Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

**Bioinformática Estrutural Aplicada à Evolução das Glutathionas
Transferases**

Andrea Guelfi

**Tese apresentada para obtenção do título de
Doutor em Agronomia. Área de concentração:
Genética e Melhoramento de Plantas**

Piracicaba
2006

Andrea Guelfi
Engenheiro Agrônomo

Bioinformática Estrutural Aplicada à Evolução das Glutathionas Transferases

Orientador:

Prof. Dr. RICARDO ANTUNES DE AZEVEDO

Tese apresentada para obtenção do título de Doutor em Agronomia.
Área de concentração: Genética e Melhoramento de Plantas

Piracicaba
2006

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

Guelfi, Andrea

Bioinformática estrutural aplicada à evolução das Glutathionas transferases /
Andrea Guelfi. - - Piracicaba, 2006.
107 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura Luiz de Queiroz, 2006.
Bibliografia.

1. Bioinformática 2. Glutathione transferase 3. Proteínas – Evolução I. Título

CDD 574.1925

“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”

*Aos meus pais Walter e Selles, pelos
vinhos, jantares e programas
inesquecíveis*

Agradecimentos

Ao Prof. Dr. Ricardo A. Azevedo pela orientação e apoio durante a realização deste trabalho;

À Profa. Dra. Silvia M. G. Molina pela amizade;

Ao Dr. Goran Neshich e Dra. Paula K. Falcão, pesquisadores da Embrapa, por todo o suporte na área de bioinformática estrutural;

À Dra. Sabrina M. Chabregas, pesquisadora do Centro de Tecnologia Canavieira, pelo irrestrito apoio. Bem como aos doutores Eugênio C. Ulian e Maria Cristina Falco e às técnicas Rose e Daniela;

Ao Prof. Dr. Giancarlo C. X. Oliveira pela orientação e pelas impagáveis conversas;

Ao Prof. Dr. Márcio de Castro Filho pelas enriquecedoras discussões;

Ao Prof. Dr. Cláudio Lopes, pela oportunidade e apoio fornecido nos momentos difíceis;

À Dra. Salete Gaziola do Laboratório de Genética Bioquímica de Plantas da ESALQ;

Às amigas Lígia, Ana Paula e Luana do Lab. de Ecologia Evolutiva e aos amigos do Lab. de Genética de Microrganismos, Wellington, Joelma, Cris Maki, Ágata, Júlia, Lacava, Carol, Pipa, Fernando, Bia, Pricila e Zezo;

Aos professores do Departamento de Genética pela colaboração. Aos funcionários da ESALQ/USP que de alguma forma contribuíram para a realização deste trabalho, particularmente Léia, Neusa, Glória, Valdir e Berdan;

Ao CNPq pela concessão da bolsa de estudo.

SUMÁRIO

RESUMO	7
ABSTRACT	8
LISTA DE FIGURAS	9
LISTA DE TABELAS	11
LISTA DE ABREVIATURAS.....	12
LISTA DE SÍMBOLOS	14
1 INTRODUÇÃO.....	15
2 REVISÃO BIBLIOGRÁFICA.....	17
2.1 A superfamília das glutathionas transferases.....	17
2.2 Evolução das proteínas	25
2.3 Filogenia e a estrutura tridimensional das proteínas	28
3 MATERIAL E MÉTODOS.....	30
3.1 Bancos de dados consultados	30
3.2 Alinhamentos de seqüência e de estrutura tridimensional.....	31
3.3 Reconstrução filogenética.....	32
3.4 Modelagem de proteínas por homologia	32
3.5 Modelagem molecular	33
3.5.1 Docking molecular	34
3.5.2 Dinâmica molecular.....	35
3.6 Contatos da interface enzima - substratos	36
3.7 Visualização das moléculas.....	37
3.8 Construções gênicas	37
3.9 Cepas bacterianas	42
3.10 Análise dos transformantes (“Screening” por PCR e restrição enzimática).....	42
3.11 Seqüenciamento das construções gênicas	43
3.12 Meios de cultura	44
4 RESULTADOS	45
4.1 Amostragem das GSTs nos diversos filos taxonômicos.....	45
4.2 Reconstrução filogenética baseada no alinhamento de seqüências de GSTs	45
4.3 Reconstrução filogenética baseada no alinhamento tridimensional.....	48
4.4 Avaliação geral do alinhamento baseado na estrutura tridimensional	48

4.5 Diferenças estruturais entre as isoformas da GST classe Phi.....	52
4.6 Alinhamento estrutural da GST classe Phi e classe Pi	54
4.7 Re-docking da glutatona na estrutura terciária da classe Pi	56
4.8 Docking da glutatona na estrutura terciária da classe Phi	57
4.9 Transferência de elétrons nas GSTs classes Phi e Pi.....	59
4.10 Re-docking do ácido etacrínico na estrutura terciária da classe Pi.....	60
4.11 Docking do ácido etacrínico na estrutura terciária da classe Phi	61
4.12 Modelagem da GST de cana-de-açúcar e escolha dos mutantes	64
4.13 Dinâmica molecular na GST de cana-de-açúcar e nos mutantes.....	67
4.14 Construções gênicas em vetor para expressão das proteínas heterólogas em bactéria.....	72
5 DISCUSSÃO	74
5.1 Padrão da superfamília das GSTs.....	74
5.2 Evolução da superfamília das GSTs.....	77
5.3 Convergência funcional e a relação tridimensional entre as classes Phi e Pi.....	81
5.3.1 Atividade enzimática relativa às classes Phi e Pi	81
5.3.2 Diferenças estruturais entre as GSTs classe Phi e classe Pi	85
5.4 Modelagem molecular da GST de cana-de-açúcar.....	90
6 CONCLUSÕES.....	91
REFERÊNCIAS	92

RESUMO

Bioinformática Estrutural Aplicada à Evolução das Glutathionas Transferases

As glutathionas transferases compõem uma superfamília de proteínas que atuam na fase II do sistema de desintoxicação das células. Participam principalmente através do processo de conjugação da glutathione com moléculas hidrofóbicas e eletrofílicas, como por exemplo os herbicidas. No entanto, outras funções foram descritas como a tolerância ao estresse oxidativo, inseticidas, antibióticos microbianos, transporte de produtos secundários tóxicos, sinalização da célula durante as respostas ao estresse e fenômenos de resistência envolvendo agentes de quimioterapia contra o câncer. Nesta tese procurou-se estabelecer uma relação entre a sequência, estrutura, função e afinidade das GSTs. A estrutura, de modo geral, determina a função da enzima, mas por si só, não dita sua especificidade. Esta última informação é fundamental para o desenvolvimento de novos agroquímicos ou para o desenho racional de novas proteínas. A relação entre a sequência, estrutura, função e afinidade mostra que o paradigma estrutura-função deveria ser ampliado para incluir a sequência de aminoácidos e a afinidade da enzima. Apesar da grande diversidade de substratos e sequências encontradas nas GSTs há pelo menos um caso de convergência funcional em duas classes distintas desta superfamília. Uma encontrada apenas no reino Animalia (classe Pi) e outra exclusiva do reino Plantae (classe Phi). Ferramentas da bioinformática estrutural, como docking molecular e minimização de energia foram utilizadas para analisar as interações entre a enzima e o substrato. Estas ajudam a explicar como duas proteínas com aproximadamente 22% de identidade de sequência apresentam afinidades semelhantes. Finalmente, foram propostos mutantes da GST de *Saccharum officinarum* utilizando a informação estrutural da enzima, visando uma alteração na afinidade da mesma.

Palavras-chave: Reconstrução Filogenética; Alinhamento Estrutural; Superfamília de Proteínas; Interação Proteína-Ligante.

ABSTRACT

Structural Bioinformatics Applied to the Evolution of Glutathione Transferases

Glutathione Transferases comprehend a superfamily of proteins that plays the phase II of the detoxification system of the cells. Their major catalysis is the conjugation of glutathione with hydrophobic and eletrophilic molecules, for example herbicides. However, other functions were described like oxidative stress, insecticides, microbial antibiotics, transport of secondary products, cells signalization during response to stress and resistance of chemotherapy drugs against cancer. This work aimed to establish a relation between sequence, structure, function and affinity of GSTs. The structure, in general, determines the function, but alone, can not determine the enzyme specificity. This last information is essential to the development of new agrochemicals or for the rational design of proteins. The relation between sequence, structure, function and affinity shows that the paradigm of structure-function should be enlarged in order to include the information of amino acid sequences and the enzyme affinities. Despite the wide variety of substrates and sequences found in the GSTs, there is at least one case of functional convergence between two distinct classes in this superfamily. One is found in the Animalia kingdom (class Pi) and the other is exclusively found in Plantae (class Phi). The structural bioinformatic tools, such as molecular docking and energy minimization were used to analyze the interactions between the enzyme and the substrate. These help to understand how two enzymes with approximately 22% of sequence identity can show the same affinities. Finally, GST mutants of *Saccharum officinarum* were proposed, using the enzyme structural information in order to modify the enzyme affinities.

Key-words: Phylogenetic Reconstruction, Structural Alignment; Protein Superfamily; Protein-Ligand Interaction.

LISTA DE FIGURAS

- Figura 1 - Esquema da construção pET28b-GST38
- Figura 2 - Esquema do plasmídeo pET28b.39
- Figura 3 - Reconstrução filogenética baseada na sequência de aminoácidos. Foram identificados os gêneros e filos dos organismos e as classes da enzima GST47
- Figura 4 - Reconstrução filogenética baseada na estrutura tridimensional das GSTs. Os códigos de quatro letras representados são decorrentes dos códigos PDB de cada estrutura amostrada.....49
- Figura 5 - Região N-terminal do alinhamento de GSTs proveniente dos dados estruturais das enzimas, realizado no programa PrISM (YANG; HONIG, 1999). As letras à esquerda indicam o código da estrutura no banco de dados do PDB.49
- Figura 6 - Ilustração da cadeia principal referente ao alinhamento estrutural das classes de GSTs. (A) alinhamento total; (B) classes Alpha (verde), Sigma (branco), Pi (vermelho) e Mu (laranja), as setas indicam peculiaridades das classes Alpha (seta verde) e Mu (seta laranja); (C) classes Zeta (azul), Tau (verde) e Omega (laranja), a região N-terminal foi representada com pontilhado; (D) classes Theta (azul), Delta (laranja) e Phi (verde), seta indica alpha-hélice da classe Theta.51
- Figura 7 - (A) alinhamento estrutural da região N-terminal da *ZmGSTF1* (verde), *ZmGSTF3* (azul), *AtGSTF4* (laranja). As posições e resíduos indicados se referem à *ZmGSTF1*. O retângulo indica a His40 da isoforma *ZmGSTF3*; (B) principais resíduos do sítio ativo das isoformas da *ZmGSTF1* (cinza) e *AtGSTF4* (laranja) alinhadas estruturalmente; (C) ampliação da Figura 6B, mostrando somente o sítio ativo das enzimas alinhadas; (D) alinhamento estrutural da isoforma *AtGSTF4* complexada (cadeia cor amarela e superfície marrom) e da apoforma (superfície representada em verde e cadeia na cor vermelha). O herbicida fluconazole conjugado com a GSH foi representado na cor vermelha.53
- Figura 8 - (A) Representação das cadeias das enzimas *HmGSTP1* (laranja) e *ZmGSTF1* (cinza), alinhadas estruturalmente; (B) Ampliação do sítio ativo das enzimas alinhadas. O conjugado ATA-GSH se encontra no sítio ativo, com as cores CPK....55
- Figura 9 - Alinhamento estrutural da *HmGSTP1* (laranja) e *ZmGSTF1* (azul). (A) A *HmGSTP1*, apresenta uma extensão na região C-terminal (Asn204) e a *ZmGSTF1* apresenta duas Gly (123 e 124) numa região de loop. (B) Sítio ativo das enzimas *HmGSTP1* (formato “bond”) e *ZmGSTF1* (formato “bola-linha”). Na cor branca foi representado o EA.55
- Figura 10 - Resíduos que participam da interação da GSH (representada com o formato VDW) com a GSTP1 (A₁) e GSTF1 (B₁). Os resíduos foram representados no formato “bond”. Comparação da possível transferência de elétrons na GSTP1 (A₂) e GSTF1 (B₂). As pontes de hidrogênio foram representadas com linhas pontilhadas. A GSH e os resíduos foram representados com o formato “bond”.58

- Figura 11 - (A₁ e A₂) Re-docking do ácido etacrínico (EA, representado no formato VDW) na estrutura *HsGSTP1*, na presença da GSH (representada no formato linha-bola); (B₁ e B₂) Docking seguido de minimização de energia do EA na enzima *ZmGSTF1*, com a presença da GSH. 63
- Figura 12 - Alinhamento da GST de milho (*Z. mays*) e de cana-de-açúcar (*S.offic*), os resíduos em negrito indicam participação no sítio ativo da enzima, particularmente no sítio H. 65
- Figura 13 - (A) Modelo da forma dimérica da GST de cana-de-açúcar (*SoGSTF1*). (B) Monômero da proteína modelada *SoGSTF1*, após a alocação do conjugado atrazine-glutationa (ATA-GSH) no sítio ativo da enzima. (C) Monômero da proteína modelada *SoGSTF1*, salientando os resíduos do sítio ativo e o conjugado ATA-GSH. (D) Visualização dos principais resíduos participantes do sítio ativo da enzima modelada *SoGSTF1*, com o conjugado atrazine-glutationa (destacado no quadro à direita) 66
- Figura 14 - Energia potencial (kJ.mol⁻¹) no decorrer do tempo (ps) da enzima controle (A1) e dos mutantes W12L (B1) e I118F (C1). O desvio RMS da proteína (–) e da ATA (↔) foram representados, no controle (A2) e nos mutantes W12L (B2) e I118F (C2). 68
- Figura 15 - Contatos do conjugado ATA-GSH (formato “linha”) com a enzima controle (A₁ e A₂), mutante Trp12Leu (B₁ e B₂) e Ile118Phe (C₁ e C₂) no tempo de 750 e 1000 ps. Os resíduos de cada enzima foram representados no formato “linha-bola” 71
- Figura 16 - Análise por PCR das colônias transformadas. PM, marcador molecular (1kb); colônias de 1-5 referentes aos mutantes W12L; colônias de 1-7 referentes aos mutantes I118L; CP, controle positivo; CN, controle negativo. 73
- Figura 17 - Análise de restrição dos fragmentos referentes aos mutantes selecionados. CP, controle positivo; PM, marcador molecular (1kb). 73
- Figura 18 - Relações encontradas entre as seqüências, estruturas, funções e afinidades das enzimas. Padrões de evolução encontrados na superfamília das GSTs, divergência funcional (esquerda) e convergência funcional (direita). 74
- Figura 19 - Filos da vida na Terra baseados no sistema dos cinco reinos e na teoria simbiótica da origem das células eucarióticas (MARGULIS; SCHWARTZ, 2001). 79
- Figura 20 - Possível rota evolutiva encontrada na superfamília das GSTs. As GSTs do reino Plantae foram representadas com o fundo azul, o reino Animalia, em roxo, o reino Fungi foi identificado com a cor laranja e GSTs de Eubacteria em rosa. 80
- Figura 21 - Duas hipóteses evolutivas para as classes Pi e Phi de GSTs. As letras A, B e C correspondem a genes de GST. 80

LISTA DE TABELAS

Tabela 1 - Número de genes nas diferentes classes de GST em plantas	21
Tabela 2 - Descrição das classes de GSTs de plantas, com respectivos ligantes.	24
Tabela 3 - Interações entre enzima GSTP1 com a GSH, obtida por cristalografia e pelo método de docking (LIGIN; SOBOLEV et al., 1996).....	56
Tabela 4 - Interações entre enzima <i>At</i> GSTF4 com a GSH, obtida por cristalografia e a <i>Zm</i> GSTF1 pelo método de docking (SOBOLEV et al., 1996).....	57
Tabela 5 - Interações entre enzima <i>Hs</i> GSTP1 com o EA, obtido por cristalografia e pelo método de docking (SOBOLEV et al., 1996).....	60
Tabela 6 - Comparação entre os contatos do EA com a enzima <i>Zm</i> GSTF1 utilizando-se o método de docking e docking seguido de minimização de energia (EM).....	62
Tabela 7 - Distância e tipo de interações dos átomos da ATA com os resíduos da enzima controle no tempo de 750 e 1000 ps (onde, HB = interação de hidrogênio e Arom. = interação aromática).....	70
Tabela 8 - Distância e tipo de interações dos átomos da ATA com os resíduos da enzima mutante Trp12Leu no tempo de 750 e 1000 ps	70
Tabela 9 - Distância e tipo de interações dos átomos da ATA com os resíduos da enzima mutante Ile118Phe no tempo de 750 e 1000 ps (DC = contatos desestabilizador).....	70
Tabela 10 - Ordem de grandeza da atividade das classes de GST para o herbicida atrazine em diferentes espécies, em ordem decrescente	82
Tabela 11 - Ordem de grandeza da atividade decrescente das classes de GST para o ácido etacrínico em diferentes espécies	84
Tabela 12 - Comparação dos resíduos do sítio G que interagem com os íons da GSH, nas GSTF1 e GSTP1.....	85
Tabela 13 - Posições correspondentes à observação do alinhamento estrutural das enzimas GSTF1 e GSTP1.....	87
Tabela 14 - Comparação dos contatos do EA entre as GSTP1 (<i>Hs</i> GSTP1) e GSTF1 (<i>Zm</i> GSTF1) (HB = interação de hidrogênio, Arom. = interação aromática, Hidrof. = interação de hidrofobicidade e DC = contato desestabilizador).....	88

LISTA DE ABREVIATURAS

Ala = alanina, (A);
Arg = arginina, (R);
Arom = interação aromática;
Asn = asparagina, (N);
Asp = ácido aspártico, (D);
ATA = atrazine (2-cloro-4-etilamino-6-esopropilamino-s-triazine);
CDNB = 1-cloro-2,4-dinitrobenzeno;
CF = função de complementariedade;
CPK = representação dos átomos segundo Corey–Pauling–Koltun;
Cys = cisteína, (C);
DC = contatos desestabilizadores;
DDBJ = DNA “Data Bank of Japan”;
desvio RMS = desvio do quadrado médio da raiz;
EA = ácido etacrínico;
EBI = “European Bioinformatics Institute”;
EC = Classificação Enzimática, “Enzyme Classification”;
EPNP = 3-(4-nitro)-fenoxi-1,2-epoxipropano;
Fluorodifen = herbicida nitrofenil-eter fluconazole;
Gln = glutamina, (Q);
Glu = ácido glutâmico, (E);
Gly = glicina, (G);
GPx = glutational peroxidase;
GSH = glutational reduzida (gamma-Glu-Cys-Gly);
GST(s) = glutational(s) S-transferase(s);
HB = interação de hidrogênio;
Hidrof. = interação de hidrofobicidade;
His = histidina, (H);
Ile = isoleucina, (I);
Leu = leucina, (L);
LPC = contatos do ligante, “Ligand Pocket Contacts”;

Lys = lisina, (K);

MDM = mini-cromossomos;

Met = metionina, (M);

NBD-Cl = 7-cloro-4-nitrobenzeno-2-oxa-1,3-diazol;

NMR = Ressonância Magnética Nuclear;

PDB = “Protein Data Bank”;

Phe = fenilalanina, (F);

PIR = “Protein Identification Resource”;

Pro = prolina, (P);

ROS = Espécies Reativas de Oxigênio;

Se = átomo de Selênio;

Ser = serina, (S);

SCOP = “Structural Classification of Proteins”;

Thr = treonina, (T);

Trp = triptofano, (W);

Tyr = tirosina, (Y);

Val = valina, (V);

VDW = representação dos átomos baseada nas distâncias de Van der Waals.

LISTA DE SÍMBOLOS

Unidades de tempo:

ns = 10^{-9} segundo;

ps = 10^{-12} segundo;

Unidades de distância:

nm = 10^{-9} metro;

Å = 10^{-10} metro;

Unidade de área:

Å² = $(10^{-10})^2$ metro quadrado.

1 INTRODUÇÃO

As proteínas são as principais responsáveis pelo funcionamento de um organismo, seja este um fungo, planta ou mamífero. Dentre estas estão as enzimas, as quais apresentam afinidades e especificidades diferentes nos mais diversos tipos de funções encontrados numa célula. A classificação destas pode se dar pela atividade que exercem ou de acordo com a sua estrutura tridimensional. A classificação de acordo com a função pode ser encontrada através do número de classificação das enzimas - EC (SwissProt) e pode ser do tipo isomerases, oxidases, transferases, hidrolases etc. A classificação de acordo com sua estrutura tridimensional pode ser encontrada no banco de dados do “Structural Classification of Proteins” - SCOP, com as seguintes subdivisões: classes, famílias, superfamílias e conformação geral (“fold”). De modo que há uma relação entre a classificação estrutural (superfamílias) e a classificação funcional (número EC). Por exemplo, a reação enzimática das glutathionas transferases (EC 2.5.1.18) promove a conjugação da glutathione com um substrato hidrofóbico eletrofílico, numa reação nucleofílica, liberando o produto gerado. Normalmente, o intuito desta reação é promover a solubilização do composto formado, de modo a facilitar sua eliminação da célula (metazoários) ou o armazenamento no vacúolo (plantas).

Nesta tese procurou-se estabelecer uma relação entre a sequência, estrutura, função e afinidade na superfamília das glutathionas transferases (GSTs). Deste modo, foi apresentada uma relação mais ampla do que somente o envolvimento da estrutura e função de uma enzima. A superfamília das GSTs é formada por sequências distintas de proteínas, com estruturas tridimensionais muito semelhantes, uma mesma função básica de conjugar a glutathione (GSH) com compostos xenobióticos hidrofóbicos eletrofílicos, mas apresenta especificidades e afinidades muito diversas entre as classes encontradas. Ou seja, a estrutura, de modo geral, determina a função da enzima mas, por si só, não dita sua especificidade. Esta última informação é fundamental para o desenvolvimento de novos agroquímicos ou para o desenho racional de novas proteínas. Assim, não somente a estrutura, mas também a sequência, com aminoácidos em posições-chave (muitas vezes no próprio sítio catalítico) apresenta um papel preponderante na afinidade específica da enzima. No entanto, ainda há um longo caminho a percorrer no que concerne a este tópico. São necessárias análises pontuais com cada sequência para cada substrato específico para inferir sobre uma possível relação entre enzima-substrato.

As ferramentas utilizadas nesta tese podem servir como inspiração para o desenho racional de novas enzimas, as quais poderiam ser obtidas posteriormente através da engenharia genética. Esta metodologia pode servir como opção para o demorado e caro processo de desenvolvimento de novos agroquímicos. Deste modo, ao invés de desenvolver novas moléculas químicas que atuem, por exemplo, como herbicidas, alteramos a proteína, para que a mesma aumente (ou diminua) sua atividade. Além disso, outra aplicação poderia ser na utilização de enzimas como biosensores, um papel muito relevante na área de biotecnologia, a qual poderia fazer uso da modelagem de proteínas visando uma maior afinidade pelo substrato.

Hipóteses:

- a. A superfamília das GSTs apresenta seqüências de proteínas diferentes com estruturas tridimensionais semelhantes, a mesma função básica (conjugação de um composto hidrofóbico com a glutatona), mas apresentam afinidades diferentes;
- b. Como a estrutura tridimensional é mais conservada que a seqüência de proteínas, uma menor amostragem de estruturas utilizada na reconstrução filogenética pode refletir a reconstrução filogenética baseada no alinhamento de seqüências;
- c. Há pelo menos um caso de convergência funcional em duas classes distintas da superfamília de GSTs, uma encontrada no reino Animalia (classe Pi) e outra exclusiva do reino Plantae (classe Phi);
- d. Enzimas parálogas, com seqüências apresentando apenas 22,3 % de identidade, podem apresentar afinidades semelhantes pelos mesmos substratos. Esta demonstração foi feita utilizando-se as ferramentas da bioinformática estrutural, como a alocação de substratos no sítio ativo da enzima (docking molecular) e a minimização de energia, além de outros programas capazes de analisar as interações entre a enzima e o substrato (chamado também de ligante).

2 REVISÃO BIBLIOGRÁFICA

2.1 A superfamília das glutathionas transferases

As glutathionas transferases (GSTs) foram primeiramente descobertas nos animais em 1961 (BOOTH; BOYLAND; SIMS, 1961; COMBES; STAKELUM, 1961) como responsáveis pelo metabolismo de desintoxicação de drogas (WILCE; PARKER, 1994). Nas plantas foi detectada em 1970 quando foi descoberta sua importância na conjugação da glutathiona (GSH) com a atrazine (2-cloro-4-etilamino-6-esopropilamino-s-triazine) em milho, protegendo a planta deste herbicida (FREAR; SWANSON, 1970). A desintoxicação da molécula de atrazine pela GST é uma resposta primária no processo de desintoxicação de herbicidas pelas plantas (SHIMABUKURO; SWANSON; WALSH, 1970). No entanto, outras funções foram descritas como a tolerância ao estresse oxidativo (KAMPRANIS et al., 2000), aos inseticidas (RANSON; PRAPANTHADARA; HEMINGWAY, 1997; TANG; TU, 1994) e aos antibióticos microbianos (ARCA; HARDISSON; SUAREZ, 1997), transporte de produtos secundários tóxicos (MUELLER et al., 2000), sinalização da célula durante as respostas ao estresse (LOYALL et al., 2000) e fenômenos de resistência envolvendo agentes de quimioterapia contra o câncer (McLELLAN; WOLF, 1999; TEW, 1994).

O sistema de desintoxicação das células apresenta três fases principais. A fase I do sistema é catalisada pela enzima do sistema citocromo P450. A fase II é caracterizada pela catálise das glutathionas (gamma-Glu-Cys-Gly, GSH), com a participação das GSTs. O propósito deste sistema baseia-se no aumento de solubilidade do composto hidrofóbico de modo a facilitar seu transporte. A fase III corresponde à eliminação do composto tóxico para fora da célula, em animais, ou para o vacúolo em plantas (SHEEHAN et al., 2001).

