

## Modelagem molecular de isoformas e docagem molecular da enzima hipoxantina-guanina fosforibosiltransferase (HGPRT) de *Schistosoma mansoni*

**Cassiano Vicente de Lima**

Universidade Estadual do Paraná, campus de União da Vitória.  
Contato: cassianovdl@gmail.com

**Deise Borchhardt Moda**

Universidade Estadual do Paraná, campus de União da Vitória.  
Contato: deise-borchhardt@hotmail.com

**Larissa Romanello**

Universidade do Estado de Minas Gerais, Unidade Ituiutaba.  
Contato: larissa.romanello@uemg.br

**Resumo:** A esquistossomose é uma doença parasitária que afeta cerca de 240 milhões de pessoas em todo o mundo. Diferente de seu hospedeiro humano, o *Schistosoma mansoni* - parasito causador dessa doença, não possui a via metabólica de síntese de purinas e depende integralmente da via de salvação para suprimento dessas bases. Assim, esta via tem sido citada na literatura como alvo potencial para o desenvolvimento de novos fármacos, uma vez que para o tratamento da doença, apenas um fármaco é utilizado, o praziquantel, e diversos casos de resistência do parasito a esse medicamento já foram reportadas. A HGPRT é uma das enzimas dessa via e catalisa a fosforibosilação reversível de hipoxantina e guanina para IMP (inosina monofosfato) ou GMP (guanosina monofosfato). Uma vez que no genoma do parasito foram identificadas três sequências que codificam para HGPRT e apenas uma das isoformas foi resolvida experimentalmente, a modelagem molecular por homologia foi realizada como forma preditiva para o estudo das estruturas ainda não resolvidas. A construção dos modelos foi realizada pelo servidor Swiss-Model; as visualizações, análises e validações pelos softwares UFSC-Chimera, PyMol e Procheck. Os modelos obtidos demonstraram qualidade satisfatória para as análises e enovelamento conservado entre as isoformas comparadas. Como ferramenta preditiva e de busca por possíveis inibidores para a enzima, foi realizada a docagem molecular na estrutura cristalográfica, elucidando comportamento e interação de pequenas moléculas no sítio ativo da enzima. Uma vez que esta foi resolvida complexada com IMP, a busca por ligantes se limitou a estruturas análogas ao mesmo e que possuíam potencial farmacológico. Para isso, foram utilizados quatro compostos da classe dos aciclonucleosídeos (aciclovir, adefovir, penciclovir e tenofovir) obtidos através do banco de dados PubChem. Os resultados da docagem foram avaliados quanto a sua pontuação e sua similaridade com o IMP, sugerindo que o penciclovir tem a maior afinidade com o sítio ativo e foi o segundo composto melhor pontuado pelo programa entre os testados. Nesse contexto, os resultados obtidos contribuem para o planejamento de ensaios biológicos de atividade enzimática com os compostos docados e a resolução experimental das estruturas modeladas.

**Palavras-chave:** Biologia estrutural; Esquistossomose; Bioinformática; Doença tropical negligenciada.

### Molecular modelling of isoforms and molecular docking of hypoxanthine-guanine phosphoribosyltransferase from *Schistosoma mansoni*

**Abstract:** *Schistosoma mansoni*, the parasite responsible for schistosomiasis, a disease that affects around 240 million people worldwide, does not have a purine synthesis pathway, depending entirely on the salvage pathway to supply its bases. This pathway has been cited as a potential target for development of new drugs against the parasite, because for treatment of the disease just one drug is used, praziquantel, and several cases of parasite resistance to this drug have been reported. HGPRT is one of the key enzymes of the parasite purine salvage pathway, catalyzing reversible phosphoribosylation of

hypoxanthine and guanine to IMP (inosine monophosphate) or guanosine monophosphate. Since three sequences coding for HGPRT were identified in the parasite genome and just one of the isoforms was resolved, homology molecular modeling was performed as a predictive method for the study of unsolved structures. The model construction was performed by the Swiss-Model server, visualizations, analyzes and validations by the software UFSC-Chimera, PyMol and Procheck. As a predictive and search tool for possible inhibitors for the enzyme, molecular docking was performed in the crystallographic structure, elucidating behavior and interaction of small molecules in the active site of the protein, which, because the crystallographic structure was complexed with IMP, the search for ligands was limited to structures similar to it and also have pharmacological potential. For this, four compounds of the acyclonucleoside class (acyclovir, adefovir, penciclovir and tenofovir) obtained through the PubChem database were used. The docking results were evaluated by the final score and drug-likeness, suggesting that penciclovir has the highest affinity to the active site and second best drug-likeness of program score among the tested compounds. The evaluations of the generated models demonstrated a satisfactory quality, the future prospects are in realizing biological assays of the activities of the docked compounds and experimental resolution of the modeled structures.

**Key-words:** Structural biology; Schistosomiasis; Bioinformatics; Neglected tropical disease.

Como citar este artigo:

LIMA, C.V.; MODA, D.B.; ROMANELLO, L. Modelagem molecular de isoformas e docagem molecular da enzima hipoxantina-guanina fosforibosiltransferase (HGPRT) de *Schistosoma mansoni*. **Luminária**, União da Vitória, v.22, n.01, p. 35 – 49, 2020.

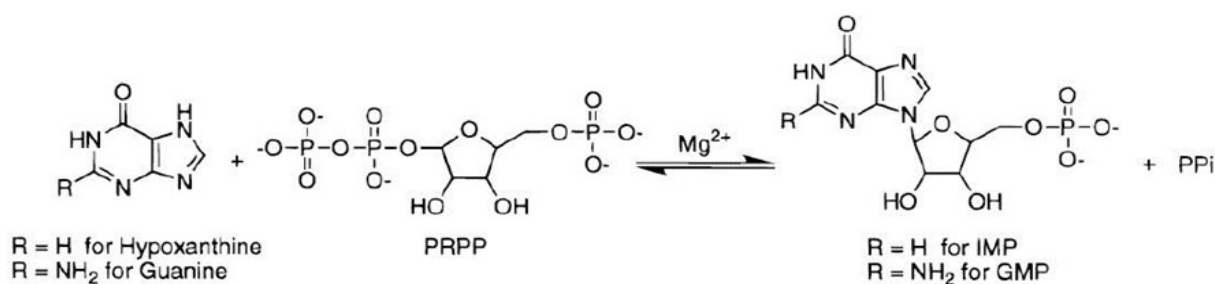
## INTRODUÇÃO

*Schistosoma mansoni* é o parasito responsável pela esquistossomose, doença que afeta cerca de 240 milhões de pessoas em todo o mundo. A infecção prevalece em áreas tropicais e subtropicais, em comunidades sem água potável e que não possuem saneamento básico adequado (WHO, 2017). Para controle e tratamento da doença o principal fármaco utilizado é o Praziquantel (PZQ), pois apresenta baixo custo e efeitos colaterais não muito significativos. Entretanto, a taxa de cura da doença por meio da utilização do PZQ não é superior a 60% e dessa maneira uma considerável parcela dos indivíduos tratados podem permanecer infectados, correndo risco de desenvolver a doença na sua forma crônica. Além disso, o parasito tem apresentado resistência ao PZQ e já são descritos possíveis aumentos no risco de doenças hepáticas nos indivíduos que passaram pelo tratamento (BARSOUM et al., 2013).

Dentro desse contexto, surge a necessidade da busca de novos alvos para o desenvolvimento de fármacos mais eficazes e seguros contra o parasito, que ao contrário de seu hospedeiro humano, não possui a via de síntese de purinas e depende integralmente da via de salvação de purinas, utilizando precursores provenientes do hospedeiro para suprir a ne-

cessidade dessas bases (CRAIG et al., 1988 *apud* ROMANELLO, 2011). Pela dependência do parasito dessa via, ela tem sido citada como um alvo em potencial para o desenvolvimento de novos fármacos contra a doença. Uma das enzimas chave dessa via é a Hipoxantina-Guanina Fosforibosiltransferase (HGPRT), fundamental para o parasito por suprir as atividades de hipoxantina e guanina, atuando na formação desses nucleotídeos essenciais (YUAN et al., 1990). A enzima funciona catalisando a fosforibosilação reversível de hipoxantina e guanina para respectivamente: inosina monofosfato (IMP) ou guanosina monofosfato (GMP) e pirofosfato, tendo como doador de fosfato e ribose o 5-fosforribosil 1-difosfato (PRPP) (KEOUGH et al., 2005 *apud* ROMANELLO, 2011) (Figura 1).

