

Ministério da Saúde

**FIOCRUZ**  
**Fundação Oswaldo Cruz**

**INSTITUTO OSWALDO CRUZ**  
**Pós-Graduação em Biologia Celular e Molecular**

*Ana Carolina Rennó Soderó*

Modelagem Molecular de Aspartil Proteases de *Schistosoma mansoni* (SmAPs) para o Desenvolvimento de Potenciais Moléculas Esquistossomicidas

Tese apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Doutor em Ciências.

**Orientador:** Prof. Dr. Floriano Paes Silva Junior

**Co-orientador:** Prof. Dr. Salvatore Giovanni De Simone

**RIO DE JANEIRO**

2011

Ficha catalográfica elaborada pela  
Biblioteca de Ciências Biomédicas/ ICICT / FIOCRUZ - RJ

S679

Sodero, Ana Carolina Rennó.

Modelagem Molecular de Aspartil Proteases de *Schistosoma mansoni* (SmAPs) para o Desenvolvimento de Potenciais Moléculas Esquistossomicidas. / Ana Carolina Rennó Sodero. – Rio de Janeiro, 2011.

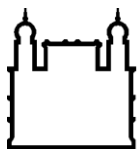
xvi, 144 f. il. ; 30 cm.

Tese (doutorado) – Instituto Oswaldo Cruz, Pós-Graduação em Biologia Celular e Molecular, 2011.

Bibliografia: f. 108-128

1. Modelagem Molecular. 2. Aspartil protease. 3. Esquistossomose. I.  
Título.

CDD 616.963



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

**INSTITUTO OSWALDO CRUZ**  
**Pós-Graduação em Biologia Celular e Molecular**

**AUTOR: Ana Carolina Rennó Soderó**

**MODELAGEM MOLECULAR DE ASPARTIL PROTEASES DE *Schistosoma mansonii* (SMAPS) PARA O DESENVOLVIMENTO DE POTENCIAIS MOLÉCULAS ESQUISTOSSOMICIDAS**

**ORIENTADOR: Prof. Dr. Floriano Paes Silva Junior**

**CO-ORIENTADOR: Prof. Dr. Salvatore Giovanni De Simone**

**Aprovada em: 30/09/2011**

**EXAMINADORES:**

**Prof. Dr. Ernesto Raúl Caffarena - Presidente**

**Prof. Dr. Osmar Norberto de Souza**

**Prof. Dr. Jose Daniel Figueroa Villar**

**Prof. Dr. Monique Araújo de Brito**

**Prof. Dr. Camila Silva de Magalhães**

Rio de Janeiro, 30 de setembro de 2011

“Amor da minha vida  
Daqui até a eternidade  
Nossos destinos  
Foram traçados na maternidade”  
(Cazuza / Ezequiel Neves / Leoni)

Dedico este trabalho à  
*Isabela Sodero Fleming*

## AGRADECIMENTOS

Foram muitas as pessoas que de alguma forma me ajudaram a concluir esse trabalho. Por isso, ficam aqui meus sinceros agradecimentos.

Início com aquela que mudou para sempre minha vida. Hoje eu entendo claramente a imensidão do amor de mãe. Isabela, você me faz uma pessoa melhor e mais feliz.

Ao Felipe Fleming, por ser meu amigo, companheiro e, especialmente, um grande pai. Obrigada pelas inúmeras discussões científicas, pelas críticas construtivas e pelo apoio constante durante todos os momentos. Certamente, você foi um grande incentivador para o término dessa tese.

Aos meus pais, João Carlos e Maria Aida, que sempre me deram suporte para a continuidade dos meus estudos. Vocês me ajudaram nos momentos mais difíceis e estavam presentes nos mais felizes. Sem vocês, eu não chegaria até aqui.

Às minhas irmãs e aos que se fizeram meus irmãos: Daniela e Marcelo, Fabrícia e Gustavo. Minhas irmãs são minhas primeiras e melhores amigas. Eu não poderia imaginar minha vida sem vocês.

Aos meus afilhados tão amados, João Marcelo e Paulo Victor.

À Julia, que mesmo que não nos vejamos como gostaríamos, é uma sobrinha muito especial e querida.

Ao Tutty, meu companheiro que me faz tanta falta.

À minha família que eu amo: tios, tias, primos e primas. Em especial, aos primos Guigui, Renata, Gugu e tia Maria Helena.

À extraordinária e querida família que eu ganhei: Fabio, Fernanda, Flavia, Norma, Marietta, Thomaz, Jorge, Carmem, Luciana, Doralice e todos os tios e tias, primos e primas. Obrigada por me receberem de forma tão acolhedora e pelo apoio constante.

À minha amiga-irmã Alessandra, sempre pronta a ajudar e oferecer um ombro amigo. Obrigada também às amigas muito queridas Nina, Thaísa, Júllia e Maria Augusta. Agradeço ainda a amizade real e duradoura da Ingrid, uma “grande” amiga duende!

Aos meus orientadores, Dr. Floriano Paes Silva Jr e Dr. Salvatore Giovanni De Simone, obrigada pela oportunidade e pela troca de conhecimento.

Aos amigos do LBPP: Amanda, André, Adriana, Babi, Bruna, Carol, Carol Bottino, Ciça (companhia constante durante toda minha gravidez), Cris, Dani Cunha (minha guru), Dani Marchi, Fátima, Genival, Gilda, Herbert, Juzinha, Kele, Luãna, Luiz, Marilúcia, Mario, Michelle, Natália, Naty, Paloma, Patrícia, Pina, Thatiane, Thaysa, Tonha, William. Nossa convivência durante essa longa jornada foi incrível!

À amiga Larissa, que mesmo longe, está sempre pronta para conversar e desabafar, ou até mesmo discutir um protocolo ou um artigo.

À minha amiga de “bancada” Aline, que me avisou, mas eu não escutei! Obrigada pela amizade, pelo “suporte técnico” e pela “vizinhança” engraçada e feliz.

Às minhas “filhas” Suzanna e Bel, obrigada pela paciência e ajuda na realização desse trabalho. Vocês são insubstituíveis!

Aos amigos do IOC Marcos, Carol e Monete.

Aos amigos do PROCC, em especial ao João Herminio e à Amanda Sutter. Sou muito grata pela ajuda para a realização dessa tese.

Aos amigos do LabMMol da UFRJ.

Às minhas amigas Isabella e Monique, obrigada pela amizade, pela ajuda e contribuição para a concretização desse trabalho.

Ao amigo Bruno Horta, que mesmo longe não esquece seus amigos.

Ao Dr. Ernesto Caffarena e ao Dr. Hugo Verli, os quais eu conheci logo no início do doutorado e que foram mais que mestres ao longo desse processo. Obrigada pela compreensão, pelos ensinamentos e pela presteza diante das dificuldades encontradas no trabalho.

Aos membros da banca examinadora, pela paciência e pela certeza da ajuda crítica e construtiva ao trabalho.

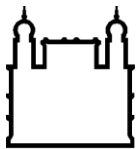
Às agências de fomento, por entenderem a importância da licença-maternidade na vida de uma mãe-bolsista.

Aos funcionários IOC/FIOCRUZ por tornarem nossa vida muito mais agradável.

A todos aqueles que, de alguma forma, contribuíram para minha formação e que, por culpa da minha falha memória, esqueci de citar nesse agradecimento. Muito obrigada.

“Cada dia sabemos mais e entendemos menos.”

(Albert Einstein)



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

## INSTITUTO OSWALDO CRUZ

### MODELAGEM MOLECULAR DE ASPARTIL PROTEASES DE *Schistosoma mansoni* (SmAPS) PARA O DESENVOLVIMENTO DE POTENCIAIS MOLÉCULAS ESQUISTOSSOMICIDAS

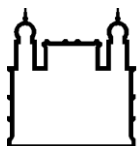
#### RESUMO

#### TESE DE DOUTORADO

**Ana Carolina Rennó Sodero**

Peptidases de parasitas são descritas como potenciais alvos terapêuticos. Novas proteases catépsina-D símile de *Schistosoma mansoni* foram identificadas, em adição à SmCD1, uma aspartil protease envolvida no metabolismo da hemoglobina. Isto indica que uma família de aspartil proteases está presente no parasita. Foram investigados os aspectos estruturais de inibição de aspartil proteases *S. mansoni* pela pepstatina e por inibidores peptidomiméticos previamente sintetizados e avaliados experimentalmente. As estruturas tridimensionais das enzimas SmCD1, SmCD2 e SmCD3 foram construídas por modelagem comparativa, seguida da metodologia de atracação molecular para gerar os complexos entre as enzimas e os inibidores. Simulações por Dinâmica Molecular foram realizadas com os complexos enzimas/pepstatina. Ligações de hidrogênio, pontes salinas e interações hidrofóbicas foram analisadas para caracterizar a relação estrutura-atividade. A enzima SmCD1 apresentou grande alteração conformacional na alça que conecta as fitas s10C e s11C, próxima ao resíduo de isovalina da pepstatina, que resultou em contatos favoráveis entre a SmCD1 e a pepstatina. Esta interação resultou na formação de um subsítio hidrofóbico em S4. Tais mudanças conformacionais não são observadas nos complexos com a SmCD2 e a SmCD3. Os resultados indicaram que a especificidade de cada enzima pode estar associada aos resíduos externos à tríade catalítica e que inibidores específicos para aspartil proteases de *S. mansoni* podem ser futuramente desenvolvidos.





Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

## INSTITUTO OSWALDO CRUZ

### MOLECULAR MODELING OF ASPARTIC PROTEASES FROM *Schistosoma mansoni* (SmAPS) FOR DEVELOPMENT OF POTENTIAL ANTI-SCHISTOSOMAL DRUGS

#### ABSTRACT

#### TESE DE DOUTORADO

**Ana Carolina Rennó Sodero**

Peptidases of parasites have been described as potential targets for chemotherapy. Novel cathepsin D-like proteases of *Schistosoma mansoni* were identified, in addition to SmCD1, an aspartyl protease required for hemoglobin metabolism. This indicates that a family of aspartyl proteases is present in this parasite. It was investigated the structural aspects of the SmCD proteases inhibition by pepstatin and peptidomimetic inhibitors, previously synthesized and experimentally evaluated. Tridimensional structures of SmCD1, SmCD2 and SmCD3 were modeled by comparative modeling, followed by molecular docking methodology to generate enzyme-inhibitor complexes. Molecular dynamics simulations were performed on complexes with pepstatin. Hydrogen bonds, salt bridges and hydrophobic interactions were analyzed to characterize structure-function relationships within the enzymes. SmCD1 showed a large conformational change in the loop connecting the s10C and s11C strands, near the isovaline of pepstatin, rendering favorable contacts with pepstatin. This interaction resulted in a hydrophobic pocket within the S4 subsite. Such conformational changes were not observed in the SmCD2 and SmCD3 complexes with pepstatin. The simulation results indicated that the specificity of each enzyme may be determined by non-active site residues and that specific inhibitors for aspartic proteases of *S. mansoni* can be further developed.

## Índice

<b>AGRADECIMENTOS.....</b>	<b>iii</b>
<b>RESUMO.....</b>	<b>vi</b>
<b>ABSTRACT.....</b>	<b>vii</b>
<b>ABREVIATURAS.....</b>	<b>xi</b>
<b>ÍNDICE DE FIGURAS.....</b>	<b>xii</b>
<b>ÍNDICE DE TABELAS.....</b>	<b>xvi</b>
<b>1 INTRODUÇÃO.....</b>	<b>1</b>
1.1 Esquistossomose.....	1
1.1.1 Ciclo Biológico da Doença.....	2
1.1.2 Controle da Doença.....	3
1.1.2.1 Perspectivas para o Tratamento da Esquistossomose.....	4
1.1.3 Aspartil Proteases.....	6
1.1.3.1 Função Biológica.....	6
1.1.3.2 Estrutura Tridimensional de Aspartil Proteases.....	7
1.1.3.3 Mecanismo Catalítico de Aspartil Proteases.....	8
1.1.3.3.1 Estudos de Difração de Raios X de Complexos Enzima-Inibidor.....	8
1.1.3.3.2 Estruturas de Raios X de Alta Resolução.....	11
1.1.3.3.3 Evidências Experimentais Baseadas em Estudos de RMN.....	12
1.1.3.3.4 Estudos de Difração de Nêutrons.....	13
1.1.3.3.5 Ligações de Hidrogênio de Baixa Barreira.....	14
1.1.3.4 Aspartil Proteases de <i>S. mansoni</i> .....	16
1.1.3.5 Inibidores de Aspartil Proteases.....	17
1.1.3.5.1 Inibidores de Aspartil Proteases de <i>S. mansoni</i> .....	18
1.2 Desenvolvimento Racional de Fármacos.....	20
1.2.1 Conceitos Básicos.....	22
1.2.2 Metodologias de Modelagem Molecular Aplicadas ao Planejamento de Fármacos.....	23
1.2.2.1 Modelagem Comparativa ou Modelagem por Homologia.....	23
1.2.2.1.1 Modeller.....	27
1.2.2.2 Atracação Molecular (“Docking”).....	27
1.2.2.2.1 AutoDock.....	29
1.2.2.2.2 Atracação Molecular de Peptídeos e Peptidomiméticos.....	31

1.2.2.3 Simulação por Dinâmica Molecular (DM).....	32
1.2.2.3.1 O Pacote de Programas GROMACS.....	33
1.2.2.3.2 Simulação por Dinâmica Molecular de Aspartil Proteases de Eucariotos.....	34
<b>2 OBJETIVOS.....</b>	<b>36</b>
2.1 Objetivos Específicos.....	36
<b>3 METODOLOGIA.....</b>	<b>37</b>
3.1 Modelagem Comparativa.....	37
3.1.1 Análise da Seqüência.....	37
3.1.2 Identificação e Seleção de Proteínas de Referência.....	37
3.1.3 Alinhamento das Seqüências de Resíduos de Aminoácidos.....	37
3.1.4 Construção dos Modelos.....	38
3.1.5 Validação dos Modelos.....	38
3.2 Atracação Molecular.....	38
3.2.1 “Re-docking” e “Cross-docking”.....	38
3.2.2 Preparação dos Ligantes.....	39
3.2.3 Preparação das Macromoléculas.....	41
3.2.3.1 Protonação.....	42
3.2.4 Parâmetros da Atracação Molecular.....	42
3.2.5 Análise dos Resultados.....	44
3.3 Análise e Composição de Subsítios.....	44
3.4 Simulação por Dinâmica Molecular.....	44
3.4.1 Análise dos Resultados.....	46
3.5 Construção de Figuras e Gráficos.....	46
<b>4 RESULTADOS E DISCUSSÃO.....</b>	<b>47</b>
4.1 Modelagem Comparativa.....	47
4.1.1 Seleção das Proteínas de Referência.....	47
4.1.2 Alinhamento das Seqüências de Aminoácidos.....	49
4.1.3 Construção e Validação dos Modelos.....	53
4.1.4 Estrutura Tridimensional dos Modelos.....	56
4.2 ATRACAÇÃO MOLECULAR.....	61
4.2.1 “Re-docking” e “Cross-docking”.....	61
4.2.2 A Pepstatina e os Ligantes da Série LBPP.....	64
4.2.3 Subsítios S1 e S3.....	67

4.2.4 Subsítio S2 e S1'.....	72
4.2.5 Subsítios S2' e S3'.....	77
4.2.6 Subsítio S4.....	81
<b>4.3 SIMULAÇÃO POR DINÂMICA MOLECULAR.....</b>	<b>84</b>
4.3.1 Estabilidade das Trajetórias.....	84
4.3.2 Análise Estrutural.....	87
4.3.3 O Papel das Pontes Salinas na Estabilidade dos Complexos.....	90
4.3.4 O Papel das Ligações de Hidrogênio nas Afinidades dos Complexos....	94
4.3.5 O Papel das Interações Hidrofóbicas no Processo de Inibição.....	100
4.4 Proposta de Modificações Estruturais para Novos Inibidores Seletivos.....	103
<b>5 CONCLUSÕES E PERSPECTIVA.....</b>	<b>106</b>
<b>6 REFERÊNCIAS.....</b>	<b>108</b>
<b>Anexo I: Peptídeo Sinal.....</b>	<b>129</b>
<b>Anexo II: Pró-Peptídeo.....</b>	<b>131</b>
<b>Anexo III: Tabelas de Valores de pKa.....</b>	<b>133</b>
<b>Anexo IV: Subsítios das Enzimas Estudadas.....</b>	<b>136</b>
<b>Anexo V: Tabelas de Pontes Salinas.....</b>	<b>138</b>
<b>Anexo VI: Tabelas de Ligações de Hidrogênio.....</b>	<b>141</b>

## ABREVIATURAS

<b>3D</b>	Tridimensional
<b>AP</b>	Aspartil Protease
<b>CADD</b>	“Computer-Assisted Drug Design” (Planejamento de Fármacos Auxiliado por Computador)
<b>Hb</b>	Hemoglobina
<b>IC<sub>50</sub></b>	Concentração Inibitória de 50%
<b>K<sub>i</sub></b>	Constante de inibição
<b>GA</b>	Algoritmo Genético
<b>IET</b>	Isómero do Estado de Transição
<b>LBHB</b>	“Low Barrier Hydrogen Bond” (Ligação de Hidrogênio de Baixa Barreira)
<b>LBPP</b>	Laboratório de Bioquímica de Proteínas e Peptídeos
<b>LGA</b>	Algoritmo Genético Lamarckiano
<b>PBC</b>	“Periodic Boundary Conditions” (Condições Periódicas de Contorno)
<b>PDB</b>	“Protein Data Bank” (Banco de Dados de Proteína)
<b>PME</b>	“Particle Mesh Ewald” (Método de Ewald para Tratamento de Interações Eletrostáticas)
<b>RG</b>	Raio de Giro
<b>RMN</b>	Ressonância Magnética Nuclear
<b>RMSD</b>	“Root Mean Square Deviation” (Desvio Quadrático Médio das Distâncias)
<b>RMSF</b>	“Root Mean Square Fluctuation” (Flutuação do Desvio Quadrático Médio)
<b>SAS</b>	Superfície de Acesso ao Solvente
<b>SBDD</b>	“Structure-Based Drug Design” (Desenho de Fármacos Baseado na Estrutura)
<b>DM</b>	Simulação por Dinâmica Molecular
<b>SmAPs</b>	Aspartil Proteases de <i>Schistosoma mansoni</i>
<b>SmCD</b>	Catepsina D de <i>Schistosoma mansoni</i>
<b>TGR</b>	“Thioredoxin-Glutathione Reductase” (Tioredoxina-Glutationa Redutase)

## ÍNDICE DE FIGURAS

<b>FIGURA 1.1.</b>	Representação do ciclo de transmissão do <i>S. mansoni</i> .....	2
<b>FIGURA 1.2.</b>	Estrutura química da oxamniquina <b>1</b> e do praziquantel <b>2</b> .....	4
<b>FIGURA 1.3.</b>	Representação esquemática do sítio de ligação de um substrato na protease.....	8
<b>FIGURA 1.4.</b>	Estrutura do composto CP-81292 <b>3</b> .....	9
<b>FIGURA 1.5.</b>	Mecanismo catalítico de aspartil proteases proposto por Veerapandian e colaboradores em 1992 [47].....	9
<b>FIGURA 1.6.</b>	Estrutura do sítio catalítico da endotiapepsina complexada ao inibidor análogo ao estado de transição H271.....	14
<b>FIGURA 1.7.</b>	Estrutura química da pepstatina <b>4</b> .....	18
<b>FIGURA 1.8.</b>	Valores de IC <sub>50</sub> e K <sub>i</sub> para os inibidores peptidomiméticos utilizados nesse trabalho.....	20
<b>FIGURA 1.9.</b>	Estratégias para a descoberta de novos fármacos.....	21
<b>FIGURA 1.10.</b>	Dados de seqüências e estruturas depositadas nos principais bancos de dados de proteínas.....	24
<b>FIGURA 1.11.</b>	Representação do gráfico de Ramachandran.....	26
<b>FIGURA 1.12.</b>	Sobreposição da cadeia principal de <b>(A)</b> aspartil proteases em complexo com inibidores peptidomiméticos; <b>(B)</b> inibidores de aspartil proteases, quando ancorados em seus respectivos sítios ativos.....	32
<b>FIGURA 3.1.</b>	Estrutura química dos inibidores pepstatina <b>4</b> , PHO <b>5</b> , CH-66 <b>6</b> e a série LBPP.....	40
<b>FIGURA 3.2.</b>	Representação da grade calculada pelo programa AutoGrid.....	43
<b>FIGURA 4.1.</b>	Alinhamento entre as seqüências da enzimas SmCD1 e as proteínas de referência utilizadas na modelagem comparativa.....	51
<b>FIGURA 4.2.</b>	Alinhamento entre a seqüência da enzima SmCD2 e as proteínas de referência utilizadas na modelagem comparativa.....	52
<b>FIGURA 4.3.</b>	Alinhamento entre a seqüência da enzima SmCD3 e as proteínas de referência utilizadas na modelagem comparativa.....	53
<b>FIGURA 4.4.</b>	Gráfico de Ramachandran para as enzimas <b>(A)</b> SmCD1, <b>(B)</b> SmCD2 e <b>(C)</b> SmCD3, obtido pelo programa PROCHECK [118]..	54

<b>FIGURA 4.5.</b>	Gráfico obtido pelo programa PROVE [120]: Z-score e Z-score RMS versus resolução do arquivo PDB para enzimas <b>(A)</b> SmCD1, <b>(B)</b> SmCD2 e <b>(C)</b> SmCD3.....	55
<b>FIGURA 4.6.</b>	Estruturas 3D da enzima SmCD1 construída por modelagem comparativa.....	57
<b>FIGURA 4.7.</b>	Estruturas 3D da enzima SmCD2 construída por modelagem comparativa.....	58
<b>FIGURA 4.8.</b>	Estruturas 3D da enzima SmCD3 construída por modelagem comparativa.....	58
<b>FIGURA 4.9.</b>	Alinhamento construído pelo programa T-Coffee [187] entre as seqüências da catepsina D humana e as enzimas SmCD1, SmCD2 e SmCD3.....	60
<b>FIGURA 4.10.</b>	Sobreposição das soluções selecionadas na metodologia de atracação molecular e as conformações obtidas pela difração de raios X.....	62
<b>FIGURA 4.11.</b>	Comparação entre as conformações selecionadas pela metodologia de atracação molecular para a pepstatina e a série de ligantes LBPP no sítio ativo das enzimas estudadas.....	67
<b>FIGURA 4.12.</b>	Comparação entre as diferentes interações dos ligantes estudados e os subsítios S1 e S3 da SmCD1.....	69
<b>FIGURA 4.13.</b>	Superfície molecular da SmCD1, evidenciando a interação hidrofóbica dos resíduos Phe115, Phe120 e Ile123 com os ligantes estudados.....	72
<b>FIGURA 4.14.</b>	Comparação entre as diferentes interações dos ligantes estudados e os subsítios S2 e S1' da SmCD1.....	74
<b>FIGURA 4.15.</b>	Comparação entre as diferentes interações dos ligantes estudados e os subsítios S2 e S1' da SmCD2.....	74
<b>FIGURA 4.16.</b>	Sobreposição da série de ligantes LBPP e a pepstatina, no subsítio S2' da SmCD1.....	78
<b>FIGURA 4.17.</b>	Comparação entre as diferentes interações dos ligantes estudados e os subsítios S2' e S3' da SmCD2.....	79
<b>FIGURA 4.18.</b>	Comparação entre as diferentes interações dos ligantes estudados e o subsítio S4 da SmCD1.....	83

<b>FIGURA 4.19.</b>	Comparação entre as diferentes interações dos ligantes estudados e o subsítio S4 da SmCD3.....	84
<b>FIGURA 4.20.</b>	RMSD dos átomos da cadeia principal de: SmCD1 livre (azul), SmCD1/pepstatina (amarelo), SmCD2 livre (verde), SmCD2/pepstatina (laranja), SmCD3 livre (vermelho) e SmCD3/pepstatina (cinza).....	85
<b>FIGURA 4.21.</b>	Sobreposição entre as estruturas iniciais (em rosa) e as estruturas médias de: SmCD1 livre (azul), SmCD1/pepstatina (amarelo), SmCD2 livre (verde), SmCD2/pepstatina (laranja), SmCD3 livre (vermelho) e SmCD3/pepstatina (cinza).....	86
<b>FIGURA 4.22.</b>	Sobreposição das estruturas médias de: SmCD1 livre (azul), SmCD1/pepstatina (amarelo), SmCD2 livre (verde), SmCD2/pepstatina (laranja), SmCD3 livre (vermelho) e SmCD3/pepstatina (cinza).....	88
<b>FIGURA 4.23.</b>	Gráficos de RMSF e em salsicha.....	89
<b>FIGURA 4.24.</b>	Gráficos de evolução temporal das distâncias entre as principais pontes salinas formadas pela SmCD1 livre (azul) e pelo complexo SmCD1/pepstatina (amarelo).....	91
<b>FIGURA 4.25.</b>	Gráficos de evolução temporal das distâncias entre as principais pontes salinas formadas pela SmCD3 livre (vermelho) e pelo complexo SmCD3/pepstatina (cinza).....	93
<b>FIGURA 4.26.</b>	Representação das principais pontes salinas importantes para a interação com a pepstatina.....	94
<b>FIGURA 4.27.</b>	Gráficos de evolução temporal das distâncias entre as ligações de hidrogênio do “fireman’s grip”.....	96
<b>FIGURA 4.28.</b>	Representação gráfica dos complexos entre pepstatina e <b>(A)</b> SmCD1, <b>(B)</b> SmCD2 e <b>(C)</b> SmCD3.....	97
<b>FIGURA 4.29.</b>	Gráficos de evolução temporal das distâncias de algumas ligações de hidrogênio entre a pepstatina e as enzimas SmCD1 SmCD1 (amarelo), SmCD2 (laranja) e SmCD3 (cinza).....	98
<b>FIGURA 4.30.</b>	RMSD de todos os átomos da pepstatina ligada à: SmCD1 (amarelo), SmCD2 (laranja) e SmCD3 (cinza).....	99
<b>FIGURA 4.31.</b>	Gráficos de evolução temporal das distâncias.....	101



<b>FIGURA 4.32.</b> Representações da superfície das enzimas livres e dos complexos.....	102
<b>FIGURA 4.33.</b> Representação esquemática das modificações estruturais propostas para síntese de inibidores seletivos das enzimas SmCD1, SmCD2 e SmCD3.....	105

## ÍNDICE DE TABELAS

<b>TABELA 3.1.</b>	Complexos utilizados no “re-docking” e “cross-docking”.....	39
<b>TABELA 3.2.</b>	Variáveis investigadas no processo de “re-docking” e “cross-docking”.....	43
<b>TABELA 3.3.</b>	Condições usadas nas Simulações por Dinâmica Molecular (DM).....	46
<b>TABELA 4.1.</b>	Proteínas utilizadas como referência para a construção dos modelos das enzimas SmCD1, SmCD2 e SmCD3.....	49
<b>TABELA 4.2.</b>	Valores selecionados das variáveis investigadas no processo de “re-docking” e “cross-docking”.....	61
<b>TABELA 4.3.</b>	Resultados do cálculo de “re-docking” e “cross-docking”.....	64
<b>TABELA 4.4.</b>	Resultados dos cálculos de atracção molecular da SmCD1 com os inibidores peptidomiméticos estudados.....	65
<b>TABELA 4.5.</b>	Resultados dos cálculos de atracção molecular da SmCD2 com os inibidores peptidomiméticos estudados.....	65
<b>TABELA 4.6.</b>	Resultados dos cálculos de atracção molecular da SmCD3 com os inibidores peptidomiméticos estudados.....	66
<b>TABELA 4.7.</b>	Interações entre os ligantes e as enzimas estudadas no subsítio S1 e S3, obtidas pelos resultados de atracção molecular.....	70
<b>TABELA 4.8.</b>	Interações entre os ligantes e as enzimas estudadas no subsítio S2, obtidas pelos resultados de atracção molecular.....	75
<b>TABELA 4.9.</b>	Interações entre os ligantes e as enzimas estudadas nos subsítios S2’ e S3’, obtidas pelos resultados de atracção molecular.....	80
<b>TABELA 4.10.</b>	Interações entre os ligantes e as enzimas estudadas no subsítio S4, obtidas pelos resultados de atracção molecular.....	82
<b>TABELA 4.11.</b>	Ligações de hidrogênio entre os resíduos de aspartato catalíticos, Asp33 e Asp219, durante a DM das enzimas estudadas.....	99

## 1 INTRODUÇÃO

A pesquisa e desenvolvimento de novos fármacos é um desafio, pois necessita de grande capital e é largamente conduzida por incentivos do mercado farmacêutico. Embora muitas terapias inovadoras tenham sido geradas, as mesmas não suprem as necessidades de doenças com as quais os incentivos comerciais não são suficientes para alavancar o setor privado em investimentos de pesquisa e desenvolvimento [1, 2]. Essas doenças são comumente denominadas “negligenciadas”, mas também são conhecidas como tropicais, endêmicas, órfãs ou pobres [1, 3]. Nos países desenvolvidos, as doenças “órfãs” afetam pequeno número de pessoas. Em países menos desenvolvidos, onde algumas doenças podem afetar milhões de pacientes, a indústria farmacêutica é relutante em investimentos no desenvolvimento de novos fármacos, pois o retorno financeiro não é garantido [1, 4]. Este é o caso da esquistossomose, uma doença endêmica no Brasil [5].

Dessa forma, torna-se essencial a pesquisa intensa dos órgãos de saúde públicos no combate a essas enfermidades.

### 1.1 ESQUISTOSSOMOSE

A esquistossomose é uma doença causada por trematódeos do gênero *Schistosoma*, que afeta cerca de 200 milhões de pessoas no mundo todo [6]. Os principais agentes etiológicos são as espécies *S. mansoni*, *S. haematobium* e *S. japonicum* [5].

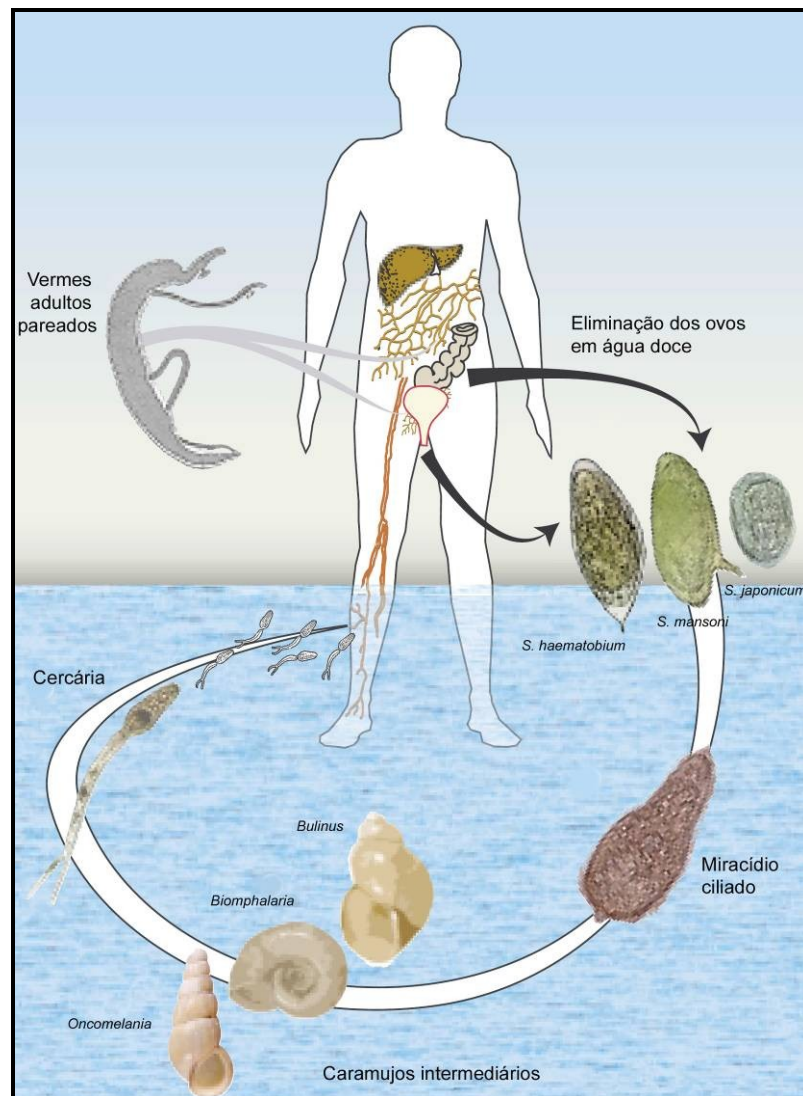
No mundo, a esquistossomose ocorre em 54 países, enfatizando-se aqueles pertencentes ao continente africano, leste do Mediterrâneo, América do Sul e Caribe. No Brasil, essa parasitose é causada pelo agente etiológico *Schistosoma mansoni* e apresenta ampla distribuição geográfica, sendo endêmica nos estados do Maranhão até Minas Gerais, com focos no Pará, Piauí, Rio de Janeiro, São Paulo, Paraná, Santa Catarina, Goiás, Distrito Federal e Rio Grande do Sul [7, 8].

A esquistossomose tem se tornado menos comum devido aos tratamentos e prevenção existentes, mas é importante advertir que a morbidade sutil ou indireta devido à fadiga e à diminuição da capacidade física e cognitiva do paciente ainda é um problema [9]. Embora o efeito da morbidade global da doença seja controverso, King e colaboradores [10] demonstraram associações significantes entre a

esquistossomose e anemia, dor crônica, diarreia, intolerância ao exercício e subnutrição. Esses resultados indicam a necessidade de reavaliar as prioridades nas pesquisas desta doença silenciosamente endêmica.

### 1.1.1 CICLO BIOLÓGICO DA DOENÇA

O *S. mansoni* apresenta um ciclo de vida complexo (FIGURA 1.1), envolvendo hospedeiros definitivos (mamíferos) e intermediários (molusco), além do ambiente externo. No Brasil, o principal hospedeiro intermediário do *S. mansoni* é o caramujo da espécie *Biomphalaria glabrata* [11].



**FIGURA 1.1.** Representação do ciclo de transmissão do *S. mansoni* (adaptado de [9]).

O ciclo biológico da esquistossomose [5] se inicia quando os ovos do *S. mansoni* são eliminados pelas fezes do hospedeiro humano infectado. Caso as fezes sejam disseminadas em água doce, os ovos eclodem liberando uma larva ciliada, denominada *miracídio*. Esse miracídio é o agente responsável por infectar o hospedeiro intermediário, o caramujo. Após quatro a seis semanas, as larvas são liberadas pelo caramujo e ficam livres na água, na forma de *cercárias*. A cercária é a forma infectante para o homem. O verme se desenvolve no organismo humano durante duas a seis semanas após a penetração das cercárias. Posteriormente, o homem infectado pode transmitir a doença ao eliminar os ovos de *S. mansoni* nas fezes, por muitos anos.

A deposição de ovos no fígado e outros órgãos é responsável pela resposta inflamatória [12]. Muitas das pessoas infectadas podem, por sua vez, permanecer assintomáticas [5].

---

### 1.1.2 CONTROLE DA DOENÇA

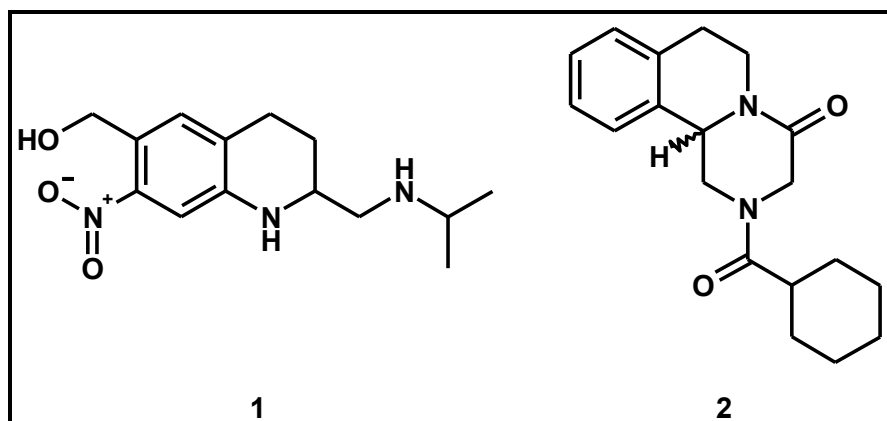
O controle da esquistossomose é um dos pontos mais importantes e difíceis para os serviços de saúde pública. Além do controle farmacoterapêutico, deve-se considerar também o controle dos moluscos, as condições de moradia e o saneamento básico da população, a educação sanitária e a adesão aos programas de controle [5, 13].

Atualmente, dois medicamentos estão disponíveis para o tratamento da esquistossomose, o praziquantel e a oxamniquina.

A oxamniquina é um composto tetraidroquinolínico (FIGURA 1.2, 1) semi-sintético, efetivo apenas em infecções por *S. mansoni*, mas sem atividade sobre o *S. japonicum* e o *S. hematobium*. Esse fármaco provoca efeitos adversos pronunciados, como sonolência e ataques epiléticos [9]. Além disso, é mais ativo em vermes machos que em fêmeas e é menos ativo em estágios imaturos do parasita [14]. Adicionalmente, a resistência em pacientes está bem documentada, principalmente no Brasil [15].

O praziquantel é um composto quinolina-pirazina acilado (FIGURA 1.2, 2) ativo em todas as espécies de *Schistosoma*. Esse composto se tornou o fármaco de escolha no tratamento da esquistossomose devido a uma série de razões, entre elas: efeitos colaterais passageiros e de pouca intensidade, administração em dose

única oral e baixo custo [9, 16] e amplo espectro de ação contra todas as espécies de *Schistosoma* e contra cestódeos [14]. No entanto, parasitas resistentes ao praziquantel foram identificados em estudos de campo [17-20] e em culturas de laboratório [21-23]. Sua atuação sobre ovos e formas imaturas do parasito é praticamente nula [9, 24] e seu mecanismo de ação não está totalmente elucidado [25], impedindo modificações químicas racionais deste fármaco no intuito de melhorar sua ação farmacológica.



**FIGURA 1.2.** Estrutura química da oxamniquina (1) e do praziquantel (2).

Além disso, a existência de um único tratamento efetivo no controle da esquistossomose não é adequada e o desenvolvimento de outros fármacos seria a estratégia mais direta para evitar e contornar o problema da resistência. Assim, consolida-se a necessidade de se encontrar alvos terapêuticos para o desenvolvimento de novas moléculas esquistossomicidas.

#### 1.1.2.1 PERSPECTIVAS PARA O TRATAMENTO DA ESQUISTOSSOMOSE

A esquistossomose crônica, causada pelos parasitas *Schistosoma haematobium*, *S. mansoni*, *S. intercalatum*, *S. mekongi*, ou *S. japonicum*, é uma das doenças infecciosas mais prevalentes no mundo [10]. O praziquantel, fármaco de escolha para o tratamento da esquistossomose (detalhes na seção 1.1.2), é uma molécula quiral administrada como racemato [26]. Um dos enantiômeros, a forma (R), é ativa (eutômero), enquanto que a forma (S) é inativa (distômero). Ou seja, o praziquantel está disponível no mercado em comprimidos com 600 mg, mas com

apenas 300 mg do componente ativo. Além disso, sugere-se que o composto sem atividade seja o maior contribuinte para o gosto desagradável e amargo do medicamento [27]. Dessa forma, tornou-se desejável a produção do enantiômero ativo puro do praziquantel.

Existem duas abordagens básicas para a produção do praziquantel como um enantiômero puro [26]: (i) a síntese do racemato seguida da separação dos enantiômeros e (ii) a síntese exclusiva do enantiômero desejado. Ambas são viáveis para produzir o praziquantel na sua forma ativa em larga escala, mas elevariam os custos do medicamento, sendo mais dispendioso que a produção racêmica do composto [26].

Dessa forma, uma estratégia para o avanço da quimioterapia da doença seria a busca por tratamentos alternativos, o que adicionalmente poderia evitar o desenvolvimento de resistência. Em 2007, Sayed e colaboradores [28] identificaram uma série de promissores inibidores de tioredoxina-glutationa redutase (TGR, “Thioredoxin-Glutathione Reductase”). Alguns inibidores apresentaram atividades na faixa de micro a nanomolar e a incubação dos esquistossomos na presença destes levou à rápida inibição da atividade da TGR e a morte do parasita. Além disso, o tratamento de ratos infectados com o composto 4-fenil-1,2,5-oxadiazol-3-carbonitrila-2-óxido levou a reduções significativas dos parasitas em múltiplos estágios, sendo sua atividade igual ou superior à do praziquantel. O composto foi ativo ainda contra as três maiores espécies infectantes para o homem: *S. mansoni*, *S. haematobium* e *S. japonicum*.

Outra estratégia em constante desenvolvimento é a produção de vacinas para evitar a infecção pelo parasita. A compreensão da resposta imunológica à infecção do esquistossomo, tanto em modelos animais quanto em humanos, sugere que o desenvolvimento de uma vacina é possível [29].

Durante as duas últimas décadas, muitos laboratórios tiveram suas atenções voltadas para a identificação de antígenos que induziam, pelo menos parcialmente, a resposta imune à esquistossomose [30]. O desenvolvimento de vacinas para *S. mansoni* resultou na proteína de 14kDa, denominada Sm14, que está na Fase 1 de testes clínicos [31]. Outra promissora vacina em testes clínicos é a BILHVAX, ou Sh28GST, derivada do *S. haematobium* [32]. Alguns candidatos a vacina também estão em progresso para *S. japonicum* [29, 33].

O tratamento ideal seria uma combinação entre vacinação e quimioterapia, reduzindo a morbidade total e limitando o impacto da reinfeção [33].

---

### 1.1.3 ASPARTIL PROTEASES

As proteases constituem uma das maiores famílias de enzimas codificadas pelo genoma humano com mais de 500 membros, segundo o banco de dados de proteases MEROPS [34]. A classificação das proteases foi descrita por Rawling e Barrett em 1993 [35] e se baseia na estrutura molecular, homologia e função das enzimas. Atualmente, esta abordagem é considerada a mais relevante para distinguir e agrupar as peptidases [36].

De acordo com essa classificação, as aspartil proteases (APs) pertencem ao clã AA das peptidases, sendo caracterizadas por empregarem uma molécula de água ativada por dois resíduos de aspartato como nucleófilo no ataque à ligação peptídica [34, 37, 38]. Todas são endopeptidases e são encontradas em organismos diversos como vertebrados, plantas e vírus.

O clan AA apresenta duas famílias: A1 e A2. A família A2 inclui as proteases integradas por retrovírus no genoma humano, como a HIV protease. As peptidases da família A1 são descritas apenas em eucariotos e compreendem todas as enzimas homólogas à pepsina e quimosina e suas homólogas lisossomais – como as catepsinas D. A grande maioria das APs pepsina-símile apresenta melhor atividade catalítica em pH ácido, com exceção da renina [34, 39]. A pepsina é amplamente conhecida por seu papel na digestão de proteínas de alimentos no estômago de vertebrados e por ser a segunda enzima a ter sido cristalizada [38].

As enzimas pepsina-símile estão envolvidas em um grande número de processos biológicos e, por este motivo, muitas são atualmente consideradas para o planejamento de fármacos.

---

#### 1.1.3.1 FUNÇÃO BIOLÓGICA

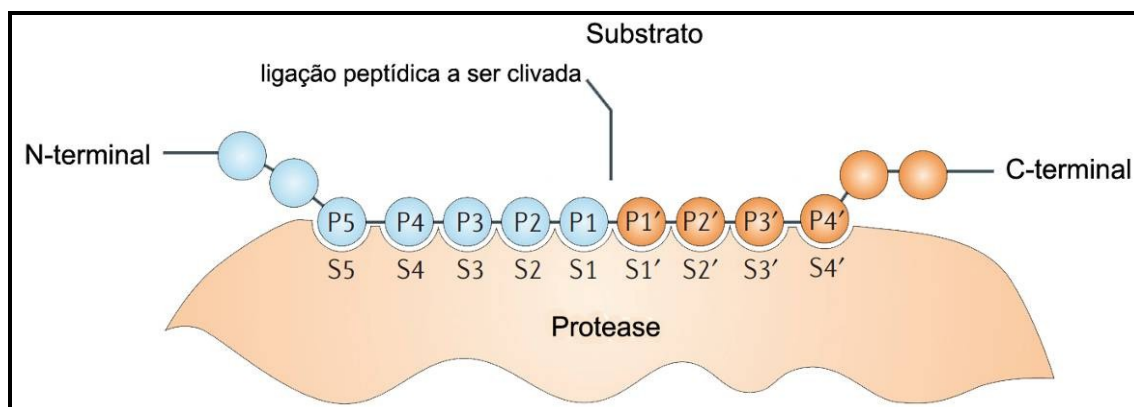
Aspartil proteases são enzimas envolvidas em diversos processos fisiológicos importantes: como na regulação da pressão arterial, doença amilóide, esporulação de fungos e na digestão de hemoglobina pelos parasitos causadores da malária e da esquistossomose [40]. Catepsinas D e E de mamíferos são aspartil proteases que



participam, dentre outros eventos, do catabolismo da Hb oriunda de eritrócitos [41-43]. Além disso, parasitas como o *Schistosoma* parecem empregar a aspartil proteases catepsinas D-símilias nos estágios iniciais de clivagem da Hb [44].

### 1.1.3.2 ESTRUTURA TRIDIMENSIONAL DE ASPARTIL PROTEASES

Grande parte das aspartil proteases de eucariotos apresenta uma única cadeia polipeptídica que forma dois domínios com enovelamentos semelhantes. O sítio catalítico está localizado na junção dos dois domínios e contém dois resíduos de ácido aspártico (Asp32 e 215, numeração da pepsina), um em cada domínio da proteína, que são mantidos coplanares por uma rede de ligações de hidrogênio conhecida como “fireman’s grip”, envolvendo as cadeias principais e laterais de resíduos vizinhos bastante conservados através da família A1 [34, 37]. A estrutura tridimensional (3D) da catepsina D humana e de outras aspartil proteases da família da pepsina compreende dois lóbulos semelhantes, basicamente com o mesmo enovelamento [45]. Cada lóbulo apresenta uma tríade catalítica homóloga caracterizada pelos resíduos Asp-Thr-Gly, com o sítio de ligação ao substrato localizado entre os dois lóbulos. Nestas proteases, o nucleófilo que ataca a ligação lábil no substrato é uma molécula de água ativada pelos dois aspartatos [46, 47]. A superfície da protease que é capaz de acomodar uma cadeia lateral de um resíduo do substrato é chamada subsítio [48] (FIGURA 1.3). Os subsítios são numerados S1-Sn em direção ao N-terminal e S1'-Sn' em direção ao C-terminal [48]. Os resíduos do substrato que a proteína acomoda são numerados como P1-Pn e P1'-Pn', respectivamente [48-50]. A estrutura do sítio ativo de cada protease define a especificidade de determinado substrato.



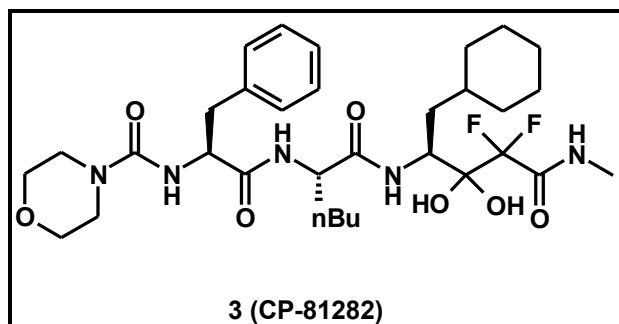
**FIGURA 1.3.** Representação esquemática do sítio de ligação de um substrato na protease. Adaptado de [50].

### 1.1.3.3 MECANISMO CATALÍTICO DE ASPARTIL PROTEASES

Já é bem estabelecido que uma molécula de água ligada às carboxilas ativas é um membro essencial do sítio ativo, sendo uma característica fundamental de todas as enzimas semelhantes à pepsina, seus complexos com inibidores e zimogênios [37, 39, 45]. Por outro lado, o estado de protonação dos resíduos catalíticos ainda é controverso, apesar de seu conhecimento ser essencial para um melhor entendimento do mecanismo catalítico desta importante classe de enzimas.

