

UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”

INSTITUTO DE BIOCÊNCIAS, LETRAS E CIÊNCIAS EXATAS

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOFÍSICA MOLECULAR

CARACTERIZAÇÃO BIOFÍSICA e ESTRUTURAL da
METALOPROTEINASE NÃO-HEMORRÁGICA do VENENO de
Bothrops moojeni* e da ENDO- β -GLICANASE de *Bacillus subtilis

PATRICIA KIMI AKAO

SÃO JOSÉ DO RIO PRETO

2011

PATRICIA KIMI AKAO

**CARACTERIZAÇÃO BIOFÍSICA e ESTRUTURAL da
METALOPROTEINASE NÃO-HEMORRÁGICA do VENENO de
Bothrops moojeni e da ENDO- β -GLICANASE do *Bacillus subtilis***

Dissertação a ser apresentada para obtenção do título de Mestre em Biofísica Molecular junto ao programa de Pós-Graduação em Biofísica Molecular do Instituto Física da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Orientador: Prof. Dr. Mário Tyago Murakami
Co-Orientador: Prof. Dr. João Ruggieiro Neto

SÃO JOSÉ DO RIO PRETO

2011

Akao, Patricia Kimi.

Caracterização biofísica e estrutural da metaloproteinase não-hemorrágica do veneno de *Bothrops moojeni* e da endo- β -glicanase de *Bacillus subtilis* / Patricia Kimi Akao. - São José do Rio Preto : [s.n.], 2011.

113 f. : il. ; 30 cm.

Orientador: Mario Tyago Murakami

Co-orientador: João Ruggiero Neto

Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas

1. Biofísica molecular. 2. Enzimas. 3. Venenos - Purificação. 4. Metaloproteínas. I. Murakami, Mario Tyago. II. Ruggiero Neto, João. III. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. IV. Título.

CDU – 577.32

PATRICIA KIMI AKAO

**CARACTERIZAÇÃO BIOFÍSICA e ESTRUTURAL da
METALOPROTEINASE NÃO-HEMORRÁGICA do VENENO de
Bothrops moojeni e da ENDO- β -GLICANASE do *Bacillus subtilis***

BANCA EXAMINADORA

Presidente e Orientador: _____
Prof. Dr. Mário Tyago Murakami
LNBio – Laboratório Nacional de Biociências

2º Examinador: _____
Prof. Dr. Fernanda Canduri
USP – São Carlos

3º Examinador: _____
Prof. Dr. Marinonio Lopes Cornélio
UNESP – São José do Rio Preto

São José do Rio Preto, 08 de abril de 2011.

*Dedico este trabalho aos meus pais, **Rui e Marlei**, aos meus irmãos, **Renata e João** e à minha avó, **Kimi**, pela força e paciência.*

Agradecimentos

É sempre muito difícil citar todas aquelas pessoas que nos ajudaram a chegar onde estamos, são inúmeras pessoas que cruzam nosso caminho todos os anos, meses, semanas. Essas pessoas sempre levam um pouco de nós e deixam um pouco de si, ajudando-nos de forma direta ou indireta a construir quem somos.

Meus pais, Rui e Marlei, são peças fundamentais da minha vida, que me ensinaram seus valores, a nunca desistir dos meus objetivos, a ser justa e honesta comigo e com todos ao meu redor.

Meus avós, Keissuke, Nemetala, Marlene e principalmente Kimi, que mesmo convivendo nada ou pouco ou muito me passaram suas essências, me amando, amparando, respeitando...

Meus irmãos, Renata e João, companheiros de jornada, que seguem ao meu lado durante anos e sei que assim continuarão, aconteça o que acontecer.

Minha família, tios, tias, primos e primas, que fazem parte constantemente da minha vida, iluminando meus dias, aconselhando, ou que seja apenas ouvindo.

As amigas, Caroline, Luana, Viviane, Rafaela, Vivian, Renata, Juliana, Camila, Mariana e Maria Luciana por serem como irmãs pra mim, tanto na convivência, quanto nos momentos difíceis não me deixando desanimar nem cair, mesmo ausentes ou presentes fisicamente estão lá em algum lugar torcendo pela minha felicidade.

Aos amigos, Felipe, Marcelo, Natália, Nayara, Joice, Carla e Aline, pela atenção despendida todos esses anos, por me passarem seus conhecimentos e tornarem minha jornada menos árdua, pela paciência nos momentos ruins, e pelo sorriso nos momentos felizes.

Aos colegas de laboratório, Celisa, Camila, Daniel, Carol e Gi, que nunca mediram esforços para ajudar no que fosse preciso, auxiliando nas horas de desespero quando tudo parecia dar errado... Aos funcionários do departamento de Física e do LNBio pela paciência, atenção e solicitude.

Ao meu orientador, Mário, por aceitar dividir comigo seus conhecimentos e tornar possível a conclusão de mais essa etapa da minha vida. Obrigado pela sua paciência e comprometimento.

A Deus por me conceder a graça de poder ter vivido ao lado de pessoas tão maravilhosas, que me completaram de diversas formas. Obrigado por me dar forças para concluir mais essa etapa e ferramentas para alcançar meus objetivos!

*Bom mesmo é ir a luta com determinação,
abraçar a vida com paixão, perder com
classe e vencer com ousadia, pois o triunfo
pertence a quem se atreve...A vida é muito
para ser insignificante!!! **Charlie Chaplin***

Índice

LISTA DE FIGURAS	i
LISTA DE TABELAS	iii
LISTA DE ABREVIACÕES	iv
RESUMO	v
ABSTRACT	vi
CAPÍTULO 1 - INTRODUÇÃO	1
1.1 BmooMP α -1	2
1.1.1. Proteínas de veneno de serpentes	2
1.1.2. Metaloproteinases de veneno de serpentes (SVMPs).....	3
1.1.3. Classificação das SVMPs	7
1.1.4. Mecanismo de catálise proposto para as SVMPs	9
1.1.5. BmooMP α -I.....	10
1.2 Endo-1,4- β -glicanase	11
1.2.1. Estrutura da Celulose.....	11
1.2.2. Celulose vs Amilose	12
1.2.3. Aplicações da Celulose.....	13
1.2.4. Características macromoleculares da celulose	14
1.2.5. Glicosidases	15
1.2.5.1. Mecanismo de catálise e modo de ligação das glicosidases	16
1.2.6. β -Glicanases	17
1.2.7. Celulases E.C. 3.2.1.4 (Família 5).....	18
1.2.8 Módulos de ligação ao carboidrato (CBM)	19
1.2.9 Endo-1,4- β -glicanase de <i>Bacillus subtilis</i>	21
CAPÍTULO 2 - OBJETIVOS	22
CAPÍTULO 3 – PROCEDIMENTOS EXPERIMENTAIS	23
3.1 BmooMP α -I.....	24
3.1.1. Primeiro passo de purificação: Troca Aniônica	24
3.1.2. Concentração	24
3.1.3. Segundo passo de purificação: Exclusão molecular.....	25
3.1.4. Terceiro passo de purificação: Afinidade.....	25
3.1.5. Quarto passo de purificação: Exclusão molecular.....	26
3.1.6. Gel de poliacrilamida desnaturante	26
3.1.7. Quantificação.....	27
3.1.8. Sequenciamento da proteína utilizando espectrometria de massa por ionização de eletrospray (ESI-MS/MS).....	28
3.1.9. Espalhamento Dinâmico de Luz (<i>Dynamic Light Scattering</i> - DLS).....	29
3.1.10. Ensaio de cristalização	30
3.1.11. Coleta de dados.....	30
3.1.12. Determinação da estrutura: Substituição Molecular.....	31
3.1.13. Refinamento e análise da estrutura.....	31
3.1.14. <i>Docking</i> molecular com ligante flexível.....	31
3.2 Endo-1,4- β -glicanase	33
3.2.1. Preparação de células competentes (Hanahan, 1983).....	33
3.2.2. Transformação por eletroporação	34
3.2.3. Teste de expressão da celulase	34
3.2.4. Maxi-Expressão da celulase	35
3.2.5. Extração da proteína da fração celular (lise)	36
3.2.6. Gel de poliacrilamida desnaturante	37
3.2.7. Primeiro passo de purificação: Afinidade	37

3.2.8. Diálise.....	38
3.2.9. Segundo passo de purificação: Troca catiônica.....	38
3.2.10. Concentração	39
3.2.11. Terceiro passo de purificação: Exclusão Molecular.....	39
3.2.12. Quantificação.....	40
3.2.13. Dicroísmo circular	40
3.2.14. Cristalização	40
CAPÍTULO 4 – RESULTADOS E DISCUSSÃO	42
4.1 BmooMP α -I.....	43
4.1.1. Identificação da proteína por espectrometria de massa.....	43
4.1.2. Homogeneidade química e estrutural	43
4.1.3. Aquisição dos dados e determinação da estrutura.....	44
4.1.4. Descrição da estrutura	47
4.1.5. Sítio de ligação ao zinco.....	48
4.1.6. Sítio de ligação ao cálcio.....	49
4.1.7. Predição das interações moleculares entre a BmooMP α -I e alguns inibidores.....	50
4.1.8. Base estrutural para a atividade hemorrágica.....	54
4.1.9. ADAMTs e SVMPs compartilham características estruturais	55
4.2 Endo-1,4- β -glicanase.....	57
4.2.1. Teste de expressão	57
4.2.2. Extração e degradação da proteína	58
4.2.3. Purificação da Endo-1,4- β -glicanase.....	59
4.2.4. Rendimento da purificação da Endo-1,4- β -glicanase.....	62
4.2.5. Espalhamento Dinâmico de Luz (DLS)	63
4.2.6. Análises de dicroísmo circular e desnaturações térmicas.....	64
4.2.7. Desnaturação química por agentes caotrópicos.....	67
4.2.8. Ensaios de cristalização	69
CAPÍTULO 5 – CONCLUSÃO	71
5.1 BmooMP α -I.....	72
5.2 Endo-1,4- β -glicanase.....	73
REFERÊNCIAS BIBLIOGRÁFICAS	75
APÊNDICE.....	86

LISTA DE FIGURAS

Figura 1. Compostos encontrados nos venenos Viperídicos	2
Figura 2. Família de zinco metaloproteinases	4
Figura 3. Esquema do motivo de ligação ao zinco e o “ <i>Met-turn</i> ”	5
Figura 4 Domínios das SVMPS	8
Figura 5. Mecanismo de catálise para as SVMPS	9
Figura 6. Esquema do motivo de ligação ao zinco e o “ <i>Met-turn</i> ” das metzincinas	10
Figura 7. Esquema das unidades de D-glicose	11
Figura 8. (A)Glicose na forma linear e piranosídica(B) Comparação entre celulose e amilose ..	12
Figura 9. (A) Estrutura da celulose, as microfibrilas se combinam para formar a fibra de celulose (B) Estrutura da amilose, as ligações $\alpha(1-4)$ conferem a amilose uma estrutura helicoidal	13
Figura 10. Mecanismos enzimáticos da hidrólise da ligação glicosídica	16
Figura 11. Esquema evidenciando as extremidades não-redutoras e redutoras	17
Figura 12. Representação esquemática da amorfogênese das fibras de celulose mediadas pelo CBM da celobiohidrolase I	20
Figura 13. Esquema do espectrômetro de massa quadrupolo <i>time-of-flight</i>	29
Figura 14. Modelo do sistema utilizado no <i>docking</i> molecular	32
Figura 15. SDS-PAGE sob condições desnaturante da BmooMP α -I purificada	43
Figura 16. Fotomicrografia do cristal da BmooMP α -I	44
Figura 17. Diagrama de Ramachandran obtido na validação da estrutura da BmooMP α -I	45
Figura 18. Representação esquemática da estrutura da BmooMP α -I	47
Figura 19. Representação do sítio de ligação ao íon zinco no sítio catalítico da BmooMP α -I	48
Figura 20. Comparação da geometria de coordenação do zinco da BmooMP α -I	49
Figura 21. A) Representação do sítio de ligação ao íon cálcio da BmooMP α -I. B) Estrutura da BmooMP α -I evidenciando os locais de ligação dos íons Zn^{2+} e Ca^{2+}	50
Figura 22. Comparação dos modos de ligação experimental e <i>in silico</i>	52
Figura 23. Comparação dos modos de ligação experimental e <i>in silico</i>	53
Figura 24. Imagens estereo da representação da cadeia principal ($C\alpha$), comparando a BmooMP α - I com outras SVMPS e as ADAMTS	55
Figura 25. Teste de expressão. SDS-PAGE 13 %	57

Figura 26. A) SDS-PAGE de diferentes volumes da amostra logo após ser concentrada B) SDS-PAGE da amostra duas semanas após ser concentrada (1); Amostra duas semanas após ser concentrada com adição de benzamidina	59
Figura 27. A) Cromatograma obtido a partir da purificação da endo-1,4- β -glicanase em coluna Hitrap Chelating HP B) SDS-PAGE	60
Figura 28. A) Cromatograma obtido a partir da purificação da endo-1,4- β -glicanase em coluna Hitrap HP SP B) SDS-PAGE	61
Figura 29. A) Cromatograma obtido a partir da purificação da endo-1,4- β -glicanase em Superdex 75 B) SDS-PAGE	62
Figura 30: Espectros de CD no UV distante da Endo-1,4- β -glicanase em 30 mM de tampão citrato fosfato a temperatura de 20°C	65
Figura 31: Espectros de CD no UV distante da Endo-1,4- β -glicanase em 30 mM de tampão citrato fosfato a temperatura de 65°C	65
Figura 32: Desnaturação térmica da Endo-1,4- β -glicanase monitoradas por mudanças no UV distante em 222 nm em 30 mM de tampão citrato fosfato	66
Figura 33: Desnaturação química da Endo-1,4- β -glicanase por uréia A) e guanidina-HCl (B) monitoradas por mudanças no UV distante em 222 nm	68
Figura 34: Gráficos de ΔG versus [desnaturante] (A) guanidina (B) uréia	69
Figura 35: Fotomicrografia das gotas de cristalização	70

LISTA DE TABELAS

Tabela 1. Descrição das funções dos domínios das SVMPs	8
Tabela 2. Classificação das famílias dentro dos clãs	15
Tabela 3. Estatística do processamento de dados e refinamento cristalográfico	46
Tabela 4. Descrição dos ligantes acoplados e suas respectivas energias de ligação calculadas ...	51
Tabela 5. Diferença de rendimento entre diferentes concentradores	63
Tabela 6. Resultados do DLS em diferentes pHs	64
Tabela 7. Temperaturas de transição (Mt) nos respectivos pHs	66
Tabela 8. Composição de estrutura secundária endo-1,4- β -glicanase	67

LISTA DE ABREVIACÕES

- **ADAM** - *A disintegrin-and metalloproteinase*
- **ADAMT** - *A disintegrin and metalloproteinase with thrombospondin motif*
- **BaP1** – Metalloproteinase da classe PI do veneno de *Bothrops asper*
- **CBM** - *Carbohydrate-binding module/ Módulo de ligação a carboidrato*
- **CD** – *Circular Dichroism/Dicroísmo Circular*
- **CMC** – Concentração Molar Crítica
- **DLS** – *Dynamic light scattering/Espalhamento Dinâmico de Luz*
- **E. coli** - *Escherichia coli*
- **ESI-MS/MS** - *Electrospray ionization tandem mass spectrometry*
- **FPLC** - *Fast Protein Liquid Chromatography*
- **FT** – *Flow-Through*
- **IPTG** - Isopropil-tio- β -galactosídeo
- **IUBMB** - *International Union of Biochemistry and Molecular Biology Nomenclature Comitee*
- **LNLS** – Laboratório Nacional de Luz Síncrotron
- **MMPs** - *Matrix-degrading Metalloproteinases*
- **Mt** – Melting temperature /Temperatura de transição
- **NCBI** - *The National Center for Biotechnology Information*
- **Ni-NTA** - Ácido níquel-nitrilotriacético
- **O.D.** – *Optical density/densidade ótica*
- **PMSF** - Fluoreto de fenilmetil sulfonil
- **PDB** - *Protein Data Bank*
- **R.M.S.D** – *Root Mean Square Deviation/Desvio da raiz média quadrática*
- **SDS** - Dodecil sulfato de sódio
- **SDS-PAGE** - *Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis*
- **SVMPs** - *Snake Venom Metalloproteinases;*
- **TEMED** - N, N, N', N'-tetrametiletilediamina
- **UV** - Ultravioleta

RESUMO

Metaloproteinases presentes no veneno de serpentes são toxinas hemostaticamente ativas que interferem em diferentes níveis na cascata de coagulação e têm sido exploradas como ferramenta bioquímica nas pesquisas de coagulação, diagnósticos, modelos de inibidores de metaloproteinases e compreensão do mecanismo de ação de venenos. Estudos estruturais complementados por ensaios de *docking* molecular têm ajudado muito na compreensão de suas especificidades e mecanismo de ação. Assim, neste trabalho a metaloproteinase não-hemorrágica fibrinogenolítica da classe PI, isolada do veneno de *Bothrops moojeni* (BmooMP α -I) foi cristalizada e sua estrutura resolvida a 1,76 Å de resolução. O íon zinco catalítico possui uma coordenação octaédrica formada por três histidinas canônicas (His¹⁴⁰, His¹⁴⁴ e His¹⁵⁰) e por três moléculas de solvente. Uma comparação seqüencial e estudos estruturais indicaram que os motivos que compreendem os resíduos 153-164 e 167-176 possuem uma característica saliente, que diferencia as outras SVMPs hemorrágicas das não-hemorrágicas de classe PI, podendo estes motivos estarem diretamente envolvidos no desenvolvimento da atividade hemorrágica. Além do projeto principal, foram realizados estudos com a proteína endo-1,4- β -glicanase de *Bacillus subtilis* (Cel5A), que hidrolisa a ligação glicosídica β (1-4) do polímero de celulose. Esse polímero tem recebido grande atenção do meio acadêmico e industrial devido ao seu potencial na produção de açúcares solúveis, produtos químicos e biocombustíveis. Porém para o aproveitamento da biomassa celulósica é necessário a ação sinérgica do coquetel enzimático formado por endo-1,4- β -glicanases, celobiose-hidrolases ou exo-1,4- β -glicanases e β -glicosidases. Assim, o gene *Cel5A* constituído pelo domínio catalítico e o módulo de ligação à carboidratos (CBM) foi clonado em vetor pET28a e expresso em células BL21(DE3) a 37°C induzido com 1 mM de IPTG pelo período de 4 horas. A proteína foi purificada por três passos cromatográficos: afinidade, troca catiônica e exclusão molecular. A pureza foi confirmada por SDS-PAGE e a proteína se apresentou monodispersa através de análises por DLS, o que indica uma amostra adequada para os estudos biofísicos e cristalográficos. A proteína foi submetida a experimentos de desnovelamento térmico e químico monitorados por CD que indicaram uma M_t de 63,5°C. Os ensaios de cristalização foram realizados e resultaram no aparecimento de microcristais e esferolitas que estão sob refinamento. A obtenção da estrutura cristalográfica combinada com estudos de desnaturação e bioquímicos irá fornecer informações preciosas sobre o funcionamento, termoestabilidade e especificidade dessa classe de celulases.

ABSTRACT

Snake venom metalloproteinases are hemostatically active toxins that interfere in different levels on the blood coagulation cascade. Therefore, they have been extensively investigated as biochemical tools for coagulation studies and diagnostics, and as model for design of metalloproteinases inhibitors and understanding of venom action mechanism. Structural studies complemented by molecular docking experiments have been instrumental to shed light on the stereo specificity and action of enzymes. In this study, a fibrin(ogen)olytic non-hemorrhagic PI class metalloproteinase isolated from *Bothrops moojeni* (BmooMP α -I) was crystallized and its structure determined at 1,76 Å resolution. The catalytic zinc ion displays an octahedral coordination formed by three canonical histidines (His¹⁴⁰, His¹⁴⁴ e His¹⁵⁰) and three solvent molecules. Comparative sequence and structural studies indicate that the motif comprising amino acid segments 153-164 e 167-176 is a salient feature that differentiates non- and hemorrhagic PI class SVMPs and it might be directly involved in the development of the hemorrhagic activity. In parallel to the main project, a biophysical study was carried out on the thermophilic endo-1,4- β -glucanase from *Bacillus subtilis*. Endo-cellulases are responsible for the cleavage of β (1-4) glycosidic linkages from cellulose. This polymer has been receiving great attention from both industrial and academic community due to its applications in food, chemical and biofuels industries. However, for its use is needed to reduce the cellulose into fermentable sugars and this process involves the synergistic action of endo-1,4- β -glucanases, cellobiose-hydrolases or exo- β -1,4-glucanases and β -glucosidases. In this project, the gene of the Cel5A consisting of the parental catalytic domain and its carbohydrate-binding module was cloned in pET28a vector and expressed in BL21(DE3) strain at 37°C and induced with 1 mM of IPTG for the period of 4 hours. The purification of the protein was accomplished in three chromatographic steps: affinity, cation exchange and size-exclusion chromatographies. The purity was confirmed by SDS-PAGE and the protein was characterized as monodisperse in solution as assessed by DLS, which indicates an appropriate sample for biophysical and crystallographic studies. Thermal unfolding experiments showed a melting temperature of 63.5°C. Crystallization experiments were performed and results obtained indicated the formation of microcrystals and spherulites that are undergoing refinement. The crystallographic structure combined with unfolding studies will provide valuable information on both specificity and thermostability of endo-cellulases.

CAPÍTULO 1 - INTRODUÇÃO

Esta dissertação aborda o estudo de duas proteínas envolvidas em sistemas não relacionados, a coagulação sanguínea e a hidrólise enzimática de biomassa lignocelulósica. A metaloproteína não-hemorrágica foi extraída do veneno de *Bothrops moojeni* e pertencente à classe I das metaloproteases zinco-dependentes de veneno de serpentes (SVMPs). Nesta parte do projeto foram investigados os aspectos estruturais do funcionamento das SVMPs não-hemorrágicas. A outra proteína é uma endo-1,4- β -glicanase (EC 3.2.1.4) de *Bacillus subtilis*, que apresenta um peso molecular de 52,3 kDa e é formada por um domínio catalítico GH5 e um domínio de ligação ao carboidrato da família 3 (CBM3). Esta enzima sinergicamente com exo-1,4- β -glicanases (EC 3.2.1.91) e β -glicosidases (EC 3.2.1.21) participa da redução de biomassa lignocelulósica em açúcares fermentáveis. Nas seções seguintes desse capítulo será feita uma descrição mais detalhada de cada uma das proteínas estudadas.

1.1 BmooMP α -1

1.1.1. Proteínas de veneno de serpentes

Venenos de serpentes são constituídos por um complexo conjunto de proteínas (mais de 90% do seu peso seco), compostos orgânicos com baixo peso molecular e compostos inorgânicos (Bieber, 1979). Dentre as proteínas consideradas não-tóxicas, estão as lectinas, fatores de crescimento e enzimas com pouca ou nenhuma toxicidade, como as desidrogenases, amino-transferases, desoxiribonucleases e amilases. Outros componentes incluem carboidratos, lipídios (fosfolipídios), amins, ácidos orgânicos, íons, agentes quelantes (citrato) e nucleotídeos (Figura 1) (Bjarnason e Fox, 1994; Meier et al., 1995, Odell et al., 1999). A quantidade e a composição do veneno podem variar de acordo com a espécie da serpente, idade, período sazonal e dieta (Daltry et al., 1996).

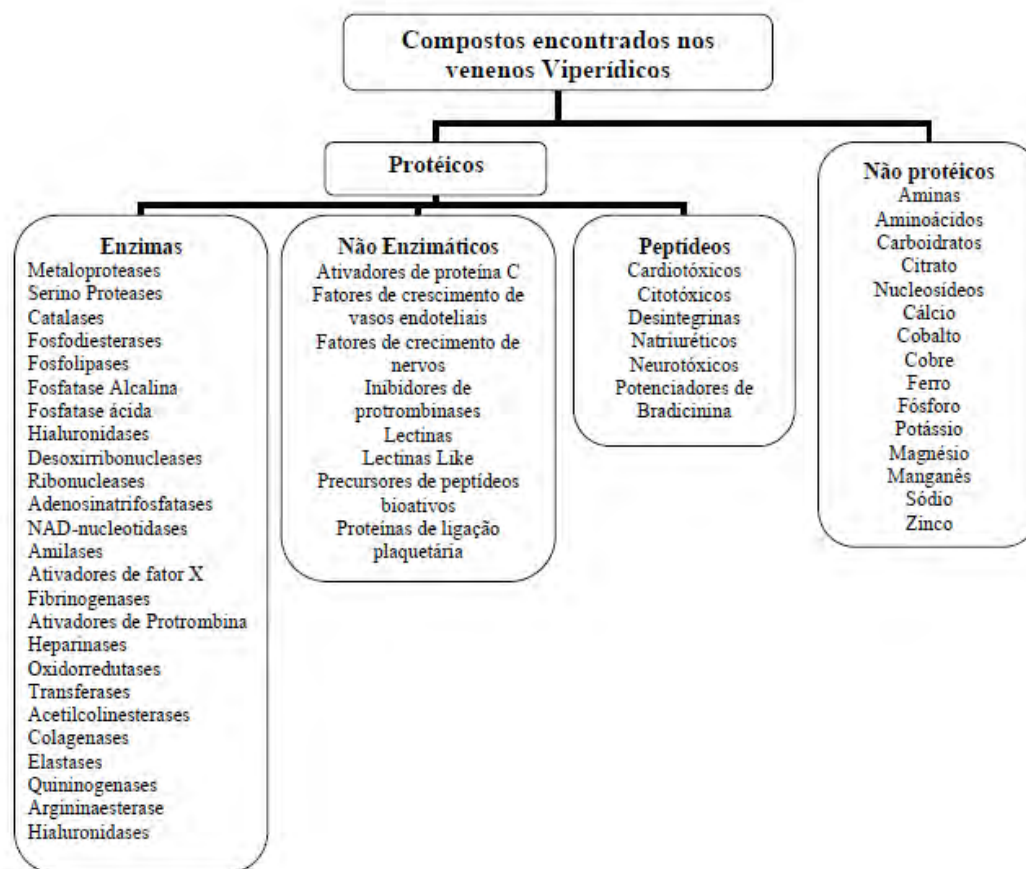


Figura1. Compostos encontrados nos venenos Viperídicos (adaptado de Sierra e Pérez, 2001).

As serpentes da família Viperidae, com algumas exceções, induzem envenenamentos caracterizados por hemorragia, além de uma infinidade de manifestações fisiopatológicas, que podem incluir necrose local, formação de bolhas e edemas, assim como efeitos sistêmicos: coagulopatia, inibição da agregação plaquetária, nefrotoxicidade, alterações hemodinâmicas e, em alguns casos, neurotoxicidade e cardiotoxicidade (Warrell, 1996). A maioria das espécies venenosas da América Latina pertencem a subfamília Crotalinae, e no Brasil, mais de 70% dos envenenamentos são causados por serpentes do gênero *Bothrops* (Fundação nacional da saúde, 2001).

No veneno das serpentes do gênero *Bothrops* as metaloproteinases e as serinoproteinases são os principais grupos de enzimas proteolíticas que podem levar ao desequilíbrio hemostático, pois estão relacionadas a diferentes fatores da cascata de coagulação, tais como fibrinogênio ou fibrina (Gasmi et al., 2000; Dunn e Broady, 2001; Takatsuka et al., 2001).

1.1.2. Metaloproteinases de veneno de serpentes (SVMPs)

As metaloproteinases são classificadas como hidrolases do tipo endopeptidases que dependem da ligação de um metal, geralmente o zinco (Zn^{2+}), em seu sítio catalítico para a manifestação de suas atividades (Blundell, 1994), podendo ser agrupadas em 4 grupos principais (Figura 2), de acordo com as similaridades seqüenciais (Hooper, 1994): a) zincinas: possuem a seqüência HEXXH, como motivo de ligação; b) Inverzincinas: que possuem um motivo invertido de ligação do zinco, HXXEH; c) carboxipeptidases: que possuem o motivo de ligação ao zinco, HXXE e d) DD-carboxipeptidases: que possuem o motivo de ligação do zinco, HXC.

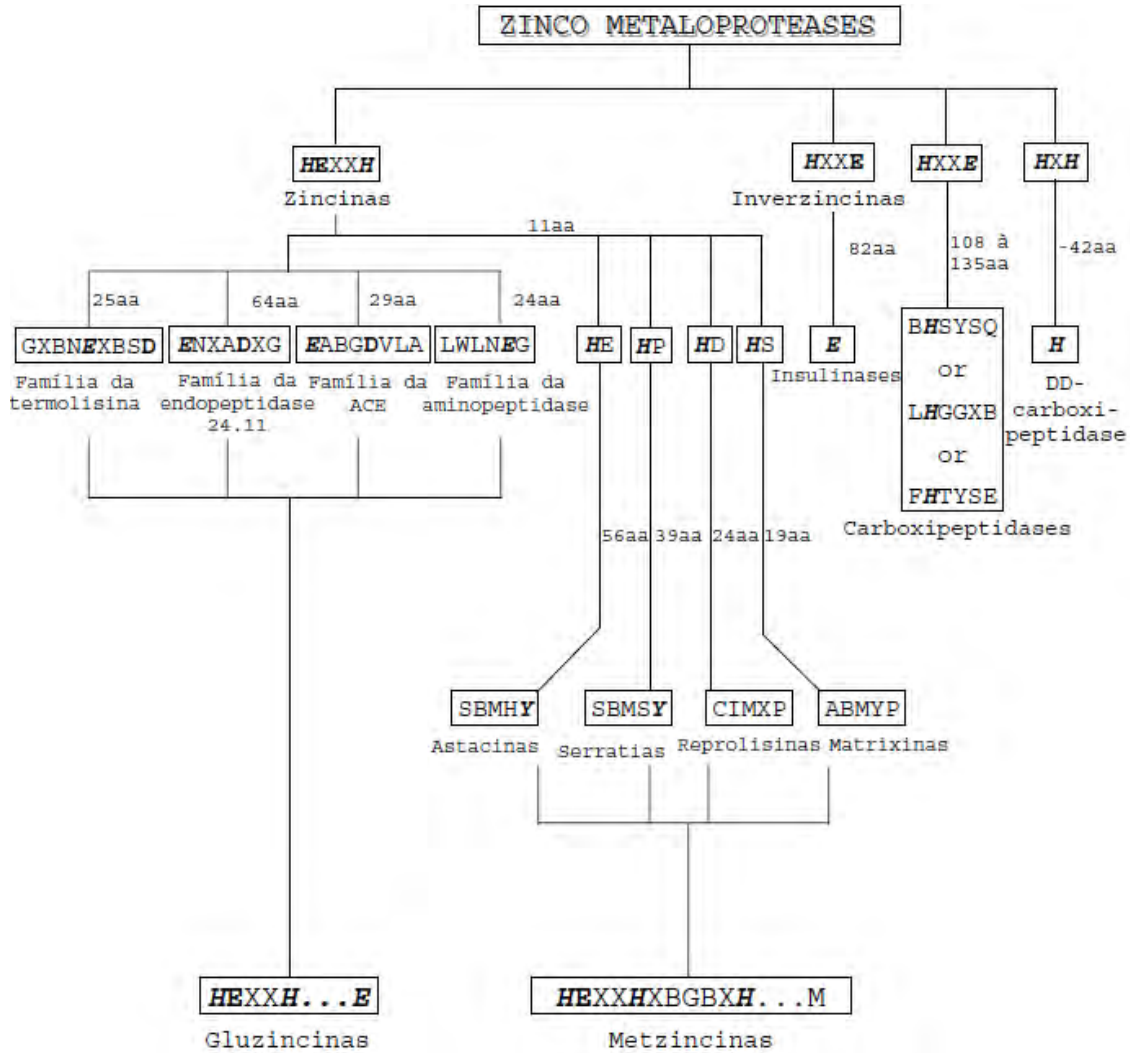


Figura 2. Família de zinco metaloproteínas (Hooper, 1994): B: resíduos polares; X: resíduos quaisquer; *Itálico*: resíduos ligados ao zinco; **Negrito**: resíduos envolvidos na catálise enzimática.

Todos os grupos possuem dois resíduos de histidinas (His) no sítio ativo, exceto as carboxipeptidases. O terceiro resíduo que coordena o metal (Zn^{2+}) pode ser His, ou um ácido glutâmico (Glu). As zincinas constituem o maior grupo, sendo subdivididas em gluzincinas (contendo Glu no sítio ativo) e metzincinas (contendo His) (Ramos, 2001). Na termolisina e carboxipeptidase o átomo de Zn^{2+} interage com o anel imidazol de dois resíduos de histidina e com a carboxila do resíduo Glu. O Zn^{2+} é também coordenado por moléculas de água que possuem um papel fundamental na atividade catalítica dessas enzimas (Creighton, 1993).

A família das metzincinas possui esse nome devido à presença do resíduo conservado de metionina que se localiza em um laço da estrutura terciária no C-terminal da

ligação do zinco com a seqüência HEXXHXXGXXH. Esse resíduo de metionina encontra-se em uma região geralmente divergente da seqüência primária, mas que é estruturalmente muito conservado. O “*Met-turn*”, como é conhecido esse laço da metionina, é localizado sob a base da pirâmide dos 3 resíduos de histidina que coordenam o íon zinco catalítico e não é envolvido em qualquer contato direto com o metal ou substrato (Figura 3) (Hege e Baumann, 2001). Incluem-se entre as metzincinas as famílias das astacinas, serratias, reprotinas e matrixinas ou MMPs (Ramos, 2001). Devido à similaridade estrutural do sítio catalítico das SVMPs e das MMPs, foi hipotetizado que inibidores das MMPs poderiam ser efetivos na neutralização das SVMPs. Essa hipótese foi comprovada no caso da BaP1, uma metaloproteinase de 22 kDa de veneno de *Bothrops asper* (Escalante et al., 2000).

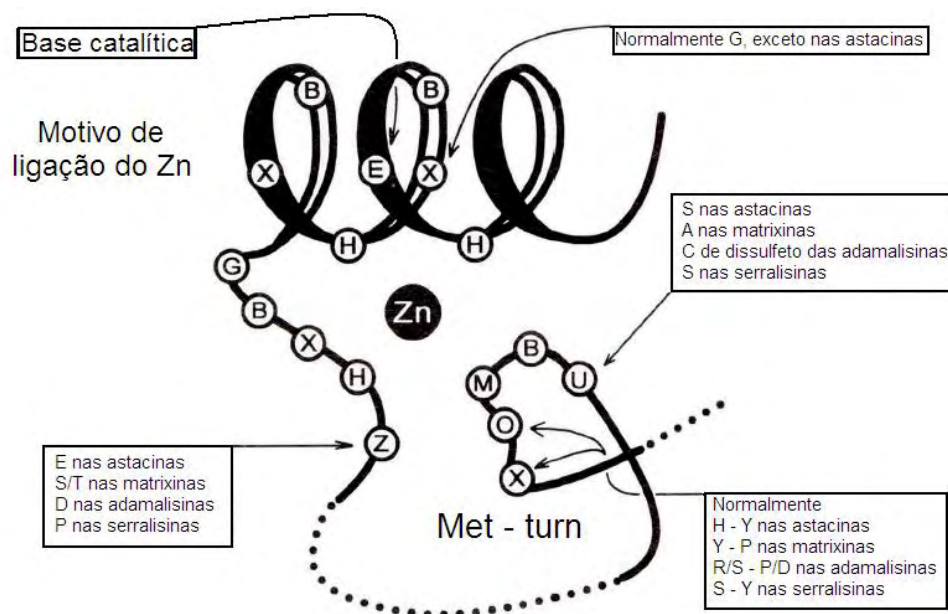


Figura 3. Esquema do motivo de ligação ao zinco e o “*Met-turn*” mostrando a seqüência comum das metzincinas. Segmentos variados são identificados pelas linhas pontilhadas (Stöcker e Bode, 1995).

As metzincinas (Bode et al., 1993) desempenham um papel fundamental nos envenenamentos causados por serpentes, sendo responsáveis por alguns efeitos patológicos tais como, hemorragia, inibição da agregação plaquetária, coagulopatia, mionecrose e resposta inflamatória (Gutiérrez et al., 2005; White, 2005; Fox e Serrano, 2005). Esses efeitos são devido à ação proteolítica das SVMPs nos componentes da matriz extracelular,

proteínas do plasma e proteínas de superfície celular (Moura-da-Silva et al., 2007). Além disso, SVMPs também são capazes de interagir com os receptores de plaquetas, células endoteliais e fibroblastos, ativando ou inibindo a resposta celular e contribuindo para a disseminação dos outros componentes do veneno (Anai et al., 2002).

O potencial hemorrágico das metaloproteínases varia largamente conforme a arquitetura estrutural (Lu et al., 2005), de forma que metaloproteínases não-hemorrágicas também foram descritas em muitos venenos (Willis e Tu, 1988; Datta et al., 1995; Terada et al., 1999). Estas enzimas possuem atividade fibrino(geno)lítica e não participam da ativação do plasminogênio. Com algumas exceções, podem ser classificadas em α - ou β -fibrinogênases de acordo com especificidade de hidrólise das cadeias do fibrinogênio. A especificidade por cadeias α ou β não é absoluta, uma vez que ocorra a degradação da cadeia alternativa com o aumento do tempo (Markland, 1998).

Vários estudos realizados para encontrar os determinantes moleculares/estruturais responsáveis pelo potencial hemorrágico das SVMPs têm sido realizados, tais como alinhamento de seqüências, reconstrução filogenética, comparação das propriedades químicas e físicas preditas, porém ainda não foi obtido nenhum resultado conclusivo (Paine et al., 1992; Manning, 1995). Watanabe et al. (2003) sugeriram que a região do laço entre o terceiro ligante do zinco (His) e o *Met-turn* poderiam estar relacionados com o fenótipo hemorrágico, devido a significativa variabilidade estrutural entre os membros das SVMPs nesses motivos. Ainda focando em características estruturais, análises de estruturas resolvidas e modeladas mostraram uma correlação indireta entre o potencial hemorrágico e a área superficial molecular polar das enzimas de classe PI, os quais podem estar contribuindo para o efeito hemorrágico (Ramos e Selistre de Araujo, 2004). Assim, um estudo mais recente propôs modificações pós-traducionais, como por exemplo, N-glicosilação na região C-terminal e resíduos de ácido aspártico (Asp¹⁴⁸) e serina (Ser¹⁷⁶) conservados no sítio catalítico, que podem potencializar a hemorragia (Ramos e Selistre de Araujo, 2004).

1.1.3. Classificação das SVMPs

Possuindo um resíduo Asp adjacente a 3^a cadeia lateral coordenante do zinco, as SVMPs e as ADAMs (*A Desintegrin and Metalloproteinases*) são classificadas como reprotinas ou subfamília M12 (Fox e Serrano, 2005). Algumas ADAMs possuem domínios adicionais como transmembrana, citoplásmico e trombospondina típico das ADAMTS (*A Disintegrin And Metalloproteinases with Thrombospondin motifs*) (Fox e Serrano, 2005).

SVMPs são sintetizadas como proteínas multidomínio e podem ser classificadas em 4 grupos de acordo com sua estrutura primária e composição de domínios (Figura 4 e Tabela 1): a) classe PI (cerca de 24 kDa): possui somente o domínio catalítico e tipicamente apresentam pouca ou nenhuma atividade hemorrágica; b) classe PII (cerca de 35 kDa): contém um domínio desintegrina em adição ao domínio catalítico; c) classe PIII (cerca de 55 kDa): é composta por um domínio catalítico, um domínio desintegrina e um domínio rico em cisteína que exercem fortes hemorragias ou efeitos de coagulação e d) classe PIV (cerca de 95 kDa): possuem um domínio adicional de tipo lectina ligado por ligações dissulfeto ao domínio da estrutura PIII (Fox e Serrano, 2005; Bjarnason e Fox 1994). O efeito hemorrágico é dependente da atividade proteolítica dessas enzimas, visto que a quelatação do zinco com o EDTA impede esse efeito (Bjarnason e Fox, 1994; Escalante et al., 2000).

É interessante notar que a atividade hemorrágica e a patologia associada com envenenamentos por serpentes da família Viperidae foram associados as SVMPs das famílias PI, PII e PIII. Assim, a presença ou a ausência de certos domínios não necessariamente denotam a capacidade hemorrágica, embora em termos de potência as proteinases da classe PIII são geralmente mais potentes (Fox e Serrano, 2005).

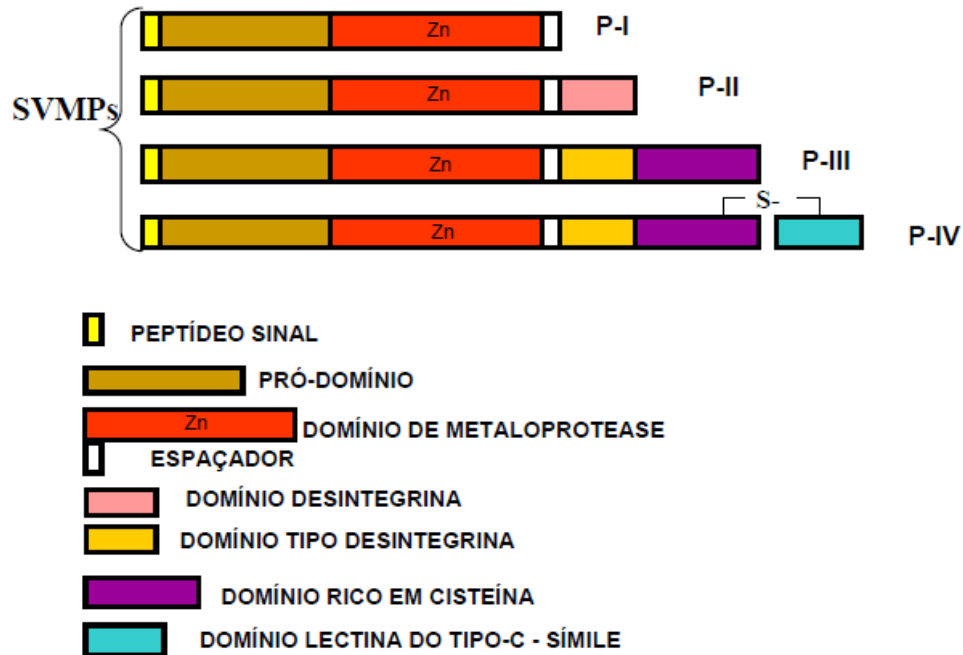


Figura 4. Domínios das SVMPs (Junqueira, 2005).

