

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE FARMÁCIA

*Planejamento in silico de novos inibidores da enzima Guanina
Fosforibosil Transferase de Giardia lamblia como possíveis
candidatos a fármacos antiparasitários*

Estudante: Gustavo Machado das Neves

Orientador: Prof. Dr. Daniel Fábio Kawano

Porto Alegre

2015

SUMÁRIO

- Lista de figuras.....	i
-Lista de Abreviaturas.....	ii
-Resumo.....	iii
1. Introdução.....	1
2. Objetivo.....	8
3. Planejamento Experimental.....	9
4. Material e Métodos.....	9
5. Resultados e Discussão.....	21
6. Conclusões.....	50
7. Referências.....	52

Lista de Figuras:

Figura 1. Estruturas químicas dos principais fármacos empregados no tratamento da giardíase.....	3
Figura 2. Conversão da guanosina em guanosina monofosfato <i>Giardia lamblia</i> através das enzimas guanosina hidrolase e guanina fosforibosil transferase (GPRTase).....	6
Figura 3. Comparação das interações realizadas entre um inibidor do estado de transição, Imucilina G fosfato, e os resíduos de aminoácido do sítio ativo da GPRTase de <i>G. lamblia</i> (A) e HGPRTase humana (B).....	7
Figura 4. Estruturas químicas dos inibidores da GPRTase de <i>G. lamblia</i>	8
Figura 5. Principais interações hidrofóbicas entre o ligante ImuGP e a enzima GPRTase.....	22
Figura 6. Visão lateral da ImuGP realizando interações hidrofóbicas e aromáticas com a enzima GPRTase.....	22
Figura 7. Principais interações por ligações de hidrogênio entre ImuGP e GPRTase.....	23
Figura 8. Avaliação da distância entre ImuGP e o resíduo Lys152 (GPRTase) e Lys165 (HGPRTase).....	24
Figura 9. Avaliação da distância entre águas HOH541 e HOH642, ImuGP e Lys152(GPRTase) e Lys165 (HGPRTase).....	24
Figura 10. Predição das 7 cavidades através do MetaPocket 2.0.....	26
Figura 11. Composição dos aminoácidos que compõem o sítio de ligação da GPRTase preditos pelo MPK1...26	26
Figura 12. Pontos farmacofóricos do primeiro modelo adotado com base na estrutura da ImuGP.....	28
Figura 13. Construção do modelo farmacofórico no ZincPharmer e obtenção de 100.211 hits.....	29
Figura 14. Modelo farmacofórico adotado para compilação das estruturas da ZINC.....	30
Figura 15. Análise das poses obtidas pelo processo de redocagem no Glide.....	31
Figura 16. Análise do procedimento de redocagem no GOLD utilizando as funções de escore (Goldscore, ChemScore, ASP e ChemPLP).....	32
Figura 17. Análise visual das poses dos ligantes ativos realizada pela docagem no Glide.....	34
Figura 18. Análise visual das poses obtidas por diferentes métodos de escore no GOLD.....	36
Figura 19. Análise da pose dos ligantes ativos na docagem realizada pelo Glide.....	38
Figura 20. Análise dos melhores ligantes da ZINC obtidos pelo Glide.....	38
Figura 21. Análise visual dos ligantes ativos pela função ASP e pela função GoldScore.....	39
Figura 22. Avaliação da pose dos ligantes da ZINC pela função ASP e pela função GoldScore.....	40
Figura 23. Análise visual das poses dos 4 ligantes da ZINC melhores ranqueados pela função ASP e as proximidades do aminoácido Trp180.....	41
Figura 24. Análise visual das poses dos 5 melhores ligantes obtidos por docagem consensual.....	43
Figura 25. Inspeção visual das melhores poses encontradas pela função GoldScore.....	44
Figura 26. Análise detalhada da pose do ligante ZINC77269479.....	45
Figura 27. Análise da pose apresentada pelo ligante ZINC77269479 e interações com o sítio da GPRTase.....	46
Figura 28. Análise das interações entre os principais ligantes e os resíduos de aminoácido da GPRTase.....	46

Lista de Abreviaturas na ordem que aparecem:

OMS: Organização mundial da Saúde

AIDS: Síndrome da Imunodeficiência Adquirida (*Acquired Immune deficiency Syndrome*)

DNA: Ácido desoxirribonucleico (*Desoxyribonucleic acid*)

RNA: Ácido ribonucleico (*Ribonucleic acid*)

PFOR: Piruvato-ferredoxina oxidorreductase

NADH: Dinucleotídeo de nicotinamida e adenina (*Nicotinamide adenine nucleotide*)

GPRTase: Guanina Fosforibosil Transferase (*Guanine Phosphoribosyltransferase*)

HGPRTase: Hipoxantina Guanina Fosforibosil Transferase (*Hypoxanthine Guanine Phosphoribosyltransferase*)

ImuGP: Imucilina G Fosfato

PDB: *Protein Data Bank*

1DQN: Complexo cristalográfico da GPRTase com ImuGP

1BZY: Complexo cristalográfico da HGPRTase humana com ImuGP

DSV: *Discovery Studio Visualizer*

SD: *Structure Data*

HOH: água

RMSD: Desvio da raiz quadrática média (*Root-mean-square deviation*)

OPLS: *Optimized Potentials for Liquid Simulations*

SP: Precisão Padrão (*Standard Precision*)

GS: GoldScore

CS: ChemScore

ASP: *Astex Statistical Potential*

ChemPLP: *Piecewise Linear Potential*

EF: Fator de enriquecimento (*Enrichment Factor*)

GA: Algoritmo genético (*Genetic algorithm*)

GMP: Guanina monofosfato

RESUMO

Introdução: A OMS estima que as infecções causadas por *G. lamblia* sejam a doença protozoária intestinal mais frequentemente relatada em todo o mundo. No entanto, um problema comum aos medicamentos usados no tratamento da giardíase diz respeito às linhagens resistentes. Sendo assim, buscam-se novos fármacos capazes de combaterem as infecções causadas por *G. lamblia*, através da utilização de enzimas validadas e específicas do parasita. A enzima guanina fosforibosil transferase (GPRTase) é essencial para a produção de nucleotídeos no parasita, uma vez que os parasitas, diferentes dos humanos, são incapazes da síntese *de novo* de nucleotídeos. **Objetivos:** Identificar novos inibidores da enzima GPRTase de *Giardia lamblia* que possuam eficácia e perfil farmacocinético adequado para serem usados no tratamento da giardíase. **Material e métodos:** Primeiramente, estudou-se o sítio de ligação das enzimas GPRTase de *G. lamblia* e de sua homóloga humana, a hipoxantina-guanina fosforibosil transferase (HGPRTase). Em seguida, construiu-se um modelo farmacofórico para se triar estruturas químicas no banco de dados da ZINC que possuam complementaridade com o sítio da GPRTase. Ademais, realizou-se o processo de docagem molecular dessas estruturas utilizando o complexo cristalográfico da GPRTase em complexo com o inibidor do estado de transição, da enzima (PDB ID: 1DQN), por meio da utilização dos programas Glide e Gold, abrangendo as etapas de validação, docagem e docagem consensual. Por fim, realizou-se a predição das propriedades farmacocinéticas dos melhores ligantes através do programa Volsurf+. **Resultados e Discussão:** O estudo do sítio de ligação da enzima GPRTase demonstrou que as interações com os resíduos Trp180, Leu181, Asp187 parecem ser importantes para a especificidade da enzima. Além disso, a possibilidade de interação com Lys152 poderia aumentar a seletividade frente à GPRTase. O modelo farmacofórico de quatro pontos resultou na obtenção de 1230 estruturas do banco de dados da ZINC. A docagem consensual utilizando Glide-Glidescore e Gold-GoldScore não apresentou consenso visual das poses. Foram escolhidas 5 estruturas avaliadas pelo GoldScore para predição farmacocinética, a qual demonstrou que os ligantes parecem não serem absorvidos pelo intestino e não ultrapassam a barreira hematoencefálica. **Conclusões:** Não foi possível obter um consenso visual das poses, devido ao tamanho dos ligantes e número de ligações rotacionáveis. Tendo em vista os resultados obtidos, novos procedimentos serão realizados visando o aumento do número de poses da docagem, docagem totalmente flexível e triagem virtual com nova base de dados através de novo modelo farmacofórico.

1. INTRODUÇÃO:

Melhorias no abastecimento de água, saneamento básico e higiene têm elevado consideravelmente os níveis de saúde e a expectativa de vida em países industrializados, embora doenças como diarreia, cólera e febre tifoide ainda permaneçam como uma das principais causas globais de mortalidade infantil. Estima-se que cerca de 880 milhões de pessoas não tenham acesso a água de qualidade, ou seja, água protegida de contaminação externa, particularmente de contaminação por matéria fecal. Todos os anos, cerca de dois milhões de pessoas morrem por desidratação associada à diarreia, como resultado da falta de acesso à água de qualidade e à infraestrutura sanitária (GRAHAM, 2011). Estima-se que quase 20% do número total de mortes infantis estejam associadas à desidratação causada por diarreia (GRAHAM, 2011) e, mesmo quando não são fatais, quadros de diarreia persistente durante a primeira infância apresentam um profundo impacto no desenvolvimento físico e mental, podendo inclusive se traduzir em prejuízos na capacidade cognitiva e na produtividade já na idade adulta (GUERRANT et al., 2002).

A diarreia é um sintoma de infecção causada por uma variedade de organismos como bactérias, vírus e protozoários. As doenças diarreicas podem ser transmitidas por água, comida ou diretamente de pessoa a pessoa via contaminação fecal-oral, sendo a sua ocorrência mais comum quando há falta de água limpa para beber, cozinhar ou para higiene básica (TULCHINSKY; VARAVIKOVA, 2009; OMS, 2010). As doenças transmitidas por água, também chamadas de doenças de veiculação hídrica, estão associadas principalmente a infecções por rotavirus, cepas patogênicas de *E. coli*, *Campylobacter jejuni* e protozoários parasitas (PALANIAPPAN et al., 2012).

Com relação às doenças de veiculação hídrica causadas por protozoários, *Giardia lamblia* e *Cryptosporidium parvum* se destacam, tendo sido responsáveis, respectivamente, por 40 e 50% dos surtos mundiais registrados entre 1954 e 2001 (KARANIS; KOURENTI; SMITH, 2007). Em função do papel preponderante destes parasitas no desenvolvimento de quadros de diarreia e de transtorno nutricional nas populações dos vários países ao redor do globo, a Organização Mundial da Saúde (OMS) reconheceu em 2004 as infecções causadas por *Giardia* e *Cryptosporidium* como parte efetiva de suas iniciativas para o combate às doenças negligenciadas (SAVIOLI; SMITH; THOMPSON, 2006).

Doenças negligenciadas são enfermidades infecciosas que, principalmente, mas não exclusivamente, afetam populações pobres dos países em desenvolvimento, onde a falta de saneamento básico e de acesso à saúde fomentam sua transmissão e a proliferação de seus vetores. Embora haja um grande interesse social no desenvolvimento de novos fármacos para

tratar ou prevenir essas doenças, poucas empresas farmacêuticas se dedicam à pesquisa e ao desenvolvimento de fármacos para atender a este mercado quando não há qualquer tipo de incentivo governamental que garanta um retorno comercial mínimo sobre seus investimentos (COHEN; DIBNER; WILSON, 2010). Neste sentido, as universidades desempenham um papel fundamental na pesquisa, no desenvolvimento e na transferência da tecnologia associada aos fármacos capazes de tratar ou prevenir doenças negligenciadas, como as doenças diarreicas causadas por protozoários parasitas (CHOKSHI, 2006).

A OMS estima que as infecções causadas por *Giardia lamblia* sejam a doença protozoária intestinal mais frequentemente relatada em todo o mundo, podendo chegar a 200 milhões de indivíduos infectados por ano. A doença se distribui geograficamente por todo o globo, dos trópicos ao ártico, e tende a ser mais comum em crianças do que em adultos (USA ENVIRONMENTAL PROTECTION AGENCY, 1999). A giardíase, doença infecciosa intestinal causada pelo protozoário flagelado *Giardia lamblia* (também chamados de *Giardia intestinalis* e *Giardia duodenalis*) se caracteriza por perda de apetite, diarreia, fezes moles ou líquidas, cólicas, náuseas e excesso de gases. Os sintomas normalmente aumentam e diminuem durante a evolução da doença e mesmo após o tratamento reinfecções podem ocorrer uma vez que a giardíase não confere imunidade. A presença de *G. lamblia* no lúmen intestinal provoca inflamação, irritação e destruição tecidual, fato que resulta no aumento da velocidade do trânsito intestinal, má absorção dos nutrientes, desidratação, perda de peso e fadiga (MCKINNEY, 2010). Nos casos da persistência crônica do parasita, deficiências na absorção de Vitamina A, B₁₂, e intolerância à lactose têm sido relatadas (CORDINGLEY; CRAWFORD, 1986; RANA; BHASIN; VINAYAK, 2005; QUIHUI-COTA et al., 2008). A giardíase pode ser fatal em pacientes com hipogamaglobulinemia e pode agravar outras doenças como AIDS e fibrose cística (MCKINNEY, 2010).

Do ponto de vista morfológico, *G. lamblia* apresenta duas formas: trofozoíto e cisto. O cisto é a forma infecciosa, normalmente ingerida junto à água contaminada. Apenas 10 cistos são necessários para se iniciar uma infestação. A transmissão por água é facilitada devido à relativa resistência do cisto à cloração, bem como à sua capacidade de sobreviver em água fresca por várias semanas. Uma vez em contato com o ácido gástrico, ocorre o excistamento do parasita, um processo mediado por cisteína proteases que resulta na liberação dos trofozoítos. Estes trofozoítos infectam o duodeno e a parte superior do intestino delgado, locais que apresentam um pH alcalino favorável à sua manutenção, e irão promover os quadros de inflamação, irritação e destruição tecidual que culminarão nos sintomas clínicos da giardíase. À medida que os trofozoítos se aproximam do cólon, a diminuição do pH inicia o

processo de encistamento, permitindo que a forma cística do parasita alcance o ambiente para a contaminação de novos indivíduos (GARDNER; HILL, 2001).

A base para o tratamento da giardíase se estabeleceu na década de 1940 com a introdução da quinacrina (**Figura 1**). Após a segunda Guerra Mundial, sua eficácia contra *G. lamblia* era superior a 90%; porém, com o aparecimento de linhagens resistentes, bem como o desenvolvimento de fármacos com um perfil de segurança mais favorável, sua produção foi interrompida nos EUA no início da década de 1990. Acredita-se que a quinacrina atue intercalando-se ao DNA (com preferência por regiões ricas em adenina ligadas a timina) inibindo assim a biossíntese de RNA e de proteínas em procariontes. A toxicidade relativa da quinacrina sobre *G. lamblia* se deve à maior velocidade de absorção do fármaco pelo parasita do que por células humanas, sendo que os casos de resistência se devem exatamente a uma redução na captação do fármaco pelo protozoário (GARDNER; HILL, 2001; MÜLLER; HEMPHILL; MÜLLER, 2011).

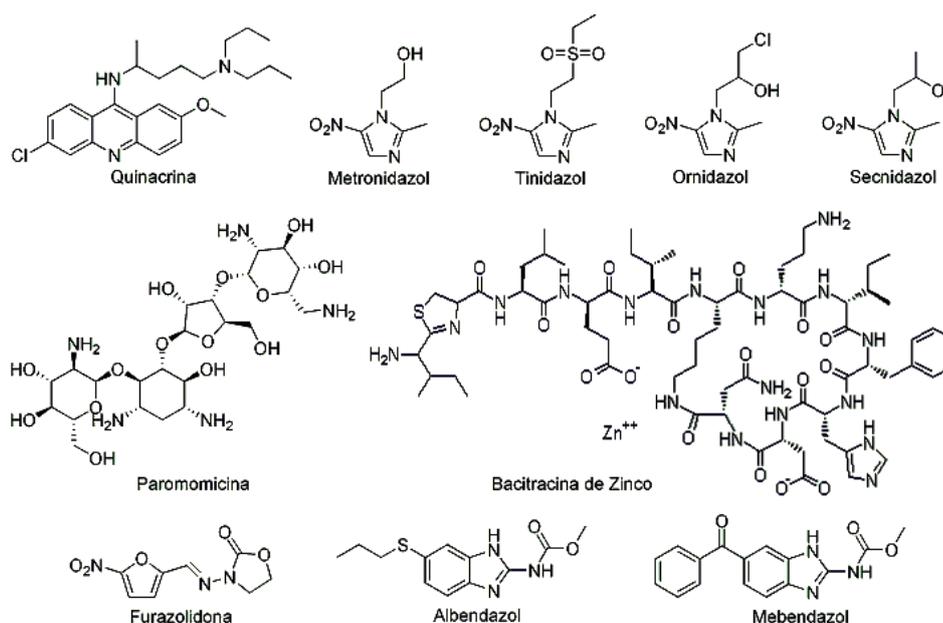


Figura 1. Estruturas químicas dos principais fármacos empregados no tratamento da giardíase.

A segunda classe de compostos a serem utilizados para o tratamento da giardíase, os nitroimidazóis, foi descoberta em 1955 e inclui fármacos como o metronidazol, tinidazol, ornidazol e secnidazol (**Figura 1**). Desde sua descoberta, metronidazol e demais nitroimidazóis são a principal ferramenta terapêutica no tratamento de infecções causadas por *G. lamblia*. De acordo com a teoria mais aceita, o metronidazol (um pró-fármaco) é reduzido a um radical nitro por elétrons provenientes da enzima piruvato-ferredoxina oxidoreductase (PFOR), enzima que não está presente em células de eucariotos superiores. Este metronidazol

reduzido atuaria como um aceptor terminal de elétrons e se ligaria covalentemente ao DNA, causando a perda da estrutura helicoidal, ruptura das hélices e subsequente morte do trofozoíto. A resistência aos nitroimidazóis correlaciona com a diminuição da atividade da PFOR, necessária para sua bioativação redutiva (GARDNER; HILL, 2001; MÜLLER; HEMPHILL; MÜLLER, 2011).

Outros fármacos utilizados para o tratamento da giardíase incluem furazolidona, benzimidazóis, paromomicina e bacitracina de zinco (**Figura 1**). A furazolidona é um dos representantes da classe dos nitrofuranos, compostos classicamente empregados como antibióticos e antimicrobianos. Como o metronidazol, o fármaco também sofre ativação redutiva no trofozoíto, porém esta redução ocorre através de uma NADH oxidase e não da PFOR. O resultado final desta bioativação também resulta em danos ao DNA e os casos de resistência estão relacionados com a diminuição da entrada do fármaco no trofozoíto ou ao aumento dos níveis de cisteína oxidases/reduzases, que são capazes de proteger o parasita contra os efeitos causados pelos radicais tóxicos (GARDNER; HILL, 2001).

Os benzimidazóis, albendazol e mebendazol, exercem seu efeito tóxico através de interação com a β -tubulina do parasita, inibindo a polimerização do seu citoesqueleto. Resistência aos benzimidazóis ocorre em função de mudanças no citoesqueleto do parasita. Já a paromomicina, um aminoglicosídeo, inibe a síntese proteica de *G. lamblia* através da ligação às subunidades 50S e 30S ribossomais, levando à leitura incorreta dos códons de RNA mensageiro. Entretanto, a suscetibilidade do parasita à paromomicina é menor do que aos nitroimidazóis, quinacrina ou furazolidona. Outro antibiótico, a bacitracina de zinco, representa uma alternativa para o tratamento de giardíase já que impede a desfosforilação de um fosfolipídio necessário para a síntese da membrana celular. Contudo, uso prolongado por via oral deste antibiótico está associado a casos de nefrotoxicidade e distúrbios gastrintestinais (GARDNER; HILL, 2001).

Um problema comum a praticamente todos os medicamentos empregados para o tratamento da giardíase diz respeito ao surgimento de linhagens resistentes. Populações de *Giardia* parecem ter mecanismos muito flexíveis para responder às alterações do meio ambiente, tais como a pressão seletiva imposta por fármacos ou à presença de anticorpos. Conforme já mencionado anteriormente, os mecanismos relacionados à resistência variam de fármaco para fármaco, mas em geral estão associados a mutações no alvo terapêutico ou à redução da captação dos fármacos pelos trofozoítos que, em ambos os casos, podem conduzir à ineficácia terapêutica (MÜLLER; HEMPHILL; MÜLLER, 2011). Pacientes que são refratários à monoterapia podem ser tratados apenas pela associação de dois ou mais fármacos, como a combinação de metronidazol com quinacrina (NASH et al., 2001).

Neste contexto, a busca por novos fármacos capazes de debelar infecções causadas por *G. lamblia* deve ser um processo contínuo, de modo à sempre permitir que o arsenal terapêutico disponível seja suficiente para tratar mesmo os pacientes intrinsecamente mais difíceis, como os imunocomprometidos (NASH et al., 2001). Idealmente, os candidatos a fármacos antiparasitários devem se ligar a uma enzima ou receptor específico do parasita (ou seja, que não possua uma proteína correspondente no hospedeiro), e que esteja relacionada a funções celulares vitais à sua sobrevivência. Este último aspecto é crucial para a viabilidade do processo pois estima-se que menos de 0,02% dos candidatos a fármacos submetidos a testes pré-clínicos cheguem a ser comercializados. Com recursos financeiros tão limitados para o desenvolvimento de fármacos nesta área, a confirmação (validação) da importância terapêutica destes alvos já no início do processo permite que os recursos disponíveis sejam empregados de modo mais produtivo. Talvez a forma mais conclusiva de se validar um alvo terapêutico seja através da produção de parasitas *knockout*, nos quais a expressão da proteína de interesse é suprimida para que se possa avaliar a verdadeira importância do alvo na viabilidade do organismo (JACKSON; PHILLIPS, 2002).

Dentre os possíveis alvos que podem ser empregados para o tratamento das infecções causadas por *G. lamblia*, a enzima guanina fosforibosil transferase (GPRTase) chamou a nossa atenção. Trata-se de uma enzima essencial para a produção de nucleotídeos no parasita e, portanto, vital para a síntese do DNA e RNA necessários à sua replicação e sobrevivência. A construção de trofozoítos *knockouts* para esta enzima pela técnica de RNA antisense mostrou que os parasitas resultantes apresentaram uma redução na capacidade de replicação de 80 a 90%, destacando a possibilidade de seu uso como alvo no desenvolvimento de novos fármacos anti-giardíase (MUNAGALA; WANG, 2002).

O papel central que a GPRTase representa na sobrevivência de *G. lamblia* se deve ao fato de que, ao contrário do que ocorre nos seres humanos, os protozoários parasitas são incapazes da síntese *de novo* de nucleotídeos purínicos e, portanto, dependem exclusivamente de mecanismos de captação e interconversão de purinas para sua sobrevivência. No caso específico da *Giardia*, o parasita apresenta uma especificidade absoluta por guanina (não sendo capaz de utilizar hipoxantina, inosina ou xantina para a síntese de nucleotídeos), a qual deve ser obtida a partir da hidrólise de nucleosídeos captados do ambiente (MEHLHORN, 2001). O reaproveitamento das bases nitrogenadas se dá através da hidrólise dos nucleosídeos pela enzima guanosina hidrolase seguida da formação do nucleotídeo correspondente pela GPRTase, conforme ilustrado na **Figura 2** (ADAM, 2001).

