



República Federativa do Brasil  
Ministério da Economia  
Instituto Nacional da Propriedade Industrial

**(11) PI 1107433-7 B1**



**(22) Data do Depósito:** 30/12/2011

**(45) Data de Concessão:** 26/01/2021

---

**(54) Título:** DESENHO COMPUTACIONAL PARA NOVOS INIBIDORES DE ALFA-AMILASES

**(51) Int.Cl.:** G06F 19/16; G06F 19/18.

**(73) Titular(es):** EMBRAPA - EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA..

**(72) Inventor(es):** GORAN NESHICH; JOSÉ GILBERTO JARDINE; IZABELA AGOSTINHO PENA NESHICH; LETÍCIA SAYURI NISCHIMURA; IVAN MANZONI.

**(57) Resumo:** DESENHO COMPUTACIONAL PARA NOVOS INIBIDORES DE ALFA-AMILASES. A presente invenção se refere a um método para realizar o desenho computacional de novos compostos com potencial função inibitória das enzimas alfa-amilases. Um enfoque especial é dado ao desenho de novos compostos inibidores das enzimas alfa-amilases produzidas por insetos que se alimentam principalmente de produtos agrícolas estocados.

Relatório Descritivo de Patente de Invenção: “DESENHO COMPUTACIONAL PARA NOVOS INIBIDORES DE ALFA-AMILASES”.

### **CAMPO DA INVENÇÃO**

O presente documento de patente refere-se a um método para o desenho  
5 computacional de novos compostos de baixo peso molecular com potencial função  
inibitória das enzimas alfa-amilases. Um enfoque especial é dado ao desenho de novos  
compostos inibidores das enzimas alfa-amilases produzidas por insetos que se  
alimentam principalmente de produtos agrícolas estocados.

### **ESTADO DA TÉCNICA**

10 Para a maior parte das nações, os grãos estão entre os principais constituintes da  
base alimentar. No entanto, sua produção é sazonal e em muitos lugares as colheitas  
ocorrem apenas uma vez ao ano. Em consequência, tanto perdas em campo como em  
estoque são muito significativas.

Para que seja possível alimentar todas as pessoas do mundo é necessário que a  
15 maior parte da produção dos principais grãos (arroz, trigo, sorgo, milho, soja e painço)  
seja estocada por até mais que um ano. O mercado de grãos é caracterizado por  
demanda relativamente estável e abastecimento muito variável ao longo do ano. A  
variação no abastecimento é consequente a variação das colheitas, que depende das  
condições climáticas. Apesar de já existirem novas variedades de plantas que crescem  
20 mais rápido e são adaptáveis a diferentes regiões climáticas, ainda assim há variação do  
abastecimento. O armazenamento de grãos tem então o papel de diminuir essa variação  
do abastecimento, sendo feito nas épocas de colheita e liberando produto para o  
mercado nas épocas de baixa produção. Consequentemente, há também menos variação  
nos preços dos grãos durante o ano, o que contribui com a estabilidade econômica  
25 (Proctor, D.L; Grain storage techniques: evolution and trends in developing countries,  
FAO agricultural services bulletin nº 109, Rome, 1994).

Para controle e minimização das perdas de grãos estocados, já existem várias  
estratégias como desinfestação térmica, baixa umidade, baixo teor de oxigênio,  
resfriamento e armazenamento hermético, por exemplo, que são delineadas segundo o  
30 clima da região, tipo de grão e tipo de inseto. Ainda assim, há grandes perdas de  
produtos estocados causadas por pragas, que perfazem cerca de 10% da produção total,

e pelo menos 50% dessa perda deve-se a insetos. Tomando o trigo como exemplo, cuja produção foi de 668 milhões de toneladas entre 2009 e 2010, das quais 186 milhões de toneladas foram estocadas e considerando que o preço da tonelada era de US\$221,99 na época, perdas de 10% desse produto em estoque alcança um prejuízo em torno de  
5 US\$4,13 bilhões, sendo metade disso devido a insetos (Embrapa, Panorama mundial do trigo no fórum nacional, Embrapa notícias, 2010). Logo, se consideramos a crescente demanda anual de estoque de grãos em resposta à crescente produção agrícola e somarmos os prejuízos de todos os outros grãos, pode-se chegar a um valor exorbitante.

As pragas causadoras de tais perdas são roedores e uma grande variedade de  
10 insetos, sendo que estes últimos são mais difíceis de controlar por serem mais resistentes, se reproduzirem mais rápido e se infiltrarem com maior facilidade nos armazéns. Muitos dos insetos que se alimentam de grãos também o fazem no campo, infestando antes da colheita e devido a isso, há certa dificuldade em separar os insetos do produto colhido para ser armazenado. As pestes de estoque mais importantes  
15 conseguem sobreviver às estratégias de controle de pragas e multiplicam-se rapidamente mesmo em condições desfavoráveis para a maioria dos insetos. Por exemplo, grão seco para menos de 12% de umidade inibe o desenvolvimento da maioria das pestes, mas ainda assim, há insetos que conseguem se alimentar e reproduzir a menos de 4% de umidade. O mesmo ocorre com o baixo teor de oxigênio controlado nos estoques  
20 visando à morte e diminuição da proliferação de insetos, pois existem linhagens resistentes a esse método. A aplicação dessas estratégias também gera alto custo ao produtor e não é acessível à maioria dos produtores de países em desenvolvimento. Os insetos que conseguem sobreviver a essas estratégias têm cada vez maior importância econômica, pois não há como criar ambiente que seja ao mesmo tempo desfavorável ao  
25 seu desenvolvimento, não prejudique a qualidade do produto e seja de baixo custo. Os pesticidas apresentam uma alternativa mais barata, mas também representam risco de produção de efeitos indesejáveis no consumidor e desenvolvimento de resistência dos insetos. Apesar disso, o desenvolvimento de pesticidas mais específicos e que ataquem estrutura altamente conservada nos insetos pode trazer uma solução mais barata e segura  
30 para implementação do controle de pestes de estoque (Proctor, D.L; Grain storage techniques: evolution and trends in developing countries, FAO agricultural services bulletin nº 109, Rome, 1994).

Para muitos insetos que se alimentam de grãos durante suas fases larval e/ou adulta o amido representa a mais importante fonte energética. O amido é um polímero de resíduos de açúcar (polissacarídeo) que tem a função de reserva energética nos vegetais. É formado de polímero linear de glicose (amilose), cujos resíduos são ligados na conformação alfa-1,4, e de polímero ramificado de glicose (amilopectina), cujas cadeias lineares de glicose alfa-1,4 são intercaladas por ligações do tipo alfa-1,6. Para obtenção de energia a partir do amido, este deve ser processado ao longo do trato digestivo para produzir glicose, que é a molécula utilizada como fonte de energia para a produção de ATP, que é então usado como fonte de energia em quase todos os processos bioquímicos. As enzimas responsáveis pela conversão do amido a resíduos de glicose são as amilases, que podem ser classificadas:

- Endoamilases (EC 3.2.1.1): que clivam as ligações alfa-1,4 na amilose, amilopectina e polissacarídeos relacionados como o glicogênio. Os produtos de hidrólise são oligossacarídeos de vários comprimentos e com configuração alfa no C1 da unidade redutora. Agem liquefazendo o amido (efeito da rápida diminuição da viscosidade).
- Exoamilases (EC 3.2.1.2): clivam as ligações alfa-1,4 da amilose, amilopectina e glicogênio das extremidades não-redutoras através da remoção de maltose/glicose. Seus produtos têm a configuração beta no C1 da unidade redutora. Agem sacrificando o amido, resultando de lenta diminuição da viscosidade.
- Enzimas desramificadoras (3.2.1.3): as ligações alfa-1,6 são resistentes a ação das alfa e beta-amilases. A enzima pululanase quebra essas ligações; a glicoamilase também, porém mais lentamente que a pululanase.

As endoamilases são o tipo mais importante para obtenção de glicose devido a sua rápida atividade despolimerizadora (Muralikrishna, G., Nirmala, M., Cereal alpha-amylases – an overview, Carbohydrate polymers, 60, 163-173, 2005). São chamadas de alfa-amilases por quebrarem as ligações do tipo alfa-1,4 entre os resíduos de açúcar em componentes do amido, glicogênio e outros carboidratos relacionados e estão presentes em vários organismos, tais como mamíferos, insetos, bactérias e fungos. Pertencem ao clã GH-H (clã H de hidrolases glicosídicas) e são classificadas em famílias segundo

similaridade entre suas sequências primárias. O clã GH-H integra cerca de 100 famílias de enzimas e as alfa-amilases compreendem as famílias 13, 70 e 77; o conjunto dessas 3 famílias também é conhecido como superfamília de alfa-amilases. Acredita-se que todas as enzimas do clã GH-H possuem o sítio catalítico com a mesma forma tridimensional

5 de barril (alfa/beta) 8, que consiste de oito segmentos de folhas-beta e oito de alfa-hélices que se alternam ao longo da cadeia polipeptídica. Essa conformação foi observada nas cristalografias de raios X de enzimas das famílias 13, 77 e 70. Cada uma dessas famílias pode conter enzimas com mecanismos diferentes de catálise: a família 77 possui glucanotransferases que clivam as ligações alfa-1,4 entre resíduos de glicose e

10 transferem uma parte do substrato para um aceptor glucano, criando uma nova ligação alfa-1,4; a família 70 possui glucanosucrases que sintetizam poliglucanos a partir de sacarose podendo formar novas ligações alfa-1,2, alfa-1,3, alfa-1,4 ou alfa-1,6 entre os resíduos de glicose; e a família 13 contém hidrolases, transferases e isomerases com mais de vinte e duas especificidades conhecidas. Acredita-se que todas as enzimas do

15 clã GH-H realizam a reação catalítica por um mecanismo comum que envolve um resíduo de aspartato na folha-beta 4 (D185, numeração da estrutura do *Tenebrio molitor*, 1JAE.pdb), um glutamato na folha-beta 5 (E222, 1JAE.pdb) e um aspartato na folha-beta 7 (D287, 1JAE.pdb) (Figura 1). A catálise se inicia com a doação de um próton do resíduo E222 ao oxigênio glicosídico induzindo a quebra da ligação. Então, a

20 porção aglicona do substrato deixa o sítio ativo da enzima. O resíduo de glicose remanescente no sítio enzimático é estabilizado por interação de hidrogênio entre seus C2 e hidroxila do C3 com o aspartato da folha-beta 7. Adicionalmente, ocorre estabilização pela formação de ligação covalente entre o C1 da glicose e o aspartato nucleófilo da folha-beta 4. Por mecanismo desconhecido, a enzima facilita a

25 aproximação de um segundo substrato com reversão das interações anteriormente explicadas: a ligação covalente se quebra, o resíduo glutamato da folha-beta 5 é reprotonado por remoção de um próton do segundo substrato, ocorre inversão da configuração do C1 no intermediário transiente formado e um novo grupo redutor é formado no C1 do substrato; algumas enzimas transferem outro resíduo de glicose,

30 fosfato, frutose ou seus isômeros (MacGregor, E.A., An overview of clan GH-H and distantly-related families, *Biologia Bratislava*, 16:5-12, 2005). Um mecanismo diferente é proposto por Mazur e colaboradores, onde há a participação de molécula de água em

ataque nucleofílico ao C1 da amilose e, após a clivagem, há formação de hemiacetal a partir da interação entre os grupos aldeído e hidroxila do resíduo glicosídico não ligado à enzima, sem formação de intermediários covalentes entre enzima e substrato (Mazur, A.K, Haser, R. & Payan, F., The catalytic mechanism of alfa-amylases based upon  
 5 enzyme crystal structures and model building calculations, Biochem. Biophys. Res. Commun., 204, 297-302, 1994).

As estruturas das alfa-amilases geralmente tem três domínios:

1. A: contém os resíduos do sítio ativo e os envolvidos na ligação ao substrato.
2. B: é o domínio menos conservado entre as alfa-amilases. Liga-se ao cálcio com  
 10 interface ao domínio A. Possui resíduos envolvidos na ligação ao substrato.
3. C: sobrepõe-se parcialmente sobre a estrutura do domínio A e possui um segundo sítio ligante ao amido independente do sítio ligante que compreende os domínios A e B.

Segundo numeração da enzima alfa-amilase do besouro *Tenebrio molitor*, a qual  
 15 será referida como TMA daqui em diante, o domínio A compreende os intervalos de resíduos 1-97 e 160-379, o domínio B compreende o intervalo 98-159, e o domínio C compreende o intervalo 380-471 (Strobl, S., Maskos, K., Betz, M., Wiegand, G., Hubre, R., Gomis-Rüth, X., Glockshuber, R., Crystal structure of yellow meal worm alpha-amylase at 1.64 Å resolution, J. Mo. Bio, 278, 617-628, 1998).

20 Nas plantas, as alfa-amilases também possuem atividade associada à maturação e germinação de sementes participando na modificação do endosperma (Muralikrishna, G., Nirmala, M., Cereal alpha-amylases – an overview, Carbohydrate polymers, 60, 163-173, 2005). No entanto, esses processos não são afetados pela ação de inibidores de alfa-amilase produzidos pelas próprias plantas. A produção de inibidores nada mais é  
 25 que um mecanismo de defesa contra a ação de insetos, bactérias e fungos que também produzem alfa-amilase e a utilizam para quebrar o amido armazenado nas estruturas vegetais. Há dois tipos principais de inibidores produzidos pelas plantas:  $\alpha$ -AI1 e  $\alpha$ -AI2, sendo o  $\alpha$ -AI1 mais comum em sementes. A inibição das enzimas da própria planta por seus inibidores não ocorre, pois as sequências de aminoácidos próximos aos sítios

catalítico e ligante são bastante diferentes das sequências de insetos, culminando em diferenças estruturais que não permitem a ligação da enzima ao inibidor. Esses inibidores, no entanto, não são efetivos contra todos os tipos de alfa-amilases produzidas pelos insetos (Bompard-Gilles, C., Rousseau, P., Rougé, P., Payan, F.,  
 5 Substrate mimicry in the active Center of a mammalian alpha-amylase: structural analysis of an enzyme-inhibitor complex, *Structure*, 4, 1441-1452, 1996).

Além dos inibidores produzidos por plantas, há também inibidores de baixo peso molecular conhecidos. O inibidor mais conhecido é a acarbose, utilizado como hipoglicemiante oral no tratamento profilático e em casos iniciais de Diabetes Mellitus  
 10 tipo 2 (DM2). A acarbose inibe competitivamente e reversivelmente as alfa-amilases pancreática e intestinal impedindo a formação de glicose. Atualmente é comercializada como Precose® nos EUA e Glucobay® na Europa pela Bayer, que patenteou um processo de purificação com alta pureza (US Pat. 4904769) para produção de medicamentos. Nesse processo, a acarbose é obtida por fermentação de bactérias Gram  
 15 positivas do gênero *Actinoplanes* e purificada por técnica de HPLC ("High-pressure liquid chromatograph"). Há também os inibidores isoacarbose e acarviosina-glicose, que no estudo bioquímico de Kim e colaboradores (1999) mostraram ter maior poder de inibição do que a acarbose (Kim, M.J., Lee, S.B., Lee, H.S., Lee, S.Y., Baek, J.S., Kim, D., Moon, T.W., Rhyt, J.F., Park, K.H., Comparative study of the inhibition of alpha-  
 20 glucosidase, alpha-amylase, and cyclomaltodextrin glucanotransferase by acarbose, isoacarbose, and acarviosina-glucose, *Arch Biochem Biophys*, 15, 371(2):277-283, 1999).