Canyuk et al. (2003), destacaram que enzimas que salvam 6-oxopurinas, incluindo as HGPRTs, são potenciais alvos para o desenvolvimento de fármacos no tratamento de doenças causadas por parasitos. Uma vez que muitos dos organismos parasitos perderam a capacidade de sintetizar bases de purina, tendo que as retirar das células infectadas. Isso os torna particularmente suscetíveis a fármacos que atacam enzimas que realizam essa função.



**Figura 1.** A - Esquema da reação catalisada pela enzima HGPRT. Fonte: FOCIA et al. (1998)

Estudos realizados por Pereira et al. (2008), demonstraram através da técnica de RNAi (RNA de interferência) a importância da enzima HGPRT para a sobrevivência de *Schistosoma mansoni*. No referido trabalho os RNAi foram produzidos contra a expressão de HGPRT e injetados 70 dias após a infecção com cercárias em camundongos. O estudo demonstrou uma redução de carga parasitária de aproximadamente 27% após o tratamento.

Na busca de possíveis inibidores que venham a interagir com o sítio ativo da HGPRT, inibindo ou reduzindo sua atividade, foram realizados estudos de docagem molecular na HGPRT já resolvida experimentalmente utilizando quatro diferentes compostos da classe dos acilonucleosídeos, que segundo Keough et al. (2009), são os primeiros compostos que demonstraram inibir seletivamente a HGPRT de *Plasmodium falciparum*, parasito responsável pela malária, como *S. mansoni* não possui a via de novo de síntese de nucleotídeos em comparação com a HGPRT humana. Esses estudos, realizados em HGPRT humana e de *P. falciparum*, utilizaram acilonucleosídeos, compostos análogos aos nucleotídeos produtos da reação da HGPRT, constituídos por uma base nitrogenada unida por uma unidade fosfatada. O termo aciclo refere-se à ausência de ligações C-C na porção furanosídica, tornando-a aberta. Os acilonucleosídeos são amplamente utilizados como agentes antivirais, por se demonstrarem potentes e seletivos, tanto em meios *in vitro* quanto *in vivo* (CLERCQ; HOLÝ, 2005). Três dos compostos testados impediram o crescimento de *P. falciparum* em cultura de eritrócitos, validando assim esses compostos como candidatos satisfatórios à inibição da *PfHGPRT*. A similaridade

de sequencial entre *PfHGPRT* e *SmHGPRT* é de 39.30% e um *Root Mean Square Deviation* (RMSD) estrutural médio de 1.249, indicando assim que acilonucleosídeos podem ter ação similar na enzima de *S. mansoni*.

A técnica de docagem se apresenta como uma ferramenta chave na predição da melhor orientação de ajuste de ligantes ou possíveis inibidores na proteína, permitindo caracterizar o comportamento das moléculas no sítio de ligação da proteína alvo, assim como, elucidar as interações moleculares (MENG et al., 2011). Devido a estrutura cristalográfica da enzima alvo estar depositada no *Protein Data Bank* (PDB) em complexo com o IMP como ligante, a conformação estrutural do sítio ativo fechado para o composto foi considerada, limitando a busca por ligantes que tenham similaridade estrutural com o IMP e tenham potencial para utilização como possíveis fármacos.

Em *S. mansoni* a HGPRT é codificada por um gene essencial (KEOUGH et al., 2005), sendo encontrados três genes homólogos para a síntese da HGPRT em seu genoma (Smp\_148820, Smp\_168500 e Smp\_103560), que estão relacionados com diferentes estágios de vida do parasito. Porém, somente uma das estruturas foi resolvida experimentalmente por Romanello em 2016 (ROMANELLO et al., 2019), a isoforma Smp\_103560 (PDB ID: 5IPF). O conhecimento estrutural das outras duas isoformas ainda não resolvidas é de fundamental importância para melhor compreender a via de salvação de purinas do parasito e consequentemente analisar a viabilidade do desenvolvimento de novas estratégias que utilizem essa via no combate ao parasito. A modelagem molecular se apresenta como uma forma preditiva de baixo custo para o

estudo dessas enzimas ainda não resolvidas, sendo assim, através de métodos computacionais foram preditas e avaliadas as estruturas tridimensionais das isoformas Smp\_148820 e Smp\_168500.

## METODOLOGIA

### Modelagem Molecular

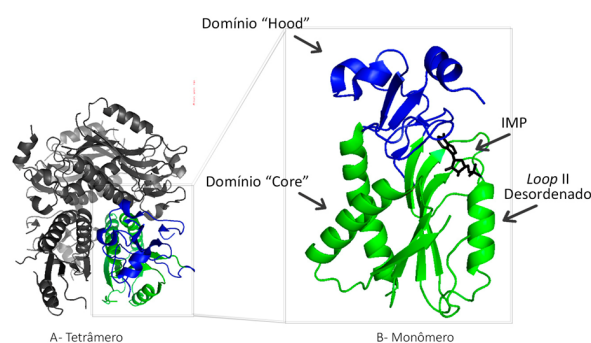
A ferramenta mais utilizada na predição de estruturas tridimensionais de proteínas é a modelagem por homologia, também conhecida como modelagem comparativa. Esta abordagem baseia-se em alguns padrões gerais que têm sido observados em nível molecular e no processo de evolução biológica (BRANDEN; TOOZE, 1991). Um modelo construído por modelagem comparativa necessita que pelo menos uma estrutura tridimensional da família em questão tenha sido elucidada por técnicas experimentais e que a sequência de aminoácidos da proteína a ser modelada tenha sido elucidada. Entretanto, é necessário que entre ambas as sequências (alvo e molde) exista certo grau de identidade. Geralmente considera-se como significativo um valor acima de 25% para a identidade sequencial entre a proteína-molde e a proteína-alvo (D'ALFONSO et al., 2001).

A proteína molde escolhida para a modelagem das duas proteínas alvo desse trabalho foi a isoforma resolvida por cristalografia de raios-X com 2.8 Å de resolução e depositada no PDB na forma de tetrâmero com a identificação 5IPF. Nessa estrutura a subunidade é composta por  $\alpha$ -hélices e folhas  $\beta$ . Podendo ser dividida em dois domínios (Figura 2). O domínio altamente conservado entre as HGPRTs chamado de “core” é composto por 4 fitas paralelas ( $\beta$ 2,  $\beta$ 3,  $\beta$ 4 e  $\beta$ 5) formando uma folha  $\beta$  flanqueada por 3  $\alpha$ -hélices ( $\alpha$ 3,  $\alpha$ 4 e  $\alpha$ 5), que liga o substrato primário PRPP. Compreende os resíduos 41 – 173 formando uma estrutura chamada de “*doublywound a/ $\beta$  sheet*” (RICHARDSON, 1981 *apud* ROMANELLO, 2016) com hélices de ambos os lados da folha  $\beta$ . Esse domínio também compreende o loop II, o qual apresenta uma região desordenada sem densidade eletrônica para os resíduos de 108 – 122. O segundo domínio, localizado acima do domínio “core” é chama-

do domínio “hood”. Este é composto por 4 pequenas fitas  $\beta$  ( $\beta$ 1,  $\beta$ 6,  $\beta$ 7 e  $\beta$ 8), divididas em 2 folhas, sendo uma paralela e outra anti-paralela, 3 pequenas  $\alpha$ -hélices ( $\alpha$ 1,  $\alpha$ 2 e  $\alpha$ 6), várias conexões de longos loops e compreende os resíduos 8 – 40 e 174 – 224. O sítio ativo de *SmpHGPRT*, assim como de suas ortólogas, está localizado na interface dos domínios “core” e “hood” e compreende 4 loops, I – IV (ROMANELLO, 2016; ROMANELLO et al., 2019).

As duas proteínas alvo neste trabalho possuem uma média de 42,3% de identidade com a proteína molde e em torno de 32% de identidade sequencial com a proteína HGPRT humana (PDB ID: 1Z7G) (Tabela 1).

Foram realizadas tentativas de modelagem utilizando-se o tetrâmero como molde. O modelo obtido não apresentou qualidade satisfatória, possivelmente devido às limitações do sistema, sendo assim, para a modelagem



**Figura 2.** A - Tetrâmero: Estrutura tetramérica de *SmpHGPRT*, conforme depositado no PDB; B - Monômero: Estrutura monomérica de *SmpHGPRT*, composta por 2 domínios: domínio “hood” (azul), domínio “core” (verde).

final foram retiradas as cadeias B, C e D da proteína molde, assim como os ligantes (IMP) e as águas, obtendo-se apenas a estrutura da cadeia A. Utilizou-se a ferramenta PMV 1.5.6 presente no pacote MGLTools versão 1.5.6 para esse procedimento.