As GSTs catalisam a reação do tripeptídeo nucleofílico glutathiona (GSH) com um eletrófilo hidrofóbico, xenobiótico ou endógeno, os quais podem ser tóxicos e/ou carcinogênicos para o organismo (AMES; PROFET; GOLD, 1990; MEISTER; ANDERSON, 1983; SALINAS; WONG, 1999). Além disso, essa família de enzimas contém classes de enzimas responsáveis pela remoção de espécies reativas de oxigênio (ROS), mais especificadamente catalisam a reação do H_2O_2 em $2H_2O$, sem a presença do Se (Selênio). Esta mesma reação é catalisada pela glutathiona peroxidase (GPx; EC. 1.11.1.9) na presença do selênio (Se) (SHEEHAN et al., 2001).

As GSTs (EC 2.5.1.18) compõem uma família de proteínas solúveis que apresentam massa molecular ao redor de 25kDa para cada subunidade. Apresentam-se normalmente na forma dimérica. São divididas em doze principais classes: Alpha, Mu, Pi, Omega (presentes em metazoários), Phi, Tau e Lambda (presentes nas plantas), Theta e Zeta (encontradas em plantas e metazoários), Sigma (Mollusca e mamíferos), Delta (insetos) e Beta (bactérias) (SHEEHAN et al., 2001). As estruturas tri-dimensionais destas classes foram determinadas excetuando-se a classe Lambda.

A nomenclatura das classes de GSTs é representada com as letras do nome do organismo, em itálico, seguida pelo nome da enzima, a classe da mesma e os números correspondendo às isoformas dos dímeros. Assim a classe Phi é representada pela letra F, Theta por T, Tau por U e Zeta por Z. Por exemplo, a enzima *GmGSTF1-1*, representa uma enzima GST de soja (*Glycine max*), da classe Phi (F) e os números representam o tipo de subunidade dos dímeros. Homodímeros são representados com o mesmo número, enquanto heterodímeros são representados com números diferentes (DIXON; LAPTHORN; EDWARDS, 2002).

Cada subunidade da estrutura dimérica apresenta um sítio ativo com cinéticas independentes (DANIELSON; MANNERVIK, 1985). Cada sítio ativo, por sua vez, é subdividido em dois outros sítios. O sítio G, ou sítio de ligação do tripeptídeo glutationa, localizado no N-terminal da proteína, o qual apresenta uma seqüência conservada de aminoácidos é altamente específico. O sítio H, ou sítio de ligação do substrato eletrofílico hidrofóbico, está localizado no C-terminal. Esta estrutura é mais flexível que a observada no sítio G. Particularmente, nessa região ligam-se diferentes substratos, como xenobióticos (p.e. herbicidas) ou compostos tóxicos endógenos (p.e. H_2O_2) (NEUEFEIND et al., 1997a; REINEMER et al., 1996). Apesar das subunidades serem independentes cataliticamente, elas existem na sua maioria na forma dimérica. A razão para este evento permanece ainda desconhecida (DIXON; LAPTHORN; EDWARDS, 2002). Aparentemente, a importância da forma dimérica consiste na manutenção da estabilidade dos dois domínios para cada subunidade da proteína (ARMSTRONG, 1997). Embora haja pouca similaridade nas seqüências de GSTs entre as diferentes classes, há uma conservação muito alta na estrutura geral da enzima (DIXON; LAPTHORN; EDWARDS, 2002).

As GSTs são organizadas em classes diferentes em função de suas características estruturais. Serão abordadas nesta revisão as GSTs citossólicas que tiveram suas estruturas tridimensionais resolvidas, seja por cristalografia (raio-X), seja por Ressonância Magnética Nuclear (NMR). De

modo geral, as GSTs são formadas na região N-terminal por três alpha-hélices e quatro folhas beta-pregueadas e na região C-terminal somente por alpha-hélices (geralmente de quatro a cinco hélices). A seguir serão resumidas algumas das principais características funcionais e estruturais das classes de GST. A classe Beta foi descrita numa ampla gama de gêneros de bactérias, como *Escherichia*, *Proteus*, *Pseudomonas*, *Klebsiella*, *Enterobacter*, *Serratia*, *Burkholderia* e *Rhodococcus*. Essa classe de GST apresenta a capacidade de se ligar a antibióticos, promovendo uma das formas de resistência das bactérias (FAVALORO et al., 1998).

A classe Sigma apresenta estruturas em *Drosophila melanogaster* (filó Mandibulata), *Ommastrephes sloani* (Mollusca) e *Schistosoma haematobium* (Platyhelminthes) e apresenta de 42 a 44% de similaridade com a proteína S-cristalino presente na lente ocular de moluscos (TOMAREV; CHUNG; PIATIGORSKY, 1995). A observação da estrutura tridimensional revelou que essa classe apresenta características próprias, especialmente com relação à interface do dímero da GST (JI et al., 1995). Esta classe também apresenta a GSH como cofator e está envolvida no mecanismo da prostaglandina-sintase (JOHNSON et al., 2003).

A classe Omega foi encontrada, até o momento, nos mamíferos, cuja estrutura foi cristalizada, e em *Caenorhabditis elegans*. A GSTO1-1 em humanos, pertencente a esta classe apresenta uma atividade tiol-transferase dependente da glutatona e catalisa a redução do deidroascorbato. Esta atividade não é comum às GSTs mas encontrada nas glutaredoxinas. A classe Omega tem uma extensão N-terminal única nas GSTs e a estrutura cristalina revelou uma cisteína (Cys) no sítio ativo diferente da tirosina (Tyr) ou serina (Ser) das demais GSTs. Resíduo semelhante foi encontrado na GST de bactéria, com uma Cys no sítio ativo. Esta classe apresenta baixa atividade na presença do 1-cloro-2,4-dinitrobenzeno (CDNB), o qual é um substrato típico das classes Alpha, Mu e Pi (BOARD et al., 2000).

As estruturas pertencentes às classes Alpha, Mu e Pi apresentam algumas semelhanças, principalmente quando comparadas com outras classes de GST (DIRR; REINEMER; HUBER, 1994). Há uma semelhança clara na estrutura destas enzimas principalmente no domínio N-terminal. A interação entre as subunidades do dímero é hidrofóbica e ocorre entre o domínio N-terminal de uma subunidade com o domínio C-terminal da outra subunidade (ARMSTRONG, 1997). A chave para esta ligação é a presença de um resíduo aromático (Phe52 na classe Alpha, Phe56 na classe Mu e Tyr49 na classe Pi). Estas classes têm uma interface de aproximadamente 25 Å X 35 Å (SINNING et al., 1993). Um total de 26 resíduos é conservado nas três classes, o

que representa de 20 a 32% de identidade de sequência (SHEEHAN et al., 2001). Lo Bello et al. (2001), Parraga et al. (1998) e Vega et al. (1998) especificaram que estas classes apresentam tirosina (Tyr7) como resíduo responsável pela catálise da enzima, a qual estabelece ponte de hidrogênio com a Cys da GSH. Embora a região N-terminal seja bem conservada nestas classes, o mesmo não ocorre na região C-terminal (DIRR; REINEMER; HUBER, 1994; SINNING et al., 1993). As classes Alpha, Mu e Pi apresentam algumas diferenças, por exemplo, na classe Mu há a presença de um “loop μ ” na região N-terminal entre a segunda folha beta (beta-2) e a segunda alpha-hélice (alpha-2) (JI et al., 1992), o que resulta numa cavidade mais profunda que a encontrada na classe Pi da GST (WILCE; PARKER, 1994). O domínio C-terminal da GST Alpha contém uma alpha-hélice extra, a qual resulta num sítio hidrofóbico menor e mais estreito comparado com o maior e mais aberto sítio das classes Mu e Pi (SINNING et al., 1993). A classe Alpha foi descrita em *Schistosoma japonicum* (Platyhelminthes), *Fasciola hepatica* (Platyhelminthes) e mamíferos e possui também atividade semelhante à enzima glutathione peroxidase (BOARD et al., 1997). Além desta, apresenta uma atividade de esteróide isomerase em ovários e testículos de ratos (SHEEHAN et al., 2001) e uma afinidade específica pelo substrato 7-cloro-4-nitrobenzeno-2-oxa-1,3-diazol (NBD-Cl) (BOARD et al., 2000). A classe Mu tem afinidade pelo CDNB e pelo hidroperóxido de cumene. Finalmente a classe Pi, a qual apresenta atividade de conjugação com a GSH e atividade de isomerase, além de contribuir na defesa do organismo contra o estresse oxidativo (SHEEHAN et al., 2001). Esta classe apresenta especificidade pelo ácido etacrínico (OAKLEY et al., 1997).

A classe Theta é encontrada em mamíferos e em *Arabidopsis*, apresenta uma serina (Ser12) como resíduo essencial para a catálise. Além disso, mostra uma especificidade única pelo substrato, pois não apresenta atividade para o CDNB, substrato típico das GSTs (SHEEHAN et al., 2001). A classe Delta, encontrada nos insetos está envolvida na desintoxicação de inseticidas organofosforados, além da deidrocloração catalisada pela GST de moléculas organocloradas, como o DDT (SHEEHAN et al., 2001; WONGSANTICHON; HARNNOI; KETTERMAN, 2003).

Em plantas existem cinco classes de GST, a saber, Phi, Tau, Zeta, Theta e Lambda (EDWARDS; DIXON, 2000). A identidade de sequência entre as classes é, via de regra, menor que 30% e dentro das classes, é maior que 30% (DIXON; LAPTHORN; EDWARDS, 2002). As GSTs, de modo geral, são capazes de desintoxicar pesticidas (THOM et al., 2002). A classe Phi é

principalmente responsável pela desintoxicação de herbicidas (DIXON; DAVIS; EDWARDS, 2002), embora também esteja envolvida na deposição de antocianina no vacúolo (MUELLER et al., 2000). As GSTs da classe Tau são importantes devido ao papel que desempenham na seletividade aos herbicidas nas culturas como soja e trigo (EDWARDS; DIXON, 2000; THOM et al., 2002), bem como na sinalização intracelular, nas respostas aos hormônios citocininas e auxinas e na deposição no vacúolo dos pigmentos de antocianina (EDWARDS; DIXON, 2000). A classe Zeta participa do catabolismo da tirosina através de uma atividade semelhante à maleilacetoacetato isomerase (BOARD et al., 1997; DIXON; DAVIS; EDWARDS, 2002), bem como apresenta atividade semelhante à glutational peroxidase, ou seja, catalisam o peróxido em oxigênio e água (BOARD et al., 1997). A Tabela 1 apresenta o número de genes de GST distribuído por classe desta enzima em plantas.

Tabela 1 - Número de genes nas diferentes classes de GST em plantas

Número total de genes de GST por organismo		Número de genes de GST por classe				
		Tau	Phi	Theta	Zeta	Lambda
<i>Arabidopsis thaliana</i>	48	28	13 ^s	3	2 ^s	2
<i>Zea mays</i>	42	28	12 ^s	--	2	--
<i>Glycine max</i>	25	20	4	--	1	--

Fonte: DIXON; LAPTHORN; EDWARDS, 2002.

^s = informação estrutural disponível no “Protein Data Bank” (PDB).

Com relação à disponibilidade e características das estruturas das GSTs, a classe Zeta apresenta duas proteínas depositadas no “Protein Data Bank” (PDB; BERMAN et al., 2000) de humano e de *Arabidopsis*, respectivamente PDB 1fw1 e PDB 1e6b. O alinhamento baseado nas seqüências de GST pertencentes à classe Zeta apresentou uma identidade de seqüência de 45% dentre organismos como *Arabidopsis thaliana*, *Gossypium hirsutum*, *Lycopersicon esculentum*, *Glycine max*, *Zea mays*, *Triticum aestivum*, *Oryza sativa*, *Homo sapiens*, *Drosophila melanogaster* e *Caenorhabditis elegans*. O alinhamento estrutural mostrou que a *HsGSTT2-2* (correspondente à classe Theta de humanos) apresenta um desvio RMS (“root square mean deviation”) de 1,47 Å em relação à *AtGSTZ1* (classe Zeta de *Arabidopsis thaliana*). A superposição das classes de GST mostrou que a *AtGSTZ1* é mais semelhante às classes Theta e

Phi do que em relação às classes Alpha e Pi (THOM et al., 2001). Dentre os resíduos conservados e cataliticamente importantes está a Ser14, a qual é responsável por estabilizar o radical tiol da GSH (BOARD et al., 1997).

A classe Tau é a mais expressa na cultura da soja e do trigo, é responsável pela ligação da GSH aos herbicidas difenil-éters e aril-oxi-fenoxi-propionatos (DIXON et al., 2003; JEPSON et al., 1994; McGONIGLE et al., 2000). Os herbicidas cloroacetamida dimetenamida (RIECHERS et al., 1997), o fenoxaprop-etil (CUMMINS; COLE; EDWARDS, 1997) e o sulfonilurea (KOEPPPE et al., 1997) são rapidamente desintoxicados via conjugação com a GSH. Tais herbicidas são usados no campo co-formulados com um herbicida “safener”, os quais são compostos que aumentam a tolerância ao herbicida nas culturas de cereais ativando rotas de desintoxicação dos xenobióticos. O herbicida “safener” provoca o aumento da expressão da GST sem, no entanto, causar danos à célula (DAVIES; CASELEY, 1999). No milho sua principal função está na desintoxicação dos cloroacetanilídeos (CUMMINS; COLE; EDWARDS, 1997; McGONIGLE et al., 2000; RIECHERS et al., 1997). A estrutura cristalizada da GST de trigo (*TaGSTU4*) permitiu a identificação de resíduos importantes que fazem parte do sítio ativo. A comparação destes resíduos nas GSTs de soja mostraram que estes apresentam mudanças significativas que podem explicar sua grande especificidade pelo substrato (THOM et al., 2002). Esta estrutura está disponibilizada no PDB com o código 1gwc (THOM et al., 2002) e mais recentemente foi obtida a estrutura tridimensional desta classe de GST em arroz, com o PDB 1oyj (DIXON et al., 2003). Esta enzima é a que mais se assemelha à classe Omega, mas sem apresentar a extensão N-terminal, nem a cisteína como resíduo responsável pela catálise, típico da classe Omega (THOM et al., 2002).

A classe Phi de GST é a classe predominante na cultura do milho e apresenta um papel fundamental na desintoxicação do cloroacetoanilídeos, tiocarbamatos e herbicidas correlatos (DIXON et al., 2003; EDWARDS; DIXON, 2000; JEPSON et al., 1994; THOM et al., 2002). Não por acaso, apresenta o maior número de estruturas depositadas no PDB, a saber: PDBs 1bye, 1axd e 1aw9 em milho e PDBs 1gnw e 1bx9 em *A. thaliana*. Existem quatro isoformas conhecidas desta classe, são: GST-I (GSTF1), GST-II (GSTF2), GST-III (GSTF3) e GST-IV (GSTF4). A GSTF1 e a GSTF3 são constitutivas, enquanto a GSTF2 e a GSTF4 são induzidas por herbicidas “safeners” (NEUEFEIND et al., 1997b). As isoenzimas GSTF1, GSTF3 e GSTF4 são homodímeros enquanto a GSTF2 é um heterodímero formado pelos monômeros da GSTF1 e

da GSTF4. GSTF1 é capaz de conjugar a atrazine, enquanto a GSTF3 está envolvida na conjugação do alachlor e metolachlor (DIXON; LAPTHORN; EDWARDS, 2002). A capacidade das GSTs de formar heterodímeros aumenta significativamente a diversidade das GSTs em plantas, segundo dados obtidos através da observação do metabolismo dos herbicidas ativados pelas GSTs em milho e trigo (DIXON; LAPTHORN; EDWARDS, 2002).

A enzima dimérica GSTF4 de *Arabidopsis thaliana* foi cristalizada (PDB 1gnw) junto com quatro moléculas da s-hexilglutathione, um inibidor da GST e um substrato análogo à GSH (REINEMER et al., 1996). Esta enzima tem um sítio H mais amplo e profundo que as GSTs de mamíferos, principalmente devido à terceira α -hélice (N-terminal) e à ausência da extensão C-terminal. Com estas características, esta enzima tem preferência por substratos maiores, por isso apresenta baixa atividade quando na presença dos substratos CDNB e EPNP (NEUEFEIND et al., 1997b; REINEMER et al., 1996). Em *Arabidopsis thaliana*, a Ser11 é responsável pela ponte de hidrogênio com o átomo de enxofre da GSH (REINEMER et al., 1996). Apresenta uma estrutura mais flexível a qual pode permitir uma variedade maior de substratos, esta flexibilidade pode ser causada pela quantidade de resíduos de glicina (Gly) na estrutura (PRADE; HUBER; BIESELER, 1998).

A GSTF3 (PDB 1aw9) apresenta um segmento exposto para o solvente de elevada flexibilidade (Val33 até Asp44). Possivelmente esta região é responsável pela adequação do ligante na enzima (NEUEFEIND et al., 1997a). Ketterman e colaboradores (2001) postularam que a enzima funciona através de um mecanismo de adequação induzido, como o observado na classe Pi, encontrada na placenta de humanos, a qual apresentou mudança conformacional quando ligada a GSH. Na presença desta, a tripsina foi incapaz de atacar a enzima, mostrando que a GSH tem a capacidade de proteger a enzima do solvente. Da mesma forma a GSTF3 apresentou elevada flexibilidade na região C-terminal (BOARD; MANNERNVIK, 1991).

A GSTF1 de milho possui dois domínios bem distintos: o N-terminal composto por 78 resíduos (1 ao 78) e o C-terminal formado pelos resíduos 88 ao 214. Há uma região intermediária formada por nove resíduos (correspondente às posições 79 até 87), a menor encontrada em plantas. A GST de *Arabidopsis* GSTF4 apresenta 15 resíduos entre os dois domínios e a GSTF3 de milho apresenta 14 resíduos (NEUEFEIND et al., 1997a). Um resíduo importante na GSTF1 é o posicionamento da Phe35, numa região central do sítio ativo, este resíduo é substituído pela Leu em outras GSTs de plantas. A Phe35 deve ser a principal responsável pela elevada

especificidade da GSTF1 pelo CDNB, já que esta enzima apresenta uma afinidade maior que a GSTF3 (milho) e GSTF4 (*Arabidopsis*) em relação a este ligante (NEUEFEIND et al., 1997b).

As estruturas das GSTs de plantas encontradas no banco de dados do PDB, contendo as informações de resolução do cristal, a respectiva classe, o ligante e a referência foram resumidas na Tabela 2. A classe Phi apresenta maior número de estruturas, cinco ao todo, sendo que estas compõem três isoformas diferentes, tanto no milho como em *Arabidopsis*. A classe Tau apresenta duas estruturas, uma em trigo e outra no arroz. Por fim, aparece a classe Zeta com uma estrutura em *Arabidopsis*.

Tabela 2 - Descrição das classes de GSTs de plantas, com respectivos ligantes

PDB	Organismo	Resolução(Å)	Classe	Ligante	Referência
1axd	Milho	2,50	GSTF1	d-ácido glutâmico + ácido láctico	NEUEFEIND et al., 1997a
1bye	Milho	2,80	GSTF1	Conjugado de atrazine com glutathione	PRADE; HUBER; BIESELER, 1998
1aw9	Milho	2,20	GSTF3	Cádmio	NEUEFEIND et al., 1997b
1gnw	<i>Arabidopsis</i>	2,20	GSTF4	s-hexilglutathione	REINEMER et al., 1996
1bx9	<i>Arabidopsis</i>	2,60	GSTF4	Flufenacet + gamma-ácido glutâmico	PRADE; HUBER; BIESELER, 1998
1gwc	Trigo	2,25	GSTU	s-hexil-glutathione + íon sulfato	THOM et al., 2002
1oyj	Arroz	1,95	GSTU	Cloro, glicerol e glutathione	DIXON et al., 2003
1e6b	<i>Arabidopsis</i>	1,65	GSTZ	Beta-mercaptoetanol	THOM et al., 2001

Há duas estruturas tri-dimensionais de GST de plantas cristalografadas conjuntamente com um herbicida e referem-se aos PDBs 1bx9 e 1bye, respectivamente, flufenacet e atrazine. O herbicida fluorodifen (herbicida nitrofenil-éter) foi modelado no programa InsightII (Accelrys) e posteriormente foi alocado no PDB 1oyj (DIXON et al., 2003), porém a informação tridimensional deste herbicida não foi disponibilizada.

2.2 Evolução das proteínas

Três são as principais correntes evolutivas, uma aborda a evolução como um processo estocástico (ou aleatório), a outra afirma que a história natural é um processo determinístico. Finalmente a terceira, a corrente mais aceita atualmente, que estabelece que a evolução da vida e da biosfera no planeta Terra pode ser considerada quase-determinística, ou seja, a partir das mesmas condições iniciais existe uma grande probabilidade de que sejam encontrados os mesmos padrões gerais de evolução tanto da biota como do clima, pois as possibilidades são limitadas (SCHWARTZMAN; LINEWEAVER, 2004).

O mecanismo mais importante na formação de novos genes e de novos processos bioquímicos é a duplicação gênica (LI, 1997). Deste modo, novas famílias de proteínas são formadas como resultado da diversificação de regiões codantes (HENIKOFF et al., 1997). Esta duplicação permite a modificação da função da proteína através do aumento do número de sítios ativos ou devido à aquisição de uma função adicional decorrente da inserção ou deleção de fragmentos repetitivos (LI, 1997). Além disso, a duplicação relaxa as restrições impostas pela seleção natural em um dos genes duplicados, e os domínios duplicados ficam livres para divergir de modo a permitir que o gene adquira uma nova especificidade ou uma função mais complexa (LI, 1997).

Após a duplicação gênica, um dos genes que codificam a proteína pode continuar a exercer sua função usual, permitindo a divergência funcional do outro gene, por causa do relaxamento da seleção natural sobre este. O conceito de ortólogo e parálogo foi introduzido originalmente por Fitch (1970). Os genes ortólogos são decorrentes do processo de especiação e os parálogos são genes que sofreram o processo de duplicação, sendo encontrados no mesmo organismo (GERLT; BABBITT, 2000). Após a duplicação, as proteínas parálogas sofrem uma menor pressão evolutiva, o que permite que sua especificidade divirja, promovendo a emergência de novas funções. A especificidade funcional de proteínas é considerada conservada entre os ortólogos e diferente entre os parálogos (MAKAROVA; GRISHIN, 1999; MIRNY; GELFAND, 2002; TATUSOV et al., 2000). Assim, as inferências sobre a função de homólogos são mais acertadas entre os ortólogos do que entre os parálogos (WHISSTOCK; LESK, 2003).

A duplicação interna na evolução de proteínas é provavelmente muito mais importante do que se pensava. Há casos nos quais os eventos de duplicação podem ser facilmente inferidos a

partir de seqüências homólogas (gene do colágeno e os genes da apolipo-proteína) (LI, 1997). Mas, há casos nos quais as regiões duplicadas se tornaram tão divergentes que não é possível identificar as seqüências homólogas. Por exemplo, os domínios variáveis da imunoglobulina não apresentam uma identidade significativa, o que poderia ser um indício de ausência de homologia, mas suas estruturas terciárias sugerem que os dois tipos de domínios foram derivados de um domínio ancestral comum (HOOD; CAMPBELL; ELGIN, 1975).

As GSTs, por sua vez, constituem uma superfamília de proteínas muito antiga, a qual é considerada como um descendente do ancestral das tiorredoxinas, responsável pela resposta ao estresse oxidativo (KOONIN et al., 1994). As classes de GST refletem grupos que divergiram em várias fases do processo evolutivo e hoje são amplamente encontrados nas mais diversas espécies (SHEEHAN et al., 2001). Atualmente são encontrados tipos específicos de GSTs em cada grupo de organismos, o que sugere que o processo de duplicação e evolução acelerou após a radiação dos filos, seguindo um padrão de especificidade relacionada com os táxons (FROVA, 2003). Esta diversificação da superfamília das GSTs citosólicas é considerada um exemplo de divergência evolutiva (ARMSTRONG, 1997).

A evolução tem modelado as proteínas de modo a torná-las específicas para cada ambiente encontrado na natureza, assim cada enzima apresenta especificidades diferentes para os mais diversos substratos encontrados. As GSTs de plantas são um bom exemplo desta especificidade. As duas classes exclusivas das plantas, classe Phi e Tau sofreram uma extensa duplicação seguida de um processo de divergência, durante a evolução (FROVA, 2003). Cada uma apresentando várias isoformas com especificidades diferentes, embora sua estrutura terciária apresente uma alta similaridade.

Seqüências diferentes podem apresentar estruturas semelhantes, isto ocorre em função da estrutura ser mais conservada que a seqüência (PONTING; RUSSELL, 2002). Não é por acaso que a modelagem de proteínas por homologia se baseia nesta premissa para obter a estrutura tridimensional das proteínas. Há vários trabalhos abordando esta questão e dentre as proteínas estudadas podem ser citadas a adenilato-ciclase e a DNA-polimerase, as quais não apresentam uma similaridade evidente, mas mostram que resíduos conservados estão envolvidos numa reação semelhante (ARTYMIUK et al., 1997). A estrutura tridimensional do fragmento N-terminal das proteínas de manutenção dos mini-cromossomos (MDM) encontrados em Arcaea apresentam uma alta conservação, apesar de uma significativa divergência de seqüências (FLETCHER et al.,

2003). Um outro exemplo é a própria superfamília das GSTs, na qual existem tipos diferentes de substratos eletrofilicos que são catalisados por diferentes seqüências desta enzima, mas apesar disso, apresentam uma conformação terciária muito semelhante (ARMSTRONG, 1997; BOARD et al., 1997; SHEEHAN et al., 2001).

Se por um lado é comum encontrar seqüências divergentes com estruturas similares, é possível encontrar também estruturas semelhantes com divergência de função. Existem exemplos que ilustram a convergência estrutural de algumas enzimas embora suas funções não sejam relacionadas (DODSON; WLODAWER, 1998; MAKAROVA; GRISHIN, 1999; RENARD et al., 2005; RUSSELL, 1998; SCHULTZ et al., 2005; UEDA et al., 1999; YOON et al., 2005; ZARRINE-AFSAR; LARSON; DAVIDSON, 2005).