Até o momento, a alfa-amilase da larva do *Tenebrio molitor* (TMA) é a mais bem estudada dentre as alfa-amilases de insetos. Sua estrutura foi cristalizada por Strobl  
 25 e colaboradores a 1,64 Å de resolução; código PDB (*Protein Data Bank*): 1JAE (Strobl, S., Maskos, K., Betz, M., Wiegand, G., Huber, R., Gomis-Rüth, F.X., Glockshuber, R., Crystal structure of yellow meal worm alpha-amylase at 1,64 Å resolution, *J Mol Biol*, 278, 617-628, 1998). O *T. molitor* é um besouro que está entre as pragas mais importantes e é causador de infestação de grãos e farináceos estocados. Suas larvas são  
 30 altamente resistentes e vivem até 80 dias a -5°C. Há também estudos de interação entre TMA e inibidores com informações importantes para o desenho racional de inibidores.

No trabalho de Pereira e colaboradores foi feita a cristalografia da estrutura do TMA em complexo com inibidor protéico produzido pelas sementes de *Amaranthus hypocondriacus*, 2,00 Å de resolução, sem grandes diferenças estruturais entre a forma livre e a complexada do TMA. Esse inibidor é específico contra alfa-amilases de insetos, sendo inativa para enzimas de mamíferos. Segundo o trabalho, isso se deve à menor interface de contato entre as proteínas, que faz com que haja menos interações de hidrofóbicas e de hidrogênio (Pereira, P.J.B., Lozanov, V., Patthy, A., Huber, R., Bode, W., Pongor, S., Strobl, S., Specific inhibition of insect alpha-amylases: yellow meal worm alpha-amylase in complex with Amaranth alpha-amylase inhibitor at 2,0 Å resolution, *Structure*, 7:1079-1088, 1999). Em outro estudo de Strobl e colaboradores a TMA foi cocrystalizada com inibidor protéico produzido pelo cereal *Eurecine coracana*, a 2,5 Å de resolução, sem diferenças estruturais entre a TMA livre e em complexo com esse inibidor. Neste trabalho também fica evidente a relevância das interações hidrofóbicas e de hidrogênio entre as proteínas (Strobl, S., Maskos, K., Wiegand, G., Huber, R., Gomis-Rüth, F.X., Glockshuber, R., A novel strategy for inhibition of alpha-amylases: yellow meal worm alpha-amylase in complex with the Ragi bifunctional inhibitor at 2,5 Å resolution, *Structure*, 6:911-921, 1998).

Com o intuito de se combater perdas de produtos agrícolas em estoque causadas por insetos, através do desenvolvimento e produção de estruturas com potencial para inibição dessas enzimas, ao longo do trabalho foram utilizadas técnicas computacionais para desenho de compostos (“computer-aided drug design” – CADD) (W. Graham Richards, Computer Aided-Drug Design, *Pure & Appl. Chem.*, 66,8: 1589-1596, 1994), através do desenho computacional de fármacos baseado na estrutura do ligante (“Ligand-Based Drug Design”-LBDD) (Acharya, C., Coop, A., Polli, E. J., MacKerrel, D., Alexander., Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach, *Current Computer – Aided Drug Design*, 7, 1:10-22, 2011), sendo o ligante a amilose. Avanços importantes na produção de fármacos já foram obtidos por meio de abordagens computacionais. O CADD pode prever resultados experimentais com razoável precisão em tempo reduzido se comparado com os métodos mais clássicos. Métodos de CADD já são amplamente utilizados pela indústria farmacêutica com a finalidade de identificar novos compostos ou aperfeiçoar compostos já existentes que apresentam atividade contra um

alvo biológico. Através de simulações de docking, que propõem predizer as interações proteína-ligante de interesse, levando em consideração parâmetros físico-químicos que interferem na ligação (D.B. Kitchen, H. Decornez, J.R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications, Nature Reviews – Drug Discovery, 3, 935-949, 2004), é possível identificar compostos que

5 potencialmente servirão como ligantes na proteína alvo.

O documento WO 2001097098 descreve um método para obtenção de moléculas com potencial inibitório a partir de bancos de dados contendo as estruturas dos compostos químicos (PubChem), potenciais efetores da função das enzimas cuja

10 atividade deve ser alterada (inibição ou potencialização) utilizando docking para selecionar entre as moléculas encontradas no PubChem, baseando-se em *scores* fornecidos pelos programas, tais como GOLD (Verdonk, M.L., Cole, J.C., Harston, M.J., Murray, C.W., Taylor, R.D., Improved protein-ligand docking using GOLD, Proteins, Structure, Function and Genetics, 52: 609-623, 2003) que simulam o

15 atracamento do ligante em sítio de ligação – alvo escolhido.

O documento RPI 2036 descreve o desenvolvimento de herbicidas contra uma enzima específica através da técnica de docking seguido de modificação molecular dos ligantes selecionados com melhores *scores*, considerando propriedades estéricas e eletrostáticas na interação ligante-proteína. Por fim, os resultados de docking foram

20 corroborados pelos resultados experimentais.

O documento US 2008/7383135 refere-se ao desenvolvimento de inibidores que agem em uma determinada região conservada entre as proteínas de uma família de quinases. O trabalho utiliza uma isoforma com estrutura tridimensional resolvida como molde para resolver a estrutura tridimensional de proteínas homólogas e para o desenho

25 de compostos com potencial para se ligarem à isoforma ou aos homólogos.

Ressalta-se a importância econômica dos organismos em questão e a dificuldade em se produzir vegetais totalmente resistentes a essas pragas é válido que se estabeleça um esforço no sentido de desenvolver compostos com o fim de preencher a necessidade de proteção eficaz contra as pragas. O desafio deste trabalho não reside apenas em

30 produzir moléculas com potencial inibitório das alfa-amilases de insetos, mas também

em produzir compostos de baixo risco para a população exposta e para o meio ambiente. Pensando nisso, propomos o desenho de novos compostos que atuariam como inibidores da enzima alfa-amilase, descrita como essencial para o correto desenvolvimento dos organismos que a contem.

## SUMÁRIO DA INVENÇÃO

5

A invenção refere-se a um método para o desenho computacional de novos compostos com potencial função inibitória das enzimas alfa-amilases. As alfa-amilases são enzimas presentes no trato digestivo de insetos, animais e humanos, e também produzidas por fungos e bactérias. Participam da hidrólise das ligações alfa-1,4 da amilose. A hidrólise da amilose resulta em resíduos de açúcar que são utilizados como fonte de energia pelos organismos. Com o objetivo de diminuir as perdas agrícolas estocadas causadas pela ação de insetos, pode-se desenhar estruturas que se ligam à alfa-amilase com alta afinidade teórica. Um enfoque especial, mas não limitante da invenção, é dado ao desenho de novos compostos inibidores das enzimas alfa-amilases produzidas por insetos que se alimentam principalmente de produtos agrícolas estocados. Dessa forma, os insetos cujas alfa-amilases são inibidas, não conseguindo obter os resíduos de açúcar necessários para obtenção de energia e, em consequência, morrem de privação energética. O desenho computacional tem como base metodologias de “Ligand-Based Drug Design” e as interações e estruturas do ligante natural (amilose) e de inibidores naturais. Também são utilizadas metodologias de “Structure-Based Drug Design”, onde as interações entre as moléculas desenhadas e a proteína alvo são avaliadas por meio de docking.

10

15

20

## BREVE DESCRIÇÃO DAS FIGURAS

**Figura 1.** Visão da tríade catalítica da alfa-amilase do *Tenebrio molitor* (TMA, 1JAE.pdb) com os resíduos de Aspartato (D185 e D287) e Glutamato (E222) destacados em preto. Imagem gerada pelo Molegro Virtual Docker.

25

**Figura 2.** Alinhamento das estruturas primárias de enzimas alfa-amilases. As sequências de insetos usadas neste alinhamento, com seus respectivos identificadores do NCBI GenPept e códigos, foram: *Tenebrio molitor*, 3891490, TMA; *Tribolium castaneum*, 167466185, TcA. As sequências de mamíferos: alfa-amilase salivar

30

humana, 215261156, HSA; alfa-amilase pancreática humana, 14719495, HPA; alfa-amilase pancreática suína, 21730556, PPA; alfa-amilase pancreática de camundongo, 111607467, MPA; alfa-amilase pancreática de ratazana, 149025954, RPA. E as sequências de plantas: *Physcomitrella patens ssp patens*, 168000555, PpA; *Glycine max*, 255638753, GmA. As sequências de alfa-amilases vegetais foram usadas com o intuito de verificar a conservação dos resíduos mais importantes envolvidos na catálise e a existência de outros resíduos que sejam exclusivos insetos e estejam espacialmente próximos ao sítio catalítico/ligante, o que poderia indicar a existência de alvos adicionais e preferenciais, pois um composto que se ligue fortemente a um resíduo específico insetos provavelmente não causará efeitos quanto à ligação à alfa-amilase de planta. Os resíduos ausentes em plantas, porém, não são exclusivos de insetos, estando presentes também em enzimas de mamíferos, cujas sequências possuem alta similaridade e identidade com as enzimas de insetos. Os resíduos mais importantes da tríade catalítica composta pelos Aspartatos 185 e 287, e pelo Glutamato 222 estão destacados por barras cinzas contornadas em preto. Os resíduos destacados por barras cinzas sem contorno (W56, W57, Q61, V151, I224, N331 e D332) são os resíduos que participam da ligação ao substrato e estão ausentes em vegetais. Em uma análise de homologia com 150 sequências, sendo 50 de insetos, 50 de mamíferos e 50 de plantas (dado não mostrado) não foram encontrados resíduos exclusivos de insetos que poderiam participar na interação com o substrato. Praticamente todos os resíduos importantes para ligação ao substrato existentes nas enzimas de insetos também estão presentes nas enzimas de mamíferos ou possuem um similar. Vale salientar que este alinhamento mostra de forma resumida a análise que foi feita entre várias sequências para facilitar a visualização da análise. Devido a isso, alguns resíduos não destacados que nessa figura parecem ser exclusivos de insetos, em análise mais profunda não o são. O resíduo G210, por exemplo, apesar de estar presente nas sequências de TMA e TcA, não é muito conservado em insetos e alguns mamíferos apresentam resíduos com características similares.

**Figuras 3A e 3B.** Alinhamento da estrutura da TMA (1JAE.pdb) (em preto) com estruturas que foram cocristalizadas com inibidor (em cinza), obtidas do Protein Data Bank. Na figura 3A, a TMA larval (1JAE.pdb) foi alinhada com outro cristal de TMA (1CLV.pdb) que possui inibidor protéico Amaranth complexado a sua estrutura. Na

figura **3B**, a TMA está alinhada com uma alfa-amilase pancreática suína (PPA, 1HX0.pdb) que foi cocrystalizada com inibidor de baixo peso molecular acarviosina-glicose. Os inibidores foram removidos das estruturas para melhor visualização. Os alinhamentos demonstram alta similaridade estrutural no que diz respeito às posições dos resíduos do sítio ligante (W56, W57, Y60, Q61, H99, H189, R183, H286). Isso demonstra que mesmo que as enzimas tenham sido cocrystalizadas com inibidores, as posições dos resíduos do sítio ligante ficam preservadas, bem como toda a estrutura restante das enzimas, o que está de acordo com os resultados mostrados na Tabela 3 (TM-Scores próximos de 1 e baixos RMSDs). Os alinhamentos foram feitos com o TM-align, usado para gerar as imagens e para fornecer dados de RMSD e TM-score sobre a qualidade dos alinhamentos (Yang Zhang, J. Skolnick, TM-align: a protein structure alignment algorithm based on the TM-score, Nucleic Acid Research, 33(7): 2302-2309, 2005).

**Figura 4.** Imagens da superfície dos sítios ligante e catalítico feitas pelo programa Molegro Virtual Docker. Em **4A1** e **4A2**, está visível a alfa-amilase suína (PPA, 1Hx0.pdb) e em **4B1** e **4B2**, está visível a TMA (1JAE.pdb). Comparando as duas enzimas (**4A1** e **4B1**), observa-se que o sítio ativo (círculo de linha mais espessa) tem poucas alterações. As regiões vizinhas ao sítio ativo e ligante (círculos pontilhados), no entanto, apresentam diferenças estruturais que podem estar envolvidas na especificidade de inibidores protéicos e também podem influenciar a afinidade com diferentes ligantes de baixo peso molecular. As duas regiões destacadas por círculos pontilhados em **4B1**, mostram as diferenças mais significativas entre essas enzimas: duas cavidades ao lado do sítio ligante que são inexistentes na PPA. Essas cavidades podem ser exploradas para o desenho de inibidores preferenciais às enzimas de insetos. Em **4A2** e **4B2**, observa-se moléculas de inibidor acarviosina-glicose ligadas às enzimas. O inibidor da PPA foi cocrystalizada com a enzima, e o inibidor presente na TMA foi posicionado por docking. Nota-se que as diferenças estruturais entre as enzimas alteram a conformação do inibidor ligado à enzima.

**Figura 5.** Sobreposição das estruturas do inibidor acarviosina-glicose no sítio ligante da PPA (1HX0.pdb) mostrando o inibidor cocrystalizado (em cinza claro) e o posicionado por docking (cinza escuro). Foi verificado pelo Ligand Energy Inspector do MVD e

pelo Java Protein Dossier do Sting\_DB que tanto no docking quanto na cocrystalização os inibidores interagem com os mesmos resíduos. Na visualização do MVD, nota-se também que ambos ocupam praticamente o mesmo espaço com pouca diferença de conformação, o que valida a técnica empregada para os dockings realizados neste trabalho. Imagem gerada pelo MVD.

**Figura 6.** Estruturas que serviram de núcleo para a ligação de grupos (“Rs”) no desenho de ligantes da série “trip”. Essas estruturas são semelhantes ao Triptofano visando interação do tipo stacking com os resíduos W56, W57 e Y60, pertencentes ao sítio ligante e próximos ao sítio ativo. Como pode ser visto em **B** e **C**, os anéis foram mantidos e as principais alterações foram: troca do Nitrogênio pelo Oxigênio (**B** e **C**), formando a estrutura do benzofurano (**B**); e o deslocamento da instauração do anel de furano para a ligação com o R1 (**C**).

**Figura 7.** Estruturas que serviram de núcleo para a ligação de grupos (“Rs”) no desenho de ligantes da série “T”. Assim como as da série “trip”, são derivadas da estruturado Triptofano e análogas ao benzofurano. Em **A**, temos uma estrutura derivada do benzofurano sem uma instauração no anel furano e sem uma instauração no anel benzeno. Em **B**, é apresentada uma estrutura semelhante à mostrada em **A**, mas sem o átomo de Oxigênio que caracteriza o anel furano. E em **C**, é mostrada uma estrutura semelhante ao benzofurano, mas sem uma instauração no anel benzeno e sem o átomo de Oxigênio.

**Figura 8.** Estrutura do núcleo para ligação de grupos (“Rs”) no desenho dos ligantes da série “T3”. É um mero aumento no número de anéis visando aumentar a abrangência de interações do tipo stacking na região do sítio ligante.

**Figuras 9.1 a 9.5.** Representação estrutural dos ligantes desenhados, por série. **Figura 9.1:** série “A”, ligantes derivados da acarbose. **Figura 9.2:** série “malt”, ligantes derivados da maltose. **Figura 9.3:** série “trip”, ligantes derivados da estrutura do triptofano com alguns grupos derivados da maltose. **Figura 9.4:** série “T”, ligantes derivados da estrutura do triptofano com grupos derivados da acarbose e acarviosina-glicose. **Figura 9.5:** série “T3”, ligantes derivados da série “T”, com grupos derivados

de acarbose e acarviosina-glicose e núcleos com maior extensão para realização de interação do tipo stacking.

**Figuras 10A e 10B.** “Poses” do ligante T3B posicionadas no sítio alvo da TMA (1JAE.pdb), figura 10A, e posicionado no sítio alvo da PPA (1HX0.pdb), figura 10B.

