



**UNIVERSIDADE FEDERAL DE PERNAMBUCO
DEPARTAMENTO DE ZOOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA ANIMAL**

RAFAEL TRINDADE MAIA

ANÁLISE *IN SILICO* E POLIMORFISMO GENÉTICO DAS GLUTATIONA S-TRANSFERASES DA CLASSE EPSILON DE *Anopheles gambiae* (DIPTERA: CULICIDAE): POSSÍVEIS IMPLICAÇÕES NA RESISTÊNCIA A INSETICIDAS QUÍMICOS.

**RECIFE
2013**

RAFAEL TRINDADE MAIA

ANÁLISE *IN SILICO* E POLIMORFISMO GENÉTICO DAS GLUTATIONA S-TRANSFERASES DA CLASSE EPSILON DE *Anopheles gambiae* (DIPTERA: CULICIDAE): POSSÍVEIS IMPLICAÇÕES NA RESISTÊNCIA A INSETICIDAS QUÍMICOS.

Tese de Doutorado apresentada ao Programa de Pós-graduação em Biologia Animal da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de Doutor.

Área de concentração: Zoologia aplicada.

Orientadora: Dra. Constância Flávia Junqueira Ayres Lopes

Co-Orientadora: Prof. Dra. Thereza Amélia Soares

**RECIFE
2013**

Catálogo na fonte

Elaine Barroso

CRB 1728

Maia, Rafael Trindade

Análise *in silico* e polimorfismo genético das glutathione S-transferases da classe epsilon de *Anopheles gambiae* (Diptera: Culicidae): possíveis implicações na resistência a inseticidas químicos/ Rafael Trindade Maia– Recife: O Autor, 2013.

153 folhas : il., fig., tab.

Orientadora: Constância Flávia Junqueira Ayres Lopes

Coorientadora: Thereza Amélia Soares

Tese (doutorado) – Universidade Federal de Pernambuco, Centro de Ciências Biológicas. Biologia Animal, 2013.

Inclui bibliografia e apêndices

- 1. *Anopheles gambiae* 2. Dinâmica molecular 3. Resistência aos inseticidas I. Lopes, Constância Flávia Junqueira Ayres (orientadora) II. Soares, Thereza Amélia (coorientadora) III. Título**

595.772

CDD (22.ed.)

UFPE/CCB- 2013- 351

RAFAEL TRINDADE MAIA

ANÁLISE *IN SILICO* E POLIMORFISMO GENÉTICO DAS GLUTATIONA S-TRANSFERASES DA CLASSE EPSILON DE *Anopheles gambiae* (DIPTERA: CULICIDAE): POSSÍVEIS IMPLICAÇÕES NA RESISTÊNCIA A INSETICIDAS QUÍMICOS.

Tese de Doutorado apresentada ao Programa de Pós-graduação em Biologia Animal da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de Doutor.

Área de concentração: Zoologia aplicada.

BANCA EXAMINADORA:

**Dra. Constância Flávia Junqueira Ayres
Lopes
(Centro de Pesquisas Aggeu Magalhães)**

**Dr. Roberto Dias Lins Neto
(Universidade Federal de Pernambuco)**

**Dr. Marcelo Henrique Santos Paiva
(Centro de Pesquisas Aggeu Magalhães)**

**Dra. Maria Alice Varjal de Melo-Santos
(Centro de Pesquisas Aggeu Magalhães)**

**Dr. Rafael Dhalia
(Centro de Pesquisas Aggeu Magalhães)**

**Dra. Claudia Maria Fontes de Oliveira
(Centro de Pesquisas Aggeu Magalhães)**

**APROVADO EM: ____/____/____
RECIFE - PE**

***À Cybelle Cristina Carvalho Rocha,
Dedico.
In memorian***

AGRADECIMENTOS

Às minhas orientadoras, Constância Flávia Junqueira Ayres Lopes e Thereza Amélia Soares, pela dedicação, paciência e toda competência demonstrada.

À Frederico Pontes, pela ajuda operacional, pelas idéias discutidas e pelo agradável convívio.

À Daniela Nadvorny, pelas idéias discutidas e pela amizade dispensada.

À CAPES, pela concessão da bolsa.

Ao PPGBA, pela oportunidade de ingressar e executar um curso de alta qualidade.

À Iêda Ferreira, pela ajuda no sequenciamento e nas análises dos genes.

Aos integrantes do Departamento de Entomologia do Centro de Pesquisas Aggeu Magalhães, pelo apoio e solicitude.

Aos membros do DQF (Departamento de Química Fundamental da UFPE), em especial à equipe do *Biomaterial Modelling Group*, pelo grande acolhimento e pelo aprendizado que me proporcionaram.

À professora Cleide Maria Ribeiro de Albuquerque, pelo apoio prestado durante o estágio a docência.

Aos professores e colegas do PPGBA, pelo convívio e por todo aprendizado.

“Mais amor à verdade do que simples apego às suas próprias convicções.”

Lema do cientista.

RESUMO

O mosquito *Anopheles gambiae* (Diptera: Culicidae) é considerado o principal vetor do *Plasmodium*, o agente etiológico da malária, a doença parasitária que mais leva ao óbito em todo o mundo. O uso intensivo de alguns inseticidas químicos, entre os quais o DDT, direcionados para o controle desse vetor, levou à seleção de linhagens resistentes de *An. gambiae*. Desta forma, os mecanismos de resistência aos inseticidas vêm sendo amplamente estudados com o intuito de desenvolver novas estratégias de controle populacional do vetor. As glutathione S-transferases (GSTs) são enzimas de detoxificação celular que desempenham um importante papel biológico no metabolismo de xenobióticos através da conjugação da glutathione reduzida (GSH), tornando-os mais solúveis e facilmente excretados da célula. As GSTs da classe epsilon em *An. gambiae* (AgGSTE) apresentam atividade anti-DDT, especialmente a AgGSTE2, cuja estrutura encontra-se disponível no PDB. Também já foi demonstrado que a enzima AgGSTE5, cuja estrutura tridimensional ainda não foi elucidada, apresentou super expressão em presença do DDT. Assim, o objetivo do presente trabalho foi construir e validar um modelo tridimensional para eliciação da estrutura da AgGSTE5 através da modelagem comparativa e simular a dinâmica molecular da AgGSTE2 e AgGSTE5 e de uma isoforma mutante (AgGSTE2mut). Nas simulações de Dinâmica Molecular (DM) foram feitas por um período de 50 nanossegundos com e sem o ligante (GSH). Após a dinâmica, as três enzimas foram submetidas ao *docking* molecular contra os compostos DDT, CDNB, carbaril, malation e cipermetrina. Também foi analisado o polimorfismo genético e a taxa de mutação para os genes *AgGSTE2* e *AgGSTE5*. A análise da sequência dos genes apontou seleção purificadora para o *AgGSTE2* e seleção positiva para o *AgGSTE5* em populações de sete países da África. Os resultados demonstraram que as proteínas têm dinâmicas diferentes e interagem com os substratos de modo diferente. As mutações da AgGSTE2mut alteram sua dinâmica e modo de ação, sendo esta enzima particularmente capaz de se ligar ao DDT, com energia de ligação menor que as outras. Finalmente, os resultados do presente trabalho sugerem que estas enzimas desempenharam um papel crucial na adaptação de *An. gambiae* ao seu habitat e possivelmente a evolutibilidade destas GSTs teve participação neste processo evolutivo, sendo portanto alvos potenciais para o desenvolvimento de novas ferramentas de controle. Conclui-se que o papel da AgGSTE2 e AgGSTE5 na metabolização de inseticidas é importante para a adaptação do *An. gambiae* e o modo de ação destas enzimas deve ser entendido como uma importante via metabólica a ser interferida com o propósito de melhorar os inseticidas e métodos de controle. Os resultados permitem concluir que o modelo teórico é válido e que a AgGSTE2, AgGSTE2mut e AgGSTE5 apresentam diferenças na dinâmica e no modo de ligação aos compostos químicos estudados, o que provavelmente reflete em uma divergência funcional destas enzimas.

Palavras-chave: GST. *Anopheles gambiae*. Resistência a inseticidas. Dinâmica molecular.

ABSTRACT

Anopheles gambiae mosquito is considered the primary *Plasmodium* vector, which is etiological malaria agent, the parasitic disease that is a major cause of death in the world. Intensive use of pesticide to control mosquitoes in the last decades, especially by DDT application, has led to the emergence of resistant *An. gambiae* populations. Such resistance mechanisms have been widely studied with the goal of develop new strategies for the vector population control. Glutathione S-transferases (GSTs) are detoxification enzymes that play an important role in the metabolism of xenobiotics through the conjugation of the reduced glutathione, resulting in a more hydro soluble molecule and readily extractable from the cell. The epsilon class GSTs in *Anopheles* (AgGSTE) has already shown anti-DDT activity, especially the AgGSTE2, which its crystal structure is available in the PDB (Protein Data Bank) database. It has also been demonstrated that the AgGSTE5, which its three-dimensional structure has not yet been elucidated, has been over expressed in DDT presence. Therefore, the goal of the present study was to build and to validate a three-dimensional structural model for AgGSTE5 protein using comparative modeling and compare the molecular dynamics of these two proteins. Besides that, a mutant isoform (AgGSTE2mut) was also evaluated. The Molecular Dynamic (DM) was performed for a period of 50 ns, with and without the ligant (GSH). Those three enzymes were used for molecular docking against DDT compounds, CDNB, carbaryl, malathion and cypermethrin. In addition, the genetic polymorphism and mutation rate for the genes *AgGSTE2* and *AgGSTE5* were also analyzed. The gene sequence analysis showed purifying selection for *AgGSTE2* and positive selection for *AgGSTE5*. The results showed that each protein has different dynamics and interact with the substrate differently, which means that it is resulting in functional divergence. Mutations on AgGSTE2mut alter its dynamics, which makes this enzyme particularly capable of binding with DDT, with lower biding energy than the others. The results lead to the conclusion that these enzymes have played a crucial role in the *An. gambiae* adaptation to its habitat and possibly the evolvability of these GSTs had a role in the evolutionary process, thus being potential targets for the development of biotechnological tools for mosquito control and the widening of the knowledge concerning this enzyme family is necessary to reach this goal.

Keywords: GST. *Anopheles gambiae*. insecticide resistance. Molecular dynamics.

LISTA DE ILUSTRAÇÕES

Número	Descrição	Página
Figura 1	Esquema da ação das Glutathione S-transferases. Formação do conjugado GSH – Xenobiótico.	19
Figura 2	Representação química da glutathione, formada pelos aminoácidos: ácido glutâmico (verde), cisteína (preto) e glicina (vermelho).	23
Figura 3	Esquema de reação de dehidrocloração do DDT. Nesta reação observa-se que não ocorre o GSH-conjugado.	24
Figura 4	Estrutura tridimensional da AgGSTE2. A estrutura secundária da proteína está representada em <i>New Cartoon</i> nas cores azul e violeta. A glutathione (ligante) está representada em esferas com os átomos de carbono em verde, de oxigênio em vermelho, nitrogênio em azul e o átomo de enxofre em dourado.	25
Figura 5	Ciclo de vida dos culicídeos.	30
Figura 6	Esquema de dinamismo conformacional e promiscuidade de uma proteína.	36
Figura 7	Mapa da África mostrando os locais de coleta das amostras usadas na análise de polimorfismo genético.	40
Figura 8	Representação dos termos dos campos de forças utilizados em Dinâmica Molecular clássica.	44
Figura 9	Esquema de <i>docking</i> entre proteína-ligante.	53
Figura 10	Resultado do alinhamento entre as sequências primárias da AgGSTE2 (<i>Query</i>) e AgGSTE5 (<i>Sbjct</i>).	57
Figura 11	Diagrama de Ramachandran da AgGSTE2 e AgGSTE2mut com 92% dos resíduos em regiões de conformação mais favorável (vermelho), 7,5 % em regiões permitidas, 0,3% em regiões desfavoráveis (bege) e 0,3% em regiões proibidas (branco).	58

Número	Descrição	Página
Figura 12	Diagrama de Ramachandran da AgGSTe5 com 88,7% dos resíduos em regiões de conformação mais favorável (vermelho), 9,7 % em regiões permitidas, 0,5% em regiões desfavoráveis (bege) e 1,0% em regiões proibidas (branco).	59
Figura 13	Gráficos do Anolea, Qmean e Gromos em função do teor da estrutura secundária (DSSP) para AgGSTe2 e AgGSTe2mut.	60
Figura 14	Gráficos do Anolea, Qmean e Gromos em função do teor da estrutura secundária (DSSP) para AgGSTe5.	61
Figura 15	Comparação do <i>Z-score</i> da AgGSTe2 com acessos não-redundantes do PISA.	62
Figura 16	Valores de RMSD para as três proteínas sem (A) e com (B) o ligante GSH.	69
Figura 17	Gráfico da distribuição do RMSD para AgGSTe2 (sem ligante).	71
Figura 18	Gráfico da distribuição do RMSD para AgGSTe2mut (sem ligante).	71
Figura 19	Gráfico da distribuição do RMSD para AgGSTe5 (sem ligante).	72
Figura 20	Gráfico da distribuição do RMSD para AgGSTe2 (com ligante).	73
Figura 21	Gráfico da distribuição do RMSD para AgGSTe2mut (com ligante).	73
Figura 22	Gráfico da distribuição do RMSD para AgGSTe5 (com ligante).	74
Figura 23	RMSF para o carbono-alfa nas proteínas com (A) e sem (B) GSH.	75
Figura 24	Raio de giro nas proteínas simuladas com (A) e sem (B) o ligante.	76
Figura 25	Superfície de acessibilidade ao solvente. A: sistemas simulados sem a glutatona. B: os mesmos sistemas simulados com a glutatona.	78

Número	Descrição	Página
Figura 26	Teor da estrutura secundária dos sistemas simulados: A) AgGSTE2 sem ligante; B) AgGSTE2 com ligante; C) AgGSTE5 sem ligante; D) AgGSTE5 com ligante; E) AgGSTE2mut sem ligante; F) AgGSTE2mut com ligante.	78
Figura 27	Autovalores (<i>Eigenvalues</i>) x autovetores (<i>p</i>) para átomos do <i>backbone</i> nos sistemas sem GSH.	79
Figura 28	Autovalores (<i>Eigenvalues</i>) x autovetores (<i>Eigenvectors</i>) para átomos do <i>backbone</i> nos sistemas com GSH.	82
Figura 29	RMSF para o primeiro autovetor nos sistemas sem (A) e com (B) GSH.	83
Figura 30	Regiões de maior contribuição para o primeiro autovetor das enzimas AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C) sem o ligante.	84
Figura 31	Regiões de maior contribuição para o primeiro autovetor das enzimas AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C) com o ligante.	85
Figura 32	Número de estruturas (populações) por cluster nos sistemas sem (A) e com (B) GSH.	88
Figura 33	Número de transições entre os clusters para as proteínas sem (A) e com (B) GSH.	92
Figura 34	Representação das melhores conformações dos complexos AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C). Em azul, resíduos dentro de um raio de 4 angstroms; em verde o DDT; em vermelho o GSH.	93
Figura 35	Representação das melhores conformações dos complexos AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C). Em azul, resíduos dentro de um raio de 4 angstroms; em verde o CDNB; em vermelho o GSH.	96

Número	Descrição	Página
Figura 36	Representação das melhores conformações dos complexos AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C). Em azul, resíduos dentro de um raio de 4 angstroms; em amarelo o carbaril; em vermelho o GSH.	97
Figura 37	Representação das melhores conformações dos complexos AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C). Em azul, resíduos dentro de um raio de 4 angstroms; em amarelo a cipermetrina; em vermelho o GSH.	97
Figura 38	Representação das melhores conformações dos complexos AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C). Em azul, resíduos dentro de um raio de 4 angstroms; em púrpura o malation; em vermelho o GSH.	98
Figura 39	Representação das melhores conformações dos complexos AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C). Em azul, resíduos dentro de um raio de 4 angstroms; em laranja o ácido úsnico aniônico; em vermelho o GSH.	98
Figura 40	Representação das melhores conformações dos complexos AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C). Em azul, resíduos dentro de um raio de 4 angstroms; em laranja o ácido úsnico neutro; em vermelho o GSH.	99
Figura 41	Representação da melhor conformação do complexo AgGSTE2-cipermetrina. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde a cipermetrina.	99
Figura 42	Representação da melhor conformação do complexo AgGSTE2mut-cipermetrina. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde a cipermetrina.	100

Número	Descrição	Página
Figura 43	Representação da melhor conformação do complexo AgGSTE5-cipermetrina. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde a cipermetrina.	100
Figura 44	Representação da melhor conformação do complexo AgGSTE2-DDT. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o DDT.	101
Figura 45	Representação da melhor conformação do complexo AgGSTE2mut-DDT. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o DDT.	101
Figura 46	Representação da melhor conformação do complexo AgGSTE5-DDT. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o DDT.	102
Figura 47	Representação da melhor conformação do complexo AgGSTE2-malation. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o malation.	102
Figura 48	Representação da melhor conformação do complexo AgGSTE2mut-malation. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o malation. O GSH	103
Figura 49	Representação da melhor conformação do complexo AgGSTE5-malation. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o malation.	103

LISTA DE TABELAS

Número	Descrição	Página
Tabela 1	Sistemas submetidos à simulação de DM.	46
Tabela 2	Compostos utilizados como ligantes para o cálculo de <i>docking</i> .	52
Tabela 3	Diversidade genética e teste de neutralidade de Tajima em seqüências de nucleotídeo do <i>AgGSTE2</i> e <i>AgGSTE5</i> . m = número de seqüências, S = número de sítios segregantes, p_s = número de sítios segregantes por sítios de nucleotídeo (S/m), $\theta = (p_s/a_1)$, π = diversidade nucleotídica, e D é o teste de Tajima.	55
Tabela 4	Teste D de Tajima para seqüências de nucleotídeo no gene <i>AgGSTE2</i> . #: $P < 0.10$; *: $P < 0.05$; **: $P < 0.01$.	55
Tabela 5	Teste D de Tajima para seqüências de nucleotídeo no gene <i>AgGSTE5</i> . #: $P < 0.10$; *: $P < 0.05$; **: $P < 0.01$.	56
Tabela 6	Resultados do <i>Qmean</i> para a <i>AgGSTE2</i> .	63
Tabela 7	Resultados do <i>Qmean</i> para a <i>AgGSTE5</i> .	65
Tabela 8	Energias de ligação (Kcal/mol) para as melhores conformações no ranking de cada complexo.	66
Tabela 9	Resíduos e átomos que constituem o sítio catalítico da <i>AgGSTE5</i> de acordo com o <i>Q-sitefinder</i> . Cada coluna representa a posição do átomo, o tipo do átomo, o tipo de aminoácido, a subunidade a qual o átomo pertence e o número do aminoácido	68
Tabela 10	Energias de ligação (Kcal/mol) para as melhores conformações no ranking de cada	95

LISTA DE ABREVIATURAS E SIGLAS

AgGSTE2 = enzima 2 da classe epsilon da Glutathione S-transferase de *Anopheles gambiae*.**

AgGSTE2 = gene 2 da classe epsilon da Glutathione S-transferase de *Anopheles gambiae*.***

AgGSTE2mut = isoforma mutante da enzima 2 da classe epsilon Glutathione S-transferase de *Anopheles gambiae*.

AgGSTE5 = enzima 5 da classe epsilon da Glutathione S-transferase de *Anopheles gambiae*.**

AgGSTE5 = gene 5 da classe epsilon da Glutathione S-transferase de *Anopheles gambiae*.***

ANOLEA = Atomic Non-Local Environment Assessment.

AU = ácido úsnico.

BLAST = *Basic Local Alignment Sequence Tool*.

CDNB = 1-Chloro-2,4-dinitrobenzene

DDE = Diclorodifenilcloroetano.

DDT = Dicloro-Difenil-Tricloroetano

DM = Dinâmica Molecular

DNA_{SP} = DNA Sequence polymorphism software.

DSSP = Dictionary of Secondary Structure of Proteins.

EM = Minimização de energia.

FW = *primer Forward*

GROMOS = um campo de forças para dinâmica molecular.

GSH = Glutathione reduzida.#

GST = Glutathione S-transferase.*

kDA = Kilodaltons.

LGA = Algoritmo Genético Lamarckiano.

OMS = Organização Mundial de Saúde.

PCA = Análise dos Componentes Principais.

PCR = Reação em Cadeia da Polimerase.

PDB = *Protein Data Bank*.

PISA = *Proteins, Interfaces, Surfaces and Assemblies*.

P-LINCS = Parallel Linear Constraint Solver algorithm.

PMDB = Protein Model Data Base.

PROCHECK = Protein Check software.

PSQS = Protein Structure Quality Score.

QMEAN = *Quality mean score*.

RG = raio de giro.

RMSD = *root mean square deviation*, ou “desvios médios quadráticos”.

RMSF = *root mean square fluctuations*, ou “flutuações médias quadráticas”.

RV = *Primer Reverse*

SASA = superfície de acessibilidade ao solvente.

SEATTLE = Seattle algorithm.

SPC = modelo “Single Point charge”.

VMD = Visual Molecule Dynamics software.

De acordo com a nomenclatura da IUPAC.

*De acordo com as regras IUBMB Enzyme Nomenclature EC 2.5.1.18

**De acordo com a proposta presente em [Che-Mendonza et al., 2009], em *Insecticide Resistance and glutathione S-transferase in mosquitoes: a review*, 8 (8) 1386-1387.

***De acordo com as regras e convenções para nomes de genes em Anofelinos (<http://skonops.imbb.forth.gr/names.html>). Parte destas regras foram publicadas no material suplementar de [Christophides et al., 2002, *Science* 298: 159-165], enquanto regras adicionais foram adaptadas das normas de nomenclatura de *Drosophila*.

LISTA DE SÍMBOLOS

α	Letra grega alfa, encontrada em “ α -hélice”.
β	Letra grega beta, encontrada em “Folha- β ”.
Å	Angstrom.
Φ	Letra grega phi, ângulo diedral.
Ψ	Letra grega psi, ângulo diedral.
h	Número de haplótipos
S	Número de sítios de segregação
S/m	Número de sítios segregantes por nucleotídeo
π	Diversidade nucleotídica
θ/w	Taxa evolutiva por sítio
nm	nanômetros.
ns	nanossegundos.

SUMÁRIO

1	INTRODUÇÃO.....	17
1.1	A MALÁRIA COMO PRINCIPAL PARASITOSE TROPICAL.....	17
1.2	BIOLOGIA DO <i>Anopheles gambiae</i> GILES (DIPTERA: CULICIDAE).....	18
1.3	O PAPEL DAS GLUTATIONA S-TRANSFERASES NA RESISTÊNCIA A INSETICIDAS QUÍMICOS.....	20
1.4	DETOXIFICAÇÃO DE INSETICIDAS PELAS GSTs.....	25
1.4.1	Organofosforados (OF).....	26
1.4.2	Organoclorados (OC).....	26
1.4.3	Carbamatos (CAR).....	28
1.4.4	Piretróides (PIR).....	28
1.5	DOCKING MOLECULAR	30
1.6	ENZIMAS PROMÍSCUAS, EVOLUTIBILIDADE E ATIVIDADE DETOXIFICADORA.....	31
1.7	RESISTÊNCIAS A INSETICIDAS QUÍMICOS, PRESSÃO SELETIVA E IMPLICAÇÕES PARA O CONTROLE DA MALÁRIA.....	34
1.8	A IMPORTÂNCIA DA BIOINFORMÁTICA ESTRUTURAL NA COMPREENSÃO DE ENZIMAS.....	36
2	OBJETIVOS.....	38
2.1	OBJETIVO GERAL.....	38
2.2	OBJETIVOS ESPECÍFICOS.....	38
3	MATERIAL E MÉTODOS.....	39
3.1	POLIMORFISMO GENÉTICO NOS GENES <i>AgGSTE2</i> E <i>AgGSTE5</i>	39
3.2	PROCEDIMENTO COMPUTACIONAL.....	41
3.2.1	Modelagem da <i>AgGSTE5</i>	41
3.2.2	Dinâmica molecular.....	42
3.3	ANÁLISE DOS DADOS.....	47
3.4	DOCKING MOLECULAR.....	51
4	RESULTADOS.....	54
4.1	POLIMORFISMO GENÉTICO DOS GENES <i>AgGSTE2</i> E <i>AgGSTE5</i>	54
4.2	ALINHAMENTO E MODELAGEM ESTRUTURAL DA <i>AgGSTE5</i>	56
4.2.2	Validação do modelo.....	58

4.2.3	Volume dos sítios de ligação.....	64
4.3	DINÂMICA MOLECULAR.....	69
4.3.1	Estabilidade conformacional.....	69
4.3.2	Conteúdo da estrutura secundária ao longo do tempo.....	78
4.3.3	Análise de componentes principais.....	81
4.3.4	Análise dos clusters.....	91
4.4	DOCKING MOLECULAR.....	94
4.4.1	Predição do modo de ligação.....	94
5	DISCUSSÃO.....	104
6	CONCLUSÕES.....	115
	REFERÊNCIAS.....	116
	APÊNDICES.....	132

1 INTRODUÇÃO

1.1 A MALÁRIA COMO PRINCIPAL PARASITOSE TROPICAL

A malária ou paludismo é uma parasitose causada por protozoários do gênero *Plasmodium* e é responsável pelo óbito de três milhões de pessoas por ano e afeta anualmente mais de 500 milhões de pessoas (SACHS; MALANEY, 2002). É a principal causa de morte no continente africano, e no mundo mais de dois bilhões de pessoas vivem em áreas de risco (BREMEN, 2001). A malária está distribuída em praticamente todos os países localizados nas regiões tropicais e alguns subtropicais, sendo que mais de 90% dos casos de morte prevalecem na África subsaariana (WHO, 2005).

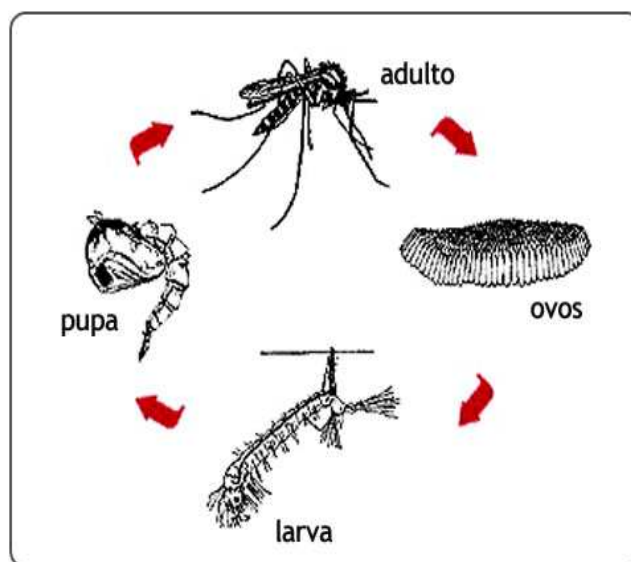
O principal modo de transmissão do agente etiológico da malária é através das picadas das fêmeas dos mosquitos do gênero *Anopheles*, podendo também ser transmitida por outras formas mais raras tais como transfusão sanguínea, uso de seringas contaminadas, acidentes de laboratório e por ocasião de parto (SACHS; MALANEY, 2002).

Devido à letalidade e alta morbidade, o impacto sócio-econômico da malária, tanto nos países endêmicos como no mundo, vem preocupando os especialistas e estudos em 31 países africanos apontam que o prejuízo causado pela malária totaliza mais de 73 bilhões de dólares, e vários países chegam a ter um déficit anual na sua economia de até 18% devido à doença (SACHS; MALANEY, 2002). Entretanto, apesar dos resultados dos programas anti-malaria serem óbvios, o nível de investimento internacional para o controle desta doença é ainda muito baixo, girando em torno dos 100 milhões de dólares anuais, enquanto estimativas da Comissão de Macroeconomia e Saúde da OMS (Organização Mundial da Saúde) estabeleceram que são necessários anualmente 2,5 bilhões de dólares para um programa de prevenção e tratamento

realmente efetivo no mundo (SACHS; MELANEY, 2002). Além disso, diversos efeitos negativos e indiretos vêm sendo anotados, como tornar os indivíduos mais suscetíveis a outras infecções, e aumentar o risco de contrair outras doenças, como acontece com crianças que contraem HIV devido a transfusões realizadas durante o tratamento da malária. Também é relatada como um importante fator causal da anemia, que interfere diretamente na produtividade, capacidade física e saúde dos trabalhadores (BASTA *et al.*, 1979; SCHOLZ, *et al.*, 1997).

1.2 BIOLOGIA DO *Anopheles gambiae* GILES (DIPTERA: CULICIDAE)

O mosquito *Anopheles gambiae* é considerado o principal vetor envolvido na transmissão da malária no continente africano (HEMINGWAY; RANSON, 2000; HEMINGWAY *et al.*, 2004). Trata-se de uma espécie adaptada ao domicílio e peridomicílio, com adultos que frequentemente habitam o interior das casas. As larvas proliferam em pequenas coleções de águas rasa e limpas no solo, em geral pobres em vegetação e expostas à intensa luz solar (CONSOLI; LOURENÇO-DE-OLIVEIRA, 1994; MINAKAWA, 2004). O *An. gambiae* pertence à família Culicidae e subfamília Anophelinae, cujos indivíduos adultos distinguem-se dos demais culicídeos por pousarem em posição perpendicular e as larvas por apresentarem orifícios respiratórios abdominais, ao invés de sifão respiratório (FORATTINI, 1996). São insetos holometábolos, com ciclo de vida médio de 10-12 dias, que compreende os estágios de ovo, larva, pupa e adulto (Figura 1) cujas fêmeas apresentam hábitos hematofágicos antropofílicos e zoofílicos, sendo o homem o hospedeiro preferencial. Os machos não realizam o repasto sanguíneo e por isso não estão envolvidos na transmissão de doenças.

Figura 1 - Ciclo de vida dos culicídeos.

Fonte: www.mosquito-pictures.com

Devido à sua importância no ciclo da malária, o *An. gambiae* é alvo de diversos estudos visando a implementação de novas estratégias de controle. Existem dois principais esquemas de controle da malária: a) diagnóstico e tratamento dos indivíduos infectados; e b) controle populacional dos mosquitos vetores (RANSON *et al.*, 2009). O tratamento dos indivíduos (humanos) é realizado através de medicamentos diversos, sendo os mais utilizados os derivados de artemisinina e de quinina (RANSON *et al.*, 2009). Já o controle de vetores é realizado por várias maneiras, como a eliminação de criadouros, o controle biológico, controle químico e o controle integrado (RANSON *et al.*, 2009). No entanto, nas últimas décadas vem sendo observado o surgimento de várias populações de *An. gambiae* resistentes à inseticidas químicos, resultantes do uso intensivo destes compostos no passado (RANSON *et al.*, 2009).

1.3 O PAPEL DAS GLUTATIONA S-TRANSFERASES NA RESISTÊNCIA A INSETICIDAS QUÍMICOS

Dos mecanismos de resistência aos inseticidas químicos destaca-se as vias de metabolização de compostos tóxicos, em decorrência da superexpressão de enzimas de detoxificação ou de modificações estruturais nestas enzimas. As glutathione S-transferases (GSTs) constituem uma das mais importantes famílias de enzimas envolvidas com esse tipo de resistência é a família, e compreendem enzimas que catalisam reações que transformam vários compostos xenobióticos em produtos solúveis (HAYES *et al.*, 2005). Em organismos eucariontes estas enzimas são classificadas em: GSTs citosólicas, GSTs microsossomais (associadas às membranas) e GSTs mitocondriais (ROBINSON *et al.*, 2004; SHEEHAN *et al.*, 2001). Em insetos, foi constatada a presença de apenas duas dessas classes: as citosólicas e microsossomais, não sendo encontrada nenhuma GST da classe mitocondrial em insetos até o momento (LUMJUAN *et al.*, 2007; ENAYATI *et al.* 2005), e portanto, sem evidências de GSTs mitocondriais em dípteros (DING *et al.*, 2003). As GSTs microsossomais catalisam reações muito similares às citosólicas, com estrutura trimérica e estando associadas às membranas plasmáticas, Embora tenham estruturas e origens diferentes das citosólicas (PRABHU *et al.*, 2004; PEARSON *et al.*, 2005). No entanto, as GSTs citosólicas já foram apontadas como importantes para a resistência aos inseticidas químicos (HEMINGWAY *et al.*, 2004; ENAYATI *et al.*, 2005; RANSON; HEMINGWAY, 2005) enquanto as GSTs microsossomais não foram até o momento relacionadas à resistência a inseticidas (ENAYATI *et al.*, 2005).

Antes da publicação do genoma e dos mapas gênicos de *Anopheles gambiae* (RANSON *et al.*, 2002; HOLT *et al.*, 2002) o principal critério para a classificação das GSTs era baseado na homologia das sequências de aminoácidos e em propriedades imunológicas (TOUNG *et al.*, 1990; BEALL *et al.*, 1992; FOURNIER *et al.*, 1992). Devido à ampla capacidade das GSTs de catalisarem vários substratos diferentes é comum que mais de um tipo de GSTs metabolize os mesmos compostos, resultando em sobreposição da especificidade. Portanto, outros aspectos são

atualmente utilizados para incluir GSTs em uma mesma classe, tais como: identidade de sequência de aminoácidos acima de 40%, propriedades imunológicas, relações filogenéticas, estrutura terciária, a habilidade de formar heterodímeros e sua localização nos cromossomos (DING *et al.*, 2003; HEMINGWAY *et al.*, 2004; RANSON e HEMINGWAY, 2005).

Nos insetos, as GSTs citosólicas estão representadas, no mínimo, por seis classes: delta, epsilon, omega, sigma, theta e zeta (RANSON *et al.*, 2002; ENAYATI *et al.* 2005; CHEMENDOZA *et al.*, 2009), sendo encontrados membros destas seis classes de GST em culicídeos, como o vetor da malária *An. gambiae*, e na mosca-da-fruta *Drosophila melanogaster* (RANSON *et al.*, 2002). As classes delta e epsilon são artrópode-específicas e representam mais de 65% do total de GSTs citosólicas encontradas nestes organismos (RANSON *et al.*, 2002). A maior parte de GSTs encontradas em insetos e envolvidas no metabolismo de compostos xenobióticos pertence a essas duas classes, enquanto que as demais classes (omega, sigma, theta e zeta) possuem uma distribuição bem mais ampla entre os grupos taxonômicos, desde bactérias até vertebrados (BOARD *et al.*, 2000; WILDENBURG *et al.*, 1998).

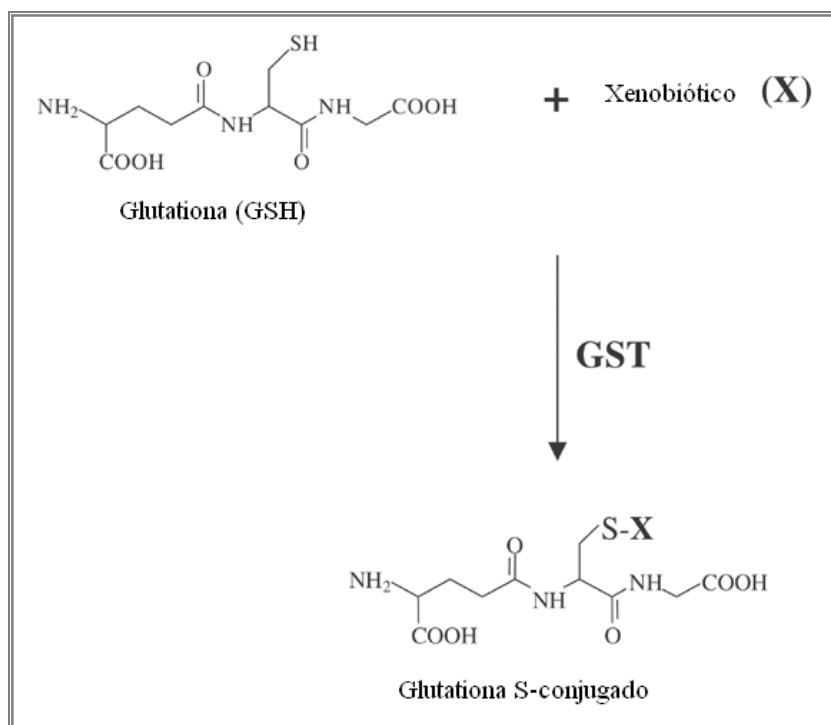
Os membros das classes delta, sigma e epsilon inicialmente eram chamados de classes I, II e III respectivamente, e posteriormente, com o aumento no número de sequências depositadas nas bases de dados e de estudos de classificação, foi adotada a nomenclatura baseada no alfabeto grego em acordo com o sistema de nomenclatura de GSTs de mamíferos (CHELVANAYAGAM *et al.*, 2001). Esta classificação teve o suporte de análises filogenéticas tanto em GSTs de mamíferos como de insetos (RANSON *et al.*, 2001, DING *et al.*, 2003). Atualmente a nomenclatura das GSTs de insetos consiste de três partes: o nome da espécie da qual a GST pertence; a classe específica da GST e o número que especifica a ordem em que a proteína foi descoberta. Deste modo utiliza-se, por exemplo, o nome AgGSTD1 para designar uma GST de

An. gambiae, membro da classe delta, sendo esta a primeira proteína desta classe a ser descoberta (CHE-MENDONZA *et al.*, 2009).

As GSTs citosólicas são compostas de duas subunidades de aproximadamente 25 kDa cada, podendo ser homodiméricas ou heterodiméricas. Cada subunidade apresenta um sítio de ligação específico para glutathione (sítio G), próximo a um sítio eletrofílico não conservado (sítio H). O sítio G fica situado na terminação N da proteína e apresenta-se como uma região altamente conservada nas GSTs. Já os resíduos do sítio H, que interagem com os substratos hidrofóbicos, são encontrados na terminação C. A diversidade no sítio H faz as GSTs apresentarem diferentes especificidades em relação aos substratos que metabolizam (MANNERVIK; DANIELSON, 1988; RANSON *et al.*, 1998). A reação catalisada pelas GSTs consiste em promover a conjugação do tripeptídeo glutathione reduzida (GSH) a um composto específico e geralmente citotóxico, que ao ligar-se a tal grupamento eletrofílico, passará do estado reduzido para o estado oxidado, e formará um composto mais solúvel e mais fácil de ser excretado da célula (Figura 2). Esta fase de conjugação representa a fase II do processo de desintoxicação celular e as GSTs representam as enzimas mais importantes desta fase, embora hajam outras envolvidas.

O resíduo ativo do sítio catalítico das GSTs tende a ser conservado dentro de uma mesma classe, porém difere entre classes. Na maioria das GSTs humanas, por exemplo, o resíduo responsável pela ativação do grupo tiol aparenta ser uma tirosina (SHEEMAN *et al.*, 2001). Já nas classes delta e epsilon esta atividade parece estar atribuída a uma serina (Ranson e HEMINGWAY *et al.*, 2005; UDOMSINPRASERT, 2005). Embora cada subunidade tenha um sítio ativo cineticamente independente, por se tratarem de proteínas diméricas as GSTs dependem essencialmente da estrutura quaternária para a catálise (DANIELSON; MANNERVIK, 1985).

Figura 2 - Esquema da ação das Glutathione S-transferases. Formação do conjugado GSH – Xenobiótico.



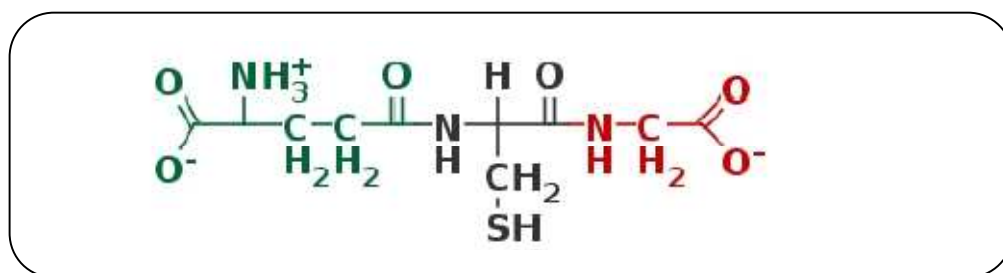
Fonte: Townsend e Tew (2003).

A glutathione é um tripeptídeo solúvel em água composto pelos aminoácidos ácido glutâmico, cisteína e glicina (Figura 3). O seu grupo tiol é um potente agente redutor, fazendo da glutathione o grupo de pequenas moléculas tióis mais abundantes no meio intracelular, alcançando geralmente concentrações na ordem de milimolares em alguns tecidos. Como é um importante agente antioxidante, a glutathione desempenha diversos papéis na detoxificação de inúmeros compostos eletrofílicos através da ação catalítica das GSTs (TOWNSEND *et al.*, 2003).

As GSTs da classe epsilon (GSTE) constituem uma das classes citosólicas mais importantes no papel de conferir resistência aos organoclorados e organofosforados, e também foi comprovado um papel na resistência à piretróides em culicídeos, como *Anopheles culicifacies*,

An. stephensi e *An. fluviatilis* (DJADID *et al*, 2006). Estudos de modelagem e *docking* molecular foram utilizados para desvendar o papel de proteínas detoxificadoras na resistência a inseticidas, identificando enzimas com capacidade de metabolização para o DDT e furanocumarinas (metabólitos secundários de plantas), assinalando proteínas-alvo para o desenvolvimento de inibidores específicos nas monooxigenases (CHIU *et al*, 2008, BAUDRY *et al*, 2003).

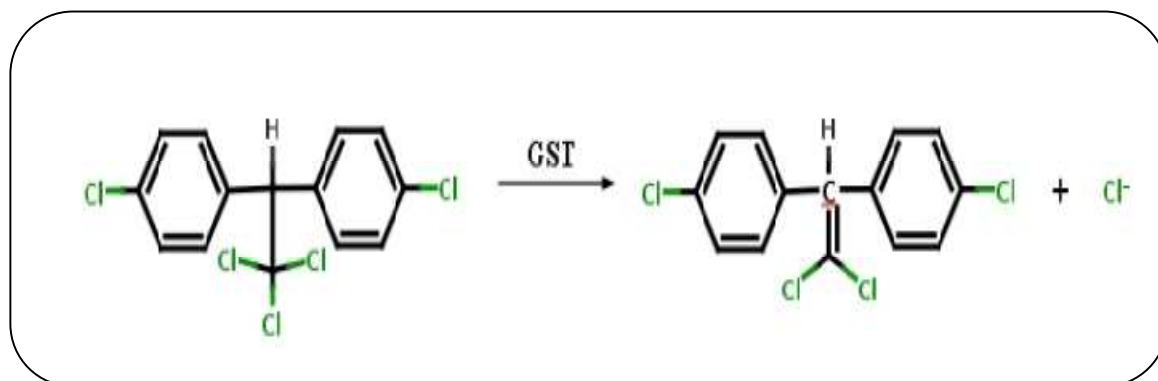
Figura 3 - Representação química da glutatona, formada pelos aminoácidos: ácido glutâmico (verde), cisteína (preto) e glicina (vermelho).



Fonte: Ilustração produzida pelo autor.

Outra reação que as GSTs podem apresentar é a de dehidroclorinação, também chamada de atividade DDTase. A reação ocorre quando uma GST remove um átomo de hidrogênio do DDT e conseqüentemente resulta na perda de um átomo de cloro, convertendo o DDT em DDE (Diclorodifenil Dicloroetano), um metabólito não tóxico (Figura 4) (HEMINGWAY *et al*, 2004). Apesar de nunca ter sido identificado o conjugado entre o GSH e o DDT, este tripeptídeo é um indispensável co-fator na reação de dehidroclorinação (HEMINGWAY 2000; SHEEHAN *et al.*, 2001; VONTAS *et al.*, 2001; ENAYATI *et al.*, 2005; FROVA 2006; LI *et al*, 2007; LUMJUAN *et al.*, 2007).

Figura 4 - Esquema de reação de dehidrocloração do DDT pela GST. Nesta reação observa-se que não ocorre o GSH-conjugado.



Fonte: Ilustração produzida pelo autor.

1.4 DETOXIFICAÇÃO DE INSETICIDAS PELAS GSTs

A notável capacidade de detoxificação das GSTs tem implicação direta na metabolização de vários xenobióticos com atividade inseticida. Os principais grupos de inseticidas nos quais já foi observada a ação de GSTs são os organoclorados (OC), os organofosforados (OF) e os piretróides (PIR).

1.1.1 Organofosforados (OF)

A detoxificação dos OFs pelas GSTs pode ocorrer por duas vias metabólicas distintas: por uma O-desalquilação ou por uma reação O-desarilação. Na O-desalquilação o conjugado GSH é formado por uma ligação no grupamento alquil do inseticida. Na O-desarilação o GSH reage com

o grupo de saída. Estas reações já foram reportadas em espécies de insetos como *Musca domestica* (OPPENORTH *et al.*, 1979; UGAKI *et al.*, 1985) e em *Plutella xilostella* (CHING e SUN, 1993).

As GSTs atuam como um mecanismo secundário de resistência em conjunção com as P450 e esterases (HEMINGWAY, 1991 *apud* CHE-MEDOZA *et al.*, 2009). De modo geral, os OFs são aplicados em uma forma que não contem atividade inseticida, e uma vez absorvidos nos organismos, são ativados para uma forma de ação inseticida organofosfatada (um oxón-análogo), pela ação das P450 nos insetos (enzimas da fase I do processo de detoxificação). Estes óxons-análogos são ainda mais neurotóxicos que os tiolados-análogos e funcionam como potentes inibidores da acetilcolinesterase. Estudos comprovaram a detoxificação de oxón-análogos do inseticida fenitrothion no mosquito *Anopheles subpictus* (HEMINGWAY *et al.*, 1991 *apud* CHE-MEDOZA *et al.*, 2009). Tais sistemas de cooperação das enzimas para detoxificação tendem a ser mais rápidos e eficientes que mecanismos independentes sendo, portanto, vias metabólicas importantes para a resistência a inseticidas químicos (Bogwitz, 2005).

1.1.2 Organoclorados (OC)

Em relação aos organoclorados as GSTs catalisam duas reações de hidrocarbonetos halogenados: dehidrocloração e conjugação com o GSH (TANG; TU, 1994). Para o DDT, a reação de dehidrocloração consiste na principal via de detoxificação (HAYES; WOLF, 1988) e talvez seja o mecanismo de resistência mais frequente em mosquitos (BROWN, 1986; HEMINGWAY, 2000). Neste grupo, encontram-se os compostos sintéticos pioneiros no uso como inseticidas, porém a maioria dos organoclorados tem atualmente seu uso proibido no Brasil (ALMEIDA *et al.*, 2007). A proibição é devido a estes compostos apresentarem efeito residual

prolongado e manterem alta persistência no ambiente aumentando o risco de acumulação nos seres vivos, inclusive em humanos.

Nas reações de dehidroclorinação catalisadas pelas GSTs o grupamento tiol (S^-) gerado no sítio ativo do GSH remove um átomo de hidrogênio do DDT. Isto resulta diretamente na eliminação de um átomo de cloro e conseqüentemente gera um composto não tóxico, o DDE (Diclorodifenilcloroetano). O DDT foi um dos organoclorados mais utilizados em controle de insetos e vários estudos revelaram populações de insetos resistentes a este composto, sendo considerado em algumas localidades não mais eficaz no controle do vetor, fato que contribuiu para o reaparecimento da doença (MILLER, 2004). Nesta reação não ocorre alterações nos níveis de GSH ao final da catálise, porque ele atua mais como um co-fator do que como um conjugado (CLARK; SHAMAAN, 1984; LIPKE; CHALKLEY, 1962). O aumento da resistência ao DDT está associado a um aumento na taxa de dehidroclorinação por GSTs em *Aedes aegypti* (GRANT *et al.*, 1991; LUMJUAN *et al.*, 2005), *Anopheles dirus* (PRAPANTHADARA *et al.*, 1996, 2000b) e *An. gambiae* (PRAPANTHADARA *et al.*, 1993, 1995; RANSON *et al.*, 2001; ORTELLI *et al.*, 2003). Há também OCs, como o lindane, que são detoxificados inicialmente pela dehidroclorinação (TANAKA *et al.*, 1981) e em seguida conjugado à glutatona pela ação das GSTs (CLARK *et al.*, 1986; BLOOMQUIST, 1998; WEI *et al.*, 2001).

1.4.3 Carbamatos (CAR)

Os carbamatos são compostos derivados dos ácidos carbâmicos, e foram desenvolvidos durante a década de 1950 (CASIDA; QUISTAD, 1998), entre os quais o carbaril é um dos mais utilizados. São potentes inibidores das acetilcolinesterases e apresentam seletividade alta contra

as AChE de certas espécies (BRAGA; VALLE, 2007). O mecanismo primário da resistência ao carbamato é relacionado à mutação no gene *kdr* e (ODUOLA *et al.*, 2012). Contudo as GSTs contribuem também para a resistência multifatorial a este composto. Estudos apontam comprovações de que as GSTs desempenham um papel na metabolização dos carbamatos. Já foi reportado, por exemplo, o aumento na expressão de GSTs em *Octodrilus compalhnatus* (IBTISSEM *et al.*, 2012). Em um recente estudo com GSTs em *Locusta migratoria* foram testadas por recombinação e ensaios sugerindo que esta classe desempenha um papel significativo na metabolização do carbaril (QIN *et al.*, 2013). Ensaio bioquímicos em *An. funestus* níveis significativos de GSTs foram quantificados, o que indicou que há uma participação secundária na detoxificação deste agente (CUAMBA *et al* 2010).

1.4.4 Piretróides (PIR)

O papel das GSTs em detoxificar inseticidas PIRs reside na capacidade destas enzimas de reduzir o dano peroxidativo que são causados pelos PIRs, catalisando lipídeos originados de reações de peroxidação (VONTAS *et al.*, 2001). Este fato foi observado em uma GST da classe delta para o hemíptero *Nilaparvata lugens* (VONTAS *et al.*, 2002) e da GST classe epsilon 2 (AaGSTe2) em *Ae. aegypti* (LUMJUAN *et al*, 2005).

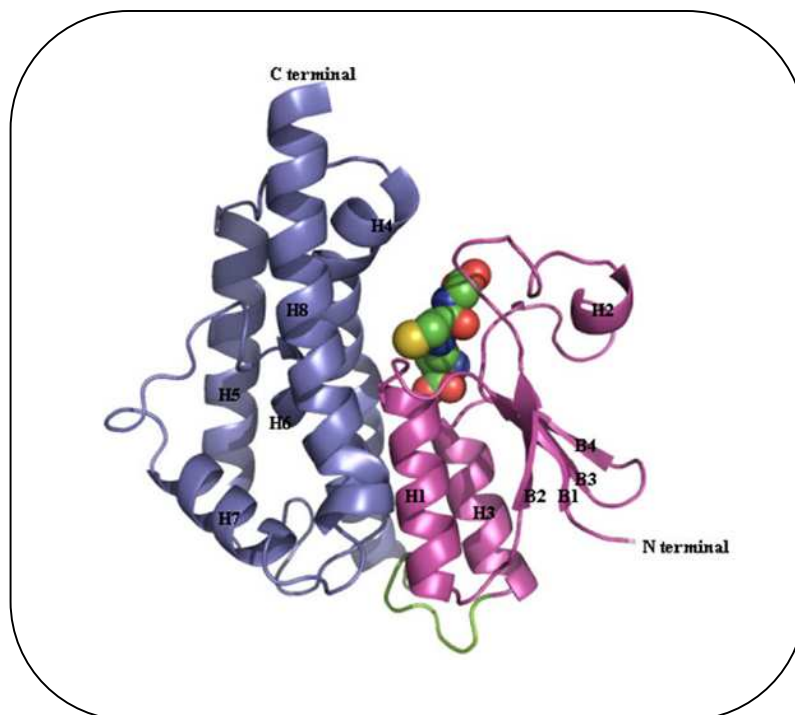
Já foi comprovado que GSTs podem proteger contra a toxicidade dos PIRs em insetos através de um processo de sequestro passivo (KOSTAROPOULOS *et al*, 2001). Vários trabalhos mostram a capacidade de algumas GSTs se ligarem a PIRs em *An. dirus* (PRAPANTHADARA *et al.*, 1998, 2000b; JIRAJAROENRAT *et al*, 2001; UDOMSINPRASERT; KETTERMAN, 2002 *apud* CHE-MEDOZA *et al*, 2009). A utilização de inibidores para GSTs em cepas de *Culex* resistentes a PIRs indicou que estas enzimas conferem resistência a estes inseticidas (XU *et al*,

2005). Em estudos de populações de campo de *Anopheles albimanus* foram detectados aumentos da expressão de GSTs quando sob seleção contínua por PIRs (PENILLA *et al*, 2006).

