

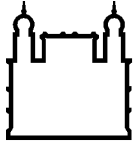
MINISTÉRIO DA SAÚDE
FUNDAÇÃO OSWALDO CRUZ
INSTITUTO OSWALDO CRUZ

Mestrado em Programa de Pós-Graduação Biologia Computacional e Sistemas

MODELAGEM MOLECULAR DE INIBIDORES DE ASPARTIL
PROTEASE: POTENCIAIS NOVOS COMPOSTOS ANTIMALARIAIS

AMANDA SUTTER DE OLIVEIRA HAMMES

Rio de Janeiro
Julho de 2012



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

AMANDA SUTTER DE OLIVEIRA HAMMES

Modelagem Molecular de Inibidores de Aspartil Protease: Potenciais Novos
Compostos Antimalariais

Dissertação apresentada ao Instituto Oswaldo
Cruz como parte dos requisitos para obtenção do
título de Mestre em Biologia Computacional e
Sistemas

Orientador (es): Prof. Dr. Ernesto Raúl Caffarena
Prof^a. Dr^a. Camila Silva de Magalhães

RIO DE JANEIRO

Julho de 2012

ii

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

H224 Hammes, Amanda Sutter de Oliveira

Modelagem molecular de inibidores de aspartil protease: potenciais novos compostos antimalariais/ Amanda Sutter de Oliveira Hammes. – Rio de Janeiro, 2012.

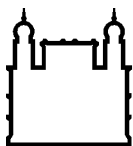
xviii, 106 f. : il. ; 30 cm.

Dissertação (Mestrado) – Instituto Oswaldo Cruz, Pós-Graduação em Biologia Computacional e Sistemas, 2012.

Bibliografia: f. 80-86

1. Malária. 2. Plasmeppsina. 3. Modelagem molecular. 4. Docking. 5. Dinâmica molecular. 6. LIE. I. Título

CDD 616.936206



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Computacional e Sistemas

AUTOR: AMANDA SUTTER DE OLIVEIRA HAMMES

MODELAGEM MOLECULAR DE INIBIDORES DE ASPARTIL PROTEASE: POTENCIAIS NOVOS COMPOSTOS ANTIMALARIAIS

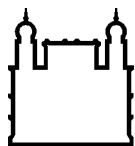
**ORIENTADORES: Prof. Dr. Ernesto Raúl Caffarena
Prof^a. Dr^a. Camila Silva de Magalhães**

Aprovada em: ____/____/____

EXAMINADORES:

**Prof. Dr. Floriano Paes Silva Junior – Presidente (Fiocruz/RJ)
Prof. Dr. Carlos Mauricio Rabello de Sant'Anna (UFRRJ)
Prof^a. Dr^a. Nelilma Correia Romeiro (UFRJ/Macaé)
Prof. Dr. Carlos Roberto Alves (Fiocruz/RJ) – Suplente e Revisor
Prof^a. Dr^a. Ana Carolina Rennó Soderó (UFRJ) – Suplente**

Rio de Janeiro, 23 de julho de 2012



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

Anexar a cópia da Ata que será entregue pela SEAC já assinada.

Agradecimentos

Aos meus pais, José Luis e Luciane, que sempre me deram a base e o incentivo para a continuidade dos meus estudos, por me apoiarem nos momentos difíceis e pela compreensão da minha ausência em vários momentos.

Às minhas irmãs, Camila, Grazi e Karol, pela amizade e suporte de sempre. Vocês são minhas amigas acima de tudo.

Ao Ernesto, meu grande orientador. Não tenho palavras para agradecer por tudo: a exemplar orientação, a confiança depositada, o constante entusiasmo e, sobretudo, a prestimosa amizade. Muito obrigada! :-). É um prazer ter trabalhado ao lado de um pesquisador sério, que está sempre disponível, preocupado não somente com as questões acadêmicas como também com o bem estar do aluno. Com certeza, isso fez toda a diferença.

À Camila Magalhães, pela orientação neste trabalho, muitas ajudas nas análises dos dockings, paciência, puxões de orelha, cobrança e, também pela amizade. Foi um grande aprendizado trabalhar com uma pessoa tão competente e dedicada à pesquisa. Obrigada por me receber de forma tão acolhedora desde o início do mestrado e pelo apoio constante.

À Dra. Cláudia Regina Brandão Gomes, por acreditar no trabalho disponibilizando as moléculas para os testes computacionais e também dando suporte sempre que foi preciso.

Aos amigos do CEDERJ, Demilson Castro, pelas caronas e pelos inúmeros conselhos dados, Michele Maria, pela animação e alegria de sempre, Ângelo Calvão (Salve Santa Clara!), Thiago Crispim, por fazer os sábados mais alegres implicando com todos a sua volta, Marcos Roboredo, por ser o grande incentivador da natação do polo, Dayanne Fernandes, pelas conversas durante os horários de coordenação, e, Marcelo e Marli, por sempre nos fazer rir com suas histórias. Gostaria de fazer um agradecimento especial à diretora, Eliane Viana, por entender minha ausência no final do mestrado e pela confiança depositada desde minha entrada como tutora no curso. A todos os outros tutores e funcionários do Cederj/Polo São Gonçalo, obrigada pela convivência. Vocês são a extensão da minha família.

Aos amigos do Laboratório de Biofísica e Modelagem Molecular do PROCC: Artur Brandt, Amanda Maia, Milene, Gilberto e aos alunos de IC: Joyce, Daniel, Eder e Vanessa. Aprendi muito na convivência com vocês. Desejo sorte e sucesso na vida acadêmica de cada um. Especialmente gostaria de agradecer ao jovem e recente doutor João Hermínio, vulgo Johnny, pelo companheirismo, apoio, dicas, brincadeiras e por fazer o clima do laboratório ser tão agradável. Espero que no futuro possamos trabalhar juntos e publicar muitos papers!! :-)

Aos amigos do PROCC, Ana, Elaine, Thais, Jéssica, Paulinho, Oswaldo, Marília, Aline Nobre, Claudinha e todos os outros, obrigada pela harmoniosa convivência durante meus 2 primeiros anos na Fiocruz. :-)

Aos amigos que tive a oportunidade de conhecer na Fiocruz e que fizeram as aulas serem mais agradáveis, as colaborações com trabalhos em grupo serem mais produtivas e pela união que nos tornou tão próximos que podemos ser chamados de família. É claro que não dá pra colocar o nome de todos aqui mas vou citar os que estiveram mais próximos ao longo deste trabalho: Pininha, Adriana Froes, Kele Belloze, Milene Jezuz, Bruno Manuel, Joana Lima, Bruno Gabriel, Artur Brandt, André Torres, Alberto Junior, Márcio Pavan, Lenilton Silveira, Cíntia Palu e Noemia

Lima. Não posso deixar de agradecer a Aninha Sodero pelas dicas com o PyMOL para a geração das figuras.

Ao meu best friend, Felipe Figueiredo, quem está sempre disposto a me ouvir, me dar conselhos sobre o universo masculino e por ser um grande incentivador para a escrita dessa dissertação. Foram muitos boosts regados ao café, apoio e aplicação do método pomodoro na reta final da escrita.

Às verdadeiras amizades que fiz ao longo desses 2 anos na cidade maravilhosa, Bel Melo e Claudinha Lima, pelos happy hours, shows de rock e cinema pra aliviar a tensão da pesquisa, Mariana Santanna, pela companhia online durante as noites dissertantes, Fernanda Maia, por ter se tornado praticamente uma irmã e Rafael Almada, um amigo que gostei imediatamente logo que conheci e por ser uma pessoa super agradável e uma ótima companhia para todos os programas culturais.

Ao companheiro de sala que mais entende de R, Luiz Max Carvalho, pela ajuda com o script para calcular a regressão linear.

À quem esteve presente durante alguns anos em minha vida, me apoiando e participando de todos os momentos de alegria e tristeza. Muito obrigada. Você foi um grande incentivador para o ingresso na pesquisa e no mundo acadêmico.

Ao Oswaldo, por sempre contar histórias e experiências vividas durante as noites no PROCC. Muitas vezes foi minha motivação. Obrigada.

Aos docentes deste programa de pós-graduação, só tenho a agradecer pela vivência que obtive nesses anos de experiência acadêmica e científica onde desenvolvi os créditos necessários para a conclusão da minha dissertação.

À PG-BCS pela ajuda de custo concedida para participar do Brazilian Symposium on Medicinal Chemistry, ocorrido em Novembro de 2010 na cidade de Ouro Preto – MG.

Às secretárias da PG, Márcia Verônica e Alessandra por amenizarem a burocracia dos processos internos à secretaria.

Ao comitê organizador dos Seminários Laveran & Deane sobre Malária, por me permitir participar na qualidade de ouvinte desta importante reunião entre pesquisadores altamente qualificados e alunos que desenvolvem teses em malária em todo o Brasil. Com certeza foi muito proveitoso estar presente nas discussões dos projetos em andamento com um objetivo em comum: erradicação da doença no Brasil. Gostaria de agradecer principalmente ao Prof. Dr. Cláudio Tadeu Daniel-Ribeiro pelo carinho, atenção, incentivo e por acreditar que os trabalhos computacionais também são de grande importância para o conhecimento das interações moleculares e auxílio aos resultados obtidos nos laboratórios.

Aos professores, Dr. Floriano Paes e Dr. Oswaldo Cruz, pela avaliação e sugestões dadas ao longo da minha apresentação de projeto no seminário discente.

Ao Dr. Floriano Paes, por aceitar fazer a revisão deste trabalho.

Aos membros da banca examinadora, por terem aceitado o convite da avaliação desse trabalho e pela certeza da ajuda crítica e construtiva do mesmo.

À CAPES, pelo investimento através da bolsa de estudos recebida.

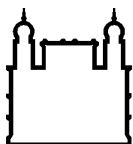
À todos aqueles que participaram direta ou indiretamente e que de alguma forma contribuíram para minha formação.

Provavelmente eu me esqueci de citar alguém, mas ficam aqui meus sinceros agradecimentos.

“É preciso recordar sempre que, por detrás de cada tabela, de cada relatório, ou de cada material de exame, existe vida, existe gente, existe sofrimento, à espera do nosso esforço e da nossa solidariedade humana”.

Carlyle Macedo

(Diretor da Organização Pan-Americana da Saúde)



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

MODELAGEM MOLECULAR DE INIBIDORES DE ASPARTIL PROTEASE: POTENCIAIS NOVOS COMPOSTOS ANTIMALARIAIS

RESUMO

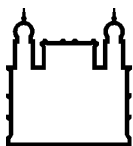
DISSERTAÇÃO DE MESTRADO EM BIOLOGIA COMPUTACIONAL E SISTEMAS

Amanda Sutter de Oliveira Hammes

Uma família de enzimas do tipo aspartil proteases, conhecida como Plasmepsinas, tem sido descrita como alvo atrativo para pesquisa e desenvolvimento de novos compostos terapêuticos para o tratamento da malária. Isto se deve ao fato de que no vacúolo alimentar do parasita existem quatro Plasmepsinas ativas e a inibição destas impede que o parasita degrade a hemoglobina, sua fonte de nutrientes para crescimento e maturação. Sabendo-se que os parasitas *Plasmodium falciparum* spp. adquirem uma rápida resistência aos atuais fármacos antimalariais, se faz necessário que novos e mais efetivos fármacos sejam descobertos para tratamento desta doença. A área de planejamento de novos fármacos baseado em estrutura (Structure Based Drug Design - SBDD) é considerada estratégica pois se baseia em uma maior compreensão dos mecanismos de reconhecimento molecular através da utilização de métodos computacionais.

O objetivo deste trabalho foi estudar, através da identificação dos modos de ligação e da determinação da energia livre de ligação, uma nova série de compostos planejados pelo Departamento de Síntese Orgânica de Farmanguinhos-Fiocruz. Métodos de *docking* e dinâmica molecular foram combinados e mostraram vantagens na utilização conjunta dessas técnicas apresentando uma boa correspondência entre o valor de energia livre de ligação, calculado através do método LIE, e o valor da afinidade de ligação obtida experimentalmente para os compostos estudados. Os resultados foram satisfatórios pois mostraram que os métodos usados no estudo de interações receptor-ligante envolvendo moléculas da família de aspartil proteases são interessantes na determinação dos plausíveis modos de ligação dos protótipos no sítio de ligação do alvo molecular.

Palavras Chave: Malária, Plasmepsina, Modelagem molecular, *docking*, dinâmica molecular, LIE.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

MOLECULAR MODELING OF ASPARTYL PROTEASE INHIBITORS: NEW POTENTIAL ANTIMALARIAL COMPOUNDS

ABSTRACT

MASTER DISSERTATION IN COMPUTATIONAL SYSTEMS BIOLOGY

Amanda Sutter de Oliveira Hammes

A family of aspartyl proteases enzyme, known as Plasmepsins, has been described as attractive target for research and development of new therapeutic compounds for treatment of malaria. This is because within the digestive vacuole of the parasite there exist four active Plasmepsins whose inhibition prevents these parasites to degrade the hemoglobin which is the source of nutrients for their growth and maturation. Bearing in mind that the parasites of the genre *Plasmodium falciparum* spp. acquire a rapid resistance to the antimalarial drugs currently on the market, it is necessary that new and more effective compounds are discovered to treat the disease. The field of Structure Based Drug Design – SBDD is nowadays considered crucial because it relies on the profound understanding of the mechanisms of molecular recognition through the use of computational methods.

Thus, the goal of this study was to identify the binding modes and determining the free energy of binding of a new series of compounds planned by the Department of Organic Synthesis of Farmanguinhos – FIOCRUZ. The joint application of docking and molecular dynamics methodologies showed advantages in using these techniques together. The results presented a good correspondence between the calculated free energy values using the Linear Interaction Energy – LIE method of compounds and their experimental values. The results may be considered satisfactory in the context of this job and show that the approach here applied to study receptor-ligand interactions involving molecules of the aspartyl proteases family was adequate in determining the plausible binding modes of prototypes molecules in the target binding site.

Keywords: Malaria, Plasmepsin, Molecular modelling, docking, Molecular dynamics, LIE.

ÍNDICE

1 INTRODUÇÃO	1
1.1 MALÁRIA	2
1.1.1 <i>O parasita e seu ciclo biológico</i>	4
1.1.2 <i>Fármacos em uso</i>	7
1.1.3 <i>Mecanismo de ação enzimática</i>	8
1.1.4 <i>Proteases Aspárticas</i>	10
1.1.4.1 <i>Plasmepsinas</i>	10
1.1.5 <i>Novos inibidores planejados por Farmanguinhos</i>	15
1.2 PLANEJAMENTO RACIONAL DE FÁRMACOS BASEADO EM ESTRUTURA.....	17
1.3 INTERAÇÕES RECEPTOR LIGANTE	18
1.4 CAMPOS DE FORÇA CLÁSSICOS	21
1.5 MÉTODOS DE ATRACAMENTO MOLECULAR (<i>DOCKING</i>)	22
1.5.1 <i>Programas de Atracamento Molecular</i>	24
1.6 MÉTODOS PARA CÁLCULO DA ENERGIA LIVRE DE LIGAÇÃO E ENERGIA DE INTERAÇÃO LINEAR (LIE)	26
1.7 REGRESSÃO LINEAR.....	29
2 OBJETIVOS	30
2.1 OBJETIVO GERAL	30
2.2 OBJETIVOS ESPECÍFICOS	30
3 MATERIAIS E MÉTODOS	31
3.1 MÉTODOS	31
3.1.1 <i>Preparação do Alvo Molecular (receptor)</i>	31
3.1.2 <i>Preparação dos Ligantes</i>	33
3.1.3 <i>Redocking</i>	34
3.1.4 <i>Dinâmica Molecular aplicada ao Cálculo de Energia Livre de Ligação</i>	37
3.1.4.1 <i>Preparação do receptor</i>	38
3.1.4.2 <i>Preparação dos ligantes</i>	40
3.1.5 <i>Métodos para Análise dos Resultados</i>	41
3.1.5.1 <i>RMSD</i>	41

3.1.5.2 Seleção das Conformações.....	41
3.1.5.3 Taxa de Sucesso	42
4 RESULTADOS E DISCUSSÃO	43
4.1 REDOCKING DO LIGANTE EH58	43
4.1.1 <i>Calibração dos parâmetros utilizando os programas Autodock 4 e Autodock Vina.....</i>	<i>43</i>
4.1.2 <i>Modelos de Carga e Protonação do resíduo Asp214.....</i>	<i>51</i>
4.1.3 <i>Utilização do programa Dockthor.....</i>	<i>60</i>
4.2 DOCKING DE NOVOS INIBIDORES DE ASPARTIL PROTEASES	65
4.3 CÁLCULO DA ENERGIA LIVRE	69
4.3.1 <i>Simulação entre a enzima PlmII e seus respectivos ligantes.....</i>	<i>70</i>
4.3.2 <i>Resultado da Regressão Linear para complexos PlmII do PDB.....</i>	<i>71</i>
4.3.3 <i>Compostos Farmanguinhos.....</i>	<i>73</i>
4.3.4 <i>Resultado da Regressão Linear para os compostos de Farmanguinhos... </i>	<i>73</i>
4.4 RANQUEAMENTO DA SÉRIE PLIP	75
4.4.1 <i>Análise da ligação hidrogênio entre os ligantes Plip e os resíduos aspárticos.....</i>	<i>76</i>
5 CONCLUSÕES.....	78
6 REFERÊNCIAS	80
7 ANEXO I.....	87
7.1 PARÂMETROS DA DINÂMICA MOLECULAR UTILIZADOS NO PROGRAMA Q.....	87
7.1.1 <i>Equilíbrio do Complexo</i>	<i>87</i>
7.1.2 <i>Dinâmica Molecular do Complexo</i>	<i>87</i>
7.1.3 <i>Dinâmica Molecular da Solução</i>	<i>88</i>

ÍNDICE DE FIGURAS

Figura 1. Áreas com transmissão da malária no mundo, 2008. Fonte: OMS	3
Figura 2. Incidência Parasitária Anual (IPA), Amazônia Legal, 2007.	4
Figura 3. Representação esquemática do ciclo biológico do Plasmodium spp.	6
Figura 4. Estrutura química de fármacos que atuam no metabolismo do heme utilizados no tratamento da malária.....	8
Figura 5. Alinhamento entre as quatro sequências de plasmepsinas utilizando a ferramenta ClustalW.....	12
Figura 6. Grau de identidade entre as sequências de plasmepsinas utilizando o servidor online blastp e utilizando a proteína Plm II como referência.....	12
Figura 7. Ligante EH58 (PDB id: 1LF3) com todas as possíveis torções	13
Figura 8. Mapa de interações entre o ligante EH58 e a enzima Plasmepsina II.....	14
Figura 9. Representação espacial das quatro estruturas de plasmepsinas indicando através do círculo que todas compartilham o mesmo sítio de ligação..	14
Figura 10. Estrutura química da molécula Plip01	16
Figura 11. Estrutura química das moléculas Plip03, Plip11, Plip27 e Plip41	16
Figura 12. Estrutura química das moléculas Plip04, Plip13, Plip31 e Plip47.....	16
<i>Figura 13. Etapas do desenho racional de fármacos baseado em estrutura Fonte: (31).....</i>	<i>17</i>
Figura 14. Distância entre os átomos doador-aceitador da ligação hidrogênio	20
Figura 15. Aspartatos catalíticos com protonação no Asp214/O1	32
Figura 16. Representação da grade calculada pelo programa Autogrid.	36
Figura 17. RMSD ao longo do tempo na simulação do complexo de código 1LF3 ...	38
Figura 18. Modelo representativo do complexo construído para os cálculos de energia livre.....	39
Figura 19. Interação do ligante EH58 com os resíduos Asp34 e Asp214..	42
Figura 20. Sobreposição da estrutura experimental do ligante EH58 e a solução de menor energia..	45

Figura 21. Sobreposição da estrutura experimental do ligante EH58 com a solução de menor energia obtida após filtro das conformações.....	47
Figura 22. Sobreposição da estrutura experimental do ligante EH58 com as 10 soluções de menor RMSD.....	49
Figura 23. Ligação hidrogênio entre o hidrogênio da hidroxila do carbono central do ligante e o oxigênio do resíduo Asp214.	50
Figura 24. Sobreposição da solução de menor energia da configuração C1 – Carga Ptn: Gromos96; Carga Lig: MMFF94, C2 – Carga Ptn: Gasteiger; Carga Lig: Gasteiger, C3 – Carga Ptn: Gasteiger; Carga Lig: MMFF94 e C4 – Carga Ptn: MMFF94; Carga Lig: MMFF94, e estrutura experimental do ligante EH58.	52
Figura 25. Sobreposição das conformações que interagem com os resíduos Asp34 e Asp214 e o ligante EH58.....	53
Figura 26. Solução de menor energia da configuração C1, C2, C3 e C4 com a estrutura experimental do ligante EH58.	55
Figura 27. Solução de menor energia da configuração C2 com resíduo Asp214 não protonado e com protonação neste aminoácido.....	56
Figura 28. Solução de menor RMSD da configuração que apresenta o resíduo Asp214 não protonado e com protonação neste resíduo.....	58
Figura 29. Solução de menor energia e solução de menor RMSD da configuração que apresenta o resíduo Asp214 não protonado e com protonação neste resíduo sobrepostos a estrutura tridimensional do ligante EH58 obtido do PDB	59
Figura 30. A estrutura representada em verde é a estrutura obtida do experimento sem protonação no resíduo Asp214 e a estrutura em azul é a estrutura obtida do experimento com Asp214 carregado.....	62
Figura 31. Solução de menor energia de interação do experimento sem o resíduo Asp214 protonado e com protonação, sobrepostos com o ligante EH58.....	63
Figura 32. Solução de menor RMSD do experimento sem o resíduo Asp214 protonado e com protonação, sobrepostos com o ligante EH58. Ambas soluções interagem com os resíduos aspárticos do sítio catalítico	64
Figura 33. Taxa de sucesso obtida por cada programa de docking.	65
Figura 34. Solução de menor energia total sobreposta com o ligante EH58 obtido do PDB. As moléculas amostradas são Plip01, Plip03, Plip04, Plip11.....	67
Figura 35. Solução de menor energia total sobreposta com o ligante EH58 obtido do PDB. As moléculas amostradas são Plip13, Plip27, Plip31, Plip41 e Plip47.....	68

Figura 36. Grupamentos do ligante EH58.	69
Figura 37. Comparação entre ΔG experimental e ΔG calculado da energia livre de ligação dos novos ligantes planejados pelo DSO e dos ligantes complexados com a enzima Plasmepsina II no primeiro caso da regressão linear.	74
Figura 38. Potenciais eletrostático (ΔU ele) e Lennard-Jones (ΔU vdw) dos ligantes planejados pelo DSO e dos ligantes complexados com a enzima Plasmepsina II....	75

LISTA DE TABELAS

Tabela 1. Complexos de Plasmeptinas disponíveis no PDB.	11
Tabela 2. Compostos planejados pelo Instituto de Fármacos – Farmanguinhos	15
Tabela 3. Valores dos parâmetros usados nos experimentos de redocking com Autodock 4	35
Tabela 4. Valores dos parâmetros usados nos experimentos de redocking com Autodock Vina.	36
Tabela 5. Parâmetros usados nos experimentos de redocking com o programa Dockthor.	36
Tabela 6. Configuração do experimento em relação ao tipo de carga utilizados em redocking com Autodock 4.	37
Tabela 7. Redocking do ligante EH58 com o programa Autodock 4 com $1,5 \times 10^6$ de avaliações de energia.	44
Tabela 8. Redocking do ligante EH58 com o programa Autodock 4 com $2,5 \times 10^6$ de avaliações de energia.	46
Tabela 9. Redocking do ligante EH58 utilizando o programa Autodock Vina.	48
Tabela 10. Redocking utilizando diferentes configurações de campos de força com o programa Autodock 4 sem a protonação no resíduo Asp214.	51
Tabela 11. Redocking utilizando diferentes configurações de campos de força com o programa Autodock 4 com protonação no resíduo Asp214/O1.	54
Tabela 12. Redocking utilizando o programa Autodock 4 com a nova configuração de parâmetros.	57
Tabela 13. Redocking utilizando o programa Autodock Vina com o resíduo Asp214/O1 do receptor protonado e sem protonação com a melhor configuração de parâmetros obtido.	59
Tabela 14. Redocking utilizando o programa Dockthor com o resíduo Asp214/O1 do receptor protonado e sem protonação.	61
Tabela 15. Análise da estrutura de menor energia de interação utilizando o programa Dockthor com o resíduo Asp214/O1 do receptor protonado e sem protonação.	63
Tabela 16. Análise da estrutura de menor RMSD utilizando o programa Dockthor com o resíduo Asp214/O1 do receptor protonado e sem protonação.	64

Tabela 17. Resultado do docking das moléculas de Farmanguinhos utilizando o programa Dockthor.....	66
Tabela 18. Resultado do cálculo de energia livre de 4 complexos de PlmII do PDB.	70
Tabela 19. Resultado dos valores de gamma obtidos da regressão linear	72
Tabela 20. Resultado do ΔG calculado para os complexos da enzima PlmII do PDB com os valores obtidos da regressão linear.	72
Tabela 21. Resultado do cálculo de energia livre dos ligantes de Farmanguinhos. ...	73
Tabela 22. Resultado do ΔG calculado para a série Plip com os valores obtidos da regressão linear.....	74
Tabela 23. Energia livre de ligação calculada com o primeiro caso da regressão linear e o resultado experimental dos ligantes planejados pelo DSO – Farmanguinhos.	76
Tabela 24. Porcentagem da ligação hidrogênio dos ligantes planejados pelo DSO – Farmanguinhos e os resíduos aspárticos Asp34/O1, Asp34/O2, Asp214/O1 e Asp214/O2 durante a simulação de dinâmica molecular.	77

LISTA DE SIGLAS E ABREVIATURAS

ATB – Construtor de Topologia Automática, na sigla em inglês

DNDi – iniciativa Medicamentos para Doenças Negligenciadas, na sigla em inglês

DSO – Departamento de Síntese Orgânica

FEP – Perturbação da Energia Livre, na sigla em inglês

GA – Algoritmo Genético, na sigla em inglês

IC₅₀ – Concentração Inibitória de 50% da atividade

K_i – Constante de inibição

LGA – Algoritmo Genético Lamarckiano, na sigla em inglês

LIE – Energia de Interação Linear, na sigla em inglês

MMFF94 – Merck Molecular Force Field

OMS – Organização Mundial da Saúde

PDB – Banco de Dados de Proteína, na sigla em inglês

Plm – Plasmepsina

QSAR –Relação quantitativa estrutura-atividade, na sigla em inglês

RMN – Ressonância Magnética Nuclear

RMSD – Raiz Quadrada do Desvio Quadrático Médio das Distâncias, na sigla em inglês

SBDD – Planejamento Racional de Fármacos Baseado em Estruturas, na sigla em inglês

1 Introdução

Embora exista atualmente um grande interesse na caracterização de sistemas receptor-ligante, pouco se sabe sobre os mecanismos de ação destes sistemas biológicos. A área de pesquisa e desenvolvimento de fármacos é hoje um avanço na tecnologia pois reduz os gastos com medicamentos importados e possibilita um planejamento mais rápido de medicamentos para diversas doenças, dentre elas as doenças negligenciadas que afetam quase um sexto da população mundial e provocam a morte de aproximadamente três mil pessoas a cada dia no mundo. (1)

Um agravante dessa situação consiste em que a indústria farmacêutica acredita que investir em pesquisas sobre essas doenças não resulta em aproveitamento financeiro e por tal motivo, pesquisas acerca de fármacos menos agressivos e com ações mais eficazes recebem apenas 5% dos recursos globais para pesquisa e desenvolvimento (2). A situação se complica mais ainda se considerarmos que o desenvolvimento de um fármaco através da metodologia tradicional chamada de *screening* cego é um processo dispendioso. Os esforços em purificar, caracterizar e sintetizar princípios ativos, além da necessidade de se ter uma grande diversidade de material biológico para testes, tornam o processo inteiro trabalhoso, ineficiente, cujo custo envolve investimentos de bilhões de dólares por parte das grandes companhias farmacêuticas mundiais. Porém essa situação tem mudado nas últimas décadas, devido ao desenvolvimento do poder computacional (velocidade e capacidade de processamento) e novas tecnologias aplicadas à computação. Uma abordagem mais racional, denominada desenho racional de fármacos (do inglês, *rational drug design*), que visa à identificação e a uma maior compreensão das interações moleculares entre receptor e ligante, envolve a utilização de métodos computacionais baseados nas estruturas tridimensionais de moléculas interagentes. Entre elas se destaca a metodologia de *docking* molecular. Com esta, trata-se de determinar a geometria de complexos receptor-ligante a partir da estrutura do receptor livre (ou em complexo com uma outra molécula) e a posterior obtenção de uma estimativa da afinidade de ligação entre o receptor e o ligante. Essa abordagem é amplamente usada na descoberta de compostos protótipos obtidos a partir de uma varredura em bancos de dados de estruturas moleculares tridimensionais sendo o custo computacional gasto com

cada simulação extremamente reduzido se comparado com os testes *in vitro*.

Em geral, os programas disponíveis no âmbito científico são de ampla abrangência e podem ser utilizados para a maioria dos sistemas biológicos, porém em alguns casos o sucesso do experimento computacional é fortemente dependente das características específicas do sistema em estudo. Neste último caso é necessária uma análise exaustiva dos parâmetros de controle dos programas para se obter resultados confiáveis.

Com a disponibilidade das estruturas tridimensionais de receptores específicos para malária, obtidas no banco de dados de proteínas (PDB) (3), será possível uma aplicação de desenho racional de fármacos para a obtenção de novos compostos específicos. A colaboração com o Departamento de Síntese Orgânica (DSO) de Farmanguinhos abre perspectivas promissoras para a identificação de novos compostos protótipos que possam ser utilizados como base para o desenvolvimento de novos fármacos para o tratamento quimioterápico de doenças.

1.1 Malária

A malária é uma doença infecciosa presente nas regiões tropicais e subtropicais de todo o mundo. Segundo a iniciativa Medicamentos para Doenças Negligenciadas (do inglês, *Drugs for Neglected Diseases initiative* - DNDi), a malária causa mais de um milhão de mortes por ano e cerca de 91% das mortes ocorrem no continente africano, sendo a maioria delas em crianças com menos de 5 anos de idade (4). Quarenta por cento da população mundial vive em áreas de transmissão e é exposta ao risco de contrair a doença, que é de difícil controle devido a vários fatores, como a resistência dos parasitas aos fármacos e dos vetores aos inseticidas (5). A Figura 1 mostra a distribuição da doença no cenário mundial.