5 Comparando as duas figuras, observa-se que no docking com a TMA o ligante apresentou mais “poses” semelhantes do que em docking com PPA. Visto que também foi observado durante docking com inibidores conhecidos, que estes apresentaram várias “poses” semelhantes. Isto sugere que “poses” mais parecidas podem implicar em interação mais estável. Considerando ainda que o MolDock é um algoritmo baseado em  
10 evolução e que as “poses” finais foram selecionadas por realizarem interações energeticamente mais estáveis e melhor complementariedade geométrica, é então importante analisar também a disposição das “poses” no sítio alvo. Além disso, também observa-se que as “poses” do ligante interagem com cavidades na TMA que não estão presentes na PPA (conforme o mostrado na Figura 4). Baseado nisso, a maior  
15 semelhança entre as posições das poses no docking com a TMA sugere que a interação do ligante é mais favorável com essa enzima do que com a PPA.

### **DESCRIÇÃO DETALHADA DA INVENÇÃO**

A presente invenção refere-se a um método para o desenho computacional de novos compostos com potencial função inibitória das enzimas alfa-amilases. Um  
20 enfoque especial é dado ao desenho de novos compostos inibidores das enzimas alfa-amilases produzidas por insetos que se alimentam principalmente de produtos agrícolas estocados.

O método proposto consiste das seguintes etapas:

1 – Obter arquivos de coordenadas referentes às estruturas tridimensionais (a  
25 partir do banco PDB) da estrutura da alfa-amilase da enzima de interesse.

2 – Fazer busca por sequências homólogas à estrutura primária da enzima do de interesse usando o BLASTp contra o banco de dados de sequências primárias do PDB e NR (“non-redundant protein sequences”) do NCBI. O BLAST fornece os dados referentes às características do alinhamento de sequência primária, como: identidades  
30 (porcentagem de aminoácidos idênticos em posições correspondentes), positivos (são os números de aminoácidos idênticos somado ao número de aminoácidos que não são idênticos entre duas sequências numa determinada posição, mas que possuem

propriedades similares, considerado como um “Positivo” por indicar substituições nas quais a matriz BLOSUM-62 pontua positivamente segundo Altshul e colaboradores (1997). Os gaps indicam o número de regiões que representam lacunas no alinhamento (seja devido a inserções e deleções no histórico evolutivo dos genes) e, por último, a

5 coluna e-value (Expect Value) representa um parâmetro estatístico do programa BLAST que indica o número de diferentes alinhamentos que ocorreriam em um banco de dados por mero acaso com Score equivalentes ou melhores que o obtido para um determinado alinhamento. Quanto menor o e-value, mais significativo é o score obtido no alinhamento em questão (S.F. Altschul, T.L. Madden, A.A. Schäffer, Z. Zhang, W.

10 Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acid Res*, 25(17):3389-3402, 1997) para análise de homologia e com o programa ClustalW (M.A Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettingan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and Clustal X version 2.0,

15 *Bioinformatics*, 23, 2947-2948, 2007).

3 – Alinhar a enzima de interesse às estruturas de outras alfa-amilases depositadas no PDB para análise da sobreposição das estruturas terciárias e para analisar os desvios (RMSD – “root-mean-square deviation”) e TM-score. O RMSD fornece em Å a média quadrática das distâncias entre pares de resíduos correspondentes, depois da

20 sobreposição de uma estrutura sobre a outra. O TM-score utiliza o fator de Levitt-Gerstein que considera os pares de resíduos em menores distâncias como mais relevantes do que aqueles a maiores distâncias e também é mais sensível à topologia geral do que a variações estruturais localizadas. O TM-score é normalizado de forma a não ser dependente do tamanho da proteína, seu valor pode ir de 0.0 a 1.0, sendo valores

25 maiores que 0.17 indicativos de similaridade estrutural elevada (Yang Zhang, J. Skolnick, TM-align: a protein structure alignment algorithm based on the TM-score, *Nucleic Acid Research*, 33(7): 2302-2309, 2005). As estruturas terciárias são visualizadas através do PyMol para análise da sobreposição das estruturas terciárias. (W.L Delano, The PyMol molecular graphics system Delano Scientific, San

30 Carlos, CA, USA, <http://www.pymol.com.org>).

4 – Fazer alinhamentos entre a estrutura primária de proteínas homólogas à proteína de interesse através do programa ClustalW 2.0 (Larkin, M. A., Blackshields,

G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J., D., Gibson, T. J., Higgins, D. G. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23, 2947-2948, 2007) e incluir proteínas de outros organismos, evidenciando as similaridades e diferenças entre estes dois conjuntos de proteínas e buscando as correspondências no alinhamento de estrutura primária dos resíduos importantes. Seguindo certos critérios como: presença exclusivamente nas sequências de insetos, escolher os alvos terapêuticos preferenciais.

5        5        –        Buscar        estruturas        de        compostos        no        PubChem (http://pubchem.ncbi.nlm.nih.gov/) por similaridade maior que 90% com o ligante natural (amilose) e proceder simulações de docking empregando o programa Molegro Virtual Docker (MVD). O docking consiste na predição da orientação de uma molécula em relação a uma segunda quando elas se ligam para formar um complexo ligante-proteína. Há vários programas que realizam docking molecular, tais como DOCK, AUTODOCK, GOLD, FLEXX, ZDOCK, M-ZDOCK, MS-DOCK, Surflex, MCDOCK, MolDock, GemDock, entre outros. Os métodos usados por cada programa diferem basicamente na busca pela conformação que melhor se liga à proteína. Há três principais tipos: Estocástico e Sistemático, por Dinâmica Molecular e Algoritmos Genéticos/Evolucionários. O método de busca Estocástica e Sistemática por ligantes é bastante empregado e variado. O programa DOCK, por exemplo, utiliza um método baseado em conformações randômicas; FLEXX é baseado no espaço conformacional do sítio ativo. O método por Dinâmica Molecular, utilizado pelo AUTODOCK, calcula as várias conformações que a proteína alvo pode adotar durante a interação com o ligante. O MVD utiliza o algoritmo MolDcok que é baseado em Algoritmo Evolucionário (uma variedade de Método Estocástico e Sistemático) para simular as interações ligante-proteína. Os algoritmos evolucionários podem ser definidos como um grupo de aproximações computacionais baseadas nos conceitos da Teoria da Evolução de Darwin. O MolDock é uma implementação do algoritmo evolucionário focado em simulações de docking molecular, onde aproximações computacionais de um processo evolutivo são aplicadas para simular a permanência das características mais favoráveis (Thomsen R, Christensen, MH., MolDock: a new technique for high-accuracy molecular docking. *J Med Chem*. Jun 1;49(11):3315-21, 2006; W.F. De Azevedo Jr, MolDock applied to structure-based virtual screening, *Current Drug Targets*, 11(3):327-

10

15

20

25

30

334, 2010). Trabalhos que comparam os programas mais utilizados indicam que programas baseados em algoritmo evolucionário tais como GemDock e MolDock apresentam melhor performance geral do que Flexx, GOLD e Surflex (R. Thomsen, M.H, Christensen, MolDock: a new technique for high-accuracy molecular docking, J Med Chem, 49, 3315-21, 2006; J.M. Yang, C.C. Chen, GemDock: a generic evolutionary method for molecular docking, Proteins, 55, 288-304, 2004). Neste trabalho, os dockings foram feitos entre os ligantes obtidos no PubChem e a TMA para seleção de estruturas com melhor afinidade com os resíduos do sítio alvo principalmente em relação aos aminoácidos conservados.

6 – Uma análise detalhada do nano-ambiente do sítio alvo é realizada usando-se o programa STING Java Protein Dossier (Neshich G., Rocchia W., Mancini A.L., Yamagishi M.E., Kuser P.R., Fileto R., Baudet C., Pinto I.P., Montagner A.J., Palandrani J.F., Krauchenco J.N., Torres R.C., Souza S., Togawa R.C., Higa R.H. 2004. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. Nucleic Acids Res. 1;32 (Web Server issue):W595-601, 2004) e a ferramenta “select”. Desta forma, são determinados conjuntos de parâmetros físico-químicos e estruturais que descrevem o nano-ambiente que é composto pelos resíduos que compõem o sítio-alvo, no qual se espera que os ligantes atraiam. Estes conjuntos são usados não apenas para identificação nas estruturas estudadas, como também para melhor compreender o ambiente para qual o inibidor seria construído, tendo a eficácia necessária para sua aplicação final. Portanto estes parâmetros indicam as características físico-químicas do alvo terapêutico para o adequado desenho computacional e modificação dos novos compostos.

7 – As estruturas selecionadas no item “5” devem ser modificadas racionalmente, de acordo com as características físico-químicas e estruturais encontradas no item “6” (por exemplo, se o potencial eletrostático local é alto, adicionam-se os átomos adequados para diminuir este potencial no composto inicial), a fim de aumentar os valores de afinidade de ligação. As alterações nos compostos são feitas computacionalmente através do programa ChemBioDraw, no qual também são feitas minimização de energia e dinâmica das estruturas de ligantes desenhadas (Zhenjiang Li, Honggui Wan, Yuhu Shi, Pingkai Ouyang, Personal experience with four kinds of chemical structure drawing software: review on ChemDraw, ChemWindows,

ISIS/Draw, and ChemSketch, J. Chem. Inf. Comput. Sci., 44, 1886-1890, 2004). Além da busca por ligantes baseados na estrutura de amilose no PubChem, também é feita busca por ligantes baseados na estrutura dos inibidores naturais. Os demais passos de docking e modificações estruturais também são aplicados a esses ligantes.

- 5           8 – Realizar predição de parâmetros moleculares que interferem na Absorção, Distribuição, Metabolismo, Excreção e Toxicidade utilizando a ferramenta online Osiris Property Explorer (Thomas Sander, Actelion Pharmaceuticals Ltd., <http://www.organic-chemistry.org/prog/peo/>) para orientar no desenho dos compostos. O programa Osiris fornece dados de risco de toxicidade: mutagenicidade, tumorigenicidade, irritante e com efeito reprodutivo, baseados em fragmentos de moléculas conhecidas depositados no banco de dados RTECS. Também prediz valores de cLogP, solubilidade em água, peso molecular e fornece o Drug-score que é um valor geral resultante da combinação de todos os parâmetros anteriores mais o Druglikeness, que compara fragmentos do composto desenhado com base de dados de compostos comercializados e com base de dados de compostos que não servem como fármacos (Fluka).

### **EXEMPLOS**

A invenção será agora descrita em maiores detalhes por meio dos exemplos a seguir, o qual não devem ser interpretados como limitativos do escopo da invenção.

- Exemplo 1 – *Análise de similaridade e identidade entre as alfa-amilases de diferentes espécies.*

- Foram obtidos os arquivos de coordenadas referentes às estruturas tridimensionais existentes do *Tenebrio molitor* através do Protein Data Bank (<http://www.pdb.org>). As alfa-amilases de insetos, cujas estruturas tridimensionais estão resolvidas e depositadas no PDB, são: enzima da larva do *T. molitor* (1JAE), *T. molitor* em complexo com inibidor Amaranth (1CLV), *T. molitor* em complexo com inibidor Ragi (1TMQ) e *T. molitor* em complexo com inibidor do *Phaseolus vulgaris* (1VIW). Como o trabalho tem como alvo principal o estudo de alfa-amilase de insetos para o desenvolvimento de inibidores, foi executado o programa BLASTp usando como query a sequência da alfa-amilase de *T. molitor* (TMA) contra a database NR do NCBI filtrada para proteínas de outros insetos. Foram evidenciados altos valores de identidade e similaridade, ambos acima de 60% e 75%, respectivamente, para a maioria das sequências de insetos obtidas como “hits” no BLASTp (Tabela 1).

**Tabela 1** – Resultados obtidos com o BLASTp (SF Altschul, TL Madden, AA Schäffer, Z Zhang, W Miller, DJ Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acid Res. 1997 Sep 1;25(17):3389-402.) entre a alfa-amilase do *Tenebrio molitor* (TMA, 1JAE.pdb) e

5 sequências do banco NR do NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)), sendo reportados os três melhores “hits” obtidos contra insetos, mamíferos, fungos e plantas. São fornecidos os dados referentes às características do alinhamento de sequência primária realizado pelo BLAST, como: identidades (porcentagem de aminoácidos idênticos em posições correspondentes), positivos (são o número aminoácidos idênticos somado ao número de

10 aminoácidos não idênticos entre duas sequências numa determinada posição, mas que possuem propriedades similares, considerado como um “Positivo” por indicar substituições nas quais a matriz BLOSUM-62 pontua positivamente segundo Altschul e colaboradores (1997), a coluna gaps indica o número de regiões que representam lacunas no alinhamento (seja devido a inserções e deleções no histórico evolutivo dos

15 genes) e, por último, a coluna e-value (Expect Value) que representa um parâmetro estatístico do programa BLAST que indica o número de diferentes alinhamentos que ocorreriam em um banco de dados por mero acaso com Score equivalentes ou melhores que o obtido para um determinado alinhamento. Quanto menor o e-value, mais

significante é o score obtido no alinhamento em questão.

| Resultados de BLASTp: TMA (1JAE.pdb) contra NR do NCBI |            |           |      |         |
|--------------------------------------------------------|------------|-----------|------|---------|
| espécie                                                | %          |           |      | e-value |
|                                                        | identidade | positivos | gaps |         |
| insetos (taxid 6960)                                   |            |           |      |         |
| <i>Blaps mucronata</i> (churchyard beetle)             | 81         | 89        | 0    | 0.0     |
| <i>Tribolium castaneum</i> (besouro castanho)          | 79         | 87        | 0    | 0.0     |
| <i>Diabrotica virgifera virgifera</i> (larva alfinete) | 65         | 78        | 2    | 0.0     |
| mamíferos (taxid 40674)                                |            |           |      |         |
| <i>Mus musculus</i> (camundongo)                       | 55         | 68        | 5    | 2e-146  |
| <i>Rattus norvegicus</i> (ratazana)                    | 54         | 68        | 5    | 3e-144  |
| <i>Bos taurus</i> (bovino)                             | 53         | 67        | 5    | 2e-142  |
| fungos (taxid 4751)                                    |            |           |      |         |
| <i>Chaetomium globosum</i>                             | 42         | 57        | 7    | 1e-88   |
| <i>Schizophyllum commune</i>                           | 41         | 57        | 9    | 8e-88   |
| <i>Melampsora larici-populina</i>                      | 37         | 50        | 10   | 3e-62   |
| plantas (taxid 3139)                                   |            |           |      |         |
| <i>Physcomitrella patens ssp patens</i> (musgo)        | 23         | 38        | 13   | 1e-08   |
| <i>Glycine max</i> (soja)                              | 24         | 37        | 17   | 2e-06   |
| <i>Nicotiana glauca</i> x <i>N. glauca</i> (tabaco)    | 25         | 37        | 17   | 4e-05   |

No BLASTp contra NR de insetos (Tabela 1), dos três melhores “hits” apenas um (*Melampsora larici-populina*) tem importância agrícola. As enzimas fúngicas não serão consideradas neste trabalho devido às baixas identidade e similaridade com a TMA. Foi também feito o BLASTp com enzimas de plantas (Tabela 1) cujos alinhamentos com a sequência query TMA (1JAE\_A) apresentaram valores de similaridade e identidade muito menores. Não foi encontrada nenhuma estrutura disponível no PDB de alfa-amilase vegetal; e as estruturas de insetos disponíveis em PDB apresentaram identidade e similaridade muito baixas (entre 30 e 40%) em relação à TMA (dado não mostrado), além de não serem de insetos caracterizados como pragas agrícolas. Em relação às enzimas de mamíferos, foi feito BLASTp contra NR do NCBI (Tabela 1) para sequências de mamíferos, que apresentou identidade e similaridade maiores do que as vegetais. No BLASTp contra PDB (Tabela 2) apenas resultaram “hits” de enzimas de mamíferos.