AYRES *et al*. (2011) caracterizaram seis membros da classe epsilon em cinco diferentes espécies de *Anopheles*. Neste trabalho foram feitos testes para detecção de seleção positiva sobre estes genes e os autores identificaram que o gene *AgGSTE5* continha códons que haviam sofrido efeito de seleção positiva, tipo de seleção que atua fixando as mutações não-sinônimas (que alteram o aminoácido) e portanto provoca o surgimento de novos fenótipos. Isto sugere que possivelmente, a proteína codificada por este gene, pode estar atuando em substratos diferentes da proteína *AgGSTE2*, a qual se mostrou bem mais conservada entre as diferentes espécies estudadas.

Estudos realizados com membros individuais de GSTs de *An. gambiae* evidenciaram eventos de duplicação e diversificação das sequências, que provavelmente estão relacionados com a adaptação ao seu nicho. Também foi constatada super-expressão de cinco das oito GSTs-epsilon em uma cepa de *An. gambiae* resistente ao DDT (DING *et al*, 2003). Destas, apenas a *AgGSTE2* teve a atividade anti-DDT confirmada (ORTELLI *et al*, 2003). A função destas enzimas é, portanto, um campo valioso de estudo a ser investigado. A caracterização funcional de uma proteína pode ser feita por experimentos laboratoriais ou por meio de técnicas computacionais sofisticadas como a dinâmica molecular (DM) e o docking molecular.

Figura 5 - Estrutura tridimensional da AgGSTE2. A estrutura secundária da proteína está representada em fita nas cores azul e violeta. A glutathiona (ligante) está representada em esferas com os átomos de carbono em verde, de oxigênio em vermelho, nitrogênio em azul e o átomo de enxofre em dourado.



Fonte: (WANG *et al.*, 2008)

1.5 DOCKING MOLECULAR

O *docking* molecular é uma técnica que visa prever a interação entre um receptor (proteína) e seu ligante por meio de simulações computacionais otimizadas e busca a melhor orientação do complexo proteína-ligante, proporcionando uma maior compreensão da função das enzimas (MORRIS *et al.*, 1998). Esta técnica mostra-se extremamente útil para compreender as GSTs em mosquitos e seu papel na resistência a inseticidas. Em um recente estudo com a AgGSTE2 e com o DDT, SETZER (2011) identificou os resíduos nas posições 112, 120, 36 e

116 (respectivamente Arginina, Fenilalanina, Leucina e Ácido glutâmico) como componentes do sítio ativo da AgGSTe2 que interagem diretamente com o DDT. Estes resíduos atuam formando um *pocket* que envolvem e interagem hidrofobicamente com o DDT. Também neste estudo o mesmo sítio de ligação da AgGSTe2 para o DDT mostrou-se um escore de energia de ligação ainda menor ao realizar o *docking* com o DDE, um metabólito do DDT com menor toxicidade (SETZER 2011).

1.6 ENZIMAS PROMÍSCUAS, EVOLUTIBILIDADE E ATIVIDADE DETOXIFICADORA

Enzimas são proteínas que apresentam atividade catalisadora e especificidade aos substratos. As enzimas podem ser de alta seletividade, ou seja, muito específicas ao substrato (FISHER, 1894), ou podem ser cataliticamente promíscuas, que é o caso de enzimas que catalisam substratos diferentes daqueles para qual foram selecionadas. A compatibilidade entre uma enzima e seu substrato é determinada em grande parte pela sua capacidade de adquirir novas conformações. A capacidade de um sistema protéico apresentar evolução adaptativa é chamada evolutibilidade (do inglês *evolvability*) e essa característica demonstra ser fundamental para o entendimento de enzimas promíscuas. Em outros termos, a evolutibilidade de uma enzima é a sua habilidade de adotar rapidamente novas funções (TOKURIKI; TAWFIK, 2009). Este fenômeno pode ocorrer devido à diversificação de confôrmeros pré-existentes ou a confôrmeros completamente novos (TOKURIKI; TAWFIK, 2009).

A variabilidade conformacional muitas vezes é a base da evolutibilidade das enzimas, e pode ser entendida como um conjunto de subestruturas alternativas, ou confôrmeros, que estejam

em equilíbrio com o estado nativo da proteína (TOKURIKI; TAWFIK, 2009). Esta diversidade conformacional torna a enzima alternar subestados conformacionais e funções além das do seu estado nativo. Atualmente admite-se que a variação conformacional é uma propriedade inerente a qualquer cadeia polimérica e pode variar desde flutuações das cadeias laterais dos resíduos a movimentos de sítios ativos e mudanças na estrutura secundária ou mesmo no *fold* como um todo (TOKURIKI; TAWFIK, 2009). Neste contexto, observa-se uma correlação entre a variabilidade conformacional de uma enzima e sua promiscuidade funcional, sendo esta flexibilidade estrutural provavelmente o mais importante mecanismo para uma proteína tornar-se promíscua (NOBELI *et al*, 2009).

Diversificação funcional de famílias e superfamílias de proteínas tem sido debatida na literatura em função da homologia (TODD *et al*, 2001; NOBELI *et al*, 2005, DEVOS; VALENCIA, 2000; GERLT; BABBITT, 2000). Há um consenso no fato de que enzimas homólogas com identidade de sequência acima de 40%, em casos de variação funcional, frequentemente apresentam a mesma reação química (mecanismo de reação), porém atuam em substratos diferentes (GERLT *et al*, 2005). Por outro lado, proteínas homólogas com baixa identidade entre sequências (< 30%) tendem a divergir tanto em atividade química quanto ao substrato. Algumas enzimas são conhecidas por terem conservado apenas parte da reação química, e, portanto são conhecidas como superfamílias de proteínas “mecanicamente diversas” (GERLT *et al*, 2005; GLASNER *et al*, 2006). O padrão de conservação da função de uma proteína aparentemente está associado à subfamília que pertence, e a variação existente está no quanto uma determinada subestrutura conservada dentro de uma superfamília faz parte do *motif* funcional desta mesma subfamília (CHIANG *et al*, 2008). Em geral, a diversidade de funções dentro de superfamílias de proteínas mostra uma distribuição assimétrica, a qual algumas poucas famílias exibem muita diversidade funcional, enquanto a maioria não apresenta (NOBELI

et al, 2009). Atualmente é possível fazer apenas uma subestimativa de variação funcional de uma superfamília protéica, devido à falta de dados e limitações dos métodos empregados para determinação de funções (GLASNER *et al*, 2006). Consequentemente não há como saber se o fato de algumas famílias apresentarem mais promiscuidade que outras se deve a restrições bioquímicas e biofísicas ou simplesmente resulta de um acidente evolutivo.

De modo geral, enzimas com atividade de desintoxicação celular resultam de seleção evolutiva para promiscuidade funcional (NATH; ATKINS, 2008). É característico nas enzimas detoxificadoras tais como citocromo P450 (CYPs), uridil transferases (UGTs), P-glicoproteínas, glutathione s-transferases (GSTs) e outras a notável capacidade de metabolizar compostos com propriedades químicas bem distintas (EKROOS; SJOGREN, 2006; GRISWOLD *et al*, 2006; REDINBO 2004). Ademais, as enzimas de detoxificação estão entre as enzimas mais promiscuas que se conhecem (GRISWOLD *et al*, 2006) e teoricamente estão no patamar de promiscuidade funcional dentro das restrições de um determinado *fold* (NATH; ATKINS, 2008). Particularmente a superfamília das glutathione s-transferases (GSTs) exibem membros com diversos graus de promiscuidade e, portanto são ótimos modelos de evolução de promiscuidade e especificidade funcional (EKROOS; SJOGREN, 2006; GRISWOLD *et al*, 2006). Já foi observada uma relação entre promiscuidade funcional e heterogeneidade conformacional em membros de GSTs classe alfa, em humanos (NATH; ATKINS, 2008). No entanto, até o momento não há nenhuma comparação explícita entre flexibilidade estrutural e promiscuidade funcional em GSTs da classe epsilon.

1.7 RESISTÊNCIAS A INSETICIDAS QUÍMICOS, PRESSÃO SELETIVA E IMPLICAÇÕES PARA O CONTROLE DA MALÁRIA

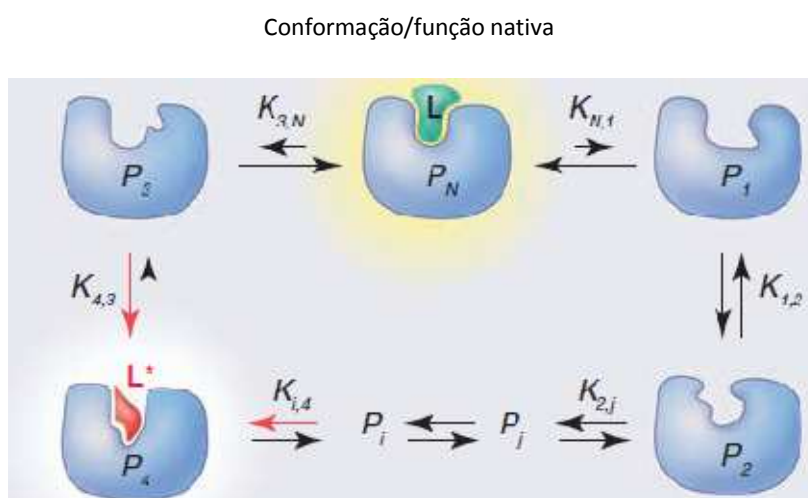
Uma das principais estratégias de controle da malária é o combate ao vetor com uso de inseticidas químicos, com frequente uso de redes e mosquiteiros impregnados com piretróides e a aplicação de DDT (Dicloro-Difenil-Tricloroetano) borrifado no interior dos domicílios. Aliado ao tratamento dos pacientes esta estratégia reduziu de modo significativo o número de casos de malária nos anos 80 (COHEN, 1982), porém foi registrada a ocorrência de resistência a estes e outros inseticidas ao longo dos anos (HEMINGWAY; RANSON, 2000; HEMINGWAY; BATES, 2003; HEMINGWAY, 2004). O DDT vem sendo usado no controle da malária por várias décadas, e apesar de ter sido relatado um elevado número de populações de espécies de *Anopheles* resistentes a este composto, já foi demonstrado que seu uso acarreta em diminuição significativa a mortalidade pela malária em alguns países africanos (WHO, 2010).

Portanto, a resistência a inseticidas químicos é uma preocupação constante para os órgãos de controle da malária, pois a utilização contínua e desordenada de inseticidas e repelentes promove, por seleção natural, a proliferação dos indivíduos mais resistentes de uma determinada população e pode diminuir as chances de sucesso dos programas de controle da doença. A seleção natural é um processo evolutivo em que os indivíduos mais aptos em uma população, ou em uma espécie, sobrevivem e se reproduzem, passando suas características aos seus descendentes, enquanto que indivíduos menos aptos têm menores chances de sobreviverem e se reproduzirem e, portanto são menos propensos a passarem suas características para as futuras gerações. Trata-se de um fenômeno que atua diretamente nas mutações ocorridas nos genes dos organismos vivos, que selecionam as mutações que originam características fenotípicas favoráveis e eliminando aquelas mutações que geram efeitos deletérios no organismo. As mutações que ocorrem nos genes podem por sua vez resultar em mudanças na sequência de aminoácidos (mutações não-silenciosas) ou manter os mesmos aminoácidos, uma vez que o código genético é redundante e degenerado (mutações silenciosas). De modo geral, mudanças no

terceiro par de bases de um códon não ocasionam alterações no peptídeo e, portanto não tem consequências nos traços fenotípicos. A seleção natural pode ser classificada como: a) seleção darwiniana ou positiva – ocorre quando a razão entre mutações não-silenciosas e mutações silenciosas é significativamente maior que 1,0 – indicando que a seleção está favorecendo a fixação das mutações que alteram a sequência primária da proteína; b) seleção purificadora ou negativa – quando esta razão é significativamente menor que 1,0 – que indica que a seleção promove a fixação de mutações que alteram a sequência primária da proteína; e c) neutra-quando a razão é próxima a 1,0 – o que significa que a seleção natural não favorece nem desfavorece a fixação das mutações que alteram a sequência primária da proteína (YANG; NIELSEN, 2000). No processo seletivo, considerando o mecanismo de resistência metabólica, os indivíduos com maior potencial para metabolizar inseticidas teriam, sob pressão seletiva exercida pelo uso intenso destes compostos, maiores chances de sobrevivência e reprodução e conseqüentemente a proporção de indivíduos resistentes seria aumentada ao longo das gerações.

Outra questão que hipoteticamente pode favorecer a resistência é a capacidade que algumas proteínas têm de apresentarem *evolutibilidade*, que é a capacidade que várias enzimas têm de realizarem rearranjos estruturais, alterar a especificidade de sua função e desenvolver a capacidade de se ligar a novos compostos (TOKURIKI; TAWFIK, 2009). Já foram apontadas várias evidências de que esta promiscuidade funcional das enzimas está diretamente relacionada com a flexibilidade de sua estrutura (TOKURIKI; TAWFIK, 2009). Deste ponto de vista, portanto, certas proteínas em seu estado nativo têm a capacidade de se enovelarem por diferentes caminhos a fim de originarem um conjunto de subestruturas alternativas capazes de se ligarem a diferentes ligantes (Figura 6).

Figura 6 - Esquema de dinamismo conformacional e promiscuidade de uma proteína. No modelo admite-se que a proteína existe como um conjunto de conformações denominado estado nativo (P_N , que interage com o ligante nativo L). Conformêmeros alternativos estão relacionados à variações estruturais que vão desde de rotâmeros da cadeia lateral e rearranjos de loops no sítio ativo até transições mais profundas. Conformêmeros menores (por exemplo, P_4) podem mediar funções alternativas, tais como a interação promíscua L^* (sendo L^* um ligante que a proteína não foi selecionada para se ligar).



Fonte: (TOKURUKI; TAWFIK, 2009).

1.8 A IMPORTÂNCIA DA BIOINFORMÁTICA ESTRUTURAL NA COMPREENSÃO DE ENZIMAS.

As enzimas são produtos da expressão de genes e suas atividades representam o fenótipo dos seres vivos. As enzimas de modo geral são proteínas complexas que desempenham atividade catalítica, ou seja, catalisam reações (GRISHAM; GARRETT, 1999). Compreender o modo de ação destas moléculas implica em conhecer detalhes de sua estrutura e de sua dinâmica molecular, uma vez que sua estrutura tridimensional e sua dinâmica determinarão o seu mecanismo de ação nos substratos (AGARWAL, 2006). Uma visão integrada da estrutura, dinâmica e função de proteínas está em grande evidência, onde as proteínas são consideradas

como máquinas dinamicamente ativas e que seus movimentos internos estão intimamente ligados a funções como catálise enzimática. Há grande interesse, tanto de grupos experimentais como computacionais, em investigar esta interconexão (AGARWAL, 2006).

Mudanças nos códons podem alterar os aminoácidos de uma enzima e resultar em diferenças na estrutura e dinâmica molecular, o que implica em mudanças no fenótipo. Segundo a teoria neutra da evolução molecular a maior parte das substituições peptídicas não deveria gerar consequências adaptativas. Contudo, certas substituições de aminoácidos podem acarretar em importantes alterações na estrutura da enzima e interferir no mecanismo de ação, com repercussões no estado adaptativo do organismo (GONZALES *et al*, 2012).

Diversos estudos relacionam a divergência sequencial e a evolução das proteínas com sua dinâmica. Mudanças na estrutura secundária das proteínas e previsões de desordem estrutural são linearmente proporcionais a mudanças nas sequências de aminoácidos (SCHAEFER *et al*, 2010). Em um estudo com um conjunto de 900 proteínas do PCDB (*Protein Conformational Database*), uma base de dados de confôrmeros, foi analisada pelo modelo SCPE (*Structurally Constrained Protein Evolutionary*) para explorar a influência de diferentes conformações na divergência sequencial das proteínas. A existência de um padrão de substituição de sítio confôrmero-específica indicou que a diversidade conformacional desempenha um papel central na modulação da evolução da proteína (JURITZ *et al*, 2012). Em um trabalho com proteínas de *Saccharomyces cerevisiae* foi constatada que a correlação linear entre aumento da superfície de acessibilidade ao solvente e taxa evolutiva (razão entre mutações não-sinônimas e mutações sinônimas) refletia uma pressão seletiva nos aminoácidos (RAMSEY *et al*, 2011).

Deste modo, os estudos computacionais das enzimas e investigações das alterações nas sequências primárias são apontados como algumas das técnicas mais promissoras para a compreensão das funções biológicas sob um ponto de vista evolutivo e adaptativo.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Analisar a dinâmica molecular, estrutura e função das enzimas AgGSTE2 e AgGSTE5 através da aplicação de metodologias computacionais.

2.2 OBJETIVOS ESPECÍFICOS

- Determinar o polimorfismo genético e a taxa evolutiva dos genes *AgGSTE2* e *AgGSTE5*.
- Construir e validar um modelo tridimensional da AgGSTE5, utilizando a estrutura cristalográfica da AgGSTE2 como molde.
- Caracterizar e comparar a dinâmica estrutural da AgGSTE5 e da AgGSTE2 através da técnica de dinâmica molecular.
- Comparar o modo de ligação das enzimas AgGSTE5 e AgGSTE2 com compostos químicos com atividade inseticida (DDT, carbaril, malation, CDNB e cipermetrina) através do *docking* molecular.

3 MATERIAL E MÉTODOS

3.1 POLIMORFISMO GENÉTICO NOS GENES *AgGSTE2* E *AgGSTE5*

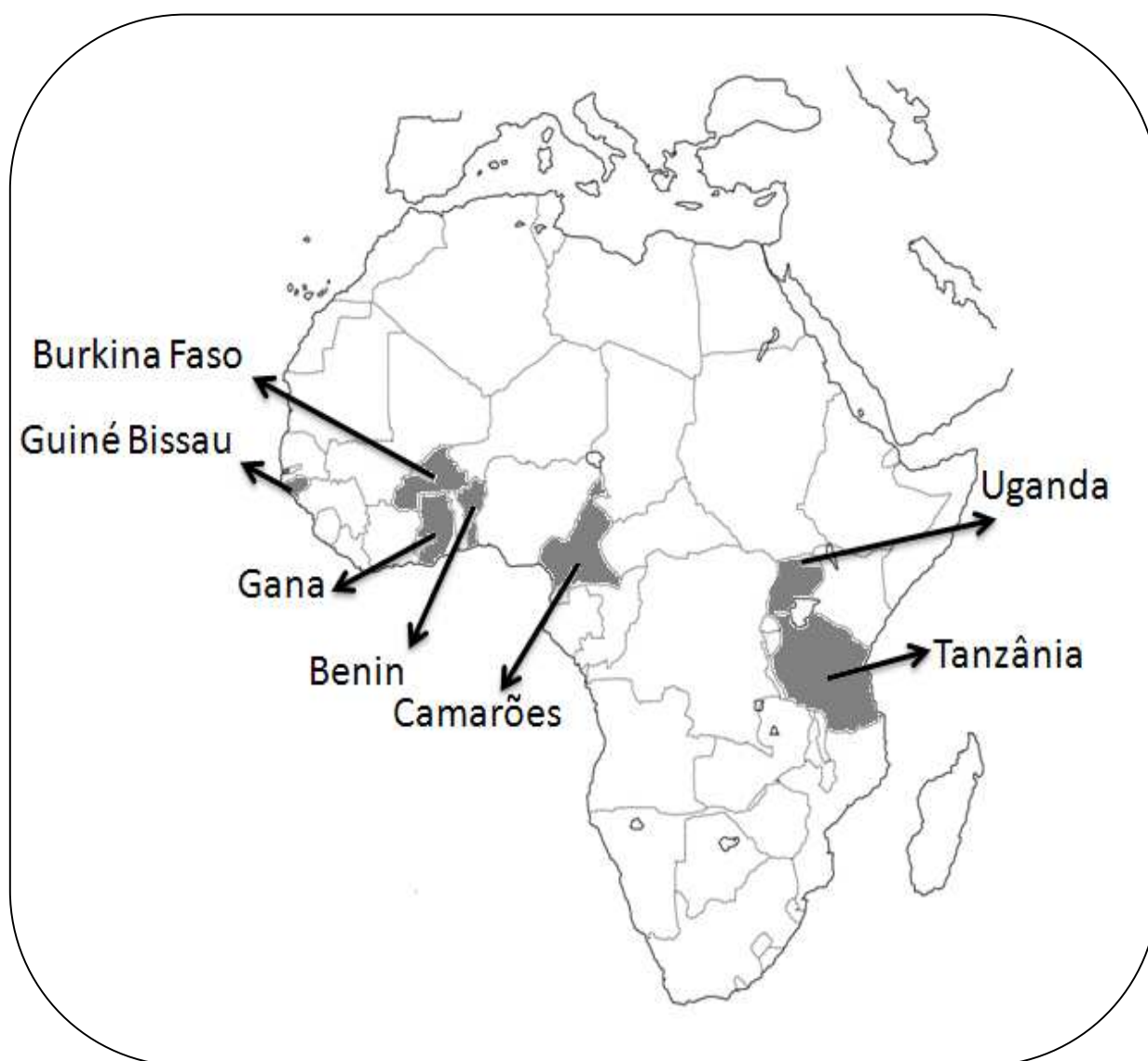
Com o objetivo de investigar a diversidade genética nos genes *AgGSTE2* e *AgGSTE5*, foram amplificados e sequenciados fragmentos destes genes em indivíduos de *An. gambiae* e *An. arabiensis*. Amostras de DNA foram gentilmente cedidas pelo Dr. Martin Donnelly da Liverpool School of Tropical Medicine. As amostras são procedentes de Uganda, Burkina Faso, Gana, Camarões, Benin, Tanzânia e Guiné-Bissau (Figura 7).

Foram utilizados os seguintes *primers* para a amplificação dos fragmentos via Reação em Cadeia da Polimerase (PCR): FW 5'- AGTTCGCTGCGAAAATGTCC -3' e REV 5'- CCAAATGCTTCCAAATTAACTC -3' para *AgGSTE2* e FW 5'- ATGGCAACGAACCCCATCATC -3' e REV 5'- TTACTTCGCAGCGACCGAC -3' para *AgGSTE5*. Uma mistura comercial para PCR (Promega) foi utilizada e seguiu-se o programa: 94° C por 2 min; 35 ciclos de 94° C por 5 min, 52° C (*AgGSTE2*) e 56° C (*AgGSTE5*) por 1 min, 72° C por 1 min; e uma extensão final de 72 °C por 5 min. Os fragmentos abrangiam os dois intróns e o exón dos genes e foram amplificados em ambas as direções. Os fragmentos amplificados foram purificados e em seguida sequenciados na Plataforma Tecnológica do Centro de Pesquisas Aggeu Magalhães – FIOCRUZ da rede PDTIS usando um sequenciador automático (Applied Biosystems). O alinhamento das sequências e a verificação da qualidade dos *contigs* foi feito com o auxílio do programa CodonCode Aligner (v 1. 2. 4 Codon Code Corporation).

Para as análises de diversidade genética foram estimados o número de haplótipos (h), a diversidade haplotípica, número de sítios segregantes (S), número de sítios segregantes por nucleotídeo (S/m), a diversidade nucleotídica (π) (Nei, 1987) e a taxa evolutiva por sítio (θ_w) (Watterson, 1975) e teste D de Tajima (D). Para saber se as sequências estão sob evolução neutra

foi realizado o teste de Tajima (Tajima, 1989). Os índices de polimorfismos foram estimados pelo programa *DNA Sequence Polymorphism* – DNA_{SP} (LIBRADO; ROZAS, 2009).

Figura 7 - Mapa da África mostrando os locais onde foram coletados mosquitos utilizados na análise do polimorfismo genético.



Fonte: Ilustração produzida pelo autor da pesquisa.

3.2 PROCEDIMENTO COMPUTACIONAL

3.2.1 Modelagem da AgGSTE5

O modelo tridimensional da proteína AgGSTE5 foi construído através da modelagem estrutural comparativa, ou modelagem por homologia estrutural. A modelagem por homologia é uma metodologia adequada à predição de estruturas tridimensionais de proteínas a partir da estrutura cristalográfica de uma proteína homóloga com identidade seqüencial >30 % e boa resolução. Esta técnica computacional baseia-se nas observações empíricas que proteínas compartilhando seqüências com uma alta identidade de aminoácidos exibem estruturas tridimensionais com o mesmo enovelamento e que a estrutura tridimensional de proteínas é mais conservada que a estrutura primária (SANTOS FILHO; ALENCASTRO, 2003). Sendo assim, uma proteína de estrutura conhecida pode ser usada como molde para a construção da estrutura tridimensional de uma proteína homóloga, a partir de sua seqüência de aminoácidos (SANTOS FILHO; ALENCASTRO, 2003). A seqüência de aminoácidos da AgGSTE5 foi previamente obtida do banco de dados *VectorBase* (<http://www.vectorbase.org/>; número de acesso: AGAP009192-PA) e comparada com a *AgGSTE2* no resultado de um *BLAST 2p sequence* (algoritmo de alinhamento local entre duas sequencias de aminoácidos), um algoritmo que executa um alinhamento entre seqüências primárias de duas proteínas. A estrutura cristalográfica da *AgGSTE2* (PDB ID: 2IMI, resolução 1.4 Å) foi utilizada como molde (número de acesso: 2IMI, *Protein Data Bank-PDB*) por apresentar uma alta similaridade de seqüência (76%) com a *AgGSTE5*. Através do servidor *Swiss-Model* (ARNOLD *et al*, 2006; KIEFER *et al*, 2009; PEITSCH, 1995) foi realizado o alinhamento de seqüências e a predição do modelo estrutural para a *AgGSTE5*. O tripeptídeo glutathiona, em sua forma reduzida, foi adicionado ao sítio ativo do modelo estrutural da *AgGSTE5*. Este modelo tridimensional foi inserido em uma caixa cúbica com moléculas de água e íons, de acordo com o modelo de água SPC (*Simple Point Charge*) para

simulações de líquidos (BERENDSEN *et al.*, 1981). Este sistema foi submetido às simulações de dinâmica molecular visando à avaliação da estabilidade e qualidade estrutural do mesmo.

3.2.2 Dinâmica Molecular

O método de dinâmica molecular (DM) descreve o movimento dos átomos de um sistema molecular e permite reproduzir propriedades médias de sistemas líquidos e sólidos (LEACH, 1991). A dinâmica molecular utiliza-se da mecânica clássica Newtoniana para calcular o movimento de um sistema de N átomos em função do tempo. A força é obtida derivando a energia potencial como a seguir:

$$\vec{F}_i = -\frac{dV_{total}}{d\vec{r}_i},$$

Onde $i = 1...N$ e i é o índice dos N átomos no sistema. Sabendo-se a força e a massa m_i é possível calcular a aceleração $\vec{a}$ usando a segunda lei de Newton apresentada na próxima equação,

$$\vec{F}_i = m_i \vec{a}_i,$$

onde $\vec{a}$ é definido segundo a equação:

$$\vec{a}_i = \frac{d\vec{v}_i}{dt},$$

Nesta derivada t é o tempo e v_i é a velocidade que é dada pela seguinte fórmula:

$$\vec{v}_i = \frac{d\vec{r}_i}{dt},$$

Onde r_i designa a posição do átomo i . Considerando a definição de velocidade e da aceleração, pode-se descrever a equação da força na seguinte forma:

$$\vec{F}_i = m_i \frac{\partial^2 \vec{r}_i}{\partial t^2}$$

Na dinâmica molecular clássica para solucionar as equações de movimento é necessária a utilização dos campos de forças, que consiste num conjunto de parâmetros que descrevem as forças que atuam nos átomos. Um campo de forças geralmente abrange a energia potencial das interações ligantes com as interações não ligantes (figura 8), que é descrita na seguinte equação:

$$E_{CF} = E\beta + E\theta + E\omega + E\varphi + E_{vdw} + E_{el} + E_{acl}$$

Onde,

$E\beta$ = Termo da energia potencial de ligação.

$E\theta$ = Termo da energia potencial angular.

$E\omega$ = Termo da energia potencial dos diedros impróprios.

$E\varphi$ = Termo da energia potencial torsional.

E_{vdw} = Termo da energia potencial de van der Waals.

E_{el} = Termo da energia potencial eletrostática.

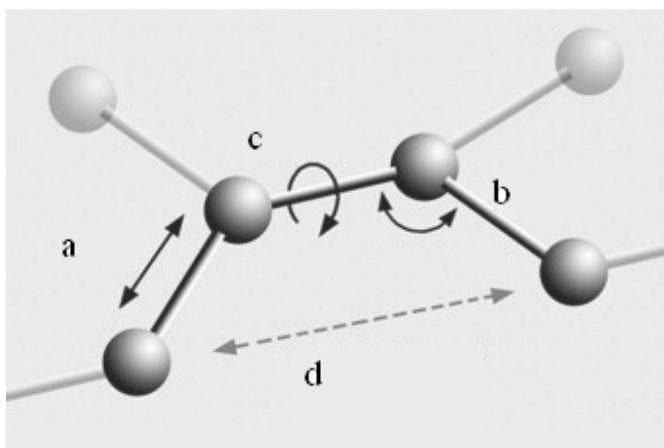
E_{acl} = Termo de acoplamento entre os potenciais dos átomos ligados.

Detalhando todos os termos, a equação pode ser descrita da seguinte maneira:

$$E_{CF} = \frac{1}{2} \sum_{n=1}^N k_{\beta n} (r_n - r_{on})^2 + \frac{1}{2} \sum_{n=1}^N k_{\theta n} (\theta_n - \theta_{on})^2 + \frac{1}{2} \sum_{n=1}^N k_{\omega n} (\omega_n - \omega_{on})^2 +$$

$$\sum_{n=1}^N k_{\varphi n} [1 + \cos (m_n \varphi_n - \delta_n)] + \sum_{i < j}^{N'} k \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \sum_{i < j}^{N'} \frac{q_i q_j}{4\pi\epsilon\epsilon_o r_{ij}}$$

Figura 8 - Representação dos termos dos campos de forças utilizados em dinâmica molecular clássica. Em (a): energia potencial da distância de ligação, (b): energia potencial das flexões angulares, (c): energia potencial das torsões, (d): interações não-ligantes.



Fonte: Ilustração produzida pelo autor da pesquisa

O modelo estrutural foi validado antes de ser submetido à simulação de dinâmica molecular para verificar sua qualidade estrutural e termodinâmica. A qualidade estereoquímica das estruturas foi avaliada através da análise da percentagem de aminoácidos localizados em regiões energeticamente favoráveis do diagrama de Ramachandran. Os gráficos de Ramachandran foram elaborados com o uso do programa PROCHECK (LASKOWSKI *et al*, 1993; LASKOWSKI *et al* 1996). A estabilidade energética do modelo foi estimada através de dois potenciais de energia ANOLEA (MELO, 1998) e GROMOS (SCOTT *et al*, 1999). A estrutura do modelo foi depositada na base de dados pública *Protein Model Data Base (PMDb)*

disponível online através do sítio <http://mi.caspar.it/PMDB/> sob o código de identidade PMDB ID: **PM0077788**.

As coordenadas atômicas da AgGSTe2 e AgGSTe5 foram obtidas dos seus respectivos arquivos PDB (Tabela 1), assim como o do seu ligante, o tripeptídeo glutathione ou GSH ($C_{10}H_{17}N_3O_6S$). A geometria do ligante também foi obtida da base de dados PDB. A partir destes arquivos foram gerados os arquivos das topologias e preparados os sistemas para a DM.

Também foi submetida às simulações uma isoforma da AgGSTe2 (AgGSTe2mut) contendo duas mutações, I114T e F120L (Isoleucina para Treonina no resíduo 114, Fenilalanina para Leucina no aminoácido 120). As três proteínas (AgGSTe2, AgGSTe2mut e AgGSTe5) foram simuladas com e sem o ligante GSH. Para a construção do mutante (AgGSTe2mut) foi utilizada as geometrias da proteína não mutante (AgGSTe2) e substituídos os resíduos no arquivo PDB, manualmente nas duas subunidades.

Os respectivos sistemas (AgGSTe2, AgGSTe2mut e AgGSTe5) foram centralizados dentro de uma caixa retangular com dimensões de $9,5 \times 8,5 \times 9,5 \text{ nm}^3$, e solvatados por meio da adição de moléculas de solvente explícito (modelo SPC). A carga total da proteína foi então neutralizada pela adição de contra-íons até atingir uma força iônica de 0.15 M. Os sistemas foram mantidos em condições isotérmicas (300 K) e isobáricas (1 atm) através dos métodos de *Nosé-Hoover thermostat* (NOSÉ; KLEIN, 1983; HOOVER, 1985) e *Parrinello-Rahman barostat* (PARRINELLO; RAHMAN, 1981).

Tabela 1 - Sistemas Submetidos À Simulação De Dinâmica Molecular.

Sistema	Acesso das Coordenadas	Duração
<i>Com ligante</i>		
AgGSTE2	PDB-2IMI	50 ns.
AgGSTE2mut	PDB-2IMI	“
AgGSTE5	Modelo teórico	“
<i>Sem ligante</i>		
AgGSTE2	PDB-2IL3	“
AgGSTE2mut	PDB-2IL3	“
AgGSTE5	Modelo teórico	“
Dados da pesquisa		

As restrições de distâncias e ângulos entre os átomos das proteínas e ligantes foram estabelecidas utilizando-se o algoritmo P-LINCS (HESS *et al*, 1997), e as restrições da geometria das moléculas de água foram feitas pelo algoritmo SETTLE (MIYAMOTO,1992). O campo de forças GROMOS (VAN GUNSTEREN *et al*, 1996) com o conjunto de parâmetros G53A6 (OOSTENBRINK *et al*, 2004) e o modelo SPC de água (BERENDSEN *et al*, 1987) foram escolhidos para descrever as interações inter-atômicas do sistema. O programa GROMACS versão 4.5.3 foi utilizado para realizar as simulações e para as análises de dados. Antes das simulações, os sistemas foram submetidos à etapa de minimização de energia, que é uma técnica

que visa encontrar um conjunto de coordenadas que minimizem a energia potencial do sistema e evitem o mau contato entre os átomos. Para este propósito utilizou-se o algoritmo *steepest descent*, que utiliza a derivada primeira para determinação da direção para o mínimo (ARFKEN, 1985; MORSE; FESHBACH, 1953). Após a equilibração, os sistemas foram simulados por 50 ns e as trajetórias resultantes foram armazenadas a cada 0.5 ps.

O conjunto de parâmetros 53A6 do campo de forças GROMOS foi escolhido para este estudo onde os átomos de hidrogênio apolares são tratados implicitamente, reduzindo assim o número de interações a serem calculadas em sistemas compostos por um elevado número de partículas.

3.3 ANÁLISES DOS DADOS

As mudanças conformacionais das proteínas estudadas com relação à configuração inicial do sistema e ao longo do tempo ao longo foram quantificadas através do cálculo dos desvios quadráticos médios ou RMSD (*Root-mean square deviation*) dado pela expressão:

$$RMSD = \left[\frac{1}{N} \sum_{i=1}^N [r_i(t) - r_i(0)]^2 \right]^{1/2}$$

Onde: Onde $r_i(t)$ e $r_i(0)$ representam as coordenadas do i -ésimo átomo no tempo t e 0 , respectivamente, e N é o número de átomos no domínio de interesse (C α).

Para se obter as oscilações de estrutura média das proteínas foi computada as flutuações quadráticas médias atômicas, ou RMSF (*Root-mean square fluctuation*) descrito pela equação:

$$RMSF = \left[\frac{1}{\Im} \sum_{j=0}^{\Im} [r_i(j) - \bar{r}_i]^2 \right]^{1/2}$$

Onde $r_i(j)$ são as coordenadas do i -ésimo átomo no passo de tempo j ; r_i suas posições médias; e $\Im$ é o tempo de simulação, expresso como o número total de passos de tempo (*time steps*) coletados.

O raio de giro (RG) foi calculado com a finalidade de se observar processos de expansão ou compactação das enzimas durante a simulação. É usado para descrever as dimensões de uma cadeia polimérica e é dado pela equação:

$$R_g^2 \stackrel{def}{=} \frac{1}{N} \sum_{k=1}^N (r_k - r_{mean})^2$$

Onde r_{mean} é a posição média dos monômeros. O raio de giro dá uma idéia de quão compacta é a estrutura de uma proteína.

A área da superfície de acessibilidade ao solvente (SASA) foi computada para obter uma idéia da hidrofobicidade das enzimas e se houve aumento da exposição ao solvente. A SASA total de uma molécula de N átomos é dada pela soma:

$$SASA = \sum_{i=1}^{N_{atoms}} A_i$$

Onde A_i é a SASA do i ésimo átomo, que por sua vez é obtida pela fórmula:

$$A_i(r^N) = S \prod_{i=1}^{N_{atoms}} 1 - \frac{p_i p_{ij} b_{ij}(r_{ij})}{S_i}$$

Sendo $S_i = 4\pi (R_i + R_{solv})^2$ a SASA do átomo isolado i de raio R_i e um solvente de raio R_{solv} . O termo $b_{ij}(r_{ij})$ representa o SASA removido S_i pela sobreposição dos átomos i e j , a uma distância $r_{ij} = |r_i - r_j|$. O parâmetro empírico p_i depende do tipo de átomo, enquanto o parâmetro empírico p_{ij} serve como um fator de adicional discriminação que distingue entre o primeiro e o próximo átomo vizinho ligado covalentemente e os átomos ligados de forma não covalente.

Foi calculado o teor da estrutura secundária para as proteínas pelo algoritmo DSSP (*Dictionary of Secondary Structure of Proteins*) ao longo das simulações com o propósito de identificar regiões onde há formação ou desnaturação da estrutura secundária. O algoritmo DSSP classifica os aminoácidos de acordo com a estrutura secundária que o mesmo faz parte, baseando-se nas ligações de hidrogênio dentro das cadeias principais. As ligações podem ser identificadas pela seguinte equação quando E é menor que -0.5 kcal/mol:

$$E = 0.084 \left\{ \frac{1}{r_{ON}} + \frac{1}{r_{CH}} + \frac{1}{r_{OH}} + \frac{1}{r_{CN}} \right\} . 332 \text{ kcal / mol}$$

A análise dos componentes principais (PCA) foi realizada para identificar as regiões que mais contribuem para os movimentos mais significativos para as mudanças conformacionais nas proteínas. A idéia básica do PCA é encontrar os quadrados mínimos para o vetor na fórmula:

$$\vec{q}(t) \approx \sum_{\alpha=1}^M a_{\alpha}(t) \vec{\eta}_{\alpha}$$

onde M é o número de graus de liberdade os vetores $\vec{\eta}_{\alpha}$ são vetores de base ortonormal, os quais são definidos pelo autovalor

$$\lambda_{\alpha} \vec{\eta}_{\alpha} = K_T \vec{\eta}_{\alpha}$$

Nesta equação K_T é a matriz de correlação determinada pela trajetória da dinâmica molecular $\vec{q}(t)$:

$$K_T = \left[\left(\vec{q}(t) - [\vec{q}(t)]_{t_{tot}} \right) \left(\vec{q}(t) - [\vec{q}(t)]_{t_{tot}} \right)' \right]_{t_{tot}},$$

onde $[\]_{t_{tot}}$ significa o tempo médio entre $0 \leq t \leq t_{tot}$. O autovalor λ_{α} é o tempo médio de $|a_{\alpha}(t)|^2$ (BALSERA *et al*, 1996).

As diferentes conformações amostradas durante a dinâmica molecular das proteína em consideração foram comparadas e agrupadas de acordo com um índice de similaridade conformacional. Esta análise permite avaliar e comparar a heterogeneidade conformacional das proteínas estudadas para determinar quais sistemas apresentam maior variação conformacional. Deste modo, esta análise classifica as estruturas de uma trajetória em grupos de estruturas similares de acordo com os valores de RMSD. O valor de corte para a classificação entre pares de

estruturas de um mesmo cluster foi de 0,1 nm entre os átomos da cadeia principal. Define-se a diferença D_{a+b} entre duas conformações, a e b , como:

$$D_{a+b} = \left(\frac{2}{N(N-1)} \sum_{i < j}^N (d_{ij}^a - d_{ij}^b)^2 \right)^{1/2},$$

onde a soma abrange todos os pares ij de N átomos, sendo consideradas as configurações a e b ; e d_{ij} é a distância entre os átomos i e j (TORDA; VAN GUNSTEREN, 1994).

Os cálculos dos desvios quadráticos médios (RMSD), os desvios quadráticos médios atômicos (RMSF), o teor da estrutura secundária (DSSP), o raio de giro (RG), a superfície de acessibilidade ao solvente (SASA), a análise dos componentes principais (PCA) e a análise por agrupamento (clustering) foram realizados com o pacote de programas GROMACS versão 4.5.3 (HESS *et al*, 1997).

3.4 DOCKING MOLECULAR

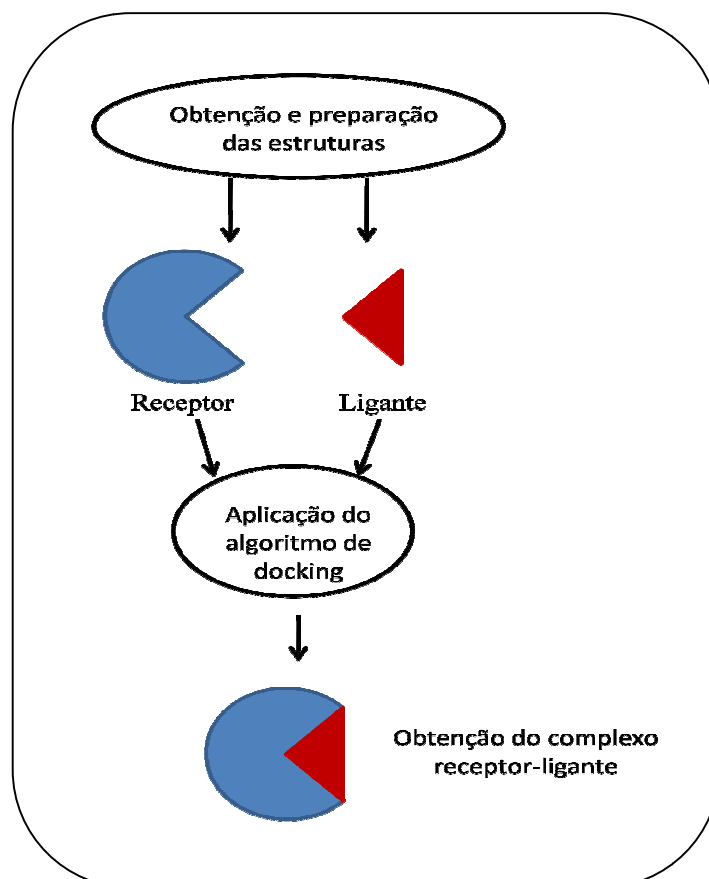
Os receptores usados nas análises de *docking* foram a estrutura cristalográfica da AgGSTE2 e seu mutante (AgGSTE2mut) e a estrutura do modelo construído para a AgGSTE5. Os ligantes utilizados foram os inseticidas DDT, carbaril, cipermetrina e o malation sendo todos estes inseticidas orgânicos sintéticos e de uso comercial, normalmente empregados no controle de culicídeos vetores (tabela 2); e o composto orgânico CDNB, molécula modelo para estudos de atividade de GSTs. As coordenadas atômicas dos compostos foram obtidas na base de dados ZINC (<http://zinc.docking.org/>).

Tabela 2 - Compostos utilizados como ligantes para o cálculo de *docking*.

Sigla/Nome comercial	Nome	Fórmula	Número de acesso
DDT	Dicloro-Difenil-Tricloroetano	C ₁₄ H ₉ Cl ₅	ZINC01530011
CDNB	2,4-Dinitroclorobenzeno	C ₆ H ₃ ClN ₂ O ₄	ZINC1540301
Carbaril	metilcarbamato de 1-naftilo	C ₁₂ H ₁₁ NO ₂	ZINC00001090
Cipermetrina	Cipermetrina	C ₂₂ H ₁₉ Cl ₂ NO ₃	ZINC71789490
Malation	Malation	C ₁₀ H ₁₉ O ₆ PS ₂	ZINC1530800
Dados da pesquisa			

O *docking* molecular é uma técnica computacional que visa calcular interações atômicas entre uma pequena molécula ligante e uma macromolécula, em busca da conformação de menor energia (Figura 9). Para o cálculo foi utilizado o programa AutoDock 4.2.2 (Morris *et al.* 2009), convertendo os arquivos em formato PDB para o formato pdbqt, que é o arquivo utilizado pelo AutoDock. Os ligantes foram assinalados com parâmetros de carga Gasteiger e apenas os hidrogênios apolares representados explicitamente. As simulações foram executadas com o algoritmo Genético Lamarckiano (LGA). A caixa foi estabelecida nas dimensões 52 X 52 X 52 Å centrado no ligante e no sítio ativo e o LGA foi submetido a cálculos de 10.000 réplicas com populações de 150 indivíduos a um máximo de 27.000 gerações e taxas de mutação *crossover* de 0,02 e 0,08 respectivamente.

Figura 9 - Esquema de *docking* entre proteína-ligante.



Fonte: Ilustração produzida pelo autor da pesquisa

4 RESULTADOS

4.1 POLIMORFISMO GENÉTICO DOS GENES *AgGSTE2* E *AgGSTE5*

Foram analisadas 62 sequências de 516 pb para o gene *AgGSTE2* e 69 sequências de 518 pb para o gene da *AgGSTE5*. Para *AgGSTE2* foram observados 55 haplótipos distintos, sendo 20 exclusivos. No total 117 sítios mostraram-se polimórficos e 401 não apresentaram polimorfismo. Em relação a *AgGSTE5*, 71 haplótipos foram encontrados, dos quais 26 foram exclusivos. A diversidade haplotípica foi similar entre os dois genes ($H_d=0,9713$ para *AgGSTE2* e $H_d=0,9861$ para *AgGSTE5*). Na tabela 3 encontram-se os valores de diversidade nucleotídica para *AgGSTE2* e *AgGSTE5*. Também foram estimados os testes D de Tajima e os limites de confiança para ambos os genes através do método *sliding window* (Tabelas 4 e 5). Neste método as estatísticas são calculadas por pequenos quadros ou janelas de dados. as estatísticas apresentadas são calculados para os dados de polimorfismo contidos dentro da janela. Desta maneira, a variação de cada estatística ao longo da região pesquisada pode ser medida. Este tipo de análise permite investigar como os padrões de variação mudança através de um segmento genômico pesquisados, é normalmente aplicado à detecção de locais de assinaturas seleção natural ou para a identificação de alterações locais na mutação ou taxas de recombinação. No gene *AgGSTE2* observa-se valores negativos e similares para todas as regiões, indicando seleção purificadora (negativa). Por outro lado, a região correspondente aos aminoácidos 54-87 (nucleotídeos 76-175) no gene *AgGSTE5* apresentou um valor positivo e significativo ($P < 0.01$). Levando-se em consideração a sequência gênica completa, o teste D mostra um valor entre 0 e -1, sendo este efeito condizente com a seleção positiva, tipo de seleção que leva a uma maior diversidade funcional da proteína.

Tabela 3 - Diversidade genética e teste de neutralidade de Tajima em sequências de nucleotídeo de *AgGSTE2* e *AgGSTE5*. m = número de sequências, S = número de sítios de segregantes, p_s = número de sítios segregantes por sítios de nucleotídeo (S/m), $\theta = (p_s/a_1)$, π = diversidade nucleotídica, e D é o teste de Tajima.

Gene	m	p_s	θ	π	D
<i>AgGSTE2</i>	62	0.177551	0.037807	0.015975	-1.990529
<i>AgGSTE5</i>	69	0.118577	0.024683	0.020479	0.570257

Dados da pesquisa

Tabela 4 -. Teste D de Tajima para sequências de nucleotídeo no gene *AgGSTE2*. #: $P < 0.10$; *: $P < 0.05$; **: $P < 0.01$.

Blocos de nucleotídeos	Ponto intermediário	D	Significância
1-100	50	-1,8820	*
26-125	75	-2,0446	*
51-150	100	-1,9335	*
76-175	125	-1,7445	#
101-200	150	-1,7657	#
126-225	175	-1,8341	*
151-250	200	-1,8113	*
176-275	225	-1,6777	#
201-300	250	-1,4580	
226-325	275	-1,0260	
251-350	300	-1,1147	
276-375	325	-0,7784	
301-400	350	-1,0197	
326-425	375	-1,2919	
351-450	400	-1,5354	
376-475	425	-1,9556	*
401-500	450	-1,5231	
426-516	471	-1,5252	

Tabela 5 - Teste D de Tajima para sequências de nucleotídeo no gene *AgGSTE5*. #: $P < 0.10$; *: $P < 0.05$; **: $P < 0.01$.

<i>Blocos de nucleotídeos</i>	<i>Ponto intermediário</i>	<i>D</i>	<i>Significância</i>
1-100	50	-0,8619	
26-125	75	0,0076	
51-150	100	0,5061	
76-175	125	2,9518	**
101-200	150	1,6795	
126-225	175	1,6787	
151-250	200	1,1376	
176-275	225	0,7546	
201-300	250	1,2107	
226-325	275	0,1579	
251-350	300	-0,1855	
276-375	325	-0,2588	
301-400	350	-0,3451	
326-425	375	-0,2916	
351-450	400	-0,2784	
376-475	425	-1,2272	
401-500	450	-0,1978	
426-516	472	0,0732	
Dados da pesquisa			

4.2 ALINHAMENTO E MODELAGEM ESTRUTURAL DA *AgGSTE5*

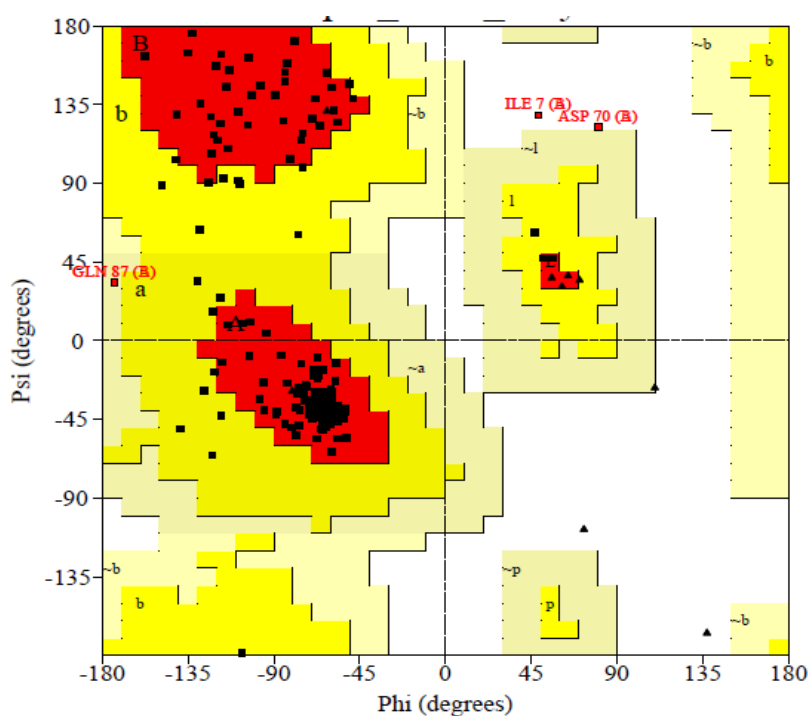
O alinhamento de seqüências para *AgGSTE5* e *AgGSTE2* (PDB-2IMI) via o algoritmo BLAST2P revelou uma similaridade de 76% e identidade de 52% entre as duas seqüências. A homologia entre as duas sequencias possibilitou o uso da estrutura cristalográfica em alta resolução (1.4 Å) da *AgGSTE2* (PDB ID 2IMI) como molde para a construção do modelo

4.2.2 Validação do modelo

O modelo obtido para a AgGSTe5 demonstrou através do diagrama de Ramachandran (Figura 11) que a maioria dos resíduos (88,7%) situavam-se nas regiões de conformações energeticamente mais favoráveis, e apenas 1% dos resíduos encontravam-se em regiões proibidas. Resultados semelhantes foram observados na estrutura cristalográfica da AgGSTe2, com 92% dos resíduos locados em regiões mais favoráveis (Figura 12).