O alinhamento de múltiplas sequências organizou as sequências de proteínas em uma matriz retangular, para que os resíduos em uma determinada coluna sejam homólogos (derivados de uma única posição em uma sequência ancestral), superponíveis (em um local rígido de alinhamento estrutural) ou desempe-

nam um papel funcional comum. Embora estes três critérios sejam essencialmente equivalentes, para proteínas, sequência, estrutura e função divergem em tempo evolutivo e diferentes critérios podem resultar em alinhamentos diferentes. Os alinhamentos refinados manualmente tendem ainda a serem superiores aos métodos puramente automatizados; existe, portanto, um esforço contínuo para melhorar a precisão biológica das ferramentas de alinhamento (KEITH, 2017). O alinhamento das sequências foi realizado utilizando-se a ferramenta *web* T-Coffee (NOTREDAME; HIGGINS; HERINGA, 2000). Das sequências utilizadas no alinhamento, Smp\_168500 possui 214 aminoácidos, Smp\_148820 possui 216 e Smp\_103560, 201 aminoácidos na cadeia A (Figura 3).

**Tabela 1.** Identidade sequencial entre as proteínas modeladas (Smp\_148820 e Smp\_168500), a proteína molde (5IPF) e a HGPRT humana (1Z7G).

ID. SEQ.	Smp_168500	43.93% ↔	Smp_148820
5IPF	46.77%		37.81%
↑ 43.72%			
1Z7G	36.18%		28.14%

Para a modelagem foi utilizada a ferramenta SWISS-MODEL, um espaço de trabalho especializado em modelagem de proteínas utilizando servidores *web*. A partir da inserção da sequência alvo, uma biblioteca de estruturas de proteínas (PDB) é pesquisada automaticamente para identificar modelos adequados para servirem de molde, ou o usuário pode selecionar um molde e alinhamento manualmente, como foi realizado neste trabalho. Com base no alinhamento de uma sequência entre a proteína alvo e a estrutura modelo, é gerado um modelo tridimensional para a proteína alvo. Ferramentas de avaliação da qualidade do modelo são usadas integradamente à ferramenta para estimar a confiabilidade dos modelos resultantes (WATERHOUSE et al., 2018).

Em um primeiro momento, os modelos foram inspecionados visualmente, sobrepondo e comparando as estruturas, na etapa seguinte, os modelos foram avaliados pelo diagrama de Ramachandran (RAMACHANDRAN, 1963), gerado pelo Procheck (LASKOWSKI et al., 1996). A obtenção de dados de estrutura, como os *plots* e informações de estrutura secundária a partir da estrutura terciária modelada foi realizada



**Figura 3.** Alinhamento das sequências com a ferramenta T-Coffee. O algoritmo difere as regiões em vermelho considerando-as “boas”, em amarelo as “medianas” e em verde e azul as “ruins”. A linha *cons* indica o índice de conservação dos aminoácidos, sendo um “\*” representando conservação total (mesmos aminoácidos); um “:” indicando similaridade média (aminoácidos com características semelhantes) e uma baixa similaridade para os campos sem nenhum caractere ou com um “.”.

pelo PDBsum (LASKOWSKI et al., 2005). A visualização e a manipulação das estruturas tridimensionais das proteínas foram realizadas nos programas PyMOL 2.0 (SCHRÖDINGER, 2018) e UCSF-Chimera versão 1.12 (PETTERSEN et al., 2004).

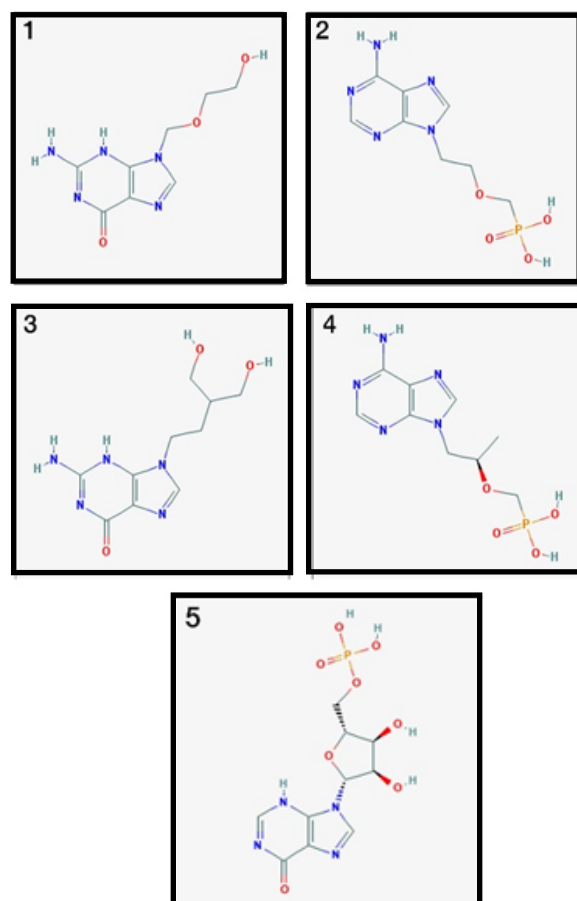
### Docagem molecular

O processo de docagem envolve quatro etapas básicas: (1) Preparo da estrutura alvo; (2) Escolha dos ligantes; (3) Predição da conformação, posição, orientação do ligante dentro dos sítios e (4) Avaliação da afinidade de ligação. A docagem também pode ser utilizada para realizar triagem virtual de grandes bibliotecas de compostos, classificar resultados e propor hipóteses estruturais de como prováveis ligantes podem interagir com os alvos. Esses cálculos podem auxiliar a aumentar a probabilidade de encontrar novos inibidores de proteínas-chave de patógenos, ao prever quão bem um composto pode ligar-se ao alvo, onde ele pode ligar-se, e que tipos de interações poderia formar no local de ligação. Também identificar candidatos promissores para testes de laboratório (MENG et al., 2011). A obtenção dos ligantes foi realizada utilizando os bancos de dados de compostos e substâncias virtuais, ZINC15 (STERLING; IRWIN, 2015) e PubChem (KIM et al., 2015), onde foram escolhidas duas classes de aciclonucleosídeos, os de bases púricas fosfatados: adefovir (ADV) e tenofovir (TFV). Assim como os aciclonucleotídeos de bases púricas sem o grupo fosfato: aciclovir (ACV) e penciclovir (PCV) (Figura 4).

Os compostos foram obtidos na forma de arquivo SDF (*structure-data file*) e convertidos para o formato PDB (formato aceito pelo *software* AutoDockVina). Após a conversão, todas as moléculas foram inspecionadas visualmente a fim de detectar possíveis erros de conversão. As ligações flexíveis dos ligantes foram selecionadas e a estrutura foi convertida para o formato PDBQT (Protein Data Bank, Partial Charge (Q), & AtomType (T)).

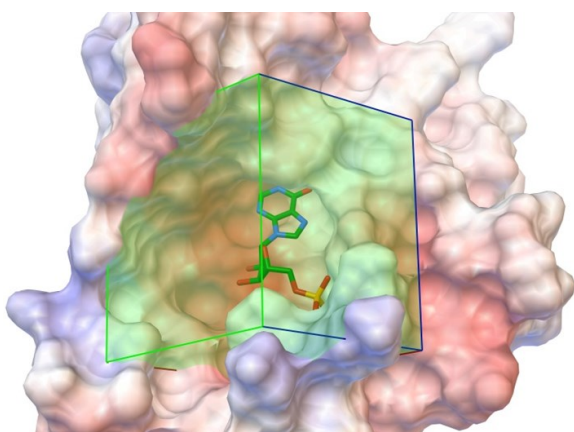
A proteína alvo teve suas cadeias B, C e D removidas, assim como os ligantes IMP e as moléculas de água (limitação do AutoDockVina), Cargas de Gasteier foram adicionadas e os hidrogênios polares foram removi-

dos. A área de docagem molecular (*grid box*) foi definida usando as coordenadas cartesianas: X = -7.278 Å; Y=13.277 Å; e Z=-10.160 Å com as dimensões de 16 × 16 × 16 Å (Figura 5). Tendo no centro o sítio e local da ligação do IMP na proteína cristalográfica (5IPF), os *scripts* do AutoDockVina foram executados utilizando o pacote de softwares AutoDockTools 1.5.6. Cada composto foi testado individualmente, e para cada teste foram computadas 10 conformações diferentes, porém, apenas a conformação com melhor pontuação final para cada ligante foi selecionada. Também foram realizados testes de docagem utilizando o próprio ligante IMP. O objetivo foi comparar a conformação obtida com resultado dos testes e a conformação cristalográfica desse mesmo ligante no sítio ativo, sendo assim a validação da técnica.