Segundo Ponting e Russell (2002) é possível ainda encontrar casos de convergência da função e dessemelhança na estrutura em famílias de proteínas homólogas. Nestes casos, a natureza inventou diferentes mecanismos para resolver o mesmo problema, ou seja, a especificidade por determinado substrato, bem como a manutenção da atividade da enzima. Como exemplos, podem ser citados a convergência funcional da lactato desidrogenase e da malato desidrogenase em *Trichomonas vaginalis* (WU et al., 1999). Além da convergência funcional da hexoquinase, riboquinase e galactoquinase a partir de diferentes conformações tridimensionais (BORK; SANDER; VALENCIA, 1993) e a evolução convergente de defensinas de invertebrados e os fatores antibacterianos de nematóides (FROY, 2005).

As enzimas podem apresentar também uma promiscuidade catalítica, a qual se refere à capacidade que uma mesma enzima apresenta de catalisar um substrato alternativo, normalmente similar ao substrato usual. O exemplo típico é a quimotripsina, a qual catalisa as reações de amidase e de fosfotriesterase (O'BRIEN; HERSCHLAG, 1999). Além deste, há outros casos, como a enzima metano-oxigenase, a qual hidrolisa 150 substratos diferentes além do metano, ou a serotonina N-acetiltransferase, a qual atua também como uma alquiltransferase (COPLEY, 2003). Entretanto, em alguns casos a habilidade fortuita que uma enzima venha a apresentar pode ser produto de condições de laboratório e não serem efetivamente encontradas na natureza; como exemplo, a adenosina deaminase, a qual catalisa a deaminação de aminopurinas, mas pode também catalisar a substituição nucleofílica aromática de anéis de purina (WACKETT, 2004). É importante salientar que a promiscuidade catalítica é diferente da capacidade que algumas enzimas apresentam, como o citocromo P450 e as GSTs, de catalisarem uma ampla gama de

substratos (também denominado de baixa especificidade) dentro de cada classe destas enzimas (COPLEY, 2003). Por outro lado, a classe Zeta das GSTs é caracterizada por apresentar uma notória promiscuidade catalítica, devido à sua capacidade de oxidação do ácido dicloro-acético e da isomerização do maleilacetoacetato em fumarilacetoacetato (POLEKHINA et al., 2001).

2.3 Filogenia e a estrutura tridimensional das proteínas

A relação entre os processos evolutivos moleculares e as estruturas e funções das macromoléculas foram reconhecidos há um longo tempo atrás. A contribuição da informação filogenética para a biologia estrutural vai além da identificação de resíduos conservados e variáveis no alinhamento (BLOUIN; BOUCHER; ROGER, 2003). A compreensão dos fundamentos físicos dos processos que ocorrem com as proteínas não apenas identificam fatores importantes na emergência de novas funções, mas também ajudam a construir os modelos filogenéticos baseados na estrutura das proteínas (LIO et al., 1998).

A análise de seqüências de RNA apresentou um sinergismo entre as inferências filogenéticas, o alinhamento e a estrutura da molécula. Este é o começo para que o mesmo seja considerado verdade para as proteínas (FELSENSTEIN, 1996). Johnson; Sali e Blundell (1990) afirmaram que a estrutura tridimensional da proteína pode ser usada para inferir relações filogenéticas. Os alinhamentos estruturais foram claramente mais exatos que a comparação de seqüências, mesmo quando a identidade de seqüência estava por volta dos 10% (JOHNSON et al., 1996). Os resíduos variáveis correlacionados no alinhamento estrutural puderam ser identificados de modo a aumentar a exatidão dos modelos usados para inferir a filogenia. Além disso, a variação de um aminoácido em um alinhamento fornece dados a respeito dos resíduos importantes para a formação das estruturas secundárias e terciárias da molécula (GRIFFITHS, 1998).

Devido ao fato da divergência evolutiva estrutural ocorrer mais lentamente que a divergência de seqüências das proteínas, as limitações seletivas podem preservar a estrutura da mesma. Conseqüentemente, a informação estrutural apresenta maior confiabilidade que a informação da seqüência (THORNE; GOLDMAN; JONES, 1996). Wood e Pearson (1999), estudaram 36 famílias de proteínas, avaliaram que a correlação entre a similaridade de seqüência e da estrutura foram consideravelmente elevadas, de forma que 18 das 36 famílias analisadas apresentaram um

coeficiente de correlação maior que 87,8 % (dentre elas estavam as GSTs). Apenas nove tiveram um coeficiente de correlação menor que 81,5 %.

3 MATERIAL E MÉTODOS

3.1 Bancos de dados consultados

Foi realizada uma busca nos três principais bancos de dados de seqüências de proteínas disponíveis, a saber: “European Bioinformatics Institute” (EBI; responsável pelo banco de dados da Europa, “European Molecular Biology Laboratory” - EMBL; KANZ et al., 2005), “Protein Identification Resource” (PIR; o banco de dados dos EEUU; SIDMAN et al., 1988) e “DNA Data Bank of Japan” (DDBJ; o banco de dados do Japão; TATENO et al., 2002). A busca foi feita utilizando as palavras-chave “glutathione transferase” e “glutathione s-transferase” ou o código da enzima 2.5.1.18 (“Enzyme Classification”, EC - “Nomenclature Committee of the International Union of Biochemistry and Molecular Biology”). Primeiramente fez-se uma busca por taxons nos cinco reinos descritos por Margulis e Schwartz (2001), cada um dos 98 filos foram pesquisados de acordo com o critério da palavra-chave explicitada acima. Para os filos Craniata e Magnoliophyta foram utilizados o banco de dados do SwissProt (BOECKMANN et al., 2003). No segundo passo, foram retiradas as seqüências repetidas e as incompletas. Num terceiro momento as seqüências muito semelhantes (mais que 90% de identidade) foram também eliminadas.

A principal fonte de dados para a bioinformática estrutural é obtida no “Protein Data Bank” (PDB; BERMAN et al., 2000; <http://www.rcsb.org/pdb/>). Este banco de dados começou em 1971 com sete arquivos de estruturas tridimensionais de proteínas e até dezembro de 2005 conta com mais de 33.930 estruturas descritas. Dentre os bancos de dados que classificam as proteínas de acordo com a conformação tridimensional, destaca-se o “Structural Classification of Proteins” (SCOP, <http://scop.mrc-lmb.cam.ac.uk/scop/>), o qual foi criado em 1995 e classifica as enzimas em domínios, famílias, superfamílias e folds (MURZIN et al., 1995). A hierarquia mais baixa é formada pelos domínios individuais das proteínas, extraídos do banco de dados do PDB. Um conjunto de domínios compõe as famílias, para os quais as semelhanças da seqüência, estruturais e algumas vezes a função, implicam em uma origem evolutiva comum. As famílias que contém proteínas com estruturas similares compõem as superfamílias. Estas podem ainda ser agrupadas em "folds", as quais compartilham uma conformação geral.

O Swiss-Prot (BOECKMANN et al., 2003; <http://us.expasy.org/sprot/>) foi estabelecido em 1986 com o objetivo de organizar as informações de sequência e estrutura das proteínas. São quatro os critérios que o distinguem de um banco de dados convencional, a seguir: a) anotação da data de depósito, informação a respeito da citação e descrição biológica da origem da proteína; b) mínima redundância dos dados; c) integração com outros bancos de dados; d) documentação a respeito da organização e informações pertinentes a este banco de dados.

3.2 Alinhamentos de sequência e de estrutura tridimensional

O alinhamento de seqüências foi feito utilizando-se o ClustalX (THOMPSON et al., 1997). Os alinhamentos estruturais das GSTs foram realizados nos programas PrISM (YANG; HONIG, 1999) e no programa Combinatorial Extension (CE; SHINDYALOV; BOURNE, 1998). O PrISM foi desenvolvido para analisar seqüências e estruturas de proteínas, embora seu papel principal seja de prever estruturas tridimensionais de proteínas a partir de suas seqüências. Neste trabalho o PrISM foi usado apenas para o alinhamento estrutural múltiplo. Este programa gera um arquivo texto contendo as seqüências alinhadas das proteínas, as quais podem posteriormente ser submetidas à análise filogenética.

O programa CE realiza o alinhamento de duas estruturas de cada vez, utilizando um algoritmo baseado no alinhamento de pares de fragmentos das estruturas avaliadas. Os alinhamentos foram realizados no endereço <http://cl.sdsc.edu/ce.html>.

As GSTs classe Kappa e microssomais foram desconsideradas das análises devido à grande divergência encontrada com as GSTs analisadas nesta tese, especialmente com relação à estrutura tridimensional da enzima. Além disso, recentemente propôs-se que a classe Kappa pertence à família da hidroxicromeno-carboxilato-isomerase e não mais às GSTs (MOREL et al., 2004; ROBINSON et al., 2004). As GST microssomais (MGST1), por sua vez, foram agrupadas numa nova superfamília denominada de proteínas associadas a membranas, as quais participam do metabolismo da glutatona (MAPEG) (JAKOBSSON et al., 1999).

3.3 Reconstrução filogenética

A reconstrução filogenética foi feita no programa Phylip3.65 (FELSENSTEIN, 1993) utilizando-se o método Neighbor-Joining (SAITOU; NEI, 1987). O alinhamento proveniente das estruturas de GST foi submetido ao mesmo método que as seqüências de proteínas, exceto pelo alinhamento propriamente dito, o qual foi realizado no programa PrISM (YANG; HONIG, 1999).

3.4 Modelagem de proteínas por homologia

A modelagem de proteínas a partir de estruturas cristalografadas depositadas no PDB baseia-se nos seguintes pontos: a) crescente importância das novas técnicas de bioinformática disponíveis para a ciência contemporânea (BAKER; SALI, 2001; KOONIN; WOLF; KAREV, 2002; SAMUDRALA; LEVITT, 2002); b) contínuo aumento de seqüências de DNA (provenientes dos projetos de seqüenciamento); c) elevado custo e tempo para determinação experimental da estrutura (1 a 2,5 anos); d) impossibilidade das estruturas acompanharem o número de seqüências obtidas. Por isso, o New York Structural Genomics Research Consortium (2001) determinou que para cada nova estrutura depositada no PDB, em média, 100 seqüências de proteínas deverão ser modeladas.

A modelagem por homologia é um método que baseia-se na predição tridimensional da proteína a partir de uma estrutura conhecida de uma ou mais proteínas correlatas. O resultado é um conjunto de coordenadas, tanto da cadeia principal como das cadeias laterais, dos aminoácidos que compõem a proteína. A modelagem pode ser sub-dividida em cinco etapas, a saber: a) alinhamento das seqüências de aminoácidos da proteína alvo com a proteína molde, cuja estrutura é conhecida; b) determinar os segmentos que representam as regiões contendo as inserções e deleções. Alinhar estruturalmente estes segmentos na cadeia principal da proteína molde, criando assim um modelo completo para a cadeia principal da proteína alvo; c) substituir os resíduos da cadeia lateral que foram mutados em relação à proteína molde. Para os resíduos que não foram mutados, perfaz-se a manutenção conformacional da cadeia lateral da proteína molde. Alguns programas buscam conformações possíveis para a cadeia lateral, de modo a evitar as colisões entre os átomos; d) avaliar se foram encontradas colisões no modelo gerado. Se encontradas devem ser corrigidas; e) refinar o modelo através de minimização de energia. Para

avaliar as diferenças estruturais entre duas proteínas avalia-se a distância média entre os átomos correspondentes de cada estrutura. Convencionalmente esta distância é reportada como desvio do quadrado médio da raiz (desvio RMS; “root-mean-square”) e apresenta a seguinte fórmula: eq. (1),

$$\text{desvio RMS} = \sqrt{\sum d_i^2 / n} \quad (1)$$

Onde: d_i é a distância entre os pares de pontos (i) após o alinhamento ótimo e n corresponde ao número total de pontos (LESK, 2002).

Para a análise da proteína de cana-de-açúcar e dos mutantes, as estruturas tridimensionais foram obtidas através de modelagem por homologia utilizando o programa Modeller7v6 (SALI; BLUNDELL, 1993). A estrutura molde foi a GST de milho da classe Phi (ZmGSTF1) com os PDBs 1axd e 1bye. A alocação do conjugado ATA-GSH foi realizada a partir do alinhamento estrutural da ZmGSTF1 com o modelo recém obtido da SoGSTF1. Desta forma, o ligante que foi obtido através da cristalização da GST de milho (PDB 1bye) passou a fazer parte da GST de cana-de-açúcar.

3.5 Modelagem molecular

A modelagem molecular trata dos processos que descrevem os complexos sistemas químicos de modo a definirem um modelo atômico o mais exato possível. Tem o objetivo de compreender e prever as propriedades macroscópicas baseadas no conhecimento detalhado da escala atômica. Para tanto, é importante compreender os tipos de interações entre os átomos, particularmente as não-covalentes. Embora as ligações covalentes sejam consideradas ligações fortes, são as interações não-covalentes as responsáveis por estabilizar tanto a estrutura tridimensional como a ligação da enzima com seu ligante. Estas interações são individualmente fracas, mas como estão presentes em grande número, sua importância faz-se presente. São as ligações de hidrogênio e as interações de Van der Waals. Estas interações são fracas e transitórias, individualmente são formadas e quebradas a cada fração de segundo. A natureza transitória destas ligações é que proporciona a flexibilidade necessária para que a enzima exerça

sua função (NELSON; LEHNINGER; COX, 2000). Como a estrutura da proteína e sua função são governadas pelas interações inter-atômicas sua avaliação e análise devem contribuir para que asserções possam ser feitas com relação à função das mesmas. Resumidamente pode-se afirmar que as forças de Van der Waals são interações intermoleculares que ocorrem em moléculas eletricamente neutras (ou apolares). Há oito aminoácidos capazes de interagir entre si através desta forças, são: Gly, Ala, Val, Leu, Ile, Phe, Pro e Met. Com relação às ligações de hidrogênio, são interações fracas que ocorrem entre o hidrogênio de um grupo -OH ou -NH e um átomo de oxigênio ou nitrogênio de outra molécula. Os aminoácidos capazes de apresentarem ligações de hidrogênio são os hidrofílicos: Ser, Thr, Cys, Asn, Gln, His, Tyr, Trp, Asp, Glu, Lys e Arg. Um terceiro tipo de interação são as eletrostáticas, as quais são fortes ligações que ocorrem entre íons e podem ser atrativas (pólos opostos) ou repulsivas (pólos iguais). Esta força é inversamente proporcional à distância. Finalmente o quarto tipo de interação se refere às ligações dissulfeto promovidas pelo radical tiol (-SH) de dois aminoácidos cisteína. Esta ligação tem maior importância em moléculas pequenas (PATRICK, 2001). Ainda, é importante lembrar que o anel aromático apresenta uma interação de Van der Waals do tipo dipolo induzido, a qual resulta numa nuvem de elétrons com as faces do anel ricas em elétrons e as bordas deficientes em elétrons (PATRICK, 2001).

Com relação às propriedades físicas macroscópicas encontradas entre os átomos podem ser de dois tipos, a primeira é o equilíbrio estático, tal como a ligação de um inibidor em uma enzima. O segundo tipo é o equilíbrio dinâmico, por exemplo, as reações cinéticas ou os processos de difusão em membranas. Basicamente há dois métodos que avaliam estes dois estados de equilíbrio, respectivamente, o docking e a dinâmica molecular.

3.5.1 Docking molecular

O docking molecular é mais simples e conseqüentemente mais rápido, pois não envolve a introdução do campo de forças, nem a solvatação do sistema, além de considerar as moléculas como estruturas rígidas. Estas características possibilitam a redução significativa dos graus de liberdade do sistema. Embora, na natureza o princípio da chave-fechadura seja válido para chaves e fechaduras flexíveis, os algoritmos de docking rígidos apresentam um êxito considerável como ferramenta de alocação do ligante no sítio ativo da enzima.

O método de docking utilizado nesta tese foi o LIGIN (SOBOLEV et al., 1996). Este método baseia-se na maximização da função de complementariedade a qual é dependente da superfície de contato e das propriedades químicas entre os átomos envolvidos. A função de complementariedade (CF) é usada para caracterizar a "alocação" do substrato no sítio ativo da enzima, é definido pela seguinte fórmula: eq.(2),

$$CF = S_l - S_i - E, \text{ onde} \quad (2)$$

S_l e S_i são respectivamente os contactos legítimos e ilegítimos entre as superfícies dos átomos do ligante e da enzima, a letra E se refere à força de repulsão de átomos muito próximos, ou seja, cuja distância seja menor que o raio de Van der Waals.

As estruturas envolvidas na análise de modelagem molecular foram: a *ZmGSTF1* de milho (PDB 1axd), correspondente à classe Phi, e a *HsGSTP1* de humanos (PDB 3gss), correspondendo à classe Pi das GSTs. Foi utilizado um volume de $5 \times 5 \times 5 \text{ \AA}^3$, no qual a molécula foi centrada na posição inicial fornecida. O programa gerou 1000 posições aleatórias para o ligante e posteriormente maximizou a função de complementariedade.

3.5.2 Dinâmica molecular

Para os casos de equilíbrio dinâmico o método mais apropriado é a dinâmica molecular, a qual é baseada na segunda lei de Newton ($F = m.a$), onde as forças são obtidas como gradientes da energia potencial de um sistema. Este método possibilita a avaliação de uma molécula de proteína complexada com seu ligante, solvatada numa caixa com água a determinada pressão e temperatura. Um dos pontos principais consiste na escolha dos parâmetros que irão definir o tipo de interações entre os átomos encontrados no sistema. Sejam estes átomos da própria proteína ou entre o ligante e a proteína.

Uma fórmula simples que define o campo de forças pode ser representada a seguir: eq. (3),

$$V(r) = V_l + V_o + V_{\text{tor}} + V_{\text{vdw}} + V_{\text{hb}} + V_{\text{elet}}, \quad (3)$$

$V(r)$ é a energia potencial do sistema na posição r , V_l é a energia resultante da extensão ou encurtamento da ligação covalente, V_o é a energia de deformação dos ângulos das ligações, V_{tor} é a energia associada aos ângulos de torção; V_{vdw} é a energia não covalente resultante das forças de Van der Waals, V_{hb} é a energia das ligações de hidrogênio e finalmente V_{elet} , a qual é resultante das forças eletrostáticas entre os átomos (MORAITAKIS; PURKISS; GOODFELLOW, 2003).

O complexo atrazina-glutationa (ATA-GSH) foi alocado no sítio catalítico da classe Phi, *ZmGSTF1* e os resultados foram analisados após a minimização de energia e dinâmica molecular do complexo GST + ligantes (ATA + GSH). Estas análises foram realizadas no programa Gromacs (Groningen Machine for Chemical Simulations; SPOEL et al., 2004). A dinâmica molecular da GST complexada com o ligante foi realizada na forma dimérica da enzima, pois apresenta um complexo mais estável ao sistema (ARMSTRONG, 1997).

Para obter o campo de forças do ligante foi utilizado o programa PRODRG (SCHUETTEKOPF; VAN AALTEN, 2004). A dinâmica molecular foi realizada no programa Gromacs 3.2, utilizando o campo de forças do gromos96. Com as seguintes condições: O sistema foi solvatado em água numa caixa cúbica, com 1,0 nm de solvente em todos os lados. Para as moléculas de água foi utilizado o modelo SPC/E. Íons de sódio foram adicionados para neutralizar a simulação. A temperatura foi mantida constante a 300K e a pressão foi mantida constante, 1bar. O algoritmo LINCS foi usado para limitar ("constrain") o comprimento das ligações covalentes ("all-bond"). Para as moléculas de água foi usado o algoritmo SETTLE para limitar o comprimento das ligações e dos ângulos. Para as interações não covalentes foi utilizado o método PME com uma distância de 0,9 nm, calculado a cada passo ("step") e atualizado a cada 10 passos. Foi utilizado um intervalo de tempo de 0,002 ps para cada passo. A energia potencial do sistema foi minimizada utilizando o algoritmo "steepest descent". As moléculas de água foram ajustadas ao sistema utilizando uma dinâmica molecular com restrição nas posições "position restraints" por 20ps. O tempo total da dinâmica foi de um nanosegundo.

3.6 Contatos da interface enzima - substratos

Para a análise dos contatos dos ligantes com a enzima foi utilizado o programa Ligand Pocket Contacts (LPC; SOBOLEV, et al., 1999). Este programa especifica as distâncias entre os átomos, bem como especifica a superfície de contato entre os mesmos. Além disso, discrimina o tipo de

interação entre as moléculas participantes da interação, a saber: interação de hidrogênio (HB), interação aromática (Arom.), interação hidrofóbica e se apresenta contatos desestabilizadores (DC).

3.7 Visualização das moléculas

Para a visualização das moléculas foi utilizado o VMD (HUMPHREY; DALKE; SCHULTEN, 1996) e para a análise das ligações de hidrogênio foi realizado o programa DeepView 3.7 (Swiss-PdbViewer; GUEx; PEITSCH, 1997).

3.8 Construções gênicas

As construções gênicas foram realizadas no laboratório de Biologia Molecular do Centro de Tecnologia Canavieira, sob a orientação da Dra. Sabrina Moutinho Chabregas. Procedimentos padrões foram utilizados na realização das construções gênicas (SAMBROOK; FRITSCH; MANIATIS, 1989). Foram realizadas mutações sítio-dirigidas no gene que codifica para a GST em cana-de-açúcar (clone SCJLRT1020C09 do SUCEST) para obtenção dos mutantes W12L e I118F. A necessidade de se adequar os fragmentos de DNA amplificados a fim de facilitar a clonagem e garantir a fusão traducional das construções, exigiu modificações nas extremidades (via PCR). Foram sintetizados os oligonucleotídeos providos com sítios de restrição (*Nco*I na extremidade 5' e *Bam*HI na extremidade 3' do fragmento amplificado) (Figura 1) e a clonagem foi feita nos respectivos sítios do vetor de expressão em bactéria pET28b (Novagen) (Figura2). O gene que codifica para a proteína GST (mutada e controle) foi fusionado a uma cauda de histidina, disponível no vetor de clonagem, para possibilitar a purificação das proteínas heterólogas expressas.

Construção pET28b-GST:**GST:**

NcoI *BamHI*

5' CCC CC ATG GAA ATG GCT CCG ATG AAG CTG TAC GGG GCG GTG ... AAG CCA TCT GCT TGG GAT CCG GG 3'

3' GGG GG TAC CTT TAC CGA GGC TAC TTC GAC ATG CCC CGC CAC ... TTC GGT AGA CGA ACC CTA GGC CC 5'

↓
NcoI e BamHI

5' C ATG GAA ATG GCT CCG ATG AAG CTG TAC GGG GCG GTG CTG ... AAG CCA TCT GCT TTG 3'

3' CTT TAC CGA GGC TAC TTC GAC ATG CCC CGC CAC GAC ... TTC GGT AGA CGA AAC CTA G 5'

pET28b:

NcoI *BamHI*

5' TAT ACC ATG GGC AGC ... GGT CGG GAT CCG AAT TCG AGC TCC GTC GAC AAG ...3'

3' ATA TGG TAC CCG TCG ... CCA GCC CTA GGC TTA AGC TCG AGG CAG CTG TTC ...5'

↓
NcoI e BamHI

5' TAT AC GAT CCG AAT TCG AGC TCC GTC GAC AAG ...3'

3' ATA TGG TAC GC TTA AGC TCG AGG CAG CTG TTC ...5'

Ligação:

5' TAT ACC ATG GAA ATG GCT CCG ... TGG GAT CCG AAT TCG AGC TCC GTC GAC AAG ...3'

3' ATA TGG TAC CTT TAC CGA GGC ... ACC CTA GGC TTA AGC TCG AGG CAG CTG TTC ...5'

Figura 1 - Esquema da construção pET28b-GST

Figura 2 - Esquema do plasmídeo pET28b

Para a mutação W12L, o gene que codifica para GST foi amplificado com os oligonucleotídeos:

W12L_{for}:

5' GGG GCC ATG GAA ATG GCT CCG ATG AAG CTG TAC GGG GCG GTG CTG TCG **TTG** AAC GTG 3'
 NcoI Met Ala Pro Met Lys Leu Tyr Gly Ala Val Leu Ser **Leu** Asn Val

e

GST_{rev}:

5' CCC GGA TCC TCA AGC AGA TGG CTT CAT CAT GGC 3'
 BamHI

Em **negrito** está o códon mutado, os sítios de restrições inseridos para clonagem foram sublinhados (*Nco*I e *Bam*HI). A reação padrão de amplificação utilizada foi a seguinte:

DNA (clone SUCEST SCJLRT1020C09)	0,1 mg
Tampão com MgSO ₄ (10X concentrado)	5,0 ml
Oligo 1	30,0 pmoles
Oligo 2	30,0 pmoles
dNTP	0,2 mM
Pfu DNApolimerase (marca Fermentas)	1,5 unidades

O volume da reação foi completado a 50 ml com água.

O ciclo utilizado consiste de uma pré-desnaturação de 6 minutos a 95°C, a seguir:

- desnaturação: 95°C por 1 minuto,
 - anelamento: 60°C por 1 minuto,
 - extensão: 72°C por 2 minutos,
- por 30 ciclos. Posteriormente, foi realizada uma extensão final de 6 minutos a 72°C.

Para mutação I118F, a amplificação foi realizada em três etapas, segundo o protocolo de mutação sítio-dirigida adaptado a partir de Strachan e Read, 1999.

ETAPA 1 - O primeiro fragmento amplificado, correspondente à porção 5' do gene, utilizou os seguintes oligonucleotídeos:

PETforw:

5' GGG GCC ATG GAA ATG GCT CCG ATG AAG CTG TA 3'

e

I118Fmut1rev:

5' TCC CCC AAG CAT CGG ACT **GAA** GAG GAC CTG GAA GAG GAT 3'

Em **negrito** está o códon mutado e sublinhado encontra-se o sítio de restrição *Nco*I inserido na extremidade 5' do fragmento amplificado.

O segundo fragmento corresponde à porção 3' do gene e foi amplificado com os seguintes oligonucleotídeos:

I118Fmut2forw:

5' ATC CTC TTC CAG GTC CTC **TTC** AGT CCG ATG CTT GGG GGA 3'

e

PETrev:

5' CCC GGA TCC TCA AGC AGA TGG CTT CAT CAT GGC 3'

Em negrito está o códon mutado. O oligonucleotídeo PETrev, descrito anteriormente, contém o sítio *Bam*HI na extremidade. Pode-se observar que os oligonucleotídeos I118Fmut1rev e I118Fmut2forw são complementares e ambos encarregam-se de inserir o códon modificado nas duas porções amplificadas do gene. A reação de amplificação utilizada é a mesma descrita anteriormente, com o seguinte ciclo:

Pré-desnaturação de 6 minutos a 95°C, a seguir:

- desnaturação: 95°C por 1 minuto
 - anelamento: 60°C por 1 minuto
 - extensão: 72°C por 1 minuto,
- por 30 ciclos. Posteriormente, foi realizada uma extensão final de 6 minutos a 72°C.