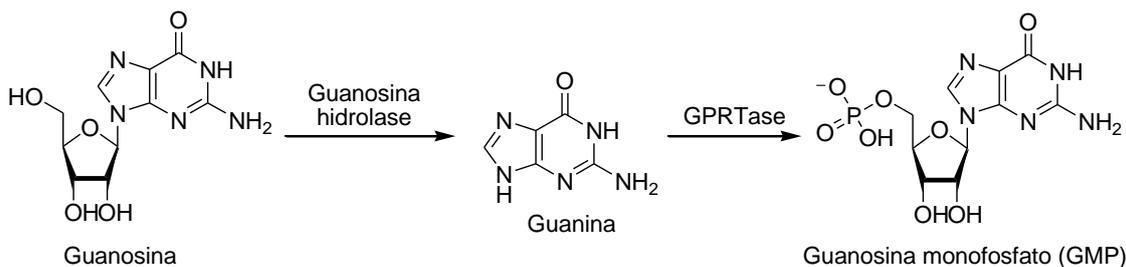


Figura 2. Conversão da guanosina em guanosina monofosfato *Giardia lamblia* através das enzimas guanosina hidrolase e guanina fosforibosil transferase (GPRTase). Adaptado de Adam (2001).

Apesar dos seres humanos serem capazes de realizar a síntese *de novo* de nucleotídeos purínicos (na qual o anel de purina é sintetizado a partir de diferentes fontes como glicina, ácido aspártico, glutamina e CO₂), a enzima humana homóloga à GPRTase de *G. lamblia* chamada de hipoxantina-guanina fosforibosil transferase (HGPRTase) tem papel fundamental na reutilização destas bases (BHAGAVAN; HA, 2015). A HGPRTase humana catalisa a formação de monofosfato de inosina e monofosfato de guanosina a partir de hipoxantina e guanina, respectivamente, sendo que sua deficiência parcial resulta na superprodução de ácido úrico, se traduzindo clinicamente como quadros precoces de gota e nefrolitíase. A ausência completa da ação desta enzima promove a chamada síndrome de Lesch-Nyhan, que se caracteriza por elevados níveis de ácido úrico no sangue e urina, além de disfunções motoras e neurológicas severas (DAVIDSON; PALELLA; KELLEY, 1988).

Pacientes com a síndrome de Lesch-Nyhan podem manifestar desordens motoras, cognitivas e/ou comportamentais, apresentando sintomas que incluem: hipotonia; atraso motor; distonia (contrações musculares mantidas); sinais de envolvimento piramidal incluindo espasticidade (rigidez muscular) e hiperreflexia; agressividade e automutilação (TORRES; PUIG, 2007; NYHAN *et al.*, 2000). Originalmente, a síndrome de Lesch-Nyhan se relaciona a um distúrbio genético herdado através de gene recessivo ligado ao braço longo do cromossomo X (Xq26), resultando assim na deficiência da expressão da HGPRTase (TORRES; PUIG, 2007; NYHAN *et al.*, 2000). Neste sentido, uma possível inibição farmacológica da atividade desta enzima poderia resultar em sintomas semelhantes aos observados na síndrome de Lesch-Nyhan, fato que ressalta a importância de se planejar inibidores específicos para a GPRTase de *G. lamblia*, com a menor afinidade possível pela enzima HGPRTase humana.

A partir da análise dos complexos cristalográficos destas enzimas, observou-se que existem diferenças estruturais entre o sítio ativo de ambas que podem ser exploradas para o planejamento de inibidores seletivos da GPRTase. Por exemplo, a presença de um resíduo de

tirosina (Y127) na GPTRase em lugar de uma isoleucina (I135) na HGPRTase resulta em impedimento estérico durante a ligação da purina no sítio ativo (MUNAGALA; SARVER; WANG, 2000). Outro resíduo importante corresponde a uma lisina presente tanto na GPTRase (K152) quanto na HGPRTase (K165). Apesar de este aminoácido ser comum às duas enzimas, na isoforma humana o grupo amino da lisina se encontra a uma distância de 3Å da carbonila da purina, de modo a estabelecer uma ligação de hidrogênio importante para a ligação do substrato. Já no caso da GPRTase, a distância entre os grupos é superior a 6Å, sendo somente possível a interação entre o grupo resíduo e a carbonila da guanina de forma indireta, utilizando duas molécula de água como ponte (Figura 3). É justamente a ausência desta interação que confere a especificidade da GPRTase por guanina em detrimento a hipoxantina, inosina ou xantina (MUNAGALA; SARVER; WANG, 2000).

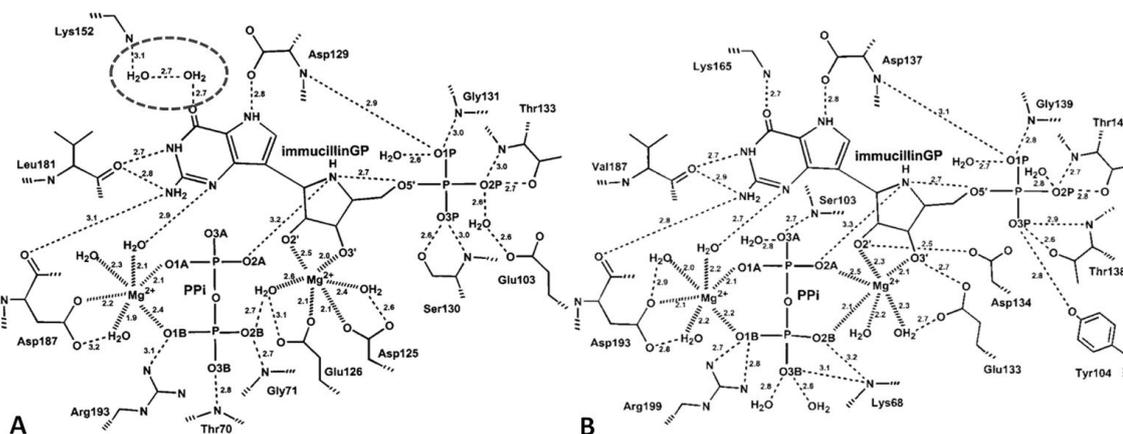


Figura 3. Comparação das interações realizadas entre um inibidor do estado de transição, Imucilina G fosfato, e os resíduos de aminoácido do sítio ativo da GPRTase de *G. lamblia* (A) e HGPRTase humana (B). Em pontilhado destacam-se as duas moléculas de água que servem como ponte entre o grupo amino da Lisina 152 e a carbonila da imucilina-GP. Adaptado de SHI et al., 1999 e SHI et al., 2000.

Neste sentido, Aronov e col. (2001) identificaram uma série de inibidores da GPRTase de *G. lamblia* realizando uma triagem virtual em base de dados contendo de derivados de ftalimidas, dentre as quais 16 foram sintetizadas e 7 apresentaram valores de IC₅₀ para a inibição da GPRTase de *G. lamblia* inferiores a 100 µM (Figura 4). Alguns destes compostos foram também avaliados frente à capacidade de inibição da HGPRTase humana e alcançaram níveis de seletividade para a GPRTase de *G. lamblia* de até sete vezes, mostrando que seria possível explorar as diferenças estruturais entre as duas enzimas para desenvolver inibidores da GPTRase com afinidade seletiva para *Giardia* (ARONOV et al., 2001).

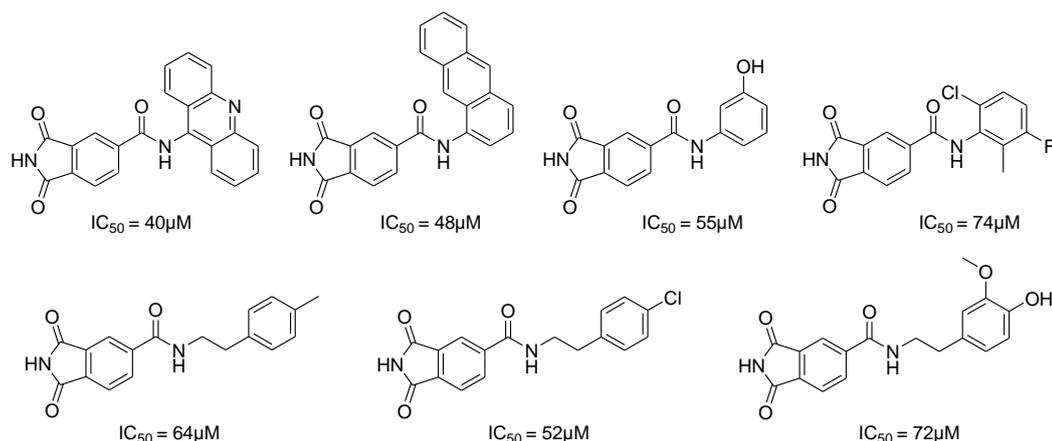


Figura 4. Estruturas dos inibidores da GPRTase de *G. lamblia* ensaiados por Aronov e col. (2001).

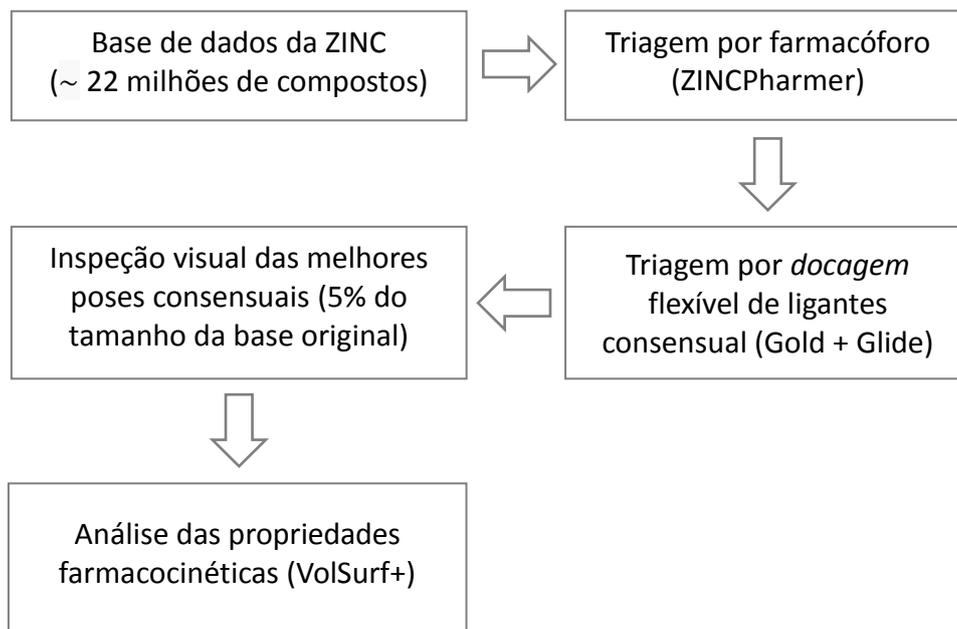
Seguindo esta linha de pensamento, a ideia central deste trabalho consiste na utilização de técnicas relacionadas ao planejamento de fármacos assistido por computador para a identificação de novos inibidores da GPRTase de *G. lamblia* com potencial uso terapêutico. Diversas iniciativas têm empregado métodos *in silico* para a descoberta de compostos bioativos com potencial terapêutico sobre doenças negligenciadas, principalmente malária, tripanossomíase africana e tuberculose (KORTAGERE, 2013). Desse modo, a descoberta de novos inibidores seletivos para a GPRTase de *G. lamblia* consistiria em um passo importante no desenvolvimento de novos fármacos capazes de combater linhagens de *Giardia* já resistentes à quimioterapia antiparasitária tradicional.

2. OBJETIVO:

O presente trabalho tem o objetivo de identificar novos inibidores da enzima Guanina Fosforibosil Transferase de *Giardia lamblia* que possuam a eficácia e um perfil farmacocinético adequado para serem empregados no tratamento da giardíase. Estes compostos serão planejados e selecionados frente a uma base de substâncias disponíveis comercialmente com o auxílio de diferentes técnicas de quimioinformática e modelagem molecular, permitindo a seleção dos compostos que teoricamente possuam uma maior afinidade pela enzima bem como características adequadas para serem veiculados por via oral.

3. PLANEJAMENTO EXPERIMENTAL:

O planejamento e a seleção dos compostos foram realizados empregando-se diferentes técnicas *in silico* de modo a se reduzir a quantidade de substâncias que serão avaliadas “*in vitro*” sobre a capacidade de inibição do crescimento de trofozoítos de *G. lamblia*. Estes métodos são representados de forma esquemática no fluxograma abaixo:



4. MATERIAIS E MÉTODOS:

4.1. Obtenção das estruturas cristalográficas e estudos do sítio de ligação das enzimas GPRTase de *Giardia* e HGPRTase humana:

A estrutura cristalográfica da enzima GPRTase de *Giardia lamblia* complexada ao inibidor de estado de transição Imucilina G fosfato (ImuGP) foi obtida do banco de dados *Protein Data Bank* (PDB) (www.rcsb.org) (BERMAN et al., 2000) (PDBID: 1DQN; complexo resolvido por difração de Raios-X com uma resolução de 1,75Å). Com a finalidade de se comparar as interações realizadas por um mesmo ligante entre a enzima do parasita e a enzima homóloga de *Homo sapiens*, também se obteve a partir do PDB a estrutura cristalográfica da enzima HGPRTase humana complexada à ImuGP (PDBID: 1BZY, complexo resolvido por difração de Raios-X com uma resolução de 2,00Å).

Em seguida, as interações intermoleculares mais relevantes realizadas entre a ImuGP e os resíduos de aminoácidos do sítio de ligação de cada enzima foram estudadas através dos programas Discovery Studio Visualizer (DSV) 4.0 (ACCELRY'S, 2013) e Ligand Explorer (MORELAND et al., 2005) e confrontadas com os dados disponíveis na literatura (SHI et al., 1999; SHI et al., 2000) para a seleção dos grupos farmacofóricos mais representativos para triagem virtual (ver seção seguinte). Além disso, realizou-se uma análise dos estados de ionização dos resíduos de aminoácido associados ao sítio de ligação da enzima GPRTase, tendo como referência a estrutura cristalográfica complexada à ImuGP (PDBID: 1DQN). A análise foi feita através do servidor DEPTH (http://mspc.bii.a-star.edu.sg/tankp/run_depth.html) (TAN et al., 2013) e PropKA (LI; ROBERTSON; JENSEN, 2005) selecionado o campo de força AMBER, conforme disponível no servidor PDB2PQR (http://nbc-222.ucsd.edu/pdb2pqr_2.0.0/) (DOLINSKY et al., 2004).

A fim de mensurar o tamanho e a localização da cavidade relacionada ao sítio de ligação da ImuGP na enzima GPRTase utilizou-se o servidor MetaPocket 2.0 (<http://projects.biotec.tu-dresden.de/metapocket/>), capaz de prever a localização e o tamanho de diversas cavidades na proteína pelo uso de um algoritmo consensual que emprega oito métodos distintos: LIGSITEcs, PASS, Q-SiteFinder, SURFNET, Fpocket, GHECOM, ConCavity e POCASA (ZHANG et al., 2011). Os resultados obtidos foram visualizados através do programa DSV 4.0 (ACCELRY'S, 2013).

4.2. Construção de um modelo farmacofórico para triagem virtual:

As técnicas de triagem virtual utilizam métodos computacionais para selecionar compostos farmacologicamente promissores em bases de dados contendo milhares ou mesmo milhões de compostos com ampla diversidade química e que geralmente possuem propriedades físico-químicas semelhantes às de fármacos, preenchendo assim os requisitos mínimos para garantir uma ampla gama de possíveis interações intermoleculares com os mais diversos alvos moleculares. As principais abordagens das estratégias de triagem virtual compreendem a triagem baseada na estrutura do alvo molecular, as quais consideram a estrutura tridimensional do alvo farmacológico no processo de seleção dos compostos (ex: docagem molecular) e a triagem baseada em ligantes que utilizam moléculas com atividade biológica conhecida como um “molde” para selecionar novas entidades químicas com as quais compartilham algum nível de similaridade e, conseqüentemente, possivelmente a mesma atividade biológica (ex: farmacóforo gerado por sobreposição de ligantes) (RODRIGUES et al., 2012).

Apesar do crescente aumento do poder de processamento dos computadores modernos, o ato de realizar triagens virtuais por docagem molecular em bases de dados contendo milhões de compostos pode ser considerado um desperdício de tempo e de recursos computacionais, além de impor um maior nível de dificuldade na posterior seleção dos compostos para os ensaios biológicos. Desse modo, é quase sempre interessante remover compostos indesejáveis e selecionar apenas os relevantes a partir de uma base de dados antes de um estudo de docagem intensivo (CHENG et al., 2012). Considerando que um farmacóforo pode ser visto como um denominador comum máximo de um grupo de moléculas que exibem perfis farmacológicos semelhantes por interagirem no mesmo sítio de ligação de um alvo terapêutico (LANGER; HOFFMANN, 2006), a triagem virtual por descritores farmacofóricos pode ser utilizada como uma pré-filtragem para as triagens virtuais por docagem molecular, reduzindo assim o uso de recursos computacionais sem comprometer a seleção de ligantes capazes de interagir com o alvo farmacológico (PEACH; NICKLAUS, 2009).

A partir da análise das interações intermoleculares mais relevantes entre o ligante e a GPRTase, definiu-se um farmacóforo contendo quatro grupos distintos, sendo a distância entre os grupos calculada utilizando-se o DSV 4.0 (ACCELRY, 2013). Em seguida, as características químicas tridimensionais destes grupos (ex: acceptor/doador de ligação de hidrogênio, aromático, hidrofóbico, cátion ou ânion) bem como suas coordenadas (x, y z) foram transcritas para o site do ZINCPharmer (<http://zincpharmer.csb.pitt.edu>) (KOES; CAMACHO, 2012), uma interface para triagem virtual por farmacóforo de compostos presentes no banco de dados da ZINC (<http://zinc.docagem.org/>) que conta atualmente com mais de 35 milhões de substâncias comercialmente disponíveis (IRWIN et al., 2012). As estruturas tridimensionais dos compostos selecionados foram então automaticamente compiladas em um arquivo de formato SD (*Structure Data*) para serem utilizadas no processo subsequente de triagem virtual por docagem molecular.

4.3. Docagem molecular:

A abordagem de docagem molecular pode ser utilizada para modelar a interação entre uma molécula pequena (ligante) e uma proteína em nível atômico, o que permite caracterizar o comportamento de moléculas pequenas no sítio de ligação de proteínas-alvo, bem como para elucidar seus processos bioquímicos fundamentais. Do ponto de vista metodológico, três grandes categorias de docagem podem ser estabelecidas: docagem com ligantes e receptores

rígidos, docagem com ligantes flexíveis e receptor rígido e docagem com ligantes e receptores flexíveis. O terceiro caso geralmente corresponde ao melhor modelo para se representar o que ocorre em meio biológico pois tanto o ligante quanto a macromolécula são tratados como entidades flexíveis, capazes de alterar suas conformações para um encaixe perfeito, formando assim o complexo ligante-macromolécula com o menor nível de energia. As grandes desvantagens da docagem flexível são a dificuldade de encontrar a melhor conformação dos aminoácidos da proteína e o tempo necessário para cada análise quando comparado à docagem com receptor rígido, no qual o espaço de busca conformacional é muito mais limitado (MENG et al., 2011). Desse modo, utilizou-se neste trabalho uma estratégia baseada em docagem com ligantes flexíveis e receptor rígido já que esta apresenta uma boa relação entre precisão e custo computacional. O processo de docagem envolveu a utilização de dois programas distintos, GOLD (http://www.ccdc.cam.ac.uk/products/life_sciences/gold/) e Glide (<http://www.schrodinger.com/productpage/14/5/>), tomando como base a estrutura cristalográfica da GPRTase (PDBID: 1DQN).

Programas de docagem molecular utilizam os chamados algoritmos de amostragem, os quais permitem ao *software* testar apenas as conformações mais importantes de um ligante no sítio ativo do alvo molecular em lugar de analisar todas as conformações possíveis, o que seria muito demorado e demandaria um enorme custo computacional (MENG et al., 2011). No caso do GOLD estes algoritmos de amostragem são baseados nos chamados algoritmos genéticos, os quais se baseiam na teoria da evolução de Darwin. Neste sistema, os diferentes graus de liberdade do ligante são codificados como números binários chamados “genes”. Estes genes são agrupados para formar um “cromossomo” que representa uma pose de um dado ligante no sítio ativo da proteína. Mutações e *crossover* entre cromossomos são realizadas para gerar novos genes e uma função de escore estabelece quais seriam os cromossomos mais aptos a sobreviver enquanto os demais seriam descartados. Em outras palavras esta função de escore determina quais as conformações do ligante seriam mais prováveis dentro do sítio ativo da proteína. Vale ressaltar que cada cromossomo no GOLD contém ainda informações sobre algumas partes flexíveis da proteína bem como sobre as ligações de hidrogênio e interações lipofílicas entre ligante-receptor, caracterizando assim um método de docagem com ligantes flexíveis e com receptor parcialmente flexível (KROEMER, 2007; MENG et al., 2011).

Já o *software* Glide utiliza um modelo de função de energia que combina termos empíricos e baseados em campos de força para realizar uma busca sistemática sobre a conformação, orientação e posição espacial do ligante no sítio ativo da proteína. Para tanto,

primeiro o ligante é colocado no sítio ativo do receptor através de um posicionamento aproximado e depois submetido a uma otimização de energia flexível no sítio utilizando uma função de energia padrão de mecânica molecular. Os melhores candidatos são então refinados por um processo de amostragem conformacional do tipo Monte Carlo, o qual permite a correta orientação de grupos periféricos e a alteração de ângulos torcionais internos (FRIESNER et al., 2004). Como GOLD e Glide operam com métodos completamente distintos para prever as conformações dos ligantes no sítio ativo da proteína e depois ranquear estas conformações via função de escore, acredita-se que os compostos apontados consensualmente como ativos pelos dois métodos apresentem uma boa probabilidade de serem ativos quando ensaiados experimentalmente.

