARA O COMPLEXO BJ46A-JARARAGINA UTILIZANDO A ABORDAGEM XPLEX. ....  | 102 |
| FIGURA 39: MAPA BIDIMENSIONAL DAS RESTRIÇÕES DE DISTÂNCIA ENCONTRADAS PARA O COMPLEXO BJ46A – JARARAGINA UTILIZANDO O ALC BS <sup>3</sup> . ....  | 103 |
| FIGURA 40: MAPA BIDIMENSIONAL DAS RESTRIÇÕES DE DISTÂNCIA ENCONTRADAS PARA O COMPLEXO BJ46A – JARARAGINA UTILIZANDO A ABORDAGEM XPLEX. ....   | 104 |
| FIGURA 41: PADRÃO PREDITO DE PONTES DISSULFETO EM BJ46A, COM BASE EM FETUÍNAS. ....   | 106 |
| FIGURA 42: ANÁLISE DAS PONTES DISSULFETO DE BJ46A PELO SOFTWARE SIM-XL. ....  | 108 |
| FIGURA 43: ESPECTROS DE MASSAS DAS PONTES DISSULFETO VALIDADAS MANUALMENTE (C9-C313 E C124-C127).....   | 109 |
| FIGURA 44: SISTEMA DE HDX-MS IMPLEMENTADO NO LABORATÓRIO DE TOXINOLOGIA, IOC/FIOCRUZ.....   | 112 |
| FIGURA 45: VISÃO INTERNA DO TANQUE DO SISTEMA DE HDX-MS.....  | 113 |
| FIGURA 46: VISÃO DAS CONEXÕES E PARTES SUBMERSAS DO SISTEMA DE HDX-MS.....  | 114 |
| FIGURA 47: VISÃO SUPERIOR DO SISTEMA DE HDX-MS. ....  | 115 |
| FIGURA 48: DIAGRAMAS DE VENN RELATIVO AO NÚMERO DE PEPTÍDEOS IDENTIFICADOS DE BJ46A (PAINEL A) E PEPSINA (PAINEL B) EM DIFERENTES CONCENTRAÇÕES DE GUANIDINA-HCL. ....  | 117 |
| FIGURA 49: COBERTURA DE SEQUÊNCIA DE BJ46A UTILIZANDO UMA SOLUÇÃO QUENCH CONTENDO 0,5 M GUANIDINA-HCL, 0,2 M TCEP E 0,2 M FOSFATO DE SÓDIO, PH 2,5, E COM A VOLTAGEM E A TEMPERATURA DO CAPILAR DA FONTE CONFIGURADOS PARA 3,5 kV E 100°C, RESPECTIVAMENTE..... | 119 |
| FIGURA 50: MAPA DE COBERTURA DE SEQUÊNCIA PARA JARARAGINA, NA FAIXA DE <i>m/z</i> 300 – 2000. ....  | 121 |
| FIGURA 51: COBERTURA DE SEQUÊNCIA (A) E NÚMERO DE PEPTÍDEOS (B) DETECTADOS PARA JARARAGINA (BARRAS EM AZUL) E PEPSINA (BARRAS EM VERDE) UTILIZANDO UMA SOLUÇÃO QUENCH CONTENDO UREIA 8 M E TCEP 1 M, EM DIFERENTES INTERVALOS DE INCUBAÇÃO A -80°C.....         | 122 |
| FIGURA 52: COBERTURA DE SEQUÊNCIA PARA JARARAGINA E PEPSINA EM TEMPOS DE INCUBAÇÃO DE ATÉ 24H A -80°C.....  | 123 |

|   |     |
|---|-----|
| FIGURA 53: COBERTURAS DE SEQUÊNCIA OBTIDAS PARA JARARAGINA E PEPSINA EM DIFERENTES TEMPOS DE INCUBAÇÃO E APÓS DILUIÇÃO DA SOLUÇÃO <i>QUENCH</i> IMEDIATAMENTE ANTES DA INJEÇÃO DA AMOSTRA NO SISTEMA DE HDX-MS..... | 124 |
| FIGURA 54: COBERTURA DE SEQUÊNCIA OBTIDA PARA JARARAGINA NA ETAPA FINAL DA OTIMIZAÇÃO DO PROTOCOLO DE HDX-MS. ....  | 125 |
| FIGURA 55: COBERTURA DE SEQUÊNCIA OBTIDA PARA BJ46A NA ETAPA FINAL DA OTIMIZAÇÃO DO PROTOCOLO DE HDX-MS. ....   | 126 |
| FIGURA 56: CROMATOGRAMAS DE ÍONS REPRESENTATIVOS OBTIDOS PARA A ANÁLISE DA CURVA DE DEUTERAÇÃO DE BJ46A LIVRE.....  | 129 |
| FIGURA 57: CROMATOGRAMAS DE ÍONS REPRESENTATIVOS OBTIDOS PARA A ANÁLISE DA CURVA DE DEUTERAÇÃO DE JARARAGINA LIVRE. ....  | 130 |
| FIGURA 58: CROMATOGRAMAS DE ÍONS REPRESENTATIVOS OBTIDOS PARA A ANÁLISE DA CURVA DE DEUTERAÇÃO DO COMPLEXO BJ46A - JARARAGINA.....  | 131 |
| FIGURA 59: MAPA DE DEUTERAÇÃO ( <i>HEATMAP</i> ) DE BJ46A LIVRE. ....   | 134 |
| FIGURA 60: ANÁLISE SUBTRATIVA DA INCORPORAÇÃO ISOTÓPICA EM BJ46A LIVRE AO FINAL DE 10.000S DE EXPOSIÇÃO A D <sub>2</sub> O. ....  | 135 |
| FIGURA 61: MAPA DE DEUTERAÇÃO ( <i>HEATMAP</i> ) DE JARARAGINA LIVRE. ....  | 139 |
| FIGURA 62: ANÁLISE SUBTRATIVA DA INCORPORAÇÃO ISOTÓPICA EM JARARAGINA LIVRE, AO FINAL DE 10.000S DE EXPOSIÇÃO A D <sub>2</sub> O. ....  | 140 |
| FIGURA 63: GRÁFICO RESIDUAL DA INCORPORAÇÃO ISOTÓPICA DE BJ46A EM COMPLEXO. ....  | 144 |
| FIGURA 64: CURVAS DE INCORPORAÇÃO REPRESENTATIVAS PARA REGIÕES DO PRIMEIRO DOMÍNIO CISTATINA DE BJ46A.....  | 145 |
| FIGURA 65: CURVAS DE INCORPORAÇÃO REPRESENTATIVAS PARA REGIÕES DO SEGUNDO DOMÍNIO CISTATINA DE BJ46A.....   | 146 |
| FIGURA 66: ANÁLISE DETALHADA DA INCORPORAÇÃO ISOTÓPICA DO INIBIDOR BJ46A EM SUA FORMA LIVRE E EM COMPLEXO COM JARARAGINA, APÓS 10.000 S DE EXPOSIÇÃO A D <sub>2</sub> O. ....                                       | 147 |
| FIGURA 67: REGIÕES PONTUAIS DE DIFERENÇA EM INCORPORAÇÃO ISOTÓPICA PARA BJ46A. ....   | 148 |
| FIGURA 68: GRÁFICO RESIDUAL DA INCORPORAÇÃO ISOTÓPICA DE JARARAGINA EM COMPLEXO.....  | 150 |
| FIGURA 69: CURVAS DE INCORPORAÇÃO REPRESENTATIVAS PARA AS REGIÕES DIFERENCIAIS DE INCORPORAÇÃO EM JARARAGINA. ....  | 151 |

|  |     |
|--|-----|
| FIGURA 70: ANÁLISE DETALHADA DA INCORPORAÇÃO ISOTÓPICA DE JARARAGINA EM SUA FORMA LIVRE E EM COMPLEXO COM BJ46A, APÓS 10.000 S DE EXPOSIÇÃO A D <sub>2</sub> O... 152    | 152 |
| FIGURA 71: CURVAS DE INCORPORAÇÃO REPRESENTATIVAS DE MUDANÇAS PONTUAIS DE INCORPORAÇÃO ISOTÓPICA NOS DOMÍNIOS DESINTEGRINA E RICO EM CISTEÍNA DE JARARAGINA. .... 153    | 153 |
| FIGURA 72: ALINHAMENTO DE SEQUÊNCIAS ENTRE JARARAGINA (JARAR) E BOTROPASINA (BOTHR). .... 155  | 155 |
| FIGURA 73: O DOMÍNIO METALOENDOPEPTIDASE (M) DE BOTROPASINA (PDB ID: 3DSL). .... 158   | 158 |
| FIGURA 74: A ESTRUTURA TRIDIMENSIONAL COMPLETA DA SVMP BOTROPASINA (PDB ID: 3DSL). .... 159  | 159 |
| FIGURA 75: MODELO ESTRUTURAL DA SVMP JARARAGINA, TENDO A ESTRUTURA 3DSL (BOTROPASINA) COMO MOLDE. .... 161   | 161 |
| FIGURA 76: <i>CROSS-LINKS</i> BS <sup>3</sup> E <i>ZERO-LENGTH</i> VALIDADOS NO MODELO ESTRUTURAL DE JARARAGINA. .... 167  | 167 |
| FIGURA 77: ANÁLISE DA ESTRUTURA PRIMÁRIA DE BJ46A PELO ALGORITMO PSIPRED... 168  | 168 |
| FIGURA 78: PREDIÇÃO DE REGIÕES INTRINSECAMENTE DESORDENADAS NO INIBIDOR BJ46A PELO SERVIDOR META-DISORDER. .... 170  | 170 |
| FIGURA 79: ALINHAMENTO DE SEQUÊNCIAS ENTRE BJ46A E A ESTRUTURA 2CH9 (CISTATINA F HUMANA DIMÉRICA). .... 172  | 172 |
| FIGURA 80: MODELO ESTRUTURAL DO DOMÍNIO C1 DE BJ46A. .... 173  | 173 |
| FIGURA 81: NÚMERO DE OCORRÊNCIAS DOS RESÍDUOS DE AMINOÁCIDOS DO DOMÍNIO C1 DE BJ46A PARA A ESTRATÉGIAS BS <sup>3</sup> E XPLEX (DIAMINA E <i>ZERO-LENGTH</i> ). .... 177 | 177 |
| FIGURA 82: <i>CROSS-LINKS</i> VALIDADOS NO MODELO FINAL DO DOMÍNIO C1 DE BJ46A..... 178  | 178 |
| FIGURA 83: ALINHAMENTO DAS SEQUÊNCIAS DO MOLDE 2CH9 E DO DOMÍNIO C2 DE BJ46A. .... 180   | 180 |
| FIGURA 84: MODELO ESTRUTURAL DO DOMÍNIO C2 DE BJ46A. .... 181  | 181 |
| FIGURA 85: NÚMERO DE OCORRÊNCIAS DOS RESÍDUOS DE AMINOÁCIDOS DO DOMÍNIO C2 DE BJ46A PARA A ESTRATÉGIAS BS <sup>3</sup> E XPLEX (DIAMINA E <i>ZERO-LENGTH</i> ). .... 184 | 184 |
| FIGURA 86: <i>CROSS-LINKS</i> VALIDADOS NO MODELO FINAL DO DOMÍNIO C2 DE BJ46A..... 185  | 185 |
| FIGURA 87: RESÍDUOS ENVOLVIDOS EM <i>CROSS-LINKS</i> INTER-PROTEÍNA NO DOMÍNIO C1 DE BJ46A. .... 187   | 187 |
| FIGURA 88: MAPA REPRESENTATIVO DA PROTEÇÃO/EXPOSIÇÃO À TROCA HIDROGÊNIO-DEUTÉRIO NO DOMÍNIO C1 DE BJ46A. .... 189  | 189 |

|  |     |
|--|-----|
| FIGURA 89: CLASSIFICAÇÃO DO GRUPO DE MODELOS COM 11 VIOLAÇÕES DE ACORDO COM A VALIDAÇÃO/VIOLAÇÃO DAS RESTRIÇÕES DE DISTÂNCIA. .... | 193 |
| FIGURA 90: DIFERENTES ORIENTAÇÕES DO DOMÍNIO C1 DE BJ46A EM RELAÇÃO À JARARAGINA. ....   | 194 |
| FIGURA 91: <i>CROSS-LINKS</i> VALIDADOS NO MODELO FINAL DO <i>DOCKING</i> C1-BJ46A/JARARAGINA.....                                 | 196 |
| FIGURA 92: REPRESENTAÇÃO DAS REGIÕES DIFERENCIAIS DE PROTEÇÃO/EXPOSIÇÃO A DEUTÉRIO NO MODELO ESTRUTURAL DE JARARAGINA. ....        | 199 |
| FIGURA 93: A INTERAÇÃO ENTRE O DOMÍNIO C1 DE BJ46A E JARARAGINA.....   | 200 |
| FIGURA 94: INTERAÇÃO ENTRE O DOMÍNIO C1 DE BJ46A E JARARAGINA, COM INCORPORAÇÃO DOS DADOS DE XL-MS E HDX-MS. ....                  | 201 |
| FIGURA 95: <i>CROSS-LINKS</i> MAPEADOS NO DOMÍNIO C2-BJ46A.....  | 203 |
| FIGURA 96: MAPA REPRESENTATIVO DA PROTEÇÃO/EXPOSIÇÃO À TROCA HIDROGÊNIO-DEUTÉRIO NO DOMÍNIO C2 DE BJ46A. ....                      | 205 |
| FIGURA 97: ANÁLISE DO GRUPO CONTENDO 11 VIOLAÇÕES DE ACORDO COM A VALIDAÇÃO/VIOLAÇÃO DE <i>CROSS-LINKS</i> .....                   | 206 |
| FIGURA 98: DIFERENTES ORIENTAÇÕES DE C2-BJ46A QUANTO AO MODELO C1-JARARAGINA. ....   | 208 |
| FIGURA 99: INTERAÇÃO C2-BJ46A COM O COMPLEXO C1-BJ46A/JARARAGINA. ....   | 209 |
| FIGURA 100: <i>CROSS-LINKS</i> VALIDADOS NO <i>DOCKING</i> DOS DOMÍNIOS C2-C1 DE BJ46A COM A METALOENDOPEPTIDASE JARARAGINA.....   | 211 |

## LISTA DE TABELAS

|   |     |
|---|-----|
| TABELA 1: ACIDENTES OFÍDICOS NOTIFICADOS NAS REGIÕES BRASILEIRAS EM 2016.....   | 2   |
| TABELA 2 : CARACTERÍSTICAS DOS SOROS ANTIOFÍDICOS PRODUZIDOS NO BRASIL.....   | 10  |
| TABELA 3: AGENTES DE LIGAÇÃO CRUZADA MAIS UTILIZADOS NA TÉCNICA DE XL-MS. ....  | 24  |
| TABELA 4: DISTÂNCIAS EUCLIDIANAS CONSIDERADAS PARA VALIDAÇÃO DE <i>CROSS-LINKS</i> ...  | 57  |
| TABELA 5: CÓDIGOS DE RESULTADO DO <i>SOFTWARE</i> TOPOLINK, UTILIZANDO O ALGORITMO CONSIDERANDO APENAS OS <i>LINKS</i> OBSERVADOS EXPERIMENTALMENTE.....  | 58  |
| TABELA 6: PARÂMETROS UTILIZADOS NAS SIMULAÇÕES DE <i>DOCKING</i> MOLECULAR.....   | 59  |
| TABELA 7: RENDIMENTO PURIFICATIVO DO NOVO MÉTODO DE PURIFICAÇÃO DE BJ46A E COMPARAÇÃO COM O MÉTODO ORIGINALMENTE PUBLICADO.....   | 68  |
| TABELA 8: ALINHAMENTOS DOS PEPTÍDEOS SEQUENCIADOS <i>DE NOVO</i> PELA FERRAMENTA PEPEXPLORER. ....  | 69  |
| TABELA 9: COBERTURA DE SEQUÊNCIA E NÚMERO DE PEPTÍDEOS DETECTADOS PARA O INIBIDOR BJ46A EM DUAS CONFIGURAÇÕES DA FONTE DE IONIZAÇÃO $\mu$ ESI. ....   | 118 |
| TABELA 10: COBERTURA DE SEQUÊNCIA E NÚMERO DE PEPTÍDEOS IDENTIFICADOS PARA JARARAGINA UTILIZANDO A SOLUÇÃO <i>QUENCH</i> 0.5 M GUANIDINA-HCL, 0,2 M TCEP, 0,2 M FOSFATO DE SÓDIO PH 2,5, EM DUAS FAIXAS DE <i>M/Z</i> PARA O ESPECTRO MS1. .... | 119 |
| TABELA 11: ANÁLISE PELO <i>SOFTWARE</i> TOPOLINK DOS <i>CROSS-LINKS</i> ENCONTRADOS PARA JARARAGINA LIVRE USANDO O ALC BS <sup>3</sup> .....  | 163 |
| TABELA 12: ANÁLISE PELO <i>SOFTWARE</i> TOPOLINK DOS <i>CROSS-LINKS</i> ENCONTRADOS PARA JARARAGINA LIVRE USANDO A ABORDAGEM XPLEX / <i>ZERO-LENGTH</i> .....   | 165 |
| TABELA 13: DEMONSTRAÇÃO DOS 10 MELHORES ALINHAMENTOS OBTIDOS PELO ALGORITMO LOMETS, ORDENADOS PELO <i>Z-SCORE</i> NORMALIZADO (NORM. <i>Z-SCORE</i> ).....  | 171 |
| TABELA 14: ANÁLISE POR <i>CLUSTERIZAÇÃO</i> DOS MODELOS GERADOS PARA O DOMÍNIO C1 DE BJ46A .....  | 172 |
| TABELA 15: ANÁLISE PELO <i>SOFTWARE</i> TOPOLINK DAS RESTRIÇÕES DE DISTÂNCIA GERADAS PELO ALC BS <sup>3</sup> E VALIDADAS NO MODELO DO DOMÍNIO C1 DE BJ46A. ....  | 174 |
| TABELA 16: ANÁLISE PELO <i>SOFTWARE</i> TOPOLINK DAS RESTRIÇÕES DE DISTÂNCIA GERADAS PELA ABORDAGEM XPLEX (DIAMINA) E VALIDADAS NO MODELO DO DOMÍNIO C1 DE BJ46A. ....  | 175 |
| TABELA 17: ANÁLISE PELO <i>SOFTWARE</i> TOPOLINK DAS RESTRIÇÕES DE DISTÂNCIA GERADAS PELA ABORDAGEM XPLEX ( <i>ZERO-LENGTH</i> ) E VALIDADAS NO MODELO C1 DE BJ46A..  | 175 |

|   |     |
|---|-----|
| TABELA 18: MELHORES ALINHAMENTOS OBTIDOS PELO ALGORITMO LOMETS, ORDENADOS PELO Z-SCORE NORMALIZADO.....   | 179 |
| TABELA 19: ANÁLISE DA MODELAGEM MOLECULAR DE C2-BJ46A POR CLUSTERIZAÇÃO..   | 180 |
| TABELA 20: ANÁLISE PELO SOFTWARE TOPOLINK DAS RESTRIÇÕES DE DISTÂNCIA GERADAS PELO ALC BS <sup>3</sup> E VALIDADAS NO MODELO ESTRUTURAL DE C2 – BJ46A. ....           | 182 |
| TABELA 21: ANÁLISE PELO SOFTWARE TOPOLINK DAS RESTRIÇÕES DE DISTÂNCIA GERADAS PELA ABORDAGEM XPLEX – DIAMINA E VALIDADAS NO MODELO ESTRUTURAL DE C2 – BJ46A.....      | 182 |
| TABELA 22: ANÁLISE PELO SOFTWARE TOPOLINK DAS RESTRIÇÕES DE DISTÂNCIA GERADAS PELA ABORDAGEM XPLEX – ZERO-LENGTH E VALIDADAS NO MODELO ESTRUTURAL DE C2 – BJ46A. .... | 183 |
| TABELA 23: RESTRIÇÕES DE DISTÂNCIA INTER-DOMÍNIO DERIVADAS DE CROSS-LINKS PARA O DOCKING MOLECULAR DE C1-BJ46A EM JARARAGINA. ....                                    | 186 |
| TABELA 24: ANÁLISE DO DOCKING JARARAGINA / C1-BJ46A, DE ACORDO COM O NÚMERO DE VIOLAÇÕES CALCULADAS PELO SOFTWARE TOPOLINK. ....                                      | 190 |
| TABELA 25: ANÁLISE DO GRUPO CONTENDO 11 VIOLAÇÕES, REPRESENTADOS PELOS MODELOS DE MENOR ENERGIA DE CADA SUBGRUPO. ....  | 191 |
| TABELA 26: ANÁLISE PELO SOFTWARE TOPOLINK DO NOVO DOCKING C1-BJ46A – JARARAGINA, EXCLUINDO RESTRIÇÕES CONFLITANTES. ....  | 192 |
| TABELA 27: VALIDAÇÃO DE CROSS-LINKS NO MODELO SELECIONADO DE DOCKING DO DOMÍNIO C1 - JARARAGINA.....  | 195 |
| TABELA 28: RESTRIÇÕES DE DISTÂNCIA INTER-DOMÍNIO DERIVADA DE XL-MS PARA O DOCKING ENTRE O SEGUNDO DOMÍNIO CISTATINA DE BJ46A – JARARAGINA.....                        | 202 |
| TABELA 29: ANÁLISE DO DOCKING C2 – BJ46A E JARARAGINA PELO SOFTWARE TOPOLINK .....  | 206 |
| TABELA 30: MAPA DE RESTRIÇÕES VALIDADAS DOS MODELOS DE MENOR ENERGIA NOS GRUPOS DE 10 E 11 VIOLAÇÕES.....   | 207 |
| TABELA 31: MAPA DOS CROSS-LINKS VALIDADOS NO MODELO DE DOCKING C2 - JARARAGINA .....  | 210 |

## LISTA DE ABREVIATURAS E SIGLAS

As nomenclaturas de termos de espectrometria de massas adotadas nesta tese estão de acordo com o descrito em (Vessecchi *et al.*, 2011).

A – Ampère

Å - Ångström =  $10^{-10}$  metro

ACS – *American Chemical Society* (Sociedade Americana de Química)

ADAM – *A disintegrin and metalloproteinase* (metaloendopeptidase com domínio desintegrina)

ALC – agente de ligação cruzada

AU – unidade de absorvância

BSA – *bovine serum albumin* (soroalbumina bovina)

BS3 – bis (sulfosuccinimidil) suberato

cDNA – ácido desoxirribonucleico complementar

CID – *collision-induced dissociation* (dissociação induzida por colisão)

Da - Dalton

DAB – 3,3-diaminobenzidina

DAMP – *danger / Damage-associated molecular patterns* (padrões moleculares associados ao perigo/dano)

DDA - *data-dependent acquisition* (aquisição dependente de dados)

DLLS – *dynamic laser light scattering* (espalhamento dinâmico de luz)

DSS – disulfosuccinimidil suberato

DTT – Ditioneitol

EDTA- $\text{Na}_2 \cdot 2\text{H}_2\text{O}$  – *Ethylenediaminetetraacetic acid disodium salt dihydrate* (Sal dissódico de ácido etilenodiaminotetracético diidratado)

EDC – *N*-(3-Dimethylaminopropyl) -*N'*-ethylcarbodiimide hydrochloride (cloreto de 1-(3-dimetilaminopropil) etilcarbodiimida)

ESI – *electrospray ionization* (ionização por eletrospray)

FA – *formic acid* - ácido fórmico

FWHM - *full width at half maximum* – largura total do pico à meia-altura

HCD - *Higher-energy collision dissociation* - dissociação induzida por colisão de maior energia

HDX-MS – *hydrogen-deuterium exchange mass spectrometry* - Troca isotópica hidrogênio-deutério monitorada por espectrometria de massas

HEPES – ácido 4-(2-hidroxietil) piperazina etanosulfônico

HOBt- N-hidroxibenzotriazol

HPLC – *high performance liquid chromatography* (cromatografia líquida de alta performance)

i.d. - *inner diameter* – diâmetro interno

IDR – *intrinsically disordered region* – região intrinsecamente desordenada

IgG - Imunoglobulina G

IM-MS – *ion mobility mass spectrometry* – espectrometria de massas por mobilidade iônica

ITC - *isothermal titration calorimetry* - calorimetria de titulação isotérmica

LC-MS – *liquid chromatography-mass spectrometry* – cromatografia líquida acoplada a espectrometria de massas

MMP – *matrix metalloendopeptidases* – metaloendopeptidases de matriz

XX

MOPS – ácido 3-*N*-morfolinopropanossulfônico

MS – Espectrometria de massas / Espectrômetro de Massas

*m/z* – razão entre massa e carga

NCE – *normalized collision energy* (Energia de colisão normalizada)

NHS – N-hidroxisuccinimida

PAGE – *polyacrylamide gel electrophoresis* (eletroforese em gel de poliacrilamida)

pD - potencial deuteriônico = pH + 0,4

PDB - *Protein data bank* (repositório de estruturas proteicas)

pH – potencial hidrogeniônico = -log[H<sup>+</sup>]

pI – ponto isoelétrico

PLI - *phospholipase A<sub>2</sub> inhibitor* (inibidor de fosfolipase A<sub>2</sub>)

PSM – *peptide-spectrum match* - correspondência peptídeo-espectro

REU – *Rosetta energy units* – unidades de energia Rosetta

RMN – Ressonância Magnética Nuclear

RMSD – *Root-mean square deviation* – desvio médio quadrático

SAXS – *small-angle X-ray scattering* – espalhamento de raios-X a baixo ângulo

SCX – *strong cation exchange* - cromatografia de troca catiônica forte

SDS – *sodium dodecyl sulfate* (dodecil sulfato de sódio)

SEC – *size-exclusion chromatography* – cromatografia de exclusão molecular

SPR - *surface plasmon resonance* – Ressonância plasmônica de superfície

SVMP - *snake venom metalloendopeptidase* (metaloendopeptidase de veneno de serpente)

SVMPI - *snake venom metalloendopeptidase inhibitor* (inibidor de metaloendopeptidase de veneno de serpente)

TCA – *trichloroacetic acid* (ácido tricloroacético)

TCEP - tris (2-carboxietil) fosfina

TFA – *trifluoroacetic acid* (ácido trifluoroacético)

Tris – 2-amino-2-(hidroximetil) propano-1,3-diol

V – Volts

VDW – Van der Waals

vWF – *von Willebrand Factor* (fator de von Willebrand)

XL – *cross-link* (ligação cruzada)

XL-MS – *cross-linking mass spectrometry* (ligação cruzada acoplada à espectrometria de massas)

**Notações de única letra e valores de massa monoisotópica de cada resíduo de aminoácido, empregados nesta tese:**

| <b>Resíduo</b>         | <b>Notação</b> | <b>Massa monoisotópica (Da)</b> |
|------------------------|----------------|---------------------------------|
| <b>Alanina</b>         | A              | 71,03                           |
| <b>Arginina</b>        | R              | 156,10                          |
| <b>Asparagina</b>      | N              | 114,04                          |
| <b>Ácido aspártico</b> | D              | 115,02                          |
| <b>Ácido glutâmico</b> | E              | 129,04                          |
| <b>Cisteína</b>        | C              | 103,00                          |
| <b>Fenilalanina</b>    | F              | 147,06                          |
| <b>Glutamina</b>       | Q              | 128,05                          |
| <b>Glicina</b>         | G              | 57,02                           |
| <b>Histidina</b>       | H              | 137,05                          |
| <b>Isoleucina</b>      | I              | 113,08                          |
| <b>Leucina</b>         | L              | 113,08                          |
| <b>Lisina</b>          | K              | 128,09                          |
| <b>Metionina</b>       | M              | 131,04                          |
| <b>Prolina</b>         | P              | 97,05                           |
| <b>Serina</b>          | S              | 87,03                           |
| <b>Treonina</b>        | T              | 101,40                          |
| <b>Triptofano</b>      | W              | 186,07                          |
| <b>Tirosina</b>        | Y              | 163,06                          |
| <b>Valina</b>          | V              | 99,06                           |

## Sumário

|   |           |
|---|-----------|
| 1 – INTRODUÇÃO .....  | 1         |
| 1.1– EPIDEMIOLOGIA DO OFIDISMO .....  | 1         |
| 1.2– ACIDENTES BOTRÓPICOS .....   | 4         |
| <b>1.2.1 - Metaloendopeptidases de venenos de serpentes (SVMP) .....</b>                                    | <b>5</b>  |
| 1.3–TERAPIA ANTIOFÍDICA: SITUAÇÃO ATUAL E NOVAS ESTRATÉGIAS .....   | 9         |
| <b>1.3.1 – Inibidores Naturais de Venenos de Serpentes .....</b>  | <b>13</b> |
| 1.4 – MÉTODOS BIOFÍSICOS PARA A DETERMINAÇÃO ESTRUTURAL DE<br>PROTEÍNAS .....                               | 17        |
| <b>1.4.1 – A espectrometria de massas na biologia estrutural integrativa .....</b>                          | <b>21</b> |
| 2 - OBJETIVOS .....   | 35        |
| 2.1 - OBJETIVO GERAL .....  | 35        |
| 2.2 - OBJETIVOS ESPECÍFICOS .....   | 35        |
| 3 - METODOLOGIA .....   | 36        |
| 3.1- MATERIAL .....   | 36        |
| 3.2 – MÉTODOS .....   | 38        |
| <b>3.2.1 – Purificação do inibidor de metaloendopeptidases BJ46a .....</b>                                  | <b>38</b> |
| <b>3.2.2 - Purificação da metaloendopeptidase de veneno de serpente<br/>        (jararagina).....</b>       | <b>39</b> |
| <b>3.2.3 - Quantificação de proteínas .....</b>   | <b>40</b> |
| <b>3.2.4 - Eletroforese em gel de poliacrilamida.....</b>   | <b>40</b> |
| <b>3.2.5 - Coloração de géis por Coomassie R-250 / G-250.....</b>   | <b>40</b> |
| <b>3.2.6 – Revelação de géis por impregnação por nitrato de prata .....</b>                                 | <b>41</b> |
| <b>3.2.7 – <i>Imunoblotting</i> do plasma de <i>B. jararaca</i> e do inibidor BJ46a .....</b>               | <b>41</b> |
| <b>3.2.8 - Aquisição de imagens de géis de poliacrilamida e membranas de<br/>        nitrocelulose.....</b> | <b>42</b> |
| <b>3.2.9 – Digestão de proteínas em solução para análise por espectrometria<br/>        de massas.....</b>  | <b>42</b> |

|   |     |
|---|-----|
| 3.2.10 – Sequenciamento <i>de novo</i> de BJ46a por espectrometria de massas .....  | 43  |
| 3.2.11 - Experimento de formação de complexos entre o inibidor BJ46a e a metaloendopeptidase jararagina .....   | 45  |
| 3.2.12 – Ensaio de inibição da atividade proteolítica de jararagina por BJ46a .....   | 45  |
| 3.2.13 – Determinação experimental das pontes dissulfeto de BJ46a por espectrometria de massas .....  | 45  |
| 3.2.14 – <i>Cross-linking</i> entre resíduos básicos (BS <sup>3</sup> ) o inibidor BJ46a e da metaloendopeptidase jararagina monitorado por espectrometria de massas (XL-MS)..... | 47  |
| 3.2.15 – <i>Cross-linking</i> entre resíduos ácidos (XPLex) entre o inibidor BJ46a e a metaloendopeptidase jararagina monitorado por espectrometria de massas (XL-MS) .....       | 47  |
| 3.2.16 – Análise por espectrometria de massas de amostras de <i>cross-linking</i> .....   | 48  |
| 3.2.17 - Análise bioinformática de dados de XL-MS.....  | 49  |
| 3.2.18 - Troca isotópica hidrogênio-deutério monitorada por espectrometria de massas (HDX-MS) .....   | 50  |
| 3.2.19 - Análise bioinformática dos dados de HDX-MS.....  | 53  |
| 3.2.20 – Análises computacionais .....  | 56  |
| 3 – RESULTADOS E DISCUSSÃO .....  | 61  |
| 4.1 – CARACTERIZAÇÃO FÍSICO-QUÍMICA DO INIBIDOR DE METALOENDOPEPTIDASES BJ46A .....   | 61  |
| 4.2 – ENSAIOS DE INTERAÇÃO ENTRE O INIBIDOR BJ46A E A METALOENDOPEPTIDASE JARARAGINA.....   | 76  |
| 4.3 – MAPEAMENTO DE RESTRIÇÕES DE DISTÂNCIA PARA BJ46A, JARARAGINA E O COMPLEXO BJ46A - JARARAGINA .....  | 82  |
| 4.4 - ANÁLISE DO COMPLEXO BJ46A – JARARAGINA POR TROCA ISOTÓPICA HIDROGÊNIO/DEUTÉRIO (HDX-MS).....  | 110 |
| 4.4.1 – Construção de um sistema especializado para aquisição de dados de HDX-MS .....  | 110 |
| 4.4.2 – Otimização das condições experimentais para a curva de incorporação de deutério .....   | 116 |

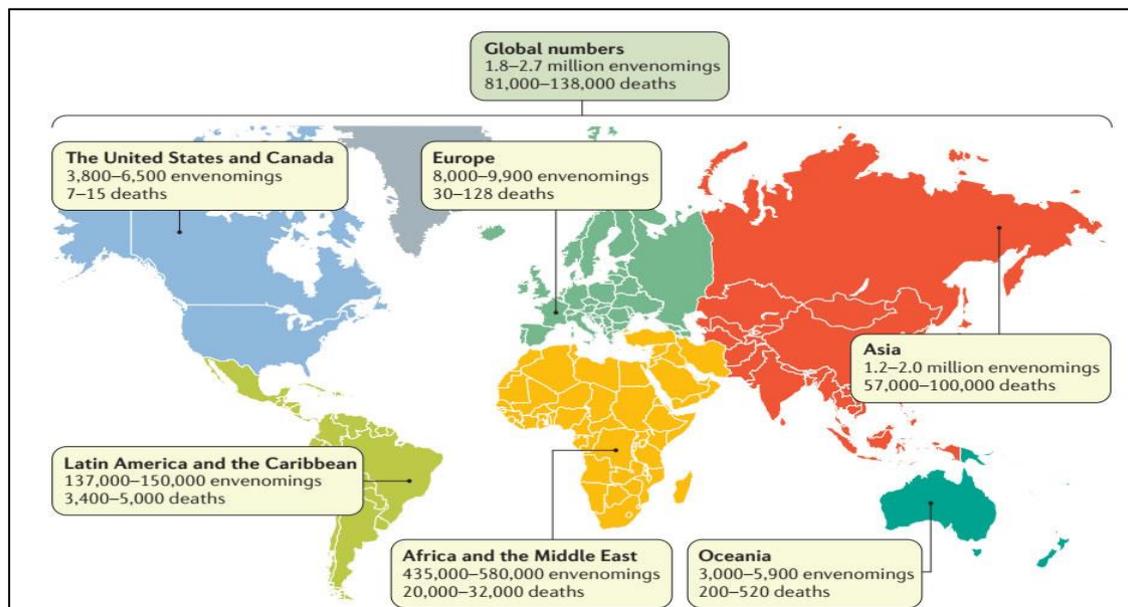
|   |     |
|---|-----|
| 4.4.3 – Curva experimental de incorporação de deutério .....  | 127 |
| 4.5 - INTEGRAÇÃO DOS DADOS DE XL-MS E HDX-MS NA MODELAGEM MOLECULAR E <i>DOCKING</i> DO COMPLEXO BJ46A-JARARAGINA ..... | 154 |
| 4.5.1 – Modelagem molecular da metaloendopeptidase jararagina.....  | 154 |
| 4.5.2 – Modelagem molecular do inibidor BJ46a.....  | 168 |
| 4.5.3 – <i>Docking</i> molecular de BJ46a – jararagina .....  | 186 |
| 5 - CONCLUSÕES .....  | 212 |
| 6 – REFERÊNCIAS BIBLIOGRÁFICAS .....  | 213 |

## 1 – INTRODUÇÃO

### 1.1– EPIDEMIOLOGIA DO OFIDISMO

O envenenamento por serpentes é considerado uma das condições de saúde mais negligenciadas em todo o mundo (Fry, 2018). As estatísticas mais recentes apontam para uma incidência anual de 1,8 a 2,7 milhões de acidentes, com estimativas na faixa de 81.000 a 138.000 mortes nos cinco continentes, como demonstra a **Figura 1** (Gutiérrez *et al.*, 2017).

Cerca de 95% dos casos de envenenamento ofídico ocorrem em países tropicais e/ou em desenvolvimento, afetando a parcela mais pobre da população residente em áreas rurais e remotas, onde o acesso aos sistemas de saúde é limitado e a infraestrutura dos centros hospitalares é frequentemente inadequada, o que contribui para a subestimação dos índices em todo o mundo (Harrison *et al.*, 2009). A situação se agrava porque os acidentes ofídicos não recebem a atenção necessária das indústrias farmacêuticas, tampouco das autoridades de saúde pública – em 2013, a Organização Mundial de Saúde removeu os acidentes ofídicos da lista de doenças tropicais negligenciadas, reintroduzindo-os apenas em 2017, agora na categoria A (Bagcchi, 2015; Chippaux, 2017).



**Figura 1: Incidência mundial de acidentes ofídicos.**

Reproduzido de Gutiérrez (2017)

No Brasil, os envenenamentos ofídicos são um problema de saúde desde o período colonial – a carta do jesuíta espanhol José de Anchieta ao padre Diego Laynes, em Roma, datada de 1560, descreve acidentes ofídicos causados por diversos gêneros de serpentes peçonhentas existentes no país (Bochner, 2003).

Atualmente, estão catalogadas 442 espécies e subespécies de serpentes no Brasil, das quais 76 são peçonhentas. Estas estão distribuídas por todo o território nacional e algumas espécies são endêmicas (Caldeira Costa e Bérnils, 2018). As serpentes de importância médica pertencem às famílias *Elapidae* e *Viperidae*, sendo os acidentes divididos em quatro grandes tipos: acidentes botrópicos (causados por serpentes dos gêneros *Bothrops* e *Bothrocophias*, conhecidas popularmente como jararaca, jararacuçu, urutu, caiçaca, comboia), acidentes crotálicos (gênero *Crotalus*, cascavéis), acidentes laquéuticos (gênero *Lachesis*, surucucu-pico-de-jaca) e acidentes elapídicos (gêneros *Micrurus* e *Leptomicrurus*, corais-verdadeiras) (Ministério Da Saúde, 2018).

O Brasil é o país latino-americano com maior incidência de acidentes ofídicos: somente em 2016, o país registrou 21.665 casos no Sistema Nacional de Agravos de Notificação do Ministério da Saúde (SINAN-MS), como demonstra a **Tabela 1** (Gutiérrez e Pereañez, 2016; Sinan, 2018).

**Tabela 1: Acidentes ofídicos notificados nas regiões brasileiras em 2016.**

Fonte: SINAN/MS. Dados atualizados em 25/01/2018 e sujeitos à revisão.

| Região de Notificação | Gênero da Serpente |                 |                 |                 | TOTAL         |
|-----------------------|--------------------|-----------------|-----------------|-----------------|---------------|
|                       | <i>Bothrops</i>    | <i>Crotalus</i> | <i>Micrurus</i> | <i>Lachesis</i> |               |
| Região Norte          | 7270               | 192             | 39              | 469             | 7970          |
| Região Nordeste       | 3957               | 843             | 127             | 28              | 4955          |
| Região Sudeste        | 3493               | 767             | 47              | 3               | 4310          |
| Região Centro-Oeste   | 2190               | 277             | 11              | 19              | 2497          |
| Região Sul            | 1797               | 109             | 17              | 0               | 1923          |
| <b>TOTAL GERAL</b>    |                    |                 |                 |                 | <b>21.655</b> |

Os acidentes ofídicos prevalecem na região Norte do país, afetando principalmente as populações indígenas e ribeirinhas, que vivem basicamente de atividades extrativistas, que as leva a passar um tempo prolongado no interior da floresta amazônica (Waldez e Vogt, 2009). Os envenenamentos apresentam uma sazonalidade muito característica na região, ocorrendo principalmente no período da cheia dos rios, quando as serpentes procuram ambientes mais secos, frequentemente situados próximos a assentamentos humanos. O desflorestamento também é um fator importante para a alta incidência de acidentes, destruindo *habitats* naturais e favorecendo a migração das serpentes para áreas habitadas. É importante ressaltar que os dados referentes ao Norte do país são subestimados, pois uma parcela significativa das vítimas se vale de tratamentos alternativos baseados em conhecimento tradicional, como a aplicação de extratos de ervas no local da mordedura. Dessa forma, as vítimas dispensam o tratamento médico, que muitas vezes inexistente nas comunidades ou é de difícil acesso, o que frequentemente leva ao agravamento do quadro clínico (Oliveira *et al.*, 2018).

Nas demais regiões do Brasil, os envenenamentos ofídicos estão associados com a agropecuária, devido às alterações ambientais decorrentes desta atividade e à presença do trabalhador por muitas horas no campo, para o plantio e a colheita da safra (Wen *et al.*, 2002; Lemos *et al.*, 2009). Na região Nordeste, os trabalhadores do campo apresentaram um risco quatro vezes maior de sofrer acidentes ofídicos quando comparados com os demais trabalhadores, corroborando a visão de que os envenenamentos por serpentes também podem ser considerados uma doença ocupacional, com consequências socioeconômicas significativas (Gutierrez *et al.*, 2006; Mise, 2014).

O perfil epidemiológico dos acidentes ofídicos no Brasil indica que o acidente botrópico, frequentemente envolvendo a serpente *Bothrops jararaca*, prevalece em todas as regiões do país, respondendo por 86% dos casos notificados em 2016, um cenário que se mantém inalterado há mais de 100 anos (Bochner e Struchiner, 2003).

## 1.2– ACIDENTES BOTRÓPICOS

A serpente *Bothrops jararaca* (Wied-Neuwied, 1824) é encontrada em toda a região Centro-Sul do país, nos estados de Goiás, Bahia, Espírito Santo, Minas Gerais, Rio de Janeiro, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul (Sazima, 1992; Caldeira Costa e Bérnils, 2018). As jararacas habitam preferencialmente ambientes úmidos (matas e áreas cultivadas), zonas rurais e periurbanas; apresentam hábitos noturnos e são consideradas as serpentes mais agressivas encontradas no Brasil (França e Fan, 1992).

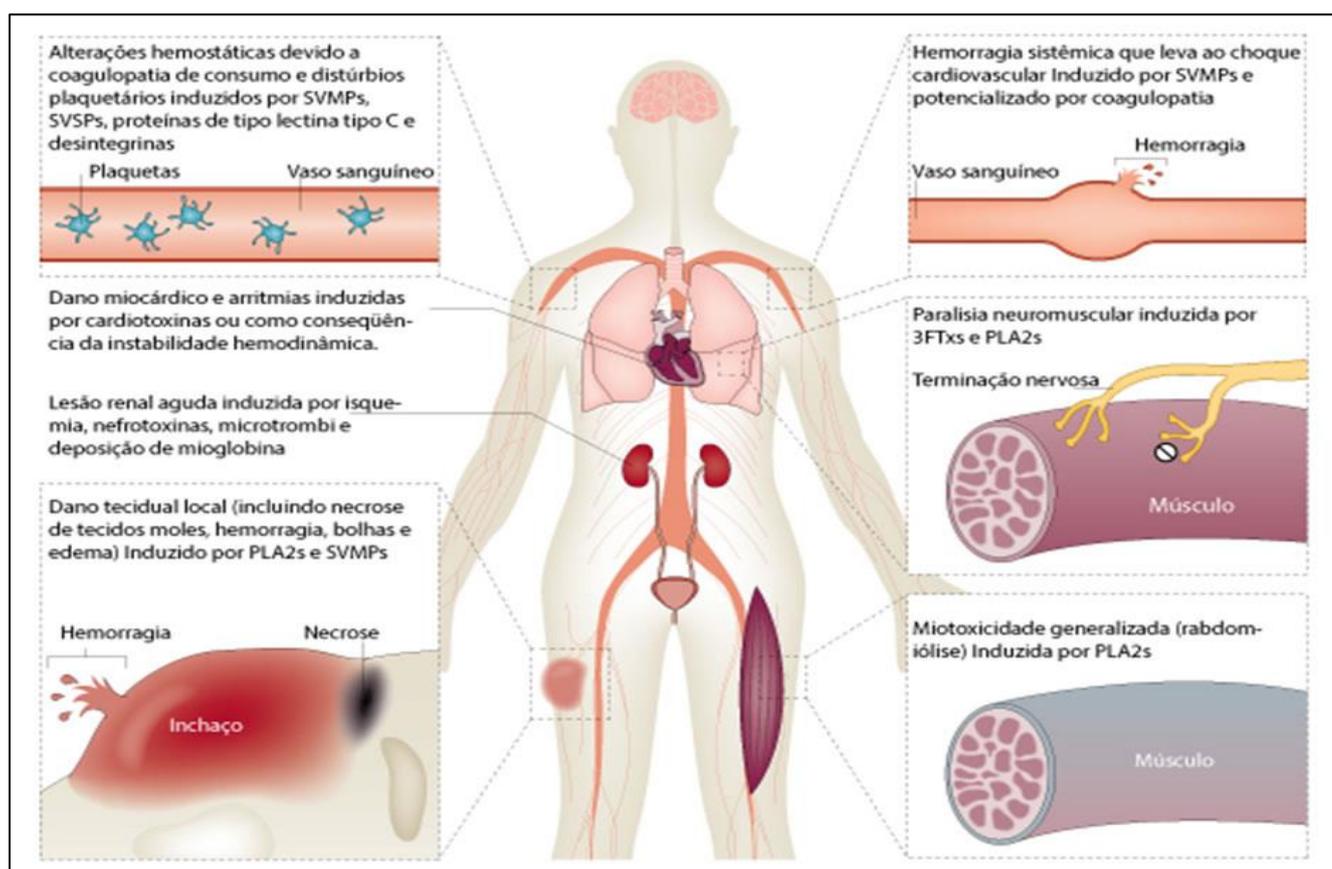
O veneno destas serpentes é uma mistura complexa, moldada pela seleção natural para imobilizar, matar e auxiliar na digestão da presa (Chippaux *et al.*, 1991; Calvete *et al.*, 2009). Há ainda uma extensa variabilidade do veneno, que pode ocorrer em função da ontogenia (serpentes juvenis / adultas), localização geográfica, dieta e sexo (Mackessy, 2010).

O veneno é produzido por uma glândula especializada, localizada na maxila superior, e estocado extracelularmente em um grande lúmen. Para prevenir a degradação do veneno no interior da glândula, as toxinas são estabilizadas por diferentes estratégias, como a presença de pequenos inibidores peptídicos e íons citrato, a síntese de toxinas na forma de zimógenos e a manutenção de um pH baixo do microambiente glandular (Mackessy e Baxter, 2006; Mackessy, 2010). Durante a inoculação, músculos compressores ao redor da glândula (*compressor glandulas*) comprimem-na e o veneno é direcionado para os ductos primário e secundário e então para o canal interno da presa inoculadora (Vitt e Caldwell, 2009). Há evidências de que a serpente é capaz de regular a quantidade de veneno inoculada, dependendo do tamanho da presa e da quantidade de veneno disponível (Hayes, 1995; Vitt e Caldwell, 2009).

As toxinas tornam-se ativas imediatamente após a injeção do veneno na presa. De uma maneira geral, as toxinas de venenos de serpentes atuam nos sistemas nervosos (central e periférico), cardiovascular e hemostático (Calvete *et al.*, 2009). O acidente botrópico é caracterizado por uma intensa ação proteolítica, coagulante e hemorrágica, com danos locais e sistêmicos (Fundação Nacional De Saúde, 2001).

Os danos locais iniciam-se momentos após o envenenamento e são caracterizados pelo surgimento de dor, edema, bolhas, equimose, dermo- e mionecrose, que podem acarretar desde perda tecidual até a amputação do membro afetado (Gutierrez *et al.*,

2009; Warrell, 2010; Wen e Malaque, 2013). Os danos sistêmicos incluem hemorragia e coagulopatia generalizada, podendo haver sangramentos espontâneos nasais e gengivais; hematúria, bradicardia ou taquicardia, hipotensão arterial e choque, como demonstrado na **Figura 2** (Gutiérrez *et al.*, 2017); esta sintomatologia resulta de uma ação sinérgica de várias toxinas. No veneno de *B. jararaca*, as metaloendopeptidases de venenos de serpentes (SVMP) são as toxinas mais abundantes, sendo parte central da patogênese deste envenenamento (Fox e Serrano, 2005; Nicolau *et al.*, 2017).



**Figura 2: Representação esquemática dos efeitos locais e sistêmicos característicos do acidente botrópico.** Reproduzido de Gutiérrez (2017) e adaptado por Soares (2018).

### 1.2.1 - Metaloendopeptidases de venenos de serpentes (SVMP)

As metaloendopeptidases de venenos de serpentes são peptidases dependentes de zinco, pertencentes à subfamília M12 (adamalinas) das metaloendopeptidases, junto com as ADAM (*a disintegrin and metalloendopeptidase*) e as ADAMTS (ADAM com motivos trombospondina), proteínas que exercem papéis relevantes na adesão celular, reprodução e processos inflamatórios (Fox e Serrano, 2005; Giebeler e Zigrino, 2016).

As metaloendopeptidases de serpentes parecem ter evoluído das ADAM, devido à pressão evolutiva nas serpentes para produzir venenos apropriados para a captura de suas presas (Fox e Serrano, 2010).

Baseado nas características dos seus domínios estruturais, as SVMP foram divididas em três grandes classes: PI, PII e PIII. Todas as SVMP compartilham um domínio metaloendopeptidase, caracterizado pelo motivo estrutural HEXXHXXGXXH responsável pela coordenação de um átomo de zinco essencial para a atividade enzimática e uma alça estrutural conservada, contendo um resíduo de metionina (*Met-turn*), que serve como um arcabouço para orientar espacialmente os três resíduos de histidina envolvidos na catálise. As pontes dissulfeto também contribuem para a estabilidade estrutural (Bode *et al.*, 1993). Além disso, são sintetizadas na forma de zimógenos contendo pré e pró-domínios: o pré-domínio (ou peptídeo sinal) é responsável pelo endereçamento das toxinas nascentes ao retículo endoplasmático no interior das células secretoras da glândula, enquanto que o pró-domínio está envolvido na manutenção da latência da toxina (Fox e Serrano, 2010; Moura-Da-Silva *et al.*, 2016)

As metaloendopeptidases de classe PI são compostas de apenas um domínio metaloendopeptidase; as de classe PII, de um domínio metaloendopeptidase acrescido de um domínio desintegrina-símile; por fim, as SVMP de classe PIII apresentam domínios metaloendopeptidase, desintegrina e rico em cisteína; eventualmente, domínios lectina-símile podem estar presentes (Fox e Serrano, 2005). As SVMP de classe PIII teriam sido as primeiras metaloendopeptidases a serem evolutivamente recrutadas para o veneno de viperídeos, enquanto que as PII e PI teriam surgido posteriormente, por meio da deleção de domínios (Casewell *et al.*, 2011). Há também subclasses de SVMP em cada classe descrita, de acordo com a presença ou não de dimerização e padrões de clivagem pós-traducional (Fox e Serrano, 2010).

O mecanismo de ativação e processamento das metaloendopeptidases de venenos de serpentes envolve a hidrólise do pró-domínio, da mesma forma que ocorre com metaloendopeptidases de matriz (MMP) e as ADAM (Gomis-Rüth, 2003). O pró-domínio apresenta uma sequência conservada (PKMCGVT) e o resíduo de cisteína presente neste motivo é coordenado com o átomo de zinco, bloqueando o sítio ativo da enzima, um mecanismo denominado *cysteine-switch* (Grams *et al.*, 1993). Trabalhos recentes indicam que o processamento das SVMP se inicia durante a secreção destas toxinas para o lúmen da glândula de veneno, sendo o pró-domínio hidrolisado em

pequenos peptídeos, que podem eventualmente apresentar propriedades inibitórias (Portes-Junior *et al.*, 2014; Moura-Da-Silva *et al.*, 2016).

No contexto do envenenamento, a injeção local de SVMP induz um efeito hemorrágico devido a lesões nas paredes de capilares sanguíneos (Kamiguti *et al.*, 1996). Os capilares são estruturados por células endoteliais, apoiadas por uma membrana basal que serve de apoio estrutural, fornece fatores de crescimento e modula a atividade destas células (Gutierrez *et al.*, 2010).

A análise do exsudato de camundongos injetados intramuscularmente com SVMP e coletado em dois intervalos de tempo (15 min e 1 h) indicou a presença de vários componentes da membrana basal, como laminina, nidogênio, perlecan, colágeno do tipo IV e proteoglicanos específicos, indicando que o primeiro alvo das SVMP é a membrana basal dos capilares e que a hidrólise é extremamente rápida (Gutiérrez *et al.*, 2016). Esta hidrólise extensiva também lesiona o endotélio, tanto pelo disparo de vias apoptóticas, devido ao descolamento das células da membrana basal, quanto pelo aumento da distensibilidade dos vasos pela falta da estabilidade mecânica conferida pela membrana basal íntegra (Gutierrez *et al.*, 2005; Tanjoni *et al.*, 2005). A hemorragia ocorre então via extravasamento dos eritrócitos por lesões nas próprias células endoteliais, um processo denominado hemorragia *per rhexis* (Gutierrez *et al.*, 2016). Este extravasamento de sangue também promove a isquemia dos músculos locais, contribuindo para o agravamento da mionecrose observada nos acidentes botrópicos (Gutierrez e Rucavado, 2000).

Concomitante à hemorragia local, as SVMP também podem induzir a incoagulabilidade sanguínea sistêmica, por meio da interação com diversos componentes do sistema hemostático. Em um processo normal de hemostasia, há formação inicial de um tampão plaquetário e os tecidos lesados liberam uma molécula chamada fator tecidual (tromboplastina tecidual). Este fator complexa-se com o fator VII da cascata de coagulação sanguínea na presença de íons cálcio, ativando-o (fator VIIa). Por sua vez, o fator VIIa ativa o fator X (Xa) que, na presença de cálcio, fosfolipídeos e fator V converte protrombina (fator II) em trombina (IIa). A trombina então catalisa a formação das redes de fibrina a partir do fibrinogênio, consolidando o tampão plaquetário inicial em um tampão hemostático (Tortora e Grabowski, 2008).

As metaloendopeptidases de venenos de serpentes alteram profundamente a função plaquetária, impedindo a agregação mediada por receptores de superfície, pelo

fator de von Willebrand (vWF), glicoproteína auxiliar na agregação plaquetária, ou pela via dependente de colágeno. O domínio desintegrina, presente em SVMP de classe PII, apresenta um motivo RGD que apresenta uma alta afinidade pela integrina  $\alpha_{IIb}\beta_3$  e inibe a agregação plaquetária por bloquear a interação do fibrinogênio a este receptor (Kamiguti, 2005). Outros motivos com o mesmo efeito biológico já foram reportados: KGD, MDV, MLD, VGD, ECD, MDG, KTS (Calvete *et al.*, 2005). As SVMP de classe PIII possuem um domínio desintegrina-símile (não possuem o motivo RGD) e, portanto, medeiam a inibição da agregação plaquetária por outras vias, como as vias mediadas pelo fator de von Willebrand e colágeno (Moura-Da-Silva e Tanjoni, 2007). Nestas vias, as SVMP conseguem hidrolisar o fator de von Willebrand e impedir a ação das integrinas  $\alpha_2\beta_1$  e GPVI, esta última responsável pela agregação mediada por colágeno, por bloqueio do receptor ou por hidrólise (Sajevic *et al.*, 2011).

Paralelamente à perturbação na atividade plaquetária, as SVMP também interferem com a atividade dos fatores de coagulação, principalmente ativando os fatores II e X (Sajevic *et al.*, 2011). Tais alterações são frequentemente utilizadas na prática clínica para avaliar a severidade do envenenamento e a eficácia da terapia (White, 2005). Além dos fatores de coagulação, as SVMP também consomem o fibrinogênio disponível, degradando preferencialmente as cadeias A $\alpha$  e B $\beta$  (Swenson e Markland, 2005). As redes de fibrina também são hidrolisadas por SVMP (Moura-Da-Silva e Baldo, 2012). A coagulopatia induzida pelos venenos botrópicos é tão intensa que 59% dos pacientes admitidos no Hospital Vital Brazil (Instituto Butantan, SP) no período de 1990-2004 apresentavam tempos de coagulação alterados e 40,8% apresentavam sangue incoagulável (Nicoletti *et al.*, 2010).

Além da indução da síndrome hemorrágica, as metaloendopeptidases de venenos de serpentes também estão relacionadas com o quadro inflamatório descrito nos envenenamentos por serpentes viperídeas, que incluem extravasamento de plasma, edema, migração leucocitária e liberação de mediadores inflamatórios, como IL-1 $\beta$  e IL-6 (Gutierrez *et al.*, 2010; Teixeira *et al.*, 2018). Jararagina, uma SVMP de classe PIII isolada do veneno de *B. jararaca*, induz a expressão e o processamento do fator de necrose tumoral  $\alpha$  (TNF- $\alpha$ ) em camundongos e a expressão desta citocina está relacionada diretamente com a indução de dermonecrose local (Laing *et al.*, 2003; Moura-Da-Silva e Baldo, 2012).

Recentemente, foi visto que a hidrólise de diversas moléculas e o dano celular induzido pelas SVMP e outras toxinas presentes no veneno podem gerar DAMP (*damage / danger-associated molecular patterns*), que foram detectados no exsudato local (Rucavado *et al.*, 2016). DAMP são moléculas e/ou fragmentos endógenos que, quando liberados para o meio extracelular, são interpretados por células do sistema imune inato como um sinal de perigo, de injúria tecidual. Os DAMP são reconhecidos por células do sistema imune inato por meio de receptores do tipo Toll (TLR) localizados na superfície destas células (Kato e Svensson, 2015). Rucavado e colaboradores demonstraram que a resposta aos DAMP gerados pelo veneno é mediada pela via de TLR4, que reconhece fragmentos de proteoglicanos (ácido hialurônico), fibronectina, heparan sulfato e fibrinogênio, componentes de matriz extracelular também degradados por SVMP. A injeção do exsudato contendo DAMP em camundongos saudáveis aumentou significativamente a permeabilidade vascular, sendo possível que os DAMP gerados no curso do envenenamento estejam relacionados também com uma atividade procoagulante, por meio da estimulação direta de células endoteliais (Rucavado *et al.*, 2016).

### 1.3–TERAPIA ANTIOFÍDICA: SITUAÇÃO ATUAL E NOVAS ESTRATÉGIAS

A administração endovenosa do soro antiofídico é o único tratamento específico para os acidentes ofídicos (Gutierrez *et al.*, 2006). Desenvolvido em 1894 por Albert Calmette, Césaire Phisalix e Gabriel Bertrand, o soro antiofídico é produzido através da imunização de animais de grande porte com uma mistura contendo doses subletais de venenos de serpentes e adjuvantes para induzir uma resposta imunológica adequada (Espino-Solis *et al.*, 2009). Os soros podem ser mono-específicos (contra venenos de um único gênero de serpente) ou poli-específicos, produzidos com misturas imunizantes mais complexas e que neutralizam venenos oriundos de mais de um gênero de serpente (Girish e Kemparaju, 2011; Silva *et al.*, 2013).

O Brasil é autossuficiente na produção de soros antiofídicos, contando com quatro centros principais de produção: Instituto Butantan (São Paulo/SP), Instituto Vital Brazil (IVB, Niterói / RJ), Fundação Ezequiel Dias (FUNED, Belo Horizonte/MG) e o Centro de Produção e Pesquisa em Imunobiológicos (CPPI, Piraquara/PR). Não há venda comercial de soros antiofídicos: toda a produção é revertida exclusivamente para o

Ministério da Saúde, que efetua a distribuição por todo o país, e as Secretarias Estaduais de Saúde definem os locais de referência para a aplicação dos soros (Silva *et al.*, 2013).

A produção dos soros antiofídicos demanda um protocolo rígido para garantir a segurança e a eficácia da terapia antiofídica. Os soros podem ser compostos de imunoglobulinas G (IgG) íntegras, fragmentos divalentes [F(ab')<sub>2</sub>] ou monovalentes [F(ab)], obtidos a partir da digestão das IgG com pepsina e papaína, respectivamente (Gutierrez *et al.*, 2007). A OMS recomenda a utilização de soros contendo fragmentos F(ab') para minimizar eventuais efeitos adversos (Organização Mundial Da Saúde, 2016). A Portaria do Ministério da Saúde nº174, de 11 de novembro de 1996, estabelece as condições para a produção dos soros antiofídicos brasileiros. As misturas imunizantes-padrão para cada tipo de soro produzido estão descritas na **Tabela 2**.

**Tabela 2 : Características dos soros antiofídicos produzidos no Brasil.**

Reproduzido da Portaria do Ministério da Saúde nº174, de 11 de novembro de 1996.

| <b>Soro</b>                | <b>Composição de venenos da mistura imunizante padrão</b>  | <b>Título neutralizante mínimo<sup>1</sup></b> |
|----------------------------|--|--|
| <b>Soro antibotrópico</b>  | <i>Bothrops jararaca</i> = 50%<br><i>Bothrops moojeni</i> = 12,5%<br><i>Bothrops jararacussu</i> = 12,5%<br><i>Bothrops alternatus</i> = 12,5%<br><i>Bothrops neuwiedi</i> = 12,5% | 5 mg/mL  |
| <b>Soro anticrotálico</b>  | <i>Crotalus durissus terrificus</i> ,<br>(crotamina positivo)  | 1,5 mg/mL                                      |
| <b>Soro antielapídico</b>  | <i>Micrurus corallinus</i><br><i>Micrurus frontalis</i>  | 1,5 mg/mL                                      |
| <b>Soro antilaquéutico</b> | <i>Lachesis muta</i>   | 3,0 mg/mL                                      |

<sup>1</sup> O título neutralizante mínimo é a quantidade (em miligramas) de veneno de referência neutralizado por 1 mililitro de soro antiofídico. O título é então expresso em mg/mL. Admite-se um coeficiente de variação de 10% devido à variabilidade inerente aos ensaios com animais de laboratório (Agência Nacional De Vigilância Sanitária e Fundação Oswaldo Cruz, 2010).

A partir da coleta do plasma hiperimune, as imunoglobulinas (ou seus fragmentos) são purificados por precipitações em grandes tanques com sulfato de amônio, seguido de centrifugação, ultrafiltração e cromatografia. Em seguida, o antissoro é submetido a diversos testes de qualidade para verificar a atividade biológica, a esterilidade, a inocuidade e a presença de pirógenos. Toda a rotina de controle de qualidade é realizada pelo Instituto Nacional de Controle de Qualidade em Saúde (INCQS), órgão ligado à Fundação Oswaldo Cruz (Ministério Da Saúde, 1996; Silva *et al.*, 2013; Cunha, 2017).

Desde a introdução do soro antiofídico na terapêutica dos acidentes ofídicos, as taxas de mortalidade foram reduzidas significativamente (Gutierrez *et al.*, 2007). Quando administrado de maneira adequada, o soro é bastante eficaz na neutralização de efeitos sistêmicos; para os acidentes com serpentes viperídeas, a interrupção da hemorragia se dá entre 30 min e 6 h do início da terapia e a normalização dos tempos de coagulação, em 6 a 24 h (Gutierrez *et al.*, 2007). Por outro lado, o soro é ineficaz na neutralização dos danos locais – experimentos de microscopia intravital em ratos demonstraram que o soro antiofídico não foi capaz de reverter o quadro, mesmo quando administrado rapidamente, apenas 15 minutos após a injeção do veneno (Battellino *et al.*, 2003). Dentre os motivos para esta neutralização ineficaz estão o título de anticorpos ineficiente para neutralização de algumas classes de toxinas de baixa massa molecular e a progressão extremamente rápida do dano local, que compromete a distribuição tecidual das imunoglobulinas e/ou de seus fragmentos (Gutiérrez *et al.*, 1998; Harrison *et al.*, 2011). Na prática, a situação é ainda mais grave: estudos epidemiológicos indicam que, no Brasil, a demora na administração do soro antiofídico (mais de 6 h após o acidente) ocorre em 42% dos casos brandos e moderados de envenenamento e está estatisticamente correlacionada com um mau prognóstico (Bochner *et al.*, 2014; Mise *et al.*, 2018).

Além da neutralização ineficiente dos danos locais, a terapia antiofídica baseada no soro antiofídico também apresenta outras limitações graves: a possível ocorrência de reações adversas, os desafios na produção, a logística de distribuição dos antissoros e o desinteresse das indústrias farmacêuticas para sua produção (Gutierrez *et al.*, 2007).

As reações adversas associadas à soroterapia advêm muitas vezes da administração de grande quantidade de soro antiofídico. A injeção excessiva de proteínas exógenas pode levar a quadros de hipersensibilidade (Girish e Kemparaju, 2011). As reações de hipersensibilidade podem ocorrer de maneira imediata ou tardia.

As reações imediatas (horas após o início da administração do soro) incluem o surgimento de urticária, náusea, cólicas e vômitos, enquanto que as reações tardias (5 a 24 dias após a exposição) incluem a “doença do soro”, com um quadro de febre, artralgia, urticária e adenopatia, agravando ainda mais o quadro clínico do paciente (Morais e Massaldi, 2009; Gutierrez, 2012; Wen e Malaque, 2013).

A produção de antissoro é bastante custosa, pois envolve etapas de imunização, purificação, testes de qualidade e envasamento, tudo em um ambiente controlado, seguindo as boas práticas de produção e de laboratório e obedecendo às legislações locais (Chippaux, 2017). Além disso, o produto é instável e demanda condições específicas para a sua preservação e transporte, como refrigeração contínua, sob risco de perda de atividade biológica (Morais e Massaldi, 2009).

Desta forma, a produção dos soros antiofídicos não é atraente para as grandes indústrias farmacêuticas, pois o antissoro é uma “droga órfã” clássica: um produto caro, com um prazo de validade (*shelf life*) curto e frequentemente destinado a pessoas que não possuem poder aquisitivo para obtê-lo (Fry, 2018). Algumas indústrias já descontinuaram a produção de soros antiofídicos, caso da Sanofi-Pasteur, que decidiu encerrar a produção do soro Fav-Afrique em 2015 devido a uma queda brusca das vendas, levando a uma crise significativa de abastecimento de antissoros na África (Chippaux e Habib, 2015; Arnold, 2016).

Diversas estratégias têm sido propostas na literatura para contornar as limitações do soro antiofídico, como o uso de anticorpos monoclonais e o emprego de formulações mais estáveis, a otimização da mistura imunizante para a neutralização da maior quantidade possível de toxinas e o uso de inibidores que possam atuar localmente, contendo os danos locais (Gutierrez *et al.*, 2007; Harrison *et al.*, 2011). Inibidores sintéticos de metaloendopeptidases de matriz, como os peptideomiméticos marimastat e o batimastat apresentaram resultados animadores de inibição de SVMP, mas ainda esbarram nas limitações de baixa especificidade e biodisponibilidade (Sanchez e Rodriguez-Acosta, 2008; Preciado e Pereañez, 2017). Neste contexto, o estudo de inibidores naturais de metaloendopeptidases de venenos de serpentes torna-se promissor, uma vez que são moléculas específicas para a interação / inibição desta classe de toxinas.

### 1.3.1 – Inibidores Naturais de Venenos de Serpentes

O estudo do fenômeno de resistência natural aos venenos de serpentes teve início com o trabalho pioneiro de Felice Fontana (1730-1805), um abade e eminente fisiologista italiano, atualmente considerado o pai da Toxinologia moderna (Garrison, 1935; Hawgood, 1995). Em seu tratado *Ricerche fisiche sopra il veleno della vipera* (“Pesquisa física sobre o veneno da víbora”, em tradução livre), Fontana relata que induziu uma víbora europeia (*Vipera berus*) a injetar seu veneno em outra víbora. Após meticulosa observação, Fontana descreveu que não houve nenhum sintoma de envenenamento na serpente atingida, mesmo depois de 36 horas, levando-o a redigir seu famoso aforismo “*Il veleno della Vipera non è veleno per la propria specie*”, traduzido livremente como “o veneno da víbora não é veneno para sua própria espécie” (Fontana, 1787).

O trabalho de Fontana lançou as bases do estudo experimental da imunidade natural aos venenos de serpentes. Após seu tratado, evidências experimentais de resistência natural em mamíferos, como o mangusto (*Herpestes edwardsii*) e o ouriço (*Erinaceus europaeus*) também foram relatadas, por Albert Calmette, Césaire Phisalix e Gabriel Bertrand (Phisalix, 1895; Calmette, 1896), seguido de outros estudos e relatos em outros gêneros de serpentes e mamíferos (Bastos *et al.*, 2016).

A resistência natural que estes animais apresentam aos efeitos deletérios do veneno de serpentes pode ser conferida por dois mecanismos não-excludentes: no primeiro, o animal resistente apresenta modificações estruturais em receptores ou proteínas que impedem a ligação de toxinas. Tal mecanismo já foi descrito para receptores nicotínicos de acetilcolina em mamíferos (Barchan *et al.*, 1992; Drabeck *et al.*, 2015) e na serpente *Natrix tessellata* (Neumann *et al.*, 1989). Há também evidências de modificação do fator de von Willebrand em mamíferos dos gêneros *Didelphis*, *Lutreolina* e *Philander*, onde a substituição pontual de alguns resíduos levariam a alterações no padrão de cargas e hidrofobicidade desta proteína, potencialmente interferindo com a especificidade da ligação deste fator com lectinas do tipo C presentes nos venenos (Jansa e Voss, 2011).

O segundo mecanismo envolve a existência de proteínas séricas que neutralizam a ação das toxinas do veneno, por meio da formação de um complexo estável e mantido por forças de natureza não-covalente. Essas proteínas inibitórias são consideradas componentes da imunidade inata destes animais; elas apresentam similaridade

estrutural com outras proteínas importantes do sistema imunológico e atuam como aceptores solúveis prontos para neutralizar eventuais toxinas presentes na corrente sanguínea, sendo uma primeira linha de defesa contra os efeitos tóxicos destas proteínas (Perales e Domont, 2002; Perales *et al.*, 2005). Foram descritos na literatura pelo menos 45 inibidores isolados do plasma de mamíferos e de serpentes resistentes aos efeitos deletérios do envenenamento (Neves-Ferreira *et al.*, 2010; Biardi *et al.*, 2011; Palacio *et al.*, 2017; Serino-Silva *et al.*, 2018).

Há duas subclasses de inibidores: os inibidores de fosfolipase A<sub>2</sub> (PLI – *phospholipase A<sub>2</sub> inhibitor*), que inibem os efeitos miotóxicos e/ou neurotóxicos associados às fosfolipases A<sub>2</sub> de venenos de serpentes, e os inibidores de metaloendopeptidases (SVMPI – *Snake venom metalloendopeptidase inhibitor*), que neutralizam os efeitos hemorrágicos causados por estas toxinas. Os SVMPI isolados do plasma e do tecido muscular de mamíferos foram respectivamente agrupados na superfamília (gênica) das imunoglobulinas e da ficolina/opsonina P35, enquanto aqueles isolados do sangue de serpentes foram classificados como membros da superfamília das cistatinas (Neves-Ferreira *et al.*, 2014). Nesta última, os inibidores mais extensivamente caracterizados são o HSF (*habu serum factor*), isolado do plasma da serpente *Protobothrops (Trimeresurus) flavoviridis* e a proteína BJ46a, isolada da serpente *Bothrops jararaca*, que apresenta 85% de identidade sequencial com o HSF (Valente *et al.*, 2001; Neves-Ferreira *et al.*, 2014).

O inibidor HSF foi o primeiro da superfamília das cistatinas a ser purificado e caracterizado, em 1972. HSF apresenta uma massa molecular de 70 kDa, determinada por ultracentrifugação (sedimentação por equilíbrio) e um ponto isoelétrico de 4,0 (Omori-Satoh *et al.*, 1972). Por espectrometria de massas, HSF apresentou uma massa de 47.810 Da. Ao compararmos este valor com a massa calculada de 70 kDa por ultracentrifugação, é possível que HSF apresente uma natureza dimérica; entretanto, os autores não cogitaram esta possibilidade em seu trabalho (Deshimaru *et al.*, 2003).

A análise da estrutura primária de HSF revelou que o inibidor apresenta dois domínios cistatina contíguos e um domínio rico em histidina em sua porção C-terminal, além de três sítios confirmados de N-glicosilação nas posições N123, N185 e N263 (Yamakawa e Omori-Satoh, 1992). Com relação à atividade biológica, HSF foi capaz de inibir várias metaloendopeptidases das classes PI, PII e PIII do veneno de *P. flavoviridis* (HR1A, HR1B, HR2a, HR2b e H2) e do veneno de *Gloydus halys brevicaudus*

(brevilisin H3, H4, H6 e L4), em uma relação molar de 1:1 (inibidor:toxina). Portanto, HSF apresenta um amplo espectro de inibição, independente da classe de SVMP (Deshimaru *et al.*, 2005). Além de interagir com diversas SVMP, HSF também é capaz de interagir com pequenas proteínas séricas (SSP, *small serum proteins*) presentes no plasma de *P. flavoviridis*. A função intrínseca das SSP ainda não foi descrita, mas, ao interagirem com HSF, as SSP aumentam o espectro de inibição de HSF – a interação do inibidor com SSP-1 permitiu a inibição da atividade catalítica de HV1, uma SVMP de classe PIII isolada do veneno da mesma serpente. Quando HSF ou SSP-1 são incubadas isoladamente com HV1, a atividade catalítica da metaloendopeptidase não é inibida (Shioi *et al.*, 2013).

O mecanismo de inibição de metaloendopeptidases por HSF ainda não está completamente elucidado, mas já há na literatura algumas evidências experimentais que podem contribuir nesta direção. Deshimaru e colaboradores alteraram quimicamente os resíduos de triptofano e lisina de HSF com ácido trinitrobenzenosulfônico e N-bromosuccinimida, respectivamente, e a atividade inibitória de HSF foi eliminada, destacando a importância destes resíduos para o mecanismo de inibição. Modificações químicas nos resíduos de arginina, tirosina e histidina não tiveram efeito sobre a atividade do inibidor (Deshimaru *et al.*, 2005). Posteriormente, Aoki e colaboradores demonstraram, por sucessivas hidrólises químicas, que a primeira metade do primeiro domínio cistatina de HSF é fundamental para a inibição de metaloendopeptidases. Quando estes resultados foram incorporados a um modelo tridimensional preliminar de HSF, os dados sugeriram um possível *cluster* de resíduos (W17, W48, K15 e K41) localizado no primeiro domínio cistatina que poderia estar envolvido na atividade inibitória da molécula (Aoki, Deshimaru, *et al.*, 2007).

O inibidor BJ46a, alvo de estudo desta tese, foi o segundo inibidor da superfamília das cistatinas a ser caracterizado detalhadamente (Valente *et al.*, 2001). BJ46a é uma glicoproteína ácida (pI 4,55) que, assim como HSF, apresenta dois domínios cistatina (C1/C2) e um domínio rico em histidina em sua porção C-terminal. BJ46a é sintetizada no fígado da serpente com um peptídeo sinal de 19 resíduos; este é processado, gerando a estrutura primária madura de BJ46a, com 322 resíduos, incluindo 12 resíduos de cisteína em 6 pontes dissulfeto preditas. A forma madura de BJ46a é então exportada para a corrente sanguínea, protegendo a serpente da atividade proteolítica de metaloendopeptidases eventualmente circulantes no sangue. Por espectrometria de

massas (MALDI-TOF), BJ46a apresentou uma massa molecular de 46.101 Da e, por cromatografia de exclusão molecular e espalhamento dinâmico de luz (DLS, *dynamic laser light scattering*), a massa molecular calculada de BJ46a foi de 79 kDa, indicando uma possível estrutura dimérica do inibidor (Valente *et al.*, 2001). Com relação à porção glicídica de BJ46a, há quatro sítios confirmados de N-glicosilação (N76, N185, N263 e N274) ocupados por glicanos complexos e sialilados; os glicanos respondem por cerca de 17% da massa do inibidor (Valente *et al.*, 2001; Bastos, 2014).

O mecanismo de inibição de metaloendopeptidases de venenos de serpentes por BJ46a envolve a formação de um complexo estável, mantido por forças de natureza não-covalente, entre o inibidor e a toxina. Medições por ressonância plasmônica de superfície (SPR) indicaram que a interação entre BJ46a e jararagina, uma metaloendopeptidase de classe PIII do veneno de *B. jararaca*, apresenta uma constante de dissociação no equilíbrio ( $K_D$ ) de 13,8 nM<sup>1</sup>. Este  $K_D$  é comparável às constantes mensuradas para interações antígeno-anticorpo, indicando que se trata de uma interação não-covalente forte (Drake *et al.*, 2004). Uma vez que o complexo é estabelecido, a toxina não é capaz de exercer sua atividade catalítica. BJ46a foi capaz de inibir a atividade catalítica de jararagina e de atrolisina-C (PI-SVMP isolada do veneno de *Crotalus atrox*) sobre o substrato fluorogênico Abz-Ala-Gly-Leu-Ala-Nbz, em uma relação molar de 1:2 (monômero do inibidor: toxina). BJ46a também inibiu a atividade hemorrágica do veneno de *B. jararaca in vivo* (Valente *et al.*, 2001). Entretanto, BJ46a não foi capaz de interagir com jararagina-C, uma forma processada de jararagina contendo apenas os domínios desintegrina-símile e rico em cisteína, mesmo em um excesso molar do inibidor de três vezes, o que sugere que o domínio metaloendopeptidase é essencial para a interação toxina-inibidor (Valente *et al.*, 2001).

Porém, assim como para HSF, o conhecimento acerca do mecanismo de inibição de BJ46a é muito limitado. BJ46a, mesmo após desglicosilação extensiva com diferentes exoglicosidases (neuraminidase,  $\beta$  (1→4) galactosidase,  $\beta$ -N-acetilglicosaminidase) e a glicoamidase PNGase F, ainda foi capaz de interagir e inibir a atividade proteolítica de jararagina *in vitro*, indicando que sua porção proteica é fundamental para a inibição de SVMP por BJ46a (Bastos, 2014). Um modelo estrutural preliminar de BJ46a foi construído, tendo como molde o modelo publicado de HSF. Os resultados indicaram

---

<sup>1</sup> Valente, RH. Dados não publicados.

que, além do *cluster* de resíduos W17, W48, K15 e K41 encontrados no modelo de HSF, BJ46a apresentaria um *cluster* adicional, formado pelos resíduos W52 e K58 no primeiro domínio cistatina, possivelmente envolvido com a interação e/ou com a ligação de uma segunda molécula de SVMP (Valente *et al.*, 2009). Até o momento, não há na literatura dados experimentais que possam confirmar esta hipótese.

Recentemente, BJ46a foi expressa de maneira recombinante utilizando-se a levedura *Pichia pastoris* (Shi *et al.*, 2012). Esta forma recombinante de BJ46a (rBJ46a) foi capaz de suprimir a invasão de células de melanoma da linhagem B16F10 e de células de carcinoma hepatocelular da linhagem MHCC97H *in vitro*. *In vivo*, rBJ46a inibiu a colonização tecidual em um modelo murino de câncer pulmonar e reduziu a ocorrência de metástase em camundongos BALB/c infectados com células MHCC97H; os autores atribuíram as propriedades antitumorais de rBJ46a à inibição da atividade catalítica de metaloendopeptidases de matriz (MMP) 2 e 9, mas não demonstraram experimentalmente esta hipótese (Ji *et al.*, 2013).

As evidências na literatura suportam potenciais aplicações de BJ46a tanto na terapia antiofídica quanto na supressão de neoplasias, mas os dados escassos acerca do mecanismo de ação e da estrutura tridimensional de BJ46a são um gargalo para o desenvolvimento de biofármacos baseados neste inibidor.

#### 1.4 – MÉTODOS BIOFÍSICOS PARA A DETERMINAÇÃO ESTRUTURAL DE PROTEÍNAS

A determinação da estrutura e dinâmica proteica é crítica para o desenvolvimento de biofármacos, não só nas etapas de descoberta e validação, como também nas etapas de formulação, que visam as estabilidades estrutural e funcional da molécula a longo prazo (Houde e Berkowitz, 2014).

O início do campo da biologia estrutural se entrelaça com o desenvolvimento de uma das técnicas mais aplicadas para a determinação estrutural de proteínas, a cristalografia de raios X. Em 1912, Max von Laue demonstrou os primeiros padrões de difração de raios X de cristais e, em 1913, Wilhelm Lawrence Bragg derivou a equação fundamental da formação do padrão de difração dos raios X, conhecida hoje como Lei de Bragg (Shi, 2014). Mas foi apenas em 1954 que David Green, Vernon Ingram e Max Perutz descreveram que, em princípio, a cristalografia de raios X também poderia ser aplicada para a determinação da estrutura tridimensional de proteínas (Green *et al.*,

1954; Shi, 2014). Quatro anos depois, em 1958, Kendrew e colaboradores publicaram a primeira estrutura cristalográfica de uma proteína, mioglobina, em uma resolução de 6 Å (Kendrew *et al.*, 1958), marcando o início da aplicação da cristalografia de raios X para determinação estrutural de proteínas.

No repositório de estruturas proteicas *Protein Data Bank* (PDB, [www.rcsb.org](http://www.rcsb.org)), a quase totalidade das estruturas depositadas foi elucidada pelos métodos de cristalografia de raios X (90,3% das estruturas resolvidas), espectroscopia de ressonância magnética nuclear (RMN, 8,1%) e microscopia eletrônica 3D (1,3%) (Protein Data Bank, 2018).

A técnica de cristalografia de raios X baseia-se na exposição de um cristal proteico a um feixe de raios X, que gera um padrão de difração característico da estrutura cristalina. Uma vez que os raios X interagem com a nuvem eletrônica em volta dos átomos, a imagem obtida experimentalmente representa um mapa de densidade eletrônica. Algoritmos computacionais específicos auxiliam na interpretação deste mapa e na construção de modelos tridimensionais de proteínas por diferentes abordagens, como o método de substituição molecular (Parker, 2003). A resolução final do modelo é dependente da qualidade do cristal analisado e, portanto, a etapa de cristalização é fundamental para o sucesso da técnica. Cristalizar proteínas é, até hoje, um processo de tentativa e erro, sensível a vários parâmetros, tais como o tipo de precipitante para a cristalização, força iônica, temperatura, concentração e homogeneidade proteica (Ilari e Savino, 2008; Baettig e Berghuis, 2018).

Diferentemente da técnica de cristalografia de raios X, a espectroscopia de ressonância magnética nuclear (RMN) não depende da formação de um cristal proteico e as medições são feitas em solução. Isso permite não só a determinação estrutural, como também fornece informações importantes sobre a dinâmica molecular da proteína em resolução atômica (Srivastava, 2017). A técnica se baseia no fato de alguns núcleos como  $^1\text{H}$ ,  $^{13}\text{C}$  e  $^{15}\text{N}$  possuírem momentos angulares intrínsecos (*spin* nuclear). Sob a influência de um campo magnético, estes núcleos podem adotar estados de energia e orientação diferentes e podem ser induzidos a intercambiar entre estes estados pela aplicação de pulsos de radiofrequência (Atkins e De Paula, 2008).

O ambiente químico de um núcleo em uma molécula determina suas propriedades em RMN, como a diferença de energia entre as orientações do *spin* e a frequência na qual o núcleo absorve energia (frequência de ressonância). Além da influência do campo

magnético aplicado, cada núcleo magnético é afetado por outros núcleos, tanto por ligações químicas quanto por interações espaciais de curta distância, que podem ser exploradas para o assinalamento das linhas de ressonância e derivação das restrições de distância que os separam, informações estas que fundamentam todo o processo de criação do modelo estrutural por RMN (Pellecchia *et al.*, 2002).

A microscopia eletrônica tridimensional, particularmente a crio-eletromicroscopia (crio-EM), vem ampliando seu leque de aplicações na área de biologia estrutural, principalmente depois de recentes avanços na instrumentação e nos *softwares* de processamento de imagens de crio-EM. Nesta técnica, um feixe de elétrons é acelerado no vácuo por toda a coluna do microscópio. Ao atravessarem o espécime, vitrificado em uma fina lâmina, elétrons difratados são então refocados pelas lentes eletromagnéticas do equipamento para a geração da imagem (Milne *et al.*, 2013). Para cada plano espacial, um conjunto de imagens é registrado e *softwares* de processamento de imagem são utilizados para reconstruir a distribuição de densidade. A partir da reconstrução tridimensional do dado, outros algoritmos são empregados para o *fitting* de modelos estruturais no mapa de crio-EM (Roseman, 2000). A crio-eletromicroscopia ainda é aplicada majoritariamente a proteínas e complexos proteicos de massa molecular igual ou superior a 200 kDa, mas já há alguns trabalhos na literatura empregando crio-EM em proteínas com massa molecular abaixo deste limite, como a estrutura da lactato desidrogenase, uma proteína de 145 kDa, em uma resolução de 2,8 Å (Merk *et al.*, 2016).

Hoje, o PDB contabiliza 147.708 estruturas resolvidas por estes métodos; entretanto, é importante ressaltar que há redundância no banco de dados (*e.g.*, uma proteína resolvida por mais de um método experimental). Ainda assim, esta cifra representa apenas 26% do número de entradas encontradas no banco de dados Swiss-Prot, um repositório de sequências proteicas manualmente curado. Entender a estrutura tridimensional de algumas proteínas e seus complexos por estas técnicas ainda é bastante desafiador, seja por limitações metodológicas ou pela própria natureza da interação – a maioria dos complexos proteína-proteína conhecidos ainda não possui um modelo estrutural (Rodrigues e Bonvin, 2014); o estudo estrutural dos inibidores naturais de venenos de serpentes se encaixa neste cenário.

Para BJ46a, as tentativas de cristalização para estudos estruturais por cristalografia de raios-X foram infrutíferas; a heterogeneidade glicídica oriunda dos

quatro sítios de N-glicosilação pode ter impedido a cristalização (Nettlehip, 2012). Mesmo na eventualidade de formação de um cristal com qualidade aceitável e obtenção dos dados de difração, a ausência de proteínas com uma identidade sequencial alta ( $\geq 40\%$ ) com BJ46a nos bancos de dados torna impossível a resolução estrutural pelo método de substituição molecular. Além disso, a heterogeneidade glicídica também pode fazer com que a resolução final da estrutura seja comprometida, devido à interferência destrutiva destes diferentes estados no padrão de difração dos raios X (Bastos *et al.*, 2016). Em RMN, a massa molecular de BJ46a (dímero, ca. 80 kDa) e do complexo formado entre o inibidor e metaloendopeptidases de venenos de serpentes supera o limite de massa molecular para protocolos-padrão da técnica, requerendo marcação isotópica tripla ( $^{15}\text{N}$ ,  $^{13}\text{C}$  e  $^2\text{H}$ ) paralelamente à marcação seletiva de aminoácidos (Bastos *et al.*, 2016). Tais marcações só seriam viabilizadas por um processo de expressão heteróloga do inibidor, com um rendimento proteico satisfatório, o que ainda não foi tentado. Esses requerimentos resultam em um alto custo de produção de amostra e um tempo de análise longo para o assinalamento dos espectros adquiridos.

Nos anos recentes, vem se consolidando na área uma nova abordagem para o estudo estrutural de proteínas, a Biologia Estrutural Integrativa, que consiste na aplicação de métodos híbridos para a construção de modelos tridimensionais de proteínas e complexos proteicos que não são passíveis de estudo pelas técnicas clássicas de biologia estrutural de alta resolução (Ward *et al.*, 2013). A biologia estrutural integrativa caracteriza-se pela utilização adicional de técnicas computacionais e informações estruturais complementares derivadas de diferentes técnicas de biologia molecular, bioquímica e biofísica, tais como SAXS / SANS (espalhamento de raios-X / nêutrons a baixo ângulo), FRET (transferência de energia de ressonância por fluorescência), microscopia eletrônica, ensaios de mutagênese, cálculos computacionais e técnicas baseadas em espectrometria de massas – *cross-linking* químico e troca isotópica hidrogênio-deutério (Alber *et al.*, 2008).

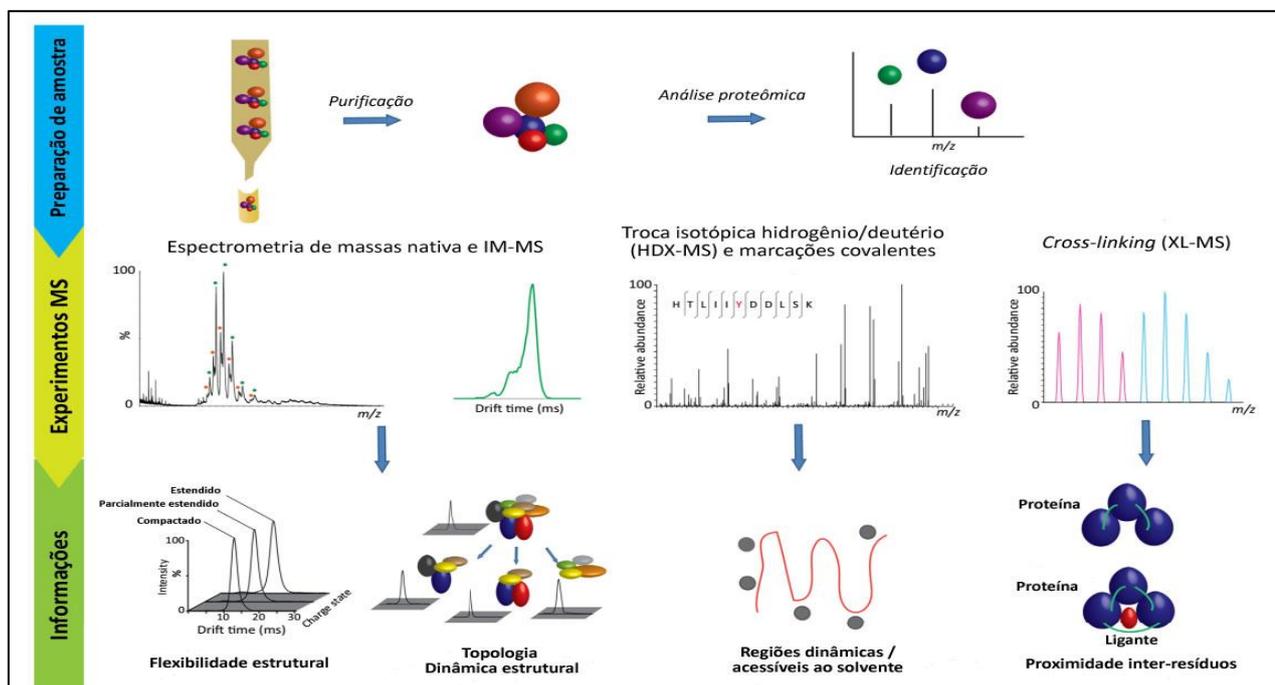
Atualmente, a biologia estrutural integrativa é a melhor estratégia quando os métodos convencionais de determinação estrutural não são capazes de resolver sozinhos os desafios estruturais existentes. Os dados gerados pelas abordagens complementares geram novas hipóteses e aumentam o entendimento da função biológica; para a modelagem molecular, a inserção de dados experimentais complementares minimiza o espaço de busca computacional, a taxa de falsos-positivos

e a dificuldade de analisar a acurácia dos modelos gerados pelos algoritmos de modelagem molecular (Karaca e Bonvin, 2013). Esta abordagem também tem sido cada vez mais utilizada para o mapeamento de grandes complexos macromoleculares, como a subunidade 39S ribossomal, RNA polimerase II e o capsídeo do vírus HIV, entre outras estruturas (Sali *et al.*, 2015).

#### 1.4.1 – A espectrometria de massas na biologia estrutural integrativa

Nos últimos 30 anos, a espectrometria de massas passou por um desenvolvimento notável, tanto em termos de inovações tecnológicas em instrumentação quanto no tocante ao leque de aplicações nas mais diversas áreas do conhecimento. A introdução de técnicas de ionização brandas, como o MALDI (*matrix-assisted laser desorption ionization*, dessorção a laser assistida por matriz) e a ionização por *electrospray* (ESI), no final dos anos 80, por Koichi Tanaka, M. Karas, F. Hillenkamp e John Fenn, permitiram a análise de biomoléculas por espectrometria de massas. Nos anos seguintes, a espectrometria de massas se estabeleceu como uma técnica analítica importante para a determinação da massa molecular de proteínas, sequenciamento, quantificação, identificação de modificações pós-traducionais e, mais recentemente, a análise estrutural de proteínas e peptídeos (Tanaka *et al.*, 1988; Fenn *et al.*, 1989; Karas *et al.*, 1989; Kaltashov *et al.*, 2013).

As técnicas baseadas em espectrometria de massas mais utilizadas para derivar informações estruturais sobre proteínas e complexos proteicos são a espectrometria de massas nativa (*native MS*), espectrometria por mobilidade iônica (*ion mobility MS*, IM-MS) e as abordagens de marcações covalentes (*covalent labeling-MS*), *cross-linking* (XL-MS) e troca isotópica hidrogênio/deutério monitorada por espectrometria de massas (HDX-MS) (Marcoux e Cianferani, 2015). Estas técnicas fornecem informações valiosas sobre estequiometria e topologia de proteínas e seus complexos, forma tridimensional de íons em fase gasosa, acessibilidade ao solvente e regiões de interação proteína-proteína, como demonstra a **Figura 3**.



**Figura 3: Abordagens baseadas em espectrometria de massas aplicadas à biologia estrutural integrativa.**

A espectrometria de massas permite identificar proteínas pertencentes a um complexo de interesse e derivar informações sobre topologia, dinâmica, acessibilidade ao solvente e proximidade inter-resíduos por diferentes experimentos. Adaptado de (Ahdash *et al.*, 2016).