**Figura 4.** Aciclonucleosídeos de base púrica selecionados para utilização nos testes de docagem na HGPR1 de *Schistosoma mansoni*: 1: Aciclovir; 2: Adefovir; 3: Penciclovir; 4: Tenofovir; 5: IMP.

Os ligantes foram então avaliados quanto a sua classificação por pontuação gerada pelo próprio AutoDockVina, que leva em conta, além de uma série de funções, as superfícies hidrofóbicas, repulsões e ligações de hidrogênio (TROT; OLSON, 2010). Os prováveis ligantes testados também foram avaliados quanto a pontuação (*drug-score*) pelo *software* Molinspiration, uma função complexa baseada em várias propriedades moleculares e estruturais que determinam se uma molécula apresenta características similares à fármacos conhecidos, como fatores de hidrofobicidade, distribuição eletrônica, características das ligações de hidrogênio, tamanho e flexibilidade da molécula e presença de várias características farmacológicas que podem influenciar o comportamento da molécula no organismo vivo, incluindo biodisponibilidade, propriedades de transporte, afinidade às proteínas, reatividade, toxicidade, estabilidade metabólica, entre outros (MOLINSPIRATION, 2018).

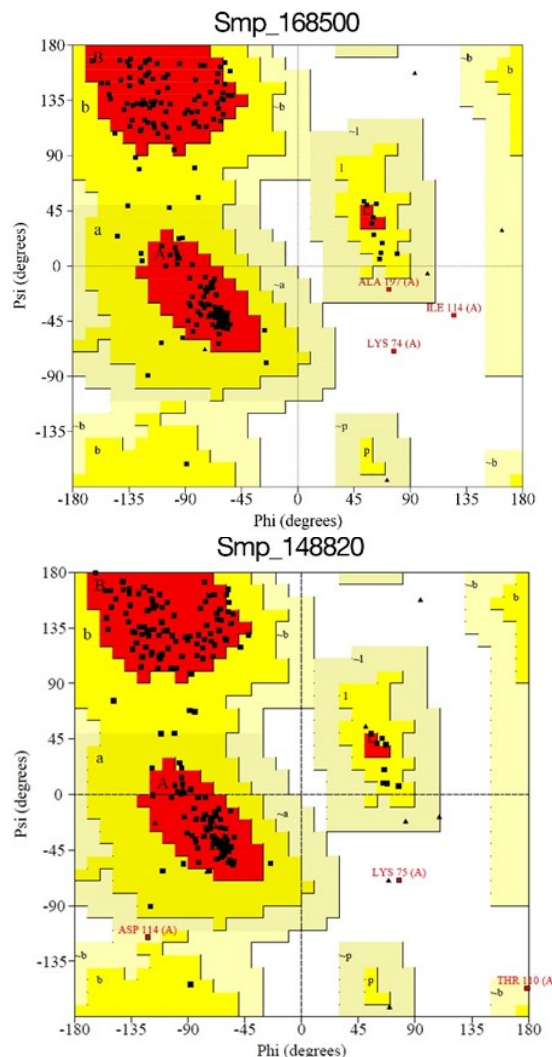


**Figura 5.** Em destaque, área de grid box para o *docking* na proteína 5IPF representada como superfície, no centro o ligante IMP. Coordenadas cartesianas: X = -7.278 Å; Y=13.277 Å; e Z=-10.160 Å; Dimensões: 16 × 16 × 16 Å.

## RESULTADOS E DISCUSSÃO

### Modelagem molecular

A primeira forma de validação utilizada foi a avaliação dos ângulos diédricos  $\Phi$  (phi) e  $\Psi$  (psi) dos aminoácidos das estruturas no diagrama de Ramachandran gerado pelo programa Procheck (LASKOWSKI et al., 1996). Os resultados dos gráficos de Ramachandran podem ser visualizados na Figura 6 para as estruturas Smp\_168500 e Smp\_148820.



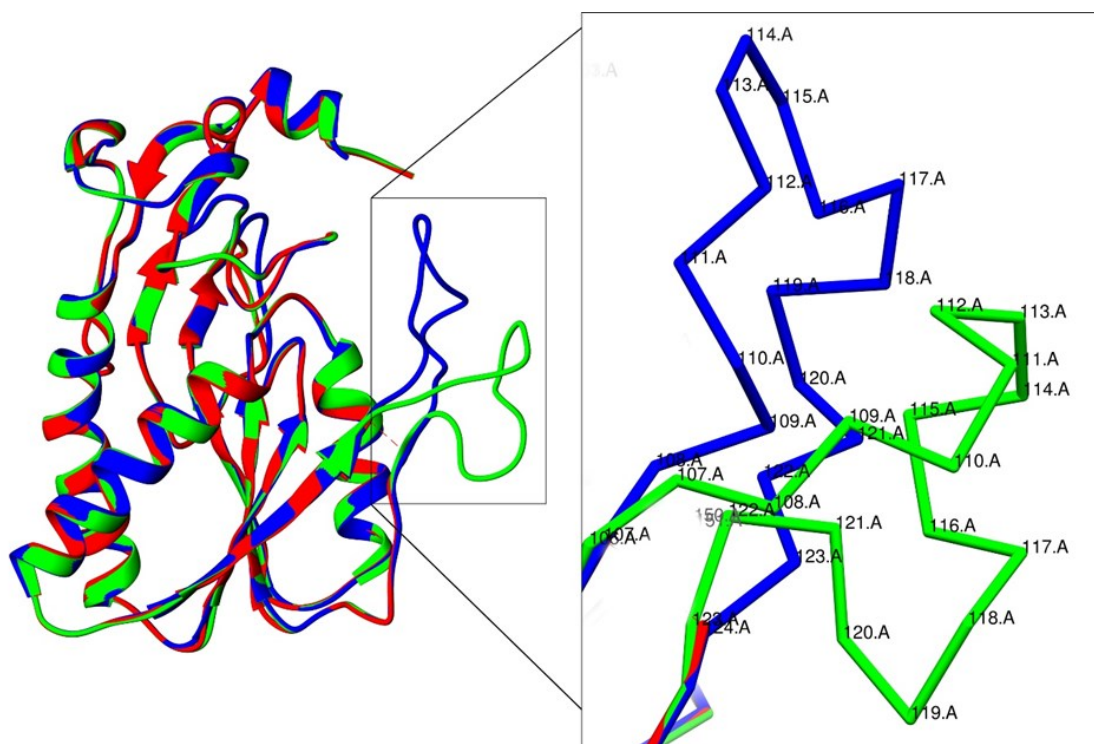
**Figura 6.** Diagramas de Ramachandran com as distribuições dos ângulos diédros para as estruturas *Sm* HGPRTSmp\_168500 e Smp\_148820. As regiões representadas em vermelho correspondem às regiões mais favoráveis para os valores dos ângulos, as regiões em amarelo escuro são as adicionalmente permitidas, amarelo claro são as generosamente permitidas e as regiões em branco são as regiões não permitidas. As glicinas são representadas por triângulos.

Na estrutura de Smp\_168500, os resíduos em regiões mais favoráveis (vermelho) correspondem a 85.6%, adicionalmente permitida (amarelo escuro) a 12.9%, generosamente permitida (amarelo claro) a 0,5% e não permitida a 1,0%. Já em Smp\_148820, os resíduos em regiões mais favoráveis (vermelho) correspondem a 86.9%, adicionalmente permitida (amarelo escuro) a 11.5%, generosa-

mente permitida (amarelo claro) a 1,0% e não permitida a 0,5%.

As estruturas modeladas apresentaram um RMSD médio para a estrutura molde (5IPF) de 0.125 para Smp\_168500 e de 0.160 para Smp\_148820, a diferença estrutural mais notável entre os modelos foi na região que compreende o loop II, que encontra-se desor-

denado e sem densidade eletrônica entre os resíduos de 108 – 122 na estrutura molde, assim como em outras HGPRs depositadas no PDB, essa desordem é comumente observada uma vez que esse loop é móvel (HÉROUX *et al.*, 1999), desse modo, os modelos obtidos tiveram variação nessa região (Figura 7).



**Figura 7.** Região do loop II, em vermelho a estrutura de 5IPF, sem estrutura definida. Em verde, loop modelado para Smp\_168500 e em azul para Smp\_148820.