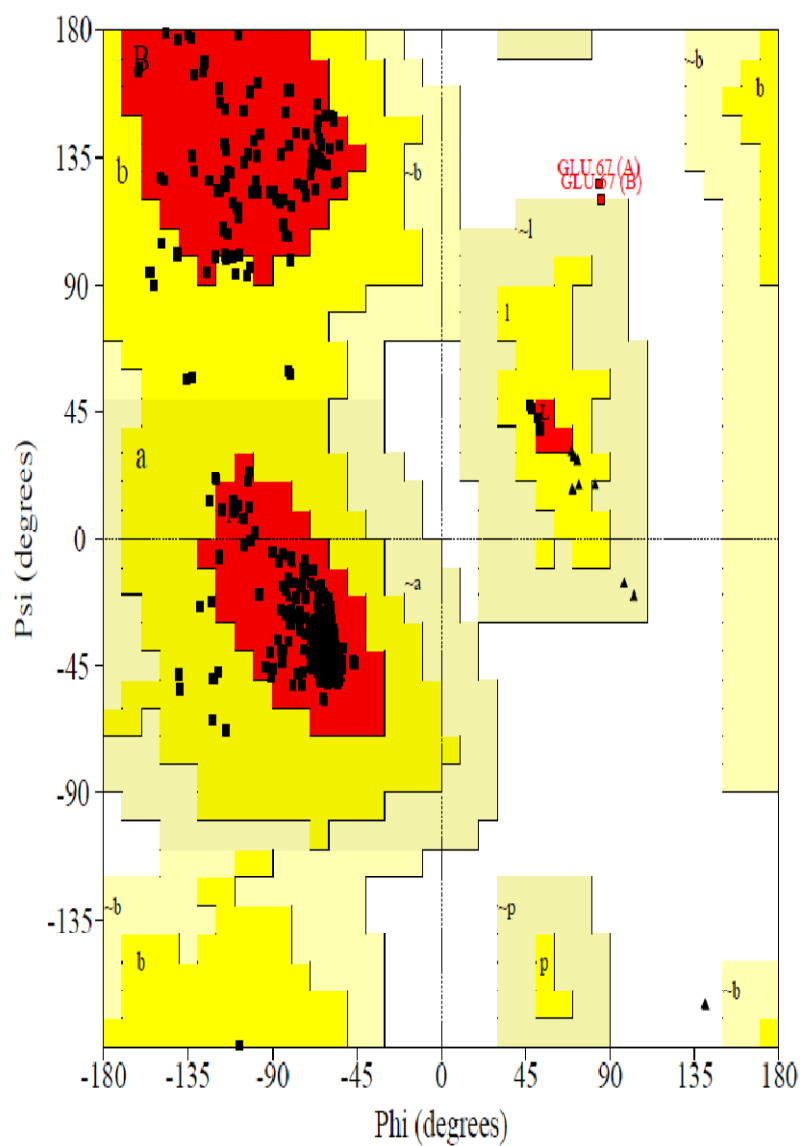
A qualidade local foi constatada também pelos resultados do *PSQS-Protein Structure Quality Score* (Tabela A1) e do ANOLEA, GROMOS e QMEAN (Figuras 13 e 14).

Figura 11 - Diagrama de Ramachandran da AgGSTe5 com 88,7% dos resíduos em regiões de conformação mais favorável (vermelho), 9,7 % em regiões permitidas, 0,5% em regiões desfavoráveis (bege) e 1,0% em regiões proibidas (branco).



Fonte: Ilustração produzida pelo autor da pesquisa

Figura 12 - Diagrama de Ramachandran da AgGSTE2 e AgGSTE2mut com 92% dos resíduos em regiões de conformação mais favorável (vermelho), 7,5 % em regiões permitidas, 0,3% em regiões desfavoráveis (bege) e 0,3% em regiões proibidas (branco).



Fonte: Ilustração produzida pelo autor da pesquisa

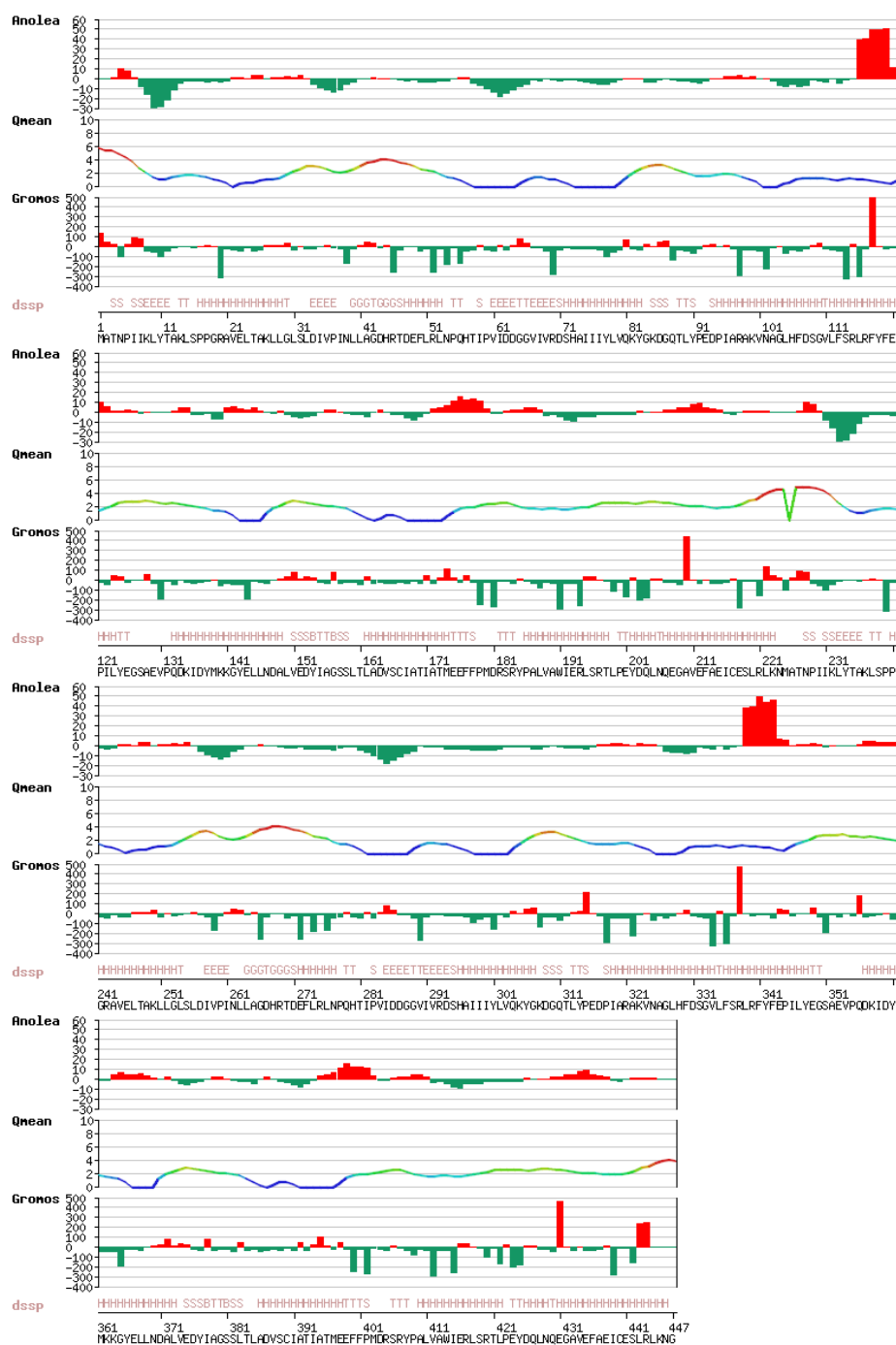
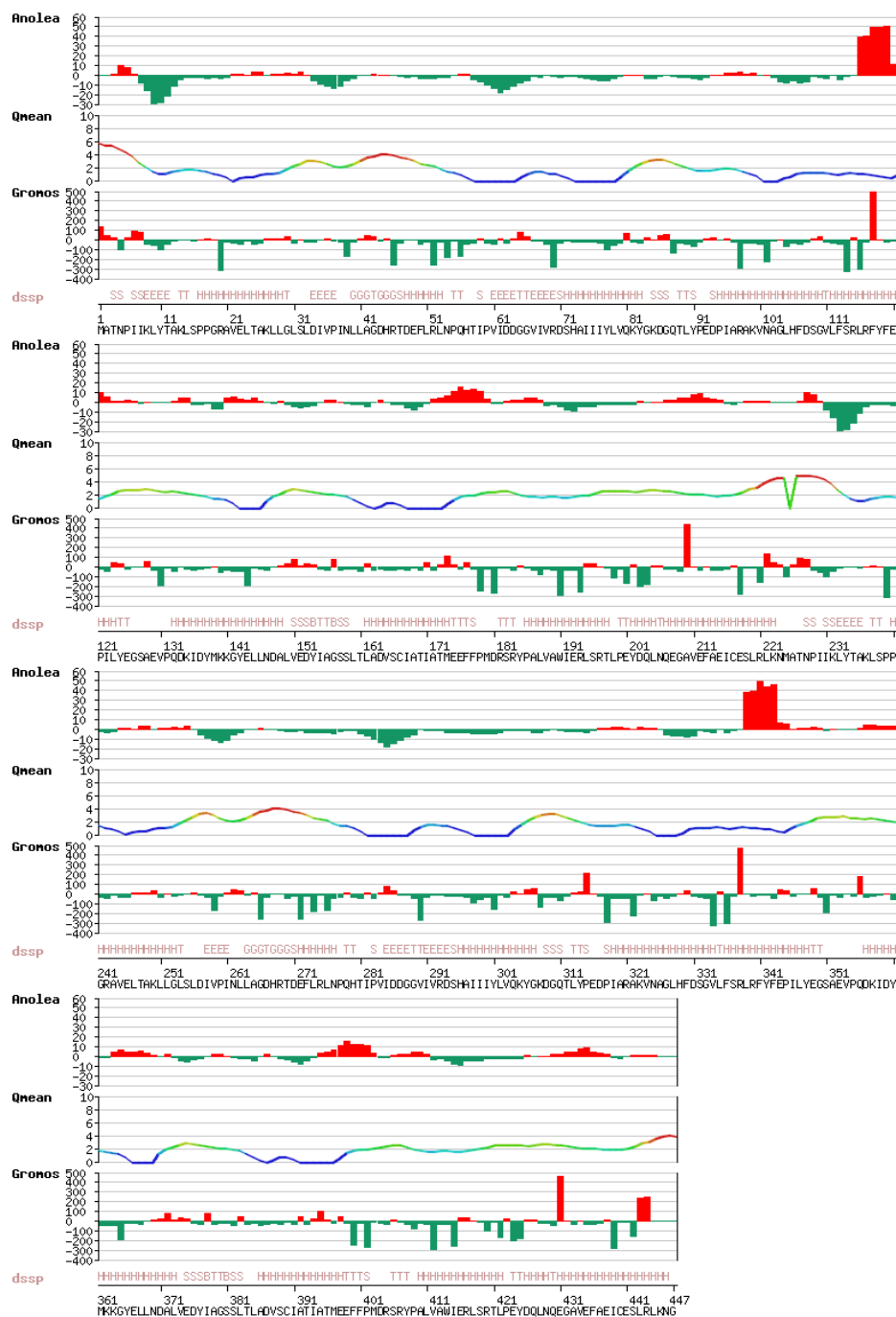
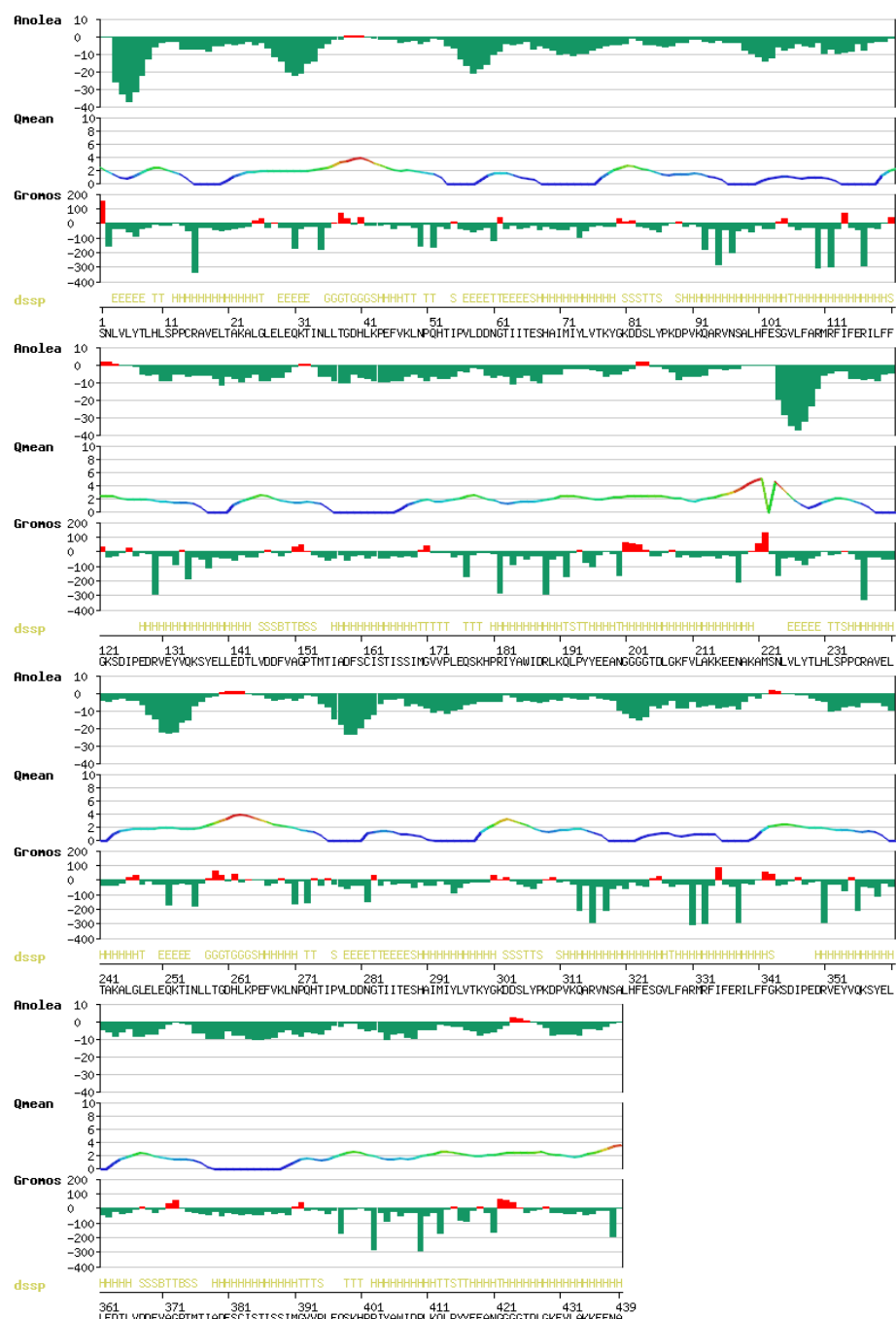


Figura 14 - Gráficos do Anolea, Qmean e Gromos em função do teor da estrutura secundária (DSSP) para AgGSTE2mut. Para o Gromos e o Anolea em verde destacam-se os resíduos em conformação favorável, o vermelho sinaliza os resíduos em conformações desfavoráveis. Para o Qmean a curva segue um gradiente de cores de azul (baixa energia) a vermelho (energia alta).



Fonte: Ilustração produzida pelo autor da pesquisa

Figura 15 - Gráficos do Anolea, Qmean e Gromos em função do teor da estrutura secundária (DSSP) para AgGSTE5. Para o Gromos e o Anolea em verde destacam-se os resíduos em conformação favorável, em vermelho sinaliza os resíduos em conformações desfavoráveis. Para o Qmean a curva segue um gradiente de cores de azul (baixa energia) a vermelho (energia alta).



Fonte: Ilustração produzida pelo autor da pesquisa

O score global do QMEAN (BENKERT *et al.*, 2008) é uma combinação linear de seis termos estruturais (tabela 6) utilizando potenciais estatísticos. O QMEAN6 é um escore confiável para todo o modelo, que pode ser utilizado a fim de comparar e classificar qualitativamente os modelos experimentais e computacionais. A estimativa de qualidade varia entre 0 e 1, com valores mais elevados para os melhores modelos. Sendo assim, os escores obtidos pelo QMEAN6 suportam a hipótese de que o modelo construído é confiável.

Tabela 6 - Resultados do *Qmean* para a AgGSTE2 e AgGSTE5.

Score	AgGSTE2	AgGSTE5
Energia de interação C_beta	-150.12	-147.20
Energia de todos os pares de átomos	-13486.92	-11780.50
Energia de solvatação	-50.75	-37.62
Energia das torções angulares	-109.05	-114.56
Conformação de Estrutura secundária	90.7%	90.1%
Acessibilidade ao solvente	83.2%	81.8%
QMEAN6 score	0.816	0.805
Dados da pesquisa		

Para fins comparativos, os mesmos protocolos utilizados para avaliar a qualidade do modelo estrutural da AgGSTE5 foram aplicados à estrutura cristalográfica de alta-resolução da AgGSTE2 (PDB-2IMI). Resultados semelhantes foram obtidos para esta estrutura experimental onde 92% dos resíduos adotam conformações energeticamente favoráveis do diagrama de Ramachandran (Figura 12). Desta forma, o modelo estrutural construído para AgGSTE5 é energeticamente comparável à estrutura da AgGSTE2. Este modelo foi utilizado como configuração inicial para as simulações de dinâmica molecular em solvente explícito.

Os resultados obtidos nas análises da qualidade estrutural demonstraram que o modelo proposto neste trabalho para a AgGSTE5 é confiável e representativo. A variação mutante da AgGSTE2 (AgGSTE2mut) apresentou resultados idênticos aos da sua isoforma tipo selvagem,

sendo portanto os gráficos e números apresentados válidos como referências para as duas formas. A estrutura tridimensional do modelo aqui construído pode ser acessada no *PMDB – Protein Model Data Base*, **ID PM0077788**.

4.2.3 Volume dos sítios de ligação

As estruturas das enzimas AgGSTe2, AgGSTe2mut e AgGSTe5 foram submetidas ao servidor do *Q-SiteFinder* (LAURIE; JACKSON, 2005), que funciona através do uso de sondas de ligações hidrofóbicas para a proteína e procura *clusters* de sondas com a energia de ligação mais favorável. Esses *clusters* são ranqueados em ordem segundo a probabilidade de ser um sítio de ligação, de acordo com a soma total das energias de ligação para cada cluster. Nesta análise, as proteínas foram submetidas com o GSH que foi tratado como ligante. Foi calculado o volume ($\text{\AA}^3$) das proteínas e do primeiro *pocket* do ranking e os átomos e resíduos foram identificados (tabela 7-9). Em todas as proteínas (AgGSTe2, AgGSTe2mut e AgGSTe5) o primeiro *pocket* do ranking incluía o sítio G (onde liga-se a glutatona) e o sítio H (substrato específico). O volume de toda a proteína da AgGSTe5 foi um pouco menor ($42977 \text{\AA}^3$) que o da AgGSTe2 ($43638 \text{\AA}^3$). O volume dos sítios de ligação da AgGSTe5 ($1120 \text{\AA}^3$) foi maior que o da AgGSTe2 ($777 \text{\AA}^3$). A isoforma mutante, a AgGSTe2mut, apresentou um volume ligeiramente menor ($43491 \text{\AA}^3$) que a da sua isoenzima, não obstante ter apresentado volume dos sítios de ligação maior ($919 \text{\AA}^3$) que a isoforma selvagem.

Tabela 7 - Resíduos e átomos que constituem o sítio catalítico da AgGSTe2 de acordo com o Q-sitefinder. Cada coluna representa a posição do átomo, o tipo do átomo, o tipo de aminoácido, a subunidade a qual o átomo pertence e o número do aminoácido.

62	CG	LEU	A	9	64	CD2	LEU	A	9	75	N	LEU	A	11
76	CA	LEU	A	11	77	C	LEU	A	11	78	O	LEU	A	11
79	CB	LEU	A	11	80	CG	LEU	A	11	81	CD1	LEU	A	11
83	N	SER	A	12	84	CA	SER	A	12	85	C	SER	A	12
87	CB	SER	A	12	88	OG	SER	A	12	89	N	PRO	A	13
90	CA	PRO	A	13	91	C	PRO	A	13	92	O	PRO	A	13
93	CB	PRO	A	13	94	CG	PRO	A	13	95	CD	PRO	A	13
96	N	PRO	A	14	97	CA	PRO	A	14	98	C	PRO	A	14
99	O	PRO	A	14	100	CB	PRO	A	14	101	CG	PRO	A	14
102	CD	PRO	A	14	120	N	ALA	A	17	121	CA	ALA	A	17
124	CB	ALA	A	17	261	ND2	ASN	A	35	262	N	LEU	A	36
263	CA	LEU	A	36	264	C	LEU	A	36	266	CB	LEU	A	36
267	CG	LEU	A	36	268	CD1	LEU	A	36	269	CD2	LEU	A	36
270	N	LEU	A	37	271	CA	LEU	A	37	274	CB	LEU	A	37
275	CG	LEU	A	37	276	CD1	LEU	A	37	277	CD2	LEU	A	37
301	CB	HIS	A	41	302	CG	HIS	A	41	303	ND1	HIS	A	41
304	CD2	HIS	A	41	305	CE1	HIS	A	41	306	NE2	HIS	A	41
381	OD1	ASN	A	50	391	CA	GLN	A	52	392	C	GLN	A	52
393	O	GLN	A	52	394	CB	GLN	A	52	395	CG	GLN	A	52
396	CD	GLN	A	52	397	OE1	GLN	A	52	398	NE2	GLN	A	52
401	C	HIS	A	53	402	O	HIS	A	53	404	CG	HIS	A	53
405	ND1	HIS	A	53	407	CE1	HIS	A	53	408	NE2	HIS	A	53
409	N	THR	A	54	410	CA	THR	A	54	411	C	THR	A	54
412	O	THR	A	54	413	CB	THR	A	54	414	OG1	THR	A	54
415	CG2	THR	A	54	416	N	ILE	A	55	417	CA	ILE	A	55
418	C	ILE	A	55	419	O	ILE	A	55	420	CB	ILE	A	55
421	CG1	ILE	A	55	423	CD1	ILE	A	55	424	N	PRO	A	56
425	CA	PRO	A	56	426	C	PRO	A	56	427	O	PRO	A	56
429	CG	PRO	A	56	430	CD	PRO	A	56	504	N	GLU	A	67
505	CA	GLU	A	67	506	C	GLU	A	67	507	O	GLU	A	67
508	CB	GLU	A	67	509	CG	GLU	A	67	510	CD	GLU	A	67
511	OE1	GLU	A	67	512	OE2	GLU	A	67	513	N	SER	A	68
514	CA	SER	A	68	515	C	SER	A	68	517	CB	SER	A	68
518	OG	SER	A	68	519	N	HIS	A	69	524	CG	HIS	A	69
525	ND1	HIS	A	69	526	CD2	HIS	A	69	527	CE1	HIS	A	69
528	NE2	HIS	A	69	549	CE	MET	A	72	798	O	GLU	A	103
800	CG	GLU	A	103	801	CD	GLU	A	103	802	OE1	GLU	A	103
803	OE2	GLU	A	103	804	N	SER	A	104	805	CA	SER	A	104
806	C	SER	A	104	807	O	SER	A	104	808	CB	SER	A	104
809	OG	SER	A	104	833	CB	PHE	A	108	834	CG	PHE	A	108
835	CD1	PHE	A	108	836	CD2	PHE	A	108	837	CE1	PHE	A	108
838	CE2	PHE	A	108	839	CZ	PHE	A	108	862	SD	MET	A	111
863	CE	MET	A	111	864	N	ARG	A	112	865	CA	ARG	A	112
866	C	ARG	A	112	867	O	ARG	A	112	868	CB	ARG	A	112
869	CG	ARG	A	112	870	CD	ARG	A	112	871	NE	ARG	A	112
872	CZ	ARG	A	112	873	NH1	ARG	A	112	874	NH2	ARG	A	112
895	CA	PHE	A	115	896	C	PHE	A	115	897	O	PHE	A	115
898	CB	PHE	A	115	899	CG	PHE	A	115	901	CD2	PHE	A	115
903	CE2	PHE	A	115	905	N	GLU	A	116	906	CA	GLU	A	116
907	C	GLU	A	116	908	O	GLU	A	116	909	CB	GLU	A	116

Tabela 7 – Continuação

910	CG	GLU	A	116	911	CD	GLU	A	116	912	OE1	GLU	A	116
913	OE2	GLU	A	116	937	CB	LEU	A	119	938	CG	LEU	A	119
939	CD1	LEU	A	119	940	CD2	LEU	A	119	945	CB	PHE	A	120
946	CG	PHE	A	120	947	CD1	PHE	A	120	948	CD2	PHE	A	120
949	CE1	PHE	A	120	950	CE2	PHE	A	120	951	CZ	PHE	A	120
958	CD1	PHE	A	121	960	CE1	PHE	A	121	962	CZ	PHE	A	121
1272	N	SER	A	161	1273	CA	SER	A	161	1274	C	SER	A	161
1275	O	SER	A	161	1276	CB	SER	A	161	1277	OG	SER	A	161
1296	CB	SER	A	164	1297	OG	SER	A	164	1303	OG1	THR	A	165
1633	CG	LEU	A	207	1634	CD1	LEU	A	207	1635	CD2	LEU	A	207
1655	CD1	PHE	A	210	1657	CE1	PHE	A	210	1659	CZ	PHE	A	210
2509	O	SER	B	98	2511	OG	SER	B	98	2526	CA	HIS	B	101
2527	C	HIS	B	101	2528	O	HIS	B	101	2529	CB	HIS	B	101
2535	N	PHE	B	102	2536	CA	PHE	B	102	2537	C	PHE	B	102
2538	O	PHE	B	102	2539	CB	PHE	B	102	2540	CG	PHE	B	102
2541	CD1	PHE	B	102	2543	CE1	PHE	B	102	2561	N	GLY	B	105
2562	CA	GLY	B	105	2563	C	GLY	B	105	2564	O	GLY	B	105
2565	N	VAL	B	106	2566	CA	VAL	B	106	2569	CB	VAL	B	106
2570	CG1	VAL	B	106	2571	CG2	VAL	B	106	2868	CB	LEU	B	140
2869	CG	LEU	B	140	2870	CD1	LEU	B	140	2871	CD2	LEU	B	140

Dados da pesquisa**Tabela 8** - Resíduos e átomos que constituem o sítio catalítico da AgGSTe2mut de acordo com o *Q-sitefinder*. Cada coluna representa a posição do átomo, o tipo do átomo, o tipo de aminoácido, a subunidade a qual o átomo pertence e o número do aminoácido.

62	CG	LEU	A	9	64	CD2	LEU	A	9	75	N	LEU	A	11
76	CA	LEU	A	11	77	C	LEU	A	11	78	O	LEU	A	11
79	CB	LEU	A	11	80	CG	LEU	A	11	81	CD1	LEU	A	11
83	N	SER	A	12	84	CA	SER	A	12	85	C	SER	A	12
87	CB	SER	A	12	88	OG	SER	A	12	89	N	PRO	A	13
90	CA	PRO	A	13	91	C	PRO	A	13	92	O	PRO	A	13
93	CB	PRO	A	13	94	CG	PRO	A	13	95	CD	PRO	A	13
96	N	PRO	A	14	97	CA	PRO	A	14	98	C	PRO	A	14
99	O	PRO	A	14	100	CB	PRO	A	14	101	CG	PRO	A	14
102	CD	PRO	A	14	120	N	ALA	A	17	121	CA	ALA	A	17
124	CB	ALA	A	17	261	ND2	ASN	A	35	262	N	LEU	A	36
263	CA	LEU	A	36	264	C	LEU	A	36	265	O	LEU	A	36
266	CB	LEU	A	36	267	CG	LEU	A	36	268	CD1	LEU	A	36
269	CD2	LEU	A	36	270	N	LEU	A	37	271	CA	LEU	A	37
274	CB	LEU	A	37	275	CG	LEU	A	37	276	CD1	LEU	A	37
277	CD2	LEU	A	37	301	CB	HIS	A	41	302	CG	HIS	A	41
303	ND1	HIS	A	41	304	CD2	HIS	A	41	305	CE1	HIS	A	41
306	NE2	HIS	A	41	381	OD1	ASN	A	50	391	CA	GLN	A	52
392	C	GLN	A	52	393	O	GLN	A	52	394	CB	GLN	A	52
395	CG	GLN	A	52	396	CD	GLN	A	52	397	OE1	GLN	A	52
398	NE2	GLN	A	52	401	C	HIS	A	53	402	O	HIS	A	53
404	CG	HIS	A	53	405	ND1	HIS	A	53	407	CE1	HIS	A	53
408	NE2	HIS	A	53	409	N	THR	A	54	410	CA	THR	A	54
411	C	THR	A	54	412	O	THR	A	54	413	CB	THR	A	54
414	OG1	THR	A	54	415	CG2	THR	A	54	416	N	ILE	A	55

Tabela 7 – Continuação

417	CA	ILE	A	55	418	C	ILE	A	55	419	O	ILE	A	55
420	CB	ILE	A	55	421	CG1	ILE	A	55	423	CD1	ILE	A	55
424	N	PRO	A	56	425	CA	PRO	A	56	426	C	PRO	A	56
427	O	PRO	A	56	429	CG	PRO	A	56	430	CD	PRO	A	56
504	N	GLU	A	67	505	CA	GLU	A	67	506	C	GLU	A	67
507	O	GLU	A	67	508	CB	GLU	A	67	509	CG	GLU	A	67
510	CD	GLU	A	67	511	OE1	GLU	A	67	512	OE2	GLU	A	67
513	N	SER	A	68	514	CA	SER	A	68	515	C	SER	A	68
517	CB	SER	A	68	518	OG	SER	A	68	519	N	HIS	A	69
524	CG	HIS	A	69	525	ND1	HIS	A	69	526	CD2	HIS	A	69
527	CE1	HIS	A	69	528	NE2	HIS	A	69	549	CE	MET	A	72
798	O	GLU	A	103	800	CG	GLU	A	103	801	CD	GLU	A	103
802	OE1	GLU	A	103	803	OE2	GLU	A	103	804	N	SER	A	104
805	CA	SER	A	104	806	C	SER	A	104	807	O	SER	A	104
808	CB	SER	A	104	809	OG	SER	A	104	833	CB	PHE	A	108
834	CG	PHE	A	108	835	CD1	PHE	A	108	836	CD2	PHE	A	108
837	CE1	PHE	A	108	838	CE2	PHE	A	108	839	CZ	PHE	A	108
862	SD	MET	A	111	863	CE	MET	A	111	864	N	ARG	A	112
865	CA	ARG	A	112	866	C	ARG	A	112	867	O	ARG	A	112
868	CB	ARG	A	112	869	CG	ARG	A	112	870	CD	ARG	A	112
871	NE	ARG	A	112	872	CZ	ARG	A	112	873	NH1	ARG	A	112
874	NH2	ARG	A	112	895	CA	PHE	A	115	896	C	PHE	A	115
897	O	PHE	A	115	898	CB	PHE	A	115	899	CG	PHE	A	115
901	CD2	PHE	A	115	903	CE2	PHE	A	115	905	N	GLU	A	116
906	CA	GLU	A	116	907	C	GLU	A	116	908	O	GLU	A	116
909	CB	GLU	A	116	910	CG	GLU	A	116	911	CD	GLU	A	116
912	OE1	GLU	A	116	913	OE2	GLU	A	116	937	CB	LEU	A	119
938	CG	LEU	A	119	939	CD1	LEU	A	119	940	CD2	LEU	A	119
945	CB	LEU	A	120	946	CG	LEU	A	120	947	CD1	LEU	A	120
948	CD2	LEU	A	120	958	CD1	PHE	A	121	960	CE1	PHE	A	121
962	CZ	PHE	A	121	1272	N	SER	A	161	1273	CA	SER	A	161
1274	C	SER	A	161	1275	O	SER	A	161	1276	CB	SER	A	161
1277	OG	SER	A	161	1296	CB	SER	A	164	1297	OG	SER	A	164
1303	OG1	THR	A	165	1633	CG	LEU	A	207	1634	CD1	LEU	A	207
1635	CD2	LEU	A	207	1655	CD1	PHE	A	210	1657	CE1	PHE	A	210
1659	CZ	PHE	A	210	2509	O	SER	B	98	2511	OG	SER	B	98
2526	CA	HIS	B	101	2527	C	HIS	B	101	2528	O	HIS	B	101
2529	CB	HIS	B	101	2535	N	PHE	B	102	2536	CA	PHE	B	102
2537	C	PHE	B	102	2538	O	PHE	B	102	2539	CB	PHE	B	102
2540	CG	PHE	B	102	2541	CD1	PHE	B	102	2543	CE1	PHE	B	102
2561	N	GLY	B	105	2562	CA	GLY	B	105	2563	C	GLY	B	105
2564	O	GLY	B	105	2565	N	VAL	B	106	2566	CA	VAL	B	106
2569	CB	VAL	B	106	2570	CG1	VAL	B	106	2571	CG2	VAL	B	106
2868	CB	LEU	B	140	2869	CG	LEU	B	140	2870	CD1	LEU	B	140
2871	CD2	LEU	B	140										

Dados da pesquisa

Tabela 9 - Resíduos e átomos que constituem o sítio catalítico da AgGSTe5 de acordo com o *Q-sitefinder*. Cada coluna representa a posição do átomo, o tipo do átomo, o tipo de aminoácido, a subunidade a qual o átomo pertence e o número do aminoácido

89	CA	ALA	A	12	90	C	ALA	A	12	91	O	ALA	A	12
92	CB	ALA	A	12	93	N	LYS	A	13	102	N	LEU	A	14
103	CA	LEU	A	14	104	C	LEU	A	14	105	O	LEU	A	14
106	CB	LEU	A	14	107	CG	LEU	A	14	109	CD2	LEU	A	14
110	N	SER	A	15	111	CA	SER	A	15	112	C	SER	A	15
114	CB	SER	A	15	115	OG	SER	A	15	116	N	PRO	A	16
120	CB	PRO	A	16	121	CG	PRO	A	16	122	CD	PRO	A	16
127	CB	PRO	A	17	128	CG	PRO	A	17	129	CD	PRO	A	17
270	O	ILE	A	37	276	CA	ASN	A	38	277	C	ASN	A	38
278	O	ASN	A	38	279	CB	ASN	A	38	280	CG	ASN	A	38
281	OD1	ASN	A	38	282	ND2	ASN	A	38	283	N	LEU	A	39
284	CA	LEU	A	39	285	C	LEU	A	39	287	CB	LEU	A	39
288	CG	LEU	A	39	289	CD1	LEU	A	39	290	CD2	LEU	A	39
291	N	LEU	A	40	292	CA	LEU	A	40	295	CB	LEU	A	40
296	CG	LEU	A	40	297	CD1	LEU	A	40	298	CD2	LEU	A	40
320	CB	HIS	A	44	321	CG	HIS	A	44	322	ND1	HIS	A	44
323	CD2	HIS	A	44	324	CE1	HIS	A	44	325	NE2	HIS	A	44
416	C	GLN	A	55	417	O	GLN	A	55	425	C	HIS	A	56
426	O	HIS	A	56	428	CG	HIS	A	56	429	ND1	HIS	A	56
431	CE1	HIS	A	56	432	NE2	HIS	A	56	433	N	THR	A	57
434	CA	THR	A	57	435	C	THR	A	57	437	CB	THR	A	57
438	OG1	THR	A	57	439	CG2	THR	A	57	440	N	ILE	A	58
441	CA	ILE	A	58	442	C	ILE	A	58	443	O	ILE	A	58
444	CB	ILE	A	58	445	CG1	ILE	A	58	447	CD1	ILE	A	58
854	CA	PHE	A	112	856	O	PHE	A	112	858	CG	PHE	A	112
859	CD1	PHE	A	112	861	CE1	PHE	A	112	862	CE2	PHE	A	112
863	CZ	PHE	A	112	882	CA	LEU	A	115	883	C	LEU	A	115
884	O	LEU	A	115	885	CB	LEU	A	115	886	CG	LEU	A	115
888	CD2	LEU	A	115	889	N	ARG	A	116	890	CA	ARG	A	116
891	C	ARG	A	116	892	O	ARG	A	116	893	CB	ARG	A	116
894	CG	ARG	A	116	895	CD	ARG	A	116	896	NE	ARG	A	116
897	CZ	ARG	A	116	898	NH1	ARG	A	116	899	NH2	ARG	A	116
923	N	PHE	A	119	924	CA	PHE	A	119	925	C	PHE	A	119
926	O	PHE	A	119	927	CB	PHE	A	119	928	CG	PHE	A	119
930	CD2	PHE	A	119	932	CE2	PHE	A	119	934	N	GLU	A	120
935	CA	GLU	A	120	936	C	GLU	A	120	937	O	GLU	A	120
938	CB	GLU	A	120	939	CG	GLU	A	120	940	CD	GLU	A	120
941	OE1	GLU	A	120	942	OE2	GLU	A	120	962	CB	LEU	A	123
963	CG	LEU	A	123	964	CD1	LEU	A	123	970	CB	TYR	A	124
971	CG	TYR	A	124	972	CD1	TYR	A	124	973	CD2	TYR	A	124
974	CE1	TYR	A	124	975	CE2	TYR	A	124	976	CZ	TYR	A	124
977	OH	TYR	A	124	1313	OG1	THR	A	169	1314	CG2	THR	A	169
1332	CB	THR	A	172	1333	OG1	THR	A	172	1669	CG	PHE	A	212
1670	CD1	PHE	A	212	1671	CD2	PHE	A	212	1672	CE1	PHE	A	212
1673	CE2	PHE	A	212	1674	CZ	PHE	A	212	1696	CD1	ILE	A	215
880	NH2	ARG	B	114	1135	CG	LEU	B	144	1136	CD1	LEU	B	144
1137	CD2	LEU	B	144										

Dados da pesquisa

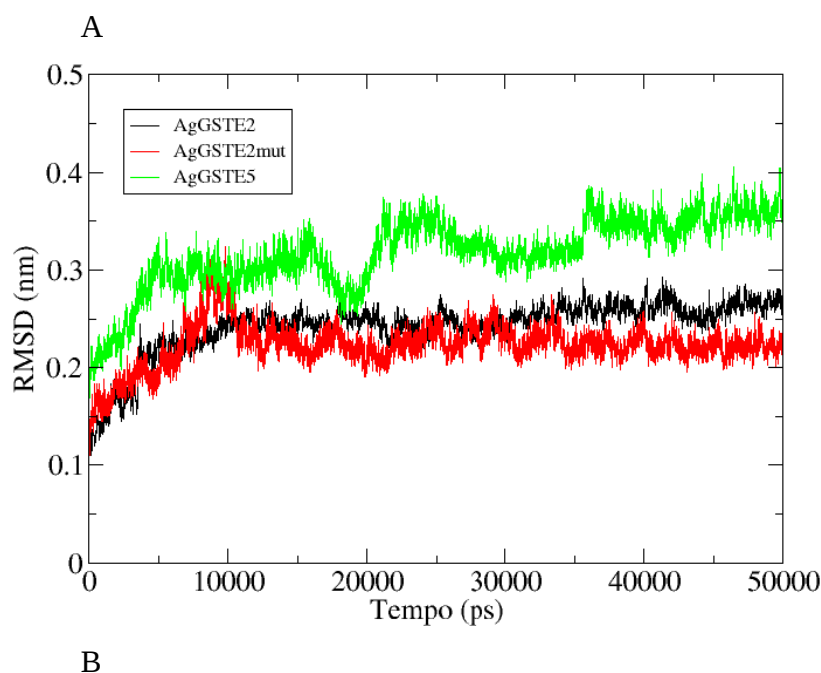
4.3 DINÂMICA MOLECULAR

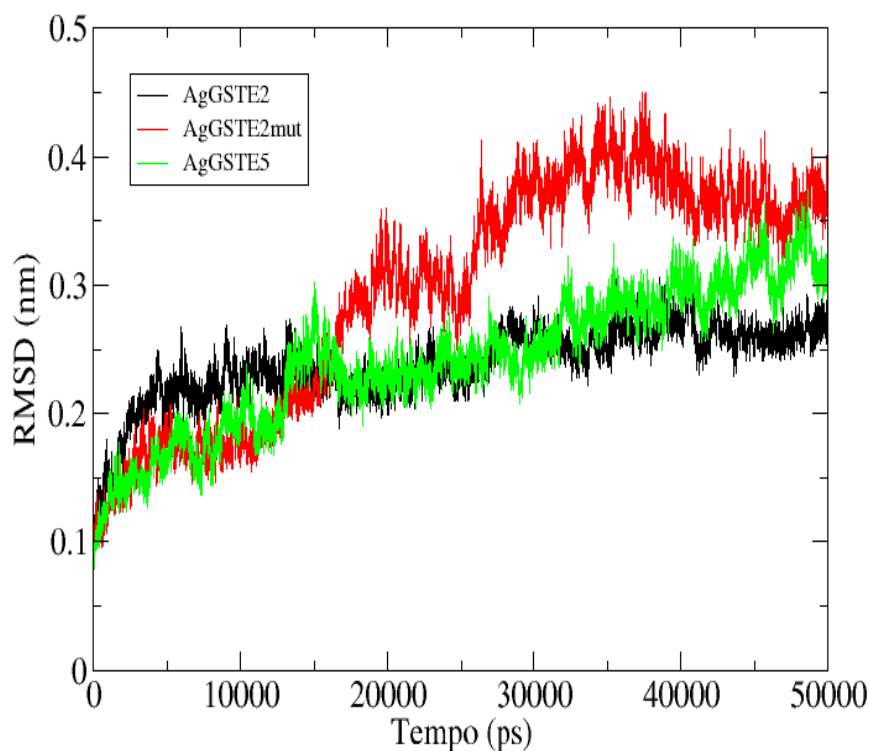
4.3.1 Estabilidade conformacional

Em simulações de dinâmica molecular de proteínas duas grandezas podem ser utilizadas na análise da estabilidade conformacional em função do tempo: os desvios quadráticos médios (RMSD) e o raio de giro (RG).

Os valores de RMSD para o carbono-alfa de cada sistema demonstram que há mudanças de conformação nos sistemas analisados. A AgGSTE5 aparenta ter mais estados conformacionais que suas homólogas, ou seja apresenta maior diversidade conformacional. Foram observadas poucas alterações no padrão das curvas de AgGSTE2 e AgGSTE5 quando simuladas com e sem o ligante (Figura 20-A). Porém, na isoforma mutante AgGSTE2mut quando simulada com o GSH (Figura 20-B) observa-se uma notável alteração da curva aos 15 nanossegundos, seguindo estável até os 25 nanossegundos quando novamente torna a subir e se estabiliza. Os valores de RMSD da AgGSTE5 também indicam que o modelo construído é estável e válido.

Figura 16 - Valores de RMSD para as três proteínas sem (A) e com (B) o ligante GSH.

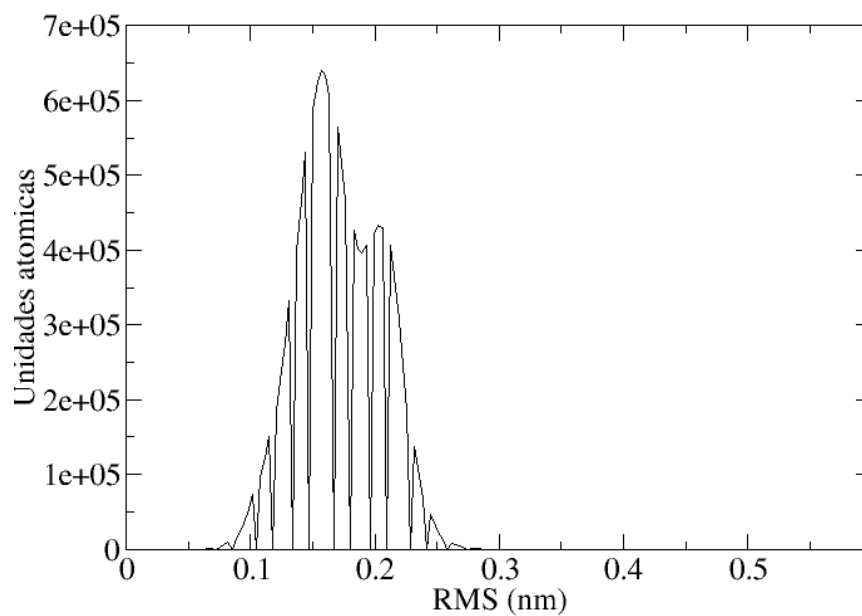




Fonte: Ilustração produzida pelo autor da pesquisa.

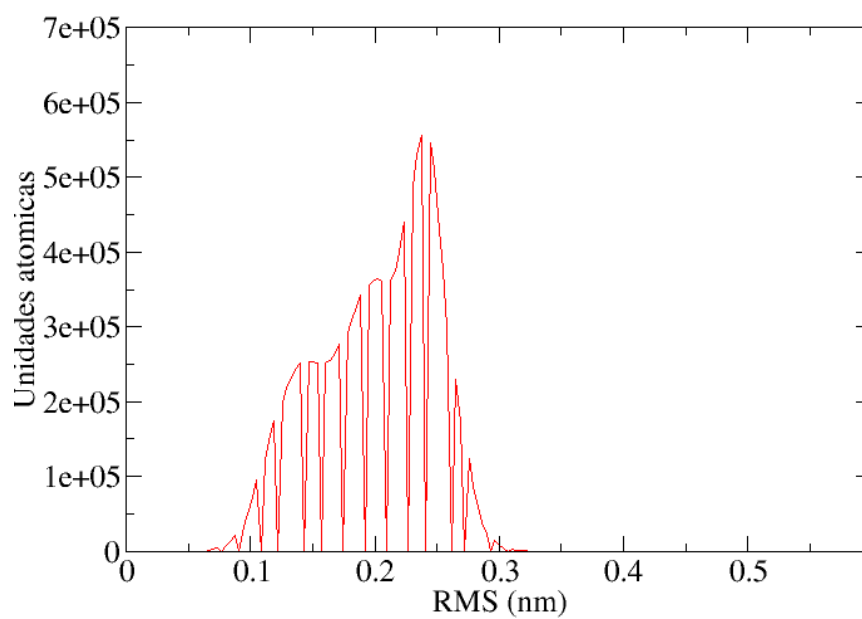
Os gráficos de distribuição do RMSD (Figuras 21 a 26) mostram que a AgGSTE2 manteve o mesmo padrão de distribuição (curvas similares) não havendo diferenças entre a estrutura simulada com e sem o ligante, o que significa que durante a simulação a presença da glutathione não altera a conformação desta proteína, e que portanto apresenta a maior homogeneidade conformacional entre os sistemas simulados. Na AgGSTE5 sem o ligante foi possível perceber duas curvas de distribuição com um pequeno período de transição entre elas sugerindo que este sistema apresenta dois estados conformacionais, porém não muito distintos entre si.

Figura 17 - Gráfico da distribuição do RMSD para AgGSTE2 (sem ligante).



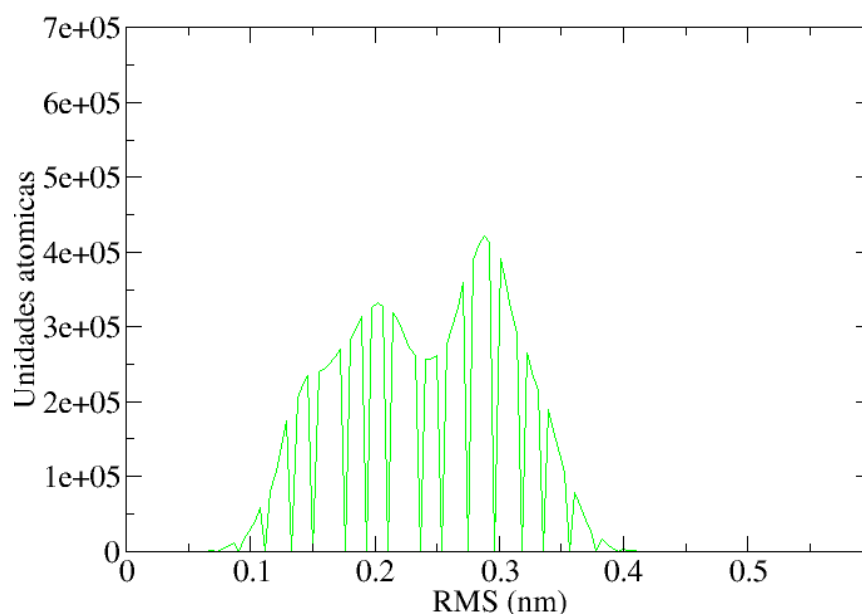
Fonte: Ilustração produzida pelo autor da pesquisa

Figura 18 - Gráfico da distribuição do RMSD para AgGSTE2mut (sem ligante).



Fonte: Ilustração produzida pelo autor da pesquisa

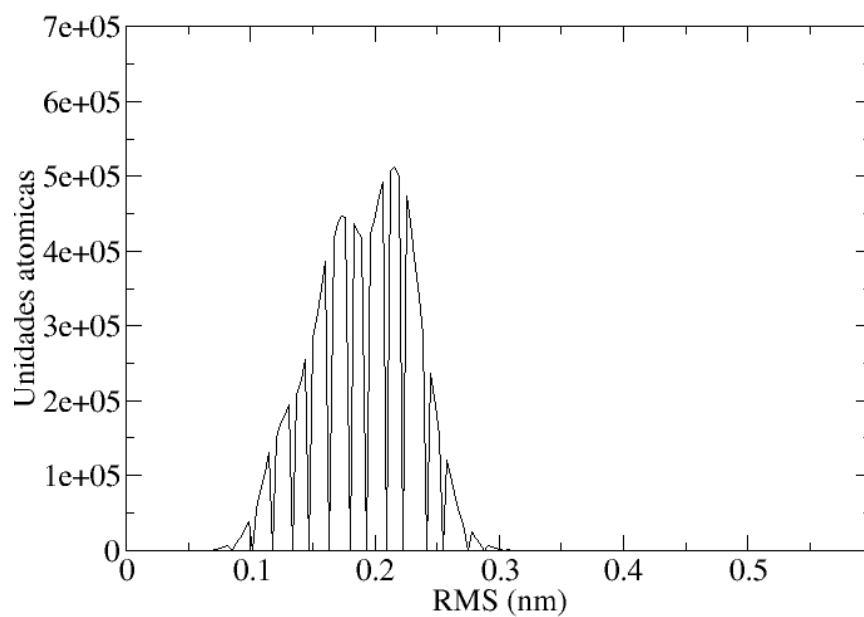
Figura 19 - Gráfico da distribuição do RMSD para AgGSTe5 (sem ligante).



Fonte: Ilustração produzida pelo autor da pesquisa.

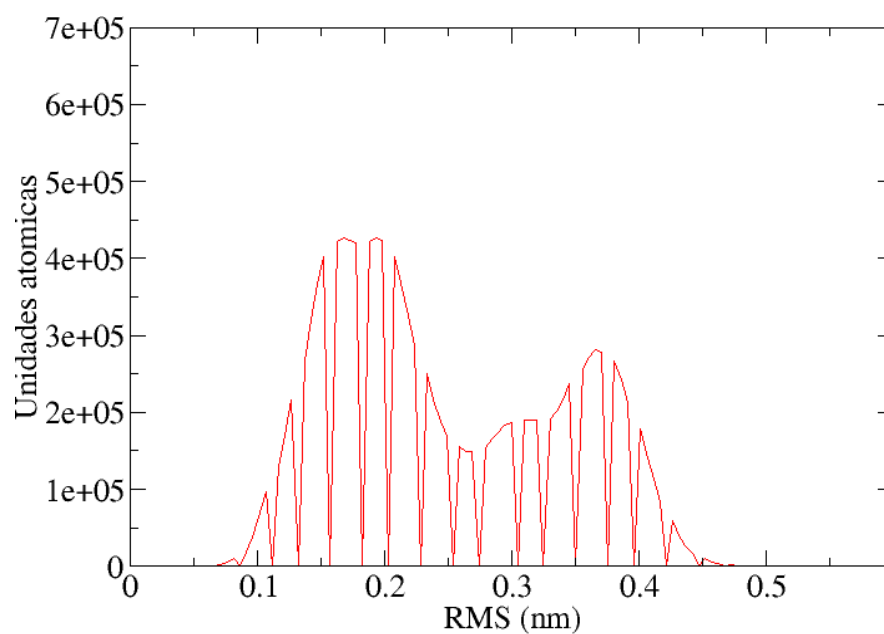
Quando simulada com o ligante, a AgGSTe5 apresentou apenas uma curva bem definida (Figura 26). Estes resultados apontam que a AgGSTe5 tem uma maior homogeneidade conformacional quando simulada com o GSH. Para AgGSTe2mut observou-se que quando ela está sem o ligante a distribuição segue uma curva assimétrica apesar do gráfico do RMSD (Figura 22) apresentar valores constantes e indicar a estabilidade da proteína. No entanto, a dinâmica da AgGSTe2mut sofre uma mudança ao ligar-se com a glutatona, onde pode-se observar duas curvas nítidas na distribuição e um período de transição entre elas, indicando uma heterogeneidade conformacional deste sistema (Figura 25).

Figura 20 - Gráfico da distribuição do RMSD para AgGSTE2 (com ligante).