#### 1.4.1.1 – *Cross-linking* de proteínas associado à espectrometria de massas (XL-MS)

Na última década, o *cross-linking* químico de proteínas associado à espectrometria de massas se consolidou como uma estratégia alternativa para auxiliar no estudo estrutural de proteínas e complexos proteicos, fornecendo restrições espaciais importantes entre grupos funcionais proteicos que podem ser exploradas por algoritmos de modelagem molecular (Sharon e Sinz, 2015).

O fundamento da técnica de *cross-linking* consiste na estabilização covalente de proteínas e complexos proteicos por agentes de ligação cruzada (ALC), pequenos compostos orgânicos com grupos funcionais reativos e separados por um braço espaçador de distância definida, medida em ângströms. Assim, apenas os resíduos que estão em proximidade espacial ( $\leq 20 \text{ \AA}$ ) – seja em uma mesma cadeia polipeptídica ou em cadeias polipeptídicas distintas – podem ser conectados covalentemente pelo ALC (Leitner *et al.*, 2010).

Os ALC podem apresentar reatividade frente a grupos amino, sulfidril, carboxila e/ou hidroxila. Agentes de ligação cruzada que apresentam grupos reativos idênticos em cada extremo do braço espaçador são chamados de homobifuncionais, enquanto que

os ALC que possuem grupos funcionais distintos em sua estrutura são denominados heterobifuncionais (Paramelle *et al.*, 2013). Já os ALC trifuncionais têm como objetivo permitir a purificação e o enriquecimento das espécies covalentemente ligadas através da interação específica de um terceiro grupo funcional (Sinz, 2007). Por fim, os ALC chamados de *zero-length* não apresentam braço espaçador e conectam dois grupos funcionais na proteína em curta distância ( $< 3 \text{ \AA}$ ) (Sinz, 2006).

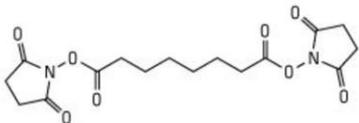
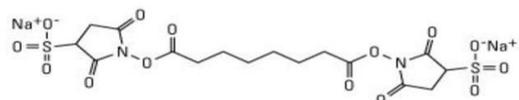
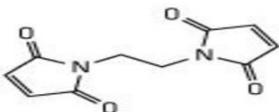
Recentemente, vem se ampliando o uso de ALC cliváveis, que conectam dois grupamentos funcionais em uma proteína por meio de uma ligação lábil, passível de ser clivada durante a etapa de fragmentação dos íons-precusores no espectrômetro de massas. O uso de ALC cliváveis favorece a fragmentação do *backbone* peptídico por dissociação induzida por colisão (CID) e facilita a análise devido à presença de íons-repórter (Sinz, 2017). Outra aplicação que vem se consolidando na área é a execução de estudos de *cross-linking in vivo*, utilizando-se ALC capazes de penetrar a membrana plasmática das células. Esta abordagem permite a análise de interações proteína-proteína transientes que ocorrem no citoplasma celular, ampliando o entendimento da biologia molecular de diversos tecidos, em condições normais ou patológicas (Sinz, 2018).

Atualmente, os ALC mais utilizados na técnica de XL-MS são baseados no grupo funcional NHS (N-hidroxissuccinimida), tais como o DSS (disuccinimidil-suberato) e o BS<sup>3</sup> (bis (sulfosuccinimidil) suberato), disponíveis comercialmente. Os ésteres de NHS reagem com grupamentos nucleófilos, formando ligações amida estáveis com grupamentos  $\epsilon$ -amino das cadeias laterais de resíduos de lisina,  $\alpha$ -amino da extremidade N-terminal da proteína ou, em menor grau, grupamentos hidroxila nas cadeias laterais de serina, treonina e tirosina (Madler *et al.*, 2009; Sharon e Sinz, 2015). Os resíduos de cisteína também podem ser explorados na técnica de XL-MS; após redução de eventuais pontes dissulfeto, o grupamento tiol das cadeias laterais torna-se reativo a maleimidias em pH 6,5-7,5, formando ligações tioéter (Paramelle *et al.*, 2013). Recentemente, Fioramonte e colaboradores desenvolveram uma abordagem inovadora em XL-MS, denominada XPLex, para *cross-links* entre resíduos ácidos (ácido aspártico e ácido glutâmico) espaçados por uma diamina como *linker*. A reação também leva à formação de *cross-links zero-length* após a ativação dos ácidos carboxílicos e reação com grupamentos nucleofílicos da proteína, sem introdução de um *linker* espaçador. A ocorrência de diversas possibilidades de *cross-links* em uma mesma reação aumenta

significativamente a diversidade de restrições de distância que podem ser geradas a partir de experimentos de XL-MS (Fioramonte *et al.*, 2018). A **Tabela 3** resume os ALC mais utilizados na técnica de XL-MS.

**Tabela 3: Agentes de ligação cruzada mais utilizados na técnica de XL-MS.**

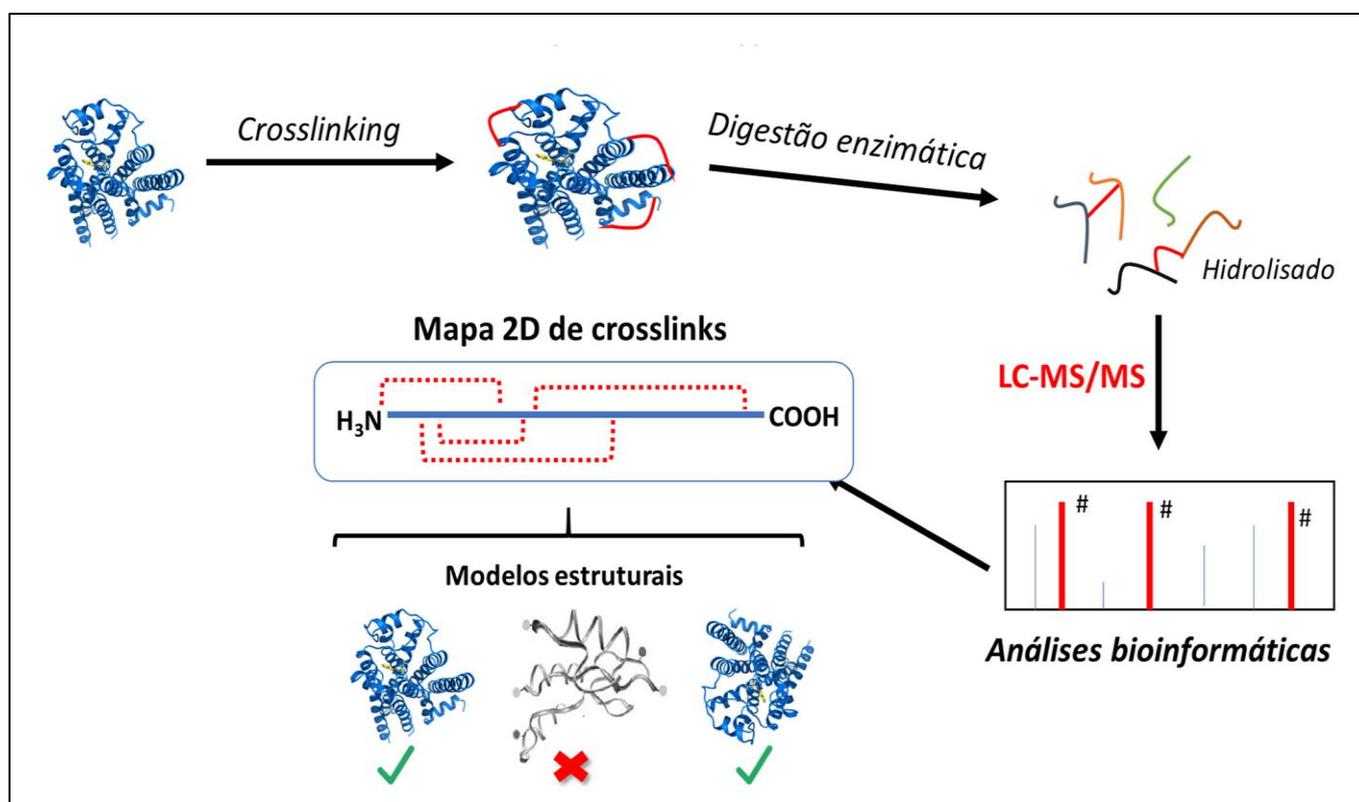
Dados obtidos de (Thermo Scientific; Fioramonte *et al.*, 2018).

| Agente de ligação cruzada                                  | Estrutura química   | Comprimento do braço espaçador | Possibilidades de <i>cross-links</i>   |
|--|---|--------------------------------|--|
| <b>DSS</b><br>(disulfosuccinimidil suberato)               |    | 11,4 Å                         | K – K  |
|  |   |                                | K – S  |
|  |   |                                | S – S  |
|  |   |                                | K – N-terminal<br>S – N-terminal   |
| <b>BS<sup>3</sup></b><br>(bis(sulfosuccinimidil) suberato) |   | 11,4 Å                         | Idênticos aos formados com ALC DSS   |
| <b>Maleimidas</b><br>(e.g. bismaleimidoetano)              |  | 8,0 Å                          | C-C  |
| <b>XPLex – links ácido-ácido</b><br>(diaminas)             |  | 11,5 Å<br>(1,6-hexanodiamina)  | D-D<br>E-E<br>D-E  |
| <b>XPLex – links zero-length</b>                           | <b>sem cadeia espaçadora</b>  | 0                              | D-K<br>D-S<br>E-K<br>E-S<br>D – N-terminal<br>E – N-terminal<br>K – C-terminal<br>S – C-terminal |

Nos experimentos de XL-MS, normalmente emprega-se a estratégia de *bottom-up proteomics*, representada esquematicamente na **Figura 4**. Esta abordagem é particularmente vantajosa, visto que não há limitações com relação à massa molecular da proteína ou do complexo proteico investigado (Sinz, 2007).

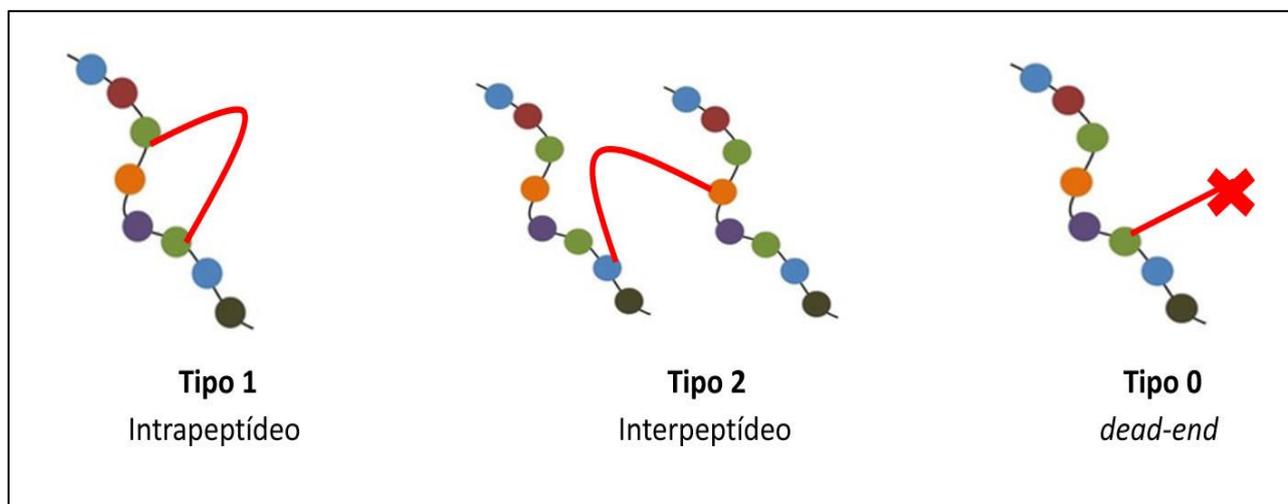
A primeira etapa experimental consiste na estabilização da proteína ou complexo proteico com o ALC, seguido de digestão enzimática. Após a hidrólise da amostra estabilizada, três tipos principais de peptídeos modificados pelo ALC podem ser encontrados na amostra: *cross-links* do tipo 1, tipo 2 e tipo 0, como demonstra a **Figura**

5 (Schilling *et al.*, 2003). Nos *cross-links* do tipo 1, ou *cross-links* intrapeptídeo, os resíduos covalentemente ligados pelo ALC encontram-se no mesmo peptídeo; já nos *cross-links* do tipo 2, o ALC conecta dois resíduos localizados em peptídeos distintos, oriundos da mesma proteína ou de proteínas diferentes. Por fim, os *cross-links* do tipo 0 modificam pontualmente apenas um resíduo de um peptídeo, devido à hidrólise do segundo grupamento reativo do ALC. Estes *links*, conhecidos também como “*dead-ends*”, podem fornecer informações sobre acessibilidade ao solvente de determinados resíduos. Em seguida, as amostras são submetidas à análise por LC-MS/MS para identificação destas espécies modificadas pelo ALC (etapa LC-MS/MS, **Figura 4**).



**Figura 4:** Fluxograma esquemático de uma reação de XL-MS.

Figura baseada em <http://daltonlab.iqm.unicamp.br/research.html>.



**Figura 5: Representação esquemática de modificações encontradas em peptídeos na técnica de XL-MS.**

*Cross-links* do tipo 1 ocorrem em resíduos localizados em um mesmo peptídeo; *cross-links* do tipo 2, acontecem em resíduos pertencentes a peptídeos distintos. Por fim, *cross-links* do tipo 0 modificam apenas um único resíduo de um peptídeo, pois a outra extremidade reativa do ALC sofre hidrólise.

A análise por espectrometria de massas dos produtos da reação de *cross-linking* é desafiadora, principalmente porque os peptídeos covalentemente estabilizados são minoritários em relação aos peptídeos lineares, ou seja, não modificados pelo ALC. Para facilitar a localização dos peptídeos ligados cruzadamente, pode-se usar ALC isotopicamente marcados ou cliváveis; outra estratégia é a utilização de métodos cromatográficos, como a cromatografia de troca catiônica forte (SCX) e a cromatografia de exclusão molecular (SEC), que leva ao enriquecimento de peptídeos ligados cruzadamente (e.g., maiores e com maior estado de carga) nas amostras. Alguns ALC trifuncionais apresentam ainda *tags* de biotina que permitem o enriquecimento dos peptídeos modificados por cromatografia de afinidade por streptoavidina (Singh *et al.*, 2010).

A modificação de resíduos de lisina por alguns ALC, como BS<sup>3</sup>/DSS, aumentam a taxa de clivagens perdidas para enzimas com especificidade de clivagem em lisina, como a tripsina, o que resulta em peptídeos grandes. Portanto, para a análise de amostras de XL-MS, é necessário o uso de espectrômetros de massas capazes de gerar dados de alta resolução tanto para espectros MS1, quanto para MS2, e que sejam capazes de realizar fragmentações eficientes, permitindo o sequenciamento dos produtos da reação de *cross-linking* (Sinz, 2007). Atualmente, o analisador de massas Orbitrap é o estado-da-arte para a aquisição de espectros MS1 e MS2 de alta resolução

e a fragmentação por HCD (*Higher-energy collision dissociation*) permite a fragmentação de íons-precusores com relações massa/carga altas (Rose *et al.*, 2012; Zubarev e Makarov, 2013).

Após a aquisição dos espectros de massas, estes precisam ser analisados com uma ferramenta bioinformática própria para a busca de espectros de *cross-linking* (análises bioinformáticas, **Figura 4**). Assim como na etapa de aquisição dos espectros de massas, a análise destes dados não é trivial devido ao tamanho do espaço de busca a ser explorado por estas ferramentas – à medida em que a complexidade da amostra aumenta (e, conseqüentemente, o número de resíduos passíveis de ligação cruzada) o número de *cross-links* potenciais segue uma expansão quadrática (Sharon e Sinz, 2015). Em 2015, nosso grupo de pesquisa participou do esforço colaborativo, coordenado pelo Dr. Paulo Costa Carvalho, que desenvolveu o *software Spectrum Identification Machine for Crosslinked Peptides* (SIM-XL), com um motor de busca otimizado para a análise de amostras de XL-MS. O SIM-XL pode executar uma redução dinâmica do espaço de busca, considerando apenas combinações que envolvam exclusivamente peptídeos que tenham sido também observados na forma linear, modificados por um *dead-end* e/ou valendo-se de íons-repórteres característicos de determinados ALC e que são gerados durante a fragmentação dos peptídeos ligados cruzadamente (Iglesias *et al.*, 2009; Lima *et al.*, 2015).

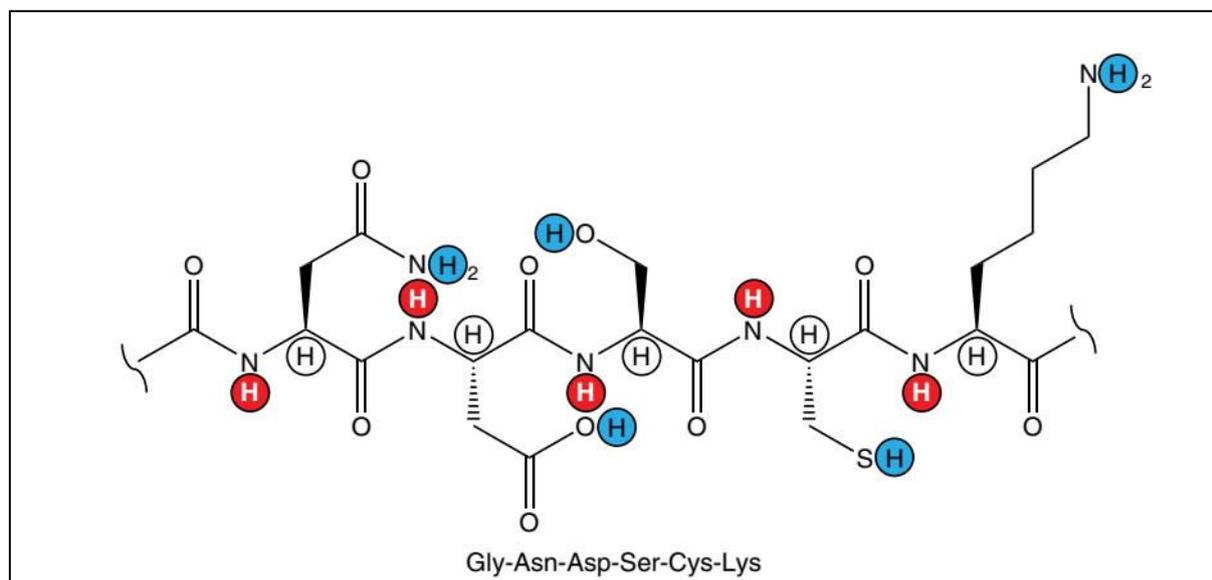
Ao final da análise, um mapa bidimensional contendo as restrições de distância detectadas é gerado para cada proteína (mapa 2D, **Figura 4**). As restrições espaciais definidas pelos *cross-links* detectados podem ser submetidas a algoritmos de modelagem / *docking* molecular *a priori*, de modo a guiar a modelagem e as simulações de *docking*, reduzindo o espaço de busca conformacional. Alternativamente, podem servir como um filtro *a posteriori*, para seleção dos modelos mais convergentes com as restrições espaciais geradas empiricamente (última etapa, **Figura 4**) (Karaca e Bonvin, 2013).

#### 1.4.1.2 – Troca isotópica hidrogênio/deutério monitorada por espectrometria de massas (HDX-MS)

A troca isotópica hidrogênio/deutério monitorada por espectrometria de massas é uma poderosa técnica analítica para a investigação de conformação proteica e pode ser aplicada a estudos de dinâmica proteica, interações proteína-proteína e proteína-pequenos ligantes, mapeamento de epítomos e desenvolvimento racional de fármacos (Oganesyan *et al.*, 2018).

Os átomos de hidrogênio correspondem a 50% de todos os átomos encontrados nas proteínas (Engler *et al.*, 2003). Os resíduos de aminoácidos que possuem hidrogênios ligados a heteroátomos (-OH, -SH, -NH, excetuando-se prolina) trocam continuamente hidrogênios com o solvente devido à presença do heteroátomo mais eletronegativo, como demonstra a **Figura 6**; os átomos de hidrogênio covalentemente ligados a carbono não sofrem este fenômeno (Jensen e Rand, 2016).

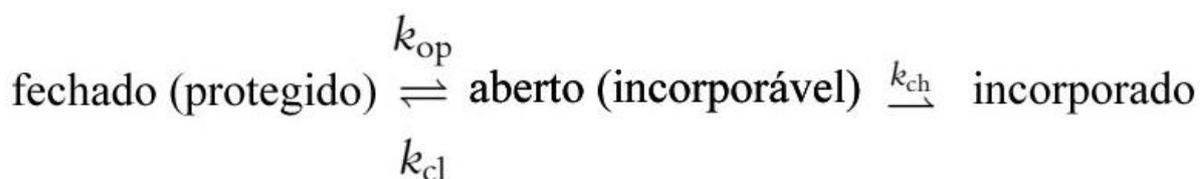
Dessa forma, quando uma proteína é exposta a um solvente que contenha um isótopo de hidrogênio (deutério ou trítio), a incorporação isotópica ocorre nestes hidrogênios lábeis, em diferentes taxas: hidrogênios de cadeia lateral e do grupamento amina da extremidade N-terminal sofrem uma troca extremamente rápida, com uma meia-vida aproximada de 0,01-1 ms, enquanto que os hidrogênios amídicos localizados no *backbone* proteico trocam mais lentamente, com meias-vidas em torno de 5 s a 60 dias, portanto passíveis de serem medidas experimentalmente (Gallagher e Hudgens, 2016).



**Figura 6: Localização dos átomos de hidrogênio em proteínas.**

Os hidrogênios de cadeia lateral (representados em esferas azuis) e os hidrogênios amídicos do *backbone* proteico (representados em vermelho) trocam hidrogênios continuamente com o solvente, enquanto que os hidrogênios covalentemente ligados a carbono (em branco) não trocam. Reproduzido de (Jensen e Rand, 2016).

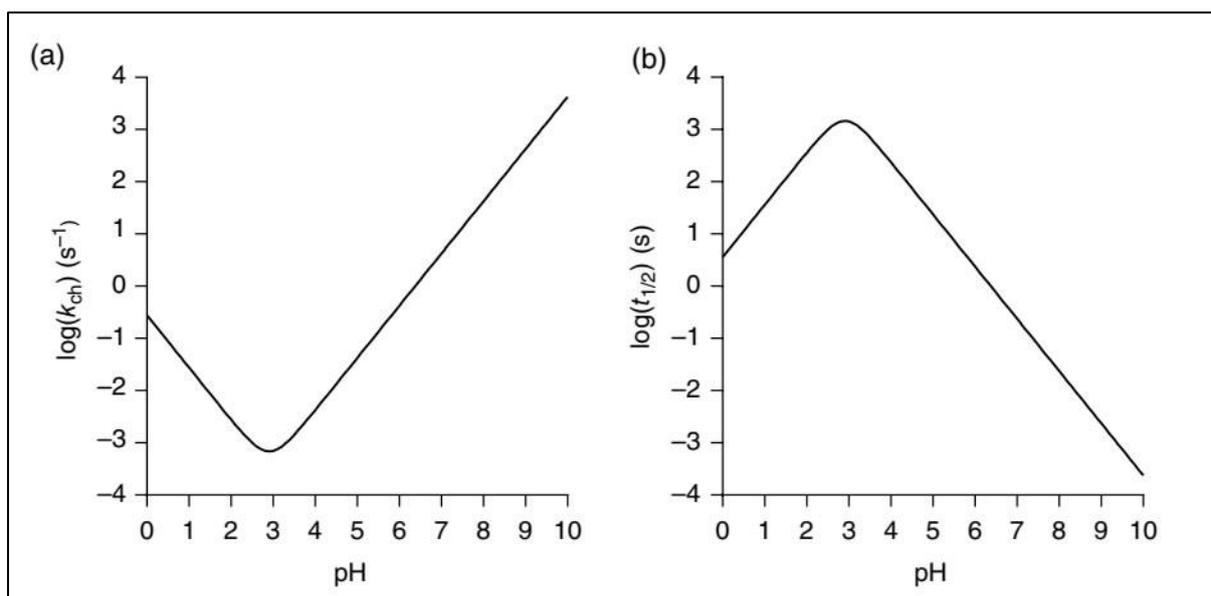
Os fundamentos da teoria da troca hidrogênio-deutério foram desenvolvidos pelo cientista dinamarquês Kaj Ulrik Linderstrøm-Lang e colaboradores no início dos anos 50 (Englander *et al.*, 1997). Na teoria de Linderstrøm-Lang, cada hidrogênio amídico do *backbone* proteico pode estar em dois estados principais: um estado fechado, incompetente para a troca isotópica e um estado aberto, passível de incorporação. Como proteínas não são estáticas, cada hidrogênio amídico pode intercambiar entre esses dois estados por meio de um ou mais eventos de flutuação estrutural local (*breathing motions*) (Mayne, 2016). Este comportamento é então representado pela equação abaixo:



Nesta equação,  $k_{op}$  e  $k_{cl}$  são as constantes de abertura e fechamento estruturais da proteína que expõem e protegem o hidrogênio amídico ao solvente, respectivamente. Já a constante  $k_{ch}$  é referida como a constante química de troca, que rege as conversões

entre N – H e N – D (Konnermann *et al.*, 2011). Por sua vez,  $k_{ch}$  é modulada principalmente pelo pH/pD da solução e a temperatura (Konnermann *et al.*, 2014).

A dependência de  $k_{ch}$  com relação ao pH da solução é fundamental para toda a metodologia de HDX-MS. Para um pequeno peptídeo a 20°C, o gráfico do logaritmo de  $k_{ch}$  em função do pH é ilustrado no painel **a** da **Figura 7**.



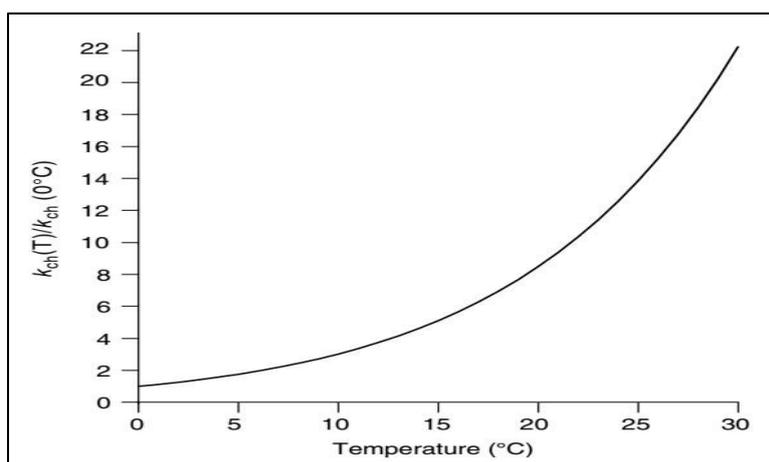
**Figura 7:** A influência do pH na constante química de troca ( $k_{ch}$ , painel a) e na meia-vida da incorporação isotópica de um hidrogênio amídico (painel b) em um peptídeo poli-DL-alanina a 20°C.

Reproduzido de Jensen e Rand, 2016.

A curva de  $k_{ch}$  em função do pH apresenta uma inflexão no intervalo de pH 2,5 – 3,0. As marcações isotópicas em estudos de HDX-MS frequentemente são feitas em soluções de pH > 4, onde a incorporação isotópica é primariamente catalisada por base ( $OD^-$ ) e, a cada unidade de pH acima de 4,  $k_{ch}$  aumenta uma ordem de magnitude (Konnermann, L. *et al.*, 2011). Quando o pH desta solução, mantido em 7 por exemplo, é rapidamente reduzido para  $\approx 2,5$ , a taxa de incorporação de deutério é reduzida em aproximadamente 4 ordens de magnitude e a meia-vida da incorporação isotópica atinge seu ponto máximo, aproximadamente 25 minutos (**Figura 7**, painel **b**), permitindo as medições dos níveis de incorporação de deutério por espectrometria de massas (Jensen e Rand, 2016).

Com relação à temperatura, o gráfico da constante química  $k_{ch}$  em função da temperatura apresenta uma curva exponencial, como demonstra a **Figura 8**. Nesta

curva, a cada 10°C de incremento de temperatura,  $k_{ch}$  aumenta em um fator aproximado de 3 vezes (Mayne, 2016). A curva apresenta um ponto mínimo a 0°C; quando a temperatura de um sistema, mantido a 20°C por exemplo, é resfriada para 0°C, há um decréscimo em  $k_{ch}$  de cerca de 10 vezes (Englander, 2006).



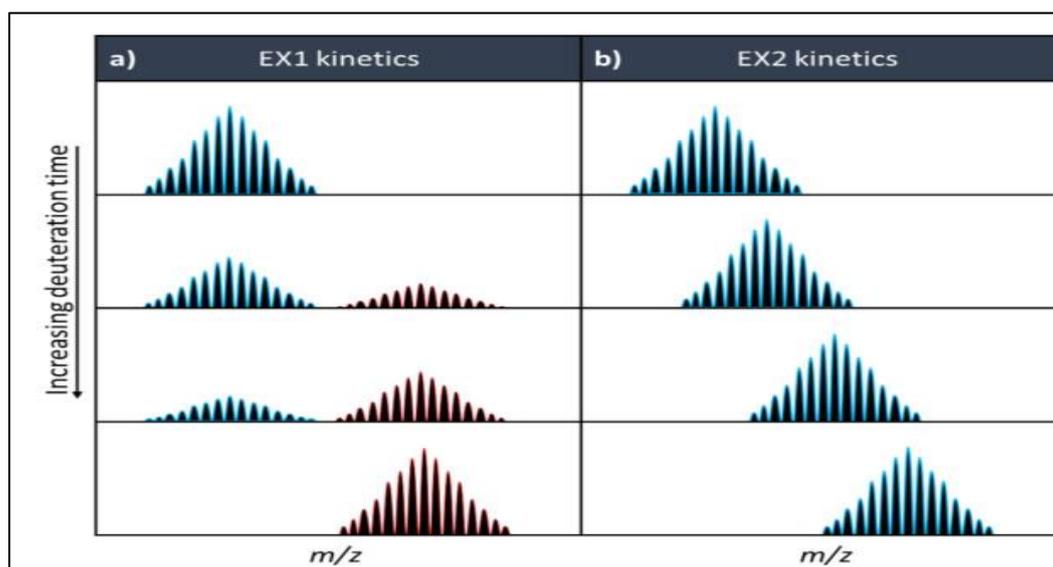
**Figura 8:** A constante química de troca  $k_{ch}$  em função da temperatura.

O valor da constante  $k_{ch}$  foi calculado para o intervalo de 0-30°C e normalizado para o valor de  $k_{ch}$  a 0°C. Reproduzido de Jensen e Rand, 2016.

É importante ressaltar que a troca isotópica completamente dependente de  $k_{ch}$  só é válida na condição onde  $k_{op} \gg k_{cl}$ , onde a proteína permanece em um estado desenovelado. Entretanto, para proteínas em seu estado nativo, os eventos de abertura/fechamento são transientes; regiões enoveladas possuem uma rede de ligações de hidrogênio e a acessibilidade ao solvente de muitos hidrogênios amídicos será limitada, de forma que podemos dizer que a constante observada de troca,  $k_{HDX}$ , será menor do que  $k_{ch}$ , e que  $k_{cl} \gg k_{ch}$  (Konermann, L. *et al.*, 2011; Konermann *et al.*, 2014; Oganessian *et al.*, 2018). Os fatores de proteção (P), definidos como a razão entre  $k_{ch}$  e  $k_{HDX}$  podem exceder  $10^6$  para algumas amidas em regiões enoveladas; porém, mesmo para hidrogênios amídicos extremamente protegidos, a incorporação isotópica ainda pode ser mensurável (Konermann, L. *et al.*, 2011). Logo, podemos distinguir dois regimes distintos de incorporação de deutério em proteínas baseados nas constantes  $k_{op}$ ,  $k_{cl}$  e  $k_{ch}$ , denominados de EX1 e EX2 (Jensen e Rand, 2016).

Em um regime EX1, também chamado de troca correlata, temos a condição onde  $k_{ch} \gg k_{cl}$ , e a constante  $k_{HDX}$  neste caso é definida como  $k_{HDX} = k_{op}$ , o que implica que a incorporação de deutério ocorre no primeiro evento de exposição dos hidrogênios

amídicos, antes que a proteína volte ao estado fechado, incompetente para a troca. O regime EX1 é frequentemente observado em valores altos de pH ou de menor estabilidade estrutural, ou pela adição de desnaturantes (Weis *et al.*, 2006; Mayne, 2016). Em um espectro de massas, a incorporação isotópica via EX1 é facilmente detectada analisando-se o envelope isotópico ao longo do experimento, como demonstra o painel **a** da **Figura 9**. A distribuição do envelope isotópico em um regime EX1 é bimodal, com dois envelopes muito distintos: um primeiro envelope, representado em azul, de menor relação massa/carga ( $m/z$ ), que ainda não sofreu o evento de abertura (e consequente incorporação) e um segundo envelope, vermelho, de maior  $m/z$ , que já incorporou o isótopo. À medida em que se aumenta o tempo de exposição ao deutério, o segundo envelope começa a se tornar prevalente no espectro de massas, indicando a ocorrência do evento de incorporação de deutério.



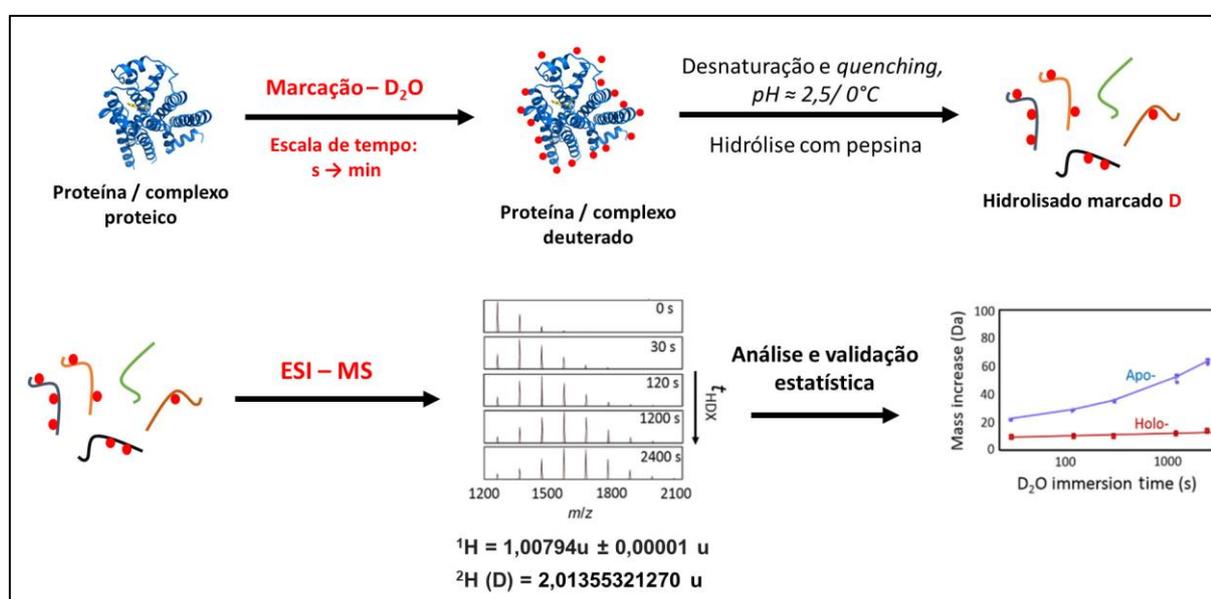
**Figura 9:** Caracterização dos regimes EX1 e EX2 em HDX-MS pela análise dos envelopes isotópicos.

Reproduzido de Oganessian *et al.*, 2018.

No regime EX2, denominado de troca não-correlata, temos uma situação onde  $k_{cl} \gg k_{ch}$ ; nestas condições, a taxa de incorporação é menor do que a taxa de fechamento, o que indica que a probabilidade da incorporação de deutério no primeiro evento de abertura é muito pequena. Isso indica que múltiplos eventos de abertura-fechamento devem acontecer até que ocorra efetivamente a incorporação isotópica. Isso resulta em um acréscimo gradual ( $m/z$ ) à centroide do envelope isotópico, que por sua

vez representa a média dos estados abertos/fechados, como demonstra o painel **b** da **Figura 9** (Gallagher e Hudgens, 2016; Mayne, 2016). Em condições fisiológicas, o regime EX2 é prevalente sobre o regime EX1 (Koneremann, L. *et al.*, 2011).

Para a obtenção dos dados experimentais de HDX-MS, há na literatura um conjunto de abordagens, de acordo com o tipo de marcação (contínua ou pulsada) e conforme a análise por espectrometria de massas (*global e local exchange*) (Wang *et al.*, 2016). A maior parte dos experimentos de HDX-MS adota a estratégia da marcação contínua e a análise por *local exchange*; o fluxograma desta abordagem está representado na **Figura 10**:



**Figura 10:** Fluxograma esquemático de um experimento de HDX-MS.

Nesta estratégia, uma proteína ou complexo proteico de interesse é exposto à água deuterada por intervalos de tempo controlados, iniciando o processo de incorporação isotópica nos hidrogênios amídicos e cadeias laterais. Após cada intervalo, a amostra é rapidamente misturada em uma solução a pH 2,5 e mantida a 0°C. Estas condições desaceleram em várias ordens de magnitude a taxa de incorporação de deutério, como discutido anteriormente neste capítulo, e minimizam a perda da marcação isotópica (*back-exchange*). Entretanto, uma vez que as cadeias laterais apresentam uma taxa de troca extremamente rápida, a marcação de deutério nestas posições é perdida durante as etapas subsequentes do experimento, e apenas a marcação nas amidas do *backbone* proteico é mantida (Gallagher e Hudgens, 2016).

A amostra deuterada então é digerida rapidamente por uma enzima apropriada (usualmente pepsina) em baixas temperaturas, e o hidrolisado é submetido a uma separação cromatográfica a baixas temperaturas (0°C) utilizando gradientes de separação curtos. Os espectros de massas então são registrados para cada amostra, e a incorporação de deutério pode ser verificada comparando-se o espectro de massas de amostras controle (sem deuteração) e amostras deuteradas após diferentes intervalos de tempo, uma vez que cada átomo de deutério incorporado leva a um acréscimo em massa no peptídeo de 1 Da.

Em um estudo de interação proteína-proteína por HDX-MS, curvas de incorporação de deutério são feitas para as proteínas isoladas e em complexo, com o objetivo de mapear as diferenças significativas em incorporação isotópica nos dois estados. Esta informação é bastante útil na delimitação de regiões que permanecem mais ou menos protegidas da troca como consequência da formação do complexo. Muitos complexos proteicos já foram estudados com a técnica, e a integração dos dados de HDX-MS em rotinas de modelagem e *docking* molecular tem sido cada vez mais utilizada (Faini *et al.*, 2016; Lössl *et al.*, 2016).

Nesta tese, empregamos as técnicas de XL-MS e HDX-MS para o estudo estrutural do inibidor natural de metaloendopeptidases BJ46a. tanto em sua forma livre quando em complexo com a metaloendopeptidase jararagina, uma PIII-SVMP isolada do veneno de *Bothrops jararaca*. Juntas, estas abordagens forneceram informações importantes para a modelagem molecular do inibidor e da interação BJ46a-jararagina, contribuindo para a proposição de um mecanismo de inibição de metaloendopeptidases de venenos de serpentes por BJ46a.

## 2 - OBJETIVOS

### 2.1 - Objetivo Geral

Caracterizar estruturalmente a interação entre o inibidor de metaloendopeptidases BJ46a e a toxina jararagina por técnicas aliadas à espectrometria de massas, aplicando-se a abordagem de biologia estrutural integrativa.

### 2.2 - Objetivos Específicos

1. Estabilizar os complexos formados entre BJ46a e jararagina pela técnica de *cross-linking* e mapear as regiões de interação por espectrometria de massas (XL-MS);
2. Analisar a dinâmica conformacional de BJ46a isolada e em complexo com jararagina por troca isotópica hidrogênio-deutério monitorada por espectrometria de massas (HDX-MS);
3. Integrar as informações estruturais geradas pelos experimentos descritos nos objetivos específicos 1 e 2 na modelagem molecular do inibidor BJ46a e da toxina jararagina, além do *docking* entre o inibidor e a metaloendopeptidase.

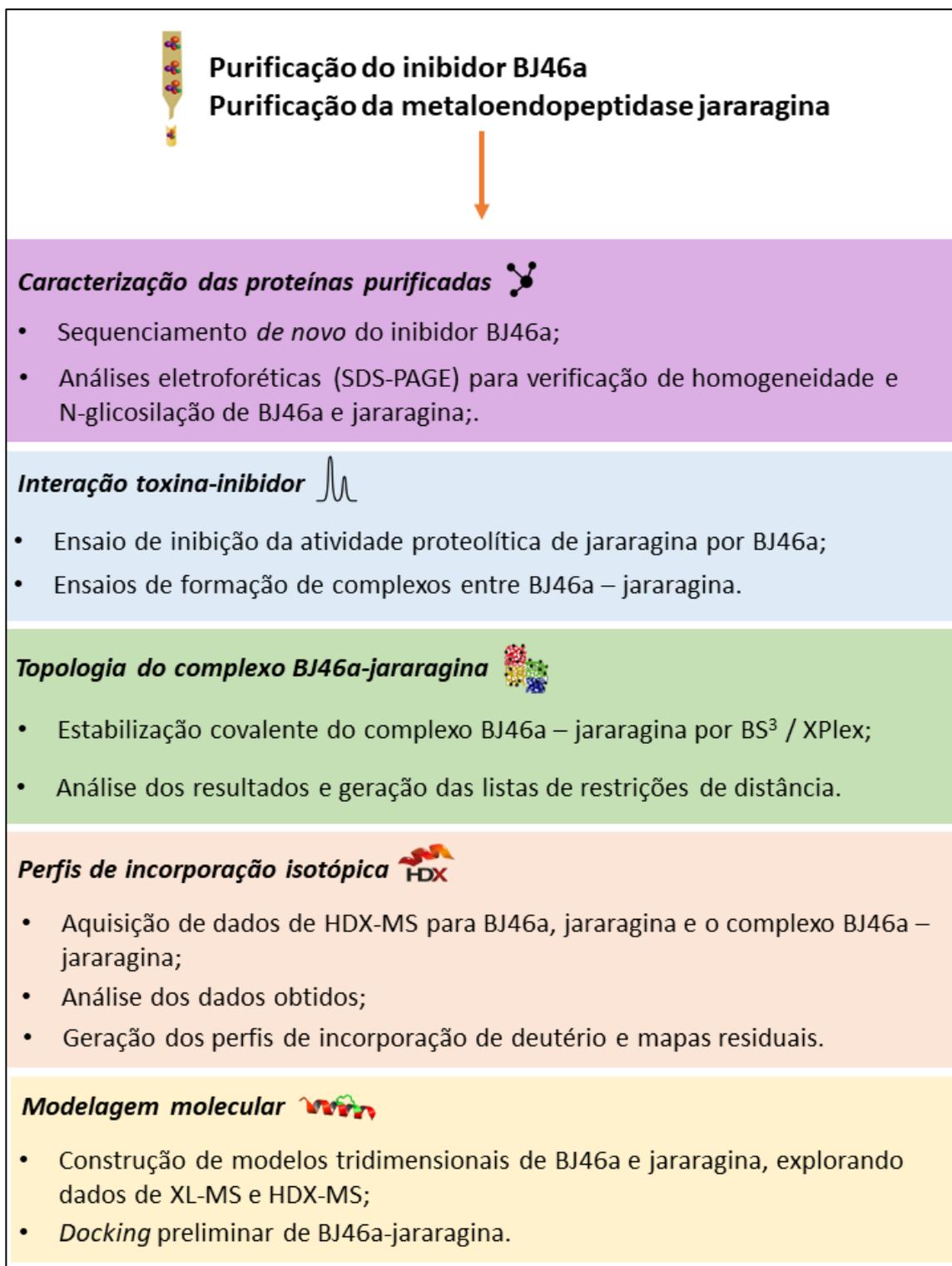
### 3 - METODOLOGIA

#### 3.1- MATERIAL

O plasma e o veneno de *Bothrops jararaca* foi obtido do Laboratório de Herpetologia do Instituto Butantan (SP), através de projeto previamente autorizado pelo Comitê de Ética no Uso de Animais da instituição, de número 1643020615. Em conformidade com a Lei 13.123/2015, a atividade de acesso ao patrimônio genético (veneno e plasma de *B.jararaca*) foi cadastrada no Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado (SisGen), sob o número de cadastro AF0A111.

Óxido de deutério (D<sub>2</sub>O, 99%) foi adquirido da empresa Cambridge Isotope Laboratories; todos os demais reagentes utilizados foram de grau ACS, ou superior. Toda a água utilizada foi purificada e deionizada em sistema Milli-Q (Millipore) e todas as soluções e tampões utilizados nos experimentos foram filtrados em membranas de 0,45 µm.

## FLUXOGRAMA GERAL DAS METODOLOGIAS UTILIZADAS NESTA TESE



## 3.2 – MÉTODOS

### 3.2.1 – Purificação do inibidor de metaloendopeptidases BJ46a

#### 3.2.1.1 – Preparo da coluna de imunoafinidade anti-BJ46a

Para o preparo da coluna de imunoafinidade anti-BJ46a, purificamos anticorpos policlonais anti-BJ46a do soro de coelhos imunizados com uma solução de BJ46a nativa e homogênea (> 95% de pureza por SDS-PAGE). A imunização dos animais e a coleta do soro foram feitas pela empresa Rheabiotech (Campinas, SP). O soro foi então centrifugado a 16.000 xg a 4°C e filtrado em filtros Whatman GD/X de 0,45 µm para remover qualquer material em suspensão e posteriormente estocado a -20°C até o uso. A purificação das imunoglobulinas foi feita utilizando-se uma coluna de afinidade HiTrap Proteína A (1,6 x 2,5 cm, GE Healthcare) a uma taxa de fluxo de 2,5 mL/min. A fração proteica que não interagiu com a fase estacionária (fração não-ligada) foi eluída isocriticamente por cinco volumes de coluna (25 mL), utilizando-se o tampão fosfato de sódio 20 mM pH 7,0. As imunoglobulinas G retidas na coluna (fração ligada) foram eluídas com uma solução de glicina-HCl 0,1 M pH 2,7 também por 5 volumes de coluna. As frações ligadas foram coletadas em tubos previamente refrigerados no gelo e o pH foi neutralizado imediatamente com uma solução de Tris 1 M, na proporção de 40 µL de Tris /mL de amostra coletada. A avaliação da purificação foi feita por eletroforese em condições desnaturantes e redutoras, como descrito no item 3.2.4 deste capítulo. Em seguida, o *pool* de IgG homogênea foi concentrado e o tampão da amostra foi trocado para uma solução-tampão de bicarbonato de sódio 0,2 M e NaCl 0,5 M pH 8,3 utilizando-se sistemas de ultrafiltração Amicon Ultra-15 (*cutoff* 10 kDa, Millipore), em centrífuga previamente refrigerada a 4°C. Após a concentração da amostra, 50 mg de IgG foram então acoplados covalentemente a uma coluna HiTrap NHS HP (1,6 x 2,5 cm, GE Healthcare), de acordo com as instruções do fabricante, com um rendimento final de acoplamento de 91%.

#### 3.2.1.2 - Metodologia de purificação de BJ46a

O plasma de *B.jararaca* foi filtrado em filtros Whatman GD/X de 0,45 µm para remover qualquer material em suspensão e posteriormente centrifugado a 14.000xg por 5 min a 4°C. Em seguida, o plasma foi diluído em tampão Tris-HCl 20 mM pH 7,5 e injetado em uma coluna HiTrap NHS de 5mL (GE Healthcare), com anticorpos policlonais

anti-BJ46a covalentemente imobilizados na matriz cromatográfica. A coluna foi previamente equilibrada com o mesmo tampão. A fração proteica que não interagiu com a fase estacionária (fração não-ligada) foi eluída isocriticamente, por cinco volumes de coluna a uma taxa de fluxo de 5 mL/min. A fração proteica que interagiu com a coluna (fração ligada) foi eluída com uma solução de glicina-HCl 0,1 M pH 2,1, também por cinco volumes de coluna, a 5 mL/min. A absorvância foi monitorada a 280 nm, em sistema cromatográfico Äkta Purifier (Amersham Biosciences). O pH das frações ligadas foi imediatamente neutralizado com uma solução de Tris 1 M, na proporção de 80 µL de Tris por mL de amostra coletada; as amostras foram liofilizadas. Posteriormente, as frações ligadas foram submetidas à cromatografia líquida de alta performance (HPLC) em coluna de fase reversa C4 (10 x 250 mm – Grace Vydac) a 2,5 mL/min, sendo a absorvância monitorada em 215 e 280 nm em sistema cromatográfico LC-10AS (Shimadzu). O sistema de solventes utilizado foi composto de TFA 0,1% (v/v) em água Milli-Q (fase móvel A) e TFA 0,08% (v/v) em acetonitrila (fase móvel B). As condições isocráticas iniciais foram de 5% de solvente B por 5 min, seguido do gradiente: 25% B em 5 min, até 42% B em 10 min, até 47% B em 25 min, até 60% B em 5 min e até 80% B em 10 min. As frações coletadas foram imediatamente neutralizadas com Tris 1 M, na proporção de 20 µL de Tris por mL de amostra e liofilizadas. Em seguida, a fração oriunda da cromatografia de fase reversa correspondente a BJ46a foi dessalinizada em HEPES 10 mM pH 7,5 por cromatografia de exclusão molecular, utilizando a coluna HiTrap Desalting 5 mL (GE Healthcare). Após a dessalinização, a amostra foi aliquoteada, novamente liofilizada e estocada a -20°C.

### **3.2.2 - Purificação da metaloendopeptidase de veneno de serpente (jararagina)**

A purificação da metaloendopeptidase de veneno de serpente jararagina seguiu o descrito na literatura (Paine *et al.*, 1992), com algumas modificações. O veneno de *B. jararaca* (200 mg) foi solubilizado em tampão Tris-HCl 20 mM, CaCl<sub>2</sub> 10 mM, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 1,2 M pH 7,4 e centrifugado a 14.000xg por 5 min a 4°C. O sobrenadante foi submetido à cromatografia de interação hidrofóbica em coluna HiTrap Phenyl HP (16 x 25 mm, GE Healthcare), previamente equilibrada com o mesmo tampão. A separação foi realizada por meio de um gradiente linear decrescente de sulfato de amônio por 5 volumes de coluna, a uma taxa de fluxo de 2 mL/min, sendo a absorvância monitorada a 280 nm em sistema cromatográfico Äkta Purifier (Amersham Biosciences). As frações enriquecidas

em jararagina foram unidas em um único tubo, precipitadas com sulfato de amônio a 80% de saturação e dessalinizadas em coluna PD-10 (Sephadex G-25, GE Healthcare), com o tampão Tris-HCl 20 mM + CaCl<sub>2</sub> 10 mM pH 6,8. Posteriormente, a amostra foi submetida à cromatografia de troca iônica em coluna Mono-Q HR 5/5 (GE Healthcare). A separação cromatográfica foi realizada por meio de um gradiente crescente de cloreto de sódio (0-1 M NaCl), por cinco volumes de coluna a uma taxa de fluxo de 1 mL/min, sendo a absorvância monitorada em 280 nm em sistema cromatográfico Äkta Purifier (GE Healthcare). A homogeneidade da amostra foi analisada por eletroforese em condições desnaturantes e redutoras e a atividade proteolítica foi avaliada por meio do ensaio de hidrólise de azocaseína, como descrito no item 3.2.13 deste capítulo. Após a purificação, a amostra foi alíquotada e estocada a -80°C.

### 3.2.3 - Quantificação de proteínas

A quantificação de proteínas em amostras heterogêneas foi feita pelo método fluorométrico Qubit (Invitrogen), seguindo as instruções do fabricante. A quantificação de proteínas em amostras homogêneas foi feita por espectrofotometria a 280 nm em espectrofotômetro Nanodrop 2000 (Thermo Scientific), usando os seguintes valores de massas molecular e coeficiente de extinção molar:

Jararagina: 50 kDa;  $\epsilon_{280\text{nm}} = 54.385 \text{ M}^{-1} \cdot \text{cm}^{-1}$

BJ46a: 46 kDa;  $\epsilon_{280\text{nm}} = 29.170 \text{ M}^{-1} \cdot \text{cm}^{-1}$

### 3.2.4 - Eletroforese em gel de poliacrilamida

Todas as análises eletroforéticas em gel de poliacrilamida foram conduzidas de acordo com o método de Laemmli (Laemmli, 1970), utilizando sistemas mini-gel (10x10 cm), VERT-i10 (Loccus Biotecnologia). Para a análise em condições nativas, os géis de poliacrilamida e preparo de amostra foram feitos nas mesmas condições descritas, mas omitindo-se o passo de fervura da amostra e os reagentes SDS e DTT dos tampões de gel, de amostra e de corrida.

### 3.2.5 - Coloração de géis por Coomassie R-250 / G-250

A coloração de géis por Coomassie R-250 / G-250 seguiu o descrito por Rabilloud e Charmont (Rabilloud e Charmont, 2000).

### 3.2.6 – Revelação de géis por impregnação por nitrato de prata

O procedimento de revelação de géis de poliacrilamida por impregnação por nitrato de prata seguiu o descrito por (Heukeshoven e Dernick, 1985), com algumas modificações. Após a corrida eletroforética, fixamos os géis com 50 mL de uma solução de etanol a 40% (v/v) e ácido acético a 10% (v/v), por 30 min, sob agitação moderada. Em seguida, esta solução foi descartada e sensibilizamos os géis com 50 mL de uma solução de etanol a 30% (v/v), glutaraldeído 0,5% (v/v), tiosulfato de sódio 0,2% (p/v), acetato de sódio triidratado 12% (p/v), por 30 min. Após este período, esta solução foi descartada e lavamos os géis com água Milli-Q por 3 ciclos de 2 min. Após a etapa da lavagem, os géis foram impregnados com 50 mL de uma solução contendo nitrato de prata a 0,25% (p/v), formaldeído 0,04% (v/v) por 20 minutos, sob agitação moderada. O excesso da solução de nitrato de prata foi removido por 3 ciclos de 20 s de agitação manual com água Milli-Q. Os géis foram revelados com 50mL de uma solução de carbonato de sódio a 2,5% (p/v), formaldeído a 0,02% (v/v), tiosulfato de sódio 0,001% (p/v) e agitados manualmente até que ocorresse a revelação de todas as bandas do gel. A reação de revelação foi interrompida descartando-se a solução reveladora e adicionando-se 50mL de uma solução de EDTA- $\text{Na}_2 \cdot 2\text{H}_2\text{O}$  1,5% (v/v), mantendo-se os géis sob agitação por 10 min. Os géis foram então guardados em uma solução aquosa de ácido acético a 1% (v/v).

### 3.2.7 – *Imunoblotting* do plasma de *B. jararaca* e do inibidor BJ46a

Após a corrida eletroforética, o gel de empilhamento foi removido e o gel de separação foi rinsado em tampão de transferência [tris 25mM, glicina 192mM, metanol 20% (v/v), pH 8,3] para remover o excesso de sais e SDS do gel. A transferência foi realizada em sistema *ECL semi-dry blotter* (Amersham Biosciences), utilizando membranas de nitrocelulose (Hybond ECL, GE Healthcare) em amperagem constante de 80 mA por 2,5 h. A eficiência de transferência foi analisada corando-se a membrana de nitrocelulose com uma solução de vermelho de Ponceau a 0,2% (p/v) em ácido acético 2% (p/v) por 5 min sob agitação, e o excesso de corante foi removido com lavagens sucessivas com água Milli-Q até que ocorresse a revelação das bandas. Após a visualização, a membrana foi descorada completamente por lavagens sequenciais com tampão TBS (tris 25mM, NaCl 140mM, KCl 2mM pH 8,0) acrescido de 0,05% (v/v) Tween-20. Em seguida, os sítios inespecíficos da membrana foram bloqueados

*overnight* com uma solução de bloqueio [leite em pó desnatado 5% (p/v) em TBS+Tween-20] sob agitação moderada a 4°C. Após o bloqueio, a membrana foi lavada por 3 ciclos de 10 min sob agitação com TBS+Tween-20 para remover todo o excesso de solução de bloqueio. Posteriormente, a membrana foi incubada com soro de coelho contendo anticorpos policlonais anti-BJ46a (anticorpo primário), diluído na proporção de 1:1000 em TBS+Tween-20, por 2h, sob agitação moderada, à temperatura ambiente. Após esta incubação, a membrana foi lavada por 3 ciclos de 10 min com TBS + Tween-20 em agitação moderada, para remover o excesso de solução de anticorpo primário. Em seguida, a membrana foi incubada com uma solução de anticorpos secundários (anticorpos anti-IgG de coelhos conjugados à peroxidase, Sigma-Aldrich) diluído 1:5000 em TBS+Tween-20 por 2h sob agitação moderada a temperatura ambiente. Após a incubação com anticorpo secundário, a membrana foi lavada com TBS+Tween por 3 ciclos de 10 min sob agitação moderada, para remover o excesso de anticorpo secundário. Por fim, a membrana foi revelada utilizando-se 10 mL de uma solução de 3,3-diaminobenzidina (DAB) em peróxido de hidrogênio (pastilhas Sigmafast - Sigma-Aldrich), com agitação manual, até que ocorresse a revelação das bandas. A reação de revelação foi interrompida por lavagens sucessivas com água Milli-Q. A membrana foi colocada para secar em papel filtro por 24 h ao final do experimento, antes da digitalização das imagens.

### **3.2.8 - Aquisição de imagens de géis de poliacrilamida e membranas de nitrocelulose**

A aquisição das imagens de géis de poliacrilamida foi feita com o auxílio do programa Adobe Photoshop Elements (Adobe Systems Incorporated) conectado a um *scanner* ImageScanner III (GE Healthcare) no modo transmissivo. As digitalizações foram feitas em resolução de 300 dpi (*dots per inch*, pontos por polegada) e as imagens foram salvas em formato .tiff para minimizar a perda de dados devido à compressão das imagens. Para as membranas de nitrocelulose, o *scanner* foi configurado no modo refletivo, com a mesma resolução e formato de imagem para os géis de poliacrilamida.

### **3.2.9 – Digestão de proteínas em solução para análise por espectrometria de massas**

A amostra a ser digerida (até 50 µg de proteína) foi solubilizada em 20 µL de NH<sub>4</sub>HCO<sub>3</sub> 0,4 M e Ureia 8 M. Adicionaram-se 5 µL de DTT 50 mM e a amostra foi

incubada por 3h a 37°C, para redução das pontes dissulfeto. Após este período, as amostras foram resfriadas à temperatura ambiente e adicionou-se 5 µL de iodoacetamida 150 mM para alquilação das cisteínas, seguido de incubação por 15 min à temperatura ambiente, protegido da luz. Em seguida, adicionaram-se 3 µL de DTT 50 mM para minimizar o excesso de iodoacetamida na amostra, seguido de nova incubação por 15 min à temperatura ambiente, protegido da luz. Posteriormente, adicionaram-se 127 µL de água Milli-Q, para diluir a concentração da ureia na amostra para 1 M. Após a diluição, adicionou-se tripsina (*Sequencing-grade modified porcine trypsin*, Promega) na proporção enzima:substrato de 1:50 (m/m). A amostra foi incubada a 37°C por 20 h e a reação de hidrólise foi interrompida pela adição de 20 µL de ácido trifluoroacético a 10% (v/v, qsp 1%). Os peptídeos resultantes foram dessalinizados e concentrados utilizando microcolunas de fase reversa POROS R2 (Applied Biosystems) e estocados a – 20°C até a análise por espectrometria de massas.

### **3.2.10 – Sequenciamento *de novo* de BJ46a por espectrometria de massas**

Para o sequenciamento de BJ46a, 40 µg do inibidor foram solubilizados gentilmente com uma solução do surfactante RapiGest SF (Waters) a 0,5% (v/v) em bicarbonato de amônio 40 mM pH 8,0, para uma concentração final de surfactante de 0,063% (v/v) em um volume de 50 µL. A amostra foi incubada por 10 min à temperatura ambiente e então 12,5 µL uma solução de DTT 50 mM foram adicionados à amostra para redução das pontes dissulfeto, seguido de incubação por 3 h a 37°C. Em seguida, as cisteínas foram alquiladas adicionando-se 12,5 µL de iodoacetamida 150 mM por 15 min à temperatura ambiente, protegendo as amostras da luz, e o excesso de iodoacetamida foi minimizado adicionando-se 7,5 µL de DTT 50 mM, seguido de nova incubação por 15 minutos à temperatura ambiente, protegida da luz. A partir desta etapa, a amostra foi dividida igualmente em dois tubos: um tubo foi submetido à desglicosilação enzimática pela glicoamidase PNGase F (0,5 U/µg de glicoproteína, Sigma-Aldrich) e incubada por 2 h a 37°C, enquanto a amostra contida no segundo tubo não foi submetida à desglicosilação (controle experimental). Após este período, os tubos foram resfriados à temperatura ambiente e aos dois tubos adicionamos tripsina (*Sequencing-grade modified porcine trypsin*, Promega) na proporção enzima:substrato de 1:50 (m/m), incubando-se por 20 h a 37°C. Em seguida, as amostras foram digeridas com endopeptidase Lys-C (Wako) nas mesmas condições. A digestão proteolítica foi

interrompida pela adição de 7,5  $\mu\text{L}$  de uma solução de ácido trifluoroacético 10% (v/v, qsp 1% - a adição de TFA também promove a hidrólise do surfactante). As amostras foram agitadas rapidamente no vórtex e centrifugadas a 14.000xg por 10 min à temperatura ambiente. O sobrenadante de cada amostra foi então transferido para um novo tubo. Os peptídeos resultantes foram concentrados e dessalinizados em microcolunas de fase reversa POROS R2 (Applied Biosystems). O eluído da coluna [em 70% (v/v) de acetonitrila e 0,1% (v/v) de TFA] foi seco completamente em centrífuga a vácuo (Speed Vac, Savant) e ressuspenso em uma solução de ácido fórmico 1% (v/v) para uma concentração estimada de peptídeos de 0,5 mg/mL (A280, considerando 1 AU = 1 mg/mL), seguido de agitação vigorosa em vórtex e sonicação por 10 min. A análise por LC-MS/MS foi feita no espectrômetro de massas Q-Exactive Plus (Thermo Scientific). A nanocromatografia em fase reversa foi feita em coluna capilar de 30 cm de comprimento (75  $\mu\text{m}$  de diâmetro interno) empacotada com resina Magic C18 (3  $\mu\text{m}$  / 200  $\text{\AA}$  – Michrom Bioresources) utilizando o cromatógrafo Ultimate 3000 (Dionex). A separação cromatográfica foi conduzida a 200 nL/min, utilizando sistema de solventes composto de ácido fórmico a 0,1% (v/v) em água Milli-Q (fase móvel A) e ácido fórmico a 0,1% (v/v) em acetonitrila (fase móvel B) em um gradiente linear de 0-60% B em 162 min. Os espectros MS1 e MS2 foram adquiridos no analisador Orbitrap, aceitando-se íons precursores de todas as cargas ou íons precursores com  $z \geq 2$ . Até 12 íons mais intensos por espectro MS1 foram submetidos à fragmentação por HCD (*higher-energy collision dissociation*). Nas análises onde todos os estados de carga dos íons precursores foram aceitos, configuramos uma fragmentação por HCD com uma energia de fragmentação (sNCE) de 25-35-45%, enquanto que nas análises onde apenas íons precursores com  $z \geq 2$  foram selecionados, utilizamos uma NCE de 30%. A análise dos dados obtidos foi feita com o auxílio dos softwares PEAKS Studio versão 8,5 (Bioinformatics Solutions Inc.) e PepExplorer (Leprevost *et al.*, 2014), este último um algoritmo integrante do *software* PatternLab for Proteomics, v.4.1 (disponível em <http://patternlabforproteomics.org/>). O sequenciamento *de novo* dos peptídeos foi realizado pelo PEAKS; admitimos um erro máximo dos íons-precursos de 10 ppm e 0,02 Da para os íons-fragmento; consideramos a carbamidometilação de cisteína como uma modificação fixa e até 3 modificações por peptídeo sequenciado. Após o sequenciamento *de novo*, os peptídeos sequenciados que apresentaram uma confiança  $\geq 90\%$  (ALC, *average local confidence*) foram exportados em uma tabela, que foi

incorporada ao algoritmo PepExplorer para alinhamento dos peptídeos sequenciados com um banco de dados contendo a sequência de BJ46a e de outros inibidores membros da superfamília das cistatinas, além de conter entradas para *small serum proteins* (SSP) e sequências correspondentes a contaminantes comuns em análises por LC-MS/MS (cRAP: *common repository of adventitious proteins*, disponível em <https://www.thegpm.org/crap/>). Aceitamos uma identidade mínima de 60% para o alinhamento e um tamanho mínimo do peptídeo de 6 resíduos, além de não distinguirmos os resíduos I/L para o cálculo de identidade.

### **3.2.11 - Experimento de formação de complexos entre o inibidor BJ46a e a metaloendopeptidase jararagina**

A interação entre o inibidor BJ46a e jararagina foi analisada por cromatografia de exclusão molecular, de acordo a descrição prévia (Valente *et al.*, 2001), com algumas modificações. Diferentes proporções molares de subunidade de BJ46a para jararagina (0.5:1, 1:1, 2:1), em tampão HEPES 64 mM + CaCl<sub>2</sub> 10 mM + NaCl 150 mM pH 7,5 foram incubadas a 25°C por 15 min, sob agitação constante a 500 rpm (Thermomixer C, Eppendorf). Após este período, as amostras foram centrifugadas a 14.000 xg por 5 min à temperatura ambiente e injetadas em uma coluna de exclusão molecular Superdex 200 Increase 5/150 GL (5 x 150 mm, GE Healthcare) a 0,2 mL/min, em tampão HEPES 64 mM + CaCl<sub>2</sub> 10 mM + NaCl 150 mM pH 7,5 em sistema cromatográfico Ettan LC (Amersham Biosciences), sendo a absorvância monitorada a 220 e 280 nm. As áreas dos picos resultantes de cada análise foram integradas utilizando-se o *software* UNICORN (GE Healthcare). Todo o experimento foi realizado em duplicata.

### **3.2.12 – Ensaio de inibição da atividade proteolítica de jararagina por BJ46a**

O ensaio de hidrólise de azocaseína por jararagina seguiu o descrito em (Cominetti, 2007). Todos os ensaios foram realizados em duplicata e expressos como a média de absorvância de cada amostra, subtraídas da média de absorvância registrada para a amostra branco.

### **3.2.13 – Determinação experimental das pontes dissulfeto de BJ46a por espectrometria de massas**

A determinação das pontes dissulfeto de BJ46a seguiu o descrito por (Lu *et al.*, 2015), com algumas modificações. Cinquenta microgramas de amostra foram

ressuspensos em 25  $\mu\text{L}$  de uma solução de Ureia 8 M + 2 mM de N-etilmaleimida em 100mM de Tris-HCl, pH 6,5 e incubados por 2 h a 37°C para desnaturação proteica e bloqueio de eventuais cisteínas livres. Após este período, a amostra foi resfriada à temperatura ambiente e adicionaram-se 175  $\mu\text{L}$  de Tris-HCl 100 mM pH 6,5 para diluir a concentração de ureia para 1 M. A amostra então foi subdividida igualmente em um outro tubo e prosseguiu-se com as digestões enzimáticas. Um tubo foi digerido com tripsina + endopeptidase Lys-C (Wako) na proporção enzima:substrato de 1:50 (m/m) por 20 h a 37°C. O segundo tubo foi digerido com endopeptidase Asp-N (Roche), na mesma proporção enzima:substrato e tempo de incubação. Posteriormente, adicionamos 500 U de PNGase F (New England Biolabs) para desglicosilar os peptídeos gerados por 2 h a 37°C. Após a desglicosilação, a reação foi encerrada pela adição de 13  $\mu\text{L}$  de uma solução de ácido trifluoroacético a 10% (v/v) para uma concentração final de 1% (v/v). Os peptídeos foram então dessalinizados e concentrados, utilizando-se microcolunas de fase reversa POROS R2. O eluído da coluna (em 70% ACN / 0,1% TFA) foi seco completamente no Speed Vac e ressuspensão em ácido fórmico 1% (v/v) para uma concentração de amostra aproximada de 0,5 $\mu\text{g}/\mu\text{L}$ . A amostra foi estocada a -20°C até o momento da análise por espectrometria de massas. O método de análise das amostras seguiu o descrito para amostras estabilizadas com BS<sup>3</sup> (subitem 3.2.15 deste capítulo), com exceção da primeira massa fixa em  $m/z$  200.

A análise dos dados foi conduzida utilizando-se o programa *Spectrum Identification Machine for Crosslinked Peptides* (SIM-XL, disponível em <http://patternlabforproteomics.org/sim-xl>), versão 1.5 (Lima *et al.*, 2015). Aceitamos para as buscas um erro máximo nos íons precursores e íons fragmento de 10 ppm e assumimos a clivagem enzimática como totalmente específica; também foi considerada modificação variável em asparagina (desamidação:  $\Delta m = +0.98\text{Da}$ ). Após filtragem inicial dos espectros pelo “score” primário do SIM-XL, prosseguimos com a análise manual de todas as pontes dissulfeto encontradas.

### **3.2.14 – Cross-linking entre resíduos básicos (BS<sup>3</sup>) o inibidor BJ46a e da metaloendopeptidase jararagina monitorado por espectrometria de massas (XL-MS)**

Inicialmente, amostras do inibidor BJ46a e da toxina jararagina foram dessalinizadas usando colunas *Zeba spin desalting* (0,5 mL, 7K MWCO – Thermo Scientific) em tampão HEPES 64 mM + CaCl<sub>2</sub> 10 mM + NaCl 150 mM pH 7,5. BJ46a (20µg) foi incubada com a toxina jararagina em uma proporção molar de 1:1 por 15 min a 27°C, sob agitação moderada, a 500 rpm (Thermomixer C, Eppendorf). O complexo foi então estabilizado pela adição do crosslinker BS<sup>3</sup> em um excesso em peso de 20 vezes em relação à quantidade de proteína, por 3 h a 27°C, sob agitação a 500 rpm. A reação foi interrompida pela adição de uma solução de bicarbonato de amônio 500 mM, para uma concentração final de 20 mM. O excesso de BS<sup>3</sup> foi removido das amostras por cromatografia de exclusão molecular em uma coluna Superdex 200 HR 10/30 (10 x 300mm, GE Healthcare) a um fluxo de 0,5 mL/min, em tampão HEPES 20 mM pH 7,5. A absorvância foi monitorada a 220 nm em sistema cromatográfico Äkta Purifier (Amersham Biosciences). As frações relativas ao complexo estabilizado foram unidas em um único tubo e secas em uma centrífuga a vácuo (Speed Vac – Savant) e ressuspensas em 60 µL de bicarbonato de amônio 0,4 M e ureia a 8 M e submetidas à digestão trípica, como descrito no item 3.2.10 deste capítulo. Após a hidrólise, as amostras foram concentradas e dessalinizadas por micro-colunas de fase reversa POROS R2 (Applied Biosystems). Os peptídeos eluídos foram secos em centrífuga a vácuo e ressuspensos em uma solução de ácido fórmico 1% (v/v) para uma concentração aproximada de 0,5 µg/µL, medido por absorvância a 280 nm, utilizando um espectrofotômetro Nanodrop 2000 (Thermo Scientific) e estocados a -20°C para posterior análise por espectrometria de massas.

### **3.2.15 – Cross-linking entre resíduos ácidos (XPlex) entre o inibidor BJ46a e a metaloendopeptidase jararagina monitorado por espectrometria de massas (XL-MS)**

A reação de estabilização do complexo BJ46a e jararagina pela abordagem XPlex seguiu a literatura (Fioramonte *et al.*, 2018), com algumas modificações. Inicialmente, submetemos as amostras de BJ46a e jararagina a dessalinização em tampão MOPS 100mM pH 6,5, utilizando colunas *Zeba spin desalting 7K* (0,5 mL – 7K MWCO - Thermo

Scientific), de acordo com as instruções do fabricante. BJ46a (40 µg) foi incubada com jararagina na proporção molar de 1:1 por 15 min a 30°C sob agitação moderada, a 500 rpm (Thermomixer C, Eppendorf). Em seguida, o complexo foi estabilizado covalentemente pela adição de N-hidroxibenzotriazol (HOBt – ativador de ácidos carboxílicos) em excesso molar de 500 vezes em relação à proteína; adicionaram-se 1,6-diaminohexano (*linker*) em excesso molar de 1000 vezes e cloridrato de 1-etil-3-(diaminopropil)carbodiimida (EDC – ativador de ácidos carboxílicos) em excesso molar de 2000 vezes. A reação de estabilização transcorreu por 2 h a 30°C, sob agitação moderada a 500 rpm (Thermomixer C, Eppendorf). Antes do encerramento da reação, uma alíquota de cada amostra foi removida para análise por eletroforese em condições nativas e eletroforese em condições desnaturantes e redutoras. Em seguida, a reação foi interrompida por precipitação utilizando uma mistura de etanol/acetona (4 volumes de etanol + 4 volumes de acetona por volume de amostra) e incubada *overnight* a -20°C. Após centrifugação a 16.000 xg, o sobrenadante foi descartado e o precipitado foi ressuspensão em 20 µL de bicarbonato de amônio 0,4 M e ureia a 8 M e prosseguiu-se à redução das pontes dissulfeto e à alquilação de cisteínas, seguido de digestão trípica, como descrito no item 3.2.10 deste capítulo. Após a digestão trípica, os peptídeos foram concentrados e dessalinizados utilizando micro-colunas de fase reversa POROS R2 (Applied Biosystems). O eluído da coluna foi seco e ressuspensão em solução de ácido fórmico 1% para uma concentração aproximada de 0,5 µg/µL (A280) e as amostras foram então estocadas a -20°C para análise por espectrometria de massas.

### **3.2.16 – Análise por espectrometria de massas de amostras de *cross-linking***

As amostras de *cross-linking* previamente dessalinizadas foram submetidas à nanocromatografia em fase reversa em coluna capilar de 30 cm de comprimento (75µm de diâmetro interno) empacotada com resina Magic C18 (3 µm / 200 Å – Michrom Bioresources) utilizando o cromatógrafo Ultimate 3000 (Dionex). A separação cromatográfica foi conduzida a 200 nL/min, utilizando sistema de solventes composto de ácido fórmico a 0,1% (v/v) em água Milli-Q (fase móvel A) e ácido fórmico a 0,1% (v/v) em acetonitrila (fase móvel B) em um gradiente linear de 0 – 50% B em 162 min, seguido de gradiente até 80% de solvente B em 4 min e eluição isocrática nesta condição por 2 minutos adicionais. Os peptídeos eluídos foram introduzidos diretamente no espectrômetro de massas Q-Exactive Plus (Thermo Scientific) utilizando uma fonte

*nanoelectrospray*. A aquisição de dados foi conduzida no modo positivo, em modo de análise dependente de dados (DDA) em condições otimizadas por nosso laboratório para amostras de *cross-linking*. Os espectros MS1 foram adquiridos no analisador Orbitrap em modo *profile*, com resolução de 70.000 (FWHM,  $m/z$  200) na faixa de  $m/z$  de 200 – 1800. *Automatic gain control* (AGC) foi configurado para  $1 \times 10^6$  e o tempo máximo de injeção (IT) para 100 ms. Os íons-precusores mais intensos (até seis íons por espectro MS1, com  $z \geq 3$ ) foram submetidos à fragmentação por HCD (*Higher-energy collision dissociation*) com uma energia de colisão (*stepped normalized collision energy*, sNCE) de 30-40% para amostras estabilizadas com BS<sup>3</sup> e 30-40% / 28-33% para amostras estabilizadas com a abordagem XPLex. Os espectros MS2 foram adquiridos no modo centroide, em uma resolução de 35.000 FWHM, AGC  $5 \times 10^5$ , IT  $\leq 100$  ms. Ao analisarmos amostras estabilizadas com BS<sup>3</sup>, fixamos a primeira massa do espectro em  $m/z$  200 para analisarmos íons-repórteres comumente detectados quando se utiliza este *cross-linker* (Santos *et al.*, 2011).

### 3.2.17 - Análise bioinformática de dados de XL-MS

A análise dos dados de XL-MS foi conduzida através do software SIM-XL, versões 1.4 e 1.5. Assumimos, para todas as buscas, um erro de massa dos íons precusores e dos íons-fragmento de 10 ppm e um máximo de 3 clivagens perdidas para a enzima tripsina. Além disso, consideramos a carbamidometilação de cisteína como uma modificação fixa. Após a busca dos espectros, realizamos uma filtragem inicial com relação ao *score* e a quantidade mínima de picos assinalados por cadeia. Apenas peptídeos que apresentassem um *score* primário mínimo de 2,0 e um mínimo de 3 picos assinalados por cadeia foram analisados manualmente. Na análise manual, verificamos a qualidade da cobertura de sequência das cadeias  $\alpha$  e  $\beta$ , a relação sinal/ruído e, principalmente, se a posição do *linker* no espectro era inequívoca – isto é, se a cobertura dos íons-fragmento no espectro de massas permitia uma única interpretação. Apenas espectros considerados confiáveis segundo os critérios apresentados acima foram utilizados nas etapas seguintes de modelagem de BJ46a e jararagina.

### **3.2.18 - Troca isotópica hidrogênio-deutério monitorada por espectrometria de massas (HDX-MS)**

#### 3.2.18.1 – Descrição do esquema experimental para HDX-MS

Para esta tese, foi construído, no Laboratório de Toxinologia, um sistema experimental exclusivo para as análises de HDX-MS. Este sistema foi baseado na infraestrutura experimental de HDX-MS do *Protein Research Group*, Universidade do Sul da Dinamarca (Trelle *et al.*, 2012). A montagem do sistema foi feita com a colaboração da empresa Nova Analítica e do Dr. André Ferreira, tecnologista do laboratório, com base no esquema ilustrado na **Figura 11**.

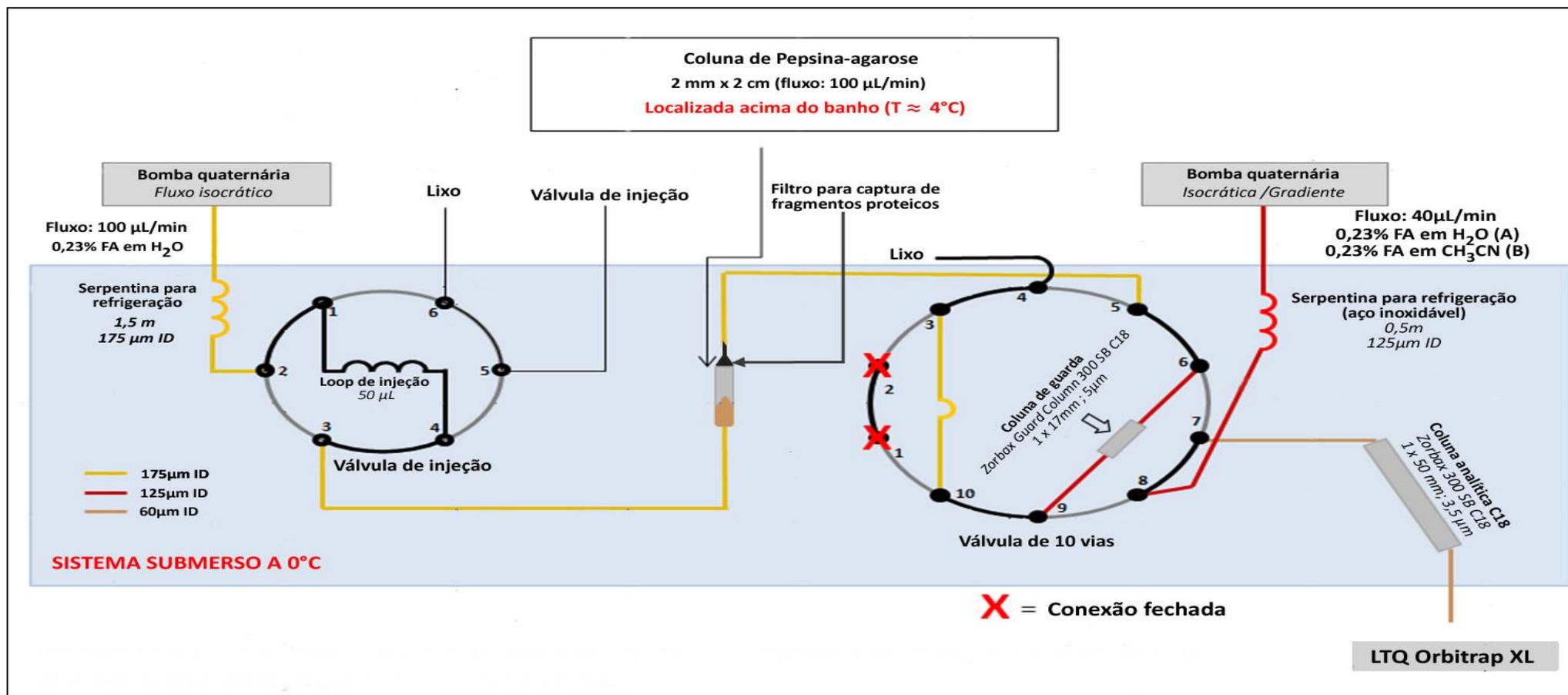


Figura 11: Desenho esquemático do sistema experimental para HDX-MS. Adaptado a partir de Mysling (2013)<sup>2</sup>

<sup>2</sup> Mysling, Simon (2013). – Arquivo pessoal.

### 3.2.18.2 – Curva experimental de incorporação de deutério

Para as curvas de incorporação de deutério, a amostra de jararagina foi concentrada e dessalinizada em tampão HEPES 64 mM + CaCl<sub>2</sub> 10 mM pH 7,5 em sistemas de ultrafiltração Amicon Ultra 0,5 (*cutoff* 10 kDa, Millipore) para uma concentração de proteína equivalente a 80 µM (4 mg/mL). Para BJ46a, alíquotas liofilizadas do inibidor (que já continham HEPES) foram ressuspensas em tampão HEPES 5 mM + CaCl<sub>2</sub> 10 mM pH 7,5, unidas em um único tubo e a concentração de proteína foi ajustada para o equivalente a 80 µM (3,66mg/mL, concentração de HEPES ≈ 88 mM). As soluções-estoque de jararagina e BJ46a foram alíquotadas e mantidas a -80°C.

As curvas de incorporação de deutério foram feitas em triplicata técnica. No início do experimento, procedemos com a formação de complexo entre BJ46a e jararagina na proporção molar de 1:1 por 15 min a 25°C sob agitação constante a 500 rpm (Thermomixer C, Eppendorf). A molaridade final do complexo foi de 40 µM. Uma alíquota dessa amostra foi submetida à cromatografia de exclusão molecular em coluna Superdex 200 Increase 5/150 (conforme descrito no item 3.2.12) para verificamos a saturação do complexo, isto é, se não havia jararagina ou BJ46a em suas formas livres. Para os controles (BJ46a livre e jararagina livre), preparamos uma solução a 40 µM a partir das soluções-estoque de cada proteína e incubamos nas mesmas condições.

Para o ponto zero da curva de incorporação de deutério, retiramos uma alíquota equivalente a 50 pmol de proteína da solução a 40 µM e adicionamos a alíquota a um tubo contendo 19 µL de uma solução de ureia 8M, TCEP 1M, pH 2,4, previamente resfriado a 0°C e a amostra foi congelada imediatamente em N<sub>2</sub> líquido.

A reação de troca hidrogênio → deutério foi iniciada por meio da diluição das amostras em tampão preparado com água deuterada (HEPES 64 mM + CaCl<sub>2</sub> 10 mM, pH 7,5), de modo que a concentração final de D<sub>2</sub>O fosse de 80% (v/v) e a reação foi mantida a 25°C, sob agitação, a 500 rpm (concentração final de proteína de 8 µM, concentração de HEPES ≈ 68mM). Alíquotas equivalentes a 50 pmol de proteína (6,3 µL) foram retiradas nos tempos de exposição de 10s, 100s 1000s e 10.000s. Ao término de cada ponto experimental, a taxa de incorporação de deutério foi desacelerada rapidamente (*quench*), adicionando-se a amostra a um tubo contendo 19µL de uma solução de ureia 8M, TCEP 1 M, pH 2,4 previamente resfriado a 0°C. Em seguida, a

amostra foi imediatamente congelada em N<sub>2</sub> líquido. Todas as amostras foram estocadas a -80°C por 24h.

No momento da análise por espectrometria de massas, a amostra correspondente a cada ponto experimental da curva foi rapidamente descongelada, centrifugando-se o tubo por 1 min a 6000 rpm em temperatura ambiente em uma pequena centrífuga de bancada (Gilson GMCLab) e colocada rapidamente no gelo por 6 min para completar a desnaturação e a redução das pontes dissulfeto. Após este período, a amostra foi diluída 1:1 com uma solução de ácido fórmico 0,23% (v/v) pH 2,4 e imediatamente injetada no sistema de HDX-MS.

Após a injeção, a amostra foi digerida em uma coluna cromatográfica contendo pepsina imobilizada em matriz de agarose (2 mm x 2 cm, Thermo Scientific) a um fluxo de 100 µL/min. Os peptídeos gerados foram acumulados e dessalinizados em uma coluna de guarda C18 (Zorbax Guard Column 300 SB C18, 1x17 mm, 5 µm, Agilent Technologies). Posteriormente, os peptídeos retidos na coluna de guarda foram submetidos a cromatografia de fase reversa em coluna C18 (Zorbax 300 SB C18, 1x50 mm, 3,5 µm, Agilent Technologies). A separação cromatográfica foi conduzida a 40 µL/min, utilizando um sistema de fases móveis composto de ácido fórmico 0,23% (v/v) em água Milli-Q, pH 2,4 (fase móvel A) e ácido fórmico 0,23% em acetonitrila (fase móvel B), com o seguinte gradiente de separação: 31%B em 5 min, 50%B em 8,5 min, 95%B em 10,5 min, 5%B em 12,5 min, 70%B em 16,5 min, 5%B em 17,5 min, 50%B em 19,5 min, 5%B em 21,5 min. Os peptídeos eluídos foram introduzidos no espectrômetro de massas LTQ Orbitrap XL (Thermo Scientific), utilizando uma fonte de ionização *microelectrospray* (µESI). A voltagem e a temperatura do capilar de entrada foram configuradas para 3,5 kV e 100°C, respectivamente, e a intensidade do gás de bainha (*sheath gas*, N<sub>2</sub>), para 10 unidades arbitrárias. Os espectros MS1 foram adquiridos no analisador Orbitrap no modo *profile* com resolução de 60.000 FWHM, na faixa de *m/z* de 300 a 1500.

### **3.2.19 - Análise bioinformática dos dados de HDX-MS**

#### **3.2.19.1 – Construção das listas de peptídeos de BJ46a e jararagina**

Para a análise dos dados de HDX-MS, uma lista de peptídeos identificados de BJ46a e jararagina foi construída a partir de análises de amostras sem deuteração em

LC-MS/MS utilizando o espectrômetro LTQ Orbitrap XL, em condições de HDX-MS. Exploramos diferentes faixas de massas nos espectros MS1 para obtermos o mapa peptídico de BJ46a e jararagina. As faixas de massas foram as seguintes:  $m/z$  300 – 2000,  $m/z$  300 -1500,  $m/z$  300 – 550,  $m/z$  500 – 750,  $m/z$  700 – 950,  $m/z$  900 – 1150,  $m/z$  1100 – 1350,  $m/z$  1300 – 1550. Para a identificação dos peptídeos, utilizamos o *software* PEAKS Studio versão 8.5 (Bioinformatics Solutions Inc.). Para todas as análises, estabelecemos um erro máximo de 15 ppm para os íons precursores e 0,6 Da para os íons-fragmento; admitimos modificações pós-traducionais em ácido glutâmico (piroglutamato: -17,03 Da) e um máximo de três modificações por peptídeo sequenciado. O banco de dados utilizado para as buscas com a ferramenta PEAKS DB foi construído com as sequências de BJ46a (gi|48428681), jararagina (gi|6248), botropasina (gi|82190823), pepsina (gi|118572685) e sequências correspondentes a contaminantes comuns em análises por LC-MS/MS (cRAP: *common repository of adventitious proteins*, disponível em <https://www.thegpm.org/crap/>). Todos os peptídeos identificados de BJ46a e jararagina foram consolidados em uma única tabela contendo a sequência, o tempo de retenção cromatográfica do peptídeo, sua relação massa/carga, o estado de carga e sua massa experimental, para a análise no programa dedicado a análise de dados de HDX-MS, HDEaminer, como descrito no próximo item.

### 3.2.19.2 – Análise dos dados de HDX-MS

A análise dos dados de HDX-MS foi conduzida utilizando o *software* HDEaminer (Sierra Analytics), versões 2.3.0 / 2.4.0. Para iniciarmos as análises, o arquivo contendo a sequência “madura” do inibidor BJ46a ou da toxina jararagina (em formato .fasta) foi carregado para o programa. Admitimos para a sequência de jararagina modificações nos resíduos E2 e E3 (piroglutamato,  $\Delta m = -17,03\text{Da}$ ). Em seguida, carregamos uma lista experimental de peptídeos detectados para cada proteína, construída segundo o descrito no subitem 3.2.19.1 deste capítulo. A partir das listas, o programa derivou um conjunto de peptídeos, eliminando eventuais repetições de um mesmo peptídeo da lista.

Os arquivos brutos (*raw files*) correspondentes a cada ponto experimental da curva de incorporação de deutério foram incorporados ao programa. A busca dos envelopes isotópicos de cada peptídeo no *software* HDEaminer é feita a partir da sobreposição de envelopes isotópicos teóricos (calculados pelo algoritmo para cada

peptídeo da lista) com os envelopes experimentais encontrados; para esta etapa da análise, consideramos uma margem de erro máxima (*m/z shift*) de 20 ppm.

Os peptídeos identificados foram analisados manualmente quanto à qualidade da sobreposição do envelope isotópico teórico com o envelope experimental, relação sinal/ruído e à presença de envelopes isotópicos vizinhos que eventualmente pudessem comprometer a detecção do envelope de interesse. Peptídeos que não apresentaram os critérios de qualidade descritos foram assinalados como de baixa confiança, e, portanto, excluídos dos cálculos de deuteração.

O cálculo de incorporação de deutério para cada peptídeo foi baseado na diferença entre a centroides derivada dos envelopes isotópicos das amostras-controle (sem deuteração) e as centroides dos envelopes relativos a cada ponto experimental de exposição a D<sub>2</sub>O. Nesta etapa, consideramos apenas o melhor estado de carga (*z*) de cada peptídeo (definido pelos *scores* resultantes das sobreposições dos envelopes teóricos e experimentais) e assumimos uma deuteração teórica máxima de 80%, excetuando-se os dois primeiros resíduos N-terminais, uma vez que estes apresentam uma taxa de *back-exchange* elevada e a marcação é frequentemente perdida durante a manipulação da amostra (Kan *et al.*, 2013). Foram retirados resíduos de prolina dos cálculos, uma vez que estes não apresentam hidrogênios passíveis de troca por deutério.

Após o cálculo das curvas de incorporação, o programa realizou um teste t de Student para avaliação de eventuais modificações no perfil de incorporação de deutério nas amostras de BJ46a e jararagina em seus estados livres ou em complexo. As mudanças estatisticamente significativas (*p*-valor < 0,05) foram ressaltadas nas curvas de incorporação de cada peptídeo com um asterisco. Com a análise estatística concluída, geramos um mapa residual para BJ46a e jararagina, indicando as diferenças em incorporação nos dois estados ao longo de toda a sequência.

A partir de todos estes dados, geramos um *heatmap* para BJ46a e jararagina livre nos diferentes tempos de exposição a D<sub>2</sub>O e realizamos uma análise subtrativa para delinear as regiões diferenciais dos perfis de incorporação isotópica de BJ46a e jararagina.

### 3.2.20 – Análises computacionais

Todas as etapas de modelagem molecular (jararagina e BJ46a) bem como as simulações de *docking* proteína-proteína foram feitas em estreita colaboração com o Dr. Francisco Gomes-Neto, pesquisador do Laboratório de Toxinologia, IOC / FIOCRUZ. As análises descritas abaixo foram executadas em ambiente Linux CentOS, em terminal *bash*.

#### 3.2.20.1 – Modelagem molecular da metaloendopeptidase jararagina e do inibidor BJ46a

A sequência correspondente à jararagina em sua forma madura (resíduos 151-571) foi obtida no banco de dados UNIPROT (www.uniprot.org/uniprot), sob o número de acesso P30431 (VM3JA\_BOTJA). A sequência correspondente ao inibidor BJ46a também foi obtida do mesmo banco de dados, sob o número de acesso Q9DGI0 (FTE46\_BOTJA). A região referente ao peptídeo sinal (resíduos 1-19) foi retirada para compor a sequência de aminoácidos da proteína madura. A modelagem foi realizada automaticamente, utilizando a suíte de aplicativos I-TASSER (Roy *et al.*, 2010), disponível em <http://zhanglab.ccmb.med.umich.edu/I-TASSER>.

#### 3.2.20.2 – Conversão dos *cross-links* em restrições de distância

Os *cross-links* identificados experimentalmente foram convertidos em restrições de distância, com base nos resíduos de aminoácidos envolvidos e no agente de ligação cruzada (ALC) utilizado. Na **Tabela 4**, apresentamos duas faixas de valores: na primeira coluna, a distância total entre os átomos de carbono  $\beta$  dos aminoácidos envolvidos, considerando cadeias laterais e espaçadora completamente estendidas; na segunda, a distância euclidiana na qual a probabilidade de encontrarmos a distância topológica entre os dois resíduos de aminoácidos é máxima (Ferrari *et al.*, 2018).