4.3.1. Validação dos métodos de docagem - redocagem:

Antes de se iniciar a triagem virtual por docagem molecular propriamente dita, os parâmetros ótimos foram selecionados através da comparação nas diferentes corridas da conformação, posição e orientação (as quais no conjunto são conhecidas como pose) do ligante docado com a pose correspondente do ligante co-cristalizado com o alvo farmacológico. Este processo, conhecido como redocagem, é considerado um requisito mínimo para determinar se os parâmetros de docagem selecionados são apropriados para a simulação com um determinado alvo ou não. Tal análise é efetuada pela comparação dos valores de desvio quadráticos médios (RMSD, do inglês *Root-Mean-Square Deviation*), uma medida da distância média entre os átomos do ligante cristalográfico de referência e do ligante docado. Para análise, utilizou-se o valor de RMSD de 2,0Å como limite já que este é amplamente reconhecido como padrão na análise dos estudos de redocagem. Deve-se ressaltar, entretanto, que este critério não é absoluto já que os programas de docagem molecular naturalmente encontram dificuldades em prever o posicionamento correto de ligantes com um alto número de ligações rotacionáveis e, nestes casos, valores de RMSD inferiores a 2,0Å raramente são observados (AROOJ et al., 2013).

a) Redocagem no programa Glide:

Iniciou-se o processo pelo preparo da proteína, o qual apresenta três etapas distintas: 1) importação e processamento, 2) revisão e de modificação e 3) refinamento. Na etapa de importação e de processamento, a estrutura cristalográfica foi importada, os hidrogênios foram adicionados, as cadeias laterais ausentes foram adicionadas e as moléculas de água

foram mantidas, num primeiro momento. Durante a etapa de revisão e modificação, a cadeia proteica B foi excluída e foi verificada a presença das águas HOH541 e HOH 642 as quais ocupam uma região da proteína que pode conferir seletividade para os ligantes de GPRTase de *Giardia lamblia* quando comparada à HGPRTase humana (ver Resultados e Discussão). O módulo Epik foi então empregado para gerar os estados de protonação e os estados de ligação a metais para os ligantes com pH igual a 7,0 +- 2,0. Finalmente, utilizou-se na etapa de refinamento o módulo do PROPKA do GLIDE para determinar o posicionamento das ligações de hidrogênio, definindo o pH igual a 7,0. Além disso, foi realizada a otimização das hidroxilas pelo ProtAssign, além da Asn, Gln e His. Por fim, foi realizada uma minimização restrita utilizando os parâmetros padrões do programa (convergência de átomos pesados ao RMSD de 0,30Å pelo campo de força OPLS_2005).

Para a simulação, foi definido o local do processo de docagem através da geração do "grid" do receptor. Esse processo foi realizado a partir da definição de um centróide tomando como referência o ligante ImuGP. O tamanho do "grid" externo foi definido como de tamanho similar à ImuGP, o tamanho do "grid" interno, por sua vez, foi definido como 16Å para as três dimensões (x, y e z). Subsequentemente, realizou-se a preparação do ligante através do módulo "Ligprep" presente no GLIDE. O ligante escolhido foi a própria ImuGP, a qual foi extraída do complexo cristalográfico (PDBID: 1DQN) utilizando o programa DSV 4.0 e posteriormente importada para o GLIDE. Os parâmetros utilizados incluíram: 1) Minimização pelo campo de força OPLS 2005; 2) Ionização através da geração de estados de protonação num pH igual a 7,0 +- 2,0 em conjunto com estado de ligação ao metal, através do módulo Epik; 3) Remoção de contraíons; 4) Geração de tautômeros; 5) Os estereoisômeros foram computados de modo a reter quiralidades específicas.

O processo de docagem, por fim, utilizou parâmetros considerados padrão pelo programa, tais como: 1) Precisão de docagem padrão (SP); 2) Ligante flexível; 3) Penalização de amidas em conformação não co-planar; 4) Adição de estados de penalidade do "Epik" ao score final; 5) Minimização pós-docagem considerando parâmetros-padrão (número de poses = 10; limiar para rejeição de poses 0,5 kcal/mol). As análises de RMSD foram realizadas no programa DSV 4.0 utilizando a ImuGP cristalográfica como referência e definindo átomos pesados (*heavy atoms*) como parâmetro analisado.

b) Redocagem no programa GOLD:

Realizou-se a validação no GOLD com auxílio do assistente (*wizard*) do programa. A primeira etapa prévia à docagem no GOLD compreende o preparo da proteína, que foi

executada utilizando-se a estrutura da proteína 1DQN modificada para conter apenas a cadeia A, obtida através do programa DSV 4.0. Assim, os átomos de hidrogênio foram adicionados à estrutura proteica, as águas mantidas num primeiro momento, e o ligante ImuGP excluído para não influenciar no processo de docagem. Na etapa subsequente definiu-se o sítio de ligação, uma esfera de raio de 16Å cujo centro foi determinado no programa DSV 4.0 através das coordenadas XYZ do centróide da ImuGP (20,278; 10,018; 50,545).

Na seleção dos ligantes para docagem selecionou-se a ImuGP, cuja estrutura cristalográfica também foi definida como referência para o cálculo dos valores de RMSD. Quanto à definição do número de ciclos/poses gerados pelo algoritmo genético (GA, do inglês *genetic algorithm*), adotou-se o número de ciclos/poses padrão do programa, ou seja, 10 ciclos de GA. Em cada corrida, foram escolhidos cada um dos quatro algoritmos de amostragem disponíveis no programa: Goldscore (GS), Chemscore (CS), ASP (*Astex Statistical Potential*) (MOOIJ; VERDONK, 2005) e ChemPLP (*Piecewise Linear Potential*) (KORB; STÜTZLE; EXNER, 2009). A seleção do melhor algoritmo de amostragem foi realizada com base nos valores de RMSD calculados em relação ao ligante cristalográfico, bem como na inspeção visual das poses.

4.3.2. Validação dos métodos de docagem - cálculo do fator de enriquecimento:

O fator de enriquecimento (EF, do inglês *Enrichment Factor*) representa uma forma de se medir a precisão de um método de docagem em triagens virtuais pela análise do número de compostos ativos recuperados por este método. Assim, o EF mostra quantas vezes mais moléculas ativas são recuperadas por um método de docagem em comparação à seleção randômica destes compostos. Portanto, um fator de enriquecimento igual a cinco significa que o método de docagem é capaz de recuperar cinco vezes mais compostos ativos do que se estes fossem aleatoriamente selecionados da base de dados (MARKT; SCHUSTER; LANGER, 2011). O fator de enriquecimento pode ser definido de acordo com a equação (1) abaixo:

$$EF = \frac{Hits_a}{N_a} / \frac{Hits_t}{N_t} \quad Eq. (1)$$

Onde,

Hits_a = é o número de compostos ativos contidos na amostra da base de dados;

Hits_t = é o número total de compostos ativos da base de dados;

N_a = é o número de compostos da amostra;

N_t = é o número total de compostos da base de dados (WEI et al., 2002).

A fim de construir uma base de dados para fins de validação e comparação dos diferentes parâmetros de triagem virtual, 300 ligantes foram aleatoriamente selecionados a partir da base de dados original oriunda da triagem virtual por farmacóforo no ZINCPharmer. Esta base foi então enriquecida com ligantes cristalográficos de "HGPRase" humana, uma vez que não existem outros ligantes descritos para GPRTase de *Giardia lamblia* além da ImuGP. Para tanto, utilizou-se o banco de compostos BindingDB (<http://www.bindingdb.org/bind/index.jsp>) (LIU et al., 2007) empregando a sigla "HGPRase" como palavra-chave e delimitando a espécie humana (*Homo sapiens*).

Dentre as 91 estruturas dispostas no BindingDB, foram escolhidas 45 estruturas que apresentavam o anel purínico característico da guanina. Dessas 45 moléculas apenas uma estrutura foi obtida através do complexo cristalográfico (PDBID: 4IJQ) em sua conformação bioativa, enquanto que as demais moléculas tiveram suas estruturas tridimensionais determinadas através do programa Corina (http://www.molecular-networks.com/online_demos/corina_demo), o qual atribui a estruturas 3-D comprimentos de ligação e ângulos pré-definidos dependendo do tipo de ligação, do tipo de átomo e do estado de hibridação. Corina também define os ângulos de torcionais mais prováveis de acordo com a natureza da estrutura (acíclica, anéis pequenos/médios, macro/policíclicos, etc.), sendo capaz de reproduzir corretamente um número variado de estruturas determinadas por cristalografia de raios-X (GASTEIGER; RUDOLPH; SADOWSKI, 1990; SADOWSKI; GASTEIGER; KLEBE, 1994). As estruturas tridimensionais obtidas foram então refinadas utilizando um método semi-empírico PM6 (do inglês *Parameterized Model 6*) (STEWART, 2007) através do programa Spartan '14 (WAVEFUNCTION, 2004).

Através da busca no banco de dados do PDB foram encontrados 15 complexos cristalográficos relacionados à enzima HGPRase humana. Desses foram escolhidos apenas três complexos cristalográficos (PDBID: 1HMP, 3GEP, 3GGJ), os quais apresentavam inibidores descritos e não possuíam mutações e ou hibridizações com outras espécies. Os ligantes cristalográficos, incluindo a ImuGP (PDBID: 1DQN), foram extraídos em suas conformações bioativas e foram abertos no programa DSV 4.0, sendo posteriormente salvos num arquivo de formato ".sdf" em conjunto com os ligantes minimizados, de modo a obter 48 estruturas no total. De maneira a simplificar os procedimentos de validação, optou-se por construir um banco de ligantes contendo 300 estruturas, as quais foram aleatoriamente selecionadas a partir do banco de dados inicial, em conjunto com os 48 ligantes ativos (Hits_t) descritos anteriormente, totalizando 348 ligantes (N_t). Esse banco de ligantes foi salvo em formato ".sdf" sendo utilizado para os procedimentos de docagem descritos a seguir.

Avaliaram-se as diferentes funções de escore de ambos os programas através dos seguintes parâmetros:

1) Cálculo do fator de enriquecimento (EF): determinado conforme a equação (1), tomando como ponto de corte amostral 20% do banco de ligantes para avaliação das poses ($N_a = 70$ ligantes). O cálculo do $EF_{20\%}$ baseou-se nos melhores escores obtidos para cada função, sendo que foi utilizado um modelo consensual para os métodos que incluíam reescore, de modo a obter um escore médio, o qual foi utilizado para reavaliar o estadiamento das poses e, com isso, resultar no $EF_{20\%}$.

2) Avaliação visual das poses com base na ImuGP docada.

a) Validação dos parâmetros de triagem virtual no programa Glide:

A etapa de triagem virtual por docagem do banco de ligantes iniciou-se do mesmo modo que a redocagem. Realizou-se, primeiramente, o preparo da proteína seguido da definição do "grid". As etapas de preparo dos ligantes e de docagem foram conduzidas através do módulo "Virtual Screening Workflow", no qual não foram definidos mecanismos de filtragem. Os parâmetros de preparação dos ligantes incluíram: 1) Ionização através da geração de estados de protonação num pH igual a 7,0 +- 2,0 em conjunto com estado de ligação ao metal, através do módulo Epik; 2) Remoção de contraíons; 3) Remoção de estados de ionização e tautomerismo de alta energia; 4) Obtenção dos estereoisômeros a partir das propriedades estéricas 2-D e 5) Determinação de um número máximo de um estereoisômero por estrutura. A etapa de docagem utilizou o campo de força OPLS2005 e o procedimento baseou-se na precisão padrão (SP) utilizando os valores padrão do programa: 1) ligante flexível; 2) penalização de amidas em conformação não co-planar; 3) adição de estados de penalidade do "Epik" ao score final; 4) executar minimização pós-docagem; 5) manter 100% dos melhores compostos e reter todos os estados de escore bons.

b) Validação dos parâmetros de triagem virtual no programa GOLD:

A etapa de docagem do banco de ligantes iniciou-se pelo preparo da proteína, através do emprego do assistente "wizard", da mesma forma que a etapa de redocagem. Posteriormente, definiu-se o sítio de ligação com base no centróide utilizado anteriormente, empregando um raio de 16 Å. Em seguida, selecionou-se o banco de ligantes contendo as estruturas a serem docadas pelo algoritmo genético e ajustou-se o número de ciclos/poses gerados para 1. Utilizou-se a própria ImuGP para ser docada e servir como referência, procedendo em seguida para a escolha das funções de escore. A fim de definir a função de escore a ser empregada, bem como definir a possibilidade da utilização de reescore pelo GOLD, foram feitos cruzamentos entre as quatro funções de escore conforme a **Tabela 1**. As

opções de busca de GA foram definidas como de maior precisão e eficiência de procura de 100%.

Tabela 1. Definição das funções de escore conforme cruzamento entre funções:

Funções	Goldscore (GS)	Chemscore (CS)	ASP	ChemPLP
Goldscore (GS)	GS sem	GS com	ASP com	ChemPLP com
	reescore	reescore CS	reescore GS	reescore GS
Chemscore (CS)	GS com	CS sem	ASP com	ChemPLP com
	reescore CS	reescore	reescore CS	reescore CS
ASP	GS com	CS com	ASP sem	ChemPLP com
	reescore ASP	reescore ASP	reescore	reescore ASP
ChemPLP	GS com	CS com	ASP com	ChemPLP sem
	reescore	reescore	reescore	
	ChemPLP	ChemPLP	ChemPLP	reescore

4.3.3. Avaliação da influência das águas no processo de docagem através dos valores EF:

A fim de verificar a influência das águas HOH541 e HOH642 no procedimento de docagem, estabeleceram-se novos processos utilizando os programas GLIDE e GOLD através do uso dos melhores parâmetros definidos na etapa de validação da docagem, utilizando o banco de ligantes. O processo realizado no GLIDE foi semelhante aos demais procedimentos realizados anteriormente, com exceção da etapa de preparo da proteína, a qual incluiu a remoção das águas HOH541 e HOH642 do sítio de ligação à enzima. Os demais passos foram mantidos iguais. Ao final da docagem, calculou-se o EF e comparou-se o resultado com o EF do procedimento anterior.

O processo realizado no GOLD iniciou-se com o preparo da proteína, o qual envolveu a adição de hidrogênios e extração dos ligantes, da mesma forma que descrito no item anterior. No entanto, as chamadas águas "ativas" foram extraídas e posteriormente recarregadas, enquanto as demais, incluindo as águas HOH541 e HOH642, foram apagadas. As águas consideradas ativas foram determinadas por meio do programa DSV 4.0, através do mapeamento das interações com os ligantes (ImuGP e pirofosfato). O programa definiu as seguintes águas que interagem com os ligantes: HOH551, HOH558, HOH588, HOH754, HOH800, HOH817, HOH849 e HOH850. O procedimento de docagem foi mantido,

utilizando a melhor função de escore obtida no procedimento anterior. Ao final da docagem, calculou-se o EF e comparou-se o resultado com o EF do procedimento anterior.

4.3.4. Docagem das estruturas obtidas do banco de dados da ZINC, avaliação visual das poses e cálculo do Fator de Enriquecimento (EF):

As estruturas contidas no banco de dados da ZINC e obtidas através do farmacóforo foram compiladas em conjunto com as 48 estruturas presentes no banco de ligantes, totalizando uma base de 1278 compostos. A base foi triada por meio dos programas GLIDE e GOLD de maneira semelhante ao descrito nas etapas anteriores de validação.

A triagem virtual realizada pelo programa GOLD sucedeu-se utilizando a melhor função de escore definida nas etapas de validação anteriores, considerando, ainda, a provável influência das águas conforme resultado obtido no procedimento descrito anteriormente (ver Resultados e Discussão). A fim de verificar o sucesso da escolha da função de escore, realizou-se uma docagem utilizando a função de escore padrão do programa, GoldScore, mantendo os demais parâmetros inalterados. Ao final dos dois procedimentos, analisaram-se as poses visualmente, tomando como base o ligante cristalográfico ImuGP, observando o modo de ligação e de disposição do ligante no sítio, de forma a envolver os principais locais de interação, sobretudo com relação às águas HOH541 e HOH642.

4.3.5. Procedimento de docagem consensual dos resultados dos programas Glide e GOLD:

O procedimento de docagem consensual entre os resultados dos programas Glide e GOLD foi realizado por meio da construção de tabelas comparativas entre as duas melhores funções de escore encontradas através dos procedimentos anteriores. Esse procedimento considerou os valores de escore de cada função em módulo, em virtude que os valores da função Glidescore são negativos, de modo a estabelecer o escalonamento das poses de forma decrescente, ou seja, o maior valor de escore corresponde a primeira e melhor pose, a partir disso, as demais poses foram ranqueadas. Em seguida, analisaram-se visualmente aproximadamente 5% da base de dados (62 estruturas), as quais estavam contidas nas duas bases de dados e presentes na amostra representativa igual a 20% da base de dados (aproximadamente 250 estruturas). O consenso foi estabelecido por meio do cálculo das médias dos ranques, resultando no ranque médio, o qual foi reordenado de modo crescente, resultando no novo escalonamento das poses. Além disso, avaliaram-se as poses com relação às interações realizadas com os resíduos de aminoácido da GPRTase através do programa DSV 4.0.

4.3.6. Análise das propriedades farmacocinéticas dos compostos mais promissores:

Estima-se que as questões relacionadas à toxicidade e a um perfil farmacocinético inadequado sejam responsáveis por aproximadamente metade de todos os casos de falha no desenvolvimento de um fármaco. Portanto, o uso de ferramentas *in silico* para a predição das propriedades farmacocinéticas, metabólicas e toxicológicas de um composto pode auxiliar na escolha de protótipos que conduzam ao desenvolvimento de fármacos seguros, efetivos e com características adequadas para serem veiculados por via oral (VAN DE WATERBEEMD; GIFFORD, 2003). Neste contexto, os compostos resultantes dos estudos de *docagem* flexível consensual tiveram suas propriedades farmacocinéticas avaliadas no programa VolSurf+ (http://www.moldiscovery.com/soft_vsplus.php) considerando os descritores de absorvidade intestinal (CACO2) e permeação da barreira hematoencefálica (LogBB).

VolSurf⁺ cria modelos preditivos de absorção, distribuição, metabolismo e excreção (ADME) dos compostos calculando os descritores relevantes para estas propriedades e empregando análises estatísticas para a construção dos modelos. Neste processo, o programa calcula os campos de interação molecular tridimensionais para cada substância e emprega métodos de processamento de imagem para convertê-los em descritores moleculares simples que são mais fáceis de entender e interpretar. Estes descritores caracterizam quantitativamente a forma, o tamanho, a polaridade e a hidrofobicidade das moléculas, bem como a relação entre eles. Em sequência, métodos estatísticos multivariados permitem a criação dos modelos que relacionam os descritores com diferentes propriedades biológicas. Tais modelos se propõem a prever a solubilidade de um composto, sua absorção intestinal passiva, permeação pela barreira hematoencefálica, ligação às proteínas plasmáticas, volume de distribuição e metabolismo. Estes modelos foram desenvolvidos a partir de dados experimentais coletados de ensaios “*in vitro*” que simulam o comportamento dos fármacos “*in vivo*” (CRUCIANI et al., 2000; CRUCIANI; PASTOR; GUBA, 2000; CRIVORI et al., 2000).

5. RESULTADOS E DISCUSSÃO

5.1. Estudos do sítio de ligação das enzimas GPRTase de *Giardia* e HGPRTase humana:

As análises do sítio de ligação da enzima GPRTase de *Giardia lamblia* (PDBID: 1DQN) com o inibidor de estado de transição, ImuGP, realizadas pelos programas Ligand Explorer e DSV mostram a existência de múltiplas interações. Essas interações incluem: interações hidrofóbicas, empilhamento π , ligações de hidrogênio e interações metálicas.

A molécula de ImuGP realiza interações hidrofóbicas com os seguintes aminoácidos: tirosina 127 (Tyr127), aspartato 129 (Asp129), treonina 133 (Thr133) e triptofano 180 (Trp180) (**Figura 5**), considerando um limiar de interação de 3,9 Å. O núcleo 9-deazaguanina da ImuGP apresenta-se interposto pelos aminoácidos aromáticos Tyr127 e Trp180, de modo a realizar interações hidrofóbicas e interações aromáticas do tipo empilhamento π em ambos os lados da molécula (**Figura 6**). Observa-se que a Tyr127 encontra-se em posição perpendicular ao anel 9-deazaguanina da ImuGP. Esses resultados encontram-se consistentes com os dados descritos na literatura (SHI *et al.*, 2000; ARONOV *et al.*, 2001; MUNAGALA *et al.*, 2000).

Estudos de bioinformática comparando diferentes fosforribosiltransferases (PRTases) demonstraram que a cadeia lateral aromática correspondente ao Trp180 da GPRTase mostra-se altamente conservada (SHI *et al.*, 2000; MUNAGALA *et al.*, 2000), com variações que abrangem desde a troca do Trp por Tyr na HGXPRTase de *Tritrichomonas foetus* ou por fenilalanina (Phe) na HGPRTase humana, até a manutenção do Trp correspondente em HGXPRTase de *Toxoplasma gondii* (MUNAGALA *et al.*, 2000). Essas interações sugerem a importância do grupamento aromático para a ligação ao sítio da enzima GPRTase

O resíduo Tyr127 da GPRTase de *G. lamblia*, por outro lado, mostra-se bastante diferenciado em comparação aos demais resíduos correspondentes em outras PRTases, nas quais há uma conservação de resíduos de isoleucinas (Ile) ou de leucinas (Leu) nessa região (SHI *et al.*, 2000; ARONOV *et al.*, 2001; MUNAGALA *et al.*, 2000). Estudos de cinética determinaram a influência da Tyr127 no modo de ligação da guanina no sítio da GPRTase. Esses estudos foram conduzidos comparando a enzima selvagem (*Wild Type*) com a enzima mutada Y127I, a qual possui uma isoleucina no lugar da tirosina na posição 127. Observou-se que as constantes enzimáticas K_m e k_{cat} da enzima mutada foram menores ($K_m = 2,8 \pm 0,2 \mu\text{M}$ e $k_{cat} = 4,8 \pm 0,1 \text{ s}^{-1}$) que os valores para a enzima selvagem ($K_m = 16,4 \pm 1,8 \mu\text{M}$ e $k_{cat} = 76,7 \pm 2,5 \text{ s}^{-1}$). Por meio desses estudos, observa-se que o resíduo Tyr127 da GPRTase de *G. lamblia* apresentaria um impedimento estérico em comparação aos demais resíduos equivalentes das outras PRTases (MUNAGALA *et al.*, 2000). Além disso, Tyr127 talvez

desempenhe um papel pouco importante para manutenção das interações dentro do sítio de ligação da GPRTase, uma vez que a mutação Y127I resultou em aumento de afinidade pelo substrato.