**Tabela 2** – Resultados de BLASTp contra o banco PDB usando-se como sequência de busca (“query sequence”) a sequência da TMA (1JAE.pdb). São mostrados os três melhores resultados “hits” encontrados. A descrição das colunas segue o padrão da tabela 1.

| Resultados de BLASTp: TMA (1JAE.pdb) contra PDB |            |           |      |         |
|-------------------------------------------------|------------|-----------|------|---------|
| organismo / origem / código PDB                 | %          |           |      | e-value |
|                                                 | identidade | positivos | gaps |         |
| porco / pancreático / 1KXQ                      | 53         | 67        | 5    | 1e-141  |
| humano / salivar / 3BLK                         | 53         | 66        | 6    | 1e-139  |
| humano / pancreático / 2CPU                     | 52         | 66        | 5    | 7e-138  |

#### Exemplo 2 – Identificação dos resíduos-alvo

Sabendo-se da importância dos insetos no estoque de grãos e pensando-se no desenho de um fármaco capaz de interagir com mais de uma espécie de inseto, foi feito um alinhamento de sequência primária usando o programa ClustalW (Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J., D., Gibson, T. J., Higgins, D. G. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23, 2947-2948, 2007), que é mostrado na Figura 2. Nesta figura também estão alinhadas duas sequências de alfa-amilases vegetais (de PpA - *Physcomitrella patens ssp patens* e GmA - *Glycine max*) com o intuito de compará-las com sequências alfa-amilases de insetos para buscar por aminoácidos que sejam específicos de insetos em relação às plantas, para não interferir

na função enzimática vegetal, e que estejam envolvidos na ligação com o substrato. São exibidas apenas as duas sequências de enzimas vegetais que apresentaram melhores resultados no BLASTp, ou seja, são exibidas as duas sequências vegetais mais similares às sequências de insetos. Da mesma forma, são exibidas as quatro sequências de mamíferos que apresentaram melhores resultados no BLASTp (HSA – alfa-amilase salivar humana, HPA – alfa-amilase pancreática humana, PPA – alfa-amilase pancreática suína, MPA – alfa-amilase pancreática e camundongo, RPA – alfa-amilase pancreática de rato). Esse alinhamento mostra de forma resumida a análise que foi feita entre várias sequências (não mostrado) para facilitar a visualização da análise. Devido a isso, alguns resíduos não destacados que nessa figura parecem ser exclusivos de insetos, em análise mais aprofundada não o são. O resíduo G210, por exemplo, apesar de estar presente nas sequências de TMA e TcA, não é muito conservado em insetos e alguns mamíferos apresentam resíduos com características similares. Isso remete à discussão sobre como identificar um alvo terapêutico ideal para o desenho computacional de novos compostos. Anderson publicou uma descrição de como deve ser realizado o procedimento de desenho computacional de fármacos baseado em estruturas protéicas, desde os critérios que devem ser adotados para a escolha do alvo terapêutico para tal desenho até os processos de desenvolvimento, docking e virtual screening (AC Anderson. The process of structure-based drug design. Chem Biol. Sep;10(9):787-97, 2003). O trabalho discute que o desenho de drogas antimicrobianas deve se basear em alvos que sejam essenciais, principalmente encontrados nos organismos que se pretende atacar, tenham uma função única nesse organismo e que sejam passíveis de sofrerem inibição por pequenas moléculas. Pensando nisso, o alvo considerado, a enzima alfa-amilase, é essencial para a sobrevivência dos insetos que se alimentam de grãos, é passível de inibição (já demonstrada praticamente através de estudos com inibidores protéicos e de baixo peso molecular) e, sugerimos a seguir resíduos a serem usados como alvos preferenciais por não estarem presentes nas estruturas das enzimas vegetais ao passo que são essenciais para que as enzimas de insetos se liguem ao substrato. Considerando as enzimas de mamíferos, estas podem ser consideradas alvos potenciais já que roedores também são considerados pragas de estoque de grãos. Em relação às enzimas humanas, em outro trabalho, poderiam ser consideradas como alvos terapêuticos no tratamento da Diabetes Mellitus tipo 2. Porém, neste trabalho tais

enzimas não podem ser consideradas alvos. Como os resíduos elencados como cruciais para ligação ao substrato estão presentes tanto em insetos como mamíferos, os novos ligantes foram desenhados para ter menor afinidade teórica com as enzimas de mamíferos e maior afinidade com enzimas de insetos. Para isso, foram consideradas as

5 poucas diferenças estruturais que circundam os sítios ligantes e catalíticos das enzimas de mamíferos e insetos, explicadas mais adiante. Espera-se que também nos teste *in vitro* e *in vivo*, os ligantes desenhados tenham maior afinidade com as enzimas de insetos do que com as de mamíferos.

Analizando o alinhamento e características estruturais de resíduos próximos aos

10 envolvidos na catálise e ligação ao substrato, pode-se perceber a existência de cinco aminoácidos importantes, presentes nas alfa-amilases de insetos e mamíferos e ausentes nas alfa-amilases vegetais: W56, W57, Q61, V151 e N331. No alinhamento, o alto grau de identidade dos resíduos especificamente envolvidos na catálise e ligação do substrato e a alta similaridade nos leva a crer que os compostos aqui desenhados para a TMA

15 seriam igualmente capazes de inibir as alfa-amilases do *Tribolium castaneum* e, possivelmente, de outros insetos e mamíferos. As estruturas da TMA cocrystalizada com inibidores protéicos (1CLV.pdb, 1TMQ.pdb, 1VIW.pdb), as estruturas da PPA complexada com inibidores (1HX0.pdb, 1KXQ.pdb, 1DHK.pdb) e a estrutura da HPA cocrystalizada com acarbose (2QV4.pdb) depositadas no PDB, foram analisadas através

20 de visualização no programa Java Protein Dossier da plataforma STING (Neshich G., Rocchia W., Mancini A.L., Yamagishi M.E., Kuser P.R., Fileto R., Baudet C., Pinto I.P., Montagner A.J., Palandrani J.F., Krauchenco J.N., Torres R.C., Souza S., Togawa R.C., Higa R.H. Java Protein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. Nucleic Acids Res. 2004 Jul 1;32 (Web

25 Server issue):W595-601, 2004), de onde foi possível extrair informações de que aminoácidos estão interagindo por meio de contatos com os inibidores complexados às estruturas. Estes são os resíduos do sítio ativo D185, E222 e D287, e sítio ligante: W56, W57, Y60, Q61, H99, R183, A186, H189 e H286, que são bastante conservados entre os insetos, segundo o alinhamento feito.

### 30 Exemplo 3 – *Análise do nanoambiente do sítio ligante e catalítico*

Para uma análise comparativa das posições estruturais dos aminoácidos relevantes ao presente estudo, foram efetuadas sobreposições estruturais entre as

diversas estruturas resolvidas de alfa-amilases depositadas no PDB (Tabela 3).

**Tabela 3** – Valores de RMSD e TM-Score resultantes de alinhamento estrutural entre a TMA (1JAE.pdb) e as estruturas de alfa-amilases de insetos e mamíferos depositadas no banco PDB. O RMSD fornece a média quadrática das distâncias entre pares correspondentes de resíduos depois da sobreposição de uma estrutura sobre a outra. O TM-score (Yang Zhang, J. Skolnick, TM-align: a protein structure alignment algorithm based on th TM-score, Nucleic Acid Research, 33(7): 2302-2309, 2005) utiliza o fator de Levitt-Gerstein que considera os pares de resíduos em menores distâncias como mais relevantes do que aqueles a maiores distâncias e também é mais sensível à topologia geral do que a variações estruturais localizadas. O TM-score é normalizado de forma a não ser dependente do tamanho da proteína, seu valor pode ir de 0.0 a 1.0, sendo valores maiores que 0.17 indicativos de similaridade estrutural elevada. Como o TM-score leva em consideração principalmente a topologia, esses resultados indicam que todas as enzimas possuem alta correlação estrutural com a TMA (1JAE.pdb), mesmo quando em complexo com inibidores.

| Alinhamento estrutural com a TMA (1JAE.pdb) |         |        |        |        |        |        |
|---------------------------------------------|---------|--------|--------|--------|--------|--------|
|                                             | 3BLK    | 2CPU   | 1HXD   | 2CLV   | 1VIW   | 1TMO   |
| TM-score                                    | 0.93395 | 0.9358 | 0.9345 | 0.9967 | 0.9896 | 0.9975 |
| RMSD                                        | 0.91    | 0.84   | 0.90   | 0.50   | 0.74   | 0.39   |

Foram calculados os valores de desvio no alinhamento entre as estruturas terciárias alinhadas através do programa TM-align (Yang Zhang, J. Skolnick, TM-align: a protein structure alignment algorithm based on th TM-score, Nucleic Acid Research, 33(7): 2302-2309, 2005). Os valores fornecidos pelo TM-align são o RMSD e o TM-score, sendo que o primeiro calcula a média quadrática da distância entre os resíduos correspondentes considerando apenas as distâncias entre pares de resíduos e sendo influenciado pelo tamanho da proteína, não levando em conta sua topologia. Já o TM-Score é calculado levando em conta a topologia das estruturas sobrepostas, o que é bastante importante uma vez que a maioria das interações entre proteínas são influenciadas pela topologia. A sobreposição das estruturas gerada pelos alinhamentos estruturais entre a TMA (1JAE.pdb) e a TMA complexada com inibidor Amaranth (1CLV.pdb), e entre a TMA (1JAE.pdb) e a PPA complexada com acarviosina-glicose

(1HX0.pdb) são mostradas nas figuras 3A e 3B<sup>\*</sup>, respectivamente, com enfoque nos resíduos importantes do sítio ligante (W56, W57, Y60, Q61, H99, R183, H189 e H286). Foi observado, nestas sobreposições, que as posições dos aminoácidos nos sítios ligantes das proteínas são bastante conservadas estruturalmente, o que reforça a idéia de que um ligante possa funcionar como inibidor da alfa-amilase ligando-se a esses aminoácidos não somente da TMA como de enzimas de outros insetos ou até de mamíferos. Apesar de a posição desses aminoácidos serem bem conservadas tanto em insetos quanto em mamíferos, há diferenças no posicionamento de outros resíduos que circundam o sítio ligante de modo a direcionar a interação com os ligantes (Figuras 4A e 4B). A visualização e confecção de imagens foram feitas com o programa MVD (Thomsen R, Christensen MH.: MolDock: a new technique for high-accuracy molecular docking. J Med Chem. Jun 1;49(11):3315-21. 2006;W.F. De Azevedo Jr, MolDock applied to structure-based virtual screening, Current Drug Targets, 11(3):327-334, 2010).

Um trabalho adicional foi realizado para a análise da conservância de parâmetros físico-químicos e estruturais dentre os resíduos importantes entre as estruturas de alfa-amilases. Sabe-se que a função da proteína é mais conservada com relação à estrutura 3D do que em relação à sequência primária (Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. EMBO J. Apr;5 (4):823-6, 1986), havendo várias proteínas com pouca ou nenhuma similaridade de sequência que, mesmo assim, têm similaridade estrutural. Conhecer a estrutura permite-nos explicar o mecanismo bioquímico pelo qual a proteína implementa sua funcionalidade. Os fatores que determinam a funcionalidade do sítio ativo de uma proteína são muito complexos e dependem da estrutura tridimensional, além das propriedades bioquímicas e biofísicas.

É possível analisar os sítios funcionais como um nano-ambiente físico-químico que acompanha uma função, ao invés de considerá-los como um grupo de resíduos fixos. Assim, buscamos uma caracterização mais genérica do sítio ativo (e/ou conjunto

---

<sup>\*</sup> São apenas mostradas duas sobreposições, referentes às figuras 3A e 3B, embora tenham sido feitas sobreposições de todos os alinhamentos para análise da posição dos resíduos citados no texto. As demais sobreposições foram omitidas pois são praticamente idênticas às apresentadas. As sobreposições das figuras 3A e 3B foram escolhidas por apresentarem estruturas co-cristalizadas com inibidor a fim de verificar se a ligação com inibidor altera a forma da alfa-amilase no sítio ligante.

dos aminoácidos que constroem o chamado sítio de ligação ao substrato), a partir de descritores estruturais, sem considerar necessariamente a conservação como o principal fator e sim a geometria específica, que sejam úteis para a previsão dos resíduos catalíticos das enzimas.

5           Aplicamos um método para seleção de descritores da estrutura contidos no STING\_DB (Neshich, G., Togawa, R., Mancini, A. L., Kuser, P. R., Yamagishi, M. E. B., Pappas Jr., G., Torres, W. V., Campos, T. F., Ferreira, L. L., Luna, F. M., Oliveira, A. G., Miura, R. T., Inoue, M. K., Horita, L. G., de Souza, D. F., Dominiquini, F., Álvaro, A., Lima, C. S., Ogawa, F. O., Gomes, B. G., Palandrani, J. C. F., dos Santos,  
10 G. F., de Freitas, E. M., Mattiuz, A. R., Costa, I. C., de Almeida, C. L., Souza, S., Baudet, C. and Higa, R. H. STING Millennium: a Web based suite of programs for comprehensive and simultaneous analysis of protein structure and sequence. *Nucleic Acids Research*, 31:13, 3386-3392, 2003), projetado especificamente para previsão do sítio catalítico e do sítio de ligação ao substrato de enzimas. Apenas alguns resíduos da  
15 enzima participam da catálise, ao passo que a grande maioria dos resíduos são não reativos. Outros resíduos são caracterizados como aqueles que servem para o atracamento de ligante ou substrato e estes também tem suas características que podem ser usadas para sua seleção e consequentemente, para descrição do nano-ambiente criado por este conjunto de resíduos que compõem o sítio de ligação. Com a  
20 identificação de um conjunto dos descritores de nano ambiente, estamos habilitados para a construção de uma tabela de assinaturas (consulta sobre os descritores estruturais dos resíduos restringindo seus intervalos de valores) para os diversos membros de famílias enzimáticas. Estas assinaturas devem funcionar em todos os membros se identificadas com rigor. No caso da proteína alfa-amilase, identificamos apenas dois  
25 parâmetros que com sucesso foram capazes de selecionar em todas as estruturas existentes no PDB o conjunto de resíduos que compõem o sítio catalítico que, como já foi mencionado anteriormente, é constituído pelos três aminoácidos: Aspartato 185, Glutamato 222 e Aspartato 287, segundo numeração da TMA (1JAE.pdb).

Foram examinadas 8 estruturas de alfa-amilases (em parêntesis a numeração dos  
30 aminoácidos da tríade catalítica correspondentes em cada estrutura) 1JAE (D185, E222, D287), 1CLV (D185, E222, D287), 1DHK (D197, E233, D3000), 1KXQ (D197, E233, D3000), 1TMQ (D185, E222, D287), 2VIW (D185, E222, D287), 2QV4 (D197, E233,

D3000) e 3BLK (D197, E233, D3000). Usando a ferramenta Select do Java Protein Dossier. O conjunto de descritores e seus intervalos de valores que caracterizam unicamente estes resíduos foram: Valor do parâmetro Conservation HSSP Relative Entropy  $\leq 2$  e do parâmetro Electrostatic Potencial no último átomo pesado da cadeia lateral de aminoácido (LHA) num intervalo de  $[-218, -180 \text{ kT/e}]$ , o que indica a característica do potencial eletrostático medido em volta de último átomo pesado da cadeia lateral, sendo este negativo e os resíduos possuem alta conservação na evolução deste aminoácido no contexto da estrutura desta enzima (indicado pelo baixo valor de “Conservation HSSP Relative Entropy”). Em adição, foram identificados os aminoácidos constituintes de sítio de ligação ao substrato para estas estruturas, de acordo com as interações com inibidores cocrystalizados indicadas pelo “Protein Java Dossier”. O conjunto de descritores e seus respectivos valores, usados para identificar o nano ambiente criado na estrutura protéica por eles, são bem mais complexos e diversos. É constituído pelos seguintes parâmetros STING, com seus respectivos intervalos: Conservation HSSP Relative Entropy  $\leq 5$ ; Reliability (x100)  $\geq 83,7$ ; Rotamers Percent  $\leq 13$ ; Unused Contacts Total  $\geq 62$ ; Unused Contacts Energy Total  $\geq 127$ ; Hydrophobic Scale Isolation  $< 1$ ; Surface Accessibility Isolation  $< 33$ . Pode-se observar que estes aminoácidos, além de conservados altamente, possuem rotâmeros alternativos, também possuem baixo valor de intervalo de potencial eletrostático médio, -39.38 a 20.75 kT/e. Outra característica observada foi o alto numero de contatos não usados (indicando alto potencial em fazer contatos com molécula de substrato) e tendo hidrofobicidade baixa e acessibilidade dos resíduos ao solvente alta. Estes conjuntos foram usados não apenas para identificação nas estruturas estudadas, como também para melhor compreender o ambiente para qual o inibidor seria construído, tendo a eficácia necessária para sua aplicação final. Portanto, estes parâmetros indicam as características físico-químicas do alvo terapêutico para o desenho adequado dos novos compostos (Da Silva, M.C.M., Grossi da Sá, M.F., Chrispeels, M.J., Togawa, R.C., Neshich, G., Analysis of structural and physico-chemical parameters involved in the specificity of binding between alpha-amylases and their inhibitors, Protein engineering, 13, 3:167-177, 2000).