Tabela 4: Distâncias euclidianas consideradas para validação de *cross-links*

| <i>Cross-link</i>                | Dist. Estendida (Å) | L <sub>MAX</sub> (0,99) * (Å) |
|----------------------------------|---------------------|-------------------------------|
| <b>BS<sup>3</sup></b>            |                     |                               |
| K-K                              | 21,8                | 17,8                          |
| K-S                              | 18,0                | 15,8                          |
| S-S                              | 14,1                | 13,4                          |
| <b>1,6-hexanodiamina (XPlex)</b> |                     |                               |
| D-D                              | 14,1                | 13,5                          |
| D-E                              | 15,4                | 14,3                          |
| E-E                              | 16,7                | 15,1                          |
| <b>Zero-length (XPlex)</b>       |                     |                               |
| K-E                              | -                   | 10,5                          |
| K-D                              | -                   | 9,7                           |
| S-E                              | -                   | 7,7                           |
| S-D                              | -                   | 7,0                           |
|                                  |                     | Cβ-Cβ                         |

\* - derivada estatisticamente (Ferrari *et al.* 2018).

### 3.2.20.3 – Análise de distâncias euclidianas e topológicas

As distâncias euclidianas e topológicas (distância acessível ao solvente) foram calculadas pelo *software* Topolink (Martínez *et al.*, 2017). A partir da descrição do experimento realizado, e tendo como base os *cross-links* que podem ser formados e suas respectivas faixas de distâncias e resultados experimentais encontrados, o programa analisa a concordância do modelo com relação ao dado experimental. O algoritmo primeiramente avalia se o par de resíduos e os átomos envolvidos estão expostos na superfície da proteína. A seguir, avalia se a distância euclidiana entre os resíduos / átomos relacionados está dentro dos limites inferior e superior permitidos, e reporta o resultado (**tabela 5**) encontrado. Utilizamos o algoritmo considerando somente os *links* observados experimentalmente, com as estratégias de *cross-links* configuradas de forma padrão e faixas de distância configuradas de acordo com a **Tabela 4** (distâncias para as formas estendidas).

Tabela 5: Códigos de resultado do *software* Topolink, utilizando o algoritmo considerando apenas os *links* observados experimentalmente.

| CÓDIGO              | DESCRIÇÃO  |
|---------------------|--|
| <b>OK:FOUND</b>     | Uma distância topológica foi encontrada e é consistente com todas as observações.  |
| <b>BAD:LONG</b>     | Uma distância topológica foi encontrada e o link foi observado. Entretanto, a distância é maior do que o limite aceitável.           |
| <b>BAD:EUCL</b>     | O <i>link</i> foi observado, mas a distância euclidiana no modelo é maior do que o limite aceitável.                                 |
| <b>BAD:NOTFOUND</b> | O <i>link</i> foi observado, a distância euclidiana está dentro dos limites, mas uma distância topológica válida não foi encontrada. |

#### 3.2.20.4 - Simulações de *docking* molecular

Os modelos obtidos na etapa de modelagem molecular foram utilizados nas simulações de *docking* molecular. A construção dos modelos, a simulação de *docking* molecular e a análise dos resultados foram realizadas utilizando a suíte de aplicativos de modelagem molecular Rosetta, versão 2017.08.59291 (Gray *et al.*, 2003; Wang *et al.*, 2005). Os modelos foram submetidos a uma etapa de relaxamento, aplicando-se restrições de posição para átomos pesados do *backbone* (N, C $\alpha$ , C, O), visando corrigir contatos desfavoráveis entre átomos, problemas geométricos e principalmente energias dos ângulos de torção das cadeias laterais na superfície das proteínas. Utilizamos o protocolo de *docking* molecular local adaptado de Ruedi Aebersold e colaboradores (Kahraman *et al.*, 2013), com parâmetros resumidos na tabela abaixo:

Tabela 6: Parâmetros utilizados nas simulações de *docking* molecular

| PARÂMETRO  | DESCRIÇÃO  |
|--|--|
| <b>-IN:FILE:S COORDENADAS.PDB</b>                | Arquivo de coordenadas dos domínios  |
| <b>-NSTRUCT 100000</b>                           | Número de modelos / docking  |
| <b>-PARTNERS A_B</b>                             | Move D em relação a E  |
| <b>-DOCK_PERT 3 8</b>                            | Perturbação translação e rotação   |
| <b>-EX1</b>                                      | Adição de rotâmeros de cadeia lateral  |
| <b>-EX2ARO</b>                                   | Adição de rotâmeros de cadeia lateral  |
| <b>-CONSTRAINTS:CST_FILE<br/>CONSTRAINTS.CST</b> | Arquivo com restrições de distâncias (XL-MS)   |
| <b>-NO_FILTERS</b>                               | Filtros do score <i>docking</i> desligado (energia total negativa, energia de interação negativa e energia de restrições $\leq 1$ ). |
| <b>-OUT:FILE:FULLATOM</b>                        | Modelo gerado em modo contendo todos os átomos.  |
| <b>-OUT:SUFFIX_A</b>                             | Sufixo adicionado ao nome dos modelos gerados.   |
| <b>-OUT:PATH:ALL./POSES_A/</b>                   | Sub-diretório de armazenamento.  |

O potencial restritivo utilizado foi o BOUNDED, com limite inferior configurado para zero e o limite superior configurado para os valores da coluna  $L_{MAX}$  da **Tabela 4**, de acordo com as cadeias laterais envolvidas e com a estratégia de *cross-link* utilizada. Foram produzidos entre 100.000 e 200.000 modelos para cada simulação. A variação foi permitida para acumular pelo menos 100 modelos com o menor número de violações dos dados de *cross-linking*.

### 3.2.20.5 – Seleção dos modelos gerados

Os modelos gerados foram analisados pelo *software* Topolink e classificados de acordo com o número total de violações das restrições experimentais de XL-MS. O grupo de modelos contendo o menor número de violações é então classificado de acordo com os conjuntos de restrições experimentais que foram validadas/violadas simultaneamente. Na sequência, os subgrupos criados tiveram seus componentes ordenados de acordo com a energia de potencial de restrição, energia de interação da superfície de *docking* e energia total expressos em unidade de energia Rosetta (REU). Os grupos foram então analisados quanto ao RMSD relativo à média das coordenadas e quanto ao RMSD relativo ao modelo de menor energia. Os grupos contendo modelos de menor energia e com maior convergência entre si (medida pelo RMSD) foram selecionados para inspeção visual. O modelo representante do grupo foi novamente analisado pelo Topolink, ativando a opção para gerar coordenadas representativas da distância topológica para os *cross-links* validados.

### 3.2.20.6 – Visualização de modelos e imagens

Os modelos obtidos foram visualizados utilizando o *software* VMD (*Visual Molecular Dynamics*). Todas as imagens foram geradas utilizando o módulo *render* em modo *tachyon* com resolução de 4000 x 4000 pixels em uma área de 16 x 16 cm.

### 3 – RESULTADOS E DISCUSSÃO

#### 4.1 – CARACTERIZAÇÃO FÍSICO-QUÍMICA DO INIBIDOR DE METALOENDOPEPTIDASES BJ46a

##### 4.1.1 – Purificação do inibidor BJ46a do plasma da serpente *Bothrops jararaca*

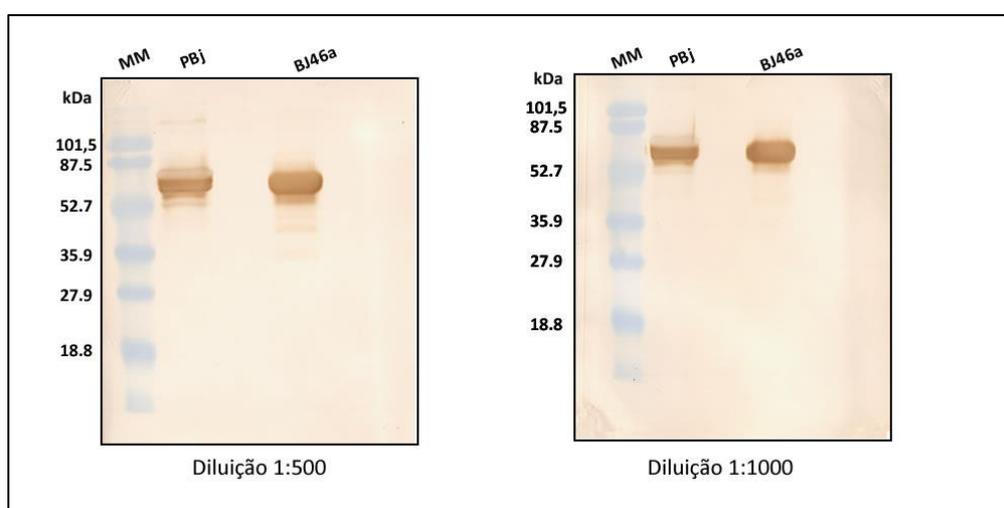
A metodologia de purificação do inibidor de metaloendopeptidases BJ46a originalmente publicada envolve uma série de etapas purificativas: precipitação fracionada com sulfato de amônio, cromatografia de interação hidrofóbica em coluna Phenyl Sepharose CL-4B e duas cromatografias sequenciais de fase reversa em coluna C4, com um rendimento purificativo final, em termos de quantidade de proteína, de 0,55% (Valente *et al.*, 2001). Em trabalho anterior, otimizamos o gradiente de separação da cromatografia de fase reversa, o que permitiu dispensar o passo de recromatografia, mas sem alteração no rendimento purificativo (Bastos, 2014).

Assim, partimos para a otimização da primeira parte do processo purificativo, substituindo a precipitação fracionada por sulfato de amônio e a cromatografia de interação hidrofóbica pela cromatografia de afinidade utilizando anticorpos policlonais anti-BJ46a, produzidos a partir da imunização de coelhos com uma solução contendo BJ46a nativa e homogênea. A ligação covalente de anticorpos em uma matriz cromatográfica específica para este fim é um método versátil e específico de isolar moléculas-alvo oriundas de misturas complexas (Nisnevitch e Firer, 2001), como é o caso do plasma de *B.jararaca*.

Inicialmente, avaliamos a reatividade do soro de coelhos contendo os anticorpos policlonais anti-BJ46a frente às proteínas do plasma de *B.jararaca*, como demonstrado na **Figura 12**. Utilizamos o soro bruto contendo as IgG anti-BJ46a como anticorpo primário e, nas duas diluições testadas (1:500 e 1:1000), os anticorpos mostraram-se reativos à proteína BJ46a (raia BJ46a), isolada de acordo com o método descrito em (Bastos, 2014). Com relação à raia que continha plasma bruto de *B.jararaca* (raia PBj), ocorreu reconhecimento em uma banda de migração compatível com o inibidor BJ46a. Não ocorreu revelação de outras bandas que correspondessem às demais proteínas

plasmáticas, o que possibilita a aplicação destes anticorpos para a cromatografia de imunoafinidade.

Para a purificação do *pool* de imunoglobulinas G a partir do soro anti-BJ46a, utilizamos uma coluna cromatográfica acoplada covalentemente com proteína A isolada de *Staphylococcus aureus*, que apresenta afinidade para a porção Fc de IgG oriundas de coelhos e de outras fontes (Hjelm *et al.*, 1972). Em seguida, procedemos ao acoplamento covalente destes anticorpos em uma coluna HiTrap NHS HP, segundo as instruções do fabricante (GE Healthcare). A coluna é composta de uma matriz cromatográfica inerte de agarose (Sephacrose) substituída com grupamentos funcionais N-hidroxissuccinimida (NHS), que reagem com aminas primárias expostas do anticorpo, bem como a extremidade N-terminal (Nisnevitch e Firer, 2001). A quantidade inicial de anticorpo acoplada foi de 50 mg; obtivemos um rendimento final de acoplamento de 91% (45,5 mg de IgG efetivamente acoplados).

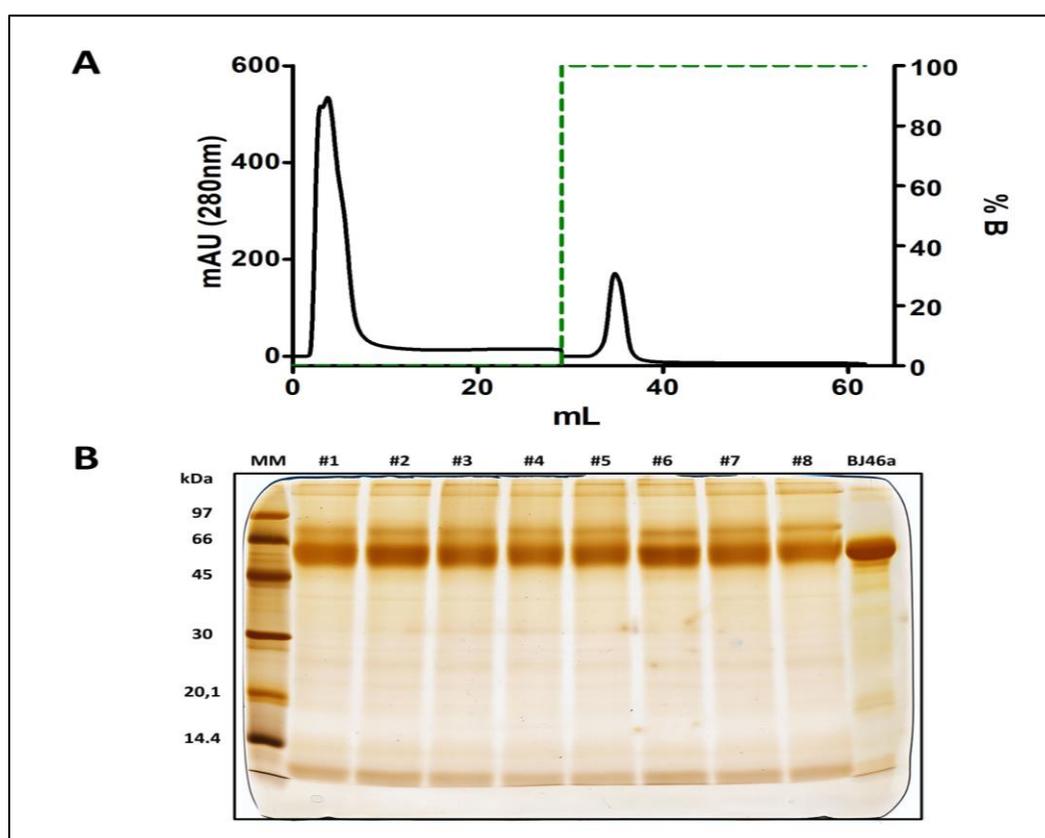


**Figura 12: Immunoblotting do plasma de *Bothrops jararaca* e do inibidor BJ46a.**

PBj = plasma bruto de *B. jararaca* (10 µg); BJ46a: BJ46a controle (1µg); MM: marcador pré-corado de massa molecular. Membranas de nitrocelulose incubadas com soro anti-BJ46a como anticorpo primário, nas diluições de 1:500 e 1:1000 e IgG anti-coelho conjugada à peroxidase como anticorpo secundário (diluição de 1:5000). As membranas foram reveladas por 3,3-diaminobenzidina (DAB).

Para a purificação de BJ46a a partir desta coluna preparada de imunoafinidade (NHS-IgG), aplicamos 2 mg de plasma filtrado de *Bothrops jararaca* diluído em tampão Tris-HCl 20mM pH 7,5 por corrida. O perfil cromatográfico está representado no painel **A** da **Figura 13**. Foram realizadas 47 corridas cromatográficas em condições idênticas.

Todas as frações proteicas que interagiram com as IgG ligadas à matriz cromatográfica (frações ligadas) foram unidas em 8 *pools* e analisadas por eletroforese em condições desnaturantes e redutoras (painel **B**, **Figura 13**). Ao compararmos estes *pools* de frações ligadas (amostras #1 - #8) com BJ46a homogênea controle (amostra BJ46a) observamos um enriquecimento na faixa da massa molecular aparente de BJ46a ( $\approx 55$  kDa), porém ainda com a presença de contaminantes de alta ( $> 97$  kDa) e baixa massas moleculares ( $< 14,4$  kDa), requerendo ao menos um passo cromatográfico adicional para a obtenção de BJ46a em sua forma homogênea.



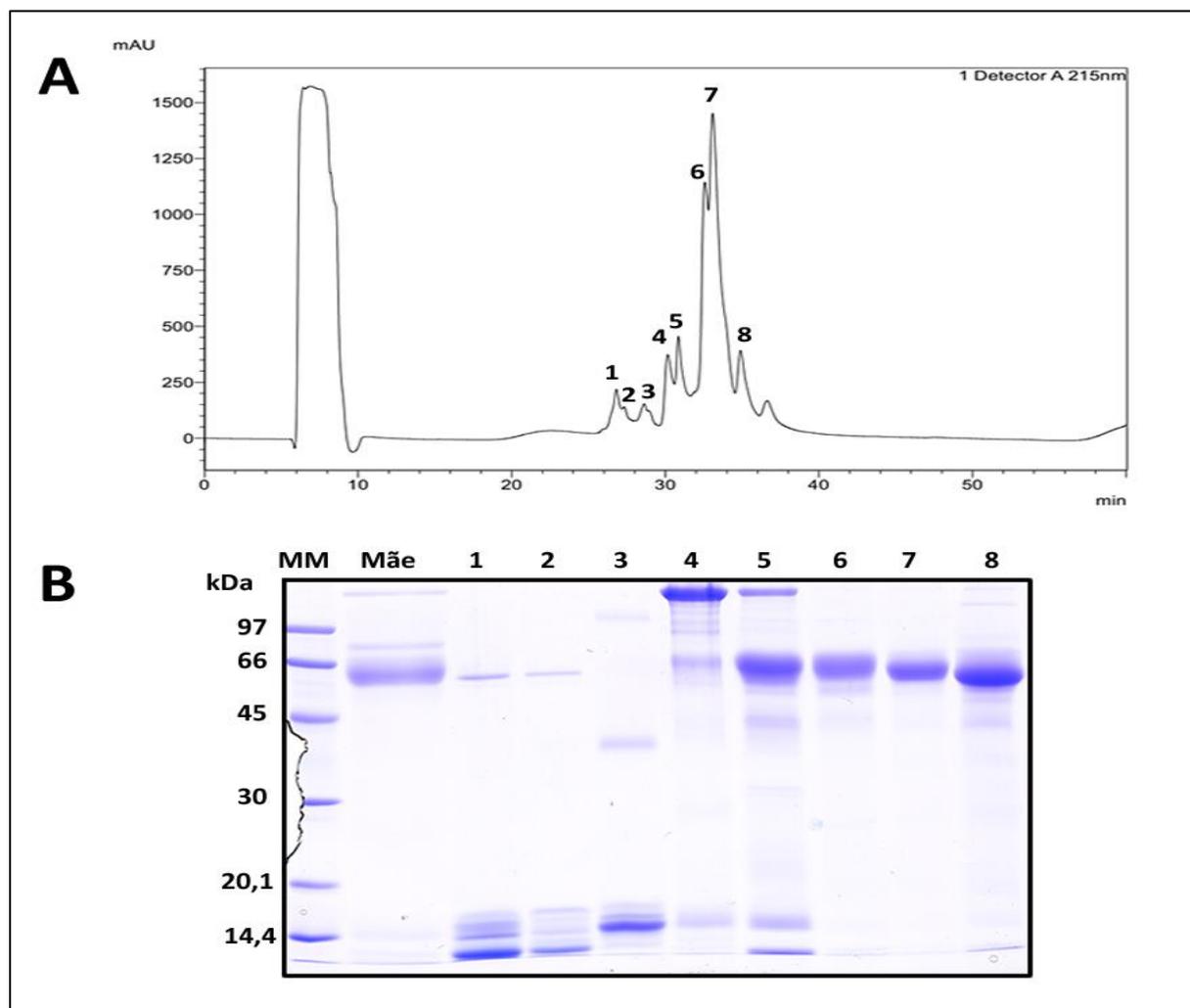
**Figura 13: Cromatografia de imunoafinidade do plasma de *B. jararaca* em coluna HiTrap NHS-IgG anti-BJ46a.**

**Painel A:** Cromatograma de imunoafinidade representativo do plasma de *B. jararaca*. Para cada corrida, injetaram-se 2 mg de plasma bruto de *B. jararaca*. A cromatografia foi realizada em sistema cromatográfico Äkta Purifier (Amersham Biosciences) e a absorvância foi monitorada a 280nm.

**Painel B:** SDS-PAGE dos *pools* de frações ligadas oriundas da cromatografia de imunoafinidade do plasma de *B. jararaca*. MM: marcador de massa molecular; #1- #8: pools de frações cromatográficas enriquecidas em BJ46a (3  $\mu$ g); BJ46a: BJ46a controle (2  $\mu$ g). Gel revelado por impregnação por prata.

Para a etapa final de purificação do inibidor, cada *pool* de fração ligada foi liofilizado e submetido à cromatografia de fase reversa em coluna C4 semi-preparativa (10 x 250mm),

gerando o perfil cromatográfico representativo representado no painel **A** da **Figura 14**. Em seguida, todas as frações coletadas foram secas em centrífuga a vácuo (*speed-vac*) e analisadas por eletroforese em condições desnaturantes e redutoras (painel **B**, **Figura 14**).



**Figura 14:** Cromatografia em fase reversa em coluna C4 semi-preparativa (10 x 250 mm) dos *pools* de frações ligadas oriundas da cromatografia de imunoafinidade.

**Painel A:** Cromatograma representativo da separação por fase reversa dos *pools* de frações ligadas. As frações coletadas estão numeradas acima de cada pico cromatográfico. Para cada corrida, injetou-se o equivalente a 1 mg de proteína.

**Painel B:** SDS-PAGE das frações coletadas da cromatografia de fase reversa. MM: marcador de massa molecular; Mãe: *pool* de fração ligada injetada na coluna de fase reversa (3 µg); 1-8: frações coletadas da cromatografia de fase reversa. Corado por Coomassie R-250.

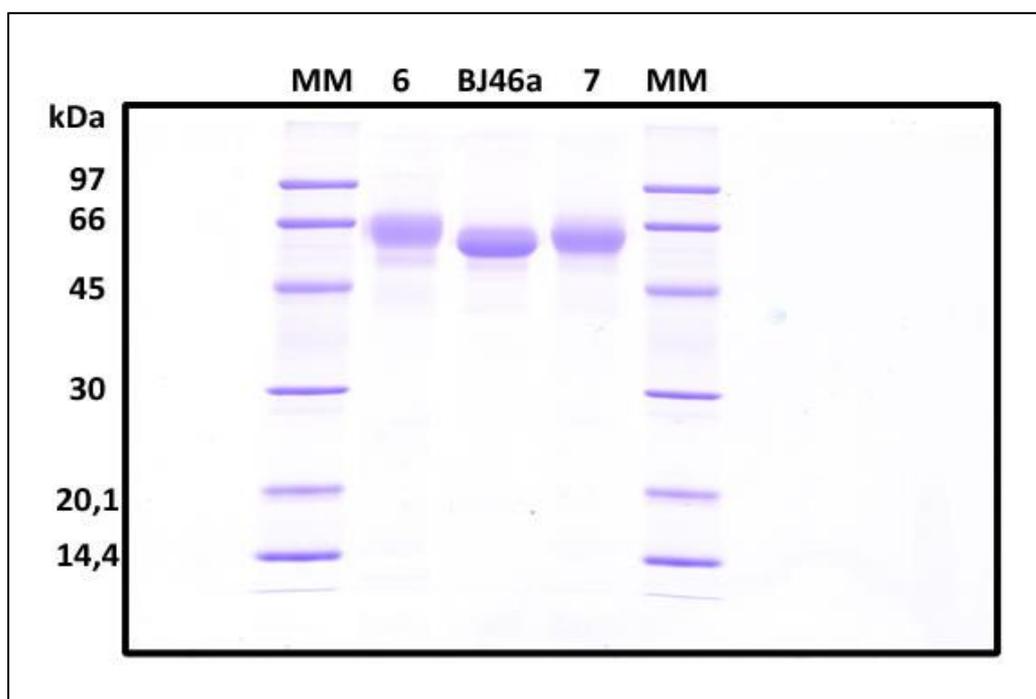
A separação dos *pools* oriundos da cromatografia de imunoafinidade por cromatografia de fase reversa resultou em dois picos majoritários (frações 6 e 7) e 6 outros picos, minoritários em relação às frações 6 e 7. Ao analisarmos o conteúdo destas frações por SDS-PAGE

(**Figura 14**, painel B) verificamos a presença de proteínas com migração eletroforética próxima à migração do marcador de massa molecular de 14.4 kDa nas frações 1, 2 e 3, além da presença de uma banda majoritária de massa molecular aparente compatível com o inibidor BJ46a nas frações 5,6,7 e 8.

Em cada uma destas frações, a banda correspondente ao inibidor BJ46a apresentou pequenas diferenças na migração eletroforética. Na literatura, é descrito que BJ46a apresenta pelo menos uma proteoforma, BJ46b, tida como menos abundante e com uma massa molecular de 46,389 Da por MALDI-TOF (Valente *et al.*, 2001). É possível que as frações 5,6,7 e 8 contenham outras proteoformas de BJ46a, que podem apresentar diferenças em sua estrutura primária e/ou em suas porções glicídicas, visto que glicídios N-ligados podem apresentar intensa micro- e macroheterogeneidade, e estas mudanças podem ser refletidas em suas migrações eletroforéticas (Marino *et al.*, 2010).

Em 2007, Aoki e colaboradores purificaram do plasma da serpente *Protobothrops (Trimeresurus) flavoviridis* proteínas de baixa massa molecular chamadas de *small serum proteins* (SSP), com massas moleculares de 6,5 – 10 kDa (Aoki, Sakiyama, *et al.*, 2007). A presença de proteínas de baixa massa molecular (< 14,4 kDa) nas frações 1, 2 e 3 levanta a hipótese de que possam corresponder a proteínas homólogas a SSP no plasma de *B. jararaca*. Em outro trabalho, demonstrou-se que o inibidor HSF, isolado da mesma serpente, pode atuar como uma proteína carreadora de SSP, o que talvez possa explicar a presença de uma banda de baixa massa molecular na fração 5, mas outros experimentos precisam ser realizados para confirmar estas hipóteses (Shioi *et al.*, 2011).

Das frações que apresentavam enriquecimento na faixa de massa molecular de BJ46a, as frações 6 e 7 mostraram-se mais abundantes e homogêneas. Realizamos uma análise comparativa destas frações com uma alíquota de BJ46a purificada previamente, como demonstra a **Figura 15**. Ao compararmos as migrações eletroforéticas das frações 6 e 7 com o controle de BJ46a, verificamos que a fração 7 apresentava uma migração mais compatível com o controle do inibidor em sua forma homogênea, além de ser a isoforma mais abundante de acordo com o cromatograma de fase reversa (**Figura 14**, painel A).

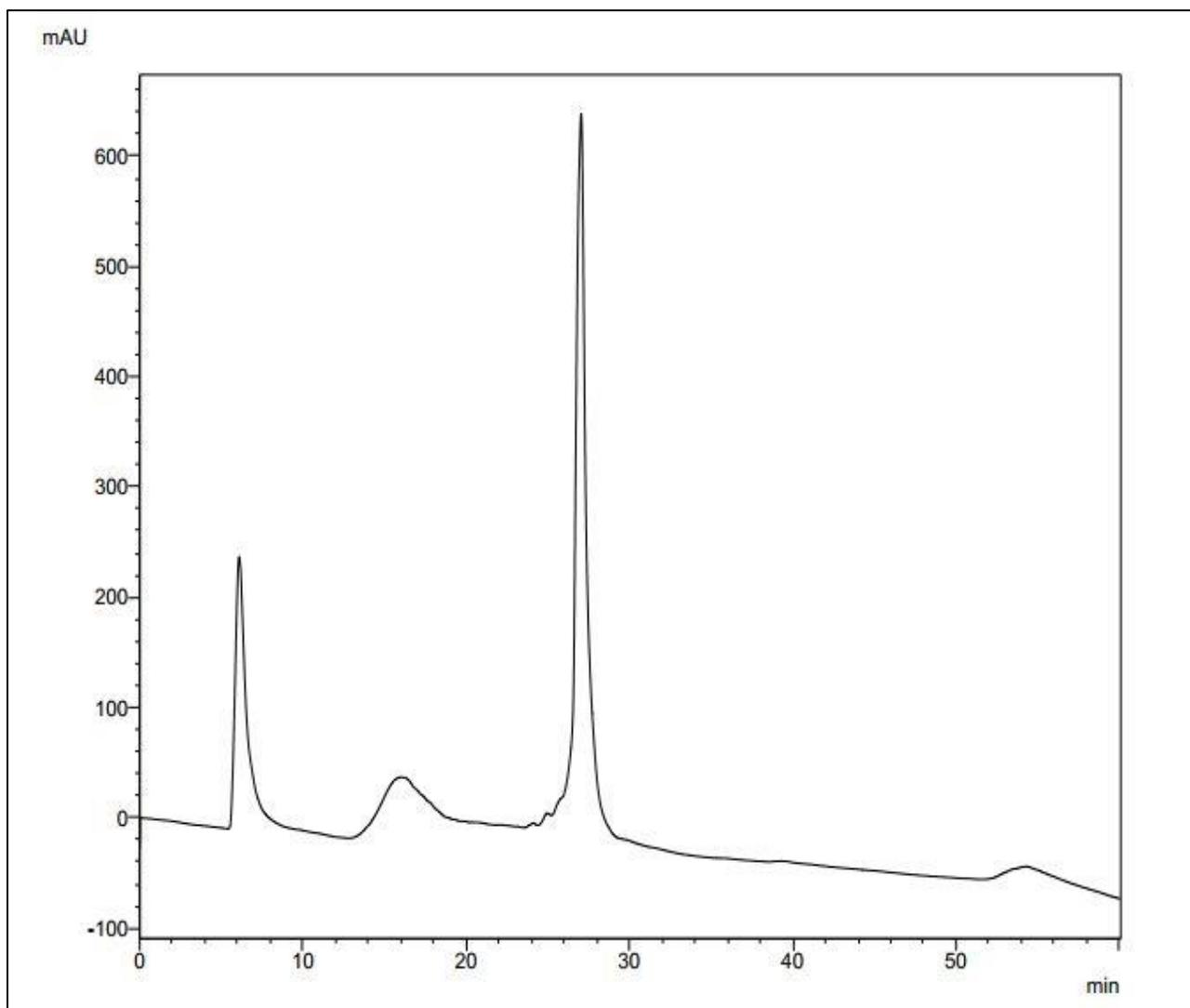


**Figura 15: SDS-PAGE das frações 6 e 7 oriundas da cromatografia de fase reversa.**

MM: marcador de massa molecular; 6: fração 6 da cromatografia de fase reversa (3 µg); 7: fração 7 da cromatografia de fase reversa (3 µg); BJ46a: BJ46a controle (3 µg). Corado por Coomassie R-250.

Por fim, uma alíquota da fração 7 foi submetida à recromatografia de fase reversa nas mesmas condições experimentais para análise da homogeneidade (**Figura 16**). Obtivemos um pico majoritário e gaussiano, indicando que não ocorreu contaminação cruzada detectável com as frações contíguas (frações 6 e 8).

Para a etapa final do processo purificativo, realizamos dessalinização da fração 7 em uma coluna de exclusão molecular HiTrap Desalting para remover o tampão Tris, utilizado para neutralizar o pH das frações coletadas da cromatografia de fase reversa, para o tampão HEPES 10 mM pH 7,5 para não comprometer as análises subsequentes de XL-MS. O rendimento purificativo final, em termos de quantidade de proteína, foi de 0,83%, como descrito na **Tabela 7**. Com a substituição dos passos de precipitação fracionada por sulfato de amônio e cromatografia de interação hidrofóbica por uma cromatografia de imunoafinidade, aumentamos o rendimento purificativo em cerca de 50%.



**Figura 16: Recromatografia em fase reversa de BJ46a purificada (fração 7) após os passos cromatográficos de imunoafinidade e fase reversa.**

Injetaram-se 100 µg de amostra em uma coluna C4 semi-preparativa (10 x 250 mm); a absorvância foi medida a 215nm.

Tabela 7: Rendimento purificativo do novo método de purificação de BJ46a e comparação com o método originalmente publicado.

| Etapa purificativa  | Quantidade de proteína (mg) | Rendimento (%) | Método original (Valente <i>et al.</i> 2001) |
|---|-----------------------------|----------------|--|
| Plasma de <i>B. jararaca</i> filtrado   | 94                          | 100            | -  |
| Cromatografia de imunoafinidade em coluna HiTrap NHS-IgG anti-BJ46a – fração ligada | 9,3                         | 9,89           | -  |
| Cromatografia em fase reversa em coluna C4 (10 x 250 mm) – fração 7                 | 1,43                        | 1,52           | -  |
| Dessalinização em coluna HiTrap Desalting   | 0,78                        | <b>0,83</b>    | <b>0,55</b>                                  |

#### 4.1.2 – Sequenciamento *de novo* do inibidor BJ46a

Para verificar eventuais modificações na estrutura primária de BJ46a purificada, realizamos um experimento de sequenciamento *de novo* da amostra do inibidor em alta resolução, utilizando o espectrômetro de massas Q-Exacte Plus (item 3.2.10, capítulo 3). A partir dos dados de sequenciamento *de novo* dos espectros obtidos pelo software PEAKS Studio, selecionamos os resultados de sequenciamento de alta confiabilidade (maior ou igual a 90% de confiança) e alinhamos estes utilizando o programa PepExplorer, pertencente à suíte de aplicativos *PatternLab for Proteomics* (Leprevost *et al.*, 2014; Carvalho *et al.*, 2016). A busca retornou 21 sequências alinhadas, em um total de 21.823 *spectral counts*, como demonstra a **Tabela 8**.

**Tabela 8: Alinhamentos dos peptídeos sequenciados *de novo* pela ferramenta PepExplorer.**

Id: identificação da entrada no banco de dados UNIPROT/TrEMBL/SwissProt; Alinhamentos: números de alinhamentos peptídicos detectados por uma dada sequência; #Espectros: contagem espectral do alinhamento; Únicos: peptídeos únicos associados à entrada; Cobertura: cobertura de sequência; Descrição: descrição da entrada no banco de dados. As entradas identificadas foram organizadas na tabela por ordem decrescente do número de alinhamentos para cada sequência.

| ID                     | Alinhamentos | #Espectros | Únicos | Cobertura | Descrição  |
|------------------------|--------------|------------|--------|-----------|--|
| Q9DGI0.1               | 1323         | 7653       | 1144   | 0.93      | RecName: Full=Antihemorrhagic factor BJ46a; [ <i>Bothrops jararaca</i> ]   |
| Q5KQS2.1               | 477          | 2818       | 347    | 0.868     | RecName: Full=Antihemorrhagic factor [ <i>Gloydius brevicaudus</i> ]   |
| Q5KQS4.1               | 341          | 2368       | 0      | 0.63      | RecName: Full=Antihemorrhagic factor cMSF [ <i>Gloydius brevicaudus</i> ]  |
| Q5KQS1.1               | 340          | 2366       | 0      | 0.62      | RecName: Full=Antihemorrhagic factor jMSF [ <i>Gloydius blomhoffii</i> ]   |
| P29695.2               | 244          | 1309       | 11     | 0.702     | RecName: Full=Antihemorrhagic factor HSF [ <i>Protobothrops flavoviridis</i> ]   |
| AAB24832.1             | 241          | 1305       | 2      | 0.774     | antihemorrhagic factor HSF=venom metalloproteinase inhibitor [ <i>Trimeresurus flavoviridis</i> =Japanese Habu snakes, serum, Peptide, 323 aa] |
| Q5KQS3.1               | 226          | 2335       | 28     | 0.398     | RecName: Full=Antihemorrhagic factor cHLP-A [ <i>Gloydius brevicaudus</i> ]  |
| Q5KQS5.1               | 153          | 1006       | 40     | 0.543     | RecName: Full=HSF-like protein; [ <i>Protobothrops flavoviridis</i> ]  |
| tr D9N568 D9N568_GLOBB | 34           | 121        | 30     | 0.523     | Small serum protein-2 OS= <i>Gloydius blomhoffii blomhoffii</i>  |
| tr D9N567 D9N567_GLOBB | 16           | 127        | 12     | 0.477     | Small serum protein-1 OS= <i>Gloydius blomhoffii blomhoffii</i>  |
| tr D9N569 D9N569_GLOBB | 13           | 53         | 11     | 0.673     | Small serum protein-3 OS= <i>Gloydius blomhoffii blomhoffii</i>  |
| tr D9N570 D9N570_GLOBB | 13           | 46         | 7      | 0.654     | Small serum protein-4 OS= <i>Gloydius blomhoffii blomhoffii</i>  |
| sp A7VN13 MSMB1_PROFL  | 13           | 73         | 0      | 0.477     | Small serum protein 1 OS= <i>Protobothrops flavoviridis</i>  |
| tr L0N0Z0 L0N0Z0_PROFL | 12           | 71         | 0      | 0.393     | Small serum protein-1 OS= <i>Protobothrops flavoviridis</i>  |
| sp A7VN14 MSMB2_PROFL  | 10           | 42         | 0      | 0.349     | Small serum protein 2 OS= <i>Protobothrops flavoviridis</i>  |
| tr L0N1W8 L0N1W8_PROFL | 10           | 42         | 0      | 0.349     | Small serum protein-2 OS= <i>Protobothrops flavoviridis</i>  |
| sp A7VN16 MSMB4_PROFL  | 6            | 14         | 1      | 0.407     | Small serum protein 4 OS= <i>Protobothrops flavoviridis</i>  |
| sp A7VN15 MSMB3_PROFL  | 6            | 21         | 5      | 0.329     | Small serum protein 3 OS= <i>Protobothrops flavoviridis</i>  |
| tr D9N571 D9N571_GLOBB | 6            | 31         | 4      | 0.22      | Small serum protein-5 OS= <i>Gloydius blomhoffii blomhoffii</i>  |
| sp P0CB15 CRIS_PROFL   | 5            | 15         | 5      | 0.036     | Serotriflin OS= <i>Protobothrops flavoviridis</i>  |
| sp A7VN17 MSMB5_PROFL  | 3            | 7          | 1      | 0.13      | Small serum protein 5 OS= <i>Protobothrops flavoviridis</i>  |

Para o inibidor BJ46a, obtivemos 1323 alinhamentos, resultando em uma cobertura de 93% em relação à sequência precursora do inibidor (entrada Q9DGI0 do banco de dados UNIPROTKB). Além disso, detectamos também peptídeos alinhados com entradas referentes às proteínas HSF, HLP, cMSF, jMSF e *small serum proteins* (SSP). Realizamos a comparação entre as sequências de BJ46a e entre estas proteínas por meio do algoritmo BLAST, como demonstrado na figura a seguir:

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

Alignments [Download](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

|                          | Description   | Max score | Total score | Query cover | E value | Ident | Accession    |
|--------------------------|---|-----------|-------------|-------------|---------|-------|--------------|
| <input type="checkbox"/> | <a href="#">P29695.2</a> RecName: Full=Antihemorrhagic factor HSF; AltName: Full=Habu serum factor; AltName: Full=Metalloproteinase inhibitor; Flags: Precursor >BAB3 | 605       | 605         | 100%        | 0.0     | 86%   | Query_242510 |
| <input type="checkbox"/> | <a href="#">AAB24832.1</a> antihemorrhagic factor HSF=venom metalloproteinase inhibitor (Trimeresurus flavoviridis=Japanese Habu snakes, serum, Peptide, 323 aa)      | 566       | 566         | 94%         | 0.0     | 85%   | Query_242511 |
| <input type="checkbox"/> | <a href="#">Q5KQS1.1</a> RecName: Full=Antihemorrhagic factor jMSF; AltName: Full=Japanese mamushi serum factor; Flags: Precursor >BAD88539.1 antihemorrhagic fa      | 546       | 546         | 100%        | 0.0     | 79%   | Query_242512 |
| <input type="checkbox"/> | <a href="#">Q5KQS4.1</a> RecName: Full=Antihemorrhagic factor cMSF; AltName: Full=Chinese mamushi serum factor; Flags: Precursor >BAD88536.1 antihemorrhagic fac      | 542       | 542         | 100%        | 0.0     | 78%   | Query_242513 |
| <input type="checkbox"/> | <a href="#">Q5KQS3.1</a> RecName: Full=Antihemorrhagic factor cHLP-A; AltName: Full=Chinese mamushi HSF-like protein A; Flags: Precursor >BAD88537.1 fetuin family    | 497       | 497         | 100%        | 0.0     | 72%   | Query_242514 |
| <input type="checkbox"/> | <a href="#">Q5KQS2.1</a> RecName: Full=Antihemorrhagic factor cHLP-B; AltName: Full=Chinese mamushi HSF-like protein B; Flags: Precursor >BAD88538.1 fetuin family    | 491       | 491         | 100%        | 1e-179  | 72%   | Query_242515 |
| <input type="checkbox"/> | <a href="#">Q5KQS5.1</a> RecName: Full=HSF-like protein; Short=HLP; Flags: Precursor >BAD88535.1 HSF-like protein precursor [Protobothrops flavoviridis]              | 478       | 478         | 100%        | 9e-175  | 70%   | Query_242516 |
| <input type="checkbox"/> | <a href="#">sp P0CB15 CRIS_PROFL</a> Serotriifin OS=Protobothrops flavoviridis PE=1 SV=1  | 17.3      | 17.3        | 13%         | 6.7     | 22%   | Query_242522 |

Figura 17: Alinhamento da sequência de BJ46a com HSF, HLP, cMSF e jMSF realizadas pelo BLAST.

O resultado do alinhamento indica que BJ46a apresenta identidade sequencial de 85% com HSF, outro inibidor de metaloendopeptidases de venenos de serpentes da família das cistatinas, como descrito anteriormente (capítulo 1, seção 1.3.1) e identidades de 79% e 78% com as proteínas jMSF e cMSF. A proteína jMSF (*japanese mamushi serum factor*) foi isolada do plasma da serpente *Gloydius blomhoffii* e cMSF, da serpente *Gloydius blomhoffii brevicaudus*. Ambas foram classificadas como inibidores de metaloendopeptidases da família das cistatinas, inibindo diferentes metaloendopeptidases purificadas do veneno de *Protobothrops flavoviridis*. Entretanto, ao comparar as sequências de cMSF e jMSF com a sequência do inibidor HSF, os autores identificaram uma deleção de 17 resíduos em suas

porções C-terminais; tal deleção não afetou as propriedades inibitórias destas proteínas, mas levou à menor estabilidade térmica de cMSF (Aoki, Tsutsumi, *et al.*, 2008). A análise do alinhamento entre BJ46a e jMSF / cMSF indicou que a maior parte das identidades observadas se deu nos domínios cistatina de BJ46a, representando, portanto, peptídeos compartilhados entre as sequências (resíduos 22-130 e 141-254, **Figura 18**).

| BJ46a - jMSF<br>(query) (subject)   |  |        |                              |           |      | BJ46a - cMSF<br>(query) (subject)   |        |        |                              |           |      |
|---|--|--------|------------------------------|-----------|------|---|--------|--------|------------------------------|-----------|------|
| Range 1: 1 to 324 <a href="#">Graphics</a> <span style="float:right">▼ Next Match ▲ Previous Match</span> |  |        |                              |           |      | Range 1: 1 to 324 <a href="#">Graphics</a> <span style="float:right">▼ Next Match ▲ Previous Match</span> |        |        |                              |           |      |
| Score   | Expect   | Method | Identities                   | Positives | Gaps | Score   | Expect | Method | Identities                   | Positives | Gaps |
| 546 bits(1407) 0.0  |  |        | Compositional matrix adjust. |           |      | 269/341(79%)  |        |        | 291/341(85%)                 |           |      |
| 17/341(4%)  |  |        |                              |           |      | 542 bits(1396) 0.0  |        |        | Compositional matrix adjust. |           |      |
| 267/341(78%)  |  |        | 291/341(85%)                 |           |      | 17/341(4%)  |        |        |                              |           |      |
| Query 1   | MNSLVALVLLGQIIGSTLSSQVRGDLCEKDAKENTDTGVRVINEHKLHGKYLINVIK    | 60     |                              |           |      |   |        |        |                              |           |      |
| M+  | LVALVLLGQIIGSTLSSQVRGDLCEC+++AKEN D VRYINEHKLH YK ALINVIK    |        |                              |           |      |   |        |        |                              |           |      |
| Sbjct 1   | MHFLVALVLLGQIIGSTLSSQVRGDLCECDREAKENADQAVRYINEHKLHEKQALINVIK | 60     |                              |           |      |   |        |        |                              |           |      |
| Query 61  | NIVVWPNQDGMVAVFLKLNLLLETECHVLDPTPVKICTVRPQHHAVENTCDVKIMFIVDT | 120    |                              |           |      |   |        |        |                              |           |      |
| NIVVWPNH+GD   | VAVFLKLNLLLETECHVLDPTPV+ CT+RPQ NHAVENDCD KIMF+V+T           |        |                              |           |      |   |        |        |                              |           |      |
| Sbjct 61  | NIVVWPNQDGMVAVFLKLNLLLETECHVLDPTPVKICTVRPQHHAVENTCDVKIMFIVDT | 120    |                              |           |      |   |        |        |                              |           |      |
| Query 121   | FKEDVFAKCHSTPDSVEIVRRNCKPCPILLPSNIPQVDSVEVYLKHKHIEKLSOHVYEV  | 180    |                              |           |      |   |        |        |                              |           |      |
| FK+DVF  | KCHSTPDSVE+VRRNCKPCPILL ++P VDSVEVYLKHKHIE+LS HVYEV          |        |                              |           |      |   |        |        |                              |           |      |
| Sbjct 121   | FKQDVFVKCHSTPDSVEDVRRNCKPCPILLSPDPHVVDSVEVYLKHKHIEQLSGHVYEV  | 180    |                              |           |      |   |        |        |                              |           |      |
| Query 181   | EISRGQHKYEPEAYVEFAIVEVICTAQELHDDHHCHPNTAGEDHIGFCRATVFRSHAS   | 240    |                              |           |      |   |        |        |                              |           |      |
| EISRGQHKYEPEA+YVEFAIVEVICTAQE   | HDDHHCHPNTAGEDHI FC+ATVFRSHAS                                |        |                              |           |      |   |        |        |                              |           |      |
| Sbjct 181   | EISRGQHKYEPEAFYVEFAIVEVICTAQEAHDDHHCHPNTAGEDHIAFCATVFRSHAS   | 240    |                              |           |      |   |        |        |                              |           |      |
| Query 241   | LEKPKDEQFESDCVILVKEGHAHSHLIQHVKEKDSISPEHNTALNIVFHPHNDTSTSHE  | 300    |                              |           |      |   |        |        |                              |           |      |
| LEKPK E FESDCVIL  | VKEGHAHSHLI+ H+ K S SP N+T                                   |        |                              |           |      |   |        |        |                              |           |      |
| Sbjct 241   | LEKPKHENFESDCVILDVKEGHAHSHLIEHIGKYSTSPGQNSTV-----            | 285    |                              |           |      |   |        |        |                              |           |      |
| Query 301   | SHEHLAEPVAFVVKELPKDISDRHTTPVKGCPCGVHFFEL                     | 341    |                              |           |      |   |        |        |                              |           |      |
| E +AE   | PVAFV KE+P DISDRHTTPVKGCPCG+ HF+L                            |        |                              |           |      |   |        |        |                              |           |      |
| Sbjct 286   | --ECVAECPVAFVVKKEVPTDISDRHTTPVKGCPCGKILHFQL                  | 324    |                              |           |      |   |        |        |                              |           |      |
| Query 301   | SHEHLAEPVAFVVKELPKDISDRHTTPVKGCPCGVHFFEL                     | 341    |                              |           |      |   |        |        |                              |           |      |
| E +AE   | PVAFV KE+P DISDRHTTPVKGCPCG+ HF+L                            |        |                              |           |      |   |        |        |                              |           |      |
| Sbjct 286   | --ECVAECPVAFVVKKEVPTDISDRHTTPVKGCPCGKILHFQL                  | 324    |                              |           |      |   |        |        |                              |           |      |

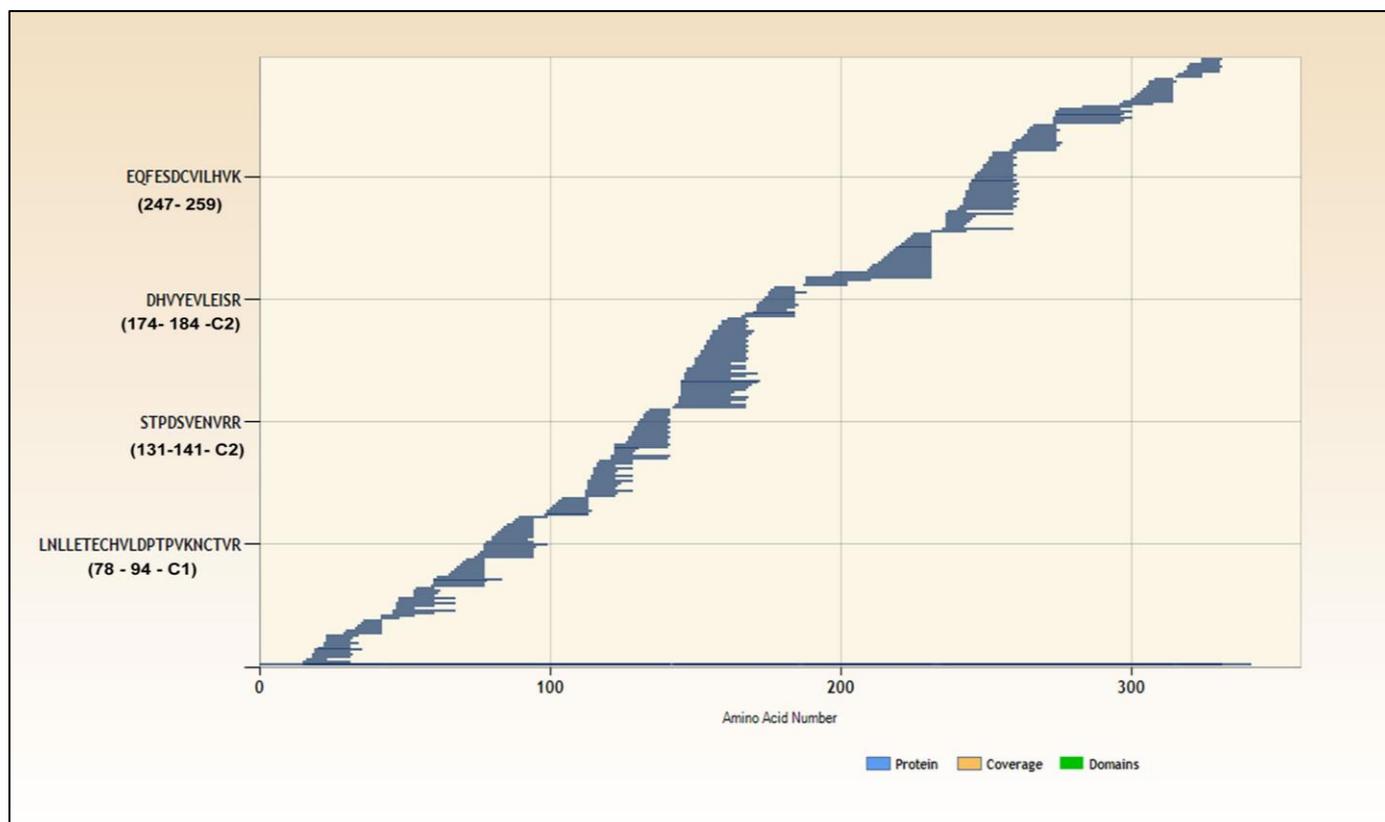
**Figura 18: Alinhamento da sequência de BJ46a com os inibidores jMSF e cMSF.**

Além de jMSF e cMSF, BJ46a também apresentou 72% de identidade sequencial com as proteínas cHLP-B e cHLP-A, isoladas da serpente *G.b. brevicaudus*. Estas proteínas apresentam, assim como HSF e BJ46a, dois domínios cistatina e uma região rica em histidina em suas estruturas, mas não apresentam atividade anti-hemorrágica. A análise filogenética das sequências de BJ46a, HSF, HLP e fetuínas verdadeiras de serpentes indica que estas proteínas evoluíram de um ancestral comum via duplicação gênica, adquirindo diferentes funções, como a resistência aos efeitos deletérios do envenenamento (Aoki *et al.*, 2009; Bastos *et al.*, 2016). Logo, é esperado que haja alinhamentos de peptídeos para as sequências de cHLP-B e cHLP-A.

Identificamos ainda em nossas análises peptídeos alinhados com as sequências referentes a pequenas proteínas séricas (*small serum proteins*, SSP) isoladas do plasma de *Gloydus brevicaudus* e *Protobothrops flavoviridis*. As SSP destas serpentes possuem uma massa molecular na faixa de 6,5 – 13 kDa e são homólogas a uma proteína prostática (PSP94) oriunda de mamíferos (Aoki, Sakiyama, *et al.*, 2007; Shioi *et al.*, 2014). A análise dos cDNA das SSP de *P. flavoviridis* indica que estas proteínas sofreram uma evolução acelerada, levando estas a adquirir funções inibitórias intrínsecas e também por meio da interação com HSF, que apresenta uma função carreadora de SSP. A constante de dissociação, medida por SPR, foi calculada em  $2,5 \pm 1,1$  nM e a redução do pH da solução para 2,0 não reduziu a afinidade de ligação de SSP à HSF (Aoki, Matsuo, *et al.*, 2008; Shioi *et al.*, 2011).

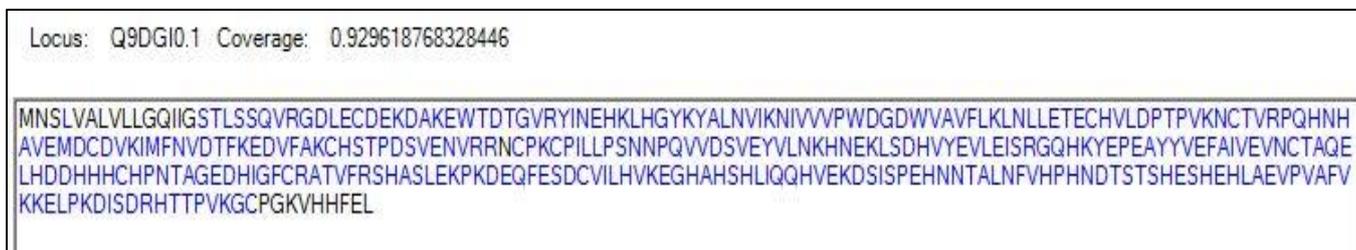
A presença de peptídeos homólogos à SSP no conjunto de dados de sequenciamento *de novo* de BJ46a reforça a hipótese de que possa haver SSP homólogas às encontradas em *P. flavoviridis* e *G. brevicaudus* no plasma de *B. jararaca*, o que ainda não havia sido descrito em estudos anteriores. Porém, ainda não é possível determinar a influência destas SSP putativas do plasma de *B. jararaca* na atividade inibitória de BJ46a. Mas, traçando-se um paralelo com o inibidor HSF, a interação do inibidor com SSP-1 de *P. flavoviridis* não influenciou a atividade inibitória de HSF frente as SVMP HR1a, HR1b, HR2a e HR2b isoladas do veneno da mesma serpente (Shioi *et al.*, 2011). Por outro lado, os dados na literatura sugerem que a interação de HSF com SSP possibilita a inibição de outras SVMP que HSF não é capaz de inibir isoladamente (como descrito no capítulo 1, seção 1.3.1).

Na etapa final da análise, verificamos todos os peptídeos alinhados com a entrada correspondente ao inibidor BJ46a e construímos uma tabela descrevendo os alinhamentos e sua respectiva contagem espectral, com o objetivo de identificar mutações pontuais na estrutura primária. Os peptídeos com maiores contagens espectrais (*De novo* Seq) em cada trecho da sequência de BJ46a (DBSeq) tiveram seus espectros manualmente validados, gerando os mapas de cobertura ilustrados na **Figura 19** e na **Figura 20**.



**Figura 19: Mapa do alinhamento dos peptídeos sequenciados *de novo* de BJ46a com a sequência depositada no UNIPROT.**

Nesta representação, a sequência de BJ46a está representada como uma linha horizontal na parte inferior do quadro. Cada peptídeo alinhado com a sequência está representado como uma barra horizontal azul. No eixo esquerdo da figura, encontram-se peptídeos representativos de cada grade horizontal do gráfico. Abaixo de cada peptídeo, encontra-se a numeração correspondente ao resíduo inicial e ao final, além da localização deste peptídeo no 1º domínio cistatina (C1) ou no 2º domínio cistatina (C2).



**Figura 20: Cobertura de sequência do inibidor BJ46a.**

Todos os resíduos marcados em azul foram confirmados por pelo menos 1 peptídeo sequenciado. Os resíduos em preto não foram detectados na análise.

Em nossas análises, os peptídeos correspondentes à sequência canônica do inibidor BJ46a foram os que apresentaram as maiores contagens espectrais. Não identificamos peptídeos correspondentes aos resíduos 1-19, região que já havia sido descrita previamente como processada após a síntese do inibidor (Valente *et al.*, 2001). Além disso, verificamos que não há peptídeos alinhados para os 8 últimos resíduos, pertencentes à região rica em histidina de BJ46a na abordagem *de novo* (Figura 20).

Para verificarmos se ocorreu uma eventual deleção destes resíduos na sequência de BJ46a, confrontamos este dado com os resultados obtidos a partir da análise da mesma amostra por uma abordagem PSM (*peptide-spectrum match*) utilizando as ferramentas PEAKS DB do *software* PEAKS studio. Nesta abordagem, identificamos 4 peptídeos para a porção C-terminal de BJ46a, em um total de 777 espectros (Figura 21).

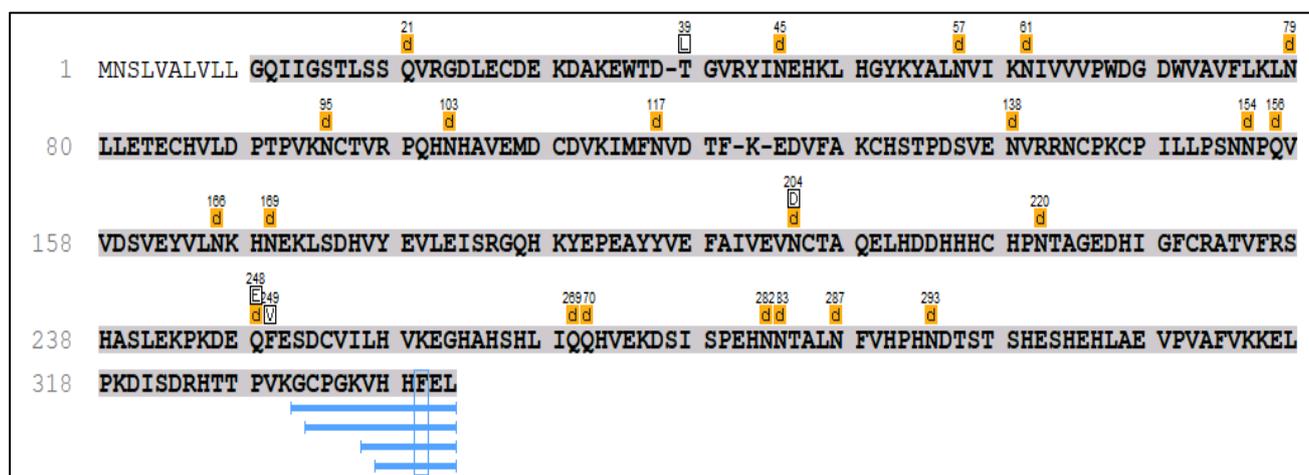


Figura 21: Cobertura de sequência para BJ46a utilizando as ferramentas PEAKS DB (abordagem PSM).

As barras em azul representam os peptídeos identificados para a região C-terminal de BJ46a.

Por fim, buscamos eventuais mutações na sequência de BJ46a no conjunto de peptídeos de maior contagem espectral alinhados com a sequência de BJ46a. Verificamos um conjunto de peptídeos que apresentaram uma mutação na posição 230 (F → V), com uma contagem espectral total de 144 espectros, que pode indicar uma mudança na estrutura primária de BJ46a. A troca ocorreu por resíduo de propriedades físico-químicas distintas: fenilalanina é um aminoácido com cadeia lateral aromática, enquanto que valina é um aminoácido apolar, de cadeia lateral alifática. Este tipo de mutação pontual é denominado

mutação *missense* e estas trocas são peças-chave na evolução molecular de proteínas (Shanthirabalan *et al.*, 2018). Evidências na literatura indicam que 50 a 80% dos resíduos em uma proteína podem ser modificados sem alterar significativamente a estrutura proteica global. Ainda assim, algumas mutações *missense* podem alterar dramaticamente a função proteica, principalmente as trocas em regiões catalíticas e em interfaces de interação, que dependem de uma estrutura tridimensional conservada (Schaefer e Rost, 2012).

Entretanto, a predição do impacto de mutações *missense* não é trivial. Há na literatura um conjunto recente de algoritmos capazes de avaliar a influência de uma dada troca, estimando-se a variação da energia livre de Gibbs ( $\Delta\Delta G$ ), tendo por referência um banco de dados, Protherm, composto de estruturas pertencentes ao PDB associadas a cálculos de  $\Delta\Delta G$ , mas estes algoritmos ainda sofrem com algumas limitações, uma vez que a comparação entre variações preditas de energia com o dado experimental pode levar a discrepâncias significativas (Shanthirabalan *et al.*, 2018). Para BJ46a, a inexistência de proteínas depositadas no PDB com uma identidade aceitável dificulta o uso destes algoritmos de predição; tal análise só seria possível por medições experimentais, como na técnica de calorimetria de titulação isotérmica (ITC) (Leavitt e Freire, 2001). Portanto, com os dados disponíveis, ainda não é possível inferir os impactos desta mutação na atividade biológica de BJ46a.

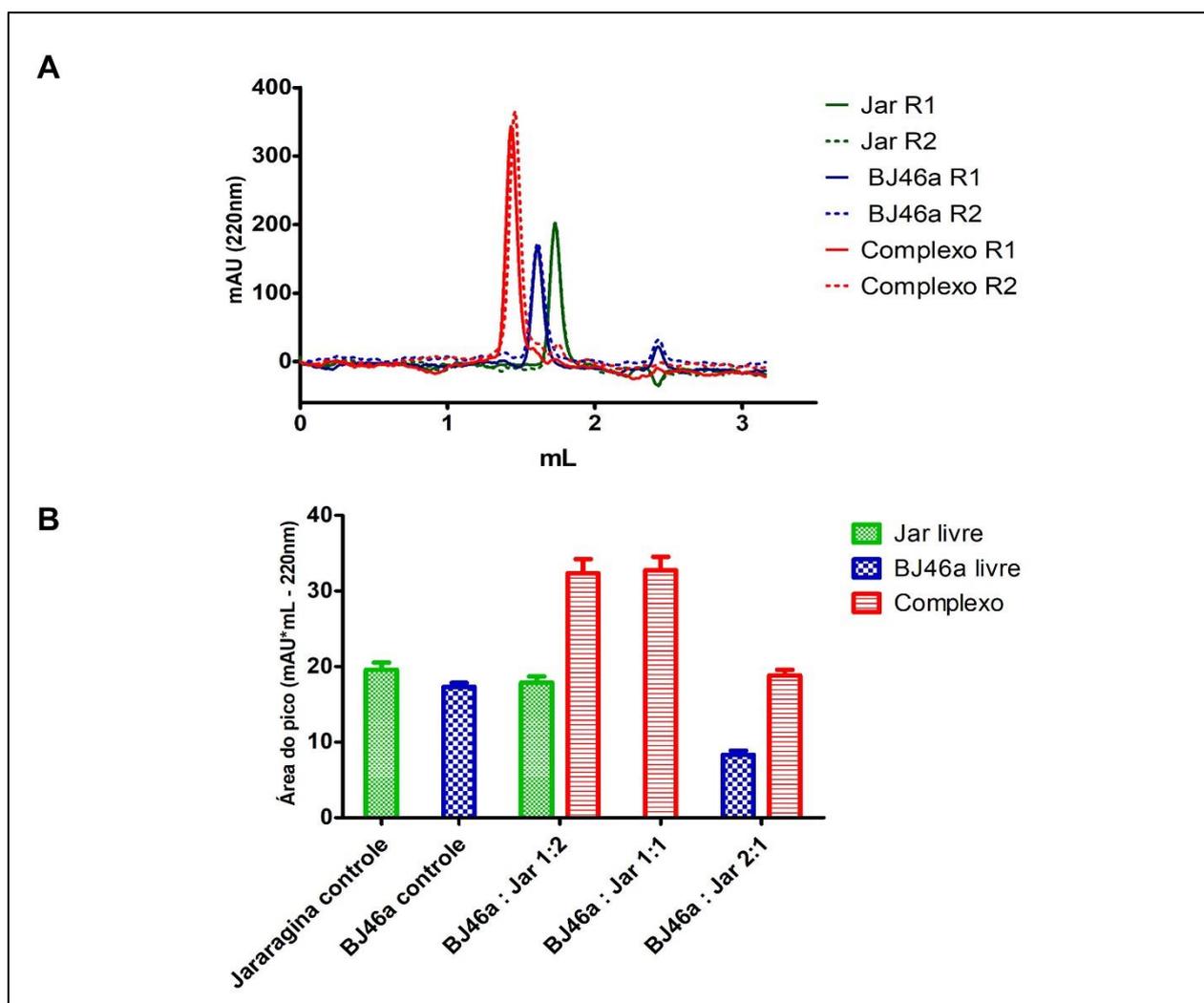
## 4.2 – ENSAIOS DE INTERAÇÃO ENTRE O INIBIDOR BJ46a E A METALOENDOPEPTIDASE JARARAGINA

Para o mapeamento das regiões de interação entre BJ46a e jararagina pelas técnicas de XL-MS e HDX-MS, a determinação da estequiometria da interação inibidor-toxina é fundamental. Em HDX-MS, as mudanças no perfil de deuteração da proteína de interesse são determinadas a partir da comparação da incorporação de deutério para as proteínas isoladas e para as proteínas em complexo. Este tipo de análise é mais confiável quando o complexo proteína-proteína representa majoritariamente a população contida na amostra (Gallagher e Hudgens, 2016; Kochert *et al.*, 2018). Para XL-MS, o excesso de proteína livre em relação ao complexo prejudica a identificação e o assinalamento de *cross-links* oriundos do complexo entre elas (Sinz, 2007).

A interação entre o inibidor BJ46a e a metaloendopeptidase jararagina foi avaliada por cromatografia de exclusão molecular em coluna Superdex 200 Increase 5/150, como descrito no item 3.2.12 do capítulo 3. Três relações molares de subunidade do inibidor: toxina (1:2, 1:1 e 1:0,5) foram incubadas por 15 minutos a 25°C em tampão HEPES 64 mM + CaCl<sub>2</sub> 10 mM pH 7,5 e injetadas na coluna previamente equilibrada com o mesmo tampão. A **Figura 22** demonstra os perfis cromatográficos dos controles de BJ46a e jararagina livres, bem como o resultado da integração das áreas dos picos obtidos para os controles e para cada relação molar testada.

Ao incubarmos BJ46a e jararagina na relação molar de 1:2 (inibidor:toxina), verificamos a presença de um pico de retenção e área cromatográfica distintas das registradas para as amostras controles, e que corresponde ao complexo formado entre inibidor e toxina; porém, a presença de um pico com retenção idêntica à registrada para o controle de jararagina livre indica que a saturação da enzima não foi alcançada nesta relação molar. Da mesma forma, a presença de um pico de retenção compatível com o controle de BJ46a livre na relação molar de 1:0,5 (inibidor:toxina) indica que esta relação molar não corresponde a estequiometria do complexo. Porém, para a relação molar de 1:1, não verificamos no perfil cromatográfico picos significativos que correspondessem ao excesso de BJ46a ou de jararagina livres, indicando que a saturação foi dada na estequiometria de 1:1 (painéis A e B, **Figura 22**).

Com base neste resultado, realizamos uma diluição seriada do complexo de BJ46a-jararagina na estequiometria de 1:1 (subunidade do inibidor:toxina) como prova de conceito para os experimentos de HDX-MS. A deuteração de uma amostra de interesse é feita diluindo-se a amostra, em tampão preparado com H<sub>2</sub>O, com o tampão deuterado correspondente, de modo que a concentração final de D<sub>2</sub>O corresponda a 80 – 95% do volume final (Gallagher e Hudgens, 2016). Para obter dados confiáveis de incorporação isotópica para um complexo proteico, é necessário então que haja a prevalência da forma complexada após a diluição da amostra (Kochert *et al.*, 2018). Preparamos uma solução-estoque de complexo em tampão aquoso a 7 µM e incubamos por 15 minutos a 25°C sob agitação moderada. Após este período, diluímos esta solução para as concentrações finais de 3,5 µM, 1,75 µM e 0,875 µM, como demonstra a **Figura 23**.

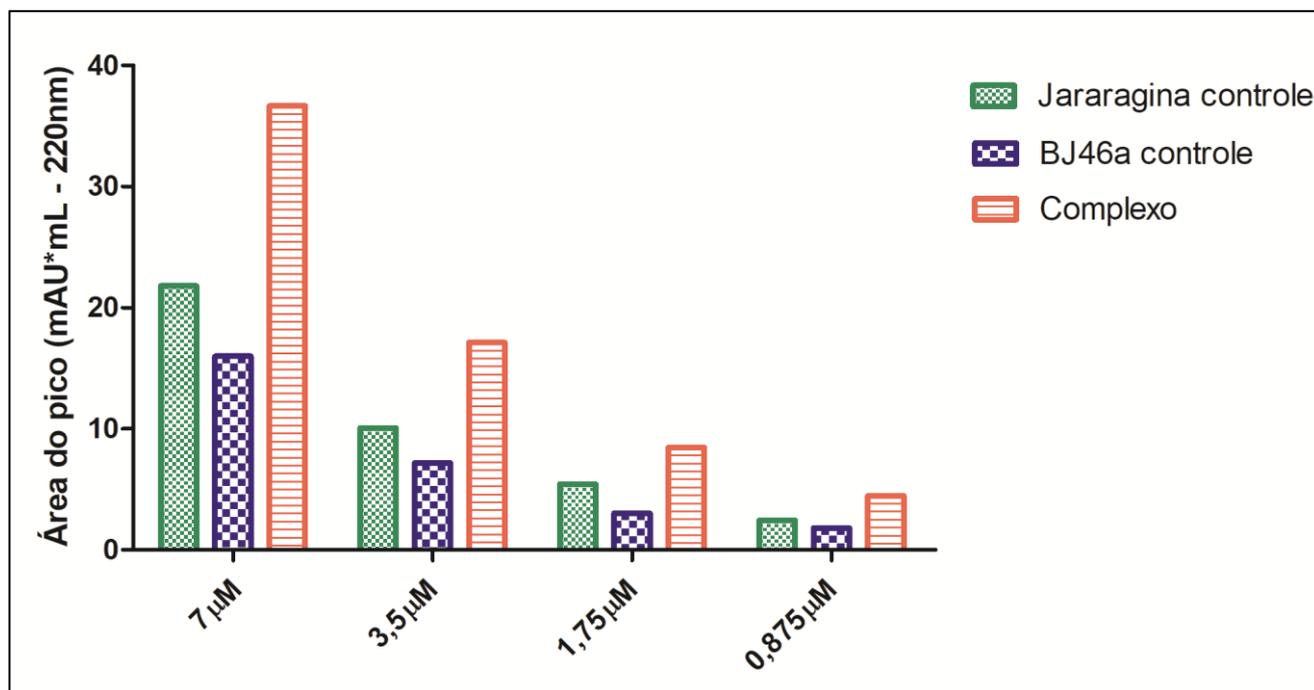


**Figura 22: Experimento de formação de complexo entre o inibidor BJ46a e a metaloendopeptidase jararagina.**

**Painel A:** Perfis cromatográficos das corridas referentes à BJ46a controle, à jararagina controle, e ao complexo BJ46a – jararagina na relação molar de 1:1. R1: replicata experimental 1; R2: replicata experimental 2.

**Painel B:** Gráfico das integrações (média e desvio-padrão) das áreas dos picos detectados a 220nm para os controles de jararagina e BJ46a livres e para cada relação molar analisada.

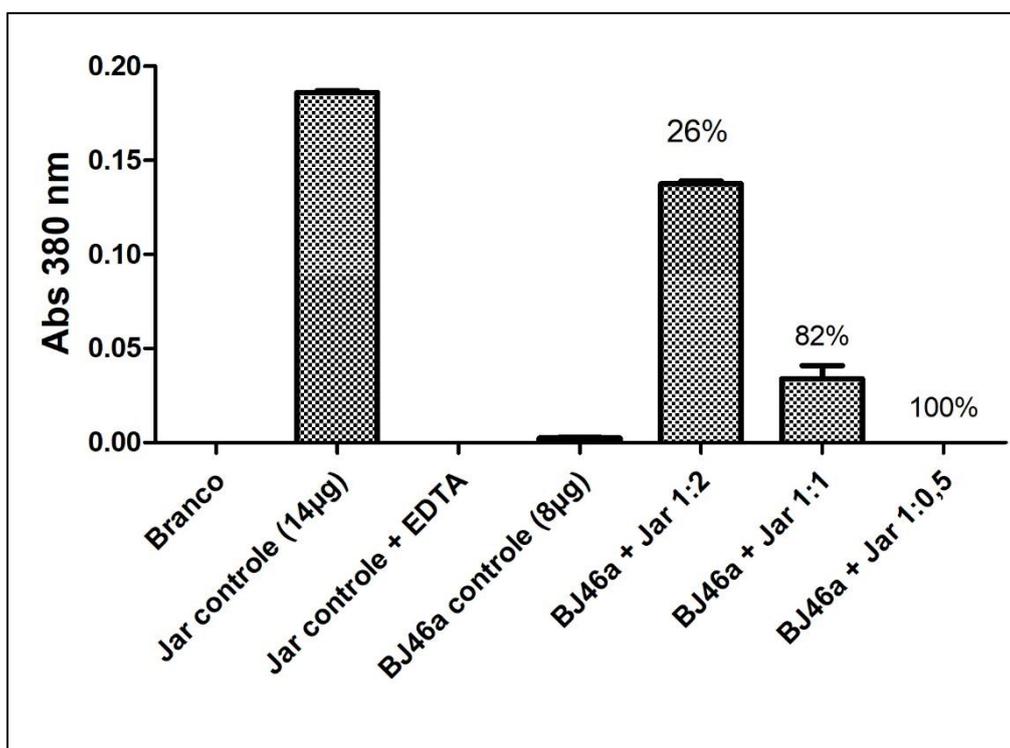
Em todas as diluições testadas, o complexo BJ46a – jararagina se manteve, com uma área do pico compatível à soma das áreas de BJ46a e jararagina controles, nas mesmas concentrações proteicas. Este resultado está em concordância com análises anteriores de ressonância plasmônica de superfície, que indicaram uma constante de dissociação ( $K_D$ ) muito baixa para a interação BJ46a – jararagina, na ordem de 13,8 nM (capítulo 1, p.16).



**Figura 23: Ensaio de diluição seriada do complexo BJ46a - jararagina como prova de conceito para os experimentos de HDX-MS.**

Neste gráfico, estão representadas as integrações das áreas (em mAU \* mL) dos picos detectados a 220 nm, tanto nos cromatogramas dos controles de BJ46a e jararagina livres, quanto nos cromatogramas do complexo inibidor-toxina (relação molar 1:1), em diferentes concentrações molares.

Com relação à inibição da atividade catalítica de jararagina por BJ46a, os dados indicam que BJ46a foi capaz de inibir apenas parcialmente a atividade proteolítica de jararagina sobre azocaseína na relação molar de 1:2 (inibidor:toxina) (**Figura 24**). Na proporção molar de 1:1, ocorreu inibição significativa da atividade da toxina, enquanto que na proporção 1:0,5, ocorreu a inibição total da proteólise de azocaseína por jararagina. Estes resultados estão compatíveis com o demonstrado no ensaio anterior, onde não há a presença de jararagina livre nas relações de 1:1 e 1:0,5 (inibidor:toxina).



**Figura 24: Ensaio de inibição de atividade azocaseinolítica de jararagina por BJ46a.**

As porcentagens de inibição em cada relação toxina:inibidor foram calculadas relativas ao controle de atividade (Jararagina 14µg).

É importante ressaltar que a saturação do complexo BJ46a-jararagina ocorreu em uma relação molar diferente da reportada na literatura. No modelo proposto por Valente e colaboradores, BJ46a é considerada um homodímero que, no momento da interação com jararagina, sofre dissociação e cada monômero consegue interagir com duas moléculas de jararagina, resultando em uma estequiometria de 1:2 (subunidade do inibidor:toxina) (Valente *et al.*, 2001). Por outro lado, nossos dados atuais são compatíveis com o descrito para HSF, para a qual experimentos de formação de complexo analisados também por cromatografia de exclusão molecular indicaram a interação em uma relação molar de 1:1 com a SVMP brevilisina H6 isolada do veneno de *Gloydius halys brevicaudus* (Deshimaru *et al.*, 2005). Os resultados da purificação do inibidor BJ46a por cromatografia de imunoafinidade, apresentados nesta tese, indicam a presença de diferentes proteoformas do inibidor no plasma, e trabalhos recentes demonstraram a presença de BJ46a no proteoma e no transcriptoma das glândulas de veneno (principal e acessória) de *B. jararaca*, levantando a hipótese de que possa haver diferentes proteoformas de BJ46 sendo sintetizadas na serpente,

cada uma exibindo diferentes afinidades de ligação com SVMP (Luna *et al.*, 2013; Valente *et al.*, 2018). Os dados de sequenciamento *de novo* apontaram uma única modificação pontual, no resíduo 230. Como discutido anteriormente, ainda não é possível inferir os impactos desta mutação, e se esta seria suficiente para justificar uma alteração na relação estequiométrica da interação entre a BJ46a com a mutação pontual detectada e a jararagina. Para os experimentos subsequentes de XL-MS e HDX-MS, assumimos a relação estequiométrica de 1:1 evidenciada pelos experimentos de formação de complexos por exclusão molecular.

### 4.3 – MAPEAMENTO DE RESTRIÇÕES DE DISTÂNCIA PARA BJ46A, JARARAGINA E O COMPLEXO BJ46A - JARARAGINA

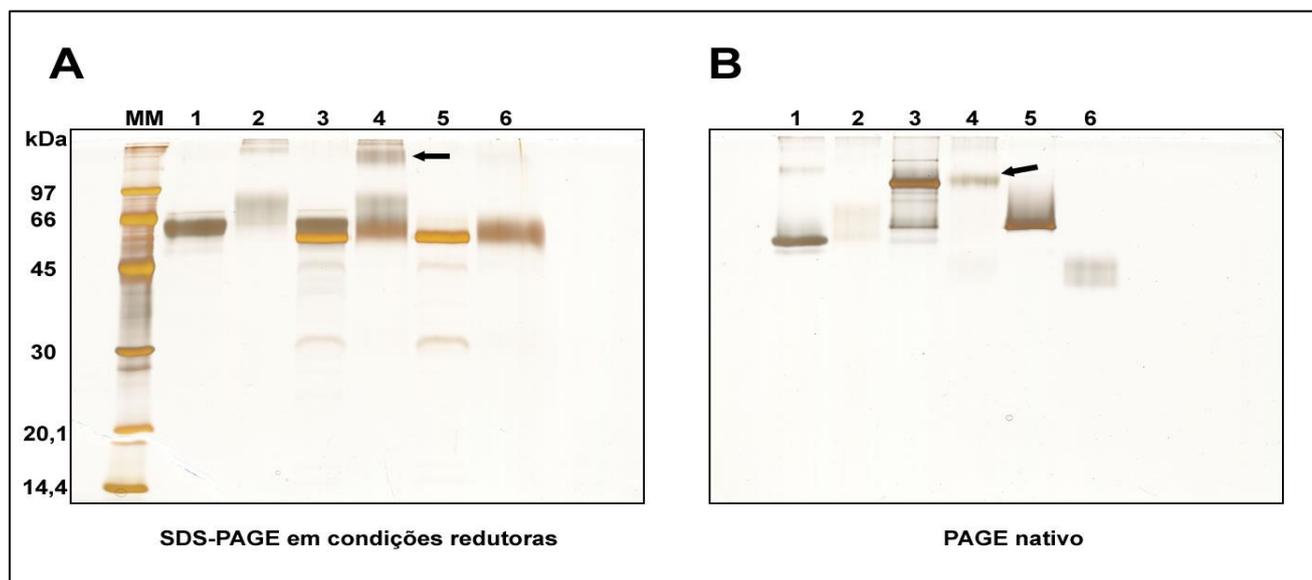
Para a primeira etapa do estudo estrutural de BJ46a, jararagina e o complexo BJ46a – jararagina, empregamos a técnica de XL-MS para derivar informações sobre restrições de distância para as proteínas isoladas e em complexo. Além disso, realizamos a determinação experimental das pontes dissulfeto em BJ46a como um conjunto adicional de restrições espaciais.

Na técnica de XL-MS, utilizamos o ALC BS<sup>3</sup> e a metodologia XPlex publicada recentemente (Fioramonte *et al.*, 2018). Como discutido anteriormente neste trabalho (capítulo 1, seção 1.4.1.1), o ALC BS<sup>3</sup> (bis (sulfosuccinimidil) suberato) apresenta em cada uma das extremidades da molécula, o grupo funcional N-hidroxisuccinimida (NHS), que apresenta reatividade preferencial para o grupamento  $\epsilon$ -amino de lisina e para o grupamento amino N-terminal; resíduos de serina também podem ser conectados covalentemente pelo ALC (Madler *et al.*, 2009). Os grupamentos NHS do BS<sup>3</sup> são conectados por um braço espaçador, uma cadeia carbônica de 11,4 Å (Paramelle *et al.*, 2013). Somando-se o tamanho da cadeia espaçadora com o tamanho da cadeia lateral de lisina (6,4 Å), o limite máximo de distância entre dois resíduos de lisina, medido a partir do carbono alfa do *backbone* proteico, é de aproximadamente 24 Å. Entretanto, simulações de dinâmica molecular indicaram que este limite pode ser um pouco maior, entre 26 e 30 Å, devido à flexibilidade inerente das cadeias laterais (Merkley *et al.*, 2014). Dessa forma, resíduos que estejam localizados em uma distância de até 30 Å poderiam ser conectados covalentemente pelo ALC BS<sup>3</sup>.

A metodologia XPlex tem como alvos os resíduos ácidos (ácido aspártico e ácido glutâmico) bem como os básicos (lisina e serina) e as porções N- e C-terminais. A reação ocorre em condições brandas e subdivide-se em duas etapas principais: inicialmente, ocorre a ativação dos ácidos carboxílicos presentes nas cadeias laterais D/E e na extremidade C-terminal pelo reagente N-(3-dietilaminopropil)-N'-etilcarbodiimida (EDC) na presença de N-hidroxibenzotriazol (HOBt), seguido de reação com uma diamina (1,6-hexametilenodiamina), formando ligação amida entre os resíduos ácidos ativados e os grupamentos amina da diamina. Há também a possibilidade da formação de links *zero-length* (sem braço espaçador), nos quais resíduos ácidos ativados podem reagir diretamente com as cadeias laterais de lisina

ou serina (Fioramonte, 2016). Esta reatividade multiplexada (*links* ácido-ácido e ácido-básico) pode cobrir, em média, 24% dos resíduos em uma dada proteína e aumenta significativamente a cobertura de sequência, uma vez que o número de *cross-links* possíveis aumenta quadraticamente com o número de resíduos reativos para a química de *cross-linking* (Fioramonte *et al.*, 2018). A complementariedade entre o ALC BS<sup>3</sup> e a abordagem XPLex pode fornecer restrições espaciais importantes para entendermos a topologia do complexo formado entre o inibidor BJ46a e a SVMP jararagina e para a filtragem de modelos tridimensionais mais compatíveis com os dados experimentais.

Para a abordagem com o ALC BS<sup>3</sup>, o complexo entre BJ46a e jararagina foi estabilizado por 3 h a 25°C sob agitação branda, utilizando um excesso em massa de BS<sup>3</sup> sobre proteína de 20 vezes. Após este período, a reação foi interrompida pela adição de uma solução de bicarbonato de amônio 500 mM (concentração final de bicarbonato de amônio: 20mM). A eficiência da reação foi analisada por eletroforese em condições nativas e em condições desnaturantes e redutoras, como demonstra a **Figura 25**, a seguir.

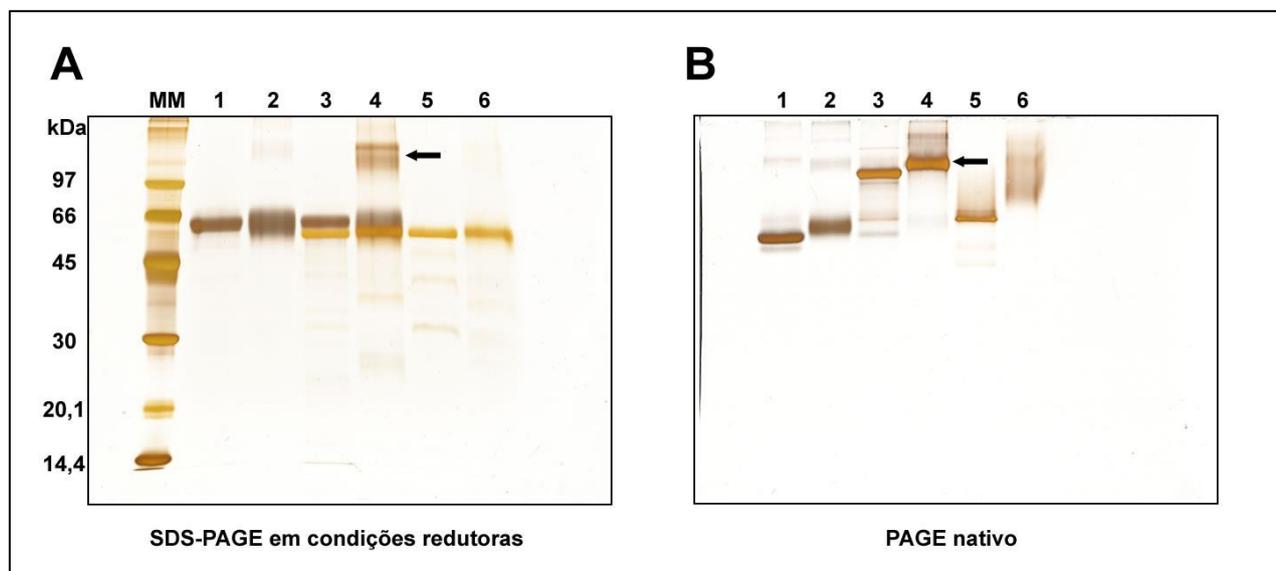


**Figura 25: Estabilização do complexo entre o inibidor BJ46a e jararagina usando o ALC BS<sup>3</sup>.**

Amostra 1: BJ46a controle (1 µg); 2: BJ46a estabilizada com BS<sup>3</sup>; 3: Complexo BJ46a – jararagina; 4: Complexo BJ46a - jararagina estabilizado com BS<sup>3</sup> (1,9 µg); 5: jararagina controle (0,9 µg); 6: jararagina estabilizada com BS<sup>3</sup>. Painel A: eletroforese em condições desnaturantes e redutoras. Painel B: eletroforese em condições nativas. As bandas referentes ao complexo BJ46a – jararagina estabilizado estão apontadas com uma seta. Géis revelados por impregnação por prata.

Quando submetemos BJ46a e jararagina à eletroforese em condições nativas (**Figura 25**, painel B) é possível visualizar a ocorrência de uma nova banda na amostra 3, de menor migração eletroforética em relação à banda de BJ46a controle (amostra 1) e de jararagina controle (amostra 5), sugerindo a formação de complexo entre duas proteínas. A mesma banda é visível quando incubamos BJ46a e jararagina na presença do ALC BS<sup>3</sup>, com uma aparência mais difusa (amostra 4). Entretanto, quando submetemos a mesma amostra à eletroforese em condições desnaturantes e redutoras (**Figura 25**, painel A), a banda referente ao complexo inibidor-toxina não é mais visualizada (amostra 3), indicando que esta interação é mantida por forças de natureza não-covalente, como previamente descrito na literatura (Valente *et al.*, 2001). Quando este mesmo complexo é estabilizado com BS<sup>3</sup>, podemos verificar a presença de uma banda de alta massa molecular (> 97 kDa, raia 4, indicada por seta) que não está presente nos controles do inibidor e toxina livres estabilizados por BS<sup>3</sup>, (amostras 2 e 6), indicando a estabilização covalente do complexo pelo ALC. Para as amostras de BJ46a e jararagina estabilizadas por BS<sup>3</sup> (amostras 2 e 6), verificamos que as bandas referentes ao inibidor e à toxina apresentaram uma mobilidade eletroforética ligeiramente menor do que quando comparados aos controles de BJ46a e jararagina sem estabilização (amostras 1 e 5), além de uma aparência difusa, o que sugere a formação de *intra-links*, isto é, *cross-links* formados dentro da própria proteína.

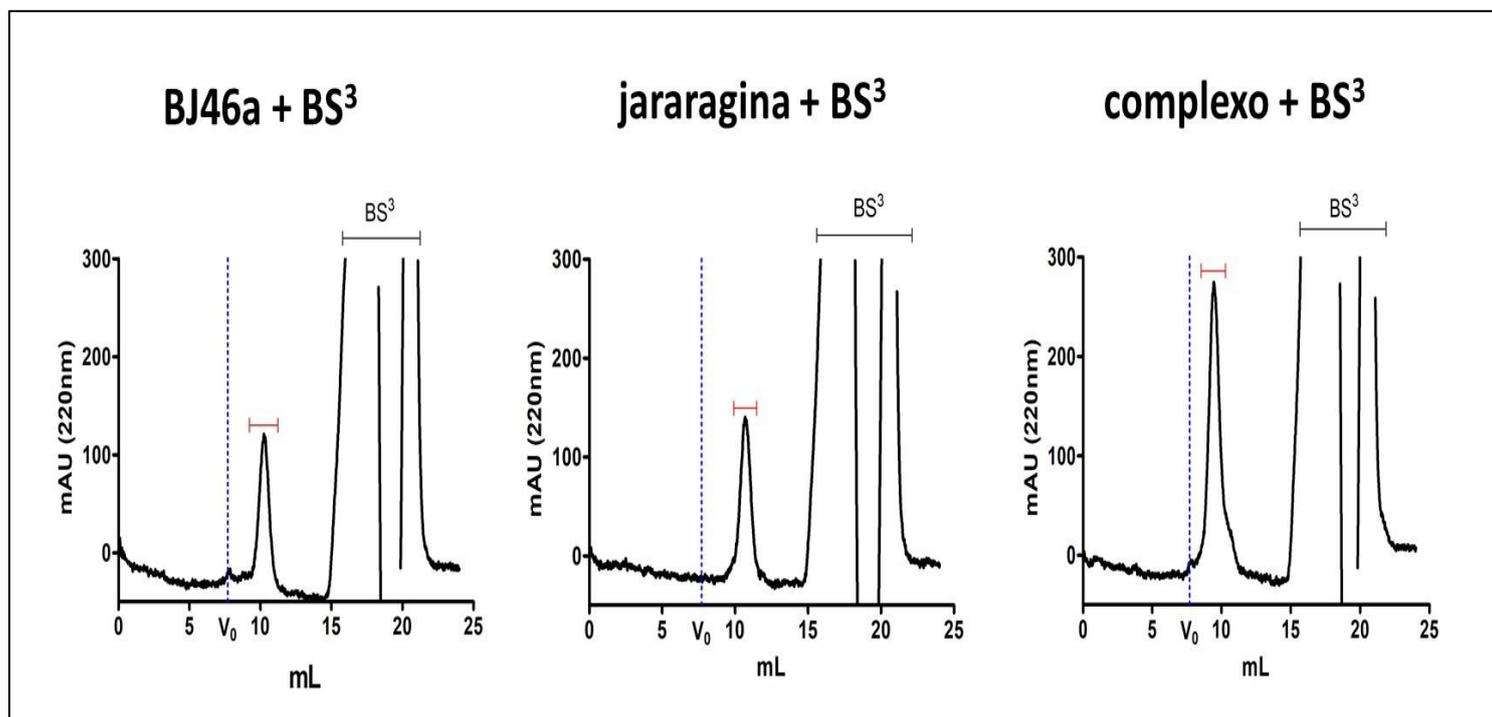
Para a abordagem XPlex, verificamos um resultado similar ao observado com o ALC BS<sup>3</sup>, onde ocorreu a estabilização covalente tanto do complexo BJ46a – jararagina, quanto do inibidor e da toxina em suas formas livres (**Figura 26**), com algumas diferenças. Para BJ46a e jararagina estabilizadas com XPlex (**Figura 26**, painel A, amostras 2 e 6) as bandas mostraram-se menos difusas do que o observado com a estabilização pelo ALC BS<sup>3</sup>; com relação ao complexo BJ46a-jararagina, evidenciamos pelo menos duas bandas, uma majoritária e outra minoritária, com mobilidade eletroforética superior a 97 kDa na amostra do complexo estabilizada com XPlex (painel A, amostra 4), o que pode refletir subpopulações do complexo estabilizado.



**Figura 26: Estabilização covalente do complexo BJ46a-jararagina pela metodologia XPlex.**

Amostra 1: BJ46a controle (2  $\mu\text{g}$ ); 2: BJ46a estabilizada por XPlex; 3: Complexo BJ46a – jararagina; 4: Complexo BJ46a - jararagina estabilizado por XPlex (3,74  $\mu\text{g}$ ); 5: jararagina controle (1,74  $\mu\text{g}$ ); 6: jararagina estabilizada por XPlex. Painel A: eletroforese em condições desnaturantes e redutoras; a banda correspondente ao complexo estabilizado entre BJ46a e jararagina encontra-se apontada por uma seta. Painel B: eletroforese em condições nativas. Géis corados por impregnação por prata.