Tabela 1. Descrição das funções dos domínios das SVMPs (Figura 4)

Domínios	Função
Peptídeo sinal	Dirige a proteína para a localização subcelular ou para ser secretada
Pró-domínio	Participa na regulação da enzima através da interação com o sítio de atividade
Catalítico de metaloproteinase	Contém o sítio de atividade enzimática de metaloproteinase
Desintegrina	Quando processado (livre do domínio metaloproteinase) possui a capacidade de inibir a agregação plaquetária e também bloqueia a adesão das plaquetas ao colágeno
Rico em cisteína	Possui grande quantidade de resíduos cisteína, e esta intimamente relacionado à inibição da agregação plaquetária induzida por colágeno
Lectina tipo-C	Domínio exclusivo das SVMPs da classe PIV, homólogo as lectinas que são proteínas que ligam-se a açúcares de forma específica. Nos venenos elas afetam também as funções plaquetárias induzindo a agregação

1.1.4. Mecanismo de catálise proposto para as SVMsPs

A região de ligação ao zinco de todos os quatro protótipos das estruturas das metzincinas são muito semelhantes. O íon zinco é tetracoordenado por uma molécula de água e 3 átomos de nitrogênio da cadeia lateral dos resíduos de histidina (distância da ligação varia entre 2,0 a 2,2 Å), na seqüência característica altamente conservada HEXXHXXGXXH. Essa molécula de água, localizada entre o grupo carboxílico da cadeia lateral do Glu e o íon zinco passa por um aumento de polarização (Figura 5A), ocorrendo uma transferência de um hidrogênio da molécula de água para o resíduo Glu. Assim o grupo carboxílico do resíduo Glu se torna um nucleófilo e ataca uma ligação peptídica acessível do substrato, fazendo com que o metal passe para um estado pentacoordenado (Figura 5B). O átomo de hidrogênio do grupo carboxílico do resíduo Glu é transferido ao átomo de nitrogênio da ligação peptídica do substrato e a ligação peptídica é rompida, liberando os produtos (Figura 5C). A enzima é então regenerada com a incorporação de novas moléculas de água. (Gong et al., 1998; Stöcker e Bode, 1995)

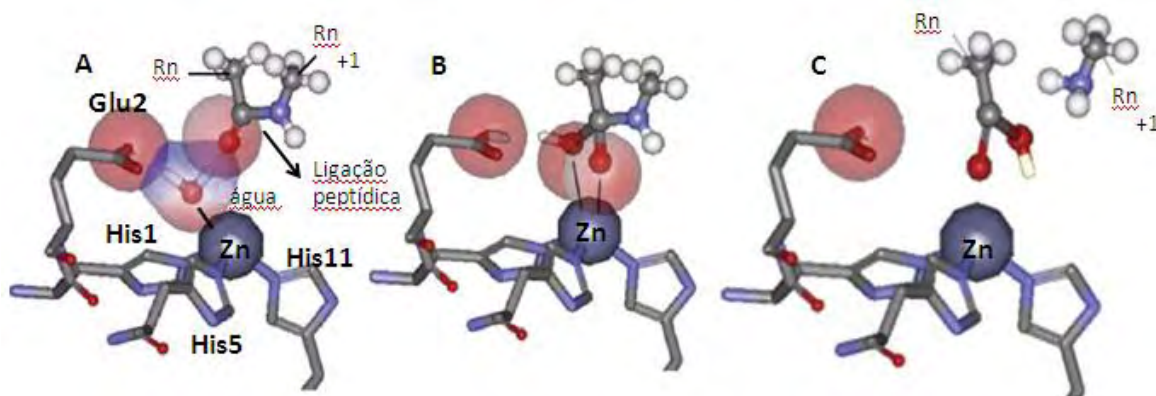


Figura 5. Mecanismo de catálise para as SVMsPs. Os números dos resíduos correspondem a suas posições no motivo de ligação do zinco (HEBXHXBGBXH). Rn e Rn +1 são, respectivamente, resíduos N- e C- terminal da ligação peptídica. (A) Aumento da polarização da molécula de água presa entre o zinco e o ácido carboxílico do resíduo Glu. (B) transferência do próton do resíduo Glu para quebrar a ligação peptídica e a formação do estado intermediário pentacoordenado do zinco. (C) Transferência do próton do resíduo Glu para o nitrogênio da ligação peptídica, quebra da ligação peptídica e liberação dos produtos. (Figura retirada e modificada de Ramos e selistre, 2006)

Na figura 6 podemos observar os quatro ligantes cercando o zinco em uma esfera de coordenação trigonal piramidal, com o nitrogênio- ϵ da segunda histidina na ponta da pirâmide, e o zinco sendo quase coplanar com os outros ligantes que formam a base.

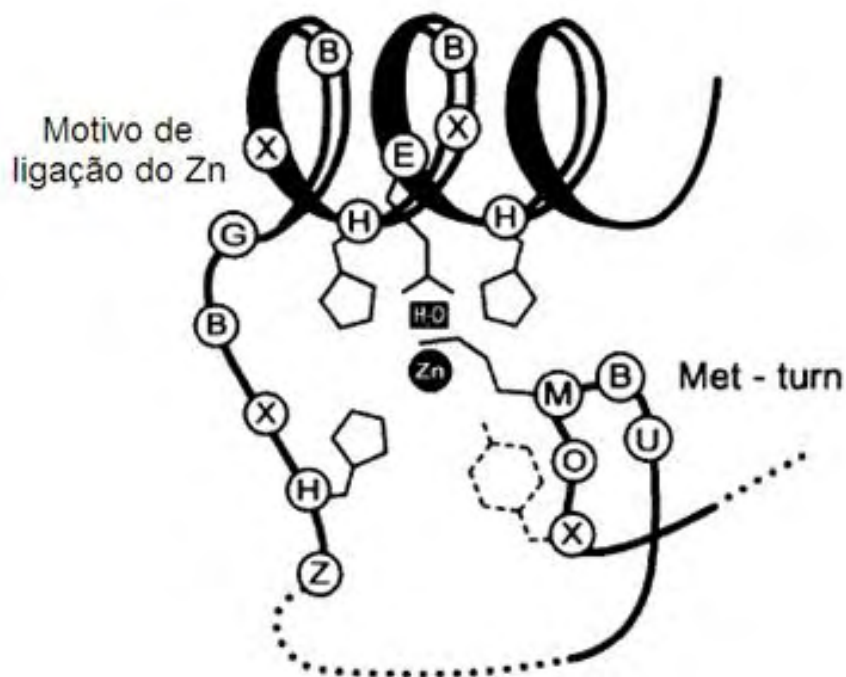


Figura 6. Esquema do motivo de ligação ao zinco e o “Met-turn” das metzincinas, evidenciando os resíduos envolvidos na catálise.

1.1.5. BmooMP α -I

BmooMP α -I é uma metaloproteinase dependente de zinco de classe PI e foi isolada do veneno de *Bothrops moojeni*, possui peso molecular de 22,6 kDa, e apresenta atividades azocaseinolítica, fibrinogenolítica e desfibrinante, porém é desprovida de atividades tipo trombina e hemorrágica (Bernardes et al., 2008). Experimentos realizados *in vivo* em camundongo mostraram que a BmooMP α -I age como um agente desfibrinogenante, removendo o fibrinogênio do sangue. Esses resultados mostram o grande potencial dessa proteína como anticoagulante e agente terapêutico no tratamento e prevenção de doenças trombóticas. Dessa forma, neste trabalho estamos interessados em compreender a base molecular da diversidade funcional apresentada pela classe PI das SVMPs, e realizar estudos estruturais comparativos com outras SVMPs hemorrágicas e não-hemorrágicas, com a finalidade de esclarecer seus efeitos hemostáticos.

1.2 Endo-1,4- β -glicanase

1.2.1. Estrutura da Celulose

A celulose é o polímero natural e biodegradável mais abundante no planeta, não é solúvel em água e é o principal constituinte da parede celular das plantas (cerca de 33% do peso da planta). Classificamos a celulose como um homopolissacarídeo linear, não-ramificado que pode possuir de 10 a 15 mil unidades de D-glicose. Essas unidades são unidas por ligações glicosídicas $\beta(1-4)$ formando unidades de celobiose e em seguida polímeros ou β -glicanas (Figura 7), que por sua vez agregam-se lateralmente por ligações de hidrogênio e forças de Van der Waals formando feixes de mais ou menos 36 deles, e esses feixes reúnem-se lado a lado para formar uma microfibrila (Nelson e Cox, 2006; Somerville et al., 2004). Trata-se de um polímero caracterizado por regiões ordenadas (cristalinas) em grande parte de seu comprimento e entrecortadas por zonas desordenadas (amorfas) (Stone e Scallan, 1965).

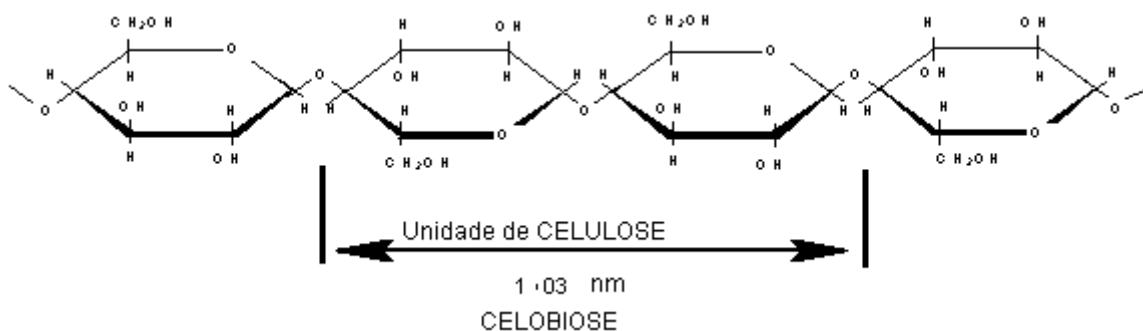


Figura 7. Esquema das unidades de D-glicose unidas por ligações $\beta(1-4)$ formando a unidade básica da celulose, a celobiose.

As regiões ordenadas são amarradas tão fortemente que nem as moléculas de água podem penetrar. O acesso limitado a essas regiões leva à alteração nas propriedades físico-químicas e na eficiência de agentes reativos, como celulases. Com este tipo de estrutura, é evidente que apenas as moléculas de celulose situadas na superfície seriam suscetíveis à ação de enzimas. Assim as celulases precisam adsorver a superfície da celulose antes da hidrólise, promovendo a desordem molecular das regiões empacotadas e expondo as cadeias de celulose enterradas enquanto elas permanecem molecularmente inalteradas

(*amorfogenesis*) (Coughlan, 1985). Quando a rede de celulose torna-se acessível a enzimas, a ação sinérgica de endo e exo glucanases causam a fragmentação da celulose liberando celo-oligossacarídeos solúveis (moléculas derivadas da celulose com um grau de polimerização menor que 6 unidades), que são rapidamente hidrolisadas, a maioria para celobiose (Coughlan, 1985).

1.2.2. Celulose vs Amilose

A amilose apresenta-se de forma muito semelhante à celulose, a única diferença é que na celulose os resíduos de glicose têm configuração β e na amilose possuem configuração α (Figura 8). Essas diferenças conferem a esses polissacarídeos estruturas tridimensionais e propriedades físicas muito diferentes entre si. Podemos representar os polímeros de glicose da celulase como uma série de anéis piranosídicos rígidos na conformação em “cadeira”, conectados por ligações de hidrogênio intermoleculares, que ocorrem entre grupos hidroxila de cadeias adjacentes e intramoleculares que ocorrem entre grupos hidroxila da mesma molécula (Figura 8A e 9A). Essas interações são responsáveis pela formação da fibra vegetal e pela rigidez da cadeia, respectivamente.

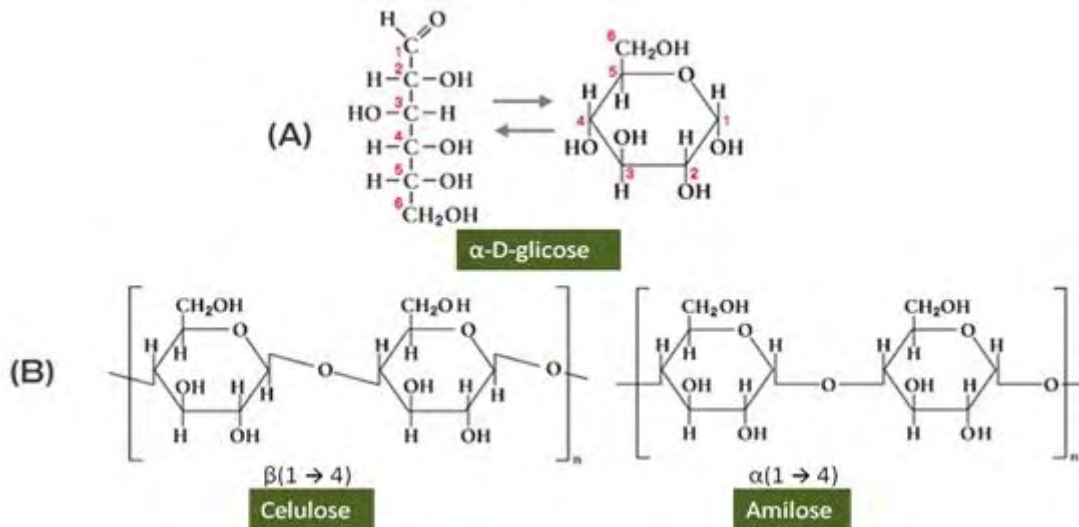


Figura 8. (A) Glicose na forma linear e piranosídica (B) Comparação entre a celulose e a amilose.

Em contraste, a conformação favorável para os polímeros da amilose é uma estrutura helicoidal estabilizada por ligações de hidrogênio (Figura 9B) (Lehninger et al., 1995).

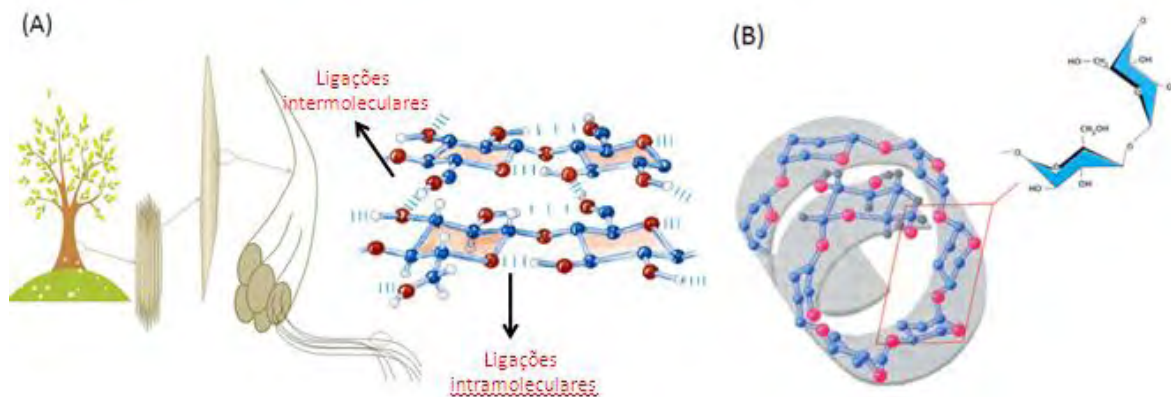


Figura 9. (A) Estrutura da celulose, as microfibrilas se combinam para formar a fibra de celulose (B) Estrutura da amilose, as ligações $\alpha(1-4)$ conferem a amilose uma estrutura helicoidal.

Diferentemente do amido e do glicogênio, que possuem ligações $\alpha(1-4)$, a celulose não pode ser digerida pelas α -amilases, ela necessita de enzimas específicas para degradá-la, as celulasas. Os bovinos e alguns ruminantes possuem protistas e bactérias em seus estômagos, que secretam celulase, assim como o cupim que possui um microrganismo simbiótico que também têm essa habilidade. O restante dos animais não consegue utilizar a celulose como fonte de energia.

1.2.3. Aplicações da Celulose

Nos últimos anos a biomassa celulósica tem recebido grande atenção do meio acadêmico e industrial devido ao seu potencial na produção de açúcares solúveis, produtos químicos e biocombustíveis. Podendo vir a servir como fonte renovável de energia se for descoberto uma forma de transformar esse polímero em açúcar simples para fermentação de forma economicamente viável.

Materiais de base celulósica também têm sido aplicados em várias áreas da medicina como hemodiálise, nomeadamente como componentes de matrizes para regeneração óssea, vasos sanguíneos artificiais e substitutos temporários de pele.

Porém, para o aproveitamento da biomassa celulósica é necessário sua redução a açúcares fermentáveis e isso envolve a ação sinérgica do coquetel enzimático formado por endo-1,4- β -glicanases (E.C. 3.2.1.4), celobiose-hidrolases ou exo-1,4- β -glicanases (E.C. 3.2.1.91) e β -glicosidases (E.C. 3.2.1.21). Essas enzimas enfrentam alguns obstáculos para efetivar a hidrólise enzimática, e o principal deles é o acesso limitado à grande parte da celulose, que está enterrada dentro da arquitetura fibrilar (Arantes e Saddler, 2010)

O processo total de conversão bioquímica da celulose a etanol envolve três etapas básicas:

1. o pré-tratamento, que causa o rompimento da hemicelulose, a transformação da lignina e o aumento do potencial de hidrólise da celulose;
2. a hidrólise da celulose por enzimas celulolíticas para a obtenção de uma solução rica em glicose
3. a fermentação alcoólica.

Porém todo esse processo é muito lento e oneroso, visa-se futuramente a busca de novas enzimas mais eficientes, um microorganismo com alto potencial de síntese dessas enzimas, melhoramento genético para aumentar a atividade específica dessas enzimas, para assim poder baratear esse coquetel enzimático e conseguir uma produção em larga escala.

1.2.4. Características macromoleculares da celulose

Apesar de possuir tantas hidroxilas livres, esse polímero tão abundante na terra, possui algumas características muito interessantes, como sua insolubilidade em água. Não é novidade que a insolubilidade de um polissacarídeo linear não está relacionada com as propriedades intrínsecas da molécula de forma direta, mas está relacionada com as interações entre moléculas semelhantes ou diferentes. Dessa forma quando duas cadeias polissacarídicas se aproximam elas possibilitam a formação de inúmeras ligações de hidrogênio entre si, diminuindo a entalpia relativa desse complexo. Essa entalpia adquirida elimina qualquer aumento de entropia ou diminuição da entalpia relativa adquirida após a inserção de moléculas de água a esse complexo (Genta, 2000).

Fan et al. (1987), em seus estudos sobre estruturas cristalinas das celulosas I e II, evidenciou que a hidroxila do carbono 3 está envolvida em uma ligação de hidrogênio com

o oxigênio do anel piranosídico da seguinte cadeia, e que as hidroxilas dos carbonos 2 e 6 estão envolvidas na formação de ligações de hidrogênio com as cadeias de celulose adjacentes. Quando se têm polímeros que não possuem emparelhamentos entre cadeias regulares, nota-se diferentes graus de solubilidade, ou seja, o número de ramificações está diretamente ligado com grau de solubilidade desses polissacarídeos (Nelson e Lewis, 1974).

1.2.5. Glicosidases

Hidrolases glicosídicas (E.C. 3.2.1.X), ou glicosidases são enzimas chave no metabolismo de carboidratos que hidrolisam as ligações glicosídicas. A nomenclatura dessas enzimas segue o padrão da *International Union of Biochemistry* (1984), onde os primeiros 3 dígitos, no caso das glicosidases, indicam enzimas que hidrolisam ligações O-glicosil e o X indica sua especificidade ao substrato e às vezes o mecanismo molecular (Henrissat, 1991). Essa classificação não está necessariamente relacionada com suas características estruturais.

Podemos classificar as glicosidases em famílias, baseado em suas similaridades sequenciais. Hoje em dia existem mais de 150 famílias catalogadas no banco de dados CAZy (*Carbohydrate-Active Enzymes* - <http://www.cazy.org/Glycoside-Hydrolases.html>). Esse banco de dados fornece atualizações contínuas das famílias.

Devido ao aumento do número de estruturas de glicosidases resolvidas e o fato do *fold* das proteínas serem bem mais conservados que suas seqüências foi possível melhorar as estratégias de comparação e assim, dividir algumas famílias em clãs (Henrissat e Bairoch, 1996):

Tabela 2. Classificação das famílias dentro dos clãs (<http://www.cazy.org/Glycoside-Hydrolases.html>)

Clãs	Estrutura	Famílias
GH – A	(β/α) ₈	1,2,5,10,17,26,30,35,39,42,50,51,53,59,72,79,86,113
GH – B	β - jelly roll	7,16
GH – C	β - jelly roll	11,12
GH – D	(β/α) ₈	27,31,36
GH – E	6-fold β -propeller	33,34,83,93
GH – F	5-fold β -propeller	43,62

GH – G	$(\alpha/\alpha)_6$	37,63
GH – H	$(\beta/\alpha)_8$	13,70,77
GH – I	$\alpha + \beta$	24,46,80
GH – J	5-fold β propeller	32,68
GH – K	$(\beta/\alpha)_8$	18,20,85
GH – L	$(\alpha/\alpha)_6$	15,65
GH – M	$(\alpha/\alpha)_6$	8,48
GH – N	β - helix	28,49

1.2.5.1. Mecanismo de catálise e modo de ligação das glicosidases

Quanto ao mecanismo de catálise dessas enzimas, podemos classificá-las como enzimas que invertem ou retêm a estereoespecificidade do carbono número 1 do resíduo de glicose envolvido na ligação glicosídica. Dessa forma, as enzimas inversoras produzem, a partir do substrato, um resíduo de α -glicose que pode estar isolado ou ainda preso na ponta redutora da cadeia. As enzimas retentoras mantêm a configuração β original (Davies e Henrissat, 1995).

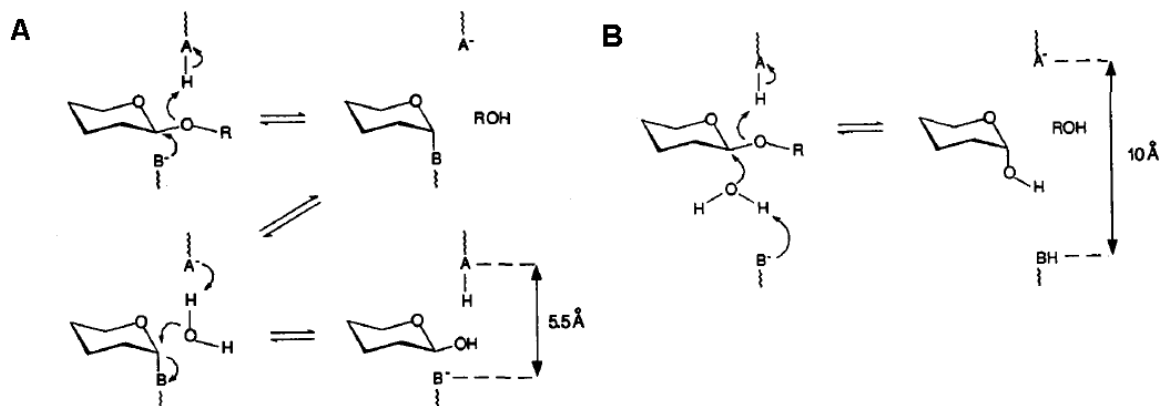


Figura 10. Mecanismos enzimáticos da hidrólise da ligação glicosídica. A) Mecanismo de retenção B) Mecanismo de inversão da configuração do carbono anomérico, propostos por Koshland (1953) e revisados por Davies e Henrissat (1995).

Ambos os tipos de enzimas utilizam um mecanismo de catálise ácida, onde dois resíduos carboxilados estão envolvidos, um agindo como nucleófilo e o outro como ácido/base. O próton doador encontra-se mais ou menos a distância de uma ligação de hidrogênio do oxigênio da ligação glicosídica. Nas enzimas retentoras, a base do ataque

nucleófilo encontra-se na vizinhança próxima do carbono anomérico, enquanto que nas enzimas inversoras há espaço para o posicionamento de uma molécula de água entre a base (nucleófilo) e o açúcar (carbono anomérico) (Figura 10). Dessa forma, nas enzimas retentoras a distância entre os dois grupos catalíticos é cerca de 5,5 Å e nas inversoras esta distância é da ordem de 10 Å. Como podemos observar pela figura 10A o carbono anomérico é atacado pelo próprio grupo nucleófilo da enzima, e na figura 10B o carbono é atacado pelo átomo de oxigênio (devidamente polarizado pelo nucleófilo da enzima) da molécula de água que se posiciona no interior do sítio ativo (Davies e Henrissat, 1995).

Outro modo de classificação dessas enzimas é baseado no modo de ataque (exo/endo). As exo clivam o substrato no final da cadeia (mais freqüentemente no terminal não-redutor) (Figura 11), já as endo clivam o substrato da cadeia. Essa diferença é perceptível facilmente no grau de viscosidade na solução de polímero, se a viscosidade de uma solução for reduzida rapidamente trata-se de uma endo-1,4-β-glicanases, porém se a enzima tiver pouca ou nenhuma influência sobre esse parâmetro trata-se de uma exo-1,4-β-glicanases (Bochkov et al., 1972).

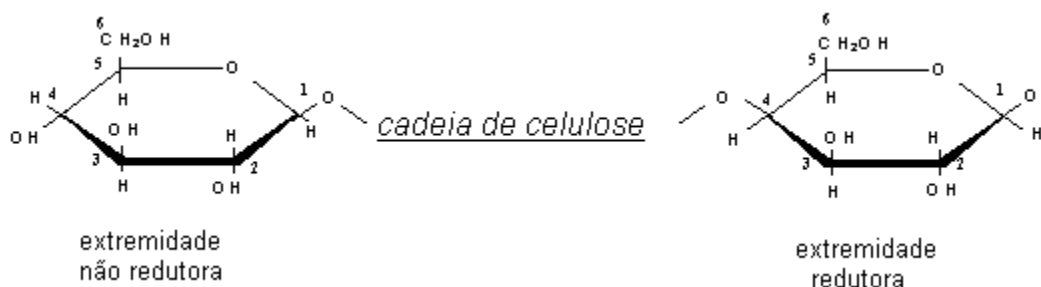


Figura 11. Esquema evidenciando as extremidades não-redutoras e redutoras.

1.2.6. β-Glicanases

Dentre essas glicosidases existem algumas famílias que possuem características de β-glicanases, ou seja, que hidrolizam especificamente as ligações glicosídicas em β-glicanas. As β-glicanases possuem diferentes mecanismos de catálise, processividade, padrão de ação e especificidade e, portanto são divididas em 9 classes enzimáticas de acordo com a *International Union of Biochemistry and Molecular Biology Nomenclature Comitee* (IUBMB), 1992:

- E.C. 3.2.1.4 Celulase – Hidrólise interna de ligações glicosídicas β (1 \rightarrow 4) em celulose, liquenana e β -glicanas cereais.
- E.C. 3.2.1.6 Endo (1-3(4)) β -D-glicanase – Hidrólise interna de ligações (1 \rightarrow 3) ou (1 \rightarrow 4) em β -glicanas quando o resíduo de glicose cujo grupo redutor está envolvido na ligação a ser hidrolizada é ligado à unidade glicosídica anterior pelo carbono 3.
- E.C. 3.2.1.39 Glican endo (1-3) β -D-glicosidase – Hidrólise de ligações glicosídicas do tipo (1 \rightarrow 3) em β -(1 \rightarrow 3)-glicanas.
- E.C. 3.2.1.58 Glican (1-3) β glicosidase – Hidrólise sucessiva de unidades de β -D-glicose da ponta não-redutora de β (1 \rightarrow 3) glicanas, liberando o anômero α da glicose.
- E.C. 3.2.1.71 Glican endo-1,2- β -glicosidase – Hidrólise randômica de ligações glicosídicas β (1 \rightarrow 2) em β (1 \rightarrow 2) glicanas.
- E.C. 3.2.1.73 Licheninase – Hidrólise de ligações glicosídicas β (1 \rightarrow 4) em β glicanas contendo ligações (1 \rightarrow 3) e (1 \rightarrow 4).
- E.C. 3.2.1.74 Glican (1-4) β -D-glicosidase – Hidrólise de ligações em β (1 \rightarrow 4) glicanas de maneira a remover unidades sucessivas de glicose.
- E.C. 3.2.1.75 Glican endo-1,6- β -glicosidase – Hidrólise randômica de ligações glicosídicas β (1 \rightarrow 6) em β (1 \rightarrow 6) glicanas.

1.2.7. Celulases E.C. 3.2.1.4 (Família 5)

Celulases pertencentes à família 5 são encontradas na natureza possuindo apenas o domínio catalítico ou o domínio catalítico em fusão a um módulo de ligação a carboidratos, incluindo as famílias 3, 5, 6, 17 e 28. Essa é uma das maiores famílias de glicosidases (Sakon et al., 1996), suas seqüências são diversas mas mantêm sete resíduos conservados próximos ao sítio catalítico (Wang et al., 1993) equivalentes a Arg⁶², His¹¹⁶, Asn¹⁶¹, Glu¹⁶², His²³⁸, Tyr²⁴⁰ e Glu²⁸² na estrutura da celulase de *Acidothermus cellulolyticus* resolvida por Sakon et al. (1996).

O mecanismo de clivagem utilizado por essas celulases é o de retenção, sendo o Glu¹⁶⁹ o próton doador e o Glu²⁵⁷ o nucleófilo.

Podemos encontrar no Protein Data Bank (PDB) diversas estruturas de celulases de vários organismos como: 1ECE (*Acidothermus cellulolyticus*), 1H1N (*Thermoascus aurantiacus*), 1BQC (*Thermomonospora fusca*), 1H5V (*Bacillus agaradharens*) e 1CEC (*Clostridium thermocellum*). Sendo que todas apresentam uma estrutura de barril (β/α)₈, onde as folhas β e as hélices α formam um esqueleto estrutural da enzima, enquanto as voltas contêm os aminoácidos responsáveis pela química catalítica. As folhas β torcidas e paralelas formam um núcleo na forma de barril e as hélices α conectam as folhas β fora do barril. As cadeias laterais hidrofóbicas das hélices α se empacotam contra as cadeias laterais hidrofóbicas das folhas β , fazendo com que o barril se torne uma concha de resíduos hidrofóbicos. O sítio ativo é um bolsão formado pela região de laços que conectam as carboxilas terminais das folhas β com os amino terminais das hélices α (Valério, 2007).

1.2.8 Módulos de ligação ao carboidrato (CBM)

Muitas enzimas que hidrolisam carboidratos, tais como celulases e xilanases, são proteínas modulares com pelo menos dois domínios distintos: o domínio catalítico e o módulo de ligação ao carboidrato (CBM) (Gilkes et al., 1991). CBMs possuem uma ou mais funções tais como: concentração da enzima na superfície do substrato/efeito de proximidade, seletividade pelo substrato e rompimento de substratos cristalinos via ação não-hidrolítica (Arante e Saddler, 2010). Esses módulos não-catalíticos (CBM) adsorvem em locais acessíveis ao substrato para formar um complexo específico não-covalente, com ligações termodinamicamente favoráveis (Lynd et al., 2002). Conseqüentemente o domínio catalítico é alinhado com o substrato para estabelecer uma alta concentração de enzima na superfície da celulose.

Vários pesquisadores têm mostrado que a remoção do CBM reduz a atividade hidrolítica do módulo catalítico em substratos insolúveis cristalinos como o Avicel, algodão e filtro de papel, entretanto sua atividade sobre celuloses amorfas ou solúveis não se altera (Ragothama et al., 2000; Irwin et al., 1993).

Trabalhos anteriores de Teeri et al. (1991) propõem que o CBM se liga às fibras de celulose e penetra na rede fibrilar da superfície descontínua, liberando fragmentos de celulose que estão não-covalentemente ligados a microfibrilas subjacentes. Também foi

sugerido que a penetração do CBM esfolia a fibra estrutural, liberando as extremidades das cadeias de celulose que permanecem ligados a fibra, resultando em uma rugosidade na superfície.

Pesquisadores russos (Klyosov e Rabinovich, 1980; Klyosov, 1990) propuseram um mecanismo para tentar explicar a dispersão da celulose (Figura 12). Eles propuseram que as celulasas são adsorvidas pelos defeitos da celulose (perturbações na estrutura cristalina da celulose) (Figura 12A), seguida por uma penetração nos espaços interfibrilares (Figura 12B). Conseqüentemente isso vai induzir uma ação mecânica (dispersão) da estrutura da celulose. Isso sugere que a presença de uma enzima grande dentro de um espaço estreito causa um aumento na pressão mecânica exercidas sobre as paredes da cavidade, inchando a estrutura da celulose e acomodando mais e mais moléculas de água entre as microfibrilas (Figura 12C). A água penetra no interior dos defeitos mais e mais dentro do espaço capilar, quebrando as ligações de hidrogênio entre as cadeias de celulose, resultando na dissociação individual das microfibrilas. Por sua vez as enzimas adsorvidas evitam que as cadeias solvatadas e as extremidades livres se realinhem e possam aderir umas as outras novamente.

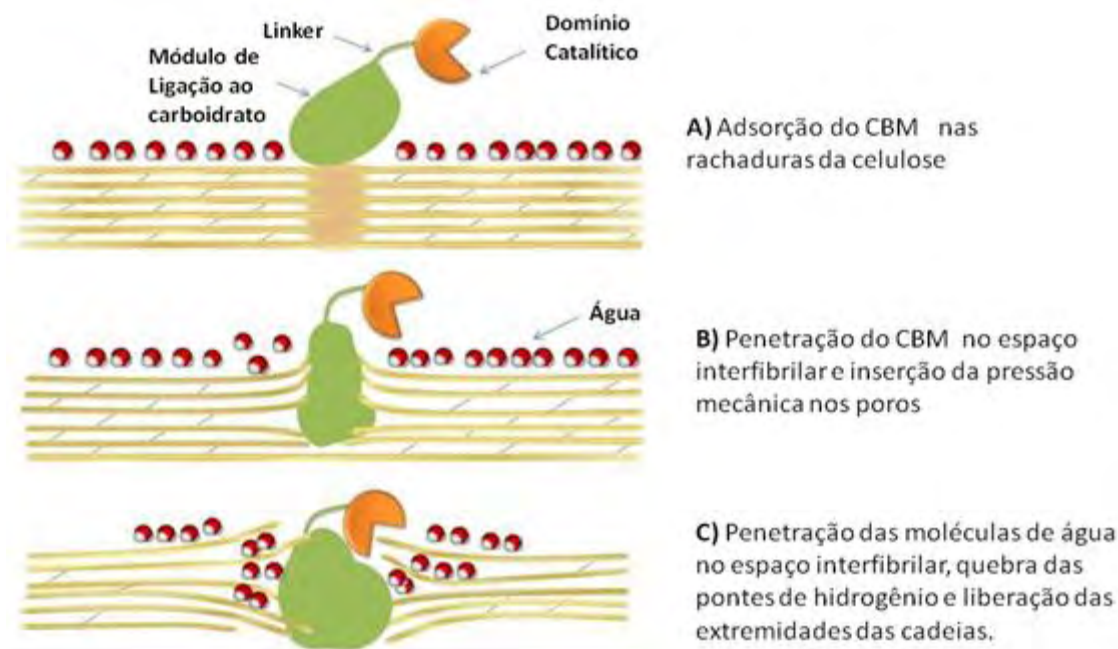


Figura 12. Representação esquemática da amorfogênese das fibras de celulose mediadas pelo CBM da celobiohidrolase I (adaptação de Esteghalian et al., 2000).

Embora as funções do CBMs durante a hidrólise enzimática da celulose ainda não tenha sido completamente elucidada e continua sendo um alvo de pesquisa parece razoável acreditar que o primeiro papel do CBM é ancorar o módulo catalítico à celulose. Esse ancoramento é geralmente aceito para aumentar a efetividade da concentração das celulasas no substrato sólido, ajudando assim a enzima na transferência de fase da fração solúvel (enzima) para a fração insolúvel (substrato) (Levy e Shoseyov, 2002).

1.2.9 Endo-1,4- β -glicanase de *Bacillus subtilis*

O gênero *Bacillus* é um dos representantes das bactérias celulolíticas, a clonagem e a caracterização de diversas exo e endo glicanases desse gênero vem sendo extensivamente pesquisadas, com a finalidade de investigar o domínio catalítico e o CBM.

Essa celulase pertencente à família 5 (E.C. 3.2.1.4) foi caracterizada por Oh et al. em 2008, possuindo peso molecular de 52,3 kDa ela cliva as ligações $\beta(1-4)$ do meio da cadeia dos polímeros de celulose. Seu mecanismo de clivagem é baseado no modelo de retenção proposto por Koshland (1953) e ela apresenta dois domínios: o domínio catalítico e o módulo de ligação a carboidratos. No presente trabalho estamos interessados em determinar a estrutura cristalográfica dessa enzima que apresenta um grande potencial biotecnológico na produção de etanol.

CAPÍTULO 2 - OBJETIVOS

Os principais objetivos deste projeto foram:

✓ A determinação da estrutura cristalográfica da metaloproteinase isolada de veneno de *Bothrops moojeni*, juntamente com os ensaios de *docking* molecular realizados com ligantes de estruturas já conhecidas. Contribuindo assim, na elucidação das bases moleculares do funcionamento das SVMPs, principalmente das não hemorrágicas que possuem grande potencial biomédico.

✓ Expressão e purificação em grande escala e alta pureza da proteína recombinante endo-1,4- β -glicanase de *Bacillus subtilis*, assim como seus ensaios cristalográficos e estudos biofísicos. Fornecendo informações muito importantes sobre seu comportamento hidrodinâmico, termoestabilidade e especificidade.

CAPÍTULO 3 – PROCEDIMENTOS EXPERIMENTAIS

Esse capítulo descreve de maneira detalhada os materiais e os métodos utilizados na execução dos projetos apresentados nessa dissertação de mestrado. Todos os experimentos foram realizados no Laboratório Nacional de Biociências (LNBio/LNLS), em Campinas.

3.1 BmooMP α -I

O veneno bruto liofilizado foi cedido pelo serpentário SanMaru (Taquaral, São Paulo, Brasil).

3.1.1. Primeiro passo de purificação: Troca Aniônica

A BmooMP α -I foi purificada de acordo protocolo previamente publicado por Bernardes et al., 2008, com pequenas modificações (aumento de um passo cromatográfico, mudança no fluxo das cromatografias por exclusão molecular e a liofilização foi trocada por concentração). A purificação consistiu em quatro passos cromatográficos sequenciais: troca aniônica, exclusão molecular, afinidade e exclusão molecular.

Materiais:

- coluna: Q-Sepharose (Amershan Biosciences),
- tampão 1: bicarbonato de amônio (AMBIC) 50 mM pH 7,8
- tampão 2: bicarbonato de amônio (AMBIC) 450 mM pH 7,8
- FPLC: GE Pharmacia AKTA Purifier,
- vazão: 1 mL/min.

Métodos:

A purificação foi realizada em coluna comercial Q-Sepharose. A cromatografia foi realizada em um sistema FPLC de baixa pressão. A coluna foi equilibrada com o tampão 1, a uma vazão de 1 mL/min em temperatura ambiente. Em seguida a amostra foi inserida no FPLC e posteriormente lavou-se a coluna com tampão 1 até que todas as proteínas não adsorvidas fossem retiradas da coluna. Após a lavagem, a eluição foi realizada em etapas, com aumento linear da concentração do tampão 2. O volume eluído foi coletado durante o experimento em frações de volume de 2 mL para análises posteriores.

3.1.2. Concentração

Após a purificação, a amostra foi concentrada a 4 °C utilizando o concentrador Amicon Ultra – 15 kDa (Millipore) a uma rotação de 4000 x g.

3.1.3. Segundo passo de purificação: Exclusão molecular

O pico correspondente foi agrupado, concentrado e submetido ao segundo passo cromatográfico.

Materiais:

- coluna: Superdex 75 16/60 (120 mL - Amershan Biosciences),
- tampão 1: bicarbonato de amônio (AMBIC) 50 mM pH 7,8,
- FPLC: GE Pharmacia AKTA Purifier,
- vazão: 0,4 mL/min.

Métodos:

A segunda purificação foi realizada em coluna comercial Superdex 75. A cromatografia foi realizada em um sistema FPLC de baixa pressão. A coluna foi equilibrada com o tampão 1, a uma vazão de 1 mL/min em temperatura ambiente. Em seguida a amostra foi inserida no FPLC, a uma vazão de 0,4 mL/min. O volume eluído foi coletado durante o experimento em frações de volume de 2 mL para análises posteriores.

3.1.4. Terceiro passo de purificação: Afinidade

Da mesma forma, o pico correspondente foi agrupado, concentrado, lavado em tampão 3 e submetido ao terceiro passo cromatográfico.

Materiais:

- coluna: Heparin Sepharose (Amershan Biosciences),
- tampão 3: Tris 10 mM pH 7,0 + CaCl₂ 50 mM,
- tampão 4: Tris 10 mM pH 7,0 + NaCl 1 M,
- FPLC: GE Pharmacia AKTA Purifier,
- vazão: 0,7 mL/min.

Métodos:

A terceira purificação foi realizada em coluna comercial Heparin Sepharose. A cromatografia foi realizada em um sistema FPLC de baixa pressão. A coluna foi equilibrada com o tampão 3, a uma vazão de 1 mL/min em temperatura ambiente. Em seguida a

amostra foi inserida no FPLC e posteriormente lavou-se a coluna com tampão 3 até que todas as proteínas não adsorvidas fossem retiradas da coluna. Após a lavagem, a eluição foi realizada em etapas, com aumento linear da concentração do tampão 4. O volume eluído foi coletado durante o experimento em frações de volume de 2 mL para análises posteriores.

3.1.5. Quarto passo de purificação: Exclusão molecular

A amostra foi novamente concentrada e submetida ao último passo cromatográfico.

Materiais:

- coluna: Superdex 75 16/60 (120 mL - Amershan Biosciences),
- tampão 5: Tris 20 mM pH 7.5 + 50 mM NaCl
- FPLC: GE Pharmacia AKTA Purifier,
- vazão: 0,4 mL/min.

Métodos:

A última purificação foi realizada em coluna comercial Superdex 75. A cromatografia foi realizada em um sistema FPLC de baixa pressão. A coluna foi equilibrada com o tampão 5, a uma vazão de 1 mL/min em temperatura ambiente. Em seguida a amostra foi inserida no FPLC, a uma vazão de 0,4 mL/min. O volume eluído foi coletado durante o experimento em frações de volume de 2 mL para análises posteriores.

3.1.6. Gel de poliacrilamida desnaturante (*Sodium dodecyl sulfate polyacrylamide gel electrophoresis* - SDS-PAGE)

Nesse trabalho, usou-se o protocolo descrito por Laemmli (com modificações), U.K., 1970 para a confecção de SDS-PAGE 10%.

Materiais:

- aparato de eletroforese da Bio-Rad (cuba, fonte, placas de vidro),
- solução de gel separador 10%: 3,45 ml de 30% acrilamida/0,8% bis-acrilamida, 5 ml de Tris-HCl/SDS pH 8.8 (1,5 M Tris-Cl contendo 0,4% de SDS), 1,5 ml de água,

50 μ l de persulfato de amônio a 10% e 15 μ l de TEMED (N, N, N', N'-tetrametiletilenediamina)

- solução do gel concentrador 4%: 0,75 ml de 30% acrilamida/0,8% bis-acrilamida, 2,5 ml de Tris-Cl/SDS pH 6.8 (0,5 M Tris-Cl contendo 0,4% de SDS), 1,83 ml de água, 30 μ l de persulfato de amônio a 10% e 6 μ l de TEMED,
- tampão de amostra 3X: 2,4 ml de Tris-HCl pH 6.8, 3 mL de SDS 20 %, 3 ml de glicerol, 0,006 g de azul de bromofenol, 1,6 mL de β -mercaptoethanol. Adicionar água para um volume de 10 ml,
- tampão de corrida: 3,02 g de Tris, 18,6 g de Glicina, 10 ml de SDS 10%, completar para 1 L de água
- solução de *Comassie Blue* R-250: 0.25g coomassie brilliant blue R-250 em 100 ml de uma solução 45% metanol e 10% ácido acético,
- solução descorante: 45 % (v/v) de etanol, 10 % (v/v) ácido acético.