#### Exemplo 4 – *Desenho computacional dos ligantes*

Para o desenho computacional de ligantes partiu-se da estrutura de fragmento do

ligante natural (maltose) e dos inibidores de baixo peso molecular já existentes (acarbose, acarviosina e isoacarbose). A partir deles foram feitas buscas por compostos que possuíam similaridade na estrutura química acima de 90% na base de dados PubChem (<http://pubchem.ncbi.nlm.nih.gov/>) através do “PubChem Structure Search” (Junguk Hur, David J Wild: PubChemSR: A search and retrieval tool for PubChem, Chemistry Central Journal. 2:11 doi:10.1186/1752-153X-2-11, 2008) e, ao todo, foram obtidas 1881 estruturas. A estratégia para o desenho desses ligantes partiu da observação da posição de todos os aminoácidos do sítio ligante e confirmação de ausência em sequência de enzimas vegetais através de vários alinhamentos feitos usando o CLUSTALw entre estruturas primárias de insetos, plantas e mamíferos. Todas as estruturas desenhadas foram então submetidas a simulações de docking com a estrutura da TMA, código PDB 1JAE. O docking foi realizado com o programa Molegro Virtual Docking (MVD) (Thomsen R, Christensen MH.: MolDock: a new technique for high-accuracy molecular docking. J Med Chem. Jun 1;49(11):3315-21, 2006), que é baseado em algoritmo evolutivo para simular as interações ligante-proteína.

As simulações de docking (com ligante flexível) usando-se o programa MVD, foram efetuadas centralizando-se nos resíduos do sítio ativo mencionado e com um raio de esfera de espaço de busca de 14Å centralizada no sítio catalítico. O algoritmo de cálculo de score usado foi o MolDock Score [GRID], usando uma resolução de Grid de 0,2Å (avaliando-se o ligante por meio de Interações internas de pontes de hidrogênio, interações eletrostáticas e torções  $Sp^2-Sp^2$ ). O algoritmo de busca e docking usado foi o MolDock Optimizer (que é uma implementação de uma variação do algoritmo evolutivo), com 20 rodadas, tamanho de população igual a 100 e com 3000 iterações. Após o docking, o MVD realizou Minimização de energia e otimização de pontes de Hidrogênio das “poses” (sendo que uma pose é um modo de ligação candidato) obtidas.

Com o MVD é possível visualizar os resultados das simulações em modelos tridimensionais, o que torna possível eliminar facilmente as “poses” que adotaram posições visualmente errôneas, não tendo de levar apenas em consideração os valores numéricos fornecidos pelo programa para isso. Dada a importância dos resíduos preferenciais, ausentes nas plantas, W56, W57, Q61, I224 N331 e D332, e os envolvidos na interação ligante-proteína Y60, H99, R183, K188, H189 e H286, os compostos foram então selecionados de acordo com os valores de energia, em relação a

esses resíduos, fornecidos pelo programa através do módulo “Ligand Energy Inspector”. Foi também considerado o MolDock score, que é a energia total da pose, como parâmetro secundário na seleção de ligantes.

5 A cada docking seguiu-se a seleção de ligantes e modificações estruturais nos mesmos a fim de melhorar a interação com os resíduos-alvo. As modificações foram feitas com o programa ChemBioDraw (<http://www.cambridgesoft.com/software/ChemBioDraw/>) e as estruturas foram energeticamente minimizadas com o mesmo.

10 No virtual screening, no docking com a maltose (produto da amilose), docking com a amilose, com a acarbose, com acarviosina e com isoacarbose, os resíduos que apresentaram melhor energia de interação foram os que realizaram interações tipo stacking com anéis dos ligantes; esses resíduos foram o W56, W57 e Y60. Além desses resíduos, também se destacaram Valina 151 e Isoleucina 224, com interações hidrofóbicas. Os inibidores (isoacarbose e acarviosina-glicose) apresentaram energia de  
15 interação menor (menor energia significa interação mais estável) do que a maltose em relação aos resíduos W56, W57 e Y60, no docking com a TMA, sugerindo que esses resíduos fazem as principais interações de ligação ao substrato. Já no docking com a PPA (alfa-amilase pancreática de porco) os resíduos com menor energia foram V163 (V151, na TMA) e W59 (W57, na TMA). Pensando-se em um desenho de estruturas  
20 que se liguem preferencialmente às enzimas de insetos, os resíduos W56, W57 e Y60 foram considerados principais alvos. Dada à inexistência de resíduos exclusivos de insetos, em relação às de mamíferos, que participem na ligação e catálise do substrato, espera-se, tanto nas simulações computacionais como nos testes *in vitro* e *in vivo*, que as interações dos ligantes com as enzimas de mamíferos sejam piores do que as  
25 interações com enzimas de insetos.

O posicionamento dos ligantes no sítio catalítico e ligante foi validado a partir de comparação com estrutura de enzima cocrystalizada com acarviosina-glicose (PPA, 1HX0.pdb). A Figura 5 mostra a sobreposição entre o docking feito com a PPA e o ligante cocrystalizado da mesma enzima.

30 Os ligantes candidatos a fármacos foram desenhados seguindo as estratégias:

- Série “A”: as estruturas dessa série foram desenhadas a partir da estrutura do inibidor acarbose, com mudanças esteroquímicas e adição de insaturações.
- Série “malt”: foram baseados na estrutura da maltose. Também foram modificados em relação à orientação dos grupos e insaturações.
- 5   • Série “trip”: as estruturas foram desenhadas a partir da estrutura do resíduo de triptofano visando a realização de interação tipo stacking com os triptofanos e tirosinas da enzima. A estrutura do anel do triptofano foi mantida e o Nitrogênio foi trocado por Oxigênio (figura 6), pois o Nitrogênio nessa posição teria repulsão com o Nitrogênio do triptofano da enzima. Nesses ligantes foram
- 10   colocados grupos (“Rs”) derivados da maltose e outros grupos.
- Série “T”: foram feitas moléculas a partir das estruturas derivadas do triptofano (Figura 7), com “Rs” derivados da acarbose e acarviosina-glicose, variando em número de “Rs” e em número de resíduos glicosídicos por “R”.
- Série “T3”: moléculas derivadas do ligante T3 mas com alterações no núcleo
- 15   derivado de triptofano (Figura 8), com aumento do número de anéis.

Todos os ligantes desenhados são mostrados nas figuras 9.1 a 9.5.

Foram obtidas 4 estruturas derivadas daquelas que apresentaram melhores resultados nas rodadas de dockings anteriores e que sugere-se que podem apresentar atividade inibitória para a(s) enzima(s) alfa-amilases(s). Nas tabelas 4 e 5 são

20   demonstrados os resultados de docking dos ligantes desenhados, de maltose e dos inibidores acarbose, isoacarbose e acarviosina, contra as enzimas TMA (1JAE.pdb) e PPA (1HX0.pdb), sendo que foi feito também docking com esta última para comparação dos valores entre enzima de inseto e enzima de mamífero (a 1HX0 possui 53% de identidade, 67% de positivos e 6% de gaps em relação à TMA, com e-value de

25   1e-145).

**Tabela 4** – Valores de energias de interação (em Kcal/mol) entre os ligantes e os resíduos conservados (W56, W57, Y60) entre as alfa-amilases de insetos e mamíferos, soma das energias de interações estabelecidas entre os ligantes e os três resíduos especificados e valores de Moldock score (energia total do ligante), Protein (total

30   energia de interação do ligante com a proteína) e Hbond (energia decorrente de

interações de Hidrogênio) fornecidos pelo Molegro Virtual Docker (Thomsen R, Christensen, MH., MolDock: a new technique for high-accuracy molecular docking. J Med Chem. 2006 Jun 1;49(11):3315-21), no módulo Ligand Energy Inspector. Estes resultados foram obtidos da simulação de docking dos ligantes desenhados com o espaço de busca centralizado no sítio catalítico da TMA. Para o desenho das novas moléculas foram especialmente consideradas as interações com os resíduos especificados. A soma das energias de interação entre os ligantes e esses resíduos foi o primeiro parâmetro de avaliação para exclusão e modificação das estruturas dos novos compostos. O segundo parâmetro utilizado foi o MolDock score. Também foram consideradas as energias de interação com proteína e de ligação de hidrogênio. Observa-se então os valores de energia total entre esses resíduos e os ligantes desenhados se comparados aos valores encontrados no docking dos inibidores, são menores ou iguais, o que sugere que a ligação dos ligantes desenhados a esses resíduos pode ser mais estável.

| Ligantes            | Energia de interação (kcal/mol) entre os ligantes e os resíduos da TMA (1JAE.pdb) |        |        |        | MolDock score | Protein | Hbond  |
|---------------------|-----------------------------------------------------------------------------------|--------|--------|--------|---------------|---------|--------|
|                     | W58                                                                               | W57    | Y62    | Soma   |               |         |        |
| maltose             | -20,05                                                                            | -31,96 | -13,66 | -45,61 | -98,10        | -125,89 | -20,34 |
| acarbose            | -7,46                                                                             | -4,28  | 0,00   | -11,74 | -180,33       | -116,76 | -11,99 |
| acarviosina-glicose | -11,33                                                                            | -39,28 | -24,90 | -75,51 | -166,69       | -250,46 | -41,73 |
| isacarbose          | -26,75                                                                            | -6,52  | -22,63 | -55,90 | -187,26       | -188,53 | -16,72 |
| A1                  | -16,37                                                                            | -26,42 | -0,60  | -43,40 | -156,63       | -99,20  | -7,29  |
| malt002             | -8,77                                                                             | -24,60 | -23,50 | -56,88 | -130,71       | -114,55 | -14,82 |
| trip005             | -8,90                                                                             | -30,25 | -21,98 | -61,13 | -81,15        | -96,80  | -0,68  |
| T3                  | -13,89                                                                            | -25,54 | -16,25 | -55,68 | -134,81       | -222,01 | -19,99 |
| T3B                 | -10,72                                                                            | -46,90 | -19,42 | -77,04 | -163,86       | -184,21 | -16,56 |

Tabela 5– Valores fornecidos pelo docking dos ligantes desenhados com espaço de busca centralizado no sítio catalítico da alfa-amilase pancreática suína (PPA, 2HX0.pdb), conforme padrão da tabela 4. Foi feito um docking com essa enzima a fim de verificar se os ligantes desenhados também poderiam se ligar às enzimas de mamíferos. Nesse docking observa-se que os valores de MolDock score e Protein são menores do que os obtidos em docking com a TMA, no entanto, esses valores são decorrentes do formato mais fechado do sítio alvo da PPA, enquanto que na TMA o sítio alvo é mais aberto e, em alguns pontos do sítio, há maior distância entre os resíduos da proteína e os ligantes. Porém, os valores das somas de energias dos três resíduos especificados (W58, W59, Y62) são mais próximos entre

si no docking com a PPA, se comparados aos valores obtidos no docking com a TMA. A maior variação dos valores de soma dos resíduos no docking com a TMA (entre -11 e -77 kcal/mol) indica maior variação de afinidade com essa enzima, enquanto que os valores obtidos no docking com a PPA (entre -55 e -66 kcal/mol) indicam afinidade muito parecida entre a proteína e os vários ligantes na simulação. Alguns ligantes desenhados (como o T3 e T3B) apresentaram diferenças maiores que 10 kcal/mol na soma das energias dos três resíduos citados, comparando-se as simulações com a TMA e a PPA, indicando maior afinidade com os três resíduos da TMA. Também foi observado durante a realização do docking que os ligantes apresentaram “poses” com posição menos coerente, provavelmente devido aos resíduos vizinhos ao sítio ligante que podem produzir algum impedimento no acesso ao sítio ligante.

| Ligantes            | Energia de interação (kcal/mol) entre os ligantes e os resíduos da PPA (1H4D.pdb) |        |        |        | MolDock score | Protein | Hbond  |
|---------------------|-----------------------------------------------------------------------------------|--------|--------|--------|---------------|---------|--------|
|                     | W59                                                                               | Y62    | Soma   |        |               |         |        |
| maltose             | -8,01                                                                             | -27,48 | -22,17 | -57,66 | -133,71       | -116,99 | -9,34  |
| acarbose            | -7,63                                                                             | -27,07 | -27,50 | -62,20 | -190,46       | -182,25 | -19,15 |
| acarviosina-glicose | -8,23                                                                             | -28,90 | -27,97 | -65,10 | -280,10       | -268,87 | -28,35 |
| isoacarbose         | -7,83                                                                             | -29,27 | -18,26 | -55,36 | -181,88       | -173,97 | -15,57 |
| A1                  | -2,30                                                                             | -48,28 | -5,96  | -56,55 | -167,29       | -125,82 | -11,91 |
| malt002             | -8,43                                                                             | -28,09 | -22,62 | -59,14 | -137,89       | -120,99 | -14,54 |
| trip005             | -5,43                                                                             | -37,08 | -19,46 | -61,97 | -91,29        | -107,35 | -5,44  |
| T3                  | -2,92                                                                             | -33,52 | -3,33  | -39,77 | -205,10       | -205,78 | -9,72  |
| T3B                 | -28,90                                                                            | -8,95  | -27,97 | -65,82 | -220,23       | -190,75 | -16,42 |

Apesar dos valores maiores de “MolDock score” e “Protein” no docking com a PPA, a contribuição da soma de energias dos 3 resíduos considerados é proporcionalmente menor. Os cinco ligantes (A1, malt002, trip005, T3, T3B) mostrados na tabela são aqueles com melhores resultados dentre os ligantes de mesma série. São também mostrados os ligantes com propriedades inibitórias conhecidas (acarbose, acarviosina-glicose e isoacarbose) e a maltose, fragmento do substrato natural. Observa-se que no docking com a PPA todos os ligantes apresentaram valores mais aproximados na soma de energias dos resíduos especificados, do que no docking com a TMA, onde há uma maior diferença nos valores, indicando que alguns dos ligantes desenhados apresentam melhor afinidade teórica com esses resíduos. Não apenas os valores numéricos foram considerados para avaliação de cada ligante como também foi

considerada a coerência das posições adotadas pelas “poses” no sítio ligante, pois muitas vezes um ligante possui várias “poses” cujas posições são muito divergentes, sugerindo pouca estabilidade da molécula no sítio. O ligante T3B, por exemplo, apesar dos valores de “MolDock score” e “Protein” serem melhores para a PPA, apresentou  
5 melhores “poses” no docking com a TMA (Figura 10A e 10B). Vale acrescentar, que o mesmo comportamento se repetiu nos docking com os outros ligantes devido às diferenças nos arredores do sítio alvo das enzimas ilustradas na Figura 4. Este é um ponto importante a se avaliar no docking já que o MolDock é um algoritmo baseado na evolução dos ligantes que elimina a cada rodada os piores candidatos, ou seja, no final  
10 do docking restam apenas as “poses” com melhores energias de interação e melhor posicionamento no sítio indicado no programa. Nos dockings com inibidores conhecidos, observou-se que a maior parte das “poses” adotavam posição e conformação muito similares entre si no sítio ligante, por isso esse critério de avaliação foi adotado para a seleção de ligantes.