Posteriormente à análise da eficiência de reação de *cross-linking* por  $\text{BS}^3$  e XPlex por eletroforese, partimos para a etapa de purificação do complexo estabilizado por  $\text{BS}^3$  por cromatografia de exclusão molecular, com o objetivo de remover o excesso de  $\text{BS}^3$  que não reagiu com a amostra e monitorar a eventual ocorrência de oligomerização artefactual das proteínas ligadas cruzadamente. Os cromatogramas obtidos para cada amostra estão representados na **Figura 27**. Para a abordagem XPlex, a reação foi encerrada por meio da precipitação das proteínas com uma mistura de etanol/acetona (descrito no item 3.2.16, capítulo 3).



**Figura 27: Purificação das amostras estabilizadas por BS<sup>3</sup> por cromatografia de exclusão molecular.**

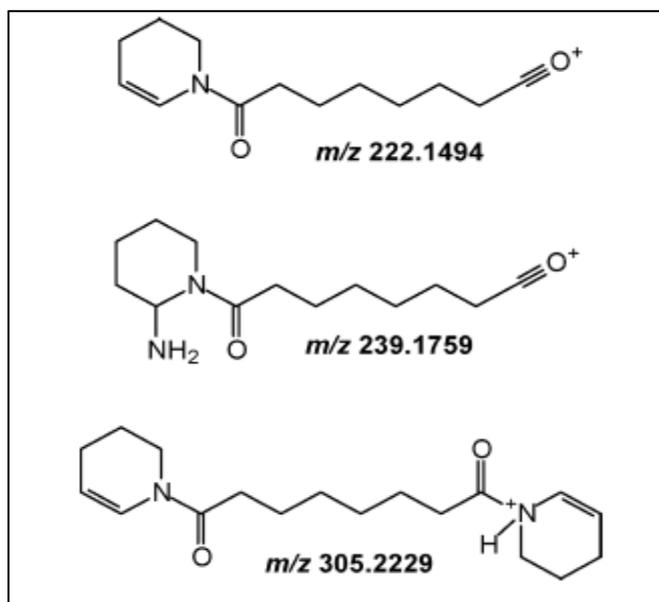
Os picos marcados por uma barra vermelha foram coletados e posteriormente hidrolisados por tripsina e submetidos à espectrometria de massas. O volume de exclusão da coluna ( $V_0$ , 7,7 mL) está representado com uma linha vertical azul.

Após a purificação, as amostras foram desnaturadas, reduzidas, alquiladas e hidrolisadas em solução por tripsina e analisadas posteriormente por nLC-nESI-HCD-MS/MS, em espectrômetro Q-Exactive Plus (Thermo Scientific). Os dados foram posteriormente analisados com auxílio do software *Spectrum Identification Machine for Crosslinked Peptides* (SIM-XL). A aquisição e a identificação de espectros de fragmentação oriundos de amostras de *cross-linking* apresentam algumas limitações. Com relação à aquisição dos dados de MS/MS, a dificuldade principal reside na diferença entre a quantidade de peptídeos lineares, majoritários na amostra, e peptídeos estabilizados covalentemente pelo ALC, presentes em quantidades subestequiométricas (Iglesias *et al.*, 2009). Além disso, os ALC que reagem com aminas primárias em resíduos de lisina (BS<sup>3</sup>/DSS) aumentam a taxa de clivagens perdidas, uma vez que a enzima tripsina, comumente empregada, não cliva na extremidade C-terminal de resíduos de lisina modificados. Dessa forma, os peptídeos ligados cruzadamente são íons-precusores de maior  $m/z$  e a fragmentação em torno da região do *linker* é desfavorecida pois a livre movimentação do próton móvel durante

a fragmentação é comprometida. Dessa forma, a fragmentação por HCD é geralmente o método de escolha para a geração de espectros de fragmentação em experimentos de XL-MS (Sinz, 2007).

A fragmentação de peptídeos ligados cruzadamente por HCD é similar ao descrito para peptídeos cíclicos para *cross-links* do tipo 1 (*intralink*) e do tipo 2 (*interlink*). Para *cross-links* do tipo 0 (*dead-end cross-links*), a fragmentação é análoga à de um peptídeo linear, com uma modificação química pontual (Schilling *et al.*, 2003). Em espectros MS/MS de amostras de *cross-linking* gerados pelo método de fragmentação CID, há o predomínio de íons de série b (N → C-terminal) e de série y (C → N-terminal). Dessa forma, em um espectro MS/MS de *cross-links* de tipo 1 e 2, temos as séries b/y de cada peptídeo covalentemente ligado em um mesmo espectro MS/MS (Iglesias *et al.*, 2009). O peptídeo que contém o maior número de resíduos e/ou de maior massa é denominado peptídeo  $\alpha$ , enquanto que o menor peptídeo (em número de resíduos e/ou de menor massa) é chamado de peptídeo  $\beta$ . Assim, cada íon-fragmento é assinalado no espectro MS/MS de acordo com sua série, posição e peptídeo a qual pertence (e.g.  $y1\beta$ ,  $b2\alpha$ ) (Schilling *et al.*, 2003).

É importante ressaltar que nos espectros de fragmentação de peptídeos covalentemente estabilizados por BS<sup>3</sup>/DSS há pelo menos um dos íons-repórter ( $m/z$  222,1 ou 239,2 ou 305,2) oriundos do rearranjo da cadeia lateral de resíduos de lisina e da cadeia espaçadora do *cross-linker*, suas estruturas estão representadas na **Figura 28** (Santos *et al.*, 2011). Estes valores de  $m/z$  não são conflitantes com nenhum valor de  $m/z$  esperado para íons-fragmento de séries b ou y derivados de qualquer aminoácido natural e, portanto, podem ser utilizados como íons-repórter, indicando a ocorrência da reação de *cross-linking*. Além disso, a presença destes íons pode fornecer informações importantes sobre a reação em um dado espectro: a presença dos íons de  $m/z$  222,1 e 239,2 indica efetivamente a ocorrência da reação de *cross-linking*; a presença do íon  $m/z$  305,2 demonstra que o espectro é proveniente de uma espécie intra- ou inter-molecular, na qual o ALC ligou covalentemente duas lisinas. Os espectros referentes a espécies *dead-end* não apresentam o íon  $m/z$  305,2 (Santos *et al.*, 2011).



**Figura 28:** Estrutura dos íons-repórter detectados em espectros MS/MS de peptídeos covalentemente estabilizados com o ALC BS<sup>3</sup>.

Reproduzido de Santos *et al.* 2011.

Em XL-MS, a análise criteriosa dos espectros de fragmentação é crítica para o sucesso da técnica. Portanto, a análise dos *cross-links* foi realizada em três etapas: na primeira, utilizamos o software *Spectrum Identification Machine for Crosslinked Peptides* (SIM-XL) para efetuar a busca dos peptídeos covalentemente ligados tanto nas amostras controle (BJ46a livre e jararagina livre) quanto na amostra referente ao complexo BJ46a-jararagina. Em seguida, realizamos uma filtragem automática utilizando os parâmetros do programa SIM-XL, como o *score* primário ( $\geq 2,5$ ), erro ( $\leq 10$  ppm), contagem espectral ( $\geq 1$ ) e número de íons-fragmento detectados por cadeia ( $\geq 3$ ). Por fim, os espectros que passaram no filtro automático do SIM-XL foram manualmente avaliados, com base nos seguintes critérios: 1) presença de um ou mais íons-repórter (apenas em espectros oriundos da abordagem com o ALC BS<sup>3</sup>); 2) relação sinal-ruído; 3) cobertura de íons-fragmento para os peptídeos  $\alpha$  e  $\beta$ ; 4) a distribuição de erro dos íons-fragmento observados no espectro e 5) a posição inequívoca do *linker*, isto é, se os íons-fragmento detectados referentes aos peptídeos não permitiam outro tipo de interpretação do espectro que não a reportada. Cada espectro foi por fim classificado em três grupos diferentes: espectros considerados excelentes, que satisfazem todos os critérios de qualidade descritos, espectros bons,

que satisfazem 4 dos 5 critérios, e espectros de qualidade média, que satisfazem pelo menos 3 critérios (Iacobucci e Sinz, 2017).

#### 4.3.1 – Análise dos espectros e mapas de restrição de distância de BJ46a e jararagina livres

Após a filtragem dos espectros obtidos, conseguimos identificar, para BJ46a livre, 20 *cross-links* utilizando o ALC BS<sup>3</sup>; para a abordagem XPlex, identificamos 15 *cross-links* entre resíduos ácidos (D/E, diamina) e 135 *cross-links zero-length* entre resíduos ácidos e básicos (D-S, D-K, E-S, E-K). A **Figura 29** e a **Figura 30** demonstram espectros representativos para cada abordagem nas categorias “excelente”, “bom” e “médio”. Tais espectros foram utilizados para a geração dos mapas bidimensionais de restrições de distância, representados na **Figura 31**, e das listas contendo todos os *cross-links* filtrados.

Analisando-se os mapas bidimensionais gerados para as duas abordagens podemos delinear quatro grandes classes de *links* encontrados na estrutura de BJ46a: 1) *cross-links* que conectam o primeiro domínio cistatina, localizado na extremidade N-terminal, com o domínio rico em histidina na porção C-terminal do inibidor; 2) *cross-links* localizados no interior de cada um dos domínios cistatina; 3) *cross-links* inter domínios cistatina e, por fim, 4) *cross-links* que ligam o segundo domínio cistatina ao domínio rico em histidina de BJ46a. Para o domínio rico em histidina, mesmo na presença de resíduos básicos passíveis de estabilização (**Figura 31**, traços pontilhados no interior da barra), não verificamos a presença de *cross-links* do tipo 2 (interpeptídeo) com o ALC BS<sup>3</sup>, porém há um grupo de 4 *dead-end cross-links* dentro do domínio, nos resíduos K296, K300, K311 e K316. Na abordagem XPlex, para a mesma região, detectamos 4 *cross-links zero-length*, entre os resíduos E241-S246, D301- K311, K296 – E321, D301 – K311 e K311 – E321.

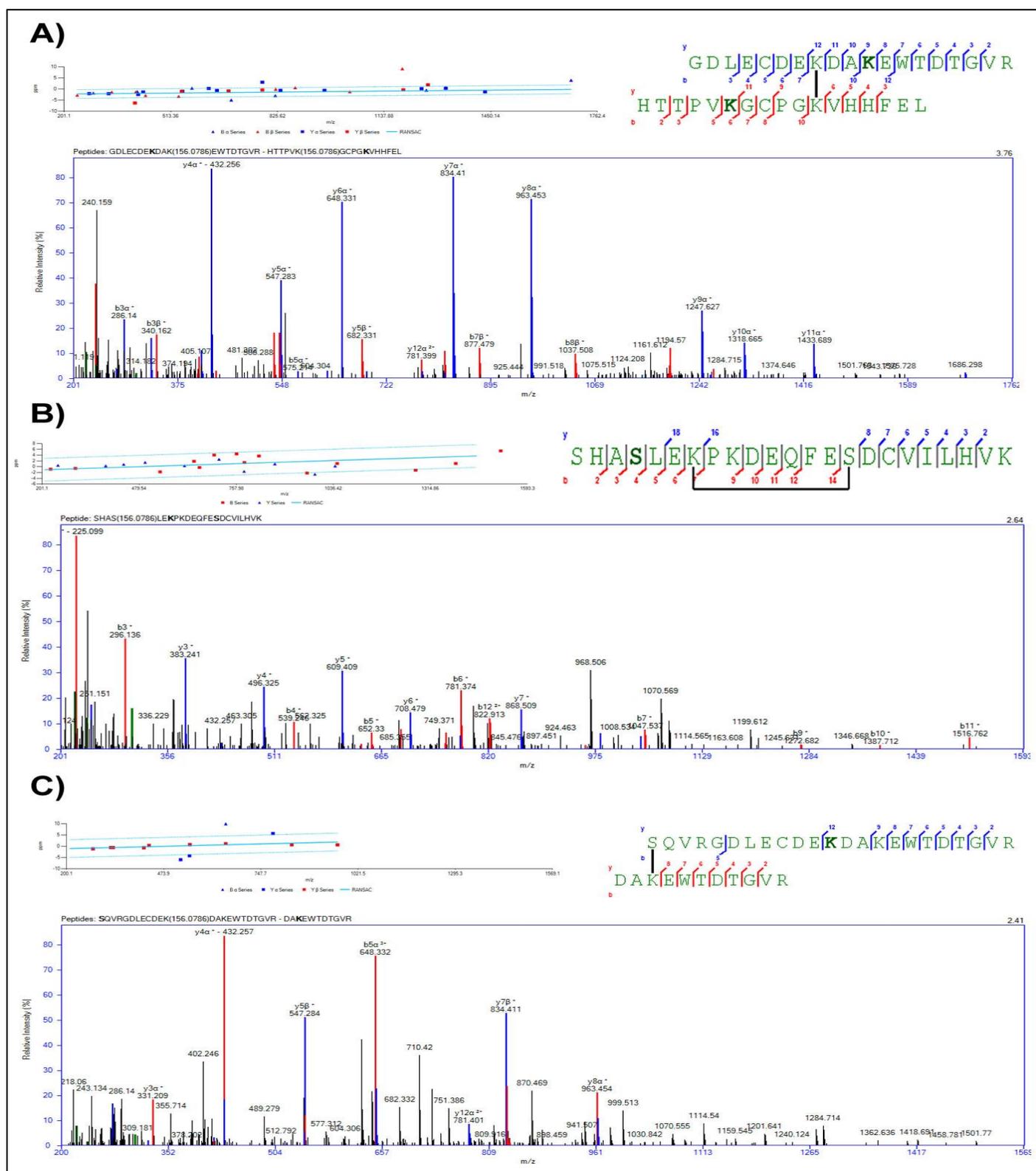
Explorando-se um pouco mais a abordagem XPlex, a **Figura 32** demonstra a distribuição dos *cross-links* ácido-ácido (D-D, D-E, E-E) e ácido-básico (D-K, D-S, E-K, E-S). Para BJ46a, detectamos significativamente mais *cross-links* ácido-básico do que *cross-links* entre resíduos ácidos, uma particularidade da reação que também foi observada para DM64, um inibidor natural de fosfolipases A<sub>2</sub> de venenos de serpentes (Soares, 2018). É importante notar que conseguimos detectar *cross-links zero-length* partindo do primeiro domínio cistatina de BJ46a, na extremidade N-terminal, com a

região final do domínio rico em histidina na porção C-terminal. Este dado indica que as porções N- e C-terminais do inibidor estão espacialmente próximas – *cross-links zero-length* não apresentam um braço espaçador e a distância máxima a ser considerada corresponde apenas a distância entre as cadeias laterais dos resíduos envolvidos. Nos nossos dados, detectamos *links* nestas regiões envolvendo resíduos de ácido aspártico / ácido glutâmico e resíduos de lisina. Para um *cross-link* D-K, a distância máxima é de 9,7 Å, enquanto que para um *cross-link* E-K, de 10,5 Å (Ferrari *et al.*, 2018).

Para a metaloendopeptidase jararagina, mapeamos 42 *cross-links* com o ALC BS<sup>3</sup>, e 28 *cross-links* com a abordagem XPlex, todos do tipo ácido-básico (*zero-length*). Os mapas bidimensionais para cada abordagem estão representados na **Figura 35**. Os *cross-links* detectados com o ALC BS<sup>3</sup> unem o domínio metaloendopeptidase ao domínio rico em cisteína e o domínio desintegrina-símile ao rico em cisteína. Há também *cross-links* intrínsecos aos três domínios estruturais da toxina.

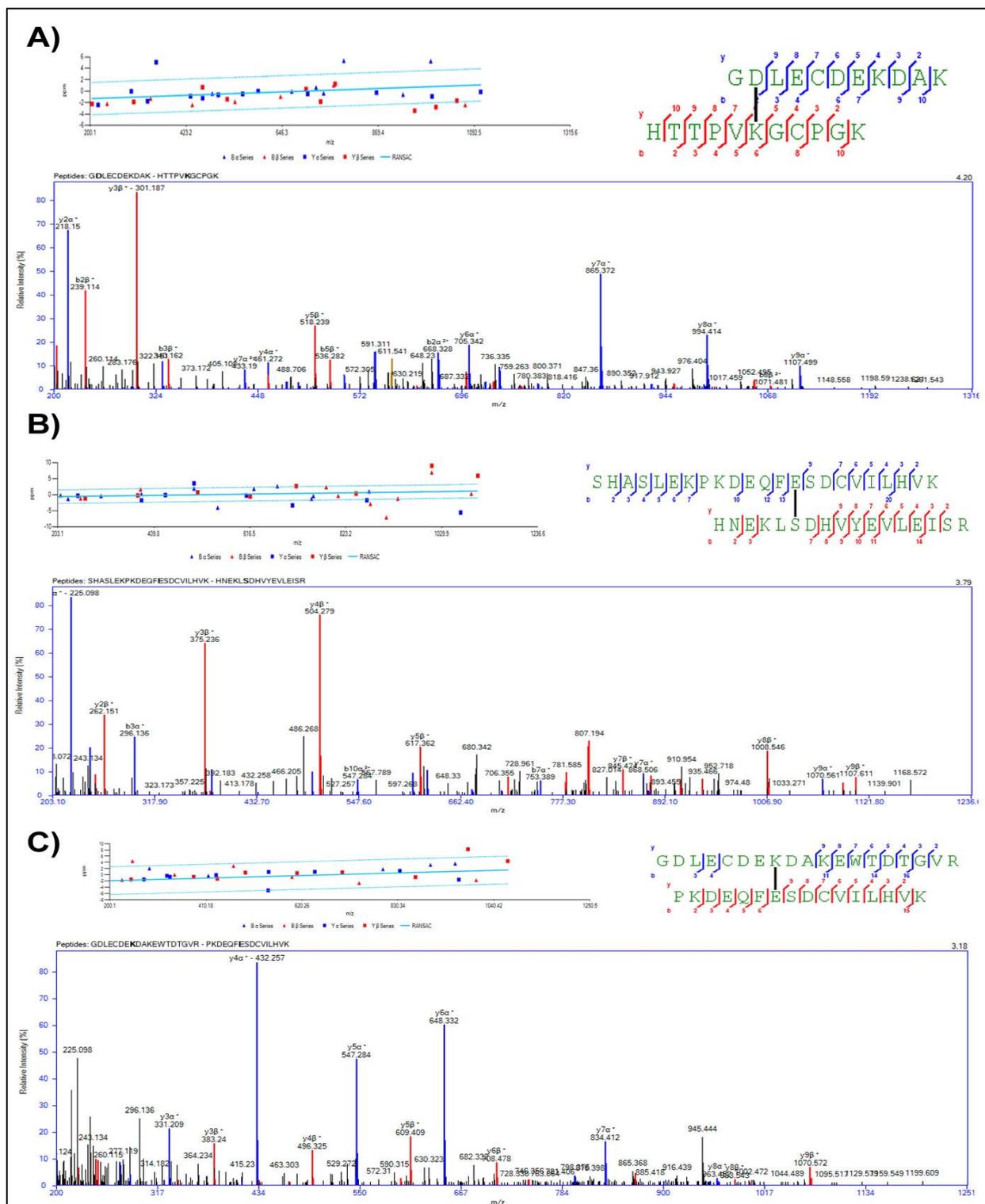
Botropasina, metaloendopeptidase isolada do veneno de *B. jararaca*, apresenta 95% de identidade sequencial com jararagina e tem estrutura cristalográfica depositada (PDB ID: 3DSL) (Muniz *et al.*, 2008). O domínio metaloendopeptidase apresenta uma folha  $\beta$  paralela com 5 fitas e 4 hélices; o domínio desintegrina-símile não apresenta elementos de estrutura secundária definidos e é composto primariamente de voltas estabilizadas por 7 pontes dissulfeto. Por fim, o domínio rico em cisteína apresenta pequenas hélices e fitas  $\beta$ , também estabilizadas por pontes dissulfeto (Muniz *et al.*, 2008).

A representação detalhada da estrutura cristalográfica de botropasina indica uma forma tridimensional em “C”, como demonstra a **Figura 36**. Traçando-se um paralelo, os padrões de *cross-links* detectados para jararagina demonstram a proximidade entre os domínios desintegrina-símile e rico em cisteína, e de regiões do domínio metaloendopeptidase com o domínio rico em cisteína na porção C-terminal da molécula, em concordância com a estrutura da botropasina. A similaridade sequencial e o padrão de *cross-links* obtido indicam que, provavelmente, jararagina e botropasina compartilham um mesmo padrão de enovelamento.



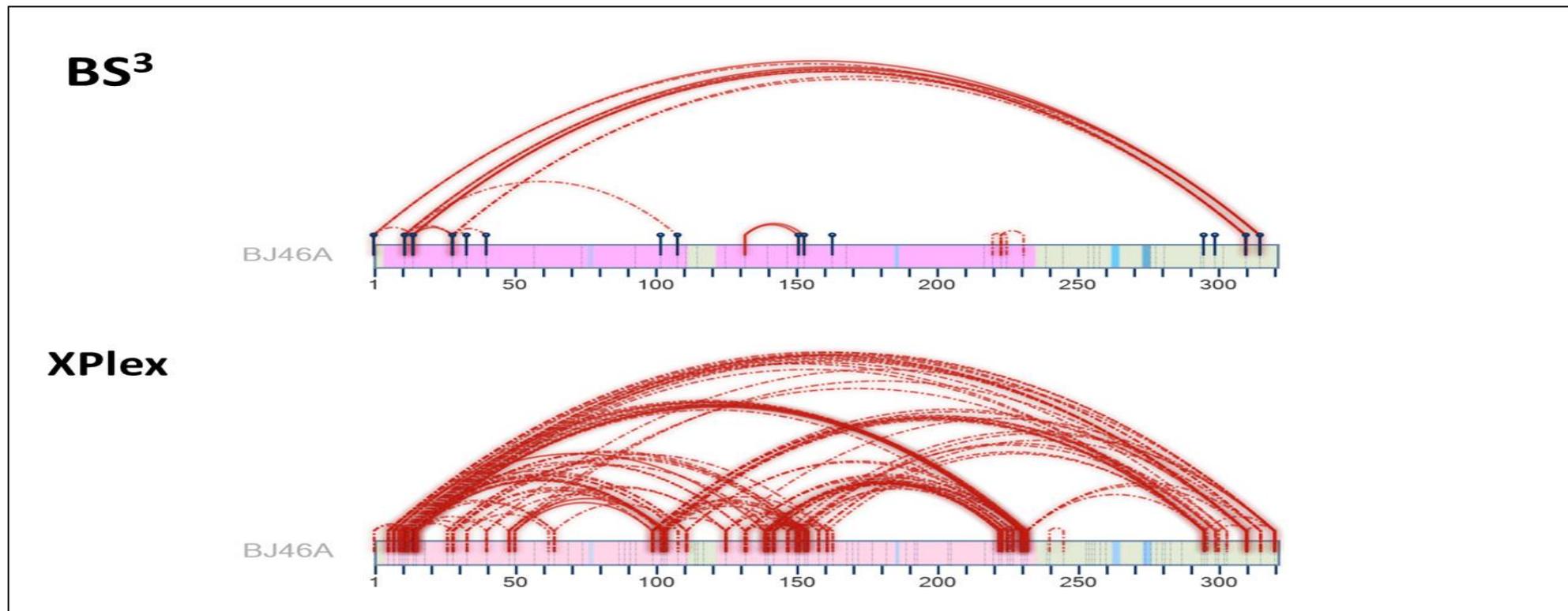
**Figura 29: Cross-links detectados para BJ46a livre utilizando o ALC BS<sup>3</sup>.**

Cada painel (A,B,C) ilustra um espectro representativo das categorias “excelente”, “bom” e “médio”, respectivamente, que foram subsequentemente empregados para gerar o mapa de restrição de distâncias, bem como a distribuição de erro dos íons-fragmento detectados e a cobertura de sequência dos peptídeos covalentemente ligados.



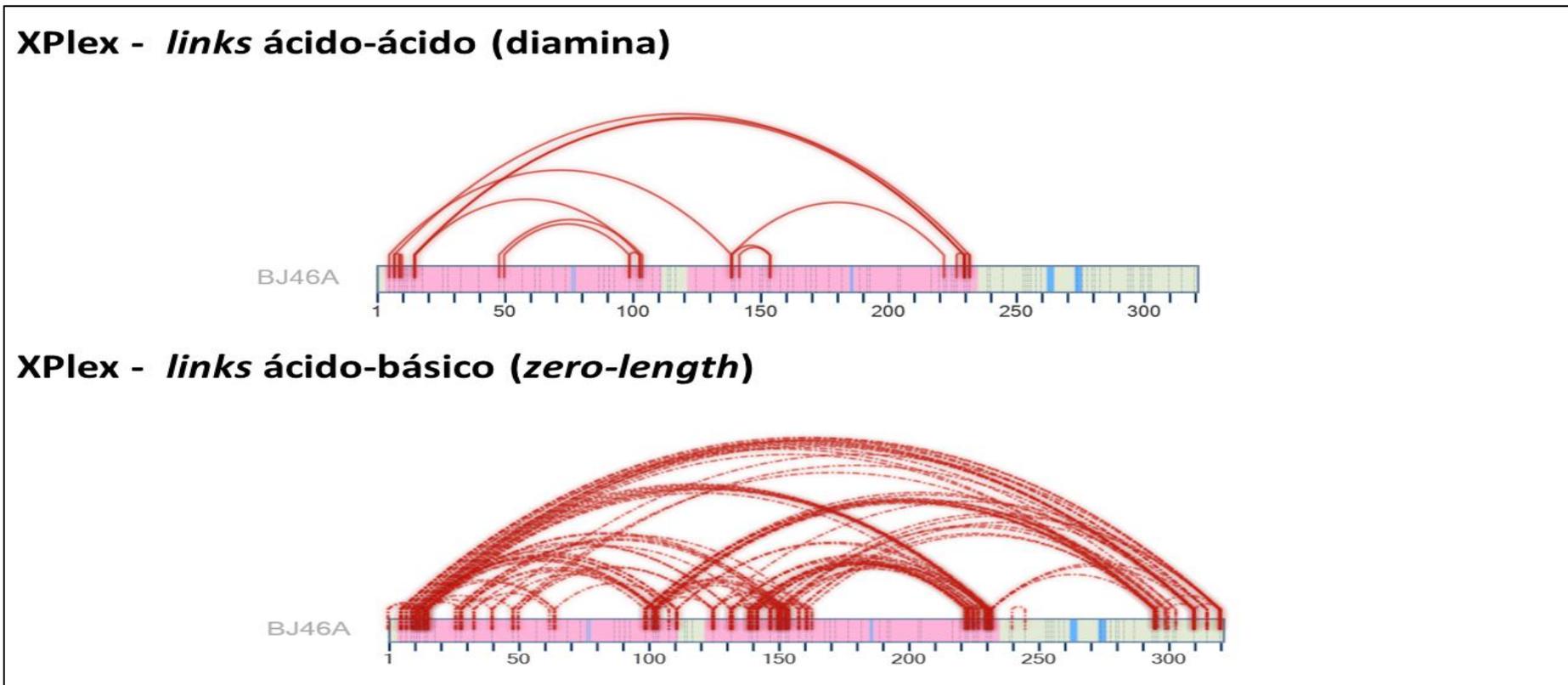
**Figura 30: Cross-links detectados para BJ46a livre utilizando a abordagem XPlax.**

Cada painel (A, B, C) ilustra um espectro representativo das categorias “excelente”, “bom” e “médio”, respectivamente, que foram subsequentemente empregados para gerar o mapa de restrição de distâncias, bem como a distribuição de erro dos íons-fragmento detectados e a cobertura de sequência dos peptídeos covalentemente ligados.



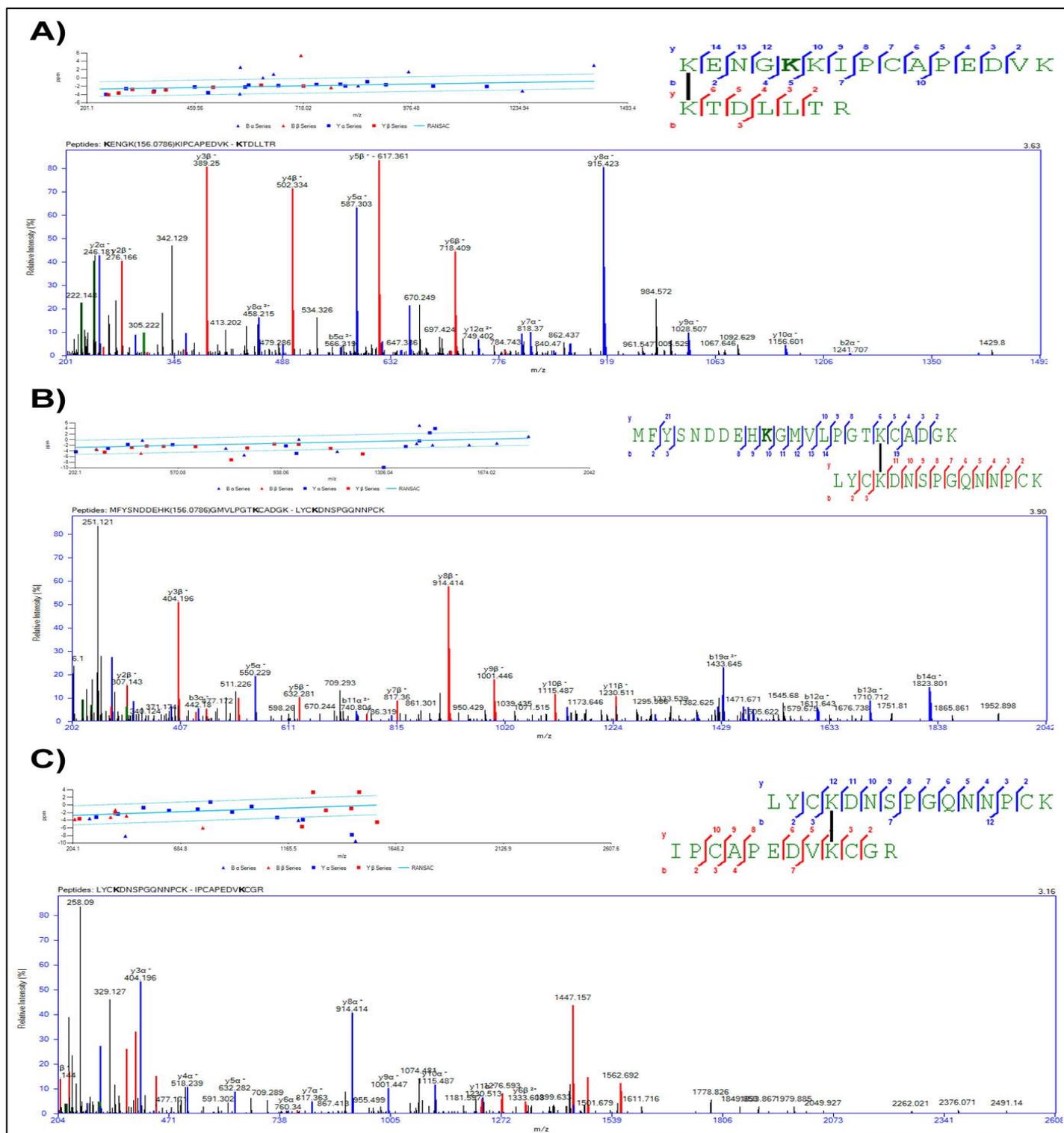
**Figura 31: Mapas bidimensionais de restrições de distância para BJ46a geradas pelas abordagens BS<sup>3</sup> e XPLex.**

Nesta imagem, a estrutura primária de BJ46a está representada como uma barra, e os domínios cistatina estão realçados em rosa, enquanto que os sítios de N-glicosilação do inibidor estão marcados em azul. Traços pontilhados em cinza no interior das barras indicam os resíduos passíveis de *cross-linking*. Cada linha vermelha indica um *cross-link* identificado (pertencentes às categorias excelente / bom / médio), enquanto que os pequenos pinos em azul, no mapa bidimensional de BS<sup>3</sup>, indicam *cross-links* do tipo 0 (*dead-end cross-links*). As linhas cheias e pontilhadas no mapa de restrições da abordagem XPLex indicam *cross-links* diamina (D-D, E-E, D-E) enquanto que as linhas pontilhadas representam *cross-links zero-length*, respectivamente. No mapa de restrições de BS<sup>3</sup>, indicam os *links* encontrados nas replicatas experimentais.



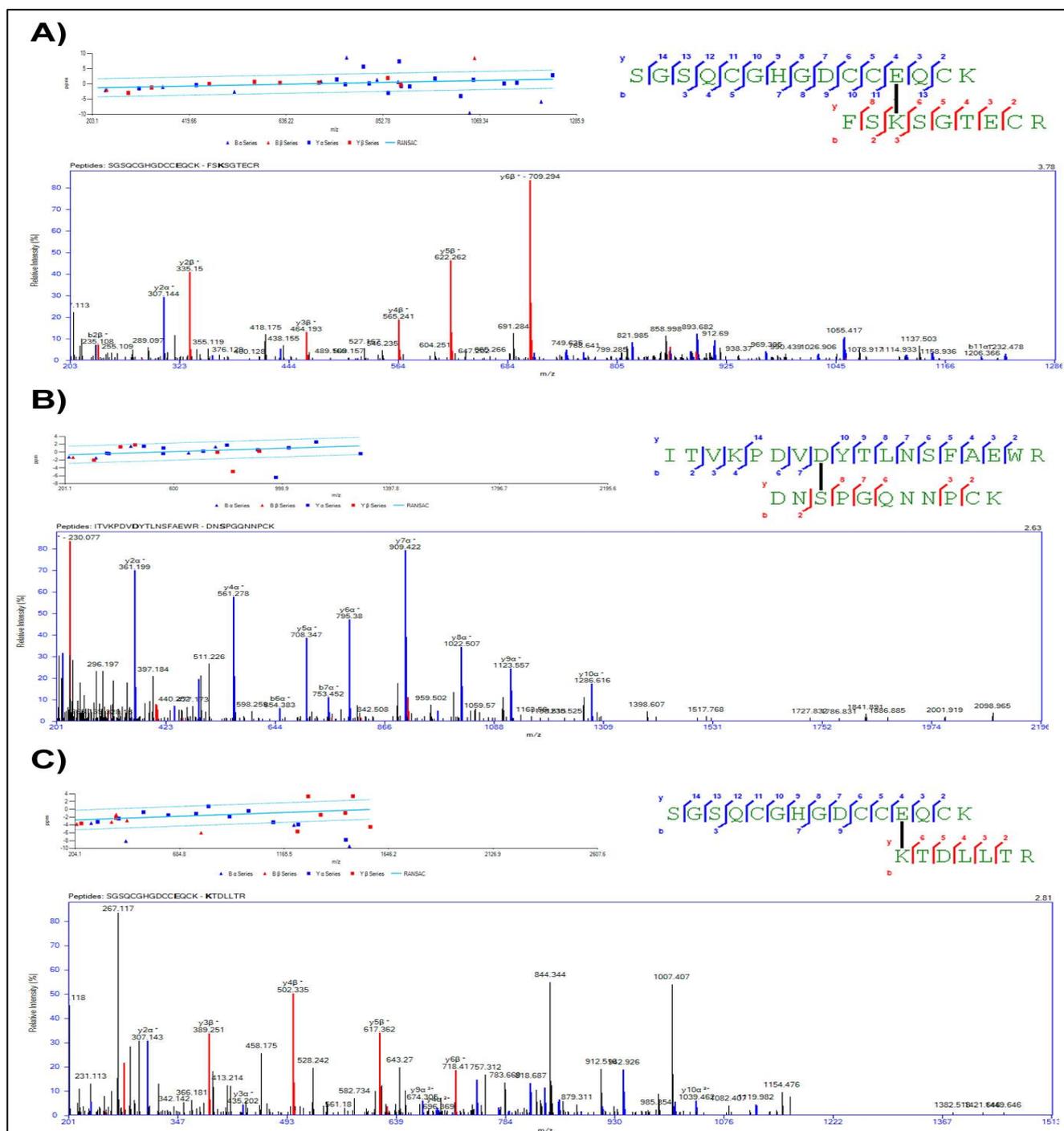
**Figura 32: Distribuição dos *cross-links* encontrados em BJ46a na abordagem XPLex, de acordo com o tipo de *cross-link* formado.**

Os *links* ácido-ácido referem-se aos *cross-links* formados entre resíduos de ácido aspártico e ácido glutâmico (D-D, D-E, E-E), enquanto que os links ácido-básico referem-se a *cross-links* formados entre um resíduo ácido (D/E) e um resíduo básico (K/S). Os domínios cistatina de BJ46a estão realçados em rosa, e os sítios de N-glicosilação, em azul. As linhas cheias e pontilhadas no mapa de restrições indicam *cross-links* diamina (D-D, E-E, D-E) enquanto que as linhas pontilhadas representam *cross-links zero-length*, respectivamente.



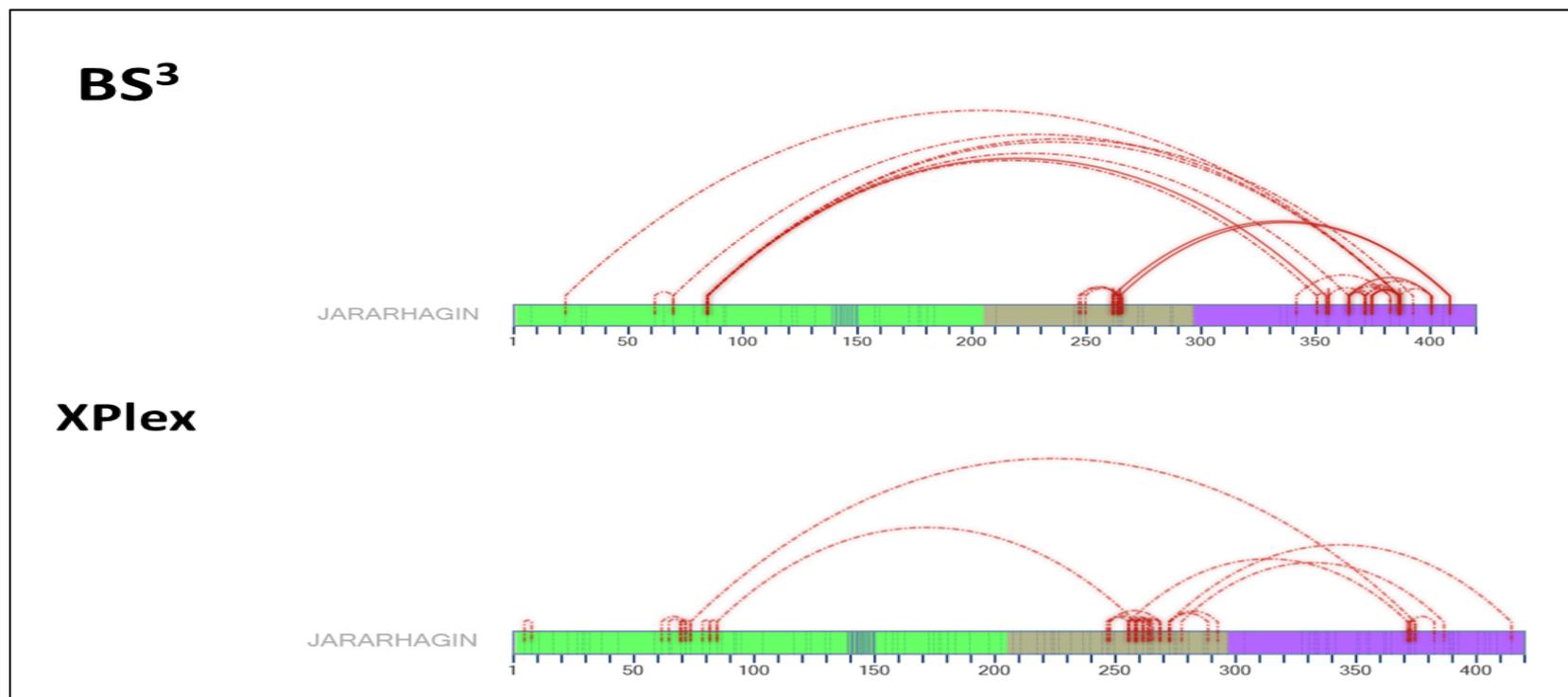
**Figura 33: Cross-links detectados para jararagina livre utilizando o ALC BS<sup>3</sup>.**

Cada painel (A, B, C) ilustra um espectro representativo das categorias “excelente”, “bom” e “médio”, respectivamente, que foram subsequentemente empregados para gerar o mapa de restrição de distâncias, bem como a distribuição de erro dos íons-fragmento detectados e a cobertura de sequência dos peptídeos covalentemente ligados.



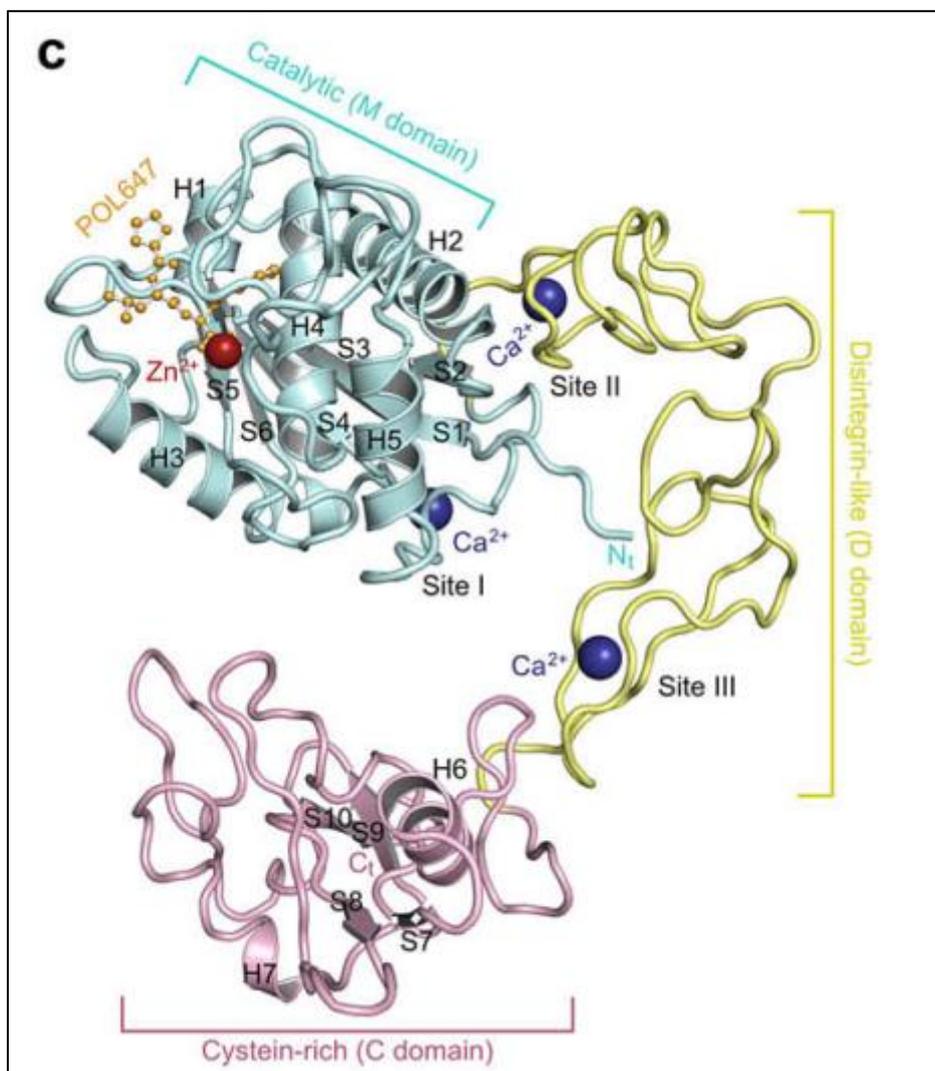
**Figura 34: Cross-links detectados para jararagina livre utilizando a abordagem XPLex.**

Cada painel (A, B, C) ilustra um espectro representativo das categorias “excelente”, “bom” e “médio”, respectivamente, que foram subsequentemente empregadas para gerar o mapa de restrição de distâncias, bem como a distribuição de erro dos íons-fragmento detectados e a cobertura de sequência dos peptídeos covalentemente ligados.



**Figura 35: Mapas bidimensionais de restrições de distância para jararagina gerados a partir das abordagens BS<sup>3</sup> e XPLex.**

Nesta representação, a estrutura primária de jararagina está representada como uma barra; o domínio metaloendopeptidase está realçado de verde, o domínio de ligação ao zinco, em azul, o domínio desintegrina, em cinza, e o domínio rico em cisteína em lilás. Cada *cross-link* detectado na estrutura de jararagina (das categorias excelente / bom / médio) foi representado como uma linha vermelha. As linhas cheias e pontilhadas no mapa de restrições da abordagem XPLex indicam *cross-links* diamina (D-D, E-E, D-E) enquanto que as linhas pontilhadas representam *cross-links zero-length*, respectivamente. No mapa de restrições de BS<sup>3</sup>, indicam os *links* encontrados nas replicatas experimentais.



**Figura 36: Representação detalhada da estrutura cristalográfica de botropasina, uma SVMP que apresenta 95% de identidade sequencial com jararagina.**

O domínio metaloendopeptidase está representado em azul, o domínio desintegrina-símile em amarelo e o domínio rico em cisteína, em lilás. Os átomos de cálcio, que estabilizam a estrutura da molécula, estão representados como esferas azuis e o átomo de zinco, essencial para a catálise enzimática, é representado como uma esfera em vermelho. Reproduzido de Muniz *et al.* 2008.

#### 4.3.2 – Análise dos espectros e mapas de restrição de distância do complexo BJ46a-jararagina

Paralelamente à análise dos *cross-links* obtidos para BJ46a e jararagina em suas formas livres, analisamos também o complexo entre a toxina e o inibidor com o ALC BS<sup>3</sup> e a abordagem XPlex, com o objetivo de identificar as regiões vizinhas à interface de interação entre estas duas proteínas. Os espectros representativos e os mapas de restrição de distância obtidos estão representados na **Figura 37** e nas **Figura 38 - 40**.

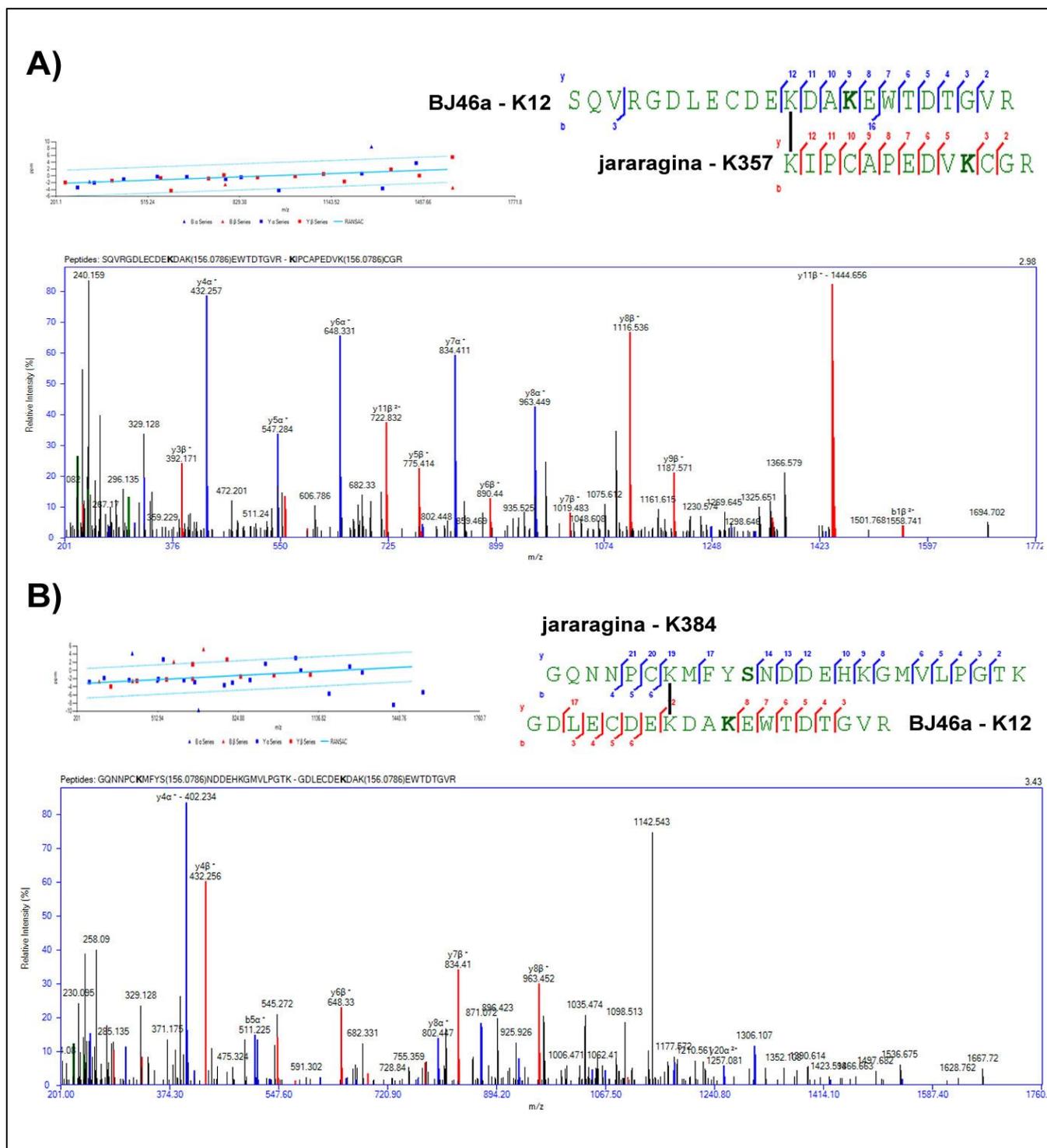
Para a interação BJ46a – jararagina (**Figura 39 e Figura 40**) pudemos identificar nos mapas 2D duas principais regiões de *cross-linking* entre estas duas proteínas: a primeira região envolve o primeiro domínio cistatina de BJ46a e o domínio rico em cisteína de jararagina, enquanto que a segunda região envolve o segundo domínio cistatina de BJ46a com o domínio metaloendopeptidase de jararagina (abordagem BS<sup>3</sup>, **Figura 39**). Outros *cross-links* também foram mapeados, conectando a porção C-terminal de BJ46a com o domínio metaloendopeptidase de jararagina e o segundo domínio cistatina de BJ46a com o domínio rico em cisteína de jararagina (abordagem XPlex, **Figura 40**). Por fim, detectamos *cross-links* conectando o domínio metaloendopeptidase de jararagina com o 1º domínio cistatina de BJ46a.

Em trabalho anterior, verificou-se que o domínio metaloendopeptidase de jararagina é essencial para a interação inibidor-toxina; BJ46a não é capaz de formar complexo com jararagina-C, uma forma processada da SVMP jararagina contendo apenas os domínios desintegrina e rico em cisteína, mesmo em um excesso molar de inibidor de três vezes (Valente *et al.*, 2001). Em nossos dados, identificamos 11 restrições de distância que partem de diferentes resíduos do inibidor BJ46a para resíduos pertencentes ao domínio metaloendopeptidase de jararagina (resíduos 1 – 205), reforçando a hipótese de que o domínio metaloendopeptidase da toxina seja parte importante da interação.

Além disso, os dados indicam um provável envolvimento dos dois domínios cistatina de BJ46a na interação inibidor-toxina, um dado inédito na literatura. Para HSF, Aoki e colaboradores demonstraram que a porção N-terminal do primeiro domínio cistatina do inibidor é indispensável para a inibição de SVMP, por meio de uma abordagem clássica de química de proteínas, incluindo modificações químicas seletivas em alguns resíduos e hidrólises sucessivas do inibidor. Entretanto, os

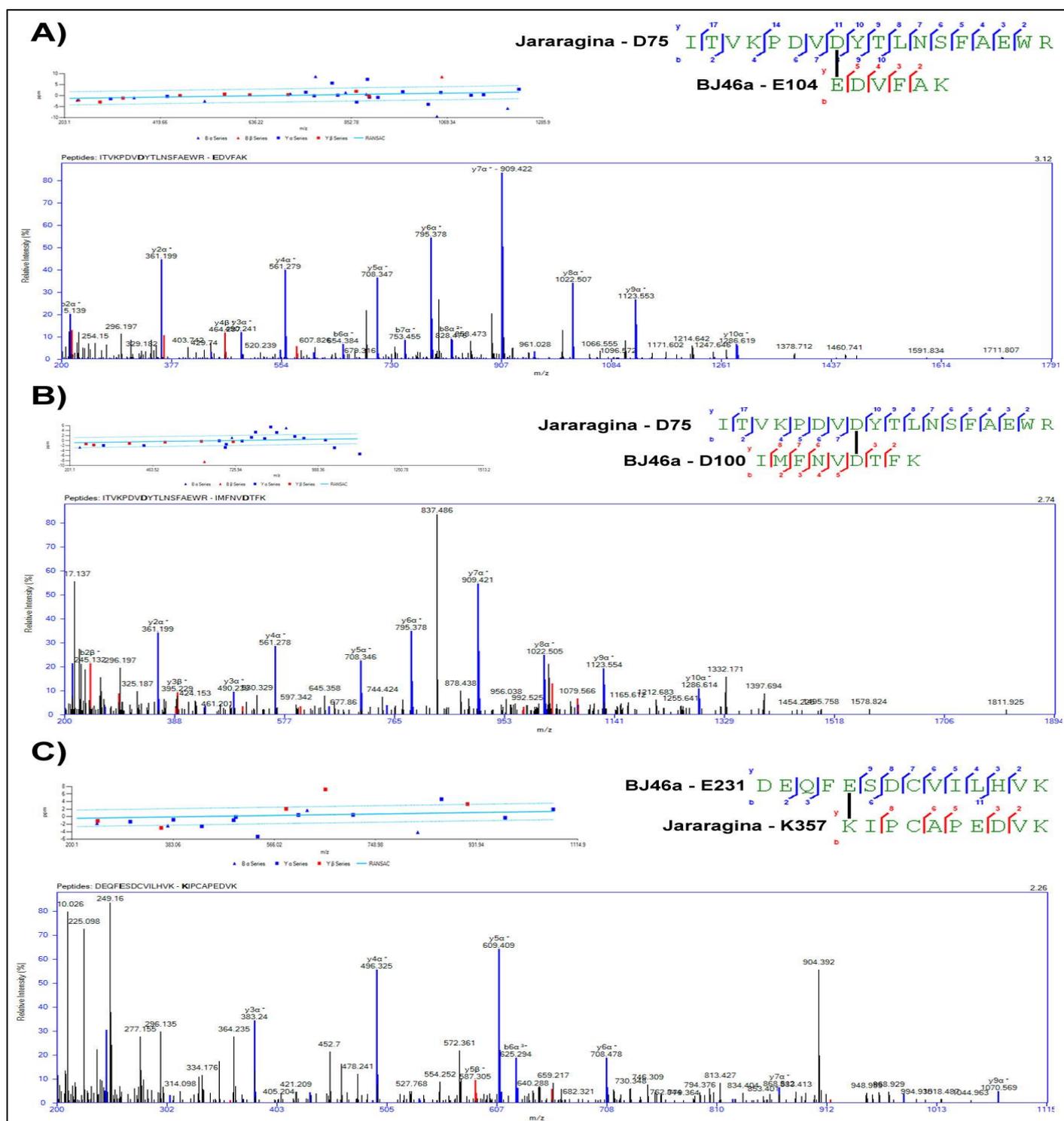
autores não reportaram nenhum fragmento correspondente ao segundo domínio cistatina de HSF que potencialmente estivesse envolvido com a interface de interação (Aoki, Deshimaru, *et al.*, 2007).

Quando comparamos os mapas 2D de BJ46a e jararagina controle (**Figura 31 e Figura 35**, respectivamente) com os *cross-links* encontrados na estrutura do inibidor e da toxina quando em complexo (**Figura 37 e Figura 38**, *cross-links* em vermelho) não verificamos mudanças significativas no perfil de restrições de distância que não a perda de alguns *links*, observáveis apenas nos controles das duas proteínas. Ainda assim, o padrão geral de restrições se mantém, o que sugere que não há uma mudança conformacional significativa em jararagina e/ou BJ46a quando da formação do complexo.



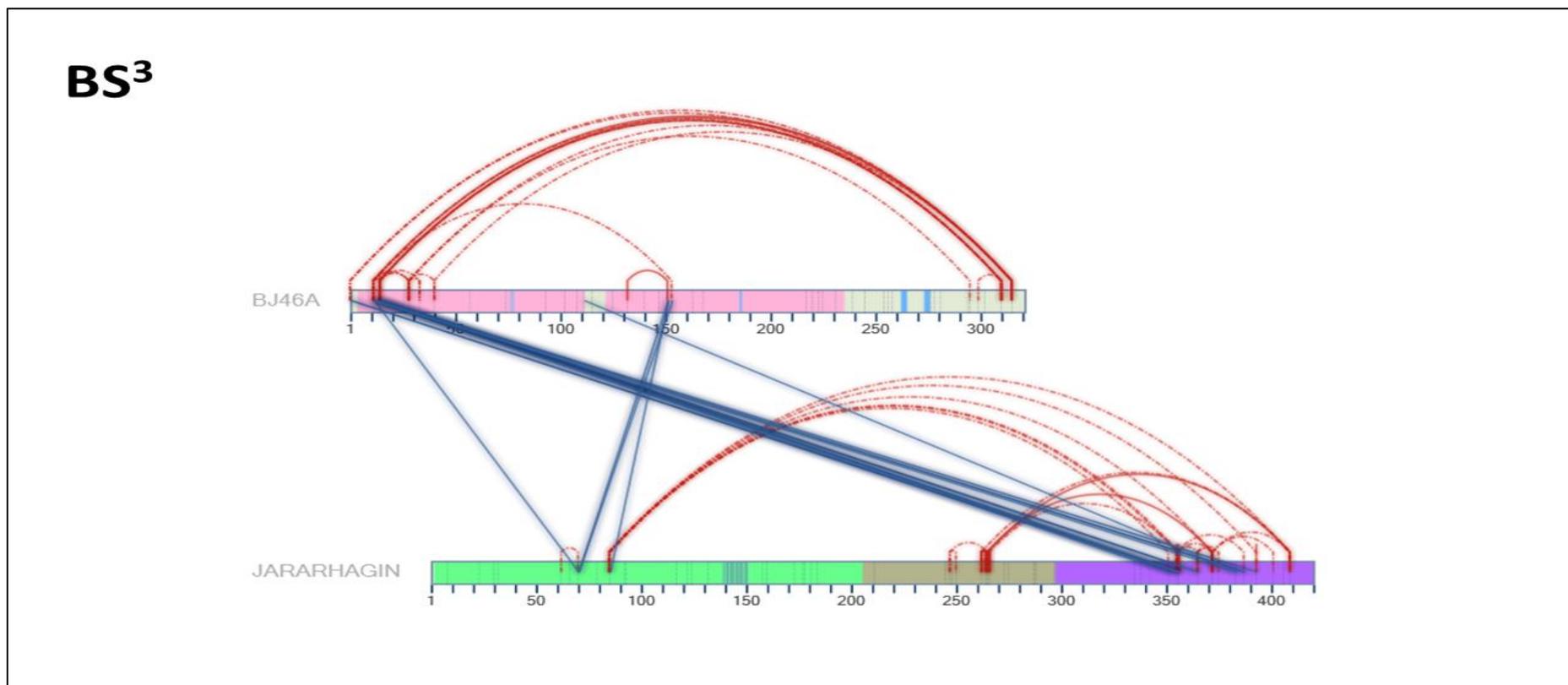
**Figura 37: Cross-links detectados para o complexo BJ46a-jararagina utilizando o ALC BS<sup>3</sup>.**

Os painéis A e B ilustram um espectro representativo das categorias “bom” e “médio”, respectivamente, que foram subsequentemente empregados para gerar o mapa de restrição de distâncias, bem como a distribuição de erro dos íons-fragmento detectados e a cobertura de sequência dos peptídeos covalentemente ligados.



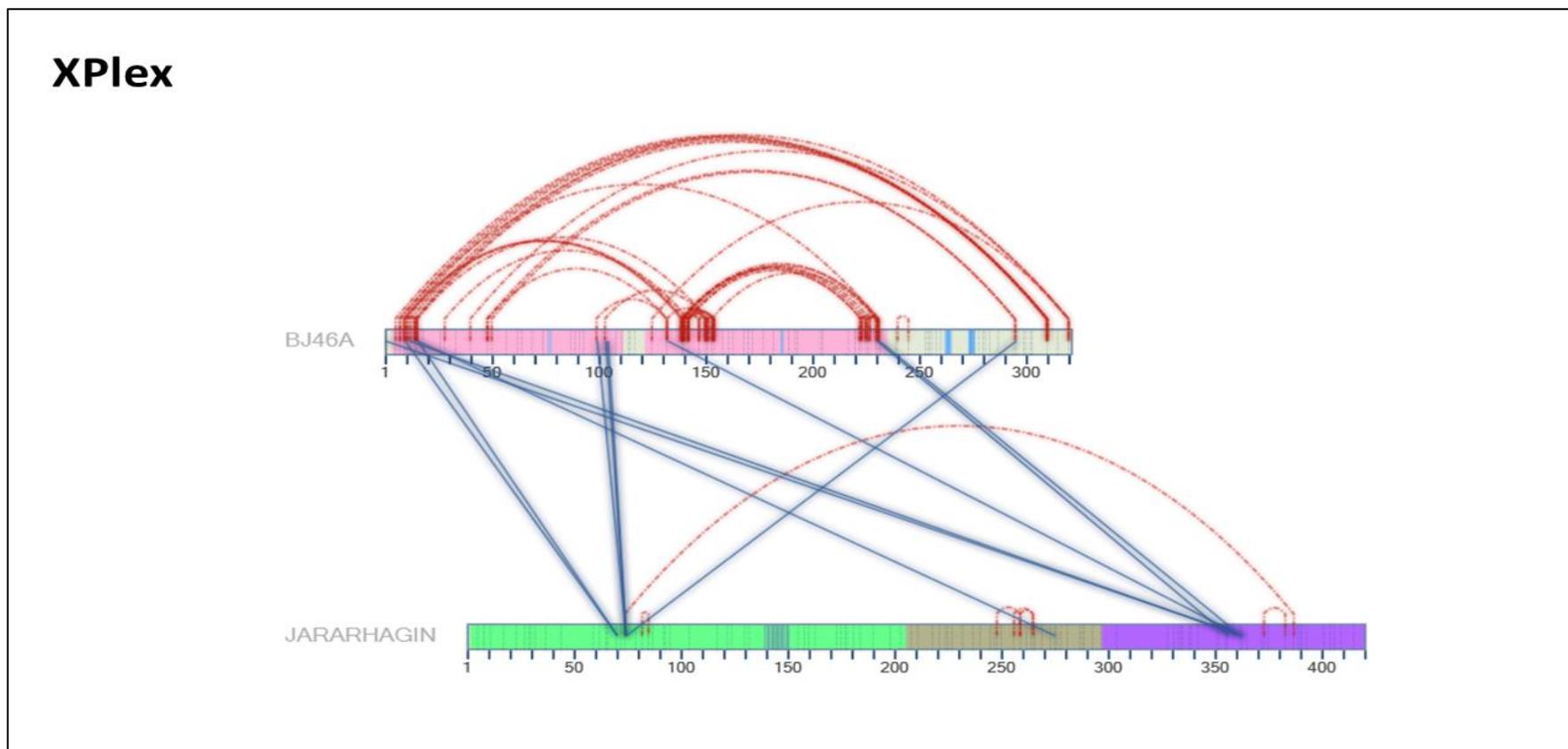
**Figura 38: Cross-links detectados para o complexo BJ46a-jararagina utilizando a abordagem XPLex.**

Cada painel (A, B, C) ilustra um espectro representativo das categorias “excelente”, “bom” e “médio”, respectivamente, que foram subsequentemente empregados para gerar o mapa de restrição de distâncias, bem como a distribuição de erro dos íons-fragmento detectados e a cobertura de sequência dos peptídeos covalentemente ligados.



**Figura 39: Mapa bidimensional das restrições de distância encontradas para o complexo BJ46a – jararagina utilizando o ALC BS<sup>3</sup>.**

Os domínios cistatina na estrutura de BJ46a estão realçados em rosa, e os sítios de N-glicosilação, em azul. Para jararagina, o domínio metaloendopeptidase está realçado em verde, o domínio desintegrina, em cinza e o domínio rico em cisteína, em lilás. Todos os *cross-links* classificados como excelentes / bons / médios estão representados neste mapa. Os *cross-links* intra-proteína estão representados em linhas vermelhas, enquanto que os *cross-links* inter-proteína, em linhas azuis. As linhas cheias e pontilhadas no mapa indicam os *links* encontrados nas replicatas experimentais.



**Figura 40: Mapa bidimensional das restrições de distância encontradas para o complexo BJ46a – jararagina utilizando a abordagem XPLex.**

Os domínios cistatina na estrutura de BJ46a estão realçados em rosa, e os sítios de N-glicosilação, em azul. Para jararagina, o domínio metaloendopeptidase está realçado em verde, o domínio desintegrina, em cinza e o domínio rico em cisteína, em lilás. Todos os *cross-links* classificados como excelentes / bons / médios estão representados neste mapa. Os *cross-links* intra-proteína estão representados em linhas vermelhas, enquanto que os *cross-links* inter-proteína, em linhas azuis.

#### 4.3.3 - Determinação experimental de pontes dissulfeto de BJ46a

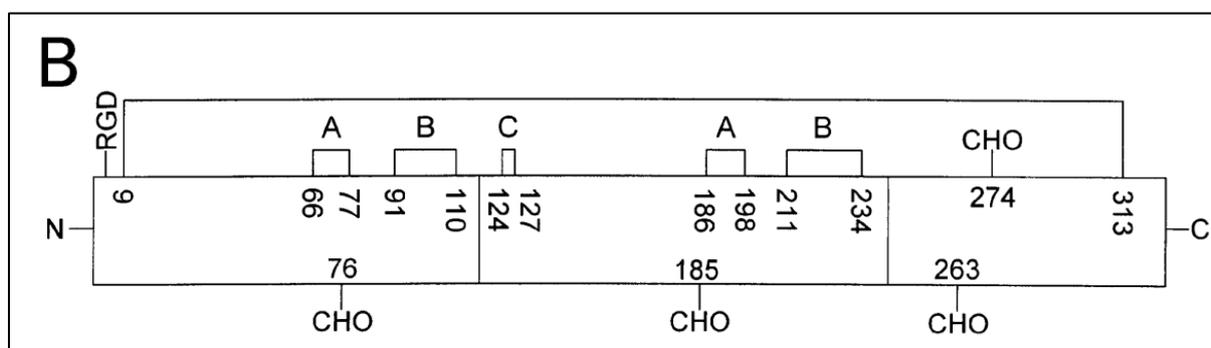
Para a modelagem integrativa de proteínas, um conjunto amplo de restrições experimentais pode ser útil para guiar a construção de modelos e/ou filtrar modelos incompatíveis com os dados disponíveis (Ward *et al.*, 2013). Dessa forma, decidimos explorar o padrão de pontes dissulfeto do inibidor BJ46a como um conjunto adicional de restrições espaciais. As pontes dissulfeto são um *cross-link* intrínseco em proteínas e, com exceção de proteínas redox, o padrão das pontes é conservado. Além disso, elas mantêm a integridade da estrutura terciária e podem auxiliar na predição de estrutura de proteínas ricas em pontes dissulfeto que não apresentem proteínas homólogas com estrutura resolvida (Kaltashov e Eyles, 2012).

Ao longo dos anos, uma série de métodos bioquímicos e biofísicos foram empregados para o mapeamento e análise de pontes dissulfeto em proteínas, como a degradação de Edman, mutagênese de resíduos de cisteína combinado com SDS-PAGE em condições redutoras e não-redutoras, cristalografia de raios-X e espectroscopia de ressonância magnética nuclear. Nos últimos anos, a determinação de pontes dissulfeto baseada em LC-MS se ampliou, devido ao baixo consumo de amostra e pela possibilidade de se determinar pontes dissulfeto em misturas complexas de proteínas (Lu *et al.*, 2018).

A formação de uma ponte dissulfeto entre dois resíduos de cisteína promove um decréscimo de massa de 2 Da, diferença passível de ser medida por espectrômetros de massas modernos; entretanto, a análise de dados é complexa, pois a maioria das ferramentas de identificação de proteínas e peptídeos por dados de LC-MS são destinadas a peptídeos lineares (Tsai *et al.*, 2013). Assim, para identificarmos as pontes dissulfeto em BJ46a, utilizamos um protocolo otimizado recentemente para LC-MS (capítulo 3, item 3.2.13) e utilizamos a ferramenta SIM-XL, explorando a ponte dissulfeto como um *cross-link* natural para facilitar a análise e representação dos dados.

O inibidor BJ46a possui 12 cisteínas em sua estrutura primária, que estariam arranjadas em 6 pontes dissulfeto preditas, ilustrada na **Figura 41**: C9 – C313, C66 – C77, C91 – C110, C124 – C127, C186 – C198, C211 – C234. A predição foi baseada no descrito para fetuínas, proteínas da superfamília das cistatinas envolvidas com diversos processos biológicos. Assim, Valente e colaboradores postularam que BJ46a possui em cada domínio cistatina *loops* do tipo A e do tipo B, e um *loop* do tipo C no

segundo domínio cistatina. BJ46a também apresentaria uma ponte ligando a porção N-terminal ao C-terminal do inibidor (Valente *et al.*, 2001).



**Figura 41: Padrão predito de pontes dissulfeto em BJ46a, com base em fetuínas.**

Nesta representação, a estrutura primária do inibidor está representada como uma barra, e linhas conectam as cisteínas envolvidas em pontes dissulfeto. Os sítios de N-glicosilação do inibidor estão marcados (CHO), assim como o motivo RGD, presente no N-terminal de BJ46a. Reproduzido de Valente *et al.* 2001.

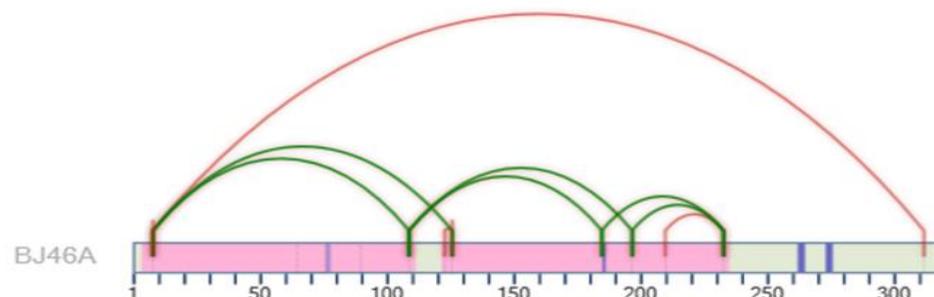
Na análise inicial das pontes dissulfeto de BJ46a pelo software SIM-XL, realizamos uma filtragem com os parâmetros do programa (score primário  $\geq 2,0$ , erro  $\leq 10$  ppm, mínimo de 1 espectro reportado e 3 picos assinalados por cadeia), gerando o padrão de pontes dissulfeto reportado na parte superior da **Figura 42**. Mapeamos as pontes C9-C313, C124-C127 e C211-C234, já previstas (*links* em vermelho), mas também detectamos pontes dissulfeto não correspondentes com a predição inicial (*links* em verde). O possível embaralhamento de pontes dissulfeto é uma limitação do experimento de assinalamento por LC-MS. Este fenômeno é um artefato experimental, que ocorre espontaneamente durante o preparo de amostra. Em valores de pH na faixa básica, os grupamentos tiol (-SH) presentes na cadeia lateral dos resíduos de cisteína convertem-se em ânions tiolato ( $S^-$ ) que podem realizar um ataque nucleofílico a ligações S-S, formando uma ponte dissulfeto embaralhada. A temperatura também é um parâmetro crítico neste processo (Glocker *et al.*, 1995; Tang e Speicher, 2004).

Frequentemente, o embaralhamento ocorre na presença de cisteínas livres, mas também pode ser observado em proteínas que apresentam todos os seus resíduos de cisteína envolvidos em pontes dissulfeto e que são expostas a condições desnaturantes brandas (Tang e Speicher, 2004). Mesmo tomando-se as precauções necessárias para minimizar o embaralhamento artefactual (digestões conduzidas em

pH 6,5 e bloqueio de eventuais cisteínas livres por N-etilmaleimida), não conseguimos evitar o embaralhamento de 4 pontes dissulfeto de BJ46a. Mas, na análise manual dos dados obtidos, estes espectros não apresentaram critérios mínimos de qualidade para a validação do dado. O uso de proteases que hidrolisam substratos em pH mais ácido, como pepsina, e tempos menores de incubação a 37°C podem minimizar o fenômeno de embaralhamento para amostras de BJ46a, possibilitando o assinalamento correto destas pontes.

Após a análise manual dos espectros, as pontes dissulfeto que conseguiram ser validadas experimentalmente envolvem as cisteínas C9-C313 e C124-C127 (**Figura 42 e Figura 43**), e serão consideradas como restrições espaciais extras para a modelagem molecular do inibidor BJ46a. A confirmação da ponte C9-C313 indica que o primeiro domínio cistatina de BJ46a está próximo espacialmente do domínio rico em histidina localizado na porção C-terminal, uma vez que uma ponte dissulfeto tem a distância aproximada de 2,04 Å (Dombkowski, 2003). É possível que esta ligação justifique o padrão de *cross-links* observado para BJ46a em sua forma livre, onde verificamos um conjunto de *links* partindo da região N-terminal ao C-terminal da proteína (**Figura 29 e Figura 30**). Esta ligação das porções N- e C-terminais por uma ponte dissulfeto é uma característica distintiva de fetuínas, reforçando a similaridade estrutural de BJ46a com este grupo de proteínas (Lee *et al.*, 2009).

### C-C – todos os *links* reportados:



### C-C – *links* validados:

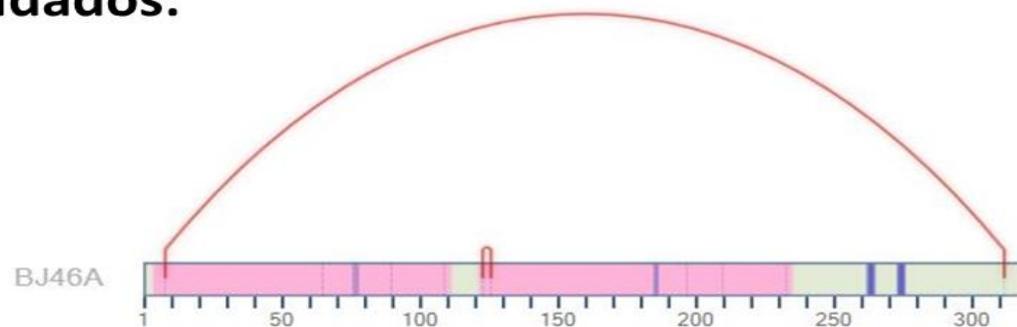


Figura 42: Análise das pontes dissulfeto de BJ46a pelo software SIM-XL.

O painel superior (todos os *links* reportados) indica todas as pontes dissulfeto reportadas pelo programa. Os *links* em verde correspondem a pontes dissulfeto não previstas e potencialmente geradas artefactualmente. Já o painel inferior indica apenas os *cross-links* que foram validados pela análise manual dos espectros. Os domínios cistatina de BJ46a estão realçados em rosa, enquanto que os sítios de N-glicosilação, em azul

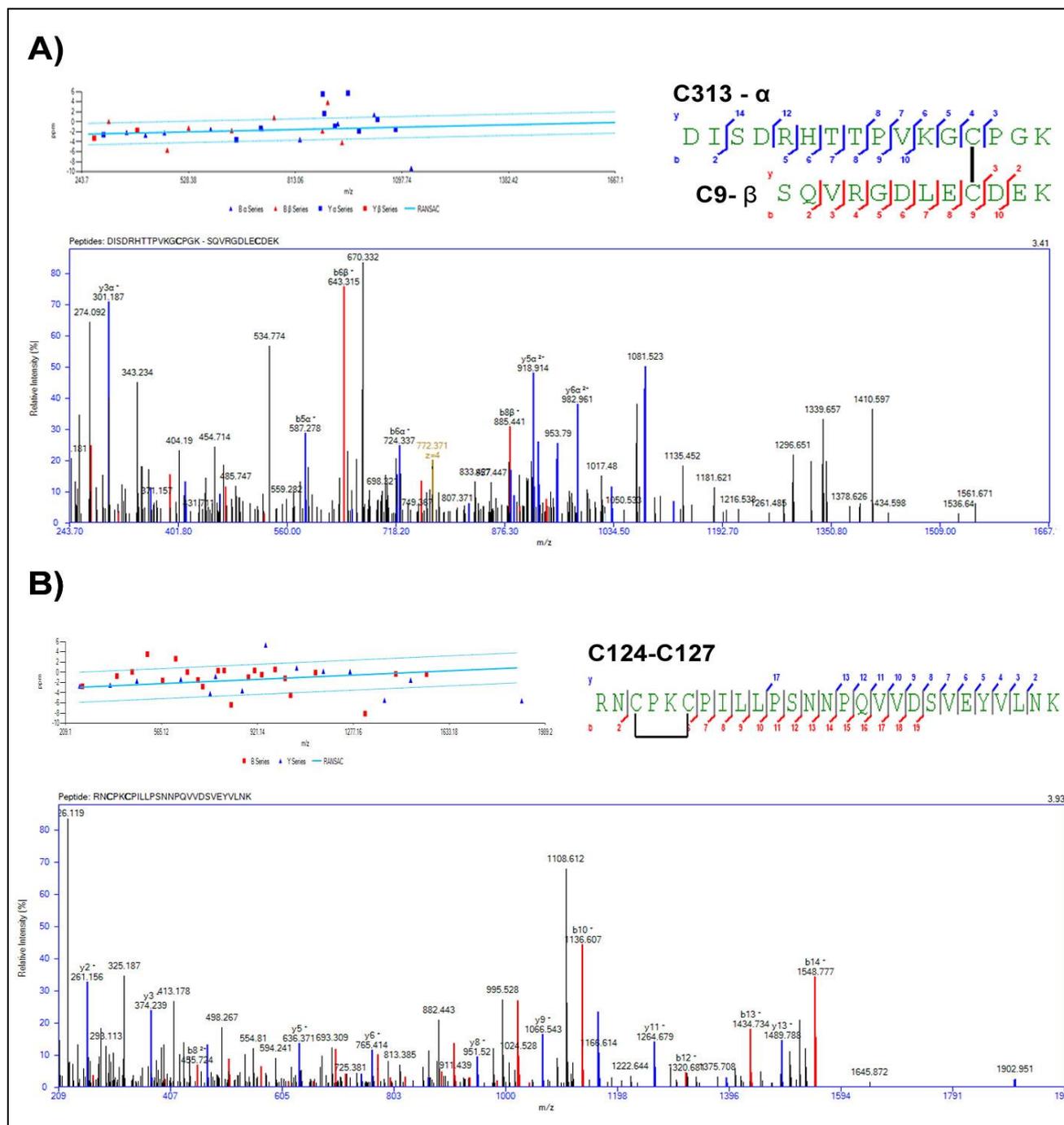


Figura 43: Espectros de massas das pontes dissulfeto validadas manualmente (C9-C313 e C124-C127).

#### 4.4 - ANÁLISE DO COMPLEXO BJ46a – JARARAGINA POR TROCA ISOTÓPICA HIDROGÊNIO/DEUTÉRIO (HDX-MS)

##### 4.4.1 – Construção de um sistema especializado para aquisição de dados de HDX-MS

O monitoramento da troca isotópica hidrogênio/deutério por espectrometria de massas é um desafio analítico. Devido à labilidade da marcação isotópica, o sistema precisa ser construído para que não haja mudanças significativas no nível de incorporação, isto é, não podem ocorrer perdas excessivas de deutério (*back-exchange*) e nem mais deutério pode ser incorporado após a finalização do experimento de incorporação (*forward-exchange*). Para isso, todas as conexões do sistema em contato com a amostra precisam ser mantidas continuamente em condições favoráveis para um valor mínimo de  $k_{ch}$ , 0°C e pH 2,5 (subseção 1.4.1.2, capítulo 1).

Para a etapa de hidrólise, o uso de colunas cromatográficas com enzimas imobilizadas é o ideal para análises de HDX-MS, pois a proporção enzima/substrato é significativamente maior devido à imobilização, o que aumenta a eficiência da digestão enzimática. A separação cromatográfica dos peptídeos resultantes deve ser feita a 0°C e com gradientes rápidos, para minimizar a taxa de *back-exchange* resultante do contato do analito deuterado com a água contida na fase móvel. Gradientes muito rápidos resultam em perda de resolução cromatográfica, mas esta limitação pode ser parcialmente contornada aumentando-se o número de pratos teóricos, pelo uso de colunas com tamanhos menores de partícula. Além disso, devemos levar em consideração o aumento da pressão devido ao aumento da viscosidade do solvente devido à baixa temperatura e ao uso de colunas com pequenas partículas. Logo, o sistema deve ser equipado com bombas cromatográficas capazes de trabalhar nestes níveis de pressão (Engen e Wales, 2015). Por fim, o sistema deve permitir a integração eficiente com o espectrômetro de massas, para o início da aquisição dos espectros MS1.

Implementamos no Laboratório de Toxinologia (Instituto Oswaldo Cruz / FIOCRUZ) um sistema para aquisição de dados de HDX-MS para esta tese. O equipamento é composto por duas bombas quaternárias Surveyor (Thermo Scientific), uma válvula de injeção manual (Rheodyne) conectada a um *loop* de amostragem de

50  $\mu$ L e uma válvula de 10 vias (Valco); as válvulas estão adaptadas em um elevador manual, que, ao ser abaixado, submerge todas as conexões do sistema em um tanque a 0°C, como demonstra a **Figura 44**. O tanque tem capacidade para 65 L e possui serpentinas acopladas por todo o interior e uma turbina de circulação de 2000 L/h localizada no fundo para a constante mistura da solução refrigerada de água/etanol (**Figura 45**). A temperatura é mantida continuamente a 0°C por um circulador termostático (Multitemp IV, GE Healthcare). Todas as fases móveis, antes de entrarem em contato com a amostra, são resfriadas a 0°C por serpentinas (**Figura 46**).

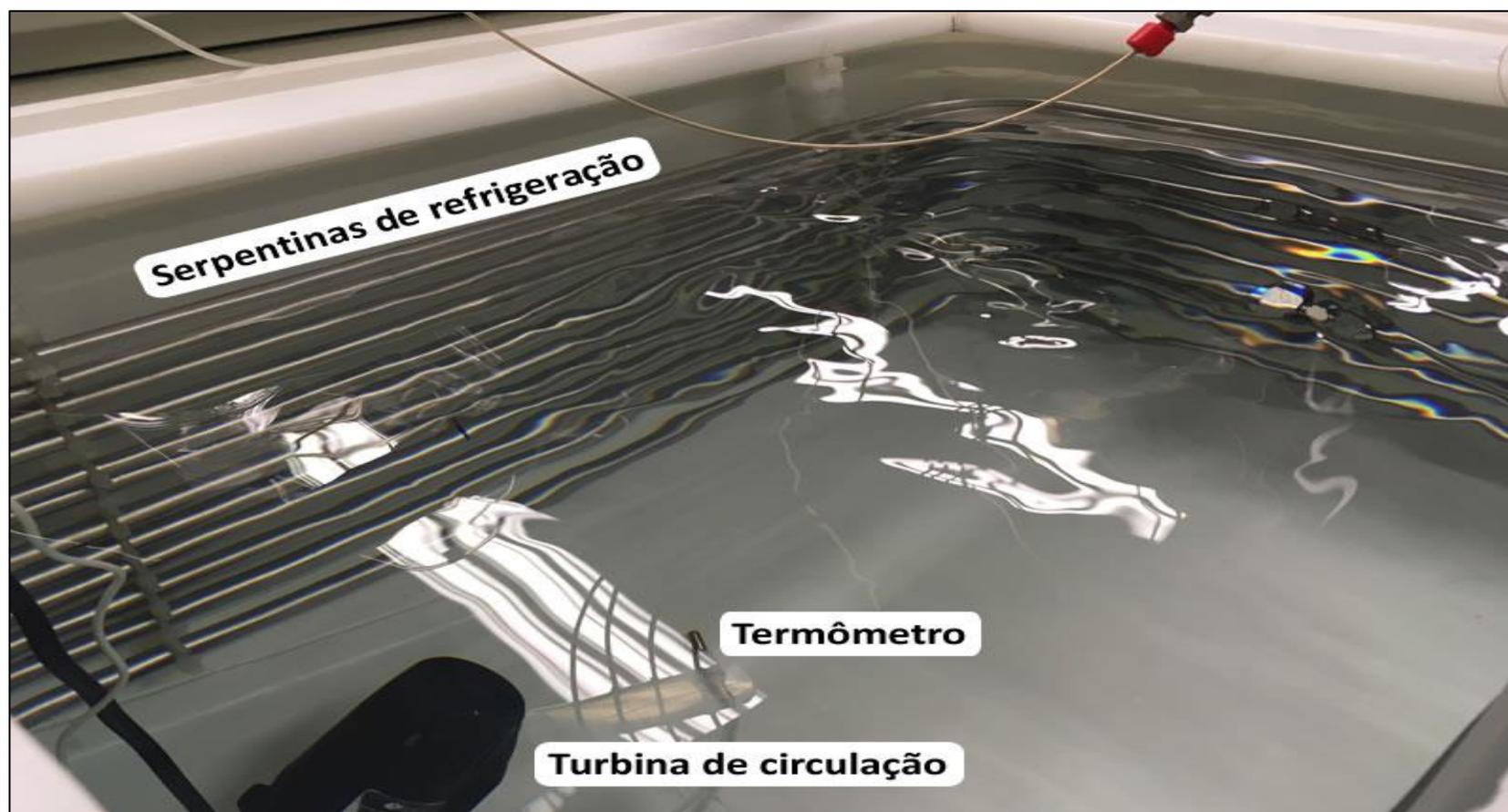
As bombas quaternárias são controladas pelo *software* Janeiro II (Flux Instruments) e a integração entre o sistema de HDX-MS e o espectrômetro de massas é feita por meio de um temporizador (relé) que é ativado manualmente imediatamente após a injeção (interruptor de partida do sistema, demonstrado na **Figura 44**). Uma vez que a amostra esteja no sistema, uma das bombas quaternárias (chamada de *sample pump*), mantida em condição isocrática (100  $\mu$ L/min em ácido fórmico 0,23% em água), desloca a amostra do *loop* de amostragem para a coluna de pepsina-agarose, para a hidrólise da amostra. Os peptídeos resultantes serão capturados e dessalinizados em coluna submersa de guarda C18 conectada à válvula de 10 vias (**Figura 46**). Eventuais fragmentos proteicos não digeridos pela pepsina imobilizada são capturados por um dispositivo posicionado logo na saída da coluna de pepsina-agarose (**Figura 47**, coluna pepsina-agarose, peça marrom).

Após 5 minutos, o relé dispara e a válvula de 10 vias troca de posição, fechando a conexão com a coluna de pepsina e agora conectando diretamente a coluna de guarda C18 à coluna analítica C18, para iniciar a cromatografia de fase reversa. A segunda bomba quaternária (chamada de *MS Pump*), mantida a 40  $\mu$ L/min, é responsável pela execução do gradiente de separação na coluna analítica para a separação dos peptídeos. A saída da coluna analítica está conectada diretamente via capilar à fonte de ionização *microelectrospray* ( $\mu$ ESI) do espectrômetro de massas LTQ Orbitrap XL (**Figura 46**).



**Figura 44: Sistema de HDX-MS implementado no Laboratório de Toxinologia, IOC/FIOCRUZ.**

Legenda: 1 – Computador do sistema de HDX-MS, que controla as bombas cromatográficas; 2: Bombas quaternárias; 3: Tanque de refrigeração; 4: Interruptor de partida do sistema; 5: Circulador termostático.



**Figura 45: Visão interna do tanque do sistema de HDX-MS.**

A solução contida no interior do tanque é uma mistura de álcool comercial e água, para impedir o congelamento. Esta mistura é continuamente refrigerada pelas serpentinas posicionadas em volta de todo o tanque e a turbina de circulação efetua a mistura desta solução, possibilitando uma temperatura uniforme, medida continuamente com um termômetro submerso.

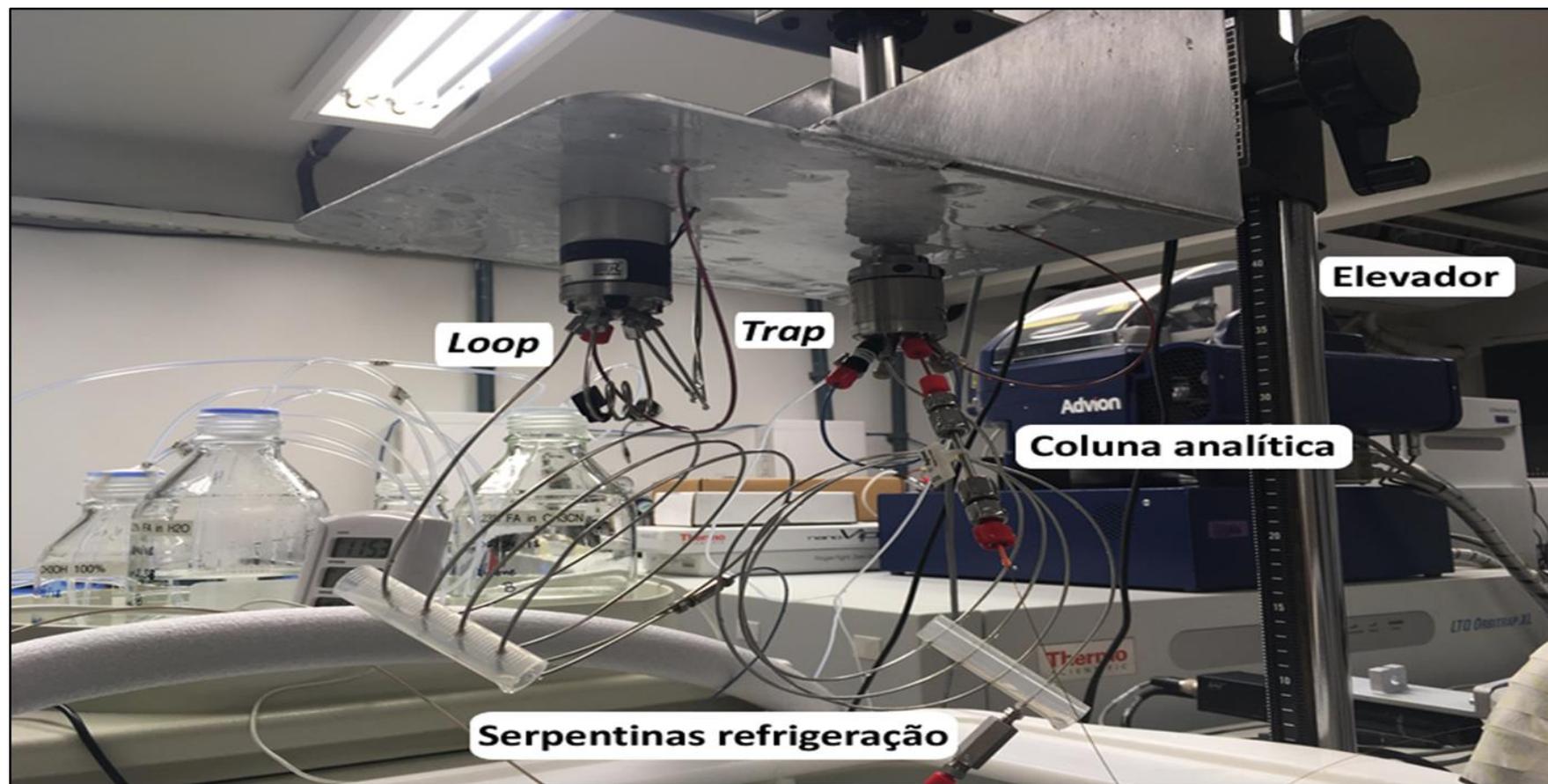
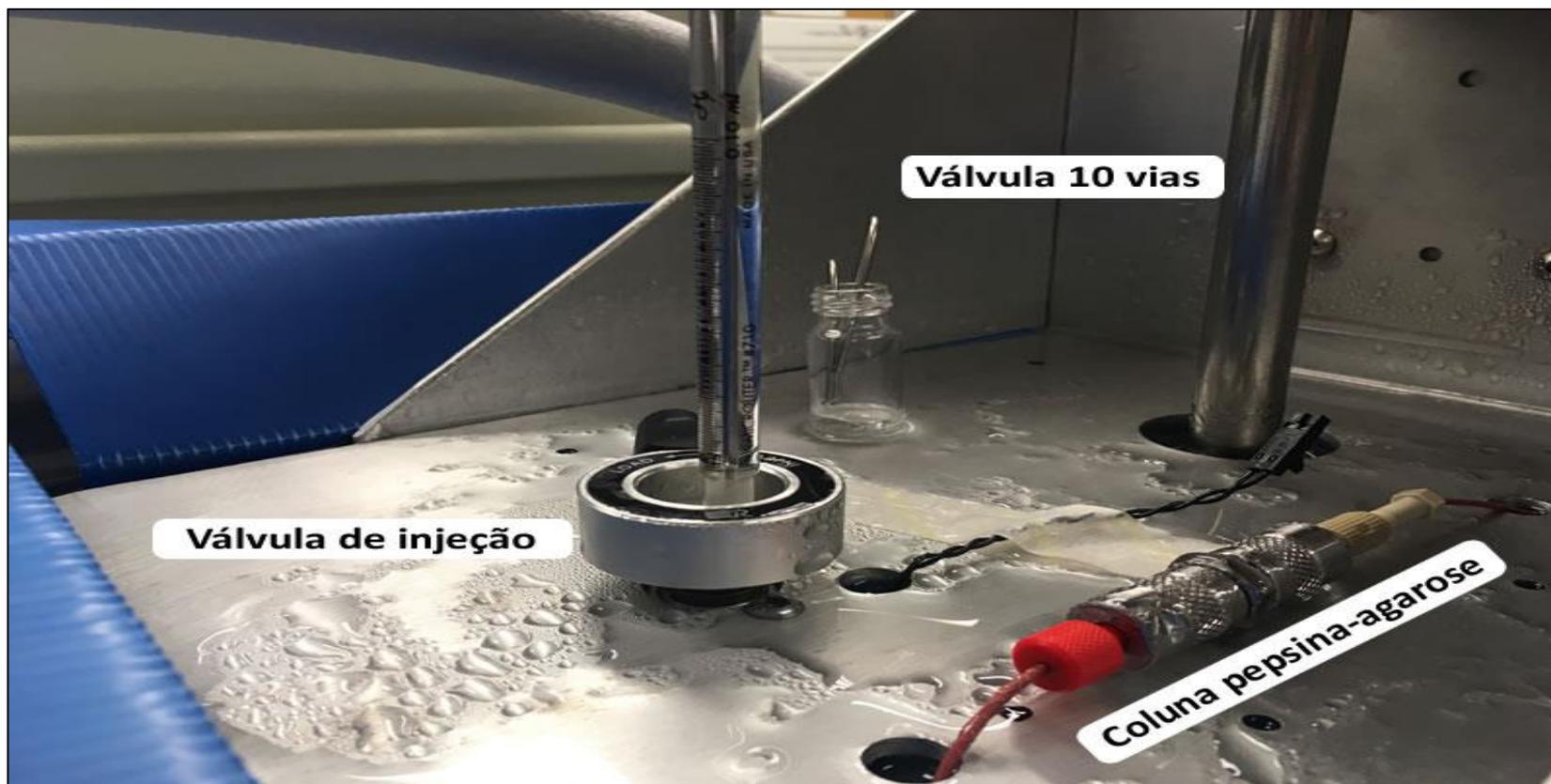


Figura 46: Visão das conexões e partes submersas do sistema de HDX-MS.

