

Ministério da Saúde  
FUNDAÇÃO OSWALDO CRUZ  
INSTITUTO OSWALDO CRUZ

Doutorado em Biologia Computacional e Sistemas

IDENTIFICAÇÃO *IN SILICO* DE POTENCIAIS INIBIDORES DA RIBOSE  
5-FOSFATO ISOMERASE DE *TRYPANOSOMA CRUZI*

VANESSA DE VASCONCELOS SINATTI CASTILHO

Rio de Janeiro  
Outubro de 2017

Castilho, Vanessa de Vasconcelos Sinatti.

Identificação in silico de potenciais inibidores da ribose 5-fosfato isomerase de *Trypanosoma cruzi* / Vanessa de Vasconcelos Sinatti Castilho. - Rio de Janeiro, 2017.

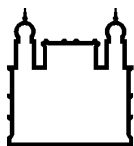
210 f.; il.

Tese (Doutorado) - Instituto Oswaldo Cruz, Pós-Graduação em Biologia Computacional e Sistemas, 2017.

Orientadora: Ana Carolina Ramos Guimarães.

Bibliografia: f. 128-149

1. Doença de Chagas. 2. Ribose 5-fosfato isomerase. 3. Modelagem molecular. 4. Atracamento molecular. 5. Dinâmica molecular. I. Título.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

## **INSTITUTO OSWALDO CRUZ**

**Programa de Pós-Graduação em Biologia Computacional e Sistemas**

*Vanessa de Vasconcelos Sinatti Castilho*

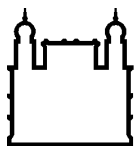
Identificação *in silico* de Potenciais Inibidores da Ribose 5-Fosfato Isomerase de *Trypanosoma cruzi*

Tese apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Doutor em Biologia Computacional e Sistemas

**Orientador:** Dra. Ana Carolina Ramos Guimarães

**RIO DE JANEIRO**

Outubro de 2017



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

## **INSTITUTO OSWALDO CRUZ**

**Programa de Pós-Graduação em Biologia Computacional e Sistemas**

***AUTOR: VANESSA DE VASCONCELOS SINATTI CASTILHO***

**IDENTIFICAÇÃO *IN SILICO* DE POTENCIAIS INIBIDORES DA RIBOSE 5-FOSFATO ISOMERASE DE *TRYPANOSOMA CRUZI***

**ORIENTADOR: Dra. Ana Carolina Ramos Guimarães**

Aprovada em: \_\_\_\_/\_\_\_\_/\_\_\_\_

### **EXAMINADORES:**

**Prof. Dr. Floriano Paes Silva Junior (IOC/FIOCRUZ) - Presidente**

**Prof. Dr. Eduardo Caio Torres dos Santos (IOC/FIOCRUZ) - membro**

**Prof. Dra. Ana Carolina Rennó Soderó (Faculdade de Farmácia/UFRJ) - membro**

**Prof. Dr. Pedro Henrique Monteiro Torres (IOC/FIOCRUZ) - suplente**

**Prof. Dra. Isabella Alvim Guedes (LNCC) - suplente**

Rio de Janeiro, 31 de outubro de 2017

Dedico essa tese à minha família, em especial à minha mãe e a minha avó.

## **AGRADECIMENTOS**

Para a realização de um grande sonho, o apoio, a confiança e a amizade são fundamentais e é por isso que tenho muito a agradecer.

Agradeço a Deus por ter me dado força e discernimento para eu alcançar meus objetivos e realizar os meus sonhos.

Aos meus pais por terem me apoiado em todas as minhas decisões, permitido que eu prosseguisse com esse sonho, principalmente a minha mãe, que sempre esteve ao meu lado me incentivando e tornando a minha vida mais feliz e amena.

À minha vó Magaly por ter sido a melhor avó do mundo e ter me dado as melhores lembranças da minha vida.

Ao meu amor, Phillippe, por ser o melhor companheiro do mundo. Como é possível ser tão feliz com alguém? Agradeço também a todos os membros da grande família que ganhei: Rosa, Chris, Kayque, Aninha, Guta e Beto! Obrigada pelos bons momentos e por todo afeto.

Aos meus lindos, Vick e Simba, por todo o companheirismo e amizade que só cachorrinhos podem dar.

Aos amigos e colegas que fiz na Fiocruz e, especialmente, às minhas amigas Claudia e a Andressa, que tornaram essa caminhada mais leve e feliz.

A todos os membros e amigos do Laboratório de Genômica Funcional e Bioinformática por fazerem parte da minha vida científica e terem me proporcionado grandes momentos.

Aos membros do PROCC, em especial ao Dr. Ernesto Caffarena, que participaram ativamente no início desse doutorado e deram muitas contribuições.

Ao Dr. Laurent Dardenne e ao Dr. João Hermínio por toda ajuda e colaboração que tornou possível a realização desse trabalho.

À minha orientadora Dra. Ana Carolina Guimarães (Carol), por ter me apresentado e me guiado por essa área incrível que é a Biologia Computacional/Bioinformática e, principalmente, por ter sido a minha “mãe”/mentora científica por esses longos anos.

Ao Dr. Marcelo A. Ferreira por toda disposição e empenho em me ajudar desde o início dessa jornada científica em 2008.

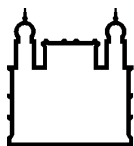
A todos aqueles que acreditaram em mim e me ajudaram de alguma forma a alcançar esse sonho. Muito obrigada!

À pós-graduação em Biologia Computacional e Sistemas por ter tornado esse doutorado possível.

Agradeço à Vice-presidência da Fiocruz pelo auxílio financeiro.

“It is impossible to convey completely the excitement of discovery, of seeing the result of an experiment and knowing that you know something new, something fundamental, and that for this moment at least, only you, in the entire world, knows it.”

Stuart Firestein



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

## INSTITUTO OSWALDO CRUZ

### IDENTIFICAÇÃO *IN SILICO* DE POTENCIAIS INIBIDORES DA RIBOSE 5-FOSFATO ISOMERASE DE *TRYPANOSOMA CRUZI*

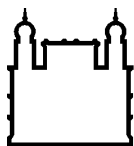
#### RESUMO

#### TESE DE DOUTORADO EM BIOLOGIA COMPUTACIONAL E SISTEMAS

Vanessa de Vasconcelos Sinatti Castilho

A doença de Chagas, causada pelo protozoário *Trypanosoma cruzi*, afeta aproximadamente sete milhões de pessoas, principalmente na América Latina, e causa 7000 mortes ao ano. Os tratamentos disponíveis são insatisfatórios e a busca por fármacos mais efetivas contra este patógeno é essencial. Nesse contexto, a enzima ribose 5-fosfato isomerase (Rpi) é um alvo de fármacos em potencial principalmente devido à sua função na via das pentoses fosfato e a sua essencialidade (anteriormente demonstrada em outros tripanossomatídeos). Neste estudo, novos inibidores potenciais foram propostos para a Rpi de *T. cruzi* (TcRpiB) com base em uma abordagem assistida por computador, incluindo a modelagem de farmacóforo baseada em estrutura e baseada em ligante. Juntamente com uma busca baseada em similaridade e subestrutura, as hipóteses de farmacóforo selecionadas foram utilizadas para a triagem virtual do subconjunto *purchasable* do banco de dados ZINC, produzindo 20.183 compostos candidatos. Estes compostos foram submetidos a estudos de atracamento molecular nos sítios ativos de TcRpiB e de Rpi de Humano (HsRpiA) visando a seleção de moléculas com alta afinidades de ligação predita em relação a enzima de *T. cruzi* mas pouca ou nenhuma atividade em relação a HsRpiA. Após o atracamento molecular e filtragens baseadas em predição de ADME-Tox (absorção, distribuição, metabolismo, excreção e toxicidade) / PAINS (*pan-assay interference compounds*), 211 moléculas foram selecionadas como potenciais inibidores de TcRpiB. Destes, três compostos - ZINC36975961, ZINC63480117 e ZINC43763931 - foram submetidos a simulações de dinâmica molecular e dois deles - ZINC36975961 e ZINC43763931 - tiveram bom desempenho e fizeram interações com resíduos importantes do sítio ativo da enzima alvo durante todo o tempo de simulação. Estes compostos podem ser considerados potenciais candidatos a inibidores de TcRpiB e também podem ser utilizados como compostos líderes para o desenvolvimento de novos inibidores de TcRpiB.





Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

## INSTITUTO OSWALDO CRUZ

### *IN SILICO* IDENTIFICATION OF POTENCIAL INHIBITORS OF RIBOSE 5-FOSFATE ISOMERASE FROM *TRYPANOSOMA CRUZI*

#### ABSTRACT

#### PHD THESIS IN COMPUTATIONAL AND SYSTEMS BIOLOGY

Vanessa de Vasconcelos Sinatti Castilho

Chagas disease, caused by the protozoan *Trypanosoma cruzi*, affects approximately seven million people, mainly in Latin America, and causes about 7,000 deaths annually. The available treatments are unsatisfactory and search for more effective drugs against this pathogen is critical. In this context, the ribose 5-phosphate isomerase (Rpi) enzyme is a potential drug target mainly due to its function in the pentose phosphate pathway and its essentiality (previously shown in other trypanosomatids). In this study, novel potential inhibitors were proposed for the Rpi of *T. cruzi* (TcRpi) based on a computer-aided approach, including structure-based and ligand-based pharmacophore modeling. Along with a substructural and similarity search, the selected pharmacophore hypotheses were used to screen the purchasable subset of the ZINC Database, yielding 20,183 candidate compounds. These compounds were submitted to molecular docking studies in the TcRpi and Human Rpi (HsRpi) active sites in order to select molecules with high predicted binding affinities towards to the *T. cruzi* enzyme but little or none to the HsRpi. After the molecular docking and ADME-Tox (absorption, distribution, metabolism, excretion and toxicity) /PAINS (pan-assay interference compounds) screenings, 211 molecules were selected as potential TcRpi inhibitors. Out of these, three compounds - ZINC36975961, ZINC63480117, and ZINC43763931 - were submitted to molecular dynamics simulations and two of them - ZINC36975961 and ZINC43763931- had Bom performance and made interactions with important active site residues over all the simulation time. These compounds could be considered potential TcRpi inhibitors candidates and also may be used as leads for developing new TcRpi inhibitors.

# SUMÁRIO

	<b>Agradecimentos</b> .....	<b>v</b>
	<b>Sumário</b> .....	<b>ix</b>
	<b>Índice de Figuras</b> .....	<b>xii</b>
	<b>Lista de Tabelas</b> .....	<b>xv</b>
	<b>Lista de Siglas e Abreviaturas</b> .....	<b>xvi</b>
<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
	1.1 <b>Doença de Chagas</b> .....	<b>1</b>
	1.2 <b>Trypanosoma cruzi – Diversidade intraespecífica e biologia celular</b> .....	<b>4</b>
	1.3 <b>Ciclo de vida do Trypanosoma cruzi</b> .....	<b>6</b>
	1.4 <b>Tratamento da doença de Chagas</b> .....	<b>8</b>
	1.5 <b>Estudos pré-clínicos e clínicos de candidatos a fármacos para a doença de Chagas</b> .....	<b>10</b>
	1.6 <b>Alvos moleculares na doença de Chagas</b> .....	<b>12</b>
	1.7 <b>Via das pentoses fosfato</b> .....	<b>13</b>
	1.7.1 <b>Ribose 5-fosfato isomerase</b> .....	<b>17</b>
	1.7.1.1 <b>A essencialidade da ribose 5-fosfato isomerase na viabilidade celular</b> .....	<b>25</b>
	1.7.1.2 <b>Inibidores de ribose 5-fosfato isomerasas</b> .....	<b>27</b>
	1.8 <b>Planejamento de novos fármacos auxiliado por computador</b> ...	<b>31</b>
	1.8.1 <b>Desenho de fármacos baseado em ligante (DFBL)</b> .....	<b>34</b>
	1.8.2 <b>Desenho de fármacos baseado em estrutura (DFBE)</b> ....	<b>35</b>
	1.8.2.1 <b>Modelagem de farmacóforo</b> .....	<b>35</b>
	1.8.2.2 <b>Atracamento molecular</b> .....	<b>38</b>
	1.8.2.3 <b>Dinâmica molecular</b> .....	<b>42</b>
	1.8.3 <b>ADME-Tox (Absorção, Distribuição, Metabolismo, Excreção e Toxicidade)</b> .....	<b>46</b>
	1.9 <b>Justificativa</b> .....	<b>50</b>
<b>2</b>	<b>OBJETIVOS</b>	<b>52</b>
	2.1 <b>Objetivo Geral</b> .....	<b>52</b>
	2.2 <b>Objetivos Específicos</b> .....	<b>52</b>
<b>3</b>	<b>MATERIAL E MÉTODOS</b>	<b>53</b>

<b>3.1</b>	<b>Conjunto de dados utilizados</b> .....	<b>54</b>
<b>3.2</b>	<b>Análises das proteínas</b> .....	<b>56</b>
3.2.1	Comparação das proteínas TcRpiB e HsRpiA.....	56
3.2.2	Análise das moléculas de águas conservadas no sítio ativo de TcRpiB .....	57
<b>3.3</b>	<b>Modelagem de farmacóforo</b> .....	<b>57</b>
3.3.1	Modelagem de farmacóforo baseado em ligante.....	57
3.3.2	Modelagem de farmacóforo baseado em estrutura .....	59
3.3.3	Triagem virtual baseada em farmacóforo e similaridade 2D .....	59
<b>3.4</b>	<b>Atracamento molecular</b> .....	<b>60</b>
3.4.1	Preparação dos ligantes .....	60
3.4.2	Preparação das proteínas.....	61
3.4.3	<i>Redocking</i> e avaliação dos programas de atracamento AutoDock Vina e Glide XP .....	62
3.4.4	Triagem virtual baseada em atracamento molecular .....	63
<b>3.5</b>	<b>Predição ADME-Tox e PAINS</b> .....	<b>64</b>
<b>3.6</b>	<b>Clusterização de compostos</b> .....	<b>65</b>
<b>3.7</b>	<b>Simulação por dinâmica molecular</b> .....	<b>66</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>68</b>
<b>4.1</b>	<b>Análises de proteínas</b> .....	<b>68</b>
4.1.1	Comparação das proteínas TcRpiB e HsRpiA.....	68
4.1.2	Análise das moléculas de águas conservadas no sítio ativo da estrutura cristalográfica TcRpiB (PDB 3K7S) ....	70
<b>4.2</b>	<b>Modelagem de farmacóforo</b> .....	<b>75</b>
4.2.1	Modelagem de farmacóforo baseado em ligante.....	75
4.2.2	Modelagem de farmacóforo baseado em estrutura .....	78
4.2.3	Triagem virtual baseada em farmacóforo e similaridade 2D .....	80
<b>4.3</b>	<b>Atracamento molecular</b> .....	<b>82</b>
4.3.1	Avaliação dos programas de atracamento AutoDock Vina e Glide XP.....	82
4.3.1.1	Estudos de <i>Redocking</i> com TcRpiB utilizando os programas AutoDock Vina e Glide XP .....	82

4.3.1.2	Ensaio de atracamento com AutoDock Vina e Glide XP utilizando a TcRpiB e o conjunto de inibidores baseados em similaridade 2D .....	86
4.3.2	Análise de desempenho do programa Glide XP com base na curva ROC .....	93
4.3.3	Triagem virtual baseada em atracamento molecular .....	94
<b>4.4</b>	<b>Predição ADME-Tox e PAINS .....</b>	<b>100</b>
<b>4.5</b>	<b>Clusterização dos compostos.....</b>	<b>106</b>
<b>4.6</b>	<b>Análises de interação molecular dos compostos mais bem ranqueados em relação a TcRpiB.....</b>	<b>115</b>
<b>4.7</b>	<b>Simulações de dinâmica molecular .....</b>	<b>121</b>
<b>5</b>	<b>CONCLUSÕES</b>	<b>126</b>
<b>6</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>128</b>
<b>7</b>	<b>APÊNDICES</b>	<b>150</b>
<b>APÊNDICE A -</b>	<b>DESCRIÇÕES DAS PROPRIEDADES DOS</b>	<b>211</b>
<b>POTENCIAIS INIBIDORES DE TCRPIB</b>		<b>150</b>
<b>APÊNDICE B -</b>	<b>ESTRUTURAS QUÍMICAS DOS COMPOSTOS CITADOS</b>	
<b>NA TESE</b>	<b>162</b>	
<b>APÊNDICE C -</b>	<b>PRODUÇÃO ACADÊMICA NO PERÍODO REFERENTE AO</b>	
<b>DOUTORADO</b>	<b>176</b>	
<b>APÊNDICE D -</b>	<b>PUBLICAÇÃO CIENTÍFICA RELACIONADA À TESE</b>	<b>179</b>

## ÍNDICE DE FIGURAS

<b>Figura 1.1:</b> Mapa com a representação da distribuição dos casos de doença de Chagas, baseado em estimativas oficiais de 2010-2013. ....	2
<b>Figura 1.2:</b> Representação do corte longitudinal das três formas evolutivas de <i>T. cruzi</i> .....	6
<b>Figura 1.3:</b> Ciclo de vida do <i>T. cruzi</i> . ....	7
<b>Figura 1.4:</b> Esquema geral da via das pentoses fosfato.....	14
<b>Figura 1.5:</b> Reações do ramo oxidativo da via das pentoses fosfato. ....	15
<b>Figura 1.6</b> Reação de isomerização da ribose 5-fosfato isomerase. ....	17
<b>Figura 1.7:</b> Mecanismo de reação de HsRpiA. ....	21
<b>Figura 1.8:</b> Alinhamento múltiplo das estruturas primárias de RpiA. ....	22
<b>Figura 1.9:</b> Mecanismo de reação de TcRpiB. ....	24
<b>Figura 1.10:</b> Alinhamento múltiplo das estruturas primárias de RpiB de alguns organismos.....	24
<b>Figura 1.11:</b> Exemplo de uma equação de campo de força. ....	43
<b>Figura 3.1:</b> Visão geral dos procedimentos para a avaliação das melhores condições do sítio ativo quanto a presença moléculas de águas conservadas e dos programas de atracamento molecular AutoDock Vina e Glide XP. ....	53
<b>Figura 3.2:</b> Visão geral das etapas desenvolvidas na identificação de potenciais inibidores de TcRpiB. ....	54
<b>Figura 3.3:</b> Inibidores de RpiBs utilizados na modelagem de farmacóforo baseado em ligante.....	56
<b>Figura 4.1:</b> Comparação entre as estruturas primárias, secundárias e terciárias de HsRpiA e TcRpiB. ....	70
<b>Figura 4.2:</b> Alinhamento múltiplo entre as sequências de aminoácidos da ribose 5-fosfato isomerase de <i>T. cruzi</i> (TCR), de <i>M. tuberculosis</i> (MTB), de <i>E. coli</i> (ECO), de <i>C. thermocellum</i> (CTH) e de <i>A. phagocytophilum</i> (APH). ....	73
<b>Figura 4.3:</b> Hipóteses de farmacóforo baseadas em ligante. ....	78
<b>Figura 4.4:</b> Hipóteses de farmacóforo baseadas em estrutura.....	80
<b>Figura 4.5:</b> (A) Pose do ligante cristalográfico R5P em relação à proteína alvo TcRpiB (PDB ID: 3K7S). ....	84
<b>Figura 4.6:</b> Interações moleculares 2D indicadas pelo MAESTRO entre o receptor TcRpiB e os mais bem classificados nos ensaios de atracamento com AutoDock Vina e Glide XP envolvendo ausência de água (A, D), com presença de moléculas	