15 As estruturas desenhadas apresentam propostas de novos candidatos a fármacos contra a ação de insetos que atacam estoques de grãos. Apresentam também uma alternativa aos atuais inibidores de alfa-amilases que são obtidos através do cultivo microbiano e então purificados. Sabendo-se que a produção por meio de síntese química é mais barata do que a aplicada por biotecnologia, o presente trabalho oferece produtos  
20 com potencial alternativa para diminuição de custos na produção de inibidores alfa-amilases.

#### Exemplo 5 – *Análise de parâmetros de Absorção, Distribuição, Metabolismo, Excreção e Toxicidade (ADMET)*

Considerando que os compostos serão para uso agrícola e seus resíduos poderão  
25 ser consumidos por humanos e animais, em sua aplicação deve-se evitar o contato com os alimentos já que possuem potencial atividade contra as enzimas de mamíferos. Isso pode ser feito com o uso de iscas de depósitos isolados do inseticida. Também é fundamental que o desenho de novos compostos seja analisado em relação ao potencial de risco de produção de efeitos indesejáveis. Pensando nisso, foi feita predição de  
30 parâmetros moleculares que influenciam na Absorção, Distribuição, Metabolismo,

Excreção e Toxicidade (ADMET) utilizando-se o Osiris Property Explorer (Thomas Sander, Actelion Pharmaceuticals Ltd., <http://www.organic-chemistry.org/prog/peo/>).

Essa ferramenta permite analisar cLogP, solubilidade em água, peso molecular e riscos de toxicidade; além disso, fornece dois parâmetros: Druglikeness e Drug-score. O

5 Druglikeness baseia-se na comparação do composto desenhado com compostos de base de dados de compostos comercializados e de base de dados de compostos sem propriedades farmacológicas (Fluka). O Drug-score fornece um valor geral que considera todos os parâmetros anteriores. Os resultados da análise constam na Tabela 6. Como não é interessante que os compostos sejam absorvidos pelo trato gastrointestinal  
10 dos consumidores, valores menores de cLogP são desejáveis. O Osiris fornece dados de toxicidade animal tais como risco de mutagenicidade, tumorigenicidade, efeito reprodutivo e irritação baseado em fragmentos moleculares com toxicidades conhecidas armazenadas em base de dados (RTECS).

**Tabela 6** – Resultados da análise de ADMET, feita pelo programa Osiris Property

15 Explorer (<http://www.organic-chemistry.org/prog/peo/>) com os dados de risco de toxicidade tais como mutagenicidade, tumorigenicidade, irritabilidade e eventuais efeitos colaterais reprodutivos, baseados em fragmentos de moléculas conhecidas depositados no banco de dados RTECS. Este programa fornece um valor de LogP, o coeficiente de partição entre água e n-octanol, que mede a hidrofobicidade (quanto maior o valor,  
20 maior a solubilidade do composto em lipídeos). No caso das plantas e tratando-se de um composto de uso superficial (não precisa ser absorvido para agir), ter um LogP baixo pode ser algo positivo pois será menos absorvido e menos distribuído em tecidos animais. O programa também fornece um valor de LogS (solubilidade), parâmetro que descreve tanto absorção e distribuição, cujo valor deve ser maior que -4, baseado no  
25 LogS médio dos medicamentos comercializados. E, por fim, o programa ainda fornece o Drug-score que é um valor geral resultante da combinação de todos os parâmetros anteriores mais o Druglikeness, que compara fragmentos do composto desenhado com base de dados de compostos comercializados e com base de dados de compostos que não servem como fármacos (Fluka). É desejável que tanto Druglikeness quanto o Drug-  
30 score sejam positivos.

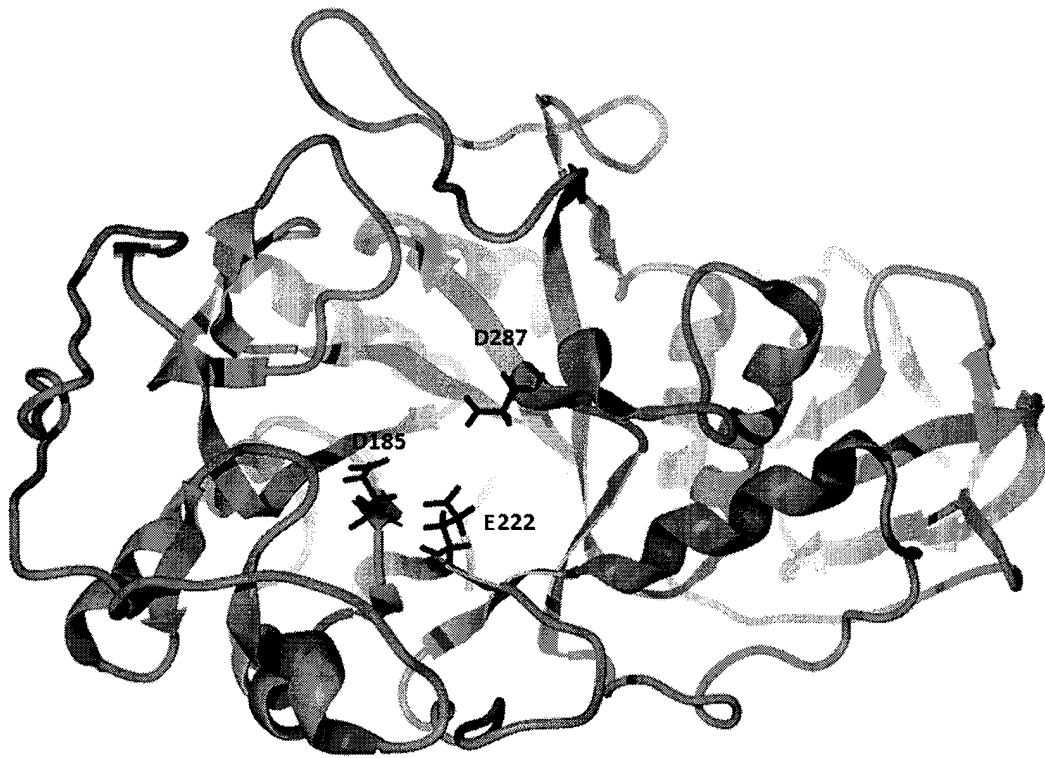
| Resultados de ADMET |           |       |              |              |            |
|---------------------|-----------|-------|--------------|--------------|------------|
| ligante             | efeito    | cLogP | solubilidade | Druglikeness | Drug-score |
| A1                  | irritante | -4,44 | -0,31        | -5,78        | 0,17       |
| malt002             | n/a       | -3,88 | -0,01        | -3,1         | 0,49       |
| trip005             | n/a       | 1,71  | -2,21        | 1,91         | 0,87       |
| T3                  | irritante | -4,92 | -1,39        | -0,48        | 0,21       |
| T3B                 | irritante | -3,44 | -3,75        | -0,73        | 0,18       |

Como pode ser observado na Tabela 6, nenhum dos 4 melhores ligantes desenhados apresentaram risco de efeito grave na análise do ADMET. Isso provavelmente se deve às características dos grupos “Rs” empregados e núcleos desenhados, que se baseiam em estruturas conhecidamente inócuas ao organismo humano. O efeito irritante provavelmente deve-se ao grande número de átomos de Oxigênio presentes nas estruturas (hidroxilas, principalmente). As estruturas que não apresentam efeito irritante são muito menores do que as outras e possuem, portanto, menos átomos de Oxigênio. No entanto, os átomos de Oxigênio em grande quantidade são necessários para realizar as várias interações de hidrogênio com os resíduos do sítio ligante e do sítio catalítico de forma a mimetizar as interações feitas pelo substrato natural e os inibidores existentes, que também possuem estrutura polissacarídica.

### REIVINDICAÇÕES

1. Método de obtenção de novos compostos com potencial função inibitória das enzimas alfa-amilases caracterizado pelas seguintes etapas: (i) obter arquivos de coordenadas referentes às estruturas tridimensionais da estrutura da alfa-amilase de interesse; (ii) fazer busca por sequências homólogas à estrutura primária da enzima de interesse; (iii) alinhar a enzima de interesse às estruturas de outras alfa-amilases, analisar quanto à sobreposição das estruturas terciárias e analisar os desvios; (iv) fazer alinhamentos entre a estrutura primária de proteínas homólogas à proteína de interesse e incluir proteínas de outros organismos; (v) evidenciar as similaridades e diferenças entre estes dois conjuntos de proteínas e buscar as correspondências no alinhamento de estrutura primária dos resíduos importantes; (vi) buscar estruturas de compostos por similaridade maior que 90% com o ligante natural (amilose) e proceder simulações de docking; (vii) analisar o nano-ambiente do sítio-alvo e determinar o conjunto de parâmetros físico-químicos e estruturais que o descrevem, composto pelos resíduos nos quais se espera que os ligantes atraiam; (viii) modificar as estruturas selecionadas no item (vi) de acordo com as características físico-químicas e estruturais encontradas no item (vii) a fim de aumentar os valores de afinidade de ligação; (ix) adicionalmente, fazer busca por ligantes baseados na estrutura dos inibidores naturais aplicando-se os passos de docking e modificações estruturais a esses ligantes; e (x) realizar predição de parâmetros moleculares que interferem na Absorção, Distribuição, Metabolismo, Excreção e Toxicidade.

2. Método de controle de insetos que utilizam as enzimas alfa-amilases para obter glicose, caracterizado pelo fato de utilizar compostos desenvolvidos pela inibição ou inativação da enzima alfa-amilase por meio de ligação parcial ou total de inibidor, sendo o inibidor obtido de acordo com a reivindicação 1.