Esta visualização corresponde à parte inferior do elevador, que abriga as partes que serão submersas no banho a 0°C: as serpentinas de refrigeração da fase móvel, o *loop* de amostragem, a *trap*, que concentra os peptídeos resultantes da digestão pela coluna pepsina-agarose, a coluna analítica, que efetua a separação destes peptídeos. A saída da coluna analítica está conectada a um capilar, que por sua vez está conectado diretamente à fonte de ionização do espectrômetro LTQ Orbitrap XL.



**Figura 47: Visão superior do sistema de HDX-MS.**

Nesta figura, encontram-se as peças posicionadas na parte superior do elevador: válvula de injeção, coluna de pepsina-agarose e a parte superior da válvula de 10 vias.

#### 4.4.2 – Otimização das condições experimentais para a curva de incorporação de deutério

Previamente à análise de proteínas e complexos proteicos por HDX-MS, alguns parâmetros precisam ser otimizados. Para a curva experimental, o preparo de uma solução adequada para a etapa de desaceleração da incorporação de deutério (*quench*) garante a retenção da marcação isotópica, favorece a posterior digestão proteolítica e minimiza a variância de *back-exchange* entre diferentes replicatas experimentais (Gallagher e Hudgens, 2016; Wang *et al.*, 2016). A solução *quench* pode ser preparada com ácidos compatíveis com LC-MS (ácido trifluoroacético / ácido fórmico), tampões com  $pK_a \approx 2,5$  e, ainda, com desnaturantes e agentes redutores (Wang *et al.*, 2016; Masson *et al.*, 2017).

Para a análise de BJ46a e jararagina por HDX-MS, é imprescindível o uso de agentes desnaturantes e redutores na solução *quench*, pois a hidrólise destas proteínas seria comprometida na ausência destes reagentes. Inicialmente, testamos para BJ46a diferentes concentrações de guanidina-HCl (0,5 M / 1 M / 2 M / 3 M) e uma concentração fixa (0,2 M) de agente redutor, TCEP, com o objetivo de identificar qual concentração experimental geraria mais peptídeos de BJ46a. Incubamos a amostra por 30 minutos a 56°C antes da injeção no sistema, sendo o espectrômetro configurado em modo MS/MS. Após a busca e identificação dos peptídeos pelo *software* PEAKS Studio, construímos um diagrama de Venn baseado no número de peptídeos identificados para analisarmos a melhor concentração de guanidina-HCl para os testes subsequentes. Os resultados estão ilustrados na **Figura 48**.

Para BJ46a, a maior parte dos peptídeos (80 peptídeos, ou 69,6% do total) são compartilhados com as quatro concentrações de guanidina-HCl testadas. Ao aumentarmos a concentração para 1 e 2 M, não verificamos aumento do número de peptídeos; o mesmo resultado foi observado para a incubação de BJ46a com 3 M de Guanidina-HCl. Mas, a cada aumento na concentração de guanidina-HCl acima de 0,5 M, detectamos um maior número de peptídeos de pepsina oriundos da coluna de pepsina-agarose, indicando que as demais concentrações promovem aumento da autólise da enzima imobilizada. Para HDX-MS, a presença de íons oriundos de outras fontes que não a amostra aumenta o ruído do espectro e prejudica as análises (Engen

e Wales, 2015). Portanto, nas etapas subsequentes, trabalhamos com a concentração de 0,5 M de guanidina-HCl e 0,2 M de TCEP como solução *quench*.

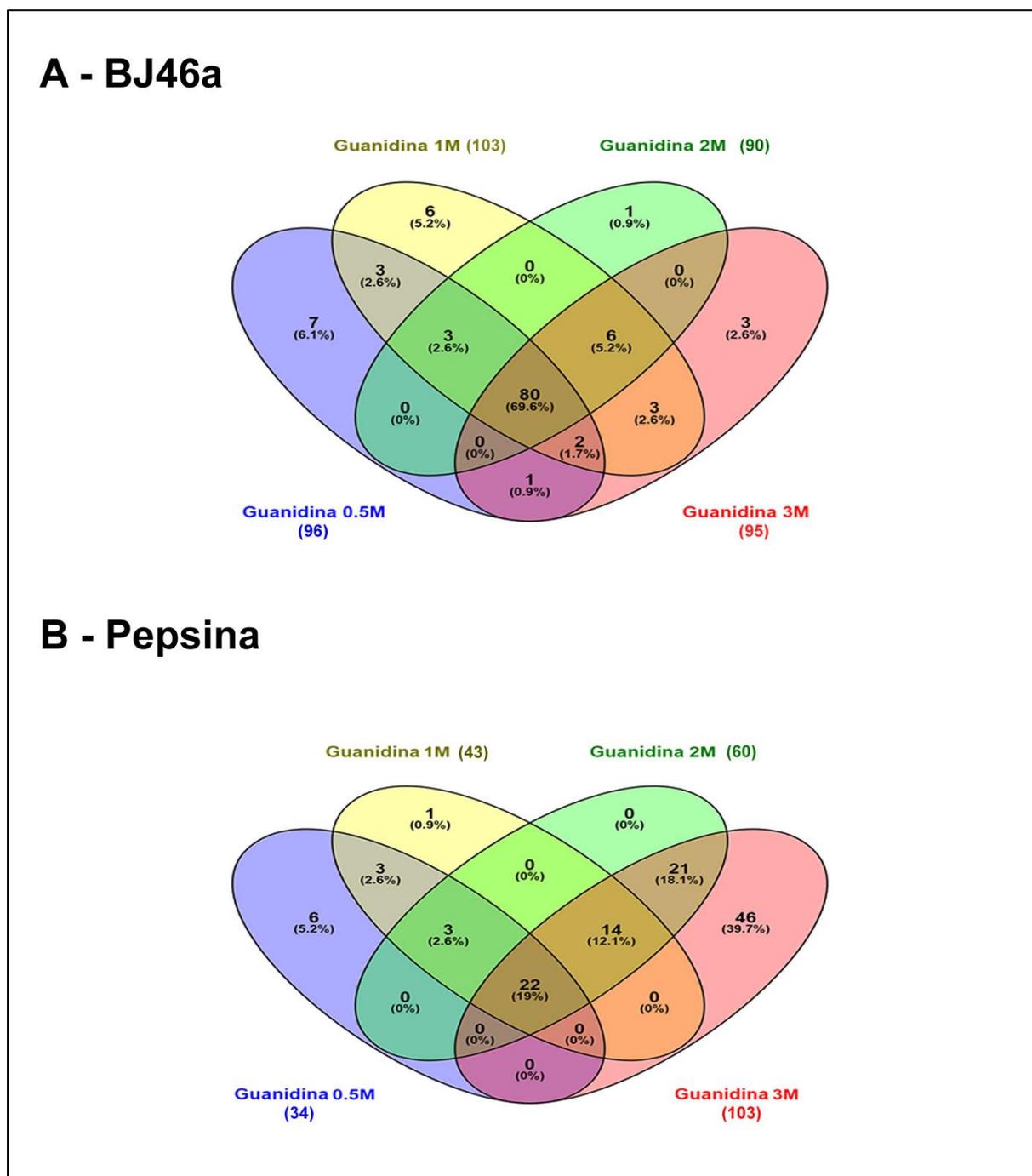


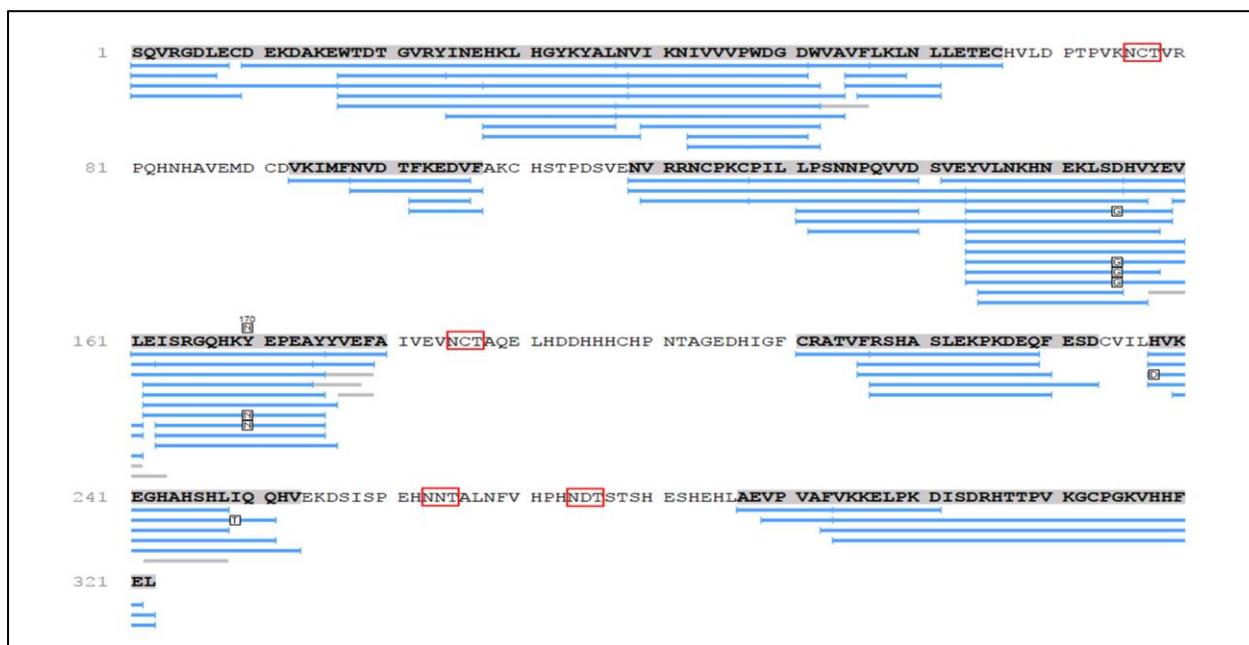
Figura 48: Diagramas de Venn relativo ao número de peptídeos identificados de BJ46a (painel A) e pepsina (painel B) em diferentes concentrações de guanidina-HCl.

Os números entre parênteses indicam o total de peptídeos detectados em cada condição experimental.

A segunda etapa de otimização do processo avaliou os parâmetros de voltagem e temperatura do capilar da fonte  $\mu$ ESI. Fontes de ionização com temperaturas um pouco mais brandas são ideais para o estudo, uma vez que menos deutério é perdido na etapa de ionização da amostra (Engen e Wales, 2015). Testamos duas condições diferentes de voltagem e temperatura de capilar: 4,5 kV e 150°C e 3,5 kV e 100°C. A solução *quench* empregada consistiu de guanidina-HCl 0,5 M, TCEP 0,2 M, com a adição de tampão fosfato de sódio 0,2 M para manter o pH da solução em 2,5 e incubação da amostra por 2 minutos a 0°C. A **Tabela 9** demonstra a cobertura de sequência e o número de peptídeos identificados para BJ46a nestas duas configurações da fonte. Não foram detectadas diferenças quanto à cobertura de sequência do inibidor, que foi mantida em 68% e ilustrada na **Figura 49**. Com exceção das regiões do inibidor que apresentam os sítios de N-glicosilação, obtivemos uma cobertura satisfatória das demais regiões, com peptídeos sobrepostos, o que é importante para análise de HDX-MS. Apesar de não verificarmos diferenças em termos de cobertura de sequência, houve uma queda (7,4%) no número de peptídeos identificados quando reduzimos a voltagem e a temperatura do capilar da fonte. É possível que essa diferença seja dada tanto pelo caráter estocástico da análise DDA, quanto condições um pouco mais brandas de ionização não foram suficientes para levar estes analitos para a fase gasosa. Mas, analisando-se a relação custo/benefício, adotamos a configuração de 3,5 kV e 100°C com vistas à maior retenção de deutério nas análises de HDX-MS.

**Tabela 9:** Cobertura de sequência e número de peptídeos detectados para o inibidor BJ46a em duas configurações da fonte de ionização  $\mu$ ESI.

| BJ46a                         | 0,5 M guanidina-HCl, 0,2 M TCEP, 0,2 M fosfato de sódio pH 2,5 |                   |
|-------------------------------|--|-------------------|
|                               | 4,5 kV, T = 150°C  | 3,5 kV, T = 100°C |
| <b>Cobertura de sequência</b> | 68%  | 68%               |
| <b>Número de peptídeos</b>    | 94   | 87                |



**Figura 49: Cobertura de sequência de BJ46a utilizando uma solução quench contendo 0,5 M guanidina-HCl, 0,2 M TCEP e 0,2 M fosfato de sódio, pH 2,5, e com a voltagem e a temperatura do capilar da fonte configurados para 3,5 kV e 100°C, respectivamente.**

Cada linha em azul, abaixo da sequência, indica um peptídeo identificado para BJ46a pela abordagem PSM (PEAKS DB) e cada linha em cinza, um peptídeo identificado por sequenciamento *de novo*. Os quatro sítios de N-glicosilação do inibidor estão realçados em vermelho.

Em seguida, analisamos a cobertura de sequência obtida para jararagina utilizando a mesma solução *quench* empregada para BJ46a, seguido de incubação por 2 minutos a 0°C. Testamos duas faixas diferentes para o espectro MS1,  $m/z$  300 – 1500 e 300 – 2000, e os resultados das buscas estão na tabela abaixo:

**Tabela 10: Cobertura de sequência e número de peptídeos identificados para jararagina utilizando a solução *quench* 0,5 M Guanidina-HCl, 0,2 M TCEP, 0,2 M fosfato de sódio pH 2,5, em duas faixas de  $m/z$  para o espectro MS1.**

| Jararagina                    | 0,5 M Guanidina-HCl, 0,2 M TCEP, 0,2 M Fosfato de Sódio pH 2,5 |                  |
|-------------------------------|--|------------------|
|                               | $m/z$ 300 - 2000   | $m/z$ 300 - 1500 |
| <b>Cobertura de sequência</b> | 50%  | 48%              |
| <b>Número de peptídeos</b>    | 83   | 81               |

Obtivemos, nestas condições, uma média de 50% de cobertura de sequência de jararagina, inferior ao que foi registrado para o inibidor BJ46a. No mapa de cobertura representativo, ilustrado na **Figura 50**, verificamos que há uma boa cobertura para a primeira metade do domínio metaloendopeptidase (1-110), mas uma baixa cobertura para os domínios desintegrina e rico em cisteína. Assumindo-se que jararagina apresente a mesma ordenação das pontes dissulfeto de botropasina, estas regiões concentram a maior parte das pontes da toxina, 13 no total (Muniz *et al.*, 2008). É possível que estas condições de desnaturação e redução não tenham sido suficientes para abrir toda a molécula e, por isso, não ocorreu uma hidrólise significativa nestas regiões, indicando que esta solução *quench* não é a adequada para analisarmos jararagina e, conseqüentemente, o complexo BJ46a - jararagina por HDX-MS.



**Figura 50: Mapa de cobertura de sequência para jararagina, na faixa de  $m/z$  300 – 2000.**

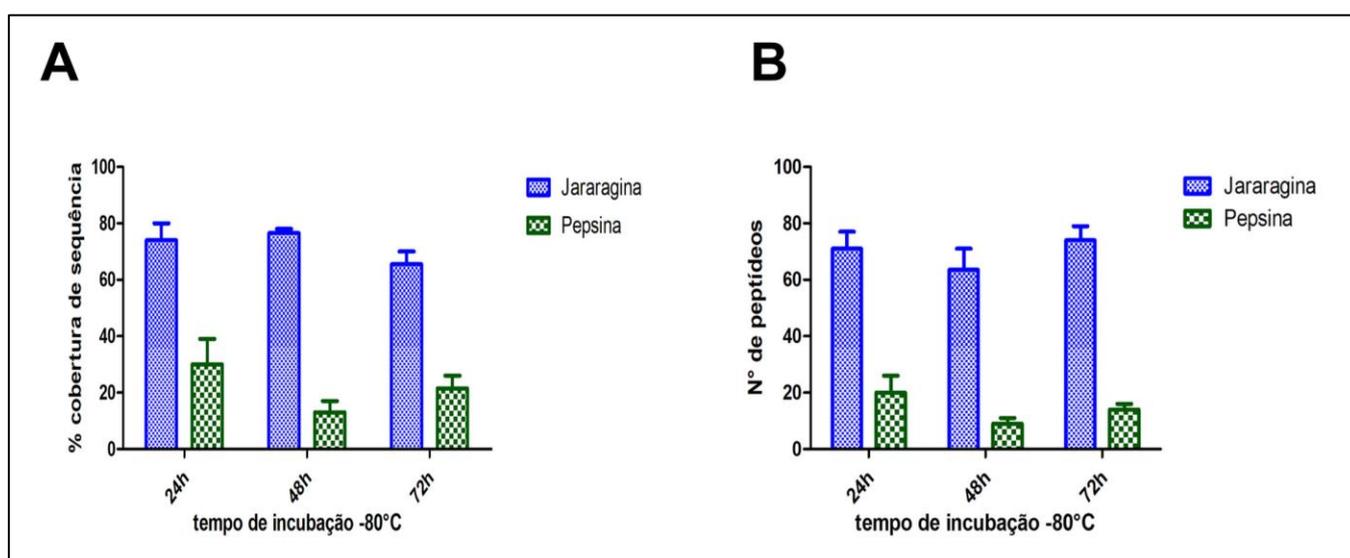
Cada linha em azul indica um peptídeo identificado para jararagina na abordagem PSM (PEAKS DB) e cada linha em cinza, um peptídeo identificado por sequenciamento *de novo*. Os peptídeos marcados com um ícone “P” em vermelho, na extremidade N-terminal da sequência, indicam peptídeos com ácido piroglutâmico ( $\Delta m = 17,03$  Da).

De fato, para que haja uma análise robusta de dados de HDX-MS, as pontes dissulfeto precisam ser reduzidas eficientemente; porém, esta redução não pode ocorrer antes ou durante a reação de incorporação de deutério, sob pena de modificação da conformação proteica e invalidação dos resultados (Zhang *et al.*, 2010). Portanto, decidimos avaliar, para BJ46a e jararagina, o método desenvolvido por Yan e colaboradores para a aquisição de dados de HDX-MS para o fator estimulante beta de colônias de macrófagos recombinante (rHM – CSF $\beta$ ), uma proteína que possui 9 pontes dissulfeto em sua estrutura (Yan *et al.*, 2002).

Neste método, ao final de cada ponto experimental da curva de incorporação de deutério, a amostra é misturada com uma solução *quench* contendo ureia a 8 M e TCEP a 1 M, pH 2,5; a própria mistura de ureia e TCEP nestas concentrações mantém

o pH ideal para a etapa de *quench*, sem a necessidade de adicionar sais para tamponar a solução. A amostra é então imediatamente congelada e incubada a  $-80^{\circ}\text{C}$  por até 72 h antes da análise. No momento da injeção no sistema de HDX-MS, a amostra é ainda incubada por 6 minutos em gelo para completar a redução das pontes dissulfeto (Yan *et al.*, 2002).

Assim, reproduzimos o método descrito para jararagina, com o objetivo de verificar se há um ganho em termos de cobertura de sequência e de número de peptídeos identificados utilizando-se esta abordagem.



**Figura 51:** Cobertura de sequência (A) e número de peptídeos (B) detectados para jararagina (barras em azul) e pepsina (barras em verde) utilizando uma solução *quench* contendo ureia 8 M e TCEP 1 M, em diferentes intervalos de incubação a  $-80^{\circ}\text{C}$ .

A **Figura 51** demonstra os resultados obtidos para jararagina neste método. Com relação à porcentagem de cobertura de sequência, demonstrada no painel A da figura, aumentamos para 74 % a cobertura já nas primeiras 24 h de incubação a  $-80^{\circ}\text{C}$ , um acréscimo de 48 % em relação à abordagem anterior. Nos tempos subsequentes, obtivemos uma média de 77 % e 66 % de cobertura para 48 h e 72 h, respectivamente. Com relação ao número de peptídeos, obtivemos uma média de 71, 64 e 74 peptídeos para 24 h, 48 e 72 h. Estes resultados iniciais indicam que, provavelmente, a maior cobertura de jararagina se dá entre 24 e 48 h de incubação a

-80°C, o que é positivo, visto que quanto menor for o tempo entre a desaceleração da incorporação de deutério (*quench*) e a análise por espectrometria de massas, menor será também a taxa de *back-exchange* (Wang et al., 2016). Realizamos um segundo experimento para análise de intervalos mais curtos de incubação, até 24 h, a -80°C (Figura 52), mas notamos um aumento significativo da cobertura de sequência de pepsina, um indicativo de que a concentração elevada de ureia e TCEP na solução *quench* degradou a coluna de pepsina, o que compromete as análises futuras.

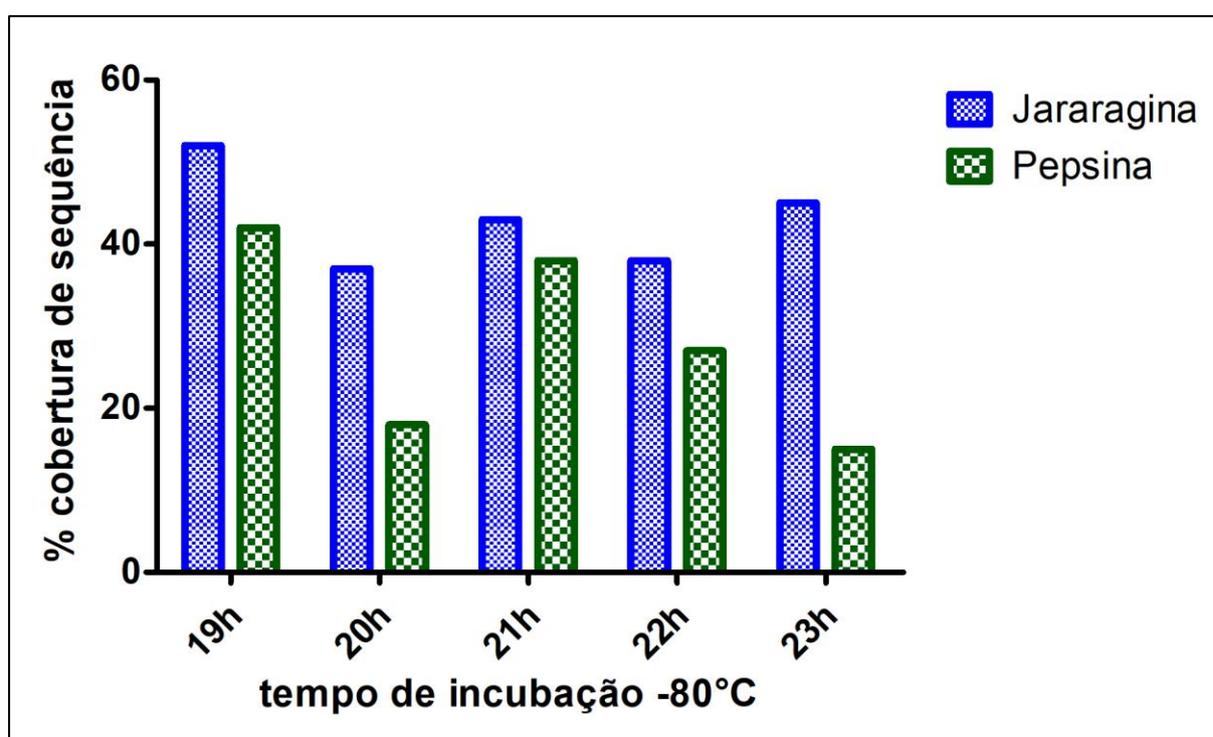
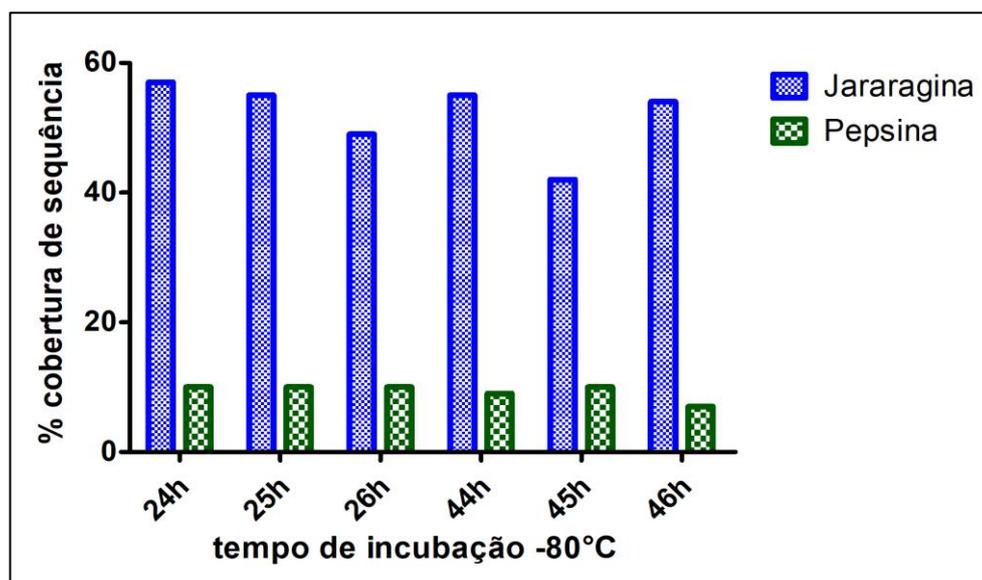


Figura 52: Cobertura de sequência para jararagina e pepsina em tempos de incubação de até 24h a -80°C.

Assim, adicionamos um passo de diluição 1:1 da amostra imediatamente antes da injeção no sistema de HDX-MS, com uma solução de ácido fórmico 0,23 % pH 2,4 resfriada a 0°C, para minimizar danos à coluna de pepsina-agarose. Realizamos então um novo experimento de incubação a -80°C com jararagina, incluindo a etapa de diluição e com uma nova coluna de pepsina-agarose, como demonstra a figura abaixo:

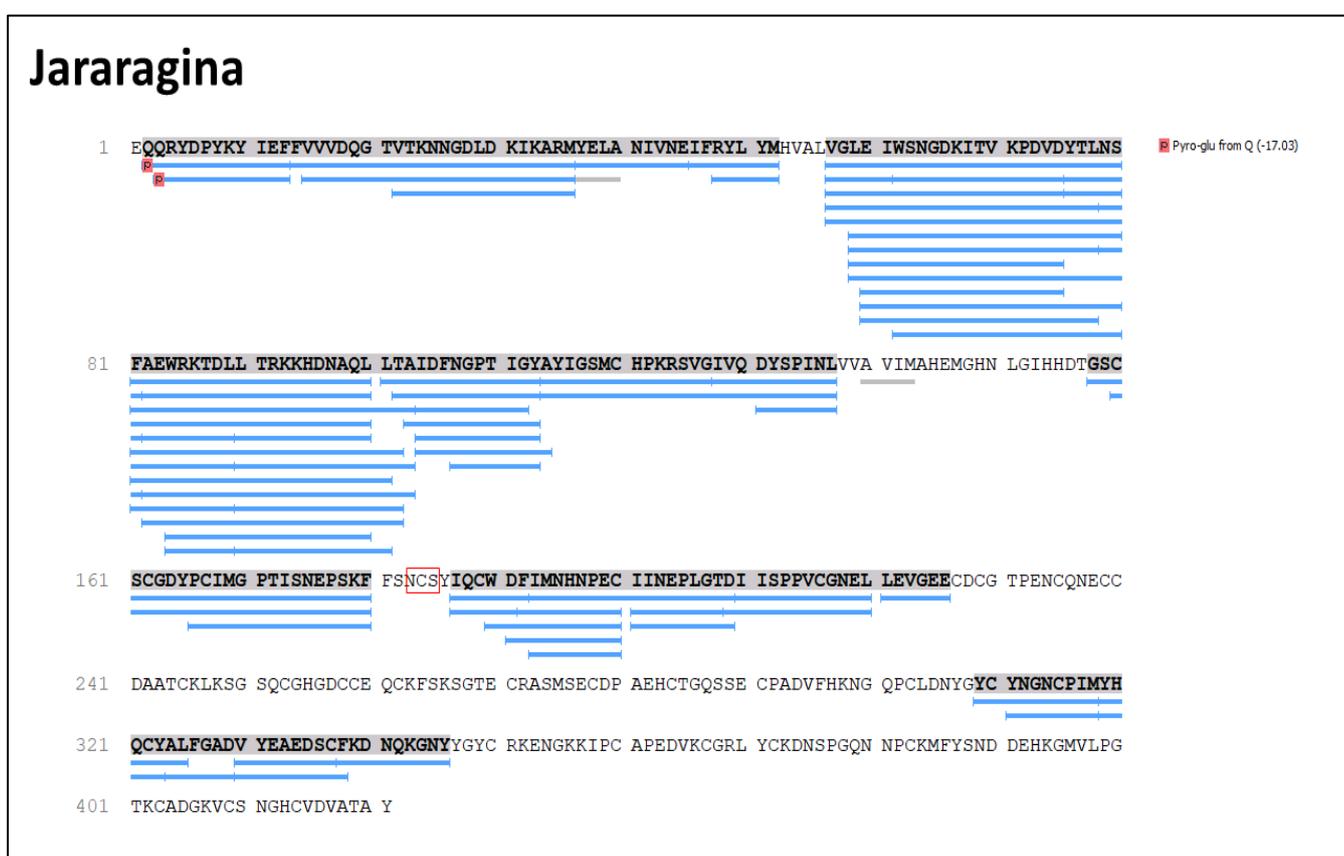


**Figura 53:** Coberturas de sequência obtidas para jararagina e pepsina em diferentes tempos de incubação e após diluição da solução *quench* imediatamente antes da injeção da amostra no sistema de HDX-MS.

Após a diluição, reduzimos significativamente a presença de pepsina nas análises; para jararagina, a cobertura de sequência atingiu 57% nas primeiras 24h de incubação, e, nos tempos subsequentes, não ocorreu um aumento significativo da cobertura, demonstrando que 24h é o intervalo adequado para a incubação das amostras. Em seguida, realizamos uma série de aquisições em diferentes intervalos de  $m/z$  (conforme descrito no capítulo 3, subitem 3.2.20.1) para a construção da lista de peptídeos para a posterior análise das curvas de deuteração. Nestas condições finais de otimização, obtivemos 55% de cobertura para a toxina jararagina, e 63% de cobertura para o inibidor BJ46a. Os mapas finais de cobertura de sequência estão ilustrados na **Figura 54** e na **Figura 55**.

Para jararagina, obtivemos mais peptídeos sobrepostos na parte final do domínio metaloendopeptidase (resíduos 161-205) e para a parte inicial do domínio desintegrina (resíduos 206 – 215) e em parte do domínio rico em cisteína (resíduos 321 – 345). Porém, mesmo com o protocolo otimizado, ainda há partes do domínio desintegrina e rico em cisteína sem cobertura de sequência, o que pode estar correlacionado com a especificidade da enzima pepsina em condições de HDX-MS, que cliva preferencialmente onde há resíduos de fenilalanina e leucina na posição P1;

resíduos de histidina, lisina, arginina e prolina nesta posição proíbem a clivagem. Resíduos de prolina também não podem assumir a posição P2 (Hamuro *et al.*, 2008), e estas regiões apresentam poucos sítios de clivagem para pepsina nestas condições. Já para o inibidor BJ46a, obtivemos uma cobertura de sequência com uma boa quantidade de peptídeos sobrepostos, excetuando-se as regiões que apresentam sítios de N-glicosilação; em muitas proteínas, a extensão, ramificação e carga dos glicanos N-ligados conferem resistência à proteólise (Costa *et al.*, 2014).



**Figura 54: Cobertura de sequência obtida para jararagina na etapa final da otimização do protocolo de HDX-MS.**

Cada linha em azul indica um peptídeo identificado para jararagina na abordagem PSM (PEAKS DB) e cada linha em cinza, um peptídeo identificado por sequenciamento *de novo*. Os peptídeos marcados com um ícone “P” em vermelho, na extremidade N-terminal da sequência, indicam peptídeos com ácido piroglutâmico ( $\Delta m = -17,03$  Da). O sítio de N-glicosilação da toxina está realçado por um retângulo vermelho.

## BJ46a



**Figura 55: Cobertura de sequência obtida para BJ46a na etapa final da otimização do protocolo de HDX-MS.**

Cada linha em azul indica um peptídeo identificado para BJ46a na abordagem PSM (PEAKS DB) e cada linha em cinza, um peptídeo identificado por sequenciamento *de novo*. Os sítios de N-glicosilação estão realçados por um retângulo vermelho.

#### 4.4.3 – Curva experimental de incorporação de deutério

Para a curva experimental de incorporação de deutério, o complexo entre BJ46a e jararagina, bem como os controles do inibidor e toxina livres, foram incubados por 15 minutos à 25°C; a amostra foi então transferida para um tampão preparado com D<sub>2</sub>O, conforme descrito no capítulo 3, subitem 3.2.19. Alíquotas das amostras deuteradas foram retiradas nos intervalos de tempo de 10, 100, 1000 e 10.000 s, rapidamente misturadas com uma solução *quench* de ureia a 8 M, TCEP 1 M a 0°C, congeladas em N<sub>2</sub> líquido e incubadas por 24 h a -80°C antes da análise por espectrometria de massas, utilizando o sistema de HDX-MS descrito no subitem 4.4.1 deste capítulo.

A **Figura 56**, **Figura 57** e a **Figura 58** demonstram os cromatogramas de íons obtidos para a análise de BJ46a livre, jararagina livre e o complexo BJ46a – jararagina, respectivamente. O perfil cromatográfico de cada amostra, em diferentes tempos de exposição ao deutério, manteve-se constante, assim como os tempos de retenção cromatográfica de cada pico, o que é importante para a análise dos dados pelo *software* específico de HDX-MS, o HDExaminer.

Para a etapa inicial da análise das curvas de deuturação, o *software* deriva, a partir de uma lista experimental previamente construída de peptídeos de BJ46a e jararagina, envelopes isotópicos teóricos para cada entrada da lista. O algoritmo busca então, nos arquivos brutos (*.raw files*) correspondentes, o envelope experimental melhor sobreposto ao envelope calculado e que esteja próximo ao tempo de retenção experimentalmente determinado da lista. Se as sobreposições forem adequadas, o algoritmo extrai as centroides relativas a cada envelope experimental e, tendo a centroide relativa ao tempo de 0 s e o valor máximo de deuturação teórica por peptídeo (calculado com base na porcentagem final de D<sub>2</sub>O na solução de marcação) como referências, o *software* calcula os níveis de deuturação para cada tempo experimental, por peptídeo detectado.

A partir destes cálculos, podemos criar um mapa de deuturação (*heatmap*), que demonstra, para cada região proteica, os níveis de deuturação ao longo de toda a estrutura primária, mapeada nos diferentes tempos de exposição ao deutério. Tais representações fornecem uma análise global dos resultados de HDX-MS (Sarpe e Schriemer, 2016). Além disso, podemos derivar curvas de incorporação que correlacionam o nível de deuturação de cada peptídeo ao tempo de exposição a D<sub>2</sub>O

em diferentes condições experimentais, o que é útil para mapear regiões diferenciais de incorporação de deutério.

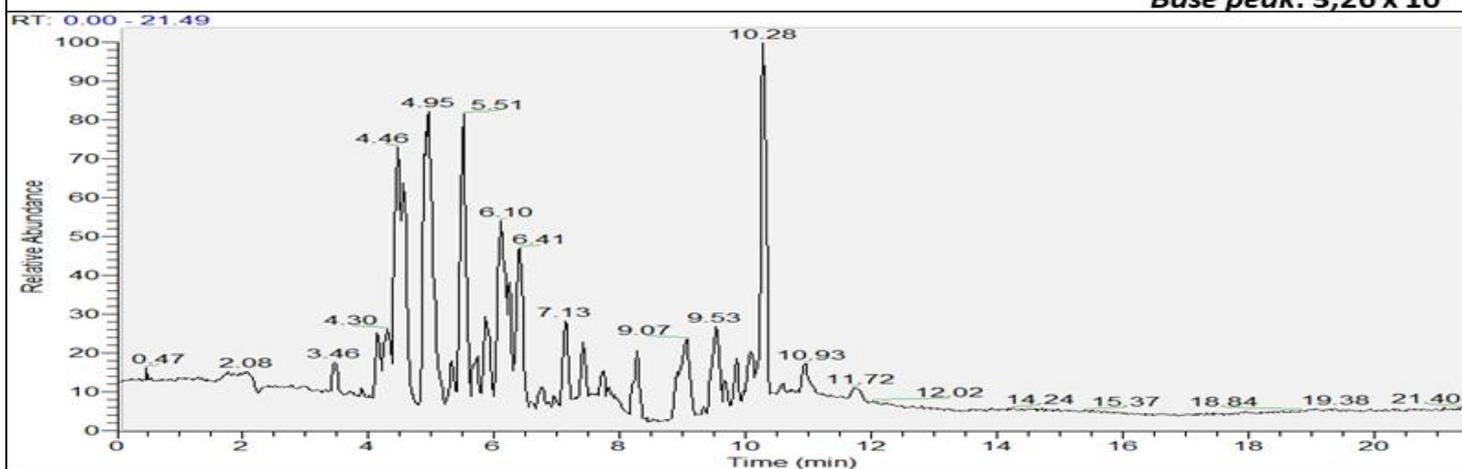
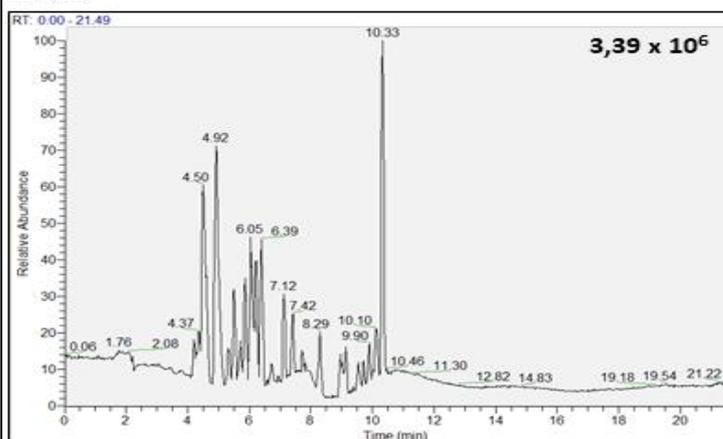
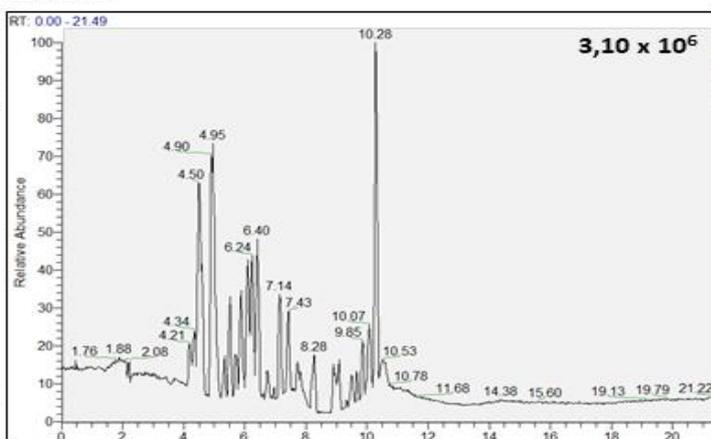
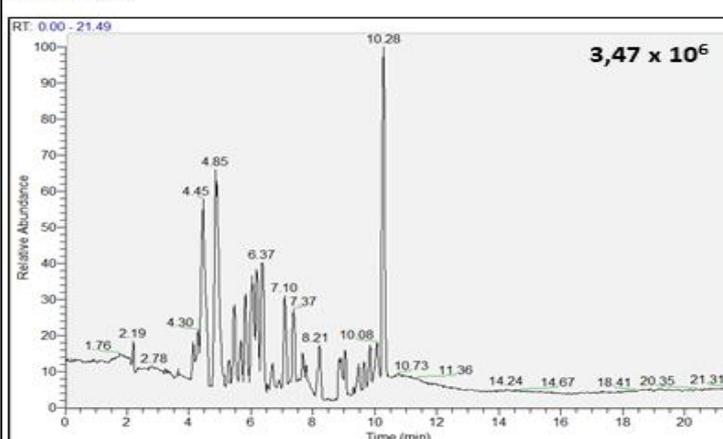
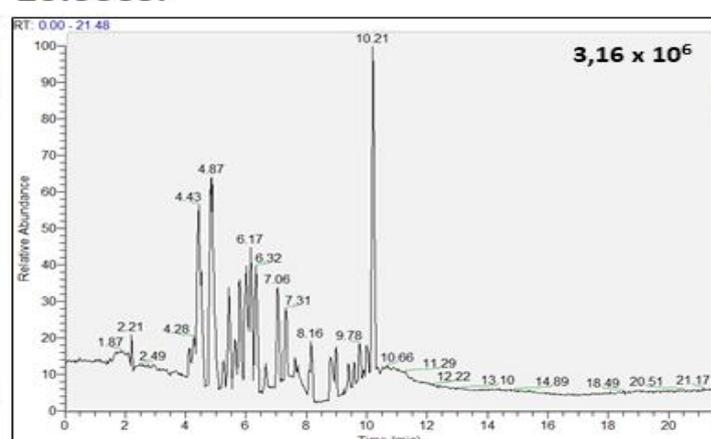
**0s:****Base peak:  $3,26 \times 10^6$** **10s:****100s:****1000s:****10.000s:**

Figura 56: Cromatogramas de íons representativos obtidos para a análise da curva de deuteração de BJ46a livre.

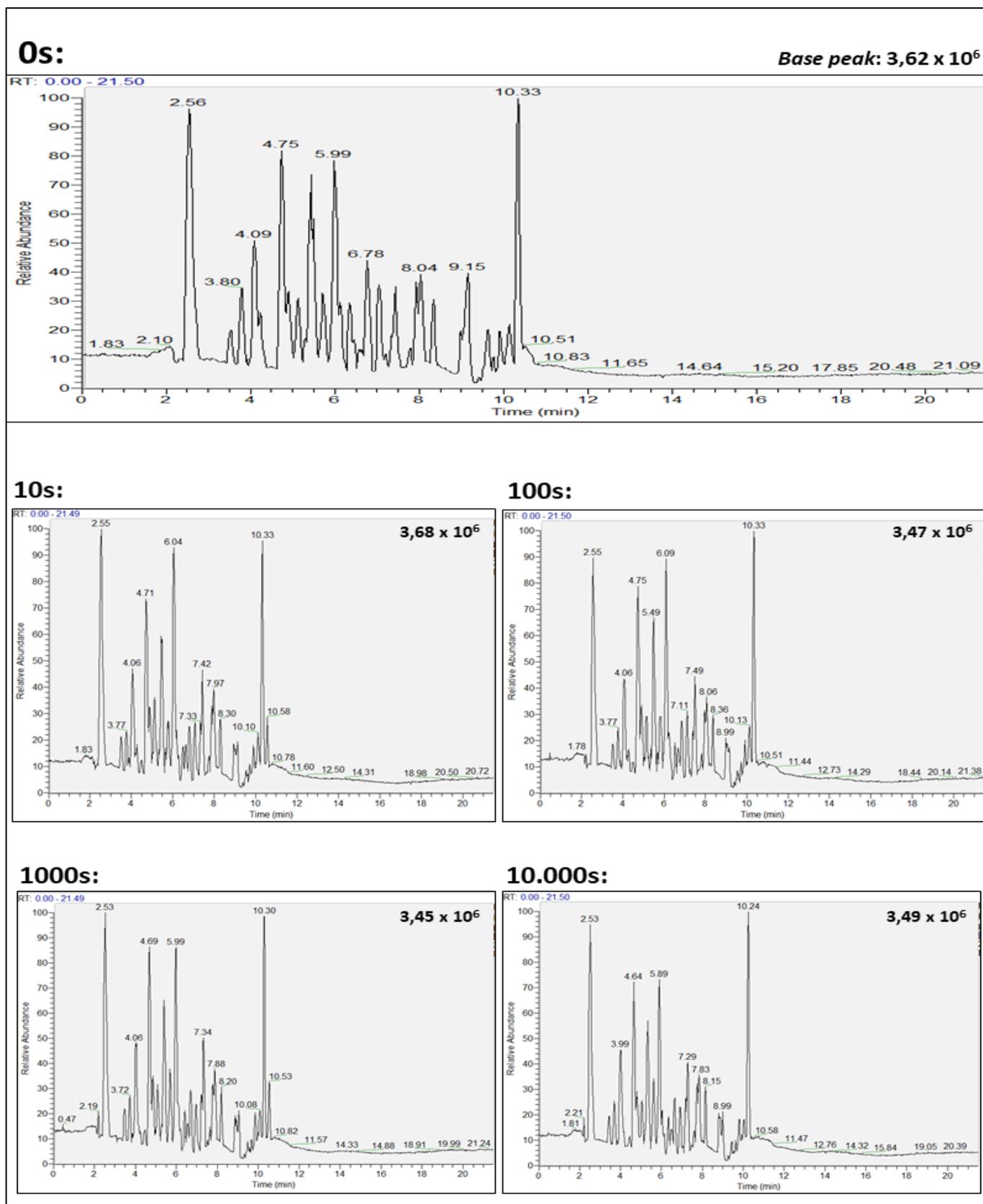
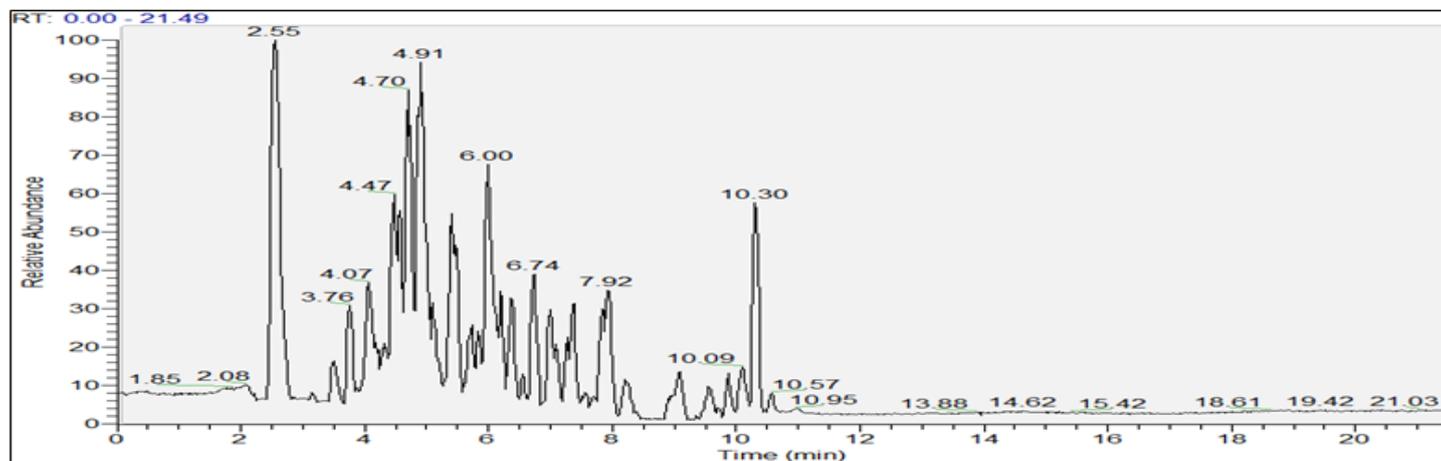
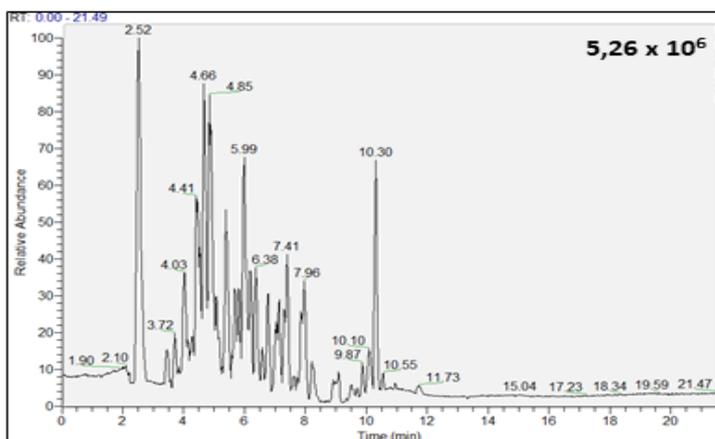


Figura 57: Cromatogramas de íons representativos obtidos para a análise da curva de deuteração de jararagina livre.

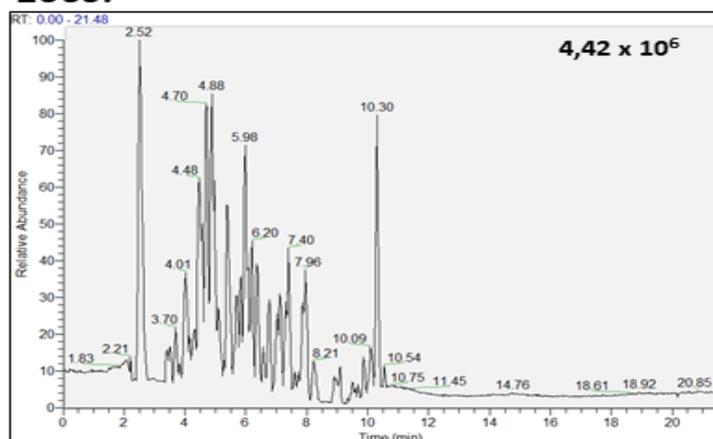
0s:

Base peak:  $5,27 \times 10^6$ 

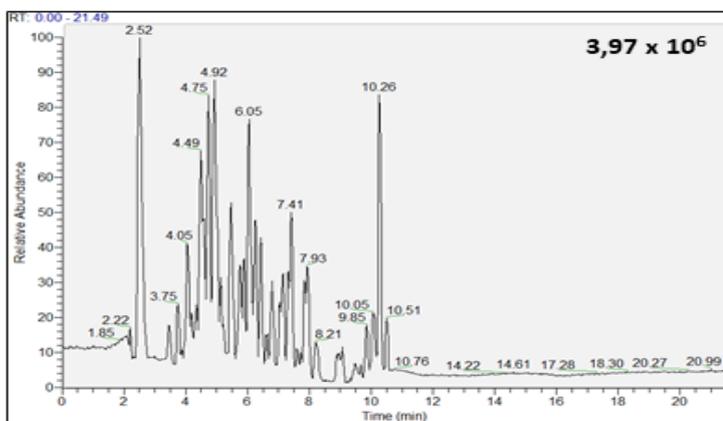
10s:

 $5,26 \times 10^6$ 

100s:

 $4,42 \times 10^6$ 

1000s:

 $3,97 \times 10^6$ 

10.000s:

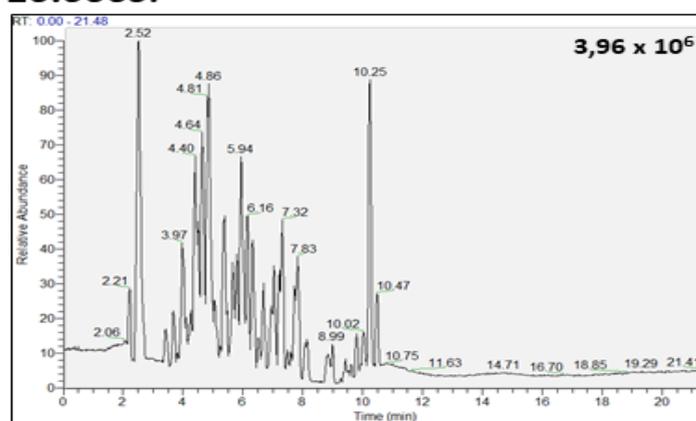
 $3,96 \times 10^6$ 

Figura 58: Cromatogramas de íons representativos obtidos para a análise da curva de deuteração do complexo BJ46a - jararagina.

A **Figura 59** ilustra o *heatmap* obtido para o inibidor BJ46a em seu estado livre, em cada um dos tempos de exposição a deutério. Com relação à cobertura de sequência (barras localizadas na parte superior da estrutura primária) obtivemos informação de incorporação isotópica para os dois domínios cistatina (resíduos 3-111 e 122 – 235) e para parte do domínio rico em histidina (236-322), com exceção das regiões que apresentam glicanos N-ligados, como já havia sido demonstrado nas etapas de otimização das condições experimentais de HDX-MS.

Com relação à incorporação isotópica, ao observarmos o *heatmap* conseguimos distinguir regiões com alta incorporação isotópica (> 80%, regiões marcadas em laranja e vermelho), média incorporação (entre 40 – 60%, em amarelo e verde) e regiões de baixa incorporação (< 30%, em azul). Conforme discutido no capítulo 1, a troca H-D em amidas do *backbone* proteico requer a abertura transiente da estrutura local para que haja a incorporação. Dessa forma, a taxa de incorporação isotópica em uma dada região proteica é influenciada por parâmetros como a acessibilidade ao solvente, redes de ligação de hidrogênio e efeitos indutivos e estéricos (Mc Allister e Konermann, 2015).

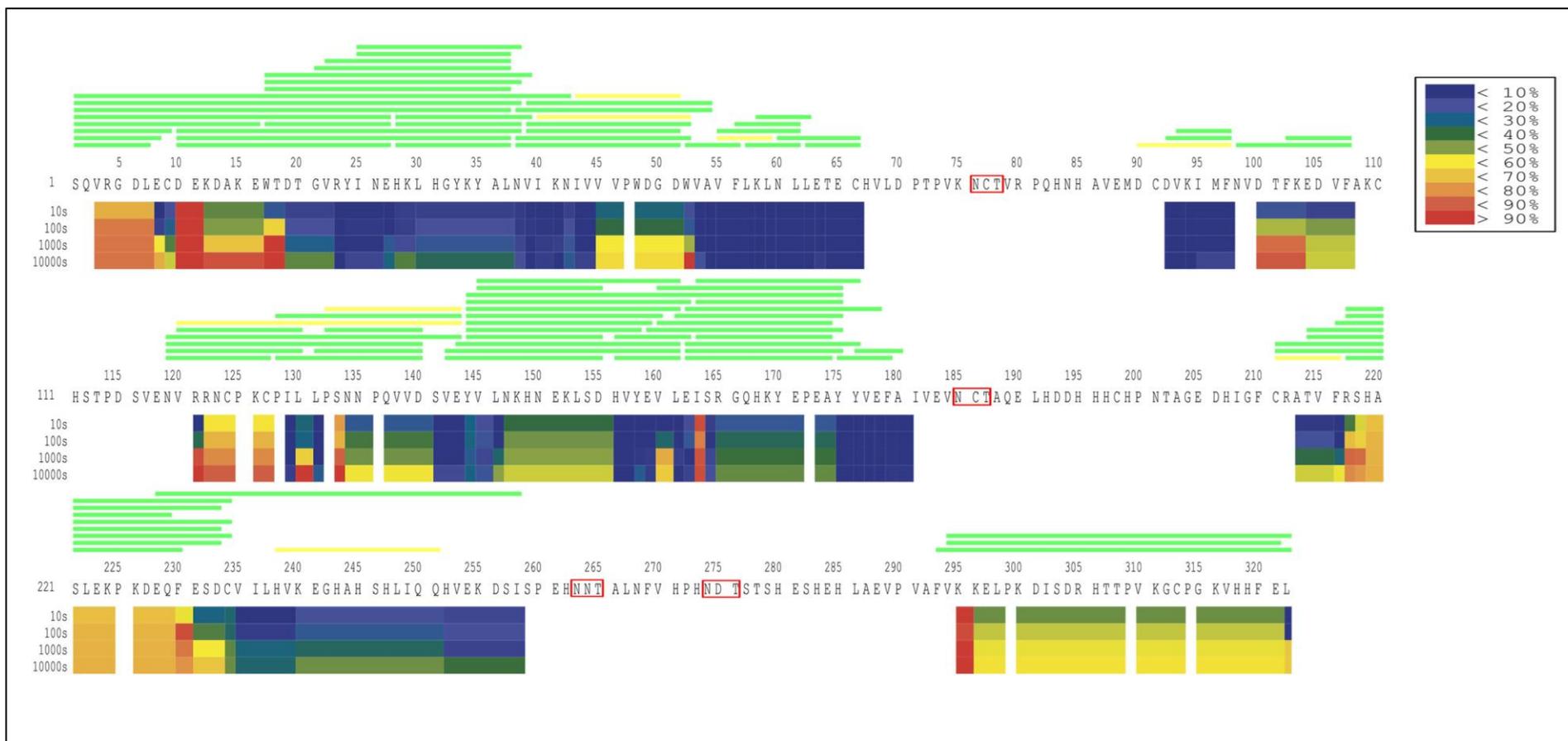
Para o primeiro domínio cistatina de BJ46a, mapeamos regiões com alta taxa de incorporação isotópica já em 10 s de exposição a D<sub>2</sub>O, localizadas na parte inicial da extremidade N-terminal do inibidor. Akashi e Takio aplicaram a técnica de HDX-MS ao estudo de uma cistatina purificada de ovo de galinha e observaram o mesmo comportamento para a porção N-terminal da proteína, com mais de 65% dos hidrogênios amídicos substituídos por deutério em 1 min de exposição a D<sub>2</sub>O (Akashi e Takio, 2000). Análises por RMN para esta mesma cistatina indicaram que os oito primeiros resíduos N-terminais são desestruturados (Dieckmann *et al.*, 1993). Visto que BJ46a apresenta o mesmo comportamento de troca isotópica nesta região, é possível que a flexibilidade inerente desta região faça com que os hidrogênios amídicos troquem mais rapidamente do que se estivessem localizados em regiões estruturadas, pois regiões desordenadas passam mais tempo em conformações que favorecem a desproteção dos hidrogênios amídicos (Weis, 2016).

A partir de uma análise subtrativa, conseguimos sublocalizar as regiões de incorporação do inibidor em sua forma livre, valendo-se dos níveis de incorporação de deutério no ponto de 10.000s para todos os peptídeos sobrepostos de cada região (Burns-Hamuro *et al.*, 2005). A **Figura 60** demonstra o resultado desta análise.

Localizamos pequenos trechos correspondentes ao primeiro domínio cistatina (23 – 27, 38 – 44, 53-66, 92-97) e no 2º domínio cistatina (131-133 / 144-146 / 156 – 159 / 161-162 / 176-180) onde não ocorreu nenhuma incorporação isotópica; provavelmente, os hidrogênios amídicos desta região podem estar envolvidos nas redes de ligação de hidrogênio que estabilizam o enovelamento característico de domínios tipo cistatina, composto de cinco fitas  $\beta$  antiparalelas cruzadas por uma  $\alpha$ -hélice (Nandy *et al.*, 2012).

Ainda assim, o padrão de incorporação isotópica obtido para o 1º domínio cistatina é ligeiramente diferente do observado para o 2º, onde comparativamente há mais regiões com maior incorporação de deutério (**Figura 59**). Isso pode indicar que, apesar dos dois domínios serem considerados do tipo cistatina – portanto compartilham de um mesmo enovelamento proteico global – estas diferenças no perfil de incorporação isotópica levantam a hipótese de que os domínios possam apresentar algumas particularidades estruturais e/ou de dinâmica que podem resultar em mudanças na acessibilidade ao solvente e em redes de ligações de hidrogênio, por exemplo, e isso se reflete no dado de HDX-MS.

Já para o domínio rico em histidina, o *heatmap* demonstra que já em 10 s de exposição a D<sub>2</sub>O há uma incorporação de 50%; em 100s, salta para 60% e este nível de deuteração se mantém constante nos tempos subsequentes de 1000 e 10.000s, indicando que, assim como o N-terminal, há também uma incorporação de deutério significativa na porção C-terminal de BJ46a. Entretanto, os dados disponíveis na literatura sobre este domínio em particular são escassos, não sendo possível correlacionar este perfil de incorporação isotópica com dados estruturais.



**Figura 59: Mapa de deuteração (*heatmap*) de BJ46a livre.**

Os níveis de deuteração do inibidor em diferentes tempos de exposição a deutério (10, 100, 1000 e 10.000 s) estão representados ao longo de toda a estrutura primária de BJ46a como barras coloridas, onde cada cor expressa um nível de deuteração específico. Os peptídeos detectados ao longo de toda a sequência do inibidor estão representados como barras em verde (alta confiança) e amarelas (média confiança). Os sítios de N-glicosilação de BJ46a estão realçados em retângulos vermelhos.

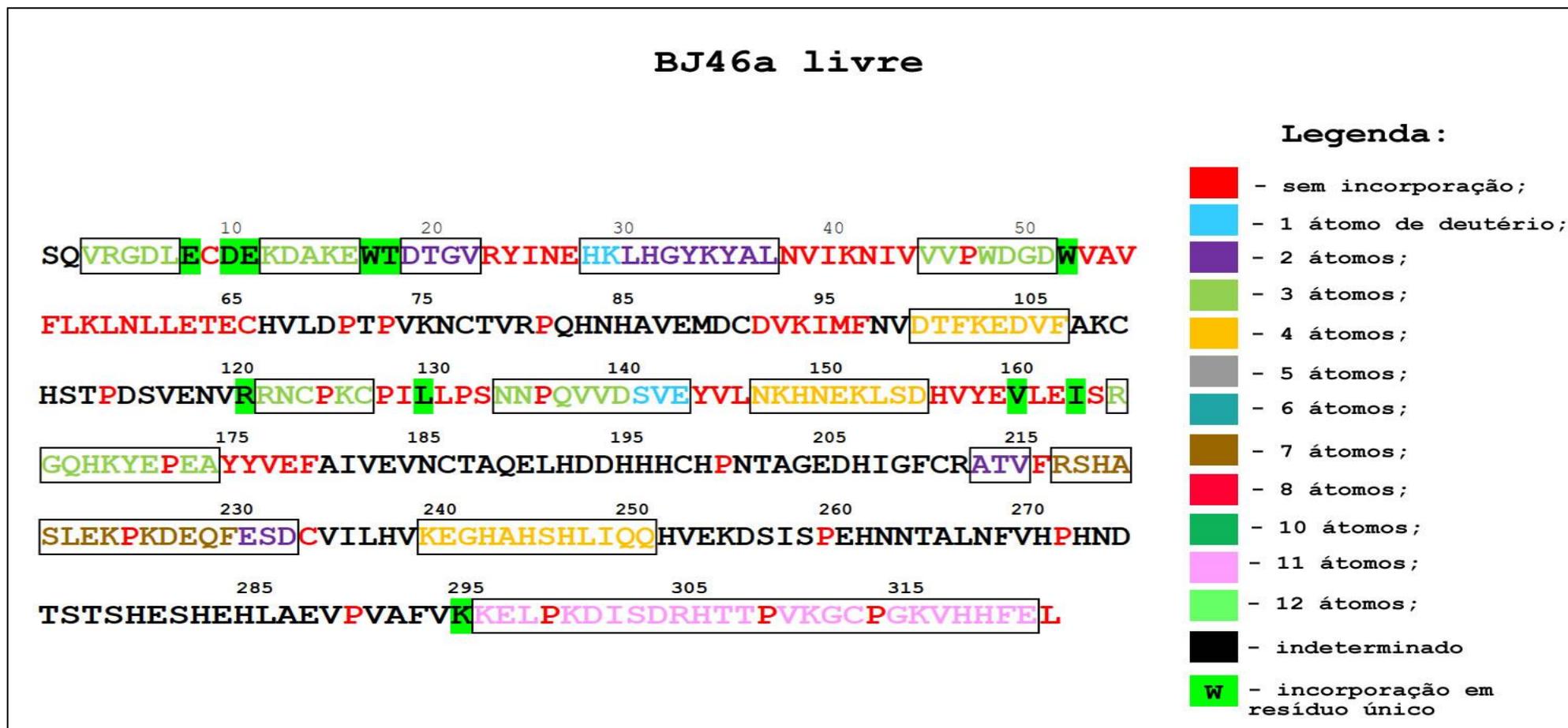


Figura 60: Análise subtrativa da incorporação isotópica em BJ46a livre ao final de 10.000s de exposição a D<sub>2</sub>O.

Nesta análise, todas as regiões onde obtivemos informação de incorporação isotópica estão realçadas em retângulos pretos, e a quantidade de átomos de deutério incorporado por região estão coloridas de acordo com o esquema. Os resíduos de prolina estão marcados em vermelho pois não há incorporação isotópica nestes resíduos.

Para a SVMP jararagina, construímos um *heatmap* (**Figura 61**) e realizamos a análise subtrativa (**Figura 62**) da mesma forma que realizamos para o inibidor BJ46a em sua forma livre. Ao analisarmos o *heatmap* de jararagina, verificamos regiões de proteção e exposição a deutério ao longo de toda a sequência da toxina.

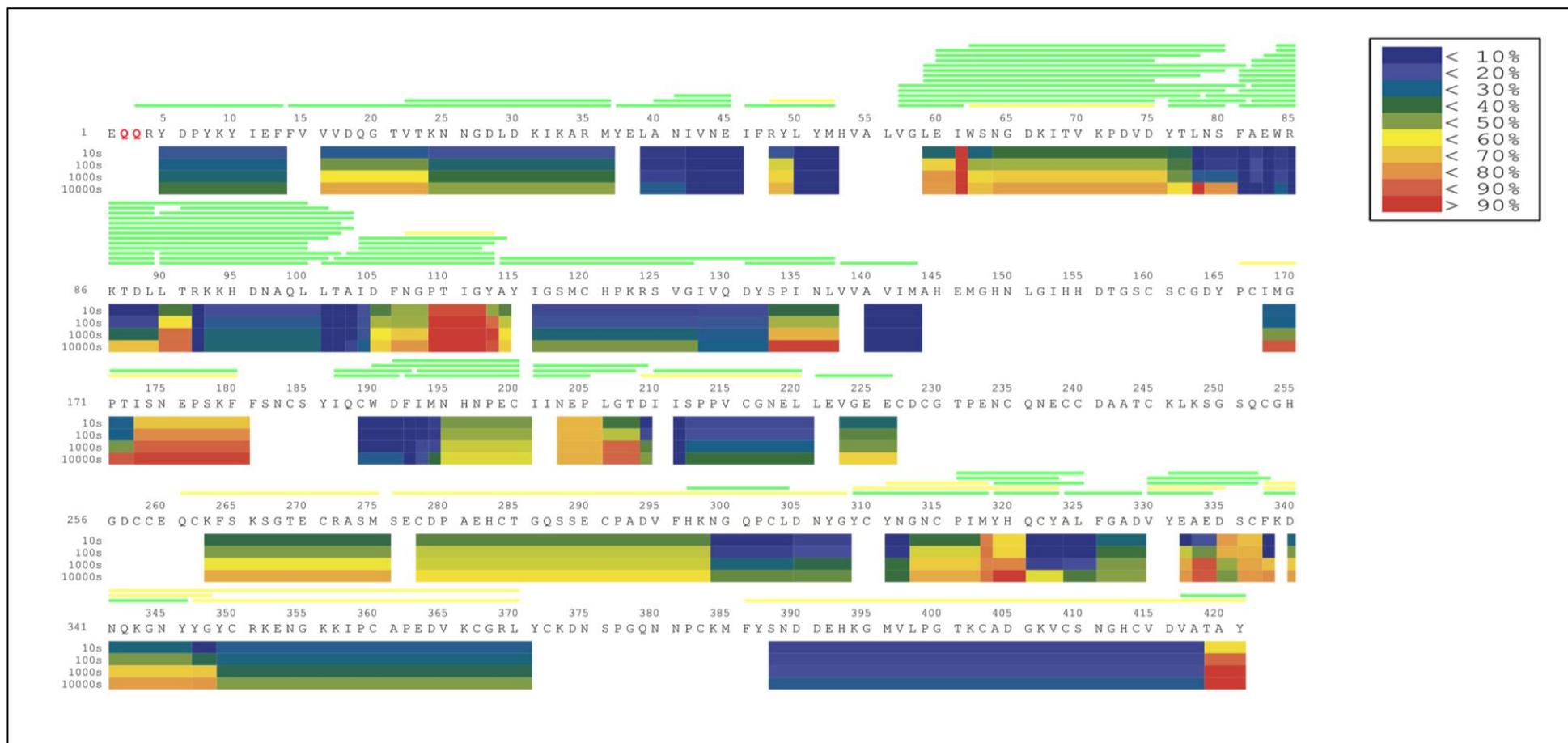
O domínio metaloendopeptidase das SVMP apresenta uma estrutura tridimensional conservada, composta de cinco fitas  $\beta$  paralelas e uma antiparalela centrais ao domínio e cinco hélices, como demonstrado para botropasina, na **Figura 36**. Nos primeiros 10 s de exposição a  $D_2O$ , a maior parte do domínio está com baixos níveis de incorporação isotópica ( $< 30\%$ ), mas já há alguns pequenos trechos da sequência que exibem incorporações maiores de deutério ( $> 30\%$ , regiões 60 – 77, 90 – 92, 105 – 115, 133 – 137, 173 – 181, 195 – 201, 203 – 209). No modelo cristalográfico de botropasina (PDB ID: 3DSL, 95% de identidade sequencial para jararagina), estas regiões correspondem majoritariamente a *loops* no domínio exposto ao solvente e porções terminais de fitas  $\beta$  e  $\alpha$ -hélices. Uma vez que HDX-MS reporta a estabilidade e dinâmica de elementos de estrutura secundária, é possível que estas partes de folhas  $\beta$  e  $\alpha$ -hélices tenham uma dinâmica diferenciada do restante da fita/hélice (Konermann, Lars *et al.*, 2011). Ao final da curva de incorporação isotópica, em 10.000 s, regiões inicialmente protegidas em 10 s de exposição sofrem o evento de exposição do hidrogênio amídico e passam a incorporar deutério, como visto nos trechos 16 – 37, 86 – 90 e 116 – 127. Há ainda regiões extremamente protegidas da troca, que não incorporaram deutério mesmo após 10.000s, como os trechos 39 – 45, 50 – 52, 81 – 85, 101 – 105, 140 – 144 (**Figura 61** e **Figura 62**). Provavelmente, estas partes do domínio estão envolvidas em uma rede estável de ligações de hidrogênio, e, portanto, o fator de proteção destes hidrogênios amídicos é elevado. Interessante notar que a região 140-144 é espacialmente próxima do domínio de ligação ao zinco, que é essencial para a catálise enzimática de jararagina (Moura-Da-Silva e Baldo, 2012). A coordenação do átomo de zinco é dada por três resíduos de histidina (H145, H149 e H155), que se orientam espacialmente a partir de um arcabouço estrutural dado por uma alça conservada (*Met-turn*) (Laing e Moura-Da-Silva, 2005). Portanto, a não-incorporação de deutério nessa região é compatível com um enovelamento estável da molécula, uma vez que estas apresentam taxas muito baixas de incorporação isotópica (Konermann, Lars *et al.*, 2011).

Para o domínio desintegrina de jararagina (resíduos 206 a 297) mapeamos uma região de troca mais lenta, com pouca incorporação de deutério (resíduos 211 – 220, 2 átomos de deutério incorporados após 10.000s de exposição) e duas regiões de troca mais rápida (264 – 275, 7 átomos de deutério incorporados e 278 – 297, 12 átomos). Na estrutura cristalográfica de botropasina, não foi possível identificar elementos de estrutura secundária para o domínio desintegrina; o domínio é essencialmente composto de *loops* estabilizados por pontes dissulfeto e dois sítios de ligação para íons  $\text{Ca}^{2+}$ , sendo um íon coordenado pelos resíduos V215, L220, E222 e D228, e outro pelos resíduos D279, D294, E282 e V295 (Muniz *et al.*, 2008). Estas particularidades do domínio desintegrina podem fazer com que a região seja estruturalmente rígida, o que minimizaria a taxa de troca isotópica (Mc Allister e Konermann, 2015; Takeda, 2016).

Porém, nossos dados demonstram uma incorporação de deutério significativa nesta região; é possível que o domínio desintegrina esteja exposto ao solvente, e com isso, a rede de ligações de hidrogênio provavelmente é o fator determinante na incorporação. O estudo das ligações de hidrogênio desta região em botropasina indicou que há uma densidade elevada de ligações de hidrogênio neste domínio, mas que são, em sua maioria, bifurcadas – ligações que apresentam mais de um doador ou acceptor e que são mais fracas, apresentando entre 50 a 60% da força de uma ligação canônica (Muniz *et al.*, 2008; Feldblum e Arkin, 2014). Ainda que não haja na literatura nenhum trabalho que avalie o impacto de ligações de hidrogênio bifurcadas na incorporação isotópica, a presença destas ligações parece não ter impedido a troca hidrogênio-deutério no domínio desintegrina de jararagina.

O domínio rico em cisteína de jararagina (resíduos 298-421) apresentou um perfil de incorporação isotópica com pequenas regiões de exposição e uma região maior de proteção (resíduos 387 – 418). Estruturalmente, em botropasina, este domínio é composto de duas pequenas  $\alpha$ -hélices e quatro fitas  $\beta$  e *loops*, estabilizados por pontes dissulfeto e ligações de hidrogênio bifurcadas, da mesma forma que o domínio desintegrina; em particular, a região 388-396 apresenta uma rede intrincada de ligações de hidrogênio que confere uma estabilidade adicional à região (Muniz *et al.*, 2008). Dessa forma, a região de proteção nos resíduos 387- 418 é compatível com o descrito para botropasina. Com relação às pequenas regiões de desproteção

detectadas no domínio, é possível que estas correspondam aos *loops* contidos no domínio.



**Figura 61: Mapa de deuteração (*heatmap*) de jararagina livre.**

Os níveis de deuteração da toxina em diferentes tempos de exposição a deutério (10, 100, 1000 e 10.000 s) estão representados ao longo de toda a estrutura primária de jararagina como barras coloridas, onde cada cor expressa um nível de deuteração específico. Os peptídeos detectados ao longo de toda a sequência da toxina estão representados como barras em verde (alta confiança) e amarelas (média confiança).

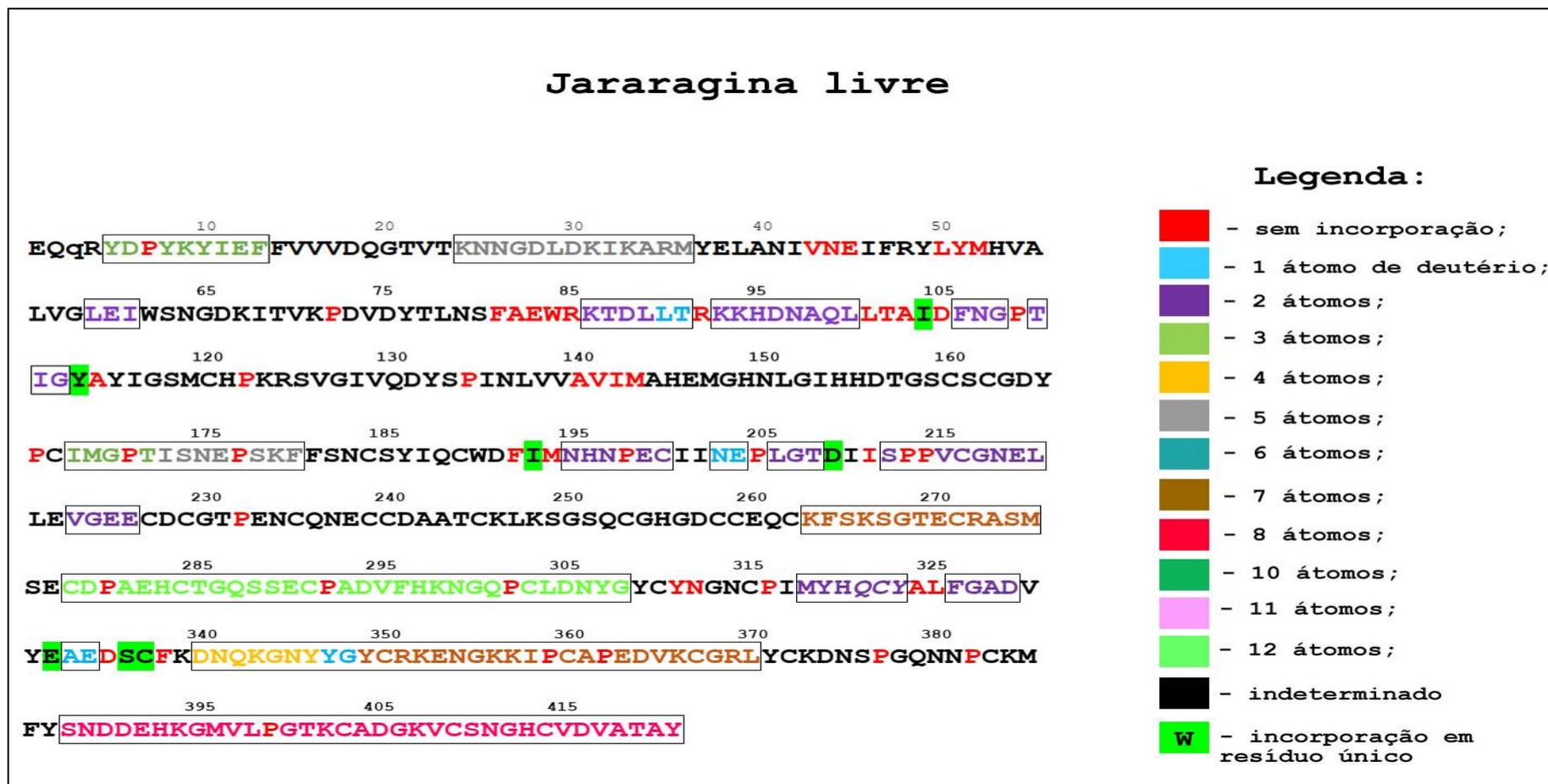


Figura 62: Análise subtrativa da incorporação isotópica em jararagina livre, ao final de 10.000s de exposição a D<sub>2</sub>O.

Nesta análise, todas as regiões onde obtivemos informação de incorporação isotópica estão realçadas em retângulos pretos, e a quantidade de átomos de deutério incorporado por região estão coloridas de acordo com o esquema. Os resíduos de prolina estão marcados em vermelho pois não há incorporação isotópica nestes resíduos.

Após a análise da incorporação isotópica de BJ46a e jararagina livres, prosseguimos para a análise do perfil de incorporação destas proteínas quando em complexo, por meio de gráficos residuais, que evidenciam as diferenças na incorporação isotópica ao longo de toda a sequência, para cada tempo de exposição a D<sub>2</sub>O. Para a construção deste gráfico, o conjunto de dados de BJ46a e jararagina nos seus estados livres e em complexo foram submetidos ao teste t de Student; apenas as mudanças estatisticamente significativas (p-valor < 0,05) foram consideradas.

Em estudos de interação proteína-proteína por HDX-MS, há três principais cenários que podemos observar em mapas comparativos de incorporação isotópica, descritos por Konermann e colaboradores (Konermann *et al.*, 2014). Em um cenário do tipo 1, há uma redução na incorporação isotópica de uma dada região quando comparado ao controle, devido à formação de ligações de hidrogênio intermoleculares e/ou pela interface de interação, que pode reduzir a acessibilidade ao solvente e, conseqüentemente, as taxas observadas de incorporação ( $k_{\text{HDX}}$ ). No cenário do tipo 2, a interação proteína-proteína promove uma maior deuteração na região, oriunda de rearranjos estruturais para acomodação do ligante que expõem os grupamentos N-H ao solvente. Por fim, há o cenário do tipo 0, no qual não há mudança alguma no perfil de incorporação isotópica; como HDX-MS mede primariamente a incorporação no *backbone* proteico, complexos proteicos que são mediados por interações eletrostáticas e/ou hidrofóbicas, em cadeias laterais de regiões estruturadas não refletirão mudanças no *backbone* (Konermann *et al.*, 2014; Kochert *et al.*, 2018).

É importante ressaltar que estes cenários não são mutuamente excludentes: os três cenários podem coexistir em uma mesma interação proteína-proteína e são tidos como cenários híbridos (Konermann *et al.*, 2014). Para a interação de BJ46a e jararagina, verificamos um cenário híbrido, onde há regiões de proteção e desproteção tanto para o inibidor quanto para a toxina (**Figura 63** e **Figura 68**).

Em BJ46a, mapeamos regiões diferenciais de incorporação no primeiro e no segundo domínio cistatina; o domínio rico em histidina não sofreu diferenças estatisticamente significativas em sua incorporação isotópica (**Figura 63** e **Figura 66**). No primeiro domínio cistatina, houve uma proteção significativa na porção N-terminal (resíduos 1-42), onde esta região incorporou 50% menos deutério do que em seu estado livre (**Figura 64**). Em 2007, Aoki e colaboradores desenvolveram uma

abordagem elegante de química de proteínas para a determinação de fragmentos ativos de HSF, envolvendo modificações químicas seletivas em alguns resíduos e hidrólises controladas. Os resultados demonstraram que a primeira metade do primeiro domínio cistatina de HSF é fundamental para a inibição da atividade catalítica de metaloendopeptidases do veneno de *P. flavoviridis*; além disso, um fragmento ativo, com atividade anti-hemorrágica significativa, foi obtido a partir da hidrólise do inibidor com papaína e brometo de cianogênio. Este fragmento corresponde a duas cadeias polipeptídicas – resíduos 5 – 89 e 312 – 317 – unidos por uma ponte dissulfeto (Aoki, Deshimaru, *et al.*, 2007). Portanto, os dados de HDX-MS de BJ46a corroboram a hipótese de que a primeira metade do domínio cistatina é importante para a inibição de SVMP por estes inibidores. Os dados de HDX-MS parecem restringir a região envolvida, o que facilita uma eventual síntese química e/ou expressão heteróloga deste fragmento para futuros testes de inibição da atividade catalítica de SVMP *in vitro*.

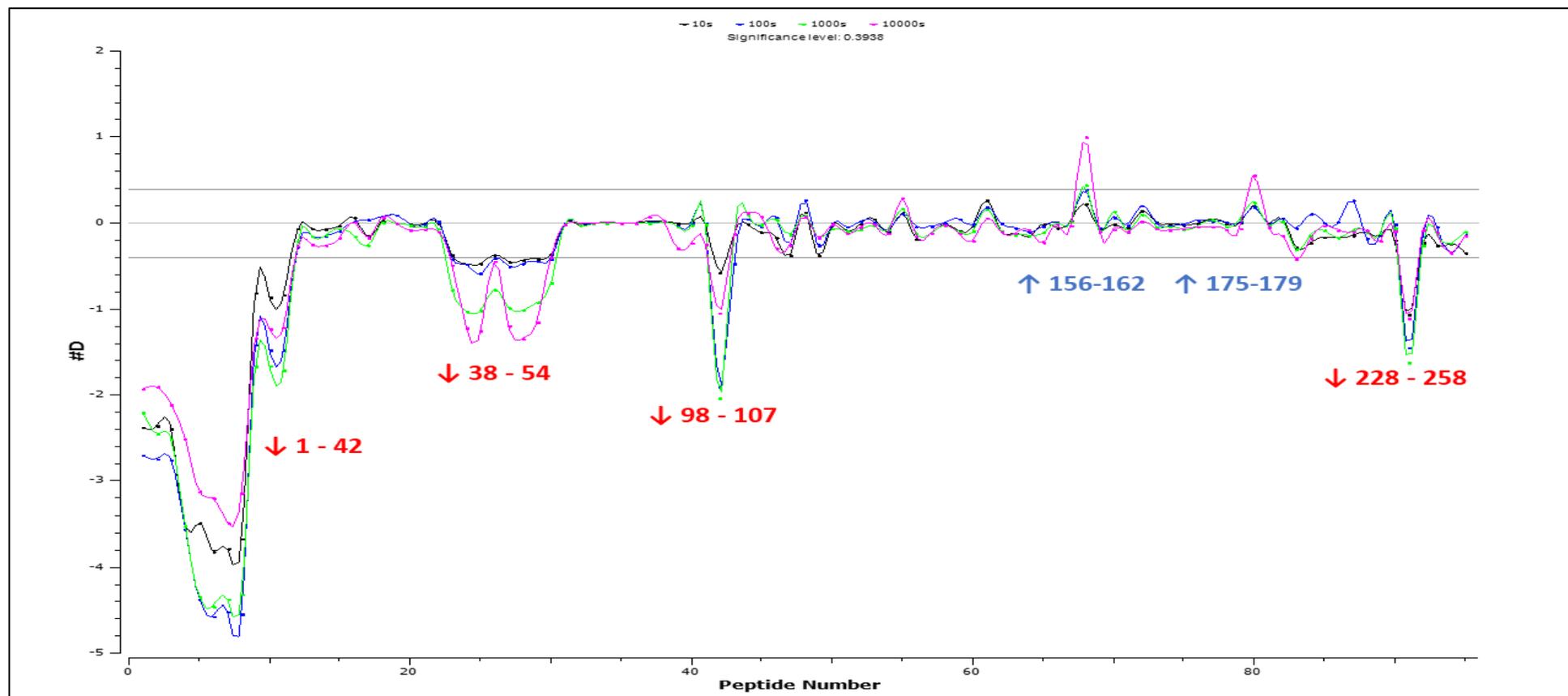
Apesar de BJ46a não ser uma cistatina canônica, os dados de HDX-MS de BJ46a são também consistentes com o descrito para uma cistatina purificada de ovo de galinha, na qual foi observada uma proteção significativa para a porção N-terminal quando em complexo com papaína; os autores atribuíram esta proteção significativa à formação de ligações de hidrogênio intermoleculares (Akashi e Takio, 2000). O mesmo fenômeno foi verificado para cistatina C humana, onde a remoção dos 10 primeiros resíduos N-terminais da proteína reduziu a afinidade de ligação da cistatina para catepsinas B e L em três ordens de magnitude e em 240 vezes para papaína (Abrahamson *et al.*, 1991). Assim, apesar de BJ46a ter como alvo específico as metaloendopeptidases de venenos de serpentes, o inibidor, assim como fetuínas e outras moléculas, evoluiu de um inibidor ancestral contendo domínios cistatina ativos (Brown e Dziegielewska, 1997). Esta relação evolutiva talvez seja a razão da importância da região N-terminal para a inibição de metaloendopeptidases (HSF / BJ46a) e para a inibição de cisteinoendopeptidases, no caso das cistatinas verdadeiras.

Além do N-terminal de BJ46a, duas outras regiões de proteção foram mapeadas no primeiro domínio cistatina de BJ46a: uma região correspondente aos resíduos 38 – 54 e outra, aos resíduos 98 – 107. No trabalho de Aoki e colaboradores, um fragmento maior, correspondente aos resíduos 5 – 147, também foi descrito como

ativo; este pôde ser truncado para o fragmento 5 – 89, sem perda de atividade hemorrágica (Aoki, Deshimaru, *et al.*, 2007).

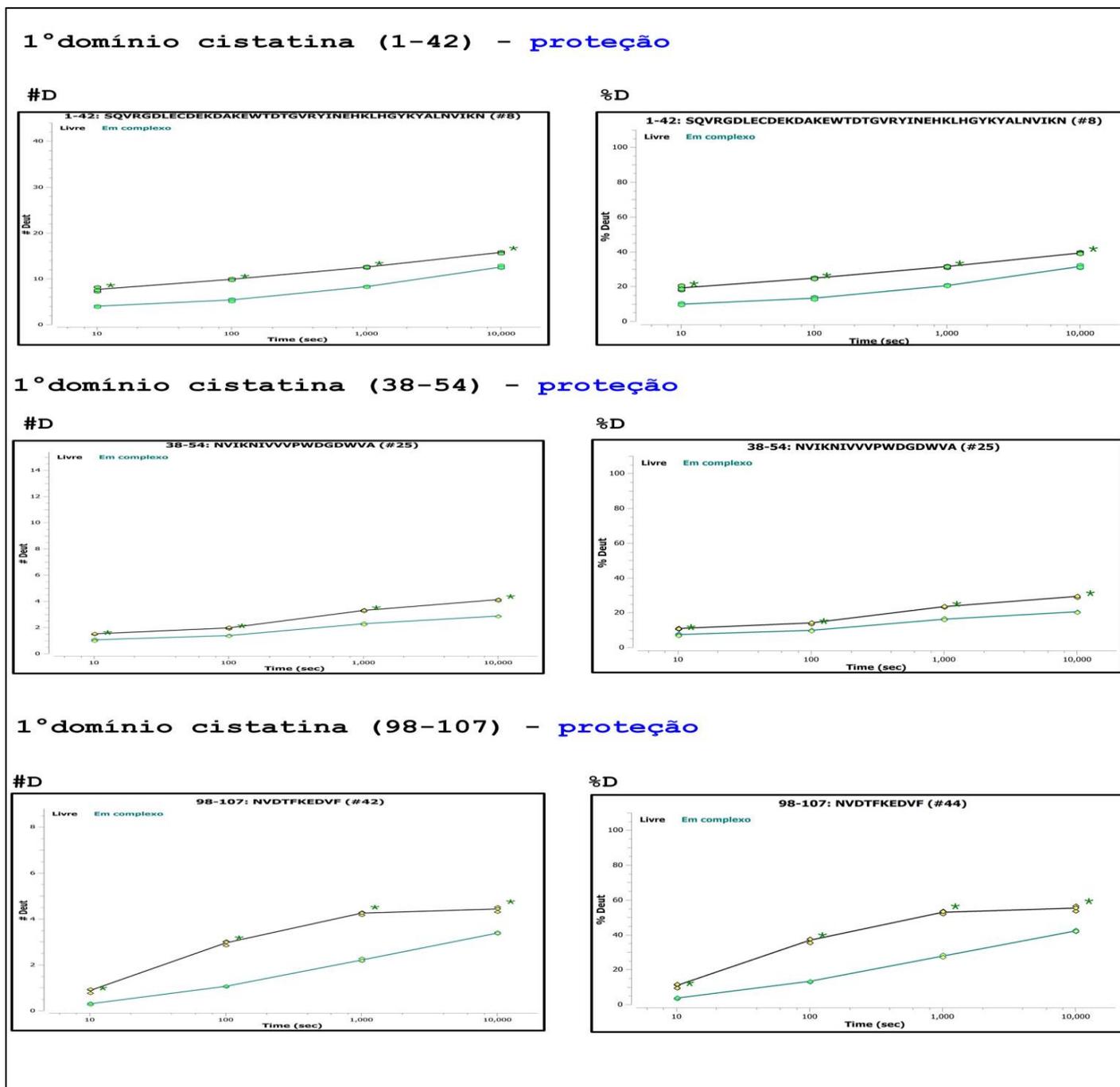
Para o segundo domínio cistatina de BJ46a, mapeamos duas regiões de exposição (156 – 162, 175 – 179) e uma região de proteção (228 - 258), como demonstra a **Figura 65**. Os dados de XL-MS (neste capítulo, subseção 4.3.2) indicam a proximidade espacial deste domínio com o domínio metaloendopeptidase de jararagina, e a detecção destas mudanças no perfil de incorporação isotópica podem indicar o envolvimento efetivo do segundo domínio cistatina na interação BJ46a-jararagina. Para HSF, não há nenhuma evidência na literatura sobre o papel do segundo domínio cistatina na inibição de metaloendopeptidases de *P. flavoviridis*. É importante ressaltar que as diferenças na incorporação isotópica para o 2º domínio cistatina de BJ46a são menos proeminentes do que as observadas para o 1º domínio cistatina (**Figura 63** e **Figura 66**), indicando que, se há efetivamente o envolvimento dos dois domínios cistatina na interação com a SVMP jararagina, o mesmo não parece equitativo. Uma outra hipótese que precisa ser considerada em estudos futuros é a ocorrência de efeitos alostéricos, onde o evento de interação proteína-proteína pode influenciar a dinâmica – e, conseqüentemente, a incorporação isotópica - de regiões distantes da interface de interação (Konermann *et al.*, 2014).

Em BJ46a, mapeamos também alguns peptídeos onde ocorreu diferença significativa de incorporação de deutério apenas em um ponto de exposição a D<sub>2</sub>O (**Figura 67**). Estas mudanças podem ser o resultado de pequenas flutuações conformacionais que devem ocorrer em BJ46a tanto em sua forma livre quanto em complexo. Como a análise de HDX-MS tem um componente temporal, as curvas de incorporação de deutério podem revelar estas flutuações estruturais locais nas diferentes condições experimentais.