Os fragmentos foram extraídos do gel de agarose utilizando-se o kit Wizard Gel and PCR Clean-up System da Promega, segundo protocolo do fabricante.

ETAPA 2 – Os fragmentos purificados foram utilizados como molde para a próxima reação:

Fragmento 1 (correspondente à porção 5' do gene)	3,0 ml
Fragmento 2 (correspondente à porção 3' do gene)	3,0 ml
Tampão com MgSO ₄ (10X concentrado)	5,0 ml
dNTP	0,2 mM
Pfu DNAPolimerase (marca Fermentas)	1,5 unidades

O volume da reação foi completado a 50 ml com água.

O ciclo utilizado consiste de uma pré-desnaturação de 2 minutos a 95°C, a seguir:

- desnaturação: 95°C por 1 minuto
 - anelamento: 60°C por 1 minuto
 - extensão: 72°C por 2 minutos,
- por 9 ciclos. Posteriormente, foi realizada uma extensão final de 2 minutos a 72°C.

ETAPA 3 – Para amplificação do fragmento completo, contendo a mutação I118F, foram adicionados 30 pmoles de cada um dos oligonucleotídeos, Petforw e PETrev, descritos anteriormente e uma reação de amplificação semelhante à anterior foi realizada, desta vez com 25 ciclos.

Para obtenção da construção contendo a GST sem mutação, que será utilizada como controle, uma reação de amplificação semelhante à descrita para a mutação W12L foi realizada com os oligonucleotídeos Petforw e PETrev, já descritos. Todos os fragmentos amplificados foram extraídos do gel de agarose e digeridos com as enzimas *Nco*I e *Bam*HI e inseridos dentro dos sítios correspondentes no vetor PET28b, em fase de leitura com uma cauda de histidina. Os vetores produzidos foram: PET28b-W12L, PET28b-I118F e PET28b-GST e serão posteriormente utilizados para expressão e purificação das formas mutantes e não mutante da proteína GST de cana-de-açúcar em bactéria, para teste de afinidade com o herbicida atrazine.

3.9 Cepas bacterianas

Durante as etapas de clonagem, foram utilizadas as cepas bacterianas *Escherichia coli* JM109 (YANISCH-PERRON; VIEIRA; MESSING, 1985).

3.10 Análise dos transformantes (“Screening” por PCR e restrição enzimática)

A análise das colônias transformantes foi realizada via PCR, as colônias foram repicadas em uma placa “Master”, antes do início da amplificação. Os clones referentes à transformação com as construções foram analisados diretamente, utilizando numa primeira etapa, os oligonucleotídeos: PETforw e I118Fmut1forw, descritos anteriormente. Numa segunda etapa, os

clones PCR positivos tiveram o DNA extraído e foram digeridos com as enzimas *Nco*I e *Eco*RI, a fim de verificar o número de fragmentos ligados ao vetor.

3.11 Seqüenciamento das construções gênicas

Alguns clones positivos foram enviados para seqüenciamento automático (seqüenciador de DNA ABI PRISM 377) no Laboratório de Genética Molecular do Departamento de Entomologia, Fitopatologia e Zoologia Agrícola da ESALQ, sob a coordenação do Prof. Dr. Luiz Eduardo Aranha Camargo. A reação de seqüenciamento, foi realizada usando-se os procedimentos descritos no manual do kit “ABI Prisma BigDye Terminator Cycle Sequencing Ready Reaction Kit” (Perkin Elmer, Applied Biosystems). Seguindo este protocolo, em tubos de 0,2 mL foram adicionados 2 mL de “mix Big Dye” (que contém dNTP, terminadores ddNTPs associados à substâncias fluorescentes e AmpliTaq DNA polimerase, FS), 3 picomoles de primer (no caso, UP descrito anteriormente), 10,5 mL de água estéril, 6 mL de tampão Tris-HCl pH 9,0 contendo cloreto de magnésio e 0,5 mL de DNA plasmidial (200-500 ng). A seguinte reação de PCR foi utilizada:

96°C por 2 minutos	
96°C por 20 segundos	25 ciclos
50°C por 10 segundos	
60°C por 4 minutos	
72°C por 7 minutos	

Após a reação de amplificação, as amostras foram transferidas para microtubos de 0,65 mL e adicionou-se uma solução contendo 20 mL de água estéril + 60 mL de isopropanol. Essa mistura foi deixada à temperatura ambiente por 15 minutos e centrifugada em uma microcentrífuga de bancada (velocidade máxima) por 20 minutos. O sobrenadante foi cuidadosamente descartado e o precipitado lavado com 250 mL de etanol 70% (preparado na hora). Centrifugou-se por mais 15 minutos sob as mesmas condições anteriores e o sobrenadante foi novamente descartado. Os tubos, com as tampas abertas, foram colocados por 1 minuto em termociclador, previamente aquecido à 90 °C para a completa evaporação da água das amostras.

No laboratório onde foi realizado o seqüenciamento, as amostras foram ressuspendidas em 4 mL de tampão contendo 5 partes de formamida e 1 parte de “loading buffer” (0,25% azul de bromofenol, 0,25% xileno cianol e 40% sacarose), o DNA foi desnaturado a 98°C por 2 minutos e 1 mL de cada amostra foi aplicado no gel por Júlia Ronzella Ottoni e Osmar Vaz de Carvalho Netto.

As colônias com a seqüência desejada foram repicadas e estocadas em glicerol a uma temperatura de -80°C.

3.12 Meios de cultura

Para crescimento de *E. coli* foi utilizado o meio de cultura Luria-Bertani (LB), composto de 1% (p/v) de bacto-triptona, 0,5% (p/v) de extrato de levedo, 1% (p/v) de NaCl, pH 7,5, acrescido de 0,01% (p/v) de canamicina (para *E. coli*). Quando necessário, 0,8% de bácto-agar foi adicionado ao meio.

4 RESULTADOS

4.1 Amostragem das GSTs nos diversos filos taxonômicos

A grande maioria dos filos analisados não apresentou seqüências de GSTs. O reino Eubacteria teve uma amostragem relativamente bem representativa de GSTs nos 10 filos, com Proteobacteria (1476 seqüências de GSTs), Spirochaetae (16 seqüências), Chlorobia (2), Actinobacteria (40) e Cyanobacteria (174). O reino Protocista apresentou somente dois filos com seqüências de GSTs, dos 30 filos analisados, Ciliophora (5) e Apicomplexa (46). O reino Fungi apresentou GSTs nos três filos, Zygomycota (2 seqüências de GST), Ascomycota (476 seqüências) e Basidiomycota (72 seqüências). O reino Plantae foi pouco representativo, dos 12 filos estudados, somente os filos Coniferophyta (7 seqüências) e Magnoliophyta (1085 seqüências) apresentaram GSTs. O reino Animalia apresentou GSTs no filo dos Platyhelminthes (83 seqüências), Nematoda (265), Mandibulata (67), Chelicerata (11), Crustacea (7), Mollusca (39), Urochordata (6), Cephalochordata (2) e Craniata (2958).

4.2 Reconstrução filogenética baseada no alinhamento de seqüências de GSTs

A análise da reconstrução filogenética baseada no método Neighbor-Joining (Figura 3), mostrou o agrupamento das onze classes de GST além do fator de elongação e de cinco seqüências encontradas em bactérias, agrupadas em três grupos. Um dos grupos contendo duas seqüências de Eubacteria foi agrupado com uma seqüência do filo Ascomycota (GST2_YEAST, *Saccharomyces cerevisiae*). O grupo que apresentou o Ascomycota foi definido por McGoldrick et al. (2005) como o grupo externo às demais classes de GST. As outras três seqüências de Eubacteria ficaram isoladas no alinhamento, o que sugere uma grande diversidade de GSTs encontradas neste organismo, afora a classe Beta. Além dos grupos citados, três grandes clusters foram identificados, a saber: o cluster formado pelas classes Alpha, Sigma, Pi e Mu, apresentaram valores significativos de bootstrap (acima de 700). O cluster formado pelas classes Tau, Omega e Zeta, com valores de bootstrap acima de 700. Finalmente o cluster formado pelas classes Theta e Delta (bootstrap 885) e a classe Phi que ficou num grupo à parte com maior semelhança para estas duas últimas classes citadas, porém, com um valor de bootstrap de 496. O

Fator de Elongação e a classe Beta formaram um outro cluster bem definido, com um valor de bootstrap de 997. Este último resultado foi também encontrado por McGoldrick e colaboradores (McGOLDRICK; O'SULLIVAN; SHEEHAN, 2005).

Figura 3 - Reconstrução filogenética baseada na seqüência de aminoácidos. Foram identificados os gêneros e filos dos organismos e as classes da enzima GST

4.3 Reconstrução filogenética baseada no alinhamento tridimensional

A análise filogenética das estruturas de GST está apresentada na Figura 4. A região N-terminal do alinhamento foi apresentada na Figura 5. Foi encontrado o mesmo padrão filogenético que o encontrado no alinhamento de seqüências, embora o número de estruturas corresponda a 21% do valor total das seqüências analisadas (32 estruturas e 153 seqüências). O cluster formado pelas classes Alpha, Sigma, Pi e Mu, apresentaram valores de bootstrap acima de 700. O cluster formado pelas classes Tau e Omega, com um valor de bootstrap de 908 e a classe Zeta, como um grupo externo a estas duas últimas classes citadas, porém com um valor de bootstrap de apenas 418. As classes Theta e Delta apresentaram um valor de bootstrap de 753 e a classe Phi foi agrupada no mesmo cluster, porém com um valor de bootstrap de 595. A classe Beta não foi relacionada estruturalmente com outras classes de GST. Em geral, os valores de bootstrap foram ligeiramente maiores do que os obtidos no alinhamento de seqüências. As exceções ficaram por conta das classes Zeta, em relação às classes Tau e Omega, cujo valor de bootstrap foi de somente 418 na reconstrução filogenética baseada no alinhamento estrutural e de 915 na reconstrução baseada no alinhamento de seqüências. Além das classes Pi e Mu cujos valores de bootstrap foram de 722 na reconstrução baseada no alinhamento estrutural e de 941 na reconstrução baseada no alinhamento de seqüências.

4.4 Avaliação geral do alinhamento baseado na estrutura tridimensional

O alinhamento estrutural das GSTs das classes Alpha, Mu, Pi, Sigma, Theta, Delta, Phi, Tau, Omega, Zeta e Beta apresentou um padrão geral muito semelhante (Figura 6A). A relação entre o alinhamento realizado com base na estrutura tridimensional e sua relação com a estrutura secundária foi mostrado na Figura 6A. Observou-se uma concordância geral nas posições das alpha-hélices, nas folhas beta-pregueadas e nos loops, apesar dos diferentes aminoácidos encontrados em cada uma das cadeias.

Figura 4 - Reconstrução filogenética baseada na estrutura tridimensional das GSTs. Os códigos de quatro letras representados são decorrentes dos códigos PDB de cada estrutura amostrada

A Figura 6B mostra o alinhamento estrutural do cluster formado pelas classes Alpha, Sigma, Pi e Mu. Há uma alpha-hélice a mais na classe Alpha, na região C-terminal da enzima (seta verde). A classe Mu é caracterizada por apresentar num formato da letra grega “Mu” na região N-terminal da enzima (seta laranja).

A Figura 6C mostra o cluster formado pelas classes Zeta, Tau e Omega, em pontilhado foi representada a região N-terminal das GSTs. Finalmente, a Figura 6D, a qual apresenta o cluster formado pelas classes Theta, Delta, neste grupo foi incluída a classe Phi, apesar dos baixos valores de bootstrap apresentados. Há uma alpha-hélice destacada na região C-terminal da classe Theta, apontada com a seta azul.

						
	5	15	25	35	45	55
2gsr	-PPYTITYFP	VRG-RCEAMR	MLLADQDQSW	KEEVV-----	-----TMETW	PPLKPSCLFR
1aqw	-PPYTVVYFP	VRG-RCAALR	MLLADQGQSW	KEEVV-----	-----TVETW	QEGSLKASCL
1gtu	--PMILGYWD	IRG-LAHAIR	LLLEYTDSSY	EEKKY-----	-----TMGDA	PDYDRSQWLN
2gtu	--PMTLGYWN	IRG-LAHSIR	LLLEYTDSSY	EEKKY-----	-----TMGDA	PDYDRSQWLN
2gst	--PMILGYWN	VRG-LTHPIR	LLLEYTDSSY	EEKRY-----	-----AMGDA	PDYDRSQWLN
1gsu	--VVTILGYWD	IRG-LAHAIR	LLLEYTETPY	QERRY-----	-----KAGPA	PDFDPDWTN
2fhe	--PAKLGYWK	IRG-LQQPVR	LLLEYLGEEKY	EEQIY-----	-----ERD--	---DGEKWFS
1gta	-MSPILGYWK	IKG-LVQPTR	LLLEYLEEKY	EEHLY-----	-----ERD--	---EGDKWRN
1fhe	--PAKLGYWK	LRG-LAQPVR	LFLEYLGEEY	EEHLY-----	-----GRD--	---DREKWS
2gsq	-PKYTLHYFP	LMG-RAELCR	FVLAAGHEEF	TDRVV-----	-----EMA--	-----DWP
1m0u	-HSYTLFYFN	VKA-LAEPLR	YLFAYGNQY	EDVRV-----	-----TRD--	-----EWP
1f3a	-GKPVLYHFN	ARG-RMECIR	WLLAAAGVEF	EEKFI-----	-----Q----	-----SPED
1k3y	-EKPKLHYFN	ARG-RMESTR	WLLAAAGVEF	EEKFI-----	-----K----	-----SAED
1gul	--RPKLHYPN	GRG-RMESVR	WVLAAGVEF	DEEFL-----	-----E----	-----TKEQ
1b48	-AKPKLYYFN	GRG-RMESIR	WLLAAAGVEF	EEEFLL-----	-----E----	-----TREQ
1pn9	---MDFYYLP	GSA-PCRAVQ	MTAAAVGVEL	NLKLTL-----	-----DL-MK	-G--EH-MK-
1jlv	---MDFYYLP	GSA-PCRAVQ	MTAAAVGVEL	NLKLTL-----	-----NL-MA	-G--EH-MK-
1jlv	---MDFYYLP	GSA-PCRAVQ	MTAAAVGVEL	NLKLTL-----	-----NL--M	AG--EH-MK-
1r5a	-T-TVLYYLP	ASP-PCRSVL	LLAKMIGVEL	DLKVL-----	-----NI-ME	-G--EQ-LK-
1v2a	---MDYYYSL	ISP-PCQSAI	LLAKKLGITL	NLKKTL-----	-----NV---	-H--DPVER-
1ljr	MG-LEFLDL	VSQ-PSRAVY	IFAKKNGIPL	ELRTV-----	-----DLVK-	-G--QH-KS-
1eem	---IRIYSMR	FCP-FAERTR	LVLKAKGIRH	EVINI-----	-----NL-K-	-N--K---P-
1gnw	-G-IKVFHGP	ASI-ATRRVL	IALHEKNLDF	ELVHV-----	-----EL-K-	-D-GE---HK
1axd	-APMKLYGAV	MSW-NLTRCA	TALEEAGSDY	EIVPI-----	-----NF----	-ATAE---H-
1aw9	-APLKLYGMP	LSP-NVVRVA	TVLNEKGLDF	EIVPV-----	-----DL-T-	-T-GA---HK
1oyj	-KELVLLDFW	VSP-FGQRCR	IAMAEGLEF	EYREE-----	-----DL--G	-N--K---S-
1gwc	-DDLKLLGAW	PSP-FVTRVK	LALALKGLSY	EDVEE-----	-----DL--Y	-K--K---S-
1pmt	---MKLYY-T	PGS-CSLSPH	IVLRETGLDF	SIERIDL---	-RTK-KTESG	KDFLAINP--
1a0f	---MKLFY-K	PGA-CSLASH	ITLRESGKDF	TLVSVDL---	-MKK-RLENG	DDYFAVNP--
1f2e	---MKLFI-S	PGA-CSLAPH	IALRETGADF	EAVKVDL---	-AVR-KTEAG	EDFLTVNP--
1fwl	--KPILYS-Y	FRSSCSWRVR	IALALKGIDY	KTVPINLIKD	GGQQFSK---	-DFQALNP--
1e6b	--KLKLYSYW	RSS-CAHRVR	IALALKGLDY	EYIPV-----	-----NL-LK	-G--DQ-FD-

Figura 5 - Região N-terminal do alinhamento de GSTs proveniente dos dados estruturais das enzimas, realizado no programa PrISM (YANG; HONIG, 1999). As letras à esquerda indicam o código da estrutura no banco de dados do PDB

A.

B.

C.

D.

Figura 6 - Ilustração da cadeia principal referente ao alinhamento estrutural das classes de GSTs. (A) alinhamento total; (B) classes Alpha (verde), Sigma (branco), Pi (vermelho) e Mu (laranja), as setas indicam peculiaridades das classes Alpha (seta verde) e Mu (seta laranja); (C) classes Zeta (azul), Tau (verde) e Omega (laranja), a região N-terminal foi representada com pontilhado; (D) classes Theta (azul), Delta (laranja) e Phi (verde), seta indica alpha-hélice da classe Theta

4.5 Diferenças estruturais entre as isoformas da GST classe Phi

As GSTs da classe Phi apresentam, no sítio G, os seguintes resíduos conservados: Arg16, Glu66, Ser67 e Arg68, os quais são responsáveis pela afinidade à glutationa. Além disso, apresentam os resíduos Phe114, Ile118, Met121 e Leu122 também conservados no sítio H da enzima, os quais participam da afinidade pelo substrato hidrofóbico eletrofílico (p.e. atrazine).

A Figura 7A mostra o alinhamento estrutural da região N-terminal (sítio de ligação da GSH) das três isoformas da classe Phi, cujas estruturas foram resolvidas. Na posição 13 há uma Asn nas isoformas *ZmGSTF1* e *ZmGSTF3* e uma Ala na *AtGSTF4*. Na posição 40 há uma His nas isoformas *ZmGSTF1* e *AtGSTF4*, mas na *ZmGSTF3* há uma Leu na posição 36 que fica estruturalmente alinhada. No alinhamento de sequência a posição 40 da isoforma *ZmGSTF3* apresenta uma His, mas esta se localiza na região de loop desta isoforma e não deve participar da interação com a GSH (Figura 7A, retângulo da cadeia de cor azul). Na posição 54 a única alteração é de uma Ile encontrada na isoforma *ZmGSTF3* ao invés da Val das demais isoformas.

A Figura 7B mostra os diferentes resíduos que participam do sítio ativo das isoformas *ZmGSTF1*, *ZmGSTF3* e *AtGSTF4*. As isoformas *ZmGSTF3* e *AtGSTF4* apresentam uma Leu na posição 35 ou invés da Phe encontrada na isoforma *ZmGSTF1* (Figura 7B e 7C). A Phe35 é responsável pela estabilização do anel aromático do substrato hidrofóbico. Além disso, apresentam uma Phe na posição 123 (*ZmGSTF1* apresenta uma Ile118), uma Ile na posição 12 (*ZmGSTF1* apresenta o Trp12) e uma Ala na posição 10 (*ZmGSTF1*, Met10). Estas diferenças são fatores determinantes para as diversas especificidades resultantes das isoformas analisadas (Figura 7C).

Há uma diferença importante entre as duas estruturas resolvidas da *AtGSTF4* em *Arabidopsis* (Figura 7D). Uma estrutura foi cristalizada com o herbicida fluconzole (PDB: 1bx9), enquanto a outra apresenta a hexilglutationa (PDB: 1gnw) em seu sítio ativo. Apesar da mesma sequência de aminoácidos, há uma diferença na estrutura tridimensional da enzima quando se compara o alinhamento estrutural das mesmas. Refere-se primordialmente ao resíduo Phe na posição 122, cuja movimentação da cadeia lateral permite uma interação com o herbicida na estrutura com o PDB 1bx9 (Figura 7D). Esta observação mostra que a Phe movimenta sua cadeia lateral de modo a aumentar a acessibilidade do substrato com o solvente, o que provavelmente facilitaria a entrada do mesmo no sítio ativo da enzima.

A.

B.

C.

D.

Figura 7 - (A) alinhamento estrutural da região N-terminal da *ZmGSTF1* (verde), *ZmGSTF3* (azul), *AtGSTF4* (laranja). As posições e resíduos indicados se referem à *ZmGSTF1*. O retângulo indica a His40 da isoforma *ZmGSTF3*; (B) principais resíduos do sítio ativo das isoformas da *ZmGSTF1* (cinza) e *AtGSTF4* (laranja) alinhadas estruturalmente; (C) ampliação da Figura 7B, mostrando somente o sítio ativo das enzimas alinhadas; (D) alinhamento estrutural da isoforma *AtGSTF4* complexada (cadeia cor amarela e superfície marrom) e da apoforma (superfície representada em verde e cadeia na cor vermelha). O herbicida fluconazole conjugado com a GSH foi representado na cor vermelha

4.6 Alinhamento estrutural da GST classe Phi e classe Pi

A comparação das Figuras 8 e 9 mostrou algumas diferenças estruturais entre as GSTs *HmGSTP1* e *ZmGSTF1*. A estrutura tridimensional geral destas duas enzimas foi muito semelhante (Figura 8A). A observação do sítio ativo alinhado com o complexo ATA-GSH (Figura 8B) e com o EA (Figura 9B), mostrou a presença da Phe na posição 35 (F35; *ZmGSTF1*) alinhado com a Phe na posição 8 (F8; *HmGSTP1*). Além disso, o resíduo responsável pela catálise na classe Phi é a Ser na posição 11, ao passo que na classe Pi é a Tyr na posição 7. Observa-se que houve uma sobreposição destes resíduos no alinhamento estrutural, apesar de estarem relativamente distantes na sequência de aminoácidos em cada uma das proteínas. A *HmGSTP1* apresentou uma Tyr na posição 108 (Y108) e uma Asn na posição 204 (N204), ao passo que a *ZmGSTF1* apresentou o Trp na posição 12 (W12), Phe na posição 114 (F114) e uma Ile na posição 118 (I118). É interessante observar que tanto a Asn, como a Tyr são resíduos polares sem carga, embora a Tyr seja também um resíduo aromático. Ou seja, estes resíduos são mais propensos a realizarem ligações de hidrogênio. Por outro lado, o Trp, Phe e Ile são resíduos apolares, ou seja, essencialmente hidrofóbicos, estes podem apresentar ligações de hidrogênio, no entanto, somente com os átomos da cadeia principal (nitrogênio e oxigênio). É importante lembrar que a Phe também apresenta uma cadeia lateral aromática. A Figura 9A mostra os resíduos que podem ser responsáveis pela flexibilidade das duas cadeias representadas, as duas glicinas (Gly123 e Gly124) na GSTF1 e a extensão C-terminal da GSTP1, em relação à classe Phi, com os resíduos 202 a 207.

Figura 8 - (A) Representação das cadeias das enzimas *HmGSTP1* (laranja) e *ZmGSTF1* (cinza), alinhadas estruturalmente; (B) Ampliação do sítio ativo das enzimas alinhadas. O conjugado ATA-GSH se encontra no sítio ativo, com as cores CPK

Figura 9 - Alinhamento estrutural da *HmGSTP1* (laranja) e *ZmGSTF1* (azul). (A) A *HmGSTP1*, apresenta uma extensão na região C-terminal (Asn204) e a *ZmGSTF1* apresenta duas Gly (123 e 124) numa região de loop. (B) Sítio ativo das enzimas *HmGSTP1* (formato “bond”) e *ZmGSTF1* (formato “bola-linha”). Na cor branca foi representado o EA

4.7 Re-docking da glutathione na estrutura terciária da classe Pi

Com o objetivo de estabelecer a metodologia utilizada, foi realizado o docking da glutathione na própria estrutura resolvida por cristalografia (*HsGSTP1*; PDB 3gss), este método é também chamado de re-docking. Não houve diferença maior que 0,1 nm entre a estrutura resolvida por cristalografia, complexada com o ligante glutathione e o resultado do re-docking da estrutura com a glutathione (Tabela 3; Figura 10A₁). As maiores diferenças entre as distâncias da enzima e da GSH foram os resíduos: Phe8 (0,6 Å), Pro53 (0,8 Å), Asn66 (0,6 Å) e Glu97 (0,8 Å). Segundo esta metodologia, os resíduos que apresentaram ligações de hidrogênio com a GSH foram: Arg13, Trp38, Lys44, Gln51, Leu52, Gln64, Ser65, Asn66.

Tabela 3 - Interações entre enzima GSTP1 com a GSH, obtida por cristalografia e pelo método de docking (LIGIN; SOBOLEV et al., 1996)

Resíduo	Posição	Cristalografia		Docking	
		Distância (Å)	Superfície (Å ²)	Distância (Å)	Superfície (Å ²)
Tyr	7	3,9	11,9	3,4	19,1
Phe	8	3,3	26,6	3,8	17,7
Arg	13	3,4	38,3	2,9	44,1
Trp	38	2,9	43,2	2,9	40,8
Lys	44	3,1	24,0	3,4	19,5
Gly	50	4,0	0,7	3,9	1,1
Gln	51	2,9	60,1	2,5	69,1
Leu	52	2,9	43,3	2,7	46,0
Pro	53	3,5	14,2	4,3	3,5
Gln	64	2,6	43,0	2,9	38,5
Ser	65	2,8	42,0	2,7	47,7
Asn	66	4,6	4,2	4,0	8,7
Glu	97	5,0	2,3	4,2	6,0

4.8 Docking da glutathione na estrutura terciária da classe Phi

Foi realizado o docking da glutathione com a estrutura tridimensional da *ZmGSTF1* (PDB 1axd), além disso, o resultado foi comparado com uma estrutura resolvida por cristalografia, complexada com a S-hexilglutathione, *AtGSTF4* (PDB 1gnw) (Tabela 4; Figura 10B₁). Exceto pelo contato com a Arg16, que foi encontrada somente na estrutura *AtGSTF4*, de modo geral, os contatos com a porção da glutathione foram semelhantes. Os resíduos que formaram ligações de hidrogênio com a glutathione foram: Asn13, His40, Lys41, Gln53, Val54, Glu66, Ser67 e Arg68. A *AtGSTF4* apresentou uma ponte de hidrogênio, com a GSH, no resíduo Arg na posição 16 ao invés do Asn13.