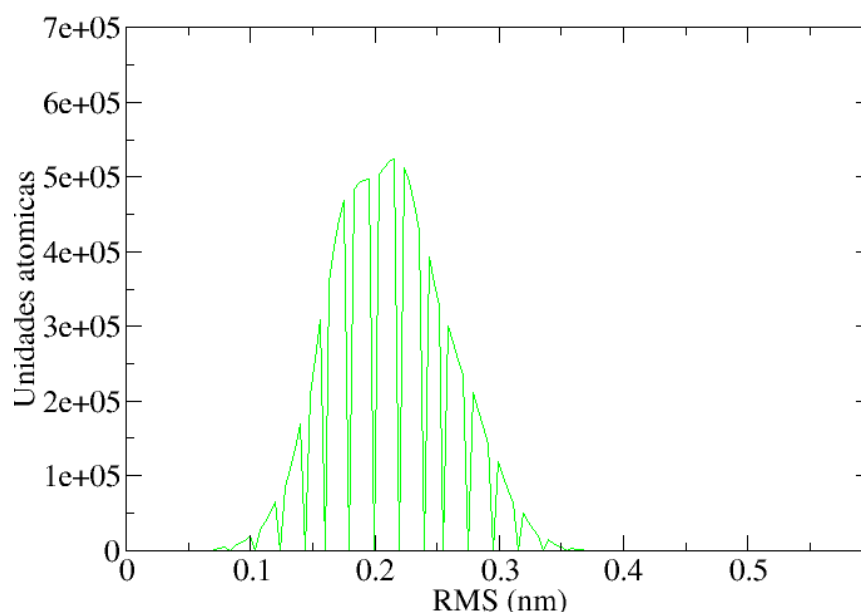
Fonte: Ilustração produzida pelo autor da pesquisa

Figura 21 - Gráfico da distribuição do RMSD para AgGSTE2mut (com ligante).



Fonte: Ilustração produzida pelo autor da pesquisa

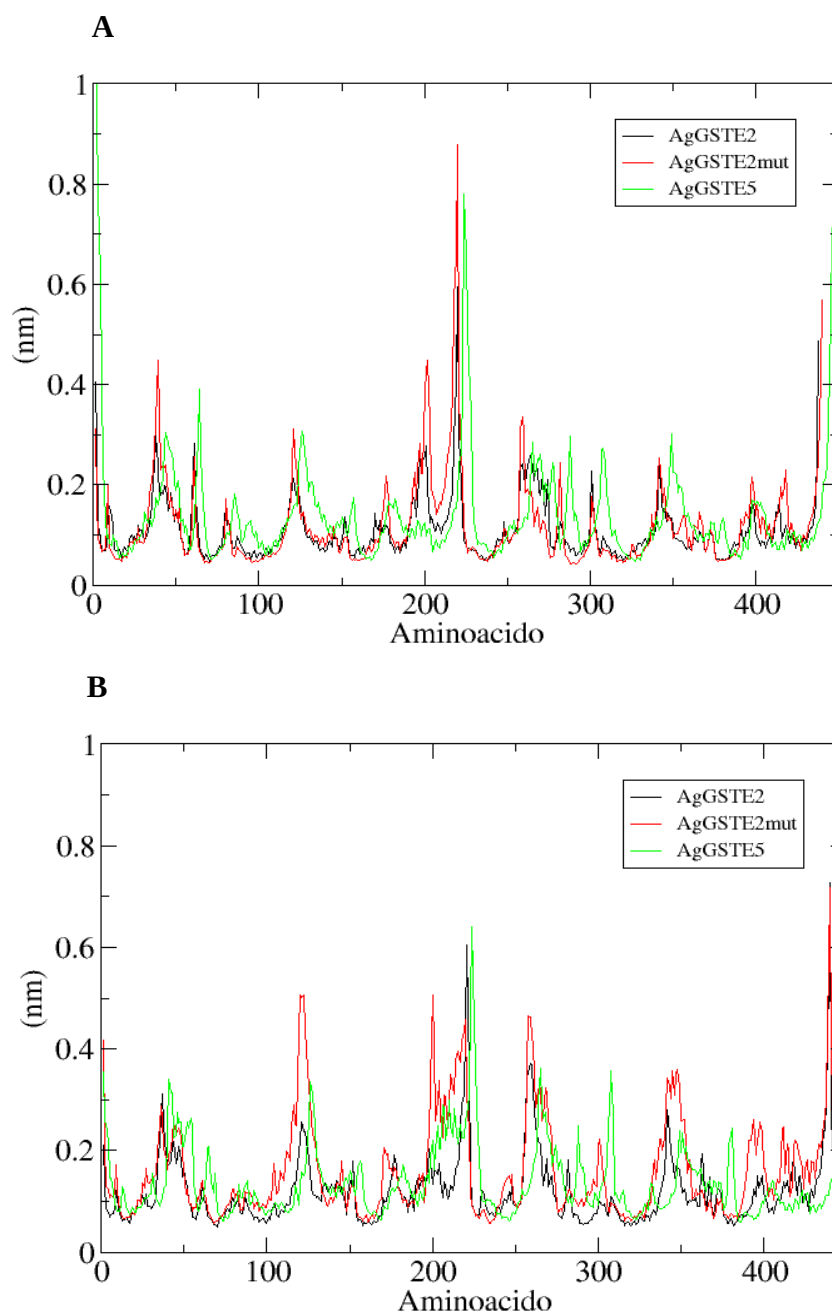
Figura 22 - Gráfico da distribuição do RMSD para AgGSTE5 (com ligante).



Fonte: Ilustração produzida pelo autor da pesquisa

Os gráficos de RMSF (Figura 27) exibem valores de flutuações médias diferentes entre todos os sistemas simulados, com e sem o GSH. No entanto, as regiões de maior flutuação (> 0.2 nm) são muito similares em todos os sistemas notando-se uma ligeira diferença entre AgGSTE2 e AgGSTE5, especificamente entre os resíduos 300-350, que na AgGSTE5 mostra-se maior. Comparando as isoformas AgGSTE2 e AgGSTE2mut percebe-se que as regiões de variações são congruentes, porém apresentam valores médios maiores na AgGSTE2mut. A presença do ligante (GSH) parece contribuir para um pequeno acréscimo nas flutuações da AgGSTE2 e AgGSTE2mut e para um decréscimo na AgGSTE5.

Figura 23 - RMSF para o carbono-alfa nas proteínas com (A) e sem (B) GSH.

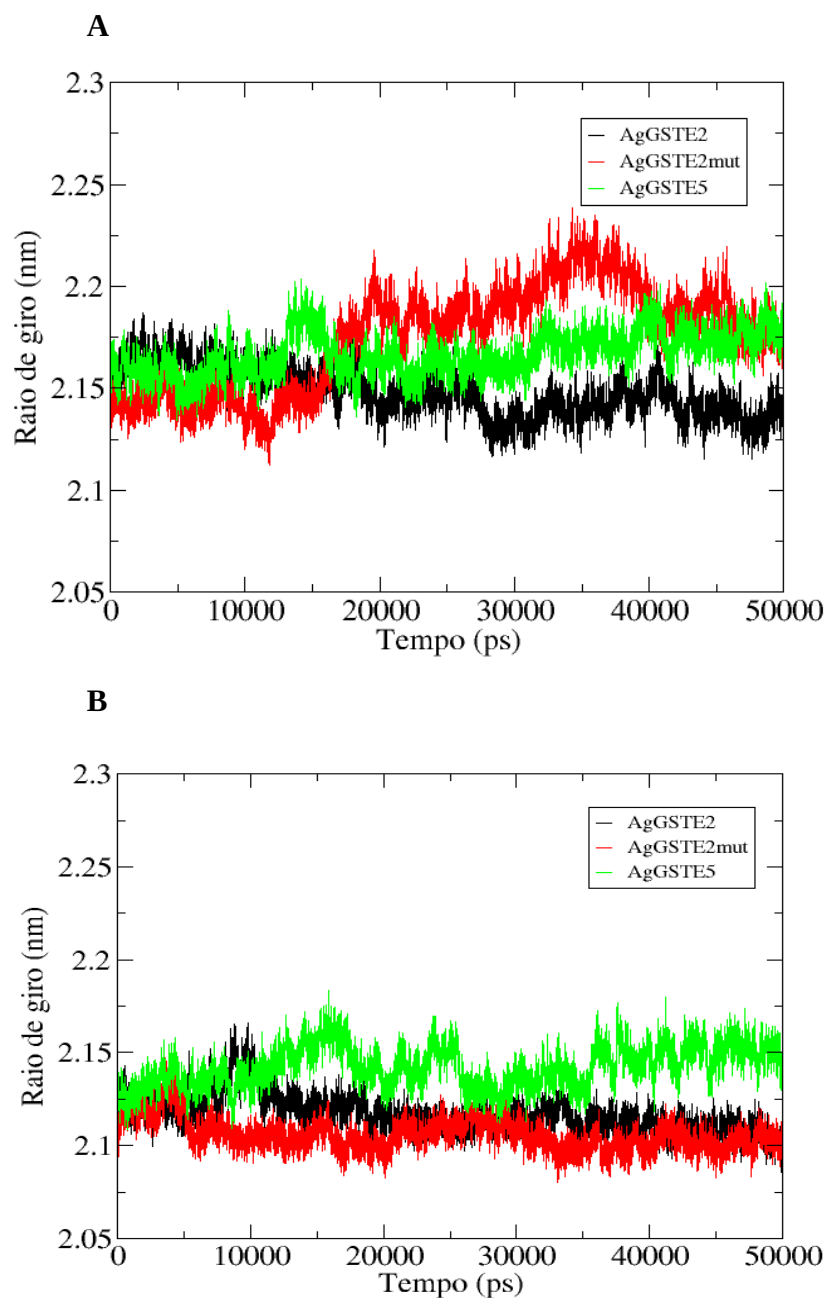


Fonte: Ilustração produzida pelo autor da pesquisa

Para saber se houve processos de compactação ou expansão estrutural das proteínas foi calculado o raio de giro para os seis sistemas (Figura 28) ao longo das simulações. Os gráficos revelam que os valores diferiram entre os três sistemas e que as proteínas tendem a aumentar seu raio de giro quando ligadas a glutathione. Na ausência do GSH os maiores valores de raio de giro

foram observados na curva referente à AgGSTE5, enquanto o menor foi na curva da AgGSTE2mut. Observa-se também uma ligeira diferença entre os raios de giro da AgGSTE2mut e AgGSTE2 no início da simulação mas que convergem ao final.

Figura 24 - Raio de giro nas proteínas simuladas com (A) e sem (B) o ligante.



Fonte: Ilustração produzida pelo autor da pesquisa

Na presença do GSH observam-se valores de raio de giro ligeiramente maiores na AgGSTe2 e na AgGSTe5 porém a curva segue o mesmo padrão e ao final da simulação observa-se um decréscimo de aproximadamente 0,25 nm na AgGSTe2 e um acréscimo também de 0,25 nm na AgGSTe5. Na AgGSTe2mut foram observadas mudanças maiores nos valores e no padrão da curva. Quando analisada sem o ligante observa-se um decréscimo dos valores de raio de giro em torno 0,25 nm e ao final, os valores convergem com os da sua isoforma do tipo selvagem (AgGSTe2). Quando simulada com o ligante GSH observa-se que AgGSTe2mut apresenta a partir dos 10 nanossegundos um acréscimo do raio de giro em 0,5 angstrom e atinge um platô aos 20 nanossegundos. A curva sofre um novo acréscimo aos 30 nanossegundos e em seguida decresce aos 40 nanossegundos estabilizando com valores de 0,25 angstroms maior que a estrutura de início da simulação, convergindo com os valores de sua homóloga AgGSTe5.

A superfície de acessibilidade ao solvente (SASA) foi calculada para verificar se houve mudanças na hidrofobicidade das enzimas (Figuras 29 e 30). A SASA mantém-se estável nos três sistemas simulados sem o ligante, e observa-se valores muito próximos nas curvas referentes às isoformas da AgGSTe2 (mutante e tipo selvagem). Quando simulada com o GSH, a AgGSTe5 apresentou uma área acessível ao solvente um pouco maior (~ 5 nm) que as duas isoformas e os seus valores convergem no final da simulação com os da AgGSTe2 e AgGSTe2mut. A SASA da AgGSTe2 e AgGSTe2mut mostraram ligeiras oscilações quando simuladas com o GSH.

Figura 25 - Superfície de acessibilidade ao solvente dos sistemas com a glutathiona

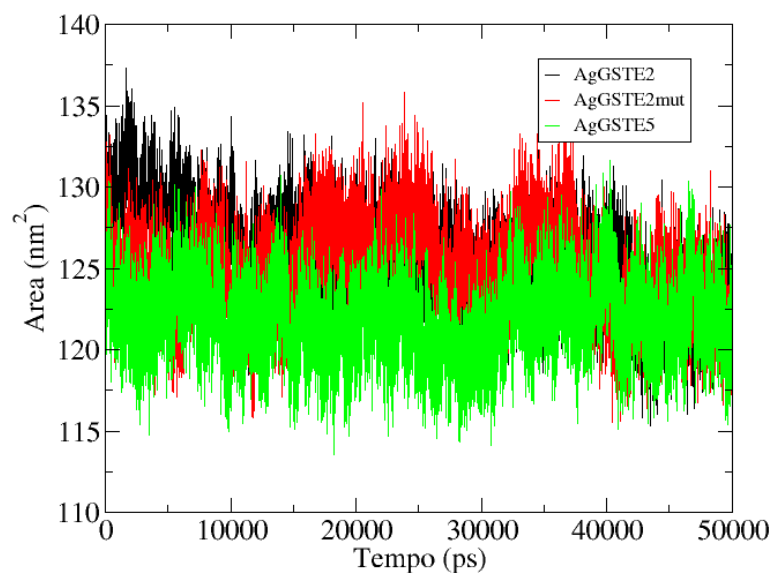
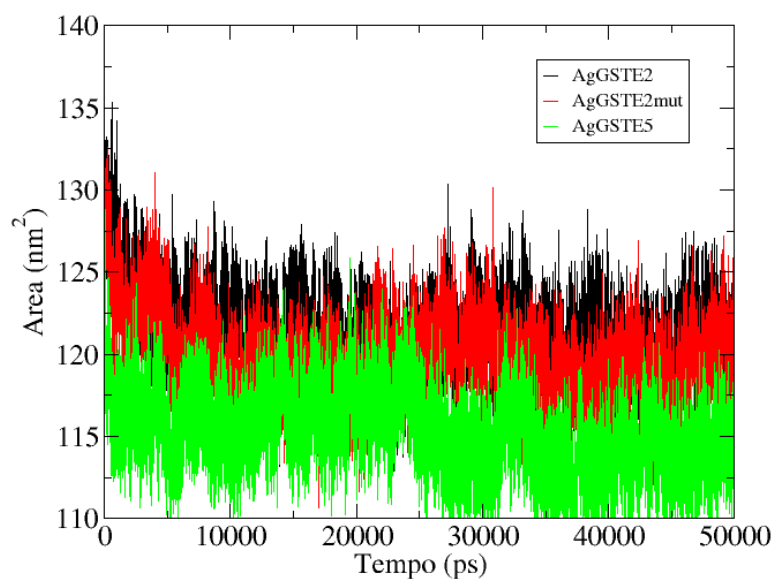


Figura 26 - Superfície de acessibilidade ao solvente dos sistemas sem a glutathiona.



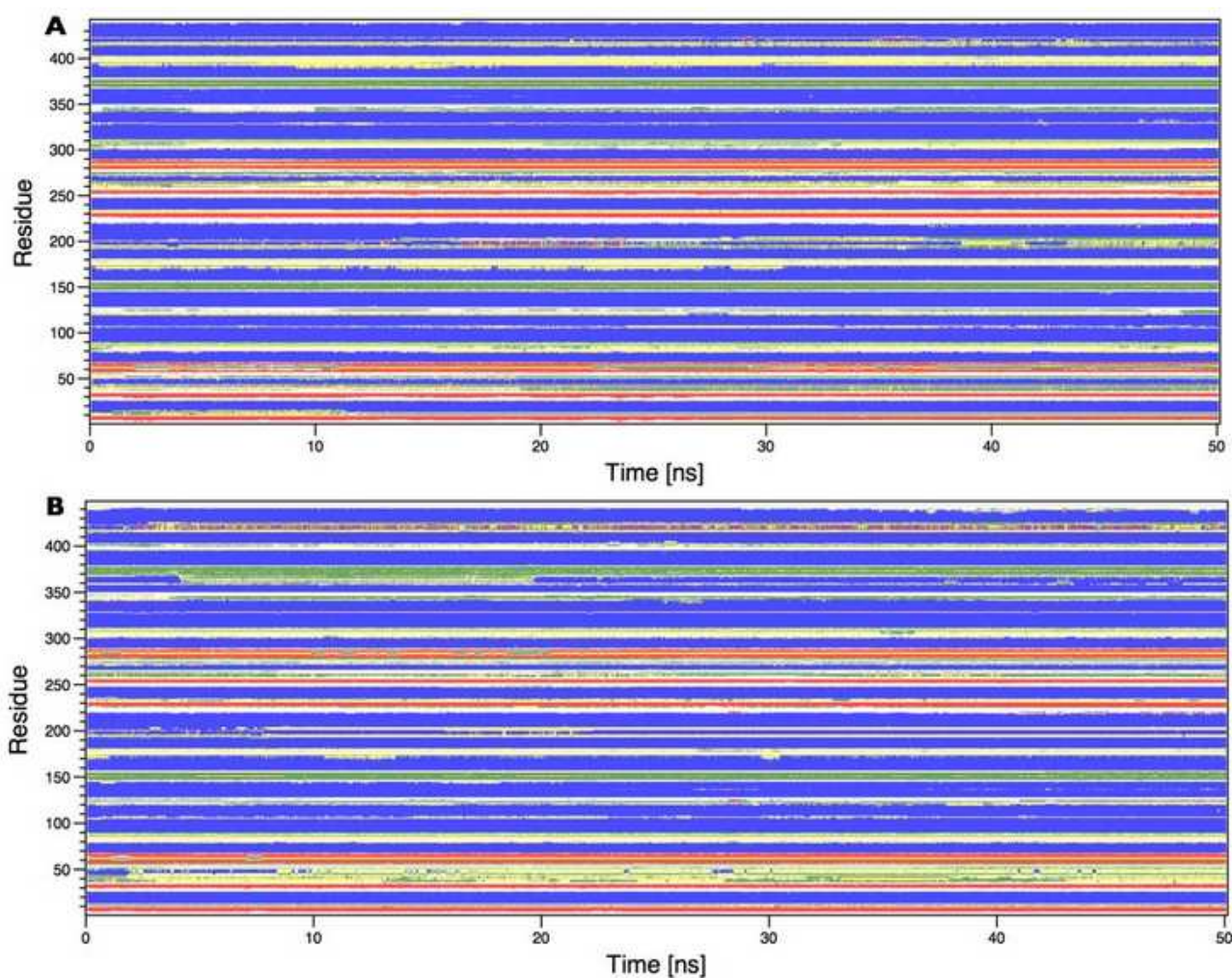
Fonte: Ilustrações 29 e 30 produzidas pelo autor da pesquisa

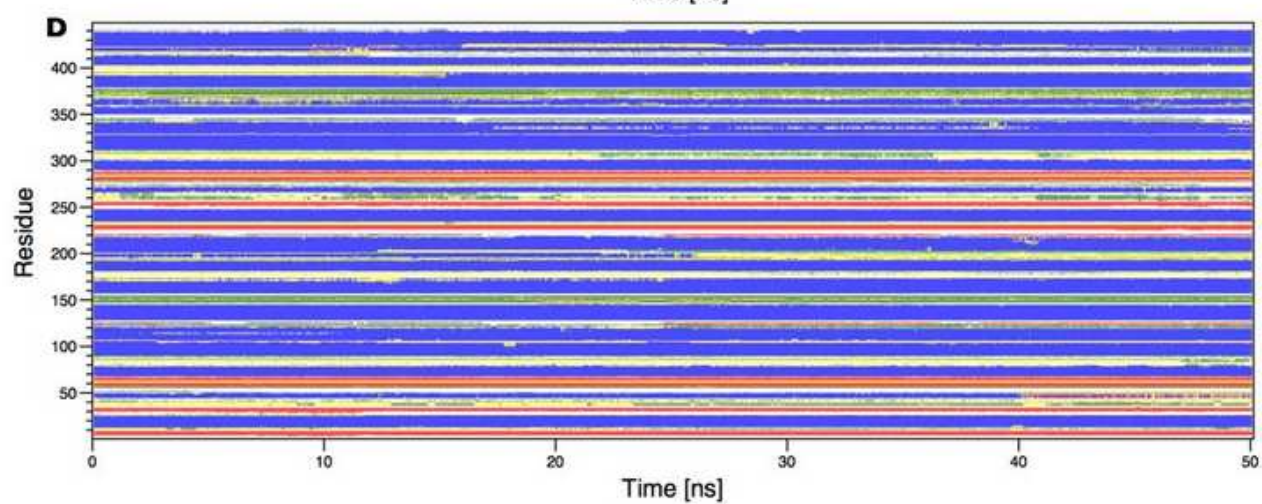
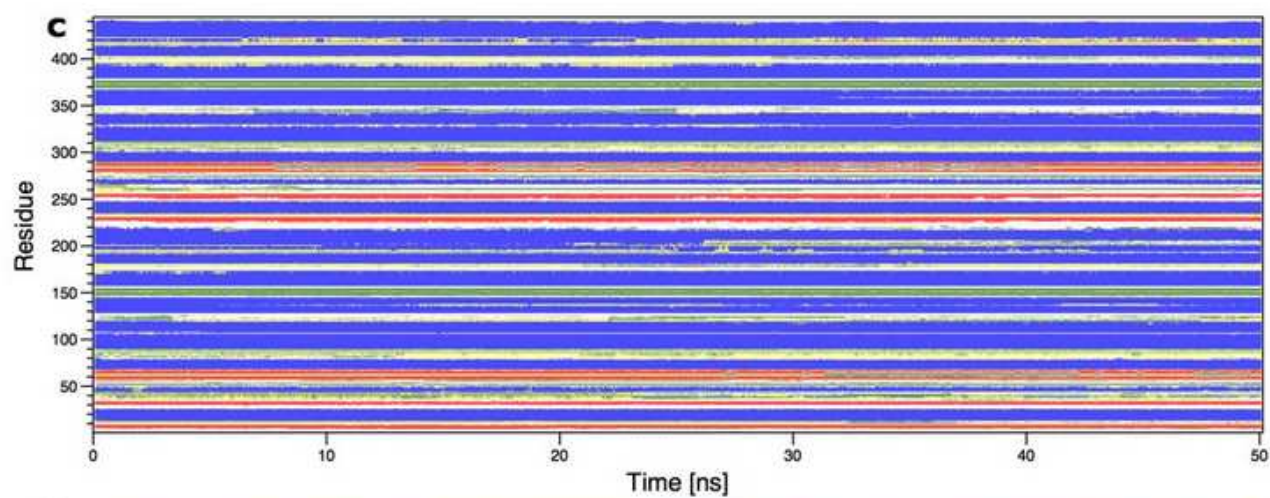
4.3.2 Conteúdo da estrutura secundária ao longo do tempo

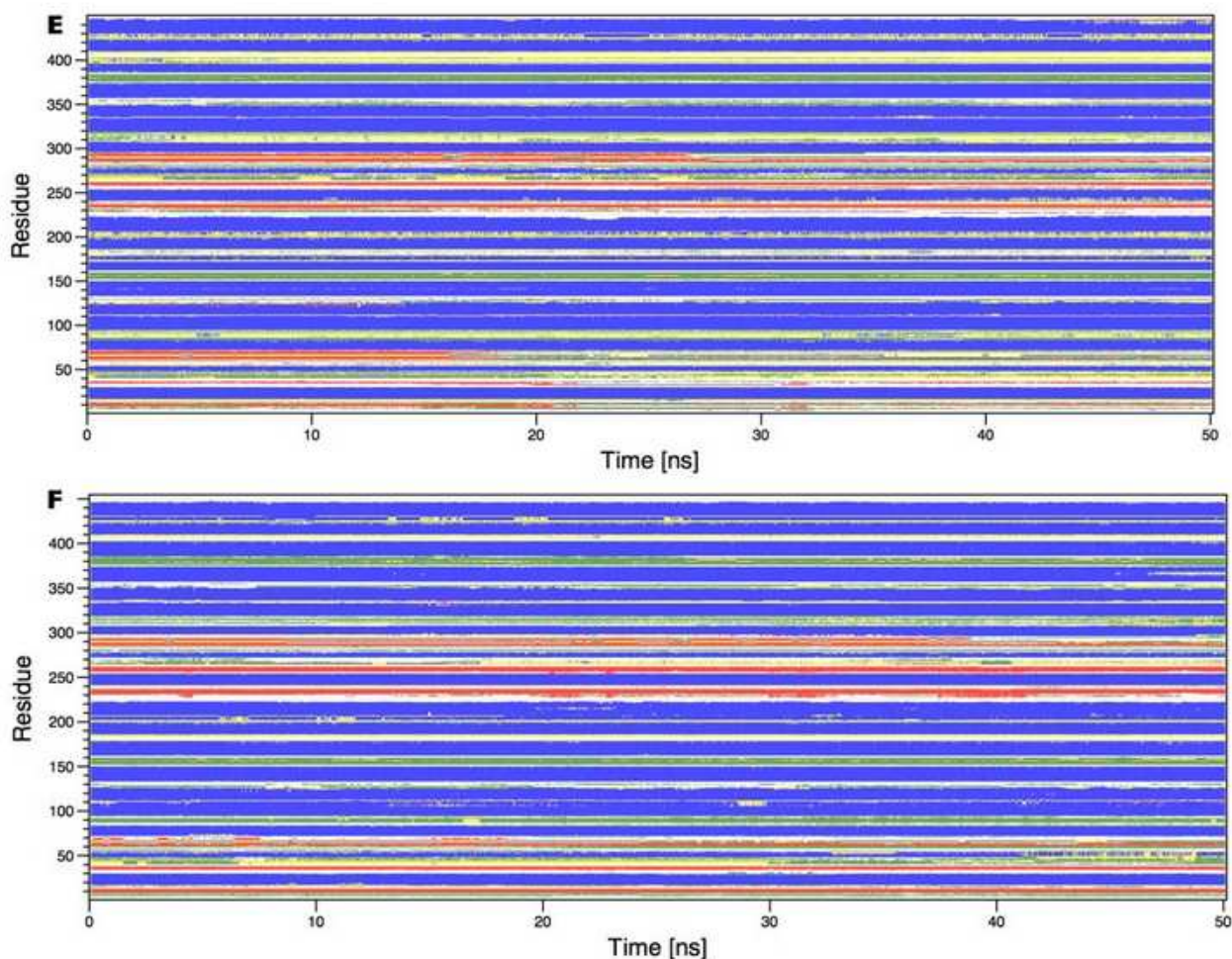
O teor da estrutura secundária permite avaliar se as proteínas apresentam alterações estruturais durante a simulação para cada tipo de estrutura secundária (hélices, folhas, alças, etc)

assim como saber quais são as regiões de transição. Os gráficos da Figura 31 mostram que a estrutura secundária dos sistemas é majoritariamente mantida durante as simulações, o que significa que não está ocorrendo desnaturação. As estruturas tendem a ser mantidas durante a dinâmica, principalmente as regiões compostas por alfa-hélices. Observam-se apenas poucas diferenças sutis entre a estrutura secundária dos sistemas com e sem ligante, não havendo, portanto, alterações significativas no nível de estrutura secundária nestas enzimas.

Figura 27 - Teor da estrutura secundária dos sistemas simulados: A) AgGSTE2 sem ligante; B) AgGSTE2 com ligante; C) AgGSTE2mut sem ligante; D) AgGSTE2mut com ligante; E) AgGSTE5 sem ligante; F) AgGSTE5 com ligante.







Fonte: Ilustração produzida pelo autor da pesquisa

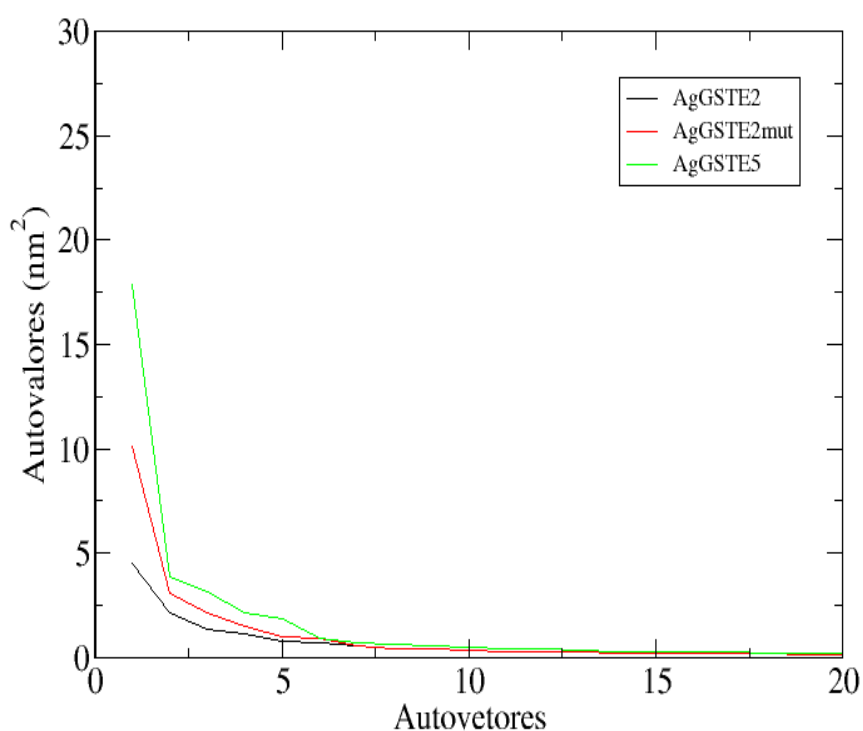
4.3.3 Análise de componentes principais.

A análise dos componentes principais (PCA) é o mais conhecido e utilizado método multivariado de redução de dados que usa técnicas matemáticas para identificar padrões em uma matriz de dados. O principal elemento deste método consiste na construção de um pequeno conjunto de novas variáveis ortogonais, ou seja, não correlacionadas, derivadas de uma combinação linear das variáveis originais (Sant'Anna 2002). Em trajetórias de dinâmica

molecular, a PCA é útil uma vez que através deste método é possível detectar flutuações de grande amplitude e distingui-las das flutuações locais de baixa amplitude.

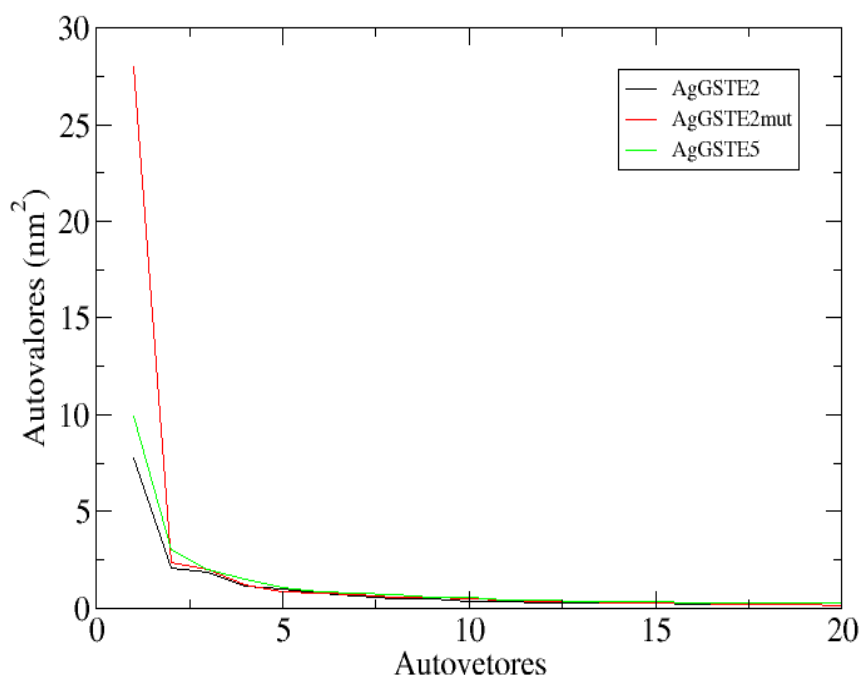
Os movimentos de grande amplitude da AgGSTE2, AgGSTE2mut e AgGSTE5 são predominantemente descritos pelos primeiros autovalores e seus respectivos autovetores (Figuras 32 e 33). Quando simuladas sem o ligante o deslocamento atômico do primeiro autovetor foi de 27 nm², 10 nm² e 8 nm² para AgGSTE2mut, AgGSTE5 e AgGSTE2 respectivamente. Na simulação com a glutathione os deslocamentos atômicos dos primeiros autovetores foram de 17 nm², 10 nm² e 4 nm².

Figura 28 - Autovalores x autovetores para átomos do *backbone* nos sistemas sem GSH.



Fonte: Ilustração produzida pelo autor da pesquisa.

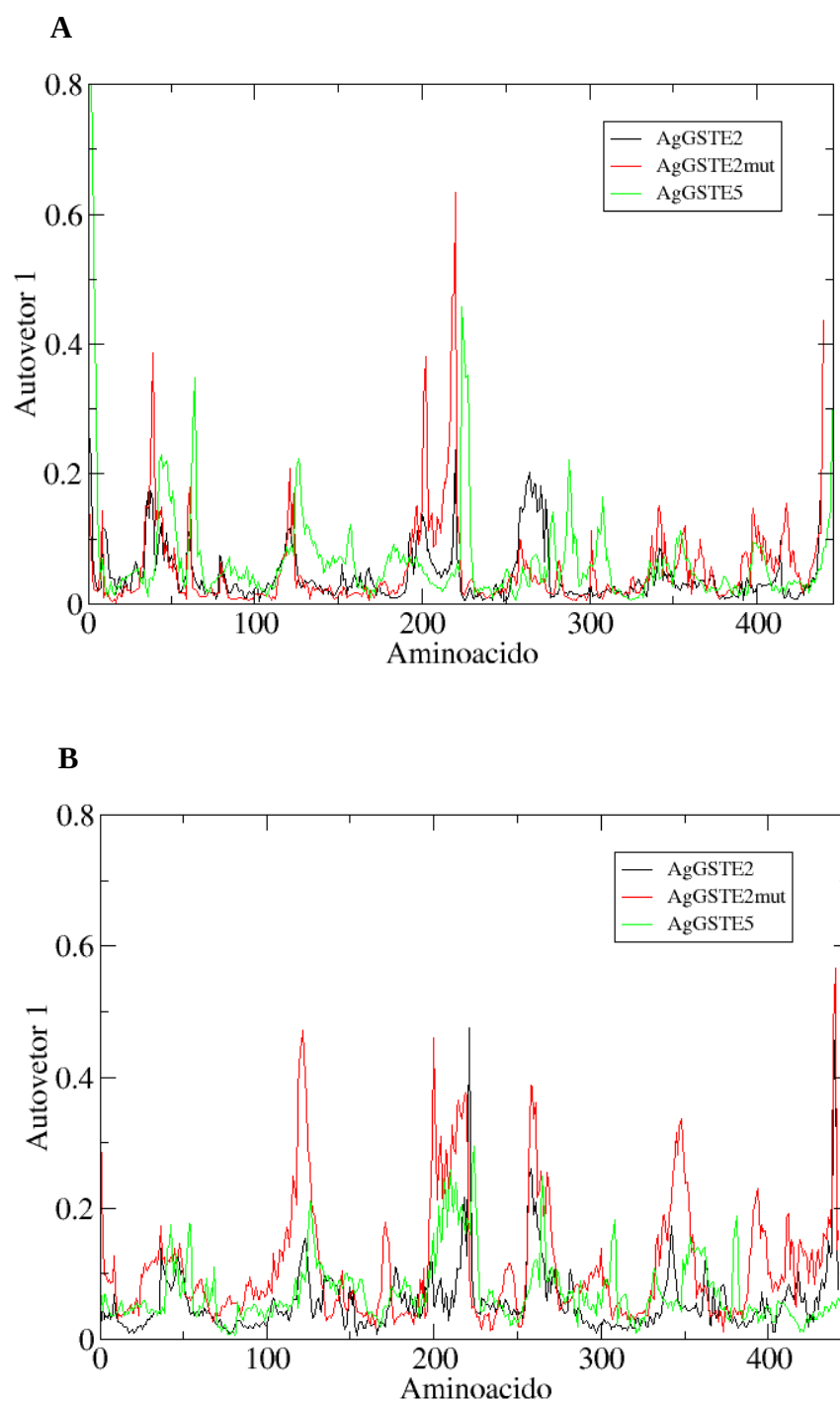
Figura 29 - Autovalores x autovetores para átomos do *backbone* nos sistemas com GSH



Fonte: Ilustração produzida pelo autor da pesquisa

Foram elaborados gráficos (Figura 34) do RMSF do primeiro autovetor de cada sistema para verificar quais regiões tem maior flutuação, e, portanto, contribuem mais para o movimento das enzimas. Na AgGSTE2 sem ligante observa-se que as regiões de maior flutuação (> 3 angstroms) situam-se nos resíduos 35-40, 118-122 e 190-210 que correspondem às hélices 2-3,7 e 11-12 e algumas regiões de alça (Figura 34-A). A AgGSTE2mut sem ligante apresentou maiores flutuações do primeiro autovetor nos resíduos 35-45, 195-205 e 210-220, que formam as hélices 2-3-4, 11-12 e 12 e compõem também algumas regiões de alça (Figura 35-B). No primeiro autovetor AgGSTE5 sem o GSH as regiões que mais contribuíram foram dos resíduos 1-10, 40-50, 120-130 e 220-240, sendo as duas primeiras regiões referentes à folha-beta 1 e as demais situam-se nas hélices 7-8 e 12, compondo parcialmente algumas regiões de alça (Figura 35-C).

Figura 30 - RMSF para o primeiro autovetor nos sistemas sem (A) e com (B) GSH.



Fonte: Ilustração produzida pelo autor da pesquisa.

Figura 31 - Regiões de maior contribuição para o primeiro autovetor das enzimas AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C) sem o ligante.

A

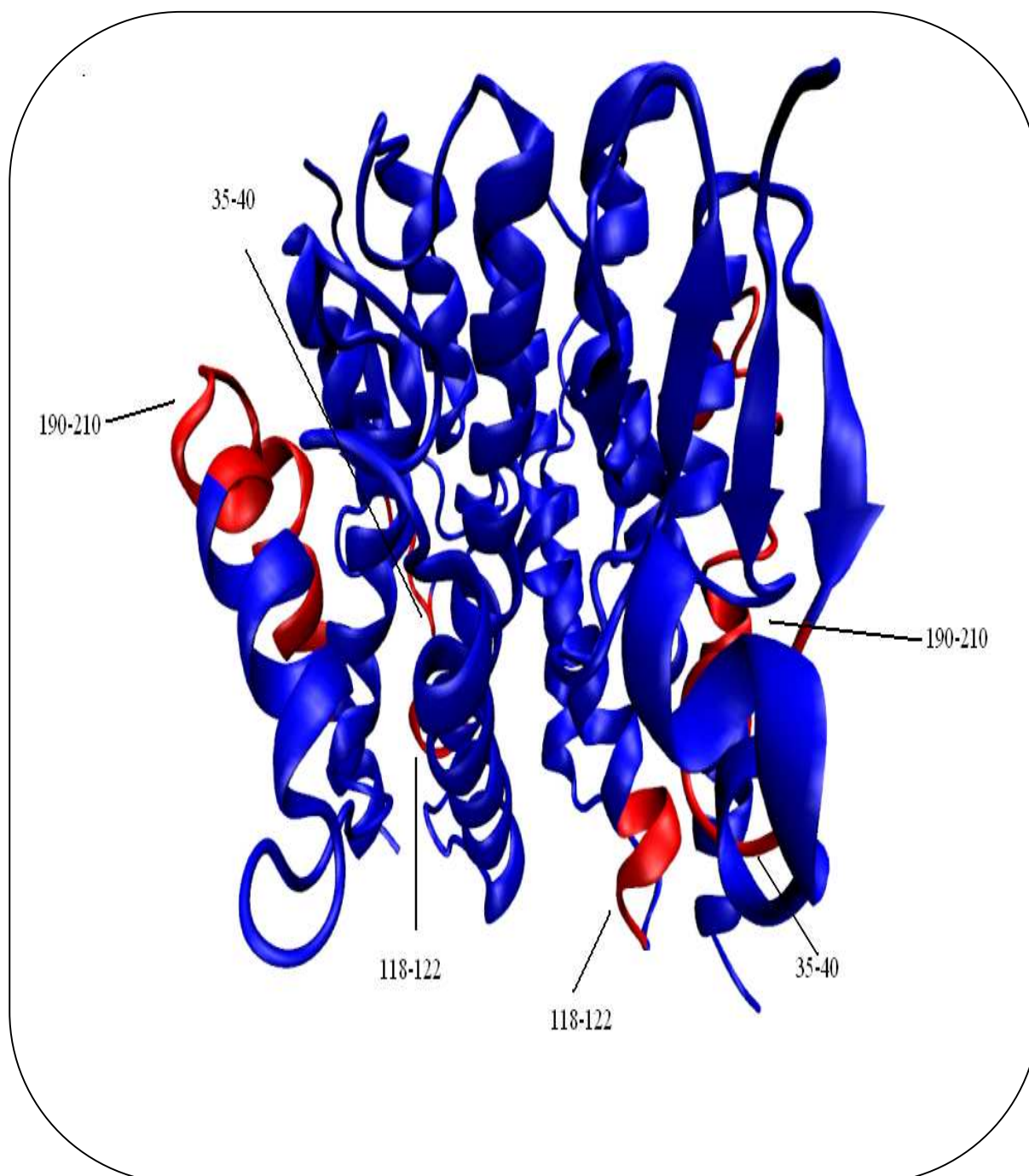
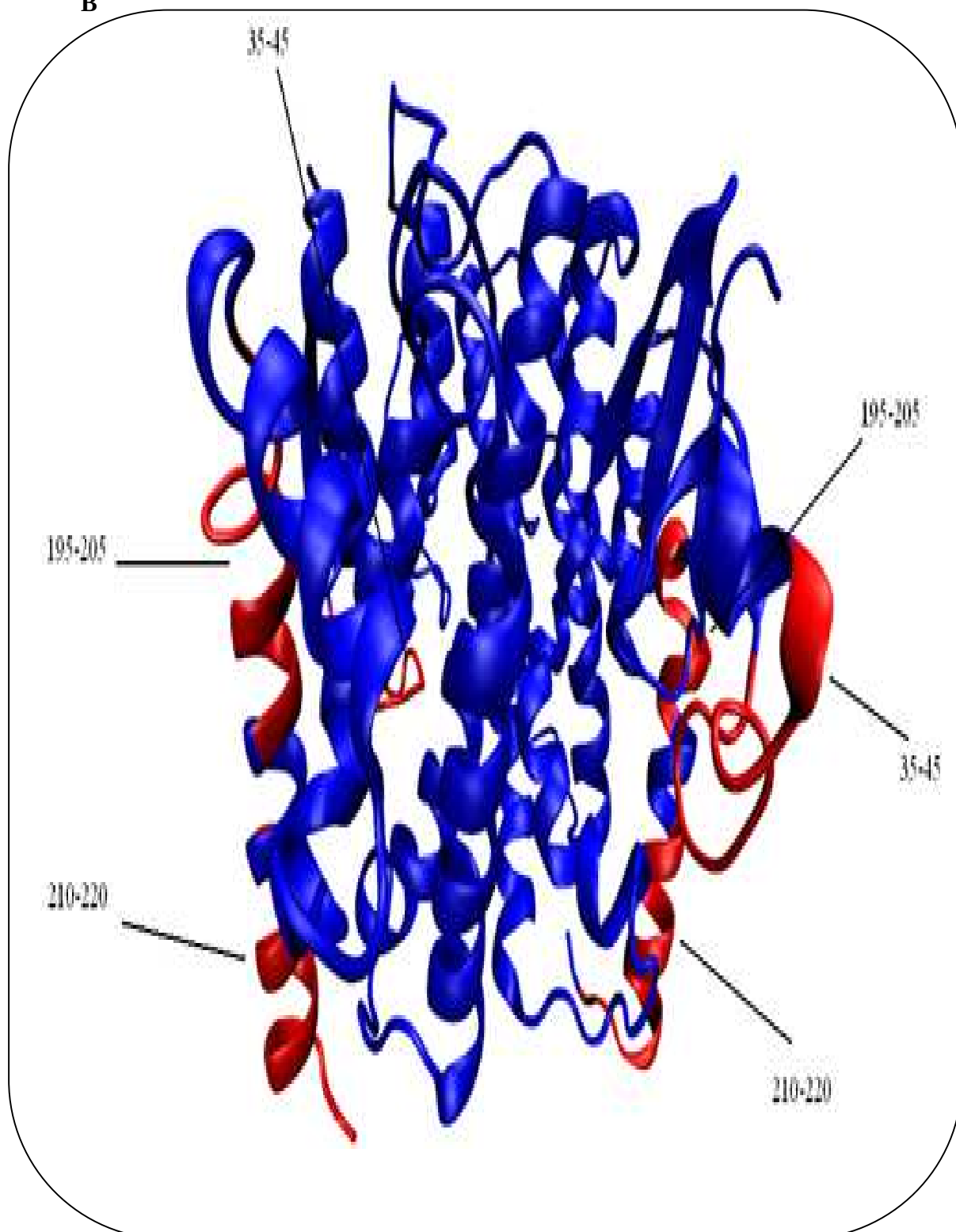


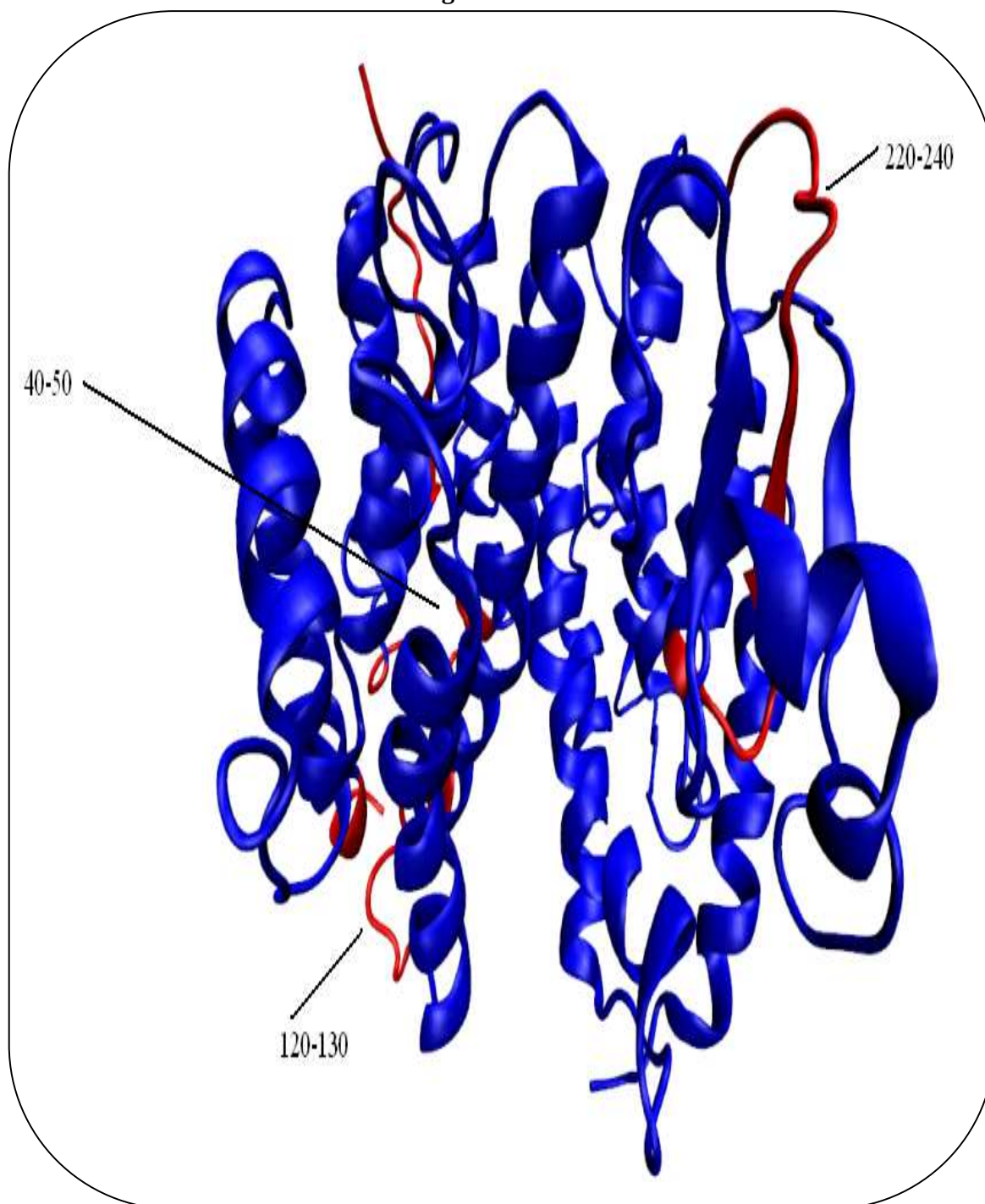
Figura 31 - Continua

B



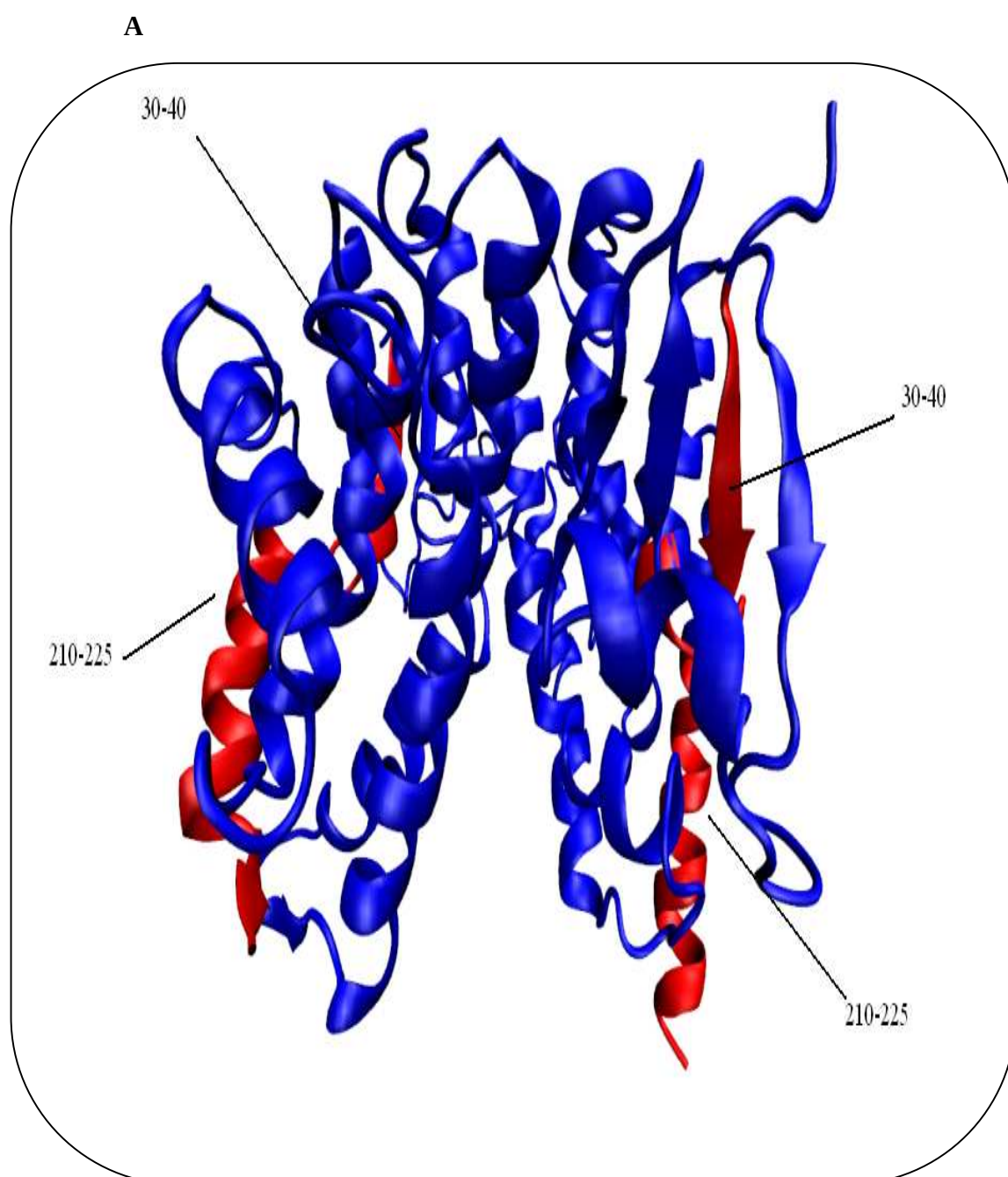
C

Figura 31 - Continua



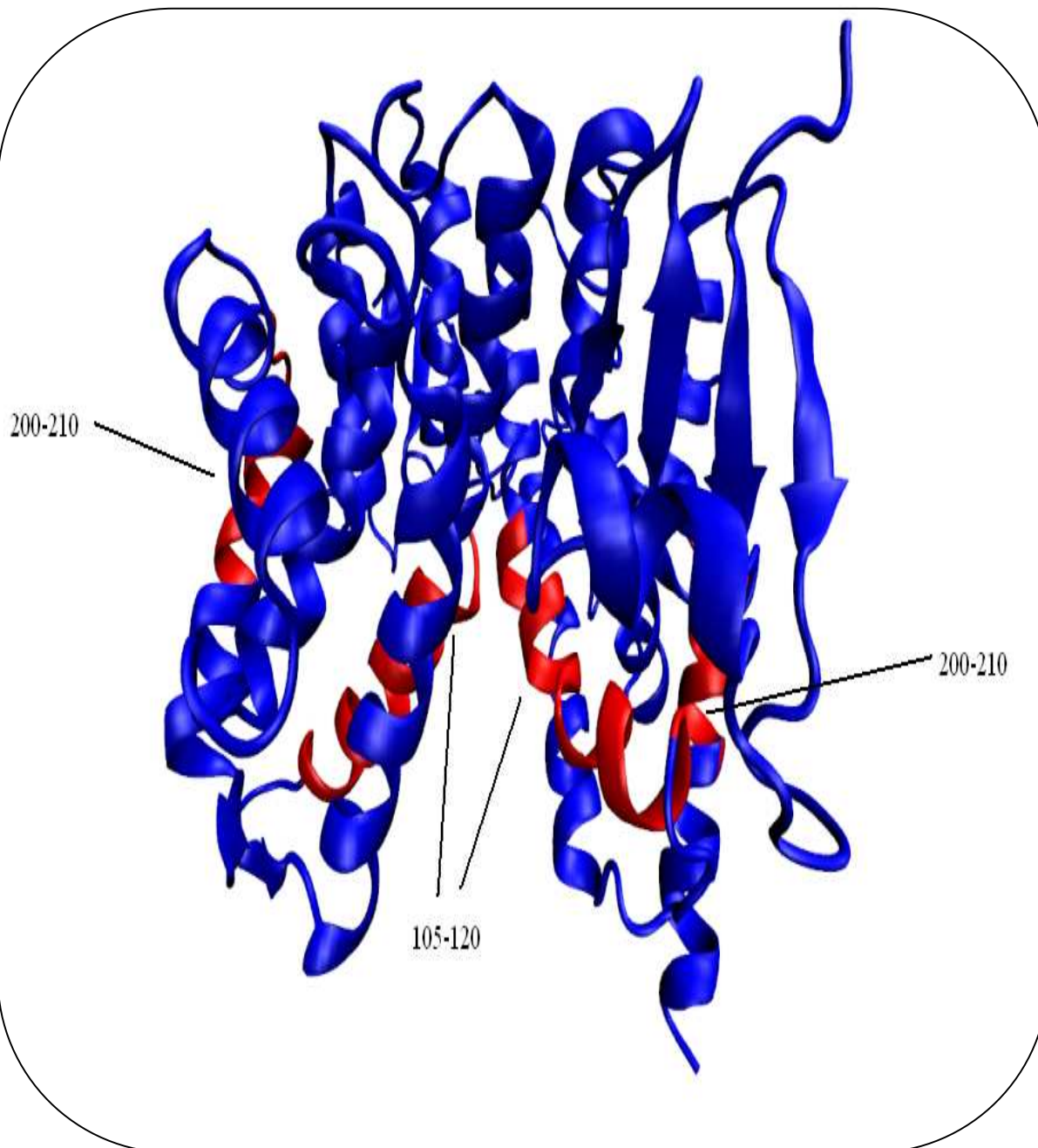
Fonte: Ilustração produzida pelo autor da pesquisa.