**Figura 63: Gráfico residual da incorporação isotópica de BJ46a em complexo.**

Neste gráfico, cada curva representa um tempo de exposição a deutério e demonstra o perfil de incorporação isotópica ao longo de toda a sequência do inibidor. As linhas horizontais em preto refletem o nível de significância. No eixo y, encontra-se a variação de incorporação de átomos de deutério ao longo de toda a sequência (eixo x, representada por números atribuídos a cada peptídeo da sequência). Em BJ46a, mapeamos três regiões de proteção no primeiro domínio cistatina (1-42, 38-54, 98-107) e duas regiões de exposição (156 – 162, 175 – 179) e uma de proteção (228-258) no segundo domínio cistatina.

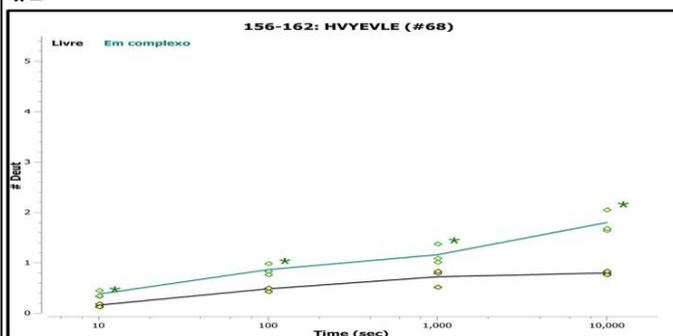


**Figura 64: Curvas de incorporação representativas para regiões do primeiro domínio cistatina de BJ46a.**

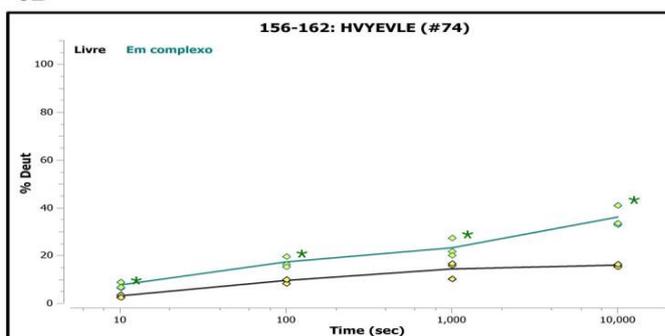
Derivamos curvas de incorporação associando o número absoluto de átomos de deutério incorporados (gráficos à esquerda) e para a incorporação relativa (gráficos à direita) ao longo dos tempos de exposição a  $D_2O$ . Nestes gráficos, a linha em preto indica o perfil de incorporação isotópica para BJ46a em sua forma livre, enquanto que a linha em azul reflete a incorporação isotópica quando em complexo. Os asteriscos em cada ponto experimental indicam significância estatística pelo teste T de Student ( $p < 0,05$ ).

## 2º domínio cistatina (156-162) - **exposição**

#D

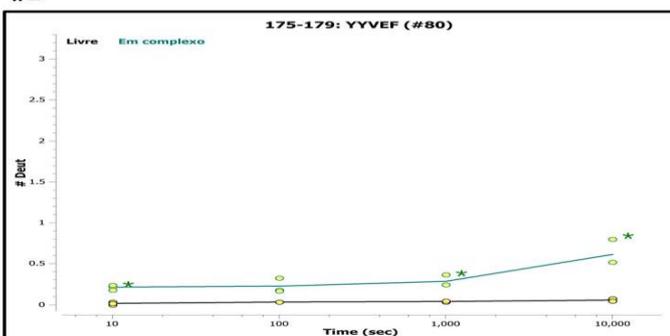


%D

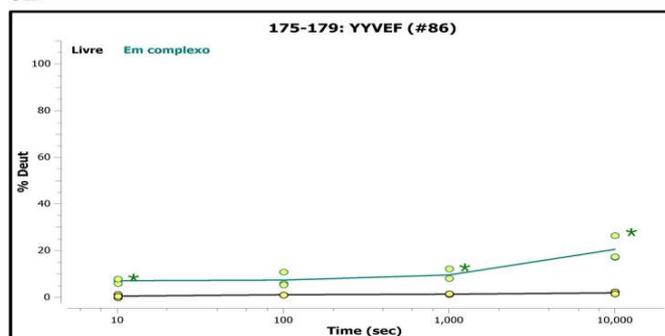


## 2º domínio cistatina (175-179) - **exposição**

#D

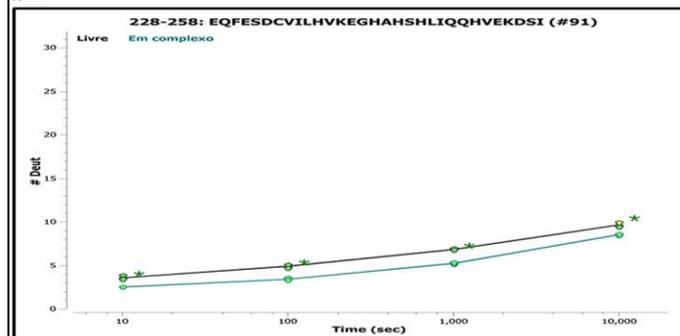


%D

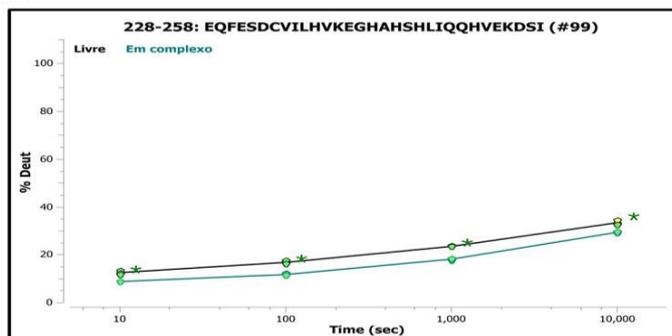


## 2º domínio cistatina (228-258) - **proteção**

#D

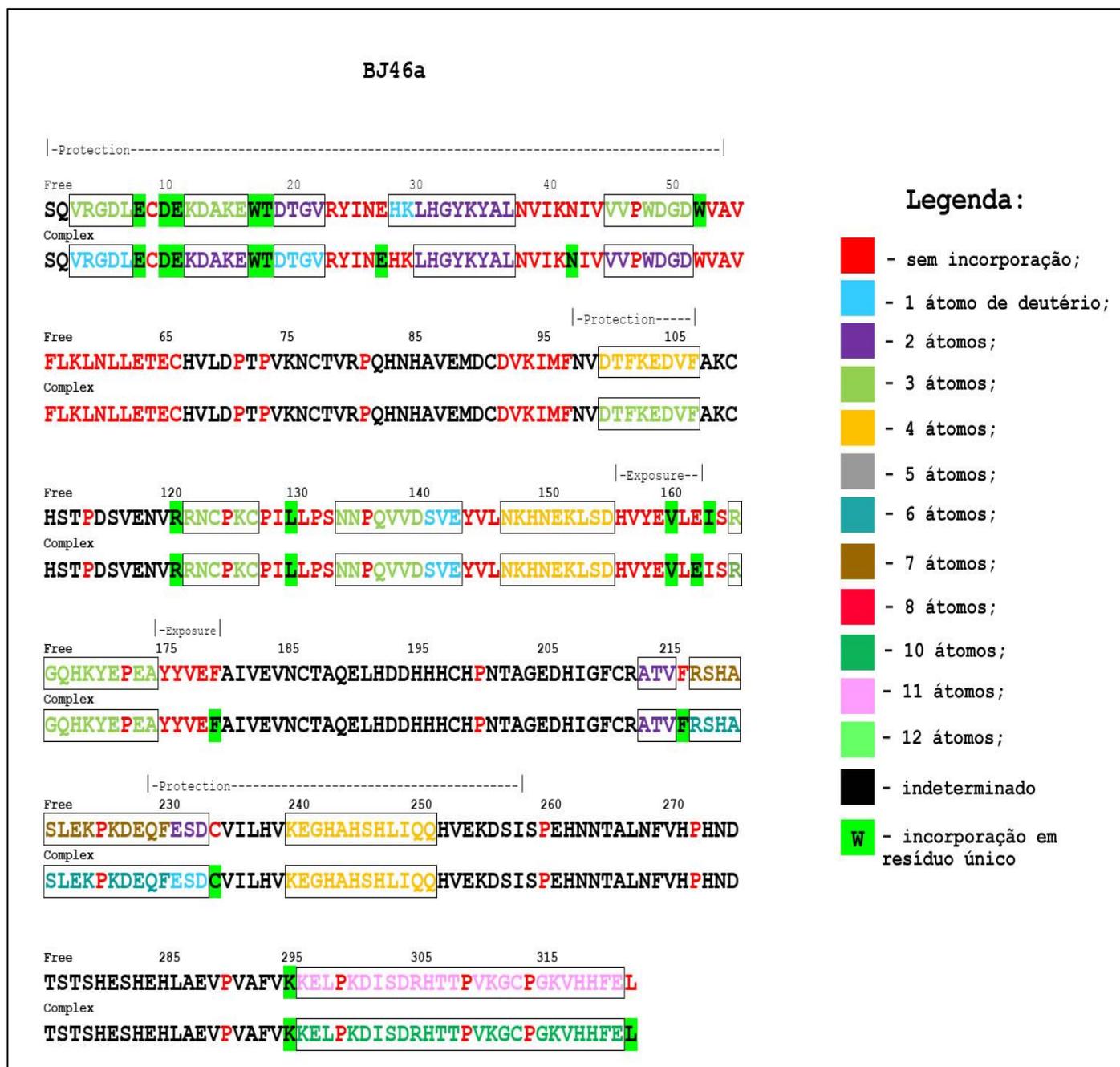


%D



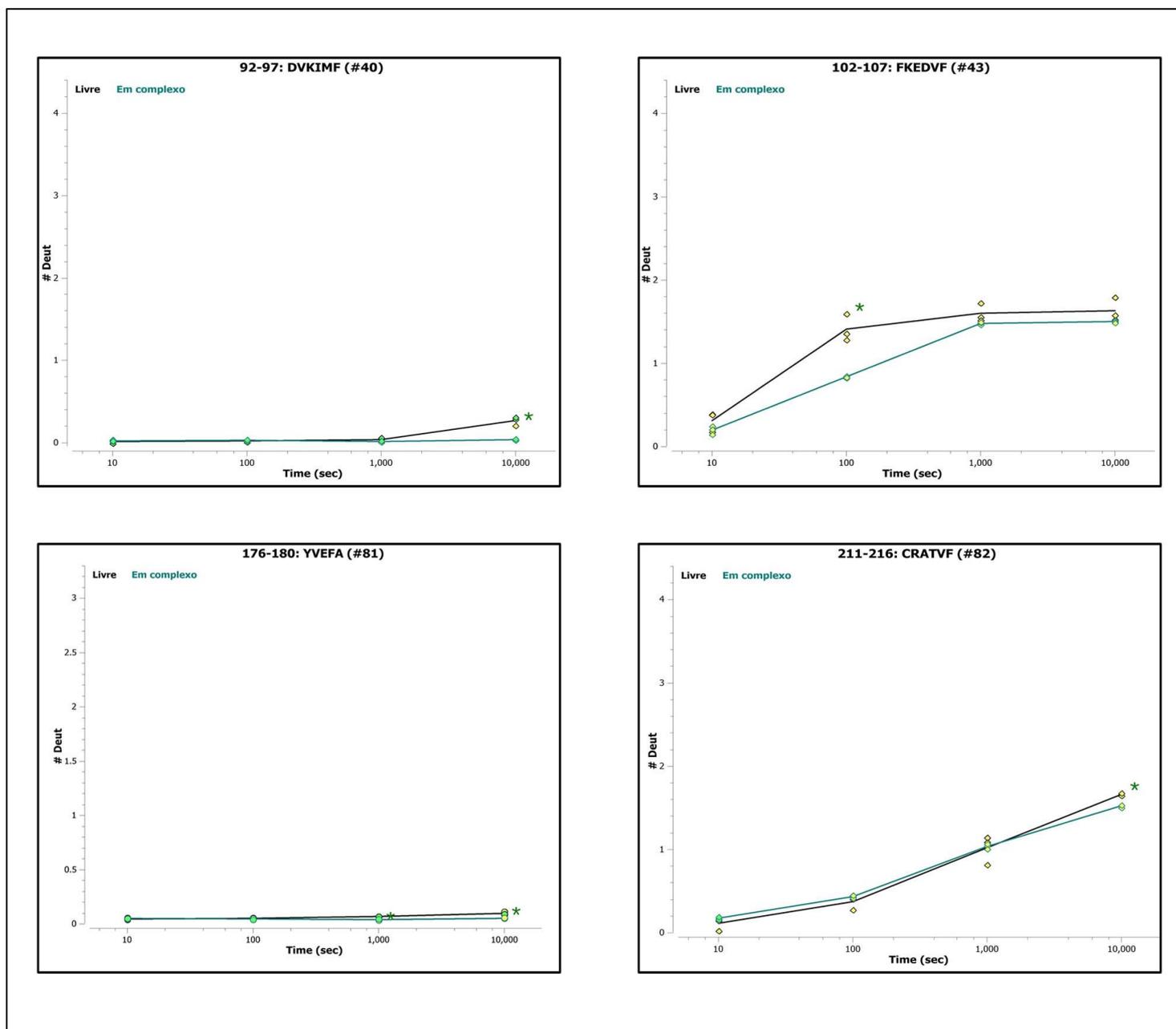
**Figura 65: Curvas de incorporação representativas para regiões do segundo domínio cistatina de BJ46a.**

Derivamos curvas de incorporação associando o número absoluto de átomos de deutério incorporados (gráficos à esquerda) e para a incorporação relativa (gráficos à direita) ao longo dos tempos de exposição a D<sub>2</sub>O. Nestes gráficos, a linha em preto indica o perfil de incorporação isotópica para BJ46a em sua forma livre, enquanto que a linha em azul reflete a incorporação isotópica quando em complexo. Os asteriscos em cada ponto experimental indicam significância estatística pelo teste T de Student ( $p < 0,05$ ).



**Figura 66: Análise detalhada da incorporação isotópica do inibidor BJ46a em sua forma livre e em complexo com jaragina, após 10.000 s de exposição a D<sub>2</sub>O.**

Nesta análise, todas as regiões onde obtivemos informação de incorporação isotópica estão realçadas em retângulos pretos, e a quantidade de átomos de deutério incorporado por região estão coloridas de acordo com o esquema. Os resíduos de prolina estão marcados em vermelho pois não há incorporação isotópica nestes resíduos.



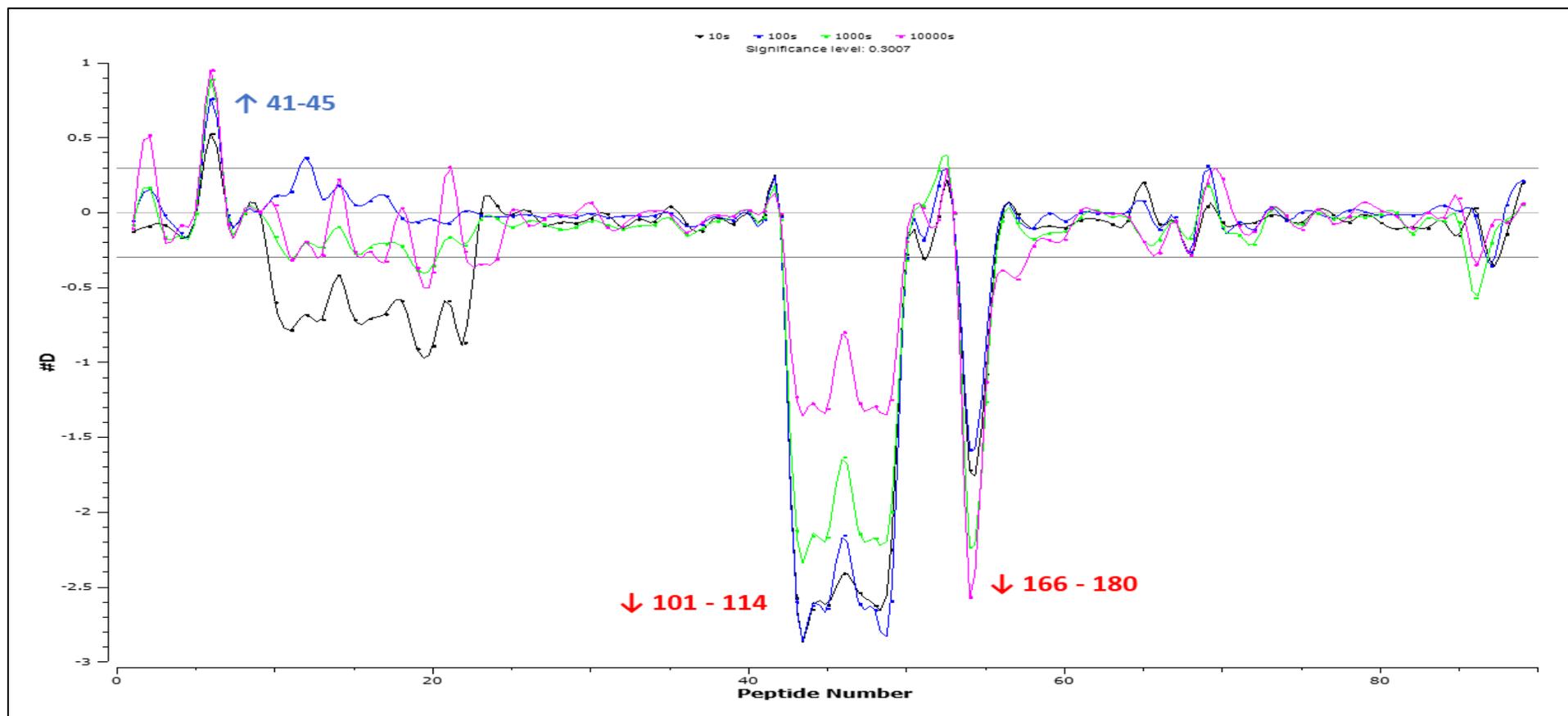
**Figura 67: Regiões pontuais de diferença em incorporação isotópica para BJ46a.**

Nestes peptídeos, ocorreram mudanças significativas na exposição de deutério entre BJ46a em sua forma livre e em complexo apenas em um único ponto de exposição a  $D_2O$ . Nestes gráficos, a linha em preto indica o perfil de incorporação isotópica para BJ46a em sua forma livre, enquanto que a linha em azul reflete a incorporação isotópica quando em complexo. Os asteriscos em cada ponto experimental indicam significância estatística pelo teste T de Student ( $p < 0,05$ ).

Para jararagina, as análises demonstraram regiões diferenciais de proteção (resíduos 101 – 114, 166 – 180) e de exposição (41-45) estatisticamente validadas no domínio metaloendopeptidase (**Figura 68** e **Figura 70**). Na estrutura cristalográfica de botropasina, a região 41-45 corresponde à parte central da hélice H2 do domínio metaloendopeptidase (**Figura 36**). No conjunto de dados para jararagina livre, esta região encontra-se protegida da incorporação isotópica em todos os tempos de exposição a D<sub>2</sub>O e, no momento que o complexo é formado, esta região incorpora um átomo de deutério (**Figura 70**). É possível que a interação com BJ46a tenha induzido uma ruptura de ligações de hidrogênio, ou uma mudança conformacional possa ter ocorrido de forma a deixar esta região mais exposta a D<sub>2</sub>O, e esta mudança é verificada logo no início da curva de incorporação (**Figura 69**).

As mudanças mais substanciais no perfil de incorporação isotópica foram registradas para as regiões 101 – 114 e 166 – 180 que tiveram suas incorporações reduzidas em aproximadamente 10% e 20%, respectivamente (**Figura 69**).

Para o domínio desintegrina e rico em cisteína de jararagina, verificamos alterações pontuais em alguns peptídeos pertencentes a estas regiões quando o complexo BJ46a – jararagina é formado, como demonstra a **Figura 71**. Como discutido anteriormente, tais alterações pontuais na curva de incorporação de deutério podem ser resultado de mudanças na dinâmica molecular, que inclusive ocorrem em regiões distantes da interface de interação proteína-proteína (Konermann *et al.*, 2014).

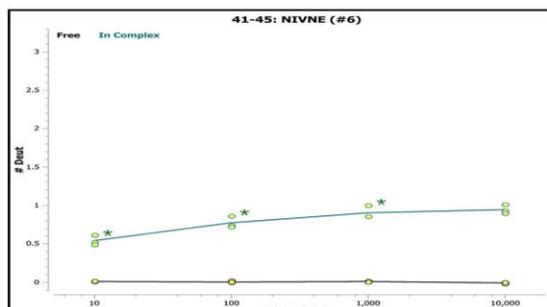


**Figura 68: Gráfico residual da incorporação isotópica de jararagina em complexo.**

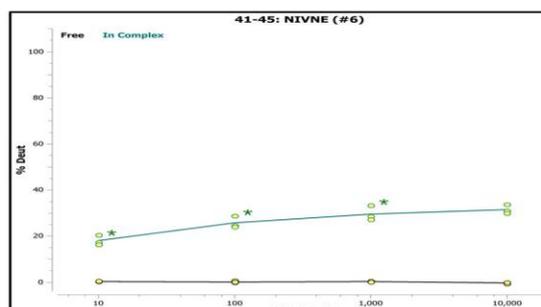
Neste gráfico, cada curva representa um tempo de exposição a deutério e demonstra o perfil de incorporação isotópica ao longo de toda a sequência da toxina. As linhas horizontais em preto refletem o nível de significância. No eixo y, encontra-se a variação de incorporação de átomos de deutério ao longo de toda a sequência (eixo x, representada por números atribuídos a cada peptídeo da sequência). Em jararagina, mapeamos duas regiões de proteção no domínio metaloendopeptidase (101 – 114 e 166-180) e uma região de exposição (41-45) no mesmo domínio.

Domínio metaloendopeptidase (41-45) - **exposição**

#D

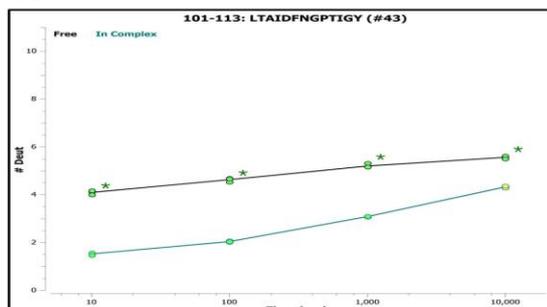


%D

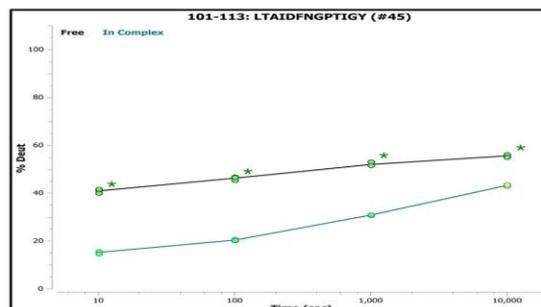


Domínio metaloendopeptidase (101-114) - **proteção**

#D

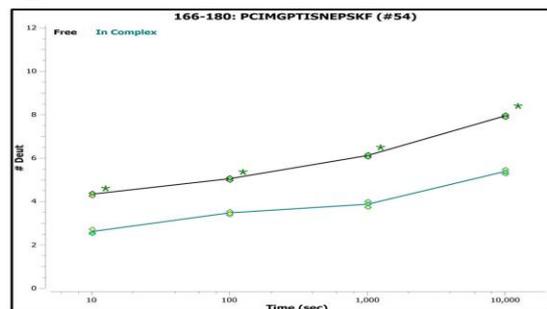


%D



Domínio metaloendopeptidase (166-180) - **proteção**

#D



%D

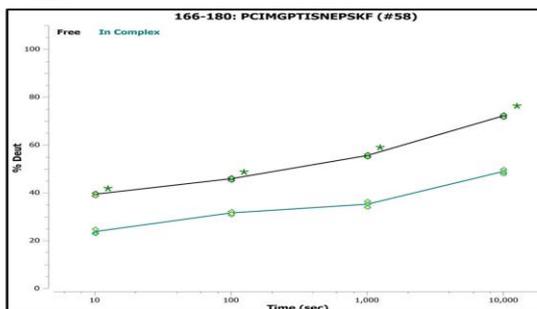


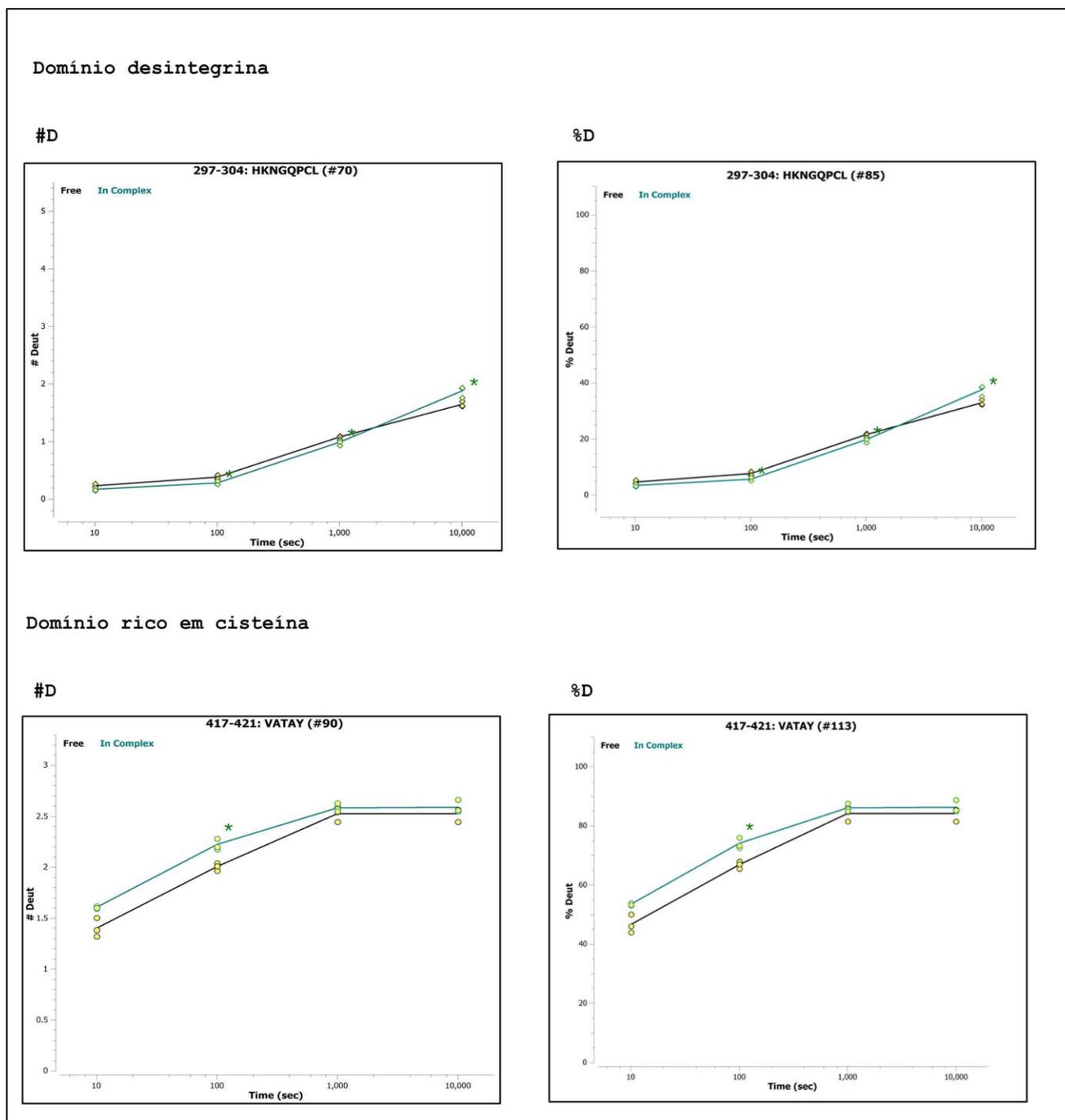
Figura 69: Curvas de incorporação representativas para as regiões diferenciais de incorporação em jararagina.

Derivamos curvas de incorporação associando o número absoluto de átomos de deutério incorporados (gráficos à esquerda) e para a incorporação relativa (gráficos à direita) ao longo dos tempos de exposição a  $D_2O$ . Nestes gráficos, a linha em preto indica o perfil de incorporação isotópica para jararagina em sua forma livre, enquanto que a linha em azul reflete a incorporação isotópica quando em complexo. Os asteriscos em cada ponto experimental indicam significância estatística pelo teste T de Student ( $p < 0,05$ ).



**Figura 70: Análise detalhada da incorporação isotópica de jararagina em sua forma livre e em complexo com BJ46a, após 10.000 s de exposição a D<sub>2</sub>O.**

Nesta análise, todas as regiões onde obtivemos informação de incorporação isotópica estão realçadas em retângulos pretos, e a quantidade de átomos de deutério incorporado por região estão coloridas de acordo com o esquema. Os resíduos de prolina estão marcados em vermelho pois não há incorporação isotópica nestes resíduos.



**Figura 71: Curvas de incorporação representativas de mudanças pontuais de incorporação isotópica nos domínios desintegrina e rico em cisteína de jararagina.**

Nestes peptídeos, ocorreram mudanças significativas na exposição de deutério entre jararagina em sua forma livre e em complexo apenas em um único ponto de exposição a  $D_2O$ . Nestes gráficos, a linha em preto indica o perfil de incorporação isotópica para jararagina em sua forma livre, enquanto que a linha em azul reflete a incorporação isotópica quando em complexo. Os asteriscos em cada ponto experimental indicam significância estatística pelo teste T de Student ( $p < 0,05$ ).

## 4.5 - INTEGRAÇÃO DOS DADOS DE XL-MS E HDX-MS NA MODELAGEM MOLECULAR E DOCKING DO COMPLEXO BJ46A-JARARAGINA

### 4.5.1 – Modelagem molecular da metaloendopeptidase jararagina

Após a obtenção da sequência madura da metaloendopeptidase jararagina do banco de dados UNIPROT (conforme descrito na subseção 3.2.21.1, capítulo 3), iniciamos a modelagem molecular da proteína utilizando o servidor I-TASSER (disponível em <http://zhanglab.ccmb.med.umich.edu/I-TASSER>), um dos melhores meta-servidores públicos de predição avaliados no CASP (*Critical Assessment of Protein Structure Prediction*, [http://predictioncenter.org/index.cgi?page=public\\_serv](http://predictioncenter.org/index.cgi?page=public_serv)).

A partir da sequência de aminoácidos da proteína-alvo (*query*), o servidor inicia a busca em bancos de dados não-redundantes por sequências similares à sequência-alvo, mesmo que distantes evolutivamente, utilizando o algoritmo PSI-BLAST (Altschul *et al.* 1997). Na segunda etapa, o PSIPRED (Jones, 1999) realiza a predição de elementos de estrutura secundária. Na terceira etapa, o meta-servidor LOMETS (Wu e Zhang, 2007) executa 11 algoritmos para efetuar o *threading* da sequência-alvo com uma biblioteca não-redundante de estruturas proteicas (Yang *et al.* 2015). Tendo como base estes alinhamentos iniciais, a sequência é dividida artificialmente entre regiões alinhadas com sucesso e regiões não-alinhadas. As primeiras são excisadas das estruturas-molde identificadas pelo LOMETS para a construção do modelo, enquanto que as últimas são modeladas *ab initio*. A construção do modelo final reúne todos os fragmentos identificados e os trechos modelados *ab initio*, por meio de uma série de simulações de Monte Carlo e guiado por um campo de força otimizado (*knowledge-based*), utilizando potenciais estatísticos, redes de ligações de hidrogênio e restrições oriundas dos fragmentos excisados das estruturas-molde (Roy *et al.* 2010; Yang *et al.* 2015).

O algoritmo LOMETS identificou a estrutura cristalográfica de botropasina (PDB ID: 3DSL, resolução de 2,7 Å) como o melhor molde para a modelagem molecular de jararagina. O alinhamento da estrutura primária de jararagina (Jarar) e botropasina (Bothr) mostra uma diferença pontual de apenas 18 resíduos de aminoácidos (**Figura 72**). Dez modificações estão contidas na porção N-terminal do domínio metaloendopeptidase e as oito restantes, em regiões próximas ao sítio ativo.

|       |     |   |     |
|-------|-----|---|-----|
| Jarar | 3   | QRYDPYKYIEFFVVDQGTVTKNNGDLDKIKARMYELANIVNEIFRYLYMHVALVGLEIW   | 62  |
|       |     | Q+Y+P++Y+E F+VVDQG VTKNNGDLDKIKARMYELANIVNEI RYLYMH ALVGLEIW  |     |
| Bothr | 1   | QKYNPFRYVELFIVVDQGMVTKNNGDLDKIKARMYELANIVNEILRYLYMHAALVGLEIW  | 60  |
| Jarar | 63  | SNGDKITVKPDVDYTLNSFAEWRKTDLLTRKKHDNAQLLTAIDFNGPTIGYAYIGSMCHP  | 122 |
|       |     | SNGDKITVKPDVDYTLNSFAEWRKTDLLTRKKHDNAQLLTAIDFNGPTIGYAYIGSMCHP  |     |
| Bothr | 61  | SNGDKITVKPDVDYTLNSFAEWRKTDLLTRKKHDNAQLLTAIDFNGPTIGYAYIGSMCHP  | 120 |
| Jarar | 123 | KRSVGVIVQDYSPINLVVAVIMAH <b>HEM</b> GHNLGIHHD <b>TG</b> SCSGDYPC <b>IM</b> GPTISNEPSKFFS                        | 182 |
|       |     | KRSV IV+DYSPINLVVAVIMAH <b>HEM</b> GHNLGIHHD T CSGDYPCIMGPTISNEPSKFFS   |     |
| Bothr | 121 | KRSVAIV <b>ED</b> YSPINLVVAVIMAH <b>HEM</b> GHNLGIHHD <b>DF</b> CSGDYPCIMGPTISNEPSKFFS                          | 180 |
| Jarar | 183 | <b>NCS</b> YIQCWDFIM <b>NHN</b> PECI <b>I</b> NEPLGTDIISPP <b>CG</b> NE <b>LE</b> VGEE <b>CD</b> CGTPENCQNECCDA | 242 |
|       |     | NCSYIQCWDFIM NP+CI+NEPLGTDI+SPPVCGNELLEVGE <b>EC</b> DCGTPENCQNECCDA  |     |
| Bothr | 181 | <b>NCS</b> YIQCWDFIM <b>KEN</b> P <b>QC</b> ILNEPLGTDIVSPPVCGNELLEVGE <b>EC</b> DCGTPENCQNECCDA                 | 240 |
| Jarar | 243 | ATCKLKSGSQCGHGDCCEQCKFSKSGTECRASMS <b>ECD</b> <b>PA</b> EHCTGQSSECPAD <b>V</b> FHKNGQP                          | 302 |
|       |     | ATCKLKSGSQCGHGDCCEQCKFSKSGTECRASM <b>SEC</b> DP <b>AE</b> HCTGQSSECPAD <b>V</b> FHKNGQP                         |     |
| Bothr | 241 | ATCKLKSGSQCGHGDCCEQCKFSKSGTECRASMS <b>ECD</b> <b>PA</b> EHCTGQSSECPAD <b>V</b> FHKNGQP                          | 300 |
| Jarar | 303 | CLDNYGYCYNGNCPIMYHQCYALFGADVVEAEDSCFKDNQKGNYYGYCRKENGKKIPCAP  | 362 |
|       |     | CLDNYGYCYNGNCPIMYHQCYALFGADVVEAEDSCFKDNQKGNYYGYCRKENGKKIPCAP  |     |
| Bothr | 301 | CLDNYGYCYNGNCPIMYHQCYALFGADVVEAEDSCFKDNQKGNYYGYCRKENGKKIPCAP  | 360 |
| Jarar | 363 | EDVKCGRLYCKDNSPGQNNPCKMFYSNDDEHKGMVLPGTKCADGKVCSSNGHCVDVATAY  | 421 |
|       |     | EDVKCGRLYCKDNSPGQNNPCKMFYSNDDEHKGMVLPGTKCADGKVCSSNGHCVDVATAY  |     |
| Bothr | 361 | EDVKCGRLYCKDNSPGQNNPCKMFYSNDDEHKGMVLPGTKCADGKVCSSNGHCVDVATAY  | 419 |

**Figura 72: Alinhamento de seqüências entre jararagina (Jarar) e Botropasina (Bothr).**

O alinhamento foi realizado utilizando-se o algoritmo PSI-BLAST (configuração padrão). Observa-se identidade sequencial de 95% (400 em 419 resíduos), sem *gaps* na seqüência.

Marcamos em verde resíduos modificados com mesma função química, em vermelho marcamos as substituições por resíduos de diferentes propriedades químicas (*missense*). A região conservada referente ao sítio ativo foi ressaltada em rosa e a seqüência-consenso referente ao sítio de N-glicosilação em amarelo. Os dois motivos ECD do domínio desintegrina foram marcados em ciano com os sítios de ligação à cálcio ressaltados em marrom.

A estrutura cristalográfica de botropasina é a única estrutura de metaloendopeptidase de classe PIII de *B. jararaca* depositada no PDB. Cristais de botropasina foram obtidos a partir da incubação da toxina com o inibidor de metaloendopeptidases POL647 e a unidade assimétrica do cristal apresenta duas moléculas de toxina; porém, botropasina é monomérica em solução (Muniz *et al.* 2008).

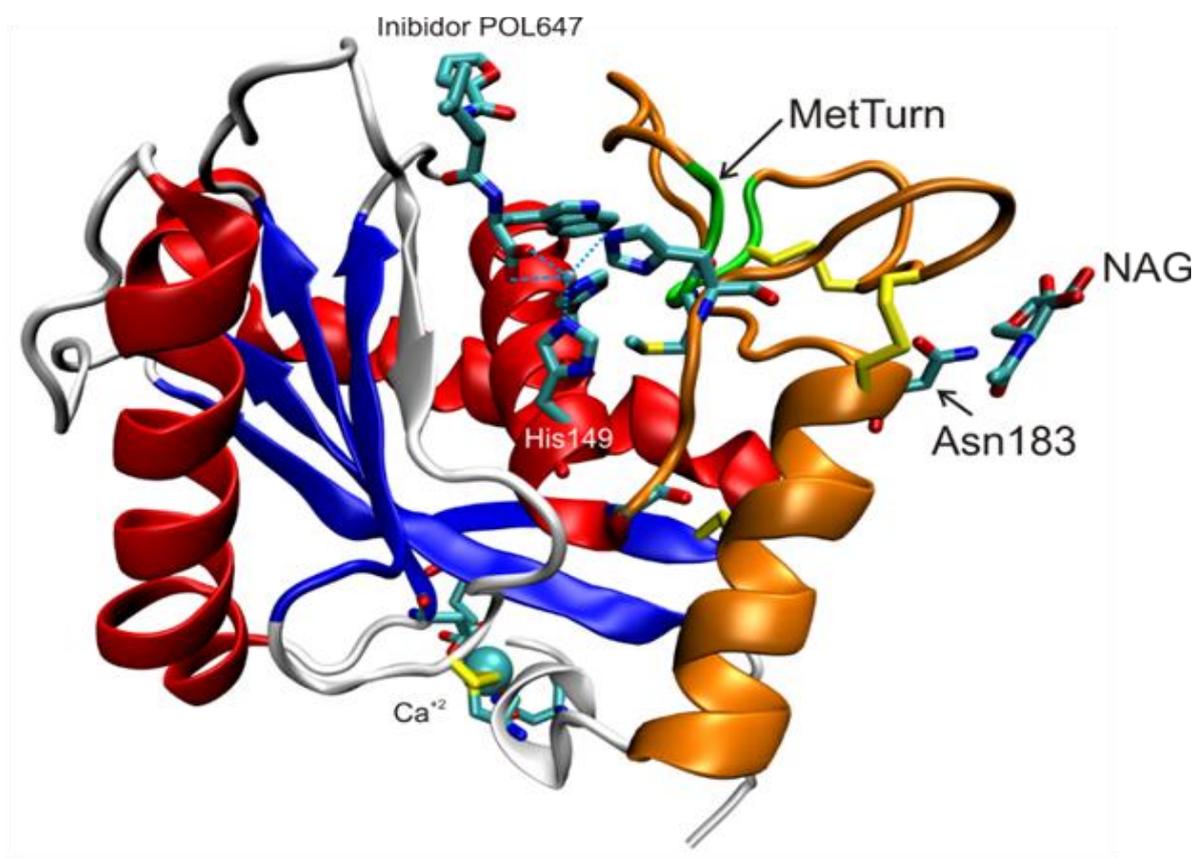
A **Figura 74** demonstra a estrutura completa da SVMP botropasina. A toxina apresenta uma estrutura tridimensional em forma de “C” e três domínios: o domínio metaloendopeptidase (M), um domínio desintegrina-símile (D) e um domínio rico em cisteína (C), na sua porção C-terminal.

O domínio metaloendopeptidase, ilustrado na **Figura 73**, é o domínio catalítico da toxina e apresenta um enovelamento do tipo  $\alpha/\beta$ , composto de uma folha  $\beta$  paralela composta por 4 fitas (S1, S2, S3, e S4), uma fita antiparalela (S5) e 4  $\alpha$ -hélices. Há um subdomínio, composto de uma  $\alpha$ -hélice e diversas voltas. O sítio ativo da enzima apresenta um átomo de zinco catalítico, coordenado pelos resíduos H145, H149 e H155 e por uma molécula de água (substituída na estrutura cristalográfica pelo inibidor POL647). A alça *Met-turn*, conservada nas metzincinas, auxilia na organização espacial dos resíduos de histidina do sítio ativo e é composta pelos resíduos C<sub>167</sub>IMXP<sub>171</sub>. A estabilidade estrutural do domínio é conferida por um domínio de ligação a cálcio (E12, D96, C200 e N203), situado do lado oposto ao sítio ativo e estabilizando o N-terminal da folha  $\beta$  e por 4 pontes dissulfeto (C120-C200, C160-C184 e C162-C167) (Muniz *et al.*, 2008). O domínio metaloendopeptidase e o domínio subsequente (desintegrina-símile) são conectados por uma região de transição, composta pelos resíduos E<sub>204</sub>PLGTDIISP<sub>213</sub>, estruturado na forma de um *loop* que promove flexibilidade na orientação dos domínios metaloendopeptidase e desintegrina-símile (Takeda, 2014).

O domínio desintegrina-símile não apresenta nenhum elemento de estrutura secundária definido, sendo composto primariamente por *loops* e voltas estabilizadas por 7 pontes dissulfeto intrínsecas ao domínio (C216-C245, C227-C240, C229-C235, C239-262, C253-259, C258-284 e C271-C291) e uma ponte adicional com o domínio rico em cisteína (C278-C310). Assim como no domínio metaloendopeptidase, há sítios de ligação ao cálcio que participam da estabilização estrutural do domínio (V215, L220, E222 e D228 e D294, V295, D279, E282, P280). Uma característica

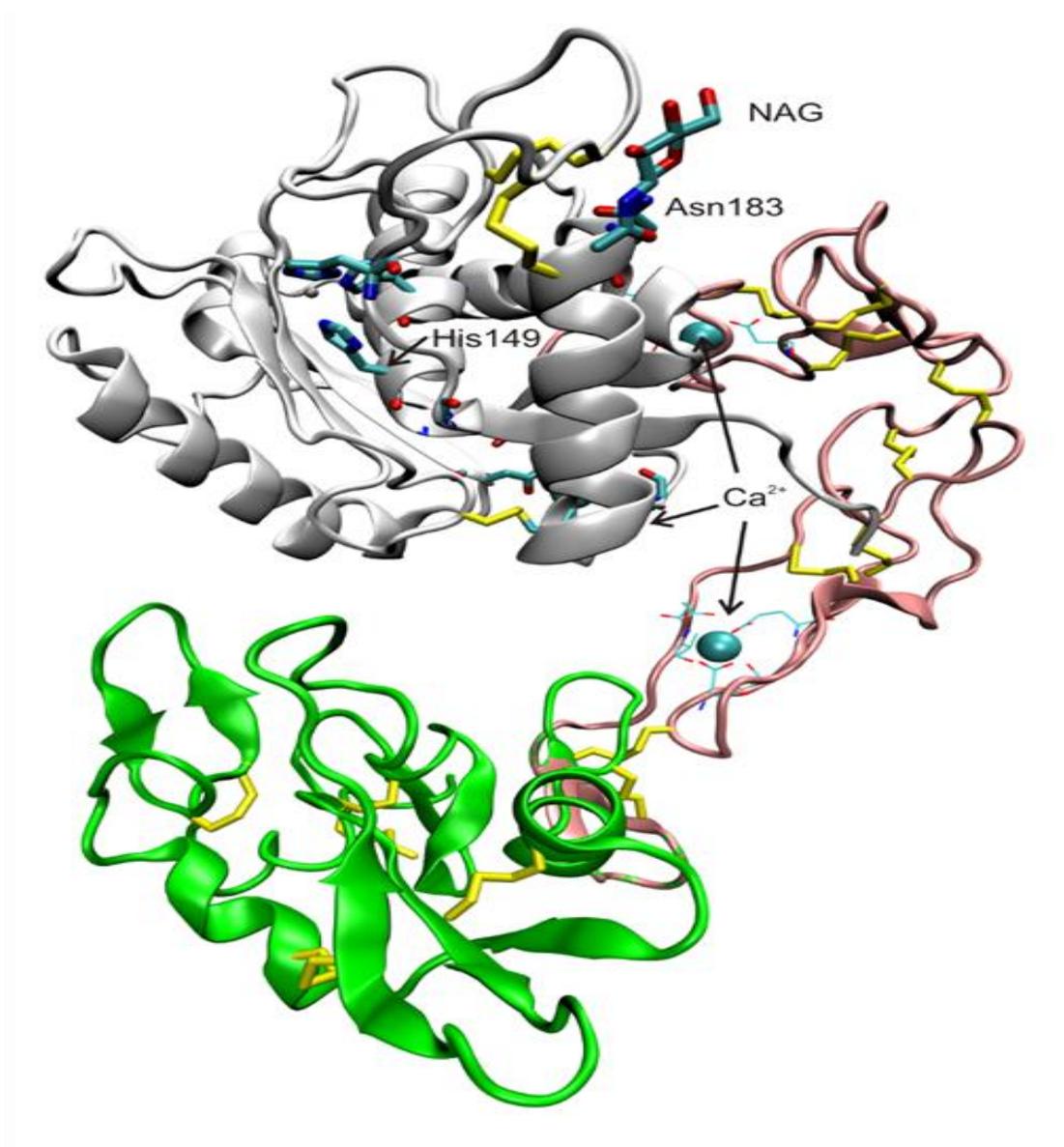
importante do domínio desintegrina-símile é a presença de dois motivos ECD, relacionados à interação das metaloendopeptidases com integrinas de superfície presentes em plaquetas e células endoteliais (Jia *et al.* 1997)

O domínio rico em cisteína (C), presente na extremidade C-terminal da toxina, apresenta também um baixo conteúdo de elementos de estrutura secundária. Há duas pequenas  $\alpha$ -hélices e quatro fitas  $\beta$ . O domínio é estruturalmente estabilizado por 6 pontes dissulfeto (C303 – C315, C322-C372, C337-C383, C350-C360, C367-C409, C403-C414) e uma rede de ligações de hidrogênio.



**Figura 73: O domínio metaloendopeptidase (M) de Botropasina (PDB ID: 3DSL).**

O domínio M foi representado na forma *cartoon*. O subdomínio 1 é composto das 4 fitas  $\beta$  paralelas representadas em azul, 1 fita  $\beta$  antiparalela (em cinza), além das 4  $\alpha$ -hélices, coloridas em vermelho. As fitas  $\beta$  paralelas e a  $\beta$  antiparalela formam a folha  $\beta$  central do domínio. O sítio ativo do domínio foi representado na forma de bastões: os resíduos H145, H149 e H155 coordenam um átomo de zinco (esfera colorida em cinza-escuro). O inibidor POL647 interage com o átomo de zinco, substituindo a água catalítica em sua forma ativa. O subdomínio 2 foi colorido em laranja, abrigando o resíduo H155 e a *Met-turn* (colorida em verde). O sítio de N-glicosilação foi representado pelo resíduo N183, próximo a uma região eletrodensa da estrutura cristalográfica, possibilitando o encaixe de uma molécula de N-acetilglicosamina (NAG). Os resíduos E12, N203, D96 e C200 (representados em bastões) coordenam o átomo de  $\text{Ca}^{2+}$  (representado pela esfera de Van der Waals, VDW) no sítio de ligação a cálcio.

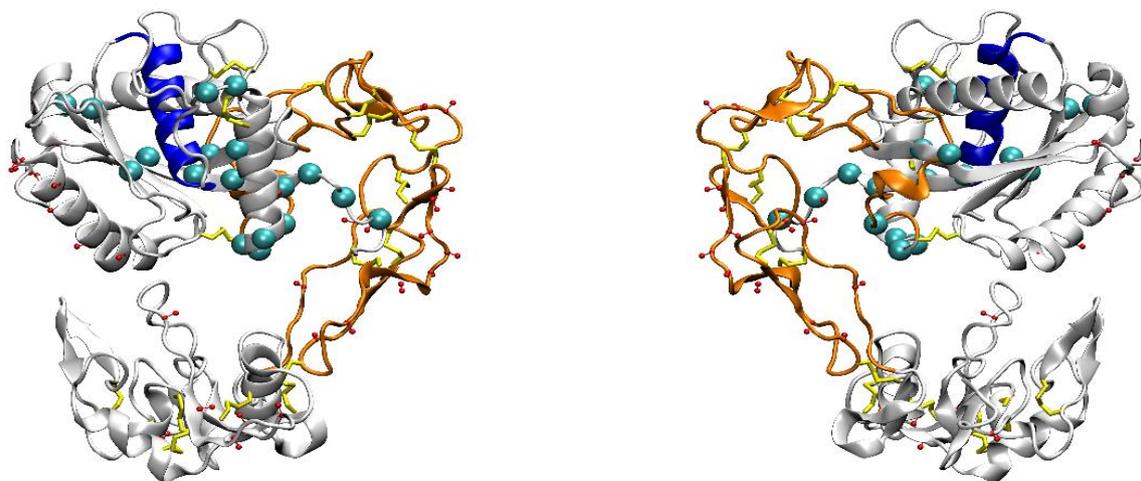


**Figura 74: A estrutura tridimensional completa da SVMP Botropasina (PDB ID: 3DSL).**

A SVMP botropasina foi representada na forma *cartoon* de acordo com o domínio: (I) Domínio metaloendopeptidase (M) colorido em cinza; (II) Domínio desintegrina-símile (D) colorido em rosa e (III) Domínio rico em cisteína (C) colorido em verde. As pontes dissulfeto foram representadas em bastões amarelos para os átomos C $\alpha$ , C $\beta$  E S $\gamma$  dos resíduos de cisteína. O sítio ativo foi representado em forma de bastões, coloridos de acordo com o tipo de átomo, pelos resíduos H145, H149 e H155 que coordenam o átomo de zinco (esfera colorida em cinza). Os sítios de ligação a cálcio foram representados na forma de linhas para as cadeias laterais dos resíduos envolvidos e na forma de esferas de VDW para o átomo de cálcio coordenado. O sítio de N-glicosilação foi representado pelo resíduo N183 e pela molécula de N-acetilglicosamina ancorada (NAG).

Após a escolha da botropasina como molde e finalização da modelagem de jararagina, os modelos de menor energia foram organizados em *clusters* e analisados pelo algoritmo SPICKER (Zhang e Skolnick, 2004). A partir desta análise, iniciou-se um novo *round* de simulações com os modelos centrais de cada *cluster*, refinando possíveis impedimentos estéricos e executando minimizações de energia (Yang *et al.* 2015). Ao final deste processo, o I-TASSER reportou até 5 grupos de modelos; a confiança de cada grupo foi refletida pelo *C-score*, uma métrica baseada na significância do alinhamento da sequência-alvo com o molde e com parâmetros de convergência na etapa de *threading* pelo LOMETS. O *C-score* varia de -5,0 a +2,0, onde quanto maior o *C-score*, maior será a confiança no modelo gerado. Modelos que apresentam um *C-score* > -1,5 frequentemente apresentam um enovelamento correto (Yang e Zhang, 2015). A identidade sequencial alta entre botropasina e jararagina fez com que o *C-score* obtido para o modelo final de jararagina fosse excelente, calculado em +1,96.

A **Figura 75** ilustra o modelo final de jararagina, com as mesmas características estruturais descritas para a estrutura cristalina de botropasina. O formato em C da enzima foi preservado, bem como a extensão dos domínios. As modificações pontuais entre a estrutura primária de botropasina e jararagina foram indicadas representando o carbono- $\alpha$  de cada resíduo modificado por esferas de Van der Waals.



**Figura 75: Modelo estrutural da SVMP jararagina, tendo a estrutura 3DSL (botropasina) como molde.**

A imagem apresenta uma visão em 180° do modelo selecionado de jararagina. Este foi representado na forma *cartoon*, colorido de acordo com os diferentes domínios: (I) domínio metaloendopeptidase em cinza, com a  $\alpha$ -hélice que abriga o sítio ativo colorida em azul; (II) domínio desintegrina-símile colorido em laranja e (III) domínio rico em cisteína colorido em cinza claro. As pontes dissulfeto foram representadas na forma de bastão para os átomos C $\alpha$ , C $\beta$  e S $\gamma$ , coloridos em amarelo. As modificações pontuais de resíduos foram representadas por esferas de Van der Waals em ciano.

#### 4.5.1.1 - Validação dos dados de XL-MS obtidos para jararagina em sua forma livre

Após a construção do modelo tridimensional de jararagina, empregamos o *software* Topolink para avaliação do modelo construído. O Topolink calcula as distâncias topológicas e euclidianas para cada restrição de distância gerada por XL-MS no modelo estrutural, permitindo uma avaliação da correspondência do modelo com os dados experimentais disponíveis (Martínez *et al.*, 2017). Para que um *cross-link* seja formado em um dado par de resíduos, as cadeias laterais precisam estar acessíveis ao solvente e, uma vez ligado à primeira cadeia lateral, o ALC precisa percorrer a superfície proteica para atingir a segunda. O cálculo da distância euclidiana entre o par de resíduos fornece um limite mínimo para o *cross-link*, mas é uma medida incompleta para a medição da distância máxima, pois não leva em consideração a geometria da superfície proteica e que partes da proteína podem obstruir a medida (Bullock *et al.*, 2018).

A **Tabela 11** demonstra os resultados da análise do Topolink para as restrições de distância geradas para jararagina livre pelo ALC BS<sup>3</sup>, enquanto que a **Tabela 12** corresponde aos resultados da análise do Topolink para os *cross-links zero-length* encontrados com a abordagem XPlex.

Para o ALC BS<sup>3</sup>, identificamos um grupo de 38 restrições espaciais, enquanto que para a abordagem XPlex, mapeamos 28 restrições. Tomando como base apenas distâncias euclidianas, validamos 63% das restrições de distância derivadas com o ALC BS<sup>3</sup> e 40% dos *cross-links zero-length* obtidos a partir da estratégia XPlex.

Ao analisarmos o conjunto de violações obtidas para a estratégia BS<sup>3</sup>, verificamos que a maior parte dos *cross-links* não-validados envolvem o domínio desintegrina-símile e rico em cisteína; uma menor parte envolve resíduos localizados em alças do domínio metaloendopeptidase com resíduos do domínio rico em cisteína (*cross-links* #3,6,7,9). Um padrão similar foi verificado para a estratégia XPlex, onde a maior parte dos *cross-links* violados ocorrem em resíduos do domínio desintegrina-símile e em parte do domínio rico em cisteína, regiões tidas como menos estruturadas do que o domínio metaloendopeptidase. Como discutido anteriormente, *cross-links zero-length* são ligações amida ou éster que resultam da ligação covalente do nitrogênio da cadeia lateral de lisina (Nζ) ou o oxigênio da cadeia lateral de resíduos de serina (Oγ) com o grupamento carbonila de um resíduo de ácido aspártico ou ácido glutâmico, seguido de eliminação de uma molécula de água (Ferrari *et al.* 2018). Dessa forma, um *cross-link zero-length* é mais restritivo do que os links observados com o ALC BS<sup>3</sup> e XPlex – diamina e são mais difíceis de serem validados em um modelo estrutural.

**Tabela 11: Análise pelo software Topolink dos cross-links encontrados para jararagina livre usando o ALC BS<sup>3</sup>.**

Legenda da tabela: Res1/Res2: resíduos envolvidos no *cross-link*; Res1/Res2 Num: numeração dos resíduos; Prot.: proteína a qual o resíduo pertence; Dist. Euclid.: distância topológica (em Å) calculada para o *cross-link*; Dist.topo.: distância topológica calculada (Å); Dist.máx.: distância máxima considerada para o *cross-link* (Å); Resultado: resultado final da análise pelo Topolink.

| CROSS-LINK # | RES1 | PROT. | RES1 NUM | RES2 | PROT. | RES2 NUM | DIST EUCLID | DIST. TOPO | DIST MAX | RESULTADO    |
|--------------|------|-------|----------|------|-------|----------|-------------|------------|----------|--------------|
| 1            | LYS  | JAR   | 31       | LYS  | JAR   | 33       | 7.182       | 13.334     | 21.800   | OK:FOUND     |
| 2            | SER  | JAR   | 63       | LYS  | JAR   | 71       | 19.932      | 26.184     | 18.000   | BAD:LONG     |
| 3            | LYS  | JAR   | 71       | LYS  | JAR   | 394      | 32.800      | >32.800    | 21.800   | BAD:EUCL     |
| 4            | LYS  | JAR   | 86       | LYS  | JAR   | 356      | 16.982      | 19.045     | 21.800   | OK:FOUND     |
| 5            | LYS  | JAR   | 86       | LYS  | JAR   | 357      | 18.817      | 19.413     | 21.800   | OK:FOUND     |
| 6            | LYS  | JAR   | 86       | LYS  | JAR   | 366      | 26.313      | >26.313    | 21.800   | BAD:EUCL     |
| 7            | LYS  | JAR   | 86       | SER  | JAR   | 388      | 22.576      | 25.858     | 18.000   | BAD:LONG     |
| 8            | LYS  | JAR   | 86       | LYS  | JAR   | 394      | 19.550      | 19.966     | 21.800   | OK:FOUND     |
| 9            | LYS  | JAR   | 93       | LYS  | JAR   | 357      | 30.428      | >30.428    | 21.800   | BAD:EUCL     |
| 10           | LYS  | JAR   | 248      | SER  | JAR   | 251      | 5.295       | 6.509      | 18.000   | OK:FOUND     |
| 11           | LYS  | JAR   | 248      | LYS  | JAR   | 266      | 21.351      | 22.685     | 21.800   | BAD:LONG     |
| 12           | SER  | JAR   | 249      | LYS  | JAR   | 266      | 21.970      | 24.575     | 18.000   | BAD:LONG     |
| 13           | SER  | JAR   | 251      | LYS  | JAR   | 266      | 16.247      | 20.469     | 18.000   | BAD:LONG     |
| 14           | LYS  | JAR   | 263      | SER  | JAR   | 265      | 7.949       | 19.771     | 18.000   | BAD:LONG     |
| 15           | LYS  | JAR   | 263      | LYS  | JAR   | 266      | 9.383       | 10.102     | 21.800   | OK:FOUND     |
| 16           | LYS  | JAR   | 263      | SER  | JAR   | 267      | 12.526      | 13.867     | 18.000   | OK:FOUND     |
| 17           | LYS  | JAR   | 263      | SER  | JAR   | 410      | 45.570      | >45.570    | 18.000   | BAD:EUCL     |
| 18           | SER  | JAR   | 265      | LYS  | JAR   | 266      | 5.452       | 7.439      | 18.000   | OK:FOUND     |
| 19           | SER  | JAR   | 265      | SER  | JAR   | 267      | 7.882       | >21.150    | 14.100   | BAD:NOTFOUND |
| 20           | LYS  | JAR   | 266      | LYS  | JAR   | 357      | 54.080      | >54.080    | 21.800   | BAD:EUCL     |
| 21           | LYS  | JAR   | 266      | LYS  | JAR   | 373      | 48.723      | >48.723    | 21.800   | BAD:EUCL     |
| 22           | LYS  | JAR   | 266      | SER  | JAR   | 410      | 41.242      | >41.242    | 18.000   | BAD:EUCL     |
| 23           | SER  | JAR   | 267      | LYS  | JAR   | 357      | 52.909      | >52.909    | 18.000   | BAD:EUCL     |
| 24           | LYS  | JAR   | 343      | LYS  | JAR   | 384      | 16.690      | 21.109     | 21.800   | OK:FOUND     |
| 25           | LYS  | JAR   | 352      | LYS  | JAR   | 357      | 5.642       | 7.207      | 21.800   | OK:FOUND     |
| 26           | LYS  | JAR   | 356      | LYS  | JAR   | 357      | 5.823       | 9.551      | 21.800   | OK:FOUND     |
| 27           | LYS  | JAR   | 356      | LYS  | JAR   | 366      | 20.739      | 33.612     | 21.800   | BAD:LONG     |
| 28           | LYS  | JAR   | 356      | LYS  | JAR   | 373      | 29.901      | >29.901    | 21.800   | BAD:EUCL     |
| 29           | LYS  | JAR   | 356      | SER  | JAR   | 388      | 19.468      | >19.468    | 18.000   | BAD:EUCL     |

## Continuação da tabela 11

| CROSS-LINK # | RES1 | PROT. | RES1 NUM | RES2 | PROT. | RES2 NUM | DIST EUCLID | DIST. TOPO | DIST MAX | RESULTADO |
|--------------|------|-------|----------|------|-------|----------|-------------|------------|----------|-----------|
| 30           | LYS  | JAR   | 357      | LYS  | JAR   | 366      | 17.194      | 33.762     | 21.800   | BAD:LONG  |
| 31           | LYS  | JAR   | 357      | LYS  | JAR   | 373      | 24.474      | >24.474    | 21.800   | BAD:EUCL  |
| 32           | LYS  | JAR   | 366      | LYS  | JAR   | 373      | 14.696      | 18.494     | 21.800   | OK:FOUND  |
| 33           | LYS  | JAR   | 366      | LYS  | JAR   | 384      | 15.717      | 30.845     | 21.800   | BAD:LONG  |
| 34           | LYS  | JAR   | 366      | LYS  | JAR   | 402      | 8.726       | 12.089     | 21.800   | OK:FOUND  |
| 35           | LYS  | JAR   | 373      | SER  | JAR   | 388      | 20.275      | >20.275    | 18.000   | BAD:EUCL  |
| 36           | LYS  | JAR   | 373      | LYS  | JAR   | 402      | 17.072      | 18.541     | 21.800   | OK:FOUND  |
| 37           | LYS  | JAR   | 373      | SER  | JAR   | 410      | 24.891      | >24.891    | 18.000   | BAD:EUCL  |
| 38           | SER  | JAR   | 376      | LYS  | JAR   | 394      | 19.137      | 20.750     | 18.000   | BAD:LONG  |

**Tabela 12: Análise pelo software Topolink dos *cross-links* encontrados para jararagina livre usando a abordagem XPlex / Zero-length**

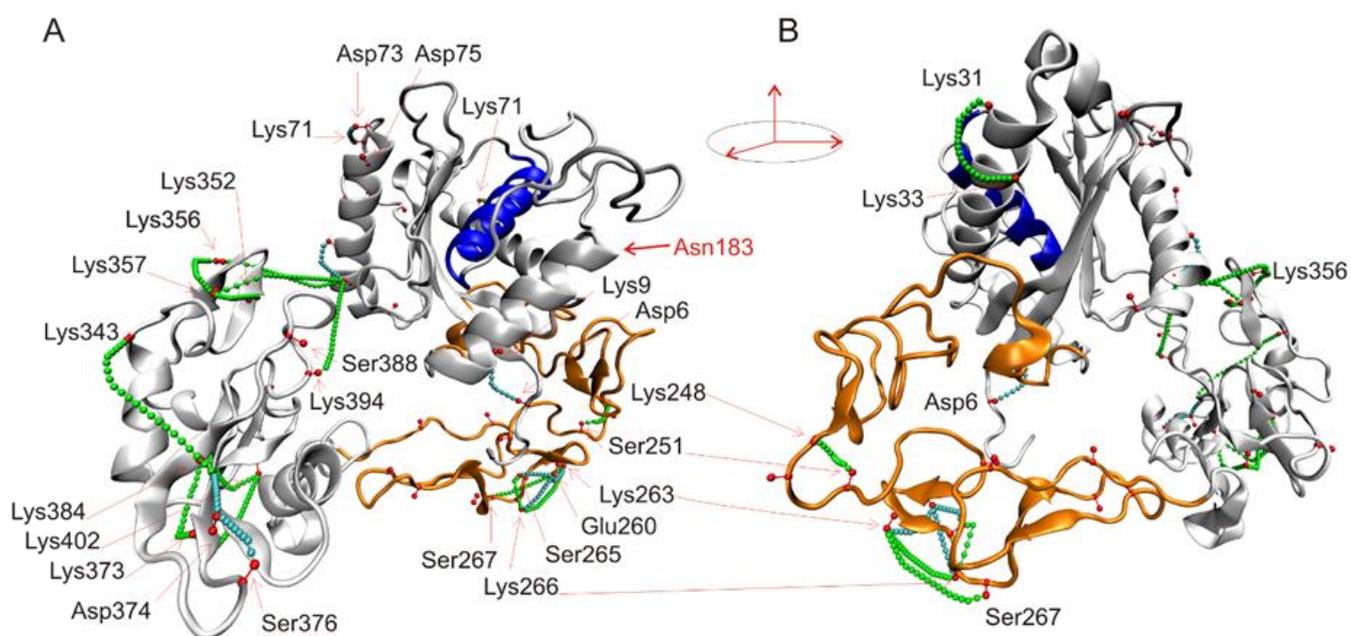
Legenda da tabela: Res1/Res2: resíduos envolvidos no *cross-link*; Res1/Res2 Num: numeração dos resíduos; Prot.: proteína a qual o resíduo pertence; Dist. Euclid.: distância topológica calculada para o *cross-link* (em Å); Dist.topo.: distância topológica calculada (Å); Dist.máx.: distância máxima considerada para o *cross-link* (Å); Resultado: resultado final da análise pelo Topolink.

| CROSS-LINK # | RES1 | PROT. | RES1 NUM | RES2 | PROT. | RES2 NUM | DIST EUCLID | DIST TOPO | DIST MAX | RESULT       |
|--------------|------|-------|----------|------|-------|----------|-------------|-----------|----------|--------------|
| 1            | ASP  | JAR   | 6        | LYS  | JAR   | 9        | 7.865       | 9.113     | 9.700    | OK:FOUND     |
| 2            | SER  | JAR   | 63       | ASP  | JAR   | 73       | 23.867      | >23.867   | 7.000    | BAD:EUCL     |
| 3            | ASP  | JAR   | 66       | LYS  | JAR   | 71       | 10.146      | 11.340    | 9.700    | BAD:LONG     |
| 4            | LYS  | JAR   | 71       | ASP  | JAR   | 73       | 4.384       | 4.678     | 9.700    | OK:FOUND     |
| 5            | LYS  | JAR   | 71       | ASP  | JAR   | 75       | 9.065       | 12.552    | 9.700    | BAD:LONG     |
| 6            | ASP  | JAR   | 75       | SER  | JAR   | 376      | 43.834      | >43.834   | 7.000    | BAD:EUCL     |
| 7            | SER  | JAR   | 80       | GLU  | JAR   | 83       | 6.013       | 7.865     | 7.700    | BAD:LONG     |
| 8            | GLU  | JAR   | 83       | LYS  | JAR   | 86       | 5.625       | 7.176     | 10.500   | OK:FOUND     |
| 9            | LYS  | JAR   | 86       | GLU  | JAR   | 260      | 40.262      | >40.262   | 10.500   | BAD:EUCL     |
| 10           | LYS  | JAR   | 248      | ASP  | JAR   | 257      | 19.727      | >19.727   | 9.700    | BAD:EUCL     |
| 11           | SER  | JAR   | 249      | GLU  | JAR   | 260      | 17.843      | >17.843   | 7.700    | BAD:EUCL     |
| 12           | SER  | JAR   | 249      | GLU  | JAR   | 270      | 29.274      | >29.274   | 7.700    | BAD:EUCL     |
| 13           | ASP  | JAR   | 257      | LYS  | JAR   | 263      | 13.498      | >13.498   | 9.700    | BAD:EUCL     |
| 14           | ASP  | JAR   | 257      | LYS  | JAR   | 266      | 12.846      | >12.846   | 9.700    | BAD:EUCL     |
| 15           | ASP  | JAR   | 257      | LYS  | JAR   | 373      | 42.505      | >42.505   | 9.700    | BAD:EUCL     |
| 16           | GLU  | JAR   | 260      | LYS  | JAR   | 263      | 6.616       | 7.654     | 10.500   | OK:FOUND     |
| 17           | GLU  | JAR   | 260      | SER  | JAR   | 265      | 5.169       | 5.670     | 7.700    | OK:FOUND     |
| 18           | GLU  | JAR   | 260      | LYS  | JAR   | 266      | 8.362       | 8.501     | 10.500   | OK:FOUND     |
| 19           | GLU  | JAR   | 260      | SER  | JAR   | 267      | 12.573      | >12.573   | 7.700    | BAD:EUCL     |
| 20           | LYS  | JAR   | 263      | GLU  | JAR   | 270      | 17.127      | >17.127   | 10.500   | BAD:EUCL     |
| 21           | SER  | JAR   | 265      | GLU  | JAR   | 270      | 9.351       | >9.351    | 7.700    | BAD:EUCL     |
| 22           | SER  | JAR   | 274      | GLU  | JAR   | 290      | 18.220      | >18.220   | 7.700    | BAD:EUCL     |
| 23           | SER  | JAR   | 274      | ASP  | JAR   | 294      | 10.720      | >10.720   | 7.000    | BAD:EUCL     |
| 24           | SER  | JAR   | 274      | ASP  | JAR   | 416      | 27.968      | >27.968   | 7.000    | BAD:EUCL     |
| 25           | ASP  | JAR   | 279      | SER  | JAR   | 388      | 29.233      | >29.233   | 7.000    | BAD:EUCL     |
| 26           | LYS  | JAR   | 373      | ASP  | JAR   | 374      | 6.105       | >14.550   | 9.700    | BAD:NOTFOUND |
| 27           | ASP  | JAR   | 374      | SER  | JAR   | 376      | 5.680       | 5.958     | 7.000    | OK:FOUND     |
| 28           | ASP  | JAR   | 374      | LYS  | JAR   | 384      | 3.979       | 4.054     | 9.700    | OK:FOUND     |

A **Figura 76** ilustra, no modelo estrutural de jararagina, todos os *cross-links* validados pelo *software* Topolink (distâncias euclidianas e topológicas validadas simultaneamente) para as duas estratégias de XL-MS. Para BS<sup>3</sup>, validamos *links* em cada um dos domínios estruturais da toxina, enquanto que para a abordagem XPLex *zero-length*, os *cross-links* validados concentraram-se nos domínios desintegrina e rico em cisteína.

Apesar da expansão do uso da técnica de XL-MS em problemas relevantes de biologia estrutural, a incorporação das restrições de distância em rotinas de modelagem ainda é um desafio (Schneider *et al.*, 2018). Como discutido anteriormente, um *cross-link* pode ser formado entre um dado par de resíduos desde que a distância que separa estes resíduos seja momentaneamente menor do que o braço espaçador do ALC. Isso faz com que a técnica de *cross-linking* seja sensível à dinâmica molecular intrínseca das proteínas, que vai desde movimentos de flexibilidade local a movimentos coordenados de domínios proteicos (Henzler-Wildsman e Kern, 2017). Dessa forma, *cross-links* que excedem a distância máxima esperada podem ser explorados para obter informações sobre a dinâmica destas proteínas (Chavez e Bruce, 2019). Alguns algoritmos já têm sido propostos / desenvolvidos para lidar com flexibilidade molecular durante a validação de *cross-links*, como o DynamXL (Degiacomi *et al.*, 2017).

Outro fator que precisa ser considerado é a existência de *cross-links* conflitantes, restrições que são oriundas de diferentes conformações proteicas em solução, e que, portanto, não podem ser explicadas por um único modelo estrutural, contribuindo para a ambiguidade do dado (Ferber *et al.*, 2016). Uma possível estratégia para minimizar estas restrições conflitantes seria a adoção de ferramentas bioinformáticas capazes de analisar dados de *cross-links* de maneira quantitativa. A abordagem de XL-MS quantitativa é muito recente nos campos da proteômica estrutural e da biologia estrutural integrativa, mas que parece bastante promissora para a distinção de *cross-links* oriundos de conformações majoritárias ou minoritárias em solução (Chen e Rappsilber, 2018).



**Figura 76: Cross-links BS<sup>3</sup> e zero-length validados no modelo estrutural de jararagina.**

O modelo estrutural de jararagina foi representado na forma *cartoon* colorido em cinza. A hélice central do domínio metaloendopeptidase, que abriga o sítio ativo, foi colorida em azul e o domínio rico em cisteínas foi colorido em laranja. Os resíduos envolvidos em *cross-links* foram representados na forma esfera e linha, coloridos em vermelho, pelos átomos Cα e Cβ. Os *cross-links* validados para a estratégia BS<sup>3</sup> foram representados na forma linha e esfera coloridos em verde. Os *cross-links* validados para a estratégia XPLOR- zero length foram representados na forma linha e esfera coloridos em ciano.

## 4.5.2 – Modelagem molecular do inibidor BJ46a

### 4.5.2.1 – Análise da estrutura primária

Iniciamos a análise da estrutura primária do inibidor BJ46a utilizando os algoritmos PSI-BLAST (Altschul *et al.*, 1997) e PSIPRED (Jones, 1999). Como descrito anteriormente, BJ46a possui dois domínios do tipo cistatina, que compreende os resíduos 1-112 e 133-240, respectivamente, e um terceiro domínio denominado de rico em histidina (Valente *et al.* 2001). Os algoritmos classificaram o domínio rico em histidina como não estruturado (*coil*) em toda a sua extensão, como demonstra a **Figura 77**, a seguir.

```

Conf: 999999999957899999999999999947899776288259943489442799639999977
Pred: CCCCCCCCCCHHHHHHHHHHHHHHCCCCCCCEEEEEEEEEEEEEEEECCCCEEEEEEEE
AA: SQVRGDLECDEKDAKEWTDTGVRVYINEHKLHGKYALNVIKNIVVVPWDGDWVAVFLKLN
      10      20      30      40      50      60

Conf: 5534223668999999999989998715762199999945982488843213346740001
Pred: EEEEEEEEEEEEEEEECCCCEEEEEEEECCCCCCCCCCCC
AA: LLETECHVLDPTFVKNCTVRPQHNHAVEMDCDVKIMFNVDTFKEDVFAKCHSTPDSVENV
      70      80      90      100     110     120

Conf: 12597776469999944899999999997259982378858988887530983689999
Pred: CCCCCCCCCCCCCCHHHHHHHHHHHHHHCCCCCCCEEEEEEEEEEEEEEEECCCCEEEEEEEE
AA: RRNCPKCPILLPSNPNQVVDVSEYVVLNKHNEKLSDHVYEVLEISRGQHKYEPEAYYVEFA
      130     140     150     160     170     180

Conf: 99516797435799999979999980468799983037766899999354652342888
Pred: EEEEEEEEEEEEEEEECCCCCCCCCCCCEEEEEEEECCCCCCCCCCCCEEEEEEEECCCC
AA: IVEVNCTAQELHDDHHHCHPNTAGEDHIGFCRATVFRSHASLEKPKDEQFESDCVILHVK
      190     200     210     220     230     240

Conf: 88888886445676887899988888876689999999988877677764234445888
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: EGHASHLIQHQHVEKDSISPEHNNTALNFVHPHNDTSTSHESHEHLAEPVAVFVKKELPK
      250     260     270     280     290     300

Conf: 7889899999999996665339
Pred: CCCCCCCCCCCCCCCCCCCCCC
AA: DISDRHTTPVKGCPGVVHVFEL
      310     320

```

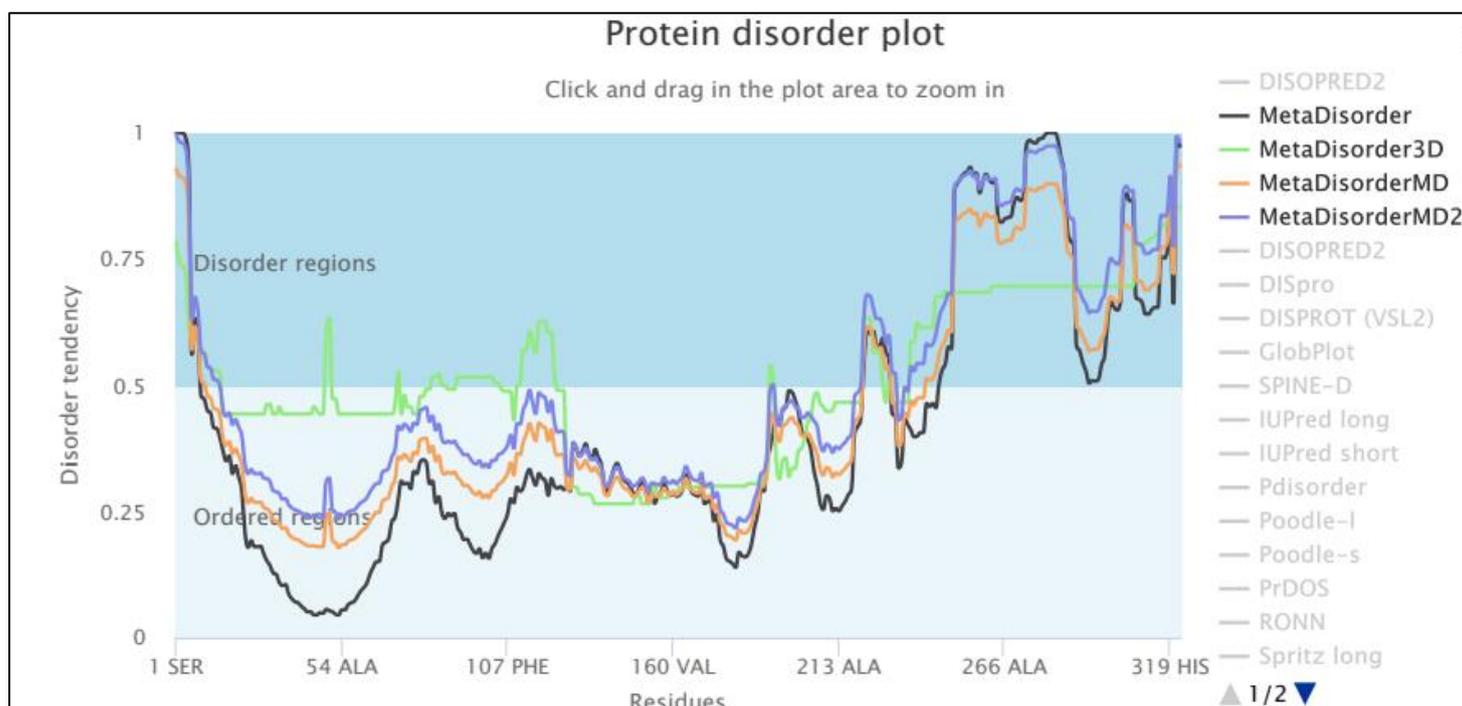
**Figura 77: Análise da estrutura primária de BJ46a pelo algoritmo PSIPRED.**

O algoritmo executa a predição de elementos de estrutura secundária e classifica pontualmente cada resíduo. A linha AA especifica a sequência de BJ46a, a linha Conf indica o *score* de confiança da predição e a linha Pred especifica o elemento de estrutura secundária predito, conforme a legenda: (C) – *coil*; (H) – hélice e (E) – fita.

A partir destas predições, avaliamos se o terceiro domínio de BJ46a poderia ser uma região intrinsecamente desordenada (IDR). Nos últimos anos, várias proteínas foram descritas como intrinsecamente desordenadas ou portadoras de regiões desordenadas. Estima-se que entre 15 a 45% das proteínas eucarióticas contenham um nível de desordem significativo, isto é, regiões de desordem com pelo menos 30 resíduos (Tompa, 2012). Regiões intrinsecamente desordenadas oferecem flexibilidade nas interações proteína-proteína, aumentando o número de estados conformacionais possíveis e permitindo a apresentação de motivos peptídicos lineares, que se ligam a outras proteínas, DNA e pequenos ligantes, aumentando a rede de interação; além disso, as IDR podem possuir taxas evolutivas (via modificações pontuais) mais aceleradas do que o observado para proteínas / regiões globulares (Buljan *et al.*, 2013; Van Der Lee *et al.*, 2014).

Atualmente, há uma série de algoritmos dedicados à predição de IDR em proteínas, baseados em diferentes princípios, como métricas de globularidade, análise de propensão de formação de elementos de estrutura secundária, reconhecimento de enovelamento (*fold recognition*) e uso de redes neurais / *machine learning* para comparação com o banco de dados de proteínas experimentalmente determinadas como desordenadas ou portadoras de regiões desordenadas (Uversky, 2017).

O servidor MetaDisorder (disponível em <http://genesilico.pl/metadisorder/>) reúne uma série de algoritmos de predição de desordem e gera um resultado-consenso a partir de todos os *outputs* dos algoritmos de predição. Submetemos a sequência de BJ46a à predição de desordem neste servidor e obtivemos o resultado ilustrado na **Figura 78**. Os domínios cistatina do inibidor (1-112 e 133-240) apresentaram uma tendência de desordem baixa, corroborando as predições iniciais de estrutura secundária, com exceção da extremidade N-terminal, marcada como desordenada. Porém, como discutido na página 132, os dados de HDX-MS obtidos para BJ46a e estudos prévios com outras cistatinas indicam uma possível desordem desta região. A maior tendência de desordem foi registrada a partir do resíduo 221 (parte final do 2º domínio cistatina) até o final da sequência, no resíduo 322, englobando todo o domínio rico em histidina.



**Figura 78: Predição de regiões intrinsecamente desordenadas no inibidor BJ46a pelo servidor MetaDisorder.**

Neste gráfico, os resíduos que compõem a estrutura primária de BJ46a estão no eixo x, e a tendência à desordem de cada resíduo, no eixo y. Uma probabilidade  $< 0,5$  indica que a região provavelmente é estruturada, enquanto que uma probabilidade  $> 0,5$  indica propensão à desordem. Cada linha colorida no gráfico mostra as probabilidades calculadas para diferentes metaservidores de desordem, que congregam diversos algoritmos de predição de desordem: MetaDisorder (linha em preto), MetaDisorder3D (linha em verde), MetaDisorderMD (linha em laranja), MetaDisorderMD2 (linha em lilás).

É interessante notar que a análise por PSIPRED indicou que a região de transição entre os dois domínios cistatina, composta por aproximadamente vinte resíduos, é predita como não-estruturada, mesmo apresentando uma ponte dissulfeto comprovada experimentalmente (C124-C127). Em sua forma estendida, esta região pode alcançar até 60 Å, o que permite inúmeros posicionamentos relativos dos dois domínios cistatina, aumentando o grau de complexidade da modelagem molecular de BJ46a. Portanto, a modelagem do inibidor não é trivial e precisa ser feita com cautela.

Inicialmente, optamos por modelar o inibidor BJ46a em sua forma completa, tentando buscar alguma conectividade entre os domínios cistatina, mas esta tentativa foi infrutífera. Os C-scores dos modelos gerados pelo I-TASSER nesta abordagem

variaram de -2,54 a -3,50, evidenciando a baixa confiabilidade das estruturas. Na etapa de validação de *cross-links* pelo software Topolink, validamos apenas 18% das restrições geradas pelo ALC BS<sup>3</sup>, e apenas 36% dos links ácido-ácido (diamina) e 10% dos *links zero-length* gerados pela estratégia XPLex, o que indica que o enovelamento obtido para esta estratégia parece não ser corroborado pelos dados experimentais. Dessa forma, optamos por seguir a estratégia de modelagem individual dos dois domínios cistatina de BJ46a; o domínio rico em histidina, predito como desordenado, não foi considerado na análise.

#### 4.5.2.2 – Modelagem molecular do primeiro domínio cistatina (C1) de BJ46a

Para a modelagem molecular do domínio C1 de BJ46a, consideramos o intervalo de resíduos 1-126, que inclui o primeiro domínio cistatina e a região de transição entre o primeiro e o segundo domínio. A **Tabela 13** ilustra os principais alinhamentos retornados pelo algoritmo LOMETS para este segmento.

**Tabela 13: Demonstração dos 10 melhores alinhamentos obtidos pelo algoritmo LOMETS, ordenados pelo Z-score normalizado (Norm. Z-score).**

Algoritmos de *threading*: (1) MUSTER; (2) FFAS-3D; (3) SPARKS-X; (4) HHSEARCH2; (5) HHSEARCH I; (6) Neff-PPAS; (7) HHSEARCH; (8) pGenTHREADER; (9) wwPPAS; (10) PROSPECT2.

| Algoritmo de Threading | PDB Hit               | Id1 Thread | Id2 total | Cobertura | Norm Z-Score |
|------------------------|-----------------------|------------|-----------|-----------|--------------|
| 6                      | <a href="#">2ch9A</a> | 0.12       | 0.17      | 0.87      | 2.97         |
| 7                      | 2ch9                  | 0.11       | 0.17      | 0.85      | 2.95         |
| 4                      | 2ch9                  | 0.12       | 0.17      | 0.90      | 2.57         |
| 1                      | <a href="#">4it7A</a> | 0.14       | 0.17      | 0.83      | 2.34         |
| 3                      | 4it7A                 | 0.14       | 0.17      | 0.84      | 2.30         |
| 2                      | 2ch9A                 | 0.12       | 0.17      | 0.83      | 2.28         |
| 9                      | 2ch9A                 | 0.12       | 0.17      | 0.90      | 2.22         |
| 5                      | <a href="#">1rn7</a>  | 0.15       | 0.17      | 0.87      | 1.85         |
| 10                     | 2ch9A                 | 0.11       | 0.17      | 0.90      | 1.41         |
| 8                      | <a href="#">3gaxA</a> | 0.17       | 0.20      | 0.83      | 1.35         |

Destes alinhamentos, identificamos 4 estruturas candidatas à molde, marcadas em azul na tabela acima. Destas, a estrutura cristalina 2CH9, da cistatina F humana dimérica, com resolução de 2,2 Å (Schuttelkopf *et al.*, 2006) apresentou o melhor Z-

score, de 2,97. A figura abaixo demonstra o alinhamento da sequência do primeiro domínio cistatina de BJ46a e a sequência da estrutura 2CH9A.

|       |    |   |    |
|-------|----|---|----|
| BJ46a | 5  | GDLECDKDAKewTDTGVRYINEHKLHGKYALNVIKNIVVVPWDWVAVFLKLNLETEC   | 65 |
| 2CH9  | 14 | KTIKINDPGVLQAARYSVEKFNCTNDMFLFKESRITRALVQIVKGLKYMLEVEIGRITC | 73 |
| BJ46a | 66 | HVLDPTPVKNCTVVRPQHNHAVEMDCDVKIMFNVDTFKEDVFAKCHSTPDS         |    |
| 2CH9  | 74 | KKNQHLRLDDCDFQTNHTLKQTLSCYSEVWVVPWLHFEVPLRCHHHHHH           |    |

Figura 79: Alinhamento de sequências entre BJ46a e a estrutura 2CH9 (cistatina F humana dimérica).

A rotina de modelagem do domínio C1 de BJ46a seguiu a mesma descrita para a SVMP jararagina; após a construção dos modelos, estes foram clusterizados e analisados pelo algoritmo SPICKER. Os *clusters* resultantes tiveram uma variação no C-score de -0,56 a -2,01 (Tabela 14). Como mencionado anteriormente, a métrica do C-score varia de -5,0 a +2,0; dessa forma, os *clusters* 4 e 5 apresentaram a menor pontuação e foram descartados da análise, e o cluster 1, de maior C-score, foi adotado como referência no alinhamento dos 5 grupos para o cálculo do RMSD (*Root-mean-square deviation*).

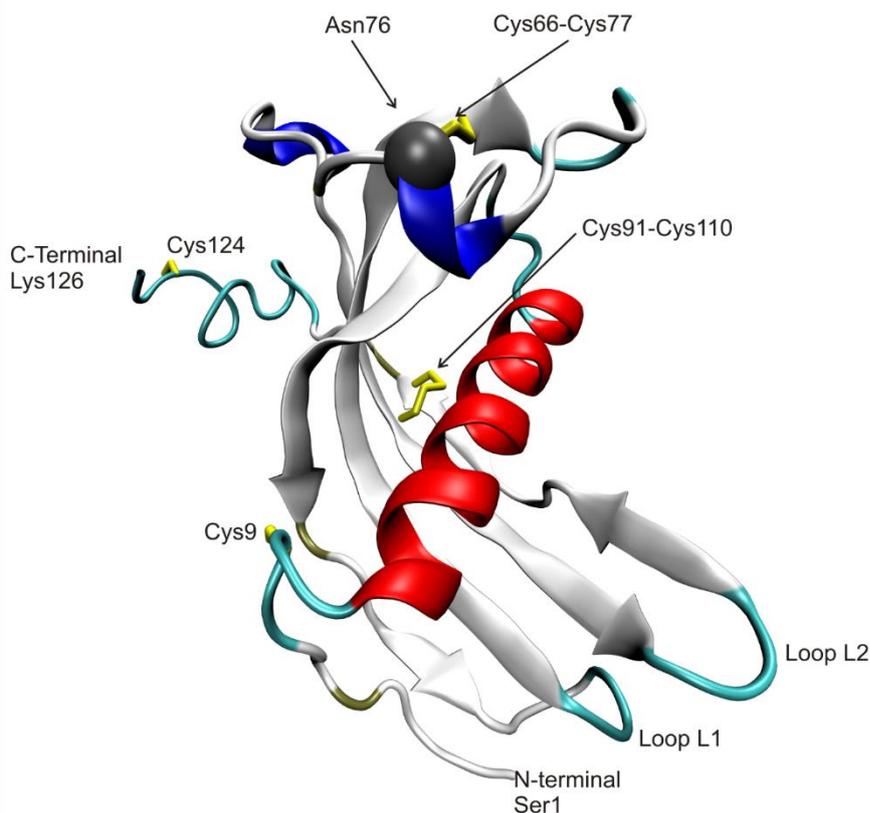
Tabela 14: Análise por *clusterização* dos modelos gerados para o domínio C1 de BJ46a

| Número do <i>Cluster</i> | C-Score | RMSD (Å) |
|--------------------------|---------|----------|
| 1                        | -0,56   | 0        |
| 2                        | -1,32   | 1,707    |
| 3                        | -1,10   | 1,819    |
| 4                        | -2,01   | 1,930    |
| 5                        | -1,87   | 1,720    |

O modelo central do *cluster* 1 está ilustrado na Figura 80. O enovelamento típico dos domínios cistatina engloba uma folha  $\beta$  antiparalela composta de 5 fitas, tendo uma  $\alpha$ -hélice central, quase perpendicular à folha, além de *hairpin loops* (L1/L2) (Nandy,2017). No modelo, este enovelamento está preservado, bem como a região acessória (AS), posicionada do lado oposto aos *loops* L1 e L2, e que abriga o sítio de N-glicosilação do domínio C1, em N76.

É interessante ressaltar a conformação estendida das porções N- e C-terminais do domínio, em concordância com as previsões iniciais de elementos de estrutura secundária pelo PSIPRED e de desordem pelo servidor Meta-Disorder. O

resíduo C9, na porção N-terminal, está envolvido em uma ligação de dissulfeto com o resíduo C313, verificada experimentalmente (subseção 4.3.3, página 108).



**Figura 80: Modelo estrutural do domínio C1 de BJ46a.**

O domínio C1 foi representado na forma *cartoon* onde a  $\alpha$ -hélice central foi colorida em vermelho, a folha  $\beta$  foi colorida em cinza, voltas foram coloridas em ciano e a hélice 3\_10 foi colorida em azul. As pontes dissulfeto e os resíduos de cisteína envolvidos foram representadas na forma de bastões pelos átomos C $\alpha$ , C $\beta$  e S $\gamma$  coloridos em amarelo. O sítio de N-glicosilação foi representado na forma de esfera pelo carbono  $\alpha$  do resíduo Asn76.

#### 4.5.2.2.1 – Validação dos dados de XL-MS no modelo do domínio C1 de BJ46a

Para o ALC BS<sup>3</sup>, obtivemos 7 *cross-links* para o primeiro domínio cistatina do inibidor; para a abordagem XPLex, obtivemos 6 restrições de *links* ácido-ácido (diamina) e 28 restrições *zero-length*. Estas restrições foram então avaliadas pelo Topolink tendo como base o modelo de C1 (**Figura 80**). Os resultados das validações pelo software estão representados nas **tabelas 15-17**. Analisando-se as distâncias euclidianas, 43% das restrições derivadas para o ALC BS<sup>3</sup>, e 38% das restrições geradas pela abordagem XPLex estão dentro dos limites aceitáveis. É importante

ressaltar que, no conjunto de violações de BS<sup>3</sup>, as distâncias topológicas calculadas excedem em muito o limite aceitável pelo *software* Topolink; porém, os resíduos envolvidos nestas restrições violadas estão localizados no N-terminal estendido do domínio e nos *loops*. Dentre o conjunto de restrições de C1, pudemos perceber *hotspots*, isto é, resíduos frequentemente envolvidos nas restrições de distância obtidas com o ALC BS<sup>3</sup> e com a estratégia XPlex. Como demonstra a **Figura 81**, os resíduos 12, 15 e 16 aparecem 6 vezes, o resíduo 104, 6 vezes e os resíduos 13 e 29, 4 vezes.