Tabela 4 - Interações entre enzima *AtGSTF4* com a GSH, obtida por cristalografia e a *ZmGSTF1* pelo método de docking (SOBOLEV et al., 1996)

Resíduo	Posição	<i>AtGSTF4</i> - cristalografia		<i>ZmGSTF1</i> - docking	
		Distância (Å)	Superfície (Å ²)	Distância (Å)	Superfície (Å ²)
Asn	13	4,0	10,1	3,0	32,3
Arg	16	3,7	4,6	--	--
Phe	35	3,4	31,8	4,7	16,3
His	40	3,6	27,2	4,7	5,0
Lys	41	2,7	32,5	4,1	15,6
Gly	52	4,1	1,3	4,7	1,1
Gln	53	3,2	52,1	2,3	81,2
Val	54	3,0	58,2	2,0	50,9
Pro	55	3,6	12,1	3,9	4,3
Glu	66	2,8	35,3	2,6	39,8
Ser	67	2,7	40,6	2,3	45,5
Arg	68	3,4	18,7	2,1	49,7

A₁. A₂. B₁. B₂.

Figura 10 - Resíduos que participam da interação da GSH (representada com o formato VDW) com a GSTP1 (A₁) e GSTF1 (B₁). Os resíduos foram representados no formato “bond”. Comparação da possível transferência de elétrons na GSTP1 (A₂) e GSTF1 (B₂). As pontes de hidrogênio foram representadas com linhas pontilhadas. A GSH e os resíduos foram representados com o formato “bond”

4.9 Transferência de elétrons nas GSTs classes Phi e Pi

Na estrutura *HsGSTP1*, os resíduos que provavelmente participam do processo de transferência de elétrons no momento da conjugação da GSH com um composto eletrofílico hidrofóbico são, a Arg13, Ser65, Asn66, Glu97, Asp98 e Cys101 (Figura 10A₂). É importante lembrar que destes, os três primeiros apresentam ligações de hidrogênio com a GSH. Esta análise foi baseada no trabalho de Winayanuwattikun et al. (2005), os quais analisaram os mecanismos de transferência de elétrons na classe Delta das GSTS. Do mesmo modo, através da avaliação do alinhamento estrutural com a *HsGSTP1*, na estrutura *ZmGSTF1*, os resíduos que fizeram ligações de hidrogênio com a GSH foram a Arg16, Ser67 e Arg68. A Cys71, Lys72 e Glu102 apresentaram potenciais ligações de hidrogênio com os resíduos citados, por isso estes seis resíduos citados podem ser fortes candidatos a participantes da transferência de elétrons no evento do ataque nucleofílico da Ser11 na GSH (Figura 10B₂). A confirmação destes dados poderá ser feita através de mutações sítio-dirigidas nestes resíduos.

4.10 Re-docking do ácido etacrínico na estrutura terciária da classe Pi

O sítio H (H: hidrofóbico) é o sítio que determina a especificidade das diversas classes das GSTs. A Tabela 5 mostra as semelhanças e diferenças encontradas nos resíduos responsáveis pela interação com o segundo substrato da enzima na GSTP1, novamente esta análise foi realizada para comparar os dados obtidos por cristalografia com a metodologia utilizada no programa de docking. Estes dados foram obtidos com o programa LIGIN (SOBOLEV et al., 1996) primeiramente através do re-docking da GSH no sítio ativo da enzima, seguido do re-docking do ácido etacrínico (EA). O programa LPC (SOBOLEV et al., 1999) foi utilizado para analisar os contatos dos ligantes com a enzima na estrutura resolvida por cristalografia, PDB 3gss, cujos ligantes são a GSH e o EA. Os principais resíduos que compõem a superfície de contato da classe Pi são: Phe8, Val10, Ile104, Tyr108, Thr109, Gly205. Os resultados dos contatos foram concordantes, observaram-se duas ligações de hidrogênio entre a enzima e o EA, uma com a hidroxila do resíduo Tyr7 e outra com a hidroxila da Tyr108 (Figuras 11A₁ e A₂).

Tabela 5 - Interações entre enzima *HsGSTP1* com o EA, obtido por cristalografia e pelo método de docking (SOBOLEV et al., 1996)

Resíduo	Posição	Cristalografia		Docking	
		Distância (Å)	Superfície (Å ²)	Distância (Å)	Superfície (Å ²)
Tyr	7	2,8	26,5	3,5	19,7
Phe	8	3,1	32,5	3,4	37,2
Val	10	2,9	41,7	3,7	24,0
Ile	104	3,3	36,7	3,6	30,9
Tyr	108	3,0	93,0	3,1	89,9
Thr	109	3,8	13,2	3,6	17,5
Gly	205	4,2	11,7	3,9	17,5
GSH	--	1,8	75,7	1,8	73,7

4.11 Docking do ácido etacrínico na estrutura terciária da classe Phi

Na classe Phi foram analisados as cinco estruturas melhores ranqueadas (maior Função de Complementariedade - CF) no método de docking. Destas, quatro posições assumiram uma posição muito semelhante com a melhor ranqueada, deste modo a estrutura com maior valor CF foi escolhida como modelo. Além disso, com o intuito de aumentar os graus de liberdade do sistema analisado (ou seja, permitir a mobilidade tanto da enzima como do ligante), a estrutura do EA escolhida juntamente com a enzima *ZmGSTF1* e a GSH foram submetidas a uma minimização de energia. A Tabela 6 mostra os contatos entre o EA e a enzima antes e após a minimização de energia. Algumas diferenças foram observadas, como a ausência de contatos do EA com a Ala8, Leu14 e Val54. Além disso, o EA apresentou um distanciamento e uma redução na superfície de contato com a Ser11, Phe35 e Phe114 e houve uma aproximação em relação aos resíduos Asn13, Asn110 e Ile118. Os resíduos envolvidos na formação da superfície de contato com o ligante EA, após a minimização foram: Met10, Ser11, Trp12, Asn13, Phe35, Asn 110, Phe114, Ile118, Met121 e Leu122. As duas ligações de hidrogênio observadas no método de docking foram com o átomo de oxigênio da Ser11 e a outra com o nitrogênio do Trp12 (Figuras 11B₁ e B₂), este tipo de ligação não foi observada após a minimização de energia. O Trp12 e a Phe 35 apresentaram interações aromáticas com o ligante EA em ambos os métodos analisados. Os demais resíduos participaram com interações hidrofóbicas (forças de Van der Waals) com o ácido etacrínico, exceto pela Asn110 a qual apresentou contatos desestabilizadores. O grupo carboxila do EA apresentou contatos com a Asn13 e Asn110. O grupo butiril do EA apresentou interações hidrofóbicas com a Ile118. Os átomos de cloro do EA fizeram contato com os resíduos Phe114 (4,1 Å) e Ile118 (3,3 Å). O grupo cetona do EA não mostrou contato com a enzima nos métodos utilizados.

Tabela 6 - Comparação entre os contatos do EA com a enzima *ZmGSTF1* utilizando-se o método de docking e docking seguido de minimização de energia (EM)

Resíduo	Posição	Docking		Docking + EM	
		Distância (Å)	Superfície (Å ²)	Distância (Å)	Superfície (Å ²)
Ala	8	3,6	29,2	--	--
Met	10	4,7	6,1	4,0	48,0
Ser	11	2,2	77,2	4,3	13,7
Trp	12	3,7	33,6	4,0	31,9
Asn	13	3,2	22,2	2,7	37,5
Leu	14	3,4	8,3	--	--
Phe	35	4,0	30,7	5,3	17,5
Val	54	3,0	26,9	--	--
Asn	110	4,4	20,6	2,7	55,1
Phe	114	2,2	55,4	4,1	28,5
Ile	118	3,7	24,8	3,3	79,4
Met	121	--	--	4,3	20,0
Leu	122	4,1	10,6	4,9	12,8

A₁. A₂. B₁. B₂.

Figura 11 - (A₁ e A₂) Re-docking do ácido etacrínico (EA, representado no formato VDW) na estrutura *HsGSTP1*, na presença da GSH (representada no formato linha-bola); (B₁ e B₂) Docking seguido de minimização de energia do EA na enzima *ZmGSTF1*, com a presença da GSH

4.12 Modelagem da GST de cana-de-açúcar e escolha dos mutantes

Com o intuito de engenheirar uma enzima da superfamília das GSTs, visando uma alteração na afinidade da enzima, foi escolhido um gene proveniente de uma das culturas de maior importância econômica para o Brasil e cujo transcriptoma encontra-se sequenciado, a GSTF1 de cana-de-açúcar. A busca da GST de cana-de-açúcar foi feita a partir da *ZmGSTF1*, de milho, enzima reconhecida pela literatura como amplamente capaz de desintoxicar, com eficiência, o herbicida atrazine (ATA). Este herbicida foi escolhido devido à quantidade de informações estruturais que o mesmo apresenta, bem como devido ao papel que a atrazine representa e principalmente representou na agricultura brasileira. Devido a estes fatores foi considerado um modelo adequado para o desenho racional de uma enzima de cana-de-açúcar, partindo das informações tridimensionais da enzima molde e do ligante. A *ZmGSTF1* foi utilizada como molde para a busca de GSTs no banco de dados de seqüências de cana-de-açúcar (Projeto Genoma da Cana-de-Açúcar, SUCEST). A seqüência encontrada em cana-de-açúcar com maior semelhança com a *ZmGSTF1* apresentou 94% de identidade de seqüência, ou seja, 201 aminoácidos, dos 213 encontrados na proteína, foram idênticos à seqüência molde (Figura 12). Foi obtida uma estrutura quaternária da enzima *SoGSTF1* de cana-de-açúcar a partir da estrutura tridimensional da *ZmGSTF1* (Figura 13A). Em seguida, foi realizado um “hard-docking” da ATA no sítio ativo da *SoGSTF1*, ou seja, como as enzimas molde e modelo são muito semelhantes pôde-se alocar o ligante da *ZmGSTF1* na *SoGSTF1* após o alinhamento das mesmas, sem que houvesse um impedimento estéreo-químico (superposição de átomos). Desta forma, o sítio ativo da enzima *SoGSTF1* pode ser analisado, bem como comparado com os dados existentes na literatura com relação à GST de milho, assim os resíduos participantes do sítio ativo foram: Met10, Trp12, Phe35, Phe114, Ile118 e Met121 (Figura 13). A diferença encontrada no sítio ativo das duas enzimas em questão foi o resíduo encontrado na posição 10, na *ZmGSTF1* este resíduo é uma metionina e na *SoGSTF1* uma leucina (Figura 12). A partir das informações dos resíduos foram propostos mutantes para cada uma destas posições. O critério de alteração do aminoácido baseou-se na mudança de um resíduo de cadeia lateral alifática para uma fenilalanina, assim Met10 foi mutada para Met10Phe, Ile118 para Ile118Phe e Met121 para Met121Phe. Por outro lado, os resíduos de cadeia aromática (triptofano e fenilalanina) foram alterados para uma leucina, do mesmo modo, Trp12 foi mutado para Trp12Leu e Phe35 e Phe114

ficaram Phe35Leu e Phe114Leu. Cada mutação foi realizada individualmente. Para analisar a interação dos mutantes com o herbicida atrazine foram efetuadas dinâmicas moleculares para cada situação e também para o controle (*SoGSTF1*).

```

S.offic:1      APMKLYGAVLSWNVTRCATALEEAGSDYEIVPINFATAEHKSPEHLARNPFGQVPALQDG 60
Z. mays:1      APMKLYGAVMSWNLTRCATALEEAGSDYEIVPINFATAEHKSPEHLVRNPFGQVPALQDG 60

S.offic:61     DLYLFESRAICKYAARKNKPELLREGNLEESAMVDVWMEVEANQYTSVLNPILFQVLISP 120
Z. mays:61     DLYLFESRAICKYAARKNKPELLREGNLEEAAMVDVWIEVEANQYTAALNPILFQVLISP 120

S.offic:121    MLGGTTDQKVVDENLEKLKKVLEVYEARLTKSKYLAGDFLSLADVNHSVTLCLLATPYA 180
Z. mays:121    MLGGTTDQKVVDENLEKLKKVLEVYEARLTKCKYLAGDFLSLADLNHSVTLCLFATPYA 180

S.offic:181    SVLDAYPHVKAWWSSLMERPSVQKVAAMMKPSA 213
Z. mays:181    SVLDAYPHVKAWWSGLMERPSVQKVAALMKPSA 213

```

Figura 12 - Alinhamento da GST de milho (*Z. mays*) e de cana-de-açúcar (*S.offic*), os resíduos em negrito indicam participação no sítio ativo da enzima, particularmente no sítio H

Figura 13 - (A) Modelo da forma dimérica da GST de cana-de-açúcar (*SoGSTF1*). (B) Monômero da proteína modelada *SoGSTF1*, após a alocação do conjugado atrazine-glutathione (ATA-GSH) no sítio ativo da enzima. (C) Monômero da proteína modelada *SoGSTF1*, salientando os resíduos do sítio ativo e o conjugado ATA-GSH. (D) Visualização dos principais resíduos participantes do sítio ativo da enzima modelada *SoGSTF1*, com o conjugado atrazine-glutathione (destacado no quadro à direita)

4.13 Dinâmica molecular na GST de cana-de-açúcar e nos mutantes

Dentre os resultados obtidos utilizando-se o método da dinâmica molecular, efetuada nos mutantes e no controle, duas situações apresentaram resultados mais significativos em relação ao controle, a saber: o mutante Ile118Phe (I118F) e o Trp12Leu (W12L). A Figura 14 mostra os gráficos de energia potencial do controle (A_1) e dos mutantes W12L (B_1) e I118F (C_1). Estes gráficos mostram que o sistema apresentou uma redução seguida de uma estabilização na energia após aproximadamente 600 ps. Com base nestes dados foram analisados os gráficos de desvio RMS no controle (A_2) e dos mutantes W12L (B_2) e I118F (C_2). O desvio RMS corresponde à distância entre os átomos da cadeia principal de duas estruturas. Como se trata de uma dinâmica, as estruturas correspondem à posição dos átomos no tempo i comparada com a posição no tempo $i-1$. A dinâmica molecular das proteínas, de modo geral, apresentou um desvio RMS constante, aproximadamente de 0,25 nm. No entanto, a atrazine apresentou um comportamento diferenciado em cada situação. No controle houve uma estabilização da ATA entre 500 e 850 ps, seguida por uma movimentação média de 0,31 nm após 850 ps. O mutante W12L não apresentou a mesma estabilização da ATA encontrada no controle e apresentou uma movimentação significativa no decorrer da dinâmica, principalmente a partir de 600 ps, com um desvio RMS de 0,35 nm. Por outro lado, o mutante I118L apresentou uma estabilização da molécula de ATA no sítio ativo da enzima, com um desvio RMS de 0,26 nm.

A₁ A₂ B₁ B₂ C₁ C₂

Figura 14 - Energia potencial (kJ.mol^{-1}) no decorrer do tempo (ps) da enzima controle (A1) e dos mutantes W12L (B1) e I118F (C1). O desvio RMS da proteína (—) e da ATA (—◆—) foram representados, no controle (A2) e nos mutantes W12L (B2) e I118F (C2)

A Tabela 7 apresenta os tipos de interações entre a ATA e a enzima controle no tempo de 750 e 1000 ps. Foram escolhidas a posição final (1000 ps) e uma estrutura intermediária de forma a visualizar a movimentação do ligante no sítio ativo da enzima. A Ser11 apresentou potenciais ligações de hidrogênio com o herbicida no tempo de 1000 ps. Os resíduos Trp12, Phe35 e Phe114 apresentaram interações aromáticas e a Ile118 apresentou interação hidrofóbica com a ATA. A Val na posição 54 não apresentou interação, apesar de estar a 3,3 Å da ATA. A Asn13 apresentou uma interação neutra com a ATA. Estas interações são consideradas potenciais devido à própria limitação da metodologia, a qual permite uma avaliação menos rigorosa com relação às distâncias encontradas entre os átomos, comparando com uma estrutura obtida por cristalografia. Assim, as distâncias entre os átomos analisados podem ser ligeiramente maiores que as encontradas na prática.

A Tabela 8 apresenta os contatos da ATA com a enzima mutante Trp12Leu. A ATA apresentou três potenciais ligações de hidrogênio, uma com a Ser11, com uma distância de 4,3 Å, outra com a Leu12 numa distância de 3,8 Å e a terceira com a Ile118 à 4,2 Å. Houve uma interação do tipo aromática com a Phe114 e uma interação hidrofóbica com a Gln53 (oxigênio da cadeia lateral deste resíduo com o grupo propil da ATA). Nota-se que neste mutante não há a interação aromática do Trp12 devido à mutação do mesmo e nem com a Phe35.

A Tabela 9 mostra interações aromáticas da ATA com o Trp12, a Phe35 e a Phe118. Há possíveis ligações de hidrogênio com a Ser11, o Trp12 e a Val54, além de interações hidrofóbicas com a Leu10, Asn13, Phe114, Met121 e Leu122. A observação dos contatos entre o ATA e as enzimas mostram um maior número de contatos com o mutante I118F em relação ao controle. Este dado pode resultar numa afinidade maior desta enzima pela ATA. Como a conjugação já foi efetuada, a GSH e a ATA estão ligadas através de uma ligação covalente, em última estância esta afinidade poderia resultar numa redução da velocidade catalítica da enzima. Neste ponto é importante lembrar o papel catalítico das GSTs, as quais catalisam a conjugação de um composto eletrofílico com a GSH de modo a aumentar a hidrofilicidade do produto. Este produto, por sua vez, deve sair do sítio ativo e dirigir-se para um vacúolo, no caso das plantas. Se há uma elevada afinidade pelo produto gerado com a enzima, este teoricamente deve sair mais lentamente do sítio ativo da mesma, impedindo ou reduzindo o número de novos substratos que serão conjugados e transformados em produtos, por unidade de tempo. A Figura 15 ilustra os contatos do conjugado ATA-GSH com a enzima controle (A₁ e A₂), mutante Trp12Leu (B₁ e B₂) e Ile118Phe (C₁ e C₂).

Tabela 7 - Distância e tipo de interações dos átomos da ATA com os resíduos da enzima controle no tempo de 750 e 1000 ps (onde, HB = interação de hidrogênio e Arom. = interação aromática)

Controle		750 ps		1000 ps	
		Distância (Å)	Tipo de Interação	Distância (Å)	Tipo de Interação
11	Ser	3,5	HB	3,3	HB
12	Trp	4,3	HB/Arom.	4,6	Arom.
13	Asn	4,0	Hidrofóbica	2,8	--
35	Phe	3,5	HB/Arom.	3,3	Arom.
54	Val	3,3	--	3,5	--
114	Phe	5,1	Hidrofóbica	3,3	Arom.
118	Ile	4,1	HB	4,0	Hidrofóbica

Tabela 8 - Distância e tipo de interações dos átomos da ATA com os resíduos da enzima mutante Trp12Leu no tempo de 750 e 1000 ps

W12L		750 ps		1000 ps	
		Distância (Å)	Tipo de Interação	Distância (Å)	Tipo de Interação
11	Ser	--	--	4,3	HB
12	Leu	--	--	3,8	HB
35	Phe	3,6	HB	4,8	--
53	Gln	--	--	3,6	Hidrofóbica
114	Phe	4,8	Hidrofóbica	3,5	Arom.
118	Ile	5,7	Hidrofóbica	4,2	HB
122	Leu	5,2	Hidrofóbica	--	--

Tabela 9 - Distância e tipo de interações dos átomos da ATA com os resíduos da enzima mutante Ile118Phe no tempo de 750 e 1000 ps (DC = contatos desestabilizador)

I118F		750 ps		1000 ps	
		Distância (Å)	Tipo de Interação	Distância (Å)	Tipo de Interação
8	Ala	4,0	DC	3,8	--
10	Leu	3,4	Hidrofóbica	3,1	Hidrofóbica
11	Ser	3,4	HB	3,3	HB
12	Trp	3,4	Hidrofóbica	4,3	HB/Arom.
13	Asn	3,4	Hidrofóbica	2,8	Hidrofóbica
35	Phe	3,5	HB/Arom.	3,0	Arom.
54	Val	3,6	HB	3,8	HB
114	Phe	3,8	Hidrofóbica	3,9	Hidrofóbica
118	Phe	3,1	Arom.	3,2	Arom.
121	Met	5,4	Hidrofóbica	3,7	Hidrofóbica
122	Leu	--	--	3,8	Hidrofóbica

A₁ A₂ B₁ B₂ C₁ C₂

Figura 15 - Contatos do conjugado ATA-GSH (formato “linha”) com a enzima controle (A₁ e A₂), mutante Trp12Leu (B₁ e B₂) e Ile118Phe (C₁ e C₂) no tempo de 750 e 1000 ps. Os resíduos de cada enzima foram representados no formato “linha-bola”

4.14 Construções gênicas em vetor para expressão das proteínas heterólogas em bactéria

Com o objetivo de checar a atividade das mutações realizadas *in silico* foram efetuadas construções através de mutações sítio-dirigidas. Duas foram as mutações propostas, Trp12Leu e Ile118Phe, conforme especificado no item anterior. As construções gênicas realizadas foram esquematizadas abaixo:

Trp12Leu:

5' GGG ^{NcoI} GCC ATG GAA ATG GCT CCG ATG AAG CTG TAC GGG GCG GTG CTG TCG **TTG** AAC ...
 Met Ala Pro Met Lys Leu Tyr Gly Ala Val Leu Ser **Leu** Asn
 1 2 3 4 5 6 7 8 9 10 11 **12** 13

e

Ile118Phe:

5' AAT CCC ATC CTC TTC CAG GTC CTC **TTC** AGT CCG ... 3'
 Asn Pro Ile Leu Phe Gln Val Leu **Phe** Ser Pro
 110 111 112 113 114 115 116 117 **118** 119 120

ambas as construções foram seguidas da terminação abaixo, correspondente ao pET28b:

5' ... GAT CCG AAT TCG AGC TCC GTC GAC AAG CTT GCG GCC GCA CTC GAG CAC CAC CAC CAC CAC CAC TGA
 Asp Pro Asn Ser Ser Ser Val Asp Lys Leu Ala Ala Ala Leu Glu **His His His His His His** stop

Foram selecionadas duas colônias referentes ao mutante Trp12Leu e seis colônias referentes ao mutante Ile118Leu (Figura 16). A extração do plasmídeo seguida de tratamento com as enzimas de restrição mostram um inserto com aproximadamente 600kb (Figura 17). A confirmação das mutações foi verificada através de sequenciamento.

Figura 16 - Análise por PCR das colônias transformadas. PM, marcador molecular (1kb); colônias de 1-5 referentes aos mutantes W12L; colônias de 1-7 referentes aos mutantes I118L; CP, controle positivo; CN, controle negativo

Figura 17 - Análise de restrição dos fragmentos referentes aos mutantes selecionados. CP, controle positivo; PM, marcador molecular (1kb)

5 DISCUSSÃO

5.1 Padrão da superfamília das GSTs

Um dos objetivos desta tese foi tentar estabelecer uma relação entre a seqüência, estrutura tridimensional, função e afinidade das proteínas, particularmente dentro da superfamília das glutatonas transferases (GSTs). Neste contexto consideramos duas situações (Figura 18).

Figura 18 - Relações encontradas entre as seqüências, estruturas, funções e afinidades das enzimas. Padrões de evolução encontrados na superfamília das GSTs, divergência funcional (esquerda) e convergência funcional (direita)

A primeira situação aborda a divergência funcional, as seqüências de aminoácidos, dentre as classes de GST, são caracterizadas por apresentar identidades muito baixas (por volta de 20 % de identidade de seqüência), no entanto, de acordo com o alinhamento estrutural apresentado nesta tese, bem como com o descrito na literatura (ARMSTRONG, 1997; HAYES, 2005; McGOLDRICK; O'SULLIVAN; SHEEHAN, 2005; SHEEHAN et al., 2001) a conformação geral da cadeia principal das proteínas apresenta uma alta similaridade (menos que 3 Å de diferença do desvio RMS). Além disso, as GSTs são classificadas como transferases, portanto sua função básica é de transferir, particularmente a GSH para um composto eletrofílico, hidrofóbico, deste modo é caracterizada por apresentar uma função clara e bem definida. No entanto, as

afinidades dentre as classes de GST é bastante distinta, como ficou esclarecido na revisão bibliográfica citada nesta tese.