de águas muito conservadas (B,E) e com presença de moléculas de águas conservadas (C, F).....	92
<b>Figura 4.7:</b> Curva ROC do desempenho da função de pontuação do Glide XP em relação a proteína TcRpiB com ausência de moléculas de águas cristalográficas no sítio ativo. ....	94
<b>Figura 4.8:</b> Interação 2D dos primeiros colocados nas triagens baseadas em atracamento molecular com o programa Glide XP para as proteínas TcRpiB e HsRpiA. ....	99
<b>Figura 4.9:</b> Modo de ligação do composto ZINC33832844, 2 <sup>o</sup> e 3 <sup>o</sup> colocado das triagens com TcRpiB e HsRpiA, respectivamente, no sítio ativo dessas enzimas..	100
<b>Figura 4.10:</b> <i>Heat map</i> e dendograma referente a clusterização dos 211 potenciais inibidores de TcRpiB. ....	109
<b>Figura 4.11:</b> Resultado da clusterização dos 211 potenciais inibidores de TcRpiB. ....	110
<b>Figura 4.12:</b> Continuação da Figura 4.11 com compostos representativos de cada grupo. ....	111
<b>Figura 4.13:</b> Ramificação de arcabouços moleculares, com até dois níveis, representando o conjunto final de potenciais inibidores de TcRpiB. ....	112
<b>Figura 4.14:</b> Continuação da Figura 4.13 contendo ramificação de arcabouços moleculares, com até dois níveis, representando o conjunto final de potenciais inibidores de TcRpiB. ....	113
<b>Figura 4.15:</b> Arcabouços moleculares do conjunto final de potenciais inibidores de TcRpiB. ....	114
<b>Figura 4.16:</b> Modo de ligação predito entre 4PEH e proteína TcRpiB. (A) O modo de ligação de 4PEH no sítio ativo de TcRpiB predito na abordagem de atracamento molecular.....	116
<b>Figura 4.17:</b> Modo de ligação predito entre ZINC36975961 e a proteína TcRpiB..	118
<b>Figura 4.18:</b> Modo de ligação predito entre ZINC63480117 e a proteína TcRpiB..	119
<b>Figura 4.19:</b> Modo de ligação predito entre ZINC43763931 e a proteína TcRpiB..	120
<b>Figura 4.20:</b> Resultados dos estudos de dinâmica molecular dos novos inibidores potenciais e o 4PEH no sítio ativo de TcRpiB.. ....	123
<b>Figura 4.21:</b> Resultados dos estudos de dinâmica molecular dos novos inibidores potenciais e o 4PEH no sítio ativo de TcRpiB. ....	124
<b>Figura 7.1</b> Estruturas químicas dos 211 compostos finais de TcRpiB.....	162

**Figura 7.2** Estruturas químicas de alguns dos compostos químicos citados na tese..

.....**173**

## LISTA DE TABELAS

Tabela 1.1: Estruturas tridimensionais de Rpis resolvidas experimentalmente.....	19
Tabela 1.2: Inibidores testados nas RpiBs .....	29
Tabela 1.3: Inibidores testados nas RpiAs .....	31
Tabela 4.1: Estruturas cristalográficas de RpiBs utilizadas nas análises de moléculas de águas estruturais no sítio ativo de TcRpiB .....	71
Tabela 4.2: Análise das moléculas de águas conservadas no sítio ativo das estruturas cristalográficas com similaridade maior que 50% com a proteína alvo TcRpiB (PDB 3K7S). .....	74
Tabela 4.3: Hipóteses de farmacóforo baseadas em ligante geradas a partir do conjunto de inibidores selecionados de RpiB .....	76
Tabela 4.4: Resultados dos <i>redockings</i> realizados utilizando diferentes preparações do receptor. ....	85
Tabela 4.5: Resultado do ensaio de atracamento com o programa AutoDock VINA utilizando o receptor em três condições em relação à presença de moléculas de água no sítio ativo. ....	90
Tabela 4.6: Resultado do ensaio de atracamento com o programa Glide XP utilizando o receptor em três condições em relação à presença de moléculas de água no sítio ativo .....	91
Tabela 4.7: Resultado da triagem virtual baseada em atracamento molecular com o programa Glide XP para as proteínas TcRpiB e HsRpiA .....	98
Tabela 4.8: Propriedades ADME preditas, filtros PAINS e DS dos controles positivos e dos dez compostos mais bem ranqueados para TcRpiB em relação a HsRpiA...	103
Tabela 7.1: Propriedades ADME preditas, filtros PAINS e DS dos 211 compostos finais para TcRpiB em relação a HsRpiA e seus respectivos grupos (clusters). ....	150
Tabela 7.2: Compostos com propriedades farmacológicas conhecidas, ou com mesmo arcabouço molecular de tais compostos (somente quiralidade distinta) entre o conjunto final após filtro ADME-Tox e PAINS (exceto metabólitos humanos).....	159



## LISTA DE SIGLAS E ABREVIATURAS

DTU	Unidade discreta de tipagem
R5P	D-ribose-5-fosfato
Ru5P	D-ribulose-5-fosfato
All6P	D-alose 6-fosfato
Allu6P	D-alulose 6-fosfato
Rpi	ribose 5-fosfato isomerase
RpiA	ribose 5-fosfato isomerase tipo A
RpiB	ribose 5-fosfato isomerase tipo B
TcRpiB	Rpi de <i>T. cruzi</i>
HsRpiA	Rpi de humanos
EcRpiA	Rpi de <i>E. coli</i>
SoRpiA	Rpi de <i>S. oleracea</i>
MtRpiB	RpiB de <i>M. tuberculosis</i>
LdRpiB	RpiB de <i>L. donovani</i>
LiRpiB	RpiB de <i>Leishmania infantum</i>
TbRpiB	RpiB de <i>T. Brucei</i>
LmRpiB	RpiB de <i>L. Major</i>
4PEA	4-fosfo-D-eritronato
4PEH	ácido 4-fosfo-D-eritronohidroxâmico
4PEHz	4-fosfo-D-eritronohidrazida
4PEAm	4-fosfo-D-eritronamida
4PMEA	4-fosfonometil-D-eritronato
5PRH	ácido 5-fosfo-D-ribonohidroxâmico
5PRA	5-fosfo-D-ribonato
5PRAm	5-fosfo-D-ribonamida
5PRHz	N-(5-fosfo-D-ribonil)-hidrazina
5PRMA	N-(5-fosfo-D-ribonil)-metilamina
5PRGly	N-(5-fosfo-D-ribonil)-glicina
5PRGABA	N-(5-fosfo-D-ribonil)-c-aminobutanoato
CADD	<i>Computer-Aided Drug Design</i>
DFBL	Desenho de fármacos baseado em ligante
DFBE	Desenho de fármacos baseado em estrutura
DM	Dinâmica molecular

ROC	<i>receiver operating characteristic)</i>
AUC	área sob curva
DS	<i>Docking score</i>
AC	Águas conservadas
AMC	Águas muito conservadas
ADME-Tox	Absorção, distribuição, metabolismo, excreção e toxicidade
PAINS	<i>Pan-assay interference compounds</i>