Métodos:

A corrida de eletroforese foi realizada por aproximadamente duas horas, à temperatura ambiente, com voltagem de 120 V. Após a corrida, o gel de poli-acrilamida foi colocado em uma solução de *coomassie-blue* R-250 para coloração por 2 horas à temperatura ambiente. A remoção do excesso de corante foi obtida em solução descorante até o aparecimento nítido das bandas de proteínas.

3.1.7. Quantificação

A concentração da proteína foi determinada através da medida de absorvância da mesma a 280 nm. Os espectros de absorvância foram obtidos em um espectrofotômetro *Jasco V-530*, com varredura de comprimento de onda de 340-200 nm, utilizando-se uma cubeta de quartzo de 1 cm de caminho ótico. O espectro de absorvância da proteína foi corrigido com a subtração do espectro de absorvância do tampão correspondente (“branco”) contendo Tris 20 mM pH 7,5 + 50 mM de NaCl. A partir dos valores dos coeficientes de extinção molar (ϵ) obtidos através do programa *ProtParam* (<http://ca.expasy.org>) e dos valores de absorvância a 280 nm, determinou-se a concentração molar da proteína através da fórmula:

$$A^{280} = \varepsilon \cdot L \cdot C$$

onde A^{280} é a absorvância medida a 280 nm, ε é o coeficiente de extinção molar em 280 nm ($M^{-1}cm^{-1}$), L é o caminho ótico em centímetros (cm) e C é a concentração molar da amostra de proteína. A amostra de proteína foi centrifugada a 5000 x g por 10 minutos antes de ser utilizada para medir a concentração.

3.1.8. Sequenciamento da proteína utilizando espectrometria de massa por ionização de eletrospray (ESI-MS/MS)

A espectrometria de massas (MS) utiliza o movimento de íons em campos elétricos e magnéticos para classificá-los de acordo com sua relação massa-carga. Desta maneira, a espectrometria de massas é uma técnica analítica por meio da qual as substâncias químicas, peptídeos se identificam, separando os íons gasosos em campos elétricos e magnéticos. Assim, a confirmação da seqüência da proteína foi obtida por espectrometria de massa.

Métodos:

A proteína BmooMP α -I presente nas bandas do gel de poliacrilamida coradas com coomassie, foram cortadas, extraídas do gel e digeridas com as proteases: tripsina, Glu-C (SV-8) e quimotripsina, de acordo com protocolo de Hellman et al., 1995. Em seguida, a amostra foi submetida a uma análise em um espectrômetro de massa por ionização de eletrospray (ESI), utilizando um refletor quadrupolo *time-of-flight* (Q-Tof Ultima – Waters/Micromass) (Figura 14) acoplado a um sistema de Ultra Performance Liquid Chromatography (UPLC) (NanoAcquity – Waters) em uma coluna Waters Symmetry C18 (5 μ m x 180 μ m x 20 mm) (Figura). A eluição foi realizada em coluna BEH 130 C18 (1,7 μ m x 75 μ m x 100 mm) utilizando um gradiente de 0-60 % (v/v) de acetonitrila por 1 hora. O espectro foi adquirido utilizando o programa Data Directed Analysis e selecionando peptídeos duplamente e triplamente carregados para experimentos de MS/MS. Todos os espectros MS/MS foram processados utilizando o software Mascot Distiller e a ferramenta de busca MASCOT.

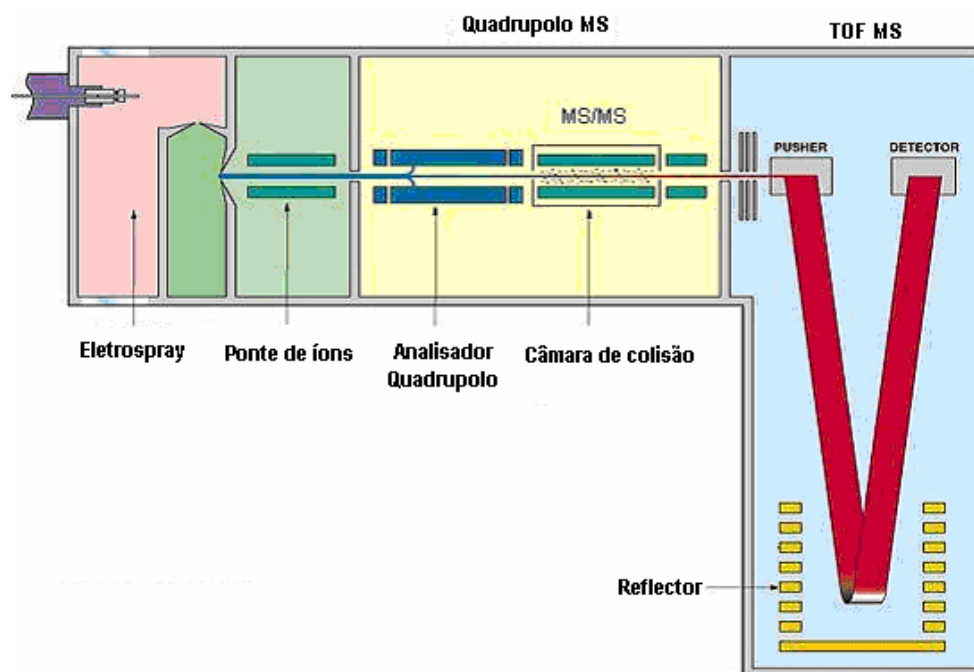


Figura 13. Esquema do espectrômetro de massa quadrupolo *time-of-flight*.

3.1.9. Espalhamento Dinâmico de Luz (*Dynamic Light Scattering - DLS*)

O espalhamento dinâmico de luz (DLS) é uma técnica que mede o seu raio hidrodinâmico, o qual é geralmente calculado baseado em modelos esféricos e permite verificar a homogeneidade da amostra. A maior vantagem do espalhamento de luz, além de fornecer informações estruturais, é o curto tempo necessário para a obtenção dos dados, a modesta quantidade de proteína utilizada (de 1 a 10 mg/ml), os pequenos volumes de amostra requeridos (30 μ l), e a relativa simplicidade de manipulação.

Métodos:

As medidas de DLS foram realizadas em cubeta de quartzo de 1 cm de caminho óptico a 20 °C utilizando-se o equipamento Dynapro 810 (Protein Solutions), portando um estabilizador de temperatura. O programa usado para a execução das medidas do raio hidrodinâmico foi o Dynamics V 6.3.4. Uma curva padrão de albumina do serum bovino foi utilizada para a calibração e os experimentos foram realizados a 291 K.

3.1.10. Ensaios de cristalização

Os ensaios iniciais de cristalização foram realizados com a proteína BmooMP α -I a uma concentração de 10 mg/mL em tampão Tris 20 mM pH 7,4. A técnica utilizada foi difusão de vapor em gota sentada utilizando o robô HoneyBee 963 (Genomic Solutions) a temperatura de 291 K. As gotas foram formadas por 0,5 μ L de proteína com o mesmo volume de solução mãe e 80 μ L de solução teste na placa de cristalização. Os testes iniciais foram realizados pelo método da matriz esparsa, com algumas modificações, utilizando os kits comerciais: Crystal Screen, Crystal Screen 2 (Hampton Research), Wizard Screens I e II (Emerald BioStructures) e Personal Structure Screens 1 e 2 (Molecular Dimensions).

3.1.11. Coleta de dados

Após o crescimento dos cristais alguns foram escolhidos e colhidos, com o auxílio de loops de nylon (*Hampton Research*), esses cristais da BmooMP α -I foram submetidos a uma crioproteção com 20 % (v/v) de glicerol e em seguida foi rapidamente congelado e mantido sob um fluxo constante de nitrogênio gasoso a temperatura de 100 K. O fluxo de nitrogênio gasoso ajuda a minimizar os danos causados pela radiação ionizante, e a solução crioprotetora serve para impedir a formação de cristais de gelo durante o resfriamento do cristal.

Os dados de difração de raios X foram coletados utilizando radiação síncrotron, na linha DB03-MX1. O comprimento de onda da fonte de radiação foi definido em 1,428 Å e o detector utilizado foi o MARCCD (165 mm). Foram coletadas 210 imagens a uma distância cristal-detector de 80 mm, cada imagem de difração foi coletada em um intervalo de 1° com tempo de exposição de 30 segundos. Os dados foram indexados, integrados, reduzidos e escalonados utilizando os programas DENZO e SCALEPACK, implementados no pacote HKL2000 (Otwinowski e Minor, 1997).

3.1.12. Determinação da estrutura: Substituição Molecular

O método de Substituição Molecular é aplicado quando a proteína estudada apresenta alta homologia com outras proteínas cuja estrutura já tenha sido determinada. Este procedimento é baseado na observação de que macromoléculas homólogas na sua seqüência possuem um enovelamento muito similar. Assim a posição do modelo molecular dentro do cristal é determinada pela rotação R e pela translação T que move o modelo de uma posição inicial de referência, especificada pelos vetores atômicos (r_o), para a posição atual, especificada pelos vetores atômicos (r):

$$r = Rr^o + T$$

A estrutura da BmooMP α -I foi resolvida por técnicas de substituição molecular com o Programa AMoRe (Navaza, 1994), utilizando como modelo de busca as coordenadas atômicas da acutolisina A, uma metaloproteinase zinco dependente isolada do veneno de *Agkistrodon acutus* (código do PDB: 1BSW, Gong et al., 1998).

3.1.13. Refinamento e análise da estrutura

Fatores térmicos isotrópicos posicionais e individuais foram refinados utilizando o programa REFMAC5 (Murshudov et al., 1997) incorporado no pacote CCP4 (CCP4, 1994). Os mapas de densidade eletrônica $2F_o - F_c$ e $F_o - F_c$ da proteína modelo foram examinados e ajustados manualmente após cada ciclo de refinamento utilizando o programa COOT (Emsley and Cowtan, 2004). As moléculas de água foram incorporadas nos ciclos finais utilizando o programa ARP/wARP (Perrakis et al., 1999). Nessa fase uma diferença no mapa de densidade eletrônica indicou a presença de dois fortes picos, que correspondem aos íons zinco, no sítio ativo, e o íon cálcio localizado na região de cruzamento do N- e C-terminais. A qualidade estereoquímica global do modelo final foi observada pelo programa PROCHECK (Laskowski et al., 1993).

3.1.14. Docking molecular com ligante flexível

Os estudos de *docking* molecular foram realizados utilizando um receptor rígido com ligantes flexíveis (Figura 14), para tanto fizemos uso do software ICM-Pro (Molsoft, La Jolla, Ca) que utiliza uma otimização global da função de energia, baseada na minimização de Monte-Carlo (Abagyan et al., 1994). Para cada ligante uma grande quantidade de modelos acoplados foram gerados, possuindo conformações aleatórias dos ligantes. No caso, foram utilizados 7 ligantes, que foram complexados com peptidases zinco-dependentes e que tiveram suas estruturas cristalográficas determinadas (Códigos do PDB: 1ZXC, 2A8H, 2FV5, 2FV9, 1KUI, 1 KUK e 1KUG). Esses ligantes foram submetidos a cálculos de *docking* molecular para que seus modos de ligação com a BmooMP α -I pudessem ser caracterizados.

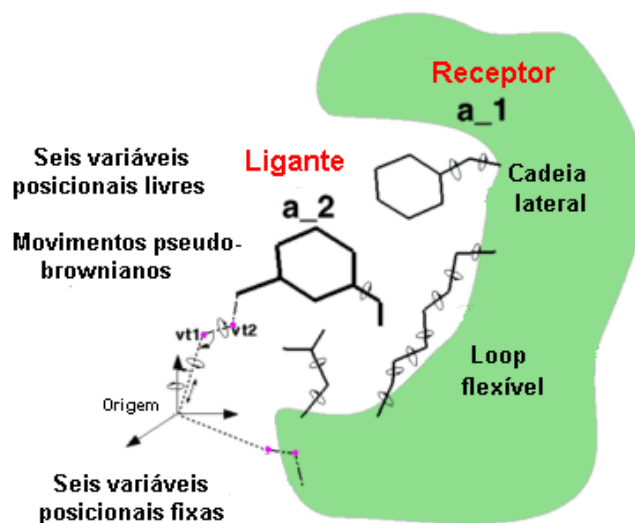


Figura 14. Modelo do sistema utilizado no *docking* molecular (<https://www.molsoft.com/gui/docking.html>).

As ferramentas do pacote ICM foram utilizadas para converter as estruturas dos ligantes e da proteína em um formato adequado para os experimentos *in silico*. Para a realização dos cálculos foi necessário a remoção das moléculas de água da proteína e assim átomos de hidrogênio foram adicionados e as cargas de Gasteiger foram calculadas para cada átomo da proteína. Adicionalmente, um mapa eletrostático foi calculado. Cada resultado foi analisado com base na energia de acoplamento, na estimativa individual das ligações de hidrogênio entre o ligante e os resíduos de aminoácidos da proteína e as interações com o íon zinco catalítico.

3.2 Endo-1,4- β -glicanase

3.2.1. Preparação de células competentes (Hanahan, 1983)

Materiais:

- tubos de cultivo estéreis com capacidade para 15 ml,
- meio de cultivo Luria Broth (LB): 1,0 % triptona, 1 % de extrato de levedura e 0,5 % de NaCl,
- incubadora com temperatura (37 °C) e agitação controladas,
- banho de gelo,
- espectrofotômetro para leitura da densidade óptica (D.O.) em 600 nm,
- centrífuga *Sorvall* modelo RC 26 Plus,
- solução 1: 100 mM de MgCl₂ estéril,
- solução 2: 100 mM de CaCl₂ estéril,
- solução 3: 15% de glicerol, 100% e 85% de 100 mM de CaCl₂,
- tubos estéreis com capacidade de 1,7 ml,
- nitrogênio líquido,
- freezer - 80°C

Métodos:

Transferir 100 μ l de um estoque de célula competente BL21 (DE3) para um tubo de cultura contendo 3 ml de meio LB. Essa cultura foi mantida por 16 horas em incubadora. A seguir transferiu-se esse conteúdo para um frasco contendo 200 ml de meio LB que foi incubado até atingir um valor de densidade ótica - D.O._(600 nm) em torno de 0,3. Nesse instante, interrompeu-se o crescimento celular colocando o frasco em banho de gelo. Centrifugou a cultura a 4000 x g por 15 minutos à 4 °C para decantar as células. Descartou-se o sobrenadante e as células foram ressuspensas em 50 ml da solução 1, onde permaneceu no gelo por 5 minutos e, a seguir, centrifugou-se nas mesmas condições. As células foram resolubilizadas em 10 ml da solução 2, em banho de gelo, e permaneceu por 20 minutos. Centrifugou-se nas mesmas condições. Novamente o sobrenadante foi descartado e ressuspensado em 1,5 ml da solução 3. Aliquotou 100 μ l em tubos com capacidade de 1 ml e congelou em nitrogênio líquido. Estocar em freezer – 80 °C.

3.2.2. Transformação por eletroporação

Materiais:

- DNA plasmidial,
- linhagem BL21 (DE3) competentes de *Escherichia coli*,
- banho de gelo,
- eletroporador – *GenePulser* (Biorad) – 2,5 mV e 10 F
- incubadora com temperatura (37 °C) e agitação controladas,
- centrífuga *eppendorf* modelo 5415D,
- placas de petri estéreis,
- LB-ágar (1 % triptona, 0,5 % extrato de levedura, 1 % NaCl, pH 7,5 e 1,8 % ágar p/v,
- antibióticos: Canamicina na concentração de 50 µg/ml,
- estufa de incubação com temperatura controlada à 37°C,

Métodos:

O vetor contendo o gene codificante foi utilizado para transformar, por eletroporação, linhagens BL21 (DE3) competentes de *Escherichia coli* com o fim de expressar a proteína. Para tanto, 1 µl do DNA plasmidial foi incubado com 40 µl de bactérias competentes e inseridos na cubeta apropriada sem formar bolhas, em gelo. O eletroporador foi ajustado nas condições indicada, a cubeta foi colocada no eletroporador e o pulso foi dado. Adicionou-se 1 mL de LB na cubeta e o meio foi transferido para um eppendorf e incubado por 1 h, à 37°C. A seguir, centrifugou-se a amostra por um minuto a 4000 x g com a finalidade de precipitar as células bacterianas, descartou-se o sobrenadante deixando apenas 100 µl para ressuspender as células. Esse material foi plaqueado em meio LB-ágar contendo canamicina e incubado em estufa por 12 horas.

3.2.3. Teste de expressão da celulase

Essa etapa foi feita com o objetivo de verificar a melhor temperatura de expressão da celulase por análise em SDS-PAGE 10% da fração não-induzida e induzida com

isopropil-tio-β-galactosídeo (IPTG) e para analisar a quantidade de proteína na fração solúvel e insolúvel.

Materiais:

- alça estéril,
- tubos de cultivo estéreis,
- meio de cultivo LB,
- antibióticos: canamicina na concentração de 50 µg/mL,
- incubadora com temperatura (37 °C) e agitação controladas,
- espectrofotômetro para leitura da densidade óptica (D.O.) em 600nm,
- solução estoque de IPTG (isopropil-tio-β-galactosídeo – Invitrogen) a 1 M.

Métodos:

Uma colônia da placa transformada foi transferida, com o auxílio de uma alça, devidamente estéril, para um inóculo de 5 ml de LB contendo canamicina. O crescimento foi feito em incubadora por um período de 12 horas. Posteriormente, uma alíquota de 500 µl desse inóculo foi usado para inocular 50 ml de meio LB, contendo canamicina e deixado em incubadora até que se atingisse um valor de D.O. em torno de 0,6-0,8. Nesse instante, adicionou-se IPTG a uma concentração final de 1 mM e incubou-se a cultura por duas e quatro horas, à 37 °C, 30°C e 25°C sob agitação, 200 rpm.

3.2.4. Maxi-Expressão da celulase

Materiais:

- meio de cultivo LB,
- tubos de cultivo estéreis,
- antibióticos: canamicina 50 µg/ml,
- incubadora com temperatura (37 °C) e agitação controladas,
- centrifuga Sorvall modelo RC 26 Plus,
- solução estoque de IPTG (isopropil-tio-β-galactosídeo – Invitrogen) a 1 M,
- tampão A: fosfato 50 mM pH 7,2 + 100 mM de NaCl + 5 mM de benzamidina,
- freezer - 20°C

Métodos:

Uma colônia da placa de petri obtida da transformação foi usada para inocular um meio de 10 ml de LB, contendo canamicina, para ser incubada por 12 horas. Esse meio foi usado para inocular a maxi-expressão que se deu em 1 L de LB contendo o mesmo antibiótico. A cultura foi mantida na incubadora por aproximadamente 3 horas até que a D.O. ($_{600\text{nm}}$) atingisse valor próximo de 0,6-0,8 seguindo-se, então, a adição de 1 mM de IPTG. Todos os parâmetros foram mantidos iguais para garantir que a proteína alvo fosse expressa nas mesmas condições.

Após a indução por quatro horas, a cultura foi centrifugada a 4000 x g, por quinze minutos a 4 °C. Descartou-se o sobrenadante e a fração celular foi ressuspensa em 30 ml tampão fosfato e transferida para um tubo de centrifugação. Foi feita uma nova centrifugação a 4000 x g, por vinte minutos para descartar o sobrenadante e congelar a fração celular.

3.2.5. Extração da proteína da fração celular (lise)

Materiais:

- tampão A: fosfato 50 mM pH 7,2 + 100 mM de NaCl + 5 mM de benzamidina,
- tubos com capacidade de 50 ml,
- sonicador: VirSonic
- banho de gelo
- centrifuga Sorvall modelo RC 26 Plus,
- lisozima 10 mg/mL
- PMSF (fluoreto de fenilmetil sulfonil) 0,1 M

Métodos:

O primeiro passo foi ressuspender a fração celular em 20 ml de tampão A e transferir para um tubo com capacidade de 50 mL. Logo após foi adicionado 0,023 mg/mL de lisozima e $2,3 \times 10^{-4}$ M de PMSF e foram incubados em gelo por 30 minutos. A seguir as células foram sonicadas em aparelho VirSonic no gelo em 15 ciclos de 10 segundos com intervalos de 30 segundos. Centrifugou-se a amostra por 30 minutos a 10000 x g à 4°C. O pellet foi descartado e a fração celular solúvel foi conduzida para os passos de purificação.

3.2.6. Gel de poliacrilamida desnaturante (*Sodium dodecyl sulfate polyacrylamide gel electrophoresis* - SDS-PAGE)

Os géis de poliacrilamida foram feitos como descrito anteriormente na seção 3.1.6.

3.2.7. Primeiro passo de purificação: Afinidade

Após a sonicação e sua respectiva centrifugação foi realizado o primeiro passo de purificação, que consistiu em uma cromatografia por afinidade, utilizando resina Ni²⁺-NTA.

Materiais:

- coluna: Ni-NTA (Hitrap Chelating HP, 5 mL - Amershan Biosciences)
- FPLC: GE Pharmacia AKTA Purifier
- tampão A: fosfato 50 mM pH 7,2 + 100 mM de NaCl + 5 mM benzamidina,
- tampão B: fosfato 50 mM pH 7,2 + 100 mM de NaCl + 500 mM de imidazol,
- vazão de injeção da amostra: 0,5 mL/min
- vazão de lavagem e eluição: 1 mL/min

Método:

Esse passo de purificação foi realizado com a coluna comercial Ni-NTA que consiste em uma coluna contendo 5 ml de agarose com níquel imobilizado. A cromatografia foi realizada em um sistema FPLC de baixa pressão. A coluna foi equilibrada com o tampão A, a uma vazão de 1 mL/min em temperatura ambiente. Em seguida a amostra foi inserido no FPLC a uma vazão de 0,5 mL/min, e posteriormente lavou-se a coluna com tampão A, na vazão de 1 mL/min, até que todas as proteínas não adsorvidas fossem retiradas da coluna. Após a lavagem, a eluição foi realizada em etapas, com aumento linear da concentração de imidazol no tampão de lavagem (tampão B). O volume eluído foi coletado durante o experimento em frações de volume de 2 mL para análises posteriores.

3.2.8. Diálise

Após o primeiro passo cromatográfico foi verificado que a amostra ainda apresentava contaminantes, necessitando assim de mais um passo cromatográfico, no caso uma troca catiônica.

Materiais:

- membrana de diálise: MWCO 12000 – 14000 Spectra/Por* 2 - Spectrum
- tampão C: fosfato 50 mM pH 7,2 + 5 mM benzamidina,

Método:

A solução foi colocada em membrana de diálise para a retirada do sal e do imidazol, sendo mantida sempre a 4 °C. A proporção do tampão C para a amostra foi de 1 para 10, ou seja, 1 mL da amostra para 10 mL de tampão C.

3.2.9. Segundo passo de purificação: Troca catiônica

Após a diálise a amostra foi submetida ao segundo passo de purificação que consistiu em uma troca catiônica, utilizando resina Hitrap HP SP.

Materiais:

- coluna: Hitrap SP HP (5 mL - Amershan Biosciences)
- FPLC: GE Pharmacia AKTA Purifier
- tampão C: fosfato 50 mM pH 7,2 + 5 mM benzamidina,
- tampão D: fosfato 50 mM pH 7,2 + 1 M de NaCl,
- vazão: 1 mL/min

Métodos:

A troca catiônica foi realizada com a coluna comercial Hitrap SP HP que consiste em uma coluna contendo 5 ml de agarose. A cromatografia foi realizada em um sistema FPLC de baixa pressão. A coluna foi equilibrada com o tampão C, a uma vazão de 1 mL/min em temperatura ambiente. Em seguida a amostra foi inserida no FPLC, e

posteriormente lavou-se a coluna com tampão C, na vazão de 1 mL/min, até que todas as proteínas não adsorvidas fossem retiradas da coluna. Após a lavagem, a eluição foi realizada em etapas, com aumento linear da concentração de NaCl no tampão de lavagem (tampão D). O volume eluído foi coletado durante o experimento em frações de volume de 2 mL para análises posteriores.

3.2.10. Concentração

Métodos:

Após a diálise, a amostra foi concentrada a 4 °C utilizando o concentrador Ultrafree CL – 30 kDa (Millipore) a uma rotação de 4000 x g, até atingir um volume final de aproximadamente 1 mL, para que pudesse ser realizado o terceiro passo cromatográfico e os estudos biofísicos.

3.2.11. Terceiro passo de purificação: Exclusão Molecular

Dilui-se 200 µL da amostra concentrada em 600 µL de tampão, em seguida essa solução foi submetida à exclusão molecular.

Materiais:

- coluna: Superdex 75 16/60 (120 mL - Amershan Biosciences)
- FPLC: GE Pharmacia AKTA Purifier
- tampão E: fosfato 50 mM pH 6,0 + 250 mM de NaCl + 5 mM benzamidina,
- vazão: 0,3 mL/min

Métodos:

A exclusão molecular foi realizada com a coluna Superdex 75 16/60 que consiste em uma coluna contendo 120 ml de agarose e dextran. A cromatografia foi realizada em um sistema FPLC de baixa pressão. A coluna foi equilibrada com o tampão E, a uma vazão de 1 mL/min em temperatura ambiente. Em seguida a amostra foi inserida no FPLC, a uma vazão de 0,3 mL/min. O volume eluído foi coletado durante o experimento em frações de volume de 2 mL para análises posteriores.

3.2.12. Quantificação

A quantificação foi realizada como descrito anteriormente na seção 3.1.7.

3.2.13. Dicroísmo circular

Os experimentos de dicroísmo circular foram realizados no espectropolarímetro *CD* Jasco J-810, com um controlador de temperatura Peltier PFD-425S. O programa utilizado para registro dos dados foi o Spectra manager[®] (JASCO). Os espectros de dicroísmo circular no UV distante (260-200 nm) foram adquiridos utilizando-se uma cubeta de quartzo de 2 mm de caminho ótico, em temperatura ambiente, com velocidade de varredura de 50 nm/minuto, tempo de resposta de 4 segundos. O espectro final foi obtido pela acumulação de 10 varreduras. A contribuição do solvente foi eliminada subtraindo seu espectro daquele da proteína. A proteína foi utilizada na concentração de 8 μM para as desnaturações químicas e 3 μM para as desnaturações térmicas. Todas as amostras foram centrifugadas a 14000 x g por 10 minutos antes das medidas.

Os valores obtidos na leitura de CD (mdeg) foram convertidos em elipticidade molar residual θ através da expressão (ADLER et al., 1973):

$$\theta = \frac{\text{CD (mdeg)}}{\text{L(mm) C(M) aa}}$$

onde CD é a elipticidade medida (mdeg), L é o caminho ótico em milímetros (mm), C é a concentração molar (M) e aa é o número de aminoácidos da proteína. Os dados convertidos foram graficados utilizando-se o programa ORIGIN 6.2.

3.2.14. Cristalização

Os ensaios iniciais de cristalização foram realizados com a endo-1,4-β-glicanase a uma concentração de 69,5 mg/mL em tampão MES 50 mM pH 6,0. A técnica utilizada foi difusão de vapor em gota sentada utilizando o robô *HoneyBee* 963 (Genomic Solutions) a temperatura de 291 K. As gotas foram formadas por 0,5 μL de proteína com o mesmo volume de solução mãe e 80 μL de solução teste na placa de cristalização. Os testes iniciais

foram realizados pelo método da matriz esparsa, com algumas modificações, utilizando os kits comerciais: Crystal Screen, Crystal Screen 2 (Hampton Research), Wizard Screens I e II (Emerald BioStructures) e Personal Structure Screens 1 e 2 (Molecular Dimensions).

CAPÍTULO 4 – RESULTADOS E DISCUSSÃO

Esse capítulo tem como objetivo apresentar os resultados dos experimentos realizados tanto da proteína BmooMP α -I de *Bothrops moojeni* quanto da proteína endo-1,4- β -glicanase. Na primeira parte da apresentação dos resultados, destaca-se a resolução da estrutura e análises de *docking* molecular da proteína BmooMP α -I. Os resultados das purificações não serão apresentados neste trabalho, mas podem ser encontrados em Bernardes et al., 2008. Na seção seguinte do capítulo, descrevemos os experimentos realizados para a obtenção da proteína endo-1,4- β -glicanase em altas quantidades e pureza para a realização de estudos estruturais. Para a complementação dos experimentos, várias consultas bibliográficas foram feitas, paralelamente, com o propósito de enriquecer a discussão dos resultados. Essa etapa, na execução do projeto, foi extremamente importante para o entendimento do processo e conclusão dos experimentos.

4.1 BmooMP α -I

4.1.1. Identificação da proteína por espectrometria de massa

A identificação da proteína foi realizada pela busca em um banco de dados não-redundante (NCBI) utilizando a ferramenta de busca MASCOT (<http://matrixscience.com>). Foram utilizados dois parâmetros para essa busca: uma clivagem com possíveis modificações de carbamidometilação da cisteína (completa) e oxidação da metionina (parcial). BmooMP α -I foi digerida pelas enzimas: tripsina, quimotripsina e SV-8 e claramente identificada com 81% do total de peptídeos digeridos compatíveis. A probabilidade dessa compatibilidade ter ocorrido randomicamente é menos que 0,5 %.

4.1.2. Homogeneidade química e estrutural

Miligramas da proteína BmooMP α -I foi obtida pelos três passos cromatográficos descritos anteriormente. Em seguida sua pureza foi analisada por uma eletroforese em gel de poliacrilamida sob condições desnaturantes, onde se pode observar uma única banda com peso molecular em torno de 22 kDa (Figura 15).

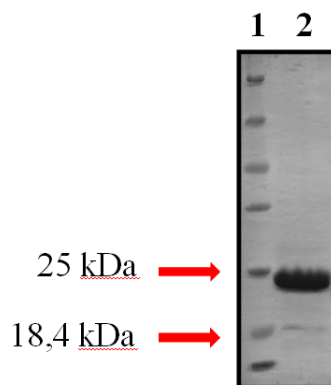


Figura 15. SDS-PAGE sob condições desnaturantes da BmooMP α -I purificada. O gel foi corado com coomassie brilliant blue R-250. A linha 1 mostra o marcador de peso molecular (β -galactosidase 116 kDa, bovine serum albumina 66,2 kDa, ovalbumina 45 kDa, lactate dehydrogenase 35 kDa, REase Bsp98I 25 kDa, β -lactoglobulina 18,4 kDa, lisozima 14,4 kDa) e a linha 2 mostra a BmooMP α -I com alto grau de pureza.

Para checar a homogeneidade estrutural foram realizados experimentos de DLS com a mesma concentração de proteína utilizada nos ensaios cristalográficos (10 mg/mL). O DLS é uma técnica muito usada como diagnóstico para determinar se uma amostra de proteína é adequada para cristalização. O DLS fornece informações sobre a distribuição de tamanhos e estado de

agregação da proteína em solução. Essas informações são importantes, pois estão intimamente relacionadas ao processo de cristalização, já que quanto mais homogênea (monodispersa) for a amostra, maior é a probabilidade de nucleação e crescimento dos cristais.

Esse experimento mostrou que a proteína está, majoritariamente em um estado monodisperso (monômero em solução), com um índice de polidispersividade em torno de 5%. Ambos os parâmetros analisados nessa etapa indicaram um estado favorável da proteína em solução para a realização dos ensaios cristalográficos e posteriores otimizações.

4.1.3. Aquisição dos dados e determinação da estrutura

Duas das condições de cristalização forneceram cristais pequenos e alongados, portanto essas condições foram selecionadas para uma etapa de refinamento. As soluções utilizadas na etapa de refinamento foram:

- tampão Tris 100 mM + 15 % (v/v) etanol pH 7,0
- tampão Hepes 100 mM + 20 % (v/v) polietilenoglicol (PEG) 8000 pH 7,5

A partir dessas condições de cristalização foi elaborada uma matriz de otimização, na tentativa de melhorar os cristais, que consistiu na variação do pH de 0,1 em 0,1 em um intervalo de pH 4,0 até 9,0, e na mistura do tampão e das substâncias volátil e não-volátil. Depois de 7-9 dias, cristais retangulares com dimensão aproximada de 0,4 x 0,3 x 0,2 mm foram obtidos da condição: tampão Tris 100 mM + 20 % (v/v) PEG 8000 pH 8,3, a 20°C (Figura 16).

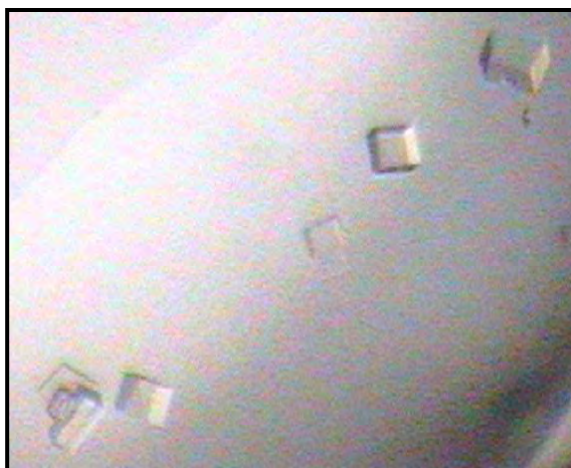


Figura 16. Fotomicrografia do cristal da BmooMP α -I

Os cristais da BmooMP α -I pertencem ao grupo espacial monoclinico P2₁, com parâmetros de dimensão celular $a = 40,1$, $b = 48,7$, $c = 54,4$ e $\beta = 91,56^\circ$. A estrutura foi determinada a uma

resolução de 1,76 Å, e o problema de fases foi resolvido utilizando a técnica de substituição molecular, onde foi utilizada a estrutura da acutolisina A como modelo de busca.

Densidades eletrônicas residuais acima de 3σ , na diferença do mapa eletrônico de Fourier, foram encontradas no bolsão do sítio ativo e na região de cruzamento do N- e C- terminais, as quais foram caracterizadas como íons zinco e cálcio, respectivamente.

Com as etapas de refinamento o residual cristalográfico convergiu para 17,5 % (R_{free} 21,6 %) para todos os dados entre 20,8 e 1,76 Å (Tabela 3). O modelo final consiste em 200 resíduos de aminoácidos, 151 moléculas de solvente, 1 íon zinco e 1 íon cálcio, apresentando uma excelente estereoquímica total, com valores de RMSD de 0,023 Å e 1,9° para o comprimento das ligações e ângulos, respectivamente. A média do fator de temperatura (*B-factor*) para todos os átomos é de 25,7 Å² (Tabela 3).

Os diagramas de Ramachandran demonstram que 98,5 % dos ângulos diedrais estão situados nas regiões mais favoráveis, os outros 1,5 % dos resíduos estão em regiões permitidas e nenhum resíduo foi encontrado em regiões não-permitidas (Figura 17).

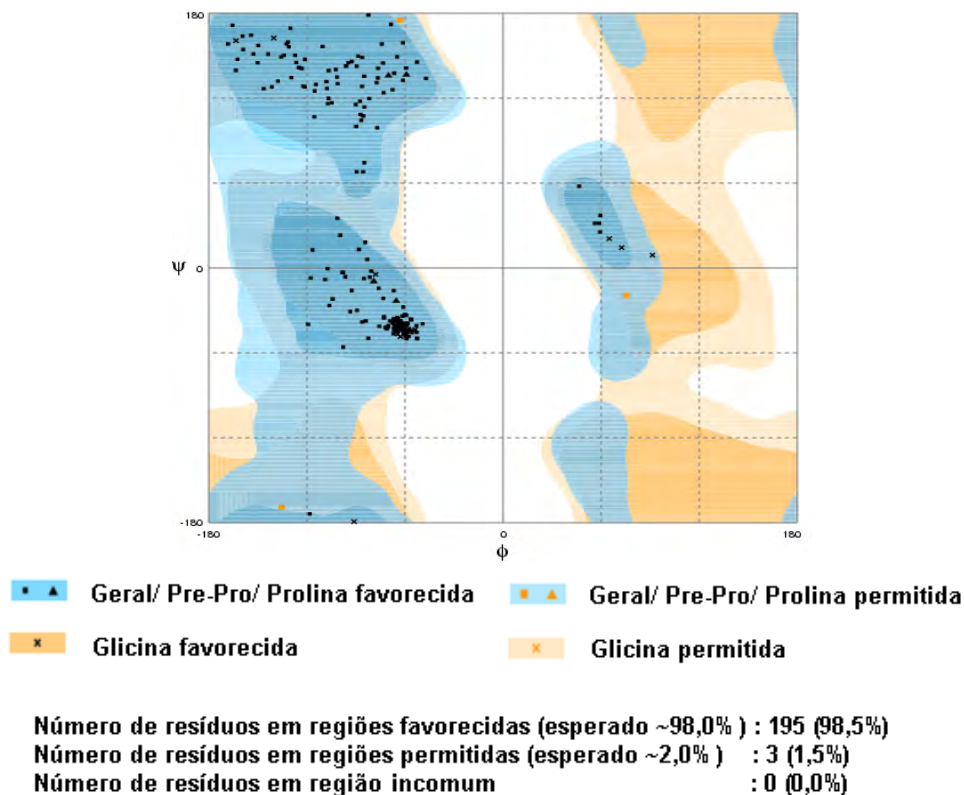


Figura 17. Diagrama de Ramachandran obtido na validação da estrutura da BmooMP α -I.

Tabela 3. Estatística do processamento de dados e refinamento cristalográfico.

<i>Configurações do feixe de luz</i>	
Fonte de radiação	Luz Síncrotron
Linha	DB03-MX1, LNLS-Brasil
Comprimento de onda (Å)	1,437
Temperatura (K)	100
Detector	MARCCD 165 mm
<i>Preparação do cristal</i>	
Solução crioprotetora	Solução mãe + 20 % glicerol
Tempo de <i>soaking</i> (s)	30
<i>Coleta de dados</i>	
Grupo espacial	P2 ₁
a (Å)	40,07
b (Å)	48,73
c (Å)	54,39
β (°)	91,56
Faixa de resolução (Å)	20,81-1,76 (1,83-1,76)
Reflexões únicas	19477
Multiplicidade ^a	5,88 (2,35)
Completeza (%)	93,7 (66,70)
I/ σ_1	14,7 (2,87)
R _{merge} (%)	5,4 (26,70)
Coefficiente de Matthews (Å ³ Da ⁻¹)	2,36
Conteúdo de solvente (%)	47,91
Números de moléculas na unidade assimétrica	1
<i>Refinamento do Modelo</i>	
Código PDB (Protein data Bank)	3GBO
Número de átomos na proteína (não hidrogênios)	1590
Número de moléculas de água	144
R _{model} (%)	17,18
R _{free} (%)	21,8
R.M.S.D. do comprimento de ligação de átomos ideais (Å)	0,023
R.M.S.D. dos ângulos de ligação de átomos ideais (°)	1,898
Média do fator-B (Å ²)	25,7
<i>Diagrama de Ramachandran</i>	
Regiões favoráveis (%)	98,5
Regiões permitidas (%)	1,5
Regiões não-permitidas (%)	0

4.1.4. Descrição da estrutura

A topologia molecular da BmooMP α -I conserva todas as principais características das SVMPs de classe PI, pois é formada por cinco folhas β , quatro longas hélices e uma hélice α curta na região N-terminal estabilizada por 3 ligações dissulfeto (Cys¹¹⁵–Cys¹⁹⁵, Cys¹⁵⁵–Cys¹⁷⁹ e Cys¹⁵⁷–Cys¹⁶²) localizadas na região C-terminal. A figura 18 apresenta o modelo cristalográfico da BmooMP α -I, onde podemos observar que a molécula possui uma estrutura elipsoidal, onde as folhas β são curvadas e encaixadas entre as hélices α 2 e α 4 no lado côncavo, e entre a α 3 no lado convexo. As folhas β 1 (resíduos 4 ao 12), β 2 (resíduos 51 ao 56), β 3 (resíduos 92 ao 97) e β 5 (resíduos 121 ao 125) são paralelas e a β 4 (resíduos 107 ao 109) é antiparalela, sendo que a β 1 é conectada a hélice α 2 pela hélice α 1.

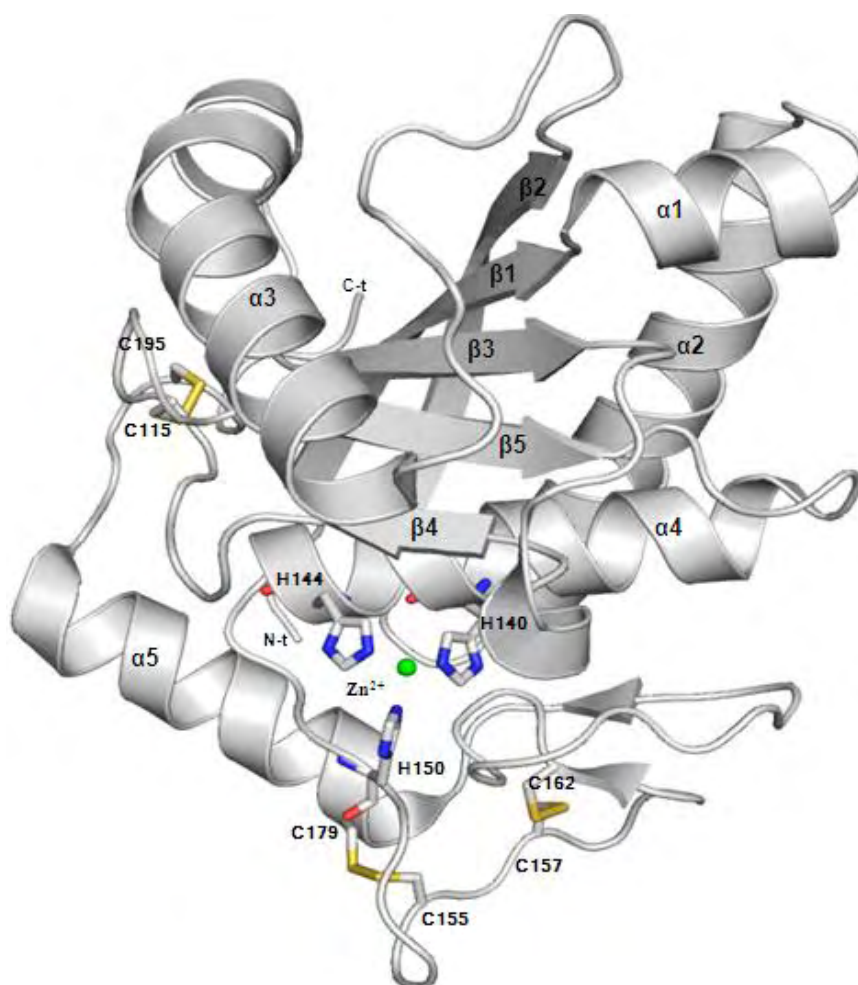


Figura 18. Representação esquemática da estrutura da BmooMP α -I. As ligações dissulfeto e as 3 histidinas catalíticas são representadas por bastões. O íon zinco é representado pela esfera verde. Carbono, enxofre e nitrogênio são representados pelas cores cinza, amarelo e azul, respectivamente. As letras e algarismos romanos representam as estruturas secundárias. As cores dos átomos estão padronizadas para todas as outras figuras. Figura feita no programa Pymol (<http://www.pymol.org/>).