Com relação à divergência funcional (Figura 18), esta poderá ser explorada como a situação mais comum encontrada na natureza dentro da superfamília de proteínas. Foi subdividido em quatro etapas para facilitar a explanação. O passo a_1 aborda as seqüências diferentes com estruturas semelhantes. O paradigma atual considera que a estrutura é mais conservada que a seqüência, devido a este mesmo motivo esta informação é utilizada como um critério para modelagem de proteínas por homologia (GUERX; DIEMAND; PEITSCH, 1999; HOLM; SANDER, 1995; MARTÍ-RENOM et al., 2000; PONTING; RUSSELL, 2002; YANG; HONIG, 1999). No passo b_1 as seqüências diferentes apresentam afinidades diferentes. Fato que não apresenta surpresa na sua afirmação considerando que esta pode ser uma condição para explicar as diferentes atividades específicas das enzimas. O passo c_1 avalia estruturas semelhantes com funções semelhantes, o que é considerado um paradigma na bioinformática estrutural (SHINDYALOV; BOURNE, 1998; THOMPSON et al., 1997; WHISSTOCK; LESK, 2003). Finalmente no passo d_1 , as estruturas são semelhantes, no entanto, as afinidades são diferentes. Este será levantado como um padrão encontrado na natureza, devido principalmente a dois fatores. Primeiro, o número de conformações estruturais é inferior às diferentes atividades específicas encontradas. Atualmente especula-se que deverá haver por volta de 30.000 genes nos humanos (no mínimo), sem considerar os alelos diferentes para cada locus, os quais apresentam especificidades diferentes, para um número de conformações espaciais ao redor de 1000 superfamílias de proteínas, classificadas pelo SCOP (MURZIN et al., 1995). Devem existir, em média, 30 genes diferentes para cada superfamília de proteína, lembrando que as superfamílias são agrupadas de acordo com sua semelhança estrutural. Levando-se em consideração que algumas das superfamílias descritas no SCOP podem não ser encontradas em humanos e que o número de genes deve estar sub-estimado. O número de genes por superfamília de proteína deve ser muito superior ao calculado acima. Além disso, dentro das superfamílias as proteínas são classificadas em famílias, as quais apresentam funções similares, mas afinidades diferentes, o que corrobora para a hipótese de que a estrutura tridimensional determina a função (geral e inespecífica) da enzima, no entanto há ainda um longo caminho a percorrer no que se refere à determinação da afinidade de uma enzima. Ou seja, são detalhes nas cadeias laterais dos resíduos de uma proteína, que faça parte, ou não, do sítio ativo da enzima, que determinará sua afinidade

por determinado substrato. Certamente, esta é a grande meta da área relacionada com o desenho racional de drogas (área farmacêutica) ou de proteínas (área da biologia molecular). Esta linha de pesquisa ganhou maior ênfase a partir do ano de 1996, com o surgimento de computadores cada vez mais velozes e programas de modelagem moleculares cada vez mais acurados. Um dado que podemos concluir desta análise é o fato de que para estudar as relações de afinidade de uma enzima é necessário avaliar as interações proteína-ligante individualmente. Ou seja, certamente não será uma simples semelhança na seqüência de uma proteína ou na sua estrutura que definirá a afinidade de uma enzima.

Um outro argumento que pode contribuir para a asserção de que a relação entre a estrutura e a afinidade não é direta, baseia-se no fato de que algumas enzimas apresentam uma característica promíscua ou uma baixa especificidade. Ou seja, a enzima catalisa reações com substratos diferentes e não apresenta uma especificidade exclusiva por um determinado ligante. Neste ponto podemos lembrar a afirmação de Darwin sobre o fato de que a evolução ocorre de forma gradativa (DARWIN, 2004). A evolução gradativa poderia resultar numa promiscuidade da enzima. Esta poderia, ainda, ser explicada através da co-evolução dos aminoácidos numa proteína (SOCOLISH et al., 2005), na qual os autores concluem que a maioria dos aminoácidos muda, mas a estrutura tridimensional é mantida. Ou seja, há um pequeno número de aminoácidos cujas interações proporcionam a estrutura tridimensional da proteína. Estas modificações na seqüência da enzima podem promover pequenas alterações na especificidade da enzima sem provocar uma alteração estrutural, de modo que resultem numa evolução gradual. Outro fato a ser considerado é que a pressão de seleção ocorre na função da enzima e não em sua estrutura (RUSS et al., 2005). Posto isso, os eventos de duplicação de genes (parcial ou completa), as mutações e recombinações do genoma e a seleção natural, ao se encarregarem de selecionar os organismos mais adaptados, poderão, ou não, proporcionar uma especialização dos parálogos no decorrer do tempo.

A segunda situação encontrada trata-se da convergência funcional das enzimas (Figura 18), na qual seqüências diferentes podem apresentar afinidades semelhantes. Este último, será explorado no decorrer desta discussão, de modo a demonstrar a evolução convergente funcional das GSTs classes Phi (plantas) e Pi (metazoários).

5.2 Evolução da superfamília das GSTs

Comparando as reconstruções filogenéticas baseadas nos alinhamentos de seqüências de proteínas e de estruturas, observou-se que as classes de GST Alpha, Sigma, Mu e Pi formaram um cluster bem definido. O mesmo relato foi feito por Sheehan et al. (2001) e Snyder e Maddison (1997), no entanto, nesta tese a classe Alpha foi considerada o grupo externo (Figuras 3 e 4) e não a classe Sigma como considerada pela literatura (SHEEHAN et al., 2001 e SNYDER; MADDISON, 1997). Outro cluster foi composto pelas classes Tau, Omega e Zeta, sendo esta última o grupo externo. McGoldrick também encontrou este cluster, embora tenha obtido a classe Omega como grupo externo. Neste aspecto é importante salientar que McGoldrick; O'Sullivan e Sheehan trabalharam com uma amostragem de apenas 10 seqüências relativas a estas classes de GST (McGOLDRICK; O'SULLIVAN; SHEEHAN, 2005). As classes Theta e Delta formaram um terceiro cluster e a classe Phi ficou como um grupo externo a este cluster, embora o valor de bootstrap tenha sido baixo (496). McGoldrick; O'Sullivan e Sheehan (2005) encontraram um cluster formado pelas classes Phi, Theta e Delta. No entanto, a metodologia utilizada foi o programa ClustalX tanto para o alinhamento como para a reconstrução filogenética, além disso, o foco visava as GSTs de fungos, assim a amostragem das demais classes de GST foi bastante reduzida. Ding e colaboradores (2003) estudaram as GSTs de insetos e chegaram a um agrupamento da classe Theta com a classe Delta, no entanto, não incluíram as GSTs classe Phi.

O cluster formado pelas classes Alpha, Mu e Pi, antes caracterizadas pelo grupo das GST de mamíferos (SHEEHAN et al., 2001), apresenta representantes dos filos Craniata, Echinodermata, Chelicerata, Mandibulata, Nematoda, Platyhelminthes, ou seja, o reino Animalia. Relembrando que no presente trabalho a classe Sigma foi incluída neste cluster, filogeneticamente mais próxima das classes Mu e Pi do que a classe Alpha.

O cluster formado pelas classes Tau, Omega e Zeta apresenta representantes de diferentes reinos. A classe Tau está presente nos filos Coniferophyta e Magnoliophyta, ou seja, representantes do reino Plantae. A classe mais próxima filogeneticamente da classe Tau é a classe Omega com representantes do filo Echinodermata e Craniata. Finalmente a classe Zeta a qual apresenta organismos do reino Plantae (Magnoliophyta) e Animalia (Craniata, Echinodermata e Nematoda). Assim, neste cluster tanto a classe Tau, como a classe Omega devem ter sofrido o processo de duplicação, provavelmente da classe Zeta. No entanto, esta classe pode ter se

diferenciado e até mesmo sofrido novas duplicações após o evento de divergência entre as plantas e os animais. Deste modo, as classes Tau e Omega seriam mais recentes que a classe Zeta (Figura 20).

A classe Zeta pode ter sido a GST ancestral do cluster das classes Alpha, Sigma, Mu e Pi, no entanto, o evento da duplicação e diversificação deste cluster deve ter ocorrido após a divergência das plantas e animais, pois não são encontradas GST de plantas no cluster da classe Alpha. Novamente, neste caso o cluster formado deve ser mais recente que a classe Zeta. Provavelmente o ancestral do filo Animalia, já continha as quatro classes Alpha, Sigma, Mu e Pi. Neste caso, a classe Alpha, pode ter sido perdida nos filos Platyhelminthes, Chelicerata e Nematoda. O mesmo pode ter ocorrido com a classe Pi para os filos Platyhelminthes e Chelicerata. Há ainda a hipótese do ancestral da classe Zeta apresentar representantes do reino Fungi, mas as seqüências disponíveis ainda não são suficientes para a sua comprovação.

No caso das classes Phi, Theta e Delta, é mais difícil tentar determinar a classe da GST ancestral. Deve ter sido uma GST ancestral das Eubacteria e Archaea, a qual sofreu o processo de duplicação até se diferenciar nas classes Phi, Theta e Delta. No entanto, pode-se inferir que a classe Phi surgiu no evento da formação do reino Plantae, pois não existem organismos do reino Animalia ou Fungi nesta classe de GST. Como não há, até o momento, seqüências de GST nos filos inferiores do reino Plantae, como Cycadophyta e Bryophyta, não podemos inferir ao certo em que momento a classe Phi surgiu e se diferenciou. Do mesmo modo, a classe Delta deve ter surgido com o evento de diferenciação dos insetos, já que esta classe de GST é exclusiva deste grupo. Na classe Theta, por outro lado, o evento da duplicação e diversificação deve ter ocorrido antes da divergência entre plantas e animais, de modo que esta classe é encontrada nestes filos. A Figura 19 apresenta os filos da vida segundo Margulis e Schwartz (2001) o qual foi usado como critério para a construção da Figura 20. Esta mostra uma possível rota evolutiva das classes de GST. A relação entre as classes Phi e Theta não foi estabelecida devido aos baixos valores de bootstrap encontrados.

Figura 19 – Filos da vida na Terra baseados no sistema dos cinco reinos e na teoria simbiótica da origem das células eucarióticas (MARGULIS; SCHWARTZ, 2001)

Com relação às classes Pi e Phi, vejamos os dois casos de evolução possíveis a partir do provável ancestral destas GSTs (Figura 21). No caso 1, o evento da duplicação ocorreria no próprio ancestral de plantas e metazoários, assim os descendentes herdariam os genes A e B e estes divergiriam no decorrer do processo evolutivo. Este parece ter sido o caso das classes Zeta e Theta, as quais são encontradas tanto em plantas quanto nos animais (Figura 20). No caso 2, o evento de duplicação desta classes ocorreu após o evento de especiação, deste modo as classes Pi e Phi (genes B e C) seriam enzimas parálogas e não ortólogas. O que parece mais plausível devido ao fato de não serem encontradas GSTs classe Pi em plantas ou classe Phi em animais.

Figura 20 - Possível rota evolutiva encontrada na superfamília das GSTs. As GSTs do reino Plantae foram representadas com o fundo azul, o reino Animalia, em roxo, o reino Fungi foi identificado com a cor laranja e GSTs de Eubacteria em rosa

Figura 21 - Duas hipóteses evolutivas para as classes Pi e Phi de GSTs. As letras A, B e C correspondem a genes de GST

5.3 Convergência funcional e a relação tridimensional entre as classes Phi e Pi

5.3.1 Atividade enzimática relativa às classes Phi e Pi

Uma revisão na atividade específica das classes de GST mostra que as classes Pi e Phi compartilham as mesmas reações catalíticas (Tabelas 10 e 11). Cada organismo apresenta sua própria isoenzima com atividade específica para a atrazine e para o ácido etacrínico (EA). Nestas comparações, a ordem de grandeza foi usada, em vez do valor exato da estimativa, de modo a minimizar os eventuais erros provenientes de diferentes metodologias adotadas.

Egaas et al. (1995) mostraram que a conjugação da atrazine com a glutathione foi observada na classe Pi de camundongos e as classes Alpha e Mu não apresentaram atividade para este substrato. Islam; Hara; Miyake (2002) estudaram os efeitos da atrazine administrada no intra-peritônio hepático de ratos. Estes autores concluíram que a atrazine é um indutor da classe Pi de GST. As atividades específicas das classes Alpha, Mu, Pi e Theta de humanos e camundongos de GST foram testadas com o herbicida atrazine (ABEL et al., 2004). Estes autores demonstraram que apenas *HsGSTP1-1* e *ZmGSTP1-1* apresentaram uma atividade significativa para este substrato (7,1 nmol/min/mg de proteína em humanos e 7,3 nmol/min/mg de proteína em camundongos). Deste modo, ficou demonstrado que a atrazine é um substrato específico para a classe Pi de GSTs.

Em milho, a GSTF1 (29 kDa) apresentou especificidade pela atrazine, enquanto a GSTF3 (27 kDa) foi responsável pela conjugação da GSH com o alachlor e o metolachlor (NEUEFEIND et al., 1997a; O'CONNELL; BREAU; FRALEY, 1988; SHIMABUKURO; SWANSON; WALSH, 1970). A GSTF1 foi a única isoforma de GST encontrada em milho que apresentou atividade significativa para a atrazine (PRADE; HUBER; BIESELER, 1998). A atividade específica da GSTF1-1 (homodímero da isoforma 1) de milho com relação à atrazine foi de 6,6 nmol.min⁻¹.mg⁻¹ de proteína enquanto a mesma apresentou baixa atividade para outros herbicidas comparados com as demais GSTs de milho. Além disso, as formas heterodiméricas da GSTF1-3 (29/26 kDa) e GSTF1-2 (29/27 kDa) apresentaram atividades específicas intermediárias em relação ao homodímero, os valores foram de 4,2 nmol.min⁻¹.mg⁻¹ de proteína para cada heterodímero (DIXON; COLE; EDWARDS, 1997). Andreou e Clonis (2002) investigaram o

potencial da GSTF1 para o uso na química analítica como um biosensor para a detecção do herbicida atrazine. Os autores concluíram que o método foi bem sucedido, pois esta isoforma apresentou elevada especificidade pela atrazine e não houve interferência com outros herbicidas como o alachlor ou o carbamil, mesmo na ausência da atrazine.

Tabela 10 - Ordem de grandeza da atividade das classes de GST para o herbicida atrazine em diferentes espécies, em ordem decrescente

GST	Organismo	Classe	Atividade Atrazine (nmol.min ⁻¹ .mg ⁻¹ de proteína)	Referências
<i>MmGSTP1-1</i>	Camundongo	Pi	7 x 10 ⁰	ABEL et al., 2004
<i>HsGSTP1-1</i>	Humano	Pi	7 x 10 ⁰	ABEL et al., 2004
<i>ZmGSTF1-1</i>	Milho	Phi	7 x 10 ⁰	DIXON; COLE; EDWARDS, 1997
<i>ZmGSTF1-2</i>	Milho	Phi Heterodimero	4 x 10 ⁰	DIXON; COLE; EDWARDS, 1997
<i>ZmGSTF1-3</i>	Milho	Phi Heterodimero	4 x 10 ⁰	DIXON; COLE; EDWARDS, 1997
<i>HsGSTM1-1</i>	Humano	Mu	2 x 10 ⁻¹	ABEL et al., 2004
<i>HsGSTA1-1</i>	Humano	Alpha	9 x 10 ⁻²	ABEL et al., 2004
<i>HsGSTT1-1</i>	Humano	Theta	4 x 10 ⁻²	ABEL et al., 2004

As classes Alpha, Mu e Pi da enzima GST são capazes de conjugar o EA, no entanto a classe Pi é a que apresenta maior afinidade (PLOEMEN; VAN OMMEN; VAN BLADEREN, 1990). O EA é considerado o substrato representativo da classe Pi e é indicado para o tratamento de pressão alta em humanos (MANNERVIK; DANIELSON, 1988; OAKLEY et al., 1997). Nuccetelli et al. (1998) determinaram que uma mutação na GST classe Alpha de humano foi capaz de converter esta enzima para a classe Pi. Esta simples mutação de uma Tyr na posição 108 para uma Val alterou a especificidade da enzima. A atividade específica para o EA na *HsGSTP1-1* foi de 1000 nmol.min⁻¹.mg⁻¹ de proteína (Tabela 11) e na *HsGSTA1-1* foi de 100 nmol.min⁻¹.mg⁻¹ de proteína (NUCCETELLI et al., 1998). Harwaldt, Rahlfs e Becker (2002) caracterizaram a GST do parasita da malária o *Plasmodium falciparum* (*PfGST*), a qual apresentou uma estrutura semelhante à GST classe Mu, embora o alinhamento de seqüências tenha se agrupado juntamente com a classe Alpha. A atividade específica desta enzima pelo EA foi de 190 nmol.min⁻¹.mg⁻¹ de

proteína (HARWALDT; RAHLFS; BECKER, 2002). Micaloni e colaboradores (2003), construíram a região C-terminal da *HsGSTP1* a qual apresenta contato com o sítio-H, de modo a alterar as propriedades catalíticas da enzima. Esta enzima teve sua especificidade alterada da classe Pi para a classe Alpha. A atividade original da *HsGSTP1* foi de $1000 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ de proteína, enquanto a atividade da *HsGSTA1* foi de $100 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ de proteína e a Pi-químera mostrou uma atividade de $200 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ de proteína. Nóvoa-Valiñas e colaboradores (2004) estudaram as GSTs de lampréia (*Petromyzon marinus*) e determinaram que a isoenzima relacionada com a classe Pi apresentou uma atividade específica de $1250 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ de proteína. Em 2004, Yang et al. isolaram a GST do hepatopâncreas de *Mytilus edulis*, o qual codifica para uma proteína de 206 aminoácidos com uma massa molecular de 23,7 kDa. Esta proteína apresentou elevada identidade de sequência com a GST classe Pi e ainda apresentou elevada atividade para o ácido etacrínico (EA) e o 1-cloro-2,4-dinitrobenzeno (CDNB).

Em milho, a classe Phi isoforma 1 (GSTF1) mostrou uma atividade específica pelo EA de $1620 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ de proteína (DIXON; COLE; EDWARDS, 1997). A atividade da classe Tau em soja foi de $109 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ de proteína (*GmGST4*) e em milho de $117 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ de proteína (*ZmGST19*) (McGONIGLE et al., 2000). Em soja, Gronwald e Plaisance (1998), compararam as sequências da região N-terminal das GSTs cujas subunidades foram denominadas por A1, B1 e B2. Os resultados indicaram que há um elevado nível de similaridade com a *ZmGSTF1*. A GST A1/A1, um homodímero expresso constitutivamente, com uma massa molecular de 26 kDa, apresentou uma atividade específica pelo EA de $2340 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ de proteína. Numa menor extensão, o heterodímero contendo as subunidades B1 e B2 (GST B1/B2) também exibiram atividade pelo EA (GRONWALD; PLAISANCE, 1998).

Tabela 11 - Ordem de grandeza da atividade decrescente das classes de GST para o ácido etacrínico em diferentes espécies

GST	Organismo	Classe	Atividade ácido etacrínico (nmol/min/mg proteína)	Referências
<i>SbGSTF1-1</i>	Sorgo	Phi	2×10^3	GRONWALD; PLAISANCE, 1998
<i>ZmGSTF1-1</i>	Milho	Phi	2×10^3	DIXON; COLE; EDWARDS, 1997
<i>HsGSTP1-1</i>	Humano	Pi	1×10^3	NUCCETELLI et al., 1998 MICALONI et al., 2003
<i>PmGSTP1-1</i>	Lampréia	Pi	1×10^3	NÓVOA-VALIÑAS et al., 2004
<i>SbGSTF1-2</i>	Sorgo	Phi Heterodímero	4×10^2	GRONWALD; PLAISANCE, 1998
<i>PfGST</i>	Plasmodio	Alpha/ Mu	2×10^2	HARWALDT; RAHLFS; BECKER, 2002
<i>ZmGST19</i>	Milho	Tau	1×10^2	McGONIGLE et al., 2000
<i>GmGST4</i>	Soja	Tau	1×10^2	McGONIGLE et al., 2000
<i>HsGSTA1-1</i>	Humano	Alpha	1×10^2	NUCCETELLI et al., 1998 MICALONI et al., 2003

Deste modo, a classe Pi, encontrada em metazoários e a classe Phi, exclusiva das plantas, apesar da baixa identidade de seqüências entre as enzimas destas classes (22,3 %), apresentaram especificidades pelos mesmos substratos, com afinidades na mesma ordem de grandeza. A explicação para este fato pode ser compreendida analisando-se tanto as diferenças no sítio ativo das enzimas, como as interações destes resíduos com os substratos atrazine e EA, descrito no próximo item.

5.3.2 Diferenças estruturais entre as GSTs classe Phi e classe Pi

A análise das ligações de hidrogênio entre as classes Phi (pdb: 1axd) e Pi (pdb: 3gss) nos mostra que os principais resíduos envolvidos na interação com a glutatona, classe Phi (classe Pi), são: Asn13 (Arg13), His40 (Trp38), Lys41 (Lys44), Gln53 (Gln51), Val54 (Leu52), Glu66 (Gln64), Ser67 (Ser65), Arg68 (Asn66). Comparando-se os contatos realizados entre nestas estruturas, encontramos um padrão muito semelhante no sítio G, ou sítio de ligação da glutatona (Tabela 12). Isto era esperado com base na literatura (ARMSTRONG, 1997; PRADE; HUBER; BIESELER, 1998; SHEEHAN et al., 2001), pois a região N-terminal das proteínas da superfamília das GSTs é muito conservada e apresenta afinidade pelo mesmo substrato, a glutatona (Figura 10).

Tabela 12 - Comparação dos resíduos do sítio G que interagem com os íons da GSH, nas GSTF1 e GSTP1

Glutaciona	Íons da GSH	Resíduos da GSTF1	Resíduos da GSTP1
gamma-Glu	NH ₃ ⁺	Glu 66	Gln 64
		Asn 13	Arg 13
		Pro 55	Pro 53
	COO ⁻	Glu 66	Gln 64
		Ser 67	Ser 65
Cys	C=O	Asn 13	Gln 51
	NH	Gln 53	Leu 52
		Gln 53	--
	C=O	Val 54	Leu 52
		--	Gln 51
Gly	NH	His 40	Trp 38
		Lys 41	Lys 44
	COO ⁻	Gln 53	Gln 51

A GSTF1 apresentou os resíduos His40 e Lys41 fazendo contato com o grupo carboxilato da porção Gly da GSH e o Glu66 fez contato com o grupo gamma-glutamato da GSH (NEUEFEIND et al., 1997a; PRADE; HUBER; BIESELER, 1998). A Lys41 contribuiu diretamente para a formação do campo eletrostático do sítio ativo (AXARLI; RIGDEN;

LABROU, 2004; LABROU; MELLO; CLONIS, 2001b; PRADE; HUBER; BIESELER, 1998). Resultado semelhante foi obtido utilizando-se o método de docking, realizado nesta tese, o qual mostrou ligações de hidrogênio do grupo Gly da GSH com os resíduos His40 e Lys41. O grupo gamma-glutamil da GSH apresentou ligações de hidrogênio com o Glu66, Ser67 e Arg68.

Na GSTP1, não houve diferença maior que 0,1 nm entre a estrutura resolvida por cristalografia, complexada com o ligante glutatona e o resultado do re-docking da mesma estrutura com a GSH. Este resultado era esperado segundo o próprio autor do programa de docking (Ligin, SOBOLEV et al., 1996). Na GSTP1, o átomo de enxofre da GSH realiza uma ligação de hidrogênio com a hidroxila da Tyr7 e com uma molécula de água. Esta molécula de água pode servir como aceptora de elétrons do grupo NH da Arg13 (PRADE; HUBER; BIESELER, 1998; OAKLEY et al., 1997). O docking molecular do estado de transição da *HsGSTP1* e do conjugado atrazine-GSH mostrou que tanto uma molécula de água como a Arg13 devem assistir à saída do átomo de cloro da atrazine, o segundo passo na reação de substituição nucleofílica promovida pela GST (ABEL et al., 2004). Na GSTF1, a análise dos resíduos que participam da transferência de elétrons mostra que a Arg16 pode apresentar este mesmo papel na classe Phi.

O alinhamento estrutural das classes Pi e Phi mostra que o resíduo responsável pela catálise na classe Pi, a Tyr na posição 7, está alinhado com a Ser11 da classe Phi (Figura 8). Este fato está amplamente divulgado na literatura de modo a confirmar estes resíduos como responsáveis pela atividade catalítica da enzima GST (ARMSTRONG, 1997; LABROU; MELLO; CLONIS, 2001b; PRADE; HUBER; BIESELER, 1998; OAKLEY et al., 1997; SHEEHAN et al., 2001). A Tabela 13 apresenta um alinhamento que foi obtido primeiramente através do alinhamento estrutural das classes Pi (*HsGSTP1*) e Phi (*ZmGSTF1*). Segundo, foi efetuada uma observação no alinhamento estrutural e foram anotados os resíduos que efetivamente se encontraram na mesma posição espacial levando em consideração a posição da cadeia lateral dos mesmos. Deste modo, o alinhamento obtido nesta tabela não reflete necessariamente o alinhamento da sequência ou da estrutura. Observação que pode ser feita particularmente na Phe8 da classe Pi e a Phe35 da classe Phi e no Trp12 (classe Phi) e Asn204 (classe Pi), este último contato é mediado por uma molécula de água (Figura 9). Observação semelhante foi feita por Axarli; Rigden e Labrou (2004), os quais afirmam que embora o sítio H da GST classe Pi de humanos apresente um volume mais estreito e aprofundado, encontra similaridades com relação à natureza dos resíduos

que compõem o sítio H. A Phe8, por exemplo, é substituída por uma Ala na GSTF1, no entanto, o alinhamento estrutural mostra que a Phe35 ocupa a mesma posição. Da mesma maneira, a superfície hidrofóbica na GSTP1 consiste dos resíduos Tyr 108 e Gly 205, os quais são substituídos pelas superfícies hidrofóbicas das cadeias laterais da Phe 114 e Ile 118.

Tabela 13 - Posições correspondentes à observação do alinhamento estrutural das enzimas GSTF1 e GSTP1

	Resíduos (número correspondente a cada estrutura)											
<i>ZmGSTF1</i>	Met 10	Ser 11	Trp 12	Asn 13	Phe 35	Lys 41	Gln 53	Val 54	Pro 55	Glu 66	Asn 110	Ile 118
<i>HsGSTP1</i>	Val 10	Tyr 7	Asn 204*	Arg 13	Phe 8	Lys 44	Gln 51	Leu 52	Pro 53	Gln 64	Ile 104	Tyr 108

* interação mediada por uma molécula de água.

Sítio Hidrofóbico

No sítio H (H: hidrofóbico), ou sítio que determina a especificidade das diversas classes das GSTs (Tabela 14), os principais resíduos que compuseram a superfície de contato da classe Pi foram: Tyr7, Phe8, Val10, Ile104, Tyr108, Thr109, Gly205. Segundo Oakley e colaboradores (1997), o EA apresentou interações primordialmente com as cadeias laterais dos resíduos Tyr7, Phe8, Val10, Val35, Ile104, Tyr108 e Asn204 (mediado por uma molécula de água). O anel aromático do ligante apresentou uma importante interação com as cadeias laterais dos resíduos Tyr108 e Phe8 (OAKLEY et al., 1997). Resultado semelhante foi observado no docking molecular do estado de transição da GSTP1 e o conjugado atrazine-GSH (ABEL et al., 2004), no qual a atrazine estabeleceu interações aromáticas entre a Phe8 e a Tyr108, além de apresentar uma típica interação de van der Waals, dentro da acurácia do modelo de docking (3,9 Å), com a cadeia lateral da Ile104.