Na **Figura 82** apresentamos os 11 *cross-links* validados (distâncias euclidianas e topológicas dentro dos limites), identificados pelo *software* Topolink, no modelo estrutural de C1-BJ46a. Os *cross-links* validados para BS<sup>3</sup> foram coloridos em verde, XPlex-diamina coloridos em ciano e XPlex-zero-length coloridos em azul. Observamos que os *cross-links* ocorrem principalmente no N-terminal do domínio, com o maior número de *cross-links* validados nos resíduos K12, E16 e K15. Os dois *loops* L1 e L2 apresentaram 1 *cross-link* validado (D51-D100) e um *cross-link* envolvendo o resíduo D103, mais próximo ao C-terminal do domínio.

**Tabela 15: Análise pelo *software* Topolink das restrições de distância geradas pelo ALC BS<sup>3</sup> e validadas no modelo do domínio C1 de BJ46a.**

Legenda da tabela: Res1/Res2: resíduos envolvidos no *cross-link*; Res1/Res2 Num: numeração dos resíduos; Prot.: proteína a qual o resíduo pertence; Dist. Euclid.: distância topológica calculada para o *cross-link*; Dist.topo.: distância topológica calculada; Dist.máx.: distância máxima considerada para o *cross-link*; Resultado: resultado final da análise pelo Topolink.

| <i>Links</i> | RES1 | PROT. | RES1 NUM | RES2 | PROT. | RES2 NUM | DIST EUCLID | DIST TOPO | DIST MAX | RESULTADO           |
|--------------|------|-------|----------|------|-------|----------|-------------|-----------|----------|---------------------|
| 1            | SER  | BJ46a | 1        | LYS  | BJ46a | 15       | 25.887      | >25.887   | 18.000   | <b>BAD:EUCL</b>     |
| 2            | LYS  | BJ46a | 12       | LYS  | BJ46a | 15       | 5.132       | 6.885     | 21.800   | <b>OK:FOUND</b>     |
| 3            | LYS  | BJ46a | 12       | LYS  | BJ46a | 29       | 25.660      | >25.660   | 21.800   | <b>BAD:EUCL</b>     |
| 4            | LYS  | BJ46a | 12       | LYS  | BJ46a | 109      | 23.732      | >23.732   | 21.800   | <b>BAD:EUCL</b>     |
| 5            | LYS  | BJ46a | 15       | LYS  | BJ46a | 29       | 21.001      | >32.700   | 21.800   | <b>BAD:NOTFOUND</b> |
| 6            | LYS  | BJ46a | 29       | LYS  | BJ46a | 34       | 9.475       | 10.606    | 21.800   | <b>OK:FOUND</b>     |
| 7            | LYS  | BJ46a | 29       | LYS  | BJ46a | 41       | 19.531      | 31.735    | 21.800   | <b>BAD:LONG</b>     |

**Tabela 16: Análise pelo software Topolink das restrições de distância geradas pela abordagem XPlex (diamina) e validadas no modelo do domínio C1 de BJ46a.**

Legenda da tabela: Res1/Res2: resíduos envolvidos no *cross-link*; Res1/Res2 Num: numeração dos resíduos; Prot.: proteína a qual o resíduo pertence; Dist. Euclid.: distância topológica calculada para o *cross-link*; Dist.topo.: distância topológica calculada; Dist.máx.: distância máxima considerada para o *cross-link*; Resultado: resultado final da análise pelo Topolink.

| LINKS | RES1 | PROT. | RES1 NUM | RES2 | PROT. | RES2 NUM | DIST EUCLID | DIST TOPO | DIST MAX | RESULTADO |
|-------|------|-------|----------|------|-------|----------|-------------|-----------|----------|-----------|
| 1     | GLU  | BJ46a | 8        | ASP  | BJ46a | 10       | 5.753       | 6.642     | 15.400   | OK:FOUND  |
| 2     | ASP  | BJ46a | 10       | GLU  | BJ46a | 11       | 6.106       | 9.180     | 15.400   | OK:FOUND  |
| 3     | GLU  | BJ46a | 16       | GLU  | BJ46a | 104      | 13.455      | 20.258    | 16.700   | BAD:LONG  |
| 4     | ASP  | BJ46a | 49       | ASP  | BJ46a | 105      | 15.876      | 20.114    | 14.100   | BAD:LONG  |
| 5     | ASP  | BJ46a | 51       | ASP  | BJ46a | 100      | 8.757       | 8.883     | 14.100   | OK:FOUND  |
| 6     | ASP  | BJ46a | 100      | GLU  | BJ46a | 104      | 13.489      | 16.544    | 15.400   | BAD:LONG  |

**Tabela 17: Análise pelo software Topolink das restrições de distância geradas pela abordagem XPlex (zero-length) e validadas no modelo C1 de BJ46a.**

Legenda da tabela: Res1/Res2: resíduos envolvidos no *cross-link*; Res1/Res2 Num: numeração dos resíduos; Prot.: proteína a qual o resíduo pertence; Dist. Euclid.: distância topológica calculada para o *cross-link*; Dist.topo.: distância topológica calculada; Dist.máx.: distância máxima considerada para o *cross-link*; Resultado: resultado final da análise pelo Topolink.

| Links | RES1 | PROT. | RES1 NUM | RES2 | PROT. | RES2 NUM | DIST EUCLID | DIST TOPO | DIST MAX | RESULTADO |
|-------|------|-------|----------|------|-------|----------|-------------|-----------|----------|-----------|
| 1     | SER  | BJ46a | 1        | ASP  | BJ46a | 10       | 27.396      | >27.396   | 7.000    | BAD:EUCL  |
| 2     | SER  | BJ46a | 1        | ASP  | BJ46a | 13       | 19.433      | >19.433   | 7.000    | BAD:EUCL  |
| 3     | ASP  | BJ46a | 6        | LYS  | BJ46a | 12       | 16.223      | >16.223   | 9.700    | BAD:EUCL  |
| 4     | ASP  | BJ46a | 6        | LYS  | BJ46a | 15       | 15.455      | >15.455   | 9.700    | BAD:EUCL  |
| 5     | GLU  | BJ46a | 8        | SER  | BJ46a | 112      | 26.966      | >26.966   | 7.700    | BAD:EUCL  |
| 6     | ASP  | BJ46a | 10       | LYS  | BJ46a | 15       | 6.792       | 7.486     | 9.700    | OK:FOUND  |
| 7     | GLU  | BJ46a | 11       | LYS  | BJ46a | 12       | 5.896       | 8.200     | 10.500   | OK:FOUND  |
| 8     | GLU  | BJ46a | 11       | LYS  | BJ46a | 15       | 7.140       | 12.887    | 10.500   | BAD:LONG  |
| 9     | LYS  | BJ46a | 12       | ASP  | BJ46a | 13       | 5.874       | 8.036     | 9.700    | OK:FOUND  |
| 10    | LYS  | BJ46a | 12       | GLU  | BJ46a | 16       | 6.456       | 7.746     | 10.500   | OK:FOUND  |
| 11    | LYS  | BJ46a | 12       | GLU  | BJ46a | 65       | 26.189      | >26.189   | 10.500   | BAD:EUCL  |
| 12    | LYS  | BJ46a | 12       | GLU  | BJ46a | 104      | 19.178      | >19.178   | 10.500   | BAD:EUCL  |

## Continuação da tabela 17

| Links | RES1 | PROT. | RES1<br>NUM | RES2 | PROT. | RES2<br>NUM | DIST<br>EUCLID | DIST<br>TOPO | DIST<br>MAX | RESULTADO           |
|-------|------|-------|-------------|------|-------|-------------|----------------|--------------|-------------|---------------------|
| 13    | ASP  | BJ46a | 13          | LYS  | BJ46a | 109         | 20.575         | >20.575      | 9.700       | <b>BAD:EUCL</b>     |
| 14    | ASP  | BJ46a | 13          | SER  | BJ46a | 112         | 27.508         | >27.508      | 7.000       | <b>BAD:EUCL</b>     |
| 15    | LYS  | BJ46a | 15          | GLU  | BJ46a | 16          | 5.321          | 7.342        | 10.500      | <b>OK:FOUND</b>     |
| 16    | LYS  | BJ46a | 15          | GLU  | BJ46a | 65          | 21.875         | >21.875      | 10.500      | <b>BAD:EUCL</b>     |
| 17    | LYS  | BJ46a | 15          | GLU  | BJ46a | 104         | 17.578         | >17.578      | 10.500      | <b>BAD:EUCL</b>     |
| 18    | GLU  | BJ46a | 16          | LYS  | BJ46a | 29          | 20.394         | >20.394      | 10.500      | <b>BAD:EUCL</b>     |
| 19    | GLU  | BJ46a | 16          | LYS  | BJ46a | 34          | 17.787         | >17.787      | 10.500      | <b>BAD:EUCL</b>     |
| 20    | GLU  | BJ46a | 16          | LYS  | BJ46a | 41          | 16.622         | >16.622      | 10.500      | <b>BAD:EUCL</b>     |
| 21    | GLU  | BJ46a | 16          | LYS  | BJ46a | 103         | 8.297          | 9.706        | 10.500      | <b>OK:FOUND</b>     |
| 22    | LYS  | BJ46a | 29          | ASP  | BJ46a | 49          | 27.962         | >27.962      | 9.700       | <b>BAD:EUCL</b>     |
| 23    | LYS  | BJ46a | 29          | ASP  | BJ46a | 100         | 29.947         | >29.947      | 9.700       | <b>BAD:EUCL</b>     |
| 24    | LYS  | BJ46a | 29          | GLU  | BJ46a | 104         | 18.136         | >18.136      | 10.500      | <b>BAD:EUCL</b>     |
| 25    | LYS  | BJ46a | 34          | GLU  | BJ46a | 63          | 8.295          | >15.750      | 10.500      | <b>BAD:NOTFOUND</b> |
| 26    | LYS  | BJ46a | 34          | GLU  | BJ46a | 104         | 21.288         | >21.288      | 10.500      | <b>BAD:EUCL</b>     |
| 27    | LYS  | BJ46a | 41          | GLU  | BJ46a | 104         | 21.507         | >21.507      | 10.500      | <b>BAD:EUCL</b>     |
| 28    | GLU  | BJ46a | 65          | SER  | BJ46a | 112         | 18.290         | >18.290      | 7.700       | <b>BAD:EUCL</b>     |

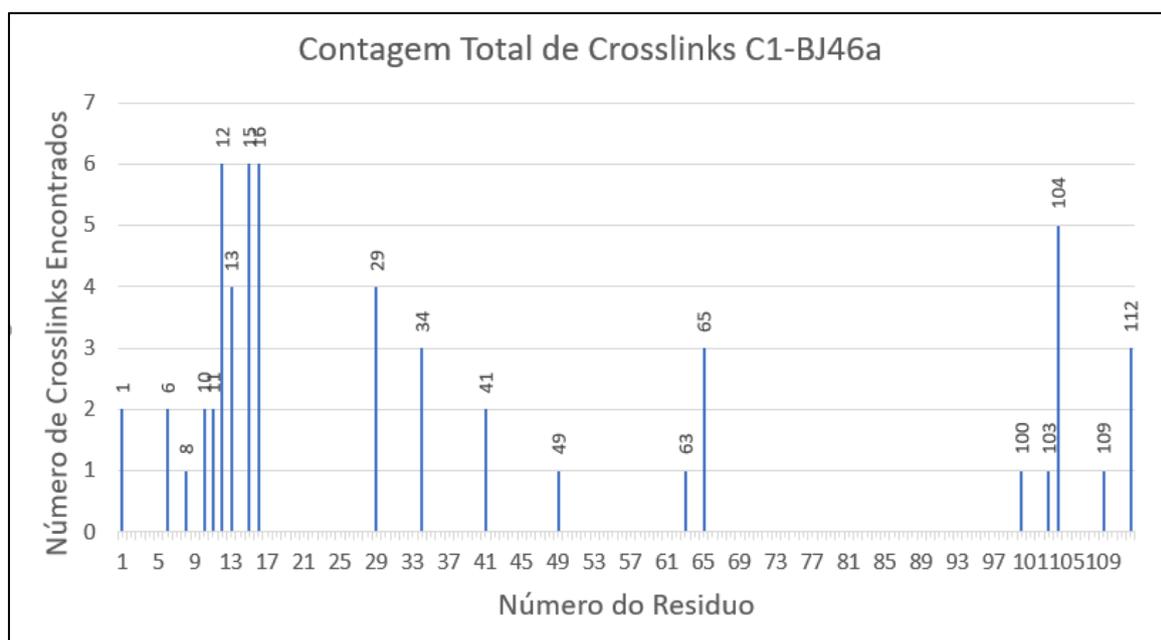
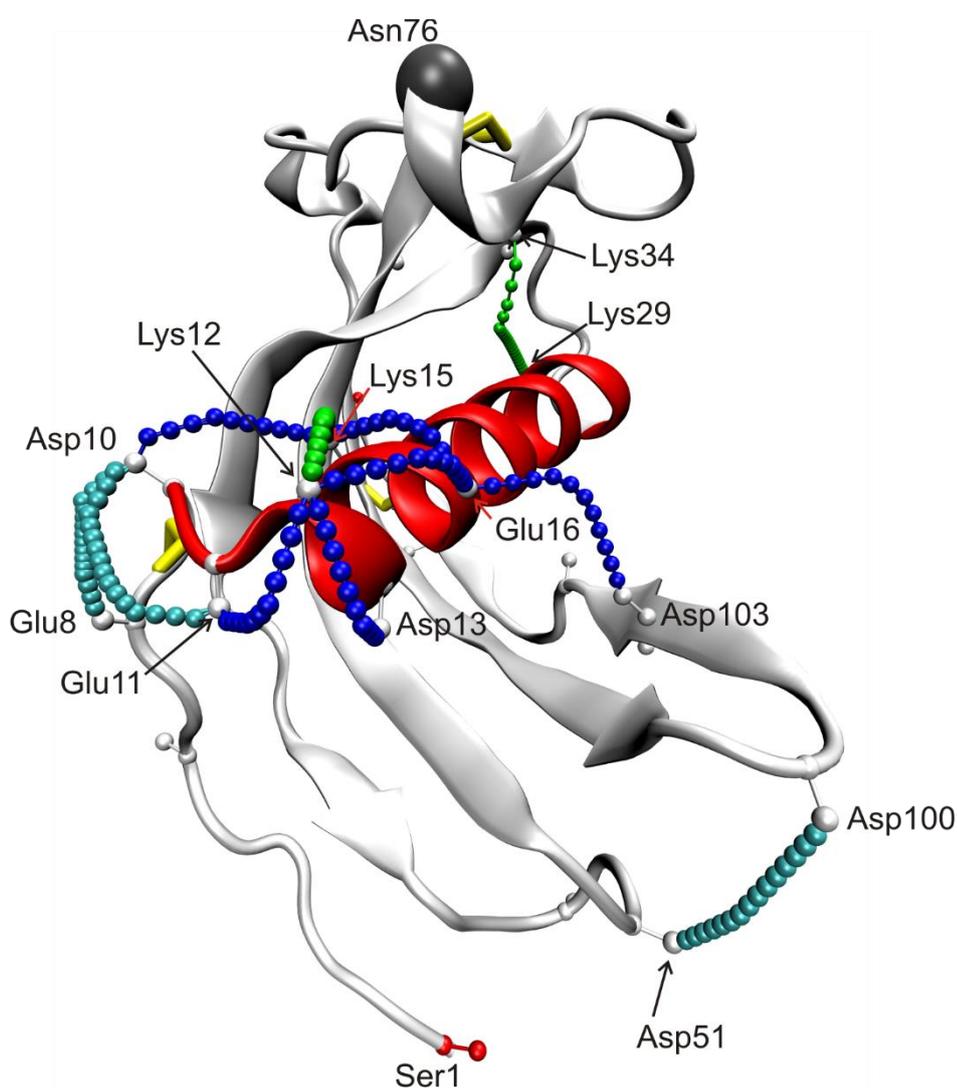


Figura 81: Número de ocorrências dos resíduos de aminoácidos do domínio C1 de BJ46a para a estratégias BS<sup>3</sup> e XPLex (diamina e zero-length).



**Figura 82: *Cross-links* validados no modelo final do domínio C1 de BJ46a.**

O domínio C1 foi representado na forma *cartoon* e colorido em cinza, com a  $\alpha$ -hélice central colorida em vermelho. As pontes dissulfeto foram representadas pelos átomos  $C\alpha$ ,  $C\beta$  e  $S\gamma$  das cisteínas em forma de bastões coloridos em amarelo. O sítio de N-glicosilação foi representado pelo carbono  $\alpha$  do resíduo N76. Os resíduos envolvidos em *cross-links* foram representados pelos carbonos  $\alpha$  e  $\beta$ , na forma linha e esfera, coloridos em branco. Os *cross-links* validados pelo Topolink foram representados na forma esfera e linha e coloridos de acordo com a estratégia: BS<sup>3</sup> colorido em verde, XPLex-diamina em ciano e XPLex-zero-length em azul escuro.

#### 4.5.2.3 – Modelagem molecular do segundo domínio cistatina (C2) de BJ46a

Para a modelagem molecular do domínio C2 de BJ46a, consideramos os resíduos 133 – 240, que compreendem o segundo domínio cistatina e a região de transição para o primeiro domínio cistatina. Os principais alinhamentos identificados pelo algoritmo LOMETS foram resumidos na tabela abaixo:

**Tabela 18: Melhores alinhamentos obtidos pelo algoritmo LOMETS, ordenados pelo Z-score normalizado.**

Algoritmo de *threading*: (1) MUSTER; (2) FFAS-3D; (3) SPARKS-X; (4) HHSEARCH2; (5) HHSEARCH I; (6) Neff-PPAS; (7) HHSEARCH, (8) pGenTHREADER; (9) wdPPAS; (10) PROSPECT2.

| <b>Rank</b> | <b>PDB Hit</b>        | <b>Id1 Thread</b> | <b>Id2 total</b> | <b>Cobertura</b> | <b>Norm Z-Score</b> |
|-------------|-----------------------|-------------------|------------------|------------------|---------------------|
| 6           | <a href="#">2ch9A</a> | 0.20              | 0.23             | 0.96             | 3.38                |
| 7           | <a href="#">4ccv</a>  | 0.19              | 0.30             | 0.83             | 3.14                |
| 4           | 2ch9                  | 0.19              | 0.23             | 0.97             | 2.52                |
| 3           | <a href="#">4n6oB</a> | 0.15              | 0.17             | 0.92             | 2.37                |
| 9           | 4n6oB                 | 0.15              | 0.17             | 0.92             | 2.24                |
| 5           | 4ccv                  | 0.19              | 0.30             | 0.84             | 2.08                |
| 2           | 2ch9A                 | 0.21              | 0.21             | 0.92             | 2.08                |
| 1           | 2ch9A                 | 0.21              | 0.23             | 0.97             | 1.99                |
| 10          | 2ch9A                 | 0.20              | 0.23             | 0.97             | 1.46                |
| 8           | <a href="#">1cewl</a> | 0.16              | 0.19             | 0.85             | 1.35                |

O algoritmo identificou 4 estruturas no PDB candidatas à molde para a modelagem do segundo domínio cistatina de BJ46a: (I) a estrutura cristalográfica 1CEW, da cistatina de ovo de galinha (*Gallus gallus*) com resolução de 2,0 Å (Bode *et al.*, 1988), apresentando 19% de identidade sequencial e 85% de cobertura; (II) A estrutura cristalográfica 4N6O, referente à proteína legumina humana em complexo com a cistatina E/M (Dall *et al.*, 2015), com resolução de 1,8 Å e apresentando 17% de identidade sequencial e 92% de cobertura; (III) a estrutura cristalográfica 4CCV, referente ao domínio N2 da glicoproteína rica em histidinas de *Oryctolagus cuniculus* em uma resolução de 1,93 Å (Kassar *et al.*, 2014) com 30% de identidade sequencial e 83% de cobertura e, por fim, (IV) a estrutura cristalográfica 2CH9, referente à cistatina F humana dimérica, em uma resolução de 2,1 Å (Schuttelkopf *et al.*, 2006), com 23% de identidade sequencial e 96% de cobertura.

A estrutura que apresentou o maior *Z-score* foi a estrutura 2CH9, sendo então o molde escolhido para modelarmos o segundo domínio cistatina de BJ46a. A **Figura 83**, abaixo, demonstra o alinhamento das duas sequências.

|          |   |
|----------|---|
| C2-BJ46a | RRNCPKCPILLPSNNPQVVDSVEYVLNKHNEKLSDHVYEVLEISRGQHKYEPEAYYVEFAIVEVNCTAQELHDDH |
| 2Ch9A    | -RVKPGFPKTIKTNDPGVLQAARYSVEKFNNCTNMFLFKESRITRALVQIVGLKYMLEVEIGRTTCKKNQ-HLRL |
| C2-BJ46a | HHCHPNTAGEDHIGFCRATVFRSHASLEKPKDEQFESDCVILHVK                               |
| 2Ch9A    | DDCDFQTNHTKQTLSCYSEVWVVPWLQ---HFVVPVLRCHHHHHH                               |

**Figura 83:** Alinhamento das sequências do molde 2CH9 e do domínio C2 de BJ46a.

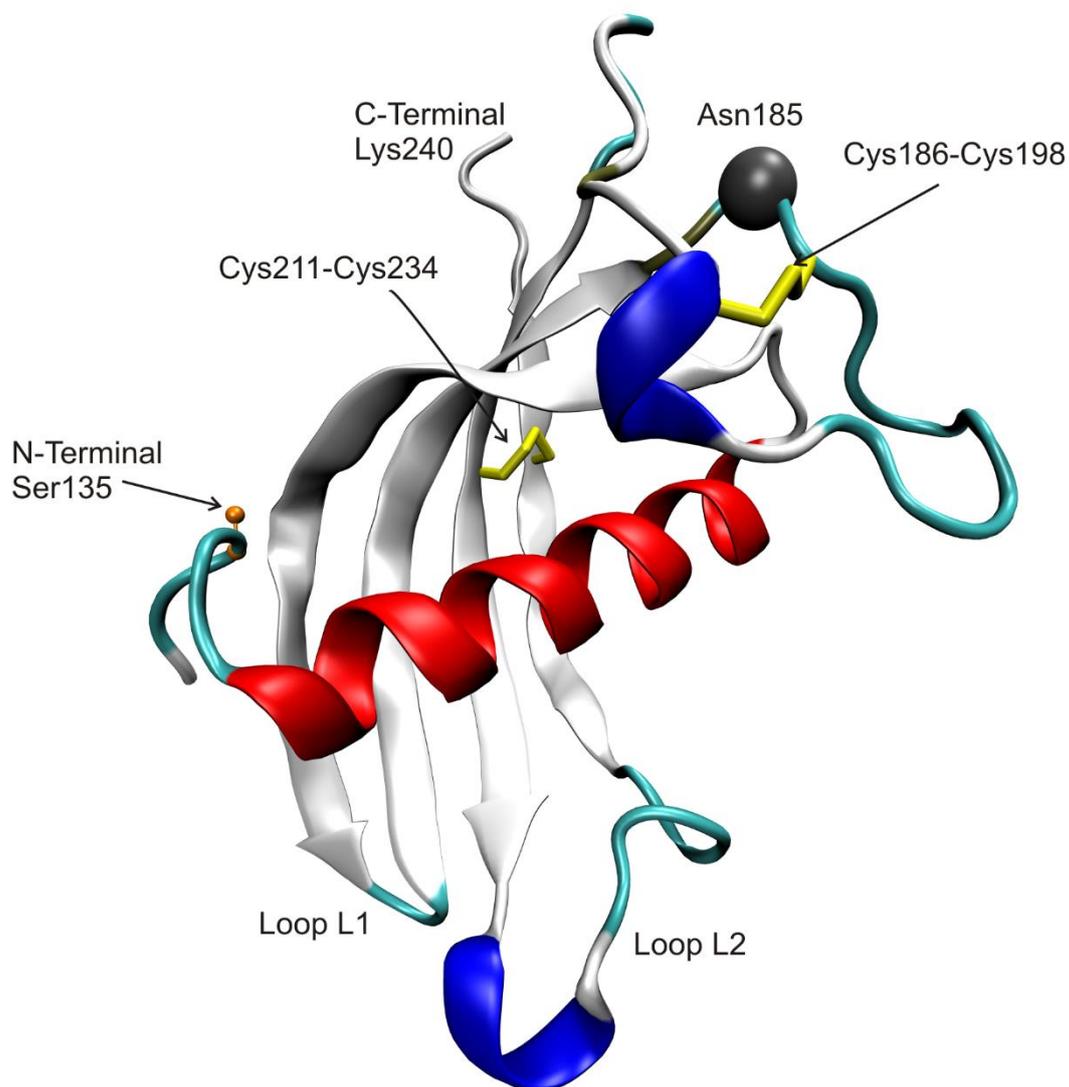
A análise de clusterização pelo algoritmo SPICKER gerou 5 *clusters* (**tabela 19**). Com exceção do *cluster 1*, todos os outros receberam um *C-score* de -5,0, a menor pontuação permitida por esta métrica e, portanto, não foram considerados nas análises subsequentes. Ao compararmos os *C-scores* obtidos para os modelos dos domínios cistatina, o domínio C2 apresentou uma pontuação superior à obtida para o domínio C1, o que se refletiu nos melhores resultados de alinhamento e modelagem molecular final.

**Tabela 19:** Análise da Modelagem Molecular de C2-BJ46a por clusterização.

O *cluster 1* foi tomado como referência para o cálculo do RMSD.

| Número do <i>Cluster</i> | <i>C-Score</i> | RMSD (Å) |
|--------------------------|----------------|----------|
| 1                        | 0.54           | 0        |
| 2                        | -5             | 5.643    |
| 3                        | -5             | 6.413    |
| 4                        | -5             | 3.212    |
| 5                        | -5             | 6.571    |

A **Figura 84** ilustra o modelo estrutural obtido para o domínio C2 de BJ46a, oriundo do *cluster 1*. Assim como descrito para o domínio C1, há a preservação do enovelamento típico de domínios cistatina também em C2, com a folha  $\beta$  e a  $\alpha$ -hélice centrais, os *loops* L1 e L2 e a região AS, oposta aos *loops*, abrigoando o sítio de N-glicosilação do domínio.



**Figura 84: Modelo estrutural do domínio C2 de BJ46a.**

O domínio C2 foi representado na forma *cartoon*, onde a  $\alpha$ -hélice central foi colorida em vermelho, a folha  $\beta$  foi colorida em cinza, *loops* coloridos em ciano e a hélice 3<sub>10</sub> em azul. As pontes dissulfeto e os resíduos de cisteína foram representados na forma de bastões pelos átomos C $\alpha$ , C $\beta$  e S $\gamma$  coloridos em amarelo. O sítio de N-glicosilação em N185 na região AS foi representado na forma de esfera pelo C $\alpha$  do resíduo.

## 4.5.2.4.1 – Validação dos dados de XL-MS no domínio C2 de BJ46a

Para o domínio C2, as abordagens de XL-MS geraram 50 restrições totais, sendo 5 restrições para o ALC BS<sup>3</sup> e 46 para a abordagem XPlex (5 *links* diamina e 41 *links zero-length*). As distâncias euclidianas encontraram-se dentro do limite para 60% dos *links* BS<sup>3</sup> e XPlex-diamina, e para 34% dos *links* XPlex – *zero length*.

**Tabela 20: Análise pelo software Topolink das restrições de distância geradas pelo ALC BS<sup>3</sup> e validadas no modelo estrutural de C2 – BJ46a.**

Legenda da tabela: Res1/Res2: resíduos envolvidos no *cross-link*; Res1/Res2 Num: numeração dos resíduos; Prot: proteína a qual o resíduo pertence; Dist. Euclid.: distância topológica calculada para o *cross-link*; Dist.topo.: distância topológica calculada; Dist.máx.: distância máxima considerada para o *cross-link*; Resultado: resultado final da análise pelo Topolink.

| LINKS | RES1 | PROT. | RES1 NUM | RES2 | PROT. | RES2 NUM | DIST EUCLID | DIST TOPO | DIST MAX | RESULTADO       |
|-------|------|-------|----------|------|-------|----------|-------------|-----------|----------|-----------------|
| 1     | SER  | BJ46a | 133      | LYS  | BJ46a | 152      | 23.662      | >23.662   | 18.000   | <b>BAD:EUCL</b> |
| 2     | SER  | BJ46a | 133      | SER  | BJ46a | 154      | 28.865      | >28.865   | 14.100   | <b>BAD:EUCL</b> |
| 3     | SER  | BJ46a | 221      | LYS  | BJ46a | 224      | 8.771       | 13.824    | 18.000   | <b>OK:FOUND</b> |
| 4     | LYS  | BJ46a | 224      | LYS  | BJ46a | 226      | 8.119       | 13.904    | 21.800   | <b>OK:FOUND</b> |
| 5     | LYS  | BJ46a | 224      | SER  | BJ46a | 232      | 16.832      | 27.348    | 18.000   | <b>BAD:LONG</b> |

**Tabela 21: Análise pelo software Topolink das restrições de distância geradas pela abordagem XPlex – diamina e validadas no modelo estrutural de C2 – BJ46a**

Legenda da tabela: Res1/Res2: resíduos envolvidos no *cross-link*; Res1/Res2 Num: numeração dos resíduos; Prot: proteína a qual o resíduo pertence; Dist. Euclid.: distância topológica calculada para o *cross-link*; Dist.topo.: distância topológica calculada; Dist.máx.: distância máxima considerada para o *cross-link*; Resultado: resultado final da análise pelo Topolink.

| LINKS | RES1 | PROT  | RES1 NUM | RES2 | PROT. | RES2 NUM | DIST EUCLID | DIST TOPO | DIST MAX | RESULTADO       |
|-------|------|-------|----------|------|-------|----------|-------------|-----------|----------|-----------------|
| 1     | ASP  | BJ46a | 140      | ASP  | BJ46a | 155      | 23.359      | >23.359   | 14.100   | <b>BAD:EUCL</b> |
| 2     | ASP  | BJ46a | 140      | GLU  | BJ46a | 223      | 12.150      | 14.792    | 15.400   | <b>OK:FOUND</b> |
| 3     | GLU  | BJ46a | 143      | ASP  | BJ46a | 155      | 17.922      | >17.922   | 15.400   | <b>BAD:EUCL</b> |
| 4     | GLU  | BJ46a | 228      | GLU  | BJ46a | 231      | 10.631      | 14.026    | 16.700   | <b>OK:FOUND</b> |
| 5     | GLU  | BJ46a | 231      | ASP  | BJ46a | 233      | 6.896       | 7.588     | 15.400   | <b>OK:FOUND</b> |

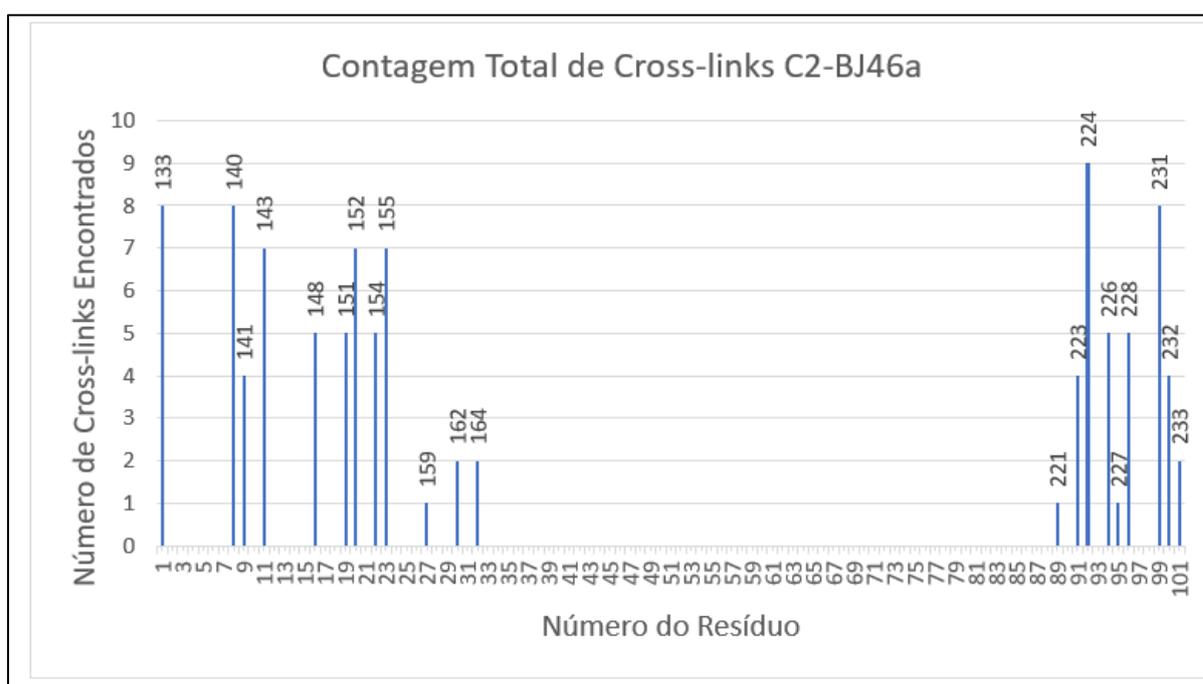
**Tabela 22: Análise pelo software Topolink das restrições de distância geradas pela abordagem XPlex – zero-length e validadas no modelo estrutural de C2 – BJ46a.**

Legenda da tabela: Res1/Res2: resíduos envolvidos no *cross-link*; Res1/Res2 Num: numeração dos resíduos; Prot: proteína a qual o resíduo pertence; Dist. Euclid.: distância topológica calculada para o *cross-link*; Dist.topo.: distância topológica calculada; Dist.máx.: distância máxima considerada para o *cross-link*; Resultado: resultado final da análise pelo Topolink.

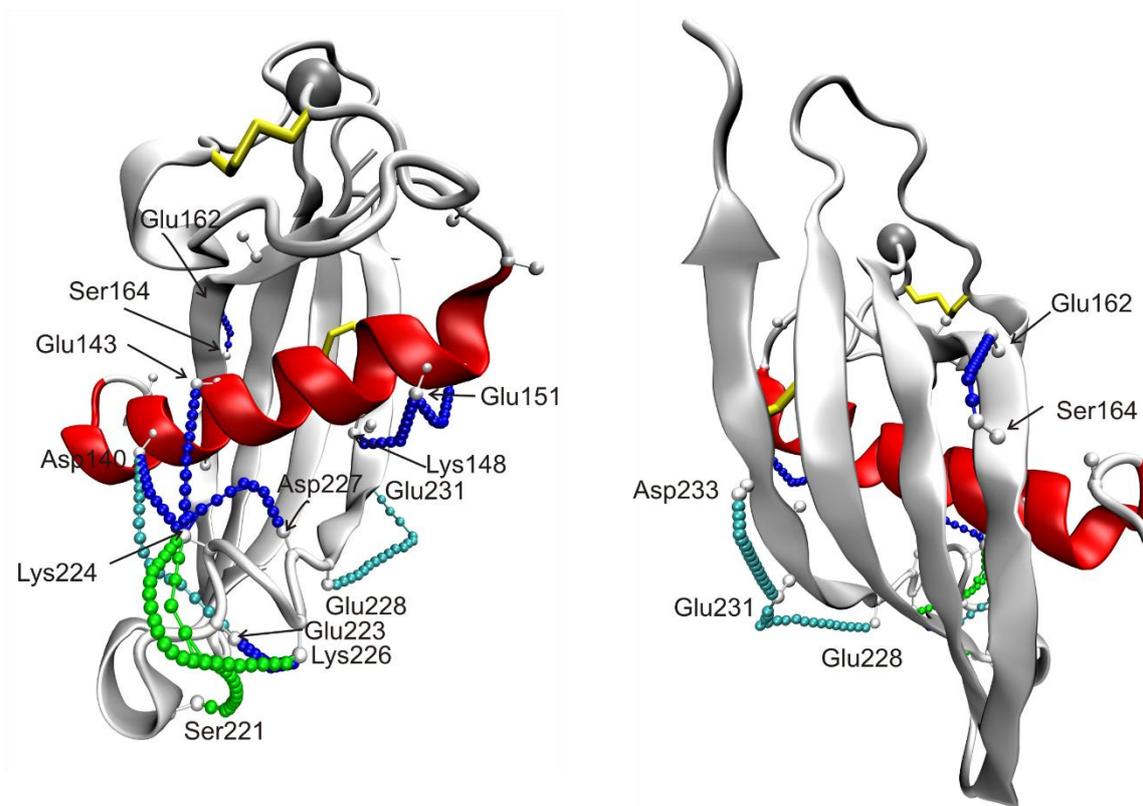
| Links | RES1 | PROT. | RES1 NUM | RES2 | PROT. | RES2 NUM | DIST EUCLID | DIST TOPO | DIST MAX | RESULTADO |
|-------|------|-------|----------|------|-------|----------|-------------|-----------|----------|-----------|
| 1     | SER  | BJ46a | 133      | ASP  | BJ46a | 140      | 12.558      | >12.558   | 7.000    | BAD:EUCL  |
| 2     | SER  | BJ46a | 133      | GLU  | BJ46a | 143      | 13.464      | >13.464   | 7.700    | BAD:EUCL  |
| 3     | SER  | BJ46a | 133      | ASP  | BJ46a | 155      | 25.470      | >25.470   | 7.000    | BAD:EUCL  |
| 4     | SER  | BJ46a | 133      | GLU  | BJ46a | 223      | 22.987      | >22.987   | 7.700    | BAD:EUCL  |
| 5     | SER  | BJ46a | 133      | GLU  | BJ46a | 228      | 21.706      | >21.706   | 7.700    | BAD:EUCL  |
| 6     | SER  | BJ46a | 133      | GLU  | BJ46a | 231      | 19.018      | >19.018   | 7.700    | BAD:EUCL  |
| 7     | ASP  | BJ46a | 140      | LYS  | BJ46a | 148      | 13.129      | >13.129   | 9.700    | BAD:EUCL  |
| 8     | ASP  | BJ46a | 140      | LYS  | BJ46a | 152      | 19.597      | >19.597   | 9.700    | BAD:EUCL  |
| 9     | ASP  | BJ46a | 140      | SER  | BJ46a | 154      | 24.898      | >24.898   | 7.000    | BAD:EUCL  |
| 10    | ASP  | BJ46a | 140      | LYS  | BJ46a | 224      | 7.805       | 7.921     | 9.700    | OK:FOUND  |
| 11    | ASP  | BJ46a | 140      | LYS  | BJ46a | 226      | 14.846      | >14.846   | 9.700    | BAD:EUCL  |
| 12    | ASP  | BJ46a | 140      | SER  | BJ46a | 232      | 16.058      | >16.058   | 7.000    | BAD:EUCL  |
| 13    | SER  | BJ46a | 141      | GLU  | BJ46a | 151      | 15.407      | >15.407   | 7.700    | BAD:EUCL  |
| 14    | SER  | BJ46a | 141      | ASP  | BJ46a | 155      | 21.932      | >21.932   | 7.000    | BAD:EUCL  |
| 15    | SER  | BJ46a | 141      | GLU  | BJ46a | 223      | 12.080      | >12.080   | 7.700    | BAD:EUCL  |
| 16    | SER  | BJ46a | 141      | GLU  | BJ46a | 228      | 11.048      | >11.048   | 7.700    | BAD:EUCL  |
| 17    | GLU  | BJ46a | 143      | LYS  | BJ46a | 148      | 9.864       | 12.580    | 10.500   | BAD:LONG  |
| 18    | GLU  | BJ46a | 143      | LYS  | BJ46a | 152      | 15.429      | >15.429   | 10.500   | BAD:EUCL  |
| 19    | GLU  | BJ46a | 143      | SER  | BJ46a | 154      | 19.840      | >19.840   | 7.700    | BAD:EUCL  |
| 20    | GLU  | BJ46a | 143      | LYS  | BJ46a | 224      | 9.264       | 9.387     | 10.500   | OK:FOUND  |
| 21    | GLU  | BJ46a | 143      | LYS  | BJ46a | 226      | 15.746      | >15.746   | 10.500   | BAD:EUCL  |
| 22    | LYS  | BJ46a | 148      | GLU  | BJ46a | 151      | 5.340       | 6.676     | 10.500   | OK:FOUND  |
| 23    | LYS  | BJ46a | 148      | ASP  | BJ46a | 155      | 13.844      | >13.844   | 9.700    | BAD:EUCL  |
| 24    | LYS  | BJ46a | 148      | GLU  | BJ46a | 159      | 14.123      | >14.123   | 10.500   | BAD:EUCL  |
| 25    | GLU  | BJ46a | 151      | LYS  | BJ46a | 152      | 5.398       | 7.297     | 10.500   | OK:FOUND  |
| 26    | GLU  | BJ46a | 151      | SER  | BJ46a | 164      | 22.464      | >22.464   | 7.700    | BAD:EUCL  |
| 27    | GLU  | BJ46a | 151      | LYS  | BJ46a | 224      | 14.229      | 14.559    | 10.500   | BAD:LONG  |
| 28    | LYS  | BJ46a | 152      | ASP  | BJ46a | 155      | 9.400       | 12.250    | 9.700    | BAD:LONG  |
| 29    | LYS  | BJ46a | 152      | GLU  | BJ46a | 162      | 20.118      | >20.118   | 10.500   | BAD:EUCL  |
| 30    | LYS  | BJ46a | 152      | GLU  | BJ46a | 231      | 12.112      | 13.117    | 10.500   | BAD:LONG  |
| 31    | SER  | BJ46a | 154      | ASP  | BJ46a | 155      | 5.701       | 8.599     | 7.000    | BAD:LONG  |
| 32    | SER  | BJ46a | 154      | GLU  | BJ46a | 231      | 19.601      | >19.601   | 7.700    | BAD:EUCL  |
| 33    | GLU  | BJ46a | 162      | SER  | BJ46a | 164      | 6.091       | 6.605     | 7.700    | OK:FOUND  |
| 34    | GLU  | BJ46a | 223      | LYS  | BJ46a | 226      | 3.790       | 4.425     | 10.500   | OK:FOUND  |
| 35    | LYS  | BJ46a | 224      | ASP  | BJ46a | 227      | 5.436       | 8.039     | 9.700    | OK:FOUND  |
| 36    | LYS  | BJ46a | 224      | GLU  | BJ46a | 228      | 10.511      | >10.511   | 10.500   | BAD:EUCL  |
| 37    | LYS  | BJ46a | 224      | GLU  | BJ46a | 231      | 18.488      | >18.488   | 10.500   | BAD:EUCL  |
| 38    | LYS  | BJ46a | 226      | GLU  | BJ46a | 231      | 17.367      | >17.367   | 10.500   | BAD:EUCL  |
| 39    | GLU  | BJ46a | 228      | SER  | BJ46a | 232      | 11.518      | >11.518   | 7.700    | BAD:EUCL  |
| 40    | GLU  | BJ46a | 231      | SER  | BJ46a | 232      | 5.527       | 9.341     | 7.700    | BAD:LONG  |
| 41    | SER  | BJ46a | 232      | ASP  | BJ46a | 233      | 6.127       | 10.917    | 7.000    | BAD:LONG  |

Com relação à distribuição dos *cross-links* na estrutura de C2, o perfil observado no domínio C1 repete-se no domínio C2, com *cross-links* ocorrendo principalmente na porção N-terminal, centrado na  $\alpha$ -hélice (resíduos 135 a 154) e nas duas últimas fitas  $\beta$  (resíduos 207 a 234). O resíduo 224 ocorreu 9 vezes na lista de restrições de distância, enquanto que os resíduos 133,140,231 ocorreram 8 vezes, e os resíduos 143, 152,155 e 231, 7 vezes (**Figura 85**).

Na **Figura 86** apresentamos os 12 *cross-links* validados (distância euclidiana e topológica dentro dos limites) no modelo estrutural de C2. As restrições validadas ocorrem principalmente na porção N-terminal do domínio, sendo que o maior número de *cross-links* validados ocorre entre a porção N-terminal de C2 e os *loops* L1 e L2. Analisando-se o grupo de violações nas duas estratégias de *cross-linking*, os resíduos envolvidos nestas restrições estão localizados principalmente em regiões não-estruturadas do domínio.



**Figura 85:** Número de ocorrências dos resíduos de aminoácidos do domínio C2 de BJ46a para a estratégias BS<sup>3</sup> e XPLex (diamina e zero-length).



**Figura 86: Cross-links validados no modelo final do domínio C2 de BJ46a.**

A imagem apresenta uma visão em 180° do modelo selecionado do domínio C2 de BJ46a. Este foi representado na forma *cartoon* colorido em cinza, com a  $\alpha$ -hélice central colorida em vermelho. As pontes dissulfeto foram representadas pelos átomos  $C\alpha$ ,  $C\beta$  e  $S\gamma$  da cisteína na forma de bastões coloridos em amarelo. O sítio de N-glicosilação foi representado pelo  $C\alpha$  do resíduo N185. Os resíduos envolvidos em *cross-links* foram representados pelos carbonos  $\alpha$  e  $\beta$ , na forma linha e esfera, coloridos em cinza. Os *cross-links* validados foram representados na forma linha e esfera e coloridos de acordo com a estratégia de XL envolvida: BS<sup>3</sup> colorido em verde, XPLex-diamina colorido em ciano e XPLex-zero-length colorido em azul escuro.

### 4.5.3 – Docking molecular de BJ46a – jararagina

#### 4.5.3.1 – Docking molecular entre o primeiro domínio cistatina de BJ46a (C1) e a metaloendopeptidase jararagina

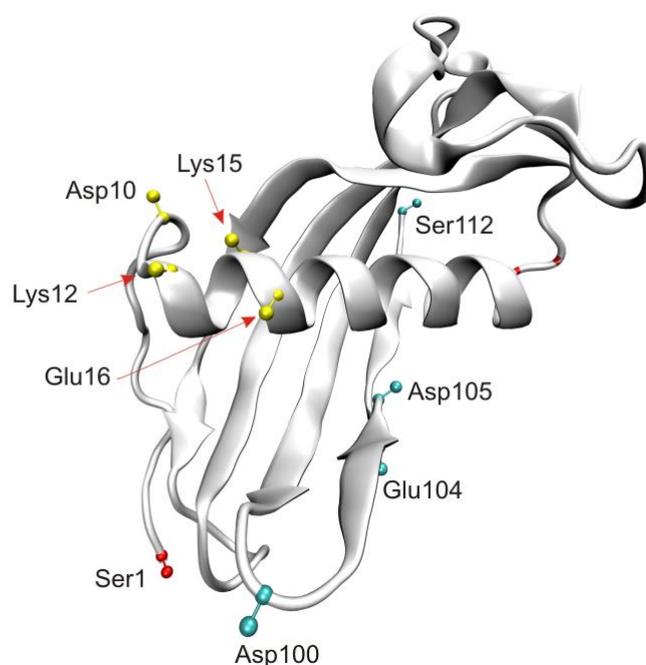
Para iniciarmos as simulações de *docking* entre o primeiro domínio cistatina de BJ46a e jararagina, buscamos nos dados de XL-MS e HDX-MS informações para guiar o *docking* e a interpretação dos modelos gerados. Derivamos dos dados de XL-MS um grupo de 23 restrições de distância, detalhadas na tabela abaixo:

**Tabela 23: Restrições de distância inter-domínio derivadas de *cross-links* para o *docking* molecular de C1-BJ46a em jararagina.**

| Link | Resíduo 1 | Prot. | Número resíduo | Resíduo 2 | Prot. | Domínio      | Número resíduo | Dist (Å) | Crosslinker     |
|------|-----------|-------|----------------|-----------|-------|--------------|----------------|----------|-----------------|
| 1    | SER       | BJ46a | 1              | LYS       | Jar   | Rico em Cys  | 356            | 15.800   | BS <sup>3</sup> |
| 2    | SER       | BJ46a | 1              | LYS       | Jar   | Rico em Cys  | 357            | 15.800   | BS <sup>3</sup> |
| 3    | SER       | BJ46a | 1              | ASP       | Jar   | Rico em Cys  | 364            | 7.000    | XPLex -ZL       |
| 4    | ASP       | BJ46a | 10             | LYS       | Jar   | Metalo       | 71             | 9.700    | XPLex -ZL       |
| 5    | LYS       | BJ46a | 12             | LYS       | Jar   | Metalo       | 71             | 17.800   | BS <sup>3</sup> |
| 6    | LYS       | BJ46a | 12             | LYS       | Jar   | Rico em Cys  | 352            | 17.800   | BS <sup>3</sup> |
| 7    | LYS       | BJ46a | 12             | LYS       | Jar   | Rico em Cys  | 356            | 17.800   | BS <sup>3</sup> |
| 8    | LYS       | BJ46a | 12             | LYS       | Jar   | Rico em Cys  | 357            | 17.800   | BS <sup>3</sup> |
| 9    | LYS       | BJ46a | 12             | GLU       | Jar   | Rico em Cys  | 363            | 10.500   | XPLex -ZL       |
| 10   | LYS       | BJ46a | 12             | LYS       | Jar   | Rico em Cys  | 366            | 17.800   | BS <sup>3</sup> |
| 11   | LYS       | BJ46a | 12             | LYS       | Jar   | Rico em Cys  | 384            | 17.800   | BS <sup>3</sup> |
| 12   | LYS       | BJ46a | 12             | SER       | Jar   | Rico em Cys  | 388            | 15.800   | BS <sup>3</sup> |
| 13   | LYS       | BJ46a | 15             | LYS       | Jar   | Rico em Cys  | 356            | 17.800   | BS <sup>3</sup> |
| 14   | LYS       | BJ46a | 15             | LYS       | Jar   | Rico em Cys  | 357            | 17.800   | BS <sup>3</sup> |
| 15   | LYS       | BJ46a | 15             | LYS       | Jar   | Rico em Cys  | 366            | 17.800   | BS <sup>3</sup> |
| 16   | LYS       | BJ46a | 15             | SER       | Jar   | Rico em Cys  | 388            | 15.800   | BS <sup>3</sup> |
| 17   | LYS       | BJ46a | 15             | LYS       | Jar   | Rico em Cys  | 394            | 17.800   | BS <sup>3</sup> |
| 18   | GLU       | BJ46a | 16             | LYS       | Jar   | Metalo       | 71             | 10.500   | XPLex -ZL       |
| 19   | GLU       | BJ46a | 16             | SER       | Jar   | Desintegrina | 276            | 7.700    | XPLex -ZL       |
| 20   | ASP       | BJ46a | 100            | ASP       | Jar   | Metalo       | 75             | 13.500   | XPLex -DIA      |
| 21   | GLU       | BJ46a | 104            | ASP       | Jar   | Metalo       | 75             | 14.300   | XPLex -DIA      |
| 22   | ASP       | BJ46a | 105            | ASP       | Jar   | Metalo       | 75             | 13.500   | XPLex -DIA      |
| 23   | SER       | BJ46a | 112            | LYS       | Jar   | Rico em Cys  | 384            | 15.800   | BS <sup>3</sup> |

Os *cross-links* que conectam o domínio C1 à metaloendopeptidase de jararagina estão distribuídos em três principais regiões: (I) a porção N-terminal do domínio, com três *cross-links* (13%) direcionados à região 356-364 do domínio rico

em cisteína de jararagina; (II) a porção N-terminal da  $\alpha$ -hélice central do domínio C1 para os resíduos 352-394 do domínio rico em cisteína de jararagina e (III) a última fita da folha  $\beta$  de C1, que conecta tanto o domínio metaloendopeptidase quanto o domínio rico em cisteínas. Notamos também uma ambiguidade do dado de *cross-linking*, particularmente para os resíduos K12, K15 e E16 de C1, que apontam para diferentes resíduos do domínio rico em cisteína de jararagina. Possivelmente, algumas destas restrições são conflitantes, oriundas de outras conformações proteicas, o que pode dificultar a validação simultânea destas restrições de distância nas simulações de *docking*.



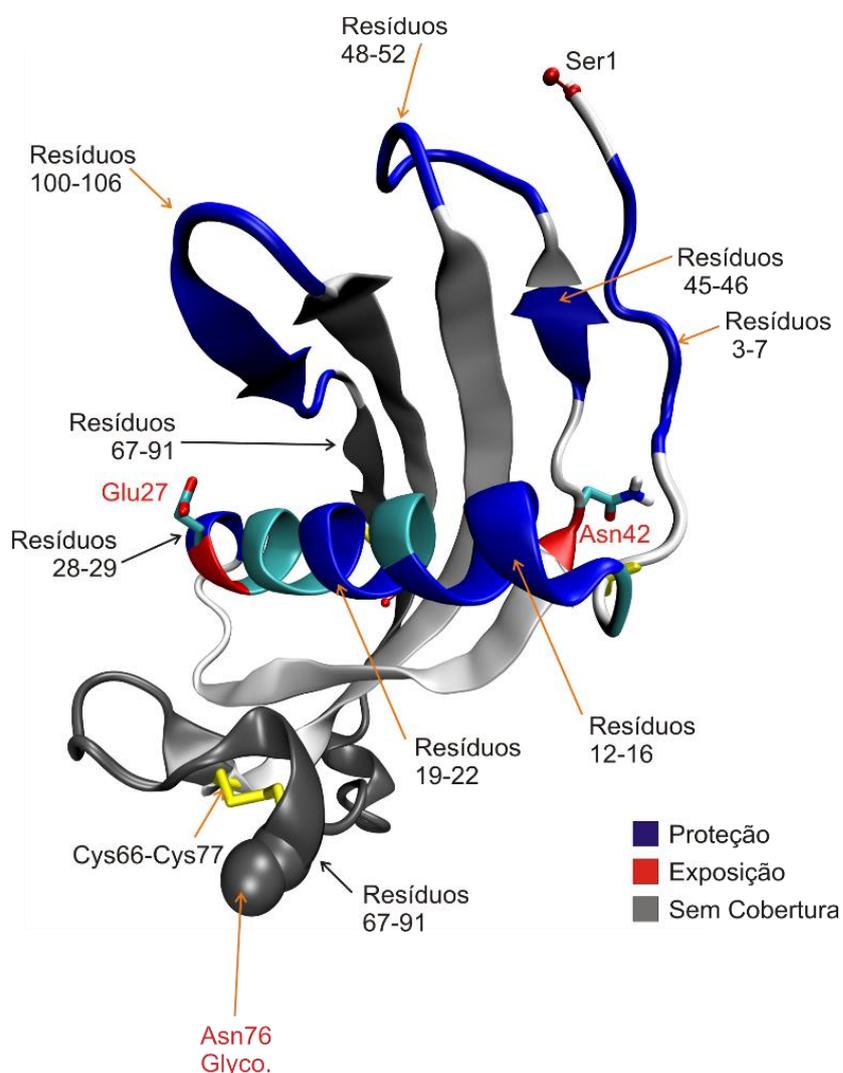
**Figura 87: Resíduos envolvidos em *cross-links* inter-proteína no domínio C1 de BJ46a.**

O domínio C1 é representado na forma *cartoon* colorido em cinza. Os resíduos envolvidos em *cross-links* são identificados pela legenda e representados por esfera e linha coloridos de acordo com a sua região de concentração estrutural (amarelo na  $\alpha$ -hélice, azul na última fita da folha  $\beta$  (resíduos 100-105) e vermelho no resíduo N-terminal do domínio).

Com relação ao perfil de incorporação isotópica de C1, as análises de HDX-MS demonstraram três diferentes regiões de proteção, que envolvem a porção N-terminal do domínio (resíduos 3 a 7), a  $\alpha$ -hélice central (resíduos 12-16, 19-22, 28-29)

e os *loops* L1 (resíduos 48-52) e L2 (resíduos 100 a 106). Há também dois resíduos que sofreram desproteção pontual: o resíduo E27, situado ao final da  $\alpha$ -hélice central do domínio, e o N42, ao final da fita  $\beta$ 1, descontinuada por perda de ângulos diedrais ideais (**Figura 88**). Como demonstrado na seção 4.4.3, as regiões vicinais ao sítio de N-glicosilação não apresentaram cobertura de sequência para a determinação de troca (resíduos 66 – 91 e 108 – 112).

É interessante notar que este perfil de incorporação isotópica é compatível tanto com estudos prévios com o inibidor HSF (discutido nas páginas 143-145) quanto para outras cistatinas canônicas. Para HSF, Aoki e colaboradores demonstraram que, após modificação química dos resíduos de lisina do complexo toxina-inibidor com ácido 2,4,6-trinitrobenzenossulfônico, um *cluster* dos resíduos K15, K41 e K103 não foram modificados, indicando que estes resíduos podem estar envolvidos com a interface de interação (Aoki *et al.* 2007). Excetuando-se o resíduo K41, para o qual não obtivemos informações de incorporação isotópica, os resíduos K15 e K103 estão em regiões protegidas quando o complexo BJ46a- jararagina é formado. Para cistatinas canônicas como a cistatina-A, que inibe endopeptidases (e.g., papaína), os *hairpin loops* L1 e L2 participam da interação e a região N-terminal estabiliza o complexo cistatina-endopeptidase, reduzindo a taxa de dissociação (Estrada *et al.* 1999).



**Figura 88: Mapa representativo da proteção/exposição à troca hidrogênio-deutério no domínio C1 de BJ46a.**

O domínio C1 foi representado na forma *cartoon* colorido em cinza. As pontes dissulfeto (C66-C77 e C91-C110) foram representadas pelos átomos C $\alpha$ , C $\beta$ , S $\gamma$ , na forma de bastões coloridos em amarelo. O sítio de N-glicosilação em N76 foi representado na forma Van der Waals para os átomos C $\alpha$  e C $\beta$ . Dois resíduos apresentaram uma desproteção (exposição) pontual: E27 e N42, representados na forma de bastões coloridos de acordo com o tipo de átomo e marcados em vermelho. As regiões que se apresentaram diferencialmente protegidas foram coloridas em azul: resíduos 3-7, 12-16, 19-22, 28-29, 45-46, 48-52, 100-106.

Na primeira simulação de *docking* C1-jararagina, adicionamos todas as restrições ilustradas na **tabela 23**. Foram gerados 136.390 modelos e todos foram analisados pelo Topolink e agrupados conforme o número de violações das restrições, como demonstra a tabela abaixo:

**Tabela 24: Análise do *docking* jararagina / C1-BJ46a, de acordo com o número de violações calculadas pelo software Topolink.**

| VIOLAÇÕES | NÚMERO DE MODELOS |
|-----------|-------------------|
| 11        | 19                |
| 12        | 3091              |
| 13        | 11168             |
| 14        | 11620             |
| 15        | 12388             |
| 16        | 9817              |
| 17        | 12465             |
| 18        | 9161              |
| 19        | 6073              |
| 20        | 5112              |
| 21        | 8376              |
| 22        | 10485             |
| 23        | 36615             |

Em um universo de 136.390 modelos, produzimos 19 modelos com menor número de violações das restrições experimentais (11 violações de um grupo de 23 restrições totais). Reagrupamos estes modelos de acordo com as restrições violadas/validadas em comum, e os resultados estão descritos na **Tabela 25**. Observamos 5 diferentes grupos de modelos, onde as diferenças estão na violação das restrições de índice 2 (BS<sup>3</sup>, S1 BJ46a – K357 jar), 10 (K12 BJ46a – 366 jar) e 13 (K15 BJ46a – 356 jar) e analisamos os modelos obtidos com relação à energia das restrições experimentais, energia da superfície de interação e energia, total, expressos em unidades de energia Rosetta (REU). Por fim, os modelos foram comparados utilizando o RMSD como medida de convergência.

O grupo 4 contém o modelo que representa a média (centro) de coordenadas do grupo e a menor dispersão com relação à média das coordenadas. Internamente, o grupo apresenta uma dispersão de RMSD de 0,320 Å até 0,800 Å com um modelo *outlier* de RMSD 3,046 Å, constituindo um grupo extremamente convergente. Dentro deste grupo, escolhemos o modelo de menor energia de restrições experimentais.

Este modelo selecionado não apresenta o melhor resultado de energia de interação de superfície; além disso, esta solução assume um posicionamento afastado do domínio C1, fato que não é corroborado pelos dados diferenciais de HDX-MS para o domínio, discutidos anteriormente. Portanto, decidimos reavaliar o grupo de restrições de distância utilizados e realizar um novo *docking*.

**Tabela 25: Análise do grupo contendo 11 violações, representados pelos modelos de menor energia de cada subgrupo.**

| CROSSLINKER     | ÍNDICE | REST.       | SUBGRUPOS     |               |               |               |               |
|-----------------|--------|-------------|---------------|---------------|---------------|---------------|---------------|
|                 |        |             | 1<br>012159_A | 2<br>031392_B | 3<br>021196_A | 4<br>029826_A | 5<br>039145_A |
| BS <sup>3</sup> | 1      | (S1-K356)   | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| BS <sup>3</sup> | 2      | (S1-K357)   | BAD:LONG      | BAD:LONG      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| ZL              | 3      | (S1-D364)   | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      |
| ZL              | 4      | (D10-K71)   | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      |
| BS <sup>3</sup> | 5      | (K12-K71)   | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      |
| BS <sup>3</sup> | 6      | (K12-K352)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| BS <sup>3</sup> | 7      | (K12-K356)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| BS <sup>3</sup> | 8      | (K12-K357)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| ZL              | 9      | (K12-E363)  | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      |
| BS <sup>3</sup> | 10     | (K12-K366)  | BAD:EUCL      | BAD:NOTFOUND  | BAD:EUCL      | BAD:EUCL      | BAD:NOTFOUND  |
| BS <sup>3</sup> | 11     | (K12-K384)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| BS <sup>3</sup> | 12     | (K12-S388)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| BS <sup>3</sup> | 13     | (K15-K356)  | OK:FOUND      | OK:FOUND      | BAD:LONG      | BAD:NOTFOUND  | BAD:NOTFOUND  |
| BS <sup>3</sup> | 14     | (K15-K357)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| BS <sup>3</sup> | 15     | (K15-K366)  | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      |
| BS <sup>3</sup> | 16     | (K15-S388)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| BS <sup>3</sup> | 17     | (K15-K394)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| ZL              | 18     | (E16-K71)   | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      |
| ZL              | 19     | (E16-S276)  | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      |
| DIA             | 20     | (D100-D75)  | BAD:EUCL      | BAD:LONG      | BAD:LONG      | BAD:LONG      | BAD:LONG      |
| DIA             | 21     | (E104-D75)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| DIA             | 22     | (D105-D75)  | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      | OK:FOUND      |
| BS <sup>3</sup> | 23     | (S112-K384) | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      | BAD:EUCL      |
|                 |        | grupo       | 1             | 2             | 3             | 4             | 5             |
|                 |        | população   | 2             | 1             | 1             | 12            | 3             |
|                 |        | RMSD        | 2,476         | 1,880         | 0,676         | 0,320         | 0,613         |

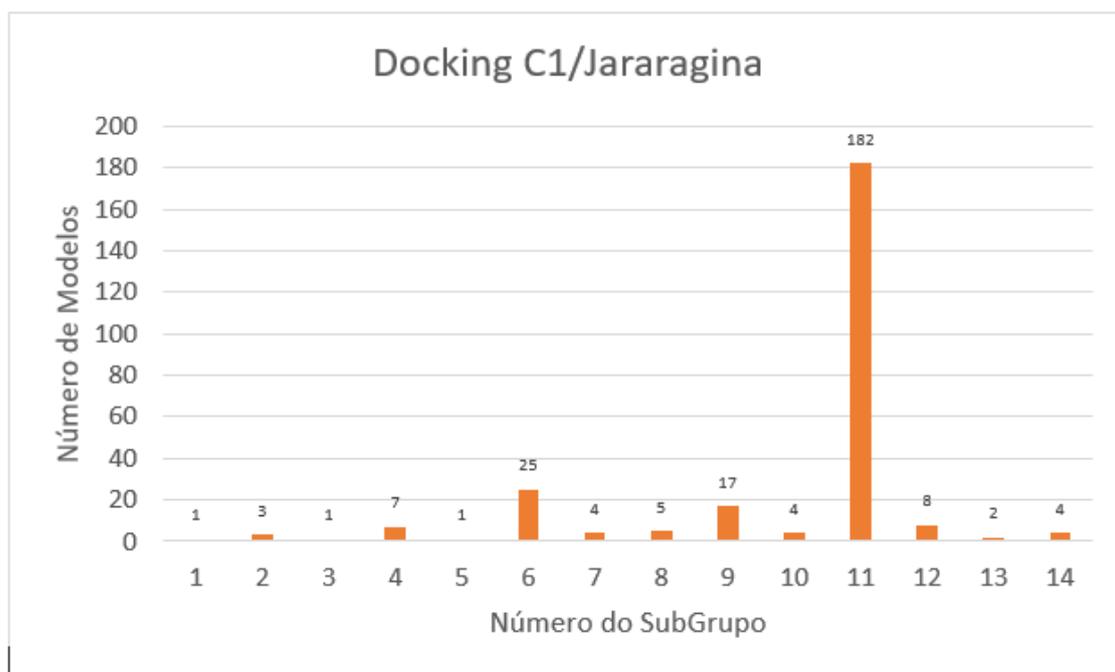
Conforme discutido anteriormente, uma parte dos *cross-links* obtidos concentraram-se nos resíduos 12, 15 e 16 do primeiro domínio cistatina de BJ46a, oriundos da estabilização do complexo com o ALC BS<sup>3</sup>, ligando diferentes resíduos em jararagina, sendo potenciais *links* conflitantes. Assim, retiramos estas restrições deste novo *docking*. Após a exclusão, restou um grupo de 12 restrições ativas. Para este novo *docking*, foram produzidos 215.822 modelos, analisados pelo software Topolink considerando-se o universo total de restrições. No universo de modelos calculados, produzimos 264 modelos com 11 violações, correspondendo a 0,12% do total, como demonstra a tabela abaixo.

**Tabela 26: Análise pelo software Topolink do novo *docking* C1-BJ46a – jararagina, excluindo restrições conflitantes.**

| VIOLAÇÕES | NÚMERO DE MODELOS |
|-----------|-------------------|
| 11        | 264               |
| 12        | 10532             |
| 13        | 24247             |
| 14        | 16334             |
| 15        | 10988             |
| 16        | 13699             |
| 17        | 15748             |
| 18        | 11136             |
| 19        | 11089             |
| 20        | 9891              |
| 21        | 15004             |
| 22        | 16758             |
| 23        | 60132             |

Embora seja difícil comparar as duas simulações de *docking* contendo conjuntos diferentes de restrições experimentais e número total de modelos calculados, podemos fazer uma extrapolação: assumindo-se um crescimento linear no cálculo, ao produzirmos 215.886 modelos, teríamos apenas 30 modelos na abordagem prévia de *docking*, que considerou todas as restrições experimentais. A retirada das restrições potencialmente conflitantes permitiu um aumento de 8,8 vezes no número de modelos, gerando 264 modelos contendo 11 violações.

Seguindo a mesma estratégia adotada anteriormente, analisamos o grupo contendo o menor número de violações com o objetivo de agrupar os modelos com as mesmas violações/validações. Foram criados 14 subgrupos, como demonstra a **Figura 89**:

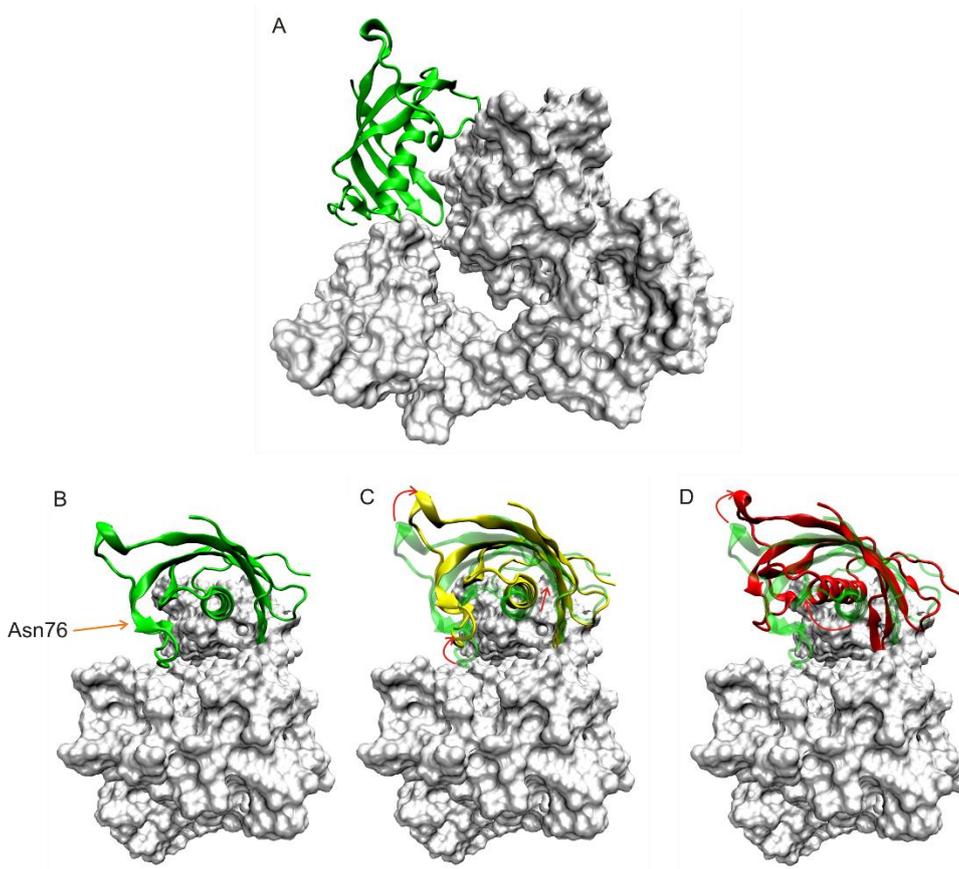


**Figura 89: Classificação do grupo de modelos com 11 violações de acordo com a validação/violação das restrições de distância.**

O grupo de modelos contendo 11 violações foi reagrupado de acordo com as restrições violadas ou validadas em comum, gerando 14 subgrupos, onde o grupo 11 é o mais populado.

Quando analisamos os 264 modelos tendo por base o RMSD do *backbone*, observamos que o modelo com menor valor de RMSD está contido no subgrupo 11, que engloba 182 modelos. Além disso, este subgrupo abriga a maior parte dos modelos de menor energia. Alinhamos as coordenadas de todos os modelos do grupo utilizando como referência o modelo de menor energia de restrições experimentais do grupo (modelo pc1\_jara\_A\_082187.pdb), e obtivemos uma dispersão de RMSD de até 3,352 Å. Neste conjunto, 59 modelos apresentaram valores de RMSD entre 0 a 1 Å, 11 modelos entre 1 a 2 Å, 108 modelos entre 2 a 3 Å e 4 modelos apresentaram RMSD entre 3 a 4 Å.

Na **Figura 90** (painel A) representamos o modelo de menor energia total e melhor resultado de energia de interação de superfície do grupo, modelo 082187. A variação de RMSD encontrada dentro do grupo pode ser visualizada utilizando os modelos que apresentaram 2 Å e 4 Å de RMSD sobrepostos (painéis C e D).



**Figura 90: Diferentes orientações do domínio C1 de BJ46a em relação à jararagina.**

O resultado do *docking* C1/jararagina foi representado por 3 diferentes modelos onde C1 interage com a jararagina na abertura do formato “C” da toxina. A jararagina foi representada pela sua superfície colorida em cinza e o domínio C1 de BJ46a foi representada na forma *cartoon*, colorido de acordo com sua orientação. (A) Visão lateral do complexo C1 – jararagina do modelo de menor energia (pc1\_jara\_082187), onde C1 foi colorido em verde. (B) Visão superior da interação onde o modelo de menor energia pousa sobre o domínio rico em cisteínas de jararagina, com sua  $\alpha$ -hélice direcionada para a abertura do formato C da enzima, o sítio de glicosilação é apontado pela seta e legenda para o resíduo Asn76. (C) A sobreposição do modelo pc1\_jara\_A\_034127 ( $1 < \text{rmsd} \leq 2$ , colorido em amarelo) com o modelo de menor energia (verde transparente). (D) A sobreposição do modelo pc1\_jara\_A\_023714 ( $3 < \text{rmsd} \leq 4$ , colorido em vermelho) com o modelo de menor energia (verde transparente). Em (C) e (D) as setas em vermelho indicam a direção do deslocamento dos modelos com relação a referência.

No modelo pc1\_jara\_082187 (**figura 90**, painéis A e B), o primeiro domínio cistatina de BJ46a interage com a jararagina direcionando a face da  $\alpha$ -hélice central do domínio para a abertura do formato “C” da enzima (representada pela superfície colorida em branco, painéis A e B).

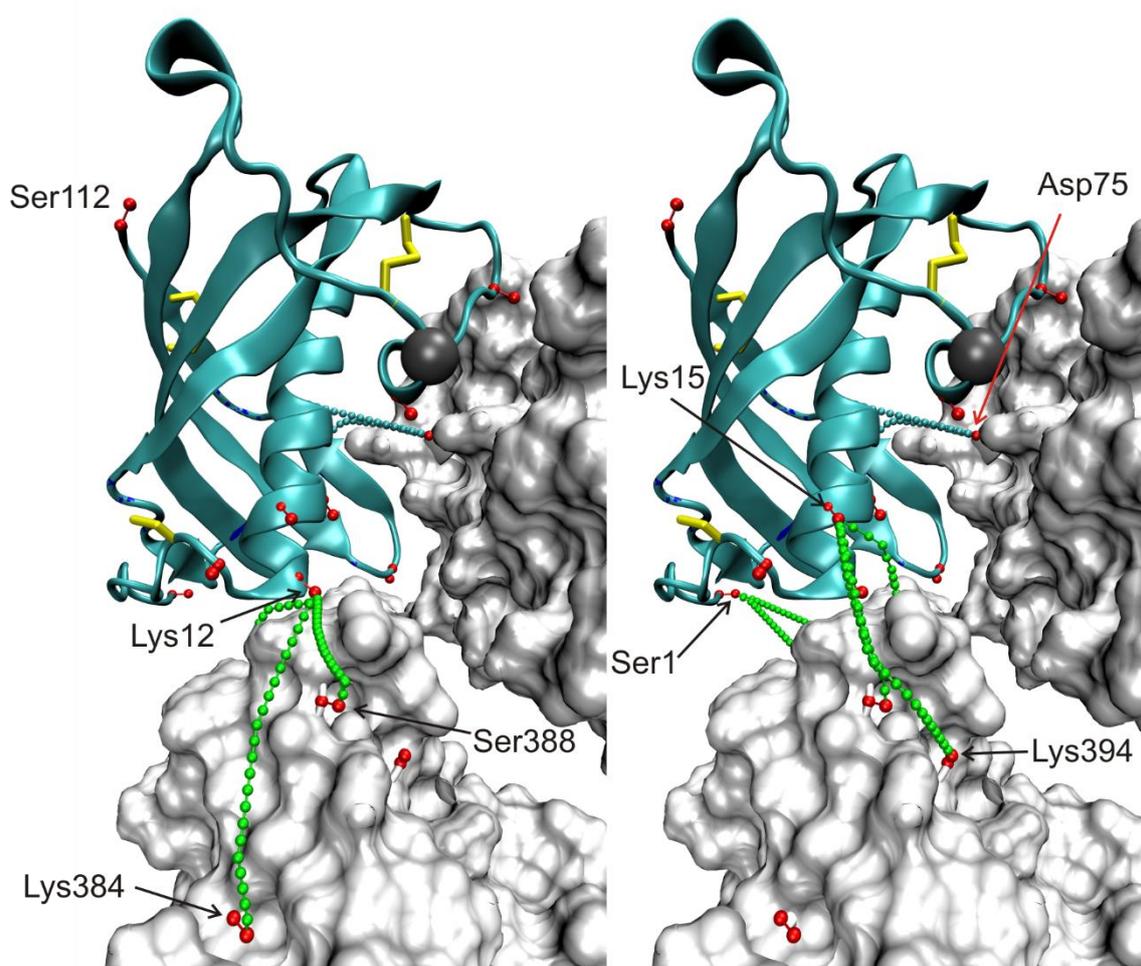
O modelo pc1\_jara\_082187 foi escolhido como representante final deste *docking* molecular. Analisamos as restrições de *cross-links* especificamente para este modelo, conforme demonstra a **tabela 27**.

**Tabela 27: Validação de *cross-links* no modelo selecionado de *docking* do domínio C1 - jararagina**

| Indices | Link | Res1 | Prot. | Res1 Num | Res2 | Prot. | Res2 Num | Dist Euclid | Dist Topo | Dist Max | Result       |
|---------|------|------|-------|----------|------|-------|----------|-------------|-----------|----------|--------------|
| 1       | BS3  | SER  | BJ46a | 1        | LYS  | JAR   | 356      | 12.321      | 12.354    | 18.000   | OK:FOUND     |
| 2       | BS3  | SER  | BJ46a | 1        | LYS  | JAR   | 357      | 15.228      | 15.451    | 18.000   | OK:FOUND     |
| 3       | ZL   | SER  | BJ46a | 1        | ASP  | JAR   | 364      | 23.709      | >23.709   | 7.000    | BAD:EUCL     |
| 4       | ZL   | ASP  | BJ46a | 10       | LYS  | JAR   | 71       | 33.513      | >33.513   | 9.700    | BAD:EUCL     |
| 5       | BS3  | LYS  | BJ46a | 12       | LYS  | JAR   | 71       | 27.577      | >27.577   | 21.800   | BAD:EUCL     |
| 6       | BS3  | LYS  | BJ46a | 12       | LYS  | JAR   | 352      | 15.174      | 15.245    | 21.800   | OK:FOUND     |
| 7       | BS3  | LYS  | BJ46a | 12       | LYS  | JAR   | 356      | 17.788      | 21.449    | 21.800   | OK:FOUND     |
| 8       | BS3  | LYS  | BJ46a | 12       | LYS  | JAR   | 357      | 14.570      | 14.811    | 21.800   | OK:FOUND     |
| 9       | ZL   | LYS  | BJ46a | 12       | GLU  | JAR   | 363      | 30.732      | >30.732   | 10.500   | BAD:EUCL     |
| 10      | BS3  | LYS  | BJ46a | 12       | LYS  | JAR   | 366      | 26.550      | >26.550   | 21.800   | BAD:EUCL     |
| 11      | BS3  | LYS  | BJ46a | 12       | LYS  | JAR   | 384      | 20.481      | 21.609    | 21.800   | OK:FOUND     |
| 12      | BS3  | LYS  | BJ46a | 12       | SER  | JAR   | 388      | 7.498       | 8.847     | 18.000   | OK:FOUND     |
| 13      | BS3  | LYS  | BJ46a | 15       | LYS  | JAR   | 356      | 20.638      | >32.700   | 21.800   | BAD:NOTFOUND |
| 14      | BS3  | LYS  | BJ46a | 15       | LYS  | JAR   | 357      | 18.302      | 20.916    | 21.800   | OK:FOUND     |
| 15      | BS3  | LYS  | BJ46a | 15       | LYS  | JAR   | 366      | 31.577      | >31.577   | 21.800   | BAD:EUCL     |
| 16      | BS3  | LYS  | BJ46a | 15       | SER  | JAR   | 388      | 12.325      | 13.034    | 18.000   | OK:FOUND     |
| 17      | BS3  | LYS  | BJ46a | 15       | LYS  | JAR   | 394      | 17.261      | 20.902    | 21.800   | OK:FOUND     |
| 18      | ZL   | GLU  | BJ46a | 16       | LYS  | JAR   | 71       | 21.919      | >21.919   | 10.500   | BAD:EUCL     |
| 19      | ZL   | GLU  | BJ46a | 16       | SER  | JAR   | 276      | 36.682      | >36.682   | 7.700    | BAD:EUCL     |
| 20      | DIA  | ASP  | BJ46a | 100      | ASP  | JAR   | 75       | 13.568      | 17.471    | 14.100   | BAD:LONG     |
| 21      | DIA  | GLU  | BJ46a | 104      | ASP  | JAR   | 75       | 10.388      | 11.457    | 15.400   | OK:FOUND     |
| 22      | DIA  | ASP  | BJ46a | 105      | ASP  | JAR   | 75       | 11.999      | 12.052    | 14.100   | OK:FOUND     |
| 23      | BS3  | SER  | BJ46a | 112      | LYS  | JAR   | 384      | 46.253      | >46.253   | 18.000   | BAD:EUCL     |

No modelo selecionado, 12 restrições totais foram validadas: 66,66% das restrições BS<sup>3</sup> e XPlex – diamina foram validadas (distâncias euclidianas e topológicas dentro dos limites), mas nenhuma restrição XPlex-zero-length foi validada. BS<sup>3</sup> apresentou 10 validações onde percebe-se 3 *hotspots* em BJ46a: 5 restrições são referentes ao resíduo 12 (50%) e 3 restrições são referentes ao resíduo 15 (30%). As duas restrições restantes envolvem o resíduo N-terminal S1. Diamina apresentou dois *cross-links* validados, referentes ao resíduo 75 de jararagina (domínio

metaloendopeptidase) para os resíduos 104 e 105 de BJ46a. Para facilitar a visualização dos dados, dividimos os dois *hotspots* em duas visualizações separadas. A **Figura 91** (painel A) apresenta os *cross-links* envolvendo o resíduo 12 de BJ46a e o painel B apresenta os *cross-links* envolvendo o resíduo 15 de BJ46a. Ambos painéis têm os *cross-links* E104-D75 e D105-D75.



**Figura 91: Cross-links validados no modelo final do docking C1-BJ46a/Jararagina.**

O domínio C1 foi representado na forma *cartoon* (colorido em ciano). As ligações de dissulfeto foram representadas pelos átomos C $\alpha$ , C $\beta$  e S $\gamma$  da cisteína na forma de bastões coloridos em amarelo. O sítio de glicosilação foi representado pelo carbono alfa do resíduo Asn76. Os resíduos envolvidos em *cross-links* foram representados pelos carbonos alfa e beta, na forma linha e esfera, coloridos em vermelho. Os *cross-links* validados foram representados na forma esfera e linha e coloridos de acordo com a estratégia envolvida: BS<sup>3</sup> colorido em verde e XPLex-diamina colorido em ciano. Separamos os *cross-links* validados em duas figuras para evitar a sobreposição excessiva. O painel A (esquerda) apresenta os *cross-links* originados no resíduo Lys12 e o painel B (direita) os *cross-links* originados no resíduo Lys15 de BJ46a.

Analizamos o modelo de *docking* gerado para a interação do domínio C1 de BJ46a com os dados diferenciais de HDX-MS para esta região. As **Figura 93 e 94** representam a superfície do complexo entre C1 e jararagina. Como mencionado anteriormente, os dados de HDX-MS diferenciais apontam para regiões de proteção localizadas no N-terminal (resíduos 3 a 7), na  $\alpha$ -hélice central do domínio (resíduos 12-16 / 19-22 / 28-29), nos *loops* L1 (48-52), L2 (100-106) e, ainda, regiões de desproteção pontuais nos resíduos E27 e N42.

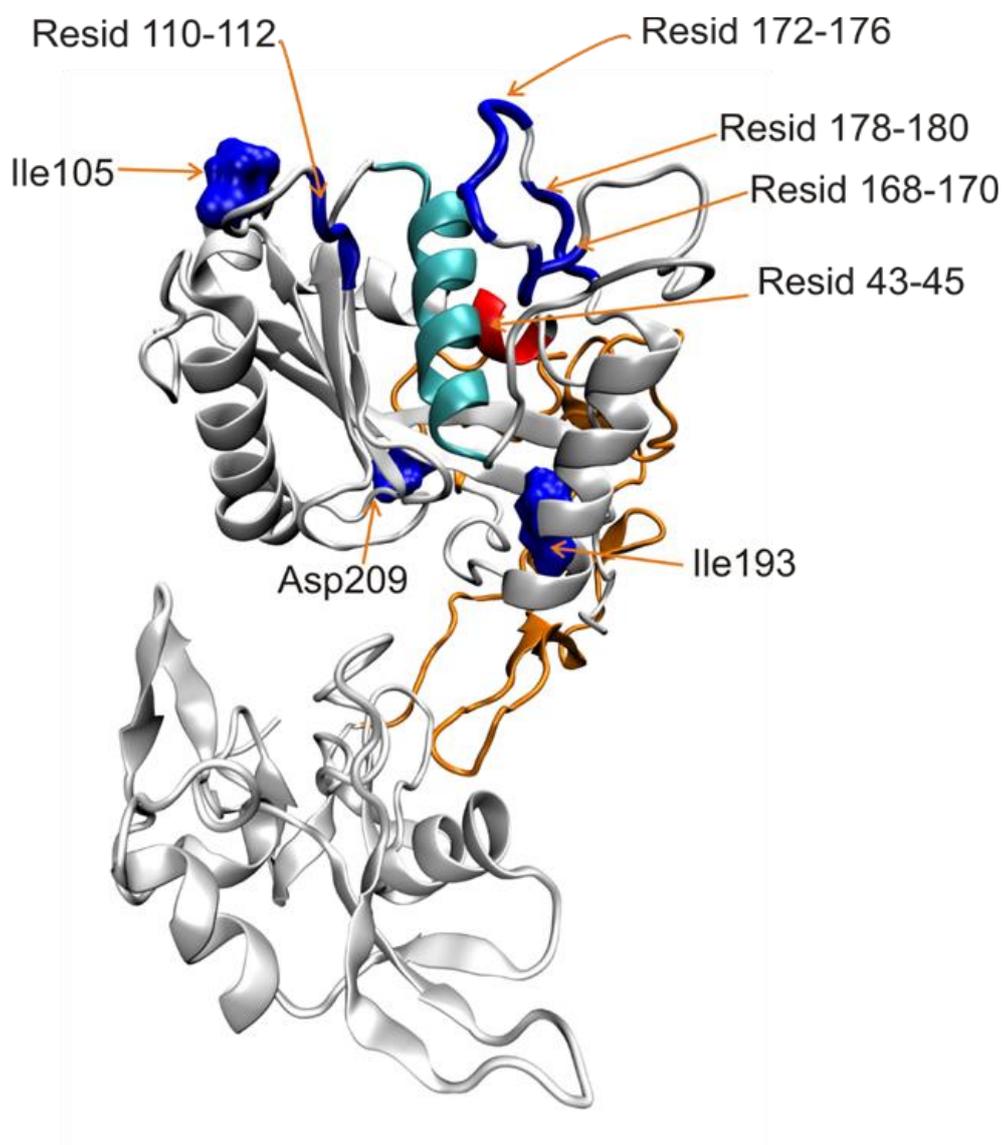
No modelo, os *loops* L1 e L2 do domínio C1 estão direcionados para a superfície de jararagina; os segmentos 12-16, 19-22 e 28-29 estão parcialmente protegidos no interior da superfície de interação. O resíduo E27 encontra-se exposto na parte superior da  $\alpha$ -hélice central, antecedendo a região protegida dos resíduos 28-29. O resíduo N42 está localizado no lado oposto, antecedendo a região protegida (**Figura 94**).

Entretanto, a região N-terminal, principalmente os resíduos 3 – 7, estão expostos no modelo, sem qualquer sinal de proteção mesmo que parcial (**Figura 94**). Como mencionado anteriormente, a porção N-terminal é extremamente importante para o estabelecimento e manutenção do complexo para cistatinas canônicas. Quando esta região é truncada, a afinidade de ligação da cistatina à papaína é 5.000 a 10.000 vezes menor do que em sua forma nativa (Nandy, 2017). É possível que o domínio cistatina 2 de BJ46a (C2) tenha um papel na orientação do N-terminal do domínio C1 quando em complexo com jararagina, impondo uma restrição adicional ao posicionamento desta região. Além disso, o método de corpos rígidos adotado pela simulação de *docking* não efetua a modificação / flexibilização de regiões não-estruturadas como *loops* e voltas.

Para jararagina, a análise diferencial dos dados de HDX-MS por mapas residuais indicou que as regiões 101-114 e 166-180 apresentaram uma proteção isotópica, isto é, incorporaram menos átomos de deutério comparado ao controle (jararagina em sua forma livre), conforme demonstrado na seção 4.4.3 deste capítulo (p.154). Já a região 43-45 incorporou mais deutério quando em complexo, uma exposição maior ao isótopo. A análise subtrativa (**Figura 70**) permitiu a sublocalização das regiões de proteção em  $^{110}\text{TIG}_{122}$  e a região

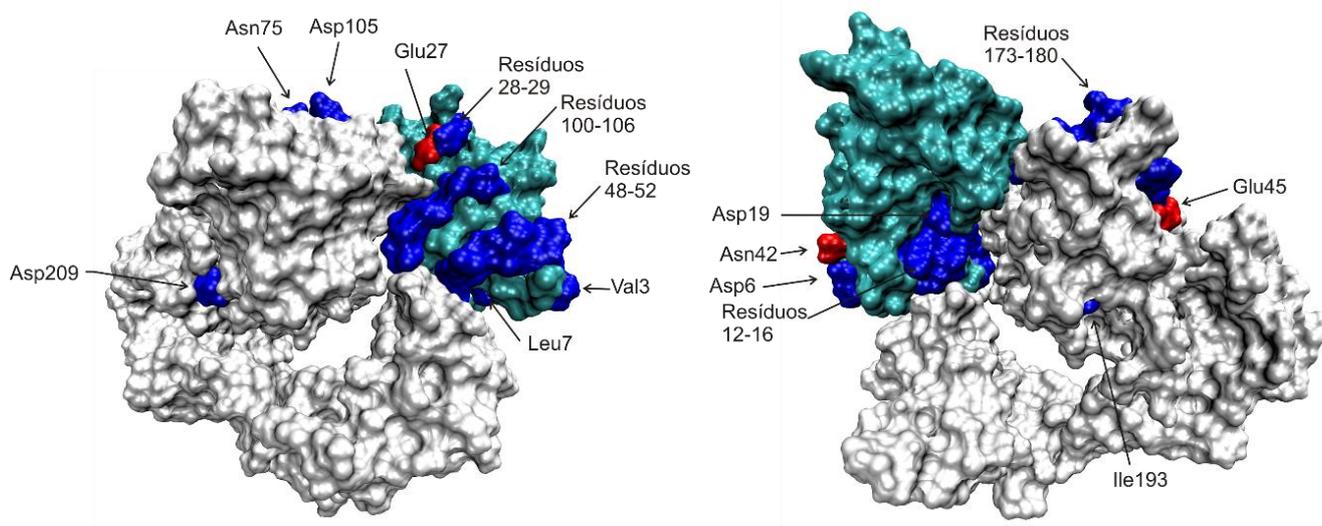
<sup>168</sup>IMG<sup>170</sup>P<sup>172</sup>TISNE<sup>176</sup>P<sup>178</sup>SKF<sup>180</sup>. A análise apontou ainda três resíduos de aminoácidos protegidos pontualmente (*single-residue resolution*): I105, I93 e D209.

A **Figura 92** ilustra as regiões diferenciais de incorporação em jararagina quando em complexo com o inibidor BJ46a. As regiões de proteção (em azul) e a região de exposição (em vermelho) margeiam a  $\alpha$ -hélice, representada em ciano, que contém os resíduos de histidina 145 e 149, parte da tríade de resíduos de histidina que coordenam o átomo de zinco, essencial para a catálise. A região 168-180 também compreende a alça *Met-turn*, parte da coordenação espacial do sítio ativo. É possível que, no momento da interação com BJ46a, ocorram rearranjos estruturais no domínio metaloendopeptidase de jararagina para acomodar o inibidor, levando à proteção ou exposição de regiões no domínio. Outra possibilidade que deve ser considerada é a formação de novas ligações de hidrogênio intra- e inter-moleculares, que afetam diretamente a incorporação isotópica nos hidrogênios amídicos do *backbone* (Kochert *et al.* 2018).



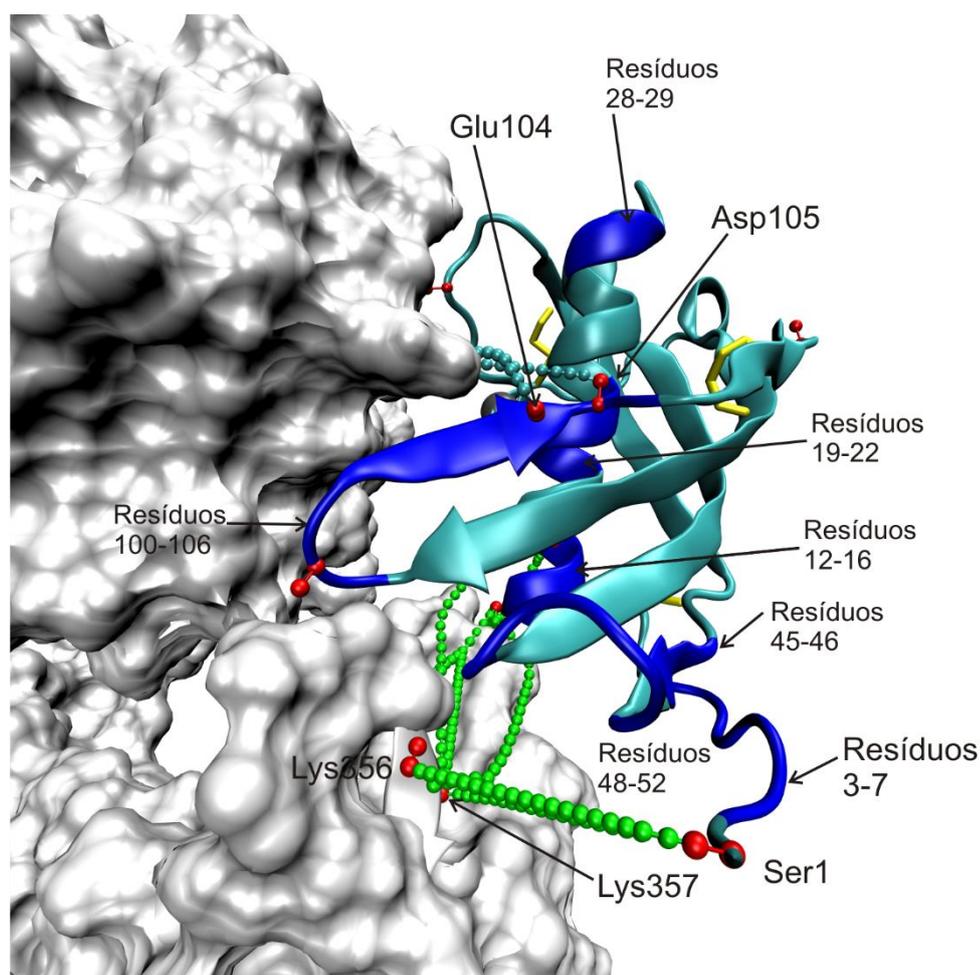
**Figura 92: Representação das regiões diferenciais de proteção/exposição a deutério no modelo estrutural de jararagina.**

A SVMP jararagina foi representada na forma *cartoon*, onde o domínio metaloendopeptidase foi colorido em cinza com a hélice que abriga o sítio ativo (resíduos 135 a 152) colorida em ciano. O domínio desintegrina-símile foi colorido em laranja e o domínio rico em cisteína foi colorido em branco. As regiões que apresentaram proteção de incorporação isotópica foram coloridas em azul. A região em vermelho representa a região que apresentou aumento da incorporação de deutério. Os resíduos para os quais a análise dos dados permitiu inferir incorporação isotópica pontual (*single-residue resolution*) foram destacados em azul.