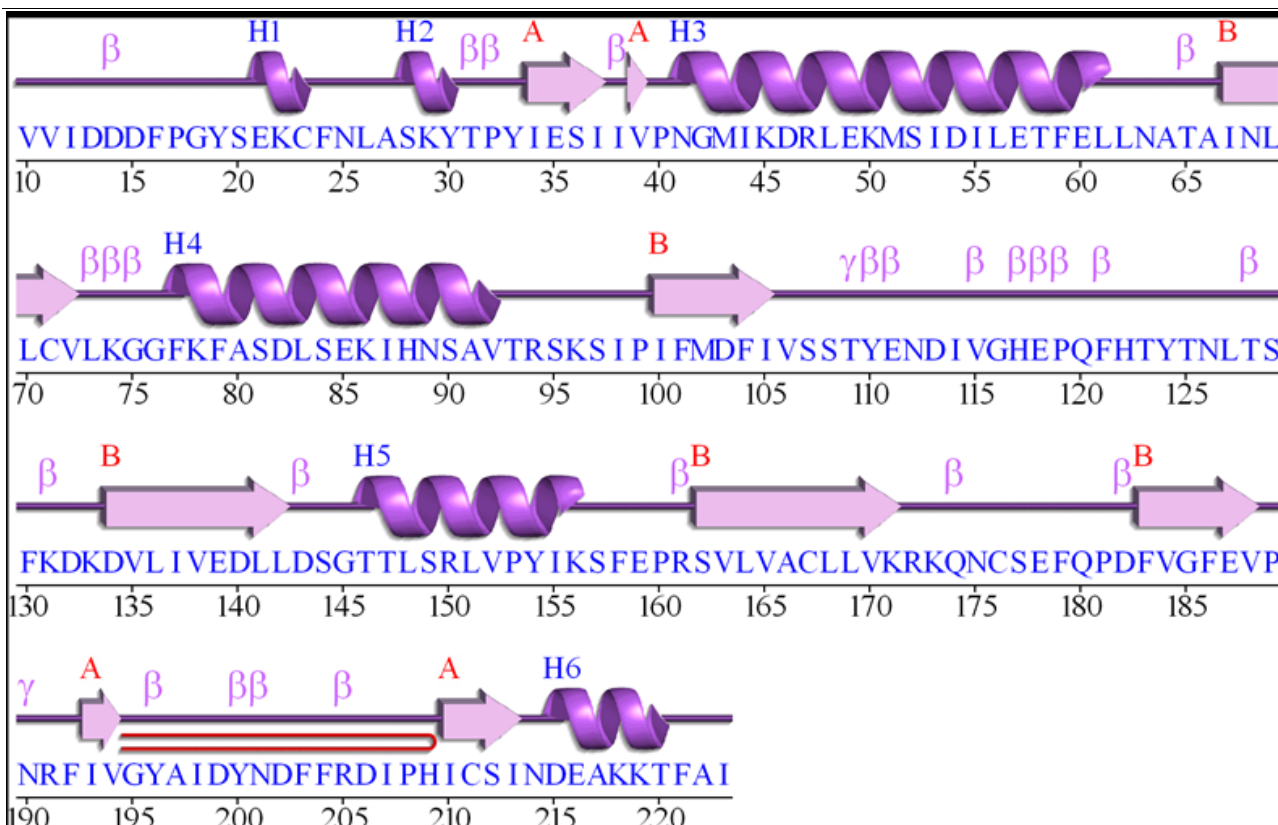
A estrutura secundária das enzimas modeladas é muito similar à enzima molde, mantendo altamente conservado o domínio “core”, composto por 4 fitas paralelas ( $\beta_2$ ,  $\beta_3$ ,  $\beta_4$  e  $\beta_5$ ) formando uma folha  $\beta$  flanqueada por 3  $\alpha$ -hélices ( $\alpha_3$ ,  $\alpha_4$  e  $\alpha_5$ ), a qual liga o substrato primário PRPP, esse domínio na estrutura de Smp\_168500 (Figura 08) compreende os resíduos de 40–172, e em Smp\_148820 os resíduos 41 – 173. O segundo domínio, localizado acima do domínio “core” é chamado domínio “hood”, este é composto por 4 pequenas fitas  $\beta$  ( $\beta_1$ ,  $\beta_6$ ,  $\beta_7$  e  $\beta_8$ ), divididas em 2 folhas, sendo uma paralela e outra antiparalela, 3 pequenas  $\alpha$ -hélices ( $\alpha_1$ ,  $\alpha_2$  e  $\alpha_6$ ) e várias conexões de longos loops. Esse domínio compreende na estrutura de Smp\_168500 os resíduos 10 – 39 e 173 – 223, e em

Smp\_148820 (Figura 8) os resíduos 8 – 40 e 174 – 223.

Em Smp\_168500, o loop I corresponde aos resíduos de 73 à 78 (ILKGGF), já em Smp\_148820, corresponde aos resíduos de 71 à 77 (VLKGGF) localizados entre a fita  $\beta_2$  e a hélice  $\alpha_4$ , que fazem parte do motivo  $\alpha/\beta/\alpha$  no domínio core.

O loop II corresponde em 5IPF aos resíduos de 107 à 134, mas não está completamente ordenado, por ausência de densidade eletrônica para os resíduos de 108 à 124. No modelo de Smp\_168500 corresponde aos resíduos de 106 à 133 e em Smp\_148820 de 107 à 135, estão localizados entre as fitas  $\beta_3$  e  $\beta_4$  do motivo  $\alpha/\beta/\alpha$  do domínio core. Este loop foi caracterizado como flexível previamente em outras estruturas (EADS *et al.*, 1994 *apud* ROMANELLO, 2016).





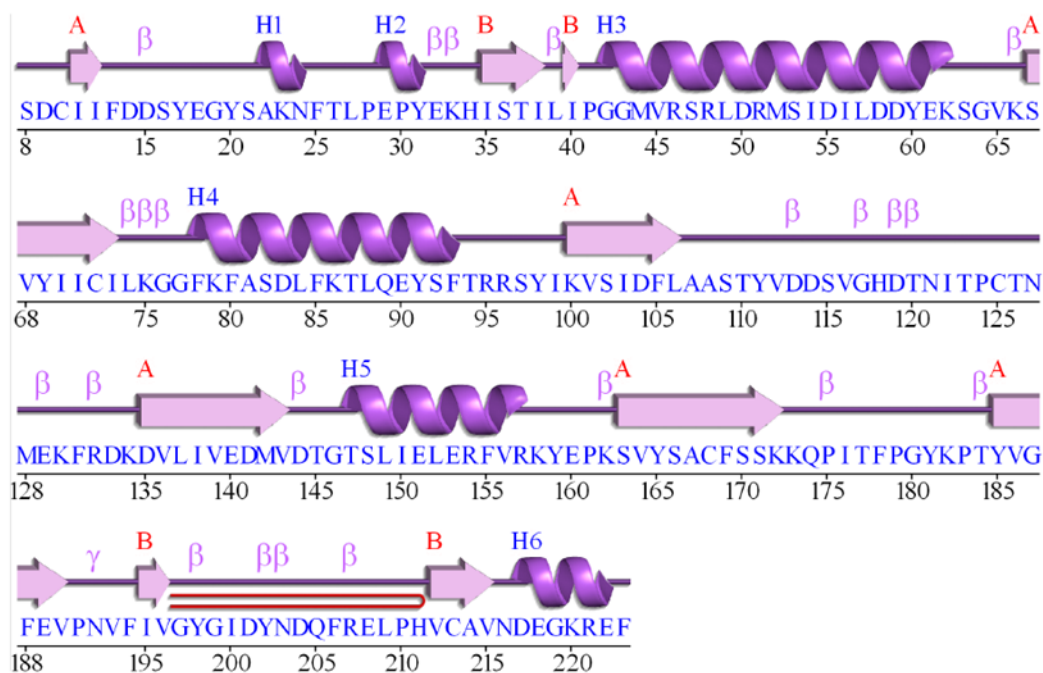
**Figura 8.** Estrutura secundária de Smp\_168500 em função da sequência de aminoácidos. As fitas  $\beta$  são representadas por setas, as  $\alpha$ -hélices por espirais com a respectiva numeração em cima (H1, H2...H25), as 2 folhas  $\beta$  são classificadas como A e B, sendo essas letras colocadas acima das fitas identificando a qual folha pertencem. Os símbolos:  $\beta$  - beta turn,  $\gamma$  - gamma turn e - beta hairpin.