#### 1.1.3.3.1 ESTUDOS DE DIFRAÇÃO DE RAIOS X DE COMPLEXOS ENZIMA-INIBIDOR

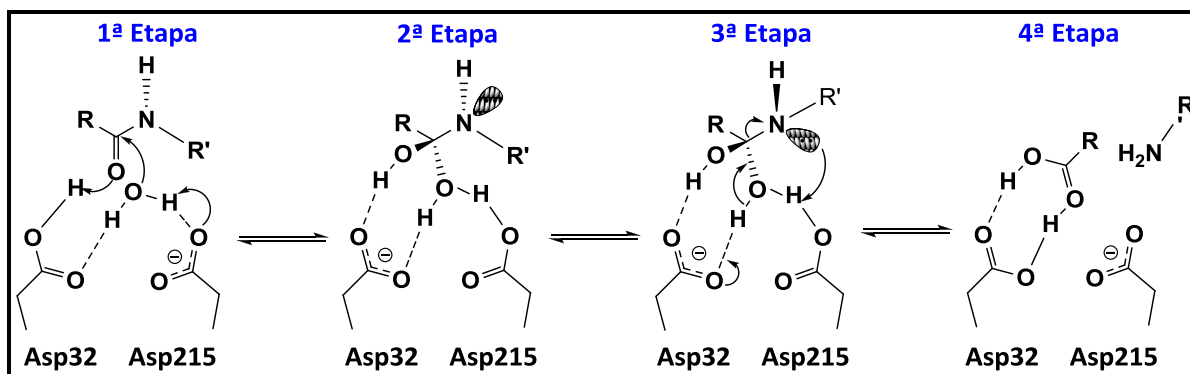
Em 1992, Veerapandian e colaboradores [47] propuseram um mecanismo de catálise de aspartil proteases, baseando-se na estrutura cristalina da endotiapepsina complexada a isómeros do estado de transição (IETs), como a substância CP-81292 (FIGURA 1.4, 3). Tanto Veerapandian e colaboradores [47] quanto Bailey & Cooper [51] demonstraram que a hidroxila presente neste inibidor e em inibidores semelhantes substitui a molécula de água encontrada no sítio catalítico da enzima livre.



**FIGURA 1.4.** Estrutura do composto CP-81292 (**3**).

A primeira etapa do mecanismo proposto inicia-se com o ataque nucleofílico da água à carbonila da ligação peptídica do substrato (FIGURA 1.5). Este ataque é acompanhado por um rearranjo do próton carboxílico do Asp32 com o oxigênio da carbonila do substrato e do ataque ao próton da água pelo oxigênio da carboxila do resíduo Asp215, formando um intermediário tetraédrico.

Na terceira etapa, o intermediário *gem*-diolamina é formado pela inversão do nitrogênio e rotação de 60° da ligação C–N, para que o mesmo fique exposto para a subsequente protonação pelo Asp215 agora protonado. Este mecanismo [47] sugere que o estado de transição é estabilizado pela carga negativa do resíduo de Asp32, baseando-se na capacidade em formar ligação de hidrogênio quando comparado ao resíduo de Asp215 em complexo com o intermediário *gem*-diolamina.



**FIGURA 1.5.** Mecanismo catalítico de aspartil proteases proposto por Veerapandian e colaboradores em 1992 [47]. Note que na passagem da 2ª para a 3ª etapa ocorre a inversão do nitrogênio e uma rotação de 60° na ligação C–N, orientando assim o orbital contendo o par de elétrons livre do nitrogênio para a captura do próton em Asp215.

O último passo consiste na protonação do nitrogênio do substrato, acompanhada pelo rearranjo de um dos prótons da hidroxila do intermediário *gem*-diolamina para o carboxilato do Asp32 e quebra da ligação C–N do intermediário. O produto ácido e a amina são formados e a configuração inicial do centro catalítico é restaurada, onde o Asp32 está protonado e o Asp215 está dissociado.

Na faixa de pH ótimo dessas enzimas é esperado que os resíduos de aspartato compartilhem uma única carga negativa [52]. O mecanismo apresentado acima pressupõe que, inicialmente, o resíduo Asp215 (que age como uma base) deve estar carregado negativamente, enquanto o Asp32 (que age como um ácido) deve estar protonado [47]. Adicionalmente, um importante papel dos resíduos adjacentes ao centro catalítico é preservar este estado dos ácidos aspárticos [45].

Rodriguez e colaboradores [53] realizaram um estudo com um substrato peptídico de HIV-1 protease (Ac-Ser-Gln-Asn-Tyr-Pro-Val-Val-NH<sub>2</sub>) e propuseram um mecanismo químico semelhante ao descrito por Veerapandian e colaboradores [47]. Os resultados mostraram coerência com os dados de cristalografia e dados cinéticos da enzima [54, 55]. Durante a ligação do substrato, o resíduo de Asp25 interage por ligação de hidrogênio com o oxigênio da carbonila do substrato, tornando o nitrogênio da prolina carregado. Com isso, o carbono da carbonila estaria acentuadamente eletrofílico para então sofrer o ataque da molécula de água. Ou seja, a transferência do próton ocorre do aspartato protonado para o oxigênio da carbonila antes do ataque nucleofílico da água, gerando uma amida O-protonada [53]. Além disso, os autores sugerem a existência de uma ligação de hidrogênio de baixa barreira (LBHB, do inglês “Low Barrier Hydrogen Bond”; ver seção 1.1.4.3.5) para a ligação O–H–O entre a carbonila e o Asp32.

Também foi proposto por Pearl & Blundell em 1984 [39], por meio de estudos de difração de raios X da endotiapepsina, que um próton é partilhado entre os ácidos aspárticos catalíticos, com a carga negativa deslocalizada em ambos os carboxilatos. Contudo, quando o substrato se liga à enzima, a simetria local é quebrada, causando a localização preferencial do próton em um dos aspartatos catalíticos, o qual sofre uma rotação e produz uma ligação de hidrogênio com a carbonila do substrato peptídico durante a reação [56].

Entretanto, os átomos de hidrogênio não podem ser localizados por análises ordinárias de estruturas obtidas por difração de raios X de proteínas. Dessa forma, embora os mecanismos descritos baseados em estruturas cristalográficas fossem

consistentes em termos gerais, faltavam evidências experimentais que comprovassem o estado de protonação dos grupos do sítio ativo. Em geral, a localização dos átomos de hidrogênio no sítio catalítico de APs pôde ser inferida por estudos de criocristalografia de raios X para regiões particularmente bem ordenadas das cadeias principais e das cadeias laterais [57].

#### 1.1.3.3.2 ESTRUTURAS DE RAIOS X DE ALTA RESOLUÇÃO

---

Uma das características principais do mecanismo proposto por Veerapandian e colaboradores em 1992 [47] é o estado de protonação dos resíduos de ácido aspártico do sítio catalítico ligados ao intermediário tetraédrico. A cristalografia de raios X de alta resolução fornece uma alternativa para a definição dos estados de protonação de carboxilatos. Grupos carboxílicos neutros apresentam diferenças entre os comprimentos de ligação C–OH e C=O (1,33 Å e 1,21 Å, respectivamente) enquanto carboxilatos ionizados apresentam comprimentos de ligação C–O iguais (1,28 Å) devido à simetria deste grupo [52, 57, 58]. Contudo, esta análise é complicada quando as regiões são desordenadas e apresentam altos fatores de temperatura. Além disso, os átomos de hidrogênio envolvidos em LBHBs estarão distribuídos entre duas posições, distantes 0,5 Å, podendo apresentar fatores de temperaturas aparentes altos.

Em 1987, foi proposto o envolvimento da transferência de um próton entre os átomos de oxigênio internos da carboxila da díade catalítica no mecanismo de catálise [59]. Entretanto, estudos de estruturas de alta resolução da endotiapepsina complexada a análogos do estado de transição demonstraram que ambos os ácidos aspárticos catalíticos apresentam diferenças significativas no comprimento de ligação C–O, característico de grupos carboxílicos protonados. Este resultado indica apenas duas possíveis posições para o próton nos resíduos de ácido aspártico catalítico. Estes podem estar posicionados no oxigênio interno do Asp32 ou no átomo de oxigênio externo do Asp215. A presença do próton nas demais posições (*i.e.*, no átomo de oxigênio externo do Asp32 ou no átomo de oxigênio interno do Asp215) não é possível, pois estas estão muito distantes para formar ligações de hidrogênio com a hidroxila do inibidor [58].

O refinamento da estrutura de três complexos com inibidores baseados na estrutura da estatina forneceram resultados sobre o comprimento de ligações muito

semelhantes [58]. Segundo os autores, o resíduo de Ser36 forma uma LBHB com o oxigênio externo de Asp32 (distância de c.a. 2,6 Å). Outra provável LBHB seria formada entre o oxigênio interno de Asp33 e a hidroxila do grupo estatina do inibidor. Para uma maior discussão sobre LBHBs em aspartil proteases, ver seção 1.1.4.3.5.

Além disso, outros estudos de resolução atômica da endotiapepsina sugeriram que a distribuição de carga no centro catalítico na ausência de substrato é simétrica - um efeito predito anteriormente por Pearl & Blundell em 1984 [39], mas não demonstrado até então. Contudo, vale destacar que o comprimento de ligação pode ser influenciado pelos fatores de temperatura dos átomos envolvidos e, além disso, é improvável que ambos os ácidos aspárticos estejam protonados devido à proximidade entre os grupos carboxílicos [57]. No complexo enzima-substrato, assume-se que os prótons da molécula de água são orientados em direção aos dois aspartatos de forma a que os pares de elétrons livres dos oxigênios da água estejam dirigidos para o substrato, favorecendo o ataque nucleofílico [52].

#### 1.1.3.3.3 EVIDÊNCIAS EXPERIMENTAIS BASEADAS EM ESTUDOS DE RMN

---

Complementando os estudos de cristalografia de raios X, importantes estudos de RMN foram realizados no intuito de elucidar o mecanismo catalítico de aspartil proteases. Em 1982, Rich e colaboradores [60] utilizou um inibidor contendo uma função cetona como análoga do estado de transição e sugeriu que esta substância se liga à enzima de forma semelhante ao intermediário *gem*-diolamina. Além disso, outros estudos demonstraram que o mecanismo catalítico das aspartil proteases envolve um intermediário que se liga de forma não-covalente à enzima [61].

Em 1991, um estudo cinético utilizando RMN de  $^{18}\text{O}$  deu suporte para o mecanismo de múltiplas etapas e simultaneamente sugeriu que o último passo, isto é, a cisão da *gem*-diolamina seria a etapa lenta da reação [55].

Cosgrove e colaboradores [62] realizaram um estudo em 1997, envolvendo a pepsina e inibidores glioxais. Concluiu-se que o principal papel das ligações de hidrogênio para os intermediários tetraédricos em aspartil proteases é orientar corretamente este intermediário para que ocorra a reação. Os autores não encontraram evidências sobre a existência das LBHBs na formação dos complexos entre a enzima e os inibidores. Entretanto este assunto é bastante controverso.

Coates e colaboradores [58] sugeriram a ocorrência de LBHBs por meio de estudos de  $^1\text{H}$ -RMN. Para mais detalhes, ver seção 1.1.4.3.5.

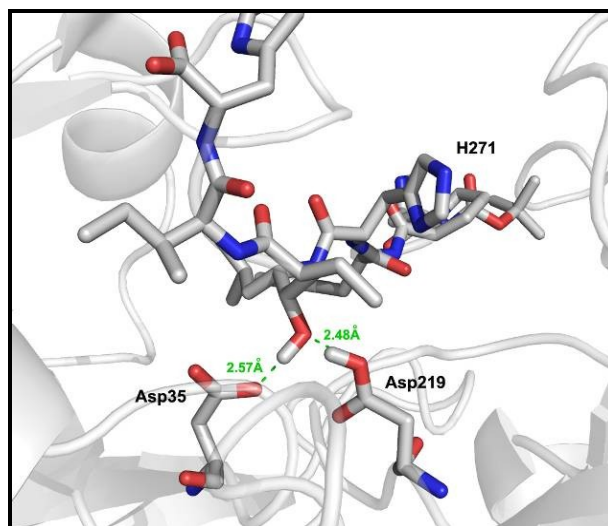
#### 1.1.3.3.4 ESTUDOS DE DIFRAÇÃO DE NÊUTRONS

---

Cristalografia de proteínas por difração de nêutrons é realizada em amostras embebidas em  $\text{D}_2\text{O}$  para permitir a troca de hidrogênio por deutério. Este último pode ser observado na difração de nêutrons, indicando a posição do próton no complexo [63]. Desta forma, esta técnica consiste em uma poderosa ferramenta na análise por raios X, pois possibilita apontar detalhes nas posições dos átomos de hidrogênio, além de revelar a estrutura do solvente [57, 64, 65]. Esta habilidade em visualizar e discriminar átomos de hidrogênio é de grande valor no estudo de mecanismos enzimáticos e o trabalho de Kossiakoff (1980) foi pioneiro nesta área [64].

A primeira evidência experimental para o mecanismo de aspartil proteases surgiu através de estudos de difração de nêutrons com cristais de endotiapepsina ligada ao inibidor H271 [57, 66]. Este inibidor apresenta um IET, o grupo hidróxietileno ( $-\text{CH}(\text{OH})\text{CH}_2-$ ), que mimetiza um dos grupamentos hidroxila do intermediário tetraédrico e substitui a molécula de água localizada no centro catalítico da enzima nativa.

O resultado da cristalografia por difração de nêutrons [57] mostrou uma densidade diferente para os átomos de deutério na hidroxila do inibidor e no oxigênio externo da carboxila de Asp219 (FIGURA 1.6). A hidroxila parece formar uma ligação de hidrogênio curta com o átomo de oxigênio externo do resíduo de Asp219, indicando que o próton (assinalado pela posição do deutério) deve se localizar entre estes átomos, o que permitiria a formação de uma ligação de hidrogênio [57]. Este estudo forneceu a primeira evidência experimental direta apoiando a distribuição de átomos de hidrogênio presentes no sítio catalítico da enzima no estado ligado ao intermediário tetraédrico, tal como aparece no mecanismo proposto por Veerapandian e colaboradores [47].



**FIGURA 1.6.** Estrutura do sítio catalítico da endotiapepsina complexada ao inibidor análogo ao estado de transição H271. Fonte: PDB, Protein Data Bank [67], código 1GKT [52].

Recentemente, Tuan e colaboradores [68] realizaram um estudo de difração de nêutrons da endotiapepsina formando um complexo com o inibidor *gem*-diol PD-135040, demonstrando que o resíduo de Asp215 apresenta o oxigênio externo protonado e que o Asp32 se encontra carregado e possivelmente fazendo uma LBHB com a hidroxila mais externa do grupamento *gem*-diol do inibidor [69]. Este resultado confirma aqueles obtidos por Coates e colaboradores em 2001 [57].

#### 1.1.3.3.5 LIGAÇÕES DE HIDROGÊNIO DE BAIXA BARREIRA

---

*Ligação de Hidrogênio de Baixa Barreira* (LBHB, do inglês “Low Barrier Hydrogen Bond”) é um tipo peculiar de ligação de hidrogênio, caracterizada principalmente pela curta distância (tipicamente menores que 2,5 Å) entre os átomos aceitadores e doadores. A proximidade entre esses tipos de átomos reduz a barreira de energia que normalmente previne a transferência do átomo de hidrogênio do grupo doador para o grupo receptor, o que facilita a troca rápida do próton entre esses grupos [70, 71]. Duas características peculiares distinguem as LBHBs das ligações de hidrogênio normais: a curta distância entre grupos doadores e aceptores, conforme determinada por cristalografia de raios X [70] e o sinal típico detectado pelo  $^1\text{H}$ -RMN entre 16 e 21 ppm [52, 58].

Com base no conceito de LBHB, Northrop [72] propôs um mecanismo catalítico para APs cuja principal característica é a presença de estrutura cíclica de



10 membros envolvendo os oxigênios externos dos aspartatos catalíticos na etapa inicial de protólise da água [72]. Acredita-se que a alta penalidade entrópica para a formação deste anel é compensada pela estabilização ocasionada pela formação da LBHB. Northrop também sugeriu uma etapa de isomerização final que retorna a enzima ao estado em que se forma a LBHB. Contudo, a estrutura proposta para a forma da enzima ligada ao intermediário tetraédrico neste mecanismo mostrou-se incompatível com os dados mais recentes de difração de nêutrons, conforme já descrito.

Rodriguez e colaboradores [53] sugeriram a existência de LBHBs para a formação do complexo enzima-intermediário durante o mecanismo de catálise. Coates e colaboradores [58] também apresentaram a hipótese da ocorrência de LBHBs entre os carboxilatos catalíticos da endotiapepsina complexada a diversos inibidores, devido à presença de inúmeros picos entre a região de 15,5 e 18,5 ppm no espectro de RMN-<sup>1</sup>H. Contudo, estes resultados foram criticados por Cosgrove e colaboradores [62], que não detectaram nenhum sinal acima de 15 ppm, sugerindo que este tipo de interação não é formado no complexo entre pepsina e pepstatina. Segundo os autores, a ausência deste tipo de interação indica que a catálise pela pepsina ocorre por um mecanismo concertado envolvendo um intermediário neutro. Logo, o aduto tetraédrico neutro forma ligações de hidrogênio mais fracas que aquelas envolvendo grupos carregados.

Embora a pepsina tenha sido uma das primeiras proteínas cuja estrutura tridimensional foi desvendada em nível atômico, os detalhes de seu mecanismo catalítico, e das aspartil proteases em geral, permanecem alvo de intensa investigação. Algumas evidências experimentais já foram alcançadas, mas muitas contradições ainda são encontradas, principalmente em relação à existência das LBHBs. Vale ressaltar que o mecanismo catalítico proposto por Veerapandian e colaboradores em 1992 [47] apresentou algumas de suas principais características confirmadas somente 10 anos mais tarde por Coates e colaboradores [57] em um estudo de difração de nêutrons.

Há uma quantidade razoável de dados experimentais que mostram que o estado de protonação da enzima livre é alterado mediante a ligação do substrato ou de IETs [47, 56, 57, 60, 69, 73-76]. Desta forma, o mecanismo de ação consistente com a maioria dos dados disponíveis é um em que um dos resíduos de aspartato catalíticos atua como ácido/base geral, ou seja, atuando num primeiro momento do

ciclo catalítico como uma base e num segundo momento como um ácido (o termo "geral" é para distinguir estratégias catalíticas que utilizam uma molécula de água, especificamente, como doador/aceptor de prótons). A característica chave deste mecanismo é a ativação da molécula de água por um dos aspartatos e subsequente protonação do átomo de nitrogênio do grupo de saída, enquanto que o outro aspartato, protonado, faz uma forte ligação de hidrogênio com o átomo de oxigênio da carbonila do substrato, onde uma carga negativa se desenvolve.

Entretanto, é cada vez mais clara a noção de que enzimas diferentes da família da pepsina podem apresentar pequenas, porém importantes variações deste mecanismo geral. Ainda, há indicações de que diferentes ligantes podem induzir diferentes estados de protonação do sítio ativo [77].

Toda essa diversidade abre um campo fértil na fronteira da pesquisa em biologia estrutural e química medicinal, pois a elucidação detalhada da catálise realizada por APs pepsina-símile é de alta relevância para o desenvolvimento de inibidores, muitos dos quais podem ser potenciais fármacos para o tratamento de doenças infecciosas, como malária e esquistossomose, ou doenças degenerativas, como Alzheimer.

---

#### 1.1.3.4 ASPARTIL PROTEASES DE *S. mansoni*

Sugere-se que uma cascata proteolítica compreendendo aspartil e cisteinil proteases esteja envolvida na degradação da hemoglobina como fonte de nutrição de helmintos como o *Schistosoma* [78, 79]. Neste contexto, propõe-se que as aspartil proteases sejam possíveis alvos para o desenvolvimento de novos fármacos contra a esquistossomose.

Em 1997, Wong e colaboradores [80] identificaram uma aspartil protease de *S. mansoni* similar à catepsina D (denominada aqui de SmCD1). A catepsina D, por sua vez, é uma aspartil protease semelhante à pepsina, pertencente à família A1 das peptidases [34]. Essa enzima é expressa em diversas células e tecidos de mamíferos e está localizada principalmente nos lisossomas.

Abordagens utilizando RNA de interferência para avaliar a função da SmCD1 foram empregadas por Delcroix e colaboradores em 2006 [81] e por Morales e colaboradores em 2008 [82]. Importantes alterações fenotípicas foram encontradas, como retardo de crescimento *in vitro* e supressão da atividade enzimática da

SmCD1 [82]. Adicionalmente, o pigmento negro de hemozoína (produto final da proteólise da hemoglobina que acumula no intestino do parasita) não estava aparente no intestino dos esquistossomulos tratados. Esses órgãos apresentaram-se com uma coloração avermelhada, indicando a presença da hemoglobina intacta. Os resultados sugeriram que a catepsina D de *S. mansoni* está envolvida nos estágios iniciais da clivagem da hemoglobina [81] e que essa enzima é essencial nas fases em que o parasita se encontra no hospedeiro humano [82]. Além disso, Brindley e colaboradores [44] demonstraram que a SmCD1 cliva sítios da hemoglobina humana diferentes dos sítios atacados pela catepsina D humana.

Estes resultados indicam que a SmCD1 é um possível alvo para intervenções anti-esquistossomicidas e destacam a especificidade para o desenvolvimento de inibidores capazes de discriminar a protease do parasito da protease homóloga mais relevante no hospedeiro humano.

Entretanto, análises de bancos de etiquetas transcritas e do genoma de *S. mansoni* indicaram que uma família de aspartil proteases está presente neste parasito (denominadas coletivamente como SmAPs), embora apenas a enzima SmCD1 seja descrita na literatura. Dentre os novos genes identificados pelo grupo do Dr. Floriano Paes Silva Jr, do Laboratório de Bioquímica de Proteínas e Peptídeos (IOC/FIOCRUZ), foram seqüenciadas integralmente duas novas SmAPs denominadas SmCD2 e SmCD3 (dados não publicados). Avanços na purificação, expressão e caracterização, no que se refere à estrutura e função destas novas enzimas, estão em constante desenvolvimento.

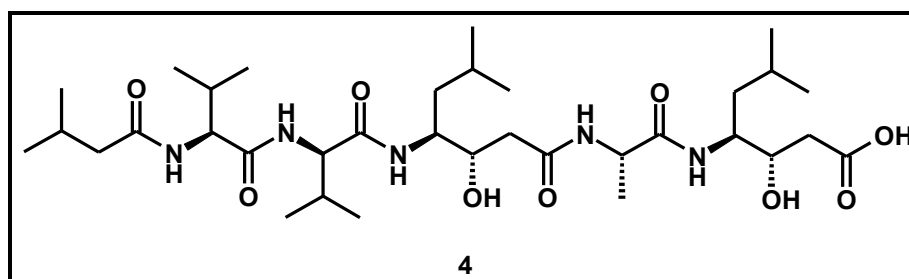
---

#### 1.1.3.5 INIBIDORES DE ASPARTIL PROTEASES

Os mecanismos pelos quais as proteases reconhecem seletivamente moléculas pequenas são conhecidos por experimentos com substratos/inibidores peptídicos [83]. Tradicionalmente, os inibidores de proteases são desenvolvidos pela triagem de produtos naturais para a identificação de compostos líderes, com subsequente otimização, ou por métodos empíricos baseados no substrato, envolvendo a redução de substratos polipeptídicos em peptídeos menores (< 10 aminoácidos) e a substituição de ligações amídicas por isósteros que não podem ser clivados. Esses tipos de inibidores, conhecidos como isósteros do estado de transição (IETs), foram utilizados como modelos para o estado de transição

tetraédrico e para auxiliar na elucidação do mecanismo catalítico de aspartil proteases [47, 56, 57, 60, 69, 73-76]. Todas as estruturas mimetizam o intermediário tetraédrico que ocorre durante a hidrólise de um dipeptídeo.

A pepstatina (FIGURA 1.7, 4) é um inibidor específico das enzimas da família A1, isolado primeiramente por Aoyagi e colaboradores em 1971 [84]. Este inibidor é um pentapeptídeo hidrofóbico que apresenta dois resíduos do aminoácido estatina (Sta). A estatina central da pepstatina foi identificada como a porção estrutural responsável por sua ação biológica [85, 86]. A hidroxila desse resíduo interage com os átomos de oxigênio OD2 de ambos os resíduos de aspartato (32 e 215), substituindo a molécula de água observada nesta posição das enzimas nativas.



**FIGURA 1.7.** Estrutura química da pepstatina (4).

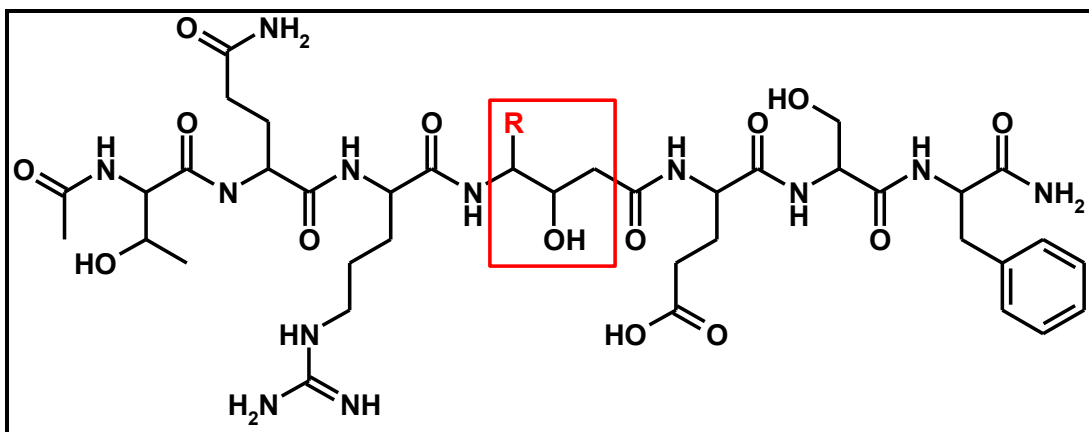
Segundo dados experimentais, os inibidores peptidomiméticos de aspartil proteases se ligam ao sítio ativo numa conformação estendida com 8-10 resíduos de aminoácidos ocupando o sítio ativo [56, 87-90]. Os inibidores que apresentam maior afinidade pelo sítio ativo são aqueles que possuem uma hidroxila capaz de mimetizar o arranjo tetraédrico do estado de transição. Esta hidroxila interage por uma ligação de hidrogênio com ambos os aspartatos catalíticos na mesma posição que uma molécula de solvente na enzima nativa [63, 91, 92], e a maioria dos análogos do estado de transição, como a estatina, mimetiza este arranjo [57]. O desenho de drogas baseado no substrato vem sendo continuamente aperfeiçoado pela disponibilidade da estrutura 3D de diversas proteases.


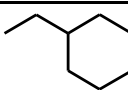
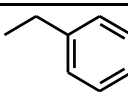
#### 1.1.3.5.1 INIBIDORES DE ASPARTIL PROTEASES DE *S. mansoni*

Em 2007, a equipe do Dr. Floriano Silva Paes Jr do Laboratório de Bioquímica de Proteínas e Peptídeos - LBPP (IOC/FIOCRUZ) desenhou um conjunto de

compostos peptidomiméticos com base no sítio de hidrólise F41-F42 da cadeia  $\beta$  da hemoglobina humana pela enzima SmCD1 [44] e na análise computacional prévia da especificidade do substrato [93]. A síntese dos compostos foi realizada em colaboração com o grupo da Dra. Maria Aparecida Juliano, da UNIFESP. Dentre todos os inibidores sintetizados, três inibidores, denominados LBPP1, LBPP37 e LBPP38, além da pepstatina, já foram avaliados experimentalmente e utilizados em estudos computacionais nesse trabalho. Os compostos da série LBPP são peptidomiméticos que diferem no resíduo central derivado de estatina (FIGURA 1.8). Os ensaios experimentais *in vitro* foram realizados no Laboratório de Bioquímica de Proteínas e Peptídeos (IOC/FIOCRUZ), quanto às suas potências de inibição da atividade de aspartil proteases presentes no extrato aquoso de vermes adultos de *S. mansoni* (SmAPs). Os ensaios foram realizados em pH ótimo para catálise (pH 3,5) [81].

Como as enzimas cisteinil proteases estão presentes nesse extrato e permanecem ativas nas condições do ensaio (pH 3,5) [79], a atividade das aspartil proteases foi selecionada adicionando-se às reações o inibidor E-64, específico para a classe de cisteinil proteases.



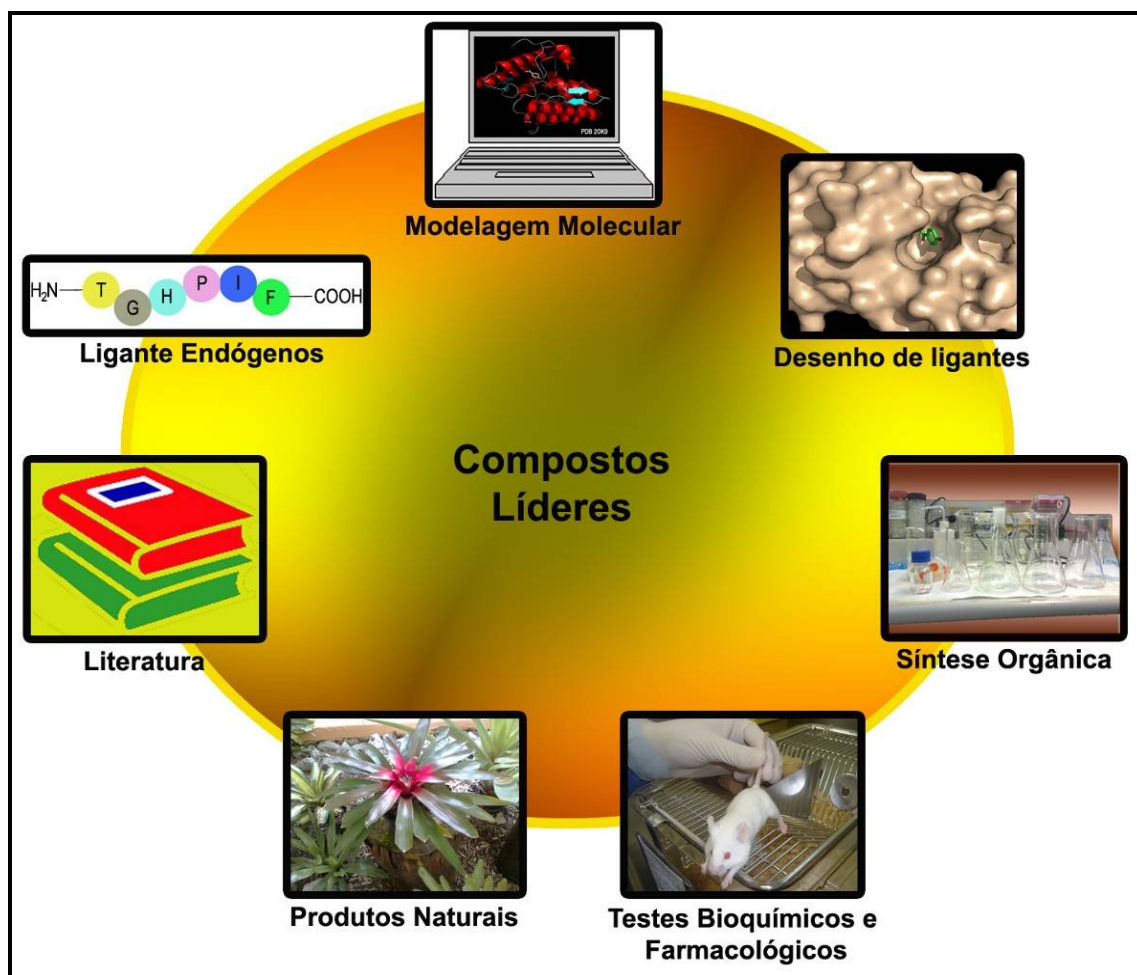
Inibidor	R	IC <sub>50</sub> ± SE (μM)	K <sub>i</sub> (nM)
Pepstatina	NA*	0,007 ± 0,002	6,6
LBPP1	 Isobutil	0,05 ± 0,02	38,6
LBPP37	 Metilciclohexil	0,94 ± 0,23	726,8
LBPP38	 Benzil	0,33 ± 0,04	251,8

\*NA: não se aplica.

**FIGURA 1.8.** Valores de IC<sub>50</sub> e K<sub>i</sub> para os inibidores peptidomiméticos utilizados nesse trabalho. Os ensaios experimentais *in vitro* foram realizados quanto às suas potências de inibição da atividade de aspartil proteases presentes no extrato aquoso de vermes adultos de *S. mansoni* (SmAPs).

## 1.2 DESENVOLVIMENTO RACIONAL DE FÁRMACOS

O processo de descoberta de novos fármacos se inicia com o conceito terapêutico que geralmente é baseado no conhecimento prévio da doença [94]. Este processo é iterativo e multidisciplinar, envolvendo diversos tipos de dados experimentais disponíveis [95], como ilustrado na FIGURA 1.9.



**FIGURA 1.9.** Estratégias para a descoberta de fármacos. As estratégias mais utilizadas são aquelas baseadas no conhecimento, derivadas da literatura, de ligantes endógenos, de informações bioestruturais ou de estudos bioquímicos e farmacológicos.

A química medicinal tem um papel inicial crucial na descoberta de novos fármacos, ao preparar novos compostos químicos e, com o adequado conhecimento da doença-alvo e de terapias farmacológicas, propor novos fármacos [96].

Durante os anos 80, os avanços na simulação computacional de sistemas biomoleculares levaram à expectativa de que tais metodologias seriam de grande utilidade para guiar o desenho molecular [97]. O planejamento de fármacos auxiliado por computador (CADD, “Computer-Assisted Drug Design”) utiliza ferramentas computacionais para auxiliar na análise das interações químicas entre um ligante e seu receptor e na exploração dos fatores estruturais relacionados ao efeito biológico [98]. O papel da modelagem molecular no desenho de fármacos está dividido em dois paradigmas distintos [99]. O primeiro está relacionado à relação estrutura-

atividade de moléculas pequenas (ligantes) que tenta racionalizar a atividade biológica na ausência da estrutura tridimensional do receptor (LBDD, do inglês “Ligand-Based Drug Design”). O segundo está focado na compreensão de interações entre receptor e ligante, utilizando-se a estrutura tridimensional do alvo terapêutico (SBDD, do inglês “Structure-Based Drug Design”).

A modelagem molecular e suas representações gráficas permitem prever e explorar aspectos tridimensionais do reconhecimento molecular e gerar hipóteses que levem ao planejamento racional e à síntese de novos ligantes, como também podem auxiliar na otimização de compostos já existentes e com atividade inibitória conhecida [100, 101]. Os métodos *in silico* podem ser aplicados a estudos de: (a) alvos como seqüência de proteínas e DNA, genomas, vias, tecidos e órgãos, (b) moléculas pequenas, e (c) complexos de ligantes com proteínas, DNA ou RNA [95].

Uma série de abordagens computacionais pode ser utilizada, tais como hipóteses geradas por farmacóforos, dinâmica molecular, mecânica quântica (QM), métodos híbridos de mecânica quântica e mecânica molecular (QM/MM), atracação molecular (muito conhecido pelo nome original, “docking”), triagem virtual, relação quantitativa estrutura-atividade (QSAR, do inglês “Quantitative Structure Activity Relationship”) etc. A seguir, estão descritos os conceitos básicos de modelagem molecular e os métodos utilizados ao longo desse trabalho.

---

### 1.2.1 CONCEITOS BÁSICOS

Modelagem Molecular, segundo o “Glossário de Termos Usados no Planejamento de Fármacos” [102], é a investigação das estruturas e das propriedades moleculares utilizando a química computacional e as técnicas de visualização gráfica, visando fornecer uma representação tridimensional, sob um dado conjunto de circunstâncias.

Em geral, pode-se dividir a modelagem molecular em métodos quânticos e métodos clássicos. Os métodos quânticos representam explicitamente os elétrons nos cálculos, sendo a energia obtida através das equações da mecânica ondulatória. Isto torna possível derivar propriedades que dependam da distribuição eletrônica e, particularmente, investigar reações químicas onde ligações são quebradas e formadas [103]. Duas classes de métodos podem ser derivadas da mecânica



quântica: *ab initio* e semi-empírico. O método semi-empírico envolve parametrizações com dados experimentais ou previamente calculados [104].

Entretanto, muitos sistemas não são viáveis de serem considerados por métodos de mecânica quântica devido ao tamanho do problema [103, 105]. Dessa forma, utilizam-se os métodos de mecânica clássica, que ignora os movimentos dos elétrons e calcula a estrutura e a energia de um sistema apenas como função das posições dos núcleos [103, 106]. Nesta metodologia, emprega-se uma função de interação denominada “campo de força” que fornece uma expressão para a energia potencial como função das coordenadas espaciais [103, 104, 106].

A escolha entre os métodos de cálculo existentes deve se basear nas propriedades que se deseja avaliar, na precisão desejada dos resultados e na capacidade computacional disponível para a realização dos cálculos [104]. Uma breve descrição das metodologias utilizadas nesse trabalho será fornecida a seguir.

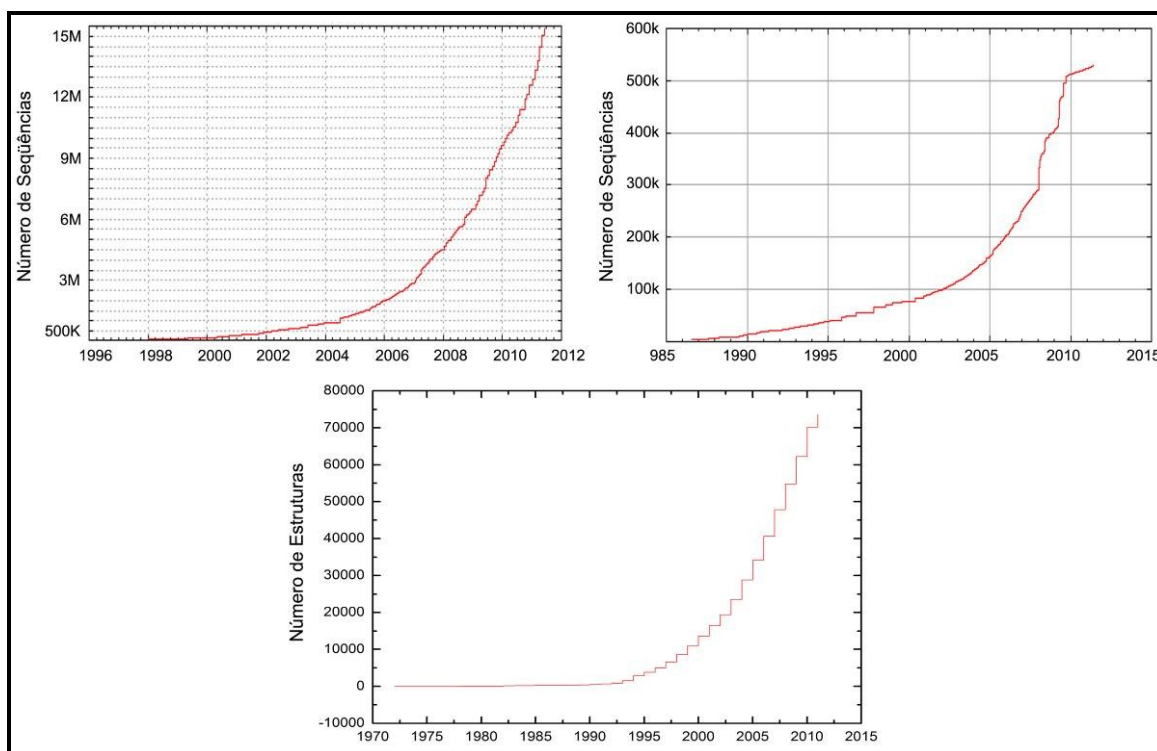
---

## 1.2.2 METODOLOGIAS DE MODELAGEM MOLECULAR APLICADAS AO PLANEJAMENTO DE FÁRMACOS

---

### 1.2.2.1 MODELAGEM COMPARATIVA OU MODELAGEM POR HOMOLOGIA

A modelagem comparativa, também conhecida como modelagem por homologia, é uma metodologia utilizada para criar modelos de estrutura de uma proteína, baseando-se no princípio de que proteínas com seqüências semelhantes apresentam estruturas tridimensionais similares [107, 108]. Como o número de estruturas resolvidas experimentalmente aumenta lentamente quando comparado ao número de novas seqüências depositadas em bancos de dados, esses métodos são de enorme interesse [109] (FIGURA 1.10). As estruturas tridimensionais de novos alvos moleculares, principalmente proteínas, enzimas e receptores são de grande importância para a compreensão dos fenômenos biológicos a nível molecular.



**FIGURA 1.10.** Dados de seqüências e estruturas depositadas nos principais bancos de dados de proteínas. **(A)** TrEMBL (Uniprot [110]) e **(B)** Swiss-Prot (Uniprot [110]); **(C)** Número de estruturas resolvidas experimentalmente no PDB [67]. Acesso aos dados: junho de 2011.

O processo de construção e validação dos modelos consiste basicamente em quatro etapas [107, 111]:

(1) Identificar estruturas tridimensionais resolvidas experimentalmente que serão utilizadas como estruturas de referência. Em geral, os modelos confiáveis podem ser gerados se apresentarem 30% ou mais de identidade com as proteínas de referência [107, 111, 112].

(2) Alinhar as seqüências-alvo e as seqüências das estruturas de referência. A etapa de alinhamento é crucial, pois erros na identificação de domínios, tais como alças e sítios de ligação do ligante, geram modelos de baixa qualidade, limitando sua utilização no processo de descoberta de novos fármacos [107].

(3) Construir modelos baseando-se no alinhamento, realizado na etapa anterior, e nas estruturas tridimensionais das proteínas de referência. A construção da parte interna da proteína-alvo (regiões conservadas) fundamenta-se na idéia de que a conformação da cadeia principal da estrutura-molde pode ser transferida para aquela. Posteriormente, as conformações das alças são avaliadas, seguida da

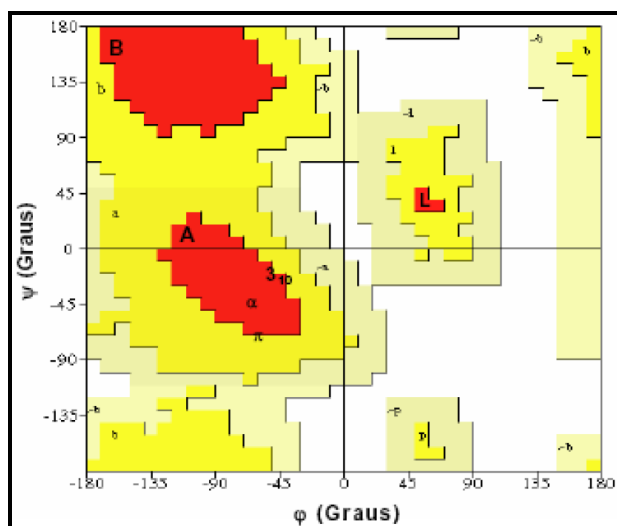
adição das cadeias laterais [111]. Os métodos principais de construção dos modelos são [109]:

- Modelagem por corpos rígidos (“modeling by rigid-body assembly”): um exemplo de programa que utiliza esse método é *Swiss-Model* [113, 114]
- Modelagem por combinação de segmentos (“modeling by segment matching”): utilizado pelo programa *Composer* [115]
- Modelagem por satisfação de restrições espaciais (“modeling by satisfaction of spatial restraints”): implementado no programa *Modeller* [116]

(4) Validar ou rejeitar os modelos construídos na terceira etapa. Diversos programas podem ser utilizados na validação de modelos obtidos por modelagem comparativa. Para validação da qualidade estereoquímica da estrutura, os softwares mais utilizados são o WHATIF [117] e o PROCHECK [118].

O software PROCHECK [118] avalia diversos parâmetros estereoquímicos e fornece gráficos, como por exemplo, o gráfico de Ramachandran (FIGURA 1.11). As estruturas secundárias de uma proteína podem ser descritas pelos ângulos de torção  $\Phi$  e  $\Psi$ , que são repetidos a cada aminoácido, ao longo da cadeia polipeptídica. Os valores permitidos desses ângulos, ou aceitáveis para estruturas tridimensionais das proteínas, são demonstrados pelo gráfico de Ramachandran.

Este gráfico é baseado na análise de 118 estruturas com resolução mínima de 2Å. Para um modelo considerado de boa qualidade, espera-se que mais de 90% dos resíduos de aminoácidos estejam nas regiões mais favoráveis.



**FIGURA 1.11.** Representação do gráfico de Ramachandran. As regiões permitidas estão mostradas para os aminoácidos: (a) regiões de hélices  $\alpha$  mais favoráveis; (a) regiões de hélices  $\alpha$  adicionalmente permitidas; ( $\sim$ a) regiões de hélices  $\alpha$  generosamente permitidas; (b) regiões de folhas  $\beta$  mais favoráveis; (b) regiões de folhas  $\beta$  adicionalmente permitidas; ( $\sim$ b) regiões de folhas  $\beta$  generosamente permitidas; (l) regiões de hélices  $\alpha$  de mão esquerda mais favoráveis; (i) regiões de hélices  $\alpha$  de mão esquerda adicionalmente permitidas; ( $\sim$ l) regiões de hélices  $\alpha$  de mão esquerda generosamente permitidas; (p) regiões para glicinas adicionalmente permitidas; ( $\sim$ p) regiões para glicina generosamente permitidas (modificado de [119]).

Outras propriedades também são usadas para a validação de modelos construídos por modelagem comparativa. Como exemplo, pode-se citar o programa PROVE [120], que avalia o volume dos átomos nas proteínas através de dados estatísticos obtidos de estruturas tridimensionais de proteínas com alta resolução. Este conjunto de dados de alta qualidade é utilizado para estabelecer volumes esperados. Os átomos das proteínas são classificados através do *Zscore*, que é o quanto o desvio padrão do volume está da média do tipo de átomo em questão.

O programa VERIFY-3D [121] determina a compatibilidade de um modelo atômico tridimensional com sua seqüência de aminoácidos pelo assinalamento de sua classe estrutural baseado no seu ambiente (hélices  $\alpha$ , folhas  $\beta$ , alças, polar e apolar etc). Com isso, o software compara esses resultados com estruturas de alta resolução.

### 1.2.2.1.1 MODELLER

---

De posse do alinhamento correto, vários métodos podem ser utilizados para a construção dos modelos tridimensionais [122, 123]. Entre eles, destaca-se o método denominado “satisfaction of spatial restraints” implementado pelo programa Modeller [116], utilizado neste trabalho. Esse programa inicia a construção do modelo utilizando restrições de distância e ângulos diedros na seqüência alvo derivadas do alinhamento com as estruturas 3D de referência [124]. Em seguida, as restrições espaciais e os termos do campo de força CHARMM22, que impõem a estereoquímica apropriada, são combinados na função objetiva. Finalmente, o modelo é gerado pela otimização da função objetiva no espaço Cartesiano [123].

---

### 1.2.2.2 ATRACAÇÃO MOLECULAR (“DOCKING”)

O processo de atracação molecular (do inglês, “docking”) consiste na avaliação da conformação e orientação de um ligante no sítio de ligação da molécula-alvo, propondo a estrutura do complexo receptor-ligante [125, 126]. Os principais objetivos desta metodologia são dois: a modelagem estrutural precisa e a estimativa da atividade de um ligante [126].

Os programas de atracação molecular contam com dois componentes complementares: um método para explorar o espaço conformacional do ligante e/ou da proteína-alvo e outro para avaliar os modos de ligação propostos (funções de pontuação) [126, 127].

Várias metodologias são empregadas na busca conformacional para considerar o modo de ligação das moléculas. Elas podem ser divididas em quatro tipos [125, 128]:

(1) Busca Sistemática: São baseados numa grade de valores para cada grau de liberdade e cada um desses valores é explorado de forma combinatória durante a busca. Um exemplo de busca sistemática é a construção incremental utilizada pelo programa FlexX [129]. Este método divide o ligante em fragmentos, um dos quais é selecionado como âncora e rigidamente colocado no sítio de ligação [125, 128]. As demais regiões são adicionadas seqüencialmente, por uma varredura dos ângulos torsionais.

(2) Busca Estocástica: Explora a conformação do ligante de forma aleatória, com alterações randômicas. Os métodos mais usuais são os algoritmos

genéticos e a busca por Monte Carlo. Os algoritmos genéticos são baseados na teoria de evolução de Darwin. As soluções de um ligante são representadas por um cromossomo e os genes codificam cada ângulo torsional [128]. Cada solução envolve a transmissão da informação genética através da reprodução, que é alterada por operadores genéticos como mutação e cruzamento. Modificações dos algoritmos genéticos foram implementadas eficientemente por alguns programas, como por exemplo, o AutoDock [130].

No método de Monte Carlo, a solução de um ligante é modificada aleatoriamente pela rotação das ligações, translação e/ou rotação do ligante, sendo a nova conformação avaliada [125, 128]. Se a nova conformação apresenta menor energia, ela é mantida. Se a conformação tiver maior energia, ela pode ser rejeitada ou pode ser usado um filtro como, por exemplo, o critério de Metrópolis. Este filtro leva em consideração a temperatura, permitindo que conformações de mais alta energia existam. MCDOCK [131] é um exemplo de programa que utiliza a metodologia de Monte Carlo como busca conformacional.

(3) Busca Determinística: Neste método, a conformação inicial determina os próximos passos para gerar uma nova conformação, que em geral é igual ou menor em energia que a inicial. Exemplos de métodos determinísticos são os métodos de minimização de energia e as simulações por dinâmica molecular [125].

(4) *Consensus*: A combinação de um ou mais métodos de busca conformacional também pode ser utilizada [128]. Por exemplo, o programa AutoxX [132] combina as funções de pontuação dos programas AutoDock e FlexX.

As funções de pontuação (“score”) são usadas como estimativas da afinidade de ligação para a molécula submetida à atração molecular [133]. Estas funções são necessárias, pois servem como ajuste na otimização da orientação e conformação do ligante. Elas são classificadas em quatro tipos [126, 128, 134]:

(1) Funções baseadas em campo de força: São derivadas dos campos de força, que originalmente foram desenvolvidos para reproduzir as propriedades conformacionais, termodinâmicas e cinéticas de complexos proteína-ligante [126, 128]. Em geral, as funções baseadas em campo de força são combinadas com os termos das funções empíricas, como termos de solvatação e entropia de ligação no AutoDock [130].