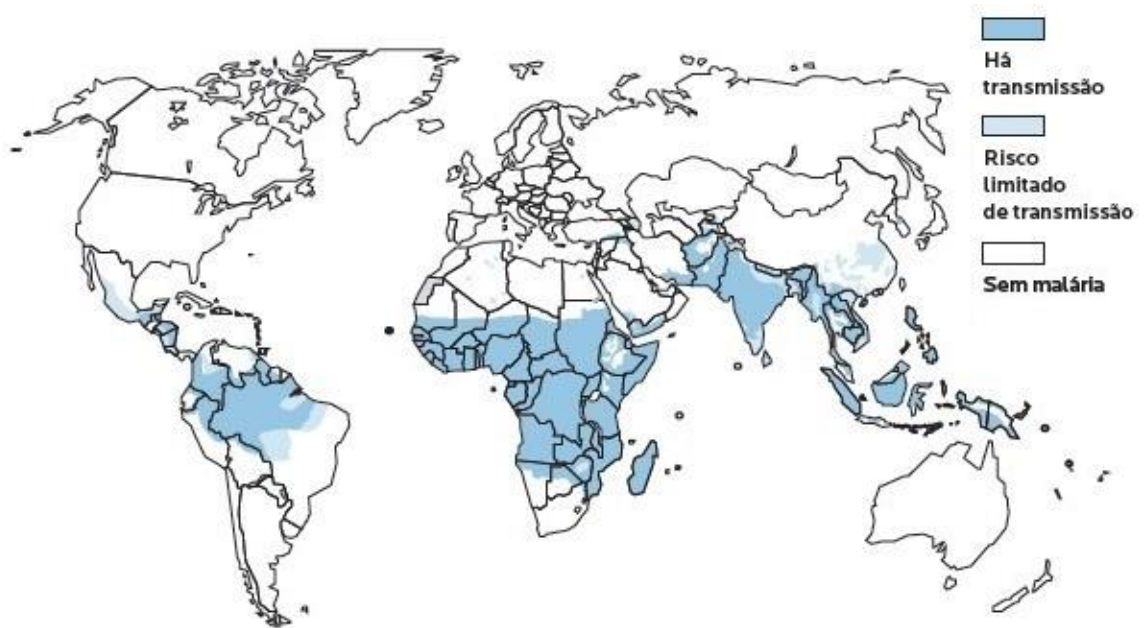


Figura 1. Áreas com transmissão da malária no mundo, 2008. Fonte: OMS

Observa-se que nas regiões tropicais e subtropicais, incluindo partes das Américas, Ásia, África e Oceania, concentra-se a transmissão da doença. As áreas de maior transmissão são também as menos desenvolvidas financeiramente, gerando o desinteresse das indústrias farmacêuticas em investir na pesquisa e desenvolvimento de medicamentos para a cura da malária.

No Brasil, a transmissão ocorre principalmente na região da Amazônia Legal composta pelos estados do Acre, Amapá, Amazonas, Pará, Rondônia, Roraima, Tocantins, Mato Grosso e Maranhão, onde de 10 a 15% da população está em risco, conforme pode ser visto no mapa de incidência parasitária anual (IPA) mostrado na Figura 2.

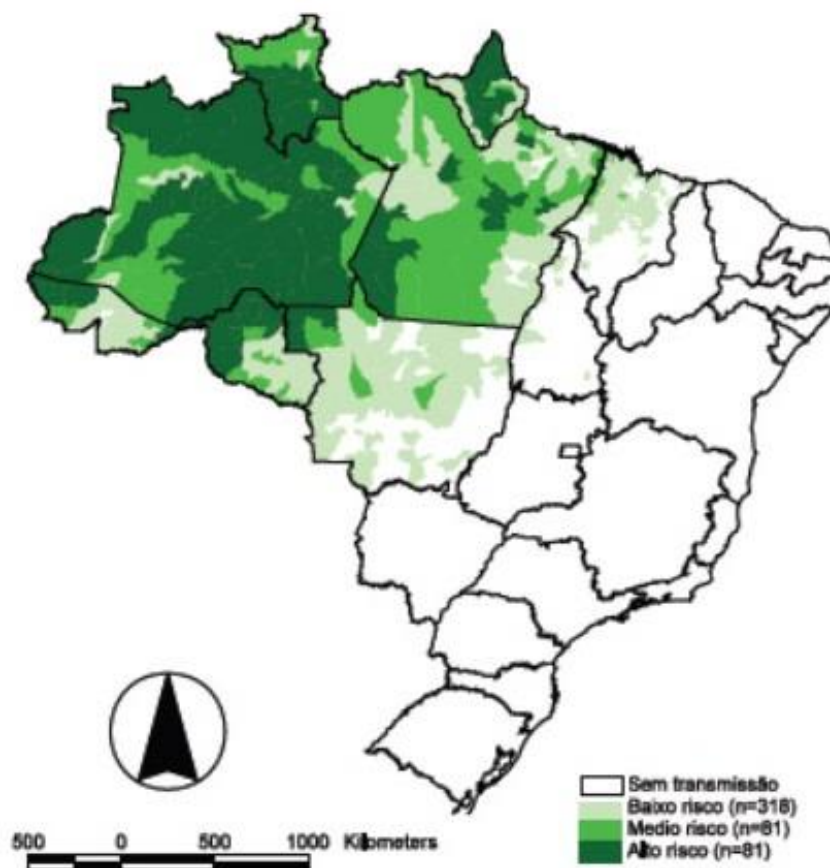


Figura 2. Incidência Parasitária Anual (IPA), Amazônia Legal, 2007.
 Fonte: Sivep-Malária/SVS/MS¹

A malária é causada por parasitas da família *Plasmodiidae*, do gênero *Plasmodium*, transmitida na natureza pela picada de fêmeas de mosquitos do gênero *Anopheles* infectadas. Dentre as espécies que parasitam o homem estão: *P. vivax*, *P. ovale*, *P. malariae*, *P. knowlesi* e *P. falciparum*, sendo este último o mais agressivo e o responsável pela forma grave da doença. (6).

1.1.1 O parasita e seu ciclo biológico

O ciclo biológico do *Plasmodium spp.* é complexo, caracterizado por uma fase sexuada, chamada de esporogonia, que ocorre nas fêmeas do mosquito *Anopheles*, e, duas fases assexuadas no hospedeiro secundário (pré-eritrocítica e eritrocítica). A Figura 3 apresenta o ciclo biológico do *Plasmodium spp.* (7, 8).

¹ <http://www.who.int/malaria/publications>

No ciclo esporogônico, representado pelas setas em vermelho, o mosquito ingere, através da tomada de sangue de um hospedeiro infectado, parasitas no estágio sexual (gametócitos machos e fêmeas) que permanecem em seu sistema digestivo por algum tempo onde ocorrerá fecundação, formação do oocisto e, quando este se torna maduro, romperá liberando esporozoítas que migram para as glândulas salivares permanecendo neste local até a próxima alimentação sanguínea do mosquito, infectando assim um novo hospedeiro vertebrado (9).

Já no ciclo assexuado, representado pelas setas em azul, a infecção se inicia através da picada de uma fêmea do mosquito *Anopheles* infectada onde é introduzido o parasita, que se acha sob a forma de esporozoíto na glândula salivar deste inseto. Após o inóculo, os esporozoítos atingem a corrente circulatória, migram para o fígado, invadem os hepatócitos onde amadurecem em esquizontes através da reprodução assexuada (esquizogonia) liberando merozoítas na corrente sanguínea. Estes merozoítas invadem eritrócitos dentro dos quais sofrem novo ciclo de reprodução assexuada, dando origem a novos merozoítas que invadirão novos eritrócitos. Após diversos ciclos sanguíneos, ocorre a formação de gametócitos, que são ingeridos pelo mosquito durante a próxima alimentação (10).

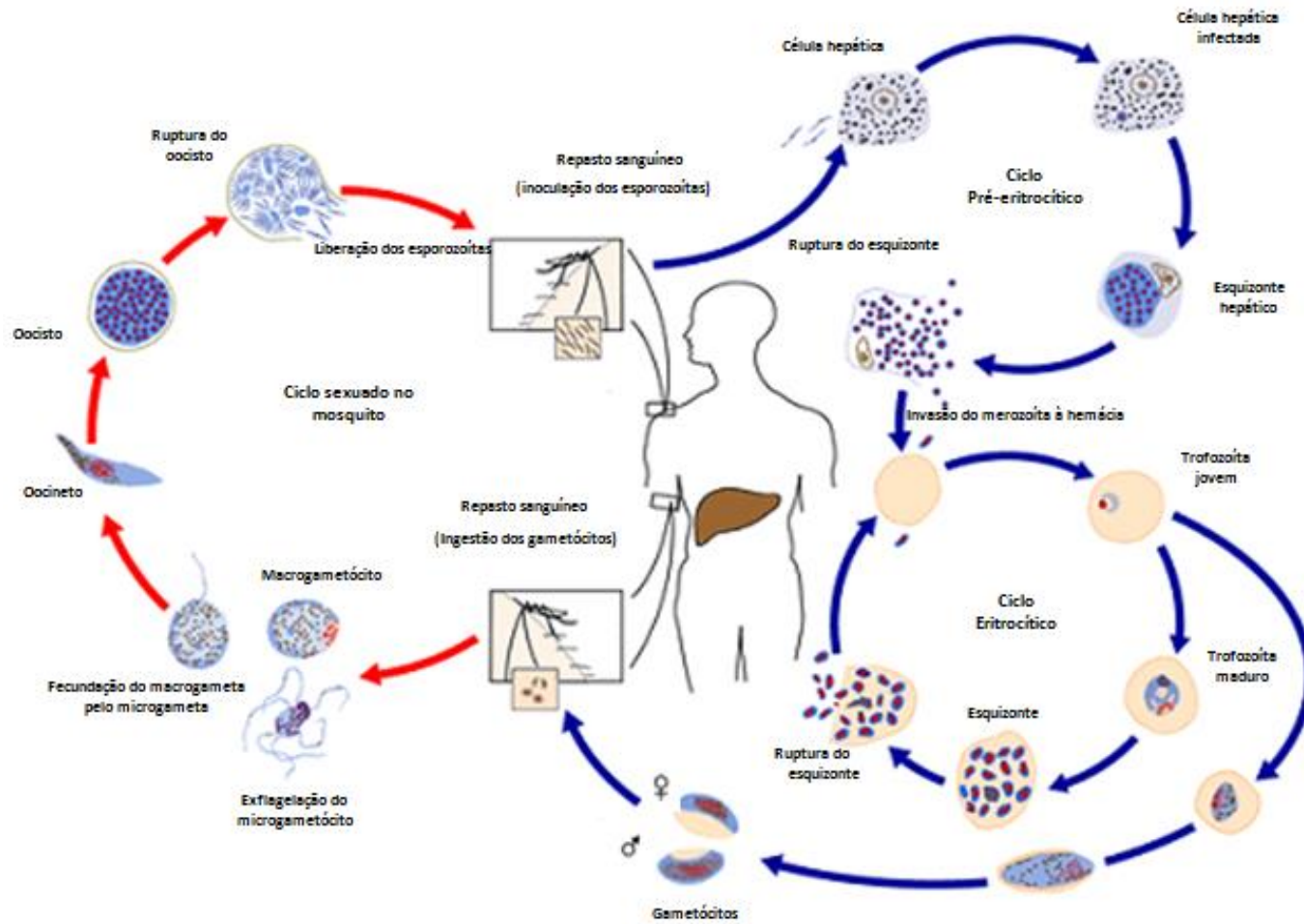


Figura 3. Representação esquemática do ciclo biológico do Plasmodium spp.
 Fonte: Adaptado de <http://www.dpd.cdc.gov/dpdx/HTML/Malaria.htm>

No genoma do *P. falciparum*, foram encontradas sequências codificantes de 10 proteinases aspárticas enumeradas de *Plm I* a *Plm X* (11). Destas, 4 estão ativas no vacúolo alimentar do parasita: *Plasmepsina I (Plm I)*, *Plasmepsina II (Plm II)*, *Histoaspartic proteinase (HAP ou Plm III)* e *Plasmepsina IV (Plm IV)*, estando envolvidas nas etapas iniciais da degradação da hemoglobina humana (12, 13).

1.1.2 Fármacos em uso

Os fármacos antimaláricos utilizados atualmente são baseados em produtos naturais ou compostos sintéticos produzidos a partir da década de 40. Inicialmente, os índios usavam a casca da Cinchona, uma planta nativa da América do Sul, como um remédio tradicional para a malária muito antes que qualquer tratamento estivesse disponível na Europa. Os jesuítas a levaram para o continente europeu mas esta se tornou o tratamento principal para a febre. Apenas em 1820, os químicos franceses Pierre Joseph Pelletier e Joseph Bienaime Caventou identificaram o alcalóide quinina como o ingrediente ativo da casca da cinchona (14).

Com a descoberta da quinina, rapidamente a demanda e uso espalharam-se pela Europa, América do Norte e Ásia, e, até meados do último século era o principal fármaco utilizado no combate à malária. Devido a sua alta toxicidade e com o aparecimento de cepas resistentes de *plasmodium falciparum*, seu uso foi reduzido mas sua importância voltou a aumentar na atualidade em função do surgimento de resistência aos outros fármacos (15).

Para cada etapa do ciclo de vida do parasita, existem fármacos que atuam especificamente em uma das vias fisiológicas. Cloroquina, quinina, mefloquina e artemisinina, atuam no metabolismo do heme, Primaquina e atovaquona, atuam no transporte de elétrons, Doxiciclina, tetraciclina e clindamicina atuam na tradução de proteínas e sulfadoxina-pirimetamina e proguanil atuam no metabolismo do fosfato (16). Na figura 4 é apresentada a estrutura química dos principais fármacos utilizados no tratamento da doença.

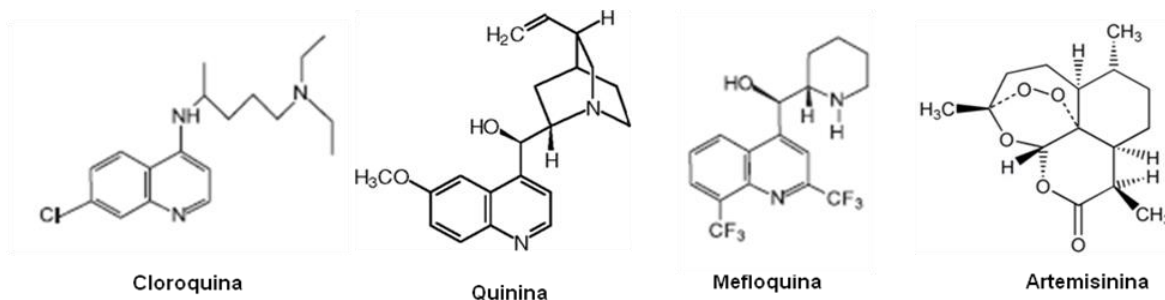


Figura 4. Estrutura química de fármacos que atuam no metabolismo do heme utilizados no tratamento da malária

Os medicamentos para malária são específicos para a fase em que se encontra a espécie de plasmodium. Em geral, o fármaco para a forma sanguínea da doença age também sobre os gametócitos que irão infectar o mosquito vetor da doença. O *Plasmodium falciparum* é o maior problema ao se falar de malária pois o mesmo tem-se mostrado resistente à cloroquina e atualmente, a opção é usar o derivado de uma planta chinesa chamada ginghaosu, cujo princípio ativo é a artemisinina, associado a drogas novas como a mefloquina ou a lumefrantina (17). No Brasil, o tratamento da doença ainda é feito com a mefloquina, porém está sendo discutido a mudança para o esquema atual sugerido pela OMS (18).

A resistência dos parasitas aos medicamentos é um fator que limita seu uso no tratamento da malária falciparum. Por isso, é necessário que novos e efetivos fármacos que apresentem características menos tóxicas sejam planejados.

1.1.3 Mecanismo de ação enzimática

Após a entrada dos merozoítos nas células sanguíneas, sua sobrevivência depende da ingestão da hemoglobina que se encontra no interior do vacúolo digestivo (VD) do parasita. Este vacúolo digestivo mantém um ambiente intravacuolar ácido pela ação de uma próton ATPase na membrana vacuolar. A hemoglobina humana sofre proteólise a aminoácidos através da ação de várias enzimas proteolíticas derivadas do plasmódio, incluindo plasmepsinas, falcipaina e falcilina (19).

Durante o catabolismo da hemoglobina, a fração heme (ferriprotoporfirina

IX) é liberada. A ferriprotoporfirina IX pode reagir com o oxigênio, produzindo superóxido e as enzimas de defesa oxidantes convertem o superóxido potencialmente citotóxico em HO. Os plasmódios polimerizam a ferriprotoporfirina IX no derivado atóxico, hemozoína. Evidências sugerem que a polimerização exige a atividade de proteínas ricas em histidina de carga positiva (15).

A resistência dos parasitos da malária varia de acordo com o antimalárico. O mecanismo de ação desses fármacos, de acordo com Woodrow & Krishna (2006), podem ocorrer por 2 processos: o primeiro, por mutações nos alvos onde atuam as drogas (como antifolatos e atovaquona) e o segundo, por mutações em transportadores envolvidos na distribuição da droga no parasito dentro do eritrócito (gene *ppfcr1* para cloroquina e *pfmdr1* para mefloquina) (20).

Cepas de *P. falciparum* apresentam resistência à vários fármacos. A cloroquina, por exemplo, ainda não tem seu mecanismo de ação totalmente esclarecido, sendo fonte de investigação na pesquisa. Acredita-se que as quinolinas atuem dentro do vacúolo digestivo promovendo uma alteração do gradiente de pH entre os meios interno e externo deste (21).

Assim como a cloroquina, o mecanismo de ação da quinina também tem sido estudada porém permanece incompleta. Pesquisas mostram que a quinina age nos fosfolípidos da membrana do vacúolo digestivo. Essa interação com a membrana do VD inibe algumas vias de transporte de íons comprometendo a degradação da hemoglobina e então, levando o parasita à morte (22).

Outro fármaco utilizado no tratamento da malária e que apresenta resistência é a artemisinina e seus derivados. Seu mecanismo de ação parece estar relacionado com sua capacidade química de gerar radicais livres que são prejudiciais aos parasitos. Recentemente, Woodrow et al, sugeriram um mecanismo de ação alternativo, que consiste na inibição de cálcio ATPase do parasito (23).

Assim como os fármacos já citados anteriormente, outros fármacos utilizados no tratamento da doença também não possuem mecanismo de ação totalmente elucidado. Sendo assim, a busca de novos medicamentos antimaláricos que possam melhorar ou substituir os fármacos atuais tem sido uma necessidade na saúde pública e tem se tornado prioridade da pesquisa em malária (24). Dentro dessa linha, a busca por substâncias antimaláricas contra as formas sanguíneas do parasito tem ganhado ênfase na pesquisa.

1.1.4 Proteases Aspárticas

As proteases aspárticas são uma família de enzimas amplamente distribuídas na natureza sendo identificadas em seres vivos desde microorganismos até mamíferos. Dentre os fenômenos fisiológicos e patológicos podemos identificar o controle da pressão sanguínea (*renina*), proteólise intracelular (*cathepsina D*), digestão (*pepsina*), infecção retroviral (proteínase do HIV) e degradação da hemoglobina na malária (*plasmepsina*), sendo esta última a enzima utilizada neste trabalho.

Essas enzimas atuam em pH ácido, com exceção da renina, e são caracterizadas pela presença de dois resíduos de ácido aspártico localizados na região catalítica e pela preferência na clivagem de ligações peptídicas entre resíduos hidrofóbicos (25).

Durante a infecção em humanos, o parasita degrada a hemoglobina, que é utilizada para suprir suas necessidades por nutrientes no seu crescimento e maturação, debilitando o hospedeiro (26). Dessa forma, as *plasmepsinas* são alvos atrativos para pesquisa e desenvolvimento de novos compostos para o tratamento da malária (27).

1.1.4.1 Plasmepsinas

Dentre as *plasmepsinas* encontradas no genoma do parasita e, mais especificamente as presentes no vacúolo alimentar (*Plm I* à *Plm IV*), iniciamos a busca sobre qual das enzimas utilizar neste trabalho.

No PDB (do inglês, *Protein Data Bank*), banco de dados público de proteínas, encontramos uma lista de complexos proteína-ligante relacionados ao organismo do *P. falciparum*, que pode ser visto na Tabela 1.

Tabela 1. Complexos de Plasmepsinas disponíveis no PDB.

Molécula	Código PDB	Ligante	# Torções do Ligante	Resolução ^(a)
Plm I	3QRV	-	-	2,4
	3QS1	PEPSTATINA ANÁLOGA	10	3,1
Plm II	1LEE	RS367	13	1,9
	1LF2	RS370	15	1,8
	1LF3	EH58	19	2,7
	1M43	PEPSTATINA A	24	2,4
	1SME	PEPSTATINA A	24	2,7
	1ME6	ESTATINA	17	2,7
	1W6H	TIT*	19	2,24
	1W6I	PEPSTATINA A	24	2,7
	1XDH	PEPSTATINA A	24	1,7
	1XE5	PEPSTATINA ANÁLOGA	25	2,4
	1XE6	PEPSTATINA ANÁLOGA	22	2,8
	2BJU	IH4*	14	1,56
	2R9B	DCL*	32	2,8
Plm III	3QVC	-	-	2,1
	3QVI	PEPSTATINA ANÁLOGA	13	2,5
	3FNS	-	-	2,5
Plm IV	1LS5	PEPSTATINA A	24	2,8
	2ANL	PEPSTATINA ANÁLOGA	9	3,3

* Nomenclatura dada pelo autor; ^(a) Resolução em Ångstroms.

Foi escolhido um complexo de cada tipo de plasmepsina, com critério de menor resolução ou ligante complexado à proteína, e feito um alinhamento entre as sequências para saber se as estruturas são similares e possuem identidade entre si. Na Figura 5 é possível visualizar o alinhamento dos complexos de código PDB: 3QRV, 1LF3, 3FNS e 1LS5, respectivamente plasmepsina I, plasmepsina II, plasmepsina III e plasmepsina IV. Os aminoácidos marcados na figura

representam os aminoácidos importantes na interação proteína-ligante descritos na literatura para essa família de enzimas.



Figura 5. Alinhamento entre as quatro sequências de plasmepsinas utilizando a ferramenta ClustalW²

Dentre as quatro *plasmepsinas* escolhidas no alinhamento, verifica-se que as mesmas possuem de 328 à 336 resíduos de aminoácidos. Ao inserirmos as sequências no servidor online *blastp*³, obtemos a porcentagem de similaridade entre essas sequências utilizando como referência a estrutura de código 1LF3, complexo da PlmII, que é o grupo de enzimas de alvo mais atrativo (27). Essa comparação é mostrada na Figura 6.

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
22852	1LF3:A PDBID CHAIN SEQUENCE Plm II	677	677	100%	0.0	100%
22853	1LS5:A PDBID CHAIN SEQUENCE Plm IV	485	485	98%	1e-176	69%
22854	3QRV:A PDBID CHAIN SEQUENCE Plm I	513	513	100%	0.0	73%
22855	3FNS:A PDBID CHAIN SEQUENCE HAP	431	431	99%	3e-155	61%

Figura 6. Grau de identidade entre as sequências de plasmepsinas utilizando o servidor online *blastp* e utilizando a proteína Plm II como referência

² <http://www.genome.jp/tools/clustalw/>

³ <http://blast.ncbi.nlm.nih.gov/Blast.cgi>

Tanto o cobertura quanto a identidade entre as sequências obtiveram valores maiores que 60% e por este motivo, a escolha da proteína utilizada para as simulações deste trabalho foi baseada no ligante ao qual a mesma está complexada.

Sendo assim, de acordo com os graus de liberdade (torções) dos compostos avaliados nesse estudo, que varia de 8 à 15, escolhemos a estrutura experimental da enzima *Plasmepsina II* complexada com o ligante EH58 (código PDB: 1LF3) (28), que embora tenha resolução de 2,7 Å, está complexada com um ligante grande, que apresenta 19 torções e possui agrupamentos similares às moléculas estudadas neste trabalho, permitindo uma melhor acomodação no sítio ativo da proteína. Este é um fator importante a ser considerado, uma vez que utilizamos a aproximação de receptor rígido. O ligante EH58 é representado na Figura 7 com todas as possíveis torções.

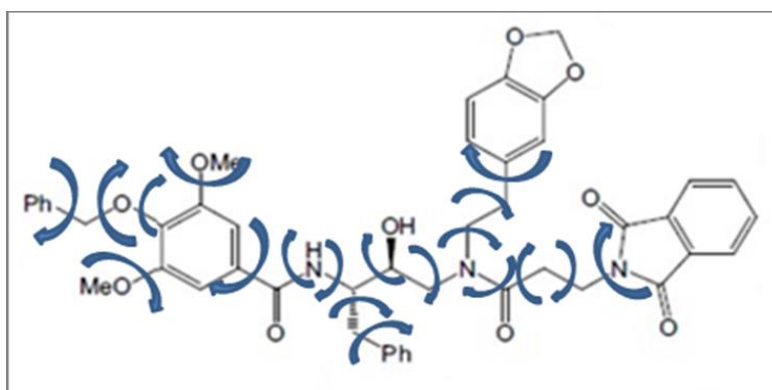


Figura 7. Ligante EH58 (PDB id: 1LF3) com todas as possíveis torções

Com a obtenção do arquivo PDB deste complexo, é possível conhecer o modo de ligação da molécula ligante no sítio ativo da molécula receptora através das coordenadas cartesianas dos átomos. Este fato permite que os programas de atracamento molecular sejam testados e posteriormente utilizados para prever o modo de ligação de outras moléculas ligantes. A Figura 8 mostra o mapa de interações entre a enzima *Plasmepsina II* e o ligante EH58 a partir da configuração cristalográfica.

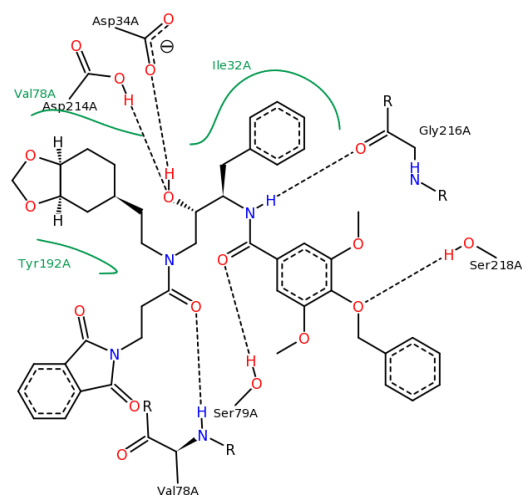


Figura 8. Mapa de interações entre o ligante EH58 e a enzima Plasmepsina II. As linhas pontilhadas indicam uma possível ligação hidrogênio e o traço em verde mostra a interação hidrofóbica entre o ligante e os resíduos do sítio ativo da proteína.

Fonte: <http://www.pdb.org/pdb/explore/explore.do?structureId=1LF3>

Pode-se observar que este ligante interage com os resíduos Ile32, Asp34, Val78, Ser79, Tyr192, Asp214, Gly216 e Ser218, que são considerados resíduos relevantes da enzima *Plm II* (27). Esses resíduos também são mostrados como esferas na Figura 9 e indicam, na estrutura tridimensional das quatro *plasmepsinas* alinhadas, que compartilham o mesmo sítio de ligação na proteína.

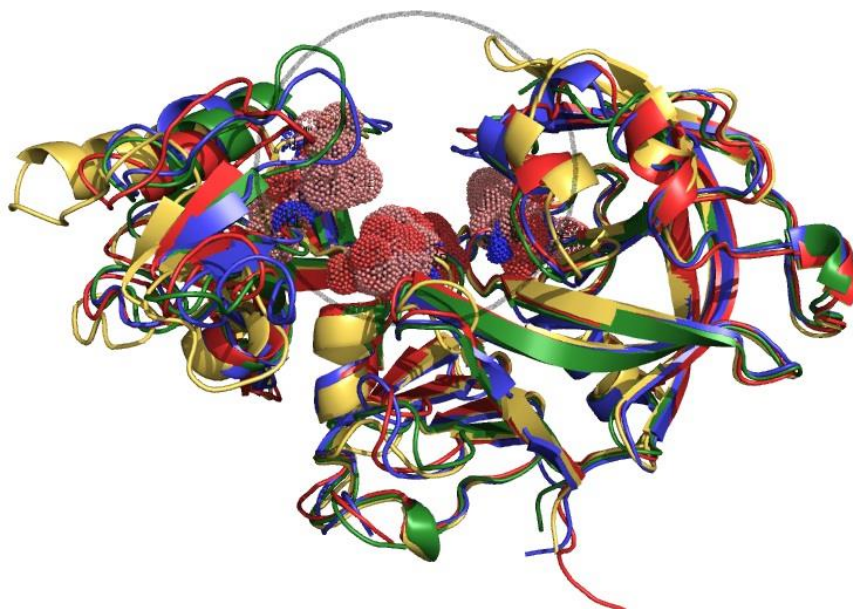


Figura 9. Representação espacial das quatro estruturas de plasmepsinas indicando através do círculo que todas compartilham o mesmo sítio de ligação. As esferas representam os aminoácidos importantes na ligação proteína-ligante.

1.1.5 Novos inibidores planejados por Farmanguinhos

O Instituto de Tecnologia em Fármacos de Farmanguinhos tem trabalhando na síntese e na avaliação antimalarial de análogos de inibidores da enzima *aspartil protease*. Esses análogos são derivados de *hidroxietilpiperazinas* que foram planejados de acordo com estudos da relação estrutura e atividade de uma série anterior (29). Além disso, resultados da avaliação *in vitro* contra a cepa W2 de *P. falciparum* comprovam atividade comparável à inibidores já conhecidos como *lopinavir*, que é o inibidor de protease do HIV mais ativo contra o *P. falciparum*, reforçando a nova série de inibidores como uma classe promissora de novos antimalariais.

A nova série é composta de 9 moléculas nomeadas como Plip 01, Plip 03, Plip 04, Plip 11, Plip 13, Plip 27, Plip 31, Plip 41 e Plip 47. A Tabela 2 detalha o número total de graus de liberdade (torções), peso molecular, valor experimental obtido e uma informação referente à atividade da molécula obtida através de testes experimentais.

Tabela 2. Compostos planejados pelo Instituto de Fármacos – Farmanguinhos

Molécula	Número de Torções	Peso Molecular (Da)	ΔG experimental (a)*	Atividade (b)
Plip01	8	314	-7,28	A
Plip03	10	532	-7,73	A
Plip04	13	750	-6,16	A
Plip11	11	484	-7,90	A
Plip13	15	654	-6,60	PA
Plip27	11	499	-7,13	A
Plip31	15	684	-6,73	A
Plip41	10	472	-8,27	A
Plip47	13	630	-6,44	PA

(a) ΔG obtido através do cálculo: $\ln(\text{valor experimental (IC}_{50}) \text{ em } \mu\text{M} * 10^{-6}) * 0,616$; (b) A: molécula apresenta atividade / PA: molécula parcialmente ativa; *Valores em kcal/mol

A estrutura dos compostos, planejados pelo DSO de Farmanguinhos, e os substituintes da posição R1 e R2 são apresentadas nas Figuras 10, 11 e 12 abaixo.