O loop III compreende os resíduos de 140 à 145 em Smp\_168500 e 141 à 146 em Smp\_148820, fica localizado entre as fitas  $\beta 4$  e hélice  $\alpha 5$ , também pertencente ao domínio core. Na estrutura de 5IPF este loop está bem definido. Todas as estruturas de HGPRT apresentam no loop III um Asp141 invariável. Estudos com mutante desse resíduo em *T. cruzi* confirmaram que esse resíduo atua na catálise, mas não é requerido como base geral nessa posição, como era proposto previamente para a HGPRT humana (CANYUK; FOCIA; EAKIN, 2001), na estrutura de Smp\_168500 esse resíduo encontra-se na posição Asp140 (Figura 9).

O loop IV é o único dos quatro que pertence ao domínio hood, é o maior dos quatro e compreende os resíduos de 189 à 210 em Smp\_168500 e 191 à 212 em Smp\_148820. Esta fita está envolvida na ligação da base purina no sítio ativo (MONZANI et al., 2007). Observa-se a diferenciação de alguns aminoácidos nessa região,

como a Val194 em 5IPF, qual atua na ligação de hidrogênio com o N1 da hipoxantina (ROMANELLO, 2016), que em Smp\_168500 e Smp\_148820 uma Isoleucina toma essa posição (Figura 10). Outros aminoácidos que atuam na ligação do IMP no sítio ativo da enzima molde também tiveram diferenciação quanto a posição estrutural nas enzimas modeladas (Figura 11), como em: Ile149 substituídas por Leucinas; Lys147 substituídas por Treoninas; Thr145 substituída por uma Serina em Smp\_168500 e Thr148 sendo substituída por uma Serina em Smp\_148820.

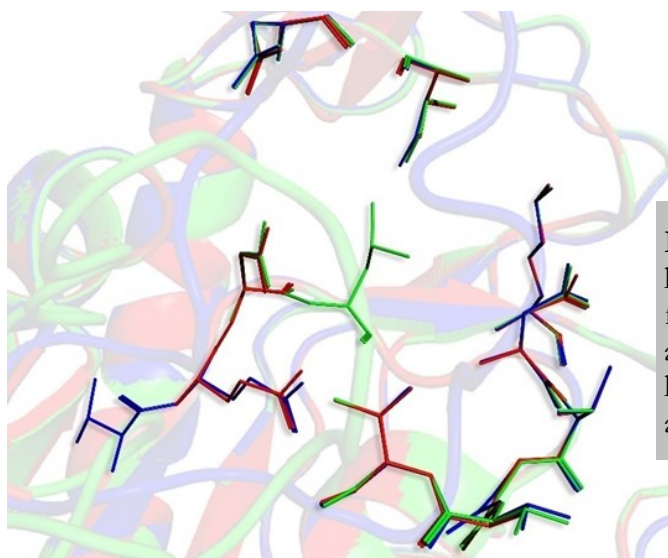
As sobreposições estruturais de *SmHPRT* com sua correspondente humana mostraram regiões extremamente conservadas entre ambas. Principalmente nas regiões entre  $\beta 3$  e  $\alpha 2$ ,  $\beta 5$ ,  $\beta 6$ ,  $\beta 7$ ,  $\beta 8$  e  $\beta 9$ . As regiões conservadas podem ser observadas nas estruturas apresentadas na Figura 12.



**Figura 9.** Estrutura secundária de Smp\_148820 em função da sequência de aminoácidos. As fitas  $\beta$  são representadas por setas, as  $\alpha$ -hélices por espirais com a respectiva numeração em cima (H1, H2...H25), as 2 folhas  $\beta$  são classificadas como A e B, sendo essas letras colocadas acima das fitas identificando a qual folha pertencem. Os símbolos:  $\beta$  - beta turn,  $\gamma$  - gammaturn e - beta hairpin.

		LOOP III
Smp_148820	140	D M V D T G
Smp_168500	141	D L L D S G
5IPF	141	D I I D T G
		LOOP IV
Smp_148820	191	P N V F I V G Y G I D Y N D Q F R E L P H V
Smp_168500	189	P N R F I V G Y A I D Y N D F F R D I P H I
5IPF	190	P N R F V V G Y A L D Y N D N F R D L H H I

**Figura 10.** Alinhamento estrutural dos loops III e IV para as três isoformas de SmHGPRT.



**Figura 11.** Sobreposição de aminoácidos que realizam ligação com o IMP na estrutura cristalográfica (vermelho) com os aminoácidos que ocupam a mesma posição estrutural nas estruturas modeladas, Smp\_168500 em verde e Smp\_148820 em azul.

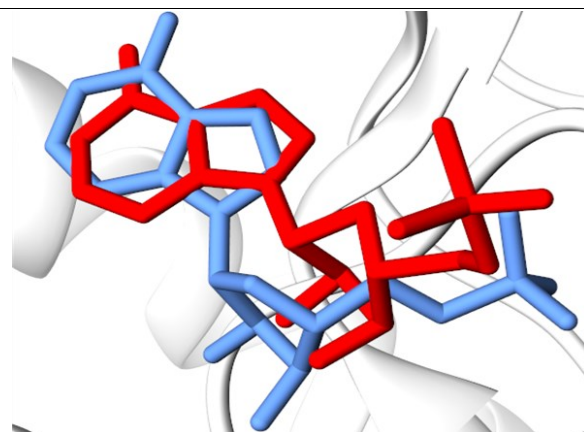


**Figura 12.** Sobreposição de estruturas de HGPR1, humana (amarelo) e de *S. mansoni*-cristalográfica (vermelho), e modelos Smp\_168500 em verde e Smp\_148820 em azul.

### Docking Molecular

Os testes utilizando o IMP como ligante resultaram em uma pontuação (*score*) de -5.5, mostrando que o resultado apontado pelo programa é similar ao demonstrado pela estrutura cristalográfica, com um RMSD de 2,57 (Figura 13). Comparando estrutura e posição dos ligantes testados no sítio ativo com o IMP na estrutura cristalográfica, obteve-se o menor RMSD para o ligante TFV, indicando que a sua posição no sítio ativo, assim como a sua estrutura é muito similar ao do IMP na estrutura cristalográfica (5IPF).

Nas interações dos acilonucleosídeos fosfatados com o sítio ativo da *SmHGPR1* (Figura 15), o ADV obteve o melhor resultado como ligante (*score*), fazendo ligações de hidrogênio com seis aminoácidos (Asp144; Thr145; Gly146; Thr148; Lys172; Val194), quatro deles também fazem ligações com o outro acilonucleosídeo fosfatado, TFV (Asp144, Gly146, Lys172 e Val194). No entanto o ADV apresentou nos testes uma pontuação de *drug-score* inferior à desejada (Tabela 2).



**Figura 13.** Sobreposição de IMPs. Em azul, conformação do IMP na estrutura cristalográfica (5IPF), em vermelho, conformação resultante do *docking* do IMP no sítio ativo da proteína.

Na estrutura de 5IPF em complexo com IMP, o fosfato do IMP faz ligação de hidrogênio com Asp144, Thr145, Gly146 e Lys147 do loop III, Thr148 da hélice  $\alpha 5$  e uma molécula de água. O O2 da ribose faz duas ligações de hidrogênio com Asp141 do loop III (ROMANELLO, 2016).

A interação da Asp141 do loop III com o IMP na estrutura cristalográfica envolve uma molécula de água, o mesmo padrão pode ser observado nas estruturas de *Trypanosoma cruzi*, *Escherichia coli*, humana e de *Toxoplasma gondii*, onde a cadeia lateral aponta para o sítio ativo. No entanto, a cadeia lateral da Asp141 aponta para fora do sítio ativo e não interage com o ligante na estrutura de *Leishmania tarentolae*, *Plasmodium falciparum*, *Tritrichomonas foetus* e humana em complexo com GMP (MONZANI et al., 2007 *apud* ROMANELLO, 2016). Nos ligantes docados, apenas um faz ligação de hidrogênio direta com Asp141, PCV (Figura 14), possivelmente devido à ausência de águas na estrutura utilizada para os testes de docagem.

É possível observar que o loop III é o maior responsável por ancorar a maioria dos ligantes no sítio ativo através de ligações de hidrogênio. A maior parte dos contatos ocorre entre a *SmHGPR1* e o grupo fosfato de TFV e ADV e os oxigênios da região afosfatada de ACV e PCV.

**Tabela 2.** Dados de composição e interação entre aciclonucleosídeos e *Sm*HGPRT. *Score* se refere à uma série de funções envolvendo as superfícies hidrofóbicas, repulsões e ligações de hidrogênio. *APSM* se refere à área polar de superfície molecular. *Drug-score* se refere a uma função baseada em parâmetros físicos, químicos e biológicos que predizem a aptidão de um componente em se tornar um fármaco.