**Figura 1**

### Figura 2

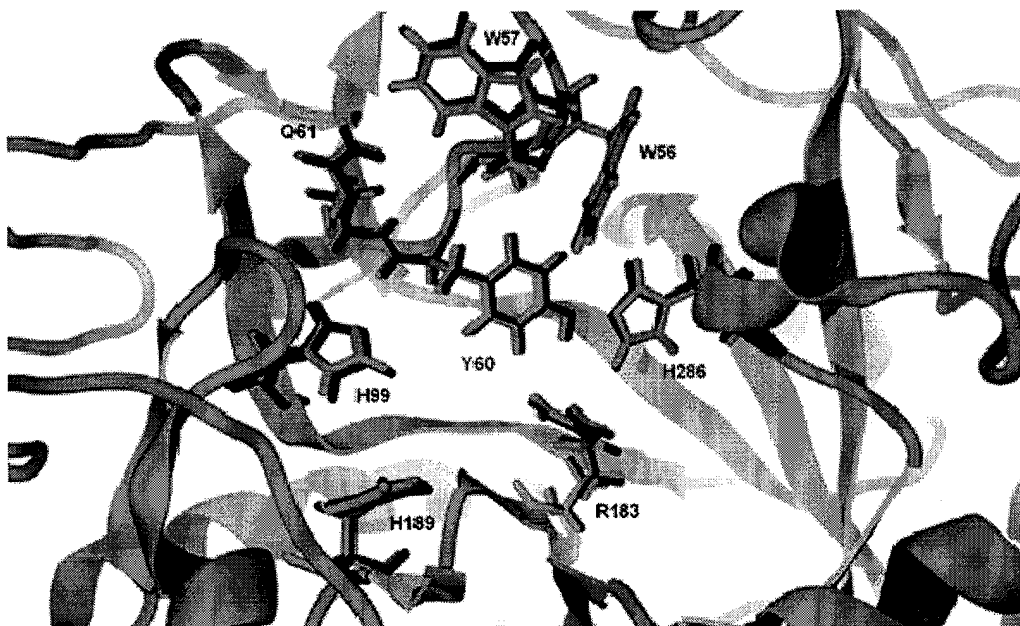


Figura 3A

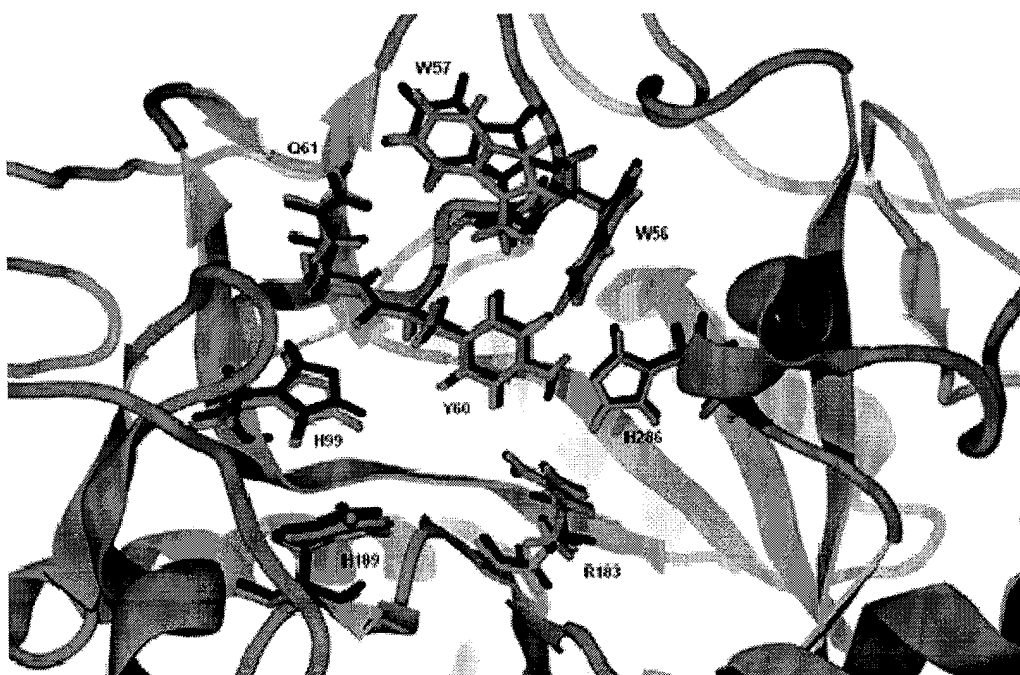


Figura 3B

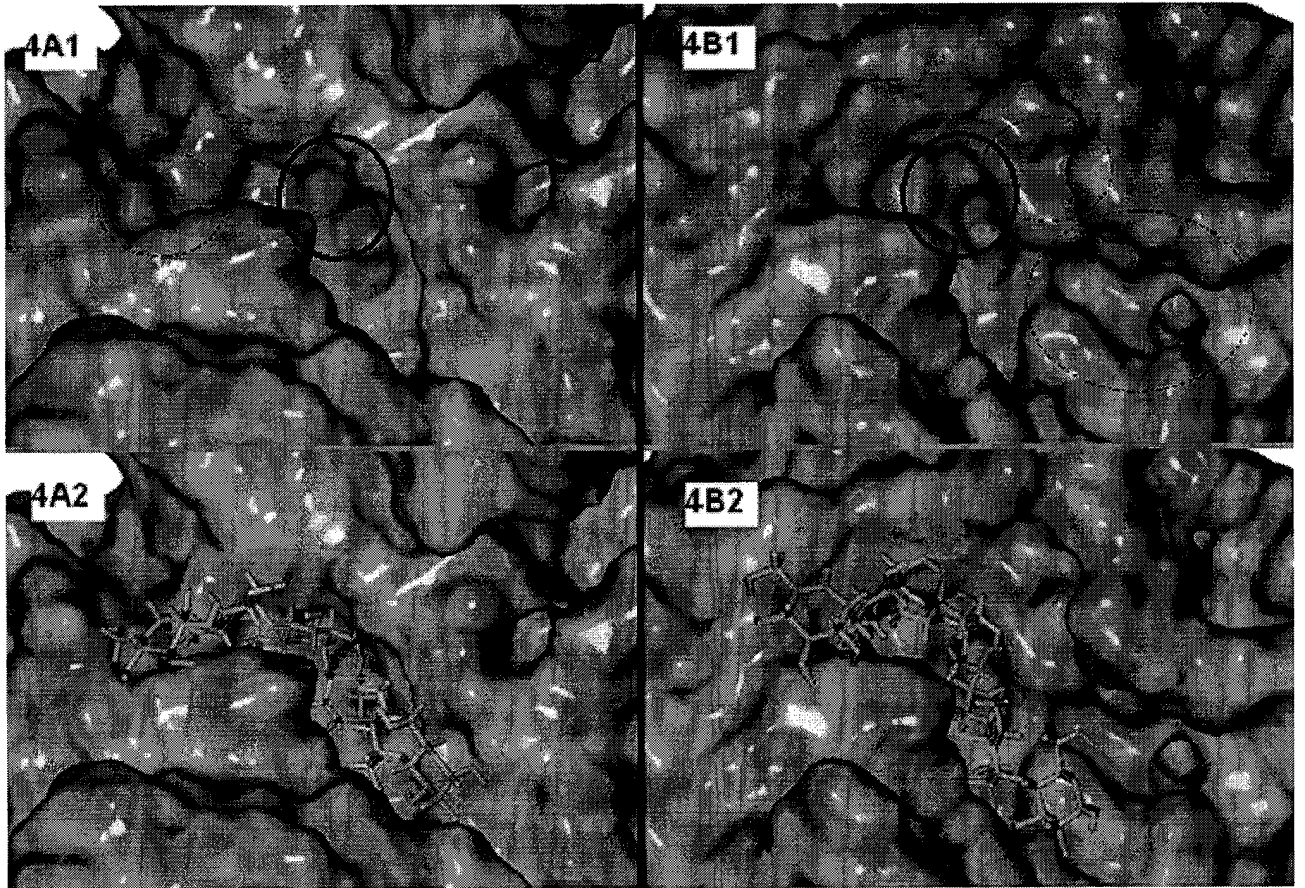
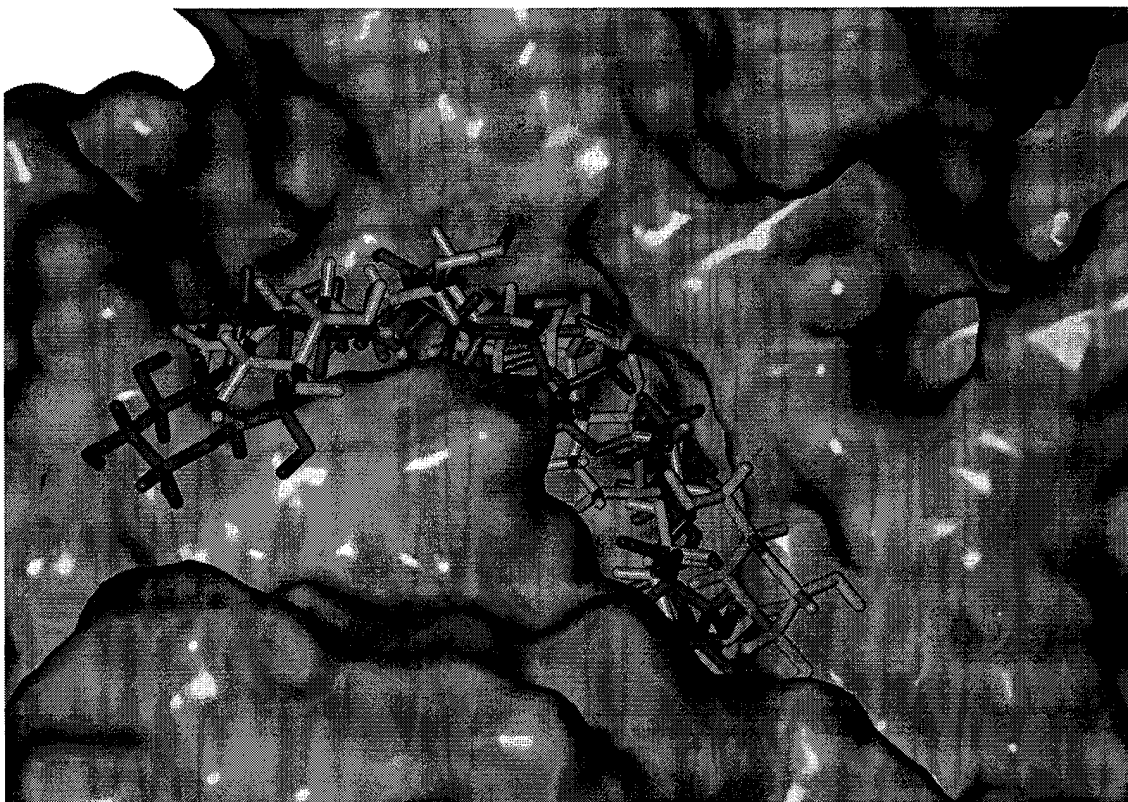
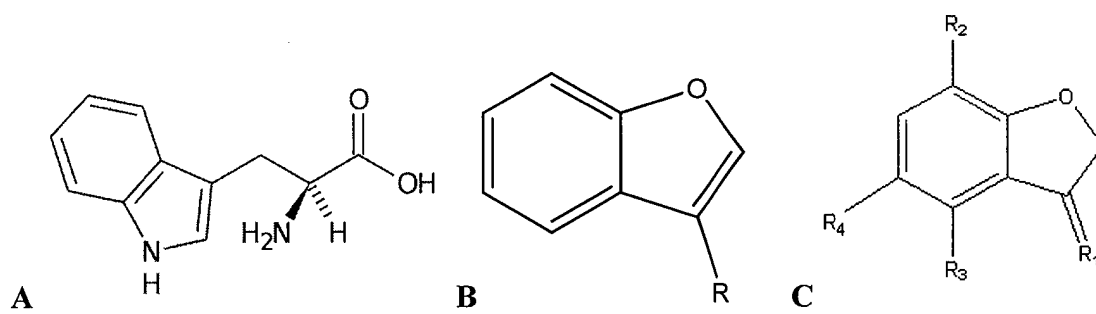
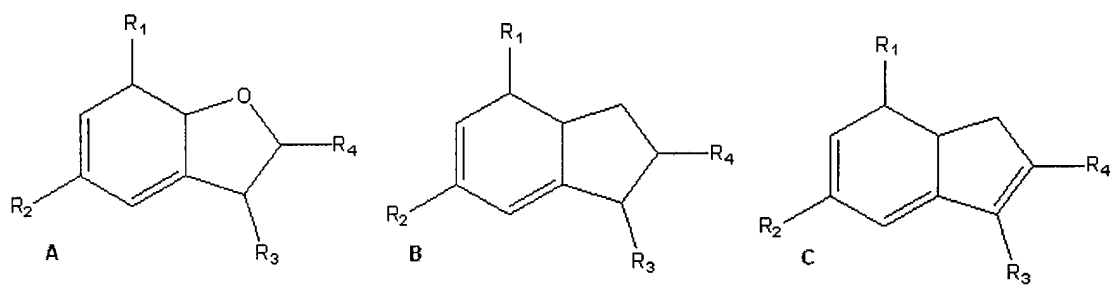
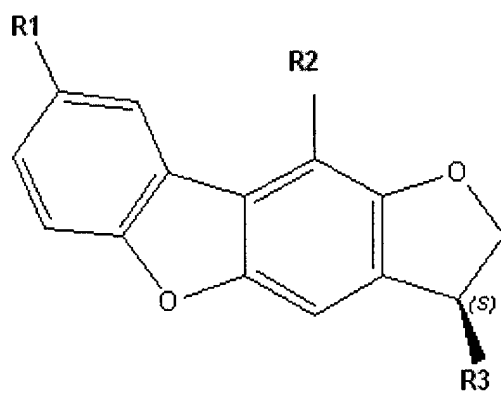


Figura 4

**Figura 5**

5

**Figura 6**

**Figura 7****Figura 8**

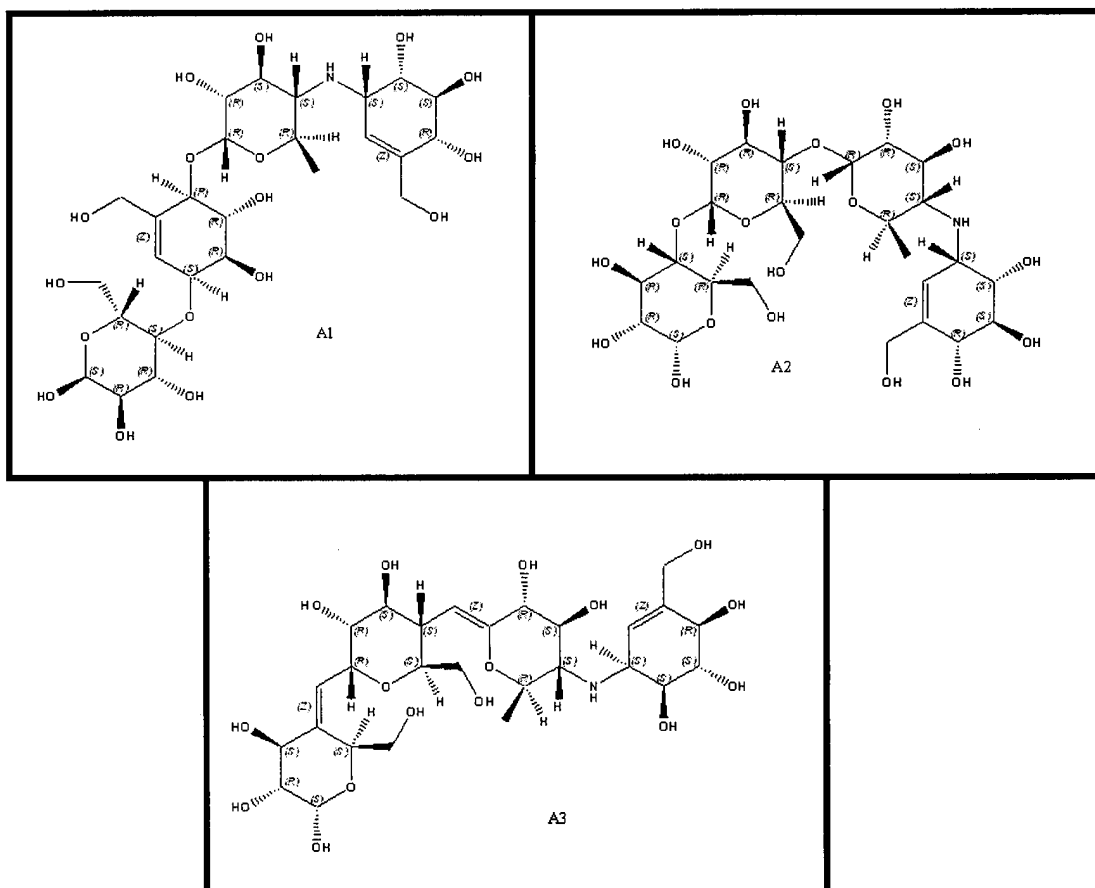
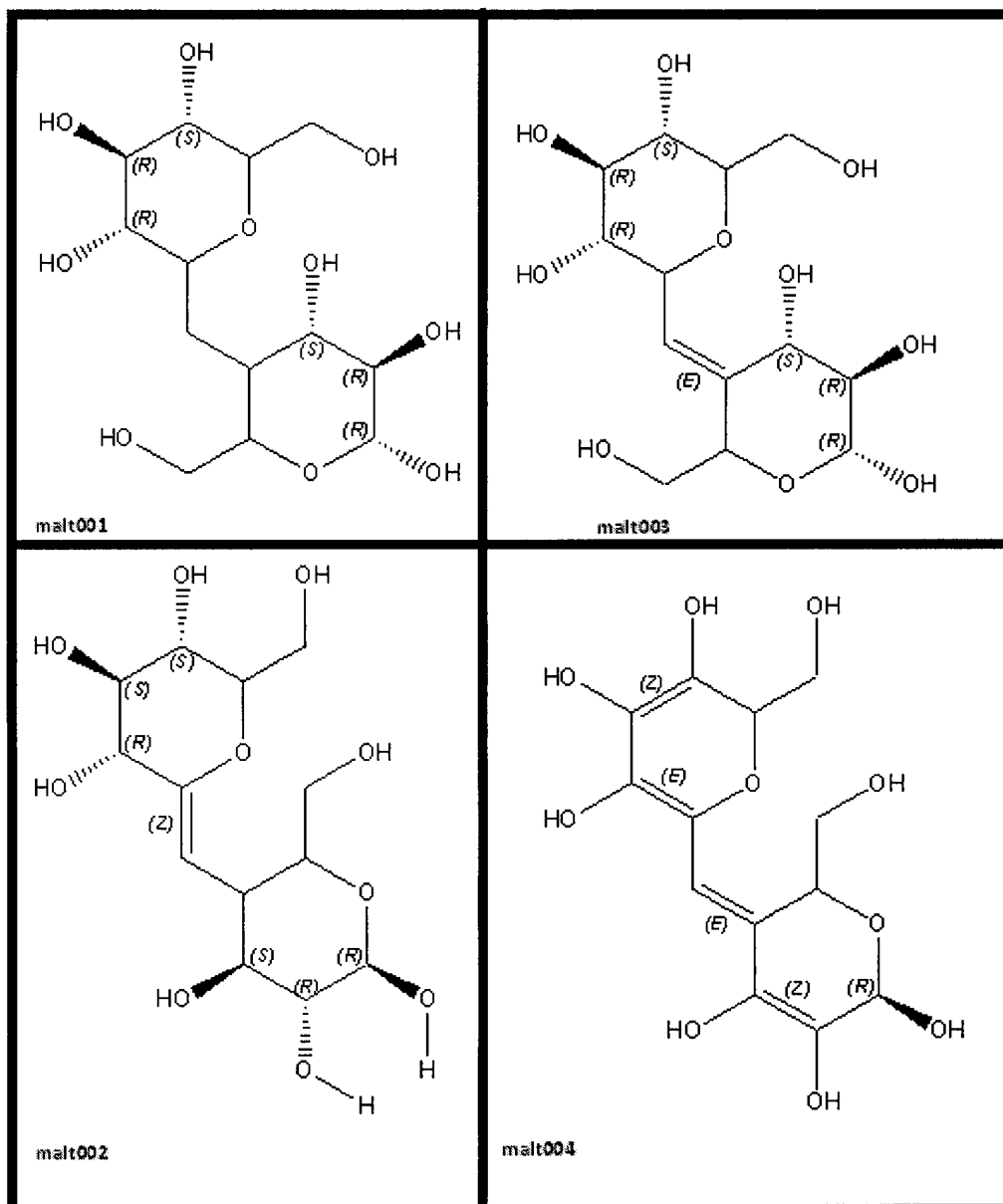


Figura 9.1



9.