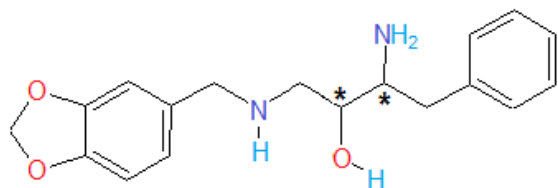


Figura 10. Estrutura química da molécula Plip01

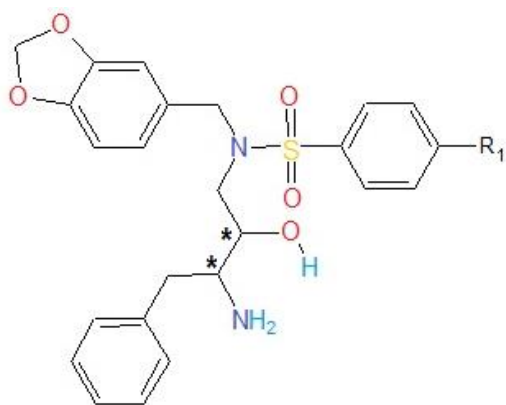


Figura 11. Estrutura química das moléculas Plip03, Plip11, Plip27 e Plip41

	Plip 03	Plip 11	Plip 27	Plip 41
R1	Br	O Me	NO2	F

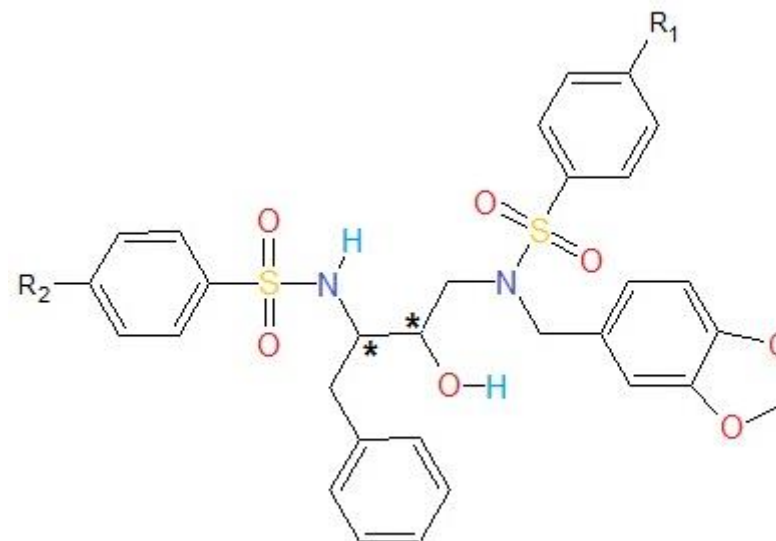


Figura 12. Estrutura química das moléculas Plip04, Plip13, Plip31 e Plip47

	Plip 04	Plip 13	Plip 31	Plip 47
R1	Br	O Me	NO2	F
R2	Br	O Me	NO2	F

* indicam o carbono quiral da molécula

1.2 Planejamento Racional de Fármacos Baseado em Estrutura

O Desenho Racional de Fármacos Baseado em Estrutura (SBDD) fundamenta-se no estudo de estruturas moleculares tridimensionais da molécula receptora para o desenho de compostos protótipos (moléculas ligantes candidatas a fármaco) tomando como base as informações estruturais e as interações envolvidas no processo de reconhecimento molecular receptor-ligante (30).

Metodologias de *docking* receptor-ligante são amplamente utilizadas dentro do SBDD, tanto para a descoberta de novas substâncias bioativas – através de técnicas conhecidas como triagem virtual (*virtual screening*) – quanto para o refinamento e otimização de compostos bioativos previamente identificados. Esta metodologia está dividida em várias etapas que podem ser observadas na Figura 13.

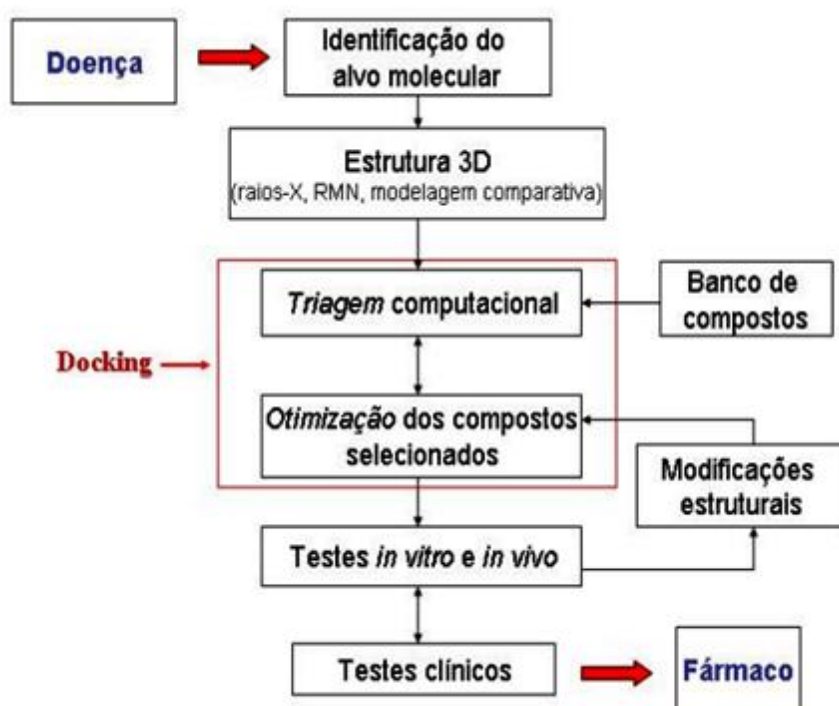


Figura 13. Etapas do desenho racional de fármacos baseado em estrutura
Fonte: (31)

A primeira etapa do SBDD é a identificação do alvo molecular relacionado à doença para a qual se deseja obter a cura. Após essa seleção, a estrutura tridimensional do alvo molecular precisa ser obtida. Isto pode ser feito através de técnicas experimentais como Ressonância Magnética Nuclear (RMN), difração de

raios-X em cristais ou por métodos teóricos como a modelagem comparativa (32). Além disso, existem várias estruturas tridimensionais depositadas em grandes bancos de estruturas moleculares de acesso público como, por exemplo, o Protein Data Bank (PDB) (3).

Com a estrutura 3D definida, metodologias de *docking* podem ser empregadas com uma triagem *in silico*, ou seja, através do *virtual screening* (triagem virtual). Com a utilização destes métodos, grandes bancos de estruturas moleculares de ligantes podem ser testados contra o alvo molecular com o objetivo de identificar as substâncias biologicamente ativas candidatas a novos fármacos.

Quando um composto promissor é encontrado, metodologias de *docking* mais acuradas são utilizadas para prever os modos de ligação mais prováveis, que são analisados pelo químico medicinal, visando à inclusão de modificações que possam ser efetuadas na molécula para que a resposta biológica seja aumentada. Após a otimização dos compostos protótipos, testes *in vitro* e *in vivo* são realizados para que dentre outras características, toxidez e absorção do fármaco no organismo, sejam analisadas.

1.3 Interações Receptor Ligante

As interações moleculares são responsáveis pelo arranjo de muitas estruturas biológicas. Ligações hidrogênio e interações hidrofóbicas são principalmente responsáveis pela estrutura tridimensional de biopolímeros, tais como proteínas, ácidos nucleicos, e membranas celulares. A interação de uma molécula ligante a um biopolímero (receptor) também é governada por interações moleculares (33). Exemplos de complexos receptor-ligante incluem complexos enzima-substrato, antígeno-anticorpo, e complexos de receptores proteicos como fármacos.

Em todos esses casos, um sítio no receptor contém grupos químicos funcionais que podem interagir de forma complementar com outros grupos específicos do ligante. É geralmente verdade que muitos contatos específicos intermoleculares devem se formar num complexo biológico receptor-ligante e como resultado desta especificidade, um dado ligante só se liga em receptores que apresentem estruturas similares. Assim, essas regras que governam o

reconhecimento molecular controlam todos os processos biológicos, desde metabolismo a respostas imunes, além de providenciar importantes informações para o desenho racional de fármacos efetivos para o tratamento das doenças (34).

Interações entre grupos químicos apolares podem ser importantes no processo de ligação. Por exemplo, muitos sítios ativos de enzimas possuem regiões hidrofóbicas que ligam grupos apolares do substrato. As interações de empilhamento π - π são também possíveis, além das conhecidas forças de dispersão, forças repulsivas e interações hidrofóbicas (35). Todas essas interações podem ser modeladas matematicamente através do potencial de Lennard-Jones (eq. 1).

$$V_{LJ}(i, j) = \frac{C_{12}(i, j)}{r_{ij}^{12}} - \frac{C_6(i, j)}{r_{ij}^6} \quad (1)$$

onde o termo $(1/r_{ij})^{12}$ quantifica a repulsão entre as nuvens eletrônicas dos átomos interagentes e o termo $(1/r_{ij})^6$, é o termo de dispersão, (atrativo) entre pares de átomos. r_{ij} é a distância entre os núcleos dos átomos envolvidos (36). Os parâmetros $C_{12}(i, j)$ e $C_6(i, j)$ são ajustados para reproduzir dados experimentais.

As interações entre grupos carregados eletricamente, descritas através do potencial de Coulomb (eq. 2), podem ser importantes no interior de uma proteína, onde a permissividade relativa ϵ pode ser muito mais baixa que o ambiente aquoso externo.

$$V_c = \sum_{i=1}^{NA} \sum_{j=1}^{NB} \frac{q_i q_j}{4\pi\epsilon_0\epsilon r_{ij}} \quad (2)$$

Sendo q_i e q_j os valores das cargas parciais dos átomos, r_{ij} é a distância entre os núcleos desses átomos, ϵ_0 , é a permissividade do vácuo, ϵ a constante dielétrica do meio e NA e NB são os números de pontos de carga de cada molécula.

Por exemplo, no pH fisiológico, as cadeias laterais de resíduos de aminoácidos que contêm ácidos carboxílicos ou grupos aminos estão carregadas negativa e positivamente, respectivamente, e podem se atrair gerando uma ponte

salina. As interações dipolo-dipolo são também possíveis dado que muitos dos elementos que estabilizam a estrutura de biopolímeros são polares, incluindo a ligação peptídica -CONH-.

Porém, as ligações hidrogênio são as mais importantes em complexos de interação receptor-ligante (37). Muitos fármacos efetivos se ligam fortemente e inibem a ação de enzimas que estão associadas ao progresso da doença. Na maioria dos casos, um inibidor bem sucedido será capaz de formar ligações hidrogênio com resíduos de aminoácidos do sítio ativo, que o substrato também pode formar (38).

A ligação hidrogênio é uma interação que ocorre entre moléculas que apresentam átomos eletronegativos (doadores) ligados covalentemente a um átomo de hidrogênio e outros átomos eletronegativos como oxigênio e nitrogênio.

Para a existência da ligação hidrogênio é necessária a presença simultânea de um átomo de hidrogênio ácido e de um receptor e doador básico. Hidrogênio ácido é aquele ligado a um átomo mais eletronegativo do que ele, de maneira que os seus elétrons sofram um afastamento parcial. Receptor básico é uma espécie química que possui um átomo ou grupo de átomos com alta densidade eletrônica, sendo que o ideal é a presença de pelo menos um par de elétrons livres (39).

A intensidade destas interações varia de 4 até 25 KJ/mol e a distância entre os átomos doador (D) –aceitador (A) para que haja interação é de no máximo 3,5 Å (Figura 14), formando um ângulo D-H-A entre 145° e 225°, onde D e A são os átomos doador e aceitador, respectivamente (40). Embora a energia de interação de ligações hidrogênio seja apenas cerca de 12% de uma ligação química covalente, elas são significativamente mais fortes do que as interações dipolo-dipolo típicas e de dispersão de London.

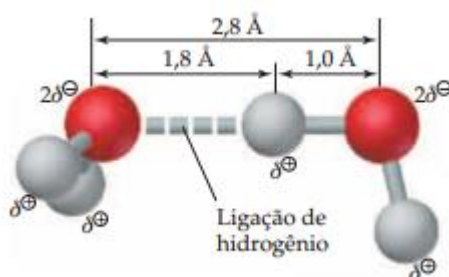


Figura 14. Distância entre os átomos doador-aceitador da ligação hidrogênio
Fonte: (40) (Pag. 382)

1.4 Campos de Força Clássicos

A representação física de um sistema molecular em simulações computacionais pode ser feita através de uma função potencial ou campo de força. Genericamente atribui-se a denominação de Campo de Força à descrição de um sistema de muitas partículas pela sobreposição de termos simples, que descrevem a interação entre duas, três ou quatro partículas.

Para o tratamento de sistemas com centenas ou milhares de átomos como proteínas e ácidos nucleicos são introduzidas funções potenciais empíricas cujos parâmetros são calibrados a partir de informações experimentais e/ou cálculos quânticos sobre pequenas moléculas. Ao conjunto de parâmetros de ajuste e às suas respectivas funções atribui-se o nome de Campo de Força Molecular (41-43).

A maioria dos campos de força clássicos é descrita através da função de energia potencial para um sistema constituído de N átomos, com vetores de posições \mathbf{r}_i . A forma funcional é expressa por:

$$\begin{aligned} \mathbf{V}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = & \sum_{n=1}^{N_b} \frac{1}{2} K_{b_n} (b_n - b_{0_n})^2 + \sum_{n=1}^{N_\theta} \frac{1}{2} K_{\theta_n} (\theta_n - \theta_{0_n})^2 + \sum_{n=1}^{N_\xi} \frac{1}{2} K_{\xi_n} (\xi_n - \xi_{0_n})^2 + \\ & \sum_{n=1}^{N_\varphi} K_{\varphi_n} [1 + \cos(n_n \varphi_n - \delta_n)] + \\ & \sum_{i < j}^N \left[\frac{C_{12}(i, j)}{r_{ij}^{12}} - \frac{C_6(i, j)}{r_{ij}^6} \right] + \sum_{i < j}^N \frac{q_i q_j}{4\pi\epsilon_0 \epsilon r_{ij}} \end{aligned} \quad (3)$$

onde os quatro primeiros termos se referem especificamente a interações entre átomos ligados (comprimento de ligação, ângulo de ligação, ângulos diedrais impróprios e próprios respectivamente). Os dois últimos termos, de van der Waals e eletrostático respectivamente, estão relacionados com interações de longo alcance entre átomos não ligados entre si. A constante dielétrica do meio é dada pelo valor do parâmetro ϵ , a carga parcial atômica no átomo i é dada pelo parâmetro q_i e a distância entre dois átomos i e j é dada por r_{ij} .

1.5 Métodos de Atracamento Molecular (*Docking*)

Os métodos de atracamento molecular visam predizer o modo de ligação de moléculas ligantes no sítio ativo de uma molécula receptora, além de estimar a afinidade de ligação entre o receptor e o ligante (44).

Esse mecanismo pode ser ilustrado através do modelo chave-fechadura sugerido por Emil Fischer em 1894 (34). Neste modelo, o receptor é representado pela “fechadura” sendo o “buraco da fechadura”, o sítio ativo. Já a chave, é representada pela molécula ligante que irá “abrir ou trancar a porta”. Com isso, teremos três tipos de ligantes que podem ser relacionados a duas atividades biológicas: (1) o ligante natural (agonista natural) e o ligante modificado (agonista modificado) desencadeiam uma resposta biológica e neste modelo representariam, respectivamente, a chave original da fechadura e a chave modificada. Ambas as chaves são capazes de “abrir a porta”; (2) o ligante antagonista, que bloqueia uma resposta biológica, é representado por uma chave falsa, que apesar de ter acesso à “fechadura”, não abre a porta, ou seja, não desencadeia a resposta biológica (45).

Porém, este modelo não representa de forma adequada a realidade pois no caso biológico, tanto “chave” quanto “fechadura” são flexíveis, ou seja, podem modificar sua conformação durante o processo de encaixe da molécula ligante na região de ligação do receptor. Na realidade, a molécula ligante interage com a proteína em uma conformação específica através de interações físico-químicas levando em consideração que esta conformação provavelmente é a que possui maior afinidade de ligação com o ligante.

Desde o início da década de 80, vários algoritmos de *docking* vêm sendo desenvolvidos. Atualmente, existem diversos programas de atracamento molecular disponíveis tanto gratuitamente quanto em forma comercial e eles diferem entre si principalmente em relação a fatores tais como (i) função de avaliação implementada; (ii) flexibilidade molecular permitida para o receptor e para o ligante e (iii) método de otimização usado para explorar a hipersuperfície de energia.

A flexibilidade das moléculas é uma questão importante a ser levada em conta, porém torna-se inviável computacionalmente simular um sistema em que a molécula receptora tenha todos seus graus de liberdade ativos. Os primeiros

algoritmos que faziam simulações de atracamento molecular, tratavam tanto o receptor quanto o ligante como moléculas rígidas considerando apenas os graus de liberdade translacionais e rotacionais da molécula ligante. Atualmente, a maioria dos programas de atracamento molecular considera a flexibilidade total do ligante, com a inclusão dos graus de liberdade conformacionais (46).

A complexidade do problema de *docking* requer métodos computacionais com potencial para investigar efetivamente um grande número de soluções possíveis, para que a solução ótima seja encontrada. Para isso, algoritmos de *docking* utilizam métodos de busca divididos em 3 categorias, de acordo com Xu, 2007 (47): busca sistemática, busca determinística e busca estocástica. A busca sistemática consiste em explorar combinatoriamente todos os graus de liberdade da molécula. Um exemplo desta busca são os algoritmos conhecidos como construção incremental ou baseados em fragmentos (48). A busca determinística é caracterizada por conseguir reproduzir sempre a mesma conformação de saída dados os mesmos parâmetros e coordenadas de entrada; esta busca pode ser exemplificada pelos métodos clássicos de simulação por dinâmica molecular e minimização de energia. Já a busca estocástica, como o próprio nome sugere, é caracterizada por utilizar variáveis aleatórias para produzir diferentes conformações de saída para um mesmo conjunto de parâmetros de entrada. Isso ocorre pois os algoritmos utilizam “sementes” diferentes para a geração de números pseudo-aleatórios. Essa busca pode ser exemplificada pelos métodos de Monte Carlo, *Simulated Annealing* e algoritmos evolucionários (45).

Nas últimas duas décadas, diversos programas de *docking* têm sido propostos tanto para uso comercial quanto para uso acadêmico. Dentre os programas existentes, podem-se citar os mais popularmente conhecidos atualmente: Autodock 4 (49), Autodock Vina(50), Dockthor (51), Glide (52), GOLD (53), FlexX (54) e Surflex (55). Dos programas mencionados, somente o Autodock 4, sua variante, Autodock Vina e Dockthor são disponibilizados gratuitamente à comunidade científica para serem usados sem fim comercial.

1.5.1 Programas de Atracamento Molecular

Dos programas de atracamento molecular disponíveis atualmente, escolhemos utilizar neste trabalho os programas Autodock 4, Autodock Vina e DockThor. Cada um desses programas será descrito a seguir.

Autodock 4

O programa Autodock foi desenvolvido na Universidade da Califórnia, em San Diego pelo grupo do Prof. Arthur J. Olson. Este programa utiliza um algoritmo genético Lamarckiano (LGA) para encontrar a melhor conformação do ligante no sítio ativo do receptor. O algoritmo genético Lamarckiano é baseado na teoria da evolução de Lamarck, onde aos descendentes são transmitidas as mudanças adquiridas durante a vida de seus antecessores (49).

Este algoritmo utiliza, para avaliação das conformações, uma função empírica que determina a interação através da energia livre de ligação receptor-ligante. A equação 4 descreve esta função implementada neste programa.

$$\Delta G = \Delta G_{vdw} + \Delta G_{hbond} + \Delta G_{ele} + \Delta G_{conf} + \Delta G_{solv} \quad (4)$$

onde ΔG_{vdw} é o termo que representa a contribuição de van der Waals, ΔG_{hbond} é o termo que descreve a contribuição das ligações hidrogênio, ΔG_{ele} , representa as contribuições do potencial eletrostático, ΔG_{conf} , descreve a contribuição entre os ângulos dos átomos (termo torcional) e ΔG_{solv} , a contribuição de solvatação.

Autodock Vina

O programa Autodock Vina herda algumas idéias do Autodock 4 tais como o tratamento do docking com a otimização estocástica global da função custo (função scoring) e pré-cálculo dos mapas de interação da grade, onde são armazenados valores das interações receptor-ligante. A vantagem da utilização deste programa é deste possuir uma maior velocidade comparada aos outros programas, a partir do paralelismo do software, usando multithreading em

máquinas com vários núcleos, além de automatizar vários passos levando o processo a ser mais transparente para o usuário (50).

De acordo com os autores, sua função de avaliação combina certas vantagens do potencial baseado em conhecimento e da função score empírica. Ela extrai informação empírica de ambas preferências conformacionais do complexo receptor-ligante e da medida de afinidade experimental. Sua função é inspirada na função score X (56) porém diferenciando em alguns termos e na parametrização do método. Sua principal desvantagem consiste em não disponibilizar os termos de energia separadamente (van der Waals, eletrostática, torcional e outras) nos arquivos de saída. A derivação da função scoring utilizada no programa Autodock Vina será disponibilizada separadamente em uma nova publicação.

Dockthor

O programa Dockthor foi desenvolvido durante a tese de doutorado de Magalhães, 2006 (31). Na primeira versão deste programa, o algoritmo utilizava uma busca estocástica através de um algoritmo genético (GA) não geracional. Este programa foi aperfeiçoado na dissertação de mestrado de Bellini, 2008 (57) com relação à distribuição da população inicial do algoritmo genético. A última versão deste programa, versão utilizada neste trabalho, foi aperfeiçoada por Marinho, 2011 (58) em sua dissertação de mestrado onde o autor portou algumas rotinas em Fortran para linguagem de programação C, tornando o tempo de processamento menor, e também implementando o campo de força MMFF94.

A função de avaliação do programa é baseada na energia de interação proteína-ligante. Ela é calculada usando o campo de força molecular clássico Gromos96 (59), implementado no programa de mecânica/dinâmica molecular THOR (60-62). A função aptidão utilizada compreende os termos para átomos não-ligados da energia de interação proteína-ligante e da energia intramolecular do ligante, além do termo relativo aos ângulos diedrais da molécula ligante. A equação 5 descreve a forma funcional da função:

$$\begin{aligned}
& \sum_{\text{proteina}} \sum_{\text{ligante}} \left\{ p \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \frac{q_i q_j}{D(r_{ij}) r_{ij}} \right\} + \\
& \sum_{\text{ligante}} \sum_{\text{ligante}} \left\{ \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{D(r_{ij}) r_{ij}} \right\} + \sum_{\text{diedrais}} \gamma_k (1 + \cos(\omega_k \theta_k - \theta_{0k}))
\end{aligned} \tag{5}$$

Onde r_{ij} é a distância entre os átomos i e j , A_{ij} e B_{ij} são os parâmetros de Lennard-Jones, q_i e q_j são as cargas parciais atômicas dos átomos i e j respectivamente, e D é a função dielétrica sigmoidal dependente da distância r_{ij} , que modela o efeito de blindagem do solvente nas interações eletrostáticas. O parâmetro γ_k é a constante de energia associada com a rotação de uma ligação química, θ_k é o ângulo de torção, ω_k é a periodicidade e θ_{0k} é o ângulo de fase. Os dois primeiros termos da equação (entre chaves) se referem especificamente às interações entre átomos não-ligados, sendo o primeiro termo de interação entre a proteína e o ligante e o segundo termo entre os átomos do ligante. Esses mesmos termos são usados para avaliação da energia de interação intramolecular do ligante (segundo termo entre chaves). O terceiro termo é o potencial torcional, que se refere à torção dos ângulos diedrais (ligações com quatro átomos) da molécula ligante (31).

O programa dá como resultado dois valores de energia: energia total e energia de interação. A energia total é calculada com base em todas as interações entre ligante e receptor e a energia de interação, leva em consideração somente a soma dos termos de van der Waals e eletrostático.

1.6 Métodos para Cálculo da Energia Livre de Ligação e Energia de Interação Linear (LIE)

A energia livre de Gibbs (ΔG), é um potencial termodinâmico que relaciona a entalpia (H) e a entropia de um sistema, quando as variáveis temperatura (T) e pressão (P) se mantêm constante no processo. Essas variáveis se relacionam através da equação 6:

$$G = H - TS \tag{6}$$

onde H é a entalpia, T a temperatura absoluta em Kelvin e S a entropia do sistema. Esse potencial termodinâmico permite prever se uma reação química do ponto de vista termodinâmico ocorrerá de forma espontânea ou não.

Do ponto de vista teórico, SBDD possui como objetivo final a determinação de diferenças de energia livre, ΔG , entre dois estados de equilíbrio termodinâmico (63, 64), como no caso onde dois ligantes distintos estão complexados a um mesmo receptor proteico. Neste sentido a determinação teórica de diferenças de energia livre através de metodologias computacionais é de grande importância na área da química medicinal além de poderem ser relacionadas diretamente com as constantes de inibição K_i receptor-ligante determinadas experimentalmente por

$$\Delta G_{obs} = RT \ln K_i \quad (7)$$

onde R é a constante dos gases e T é a temperatura absoluta.

A obtenção de estimativas mais precisas de ΔG envolve um custo computacional muitas vezes proibitivo. Atualmente, os métodos baseados na mecânica molecular, como modelagem comparativa ou dinâmica molecular são abordagens confiáveis para simular os efeitos de mutações específicas da afinidade entre ligantes e receptores.

O método de Energia de Interação Linear (LIE) (65) foi originalmente desenvolvido para a predição da afinidade de ligação entre uma família de inibidores estruturalmente relacionados com a endotiapepsina (EP), uma proteína pertencente à família de aspartil proteases. O método se baseia na combinação linear de variações entre valores médios de energia de interação do ligante com o receptor obtidas a partir de simulações por dinâmica molecular para o cálculo da energia livre.

Esta metodologia é considerada útil devido a sua simplicidade, redução do tamanho do sistema a ser simulado, representação através de modelos de água explícita, possibilidade de calcular valores absolutos da energia de ligação e sua utilização em curtos tempos de simulação da ordem de dezenas de nanosegundos.

Mais especificamente, esse método empírico se baseia no cálculo da energia de interação entre o ligante e o meio (proteína, solvente, etc). As contribuições à energia livre podem ser divididas em dois termos, eletrostático e de Lennard-Jones. A principal idéia do método é dividir as contribuições polares (eletrostáticas) e apolares (Lennard-Jones) para o cálculo de energia livre de ligação (ΔG) aproximando-o por uma medida linear dos termos correspondentes às médias temporais dessas interações intermoleculares nos dois estados termodinâmicos do equilíbrio: ligado (*bound*) e não ligado no receptor (*free*).

$$\Delta G = \alpha \left(\langle U_{vdw}^{l-s} \rangle - \langle U_{vdw}^{l-w} \rangle \right) + \beta \left(\langle U_{el}^{l-s} \rangle - \langle U_{el}^{l-w} \rangle \right) + \gamma \quad (8)$$

onde os colchetes representam valores médios, vdw e el, respectivamente representam as contribuições das energias do potencial de Lennard-Jones e eletrostáticas do ligante com o seu entorno (l-s), obtidas por simulações de dinâmica molecular ou Monte Carlo. A diferença entre os termos dentro dos colchetes indica variação entre estados, do não-ligado (s) (em solução) para o ligado (l) (em complexo).

Os coeficientes de calibração do método LIE, α , β e γ (eq. 8) são dependentes do sistema de estudo. Nesse caso particular, o coeficiente da contribuição eletrostática para a variação da energia livre foi $\beta = 0,37$ de acordo com a aproximação linear, porém valores diferentes para esse coeficiente têm sido encontrados dependendo da estrutura química do ligante: 0,33 (com dois ou mais grupos hidroxila), 0,37 (um grupo hidroxila), 0,43 (polar sem grupos hidroxila) e 0,50 (carregado) (66). A contribuição apolar (correspondente ao coeficiente α) é determinada de forma empírica e mantém uma boa aproximação linear com medidas relacionadas ao tamanho do ligante: área, volume e contribuição de átomos de características hidrofóbicas. Em estudos realizados previamente, (66) o valor de $\alpha = 0,181$ tem reproduzido energias livre de ligação de vários sistemas, e tem-se demonstrado que esse coeficiente não é fortemente dependente do sistema nem do campo de força utilizado. A constante γ reflete a hidrofobicidade do sítio de ligação que pode ser necessária para reprodução das energias livres absolutas, mas não é necessária para o cálculo da energia livre relativa (67).

1.7 Regressão Linear

A área de modelagem estatística de regressão recebeu grande impulso desde a criação dos modelos lineares generalizados no início da década de 70. No Brasil, a área começou a se desenvolver a partir de meados da década de 80 (68). A análise de regressão linear consiste na realização de uma análise estatística com o objetivo de verificar a existência de uma relação funcional entre uma variável dependente com uma ou mais variáveis independentes. Para tentar estabelecer uma equação que representa o fenômeno em estudo, um gráfico, chamado de diagrama de dispersão, pode ser construído para verificar como se comportam os valores da variável dependente (Y) em função da variação da variável independente (X) (69).

Para a realização desse estudo, escolhemos utilizar o programa estatístico R (70). Os comandos utilizados no programa assim como o significado das funções são explicados a seguir.

```
gp <- gm - (alfa * w + beta * e)
lm1 <- glm(gm ~ offset(l(alfa*w)) + offset(l(beta*e)))
lm2 <- glm(gm ~ offset(l(alfa*w)) + e)
lm3 <- glm(gm ~ w + offset(l(beta*e)))
lm4 <- glm(gm ~ w + e)
```

onde w indica os valores da interação de Van der Waals, e indica os valores da interação eletrostática, gm indica a função que recebe os valores experimentais dos complexos estudados e alfa e beta possuem valores constantes de 0,181 e 0,37, respectivamente. A função gp retorna o ΔG predito, a função lm1 mostra o resultado de gama quando alfa e beta estão fixos, lm2, quando alfa está fixo e beta está variando, lm3, quando alfa está variando e beta está fixo e lm4, quando alfa e beta estão variando.

2 Objetivos

2.1 Objetivo Geral

O objetivo deste trabalho é o estudo de novos compostos antimalariais, potenciais inibidores de *aspartil proteases*, que possam bloquear a atividade da enzima *Plasmepsina II (Plm II)*. Para atingir este objetivo, os itens da seção 2.2 foram planejados.

2.2 Objetivos Específicos

(1) Identificação do modo de ligação da série de compostos planejados pelo Departamento de Síntese Orgânica de Farmanguinhos (Instituto de Tecnologia em Fármacos – Fiocruz);

(2) Análise dos modos de ligação e das interações receptor-ligante dos compostos estudados;

(3) Avaliação da energia livre de ligação e determinação da afinidade de ligação dos compostos estudados.

3 Materiais e Métodos

Este trabalho foi realizado no Laboratório de Biofísica Computacional e Modelagem Molecular do Programa de Computação Científica na Fundação Oswaldo Cruz (Fiocruz). Para a primeira parte do estudo foram consideradas todas as estruturas tridimensionais da enzima *Plasmepsina II*, disponíveis no banco de dados de proteínas PDB (Protein Data Bank) (3), para alinhamento e análise da identidade estrutural. Para os experimentos de atracamento molecular foram utilizados os programas Autodock 4 (71), Autodock Vina (50) e DockThor (31, 51, 58), e para visualização dos complexos e elaboração das figuras foram usados os programas PyMOL (72), Rasmol (73) e VMD (74).