A classe Phi, por sua vez, apresentou os seguintes resíduos envolvidos na formação da superfície de contato com o ligante EA: Met10, Trp12, Asn13, Phe35, Phe114 e Ile118. Duas foram as ligações de hidrogênio observadas. Uma com o átomo de oxigênio da Ser11 (-OH) e a outra com o nitrogênio do Trp12 (Figura 11, B₁ e B₂). É importante salientar que embora o docking do EA na estrutura *HsGSTP1* tenha concordado com os contatos obtidos na própria estrutura resolvida, o mesmo não é diretamente esperado na classe Phi, pois o método de docking

utilizado aloca o ligante no sítio da enzima sem permitir mobilidade entre os átomos, seja através da modificação do comprimento da ligação entre dois átomos numa mesma molécula ou entre os ângulos diedrais da molécula. Este método analisa as várias posições que o ligante pode assumir, de acordo com a função de complementariedade (CF), sem que haja flexibilidade do mesmo. No entanto, segundo Sobolev e colaboradores (1996), embora a mobilidade seja esperada numa enzima quando em contato com um ligante, o modelo de docking rígido é capaz de nos fornecer os potenciais contatos da enzima com o ligante.

Tabela 14 - Comparação dos contatos do EA entre as GSTP1 (*HsGSTP1*) e GSTF1 (*ZmGSTF1*) (HB = interação de hidrogênio, Arom. = interação aromática, Hidrof. = interação de hidrofobicidade e DC = contato desestabilizador)

Classe Pi (cristalografia)				Classe Phi (docking + EM)			
Resíduo	Número	Distância (Å)	Tipo de Interação	Resíduo	Número	Distância (Å)	Tipo de Interação
Tyr	7	2,8	HB	Met	10	4,0	Hidrof.
Phe	8	3,1	Hidrof.	Ser	11	4,3	--
Val	10	2,9	Hidrof.	Trp	12	4,0	Arom.
Ile	104	3,3	Hidrof.	Asn	13	2,7	Hidrof.
Tyr	108	3,0	HB/Arom.	Phe	35	5,3	Arom.
Thr	109	3,8	Hidrof.	Asn	110	2,7	DC
Gly	205	4,2	--	Phe	114	4,1	Hidrof.
				Ile	118	3,3	Hidrof.
				Met	121	4,3	Hidrof.
				Leu	122	4,9	Hidrof.

Na *ZmGSTF1* conjugada com o Cibacron Blue 3GA (CB3GA), o conjunto de interações com a enzima envolveu os resíduos Met10, Trp12, Phe35, Phe114, Ile118 e Met121 (AXARLI; RIGDEN; LABROU, 2004). Ainda na *ZmGSTF1*, o complexo conjugado atrazine-GSH apresentou contatos com a enzima com os resíduos Met10 e Phe35 de um lado e o Trp12 e Ile118 no outro lado do conjugado (PRADE; HUBER; BIESELER, 1998). Segundo Prade e colaboradores (1998), em plantas, o “sanduíche” observado na classe Pi composto pela Phe8-atrazine-Tyr108, é parcialmente visto na classe Phi, na qual a porção da atrazine estabelece contatos entre a Phe35 e a Ile118. Em Axarli; Rigden e Labrou (2004), a *ZmGSTF1* complexada

com o conjugado atrazine-GSH, apresenta interação aromática paralela com a Phe35 enquanto a interação aromática com o Trp12 é perpendicular.

Elevada flexibilidade

A elevada flexibilidade encontrada nas classes Pi e Phi podem facilitar a conjugação da GSH por substratos similares nestas duas enzimas. Na classe Phi dois resíduos de Gly foram encontrados numa região de loop (Gly123 e Gly124), entre as hélices alpha 4 e alpha 5, o que pode favorecer a flexibilidade nas alpha-hélices destacadas. Prade e colaboradores (1998) encontraram valores altos do desvio RMS deste loop entre os quatro monômeros obtidos por cristalografia, demonstrando uma alta flexibilidade desta região. Por outro lado, a classe Pi apresenta uma região C-terminal estendida (resíduos 201 a 209), em relação à GST classe Phi, o que inclui uma Asn na posição 204. Este resíduo, por exemplo, realiza uma interação de hidrogênio com o ligante através de uma molécula de água (Figura 9A). Além disso, Micaloni e colaboradores (2003), incluíram uma alpha-hélice na região C-terminal da enzima, de forma que a classe Pi se assemelhasse com a classe Alpha. Deste modo, a flexibilidade encontrada na classe Pi foi reduzida e esta enzima-quimera começou a apresentar atividade da classe Alpha.

Convergência funcional propriamente dita

As GSTs classe Pi (reino Animalia) e Phi (reino Plantae) apresentaram uma baixa identidade de seqüências de proteínas, fruto da própria divergência evolutiva encontrada nos reinos em questão. No entanto, as estruturas tridimensionais destas enzimas permaneceram conservadas no decorrer da história. Por exemplo, o sítio H de ambas as classes apresentou resíduos capazes de permitir uma catálise eficiente com os substratos avaliados, tais como: a Phe8 e Tyr108 da classe Pi e a Phe35 e Ile118 da classe Phi. Além disso, as estruturas mostraram uma certa flexibilidade em suas respectivas cadeias, seja devido à presença de resíduos de Gly na classe Phi ou através de uma extensão na região C-terminal na classe Pi.

Para compreender como a convergência funcional pode ter ocorrido é necessário mencionar que a seleção natural ocorre de forma a privilegiar os organismos mais adaptados e capazes de gerar descendentes. Deste modo, os organismos cujo mecanismo de desintoxicação foi mais eficiente para moléculas semelhantes ao EA e à ATA foram privilegiados, os seja, sobreviveram mais tempo e deixaram mais descendentes. É possível que o substrato “original” para o qual estes genes foram se especializando, através da seleção natural, fosse o mesmo. No entanto, torna-se

difícil inferir sobre o papel fisiológico destas classes de GST durante o processo de seleção e evolução, mesmo porque este é um mecanismo dinâmico e as alterações climáticas e ambientais se alteram no decorrer dos tempos. Certamente as moléculas de ATA e EA não existiam nos primórdios da vida no planeta, mas provavelmente alguns compostos semelhantes eram formados, resultado do metabolismo de outros organismos ou do próprio processo de formação da vida na Terra. Deste modo, e por algum motivo especial, a natureza selecionou moléculas diferentes mas que fossem capazes de resolver o mesmo problema. Deste modo, as GSTs classes Pi e Phi compõem mais um exemplo de evolução convergente funcional, assim como os descritos por Ponting e Russell (2002) e Wu et al., (1999).

5.4 Modelagem molecular da GST de cana-de-açúcar

A modelagem molecular de proteínas é uma ferramenta fundamental para o desenho racional de moléculas, sejam estas pequenas moléculas ou enzimas. Dentre as técnicas mais comuns, estão envolvidas a modelagem por homologia (ROBINSON et al., 2004), seguida de docking molecular (ABEL et al., 2004; NEUMANN et al., 2004) ou dinâmica molecular (LABROU; KOTZIA; CLONIS, 2004; NEUMANN et al., 2004; SOUZA; ORNSTEIN, 1999; OROZCO et al., 1997; OYEDOTUN; LEMIRE, 2003; STELLA et al., 1999a; STELLA et al., 1999b; WONGTRAKUL; SRAMALA; KETTERMAN, 2003). Alguns autores mostram que é possível alterar a afinidade da enzima por determinado substrato (HEDEROS et al., 2004; KUMAR et al., 2002; NUCCETELLI et al., 1998), modificar o pKa da enzima (LABROU; MELLO; CLONIS, 2001a; LABROU; RIGDEN; CLONIS, 2004) e desestabilizar a mesma alterando por exemplo um único resíduo (WONGTRAKUL; SRAMALA; KETTERMAN, 2003). As técnicas utilizadas nesta tese permitiram especificar os resíduos que fazem parte do sítio catalítico da GST de cana-de-açúcar, além disso, as análises da dinâmica molecular sugerem que dois resíduos mutados poderiam afetar mais significativamente a afinidade da enzima pelo herbicida atrazine. Neste caso, é necessário comprovação, através de experimentação *in vitro* nos mutantes obtidos por mutação sítio-dirigida, seguido de expressão, purificação e avaliação da atividade da GST clonada em *Escherichia coli*. Esta parte do trabalho está sendo desenvolvida em parceria com o Centro de Tecnologia Canavieira - CTC, no entanto, os experimentos ainda estão em andamento e não houve tempo hábil para que os resultados fossem incluídos nesta tese.

6 CONCLUSÕES

Padrão das superfamílias de proteínas.

A relação entre a sequência, estrutura, função e afinidade descritas nesta tese não procuraram inovar os conhecimentos atuais sobre estes quatro elementos, mas, organizar as notações de modo a facilitar a explanação deste assunto. Além disso, procura mostrar que o paradigma estrutura-função deveria ser ampliado para incluir a sequência de aminoácidos e a afinidade da enzima. Já que controlar a atividade específica da enzima é o assunto que realmente se encontra na fronteira da ciência.

Evolução convergente entre as classes Pi e Phi.

As classes Phi e Pi de GST apresentaram apenas 22,3% de identidade de sequência, com o sítio H contendo as maiores diferenças, mas mantendo de modo geral, as características bioquímicas dos resíduos. Embora o sítio G tenha demonstrado maior semelhança, as análises mostraram que o mecanismo catalítico deve ser diferente devido principalmente ao processo de transferência de elétrons na catálise. A avaliação das GSTs nos mais diversos filos mostrou que as classes Phi e Pi devem ser genes parálogos. Isto significa que elas não compartilham um ancestral comum direto, ou seja, as análises sugerem que uma, ou mais, duplicações dos genes das classes Phi e Pi surgiram após o evento de diferenciação dos filos Plantae e Animalia. Como as demais classes de GST não apresentaram as mesmas afinidades pelos substratos analisados, é mais parcimonioso pensar que as classes Pi e Phi convergiram ao invés de todas as demais classes perderem esta habilidade de conjugar a atrazine e o EA. Por isso, as GSTs classe Pi e Phi podem ser consideradas como um exemplo de convergência evolutiva.

Desenho racional

Embora a imprescindível experimentação não tenha ocorrido a tempo de ser publicada nesta tese, a proposta e análise dos mutantes da GST de *Saccharum officinarum* procuraram mostrar que as ferramentas da bioinformática estrutural podem contribuir para uma maior compreensão dos processos bioquímicos que ocorrem nas células. Além disso, o objetivo desta metodologia consiste numa redução do tempo experimental, na busca por novas moléculas ou moléculas melhoradas.

REFERÊNCIAS

ABEL, E.L.; OPP, S.M.; VERLINDE, C.L.M.J.; BAMMLER, T.K.; EATON, D.L. Characterization of atrazine biotransformation by human and murine glutathione s-transferases. **Toxicological Sciences**, Orlando, v. 80, p. 230-238, 2004.

AMES, B.N.; PROFET, M.; GOLD, L.S. Nature's chemicals and synthetic chemicals: Comparative toxicology. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 87, p. 7782-7786, 1990.

ANDREOU, V.G.; CLONIS, Y.D. A portable fiber-optic pesticide biosensor based on immobilized cholinesterase and sol-gel entrapped bromocresol purple for in-field use. **Biosensors & bioelectronics**, Oxford, v. 17, n. 1-2, p. 61-9, Jan. 2002.

ARCA, P.; HARDISSON, C.; SUAREZ, J.E. Purification of a glutathione s-transferase that mediates fosfomycin resistance in bacteria. **Antimicrobial agents and chemotherapy**, Detroit, v. 34, p. 844-848, 1997.

ARTYMIUK, P.J.; POIRRETTE, A.R.; RICE, D.W.; WILLETT, P. A polymerase I palm in adenylyl cyclase? **Nature**, London, v. 388, p. 33-34, 1997.

ARMSTRONG, R.N. Structure, catalytic mechanism, and evolution of the glutathione transferases. **Chemical research in toxicology**, Washington, v. 10, p. 2-18, 1997.

AXARLI, I.A.; RIGDEN, D.J.; LABROU, N.E. Characterisation of the ligandin site of maize glutathione transferase I. **Biochemical Journal**, Colchester, v. 382, p. 885-893, Jun., 2004.

BAKER, D.; SALI, A. Protein structure prediction and structural genomics. **Science**, Washington, v. 294, n. 5540, p. 93-96, 2001.

BERMAN, H.M.; WESTBROOK, J.; FENG, Z.; GILLILAND, G.; BHAT, T.N.; WEISSIG, H.; SHINDYALOV, I.N.; BOURNE, P.E. The Protein Data Bank. **Nucleic Acids Research**, London, v. 28, p. 235-242, 2000. Disponível em: <<http://www.rcsb.org/pdb/>>. Acesso em: 10 outubro 2005.

BLOUIN, C.; BOUCHER, Y.; ROGER, A.J. Inferring functional constraints and divergence in protein families using 3D mapPing of phylogenetic information. **Nucleic Acids Research**, London, v. 31, n. 2, p. 790-797, 2003.

BOARD, P.G.; BAKER, R.T.; CHELVANAYAGAM, G.; JERMIIN, L.S. Zeta, a novel class of glutathione transferases in a range of species from plants to humans. **Biochemical Journal**, Colchester, v. 328, p. 929-35, 1997.

BOARD, P.G.; COGGAN, M.; CHELVANAYAGAM, G.; EASTEAL, S.; JERMIN, L.S.; SCHULTE, G.K.; DANLEY, D.E.; HOTH, L.R.; GRIFFOR, M.C.; KAMATH, A.V. *et al.* Identification, characterization and crystal structure of the omega class glutathione transferases. **Journal of Biological Chemistry**, Baltimore, v. 275, p. 24798-24806, 2000.

BOARD, P.G.; MANNERVIK, B. The contribution of the C-terminal sequence to the catalytic activity of GST2, a human alpha-class glutathione transferase. **Biochemical Journal**, Colchester, v. 275, n. 1, p. 171-174, 1991.

BOECKMANN, B.; BAIROCH, A.; APWEILER, R.; BLATTER, M.-C.; ESTREICHER, A.; GASTEIGER, E.; MARTIN, M.J.; MICHOD, K.; O'DONOVAN, C.; PHAN, I.; PILBOUT, S.; SCHNEIDER, M. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. **Nucleic Acids Research**, London, v. 31, p. 365-370, 2003.

BOOTH, J.; BOYLAND, E.; SIMS, P. An enzyme from rat liver catalyzing conjugation with glutathione. **Biochemical Journal**, Colchester, v. 79, p. 516-524, 1961.

BORK, P.; SANDER, C.; VALENCIA, A. Convergent evolution of similar enzymatic function on different protein folds: The hexokinase, ribokinase, and galactokinase families of sugar kinases. **Protein Science**, New York, v. 2, n. 1, p. 31-40, 1993.

COMBES, B.; STAKELUM, G.S. A liver enzyme that conjugates sulfobromophthalein sodium with glutathione. **Journal of Clinical Investigation**, New Haven, v. 40, p. 981-988, 1961.

COPLEY, S.D. Enzymes with extra talents: moonlighting functions and catalytic promiscuity. **Current Opinion in Chemical Biology**, Boulder, v. 7, p. 265-272, 2003.

CUMMINS, I.; COLE, D.J.; EDWARDS, R. Purification of multiple glutathione S-transferases involved in herbicide detoxification from wheat (*Triticum aestivum* L.) treated with the safener fenchlorazole-ethyl. **Pesticide Biochemistry and Physiology**, New York, v. 59, p. 35-49, 1997.

DANIELSON, U.; MANNERVIK, B. Kinetic independence of the subunits of cytosolic glutathione transferase from the rat. **Biochemical Journal**, Colchester, v. 231, p. 263-267, 1985.

DARWIN, C. **A origem das espécies**. Rio de Janeiro: Ediouro Publicações S.A., 2004. 517p.

DAVIES, J.; CASELEY, J.C. Herbicide safeners: a review. **Pesticide science**, London, v. 55, p. 1043-1058, 1999.

DING, Y.; ORTELLI, F.; ROSSITER, L.C.; HEMINGWAY, J.; RANSON, H. The *Anopheles gambiae* glutathione transferase supergene family: annotation, phylogeny and expression profiles. **BMC Genomics**, London, v. 4, n. 1, p.35, Aug. 2003.

DIRR, H.W.; REINEMER, P.; HUBER, R. X-ray crystal structures of cytosolic glutathione S-transferases. Implications for protein architecture, substrate recognition and catalytic function. **European Journal of Biochemistry**, Oxford, v. 220, p. 645-661, 1994.

DIXON, D.; COLE, D.J.; EDWARDS, R. Characterisation of multiple glutathione transferases containing the GST I subunit with activities toward herbicide substrates in maize (*Zea mays*). **Pesticide Science**, London, v. 50, p. 72-82, 1997.

DIXON, D.P.; DAVIS, B.G.; EDWARDS, R. Functional divergence in the glutathione transferase superfamily in plants. **Journal of Biological Chemistry**, Baltimore, v. 277, n. 34, p. 30859-30869, 2002.

DIXON, D.P.; LAPTHORN, A.; EDWARDS, R. Plant glutathione transferases. **Genome Biology**, London, v. 3, n. 3, p. 3004.1-3004.10, 2002.

DIXON, D.P.; MCEWEN, A.G.; LAPTHORN, A.J.; EDWARDS, R. Forced evolution of a herbicide detoxifying glutathione transferase. **Journal of Biological Chemistry**, Baltimore, v. 278, p. 23930-23935, 2003.

DODSON, G.; WLODAWER, A. Catalytic triads and their relatives. **Trends in Biochemical Sciences**, Amsterdam, v. 23, p. 347-352, 1998.

EDWARDS, R.; DIXON, D.P. The role of glutathione transferases in herbicide metabolism. In: Cobb, A.H., Kirkwood, R.C. (Ed.). **Herbicides and their mechanisms of action**. Sheffield: Sheffield Academic Press, 2000, p. 38-71.

EGAAS, E.; FALLS, J.G.; SVENDSEN, N.O.; RAMSTAD, H.; SKAARE, J.U.; DAUTERMAN, W.C. Strain and sex specific differences in the glutathione S-transferase class Pi in the mouse examined by gradient elution of the glutathione affinity matrix and reverse-phase high performance liquid chromatography. **Biochimica et Biophysica Acta**, Amsterdam, v. 1243, p. 256-264, 1995.

FAVALORO, B.; TAMBURRO, A.; ANGELUCCI, S.; DE LUCA, A.; MELINO, S.; DI ILIO, C.; ROTILIO, D. Molecular cloning, expression and site-directed mutagenesis of glutathione S-transferase from *Ochrobactrum anthropi*. **Biochemical Journal**, Colchester, v. 335, p. 573-579, 1998.

FELSENSTEIN, J. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. **Methods in Enzymology**, New York, v. 266, p.418-27, 1996.

FELSENSTEIN, J. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author: Department of Genetics, University of Washington, Seattle. 1993.

FITCH, W.M. Further improvements in the method of testing for evolutionary homology among proteins. **Journal of Molecular Biology**, London , v. 49, n. 1, p. 1-14, 1970.

FLETCHER, R.J.; BISHOP, B.E.; LEON, R.P.; SCLAFANI, R.A.; OGATA, C.M.; CHEN, X.S. The structure and function of MCM from archaeal *M. Thermoautotrophicum*. **Nature structural biology**, New York, v. 10, n. 3, p. 160-167, Mar. 2003.

FREAR, D.S.; SWANSON, H.R. Biosynthesis of S-(4-ethylamino-6-isopropylamino-2-s-triazino) glutathine partial purification and properties of glutathione S-transferase from corn. **Phytochemistry**, Oxford, v. 9, p. 2123-2132, 1970.

FROY, O. Convergent Evolution of intertebrate defensins and nematode antibacterial factors. **Trends in Microbiology**, Cambridge, v. 13, n. 7, p. 314-319, Jul. 2005.

FROVA, C. The plant glutathione transferase gene family: genomic structure, functions, expression and evolution. **Physiologia Plantarum**, Lund, v. 119, p. 469-479, 2003.

GERLT, J.A.; BABBITT, P.C. Can sequence determine function? **Genome biology**, London, v. 1, n. 5, p. reviews0005, Nov. 2000.

GRIFFITHS, C.S. The correlation of protein structure and evolution of a protein-coding gene: phylogenetic inference using cytochrome oxidase III. **Molecular Biology and Evolution**, Chicago, v. 15, n. 10, p. 1337-1345, 1998.

GRONWALD, J.W.; PLAISANCE, K.L. Isolation and characterization of glutathione S-transferase isozymes from sorghum. **Plant Physiology**, Rockville, v. 117, n. 3, p. 877-92, Jul. 1998.

GUEX, N.; DIEMAND, A.; PEITSCH, M.C. Protein modelling for all. **Trends in Biochemical Sciences**, Amsterdam, v. 24, p. 364-367, 1999.

GUEX, N.; PEITSCH, M.C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. **Electrophoresis**, Wheiwheim, v. 18, p. 2714-2723, 1997.

HARWALDT, P.; RAHLFS, S.; BECKER, K. Glutathione S-transferase of the malarial parasite *Plasmodium falciparum*: characterization of a potential drug target. **Biological Chemistry**, Berlin, v. 383, p. 821-830, May. 2002.

HAYES, J.D. Glutathione transferases. **Annual Review of Pharmacology and Toxicology**, Palo Alto, v. 45, p. 51-88, 2005.

HEDEROS, S.; BROO, K.S.; JAKOBSSON, E.; KLEYWEGT, G.J.; MANNERVIK, B.; BALTZER, L. Incorporation of a single His residue by rational design enables thiol-ester hydrolysis by human glutathione transferase A1-1. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 101, n. 36, p. 13163-13167, Sep. 2004.

HENIKOFF, S.; GREENE, E.A.; PIETROKOVSKI, S.; BORK, P.; ATTWOOD, T.K.; HOOD, L. Gene families: The taxonomy of protein paralogs and chimeras. **Science**, Washington, v. 278, p. 609-614, 1997.

HOOD, L.; CAMPBELL, J.H.; ELGIN, S.C. The organization, expression, and evolution of antibody genes and other multigene families. **Annual review of genetics**, Palo Alto, v. 9, p. 305-353, 1975.

HOLM, L.; SANDER, C. Dali: a network tool for protein structure comparison. **Trends in Biochemical Sciences**, Amsterdam, v. 20, p. 478-780, 1995.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. "VMD - Visual Molecular Dynamics", **Journal of Molecular Graphics**, Stoneham, v. 14, p. 33-38, 1996.

ISLAM, M.O.; HARA, M.; MIYAKE, J. Induction of p-glycoprotein, glutathione S-transferase and cytochrome P450 in rat liver by atrazine. **Environmental Toxicology and Pharmacology**, Amsterdam, v. 12, p. 1-6, 2002.

JAKOBSSON, P.J.; MORGENSTERN, R.; MANCINI, J.; FORD-HUTCHINSON, A.; PERSSON, B. Common structural features of MAPEG-a widespread superfamily of membrane associated proteins with highly divergent functions in eicosanoid and glutathione metabolism. **Protein Science**, New York, v. 8, p. 689-692, 1999.

JEPSON, I.; LAY, V.J.; HOLT, D.C.; BRIGHT, S.W.J.; GREENLAND, A.J. Cloning and characterization of maize herbicide safener-induced cDNAs encoding subunits of glutathione S-transferase isoforms I, II and IV. **Plant Molecular Biology**, Dordrecht, v. 26, n. 6, p. 1855-1866, 1994.

JI, X.; VON ROSENVINGE, E.C.; JOHNSON, W.W.; TOMAREV, S.I.; PIATIGORSKY, J.; ARMSTRONG, R.N.; GILLILAND, G.L. 3-Dimensional structure, catalytic properties, and evolution of a sigma-class glutathione transferase from squid, a progenitor of the lens S-crystallins of cephalopods. **Biochemistry**, Washington, v. 34, p. 5317-5328, 1995.

JI, X.; ZHANG, P.; ARMSTRONG, R.N.; GILLILAND, G.L. The three-dimensional structure of a glutathione S-transferase from the mu class gene. Structural analysis of the binary complex of isoenzyme 3-3 and glutathione at 2.2 Å resolution. **Biochemistry**, Washington, v. 31, p. 10169-10184, 1992.

JOHNSON, K.A.; ANGELUCCI, F.; BELLELLI, A.; HERVE, M.; FONTAINE, J.; TSERNOGLOU, D.; CAPRON, A.; TROTTEIN, F.; BRUNORI, M. Crystal structure of the 28 kDa glutathione S-transferase from *Schistosoma haematobium*. **Biochemistry**, Washington, v. 42, n. 34, p. 10084-94, Sep. 2003.

JOHNSON, M.S.; MAY, A.C.; RODIONOV, M.A.; OVERINGTON, J.P. Discrimination of common protein folds: application of protein structure to sequence/structure comparisons. **Methods in Enzymology**, New York, v. 266, p. 575-98, 1996.

JOHNSON, M.S.; SALI, A.; BLUNDELL, T.L. Phylogenetic relationships from three-dimensional protein structures. **Methods in Enzymology**, New York, v. 183, p. 670-90, 1990.

KAMPRANIS, S.C.; DAMIANOVA, R.; ATALLAH, M.; TOBY, G.; KONDI, G.; TSICHLIS, P.N.; MAKRIS, A.M. A novel plant glutathione s-transferase/peroxidase suppresses bax lethality in yeast. **Journal of Biological Chemistry**, Baltimore, v. 275, p. 29207-29216, 2000.

KANZ, C.; ALDEBERT, P.; ALTHORPE, N.; BAKER, W.; BALDWIN, A.; BATES, K.; BROWNE, P.; BROEK, A.; CASTRO, M.; COCHRANE, G.; DUGGAN, K.; EBERHARDT, R.; FARUQUE, N.; GAMBLE, J.; DIEZ, J.G.; HARTE, N.; KULIKOVA, T.; LIN, Q.; LOMBARD, V.; LOPEZ, R.; MANCUSO, R.; MCHALE, M.; NARDONE, F.; SILVENTOINEN, F.; SOBHANY, S.; STOEHR, P.; TULI, M.A.; TZOUVARA, K.; VAUGHAN, R.; WU, D.; ZHU, W.; APWEILER, R. The EMBL Nucleotide Sequence Database. **Nucleic Acids Research**, London, v. 33, p. D29-D33, 2005. Disponível em: <<http://www.ebi.ac.uk/embl>>. Acesso em: 20 outubro 2005.