Figura 32 - Regiões de maior contribuição para o primeiro autovetor das enzimas AgGSTE2 (A), AgGSTE2mut (B) e AgGSTE5 (C) com o ligante.



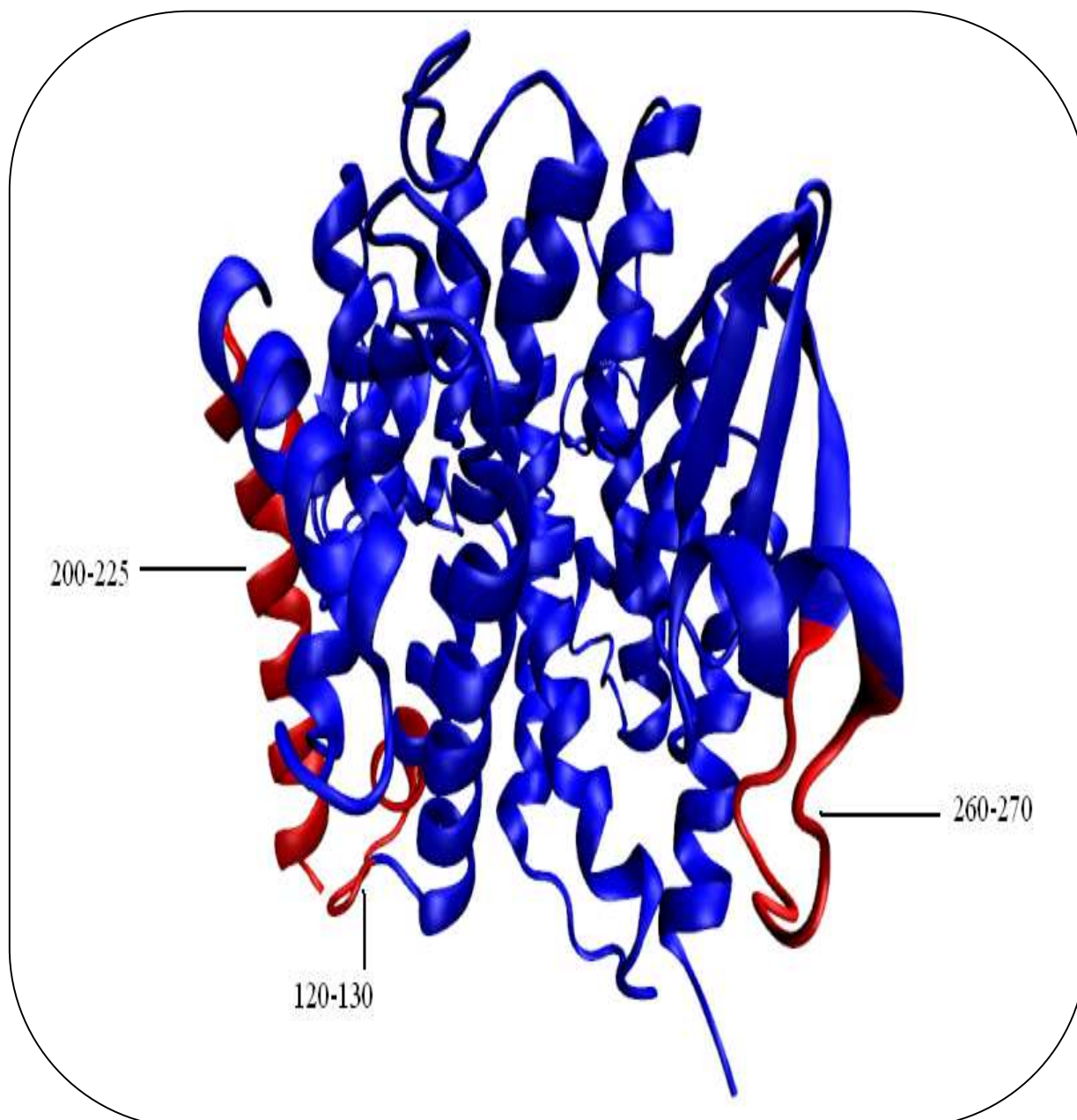
B

Figura 32 - Continua



C

Figura 32 – Continua.



Fonte: Ilustração produzida pelo autor da pesquisa

4.3.4 Análise dos clusters

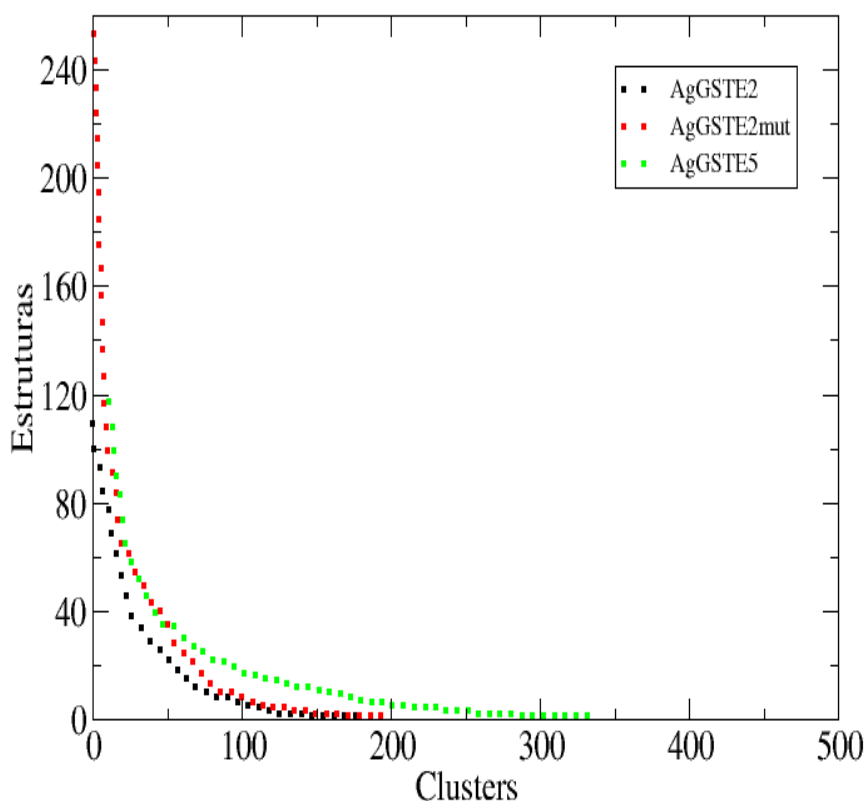
A análise dos *clusters* das simulações de DM foi realizada para identificar o grau de sobreposição de conformação entre as estruturas amostradas dentro dos conjuntos AgGSTe2, AgGSTe2mut e AgGSTe5 para avaliar a heterogeneidade conformacional dos sistemas. Esta análise classifica as estruturas de uma trajetória em grupos (*clusters*) de estruturas semelhantes como uma função do *cut-off* do RMSD (0,1 nm) dos átomos de *backbone* entre os pares de estruturas da trajetória. Uma vez que a amostragem conformacional em simulações MD evolui de acordo com uma distribuição de Boltzmann, os *clusters* de conformações derivados da MD apresentam diferentes tamanhos onde sub-estados de maior energia são menos populados que os sub-estados de menor energia.

Ao comparar o número de clusters em relação aos números de estruturas por *clusters* (figura 37) observa-se que os sistemas apresentam *clusters* mais populados quando estão sem o GSH. A AgGSTe5 foi a enzima que apresentou o maior número de *clusters* quando simulada com o ligante e sem ele. No entanto, mais da metade dos *clusters* foram populados com poucas estruturas (< 20), inclusive observa-se vários *clusters* compostos por uma única conformação. Na AgGSTe2 observou-se também mais clusters quando ligada ao GSH, porém os *clusters* foram menos populados que na estrutura simulada sem o ligante. O mesmo padrão pode ser observado na AgGSTe2mut, com a diferença de que este sistema apresentou os *clusters* com maior número de estruturas quando simulada com o ligante (> 250 estruturas) assim como sem o ligante (> 160 estruturas). Por outro lado, o número de transições entre os *clusters* foi menor na AgGSTe5, onde não observa-se nenhum *cluster* com mais de 50 transições. Admitindo que o sistema está em equilíbrio estes resultados sugerem um número equivalente de subestados de baixa energia, conectados por conformações transitórias dos cluster menos populados. Apesar de apresentar um

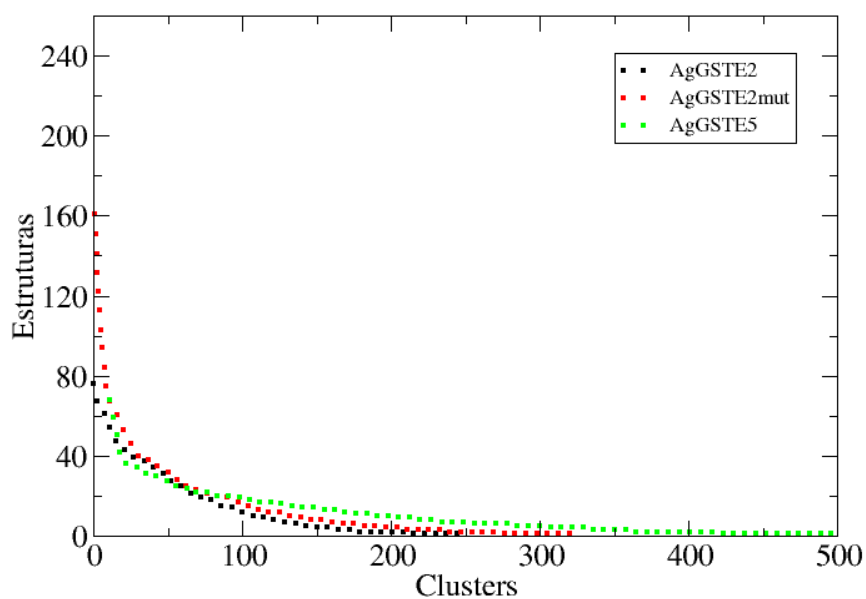
maior número de *clusters*, a maior parte destes *clusters* apresenta poucas transições (<10). Na AgGSTE2 e AgGSTE2mut o número de transições foram próximos, sendo que a AgGSTE2 mostrou *clusters* com maior numero de transições em comparação a sua isoforma, enquanto AgGSTE2mut obteve mais *clusters* com transições. Para os três sistemas constata-se que o número de transições por *clusters* diminui quando simulados com a glutathione. Ao mesmo tempo é possível constatar um aumento no numero de *clusters*, embora estejam populados com menos transições, nos sistemas com o GSH.

Figura 33 - Número de estruturas (populações) por cluster nos sistemas sem (A) e com (B) GSH.

A



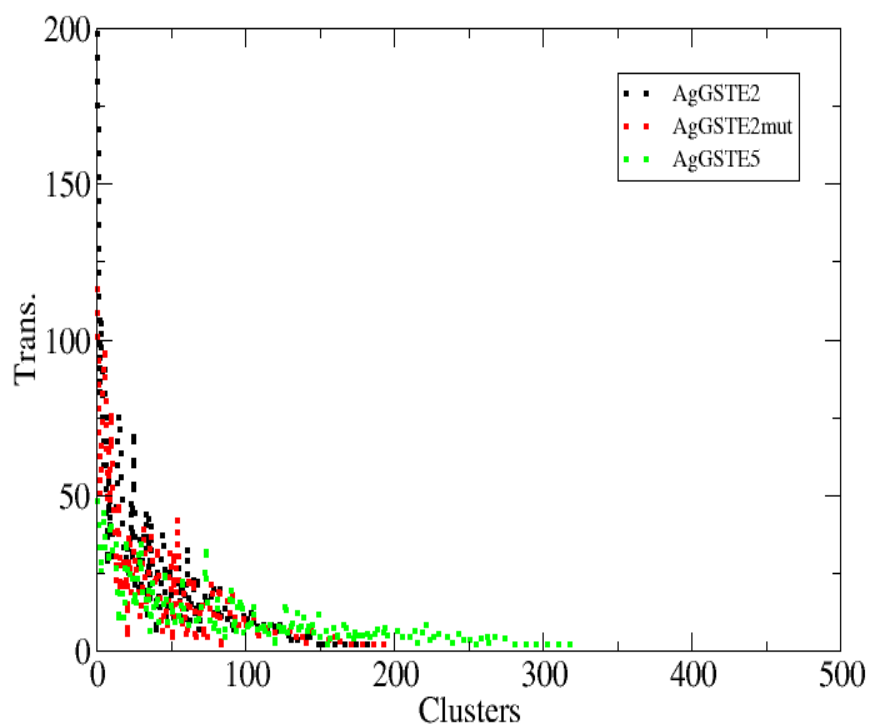
B

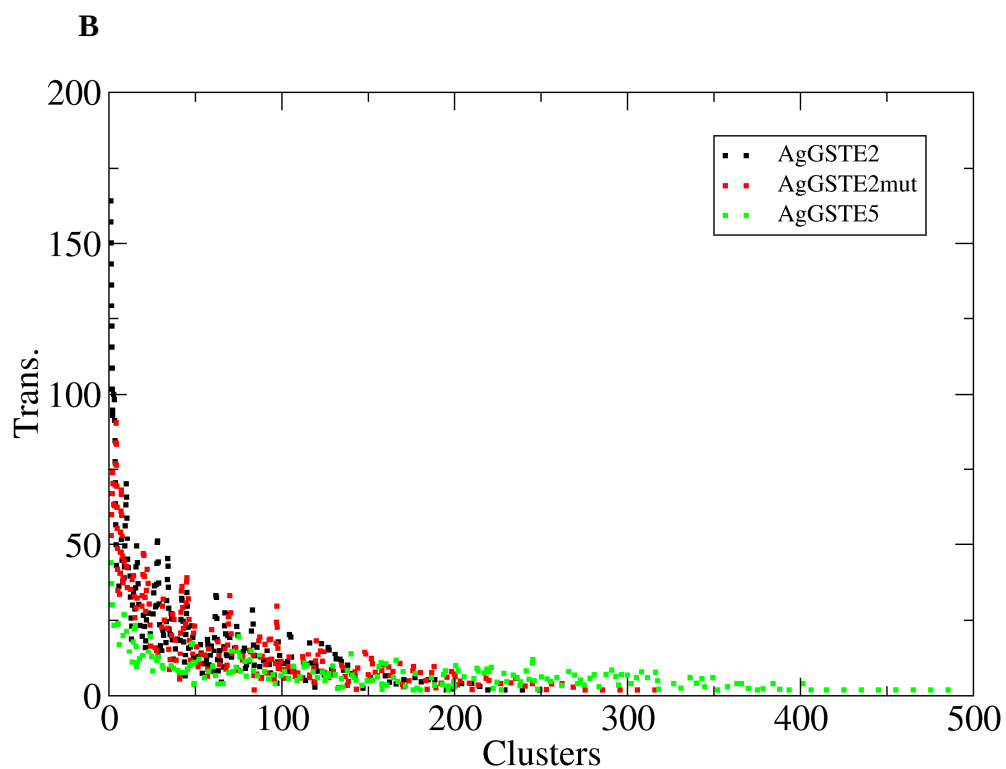


Fonte: Ilustração produzida pelo autor da pesquisa

Figura 34 – Número de transições entre os clusters para as proteínas sem (A) e com (B) GSH

A





Fonte: Ilustração produzida pelo autor da pesquisa

4.4 DOCKING

4.5 MOLECULAR

4.4.1 Predição do modo de ligação

As energias de ligação entre as três proteínas e os cinco diferentes compostos estudados foram calculadas e estão disponíveis na tabela 11. As conformações de menor energia de cada

complexo foram analisadas visualmente (VMD) e foram medidas as menores distâncias entre o GSH e o ligante (Figuras 39 a 44).

Tabela 10 - Energias de ligação (Kcal/mol) para as melhores conformações no ranking de cada complexo

Ligantes	DDT	CDNB	Carbaril	Cipermetrina	Malation
Proteínas					
AgGSTE2	-5.13	-7.32	-5.85	-8.37	-3.31
AgGSTE2mut	-9.16	-6.99	-6.09	-8.81	-3.67
AgGSTE5	-7.68	-7.17	-6.42	-8.64	-3.24

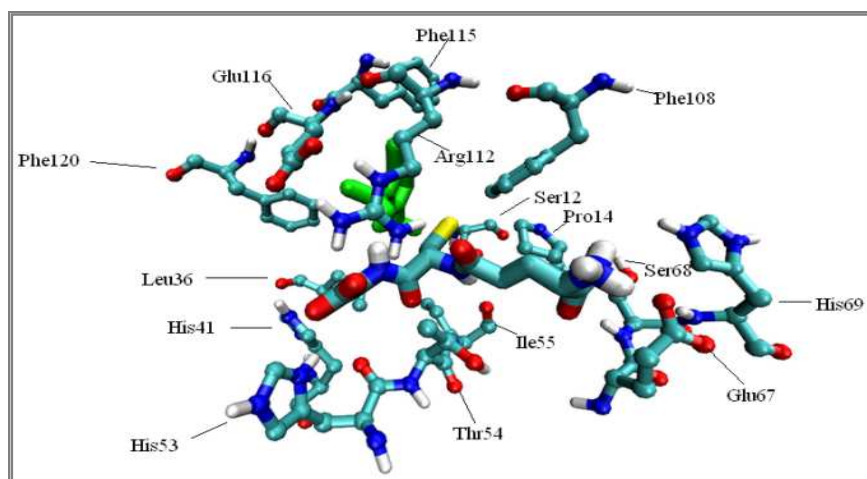
A menor energia foi observada no complexo AgGSTE2mut-DDT, indicando uma maior afinidade entre esta enzima e este inseticida. A distância observada (Figura 38-B) entre o DDT e o GSH (< 4 angstroms) e a posição demonstra que este confôrmero é um candidato potencial a metabolizar o DDT. A energia de ligação deste complexo foi a menor entre todas as comparações, inclusive para o CDNB, que é um composto orgânico sintético benzeno-derivado e é utilizado como modelo padrão para pesquisas com GSTs de diversas classes (Ministério da Saúde, 2006). No *docking* com o DDT observou-se energias um pouco maiores para AgGSTE2 e AgGSTE5 quando comparadas com os valores da AgGSTE2mut, porém os valores em ambas foram negativos. As distâncias entre DDT e GSH nestes confôrmeros mostram uma proximidade que possibilita interações, sendo na AgGSTE5 a menor distância (2.91 angstroms) observada nos complexos simulados com o DDT. Em todas as três enzimas observou-se uma aproximação entre o grupo triclorometil do DDT com GSH, evidenciando a capacidade destas enzimas de se ligarem a este inseticida.

Para o carbaril, a enzima com menor energia de ligação foi a AgGSTe5, seguida da AgGSTe2mut e AgGSTe2. No entanto foi a AgGSTe2mut que mostrou a conformação com a menor distância entre os ligantes. A proximidade do carbaril com a glutathione sugere que os três sistemas podem formar o GSH-conjugado com este inseticida.

Nos complexos simulados com a cipermetrina foram observados os menores valores de energia, exceto na AgGSTe2mut cujo menor score de energia foi para a simulação com o DDT. Nas conformações da AgGSTe2 e AgGSTe2mut as distâncias de ligação entre a cipermetrina e o GSH foram de 3,39 e 2,74 angstroms respectivamente, evidenciando um potencial destas enzimas para metabolizar a cipermetrina. Na AgGSTe5, a distância entre os ligantes foi de 4,81 angstroms, o que indica que apesar da enzima ter afinidade de ligação com o inseticida a probabilidade de ocorrer a reação de conjugação à glutathione é baixa.

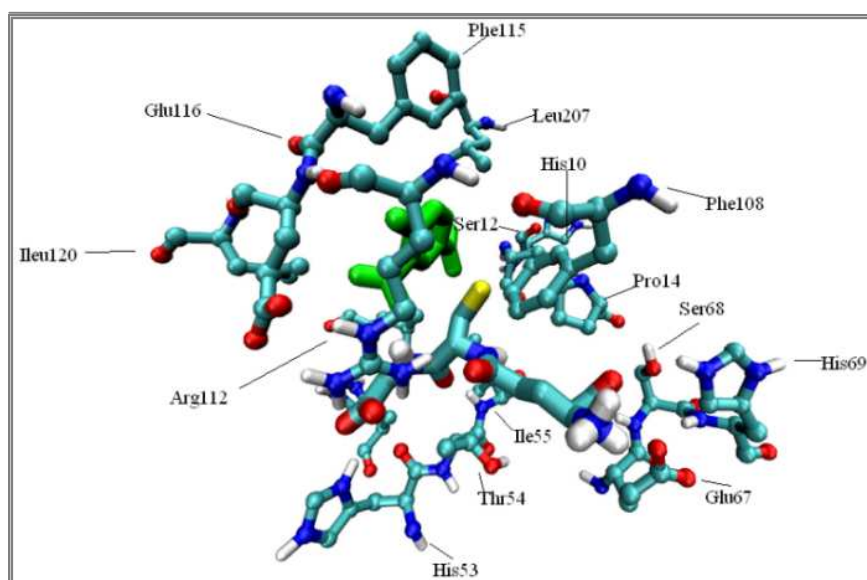
O malation, apesar de ter demonstrado valores negativos ao ser complexado com as enzimas foi o composto que mostrou os maiores valores de energia para os três sistemas. Além disso, em nenhum dos conformêros foi observado uma proximidade razoável do GSH (AgGSTe2 = 8,10 angstroms; AgGSTe2mut = 9,57 angstroms; AgGSTe5 = 5,26 angstroms), o que afasta a possibilidade de uma destas enzimas ser capaz de metabolizar o malation.

Figura 35 - Representação da melhor conformação do complexo AgGSTE2-CDNB. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o CDNB



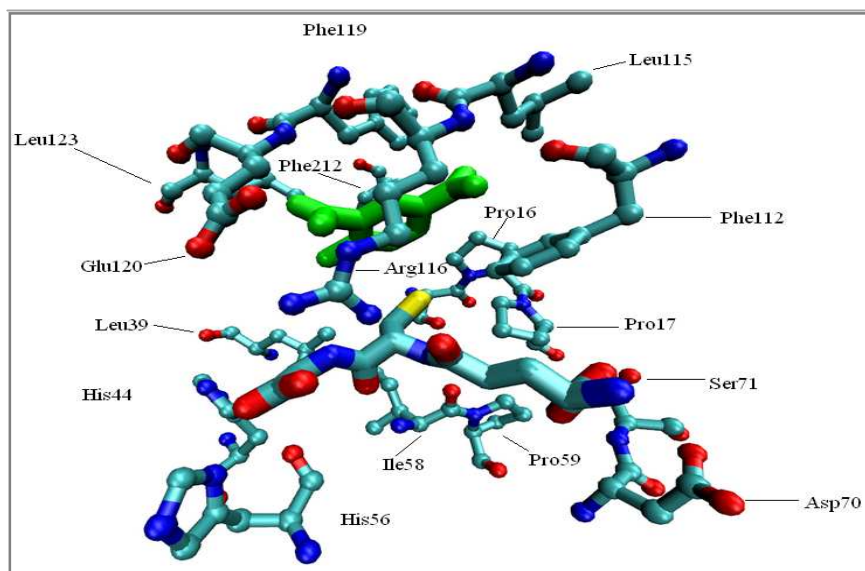
Fonte: Ilustração produzida pelo autor da pesquisa

Figura 36 - Representação da melhor conformação do complexo AgGSTE2mut-CDNB. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o CDNB.



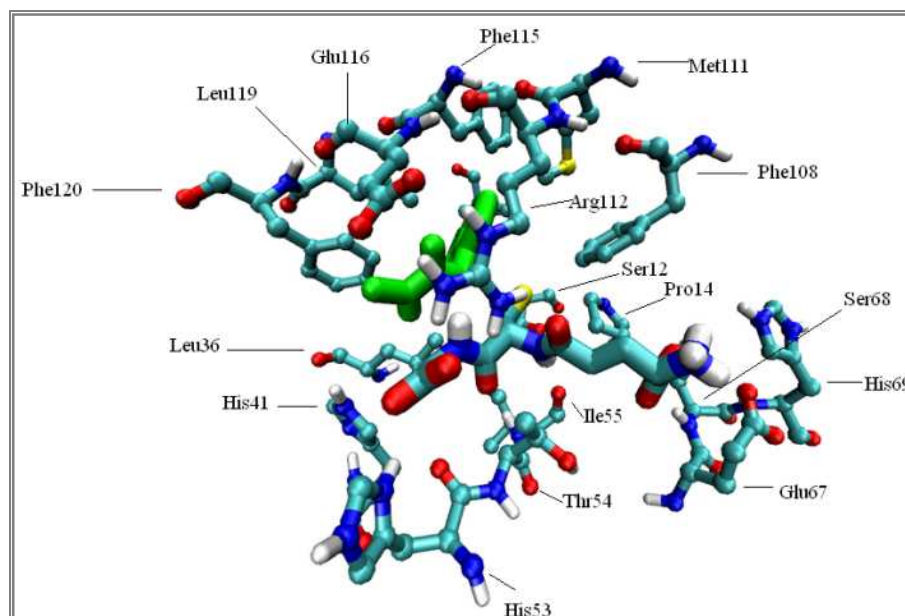
Fonte: Ilustração produzida pelo autor da pesquisa

Figura 37 - Representação da melhor conformação do complexo AgGSTE5-CDNB. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o CDNB.



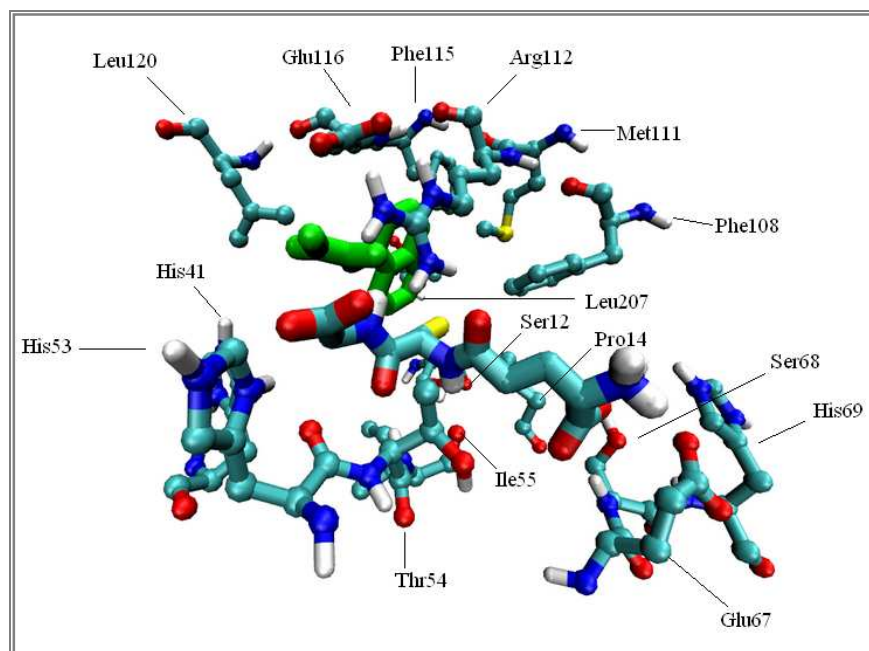
Fonte: Ilustração produzida pelo autor da pesquisa

Figura 38 - Representação da melhor conformação do complexo AgGSTE2-carbaril. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o carbaril.



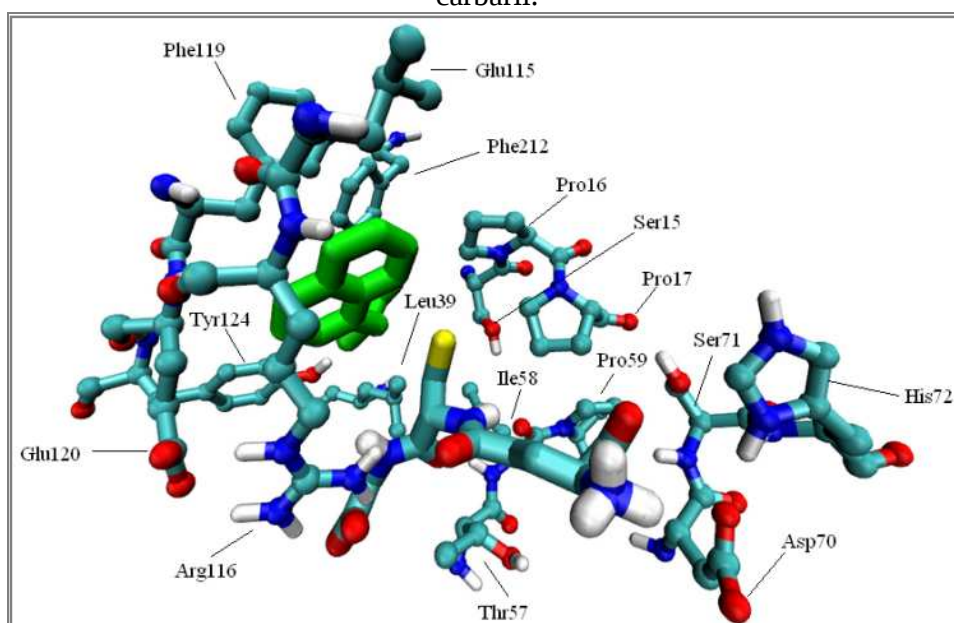
Fonte: Ilustração produzida pelo autor da pesquisa

Figura 39 - Representação da melhor conformação do complexo AgGSTE2mut-carbaril. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o carbaril.



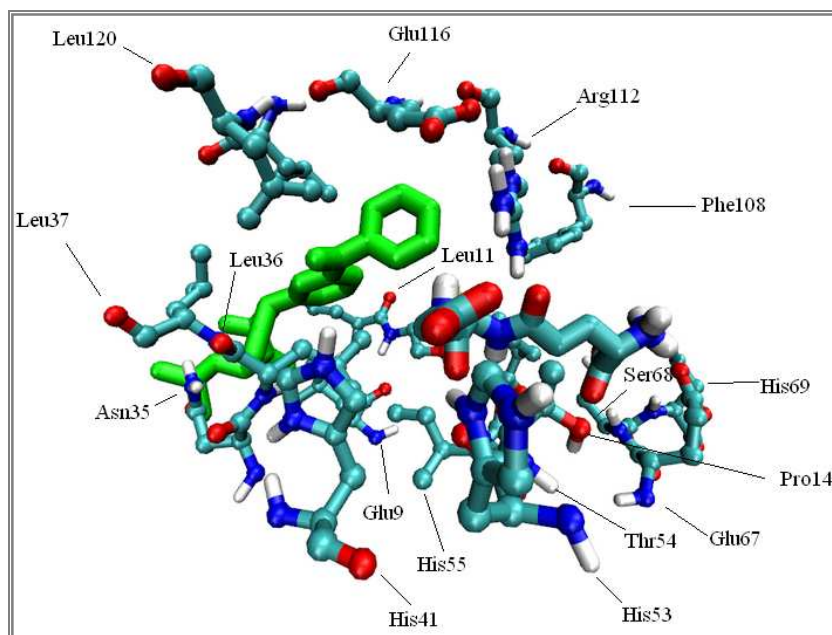
Fonte: Ilustração produzida pelo autor da pesquisa.

Figura 40 - Representação da melhor conformação do complexo AgGSTE5-carbaril. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o carbaril.



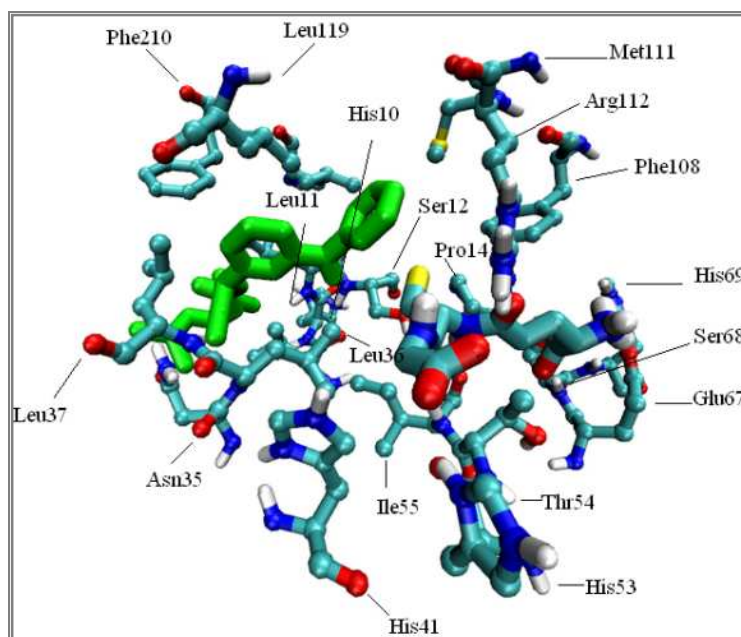
Fonte: Ilustração produzida pelo autor da pesquisa.

Figura 41 - Representação da melhor conformação do complexo AgGSTE2-cipermetrina. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde a cipermetrina.



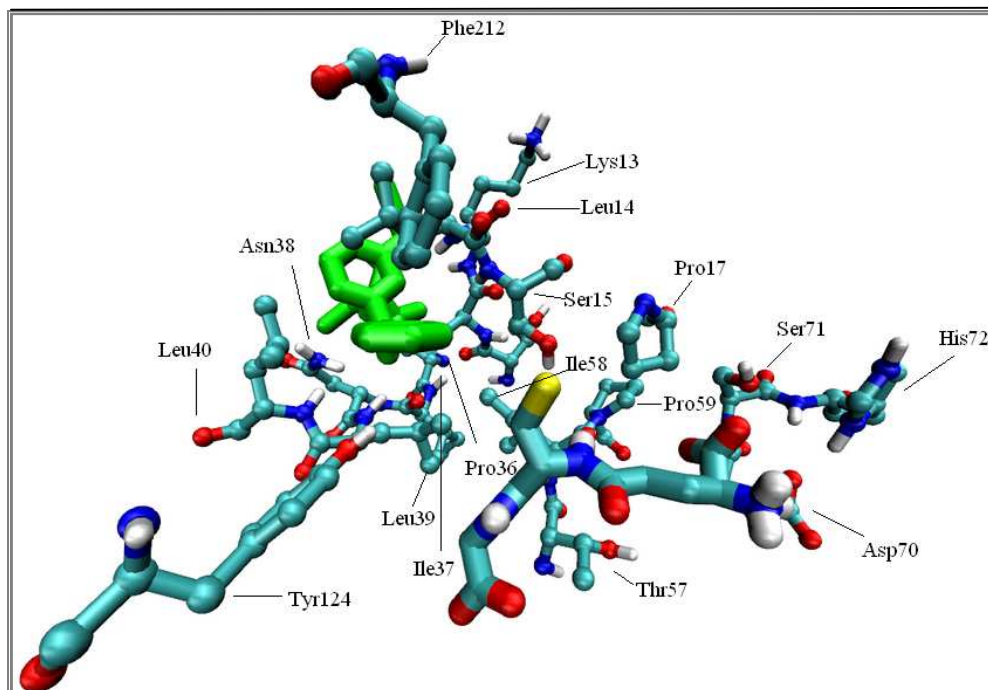
Fonte: Ilustração produzida pelo autor da pesquisa

Figura 42 - Representação da melhor conformação do complexo AgGSTE2mut-cipermetrina. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde a cipermetrina.



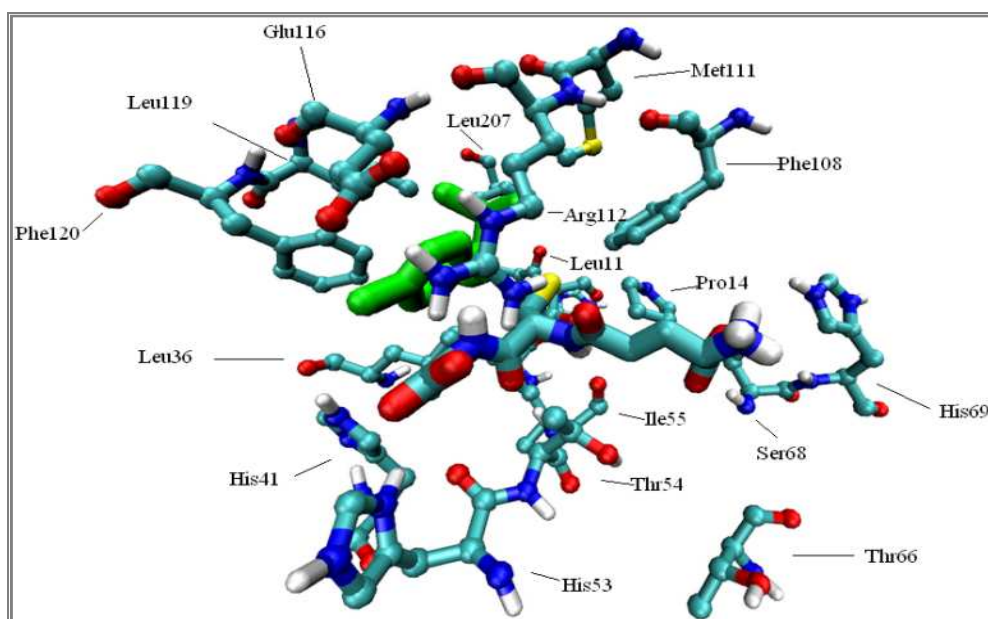
Fonte: Ilustração produzida pelo autor da pesquisa

Figura 43 - Representação da melhor conformação do complexo AgGSTE5-cipermetrina. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde a cipermetrina.



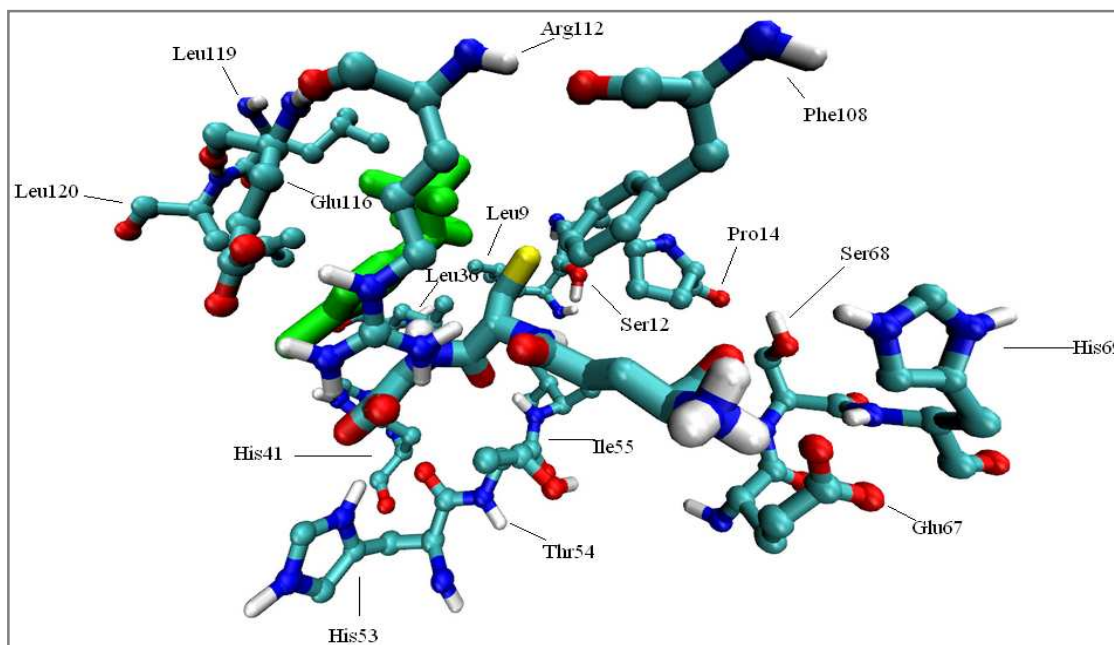
Fonte: Ilustração produzida pelo autor da pesquisa

Figura 44 - Representação da melhor conformação do complexo AgGSTE2-DDT. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o DDT.



Fonte: Ilustração produzida pelo autor da pesquisa

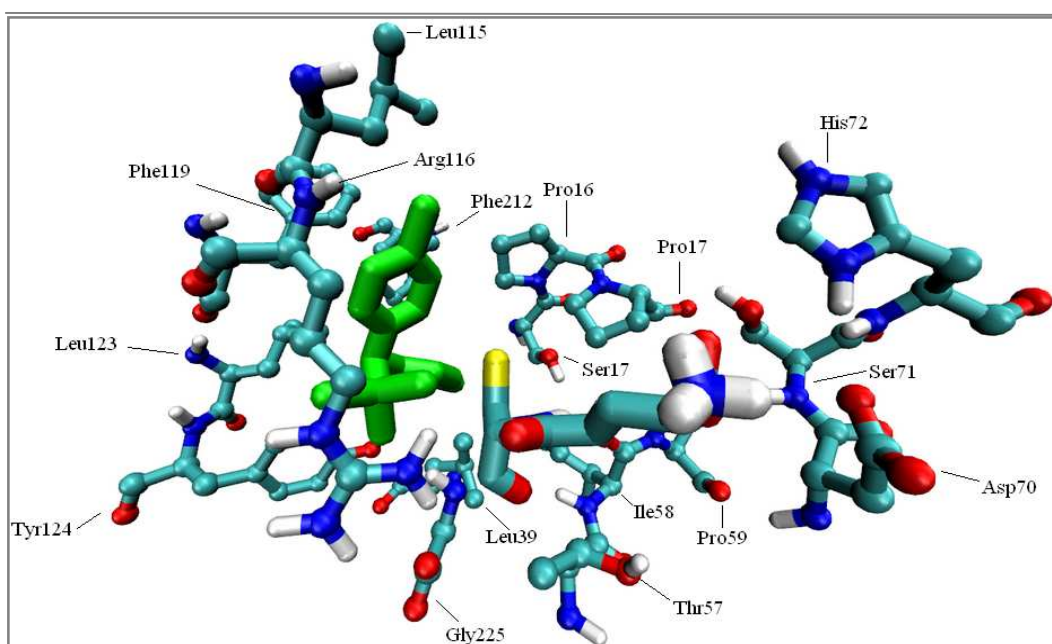
Figura 45 - Representação da melhor conformação do complexo AgGSTE2mut-DDT. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o



DDT.

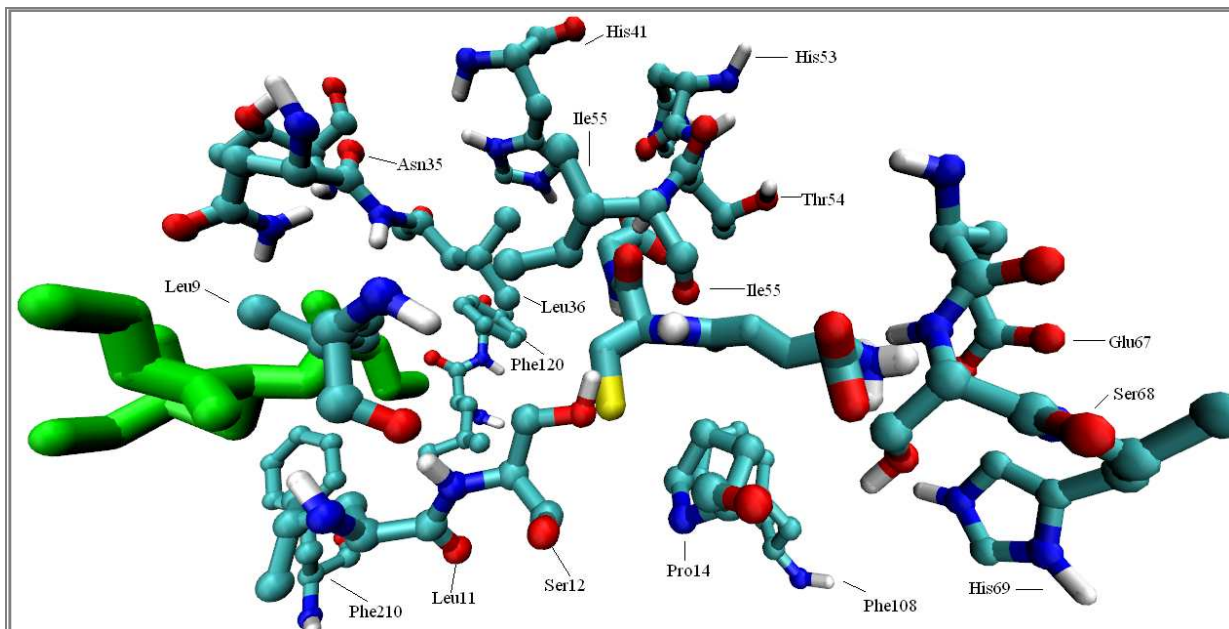
Fonte: Ilustração produzida pelo autor da pesquisa.

Figura 46 - Representação da melhor conformação do complexo AgGSTE5-DDT. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o DDT.



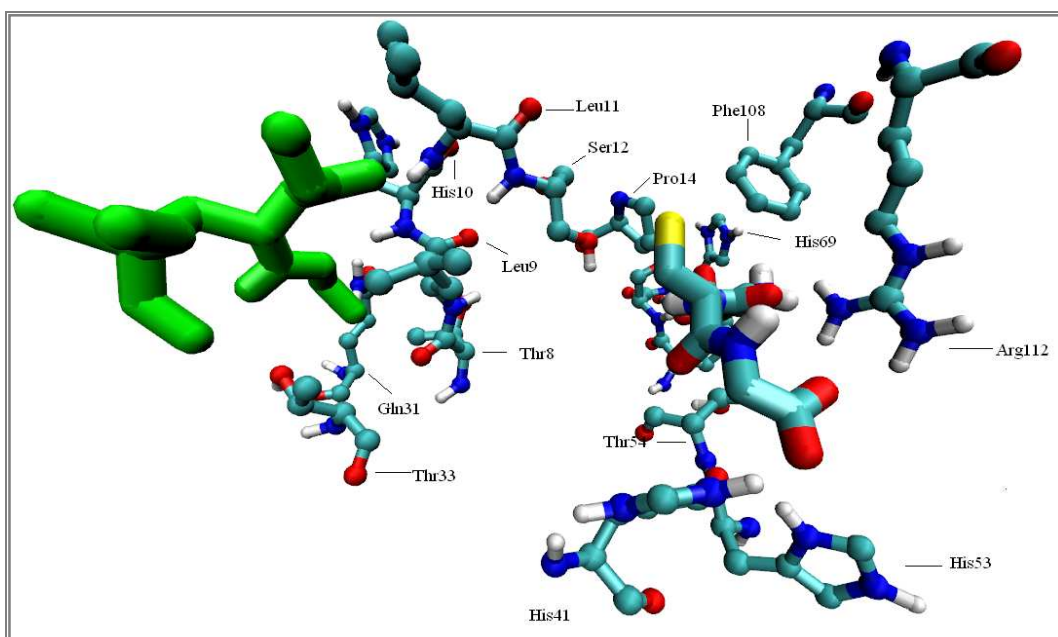
Fonte: Ilustração produzida pelo autor da pesquisa.

Figura 47 - Representação da melhor conformação do complexo AgGSTE2-malation. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o malation.



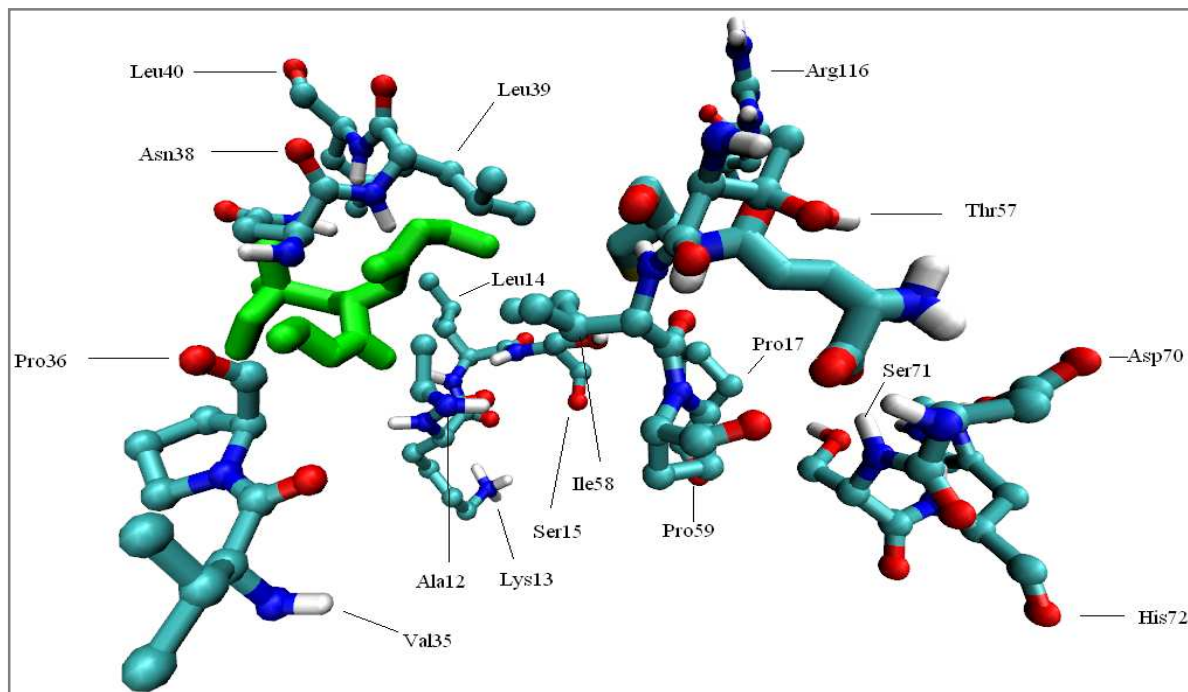
Fonte: Ilustração produzida pelo autor da pesquisa.