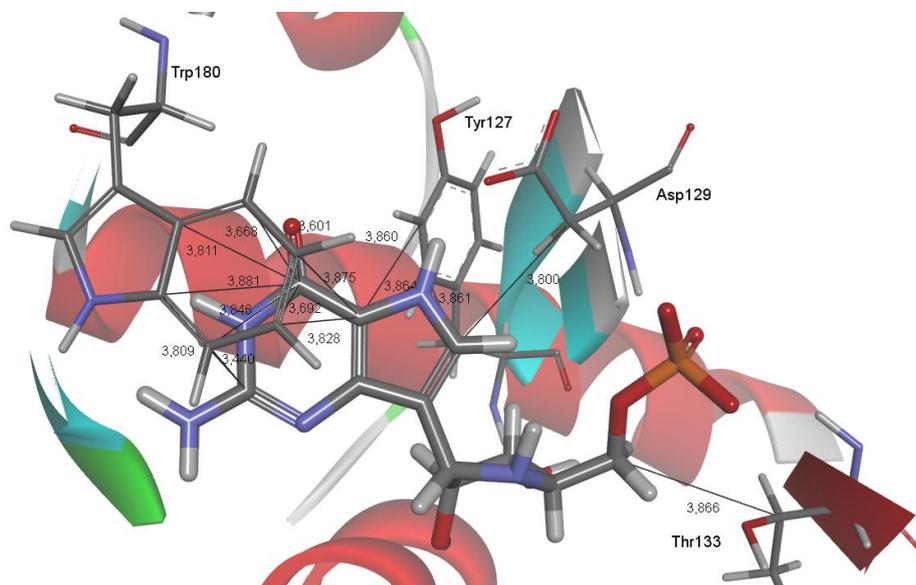


Figura 5. Principais interações hidrofóbicas entre o ligante ImuGP e a enzima GPRTase

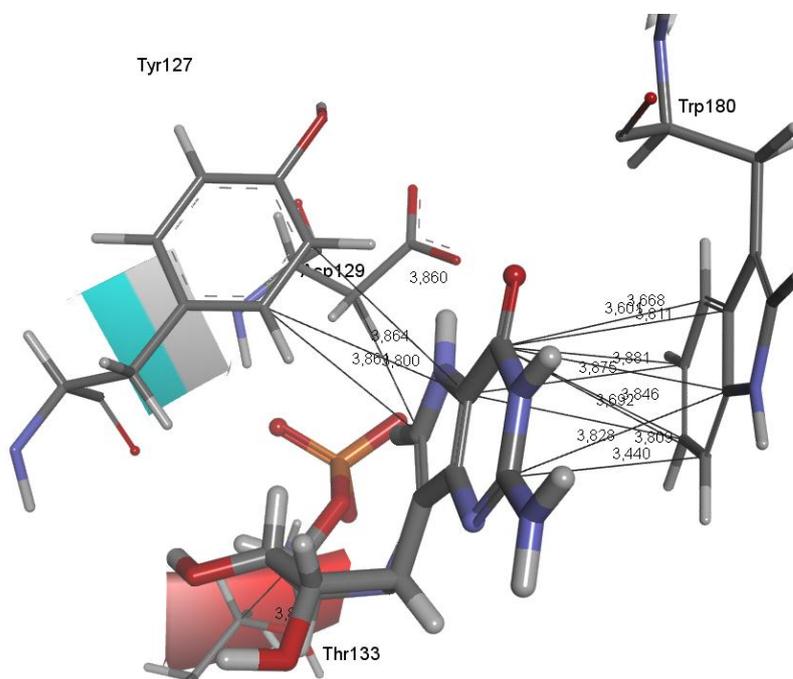


Figura 6. ImuGP realizando interações hidrofóbicas e aromáticas com a enzima GPRTase

As interações do tipo ligações de hidrogênio realizadas entre a ImuGP e o sítio de ligação à GPRTase de *G. lamblia* incluem os seguintes aminoácidos (**Figura 7**): aspartato 125 (Asp125), glutamato 126 (Glu126), tirosina 127 (Tyr127), aspartato 129 (Asp129), serina 130 (Ser130), glicina 131 (Gly131), treonina 133 (Thr133), leucina 181 (Leu181) e aspartato 187 (Asp187), considerando um limiar de interação de 3,3 Å. Esses dados encontram-se de acordo com o relatado na literatura. De acordo com Shi e colaboradores, as interações realizadas entre o grupamento amina exociclo (N2) e os grupos carbonílicos presentes nos resíduos Leu181 e Asp187, distantes respectivamente 2,8 Å e 3,1 Å, desempenham um papel fundamental na seletividade da enzima GPRTase pela guanina, em detrimento de outros derivados purínicos como xantina e hipoxantina (SHI *et al.*, 2000).

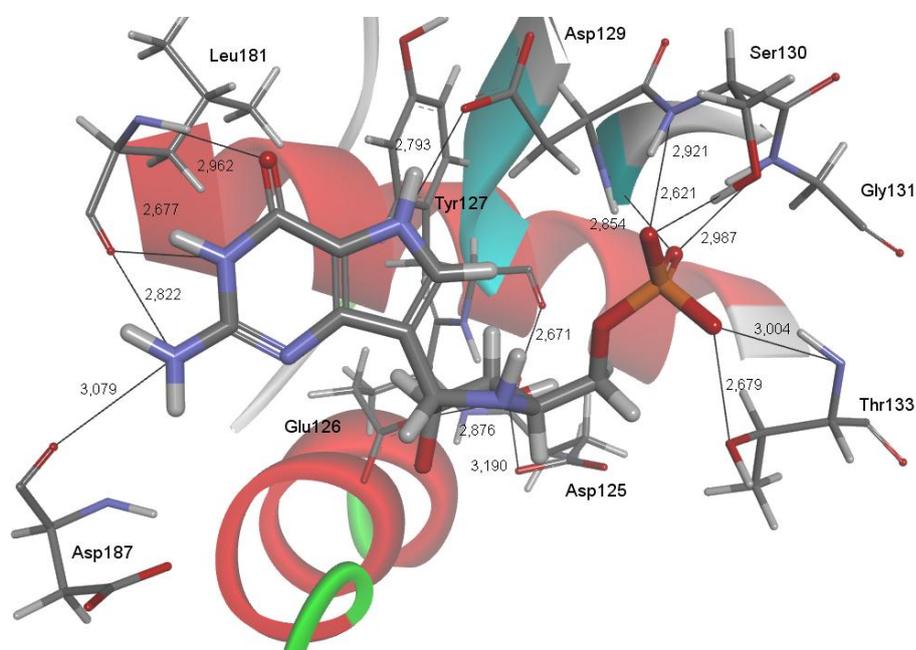


Figura 7. Principais interações por ligações de hidrogênio entre ImuGP e GPRTase.

Ao analisar as estruturas cristalográficas de outras PRTases, pode-se perceber a conservação das interações entre o grupo ϵ -amino da lisina e o oxigênio carbonílico presente na posição 6 do anel purínico, as quais distam em média cerca de 3 Å (2,7 Å para a HPRTase humana) (SHI *et al.*, 2000; ARONOV *et al.*, 2001; MUNAGALA *et al.*, 2000). Por outro lado, a lisina correspondente presente no complexo cristalográfico da GPRTase (Lys152) dista 6,3 Å do oxigênio carbonílico (**Figura 8**) (SHI *et al.*, 2000; ARONOV *et al.*, 2001; MUNAGALA *et al.*, 2000).

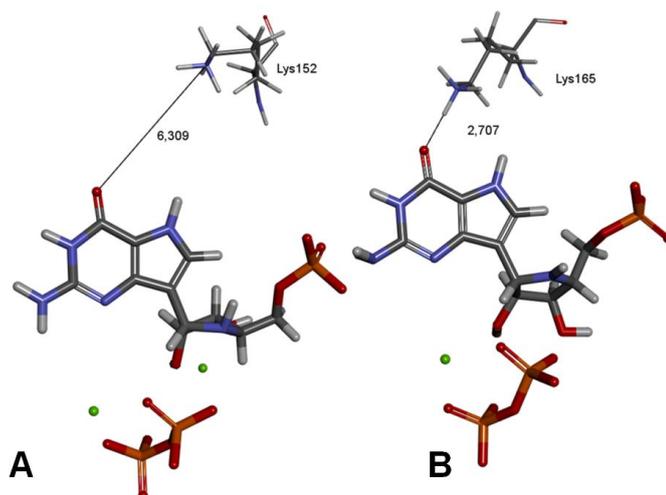


Figura 8. **A)** Avaliação da distância entre ImuGP e o resíduo Lys152 (GPRTase); **B)** Avaliação da distância entre ImuGP e o resíduo Lys165 (HGPRTase).

Entretanto, ao avaliar as interações de hidrogênio realizadas entre a ImuGP e as moléculas de água presentes no sítio de ligação da enzima GPRTase de *G. lamblia*, nota-se a existência de interações importantes entre o oxigênio carbonílico da posição 6 e duas moléculas de água, HOH541 e HOH642, as quais se encontram a uma distância de 3,0 Å e 2,7 Å, respectivamente, do grupo funcional do ligante. As águas HOH541 e HOH642 também se encontram distantes 3,1 Å e 4,1 Å, respectivamente, do grupo ϵ -amino da Lys152. Por outro lado, a análise das ligações de hidrogênio realizadas entre a ImuGP e as moléculas de água no sítio de ligação da enzima HPRTase humana demonstra a existência de apenas uma molécula de água nas proximidades da Lys165 (HOH651), a qual se encontra distante 5,2 Å do oxigênio carbonílico da ImuGP e 3,0 Å do grupo ϵ -amino da Lys165 (**Figura 9**).

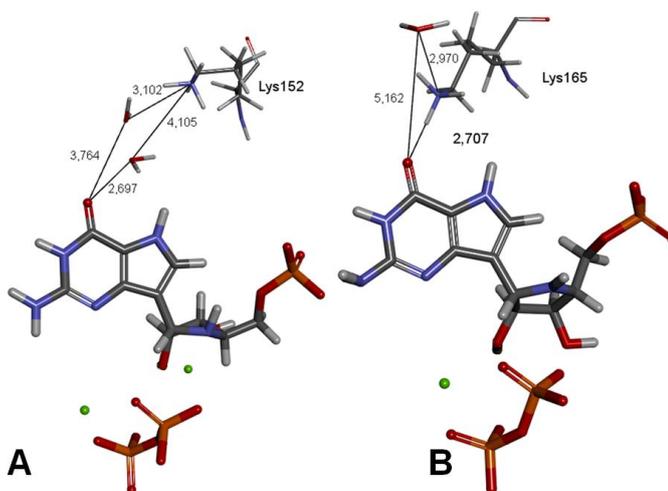


Figura 9. **A)** Avaliação da distância entre águas HOH541 e HOH642, ImuGP e Lys152(GPRTase); **B)** Avaliação da distância entre águas, ImuGP e Lys165 (HGPRTase).

Assim sendo, ligantes contendo grupos capazes de realizar ligações de hidrogênio diretamente com a Lys152 da GPRTase de *G. lamblia* poderiam constituir a base do desenvolvimento de ligantes seletivos para a enzima uma vez que estes mesmos grupos poderiam sofrer repulsão estérica com a Lys165 da HGPRTase humana. A fim de se delimitar a geometria deste e de outros possíveis subsítios presentes na GPRTase que pudessem ser explorados para o desenvolvimento de novos inibidores da enzima, realizou-se a predição dos sítios putativos de ligação empregando-se o servidor Metapocket (HUANG *et al.*, 2009).

Os métodos de predição de sítio de ligação proteína-ligante podem ser divididos em dois grandes grupos, de acordo com a estratégia adotada: 1) métodos de predição geométrica e 2) métodos de predição energética (KRIVAK & HOKSZA, 2015; ZHANG *et al.*, 2011; PÉROT *et al.*, 2010; LAURIE & JACKSON, 2006). Os métodos de predição geométrica do sítio de ligação proteína-ligante abrangem diversas metodologias, desde a disposição de esferas virtuais arranjadas entre a proteína e a região acessível pelo solvente até a utilização de triângulos de Delaunay. De modo geral, esses métodos buscam prováveis cavidades e fendas na superfície da proteína-alvo, através de algoritmos, os quais predizem regiões correspondentes aos átomos pertencentes à proteína e regiões sem interações que correspondem ao solvente. Encaixam-se aqui alguns dos métodos utilizados pelo servidor Metapocket 2.0: GHECOM (KAWABATA, T., 2006), SURFNET (LASKOWSKI, R. A., 1995), LigSite^{CS} (HUANG & SCHROEDER, 2006), POCASA (YU *et al.*, 2010) e FPocket (LE GUILLOUX *et al.*, 2009).

O servidor Metapocket analisa de modo consensual os resultados gerados individualmente por cada ferramenta, comparando os dados de função de escore gerados por cada preditor através do cálculo de uma função chamada Z-score. Em seguida, o programa escalona os resultados de cada sítio de ligação gerado por cada preditor, conforme os maiores valores de Z-score. Depois disso, seleciona os três melhores resultados e os agrupa de acordo com a similaridade espacial dos sítios, ranqueando novamente conforme o valor total do Z-score. Finalmente, a última etapa identifica os resíduos de aminoácido que formam os sítios de ligação preditos pelo Metapocket (HUANG *et al.*, 2009).

Assim sendo, a análise consensual realizada através do servidor Metapocket 2.0 utilizando as ferramentas GHECOM, SURFNET, LigSite^{CS}, POCASA, FPocket, ConCavity e PASS11, indicou a existência de 7 sítios de ligação putativos na enzima GPRTase de *Giardia lamblia* (**Figura 10**).

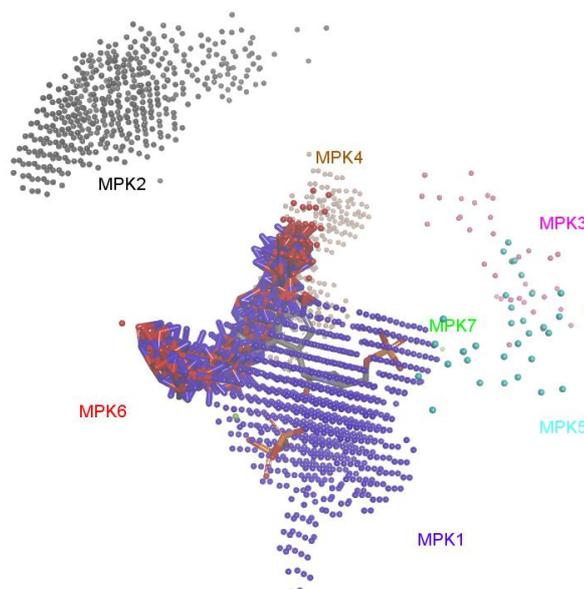


Figura 10. Predição de sete cavidades putativas na GPRTase de *G. lamblia* através do servidor MetaPocket. Destaca-se a sobreposição entre a ImuGP e as predições dos sítios MPK1 e MPK6.

Conforme o z-score gerado pelo servidor Metapocket, observa-se que o sítio MPK1 apresenta o melhor escore dentre os demais. Pode-se observar, ainda, que o sítio previsto delinea a região ocupada pelo ligante ImuGP, predizendo a composição da cavidade pelos seguintes aminoácidos (**Figura 11**): Leu68, Leu69, Thr70, Gly71, Ala72, Asp125, Glu126, Asp187, Arg193, Leu186, Tyr127, Leu181, Trp180, Asp188, Asn189, Gly190, Asp129, Ser179, Lys152, Thr133, Val128, Val94, Ile201, Gly178, Ile134, Ser130, Gly131, Lys93, Ser95, His132, Ser96, Val105, Ile157, Glu103, Lys98, Ser104 e Val106.

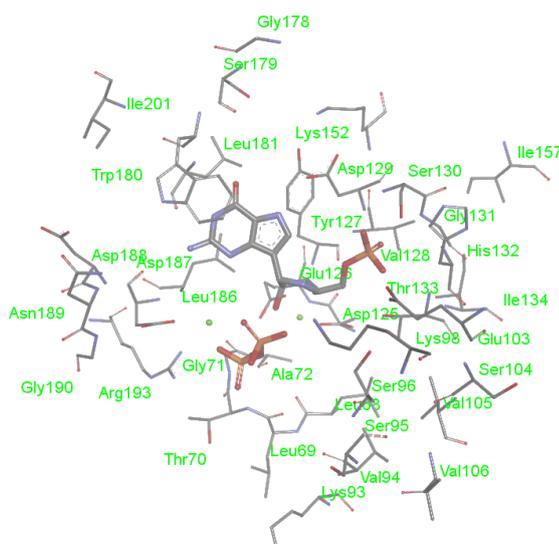


Figura 11. Aminoácidos que compõem o sítio de ligação da GPRTase preditos pelo MPK1.

Esta lista inclui todos os resíduos de aminoácido que interagem com o ligante ImuGP, seja por ligações de hidrogênio, seja por interações aromáticas/hidrofóbicas. Além disso, cabe salientar que a predição pelo Metapocket inclui a Lys152, fato este que corrobora com a hipótese adotada da possibilidade de ampliar as interações do ligante com a enzima através de interações de hidrogênio com o grupo ϵ -amino desse resíduo de aminoácido.

Considerando que estruturas proteicas resolvidas por cristalografia de raios-X geralmente não revelam as posições dos átomos de hidrogênio, pode-se concluir que os estados tautoméricos dos resíduos de histidina e de outros grupos presentes nos ligantes também não poderão ser diretamente determinados. De modo análogo, as posições relativas dos átomos de nitrogênio e de oxigênio na glutamina e as cadeias laterais da asparagina também podem ser muito difíceis de se atribuir nestas situações. O mesmo pode ser dito para os estados de ionização dos resíduos de aminoácidos localizados no sítio ativo da proteína, fato que pode afetar significativamente as simulações de docagem molecular. Embora possa parecer razoável considerar nestas simulações todos os possíveis estados de protonação dos resíduos de aminoácido, na prática isto acarretaria num custo computacional proibitivo por conta da explosão combinatória de possibilidades. Assim sendo, as simulações de docagem molecular exigem que sejam selecionados *a priori* os estados de protonação de cada um destes resíduos de aminoácido (SPYRAKIS *et al.*, 2007).

Assim sendo, realizou-se a análise dos grupos ionizáveis do sítio de ligação da enzima GPRTase conduzidas através do servidor DEPTH, a qual indicou que os seguintes resíduos de aminoácidos encontram-se ionizados: Asp125, Glu126, Asp129, Lys152 e Asp187. O servidor DEPTH emprega o uso de funções como interações eletrostáticas, ligações de hidrogênio e área superficial acessível pelo solvente em conjunto com o algoritmo DEPTH para prever o pKa dos resíduos de aminoácidos. O algoritmo DEPTH realiza a medida da distância mais próxima entre um resíduo de aminoácido ou um átomo e o solvente livre. Para isso o algoritmo determina a profundidade dos resíduos da proteína através da solvatação virtual, utilizando o modelo molecular SPC216 presente no pacote GROMACS, de modo a obter as águas livres (*bulk waters*) e as águas não livres (*non-bulk waters*) (TAN *et al.*, 2013; CHAKRAVARTY, S. & VARADARAJAN, R., 1999).

O módulo PropKa do servidor PDB2PQR, indicou que os mesmos aminoácidos avaliados pelo servidor DEPTH (Asp125, Glu126, Asp129, Lys152 e Asp187) encontram-se ionizados. Cabe ressaltar que, de acordo com o PropKa, o resíduo Asp125 apresenta um valor de pKa predito de 7,43, devido à ionização em conjunto com o Glu126. O PropKa estima o pKa através de 5 passos: 1) Leitura do arquivo contendo o complexo cristalográfico ".pdb" e

identificação dos grupos ionizáveis; 2) Determinação da posição dos prótons das amidas do esqueleto proteico nos resíduos de Asparagina (Asn), Glutamina (Gln), Trp, Histidina (His) e Arg; 3) Obtenção de valores de pK_a temporários para grupos ionizáveis através do uso de estados de protonação de outros grupos ionizáveis de fácil determinação; 4) Determinação do pK_a através de repetição; 5) Obtenção dos valores de pK_a preditos (LI *et al.*, 2005).

5.2. Construção de um modelo farmacofórico para triagem virtual:

A partir das análises das interações que a ImuGP realiza com os aminoácidos do sítio de ligação da enzima GPRTase, definiu-se um farmacóforo que abrangesse as interações mais relevantes. O farmacóforo construído tem como intuito preservar a atividade de inibição frente à enzima presente em *Giardia lamblia*, bem como aumentar a especificidade do ligante frente ao parasita, minimizando, assim, potenciais interações com a enzima humana e o desenvolvimento de reações adversas indesejáveis, além de possibilitar a obtenção de um número razoável de estruturas do banco de dados da ZINC, através da ferramenta ZINCPharmer associada ao banco de dados da ZINC.

Os grupos farmacofóricos originalmente escolhidos encontram-se descritos na **Figura 12**. Primeiramente, optou-se por escolher apenas três pontos farmacofóricos, sendo dois grupos doadores de ligações de hidrogênio e um grupo aceptor de ligações de hidrogênio.

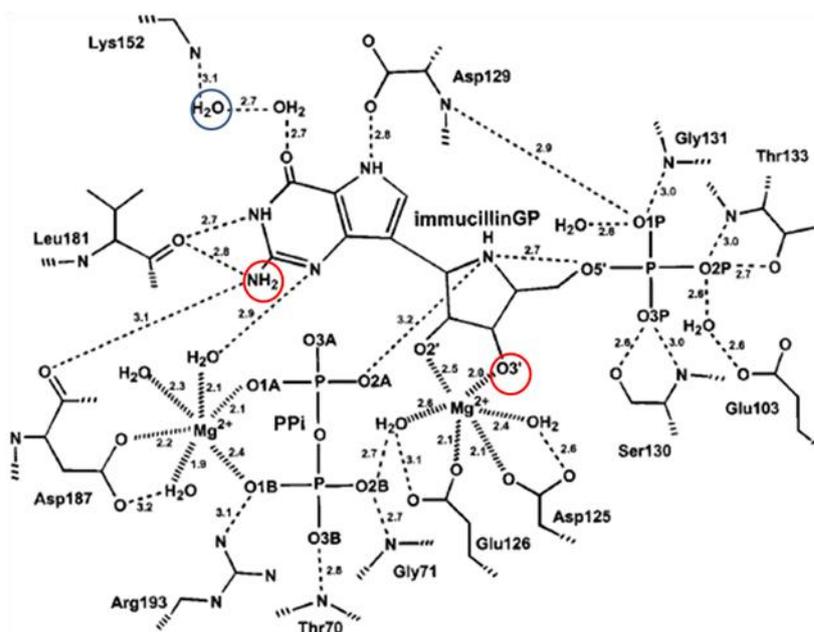


Figura 12. Grupos farmacofóricos do primeiro modelo construído com base na estrutura da ImuGP.

Os grupos doadores de ligações de hidrogênio escolhidos, foram: 1) Nitrogênio N2 exociclo (coordenadas XYZ: 17.30 9.09 55.33), o qual se encontra envolvido com interações de hidrogênio com os resíduos Leu181 e Asp187 que podem estar associadas à seletividade da enzima à guanina; 2) Oxigênio 3' da ribose (coordenadas: 18.78 8.46 47.36), o qual interage com o magnésio utilizado como cofator da enzima. O grupo acceptor escolhido foi determinado conforme a posição da água HOH541 (coordenadas: 19,18; 16,76; 52,76), a qual pode fazer ligações de hidrogênio com o grupo ϵ -amino da Lys152, que se encontra em estado ionizado. A adição de um grupo prolongador contendo um grupo acceptor de ligações de hidrogênio pode conferir uma especificidade maior à enzima GPRTase de *G. lamblia*, uma vez que a Lys152 encontra-se a uma distância de 6,3 Å da ImuGP, enquanto que o resíduo correspondente na HGPRase humana, Lys165, encontra-se a cerca de 2,7 Å do ligante, resultando, portanto, num impedimento estérico na enzima humana.

No entanto, a escolha desses três grupos farmacofóricos não foi restritiva o suficiente no ZincPharmer, gerando uma base de dados de 100.211 ligantes na época da análise (**Figura 13, A**). Optou-se pela adição de um quarto grupo farmacofórico que possibilitasse a restrição para formação de uma base menor, em torno de 1.000 ligantes. Sendo assim, escolheu-se a interação aromática do anel de seis membros da purina (coordenadas: 18.72 10.65 53.62), a qual promove interações hidrofóbicas e de empilhamento π com o resíduo de Trp180, aumentando a afinidade dos ligantes pelo receptor (**Figura 13, B**). Foram obtidas 3.705 estruturas, as quais foram refinadas, de modo a se obter 1.230 estruturas não-redundantes. Os grupos farmacofóricos escolhidos para triagem e as distâncias entre estes grupos encontram-se dispostos na **Figura 14**, gerada utilizando-se o DSV 4.0 (ACCELRYS, 2013).