Para a segunda parte deste estudo, que aborda o cálculo da energia livre de ligação através de dinâmica molecular, foi utilizado o programa Q (75). O cálculo da energia livre, executado através do programa escolhido, foi realizado nas estações de trabalho do laboratório, que possuem processador Intel Core Quad (CPU Q6600, 2.40GHz) com quatro núcleos para processamento e 3Gb DDR3. O sistema operacional instalado nessas máquinas é o Ubuntu 11.04.

3.1 Métodos

Nesta seção será apresentada a metodologia adotada neste trabalho para a preparação das estruturas do receptor e dos ligantes, assim como os métodos de simulação de atracamento molecular e estudo da dinâmica dos complexos. A primeira parte abordará a preparação das moléculas. Na segunda parte será descrito a metodologia utilizada no método Linear Interaction Energy (LIE) (65) usado para o cálculo da energia livre de ligação.

3.1.1 Preparação do Alvo Molecular (receptor)

Com a estrutura experimental escolhida, o arquivo obtido do PDB foi separado em dois arquivos: um contendo somente as informações do receptor e outro somente as informações do ligante. Esses arquivos foram preparados separadamente e submetidos às simulações computacionais. Uma vez que o

estado de protonação dos aspartatos catalíticos é uma questão não resolvida (76), neste trabalho foram realizados experimentos com o resíduo Asp214 protonado e Asp34 desprotonado, de acordo com os estudos prévios de dinâmica molecular de aspartil proteases (77, 78). A figura 15 ilustra o resíduo Asp214 com protonação no oxigênio interno à proteína (O1).

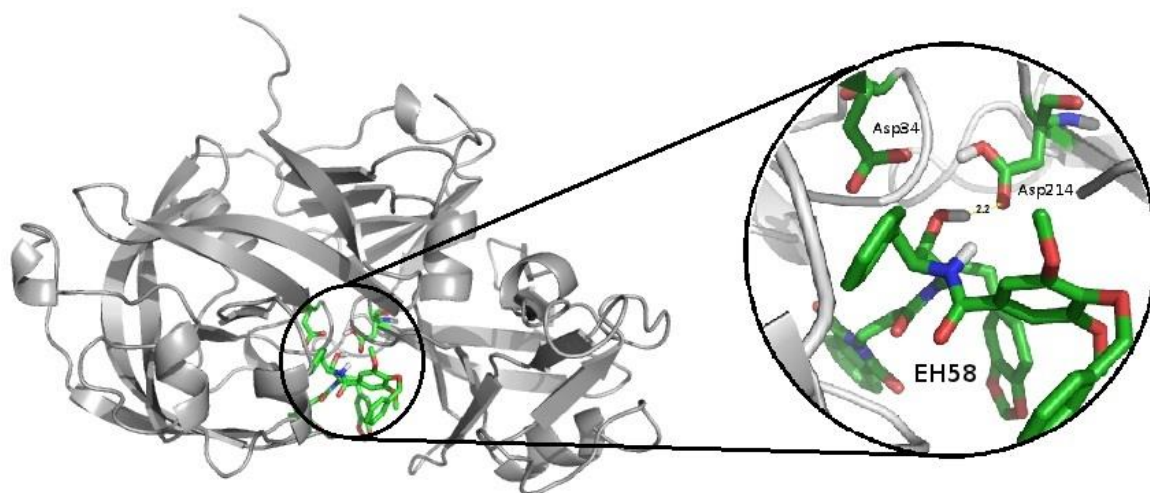


Figura 15. Aspartatos catalíticos com protonação no Asp214/O1

Os arquivos do receptor foram preparados de duas formas:

- 1 – Receptor com o resíduo Asp214 não protonado
- 2 – Receptor com átomo de oxigênio O1 do resíduo Asp214 protonado

Para cada uma das formas acima, foram gerados três arquivos com diferentes conjuntos de cargas parciais atômicas, de acordo com os campos de força abaixo:

- 1 – Gromos96 (79)
- 2 – Gasteiger (80)
- 3 – MMFF94 (81)

O conjunto de cargas Gromos96 do receptor foi gerada com o programa GROMACS (82). Para cargas Gasteiger, as conformações foram geradas pelo programa Autodock 4 através da ferramenta MGLTools (83). Com o programa MOE (84), que, entre outras funções, calcula a carga dos átomos e adiciona todos

os hidrogênios na molécula, o arquivo com cargas MMFF94 foi gerado para utilização nos programas Autodock 4 e Autodock Vina. A preparação do arquivo com cargas MMFF94 para simulações no DockThor, foi realizada com o programa Pdbthorbox (58, 85)⁴.

Para o receptor com os resíduos de ácidos aspárticos protonados, foi utilizado o script *prepare_receptor* do Autodock 4. Os arquivos do receptor com protonação no resíduo Asp214 e com os campos de força usados nas simulações foram gerados da maneira descrita anteriormente, sendo a protonação selecionada através do script *pdb2gmx* do programa GROMACS, para as simulações nos programas Autodock 4 e Autodock Vina. Para as simulações no programa DockThor, o estado de protonação foi definido através do programa Pdbthorbox.

3.1.2 Preparação dos Ligantes

Para os experimentos de *redocking* do ligante EH58 no sítio da Plm II foi utilizada sua conformação original na enzima à qual estava complexado. Foram preparados dois sistemas com diferentes conjuntos de cargas parciais atômicas: o primeiro com cargas Gasteiger e o segundo, com cargas MMFF94.

Para o primeiro sistema, a adição de todos os átomos de hidrogênio e o cálculo de cargas Gasteiger foram feitos com o programa Autodock 4, através da ferramenta MGLTools. Para o segundo, a adição de hidrogênios e de cargas MMFF94 foram feitas com o programa MOE. Entretanto, este programa não gera uma saída no formato requerido em programas de atracamento molecular (extensão *.pdbqt*). Para isso, foi utilizado o script *prepare_ligand* do programa Autodock 4. Para ambos arquivos, todas as torções disponíveis do ligante foram consideradas ativas.

As estruturas de todos os compostos planejados pelo Departamento de Síntese Orgânica de Farmanguinhos foram construídas e otimizadas utilizando o campo de força MMFF94 no programa MOE até a norma do gradiente atingir um valor inferior a 0,05 kcal/(mol.Å). A estrutura minimizada foi usada como modelo inicial para o *docking*.

⁴ A versão original do programa PDBTHORBOX foi criada por Laurent E. Dardenne, Ernesto R. Caffarena, Isabella Ortman e Michel Loos.

Para o programa DockThor, a preparação da topologia do ligante EH58 para *redocking* e da nova série de ligantes planejados pelo DSO foi feita com o programa mmffligand, implementado na dissertação de Almeida, 2011 (58). Este programa adiciona todos os átomos de hidrogênio na molécula, ativa todas as possíveis torções do ligante e calcula cargas MMFF94, já deixando o arquivo de saída no formato de entrada para as simulações (extensão .top).

3.1.3 Redocking

Para obtenção de um melhor conjunto de parâmetros a ser utilizado nos programas de atracamento molecular para o sistema escolhido, foram feitas análises de *redocking* do complexo *PlmII*-EH58 (código PDB: 1LF3). Nesses experimentos, o ligante, que foi retirado da região de ligação da conformação cristalográfica, é “recolocado” no sítio ativo da macromolécula através de programas de atracamento molecular. A partir da escolha do conjunto de parâmetros usados nesta simulação, é calculada a taxa de sucesso, definida como a razão entre o número de vezes que foi obtida a conformação próxima da posição do ligante cristalográfico e o número total de experimentos.

Os experimentos de *redocking* foram utilizados para avaliar 2 aspectos:

- Calibração de parâmetros;
- Modelos de carga e protonação/não protonação do resíduo Asp214;

Para o estabelecimento de um protocolo padrão, alguns parâmetros foram combinados em diversas simulações de atracamento molecular até que fosse obtido um conjunto de parâmetros que gerassem conformações mais próximas da estrutura obtida do PDB e que a média da energia de interação do conjunto de soluções fosse próxima à de menor energia obtida pelo algoritmo.

Para que o resultado do programa de *docking* seja “aceitável”, a estrutura de menor energia encontrada pelo algoritmo deveria apresentar um valor de raiz quadrada do desvio médio quadrático (RMSD) de até 3 Å com a estrutura cristalográfica do ligante. Quanto menor o valor encontrado, maior a chance dos parâmetros definidos no programa conseguirem reproduzir os dados conhecidos

experimentalmente. Esta análise de RMSD somente pode ser realizada em experimentos de *redocking* onde a estrutura experimental do complexo é conhecida. Portanto, o objetivo desses experimentos é a obtenção de um conjunto de parâmetros a serem utilizados nos experimentos de *docking* posteriores.

Para o programa Autodock 4, os parâmetros testados foram: tamanho da população, elitismo e número máximo de avaliações de energia. Os valores se detalham na Tabela 3. Os outros parâmetros foram mantidos com seus respectivos valores usuais. Para cada experimento de *redocking*, a dimensão da grade de energia foi de 25 Å em cada direção (x, y e z) com discretização de 0,375 Å estando a grade centralizada no ligante (Figura 16). Esse tamanho foi necessário para que todos os ligantes pudessem se acomodar no sítio da enzima.

Com o programa Autodock Vina, os parâmetros testados foram: *energy_range*, que é o valor máximo de diferença de energia em kcal/mol entre o melhor e o pior modos de ligação (menor energia) amostrados, e *exhaustiveness*, que é equivalente ao esforço empenhado pelo algoritmo de busca para encontrar um mínimo global próximo da conformação nativa da molécula ligante. Os valores são mostrados na Tabela 4.

Nos experimentos realizados com o programa DockThor, todas as variáveis mantiveram seus respectivos valores padrão para análise do estado de protonação do resíduo Asp214. As cargas MMFF94 foram utilizadas para proteína e ligante e o tamanho da grade foi definido por análise visual do ligante EH58 no sítio ativo do receptor sendo a caixa centralizada no ligante e tendo valores de 9,0 em cada direção X, Y e Z, sendo o tamanho total de 27Å. Os valores dos parâmetros são mostrados na Tabela 5.

Tabela 3. Valores dos parâmetros usados nos experimentos de *redocking* com Autodock 4

Variável	Valores analisados
População	150; 300; 500; 1000
Elitismo	1%; 5%; 10%
Número Máximo de Avaliações de Energia	$1,5 \times 10^6$; $2,5 \times 10^6$
Número de Execuções	150
Taxa de Mutação	0,02
Taxa de Crossover	0,8
Algoritmo	Algoritmo Genético Lamarckiano

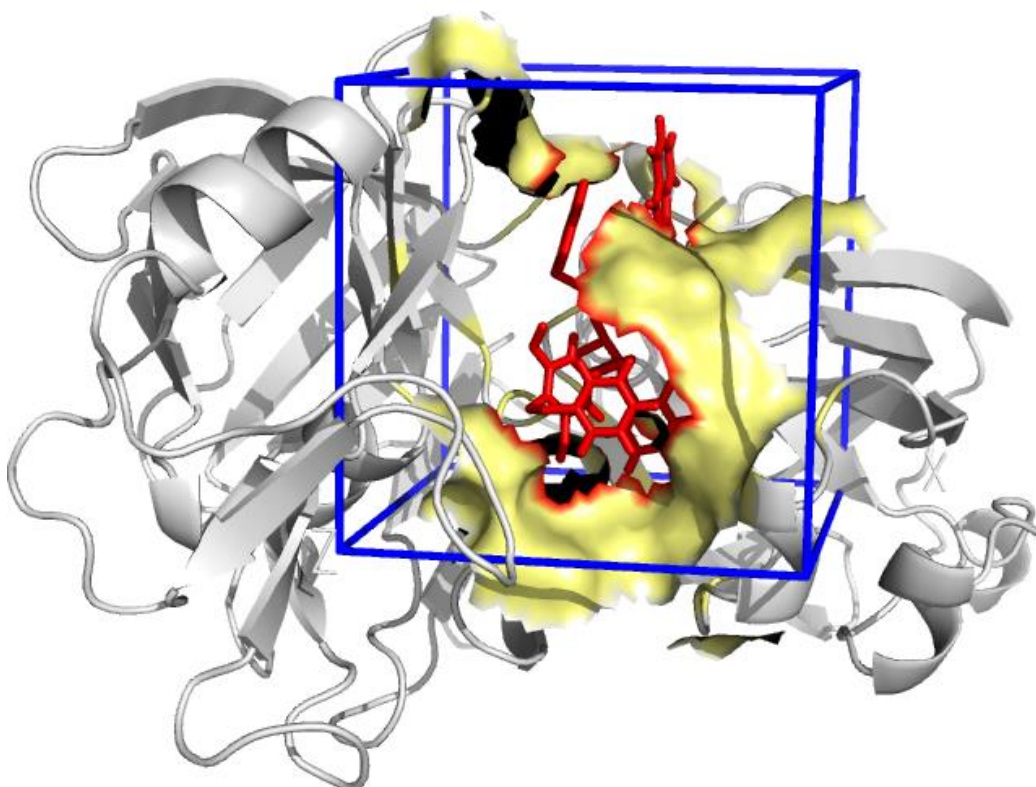


Figura 16. Representação da grade calculada pelo programa Autogrid. Em amarelo encontra-se a superfície acessível ao solvente da Plasmepsina II com até 4 Å de distância do ligante EH58, representado em vermelho.

Tabela 4. Valores dos parâmetros usados nos experimentos de redocking com Autodock Vina.

Variável	Valores analisados
Energy range	3; 5; 15; 50 (Kcal/mol)
Exhaustiveness	4; 8; 20; 50; 100; 120; 150; 170; 200
Número de Execuções	1
Número de conformações por execução	20

Tabela 5. Parâmetros usados nos experimentos de redocking com o programa Dockthor.

Variáveis	Valores utilizados
Número de Execuções	150
Tamanho da População	1000
Número de Avaliações de Energia	1000000
Tamanho da Caixa (x,y,z)	9.0 9.0 9.0
Semente	-1000

Para avaliação do segundo aspecto, foram realizados experimentos com diferentes modelos de carga tanto com a protonação do resíduo Asp214 quanto sem protonação. Com o programa Autodock 4, foram realizados quatro experimentos de *redocking* com diferentes configurações de campos de força para o receptor e o ligante conforme mostrado na Tabela 6. Para o programa Autodock Vina, foram realizados experimentos para avaliar a influência do estado de protonação no resíduo Asp214.

Tabela 6. Configuração do experimento em relação ao tipo de carga utilizados em redocking com Autodock 4.

Configuração	Carga Proteína	Carga Ligante
C1	Gromos96	MMFF94
C2	Gasteiger	Gasteiger
C3	Gasteiger	MMFF94
C4	MMFF94	MMFF94

Para o experimento com os diferentes modelos de cargas parciais atômicas, foi utilizado o melhor conjunto de parâmetros obtidos com a primeira análise de calibração, com exceção do número de execuções do programa, avaliando só dois valores: 20 e 150 e o número de avaliações de energia sendo usado os valores $1,5 \times 10^6$ e $2,5 \times 10^6$. A partir dos resultados destes experimentos, o conjunto de parâmetros que alcançou a maior taxa de sucesso foi utilizado para realização de um novo experimento com os mesmos parâmetros, porém considerando o resíduo Asp214 protonado e sem protonação.

3.1.4 Dinâmica Molecular aplicada ao Cálculo de Energia Livre de Ligação

As simulações de dinâmica molecular foram realizadas com a versão 5 do programa Q (86). Para cada estimativa de energia livre de ligação utilizando o método LIE (65) são necessárias duas simulações: uma delas com ligante em solução (ligante+solvente) e a outra no sítio da proteína formando o complexo (ligante+receptor+solvente).

Nesta seção será descrito como foram preparados os arquivos do receptor e ligantes para as simulações. Para o complexo, realizamos 900 ps de simulação

variando a temperatura de 50 a 310 K. Após a equilibração do sistema, foram realizados 5 ns de simulação à temperatura de 310 K completando a fase de produção. Para o cálculo da energia livre de ligação foram utilizados os últimos 2,5 ns da simulação para garantir que o cálculo fosse efetuado com o sistema em equilíbrio. A Figura 17 mostra o RMSD da proteína de código 1LF3 durante o tempo simulado. As trajetórias foram salvas a cada 0,5 ps. O sistema em solução também foi simulado em 5 ns à temperatura de 310 K.

Os coeficientes de calibração α , β e γ utilizados no método LIE para o cálculo da energia livre de ligação foi 0,181 e 0,37 para alfa e beta, respectivamente, sendo o gama ajustado para reproduzir o resultado experimental. Contudo, o valor de gama varia de acordo com os parâmetros definidos na regressão linear.

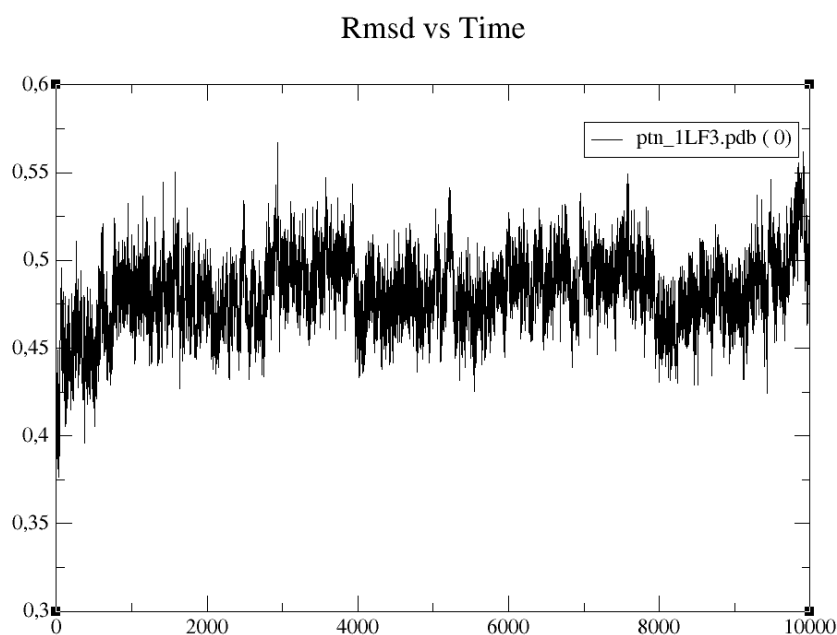


Figura 17. RMSD ao longo do tempo na simulação do complexo de código 1LF3

3.1.4.1 Preparação do receptor

Para que o custo computacional envolvido no cálculo do potencial eletrostático de longo alcance seja menor, o programa Q utiliza uma abordagem que reduz o número de átomos simulados confinando o sistema em uma esfera

de raio delimitado, considerando apenas as interações ligadas dentro do raio dessa esfera.

A esfera usada para as simulações foi delimitada em 18 Å sendo centralizada no átomo central do ligante. A Figura 18 representa esta abordagem. Foi verificado quais resíduos de aminoácidos carregados estavam presentes dentro do raio de 10 Å para que os mesmos se mantivessem carregados. O resíduo Asp214 foi mantido protonado para todas as simulações e foi definido como ASH no arquivo de topologia do programa. Os aminoácidos Arg, Asp, Lys e Glu foram modificados de acordo com cada sistema para AR+, AS-, LY+ e GL- mantendo-as na sua forma neutra àqueles que se encontravam entre a esfera de raio 10 Å e a esfera de raio 18 Å.

Após a identificação e alteração dos resíduos carregados no interior da esfera, foi utilizado o script *Qprep* do programa Q para preparação do arquivo de topologia do receptor sendo usado o campo de força Gromos96. O parâmetro *solvent_pack*, que é a distância mínima entre soluto e o solvente de átomos pesados durante a adição de solvente, foi fixado em 2,3 Å. O parâmetro *boundary* foi definido para simulação do sistema em uma esfera.

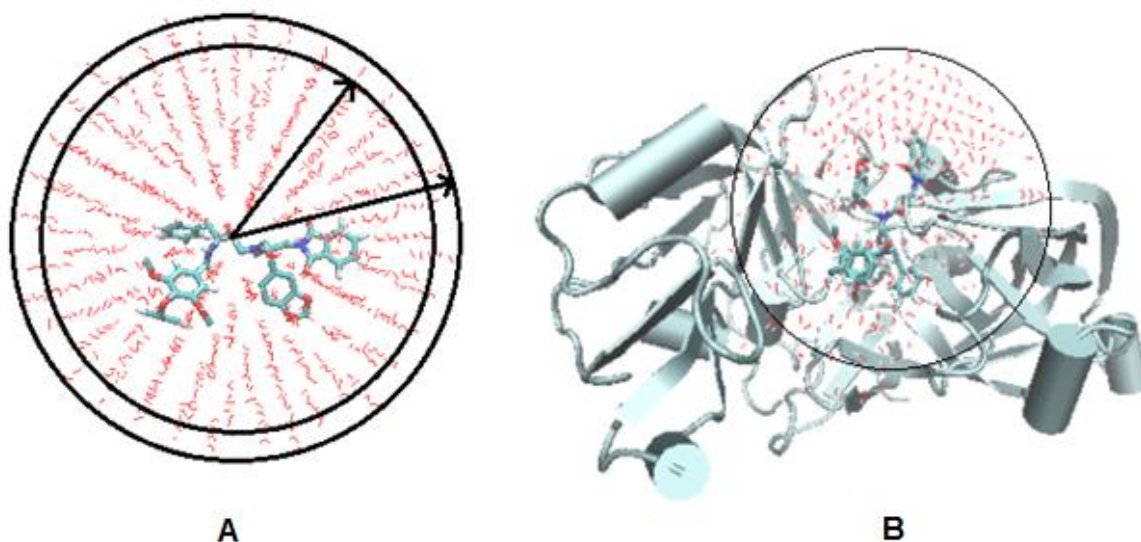


Figura 18. Modelo representativo do complexo construído para os cálculos de energia livre. Em A, o sistema ligante+solvente. O resíduo representado por bastões no centro da esfera representa o ligante de referência. As setas indicam o raio externo da esfera, onde os átomos foram mantidos fixos por restrições harmônicas. Em B, o sistema ligante+receptor+solvente.

3.1.4.2 Preparação dos ligantes

A preparação do ligante EH58 foi realizada a partir da estrutura experimental obtida do PDB. Para os ligantes da série Plip, a estrutura inicial usada nas simulações foi a estrutura de menor energia obtida através dos experimentos de *docking* molecular. Para a adição de cargas atômicas foi utilizado o servidor Automated Topology Builder (ATB) (87), que faz a construção automática da topologia da molécula a partir do método escolhido para o cálculo das cargas. Este servidor disponibiliza 3 métodos para otimização da molécula: PM3 (88), AM1 (89) e HF/STO-3G (90). De acordo com trabalhos publicados por Sant'anna, C. M. R. (91), resultados demonstraram que o desempenho do método PM3 é melhor aos demais citados.

PM3 foi introduzido em 1985 como um modelo semiempírico mais acurado apresentando baixos resultados de média de erros comparados à AM1 principalmente na entalpia de formação e, portanto, foi o escolhido para utilização nos ligantes deste trabalho já que também é o método recomendado para cálculo de cargas para ligantes parametrizados de acordo ao campo de força Gromos (87). Além disso, o método PM3 é melhor para ligações hidrogênio.

O servidor ATB possui uma limitação com relação a quantidade de átomos da molécula aceitando somente moléculas com até 99 átomos. Porém, a partir de 40 átomos, o servidor só estima a topologia inicial da molécula. Sendo assim, todos os ligantes foram divididos em partes e submetidos ao servidor para cálculo das cargas parciais atômicas. Com o resultado do servidor, foi realizada a junção das partes da molécula e a divisão em pequenos grupos de carga. Para que o ligante se mantivesse neutro, os valores atribuídos aos átomos de hidrogênio foram ajustados. Após estes ajustes, o arquivo .lib com as informações de cargas para cada átomo, lista das ligações atômicas e a separação dos átomos em grupos de cargas está pronto para ser executado pelo script *Qprep*. Os parâmetros usados neste script como, *solvent_pack* e *boundary*, foram mantidos iguais à preparação do arquivo do receptor.

3.1.5 Métodos para Análise dos Resultados

3.1.5.1 RMSD

O RMSD é a medida da similaridade entre estruturas moleculares. Este cálculo é realizado através da média da distância entre duas estruturas dado como entrada as coordenadas dos átomos da molécula. A equação 9 mostra a fórmula usada neste cálculo.

$$RMSD(v, w) = \sqrt{\frac{1}{n} \sum_{i=1}^n (v_{ix} - w_{ix})^2 + (v_{iy} - w_{iy})^2 + (v_{iz} - w_{iz})^2} \quad (9)$$

onde n é o número total de átomos e v e w representam as estruturas envolvidas. Para calcular a distância entre a estrutura tridimensional do ligante EH58 e a estrutura das moléculas dadas como solução do *redocking*, foi desenvolvido um script na linguagem Perl (92), que faz a leitura das coordenadas X, Y e Z de cada molécula e armazena os valores de cada átomo em uma estrutura de vetores. Para que a solução seja válida, o valor resultado do cálculo do RMSD deve ser menor ou igual a 3 Å.

3.1.5.2 Seleção das Conformações

Para selecionar as soluções resultantes do *docking*, foram definidos 2 critérios:

1 – Distância de até 3,5 Å entre o átomo de oxigênio da hidroxila do ligante e um dos oxigênios dos resíduos aspárticos;

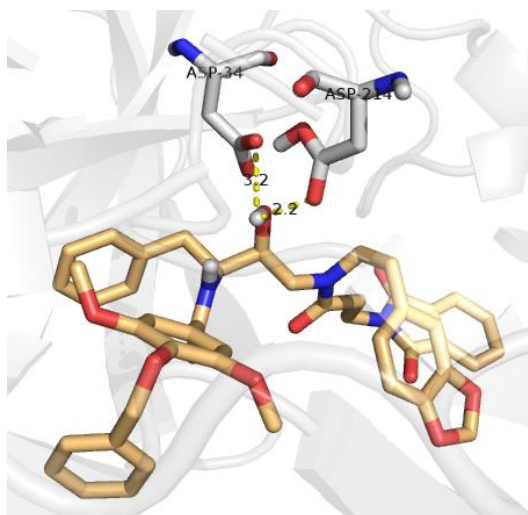


Figura 19. Interação do ligante EH58 com os resíduos Asp34 e Asp214. A distância entre os átomos é dada em Ångstrom.

Este filtro é realizado para selecionar as conformações que possuem o grupo hidroxila próximo a esses resíduos, já que são os resíduos com maior importância na interação receptor-ligante nesta família.

2 – Separação das soluções em grupos (clusterização).

Após a separação das soluções que possuem interação com pelo menos um dos resíduos aspárticos, estas soluções são organizadas por ordem de menor energia. A primeira solução, ou seja, aquela conformação que interage com os resíduos aspárticos e possui a menor energia dentre todas as moléculas é considerada líder do grupo. Tendo esta conformação como modelo, é calculado o RMSD entre esta e as demais estruturas. O grupo será definido por todas as estruturas que possuem $\text{RMSD} < 2,0 \text{ \AA}$ em relação ao líder. Das conformações remanescentes é escolhido o líder do próximo grupo e o mesmo procedimento é realizado até que todas as conformações sejam agrupadas. A análise foi realizada somente com os líderes de cada grupo.

3.1.5.3 Taxa de Sucesso

A taxa de sucesso é calculada pela divisão entre a quantidade de conformações que obtiveram RMSD menor que 3 \AA sobre todas as conformações obtidas.

4 Resultados e Discussão

Este capítulo está dividido em duas seções. Na primeira seção abordaremos os resultados dos experimentos de *redocking* do ligante EH58 para avaliação em três aspectos: (1) conjunto de parâmetros dos programas de atracamento molecular que melhor se adequa ao sistema em estudo (seção 4.1.1), (2) o modelo de carga utilizado pelos programas de atracamento molecular para a proteína e o ligante (seção 4.1.2) e (3) protonação do aspartato (seção 4.1.2).

Na segunda seção apresentaremos os resultados obtidos com o *docking* dos novos compostos planejados pelo DSO utilizando o protocolo validado pelos experimentos anteriores. Finalmente, apresentaremos os resultados da dinâmica molecular com a ordenação dos novos compostos juntamente com uma comparação dos valores experimentais de ΔG com a energia livre de ligação calculada.

4.1 Redocking do ligante EH58

Esta seção está dividida em três partes: (1) Calibração dos parâmetros utilizando os programas Autodock 4 e Autodock Vina; (2) Modelos de Carga e Protonação do resíduo Asp214 e (3) Uso do programa Dockthor.

4.1.1 Calibração dos parâmetros utilizando os programas Autodock 4 e Autodock Vina

Os primeiros experimentos de *redocking* para calibração dos parâmetros foram realizados com a protonação no O1 do resíduo Asp214 de acordo com Friedman et al., (93). A Tabela 7 mostra os resultados obtidos com o programa Autodock 4 utilizando $1,5 \times 10^6$ avaliações de energia com a variação do tamanho da população e % de elitismo (Tabela 3).

Para cada experimento, o programa foi executado 150 vezes com diferentes sementes. Para análise deste protocolo, a taxa de sucesso foi calculada em relação ao valor de RMSD entre a estrutura de menor energia e a estrutura experimental do ligante EH58. A taxa de sucesso foi calculada com base

no número de soluções encontradas com RMSD menor que 3,0 Å. Em todos os experimentos, só foram consideradas válidas para cômputo da taxa de sucesso, soluções capazes de formar ligações hidrogênio entre o ligante e pelo menos um dos resíduos aspárticos catalíticos (Asp34 ou Asp214).

O resultado deste cálculo mostrou, para todos os experimentos, que a solução de menor energia obtida não conseguiu encaixar o ligante na posição da estrutura obtida experimentalmente, visto que os valores de RMSD obtidos foram maiores que 3 Å. Com relação ao critério de ligação hidrogênio com os resíduos aspárticos catalíticos, menos de 5% de todas as soluções encontradas satisfizeram este critério.

Tabela 7. Redocking do ligante EH58 com o programa Autodock 4 com $1,5 \times 10^6$ de avaliações de energia.