A estrutura é dividida em dois subdomínios, o superior (resíduos 1 ao 146) e o inferior (resíduos 147 ao 200). O subdomínio inferior é formado por um complexo de vários loops (resíduos 147 ao 176) e a longa hélice α 5. Esse subdomínio abriga o *Met-turn* (Met¹⁶⁴) e juntamente com a hélice α 4 do domínio superior, formam o sítio ativo. A cavidade do sítio ativo é localizada na interface do domínio superior e o domínio inferior como observado em outras metzincinas, como a adamalisina-II (Gomis-Rüth et al., 1994), H₂-proteinase (Kumasaka et al., 1996) e acutolisina A (Gong et al., 1998).

4.1.5. Sítio de ligação ao zinco

As posições dos resíduos catalíticos importantes no sítio ativo são completamente conservadas incluindo o *Met-turn* e o motivo de ligação do zinco H₁₄₀E₁₄₁XXH₁₄₄XXG₁₄₇XXH₁₅₀

A figura 19 mostra o ambiente de ligação do zinco, que apresenta uma coordenação octaédrica distorcida formada por átomos de N ϵ 2 dos anéis imidazólicos das três histidinas catalíticas (His¹⁴⁰, His¹⁴⁴ e His¹⁵⁰) e adicionalmente três moléculas de água (OW¹³¹, OW¹⁴³ e OW¹⁴⁴).

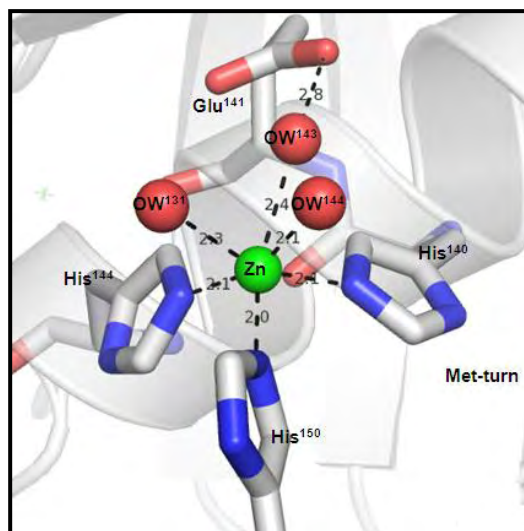


Figura 19. Representação do sítio de ligação ao íon zinco no sítio catalítico da BmooMP α -I. Figura feita no programa Pymol (<http://www.pymol.org/>).

Em outras SVMPs como, adamalisina-II, H₂-proteinase e acutolisina A, o íon zinco é coordenado tetraedricamente pelas mesmas três histidinas e por somente uma molécula de água (Figura 20B), que é estruturalmente relacionada com o resíduo Glu¹⁴¹, que polariza a molécula de água para realizar o ataque nucleofílico na ligação peptídica da cadeia do substrato.

A molécula de água OW¹⁴³ é coordenada por uma ligação curta (distância = 2,8 Å) quando comparada com outras metzincinas, nas quais essa distância varia entre 3,5 e 4,3 Å. Essa distancia entre a água OW¹⁴³ e o resíduo Glu parece crucial para a atividade proteolítica. Várias evidências estruturais sugerem que essa distância de contato e a atividade proteolítica está correlacionada com o pKa do grupo carboxílico do resíduo Glu, que é 4,3 (Gomis-Rüth et al., 1993). O grupo carboxilato desse Glu, sob condições ácidas, apresentaria pouca habilidade de polarizar a molécula de água, aumentando suas distâncias interatômicas e diminuindo a atividade proteolítica. Enquanto que em condições básicas, diminuiriam suas distâncias interatômicas, favorecendo o ataque nucleofílico (Zhu, et al., 1999). Assim, distâncias interatômicas maiores que 4 Å (faixas de pH ácidas) poderiam afetar a capacidade de polarização do resíduo Glu sobre a molécula de água. Isso explicaria o fato de que algumas SVMPs apresentam baixas atividades ou estão inativas a baixos valores de pH (Gong et al., 1998).

Outro modo de coordenação do zinco foi descrito na estrutura da astacina (Bode et al., 1992), onde a molécula de água foi substituída por um grupo hidroxila da cadeia lateral da Tyr¹⁶⁹ (Figura 20C), que possivelmente, auxilia na fixação do substrato no sítio ativo (Stöcker et al., 1995).

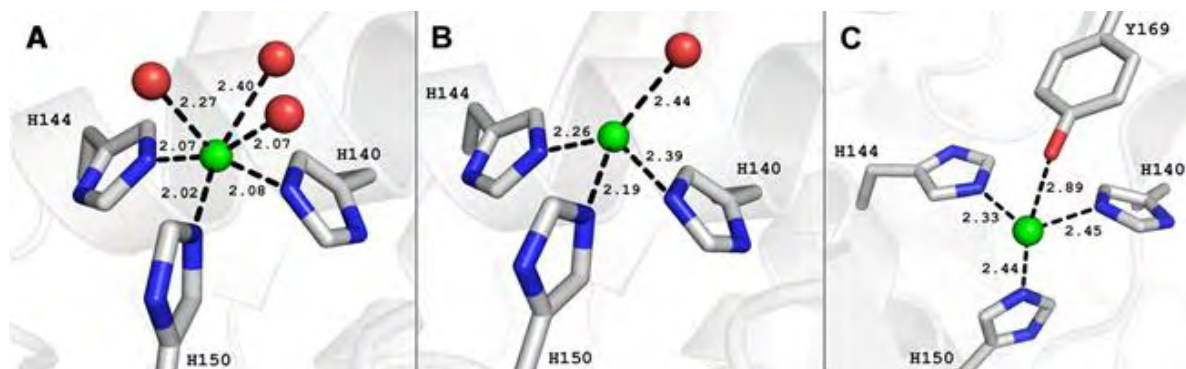


Figura 20. Comparação da geometria de coordenação do zinco da BmooMP α -I (A), adamalisina-II (B) e astacina (C). As interações são apresentadas em linhas pretas pontilhadas. Figura feita no programa Pymol (<http://www.pymol.org/>).

4.1.6. Sítio de ligação ao cálcio

Como na acutolisina A, na adamalisina-II e atrolisina-C o íon cálcio localiza-se na região de cruzamento do N- e C-terminal no lado oposto ao bolsão do sítio ativo. Como podemos observar da figura 21, o íon cálcio é heptacoordenado pelos átomos O δ 1Asn¹⁹⁸, OCys¹⁹⁵, O δ 1Asp⁹¹, O δ 2Asp⁹¹, O ϵ 1Glu⁹, OW⁵ e OW¹⁰, e têm sido considerado muito importante

estruturalmente, pois possui um papel estabilizador em estruturas de SVMPs de classes PII, PIII e PIV.

Devido a sua localização na porção C-terminal, o íon não só manteria a rigidez e estabilidade da região, mas também desempenha um papel de “fecho”, que conectaria o domínio metaloproteinase com o domínio desintegrina. Porém, nenhuma evidência foi encontrada para seu papel na atividade enzimática.

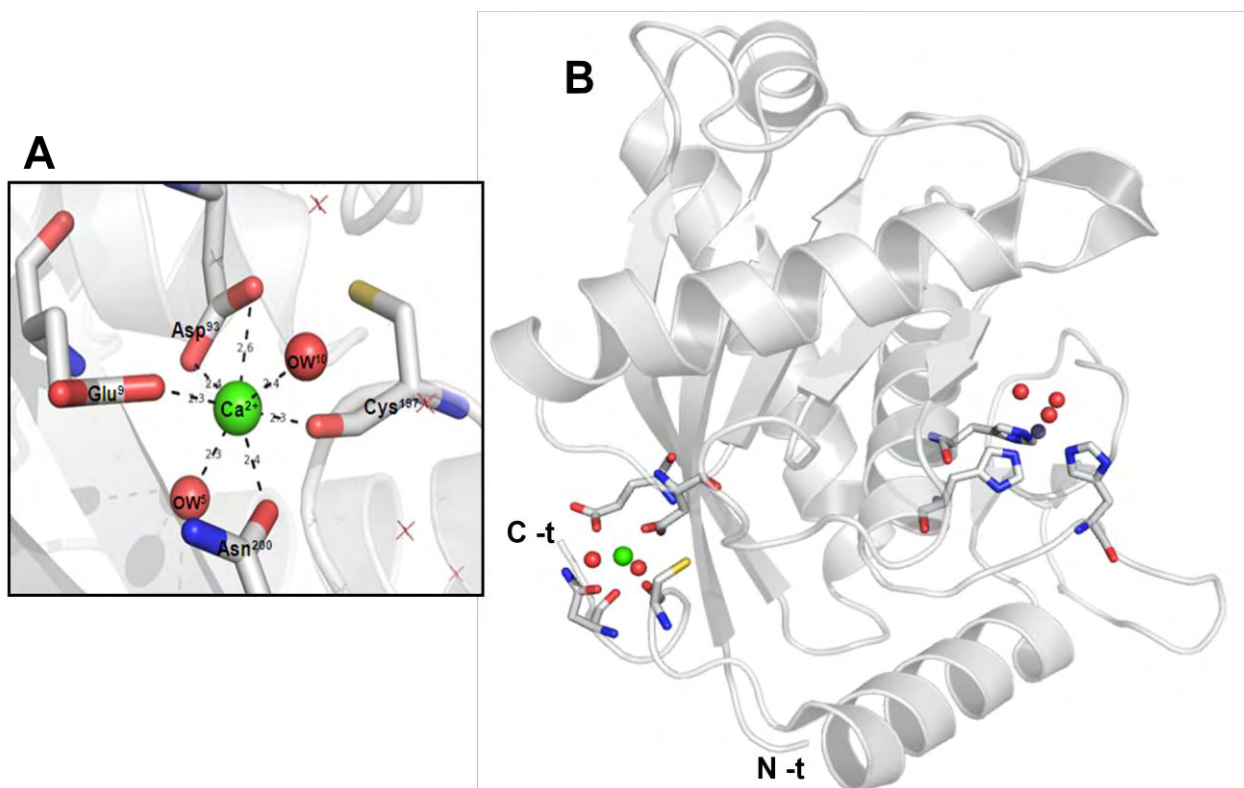


Figura 21. A) Representação do sítio de ligação ao íon cálcio da BmooMP α -I. B) Estrutura da BmooMP α -I evidenciando os locais de ligação dos íons Zn²⁺ e Ca²⁺. Figuras feitas no programa Pymol (<http://www.pymol.org/>).

4.1.7. Predição das interações moleculares entre a BmooMP α -I e alguns inibidores

Muitos inibidores conhecidos de SVMPs foram complexados com a BmooMP α -I e submetidos a ensaios de cristalização, porém não foi possível obter nenhum cristal adequado para as difrações de raios X. Por isso, na tentativa de elucidar o modo de interação da BmooMP α -I, foram feitos estudos de *docking* molecular com ligantes flexíveis. As metodologias de *docking* são utilizadas para determinar a geometria de complexos receptor-ligante a partir da estrutura do receptor e a posterior obtenção de uma estimativa da afinidade de ligação entre o receptor e o ligante.

Na tabela 4 são fornecidas as descrições dos ligantes utilizados nos estudos de *docking*, juntamente com seus PDBIDs, suas energias de ligação e suas fórmulas e pesos moleculares.

Tabela 4. Descrição dos ligantes acoplados e suas respectivas energias de ligação calculadas. Os valores da energia de ligação são valores relativos.

Molécula	Nome da molécula	Formula molecular	Massa Molecular (g/mol)	PDBID	Energia de Ligação (kcal/mol)
1	pyroGlu-Gln-Trp (pEQW)	C ₂₁ H ₃₂ N ₅ O ₆	450.03	1KUI	-37.27
2	pyroGlu-Lys-Trp (pEKW)	C ₂₂ H ₃₄ N ₅ O ₅	448.01	1KUK	-36.44
3	pyroGlu-Asn-Trp (pENW)	C ₂₀ H ₃₀ N ₅ O ₆	436.0	1KUG	-35.87
4	(2R)-N-Hydroxy-2-[(3S)-3-Methyl-3-{4-[(2-Methylquinolin-4-YL)Methoxy]Phenyl}-2-Oxopyrrolidin-1-YL]Propanamide	C ₂₅ H ₂₇ N ₃ O ₄	433.50	2FV5	-34.23
5	3,N(D,L-[2-(Hydroxyamino-Carbonyl)Methyl]-4-Methyl Pentanoyl)L-3-(Tert-Butyl)Glycyl-L-Alanine	C ₁₉ H ₃₇ N ₅ O ₅	415.53	2FV9	-30.88
6	4-{4-[(4-Aminobut-2-Ynyl)Oxy]Phenyl}Sulfonyl-N-Hydroxy-2,2-Dimethylthiomorpholine-3-Carboxamide	C ₁₇ H ₂₃ N ₃ O ₅ S ₂	413.51	2A8H	-28.85
7	(3S)-4-{[4-(But-2-Ynyloxy)Phenyl]Sulfonyl}-N-Hydroxy-2,2-Dimethylthiomorpholine-3-Carboxamide	C ₁₇ H ₂₂ N ₂ O ₅ S ₂	398.50	1ZXC	-22.92

O desvio quadrático médio (RMSD) entre as melhores configurações *in silico* e as estruturas cristalinas obtidas por difração de raios X foram 1ZXC 1.937 Å, 2A8H 1.823 Å, 2FV5 3.394 Å, 2FV9 4.815 Å, 1KUI 2.31 Å, 1KUK 1.80 Å e 1KUG 1.99 Å. Todos os *dockings* mostraram uma boa concordância com relação aos modos de ligação experimentais observados nas estruturas cristalinas, sendo que as mudanças conformacionais mais significantes ocorreram devido a substituições de aminoácidos ao redor do sítio ativo, resultando em distintas propriedades físico-químicas no micro-ambiente.

O ligante pEQW resultou no melhor valor de energia de ligação (-37,27 kcal.mol⁻¹) e valores muito semelhantes também foram obtidos para o pENW e o pEKW (Tabela 4). No modelo complexado, os resíduos pyroGlu e Gln ocupam uma conformação de cadeia diferente, obstruindo a entrada para o sítio ativo, enquanto que na estrutura experimental (1KUI) esses resíduos são paralelos ao sítio de ligação ao zinco (Figura 22A, 22B e 22C). Estas novas conformações adotadas pelos resíduos pyroGlu e Gln são causadas por uma mutação dupla Trp¹¹⁰Arg e Arg¹⁰⁵Gly ao redor do sítio ativo (Figura 22A, 22B e 22C). Outra diferença significativa ocorreu nos átomos que coordenam o íon zinco catalítico, onde o C-terminal do grupo carboxílico do resíduo Trp na estrutura experimental é substituído por átomos de oxigênio da carbonila do pyroGlu e Gln no modelo complexado (Figura 22A, 232 e 22C).

O composto 4 resultou em um modo completamente novo de ligação, com uma inversão da cadeia utilizando o íon zinco como referência espacial. Os determinantes estruturais para este modo de ligação são a substituição Ala¹⁷⁰Leu e o encurtamento do laço que compreende os

resíduos 111 à 115 na estrutura da BmooMP α -I (Figura 22D). O encurtamento do laço permite uma nova configuração do anel indol. O modo esperado de ligação não é estericamente permitido devido ao grupo metila ligado ao grupo indol e a presença de Leu¹⁷⁰ no subsítio esperado.

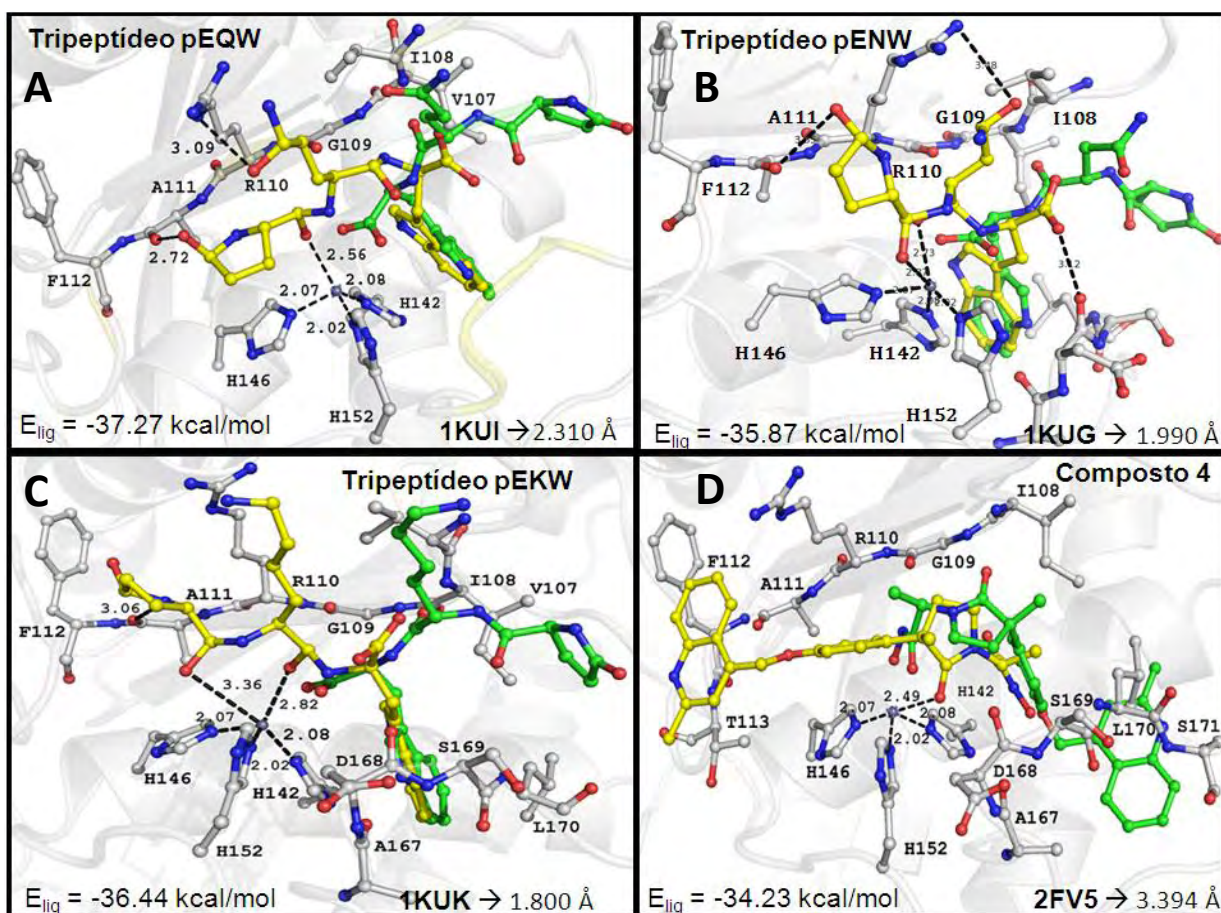


Figura 22. Comparação dos modos de ligação experimental e *in silico*. (A) Tripeptídeo pEQW da estrutura 1KUI, (B) Tripeptídeo pENW da estrutura 1KUG, (C) Tripeptídeo pEKW da estrutura 1KUK e (D) Composto 4 da estrutura 2FV5. Os átomos são representados por bastões e esferas. Os átomos de carbono dos ligantes acoplados e os átomos de carbono dos ligantes experimentais estão em amarelo e verde, respectivamente. As cores dos átomos da proteína seguem o padrão das cores descrito anteriormente. Somente as interações entre a BmooMP α -I e o ligante acoplado estão apresentadas como linhas pretas pontilhadas. Os resíduos de aminoácidos e as interações estão inseridas na própria figura. Figuras feitas no programa Pymol (<http://www.pymol.org/>).

O composto 5 exibe um modo de ligação semelhante ao do cristal 2FV9. Pequenas diferenças são observadas na posição da cadeia, porém todas as interações são conservadas, exceto por uma ligação de hidrogênio adicional entre Ser¹⁶⁹ e átomo de nitrogênio da cauda ligante (Figura 23A). Resultados similares foram observados para o composto 6, que conserva o modo de ligação com o íon zinco catalítico e somente a cauda alifática é deslocada devido à presença da Leu¹⁷⁰ no subsítio (Figura 23B).

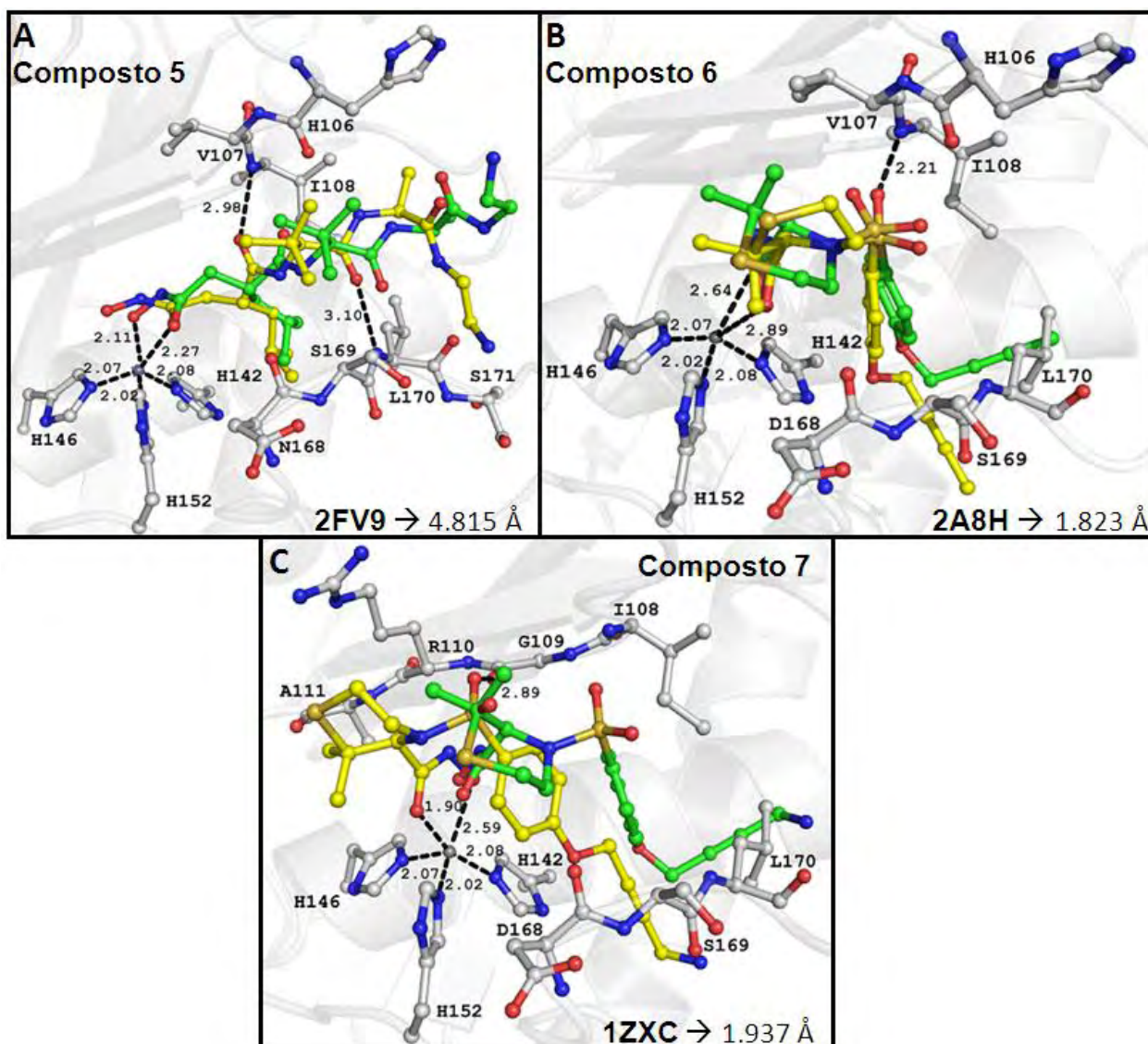


Figura 23. Comparação dos modos de ligação experimental e *in silico*. (A) Composto 5 da estrutura 2FV9, (B) Composto 6 da estrutura 2^a8H e (C) Composto 7 da estrutura 1ZXC. Os átomos são representados por bastões e esferas. Os átomos de carbono dos ligantes acoplados e os átomos de carbono dos ligantes experimentais estão em amarelo e verde, respectivamente. As cores dos átomos da proteína seguem o padrão das cores descrito anteriormente. Somente as interações entre a BmooMP α -I e o ligante acoplado estão apresentadas como linhas pretas pontilhadas. Os resíduos de aminoácidos e as interações estão inseridas na própria figura. Figuras feitas no programa Pymol (<http://www.pymol.org/>).

Um combinatório dos efeitos acima mencionados pôde ser observado no modo de ligação do composto 7, incluindo uma nova conformação da cauda do ligante devido à cadeia lateral da Leu¹⁷⁰ e a pentacoordenação do íon zinco por outro conjunto de átomos (Figura 23C).

4.1.8. Base estrutural para a atividade hemorrágica

Sobreposição das estruturas de BmooMP α -I, da adamalisina-II e H2-proteinase (SVMPs não-hemorrágicas de classe PI) sobre a acutolisina A e acutolisina C (SVMPs hemorrágicas de classe PI) mostrou que ambos os grupos conservam todos os elementos considerados essenciais para a catalise, incluindo as três histidinas catalíticas, o Glu¹⁴¹ polarizador e o *Met-turn*, nas mesmas conformações relativas.

Análises de interface com os programas Protein Interface, Surfaces e Assemblies web Server (Krissinel e Henrick, 2007), revelaram que ambos os subgrupos apresentam uma área superficial acessível ao solvente muito similar, cerca de 9500 Å² e uma energia livre de solvatação em torno de -190 Kcal.mol⁻¹. Esses programas transformam a macromolécula em varias esferas de Van der Waals rígidas e usam uma sonda esférica e sólida, para simular uma molécula de solvente, que rola sobre a superfície dessas esferas. A superfície esférica da sonda, em contato com os vários átomos da superfície da macromolécula descreve alguns segmentos, concavidades e reentrâncias que compõe toda a superfície gerada (Ferreira, 2009).

As diferenças foram significantes somente na região dos laços adjacentes do *Met-turn* (resíduos 153-164 e 167-176) e no C- e N-terminal devido a deleções de resíduos (Figura 24A). Essa variabilidade no N- e C-terminal e a distribuição de carga superficial têm sido muito estudadas, porém ainda não foram obtidos resultados que pudessem diferenciar esses subgrupos.

Assim, utilizando um algoritmo baseado na conservação evolucionária (identificar as regiões funcionais das proteínas) e análises comparativas seqüenciais e estruturais foi descoberto um motivo particular, formado pelos segmentos de resíduos 153-164 e 167-176 adjacentes ao *Met-turn*, que apresenta um padrão coerente na composição dos aminoácidos, tornando possível a diferenciação de ambas as classes e conseqüentemente pode estar envolvido na atividade hemorrágica das SVMPs de classe PI. Nessa hipótese, as SVMPs não-hemorrágicas contém aminoácidos com cadeias laterais grandes nesse motivo, o que ocasiona impedimentos estéricos prevenindo o reconhecimento de substratos. Entretanto, em SVMPs hemorrágicas esse motivo é composto por aminoácidos polares com cadeias laterais pequenas (Ser e Thr) favorecendo a ligação e o reconhecimento de substratos alvos da indução hemorrágica. Estudos anteriores também marcaram o segmento 153-176 como o principal causador das diferenças estruturais entre as SVMPs de classe PI com atividades hemorrágicas diferentes (Watanabe et al., 2003), concordando com nossas observações.

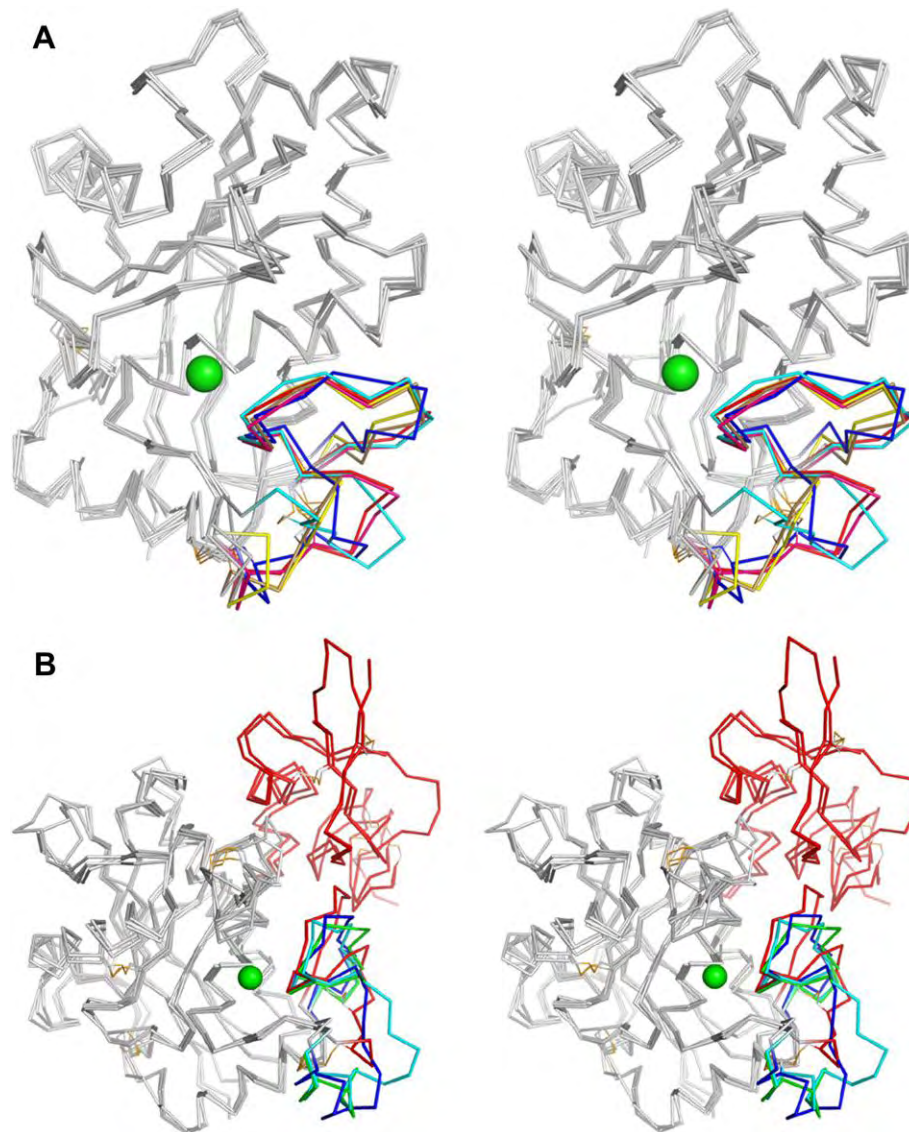


Figura 24. Imagens estereo da representação da cadeia principal ($C\alpha$), comparando a BmooMP α -I com outras SVMPs e as ADAMTS. O íon zinco é representado por uma esfera verde. (A) Resultados da sobreposição da estrutura da BmooMP α -I (vermelho) com outras SVMPs incluindo a adamalysina-II (1IAG, ciano), H2-proteinase (1WNI, amarelo), BaP1 (1ND1, rosa), acutolysina A (1BSW, azul) e acutolysina C (1QUA, areia). (B) Resultados da sobreposição da estrutura da BmooMP α -I (vermelho vivo) com proteínas da família ADAMTS incluindo ADAMTS-1 (2JIH, verde), ADAMTS-4 (2RJP, ciano) e ADAMTS-5 (3B8Z, azul). O domínio desintegrina das ADAMTS esta colorido em vermelho. Figuras feitas no programa Pymol (<http://www.pymol.org/>).

4.1.9. ADAMTs e SVMPs compartilham características estruturais

Alinhamentos múltiplos e sobreposições estruturais indicam que a BmooMP α -I apresenta maior similaridade seqüencial e estrutural com as proteínas da família ADAMTs (*a disintegrin and metalloproteinase with thrombospondin motifs*) que outras metzincinas. As ADAMTs humanas consistem em 19 genes relacionados que estão envolvidos em diversos processos biológicos, tais como adesão celular, coagulação sanguínea, inflamação e angiogênese. Mutações

nos genes das ADAMTs podem estar associadas com varias doenças, incluindo a síndrome de Ehlers-Danlos, osteoartrite, trombozes e podem estar envolvidos em diversos tipos de câncer.

Análises estruturais mostraram que os resíduos cataliticamente relevantes são totalmente conservados, tanto nas SVMs de classe PI quanto nas ADAMTs, incluindo o *Met-turn*, as histidinas catalíticas, Glu¹⁴¹ e as ligações dissulfeto. As diferenças mais significantes no domínio catalítico estão concentradas nos laços adjacentes ao *Met-turn*, que adotam uma conformação completamente diferente devido à diversidade dos aminoácidos que compõem a região (Figura 24B).

Essas observações suportam a hipótese de que o segmento 153-176 possa estar envolvido na especificidade e no reconhecimento macromolecular dessas proteinases. Entretanto, são necessários estudos mais aprofundados para que se possa elucidar o exato papel desse segmento no funcionamento das SVMs, ADAMs e ADAMTs.

4.2 Endo-1,4- β -glicanase

4.2.1. Teste de expressão

O clone (plasmídeo + gene) da endo-1,4- β -glicanase de *Bacillus subtilis* foi gentilmente cedido pelo Prof. Dr. Roberto Ruller (CTBE). Assim, a primeira etapa na execução do projeto foi descobrir qual seriam as melhores condições para expressão da proteína. Para isso, linhagem competente de *Escherichia coli*, BL21 (DE3), foi usada como hospedeiro para a produção da proteína. Após a inserção do plasmídeo, que contém o gene codificante da endo-1,4- β -glicanase, por um processo chamado transformação por eletroporação, foi possível o cultivo dessas células em meio LB, utilizando canamicina como antibiótico, nas seguintes temperaturas de 25, 30 e 37°C. Assim que uma população celular bacteriana apropriada foi alcançada (O.D. 0,6-0,8), a indução da proteína foi feita com o uso de 1 mM de IPTG e a presença da proteína foi analisada em gel de poliacrilamida 13%.

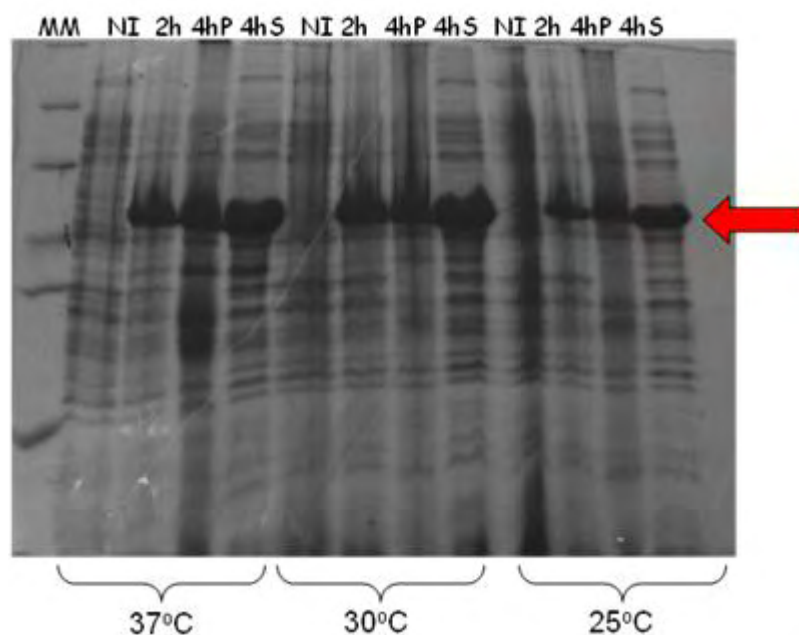


Figura 25. Teste de expressão. SDS-PAGE 13 %. NI - fração não induzida (0 horas – sem IPTG), 2h - a fração induzida após 2 horas da adição de 1 mM de IPTG, 4hP – pellet da fração induzida após 4 horas da adição de 1 mM de IPTG, 4hS – sobrenadante da fração induzida após 4 horas da adição de 1 mM de IPTG e o MM - representa o marcador de peso molecular.

A figura 25 mostra a fração celular não-induzida (NI), que é utilizada como um padrão para comparar a produção de proteínas endógenas e a produção da proteína heteróloga alvo, além da fração induzida com IPTG após 2 e 4 horas. A fração induzida após 4 horas foi submetida a

um ensaio para verificar se a proteína estava sendo expressa na forma solúvel ou em corpos de inclusão. Assim, a fração foi sonicada e o precipitado (4hP) foi separado do sobrenadante (4hS). Este ensaio é de grande importância, pois analisa os componentes protéicos que ficam em solução após o processo de rompimento das células bacterianas por sonicação.

Analisando o gel foi possível notar, que após um período de 4 horas da adição do indutor, a melhor temperatura de expressão na fração solúvel foi a 37°C, embora em outras temperaturas como 30°C também foi observado um nível significativo de expressão na fração solúvel. Desta forma, foi escolhida a temperatura de 37°C como padrão para seguintes experimentos de expressão.

A partir dessa padronização partiu-se para a expressão em larga escala, como descrito nos materiais e métodos.

4.2.2. Extração e degradação da proteína

A expressão foi realizada a partir de um volume de 1 L de LB, a massa celular foi ressuspensa em tampão fosfato 50 mM pH 7,2 + 100 mM de NaCl + 5 mM de benzamidina com adição de lisozima e PMSF, para ajudar na lise da parede celular e inibir a ação de serinoproteases endógenas, respectivamente. Essa amostra foi incubada em gelo por 30 min e em seguida sonicada e centrifugada, estando assim pronta para as purificações.

As primeiras purificações estavam sendo realizadas sem a adição de benzamidina, um inibidor competitivo de tripsina e serinoproteases, e quando a amostra era analisada por SDS-PAGE percebeu-se que a proteína estava sendo rapidamente degradada, impossibilitando a cristalização e os estudos biofísicos. Na figura 26A é possível notar o início da degradação da amostra logo após a concentração e na figura 26B nota-se a degradação da proteína após duas semanas. Assim, foi padronizado o uso de benzamidina em todos os passos, desde a lise da célula bacteriana até a concentração final.

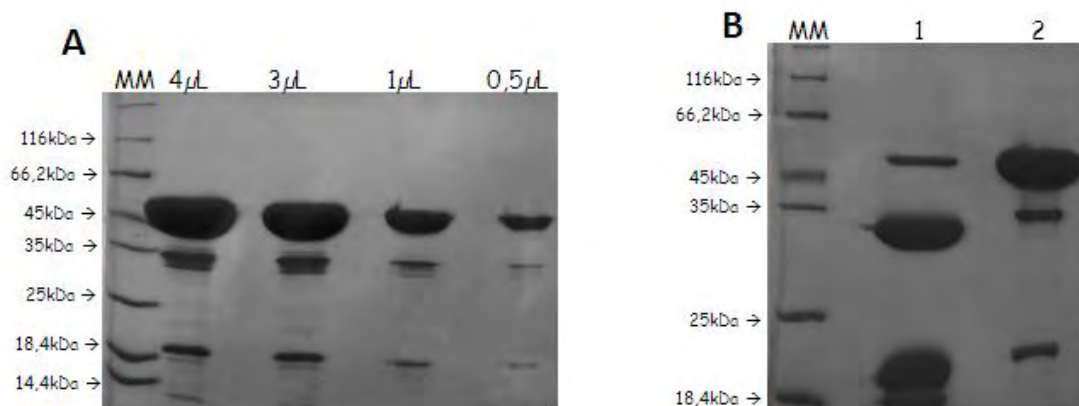


Figura 26. A) SDS-PAGE de diferentes volumes da amostra logo após ser concentrada B) SDS-PAGE da amostra duas semanas após ser concentrada (1); Amostra duas semanas após ser concentrada com adição de benzamidina.

Análises preliminares de espectrometria de massa das respectivas bandas mostradas na figura 26 indicaram que a tripsina estava clivando a endo-1,4-β-glicanase exatamente entre o domínio catalítico (~33 kDa) e o CBM (~19 kDa).

4.2.3. Purificação da Endo-1,4-β-glicanase

A endo-1,4-β-glicanase foi purificada com três passos cromatográficos (Santos et al., 2011 – submetido), sendo o primeiro uma cromatografia por afinidade em resina Hitrap Chelating HP com gradiente linear de imidazol. Inicialmente a coluna foi equilibrada com tampão A e em seguida a amostra foi aplicada, resultando na separação da amostra em 5 picos protéicos (FT, 1, 2, 3 e 4). O primeiro pico consiste na eluição das proteínas que não se ligaram a resina (FT), os picos 1 e 2 de proteínas que se ligaram a resina, porém com baixa intensidade, e picos 3 e 4 proteínas que se ligaram mais fortemente a resina, e por isso necessitaram de uma concentração maior de imidazol para serem liberadas.

Nota-se a presença da banda da proteína alvo praticamente em toda a purificação, porém a partir de 266 mM (Figura 27A pico 3 e Figura 27B a partir da fração 41) podemos perceber uma concentração muito maior da proteína. Todas as frações do pico 3 foram agrupadas e submetidas a um processo de diálise em tampão C.