(2) Funções empíricas: Baseiam-se na idéia de que a energia de ligação pode ser decomposta em contribuições individuais, como por exemplo, ligações de

hidrogênio, efeitos hidrofóbicos, interação metal-ligante, entre outros. Entre as funções de pontuação empíricas mais comuns encontra-se ChemScore [135], implementada pelo programa GOLD [136].

(3) Funções baseadas em conhecimento: Os complexos proteína-ligante são modelados utilizando-se potenciais de interação entre pares atômicos. Tal como as funções empíricas, as funções baseadas no conhecimento se propõem a reproduzir dados experimentais [126]. Exemplos de funções baseadas no conhecimento são DrugScore [137, 138] e PMF [139, 140].

(4) *Consensus*: Assim como na busca conformacional, as funções de pontuação também podem ser utilizadas juntas para aperfeiçoar a probabilidade de identificar as melhores conformações. Um exemplo é o X-CSCORE [141].

Adicionalmente, as funções de pontuação podem apresentar simplificações que negligenciam muitos termos, como contribuições de solvatação e entropia [128]. Logo, não se pode esperar que as funções façam avaliações de afinidade precisas, mas sim previsões qualitativas. Entretanto, elas devem reconhecer soluções mostrando boa complementaridade eletrostática e estérica entre receptor e ligante e dar as melhores classificações aos compostos mais ativos. Ao contrário, os ligantes menos ativos devem ter as pontuações mais baixas. Esta etapa é crítica na descoberta de novos compostos líderes.

Atualmente, os grandes desafios na metodologia de atração molecular consistem na flexibilidade das proteínas e na ausência de funções de pontuação robustas [126, 142, 143].

#### 1.2.2.2.1 AUTODOCK

---

O AutoDock é um conjunto de programas gratuitos, desenvolvido pelo grupo de Arthur J. Olson da Universidade da Califórnia (San Diego, EUA) que realiza a atração molecular de um ligante flexível com o sítio de uma macromolécula [130]. O programa utiliza, dentre outros, o algoritmo genético Lamarckiano (LGA) como método de busca conformacional. Os algoritmos genéticos são inspirados na teoria evolucionária de Darwin. Cada busca do ótimo conformacional consiste na geração de uma população de indivíduos. Cada indivíduo representa uma possível solução para o problema de atração, que é a configuração do ligante em relação à proteína. Assim, cada configuração do ligante é representada por um cromossomo,

constituído por genes que representam os graus de liberdade translacional, orientacional e conformacional. Esses indivíduos são avaliados pela energia de interação do complexo macromolecular, proteína-ligante, e pela energia intramolecular do ligante. A seleção dos indivíduos que irão se reproduzir leva em conta esses parâmetros de energia e, na seqüência, mutações e cruzamentos (“crossovers”) poderão ocorrer para gerar novos indivíduos [144]. Uma busca local pelo método “Solis and Wets” [145] é realizada para a minimização de energia do melhor indivíduo em cada geração.

Outro parâmetro importante nos algoritmos genéticos é o elitismo. Após a formação de uma nova população, há a possibilidade de perder os melhores indivíduos. O elitismo é um método para preservar os melhores indivíduos de uma geração na geração seguinte, evitando que a nova população se torne pior do que a população atual [146].

Usualmente, múltiplas execuções empregando o LGA são realizadas em cada experimento de atração. Cada execução fornece uma solução, ou seja, um modo de interação estimado. Ao final do experimento de atração, o AutoDock agrupa as diferentes soluções em aglomerados (“clusters”) de conformações, de acordo com a semelhança entre elas e ordena os aglomerados de acordo com a menor energia representativa daquele conjunto. A diferença entre as conformações é medida pelo RMSD (do inglês, *Root Mean Square Deviation*, ou seja, Desvio Quadrático Médio das Distâncias), em angstroms (Å).

Para os cálculos de energia de interação receptor-ligante, o AutoDock utiliza um programa que permite diminuir os custos computacionais através do cálculo prévio de mapas de energia, é o chamado AutoGrid. A proteína é tratada como uma malha tridimensional cúbica, com espaçamento predefinido em angstroms [130].

Para cada tipo de átomo presente no ligante um mapa de afinidade é gerado e cada ponto de interação da macromolécula englobado nesta malha recebe um valor de energia correspondente àquela interação (átomo prova do ligante/ átomo da proteína). Os mapas são usados durante o processo de atração molecular para a determinação da energia total da interação entre o ligante e a macromolécula.

O programa AutoDock utiliza ainda uma função de pontuação (score) para calcular a energia final de cada posição do ligante, processada ao longo de cada geração. São incluídos nessa função um modelo termodinâmico do processo de ligação (que permite a inclusão de termos intramoleculares) e um modelo de



dessolvatação. Além disso, incorporam também um modelo de direcionalidade das ligações de hidrogênio. A energia livre de ligação é considerada igual à diferença entre (a) a energia do ligante e da proteína separados em seus estados livres e (b) a energia do complexo proteína-ligante [147].

O método empregado para estimar a acurácia da atracção molecular baseia-se na proximidade entre a conformação da solução de menor energia ou da família mais populosa, obtida na atracção, e aquela observada experimentalmente no complexo enzima-ligante, obtida por difração de raios X. Valores menores que 2,0 Å refletem uma solução bem sucedida [147, 148].

#### 1.2.2.2.2 ATRACÇÃO MOLECULAR DE PEPTÍDEOS E PEPTIDOMIMÉTICOS

---

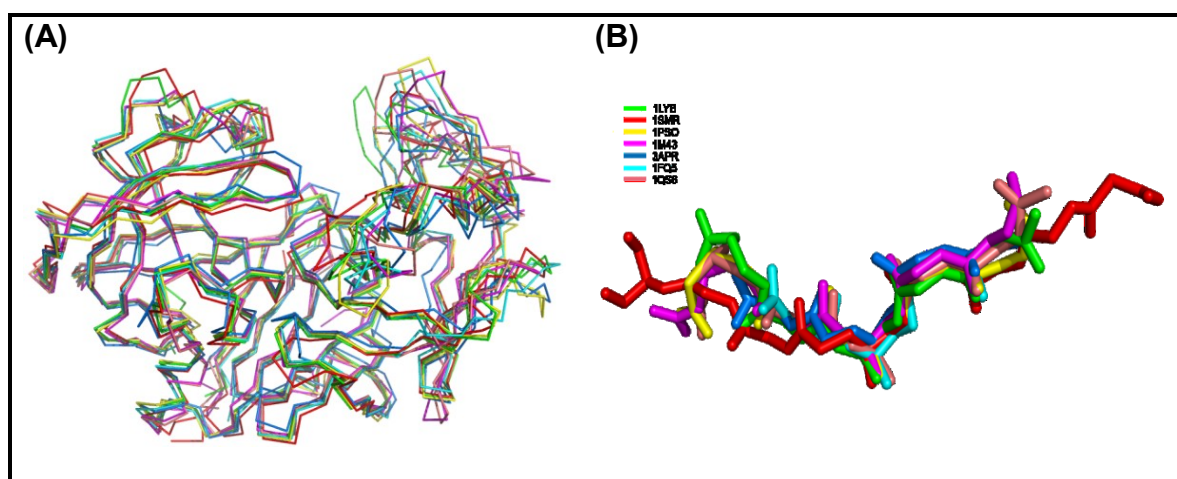
Dois aspectos contribuem para sucesso da atracção molecular no desenho de fármacos baseado na estrutura: a geração de modos de ligação razoáveis e as melhores pontuações para os compostos mais ativos [149]. Entretanto, não há ainda um programa adequado para a atracção molecular de inibidores peptidomiméticos [142]. As principais metodologias desenvolvidas para abordar esse problema serão brevemente revisadas a seguir.

Em 2004, Moitessier e colaboradores [150] desenvolveram um método híbrido entre algumas técnicas de desenhos de fármacos baseadas na estrutura e no ligante. Inicialmente, realiza-se a determinação preliminar dos farmacóforos. Em seguida, a atracção molecular desse farmacóforo na estrutura do receptor resulta num modo de ligação farmacóforo/receptor. A atracção molecular de ligantes orientados pelo farmacóforo é então obtida pela combinação de grupos de interação do ligante com os pontos farmacofóricos. A Simulação por Dinâmica Molecular (DM) em água refina os complexos propostos. O método foi validado com peptídeos e pseudopeptídeos, dentro outros compostos.

Moitessier e colaboradores [149] também desenvolveram um método de atracção molecular, orientado pelo encaixe-induzido. Segundo os autores, essa metodologia é capaz de prever a atividade de inibidores peptidomiméticos altamente flexíveis de BACE1. O método se baseia nos algoritmos genéticos, onde os cromossomos codificam o complexo como um todo – há genes que codificam a conformação do ligante e genes que codificam a conformação da proteína. A modelagem das cadeias principais e laterais é realizada durante a atracção do

inibidor. Os autores afirmaram que essa abordagem orientada é efetiva para inibidores com muitas torções livres e para o ajuste da estrutura da proteína. O método foi descrito como altamente convergente, mas com elevado custo computacional. Também foi desenvolvido no mesmo trabalho uma função de pontuação (RankScore) baseada em campo de força para a BACE1, que considera a perda entrópica da proteína, a dessolvatação do ligante e a solvatação do complexo, demonstrando capacidade para discriminar compostos ativos de inativos.

Alguns autores optaram por adotar outra abordagem: manter rígida a cadeia principal dos ligantes peptídicos [151, 152]. Uma vez que essas enzimas apresentam alta porcentagem de identidade seqüencial e estrutural entre si e os inibidores peptidomiméticos apresentam uma conformação estendida muito semelhante quando no sítio de ligação de suas respectivas enzimas [56, 87-90] (FIGURA 1.12), acredita-se que esta seja a orientação espacial mais provável do ligante.



**FIGURA 1.12.** Sobreposição da cadeia principal de **(A)** aspartil proteases em complexo com inibidores peptidomiméticos; **(B)** inibidores de aspartil proteases, quando ancorados em seus respectivos sítios ativos.

### 1.2.2.3 SIMULAÇÃO POR DINÂMICA MOLECULAR (DM)

A simulação por dinâmica molecular (DM) consiste na propagação temporal das posições atômicas de acordo com as leis da mecânica clássica [153]. Dessa forma, as simulações são utilizadas para obter informações sobre a evolução das conformações do sistema com o tempo e permitem calcular valores de propriedades estruturais e termodinâmicas do sistema [103].

Esta técnica apresenta diversas aplicações em proteínas e biomoléculas, tais como enovelamento e desnaturação proteicas; aspectos da estrutura biomolecular, reconhecimento e função; determinação e refinamento de estruturas obtidas por difração de raios X, ressonância magnética nuclear (RMN) e modelagem molecular; propriedades conformacionais; e no planejamento racional de fármacos [153-155].

A primeira simulação de uma proteína foi publicada em 1977 por McCammon e colaboradores [156]. Um modelo contendo aproximadamente 500 átomos do inibidor de tripsina bovina pancreática foi simulado num tempo total de apenas 9,2 ps. Atualmente, as simulações cobrem períodos de tempos muito superiores, com dezenas ou centenas de nanossegundos a microssegundos, e com centenas a milhares de átomos envolvidos.

A DM gera trajetórias (*i.e.* conformações como uma função do tempo) de um sistema molecular pela integração simultânea das equações de movimento de Newton para todos os átomos do sistema [155, 157]. Várias condições experimentais também podem ser consideradas na simulação, como moléculas de água explícitas e proteínas vizinhas presentes no meio cristalino [103, 153]. Comumente utilizam-se as chamadas “condições periódicas de contorno” (do inglês “Periodic Boundary Conditions”, PBC). É construída uma caixa, que pode ser cúbica, retangular, dodecaédrica ou octaédrica truncada, onde o sistema é replicado periodicamente em todas as direções para representar o sistema infinito.

Alguns programas estão disponíveis para a realização de simulação por DM, podendo-se destacar: GROMACS [158-160], TINKER [161] e AMBER [162].

#### 1.2.2.3.1 O PACOTE DE PROGRAMAS GROMACS

---

O pacote GROMACS é uma coleção de programas e bibliotecas para a simulação por DM e subsequente análise dos resultados das trajetórias geradas [158]. As equações de movimento são integradas segundo o algoritmo de *Verlet Leapfrog* e as simulações podem ser realizadas com ou sem as condições periódicas de contorno. O programa simula PBCs de geometria retangular, dodecaédrica e octaédrica truncada. Quando as condições periódicas de contorno são aplicadas, as interações de longo alcance podem ser calculadas pelo método do somatório de Ewald, raio de corte (“cut-off”) ou campo de reação (“reaction field”) [163].

O acoplamento de pressão é implementado para todos os tipos de células e a temperatura também pode ser controlada pelo termostato de Nosé-Hoover [164] ou pelo termostato de Berendsen [165]. O GROMACS também fornece suporte para restrições no processo de DM. O algoritmo mais utilizado é o SHAKE [166, 167], embora o algoritmo LINCS [168] seja mais estável.

Diversos modelos de água estão disponíveis, como por exemplo, SPC, SPC/E, TIP3P, TIP4P [158].

O campo de força GROMOS 87 [169] e GROMOS 96 [170] para simulações de biomoléculas, entre outros, são distribuídos juntamente à versão do GROMACS.

O pacote GROMACS consiste de cerca de 75 programas executáveis, sendo a maioria ferramenta para análise da trajetória ou dados de energia gerados pela DM [158]. O programa *pdb2gmx* gera os arquivos de topologia a partir de coordenadas de arquivos PDB. Esse programa assinala automaticamente as ligações, ângulos, diedros e cargas baseados no banco de dados de topologias. Como o campo de força GROMOS apresenta um conjunto de parâmetros aplicados a um número restrito de moléculas, como aminoácidos, água e íons, torna-se necessário a utilização de uma ferramenta distinta capaz de gerar os arquivos de topologia para os ligantes do complexo. Para essa finalidade, pode-se empregar o servidor *online* PRODRG [171]. Os estados de protonação de grupos ionizáveis são padronizados para pH 7, mas é possível selecionar os estados manualmente também pelo programa *pdb2gmx*.

Outro programa muito utilizado é o *genbox*, que solvata as moléculas, tanto em caixas (quando usada as condições periódicas de contorno) quanto em camadas. Além disso, o *genion* posiciona íons em potenciais eletrostáticos favoráveis.

#### 1.2.2.3.2 SIMULAÇÃO POR DINÂMICA MOLECULAR DE ASPARTIL PROTEASES DE EUCARIOTOS

---

Diversos trabalhos de DM de aspartil proteases de vírus estão descritos na literatura [172-178], mas poucos são aqueles abrangendo as aspartil proteases de eucariotos [179-181]. Em geral, os estudos envolvem os mecanismos catalíticos para a hidrólise do substrato, em especial o estado de protonação dos resíduos de aspartato catalítico, como no caso da plasmepsina IV de *Plasmodium falciparum* [181].

Em 2000, Okoniewska e colaboradores [180] investigaram o papel do resíduo Gly76 no mecanismo catalítico da pepsina, pelas mutações G76A, G76V e G76S. As DMs foram realizadas nos modelos minimizados e a flexibilidade do “flap” foi determinada baseada no desvio quadrático médio (RMSD, do inglês “Root Mean Square Deviation”) dos átomos de carbono- $\alpha$  durante os cálculos de DM. Os estudos demonstraram que o resíduo de Gly76 é essencial para a proteólise catalisada pela pepsina devido à contribuição do “flap” para a flexibilidade da região. Os autores sugeriram que essa flexibilidade é essencial para a ligação de hidrogênio do grupo hidroxila de Thr77 com o substrato. Análises das distâncias das ligações de hidrogênio revelaram que os átomos da cadeia principal de Gly76 participam da manutenção da conformação do “flap” e na geometria dos intermediários da reação. Assim, a ligação é possível pela alteração da posição do “flap”, a qual é facilitada pela Gly76. Isso sugere que Gly76 age como uma dobradiça no topo do “flap” e permite a formação de ligações de hidrogênio pelos resíduos Thr77 e Gly78. No complexo com o ligante, a posição do “flap” permite a interação do átomo de nitrogênio da cadeia principal de Gly76 com o intermediário e possivelmente contribui para a estabilização do estado de transição. Adicionalmente, os autores demonstraram que os resíduos Gly76 e Thr77 contribuem para a rede de ligações de hidrogênio responsável pelo alinhamento do substrato e pela eficiência da catálise.

Outro estudo envolvendo DM e aspartil protease de eucarioto foi publicado em 2011 por de Faria Silva e colaboradores [182]. Os autores descreveram os resultados da DM híbrida (QM/MM) entre a plasmepsina II de *P. falciparum* e o inibidor EH58, um conhecido inibidor dessa enzima. Os resultados revelaram que há forte interação entre os resíduos Asp34, Val78, Ser79, Tyr192 e Asp214, e o ligante. Além disso, os resultados indicaram o favorecimento da protonação do resíduo catalítico Asp214.

## 2 OBJETIVOS

O presente trabalho teve como objetivo avaliar as possíveis interações entre SmAPs e os inibidores da série LBPP, além da pepstatina, auxiliando o desenvolvimento racional de novos inibidores antiparasitários.

### 2.1 OBJETIVOS ESPECÍFICOS

- Obtenção das estruturas tridimensionais de SmAPs codificadas pelos novos genes identificados (SmCD2 e SmCD3) por modelagem;
- Atracção molecular nas estruturas de SmCD1, SmCD2 e SmCD3 de inibidores peptidomiméticos, previamente sintetizados e experimentalmente avaliados;
- Análise por dinâmica molecular das estruturas de SmCD1, SmCD2 e SmCD3 livres e em complexo com a pepstatina, visando a caracterização e diferenciação das interações desse inibidor com as diferentes aspartil proteases do parasita.

## 3 METODOLOGIA

### 3.1 MODELAGEM COMPARATIVA

As aspartil proteases de *Schistosoma mansoni* 1, 2 e 3 (SmCD1, SmCD2 e SmCD3, respectivamente) ainda não apresentam estrutura 3D determinada experimentalmente. Dessa forma, utilizou-se a metodologia conhecida como *Modelagem Comparativa* que pode fornecer um modelo tridimensional útil de uma proteína relacionada a pelo menos uma estrutura conhecida.

#### 3.1.1 ANÁLISE DA SEQÜÊNCIA

As seqüências das aspartil proteases de *S. mansoni* (SmCD1, 2 e 3) similares à catepsina D humana foram obtidas a partir do banco de dados de proteínas TrEMBL [110], sob códigos P91802, Q2Q0I8 e B0L5P6, respectivamente. O peptídeo sinal foi removido de acordo com o programa Signal3P Server [183] (Anexo I) e o pró-peptídeo foi eliminado baseando-se nas similaridades entre o alinhamento de seqüências obtidas no banco de dados Swissprot [110] (Anexo II).

#### 3.1.2 IDENTIFICAÇÃO E SELEÇÃO DE PROTEÍNAS DE REFERÊNCIA

Inicialmente, é necessária a identificação de uma ou mais proteínas de estrutura tridimensional conhecida, que serão utilizadas como referência na construção dos modelos tridimensionais. Para isso, as seqüências da SmCD1, SmCD2 e SmCD3 foram submetidas a uma busca contra o PDB [67] utilizando as opções padrão do programa BLASTP [184].

#### 3.1.3 ALINHAMENTO DAS SEQÜÊNCIAS DE RESÍDUOS DE AMINOÁCIDOS

Uma vez selecionadas as referências, um método especializado deve ser usado para alinhar a seqüência alvo com as referências estruturais [185]. O objetivo é alinhar resíduos equivalentes levando em conta características comuns, como

estruturas altamente conservadas nas famílias de proteínas [186]. Os alinhamentos das enzimas SmCD1, SmCD2 e SmCD3 com as proteínas de referência foram obtidos pelo programa T-Coffee Server [187] e orientados pelos aminoácidos altamente conservados na família, incluindo a tríade catalítica Asp-Gly-Thr e os pares de cisteínas [34].

---

### 3.1.4 CONSTRUÇÃO DOS MODELOS

De posse do alinhamento obtido na seção anterior, realizou-se a construção dos modelos tridimensionais pelo programa Modeller 8v2 [116], que utiliza o método de satisfação por restrição espacial. Foram construídos 50 modelos de cada enzima, dos quais apenas um foi selecionado, baseando-se na função objetiva e nos resultados de validação obtidos pelos programas descritos na seção 3.1.5.

---

### 3.1.5 VALIDAÇÃO DOS MODELOS

Em seguida à construção do modelo, é necessário verificar a qualidade dos mesmos, pois essa avaliação determina a sua utilidade [185]. A qualidade estereoquímica e geral dos modelos foi realizada usando os programas PROCHECK [118], PROVE [120] e VERIFY-3D [121].

## 3.2 ATRACAÇÃO MOLECULAR

---

### 3.2.1 “RE-DOCKING” E “CROSS-DOCKING”

Para o estudo de “re-docking” e “cross-docking”, foram utilizados complexos cujos ligantes são estruturalmente semelhantes àqueles estudados nesse trabalho e cujas enzimas apresentam as maiores identidades seqüenciais com as enzimas-alvo (*i.e.*, SmCD1, SmCD2 e SmCD3).

As enzimas catepsina D humana e renina de camundongo foram usadas para o “re-docking”. Nesses testes o ligante é “re-atracado” na conformação da proteína com a qual ele foi co-cristalizado e que já está na conformação (de ligação) encontrada no complexo obtido por difração de raios X.



O “cross-docking” foi usado para o mesmo efeito de protocolo, desta vez com apenas uma enzima, a pepsina humana, sob dois códigos diferentes no PDB, 1QRP e 1PSO, co-cristalizada respectivamente com um ligante peptídico fosfatado PHO e a pepstatina. Nos experimentos de “cross-docking”, um ligante A de um complexo proteína-ligante A é atracado a um receptor B, de um complexo proteína-ligante B, onde o receptor de ambos é uma mesma proteína [125].

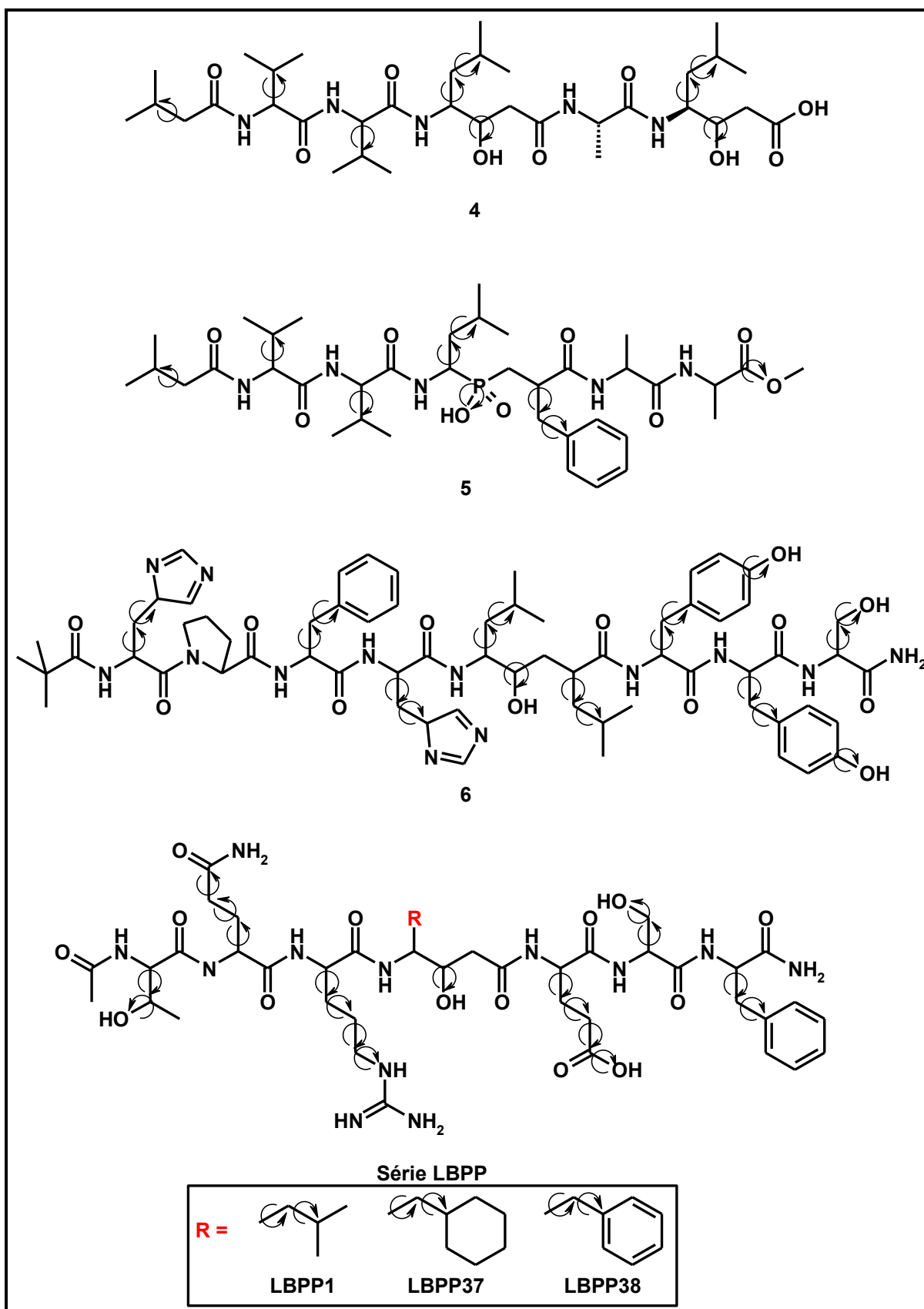
A TABELA 3.1 apresenta um resumo dos complexos obtidos do PDB.

**TABELA 3.1.** Complexos utilizados no “re-docking” e “cross-docking”.

<b>Protocolo</b>	<b>Proteína</b>	<b>Ligante</b>	<b>Código PDB</b>	<b>Resolução (Å)</b>
<b>“re-docking”</b>	Catepsina D humana	Pepstatina	1LYB	2,5
	Renina de Camundongo	CH-66	1SMR	2,0
<b>“cross-docking”</b>	Pepsina Humana	PHO	1QRP	1,9
	Pepsina Humana	Pepstatina	1PSO	2,0

### 3.2.2 PREPARAÇÃO DOS LIGANTES

Os cálculos de “re-docking” foram realizados com os inibidores peptidomiméticos pepstatina e CH-66 (FIGURA 3.1) em complexo com as enzimas catepsina D humana e renina de camundongo, respectivamente. Os complexos foram obtidos pelo banco de dados de estruturas tridimensionais de proteínas (PDB, “Protein Data Bank” [67]), sob código de acesso 1LYB e 1SMR, respectivamente. Os cálculos de “cross-docking” foram realizados a partir de complexos, também obtidos pelo PDB, da pepsina humana com dois inibidores peptidomiméticos distintos: pepstatina (código PDB 1PSO) e PHO (código PDB 1QRP) (FIGURA 3.1).



**FIGURA 3.1.** Estrutura química dos inibidores pepstatina (4), PHO (5), CH-66 (6) e a série LBPP. As ligações tratadas como flexíveis estão indicadas pelas setas.

Os três inibidores peptidomiméticos (LBPP1, LBPP37 e LBPP38) foram construídos pelo módulo *Sketch Molecule* do pacote Sybyl v8.0 (Tripos Associates, St. Louis, MO). A estrutura cristalográfica do inibidor decapeptídico CH-66, obtido pelo PDB sob código 1SMR, foi utilizada como base.

A preparação de todos os ligantes foi realizada no pacote Sybyl v8.0 (Tripos Associates, St. Louis, MO). Os átomos de hidrogênio foram adicionados à estrutura e 1000 passos de minimização de energia foram executados pelo método *Steepest Descent*, usando cargas de *Gasteiger-Hückel*, constante dielétrica 80 e campo de força *Tripos* [188]. Em seguida as estruturas foram otimizadas através do método de Gradiente Conjugado.

Também foram selecionadas as ligações flexíveis e não flexíveis dos ligantes (FIGURA 3.1). Dessa forma, durante o processo, permitiu-se mobilidade apenas às cadeias laterais, mantendo-se a cadeia principal rígida. Assim, as diferentes cadeias laterais livres poderão se posicionar no sítio ativo buscando as posições energeticamente e espacialmente mais favoráveis, durante os cálculos de atração molecular. Alguns autores já utilizaram essa estratégia para atração de ligantes peptídicos em aspartil proteases utilizando o programa AutoDock [151, 152].

---

### 3.2.3 PREPARAÇÃO DAS MACROMOLÉCULAS

As estruturas 3D da pepsina humana, da catepsina D humana e da renina de camundongo foram extraídas do PDB, sob códigos 1QRP e 1PSO; 1LYB; e 1SMR, respectivamente. As aspartil proteases de *Schistosoma mansoni* SmCD1, SmCD2 e SmCD3 foram obtidas por modelagem comparativa, como descrito na seção 1.2.2.1.

Os átomos de hidrogênio foram adicionados às estruturas das proteínas no pacote Sybyl v8.0 (Tripos Associates, St. Louis, MO) e 1000 passos de minimização de energia foram executados pelo método *Powell*, usando cargas *Amber99*, constante dielétrica 80 e campo de força *Tripos* [188]. Em seguida as estruturas foram otimizadas pelo método de Gradiente Conjugado.

---

### 3.2.3.1 PROTONAÇÃO

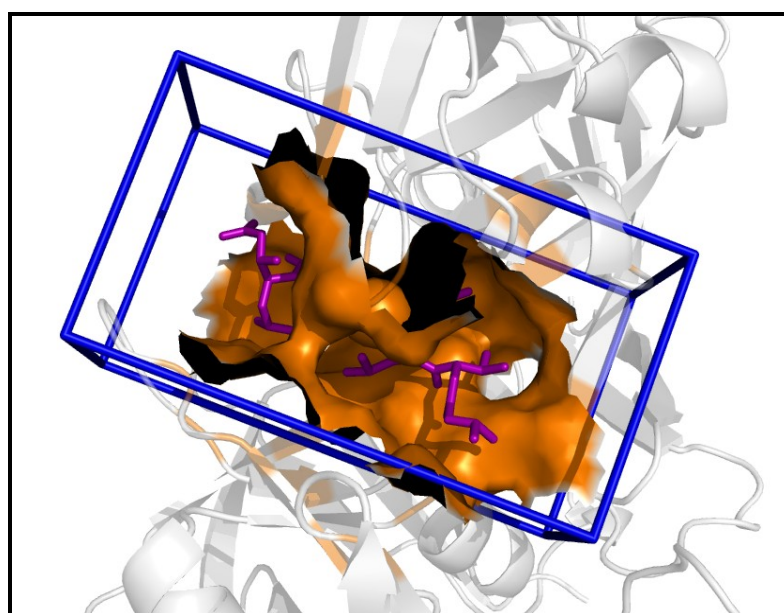
Dentre as condições dos ensaios experimentais, destaca-se o pH 3,5 necessário para a atividade de aspartil proteases em *S. mansoni* [79]. Como a protonação das cadeias laterais dos aminoácidos é um fator importante na dinâmica das moléculas, utilizou-se o programa PROPKA [189] para verificar quais os aminoácidos estariam protonados em pH 3,5. Considerou-se protonados os aminoácidos com pKa calculado acima de 5,0. Estes aminoácidos estão listados nas TABELAS 1 a 6 do Anexo III.

Outros aminoácidos cuja protonação deve ser considerada, independente do pH em questão, são os ácidos aspárticos do sítio catalítico. Para o complexo, além dos resíduos das TABELAS 1 a 6 do Anexo III, o oxigênio externo do Asp219 (Asp215 na numeração da pepsina) foi considerado protonado e o Asp33 (Asp32 na numeração da pepsina) desprotonado, de acordo com as evidências experimentais para outras aspartil proteases da mesma família [52, 57, 58, 69] que confirmam o mecanismo proposto por Veerapandian e colaboradores [47].

---

### 3.2.4 PARÂMETROS DA ATRACAÇÃO MOLECULAR

Os estudos de atracção molecular foram realizados utilizando-se o programa AutoDock 4.2 [130]. A interface gráfica do AutoDock Tools foi utilizada para editar os arquivos de entrada. Cargas Gasteiger foram assinaladas para a SmCD1 e os inibidores. Grades com 40 X 74 X 42 pontos (15,00 X 27,75 X 15,75 Å<sup>3</sup>), centradas no ligante, foram calculadas usando-se o programa AutoGrid (FIGURA 3.2). Considerou-se o espaçamento padrão entre os pontos da grade (0,375 Å).



**FIGURA 3.2.** Representação da grade calculada pelo programa AutoGrid. Em destaque a superfície de acesso ao solvente (laranja) dos resíduos da SmCD1 a até 4,0 Å de distância da pepstatina (em roxo).

Para estabelecer um protocolo, algumas variáveis foram modificadas ao longo dos inúmeros cálculos de atracção molecular (“re-docking” e “cross-docking”), até que se obtivessem os melhores resultados. Como a posição do ligante é conhecida através de dados obtidos por difração de raios X, buscaram-se variáveis que gerassem conformações mais próximas àquela obtida experimentalmente. População inicial, número de avaliações de energia, elitismo, número de execuções, taxas de mutação e de “crossover” e algoritmo de busca foram os parâmetros analisados (TABELA 3.2).

**TABELA 3.2.** Variáveis investigadas no processo de “re-docking” e “cross-docking”.

Variáveis	Faixa de valores analisados
População Inicial	100; 250; 500; 1000
Número de Avaliações de Energia	$5 \times 10^5$ ; $1 \times 10^6$ ; $5 \times 10^6$ ; $1 \times 10^7$
Elitismo	1 e 10%
Número de Execuções	100; 150; 200
Taxa de Mutação	0,1; 0,2; 0,4; 0,8; 1,0
Taxa de “Crossover”	0,4; 0,6; 0,8; 0,9
Algoritmo Genético	GA e LGA

### 3.2.5 ANÁLISE DOS RESULTADOS

As soluções encontradas pela metodologia de atracção molecular são agrupadas em um mesmo aglomerado (“cluster”) caso o RMSD não exceda 2,0 Å, quando comparadas entre si. Dentro dos aglomerados, as estruturas são ordenadas a partir daquela de menor energia.

Foram analisadas as conformações dos aglomerados mais populosos, assim como as soluções de menor energia. Também foi realizada uma inspeção visual das soluções possíveis, visando analisar os contatos entre os resíduos altamente conservados dentro do sítio ativo, como por exemplo, as ligações de hidrogênio existentes entre a hidroxila do inibidor e os aspartatos catalíticos.

### 3.3 ANÁLISE E COMPOSIÇÃO DE SUBSÍTIOS

Após a etapa de atracção molecular dos inibidores no sítio de ligação das enzimas, seguiu-se para o mapeamento dos subsítios. Os mesmos foram identificados como os resíduos das proteínas distantes até 4,0 Å dos resíduos dos ligantes.

### 3.4 SIMULAÇÃO POR DINÂMICA MOLECULAR

A Simulação por Dinâmica Molecular (DM) foi realizada nos sistemas livres e nos complexos enzima-inibidor. Para a análise das enzimas SmCD1, SmCD2 e SmCD3 livres e seus complexos com a pepstatina (FIGURA 1.7, 4) utilizou-se a trajetória obtida pela DM descrita aqui.

As DMs foram iniciadas com a estrutura 3D obtida na seção 1.2.2.1, utilizando-se o pacote de programas GROMACS [159, 190], no *ensemble* NPT e campo de força GROMOS 96. O arquivo de topologia molecular necessário para o ligante pepstatina foi gerado pelo servidor PRODRG [191].

Como o programa pdb2gmx do pacote GROMACS utiliza como padrão pH 7,0 para a protonação dos aminoácidos, também utilizou-se o método discutido na seção 3.2.2.1 para a protonação referente ao pH 3,5. Além dos resíduos indicados pelo programa PROPKA [189], o átomo de oxigênio externo do resíduo de Asp33 foi protonado para a realização dos estudos de dinâmica molecular com a enzima livre.

Para o complexo, além dos resíduos das TABELAS 1 a 6 do Anexo III, o átomo de oxigênio externo do Asp215 foi considerado protonado e o Asp33 desprotonado.

As proteínas foram solvatadas em uma caixa retangular construída com arestas distantes, pelo menos, 9,0 Å da superfície da proteína. Moléculas de água SPC [192] foram usadas para esse propósito. Os sistemas foram neutralizados pela substituição das moléculas de água por íons  $\text{Na}^+$  ou  $\text{Cl}^-$ . Restrição das ligações envolvendo átomos de hidrogênio foi realizada pelo algoritmo LINCS [163]. Condições periódicas de contorno foram aplicadas e energias de interações de longo alcance foram calculadas pelo método de Ewald (PME) [193].

A simulação das moléculas de água e íons foi realizada a 310 K e 1 atm por 100 ps com o objetivo de equilibrá-los no campo do potencial exercido pela proteína e para preencher as cavidades da superfície da proteína.

A temperatura e a pressão foram mantidas constantes e controladas usando técnicas de acoplamento [165]. As constantes de acoplamento e compressibilidade isotérmica, assim como as demais condições utilizadas nas DMs, são mostradas na TABELA 3.3.

Em seguida, foram realizadas minimizações de energia, utilizando-se 2000 passos do algoritmo *Steepest Descent*, para remover contatos de alta repulsão. As equações de movimento foram integradas utilizando-se o algoritmo de *Verlet Leapfrog* [194].

O sistema foi aquecido de 50 K a 310 K em seis blocos, nos quais a temperatura foi gradativamente aumentada. Cada bloco teve duração de 20 ps, totalizando 120 ps de aquecimento para cada sistema. Essa etapa, denominada de *termalização*, é realizada com o intuito de uniformizar as energias contidas nas estruturas iniciais, evitando a deformação da proteína [195].

Em seguida, a fase de produção da DM, na temperatura final desejada, foi iniciada com um passo de integração de 0,002 ps. Todos os sistemas foram simulados por 30 ns.

**TABELA 3.3.** Condições usadas nas Simulações por Dinâmica Molecular (DM).

<b>Condições</b>	<b>Valores</b>
<b>Tempo total</b>	30 ns
<b>Intervalo de integração</b>	0,002 ps
<b>Tratamento das interações eletrostáticas</b>	PME
<b>Raio de corte das interações eletrostáticas e de <i>van der Waals</i></b>	9,0 Å
<b>Acoplamento de temperatura</b>	Berendsen
<b>Tempo de acoplamento</b>	0,1 ps
<b>Banho térmico</b>	310 K
<b>Acoplamento de pressão</b>	Berendsen
<b>Tempo de acoplamento</b>	0,5 ps
<b>Compressibilidade isotérmica</b>	$4,5 \times 10^{-5} \text{ bar}^{-1}$
<b>Pressão de referência</b>	1 atm

### 3.4.1 ANÁLISE DOS RESULTADOS

Todas as análises foram realizadas por programas inseridos no pacote GROMACS. Os valores de RMSD foram obtidos pelas cadeias principais das enzimas. O raio de giro foi calculado para as proteínas completas, indicando a compactação da estrutura como um todo.

Para o cálculo das ligações de hidrogênio, utilizou-se uma distância máxima de 3,5 Å e um ângulo máximo de 30°. A formação de interações hidrofóbicas e pontes salinas foi avaliada pela distância máxima de 4,0 Å entre os átomos das cadeias laterais consideradas.

A partir desses conceitos, realizou-se uma análise da ocorrência (porcentagem) dessas interações na trajetória de cada simulação, indicando a estabilidade de cada uma destas.

### 3.5 CONSTRUÇÃO DE FIGURAS E GRÁFICOS

Para a edição das figuras de enzimas, ligantes e complexos enzima-ligantes utilizou-se o programa PYMOL (The PyMOL Molecular Graphics System, Version 0.99, Schrödinger, LLC). Para edição dos gráficos, os dados foram importados foram construídos utilizando-se o programa ORIGINPRO software (Origin Lab Corp., Northampton, MA).



## 4 RESULTADOS E DISCUSSÃO

### 4.1 MODELAGEM COMPARATIVA

#### 4.1.2 SELEÇÃO DAS PROTEÍNAS DE REFERÊNCIA

O primeiro fator observado para a escolha das proteínas de referência foi em relação à família de aspartil proteases, pois devem pertencer à família A1 das peptidases. Em seguida, foi analisada a identidade entre as seqüências-alvo e as seqüências de referência, assim como a resolução das estruturas, o número de “gaps” e o número de resíduos alinhados pelo programa BLASTP [184].

No caso da SmCD1, observa-se um menor número de resíduos alinhados com as seqüências das estruturas de referência, quando comparado à SmCD2 e SmCD3. Isto se deve à região C-terminal dessa enzima, a qual não existe estrutura obtida experimentalmente depositada no PDB.

Fatores estruturais também foram analisados. Todas as aspartil proteases apresentam uma estrutura flexível em forma de “flap”, que está aberto quando a enzima se encontra livre. Entretanto, essa região se fecha quando o substrato/inibidor se liga, enclausurando o ligante e formando um túnel composto pelo sítio catalítico e pelos subsítios da proteína. A função do “flap” possivelmente é excluir as águas do sítio catalítico e ligar o substrato/inibidor [196]. Dessa forma, para a etapa posterior de atracação molecular (seção 3.2), as proteínas em estudo devem estar numa conformação apropriada para a ligação de inibidores, ou seja, o “flap” deve estar estendido sobre o sítio de ligação. Assim, foram selecionadas proteínas complexadas a inibidores para a construção dos modelos das enzimas SmCD1, SmCD2 e SmCD3 (TABELA 4.1).

O alinhamento de seqüências apresenta um importante papel na geração de modelos por homologia acurados. Com o decréscimo da similaridade seqüencial, os alinhamentos passam a conter um número maior de gaps e erros de alinhamento [185]. Os métodos de alinhamentos par-a-par buscam por alinhamentos ótimos entre duas seqüências, como por exemplo, o programa BLAST [184]. Essa abordagem é útil para a busca por seqüências homólogas em banco de dados. Já os métodos de alinhamento múltiplo alinham simultaneamente várias seqüências para identificar regiões conservadas, prever sítios funcionais e funções da proteína [197]. Essa

abordagem é particularmente adequada para proteínas com baixa identidade seqüencial (< 40%). Em casos onde uma única estrutura de referência não proporciona a informação estrutural completa para o alvo, a abordagem com seqüências múltiplas poderia aperfeiçoar a qualidade do modelo [107, 123]. Além disso, foi proposto que o uso de múltiplas seqüências de referência seria útil, pois não é trivial identificar a melhor de duas ou mais proteínas de referência [198]. No caso das enzimas SmCD1 e SmCD2, embora a catepsina D humana apresente mais de 40% de identidade, é importante ressaltar que a utilização de outras proteínas como referência auxiliam na construção de regiões deficientes ou com similaridades pontuais baixas. Esse é o caso da fita s7C na SmCD1 (correspondente à s8C na SmCD2). Além disso, as demais estruturas apresentam resoluções melhores que aquela apresentada pela catepsina D humana (código PDB 1LYB).

Um número máximo de três estruturas de referência foi selecionado nesse trabalho para a construção dos modelos, pois segundo a literatura, o programa Modeller produz modelos mais deteriorados quando utilizado mais de três [199]. Além disso, o programa é capaz de selecionar as melhores regiões de mais de dois alinhamentos e combiná-las em uma estrutura aperfeiçoada.

No caso da SmCD3, foram selecionadas duas reninas pois as demais proteases apresentaram menor identidade seqüencial.

Os complexos empregados como referência são encontrados no PDB [67]: catepsina D humana (código PDB 1LYB), pepsina humana (códigos PDB 1QRP e 1HRN) e renina de camundongo (código PDB 1SMR). A TABELA 4.1 descreve as proteínas utilizadas como referência para a construção dos modelos das formas maduras das enzimas SmCD1, SmCD2 e SmCD3, bem como a identidade entre elas.

**TABELA 4.1.** Proteínas utilizadas como referência para a construção dos modelos das enzimas SmCD1, SmCD2 e SmCD3.

Proteína Alvo	Proteína de Referência/Inibidor	Código PDB	Resolução (Å)	Gaps (%)	Identidade (%)	Número de resíduos alinhados
SmCD1	Renina de camundongo/CH-66	1SMR	2,00	1	45	331/377
	Catepsina D humana/pepstatina	1LYB	2,50	2	57	331/377
	Pepsina humana/PHO	1QRP	1,96	3	47	332/377
SmCD2	Renina de camundongo/CH-66	1SMR	2,00	1	38	327/339
	Catepsina D humana/pepstatina	1LYB	2,50	1	44	332/339
	Pepsina humana/PHO	1QRP	1,96	3	42	335/339
SmCD3	Renina humana/poliidroximonoamida	1HRN	1,80	3	33	325/332
	Renina de camundongo/CH-66	1SMR	2,00	3	33	328/332
	Catepsina D humana/pepstatina	1LYB	2,50	2	38	326/332

<sup>1</sup>Número de resíduos alinhados/Número total de resíduos da enzima.

#### 4.1.3 ALINHAMENTO DAS SEQÜÊNCIAS DE AMINOÁCIDOS

Para o alinhamento múltiplo entre as seqüências-alvo e as proteínas de referência (FIGURAS 4.1, 4.2 e 4.3), utilizou-se o programa T-Coffee [187], um dos programas disponíveis mais acurados [197].

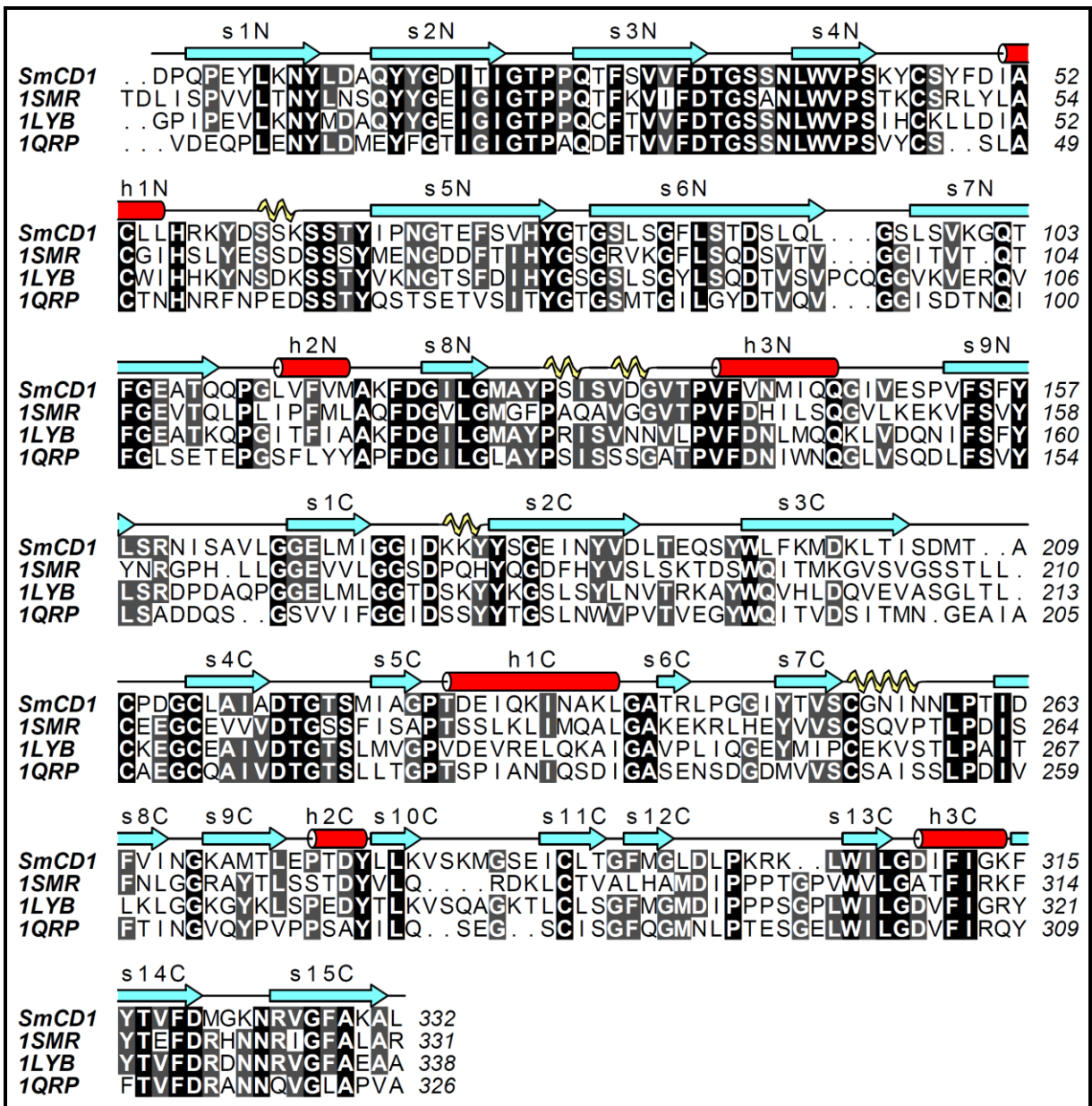
Comparou-se o alinhamento das seqüências das enzimas SmCD1, SmCD2 e SmCD3 com os dados disponíveis para proteínas da família à qual as mesmas pertencem. O sítio de ligação das aspartil proteases, composto por duas tríades catalíticas em cada lóbulo (D-T-G) foi localizado nas três enzimas estudadas. Além disso, os resíduos vizinhos ao sítio catalítico também são idênticos ou com as mesmas características físico-químicas entre as seqüências das proteínas de referência e das proteínas-alvo.