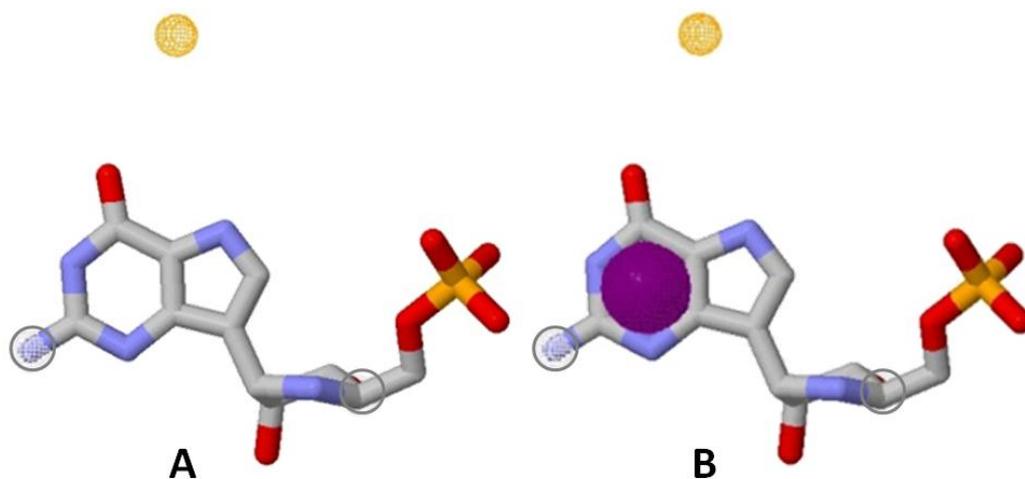


Figura 13. Construção do modelo farmacofórico no ZincPharmer. **A)** Farmacóforo formado por três grupos (100.211 *hits* obtidos). **B)** Farmacóforo formado por quatro grupos (1.230 *hits* obtidos).

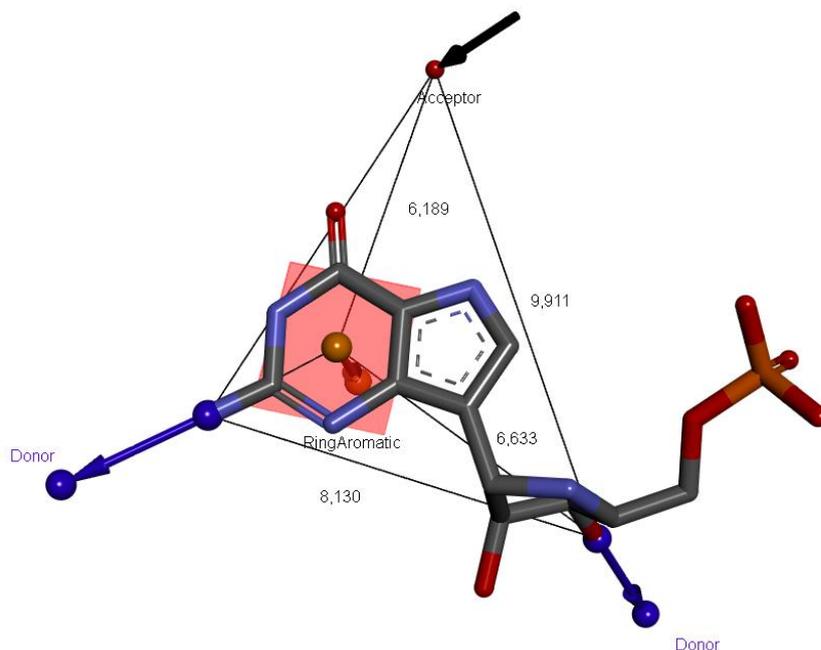


Figura 14. Modelo farmacofórico adotado para triagem de compostos da ZINC. As esferas azuis representam grupos doadores de ligações de hidrogênio, a vermelha um grupo aceptor de ligações de hidrogênio e o quadrado vermelho o anel aromático.

5.3. Docagem molecular:

5.3.1. Validação dos métodos de docagem - redocagem:

a) Redocagem no programa Glide:

O processo de redocagem realizado no programa Glide gerou nove poses como resultado. Essas poses apresentaram valores de RMSD inferiores a 0,400 Å e valores de escore inferiores a -9,100 para o "docking score", tomando como referência o ligante ImuGP obtido da estrutura cristalográfica, conforme pode ser observado na **Tabela 2**.

GlideScore, uma das funções de escore atribuída ao Glide foi desenvolvida com base na função ChemScore, apresentando algumas modificações e ampliações comparadas à sua precursora: introdução de um modelo de solvatação, incorporação das interações eletrostáticas (Coulombicas) e de Van der Waals, especificações dos termos de interações metálicas e de interações por ligações de hidrogênio (FRIESNER *et al.*, 2004). Por sua vez, a função Glide Emodel combina a função GlideScore com o escore obtido pelo "grid" energético das interações não-ligantes, podendo ser adequada para análise conformacional (SCHRÖDINGER, 2009). Por fim, a função "Docking Score" atribui penalidades energéticas ao resultado final, as quais são calculadas pelo módulo Epik a partir do custo energético para gerar cada estado, levando em consideração o pH do modelo e das interações metálicas com o receptor (SASTRY *et al.*, 2013).

Tabela 2. Tabela informativa dos valores de escore obtidos pelas poses da ImuGP durante o processo de redocagem.

Nome	Docking Score (kcal/mol)	GlideScore (kcal/mol)	Glide Emodel	RMSD (Å)
1DQN - Pose 1	-13,8012	-14,4465	-241,765	0,269414
1DQN - Pose 2	-13,361	-14,6223	-193,42	0,307979
1DQN - Pose 3	-13,3468	-14,1802	-216,417	0,269653
1DQN - Pose 4	-13,3161	-14,6549	-160,802	0,343434
1DQN - Pose 5	-12,497	-13,8523	-200,524	0,271643
1DQN - Pose 6	-12,082	-13,6153	-176,926	0,268535
1DQN - Pose 7	-10,4002	-13,4938	-221,348	0,294789
1DQN - Pose 8	-10,2167	-13,6977	-165,12	0,3587
1DQN - Pose 9	-9,12165	-12,9253	-180,728	0,292095

A análise detalhada das poses obtidas (**Figura 15**) revela que existe uma concordância entre os valores energéticos estimados e o posicionamento das estruturas no sítio de ligação se comparadas à pose cristalográfica da ImuGP, em verde, demonstrando que o método foi reprodutível uma vez que existe um grau de incerteza nos valores de RMSD como pode ser observado na **Tabela 2**.

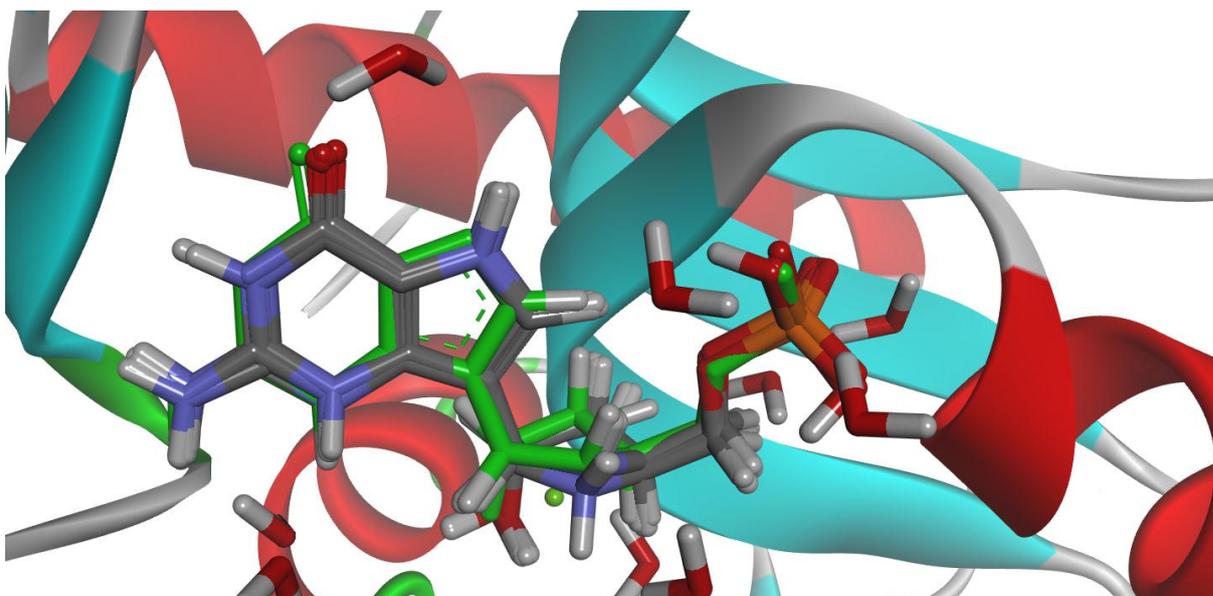


Figura 15. Análise das poses obtidas por de redocagem no Glide. ImuGP cristalográfica em verde.

b) Redocagem no programa GOLD:

O procedimento de redocagem desenvolvido através do programa GOLD conduziu à escolha da melhor função, a qual foi utilizada como padrão nos procedimentos posteriores. Essa escolha baseou-se na inspeção visual e na avaliação do RMSD para as dez poses geradas, as quais podem ser observadas na **Tabela 3** e na **Figura 16**:

Tabela 3. Valores de escore para as quatro principais funções de escore do programa GOLD relativos à redocagem da ImuGP.

Pose	GoldScore		ChemScore		ASP		ChemPLP	
	GS Fit	RMSD	CS Fit	RMSD	ASP Fit	RMSD	PLP Fit	RMSD
ImuGP Pose 1	129,011	0,379	43,9623	0,2794	70,8184	0,431	110,739	0,3184
ImuGP Pose 2	126,218	0,4025	45,9948	0,2753	71,7557	0,3113	113,336	0,238
ImuGP Pose 3	128,722	0,3855	45,3759	0,266	70,6471	0,3259	117,395	0,3065
ImuGP Pose 4	128,484	0,3967	45,8581	0,303	71,8784	0,2707	109,591	0,6232
ImuGP Pose 5	128,348	0,3769	44,9957	0,4157	71,4599	0,2423	110,111	0,3458
ImuGP Pose 6	128,961	0,3755	46,0581	0,3299	70,0172	0,2832	109,923	0,3191
ImuGP Pose 7	124,037	0,4116	47,1551	0,3513	72,2943	0,5	109,482	0,2949
ImuGP Pose 8	125,404	0,3757	46,2988	0,2252	69,8719	0,3086	111,699	0,2116
ImuGP Pose 9	125,424	0,3991	46,0021	0,295	70,4936	0,2697	116,776	0,2822
ImuGP Pose 10	129,362	0,3932	46,0221	0,2287	72,2341	0,2256	113,595	0,2329

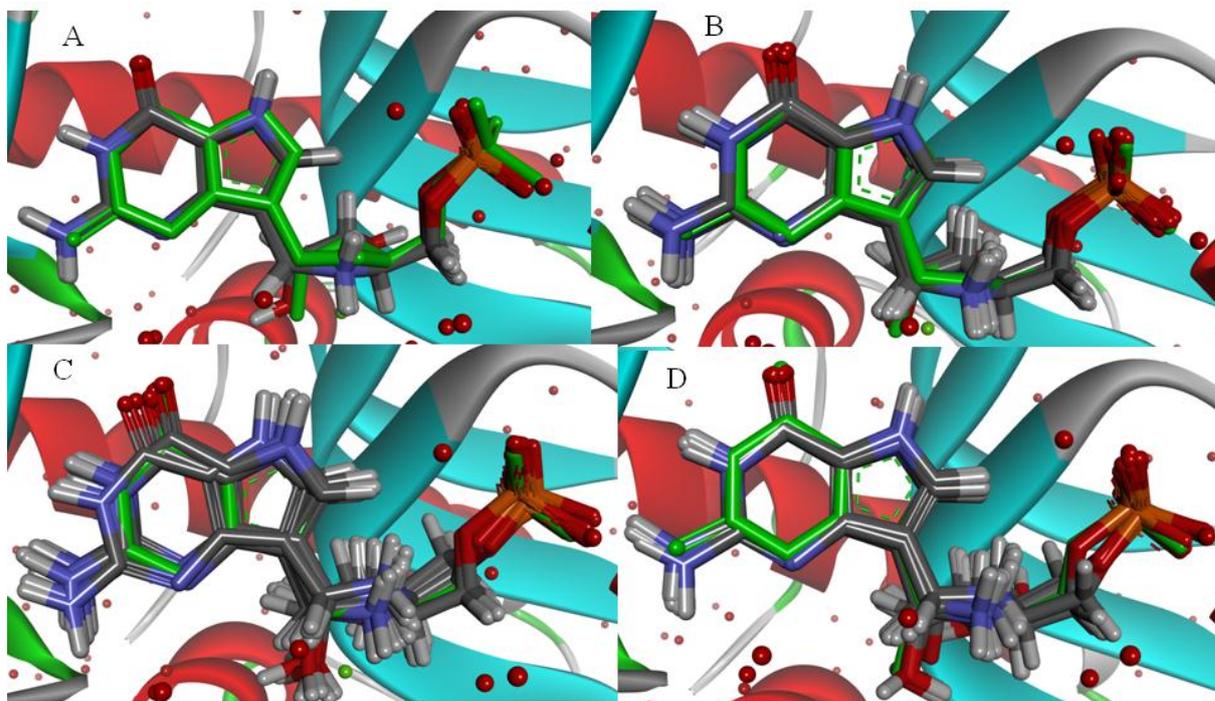


Figura 16. Análise da redocagem no GOLD utilizando as seguintes funções de escore: A: Goldscore; B: ChemScore; C; ASP; D: ChemPLP. ImuGP cristalográfica em verde.

Analisando os resultados de RMSD obtidos pelas diferentes funções de escore para as dez poses geradas, bem como a inspeção visual para cada função, podemos observar que todas as funções apresentaram valores de RMSD inferiores a 1,000 Å e mínimas diferenças conformacionais entre as poses. Por conseguinte, esperar-se-ia que todas as funções de escore, avaliadas sob os mesmos parâmetros, fossem adequadas para a triagem virtual. Resultados tão semelhantes nesta etapa de redocagem não permitem inferir qual o melhor algoritmo de escore, sendo necessária uma nova etapa de validação que avalie um maior número de ligantes, os quais devem abranger estruturas que se encaixem de modo semelhante à ImuGP e estruturas aleatórias, de modo a fornecer novos dados para subsidiar a comparação entre as diferentes funções.

5.3.2. Validação dos métodos de docagem - cálculo do fator de enriquecimento (EF):

A validação dos parâmetros de triagem virtual através do cálculo do EF foi realizada com um banco contendo 348 estruturas, construído a partir de 300 compostos da ZINC e outros 48 ligantes conhecidos. Dos 348 compostos docados, foram analisados somente os 70 compostos mais bem ranqueados, número que corresponde à aproximadamente 20% do tamanho desta base. Cabe salientar ainda que estes 300 ligantes aleatoriamente selecionados correspondem à aproximadamente 25% das 1.230 estruturas obtidas no servidor da ZINC por triagem farmacofórica.

A avaliação dos resultados obtidos através do programa Glide com relação ao posicionamento dos ligantes conhecidos, dispostos numa amostra de 70 compostos (N_a) analisados num banco de ligantes de 348 estruturas ao todo (N_t), revelou a existência de 33 ligantes conhecidos ($Hits_a$) num universo total de 48 ligantes ativos ($Hits_t$). O resultado encontrado para o $EF_{20\%}$ foi igual a 3,42 (**Equação 2**):

$$EF_{20} = \frac{Hits_a}{N_a} / \frac{Hits_t}{N_t} = EF = \frac{33}{70} / \frac{48}{348} = 3,42 \quad Eq. (2)$$

A análise visual das melhores poses dos ligantes ativos demonstra uma ótima correlação com a orientação do ligante cristalográfico. De modo geral, os compostos apresentaram uma boa sobreposição do núcleo 9-deazaguanina, enquanto a região do ciclo iminoribitol e do fosfato apresentam ligeiras alterações, decorrente das diferentes conformações adotadas pelos carbonos sp^3 (**Figura 17**).

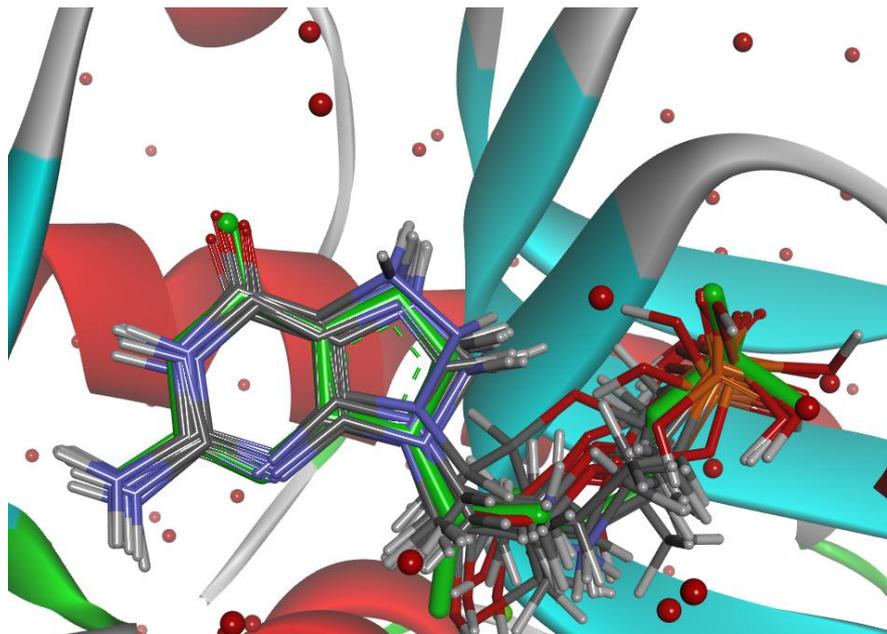


Figura 17. Análise visual das poses dos ligantes ativos realizada pela docagem no Glide. ImuGP cristalográfica em verde.

A avaliação dos resultados do posicionamento dos ligantes conhecidos (Hits_a) obtidos através do programa GOLD, bem como o EF_{20%}, calculados para cada função de escore com ou sem opção de rescore, encontram-se dispostos na **Tabela 4**:

Tabela 4. Valores calculados dos fatores de enriquecimento encontrados para as diferentes funções de escore.

Funções de Escore	Ligantes conhecidos (Hits _a)	Fator de enriquecimento (EF)
Goldscore	42	4,35
Chemscore	30	3,11
ASP	48	4,97
ChemPLP	44	4,55
Goldscore e Chemscore	35	3,63
Goldscore e ASP	43	4,45
Goldscore e ChemPLP	42	4,35
Chemscore e Goldscore	37	3,83
Chemscore e ASP	39	4,04
Chemscore e ChemPLP	36	3,73
ASP e Goldscore	46	4,76
ASP e Chemscore	30	3,11
ASP e ChemPLP	39	4,04
ChemPLP e Goldscore	41	4,25
ChemPLP e Chemscore	35	3,63
ChemPLP e ASP	44	4,56

Embora o EF seja uma ferramenta bastante útil para medir o “enriquecimento” da fração de compostos ativos quando comparados à distribuição randômica, deve-se ter em mente que este recurso apresenta algumas limitações. Como se pode prever através da análise da **Equação 2**, o EF é altamente dependente da proporção de compostos ativos em relação ao número total de compostos numa base de dados e, portanto, valores de EF obtidos utilizando-se diferentes proporções de compostos ativos não podem ser comparados entre si. Desse modo, somente valores de EF calculados para uma mesma base de dados podem ser comparados para se medir o desempenho de uma triagem virtual. Além disso, a equação de EF não confere diferentes “pesos” que favoreçam uma melhor classificação das moléculas mais ativas em detrimento daquelas com menor atividade e, portanto, não permite discriminar, por exemplo, entre uma situação em que os compostos ativos estão no topo da lista contra uma em que estes foram classificados nas últimas posições se todos estes foram recuperados dentro da amostra analisada (MARKT; SCHUSTER; LANGER, 2011).

A análise visual das poses individuais foi semelhante entre as diferentes funções de escore, de modo a dispor os ligantes ativos em orientação semelhante ao ligante cristalográfico ImuGP, diferindo em algumas conformações mais lábeis como o anel iminorbitol e o fosfato, conforme se pode observar na **Figura 18**, a qual reproduz as cinco melhores poses dos ligantes ativos em quatro funções de escore diferentes (Goldscore, ChemScore, ASP e ChemPLP). Desse modo, podemos inferir que o grande diferencial entre os algoritmos de escore encontra-se na forma de escalonamento das poses. Através da **Tabela 4**, podemos perceber que a função ASP apresentou a melhor capacidade de escalonar os ligantes ativos, sendo capaz de recuperar todos os 48 ligantes dispostos numa base de análise de 70 poses.

Conforme avaliado por Cheng e colaboradores num estudo virtual empregando 16 funções de escore testando uma base contendo 195 complexos proteína-ligante, a função ASP presente no programa GOLD apresentou a melhor "força de docagem", ou seja, a melhor capacidade de prever as melhores poses baseadas nos melhores escores dentre os demais métodos, com uma taxa de sucesso de 82,5% considerando um RMSD inferior a 2.0 Å da pose cristalográfica (CHENG *et al.*, 2009). Sendo assim, escolheu-se a função ASP como padrão para as análises posteriores. Além disso, foi escolhida a função GoldScore, com intuito de comparação, por ser a função padrão do programa.

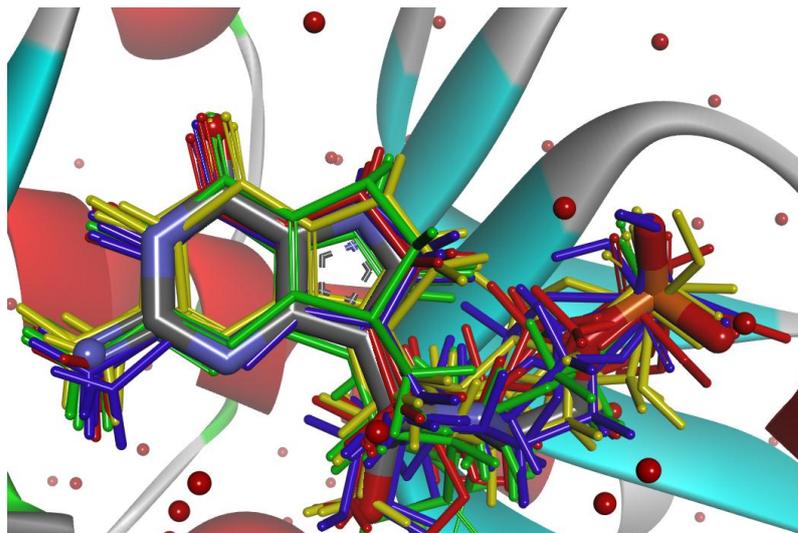


Figura 18. Análise visual das poses obtidas por diferentes métodos de escore no GOLD. ImuGP cristalográfica em cores normais; GoldScore em amarelo; ASP em azul; ChemPLP em verde; ChemScore em vermelho.