Tam. Pop.	Elitismo (%)	Menor Energia ^(a)	Energia Média ^(a)	RMSD da conf. menor energia ^(b)	Menor RMSD ^(b)	RMSD Médio ^(b)	T.S. RMSD $\leq 3,0$ ^(c)	Interação Ligante – Asp ^(d)
150	1	-4,90	-1,70 \pm 0,98	10,53	5,11	10,22 \pm 1,75	0%	2%
	5	-5,42	-2,83 \pm 0,98	11,10	5,55	10,33 \pm 1,59	0%	2%
	10	-6,19	-3,20 \pm 0,95	4,62	3,66	10,43 \pm 1,61	0%	0,67%
300	1	-4,19	-1,01 \pm 0,96	4,44	4,43	10,58 \pm 1,79	0%	1,33%
	5	-6,38	-2,43 \pm 1,16	5,04	4,42	10,40 \pm 1,84	0%	4%
	10	-5,93	-2,94 \pm 1,01	11,85	4,36	10,08 \pm 1,82	0%	2%
500	1	-4,75	-0,59 \pm 0,85	4,90	4,67	10,48 \pm 1,92	0%	4%
	5	-5,07	-1,75 \pm 1,00	10,09	4,90	10,40 \pm 1,88	0%	0,67%
	10	-4,67	-2,44 \pm 0,94	6,22	5,08	10,40 \pm 1,83	0%	4,67%
1000	1	-2,11	0,13 \pm 0,73	9,54	4,75	10,62 \pm 1,72	0%	0%
	5	-3,71	-0,81 \pm 0,84	8,87	3,85	10,51 \pm 1,94	0%	2%
	10	-4,48	-1,31 \pm 0,91	2,93	2,93	10,37 \pm 1,93	0,67%	4,67%

Experimento realizado com $1,5 \times 10^6$ de avaliações de energia. ^(a)Energia dada em kcal/mol; ^(b)RMSD dado em Ångstroms; ^{(c),(d)} Taxa de Sucesso e % de conformações que interagem com Asp34/Asp214 calculadas com as 150 soluções obtidas.

Nenhuma das configurações de parâmetros com o programa Autodock 4 obteve resultado satisfatório. Uma análise visual das conformações foi realizada sobrepondo os modos de ligação encontrados com a estrutura experimental do ligante EH58. As conformações de menor energia que foram obtidas com o valor de elitismo de 1% e tamanhos de população de 150, 300, 500 e 1000 são mostradas na Figura 20. É possível observar que as soluções se posicionaram no sítio ativo da macromolécula e sobrepuseram os grupamentos do ligante EH58, com exceção da conformação em D. Porém, o algoritmo não conseguiu encontrar

a mesma posição relativa dos grupamentos da estrutura cristalográfica, o que justifica os altos valores de RMSD mostrados na Tabela 7.

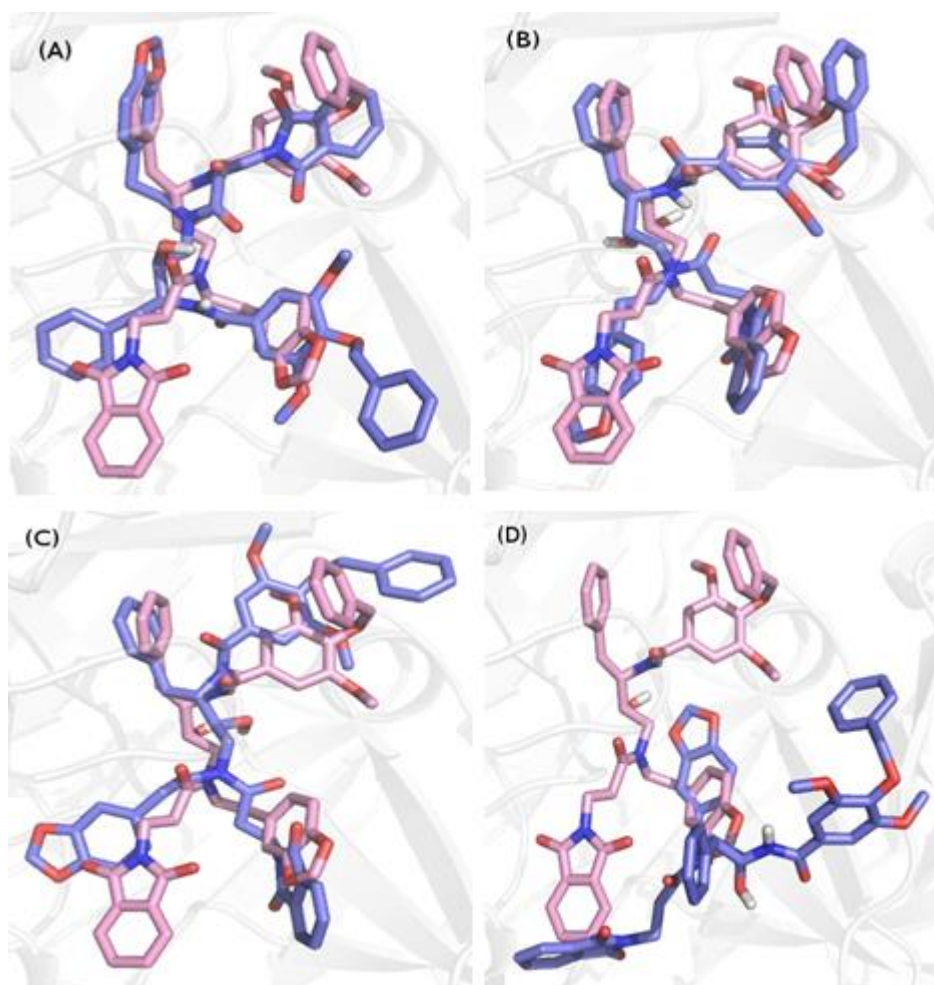


Figura 20. Sobreposição da estrutura experimental do ligante EH58 (em rosa), e a solução de menor energia (em azul). Conformações obtidas com elitismo igual a 1% e tamanho da população (A) 150, (B) 300, (C) 500 e (D) 1000. O valor de RMSD nas soluções A, B, C e D é 5,11, 4,43, 4,90 e 4,75 Å, respectivamente.

Para verificar se os parâmetros testados no programa Autodock 4, de fato, não conseguem reproduzir o sistema deste estudo, novas simulações foram realizadas aumentando-se o número máximo de avaliações de energia para $2,5 \times 10^6$. Os resultados são mostrados na Tabela 8.

Tabela 8. Redocking do ligante EH58 com o programa Autodock 4 com $2,5 \times 10^6$ de avaliações de energia.

Tam. Pop.	Elitismo (%)	Menor Energia ^(a)	Energia Média ^(a)	RMSD da conf. menor energia ^(b)	Menor RMSD ^(b)	RMSD Médio ^(b)	T.S. RMSD $\leq 3,0$ ^(c)	Interação Ligante – Asp ^(d)
150	1	-6,16	-2,15 \pm 1,00	4,18	4,17	10,85 \pm 1,78	0%	5,33%
	5	-7,02	-3,19 \pm 1,13	5,05	3,82	10,38 \pm 1,82	0%	4,67%
	10	-5,71	-3,25 \pm 1,08	11,01	4,60	10,45 \pm 1,72	0%	0%
300	1	-4,26	-1,48 \pm 0,95	5,13	4,40	10,24 \pm 1,80	0%	2%
	5	-5,67	-2,85 \pm 1,17	9,42	3,28	10,22 \pm 1,80	0%	2,67%
	10	-6,87	-3,28 \pm 1,12	4,57	4,56	10,46 \pm 1,86	0%	2%
500	1	-3,76	-1,17 \pm 0,88	5,02	4,76	10,37 \pm 1,82	0%	4%
	5	-6,02	-2,36 \pm 1,00	4,98	4,07	10,38 \pm 1,90	0%	3,33%
	10	-6,54	-2,83 \pm 1,08	11,98	4,91	10,41 \pm 1,55	0%	1,33%
1000	1	-2,68	-0,47 \pm 0,82	10,28	4,40	10,49 \pm 1,79	0%	0,67%
	5	-5,25	-1,55 \pm 1,04	11,44	5,69	10,39 \pm 1,71	0%	3,33%
	10	-4,25	-1,98 \pm 0,81	8,39	4,67	10,49 \pm 1,83	0%	1,33%

Experimento realizado com $2,5 \times 10^6$ de avaliações de energia. ^(a)Energia dada em kcal/mol; ^(b)RMSD dado em Ångstroms; ^{(c),(d)} Taxa de Sucesso e % de conformações que interagem com Asp34/Asp214 calculadas com as 150 soluções obtidas.

Estes experimentos mostraram que mesmo com o aumento do número máximo de avaliações de energia, o programa não conseguiu encontrar soluções próximas da estrutura experimental, o que é confirmado pelos valores de RMSD obtidos.

Uma análise visual foi realizada sobrepondo as soluções de menor energia com a estrutura do ligante EH58 obtido do PDB. Todas as soluções de menor energia se sobrepuseram à estrutura original do ligante, porém, não houve correspondência em relação aos grupamentos da molécula. O resultado deste experimento foi similar ao obtido anteriormente (Tabela 7). Por se tratar de um ligante com muitos graus de liberdade conformacionais, os algoritmos dos programas de atracamento molecular possuem maior dificuldade em encontrar a conformação do ligante observada experimentalmente.

A Figura 21 ilustra a solução de menor energia em que o átomo de oxigênio do grupamento hidroxil central do ligante interage com os resíduos Asp34 e Asp214.

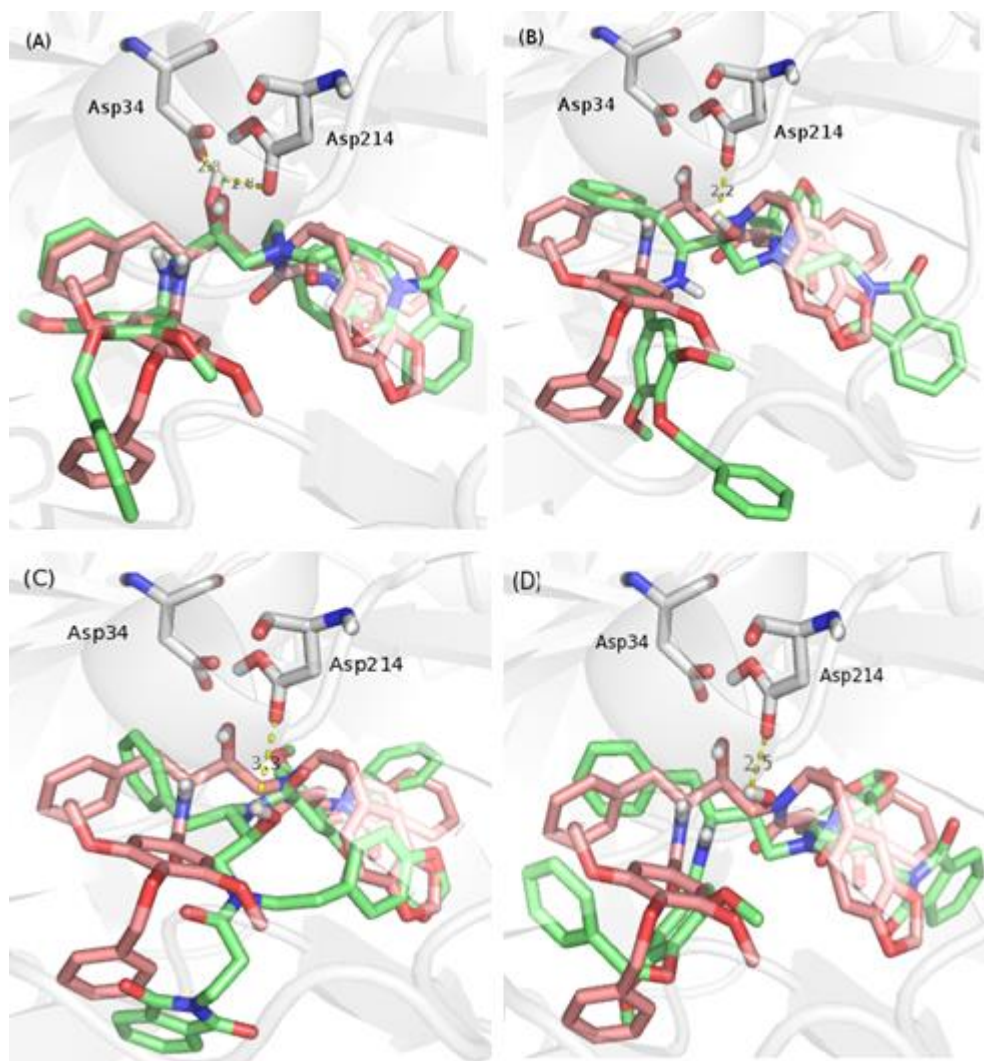


Figura 21. Sobreposição da estrutura experimental do ligante EH58 (representado em rosa) com a solução de menor energia obtida após filtro das conformações (representado em verde), i.e., possuem pelo menos 1 ligação hidrogênio com um dos resíduos aspárticos. As conformações selecionadas são as soluções de menor energia dos experimentos com a variável elitismo em 1% e tamanho da população (A) 150, (B) 300, (C) 500 e (D) 1000. O valor de RMSD entre as estruturas docadas e a cristalográfica foi (A) 4,17, (B) 5,50, (C) 10,07 e (D) 4,40 Å. A distância entre os átomos está pontilhada em amarelo e o valor é dado em Å.

Nos experimentos com o programa Autodock 4 não foi obtido um conjunto de parâmetros que gerassem soluções próximas à estrutura experimental. Sendo assim, novas simulações com variação do modelo de carga utilizado para a proteína e para o ligante foram realizadas e estão descritas na seção 4.1.2.

Para avaliar o sistema estudado com o programa Autodock Vina foram realizados experimentos com variação nos parâmetros de entrada deste programa. Os resultados são mostrados na Tabela 9.

Tabela 9. Redocking do ligante EH58 utilizando o programa Autodock Vina.

Energy range	Exhaust.	Menor Energia ^(a)	Energia Média ^(a)	RMSD da Conf. Menor energia ^(b)	Menor RMSD ^(b)	RMSD Médio ^(b)	T.S. RMSD $\leq 3,0$ ^(c)	Interação Ligante – Asp ^(d)
3	4	-8,4	-7,36 ± 0,48	11,45	6,52	9,66 ± 1,91	0%	0%
	8	-8,7	-7,93 ± 0,44	9,52	6,52	10,58 ± 1,44	0%	0%
	20	-8,4	-7,99 ± 0,19	9,67	3,84	9,24 ± 2,57	0%	25%
	50	-10,0	-9,50 ± 0,21	4,22	2,15	7,22 ± 3,83	15%	55%
	100	-10,1	-9,60 ± 0,18	4,10	1,53	6,17 ± 3,87	25%	65%
	120	-10,1	-9,59 ± 0,18	4,16	1,61	4,48 ± 2,89	35%	85%
	150	-10,1	-9,59 ± 0,18	4,16	1,61	4,48 ± 2,89	35%	85%
	170	-10,1	-9,60 ± 0,17	4,16	1,61	4,49 ± 2,89	35%	85%
5	200	-10,1	-9,60 ± 0,17	4,16	1,61	4,60 ± 2,85	30%	85%
	4	-8,4	-7,36 ± 0,48	11,45	6,52	9,66 ± 1,91	0%	0%
	8	-8,7	-7,93 ± 0,44	9,52	6,52	10,58 ± 1,44	0%	0%
	20	-8,4	-7,99 ± 0,19	9,67	3,84	9,24 ± 2,57	0%	25%
	50	-10,0	-9,50 ± 0,21	4,22	2,15	7,22 ± 3,83	15%	55%
	100	-10,1	-9,56 ± 0,20	4,16	1,61	4,42 ± 2,86	40%	85%
	120	-10,1	-9,59 ± 0,18	4,16	1,61	4,48 ± 2,89	35%	85%
	150	-10,1	-9,59 ± 0,17	4,16	1,61	4,48 ± 2,89	35%	85%
15	170	-10,1	-9,60 ± 0,17	4,16	1,61	4,49 ± 2,89	35%	85%
	200	-10,1	-9,60 ± 0,17	4,16	1,61	4,60 ± 2,85	30%	85%
	4	-8,5	-7,40 ± 0,50	11,36	6,55	9,54 ± 1,68	0%	0%
	8	-8,5	-7,82 ± 0,31	11,36	7,67	10,41 ± 1,08	0%	5%
	20	-10,1	-9,21 ± 0,31	4,18	1,76	7,04 ± 3,84	20%	55%
	50	-10,0	-9,53 ± 0,19	4,19	1,76	5,10 ± 3,31	35%	75%
	100	-10,1	-9,56 ± 0,20	4,16	1,61	4,42 ± 2,86	40%	85%
	120	-10,1	-9,59 ± 0,18	4,16	1,61	4,48 ± 2,89	35%	85%
50	150	-10,0	-9,60 ± 0,14	4,15	1,83	5,64 ± 3,61	30%	70%
	170	-10,0	-9,60 ± 0,14	4,15	1,83	5,64 ± 3,60	30%	70%
	200	-10,0	-9,64 ± 0,17	4,15	1,83	5,63 ± 3,60	30%	70%
	4	-8,5	-7,48 ± 0,50	11,36	6,55	9,54 ± 1,68	0%	0%
	8	-8,5	-7,82 ± 0,31	11,36	7,67	10,41 ± 1,08	0%	5%
	20	-10,1	-9,21 ± 0,31	4,18	1,76	7,04 ± 3,84	20%	55%
	50	-10,0	-9,53 ± 0,19	4,19	1,76	5,10 ± 3,31	35%	75%
	100	-10,1	-9,56 ± 0,20	4,16	1,61	4,42 ± 2,86	40%	85%
50	120	-10,0	-9,59 ± 0,17	4,15	2,43	6,20 ± 3,74	25%	65%
	150	-10,0	-9,60 ± 0,14	4,15	1,83	5,64 ± 3,61	30%	70%
	170	-10,0	-9,60 ± 0,14	4,15	1,83	5,64 ± 3,60	30%	70%
	200	-10,0	-9,64 ± 0,17	4,15	1,83	5,63 ± 3,60	30%	70%

^(a)Energia dada em kcal/mol; ^(b) RMSD dado em Ångstroms; ^{(c),(d)} Taxa de Sucesso e % de conformações que interagem com Asp34/Asp214 calculadas com as 20 soluções obtidas.

Os resultados mostraram que o programa encontrou soluções com valor de menor energia próximos ao valor de energia média entre as soluções de cada experimento. Porém, o algoritmo não foi capaz de encontrar uma conformação próxima da estrutura experimental (RMSD < 3,0 Å) que ao mesmo tempo fosse a solução de menor energia. Algumas soluções apresentaram troca na posição dos

grupamentos moleculares quando sobrepostas com a estrutura cristalográfica do ligante EH58, da mesma forma que as soluções obtidas com o programa Autodock 4. Ainda assim, experimentos com o programa Autodock Vina conseguiram encontrar conformações com RMSD menor que 2,0 Å quando comparadas à estrutura experimental.

Com relação aos valores definidos para o parâmetro *energy_range*, não houve alteração significativa nos resultados. Por outro lado, o parâmetro *exhaustiveness*, só apresentou conformações com RMSD menor que 3,0 Å, quando seu valor foi maior ou igual a 50, para valores de *energy_range*, 3 e 5, e, maior ou igual a 20, para valores de *energy_range* iguais a 15 e 50.

Uma análise visual das conformações foi realizada com relação à (A) posição das soluções no sítio ativo do receptor e (B) interação com os resíduos Asp34 e Asp214, mostradas nas Figuras 22 e 23.

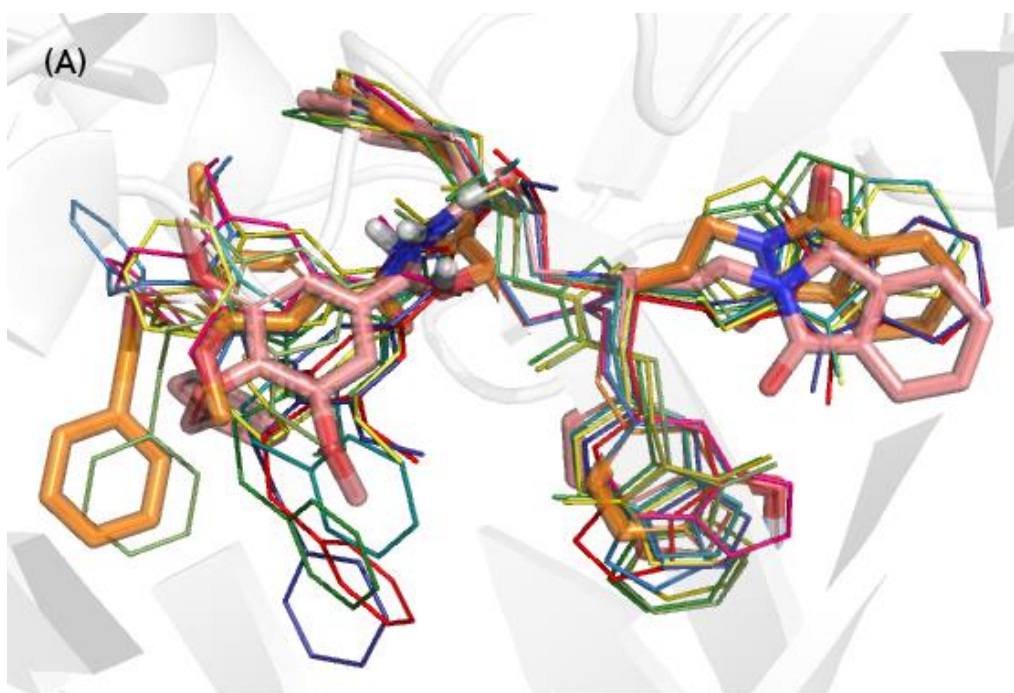


Figura 22. Sobreposição da estrutura experimental do ligante EH58 (em rosa) com as 10 soluções de menor RMSD (representados por linhas em outras cores). Experimento realizado com *energy_range* = 5 e *exhaustiveness* = 150. A estrutura representada em laranja é a conformação encontrada mais próxima da conformação experimental (RMSD de 1,61 Å).

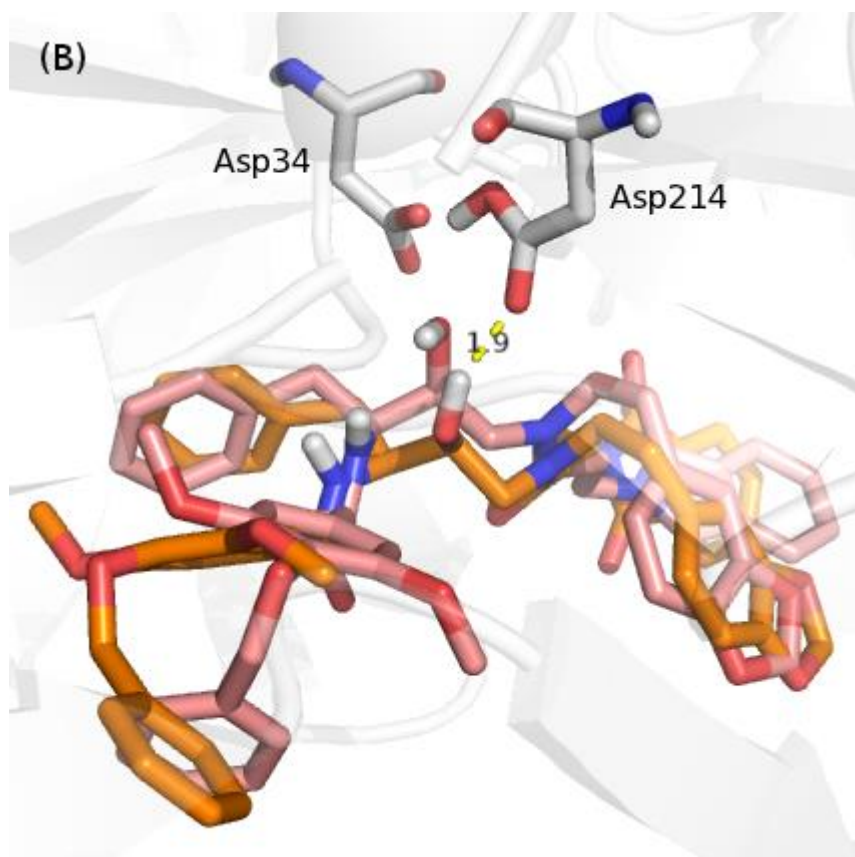


Figura 23. Ligação hidrogênio entre o hidrogênio da hidroxila do carbono central do ligante (em laranja) e o oxigênio do resíduo Asp214. Solução de menor energia, representada em laranja, e estrutura experimental do ligante EH58, representado em rosa.

A comparação entre as conformações dos ligantes que apresentaram menor energia em cada experimento de *redocking* com a estrutura experimental mostrou que o algoritmo encontrou soluções com uma boa superposição das conformações dos ligantes no sítio ativo da macromolécula tendo preenchido todos os grupamentos do ligante. Entretanto, por se tratar de um ligante altamente flexível, os valores médios de RMSD obtidos foram maiores ou iguais a 10,08 Å para o programa Autodock 4 e 4,42 Å para o programa Autodock Vina.

A comparação entre os experimentos de *redocking* realizados com a variação de parâmetros para calibração dos programas Autodock 4 e Autodock Vina mostrou que o programa Autodock Vina se adapta melhor ao sistema em estudo pois foi o que obteve taxa de sucesso de 40% (RMSD < 3,0 Å) dentre todas as conformações obtidas como resultado. Já o programa Autodock 4, apresentou 0% de sucesso em todos os experimentos realizados.

4.1.2 Modelos de Carga e Protonação do resíduo Asp214

Mesmo com a variação dos parâmetros do programa Autodock 4, não foi possível obter um bom resultado. Sendo assim, a segunda parte dos experimentos de *redocking* consistiu na avaliação deste programa com relação aos modelos de carga dos campos de força definidos para proteína e ligante. Os parâmetros escolhidos para estas simulações foram baseados nos resultados de *redocking* da seção 4.1.1. Os experimentos foram realizados com o valor de 150 para o tamanho da população, 1 % para elitismo e $1,5 \times 10^6$ avaliações de energia. Estes valores foram selecionados por apresentarem melhores resultados na análise visual em comparação aos outros conjuntos de parâmetros testados.

Quatro modelos de cargas foram escolhidos para serem avaliados com o programa Autodock 4 (Tabela 6). Para cada configuração foram realizados experimentos com o resíduo Asp214 protonado e também sem protonação. A Tabela 10 mostra o resultado obtido de cada experimento realizado sem a protonação do Asp214.

Tabela 10. *Redocking utilizando diferentes configurações de campos de força com o programa Autodock 4 sem a protonação no resíduo Asp214.*

Caso Teste ^(a)	Menor Energia ^(b)	Energia Média ^(b)	RMSD da Conf. Menor energia ^(c)	Menor RMSD ^(c)	RMSD Médio ^(c)	T.S. RMSD $\leq 3,0$ ^(d)	Interação Ligante – Asp ^(e)
C1	-3,49	-1,74 ± 0,94	9,38	5,49	10,17 ± 2,01	0%	5%
C2	-10,51	-5,93 ± 1,99	5,18	5,12	9,48 ± 2,20	0%	15%
C3	-4,19	-2,30 ± 0,87	8,60	6,09	10,11 ± 1,82	0%	0%
C4	1,00	2,53 ± 0,82	10,26	8,94	11,74 ± 1,73	0%	0%

^(a) C1 – Carga Ptn: Gromos96; Carga Lig: MMFF94 / C2 – Carga Ptn: Gasteiger; Carga Lig: Gasteiger / C3 – Carga Ptn: Gasteiger; Carga Lig: MMFF94 / C4 – Carga Ptn: MMFF94; Carga Lig: MMFF94. ^(b)Energia dada em kcal/mol; ^(c)RMSD dado em Ångstroms; ^{(d),(e)} Taxa de Sucesso e % de conformações que interagem com Asp34/Asp214 calculadas com as 20 soluções obtidas.

É possível observar que o algoritmo não conseguiu encontrar nenhuma conformação próxima da estrutura experimental tendo em vista que os valores de RMSD médios e de menor energia obtidos foram maiores que 5 Å.

Uma segunda análise desse resultado foi realizada usando o critério da ligação hidrogênio entre o OH central do ligante e os resíduos aspárticos catalíticos para avaliar todas as soluções de cada experimento. Nesse sentido,

com a configuração C2, 15% das soluções formam ligação hidrogênio com pelos menos um dos resíduos aspárticos do sítio catalítico. Além da configuração C2 apresentar soluções que interagem com os resíduos Asp34 e Asp214 e encontrar soluções com menores valores de RMSD comparados com os outros experimentos, esta também foi a configuração que obteve o menor valor de energia de interação (-10,51 kcal/mol).

Uma análise visual também foi realizada sobrepondo a solução de menor energia de cada configuração com os diferentes modelos de carga dos campos de força com o ligante EH58. As mesmas são apresentadas na Figura 24.

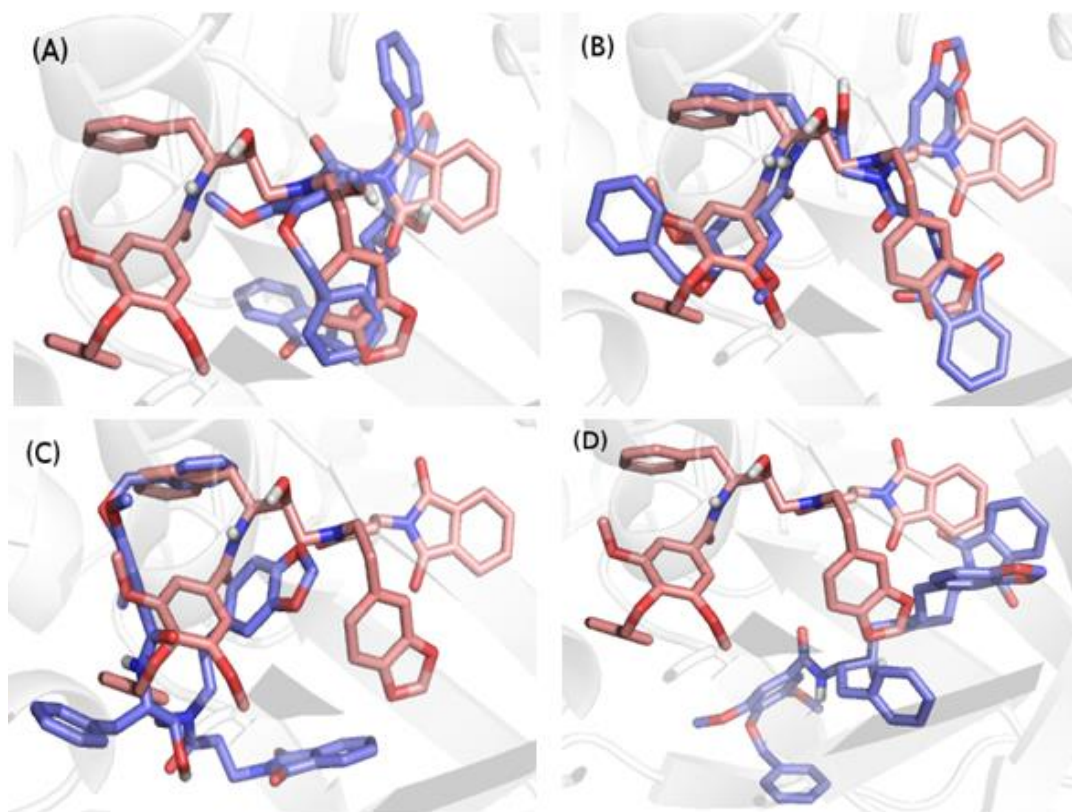


Figura 24. Sobreposição da solução de menor energia da configuração (A) C1 – Carga Ptn: Gromos96; Carga Lig: MMFF94, (B) C2 – Carga Ptn: Gasteiger; Carga Lig: Gasteiger, (C) C3 – Carga Ptn: Gasteiger; Carga Lig: MMFF94 e (D) C4 – Carga Ptn: MMFF94; Carga Lig: MMFF94, representadas em azul, e estrutura experimental do ligante EH58, representado em rosa. O valor de RMSD entre o ligante EH58 e as estruturas A, B, C e D é 9,38, 5,18, 8,60 e 10,26 Å, respectivamente.