Os três pares de cisteínas conservados na família A1 de aspartil proteases foram identificados para as enzimas SmCD1 e SmCD3. Desses três pares, apenas

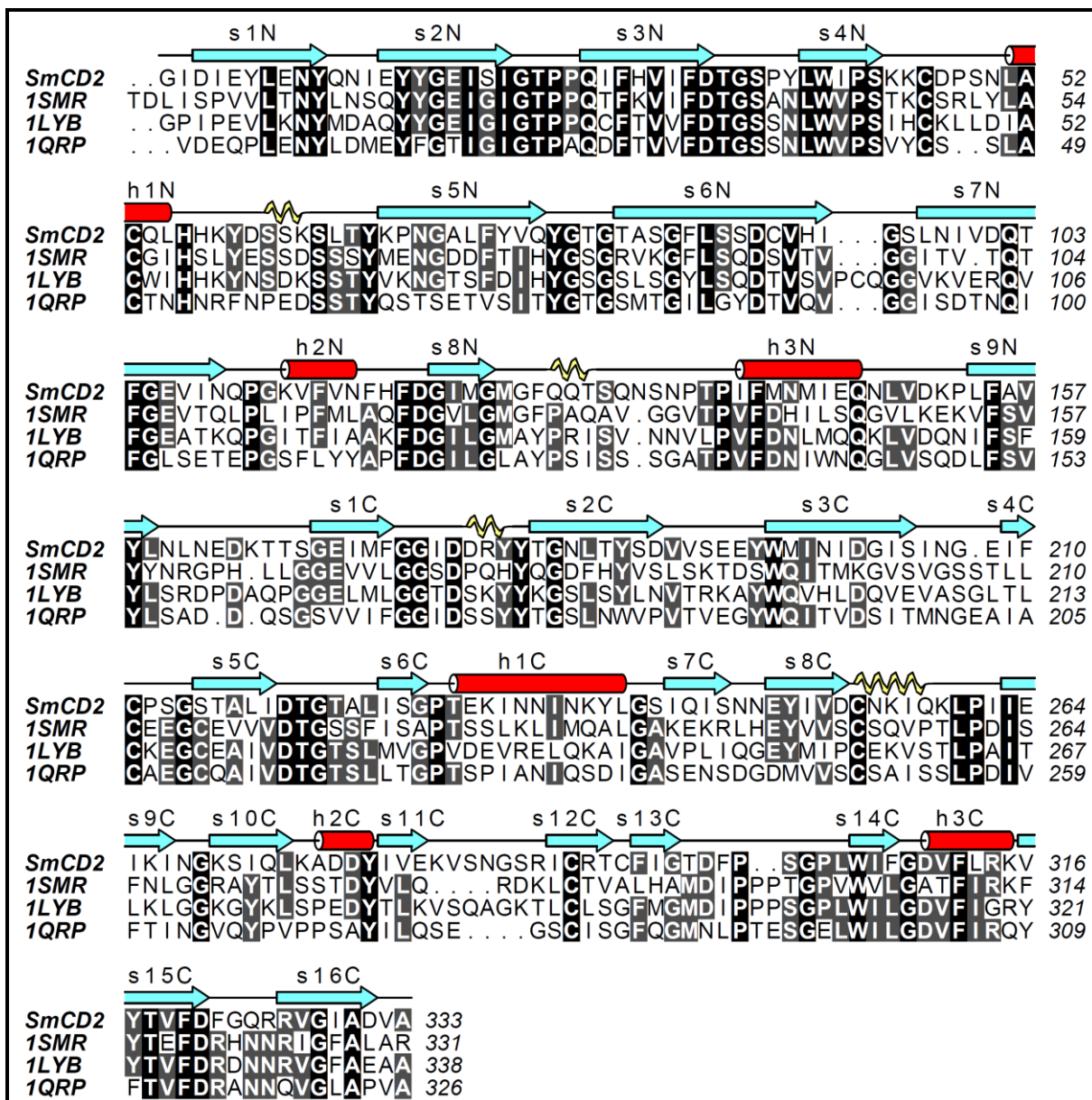
dois foram identificados na SmCD2 (C46-C53, C239-C276). Adicionalmente, dois resíduos de cisteínas (estruturalmente distantes) não participam de ligações dissulfeto no modelo da enzima SmCD2 (C196 and C279) (FIGURAS 4.1, 4.2 e 4.3).

Pela análise do alinhamento, a alça rica em prolina da catepsina D humana corresponde a resíduos básicos na enzima SmCD1, como já descrito por Silva-Jr e colaboradores [93]. Essa região está ausente (apresenta um “gap”) na enzima SmCD2 e, na SmCD3, corresponde a resíduos hidrofóbicos (Leu298, Leu299 e Pro300).

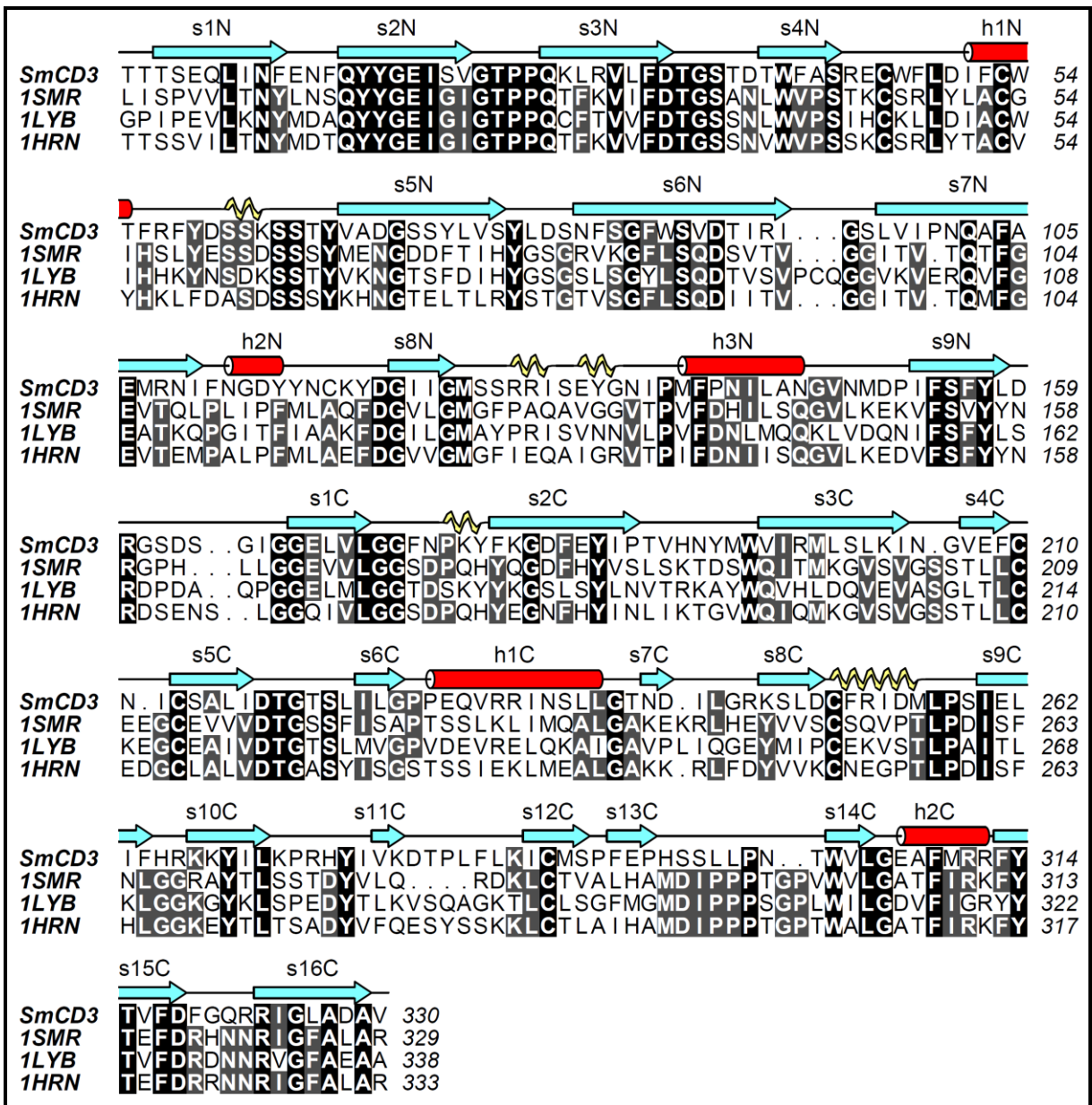
As regiões menos conservadas estão localizadas no N-terminal das enzimas (próximo à fita s1C), nas hélices h2N, h3N e h1C, e nas alças entre as fitas s2C e s3C, s3C e s4C, s10C e s11C (s11C e s12C nas enzimas SmCD2 e SmCD3). A hélice h2N e a alça entre as fitas s10C e s11C da SmCD1 apresentaram diferenças importantes no que diz respeito à sua conformação e interação com os ligantes, como apresentado na seção 3.3.5.



**FIGURA 4.1.** Alinhamento entre as seqüências da enzimas SmCD1 e as proteínas de referência utilizadas na modelagem comparativa. O alinhamento foi construído pelo programa T-Coffee [187]. Os resíduos conservados foram destacados de acordo com a identidade entre as seqüências (100% de identidade em preto e 75% em cinza). As estruturas secundárias dos modelos da SmCD1 foram assinaladas pelo programa DSSP [200]. A figura foi obtida pelo programa ALINE [201].



**FIGURA 4.2.** Alinhamento entre a seqüência da enzima SmCD2 e as proteínas de referência utilizadas na modelagem comparativa. O alinhamento foi construído pelo programa T-Coffee [187]. Os resíduos conservados foram destacados de acordo com a identidade entre as seqüências (100% de identidade em preto e 75% em cinza). As estruturas secundárias do modelo da SmCD2 foram assinaladas pelo programa DSSP [200]. A figura foi obtida pelo programa ALINE [201].



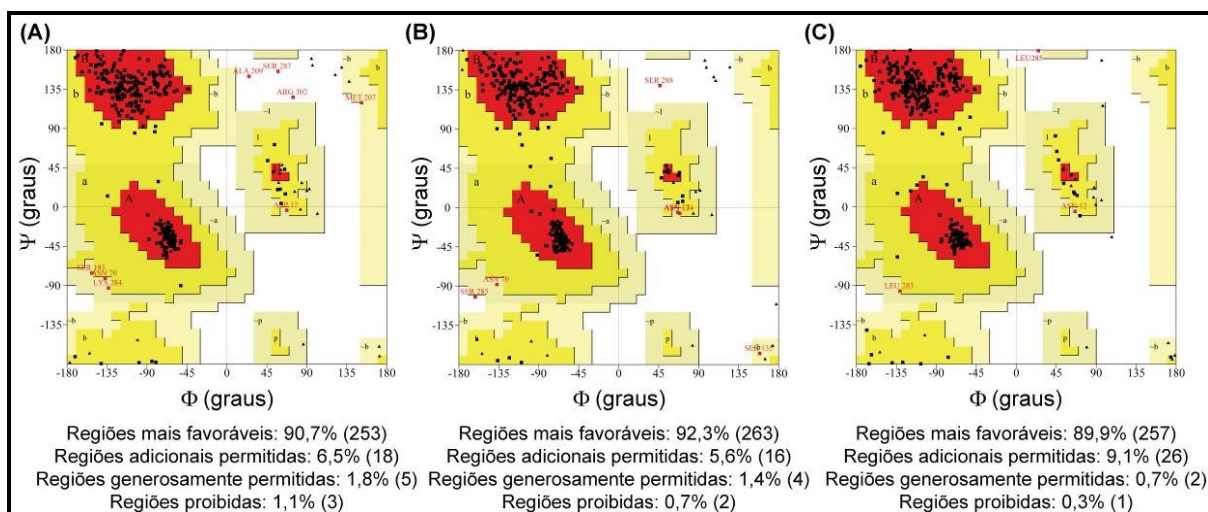
**FIGURA 4.3.** Alinhamento entre a seqüência da enzima SmCD3 e as proteínas de referência utilizadas na modelagem comparativa. O alinhamento foi construído pelo programa T-Coffee [187]. Os resíduos conservados foram destacados de acordo com a identidade entre as seqüências (100% de identidade em preto e 75% em cinza). As estruturas secundárias do modelo da SmCD3 foram assinaladas pelo programa DSSP [200]. A figura foi obtida pelo programa ALINE [201].

#### 4.1.4 CONSTRUÇÃO E VALIDAÇÃO DOS MODELOS

De posse do alinhamento obtido na seção anterior, foram construídos 50 modelos pelo programa Modeller 8v2 [116], dos quais apenas um foi selecionado,

baseando-se na função objetiva e nos resultados de validação obtidos pelos programas mencionados na seção 3.1.5. Em seguida à construção do modelo, foi verificada a qualidade dos mesmos, pois essa avaliação determina a sua utilidade [185].

O programa PROCHECK fornece, dentre outros, o gráfico de Ramachandran que assinala quais resíduos encontram-se nas regiões energeticamente mais favoráveis e desfavoráveis [109]. Pela análise do gráfico de Ramachandran (FIGURA 4.4), o modelo da enzima SmCD1 apresentou 90,7% dos resíduos nas regiões mais favoráveis no gráfico, e apenas três resíduos na região proibida. O modelo da enzima SmCD2 apresentou 92,3% nas regiões mais favoráveis e dois resíduos na região proibida. O modelo da enzima SmCD3 apresentou 89,9% nas regiões mais favoráveis e apenas um resíduo na região proibida. Dentre os resíduos encontrados nas regiões proibidas, observou-se a presença da Ser287 (SmCD1), Ser288 (SmCD2) e Leu285 (SmCD3). Esses resíduos estão localizados na mesma posição de Lys299 na estrutura da catepsina D humana utilizada como referência (código PDB 1LYB) que também se encontra na região proibida do gráfico de Ramachandran. Os demais resíduos encontrados na região não estão localizados no sítio ativo nem próximos ao sítio de ligação dos inibidores.

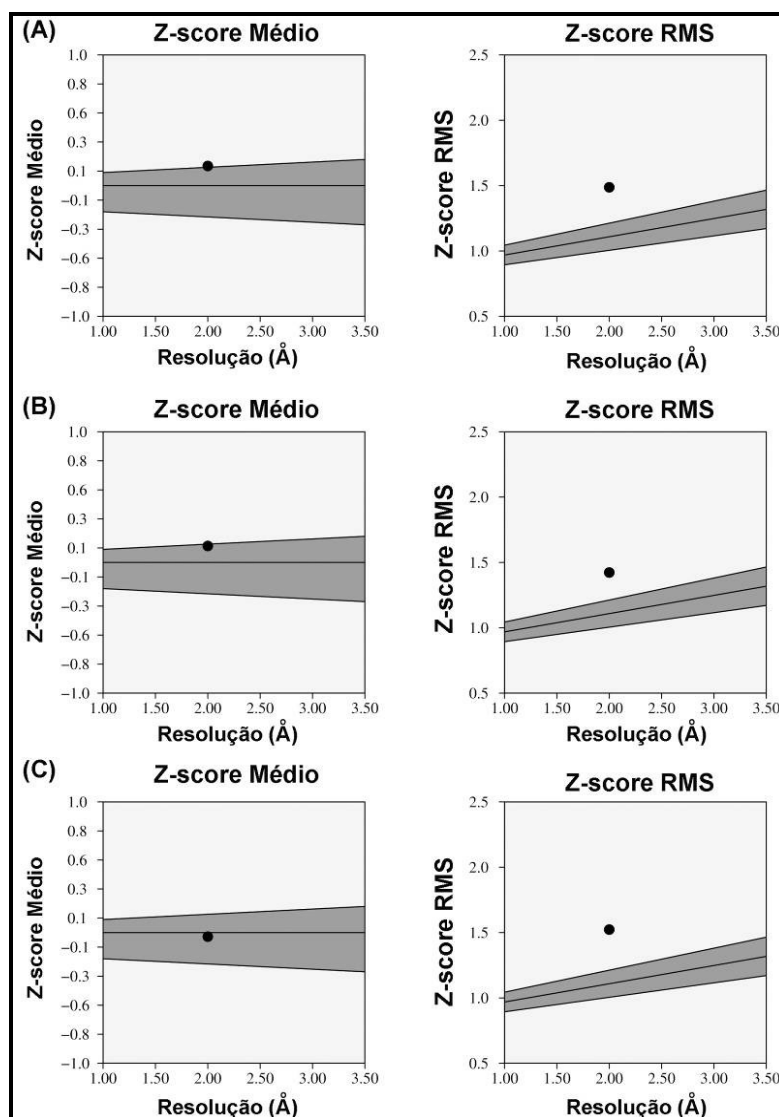


**FIGURA 4.4.** Gráfico de Ramachandran para as enzimas **(A)** SmCD1, **(B)** SmCD2 e **(C)** SMCD3, obtido pelo programa PROCHECK [118].

O programa PROVE (“Protein Volume Evaluation”) calcula o volume dos átomos ao utilizar um algoritmo que trata os átomos como esferas. O desvio entre os volumes obtidos para os átomos do modelo e os valores esperados para cada átomo



é indicado pelo Z-score. Para as enzimas SmCD1, SmCD2 e SmCD3, os valores de Z-score médios foram muito próximos do padrão. Além disso, todos os valores de Z-score RMS (escore que visa avaliar a qualidade global da estrutura) são maiores que 1,2, o que é esperado para estrutura de baixa resolução, como as obtidas por modelagem por homologia (FIGURA 4.5).



**FIGURA 4.5.** Gráfico obtido pelo programa PROVE [120]: Z-score e Z-score RMS versus resolução do arquivo PDB para enzimas (A) SmCD1, (B) SmCD2 e (C) SmCD3.

O programa Verify-3D [121] determina a compatibilidade de um modelo atômico tridimensional com sua seqüência de aminoácidos pelo assinalamento de sua classe estrutural baseado no seu ambiente. Resíduos com valores de escore baixos indicam que a estrutura predita pode estar incorreta [202]. No modelo da

enzima SmCD1, 95,2% dos resíduos apresentam valor de 3D-1D escore médio > 0,2, enquanto que para a SmCD2 este valor é de 98,8%, e na SmCD3 é de 90,3% indicando que, possivelmente, as estruturas 3D estão corretamente preditas.

Além das avaliações descritas acima, verificou-se a compatibilidade do modelo com os dados experimentais disponíveis para proteínas da família à qual as mesmas pertencem. Os modelos das enzimas SmCD1, SmCD2 e SmCD3 apresentaram dois lóbulos semelhantes, com a estrutura do “flap” localizada sobre o sítio ativo. A posição do sítio catalítico (duas tríades catalíticas formadas pelos resíduos Asp-Thr-Gly) foi identificada entre os dois lóbulos localizados no centro das proteínas, bem como a co-planaridade dos grupos carboxilatos do Asp33 e Asp219. Também foram investigadas as posições das pontes dissulfeto conservadas. Desta forma, os modelos utilizados neste trabalho foram considerados confiáveis e úteis para posteriores comparações estruturais com outras enzimas da mesma família.

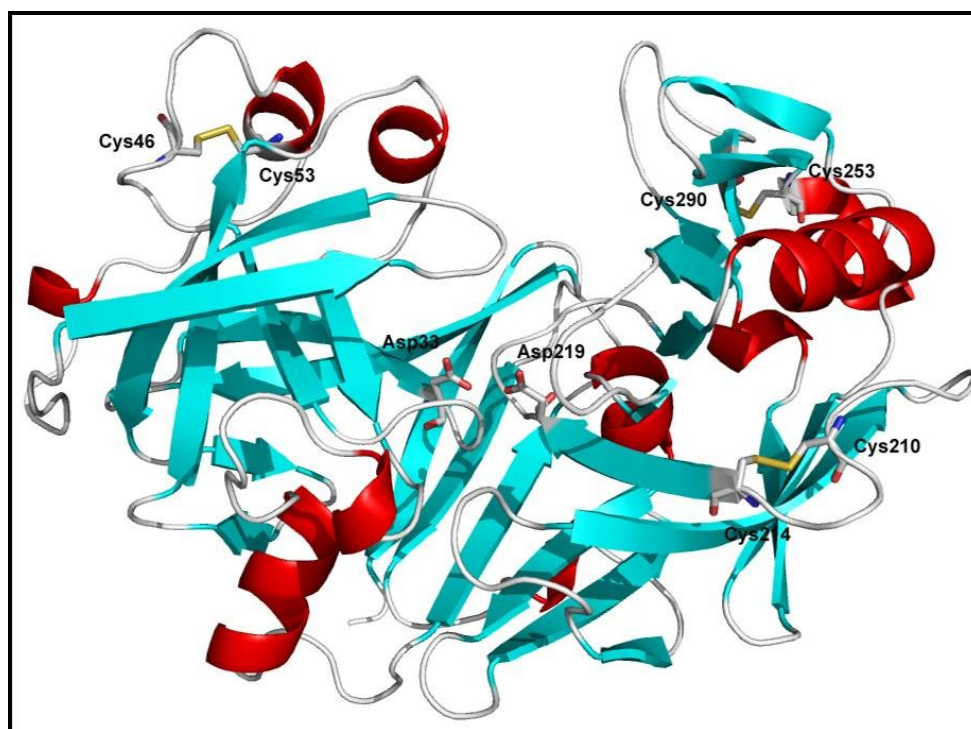
---

#### 4.1.4 ESTRUTURA TRIDIMENSIONAL DOS MODELOS

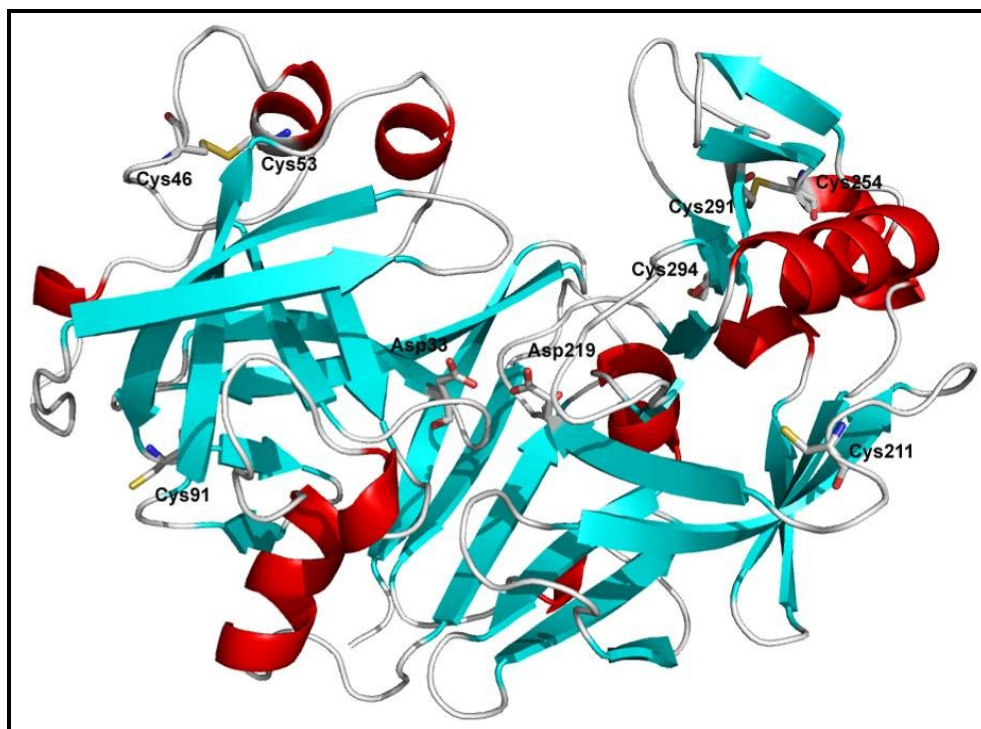
Os três modelos construídos apresentaram estruturas 3D semelhantes a outras aspartil proteases (FIGURAS 4.6, 4.7 e 4.8), pois o enovelamento é altamente conservado na família A1 do banco de dados do MEROPS [34]. SmCD1 e SmCD2 são compostas de seis hélices  $\alpha$  e quatro hélices 3/10. SmCD3 é constituída de cinco hélices  $\alpha$  e cinco hélices 3/10. Também foram encontradas 24 fitas- $\beta$  na enzima SmCD1 e 25 na SmCD2 e SmCD3. A estrutura secundária foi assinalada pelo programa DSSP [200] e é encontrada nas FIGURAS 4.6, 4.7 e 4.8. Assim como todas as aspartil proteases da família, a estrutura é formada por dois lóbulos conectados por seis fitas  $\beta$ , que compõem a base do sítio ativo. O sítio catalítico está localizado em duas alças localizadas no interior da enzima – uma entre as fitas s3N e s4N e outra entre as fitas s4C e s5C, na SmCD1, e entre as fitas s5C e s6C na SmCD2 e SmCD3. Este sítio é formado por três resíduos, a chamada “tríade catalítica”, em cada lóbulo: Asp33 (219), Thr34 (220) and Gly35 (221). As enzimas contêm ainda uma região de fitas- $\beta$  extendidas conhecida como “flap” (resíduos 72 a 87).

Três ligações dissulfeto foram detectadas nos modelos das enzimas SmCD1 e SmCD3, as quais são altamente conservadas na família A1 de aspartil proteases (C46-C53, C210-C214 and C253-C290). Apenas duas ligações dissulfeto foram

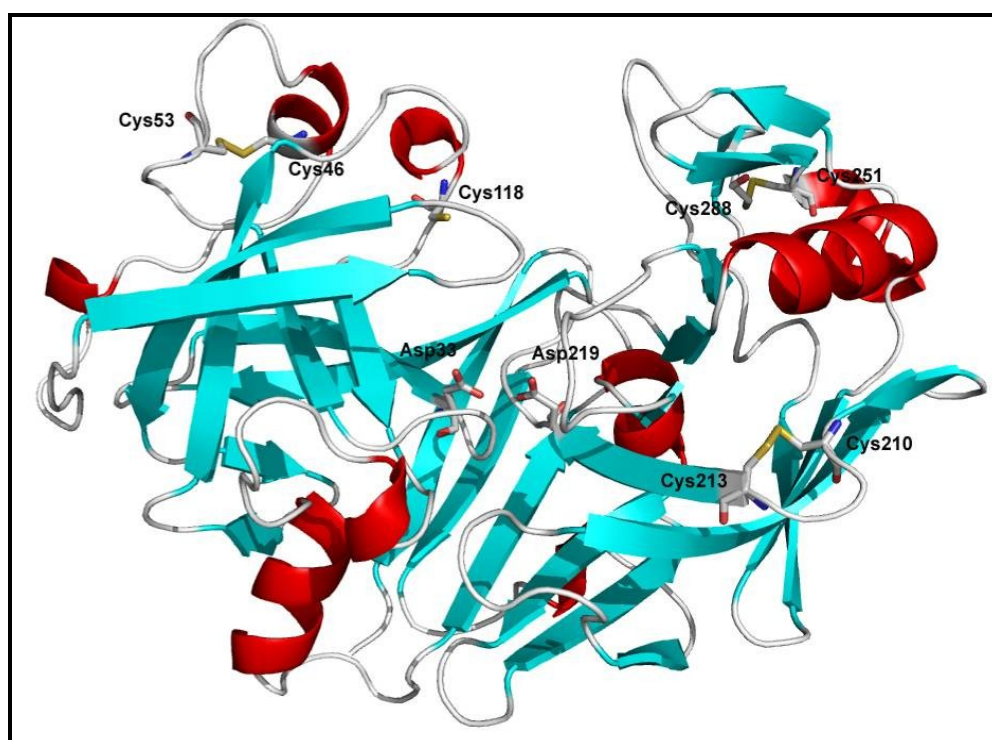
identificadas no modelo de SmCD2 (C46-C53, C239-C276). Contudo, dois resíduos de cisteínas (estruturalmente distantes) não participam de ligações dissulfeto no modelo da enzima SmCD2 (C196 and C279) (FIGURAS 4.6, 4.7 e 4.8).



**FIGURA 4.6.** Estruturas 3D da enzima SmCD1 construída por modelagem comparativa. Ligações dissulfeto, resíduos de cisteína livres e os aspartatos catalíticos (Asp33 e Asp219) estão representados em bastão.



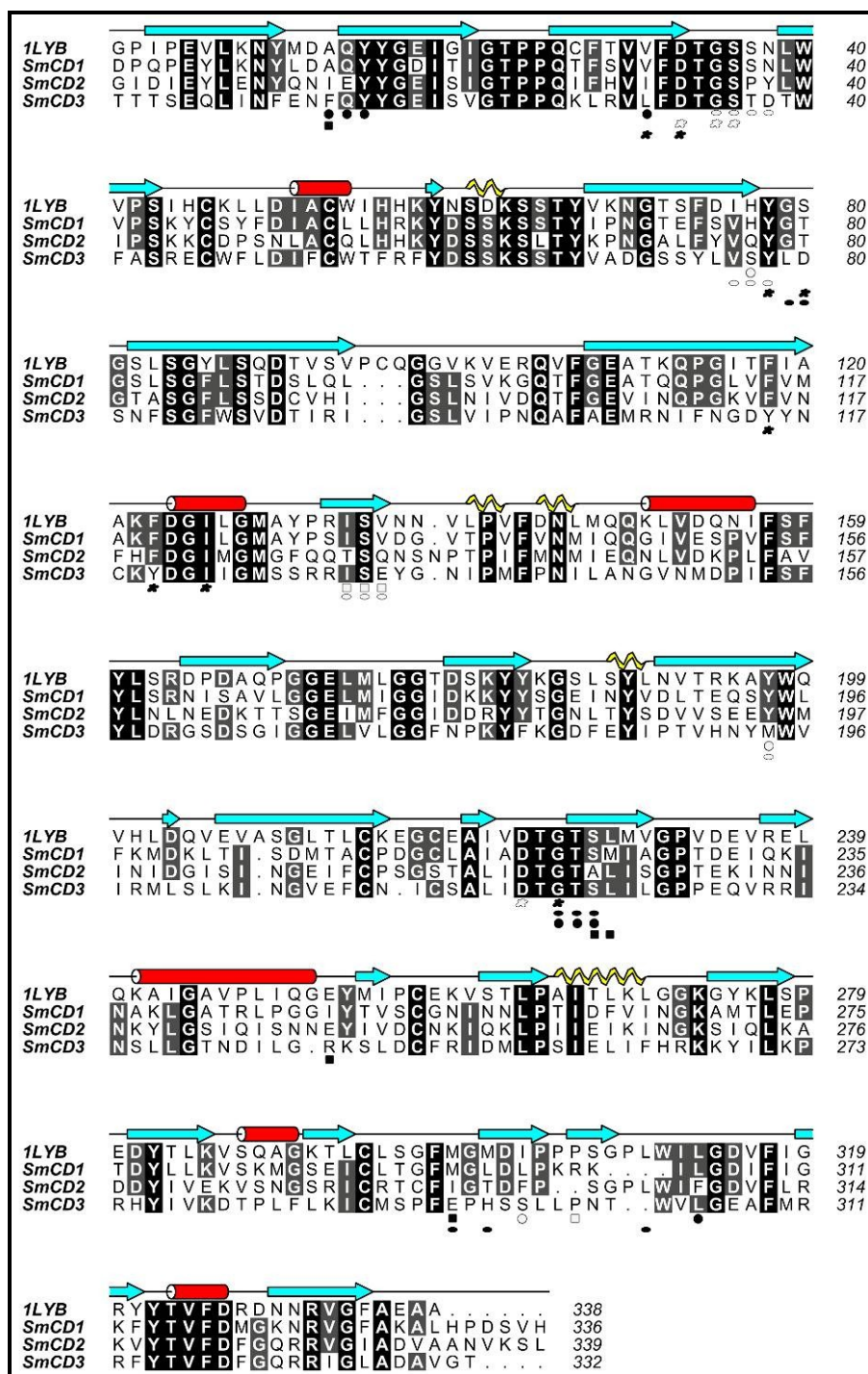
**FIGURA 4.7.** Estruturas 3D da enzima SmCD2 construída por modelagem comparativa. Ligações dissulfeto, resíduos de cisteína livres e os aspartatos catalíticos (Asp33 e Asp219) estão representados em bastão.



**FIGURA 4.8.** Estruturas 3D da enzima SmCD3 construída por modelagem comparativa. Ligações dissulfeto, resíduos de cisteína livres e os aspartatos catalíticos (Asp33 e Asp219) estão representados em bastão.

Em relação aos subsítios de interação com o substrato/inibidor, a enzima SmCD2 apresenta um subsítio S2 ainda menos hidrofóbico que na SmCD1 (FIGURA 4.9), que foi chamado de “raso e anfipático” por Silva-Jr e colaboradores em 2002 [93]. Essa redução da hidrofobicidade de S2 se deve a mutações de Ala226 por Ser227 e Leu297 por Thr297. Esses resíduos correspondem a Leu225 e His295 na enzima SmCD3, a qual também apresenta um subsítio S2 menos hidrofóbico que em SmCD1.

Silva-Jr e colaboradores [93] também descreveram a “alça básica” relacionada à alça rica em prolina da catepsina D humana (resíduos Lys301 a Lys303). Os autores relataram que essa diferença tornaria o subsítio menos capaz de acomodar cadeias laterais maiores. Essa região está ausente na enzima SmCD2, como indicado no alinhamento da FIGURA 4.9, e pode permitir a ligação de substratos/inibidores mais volumosos. Na SmCD3, esta seqüência é mutada para Leu298, Leu299 e Pro300, respectivamente, produzindo um subsítio hidrofóbico.



**FIGURA 4.9.** Alinhamento construído pelo programa T-Coffee [187] entre as seqüências da catepsina D humana e as enzimas SmCD1, SmCD2 e SmCD3. Os resíduos conservados foram destacados de acordo com a identidade entre as seqüências (100% de identidade em preto e 75% em cinza). As estruturas secundárias da catepsina D humana (código PDB 1LYB) foram assinaladas pelo programa DSSP [200]. Os símbolos correspondem aos subsítios: □ S4', ○ S3', ○ S2', ☆ S1', ♣ S1, ● S2, ● S3, ■ S4. A figura foi obtida pelo programa ALINE [201].

## 4.2 ATRACAÇÃO MOLECULAR

Os resultados dos cálculos de atracção molecular, tanto para o protocolo quanto para os ligantes construídos, seguiram um conjunto de parâmetros que foram selecionados a partir de uma faixa de variáveis propostas, conforme descrito na seção 3.2. Na TABELA 4.2 estão descritos os parâmetros selecionados.

Na análise dos resultados do “re-docking” e do “cross-docking”, foram consideradas bem sucedidas as soluções com RMSD menores que 2,0 Å, em relação à estrutura obtida por difração de raios X. Esse valor é usual para considerar estruturas corretamente avaliadas [148].

O elitismo, que é um parâmetro utilizado para preservar os melhores indivíduos de uma geração para a seguinte [146], foi o fator mais impactante nos resultados durante os cálculos de atracção molecular. Ao aumentar de 1% para 10%, a taxa de sucesso do processo (porcentagem de estruturas com RMSD < 2,0 Å) aumentou consideravelmente (dados não mostrados).

**TABELA 4.2.** Valores selecionados das variáveis investigadas no processo de “re-docking” e “cross-docking”.

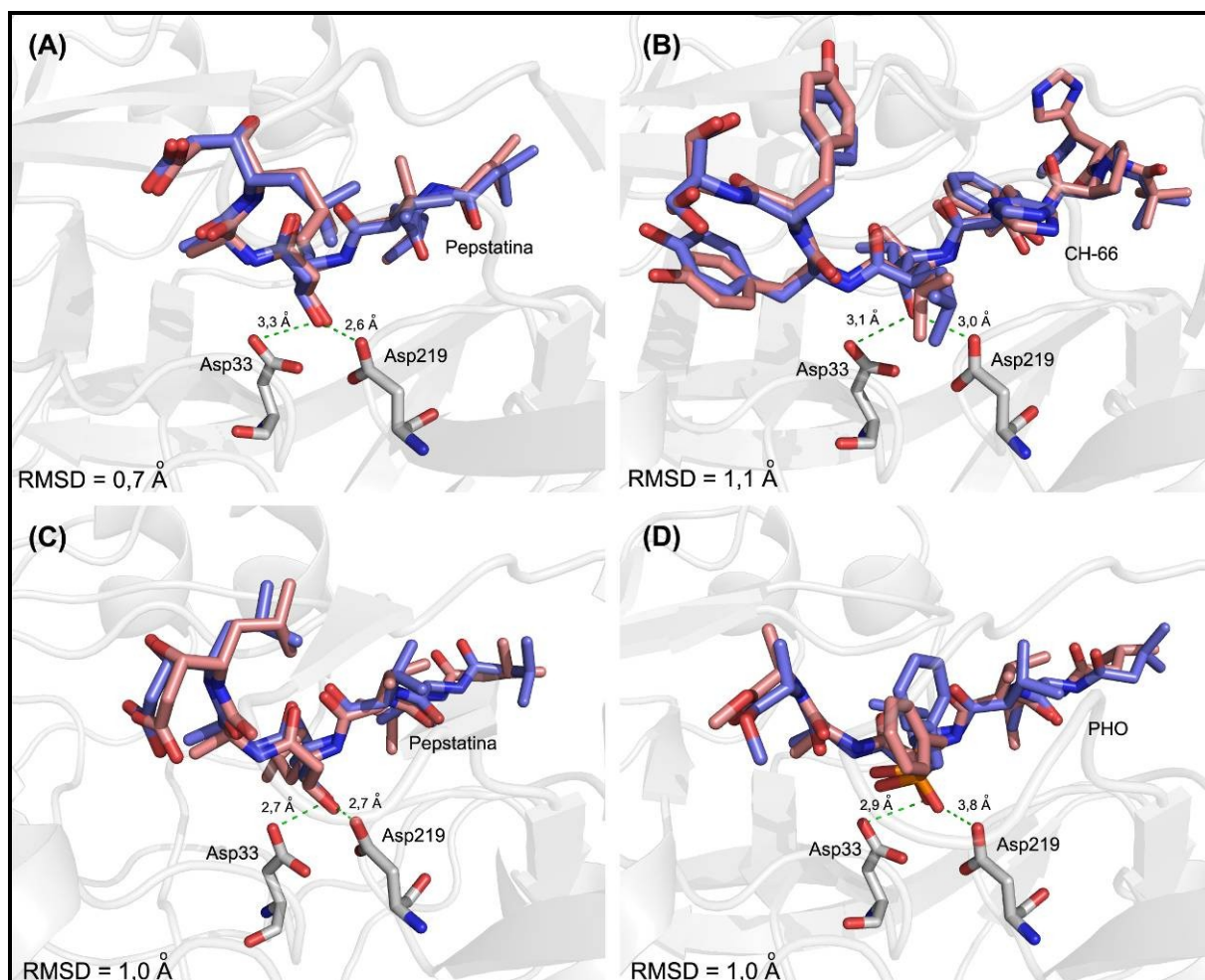
Variáveis	Valores selecionados
População Inicial	1000
Número de Avaliações de Energia	$5 \times 10^6$
Elitismo	10%
Número de Execuções	150
Taxa de Mutação	0,1
Taxa de Crossover	0,6
Algoritmo Genético	LGA

### 4.2.1 “RE-DOCKING” E “CROSS-DOCKING”

As enzimas catepsina D humana complexada à pepstatina e a renina de camundongo complexada ao ligante CH-66 foram usadas para o “re-docking” (código PDB 1LYB e 1SMR, respectivamente). Para o “cross-docking” foram utilizados dois complexos com a pepsina humana (código PDB 1PSO e 1QRP).

A comparação entre as conformações dos ligantes selecionadas (isto é, de menor energia) e aquelas obtidas por difração de raios X, após o experimento de atracção molecular, mostrou a adequada sobreposição dos ligantes no sítio ativo

de suas respectivas enzimas em ambos os protocolos (com RMSD entre as conformações próximas a 1,0 Å) (FIGURA 4.10), sugerindo que os parâmetros utilizados para os cálculos no presente estudo estão ajustados e são convergentes. Estes complexos são estabilizados por várias ligações de hidrogênio entre os átomos da cadeia principal dos inibidores e os átomos da enzima, tanto da cadeia principal quanto das cadeias laterais e que foram mantidas após a atração molecular.



**FIGURA 4.10.** Sobreposição das soluções selecionadas na metodologia de atração molecular e as conformações obtidas pela difração de raios X. **(A)** pepstatina (código PDB 1LYB) complexada à catepsina d humana, **(B)** CH-66 (código PDB 1SMR) complexado à renina de camundongo, **(C)** pepstatina (código PDB 1PSO) ligada à pepsina humana (código PDB 1QRP), e **(D)** ligante fosfatado PHO (código PDB 1QRP) complexado à pepsina humana (código PDB 1PSO). Em destaque, encontram-se os resíduos de aspartato catalítico (Asp33 e Asp219) e as



conformações dos inibidores obtidos por difração de raios X (rosa) e pela atração molecular (azul).

O “re-docking” da pepstatina na catepsina D humana (código PDB 1LYB) resultou em apenas um aglomerado, indicando que o programa AutoDock 4.2 [130] encontrou posições muito semelhantes à conformação do cristal e delas entre si. O “cross-docking” do inibidor fosfatado PHO (código PDB 1QRP) na pepsina humana (código PDB 1PSO) e o “re-docking” do CH-66 na renina de camundongo (código PDB 1SMR) resultaram em dois aglomerados com 147 conformações no aglomerado mais populoso. O “cross-docking” da pepstatina (código PDB 1PSO) na pepsina humana (código PDB 1QRP) resultou na formação de três aglomerados, sendo o mais populoso com 95 conformações (TABELA 4.3). O aglomerado mais populoso apresentou as soluções de menor energia nos dois experimentos de “re-docking” e “cross-docking” do ligante PHO (código PDB 1QRP) na pepsina humana (código PDB 1PSO). Os três inibidores superpuseram-se adequadamente aos respectivos ligantes obtidos por difração de raios X, apresentando  $\text{RMSD} \leq 1,1 \text{ \AA}$  (FIGURA 4.10).

Para o “cross-docking” da pepstatina (código PDB 1PSO) na pepsina humana, a conformação de menor energia não está inserida no aglomerado mais populoso. Esperaria-se que o programa encontrasse um aglomerado com maior número de conformações e que o mesmo fosse o de menor energia. Como esse fato não é verdade para o caso do “cross-docking” da pepstatina na pepsina humana, partiu-se para a análise visual das conformações do ligante.

As conformações do aglomerado mais populoso não se adequam aos parâmetros exigidos. O átomo de oxigênio do resíduo de estatina central da pepstatina se localiza afastado dos aspartatos catalíticos (distâncias maiores que  $3,5 \text{ \AA}$ ), impedindo a interação por ligação de hidrogênio entre os mesmos. Dessa forma, o ligante se desloca da posição inicial (conformações com RMSD maiores que  $2,2 \text{ \AA}$ ), o que afeta negativamente as contribuições de energia intermolecular dos mesmos.

Contudo, a conformação de menor energia apresenta menor RMSD em relação à conformação encontrada pela difração de raios X (RMSD de  $1,0 \text{ \AA}$  contra  $2,2 \text{ \AA}$  da solução do aglomerado mais populoso) e o átomo de oxigênio do resíduo

de estatina central localiza-se numa posição favorável para formar ligações de hidrogênio com os aspartatos catalíticos (Asp33 e Asp219).

No caso do “cross-docking” seria esperado que os ligantes interagissem de forma bastante semelhante na proteína, uma vez que se trata de um mesmo sítio de ligação. No entanto deve-se levar em conta que a flexibilidade do receptor leva a uma adaptação induzida (em inglês, “induced fit”) do mesmo em relação ao ligante e, conseqüentemente, a alterações na estrutura tridimensional de cada complexo final. São esperados resultados melhores para os processos de “re-docking”, nos quais o ligante atracado é o mesmo do complexo original.

Os diversos valores de RMSD, os valores de energia e o número de conformações do aglomerado mais populoso encontram-se na (TABELA 4.3).

**TABELA 4.3.** Resultados do cálculo de “re-docking” e “cross-docking”.

	<b>Catepsina D Humana (1LYB)</b>	<b>Renina de camundongo (1SMR)</b>	<b>Pepsina Humana (1PSO)</b>	<b>Pepsina Humana (1QRP)</b>
<b>Inibidor</b>	Pepstatina	CH-66	PHO	Pepstatina
<b>Código PDB do inibidor</b>	1LYB	1SMR	1QRP	1PSO
<b>Número de Torções</b>	9	19	9	9
<b>Número de Aglomerados</b>	1	2	2	3
<b>Número de conformações do aglomerado de menor energia</b>	150 <sup>c</sup>	147 <sup>c</sup>	147 <sup>c</sup>	48
<b>Número de conformações do aglomerado mais populoso</b>	150	147	147	95
<b>RMSD do composto de menor energia (Å)<sup>a</sup></b>	0,7	1,1	1,0	1,0
<b>Energia livre de ligação Estimada (kcal/mol)<sup>b</sup></b>	-9,5	-13,7	-16,4	-6,9

<sup>a</sup>Em relação à respectiva conformação do ligante encontrado na difração de raios X.

<sup>b</sup>Em relação à conformação de menor energia.

<sup>c</sup>Família de menor energia = Família mais populosa.

#### 4.2.2 A PEPSTATINA E OS LIGANTES DA SÉRIE LBPP

As TABELAS 4.4, 4.5 e 4.6 apresentam os resultados do processo de atracação molecular da pepstatina e dos ligantes da série LBPP nas enzimas SmCD1, SmCD2 e SmCD3, respectivamente. A solução de menor energia foi

selecionada para todos os ligantes, pois resultaram em estruturas com características peculiares estabelecidas (tanto no protocolo discutido na seção 4.2.1, quanto nas SmAPs), como a distância favorável para formar uma ligação de hidrogênio entre a hidroxila central dos inibidores e os aspartatos catalíticos.

**TABELA 4.4.** Resultados dos cálculos de atração molecular da SmCD1 com os inibidores peptidomiméticos estudados.

<b>SmCD1</b>				
<b>Inibidor</b>	<b>Pepstatina</b>	<b>LBPP1</b>	<b>LBPP37</b>	<b>LBPP38</b>
<b>Número de torções</b>	9	20	20	20
<b>Número de aglomerados</b>	2	9	9	9
<b>Número de conformações do aglomerado de menor energia</b>	132 <sup>b</sup>	121 <sup>b</sup>	78 <sup>b</sup>	77 <sup>b</sup>
<b>Número de conformações do aglomerado mais populoso</b>	132	121	78	77
<b>Energia livre de ligação estimada (kcal/mol)<sup>a</sup></b>	-14,9	-8,9	-20,5	-21,0

<sup>a</sup> Conformação de menor energia.

<sup>b</sup> Aglomerado de menor energia = Aglomerado mais populoso

**TABELA 4.5.** Resultados dos cálculos de atração molecular da SmCD2 com os inibidores peptidomiméticos estudados.

<b>SmCD2</b>				
<b>Inibidor</b>	<b>Pepstatina</b>	<b>LBPP1</b>	<b>LBPP37</b>	<b>LBPP38</b>
<b>Número de torções</b>	9	20	20	20
<b>Número de aglomerados</b>	1	4	7	8
<b>Número de conformações do aglomerado de menor energia</b>	150 <sup>b</sup>	22	16	34
<b>Número de conformações do aglomerado mais populoso</b>	150	106	85	75
<b>Energia livre de ligação estimada (kcal/mol)<sup>a</sup></b>	-15,6	-20,0	-11,3	-21,9

<sup>a</sup> Conformação de menor energia.

<sup>b</sup> Aglomerado de menor energia = Aglomerado mais populoso

**TABELA 4.6.** Resultados dos cálculos de atração molecular da SmCD3 com os inibidores peptidomiméticos estudados.

Inibidor	SmCD3			
	Pepstatina	LBPP1	LBPP37	LBPP38
Número de torções	9	20	20	20
Número de aglomerados	3	8	7	10
Número de conformações do aglomerado de menor energia	124 <sup>b</sup>	124 <sup>b</sup>	108 <sup>b</sup>	96 <sup>b</sup>
Número de conformações do aglomerado mais populoso	124	124	108	96
Energia livre de ligação estimada (kcal/mol) <sup>a</sup>	+7,7	-3,3	-4,8	-2,2

<sup>a</sup> Conformação de menor energia.