5.3.3. Avaliação da influência das águas no processo de docagem através dos valores EF:

A avaliação da influência das águas HOH 541 e HOH 642 no processo de docagem foi realizada através dos dois programas Glide e GOLD utilizando as melhores funções de escore, GlideScore e GOLD-ASP, definidas através dos procedimentos anteriores. Para tanto deletaram-se as duas águas e se avaliou a capacidade dos programas em encontrar as 48 estruturas ativas presentes no banco de ligantes, contendo 348 estruturas, da mesma forma que foi realizada no procedimento de validação.

Os procedimentos realizados no programa Glide através da função GlideScore apresentaram uma ligeira diferença nos valores do fator de enriquecimento, resultando em $EF_{20\%}$ igual a 3,42 para o procedimento considerando as águas e 3,73 para o procedimento sem as águas. Por outro lado, os procedimentos realizados no programa GOLD através da função ASP apresentaram valores de $EF_{20\%}$ iguais a 4,97 para o procedimento considerando as águas e 4,25 para o procedimento desconsiderando as águas.

A presença de águas no sítio de ligação à proteína pode levar ao aumento das interações entre o ligante e os resíduos de aminoácidos da proteína, uma vez que a água pode atuar tanto com aceptor quanto doador de ligações de hidrogênio, realizando múltiplas interações e contribuindo para o aumento da afinidade entre ligante e receptor. No entanto, a maioria dos protocolos de docagem para "screening" virtual ignora os efeitos da água no sítio de ligação, uma vez que os efeitos termodinâmicos e estruturais são difíceis de serem determinados, dificultando assim a modelagem (LADBURY, 1996).

Além disso, resultados conflitantes são relatados na literatura, demonstrando que ainda não há um consenso sobre o papel das águas nesses procedimentos (SANTOS *et al.*, 2010). Em estudo conduzido por Huang e colaboradores, a precisão de docagem foi ampliada com a adição das águas no sítio de ligação (HUANG; SHOICHET, 2008). Por outro lado, o estudo realizado por Birch e colaboradores não demonstrou melhora na precisão (BIRCH *et al.*, 2002). Tendo em vista que se espera que os ligantes extraídos do banco de dados da ZINC ocupem a posição da água HOH 541, pode-se considerar adequado excluir as duas moléculas de água do sítio ativo da enzima para as etapas de docagem posteriores.

5.3.4. Docagem das estruturas obtidas do banco de dados da ZINC, avaliação visual das poses e cálculo do Fator de Enriquecimento (EF):

O procedimento de docagem dos 1.278 compostos, sendo que desses 1.230 estruturas foram extraídas da base de dados da ZINC, foi desenvolvido utilizando os programas Glide e GOLD por meio das funções de escore GlideScore, Goldscore e ASP. Os parâmetros validados anteriormente foram utilizados e as águas HOH 541 e HOH 642 foram removidas do sítio de docagem.

Os resultados encontrados pelo programa Glide apresentaram um fator de enriquecimento $EF_{20\%}$ igual a 3,64, encontrando 35 ligantes ativos dentro de um universo amostral de 256 estruturas. Os ligantes ativos apresentaram-se em poses semelhantes ao ligante cristalográfico (**Figura 19**). Ao observar as quatro melhores poses dos ligantes da ZINC (**Figura 20, A**) percebe-se uma clara sobreposição dos grupos aromáticos nas proximidades da 9-deazaguanina da ImuGP. Além disso, as hidroxilas desses ligantes sobrepõem-se à hidroxila 3' do anel iminorribitol da ImuGP, podendo promover interações com os metais do sítio. Ademais, esses ligantes apresentam também um grupo doador de ligações de hidrogênio (amina ou hidroxila) em posição semelhante ao nitrogênio N2 exociclo, os quais podem realizar interações com os resíduos de Leu181 e Asp187 (**Figura 20, B**). Cabe salientar ainda que o ligante ZINC03882544 apresenta uma hidroxila distante 3,07 Å do grupo ϵ -amino da Lys152, a qual se encontra mais próxima que a água HOH541 que está a 3,10 Å do grupo amino (**Figura 20, C**).

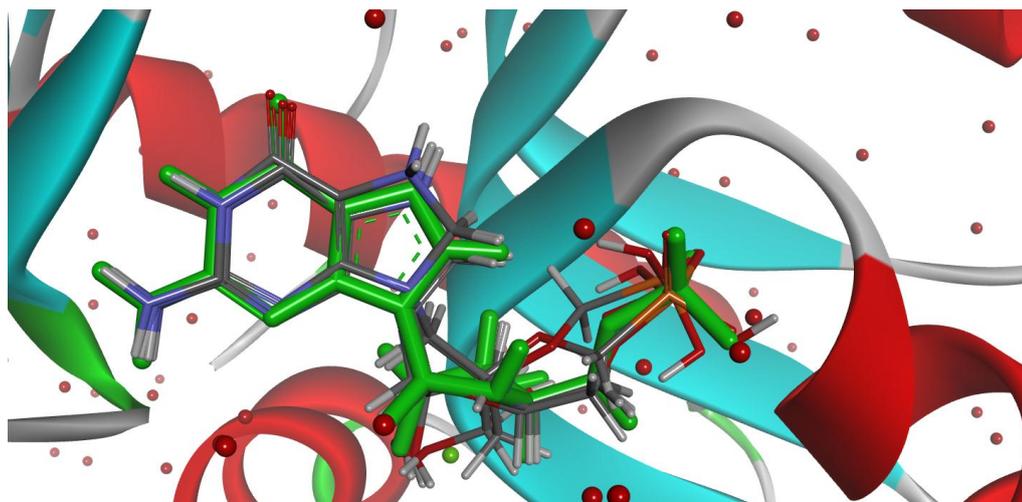


Figura 19. Análise da pose dos ligantes ativos na docagem realizada pelo Glide. Foram destacados os quatro melhores ligantes. ImuGP destacada em verde.

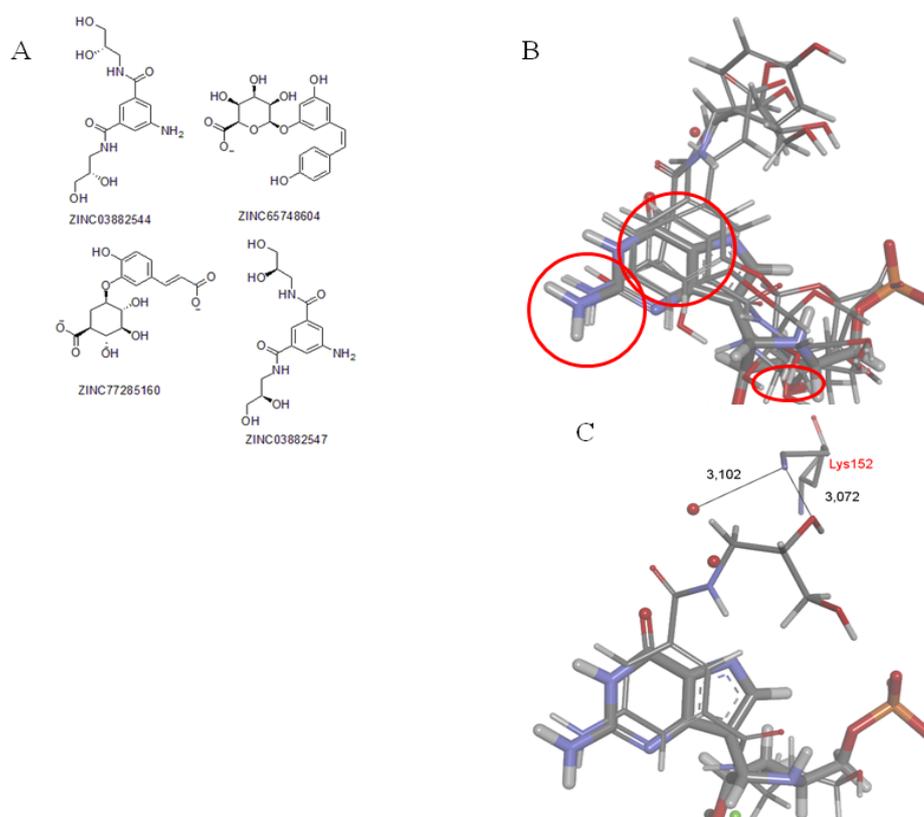


Figura 20. Análise dos melhores ligantes da ZINC obtidos pelo Glide. **A)** Estruturas químicas dos ligantes da ZINC. **B)** Sobreposição dos quatro ligantes em conjunto com ImuGP. Em destaque a sobreposição dos nitrogênios em N2, anéis aromáticos e hidroxilas na posição 3' do anel iminoribitol. **C)** Análise da distância entre a hidroxila do ligante ZINC03882544 e a Lys152.

Por outro lado, os resultados encontrados pelo GOLD utilizando duas funções de escore diferentes (ASP e GoldScore) apresentaram significativas alterações visuais. Conforme observado na **Figura 21, A**, que corresponde ao algoritmo ASP, pode-se perceber que os ligantes ativos com melhor escalonamento encontram-se, em geral, nas mesmas poses em comparação com o ligante cristalográfico ImuGP, havendo algumas pequenas alterações no anel iminorribitol e no fosfato. No entanto, ao se observar as poses dos ligantes ativos melhores ranqueados pelo algoritmo GoldScore, foram encontradas poses fora da orientação esperada, como a estrutura 3GEP e 3GGJ (**Figura 21, B**).

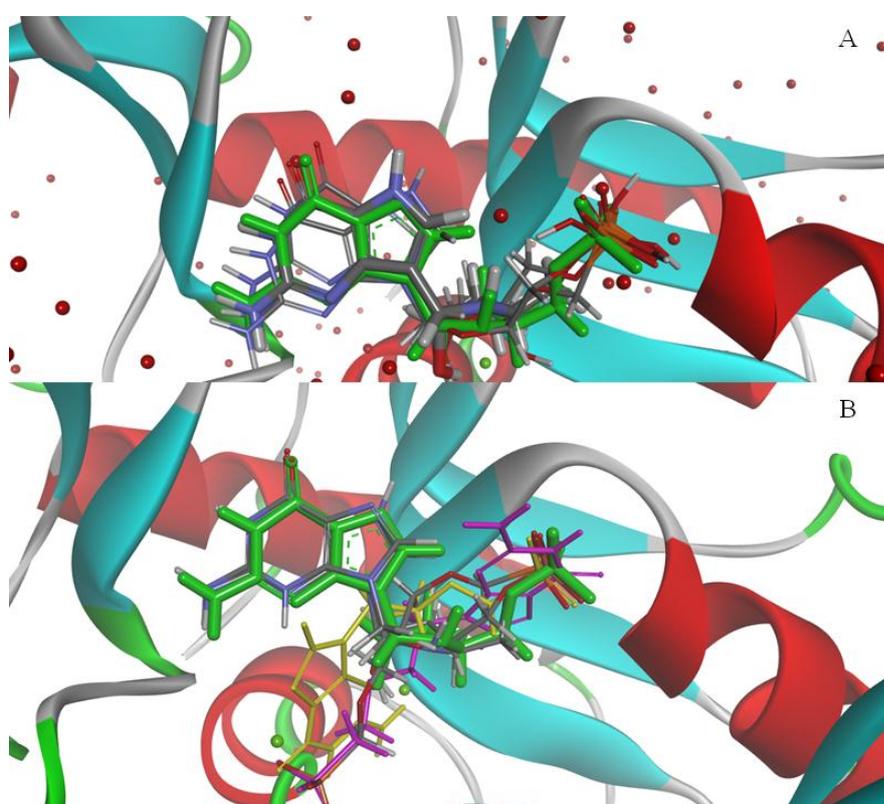


Figura 21. A) Análise visual dos ligantes ativos pela função ASP. B) Análise visual dos ligantes ativos pela função GoldScore. Amarelo = 3GEP e violeta = 3GGJ.

Os fatores de enriquecimento resultantes da docagem com ASP e Goldscore foram, respectivamente, de 4,68 (com 45 ligantes ativos encontrados) e de 2,50 (com 24 ligantes ativos encontrados). No entanto, ao avaliar a pose dos ligantes da ZINC pode-se perceber que há uma inversão na qualidade das conformações adotadas, ou seja, os ligantes da ZINC docados pela função ASP (**Figura 22, A**) encontram-se em poses não tão favoráveis quanto as encontradas pela função GoldScore (**Figura 22, B**).

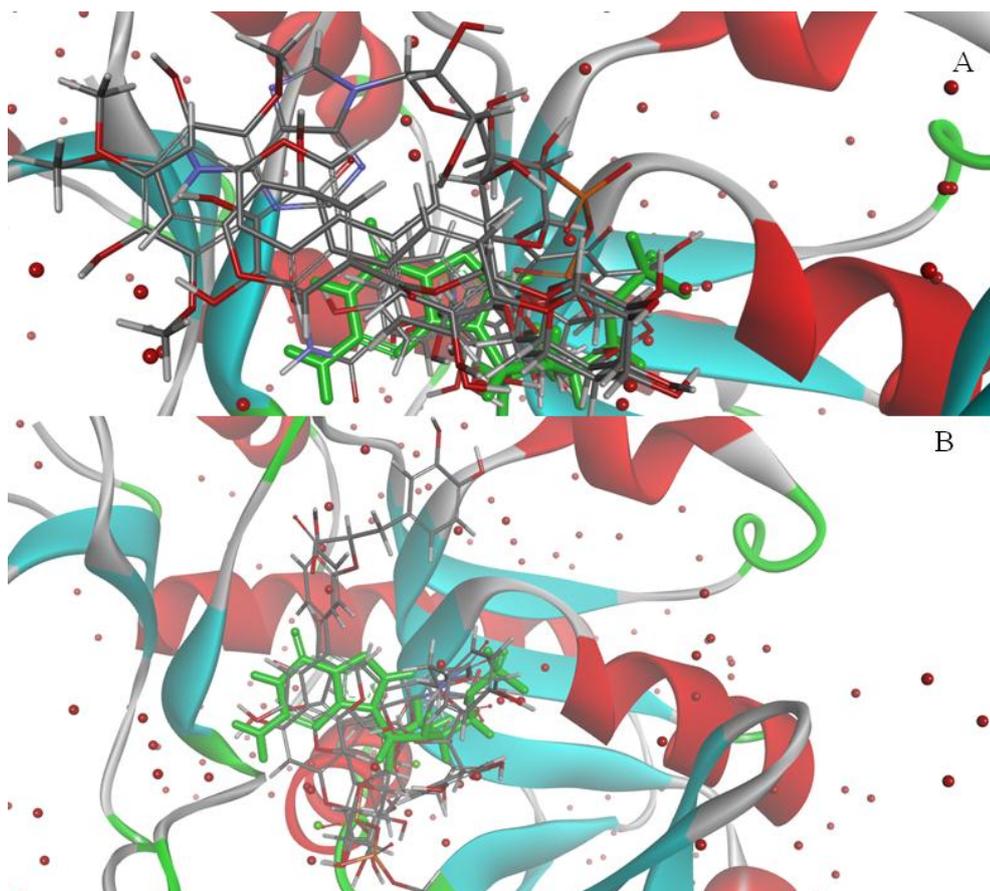


Figura 22. A) Avaliação da pose dos ligantes da ZINC pela função ASP. B) Avaliação da pose dos ligantes da ZINC pela função GoldScore.

Tal diferença pode ser dada pela maior tendência da função ASP em privilegiar as interações energéticas dos anéis aromáticos. Tal fato foi observado por Mooij e colaboradores durante o desenvolvimento do algoritmo ASP, através da utilização de diversos banco de dados, dentre eles alguns ligantes da enzima CDK2 (*Cyclin-dependent kinase 2*), os quais dispunham carbonos aromáticos nas proximidades dos resíduos de fenilalanina da enzima (MOOIJ *et al.*, 2005). Observa-se que os quatro ligantes da ZINC melhores ranqueados (ZINC08143469, ZINC67910690, ZINC67910684 e ZINC77257212) encontram-se direcionados à porção exterior da proteína, situado nas proximidades do Trp180 com sobreposição dos anéis aromáticos, corroborando com o dado encontrado por Mooij e colaboradores (**Figura 23, A e B**). Sendo assim, utilizaram-se os resultados da docagem realizado pela função de escore GoldScore.

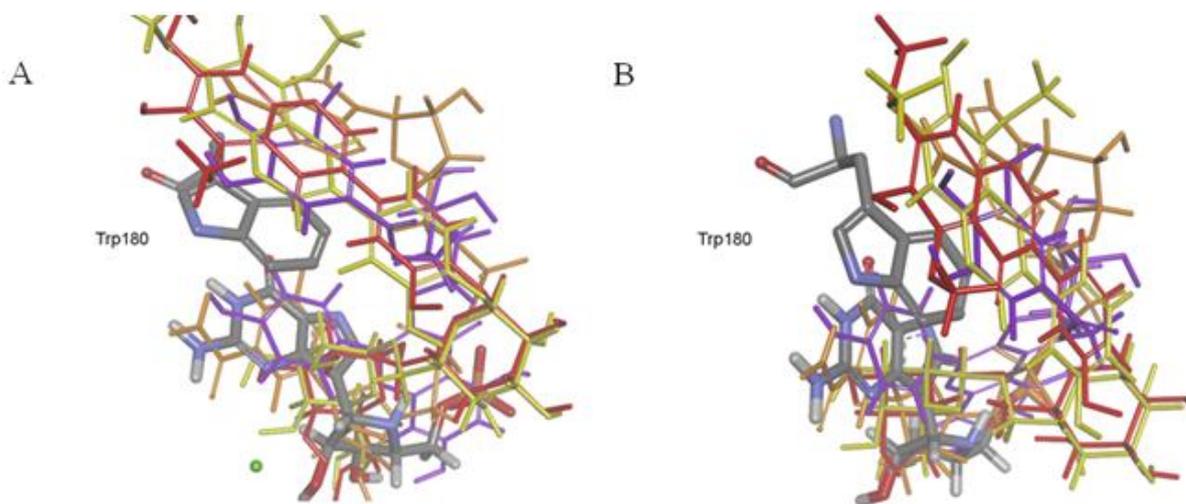


Figura 23. Análise visual das poses dos quatro ligantes da ZINC melhores ranqueados pela função ASP e as proximidades do aminoácido Trp180. **A)** Visão fronta. **B)** Visão Lateral.

5.3.5. Procedimento de docagem consensual dos resultados dos programas Glide e GOLD

O procedimento de docagem consensual levou à obtenção das 62 melhores estruturas comuns aos dois programas considerando as duas melhores funções de escore validadas anteriormente (Glidescore e GoldScore). Os resultados dos ranqueamentos obtidos pelos dois algoritmos de escore foram determinados através da média aritmética simples entre as posições de cada estrutura conforme os valores obtidos pelas funções (**Tabela 5**).

A análise visual dos ligantes revela, em geral, que não foi observado um consenso visual entre as poses geradas pelos dois programas Glide e Gold, considerando o número de poses utilizadas e utilizando os algoritmos de escore Glidescore e GoldScore, respectivamente. Uma visualização geral das poses das cinco estruturas melhores ranqueadas pode ser apreciada na **Figura 24**.

A inspeção visual das poses permitiu a identificação das melhores estruturas, conforme dispostas a seguir: ZINC77269479, ZINC77257212, ZINC67913696, ZINC67902474 e ZINC38143554. De modo geral, as poses obtidas pelo algoritmo GoldScore apresentam-se numa melhor orientação, mesmo que ainda não possuam a conformação mais plausível, se comparadas às obtidas pelo Glidescore, conforme observado na **Figura 25**. No entanto, observa-se uma maior tendência do algoritmo GoldScore em docar os ligantes na região próxima ao pirofosfato, como avaliado para o ligante ZINC77269479, discutido a seguir.

Tabela 5. Resultados gerais da docagem consensual obtidas pelos programas Glide e GOLD com determinação do ranque médio.