É possível observar que as soluções das configurações de modelos de cargas C1, C3 e C4, representadas na Figura 24 como A, C e D respectivamente, não preencheram todos os sítios do ligante EH58. A configuração do modelo de carga C2, representada pela imagem B, foi a conformação que melhor se

aproximou da estrutura experimental do ligante cristalográfico. Porém, ainda assim, esta solução apresentou troca na posição de dois de seus grupamentos químicos, comparados à estrutura de referência.

As soluções da configuração C2, selecionadas através do filtro da ligação hidrogênio, estão ilustradas na Figura 25. É possível observar que todas as soluções desta configuração também preencheram os sub-sítios do receptor.

As imagens B, C e D indicam os valores da distância entre o hidrogênio da hidroxila da estrutura da solução com os resíduos Asp34 e Asp214. Todos os valores encontrados foram menores que 2,5 Å, caracterizando a formação de ligação hidrogênio entre as moléculas.

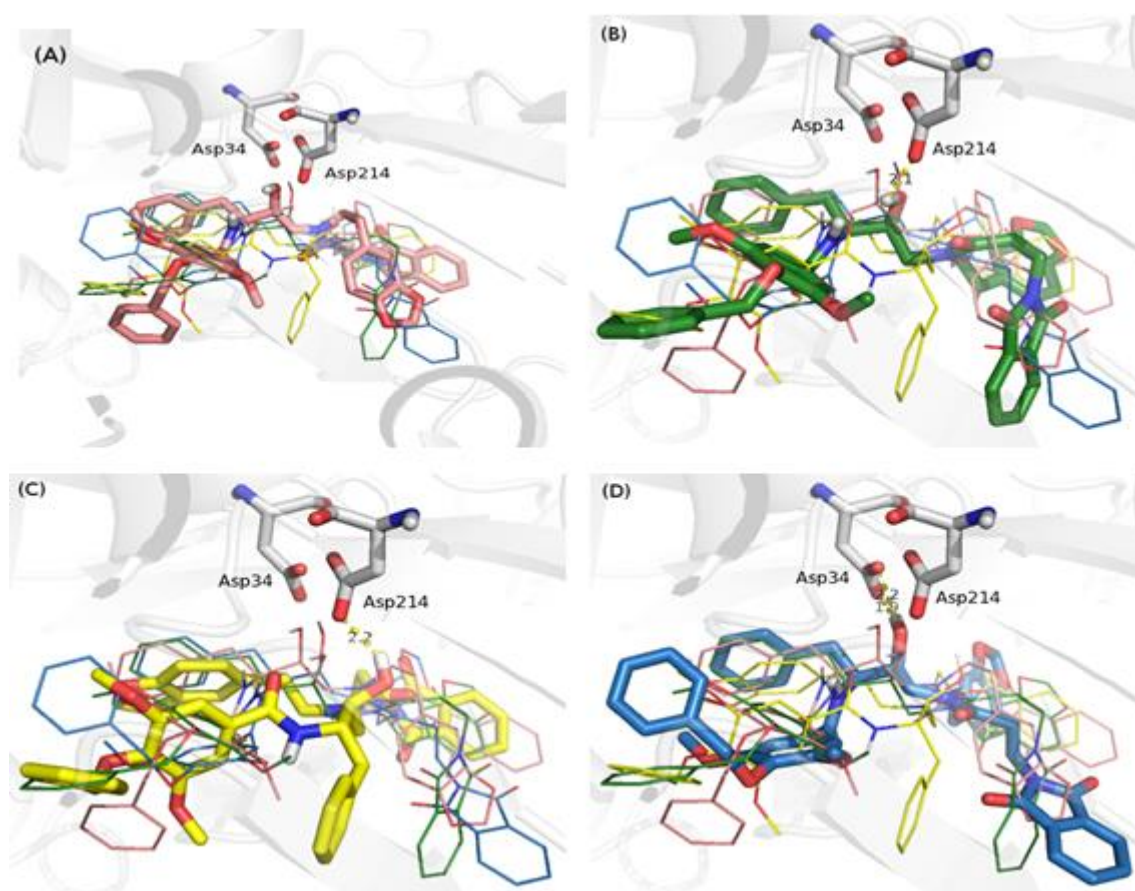


Figura 25. (A) Sobreposição das conformações que interagem com os resíduos Asp34 e Asp214, representadas por linhas de cores aleatórias, e o ligante EH58, representado de rosa. A imagem (B) mostra, em verde, a estrutura que possui menor RMSD com o ligante original e os valores de distância com os resíduos aspárticos do sítio catalítico. As imagens C e D também mostram, (em amarelo e azul, respectivamente), as soluções que interagem com os resíduos aspárticos catalíticos e as distâncias das possíveis ligações hidrogênio. A estrutura representada em azul é a estrutura de menor energia deste experimento. Todos os valores são mostrados em Å.

Os experimentos utilizando os diferentes modelos de cargas dos campos de força foram repetidos utilizando o receptor com o resíduo Asp214 protonado para verificar como o estado de protonação representado com os diferentes conjuntos de cargas parciais atômicas do complexo influenciavam no resultado. A Tabela 11 mostra os resultados deste experimento.

Tabela 11. Redocking utilizando diferentes configurações de campos de força com o programa Autodock 4 com protonação no resíduo Asp214/O1.

Caso Teste ^(a)	Menor Energia ^(b)	Energia Média ^(b)	RMSD da Conf. Menor energia ^(c)	Menor RMSD ^(c)	RMSD Médio ^(c)	T.S. RMSD <=3,0 ^(d)	Interação Ligante – Asp ^(e)
C1	-3,33	-1,68 ± 0,88	11,72	6,79	11,03 ± 1,78	0%	0%
C2	-8,58	-6,00 ± 1,59	3,63	3,63	9,49 ± 2,35	0%	10%
C3	-4,69	-2,48 ± 0,79	12,08	7,08	10,65 ± 1,70	0%	0%
C4	-0,04	1,11 ± 0,63	13,45	8,49	11,28 ± 1,85	0%	0%

^(a) C1 – Carga Ptn: Gromos96; Carga Lig: MMFF94 / C2 – Carga Ptn: Gasteiger; Carga Lig: Gasteiger / C3 – Carga Ptn: Gasteiger; Carga Lig: MMFF94 / C4 – Carga Ptn: MMFF94; Carga Lig: MMFF94. ^(b)Energia dada em kcal/mol; ^(c)RMSD dado em Ångstroms; ^{(d),(e)} Taxa de Sucesso e % de conformações que interagem com Asp34/Asp214 calculadas com as 20 soluções obtidas.

Através dos valores obtidos nesses experimentos, é possível concluir que o programa Autodock 4 somente conseguiu encontrar soluções que interagem com os resíduos aspárticos do sítio ativo quando o modelo de cargas Gasteiger foi usado tanto para proteína como para o ligante. Além disso, com este conjunto de cargas o programa conseguiu encontrar a menor energia de interação dentre todos os experimentos da Tabela 11. As soluções de menor energia obtidas nesses experimentos estão mostradas na Figura 26.

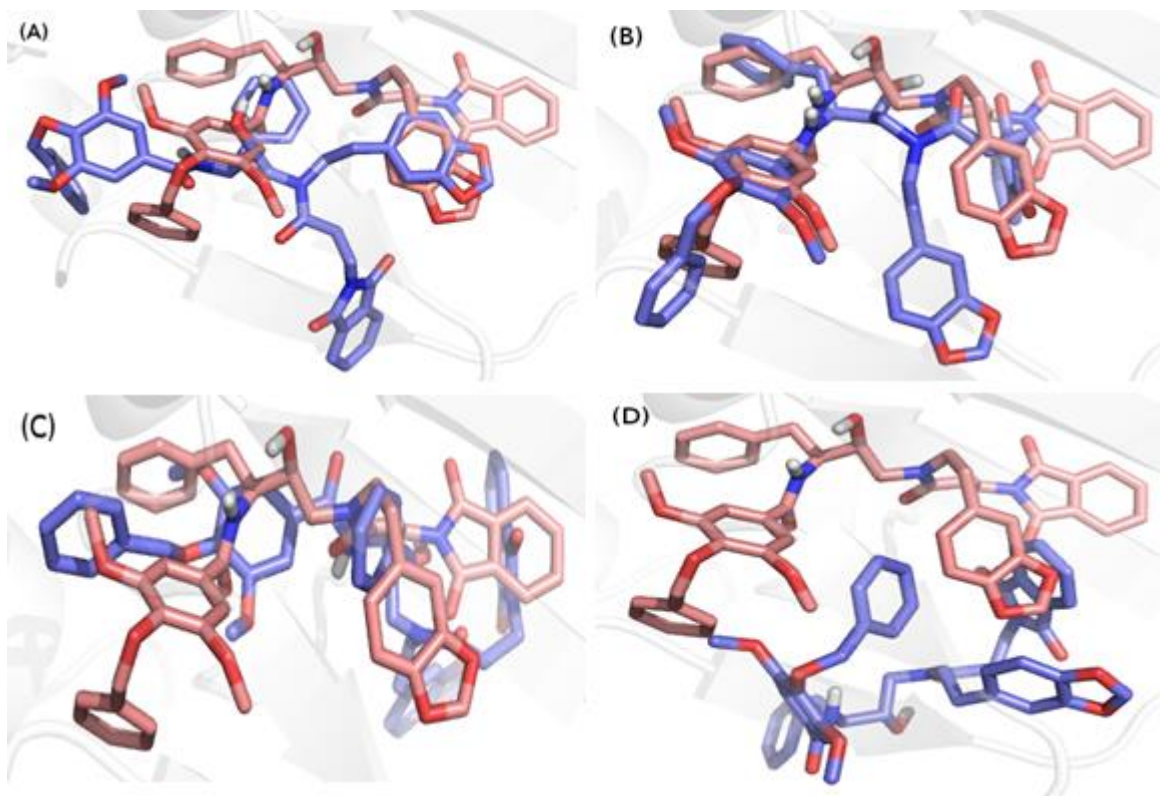


Figura 26. Solução de menor energia da configuração (A) C1, (B) C2, (C) C3 e (D) C4, representada em azul, com a estrutura experimental do ligante EH58, representado em rosa. O valor de RMSD entre o ligante EH58 e as estruturas A, B, C e D é 11,72, 3,63, 12,08 e 13,45 Å, respectivamente.

Ao compararmos a solução de menor energia para o conjunto C2, que possui o resíduo Asp214 protonado, com o correspondente sem a protonação neste aminoácido, pode-se verificar que o experimento realizado com a protonação no resíduo Asp214, consegue um preenchimento correto de todos os grupamentos do ligante EH58. Esta comparação pode ser vista na Figura 27.

Apesar da solução do primeiro experimento (sem protonação), conseguir uma conformação de menor energia menor do que a solução com o resíduo Asp214 protonado, e esta parametrização obter mais conformações que interagem com os resíduos aspárticos do sítio catalítico, o algoritmo não conseguiu encontrar uma solução de menor energia com um valor baixo de RMSD (< 3 Å), o que foi obtido com o segundo experimento.

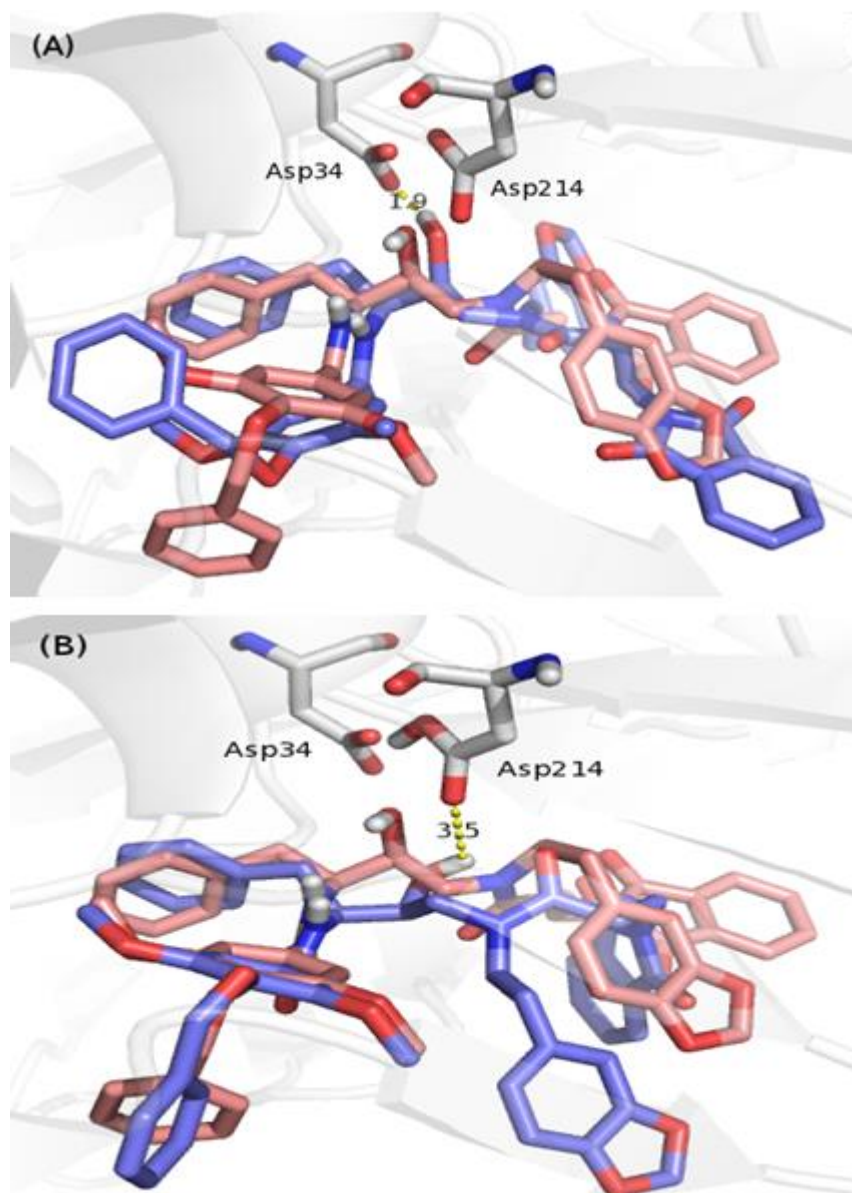


Figura 27. Solução de menor energia da configuração C2 com (A) resíduo Asp214 não protonado e (B) com protonação neste aminoácido. As soluções estão representadas em azul e a estrutura experimental do ligante EH58, representado em rosa. O valor de RMSD entre o ligante EH58 e as estruturas A e B é 5,18 e 3,63 Å, respectivamente.

Com os resultados obtidos nos experimentos de *redocking* utilizando o programa Autodock 4, foi possível concluir que o modelo de cargas Gasteiger para ambas moléculas é o modelo que melhor descreve o sistema deste estudo. Os parâmetros: tamanho da população, elitismo, número de rodadas e número de avaliações de energia foram usados com os valores 150, 1%, 20 e $1,5 \times 10^6$, respectivamente.

Um novo experimento com o programa Autodock 4 foi realizado utilizando o modelo de cargas que obteve o melhor resultado juntamente com os parâmetros pré-definidos (seção 4.1.1). Foram realizadas duas simulações utilizando cargas Gasteiger tanto na proteína quanto no ligante juntamente com o conjunto de parâmetros já definidos alterando somente o número de execuções do programa e a quantidade de avaliações de energia para aumentar a chance do algoritmo encontrar soluções próximas ao modelo experimental.

O número de execuções do programa foi alterado para 150 e o número de avaliações de energia para $2,5 \times 10^6$. Os resultados destes novos experimentos são apresentados na Tabela 12.

Tabela 12. Redocking utilizando o programa Autodock 4 com a nova configuração de parâmetros.

ASP 214 O1 Protonado?	Menor Energia ^(a)	Energia Média ^(a)	RMSD da Conf. Menor energia ^(b)	Menor RMSD ^(b)	RMSD Médio ^(b)	T.S. RMSD <=3,0 ^(c)	Interação Ligante – Asp ^(d)
Não	-12,6	-7,55 ± 1,71	11,83	3,77	9,73 ± 1,72	0%	5%
Sim	-11,3	-7,72 ± 1,50	10,10	2,01	8,34 ± 1,97	0,667%	19%

^(a) Energia dada em kcal/mol; ^(b)RMSD dado em Ångstroms; ^{(c),(d)} Taxa de Sucesso e % de conformações que interagem com Asp34/Asp214 calculadas com as 150 soluções obtidas.

Os resultados mostram que as configurações utilizadas permitiram encontrar valores de energias e RMSD muito próximos entre si, diferindo principalmente na quantidade de conformações que interagem com os resíduos aspárticos catalíticos. O experimento que utilizou a protonação no resíduo Asp214 encontrou soluções com valores menores de RMSD em relação à estrutura experimental do ligante EH58. As soluções de menor RMSD de cada configuração testada foram sobrepostas com o ligante EH58 e são mostradas na Figura 28.

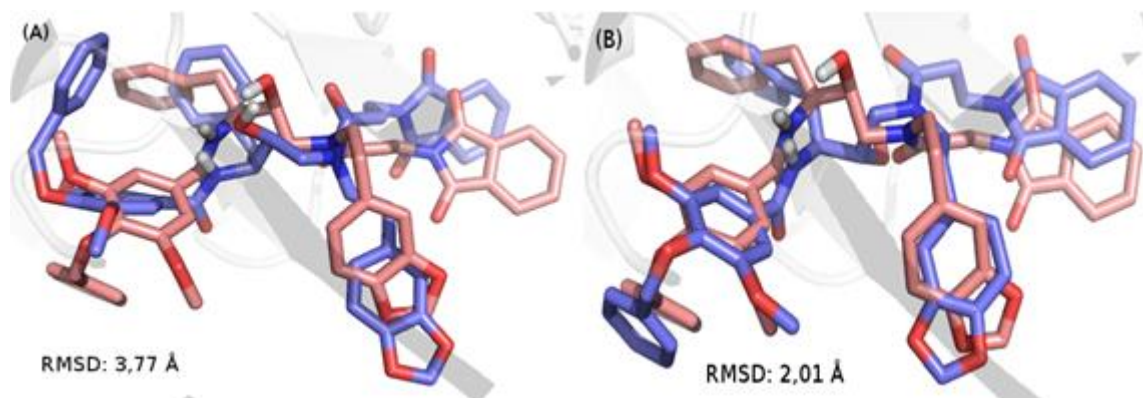


Figura 28. Solução de menor RMSD da configuração que (A) apresenta o resíduo Asp214 não protonado e (B) com protonação neste resíduo. As soluções estão representadas em azul e a estrutura experimental do ligante EH58, representado em rosa. O valor de RMSD entre as estruturas é mostrado na figura.

De acordo com os experimentos de *redocking* realizados com o programa Autodock 4, pode-se concluir que somente utilizando cargas Gasteiger para proteína e ligante, obtêm-se soluções com valores menores de energias em comparação aos outros experimentos usando os mesmos parâmetros com diferentes modelos de carga. Ainda assim, com relação ao primeiro experimento, não foi possível obter nenhuma configuração com RMSD menor que 3 Å entre a solução obtida pelo algoritmo e a conformação do ligante obtido do banco de dados PDB. O resultado mais favorável, ou seja, aquele que obteve conformações que interagem com os resíduos aspárticos, foi obtido utilizando o valor de 150 para tamanho da população, 1% para elitismo e $2,5 \times 10^6$ para avaliações de energia e apesar disso, o programa não conseguiu obter uma taxa de sucesso maior que 19%.

Também foram realizados novos experimentos com o programa Autodock Vina. Nos experimentos de calibração dos parâmetros usados neste programa, o receptor com protonação no resíduo Asp214 foi utilizado. Como este programa ignora as cargas parciais atômicas inicialmente definidas para proteína e ligante, os melhores valores de parâmetros, obtidos com este programa em experimentos anteriores, foram utilizados novamente em um novo experimento para comparar se as soluções obtidas possuem relação à protonação do resíduo Asp214. A Tabela 13 mostra o resultado obtido.

Tabela 13. Redocking utilizando o programa Autodock Vina com o resíduo Asp214/O1 do receptor protonado e sem protonação com a melhor configuração de parâmetros obtido.

ASP 214 O1 Protonado?	Menor Energia ^(a)	Energia Média ^(a)	RMSD da Conf. Menor energia ^(b)	Menor RMSD ^(b)	RMSD Médio ^(b)	T.S. RMSD $\leq 3,0$ ^(c)	Interação Ligante – Asp ^(d)
Não	-9,80	-9,27 \pm 0,27	10,91	7,66	10,83 \pm 1,19	0%	0%
Sim	-10,2	-9,68 \pm 0,15	5,18	2,03	6,86 \pm 3,73	22,75%	63%

^(a) Energia dada em kcal/mol; ^(b)RMSD dado em Ångstroms; ^{(c),(d)} Taxa de Sucesso (T.S.) e % de conformações que interagem com Asp34/Asp214 calculadas com as 400 soluções obtidas.

Pela análise da tabela acima, pode-se concluir que este programa encontrou soluções com valor de menor energia muito próximos da energia média do experimento. Porém, somente no segundo experimento, com o resíduo Asp214 protonado, obteve-se mais de 22% das soluções com RMSD menor que 3 Å da estrutura original, e mais de 60% de soluções em que o ligante interage com os ácidos aspárticos do sítio ativo deste receptor.

A Figura 29 mostra as conformações de menor energia e menor RMSD encontradas em cada experimento sobrepostas com o ligante EH58.

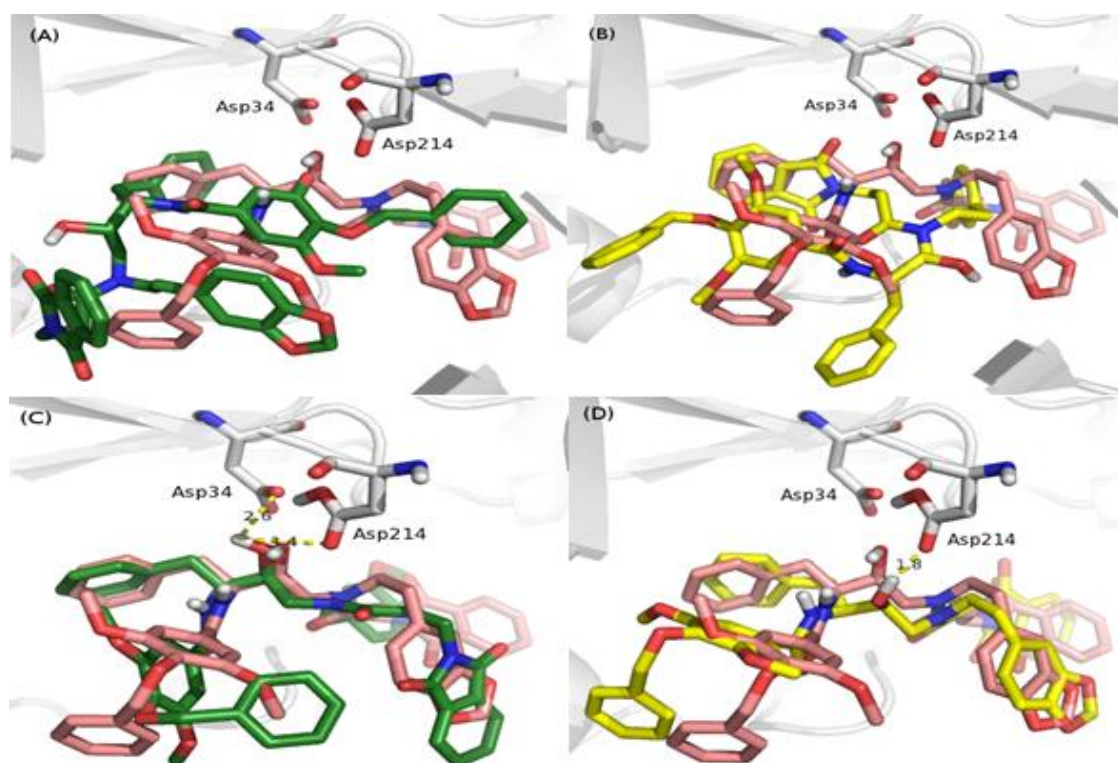


Figura 29. Solução de menor energia (verde) e solução de menor RMSD (amarelo) da configuração que apresenta o resíduo Asp214 não protonado (A e B, respectivamente) e com protonação neste resíduo (C e D, respectivamente) sobrepostos a estrutura tridimensional do ligante EH58 obtido do PDB (rosa).

Através da figura acima, podemos verificar que tanto a solução (A) de menor energia, quanto a (B) de menor RMSD, do experimento sem protonação no resíduo Asp214 ficaram próximas à estrutura experimental do ligante EH58. Já no experimento com o resíduo Asp214 protonado, a solução de menor energia (C) apresentou distância menor que 3,5 Å entre a hidroxila central da molécula com ambos resíduos aspárticos, e a solução de menor RMSD (D), distância de 1,8 Å com Asp214.

Sendo assim, podemos concluir que o programa Autodock Vina consegue representar o sistema deste estudo quando usa a protonação no resíduo Asp214, utilizando os parâmetros já testados nos experimentos iniciais, apresentando uma taxa de sucesso maior que 60% nas 20 execuções do algoritmo.

4.1.3 Utilização do programa Dockthor

Esta parte dos experimentos de *redocking* consiste na análise do programa Dockthor com a realização de experimentos com o resíduo Asp214 protonado e sem protonação. O modelo de carga do campo de força usado para proteína e para o ligante foi o MMFF94. Os parâmetros e valores associados que foram utilizados neste experimento podem ser observados na Tabela 5. Os resultados obtidos são mostrados na Tabela 14.

Tabela 14. Redocking utilizando o programa Dockthor com o resíduo Asp214/O1 do receptor protonado e sem protonação.

Asp214 O1 Protonado?	Menor Energia Total (a)	Energia de Interação* (a)	RMSD* (b)	Energia Total Média (a)	Energia de Interação Média (a)	RMSD Médio (b)	Prováveis HB***	T.S. RMSD <=3,0(c)	Interação Ligante – Asp(d)
Não	86,12	-49,88	1,51	107,82 ± 15,47	-38,08 ± 9,36	4,80 ± 2,92	Asp34; Gly36; Asn76; Ser79; Leu131; Tyr192; Asp214; Gly216; Ser218.	36,67%	80,67%
Sim	84,17	-51,79	1,19	104,34 ± 16,59	-40,47 ± 9,04	4,92 ± 2,72	Ile14; Asp34; Gly36; Asn76; Tyr77; Val78; Ser79; Leu131; Tyr192; Asp214; Gly216; Ser218; Phe294.	28,67%	80,67%

(a) Energia dada em kcal/mol; (b) RMSD dado em Ångstroms; (c),(d) Taxa de Sucesso e Filtro das conformações que interagem com Asp34/Asp214 calculadas a partir de 150 soluções.* Energia de Interação Receptor-Ligante calculado pelo somatório dos valores de energia de van der Waals e Coulomb. O valor mostrado é o da solução de menor energia total. ** RMSD da conformação de menor energia total. *** Aminoácidos que provavelmente interagem por ligação hidrogênio (HB) com a solução.

Os resultados obtidos com o programa Dockthor não mostraram diferença significativa na protonação do resíduo Asp214/O1 com relação às médias de energia total, energia de interação e RMSD. Ambos resultados conseguiram encontrar soluções próximas à estrutura cristalográfica do ligante EH58 e uma quantidade igual de conformações que interagem com os resíduos aspárticos (80%). Apesar da taxa de sucesso do experimento sem protonação no Asp214 ser maior, a conformação de menor energia total possui RMSD de 1,51 Å enquanto que a conformação de menor energia total do experimento com Asp214 protonado encontrou uma solução com 1,19 Å. A Figura 30 mostra a estrutura de menor energia total sobreposta com o ligante EH58.

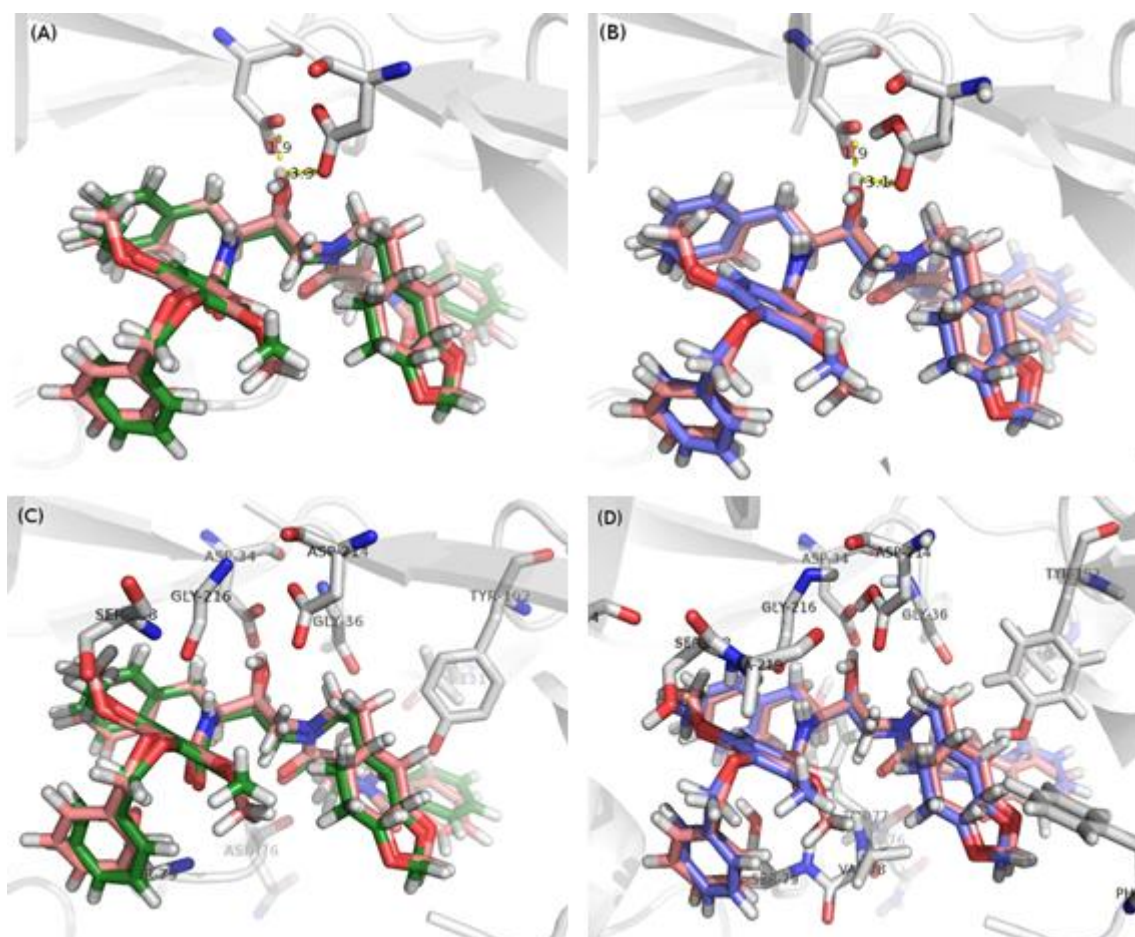


Figura 30. A estrutura representada em verde é a estrutura obtida do experimento sem protonação no resíduo Asp214 e a estrutura em azul é a estrutura obtida do experimento com Asp214 carregado. As imagens A e B indicam possíveis ligações hidrogênio entre a hidroxila central do ligante EH58 com um dos resíduos aspárticos. Já as imagens C e D, mostram as mesmas estruturas porém também é apresentado os resíduos que interagem com o ligante (distância entre os átomos de até 3,5 Å). A estrutura cristalográfica do ligante EH58 é representada em rosa em todas as imagens.