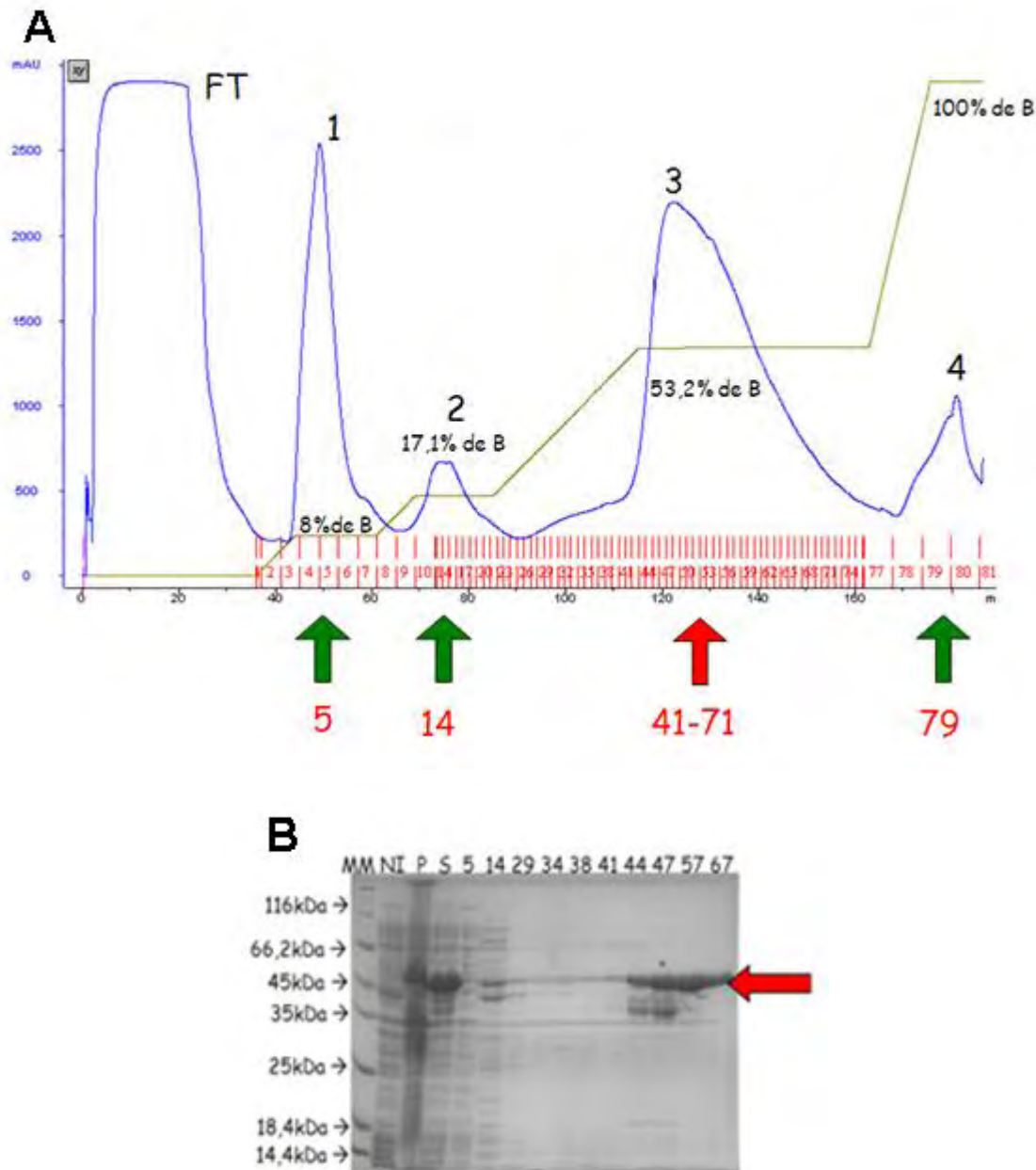


Figura 27. A) Cromatograma obtido a partir da purificação da endo-1,4-β-glicanase em coluna Hitrap Chelating HP **B)** SDS-PAGE, onde MM significa marcador molecular, NI - fração não induzida, P - pellet, S - sobrenadante e os números correspondem às respectivas frações do cromatograma.

No intuito de eliminar os contaminantes do pico 3 (Figura 27B) foi realizada uma segunda cromatografia, que consistiu em uma troca catiônica. Após a diálise o pico 3 foi aplicado em uma coluna Hitrap HP SP e eluído com NaCl em gradiente linear. Dessa purificação resultaram 3 picos: FT, 1 e 2 (Figura 28A). A proteína alvo encontra-se nos picos 1 e 2, porém o pico 2

apresenta um maior grau de pureza (Figura 28B). Os dois picos foram concentrados isoladamente e submetidos a um terceiro passo cromatográfico, exclusão molecular.

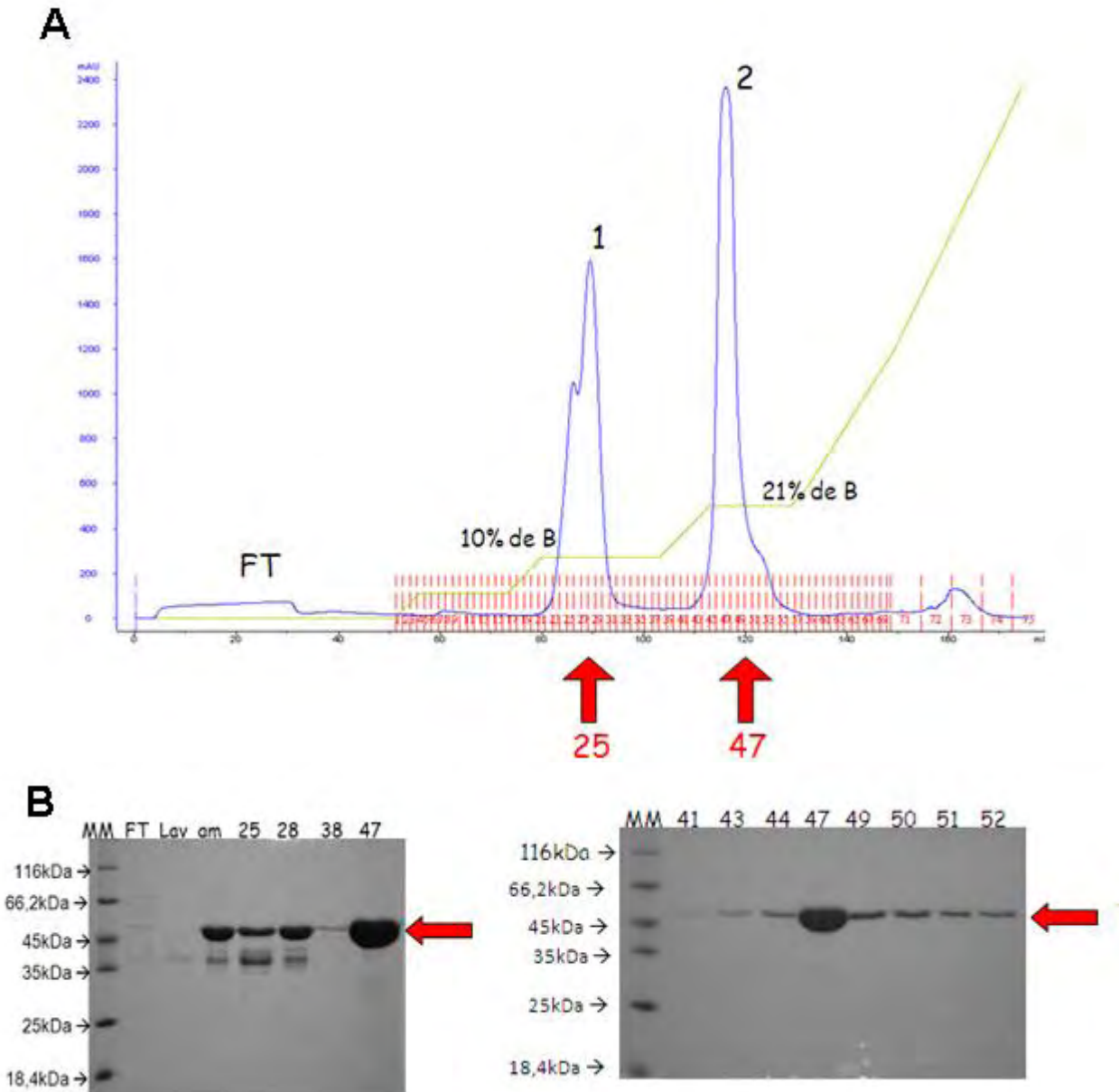


Figura 28. A) Cromatograma obtido a partir da purificação da endo-1,4-β-glicanase em coluna Hitrap HP SP B) SDS-PAGE, onde MM significa marcador molecular, FT - *flow-through*, Lav - lavagem, am - amostra aplicada na coluna e os números correspondem às respectivas frações do cromatograma.

As filtrações em gel foram realizadas em coluna superdex 75, com um fluxo de 0,3 mL/min, podemos observar que as purificações foram muito semelhantes, a proteína alvo eluiu com o mesmo volume de tampão (Figura 29A), evidenciando que ambas as proteínas se

encontram sob a forma de monômeros. As frações da purificação foram analisadas por SDS-PAGE (figura 29B), onde percebe-se que a amostra 2 (am2) apresenta maior quantidade de proteína alvo e com maior grau de pureza do que a amostra 1 (am1), e que a exclusão molecular não conseguiu eliminar os contaminantes presentes na amostra 1.

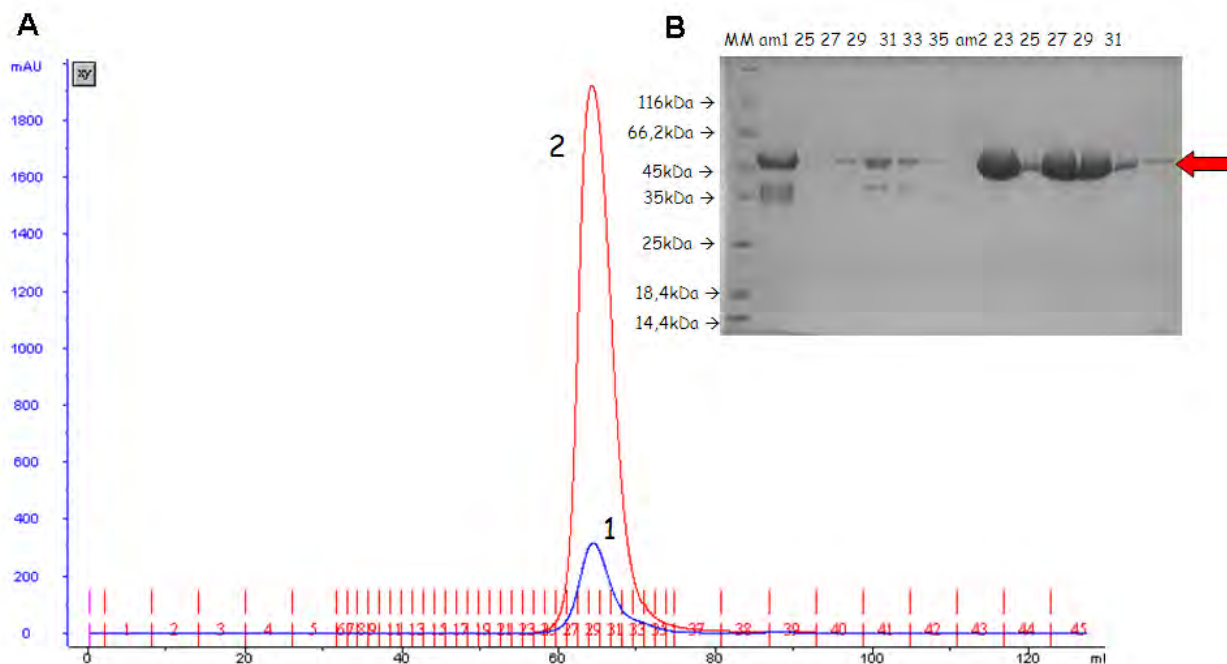


Figura 29. A) Cromatograma obtido a partir da purificação da endo-1,4-β-glicanase em Superdex 75 B) SDS-PAGE, onde MM significa marcador molecular, am1 – amostra 1, am2 – amostra 2 e os números correspondem às respectivas frações do cromatograma.

4.2.4. Rendimento da purificação da Endo-1,4-β-glicanase

Inicialmente todos os passos de concentração estavam sendo realizados com concentrador Amicon Ultra-15 com corte de 30 kDa. Entretanto, a amostra não apresentava um aumento linear da concentração, ao alcançar um certo valor a concentração estagnava. Assim mediu-se a concentração do *flow-through*, e foi comprovado que uma grande quantidade de proteína estava passando pelo filtro. Após análise do concentrador percebeu-se que essa perda de proteína estava ocorrendo porque a matéria prima do filtro do concentrador utilizado é celulose metilada regenerada. Sabe-se, pela literatura, que endo-1,4-β-glicanases são enzimas que possuem a capacidade de degradar polímeros de celulose mesmo quando não estão solúveis e possíveis

derivados de celulose e em caso de aumento da concentração da enzima, promovia a atividade descontrolada sobre o filtro do concentrador.

Depois de detectado o problema foi feita uma mudança nos concentradores, começou-se a utilizar o concentrador Ultrafree CL – 30 kDa (Millipore), que possui um filtro feito de polifluoreto de vinilideno (PVDF) hidrofílico. Pela tabela 5 percebe-se a diferença nas concentrações finais de proteína que era 18,7 mg/mL, e utilizando os novos filtros alcançou aproximadamente 70 mg/mL.

Tabela 5. Diferença de rendimento entre diferentes concentradores.

Tipo de concentrador utilizado	Rendimento (mg)
Amicon	6,7
Ultrafree	31,2

4.2.5. Espalhamento Dinâmico de Luz (DLS)

Medidas com a proteína em baixa concentração foram realizadas para determinar seu estado de agregação e os efeitos do pH na distribuição de tamanhos da proteína (monômeros e/ou oligômeros). A proteína foi diluída para 2 mg/ml, a partir da solução da proteína concentrada a 9,5 mg/ml, nos respectivos tampões (Tabela 6). A tabela 6 apresenta os resultados obtidos em termos do raio hidrodinâmico médio e da polidispersividade da amostra.

A polidispersividade de uma amostra polimérica está associada com a “largura” da sua curva de distribuição de massa molar. Para duas amostras de um mesmo polímero com mesmos valores médios de massa molar, o IP (índice de polidispersividade) é um indicativo da faixa de variação das propriedades hidrodinâmicas da amostra. O ideal é que o IP seja o menor possível para uma amostra homogênea e adequada para estudos biofísicos e estruturais.

Ao analisarmos a tabela 6 podemos perceber que todas as amostras apresentaram um índice de polidispersividade baixo. Essas informações são importantes, porque estão relacionadas ao processo de cristalização, já que quanto mais homogênea (monodispersa) for a amostra, maior é a probabilidade de nucleação e crescimento dos cristais.

Tabela 6. Resultados do DLS em diferentes pHs.

Tampão	pH	R(nm)	IP (%)
Acetato de sódio	3,6	3,1	12,0
Acetato de sódio	4,5	3,0	8,3
Citrato de sódio	5,0	2,8	9,5
MES	6,0	3,2	14,5
Hepes	7,0	3,3	13,8
Hepes	7,5	2,7	13,3

4.2.6. Análises de dicroísmo circular e desnaturações térmicas

Os espectros de CD no UV distante (200 a 260 nm) foram realizados para analisar a estrutura secundária apresentada pela proteína e sua estabilidade térmica. A proteína foi diluída para a concentração de 3 μM em tampão citrato fosfato (pHs: 5,0, 5,5, 6,0, 7,2, 8,0, 5,5 com 10 mM EDTA e 5,5 com 10 mM MnCl₂) às temperaturas de 20 e 65°C. A 20°C foi possível observar uma alta estabilidade estrutural da proteína em pHs ácidos (Figura 30). Já a 65°C, a proteína mostrou-se praticamente toda desenovelada em todos os pHs, com exceção do pH 5,5 com MnCl₂, (curva amarela da Figura 31) onde os íons Mn²⁺ ajudaram na estabilidade protéica fazendo com que somente o pico negativo a 230 nm fosse perdido, provavelmente devido à perda das interações entre o domínio catalítico e o CBM.

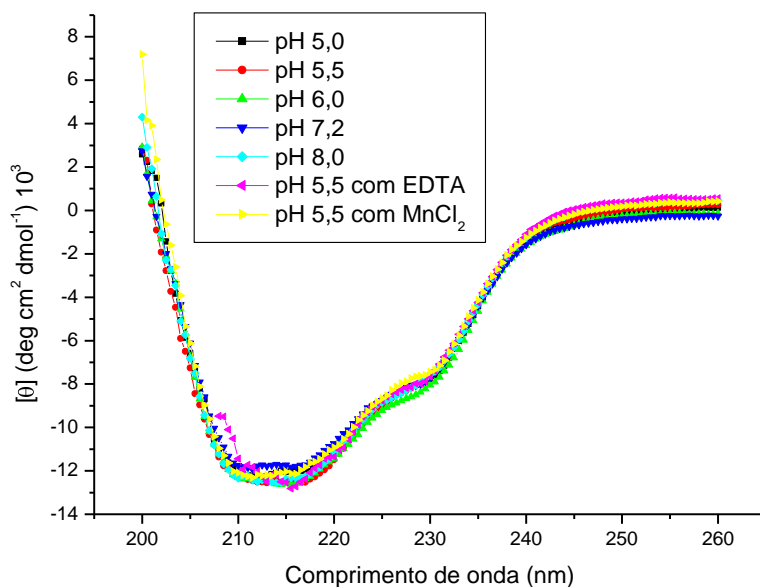


Figura 30: Espectros de CD no UV distante da Endo-1,4-β-glicanase em 30 mM de tampão citrato fosfato nos pHs 5,0, 5,5, 6,0, 7,2, 8,0 5,5 com 0,01 M de EDTA e 5,5 com 0,01 M de $MnCl_2$ a temperatura de 20°C.

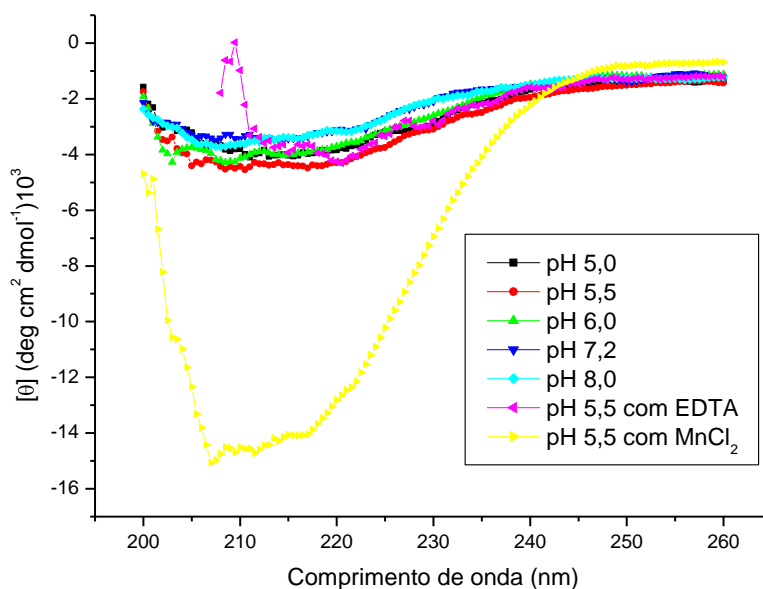


Figura 31: Espectros de CD no UV distante da Endo-1,4-β-glicanase em 30 mM de tampão citrato fosfato nos pHs 5,0, 5,5, 6,0, 7,2, 8,0 e 5,5 com 0,01 M de $MnCl_2$ a temperatura de 65°C.

No gráfico de desnaturação térmica (Figura 32) podemos perceber claramente que os íons Mn^{2+} estabilizaram significativamente a estrutura da proteína, pois sua temperatura de transição (M_t) subiu de 62,9°C para 80°C (Tabela 7). É possível perceber também que a M_t diminui com o

aumento do pH, ou seja, quanto mais básico o pH menor a capacidade da proteína de se manter enovelada.

Tabela 7. Temperaturas de transição (Mt) nos respectivos pHs.

pH	Mt (°C)
5,0	62,3
5,5	62,9
5,5 + EDTA	62,9
5,5 + MnCl ₂	80,0
6,0	63,5
7,2	59,6
8,0	55,6

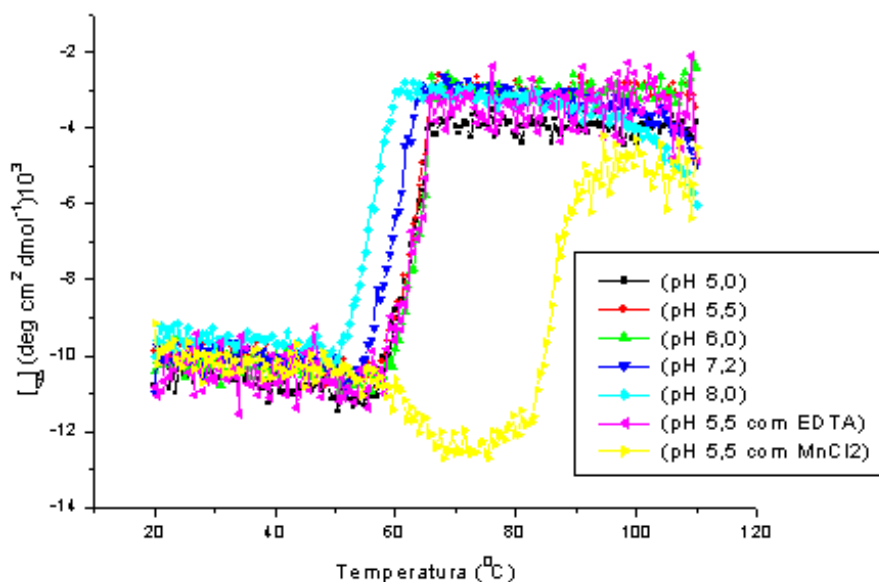


Figura 32: Desnaturação térmica da Endo-1,4-β-glicanase monitoradas por mudanças no UV distante em 222 nm em 30 mM de tampão citrato fosfato nos pHs 5,0, 5,5, 6,0, 7,2, 8,0 e 5,5 com 0,01 M de MnCl₂.

Para compreender melhor os espectros de CD foi feita uma análise quantitativa do conteúdo de estrutura secundária da proteína a 20°C em tampão citrato fosfato pH 6,0. A porcentagem de cada tipo de estrutura secundária predita foi realizada utilizando o programa CDPro (<http://lamar.colostate.edu/~sreeram/CDPro/main.html>), onde foi possível observar que a

proteína apresenta estruturas secundárias típicas como hélices α (2,9%) e folhas β (42,1%), o que é um indício da sua estruturação protéica (Tabela 8).

Tabela 8: Composição de estrutura secundária da endo-1,4-β-glicanase em pH 6,0. Os valores são dados em porcentagem (%) e o erro é de 4,6 % para as medidas adquiridas de 200 a 260 nm.

hélice α	Folhas β	Turn	Random Coil
2,9	42,1	18,5	36,4

4.2.7. Desnaturação química por agentes caotrópicos

Outro experimento realizado foi desnaturação química, com o intuito de verificar a estabilidade da proteína na presença de dois agentes desnaturantes, guanidina e uréia. Para isso as amostras foram diluídas em 10 mM de tampão fostato pH 7,2, deixadas incubar com o respectivo agente desnaturante por 1 hora e em seguida centrifugada por 10 min a 10000 x g.

A partir dos dados de elipticidade em 222 nm da proteína em função da concentração de desnaturante podemos obter alguns parâmetros termodinâmicos como energia livre de Gibbs (ΔG), energia livre de Gibbs da água (ΔG_{H_2O}), concentração molar crítica (CMC) e o valor da dependência da energia livre (m).

Os valores da fração desenovelada foram calculados de acordo com a fórmula:

$$f_D = \frac{\theta_\lambda^{obs} - \theta_\lambda^0}{\theta_\lambda^n - \theta_\lambda^0}$$

onde θ_λ^{obs} é a elipticidade observada a qualquer concentração, θ_λ^n é a elipticidade da proteína totalmente enovelada e θ_λ^0 é a elipticidade na forma completamente desenovelada. Dessa forma calculamos o valor de K (fração desenovelada pela fração enovelada):

$$K = \frac{f_D}{f_N} \Rightarrow K = \frac{f_D}{1 - f_D}$$

Ambas as desnaturações apresentaram um formato sigmoideal indicando uma transição cooperativa (Figura 33). A CMC da proteína foi encontrada calculando o ajuste de sigmóide. Para a guanidina o valor encontrado foi de 1,62 M e para a uréia é 6,99 M. Curiosamente a estabilidade dessa proteína difere consideravelmente dependendo do agente caotrópico. Como observado na figura 33 a CMC da uréia é muito maior que a obtida para a guanidina, provavelmente devido ao papel desempenhado pelas diferentes interações eletrostáticas.

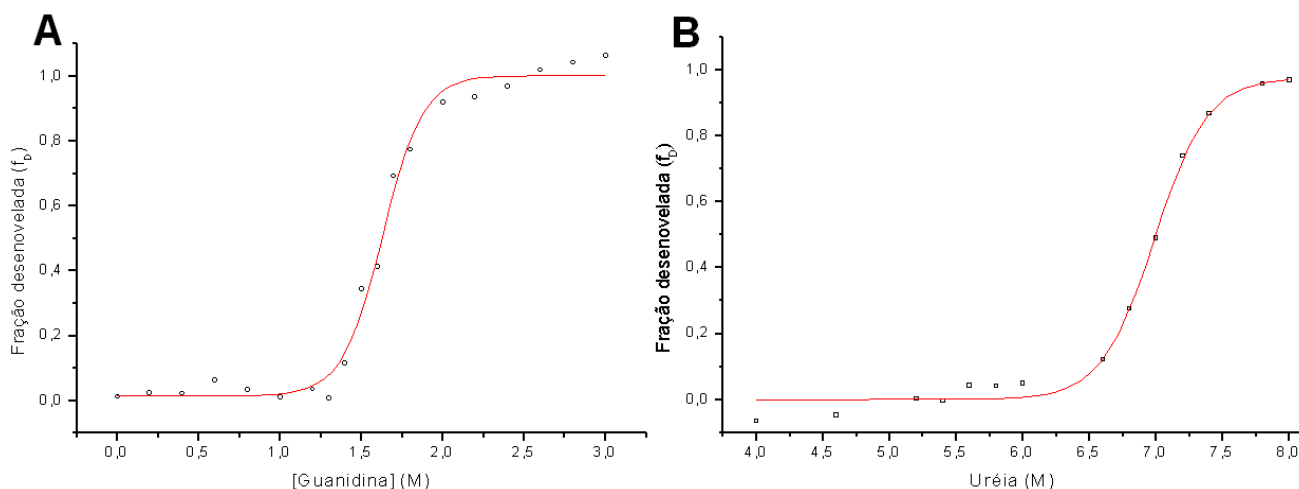


Figura 33: Desnaturação química da Endo-1,4-β-glicanase por uréia (A) e guanidina-HCl (B) monitoradas por mudanças no UV distante em 222 nm. A curva sigmoidal é representada pela linha sólida vermelha.

Utilizando os dados da figura 33 foi possível calcular o ΔG através da equação $\Delta G = -RT \ln(K)$ e assim obter os gráficos de ΔG versus a concentração do desnaturante (Figura 34). Extrapolando esses gráficos linearmente obtivemos os valores de ΔG_{H_2O} e o m :

$$\Delta G = \Delta G_{H_2O} - m \cdot [Desnaturante]$$

onde m é o valor da dependência da energia livre da proteína com a concentração de desnaturante (grau de desenhovelamento) e ΔG_{H_2O} é uma estimativa da estabilidade conformacional da proteína na ausência de desnaturantes. No ponto médio da curva ΔG versus [Desnaturante], onde $\Delta G = 0$ tem-se que a concentração do desnaturante é dada por: $[Guanidina]_{1/2} = \Delta G_{H_2O} / m = 1,63$ M e $[Uréia]_{1/2} = \Delta G_{H_2O} / m = 6,99$ M. Esses valores corroboram com os valores encontrados na sigmóide da figura 33.

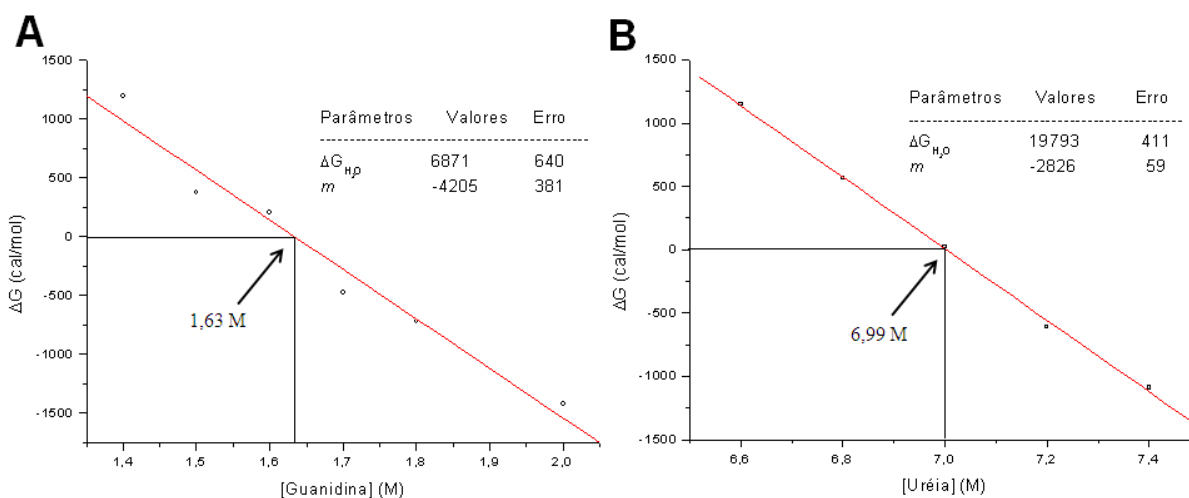


Figura 34: Gráficos de ΔG versus [desnaturante] (A) guanidina (B) uréia. A extrapolação é representada pela linha sólida vermelha.

Os desnaturantes interagem diretamente com a proteína e indiretamente perturbam as propriedades da água como solvente alterando assim a interação água/proteína. Quando a proteína se encontra desenovelada fica mais exposta ao solvente do que quando está em seu estado nativo, dessa forma os desnaturantes ajudam a estabilizar o processo desnaturação química, por isso o coeficiente de energia dos desnaturantes é positivo ($m > 0$). Cada resíduo das proteínas contribui para o valor de m de acordo com a sua fração de exposição a solventes orgânicos obtida pela desnaturação. O coeficiente é uma quantidade extensiva e é proporcionalmente maior para as proteínas maiores, com mais resíduos.

4.2.8. Ensaios de cristalização

Após a análise conjunta dos experimentos de DLS e CD, concluiu-se que a proteína apresenta-se estruturada e monodispersa, possuindo grande potencial de cristalização. Os testes iniciais de cristalização foram feitos com a proteína a 69,5 mg/mL em tampão MES 50 mM pH 6,0, utilizando o robô *HoneyBee* 963 e os kits iniciais Crystal Screen, Crystal Screen 2 (Hampton Research), Wizard Screens I e II (Emerald BioStructures) e Personal Structure Screens 1 e 2 (Molecular Dimensions).

Após uma semana foi possível obter algumas esferolitas e precipitados cristalinos (Figura 34) nas seguintes condições:

- A) Kit inicial 4 condição C1 (0,15M ácido málico + 20% PEG 3350 pH 7,0)
- B) Kit inicial 6 condição D7 (0,1M imidazol + 30% PEG 1500 0,1M CaCl_2 pH 6,5)

C) Kit inicial 1 condição E7 (0,1M acetato de sódio + 20% isopropanol + 0,2M CaCl_2 pH 4,6)

D) Kit inicial 2 condição B2 (0,1M acetato de sódio + 30% PEG 400 + 0,2M CaAc_2 pH 4,5)

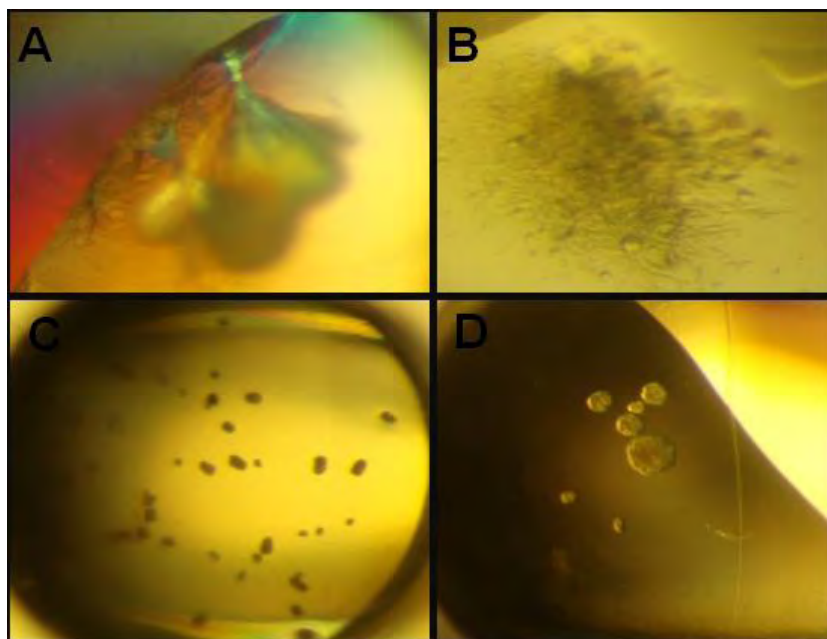


Figura 35. Fotomicrografia das gotas de cristalização.

Essas condições foram submetidas a um refinamento:

- variando o pH de 0,5 em 0,5 na condição inicial, na faixa de 3,5 a 10,0;
- variando o precipitante de 0,2 em 0,2 % na condição inicial, na faixa de 20 a 30;
- adicionando 1% de dioxano na condição inicial;
- adicionando 5% de glicerol na condição inicial.

Porém até o momento ainda não foram obtidos cristais difratáveis para os estudos estruturais.

CAPÍTULO 5 – CONCLUSÃO

Esse capítulo apresenta um resumo de todas as atividades executadas no decorrer desse trabalho com destaque para os resultados mais relevantes. Como demonstrado em todo o trabalho, a conclusão também se divide em duas, a primeira parte refere-se à proteína BmooMP α -I de *Bothrops moojeni* e a segunda conclui os resultados da proteína endo-1,4- β -glicanase. Além disso, ao final de cada conclusão são sugeridos o andamento e a perspectiva futura dos trabalhos.

5.1 BmooMP α -I

A estrutura cristalográfica da SVMP não hemorrágica de veneno de *Bothrops moojeni* foi determinada a 1,76 Å de resolução e sua estrutura revelou uma nova coordenação octaédrica do íon zinco catalítico. Análises comparativas de estruturas de SVMPs hemorrágicas e não-hemorrágicas sugerem que o motivo estrutural que inclui o segmento de aminoácidos 153-164 e 167-176 é responsável pela diferenciação funcional entre os subgrupos da classe PI das SVMPs. Neste trabalho foi proposto que quando as cadeias laterais destes segmentos são volumosas como nas SVMPs não-hemorrágicas, ocorre um impedimento estérico que previne a interação macromolecular. Enquanto que SVMPs hemorrágicas possuem resíduos de aminoácidos polares com cadeias laterais menores nessa região.

Comparações estruturais com as ADAMTs demonstraram que os resíduos cataliticamente relevantes são totalmente conservados e que as diferenças mais significantes estão neste mesmo motivo estrutural descrito como fator de diferenciação entre os subgrupos de SVMPs. Os estudos de *docking* molecular foram realizados com ligantes de metzincinas conhecidos experimentalmente, e indicaram a associação favorável desses inibidores ao bolsão do sítio ativo da BmooMP α -I, o que pode servir como modelo para entender as proteínas humanas relacionadas.

Uma possibilidade promissora para futuros experimentos seria a produção heteróloga das SVMPs de classe I em sistema procarioto e por seguinte mutação sítio dirigido neste motivo estrutural (interconversão).

5.2 Endo-1,4- β -glicanase

A padronização da expressão e a purificação da endo-1,4- β -glicanase de *Bacillus subtilis* foi estabelecida com sucesso, e a obtenção da proteína suficientemente pura e em altas concentrações possibilitou a realização dos ensaios de cristalização e espectroscópicos.

Estudos de espalhamento dinâmico de luz contribuíram para observar a presença de uma solução monodispersa formada de monômeros. A estabilidade térmica e química da endo-1,4- β -glicanase foi avaliada por dicroísmo circular, e os resultados indicaram que a proteína possui uma significativa estabilidade térmica na faixa de pH de 5,0 to 8,0. Curiosamente, a presença de íons de Mn^{2+} aumenta a estabilidade da proteína resultando em um aumento de 17°C em sua temperatura de transição. Além disso, foi observado que essa proteína apresenta um diferente padrão de estabilidade na desnaturação com uréia e guanidina. Esse comportamento anômalo pode ser originado devido ao distinto mecanismo molecular envolvendo a análise das interações eletrostáticas favoráveis sobre a superfície de proteínas globulares.

Os experimentos de cristalização, até o momento, não resultaram em cristais difratáveis, sendo apenas possível observar a presença de precipitados cristalinos e esferolitas. As placas estão sob observação e algumas condições de refinamento estão em andamento com o objetivo de obter cristas difratáveis para os estudos estruturais. A elucidação dessa estrutura juntamente com as investigações biofísicas realizadas poderão trazer informações para o melhor entendimento da termoestabilidade e especificidade das glicanases da família 5.

Como podemos ver no apêndice, outros estudos foram realizados por alunos do grupo a partir da clivagem da proteína entre o domínio catalítico e o CBM, seguida de tentativas de cristalização, estudos biofísicos e comparações dos resultados da proteína inteira com os domínios separados. Abaixo temos as principais conclusões do artigo submetido (apêndice).

- Deleção do módulo de ligação ao carboidrato (CBM) aumenta a estabilidade térmica;
- O íon manganês hiper estabiliza a BsCel5A por um efeito específico sobre o domínio catalítico;
- Estudos cristalográficos revelam a base molecular para o efeito de estabilização do íon manganês;
- CBM de *B. subtilis* representa uma nova subfamília de CBM3, que não possui o sítio de ligação ao cálcio canônico;

- Os resíduos polares no CBM desempenham um papel fundamental no reconhecimento de oligossacarídeos solúveis como avaliado por deslocamentos químicos de NMR;
- O domínio acessório associado ao domínio catalítico forma uma grande superfície planar necessária para a ligação em celulose cristalina;

REFERÊNCIAS BIBLIOGRÁFICAS

Adler, A. J., Greenfield, N. J., Fasman, G. D. Circular dichroism and optical rotatory dispersion of proteins and polypeptides. *Methods in Enzymology*, v. 27, p. 675-735, 1973.

Abagyan, R., Totrov, M., Kuznetsov, D. ICM: a new method for structure modeling and design: applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry*, v. 15, n. 5, p. 488–506, 1994.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology*, v. 215, n. 3, p. 403–410, 1990.

Anai, K., Sugiki, M., Yoshida, E., Maruyama, M. Neutralization of a snake venom hemorrhagic metalloproteinase prevents coagulopathy after subcutaneous injection of *Bothrops jararaca* venom in rats. *Toxicon*, v. 40, n. 1, p. 63-68, 2002.

Arantes, V. e Saddler, J. N. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnology for Biofuels*, v.3, n. 1, pp. 4-15, 2010.

Bernardes, C. P., Santos-Filho, N. A., Costa, T. R., Gomes, M. S., Torres, F. S., Costa, J., Borges, M.H., Richardson, M., Santos, D.M., Castro Pimenta, A. M., Homsí-Brandeburgo, M.I., Soares, A.M., Oliveira, F., Isolation and structural characterization of a new fibrin(ogen)olytic metalloproteinase from *Bothrops moojeni* snake venom. *Toxicon*, v. 51, n. 4, p. 574–584, 2008.

Bieber, A. L, In: Lee, C.Y. (Ed.). *Metal and non protein constituents in snake venoms.*, New York: Springer-Verlag. 1979.

Bjarnason, J. B. e Fox, J. W. Hemorrhagic metalloproteinases from snake venoms. *Pharmacology & Therapeutics*, v. 62, n. 3, p. 325-372, 1994.

Blundell, T. L. Metalloproteinase superfamilies and drug design. *Nature Structural Biology*, v. 1, n. 2, p.73-75, 1994.

Bochkov, A. F. Sova, V. V., Kirkwood, S. The study of the action pattern of an exo- β -(1-3)-D-glucanase. *Biochimica et Biophysica Acta*, v.258, n. 2, p. 531-540, 1972.

Bode, W., Gomis-Rüth, F. X., Huber, R., Zwillig, R., Stöcker, W. Structure of astacin and implications for activation of astacins and zinc-ligation of collagenases. *Nature*, v. 358, n. 6382, p. 164–167, 1992.

Bode, W., Gomis-Ruth, F. X., Stockler, W. Astacins, serralyins, 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 Letters*, v. 331, n. 1-2, p. 134–140, 1993.

Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallographica. Section D, Biological Crystallography*, v. 50, n. Pt 5, p. 760–763, 1994.

Coughlan, M. P. In: *Biotechnology and genetic engineering reviews*. Russell, G.E., editor. v. 3. Newcastle-upon-Tyne: Interscience. The properties of fungal and bacterial cellulases with comment on their production and application. p. 37–109, 1985.

Creighton, T. E. *Proteins: Structures and Molecular Properties*. (Second Edition ed.). New York: Freeman, 1993.

Daltry, J. C., Wüster, W., Thorpe, R. S. Diet and snake venom evolution. *Nature*, v. 379, n. 6565, p. 537-540, 1996.

Datta, G. Dong, A., Witt, J., Tu, A.T. Biochemical characterization of basilase, a fibrinolytic enzyme from *Crotalus basiliscus basiliscus*. *Archives of biochemistry and biophysics*, v. 317, n. 2, p. 365-373, 1995.

Davies, G. e Henrissat, B. Structures and mechanisms of glycosil hydrolases. *Structure*, v.15, n. 3, p. 853-859, 1995.

DeLano, W. L., The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, USA, 2002.

Dunn, R. D. e Broady, K. W. Snake inhibitors of phospholipase A2 enzymes. *Biochimica et Biophysica Acta*, v. 1533, n. 1, p. 29-37, 2001.

Emsley, P. e Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallographica. Section D, Biological Crystallography*, v. 60, n. Pt 12 Pt 1, p. 2126–2132, 2004.

Escalante, T., Franceschi, A., Rucavado, A., Gutiérrez, J. M. Effectiveness of batimastat, a synthetic inhibitor of matrix metalloproteinases, in neutralizing local tissue damage induced by BaP1, a hemorrhagic metalloproteinase from the venom of the snake *bothrops asper*. *Biochemical Pharmacology*, v. 60, n. 2, p. 269-274, 2000.

Esteghalian, A. R., Srivastava, V., Gilkes, N., Gregg, D. J., Saddler, J. N.. An overview of factors influencing the enzymatic hydrolysis of lignocellulosic feedstocks. In: *Glycosyl Hydrolases for Biomass Conversion*. Himmel, M. E., Baker, J. O., Saddler, J. N., editor. Washington DC: American Chemical Society, p. 100–111, 2000.

Fan, L. T., Gharpuray, M. M., Lee, Y.-H. Cellulose Hydrolysis. Berlin: Springer-Verlag, 202f. 1987.

Ferreira, R. N. Sequenciamento completo e estrutura cristalográfica da metaloprotease leucurolisina-A, purificada do veneno de *Bothrops leucurus*. 2009. 94f. Tese (Doutorado) - Universidade Federal de Minas Gerais, 2009.

Fox, J. W. e Serrano, S. M. T. Structural considerations of the snake venom metalloproteinases, key members of the M12 reprotolysin family of metalloproteinases. *Toxicon*, v. 45, n. 8, p. 969–985, 2005.

Fundação nacional da saúde. Manual do diagnóstico e tratamento de acidentes por animais peçonhentos. Ministério da saúde- Brasil. p.112, 2001.

Gasmi, A., Srairi, N., Karoui, H., El Ayeb, M. Amino acid sequence of VIF: identification in the C-terminal domain of residues common to non-hemorrhagic. *Biochimica et Biophysica Acta*, v.1481, n.1, p. 209-212, 2000.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, A. Protein identification and analysis tools on the ExPASy server. In: Walker, J.M. (Ed.), *The Proteomics Protocols Handbook*. New York: Human Press, p.571–607, 2005.

Genta, F.A. Purificação e caracterização de β -glucanases digestivas de *Periplaneta americana* (Dictyoptera). 2000. 158f. Tese (Mestrado) – Universidade de São Paulo, 2000.

Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C. Jr., Warren, R.A. Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families. *Microbiological reviews*, v.55, n.2, p. 303-315.