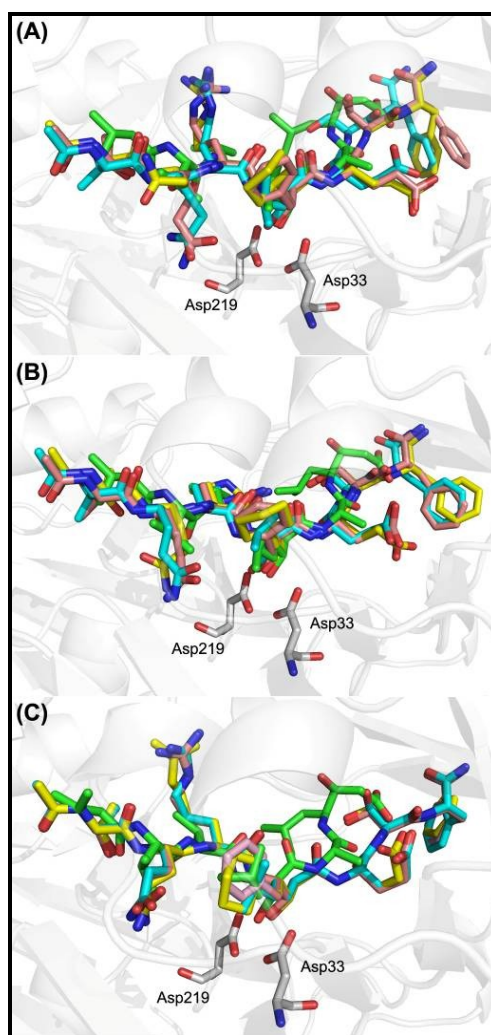
<sup>b</sup> Aglomerado de menor energia = Aglomerado mais populoso

O resultado para o complexo SmCD3/pepstatina gerou uma energia livre estimada positiva, ocasionada por energias intermoleculares muito altas. Além disso, as energias de ligação estimadas para os ligantes da série LBBP em complexo com a enzima SmCD3 também são menos favoráveis que aquelas previstas para a SmCD1 e SmCD2.

Embora, não existam colisões aparentes entre os ligantes e a enzima SmCD3 nos resultados de menor energia, o sítio de ligação dessa enzima mostra-se menor que nas demais enzimas, o que pode ocasionar as altas contribuições das energias intermoleculares. O complexo do aglomerado mais populoso/menor energia apresentou uma estrutura coerente, semelhante àquelas obtidas para SmCD1 e SmCD2 (TABELA 4.6 e FIGURA 4.11).

Apesar do inibidor LBPP1 apresentar como substituinte em P1 um radical isobutil, e os inibidores LBPP37 e LBPP38 apresentar como substituinte um metilciclohexil e um benzil, respectivamente, as conformações selecionadas para os três ligantes, pelos cálculos de atração molecular, são semelhantes entre si (FIGURA 4.11). Como os substituintes de LBPP37 e LBPP38 são radicais mais volumosos se comparados ao isobutil, pode-se resultar em uma mudança de conformação dos ligantes de modo geral. Além disso, todas as conformações encontradas para a série LBBP na SmCD2 e SmCD3 são muito semelhantes entre si.

As TABELAS 1 do Anexo IV listam os subsítios S4 a S4' das enzimas SmCD1, SmCD2 e SmCD3, que acomodam cada resíduo dos ligantes a 4,0 Å.



**FIGURA 4.11.** Comparação entre as conformações selecionadas pela metodologia de atracção molecular para a pepstatina e a série de ligantes LBPP no sítio ativo das enzimas estudadas. **(A)** SmCD1, **(B)** SmCD2 e **(C)** SmCD3. O composto em verde corresponde à pepstatina; em azul, ao ligante LBPP1; em amarelo, LBPP37; e em rosa, LBPP38.

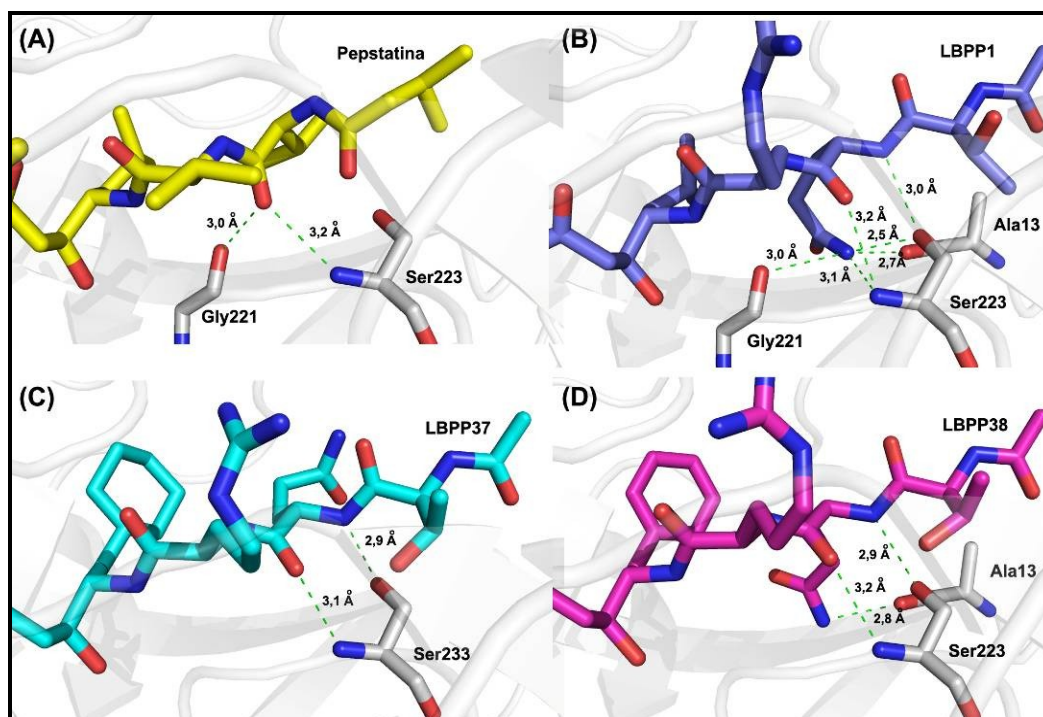
#### 4.2.1 SUBSÍTIOS S1 E S3

Estudos anteriores demonstraram que os sítios de ligação de aspartil proteases formam um canal contínuo e que a comunicação entre os subsítios pode afetar a interação de determinado resíduo em função de resíduos que interajam com a proteína em outras posições [90]. No caso das enzimas humanas pepsina, renina e catepsina D, em que existe comunicação entre os subsítios S1 e S3, há uma

alteração na especificidade da interação em S3 dependendo do resíduo presente na posição P1 [90]. A análise da interação da série de ligantes LBPP com as enzimas SmCD1, SmCD2 e SmCD3 ratificou a comunicação proposta por Fujinaga e colaboradores [90].

Na SmCD1, o átomo de nitrogênio da cadeia lateral do resíduo de glutamina de LBPP1 (P3) interage em S3 com a hidroxila da cadeia lateral do resíduo Ser223 da proteína através de uma ligação de hidrogênio (distância  $N\cdots O$  2,5 Å). LBPP1 é o único ligante da série que mantém esta interação (FIGURA 4.12). O átomo de oxigênio da cadeia lateral de Ser223 produz ainda uma interação por ligação de hidrogênio com o nitrogênio da cadeia principal de uma glutamina também em P3 com os três ligantes: LBPP1 (distância  $O\cdots N$  3,0 Å); LBPP37 (distância  $O\cdots N$  2,9 Å); LBPP38 (distância  $O\cdots N$  2,9 Å), interação esta não presente na pepstatina. Sugere-se que a presença de um substituinte menor (isobutil de LBPP1) em P1 modifique a especificidade em S3. Esse substituinte poderia alterar a conformação do resíduo de glutamina em P3, de forma a favorecer a interação por ligação de hidrogênio com Ser223.

Entretanto, na enzima SmCD3, tanto a cadeia lateral da glutamina de LBPP1 quanto a cadeia lateral de Ser222 assumem posições diferentes daquela observada na SmCD1, o que não admite a formação das ligações de hidrogênio mencionadas para SmCD1. Na enzima SmCD2, estas interações também não ocorrem devido a uma mutação da Ser para Ala224, um resíduo com cadeia lateral sem átomos aceptores e doadores de ligação de hidrogênio.



**FIGURA 4.12.** Comparação entre as diferentes interações dos ligantes estudados e os subsítios S1 e S3 da SmCD1. Complexos: **(A)** SmCD1/pepstatina, **(B)** SmCD1/LBPP1, **(C)** SmCD1/ LBPP37 E **(D)** SmCD1/ LBPP38.

O átomo de nitrogênio da cadeia principal de Ser223 (SmCD1) e Ala224 (SmCD2) produz ainda uma ligação de hidrogênio com o oxigênio da cadeia principal em P3, tanto na série de ligantes LBPP como na pepstatina (TABELA 4.7). Na SmCD3, o resíduo de Ser222 não apresenta essa interação, com exceção da pepstatina (distância  $N \cdots O$  2,8 Å).

Vale ressaltar que o átomo de nitrogênio da Ser223 da SmCD1 e da Ala222 da SmCD2 também interagem com a pepstatina em P4.

**TABELA 4.7.** Interações entre os ligantes e as enzimas estudadas no subsítio S1 e S3, obtidas pelos resultados de atracação molecular.

Enzima	Ligante	Resíduo Ligante	Átomo Ligante	Resíduo Proteína	Átomo Proteína	Distância (Å)			
SmCD1	Pepstatina	Val	O (cadeia principal)	Gly221	O (cadeia principal)	3,0			
				Ser223	N (cadeia principal)	3,2			
	LBPP1	Gln	N (cadeia lateral)	Gly221	O (cadeia principal)	3,0			
					N (cadeia principal)	3,1			
				Ser223	O (cadeia principal)	2,5			
					O (cadeia lateral)	3,2			
				LBPP37	Gln	O (cadeia principal)	Ser223	N (cadeia principal)	3,1
								N (cadeia principal)	O (cadeia lateral)
	LBPP38	Gln	N (cadeia lateral)	Ala13	O (cadeia principal)	2,8			
					O (cadeia principal)	N (cadeia principal)	3,2		
					Ser223	O (cadeia lateral)	2,9		
SmCD2	Pepstatina	Val	O (cadeia principal)	Ala224	N (cadeia principal)	3,2			
	LBPP1	Gln	N (cadeia lateral)	Ile13	O (cadeia principal)	3,0			
			O (cadeia principal)	Ala224	N (cadeia principal)	3,2			
	LBPP37	Gln	N (cadeia lateral)	Ile13	O (cadeia principal)	2,8			
				Gly222	O (cadeia principal)	2,8			
	LBPP38	Gln	O (cadeia principal)	Ala224	N (cadeia principal)	3,4			
					N (cadeia lateral)	Ile13	O (cadeia principal)	2,6	
					O (cadeia principal)	3,2			
				N (cadeia principal)	3,2				



(continuação da TABELA 4.7)

Enzima	Ligante	Resíduo Ligante	Átomo Ligante	Resíduo Proteína	Átomo Proteína	Distância (Å)
SmCD3	Pepstatina	Val	O (cadeia principal)	Ser222	N (cadeia principal)	2,8
	LBPP1	Gln	N (cadeia lateral)	Phe13	O (cadeia principal)	2,6
	LBPP37					2,9
	LBPP38					3,2
Gly220						2,8

Por se tratarem de ligantes maiores com diferentes cadeias laterais, a série LBPP adéqua-se de forma diferente à proteína quando comparados à pepstatina, e desta forma favorece interações que antes não existiam.

Na SmCD1, o átomo de nitrogênio da cadeia lateral e da cadeia principal da glutamina de LBPP1 e LBPP38 interage com Ala13 por ligação de hidrogênio, tratando-se de interações particulares destes ligantes, uma vez que na pepstatina este resíduo apresenta uma cadeia lateral menor e hidrofóbica, uma valina (FIGURA 4.12). Na SmCD2 e SmCD3, a Ala13 é mutada por Ile13 e Phe13, respectivamente. Ambos são resíduos mais volumosos que a alanina, alterando a conformação do ligante. Dessa forma, são mantidas apenas as interações com a cadeia lateral em todos os ligantes da série LBPP.

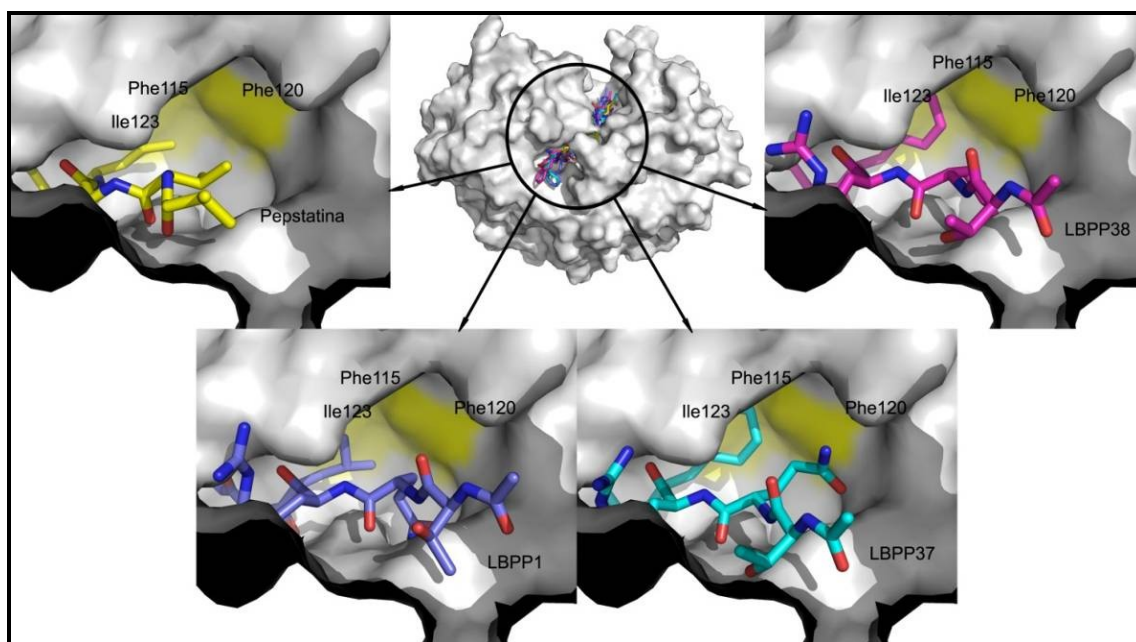
O resíduo de Gly221 da enzima SmCD1 também interage por ligação de hidrogênio com o átomo de nitrogênio da cadeia lateral da glutamina de LBPP1 e LBPP38 (3,0 Å e 3,6 Å respectivamente). Na SmCD2 e SmCD3, esse contato ocorre apenas para os ligantes LBPP37 e LBPP38, respectivamente.

Aspartil proteases humanas como a pepsina, renina e catepsina D preferem resíduos hidrofóbicos em P1. A renina prefere os resíduos hidrofóbicos maiores, enquanto as demais favorecem os resíduos menores, exceto a gastricsina que possui uma especificidade variada, favorecendo fenilalanina sobre leucina, mas não sobre um ciclohexilalanil [90]. Uma fenilalanina na posição 131 da catepsina D humana, comparada a isoleucina em outras enzimas, permite acomodar resíduos menores em P1. No caso da SmCD1 (FIGURA 4.13) e da SmCD2 esta posição é ocupada por uma fenilalanina (Phe120), enquanto que na SmCD3 é por uma tirosina (Tyr120). Estes resíduos favorecem resíduos menores, como aponta os resultados experimentais de inibição (FIGURA 1.8). Estas interações podem contribuir para a

maior atividade da pepstatina e de LBPP1 em relação aos outros ligantes, no extrato aquoso do parasita.

Os resultados obtidos nos cálculos de atracção molecular com as enzimas SmCD1 e SmCD2 demonstraram que todos os ligantes da série LBPP e a pepstatina produzem interações hidrofóbicas em P1, com destaque para os resíduos Phe115, Phe120 e Ile123 das proteínas (FIGURA 4.13). No entanto, devido à presença dos resíduos Tyr115 e Tyr120 apontados para o ligante na SmCD3, esta região se torna ligeiramente mais polar nessa enzima. Dessa forma, a SmCD3 deve favorecer substituintes polares em P1.

O ligante LBPP38 apresenta ainda uma interação *T-stacking* com Tyr78 em todas as enzimas estudadas, sendo esse um resíduo importante para a ligação do substrato/inibidor [88, 90].



**FIGURA 4.13.** Superfície molecular da SmCD1, evidenciando a interação hidrofóbica dos resíduos Phe115, Phe120 e Ile123 com os ligantes estudados.

#### 4.2.2 SUBSÍTIO S2 E S1'

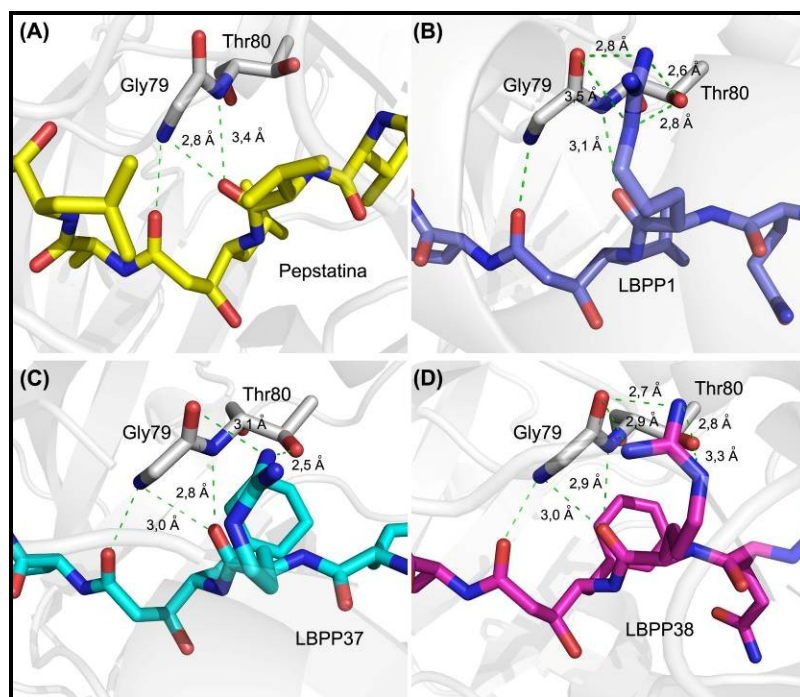
Os três ligantes da série LBPP interagem de forma semelhante neste subsítio nas enzimas SmCD1, sendo que os resíduos das proteínas (Gly79 e Thr80) se localizam próximos aos resíduos de arginina em P2 dos inibidores (FIGURA 4.14). Ocorrem interações entre a cadeia lateral dos resíduos das proteínas e das cadeias

principais dos ligantes. Na SmCD1, todos os ligantes da série LBPP produzem ligações de hidrogênio entre esses resíduos e a cadeia lateral de arginina do inibidor.

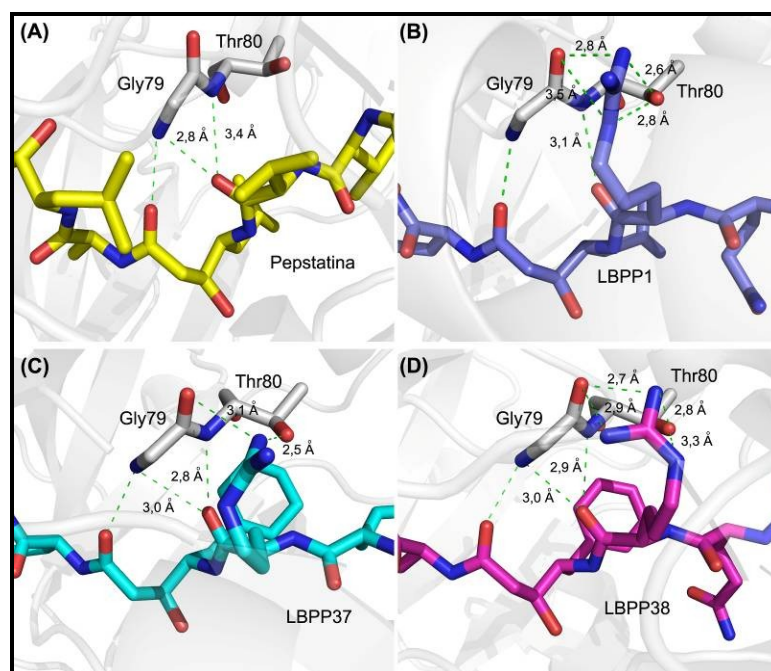
Na SmCD2, a conformação da cadeia lateral dos ligantes da série LBPP é diferente daquela encontrada para a SmCD1. Dessa forma, os resíduos Gly79 e Thr80 interagem com a cadeia principal dos inibidores e o resíduo Thr223 interage com a cadeia lateral de LBPP1 e LBPP37. Devido a essa alteração conformacional, as cadeias laterais dos inibidores na SmCD2 se localizam próximos aos resíduos Phe300 e Ile307, os quais favoreceriam cadeias hidrofóbicas (FIGURA 4.15).

Na SmCD3, os resíduos Gly79 e Thr80 da SmCD1 correspondem a Leu79 e Asp80. Assim como para SmCD1, ambos interagem por ligação de hidrogênio com a cadeia lateral e principal dos ligantes (TABELA 4.8).

Adicionalmente, Gly79 e Thr80 interagem por ligação de hidrogênio no subsítio S1' nas enzimas SmCD1 e SmCD2 (FIGURAS 4.14 e 4.15). Na SmCD3, o resíduo Leu79 provoca um impedimento para a formação da ligação de hidrogênio nesse subsítio. Entretanto, a cadeia lateral de His295 (Leu297 na SmCD1 e Thr298 na SmCD2) forma uma ligação de hidrogênio com a cadeia lateral da arginina na série LBPP (distância N...N 3,0 Å).



**FIGURA 4.14.** Comparação entre as diferentes interações dos ligantes estudados e os subsítios S2 e S1' da SmCD1. Complexos: **(A)** SmCD1/pepstatina, **(B)** SmCD1/LBPP1, **(C)** SmCD1/ LBPP37 e **(D)** SmCD1/LBPP38.



**FIGURA 4.15.** Comparação entre as diferentes interações dos ligantes estudados e os subsítios S2 e S1' da SmCD2. Complexos: **(A)** SmCD2/pepstatina, **(B)** SmCD2/LBPP1, **(C)** SmCD2/ LBPP37 e **(D)** SmCD2/LBPP38.

**TABELA 4.8.** Interações entre os ligantes e as enzimas estudadas no subsítio S2, obtidas pelos resultados de atracação molecular.

Enzima	Ligante	Resíduo Ligante	Átomo Ligante	Resíduo Proteína	Átomo Proteína	Distância (Å)	
SmCD1	Pepstatina	Val	O (cadeia principal)	Gly79	N (cadeia principal)	2,8	
					Thr80	O (cadeia principal)	3,4
	LBPP1	Arg	N (cadeia lateral)	Gly79	O (cadeia principal)	2,8	
					Thr80	O (cadeia lateral)	2,6
	LBPP37	Arg	O (cadeia principal)	Gly79	N (cadeia principal)	3,1	
			N (cadeia lateral)		O (cadeia principal)	3,1	
			O (cadeia principal)	Thr80	N (cadeia principal)	3,0	
			N (cadeia lateral)		O (cadeia lateral)	3,1	
			O (cadeia principal)		N (cadeia principal)	2,8	
			N (cadeia lateral)		O (cadeia principal)	2,7	
	LBPP38	Arg	O (cadeia principal)	Gly79	N (cadeia principal)	3,0	
			N (cadeia lateral)	Thr80	O (cadeia lateral)	2,8	
			O (cadeia principal)		N (cadeia principal)	2,9	
					O (cadeia principal)	2,8	
	SmCD2	Pepstatina	Val	O (cadeia principal)	Gly79	N (cadeia principal)	2,8
						Thr80	N (cadeia principal)
LBPP1		Arg	N (cadeia principal)	Gly79	O (cadeia lateral)	3,2	
			O (cadeia principal)		N (cadeia principal)	3,5	
			O (cadeia principal)	Thr80	N (cadeia principal)	3,0	
			N (cadeia principal)		O (cadeia lateral)	3,2	
			N (cadeia lateral)		Thr223	O (cadeia lateral)	3,0
						O (cadeia lateral)	2,8

(continuação da TABELA 4.8)

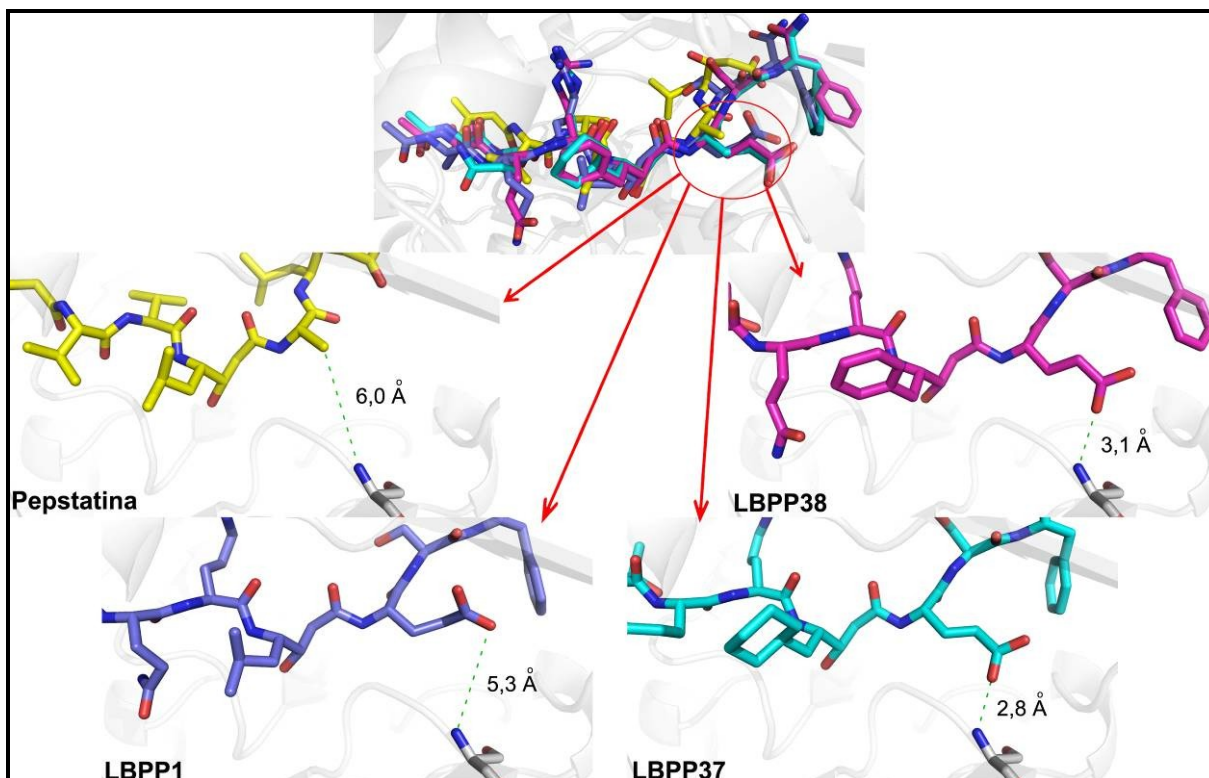
Enzima	Ligante	Resíduo Ligante	Átomo Ligante	Resíduo Proteína	Átomo Proteína	Distância (Å)	
SmCD2	LBPP37	Arg	O (cadeia principal)	Gly79	N (cadeia principal)	2,8	
				Thr80	N (cadeia principal)	2,7	
			N (cadeia lateral)	Thr223	O (cadeia lateral)	2,9	
				LBPP38	Arg	O (cadeia principal)	Gly79
	Thr80	N (cadeia principal)	2,9				
	N (cadeia principal)	O (cadeia lateral)	3,5			3,0	
			SmCD3			Pepstatina	Val
	N (cadeia principal)	Asp80		O (cadeia lateral)	2,2		
O (cadeia principal)	Leu79	N (cadeia principal)		2,6			
LBPP1	Arg	N (cadeia lateral)		Leu79	O (cadeia principal)		
		N (cadeia principal)		Asp80	O (cadeia lateral)	2,5	
				Asp80	O (cadeia lateral)	2,4	
		N (cadeia lateral)		His295	N (cadeia lateral)	3,0	
LBPP37	Arg	O (cadeia principal)		Leu79	N (cadeia principal)	2,6	
		N (cadeia lateral)	Leu79	O (cadeia principal)	2,8		
			Asp80	O (cadeia principal)	3,1		
		N (cadeia principal)	Asp80	O (cadeia lateral)	2,4		
LBPP38	Arg			O (cadeia principal)	Leu79	N (cadeia principal)	2,6
		N (cadeia lateral)	Leu79	O (cadeia principal)	2,6		
			Asp80	O (cadeia lateral)	2,7		
		N (cadeia principal)	Asp80	O (cadeia lateral)	2,4		

Este subsítio é um importante alvo para o desenvolvimento de novos inibidores, pois envolve resíduos que constituem o “flap” das enzimas, o qual é crítico nas interações enzima-ligante [180, 196, 203].

---

#### 4.2.3 SUBSÍTIOS S2' E S3'

O átomo de oxigênio da cadeia lateral do resíduo de glutamato de LBPP37 e LBPP38, em P2', interage com o átomo de nitrogênio presente no resíduo de Ser37 da SmCD1 (distâncias OE1...N 2,8 Å e OE1...N 3,1 Å, respectivamente). Já em LBPP1 esta interação não ocorre. A pepstatina, um inibidor mais potente para aspartil proteases de *S. mansoni* (FIGURA 1.8), possui uma cadeia lateral pouco volumosa em P2' (alanina) e não interage com a proteína em S2' na SmCD3. Nas enzimas SmCD2 e SmCD3, a Ser37 corresponde à Pro37 e Thr37, respectivamente. Esses resíduos não interagem com os ligantes da série LBPP e com a pepstatina em nessas enzimas. Dessa forma, sugere-se que esta interação possa contribuir para a especificidade do ligante em relação a cada enzima em particular, mas não para sua potência em relação ao extrato aquoso do parasita. A posição da cadeia lateral do glutamato dos ligantes na SmCD1 pode ser visualizada espacialmente através de uma análise dos contatos ao redor do resíduo, como mostra a FIGURA 4.16.

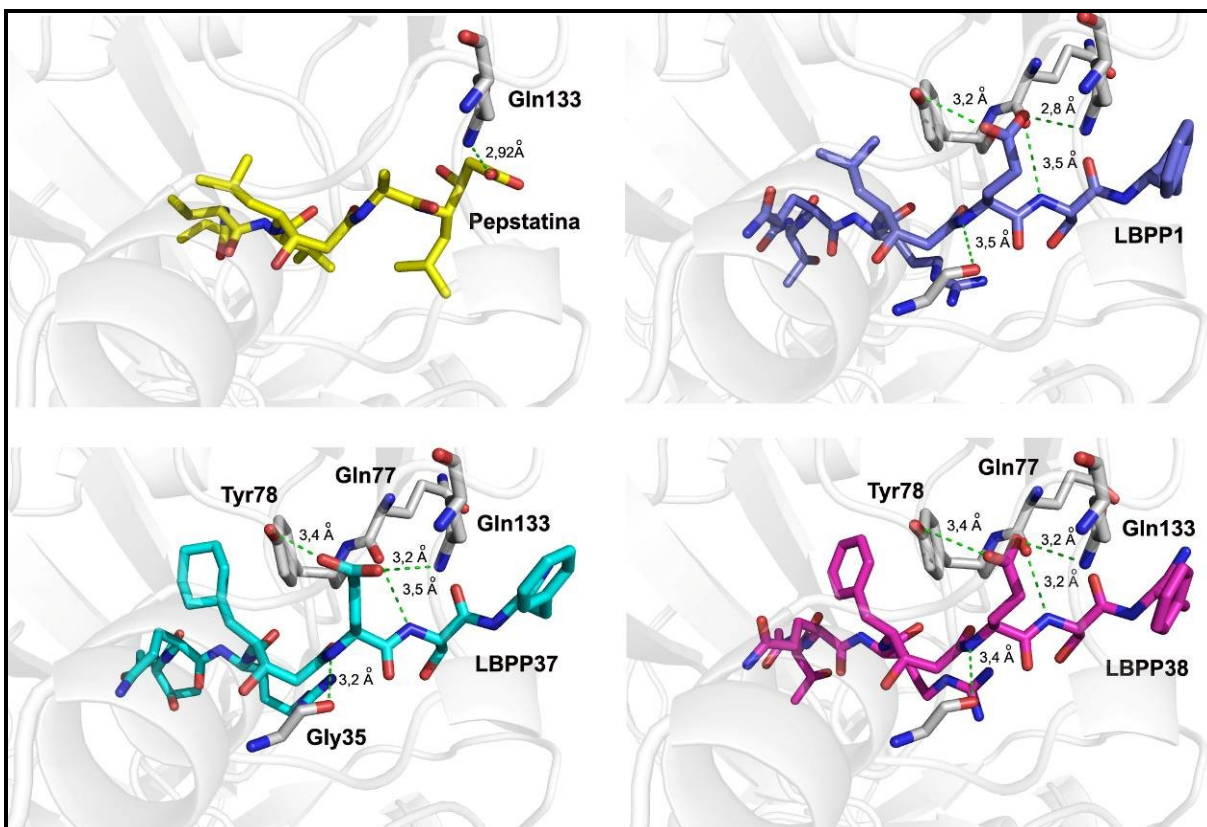


**FIGURA 4.16.** Sobreposição da série de ligantes LBPP e a pepstatina, no subsítio S2' da SmCD1. A interação de LBPP 37 e 38 neste subsítio com a Ser37 e sua distância pronunciada nos casos da pepstatina e LBPP1.

O átomo de nitrogênio da cadeia principal do resíduo de serina em P3' de todos os ligantes da série LBPP, tanto na SmCD1 quanto na SmCD2, interage por ligação de hidrogênio com o resíduo 77 (His77 na SmCD1 e Gln77 na SmCD2). Na SmCD3, essa interação não ocorre (TABELA 4.9). Entretanto, Ser77 interagem com a cadeia lateral do resíduo de Glu de todos os ligantes da série LBPP em P2'.

Outro resíduo apontado como importante para a especificidade da SmCD2 é Gln133. Esse resíduo forma ligações de hidrogênio com todos os ligantes estudados e é específico para SmCD2 (FIGURA 4.17). Na SmCD1, ele corresponde à Val133 e, na SmCD3, à Glu133. Ambos não produzem qualquer tipo de interação com os inibidores em suas respectivas enzimas.





**FIGURA 4.17.** Comparação entre as diferentes interações dos ligantes estudados e os subsítios S2' e S3' da SmCD2. Complexos: **(A)** SmCD2/pepstatina, **(B)** SmCD2/LBPP1, **(C)** SmCD2/ LBPP37 e **(D)** SmCD2/LBPP38.

Ainda em P2', o átomo de nitrogênio da cadeia principal dos três ligantes da série LBPP, em todas as enzimas, forma uma ligação de hidrogênio com Gly35 (TABELA 4.9).

Destaca-se que a pepstatina interage apenas com a SmCD1 em S2' e com a SmCD2 em S3'. Esse ligante não interagem com a SmCD3 nesses subsítios (TABELA 4.9 e TABELA 1 do Anexo IV).

**TABELA 4.9.** Interações entre os ligantes e as enzimas estudadas nos subsítios S2' e S3', obtidas pelos resultados de atracção molecular.

Enzima	Ligante	Resíduo Ligante	Átomo Ligante	Resíduo Proteína	Átomo Proteína	Distância (Å)	
SmCD1	Pepstatina	Ala	O (cadeia principal)	Tyr194	O (cadeia principal)	2,9	
		Ser	N (cadeia principal)	His77	O (cadeia principal)	3,3	
	LBPP1	Glu	N (cadeia principal)	Gly35	O (cadeia principal)	3,3	
			O (cadeia lateral)	His77	O (cadeia principal)	2,7	
		O (cadeia principal)	Tyr194	O (cadeia lateral)	2,9		
	LBPP37	Ser	N (cadeia principal)	His77	O (cadeia principal)	2,7	
			O (cadeia lateral)	His77	O (cadeia principal)	3,2	
		Glu	O (cadeia lateral)	Ser37	N (cadeia principal)	2,8	
			O (cadeia lateral)	Ile131	O (cadeia principal)	2,9	
			O (cadeia principal)	Tyr194	O (cadeia lateral)	2,5	
	LBPP38	Ser	N (cadeia principal)	His77	O (cadeia principal)	2,6	
			O (cadeia lateral)	His77	O (cadeia principal)	3,3	
		Glu	O (cadeia lateral)	Ser37	N (cadeia principal)	3,1	
			O (cadeia lateral)	Ile131	O (cadeia principal)	3,0	
			O (cadeia principal)	Tyr194	O (cadeia lateral)	2,6	
	SmCD2	Pepstatina	Sta	O (cadeia principal)	Gln133	N (cadeia lateral)	2,9
			Ser	N (cadeia principal)	Gln77	O (cadeia principal)	3,5
		LBPP1	Glu	N (cadeia principal)	Gly35	O (cadeia principal)	3,5
O (cadeia lateral)				Tyr78	O (cadeia lateral)	3,2	
O (cadeia lateral)			Gln133	N (cadeia lateral)	2,8		

(continuação da TABELA 4.9)

Enzima	Ligante	Resíduo Ligante	Átomo Ligante	Resíduo Proteína	Átomo Proteína	Distância (Å)
SmCD2	LBPP37	Ser	N (cadeia principal)	Gln77	O (cadeia principal)	3,5
				Gly35	O (cadeia principal)	3,2
		Glu	O (cadeia lateral)	Tyr78	O (cadeia lateral)	2,8
				Gln133	N (cadeia lateral)	3,1
	LBPP38	Ser	N (cadeia principal)	Gln77	O (cadeia principal)	3,2
				Gly35	O (cadeia principal)	3,4
		Glu	O (cadeia lateral)	Tyr78	O (cadeia lateral)	3,4
				Gln133	N (cadeia lateral)	3,2
SmCD3	Pepstatina	-	-	-	-	-
	LBPP1	Glu	N (cadeia principal)	Gly35	O (cadeia principal)	2,8
			O (cadeia lateral)	Ser77	N (cadeia principal)	3,4
	LBPP37	Glu	N (cadeia principal)	Gly35	O (cadeia principal)	2,7
			O (cadeia lateral)	Ser77	N (cadeia principal)	3,4
	LBPP38	Glu	N (cadeia principal)	Gly35	O (cadeia principal)	2,7
			O (cadeia lateral)	Ser77	N (cadeia principal)	3,1
			O (cadeia lateral)	Tyr78	O (cadeia lateral)	3,4

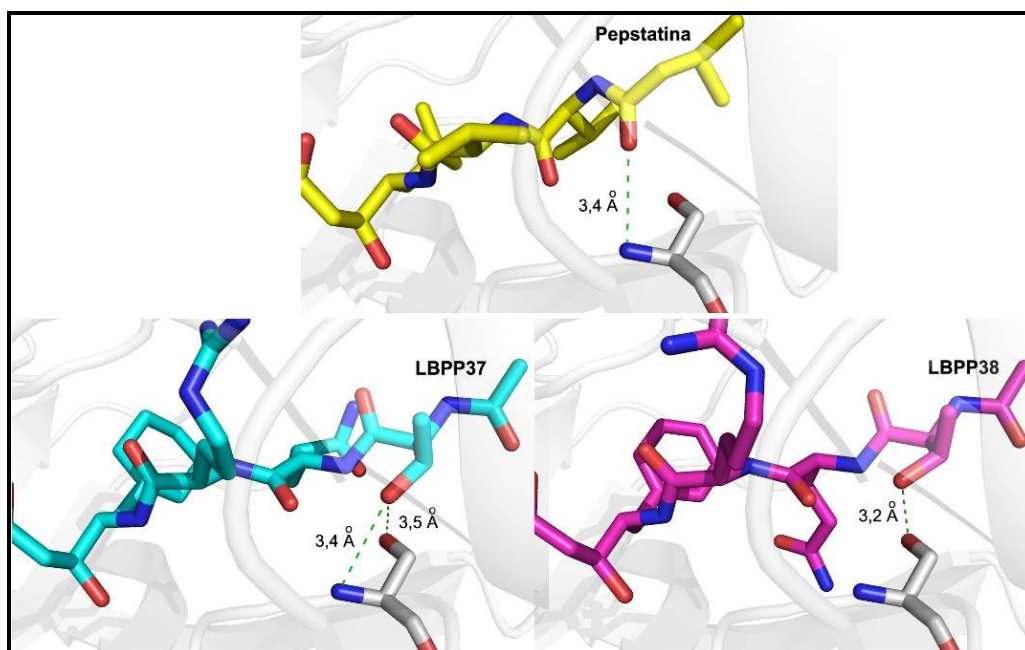
#### 4.2.6 SUB SÍTIO S4

A TABELA 4.10 apresenta as interações entre os ligantes e as enzimas estudadas no subsítio S4, obtidas pelos resultados de atração molecular.

**TABELA 4.10.** Interações entre os ligantes e as enzimas estudadas no subsítio S4, obtidas pelos resultados de atração molecular.

Enzima	Ligante	Resíduo Ligante	Átomo Ligante	Resíduo Proteína	Átomo Proteína	Distância (Å)
SmCD1	Pepstatina	Sta	O (cadeia principal)	Ser223	N (cadeia principal)	3,4
	LBPP1	-	-	-	-	-
	LBPP37	Ser	O (cadeia lateral)	Ser223	N (cadeia principal)	3,4
					O (cadeia lateral)	3,5
LBPP38	Ser	O (cadeia lateral)	Ser223	O (cadeia lateral)	3,2	
SmCD2	Pepstatina	Sta	O (cadeia principal)	Ala224	N (cadeia principal)	3,4
	LBPP1	Ser	O (cadeia lateral)	Arg292	N (cadeia lateral)	3,9
	LBPP37	-	-	-	-	-
	LBPP38	-	-	-	-	-
SmCD3	Pepstatina	Sta	O (cadeia principal)	Ser222	N (cadeia principal)	3,1
	LBPP1	Ser	O (cadeia lateral)	Leu223	N (cadeia principal)	2,5
				Glu293	O (cadeia lateral)	3,5
	LBPP37	Ace	N (cadeia lateral)	Arg246	O (cadeia lateral)	3,2
		Ser	O (cadeia lateral)	Leu223	N (cadeia principal)	2,7
	LBPP38	Ace	N (cadeia lateral)	Arg246	N (cadeia lateral)	3,2
Ser		O (cadeia lateral)	Leu223	N (cadeia principal)	2,5	

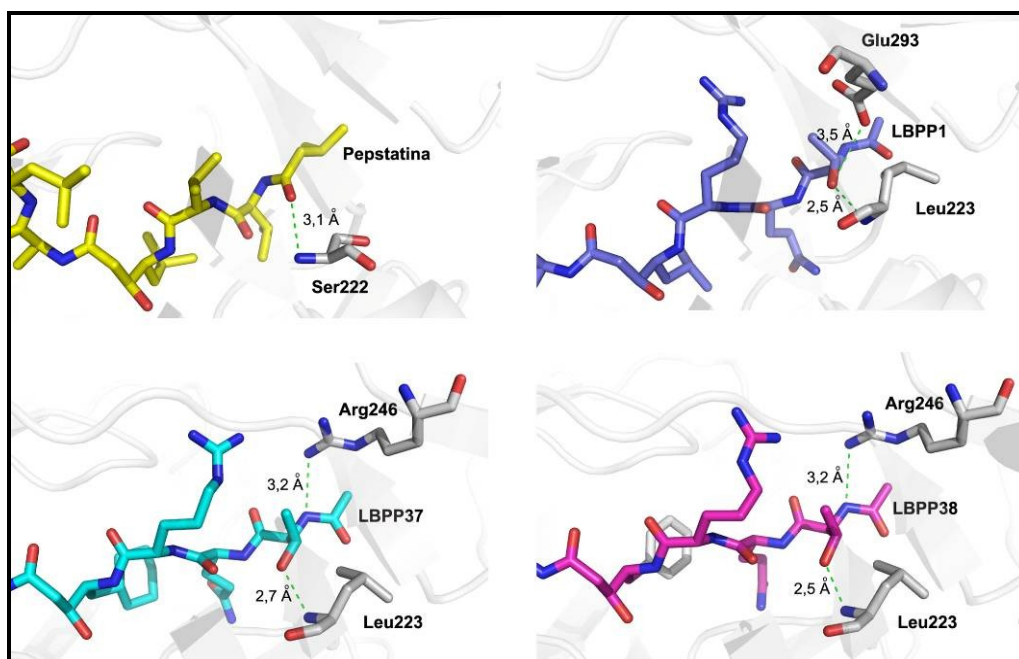
Em P4 apenas o ligante LBPP1 não interage com o resíduo de Ser223 da SmCD1. Para LBPP37 e LBPP38 a hidroxila do resíduo de treonina produz contato com um átomo de nitrogênio da cadeia lateral da serina nas distâncias O...N (3,4 Å e 3,6 Å respectivamente), assim como ocorre na pepstatina O...N 3,4 Å (FIGURA 4.18).



**FIGURA 4.18.** Comparação entre as diferentes interações dos ligantes estudados e o subsítio S4 da SmCD1. Complexos: **(A)** SmCD1/pepstatina, **(B)** SmCD1/LBPP37 e **(C)** SmCD1/LBPP38. O complexo SmCD1/LBPP1 não apresenta interações neste subsítio.

Na SmCD2, LBPP37 e LBPP38 não apresentam interações no subsítio S4. A pepstatina, entretanto, forma uma ligação de hidrogênio com Ala224 em P4 (distância O...N de 3,4 Å). Adicionalmente, no subsítio S4 da SmCD2, ocorre a interação da cadeia lateral de serina de LBPP1 com a cadeia lateral de Arg292. Na SmCD1 e SmCD3 esse resíduo corresponde a Leu291 e Met289, respectivamente, o que impede a interação por ligação de hidrogênio com a cadeia lateral e torna esse subsítio ligeiramente mais apolar que na SmCD2.

Em relação às SmCDs, a SmCD3 é a enzima que apresenta mais contatos com os ligantes. Além das ligações de hidrogênios produzidas entre os ligantes e os resíduos de Ser222, Leu223, Arg246 e Glu293, ocorrem interações hidrofóbicas entre as regiões finais dos ligantes e os resíduos de Phe13, Leu223 e Leu225 (FIGURA 4.19).



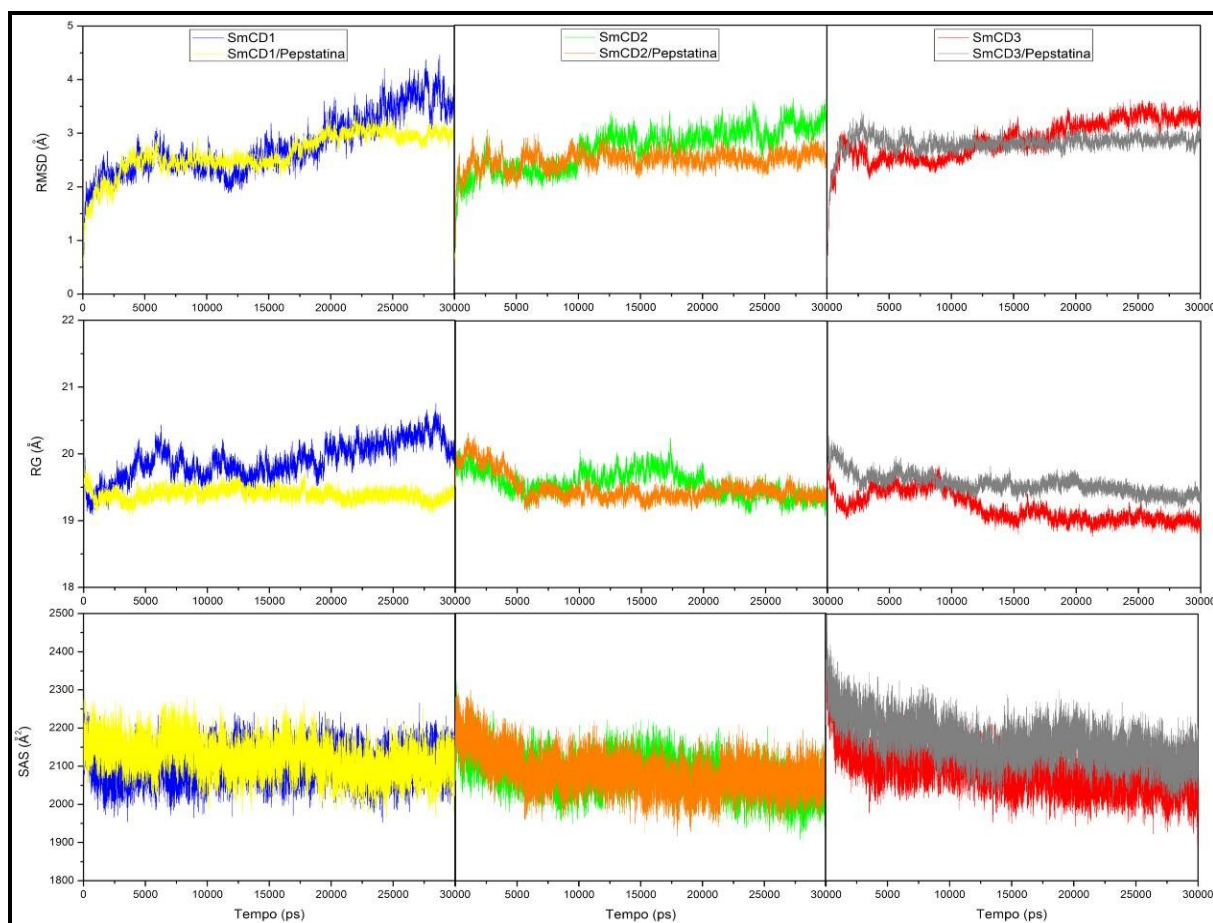
**FIGURA 4.19.** Comparação entre as diferentes interações dos ligantes estudados e o subsítio S4 da SmCD3. Complexos: **(A)** SmCD3/pepstatina, **(B)** SmCD3/LBPP37 e **(C)** SmCD3/LBPP38. O complexo SmCD3/LBPP1 não apresenta interações neste subsítio.