# 1 INTRODUÇÃO

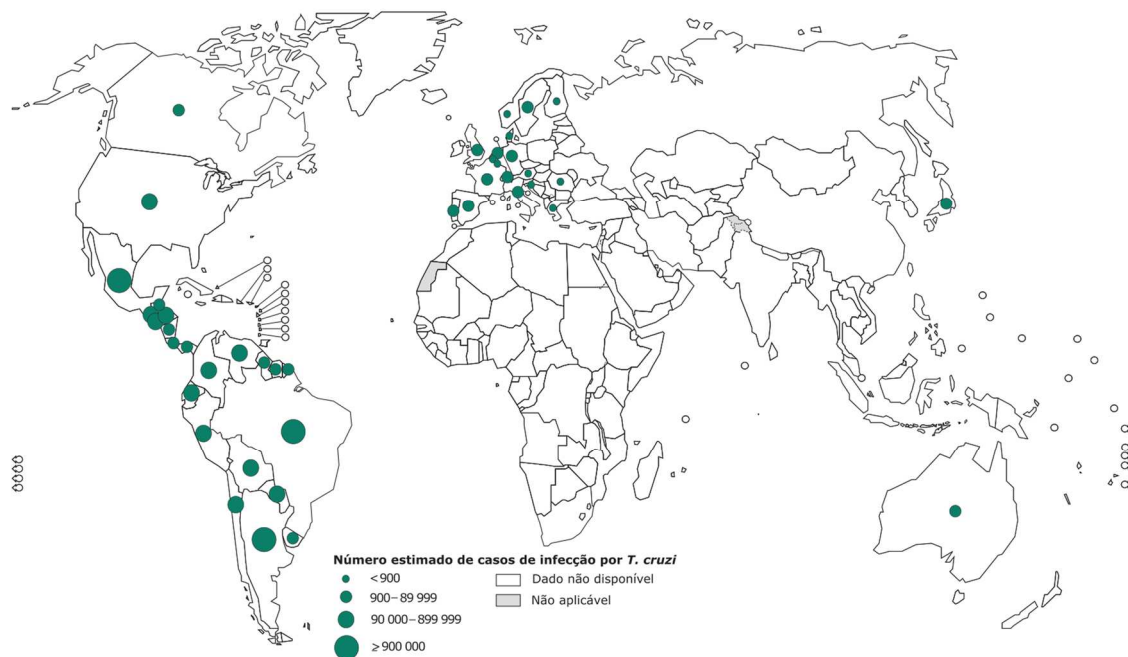
## 1.1 Doença de Chagas

A doença de Chagas, causada pelo protozoário flagelado *Trypanosoma cruzi*, é uma doença endêmica de países da América Latina, entre eles o Brasil. No entanto, cada vez mais casos de doenças de Chagas tem sido relatados em países não endêmicos, como EUA, Canadá, Austrália, Japão e alguns países da Europa, em consequência do movimento migratório de pessoas infectadas (1,2) (Figura 1.1). Estimativas da Organização Mundial de Saúde (OMS) apontam que essa doença afeta cerca de sete milhões de pessoas em todo mundo e causa aproximadamente 10.000 mortes por ano, representando um importante problema de saúde pública de alcance mundial (3,4).

Além do ser humano, outros mamíferos, tais como marsupiais, morcegos, roedores e carnívoros, podem ser contaminados com o *T. cruzi*, representando os chamados animais reservatórios. A principal forma de transmissão da doença de Chagas é a vetorial através do inseto triatomíneo (Ordem Hemiptera, Família Reduviidae, Subfamília Triatominae). Das mais de 140 espécies de triatomíneos já identificadas, aquelas pertencentes aos gêneros *Panstrongylus*, *Rhodnius* e *Triatoma* são as mais importantes em termos da transmissão vetorial da doença de Chagas. Na América do Sul, o *Triatoma infestans* representa uma das principais espécies vetoras, mas foi recentemente eliminado no Uruguai, Chile e Brasil. Outras espécies que se destacam como vetores do *T. cruzi* são *Rhodnius prolixus*, *Rhodnius ecuadoriensis*, *Triatoma dimidiata*, *Panstrongylus geniculatus*, *Rhodnius ecuadoriensis*, *Rhodnius brethesi*, *Panstrongylus megistus*, *Triatoma maculata*, *Triatoma pseudomaculata*, *Triatoma sórdida* e *Triatoma brasiliensis*, sendo que muitas dessas espécies estão presentes no Brasil (5).

A transfusão de sangue é o segundo mecanismo mais frequente de transmissão da doença de Chagas, destacando-se como a principal causa de transmissão em países não endêmicos e áreas urbanas, seguido pela transmissão por transplante de órgãos, transmissão vertical de mãe para filho (via placenta) e, mais raramente, por acidentes laboratoriais (2,5). Já a transmissão oral, através da ingestão de alimentos contaminados com o parasita, é geralmente responsável por surtos de

infecção aguda, sobretudo em regiões rurais, onde a presença de várias espécies de animais reservatórios e de triatomíneos torna esse meio de transmissão possível (5,6).



**Figura 1.1:** Mapa com a representação da distribuição dos casos de doença de Chagas, baseado em estimativas oficiais de 2010-2013. Fonte: WHO (7)

Várias iniciativas para prevenção e controle da doença de Chagas foram realizadas no Brasil e em alguns outros países da América Latina, das quais se destacaram as iniciativas do Cone Sul em 1991, Andina em 1997, a da América Central e México em 1998 e a do Amazonas em 2004 (8). Nessas iniciativas as medidas de controle incluíram combate ao vetor, controle dos bancos de sangue, melhoramento das moradias em regiões endêmicas, instruções sobre os mecanismos da doença a populações de risco e, em alguns casos, a observação da migração de indivíduos de áreas rurais endêmicas. O forte impacto dessas medidas foi evidenciado pela diminuição do número de pessoas infectadas com essa doença de 30 milhões em 1990 para cerca de sete milhões atualmente (2,8–10).

Os sintomas, o diagnóstico e o tratamento variam de acordo com a fase de manifestação da doença de Chagas. A primeira forma de manifestação da doença é a fase aguda que ocorre logo após a inoculação do parasita e pode durar até oito semanas (11). A fase aguda é geralmente assintomática. Em casos envolvendo sintomas, eles se desenvolvem após 1 a 2 semanas posteriormente a transmissão vetorial ou após 20 a 40 dias em casos de transmissão por transfusão de sangue. Os

sintomas são na maioria das vezes inespecíficos e podem incluir febre, diarreia, reações alérgicas, vômito, inchaço dos linfonodos e fígado e, mais raramente, insuficiência cardíaca aguda ou meningoencefalites. Na transmissão vetorial, na região de inoculação do parasita pode ocorrer um processo inflamatório, denominado de chagoma de inoculação (edemas cutâneos) ou sinal de Romaña (edema na conjuntiva ocular) (6,12,13). Se não for tratada, a fase aguda pode levar a óbito em 2% a 8% dos casos, especialmente crianças (14).

A fase crônica da doença se inicia de 2 a 3 meses após a infecção inicial. Entre 60% a 70% dos infectados nunca desenvolvem sintomas, representando a forma indeterminada da fase crônica. Essa forma de manifestação é caracterizada pela presença de anticorpos contra o parasita, eletrocardiograma normal e falta de evidências radiográficas de anormalidades cardíacas ou gastrointestinais. Apesar disso, lesões inflamatórias têm sido detectadas em amostras de tecido do coração, músculo esquelético e do trato gastrointestinal de pacientes assintomáticos. Somente os outros 30 a 40% dos infectados, após um período de 10 a 40 anos da infecção inicial, desenvolverão sintomas clínicos – a forma determinada da fase crônica. Esta forma determinada pode ser dividida em três grupos: forma cardíaca, digestiva e cardiodigestiva, e estão diretamente relacionadas com a infecção pelo *T. cruzi* de células ganglionares, células do músculo liso, esquelético e cardíaco de órgãos como o coração, esôfago e cólon (11,13,15,16).

A forma cardíaca da doença de Chagas é a mais severa e frequente manifestação da fase crônica, podendo ou não apresentar insuficiência ventricular. Nessa forma da doença, são encontradas alterações características no eletrocardiograma. Além disso, a radiografia do tórax frequentemente indica cardiomegalia com ou sem congestão pulmonar. Os sintomas incluem insuficiência cardíaca, arritmias, angina e tromboembolismo sistêmico e pulmonar. A primeira manifestação clínica geralmente é a insuficiência biventricular, embora a arritmia cardíaca, incluindo a morte súbita, também aconteça (11). A morte súbita representa dois terços de todas as mortes da forma cardíaca da doença de Chagas, seguida pela insuficiência cardíaca refratária e pelo tromboembolismo. É interessante ressaltar que a insuficiência cardíaca causada pela doença de Chagas está associada com uma maior mortalidade que a insuficiência cardíaca por outras causas (13,17).

Já a forma digestiva da doença de Chagas se desenvolve em 10 a 15% dos pacientes crônicos e se caracteriza por alterações nas funções motoras, secretórias e

de absorção do esôfago e gastrointestinal. Além disso, pacientes com a forma digestiva podem desenvolver também dilatações do trato gastrointestinal, como o megaesôfago, megacólon e megaestômago. Essas manifestações são em grande maioria causadas por lesões do sistema nervoso entérico ocasionados por infecção pelo parasita. O megaesôfago causa principalmente disfagia com dor epigástrica, regurgitação, ptialismo e má nutrição em casos severos. O aumento de prevalência de câncer de esôfago em casos de pacientes com megaesôfago também é observado. O megacólon frequentemente afeta o segmento simoidal, o cólon descendente ou a combinação deles. Os sintomas do megacólon envolvem prisão de ventre, distensão abdominal e, em alguns casos, obstrução do intestino grosso. A maioria dos casos de megacólon estão associados a megaesôfago (15,18).