Gomis-Rüth, F. X., Kress, L. F., Kellermann, J., Mayr, I., Lee, X., Huber, R. Refined 2.0 Å X-ray crystal structure of the snake venom zinc endopeptidase adamalysin II. Primary and tertiary structure determination, refinement, molecular structure and comparison with astacin, collagenase and thermolysin. *Journal of Molecular Biology*, v. 239, n. 4, p. 513–544, 1994.

Gomis-Rüth, F. X., Kress, L. F., Bode, W. First structure of a snake venom metalloproteinase: a prototype for matrix metalloproteinases/collagenases. *The Embo Journal*, v. 12, n. 11, p. 4151-4157, 1993.

Gong, W., Zhu, X., Liu, S., Teng, M., Niu, L. Crystal structures of acutolysin A, a three-disulfide hemorrhagic zinc metalloproteinase from the snake venom of *Agkistrodon acutus*. *Journal of Molecular Biology*, v. 283, n. 3, p. 657–668, 1998.

Gutiérrez, J. M., Rucavado, A., Escalante, T., Díaz, C. Hemorrhage induced by snake venom metalloproteinases: biochemical and biophysical mechanisms involved in microvessel damage. *Toxicon*, v. 45, n. 8, p. 997–1011, 2005.

Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, v. 166, n. 4, p. 557-580, 1983.

Hege, T. e Baumann, U. The conserved methionine residue of the metzincins: a site-directed mutagenesis study. *Journal of Molecular Biology*, v. 314, n. 2, p.181-186, 2001.

Hellman, U., Wernstedt, C., Góñez, J., Heldin, C. H. Improvement of an “In-Gel” digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Analytical Biochemistry*, v. 224, n. 1, p. 451–455, 1995.

Henrissat, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal*, v. 280 (Pt 2), p. 309-316, 1991.

Henrissat B. e Bairoch, A. Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal*. v. 316 (Pt 2), p. 695-696, 1996.

Hooper, N. M. Families of zinc metalloproteases. *FEBS Letters* v. 354, n. 1, p.1-6, 1994.

I.U.B. Enzyme Nomenclature: Recommendations of the Nomenclature Committee of the International Union of Biochemistry on the Nomenclature and Classification of Enzyme-Catalysed Reactions, Academic Press, London and New York, 1984.

International Union of Biochemistry and Molecular Biology Nomenclature Comitee. *Enzyme Nomenclature San Diego: Academic Press* 863f., 1992.

Irwin, D. C., Spezio, M., Walker, L. P., Wilson, D. B. Activity studies of eight purified cellulases: Specificity, synergism, and binding domain effects. *Biotechnology and Bioengineering*, v. 42, n. 8, p. 1002-1013, 1993.

Junqueira, M. R. Aplicação de técnicas proteômicas na caracterização do veneno da serpente *Bothrops insularis* (Viperidae). 2005. 73f. Tese (Mestrado) - Instituto Oswaldo Cruz, 2005.

Klyosov, A. A. e Rabinovich, M. L. Enzymatic conversion of cellulose to glucose: present state of the art and potential. In: *Enzyme Engineering: Future Directions*. Wingard, L.B., Berezin, I.V., Klyosov, A.A., editor. New York: Plenum Press, p. 83–165, 1980.

Klyosov, A. A. Trends in biochemistry and enzymology of cellulose degradation. *Biochemistry*, v.29, n. 47, p. 10577-10585, 1990.

Koshland, D. R. Stereochemistry and mechanism of enzymatic reactions. *Biological Reviews of Cambridge Philosophical Society*, v. 28, p. 416-436, 1953.

Krissinel, E. e Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, v. 372, n. 3, p. 774–797, 2007.

Kumasaka, T., Yamamoto, M., Moriyama, H., Tanaka, N., Sato, M., Katsube, Y., Yamakawa, Y., Omori-Satoh, T., Iwanaga, S., Ueki, T. Crystal structure of H2-proteinase from the venom of *Trimeresurus flavoviridis*. *Journal of Biochemistry*, v. 119, n. 1, p. 49–57, 1996.

Laemmli, U. K. Cleavage of structural proteins during the assembly of head of bacteriophage T4. *Nature*, v. 227, n. 5259, p. 680–685, 1970.

Laskowski, R. A., Moss, D. S., Thornton, J. M. Main-chain bond lengths and bond angles in protein structures. *Journal of Molecular Biology*, v. 231, n. 4, p. 1049–1067, 1993.

Lehninger, A. L., Nelson, D.L., Cox, M.M. *Principles of Biochemistry*. 2 ed Editora Sarvier 33 Irving place, NY, 1995.

Levy, I. e Shoseyov, O. Cellulose-binding domains: biotechnological applications. *Biotechnology Advances*, v. 20, n. (3-4), p. 191-213, 2002.

Lu, X., Lu, D. Scully, M. F., Kakkar, V. V. Snake venom metalloproteinase containing a disintegrin-like domain, its structure-activity relationships at interacting with integrins. *Current Medical Chemistry. Cardiovascular and Hematological Agents*, v. 3, n. 3, p. 249-260, 2005.

Lynd, L. R., Weimer, P. J., Zyl, W. H. van, Pretorius, I.S. Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiology and Molecular Biology Reviews*. v. 66, n.3, p. 506-577, 2002.

Manning, M. C. Sequence analysis of fibrolase, a fibrinolytic metalloproteinase from *Agkistrodon contortrix contortrix*. *Toxicon*, v. 33, n. 9, p. 1189-1200, 1995.

Marcussi, S. Caracterização Funcional e Estrutural de uma Metaloprotease Isolada do Veneno de *Bothrops jararacussu*. 2007. 140f. Tese (Doutorado) – Faculdade de Medicina de Ribeirão Preto, 2007.

Markland, F. S. Snake venom fibrinogenolytic and fibrinolytic enzymes: an updated inventory. *Journal of Thrombosis and Haemostasis*. v. 79, n. 3, p. 668–674, 1998.

Markland, F. S. Inventory of a- and b-fibrinogenases from snake venoms. *Journal of Thrombosis and Haemostasis*, v. 65, p. 438–443, 1991.

Meier, J. S. e Stöcker K. F. In: handbook of Clinical Toxicology of Animal Venoms and Posions. Ed. Meier, J. e White, J. Biology and distribution of venomous snakes of medical importance and the composition of snake venoms. Boca Raton, Florida: CRC Press Inc. p. 367, 1995.

Moreira, M. R. Natureza das interações celulose-água. 2009. 83f. Tese (Mestrado) – Universidade de São Paulo, 2009.

Moura-da-Silva, A. M., Butera, D., Tanjoni, I. Importance of snake venom metalloproteinases in cell biology: effects on platelets, inflammatory and endothelial cells. *Current Pharmaceutical Design*, v. 13, n. 28, p. 2893–2905, 2007.

Murshudov, G. N., Vagin, A. A., Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica. Section D, Biological Crystallography*, v. 53, n. Pt 3, p. 240–255, 1997.

Navaza, J. AMoRe: an automated package for molecular replacement. *Acta Crystallographica. Section A*, v. 50, n. 2, p. 157–163. 1994.

Nelson, T. E. e Lewis, B. A. Separation and characterization of the soluble and insoluble components of insoluble laminarin. *Carbohydrate Research*, v. 33 p. 63-74, 1974.

Nelson, D. L. e Cox, M. M. *Lehninger Principles of Biochemistry*. 4 ed. Editora Sarvier, 2006.

Odell, G.V., Fenton, A. W. Ownby, C.L., Doss, M.P. e Schmidt, J.O. The role of venom citrate. *Toxicon*, v. 37, n. 3, p. 404-409, 1999.

Oh, J-H., Cha, J-A., Yoon, M-H. Molecular characterization of a β -1,4-Endoglucanase gene from *Bacillus subtilis* H12. *Journal of the Korean Society for Applied Biological Chemistry*, v. 51, n. 4, p. 299-304, 2008.

Otwinowski, Z. e Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, v. 276, p. 307–326, 1997.

Paine, M. J., Desmond, H. P., Theakston, R. D., Crampton, J. M. Purification, cloning and molecular characterization of a high molecular weight hemorrhagic metalloprotease, jararhagin, from *Bothrops jararaca* venom. Insights into disintegrin gene family. *Journal of Biological Chemistry*, v. 267, n. 32, p. 22869-22876, 1992.

Perrakis, A., Morris, R., Lamzin, V. S. Automated protein model building combined with iterative structure refinement. *Nature Structural Biology*, v. 6, n. 5, p. 458–463, 1999.

Raghothama, S., Simpson, P. J., Szabó, L., Nagy, T., Gilbert, H.J., Williamson, M.P. Solution Structure of the CBM10 Cellulose Binding Module from *Pseudomonas Xylanase* A. *Biochemistry*, v. 39, n. 5 p. 978-984, 2000.

Ramos, O. H. P. Produção de uma pró-metaloproteinase recombinante em bactéria, ativação in vitro e modelagem molecular. 2001. 147f. Tese (Mestrado) – Universidade Federal de São Carlos, São Paulo, 2001.

Ramos, O.H.P. e Selistre-de-Araujo, H.S. Comparative analysis of the catalytic domain of hemorrhagic and non-hemorrhagic snake venom metalloproteinases using bioinformatic tools. *Toxicon*, v. 44, n. 5, p. 529–538, 2004.

Ramos, O.H.P. e Selistre-de-Araujo, H.S. Snake venom metalloproteinases — structure and function of catalytic and disintegrin domains B. *Comparative Biochemistry and Physiology, Part C*, v. 142, n. 3-4, p. 328–346, 2006.

Santos, C. R., Paiva, J. H., Sforça, M. L., Neves, J. L., Navarro, R.Z., Cota, J., Akao, P. K., Meza, A. N., Polikarpov, I., Xavier-Neto, J., Squina, F. M., Ward, R. J., Ruller, Zeri, A. C., Murakami, M. T. Dissecting the structure-function-stability relationships of family 5A cellulose:Portrayal of a novel ion-based mechanism for thermal stabilization and a new family 3 carbohydrate-binding module lacking the calcium-binding site. *The Journal of Biological Chemistry* – submetido para publicação em 2011.

Sierra, M.C.M. e Pérez, M. B. Serpientes exóticas: nueva moda, nueva urgência. *Revisiónes Jueves Medicina intensiva*, v.2, n.25, p. 66-75, 2001.

Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredes, A., Person, S., Raab, T., Vormerk, S., Youngs, H. Toward a systems approach to understanding plant cell walls. *Science* v. 306, n. 5705, p. 2206–2211, 2004.

Stöcker, W. e Bode, W. Structural features of a superfamily of zinc-endopeptidases: the metzincins. *Current Opinion in Structural Biology*, v. 5, n. 3, p. 383–390, 1995.

Stöcker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Rüth, F. X., McKay, D. B., Bode, W. The metzincins—topological and sequential relations between the astacins, adamalysins, serralytins, and matrixins (collagenases) define a superfamily of zinc-peptidases. *Protein Science*, v. 4, n. 5, p. 823-840, 1995.

Stone, J.E. e Scallan, A.M. Effect of component removal upon the porous structure of the cell wall of wood. *Journal Polymer Science Part C*, v. 11, p. 13–25, 1965.

Takatsuka, H., Sakurai, Y., Yoshioka, A., Kokubo, T., Usami, Y., Suzuki, M., Matsui, T., Titani, K., Yagi, H., Matsumoto, M., Fujimura, Y. Molecular characterization of L-amino acid oxidase from *Agkistrodon halys blomhoffii* with special reference to platelet aggregation. *Biochimica et Biophysica Acta*, v. 1544, n. 1-2, p. 267-277, 2001.

Teeri, T. T., Reinikainen, T., Ruohonen, L., Jones, T. A., Knowles, J. K. C. Domain function in *Trichoderma reesei* cellobiohydrolases. *Journal of Biotechnology*, v. 24, p. 169-176, 1992.

Terada, S., Hori, J., Fujiruma, S., Kimoto, E. Purification and amino acid sequence of brevilysin L6, a non-hemorrhagic metalloprotease from *Agkistrodon halys brevicaudus* venom. *Journal of Biochemistry*, v. 125, n. 1, p. 64-69, 1999.

Valério, A. A. Estudos estruturais e funcionais de duas glicosídeo hidrolases: a celulase putativa XF0810 de *Xylella fastidiosa* e a lisozima digestiva 1 de *Musca domestic*. 2007. 159f. Tese (Doutorado) – Universidade Estadual de Campinas, 2007.

Warrell, D. A. Clinical features of envenoming of snake bites. In: *Envenoming and their treatment*. Ed. C. Bon and M. Goyffon,. Lyon: Fon. Marcel Mérieux, p. 63-76, 1996.

Watanabe, L., Shannon, J.D., Valente, R.H., Rucavado, A., Alape-Girón, A., Kamiguti, A.S., Theakston, R.D.G., Fox, J.W., Gutiérrez, J.M., Arni, R.K. Amino acid sequence and crystal structure of BaP1, a metalloproteinase from *Bothrops asper* snake venom that exerts multiple tissue-damaging activities. *Protein Science*, v. 12, n. 10, p. 2273–2281, 2003.

White, J. Snake venoms and coagulopathy. *Toxicon*, v. 45, n. 8, p. 951–967, 2005.

Willis, T.W., Tu, A.T. Purification and biochemical characterization of atroxase, a nonhemorrhagic fibrinolytic protease from western diamondback rattlesnake venom. *Biochemistry*, v. 27, n. 13, p. 4769- 4777, 1988.

Zhu, X., Teng, M., Niu, L. Structure of acutolysin-C, a hemorrhagic toxin from the venom of *Agkistrodon acutus*, providing further evidence for the mechanism of the pH-dependent proteolytic reaction of zinc metalloproteinases. *Acta Crystallographica. Section D, Biological Crystallography*, v. 55, n. (Pt11), p. 1834-1841, 1999.

APÊNDICE



Structural studies of BmooMP α -I, a non-hemorrhagic metalloproteinase from *Bothrops moojeni* venom

P.K. Akao^{a,c}, C.C.C. Tonoli^a, M.S. Navarro^a, A.C.O. Cintra^b, J.R. Neto^c, R.K. Arni^{c,d}, M.T. Murakami^{a,*}

^a Center for Structural Molecular Biology, Brazilian Synchrotron Light Laboratory, Campinas, SP 13083-970, Brazil

^b Department of Clinical, Toxicological and Bromatological Analysis, FCFRP, USP, Ribeirão Preto, SP, Brazil

^c Department of Physics, IBILCE/UNESP, São José do Rio Preto, SP, Brazil

^d Center for Applied Toxinology, CAT/CEPID, Butantan Institute, São Paulo, SP, Brazil

ARTICLE INFO

Article history:

Received 17 June 2009

Received in revised form 13 August 2009

Accepted 17 August 2009

Available online 23 August 2009

Keywords:

Snake venom metalloproteinase

Bothrops moojeni

Crystal structure

ABSTRACT

Hemostatically active snake venom metalloproteinases (SVMPs) perturb the blood coagulation cascade at specific points and due to their potential application as thrombolytic agents, the fibrin(ogen)olytic non-hemorrhagic SVMPs have been employed as biochemical tools in coagulation research and diagnosis. Structural studies complemented by the design of metalloproteinase inhibitors have been instrumental in understanding their stereo specificity and action mechanism. We present here, details of the crystal structure of BmooMP α -I, a 22.6 kDa non-hemorrhagic P-I class SVMP isolated from *Bothrops moojeni* venom, determined at 1.76 Å resolution. In this structure, the catalytic zinc ion displays an unusual octahedral coordination formed by the three canonical histidines (His¹⁴², His¹⁴⁶ and His¹⁵²) and additionally, by three solvent molecules. Comparative sequence and structural studies indicate that the motif comprising amino acid segments 153–164 and 167–176 adjacent to the methionine-turn is a salient feature that differentiates both non and hemorrhagic P-I class SVMPs and could directly be involved in the development of the hemorrhagic activity.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Snake venom metalloproteinases (SVMPs) play a key role in Viperidae envenomations causing local and pathological effects such as hemorrhage, inhibition of platelet

aggregation, coagulopathy, myonecrosis and inflammatory response (Gutiérrez et al., 2005; White, 2005; Fox and Serrano, 2005). These effects are due to the proteolytic action of SVMPs on extracellular matrix components, plasma proteins and cell surface proteins (Moura-da-Silva et al., 2007). Moreover, SVMPs are also capable of interacting with platelet receptors, endothelial cells and fibroblasts, activating or inhibiting the cellular response (Moura-da-Silva et al., 2007).

SVMPs belong to the metzincin super-family of zinc-dependent peptidases, which also includes metalloproteinases important for mammalian functions like matrix-degrading metalloproteinases (MMPs) and ADAMs (A Disintegrin And Metalloproteinases) (Bode et al., 1993; Stocker and Bode, 1995). SVMPs are synthesized as multidomain zymogen proteins and can be grouped into four classes taking into account their primary structure and domain organization: P-I class comprises only a

Abbreviations: ADAMs, a disintegrin- and metalloproteinase; ADAMTS, a disintegrin and metalloproteinase with thrombospondin motifs; BaP1, P-I class metalloproteinase from *Bothrops asper* venom; BmooMP α -I, P-I class metalloproteinase from *Bothrops moojeni* venom; DLS, dynamic light scattering; ESI-MS, Electrospray Ionization Mass Spectrometry; FPLC, Fast Protein Liquid Chromatography; LNLS, Brazilian Synchrotron Light Laboratory; MMPs, matrix-degrading metalloproteinases; NCBI, The National Center for Biotechnology Information; PDB, protein data bank; PEG, polyethylene glycol; RMSD, root-mean-square deviation; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; SVMPs, snake venom metalloproteinases; UPLC, ultra performance liquid chromatography.

* Corresponding author. Tel.: +55 19 3512 1106; fax: +55 19 3512 1004.

E-mail address: mtmurakami@lnls.br (M.T. Murakami).

zinc-dependent catalytic domain in the mature form and typically display fibrin(ogen)olytic activity; P-II class contains a disintegrin domain in addition to the catalytic domain; P-III class is comprised of a catalytic domain, a disintegrin-like domain and a cysteine-rich domain, which exert potent hemorrhagic effects; and P-IV class has an additional lectin-like domain linked by disulfide bonds to a P-III class SVMP (Fox and Serrano, 2005).

BmooMP α -I is a 22.6 kDa P-I class zinc-dependent metalloproteinase isolated from *Bothrops moojeni* venom, which displays azocaseinolytic, fibrin(ogen)olytic and defibrinating activities, but is devoid of hemorrhagic and thrombin-like activities (Bernardes et al., 2008). BmooMP α -I is of great medical interest as a therapeutic agent in the treatment and prevention of thrombotic diseases due to its defibrinogenating effect observed in *in vivo* experiments in mice (Bernardes et al., 2008). In an attempt to understand the molecular basis of the specificity of BmooMP α -I, its crystal structure has been determined at 1.76 Å resolution and comparative structural analyses with other non- and hemorrhagic SVMPs have been employed to shed light on the structural determinants for their observed pharmacological effects.

2. Material and methods

2.1. Purification and biochemical characterization

Crude desiccated *B. moojeni* venom was obtained from SanMaru serpentarium (Taquaral, São Paulo, Brazil). BmooMP α -I was purified based on the previously published protocol with minor modifications (Bernardes et al., 2008). The purification steps include, sequentially, anion-exchange chromatography on Q-Sepharose (Amersham Biosciences), size-exclusion chromatography on Superdex 75 (Amersham Biosciences) and affinity chromatography on Heparin-Sepharose (Amersham Biosciences). All chromatographic steps were performed using a Fast Protein Liquid Chromatography (FPLC) system. The purity of the sample was confirmed by SDS-PAGE under reducing conditions (Laemmli, 1970). Protein concentration was determined by absorbance at 280 nm, using a theoretical extinction coefficient of $24\,325\text{ M}^{-1}\text{ cm}^{-1}$ (Gasteiger et al., 2005).

2.2. Protein sequencing by electrospray ionization mass spectrometry (ESI-MS/MS)

Coomassie-stained protein bands containing the fibrin(ogen)olytic enzyme isolated from *B. moojeni* venom were excised from the gel and in-gel digested with trypsin, Glu-C (SV-8) and chymotrypsin, according to established protocol in our laboratory (Hellman et al., 1995) and subjected to Electrospray Ionization Mass Spectrometry (ESI) analysis, using a quadrupole-time-of-flight mass spectrometer (Q-Tof Ultima – Waters/Micromass) coupled to a Ultra Performance Liquid Chromatography (UPLC) system (NanoAcquity – Waters). The peptides generated by digestion were desalted on-line using a Waters Symmetry C18 trap column ($5\ \mu\text{m} \times 180\ \mu\text{m} \times 20\ \text{mm}$). The elution was carried out in a BEH 130 C18 ($1.7\ \mu\text{m} \times 75\ \mu\text{m} \times 100\ \text{mm}$)

column using a gradient of 0–60% (v/v) acetonitrile for 1 h. The spectra were acquired using the Data Directed Analysis, selecting the doubly and triply charged peptides for MS/MS experiments. All the MS/MS spectra were processed using the Mascot Distiller software and the MASCOT search engine (Matrix Science, Boston).

2.3. Crystallization and data collection

Prior to the crystallization trials, the sample was concentrated to $10\ \text{mg ml}^{-1}$ in a 20 mM Tris–HCl buffer pH 7.4. Dynamic light-scattering experiments were performed to verify the structural homogeneity of sample using a DynaPro 810 (Protein Solutions) apparatus equipped with a temperature stabilizer. Standard curves of bovine serum albumin were used for calibration and the experiments were conducted at 291 K.

Crystallization was performed by the sitting-drop vapor-diffusion technique at 291 K using a cartesian HoneyBee 963 crystallization robot (Genomics Solutions). The drops were formed by mixing 0.5 μL of protein sample with an equal volume of mother solution. Initial trials were carried out by the sparse-matrix method with some modifications using commercially available crystallization screen kits including Crystal Screen and Crystal Screen 2 (Hampton Research), Wizard Screens I and II (Emerald BioStructures) and Personal Structure Screens 1 and 2 (Molecular Dimensions).

Initially, very small, elongated crystals were obtained from two crystallization conditions, one containing 15% (v/v) ethanol and 100 mM Tris–HCl buffer pH 7.0 and another containing 20% (v/v) polyethylene glycol (PEG) 8000 and 100 mM HEPES buffer pH 7.5. The fine-tuning of pH-shift strategy (0.1 pH unit variation ranging from 4.0 to 9.0) was employed to obtain monocrystals suitable for X-ray diffraction experiments. Rectangular crystal with approximate dimensions of $0.4 \times 0.3 \times 0.2\ \text{mm}$ was obtained after 7–9 days from the solution containing 100 mM Tris–HCl pH 8.3 and 20% (v/v) PEG 8000 by mixing equal volume (1 μL) of protein sample and mother solution.

The crystals were cryoprotected with 20% (v/v) glycerol and flash-cooled to 100 K. X-ray diffraction data were collected using synchrotron radiation at the Beamline DB03-MX1, Brazilian Synchrotron Light Laboratory (LNLS), Brazil. The wavelength of the radiation source was set to 1.428 Å and a MARCCD detector (165 mm) was used to record the X-ray diffraction data. Data sets were indexed, integrated, reduced and scaled using the programs DENZO and SCALEPACK as implemented in the HKL2000 package (Otwinowski and Minor, 1997). Data collection statistics are presented in Table 1.

2.4. Structure determination and refinement

The BmooMP α -I structure was solved by molecular replacement techniques with the program AMoRe (Navaza, 1994) using the atomic coordinates of acutolysin A, a zinc-dependent metalloproteinase isolated from the venom of *Agkistrodon acutus* (PDB code 1BSW, Gong et al., 1998) as the search model.

Table 1
Data processing and refinement statistics.

Beamline setup	
Radiation source	Synchrotron light
Beamline	DB03-MX1, LNLS-Brazil
Wavelength (Å)	1.437
Temperature (K)	100
Detector	MARCCD 165 mm
Crystal preparation	
Cryoprotectant solution	Mother liquor + 20% glycerol
Soaking time (s)	30
Data collection	
Space group	P2 ₁
a (Å)	40.07
b (Å)	48.73
c (Å)	54.39
β (°)	91.56
Resolution range (Å)	20.81–1.76 (1.83–1.76)
Unique reflections	19477
Multiplicity ^a	5.88 (2.35)
Completeness (%)	93.7 (66.70)
I/σ ₁	14.7 (2.87)
R _{merge} (%)	5.4 (26.70)
Matthews coefficient (Å ³ Da ⁻¹)	2.36
Corresponding solvent (%)	47.91
Number of molecules in asymmetric unit	1
Model Refinement	
Protein Data	3GBO
Bank identification	
Number of nonhydrogen atoms	1590
Number of water molecules	144
R _{model} (%)	17.1
R _{free} (%)	21.8
R.m.s. deviation from ideal bond lengths (Å)	0.023
R.m.s. deviation from ideal angles (°)	1.898
Average B-factor (Å ²)	25.7
Ramachandran plot	
Most favored regions (%)	98.5
Allowed regions (%)	1.5
Disallowed regions (%)	0

^a Values for the outermost resolution shell are given in parentheses.

Positional and individual isotropic thermal factor refinements were carried out using REFMAC5 (Murshudov et al., 1997) as incorporated in the CCP4 suite (CCP4, 1994). The 2F_o–F_c and F_o–F_c electron-density maps were examined; the protein model was manually adjusted after each refinement cycle using COOT (Emsley and Cowtan, 2004). During the final cycles water molecules were introduced using the program ARP/wARP (Perrakis et al., 1999) and at this stage, an electron-density difference map clearly indicated the presence of two strong peaks corresponding to a zinc ion bound to the active-site and a calcium ion located at the crossover region of the N- and C-termini.

The overall stereochemistry of the final model was assessed using the program PROCHECK (Laskowski et al., 1993). Three-dimensional structures were displayed,

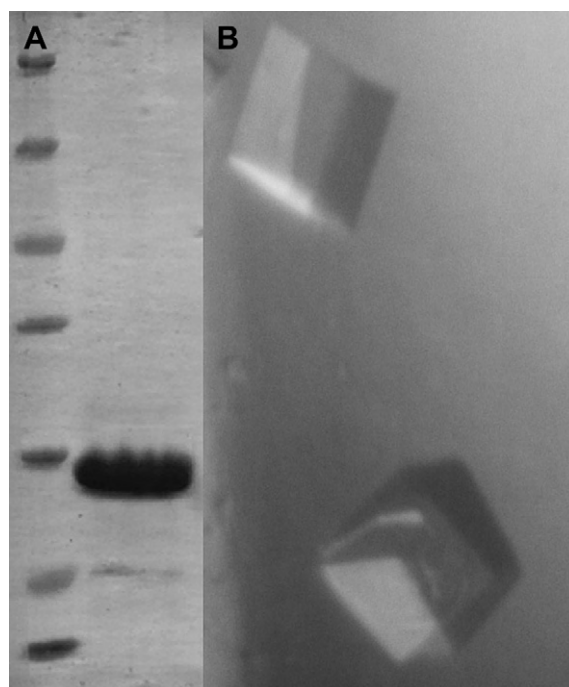


Fig. 1. (A) SDS-PAGE under reducing conditions of the purified BmooMP α -I. The gel was stained with coomassie brilliant blue R-250. Lane 1, standard molecular weight markers (β -galactosidase 116 kDa, bovine serum albumin 66.2 kDa, ovalbumin 45 kDa, lactate dehydrogenase 35 kDa, REase Bsp98I 25 kDa, β -lactoglobulin 18.4 kDa, lysozyme 14.4 kDa); and Lane 2, purified BmooMP α -I. (B) Photomicrograph of a BmooMP α -I crystal.

analyzed and compared using the program COOT (Emsley and Cowtan, 2004). All the molecular graphic figures were generated using the program PyMOL (DeLano, 2002).

2.5. Flexible molecular docking

Flexible molecular docking was performed using ICM-Pro software (Molsoft, La Jolla, CA) via global optimization of the energy function based on the Monte-Carlo minimization procedure (Abagyan et al., 1994). For each compound a stack of docked complexes with random ligand conformations and orientations was generated. In this study, 7 compounds whose crystal structures complexed with zinc-dependent peptidases (PDB codes 1ZXC, 2A8H, 2FV5, 2FV9, 1KUI, 1KUK and 1KUG) had been determined, were subjected to docking calculations in order to characterize their binding modes in BmooMP α -I structure. ICM tools package was used to convert PDB structures of compounds to a suitable format and to prepare the protein structure. For docking calculations the protein structure was stripped of water molecules and then polar hydrogen atoms were added and Gasteiger charges were calculated for each atom of the macromolecule. Additionally, an electrostatic map and a desolvation map were calculated. Each resulting complex was analyzed on the basis of docking energy and individual three-dimensional evaluation of hydrogen bonding interactions between ligand and protein amino acid residues, and interactions with catalytic zinc ion.

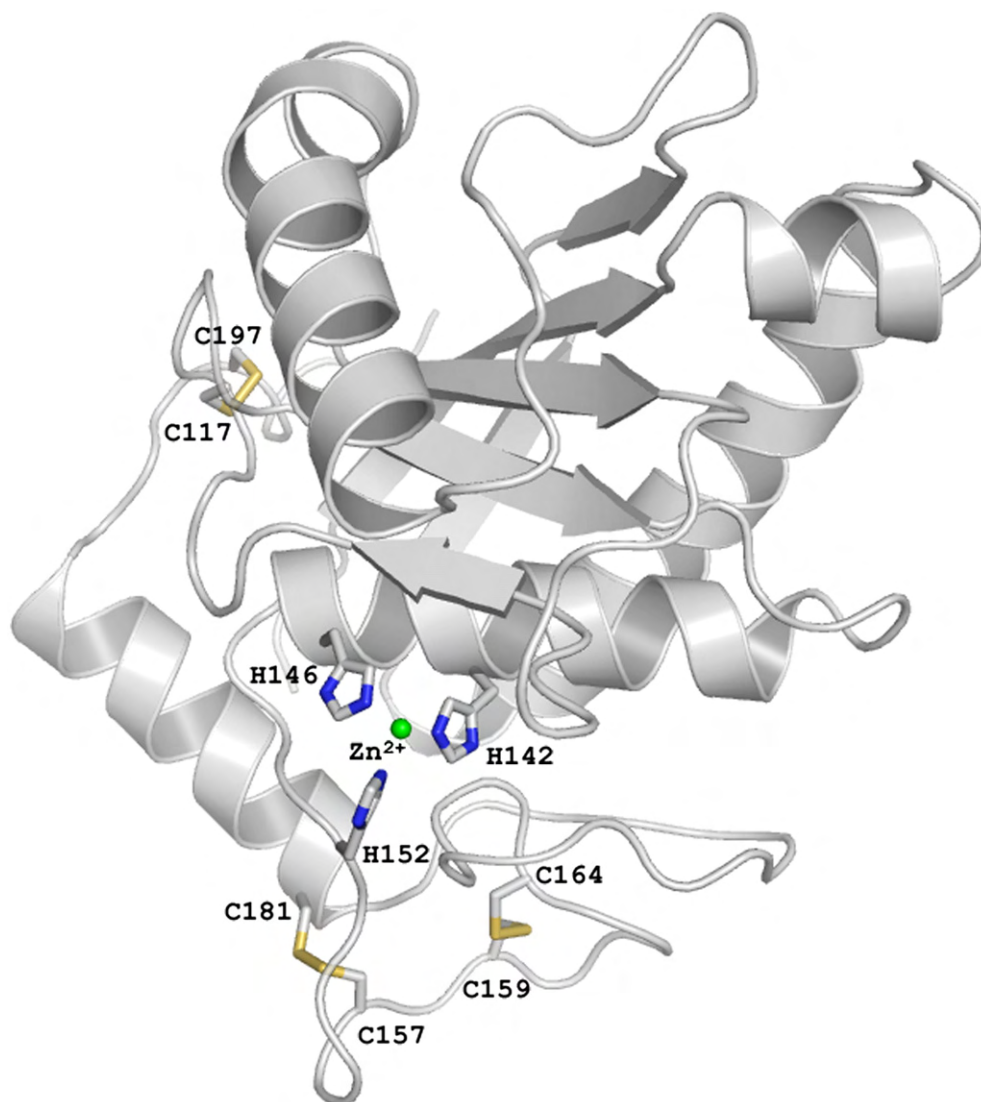


Fig. 2. Schematic cartoon representation of BmooMP α -I structure. The disulphide bridges and three catalytic histidines are shown in stick representation. The zinc ion is drawn as sphere. Carbon, sulfur, nitrogen and zinc atoms are shown in gray, yellow, blue and green, respectively. The atom colors are standard for the following pictures. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.

3. Results and discussion

3.1. Mass-spectrometric analysis

Protein identification was performed by searching in a non-redundant protein database (NCBI) using the MASCOT search engine (<http://matrixscience.com>). The following parameters were used for database searches: one missed cleavage with allowed modifications carbamidomethylation of cysteine (complete) and oxidation of methionine (partial). BmooMP α -I was clearly identified using the MASCOT search engine with 81% overall matched peptides generated by trypsin, chymotrypsin and SV-8 in-gel digestion with individual ions scores >53. The probability of such a match being produced randomly is less than 0.5%.

3.2. Chemical and structural homogeneity

Milligrams of BmooMP α -I sample were obtained by three chromatographic steps as previously published using anionic-exchange, size-exclusion and heparin-agarose affinity chromatographies (Bernardes et al., 2008). The purity was analyzed by polyacrylamide gel electrophoresis under denaturing and reducing conditions and migrated as a single band with approximate molecular weight of 22 kDa (Fig. 1A). In order to check structural homogeneity, dynamic light-scattering (DLS) experiments have been performed at the protein concentration used for the crystallization trials, which demonstrated a monomodal distribution with average polydispersity of 5% and estimated molecular weight corresponding to a monomer in solution. Both parameters used to evaluate the

crystallizability of proteins have indicated a highly favored state for initial crystallization experiments and further crystal optimization.

3.3. Structure determination

Suitable crystals for X-ray diffraction of BmooMP α -I were obtained from a mother solution containing 100 mM Tris-HCl pH 8.3 and 20% PEG 8000 after approximate 8 days at 20 °C (Fig. 1B). BmooMP α -I crystals belong to the monoclinic space group P2₁ with cell dimensions parameters $a = 40.1$, $b = 48.7$, $c = 54.4$ Å and $\beta = 91.56^\circ$. The structure has been determined at 1.76 Å resolution using molecular replacement techniques. Residual electron densities above 3σ in the Fourier difference maps have been observed at the active-site pocket and at the crossover region of the N- and C-termini, which were assigned as zinc and calcium ions, respectively. The refinement converged to a crystallographic residual of 17.5% (R_{free} 21.6%) for all data between 20.8 and 1.76 Å (Table 1). The final model (asymmetric unit) consists of 200 amino acid residues, 151 solvent molecules, one zinc ion and one calcium ion. The model displays excellent overall stereochemistry with RMSD values of 0.023 Å and 1.9° for bond lengths and angles, respectively. The average temperature factor (B-value) for all atoms is 25.7 Å² (Table 1). The Ramachandran plot demonstrates that 98.5% of the dihedral angles are situated in the most favored regions, while the remaining 1.5 % of residues lie in the permitted regions and no residues are found in the additionally allowed and disallowed regions (Table 1).

3.4. Overall structure

The molecular topology of BmooMP α -I conserves all the main features of P-I class SVMs formed by a five-stranded β -sheet, four long helices and one short N-terminal α -helix stabilized by three disulfide bridges (Cys¹¹⁷-Cys¹⁹⁷, Cys¹⁵⁷-Cys¹⁸¹ and Cys¹⁵⁹-Cys¹⁶⁴) located in the C-terminal region. In the oblate ellipsoidal structure, the β -sheet is curved and sandwiched between helices α 2 and α 4 on the concave side and helix α 3 on the convex side (Fig. 2). The active-site cavity is located at the interface of the upper domain (about

150N-terminal residues) and lower domain (about 50C-terminal residues) as commonly observed in other metzincins such as adamalysin-II (Gomis-Rüth et al., 1994), H₂-proteinase (Kumasaka et al., 1996) and acutolysin A (Gong et al., 1998).

As with acutolysin A and adamalysin-II, a calcium ion was identified on the crossover region of the N- and C-termini at the opposite side of the active-site pocket. The calcium ion is heptacoordinated by O δ 1Asn²⁰⁰, OCys¹⁹⁷, O δ 1Asp¹⁹³, O δ 2Asp¹⁹³, O ϵ 1Aps⁹, OW⁵ and OW¹⁰ atoms and has been considered structurally important for the stabilization of the proteolytic domain. No evidence has been found for its role in enzymatic activity (Gomis-Rüth et al., 1994).

3.5. Zinc-binding site

The positions of the catalytically important residues at the active-site cleft are fully conserved including the zinc-binding motif HEXXHXXGXXH and the methionine-turn at the bottom of the catalytic histidines. In the BmooMP α -I structure, the zinc ion displays a novel distorted octahedral coordination formed by the N ϵ 2 atoms of the three catalytic histidines (His¹⁴², His¹⁴⁶ and His¹⁵²), numbering based on adamalysin-II structure, (Gomis-Rüth et al., 1994) and additionally three solvent molecules (OW¹³¹, OW¹⁴³ and OW¹⁴⁴) (Fig. 3A). In other SVMs such as H₂-proteinase, adamalysin-II and acutolysin A, the zinc ion is tetrahedrally coordinated by same three catalytic histidines and only one solvent molecule, which is structurally related to the residue Glu¹⁴³, which polarizes solvent molecules prior to the nucleophilic attack on the scissile peptide bond of a polypeptide chain substrate (Fig. 3B). The corresponding solvent molecule OW¹⁴³ in the BmooMP α -I structure is coordinated by a short bond (distance = 2.79 Å) when compared with other reported metzincins, where typical distances ranging from 3.52 to 4.3 Å have been reported. A different mode of zinc coordination has also been described in the astacin structure in which the typical interaction with a water molecule is replaced by the hydroxyl group of Tyr¹⁶⁹ side chain (Bode et al., 1992) (Fig. 3C).

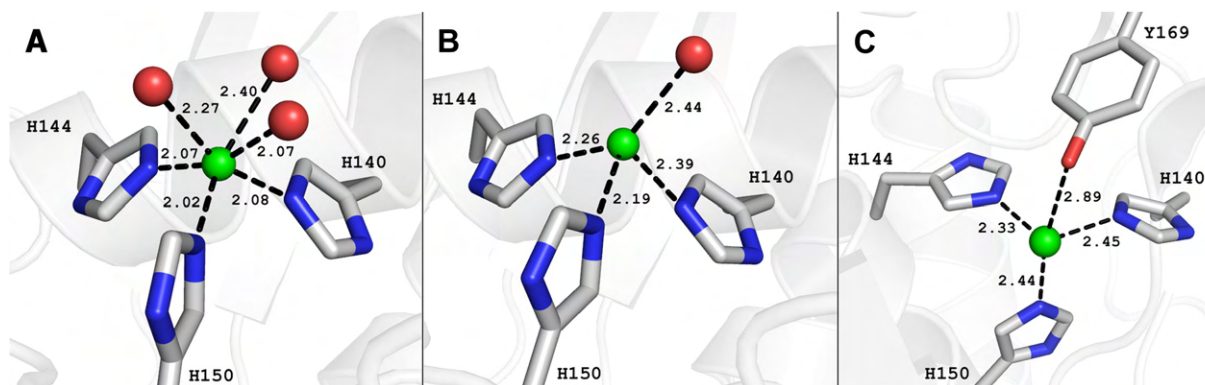


Fig. 3. Coordination geometry of the active-site zinc ion in BmooMP α -I (A), adamalysin-II (B) and astacin (C) structures. The amino acid residues are presented in stick representation and atom colors. The interactions are shown in black dashes.

Table 2

Description of docked ligands and their respective calculated binding energies.

Compound	Chemical name	Molecular formula	Molecular weight (g/mol)	PDBID	Binding-Energy (kcal/mol)
1	pyroGlu-Gln-Trp (pEQW)	C ₂₁ H ₃₂ N ₅ O ₆	450.03	1KUI	−37.27
2	pyroGlu-Lys-Trp (pEKW)	C ₂₂ H ₃₄ N ₅ O ₅	448.01	1KUK	−36.44
3	pyroGlu-Asn-Trp (pENW)	C ₂₀ H ₃₀ N ₅ O ₆	436.0	1KUG	−35.87
4	(2R)-N-Hydroxy-2-[(3S)-3-Methyl-3-[4-[(2-Methylquinolin-4-yl)methoxy]phenyl]-2-oxopyrrolidin-1-yl]propanamide	C ₂₅ H ₂₇ N ₃ O ₄	433.50	2FV5	−34.23
5	3,N(D,L-[2-(Hydroxyamino-Carbonyl)methyl]-4-Methyl Pentanoyl)-L-3-(Tert-Butyl)Glycyl-L-Alanine	C ₁₉ H ₃₇ N ₅ O ₅	415.53	2FV9	−30.88
6	4-({4-[(4-Aminobut-2-ynyl)oxy]phenyl)sulfonyl}-N-Hydroxy-2,2-Dimethylthiomorpholine-3-Carboxamide	C ₁₇ H ₂₃ N ₃ O ₅ S ₂	413.51	2A8H	−28.85
7	(3S)-4-[[4-(But-2-Ynyloxy)phenyl]sulfonyl]-N-Hydroxy-2,2-Dimethylthiomorpholine-3-Carboxamide	C ₁₇ H ₂₂ N ₂ O ₅ S ₂	398.50	1ZXC	−22.92

3.6. Prediction of molecular interactions between BmooMP α -I and its inhibitors

Several well-known inhibitors of SVMPs have been complexed with BmooMP α -I and submitted to crystallization trials, which did not result in suitable crystals for X-ray diffraction. In order to elucidate their mode of interaction to BmooMP α -I, *in silico* flexible molecular docking has been performed. The overall docking results are in good agreement with experimental binding modes observed in crystal structures of adamalysin-related proteinase complexes. Significant conformational changes are owing to amino acid substitutions around the active-site, which results in distinct physicochemical properties in the microenvironment. The root-mean-square deviation (RMSD) between the best docking poses and X-ray crystal structures were 1ZXC 1.937 Å, 2A8H 1.823 Å, 2FV5 3.394 Å, 2FV9 4.815 Å, 1KUI 2.31 Å, 1KUK 1.80 Å and 1KUG 1.99 Å. Description of docked compounds and their respective binding energies is summarized in Table 2. The best binding-energy values were obtained with pEQW tripeptide (−37 kcal mol^{−1}) and its analogous pENW and pEKW (Table 2).