Além da análise da estrutura de menor energia total, foram realizadas outras 2 análises levando-se em consideração a estrutura de (1) menor energia de interação e (2) de menor RMSD. Os resultados dessas análises são apresentados nas Tabelas 15 e 16, respectivamente.

Tabela 15. Análise da estrutura de menor energia de interação utilizando o programa DockThor com o resíduo Asp214/O1 do receptor protonado e sem protonação.

ASP 214 O1 Protonado?	Menor Energia de Interação ^(a)	Energia Total ^(a)	RMSD ^{*(b)}	Prováveis HB**
Não	-50,90	90,10	1,90	Asp34; Gly36; Asn76; Val78; Ser79; Leu131; Tyr192; Asp214; Gly216; Thr217; Ser218; Ala219.
Sim	-52,46	86,42	1,58	Ile14; Asp34; Gly36; Asn76; Tyr77; Val78; Ser79; Leu131; Tyr192; Asp214; Gly216; Ser218; Ala219; Phe294.

^(a) Energia dada em kcal/mol; ^(b)RMSD dado em Ångstroms; * RMSD da conformação de menor energia. ** Aminoácidos que provavelmente interagem por ligação hidrogênio (HB) com a solução.

Assim como na análise da estrutura de menor energia total, as soluções de menor energia de interação também apresentaram resultados similares entre si. A Figura 31 mostra a disposição tridimensional da molécula no sítio de ligação da *Plasmepsina II*, assim como os principais resíduos interagentes com a estrutura da solução desta análise.

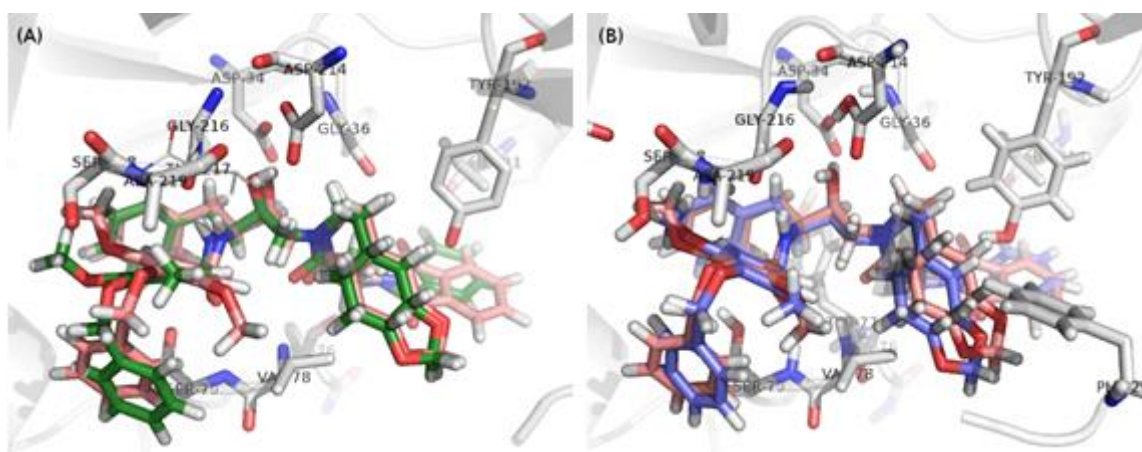


Figura 31. Solução de menor energia de interação do experimento (A) sem o resíduo Asp214 protonado e (B) com protonação, sobrepostos com o ligante EH58 (em rosa).

A terceira análise foi realizada com a estrutura de menor RMSD obtida pelos experimentos com este programa. Na Tabela 16 são apresentados os valores relacionados à esta solução, e na Figura 32, a visualização das mesmas.

Tabela 16. Análise da estrutura de menor RMSD utilizando o programa DockThor com o resíduo Asp214/O1 do receptor protonado e sem protonação.

ASP 214 O1 Protonado?	Menor RMSD ^(a)	Energia Total ^(b)	Energia de Interação ^(b)	Prováveis HB*
Não	0,75	90,97	-48,12	Asp34; Gly36; Asn76; Ser79; Leu131; Tyr192; Asp214; Gly216; Ser218.
Sim	0,84	85,97	-52,17	Asp34; Gly36; Asn76; Tyr77; Val78; Ser79; Leu131; Tyr192; Asp214; Gly216; Ser218; Phe294.

^(a) RMSD dado em Ångstroms; ^(b) Energia dada em kcal/mol. * Aminoácidos que provavelmente interagem por ligação hidrogênio (HB) com a solução.

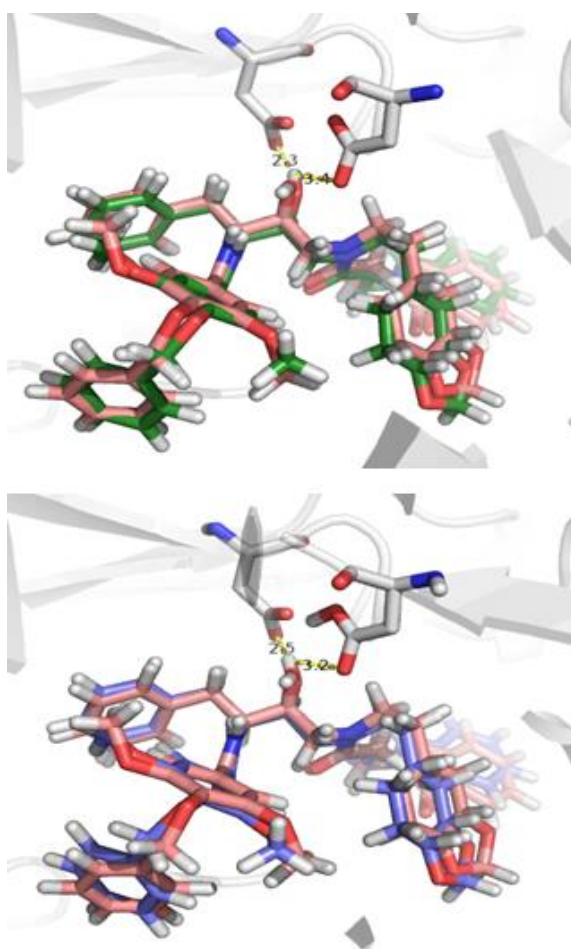


Figura 32. Solução de menor RMSD do experimento sem o resíduo Asp214 protonado (estrutura em verde) e com protonação (estrutura em azul), sobrepostos com o ligante EH58 (em rosa). Ambas soluções interagem com os resíduos aspárticos do sítio catalítico (indicado em pontilhado amarelo com seus respectivos valores de distância).

De acordo com os resultados obtidos nos experimentos de redocking com os programas Autodock 4, Autodock Vina e Dockthor e tendo em vista a taxa de sucesso obtida por cada programa, conforme demonstrado na Figura 33, os experimentos de *redocking* realizados com o programa Dockthor foram os que obtiveram a maior taxa de sucesso em comparação aos outros programas, sugerindo que os parâmetros utilizados neste programa para os cálculos do sistema estudado estão melhor ajustados. Todas as soluções obtidas estão estabilizadas por diversas ligações hidrogênio entre os principais resíduos da enzima *Plasmepsina II* e o ligante EH58. Sendo assim, este programa foi o escolhido para realização dos experimentos de *docking* da nova série de ligantes planejados pelo DSO.

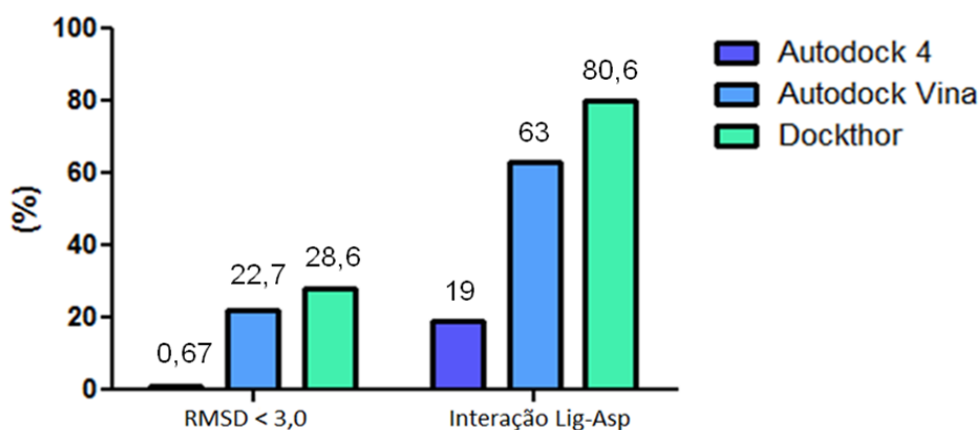


Figura 33. Taxa de sucesso obtida por cada programa de docking.

4.2 Docking de Novos Inibidores de Aspartil Proteases

Para os experimentos de *docking* com o programa Dockthor, a estrutura de menor energia total foi a escolhida para análise. Os parâmetros definidos para este experimento são os mesmos usados nos experimentos de *redocking* com este programa, mostrados na seção anterior. Apesar dos resultados com relação a protonação do resíduo Asp214 não apresentarem diferenças significativas com este programa, os novos experimentos com os ligantes planejados pelo DSO utilizaram o resíduo Asp214 protonado, tendo como base resultados obtidos em trabalhos anteriores (93). A Tabela 17 mostra os resultados deste experimento.

Tabela 17. Resultado do docking das moléculas de Farmanguinhos utilizando o programa DockThor.

Molécula	Menor Energia Total ^(a)	Energia de Interação* ^(a)	Energia Total Média ^(a)	Energia de Interação Média ^(a)	Prováveis Ligações de Hidrogênio
Plip 01	19,04	-25,82	23,92 ± 1,61	-22,75 ± 2,28	Ser79, Asp214, Gly216, Thr217, Ser218
Plip 03	30,28	-24,80	30,99 ± 2,52	-24,03 ± 3,37	Asp34, Gly36, Asn76, Tyr77, Ser79, Tyr192, Asp214, Gly216, Thr217
Plip 04	31,21	-37,50	30,44 ± 4,95	-26,69 ± 4,99	Asp34, Gly36, Asn76, Tyr77, Ser79, Asp214, Gly216, Thr217, Ser218, Phe294
Plip 11	31,92	-25,41	36,06 ± 2,92	-21,94 ± 3,84	Asp34, Gly36, Asn76, Tyr77, Val78, Ser79, Asp214, Gly216, Thr217
Plip 13	48,94	-28,65	49,23 ± 5,39	-23,54 ± 5,38	Asp34, Gly36, Ser37, Asn76, Tyr77, Val78, Ser79, Tyr192, Asp214, Gly216, Phe294
Plip 27	31,63	-36,16	41,94 ± 2,57	-23,28 ± 3,71	Gly36, Asn76, Val78, Ser79, Tyr192, Asp214, Gly216, Thr217
Plip 31	52,54	-34,49	56,89 ± 5,39	-26,40 ± 5,19	Asp34, Gly36, Asn76, Val78, Ser79, Asp214, Gly216, Ser218, Phe294
Plip 41	28,20	-33,14	29,35 ± 2,63	-24,34 ± 3,97	Asp34, Gly36, Asn76, Val78, Ser79, Asp214, Gly216, Thr217
Plip 47	15,66	-37,58	27,93 ± 4,81	-26,75 ± 4,58	Ile14, Asp34, Gly36, Val78, Ser79, Tyr192, Asp214, Gly216, Thr217, Ser218

^(a) Energia dada em kcal/mol; * O valor da energia de interação mostrado é o valor obtido da solução de menor energia total.

Para avaliar se a solução obtida é uma boa solução, foram realizadas duas análises. A primeira análise foi realizada com o objetivo de verificar se as conformações conseguiram um bom cobrimento dos grupamentos do ligante EH58 cristalográfico (Figuras 34 e 35). A segunda análise foi realizada com o objetivo de verificar as possíveis interações da solução encontrada com os resíduos importantes presentes no sítio de ligação da enzima *Plasmepsina II*. Esta segunda análise foi realizada com o *script* desenvolvido descrito na seção 3.1.7.2.

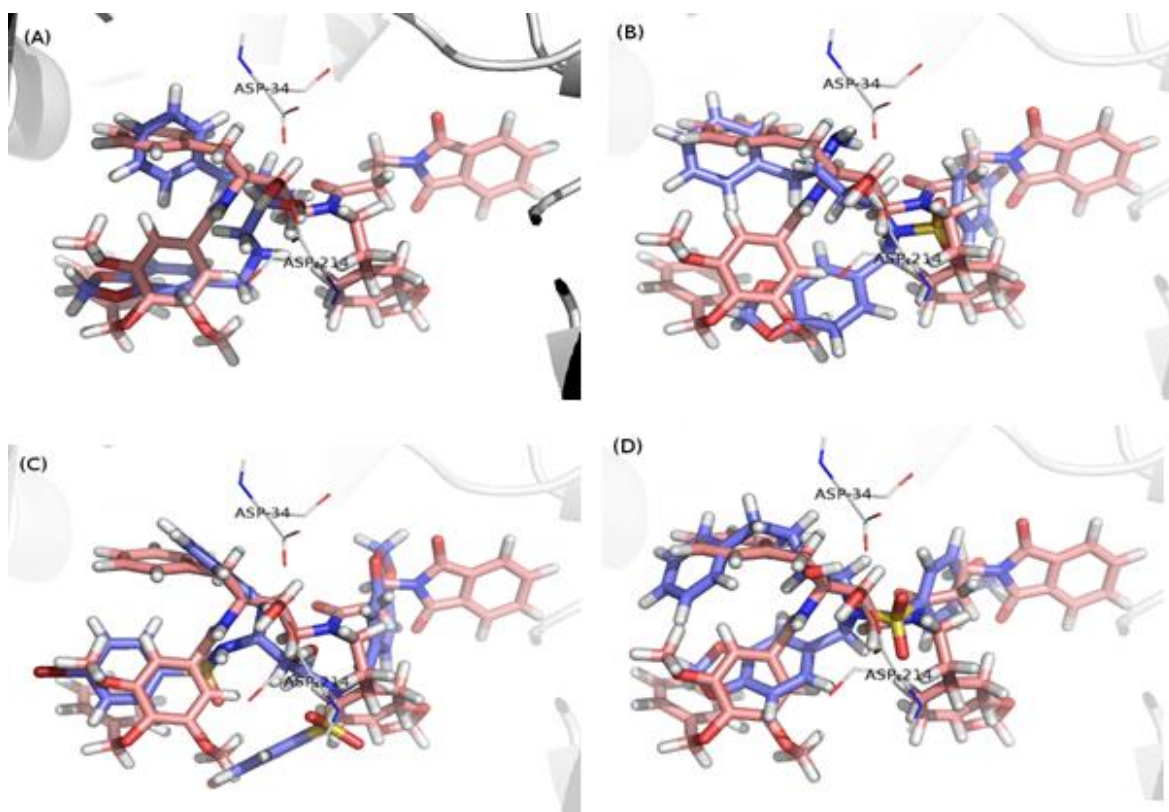


Figura 34. Solução de menor energia total (em azul) sobreposta com o ligante EH58 obtido do PDB (em rosa). As moléculas amostradas são (A) Plip01, (B) Plip03, (C) Plip04, (D) Plip11. Os resíduos aspárticos Asp34 e Asp214 são representados em bastão.

É possível observar que a maioria dos ligantes da série Plip sobrepõem a 2 ou 3 dos 4 grupamentos do ligante EH58, indicando que tais grupamentos podem manter as interações com o sub-sítio respectivo na enzima *Plasmepsina II*.

Porém, ainda assim é possível observar que o grupamento S2' (Figura 36) do ligante EH58 não é preenchido na maioria das soluções encontradas para as moléculas Plip, sugerindo que as moléculas planejadas pelo DSO – Farmanguinhos precisam de uma modificação estrutural para que este novo grupo também preencha este sub-sítio no receptor e consiga interagir com os aminoácidos presentes nessa região.

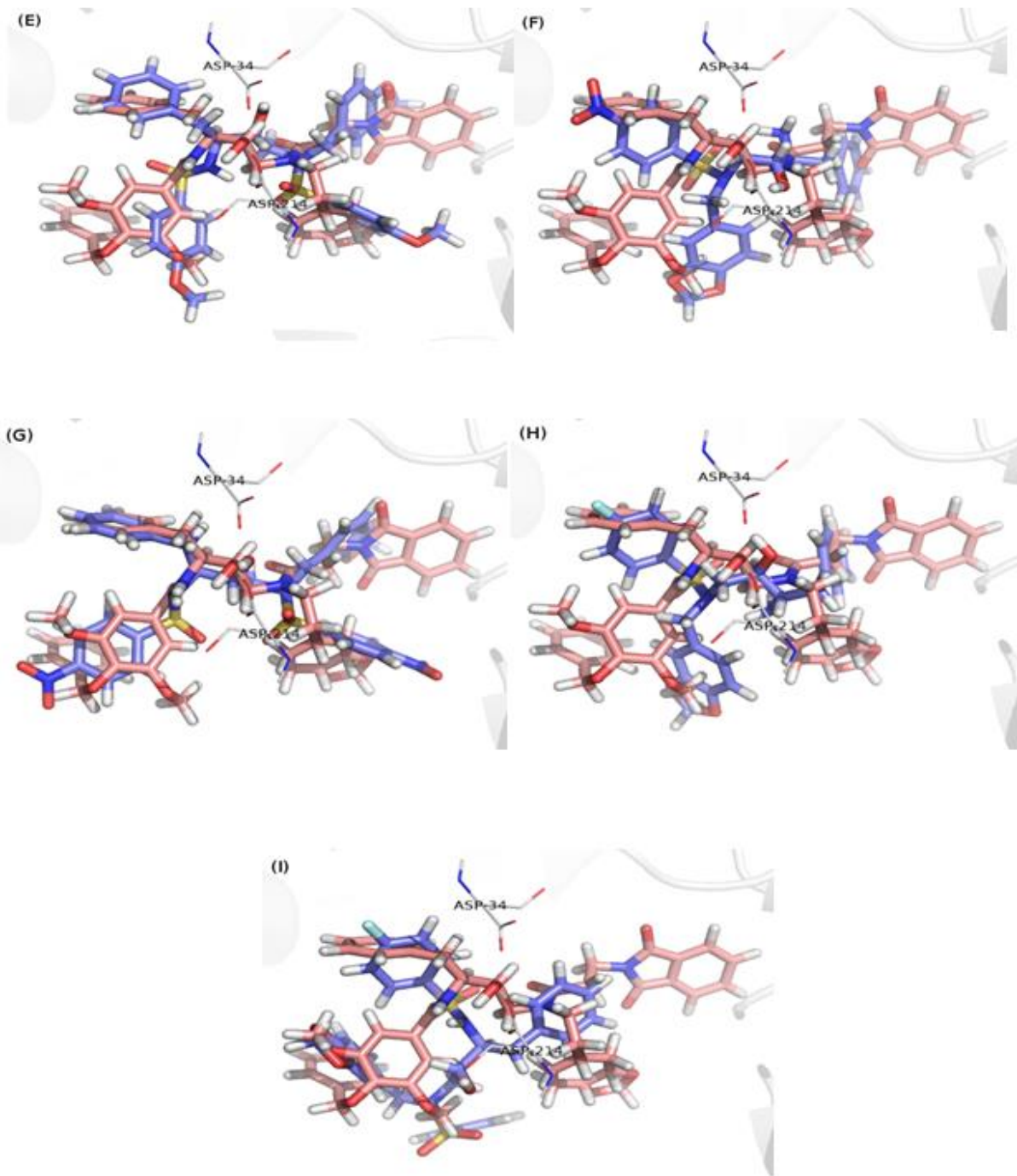


Figura 35. Solução de menor energia total (em azul) sobreposta com o ligante EH58 obtido do PDB (em rosa). As moléculas amostradas são (E) Plip13, (F) Plip27, (G) Plip31, (H) Plip41 e (I) Plip47. Os resíduos aspárticos Asp34 e Asp214 são representados em bastão.

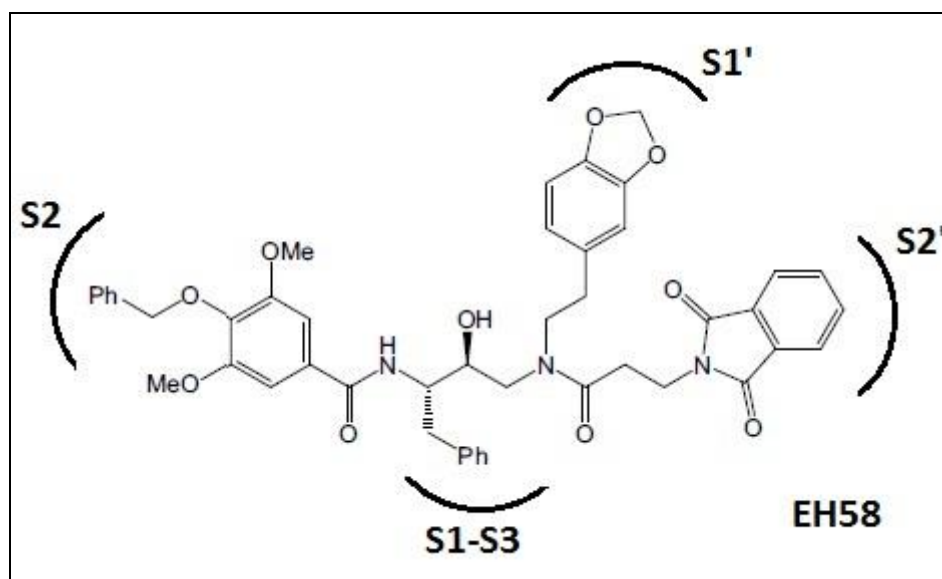


Figura 36. Grupos do ligante EH58 (adaptado) (6).

4.3 Cálculo da Energia Livre

Para o cálculo da energia livre de ligação, foram utilizadas as estruturas químicas dos ligantes cristalográficos complexados à enzima *Plasmepsina II* disponível no banco de dados de proteínas (PDB) sob os seguintes códigos: 1LF3, 1LF2, 1LEE e 2BJU e, para a série Plip, a conformação de menor energia obtida dos resultados do *docking* com o programa DockThor. Os resultados das simulações de dinâmica molecular do complexo cristalográfico *PlmII*-ligante e das novas moléculas (série Plip), planejadas pelo DSO - Farmanguinhos, são mostrados nas seções 4.3.1 e 4.3.2, respectivamente. Os valores do potencial eletrostático, potencial de Lennard-Jones e energia livre de Gibbs, obtidos através da aplicação da metodologia LIE, foram calculados utilizando as informações dos últimos 2,5 ns da DM. Os valores utilizados nos termos α e β foram 0,181 e 0,37, respectivamente. Já o termo γ foi ajustado após regressão linear dos valores experimentais dos ligantes cristalográficos e combinando os valores dos termos α e β . O erro associado ao desvio padrão total de ΔG foi calculado segundo a expressão 10:

$$\sigma_{tot} = \sqrt{\beta^2 (\sigma_{bound}^2(el) + \sigma_{free}^2(el)) + \alpha^2 (\sigma_{bound}^2(vdw) + \sigma_{free}^2(vdw))} \quad (10)$$

onde α e β correspondem aos coeficientes LIE e σ_{tot} representa desvio padrão para termos de interação em estados ligados e livres.

4.3.1 Simulação entre a enzima PlmII e seus respectivos ligantes

O resultado da variação da energia livre, segundo o método LIE, através da DM entre 4 complexos de enzimas da *Plasmepsina II* complexada com o ligante presente em sua estrutura cristalográfica, está mostrado na Tabela 18. Nesta tabela se destacam os valores médios da energia eletrostática e do potencial de Lennard-Jones da simulação do complexo (enzima+ligante+solvente) e da simulação do ligante na solução (ligante+solvente) além da diferença da energia eletrostática e van der Waals. Os respectivos valores de desvio padrão também estão indicados para cada um dos valores calculados.

Tabela 18. Resultado do cálculo de energia livre de 4 complexos de PlmII do PDB.

PDB id	$\langle U \rangle^{el} cplx^{(a)*}$	$\langle U \rangle^{el} sol^{(b)*}$	$\langle U \rangle^{vdw} cplx^{(c)*}$	$\langle U \rangle^{vdw} sol^{(d)*}$	$\Delta U^{el(e)*}$	$\Delta U^{vdw(f)*}$	ΔG_{exp}^*
1LF3	-92,95 ± 2,68	-105,57 ± 1,60	-82,51 ± 1,89	-57,26 ± 0,50	12,61 ± 3,12	-25,25 ± 1,95	-9,92
1LF2	-80,98 ± 8,83	-93,68 ± 3,04	-58,92 ± 2,44	-36,73 ± 0,30	12,70 ± 9,34	-22,19 ± 2,46	-10,67
1LEE	-81,58 ± 7,01	-88,89 ± 1,91	-56,98 ± 2,54	-36,23 ± 0,48	7,31 ± 7,27	-20,74 ± 2,59	-10,98
2BJU	-51,78 ± 1,64	-67,83 ± 0,55	-64,71 ± 1,11	-50,15 ± 0,38	16,06 ± 1,73	-14,57 ± 1,17	-10,74

(a) Média da energia eletrostática do complexo; (b) Média da energia eletrostática da solução; (c) Média da energia de van der Waals do complexo (d) Média da energia de van der Waals da solução; (e) Diferença da energia eletrostática; (f) Diferença da energia de van der Waals; * Todos os valores da tabela estão em Kcal/mol.

Os valores médios da Tabela 18 são os obtidos através da DM. Já os valores de ΔU são calculados a partir da diferença entre valores médios das energias em questão entre o complexo e a solução. A equação 11 mostra a fórmula para o cálculo do ΔU eletrostático. O mesmo cálculo é aplicado para ΔU de van der Waals.

$$\Delta U_{ele} = \langle U \rangle_{cplx}^{ele} - \langle U \rangle_{sol}^{ele} \quad (11)$$

Para o cálculo do desvio padrão foi aplicada a fórmula da equação 12, onde σ_x representa o valor obtido do desvio padrão da média calculada, sendo x complexo (cplx) ou solução (sol).

$$DP = \sqrt{(\sigma_{cplx})^2 + (\sigma_{sol})^2} \quad (12)$$

Para ajustar o valor do parâmetro gamma foi realizada regressão linear levando em consideração os valores de α e β e também os valores experimentais dos 4 complexos P1mII-ligante extraídos do PDB. Com o valor de gamma calculado, este será extrapolado para as simulações com as moléculas da série P1ip dado que o receptor é o mesmo para todos os ligantes.

4.3.2 Resultado da Regressão Linear para complexos P1mII do PDB

Com a regressão linear, quatro configurações são possíveis:

- 1º caso: variáveis α e β fixa em 0,181 e 0,37, respectivamente;
- 2º caso: variável α fixa em 0,181 e variável β livre;
- 3º caso: variável α livre e variável β fixa em 0,37;
- 4º caso: variáveis α e β livres.

Os cálculos foram realizados com o programa R. Como entrada, os valores de energia de van der Waals e energia eletrostática foram fornecidos assim como os valores de ΔG experimental de cada um dos complexos da enzima P1mII retirados do PDB e os valores fixos de alfa e beta.

```
beta <- .370
alfa <- .181
w <- c(-25.25, -22.19, -20.74, -14.57)
e <- c(12.61, 12.70, 7.31, 16.06)
gm <- c(-9.92, -10.67, -10.98, -10.74)
```

Após execução dos comandos descritos na seção 1.7, temos o valor de gamma, descrito na Tabela 19, para cada uma das possibilidades da regressão linear.

Tabela 19. Resultado dos valores de gamma obtidos da regressão linear

	Gamma	Alpha	Beta
1º Caso	-11,336	0,181	0,37
2º Caso	-6,12517	0,181	-0,058
3º Caso	-19,0796	-0,193	0,37
4º Caso	-13,6695	-0,096	0,092

Com os valores obtidos da regressão linear, podemos aplicá-los no cálculo do ΔG (fórmula da equação 8 mostrada na seção 1.6) dos complexos da enzima PlmII obtidas do PDB. Os resultados são mostrados na Tabela 20.

Tabela 20. Resultado do ΔG calculado para os complexos da enzima PlmII do PDB com os valores obtidos da regressão linear.

Molécula	ΔG_{exp} *	ΔG calc 1º Caso*	ΔG calc 2º Caso*	ΔG calc 3º Caso*	ΔG calc 4º Caso*
1LF3	-9,92	-11,24	-11,43	-9,53	-10,10
1LF2	-10,67	-10,65	-10,88	-10,09	-10,39
1LEE	-10,98	-12,39	-10,31	-12,37	-11,02
2BJU	-10,74	-8,03	-9,70	-10,32	-10,81

* Todos os valores são dados em Kcal/mol

Os valores destacados são aqueles que chegaram mais próximo ao valor do ΔG experimental. Com exceção do complexo 1LF2, todas as outras moléculas tiveram o melhor resultado com os valores calculados com a configuração do quarto caso, com alfa e beta livres. Porém, os casos 2, 3 e 4 apresentam valores negativos de alfa ou beta, no qual, preferencialmente, deveriam ser positivos. Ainda, valores que ficam livres na regressão linear não respeita as leis da física porque são previsões estatísticas.

Cabe ressaltar que todos os valores do parâmetro gamma para todos os casos analisados são até duas ordens de grandeza maiores que os correspondentes a alfa e beta, atribuindo um valor dominante a esse parâmetro com relação aos outros. Ou seja, o peso das interações em si aparece apenas como flutuação sobre um valor constante, determinado pelo parâmetro gamma.

Uma consequência importante disto é que o alto valor do parâmetro gamma em relação ao alfa e ao beta dificultou a classificação e ranqueamento dos ligantes.

Diante do exposto, os valores a serem considerados para a classificação das moléculas são os do primeiro caso da regressão linear.

4.3.3 Compostos Farmanguinhos

O resultado da simulação para obtenção da energia livre de ligação, segundo o método LIE, para as moléculas da série Plip complexadas com a enzima *Plasmepsina II*, é mostrado na Tabela 21.