Figura 48 - Representação da melhor conformação do complexo AgGSTE2mut-malation. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o malation. O GSH



Fonte: Ilustração produzida pelo autor da pesquisa

Figura 49 - Representação da melhor conformação do complexo AgGSTE5-malation. Resíduos estão representados em bastões e esferas. O GSH está representado em bastões. Em verde o malation.



Fonte: Ilustração produzida pelo autor da pesquisa.

5 DISCUSSÃO

No presente trabalho, foi construído um modelo tridimensional de uma Glutathione S-transferase da classe epsilon (AgGSTE5) de *Anopheles gambiae* para realizar os estudos de dinâmica molecular entre duas proteínas desta classe, e este estudo representa a primeira investigação sobre a dinâmica molecular em GSTs nesta espécie. O modelo desenvolvido foi o primeiro modelo teórico-computacional elaborado e validado para a classe epsilon das glutathione S-transferases e encontra-se disponível em um repositório de acesso público. Os modelos por homologia vêm se tornando uma ferramenta valiosa no estudo de proteínas de desintoxicação celular de mosquitos, elucidando alguns mecanismos destas enzimas na resistência a inseticidas. Em *Anopheles minimus* três enzimas da superfamília P450, com potencial para metabolização de piretróides, foram usadas para construir modelos, os quais foram obtidos pela técnica de modelagem comparativa com múltiplos *templates* (LERTKIATMONGKOL *et al.*, 2011) e este estudo demonstrou diferenças na topologia dos sítios ativos entre as três enzimas. Em *An. gambiae*, duas enzimas relacionadas com a metabolização do DDT, a CYP6Z1 e CYP6Z1, foram modeladas a partir de uma estrutura-molde da enzima humana CYP3A4 (CHIU *et al.*, 2008). O *template* usado mostrou identidades de 28,1% (quando comparada a CYP6Z1) e 28,7% (quando comparada a CYPZ2), menor que a identidade observada entre a AgGSTE5 e AgGSTE2, e os gráficos de Ramachandran obtidos por CHIU e seus colaboradores mostraram uma maior quantidade de resíduos em regiões energeticamente proibidas (2,3% e 2,5% respectivamente) que os observados em nosso modelo. Em ambos os estudos, foram feitas comparações entre as estruturas homólogas e diferenças na topologia dos seus sítios ativos, demonstrando que as diferenças na área e composição destes sítios resultavam em diferenças na seletividade dos substratos. Da mesma forma os resultados obtidos no presente estudo sugerem que apesar das

proteínas homólogas catalisarem as mesmas reações, evolutivamente elas divergiram de modo a metabolizarem compostos diferentes. Esses dados fornecem evidências para a diversificação funcional dos membros de GSTs da classe epsilon e revelam o seu importante papel na adaptação de *An. gambiae* aos diferentes nichos. Apesar da importância das GSTs na resistência a inseticidas nos mosquitos, muitas destas proteínas ainda não possuem estrutura tridimensional elucidada. Modelos para GSTs já foram propostos pela técnica da modelagem por homologia com sucesso em algumas espécies de insetos, como *Lygus lineolaris*, *Helicoverpa armigera*, *Manduca sexta* (HABEEB *et al.*, 2011) e para *Drosophila melanogaster* (LOW *et al.*, 2007). Este último trabalho propôs um modelo para DmGSTD1 (Glutathione S-transferase 1 da classe delta) e associou a capacidade desta enzima em metabolizar o DDT com uma mutação G171K situada em um sítio sob forte seleção positiva ($dNS/dS > 1.0$). Resultados similares foram encontrados neste estudo, uma vez que a estrutura da isoforma mutante (AgGSTE2mut) foi a que mostrou melhor conformação de ligação com o DDT. Isto sugere que mutações pontuais nos genes das GSTs podem estar sendo selecionadas para aumentar a capacidade de metabolização do DDT por estas enzimas. Este composto foi utilizado intensivamente no controle do mosquito *An. gambiae* na África, e provavelmente foi o agente seletivo das mutações na isoforma AgGSTE2mut. Esta informação é de grande relevância para o controle deste vetor, pois a detecção de haplótipos (alelos) com os polimorfismos da AgGSTE2mut indicam que a população tem indivíduos resistentes ao DDT. No mosquito *Culex quinquefasciatus*, principal vetor da filariose e do vírus do Nilo Ocidental, um modelo *in silico* também foi elaborado para AChE1 (Acetilcolinesterase 1) utilizando uma AChE de *An. gambiae* com 79,5 % de identidade como molde (LI *et al.*, 2004). Em *An. gambiae* uma estrutura foi construída para AgAChE e seu modelo refinado através de simulações múltiplas de dinâmica molecular (PANG, 2006). Em *An. plumbeus*, as enzimas ApGSTE2 e ApGSTE2B também foram modeladas por AYRES *et al.*,

(2011) cujas identidades em relação ao *template* foi de 77% e 86%, respectivamente. Assim como nas enzimas aqui estudadas, estes genes divergiram via duplicação gênica, e, portanto verificam-se valores de similaridade bem evidentes entre as sequências primárias destas enzimas. Entretanto, naquele trabalho, os modelos não foram validados. A validação e otimização dos modelos são etapas indispensáveis à modelagem molecular. Para saber se os modelos têm qualidade estrutural e são termodinamicamente estáveis, é necessário acessar a geometria da proteína e as conformações de seus resíduos e átomos. Como a similaridade entre as estruturas-alvo e o *template* foram muito altas (>70%), é provável que a estrutura obtida para os modelos teóricos se aproxime mais da estrutura dos *templates* do que da estrutura real das enzimas-alvo. Neste âmbito técnicas de otimização de estruturas, como a minimização de energia (EM) e a dinâmica molecular (DM) são extremamente providenciais para a obtenção dos melhores confôrmeros.

A presença de diversidade nas conformações estruturais de proteínas modula a divergência de sequências das mesmas (JURITZ *et al.*, 2012) e, portanto deve ser investigada. A elucidação das estruturas de enzimas de desintoxicação, em particular aquelas que estão diretamente envolvidas na metabolização de inseticidas químicos, representa novas oportunidades para a investigação de possíveis inibidores da atividade destas enzimas e consequentemente novas forma de controle de insetos vetores. A AgGSTe2, cuja estrutura cristalográfica foi elucidada (WANG *et al.*, 2008) é uma forte candidata a alvo para desenvolvimento de inibidores químicos, sintéticos ou naturais, assim como outras GSTs, pois estas são alvos em potencial para o uso de RNAs de interferência para controle de populações de *An. gambiae* resistentes ao DDT. Do mesmo modo, sua homóloga AgGSTe5 mostra-se um alvo em potencial para o desenvolvimento racional de inibidores, uma vez que no presente trabalho esta enzima teve sua estrutura elucidada e há um campo aberto a investigação sobre a interação

com compostos químicos inseticidas. A caracterização e comparação estrutural de enzimas detoxificadoras podem trazer *insights* importantes para a compreensão dos papéis que estas proteínas desempenham ao metabolizar diversos compostos. Diferenças nos sítios de ligação entre proteínas homólogas podem ser o primeiro passo para a compreensão dos mecanismos de resistência a inseticidas químicos nos mosquitos.

Estudos de genética populacional também tem se mostrado extremamente úteis para a compreensão dos processos evolutivos recentes nas espécies de mosquitos, como expansão e introdução recente de populações naturais (MAIA *et al*, 2009) de mosquitos vetores, assim como detectar mutações e seleções em genes associado à resistência a inseticidas (RAYMOND *et al.*, 2001). Os genes da AgGSTE2 e AgGSTE5 em nossas análises apresentaram valores de diversidade muito próximos ($Hd=0,9713$ para AgGSTE2 e $Hd=0,9861$ para AgGSTE5) e elevados, revelando que estes genes possuem variação genética suficiente para detectar recentes eventos de evolução. A modulação da promiscuidade funcional das enzimas vem sendo associada à ocorrência de mutações neutras, resultando em seleção balanceadora (BLOOM *et al.*, 2007, TOKURIKI; TAWFIK, 2009). A capacidade de uma enzima rapidamente adotar novas funções em um mesmo *fold*, ou adotar um *fold* completamente diferente é denominada evolutibilidade. A evolutibilidade pode ser o ponto inicial para se entender a divergência funcional entre enzimas de uma mesma família (TOKURIKI e TAWFIK, 2009). Nos testes de neutralidade foi observado que na AgGSTE2 todas as regiões da proteína estão sob forte seleção purificadora (TAJIMA, 1989), ou seja, a seleção desfavorece as mutações não sinônimas (que alteram a cadeia peptídica), indicando que os polimorfismos presentes no gene que codifica esta enzima não favorecem a diversidade fenotípica. O gene AgGSTE5 apresentou alguns sítios sob seleção positiva neste trabalho, no qual diferentes populações de *An. gambiae* foram investigadas. Estes dados corroboram os resultados encontrados por (AYRES, *et al*, 2011), que observaram sítios no

mesmo gene sob efeito de seleção positiva entre diferentes espécies de *Anopheles*. Este fato sugere que tanto nas populações como entre as espécies, este gene confere alguma vantagem adaptativa e deve ter influenciado no processo de especiação do gênero *Anopheles*, conferindo a capacidade das diferentes espécies colonizarem nichos ecológicos distintos, viabilizando a metabolização dos compostos tóxicos presentes naqueles ambientes pelos mosquitos.

Os volumes do sítio catalítico variaram entre as três enzimas. O sítio da AgGSTe2mut tem um volume acima de 18% maior que sua isoforma selvagem, a AgGSTe2. Outro fato a ser observado é que as mutações na AgGSTe2mut foram responsáveis pelo aumento do sítio de ligação. A AgGSTe5 apresentou o maior sítio catalítico, com um volume acima de 21% em relação à AgGSTe2mut e acima de 44% comparada a AgGSTe2. O aumento no volume do sítio ativo é observado em enzimas promíscuas, pois permite uma melhor acomodação dos substratos (EKROOS e SJÖGREN, 2006). Outro fato a ser observado é que as mutações na AgGSTe2mut são responsáveis pelo aumento do sítio de ligação.

Os resultados da dinâmica molecular mostraram diferenças nas enzimas entre si e quando ligadas ao GSH. Ao analisar os valores do RMSD observa-se que sem o ligante a AgGSTe5 é a enzima que mais diverge da estrutura inicial, e apresenta um perfil com três platôs que sugerem a presença de subestados conformacionais distintos. Contudo, a distribuição do RMSD neste sistema mostra um perfil com duas curvas em estado de transição entre estas. Isto indica que o segundo platô e o terceiro compõem um único subestado conformacional. As duas isoformas da AgGSTe2 (selvagem e mutante) mostram valores de RMSD muito próximos quando estão sem a glutatona, com apenas um platô sugerindo um único subestado conformacional. No entanto, os valores de RMSD para estas enzimas apresentaram perfis de distribuição diferentes, onde se observa uma curva quase harmônica na AgGSTe2 e na AgGSTe2mut vê-se uma curva assimétrica negativa (à esquerda), ou seja, mostra uma tendência maior nesta enzima de divergir

da estrutura original do que sua isoforma. Nos valores de RMSD com o ligante a AgGSTe5 apresenta uma curva crescente, na ordem de 0.1 nm. O gráfico de distribuição mostra uma curva simétrica indicando a existência de um subestado conformacional, mostrando que ao ligar-se a glutathione esta enzima apresenta uma maior homogeneidade estrutural, o que sugere que o ligante GSH é importante para a estabilidade desta proteína. Na AgGSTe2 ligada ao GSH os valores de RMSD não tiveram modificações significativas e sua curva de distribuição é quase harmônica, demonstrando não haver mudanças significativas na sua estrutura esteja ela ligada ou não ao tripeptídeo, indicando que mesmo ligada ao GSH esta enzima não diverge demasiadamente de seu estado nativo. Esta enzima, portanto, foi a que demonstrou maior homogeneidade conformacional neste estudo. A isoforma mutante, AgGSTe2mut, quando ligada ao GSH apresenta valores mais altos de RMSD, com dois platôs na curva. No gráfico de distribuição observam-se nitidamente duas curvas gaussianas indicando a existência de dois subestados conformacionais. Neste contexto pode-se afirmar que as mutações na AgGSTe2mut repercutem substancialmente na dinâmica desta enzima, principalmente no modo como ela interage com o ligante. A glutathione parece também exercer um papel importante na conformação da AgGSTe5. Mudanças conformacionais são características típicas de enzimas que apresentam promiscuidade funcional, sem, contudo perder a estabilidade estrutural ou perder a sua função inicial (TOKURIKI e TAWFIK, 2009). As mudanças de conformação observadas na AgGSTe5 e AgGSTe2mut são provavelmente resultado da evolutibilidade destas enzimas, e podem estar atreladas às mutações pontuais, no caso da AgGSTe2mut e à seleção positiva que atuou na evolução divergente da AgGSTe5.

As regiões de maior flutuação são mostradas nos gráficos de RMSF. Observou-se que nos sistemas simulados as regiões de maior flutuação são congruentes, tanto quando analisadas com ou sem o ligante. No entanto, nas enzimas sem o GSH observou-se uma elevada flutuação nas

regiões de início e fim de cadeia, o que significa que boa parte das flutuações pode ser atribuída a uma acomodação entre as duas subunidades. O RMSF das enzimas com o ligante mostrou valores menores nas regiões extremas das cadeias, porém mostraram um aumento na flutuação das demais regiões das enzimas. Vale salientar que a AgGSTE2mut apresentou os maiores valores de RMSF, sendo que uma das regiões de maior flutuação (resíduos 110-140) é onde situam-se as duas mutações (resíduos 114 e 120).

O aumento da evolutibilidade nas proteínas no *An. gambiae* pode estar associado à desordem global ou local e à diminuição da compactação da proteína (FERRADA e WAGNER, 2008). O raio de giro é uma medida que dá uma idéia de processos de compactação e expansão de macromoléculas (LOBANOV *et al.*, 2008). O raio de giro calculado para as enzimas sem o GSH foi maior na AgGSTE5, sendo esta enzima menos compacta que as outras duas. A curva oscila até os 40 ns e isto mostra que esta enzima passa por processos de compactação e expansão até atingir um volume constante. A AgGSTE2 e AgGSTE2mut têm curvas ligeiramente decrescentes, demonstrando que estas enzimas aumentam sua compactação, sendo a isoforma mutante um pouco mais compacta. Nos sistemas com o GSH observou-se que a AgGSTE5 teve um pequeno aumento no seu volume e que os valores do raio de giro oscilaram menos. Na AgGSTE2 os valores de raio de giro também foram ligeiramente maiores, apesar da curva continuar decrescente. A maior mudança foi observada na AgGSTE2mut, que apresentou valores maiores e uma curva crescente a partir dos 15 ns, chegando a atingir os maiores valores de raio de giro. No final, a curva tende a convergir com os valores da AgGSTE5. De modo geral, a presença da glutathione implica em um aumento no diâmetro das proteínas e, portanto uma perda da compactação. Porém o perfil da curva da AgGSTE2mut sugere que grande parte desta expansão estrutural esteja relacionada às mutações encontradas nesta isoforma.

A superfície de acessibilidade ao solvente (SASA) é a área superficial da proteína que fica exposta ao solvente e seu aumento tem sido relacionada com um aumento da taxa evolutiva (FRANZOSA e XIA, 2008; RAMSEY *et al.*, 2011). Em proteínas de *Escherichia coli* e *Salmonella enterica* foi constatado um decréscimo da seleção purificadora em resposta a um acréscimo da SASA (BUSTAMANTE *et al.*, 2000). Contudo, os resultados da SASA para as enzimas estudadas não corroboram com estes trabalhos. Nas enzimas sem o ligante observou-se um decréscimo da SASA ao longo da simulação, sendo os valores da AgGSTE5, que tem sítios de seleção positiva, os menores aqui obtidos, o que corrobora com a suposta correlação existente entre taxa evolutiva e SASA. O decréscimo do SASA apresentado nas simulação da AgGSTE5 provavelmente está associado a um período de equilibração devido a estrutura se tratar de um modelo comparativo. Os valores da AgGSTE2 e AgGSTE2mut foram muito próximos, e o perfil de suas curvas foram quase idênticos. Quando simuladas com a glutatona as três enzimas exibem valores de SASA um pouco maiores, apesar de também mostrarem um decréscimo ao longo da simulação. Outro fato observado é que sem o GSH os valores da AgGSTE5 são sempre menores que as demais, enquanto que quando ligadas ao GSH os valores das três enzimas convergem ao fim da simulação. Os resultados comprovam que as mutações da AgGSTE2mut não interferem significativamente na SASA.

Proteínas robustas, isto é, que apresentam robustez, tende a apresentar maior evolutibilidade (RORIK; WAGNER, 2011). A robustez de proteínas pode ser definida como a tendência de um sistema manter a sua funcionalidade sob perturbação. Esta característica é particularmente interessante quando a proteína está sob perturbação mutacional, pois são mais relevantes para a evolutibilidade, e causa resposta à seleção positiva (WAGNER; ALTENBERG, 1996; PIGLIUCCI, 2008; WAGNER, 2008). O teor da estrutura secundária é alterado quando há desordem estrutural. Alterações na estrutura secundária e desordem estrutural são linearmente

proporcionais a mudanças nas sequências (SCHAEFER *et al.*, 2010). Nos resultados da estrutura secundária observa-se que não há mudanças significantes nos sistemas simulados, e isto é uma evidência de que estas enzimas são dotadas de uma razoável robustez, especialmente a AgGSTe2mut e a AgGSTe5.

A análise da dinâmica essencial mostrou que as principais moções nas proteínas diferem entre si. No primeiro autovetor é notória a diferença de moções de larga amplitude nas proteínas. Uma das maiores diferenças é observada na AgGSTe2mut, que quando ligada ao GSH têm o maior autovalor. Nesta enzima, as movimentações de grande amplitude parecem estar associadas em grande parte às mutações, assim como parece haver uma contribuição significativa do ligante para estes movimentos. O aumento da amplitude em face ao GSH é observado também na AgGSTe2, porém com autovalores bem menores. A AgGSTe5 teve o maior autovalor sem o GSH, e teve seu primeiro autovetor diminuído com a presença do ligante. Estes resultados sugerem que a AgGSTe2 e AgGSTe2mut tenham sido adaptadas para realizar a reação de conjugação com mais eficiência que a AgGSTe5, que por sua vez é mais propensa a mutações não-sinônimas. As regiões que contribuem para estes movimentos também diferiram, inclusive entre as isoformas da AgGSTe2. Isto sugere que as mutações da AgGSTe2 produzem fortes interações alostéricas com as demais regiões da enzima.

Pela análise dos *clusters* foi possível constatar que a AgGSTe2mut apresentou uma maior heterogeneidade conformacional, uma vez que mostrou *clusters* mais populados que as outras enzimas. A AgGSTe5, apesar de ter mostrado um maior número de *clusters* em relação as demais, muitos destes estão representados por poucas estruturas, e alguns compostos por uma única conformação, que provavelmente representam estruturas de transição entre os subestados conformacionais.

Os resultados do *docking* mostraram que as três enzimas têm afinidade com compostos de natureza diferente. De fato, isto representa uma evidência *in silico* de que estas enzimas apresentam uma notável promiscuidade funcional, resultante de uma multi-especificidade ao substrato. Apesar da AgGSTe2mut ter apresentado os menores valores para cinco dos sete compostos submetidos ao cálculo de *docking*, os valores não diferiram muito. Quando comparadas as duas isoformas observou-se que para seis dos sete compostos testados a enzima mutante teve energias ligeiramente mais favoráveis que a forma selvagem. A explicação mais plausível para este resultado está no fato da AgGSTe2mut ter um sítio catalítico de maior volume, resultante das mutações nesta enzima, que provavelmente permitem uma melhor acomodação dos compostos.

A *evolutibilidade* apresentada por estas enzimas, principalmente a AgGSTe2mut, pode representar um importante aspecto na capacidade do *An. gambiae* ter populações resistentes à inseticidas químicos. Este é um conceito recente (TOKUKURI; TAWFIK, 2009) e deve ser levado em consideração nos estudos futuros da evolução molecular de superfamília de enzimas. O uso de inseticidas químico nesta espécie precisa ser repensado e reavaliado como modo de controle. Uma perspectiva futura pode estar no potencial de desenvolvimento de inibidores específicos para estas enzimas, na tentativa de diminuir a resposta aos inseticidas utilizados, especialmente ao DDT. Outro aspecto que evidencia o potencial das GSTs da classe épsilon como alvos para inibição é o fato desta classe de enzimas ser específica de artrópodes, o que possibilita o futuro desenvolvimento de compostos inibidores que não afetem outras espécies, como mamíferos por exemplo. Entender os mecanismos de evolução e adaptação destas enzimas, e detalhes de sua dinâmica e funcionamento é indispensável quando se planeja um controle racional e integrado de uma espécie vetora. Outra possível aplicação é utilizar estas enzimas

como indicadores de populações resistentes e refratárias a diversos inseticidas, e assim escolher o melhor tipo de composto a ser utilizado para cada população.

6 CONCLUSÕES

- Os genes *AgGSTE2* e *AgGSte5* estão evoluindo sob diferentes pressões seletivas.
- O modelo tridimensional da *AgGSTE5* é válido.
- As enzimas diferem em sua dinâmica.
- Mutações na *AgGSTE2* resultam em alteração da sua dinâmica. Estas mudanças geram efeito direto no modo como esta enzima se liga ao DDT.
- As enzimas apresentam multiespecificidade aos substratos, sendo portanto enzimas promíscuas.
- As três enzimas apresentam potencial de ligação a inseticidas químicos e provavelmente estão envolvidas em mecanismos de resistência no *An. gambiae*.

REFERÊNCIAS

AGARWAL, P. K. (2006). Enzymes: An integrated view of structure, dynamics and function. **Microbial cell factories**. v.5, n.2, 2006.

ALLEN, H.H; TILDESLEY, D. J. **Computer Simulation of Liquids**. Oxford: Clarendon Press, 1989. 382 p.

ALMEIDA, F. V. *et al.* Persistent toxic substances (PTS) in Brazil. **Quimica nova**. v.30, p.1976-1985, 2007.

ARFKEN, G. The Method of Steepest Descents. §7.4 In:_____. **Mathematical Methods for Physicists**. 3rd ed. Orlando - FL: Academic Press, 1985. p. 428-436.

ARNOLD, K. *et al.* (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. **Bioinformatics**. v.22, 195-201, 2006.

AYRES, C.F.J. *et al.* (2011). Comparative genomics of the anopheline glutathione S-transferase epsilon cluster. **PloS one**, v.6, n.12, 2011.

BALSERA, Manel. A. *et al.* Principal Component Analysis and Long Time Protein Dynamics. **Journal of Physical Chemistry**. 100:2567-2572, 1996.

BASTA, S.S. *et al.* Iron deficiency anaemia and the productivity of adult males in Indonesia. **American Journal of Clinical Nutrition** v.32, p.916-925, 1979.

BAUDRY J. *et al.* Molecular docking of substrates and inhibitors in the catalytic site of CYP6B1, an insect cytochrome P450 monooxygenase. **Protein Engineering**. vol.16 n.8, 2003.

BEALL, C. *et al.* Isolation of a *Drosophila* gene encoding glutathione S-transferase. *Biochem. Genet.* v.30, p. 515-527, 1992.

BENKERT, P; TOSATTO, S.C.E; SCHOMBURG, D. QMEAN: A comprehensive scoring function for model quality assessment. **Proteins: Structure, Function, and Bioinformatics.** v.71, n.1, p.261-277, 2008.

BERENDSEN, H. J. C. *et al* Interaction models for Water in Relation to Protein Hydration. **The Jerusalem Symposia on Quantum Chemistry and Biochemistry.** v. 14, 1981, pp 331-342.

BERENDSEN, H. J. C; GRIGERA, J. R; STRAATSMA, T. P. J. *Phys. Chem.* 91, Descent. In:_____. **Methods of Theoretical Physics, Part I.** New York: McGraw-Hill, 1953. p. 434-443.

BERENDSEN, H. J. C; GRIGERA, J. R; STRAATSMA, T. P. (1987).The missing term in effective pair potentials. **J. Phys. Chem.**, v.91, 1987.

BERENDSEN, H.J.C. *et al.* Interaction models for water in relation to protein hydration. In: **Intermolecular forces.** (ed. B. Pullman). Dordrecht, The Netherlands: Reidel, 1981. p. 331-342.

BERENDSEN, H.J.C. *et al.* Interaction models for water in relation to protein hydration. In: **Intermolecular forces.** (ed. B. Pullman). Dordrecht, The Netherlands: Reidel, 1981. p. 331-342.

BLOOM, J. D. *et al.* Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution. **Biology direct.** v.2, n.17, 2007.

BOARD, P.G. *et al.* (2000). Identification, characterization, and crystal structure of the Omega class glutathione transferases. **J. Biol. Che.** 275(32), 24798–806, 2000.

BOGWITZ, M. The genetics resistance to lufenuron in *Drosophila melanogaster*. Ph.D. Thesis, University of Melbourne, Melbourne,Australia, 2005.

BRAGA, I A, VALLE, D. *Aedes aegypti*: Inseticidas, mecanismos de ação e resistência. **Epidemiologia e serviços de saúde**. V.16, p.279–293,2007.

BRASIL. Ministério da Saúde. Fundação Oswaldo Cruz. **Metodologia para qualificação de atividades de enzimas relacionados com a resistência a inseticidas em aedes aegypti**. Brasília: Ministério da Saúde, 2006. (Série A. Normas e Manuais Técnicos).

BREMEN, J. The ears of hippopotamus manifestations, determinants, and estimates of the malaria burden. **Am. Trop. J. Med. Hyg.** 64, 1-11, 2001.

BROWN, A.W.A. Insecticide resistance in mosquitoes: a pragmatic review. **J. Am. Mosq. Control Assoc.** v.2, p.123-140, 1986.

BUSTAMANTE, C. D; TOWNSEND, J. P; HARTL, D. L. Solvent accessibility and purifying selection within proteins of *Escherichia coli* and *Salmonella enterica*. **Molecular biology and evolution**. v.17, n.2. 2000.

CASIDA, J.E; QUISTAD, G.B. Golden age of insecticide: Past, present, or Future? **Ann. Rev. Entomol.** v43, p.1-16, 1998.

CHELVANAYAGAM, G; PARKER, M.W; BOARD, P.G. Fly fishing for GSTs: a unified nomenclature for mammalian and insect glutathione s-transferases. **Chem. Biol. Interact.** 133: 256-260, 2001.

CHE-MENDOZA, A; PENILLA, R. P; RODRÍGUEZ, D. A. Insecticide resistance and glutathione S-transferases in mosquitoes. **A review**. v.8, n.8, p.1386–1397, 2009.

CHIANG, F.M; SUN, C.N. Glutathione transferase isozymes of diamondback moth larvae and their role in the degradation of some organophosphorus insecticides. **Pestic. Biochem. Physiol.** v.45, p. 7-14. 1993.

CHIANG, R.A; SALI, A; BABBITT, P.C. Evolutionarily conserved substrate substructures for automated annotation of enzyme superfamilies. **PLOS Comput. Biol.** v.4, 2008.

CHIU, T. *et al.* Comparative molecular modeling of *Anopheles gambiae* CYP6Z1, a mosquito P450 capable of metabolizing DDT. **Proceedings of the National Academy of Sciences of the USA**, v.105, n.26, 2008.

CLARK, A.G. *et al.* Insecticide metabolism by multiple glutathione S-transferases in two strains of the house fly, *Musca domestica* (L.). **Pestic. Biochem. Physiol.** v.25, p.169-175, 1986.

CLARK, A.G; SHAMAAN, N.A. Evidence that DDT-dehydrochlorinase from the house fly is a glutathione S-transferase. **Pestic. Biochem. Physiol.** v.22, p. 249-261, 1984.

COHEN, S. Malaria. **British Medical Bulletin.** v.38, p.115–218, 1982.

CONSOLI, R. A. G. B; LOURENÇO-DE-OLIVEIRA, R. **Principais Mosquitos de Importancia Sanitaria no Brasil.** Rio de Janeiro: Editora Fiocruz, 1994.

DANIELSON, U. H; MANNERVIK, B. Kinetic independence of the subunits of cytosolic glutathione transferase from the rat. **The Biochemical journal.** v.231, n.2, 1985.

DEVOS, D; VALENCIA, A. Practical limits of function prediction. **Proteins** v.41, p. 98–107, 2000.

DING, Y. *et al.* The *Anopheles gambiae* glutathione transferase supergene family: annotation, phylogeny and expression profiles. **BMC genomics.** v. 4, n.1, 2003.

DJADID, N. D. *et al.* Identification, Sequence Analysis, and Comparative Study on GSTe2 Insecticide Resistance Gene in Three Main World Malaria Vectors: *Anopheles stephensi*, *Anopheles culicifacies*, and *Anopheles fluviatilis*. **Journal of Medical Entomology.** 43(6), p.1171-1177, 2006.

DJADID, N.D. *et al.* Genetic analysis of rDNA-ITS2 and RAPD loci in field populations of the malaria vector, *Anopheles stephensi* (Diptera, Culicidae): Implications for the control program in Iran. **Acta Trop.** 97(1), p.65-74, 2006.

EKROOS, M., SJÖGREN, T. (2006). Structural basis for ligand promiscuity in cytochrome P450 3A4. **Proceedings of the National Academy of Sciences of the United States of America**. v.103, n.37, 2006.

ENAYATI A.A; RANSON, H; HEMINGWAY, J. (2005). Insect Glutathione transferases and insecticide resistance. **Review. Insect. Mol. Biol.** 14(1), p.3-8, 2005.

FERRADA, E., WAGNER, A. Protein robustness promotes evolutionary innovations on large evolutionary time-scales. **Proceedings. Biological sciences / The Royal Society**, 275(1643), p.1595–602, 2008.

FISCHER, E. Einfluss der Configuration auf die Wirkung der Enzyme. **Ber. Dt. Chem. Ges.** 27 (3), p. 2985–93, 1894.

FLORES, A.V. *et al.* (2004). Organoclorados: um problema de saúde pública. **Ambiente & Sociedade**. Vol. VII, nº. 2, p 111-125, 2004.

FORATTINI, Oswaldo Paulo. **Culicidologia Medica: Identificação, biologia, epidemiologia**. São Paulo, Editora da Universidade de São Paulo, 1996. v. 2.

FOURNIER, D. *et al.* (1992). Insect glutathione S-transferases. Biochemical characteristics of the major forms from houseflies susceptible and resistant to insecticides. **J.Biol. Chem.** v.267, p.1840–1845, 1992.

FRANZOSA, E; XIA, Y. Chapter 1 - Structural Perspectives on Protein Evolution. **Annual Reports in Computational Chemistry**. v. 4, p. 3–21, 2008.

FROVA, C. (2006). Glutathione transferases in the genomic era: new insights and perspectives. **Biomol Eng.** v.23, n. 4, p. 149-169, 2006.

GERLT, J.A; BABBITT, P.C. Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. **Annu. Rev. Biochem.** v.70, p.209–246, 2001.

GERLT, J.A; BABBITT, P.C; RAYMENT, I. (2005). Divergent evolution in the enolase superfamily: the interplay of mechanism and specificity. **Arch. Biochem. Biophys.** 433(1), p.59–70, 2005.

GLASNER, M.E; GERLT, J.A; BABBITT, P.C. Evolution of enzyme superfamilies. **Curr.Opin. Chem. Biol.** v.10, p.492–497, 2006.

GONZALES, Matthew J; DUGAN, Jonathan. M; SHAFER, Robert W. Synonymous–non-synonymous mutation rates between sequences containing ambiguous nucleotides (Syn-SCAN) . **Bioinformatics applications note**, v.18, n.6, p.886–887, 2002.

GRANT, D.F; DIETZE, E.C; HAMMOCK, B.D. Glutathione S-transferase isozymes in *Aedes aegypti*: purification, characterization, and isozyme specific regulation. **Insect. Biochem.** v.4, p. 421-433, 1991.

GRISHAM, M. Charles; GARRETT, Reginald H. **Biochemistry**. Philadelphia: Saunders College Pub, 1999. p. 426.

GRISWOLD, C.K. (2006). Pleiotropic mutation, modularity and evolvability. **Evol Dev.** v.8, n.1, p.81–93, 2006.

GRISWOLD, K. E. *et al.* The evolution of catalytic efficiency and substratepromiscuity in human theta class 1-1 glutathione transferase, **J.Mol. Biol.** 364, 400-410, 2006.

HABEEB, S. K. M. Comparative Molecular Modeling of Insect Glutathione S-Transferases. **International Journal of Computer Applications**, 14(5), p.16–22, 2011.

HAYES, J. D; FLANAGAN, J. U; JOWSEY, I. R. Glutathione transferases. **Annual Review of Pharmacology and Toxicology.** v. 45, p. 51-88, 2005.

HAYES, J.D; WOLF, C.R. Role of glutathione transferase in drug resistance. In: Sies H, Ketterer B (eds) **Glutathione conjugation: Mechanisms and Biological Significance**. Academic Press, London: Academic Press, 1988. p. 315-355.

HEMINGWAY, J. Mini review: The molecular basis of two contrasting metabolic mechanisms of insecticide resistance. **Insect. Biochem. Mol. Biol.** v. 30, p.1009-1015, 2000.

HEMINGWAY, J; BATES I. Malaria: Past problems and future prospects. **EMBO Reports**, june (4). (suppl 1), 2003.

HEMINGWAY, J; HAWKES, N. J; McCARROLL, Ranson, H. (2004). The molecular basis of insecticide resistance in mosquitoes. **Insect Biochem Mol Biol.** v. 34, p. 653–665, 2004.

HEMINGWAY, J; MIYAMOTO, J; HERATH, P.R.J. A possible novel link between organophosphorus and DDT insecticide resistance genes in *Anopheles*: supporting evidence from fenitrothion metabolism studies. **Pestic. Biochem. Physiol.** v.39, p.49-56, 1991.

HEMINGWAY, J; RANSON, H. (2000). Insecticide resistance in insect vectors of human disease. **Annu Rev Entomol.** v.45, p.371–391, 2000.

HESS, B. LINC: A linear constraint solver for molecular simulations. **J Comput Chem**, v.18, p.1463, 1997.

HOLT, R.A. The genome sequence of the malaria mosquito *Anopheles gambiae*. **Science** 298(5591): 129-149, 2002.

HOOVER, William. G. Canonical dynamics: Equilibrium phase-space distributions. **Physical Review A**. v.31, n.3, p.1695, 1985.

HUMPHREY, W; DALKE, A; SCHULTEN, K. VMD - Visual Molecular Dynamics, **Journal of molecular graphics**. v.14, n.1, 33-8, 27-8, 1996.

IBTISSEM, Z. *et al.* Potential toxicity of an insecticide of the family of carbamates on a bio-indicator model of the pollution the earthworm *Octodrilus complanatus*. **Annals of Biological Research**. v.3, n.11, p.5367–5373, 2012.

JIRAJAROENRAT, K. *et al.* Heterologous expression and characterization of alternatively spliced glutathione S-transferases from a single *Anopheles* gene. **Insect Biochemistry and molecular biology**. v. 31, n. 9, p. 867-875, 2001.

JURITZ, E. *et al.* Protein Conformational Diversity Modulates Sequence Divergence. **Molecular biology and evolution**. v.30, n.1, p.79-87, 2012.

KABSCH, W; SANDER, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features". **Biopolymers**. v.22, n.12, p.2577-637, 1983.

KIEFER, F. *et al.* The SWISSMODEL Repository and associated resources. **Nucleic Acids Research**. v.37, 2009.

LASKOWSKI, R. A. *et al.* (1993). PROCHECK - a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**. v.26, part 2, p.283-291, 1993.

LASKOWSKI, R. A. *et al.* AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. **Journal of Biomolecular NMR**. v.8, n.4, p. 477-486, 1996.

LAURIE, A.T; JACKSON, R.M. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. **Bioinformatics**. Oxford-England. v. 2, n.9, p.1908-1916, 2005.

LEACH, Andrew. **Molecular modelling: principles and applications**. (2001). S.l: Prentice Hall, 2001.

LI, X. *et al.* Ubiquitin- and ATP-independent proteolytic turnover of p21 by the REG gamma proteasome pathway. **Molecular Cell**. v.26, n.6, p.831-842, 2007.

LI, X; SCHULER, M.A; BERENBAUM, M.R. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. **Annual Review of Entomology**. v.52, p.231-253, 2007.

LIBRADO, P; ROZAS, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**. v.25, n. 11, p.1451-1452, 2009.

LIPKE, H; CHALKLEY, J. Glutathione, oxidized and reduced, in some dipterans treated with 1,1,1-trichloro-2,2-di-(*p*-chlorophenyl)ethane. **The Biochemical Journal**. v.85, n. 1, p.104-109, 1962.

LOBANOV, M. Y; BOGATYREVA, N. S; GALZITSKAYA, O. V. (2008). Radius of gyration as an indicator of protein structure compactness. **Molecular Biology**. v.42, n.4, p. 623–628, 2008.

LOW, W. Y. *et al* .(2007). Molecular evolution of glutathione S-transferases in the genus *Drosophila*. **Genetics**, v.177, n.3, p.1363–75, 2007.

LUMJUAN, N. *et al*. The *Aedes aegypti* glutathione transferase family. **Insect. biochemistry and molecular biology**. v.37, n.10, p. 1026–1035, 2007.

LUMJUAN, N. *et al*. (2005). Elevated activity of an Epsilon class glutathione transferase confers DDT resistance in the dengue vector, *Aedes aegypti*. **Insect biochemistry and molecular biology**. v.35, n.8, p.861–871, 2005.

MAIA, R. T. *et al*. (2009). Reduced levels of genetic variation in *Aedes albopictus* (Diptera: Culicidae) from Manaus, Amazonas State, Brazil, based on analysis of the mitochondrial DNA ND5 gene. **Genetics and molecular research: GMR**. v.8, n.3, p. 998–1007, 2009.

MANNERVIK, B; DANIELSON, U.H. Glutathione transferases: structure and catalytic activity. **Critical review in biochemistry**. v.23, n.3, p.283–337, 1988.

MELO, F; FEYTMANS, E. Assessing Protein Structures with a Non-local Atomic Interaction Energy. **Journal of molecular biology**. 277(5), p.1141, 1998.

MILLER, G.T. **Sustaining the Earth**, 6.ed. California: Thompson Learning, 2004. Chapter 9, Pages 211-216.

MINAKAWA, N. *et al*. Habitat characteristics of *Anopheles gambiae* s.s. larvae in a Kenyan highland. **Medical and veterinary Entomology**. v. 18, n.3, p.301-305, 2004.

MIYAMOTO, S; KOLLMAN, P.A. SETTLE: An analytical version of the SHAKE and RATTLE algorithms for rigid water models. **Journal of Computational Chemistry**. v. 13, n. 8, 952–962, 1992.

MORRIS, G. M. *et al.* Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. **Journal of Computational Chemistry.** v.16, p. 2785-91, 2009.

MORRIS, G. M. *et al.* Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. **Journal of Computational Chemistry.** v.19, n.14, p. 1639-1662, 1998.

MORSE, P. M; FESHBACH, H. Asymptotic Series; Method of Steepest Descent. §4.6 in: **Methods of Theoretical Physics, Part I.** New York: McGraw-Hill, 1953. p. 434-443.

MOURKIDOU, E. Glutathione S-transferase in the defence against pyrethroids in insects. **Insect biochemistry and molecular biology.** v.31, p.313-319, 2001.

NATH, A; ATKINS, W. M. A quantitative index of substrate promiscuity. **Biochemistry.** v.47, n.1, p. 157–166, 2008.

NEI, M. **Molecular Evolutionary Genetics.** (1987). New York: Columbia University Press, 1987.

NOBELI, I. *et al.* A ligand-centric analysis of the diversity and evolution of protein-ligand relationships in E.coli. **Journal of molecular biology.** 347(2), p.415–436, 2005.

NOBELI, I; FAVIA, A. D; THORNTON, J. M. Review Protein promiscuity and its implications for biotechnology. **Nature biotechnology.** v.27, n.2, p.157-167, 2009.

NOSÉ, S; KLEIN, M. L. Constant pressure molecular dynamics for molecular systems **Molecular Physics,** v.50, p. 1055, 1983.

ODUOLA, A. O. Evidence of carbamate resistance in urban populations of *Anopheles gambiae* s.s. mosquitoes resistant to DDT and deltamethrin insecticides in Lagos, South-Western Nigeria. **Parasites & vectors.** v.5, n.1, p. 116, 2012.

OOSTENBRINK, C. *et al* A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. **Journal of computational chemistry.** v.25, n. 13, p.1656–76, 2004.

OPPENORTH, F.J; VAN DER PAS, L.J.T; HOUX, N.W.H. Glutathione S-transferases and hydrolytic activity in a tetrachlorvinphos-resistant strain of housefly and their influence on resistance. **Pesticide biochemistry and physiology**. v.11, p.176-188, 1979.

ORTELLI, F. *et al.* Heterologous expression of four glutathione transferase genes genetically linked to a major insecticide resistance locus from the malaria vector, *Anopheles gambiae*. **The Biochemical Journal**. 373 (Pt 3), p.957-963, 2003.

PANG, Y. (2006). Novel Acetylcholinesterase Target Site for Malaria Mosquito Control. **PLoS ONE**. v.1, n. 1, 2006.

PARRINELLO, M. Polymorphic transitions in single crystals: A new molecular dynamics method. **Journal of Applied Physics**, v.52, n.12, p.7182, 1981.

PEARSON, W.R. . Phylogenies of glutathione transferase families. **Methods Enzymol**. 401, p.186–204, 2005.

PEITSCH, M. C. Protein modeling by E-mail Bio/Technology. *Nature biotechnology*. v.13, p.658-660, 1995.

PENILLA, R.P. *et al.* Changes in glutathione S-transferase levels in DDT resistant natural Mexican populations of *Anopheles albimanus* under different insecticide resistance management strategies. **Pesticide Biochemistry and Physiology**. v.86, p. 63-71, 2006.

PIGLIUCCI, M. Is evolvability evolvable? **Nature reviews. Genetics**, v.9, n.1, p.75–82, 2008.

PRABHU, K.S. *et al.* Characterization of a class alpha glutathione-S-transferase with glutathione peroxidase activity in human liver microsomes. **Archives of Biochemistry and biophysics**. 424(1), p.72– 80, 2004.

PRAPANTHADARA, L. *et al.* Cloning, expression and characterization of an insect class I glutathione S-transferase from *Anopheles dirus* species B. **Insect, biochemistry and molecular biology**. v.28, p.321-329, 1998.

PRAPANTHADARA, L. et al. (1996). Purification and characterization of a major glutathione transferase from the mosquito *Anopheles dirus* (species B). **Insect, biochemistry and molecular biology**. v.26, p.277-285, 1996.

PRAPANTHADARA, L. et al. Isoenzymes of glutathione S-transferase from the mosquito *Anopheles dirus* species B: the purification, partial characterization and interaction with various insecticides. **Insect, biochemistry and molecular biology**. v.30, p.395-403, 2000b.

QIN, G. et al. Characterization and Functional Analysis of Four Glutathione S-Transferases from the Migratory Locust, *Locusta migratoria*. **PLoS ONE**. v.8, n.3, 2013.

RAMSEY, D. C. et al. The Relationship Between Relative Solvent Accessibility and Evolutionary Rate in Protein Evolution. **Genetics**. v.188, n.2, p.479-488, 2011.

RANSOM, H; HEMINGWAY, J. Mosquito Glutathione Transferases (2005). **Methods in Enzymology**, v. 401, p.226–241, 2005.

RANSON, H. et al. Evolution of supergene families associated with insecticide resistance. **Science** v.298, n. 5591, p. 179–181, 2002.

RANSON, H. et al. Identification of a novel class of insect glutathione S-transferases involved in DDT resistance in the malaria vector, *Anopheles gambiae*. **Biochemical Journal**. 359(2), p. 295-304, 2001.

RANSON, H. et al. Insecticide resistance in *Anopheles gambiae*: data from the first year of a multi-country study highlight the extent of the problem. **Malaria Journal**. 17;8, p. 299, 2009.

RANSON, H; COLLINS, F.H; HEMINGWAY, J. (1998). The role of alternative mRNA splicing in generating heterogeneity within the *Anopheles gambiae* class I glutathione S-transferase family. **Proceedings of the National Academy of sciences of United States of America**. v. 95, n. 24, 1998.

RAYMOND, M. et al. Insecticide resistance in the mosquito *Culex pipiens*: what have we learned about adaptation? **Genetica**. v.112-113, n. 1, p. 287–96, 2001.

REDINBO, M. R. Promiscuity: what protects us, perplexes us. **Drug Discovery Today**. v. 9, n. 10, p. 431-432, 2004.

ROBINSON, A. *et al.* Modelling and bioinformatics studies of the human Kappa class glutathione transferase predict a novel third transferase family with homology to prokaryotic 2-hydroxychromene-2-carboxylate isomerases. **Biochemistry Journal**. 379, (pt 3), p. 541–552, 2004.

RORICK, M. M; WAGNER, G. P. Protein structural modularity and robustness are associated with evolvability. *Genome biology and evolution*. v. 3, p.456–75, 2011.

SACHS, J; MALANEY, P. (2002). The economic and social burden of malaria. *Nature*. v. 415, p.680-685, 2002.

SANTOS FILHO, Osvaldo Andrade; ALENCASTRO, Ricardo Bicca de. Modelagem de proteínas por homologia **Química nova**. v.26, n.2, São Paulo, mar/abr, 2003.

SCHAEFER, C; SCHLESSINGER, A; ROST, B. Protein secondary structure appears to be robust under in silico evolution while protein disorder appears not to be. **Bioinformatics (Oxford, England)**. v.26, n.5, p.625–631, 2010.

SCHAEFER, C; SCHLESSINGER, A; ROST, B. (2010). Protein secondary structure appears to be robust under in silico evolution while protein disorder appears not to be. **Bioinformatics (Oxford, England)**. v.26, n.5, p. 625–31, 2010.

SCHOLZ, B. D. *et al.* Anaemia is associated with reduced productivity of women workers even in less physically-strenuous tasks. *The British Journal of Nutrition*. v. 77, n. 1, p. 47–57. 1997.

SETZER, W. N. (2011). The molecular mechanism for DDT detoxification in *Anopheles gambiae*: a molecular docking study. **Journal of Biophysical Chemistry**. v.2, n. 2, p. 135-136, 2011.

SHEEHAN, D. Structure, function and evolution of glutathione transferases: implications for classification of 54 non-mammalian members of an ancient enzyme superfamily. **Biochemistry Journal**. V.360, p. 1–16, 2001.

TAJIMA, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. **Genetics**. v.123, n.3, p.585–595, 1989.

TANAKA, K; NAKAJIMA, M; KURIHARA, N. The mechanism of resistance to lindane and hexadeuterated lindane in the third in the third yumenoshima strain of house fly. **Pesticide Biochemistry and Physiology**. v.16, n.2, p.149-157, 1981.

TANG, A.H; TU, C.P. . Biochemical characterization of *Drosophila* glutathione S-transferase D1 and D2. **Journal of Biological Chemistry**. v.269, n.45, p.27876-27884, 1994.

TODD, A.E; ORENGO, C.A; THORNTON, J.M. (2001). Evolution of function in protein super- families, from a structural perspective. **Journal of Molecular Biology**. v.307, n. 4, p.1113–1143, 2001.

TORDA, A. E; WILFRED, F. Dynamics Configurations. **Journal of Computational Chemistry**. v.15, n.12, p.1331-1340, 1994.

TOUNG, Y.P.S; HSIEH, T.S; TU, C.P.D. (1990). Drosophila glutathione Stransferase1-1 shares a region of sequence homology with the maize glutathione S-transferase III. **Proceedings of the. National Academy of Sciences of the. Unites States of America**. v.87, n.1, p.31-35, 1990.

TOWNSEND, D. M; TEW, K. D; TAPIERO, H. The importance of glutathione in human disease. **Biomedicine & Pharmacotherapy**. v. 57, n. 3-4, p.145–155, 2003.

UDOMSINPRASERT, R. *et al.* Identification, characterization and structure of a new Delta class glutathione transferase isoenzyme. **Biochemical. Journal**. v.388, pt. 3, p. 763-771, 2005.

UDOMSINPRASERT, R; KETTERMAN, A.J. (2002). Expression and characterization of a novel class of glutathione S-transferase from *Anopheles dirus*. **Insect Biochemistry and Molecular Biology**. v.32, n. 4, p.425-433, 2002.

UGAKI, M; SHONO, T; FUKAMI, J. Metabolism of fenitrothion by organophosphorus-resistant and susceptible house flies *Musca domestica* L. **Pesticide Biochemistry and Physiology**. v. 23, n.1 p. 33-40, 1985.

VAN GUNSTEREN, W. F. *et al.* Biomolecular modeling: Goals, problems, perspectives. **Angewandte Chemie (International ed. in English)**. v.45, n.25, p.4064–92, 2006.

VAN GUNSTERN, W. F. *et al.* Biomolecular Simulation. The GROMOS96 Manual and User Guide; vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.y.: Zürich Groningen, 1996.

VONTAS, J.G. *et al.* Purification, molecular cloning and heterologous expression of a glutathione S-transferase involved in insecticide resistance from the rice brown planthopper, *Nilaparvata lugens*. **The Biochemical Journal**. v.362, pt. 2, p.329-337, 2002.

VONTAS, J.G; SMALL, G.J; HEMINGWAY, J. (2001). Glutathione S-transferases as antioxidant defence agents confer pyrethroid resistance in *Nilaparvata lugens*. **The Biochemical Journal**. v.357, n. pt. 1, p.65-72, 2001.

WAGNER, A. Robustness and evolvability: a paradox resolved. **Proceedings. Biological sciences / The Royal Society**. v.275, n.1630, p.91–100, 2008.

WAGNER, G. P; ALTENBERG, L. (1996.) Perspective: Complex Adaptations and the Evolution of Evolvability. **Evolution**. v. 50, n. 3. p. 967-976, 1996.

WANG, Y. *et al.* (2008). Structure of an insect epsilon class glutathione S-transferase from the malaria vector *Anopheles gambiae* provides an explanation for the high DDT-detoxifying activity. **Journal of Structural Biology**. v.164, n.2, p.228-235 2008.

WHO. World Health Organization. Malaria Report. 2005.

WHO. World Health Organization. Malaria report. 2010.