3.7. Structural basis for the hemorrhagic activity

Superpositioning of BmooMP α -I, adamalysin-II and H₂-proteinase structures (non-hemorrhagic P-I class SVMPs) on BaP1, acutolysin A and acutolysin C structures (hemorrhagic P-I class SVMPs) indicated that both subgroups conserve all elements considered essential for catalysis including three catalytic histidines, the polarizer Glu¹⁴³ and methionine-turn, in same relative conformations. Interface analysis with Protein Interfaces, Surfaces and Assemblies web server (Krissinel and Henrick, 2007) reveals that both subgroups display very similar solvent-accessible surface area of 9500 Å² and solvation free energy of around −190 Kcal mol^{−1}. Differences are only significant at adjacent loops (amino acid segments 153–164 and 167–176) to the methionine-turn and at both N- and C-termini due to amino acid residues deletions (Fig. 4A). The variability at the N- and C-termini and the surface

charge distribution have been analyzed and did not provide any information that could differentiate between both subgroups. Thus, using an algorithm based on evolutionary conservation for the identification of functional regions in proteins and extensive structural and sequence comparative analyses, a particular motif formed by the amino acid segments 153–164 and 167–176 adjacent to the methionine-turn display a coherent pattern of amino acid composition, which enable us to differentiate both classes and consequently can be involved in the development of hemorrhagic activity in P-I class SVMPs. In this hypothesis, the non-hemorrhagic SVMPs contain large amino acid residues at this motif that could provoke steric clashes preventing macromolecular substrate recognition, whereas in hemorrhagic SVMPs this motif is populated by small polar amino acid residues (Ser and Thr) enabling the binding and recognition of targets in the process of hemorrhagic induction. Previous studies have also marked the segment 153–176 as the holder of the main structural differences between P-I class SVMPs with very different hemorrhagic activities (Watanabe et al., 2003), which is in agreement with our observations.

3.8. ADAMTS and SVMPs: shared structural features

Multiple alignment and three-dimensional structure superpositioning indicate that BmooMP α -I displays high sequence and structural similarity with ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) family proteins than other metzincins. The human ADAMTS proteins consist of 19 related genes and are involved in several biological processes such as cell adhesion, blood coagulation, inflammation and angiogenesis. Disorders in ADAMTS genes have been associated with various diseases, including Ehlers–Danlos syndrome, osteoarthritis, and thrombosis and may also play a role in several types of cancer. Structural analysis showed that the overall fold and catalytically-relevant residues are fully conserved in both P-I class SVMPs and ADAMTS including methionine-turn, catalytic histidines, Glu¹⁴³ residue and disulphide bridges. The most significant differences in the catalytic

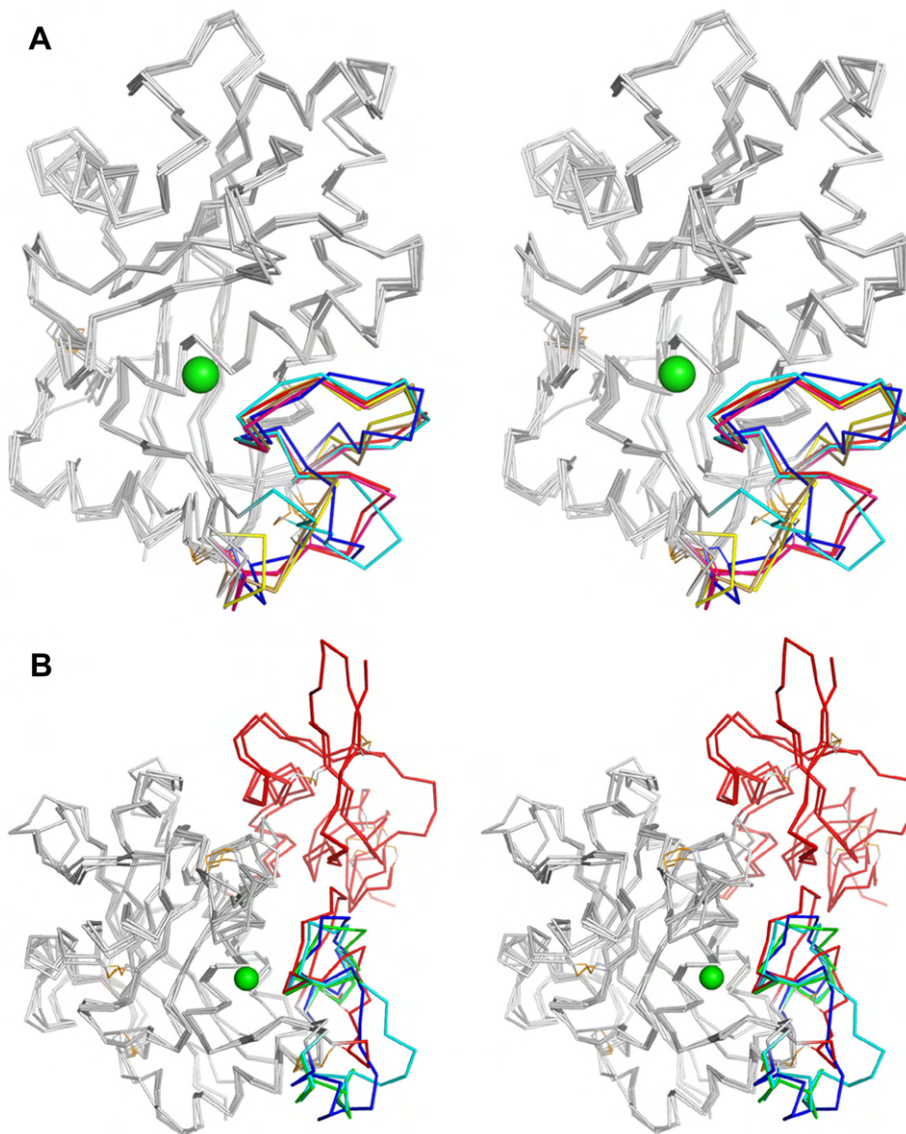


Fig. 4. Stereo images of $C\alpha$ backbone representations of comparative studies between SVMPs and ADAMTS. The zinc ion is displayed as a green sphere in both panels. (A) Results of the superpositioning of BmooMP α -I structure (red) on other SVMPs including adamalysin-II (1IAG, cyan), H₂-proteinase (1WNI, yellow), BaP1 (1ND1, pink), acutolysin A (1BSW, blue) and acutolysin C (1QUA, sand). (B) Results of the superpositioning of BmooMP α -I structure (light red) on ADAMTS family proteins including ADAMTS-1 (2JIH, green), ADAMTS-4 (2RJP, cyan) and ADAMTS-5 (3B8Z, blue). The desintegrin domain of ADAMTS is colored in red.

domain are concentrated at the adjacent loops to the methionine-turn, which adopt a completely different conformation due to the diversity of amino acid compositions and at the interfacial 105-loop that displays different lengths depending on the source (Fig. 4B). These observations support the hypothesis that the segment 153–176 could be involved in the specificity and macromolecular recognition of adamalysin-related proteinases. However, further studies are required to elucidate the exact role of this segment in the functioning of SVMPs, ADAMs and ADAMTS.

4. Concluding remarks

The crystal structure of a new non-hemorrhagic P-I class SVMP from *B. moojeni* venom has been determined at

1.76 Å resolution and its structure revealed a novel distorted octahedral coordination of the catalytic zinc ion. Comparative analyses of a number of non- and hemorrhagic P-I class SVMPs structures suggest that the motif comprising the amino acid segments 153–164 and 167–176 flanking the methionine-turn is responsible for the functional differentiation between these subgroups of P-I class SVMPs. In our hypothesis, we propose that large side-chains at this segment in non-hemorrhagic SVMPs could restrict macromolecular interaction and binding, whereas in hemorrhagic SVMPs this motif is populated by small polar amino acid residues. Additional structural comparisons with ADAMTS have also demonstrated significant differences in adjacent loops to the methionine-turn among the broad family of adamalysin-related

proteinases, which supports our observations. Flexible molecular docking studies have also been performed with experimentally known inhibitors of metzincins indicating the favorable binding of peptidic inhibitors to the BmooMP α -I active-site pocket, which could serve as model to understand human adamalysin-related proteins.

Acknowledgements

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 471192/2007-4) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, 98/14138-2). MTM and RKA acknowledge support from FAPESP, CNPq, CAPES and DAAD.

Protein Data Bank accession code

The atomic coordinates and structure factors have been deposited with the Worldwide Protein Data Bank and are available under accession code 3GBO.

Conflict of interest

None declared.

References

- Abagyan, R., Totrov, M., Kuznetsov, D., 1994. ICM: a new method for structure modeling and design: applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* 15, 488–506.
- Bernardes, C.P., Santos-Filho, N.A., Costa, T.R., Gomes, M.S., Torres, F.S., Costa, J., Borges, M.H., Richardson, M., Santos, D.M., Castro Pimenta, A. M., Homsí-Brandeburgo, M.I., Soares, A.M., Oliveira, F., 2008. Isolation and structural characterization of a new fibrin(ogen)olytic metalloproteinase from *Bothrops moojeni* snake venom. *Toxicon* 51, 574–584.
- Bode, W., Gomis-Rüth, F.X., Huber, R., Zwilling, R., Stöcker, W., 1992. Structure of astacin and implications for activation of astacins and zinc-ligation of collagenases. *Nature* 358, 164–167.
- Bode, W., Gomis-Ruth, F.X., Stockler, W., 1993. 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.* 331, 134–140.
- CCP4 Collaborative Computational Project Number 4, 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* 50, 760–763.
- DeLano, W.L., 2002. The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, USA.
- Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* 60, 2126–2132.
- Fox, J.W., Serrano, S.M.T., 2005. Structural considerations of the snake venom metalloproteinases, key members of the M12 repolysin family of metalloproteinases. *Toxicon* 45, 969–985.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R. D., Bairoch, A., 2005. Protein identification and analysis tools on the ExPASy server. In: Walker, J.M. (Ed.), *The Proteomics Protocols Handbook*. Human Press, New York, pp. 571–607.
- Gomis-Rüth, F.X., Kress, L.F., Kellermann, J., Mayr, I., Lee, X., Huber, R., 1994. Refined 2.0 Å X-ray crystal structure of the snake venom zinc-endopeptidase adamalysin II. Primary and tertiary structure determination, refinement, molecular structure and comparison with astacin, collagenase and thermolysin. *J. Mol. Biol.* 239, 513–544.
- Gong, W., Zhu, X., Liu, S., Teng, M., Niu, L., 1998. Crystal structures of acutolysin A, a three-disulfide hemorrhagic zinc metalloproteinase from the snake venom of *Agkistrodon acutus*. *J. Mol. Biol.* 283, 657–668.
- Gutiérrez, J.M., Rucavado, A., Escalante, T., Díaz, C., 2005. Hemorrhage induced by snake venom metalloproteinases: biochemical and biophysical mechanisms involved in microvessel damage. *Toxicon* 45, 997–1011.
- Hellman, U., Wernstedt, C., Góñez, J., Heldin, C.H., 1995. Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Anal. Biochem.* 224, 451–455.
- Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797.
- Kumasaka, T., Yamamoto, M., Moriyama, H., Tanaka, N., Sato, M., Katsube, Y., Yamakawa, Y., Omori-Satoh, T., Iwanaga, S., Ueki, T., 1996. Crystal structure of H2-proteinase from the venom of *Trimeresurus flavoviridis*. *J. Biochem.* 119, 49–57.
- Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of head of bacteriophage T4. *Nature* 227, 680–685.
- Laskowski, R.A., Moss, D.S., Thornton, J.M., 1993. Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* 231, 1049–1067.
- Moura-da-Silva, A.M., Butera, D., Tanjoni, I., 2007. Importance of snake venom metalloproteinases in cell biology: effects on platelets, inflammatory and endothelial cells. *Curr. Pharm. Des.* 13, 2893–2905.
- Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* 53, 240–255.
- Navaza, J., 1994. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* 50, 157–163.
- Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326.
- Perrakis, A., Morris, R., Lamzin, V.S., 1999. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* 6, 458–463.
- Stocker, W., Bode, W., 1995. Structural features of a superfamily of zinc-endopeptidases: the metzincins. *Curr. Opin. Struct. Biol.* 5, 383–390.
- Watanabe, L., Shannon, J.D., Valente, R.H., Rucavado, A., Alape-Girón, A., Kamiguti, A.S., Theakston, R.D.G., Fox, J.W., Gutiérrez, J.M., Arni, R.K., 2003. Amino acid sequence and crystal structure of BaP1, a metalloproteinase from *Bothrops asper* snake venom that exerts multiple tissue-damaging activities. *Protein Sci.* 12, 2273–2281.
- White, J., 2005. Snake venoms and coagulopathy. *Toxicon* 45 (8), 951–967.

Manuscript Title: Dissecting the structure-function-stability relationships of family 5A cellulase: portrayal of a novel ion-based mechanism for thermal stabilization and a new family 3 carbohydrate-binding module lacking the calcium-binding site

Manuscript No: JBC/2011/241075

Manuscript Type: Regular Paper

Date Submitted by the Author: 17 Mar 2011

Complete List of Authors: Camila Ramos Santos, Joice Helena Paiva, Mauricio L. Sforca, Jorge L. Neves, Rodrigo Z. Navarro, Junio Cota, Patricia K. Akao, Andreia N. Meza, Igor Polikarpov, Jose Xavier-Neto, Fabio M. Squina, Richard J. Ward, Roberto Ruller, Ana Carolina Zeri, and Mario Tyago Murakami

Keywords: Enzyme mechanisms; Enzyme structure; NMR; Protein stability; X-ray crystallography ; Carbohydrate-binding module; Cellulases

DISSECTING THE STRUCTURE-FUNCTION-STABILITY RELATIONSHIPS OF FAMILY 5A CELLULASE: PORTRAYAL OF A NOVEL ION-BASED MECHANISM FOR THERMAL STABILIZATION AND A NEW FAMILY 3 CARBOHYDRATE-BINDING MODULE LACKING THE CALCIUM-BINDING SITE

Camila Ramos dos Santos^{1,#}, Joice Helena Paiva^{1,#}, Maurício L. Sforça¹, Jorge L. Neves¹, Rodrigo Z. Navarro¹, Júnio Cota², Patrícia K. Akao¹, Andréia N. Meza¹, Igor Polikarpov³, José Xavier-Neto¹, Fábio M. Squina², Richard J. Ward⁴, Roberto Ruller², Ana Carolina Zeri¹, Mário Tyago Murakami*

¹Laboratório Nacional de Biociências, Centro Nacional de Pesquisas em Energia e Materiais, Campinas, Brazil. ²Centro de Tecnologia do Bioetanol, Centro Nacional de Pesquisas em Energia e Materiais, Campinas, Brazil. ³Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, SP, Brazil.

⁴Departamento de Química, Universidade de São Paulo, Ribeirão Preto, Brasil.

#Both authors have contributed equally to this work

Running head: Function, structure and stability of thermophilic cellulases 5A

Address correspondence to: Mário Tyago Murakami, Giuseppe Máximo Scolfaro, 10000, Campinas, SP, Brasil. Zip code: 13083-970. Tel: +55-19-35121106. Fax: +55-19-35121006.

E-mail: mario.murakami@lnbio.org.br

Cellulases participate directly or indirectly in a number of physiological events such as cell-wall remodeling in plants and bioremediation in bacteria. Their ability to depolymerize crystalline cellulose is of a great biotechnological interest for environmentally compatible production of fuels from lignocellulosic biomass. However, industrial use of cellulases is somewhat limited both by their low catalytic efficiency and stability. In this work, we conducted a detailed functional and structural characterization of the thermophilic family 5A cellulase from *Bacillus subtilis* (BsCel5A), which consists of a α/β barrel catalytic core fused to a family 3 carbohydrate-binding module (CBM3) by a 23-residue-long linker. Phylogenetic and NMR structural analyses revealed that the *Bacillus* CBM3 motifs represent a new subfamily, which lack the classical calcium-binding motif. Despite its negative effect on the enzyme stability, the accessory domain plays a pivotal role in the catalysis. Together with the catalytic core, the CBM forms a large planar surface for cellulose recognition, which conducts the substrate in a proper conformation to the active site and increases the enzymatic efficiency. Notably, the manganese metal ion has a hyper-stabilization effect on BsCel5A and by using deletion constructs we determined that this stabilizing effect maps to the catalytic core. Crystallographic studies unveiled that this novel mechanism is based on the tight

hexacoordination of the manganese metal ion by negatively charged residues located at the opposite face of the catalytic site.

The production of ethanol from lignocellulosic biomass is well-placed among other possibilities to produce energy, owing to its potential sustainability and agro-economic benefits (1). Different chemical and enzymatic strategies have been proposed for the saccharification of lignocellulosic biomass (2, 3). The later has been implemented at different organization levels, from point mutations of cellulolytic enzymes to engineering of the whole metabolic pathways in microorganisms (4). Furthermore, these two approaches have been employed synergistically to increase the yields of fermentable sugars from lignocellulosic biomass.

Although enzymatic hydrolysis has been successfully implemented as biomass-to-bioenergy technology, it is widely regarded as an expensive and wasteful link of the whole chain of bioethanol production mainly because of the need to use large amounts of enzymes to compensate their low catalytic efficiency and stability (5, 6). Driven by this need, a number of enzymes with biological and commercial value have been systematically modified through structure-based rational approaches, or by intelligent use of serendipity (7, 8, 9). In spite of these efforts, the molecular basis underlying the protein stability and its intricate correlation with catalysis is still elusive, putting a premium on studies regarding structure-function-stability relationships and the discovery of new

alternative pathways to improve thermal stability of the enzymes.

A key point concerning the increase of the catalytic efficiency of cellulases is the presence of accessory domains named as carbohydrate-binding modules (CBMs). These non-catalytic parts of the enzymes promote the disruption of the cellulose structure, thus increasing the substrate accessibility and delivering it to the catalytic domain in a configuration favored for catalysis (reviewed in 10, 11, 12). CBMs, found in enzymes from different life forms, can be appended to N- or/and C-terminus of the parental protein as a single or multiple copies. Despite previous extensive functional and structural studies of CBMs, the consequences of a fusion of a given CBM to a specific glycoside hydrolase family (GH) member to its function and stability remain to be established.

In this work, we addressed the structural basis of thermal stability and catalytic efficiency of a thermophilic cellulase 5A from the bacterium *Bacillus subtilis*. We describe a novel mechanism of thermal stabilization based on metal ion coordination and dissect the role of the family 3 CBM in the function and stability of cellulases 5A. This in-depth understanding of the modulation of enzyme stability and function at molecular level is crucial for its optimization and consequent use in the bioethanol generation from lignocellulosic biomass into a realistic economically-viable process.

EXPERIMENTAL PROCEDURES

Cloning, expression and purification – Three constructions of BsCel5A were amplified by PCR using *B. subtilis* 168 genomic DNA and specific oligonucleotides (5'-GCTAGCGCAGCAGGGACAAAACG-3' and 5'-GGATCCTTAGGTGCCGAGAATGTTTC-3' for CC; 5'-GCTAGCATTCTGTACAGTACAGAGCA-3' and 5'-GGATCCCTAATTTGGTTCTGTCCCC-3' for CBM; 5'-GCTAGCGCAGCAGGGACAAAACG-3' and 5'-GGATCCCTAATTTGGTTCTGTCCCC-3' for WT). The amplified sequences were cloned into pGEM-T Easy vector (Promega) and further subcloned into the pET28a expression vector using the restriction sites present in the oligonucleotides (underlined sequence). The proteins were expressed in BL21(DE3)*slyD* cells at 37°C for 4 hours after

induction with 0.5 mM IPTG in LB medium. For NMR experiments, the BL21(DE3)*slyD* cells transformed with pET28a-CBM plasmid were grown in M9 minimal medium supplemented with 2 g/l ¹³C-glucose and 1 g/l ¹⁵N-ammonium chloride (13) using the same conditions aforementioned. The harvested cells were resuspended in lysis buffer (50 mM sodium phosphate pH 7.2, 100 mM NaCl, 1 mM PMSF, 5 mM benzamidine) and disrupted by lysozyme treatment (80 µg/ml, 30 min, on ice), followed by sonication (Vibracell VCX 500, Sonics & Materials, Inc.). The solutions were centrifuged at 10,000g for 30 min and the supernatants were loaded onto nickel-charged 5 mL HiTrap Chelating columns (GE Healthcare) using a flow rate of 1 ml/min. The proteins were eluted using a non-linear gradient of 0-500 mM imidazole. For WT construction, an anionic-exchange chromatography using a 5 mL HiTrap SP HP column, flow rate of 1 ml/min and a non-linear gradient of 0-1 M NaCl (GE Healthcare) was used to improve sample purity. All constructions were further concentrated using Ultrafree-CL (Millipore) and submitted to size-exclusion chromatography using a Superdex 75 16/60 column (GE Healthcare) and flow rate of 0.5 ml/min. All chromatographic steps were performed using an ÄKTA FPLC system (GE Healthcare). Sample homogeneity was confirmed by polyacrylamide gel electrophoresis under denaturing conditions (14). Protein concentration was estimated by absorbance at 280 nm using molar extinction coefficients derived from constructs sequence.

Enzymatic characterization and optimization of endo-β-1,4-glucanase activity – The standard enzymatic assays for CC and WT constructs evaluation were performed in triplicate using different substrates at 0.25%(w/v) and 100 ng of enzyme in 50 mM citrate-phosphate buffer, pH 6.0. The mixture was incubated at 60°C for 60 min to determine activity in a set of 12 polysaccharides including β-glucan, carboxymethylcellulose, larch arabinogalactan, laminarin, galactomannan, xyloglucan, debranched arabinan, 1,4-β-D-mannan, rye arabinoxylan, xylan from beechwood, pectin and lichenan (purchased from Megazyme, Ireland, and Sigma–Aldrich, USA). The enzymatic activity was determined by the amount of reducing

sugar liberated from different polysaccharides using the DNS method (15). One unit (U) of endoglucanase activity was defined as the quantity of enzyme to release 1 μmol of reducing sugars per minute under standard conditions. A Response Surface Methodology was performed to optimize the reaction conditions for both CC and WT constructs of BsCel5A in absence or presence of manganese ion (10 mM). The variables analyzed here were the pH together with temperature. A Central Composite Design ($k = 2$) with four central points, totalizing 12 experiments was considered for optimization of these variables. Details concerning the statistical approaches for these experiments can be found in Myers and Montgomery (16). The regression and graphical analysis of the data were performed using the software Statistica™ version 8.0, considering a *p*-value of 0.05. The significance of the regression coefficients was given by Student's *t*-test. The second order model equation was determined by Fisher's test and the multiple coefficient of determination (R^2) gave the variance explained by the model. All glucanase activity assays were carried out in triplicate using 10 μl of enzyme (100 ng), 40 μl of 0.1 M citrate-phosphate buffer and 50 μl of β -glucan (stock 0.5%(w/v) in water). The enzymatic reaction was performed at different pHs in citrate-phosphate buffer system and various temperatures as indicated in the [supplemental Table S1](#). Metal ions, chelating agent (EDTA), surfactants (SDS and Triton X-100), and organic solvents (methanol, ethanol and DMSO) were added to the enzyme assay to test their effects on the activity. The kinetic study was performed with β -glucan as substrate at 60°C and pH 6.0 according to the method described above. The kinetic parameters k_{cat} and K_{m} were determined by a non-linear regression fit of the saturation curves using the Michaelis–Menten equation.

Circular dichroism spectroscopy and thermal unfolding studies – Far-UV CD spectra (190-260 nm) were recorded on a Jasco J-810 spectropolarimeter (Jasco International Co.) coupled to a Peltier temperature controller using a 1 mm quartz cuvette. Proteins were used in a concentration of 10 μM . The data collection parameters were set to scan rate of 50 nm/min, response time of 4 s, sensitivity of 100 mdeg, accumulation of 10, heating rate of 1°C/min and

delay time for spectrum collection of 60 s. Results were expressed as mean residue ellipticity ($\text{deg cm}^2 \text{ dmol}^{-1} \text{ residue}^{-1}$). All thermal unfolding experiments were monitored at 222 nm, except for the CBM construct, which was monitored at 230 nm.

Small angle X-ray scattering – SAXS data for CC, CBM and WT constructs were collected using a 165 mm MarCCD detector on the D02A/SAXS2 beamline at the Brazilian Synchrotron Light Laboratory. The radiation wavelength was set to 1.48 Å and the sample-to-detector distance was set to 1559.1 mm to give a scattering vector-range from 0.01 nm^{-1} to 0.23 nm^{-1} ($q = 4*\pi*\sin\theta/\lambda$, where 2θ is the scattering angle). The samples at 2, 4 or 8 mg/ml, in 150 mM NaCl and 20 mM Tris-HCl (pH 7.5), were centrifuged for 10 min at 20,000g (4°C) and then filtered to remove any aggregates. Measurements were carried out at room temperature ($\sim 23^\circ\text{C}$). Buffer scattering was recorded before each sample scattering and then subtracted from the corresponding sample scattering. The integration of SAXS patterns were performed using Fit2D software (17) and the curves were scaled by protein concentration. Data analysis was performed using GNOM (18). Low resolution envelopes of each construction were calculated from the experimental SAXS data using *ab initio* procedure implemented in DAMMIN program (19). Averaged models were generated from 10 runs using the suite of programs DAMAVER (20). The SAXS envelope and the NMR or XTAL structure were superimposed using the program SUPCOMB 2.0 (21). BUNCH software, which employs rigid body model (RBM) and simulated annealing routines, was used to find the best relative positions of the NMR and the XTAL structures (22). This solution was superposed in the SAXS envelope with SUPCOMB 2.0 (21). CRY SOL was used to calculate the scattering curve from envelopes and from high resolution structures including the WT model obtained from BUNCH (23).

Nuclear Magnetic Resonance Spectroscopy – NMR experiments were performed using a Varian Inova spectrometer at the National Biosciences Laboratory (LNBio/CNPEM), operating at a ^1H Larmor frequency of 599.887 MHz and temperature of 293 K. For these experiments, the

spectrometer was equipped with a triple resonance cryogenic probe and a Z pulse-field gradient unit. The ^{15}N - ^{13}C labeled CBM sample was dissolved in 300 μ phosphate buffer pH 7.2 containing 5% (v/v) D_2O , at a final concentration of approximately 0.3 mM. Water suppression was achieved by low-power continuous wave irradiation over the relaxation delay or using the WATERGATE method (24). All data were processed using NMRPipe and NMRVIEW software (25, 26). Prior to Fourier transformation, the time domain data were zero-filled in all dimensions. When necessary, a fifth-order polynomial baseline correction was applied after transformation and phasing. To obtain distance constraints, cross-peak volumes were estimated from the Noesy spectra. Protein backbone resonance peaks were assigned using ^1H - ^{15}N HSQC and standard triple resonance experiments including HNCACB, CBCACONH, HNCOC, HNCACO (27) of the ^{15}N - ^{13}C labeled CBM sample. All spectra for resonance assignment were recorded by setting up a resolution of 2048X128X96 points and spectral window of 8.00 kHz, 12.07 kHz and 1.95 kHz for ^1H , ^{13}C and ^{15}N , respectively. In order to assign the side chain resonances, the ^{15}N - ^{13}C labeled CBM sample was lyophilized and dissolved in 100% D_2O , followed by acquisition of HCCH-TOCSY, hCCH-TOCSY spectra. NOE-derived distance restraints were obtained from the ^{15}N -HSQC-NOESY and ^{13}C -HSQC-NOESY (separately optimized for aliphatics and aromatics) both collected with 80 ms mixing time. The structure of the CBM domain was calculated in a semi-automated iterative manner with the program CYANA version 2.1 (28), using 100 starting conformers. CYANA 2.1 protocol was applied to calibrate and assign NOE cross-peaks. After the first few rounds of automatic calculations, the NOESY spectra were analyzed again to identify additional cross-peaks consistent with the structural model and to correct misidentified NOEs. The structures obtained were further refined by restrained minimization and molecular dynamic studies using the CNS software (29). The 20 structures with the lowest target function were selected to represent the ensemble of protein structures. The quality of the structures was analyzed with PROCHECK-NMR (30). The NMR data was deposited at the BMRB data bank under

code 17399 and Protein Data Bank under code 2I8A.

Proton/Deuterium exchange – Proton-deuterium (H/D) exchange measurements were carried out with the ^{15}N - ^{13}C labeled CBM sample dissolved in 100% D_2O . In order to monitor the intensity loss of labile protons, a series of ^1H - ^{15}N -HSQC spectra were collected first at 15 minutes (time required to dissolve the protein and perform the first experiment) and subsequently in 40 minutes time intervals over two days. Additionally, a last spectrum was collected 10 days after the protein dissolution.

Backbone ^{15}N - ^1H relaxation experiments – Dynamic regions of the protein were mapped by measurements of spin-lattice relaxation rates (R_1), spin-spin relaxation rates (R_2), rotating frame relaxation rates ($R_{1\rho}$) and steady-state heteronuclear NOE at two magnetic fields, 11.7 T and 14.1 T. R_1 experiments were conducted by setting relaxation delays to 10, 210, 410, 810, 1010, 1210, 1410, 1610, 1810, 2010 ms. $R_{1\rho}$ and R_2 experiments were conducted by setting relaxation delays to 10, 30, 50, 70, 90, 110, 130, 170, 210, 250 ms. Peak intensity decays in function of relaxation delays points are fitted to mono-exponentials by using the rate analysis function of the NMRView Software (26). Steady-state heteronuclear NOE measurements were acquired with proton saturation of 3 s and recycle delay of 5 s while for measurements without proton saturation a recycle delay of 8 s was employed.

X-ray crystallography – The highly purified CC protein was concentrated to 75 mg/ml and its buffer was exchanged to water by ultra filtration using Vivaspin6 centrifugal filters (GE Healthcare). Crystallization experiments were performed by vapor diffusion method using a HoneyBee 963 robot (Genomic Solutions). Sitting drops were prepared by mixing 0.5 μl of the protein solution with an equal volume of mother liquor and equilibrated against 80 μl of the reservoir at 18°C. Formulations based on the following commercial crystallization kits were tested: SaltRX, Crystal Screen, Crystal Screen 2 (Hampton Research), Precipitant Synergy, Wizard I and II (Emerald BioSystems), PACT and JCSG+

(Qiagen/Nextal), totalizing 544 solutions. Large crystals grew within 45 days in the following conditions: 100 mM bis-tris-propane pH 6.5, 20%(w/v) PEG 3350 and 200 mM sodium nitrate (form I); 100 mM SPG (succinic acid, sodium dihydrogen phosphate monohydrate, glycine) pH 6.0 and 25%(w/v) PEG 1500 (form II). Crystals were transferred to new drops containing the mother liquor and 20%(v/v) glycerol for cryoprotection and then flash-cooled in a 100 K nitrogen gas stream. In order to obtain an ion-protein complex, crystals belonging to form II were soaked for 3 min in a cryoprotectant solution containing 50 mM of manganese chloride. X-ray diffraction data were collected at the W01B-MX2 beamline (Brazilian Synchrotron Light Laboratory, Campinas, Brazil). Data processing was carried out with HKL2000 (31) and for molecular replacement was performed with MolRep (32).

RESULTS AND DISCUSSION

Accessory domain modulates enzyme stability and kinetics – Wild-type protein and its truncated domains were expressed in *E. coli* cells and purified by nickel-affinity and size-exclusion chromatographies (supplemental Fig. S1). The thermal stability of the BsCel5A constructions was assessed by unfolding studies using CD spectroscopy (Fig. 1). As expected, CD spectra of the catalytic core (CC) and CBM3 displayed characteristic profiles of α/β and all- β proteins, respectively (Fig. 1A). BsCel5A and its truncated forms showed a single structural transition in the thermal unfolding experiments (Fig. 1B), despite the fact that BsCel5A is a two-domain protein and two transitions corresponding to each domain were expected. Probably, the 23-residue-long linker and inter-domains contacts drive the thermal denaturation to a single-transition process. Interestingly, unfolding studies revealed that the deletion of the accessory domain increases thermal stability from 64°C to 78°C. Structural analysis suggests that the short linker could restrict the freedom degree between the domains. Then, the molecular events triggered in the CBM3 may directly reflect in the CC via linker and inter-domain interactions. Additionally, the loss of CBM3 promotes the expose of polar residues at the CC surface, which becoming favorably hydrated, increasing its stability (33). As a

compensation for its negative contribution to the enzyme stability, CBM increases the catalytic efficiency by altering the enzyme affinity to the substrate (K_m) (Table 1). The wild-type protein has an extended substrate binding surface owing to the strategically 3D arrangement of CBM in relation to CC. This type of disposition favors additional protein-carbohydrate interactions, which reflects in a higher substrate affinity. It may explain the lower turnover of the wild-type enzyme (V_{max}). However, the gain in K_m compensates the lower V_{max} , and globally CBM enhances the catalytic efficiency of enzyme (k_{cat}/K_m). Moreover, biochemical characterization shows that the CBM deletion does not alter the substrate specificity (supplemental Fig. S2) and optimum conditions for catalysis (supplemental Fig. S3). These data clearly indicate the importance of CBM3 in the regulation of overall enzyme function.

BsCel5A CBM3 is a member of novel calcium-independent subfamily – In order to gain insights into its functional relationships, we determined the NMR structure of the CBM3. The ^1H - ^{15}N HSQC spectrum of CBM showed good signal dispersion (supplemental Fig. S4), indicating that the protein is folded and permitting to readily assign the backbone and side-chain resonances. The CBM domain of BsCel5A folds into a β -sandwich, a common scaffold for CBMs, encompassing a core formed by two β -sheets of four and five β -strands, and one short α -helix at the C-terminal region (Fig. 2A). A classification of CBMs, taking into account the structural and functional similarities, has grouped these modules in types A, B and C (10). The type A binds crystalline cellulose and chitin by a flat surface, the type B binds single-chain polysaccharides by an extended groove or cleft, and the type C binds mono, di or trisaccharides. The BsCel5ACBM belongs to a type A due to its ability to bind crystalline cellulose (34) and higher structural similarity with type A (PDB code 1NBC) than other type B (PDB code 1J84) or type C (PDB code 1I82) CBMs. The most common classification, based on sequence similarities, has classified the BsCel5A CBM as a member of the CBM3 superfamily (<http://www.cazy.org/Carbohydrate-Binding-Modules.html>). Canonically, the CBM3 members with solved structures have a highly conserved calcium-binding site, which intriguingly is not

present in BsCel5A CBM3 structure (Fig. 2B and 2C). So far, five members of CBM3 have been solved, sharing from 13 to 36% of identity with the BsCel5A CBM3. Structural comparisons showed that the β -sandwich core is conserved between all of them but the loops located at the molecule extremities are quite different in both composition and conformation (Fig. 2B). Search for proteins sharing sequence similarity with BsCel5ACBM in the GenBank has resulted in a set of proteins that do not have the structural requirements for calcium binding (supplemental Fig. S5). All these proteins are from *Bacillus* spp., suggesting that ancestor of this bacterium genus has lost the CBM3 calcium binding site during the evolution process. Therefore, we propose a new classification of CBM3 into 2 groups: the calcium-binding proteins, which include the already defined subfamilies CBM3a, CBM3b, CBM3b' and CBM3c; and the calcium-independent proteins, termed herein as CBM3d subfamily, with BsCel5A CBM3 being its first structurally characterized member. The classification of CBM3s in a, b and c refers to the type of substrate that is preferentially recognized by the CBMs (35). The a and b groups bind crystalline cellulose, a is specific of cellulosomal scaffolding proteins and b is found in soluble cellulases. Recently, it was shown that certain CBM3b members are not able to bind neither crystalline nor soluble cellulose and those were designated as CBM3b' (36). The c group binds single chain cellulose and it is found in cellulosomal cellulases. Thus, the new CBM3d subfamily is closer related to the CBM3a subfamily owing its substrate preference; however lacking the calcium-binding site, which seems to have a structural role.

Polar residues play a pivotal role in oligosaccharide recognition by CBM3d – The molecular surface of the CBMs has been associated with the type of substrate that they can bind, e.g. the CBM3a presents a very flat surface that seems to make sense for the crystalline cellulose binding. Despite the fact that BsCel5A binds crystalline cellulose – such as Avicel – and the binding only occurs when the CBM is present (34), the CBM surface is not plane as found in CBM3a members (Fig. 3). The lack of flatness of the BsCel5ACBM surface can be explained primarily by the absence of the CBM3a β 4' strand

and second by the large structural differences in surface loops (Fig. 3). However, the exact protein-carbohydrate interactions are not known, since structures of a CBM3a or even a type A CBM are not available. Thus, in order to map the residues involved in the recognition of cellulose, we analyzed the variations in the nuclear magnetic resonance frequencies of the BsCel5ACBM in the presence of cellotetraose. The major perturbations in chemical shifts have occurred in Asn463 and Arg455, and minor changes were observed in Gln369, Gln453 and Asp407 (Fig. 2D). The residues Arg455, Gln369, Gln453 and Asp407 are conserved in the CBM of CipA from *C. thermocellum* (CtCipACBM, PDB code 1NBC), a well characterized type A CBM (37). The Tyr408 is also conserved but it did not show a chemical shift, perhaps due to short length of the cellotetraose, which would preferentially bind the neighboring residues of the cellulose interaction surface. The molecular recognition and affinity towards carbohydrates is achieved by van der Waals interactions arising from stacking of aromatic residues with the furanose/pyranose rings and by hydrogen bonds between polar groups of saccharides and amino acid side chains. The mapped interactions of CBM from BsCel5A with cellotetraose indicate the carbohydrate binding face of this domain, which is rich in charged residues, such as Gln, Arg, Asp, Lys and Asn (results not shown). The only aromatic residue in this face is the Tyr408, while a set of aromatic residues is found pointing to the opposite side, making interactions in the core of the β -sandwich. These findings indicate that the interactions of BsCel5ACBM with the cellulose are mainly by polar side chains. The polar residues may not only work on the carbohydrate affinity/recognition but could specifically replace inter-chain hydrogen bonds, acting to disrupt the tight cellulose structure (37). Therefore, the CBM residues might provoke local separation of single chain cellulose to be delivered to the CC, thus contributing to higher catalytic efficiency of the enzyme in the presence of CBM.

Three-dimensional domain architecture of Cel5A reveals a flat cellulose-binding platform – Wild-type BsCel5A (BsCel5AWT) and its separated domains were analyzed by SAXS, showing that all forms are monomers in solution with R_g values of

1.75, 1.63 and 3.04 nm, and D_{\max} values of 4.65, 4.8, and 9.5 nm, for the CC, CBM and WT (supplemental Fig. S6 and Table 2). The CC and CBM were nicely fitted into the corresponding envelopes (Fig. 4A). The good quality of SAXS data allowed to observe the cavity corresponding to the active site of the enzyme. The averaged low-resolution model of the wild-type protein calculated from 20 independent *ab initio* models gives a good notion of the three-dimensional arrangement of the accessory domain with respect to its parental catalytic domain (Fig. 4B). The molecular envelope reveals a large continuous planar surface, which is essential for crystalline cellulose binding. The atomic structures of both domains fitted into the SAXS envelope of the wild-type protein showed that this planar surface is formed by the active-site interface of the CC in line with the cellulose recognition face of CBM. This model is in a full agreement with the crystallographic structures of the Cel9G from *Clostridium cellulolyticum* and E4 from *Thermomonospora fusca* (Fig. 4C). Interestingly, the SAXS envelope is continuous and it does not present a kind of “neck” in the linker region, indicating that, as expected from spectroscopic data, the CC and CBM interact to each other. Indeed, analytical size-exclusion chromatography showed that an equimolar mixture of the CC and CBM eluted as a single peak (results not shown) indicating that these domains interact in a linker-independent manner.

Manganese metal ion hyper-stabilizes the catalytic core of BsCel5A – Divalent ions are well-known cofactors of enzymes either activating their catalytic machinery or having a structural role. In this work, we discovered that manganese ion drastically improves the thermal stability of BsCel5A. Mn^{2+} ion elevated the T_m of BsCel5A on 23°C, from 64°C to 87°C (Fig. 5B). In order to map the domains involved in the manganese ion interactions, truncated forms of the protein were analyzed. As shown in Fig. 1B, the CC had its thermal stability improved by the Mn^{2+} ion, whereas that of CBM remained unaltered, mapping the specific interactions of Mn^{2+} ion with the catalytic core. Other 10 different divalent ions were tested and none of them affected the stability, which demonstrates that family 5A cellulases have

a very specific motif for recognition of manganese ions. .

Structural basis of cel5A stabilization by manganese ion – In order to identify the manganese-binding site and to delineate the structural basis for the enhanced thermal stability of BsCel5A by this ion, we undertook an X-ray crystallographic study. BsCel5ACC was crystallized in two different crystalline forms (form I and form II) that diffracted up to 2.90 Å and 2.10 Å, respectively. However, only the crystal form II was successfully soaked with manganese (Table 3). Crystals belonging to form I displays a very particular crystallographic symmetry in which the manganese-binding site is occluded by an aspartic residue of the neighboring molecule (Fig. 5D), which prevented the ion incorporation by soaking methods. The statistical details of data collection and structure refinement are given in Table 2. CC structure encompasses a classical $(\beta/\alpha)_8$ -barrel fold (Fig. 5A) with the active site located at the crevice formed by the C-terminal ends of β -strands. The catalytic machinery involves two acid residues Glu229 and Glu169, the aromatic gate keepers Trp207 and Trp291, the polarizer His229, Arg92 and Thr256 (Fig. 5B). From homology to other GH5 cellulases, the Glu169 is expected to be the proton donor, which protonates the glycosidic bond and catalyzes the bond fission, and the Glu257 is the nucleophile, which facilitates the reaction by stabilizing the resulting carbonium ion intermediate. Thus, in the resting state the carboxylated catalyst must be protonated and the nucleophile must be unprotonated. Several studies demonstrated the key role of histidine residue in the protonation of the acid/base glutamate residue, which is controlled by a catalytic triad, similar to the serine proteases (38). In this scheme, a third element, which could be a serine/threonine, an aspartic acid or a water molecule, would protonate the histidine, permitting to classify the glycosyl hydrolases accordingly to this histidine proton donor (38). BsCel5A contains a threonine residue in the equivalent position which allows to classify it as a type A catalytic triad. Cel5As from *Thermotoga maritima* and *Clostridium thermocellum* have a water molecule in the corresponding position and are classified as type C enzymes (supplemental Fig. S7). Water-dependent

catalytic machineries are labile at alkaline conditions, since the water molecule would deprotonate at high pH, disrupting the hydrogen network required for catalysis. On the other hand, enzymes with type A catalytic triad would remain active even in basic conditions, which is advantageous for biotechnological purposes. In the manganese-complex structure, the ion is located at the opposite face of the active site, near to the lid-like β -hairpin (Fig. 5C) and is hexacoordinated by two negatively charged residues Asp195 and Asp197, the polar Asn198, the carbonyl oxygen from Gly157 and two water molecules (Fig. 5D). The strategic location of the manganese-binding site permits simultaneous water-mediated interactions between the N-terminal and distant regions of the protein, causing increase of structure rigidity and consequently a notable gain in the protein stability. Comparison between forms II (without manganese ion) and II* (with manganese ion) shows that the ion only substituted by a water molecule in the highly negatively-charged hole without affecting neither the cleft architecture nor the conformational states of catalytically-relevant residues. This result also suggests that the region comprising the lid-like β -hairpin, which is subsequent to the N-terminal end, can be a strategic motif to be engineered aiming to improve protein stability. Indeed, in the analogous hyperthermostable Cel5A from *Thermotoga maritima* (TmCel5A, 39), the β -hairpin is replaced by an α -helical element, which permits additional polar contacts supporting the fact that the stabilization of this region may implicate in higher stability.