**Figura 93: A interação entre o domínio C1 de BJ46a e jararagina.**

O domínio C1 de BJ46a foi representado pela superfície colorida em ciano e as regiões que apresentaram proteção diferencial nos dados de HDX-MS foram coloridas em azul. Regiões que sofreram exposição pontual foram coloridas em vermelho. A metaloendopeptidase jararagina foi representada pela superfície colorida em cinza, onde as regiões protegidas diferencialmente foram coloridas em azul.



**Figura 94: Interação entre o domínio C1 de BJ46a e jaragina, com incorporação dos dados de XL-MS e HDX-MS.**

A orientação da figura 93 foi alterada em 90°. O domínio C1 de BJ46a foi representado pela superfície colorida em ciano e as regiões que apresentaram proteção diferencial nos dados de HDX-MS foram coloridas em azul. Os *cross-links* validados foram representados na forma esfera e linha e coloridos de acordo com a estratégia envolvida: BS<sup>3</sup> colorido em verde, diamina colorido em ciano e *zero-length* colorido em azul escuro. As regiões que apresentaram proteção diferencial foram coloridas em azul e identificadas pelas legendas. Observamos os *loops* L1 (resíduo 48 a 52) e L2 (resíduos 100 a 106) sendo direcionados para a superfície de jaragina, na região do domínio rico em cisteína e domínio metaloendopeptidase. A  $\alpha$ -hélice central está parcialmente protegida no interior da superfície de interação.

#### 4.5.3.3 – Docking molecular entre o segundo domínio cistatina de BJ46a (C2) e a metaloendopeptidase jararagina

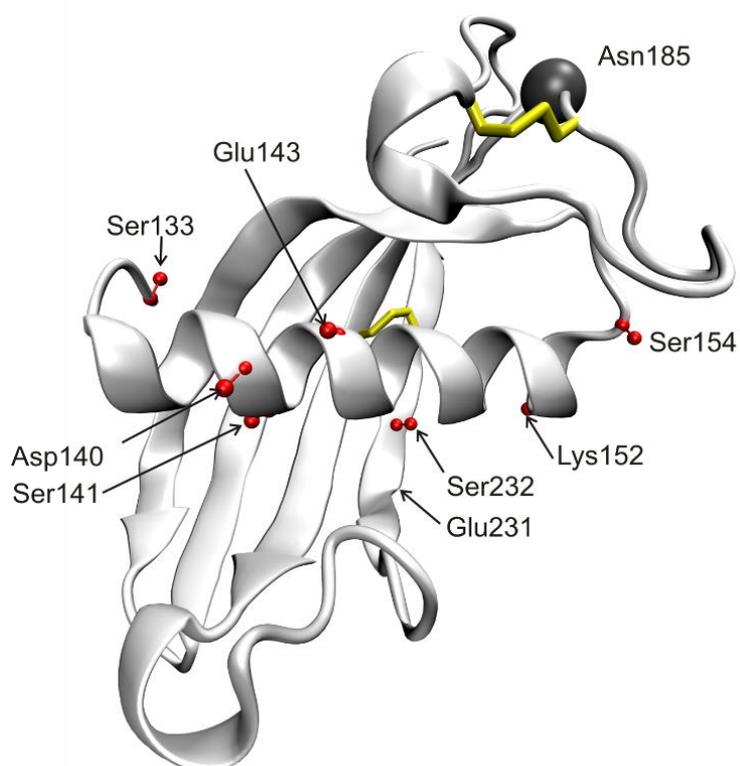
Para o *docking* do domínio C2 de BJ46a com jararagina, avaliamos inicialmente os dados de XL-MS e de HDX-MS disponíveis para este domínio. Os dados de XL-MS geraram 15 restrições espaciais inter-domínio, como demonstra a **tabela 28**:

**Tabela 28: Restrições de distância inter-domínio derivada de XL-MS para o docking entre o segundo domínio cistatina de BJ46a – jararagina.**

Legenda: C1/C2: primeiro e segundo domínio cistatina de BJ46a; (M): domínio metaloendopeptidase de jararagina; (C): domínio rico em cisteína de jararagina; ZL: *zero-length*; DIA: diamina.

| Link | Res 1 | Prot. | Res. Num | Res 2 | Prot. | Res.num | Domínio   | Dist (Å) | Crosslinker     |
|------|-------|-------|----------|-------|-------|---------|-----------|----------|-----------------|
| 1    | SER   | BJ46a | 133      | ASP   | BJ46a | 49      | C2-C1     | 7.000    | ZL              |
| 2    | SER   | BJ46a | 133      | ASP   | BJ46a | 100     | C2-C1     | 7.000    | ZL              |
| 3    | SER   | BJ46a | 133      | GLU   | Jar   | 363     | C2-Jar(C) | 7.700    | ZL              |
| 4    | ASP   | BJ46a | 140      | ASP   | BJ46a | 10      | C2-C1     | 14.100   | DIA             |
| 5    | SER   | BJ46a | 141      | GLU   | BJ46a | 11      | C2-C1     | 7.700    | ZL              |
| 6    | GLU   | BJ46a | 143      | LYS   | BJ46a | 29      | C2-C1     | 10.500   | ZL              |
| 7    | LYS   | BJ46a | 152      | GLU   | BJ46a | 16      | C2-C1     | 10.500   | ZL              |
| 8    | LYS   | BJ46a | 152      | GLU   | BJ46a | 104     | C2-C1     | 10.500   | ZL              |
| 9    | LYS   | BJ46a | 152      | LYS   | Jar   | 71      | C2-Jar(M) | 21.800   | BS <sup>3</sup> |
| 10   | LYS   | BJ46a | 152      | LYS   | Jar   | 86      | C2-Jar(M) | 21.800   | BS <sup>3</sup> |
| 11   | SER   | BJ46a | 154      | LYS   | BJ46a | 12      | C2-C1     | 18.000   | BS <sup>3</sup> |
| 12   | SER   | BJ46a | 154      | LYS   | Jar   | 71      | C2-Jar(M) | 18.000   | BS <sup>3</sup> |
| 13   | GLU   | BJ46a | 231      | LYS   | Jar   | 357     | C2-Jar(C) | 10.500   | ZL              |
| 14   | SER   | BJ46a | 232      | GLU   | BJ46a | 8       | C1-C2     | 7.700    | ZL              |
| 15   | SER   | BJ46a | 232      | GLU   | Jar   | 363     | C2-Jar(C) | 7.700    | ZL              |

Estruturalmente, existem três regiões que concentram os *cross-links* no segundo domínio cistatina de BJ46a, ilustrados na **Figura 95**. A região N-terminal do domínio (resíduo S133) apresenta duas restrições direcionadas ao domínio C1 de BJ46a e uma restrição direcionada ao domínio rico em cisteína de jararagina, concentrando 3 restrições (33% do total). Na  $\alpha$ -hélice central, os resíduos D140, S141 e E143 estão direcionados para o domínio C1 de BJ46a, enquanto que os resíduos 152 e 154 estão direcionados ao domínio C1 de BJ46a e ao domínio metaloendopeptidase de jararagina. Por fim, na última fita  $\beta$ , os resíduos 231 e 232 estão conectados aos domínios C1 de BJ46a e rico em cisteína de jararagina. O padrão de *cross-links* encontrado para o domínio C2 difere do domínio C1, no qual um conjunto de resíduos (*hotspots*) interagem com diferentes resíduos de jararagina. Em C2, há *hotspots*, mas estes interagem tanto com o domínio C1 de BJ46a e com os domínios de jararagina, e, provavelmente, não são restrições conflitantes.

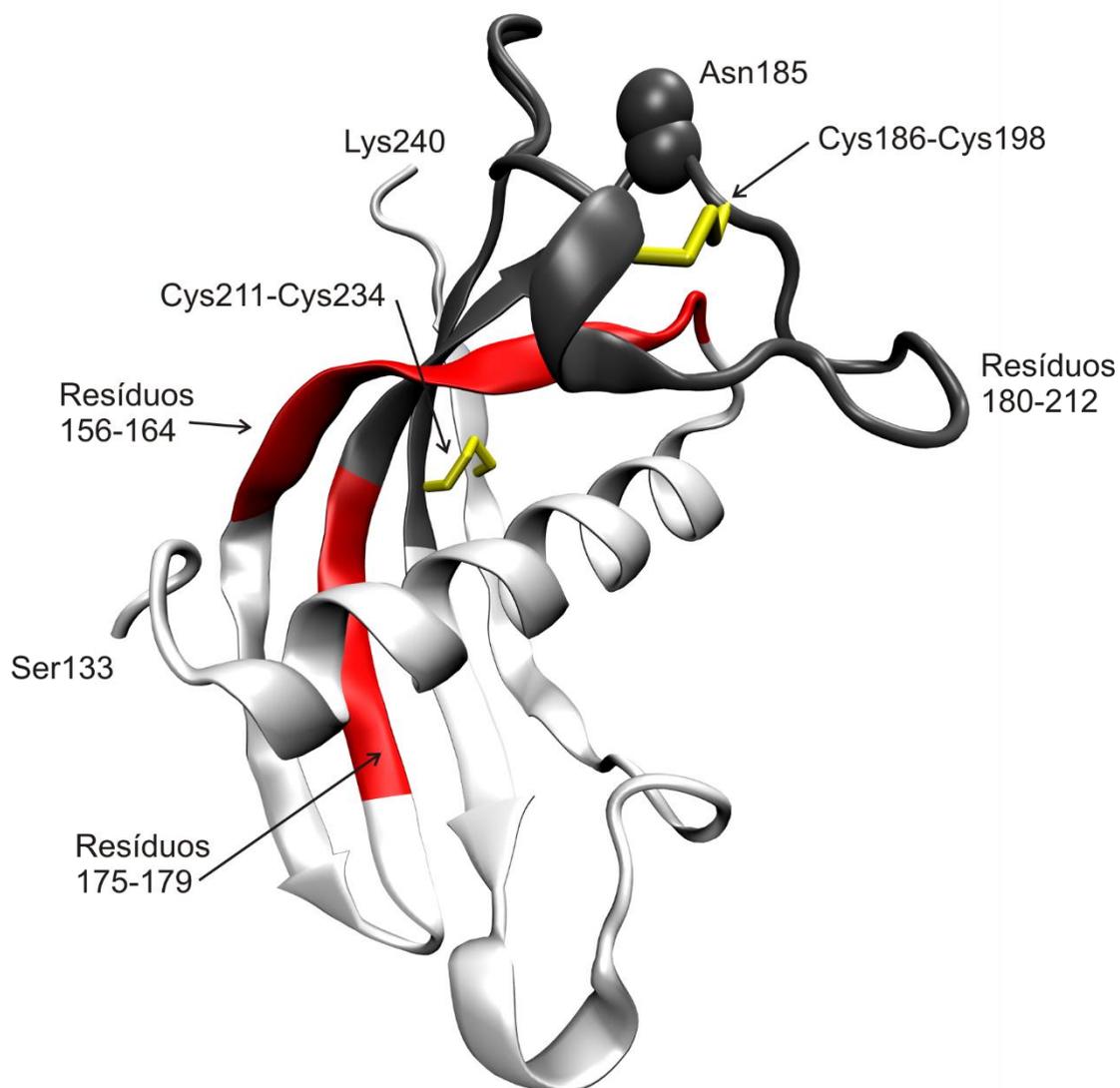


**Figura 95: Cross-links Mapeados no Domínio C2-BJ46a.**

O domínio C2 foi representado na forma *cartoon* (colorido em cinza). As ligações de dissulfeto foram representadas pelos átomos C $\alpha$ , C $\beta$  e S $\gamma$  da cisteína na forma de bastões coloridos em amarelo. O sítio de glicosilação foi representado pelo carbono alfa do resíduo Asn185. Os resíduos envolvidos em *cross-links* foram representados pelos carbonos alfa e beta, na forma linha e esfera, coloridos em vermelho. Não foram observados *cross-links* em resíduos relacionados aos *loops* L1 e L2.

Conforme discutido na subseção 4.4.3 deste capítulo, os perfis de incorporação isotópica entre os domínios cistatina de BJ46a são distintos, tanto na forma livre de BJ46a quanto em complexo com a SVMP jararagina. Para o segundo domínio cistatina de BJ46a, mapeamos duas regiões de exposição: uma compreendendo os resíduos 156 a 164, localizados na metade da primeira fita  $\beta$ , e outra compreendendo os resíduos 175 – 179, que compreende a parte central da segunda fita  $\beta$ . Os mapas residuais e a análise subtrativa também indicaram uma região de proteção para o intervalo de resíduos 228-258 e uma incorporação pontual (*single-residue resolution*) para o resíduo C234. Assim como no domínio C1, a região vicinal ao sítio de N-glicosilação (N185) não apresentou cobertura de sequência. As regiões diferenciais de incorporação isotópica foram ilustradas na **Figura 96**.

As informações acerca do papel do segundo domínio cistatina para o inibidor HSF e para cistatinas canônicas são escassas. Para HSF, nenhum fragmento anti-hemorrágico oriundo de C2 foi detectado após hidrólises sucessivas do inibidor e nenhuma proteção dos resíduos de lisina foi reportada após a trinitrofenilação dos resíduos de lisina do complexo inibidor-toxina, indicando que o domínio C2 talvez não faça parte da interface de interação (Aoki *et al.* 2007). Entretanto, verificamos nesta tese um conjunto de *cross-links* partindo do domínio C2 para o domínio C1 de BJ46a; é possível que o domínio contribua para a orientação relativa de C1 no complexo inibidor-toxina, mas outros experimentos e simulações precisam ser realizados para confirmar esta hipótese.



**Figura 96: Mapa representativo da proteção/exposição à troca hidrogênio-deutério no domínio C2 de BJ46a.**

O domínio C2 foi representado na forma *cartoon* colorido em cinza. As pontes dissulfeto (C186-C198 e C211-C234) foram representados pelos átomos C $\alpha$ , C $\beta$  e S $\gamma$ , na forma de bastões coloridos em amarelo. O sítio de N-glicosilação no resíduo N185 foi representado na forma de esferas de Van der Waals para os átomos C $\alpha$  e C $\beta$ . As regiões compreendidas entre os resíduos 156 a 164 e 175 a 176 apresentaram exposição, e foram coloridas em vermelho. A região 180 a 212 não apresentou cobertura de sequência, colorida em cinza.

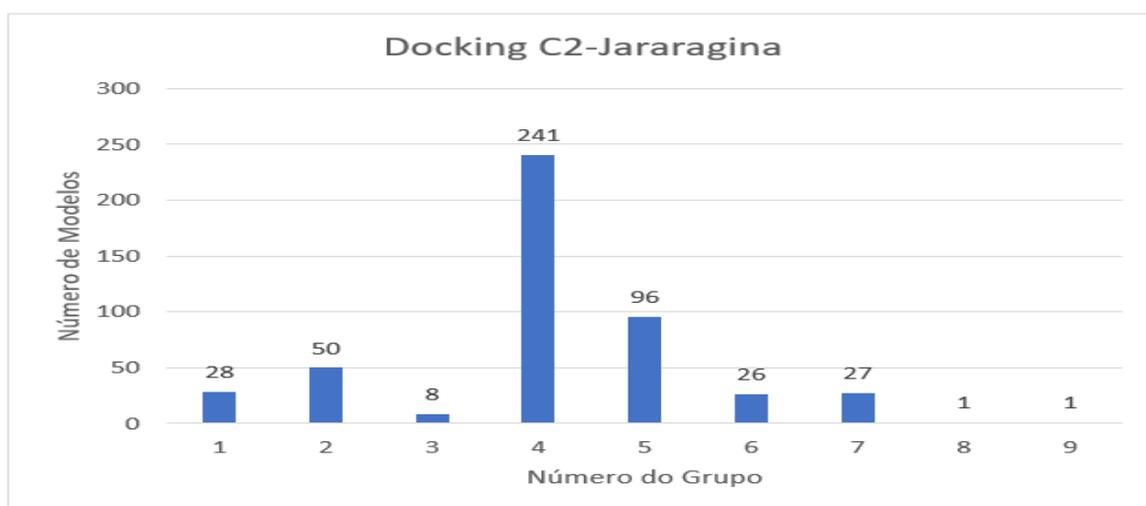
Para o *docking* C2-jararagina, foram calculados 200.082 modelos, classificados de acordo com o número de restrições experimentais violadas, analisadas pelo *software* Topolink (**tabela 29**). Obtivemos 3 modelos contendo 10 violações, o menor número no conjunto.

Tabela 29: Análise do *docking* C2 – BJ46a e jararagina pelo *software* Topolink.

| Número de Violações | Número de Modelos |
|---------------------|-------------------|
| 10                  | 3                 |
| 11                  | 478               |
| 12                  | 4012              |
| 13                  | 6311              |
| 14                  | 23890             |
| 15                  | 165388            |

Devido à baixa amostragem do grupo contendo 10 violações, analisamos também o grupo que apresentou 11 violações, contendo 478 modelos. O grupo de 10 violações apresentou validação para os *links* 8,9,10,12 e 13 (**Tabela 28**) em todos os três modelos, sendo que o modelo de menor energia de restrições de distância foi o modelo 000847.

Para o grupo contendo 11 violações, agrupamos os modelos de acordo com a validação/violação das restrições experimentais, gerando 9 diferentes subgrupos, como demonstra a **Figura 97**.



**Figura 97: Análise do Grupo contendo 11 violações de acordo com a Validação/Violação de *cross-links*.**

O grupo contendo 11 violações das restrições de *cross-links* foram agrupados gerando 9 subgrupos. O subgrupo 4 corresponde a 50,41% do total de modelos produzidos no grupo de 11 violações.

As diferenças entre os nove grupos podem ser evidenciadas pelo mapa das restrições violadas/validadas representadas pelos modelos de menor energia de restrição experimental de cada grupo. As principais restrições que causam variação neste grupo são as restrições de índice 8, 10, 11, e 13 (**Tabela 23**). Apenas um modelo apresentou variação da restrição de índice 1 e 8, como demonstra a **tabela 30**.

A partir deste mapa, alinhamos todos os modelos dos grupos de 10 e 11 violações utilizando como referência o modelo de menor energia de restrições experimentais do grupo de 10 violações (modelo c2\_c1\_jara\_B\_000847). Selecionamos então os 100 modelos de menor RMSD. A dispersão do grupo foi de 0 Å (referência) a 1,964 Å, constituindo um grupo bastante convergente.

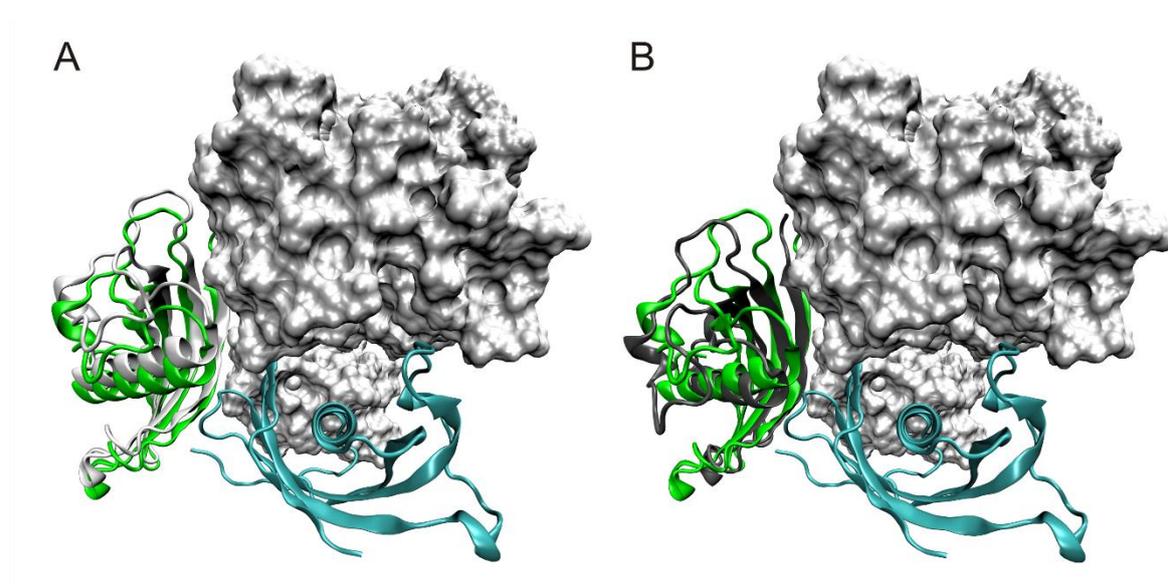
**Tabela 30: Mapa de restrições validadas dos modelos de menor energia nos grupos de 10 e 11 violações.**

| Ind | Cross | Rest        | 10viol   |          |          |          |          | 11viol   |              |          |          |          |          |
|-----|-------|-------------|----------|----------|----------|----------|----------|----------|--------------|----------|----------|----------|----------|
|     |       |             | 000847   | 021086   | 024474   | 020952   | 054103   | 048729   | 034008       | 058518   | 055928   | 004535   |          |
| 1   | ZL    | (133B-49A)  | BAD:EUCL     | BAD:EUCL | BAD:EUCL | BAD:EUCL | OK:FOUND |
| 2   | ZL    | (133B-100A) | BAD:EUCL     | BAD:EUCL | BAD:EUCL | BAD:EUCL | BAD:EUCL |
| 3   | ZL    | (133B-363D) | BAD:EUCL     | BAD:EUCL | BAD:EUCL | BAD:EUCL | BAD:EUCL |
| 4   | DIA   | (140B-10A)  | BAD:EUCL     | BAD:EUCL | BAD:EUCL | BAD:EUCL | BAD:EUCL |
| 5   | ZL    | (141B-11A)  | BAD:EUCL     | BAD:EUCL | BAD:EUCL | BAD:EUCL | BAD:EUCL |
| 6   | ZL    | (143B-29A)  | BAD:EUCL     | BAD:EUCL | BAD:EUCL | BAD:EUCL | BAD:EUCL |
| 7   | ZL    | (152B-16A)  | BAD:EUCL     | BAD:EUCL | BAD:EUCL | BAD:EUCL | BAD:EUCL |
| 8   | ZL    | (152B-104A) | OK:FOUND | BAD:EUCL | BAD:LONG | OK:FOUND | OK:FOUND | OK:FOUND | OK:FOUND     | OK:FOUND | OK:FOUND | OK:FOUND | BAD:EUCL |
| 9   | BS3   | (152B-71D)  | OK:FOUND     | OK:FOUND | OK:FOUND | OK:FOUND | OK:FOUND |
| 10  | BS3   | (152B-86D)  | OK:FOUND | OK:FOUND | OK:FOUND | BAD:LONG | OK:FOUND | OK:FOUND | OK:FOUND     | OK:FOUND | OK:FOUND | OK:FOUND | OK:FOUND |
| 11  | BS3   | (154B-12A)  | BAD:EUCL     | BAD:LONG | BAD:LONG | BAD:EUCL | BAD:EUCL |
| 12  | BS3   | (154B-71D)  | OK:FOUND     | OK:FOUND | OK:FOUND | OK:FOUND | OK:FOUND |
| 13  | ZL    | (231B-357D) | OK:FOUND | OK:FOUND | OK:FOUND | OK:FOUND | BAD:EUCL | BAD:LONG | BAD:NOTFOUND | BAD:EUCL | BAD:LONG | BAD:EUCL | BAD:EUCL |
| 14  | ZL    | (232B-8A)   | BAD:EUCL     | BAD:EUCL | BAD:EUCL | BAD:EUCL | BAD:EUCL |
| 15  | ZL    | (232B-363D) | BAD:EUCL     | BAD:EUCL | BAD:EUCL | BAD:EUCL | BAD:EUCL |
|     |       | Etotal      | -89,797  | -83,966  | -103,479 | -80,456  | -49,219  | -92,333  | -50,243      | -41,683  | -79,379  | 18,334   |          |
|     |       | E IS_C      | -0,248   | -1,245   | -3,325   | -0,819   | 0,148    | -6,899   | -2,185       | -0,701   | -3,448   | -2,068   |          |
|     |       | E Rest      | 361,081  | 382,411  | 344,527  | 438,901  | 370,294  | 364,425  | 363,974      | 376,870  | 475,981  | 543,924  |          |
|     |       | Grupo       | 10-viol  | 1,000    | 2,000    | 3,000    | 4,000    | 5,000    | 6,000        | 7,000    | 8,000    | 9,000    |          |

Na **Figura 98**, apresentamos a sobreposição dos modelos c2\_c1\_jara\_B\_000847 (referência), o modelo c2\_c1\_jara\_B\_003445 (RMSD 1,008 Å) e o modelo c2\_c1\_jara\_B\_075143 (RMSD 1,964 Å). Os modelos selecionados pela dispersão do RMSD apresentaram uma variação de energia de restrições

experimentais entre 361 a 583 REU, enquanto que a energia relativa à superfície de interação variou de -0,248 a -6,239 REU.

Ao compararmos estes modelos, observamos deslocamentos laterais do domínio C2, além de uma variação do ângulo de posicionamento relativo à jararagina. A presença de restrições de distância envolvendo resíduos de C2 com o N-terminal de C1, e com os domínios metaloendopeptidase e rico em cisteína de jararagina (**Tabela 28**) indica que o domínio C2 de BJ46a se aproxima do complexo C1-jararagina pela extremidade N-terminal de C1 e se aproxima da toxina nas vizinhanças do domínio metaloendopeptidase e rico em cisteína.

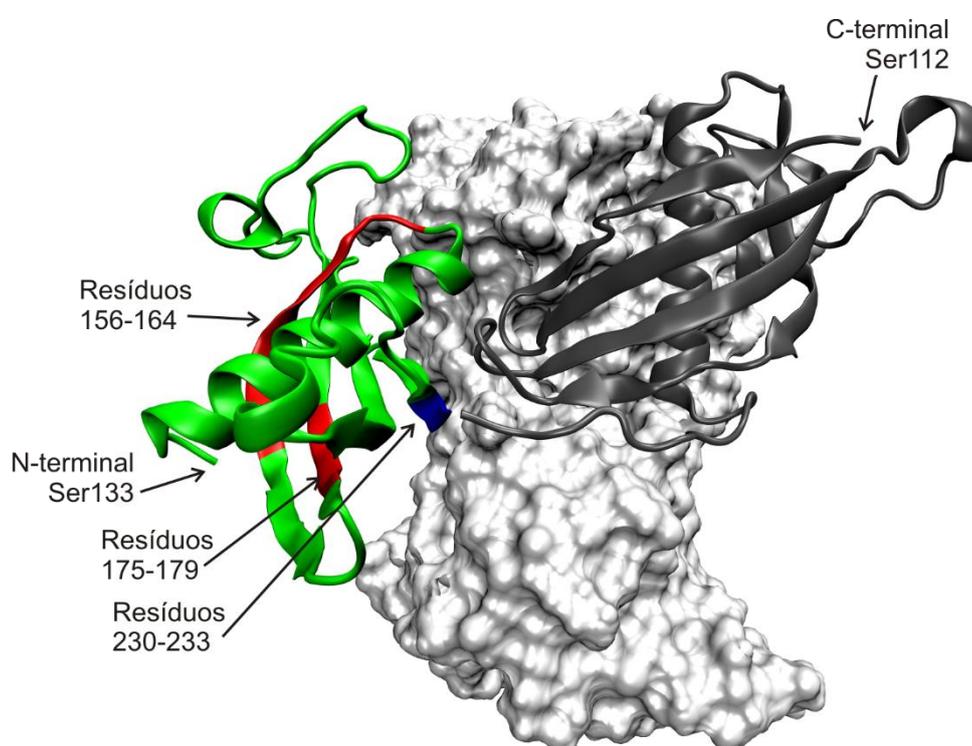


**Figura 98: Diferentes orientações de C2-BJ46a quanto ao modelo C1-Jararagina.**

A metaloendopeptidase jararagina foi representada na forma de superfície colorida em cinza. Os domínios cistatina C1 e C2 foram representados na forma *cartoon*, onde o domínio C1 foi colorido em ciano. O modelo de referência para C2-BJ46a (menor energia de restrições), modelo c2\_c1jara\_B\_000847, colorido em verde nas Figuras A e B, na mesma posição. O modelo c2\_c1jara\_B\_003445 com RMSD de 1,008 Å foi colorido em branco (Figura A) e o modelo c2\_c1jara\_B\_075143 (RMSD 1,964 Å) foi colorido em cinza.

A **Figura 99** ilustra os dados diferenciais de HDX-MS de C2 no *docking* C2-jararagina. É possível verificar nesta representação que as fitas  $\beta$  de C2 que sofreram uma maior incorporação isotópica quando em complexo com jararagina estão completamente expostas no modelo selecionado. A região tida como protegida da troca H-D está próxima à interface de interação com jararagina. Entretanto, um

aspecto que ainda precisa de ajustes é a distância entre o resíduo C-terminal de C1 (resíduo Ser112) e o resíduo N-terminal de C2 (Ser 133), calculada em 55 Å. Em uma hipotética configuração estendida, esta distância é de 60 Å. É possível que, ao adotarmos uma estratégia independente de modelagem e *docking* dos domínios de BJ46a, a informação referente à flexibilidade das porções terminais tenha sido perdida.



**Figura 99: Interação C2-BJ46a com o complexo C1-BJ46a/jararagina.**

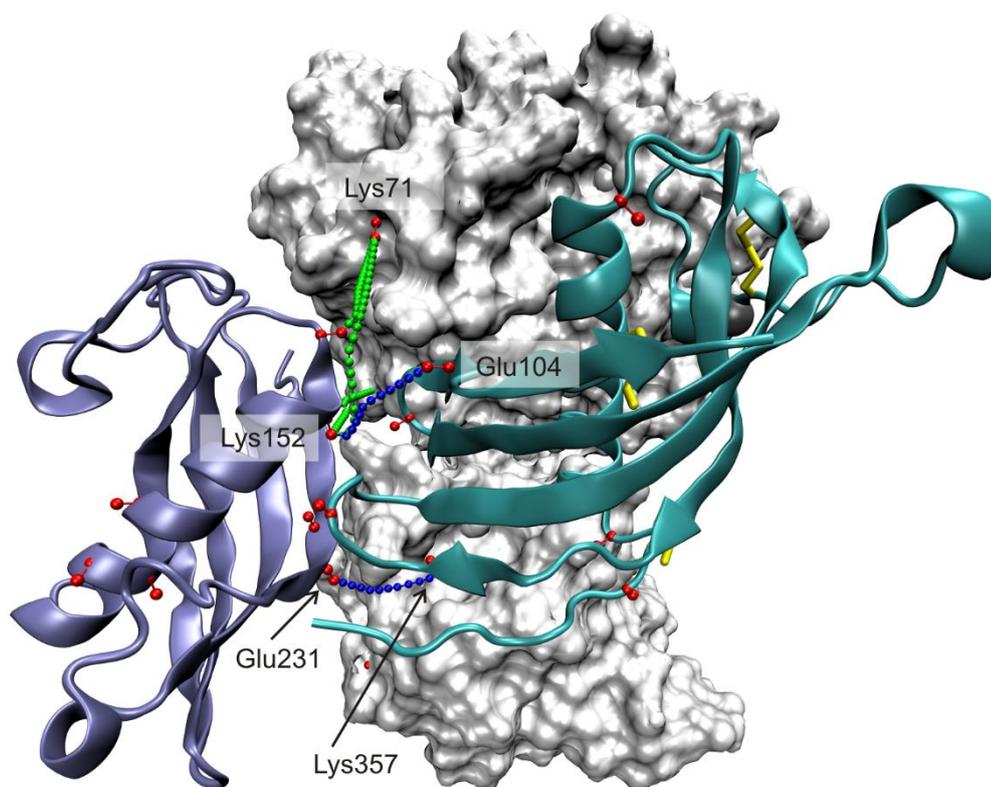
A jararagina foi representada na forma de superfície colorida em cinza. Os domínios C1 e C2 de BJ46a foram representados na forma Cartoon coloridos em cinza e verde, respectivamente. As regiões do domínio C2-BJ46a que apresentaram exposição no perfil de HDX-MS foram coloridas em vermelho, a região que apresentou proteção foi colorida em azul.

Por fim, o modelo final (*docking* C1-C2-jararagina) foi então analisado pelo *software* Topolink. Validamos 3 das 4 restrições geradas pelo ALC BS<sup>3</sup>, e 2 das 10 restrições geradas pela abordagem XPLex – *zero length*. Não foi possível validar a restrição oriunda de XPLex – diamina no modelo (**Tabela 30**). As restrições validadas encontram-se ilustradas na **Figura 100**.

Os resultados da integração de dados experimentais de XL-MS e HDX-MS para a modelagem e *docking* molecular ofereceram uma visão inédita sobre a interação do inibidor BJ46a com a metaloendopeptidase jararagina. Juntos, os dados indicam que o primeiro domínio cistatina de BJ46a interage com o domínio metaloendopeptidase de jararagina, especialmente em regiões que flanqueiam o átomo de zinco presente no domínio, essencial para a catálise enzimática. Apesar de BJ46a não ser considerada uma cistatina verdadeira (*i.e.*, que inibe cisteinoendopeptidases) devido às alterações significativas na região ativa de inibição desta classe de peptidases, os dados de HDX-MS, associados à análise estrutural, demonstraram que o primeiro domínio cistatina de BJ46a ainda retém características estruturais da interação de cistatinas verdadeiras com cisteinoendopeptidases, particularmente a importância da região N-terminal e dos *loops* L1 e L2 para a interação com a toxina (Nandy, 2017). Por outro lado, a interação do segundo domínio cistatina com a metaloendopeptidase jararagina ainda não está clara e precisa ser revisitada, com novas fontes de dados estruturais que gerem restrições adicionais para guiar uma nova simulação de *docking* molecular.

**Tabela 31: Mapa dos *cross-links* validados no modelo de *docking* C2 - jararagina**

| Indices | Link | Res1 | Cadeia | Res1 Num | Res2 | cadeia | Res2 num | Dist Euclid | Dist Topo | Dist Max | Result   |
|---------|------|------|--------|----------|------|--------|----------|-------------|-----------|----------|----------|
| 1       | ZL   | SER  | B      | 133      | ASP  | A      | 49       | 24,643      | >24,643   | 7.000    | BAD:EUCL |
| 2       | ZL   | SER  | B      | 133      | ASP  | A      | 100      | 25.907      | >25.907   | 7.000    | BAD:EUCL |
| 3       | ZL   | SER  | B      | 133      | GLU  | D      | 363      | 24.691      | >24.691   | 7.700    | BAD:EUCL |
| 4       | DIA  | ASP  | B      | 140      | ASP  | A      | 10       | 42.609      | >42.609   | 14.100   | BAD:EUCL |
| 5       | ZL   | SER  | B      | 141      | GLU  | A      | 11       | 32.620      | >32.620   | 7.700    | BAD:EUCL |
| 6       | ZL   | GLU  | B      | 143      | LYS  | A      | 29       | 40.583      | >40.583   | 10.500   | BAD:EUCL |
| 7       | ZL   | LYS  | B      | 152      | GLU  | A      | 16       | 18.359      | >18.359   | 10.500   | BAD:EUCL |
| 8       | ZL   | LYS  | B      | 152      | GLU  | A      | 104      | 9.439       | 10.182    | 10.500   | OK:FOUND |
| 9       | BS3  | LYS  | B      | 152      | LYS  | D      | 71       | 16.511      | 21.285    | 21.800   | OK:FOUND |
| 10      | BS3  | LYS  | B      | 152      | LYS  | D      | 86       | 19.286      | 20.421    | 21.800   | OK:FOUND |
| 11      | BS3  | SER  | B      | 154      | LYS  | A      | 12       | 24.435      | >24.435   | 18.000   | BAD:EUCL |
| 12      | BS3  | SER  | B      | 154      | LYS  | D      | 71       | 9.246       | 9.675     | 18.000   | OK:FOUND |
| 13      | ZL   | GLU  | B      | 231      | LYS  | D      | 357      | 9.632       | 10.464    | 10.500   | OK:FOUND |
| 14      | ZL   | SER  | B      | 232      | GLU  | A      | 8        | 25.689      | >25.689   | 7.700    | BAD:EUCL |
| 15      | ZL   | SER  | B      | 232      | GLU  | D      | 363      | 23.335      | >23.335   | 7.700    | BAD:EUCL |



**Figura 100: *Cross-links* validados no *docking* dos domínios C2-C1 de BJ46a com a metaloendopeptidase jararagina.**

A metaloendopeptidase jararagina foi representada pela sua superfície colorida em cinza. Os domínios C1 e C2 de BJ46a foram representados na forma *cartoon* coloridos em ciano e azul claro, respectivamente. Os resíduos envolvidos nos *cross-links* foram representados pelos carbonos C $\alpha$  e C $\beta$  na forma linha e esfera coloridos em vermelho. Os *cross-links* validados foram representados na forma linha e esfera coloridos de acordo com o agente de *cross-link*: BS<sup>3</sup> em verde e *zero-length* em azul.

## 5 - CONCLUSÕES

- Implementamos uma metodologia otimizada de purificação do inibidor BJ46a, com rendimento purificativo superior ao originalmente publicado;
- A aplicação da técnica de XL-MS permitiu o mapeamento da topologia do complexo BJ46a – jararagina e gerou restrições espaciais importantes para a modelagem molecular de jararagina e dos domínios cistatina de BJ46a, bem como auxiliou no posicionamento espacial dos domínios para as simulações de *docking* molecular;
- Confirmamos experimentalmente as pontes dissulfeto C9-C313 e C124-C127 em BJ46a. A ponte C9-C313, que une as extremidades N- e C-terminais do inibidor, é uma característica distintiva de fetuínas, reforçando a similaridade estrutural de BJ46a com este grupo de proteínas;
- Implementamos, no Laboratório de Toxinologia, toda a infraestrutura experimental necessária para a aquisição de dados de troca isotópica hidrogênio/deutério monitorada por espectrometria de massas;
- Os dados de HDX-MS de BJ46a e jararagina possibilitaram o mapeamento de regiões-chave de interação inibidor-toxina. Para BJ46a, a mudança mais expressiva ocorreu no N-terminal do primeiro domínio cistatina, enquanto que para jararagina, as principais alterações no perfil de incorporação isotópica ocorreram em regiões que flanqueiam o átomo de zinco catalítico, no domínio metaloendopeptidase;
- Esta tese aprofunda o conhecimento acerca do fenômeno de inibição de metaloendopeptidases de venenos de serpentes por BJ46a e tem o potencial de embasar futuras abordagens terapêuticas para os acidentes ofídicos, bem como para outras patologias associadas com a atividade e/ou expressão desregulada de metaloendopeptidases.

## 6 – REFERÊNCIAS BIBLIOGRÁFICAS

ABRAHAMSON, M. et al. Human Cystatin C - Role of the N-terminal segment in the inhibition of human cysteine proteinases and in its inactivation by leucocyte elastase. **Biochem. J.**, v. 273, p. 621-626, 1991.

AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA; FUNDAÇÃO OSWALDO CRUZ. **Farmacopeia Brasileira**. 5a Edição. . Brasília: Editora FIOCRUZ, 2010. 904.

AHDASH, Z.; PYLE, E.; POLITIS, A. Hybrid Mass Spectrometry: Towards Characterization of Protein Conformational States. **Trends in Biochemical Sciences**, v. 41, n. 8, p. 650-653, 2016.

AKASHI, S.; TAKIO, K. Characterization of the interface structure of enzyme-inhibitor complex by using hydrogen-deuterium exchange and electrospray ionisation Fourier transform ion cyclotron resonance mass spectrometry. **Protein Science**, v. 9, p. 2497-2505, 2000.

ALBER, F. et al. Integrating diverse data for structure determination of macromolecular assemblies. **Annu Rev Biochem**, v. 77, p. 443-77, 2008.

ALTSCHUL, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Res**, v. 25, n. 17, p. 3389-402.

AOKI, N. et al. Snake fetuin: isolation and structural analysis of new fetuin family proteins from the sera of venomous snakes. **Toxicon**, v. 54, n. 4, p. 481-90.

AOKI, N.; DESHIMARU, M.; TERADA, S. Active fragments of the antihemorrhagic protein HSF from serum of habu (*Trimeresurus flavoviridis*). **Toxicon**, v. 49, n. 5, p. 653-62.

AOKI, N. et al. Accelerated evolution of small serum proteins (SSPs) - The PSP94 family of proteins in a Japanese viper. **Gene**, v. 426, p. 7-14, 2008.

AOKI, N. et al. Identification of novel serum proteins in a Japanese viper: Homologs of mammalian PSP94. **Biochemical and Biophysical Research Communications**, v. 359, p. 330-334, 2007.

AOKI, N. et al. Properties and cDNA cloning of antihemorrhagic factors in sera of Chinese and Japanese mamushi (*Gloydius blomhoffi*). **Toxicon**, v. 51, n. 2, p. 251-61, Feb 2008.

ARNOLD, C. The Snakebite Fight. **Nature**, v. 537, p. 26-28, 2016.

ATKINS, P.; DE PAULA, J. **Físico-Química Biológica**. Rio de Janeiro: LTC, 2008. 511-544.

BAETTIG, O. M.; BERGHUIS, A. M. Protein Crystallization. In: NADEAU, J. L. (Ed.). **Introduction to Experimental Biophysics - Biological Methods for Physical Scientists**. Boca Raton: CRC Press, 2018. cap. 6, p.187-224.

BAGCCHI, S. Experts call for snakebite to be re-established as a neglected tropical disease. **The BMJ**, v. 351, p. h5313, 2015.

BARCHAN, D. et al. How the mongoose can fight the snake: The binding site of the mongoose acetylcholine receptor. **Proc Natl Acad Sci U S A**, v. 89, p. 7717-7721, 1992.

BASTOS, V. A. **Caracterização da Porção Glicídica do Inibidor de Metaloproteínas BJ46a**. 2014. 113 (Mestrado em Biologia Celular e Molecular). Instituto Oswaldo Cruz, Rio de Janeiro.

BASTOS, V. A. et al. Natural Inhibitors of Snake Venom Metalloendopeptidases: History and Current Challenges. **Toxins (Basel)**, v. 8, n. 9.

BATTELLINO, C. et al. Assessment of efficacy of bothropic antivenom therapy on microcirculatory effects induced by *Bothrops jararaca* snake venom **Toxicon**, v. 41, p. 583-593, 2003.

BIARDI, J. E. et al. Isolation and identification of a snake venom metalloproteinase inhibitor from California ground squirrel (*Spermophilus beecheyi*) blood sera. **Toxicon**, v. 58, n. 6-7, p. 486-93.

BOCHNER, R. **Acidentes por Animais Peçonhentos: Aspectos Históricos, Epidemiológicos, Ambientais e Sócio-Econômicos**. 2003. 145p. (Doutorado). Escola Nacional de Saúde Pública Sérgio Arouca.

BOCHNER, R.; FISZON, J. T.; MACHADO, C. A Profile of Snake Bites in Brazil, 2001 to 2012. **Journal of Clinical Toxicology**, v. 04, n. 03, 2014.

BOCHNER, R.; STRUCHINER, C. J. Epidemiologia dos Acidentes Ofídicos nos últimos 100 anos no Brasil: uma revisão. **Cad. Saúde Pública**, v. 19, n. 1, p. 7-16, 2003.

BODE, W. et al. The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. **EMBO J**, v. 7, n. 8, p. 2593-9.

BODE, W.; GOMIS-RÜTH, F. X.; STÖCKLER, W. Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the metzincins. **FEBS Lett**, v. 331, p. 134-140, 1993.

BROWN, W. M.; DZIEGIELEWSKA, K. M. Friends and relations of the cystatin superfamily--new members and their evolution. **Protein Sci**, v. 6, n. 1, p. 5-12.

BULJAN, M. et al. Alternative splicing of intrinsically disordered regions and rewiring of protein interactions. **Curr Opin Struct Biol**, v. 23, n. 3, p. 443-50, Jun 2013.

BULLOCK, J. M. A. et al. Modeling Protein Complexes Using Restraints from Crosslinking Mass Spectrometry. **Structure**, v. 26, n. 7, p. 1015-1024 e2, Jul 3 2018.

BURNS-HAMURO, L. et al. Distinct interaction modes of an AKAP bound to two regulatory subunit isoforms of protein kinase A revealed by amide hydrogen/deuterium exchange. **Protein Science**, v. 14, p. 2982-2992, 2005.

CALDEIRA COSTA, H.; BÉRNILS, R. S. Répteis do Brasil e suas Unidades Federativas: Lista de Espécies. **Herpetologia Brasileira**, v. 8, n. 1, p. 11-57, 2018.

CALMETTE, A. **Le Venin des Serpents: Physiologie de l'Envenimation, Traitement des morsures venimeuses par le sérum des animaux vaccinés** Paris: Société D'Éditions Scientifiques, 1896. 72.

CALVETE, J. J. et al. Snake venom disintegrins: evolution of structure and function. **Toxicon**, v. 45, p. 1063-1074, 2005.

CALVETE, J. J. et al. Venoms, venomics, antivenomics. **FEBS Lett**, v. 583, n. 11, p. 1736-43, Jun 5 2009.

CARVALHO, P. C. et al. Integrated analysis of shotgun proteomic data with PatternLab for Proteomics 4.0. **Nature Protocols**, v. 11, n. 1, p. 102-117, 2016.

CASEWELL, N. R. et al. Domain Loss Facilitates Accelerated Evolution and Neofunctionalization of Duplicate Snake Venom Metalloproteinase Toxin Genes. **Mol.Biol.Evol.**, v. 28, n. 9, p. 2637-2649, 2011.

CHAVEZ, J.D.; BRUCE, J.E. Chemical cross-linking with mass spectrometry: a tool for systems structural biology. **Current Opinion in Chemical Biology**, v. 48, p. 8-18, 2019.

CHEN, Z. A.; RAPPSILBER, J. Protein Dynamics in Solution by Quantitative Crosslinking/Mass Spectrometry. **Trends Biochem Sci**, v. 43, n. 11, p. 908-920, Nov 2018.

CHIPPAUX, J. P. Snakebite envenomation turns again into a neglected tropical disease! **J Venom Anim Toxins Incl Trop Dis**, v. 23, p. 38, 2017.

CHIPPAUX, J. P.; HABIB, A. G. Antivenom shortage is not circumstantial but structural. **Trans R Soc Trop Med Hyg**, v. 109, n. 12, p. 747-8.

CHIPPAUX, J. P.; WILLIAMS, V.; WHITE, J. Snake Venom Variability: Methods of Study, Results and Interpretation. **Toxicon**, v. 29, p. 1279-1303, 1991.

COMINETTI, M. R. Caracterização da atividade proteolítica de toxinas e uso de inibidores. In: SELISTRE-DE-ARAÚJO, H. S. e FERREIRA DE SOUZA, D. H. (Ed.). **Métodos em Toxinologia: toxinas de serpentes**. São Carlos: Editora UFSCar, 2007. p.45-51.

COSTA, A. R. et al. Glycosylation: impact, control and improvement during therapeutic protein production. **Crit Rev Biotechnol**, v. 34, n. 4, p. 281-99.

CUNHA, L. E. R. Soros Antiofídicos: História, Evolução e Futuro. **Journal Health NPEPS**, v. 2, n. Supl. 1, p. 1-4, 2017.

DALL, E. et al. Structure and mechanism of an aspartimide-dependent peptide ligase in human legumain. **Angew Chem Int Ed Engl**, v. 54, n. 10, p. 2917-21.

DEGIACOMI, M. T. et al. Accommodating Protein Dynamics in the Modeling of Chemical Crosslinks. **Structure**, v. 25, n. 11, p. 1751-1757 e5, Nov 7 2017.

DESHIMARU, M. et al. Properties and cDNA cloning of an antihemorrhagic factor (HSF) purified from the serum of *Trimeresurus flavoviridis*. **Toxicon**, v. 46, p. 937-945, 2005.

DESHIMARU, M. et al. Efficient Purification of an Antihemorrhagic factor in the serum of Japanese Habu (*Trimeresurus flavoviridis*). **Fukuoka Uni. Sci. Rep.**, v. 33, n. 2, p. 45-53, 2003.

DIECKMANN, T. et al. The Structures of Native Phosphorylated Chicken Cystatin and of a Recombinant Unphosphorylated Variant in Solution. **J. Mol. Biol.**, v. 234, p. 1048-1059, 1993.

DOMBKOWSKI, A. A. Disulfide by Design™: a computational method for the rational design of disulfide bonds in proteins. **Bioinformatics**, v. 19, n. 14, p. 1852-1853, 2003.

DRABECK, D. H.; DEAN, A. M.; JANSA, S. A. Why the honey badger don't care: Convergent evolution of venom-targeted nicotinic acetylcholine receptors in mammals that survive venomous snake bites. **Toxicon**, v. 99, p. 68-72.

DRAKE, A. W.; MYSZKA, D. G.; KLAKAMP, S. L. Characterizing high-affinity antigen/antibody complexes by kinetic- and equilibrium-based methods. **Anal. Biochem.**, v. 328, p. 35-43, 2004.

ENGEN, J. R.; WALES, T. E. Analytical Aspects of Hydrogen Exchange Mass Spectrometry. **Annual Review of Analytical Chemistry**, v. 8, p. 127-148, 2015.

ENGLANDER, S. W. Hydrogen Exchange and Mass Spectrometry: A Historical Perspective. **J Am Soc Mass Spectrom**, v. 17, p. 1481-1489, 2006.

ENGLANDER, S. W. et al. Hydrogen exchange: the modern legacy of Linderstrøm-Lang. **Protein Science**, v. 6, p. 1101-1109, 1997.

ENGLER, N. et al. Hydrogen atoms in proteins: position and dynamics. **PNAS**, v. 100, n. 18, p. 10243-10248, 2003.

ESPINO-SOLIS, G. P. et al. Antidotes against venomous animals: state of the art and perspectives. **J Proteomics**, v. 72, n. 2, p. 183-99.

ESTRADA, S.; RAUB-SEGALL, E.; BJÖRK, I.; OLSON, S.T. The contribution of the N-terminal region residues of cystatin a (stefin a) to the affinity and kinetics of inhibition of papain, cathepsin B, and cathepsin L. **Biochemistry**, v. 38 p. 7339-7345, 1999.

FAINI, M.; STENGEL, F.; AEBERSOLD, R. The Evolving Contribution of Mass Spectrometry to Integrative Structural Biology. **J Am Soc Mass Spectrom**, v. 27, n. 6, p. 966-74.

FELDBLUM, E. S.; ARKIN, I. T. Strength of a bifurcated H Bond. **PNAS**, v. 111, n. 11, p. 4085-4090, 2014.

FENN, J. B. et al. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. **Science**, v. 246, p. 64-71, 1989.

FERBER, M. et al. Automated structure modeling of large protein assemblies using crosslinks as distance restraints. **Nature Methods**, p. 1-9, 2016.

FERRARI, A. J. R.; GOZZO, F. C.; MARTÍNEZ, L. Statistical force-field for statistical modeling using chemical cross-linking/mass spectrometry distance constraints. **ChemRxiv, Preprint**, 2018.

FIORAMONTE, M. **Novas abordagens para ligação cruzada em proteômica estrutural**. 2016. 77 (Doutorado). Instituto de Química, Universidade Estadual de Campinas (UNICAMP), Campinas.

FIORAMONTE, M. et al. Xplex: An Effective, Multiplex Cross-Linking Chemistry for Acidic Residues. **Anal. Chem.**, v. 90, n. 10, p. 6043-6050, 2018.

FONTANA, F. **Ricerche Fisiche Sopra Il Veleno della Vipera con Alcune Osservazioni Sopra le Anguillette del Grano Sperone**. Lucca: Jacopo Giusti, 1787. 170p.

FOX, J. W.; SERRANO, S. M. Structural considerations of the snake venom metalloproteinases, key members of the M12 reprotolysin family of metalloproteinases. **Toxicon**, v. 45, n. 8, p. 969-85, Jun 15 2005.

\_\_\_\_\_. Snake Venom Metalloproteinases. In: MACKESSY, S. P. (Ed.). **Handbook of Venoms and Toxins of Reptiles**. Boca Raton: CRC Press, 2010. p.96-109.

FRANÇA, F. O. S.; FAN, H. W. Acidente Botrópico. In: SCHWARTZMAN, S. (Ed.). **Plantas Venenosas e Animais Peçonhentos**. São Paulo: Editora Sarvier, 1992. cap. 3, p.149-160.

FRY, B. G. Snakebite: When the Human Touch Becomes a Bad Touch. **Toxins** v. 10, n. 4, 2018.

FUNDAÇÃO NACIONAL DE SAÚDE. **Manual de Diagnóstico e Tratamento de Acidentes com Animais Peçonhentos**. Brasília: 120 p. 2001.

GALLAGHER, E. S.; HUDGENS, J. W. Mapping Protein-Ligand Interactions with Proteolytic Fragmentation, Hydrogen/Deuterium Exchange-Mass Spectrometry. **Methods Enzymol**, v. 566, p. 357-404, 2016.

GARRISON, F. H. Felice Fontana: A Forgotten Physiologist of the Trentino. **Bulletin of the New York Academy of Medicine**, v. 11, n. 3, p. 117-122, 1935.

GIEBELER, N.; ZIGRINO, P. A Disintegrin and Metalloprotease (ADAM): Historical Overview of Their Functions. **Toxins**, v. 8, p. 122, 2016.

GIRISH, K. S.; KEMPARAJU, K. Overlooked Issues of Snakebite Management: Time for Strategic Approach. **Current Topics in Medicinal Chemistry**, v. 11, p. 2494-2508, 2011.

GLOCKER, M. O.; ARBOGAST, B.; DEINZER, M. L. Characterization of Disulfide Linkages and Disulfide Bond Scrambling in Recombinant Human Macrophage Colony Stimulating Factor by Fast-Atom Bombardment Mass Spectrometry of Enzymatic Digests. **J Am Soc Mass Spectrom**, v. 6, p. 638-643, 1995.

GOMIS-RÜTH, F. X. Structural Aspects of the Metzincin clan of Metalloendopeptidases. **Molecular Biotechnology**, v. 24, p. 157-202, 2003.

GRAMS, F. et al. Activation of snake venom metalloproteinases by a cysteine-switch-like mechanism. **FEBS Lett**, v. 335, n. 1, p. 76-80, 1993.

GRAY, J. J. et al. Protein-Protein Docking with Simultaneous Optimization of Rigid-body Displacement and Side-chain Conformations. **J. Mol. Biol.**, v. 331, p. 281-299, 2003.

GREEN, D. W.; INGRAM, V. M.; PERUTZ, M. F. The Structure of Haemoglobin. IV. Sign determination by the isomorphous replacement method. **Royal Soc. Lond. A**, v. 225, p. 287-307, 1954.

GUTIERREZ, J. M. **Snakebite Envenoming: A Public Health Perspective**. Public Health - Methodology, Environmental and Systems Issues. MADDOCK, J.: InTechOpen 2012.

GUTIÉRREZ, J. M. et al. Snakebite envenoming. **Nat Rev Dis Primers**, v. 3, p. 17063, 2017.

GUTIERREZ, J. M. et al. Hemorrhage Caused by Snake Venom Metalloproteinases: A Journey of Discovery and Understanding. **Toxins (Basel)**, v. 8, n. 4, p. 93, Mar 26 2016.

GUTIÉRREZ, J. M. et al. A Comprehensive View of the Structural and Functional Alterations of Extracellular Matrix by Snake Venom Metalloproteinases (SVMPs): Novel Perspectives on the Pathophysiology of Envenoming. **Toxins** v. 8, p. 77-97, 2016.

GUTIÉRREZ, J. M. et al. Neutralization of local tissue damage induced by *Bothrops asper* (terciopelo) snake venom. **Toxicon**, v. 36, n. 11, p. 1529-1538, 1998. ISSN 00410101.

GUTIERREZ, J. M. et al. Trends in snakebite envenomation therapy: scientific, technological and public health considerations. **Curr Pharm Des**, v. 13, n. 28, p. 2935-50, 2007.

GUTIÉRREZ, J. M.; PEREAÑEZ, J. A. The Need for an Integrated Approach in Confronting Snakebite Envenoming in Latin America: The Relevance of Endogenous Scientific and Technological Research. **Revista Vitae**, v. 23, n. 2, p. 103-105, 2016.

GUTIERREZ, J. M.; RUCAVADO, A. Snake venom metalloproteinases: their role in the pathogenesis of local tissue damage. **Biochimie**, v. 82, n. 9-10, p. 841-50.

GUTIERREZ, J. M. et al. Experimental pathology of local tissue damage induced by *Bothrops asper* snake venom. **Toxicon**, v. 54, p. 958-975, 2009.

GUTIERREZ, J. M.; RUCAVADO, A.; ESCALANTE, T. Snake Venom Metalloproteinases - Biological Roles and Participation in the Pathophysiology of Envenomation. In: MACKESSY, S. P. (Ed.). **Handbook of Venoms and Toxins of Reptiles**. Boca Raton: CRC Press, 2010. p.115-138

GUTIERREZ, J. M. et al. Hemorrhage induced by snake venom metalloproteinases: biochemical and biophysical mechanisms involved in microvessel damage. **Toxicon**, v. 45, n. 8, p. 997-1011, Jun 15 2005.

GUTIERREZ, J. M.; THEAKSTON, R. D.; WARRELL, D. A. Confronting the neglected problem of snake bite envenoming: the need for a global partnership. **PLoS Med**, v. 3, n. 6, p. e150, Jun 2006.

HAMURO, Y. et al. Specificity of immobilized porcine pepsin in H/D exchange compatible conditions. **Rapid Commun Mass Spectrom**, v. 22, n. 7, p. 1041-6, Apr 2008.

HARRISON, R. A. et al. Research strategies to improve snakebite treatment: challenges and progress. **Journal of Proteomics**, v. 74, p. 1768-1780, 2011.

HARRISON, R. A. et al. Snake Envenoming: A disease of poverty. **PLoS Negl. Trop. Dis.**, v. 3, n. 12, p. e569, 2009.

HAWGOOD, B. J. Abbé Felice Fontana (1730-1805): Founder of Modern Toxinology. **Toxicon**, v. 33, n. 5, p. 591-601, 1995.

HAYES, W. K. Venom metering by juvenile prairie rattlesnakes, *Crotalus v. viridis*: effects of prey size and experience. **Anim. Behav.**, v. 50, p. 33-40, 1995.

HENZLER-WILDSMAN, K.; KERN, D. Dynamic Personalities of Proteins. **Nature**, v. 450, p. 964-972, 2017.

HEUKESHOVEN, J.; DERNICK, R. Simplified method for silver staining of protein in polyacrylamide gels and the mechanism of silver staining. **Electrophoresis**, v. 6, p. 103-112, 1985.

HJELM, H.; HJELM, K.; SJÖQUIST, J. Protein A from *Staphylococcus aureus*: its isolation by affinity chromatography and its use as an immunosorbent for isolation of immunoglobulins. **FEBS Lett**, v. 28, n. 1, p. 73-76, 1972.

HOUDE, D. J.; BERKOWITZ, S. A. Biopharmaceutical Industry's Biophysical Toolbox. In: HOUDE, D. J. e BERKOWITZ, S. A. (Ed.). **Biophysical Characterization of Proteins in Developing Biopharmaceuticals**. Amsterdam Elsevier, 2014. cap. 3, p.49-77.

IACOBUCCI, C.; SINZ, A. To Be or Not to Be? Five Guidelines to Avoid Misassignments in Cross-Linking Mass Spectrometry **Analytical Chemistry**, v. 89, n. 15, p. 7832-7835, 2017.

IGLESIAS, A. H.; SANTOS, L. F.; GOZZO, F. C. Collision-induced dissociation of Lys-Lys intramolecular crosslinked peptides. **J Am Soc Mass Spectrom**, v. 20, n. 4, p. 557-66, Apr 2009.

ILARI, A.; SAVINO, C. Protein Structure Determination by X-Ray Crystallography. In: KEITH, J. M. (Ed.). **Bioinformatics - Data, Sequence Analysis and Evolution**. Totowa: Humana Press, v.1, 2008. p.63-87.

JANEWAY JR., C. A. et al. **Reconhecimento de Antígenos por Receptores de Células B e T**. Porto Alegre: Artmed, 2007. p.105.

JANSA, S. A.; VOSS, R. S. Adaptive Evolution of the Venom-Targeted vWF Protein in Opossums that Eat Pitvipers **PLoS One**, v. 6, n. 6, p. e20997, 2011.

JENSEN, P. F.; RAND, K. D. Hydrogen Exchange: A Sensitive Window Into Protein Conformation and Dynamics. In: WEIS, D. D. (Ed.). **Hydrogen Exchange Mass Spectrometry of Proteins - Fundamentals, Methods and Applications**. Chichester: Wiley, 2016. cap. 1, p.1-17.

Jl, M. K. et al. Recombinant snake venom metalloproteinase inhibitor BJ46A inhibits invasion and metastasis of B16F10 and MHCC97H cells through reductions of matrix metalloproteinases 2 and 9 activities. **Anticancer Drugs**, v. 24, n. 5, p. 461-72.

JIA, L.-G. et al. Function of disintegrin-like/cysteine-rich domains of atrolysin A. Inhibition of platelet aggregation by recombinant protein and peptide antagonists. **J.Biol. Chem.**, v. 272, p. 13094-13102, 1997.

JONES, D. T. Protein secondary structure prediction based on position-specific scoring matrices. **J Mol Biol**, v. 292, n. 2, p. 195-202, Sep 17 1999.

KAHRAMAN, A. et al. Cross-link guided molecular modeling with ROSETTA. **PLoS One**, v. 8, n. 9, p. e73411, 2013.

KALTASHOV, I. A.; BOBST, C. E.; ABZALIMOV, R. R. Mass spectrometry-based methods to study protein architecture and dynamics. **Protein Sci**, v. 22, n. 5, p. 530-44.

KALTASHOV, I. A.; EYLES, S. J. **Mass Spectrometry in Structural Biology and Biophysics: Architecture, Dynamics and Interaction of Biomolecules**. Wiley, 2012. 277.

KAMIGUTI, A. S. Platelets as targets of snake venom metalloproteinases. **Toxicon**, v. 45, n. 1, p. 1041-1049, 2005.

KAMIGUTI, A. S. et al. Insights into the mechanism of haemorrhage caused by snake venom metalloproteinases. **Toxicon**, v. 34, n. 6, p. 627-642, 1996.

KAN, Z.-Y. et al. Protein hydrogen exchange at residue resolution by proteolytic fragmentation mass spectrometry analysis. **PNAS**, v. 110, n. 41, p. 16438-16443, 2013.

KARACA, E.; BONVIN, A. M. J. J. Advances in Integrative Modeling of Biomolecular Complexes. **Methods**, v. 59, p. 372-381, 2013.

KARAS, M.; BAHR, U.; HILLENKAMP, F. UV Laser Matrix Desorption/Ionization Mass Spectrometry of Proteins in the 100 000 Dalton range. **International Journal of Mass Spectrometry and Ion Processes**, v. 92, p. 231-242, 1989.

KASSAAR, O. et al. Crystal structure of histidine-rich glycoprotein N2 domain reveals redox activity at an interdomain disulfide bridge: implications for angiogenic regulation. **Blood**, v. 123, n. 12, p. 1948-55.

KATO, J.; SVENSSON, C. I. Role of Extracellular Damage-Associated Molecular Pattern Molecules (DAMPs) as Mediators of Persistent Pain. **Progress in Molecular Biology and Translational Science**, v. 131, p. 251-279, 2015.

KENDREW, J. C. et al. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. **Nature**, v. 181, p. 662-666, 1958.

KOCHERT, B. A. et al. Hydrogen-deuterium exchange mass spectrometry to study protein complexes. In: MARSH, J. A. (Ed.). **Protein Complex Assembly: Methods and Protocols**: Springer, v.1764, 2018. cap. 10, p.153-171.

KONERMANN, L.; PAN, J.; LIU, Y.-H. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. **Chem. Soc. Rev.**, v. 40, p. 1224-1234, 2011.

KONERMANN, L.; PAN, J.; LIU, Y. H. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. **Chem Soc Rev**, v. 40, n. 3, p. 1224-34, Mar 2011. ISSN 1460-4744 (Electronic)

0306-0012 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/21173980> >.

KONERMANN, L.; RODRIGUEZ, A.; SOWOLE, M. A. Type 1 and Type 2 Scenarios in Hydrogen Exchange. **Analyst**, v. 139, p. 6078, 2014.

KONERMANN, L.; PAN, J.; LIU, Y.-H. Hydrogen Exchange Mass Spectrometry for studying protein structure and dynamics. **Chem Soc Rev**, v. 40, p. 1224-1234, 2011.

KONERMANN, L.; VAHIDI, S.; SOWOLE, M. A. Mass Spectrometry Methods for Studying Structure and Dynamics of Biological Macromolecules. **Anal Chem**, v. 86, p. 213-232, 2014.

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, v. 227, n. 5259, p. 680-5, Aug 15 1970. ISSN 0028-0836 (Print) 0028-0836 (Linking).

LAING, G. D. et al. Inflammatory pathogenesis of snake venom metalloproteinase-induced skin necrosis **Eur. J. Immunol.**, v. 33, p. 3458-3463, 2003.

LAING, G. D.; MOURA-DA-SILVA, A. M. Jararhagin and its multiple effects on hemostasis. **Toxicon**, v. 45, n. 8, p. 987-96.

LEAVITT, S.; FREIRE, E. Direct measurement of protein binding energetics by isothermal titration calorimetry. **Current Opinion in Structural Biology**, v. 11, p. 560-566, 2001.

LEE, C. et al. Type 3 cystatins; fetuins, kininogen and histidine-rich glycoprotein. **Frontiers in Bioscience**, v. 14, p. 2911-2922, 2009.

LEITNER, A. et al. Probing Native Protein Structures by Chemical Cross-linking, Mass Spectrometry and Bioinformatics. **Molecular and Cellular Proteomics**, v. 9, p. 1634-1649, 2010.

LEMOS, J. D. C. et al. Epidemiologia dos Acidentes Ofídicos notificados pelo Centro de Assistência e Informação Toxicológica de Campina Grande (CEATOX-CG), Paraíba. **Rev. Bras. Epidemiol**, v. 12, n. 1, p. 50-59, 2009.

LEPREVOST, F. V. et al. PepExplorer: a similarity-driven tool for analyzing de novo sequencing results. **Mol Cell Proteomics**, v. 13, n. 9, p. 2480-9, Sep 2014.

LIMA, D. B. et al. SIM-XL: A powerful and user-friendly tool for peptide cross-linking analysis. **J Proteomics**, v. 129, p. 51-5.

LÖSSL, P.; WATERBEEMD, M. V. D.; HECK, A. J. The diverse and expanding role of mass spectrometry in structural and molecular biology. **The EMBO Journal**, p. 1-24, 2016.

LU, S. et al. Mapping disulfide bonds from sub-micrograms of purified proteins or micrograms of complex protein mixtures **Biophys. Rep.**, v. 4, n. 2, p. 68-81, 2018.

LU, S. et al. Mapping native disulfide bonds at a proteome scale. **Nat Methods**, v. 12, n. 4, p. 329-31.

- LUNA, M. S. et al. Activation of Bothrops jararaca snake venom gland and venom production: a proteomic approach. **J Proteomics**, v. 94, p. 460-72, Dec 6 2013. ISSN 1876-7737 (Electronic) 1874-3919 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24176787> >.
- MACKESSY, S. P. The Field of Reptile Toxinology: Snakes, Lizards and their Venoms. In: MACKESSY, S. P. (Ed.). **Handbook of Venoms and Toxins of Reptiles**. Boca Raton: CRC Press, 2010. cap. 1, p.3-23.
- MACKESSY, S. P.; BAXTER, L. M. Bioweapons synthesis and storage: The venom gland of front-fanged snakes. **Zoologischer Anzeiger - A Journal of Comparative Zoology**, v. 245, n. 3-4, p. 147-159, 2006. ISSN 00445231.
- MADLER, S. et al. Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. **J Mass Spectrom**, v. 44, n. 5, p. 694-706.
- MARCOUX, J.; CIANFERANI, S. Towards integrative structural mass spectrometry: Benefits from hybrid approaches. **Methods**, v. 89, p. 4-12, Nov 1 2015.
- MARINO, K. et al. A systematic approach to protein glycosylation analysis: a path through the maze. **Nat Chem Biol**, v. 6, n. 10, p. 713-23, Oct 2010.
- MARTÍNEZ, L.; FERRARI, A.; GOZZO, F. C. Topolink: A software to validate structural models using chemical crosslinking restraints. **Protocol Exchange** 2017.
- MASSON, G. R.; JENKINS, M. L.; BURKE, J. E. An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. **Expert Opin Drug Discov**, v. 12, n. 10, p. 981-994.
- MAYNE, L. Hydrogen Exchange Mass Spectrometry. **Methods Enzymol**, v. 566, p. 335-56, 2016. ISSN 1557-7988 (Electronic) 0076-6879 (Linking).
- MC ALLISTER, R. G.; KONERMANN, L. Challenges in the interpretation of Protein H/D Exchange data: a molecular dynamics simulation perspective. **Biochemistry**, v. 54, p. 2683-2692, 2015.
- MERK, A. et al. Breaking cryo-EM Resolution Barriers to Facilitate Drug Discovery. **Cell**, v. 165, p. 1698-1707, 2016.
- MERKLEY, E. D. et al. Distance restraints from crosslinking mass spectrometry: mining a molecular dynamics simulation database to evaluate lysine-lysine distances. **Protein Sci**, v. 23, n. 6, p. 747-59.
- MILNE, J. L. et al. Cryo-electron microscopy--a primer for the non-microscopist. **FEBS J**, v. 280, n. 1, p. 28-45, Jan 2013.

MINISTÉRIO DA SAÚDE. **Portaria nº174 - Normas Técnicas de Produção e Controle de Qualidade dos Soros Antiofídicos, Antitóxicos e Anti-rábcicos**. Brasília 1996.

\_\_\_\_\_. Acidentes por Animais Peçonhentos - Serpentes. 2018. Disponível em: < <http://portalms.saude.gov.br/saude-de-a-z/acidentes-por-animais-peconhentos-serpentes> >. Acesso em: 29/7/2018.

MISE, Y.; LIRA-DA-SILVA, R.; CARVALHO, F. Time to treatment and severity of snake envenoming in Brazil. **Revista Panamericana de Salud Pública**, p. 1-6, 2018.

MISE, Y. F. **Aspectos Epidemiológicos do Ofidismo no Nordeste Brasileiro**. 2014. 98 (Doutorado). Instituto de Saúde Coletiva, Universidade Federal da Bahia

MORAIS, V. M.; MASSALDI, H. Snake Antivenoms: Adverse Reactions and Production Technology. **Journal of Venomous Animals Including Tropical Diseases**, v. 15, n. 1, p. 2-18, 2009.

MOURA-DA-SILVA, A. M. et al. Processing of Snake Venom Metalloproteinases: Generation of Toxin Diversity and Enzyme Inactivation. **Toxins (Basel)**, v. 8, n. 183, p. 62-76, 2016.

MOURA-DA-SILVA, A. M.; BALDO, C. Jararhagin, a hemorrhagic snake venom metalloproteinase from *Bothrops jararaca*. **Toxicon**, v. 60, n. 3, p. 280-9, Sep 1 2012. ISSN 1879-3150 (Electronic) 0041-0101 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/22534074> >.

MOURA-DA-SILVA, A. M.; TANJONI, I. Importance of Snake Venom Metalloproteinases in Cell Biology: Effects on Platelets, Inflammatory and Endothelial Cells. **Current Pharmaceutical Design**, v. 13, p. 2893-2905, 2007.

MUNIZ, J. R. C. et al. The three-dimensional structure of bothropasin, the main hemorrhagic factor from *Bothrops jararaca* venom: Insights for a new classification of snake venom metalloprotease subgroups. **Toxicon**, v. 52, p. 807-816, 2008.

NANDY, S. K. A Brief Account of Structure-Function Relationship of the Traditional Cysteine Protease Inhibitor - Cystatin with a Special Focus on Human Family 1 and 2 Cystatins. In: CHAKRABORTI, S. e DHALLA, N. S. (Ed.). **Proteases in Physiology and Pathology**. Singapore: Springer, 2017. p.579-605.

NANDY, S. K.; BHUYAN, R.; SEAL, A. Modelling family 2 cystatins and their interaction with papain. **Journal of Biomolecular Structure and Dynamics**, p. 1-16, 2012.

NETTLESHIP, J. E. Structural Biology of Glycoproteins. In: PETRESCU, S. (Ed.). **Glycosylation**: InTechOpen, 2012. cap. 3, p.41-62.

NEUMANN, D. et al. Snake acetylcholine receptor: cloning of the domain containing the four extracellular cysteines of the alpha subunit. **Proc Natl Acad Sci U S A**, v. 86, n. 18, p. 7255-9, Sep 1989.

NEVES-FERREIRA, A. G. C. et al. Natural Inhibitors of Snake Venom Metalloproteinases. In: GOPALAKRISHNAKONE, P.; CRUZ, L. J., et al (Ed.). **Toxins and Drug Discovery**. Dordrecht: Springer v.1, 2014. cap. 4, p.53-80.

NEVES-FERREIRA, A. G. C. et al. Natural Inhibitors - Innate Immunity to Snake Venoms. In: MACKESSY, S. P. (Ed.). **Handbook of Venoms and Toxins of Reptiles**. Boca Raton: CRC Press, v.1, 2010. p.259-284.

NICOLAU, C. A. et al. An in-depth snake venom proteome characterization: benchmarking *Bothrops jararaca*. **Journal of Proteomics**, v. 151, p. 214-231, 2017.

NICOLETI, A. F. et al. Comparison of *Bothropoides jararaca* bites with and without envenoming treated at the Vital Brazil Hospital of the Butantan Institute, State of São Paulo. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 43, n. 6, p. 657-661, 2010.

NISNEVITCH, M.; FIRER, M. A. The solid phase in affinity chromatography: strategies for antibody attachment. **J. Biochem. Biophys. Methods**, v. 49, p. 467-480, 2001.

OGANESYAN, I.; LENTO, C.; WILSON, D. J. Contemporary hydrogen deuterium exchange mass spectrometry. **Methods**, Apr 26 2018. ISSN 1095-9130 (Electronic) 1046-2023 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/29704663> >.

OLIVEIRA, S. S. et al. Snakebites in the Brazilian Amazon: Current Knowledge and Perspectives. In: GOPALAKRISHNAKONE, P.; VOGEL, C.-W., et al (Ed.). **Clinical Toxicology in Australia, Europe and Americas**. Dordrecht: Springer Netherlands, 2018. cap. 3, p.73-99.

OMORI-SATOH, T. et al. Purification and characterization of an antihemorrhagic factor in the serum of *Trimeresurus flavoviridis*, a crotalid. **Biochimica et Biophysica Acta**, v. 285, p. 414-426, 1972.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **WHO Guidelines for the Production, Control and Regulation of Snake Venom Immunoglobulins**. Geneva: 146 p. 2016.

PAINE, M. J. I. et al. Purification, Cloning and Molecular Characterization of a High Molecular Weight Hemorrhagic Metalloprotease, Jararhagin, from *Bothrops jararaca* Venom. . **J Biol Chem**, v. 267, n. 32, p. 22869-22876, 1992.

PALACIO, T. Z. et al. Isolation and characterization of a novel metalloprotease inhibitor from *Bothrops alternatus* snake serum. **Int J Biol Macromol**, v. 98, p. 436-446, May 2017. ISSN 1879-0003 (Electronic) 1041-8130 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/28163123> >.

- PARAMELLE, D. et al. Chemical cross-linkers for protein structure studies by mass spectrometry. **Proteomics**, v. 13, n. 3-4, p. 438-56, Feb 2013. ISSN 1615-9861 (Electronic) 1615-9853 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/23255214> >.
- PARKER, M. W. Protein Structure from X-Ray Diffraction. **Journal of Biological Physics**, v. 29, p. 341-362, 2003.
- PELLECCHIA, M.; SEM, D. S.; WÜTHRICH, K. NMR In Drug Discovery. **Nature Reviews - Drug Discovery**, v. 1, p. 211-219, 2002.
- PERALES, J.; DOMONT, G. B. Are inhibitors of metalloproteinases, Phospholipases A2 and Myotoxins Members of the Innate Immune System? In: MÉNEZ, A. (Ed.). **Perspectives in Molecular Toxicology**. Chichester: John Wiley & Sons, 2002. p.435-456.
- PERALES, J. et al. Natural Inhibitors of Snake Venom Hemorrhagic Metalloproteinases. **Toxicon**, v. 45, p. 1013-1020, 2005.
- PHISALIX, C. Recherches sur l'immunité du hérisson contre le venin de vipère. **C.R.Soc.Biol.**, v. 47, p. 639-641, 1895.
- PORTES-JUNIOR, J. A. et al. Unraveling the processing and activation of snake venom metalloproteinases. **J Proteome Res**, v. 13, n. 7, p. 3338-48, Jul 3 2014. ISSN 1535-3907 (Electronic) 1535-3893 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/24914619> >.
- PRECIADO, L. M.; PEREAÑEZ, J. A. Low molecular mass natural and synthetic inhibitors of snake venom metalloproteinases. **Toxin Reviews**, v. Early Online, p. 1-8, 2017.
- PROTEIN DATA BANK. PDB Data Distribution by Experimental Method and Molecular Type. 2018. Disponível em: < <https://www.rcsb.org/stats/summary> >. Acesso em: 02/09/2018.
- RABILLOUD, T.; CHARMONT, S. Detecton of proteins on two-dimensional electrophoresis gels. In: RABILLOUD, T. (Ed.). **Proteome Research: Two-Dimensional Gel Electrophoresis and Identification Methods**. New York: Springer, 2000. p.107-126.
- RODRIGUES, J. P. G. L. M.; BONVIN, A. M. J. J. Integrative Computational Modeling of Protein Interactions. **FEBS J**, v. 281, p. 1988-2003, 2014.
- ROSE, R. J. et al. High Sensitivity Orbitrap mass analysis of intact macromolecular assemblies. **Nature Methods**, p. 1-5, 2012.
- ROSEMAN, A. M. Docking structures of domains into maps from cryo-electron microscopy using local correlation. **Acta Cryst.**, v. D56, p. 1332-1340, 2000.
- ROY, A.; KUCUKURAL, A.; ZHANG, Y. I-TASSER: a unified platform for automated protein structure and function prediction. **Nat Protoc**, v. 5, n. 4, p. 725-38, Apr 2010.

RUCAVADO, A. et al. Viperid Envenomation Wound Exudate Contributes to Increased Vascular Permeability via a DAMPs/TLR4 Mediated Pathway. **Toxins** v. 8, p. 199-212, 2016.

SAJEVIC, T.; LEONARDI, A.; KRIZAJ, I. Haemostatically active proteins in snake venoms. **Toxicon**, v. 57, n. 5, p. 627-45.

SALI, A. et al. Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. **Structure**, v. 23, p. 1156-1167, 2015.

SANCHEZ, E. E.; RODRIGUEZ-ACOSTA, A. Inhibitors of snake venoms and development of new therapeutics. **Immunopharmacol Immunotoxicol**, v. 30, n. 4, p. 647-78, 2008.

SANTOS, L. F.; IGLESIAS, A. H.; GOZZO, F. C. Fragmentation features of intermolecular cross-linked peptides using N-hydroxy-succinimide esters by MALDI- and ESI-MS/MS for use in structural proteomics. **J Mass Spectrom**, v. 46, n. 8, p. 742-50.

SARPE, V.; SCHRIEMER, D. C. Data Processing in Bottom-Up Hydrogen Exchange Mass Spectrometry. In: WEIS, D. D. (Ed.). **Hydrogen Exchange Mass Spectrometry of Proteins: Fundamentals, Methods and Applications**. Chichester: John Wiley & Sons, v.1, 2016. p.37-53.

SAZIMA, I. Natural history of the jararaca pitviper, *Bothrops jararaca*, in southeastern Brazil. In: CAMPBELL, J. A. e BRODIE JR, E. D. (Ed.). **Biology of the Pitvipers**. Tyler: Selva Publishing, 1992. p.199-216.

SCHAEFER, C.; ROST, B. Predict impact of single amino acid change upon protein structure. **BMC Genomics**, v. 13, n. Suppl 4, p. S4, 2012.

SCHILLING, B. et al. MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. **Journal of the American Society for Mass Spectrometry**, v. 14, n. 8, p. 834-850, 2003. ISSN 1044-0305 1879-1123.

SCHNEIDER, M.; BELSON, A.; RAPPSILBER, J. Protein Tertiary Structure by Crosslinking / Mass Spectrometry. **Trends in Biochemical Sciences**, v. 43, n. 3, p. 157-169, 2018.

SCHUTTELKOPF, A. W. et al. Structural basis of reduction-dependent activation of human cystatin F. **J Biol Chem**, v. 281, n. 24, p. 16570-5, Jun 16 2006. ISSN 0021-9258 (Print) 0021-9258 (Linking). Disponível em: < <http://www.ncbi.nlm.nih.gov/pubmed/16601115> >.

SERINO-SILVA, C. et al. Purification and characterization of the first gamma-phospholipase inhibitor (gammaPLI) from *Bothrops jararaca* snake serum. **PLoS One**, v. 13, n. 3, p. e0193105, 2018.

SHANTHIRABALAN, S.; CHOMILIER, J.; CARPENTIER, M. Structural effects of point mutations in proteins. **Proteins**, v. 86, p. 853-867, 2018.

SHARON, M.; SINZ, A. Studying Protein-Protein Interactions by combining Native Mass Spectrometry and Chemical Cross-Linking. In: KOOL, J. e NIESSEN, W. M. A. (Ed.). **Analyzing Biomolecular Interactions by Mass Spectrometry**. Weinheim: Wiley-VCH, 2015. p.57-79.

SHI, Y. A Glimpse of Structural Biology Through X-Ray Crystallography. **Cell**, v. 159, p. 995-1014, 2014.

SHI, Y. et al. High-level expression, purification, characterization and structural prediction of a snake venom metalloproteinase inhibitor in *Pichia pastoris*. **Protein J**, v. 31, n. 3, p. 212-21, Mar 2012.

SHIOI, N.; DESHIMARU, M.; TERADA, S. Structural analysis and characterization of new small serum proteins from the serum of a venomous snake (*Gloydius blomhoffii*). **Biosci Biotechnol Biochem**, v. 78, n. 3, p. 410-9, 2014.

SHIOI, N.; NARAZAKI, M.; TERADA, S. Novel Function of Antihemorrhagic Factor HSF as an SSP-binding protein in Habu (*Trimeresurus flavoviridis*) Serum. **Fukuoka Uni. Sci. Rep.**, v. 41, n. 2, p. 177-184, 2011.

SHIOI, N. et al. Small serum protein-1 changes the susceptibility of an apoptosis-inducing metalloproteinase HV1 to a metalloproteinase inhibitor in habu snake (*Trimeresurus flavoviridis*). **J Biochem**, v. 153, n. 1, p. 121-9, Jan 2013.

SILVA, A. S. T. et al. **Soros e Vacinas**. 2. São Paulo: Instituto Butantan, 2013. 60.

SINAN. Acidente por Animais Peçonhentos - Notificações registradas no Sistema de Informação de Agravos de Notificação - Brasil. 2018. Disponível em: <<http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sinannet/cnv/animaisbr.def>>. Acesso em: 29/7/2018.

SINGH, P.; PANCHAUD, A.; GOODLETT, D. R. Chemical Cross-linking and Mass Spectrometry as a Low-Resolution Protein Structural Determination Technique. **Anal. Chem.**, v. 82, p. 2636-2642, 2010.

SINZ, A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. **Mass Spectrom Rev**, v. 25, p. 663-682, 2006.

\_\_\_\_\_. Chemical Cross-linking and mass spectrometry for investigation of protein-protein interactions. In: DOWNARD, K. J. (Ed.). **Mass Spectrometry of Protein Interactions**. New Jersey: John Wiley & Sons, 2007. p.83-107.

\_\_\_\_\_. Divide and conquer: cleavable cross-linkers to study protein conformation and protein-protein interactions. **Anal Bioanal Chem**, v. 409, n. 1, p. 33-44, Jan 2017.

\_\_\_\_\_. Crosslinking Mass Spectrometry goes *in tissue*. **Cell Systems**, v. 6, p. 10-12, 2018.

SOARES, B. S. **Análise estrutural de um complexo miotoxina-antimiotoxina por cross-linking, espectrometria de massas e bioinformática**. 2018. 131 (Mestrado em Bioquímica). Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro.

SRIVASTAVA, V. K. Techniques in Biophysics. In: MISRA, G. (Ed.). **Introduction to Biomolecular Structure and Biophysics - Basics of Biophysics**. Singapore: Springer, 2017. p.205-243.

SWENSON, S.; MARKLAND, F. S., JR. Snake venom fibrin(ogen)olytic enzymes. **Toxicon**, v. 45, n. 8, p. 1021-39, Jun 15 2005.

TAKEDA, S. Structure-function Relationships of Modular Domains of PIII-Class Snake Venom Metalloproteinases. **Toxinology**, p. 1- 22, 2014.

\_\_\_\_\_. ADAM and ADAMTS Family Proteins and Snake Venom Metalloproteinases: A Structural Overview. **Toxins (Basel)**, v. 8, n. 5, May 17 2016.

TANAKA, K. et al. Protein and Polymer Analyses up to  $m/z$  100,000 by Laser Ionization Time-Of-Flight Mass Spectrometry. **Rapid Commun Mass Spectrom**, v. 2, n. 8, p. 151-153, 1988.

TANG, H.-Y.; SPEICHER, D. W. Determination of Disulfide-bond linkages in proteins. **Current Protocols in Protein Science**, v. Unit 11.11, p. 11.11.1-11.11.20, 2004.

TANJONI, I. et al. Jararhagin, a snake venom metalloproteinase, induces a specialized form of apoptosis (anoikis) selective to endothelial cells. **Apoptosis**, v. 10, n. 1, p. 851-861, 2005.

TEIXEIRA, C.; MOREIRA, V.; GUTIÉRREZ, J. M. Venoms. In: CAVAILLON, J.-M. e SINGER, M. (Ed.). **Inflammation - From Molecular and Cellular Mechanisms to the Clinic**. Weinheim: Wiley-VCH-Verlag, 2018. cap. 5, p.99-128.

THERMO SCIENTIFIC. Crosslinking Technical Handbook. Disponível em: < <https://tools.thermofisher.com/content/sfs/brochures/1602163-Crosslinking-Reagents-Handbook.pdf> >. Acesso em: 07/09/2018.

TOMPA, P. Intrinsically Disordered Proteins: A 10-year recap. **Trends in Biochemical Sciences**, v. 37, n. 12, p. 509-516, 2012.

TORTORA, G. J.; GRABOWSKI, S. R. **Princípios de Anatomia e Fisiologia**. Rio de Janeiro: Guanabara Koogan, 2008. 1047.

TRELLE, M. B. et al. Hydrogen/deuterium exchange mass spectrometry reveals specific changes in the local flexibility of plasminogen activator inhibitor 1 upon binding to the somatomedin B domain of vitronectin. **Biochemistry**, v. 51, n. 41, p. 8256-66, Oct 16 2012.

TSAI, P.-L.; CHEN, S.-F.; HUANG, S. Y. Mass Spectrometry-based strategies for protein disulfide bond identification. **Rev. Anal. Chem.**, v. 32, n. 4, p. 257-268, 2013.

UVERSKY, V. N. How to Predict Disorder in a Protein of Interest. In: ZHOU, Y.; KLOCZKOWSKI, A., *et al* (Ed.). **Prediction of Protein Secondary Structure**. New York: Humana Press, v.1, 2017. p.137-158.

VALENTE, R. H. *et al*. BJ46a, a snake venom metalloproteinase inhibitor. Isolation, characterization, cloning and insights into its mechanism of action. **Eur J Biochem**, v. 268, n. 10, p. 3042-52, 2001.

VALENTE, R. H. *et al*. *Bothrops jararaca* accessory venom gland is an ancillary source of toxins to the snake. **Journal of Proteomics**, v. 177, p. 137-147, 2018.

VALENTE, R. H. *et al*. Snake Venom Metalloproteinase Inhibitors - An Overview And Future Perspectives. In: LIMA, M. E. (Ed.). **Animal Toxins - State of the Art. Perspectives in Health and Biotechnology**. Belo Horizonte: Editora UFMG, 2009. p.547-558.

VAN DER LEE, R. *et al*. Classification of intrinsically disordered regions and proteins. **Chem Rev**, v. 114, n. 13, p. 6589-631, Jul 9 2014.

VESSECCHI, R. *et al*. Nomenclatura de Espectrometria de Massas em Língua Portuguesa. **Química Nova**, v. 34, n. 10, p. 1875-1887, 2011.

VITT, L. J.; CALDWELL, J. P. **Herpetology - An Introductory Biology of Amphibians and Reptiles - Third Edition**. San Diego: Academic Press, 2009. 697.

WALDEZ, F.; VOGT, R. C. Aspectos ecológicos e epidemiológicos de acidentes ofídicos em comunidades ribeirinhas do baixo rio Purus, Amazonas, Brasil. **Acta Amazonica**, v. 39, n. 3, p. 681-692, 2009.

WANG, C.; SCHUELER-FURMAN, O.; BAKER, D. Improved side-chain modeling for protein-protein docking. **Protein Sci**, v. 14, p. 1328-1339, 2005.

WANG, L. C.; KRISHNAWURTHY, S.; ANAND, G. S. Hydrogen Exchange Mass Spectrometry Experimental Design. In: WEIS, D. D. (Ed.). **Hydrogen Exchange Mass Spectrometry of Proteins: Fundamentals, Methods and Applications**. Chichester: Wiley, 2016. p.19-35.

WARD, A. B.; SALI, A.; WILSON, I. A. Integrative Structural Biology. **Science**, v. 339, p. 913-914, 2013.

WARRELL, D. A. Snake bite. **The Lancet**, v. 375, p. 77-88, 2010.

WEIS, D. D. Analysis of Disordered Proteins by Hydrogen Exchange Mass Spectrometry. In: WEIS, D. D. (Ed.). **Hydrogen Exchange Mass Spectrometry of Proteins: Fundamentals, Methods and Applications**. Chichester: John Wiley & Sons, 2016. p.295-321.

WEIS, D. D. *et al*. Identification and Characterization of EX1 kinetics in H/D Exchange Mass Spectrometry by Peak Width Analysis. **J Am Soc Mass Spectrom**, v. 17, p. 1498-1509, 2006.

WEN, F. H. et al. Influência das alterações ambientais na Epidemiologia dos Acidentes Ofídicos e na Distribuição Geográfica de Serpentes de Importância Médica nos Estados de São Paulo e Paraná, 1988-1997. **Informe Epidemiológico do SUS**, v. 11, n. 1, p. 45-47, 2002.

WEN, F. H.; MALAQUE, C. M. S. **Animals Envenomations in Brazil**. 1. São Paulo: Instituto Butantan, 2013. 27.

WHITE, J. Snake Venoms and Coagulopathy. **Toxicon**, v. 45, n. 1, p. 951-967, 2005.

WU, S.; ZHANG, Y. LOMETS: A local meta-threading-server for protein structure prediction. **Nucleic Acids Res**, v. 35, p. 3375-3382, 2007.

YAMAKAWA, Y.; OMORI-SATOH, T. Primary Structure of the Antihemorrhagic Factor in Serum of Japanese Habu: A Snake Venom Metalloproteinase Inhibitor with a Double-Headed Cystatin Domain. **J.Biochem.**, v. 112, p. 583-589, 1992.

YAN, X. et al. Hydrogen/deuterium exchange and mass spectrometric analysis of a protein containing multiple disulfide bonds: Solution structure of recombinant macrophage colony stimulating factor-beta (rhM-CSF $\beta$ ). **Protein Sci**, v. 11, p. 2113-2124, 2002.

YANG, J. et al. The I-TASSER suite: protein structure and function prediction. **Nature Methods**, vol. 12, n.1, p. 7-8, 2015.

YANG, J.; ZHANG, Y. Protein Structure and Function Prediction Using I-TASSER. **Curr. Protoc. Bioinform.** vol. 52, p. 5.8.1 – 5.8.15, 2015.

ZHANG, H.-M. et al. Simultaneous Reduction and Digestion of Proteins with Disulfide Bonds for Hydrogen/Deuterium Exchange Monitored by Mass Spectrometry. **Anal Chem**, v. 82, p. 1450-1454, 2010.

ZHANG, Y.; SKOLNICK, J. SPICKER: A clustering approach to identify near-native protein folds. **Journal of Computational Chemistry**, v. 25, p. 865-871, 2004.

ZUBAREV, R. A.; MAKAROV, A. Orbitrap Mass Spectrometer. **Analytical Chemistry**, v. 85, p. 5288-5296, 2013