Tabela 21. Resultado do cálculo de energia livre dos ligantes de Farmanguinhos.

Molécula	$\langle U \rangle^{\text{elcplx}}(\text{a})^*$	$\langle U \rangle^{\text{el sol}}(\text{b})^*$	$\langle U \rangle^{\text{vdw cplx}}(\text{c})^*$	$\langle U \rangle^{\text{vdw sol}}(\text{d})^*$	$\Delta U^{\text{el}}(\text{e})^*$	$\Delta U^{\text{vdw}}(\text{f})^*$
Plip 01	-5,02 ± 0,46	-8,15 ± 0,24	-42,13 ± 0,67	-28,30 ± 0,20	3,13 ± 0,51	-13,83 ± 0,70
Plip 03	-27,23 ± 5,50	-34,01 ± 0,53	-51,51 ± 1,84	-34,67 ± 0,29	6,78 ± 5,53	-16,84 ± 1,86
Plip 04	-38,51 ± 1,60	-46,29 ± 0,65	-58,99 ± 1,49	-38,40 ± 0,40	7,78 ± 1,73	-20,59 ± 1,54
Plip 11	-27,62 ± 2,10	-32,62 ± 0,59	-54,21 ± 1,80	-36,89 ± 0,30	5,00 ± 2,18	-17,32 ± 1,83
Plip 13	-36,49 ± 2,16	-44,71 ± 0,77	-73,47 ± 1,21	-50,42 ± 0,43	8,22 ± 2,30	-23,05 ± 1,28
Plip 27	-38,55 ± 2,90	-47,94 ± 0,70	-60,80 ± 1,99	-37,46 ± 0,32	9,39 ± 2,98	-23,34 ± 2,02
Plip 31	-56,45 ± 3,52	-72,81 ± 0,84	-76,74 ± 1,99	-46,17 ± 0,46	16,36 ± 3,62	-30,57 ± 2,04
Plip 41	-30,72 ± 2,55	-35,75 ± 0,66	-55,34 ± 2,77	-36,74 ± 0,27	5,03 ± 2,63	-18,61 ± 2,78
Plip 47	-48,85 ± 1,40	-56,76 ± 0,80	-64,99 ± 2,11	-42,59 ± 0,24	7,91 ± 1,61	-22,40 ± 2,13

(a) Média da energia eletrostática do complexo; (b) Média da energia eletrostática da solução; (c) Média da energia de van der Waals do complexo (d) Média da energia de van der Waals da solução; (e) Diferença da energia eletrostática; (f) Diferença da energia de van der Waals. * Todos os valores da tabela estão em Kcal/mol.

4.3.4 Resultado da Regressão Linear para os compostos de Farmanguinhos

Para obtenção do ΔG calculado através do método LIE para as moléculas da série Plip, usamos o valor de gamma, encontrado no primeiro caso da regressão linear dos complexos de *Plasmepsina II* do PDB, mostrado na Tabela 19. Os resultados da regressão para a série Plip são mostrados na Tabela 22.

Tabela 22. Resultado do ΔG calculado para a série Plip com os valores obtidos da regressão linear.

Molécula	ΔG_{exp}^*	ΔG_{calc} 1º Caso*
Plip 01	-7,28	-12,68
Plip 03	-7,73	-11,88
Plip 04	-6,16	-12,18
Plip 11	-7,90	-12,62
Plip 13	-6,60	-12,47
Plip 27	-7,13	-12,09
Plip 31	-6,73	-10,82
Plip 41	-8,27	-12,84
Plip 47	-6,44	-12,46

* Todos os valores são dados em Kcal/mol

Na Figura 37 são mostrados os valores obtidos experimentalmente e os calculados com o método LIE usando os valores de α , β , e γ do primeiro caso da regressão linear. Os valores de energia livre calculados para as moléculas Plip são mais negativos que os valores obtidos experimentalmente. As contribuições individuais dos termos de Lennard-Jones e eletrostáticos podem ser vistas na Figura 38.

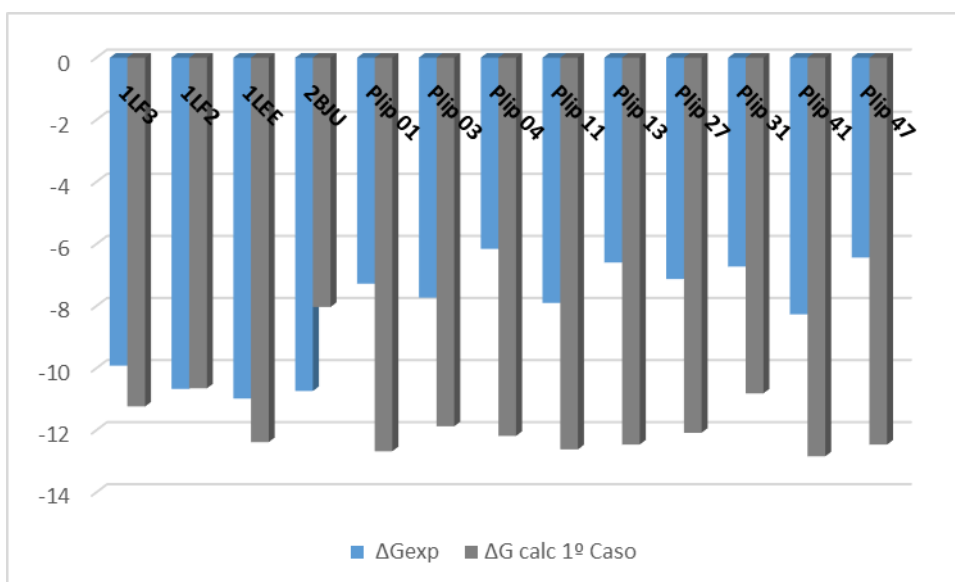


Figura 37. Comparação entre ΔG experimental e ΔG calculado da energia livre de ligação dos novos ligantes planejados pelo DSO e dos ligantes complexados com a enzima Plasmepsina II no primeiro caso da regressão linear. Os valores no eixo Y estão em Kcal/mol.

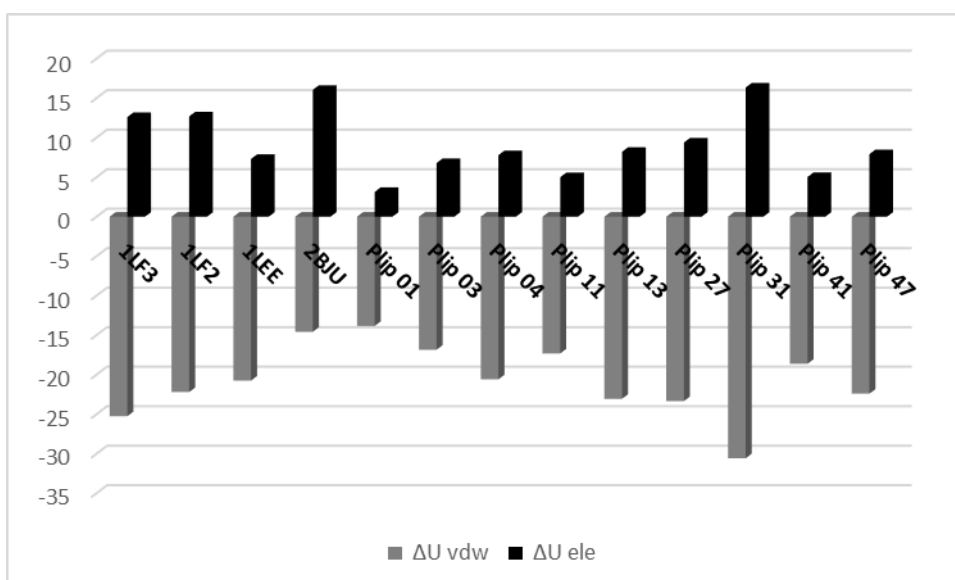


Figura 38. Potenciais eletrostático (ΔU_{ele}) e Lennard-Jones (ΔU_{vdw}) dos ligantes planejados pelo DSO e dos ligantes complexados com a enzima Plasmepsina II. Os valores são dados em Kcal/mol.

É possível observar que a variação de energia potencial eletrostática entre os estados ligado e livre gerou resultados positivos, ou seja, a interação eletrostática das moléculas em solução é mais favorável que dentro do sítio, o que faz sentido se considerarmos a característica majoritariamente hidrofóbica dos ligantes. Em contrapartida, a variação da energia potencial de Lennard-Jones entre esses mesmos estados apresentaram valores menores que 0, indicando que há uma boa interação proteína-ligante em complexo.

Em média, a contribuição eletrostática e a de Lennard-Jones da série Plip, em valor absoluto, é de 7,73 e -20,73 Kcal/mol, respectivamente. Dessa forma, podemos concluir que a maior contribuição para o cálculo da energia livre de ligação procede do termo do potencial de Lennard-Jones.

4.4 Ranqueamento da Série Plip

Para classificar os novos compostos da série Plip, organizamos os valores de energia, obtidos por métodos experimentais, em ordem crescente e comparamos com os valores de energia livre obtidos através do método LIE de acordo com o primeiro caso da regressão linear. A Tabela 23 mostra a classificação dessas moléculas.

Tabela 23. Energia livre de ligação calculada com o primeiro caso da regressão linear e o resultado experimental dos ligantes planejados pelo DSO – Farmanguinhos.

Molécula	ΔG_{exp}^*	ΔG_{calc} 1º Caso*	Atividade (a)
Plip 41	-8,27	-12,84	A
Plip 11	-7,90	-12,62	A
Plip 03	-7,73	-11,88	A
Plip 01	-7,28	-12,68	A
Plip 27	-7,13	-12,09	A
Plip 31	-6,73	-10,82	A
Plip 13	-6,60	-12,47	PA
Plip 47	-6,44	-12,46	PA
Plip 04	-6,16	-12,18	A

*Valores em kcal/mol; (a) A: molécula apresenta atividade / PA: molécula parcialmente ativa;

A Tabela 23 mostra que houve uma correspondência entre os valores de energia livre calculados e experimentais. Especialmente para as 2 primeiras moléculas, Plip41 e Plip11, o resultado indica que o método LIE foi capaz de encontrar uma correspondência com os valores experimentais.

4.4.1 Análise da ligação hidrogênio entre os ligantes Plip e os resíduos aspárticos

Com o objetivo de verificar se os ligantes que foram classificados em ordem de menor energia estão interagindo com os resíduos do sítio de ligação do receptor por ligações hidrogênio, foi calculada a distância entre o átomo de oxigênio da hidroxila do ligante com os átomos de oxigênio da cadeia lateral dos resíduos Asp34 e Asp214. A distância considerada para que haja formação da ligação hidrogênio entre as moléculas foi 3,5 Å. A Tabela 24 mostra os valores em porcentagem de soluções que interagiram com cada um dos átomos de oxigênio dos resíduos Asp34 e Asp214.

Tabela 24. Porcentagem da ligação hidrogênio dos ligantes planejados pelo DSO – Farmanguinhos e os resíduos aspárticos Asp34/O1, Asp34/O2, Asp214/O1 e Asp214/O2 durante a simulação de dinâmica molecular.

	Asp34/O1	Asp34/O2	Asp214/O1	Asp214/O2
Plip 01	0%	0%	0,01%	0%
Plip 03	13,16%	16,78	1,01%	3,02%
Plip 04	0,01%	0%	0,07%	0,16%
Plip 11	0,18%	0,03%	0,01%	0%
Plip 13	10,91%	0,92%	0,16%	0,39%
Plip 27	68,91%	33,41%	37,39%	33,37%
Plip 31	9,55%	4,50%	0,01%	0%
Plip 41	26,04%	4,35%	26,87%	29,59%
Plip 47	0%	0%	0,03%	0%

*Valor calculado em 10000 configurações

A molécula Plip41, aquela que se mostrou mais ativa dentre a série Plip, não foi a molécula que apresentou maior porcentagem de ligação hidrogênio durante os 5 ns da simulação de DM. A molécula Plip11, classificada na segunda posição das moléculas Plip que possuem menor energia, apresentou praticamente 0% de interação com os resíduos aspárticos durante a simulação.

Sendo assim, a partir dos dados da tabela, é possível concluir que não existe uma relação direta entre a formação de ligação hidrogênio do ligante e resíduos aspárticos importantes da PlmII com os valores obtidos do cálculo de ΔG .

5 Conclusões

Neste trabalho foi realizado o estudo da interação de diversas moléculas ligantes no sítio ativo da proteína da família da aspartil protease, *Plasmepsina II*. Uma das vantagens importantes das técnicas de simulação computacional é que elas podem ser usadas para obter informação detalhada no nível atômico-molecular. O grande desafio é assegurar que os resultados obtidos são um reflexo das propriedades do sistema.

Os resultados obtidos nos estudos de *redocking* do ligante EH58 no sítio de ligação da enzima *Plasmepsina II* confirmaram a dificuldade dos programas atuais de *docking* em lidar com moléculas ligantes altamente flexíveis, como já mostrado na literatura (77).

As análises com o programa Autodock 4 mostraram, através de diversas simulações utilizando vários conjuntos de parâmetros, que melhores resultados são obtidos quando são utilizadas cargas Gasteiger para proteína e para o ligante.

Já o programa Autodock Vina se mostrou melhor do que o programa Autodock 4, encontrando uma quantidade maior de soluções próximas à estrutura experimental do ligante EH58, quando o parâmetro *exhaustiveness* é utilizado a partir de 100. A modificação do parâmetro *energy_range* não alterou significativamente os resultados.

O programa DockThor apresentou as melhores taxas de sucesso dentre os três programas de *docking* utilizados, ou seja, encontrou mais conformações próximas à estrutura experimental do ligante EH58 e também uma maior quantidade de conformações que interagem com os resíduos Asp34 e Asp214. Este fato motivou a escolha deste programa para ser utilizado nas simulações de atracamento molecular da série Plip do DSO – Farmanguinhos.

Para as simulações da dinâmica molecular foram utilizadas as estruturas das moléculas da série Plip geradas com o programa DockThor. A simulação computacional se mostrou como uma importante técnica computacional para a avaliação das moléculas bioativas de Farmanguinhos, tanto na sua capacidade de predição da energia de interação entre o receptor e o ligante quanto na determinação dos plausíveis modos de ligação dos protótipos no sítio de ligação do alvo molecular.

Por fim, os resultados obtidos por este trabalho mostram as vantagens de se utilizar a combinação de técnicas de *docking* e dinâmica molecular através da aplicação do método de determinação da energia livre (LIE) para a predição dos modos mais prováveis de ligação de uma série de ligantes no sítio ativo da *Plm II*.

O método LIE mostrou que os ligantes Plip mais ativos também apresentaram energias livre de ligação mais favoráveis e a maior contribuição de energia procedeu do termo do potencial de Lennard-Jones. Esses ligantes mais efetivos não são os que formaram ligações hidrogênio com alta porcentagem de meia vida com o resíduos Asp214 e Asp34 indicando que outras interações intermoleculares estão envolvidas.

Os plausíveis modos de ligação das moléculas da série Plip, obtidos e validados neste trabalho, podem agora ser utilizados como pontos de partida para modificações estruturais nestas moléculas, visando o aumento da afinidade de ligação das mesmas e/ou a obtenção de propriedades desejáveis, tais como absorção, metabolismo, etc. Como exemplo, uma possível modificação estrutural visando aumentar a afinidade de ligação seria a inclusão de um grupamento no ligante para preencher o sub-sítio S2' da enzima *Plm II*, interagindo assim com os resíduos presentes nesta região.

6 Referências

1. Malária 2010 [14/05/2012]. Available from: <http://www.dndi.org.br/pt/doencas-negligenciadas/malaria.html>.
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. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Research*. 2000;28(1):235-42.
4. Murray CJ, Rosenfeld LC, Lim SS, Andrews KG, Foreman KJ, Haring D, et al. Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet*. 2012;379(9814):413-31.
5. Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI. The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature*. 2005;434(7030):214-7.
6. Cunico W, Gomes CR, Facchinetti V, Moreth M, Penido C, Henriques MG, et al. Synthesis, antimalarial evaluation and molecular modeling studies of hydroxyethylpiperazines, potential aspartyl protease inhibitors, part 2. *European journal of medicinal chemistry*. 2009;44(9):3816-20.
7. Prudencio M, Rodriguez A, Mota MM. The silent path to thousands of merozoites: the *Plasmodium* liver stage. *Nature reviews Microbiology*. 2006;4(11):849-56.
8. Amino R, Thiberge S, Martin B, Celli S, Shorte S, Frischknecht F, et al. Quantitative imaging of *Plasmodium* transmission from mosquito to mammal. *Nature medicine*. 2006;12(2):220-4.
9. L. R. Parasitologia. Rio de Janeiro: Guanabara-Koogan; 2001.
10. Saúde SdVe. Manual de diagnóstico laboratorial da malária. In: Saúde Md, editor. 1 ed. Brasília2005. p. 112.
11. Coombs GH, Goldberg, D.E., Klemba M., Berry, C., Kay, J, Mottram, J.C. Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends Parasitol*. 2001;1:532–7.
12. Banerjee R, Liu J, Beatty W, Pelosof L, Klemba M, Goldberg DE. Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. *Proceedings of the National Academy of Sciences*. 2002;99(2):990-5.

13. Goldberg DE, Slater A, Beavis R, Chait B, Cerami A, Henderson G. Hemoglobin degradation in the human malaria pathogen *Plasmodium falciparum*: a catabolic pathway initiated by a specific aspartic protease. *The Journal of experimental medicine*. 1991;173(4):961-9.
14. França TC, Santos MGd, Figueroa-Villar JD. Malária: aspectos históricos e quimioterapia. *Quim Nova*. 2008;31(5):1271-8.
15. Foye WO, Lemke TL, Williams DA. Foye's principles of medicinal chemistry: Lippincott Williams & Wilkins; 2008.
16. Golan D, Junior T, Armen H, Armstrong EJ, Armstrong AW. Princípios de farmacologia: a base fisiopatológica da farmacoterapia: Guanabara Koogan; 2009.
17. Li G, Guo X, Arnold K, Jian H, Fu L. Randomised comparative study of mefloquine, qinghaosu, and pyrimethamine-sulfadoxine in patients with falciparum malaria. *The Lancet*. 1984;324(8416):1360-1.
18. Saúde Md, Epidemiológica ddVeSDdV. Ações de controle da malária: manual para profissionais de saúde na atenção básica. Ministério da Saúde Brasília; 2006.
19. DUARTE CDS. ESTUDOS DE QSAR-2D EM QUINOLINA METANOL COM ATIVIDADE ANTIMALÁRICA FRENTE AO *Plasmodium falciparum*. Manaus - AM: UNIVERSIDADE DO ESTADO DO AMAZONAS – UEA; 2012.
20. Woodrow C, Krishna S. Antimalarial drugs: recent advances in molecular determinants of resistance and their clinical significance. *Cellular and Molecular Life Sciences CMLS*. 2006;63(14):1586-96.
21. Bell A. Antimalarial drug synergism and antagonism: mechanistic and clinical significance. *FEMS microbiology letters*. 2005;253(2):171-84.
22. Fitch CD. Ferriprotoporphyrin IX, phospholipids, and the antimalarial actions of quinoline drugs. *Life sciences*. 2004;74(16):1957-72.
23. Woodrow C, Haynes R, Krishna S. Artemisinins. *Postgraduate medical journal*. 2005;81(952):71-8.
24. Ridley RG. Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature*. 2002;415(6872):686-93.
25. Pearl L, Blundell T. The active site of aspartic proteinases. *FEBS letters*. 1984;174(1):96-101.
26. Omara-Opyene AL, Moura PA, Sulsona CR, Bonilla JA, Yowell CA, Fujioka H, et al. Genetic disruption of the *Plasmodium falciparum* digestive vacuole

plasmepsins demonstrates their functional redundancy. *J Biol Chem.* 2004;279(52):54088-96.

27. Ersmark K, Samuelsson B, Hallberg A. Plasmepsins as potential targets for new antimalarial therapy. *Med Res Rev.* 2006;26(5):626-66.

28. 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*. *Journal of molecular biology.* 2003;327(1):173-81.

29. Cunico W, Gomes CR, Moreth M, Manhanini DP, Figueiredo IH, Penido C, et al. Synthesis and antimalarial activity of hydroxyethylpiperazine derivatives. *European journal of medicinal chemistry.* 2009;44(3):1363-8.

30. Anderson AC. The process of structure-based drug design. *Chemistry & biology.* 2003;10(9):787-97.

31. Magalhães CSd. Algoritmos Genéticos para o Problema de Docking Proteína-Ligante: Laboratório Nacional de Computação Científica; 2006.

32. Greer J. Comparative modeling methods: Application to the family of the mammalian serine proteinases. *Proteins.* 1990;7:317-34.

33. Nelson DL, Cox MM, Lehninger AL, Cuchillo CM. *Princípios de bioquímica: Omega;* 2006.

34. Verli H. *Bioinformática da Biologia à flexibilidade molecular.* 1 ed. Porto Alegre 2014. 282 p.

35. Mati IK, Cockroft SL. Molecular balances for quantifying non-covalent interactions. *Chemical Society Reviews.* 2010;39(11):4195-205.

36. Atkins P, De Paula J. *Elements of physical chemistry: Oxford University Press;* 2012.

37. Fraga CAM. *Razões da Atividade Biológica: Interações entre micro e biomacromoléculas. Redes.* 2011.

38. Gringauz A, Muller G. *Introduction to medicinal chemistry: how drugs act and why.* 1997.

39. ROCHA WR. *Interações intermoleculares. Cadernos Temáticos de Química Nova na Escola.* 2001;4:31-6.

40. White DP. *QUÍMICA A Ciência Central.* Prentice Hall; 2005.

41. Brooks BR, Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. CHARMM: A Program for Macromolecular Energy Minimization, and Dynamics Calculations. *J Comp Chem.* 1983;4(187).
42. Brooks III CL, M. Karplus, and B. M. Pettitt. Proteins, a theoretical perspective of dynamics, structure and thermodynamics. In *Advances in Chemical Physics*. Prigogine aSAR, editor: John Wiley & Sons; 1988.
43. van Gunsteren WE, Berendsen, H. J. C. . *Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry*. Anger, Chem Inr Ed Engl. 1990;29:992-1023.
44. Kastritis PL, Moal IH, Hwang H, Weng Z, Bates PA, Bonvin AM, et al. A structure-based benchmark for protein–protein binding affinity. *Protein Science.* 2011;20(3):482-91.
45. Morgon NH, Coutinho, K. *Métodos de Química Teórica e Modelagem Molecular: Livraria da Física*; 2007.
46. Halperin I, Ma B, Wolfson H, Nussinov R. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics.* 2002;47(4):409-43.
47. Xu Y, Xu D, Liang J. *Computational methods for protein structure prediction and modeling.* 2007.
48. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. *Journal of molecular biology.* 1996;261(3):470-89.
49. 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.
50. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010;31(2):455-61.
51. de Magalhães CS, Almeida DM, Barbosa HJC, Dardenne LE. A dynamic niching genetic algorithm strategy for docking highly flexible ligands. *Information Sciences.* 2014;289:206-24.
52. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem.* 2004;47(7):1739-49.
53. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. *Proteins.* 2003;52(4):609-23.

54. 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.
55. Jain AN. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *J Med Chem.* 2003;46(4):499-511.
56. Wang R, Lai L, Wang S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *Journal of computer-aided molecular design.* 2002;16(1):11-26.
57. GONÇALVES RB. Desenvolvimento e validação de novos métodos de distribuição da população inicial em algoritmos genéticos para o problema de docking proteína-ligante: Laboratório Nacional de Computação Científica; 2008.
58. Almeida DM. Dockthor: Implementação, Aprimoramento e Validação de um Programa de Docking Receptor-Ligante: Laboratório Nacional de Computação Científica; 2011.
59. Van Gunsteren W, Berendsen H. Groningen molecular simulation (GROMOS). Library manual, Biomos, Groningen, The Netherlands. 1987:1-221.
60. Arêas E, Pascutti P, Schreier S, Mundim K, Bisch P. Molecular dynamics at a cytoplasm/membrane interface of a signal sequence from an E. coli maltoporin. *Brazilian journal of medical and biological research= Revista brasileira de pesquisas medicas e biologicas/Sociedade Brasileira de Biofisica[et al].* 1994;27(2):527-33.
61. Areas EP, Pascutti PG, Schreier S, Mundim KC, Bisch PM. Molecular dynamics simulations of signal sequences at a membrane/water interface. *The Journal of Physical Chemistry.* 1995;99(40):14885-92.
62. Caffarena ER, Bisch PM. Hydration of T-antigen Gal β (1-3) GalNAc and the isomer Gal β (1-3) GlcNAc by molecular dynamics simulations. *Journal of Molecular Graphics and Modelling.* 2000;18(2):119-25.
63. van Gunsteren WF, King PM, Mark AE. Fundamentals of drug design from a biophysical viewpoint. *Quarterly reviews of biophysics.* 1994;27(4):435-81.
64. Kollman P. Free-Energy Calculations - Applications to Chemical and Biochemical Phenomena. *Chem Rev.* 1993;93(7):2395-417.
65. Aqvist J, Medina C, Samuelsson JE. A new method for predicting binding affinity in computer-aided drug design. *Protein engineering.* 1994;7(3):385-91.
66. Hansson T, Marelus J, Aqvist J. Ligand binding affinity prediction by linear interaction energy methods. *Journal of computer-aided molecular design.* 1998;12(1):27-35.

67. Gutierrez-de-Teran H, Nervall M, Ersmark K, Liu P, Janka LK, Dunn B, et al. Inhibitor binding to the plasmepsin IV aspartic protease from *Plasmodium falciparum*. *Biochemistry*. 2006;45(35):10529-41.
68. Paula GA. Modelos de regressão: com apoio computacional: IME-USP São Paulo; 2004.
69. Charnet R, Freire CdL, Charnet EM, Bonvino H. Análise de modelos de regressão linear com aplicações. Campinas, São Paulo, Unicamp, 356p. 1999.
70. Statistical Package R. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2009.
71. Trott O, Olson AJ. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J Comput Chem*. 2010;31(2):455–61.
72. SCHRODINGER L. The PyMOL Molecular Graphics System. Version 1.3r1 ed2010.
73. Sayle RA, Milner-White EJ. RASMOL: biomolecular graphics for all. *Trends in biochemical sciences*. 1995;20(9):374.
74. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *Journal of molecular graphics*. 1996;14(1):33-8, 27-8.
75. Marelius Jea. Q: a molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems *Journal of Molecular Graphics and Modelling*. 1998;16(4-6):213-25.
76. Sodero ACRDS, S. G.; Silva-Jr, F. P. Mecanismo catalítico e estado de protonação do sítio ativo de aspartil proteases pepsina-símiles. *Rev Virtual Quim*. 2009;1(2):128-37.
77. Friedman R, Caflich A. On the orientation of the catalytic dyad in aspartic proteases. *Proteins*. 2010;78(6):1575-82.
78. Friedman R, Caflich A. The protonation state of the catalytic aspartates in plasmepsin II. *FEBS letters*. 2007;581(21):4120-4.
79. Ott KH, Meyer B. Parametrization of GROMOS force field for oligosaccharides and assessment of efficiency of molecular dynamics simulations. *Journal of Computational Chemistry*. 1996;17(8):1068-84.
80. Mario Ja. A new model for calculating atomic charges in molecules. *Tetrahedron Letters*. 1978;19(34):3181-4.

81. Halgren TA. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J Comput Chem.* 1996;17(5/6):520–52.
82. Lindahl E HB, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J Mol Model.* 2001;7(8):306-17.
83. Olson. AJ. Molecular Graphics Laboratory Tool. The Scripps Research Institute; 2011. p. Used for visualization and analysis of molecular structures.
84. Inc. CCG. Molecular Operating Environment (MOE). 2006.08 ed2006.
85. Dardenne. LE. Propriedades Eletrostáticas do Sítio Ativo de Cisteína Proteinases da Família da Papaina: Universidade Federal do Rio de Janeiro; 2000.
86. Marelius J, Kolmodin K, Feierberg I, Aqvist J. Q: a molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems. *Journal of molecular graphics & modelling.* 1998;16(4-6):213-25, 61.
87. Malde AK ZL, Breeze M, Stroet M, Poger D, Nair PC, Oostenbrink C, Mark AE. An Automated force field Topology Builder (ATB) and repository: version 1.0. *Journal of Chemical Theory and Computation.* 2011;7(12):4026–37.
88. Nunn DB. William Stewart Halsted--a profile of courage, dedication, and scientific search for truth. *Journal of vascular surgery : official publication, the Society for Vascular Surgery [and] International Society for Cardiovascular Surgery, North American Chapter.* 1989;10(3):221-9.
89. Dewar MJS, Zoebisch, E.G., Healy, E.F., Stewart, J.J.P. AM1: A New General Purpose Quantum Mechanical Molecular Model1. *J Am Chem Soc.* 1985;107.
90. Hehre WJ SR, Pople JA. Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of Slater-type atomic orbitals. *J Chem Phys.* 1969;51:2657–65.
91. Sant'Anna CMR. ESTUDO SEMI-EMPÍRICO DE ÁCIDOS HIDROXÂMICOS: ÁCIDO FORMOÍDROXÂMICO E DERIVADOS DO ALELOQUÍMICO DIMBOA. *Química Nova.* 2001;24(5).
92. O'Reilly. Programming Perl2012.
93. Friedman R, Caffisch A. The protonation state of the catalytic aspartates in plasmepsin II. *FEBS letters.* 2007;581(21):4120-4.

7 Anexo I

7.1 Parâmetros da dinâmica molecular utilizados no programa Q

7.1.1 Equilibração do Complexo

```
[MD]
steps          90000
stepsize       1.0
temperature    50.0      # temperatura variou de 50.0 a 310.0
bath_coupling  1.0      # após 30000 passos, o valor foi mudado para 10.0
random_seed    3333
initial_temperature 50.0  # temperatura variou de 50.0 a 310.0
lrf            on

[intervals]
non_bond       25
output         50
trajectory     50

[files]
topology        cplx.top
final           lie_eq.re
fep             6a.fep
trajectory      lie_eq.dcd

[sequence_restraints]
1 3335 5.0 0
```

7.1.2 Dinâmica Molecular do Complexo

```
[MD]
steps          5000000
stepsize       1.0
temperature    310.0
bath_coupling  5
lrf            on

[intervals]
non_bond       50
output         50
energy         500
trajectory     500

[files]
topology        cplx.top
restart         lie_eq.re
final          md5ns.re
energy         md5ns.en
fep            6a.fep
trajectory      md5ns.dcd
```

7.1.3 Dinâmica Molecular da Solução

```
[MD]
steps          5000000
stepsize       1.00
temperature    310.0
bath_coupling  10
lrf            on
```

```
[intervals]
non_bond       25
output         10
energy         250
trajectory     250
```

```
[files]
topology       6a.top
final          md5ns.re
energy         md5ns.en
trajectory     md01.dcd
fep           6a.fep
```

```
[solvent]
radial_force   60.0
polarisation   on
polarisation_force 20.0
```

```
[sequence_restraints]
1 61 5.0
```