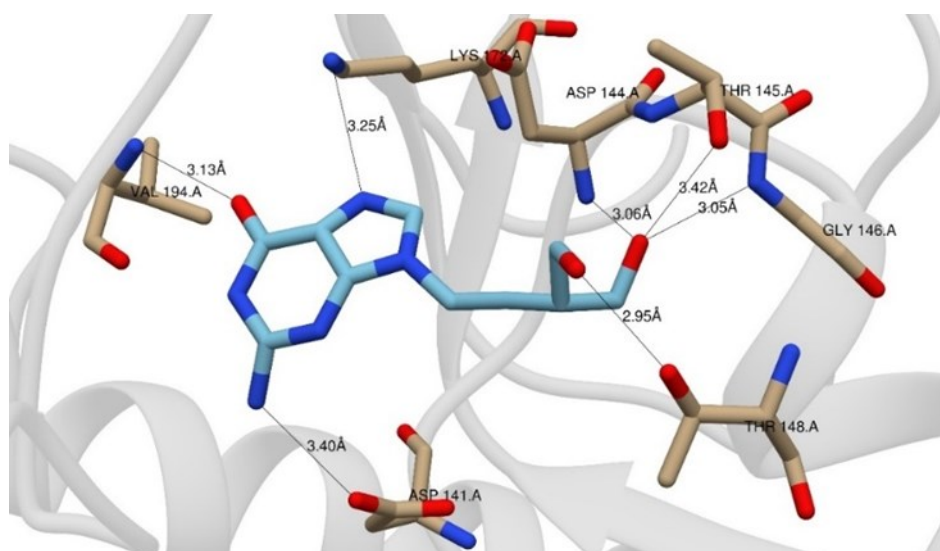
Componente	Pontuação (Score)	g/mol	Lig. H Ligante	Lig. H Recept.	APSM	Pontuação (Drug-score)	RMSD (IMP)
Aciclovir	-5.5	225.21	3	5	119.06	0.157	3.19
Adefovir	-5.6	273.19	4	8	136.39	0.287	3.63
Penciclovir	-6.1	253.26	5	7	130.06	0.308	3.35
Tenofovir	-5.7	287.22	3	5	136.39	0.473	2.60

Três dos aminoácidos presentes no sítio ativo fazem ligações com todos os ligantes testados e com o IMP na estrutura cristalográfica, são eles Asp144, Gly146 e Val194, esses aminoácidos estão descritos por Héroux et al. (1999) como atuantes no mecanismo de

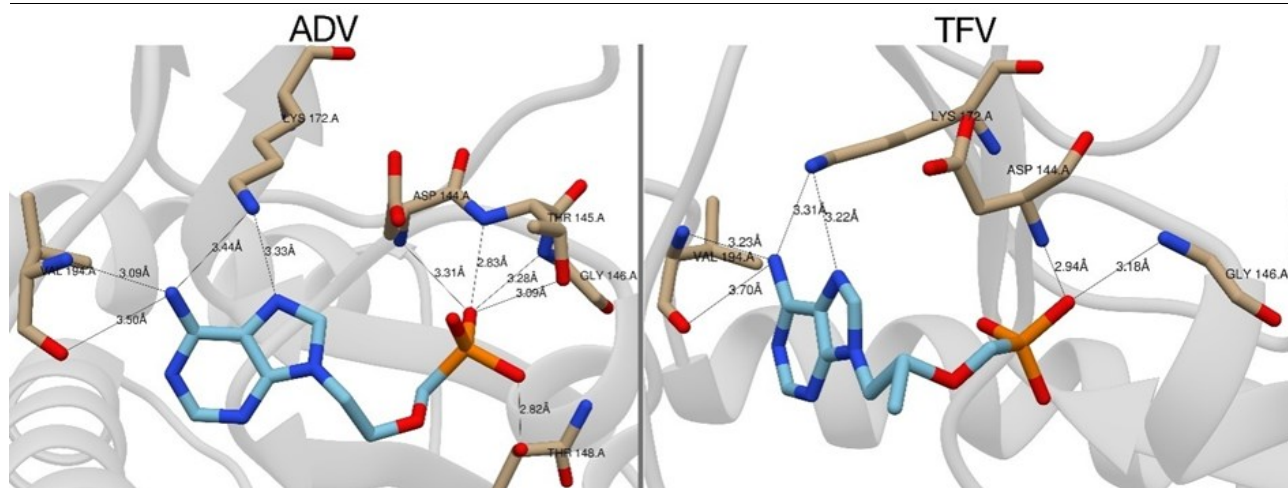
reação enzimática da HGPRT. Outros dois aminoácidos, Thr145 e Lys172 fazem ligações com três dos quatro ligantes. O aminoácido Thr145 faz ligação com três dos ligantes, exceto com TFV (Tabela 3).

**Tabela 3.** Aminoácidos que fazem ligação de hidrogênio direta entre *Sm*HGPRT e o ligante.

Componente	Aminoácidos de interação
Aciclovir	Asp144; Thr145; Gly146; Val194; Asp200
Adefovir	Asp144; Thr145; Gly146; Thr148; Lys172; Val194
Penciclovir	Asp141; Asp144; Thr145; Gly146; Thr148; Lys172; Val194
Tenofovir	Asp144; Gly146; Lys172; Val194
IMP (5IPF)	Asp144, Thr145, Gly146, Lys147, Thr148, Lys172, Val194



**Figura 14.** Interação do PCV com os aminoácidos do sítio ativo da estrutura cristalográfica (5IPF), as distâncias das ligações de hidrogênio estão marcadas como linhas.



**Figura 15.** Interação do aciclonucleotídeos fosfatados (TFV e ADV) com os aminoácidos do sítio ativo da estrutura cristalográfica (5IPF), as distâncias das ligações de hidrogênio estão marcadas como linhas.

## CONCLUSÃO

A busca de novas terapias para as doenças que acometem a humanidade é uma constante no meio científico, havendo sempre a necessidade da introdução ou reposicionamento de novos fármacos no arsenal terapêutico, seja pela falta de eficiência dos fármacos atuais, pelo alto nível de toxicidade dos mesmos, pelo surgimento de novos processos patológicos, ou até mesmo pelo aumento do número de casos de uma doença já existente em uma determinada população. Mas, talvez o maior desafio não seja a descoberta de novas moléculas promissoras, e sim a descoberta de novas propriedades e potenciais de moléculas já existentes.

A aplicação da química computacional e bioinformática tem oferecido um suporte de baixo custo para o desenvolvimento ou reposicionamento desses fármacos. Com o poder computacional e a tecnologia disponível atualmente, pode-se realizar um direcionamento nos estudos, facilitado pela capacidade de predição virtual de interações e propriedades. As técnicas de química computacional, aliadas à bioinformática, permitem uma análise criteriosa do alvo terapêutico a ser estudado, bem como a construção de modelos por homologia para os alvos que ainda não têm estrutura resolvida.

Neste trabalho, foram gerados modelos para as duas isoformas de *Sm*HGPRT ainda não resolvidas experimentalmente, cada

uma das estruturas passou por um processo de validação levando em consideração parâmetros estereoquímicos, de contatos atômicos e enovelamento. Essas avaliações validaram os modelos gerados para o estudo das diferenciações entre as isoformas. Foram observadas diferenciações na posição de aminoácidos no sítio ativo nas estruturas modeladas que poderiam acarretar em interações diferentes dos ligantes com as enzimas. Porém, apenas com métodos experimentais como cristalografia de raios-X ou microscopia crio-eletrônica pode-se confirmar tais modificações.

Após a realização das etapas no teste de docagem na *Sm*HGPRT cristalográfica, observou-se que os compostos aciclonucleotídeos apresentaram afinidade maior que o IMP docado no sítio ativo, sugerindo que os mesmos podem competir ou inibir o sítio ativo da enzima. O composto que obteve melhor resultado (*score*) foi o penciclovir, interagindo com os mesmos aminoácidos participantes da ligação do IMP na estrutura cristalográfica e docada. Assim, a grande perspectiva se encontra na realização futura de ensaios biológicos de atividade dos compostos selecionadas, e dos que ainda serão selecionados em futuros estudos de docagem utilizando *Sm*HGPRT.

## REFERÊNCIAS

- BARSOUM, R. S.; ESMAT, G.; EL-BAZ, T. Humanschistosomiasis: clinical perspective. *J. Adv. Res.*, v. 4, p. 433-444, 2013.