### 4.3 SIMULAÇÃO POR DINÂMICA MOLECULAR

Realizou-se a Simulação por Dinâmica Molecular (DM) nas enzimas livres (SmCD1, SmCD2 e SmCD3) e nos complexos com a pepstatina. Ambos foram obtidos nas etapas de modelagem comparativa (seção 3.1) e de atração molecular (seção 3.2)

#### 4.3.1 ESTABILIDADE DAS TRAJETÓRIAS

A FIGURA 4.20 apresenta os gráficos de desvio quadrático médio da cadeia principal (RMSD, *Root Mean Square Deviation*), raio de giro (RG) e superfície acessível ao solvente (SAS) para os modelos estudados.



**FIGURA 4.20.** RMSD dos átomos da cadeia principal de: SmCD1 livre (azul), SmCD1/pepstatina (amarelo), SmCD2 livre (verde), SmCD2/pepstatina (laranja), SmCD3 livre (vermelho) e SmCD3/pepstatina (cinza). O raio de giro e a superfície acessível ao solvente também estão representadas pela mesma representação de cores.

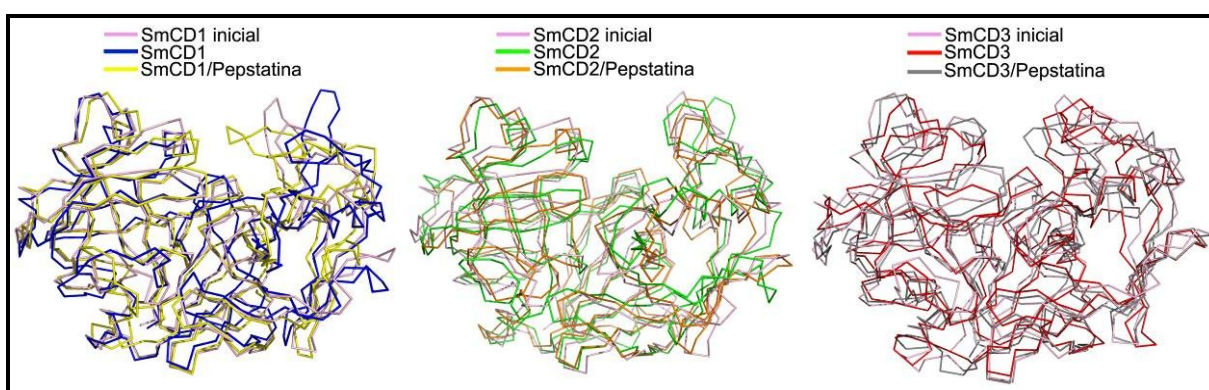
O desvio quadrático médio da cadeia principal (RMSD, *Root Mean Square Deviation*), em relação à estrutura inicial, aumenta rapidamente até estabilizar em 2 ns de simulação, para todas as enzimas e complexos. Contudo, a SmCD1 livre se mantém estável por mais 18 ns e apresenta elevado RMSD após 20 ns de simulação.

Deve-se ressaltar que os estados iniciais das enzimas livres correspondem às formas inibidas dessas enzimas, pois os modelos foram gerados a partir de estruturas de referência ligadas a inibidores (seção 3.1). Dessa forma, esperaria-se grandes alterações nas enzimas livres, pois as estruturas devem evoluir de um estado ligado para um não-ligado. O sítio ativo de aspartil proteases é coberto pelo “flap”, que em eucariotos se compõe por uma estrutura de grampo  $\beta$  que se localiza

perpendicular ao centro do sítio ativo [203]. Durante o ciclo catalítico, o “flap” deveria abrir para permitir a entrada do substrato/inibidor no sítio de ligação e a saída dos produtos de hidrólise. Entretanto, segundo os dados de estruturas cristalográficas [39, 88, 90, 204-207], o “flap” de aspartil proteases de eucariotos não sofrem grandes movimentos para a entrada do substrato/inibidor (cerca de 2,0 Å) [203, 208]. Isso também pode ser observado para as enzimas estudadas de *S. mansoni*, pois não há grandes alterações conformacionais na região do “flap” (FIGURAS 4.21, 4.22 e 4.23).

Entretanto, a sobreposição entre a estrutura inicial (obtida na seção 3.1) e as estruturas médias de cada enzima na sua forma livre e na forma ligada à pepstatina revelam que, no caso da SmCD1, há uma importante diferença na alça que conecta as fitas s10C e s11C (FIGURA 4.21). Quando a enzima está ligada ao inibidor, essa alça se comporta como se estivesse “fechada”. Quando a enzima está livre, a alça estaria “aberta” e poderia auxiliar na entrada do substrato/inibidor. A estrutura inicial parece uma estrutura intermediária entre as duas conformações (livre e ligada). Como o estado “fechado” interage com o inibidor (ver discussão adiante na seção 3.3.2 e 3.3.5), ele se torna menos flexível que o estado “aberto”.

Este efeito também é demonstrado para SmCD2 e SmCD3 livres, mas em menor intensidade. Nota-se que a sobreposição das estruturas iniciais e das estruturas médias das enzimas livres e ligadas não se diferem tão significativamente quanto na SmCD1 (FIGURA 4.21).



**FIGURA 4.21.** Sobreposição entre as estruturas iniciais (em rosa) e as estruturas médias de: SmCD1 livre (azul), SmCD1/pepstatina (amarelo), SmCD2 livre (verde), SmCD2/pepstatina (laranja), SmCD3 livre (vermelho) e SmCD3/pepstatina (cinza). As estruturas médias foram calculadas pela sobreposição dos últimos 20 ns de cada trajetória utilizando-se os átomos da cadeia principal das enzimas.



O raio de giro (RG) e a superfície acessível ao solvente (SAS) não apresentam alterações significativas durante as simulações. Contudo, pode-se observar que os complexos são relativamente mais compactos que as enzimas livres, com exceção da SmCD3. A pepstatina reduz a flexibilidade de SmCD1 e SmCD2, mas este efeito não é aparente para SmCD3 (FIGURA 4.23).

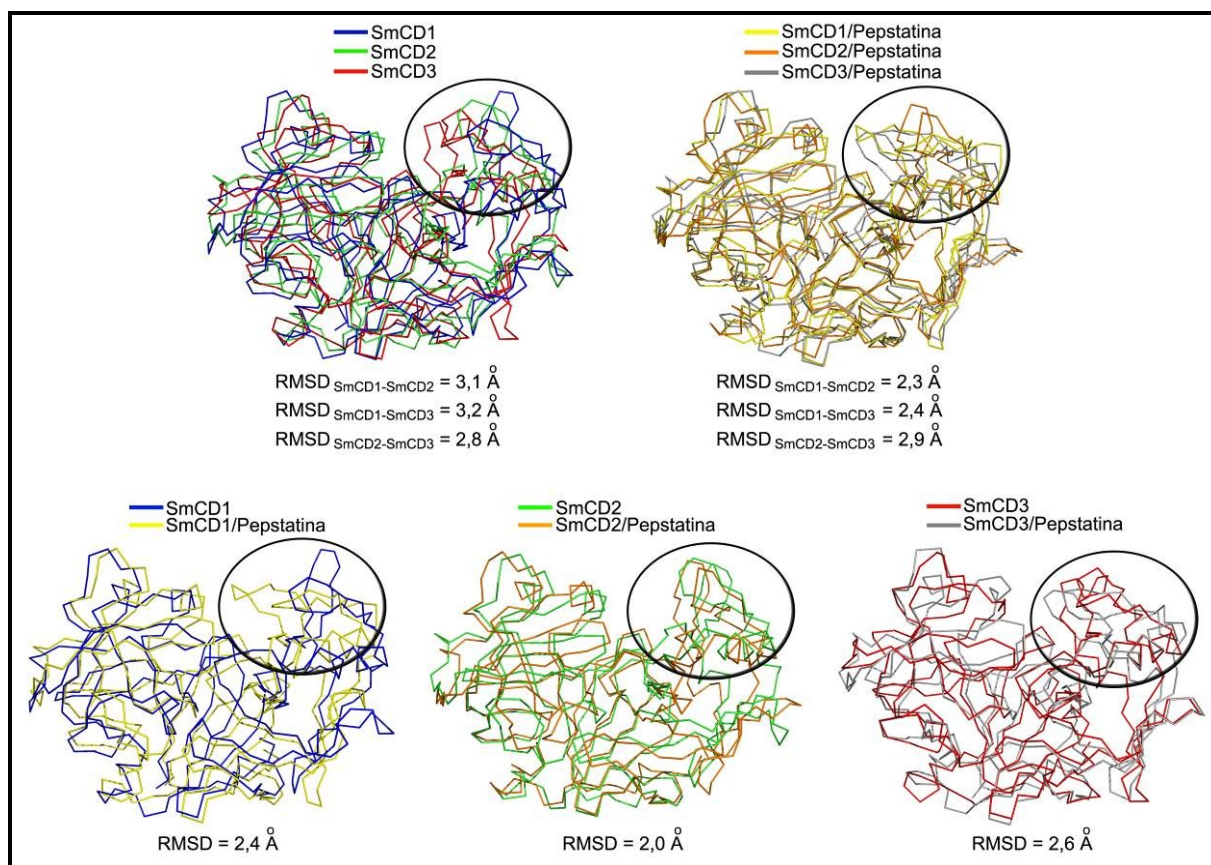
---

#### 4.3.2 ANÁLISE ESTRUTURAL

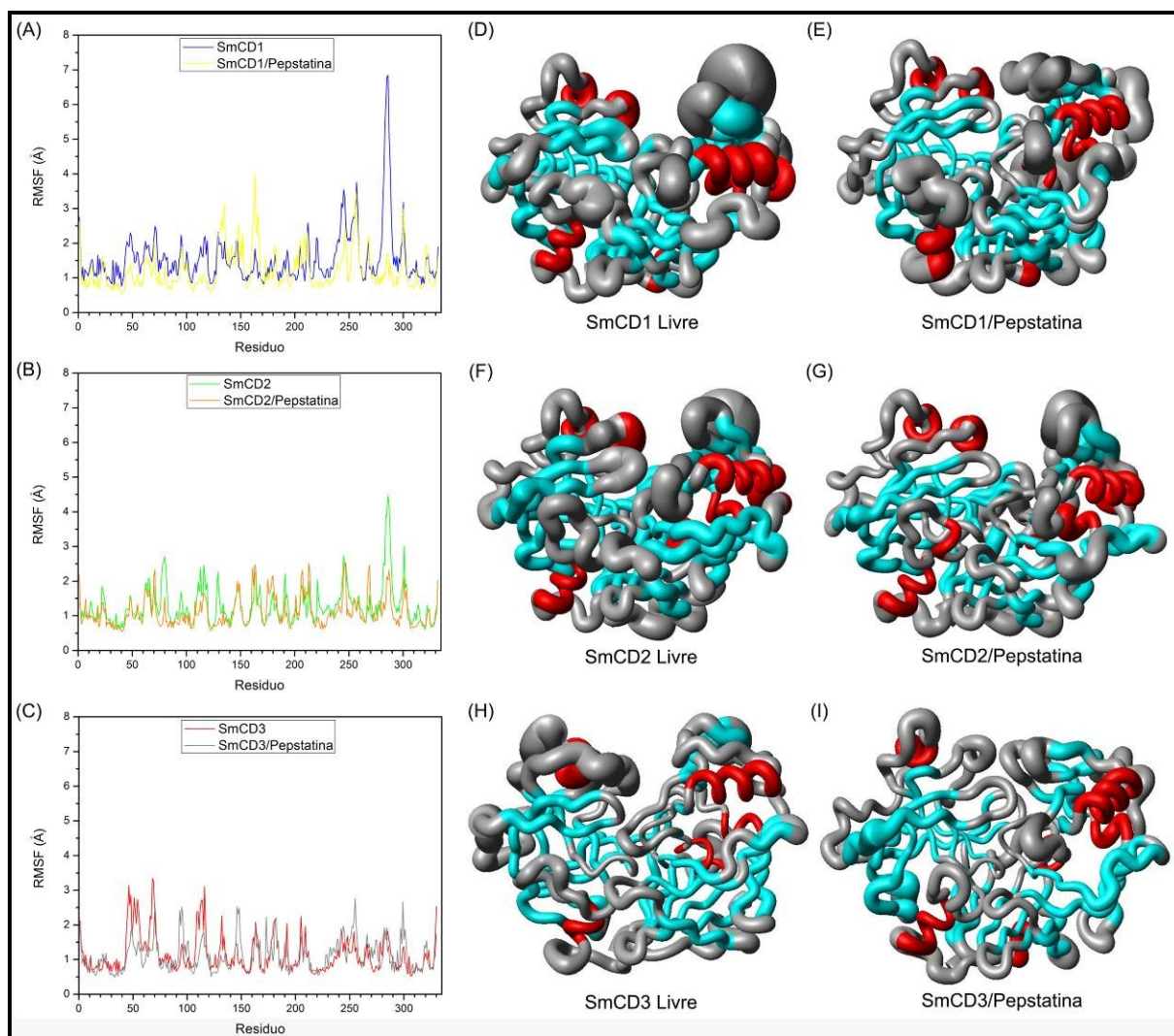
A FIGURA 4.22 apresenta a sobreposição das estruturas médias de cada enzima na sua forma livre sobre a forma ligada à pepstatina, além da sobreposição das três enzimas livres e das três enzimas ligadas à pepstatina. As estruturas médias foram consideradas pela sobreposição dos últimos 20 ns de cada trajetória utilizando-se os átomos da cadeia principal das enzimas. Os resultados indicam diferenças estruturais significativas entre as três enzimas livres ( $\text{RMSD} > 2,5 \text{ \AA}$ ). As diferenças estão localizadas principalmente nas alças, especialmente na alça que conecta as fitas s10C e s11C, as quais estão localizadas na região mais flexível das proteínas. Os RMSDs das formas ligadas à pepstatina em relação às enzimas livres foram calculados sobre todos os resíduos da cadeia principal, indicando que o complexo SmCD2/pepstatina é mais semelhante à estrutura livre da SmCD2 ( $\text{RMSD} = 2,0 \text{ \AA}$ ) que as estruturas de SmCD1/pepstatina e SmCD3/pepstatina o são de suas respectivas estruturas livres ( $\text{RMSD}$  of  $2,4 \text{ \AA}$  e  $2,6 \text{ \AA}$ , respectivamente). Em outras palavras, SmCD2 é a enzima cuja estrutura é menos perturbada pela ligação ao inibidor pepstatina. Além disso, as formas ligadas à pepstatina de SmCD1 e SmCD2 estão mais estruturalmente correlacionadas entre si do que com a SmCD3, enquanto que em suas formas livres, SmCD2 e SmCD3 apresentam melhor correlação estrutural (FIGURA 4.22).

Os valores de RMSD fornecem uma análise coletiva de todos os resíduos das proteínas, promovendo uma perspectiva global das modificações conformacionais ao longo da trajetória da DM. Assim, para análises específicas do comportamento conformacional das enzimas simuladas, empregou-se o método que descreve a flutuação estrutural de uma proteína como função do tempo e do número de cada resíduo - a análise da flutuação do desvio quadrático médio (RMSF, do inglês "Root Mean Square Fluctuation") [209]. Conforme exposto acima, a estrutura da enzima

SmCD1 sofre considerável perturbação com a ligação da pepstatina. As alterações conformacionais residem, sobretudo, na alça que conecta as fitas s10C e s11C (resíduos de Val282 a Ile289) no subsítio S4 da enzima SmCD1 (FIGURAS 4.22 e 4.23). Além disso, as enzimas SmCD1 e SmCD2 livres mostram o sítio ativo mais exposto ao solvente que a SmCD3 livre, devido ao rearranjo de duas alças (entre as folhas s7N e s8N e entre as folhas s10C e s11C) em direção à superfície. Detalhes sobre estes movimentos são discutidos na seção 3.3.5.



**FIGURA 4.22.** Sobreposição das estruturas médias de: SmCD1 livre (azul), SmCD1/pepstatina (amarelo), SmCD2 livre (verde), SmCD2/pepstatina (laranja), SmCD3 livre (vermelho) e SmCD3/pepstatina (cinza). As elipses destacam as diferenças entre as alças que conectam as fitas s11C e s12N das enzimas livres e ligadas. As estruturas médias foram calculadas pela sobreposição dos últimos 20ns de cada trajetória utilizando-se os átomos da cadeia principal das enzimas.



**FIGURA 4.23.** Gráficos de RMSF e em salsicha. **(A)** SmCD1 livre (azul) e em complexo com pepstatina (amarelo), **(B)** SmCD2 livre (verde) e em complexo com pepstatina (laranja), **(C)** SmCD3 livre (vermelho) e em complexo com pepstatina (cinza); gráfico em salsicha para **(D)** SmCD1 livre, **(E)** complexo SmCD1/pepstatina, **(F)** SmCD2 livre, **(G)** complexo SmCD2/pepstatina, **(H)** SmCD3 livre e **(I)** complexo SmCD3/pepstatina. Alças estão coloridas em cinza, folhas  $\beta$  em ciano e hélices  $\alpha$  em vermelho. As estruturas foram calculadas pela sobreposição dos últimos 20 ns de cada trajetória utilizando-se os átomos da cadeia principal das enzimas. Gráfico em salsicha realizado pelo programa MolMol [210].

### 4.3.3 O PAPEL DAS PONTES SALINAS NA ESTABILIDADE DOS COMPLEXOS

Devido à flexibilidade das proteínas, os resíduos carregados se deslocam um em relação ao outro em diferentes conformações da proteína. Conseqüentemente, espera-se que a flexibilidade geral da proteína afetar a força eletrostática entre os resíduos carregados [211].

Foram analisadas as distancias entre os pares iônicos nas enzimas livres e ligadas à pepstatina, durante todo período de simulação. Todas as pontes salinas estão descritas na TABELAS 1 a 3 do Anexo V.

A proporção dos resíduos carregados (Asp, Glu, His, Lys e Arg) é semelhante em todas as enzimas, embora seja ligeiramente maior na SmCD3 (20,4% contra 19,2% na SmCD2 e 17,4% na SmCD1). Consistente com o aumento da ocorrência de resíduos carregados está a formação de pontes salinas. SmCD3 apresenta 13 interações com ocorrência maior que 50% ao longo da trajetória enquanto SmCD2 apresenta 14 interações e SmCD1 apenas 10, as quais podem estar associadas à maior compactação da SmCD3 e SmCD2 em relação à SmCD1.

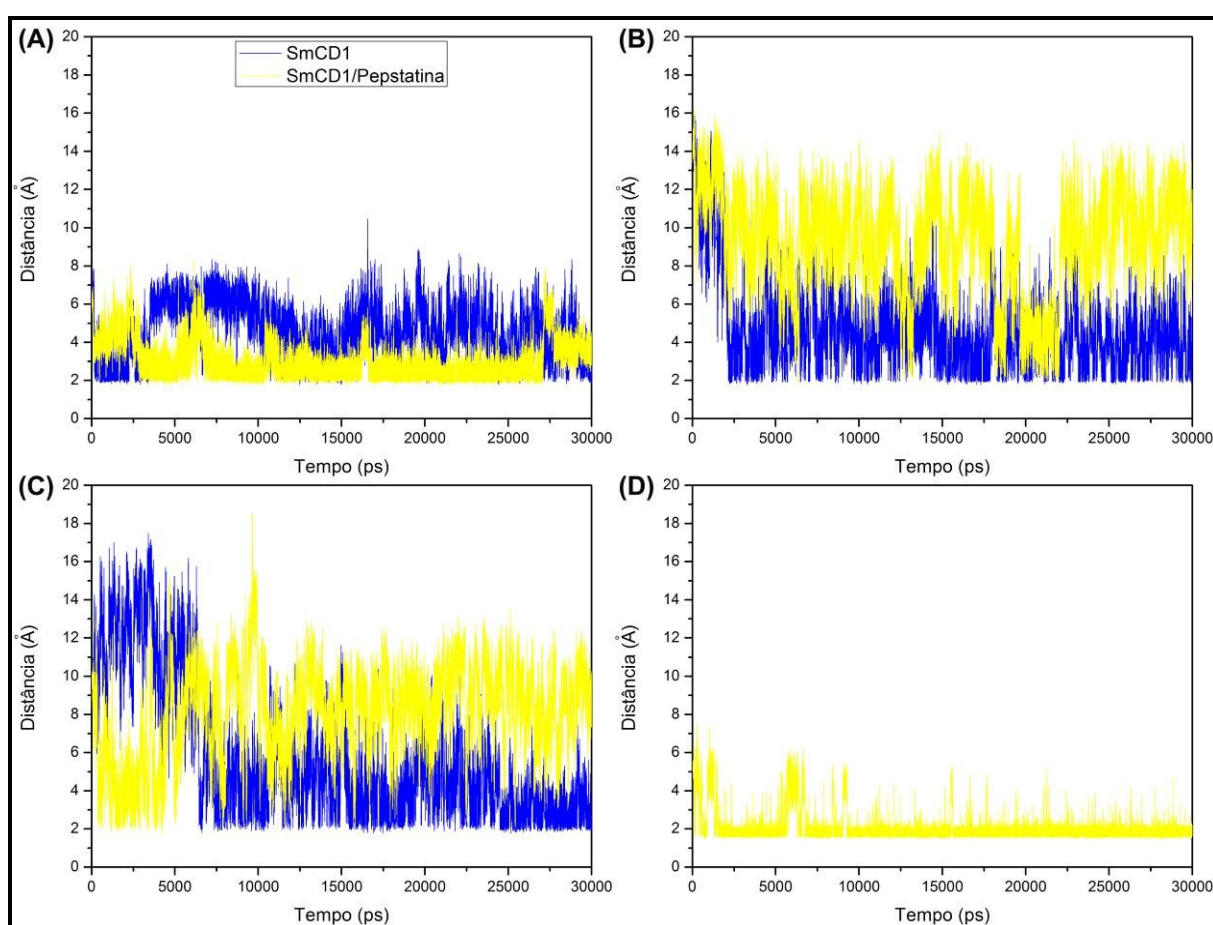
O resíduo de Asp277 está localizado na hélice h2C, próxima à fita s11C (FIGURA 4.26). A alça que conecta as fitas s10C e s11C apresentou amplo movimento (como discutido na seção 3.3.5), o qual resultou na aproximação da alça, entre as fitas s9N e s1C, e a hélice h2C. Este movimento permitiu a alta ocorrência ao longo da trajetória da ponte salina formada pelos resíduos Asp277-Arg160 no complexo SmCD1/pepstatina (FIGURA 4.24), refletindo na menor flexibilidade desta região (FIGURA 4.23). Em conjunto com o subsítio hidrofóbico S4, essa ponte salina parece ser importante para a especificidade e estabilidade da ligação do inibidor na SmCD1.

Entretanto, sugere-se que para produzir a aproximação entre alça entre as fitas s9N e s1C, e a hélice h2C, a interação entre os resíduos Lys281 e Glu288 torna-se menos freqüente no complexo SmCD1/pepstatina (TABELA 1 do Anexo V). A perda dessa interação é compensada com o aumento da ocorrência da ponte salina entre Asp277 e Arg160.

A interação entre os resíduos carregados Asp277 e Arg160 não ocorrem na SmCD2 e na SmCD3, onde os resíduos de Asp277 ou Arg160 são mutados para resíduos neutros. Já a ponte salina formada pelo par Lys281 e Glu288 na SmCD1

corresponde ao par Glu282 e Arg289 na SmCD2. Essa ponte salina não ocorre na SmCD3, pois os resíduos de Lys281 e Glu288 da SmCD1 correspondem a Lys279 e Lys286, respectivamente.

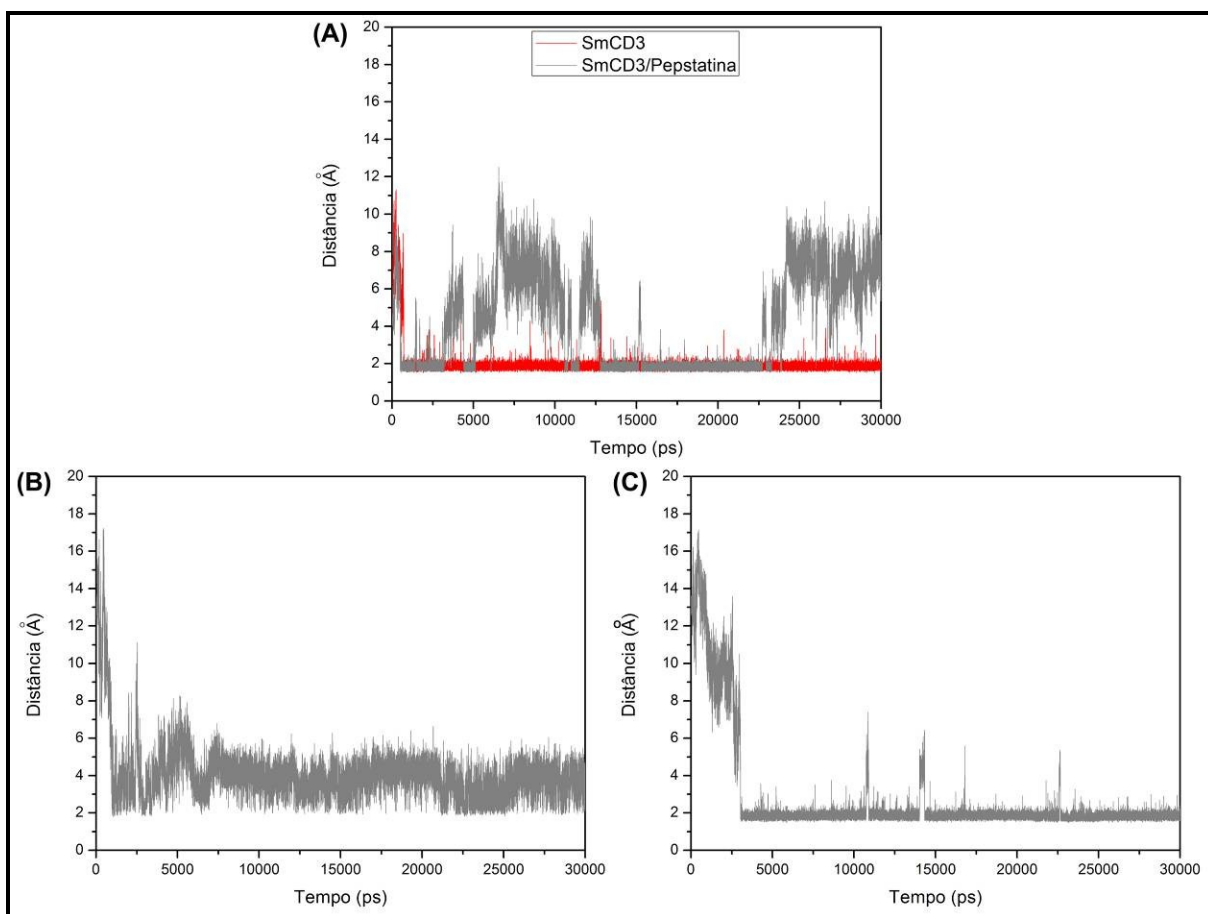
Os resíduos Asp188 e Lys198 estão posicionados nas fitas s2C e s3C, respectivamente. Sugere-se que a perda desta interação no complexo SmCD1/pepstatina (FIGURA 4.24) se deve à formação da ligação de hidrogênio entre o grupo hidroxila da cadeia lateral da Tyr194 com o átomo de oxigênio da carbonila da cadeia principal de P2' (Ala) do ligante (distância média  $O\cdots O = 3,5 \pm 1,6$  Å) (TABELA 1 do Anexo VI). Contudo, a interação entre Asp188 e Lys198 não aparece na SmCD2 e SmCD3.



**FIGURA 4.24.** Gráficos de evolução temporal das distâncias entre as principais pontes salinas formadas pela SmCD1 livre (azul) e pelo complexo SmCD1/pepstatina (amarelo). **(A)** Arg160 e Asp277; **(B)** Asp188 e Lys198; **(C)** Lys281 e Glu288; **(D)** Pepstatina e His77.

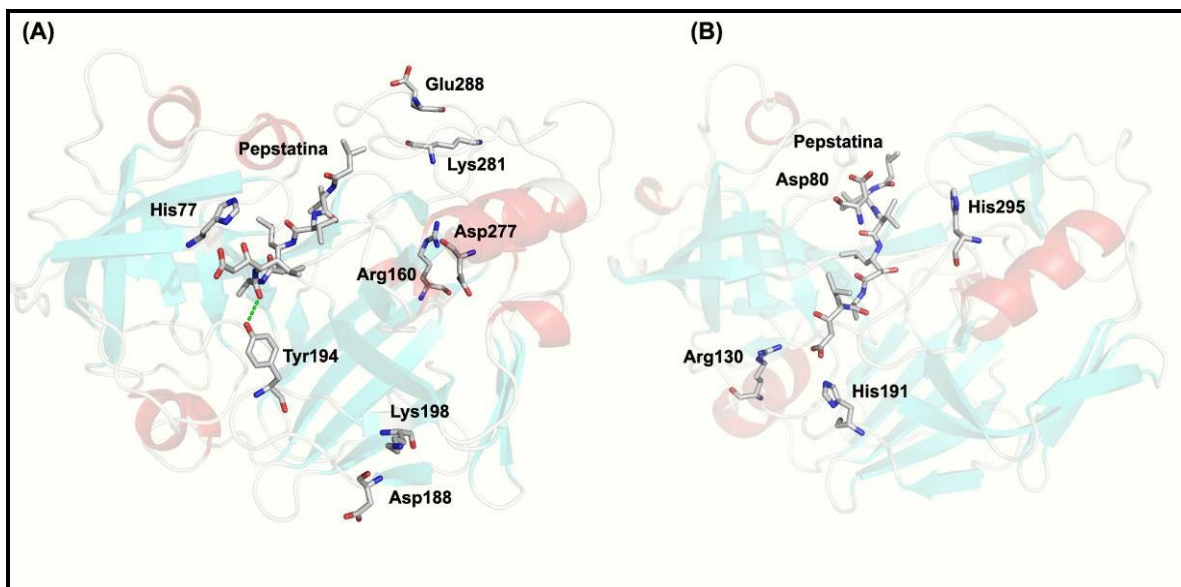
O par iônico formado pelos resíduos Asp80 e His295 na SmCD3 está localizado na região do “flap” e na alça entre as fitas s13C e s14C (FIGURA 4.25 e FIGURA 4.26). Sua ocorrência diminui substancialmente quando a pepstatina se liga à SmCD3 (TABELA 3 do Anexo V), pois o inibidor produz uma ligação de hidrogênio com o resíduo de Asp80 (TABELA 1 do Anexo VI).

A pepstatina produz apenas uma ponte salina com o subsítio S3' na enzima SmCD1 (FIGURA 4.25 e FIGURA 4.26), como observado durante a simulação por dinâmica molecular. O átomo de nitrogênio da cadeia lateral do resíduo de His77 interage com o carboxilato da pepstatina em 94,7% da simulação total (distância média  $\text{NH}\cdots\text{O} = 2,0 \pm 0,7 \text{ \AA}$ ). A enzima SmCD3 produz duas pontes salinas com a pepstatina (FIGURA 4.25 e FIGURA 4.26). Os resíduos de Arg130 e His191 interagem com o carboxilato da pepstatina (P3') em 58,7% e 88,7% da simulação (distância média  $\text{NH}\cdots\text{O} = 4,0 \pm 1,6 \text{ \AA}$  e  $2,8 \pm 2,7 \text{ \AA}$ , respectivamente). Estes contatos sugerem que um novo bolso de ligação é formado (subsítio S4'), que não está presente na SmCD2 e outras aspartil proteases em complexo com a pepstatina [88, 90].



**FIGURA 4.25.** Gráficos de evolução temporal das distâncias entre as principais pontes salinas formadas pela SmCD3 livre (vermelho) e pelo complexo SmCD3/pepstatina (cinza). **(A)** Asp80 e His295; **(B)** Pepstatina e Arg130; **(C)** Pepstatina e His191.

Outras pontes salinas estão envolvidas na estabilidade geral das proteínas SmCD1, SmCD2 e SmCD3, mas não modificam a especificidade das mesmas, pois os resíduos estão localizados externamente aos sítios de interação com o ligante.



**FIGURA 4.26.** Representação das principais pontes salinas importantes para a interação com a pepstatina. **(A)** SmCD1 e **(B)** SmCD3.

#### 4.3.4 O PAPEL DAS LIGAÇÕES DE HIDROGÊNIO NAS AFINIDADES DOS COMPLEXOS

As aspartil proteases possuem duas alças homólogas no sítio catalítico, formado pelos resíduos de Asp-Thr-Gly [39]. Cada tríade está localizada em lobos opostos semelhantes. Uma característica importante dessa família de enzimas é a interação conservada denominada “fireman’s grip” [39], que envolve ligações de hidrogênio entre os dois resíduos de treonina da tríade catalítica.

A primeira ligação de hidrogênio é formada entre o oxigênio da cadeia lateral da treonina de um lobo e a carbonila da cadeia principal do resíduo anterior ao ácido aspartico catalítico do lobo oposto. Pelas análises de DM, esta interação apresenta ocorrência maior que 50% ao longo da simulação de todas as enzimas, livres e inibidas. Contudo, a interação produzida pela Thr34 na SmCD2 livre apresenta baixa ocorrência inesperada – abaixo de 10% ao longo da trajetória. Esta interação parece estabilizar a ligação da pepstatina, pois é encontrada com ocorrência maior que 50% ao longo das trajetórias de DM dos complexos (FIGURA 4.27). Diversos estudos descrevem o resíduo de Thr34 como importante para a atividade da enzima HIV-protease [212-214], o que pode também ser apropriado para as SmAPs.

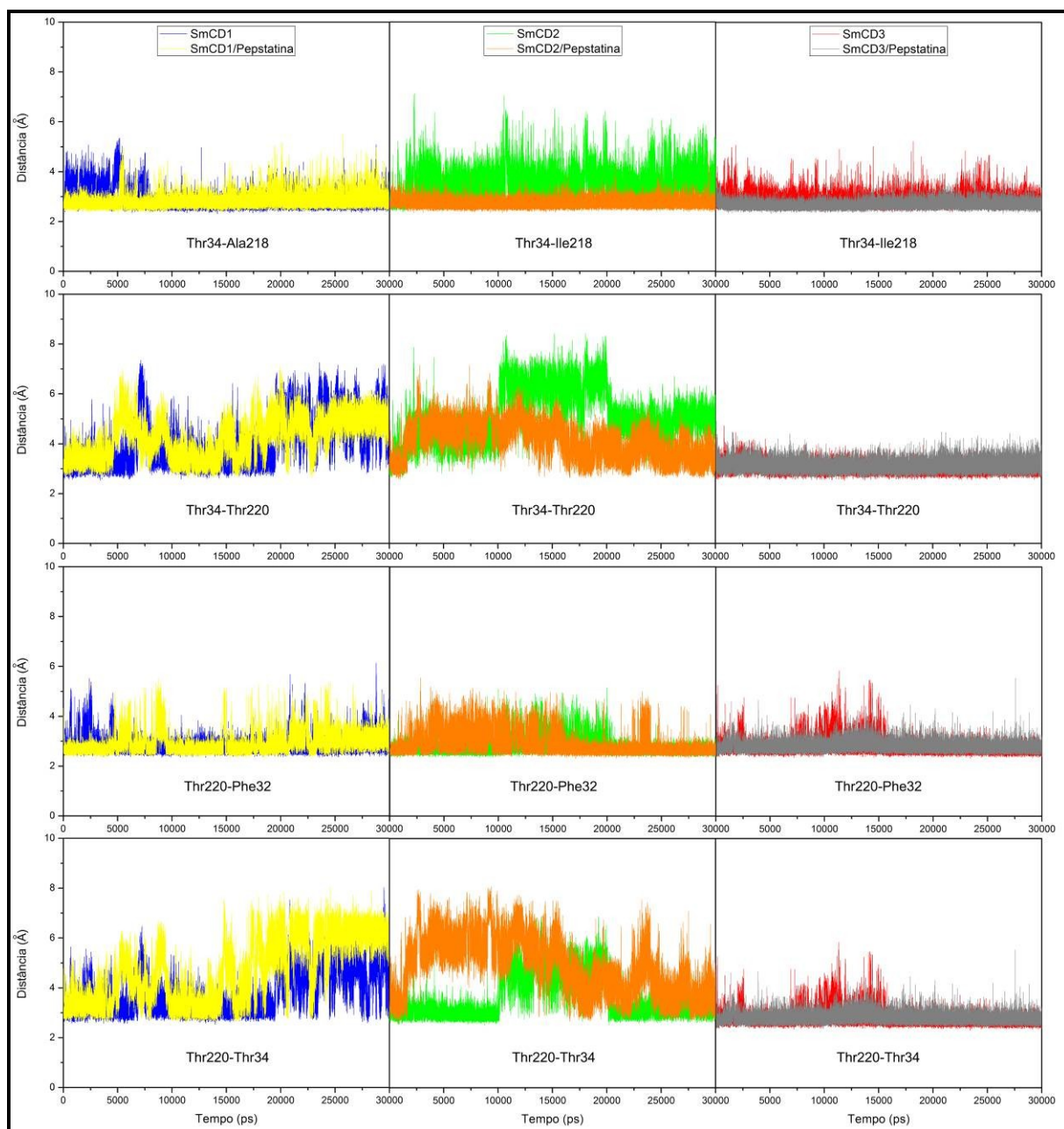
Interações semelhantes são produzidas pela hidroxila do resíduo de Thr do sítio ativo e o nitrogênio da cadeia principal do resíduo Thr do lóbulo oposto. Esta



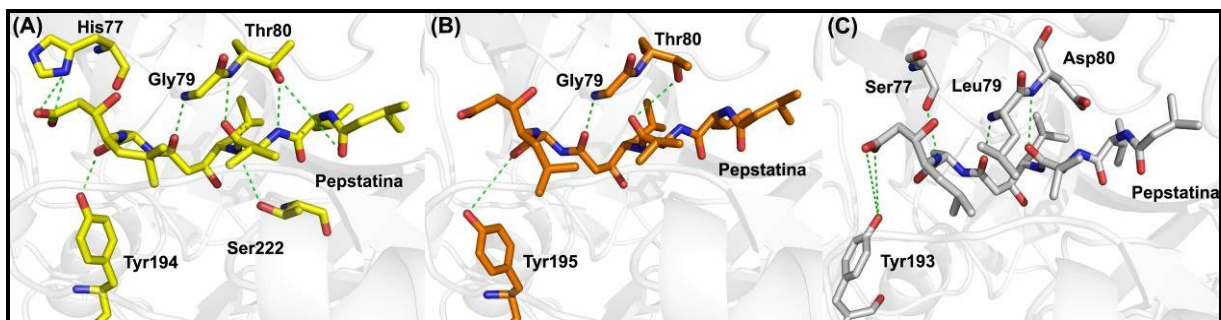
ligação de hidrogênio apresenta ocorrência abaixo de 50% ao longo das trajetórias para SmCD1 e SmCD2, especialmente quando as enzimas estão em complexo com a pepstatina (FIGURA 4.27 e TABELA 1 do Anexo VI). Além disso, esta interação deve ser importante para SmCD3, onde é encontrada com prevalência maior que 55% durante as simulações (TABELA 1 do Anexo VI).

Outra propriedade estrutural significativa de todas as aspartil proteases é a região denominada “flap”, a qual é descrita como crucial nas interações enzima-ligante [180, 196, 203]. A ligação do substrato/inibidor em aspartil proteases auxilia a manter o “flap” estável, reduzindo seu movimento [172, 203, 215]. Esse fato pode ser observado pelos resultados da DM, pois o “flap” apresenta menor flexibilidade nas enzimas ligadas à pepstatina (FIGURA 4.23).

Essa região é bastante conservada na SmCD1 e SmCD2, mas algumas diferenças foram encontradas na SmCD3. A cadeia principal da pepstatina produz ligação de hidrogênio com o resíduo de Gly79 e com a cadeia lateral do resíduo de Thr80, tanto da SmCD1 quanto na SmCD2. Embora o resíduo de Gly79 seja mutado por Leu79 na SmCD3, o átomo de nitrogênio da cadeia principal desse resíduo também produz uma ligação de hidrogênio com o resíduo de valina da pepstatina em P1' (distância média  $N \cdots O = 2,8 \pm 0,2 \text{ \AA}$ ) (TABELA 2 do Anexo VI). Já o resíduo de Thr80 é mutado por Asp80 em SmCD3, o qual produz uma ligação de hidrogênio com o oxigênio da cadeia principal da valina na pepstatina, mas esta interação apresenta baixa ocorrência durante a DM. Tais mutações no “flap” da SmCD3 compõem subsítios S1 e S3' menores, os quais podem não aceitar cadeias laterais volumosas do inibidor (FIGURA 4.28).



**FIGURA 4.27.** Gráficos de evolução temporal das distâncias entre as ligações de hidrogênio do “fireman’s grip”: SmCD1 livre (azul), SmCD1/pepstatina (amarelo), SmCD2 (verde), SmCD2/pepstatina (laranja), SmCD3 (vermelho) e SmCD3/pepstatina (cinza). Os átomos envolvidos nessas interações, bem como as distâncias médias, estão descritos na TABELA 1 do Anexo VI.

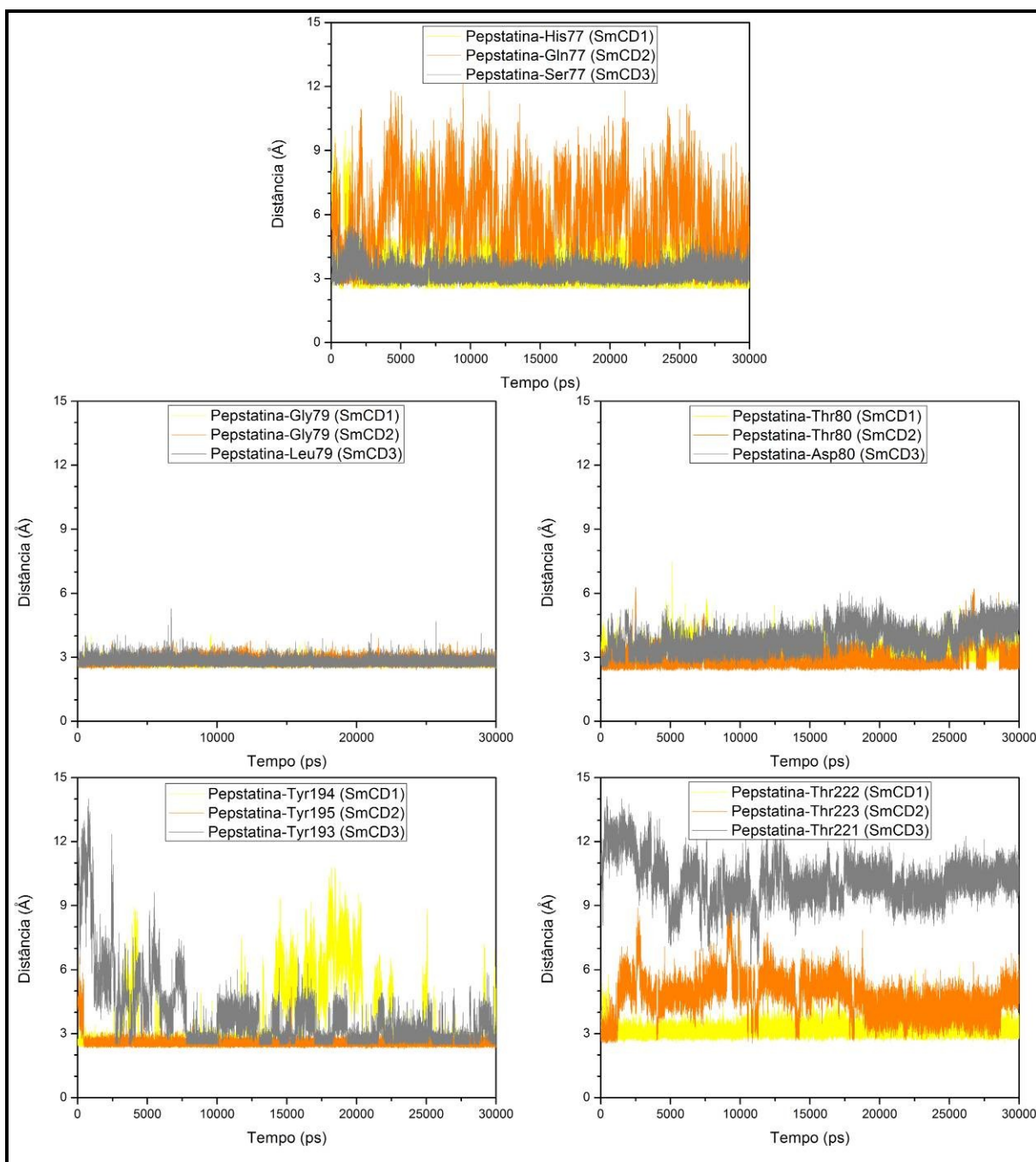


**FIGURA 4.28.** Representação gráfica dos complexos entre pepstatina e **(A)** SmCD1, **(B)** SmCD2 e **(C)** SmCD3. Destaca-se a localização das ligações de hidrogênio com mais de 50% de ocorrência ao longo da trajetória. A cadeia principal é exibida utilizando-se a representação esquemática e as cadeias laterais dos resíduos específicos são mostradas em bastão.

Tyr194 é outro resíduo importante, conservado na pepsina humana, que forma uma ligação de hidrogênio com a pepstatina [90]. O oxigênio da cadeia lateral de Tyr194 da enzima SmCD1 (Tyr195 na SmCD2) produz esta interação com a carbonila do resíduo de Ala da pepstatina em P2' (distância média  $O \cdots O = 0,35 \pm 0,16$  nm para SmCD1 e  $0,26 \pm 0,02$  nm para SmCD2). Embora o resíduo equivalente na enzima SmCD3 seja a Met194, o resíduo vizinho Tyr193 forma uma ligação de hidrogênio com o grupo carboxílico da estatina terminal da pepstatina em P3' (distância média  $O \cdots O = 3,8 \pm 1,7$  Å, para ambos os oxigênios) (FIGURA 4.28 e FIGURA 4.29).

Ambos os átomos de oxigênio da região C-terminal da pepstatina produzem uma ligação de hidrogênio com N $\delta$ 1 da cadeia lateral do resíduo His77 na SmCD1, sugerindo que este resíduo apresenta uma conformação estável no complexo, favorável à interação com o inibidor. Na enzima SmCD3, o resíduo correspondente é Ser77, que forma uma ligação de hidrogênio com o nitrogênio da cadeia principal da estatina (P3' da pepstatina). Essas interações não ocorrem com o resíduo Gln77 da SmCD2, que se localiza em direção ao exterior da proteína. Esse resíduo é mais volumoso, impedindo a interação com o inibidor. Estes resultados confirmam a presença de um novo subsítio de ligação na enzima SmCD1 (subsítio S4').

O grupo hidroxila da cadeia lateral de Thr222 produz uma ligação de hidrogênio com a cadeia principal do resíduo de estatina da pepstatina no subsítio S1. Entretanto, os resíduos correspondentes na enzima SmCD2 (Thr223) e SmCD3 (Thr221) não formam esta interação.



**FIGURA 4.29.** Gráficos de evolução temporal das distâncias de algumas ligações de hidrogênio entre a pepstatina e as enzimas SmCD1 SmCD1 (amarelo), SmCD2 (laranja) e SmCD3 (cinza). Os átomos envolvidos nessas interações, bem como as distâncias médias, estão descritos na TABELA 1 do Anexo VI.

Surpreendentemente, a ligação de hidrogênio entre os átomos de oxigênio dos aspartatos catalíticos e a hidroxila do resíduo de estatina central da pepstatina apresenta ocorrência abaixo de 25% ao longo das simulações dos complexos estudados (TABELAS 2 e 3 do Anexo VI). Os aspartatos catalíticos favorecem a

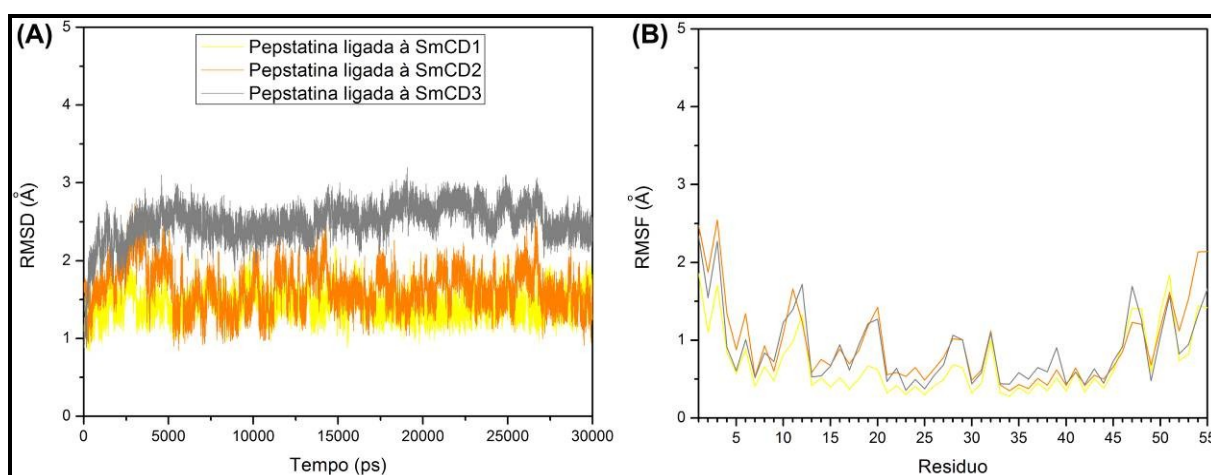
ligação de hidrogênio entre si, como observado na ocorrência dessa interação ao longo das trajetórias (TABELA 4.11), o que pode ser um indício da existência das LBHBs (do inglês, “Low Barrier Hydrogen Bond”, Ligação de Hidrogênio de Baixa Barreira) discutidas na seção 1.1.3.3.5.