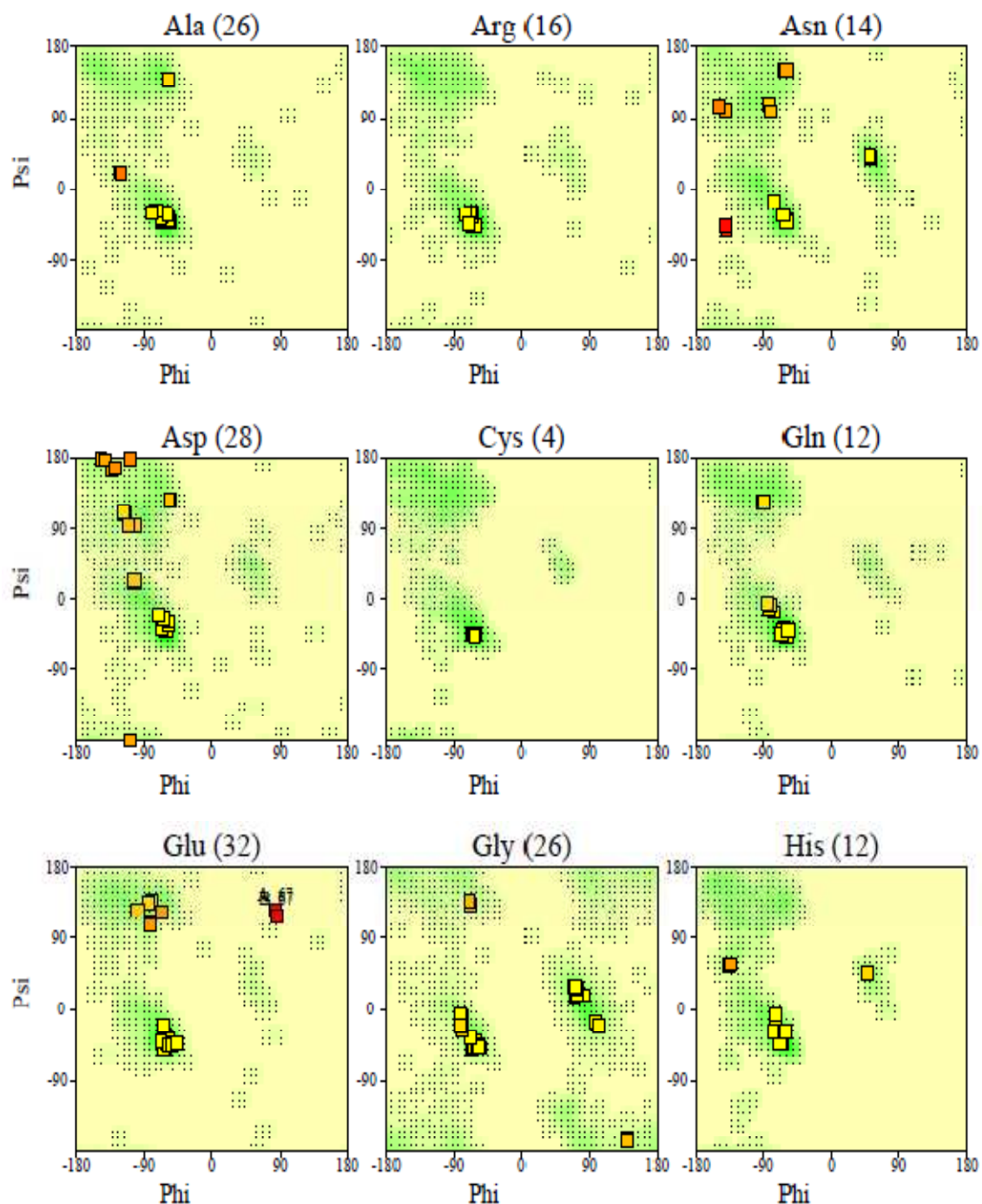
WILDENBURG, G; LIEBAU; E; HENKLE-DUHRSEN, K. (1998). *Onchocerca volvulus*: ultrastructural localization of two glutathione S-transferases. **Experimental.Parasitology**. v.88, n.1, p. 34–42, 1998.

XU, Q. *et al.* Resistance in the mosquito, *Culex quinquefasciatus*, and possible mechanisms for resistance. *Pest.Managment Science*. v. 61, n.11, p.1096-1102, 2005.

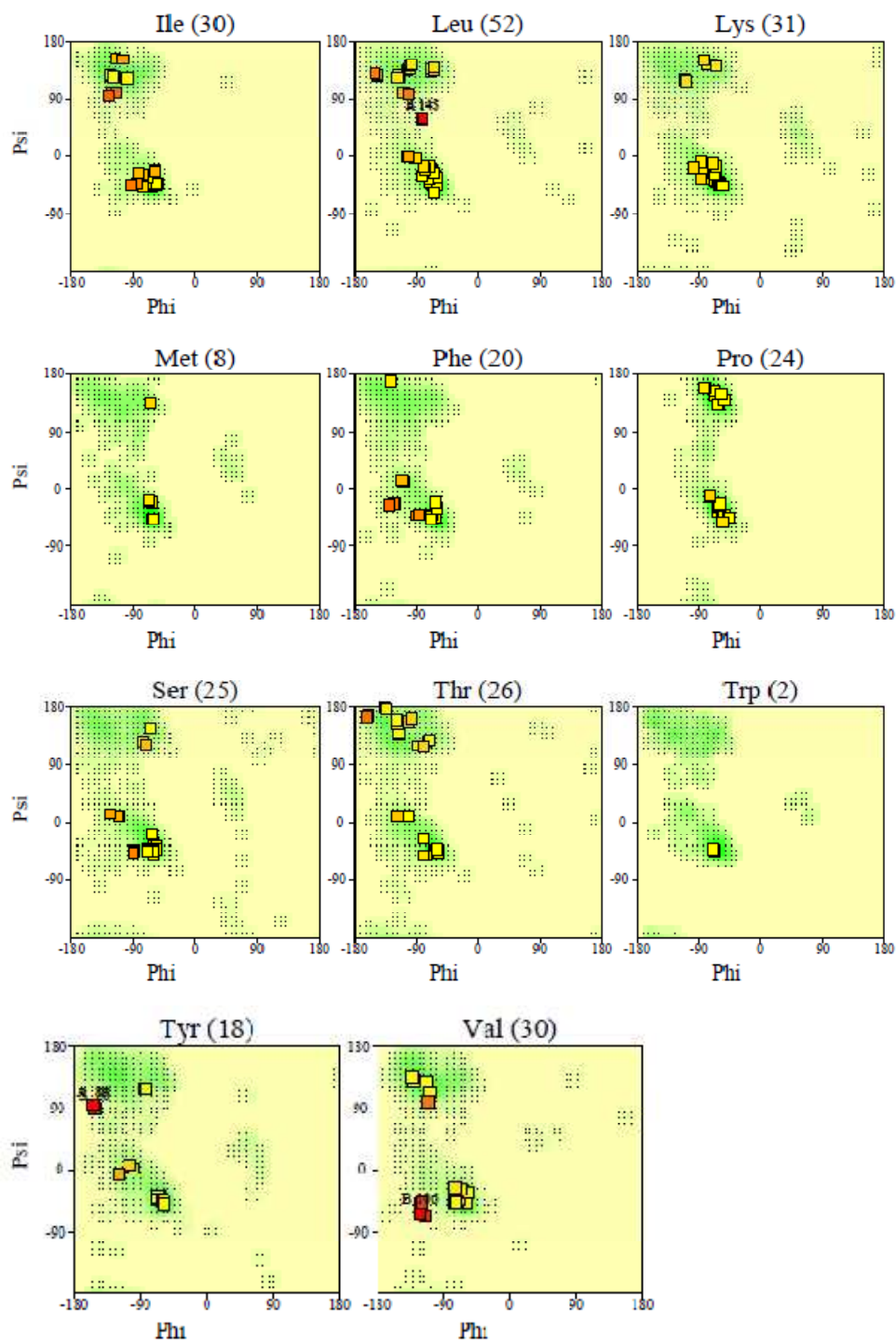
YANG, Z; NIELSEN, R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. **Molecular biology and evolution**. v.17, n.1, p. 32–43, 2000.

APÊNDICES

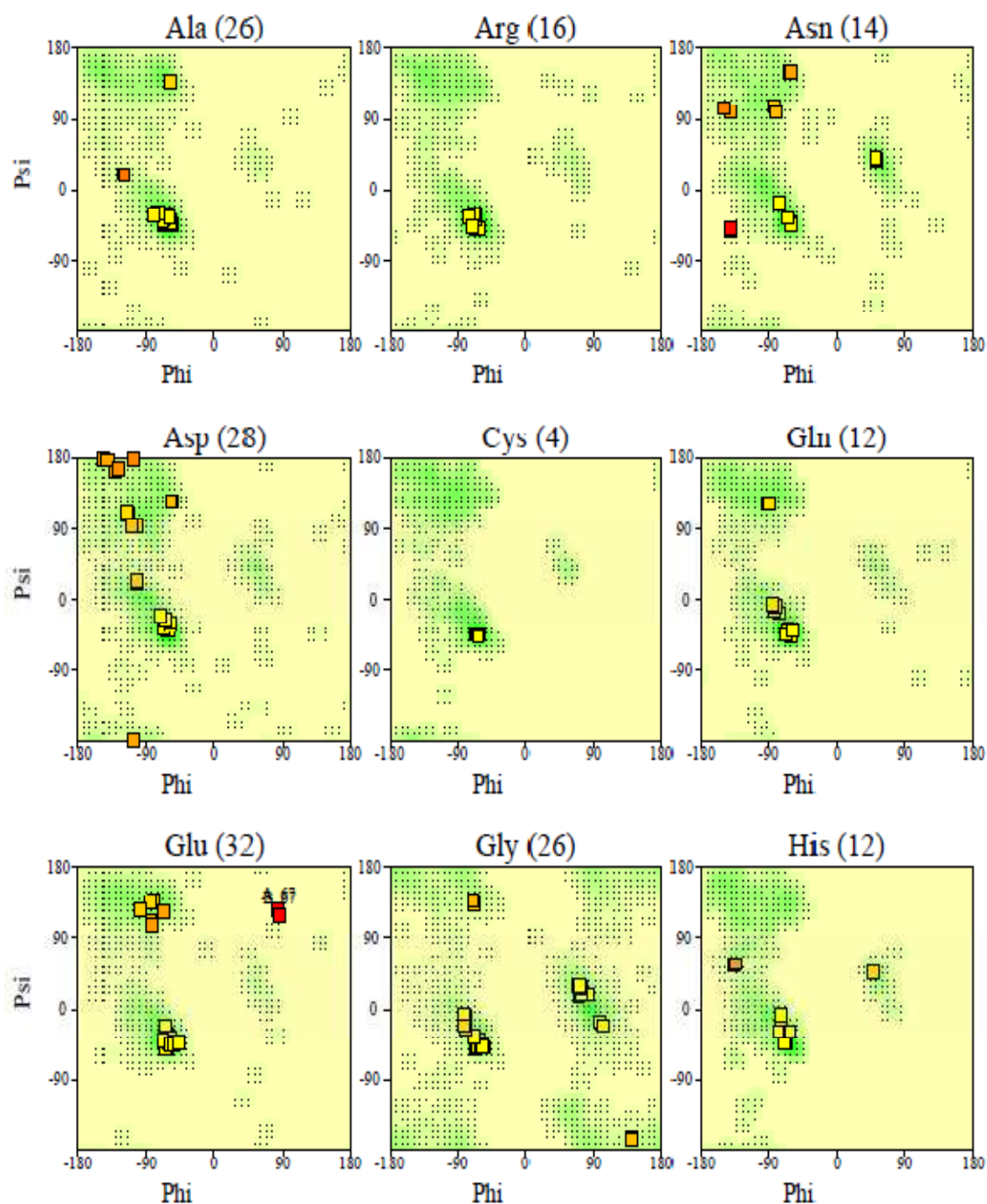
APÊNDICE A - Diagrama de Ramachandran para todos os tipos de resíduos da AgGSTE2. Os números de resíduos estão entre parêntesis. Aqueles em conformações desfavoráveis (score < -3,00) estão rotulados.



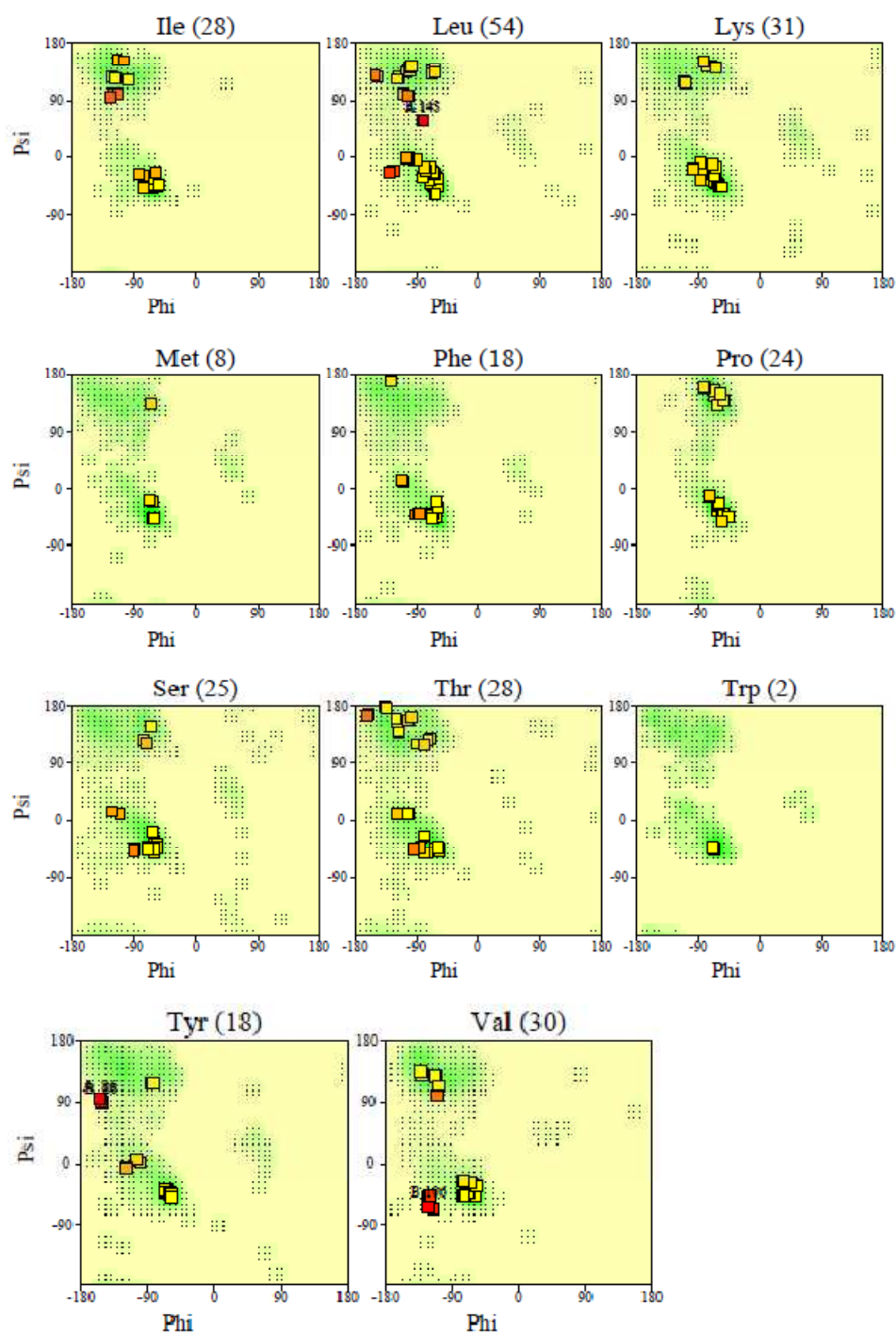
APÊNDICE A – Continuação



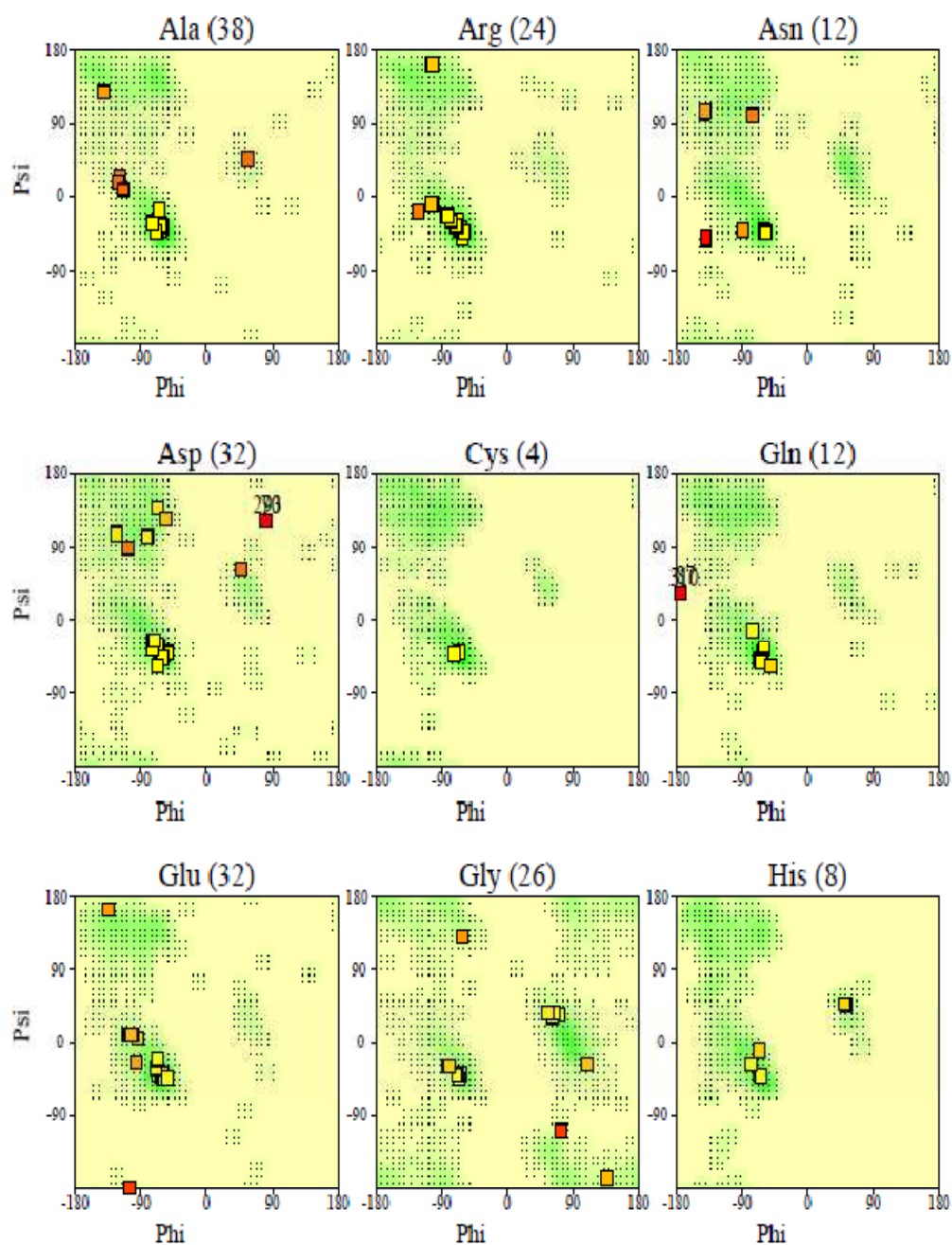
APÊNDICE B - Diagrama de Ramachandran para todos os tipos de resíduos da AgGSTE2mut. Os números de resíduos estão entre parêntesis. Aqueles em conformações desfavoráveis (score < -3,00) estão rotulados.



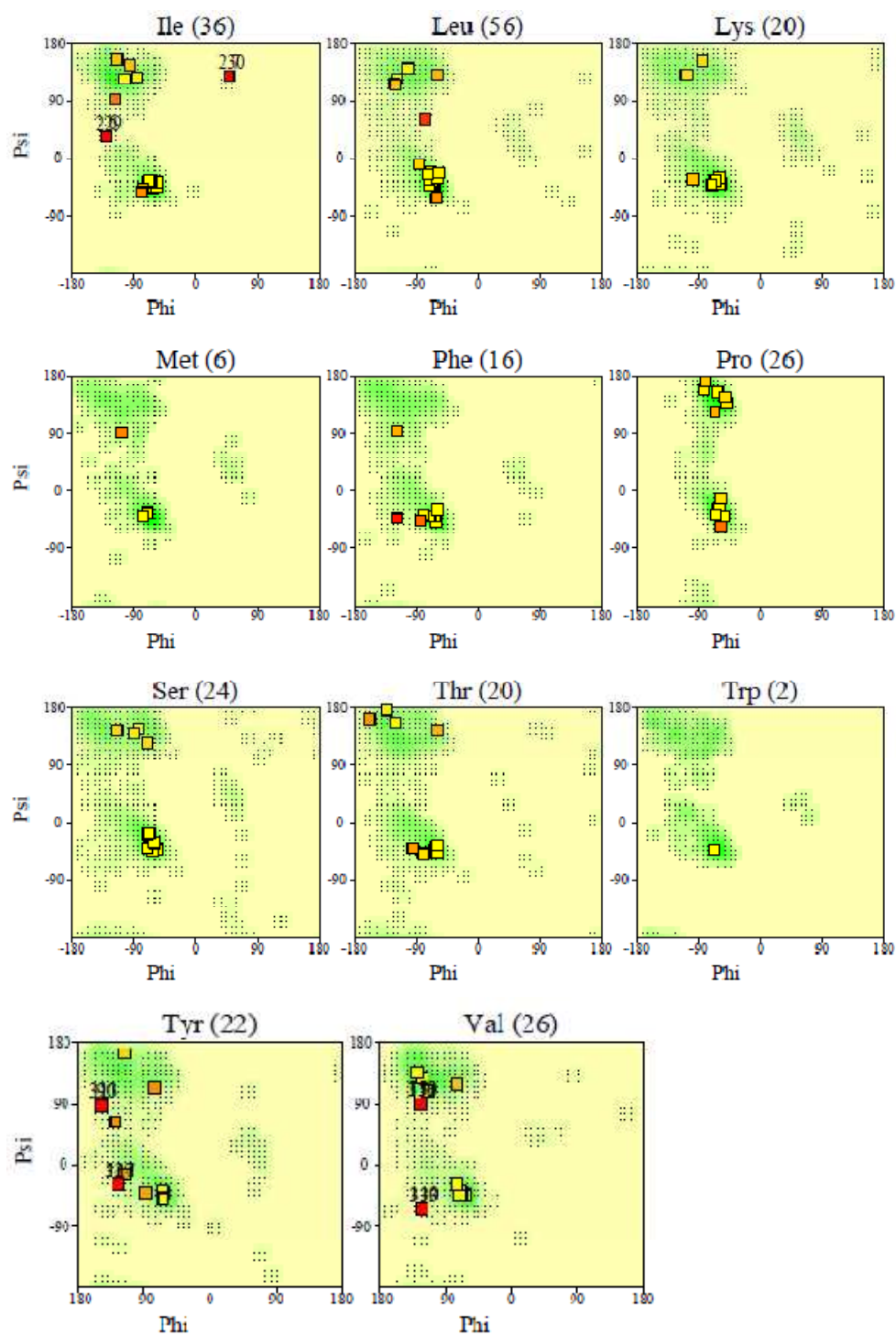
APÊNDICE B - Continuação



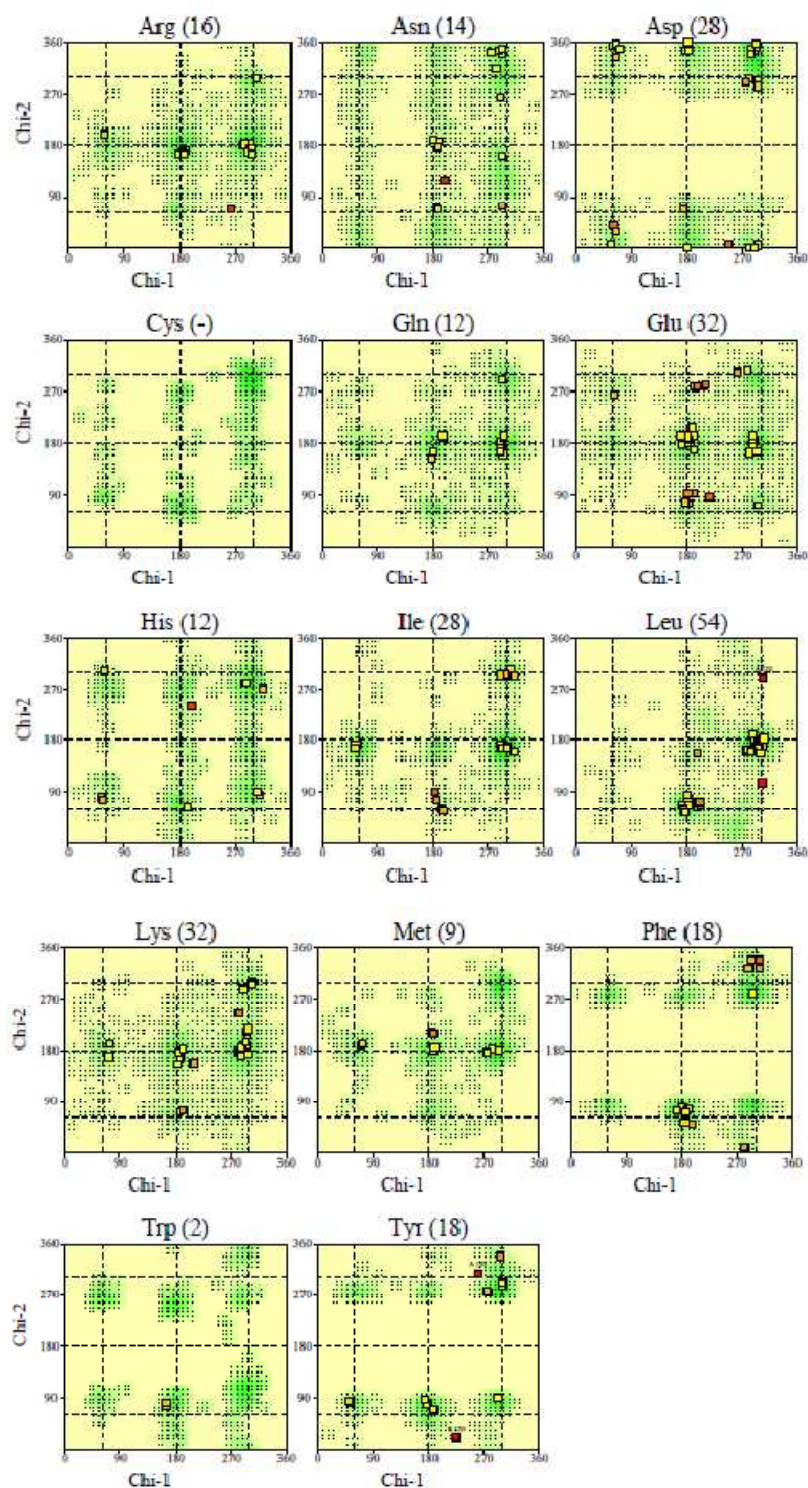
APÊNDICE C - Diagrama de Ramachandran para todos os tipos de resíduos da AgGSTe5. Os números de resíduos estão entre parêntesis. Aqueles em conformações desfavoráveis (score < -3,00) estão rotulados.



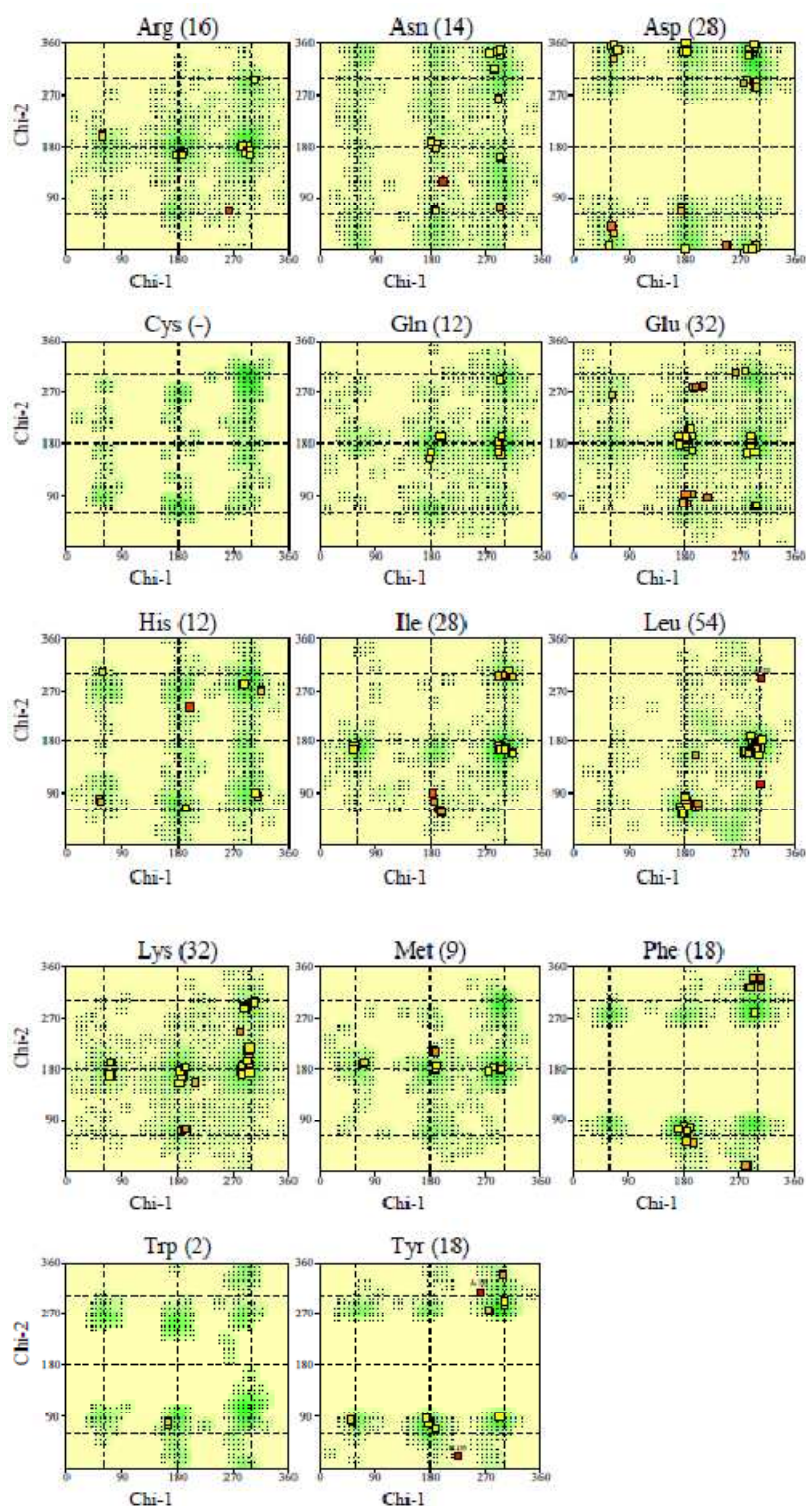
APÊNDICE C – Continuação



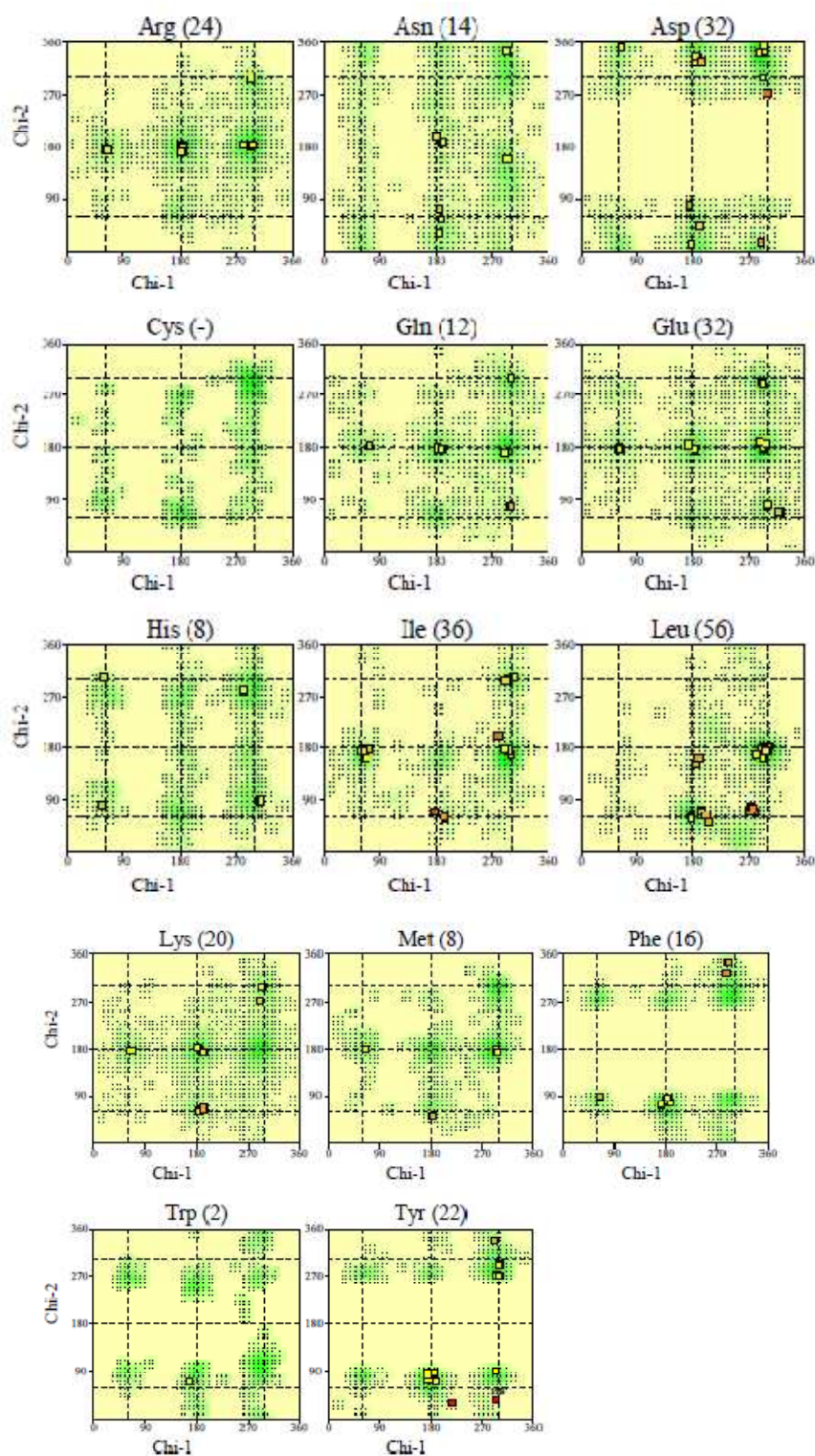
APÊNDICE D - Conformações dos ângulos Chi1-Chi2 para os resíduos da AgGSTe2. Os números de resíduos estão entre parêntesis. Aqueles em conformações desfavoráveis (score < -3,00) estão rotulados.



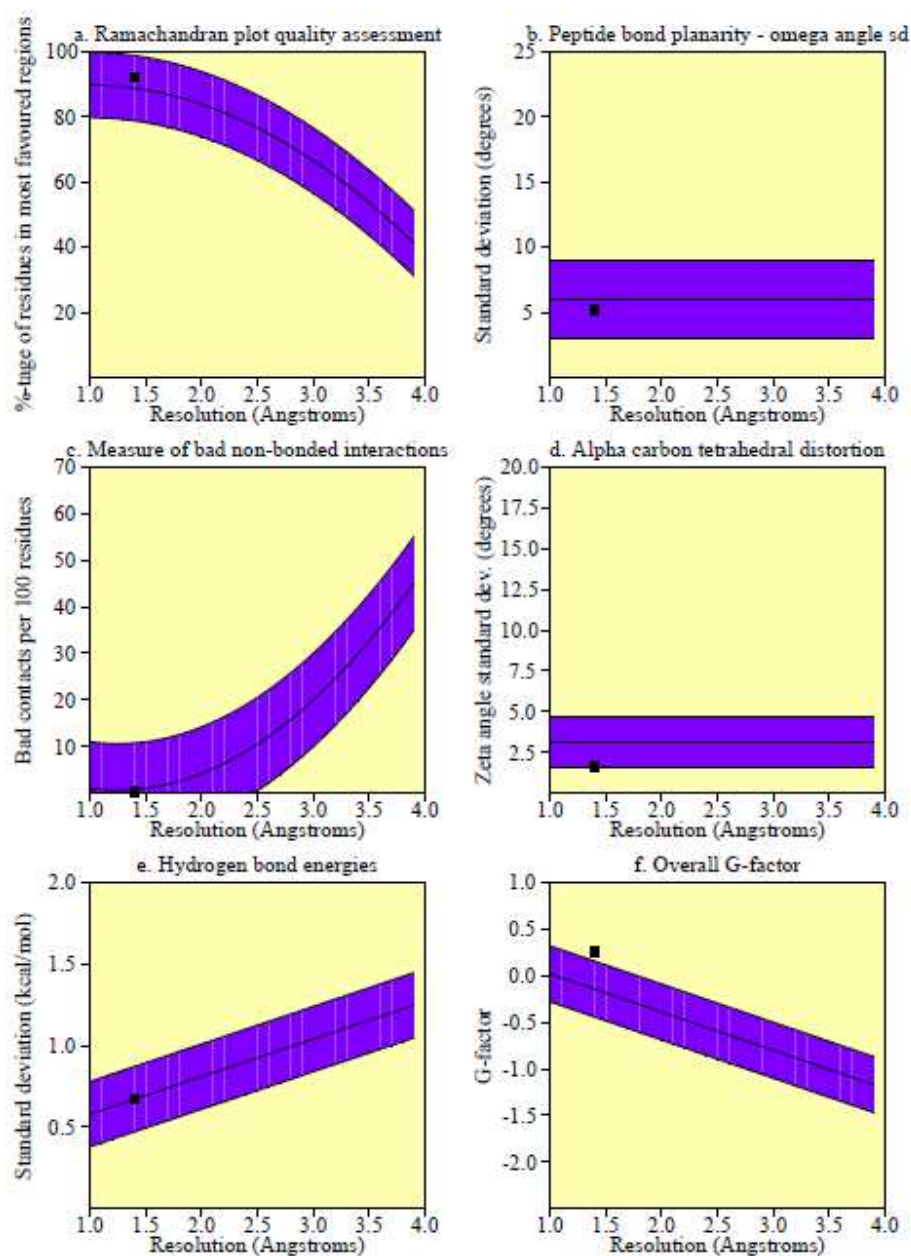
APÊNDICE E - Conformações dos ângulos Chi1-Chi2 para os resíduos da AgGSTE2mut. Os números de resíduos estão entre parêntesis. Aqueles em conformações desfavoráveis (score < -3,00) estão rotulados.



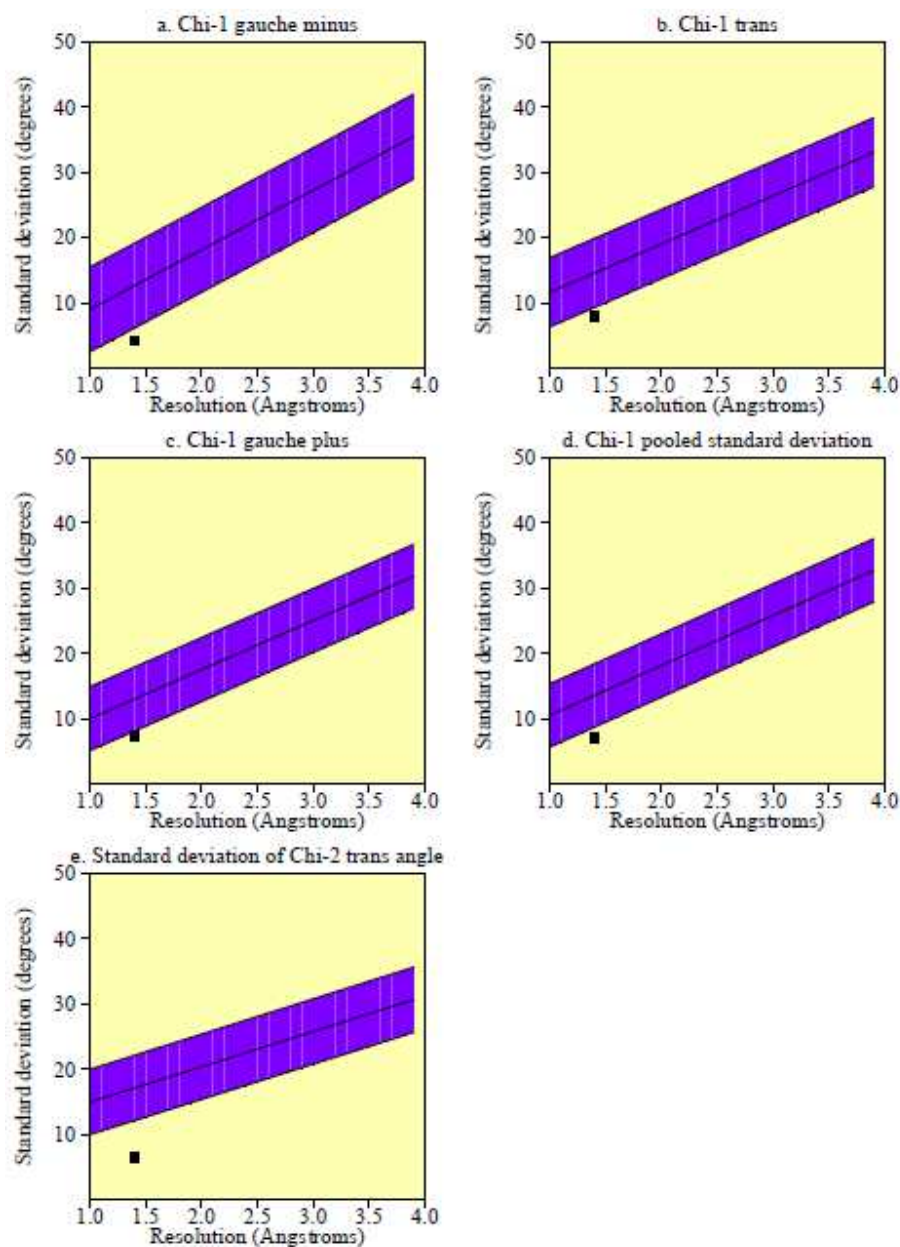
APÊNDICE F - Conformações dos ângulos Chi1-Chi2 para os resíduos da AgGSTE5. Os números de resíduos estão entre parêntesis. Aqueles em conformações desfavoráveis (score < -3,00) estão rotulados.



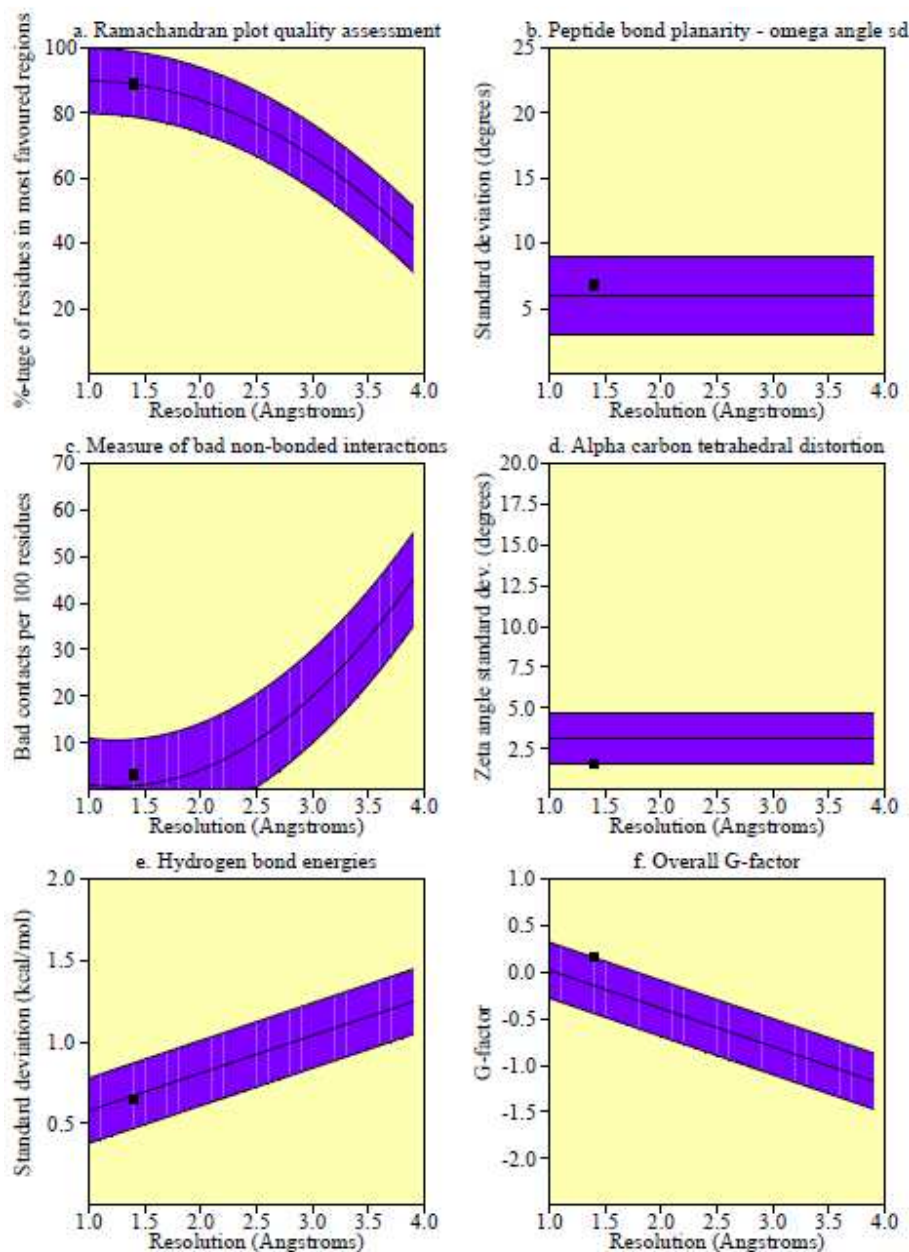
APÊNDICE G - Parâmetros das cadeias principais dos resíduos da AgGSTE2.

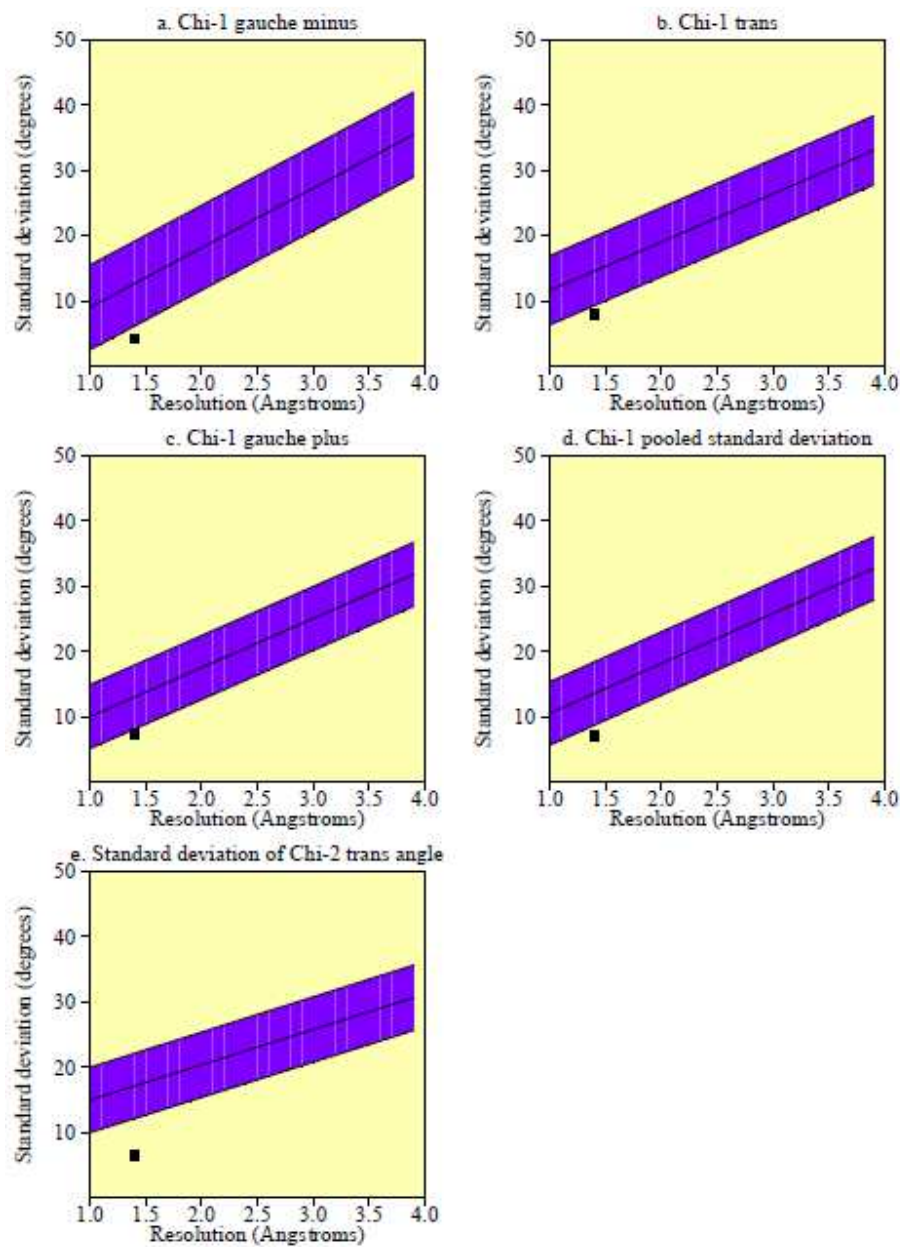


APÊNDICE H - Parâmetros das cadeias laterais dos resíduos da AgGSTE2.



APÊNDICE I - Parâmetros das cadeias principais dos resíduos da AgGSTe5.



APÊNDICE J - Parâmetros das cadeias laterais dos resíduos da AgGSTE5.

APÊNDICE K – Tabela - Dados detalhados dos potenciais de energia em cada resíduo da AgGSTE2 e AgGSTE5. **AA:** Número e tipo de resíduo; **Local:** Potencial geométrico local, **Burial:** Potencial de borda; **Contact:** Potencial de contato entre pares de resíduos.

Contato entre pares de resíduos.