## **1.2 *Trypanosoma cruzi* – Diversidade intraespecífica e biologia celular**

O *T. cruzi*, agente etiológico da doença Chagas, é uma espécie caracterizada por possuir uma grande diversidade genética e fenotípica e pode ser dividida em seis unidades discretas de tipagem (DTUs, do inglês *discrete typing units*), classificadas de TcI a TcVI, que diferem entre si também quanto a sua infectividade, tropismo por tecidos e susceptibilidade a fármacos (19,20). Recentemente, uma sétima DTU, a Tcbat, foi proposta por Marcili e colaboradores (21). Esta DTU foi primeiramente isolada em morcegos e, apesar da distância genética relevante em relação as outras DTUs, está evolutivamente mais próxima da TcI (21). As DTUs podem ser descritas como unidades compostas por cepas/indivíduos geneticamente mais relacionados entre si do que com aqueles de outras unidades, sendo estas unidades identificáveis por marcadores genéticos, moleculares ou imunológicos comuns. Apesar disso, não há um consenso sobre a melhor metodologia para identificar as diferentes DTUs nessa espécie (20,22).

Estudos sobre a história evolutiva do *T. cruzi*, organismo diplóide, e as relações evolutivas entre as DTUs indicam que a variabilidade genética desse parasita é resultante da reprodução assexuada (evoluindo com o acúmulo de mutações discretas), das trocas genéticas entre os parasitas e da hibridização. Atualmente, sabe-se que as DTUs TcV e TcVI são híbridos da TcII e TcIII, enquanto que também há indícios de que TcIII e TcIV talvez sejam híbridos de TcI e TcII. Já TcI e TcII foram

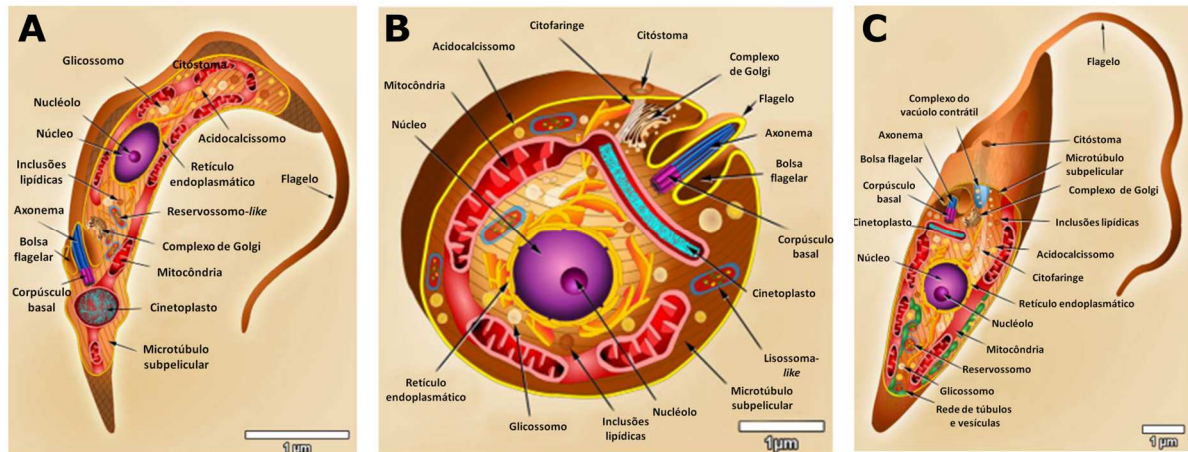
originados do mesmo ancestral comum há 1 a 3 milhões de anos, permanecendo até hoje linhagens puras (19,22).

Quanto à distribuição e ciclos de vida das DTUs, o TcI tem uma ampla distribuição na zona norte da América do Sul e está relacionada tanto a ciclos selváticos quanto domésticos. As TcII, V e VI estão relacionadas a ciclos domésticos e com grande incidência em pacientes na fase crônica da doença de Chagas na Bolívia e nos países do Cone Sul. Já a TcIII e IV estão relacionadas predominantemente a ciclos selváticos, a TcIII no Brasil e em países adjacentes e o TcIV principalmente na Venezuela. A TcBat foi encontrada recentemente em humanos, mas foram primeiramente identificados em morcegos. Diferentes DTUs podem coexistir em um mesmo vetor ou hospedeiro (19,22).

Por pertencer a ordem Kinetoplastida, o *T. cruzi* apresenta flagelo e cinetoplasto, uma estrutura formada por alta concentração de DNA extranuclear, o K-DNA, localizada dentro da matriz mitocondrial e próxima ao corpúsculo basal do flagelo. O k-DNA é feito de minicírculos e maxicírculos interligados. Enquanto os minicírculos codificam RNAs guia que editam e modificam os transcritos dos maxicírculos, os maxicírculos são tanto estrutural quanto funcionalmente semelhantes ao DNA mitocondrial de eucariotos superiores, codificando rRNAs e proteínas da cadeia respiratória (23).

Assim como outros eucariotos, o *T. cruzi* possui núcleo, retículo endoplasmático, complexo de Golgi e mitocôndria. Essa mitocôndria é única e se ramifica por toda a célula do parasita. Além disso, o *T. cruzi* apresenta organelas e estruturas características da família Trypanosomatidae, como os microtúbulos subpeliculares, a estrutura paraflagelar, os reservossomos, o cinetoplasto, o citóstoma, os glicossomos e os acidocalcisomos (24).

A fim de se adaptar aos diferentes microambientes impostos pelos seus hospedeiros mamíferos e invertebrados, o *T. cruzi* passa por alterações estruturais e metabólicas, viabilizando assim a infecção em cada hospedeiro. Dessa forma, ao longo do seu desenvolvimento o *T. cruzi* apresenta três formas evolutivas distintas: tripomastigota, amastigota e epimastigota (Figura 1.2). Essas formas evolutivas podem ser identificadas, por análise ao microscópio ótico, pela posição relativa do cinetoplasto em relação ao núcleo e a região de emergência do flagelo (25).



**Figura 1.2:** Representação do corte longitudinal das três formas evolutivas de *T. cruzi*. A. Forma tripomastigota. B. Forma amastigota. C. Forma epimastigota. Fonte: Adaptado de Teixeira *et al.*, 2012 (26).

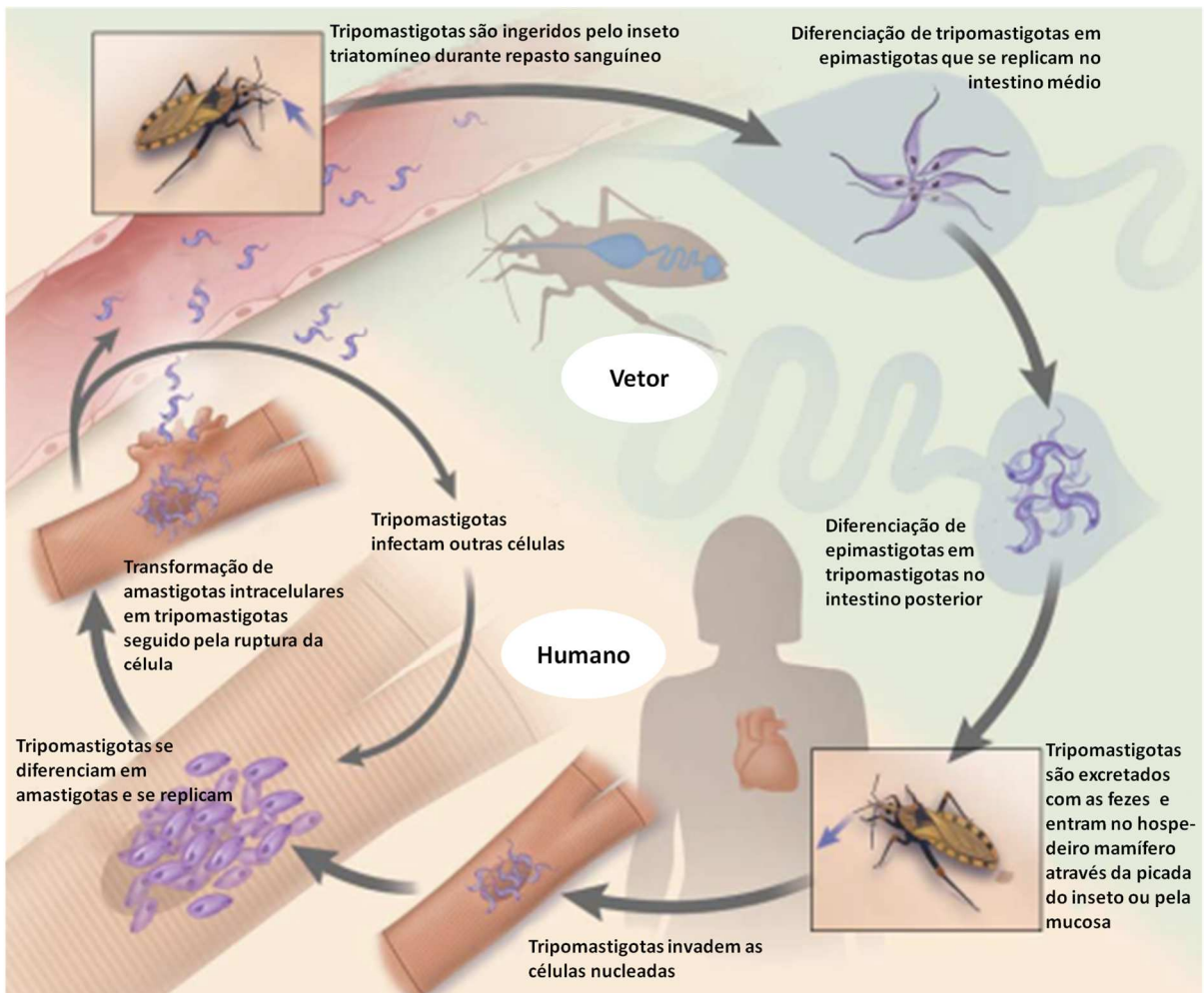
A forma tripomastigota não possui capacidade de se replicar e corresponde a forma infectiva extracelular do *T. cruzi* tanto no hospedeiro vertebrado (forma tripomastigota sanguínea) como no invertebrado (forma tripomastigota metacíclica). Apresentando de 12 a 20 µm de comprimento, esta forma evolutiva é alongada, possui um flagelo que emerge da bolsa flagelar, sendo a membrana ondulante normalmente estreita e situada em toda a extensão lateral do parasita. O tripomastigota apresenta também um cinetoplasto com alta densidade de DNA na região posterior da célula. A forma epimastigota se multiplica por divisão binária e é encontrada no trato digestivo do inseto triatomíneo e em glândulas anais de gambás. Ela apresenta de 20 a 40 µm de comprimento, o cinetoplasto se encontra junto ao núcleo, em uma posição anterior a ele. O flagelo é livre e bem desenvolvido, enquanto a membrana ondulante é curta. Já a forma amastigota é a forma intracelular do *T. cruzi*, se localizando em tecidos específicos do hospedeiro vertebrado e sendo capaz de se multiplicar por divisão binária longitudinal. Essa forma evolutiva apresenta 4 µm de diâmetro, possui formato arredondado, flagelo interno e ausência de membrana ondulante, se movimentando apenas por rotação. O cinetoplasto é um pouco menor que o núcleo e se localiza entre o flagelo e o núcleo. (25,27)