Figura 9.2

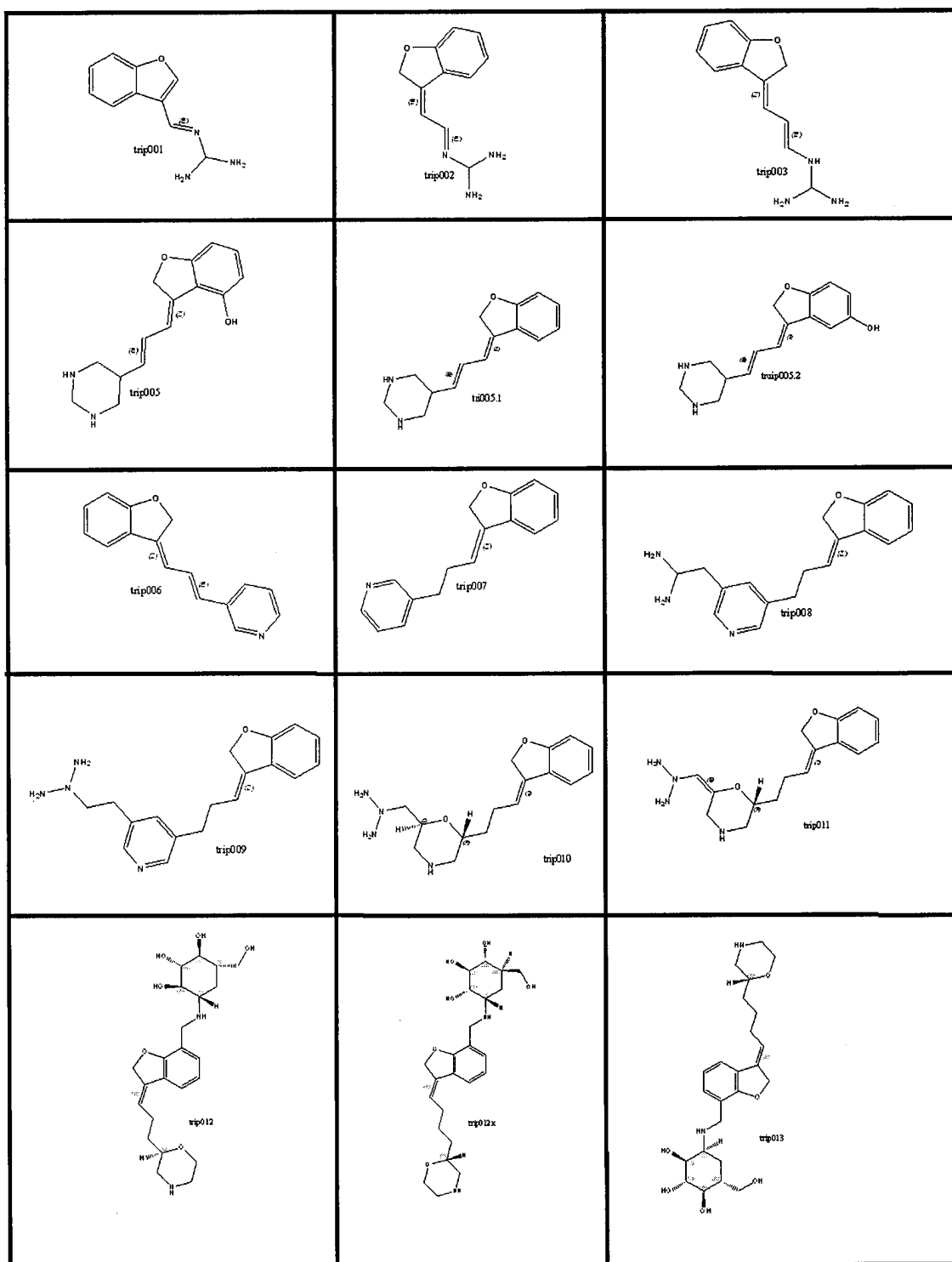


Figura 9.3

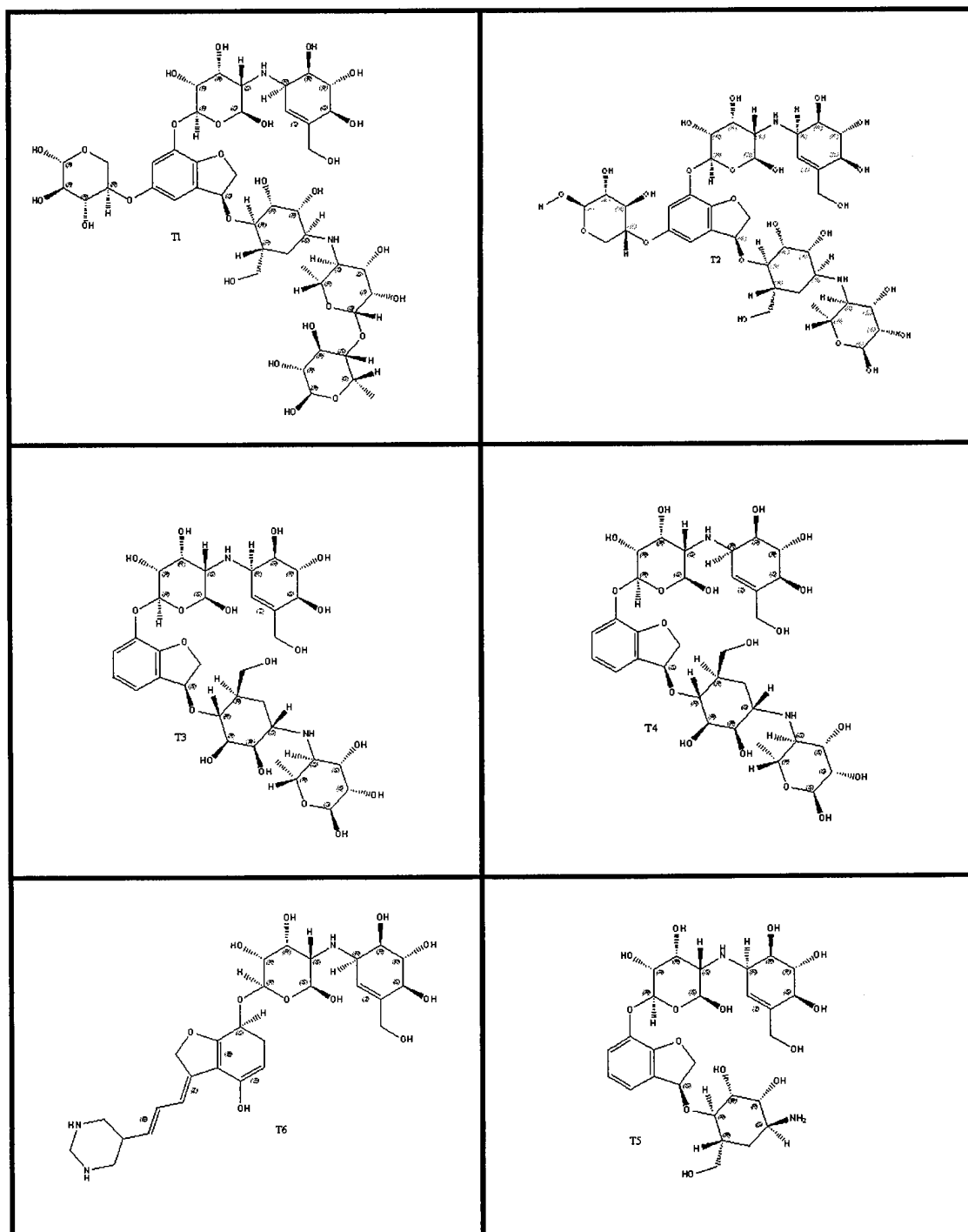


Figura 9.4

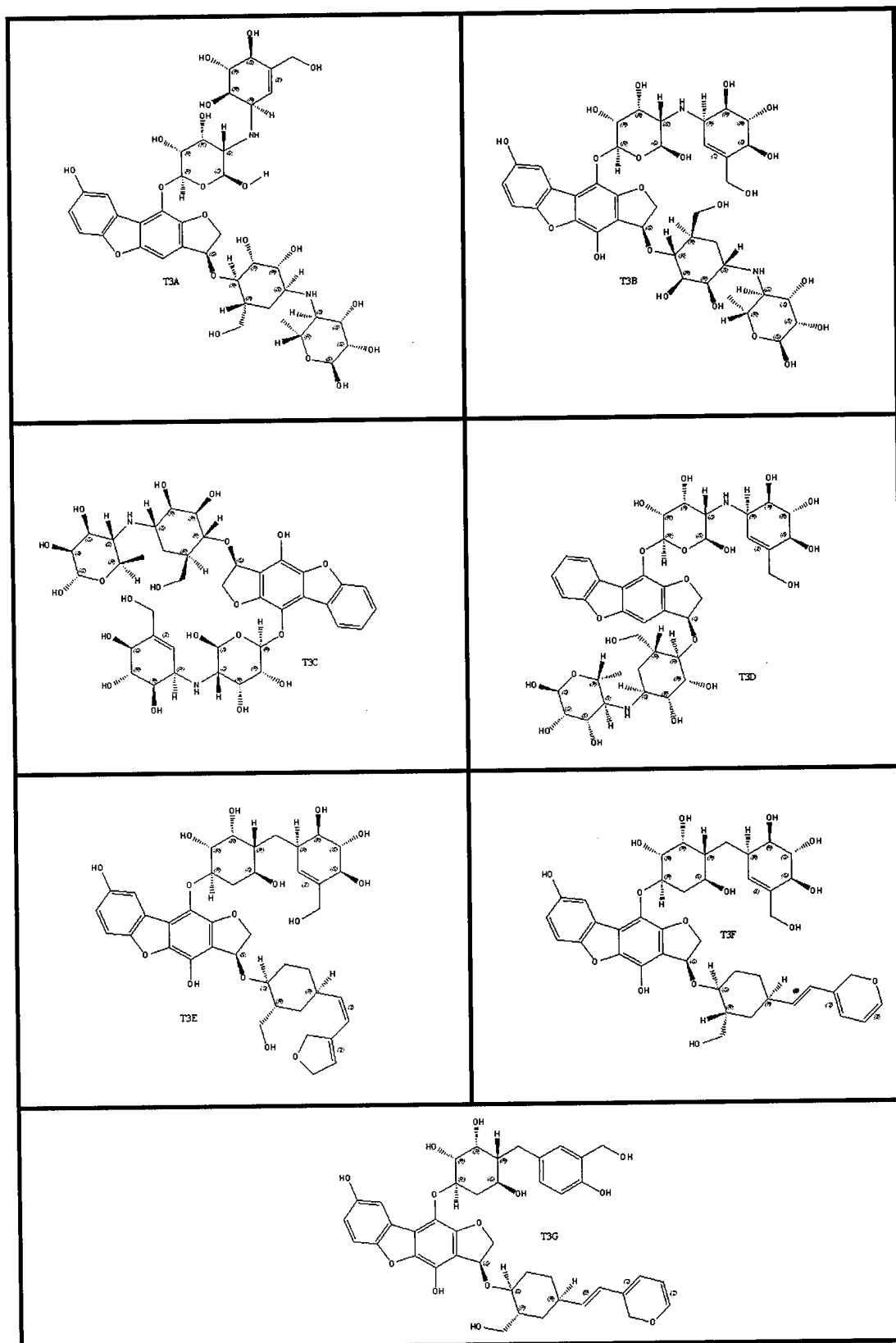


Figura 9.5

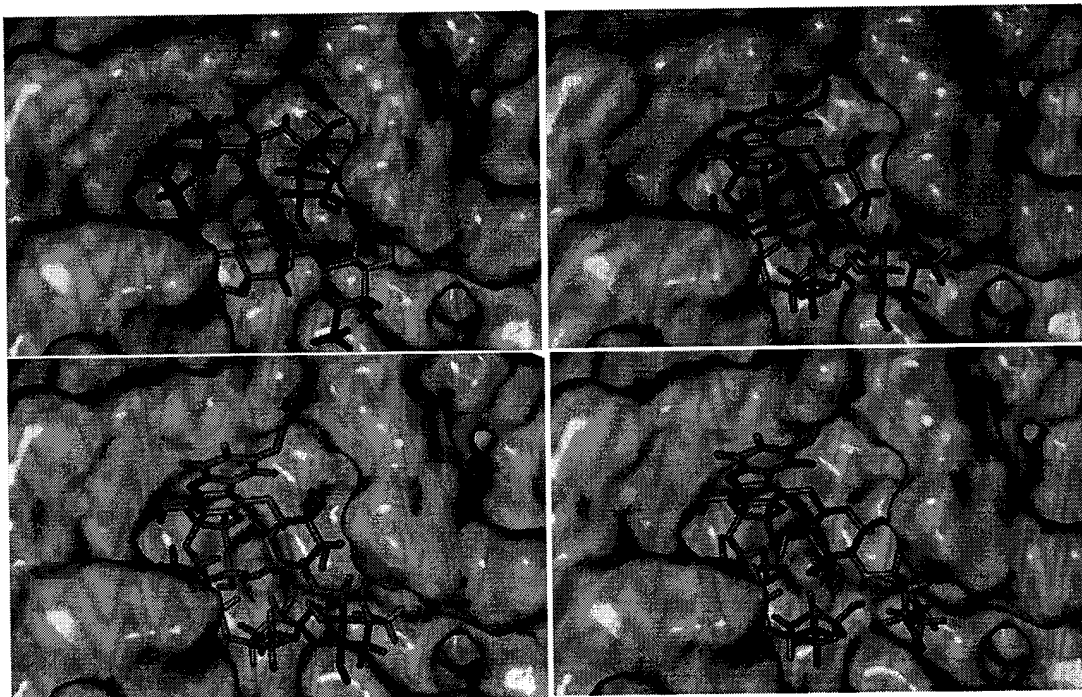


Figura 9A

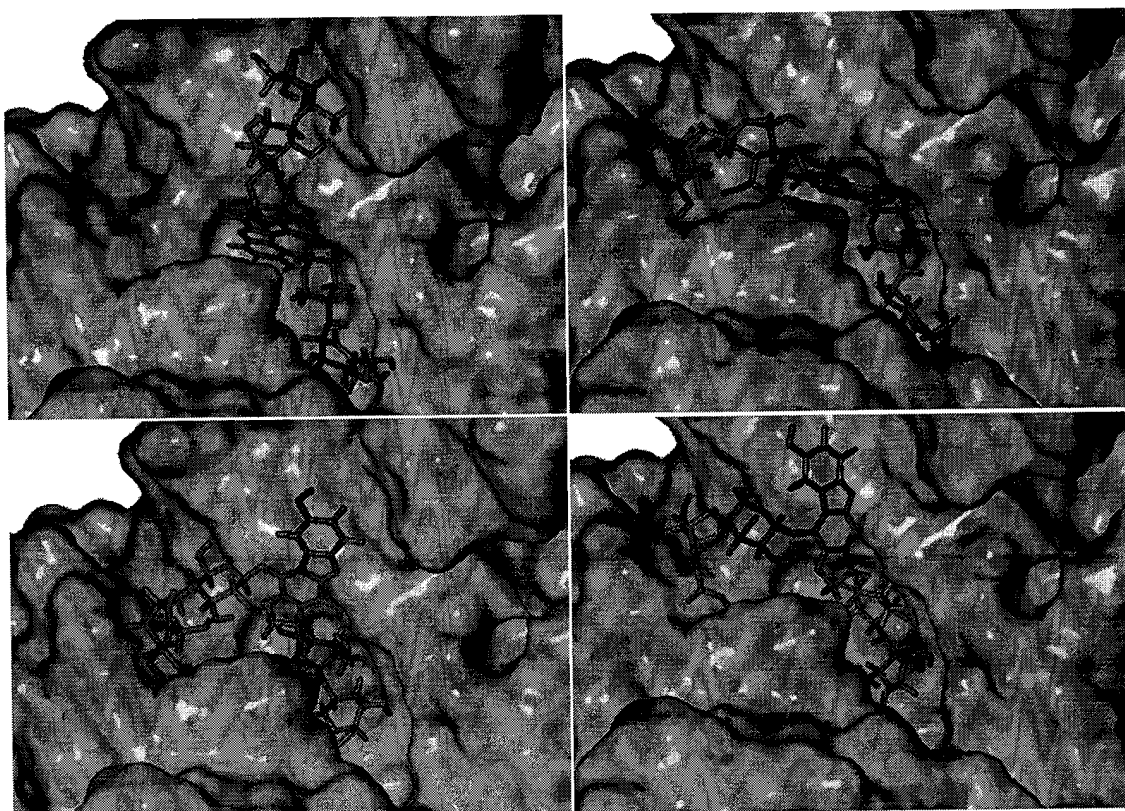


Figura 10B