AgGSTE2

AA	Local	Burial	Contact	Total
1. S	0.0000	-0.3600	-0.0750	-0.4350
2. N	0.3400	-0.3700	-0.0900	-0.1200
3. L	-0.0300	-0.0100	0.0450	0.0050
4. V	-0.5300	0.1100	0.3500	-0.0700
5. L	-0.4300	-0.4600	-0.2650	-1.1550
6. Y	1.3000	-0.2700	-0.3150	0.7150
7. T	-0.7800	0.2400	0.1850	-0.3550
8. L	-0.2300	-0.0100	-0.0800	-0.3200
9. H	0.8400	-0.0300	-0.2050	0.6050
10. L	0.3600	-0.0100	0.0250	0.3750
11. S	0.1500	0.0800	0.1600	0.3900
12. P	0.4700	0.4200	-0.1950	0.6950
13. P	-0.0700	0.4200	-0.0750	0.2750
14. C	0.0400	-0.2500	0.1800	-0.0300
15. R	0.0400	0.3400	-0.1900	0.1900
16. A	-0.8400	-0.2400	0.1300	-0.9500
17. V	-1.0200	-0.4700	-0.3800	-1.8700
18. E	-0.2800	0.6900	-0.1000	0.3100
19. L	-0.4300	-0.4600	-0.3300	-1.2200
20. T	-0.5400	0.2400	0.5700	0.2700
21. A	-0.0600	0.1300	-0.1900	-0.1200
22. K	0.2700	-0.2100	-0.0550	0.0050
23. A	0.3600	0.1300	-0.0850	0.4050
24. L	-0.2000	-0.0100	-0.2450	-0.4550
25. G	-0.1900	-0.2100	-0.0550	-0.4550
26. L	-0.2400	-0.0100	-0.3450	-0.5950
27. E	-0.1500	-0.3200	-0.1050	-0.5750
28. L	0.2400	-0.0100	-0.3100	-0.0800
29. E	0.9800	-0.1100	0.0050	0.8750
30. Q	0.3500	-0.0700	-0.0050	0.2750
31. K	-0.2700	-0.2100	-0.0850	-0.5650
32. T	-0.3100	-0.1800	0.0600	-0.4300
33. I	-0.5400	-0.4500	-0.0800	-1.0700
34. N	-0.5600	0.0100	-0.2350	-0.7850
35. L	0.1400	-0.0100	-0.4000	-0.2700
36. L	0.1700	1.0500	0.1550	1.3750
37. T	-0.2100	-0.1800	-0.1600	-0.5500
38. G	-0.6100	0.0200	0.1750	-0.4150
39. D	-0.5200	0.0200	-0.0950	-0.5950
40. H	-0.3500	0.0000	0.0800	-0.2700
41. L	0.1500	1.0500	0.0450	1.2450
42. K	0.1900	-0.2100	-0.4600	-0.4800
43. P	-0.4600	-0.4900	0.0700	-0.8800
44. E	-0.4000	-0.3200	-0.6100	-1.3300
45. F	0.4400	-0.4000	0.0850	0.1250
46. V	0.6700	0.1100	0.0250	0.8050
47. K	1.0800	-0.1500	-0.4300	0.5000
48. L	0.3200	-0.0100	-0.2050	0.1050
49. N	-0.1600	0.5200	-0.0050	0.3550
50. P	-0.4400	0.1900	0.0300	-0.2200
51. Q	-0.2800	-0.0700	-0.1700	-0.5200
52. H	-0.0700	-0.0300	-0.0050	-0.1050

53. T	0.3000	-0.0200	-0.1700	0.1100
54. I	0.0900	0.0300	-0.1350	-0.0150
55. P	-0.7100	0.4200	-0.1100	-0.4000
56. V	-0.5800	-0.4700	-0.2900	-1.3400
57. L	0.0300	-0.4600	-0.4000	-0.8300
58. D	0.5800	0.0200	0.5600	1.1600
59. D	-0.6100	0.0200	0.1950	-0.3950
60. N	-0.6300	-0.3700	-0.1400	-1.1400
61. G	-0.5600	-0.2100	0.0200	-0.7500
62. T	-0.5600	-0.0200	0.0800	-0.6600
63. I	-0.6600	0.0300	0.0750	-0.5550
64. I	-0.6000	-0.4500	-0.3400	-1.3900
65. T	0.2000	-0.0200	-0.1000	0.0800
66. E	-0.3900	0.6900	-0.0200	0.2800
67. S	-0.0200	0.3400	0.2450	0.5650
68. H	-0.3700	0.0000	0.0250	-0.3450
69. A	0.6300	-0.2400	0.0600	0.4500
70. I	-0.5400	-0.4500	-0.1550	-1.1450
71. M	-0.7200	-0.5000	-0.2900	-1.5100
72. I	-0.2300	-0.4500	-0.2150	-0.8950
73. Y	-0.3200	-0.2700	-0.2400	-0.8300
74. L	-0.0200	-0.4600	-0.4350	-0.9150
75. V	0.8000	-0.4700	-0.3100	0.0200
76. T	0.2100	-0.0200	0.1250	0.3150
77. K	0.0700	-0.2100	-0.0600	-0.2000
78. Y	-0.5000	0.5700	-0.1200	-0.0500
79. G	-0.4200	0.0200	0.1600	-0.2400
80. K	-0.5300	-0.1500	0.0600	-0.6200
81. D	-0.5500	-0.4900	-0.1350	-1.1750
82. D	-0.6500	0.0200	-0.0250	-0.6550
83. S	-0.4000	-0.3600	-0.2500	-1.0100
84. L	0.2700	-0.0100	-0.0050	0.2550
85. Y	-0.2700	-0.2700	0.1950	-0.3450
86. P	-0.3900	0.1900	-0.1200	-0.3200
87. K	-0.5900	-0.1500	-0.3150	-1.0550
88. D	-0.2800	-0.4900	-0.0800	-0.8500
89. P	-0.2500	0.1900	-0.1350	-0.1950
90. V	0.2400	0.6900	0.3450	1.2750
91. K	0.7800	-0.2100	-0.1750	0.3950
92. Q	-1.2500	-0.0700	-0.1400	-1.4600
93. A	-1.0500	-0.2400	0.1600	-1.1300
94. R	-0.6600	-0.2100	0.1050	-0.7650
95. V	-0.2600	-0.4700	0.0050	-0.7250
96. N	0.0600	0.5200	0.3450	0.9250
97. S	0.3900	0.3400	-0.0250	0.7050
98. A	-0.8000	-0.2400	-0.0800	-1.1200
99. L	-1.0400	-0.4600	0.1750	-1.3250
100. H	-0.5400	0.0000	-0.1900	-0.7300
101. F	-0.3700	0.0100	-0.0500	-0.4100
102. E	-0.1200	0.6900	0.1450	0.7150
103. S	-0.4200	0.3400	-0.0600	-0.1400
104. G	-0.5100	0.2100	0.1750	-0.1250
105. V	0.4000	0.1100	0.0200	0.5300
106. L	0.9500	-0.4600	0.0450	0.5350
107. F	0.4300	-0.4000	0.2350	0.2650
108. A	-1.0000	-0.2400	0.1200	-1.1200
109. R	-1.1800	0.3400	0.4200	-0.4200

110.	M	-1.1200	-0.5000	-0.2950	-1.9150
111.	R	-0.2800	-0.2100	0.0200	-0.4700
112.	F	-0.2500	0.0100	-0.0450	-0.2850
113.	I	-0.3200	-0.4500	-0.4000	-1.1700
114.	F	0.4900	0.0100	-0.3450	0.1550
115.	E	-0.2000	-0.1100	0.1000	-0.2100
116.	R	-1.0400	-0.2100	0.0800	-1.1700
117.	I	0.6300	-0.4500	0.2300	0.4100
118.	L	0.5000	-0.0100	-0.1650	0.3250
119.	F	0.5500	0.7300	0.0250	1.3050
120.	F	-0.2200	0.7300	0.0650	0.5750
121.	G	-0.3800	0.0200	0.0850	-0.2750
122.	K	-0.1700	-0.2100	0.0000	-0.3800
123.	S	-0.6100	0.0800	-0.1000	-0.5300
124.	D	0.3400	-0.4900	-0.3500	-0.5000
125.	I	-0.2600	0.0300	-0.3150	-0.5450
126.	P	-0.2000	0.1900	0.1500	0.1400
127.	E	-0.6800	-0.3200	0.1100	-0.8900
128.	D	-0.3900	0.0200	0.1350	-0.2350
129.	R	-0.2700	0.3400	0.1000	0.1700
130.	V	0.0900	0.1100	0.0250	0.1050
131.	E	-0.5900	-0.1100	-0.3600	-1.0600
132.	Y	-0.8400	-0.0700	0.0200	-0.8900
133.	V	-0.1100	-0.4700	-0.1900	-0.7900
134.	Q	-0.5300	-0.0700	0.0300	-0.5700
135.	K	-0.3800	-0.2100	-0.4950	-1.0850
136.	S	-0.0800	0.3400	0.5600	0.8200
137.	Y	0.2500	-0.2700	-0.0700	-0.0900
138.	E	-0.2000	-0.1100	-0.4300	-0.7400
139.	L	-1.0500	-0.4600	0.3900	-1.1200
140.	L	-0.8800	-0.4600	-0.2050	-1.5450
141.	E	0.0000	0.6900	-0.0050	0.6850
142.	D	-0.0200	0.0200	0.2650	0.2650
143.	T	0.1600	-0.0200	0.1300	0.2700
144.	L	0.8800	-0.4600	0.0650	0.4850
145.	V	0.1500	0.6900	0.1050	0.9450
146.	D	-0.2800	0.0200	0.2150	-0.0450
147.	D	-0.3900	0.0200	-0.1450	-0.5150
148.	F	0.2800	-0.4000	0.2500	0.1300
149.	V	0.6900	-0.4700	0.0550	0.3150
150.	A	-0.0500	-0.2400	0.0850	-0.2050
151.	G	-0.3600	0.0200	0.0600	-0.2800
152.	P	-0.2000	-0.4900	0.0150	-0.6750
153.	T	-0.0700	-0.1800	-0.1250	-0.3750
154.	M	0.3300	0.1100	0.0250	0.4650
155.	T	0.4900	0.2400	0.1050	0.8350
156.	I	0.5900	-0.4500	-0.4050	-0.2650
157.	A	0.1000	-0.2400	-0.0300	-0.1700
158.	D	-0.4900	0.7800	1.0150	1.3050
159.	F	0.0000	-0.4000	-0.3400	-0.7400
160.	S	0.0800	0.3400	0.1300	0.5500
161.	C	-0.7000	-0.2500	0.3250	-0.6250
162.	I	0.0300	0.4500	0.1200	-0.3600
163.	S	-0.4400	0.3400	-0.2450	-0.3450
164.	T	-0.4400	0.2400	0.0800	-0.1200
165.	T	0.8700	-0.4500	-0.5900	-0.1700
166.	S	0.2800	0.3400	-0.2100	0.4100
167.	S	0.3100	0.3400	-0.0650	0.5850
168.	I	-0.1300	-0.4500	-0.4550	-1.0350
169.	M	0.0700	0.1100	-0.2250	-0.0450
170.	G	-0.4100	0.0200	-0.0150	-0.4050
171.	V	0.3100	0.1100	-0.4900	-0.0700
172.	V	1.4000	0.1100	-0.5950	0.9150
173.	P	-0.5700	-0.4900	-0.1100	-1.1700
174.	L	0.1200	-0.0100	-0.4350	-0.3250
175.	E	0.9800	-0.1100	-0.5100	0.3600
176.	Q	-0.2300	-0.2800	0.0250	-0.4850
177.	S	-0.3500	-0.3600	-0.1300	-0.8400
178.	K	-0.5400	-0.1500	-0.5550	-1.2450
179.	H	-0.5300	0.0000	0.0350	-0.4950
180.	P	-0.3100	-0.4900	-0.1500	-0.9500
181.	K	-0.2400	-0.2100	-0.1750	-0.6250
182.	I	-0.2300	-0.4500	-0.2750	-0.9550
183.	Y	-0.1300	-0.0700	-0.2500	-0.4500
184.	A	-0.4500	0.1300	0.3250	0.0050
185.	W	-0.7300	-0.3400	-0.2950	-1.3650
186.	I	-1.1200	0.0300	-0.1200	-1.2100
187.	D	-0.5700	0.0200	-0.2150	-0.7650
188.	R	-0.9700	-0.2100	-0.2950	-1.4750
189.	L	-0.8300	-0.4600	-0.0450	-1.3350
190.	K	0.2300	-0.2100	-0.2300	-0.2100
191.	Q	0.1200	-0.2800	-0.1400	-0.3000
192.	L	0.0400	-0.0100	-0.1700	-0.1400
193.	P	0.1000	-0.4900	-0.1550	-0.5450
194.	Y	-0.3600	-0.0700	0.0000	-0.4300
195.	Y	0.4100	-0.2700	-0.2000	-0.0600
196.	E	-0.0700	-0.3200	-0.1000	-0.4900
197.	R	-0.1100	-0.3200	0.0300	-0.4000
198.	A	-0.1400	-0.2400	0.0300	-0.3500
199.	N	-0.5800	0.5200	0.0700	0.0100
200.	G	-0.5000	0.0200	-0.2000	-0.5800
201.	G	0.2100	0.0200	0.1500	-0.3400
202.	G	1.2500	0.0200	-0.1350	1.1350
203.	G	0.9200	0.2100	-0.4000	0.7300
204.	T	0.3600	-0.1800	0.2350	-0.0550
205.	D	-0.0100	0.0200	-0.1650	-0.1550
206.	L	0.8200	-0.4600	0.0800	0.4400
207.	G	0.1100	0.2100	0.2700	0.5900
208.	K	0.0000	-0.2100	-0.0950	-0.3050
209.	F	-0.6100	0.0100	0.0250	-0.5750
210.	V	0.2700	-0.4700	-0.1050	-0.3050
211.	L	-0.9400	-0.0100	0.2750	-0.6750
212.	A	-1.0100	0.1300	0.1600	-0.7200
213.	K	-0.2000	0.6300	-0.0550	0.3750
214.	K	-0.3100	0.6300	0.0150	0.3350
215.	E	-0.3800	-0.3200	-0.0850	-0.7850
216.	S	-0.6000	-0.1100	-0.7800	-1.4900
217.	N	0.4900	0.0100	-0.2750	0.2250
218.	A	-0.3700	0.1300	0.0350	-0.2050
219.	K	1.2500	-0.1500	-0.4800	0.6200
220.	A	0.4100	0.1200	-0.0250	0.5050
221.	M	0.5100	0.8000	0.0900	1.4000
222.	S	0.1600	-0.3600	0.0200	-0.1800
223.	N	-0.1500	-0.3700	-0.0900	-0.6100
224.	L	-0.0300	-0.4600	0.0300	-0.4600
225.	V	-0.5300	0.1100	0.3850	-0.0350
226.	L	0.4900	-0.4600	-0.2650	-0.2350
227.	Y	1.3000	-0.2700	-0.3150	0.7150
228.	T	-0.7800	0.2400	0.1850	-0.3550
229.	L	-0.2300	-0.0100	-0.0150	-0.2550
230.	H	0.8400	0.0500	-0.1900	0.7000
231.	L	0.3600	-0.0100	-0.0700	0.2800
232.	S	0.1500	0.0800	0.1600	0.3900
233.	P	0.4700	0.4200	0.1150	0.7750
234.	P	-0.0700	0.4200	-0.0750	0.2750
235.	C	0.0400	-0.2500	0.1800	-0.0300
236.	R	0.0400	0.3400	-0.1900	0.1900
237.	A	-0.8400	-0.2400	0.1000	-0.9800
238.	V	-1.0200	-0.4700	-0.3800	-1.8700
239.	E	-0.2800	0.6900	-0.2800	0.1300
240.	L	0.4300	0.4600	0.3000	-1.1900
241.	T	-0.5400	0.2400	0.5700	0.2700
242.	A	-0.0600	0.1300	-0.1850	-0.1150
243.	K	0.2700	-0.2100	-0.2650	-0.2050
244.	A	0.3600	0.1300	-0.0800	0.4100
245.	L	-0.2000	-0.0100	-0.2450	-0.4550
246.	G	-0.1900	-0.2100	0.0050	-0.3950
247.	L	-0.2400	-0.0100	-0.3050	-0.5550
248.	E	-0.1500	-0.3200	-0.0450	-0.5150
249.	L	0.2400	-0.0100	-0.2700	-0.0400
250.	E	0.9800	-0.1100	0.0250	0.8950

251. Q	0.3500	-0.0700	-0.0050	0.2750	322. F	-0.3700	0.0100	-0.0400	-0.4000
252. K	-0.2700	-0.2100	-0.0850	-0.5650	323. E	-0.1200	0.6900	0.1250	0.6950
253. T	-0.3100	-0.1800	0.0600	-0.4300	324. S	-0.4200	0.3400	-0.0850	-0.1650
254. I	-0.5400	-0.4500	-0.0800	-1.0700	325. G	-0.5100	0.2100	0.1700	-0.1300
255. N	-0.5600	0.0100	-0.1450	-0.6950	326. V	0.4000	0.1100	0.0350	0.5450
256. L	0.1400	-0.0100	-0.4000	-0.2700	327. L	0.9500	-0.4600	0.0450	0.5350
257. L	0.1700	1.0500	-0.1450	1.0750	328. F	0.4300	0.0100	0.2950	0.7350
258. T	-0.2100	-0.1800	-0.1600	-0.5500	329. A	-1.0000	-0.2400	0.1250	-1.1150
259. G	-0.6100	0.0200	0.1750	-0.4150	330. R	-1.1800	0.3400	0.4450	-0.3950
260. D	-0.5200	0.0200	-0.1450	-0.6450	331. M	-1.1200	-0.5000	-0.3750	-1.9950
261. H	-0.3500	0.0000	0.0800	-0.2700	332. R	-0.2800	-0.2100	0.0200	-0.4700
262. L	0.1500	1.0500	0.0750	1.2750	333. F	-0.2500	0.0100	-0.0450	-0.2850
263. K	0.1900	-0.2100	-0.4600	-0.4800	334. I	-0.3200	-0.4500	-0.3750	-1.1450
264. P	-0.4600	-0.4900	0.1450	-0.8050	335. F	0.4900	0.0100	-0.3250	0.1750
265. E	-0.4400	-0.3200	-0.6300	-1.3500	336. E	-0.2000	0.6900	0.1000	0.5900
266. F	0.4000	-0.4000	0.0300	0.0700	337. R	-1.0400	0.3400	0.0800	-0.6200
267. V	0.6700	0.1100	-0.0750	0.7050	338. I	0.6300	-0.4500	0.2850	0.4650
268. K	1.0800	-0.1500	-0.3850	0.5450	339. L	0.5000	-0.0100	-0.2350	0.2550
269. L	0.3200	-0.0100	-0.2300	0.0800	340. F	0.5500	0.7300	0.0750	1.3550
270. N	-0.1600	0.5200	0.0800	0.4400	341. F	-0.2200	0.7300	0.1050	0.6150
271. P	-0.4400	0.1900	0.0100	-0.2400	342. G	-0.3800	0.0200	0.0850	-0.2750
272. Q	-0.2800	-0.0700	-0.1700	-0.5200	343. K	-0.1700	-0.2100	-0.0350	-0.4150
273. H	-0.0700	-0.0300	-0.0050	-0.1050	344. S	-0.5100	0.0800	-0.0400	-0.4700
274. T	0.3000	-0.0200	-0.2000	0.0800	345. D	0.3400	-0.4900	-0.3600	-0.5100
275. I	0.0900	-0.4500	-0.1350	-0.4950	346. I	-0.2600	0.0300	-0.3500	-0.5800
276. P	-0.7100	0.4200	-0.0800	-0.3700	347. P	-0.2000	0.1900	0.1500	0.1400
277. V	-0.5800	-0.4700	-0.2700	-1.3200	348. E	-0.6800	-0.3200	0.1100	-0.8900
278. L	0.0300	-0.4600	-0.4000	-0.8300	349. D	-0.3900	0.0200	0.0950	-0.2750
279. D	0.5800	0.0200	0.5850	1.1850	350. R	-0.2700	0.3400	0.1000	0.1700
280. D	-0.6100	0.0200	0.2200	-0.3700	351. V	0.0900	0.1100	-0.0950	0.1050
281. N	-0.6300	-0.3700	-0.0500	-1.0500	352. E	-0.5900	-0.1100	-0.3600	-1.0600
282. G	-0.5600	-0.2100	-0.0050	-0.7750	353. Y	-0.8400	-0.0700	0.0200	-0.8900
283. T	-0.5600	-0.0200	-0.0850	-0.6650	354. V	-0.1300	-0.4700	-0.1900	-0.7900
284. I	-0.6600	0.0300	0.0700	-0.5600	355. Q	-0.5300	-0.0700	0.0650	-0.5350
285. I	-0.6000	-0.4500	-0.2250	-1.2750	356. K	-0.3800	-0.2100	-0.4950	-1.0850
286. T	0.2000	-0.0200	-0.1000	0.0800	357. S	-0.0800	0.3400	0.5600	0.8200
287. E	-0.3900	0.6900	-0.0950	0.2050	358. Y	0.2500	-0.2700	-0.0700	-0.0900
288. S	-0.0200	0.3400	0.2450	0.5650	359. E	-0.2000	-0.1100	-0.4300	-0.7400
289. H	-0.3700	0.0000	0.0250	-0.3450	360. L	-1.0500	-0.4600	0.3900	-1.1200
290. A	-0.8800	-0.2400	0.0600	-1.0600	361. L	-0.8800	-0.4600	-0.2050	-1.5450
291. I	-0.5400	-0.4500	-0.3150	-1.3050	362. E	0.0000	0.6900	-0.0050	0.6850
292. M	-0.7200	-0.5000	-0.2650	-1.4850	363. D	-0.0200	0.0200	0.2850	0.2850
293. I	-0.2300	-0.4500	-0.2150	-0.8950	364. T	0.1600	-0.0200	0.1250	0.2650
294. Y	-0.3200	-0.2700	-0.2400	-0.8300	365. L	0.8800	-0.4600	0.0650	0.4850
295. L	-0.0200	-0.4600	-0.4050	-0.8850	366. V	0.1500	0.6900	0.1050	0.9450
296. V	0.8000	-0.4700	-0.1050	0.2250	367. D	-0.2800	0.0200	0.2150	-0.0450
297. T	0.2100	-0.0200	0.1200	0.3100	368. D	-0.3900	0.0200	-0.1450	-0.5150
298. K	0.0700	-0.2100	-0.0600	-0.2000	369. F	0.2800	-0.4000	0.2500	0.1300
299. Y	-0.5000	-0.0700	-0.1300	-0.7000	370. V	0.6900	-0.4700	0.0950	0.3150
300. G	-0.4200	0.0200	0.1900	-0.2100	371. A	-0.0500	-0.2400	0.0500	-0.2400
301. K	-0.5300	-0.1500	0.0600	-0.6200	372. G	-0.3600	0.0200	0.0250	-0.3150
302. D	-0.5500	-0.4900	-0.1350	-1.1750	373. P	-0.2000	-0.4900	0.0150	-0.6750
303. D	-0.6500	0.0200	0.0950	-0.5350	374. T	-0.0700	-0.1800	-0.0800	-0.3300
304. S	-0.4000	-0.3600	-0.2350	-0.9950	375. M	0.3300	0.1100	0.0250	0.4650
305. L	0.2700	-0.0100	0.0300	0.2900	376. T	0.4900	0.2400	0.1050	0.8350
306. Y	-0.2700	-0.2700	0.1950	-0.3450	377. I	0.5900	-0.4500	-0.4500	-0.3100
307. P	-0.3900	0.1900	-0.1200	-0.3200	378. A	0.1000	-0.2400	-0.0300	-0.1700
308. K	-0.5900	-0.1500	-0.3950	-1.1350	379. D	-0.4900	0.7800	0.8650	1.1550
309. D	-0.2800	-0.4900	-0.1250	-0.8950	380. F	0.0000	-0.4000	-0.3350	-0.7350
310. P	-0.2500	0.1900	-0.1350	-0.1950	381. S	0.0800	0.3400	0.1650	0.5850
311. V	0.2400	0.1100	0.2550	0.6050	382. C	-0.7000	-0.2500	0.3400	-0.6100
312. K	-0.7100	-0.2100	-0.1650	-1.0850	383. I	-0.0300	-0.4500	0.1200	-0.3600
313. Q	-1.2500	-0.0700	-0.2200	-1.5400	384. S	-0.4400	0.3400	-0.2650	-0.3650
314. A	-1.0500	-0.2400	0.2100	-1.0800	385. T	-0.4400	0.2400	0.0800	-0.1200
315. R	-0.6600	-0.2100	0.0700	-0.8000	386. I	0.8700	-0.4500	-0.5650	-0.1450
316. V	-0.2600	-0.4700	0.0050	-0.7250	387. S	0.2800	0.3400	-0.2450	0.3750
317. N	0.0600	0.5200	0.3400	0.9200	388. S	0.3100	0.3400	-0.0950	0.5550
318. S	0.3900	0.3400	0.0050	0.7350	389. I	-0.1300	-0.4500	-0.5050	-1.0850
319. A	-0.8000	-0.2400	-0.0800	-1.1200	390. M	0.0700	0.1100	-0.1600	0.0200
320. L	-1.0400	-0.4600	0.1750	-1.3250	391. G	-0.4100	0.0200	-0.0150	-0.4050
321. H	-0.5400	0.0000	-0.1900	-0.7300	392. V	0.3100	0.1100	-0.5100	-0.0900

393. V	1.4000	0.1100	-0.6350	0.8750
394. D	-0.5700	-0.4900	-0.1100	-1.1700
395. L	0.1200	-0.0100	-0.3900	-0.2800
396. K	0.9800	-0.1100	-0.5150	0.3550
397. Q	-0.2300	-0.2800	0.0200	-0.4900
398. S	-0.3500	-0.3600	-0.1300	-0.8400
399. K	-0.5400	-0.1500	-0.5700	-1.2600
400. H	0.5300	0.0000	0.0300	-0.5000
401. P	-0.3100	-0.4900	-0.1150	-0.9150
402. K	-0.2400	-0.2100	-0.2100	-0.6600
403. I	-0.2300	-0.4500	-0.2200	-0.9000
404. Y	-0.1300	-0.0700	-0.2850	-0.4850
405. A	-0.4500	0.1300	0.2800	-0.0400
406. W	-0.7300	-0.3400	-0.2950	-1.3650
407. I	1.1200	0.0300	0.0950	-1.1050
408. D	-0.5700	0.0200	-0.2250	-0.7750
409. R	-0.9700	-0.2100	-0.2950	-1.4750
410. L	-0.8300	-0.4600	-0.1800	-1.4700
411. K	0.2300	-0.2100	-0.2300	-0.2100
412. Q	0.1200	-0.2800	-0.1400	-0.3000
413. L	0.0400	-0.0100	-0.1700	-0.1400
414. P	0.1000	-0.4900	-0.1550	-0.5450
415. Y	-0.3600	-0.0700	0.0000	-0.4300
416. Y	0.4100	-0.2700	-0.1300	0.0100
417. E	-0.0700	-0.3200	-0.1000	-0.4900
418. R	-0.1100	-0.3200	0.0700	-0.3600
419. A	-0.1400	-0.2400	0.0300	-0.3500
420. N	-0.5800	0.5200	0.0850	0.0250
421. G	-0.5000	0.0200	-0.1800	-0.6600
422. G	-0.2100	0.0200	-0.1650	-0.3550
423. G	1.2500	0.2100	-0.1700	1.2900
424. G	0.9200	0.2100	-0.5800	0.5500
425. T	0.3600	-0.0200	-0.2600	0.0800
426. D	-0.0100	0.0200	-0.1450	-0.1350
427. L	0.8200	-0.4600	0.0550	0.4150
428. G	0.1100	0.2100	0.0550	0.3750
429. K	0.0000	-0.2100	-0.0950	-0.3050
430. F	-0.6100	0.0100	0.1200	-0.4800
431. V	0.2700	-0.4700	-0.0850	-0.2850
432. L	-0.9400	-0.0100	0.2750	-0.6750
433. A	-1.0100	0.1300	0.1600	-0.7200
434. K	-0.6900	0.6300	-0.0550	-0.1150
435. K	-0.0400	0.6300	0.0150	0.6050
436. E	0.4600	-0.3200	0.0350	0.1750
437. E	-0.6000	-0.3200	-0.4150	-1.3350
438. N	-0.2600	0.0100	-0.2850	-0.5350
439. A	0.5400	0.1200	0.0350	0.6950
440. K	0.0000	-0.1500	0.0000	-0.1500
19. R	0.7100	0.3400	0.0250	1.1050
20. A	-0.8400	-0.2400	0.2650	-0.8150
21. V	-1.0200	-0.4700	-0.2200	-1.7100
22. E	-0.2800	0.6900	-0.2100	0.2000
23. L	-0.4300	-0.4600	-0.0650	-0.9550
24. T	-0.5400	0.2400	0.6500	0.3500
25. A	-0.0600	-0.2400	-0.0400	-0.3400
26. K	0.4400	0.2100	0.0900	0.1400
27. L	0.6500	-0.0100	-0.2050	0.4350
28. L	-0.1000	-0.0100	-0.2450	-0.3550
29. G	-0.1900	-0.2100	0.0450	-0.3550
30. T	-0.3900	-0.0100	-0.3550	-0.7550
31. S	-0.3700	-0.3600	0.1350	-0.5950
32. L	0.3000	0.0100	0.4200	-0.1300
33. D	-0.2400	0.0200	0.1200	-0.1000
34. I	-0.2000	0.0300	0.0050	-0.1650
35. V	-0.5900	0.1100	-0.0300	-0.5100
36. P	-0.7300	-0.4900	0.0600	-1.1600
37. T	-0.0700	0.0300	-0.2250	-0.2650
38. N	-0.5600	-0.3700	-0.0300	-0.9600
39. L	0.1400	-0.0100	-0.3550	-0.2250
40. L	0.3800	1.0500	0.1000	1.5300
41. A	-0.2800	0.1200	0.1700	0.0100
42. G	-0.4900	0.0200	0.1800	-0.2900
43. D	-0.5200	0.0200	-0.0250	-0.5250
44. H	-0.4600	0.0000	0.0250	-0.4350
45. R	0.2300	-0.2100	-0.1000	-0.0800
46. T	-0.1300	-0.1800	-0.1400	-0.4500
47. D	-0.4200	-0.4900	0.0000	-0.9100
48. E	-0.1700	-0.3200	-0.3150	-0.8050
49. F	0.4700	0.0100	-0.0900	0.3900
50. L	0.1900	-0.4600	0.6450	0.3750
51. R	0.6900	0.0400	-0.6750	0.0550
52. L	0.1700	-0.0100	-0.1300	0.0300
53. N	-0.1600	0.5200	0.0950	0.4550
54. P	-0.4400	0.4200	0.2500	0.2300
55. Q	-0.2800	0.5000	-0.0750	0.1450
56. H	-0.0700	0.0000	-0.1450	-0.2150
57. T	0.3000	-0.0200	-0.1750	0.1050
58. I	0.0900	0.0300	-0.0800	0.0400
59. P	0.7700	0.4200	-0.1550	1.0350
60. V	-0.5700	-0.4700	-0.1550	-1.1950
61. I	-0.3300	-0.4500	-0.2250	-1.0050
62. D	0.4900	0.0200	0.1900	0.7000
63. D	-0.7200	0.0200	0.3000	-0.4000
64. G	-0.6200	-0.2100	0.2100	-0.6200
65. G	-0.2200	-0.2100	-0.0050	-0.4350
66. V	-0.5100	0.1100	0.1550	-0.2450
67. I	-0.7800	0.8400	0.1500	0.2100
68. V	-0.2000	-0.4700	-0.2100	-0.8800
69. R	0.1900	0.3400	0.0650	0.5950
70. D	-0.2600	0.7800	0.2050	0.7250
71. S	-0.6800	0.0800	0.2800	-0.3200
72. H	-0.3700	0.0000	0.0100	-0.3600
73. A	0.6300	-0.2400	0.1050	0.4950
74. I	-0.9300	-0.4500	-0.1450	-1.5250
75. I	0.3400	0.4500	0.2650	-1.0550
76. I	-0.5200	0.0300	-0.2400	-0.7300
77. Y	-0.3200	-0.2700	-0.0950	-0.6850
78. L	-0.0200	-0.4600	-0.4550	-0.9350
79. V	0.7600	-0.4700	-0.2550	0.0350
80. Q	0.0400	-0.0700	0.0050	-0.0250
81. K	-0.2800	-0.2100	-0.1500	-0.6400
82. Y	-0.5000	0.5700	-0.1650	-0.0950
83. G	-0.4200	0.0200	0.2000	-0.2000
84. K	-0.5300	-0.1500	0.0000	-0.6800
85. D	-0.6100	-0.4900	0.0000	-1.1000
86. G	-0.6400	-0.2100	-0.2000	-1.0500
87. Q	-0.3000	-0.0700	0.0450	-0.3250
88. T	0.4100	-0.1800	-0.0600	0.1700

AgGSTE5

AA	Local	Burial	Contact	Total
1. M	0.0000	0.8000	0.0650	0.8650
2. A	-0.0900	0.1300	-0.0800	-0.0400
3. T	-0.0400	-0.1800	-0.0300	-0.2500
4. N	-0.4100	-0.3700	0.0750	-0.7050
5. P	-0.3100	0.1900	0.0600	-0.0600
6. I	0.6000	0.0300	0.0350	0.6650
7. I	0.7900	0.4500	0.0350	0.3750
8. K	-0.3700	0.6300	-0.3450	-0.0850
9. L	-0.1000	-0.4600	-0.5350	-1.0950
10. Y	1.3000	-0.2700	-0.1500	0.8800
11. T	-0.3100	0.2400	0.0950	0.0250
12. A	0.4100	0.1300	-0.1650	0.3750
13. K	0.4400	0.2100	0.2300	0.0000
14. L	0.6600	-0.0100	-0.3150	0.3350
15. S	0.1500	0.0800	0.0600	0.2900
16. P	0.4700	0.4200	-0.0550	0.8350
17. P	-0.6200	0.1900	-0.0050	-0.4350
18. G	-0.5900	0.2100	0.0000	-0.3800

89. L	0.5200	-0.0100	-0.2050	0.3050	160. T	0.9100	0.2400	0.1250	1.2750
90. Y	-0.2700	-0.2700	-0.0750	-0.6150	161. L	0.3800	-0.4600	-0.5650	-0.6450
91. P	0.2400	0.1200	0.1050	0.3250	162. A	0.0700	0.2400	0.0550	-0.2550
92. E	-0.6800	-0.3200	-0.4350	-1.4350	163. D	-0.2900	0.7800	0.8550	1.3450
93. D	-0.4800	-0.4900	0.0250	-0.9450	164. V	0.3900	-0.4700	-0.5600	-0.6400
94. P	-0.4200	-0.4900	-0.0550	-0.9650	165. S	0.1300	0.3400	0.2250	0.6950
95. I	0.1900	0.8400	0.1900	1.2200	166. C	-0.7000	-0.2500	0.3400	-0.6100
96. A	-1.3400	0.1300	0.0950	-1.1150	167. I	-0.6800	-0.4500	-0.1400	-1.2700
97. R	-1.0800	-0.2100	-0.3450	-1.6350	168. A	-0.4200	-0.2400	0.1750	-0.4850
98. A	-0.6100	0.1300	0.0350	-0.4450	169. T	-0.6200	0.2400	0.1850	-0.1950
99. K	-0.2900	-0.2100	0.0000	-0.5000	170. I	-0.4500	-0.4500	-0.3000	-1.2000
100. V	0.4900	-0.4700	0.1350	0.1550	171. A	-0.4200	-0.2400	0.1550	-0.5050
101. N	-0.5100	0.0100	0.1250	-0.3750	172. T	-0.7100	0.2400	0.2050	-0.2650
102. A	0.1200	0.1300	0.2700	0.5900	173. M	0.0700	0.5000	0.4500	-1.0200
103. G	-0.0300	0.2100	0.0200	0.2000	174. E	0.0700	0.6900	0.2450	1.0050
104. L	-0.3300	-0.4600	0.1800	-0.6100	175. E	-0.2900	-0.1100	0.2600	-0.1400
105. H	-0.5400	-0.0300	-0.2800	-0.8500	176. F	-0.2600	-0.4000	-0.0150	-0.6750
106. F	0.0300	0.0100	-0.0300	0.0100	177. F	-0.0400	-0.4000	-0.3200	-0.7600
107. D	0.3800	0.7800	0.2400	1.4000	178. P	-0.5900	-0.4900	0.0400	-1.0400
108. S	-0.6000	0.3400	-0.0700	-0.3300	179. M	-0.3500	0.1100	-0.0200	-0.2600
109. G	-0.5100	0.0200	-0.0400	-0.5300	180. D	0.3500	-0.4900	-0.4350	-0.5750
110. V	0.4000	0.1100	0.1950	0.7050	181. R	-0.6000	0.0400	0.1300	-0.4300
111. L	0.9500	-0.4600	0.1750	0.6650	182. S	0.2300	-0.3600	-0.1700	-0.3000
112. F	0.0500	0.0100	0.3350	0.3950	183. R	-0.3400	0.0400	-0.3250	-0.6250
113. G	0.0000	0.3400	0.0150	0.4050	184. Y	0.4500	0.0700	0.1500	-0.6700
114. R	-0.2100	0.3400	0.2950	0.4250	185. P	-0.2400	-0.4900	-0.0150	-0.7450
115. L	-0.7400	-0.4600	-0.0850	-1.2850	186. A	-0.1300	0.1300	0.1450	0.1450
116. R	-0.6200	0.3400	-0.2100	-0.4900	187. L	-0.9200	-0.4600	-0.1500	-1.5300
117. F	-0.2400	0.0100	0.1300	-0.1000	188. V	-0.7200	0.1100	0.0800	-0.5300
118. Y	-0.3600	-0.2700	-0.2400	-0.8700	189. A	-0.4500	0.1300	0.2450	-0.0750
119. F	0.2700	-0.4000	-0.3500	-0.4800	190. W	-0.7300	-0.3400	-0.1650	-1.2350
120. E	0.0100	-0.1100	0.0150	-0.0850	191. I	-0.4300	0.0300	-0.1800	-0.5800
121. P	0.0000	0.1900	-0.0850	0.1050	192. E	-0.6000	-0.1100	-0.1100	-0.8200
122. I	0.0100	-0.4500	0.4300	-0.0100	193. R	-0.9800	0.3400	-0.2550	-0.8950
123. L	0.5000	-0.0100	-0.3200	0.1700	194. L	0.0600	-0.4600	-0.1700	-0.5700
124. Y	0.3100	0.5700	0.0750	0.9550	195. G	0.4700	0.0000	0.0650	-0.3250
125. E	-0.4700	-0.1100	0.0150	-0.5650	196. R	-0.4200	0.0400	-0.0150	-0.3950
126. G	-0.5000	0.0200	0.0250	-0.4550	197. T	0.1900	-0.0200	0.2000	0.3700
127. S	-0.3000	0.0800	-0.2350	-0.4550	198. L	0.1700	-0.0100	0.1800	0.3400
128. A	-0.4400	0.1300	0.1550	-0.1550	199. P	-0.1300	-0.4900	-0.0950	-0.7150
129. E	0.1500	-0.3200	-0.2250	-0.3950	200. E	-0.5700	-0.1100	-0.0650	-0.7450
130. V	-0.6400	0.1100	0.0400	-0.4900	201. Y	-0.0200	-0.2700	-0.0900	-0.3800
131. P	-0.5900	-0.4900	0.1200	-0.9600	202. D	-0.1100	0.0200	0.1450	0.0550
132. Q	-0.5500	-0.2800	0.0500	-0.7800	203. Q	-0.1800	-0.2800	-0.1650	-0.6250
133. D	-0.6000	-0.4900	0.2800	-0.8100	204. L	0.2800	-0.4600	0.3650	0.1850
134. K	-0.7600	0.6300	0.2450	0.1150	205. N	0.0500	0.5200	0.5400	1.1100
135. I	0.2100	0.0300	0.1550	0.0250	206. Q	0.3200	0.0700	0.1050	-0.2050
136. D	-0.6500	0.0200	-0.6850	-1.3150	207. E	-0.4600	-0.1100	-0.0100	-0.5800
137. Y	-0.3600	-0.0700	0.0750	-0.3550	208. G	0.1400	0.2100	0.1250	0.4750
138. M	-0.8700	-0.5000	-0.2600	-1.6300	209. A	-0.3300	-0.2400	0.0050	-0.5650
139. K	-0.9700	0.6300	-0.6150	-0.9550	210. V	-1.0200	0.1100	0.5650	-0.3450
140. K	0.8200	-0.2100	-0.4950	0.1150	211. E	-0.0600	-0.1100	0.1850	0.0150
141. G	0.4800	0.2100	0.3150	1.0050	212. F	-0.7800	-0.4000	-0.0900	-1.2700
142. Y	0.3600	-0.2700	-0.2250	-0.1350	213. A	-0.7100	-0.2400	-0.0300	-0.9800
143. E	-0.2000	-0.1100	-0.7650	-1.0750	214. E	-0.6800	-0.1100	0.2100	-0.5800
144. L	-1.0500	-0.4600	0.2550	-1.2550	215. I	-0.6600	0.0300	-0.0200	-0.6500
145. L	-0.6000	-0.4600	-0.4650	-1.5250	216. C	-0.0200	-0.2500	0.0550	-0.2150
146. N	0.2600	0.5200	0.3050	1.1650	217. B	0.1100	0.1100	0.0200	-0.2000
147. D	0.3900	-0.4900	0.4150	0.3150	218. S	-0.1100	0.0800	0.0050	-0.0250
148. A	0.0900	0.1300	-0.1900	0.0300	219. L	0.2400	-0.0100	0.0450	0.2750
149. L	0.7000	-0.4600	0.0300	0.2700	220. R	0.6900	0.3400	-0.1050	0.9250
150. V	0.7200	0.6900	0.0800	1.4900	221. L	0.7900	1.0500	0.1850	2.0250
151. E	-0.1000	-0.1100	0.0400	-0.1700	222. K	0.9300	-0.1500	0.0800	0.8600
152. D	-0.3300	0.0200	-0.3750	-0.6850	223. N	0.3200	0.0100	0.1400	0.4700
153. Y	0.3300	-0.2700	0.2350	0.2950	224. M	0.2700	0.8000	0.0650	1.1350
154. I	0.8300	-0.4500	0.1200	0.5000	225. A	-0.0900	0.1300	-0.0800	-0.0400
155. A	-0.1500	-0.2400	0.0650	-0.3250	226. T	-0.0400	-0.1800	-0.0300	-0.2500
156. G	-0.4100	0.0200	-0.0100	-0.4000	227. N	-0.4100	-0.3700	0.0750	-0.7050
157. G	0.4400	0.3600	0.2750	-1.0750	228. P	0.3100	0.1200	0.0600	-0.0600
158. S	-0.4400	-0.3600	-0.1650	-0.9650	229. I	0.6000	0.0300	0.0350	0.6650
159. L	0.6100	-0.0100	-0.1200	0.4800	230. I	0.7900	-0.4500	0.0350	0.3750

231. K	-0.3700	0.6300	-0.3450	-0.0850
232. L	-0.1000	-0.4600	-0.5350	-1.0950
233. Y	1.3000	-0.2700	-0.1500	0.8800
234. T	-0.3100	0.2400	0.0950	0.0250
235. A	0.4100	0.1300	-0.1650	0.3750
236. K	0.4400	-0.2100	-0.2300	0.0000
237. L	0.6600	-0.0100	-0.3150	0.3350
238. S	0.1500	0.0800	0.0600	0.2900
239. P	0.4700	0.4200	-0.0550	0.8350
240. P	-0.6200	0.1900	-0.0050	-0.4350
241. G	-0.5900	0.2100	0.0000	-0.3800
242. R	0.7400	0.3400	0.0250	1.1050
243. A	-0.8400	-0.2400	0.2650	-0.8150
244. V	-1.0200	-0.4700	-0.2200	-1.7100
245. E	-0.2800	0.6900	-0.2100	0.2000
246. L	-0.4300	-0.4600	-0.0650	-0.9550
247. T	-0.5400	0.2400	0.6500	0.3500
248. A	-0.0600	-0.2400	-0.0400	-0.3400
249. K	0.4400	-0.2100	-0.0900	0.1400
250. L	0.6500	-0.0100	-0.2050	0.4350
251. L	-0.1000	-0.0100	-0.2450	-0.3550
252. G	-0.1900	-0.2100	0.0450	-0.3550
253. L	-0.3900	-0.0100	-0.3550	-0.7550
254. S	-0.3700	-0.3600	0.1350	-0.5950
255. L	0.3000	-0.0100	-0.4200	-0.1300
256. D	-0.2400	0.0200	0.1200	-0.1000
257. I	-0.2000	0.0300	0.0050	-0.1650
258. V	-0.5900	0.1100	-0.0300	-0.5100
259. P	-0.7300	-0.4900	0.0600	-1.1500
260. I	-0.0700	0.0300	-0.2250	-0.2650
261. N	-0.5600	-0.3700	-0.0300	-0.9600
262. L	0.1400	-0.0100	-0.3550	-0.2250
263. L	0.3800	1.0500	0.1000	1.5300
264. A	-0.2800	0.1200	0.1700	0.0100
265. C	-0.4900	0.0200	0.1800	-0.2900
266. D	-0.5200	0.0200	-0.0250	-0.5250
267. H	-0.4600	0.0000	0.0250	-0.4350
268. R	0.2300	0.0400	-0.1000	0.1700
269. T	-0.1300	-0.1800	-0.1400	-0.4500
270. D	-0.4200	-0.4900	0.0000	-0.9100
271. E	-0.1700	-0.3200	-0.3150	-0.8050
272. F	0.4700	0.0100	-0.0900	0.3900
273. L	0.1900	-0.0100	0.4150	0.5950
274. R	0.6900	0.0400	-0.5600	0.1700
275. L	0.1700	-0.0100	-0.1300	0.0300
276. N	-0.1600	0.5200	0.0550	0.4150
277. P	-0.4400	0.1900	0.1200	-0.1300
278. Q	-0.2800	-0.0700	-0.1250	-0.4750
279. H	-0.0700	-0.0300	-0.0750	-0.1750
280. T	0.3000	-0.0200	-0.1950	0.0850
281. I	0.0900	0.0300	-0.0800	0.0400
282. P	0.7700	0.4200	-0.1550	1.0350
283. V	-0.5700	-0.4700	-0.1550	-1.1950
284. I	-0.3300	-0.4500	-0.2250	-1.0050
285. D	0.4900	0.0200	0.1900	0.7000
286. D	-0.7200	0.0200	0.3000	-0.4000
287. G	-0.6200	-0.2100	0.2100	-0.6200
288. C	-0.2200	-0.2100	-0.0050	-0.4250
289. V	-0.5100	0.6900	0.2850	0.4650
290. I	-0.7800	0.8400	0.1500	0.2100
291. V	-0.2000	0.1100	-0.2800	-0.3700
292. R	0.1900	-0.2100	0.1100	0.0900
293. D	-0.2600	0.0200	0.0700	-0.1700
294. S	-0.6800	0.0800	0.2600	-0.3400
295. H	-0.3700	-0.0300	-0.0900	-0.4900
296. A	0.6300	-0.2400	0.1550	0.5450
297. I	-0.9300	-0.4500	-0.2050	-1.5850
298. I	-0.3400	-0.4500	-0.2650	-1.0550
299. I	-0.5200	0.0300	-0.4500	-0.9400
300. Y	-0.3200	-0.0700	-0.0700	-0.4600
301. L	-0.0200	-0.4600	-0.4550	-0.9350

302. V	0.7500	-0.4700	-0.2550	0.0350
303. Q	0.0400	-0.0700	0.0450	0.0150
304. K	-0.2800	-0.2100	-0.2350	-0.7250
305. Y	-0.5000	0.5700	-0.1650	-0.0950
306. G	-0.4200	0.0200	0.2000	-0.2000
307. K	-0.5300	-0.1500	0.0000	-0.6800
308. D	-0.6100	-0.4900	0.0000	-1.1000
309. G	-0.6400	-0.2100	-0.2000	-1.0500
310. Q	-0.3000	-0.0700	0.0450	-0.3250
311. T	0.4100	-0.1800	-0.0600	0.1700
312. L	0.5200	-0.0100	-0.2050	0.3050
313. Y	-0.2700	-0.2700	-0.0750	-0.6150
314. P	0.2400	0.1900	-0.1050	0.3250
315. E	-0.6800	-0.3200	-0.4350	-1.4350
316. D	-0.4800	-0.4900	0.0250	-0.9450
317. P	-0.4200	0.1900	-0.0500	-0.2800
318. I	0.1900	0.8400	0.1300	1.1600
319. A	-1.3400	0.1300	0.0950	-1.1150
320. R	-1.0800	0.3400	-0.3600	-1.1000
321. A	-0.6100	-0.2400	0.1000	-0.7500
322. K	-0.2900	-0.2100	0.0200	-0.4800
323. V	0.4900	-0.4700	0.1150	0.1350
324. N	-0.5100	0.5200	0.2950	0.3050
325. A	0.1900	-0.2400	0.3400	0.2900
326. G	-0.0300	0.2100	-0.0800	0.1000
327. L	-0.3300	-0.4600	0.3000	-0.4900
328. H	-0.5400	0.0000	-0.2300	-0.7700
329. F	0.0300	-0.4000	0.0950	-0.2750
330. D	0.3800	0.7800	0.2400	1.4000
331. S	-0.6000	0.3400	-0.0700	-0.3300
332. G	-0.5100	0.0200	-0.0850	-0.5750
333. V	0.4000	-0.4700	0.2350	0.1650
334. L	0.9500	-0.4600	0.1750	0.6650
335. F	0.0500	0.0100	0.3500	0.4100
336. S	0.0800	0.3400	-0.0050	0.4150
337. R	-0.2100	0.3400	0.2950	0.4250
338. L	-0.7400	-0.4600	-0.0850	-1.2850
339. R	-0.6200	-0.2100	-0.2100	-1.0400
340. F	-0.2100	0.0100	0.1300	-0.1000
341. Y	-0.3600	-0.2700	-0.2400	-0.8700
342. F	0.2700	-0.4000	-0.3500	-0.4800
343. E	0.0100	-0.1100	0.0300	-0.0700
344. P	0.0000	0.1900	-0.1100	0.0800
345. I	0.0100	-0.4500	0.4300	-0.0100
346. L	0.5000	-0.0100	-0.3200	0.1700
347. Y	0.3100	0.5700	0.0750	0.9550
348. E	-0.4700	-0.3200	0.0550	-0.7350
349. C	-0.5000	0.0200	0.0250	-0.4550
350. S	-0.3000	0.0800	-0.2600	-0.4800
351. A	-0.4400	0.1300	0.1550	-0.1550
352. E	0.1500	-0.3200	-0.2250	-0.3950
353. V	-0.6400	0.1100	0.0400	-0.4900
354. P	-0.5900	-0.4900	0.1200	-0.9600
355. Q	-0.5500	-0.2800	0.0500	-0.7800
356. D	-0.6000	0.0200	0.3050	-0.2750
357. K	-0.7600	0.6300	0.2450	0.1150
358. I	-0.2100	0.0300	0.1550	-0.0250
359. D	-0.6500	0.0200	-0.6850	-1.3150
360. Y	-0.3500	-0.2700	0.0450	-0.5850
361. M	-0.8700	-0.5000	-0.2600	-1.6300
362. K	-0.9700	0.6300	-0.6150	-0.9550
363. K	0.8200	0.6300	-0.5100	0.9400
364. G	0.4800	0.2100	0.3050	0.9950
365. Y	0.3600	-0.2700	-0.2250	-0.1350
366. E	-0.2000	0.6900	-0.7650	-0.2750
367. L	-1.0500	-0.4600	0.3600	-1.1500
368. L	-0.6000	-0.4600	-0.3700	-1.4300
369. N	0.2600	0.5200	0.3700	1.1500
370. D	0.3900	0.0200	0.3900	0.8000
371. A	0.0900	-0.2400	-0.1700	-0.3200
372. L	0.7000	-0.4600	0.0600	0.3000

373. V	0.7200	0.6900	0.1200	1.5300
374. E	-0.1000	-0.1100	0.0400	-0.1700
375. D	-0.3300	0.0200	-0.3750	-0.6850
376. Y	0.3300	-0.2700	0.2350	0.2950
377. I	0.8300	-0.4500	0.1200	0.5000
378. A	-0.1500	-0.2400	0.0650	-0.3250
379. G	-0.4100	0.0200	-0.0100	-0.4000
380. S	-0.4400	-0.3600	-0.2750	-1.0750
381. S	-0.4400	-0.3600	-0.1650	-0.9650
382. L	0.6100	-0.0100	-0.1200	0.4800
383. T	0.9100	0.2400	0.1250	1.2750
384. L	0.3800	-0.4600	-0.5650	-0.6450
385. A	-0.0700	-0.2400	0.0550	-0.2550
386. D	-0.2900	0.7800	0.8550	1.3450
387. V	0.3900	-0.4700	-0.5600	-0.6400
388. S	0.1300	0.3400	0.2250	0.6950
389. C	-0.7000	-0.2500	0.3400	-0.6100
390. I	-0.6800	-0.4500	-0.1400	-1.2700
391. A	-0.4200	-0.2400	0.1750	-0.4850
392. T	-0.6200	0.2400	0.1850	-0.1950
393. I	-0.4500	-0.4500	-0.3000	-1.2000
394. A	-0.4200	-0.2400	0.1550	-0.5050
395. T	-0.7100	0.2400	0.2050	-0.2650
396. M	-0.0700	-0.5000	-0.4500	-1.0200
397. E	0.0700	0.6900	0.2450	1.0050
398. E	-0.2900	-0.1100	0.2600	-0.1400
399. F	-0.2600	-0.4000	-0.0150	-0.6750
400. F	-0.0400	-0.4000	-0.3200	-0.7600
401. P	-0.5900	-0.4900	0.0400	-1.0400
402. M	0.3500	0.1100	0.0200	-0.2600
403. D	0.3500	-0.4900	-0.4350	-0.5750
404. R	0.6000	0.0400	0.1300	-0.4300
405. S	0.2300	-0.3600	-0.1700	-0.3000
406. R	0.3400	0.0400	0.3250	-0.6250
407. Y	-0.4500	-0.0700	-0.1500	-0.6700
408. F	0.2400	0.4200	0.0150	-0.7450
409. A	-0.1300	0.1300	0.1450	0.1450
410. L	-0.9200	-0.4600	-0.1500	-1.5300
411. V	-0.7200	0.1100	0.0800	-0.5300
412. A	-0.4500	0.1300	0.2450	-0.0750
413. W	-0.7300	-0.3400	-0.1650	-1.2350
414. I	-0.4300	0.0300	-0.1800	-0.5800
415. E	-0.6000	-0.1100	-0.1100	-0.8200
416. R	-0.9800	0.3400	-0.2550	-0.8950
417. L	0.0600	-0.4600	-0.1700	-0.5700
418. S	-0.4700	0.0800	0.0650	-0.3250
419. R	-0.4200	0.0400	-0.0150	-0.3950
420. T	0.1900	-0.0200	0.2000	0.3700
421. L	0.1700	-0.0100	0.1800	0.3400

422. F	-0.1300	-0.4900	-0.0950	-0.7150
423. E	-0.5700	-0.1100	-0.0650	-0.7450
424. Y	-0.0200	-0.2700	-0.0900	-0.3800
425. D	-0.1100	0.0200	0.1450	0.0550
426. Q	-0.1800	-0.2800	-0.1650	-0.6250
427. L	0.2800	-0.4600	0.3650	0.1850
428. N	0.0500	0.5200	0.5400	1.1100
429. Q	-0.3200	-0.0700	0.1050	-0.2850
430. E	-0.4600	-0.1100	-0.0100	-0.5800
431. G	0.1400	0.2100	0.1250	0.4750
432. A	-0.3300	-0.2400	0.0050	-0.5650
433. V	-1.0200	0.1100	0.5650	-0.3450
434. E	-0.0600	-0.1100	0.1850	0.0150
435. F	-0.7800	-0.4000	-0.0900	-1.2700
436. A	-0.7100	-0.2400	-0.0300	-0.9800
437. E	-0.6800	-0.1100	0.2100	-0.5800
438. I	-0.6600	0.0300	-0.0200	-0.6500
439. C	-0.9100	-0.2500	0.0550	-1.1050
440. E	-0.3800	-0.1100	0.0200	-0.4700
441. S	-0.6400	0.0800	0.0050	-0.5550
442. L	-0.1100	-0.0100	0.0450	-0.0750
443. R	-0.2800	0.3400	-0.1050	-0.0450
444. L	0.7900	1.0500	0.1850	2.0250
445. K	-0.2400	-0.1500	0.0800	-0.3100
446. N	0.0000	0.0100	0.1400	0.1500