### 1.3 Ciclo de vida do *Trypanosoma cruzi*

No ciclo de vida do *T. cruzi*, três formas evolutivas são observadas: epimastigota, amastigota e tripomastigota (Figura 1.3). Além disso, esse ciclo de vida



envolve a infecção de dois hospedeiros – um hospedeiro mamífero, podendo este ser o homem, e um hospedeiro invertebrado, o inseto triatomíneo hematófago (barbeiro). Esse inseto se infecta, e inicia o ciclo, ao ingerir sangue de mamíferos contaminados com as formas tripomastigotas sanguíneas circulantes do *T. cruzi*. Uma vez no intestino médio, as formas tripomastigotas se transformam em epimastigotas não infectantes que se multiplicam por fissão binária e se diferenciam nas formas infectivas tripomastigotas metacíclicas, na porção final do intestino (13,25,28,29).



**Figura 1.3:** Ciclo de vida do *T. cruzi*. Fonte: Adaptado de Bern *et al.*, 2011 (30,31).

Durante o repasto sanguíneo em mamíferos, o triatomíneo infectado ingere uma quantidade significativa de sangue em seu sistema digestório que induz a eliminação de fezes e urina, contendo os tripomastigotas metacíclicos, próximos à região da picada. A partir disso, as formas tripomastigotas metacíclicas altamente móveis invadem a pele, a mucosa ou tecido lesionado desse hospedeiro vertebrado, infectando-o. Inicialmente as formas tripomastigotas metacíclicas infectam células de tecidos locais do hospedeiro pela formação de um vacúolo endocítico, denominado

vacúolo parasitóforo. Após sucessivas fusões com lisossomas recrutados pela célula hospedeira, o vacúolo parasitóforo se acidifica. Isso induz os parasitas a produzirem proteínas líticas que rompem a membrana do vacúolo e, conseqüentemente, liberam os tripomastigotas no citoplasma. Nesse novo microambiente, os tripomastigotas diminuem de tamanho e se diferenciam em amastigotas. Os parasitas nessa forma evolutiva se multiplicam por fissão binária por aproximadamente nove gerações para, então, se diferenciarem em tripomastigotas sanguíneos. Esses tripomastigotas sanguíneos rompem a membrana plasmática, caem na corrente sanguínea e/ou linfática e podem penetrar novas células nucleadas do hospedeiro vertebrado, especialmente células nervosas e células musculares esqueléticas, cardíacas e lisas de órgãos como esôfago, intestino e coração. Essas formas tripomastigotas sanguíneas, além de invadirem novas células, podem ser ingeridas pelos insetos triatomíneos durante o repasto sanguíneo, recomeçando o ciclo de transmissão (13,25,28,29).

#### **1.4 Tratamento da doença de Chagas**

O tratamento disponível para a doença de Chagas é realizado, desde a segunda metade do século XX, com benznidazol (1; Apêndice B- Fig. 7.2) e nifurtimox (2), compostos nitroheterocíclicos que contém um grupo nitro ligado a um anel furano e a um anel imidazol, respectivamente. Eles atuam como pro-fármacos e devem passar por uma ativação mediada por enzimas (nitrorredutases) a fim de apresentarem efeitos citotóxicos para o parasita. O mecanismo de ação desses fármacos não é totalmente compreendido, mas é sugerido que após a ativação desses compostos são formados radicais livres. Em presença de oxigênio, esses radicais se oxidam formando superóxido e regenerando o composto original (6,32). Além disso, a redução do nifurtimox por nitrorredutases quebra o anel furano resultando em uma nitrila de cadeia aberta e insaturada que é tão tóxico quanto o composto de origem, indicando que a atividade tripanocida do nifurtimox não envolve necessariamente o estresse oxidativo (6). Já no caso do benznidazol, estudos mostraram que ele se liga a tióis e proteínas contendo tióis, dessa forma inativando proteínas com cisteína no sítio ativo (6).

Tanto benznidazol quanto nifurtimox são eficazes no tratamento da fase aguda da doença de Chagas, já a eficácia desses medicamentos na fase crônica permanece

controversa. Apesar disso, a OMS, baseando-se em evidências de que esses compostos poderiam atenuar o desenvolvimento da doença na fase crônica, recomenda que pacientes crônicos também usem esses medicamentos (3,33). A fim de confirmar o benefício do benznidazol no tratamento de pacientes em fase crônica com cardiomiopatia, um teste clínico em fase três, denominado BENEFIT, foi iniciado em 2005. No entanto, esse teste, finalizado em meados de 2015, demonstrou que, apesar desse composto reduzir a detecção de parasitas circulantes, não foi capaz de reduzir a progressão do dano cardíaco (34).

Nifurtimox e benznidazol, embora sejam bem tolerados em crianças, são caracterizados por possuírem alta toxicidade em adultos, levando a uma série de graves efeitos colaterais. No caso do benznidazol, as reações adversas incluem hipersensibilidade (como erupções foto-sensíveis e dermatites), supressão da medula óssea (trombocitopenia, neutropenia, agranulocitose) e neuropatia periférica. Já no caso de nifurtimox, os efeitos adversos são ainda mais sérios, fazendo com que o tratamento de escolha na maioria dos países endêmicos, entre eles o Brasil, seja o benznidazol. Entre os efeitos adversos do nifurtimox, são descritos problemas gastrointestinais e, em alguns casos, perturbações do sistema nervoso central como polineurite, confusão ou convulsões focais ou generalizadas. Foi observado também casos de aberrações cromossômicas em crianças submetidas ao tratamento (6).

Outra desvantagem desses medicamentos inclui as diferenças de susceptibilidade entre as diferentes DTUs de *T. cruzi*, com indícios de que algumas DTUs apresentam resistência a compostos nitroheterocíclicos (6). Associado a esses problemas está também a necessidade da administração desses medicamentos por períodos relativamente longos – de 1 a 4 meses para nifurtimox e de 1 a 2 meses para benznidazol – e dessa forma, o alto custo do tratamento (6). Esse conjunto de desvantagens, com destaque a pouca eficiência na fase crônica da doença e os efeitos adversos desses medicamentos, torna de grande importância o desenvolvimento de novos fármacos mais seguros e efetivos no para o tratamento da doença de Chagas.

## **1.5 Estudos pré-clínicos e clínicos de candidatos a fármacos para a doença de Chagas**

A partir da descoberta de um fármaco (processo melhor discutido no item 1.81.8), sucessivas fases de avaliação são realizadas até a sua disponibilização no mercado (35). As etapas desse processo de avaliação incluem testes pré-clínico e clínicos e procedimentos regulatórios para aprovação do medicamento. O teste pré-clínico envolve pesquisas *in vitro* e *in vivo* com animais, em condições recomendadas pela regulamentação de uso de animais vigente no país da pesquisa, visando principalmente descobrir se o composto de interesse tem efeitos tóxicos. Além disso, essa etapa abrange o desenvolvimento de formulações do fármaco, geralmente em solução ou suspensão, que podem ser administradas aos animais nos testes de toxicidade; e solução, suspensão ou comprimidos ou cápsulas que podem ser administradas a humanos nos testes clínicos. O desenvolvimento de processos de síntese do composto, que irão possibilitar que ele seja manufaturado de modo reprodutível e em larga escala, também fazem parte da etapa pré-clínica (35).