Inhibitory effect of manganese ion is by a non-competitive mechanism – Despite the remarkable gain in the protein stability by the manganese ion, it also influences the enzymatic activity. Analysis of substrate preference was carried out in the presence and absence of manganese ion indicating no differences on the selectivity; however, the enzyme displayed a decrease of approximately 40% in its total activity in the presence of Mn^{2+} ion (supplemental Fig. S2). In our crystallographic studies, we observed that the manganese ion binds to an acidic pocket that is very similar to the active site suggesting a possible secondary site involving the catalytic glutamates. In fact, structural studies with TmCel5A showed that a divalent (cadmium)

ion was found interacting with the catalytic glutamic acid residues (40). Therefore, to test if there is a competition between the manganese ion and the substrate for the active we have performed a kinetics study in the presence or absence of manganese ion. The results showed that the presence of the Mn^{2+} ion did not interfere neither in the substrate affinity (K_m) (Table 2) nor in optimal conditions for reaction (supplemental Fig. S3). This implies that the manganese ion does not bind to the active site discarding the possibility of an inhibitory effect by a competitive mechanism. However, the presence of the ion decreased the V_{max} reflecting in a lower enzymatic efficiency (Table 2). Thus, we concluded that the manganese ion causes a non-competitive inhibition. We speculate that this effect is a consequence of higher overall molecular rigidity which prevents conformational changes required for optimal catalysis. This hypothesis is supported by previous studies that described the pivotal role of ligand-induced conformational alterations in the catalytic activity of family 5A cellulases (40).

Biotechnological implications – Modern biotechnology lies on the interface between basic and applied science demanding continuously novel enzymes with improved catalytic efficiency and stability under extreme and non-natural conditions. Cellulases have many industrial applications including food processing, finishing of textiles and generation of cellulosic bioethanol, one of the sustainable energy sources. These industrial processes require cellulolytic activity under a broad range of pH, temperature, and ionic settings. In this context, the understanding of enzymes at a molecular level and the mechanisms to improve their stability and to modify their catalytic efficiency and specificity is essential for rational structure-based protein engineering aiming to ameliorate their use in biotechnological processes.

In this work we have delineated the roles exerted by the accessory domain of family 5 cellulases and revealed a novel mechanism for thermal stabilization by a combination of X-ray crystallography, NMR, SAXS, molecular biology, biophysical and enzymatic studies. We revealed that BsCel5A CBM3 interacts with CC which decreases its thermal stability but improves its catalytic efficiency. The manganese ion interacts specifically with CC and has the opposite effects.

It significantly enhances the thermal stability of the protein but simultaneously increases its structural rigidity that diminishes the enzymatic activity. The wild type protein in presence of manganese has similar efficiency than the CC in absence of manganese, however being considerably more thermostable. These data

indicate that the negative effect of manganese on the catalytic efficiency of the enzyme can be counterbalanced by the presence of CBM whereas the negative effect of CBM on the thermal stability can be fully surpassed by the presence of manganese ions.

REFERENCES

- Farrell, A. E., Plevin, R. J., Turner, B. T., Jones, A. D., O'Hare, M., and Kammen, D. M. (2006) *Science* **311**, 506–508.
- Lynd, L. R., Laser, M. S., Bransby, D., Dale, B. E., Davison, B., Hamilton, R., Himmel, M., Keller, M., McMillan, J.D., Sheehan, J., and Wyman, C.E. (2008) *Nat. Biotechnol.* **26**, 169–172.
- Binder, J. B., and Raines, R. T. (2010) *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4516–4521.
- Shaw, A. J., Podkaminer, K. K., Desai, S. G., Bardsley, J. S., Rogers, S. R., Thorne, P. G., Hogsett, D. A., and Lynd, L. R. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13769–13774.
- Margeot, A., Hahn-Hagerdal, B., Edlund, M., Slade, R., and Monot, F. (2009) *Curr. Opin. Biotechnol.* **20**, 372–380.
- Wen, F., Nair, N. U., and Zhao, H. (2009) *Curr. Opin. Biotechnol.* **20**, 412–419.
- Heinzelman, P., Snow, C. D., Wu, I., Nguyen, C., Villalobos, A., Govindarajan, S., Minshull, J., and Arnold, F. H. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5610–5615.
- Moraís, S., Heyman, A., Barak, Y., Caspi, J., Wilson, D. B., Lamed, R., Shoseyov, O., and Bayer, E. A. (2007) *J. Biotechnol.* **147**, 205–211.
- van der Veen, B. A., Skov, L. K., Potocki-Véronèse, G., Gajhede, M., Monsan, P., and Remaud-Simeon, M. (2006) *FEBS J.* **273**, 673–681.
- Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) *Biochem. J.* **382**, 769–781.
- Shoseyov, O., Shani, Z., and Levy, I. (2006) *Microbiol. Mol. Biol. Rev.* **70**, 283–295.
- Hashimoto, H. (2006) *Cell Mol. Life Sci.* **63**, 2954–2967.
- Li, M. X., Corson, D. C., and Sykes, B. D. (2002) *Methods Mol. Biol.* **173**, 255–265.
- Laemmli, U. K. (1970) *Nature* **227**, 680–685.
- Miller, G. L. (1959) *Anal. Chem.* **31**, 426–428.
- Myers, R. H., and Montgomery, D. C. (2002). *Response Surface Methodology*, second ed, John Wiley & Sons, New York.
- Hammersley, A. P., Brown, K., Burmeister, W., Claustre, L., Gonzalez, A., McSweeney, S., Mitchell, E., Moy, J. P., Svensson, S. O., and Thompson, A. W. (1997) *J. Synchrotron Radiat.* **4**, 67–77.
- Svergun, D. I. (1992) *J. Appl. Crystallogr.* **25**, 495–503.
- Svergun, D. I. (1999) *Biophys. J.* **76**, 2879–2886.
- Volkov, V. V., and Svergun, D. I. (2003) *J. Appl. Crystallogr.* **36**, 860–864.
- Kozin, M. B., and Svergun, D. I. (2000) *J. Appl. Crystallogr.* **34**, 33–41.
- Petoukhov, M. V., and Svergun, D. I. (2005) *Biophys. J.* **89**, 1237–1250.
- Svergun, D. I., Barberato, C., Koch, M. H. J. (1995) *J. Appl. Crystallogr.* **28**, 768–773.
- Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* **2**, 661–665.
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6**, 277–293.
- Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR* **4**, 603–614.
- Sattler, M., Schleucher, J., and Griesinger, C. (1999) *Prog. Nucl. Mag. Res. Sp.* **34**, 93–158.
- Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol.* **273**, 283–298.
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. -S., Kuszewski, J., Nilges, M., Panuu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, L. G. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921.

30. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8**, 477–486.
31. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326.
32. Vagin, A., and Teplyakov, A. (2010) *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25.
33. Vogt, G., Woell, S., and Argos, P. (1997) *J. Mol. Biol.* **269**, 631–643.
34. Wang, Y., Yuan, H., Wang, J., and Yu, Z. (2009) *Bioresour. Technol.* **100**, 345–349.
35. Bayer, E. A., Chanzy, H., Lamed, R., and Shoham, Y. (1998) *Curr. Opin. Struct. Biol.* **8**, 548–557.
36. Jindou, S., Xu, Q., Kenig, R., Shulman, M., Shoham, Y., Bayer, E. A., and Lamed, R. (2006) *FEMS Microbiol. Lett.* **254**, 308–316.
37. Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. (1996) *EMBO J.* **15**, 5739–5751.
38. Shaw, A., Bott, R., Vornrhein, C., Bricogne, G., Power, S., and Day, A. G. (2002) *J. Mol. Biol.* **320**, 303–309.
39. Pereira, J. H., Chen, Z., McAndrew, R. P., Sapra, R., Chhabra, S. R., Sale, K. L., Simmons, B. A., and Adams, P. D. (2010) *J. Struct. Biol.* **172**, 372–329.
40. Varrot, A., Frandsen, T. P., Driguez, H., and Davies, G. J. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58**, 2201–2204.

FOOTNOTES

This research was supported by Grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

The main abbreviations used are: Cel5A, Cellulase 5A; BsCel5A, Cellulase 5A from *Bacillus subtilis*; GH, glycoside hydrolase family; CC, catalytic core; CBM, carbohydrate binding module; WT, wild type; CD, circular dichroism; XTAL, crystal, SAXS, small angle X-ray scattering.

FIGURE AND TABLE LEGENDS

Figure 1: Biophysical characterization of BsCel5A constructions. (A) Far-UV CD spectra and (B) thermal denaturation in the absence (dashed lines) or in the presence of manganese (solid lines). Curves from CC, CBM and WT are shown in red, blue and black, respectively.

Figure 2: NMR solution structure of BsCel5ACBM. (A) View of the ensemble of 20 NMR solution structures, with β -strands in cyan, α -helix in magenta and loops in gray. (B) Structural comparison of BsCel5ACBM and CBM3 of CipA from *C. thermocellum* (CtCipACBM, PDB code 1NBC) with structural differences colored in blue and red, respectively, and the Ca^{2+} of CtCipACBM shown in green. (C) Calcium binding site of CtCipACBM (wheat) and the corresponding region in the BsCel5ACBM structure (light blue). (D) BsCel5ACBM residues involved in carbohydrate recognition: the residues that are conserved in CtCipACBM and were identified by chemical shift are in green, residues conserved but not identified are in yellow and residues identified but not conserved are in cyan.

Figure 3: Structural comparison of BsCel5ACBM with CBM3a, b' and c. The CBMs of CipA (CBM3a, PDB code 1NBC), E4 (CBM3c, PDB code 4TF4) and Cel9V (CBM3b', PDB code 2WOB) were previously superposed in the BsCel5ACBM structure and are shown as cartoon representations (left). The main differences of these proteins with the BsCel5ACBM, at the carbohydrate binding face, were colored in orange, cyan, and pink, respectively. The corresponding regions in BsCel5ACBM are colored in green. The molecular surfaces of the CipA (orange), E4 (cyan), Cel9V (pink) and BsCel5ACBM (green) are shown as side and frontal views (right).

Figure 4: SAXS analysis of the BsCel5A constructions. (A) SAXS envelopes from CC (left) and CBM (right) are shown as transparent surfaces and the fitted XTAL or NMR structures are shown as cartoon representations. (B) The CC and CBM domains arrangement into the SAXS envelope of WT. (C) The structure of E4 from *T. fusca* (light blue) was superposed in the BsCel5A model (gray); the CC bound carbohydrate of E4 is shown in green and the chemical shifted residues of BsCel5ACBM are shown in yellow.

Figure 5: Crystal structure of BsCel5ACC. (A) Representation of the protein structure as cartoon and the manganese ion as sphere. (B) Zoom of the active site showing the catalytic residues in yellow, other residues in green and two water molecules as red spheres. (C) Electrostatic surface colored by charge, from red (-) to blue (+), indicating the presence of the manganese ion in an acidic groove. (D) Interactions of protein residues (gray) and water molecules (red spheres) with manganese ion (green sphere). The 2.9 Å structure (pink) has an aspartic acid residues of a neighboring molecule (blue) occluding the manganese binding site.

Table 1 – Kinetic parameters of BsCel5A and its truncated catalytic core.

Table 2 – Summary of SAXS calculations.

Table 3 – Data collection and refinement statistics.

Table 1

	CC		WT	
	0 mM Mn ²⁺	10 mM Mn ²⁺	0 mM Mn ²⁺	10 mM Mn ²⁺
V_{\max}	143.7 ± 6.6	92.3 ± 5.9	46.2 ± 3.1	35.7 ± 1.8
K_m	1.4 ± 0.2	1.6 ± 0.3	0.32 ± 0.2	0.5 ± 0.2
k_{cat}	78.7	50.5	40.4	31.2
k_{cat}/K_m	56.2	31.6	127.8	58.8

Table 2

	CC			CBM			WT		
	exper [*]	envel [#]	xtal [◇]	exper [*]	envel [#]	NMR [◇]	exper [*]	envel [#]	assem [◇]
R _g (nm)	1.75	1.73	1.86	1.63	1.63	1.65	3.04	2.96	2.87
D _{max} (nm)	4.65			4.80			9.50		
Chi [‡]		2.13	2.89		1.89	2.75		2.37	2.01

^{*} calculated from experimental data

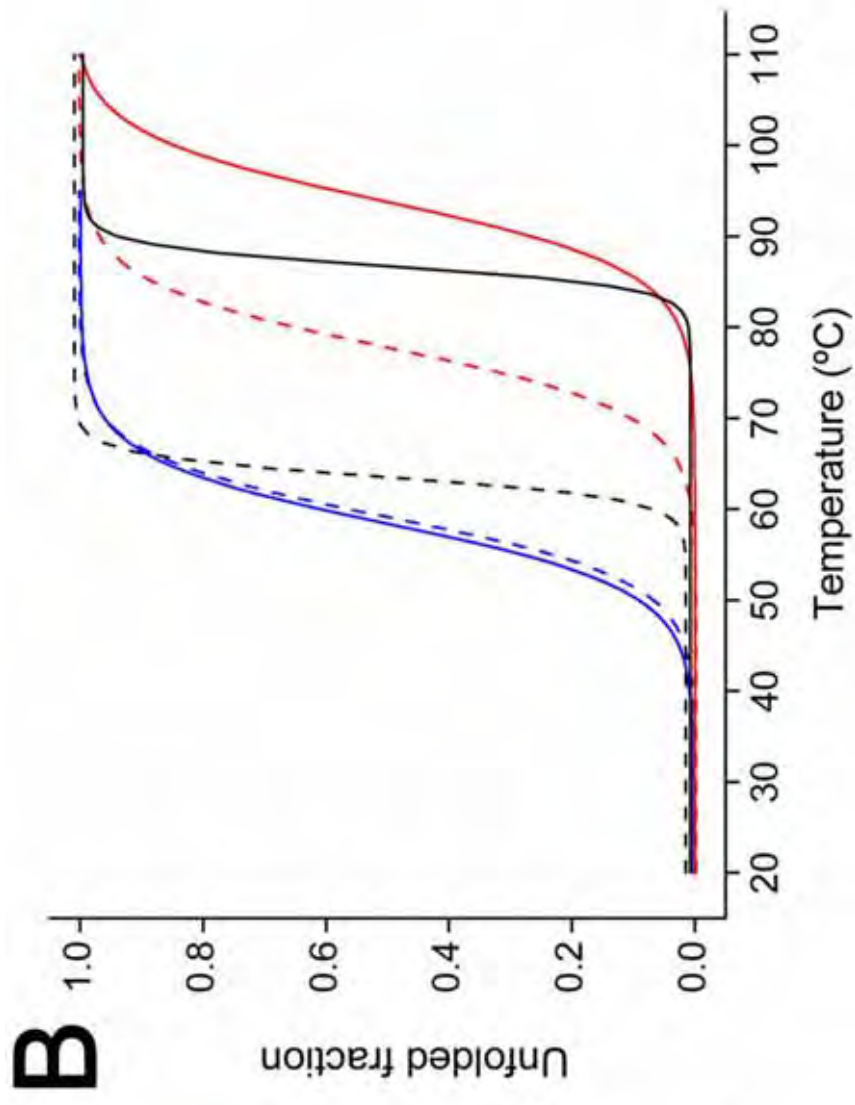
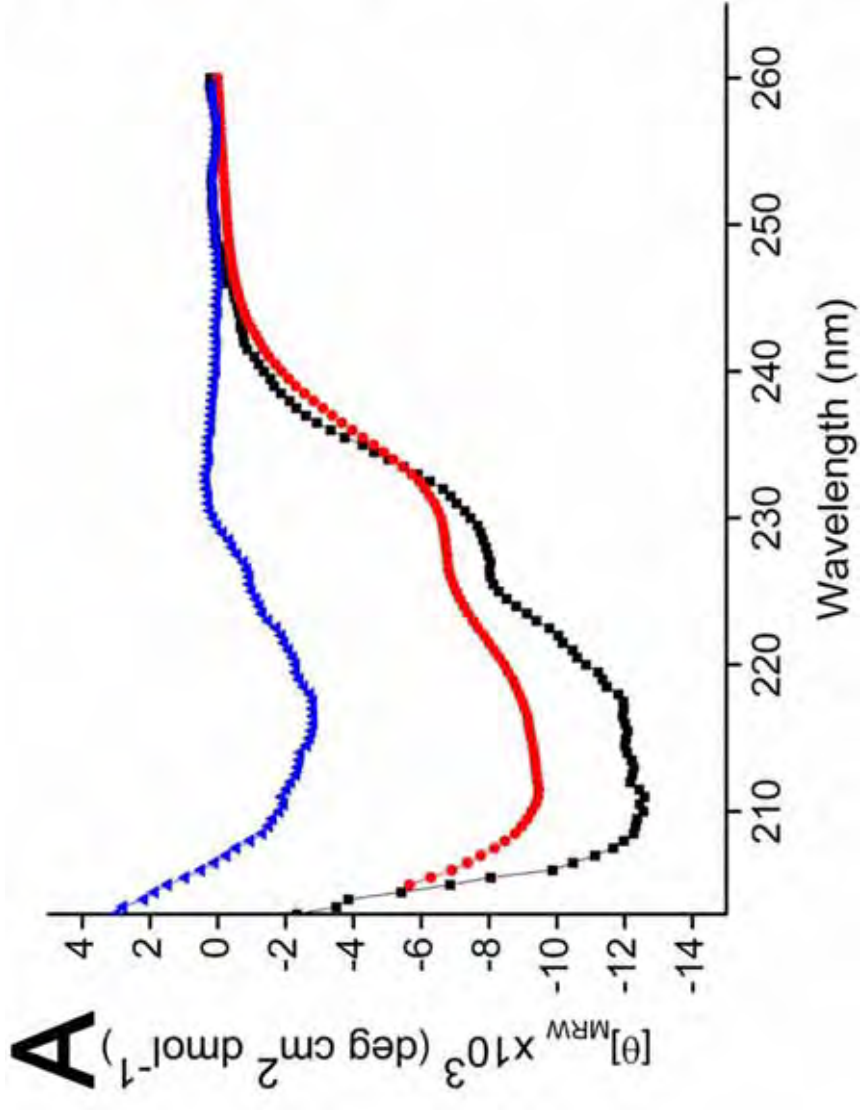
[#] calculated from the SAXS-derived envelope

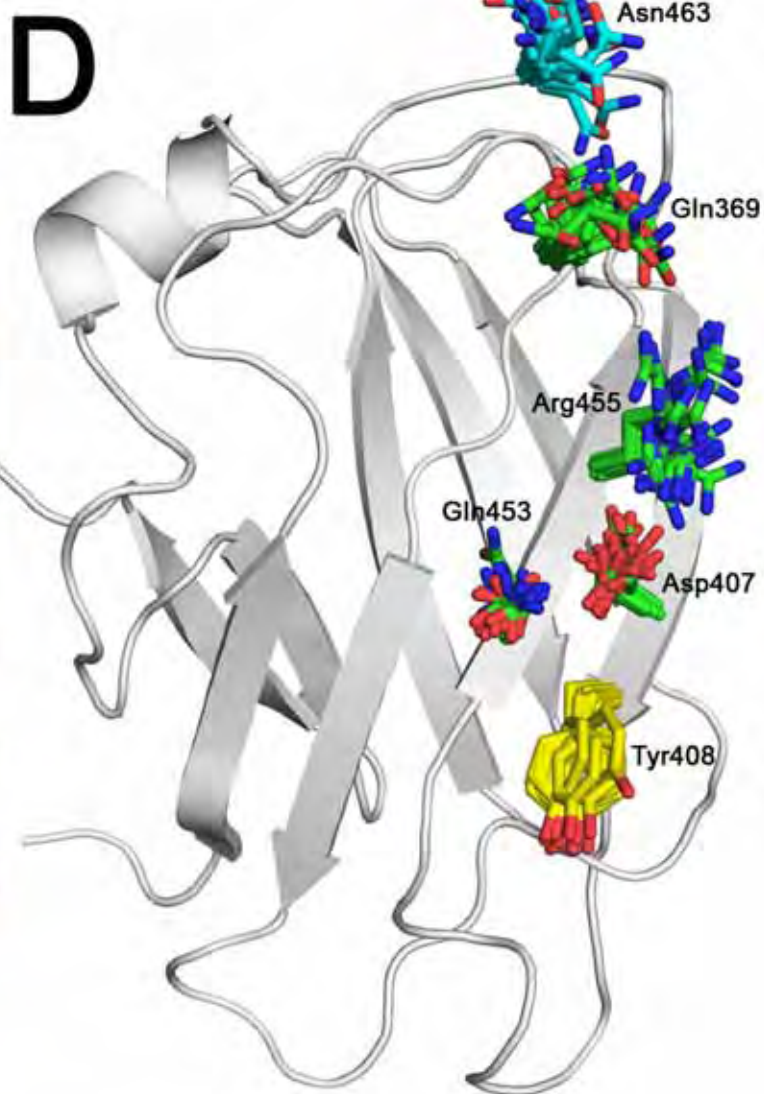
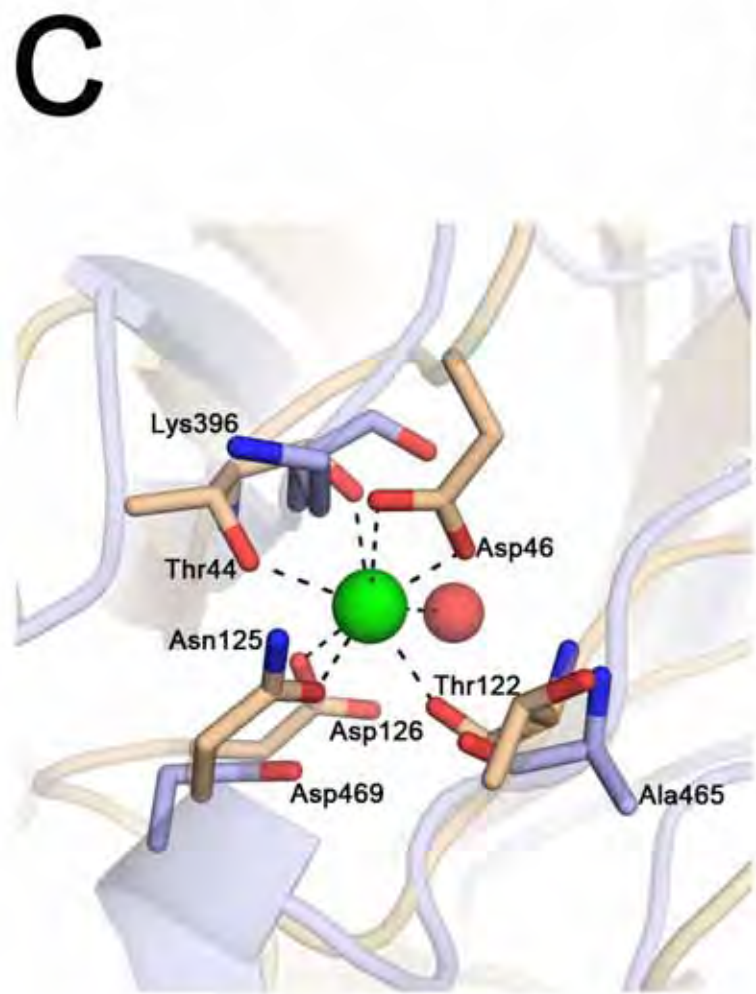
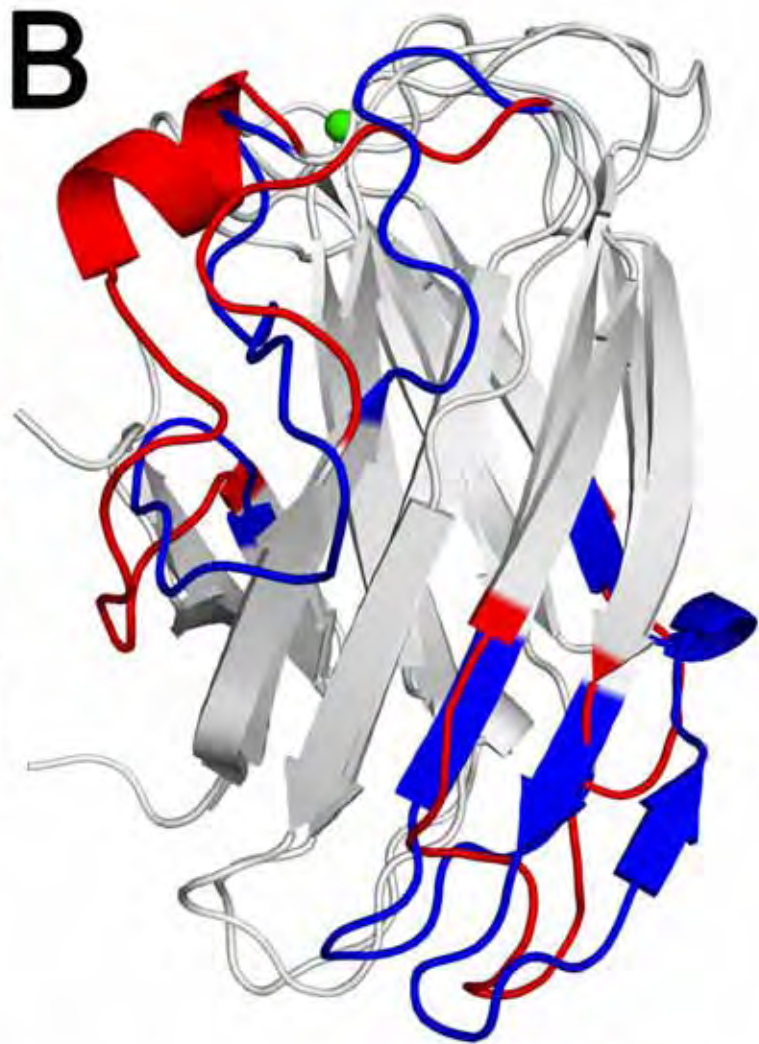
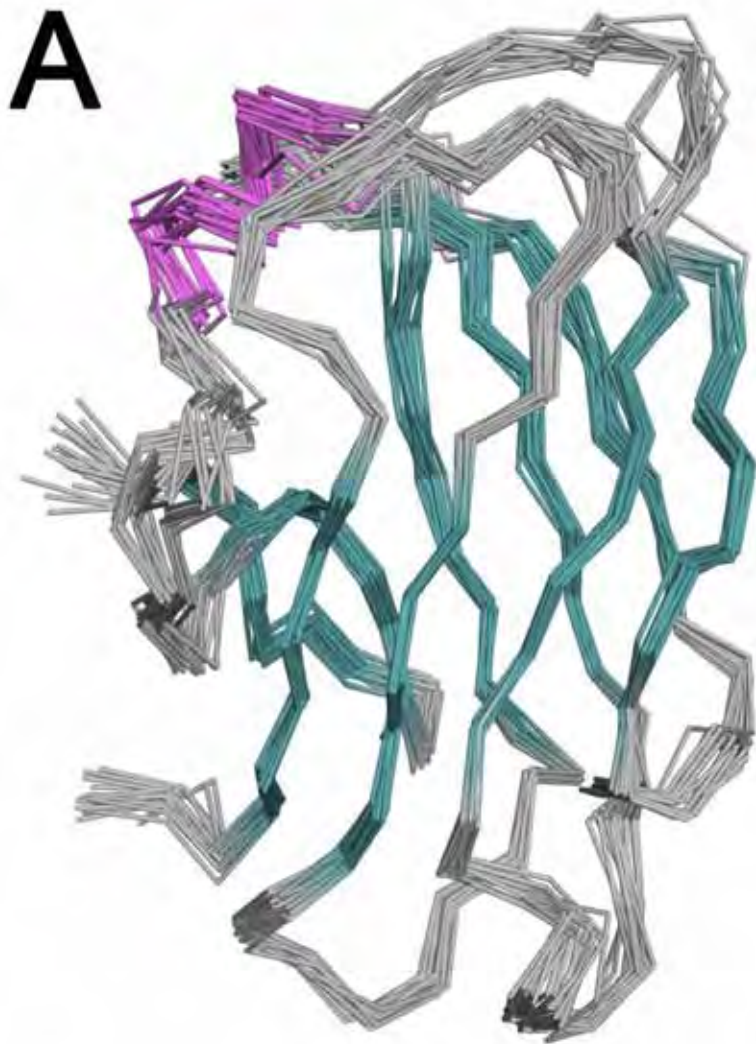
[◇] calculated from the XTAL or NMR or assembly

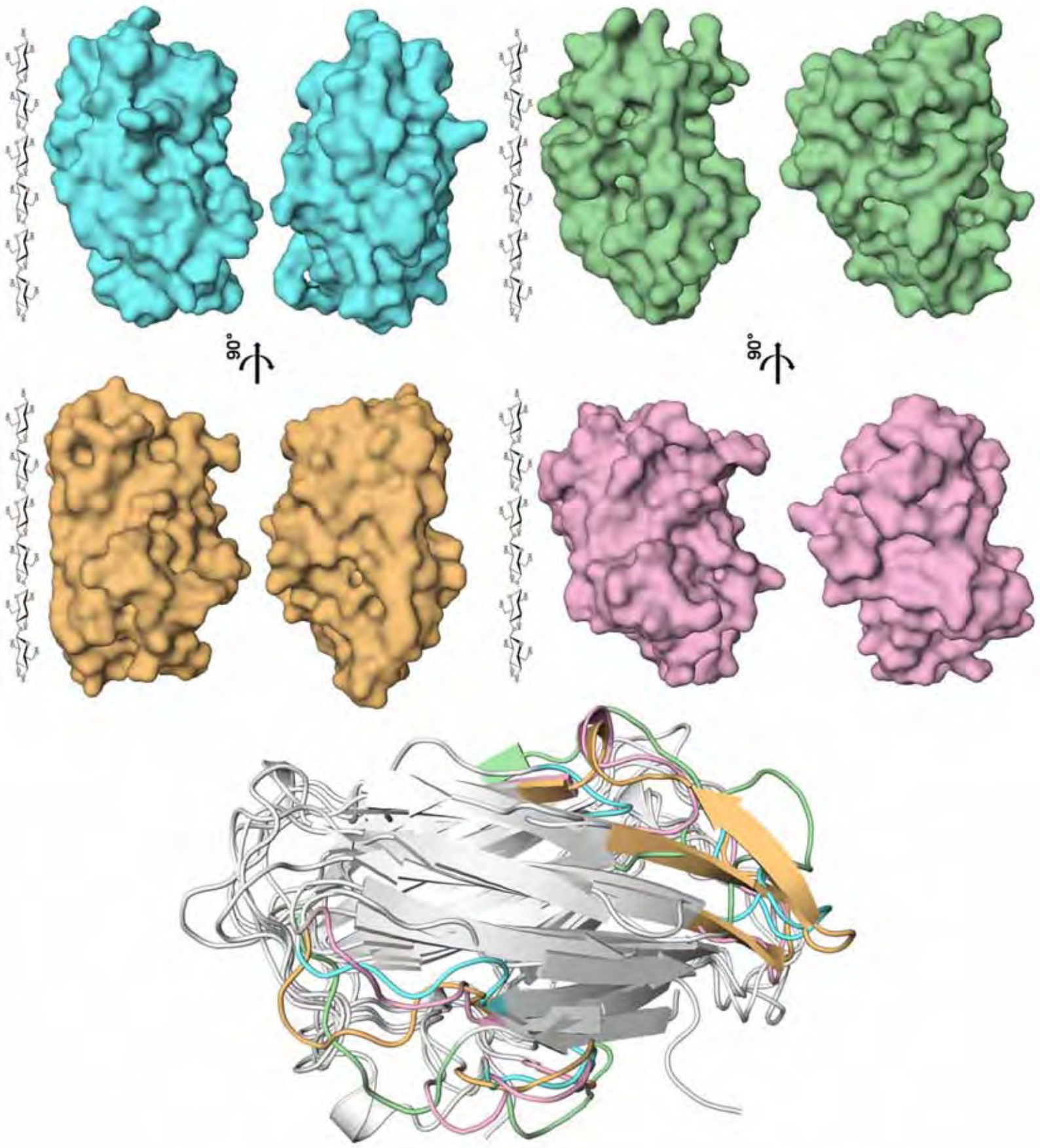
[‡] comparison between experimental and theoretical curves

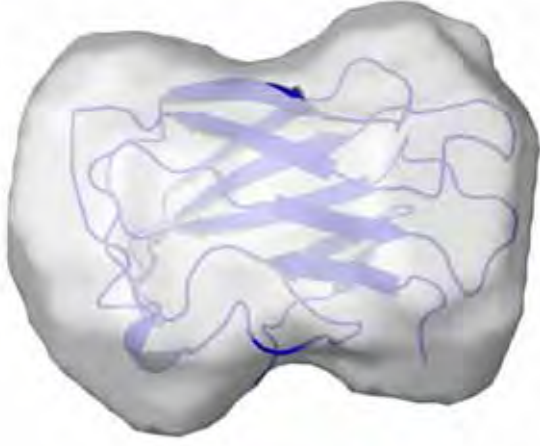
Table 3

	Form I	Form II 0 mM Mn ²⁺	Form II* 10 mM Mn ²⁺
<i>Data collection</i>			
Space group	C2	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit cell parameters (Å, °)	a=155.36, b=81.08, c=114.27, β=104.29	a=50.24, b=110.76, c=121.38	a=49.53 b=110.62 c=122.26
Resolution	50.00-.2.90 (3.00- 2.90)	50.00-2.10 (2.18-2.10)	50.00-1.97 (2.04-1.97)
N° unique reflections	30,407 (2,607)	39,069 (3,393)	47,366 (4,163)
Multiplicity	3.5 (2.7)	6.9 (6.2)	4.8 (4.3)
Completeness (%)	97.1 (83.7)	97.2 (85.9)	97.6 (87.8)
R _{merge} (%)	8.3 (43.6)	11.6 (31.9)	8.4 (37.3)
I/σ(I)	11.3 (2.0)	13.5 (5.0)	14.7 (3.8)
<i>Refinement</i>			
R _{work} /R _{free} (%)	23.0/29.6	17.1/22.1	16.4/20.6
Average B-factor	54.5	32.7	29.1
r.m.s.d. for bond lengths (Å)	0.006	0.023	0.025
r.m.s.d. for bond angles (°)	1.261	1.903	1.938
Protein chains	4	2	2
Water molecules	30	330	452
Other molecules	–	3 GOL	2 Mn, 3 GOL, 1 PO ₄
<i>Ramachandran plot</i>			
Favored region (%)	90.63	96.44	97.09
Outlier region (%)	1.35	0.17	0.00
PDB entry code	3PZV	3PZU	3PZT

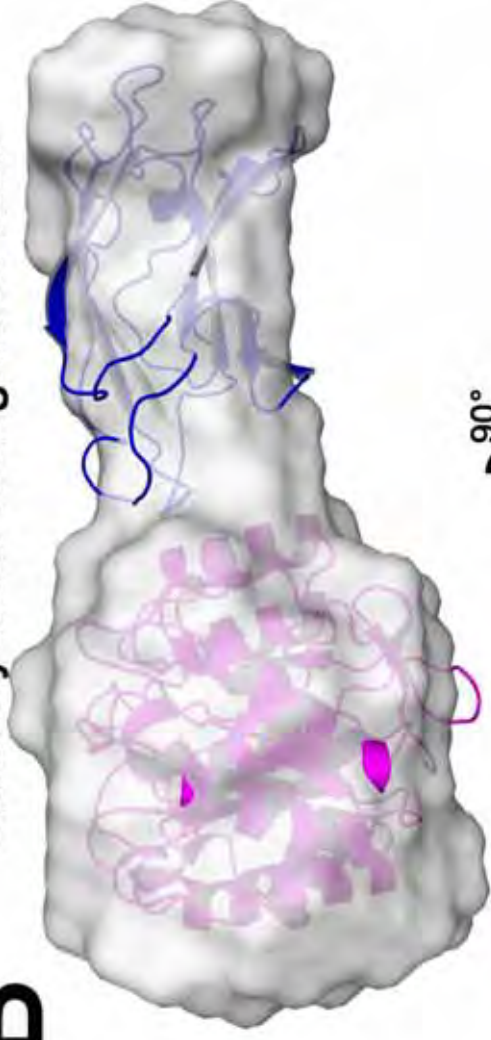
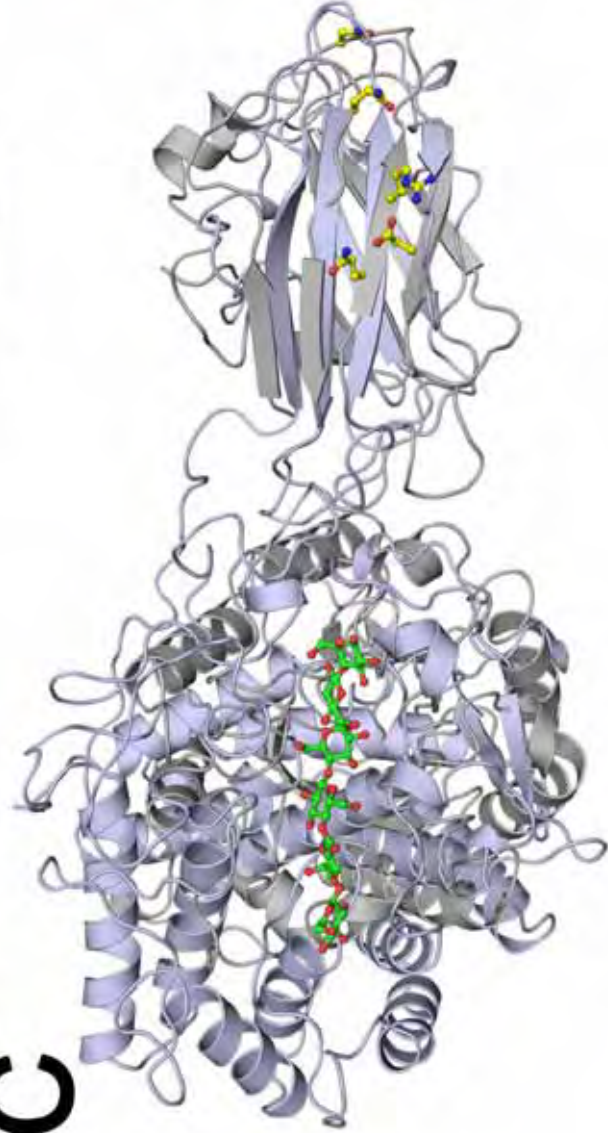


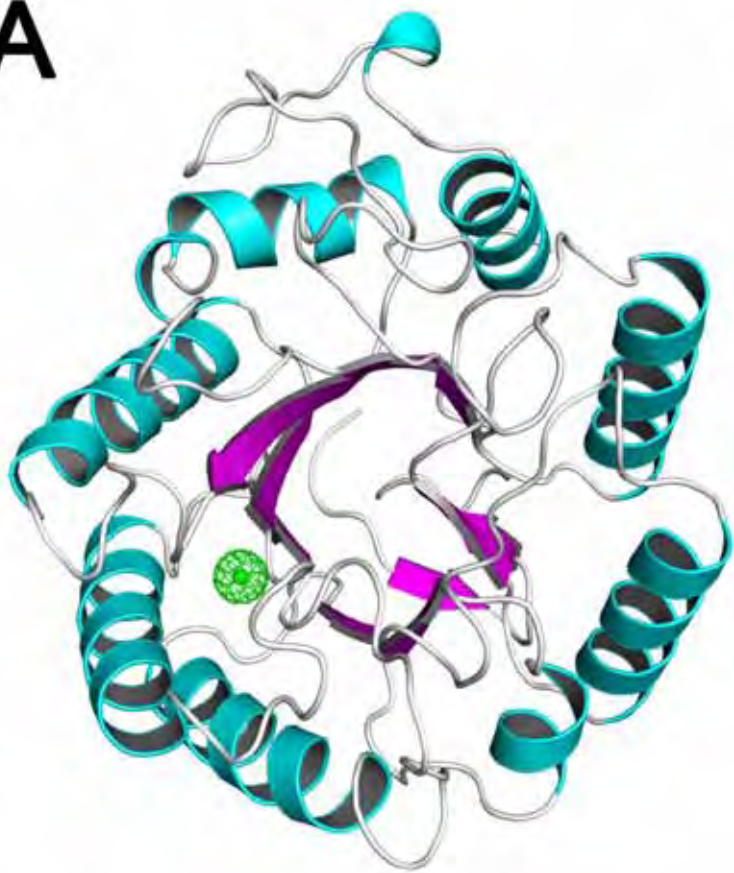
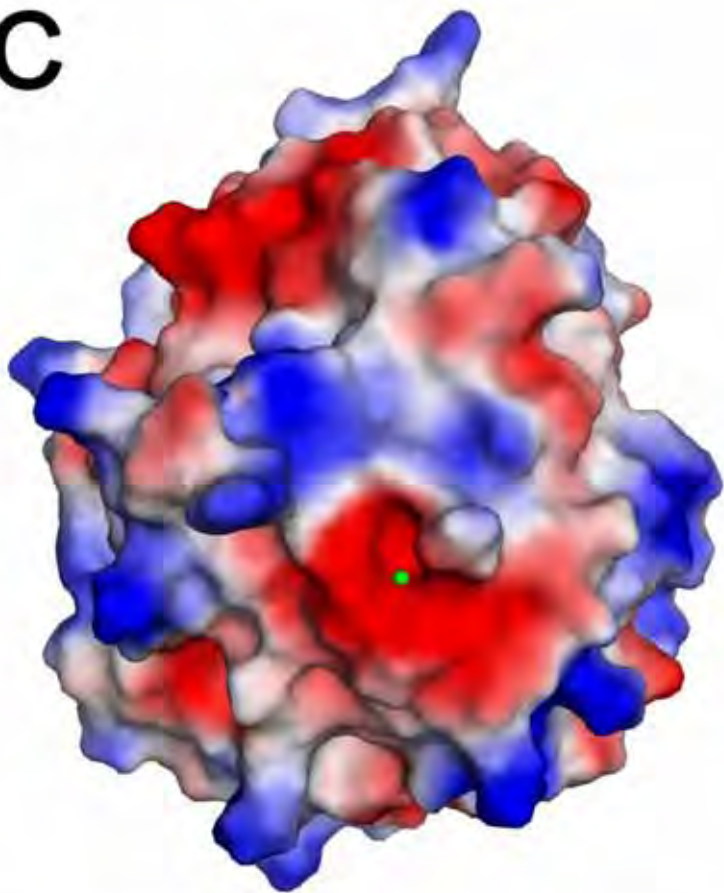
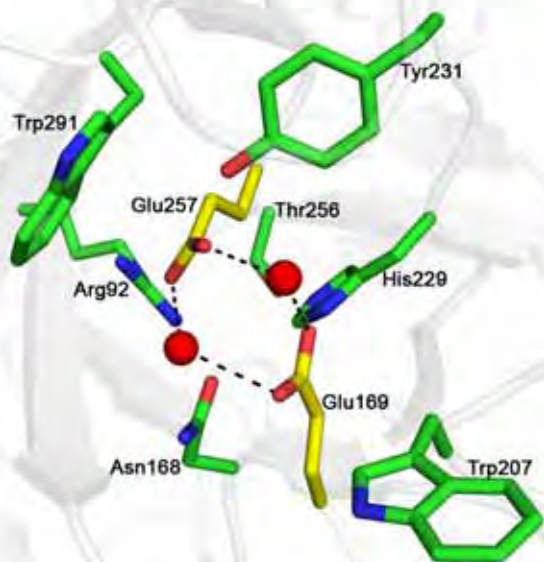




A**B**

Carbohydrate recognition face

 90°
**C**

A**C****B****D**