- BRANDEN, C.; TOOZE, J.; Introduction to Protein Structure, Gerland: New York, 1991. In SANTOS-FILHO, O. A., ALENCASTRO, R. B. Modelagem de proteínas por homologia. **Química Nova**. v. 26, n. 2, p. 253-259, 2003.
- CANYUK, B.; FOCIA, P. J.; EAKIN, A. E. The role for an invariant aspartic acid in hypoxanthine phosphoribosyltransferases is examined using saturation mutagenesis, functional analysis, and X-ray crystallography. **Biochemistry**. v. 40, n. 9, p. 2754-65, 2001.
- CANYUK, B.; MEDRANO, F. J.; WENCK, M. A.; FOCIA, P. J.; EAKIN, A. E.; CRAIG, S. P., 3RD. Interactions at the dimer interface influence the relative efficiencies for purine nucleotide synthesis and pyrophosphorolysis in a phosphoribosyltransferase. **Journal of Molecular Biology**. v. 335, n. 4, p. 905-921, 2004.
- CHEN, Q.; LIANG, Y.; SU, X.; GU, X.; ZHENG, X.; LUO, M. Alternative IMP binding in feedback inhibition of hypoxanthine-guanine phosphoribosyltransferase from *Thermoanaerobacter tengcongensis*. **Journal of Molecular Biology**. v. 348, n. 5, p. 1199-1210, 2005.
- CLERCQ, Erik De; HOLY, Antonín. Acyclic nucleoside phosphonates: a key class of antiviral drugs. **Nature Reviews Drug Discovery**. v. 4, p. 928-940, 2005.
- D'ALFONSO, G.; TRAMONTANO, A.; LAHM, A. Structural conservation in single-domain proteins: implications for homology modeling. **Journal of Structural Biology**. v. 134, p. 246-256, 2001.
- EADS, J. C.; SCAPIN, G.; XU, Y.; GRUBMEYER, C.; SACCHETTINI, J. C. The crystal structure of human hypoxanthine-guanine phosphoribosyltransferase with bound GMP. **Cell**. v. 78, n. 2, p. 325-34, 1994.
- FOCIA, P. J.; CRAIG, S. P., 3RD; NIEVES-ALICEA, R.; FLETTERICK, R. J.; EAKIN, A. E. A 1.4 Å crystal structure for the hypoxanthine phosphoribosyltransferase of *Trypanosoma cruzi*. **Biochemistry**. v. 37, n. 43, p. 15066-15075, 1998.
- HEROUX, A.; WHITE, E. L.; ROSS, L. J.; DAVIS, R. L.; BORHANI, D. W. Crystal structure of *Toxoplasma gondii* hypoxanthine-guanine phosphoribosyltransferase with XMP, pyrophosphate, and two Mg(2+) ions bound: insights into the catalytic mechanism. **Biochemistry**. v. 38, n. 44, p. 14495-506, 1999.
- KEITH, J. M.; Bioinformatics: Volume I: Data, Sequence Analysis, and Evolution, **Methods in Molecular Biology**. v. 1525, DOI 10.1007/978-1-4939-6622-6\_8, Springer Science+Business Media New York, 2017.
- KEOUGH, D. T.; HOCKOVÁ, Dana; KRECMEROVÁ, Marcela; CESNEK, Michal; HOLY, Antonín; NAESSENS, Lieve; BRERETON, Ian M.; WINZOR, Donald J.; JERSEYA John de; GUDDAT, Luke W.. *Plasmodium vivax* hypoxanthine-guanine phosphoribosyltransferase: A target for anti-malarial chemotherapy. **Molecular and Biochemical Parasitology Journal**. v. 173, p. 165-169, 2010.
- KEOUGH, D. T.; BRERETON, Ian M.; JERSEY, John de; GUDDAT, Luke W.. The crystal structure of free human hypoxanthine-guanine phosphoribosyltransferase reveals extensive conformational plasticity throughout the catalytic cycle. **Journal of Molecular Biology**. v. 351, n. 1, p. 170-181, 2005.
- KEOUGH, D. T.; HOCKOVA, D.; HOLY, A.; NAESSENS, L. M. J.; SKINNER-ADAMS, T. S.; JERSEY, J.; GUDDAT, L. Inhibition of Hypoxanthine-Guanine Phosphoribosyltransferase by Acyclic Nucleoside Phosphonates: A New Class of Antimalarial Therapeutics. **J. Med. Chem.** v. 52, p. 4391-4399, 2009.
- KIM, S.; THIESSEN, P.A.; BOLTON, E.E.; CHEN, J.; FU, G.; GINDULYTE, A.; HAN, L.; HE, J.; HE, S.; SHOEMAKER, B. A.; WANG, J.; YU, B.; ZHANG, J.; BRYANT, S.H.; PubChem Substance and Compound databases. **Nucleic Acids Res.** v. 44, n. 1, p. 1202-1213, 2016.
- LASKOWSKI, R. A., CHISTYAKOV, V.; THORNTON J. M., 2005. PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. **Nucleic Acids Res.** v. 33, p. 266-268, 2005.

- LASKOWSKI, R. A.; RULLMANN, J. A.; MACARTHUR, M. W.; KAPTEIN, R.; THORNTON, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. **Journal of Biomolecular NMR**. v. 8, n. 4, p. 477-486, 1996.
- MENG, X.; ZHANG, H.; MEZEI, M.; CUI, M. Molecular Docking: A powerful approach for structure-based drug discovery. **CurrComput Aided Drug**. v.7, n. 2, p.146–157, 2011.
- MOLINSPIRATION, Molinspiration Cheminformatics. **Calculation of Molecular Properties and Bioactivity Score**. Disponível em: <<http://www.molinspiration.com/cgi-bin/properties>>. Acessado em 20 de setembro de 2018.
- MONZANI, P. S.; TRAPANI, S.; THIEMANN, O. H.; OLIVA, G. Crystal structure of Leishmania tarentolae hypoxanthine-guanine phosphoribosyltransferase. **BMC Structural Biology**. v. 7, n. 1, p. 59, 2007.
- NOTREDAME, C.; HIGGINS, D. G.; HERRING, J. T-Coffee: A novel method for multiple sequence alignments. **Journal of Molecular Biology**. v.302, p. 205-217, 2000.
- OLEG, TROTT, A, JOLSON. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. **Journal of Computational Chemistry**. v. 31, n. 2, p. 455–461, Jan.2010.
- PEREIRA, T. C. et al. *Schistosoma mansoni*: evaluation of na RNAi-based treatment targetingHGPRTase gene. **Exp. Parasitol**. v. 118, n. 4, p. 619-623, 2008.
- RAMACHANDRAN G.N., RAMAKRISHNAN C., SASISEKHARAN V. Stereochemistryof polypeptide chain configurations. **Journal of Molecular Biology**.v. 7, p. 95-99, 1963.
- RICHARDSON, J. S. The anatomy and taxonomy of protein structure. **Advances in ProteinChemistry**. v. 34, n. 2, p. 167-339, 1981.
- ROMANELLO, L. et. al. *In vitro* and *in vivo* characterization of the multiple isoforms of *Schistosoma mansoni* hypoxanthine-guanine phosphoribosyltransferases. **Molecular and Biochemical Parasitology**. V. 229. p. 24-34, 2019.
- ROMANELLO, L. **Estudos das enzimas adenosina kinase e hipoxantina-guaninafosforibosiltransferase de Schistosoma mansoni**. 2011, 95 f. Dissertação de Mestrado – Universidade Federal de São Carlos, São Carlos, 2011.
- ROMANELLO, L. **Estudos das enzimas adenosina kinaseisoforma 1, hipoxantina-guanina fosforibosiltransferaseisoformas 1, 2 e 3, adenilsuccinato-liase, adenilsuccinatosintetase de Schistosoma mansoni**. 2016, 160 f. Tese de Doutorado - Instituto de Física de São Carlos, Universidade de São Paulo, 2016.
- SCHRÖDINGER, LLC. **ThePyMOL Molecular Graphics System, Version 2.0**.Disponível em: <<https://pymol.org/2/>>. Acessado em 20 de setembro de 2018.
- STERLING, T.; IRWIN J.; J. ZINC 15 – Ligand Discovery for Everyone. **Journal of Chemical Information and Modeling**, v. 55, n. 11, p. 2324-2337, 2015.
- TROTT, O.; A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading.**Journal of Computational Chemistry**. v. 31, p. 455-461, 2010.
- WATERHOUSE, A., BERTONI, M., BIERNERT, S., STUDER, G., TAURIELLO, G., GUMIENNY, R., HEER, F.T., de Beer, T.A.P., REMPFER, C., BORDOLI, L., LEPORE, R., SCHWEDE, T. SWISS-MODEL: homology modeling of protein structures and complexes. **NucleicAcids Res**. v. 46, n.1, p. 296-303, 2018.

Recebido em: 08/05/2020.

Aceito em: 13/07/2020.