Já a etapa de testes clínicos pode ser subdividida em quatro fases (fase 0 a III) e visa avaliar a eficácia e segurança do composto ao ser administrado em humanos para então o medicamento ser disponibilizado comercialmente. A fase 0 foi estabelecida pelo FDA (*Food and Drug Administration*) em 2006 e tem como objetivo ser um estudo exploratório de micro dosagens subterapêuticas dos compostos administradas em 10 a 15 humanos saudáveis. Essa dose é administrada uma única vez e é realizada para obtenção de dados preliminares de ADME (Absorção, Distribuição, Metabolismo e Excreção) do composto e para ranquear candidatos a fármacos com potenciais similares indicados pelos testes pré-clínicos. A fase I na maioria das vezes também é realizada com humanos sadios, nesse caso, de 20 a 100 indivíduos, a fim de avaliar a segurança, tolerabilidade (níveis de dosagem e efeitos colaterais), propriedades farmacocinéticas e efeitos farmacológicos do composto. Essa fase dura de alguns meses a um ano e, segundo o FDA, aproximadamente 70% dos candidatos a fármacos vão para a próxima fase. Na fase II são realizados testes em um número pequeno de pacientes (em torno de cem indivíduos). Esses testes visam medir a eficiência do composto contra a doença e duram de alguns meses a 2 anos. Aproximadamente 33% dos fármacos são bem-sucedidos na fase II. Na fase III, milhares de pacientes são tratados com o composto a fim de verificar a real eficácia,

efeitos adversos raros e possibilitar também a comparação do fármaco a outros fármacos similares já disponibilizados no mercado. A fase III tem duração de 2 a 6 anos e de 25 a 30% dos fármacos são aprovadas. Se aprovado nessas três fases, o novo fármaco pode ser disponibilizado comercialmente em um país após a aprovação do órgão regulamentador específico, como o FDA nos EUA e a ANVISA (Agência Nacional de Vigilância Sanitária) no Brasil. O monitoramento desse novo medicamento, após ser disponibilizado comercialmente, é denominado fase IV e tem como objetivo avaliar a eficácia, tolerabilidade e evidenciar a presença de qualquer efeito adverso raro e inesperado (35,36).

No contexto da doença de Chagas, por ser uma doença com urgência no desenvolvimento de novos medicamentos, várias tentativas recentes foram realizadas visando aprovação de novos fármacos no tratamento dessa doença. Entre essas tentativas, destacaram-se os estudos de prova de conceito recentemente finalizados com os compostos posaconazol (**3**; Apêndice B- Fig. 7.2) e ravuconazol (**4**) que são inibidores da síntese de ergosterol. Esses estudos em particular foram denominados STOP-CHAGAS (*Study of Oral Posaconazole in the Treatment of Asymptomatic Chronic Chagas Disease*), CHAGASAZOL e E1224. De maneira geral, estudos de prova de conceito têm como objetivo confirmar a viabilidade, aplicabilidade e eficácia de um medicamento a partir de um pequeno número de pacientes, podendo ser aplicado tanto em testes clínicos na fase I quanto na fase II (14).

O CHAGASAZOL foi um estudo destinado a avaliar a segurança e a eficácia do posaconazol em relação ao benznidazol (**1**) em um tratamento de 60 dias, visando a obtenção de resultados negativos contínuos da carga parasitária na fase crônica da doença de Chagas. Os resultados desse teste indicaram que apesar de o posaconazol apresentar uma atividade tripanocida em pacientes crônicos, isso não se sustentou ao longo dos tratamentos (37). Já o STOP-CHAGAS apresentava a finalidade de analisar a eficácia de posaconazol e posaconazol juntamente com benznidazol (**1**), comparado a placebo e o tratamento apenas com benznidazol em pacientes em fase indeterminada da doença de Chagas. Esse estudo demonstrou que o tratamento com benznidazol é superior ao tratamento com posaconazol ou posaconazol em combinação com benznidazol (38,39). O E1224, outro teste clínico em fase II recentemente finalizado, visava mostrar a eficácia de ravuconazol, um azol como o posaconazol, em comparação ao placebo e ao tratamento com benznidazol em pacientes em fase indeterminada. Esse estudo indicou que, apesar de efetivo em

eliminar o parasita, isso não se manteve 1 ano após o final do tratamento (14,40). Já em fase pré-clínica, o K777 (5), um inibidor de cisteína protease que inibe a cruzipaina, apresentou toxicidade em primatas não humanos e em cachorros (41). Infelizmente a falha de todos esses testes clínicos e pré-clínicos mostra que estamos apenas no início da busca por novos medicamentos mais eficazes no tratamento da doença de Chagas.

## **1.6 Alvos moleculares na doença de Chagas**

As limitações descritas para o tratamento disponível da doença de Chagas e as recentes falhas nos estudos pré-clínicos e clínicos com novos medicamentos tornam a busca por novos fármacos imprescindível. Nesse contexto, abordagens de descobertas de fármacos baseadas em alvo representam um processo favorável a identificação de compostos antitripanocidas mais seletivos e, portanto, com menores efeitos adversos ao hospedeiro. Para isso, alvos terapêuticos essenciais ao metabolismo do parasita podem ser utilizados. Além disso, o alvo ideal seria aquele sem macromoléculas estruturalmente semelhantes no organismo hospedeiro. Dessa forma, enzimas específicas ou análogas – proteínas com a mesma função enzimática, mas com estruturas tridimensionais distintas devido a diferentes origens evolutivas – do parasita em relação ao hospedeiro são alvos moleculares interessantes.

Motivado, entre outras razões, pela importância da identificação de alvos moleculares mais promissores, o nosso grupo de Bioinformática do Laboratório Genômica Funcional e Bioinformática desenvolveu uma ferramenta denominada AnEnPi (*Analogous Enzyme Pipeline*) (42). O AnEnPi possui como estratégia a identificação funcional de enzimas baseada em análises de similaridades entre sequências, permitindo a identificação de proteínas análogas, homólogas e específicas entre diferentes organismos, entre eles o *T. cruzi* e os seres humanos (42). Outra funcionalidade é o mapeamento das atividades enzimáticas no metabolismo do organismo de interesse, auxiliando, dessa forma, na identificação de enzimas essenciais.

Através dessa abordagem pós-genômica, o AnEnPi identificou a enzima ribose 5-fosfato isomerase como sendo uma enzima análoga de *T. cruzi* em relação a enzima de mesma atividade em humanos (43). Em ambos organismos, essa enzima pertence a via das pentoses fosfato, uma via importante na produção de precursores de

nucleotídeos e NADPH. Essa participação em uma via metabólica relevante de *T. cruzi* e a analogia predita em relação ao hospedeiro humano tornam a ribose 5-fosfato isomerase um alvo terapêutico promissor. Além disso, outro ponto chave é a importância (comprovada experimentalmente) dessa enzima na viabilidade celular de alguns tripanossomatídeos (44), representando indícios de que talvez seja essencial em *T. cruzi* também. Tendo isso em perspectiva, através de uma abordagem *in silico* focada no desenho de fármacos baseado em estrutura e em ligante, a ribose 5-fosfato isomerase foi a macromolécula alvo do nosso estudo visando a identificação de compostos com potencial de inibi-la e, conseqüentemente, com atividade antichagásica potencial.

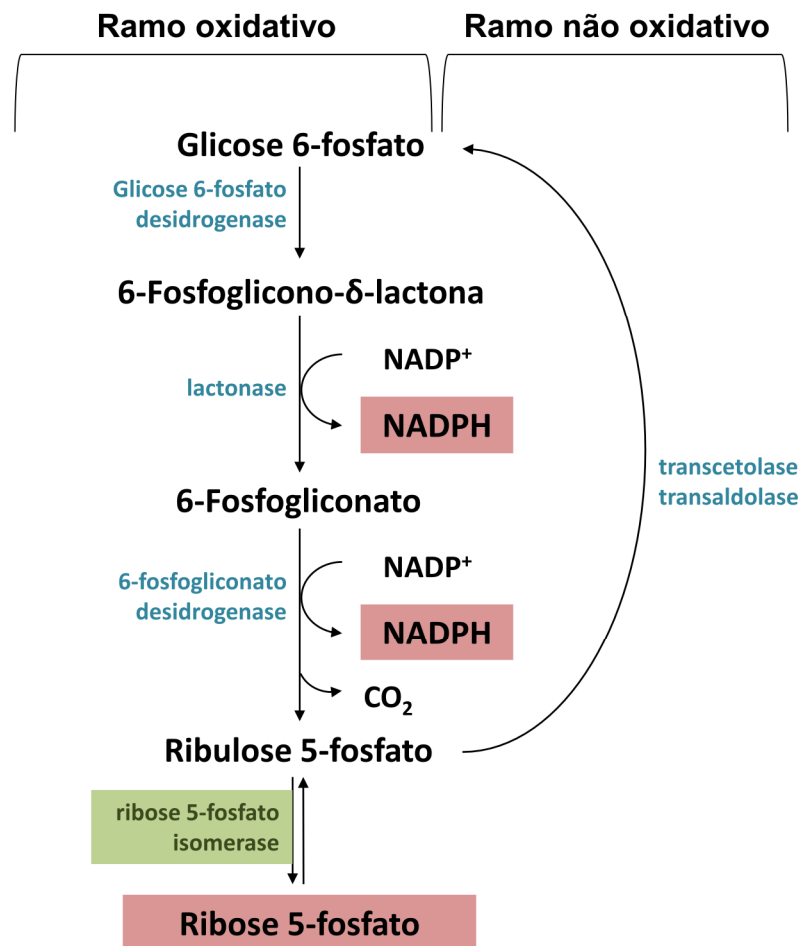
### **1.7 Via das pentoses fosfato**

A via das pentoses fosfato, a partir da metabolização de glicose 6-fosfato (6), permite a formação de dois compostos fundamentais para o metabolismo celular: NADPH e ribose 5-fosfato (7). O NADPH atua como doador de elétrons, fornecendo o poder redutor essencial para diversas vias biossintéticas. Além disso, o NADPH favorece a manutenção de um ambiente redutor nas células, prevenindo e revertendo os efeitos danosos de espécies reativas de oxigênio a moléculas sensíveis como proteínas e lipídeos. Já a ribose 5-fosfato é o precursor necessário para a síntese de nucleotídeos e ácidos nucleicos, importantes principalmente em células em divisão. Outro metabólito importante proveniente da via das pentoses fosfato é a eritrose 4-fosfato, um dos precursores na biossíntese de aminoácidos aromáticos e de várias vitaminas (45,46).