Composto	Docking score	Ranque Glide	Goldscore fitness	Ranque Gold	Ranque médio
ZINC77269479	-8,98239	9	126,15	2	5,5
ZINC08143469	-6,11325	39	123,777	3	21
ZINC77285160	-10,1469	3	106,863	68	35,5
ZINC85867123	-7,39233	24	106,568	71	47,5
ZINC77257200	-5,97743	43	107,306	62	52,5
ZINC38139496	-6,72936	32	106,157	77	54,5
ZINC04273365	-5,36885	111	112,908	15	63
ZINC67910269	-5,82587	51	106,177	76	63,5
ZINC38143831	-7,27457	26	104,69	101	63,5
ZINC38143765	-5,57885	77	107,838	54	65,5
ZINC39804757	-5,48093	91	108,408	47	69
ZINC38143554	-7,4983	23	103,485	118	70,5
ZINC04098612	-5,91221	46	104,689	102	74
ZINC31158983	-6,71729	33	103,602	115	74
ZINC77269423	-5,42746	99	108,177	50	74,5
ZINC85867118	-5,73716	55	104,79	96	75,5
ZINC49601560	-5,64787	67	105,788	86	76,5
ZINC49601757	-5,19085	134	112,43	20	77
ZINC12428286	-5,17796	137	112,276	21	79
ZINC38139803	-7,28387	25	102,809	138	81,5
ZINC33861449	-5,5622	78	105,104	92	85
ZINC67913072	-5,72183	60	103,612	114	87
ZINC67910405	-5,08662	166	117,365	9	87,5
ZINC67912780	-6,98134	29	102,062	159	94
ZINC67903538	-5,09673	159	109,784	35	97
ZINC95486360	-5,04395	182	113,469	12	97
ZINC38139723	-5,37477	110	105,641	87	98,5
ZINC77257212	-5,83762	50	102,408	151	100,5
ZINC03915684	-6,17809	38	101,839	164	101
ZINC67902474	-5,13074	148	107,792	56	102
ZINC58900430	-5,66239	65	102,391	152	108,5
ZINC77257387	-5,23542	130	104,733	99	114,5
ZINC67911931	-4,96923	203	111,417	27	115
ZINC15721409	-5,06896	173	107,438	58	115,5
ZINC09212671	-5,09185	164	106,078	79	121,5
ZINC77269667	-5,0949	161	105,02	93	127
ZINC05923208	-6,93321	30	98,5509	246	138
ZINC57761799	-5,41454	103	101,379	172	137,5
ZINC33861402	-4,89594	221	107,048	66	143,5
ZINC04252672	-5,06449	176	103,555	116	146
ZINC08681504	-5,7895	53	98,6295	243	148
ZINC04349577	-4,88958	225	106,409	73	149
ZINC13424718	-5,1493	143	102,309	155	149
ZINC71774753	-4,80664	248	107,409	59	153,5

ZINC72341223	-5,25927	122	100,696	186	154
ZINC13424741	-5,0297	186	103,171	127	156,5
ZINC67913696	-5,45923	92	99,3843	223	157,5
ZINC31167917	-4,93164	212	103,764	112	162
ZINC04098544	-5,41439	105	99,408	222	163,5
ZINC08143568	-5,44924	93	98,8435	237	165
ZINC59816865	-4,89537	222	104,417	108	165
ZINC38139797	-4,81497	246	105,856	84	165
ZINC77269228	-4,95209	208	102,815	136	172
ZINC70691536	-5,26578	121	99,1672	230	175,5
ZINC12428583	-5,30336	115	98,811	239	177
ZINC03978800	-5,01178	193	101,375	173	183
ZINC04349517	-5,1396	146	99,307	226	186
ZINC05023196	-5,03948	185	100,429	190	187,5
ZINC04349687	-4,8343	241	102,756	140	190,5
ZINC85473422	-4,8186	245	102,504	148	196,5
ZINC38160904	-4,9574	207	100,334	196	201,5
ZINC08035395	-5,01229	192	98,9679	234	213

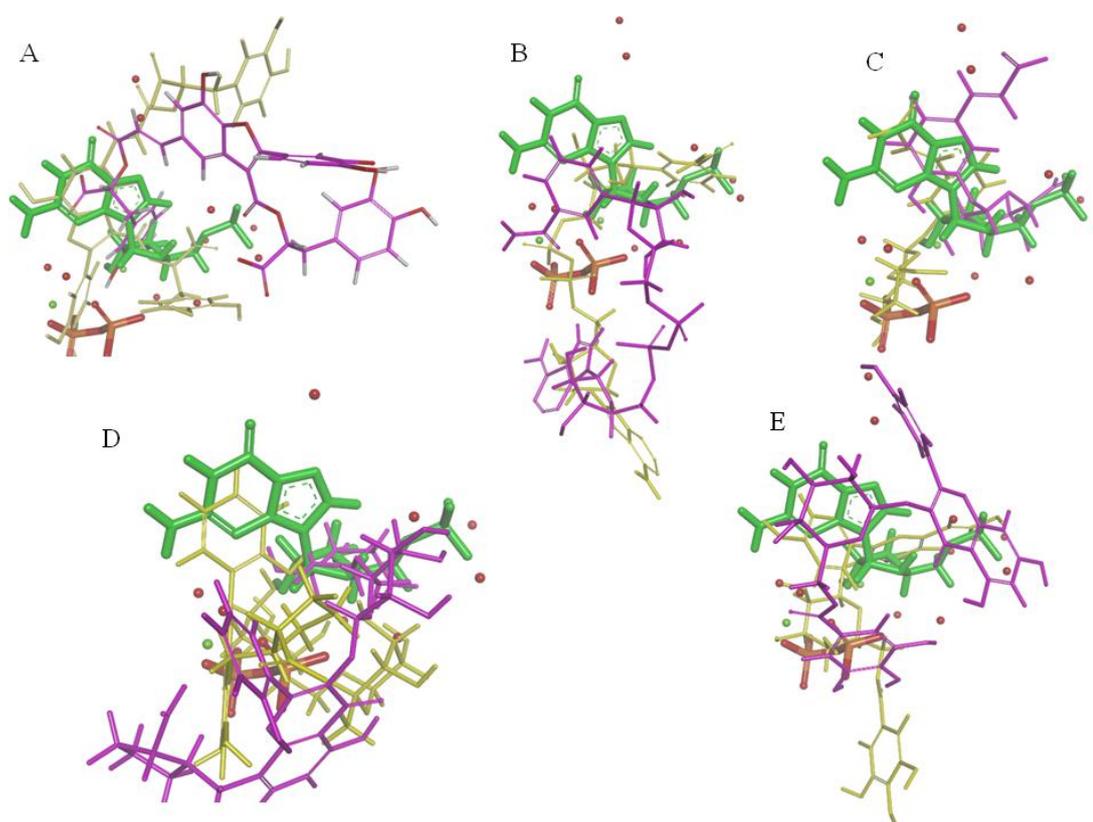


Figura 24. Análise visual das poses dos cinco melhores ligantes obtidos por docagem consensual. **A)** ZINC77269479; **B)** ZINC08143469; **C)** ZINC77285160; **D)** ZINC85867123) **E)** ZINC77257200. Verde: ImuGP; Amarelo: Pose Gold; Roxo: Pose Glide.

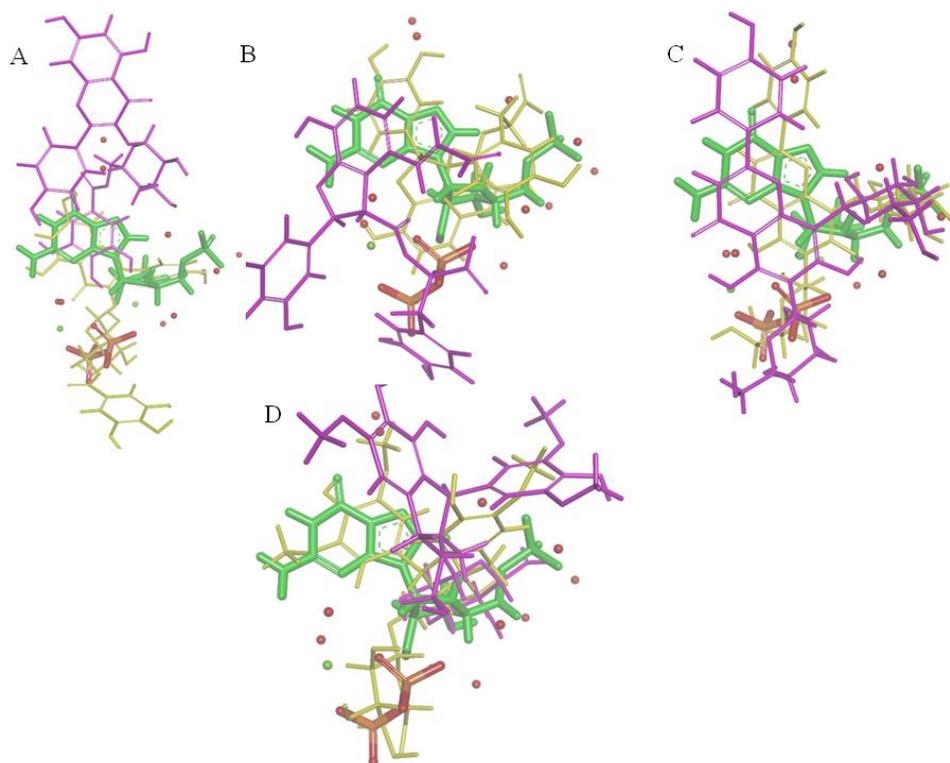


Figura 25. Inspeção visual das melhores poses encontradas pela função GoldScore. **A)** ZINC77257212; **B)** ZINC67913696; **C)** ZINC67902474; **D)** ZINC38143554. Verde: ImuGP; Amarelo: GoldScore; Roxo: Glidescore.

Uma análise detalhada da pose obtida através do GoldScore para o melhor ligante em potencial, ZINC77269479 (**Figura 26, A**), demonstra a manutenção das interações aromáticas com o Trp180, assim como das ligações de hidrogênio com Leu181 e Asp187 (**Figura 26, B**), as quais apresentam grande importância para garantia de especificidade da enzima pelo substrato. Ademais, observa-se a existência da ligação de hidrogênio entre o grupo ϵ -amino da Lys152 e o oxigênio carbonílico do ligante com distância igual a 3,30 Å. Salienta-se ainda, o surgimento de uma interação não observada anteriormente entre o mesmo grupo ϵ -amino da Lys152, o qual se encontra protonado, com o grupo aromático do ligante, caracterizando uma interação cátion- π , distante 1,5 Å (**Figura 26, C**). Tal interação pode ser de grande valia para aumento da especificidade frente à enzima do parasita, tendo em vista que, energeticamente, as interações cation- π são equiparáveis às ligações de hidrogênio usuais (ZACHARIAS; DOUGHERTY, 2002). No entanto, também se observa uma região de impedimento estérico entre um grupo aromático do ZINC77269479 e o pirofosfato, o qual é utilizado como substrato da reação de síntese do GMP. Tal impedimento pode ser desfavorável para ligação à enzima, uma vez que o grupo aromático não realiza interações semelhantes ao pirofosfato, sobretudo interações de coordenação com os metais.

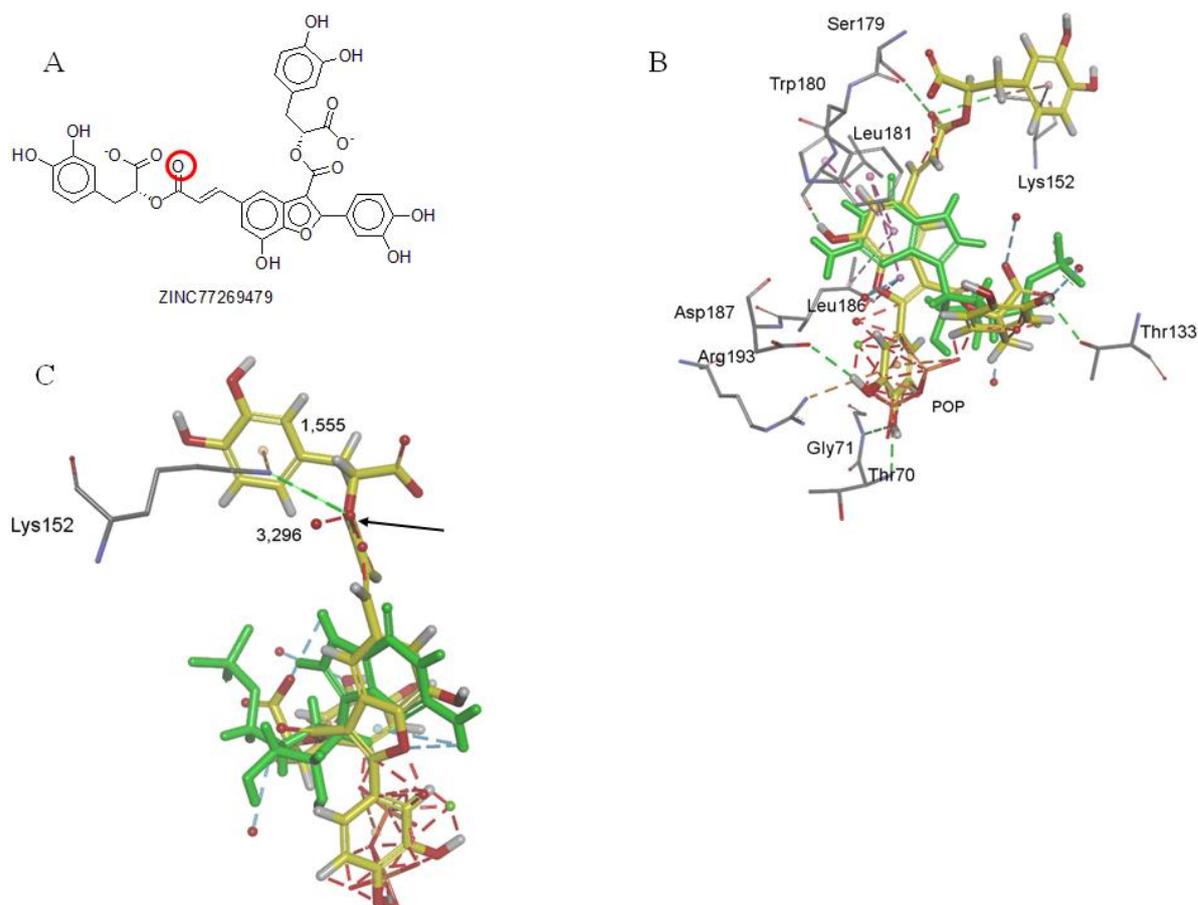


Figura 26. Análise detalhada da pose do ligante ZINC77269479 (Amarelo) em comparação à ImuGP (Verde). **A)** Estrutura química do ligante indicando a carbonila que interage com a Lys152. **B)** Análise das interações com os aminoácidos do sítio de ligação da GPR152. **C)** Visualização das distâncias entre a carbonila e o grupo ϵ -amino da Lys152.

Por outro lado, análise da pose gerada pelo Glidescore revela uma orientação diferente, voltada para o exterior da proteína. Como se pode perceber pela **Figura 27**, não há uma manutenção das interações aromáticas e das principais ligações de hidrogênio, com exceção da Leu181, como definido anteriormente (SHI *et al.*, 2000). Destacam-se algumas interações aromáticas com His132 e Tyr127, bem como interações de hidrogênio entre o grupo α -amino da Leu181 e o oxigênio carbonílico do ligante, com distância de aproximadamente 2,98 Å, e entre o grupo ϵ -amino da Lys98 com o carboxilato do ligante, com distância de aproximadamente 2,90 Å (**Figura 27**).

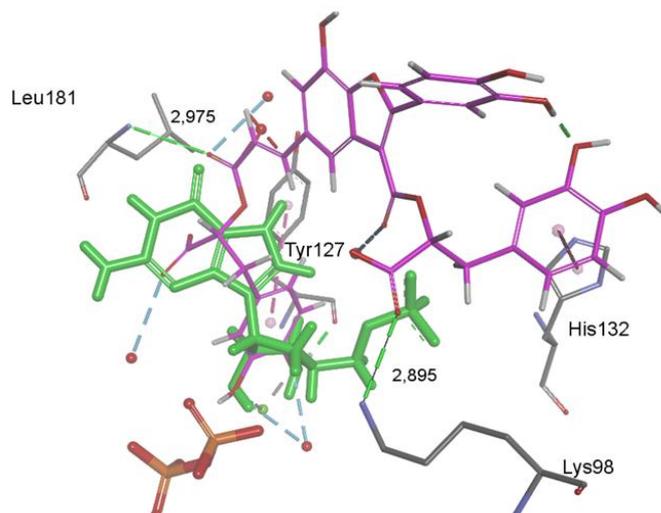


Figura 27. Análise da pose do ligante ZINC77269479 (em roxo) gerada por docagem pelo Glide e as principais interações com os resíduos de aminoácido do sítio da GPRTase. Verde = ImuGP.

Os demais ligantes obtidos através do GOLD (ZINC77257212, ZINC67913696, ZINC67902474 e ZINC38143554) apresentaram ao menos duas principais interações com o sítio de ligação à GPRTase, em geral interações aromáticas (empilhamento π) com o resíduo Trp180 e ligações de hidrogênio com Leu181. Cabe salientar que apenas o composto ZINC67902474 apresentou também interações com a Lys152. Uma análise mais detalhada pode-ser vista com base na **Figura 28**:

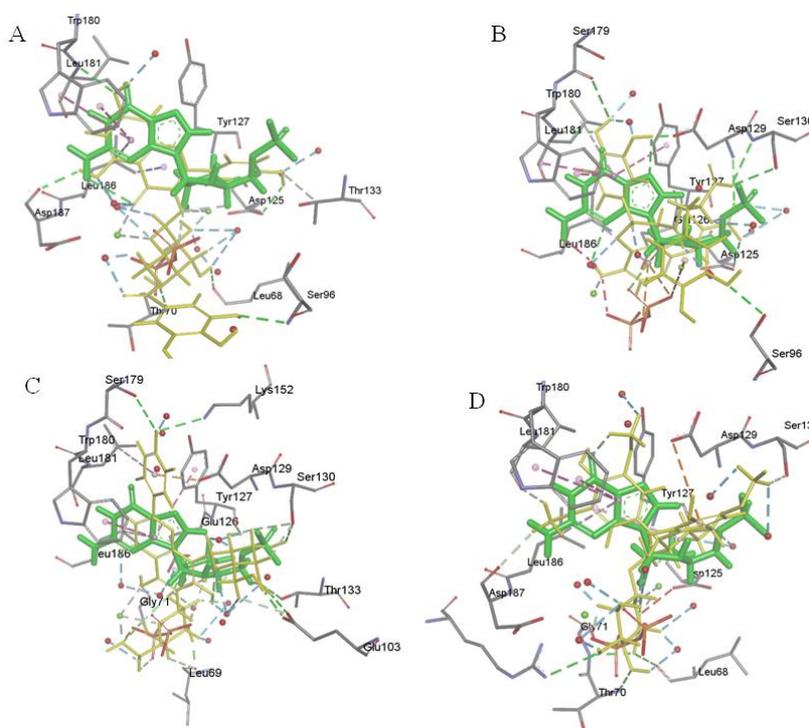


Figura 28. Análise das interações entre os principais ligantes e os resíduos de aminoácido da GPRTase. **A)** ZINC77257212; **B)** ZINC67913696; **C)** ZINC67902474; **D)** ZINC38143554.

O principal problema relacionado aos estudos de triagem virtual é que muitos dos compostos que são preditos como ativos nos ensaios de docagem acabam se mostrando inativos quando avaliados experimentalmente (PEACH; NICKLAUS, 2009). Uma análise extensiva de dez dentre os principais softwares utilizados para triagem virtual por docagem demonstrou que estes programas são eficientes em prever a conformação do ligante e sua posição/orientação no sítio ativo das proteínas mas falham na tentativa de ranquear os compostos mais ativos (WARREN et al., 2006). Isso ocorre porque as chamadas funções de escore, nas quais estes programas se baseiam para prever a afinidade de ligação, são planejadas para funcionarem em um vasto conjunto de alvos moleculares e podem não ser tão precisas para um determinado alvo em específico (KROEMER, 2007).

Uma das formas de se contornar este problema é através da chamada docagem consensual, através do qual os compostos selecionados como mais prováveis de serem ativos não são aqueles que apresentam os maiores valores de escore mas sim os que são preditos como ativos e que são comuns entre as triagens realizadas com dois ou mais programas contendo diferentes funções de escore. Consequentemente, haveria neste caso uma maior probabilidade de que os compostos selecionados sejam realmente ativos quando testados experimentalmente (PAUL; ROGNAN, 2002).

No entanto, para a base de dados de 1.230 compostos extraída da ZINC com base nos quatro grupos farmacofóricos não foi possível estabelecer um consenso visual adequado para os procedimentos de docagem empregando o algoritmo genético do Gold, em conjunto com o GoldScore, e o algoritmo empírico do Glide, em conjunto com o Glidescore. Uma das prováveis causas para a não obtenção do consenso deve-se ao número reduzido de poses que foram geradas em cada um dos programas de docagem visto que, se as funções de escore normalmente falham em ranquear corretamente os compostos mais ativos dentro de uma base de dados, estas certamente também são suscetíveis a não ordenar corretamente as poses mais prováveis para um dado composto. Ao estudar o provável modo de ligação do haloperidol frente ao receptor dopaminérgico D3 por docagem molecular, Zanatta e colaboradores (2014) geraram 1.000 poses para o mesmo ligante, agrupando as poses semelhantes ($\text{RMSD} \leq 1\text{\AA}$) em 13 categorias diferentes. É bastante provável que a adoção de uma abordagem semelhante neste trabalho venha resultar em melhor consenso das poses geradas pelos dois programas.

Por outro lado, ligantes contendo inúmeras ligações rotacionáveis, como alguns dos observados neste trabalho também impõem uma maior dificuldade aos programas de docagem no encaixe adequado das estruturas no sítio de ligação da enzima, permitindo uma variação muito ampla das conformações e orientações das poses. A adoção de uma nova base de dados

contendo ligantes menores e que apresentem uma maior facilidade de encaixe ao sítio de ligação da enzima está sendo estudada a partir do estabelecimento de um novo modelo farmacofórico.

5.3.6. Análise das propriedades farmacocinéticas dos compostos mais promissores:

A análise das propriedades farmacocinéticas dos cinco compostos mais promissores através do programa Volsurf+ avaliou os descritores relacionados à absorção intestinal em células CACO2 (CACO2) e à permeação da barreira hematoencefálica (LogBB). Os valores obtidos para esses compostos encontram-se dispostos na **Tabela 6**:

A predição da absorção intestinal em células CACO2 realizada pelo Volsurf+ considera valores de escore +1 (para uma permeação superior a 8×10^{-6} cm/s), indicando a possibilidade de permeação passiva através do modelo, e valores de escore -1 (para uma permeação inferior a 4×10^{-6} cm/s), indicando pouca possibilidade de permeação na célula epitelial. Por sua vez, a predição da permeação da barreira hematoencefálica realizada pelo Volsurf+ baseia-se num modelo quantitativo usando 500 compostos extraídos da literatura, os quais apresentaram os seguintes resultados: habilidade de passar a barreira hematoencefálica (LogBB > 0,5); permeação moderada (LogBB entre 0 e 0,5); pouca penetração (LogBB maior que -0,3) e muito pouca penetração (LogBB menor que -0,3).

Sendo assim, a análise dos resultados encontrados para predição de absorção intestinal em células CACO2 e da permeação da barreira hematoencefálica (LogBB) demonstram que tais moléculas apresentariam pouca ou nenhuma capacidade de absorção por via intestinal e nenhuma ou quase nenhuma capacidade de transpor a barreira hematoencefálica, fato o qual é desejável, uma vez que tais moléculas devem atuar a nível de lúmen intestinal não sendo, portanto, absorvidas. No entanto, cabe salientar que tais moléculas devem apresentar a capacidade de permear a membrana plasmática do parasita, para então promover a inibição enzimática proposta.

Tabela 6. Valores dos descritores farmacocinéticos de absorvidade intestinal (CACO2) e de permeação da barreira hematoencefálica (LogBB) para os cinco compostos mais promissores.

Nome	Estrutura química	CACO2	LogBB
ZINC77269479		-2,34	-4,66
ZINC77257212		-2,16	-3,48
ZINC67913696		-1,95	-4,19
ZINC67902474		-2,94	-4,13
ZINC38143554		-2,42	-4,41

6. CONCLUSÕES

O estudo do sítio de ligação da enzima GPRTase de *Giardia lamblia*, baseado no complexo cristalográfico contendo o inibidor do estado de transição ImuGP, bem como nos dados da literatura, indicou a existência de interações importantes para a especificidade da enzima frente à derivados da guanina, tais como: interações aromáticas do tipo empilhamento π com o triptofano 180 e ligações de hidrogênio com leucina 181 e aspartato 187. A análise comparativa da enzima humana HGPRTase indicou a existência de um resíduo de lisina (Lys165) distante 2,7 Å da carbonila do ligante, enquanto que o resíduo homólogo da enzima GPRTase encontra-se a aproximadamente 6,3 Å, demonstrando que haveria a possibilidade de explorar essa nova interação de modo a aumentar a seletividade frente à enzima do parasita, minimizando, assim, possíveis interações com a enzima humana.

A análise do sítio da GPRTase indica, ainda, a existência de duas águas (HOH541 e HOH642) que se encontram a 3,1 e 4,1 Å, respectivamente, da Lys152, possibilitando a utilização de tais coordenadas como ponto farmacofórico para realização de ligações de hidrogênio com o grupo ϵ -amino do aminoácido. Além disso, conforme predição através do servidor Depth e do Propka, o grupo ϵ -amino da Lys152 encontra-se ionizado em pH fisiológico, atuando possivelmente como um forte doador de hidrogênios. Sendo assim, construiu-se um modelo farmacofórico baseado em quatro pontos: um grupo aromático do anel de seis membros da purina (coordenadas XYZ: 18.72, 10.65, 53.62,); dois grupos doadores de ligações de hidrogênio, nitrogênio N2 exociclo (coordenadas XYZ: 17.30 9.09 55.33), Oxigênio 3' da ribose (coordenadas XYZ: 18.78, 8.46, 47.36); um grupo acceptor de ligações de hidrogênio determinado conforme posição da HOH541 (coordenadas XYZ: 19,18; 16,76; 52,76)

A realização da validação dos procedimentos de docagem para a enzima GPRTase possibilitou a escolha de três algoritmos de escore: Glidescore, GoldScore e ASP. Os parâmetros de escore foram adequados para a triagem inicial contendo o banco de ligantes de 348 estruturas e revelou influência da presença das águas no sítio de docagem, as quais foram eliminadas para as etapas posteriores. O procedimento de docagem final indicou que as funções de escore mais adequadas para a base de dados de 1.278 ligantes foram Glidescore e GoldScore, uma vez que parece que a função ASP apresenta a tendência de privilegiar interações aromáticas em detrimento das demais. Os resultados da docagem consensual não foram satisfatórios, uma vez que não foi possível obter um consenso visual entre as melhores poses obtidas pelos dois métodos de docagem e de escalonamento (Glide-Glidescore e Gold-

GoldScore). Tal impossibilidade talvez possa ter sido originada pelos compostos triados na base de dados, os quais, em sua maioria, apresentavam grande volume e muitas ligações rotacionáveis originando múltiplas conformações e orientações, dificultando assim o processo de docagem no sítio de ligação à enzima.