Pela análise de RMSD e RMSF da pepstatina ao longo da trajetória de DM (FIGURA 4.26), observa-se que esse inibidor não apresenta grande alteração conformacional em relação à estrutura inicial e que o complexo SmCD1/pepstatina é o mais estável durante a simulação. Além disso, os átomos que constituem a hidroxila (número 30 e 31, respectivamente) apresentam valores de RMSF abaixo de 0,6 Å.

**TABELA 4.11.** Ligações de hidrogênio entre os resíduos de aspartato catalíticos, Asp33 e Asp219, durante a DM das enzimas estudadas.

Complexo	Resíduo	Átomo	Resíduo	Átomo	Distância Média (Å) ± DP	Ocorrência (%) <sup>1</sup>
SmCD1/ pepstatina	Asp33	OD1	Asp219	OD2	2,2 ± 0,0	68,7
		OD2			2,9 ± 0,4	30,7
SmCD2/ pepstatina	Asp33	OD1	Asp219	OD2	3,1 ± 0,8	74,6
		OD2			4,0 ± 0,8	26,9
SmCD3/ pepstatina	Asp33	OD1	Asp219	OD2	2,8 ± 0,2	92,8
		OD2			2,9 ± 0,3	17,0

<sup>1</sup> Ocorrência calculada com distância máxima de 3,5 Å e um ângulo máximo de 30°, respectivamente, durante toda a simulação.



**FIGURA 4.30.** RMSD e RMSF de todos os átomos da pepstatina ligada à: SmCD1 (amarelo), SmCD2 (laranja) e SmCD3 (cinza).

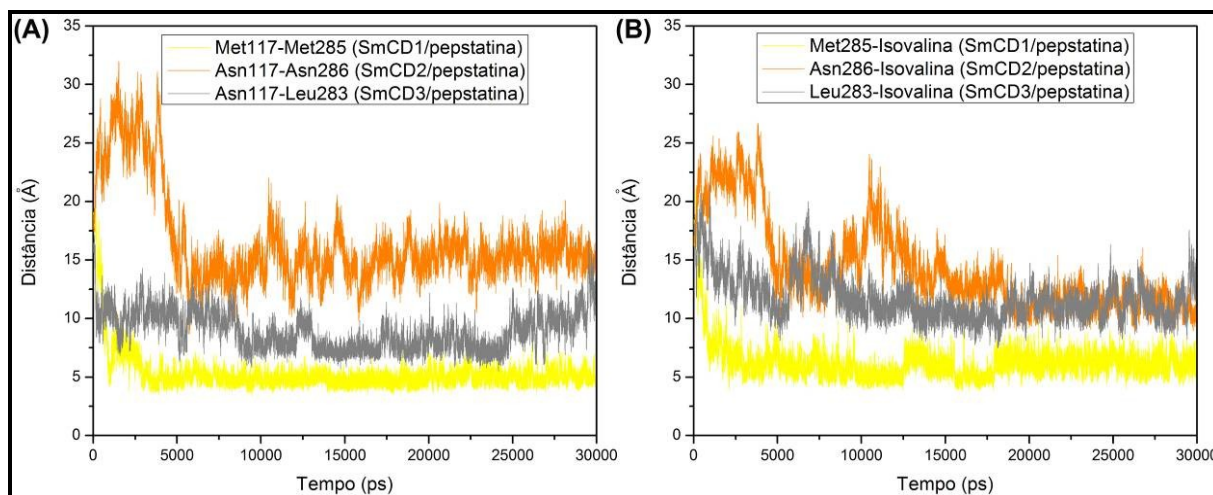
A pepstatina mostra maior número de ligações de hidrogênio com SmCD1 ( $8,2 \pm 1,5$ ) do que com SmCD2 ( $5,3 \pm 1,4$ ) e SmCD3 ( $5,2 \pm 1,5$ ). Sugere-se que as ligações de hidrogênio entre a pepstatina e a SmCD1 possam contribuir para a maior estabilização do complexo, em relação àqueles formados com a SmCD2 e a SmCD3.

### 3.3.5 O PAPEL DAS INTERAÇÕES HIDROFÓBICAS NO PROCESSO DE INIBIÇÃO

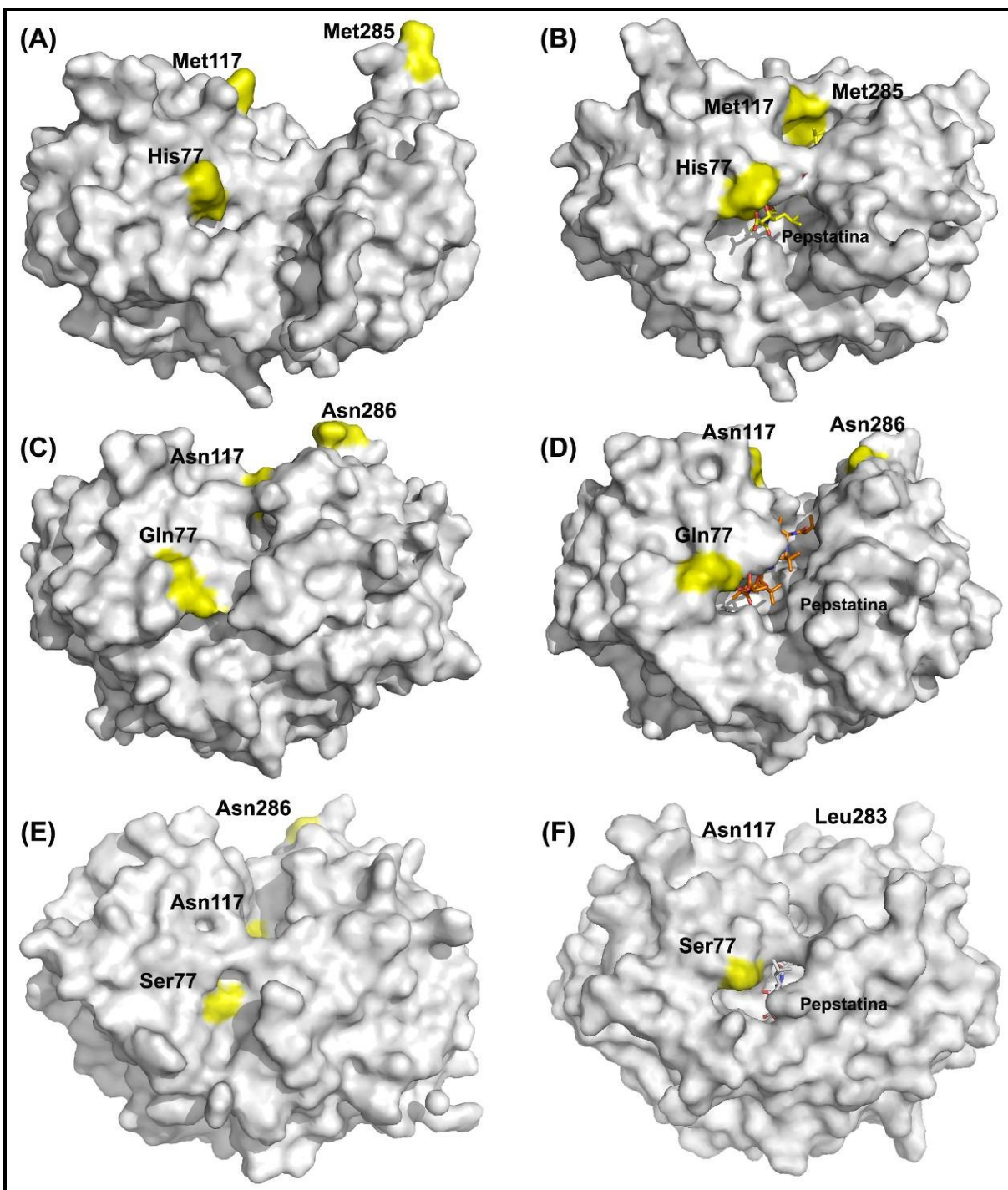
Embora a hélice h2N (resíduos 113 a 117) e a alça que conecta as fitas s10C e s11C (resíduos 282 a 288) não tenham sido descritas como relacionadas ao sítio de ligação de inibidores/substratos na SmCD1 [80, 93], as análises das simulações demonstraram uma grande alteração conformacional em direção ao resíduo de isovalina da pepstatina (P4). Nos primeiros 5 ns de simulação, foi observada uma queda significativa nas distâncias intramoleculares entre os resíduos da hélice h2N e a alça que conecta as fitas s10C e s11C (em particular, entre os resíduos Met117 e Met285) (FIGURA 4.27). Esta aproximação produziu contatos favoráveis entre a SmCD1 e a pepstatina. Esta interação caracteriza um subsítio mais compacto e hidrofóbico em S4. Dealwis e colaboradores também descreveram um subsítio mais compacto envolvendo a posição da hélice h2N na renina de camundongo [216]. Tais alterações conformacionais na enzima SmCD1 também podem ser visualizadas pelo RMSF (FIGURA 4.23). Estas regiões sofrem significativa redução da flexibilidade na SmCD1 complexada à pepstatina. Estes resultados sugerem a presença de um “segundo flap” na alça entre as fitas s10C e s11C da SmCD1, sendo importante para a especificidade e a ligação do substrato/inibidor à enzima SmCD1.

Entretanto, as mudanças conformacionais descritas acima não foram observadas no complexo SmCD2/pepstatina (correspondendo à alça entre as fitas s11C e s12N). Embora a distância entre Asn117 e Asn286 rapidamente decaia nos primeiros 5 ns na SmCD2, a mesma é muito maior que na enzima SmCD1 (FIGURA 4.31). Isso ocorre devido à composição de aminoácidos hidrofílicos dessa região na SmCD2, pois Met117 e Met285 da SmCD1 correspondem a Asn117 e Asn286 na SmCD2. Adicionalmente, Gly247 é mutado para Asn248, o que contribui para um subsítio S4 mais polar na SmCD2.

A enzima SmCD3 mostra um comportamento diferente daqueles observados para SmCD1 e SmCD2, em relação à evolução das distâncias entre os resíduos de aminoácidos na região do subsítio S4. Embora a distância final entre Asn117 e Leu283 e entre Leu283 e o resíduo de isovalina da pepstatina não revelem vastas diferenças na SmCD2, a conformação da alça entre as fitas s10C e s11C na SmCD3 livre não apresenta grande diferença da enzima inibida, o que pode ser visualizado nas FIGURAS 4.31 e 4.32.



**FIGURA 4.31.** Gráficos de evolução temporal das distâncias. **(A)** Met117 e Met285 no complexo SmCD1/pepstatina (amarelo), Asn117 e Asn286 no complexo SmCD2/pepstatina (laranja), e Asn117 e Leu283 no complexo SmCD3/pepstatina (cinza); **(B)** Met285 do complexo SmCD1/pepstatina (amarelo), Asn286 do complexo SmCD2/pepstatina (laranja), Leu283 do complexo SmCD3/pepstatina (cinza), e isovalina da pepstatina.



**FIGURA 4.32.** Representações da superfície das enzimas livres e dos complexos. **(A)** SmCD1, **(B)** SmCD1/pepstatina, **(C)** SmCD2, **(D)** SmCD2/pepstatina, **(E)** SmCD3 e **(F)** SmCD3/pepstatina. As estruturas correspondem à última fornecida pela simulação (30 ns).



### 4.3 PROPOSTA DE MODIFICAÇÕES ESTRUTURAIS PARA NOVOS INIBIDORES SELETIVOS

Para se tornarem potenciais candidatos a fármacos, os inibidores de protease não devem ter apenas alta potência e seletividade. Eles devem possuir também propriedades farmacocinéticas e farmacodinâmicas apropriadas [83]. Entretanto, informações de como aspartil proteases seletivamente reconhecem pequenas moléculas orgânicas geralmente são derivados de estudos com substratos peptídicos / inibidores peptidomiméticos. E a chance desses ligantes se tornarem fármacos aumenta na medida em que se convertem peptídeos em não-peptídeos.

Além disso, inibidores seletivos para uma determinada protease pertencente a um grupo de enzimas homólogas, como é o caso da família SmAP, podem ser muito úteis como ferramentas para estudo da função destas enzimas. Este inibidor pode ser usado para “nocautear” quimicamente uma protease específica presente numa mistura de enzimas ou num organismo inteiro, contribuindo assim para caracterizar o papel da enzima nocauteada na atividade biológica em estudo.

Dessa forma, sugere-se nesse trabalho uma estratégia para gerar inibidores seletivos para cada protease SmCD. Com base nas análises descritas, indica-se a modificação de resíduos que compõe a série LBPP (FIGURA 4.33), para a maior seletividade dos inibidores.

Para não reduzir a potência dos compostos em relação à SmCD1 e SmCD2, propõe-se manter o radical isobutil em P1, pois o resíduo de Phe120 nessas enzimas favoreceria resíduos menores no subsítio S1. Além disso, o substituinte isobutil conservaria as interações hidrofóbicas com os resíduos Phe115, Phe120 e Ile123 dessas enzimas. A enzima SmCD3, entretanto, apresenta um subsítio P1 mais polar, com a presença de alguns resíduos de tirosina em sua composição. Dessa forma, propõe-se a alteração do grupo isobutil pelo substituinte hidroximetil para permitir interações por ligação de hidrogênio com esses resíduos, especialmente Tyr120.

Em P3, o resíduo de glutamina de LBPP1 interage por ligação de hidrogênio com o resíduo de Ser223 e de Ala13 na SmCD1. Na enzima SmCD2, esses resíduos correspondem a Ala224 e Ile13, respectivamente. A modificação da glutamina, em P3, para valina poderia favorecer a interação hidrofóbica entre o

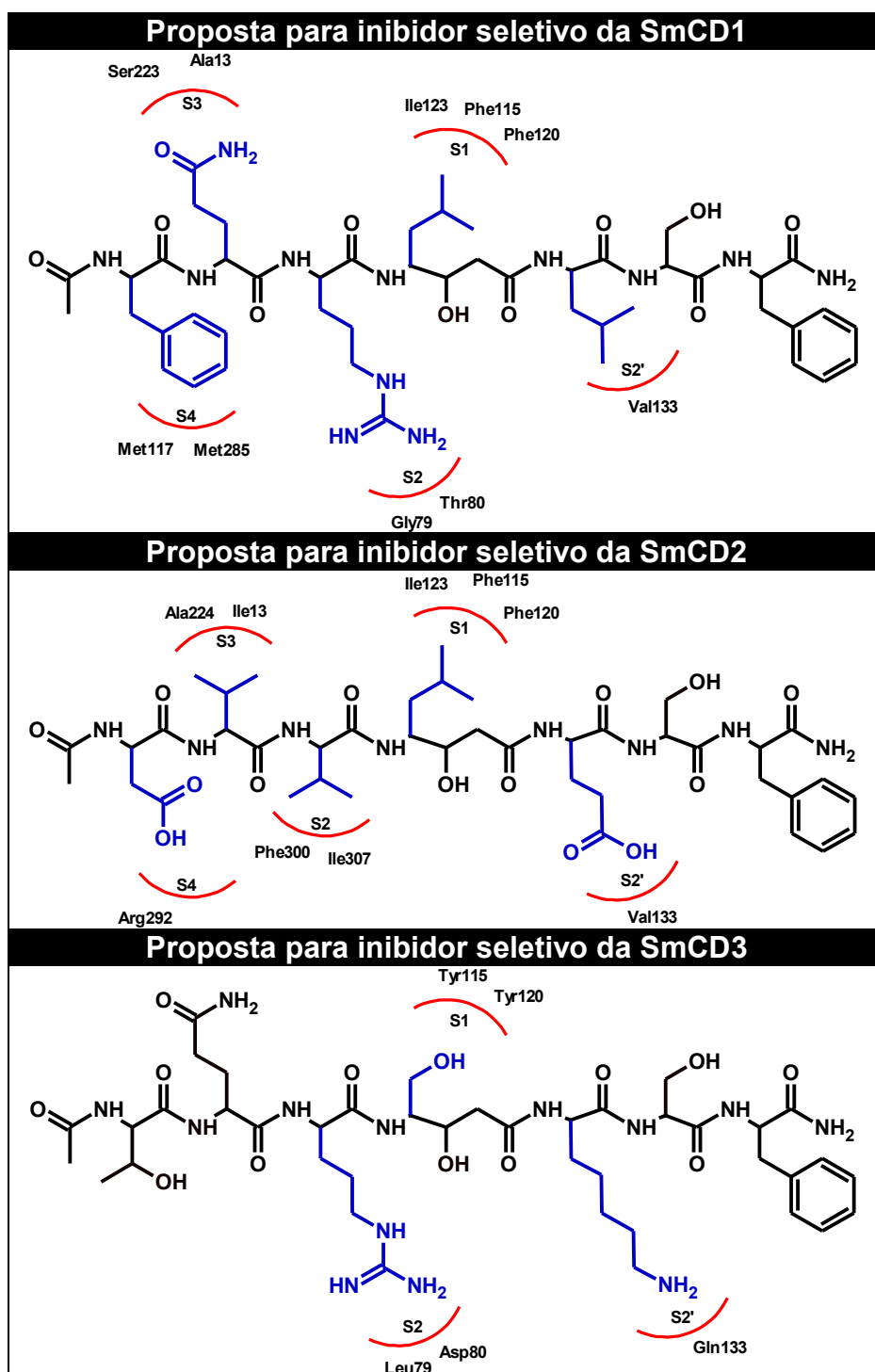
inibidor e esses resíduos na SmCD2. Além de hidrofóbico, a valina é um aminoácido menos volumoso que evitaria um impedimento estérico, especialmente com Ile13.

Sugere-se que os resíduos do inibidor em P2 sejam importantes para a ligação na enzima, pois envolvem interações com os resíduos constituintes do “flap”. Na SmCD1, todos os ligantes da série LBPP produzem ligações de hidrogênio entre os resíduos Gly79 e Thr80 e a cadeia lateral de glutamina do inibidor. Na SmCD2, a conformação da cadeia lateral dos ligantes da série LBPP é diferente daquela encontrada para a SmCD1 de forma que as cadeias laterais dos inibidores se localizam próximos aos resíduos Phe300 e Ile307. Assim, propõe-se a mutação da arginina por resíduos hidrofóbicos menos volumosos, como alanina ou valina. Esses resíduos não afetariam as interações por ligação de hidrogênio entre os resíduos do “flap” e a cadeia principal do inibidor, e ainda promoveriam interações hidrofóbicas nesse subsítio da proteína.

O subsítio S2' é uma região que também poderia ser explorada para avaliar a seletividade na enzima SmCD3. A mudança de carga do resíduo de P2' (substituindo o resíduo de Glu por Lys, por exemplo) poderia fornecer uma interação por ligação de hidrogênio com o resíduo Gln133. Ao alterar por um resíduo hidrofóbico (Ile ou Leu, por exemplo), esse substituinte poderia formar interações hidrofóbicas com o resíduo de Val133 da enzima SmCD1.

Para fornecer um ligante específico para a SmCD1, também poder-se-ia explorar a modificação do resíduo de serina em P4 da série de ligantes LBPP. Propõe-se uma mudança para um resíduo de leucina ou fenilalanina, ambos hidrofóbicos, para interação com o “segundo flap” na alça entre as fitas s10C e s11C da SmCD1.

A mutação, nesse subsítio, para um resíduo carregado negativamente (Asp ou Glu) poderia favorecer a interação iônica com Arg292, aumentando a seletividade para a enzima SmCD2.



**FIGURA 4.33.** Representação esquemática das modificações estruturais propostas para síntese de inibidores seletivos das enzimas SmCD1, SmCD2 e SmCD3.

## 5 CONCLUSÕES E PERSPECTIVA

As características estruturais das enzimas catepsina D-símile de *Schistosoma mansoni* SmCD1, SmCD2 e SmCD3 foram investigadas para permitir o futuro desenvolvimento de novos e seletivos inibidores.

Os modelos 3D das enzimas foram construídos por modelagem comparativa e adequadamente validados. Os três modelos apresentaram estruturas semelhantes às aspartil proteases, o que era esperado devido ao enovelamento altamente conservado nessa família de enzimas.

A metodologia de atracção molecular foi validada pelo “re-docking” e “cross-docking” de enzimas homólogas complexadas a peptidomiméticos. Os resultados indicaram a formação de poucos aglomerados, sendo que a conformação de menor energia estava inserida no aglomerado mais populoso, com exceção do “cross-docking” da pepstatina na pepsina humana. Nesse caso, o aglomerado mais populoso não era o de maior energia, mas as conformações contidas nesse aglomerado não se adequaram aos parâmetros requeridos nesse trabalho. Dessa forma, foram selecionadas as conformações de menor energia para todos os ligantes.

Em seguida, a metodologia foi empregada para os três inibidores peptidomiméticos da série LBPP e da pepstatina com a SmCD1, SmCD2 e SmCD3. Obtiveram-se as conformações de menor energia com os parâmetros exigidos nesse trabalho. Assim como a catepsina D humana, a SmCD1 favorece a presença de resíduos menores em P1 devido à presença de uma fenilalanina na posição 120, divergindo da renina de camundongo (código PBD 1SMR). Por se tratarem de ligantes maiores e diferentes cadeias laterais, a série LBPP adéqua-se de forma diferente à proteína quando comparados à pepstatina, e desta forma favorece interações que não existem com esta última. São sugeridas regiões específicas para cada enzima, as quais poderiam ser exploradas para o futuro desenvolvimento de novos inibidores.

Foram também realizados estudos por DM das SmAPs em complexo com a pepstatina. Os resultados sugerem que a pepstatina é capaz de provocar grandes modificações conformacionais na estrutura da SmCD1, quando comparada à ligação na SmCD2 e SmCD3. A maior diferença se encontra na alça que conecta as fitas s10C e s11C, favorecendo a formação de novas interações entre SmCD1 e a

pepstatina. Estes contatos são interações hidrofóbicas no subsítio S4, que é confirmado pela presença de pontes salinas na SmCD1, mas não fazem parte na SmCD2 e SmCD3.

A importância das interações denominadas “fireman’s grip” e do “flap” para a inibição também foi investigada pela análise das ligações de hidrogênio durante as DM das enzimas livres e ligadas à pepstatina. O resíduo de Thr33 (Th34 nas SmAPs) parece ser fundamental para a atividade enzimática. Adicionalmente, os resíduos do “flap” são mais flexíveis quando as enzimas livres são simuladas.

Os resultados indicaram que a especificidade de cada enzima pode ser determinada pelos resíduos externos à tríade catalítica. Como consequência, SmCD1, SmCD2 e SmCD3 são estruturalmente relacionadas entre si, mas apresentam especificidades distintas, o que pode indicar funcionalidade diferentes. Sugere-se ainda que o subsítio S4, em conjunto com o subsítio S2, são importantes alvos para o desenvolvimento de novos inibidores.

Testes experimentais com novas moléculas estão em andamento no Laboratório de Bioquímica de Proteínas de Peptídeos (IOC/FIOCRUZ) para identificar outros resíduos importantes na especificidade e no reconhecimento do ligante pela proteína. A caracterização bioquímica das enzimas purificadas fornecerá maior compreensão do mecanismo de inibição, o que será útil na triagem de compostos líderes para a descoberta de novas drogas esquistossomicidas.

## 6 REFERÊNCIAS

- [1] Nwaka S, Ridley RG. Virtual drug discovery and development for neglected diseases through public-private partnerships. *Nat Rev Drug Discov.* 2003;2(11):919-28.
- [2] Trouiller P, Olliaro P, Torreele E, Orbinski J, Laing R, Ford N. Drug development for neglected diseases: a deficient market and a public-health policy failure. *Lancet.* 2002;359(9324):2188-94.
- [3] Kettler HE, Modi R. Building local research and development capacity for the prevention and cure of neglected diseases: the case of India. *Bull World Health Organ.* 2001;79(8):742-7.
- [4] Trouiller P, Torreele E, Olliaro P, White N, Foster S, Wirth D, et al. Drugs for neglected diseases: a failure of the market and a public health failure? *Trop Med Int Health.* 2001;6(11):945-51.
- [5] Pordeus LC, Aguiar LR, Quinino LRM, Barbosa CS. A ocorrência das formas aguda e crônica da esquistossomose mansônica no Brasil no período de 1997 a 2006: uma revisão de literatura. *Epidemiol Serv Saúde.* 2008;17(3):163-75.
- [6] World Health Organization. Disponível em: <http://www.who.int/schistosomiasis/en/>. Acesso em: 18 de maio de 2011.
- [7] Ministério da Saúde. Disponível em: <http://portal.saude.gov.br/portal/svs/>. Acesso em: 18 de maio de 2011.
- [8] Doenças infecciosas e parasitárias: guia de bolso. 8 ed ed. Brasília: Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância Epidemiológica.  
2010.
- [9] Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. *Lancet.* 2006;368(9541):1106-18.

- [10] King CH, Dickman K, Tisch DJ. Reassessment of the cost of chronic helminthic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis. *Lancet*. 2005;365(9470):1561-9.
- [11] Neves D. *Parasitologia Humana*. 11 ed. ed. São Paulo: Atheneu 2005.
- [12] La Flamme AC, Canagasabay K, Harvie M, Backstrom BT. Schistosomiasis protects against multiple sclerosis. *Mem Inst Oswaldo Cruz*. 2004;99(5):33-6.
- [13] Loukas A, Bethony JM. New drugs for an ancient parasite. *Nat Med*. 2008;14(4):365-7.
- [14] Cioli D, Picamattoccia L, Archer S. Antischistosomal Drugs - Past, Present ... and Future. *Pharmacol Ther*. 1995;68(1):35-85.
- [15] Geerts S, Gryseels B. Anthelmintic resistance in human helminths: a review. *Trop Med Int Health*. 2001;6(11):915-21.
- [16] Cioli D, Picamattoccia L, Archer S. Antischistosomal drugs - Past, present... and future. *Pharmacology & Therapeutics*. 1995;68(1):35-85.
- [17] Ismail M, Botros S, Metwally A, William S, Farghally A, Tao LF, et al. Resistance to praziquantel: Direct evidence from *Schistosoma mansoni* isolated from Egyptian villagers. *Am J Trop Med Hyg*. 1999;60(6):932-5.
- [18] Melman SD, Steinauer ML, Cunningham C, Kubatko LS, Mwangi IN, Wynn NB, et al. Reduced Susceptibility to Praziquantel among Naturally Occurring Kenyan Isolates of *Schistosoma mansoni*. *Plos Neglect Trop Dis*. 2009;3(8):10.
- [19] Gryseels B, Mbaye A, De Vlas SJ, Stelma FF, Guisse F, Van Lieshout L, et al. Are poor responses to praziquantel for the treatment of *Schistosoma mansoni* infections in Senegal due to resistance? An overview of the evidence. *Tropical Medicine & International Health*. 2001;6(11):864-73.
- [20] Ismail M, Metwally A, Farghaly A, Bruce J, Tao LF, Bennett JL. Characterization of isolates of *Schistosoma mansoni* from Egyptian villagers that tolerate high doses of praziquantel. *Am J Trop Med Hyg*. 1996;55(2):214-8.

- [21] Pica-Mattoccia L, Doenhoff MJ, Valle C, Basso A, Troiani AR, Liberti P, et al. Genetic analysis of decreased praziquantel sensitivity in a laboratory strain of *Schistosoma mansoni*. *Acta Trop*. 2009;111(1):82-5.
- [22] Liang YS, Coles GC, Doenhoff MJ, Southgate VR. In vitro responses of praziquantel-resistant and -susceptible *Schistosoma mansoni* to praziquantel. *Int J Parasit*. 2001;31(11):1227-35.
- [23] William S, Botros S, Ismail M, Farghally A, Day TA, Bennett JL. Praziquantel-induced tegumental damage in vitro is diminished in schistosomes derived from praziquantel-resistant infections. *Parasitology*. 2001;122:63-6.
- [24] Renganathan E, Cioli D. An international initiative on praziquantel use. *Parasitology Today*. 1998;14(10):390-1.
- [25] Cioli D. Chemotherapy of schistosomiasis: An update. *Parasitology Today*. 1998;14(10):418-22.
- [26] Fenwick A, Savioli L, Engels D, Bergquist NR, Todd MH. Drugs for the control of parasitic diseases: current status and development in schistosomiasis. *Trends Parasitol*. 2003;19(11):509-15.
- [27] Meyer T, Sekljic H, Fuchs S, Bothe H, Schollmeyer D, Miculka C. Taste, A New Incentive to Switch to (R)-Praziquantel in Schistosomiasis Treatment. *Plos Neglect Trop Dis*. 2009;3(1):5.
- [28] Sayed AA, Simeonov A, Thomas CJ, Inglese J, Austin CP, Williams DL. Identification of oxadiazoles as new drug leads for the control of schistosomiasis. *Nat Med*. 2008;14(4):407-12.
- [29] McManus DP, Loukas A. Current status of vaccines for schistosomiasis. *Clin Microbiol Rev*. 2008;21(1):225-+.
- [30] Siddiqui AA, Ahmad G, Damian RT, Kennedy RC. Experimental vaccines in animal models for schistosomiasis. *Parasitol Res*. 2008;102(5):825-33.
- [31] Tendler M, Simpson AJG. The biotechnology-value chain: Development of SmA as a schistosomiasis vaccine. *Acta Trop*. 2008;108(2-3):263-6.



- [32] Capron A, Riveau G, Capron M, Trottein F. Schistosomes: the road from host-parasite interactions to vaccines in clinical trials. *Trends Parasitol.* 2005;21(3):143-9.
- [33] Bergquist R, Utzinger J, McManus DP. Trick or Treat: The Role of Vaccines in Integrated Schistosomiasis Control. *Plos Neglect Trop Dis.* 2008;2(6):4.
- [34] Rawlings ND, Morton FR, Kok CY, Kong J, Barrett AJ. MEROPS: the peptidase database. *Nucleic Acids Res.* 2008;36:D320-D5.
- [35] Rawlings ND, Barrett AJ. Evolutionary families of peptidases. *Biochemistry Journal.* 1993;290:205-18.
- [36] Kasný M, Mikes L, Hampl V, Dvorák J, Caffrey CR, Dalton JP, et al. Peptidases of Trematodes. In: Rollinson D, Hay SI, eds. *Advances in Parasitology.* Londres: Elsevier 2009:205-97.
- [37] Sodero A, De Simone S, Silva-Jr F. Mecanismo catalítico e estado de protonação do sítio ativo de aspartil proteases pepsina-símiles. *Rev Virtual Quím.* 2009;1(2):128-37.
- [38] Northrop J. Crystalline pepsin. I. Isolation and tests of purity. *J Gen Physiol.* 1931;13(6):739-66.
- [39] Pearl L, Blundell T. The Active-Site of Aspartic Proteinases. *FEBS Lett.* 1984;174(1):96-101.
- [40] Andreeva NS, Gurskaya GV. Interdomain interactions in aspartic proteases of higher organisms and their analogs in retroviral enzymes. *Mol Biol.* 2006;40(3):427-32.
- [41] Sakai H, Saku T, Kato Y, Yamamoto K. Quantitation and immunohistochemical localization of cathepsins E and D in rat tissues and blood cells. *Biochim Biophys Acta.* 1989;991(2):367-75.
- [42] Fruitier I, Garreau I, Piot JM. Cathepsin D is a good candidate for the specific release of a stable hemorphin from hemoglobin in vivo: VV-hemorphin-7. *Biochem Biophys Res Commun.* 1998;246(3):719-24.

- [43] Beyer BM, Dunn BM. Self-activation of recombinant human lysosomal procathepsin D at a newly engineered cleavage junction, 'short' pseudocathepsin D. *J Biol Chem.* 1996;271(26):15590-6.
- [44] Brindley PJ, Kalinna BH, Wong JYM, Bogitsh BJ, King LT, Smyth DJ, et al. Proteolysis of human hemoglobin by schistosome cathepsin D. *Molecular and Biochemical Parasitology.* 2001;112(1):103-12.
- [45] Andreeva NS, Rumsh LD. Analysis of crystal structures of aspartic proteinases: On the role of amino acid residues adjacent to the catalytic site of pepsin-like enzymes. *Protein Sci.* 2001;10(12):2439-50.
- [46] Sodero A, De Simone S, Silva-Jr F. Mecanismo catalítico e estado de protonação do sítio ativo de aspartil proteases pepsina-símiles. *Rev Virtual Quím.* 2009;1:128-37.
- [47] Veerapandian B, Cooper JB, Sali A, Blundell TL, Rosati RL, Dominy BW, et al. Direct observation by X-ray-analysis of the tetrahedral intermediate of aspartic proteinases. *Protein Sci.* 1992;1(3):322-8.
- [48] Schechter I, Berger AC. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun.* 1967;27(2):157-62.
- [49] Brinkworth RI, Prociv P, Loukas A, Brindley PJ. Hemoglobin-degrading, aspartic proteases of blood-feeding parasites - Substrate specificity revealed by homology models. *J Biol Chem.* 2001;276(42):38844-51.
- [50] Turk B. Targeting proteases: successes, failures and future prospects. *Nat Rev Drug Discov.* 2006;5(9):785-99.
- [51] Bailey D, Cooper JB. A Structural Comparison of 21-Inhibitor Complexes of the Aspartic Proteinase from *Endothia-Parasitica*. *Protein Sci.* 1994;3(11):2129-43.
- [52] Coates L, Erskine PT, Mall S, Gill R, Wood SP, Myles DAA, et al. X-ray, neutron and NMR studies of the catalytic mechanism of aspartic proteinases. *Eur Biophys J Biophys Lett.* 2006;35(7):559-66.

- [53] Rodriguez EJ, Angeles TS, Meek TD. Use of N-15 Kinetic Isotope Effects to Elucidate Details of the Chemical Mechanism of Human Immunodeficiency Virus-1 Protease. *Biochemistry*. 1993;32(46):12380-5.
- [54] Hyland LJ, Tomaszek TA, Meek TD. Human Immunodeficiency Virus-1 Protease .2. Use of Ph Rate Studies and Solvent Kinetic Isotope Effects to Elucidate Details of Chemical Mechanism. *Biochemistry*. 1991;30(34):8454-63.
- [55] Hyland LJ, Tomaszek TA, Roberts GD, Carr SA, Magaard VW, Bryan HL, et al. Human Immunodeficiency Virus-1 Protease .1. Initial Velocity Studies and Kinetic Characterization of Reaction Intermediates by O-18 Isotope Exchange. *Biochemistry*. 1991;30(34):8441-53.
- [56] Suguna K, Padlan EA, Smith CW, Carlson WD, Davies DR. Binding of a Reduced Peptide Inhibitor to the Aspartic Proteinase from *Rhizopus-Chinensis* - Implications for a Mechanism of Action. *Proc Natl Acad Sci U S A*. 1987;84(20):7009-13.
- [57] Coates L, Erskine PT, Wood SP, Myles DAA, Cooper JB. A neutron laue diffraction study of endothiapepsin: Implications for the aspartic proteinase mechanism. *Biochemistry*. 2001;40(44):13149-57.
- [58] Coates L, Erskine PT, Crump MP, Wood SP, Cooper JB. Five atomic resolution structures of endothiapepsin inhibitor complexes: Implications for the aspartic proteinase mechanism. *J Mol Biol*. 2002;318(5):1405-15.
- [59] Pearl LH. The Catalytic Mechanism of Aspartic Proteinases. *FEBS Lett*. 1987;214(1):8-12.
- [60] Rich DH, Bernatowicz MS, Agarwal NS, Kawai M, Salituro FG, Schmidt PG. Inhibition of Aspartic Proteases by Pepstatin and 3-Methylstatine Derivatives of Pepstatin - Evidence for Collected-Substrate Enzyme-Inhibition. *Biochemistry*. 1985;24(13):3165-73.
- [61] Hofmann T, Dunn BM, Fink AL. Cryoenzymology of Penicillopepsin - Appendix - Mechanism of Action of Aspartyl Proteinases. *Biochemistry*. 1984;23(22):5247-56.

- [62] Cosgrove S, Rogers L, Hewage CM, Malthouse JPG. NMR study of the inhibition of pepsin by glyoxal inhibitors: Mechanism of tetrahedral intermediate stabilization by the aspartyl proteases. *Biochemistry*. 2007;46(39):11205-15.
- [63] Dunn BM. Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chem Rev*. 2002;102(12):4431-58.
- [64] Kossiakoff AA. The Application of Neutron Crystallography to the Study of Dynamic and Hydration Properties of Proteins. *Annu Rev Biochem*. 1985;54:1195-227.
- [65] Niimura N. Neutrons expand the field of structural biology. *Current Opinion in Structural Biology*. 1999;9(5):602-8.
- [66] Cooper JB, Myles DAA. A preliminary neutron Laue diffraction study of the aspartic proteinase endothiapepsin. *Acta Crystallographica Section D-Biological Crystallography*. 2000;56:246-8.
- [67] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Res*. 2000;28(1):235-42.
- [68] Tuan HF, Erskine P, Langan P, Cooper J, Coates L. Preliminary neutron and ultrahigh-resolution X-ray diffraction studies of the aspartic proteinase endothiapepsin cocrystallized with a gem-diol inhibitor. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications*. 2007;63:1080-3.
- [69] Coates L, Tuan HF, Tomanicek S, Kovalevsky A, Mustyakimov M, Erskine P, et al. The catalytic mechanism of an aspartic proteinase explored with neutron and X-ray diffraction. *J Am Chem Soc*. 2008;130(23):7235-+.
- [70] Cleland WW, Frey PA, Gerlt JA. The low barrier hydrogen bond in enzymatic catalysis. *J Biol Chem*. 1998;273(40):25529-32.
- [71] Frey PA. Strong hydrogen bonding in molecules and enzymatic complexes. *Magnetic Resonance in Chemistry*. 2001;39:S190-S8.

- [72] Northrop DB. Follow the protons: A low-barrier hydrogen bond unifies the mechanisms of the aspartic proteases. *Accounts of Chemical Research*. 2001;34(10):790-7.
- [73] Erskine PT, Coates L, Mall S, Gill RS, Wood SP, Myles DAA, et al. Atomic resolution analysis of the catalytic site of an aspartic proteinase and an unexpected mode of binding by short peptides. *Protein Sci*. 2003;12(8):1741-9.
- [74] Rich DH, Bernatowicz MS. Synthesis of Analogs of the Carboxyl Protease Inhibitor Pepstatin - Effect of Structure in Subsite-P3 on Inhibition of Pepsin. *J Med Chem*. 1982;25(7):791-5.
- [75] Rich DH, Bernatowicz MS, Schmidt PG. Direct C-13 NMR Evidence for a Tetrahedral Intermediate in the Binding of a Pepstatin Analog to Porcine Pepsin. *J Am Chem Soc*. 1982;104(12):3535-6.
- [76] Schmidt PG, Bernatowicz MS, Rich DH. Spin-Labeled Pepstatin Binding to Pepsin - a Study by Electron-Spin Resonance and Nuclear Magnetic-Resonance. *Biochemistry*. 1982;21(8):1830-5.
- [77] Rajamani R, Reynolds CH. Modeling the protonation states of the catalytic aspartates in beta-secretase. *J Med Chem*. 2004;47(21):5159-66.
- [78] Williamson AL, Brindley PJ, Knox DP, Hotez PJ, Loukas A. Digestive proteases of blood-feeding nematodes. *Trends Parasitol*. 2003;19(9):417-23.
- [79] Caffrey CR, McKerrow JH, Salter JP, Sajid M. Blood 'n' guts: an update on schistosome digestive peptidases. *Trends Parasitol*. 2004;20(5):241-8.
- [80] Wong JYM, Harrop SA, Day SR, Brindley PJ. Schistosomes express two forms of cathepsin D. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology*. 1997;1338(2):156-60.
- [81] Delcroix M, Sajid M, Caffrey CR, Lim KC, Dvorak J, Hsieh I, et al. A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. *J Biol Chem*. 2006;281(51):39316-29.

- [82] Morales ME, Rinaldi G, Gobert GN, Kines KJ, Tort JF, Brindley PJ. RNA interference of *Schistosoma mansoni* cathepsin D, the apical enzyme of the hemoglobin proteolysis cascade. *Molecular and Biochemical Parasitology*. 2008;157(2):160-8.
- [83] Leung D, Abbenante G, Fairlie DP. Protease Inhibitors: Current Status and Future Prospects. *J Med Chem*. 2000;43(3):305-41.
- [84] Aoyagi T, Kunimoto S, Morishim.H, Takeuchi T, Umezawa H. Effect of Pepstatin on Acid Proteases. *J Antibiot*. 1971;24(10):687-&.
- [85] Marciniszyn J, Huang JS, Hartsuck JA, Tang J. Mechanism of intramolecular activation of pepsinogen - Evidence for an intermediate-delta and involvement of active-site of pepsin in intramolecular activation of pepsinogen. *J Biol Chem*. 1976;251(22):7095-102.
- [86] Rich DH, Sun E, Singh J. Synthesis of Dideoxy-Pepstatin - Mechanism of Inhibition of Porcine Pepsin. *Biochem Biophys Res Commun*. 1977;74(2):762-7.
- [87] Asojo OA, Gulnik SV, Afonina E, Yu B, Ellman JA, Haque TS, et al. Novel uncomplexed and complexed structures of plasmepsin II, an aspartic protease from *Plasmodium falciparum*. *J Mol Biol*. 2003;327(1):173-81.
- [88] Baldwin ET, Bhat TN, Gulnik S, Hosur MV, Sowder RC, Cachau RE, et al. CRYSTAL-STRUCTURES OF NATIVE AND INHIBITED FORMS OF HUMAN CATHEPSIN-D - IMPLICATIONS FOR LYSOSOMAL TARGETING AND DRUG DESIGN. *Proceedings of the National Academy of Sciences of the United States of America*. 1993;90(14):6796-800.
- [89] Bernstein NK, Cherney MM, Yowell CA, Dame JB, James MNG. Structural insights into the activation of *P. vivax* plasmepsin. *J Mol Biol*. 2003;329(3):505-24.
- [90] Fujinaga M, Chernaia MM, Tarasova NI, Mosimann SC, James MNG. Crystal-Structure of Human Pepsin and Its Complex with Pepstatin. *Protein Science*. 1995;4(5):960-72.

- [91] Bott R, Subramanian E, Davies DR. 3-Dimensional Structure of the Complex of the *Rhizopus-Chinensis* Carboxyl Proteinase and Pepstatin at 2.5-Å Resolution. *Biochemistry*. 1982;21(26):6956-62.
- [92] James MNG, Sielecki A, Salituro F, Rich DH, Hofmann T. Conformational Flexibility in the Active-Sites of Aspartyl Proteinases Revealed by a Pepstatin Fragment Binding to Penicillopepsin. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*. 1982;79(20):6137-41.
- [93] Silva FP, Ribeiro F, Katz N, Giovanni-De-Simone S. Exploring the subsite specificity of *Schistosoma mansoni* aspartyl hemoglobinase through comparative molecular modelling. *FEBS Lett*. 2002;514(2-3):141-8.
- [94] Cohen AF. Developing drug prototypes: pharmacology replaces safety and tolerability? *Nat Rev Drug Discov*. 2010;9(11):856-65.
- [95] Kortagere S, Ekins S. Troubleshooting computational methods in drug discovery. *J Pharmacol Toxicol Methods*. 2010;61(2):67-75.
- [96] Lombardino JG, Lowe JA. The role of the medicinal chemist in drug discovery - Then and now. *Nat Rev Drug Discov*. 2004;3(10):853-62.
- [97] Jorgensen WL. Efficient Drug Lead Discovery and Optimization. *Accounts Chem Res*. 2009;42(6):724-33.
- [98] Carvalho I, Pupo MT, Borges ADL, Bernardes LSC. Introdução à Modelagem Molecular de Fármacos no Curso Experimental de Química Farmacêutica. *Quím Nova*. 2003;26(3):428-38.
- [99] Abraham DJ. *Burger's Medicinal Chemistry and Drug Discovery*. 6 ed: John Wiley and Sons Inc. Publication 2003.
- [100] Teramoto R, Fukunishi H. Supervised consensus scoring for docking and virtual screening. *J Chem Inf Model*. 2007;47(2):526-34.
- [101] Wang RX, Lu YP, Wang SM. Comparative evaluation of 11 scoring functions for molecular docking. *J Med Chem*. 2003;46(12):2287-303.

- [102] Sant'Anna CMR. Glossário de Termos Usados no Planejamento de Fármacos (Recomendações da IUPAC para 1997). *Quím Nova*. 2002;25(3):505-12.
- [103] Leach A. *Molecular Modelling: Principles and Applications*. Harlow: Prentice Hall 2001.
- [104] Sant'Anna CMR. Métodos de modelagem molecular para estudo e planejamento de compostos bioativos: Uma introdução. *Rev Virtual Quim*. 2009;1(1):49-57.
- [105] McCammon JA, Harvey SC. *Dynamics of Proteins and Nucleic Acids* Cambridge: University Press 1987.
- [106] Rodrigues CR. Processos Modernos no Desenvolvimento de Fármacos. *Cadernos Temáticos de Química Nova na Escola*. 2001;3:43-9.
- [107] Cavasotto CN, Phatak SS. Homology modeling in drug discovery: current trends and applications. *Drug Discov Today*. 2009;14(13-14):676-83.
- [108] Nayeem A, Sitkoff D, Krystek S. A comparative study of available software for high-accuracy homology modeling: From sequence alignments to structural models. *Protein Sci*. 2006;15(4):808-24.
- [109] Santos Filho OA, Alencastro RB. Modelagem de Proteínas por Homologia. *Quím Nova*. 2003;26(2):253-9.
- [110] Apweiler R, Martin MJ, O'Donovan C, Magrane M, Alam-Faruque Y, Antunes R, et al. Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res*. 2011;39:D214-D9.
- [111] Rayan A. New tips for structure prediction by comparative modeling *Bioinformatics*. 2009;3(6):263-7.
- [112] Baker D, Sali A. Protein structure prediction and structural genomics. *Science*. 2001;294(5540):93-6.
- [113] Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*. 1997;18(15):2714-23.



- [114] Peitsch MC. Protein Modeling by E-Mail. *Bio-Technology*. 1995;13(7):658-60.
- [115] Sutcliffe MJ, Haneef I, Carney D, Blundell TL. Knowledge Based Modeling of Homologous Proteins .1. 3-Dimensional Frameworks Derived from the Simultaneous Superposition of Multiple Structures. *Protein Eng*. 1987;1(5):377-84.
- [116] Sali A, Blundell TL. Comparative Protein Modeling by Satisfaction of Spatial Restraints. *J Mol Biol*. 1993;234(3):779-815.
- [117] Vriend G. What If - a Molecular Modeling and Drug Design Program. *J Mol Graph*. 1990;8(1):52-&.
- [118] Laskowski RA, Macarthur MW, Moss DS, Thornton JM. PROCHECK - a Program to Check the Stereochemical Quality of Protein Structures. *J Appl Crystallogr*. 1993;26:283-91.
- [119] Silveira A. Análise por modelagem e dinâmica molecular da interação entre a integrina  $\alpha 6\beta 1$  e a laminina 111 humana. Petrópolis: Laboratório Nacional de Computação Científica, LNCC; 2007.
- [120] Pontius J, Richelle J, Wodak SJ. Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J Mol Biol*. 1996;264(1):121-36.
- [121] Eisenberg D, Luthy R, Bowie JU. VERIFY3D: Assessment of protein models with three-dimensional profiles. *Macromolecular Crystallography, Pt B*. San Diego: Academic Press Inc 1997:396-404.
- [122] Eswar N, John B, Mirkovic N, Fiser A, Ilyin VA, Pieper U, et al. Tools for comparative protein structure modeling and analysis. *Nucleic Acids Res*. 2003;31(13):3375-80.
- [123] Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomolec Struct*. 2000;29:291-325.
- [124] Aszodi A, Taylor WR. Homology modelling by distance geometry. *Fold Des*. 1996;1(5):325-34.