KETTERMAN, A.J.; PROMMEENATE, P.; BOONCHAUY, C.; CHANAMA, U.; LEETACHEWA, S.; PROMTET, N.; PRAPANTHADARA, L. Single amino acid changes outside the active site significantly affect activity of glutathione S-transferases. **Insect Biochemistry And Molecular Biology**, Oxford, v. 31, n. 1, p. 65-74, 2001.

KOEPPE, M.K.; BAREFOOT, A.C.; COTTERMAN, C.D.; ZIMMERMAN, W.T.; LEEP, D.C. Basis of selectivity of the herbicide flupyrsulfuron-methyl in wheat. **Pesticide Biochemistry and Physiology**, New York, v. 59, p. 105-117, 1997.

KOONIN, E.V.; MUSHEGIAN, A.R.; TATUSOV, R.L.; ALTSCHUL, S.F.; BRYANT, S.H.; BORK, P.; VALENCIA, A. eukaryotic translation elongation factor 1-gamma contains a glutathione transferase domain - Study of a diverse, ancient protein superfamily using motif search and structural modeling. **Protein Science**, New York, v. 3, p. 2045-2054, 1994.

KOONIN, E.V.; WOLF, Y.I.; KAREV, G.P. The structure of the protein universe and genome evolution. **Nature**, London, v. 420, n. 6912, p. 218-23, 2002.

KUMAR, M.B.; FUJIMOTO, T.; POTTER, D.W.; DENG, Q.; PALLI, S.R. A single point mutation in ecdysone receptor leads to increased ligand specificity: Implications for gene switch applications. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 99, n. 23, p. 14710-14715, Nov. 2002.

LABROU, N.E.; KOTZIA, G.A.; CLONIS, Y.D. Engineering the xenobiotic substrate specificity of maize glutathione s-transferase I. **Protein Engineering, Design and Selection**, Oxford, v. 17, n. 10, p. 741-748, 2004.

LABROU, N.E.; MELLO, L.V.; CLONIS, Y.D. Functional and structural roles of the glutathione-binding residues in maize (*Zea mays*) glutathione S-transferase I. **Biochemical Journal**, Colchester, v. 358, p. 101-110, 2001a.

LABROU, N.E.; MELLO, L.V.; CLONIS, Y.D. The conserved Asn49 of maize glutathione S-transferase I modulates substrate binding, catalysis and intersubunit communication. **European Journal of Biochemistry**, Oxford, v. 268, p. 3950-3957, 2001b.

LABROU, N.E.; RIGDEN, D.J.; CLONIS, Y.D. Engineering the pH-dependence of kinetic parameters of maize glutathione s-transferase I by site-directed mutagenesis. **Biomolecular Engineering**, Amsterdam, v. 21, p. 61-66, 2004.

LESK, A.M. **Introduction to Bioinformatics**. New York: Oxford University Press, 2002. 283 p.

LI, W-H. 1997. Evolution by gene duplication and domain shuffling. In: _____. **Molecular Evolution**. Massachusetts: Sinauer Associates, Inc., 1997. chap. 10, p. 269-308.

LIO, P.; GOLDMAN, N.; THORNE, J.L.; JONES, D.T. PASSML: Combining evolutionary inference and protein secondary structure prediction. **Bioinformatics**, Oxford, v. 14, p. 726-733, 1998.

Lo BELLO, M.; NUCCETELLI, M.; CACCURI, A.M.; STELLA, L.; PARKER, M.W.; ROSSJOHN, J.; MCKINSTRY, W.J.; MOZZI, A.F.; FEDERICI, G.; POLIZIO, F.; PEDERSEN, J.Z.; RICCI, G. Human glutathione transferase P1-1 and nitric oxide carriers; a new role for an old enzyme. **Journal of Biological Chemistry**, Baltimore, v. 276, n. 45, p. 42138-45, 2001.

LOYALL, L.; UCHIDA, K.; BRAUN, S.; FURUYA, M.; FROHNMEYER, H. Glutathione and a UV light-induced glutathione S-transferase are involved in signaling to chalcone synthase in cell cultures. **The Plant Cell**, Baltimore, v. 12, p. 1939-1950, 2000.

MAKAROVA, K.S.; GRISHIN, N.V. Thermolysin and mitochondrial processing peptidase: how far structure-functional convergence goes. **Protein Science**, New York, v. 8, p. 2537-2540, 1999.

MANNERVIK, B.; DANIELSON, U.H. Glutathione transferases-structure and catalytic activity. **Critical Reviews in Biochemistry**, Boca Raton, v. 23, n. 3, p. 283-337, 1988.

MARGULIS, L.; SCHWARTZ, K.V. **Cinco Reinos: Um guia ilustrado dos filós da vida na Terra**. 3ed. Rio de Janeiro: Guanabara Koogan, 2001. 497 p.

MARTÍ-RENOM, M.A.; STUART, A.C.; FISER, A.; SÁNCHEZ, R.; MELO, F.; SALI, A. Comparative protein structure modeling of genes and genomes. **Annual Review of Biophysics and Biomolecular Structure**, Palo Alto, v. 29, p. 291-325, 2000.

McGOLDRICK, S.; O'SULLIVAN, S.M.; SHEEHAN, D. Glutathione transferase-like proteins encoded in genomes of yeasts and fungi: insights into evolution of a multifunctional protein superfamily. **FEMS Microbiology Letters**, Amsterdam, v. 242, n. 1, p. 1-12, Jan. 2005.

McGONIGLE, B.; KEELER, S.J.; LAN, S-M.C.; KOEPPE, M.K.; O'KEEFE, D.P. A genomics approach to the comprehensive analysis of the glutathione S-transferase gene family in soybean and maize. **Plant Physiology**, Rockville, v. 124, p. 1105-1120, 2000.

McLELLAN, L.I.; WOLF, C.R. Glutathione and glutathione-dependent enzymes in cancer drug resistance. **Drug Resistance Updates : Reviews and Commentaries in Antimicrobial and Anticancer Chemotherapy**, Edinburgh, v. 2, n. 3, p. 153-164, Jun. 1999.

MEISTER, A.; ANDERSON, M.E. Glutathione. **Annual Review of Biochemistry**, Palo Alto, v. 52, p. 711-60, 1983.

MICALONI, C.; KONG, G.K-W.; MAZZETTI, A.P.; NUCCETELLI, M.; ANTONINI, G.; STELLA, L.; McKINSTRY, W.J.; POLEKHINA, G.; ROSSJOHN, J.; FEDERICI, G.; RICCI, G.; PARKER, M.W.; Lo BELLO, M. Engineering a new C-terminal tail in the H-site of human glutathione transferase P1-1: Structural and functional consequences. **Journal of Molecular Biology**, London, v. 325, p. 111-122, 2003.

MIRNY, L.A.; GELFAND, M.S. Using orthologous and paralogous proteins to identify specificity-determining residues in bacterial transcription factors. **Journal of Molecular Biology**, London, v. 321, p. 7-20, 2002.

MORAITAKIS, G.; PURKISS, A.G.; GOODFELLOW, J.M. Simulated dynamics and biological macromolecules. **Reports on Progress in Physics**, Bristol, v. 66, p. 383-406, 2003.

MOREL, F.; ROUCH, C.; PETIT, E.; PITON, A.; THERET, N.; COLES, B.; GUILLOUZO, A. Gene and protein characterization of the human glutathione S-transferase Kappa and evidence for a peroxisomal localization. **Journal of Biological Chemistry**, Baltimore, v. 279, n. 16, p. 16246-16253, Apr. 2004.

MUELLER, L.A.; GOODMAN, C.D.; SILADY, R.A.; WALBOT, V. AN9, a petunia glutathione S-transferase required for anthocyanin. **Plant Physiology**, Rockville, v. 123, p. 1561-1570, 2000.

MURZIN, A.G.; BRENNER, S.E.; HUBBARD, T.; CHOTHIA, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. **Journal of Molecular Biology**, London, v. 247, p. 536-540, 1995.

NELSON, D.L.; LEHNINGER, A.L.; COX, M.M. **Lehninger Principles of Biochemistry**. 3 ed. New York: Worth Publishers, 2000. 1152 p.

NEUEFEIND, T.; HUBER, R.; DASENBROCK, H.; PRADE, L.; BIESELER, B. Crystal structure of herbicide-detoxifying maize glutathione S-transferase-I in complex with lactoylglutathione: evidence for an induced-fit mechanism. **Journal of Molecular Biology**, London, v. 274, n. 4, p. 446-53, 1997a.

NEUEFEIND, T.; HUBER, R.; REINEMER, P.; KNABLEIN, J.; PRADE, L.; MANN, K.; BIESELER, B. Cloning, sequencing, crystallization and X-ray structure of glutathione S-transferase-III from *Zea mays* var. mutin: a leading enzyme in detoxification of maize herbicides. **Journal of Molecular Biology**, London, v. 274, n. 4, p. 577-87, 1997b.

NEUMANN, D.; LEHR, C-M.; LENHOF, H-P.; LOHLBACHER, O. Computational modeling of the sugar-lectin interaction. **Advanced Drug Delivery Reviews**, Amsterdam, v. 56, p. 437-457, 2004.

NÓVOA-VALIÑAS, M.C.; PÉREZ-LOPEZ, M.; MELGAR-RIOL, M.J. Hepatic glutathione s-transferases from lamprey (*Petromyzom marinus*): purification and characterization. **Biochemical Systematics and Ecology**, Elmsford, v. 32, p. 169-178, 2004.

NUCCETELLI, M.; MAZZETTI, A.P.; ROSSJOHN, J.; PARKER, M.W.; BOARD, P.; CACCURI, A.M.; FEDERICI, G.; RICCI, G.; Lo BELLO, M. Shifting substrate specificity of human glutathione transferase (from class Pi do class Alpha) by a single point mutation. **Biochemical and Biophysical Research Communications**, New York, v. 252, p. 184-189, 1998.

OAKLEY, A.J.; ROSSJOHN, J.; Lo BELLO, M.; CACCURI, A.M.; FEDERICI, G.; PARKER, M.W. The three-dimensional structure of the human Pi class glutathione transferase P1-1 in complex with the inhibitor ethacrynic acid and its glutathione conjugate. **Biochemistry**, Washington, v. 36, p. 576-585, 1997.

O'BRIEN, P.J.; HERSCHLAG, D. Catalytic promiscuity and the evolution of new enzymatic activities. **Chemistry and Biology**, London, v. 6, n. 4, p. R91-R105, 1999.

O'CONNELL, K. M.; BREAU, E.J.; FRALEY, R.T. Different rates of metabolism of two chloroacetanilide herbicides in Pioneer 3320 corn. **Plant physiology**, Lancaster, v. 86, p. 359-363, 1988.

OROZCO, M.; VEGA, C.; PARRAGA, A.; GARCIA-SAEZ, I.; COLI, M.; WALSH, S.; MANTLE, T.J. On the reaction mechanism of class Pi glutathione S-transferase. **Proteins**, Hoboken, v. 28, n. 4, p. 530-542, Aug. 1997.

OYEDOTUN, K.; LEMIRE, B.D. The quaternary structure of the *Saccharomyces cerevisiae* succinate dehydrogenase: Homology modeling, cofactor docking, and molecular dynamics simulation studies. **Journal of Biological Chemistry**, Baltimore, v. 279, n. 10, p. 9424-9431, Dec. 2003.

PARRAGA, A.; GARCIA-SAEZ, I.; WALSH, S.B.; MANTLE, T.J.; COLL, M. The three-dimensional structure of a class-Pi glutathione S-transferase complexed with glutathione: the active-site hydration provides insights into the reaction mechanism. **Biochemical Journal**, Colchester, v. 333, n. 3, p. 811-6, 1998.

PATRICK, G.L. **An Introduction to Medicinal Chemistry**. 2ed. New York: Oxford University Press, 2001. 622p.

PLOEMEN, J.H.T.M.; VAN OMMEN, B.; VAN BLADEREN, P.J. Inhibition of rat and human glutathione S-transferase isoenzymes by ethacrynic acid and its glutathione conjugate. **Biochemical pharmacology**, Oxford, v. 40, p. 1631-1635, 1990.

POLEKHINA, G.; BOARD, P.G.; BLACKBURN, A.C.; PARKER, M.W. Crystal structure of maleylacetoacetate isomerase/glutathione transferase zeta reveals the molecular basis for its remarkable catalytic promiscuity. **Biochemistry**, Washington, v. 40, n. 6, p. 1567-76, Feb. 2001.

PONTING, C.P.; RUSSELL, R.R. The natural history of protein domains. **Annual Review of Biophysics and Biomolecular Structure**, Palo Alto, v. 31, p. 45-71, 2002.

PRADE, L.; HUBER, R.; BIESELER, B. Structures of herbicides in complex with their detoxifying enzyme glutathione S-transferase - explanations for the selectivity of the enzyme in plants. **Structure**, London, v. 6, n. 11, p. 1445-52, 1998.

RANSON, H.; PRAPANTHADARA, L.; HEMINGWAY, J. Cloning and characterization of two glutathione S-transferases from a DDT-resistant strain of *Anopheles gambiae*. **Biochemical Journal**, Colchester, v. 324, p. 97-102, 1997.

REINEMER, P.; PRADE, L.; HOF, P.; NEUEFEIND, T.; HUBER, R.; ZETTL, R.; PALME, K.; SCHELL, J.; KOELLN, I.; BARTUNIK, H.D.; BIESELER, B. Three-dimensional structure of glutathione S-transferase from *Arabidopsis thaliana* at 2.2 Å resolution: structural characterization of herbicide-conjugating plant glutathione S-transferases and a novel active site architecture. **Journal of Molecular Biology**, London, v. 255, n. 2, p. 289-309, 1996.

RENARD, M.; VARELA, P.F.; LETZELTER, C.; DUQUERROY, S.; REY, F.A.; HEIDMANN, T. Crystal structure of a pivotal domain of human syncytin-2, a 40 million years old endogenous retrovirus fusogenic envelope gene captured by primates. **Journal of Molecular Biology**, London, v. 352, p. 1029-1034, 2005.

RIECHERS, D.E.; IRZYK, G.P.; JONES, S.S.; FUERST, E.P. Partial characterization of glutathione S-transferases from wheat (*Triticum* spp.) and purification of a safener-induced glutathione S-transferase from *Triticum tauschii*. **Plant Physiology**, Rockville, v. 114, n. 4, p. 1461-70, 1997.

ROBINSON, A.; HUTTLEY, G.A.; BOOTH, H.S.; BOARD, P.G. Modelling and bioinformatics studies of the human Kappa-class glutathione transferase predict a novel third glutathione transferase family with similarity to prokaryotic 2-hydroxychromene-2-carboxylate isomerases. **Biochemical Journal**, Colchester, v. 379, p. 541-552, 2004.

RUSS, W.P.; LOWERY, D.M.; MISHRA, P.; YAFFE, M.B.; RANGANATHAN, R. Natural-like function in artificial WW domains. **Nature**, London, v. 437, p. 579-583, Sep. 2005.

RUSSELL, R.B. Detection of protein three-dimensional side-chain patterns: new examples of convergent evolution. **Journal of Molecular Biology**, London, v. 279, n. 5, p. 1211-1227, Jun. 1998.

SAITOU, N.; NEI, M. The neighbour-joining method: a new method for constructing phylogenetic tree. **Molecular Biology and Evolution**, Chicago, v. 4, p. 406-425, 1987.

SALI, A.; BLUNDELL, T.L. Comparative protein modelling by satisfaction of spatial restraints. **Journal of Molecular Biology**, London, v. 234, p. 779-815, 1993.

SALINAS, A.E.; WONG, M.G. Glutathione S-transferases – A review. **Current Medicinal Chemistry**, Schiphol, v. 6, p. 279-309, 1999.

SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. **Molecular cloning: A Laboratory Manual**. New York: Cold Spring Harbor Laboratory, 1989. chap. 13.

SAMUDRALA, R.; LEVITT, M. A comprehensive analysis of 40 blind protein structure predictions. **BMC Structural Biology**, London, v. 2, p. 3, Aug. 2002.

SCHUETTELKOPF, A.W.; VAN AALTEN, D.M.F. PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. **Acta crystallographica**, Section D, Biological crystallography, Copenhagen, v. D60, p. 1355-1363, 2004.

SCHULTZ, J.; MAISEL, S.; GERLACH, D.; MULLER, T.; WOLF, M. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throuyghout the Eukaryota. **RNA**, New York, v. 11, p. 361-364, 2005.

SCHWARTZMAN, D.W.; LINEWEAVER, C.H. The hyperthermophilic origin of life revisited. **Biochemical Society transactions**, London, v. 32, n. 2, p. 168-171, Apr. 2004.

SHEEHAN, D.; MEADE, G.; FOLEY, V.M.; DOWD, C.A. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. **Biochemical Journal**, Colchester, v. 360, n. 1, p. 1-16, 2001.

SHIMABUKURO, R.H.; SWANSON, H.R.; WALSH, W.C. Glutathione conjugation: atrazine detoxification mechanism in corn. **Plant Physiology**, Rockville, v. 46, p. 103-107, 1970.

SHINDYALOV, I.N.; BOURNE, P.E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. **Protein Engineering**, Oxford, v. 11, n. 9, p. 739-747, 1998.

SIDMAN, K.E.; GEORGE, D.G.; BARKER, W.C.; HUNT, L.T. The protein identification resource (PIR). **Nucleic Acids Research**, London, v. 16, p. 1869-71, 1988.

SINNING, I.; KLEYWEGT, G.J.; COWAN, S.W.; REINEMER, P.; DIRR, H.W.; HUBER, R.; GILLILAND, G.L.; ARMSTRONG, R.N.; JI, X.; BOARD, P.G. *et al.* Structure determination and refinement of human alpha class glutathione S-transferase A1-1, and a comparison with the mu and pi class enzymes. **Journal of Molecular Biology**, London, v. 232, p. 192-212, 1993.

SNYDER, M.J.; MADDISON, D.R. Molecular phylogeny of glutathione-S-transferases. **DNA Cell Biology**, New York, v. 16, n. 11, p. 1373-84, Nov. 1997.

SOBOLEV, V.; SOROKINE, A.; PRILUSKY, J.; ABOLA, E.E.; EDELMAN, M. Automated analysis of interatomic contacts in proteins. **Bioinformatics**, Oxford, v. 15, p. 327-332, 1999.

SOBOLEV, V.; WADE, R.C.; VRIEND, G.; EDELMAN, M. Molecular docking using surface complementarity. **Proteins: Structure, Function, and Genetics**, Hoboken, v. 25, p. 120-129, 1996.

SOCOLISH, M.; LOCKLESS, S.W.; RUSS, W.P.; LEE, H.; GARDNER, K.H.; RANGANATHAN, R. Evolutionary information for specifying a protein fold. **Nature**, London, v. 437, p. 512-518, Sep. 2005.

SOUZA, O.N.; ORNSTEIN, R.L. Molecular dynamics simulations of a protein-protein dimer: particle-mesh Ewald electrostatic model yields far superior results to standard cutoff model. **Journal of Biomolecular Structure & Dynamics**, Guilderland, v. 16, n. 6, p. 1205-1218, Jun. 1999.

SPOEL, D. van der; BUUREN, A.R. van; APOL, E.; MEULENHOF, P.J.; TIELEMAN, D.P.; SIJBERS, A.L.T.M.; HESS, B.; FEENSTRA, K.A.; LINDAHL, E.; DRUNEN, R. van; BERENDSEN, H.J.C. **Gromacs User Manual version 3.2**, 2004. Disponível em <www.gromacs.org>. Acesso em: 15 agosto 2005.

STELLA, L.; DI IORIO, E.E.; NICOTRA, M.; RICCI, G. Molecular dynamics simulations of human glutathione transferase P1-1: conformational fluctuations of the apo-structure. **Proteins**, Hoboken, v. 37, n. 1, p. 10-19, Oct. 1999a.

STELLA, L.; NICOTRA, M.; RICCI, G.; ROSATO, N.; DI IORIO, E.E. Molecular dynamics simulations of human glutathione transferase P1-1: analysis of the induced-fit mechanism by GSH binding. **Proteins**, Hoboken, v. 37, n. 1, p. 1-9, Oct. 1999b.

STRACHAN, T.; READ, A.P. **Human Molecular Genetics**, BIOS Scientific Pyblishes, Ltda, Segunda edição, 1999. Disponível em: <<http://www.nap.edu/books/0309069971/html>>. Acesso em: 15 jan. 2005.

TANG, A.H.; TU, C.P.D. Biochemical characterization of Drosophila glutathione S-transferases D1 and D2. **Journal of Biological Chemistry**, Baltimore, v. 269, p. 27876-84, 1994.

TATENO, Y.; IMANISHI, T.; MIYAZAKI, S.; FUKAMI-KOBAYASHI, K.; SAITOU, N.; SUGAWARA, H.; GOJOBORI, T. DNA Data Bank of Japan (DDBJ) for genome scale research in life science. **Nucleic Acids Research**, London, v. 30, n. 1, p. 27-30, 2002.

TATUSOV, R.L.; GALPERIN, M.Y.; NATALE, D.A.; KOONIN, E.V. The COG database: a tool for genome-scale analysis of protein functions and evolution. **Nucleic Acids Research**, London, v. 28, n. 1, p. 33-36, Jan. 2000.

TEW, K.D. Glutathione-associated enzymes in anticancer drug resistance. **Cancer Research**, Baltimore, v. 54, p. 4313-20, 1994.

THOM, R.; CUMMINS, I.; DIXON, D.P.; EDWARDS, R.; COLE, D.J.; LAPTHORN, A.J. Structure of a Tau class glutathione S-transferase from wheat active in herbicide detoxification. **Biochemistry**, Washington, v. 41, n. 22, p. 7008-20, 2002.

THOM, R.; DIXON, D.P.; EDWARDS, R.; COLE, D.J.; LAPTHORN, A.J. The structure of a Zeta class glutathione S-transferase from *Arabidopsis thaliana*: characterisation of a GST with novel active-site architecture and a putative role in tyrosine catabolism. **Journal of Molecular Biology**, London, v. 308, n. 5, p. 949-62, 2001.

THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUGIN, F.; HIGGINS, D.G. The CLUSTAL_X windows interface: flexible strategies for Multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Research**, London, v. 25, n. 24, p. 4876-82, 1997.

THORNE, J.L.; GOLDMAN, N.; JONES, D.T. Combining Protein Evolution and Secondary Structure. **Molecular Biology and Evolution**, Chicago, v. 13, n. 5, p. 666-673, 1996.

TOMAREV, S.I.; CHUNG, S.; PIATIGORSKY, J. Glutathione S-transferase and S-crystallins of cephalopods: Evolution from active enzyme to lens-refractive proteins. **Journal of Molecular Evolution**, New York, v. 41, p. 1048-1056, 1995.

UEDA, K.; MATSUO, M.; TANABE, K.; MORITA, K.; KIOTA, N.; AMACHI, T. Comparative aspects of the function and mechanism of SUR1 and MDR1 proteins. **Biochimica et Biophysica Acta**, Amsterdam, v. 1461, p. 305-313, 1999.

VEGA, M.C.; WALSH, S.B.; MANTLE, T.J.; COLL, M. The three-dimensional structure of Cys-47-modified mouse liver glutathione S-transferase P1-1. **Journal of Biological Chemistry**, Baltimore, v. 273, n. 5, p. 2844-50, 1998.

WACKETT, L.P. Evolution of enzymes for the metabolism of new chemical inputs in the environment. **Journal of Biological Chemistry**, Baltimore, v. 279, p. 41259-41262, 2004.

WHISSTOCK, J.C.; LESK, A.M. Prediction of protein function from protein sequence and structure. **Quarterly Reviews of Biophysics**, Cambridge, v. 36, n. 3, p. 307-340, 2003.

WILCE, M.C.J.; PARKER, M.W. Structure and function of glutathione S-transferases. **Biochimica et Biophysica Acta**, Amsterdam, v. 1205, p. 1-18, 1994.

WONGSANTICHON, J.; HARNNOI, T.; KETTERMAN, A.J. A sensitive core region in the structure of glutathione S-transferases. **Biochemical Journal**, Colchester, v. 373, p. 759-65, 2003.

WONGTRAKUL, J.; SRAMALA, I.; KETTERMAN, A.J. A non-active site residue, cysteine 69, of glutathione s-transferase ADGST3-3 has a role in stability and catalytic function. **Protein and Peptide Letters**, Schiphol, v. 10, n. 4, p. 375-385, 2003.

WOOD, T.C.; PEARSON, W.R. Evolution of protein sequences and structures. **Journal of Molecular Biology**, London, v. 291, n. 4, p. 977-95, Aug. 1999.

WU, G.; FISER, A.; ter KUILLE, B.; SALL, A.; MULLER, M. Convergent evolution of *Trichomonas vaginalis* lactate dehydrogenase from malate dehydrogenase. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 96, p. 6285-6290, 1999.

YANG, A.S.; HONIG, B. Sequence to structure alignment in comparative modeling using PrISM. **Proteins**, Hoboken, suppl. 3, p. 66-72, 1999.

YANG, H.L.; ZENG, Q.Y.; LI, E.Q.; ZHU, S.G.; ZHOU, X.W. Molecular cloning, expression and characterization of glutathione S-transferase from *Mytilus edulis*. **Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology**, Oxford, v. 139, n. 2, p. 175-182, 2004.

YANISCH-PERRON, C.; VIEIRA, J.; MESSING, J. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19. **Gene**, Amsterdam, v. 33, n. 1, p. 103-109, 1985.

YOON, S.; JONES, B.C.; LOGSDON, N.J.; WALTER, M.R. Same structure, different function: crystal structure of the epstein-barr virus IL-10 bound to the soluble IL-10R1 chain. **Structure**, London, v. 13, p. 551-564, Apr. 2005.

ZARRINE-AFSAR, A.; LARSON, S.M.; DAVIDSON, A.R. The family feud: do proteins with similar structures fold via the same pathway? **Current Opinion in Structural Biology**, London, v. 15, p. 42-49, Jan. 2005.