A análise visual das poses indicou a existência de cinco estruturas (ZINC77269479, ZINC77257212, ZINC67913696, ZINC67902474 e ZINC38143554) que apresentaram interações semelhantes à ImuGP, com destaque para os compostos ZINC77269479 e ZINC67902474, os quais apresentaram interações de hidrogênio com a Lys152. Por fim, a predição farmacocinética realizada através do programa Volsurf+ indicou que tais compostos não apresentariam a capacidade de permear as barreiras hematoencefálicas e da membrana plasmática das células intestinais, não sendo, portanto, capazes de sofrer absorção por via oral, fato que é desejável, uma vez que tais moléculas devem atuar no lúmen intestinal.

Tendo em vista os resultados obtidos, serão realizados novos procedimentos utilizando uma nova base de dados já compilada, a qual utilizará um novo modelo farmacofórico que desconsidera o doador de ligações de hidrogênio localizado na hidroxila O3' do anel iminoribitol e considera o nitrogênio protonado do mesmo anel como grupo catiônico, de modo a triar estruturas novas que apresentem menor volume e menor número de ligações rotacionáveis, o que poderá facilitar o procedimento de docagem consensual através dos programas Glide e GOLD. Será, ainda, realizada a análise do sítio de ligação da enzima HGPRtase humana pelo MetaPocket a fim de compará-lo com o sítio da GPRTase. Além disso, ao final da docagem consensual será aumentado o número de poses obtidas dos melhores compostos, as quais serão analisadas visualmente e agrupadas através de RMSD menor que 1Å. Em seguida, realizar-se-á um estudo detalhado através de docagem totalmente flexível (*Induced Fit Docking*) dos melhores ligantes, de modo a selecionar as melhores e as piores poses e escores para aquisição e teste *in vitro*, de maneira a comparar tais ligantes.

REFERÊNCIAS:

ACCELRY'S Discovery studio modeling environment. Versão 4.0. San Diego, CA.: Accelrys Software Inc., 2013.

ADAM, R. M. Biology of *Giardia lamblia*. **Clinical Microbiology Reviews**, v. 14, n. 3, p. 447-475, 2001.

ARONOV, A. M.; MUNAGALA, N. R.; KUNTZ, I. D.; WANG, C. C. Virtual screening of combinatorial libraries across a gene family: in search of inhibitors of *Giardia lamblia* guanine phosphoribosyltransferase. **Antimicrobial Agents and Chemotherapy**, v. 45, n. 9, p. 2571-2576, 2001.

AROOJ, M.; KIM, S.; SAKKIAH, S.; CAO, G. P.; LEE, Y.; LEE, K. W. Molecular modeling study for inhibition mechanism of human chymase and its application in inhibitor design. **PLoS One**, v. 8, n. 4, 2013. DOI: 10.1371/journal.pone.0060470.

BHAGAVAN, N. V.; HA, C-E. **Essentials of medical biochemistry: with clinical cases**. San Diego: Elsevier Inc., 2015. p. 465-489.

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**, v. 28, p. 235-242, 2000.

BIRCH, L.; MURRAY, C. W.; HARTSHORN, M. J.; TICKLE, I. J.; VERDONK, M. L. Sensivity of molecular docking to induced fit effects in influenza virus neuraminidase. **Journal of Computer-Aided Molecular Design**, v. 16, n. 12, p. 855-869, 2002.

CHAKRAVARTY, S.; VARADARAJAN, R. Residue depth: a novel parameter for the analysis of protein structure and stability. **Structure**, v. 7, n. 7, p. 723-732, 1999.

CHENG, A. C.; COLEMAN, R. G.; SMYTH, K. T.; CAO, Q.; SOULARD, P.; CAFFREY, D. R.; SALZBERG, A. C.; HUANG, E. C. Structure-based maximal affinity model predicts small-molecule druggability. **Nature Biotechnology**, v. 25, p. 71-75, 2007.

CHENG, T.; LI, X.; LI, Y.; LIU, Z.; WANG, R. Comparative Assessment of Scoring Functions on a Diverse Test Set. **Journal of Chemical Information and Modeling**, v. 49, p. 1079-1093, 2009

CHENG, T.; LI, Q.; ZHOU, Z.; WANG, Y.; BRYANT, S. H. Structure-based virtual screening for drug discovery: a problem-centric review. **AAPS Journal**, v. 14, n.1, p. 133-141, 2012.

CHOKSHI, D. A. Improving access to medicines in poor countries: the role of universities. **PLoS Medicine**, v. 3, n. 6, 2006. Disponível em: <<http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.0030136>>.

COHEN, J.; DIBNER, M. S.; WILSON, A. Development of and access to products for neglected diseases. **PLoS One**, v. 5, n. 5, 2010. Disponível em: <<http://www.plosntds.org/article/info:doi/10.1371/journal.pntd.0001750>>.

CORDINGLEY, F. T.; CRAWFORD, G. P. *Giardia* infection causes vitamin B₁₂ deficiency. **Australian and New Zealand Journal of Medicine**, v. 16, n. 1, p. 78–79, 1986.

CRIVORI, P.; CRUCIANI, G.; CARRUPT, P. A.; TESTA, B. Predicting blood-brain barrier permeation from three-dimensional molecular structure. **Journal of Medicinal Chemistry**, v. 43, n. 11, p. 2204-2216, 2000.

CRUCIANI, G.; CRIVORI, P.; CARRUPT, P. A.; TESTA, B. Molecular fields in quantitative structure-permeation relationships: the VolSurf approach. **Journal of Molecular Structure: Theochem**, v. 503, p. 17-30, 2000.

CRUCIANI, G.; PASTOR, M.; GUBA, W. VolSurf: a new tool for the pharmacokinetic optimization of lead compounds. **European Journal of Pharmaceutical Sciences**, v. 11, n. 2, p. 29-39, 2000.

DAVIDSON, B. L.; PALELLA, T. D.; KELLEY, W. N. Human hypoxanthine-guanine phosphoribosyltransferase: a single nucleotide substitution in cDNA clones isolated from a patient with Lesch-Nyhan syndrome (HPRTMidland). **Gene**, v. 68, n. 1, p. 85-91, 1988.

DOLINSKY, T. J.; NIELSEN, J. E.; MCCAMMON, J. A.; BAKER, N. A. PDB2PQR: an automated pipeline for the setup, execution, and analysis of poisson-boltzmann electrostatics calculations. **Nucleic Acids Research**, v. 32, p. W665-W667, 2004.

FRIESNER, R. A.; BANKS, J. L.; MURPHY, R. B.; HALGREN, T. A.; KLICIC, J. J.; MAINZ, D. T.; REPASKY, M. P.; KNOLL, E. H.; SHELLEY, M.; PERRY, J. K.; SHAW, D. E.; FRANCIS, P.; SHENKIN, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. **Journal of Medicinal Chemistry**, v. 47, n. 7, p. 1739-1749, 2004.

GARDNER, T. B.; HILL, D. R. Treatment of giardiasis. **Clinical Microbiology Reviews**, v. 14, n. 1, p. 114–128, 2001.

GASTEIGER, J.; RUDOLPH, C.; SADOWSKI, J. Automatic generation of 3D-atomic coordinates for organic molecules. **Tetrahedron Computer Methodology**, v. 3, n. 6, p. 537-547, 1990

GRAHAM, J. Tackling the water crisis: a continuing need to address spatial and social equity. In: SELENDY, J. M. H. (ed.) **Water and sanitation related diseases and the environment: challenges, interventions and preventive measures**. Oxford: Wiley-Blackwell, 2011, p. 3-16.

GUERRANT, R. L.; KOSEK, M.; MOORE, S.; LORNTZ, B.; BRANTLEY, R.; LIMA A. A. Magnitude and impact of diarrheal diseases. **Archives of Medical Research**, v. 33, p. 351-355, 2002.

HUANG, B.; SCHROEDER, M. LIGSITEcsc: predicting ligand binding sites using the Connolly surface and degree of conservation. **BMC Structure Biology**, v. 6, n. 19, p. 1-11, 2006

HUANG, B. MetaPocket: a meta approach to improve protein ligand binding site prediction. **OMICS**, v. 13, p. 325-330, 2009.

HUANG, N; SHOICHET, B. K. Exploiting Ordered Waters in Molecular Docking. **Journal of Medicinal Chemistry**, v. 51, n. 16, p. 4862-4865, 2008.

IRWIN, J. J.; STERLING, T.; MYSINGER, M. M.; BOLSTAD, E. S.; COLEMAN, R. G. ZINC: a free tool to discover chemistry for biology. **Journal of Chemical Information and Modeling**, v. 52, n. 7, p. 1757-1768, 2012.

JACKSON, L. K.; PHILLIPS, M. A. Target validation for drug discovery in parasitic organisms. **Current Topics in Medicinal Chemistry**, v. 2, p. 425-438, 2002.

KARANIS, P.; KOURENTI, C.; SMITH, H. Waterborne transmission of protozoan parasites: a worldwide review of outbreaks and lessons learnt. **Journal of Water and Health**, v. 5, n. 1, p. 1-38, 2007.

KAWABATA, T. Detection of multiscale pockets on protein surfaces using mathematical morphology. **Proteins**, v. 78, p. 1195-1211, 2009.

KOES, D. R.; CAMACHO, C. J. ZINCPharmer: pharmacophore search of the ZINC database. **Nucleic Acids Research**, v. 40, p. W409-W414, 2012.

KORB, O.; STUTZLE, T.; EXNER, T. E. Empirical scoring functions for advanced protein-ligand docking with PLANTS. **Journal of Chemical Information and Modeling**, v. 49, n. 1, p. 84-96, 2009.

KORTAGERE, S. **In silico models for drug discovery**. New York: Humana Press, 2013. 288 p.

KRIVÁK, R.; HOKSZA, D. Improving protein-ligand binding site prediction accuracy by classification of inner pocket points using local features. **Journal of cheminformatics**, V. 7, N. 12, P. 1-13, 2015.

KROEMER, R. T. Structure-based drug design: docagem and scoring. **Current Protein and Peptide Science**, v. 8, p. 312-328, 2007.

LADBURY, J. E. Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. **Chemistry & Biology**, v. 3, n.12,p. 973-980, 1996.

LANGER, T.; HOFFMANN, R. D. **Pharmacophores and pharmacophore searches**. 1ª edição. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA, 2006. p. 3-16.

LASKOWSKI, R. A., SURFNET: A program for visualizing molecular surfaces, cavities, and intermolecular interactions. **Journal of Molecular Graphics**, v. 13, p. 323-330, 1995.

LAURIE, A. T. R.; JACKSON, R. M. Methods for the Prediction of Protein-Ligand Binding Sites for Structure-Based Drug Design and Virtual Ligand Screening. **Current Protein and Peptide Science**, n. 7, p. 395-406, 2006.

LE GUILLOUX, V.; SCHMIDTKE, P.; TUFFERY, P. Fpocket: An open source platform for ligand pocket detection. **BMC bioinformatics**, v. 10, n. 168, p. 1-11, 2009.

LI, H.; ROBERTSON, A. D.; JENSEN, J. H. Very fast empirical prediction and rationalization of protein pka values. **Proteins**, v. 61, p. 704-721, 2005.

LIU, T.; LIN, Y.; WEN, X.; JORISSEN, R. N.; GILSON, M. K. BINDINGDB: A web-accessible database of experimentally determined protein-ligand binding affinities. **Nucleic Acids Research**, v. 35, p. D198-D201, 2007.

MARKT, P.; SCHUSTER, D.; LANGER, T. Pharmacophore models for virtual screening. In: SOTRIFFER, C.; MANNHOLD, R.; KUBINYI, H.; FOLKERS, G. **Virtual screening: principles, challenges, and practical guidelines**, volume 48. Weinheim: Wiley-VCH Verlag & Co. KGaA, 2011. p. 115-152.

MCKINNEY, M. **Lippincott's guide to infectious diseases**. New York: Lippincott Williams & Wilkins, 2010. 384 p.

MEHLHORN, H. **Encyclopedic reference of parasitology, volume 1: biology, structure, function**. Berlin: Springer-Verlag, 2001, p. 543-545.

MENG, X. Y., ZHANG, H. X., MEZEI, M.; CUI, M. Molecular docking: a powerful approach for structure-based drug discovery. **Current Computer-Aided Drug Design**, v. 7, p. 146-157, 2011.

MOOIJ, W. T., VERDONK, M. L. General and targeted statistical potentials for protein-ligand interactions. **Proteins**, v. 61, n. 2, p. 272-287, 2005.

MORELAND, J. L.; GRAMADA, A.; BUZKO, O. V.; ZHANG, Q.; BOURNE, P. E. The Molecular Biology Toolkit (MBT): a modular platform for developing molecular visualization applications. **BMC Bioinformatics**, 2005. DOI: 10.1186/1471-2105-6-21.

MÜLLER, J.; HEMPHILL, A.; MÜLLER, N. Treatment of giardiasis and drug resistance. In: LUJÁN, H. D.; SVÄRD, S. (Eds) **Giardia - a model organism**. New York: Springer Verlag, 2011.

MUNAGALA, N.; SARVER, A. E.; WANG, C. C. Converting the guanine phosphoribosyl transferase from *Giardia lamblia* to a hypoxanthine-guanine phosphoribosyltransferase. **Journal of Biological Chemistry**, v. 275, n. 47, p. 37072-37077, 2000.

MUNAGALA, N.; WANG, C. C. The pivotal role of guanine phosphoribosyltransferase in purine salvage by *Giardia lamblia*. **Molecular Microbiology**, v. 44, n. 4, p. 1073-1079, 2002.

NASH, T. E.; OHL, C. A.; THOMAS, E.; SUBRAMANIAN, G.; KEISER, P.; MOORE, T. A. Treatment of patients with refractory giardiasis. **Clinical Infectious Diseases**, v. 33, n. 1, p. 22-28, 2001.

NYHAN, W. L.; O'NEILL, J. P.; JINNAH, H. A.; HARRIS, J. C. **Lesch-Nyhan Syndrome**. 2000. In: PAGON, R. A.; ADAM, M. P.; ARDINGER, H. H., *et al.* **GeneReviews**[®]. University of Washington, Seattle; 1993-2015. Disponível em: <http://www.ncbi.nlm.nih.gov/books/NBK1149/>

ORGANIZAÇÃO MUNDIAL DA SAÚDE (OMS) - Working group on water and health. **Technical guidance document on water-related disease surveillance**. Genebra, Suíça: OMS, 2010.

PALANIAPPAN, M.; GLEICK, P. H.; ALLEN, L.; COHEN, M. J.; CHRISTIAN-SMITH, J. SMITH, C. Water Quality. In: GLEICK, P. H. (Ed.). **The world's water volume 7: the biennial report on freshwater resources**. New York: Springer, 2012. p. 45-72.

PAUL, M.; ROGNAN, D. ConsDock: a new program for the consensus analysis of protein-ligand interactions. **PROTEINS: structure, function and genetics**, v. 47, p. 521-533, 2002.

PEACH, M. L.; NICKLAUS, M. C. Combining docking with pharmacophore filtering for improved virtual screening. **Journal of Cheminformatics**, v. 1, n. 6, 2009. Disponível em: <http://www.jcheminf.com/content/1/1/6>.

PÉROT, S.; SPERANDIO, O.; MITEVA, M. A.; CAMPROUX, A-C.; VILLOUTREIX, B. O. Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. **Drug Discovery Today**, v. 15, n.15/16, p. 656-667, 2010.

QUIHUI-COTA, L.; ASTIAZARÁN-GARCÍA, H.; VALENCIA, M. E.; MORALES-FIGUEROA, G. G.; LOPEZ-MATA, M. A.; VAZQUEZ ORTIZ, F. Impact of *Giardia intestinalis* on vitamin a status in schoolchildren from northwest Mexico. **International Journal for Vitamin and Nutrition Research**, v. 78, n. 2, p. 51-56, 2008.

RANA, S. V.; BHASIN, D. K.; VINAYAK, V. K. Lactose hydrogen breath test in *Giardia lamblia*-positive patients. **Digestive Diseases and Sciences**, v. 50, n. 2, p. 259-261, 2005.

RODRIGUES, R. P.; MANTOANI, S. P.; ALMEIDA, J. R.; PINSETTA, F. R.; SEMIGHINI, E. P.; SILVA, V. B.; SILVA, C. H. T. P. Estratégias de Triagem Virtual no Planejamento de Fármacos. **Revista Virtual de Química**, v. 4, p. 739-776, 2012.

SADOWSKI, J.; GASTEIGER, J.; KLEBE, G. Comparison of automatic three-dimensional model builders using 639 X-Ray structures. **Journal of Chemical Information and Modeling**, v. 34, n. 4, 1000-1008, 1994.

SANDERSON, D. M.; EARNSHAW, C. G. Computer prediction of possible toxic action from chemical structure: The DEREK system. **Human & Experimental Toxicology**, v. 10, p. 261-273, 1991.

SANTOS, R.; HRITZ, J.; OOSTENBRINK, C. Role of water in molecular docking simulations of cytochrome P450 2D6. **Journal of Chemical Information and Modeling**, v. 50, n. 1, p. 146-154.

SASTRY, G.; ADZHIGIREY, M.; DAY, T.; ANNABHIMOJU, R.; SHERMAN, W. Protein and ligand preparation: parameters, protocols and influence on virtual screening enrichments. **Journal of Computer Aided Molecular Design**, v. 27, p. 221-234, 2013.

SAVIOLI, L.; SMITH, H.; THOMPSON, A. *Giardia* and *Cryptosporidium* join the 'Neglected Diseases Initiative'. **Trends in Parasitology**, v. 22, n. 5, p. 203-208, 2006.

SHI, W.; LI, C. M.; TYLER, P. C.; FURNEAUX, R. H.; GRUBMEYER, C.; SCHRAMM, V. L.; ALMO, S. C. The 2.0 Å structure of human hypoxanthine-guanine phosphoribosyltransferase in complex with a transition-state analog inhibitor. **Nature Structural Biology**, v. 6, n. 6, p. 588-593, 1999.

SCHRÖDINGER. **Glide user manual**. GLIDE, versão 5.5. Nova York: Schrödinger LLC, 2009.

SHI, W.; MUNAGALA, N. R.; WANG, C. C.; LI, C. M.; TYLER, P. C.; FURNEAUX, R. H.; GRUBMEYER, C.; SCHRAMM, V. L.; ALMO, S. C. Crystal structures of *Giardia lamblia* guanine phosphoribosyltransferase at 1.75 Å. **Biochemistry**, v. 39, n. 23, p. 6781-6790, 2000.

SPYRAKIS, F.; KELLOGG, G. E.; AMADASI, A.; COZZINI, P. Scoring functions for virtual screening. In: RAHMAN, A.-U.; CALDWELL, G. W.; CHOUDHARY, M. I.; PLAYER, M. P. (eds.) **Frontiers in Drug Design and Discovery**, volume 3. Oak Park: Bentham Science Publishers, 2007, p. 317-379.

STEWART, J. J. P. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. **Journal of Molecular Modeling**, v. 13 n. 12, p. 1173-1213, 2007.

TAN, K. P.; NGUYEN, T. B.; PATEL, S.; VARADARAJAN, R.; MADHUSUDHAN, M. S. Depth: a web server to compute depth, cavity sizes, detect potential small-molecule ligand-binding cavities and predict the pKa of ionizable residues in proteins. **Nucleic Acids Research**, v. 41, W314-W321, 2013.

TORRES, R. J.; PUIG, J. G. Hypoxanthine-guanine phosphoribosyltransferase (HPRT) deficiency: Lesch-Nyhan syndrome. **Orphanet Journal of Rare Diseases**, v. 2, n. 48, p. 1-10, 2007.

TESTA, B.; BALMAT, A.; LONG, A. Predicting drug metabolism: concepts and challenges. **Pure and Applied Chemistry**, v. 76, n. 5, p. 907-914, 2004.

TULCHINSKY, T. H.; VARAVIKOVA, E. A. **The new public health**. 2^a edição. San diego, California: Elsevier Academic Press, 2009. 672 p.

USA ENVIRONMENTAL PROTECTION AGENCY (EPA). **Giardia: drinking water health advisory**. EPA report No. 822-R-99-008. Washington: EPA, 1999. P. 1-40.

VAN DE WATERBEEMD, H.; GIFFORD, E. ADMET in silico modelling: towards prediction paradise? **Nature Reviews Drug Discovery**, v. 2, n. 3, p. 192-204, 2003.

WARREN, G. L.; ANDREWS, C. W.; CAPELLI, A. M.; CLARKE, B.; LALONDE, J.; LAMBERT, M. H.; LINDVALL, M.; NEVINS, N.; SEMUS, S. F.; SENGER, S.; TEDESCO, G.; WALL, I. D.; WOOLVEN, J. M.; PEISHOFF, C. E.; HEAD, M. S. A critical assessment of docagem programs and scoring functions. **Journal of Medicinal Chemistry**, v. 49, n. 20, p. 5912-5931, 2006.

WAVEFUNCTION Spartan'14. Versão v.1.1.4. San Diego, CA.: Wavefunction Inc., 2014.

WEI, B. Q.; BAASE, W. A.; WEAVER, L. H.; MATTHEWS, B. W.; SHOICHET, B. K. A model binding site for testing scoring functions in molecular docking. **Journal of Molecular Biology**, v. 322, n. 2, p. 339-355, 2002.

YU, J.; ZHOU, Y.; TANAKA, I.; YAO, M. Roll: a new algorithm for the detection of protein pocket and cavities with a rolling probe sphere. **Bioinformatics**, v. 26, n. 1, p. 46-52, 2009.

ZACHARIAS, N.; DOUGHERTY, D. A. Cation- π interactions in ligand recognition and catalysis. **TRENDS in Pharmaceutical Sciences**, v. 23, n. 6, p. 281-287, 2002.

ZANATTA, G.; NUNES, G.; BEZERRA, E. M.; DA COSTA, R. F.; MARTINS, A.; CAETANO, E. W.; FREIRE, V. N.; GOTTFRIED, C. Antipsychotic haloperidol binding to the human dopamine D3 receptor: beyond docking through QM/MM refinement toward the design of improved schizophrenia medicines. **ACS Chemical Neuroscience**, v. 5, n. 10, p. 1041-1054, 2014.

ZHANG, Z.; YU LI, Y.; LIN, B.; SCHROEDER, M.; HUANG, B. Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction. **Bioinformatics**, v. 27, n. 15, p. 2083-2088, 2011.