- [125] Brooijmans N, Kuntz ID. Molecular recognition and docking algorithms. *Annu Rev Biophys Biomolec Struct.* 2003;32:335-73.
- [126] Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat Rev Drug Discov.* 2004;3(11):935-49.
- [127] Halperin I, Ma BY, Wolfson H, Nussinov R. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins.* 2002;47(4):409-43.
- [128] Moitessier N, Englebienne P, Lee D, Lawandi J, Corbeil CR. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *Br J Pharmacol.* 2008;153:S7-S26.
- [129] Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. *J Mol Biol.* 1996;261(3):470-89.
- [130] Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry.* 1998;19(14):1639-62.
- [131] Liu M, Wang SM. MCDOCK: A Monte Carlo simulation approach to the molecular docking problem. *J Comput-Aided Mol Des.* 1999;13(5):435-51.
- [132] Wolf A, Zimmermann M, Hofmann-Apitius M. Alternative to consensus scoring - A new approach toward the qualitative combination of docking algorithms. *J Chem Inf Model.* 2007;47(3):1036-44.
- [133] Stahl M, Rarey M. Detailed analysis of scoring functions for virtual screening. *J Med Chem.* 2001;44(7):1035-42.
- [134] Magalhães CS, Barbosa HJC, Dardenne LE. Métodos de Docking Receptor-Ligante para o Desenho Racional de Compostos Bioativos. In: Morgon NH, Coutinho K, eds. *Métodos de Química Teórica e Modelagem Molecular.* São Paulo: Livraria da Física 2007:539.
- [135] Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP. Empirical scoring functions .1. The development of a fast empirical scoring function to estimate the

binding affinity of ligands in receptor complexes. *J Comput-Aided Mol Des.* 1997;11(5):425-45.

[136] Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. *Proteins.* 2003;52(4):609-23.

[137] Gohlke H, Hendlich M, Klebe G. Knowledge-based scoring function to predict protein-ligand interactions. *J Mol Biol.* 2000;295(2):337-56.

[138] Velec HFG, Gohlke H, Klebe G. DrugScore(CSD)-knowledge-based scoring function derived from small molecule crystal data with superior recognition rate of near-native ligand poses and better affinity prediction. *J Med Chem.* 2005;48(20):6296-303.

[139] Muegge I. PMF scoring revisited. *J Med Chem.* 2006;49(20):5895-902.

[140] Muegge I, Martin YC. A general and fast scoring function for protein-ligand interactions: A simplified potential approach. *J Med Chem.* 1999;42(5):791-804.

[141] Wang RX, Lai LH, Wang SM. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J Comput-Aided Mol Des.* 2002;16(1):11-26.

[142] Chandrashekar IR, Rao GS, Cowsik SM. Molecular Modeling of the Peptide Agonist-Binding Site in a Neurokinin-2 Receptor. *J Chem Inf Model.* 2009;49(7):1734-40.

[143] Huang SY, Zou X. Advances and Challenges in Protein-Ligand Docking. *Int J Mol Sci.* 2010;11(8):3016-34.

[144] Magalhães C. Algoritmos genéticos para o problema de docking proteína-ligante [Doutorado]. Petrópolis: Laboratório Nacional de Computação Científica; 2006.

[145] Solis FJ, Wets RJB. Minimization by random search techniques. *Mathematics of Operations Research.* 1981;6 (1):19-30.

[146] Malaquias NGL. Uso dos Algoritmos Genéticos para a Otimização de Rotas de Distribuição [Mestrado]. Uberlândia: Universidade Federal de Uberlândia; 2006.

- [147] Huey R, Morris GM, Olson J, Goodsell DS. A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry* 2007;28(6):1145-52.
- [148] Yusuf D, Davis AM, Kleywegt GJ, Schmitt S. An alternative method for the evaluation of docking performance: RSR vs RMSD. *Journal of Chemical Information and Modeling*. 2008;48(7):1411-22.
- [149] Moitessier N, Therrien E, Hanessian S. A Method for Induced-Fit Docking, Scoring, and Ranking of Flexible Ligands. Application to Peptidic and Pseudopeptidic  $\alpha$ -secretase (BACE 1) Inhibitors. *J Med Chem*. 2006;49(20):5885-94.
- [150] Moitessier N, Henry C, Maigret B, Chapleur Y. Combining Pharmacophore Search, Automated Docking, and Molecular Dynamics Simulations as a Novel Strategy for Flexible Docking. Proof of Concept: Docking of Arginine-Glycine-Aspartic Acid-like Compounds into the  $\alpha_V\beta_3$  Binding Site. *J Med Chem*. 2004;47(17):4178-87.
- [151] Madurga S, Belda I, Llorca X, Giralt E. Design of enhanced agonists through the use of a new virtual screening method: Application to peptides that bind class I major histocompatibility complex (MHC) molecules. *Protein Sci*. 2005;14(8):2069-79.
- [152] Osterberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS. Automated docking to multiple target structures: Incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins-Structure Function and Genetics*. 2002;46(1):34-40.
- [153] Adcock SA, McCammon JA. Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chem Rev*. 2006;106(5):1589-615.
- [154] Scheraga HA, Khalili M, Liwo A. Protein-folding dynamics: Overview of molecular simulation techniques. *Annu Rev Phys Chem*. 2007;58:57-83.
- [155] Vangunsteren WF, Mark AE. On the Interpretation of Biochemical Data by Molecular-Dynamics Computer-Simulation. *Eur J Biochem*. 1992;204(3):947-61.
- [156] McCammon JA, Gelin BR, Karplus M. Dynamics of Folded Proteins. *Nat Med*. 1977;267(5612):585-90.

- [157] Vangunsteren WF, Berendsen HJC. Computer-Simulation of Molecular-Dynamics - Methodology, Applications, and Perspectives in Chemistry. *Angew Chem-Int Edit Engl.* 1990;29(9):992-1023.
- [158] Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J Mol Model.* 2001;7(8):306-17.
- [159] Berendsen HJC, Vandespoel D, Vandrunen R. Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. *Computer Physics Communications.* 1995;91(1-3):43-56.
- [160] Christen M, Hünenberger PH, Bakowies D, Baron R, Bürgi R, Geerke DP, et al. The GROMOS Software for Biomolecular Simulation: GROMOS05. *J Comput Chem.* 2005;26(16):1719–51.
- [161] Frank M, Gutbrod P, Hassayoun C, von der Lieth CW. Dynamic molecules: molecular dynamics for everyone. An internet-based access to molecular dynamic simulations: basic concepts. *J Mol Model.* 2003;9(5):308-15.
- [162] Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, et al. Amber, a Package of Computer-Programs for Applying Molecular Mechanics, Normal-Mode Analysis, Molecular-Dynamics and Free-Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput Phys Commun.* 1995;91(1-3):1-41.
- [163] Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A SMOOTH PARTICLE MESH EWALD METHOD. *J Chem Phys.* 1995;103(19):8577-93.
- [164] Nosé S. A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys.* 1984;52(2):255-68.
- [165] Berendsen HJC, Postma JPM, van Gunsteren WF, Dinola A, Haak JR. Molecular-Dynamics with Coupling to an External Bath. *J Chem Phys.* 1984;81(8):3684-90.

- [166] Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J Comput Phys.* 1977;23(3):327-41.
- [167] Vangunsteren WF, Berendsen HJC. Algorithms for Macromolecular Dynamics and Constraint Dynamics. *J Mol Phys.* 1977;34(5):1311-27.
- [168] Hess B, Bekker H, Berendsen HJC, Fraaije J. LINCS: A linear constraint solver for molecular simulations. *J Comput Chem.* 1997;18(12):1463-72.
- [169] Scott WRP, van Gunsteren WF. The GROMOS software package for biomolecular simulations. In: Corongiu ECaG, ed. *Methods and Techniques in Computational Chemistry: METECC-95*. Cagliari: STEF 1995:397-434.
- [170] Scott WRP, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, et al. The GROMOS Biomolecular Simulation Program Package  
*J Phys Chem A.* 1999;103(19):3596-607.
- [171] Schüttelkopf AW, van Aalten DMF. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D-Biological Crystallography.* 2004;60(8):1355-63.
- [172] Batista PR, Wilter A, Durham E, Pascutti PG. Molecular dynamics simulations applied to the study of subtypes of HIV-1 protease common to Brazil, Africa, and Asia. *Cell Biochem Biophys.* 2006;44(3):395-404.
- [173] Hao GF, Yang GF, Zhan CG. Computational mutation scanning and drug resistance mechanisms of HIV-1 protease inhibitors. *J Phys Chem B.* 2010 114(29):9663-76.
- [174] Karthik S, Senapati S. Dynamic flaps in HIV-1 protease adopt unique ordering at different stages in the catalytic cycle. *Proteins.* 2011 79(6):1830-40.
- [175] Purohit R, Rajendran V, Sethumadhavan R. Studies on adaptability of binding residues and flap region of TMC-114 resistance HIV-1 protease mutants.  
*J Biomol Struct Dyn.* 2011 29(1):137-52.

- [176] Sadiq SK, De Fabritiis G. Explicit solvent dynamics and energetics of HIV-1 protease flap opening and closing. *Proteins*. 2010 78(14):2873-85.
- [177] Soares RO, Batista PR, Costa MG, Dardenne LE, Pascutti PG, Soares MA. Understanding the HIV-1 protease nelfinavir resistance mutation D30N in subtypes B and C through molecular dynamics simulations. *J Mol Graph Model*. 2010 29(2):137-47.
- [178] Sousa SF, Tamames B, Fernandes PA, Ramos MJ. Detailed atomistic analysis of the HIV-1 protease interface. *J Phys Chem B*. 2011;115(21):7045-57.
- [179] Mao B. Molecular-dynamics investigation of molecular flexibility in ligand binding. *Biochem J*. 1992;288:109-16.
- [180] Okoniewska M, Tanaka T, Yada RY. The pepsin residue glycine-76 contributes to active-site loop flexibility and participates in catalysis. *Biochem J*. 2000;349:169-77.
- [181] Gutiérrez-de-Terán H, Nervall M, Dunn BM, Clemente JC, Aqvist JA. Computational analysis of plasmepsin IV bound to an allophenylnorstatine inhibitor. *FEBS Lett*. 2006;580 (25):5910-6.
- [182] de Farias Silva N, Lameira J, Alves CN. Computational analysis of aspartic protease plasmepsin II complexed with EH58 inhibitor: a QM/MM MD study. *J Mol Model*. 2011 in press.
- [183] Emanuelsson O, Brunak S, von Heijne G, Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Protocols*. 2007;2(4):953-71.
- [184] Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389-402.
- [185] Gomes D. Estudo da Falcipaina-2 em Complexo com Ligantes por Modelagem e Dinâmica Molecular como Suporte ao Desenvolvimento Racional de Novos Fármacos Antimaláricos. Rio de Janeiro: Universidade Federal do Rio de Janeiro; 2006.

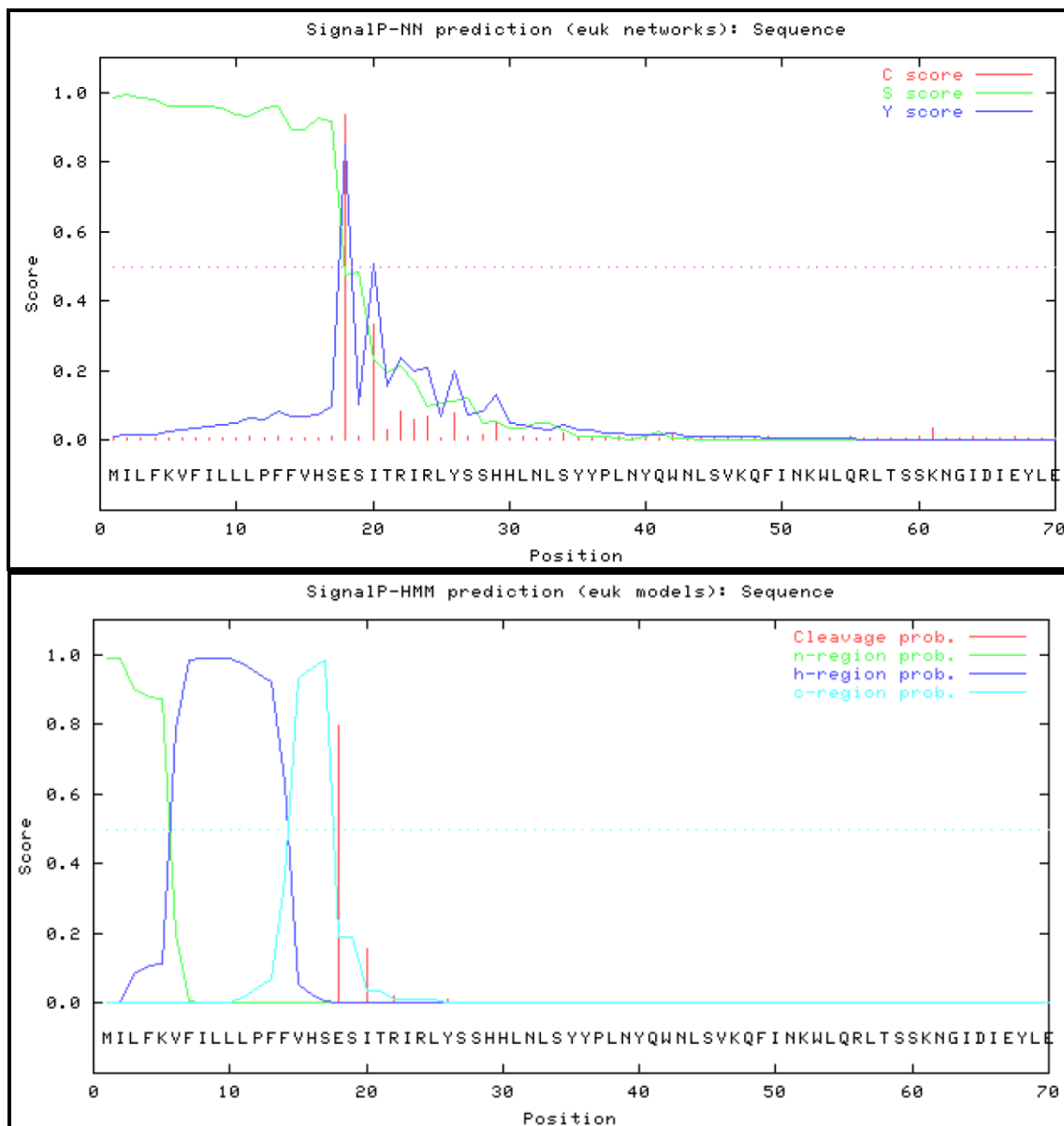
- [186] Santos Filho O, Alencastro R. Modelagem de proteínas por homologia. *Quím Nova*. 2003;26:253-9.
- [187] Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol*. 2000;302(1):205-17.
- [188] Clark M, Cramer RD, Vanopdenbosch N. Validation of the General-Purpose Tripos 5.2 Force-Field. *J Comput Chem*. 1989;10(8):982-1012.
- [189] Li H, Robertson AD, Jensen JH. Very fast empirical prediction and rationalization of protein pK(a) values. *Proteins*. 2005;61(4):704-21.
- [190] Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*. 2008;4(3):435-47.
- [191] vanAalten DMF, Bywater R, Findlay JBC, Hendlich M, Hooft RWW, Vriend G. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *Journal of Computer-Aided Molecular Design*. 1996;10(3):255-62.
- [192] Berweger CD, Vangunsteren WF, Mullerplathe F. Force-Field Parametrization by Weak-Coupling - Reengineering Spc Water. *Chem Phys Lett*. 1995;232(5-6):429-36.
- [193] Hockney RW, Goel SP, Eastwood JW. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics*. 1974;14(2):148-58.
- [194] Hockney RW. The potential calculation and some applications. . *Meth Comput Phys*. 1970;9:136-211.
- [195] Verli H. Interação Heparina-Antitrombina: Reconhecimento Molecular Caracterizado por Ferramentas de Modelagem Molecular. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2005.
- [196] Gustchina A, Weber IT. Comparison of Inhibitor Binding in Hiv-1 Protease and in Nonviral Aspartic Proteases - the Role of the Flap. *FEBS Lett*. 1990;269(1):269-72.



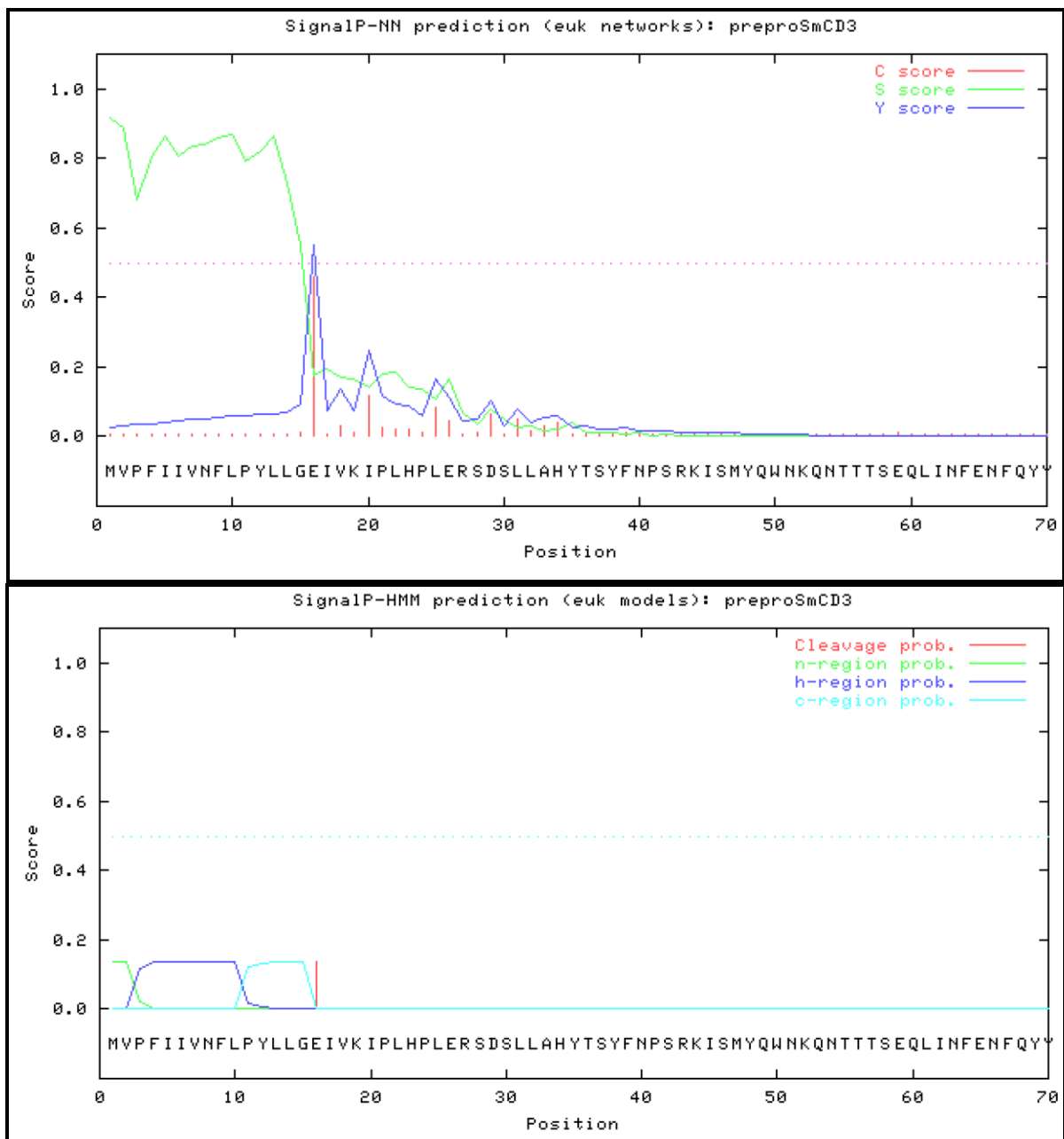
- [197] Edgar RC, Batzoglou S. Multiple sequence alignment. *Current Opinion in Structural Biology*. 2006;16(3):368-73.
- [198] Contreras-Moreira B, Fitzjohn PW, Bates PA. In silico protein recombination: Enhancing template and sequence alignment selection for comparative protein modelling. *J Mol Biol*. 2003;328(3):593-608.
- [199] Larsson P, Wallner B, Lindahl E, Elofsson A. Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Sci*. 2008;17(6):990-1002.
- [200] Kabsch W, Sander C. Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers*. 1983;22(12):2577-637.
- [201] Bond CS, Schüttelkopf AW. ALINE : a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallographica Section D-Biological Crystallography*. 2009;D65:510-2.
- [202] Luthy R, Bowie JU, Eisenberg D. Assessment of Protein Models with 3-Dimensional Profiles. *Nature*. 1992;356(6364):83-5.
- [203] Hong L, Tang J. Flap position of free memapsin 2 (beta-secretase), a model for flap opening in aspartic protease catalysis. *Biochemistry*. 2004;43(16):4689-95.
- [204] Cooper J, Foundling S, Hemmings A, Blundell T, Jones DM, Hallett A, et al. The Structure of a Synthetic Pepsin Inhibitor Complexed with Endothiapepsin. *Eur J Biochem*. 1987;169(1):215-21.
- [205] Dhanaraj V, Dealwis CG, Frazao C, Badasso M, Sibanda BL, Tickle IJ, et al. X-Ray Analyses of Peptide-Inhibitor Complexes Define the Structural Basis of Specificity for Human and Mouse Renins. *Nature*. 1992;357(6378):466-72.
- [206] James MNG, Sielecki AR. Structure and Refinement of Penicillopepsin at 1.8-Å Resolution. *J Mol Biol*. 1983;163(2):299-361.

- [207] Sielecki AR, Hayakawa K, Fujinaga M, Murphy MEP, Fraser M, Muir AK, et al. Structure of Recombinant Human Renin, a Target for Cardiovascular-Active Drugs, at 2.5 Å Resolution. *Science*. 1989;243(4896):1346-51.
- [208] Patel S, Vuillard L, Cleasby A, Murray CW, Yon J. Apo and inhibitor complex structures of BACE (beta-secretase). *J Mol Biol*. 2004;343(2):407-16.
- [209] Riboldi GP, Verli H, Frazzon J. Structural studies of the *Enterococcus faecalis* SufU [Fe-S] cluster protein. *BMC Biochemistry*. 2009;10:1-10.
- [210] Koradi R., Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph*. 1996;14:51-5.
- [211] Kumar S, Nussinov R. Close-range electrostatic interactions in proteins. *Chembiochem*. 2002;3(7):604-17.
- [212] Ingr M, Uhlikova T, Strisovsky K, Majerova E, Konvalinka J. Kinetics of the dimerization of retroviral proteases: The "fireman's grip" and dimerization. *Protein Sci*. 2003;12(10):2173-82.
- [213] Konvalinka J, Litera J, Weber J, Vondrasek J, Hradilek M, Soucek M, et al. Configurations of diastereomeric hydroxyethylene isosteres strongly affect biological activities of a series of specific inhibitors of human-immunodeficiency-virus proteinase. *Eur J Biochem*. 1997;250(2):559-66.
- [214] Rose JR, Babe LM, Craik CS. Defining the Level of Human-Immunodeficiency-Virus Type-1 (Hiv-1) Protease Activity Required for Hiv-1 Particle Maturation and Infectivity. *J Virol*. 1995;69(5):2751-8.
- [215] Freedberg DI, Ishima R, Jacob J, Wang YX, Kustanovich I, Louis JM, et al. Rapid structural fluctuations of the free HIV protease flaps in solution: Relationship to crystal structures and comparison with predictions of dynamics calculations. *Protein Sci*. 2002;11(2):221-32.
- [216] Dealwis CG, Frazao C, Badasso M, Cooper JB, Tickle IJ, Driessen H, et al. X-Ray-Analysis at 2.0 Å Resolution of Mouse Submaxillary Renin Complexed with a Decapeptide Inhibitor Ch-66, Based on the 4-16 Fragment of Rat Angiotensinogen. *J Mol Biol*. 1994;236(1):342-60.

## ANEXO I: PEPTÍDEO SINAL

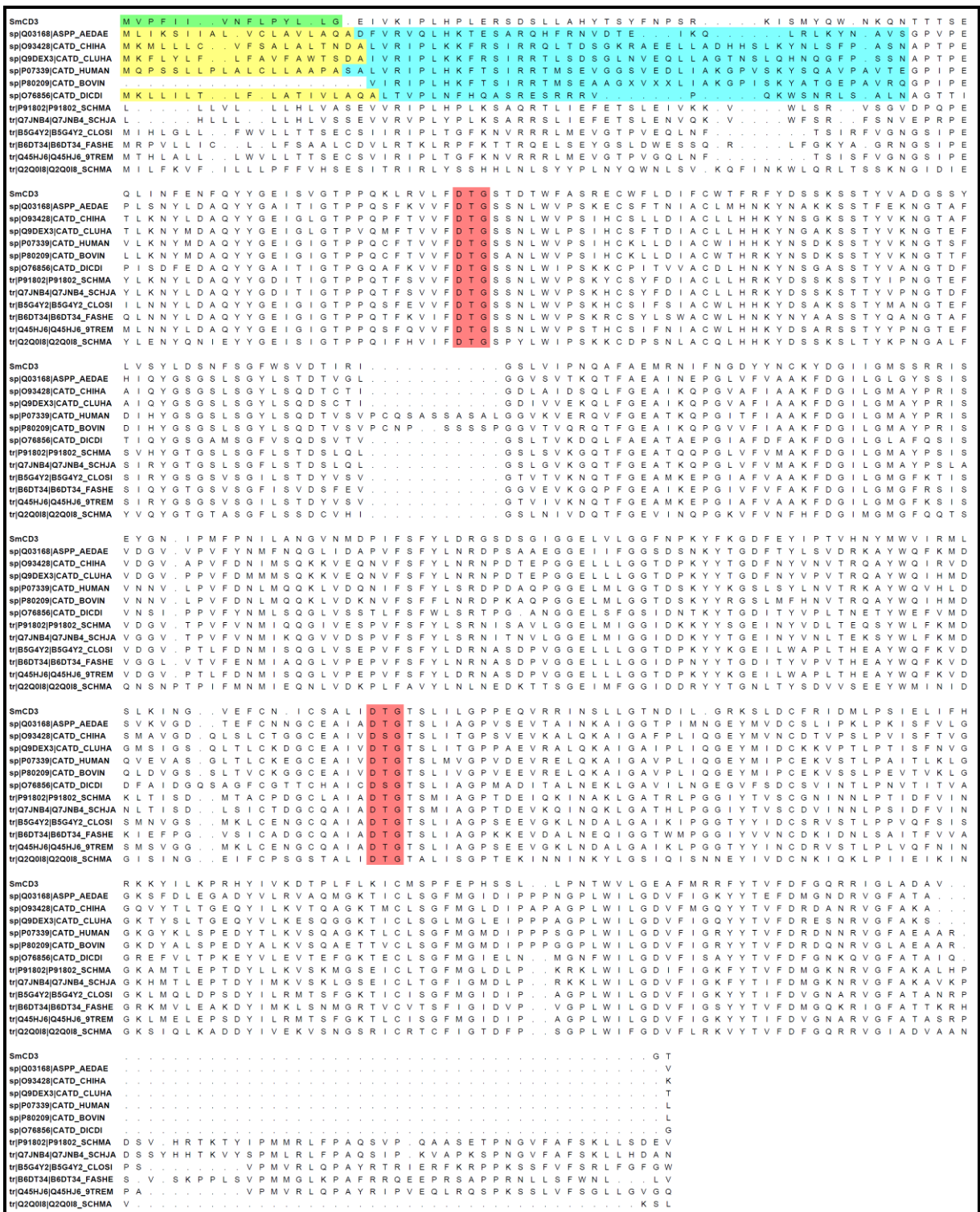


**FIGURA 1.** Análise do peptídeo sinal para a enzima SmCD2 realizado pelo programa Signal3P [183].



**FIGURA 2.** Análise do peptídeo sinal para a enzima SmCD3 realizado pelo programa Signal3P [183].





**FIGURA 2.** Alinhamento das seqüências da SmCD3 e de aspartil proteases depositadas no banco de dados Swiss-Prot [110].O pró-segmento foi eliminado baseado na similaridade entre as seqüências. As regiões em verde correspondem ao peptídeo sinal analisado pelo programa Signal3P [183]; em amarelo e azul aos peptídeos sinal e pró-peptídeo experimentalmente avaliados, respectivamente; em rosa à tríade catalítica.

**ANEXO III: TABELAS DE VALORES DE PKA**

**TABELA 1.** Resíduos com valores de pKa maiores que cinco na catepsina D humana (código PDB 1LYB), segundo a análise do programa PROPKA [189].

<b>Resíduo</b>	<b>pKa predito</b>	<b>pKa padrão</b>
<b>ASP231</b>	7,62	3,80
<b>ASP323</b>	7,35	3,80
<b>GLU5</b>	7,19	4,50
<b>GLU18</b>	5,06	4,50
<b>GLU180</b>	6,18	4,50
<b>GLU260</b>	5,31	4,50
<b>HIS45</b>	6,36	6,50
<b>HIS56</b>	6,28	6,50
<b>HIS57</b>	8,06	6,50
<b>HIS77</b>	6,50	6,50
<b>HIS209</b>	7,59	6,50

**TABELA 2.** Resíduos com valores de pKa maiores que cinco na renina de camundongo (código PDB 1SMR), segundo a análise do programa PROPKA [189].

<b>Resíduo</b>	<b>pKa predito</b>	<b>pKa padrão</b>
<b>ASP215</b>	9,83	3,80
<b>GLU164</b>	5,96	4,50
<b>HIS74</b>	6,50	6,50
<b>HIS139</b>	5,08	6,50
<b>HIS159</b>	6,50	6,50
<b>HIS174</b>	6,63	6,50
<b>HIS180</b>	6,91	6,50
<b>HIS243</b>	6,29	6,50
<b>HIS316</b>	6,50	6,50

**TABELA 3.** Resíduos com valores de pKa maiores que cinco na pepsina humana (código PDB 1PSO e 1QRP), segundo a análise do programa PROPKA .

<b>Resíduo</b>	<b>pKa predito</b>	<b>pKa padrão</b>
<b>ASP32</b>	9,38	3,80
<b>ASP96</b>	5,32	3,80
<b>ASP303</b>	6,65	3,80
<b>GLU13</b>	5,80	4,50
<b>GLU107</b>	5,52	4,50
<b>HIS53</b>	5,51	6,50

**TABELA 4.** Resíduos com valores de pKa maiores que cinco na SmCD1, segundo a análise do programa PROPKA [189].

<b>Resíduo</b>	<b>pKa predito</b>	<b>pKa padrão</b>
<b>ASP309</b>	7,29	3,80
<b>GLU5</b>	5,74	4,50
<b>GLU106</b>	5,17	4,50
<b>GLU169</b>	5,51	4,50
<b>HIS56</b>	6,12	6,50
<b>HIS77</b>	6,50	6,50

**TABELA 5.** Resíduos com valores de pKa maiores que cinco na SmCD2, segundo a análise do programa PROPKA [189].

<b>Resíduo</b>	<b>pKa predito</b>	<b>pKa padrão</b>
<b>Glu5</b>	5,40	4,50
<b>Glu14</b>	5,59	4,50
<b>His29</b>	5,41	6,50
<b>Asp33</b>	9,53	3,80
<b>His56</b>	6,44	6,50
<b>His57</b>	9,36	6,50
<b>His93</b>	6,34	6,50



**TABELA 6.** Resíduos com valores de pKa maiores que cinco na SmCD3, segundo a análise do programa PROPKA [189].

<b>Resíduo</b>	<b>pKa predito</b>	<b>pKa padrão</b>
<b>Glu106</b>	5,06	4,50
<b>His119</b>	6,36	6,50
<b>Glu170</b>	5,83	4,50
<b>Asp310</b>	6,98	3,80
<b>Glu5</b>	4,50	7,48
<b>Glu11</b>	4,50	5,11
<b>Asp38</b>	3,80	6,88
<b>Asp50</b>	3,80	6,71
<b>Glu106</b>	4,50	6,26
<b>Asp114</b>	3,80	5,22
<b>Asp121</b>	3,80	6,13
<b>Glu133</b>	4,50	5,19
<b>Glu169</b>	4,50	5,94
<b>Glu185</b>	4,50	5,34
<b>His191</b>	6,50	6,02
<b>Glu261</b>	4,50	6,04
<b>His265</b>	6,50	5,91
<b>Asp280</b>	3,80	5,74
<b>Glu293</b>	4,50	6,81
<b>His295</b>	6,50	5,22
<b>Glu307</b>	4,50	9,83

**ANEXO IV: SUBSÍTIOS DAS ENZIMAS ESTUDADAS**
**TABELA 1.** Substítios S4 a S4'das enzimas SmCD1, SmCD2 e SmCD3 que acomodam os resíduos dos ligantes a 4 Å de distância.

Enzima	Ligante												
	S4'	S3'	S2'	S1'	S1	S2	S3	S4					
SmCD1	Pepstatina	-	Tyr194 Leu299	Ser36 Val76 His77 Tyr78	Gly35 Ser36 Asp219	Asp33 Tyr78 Thr80 Ile123	Thr80 Gly221 Thr222 Ser223 Met295 Ile306	Gly221 Thr222 Ser223	Ser223 Ile248 Met295				
		LBPP1	Tyr194	Gly35 Ser36 Val76 His77 Tyr78 Ile131 Tyr194	Asp33 Asp219	Val31 Asp33 Thr80 Phe120 Leu123 Gly221	Gly79 Thr80 Gly221 Thr222 Met295 Leu297	Ala13 Gln14 Gly221 Thr222 Ser223 Asp309	Ala13 Ser223				
				LBPP37		Ile131 Arg302	His77 Tyr194	Ser36 Ser37 Asn38 His77 Tyr78 Ile131 Val133	Gly35 Asp219	Asp33 Tyr78 Thr80 Phe115 Phe120 Ile123 Gly221	Gly79 Thr80 Thr222 Met295 Leu297	Ala13 Gln14 Met117 Gly221 Ser223	Ser223 Met224 Met295
								LBPP38		Ile131 Ser132 Val133	His77 Tyr194	Ser36 Ser37 Asn38 Val76 His77 Tyr78 Ile131 Val133	Asp219
SmCD2	Pepstatina	-	Phe300	Ser36 Gln77 Tyr78 Gln133	Asp220	Asp33 Tyr78 Thr80 Ile123 Gly222	Thr223	Ile13 Thr80 Phe118 Gly222	Ile13 Ala224 Arg292				
		LBPP1	Gln130 Thr131 Ser132 Gln133	Gln77	Gly35 Ser36 Val76 Gln77 Tyr78 Gln133	Asp220 Thr223	Asp33 Tyr78 Thr80 Phe120 Ile123 Gly222	Thr80 Asp220 Thr223 Phe300 Ile307	Ile13 Glu14 Phe118 Phe120 Gly222	Ile13 Ala224 Leu225 Arg292			

(continuação da TABELA 1)

Enzima	Ligante	(continuação da TABELA 1)							
		S4'	S3'	S2'	S1'	S1	S2	S3	S4
SmCD2	LBPP37	Thr131 Ser132 Gln133	Gln77	Gly35 Ser36 Pro37 Tyr38 Gln77 Tyr78 Thr131 Gln133	Asp220	Asp33 Tyr78 Thr80 Phe115 Phe118 Phe120 Ile123	Gly79 Thr80 Asp220 Thr223 Phe300 Ile307	Ile13 Glu14 Thr80 Phe118 Gly222 Thr223 Ala224 Asp370	Ile13 Thr223 Ala224 Leu225 Arg292 Ile296
	LBPP38	Gln130 Thr131 Ser132 Gln133	Gln77	Gly35 Ser36 Val76 Gln77 Tyr78 Gln133	Asp220	Asp33 Tyr78 Thr80 Phe115 Phe118 Phe120 Ile123	Gly79 Thr80 Phe223 Phe300 Ile307	Ile13 Glu14 Ile31 Phe118 Phe120 Gly221	Ile13 Ala224 Leu225 Arg292
SmCD3	Pepstatina	-	Leu79 Ser297 Asn301	Ser36 Ser77 Tyr78	Asp33 Asp218	Leu31 Asp33 Tyr78 Tyr115 Ile123 Gly220	Leu79 Asp80 Thr221 Leu225 Val304	Phe13 Asp80 Tyr120 Gly220 Thr221 Ser222	Phe13 Leu223 Arg246 Glu293
	LBPP1	Pro300	Ser77	Gly35 Ser36 Val76 Ser77 Tyr78 Ile131	Gly35 Asp218	Leu31 Asp33 Tyr78 Asp80 Tyr115 Tyr120 Ile123 Gly220	Leu79 Asp80 Thr221 Leu225 His295	Phe13 Gln14 Leu31 Asp80 Tyr120 Gly220 Ser222	Phe13 Ser222 Leu223 Leu225 Glu293 His295
	LBPP37	Pro300	-	Gly35 Ser36 Val76 Ser77 Tyr78 Ile131	Gly35 Asp218	Leu31 Asp33 Tyr78 Asp80 Tyr115 Tyr120 Ile123 Gly220	Leu79 Asp80 Thr221 Leu225 His295	Phe13 Gln14 Leu31 Asp80 Tyr120 Gly220 Ser222	Phe13 Ser222 Leu223 Leu225 Glu293 His295
	LBPP38	Pro300	Asn301	Gly35 Ser36 Val76 Ser77 Tyr78 Ile131	Gly35 Asp218	Leu31 Asp33 Tyr78 Asp80 Tyr115 Tyr120 Ile123 Gly220	Leu79 Asp80 Thr221 Leu225 His295	Phe13 Gln14 Leu31 Asp80 Tyr120 Gly220 Ser222	Phe13 Ser222 Leu223 Leu225 Glu293 His295

**ANEXO V: TABELAS DE PONTES SALINAS****TABELA 1.** Ocorrência e distâncias médias das principais pontes salinas de SmCD1 e do complexos SmCD1/pepstatina.

Resíduos	SmCD1		SmCD1/pepstatina	
	Distância média (Å) ± DP	Ocorrência durante a simulação (%)	Distância média (Å) ± DP	Ocorrência durante a simulação (%)
Arg160-Asp277	4,5 ± 1,4	38,7	3,0 ± 1,0	84,9
Arg57-Asp121	2,6 ± 0,9	90,2	2,6 ± 0,9	88,6
Asp12-Arg160	2,6 ± 0,8	92,6	3,9 ± 1,5	67,9
Asp188-Lys198	4,7 ± 2,3	43,5	9,2 ± 2,6	4,7
Asp18-Arg57	4,8 ± 2,5	52,1	3,9 ± 1,8	65,9
Asp206-Lys238	4,2 ± 1,7	51,7	3,7 ± 1,3	68,4
Glu231-Lys303	3,6 ± 1,5	68,4	4,0 ± 1,7	57,9
Lys281-Glu288	5,8 ± 3,5	42,9	8,1 ± 2,4	7,1
Lys44-Asp60	3,0 ± 0,9	85,7	2,9 ± 0,8	90,6

**TABELA 2.** Ocorrência e distâncias médias das principais pontes salinas de SmCD2 e do complexos SmCD2/pepstatina.

Resíduos	SmCD2		SmCD2/pepstatina	
	Distância média (Å) ± DP	Ocorrência durante a simulação (%)	Distância média (Å) ± DP	Ocorrência durante a simulação (%)
Asp189-Arg326	3,0 ± 1,1	83,7	3,2 ± 1,5	78,9
His57-Asp121	1,8 ± 0,1	99,9	1,8 ± 0,2	99,7
Asp278-Arg314	4,0 ± 1,1	54,4	6,9 ± 1,4	5,1
Asp278-Lys315	2,8 ± 1,1	86,0	2,6 ± 0,7	94,7
Glu18-His29	2,9 ± 0,7	95,2	2,4 ± 0,9	94,2
Glu193-Arg325	3,9 ± 3,9	57,5	4,3 ± 1,0	42,9
Glu249-Arg292	3,7 ± 0,9	69,0	3,5 ± 1,3	70,8
Glu282-Arg289	4,0 ± 2,4	68,0	10,6 ± 3,2	3,1
Lys44-Asp60	3,0 ± 0,8	88,8	3,0 ± 0,8	88,0
His29-Asp121	3,7 ± 0,7	71,7	3,4 ± 0,7	80,7
His56-Asp121	3,8 ± 0,4	76,4	3,2 ± 0,7	90,8
Lys270-Asp331	3,2 ± 1,3	77,0	3,5 ± 2,1	76,9
Lys45-Glu106	2,7 ± 1,0	94,8	3,2 ± 1,8	85,1
Lys68-Asp101	3,2 ± 1,3	78,3	3,8 ± 1,8	64,8

**TABELA 3.** Ocorrência e distâncias médias das principais pontes salinas de SmCD3 e do complexos SmCD3/pepstatina.

Resíduos	SmCD3		SmCD3/pepstatina	
	Distância média (Å) ± DP	Ocorrência durante a simulação (%)	Distância média (Å) ± DP	Ocorrência durante a simulação (%)
<b>Arg232-Asp242</b>	3,3 ± 1,2	78,3	11,3 ± 2,,8	1,3
<b>Arg312-Asp328</b>	3,9 ± 1,6	58,4	5,9 ± 2,0	24,5
<b>Arg44-Asp60</b>	5,2 ± 1,9	32,5	4,0 ± 1,1	51,8
<b>Asp255-Arg274</b>	3,9 ± 1,9	59,5	8,5 ± 3,6	15,1
<b>Asp318-Arg323</b>	4,0 ± 1,4	49,6	3,2 ± 1,0	77,6
<b>Asp60-Lys63</b>	3,8 ± 1,2	58,1	5,3 ± 1,3	17,7
<b>Asp80-His295</b>	2,0 ± 0,8	97,6	4,1 ± 2,5	53,0
<b>Glu18-Arg29</b>	3,3 ± 0,9	78,1	2,8 ± 0,7	95,4
<b>Lys181-Asp328</b>	5,4 ± 2,3	36,2	7,2 ± 2,4	9,5
<b>Lys203-Glu208</b>	4,3 ± 1,1	48,6	38 ± 1,1	63,3
<b>Asp159-Arg311</b>	5,6 ± 0,8	1,2	37 ± 0,7	69,6

## ANEXO VI: TABELAS DE LIGAÇÕES DE HIDROGÊNIO

**TABELA 1.** Ligações de hidrogênio da região “fireman’s grip” nas enzimas livres e inibidas, calculadas durante as simulações por dinâmica molecular.

	Resíduo	Átomo Resíduo	Resíduo	Átomo Resíduo	Distância Média (Å) ± DP	Ocorrência (%) <sup>1</sup>
<b>SmCD1</b>	Thr34	O (cadeia lateral)	Ala218	O (cadeia principal)	2,9 ± 0,4	69,1
	Thr34	O (cadeia principal)	Thr220	N (cadeia principal)	4,0 ± 1,0	31,8
	Thr220	O (cadeia lateral)	Phe32	O (cadeia principal)	2,8 ± 0,3	84,8
	Thr220	O (cadeia principal)	Thr34	N (cadeia principal)	3,7 ± 0,8	43,8
<b>SmCD1/ pepstatina</b>	Thr34	O (cadeia lateral)	Ala218	O (cadeia principal)	2,8 ± 0,2	88,6
	Thr34	O (cadeia principal)	Thr220	N (cadeia principal)	4,3 ± 0,8	17,3
	Thr220	O (cadeia lateral)	Phe32	O (cadeia principal)	2,9 ± 0,4	58,3
	Thr220	O (cadeia principal)	Thr34	N (cadeia principal)	4,7 ± 1,3	17,5
<b>SmCD2</b>	Thr34	O (cadeia lateral)	Ile218	O (cadeia principal)	3,5 ± 0,6	7,7
	Thr34	O (cadeia principal)	Thr220	N (cadeia principal)	5,1 ± 1,0	3,4
	Thr220	O (cadeia lateral)	Phe32	O (cadeia principal)	2,7 ± 0,3	89,8
	Thr220	O (cadeia principal)	Thr34	N (cadeia principal)	3,6 ± 0,9	57,0
<b>SmCD2/ pepstatina</b>	Thr34	O (cadeia lateral)	Ile218	O (cadeia principal)	2,7 ± 0,2	93,9
	Thr34	O (cadeia principal)	Thr220	N (cadeia principal)	4,1 ± 0,7	19,6
	Thr220	O (cadeia lateral)	Phe32	O (cadeia principal)	2,9 ± 0,4	72,2
	Thr220	O (cadeia principal)	Thr34	N (cadeia principal)	4,9 ± 1,1	3,1

<sup>1</sup> Ocorrência calculada com distância máxima de 3,5 Å e um ângulo máximo de 30°, respectivamente, durante toda a simulação.

(continuação da TABELA 1)

	Resíduo	Átomo Resíduo	Resíduo	Átomo Resíduo	Distância Média (nm) ± DP	Ocorrência (%) <sup>1</sup>
<b>SmCD3</b>	Thr34	O (cadeia lateral)	Ile218	O (cadeia principal)	2,9 ± 0,2	87,3
	Thr34	O (cadeia principal)	Thr220	N (cadeia principal)	3,0 ± 0,2	82,7
	Thr220	O (cadeia lateral)	Phe32	O (cadeia principal)	2,8 ± 0,3	90,5
	Thr220	O (cadeia principal)	Thr34	N (cadeia principal)	3,1 ± 0,2	53,2
<b>SmCD3/pepstatina</b>	Thr34	O (cadeia lateral)	Ile218	O (cadeia principal)	2,7 ± 0,1	97,0
	Thr34	O (cadeia principal)	Thr220	N (cadeia principal)	3,1 ± 0,2	55,1
	Thr220	O (cadeia lateral)	Phe32	O (cadeia principal)	2,8 ± 0,2	87,4
	Thr220	O (cadeia principal)	Thr34	N (cadeia principal)	3,0 ± 0,2	79,4

<sup>1</sup> Ocorrência calculada com distância máxima de 3,5 Å e um ângulo máximo de 30°, respectivamente, durante toda a simulação.



**TABELA 2.** Ligações de hidrogênio entre a pepstatina e as enzimas, durante a DM.

Subsítio	Resíduo Proteína	Átomo Proteína	Resíduo Pepstatina	Átomo Pepstatina	Distância Média (Å) ± DP	Ocorrência (%) <sup>1</sup>	
SmCD1/ pepstatina	P1	Asp33	O (cadeia lateral)	Sta	OH	3,4 ± 0,04	11,8
			O (cadeia lateral)	Sta	OH	3,5 ± 0,06	22,1
	P3'	His77	N (cadeia lateral)	Sta	O (cadeia principal)	3,5 ± 0,11	51,8
			N (cadeia lateral)	Sta	O (cadeia principal)	3,5 ± 0,11	53,1
	P1'	Gly79	N (cadeia lateral)	Sta	O (cadeia principal)	2,8 ± 0,01	98
	P2	Thr80	N (cadeia principal)	Val	O (cadeia principal)	2,9 ± 0,02	92,6
			O (cadeia lateral)	Val	N (cadeia principal)	3,2 ± 0,03	63,8
	P4		O (cadeia lateral)	Iva	O (cadeia principal)	3,6 ± 0,12	55,3
	P2'	Tyr194	O (cadeia lateral)	Ala	O (cadeia principal)	3,5 ± 0,16	69,4
	P1	Asp219	O (cadeia lateral)	Sta	OH	4,0 ± 0,07	1,4
O (cadeia lateral)			Sta	OH	3,5 ± 0,07	6,1	
P1	Thr222	O (cadeia lateral)	Sta	N (cadeia principal)	3,2 ± 0,03	34,1	
		O (cadeia lateral)	Sta	O	3,3 ± 0,03	20,3	

<sup>1</sup> Ocorrência calculada com distância máxima de 3,5 Å e um ângulo máximo de 30°, respectivamente, durante toda a simulação.

(continuação da TABELA 2)

Subsítio	Resíduo Proteína	Átomo Proteína	Resíduo Pepstatina	Átomo Pepstatina	Distância Média (Å) ± DP	Ocorrência (%) <sup>1</sup>	
SmCD2/ pepstatina	P1	Asp33	O (cadeia lateral)	Sta	OH	4,0 ± 0,08	1,9
			O (cadeia lateral)	Sta	OH	4,9 ± 0,09	0,8
	P1'	Gly79	N (cadeia lateral)	Sta	O (cadeia principal)	2,8 ± 0,01	98,8
	P2	Thr80	O (cadeia lateral)	Val	O (cadeia principal)	2,9 ± 0,05	54,6
	P2'	Tyr195	O (cadeia lateral)	Ala	O (cadeia principal)	2,6 ± 0,02	98,3
	P1	Asp219	O (cadeia lateral)	Sta	OH	3,4 ± 0,04	16,3
O (cadeia lateral)			Sta	OH	3,3 ± 0,03	20,3	
SmCD3/ pepstatina	P1	Asp33	O (cadeia lateral)	Sta	OH	5,6 ± 0,10	0,3
			O (cadeia lateral)	Sta	OH	5,0 ± 0,10	0,8
	P3'	Ser77	O (cadeia principal)	Sta	N (cadeia principal)	3,3 ± 0,04	65,7
	P1'	Leu79	N (cadeia principal)	Sta	O (cadeia principal)	2,8 ± 0,02	98,7
	P2	Asp80	N (cadeia principal)	Val	O (cadeia principal)	3,8 ± 0,06	28,4
	P1	Asp219	O (cadeia lateral)	Sta	OH	3,5 ± 0,08	14,4
O (cadeia lateral)			Sta	OH	3,9 ± 0,06	5,0	

<sup>1</sup> Ocorrência calculada com distância máxima de 3,5 Å e um ângulo máximo de 30°, respectivamente, durante toda a simulação.