A via das pentoses fosfato é encontrada tanto em organismos procariontes quanto em eucariontes e as suas reações geralmente ocorrem apenas no citoplasma. Essa via metabólica pode ser dividida em duas fases: uma fase oxidativa e outra não oxidativa (Figura 1.4).

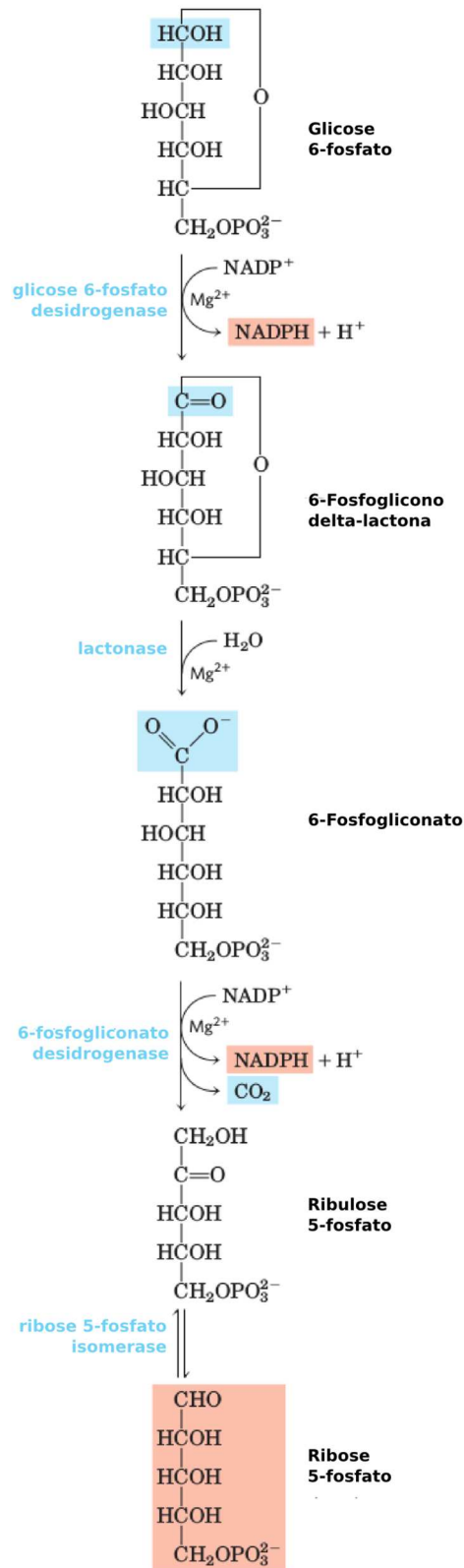
De maneira geral, na fase oxidativa (Figura 1.5), o NADPH é formado num processo envolvendo três reações irreversíveis que resultam na oxidação de glicose 6-fosfato (6) a ribose 5-fosfato. Essa fase oxidativa se inicia com a oxidação de glicose 6-fosfato pela glicose 6-fosfato desidrogenase, reduzindo NADP<sup>+</sup>. O produto dessa reação, o 6-fosfoglicono-delta-lactona (8), é hidrolisado pela lactonase a 6-fosfogliconato (9). Em seguida, a 6-fosfogliconato desidrogenase oxida e descarboxila

o 6-fosfogliconato, levando a formação final de ribulose 5-fosfato (10) e de mais uma molécula de NADPH. A ribulose 5-fosfato é convertida pela ribose 5-fosfato isomerase a ribose 5-fosfato, que será utilizada nas reações de síntese de nucleotídeos. O resultado dessa fase, portanto, é a produção de duas moléculas de NADPH para reações biossintéticas e uma molécula de ribose 5-fosfato para síntese de nucleotídeos e ácidos nucleicos (45). É interessante notar que a glicose 6-fosfato, que inicia a fase oxidativa, também pode ser metabolizada pela via glicolítica para a geração de piruvato (11). O que determina a entrada de glicose-6-fosfato na via glicolítica ou na via das pentoses-fosfato é basicamente as concentrações relativas de  $\text{NADP}^+$  e NADPH (45,47).



**Figura 1.4:** Esquema geral da via das pentoses fosfato. Em rosa estão representados os principais produtos dessa via metabólica, em azul estão representadas as enzimas envolvidas na via, com o destaque para a enzima ribose 5-fosfato isomerase em verde. Fonte: Adaptado de Nelson *et al.*, 2011 (45).





**Figura 1.5:** Reações do ramo oxidativo da via das pentoses fosfato. Em rosa estão representados os principais produtos dessa via metabólica, em azul estão representadas as enzimas envolvidas na via. Fonte: Adaptado de Nelson *et al.*, 2011 (45).

Já na fase não oxidativa, as pentoses fosfato são reconvertidas em glicose 6-fosfato ou a frutose 6-fosfato (**12**) e a gliceraldeído 3-fosfato (**13**), sendo as reações envolvidas nessa fase irreversíveis (45,47). Essa fase ocorre quando certas células, em dado momento, precisam apenas do poder redutor de NADPH. Nesses casos, é iniciada a fase não oxidativa, a fim de permitir que as pentoses fosfato produzidas na fase oxidativa sejam recicladas em glicose 6-fosfato ou a frutose-6-fosfato e gliceraldeído-3-fosfato. Para que essa reciclagem aconteça, primeiramente a ribulose 5-fosfato é epimerizada a xilulose 5-fosfato (**14**) e em seguida as enzimas transcetolase e a transaldolase catalisam a conversão reversível de seis pentoses-fosfato a cinco hexoses-fosfato. Ao final desses rearranjos, tanto a glicose-6-fosfato quanto a frutose-6-fosfato e o gliceraldeído-3-fosfato produzidos podem entrar na via glicolítica para produção de energia ou, no caso da glicose-6-fosfato, reiniciar a via das pentoses fosfato (45).

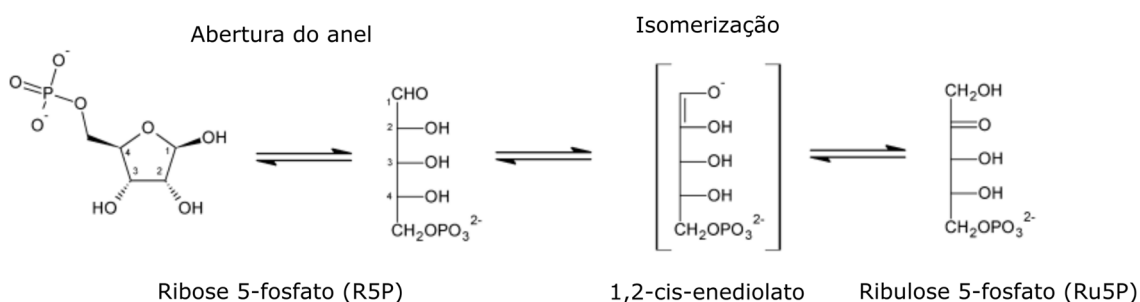
Em tripanossomatídeos, as enzimas dessa via tem sido largamente estudadas, o que tem revelado que algumas delas, como a ribose 5-fosfato isomerase, são essenciais a esses parasitas (44). Nesses organismos, além da clássica localização das enzimas no citoplasma, algumas enzimas se localizam nos glicossomos, organelas características desses parasitas e responsáveis por diversos processos metabólicos importantes como a glicólise e a via de salvamento de purinas (46,48,49). Todas as atividades enzimáticas da via das pentoses fosfato foram confirmadas nas três formas evolutivas de *T. cruzi* (50) e nas formas promastigotas de *Leishmania mexicana*, *Leishmania braziliensis*, *Leishmania donovani* e *Leishmania tropica* (46,51). Já em *Trypanosoma brucei*, demonstrou-se que todas as enzimas dessa via estão presentes na forma procíclica, enquanto que a transcetolase e ribulose 5-fosfato epimerase estão ausentes na forma sanguínea (46).

Nesses eucariotos em particular, a via das pentoses fosfato é de vital importância na proteção contra o estresse oxidativo causado pelos mamíferos hospedeiros, que produzem espécies reativas de oxigênio como um de seus mecanismos de defesa no combate a infecção. Nesse processo de proteção dos tripanossomatídeos, o NADPH é fundamental para que a tripanotiona redutase mantenha reduzida a tripanotiona, um reagente anti-oxidante característico desses parasitas (46,52). A tripanotiona possui função similar a glutatona presente em vários organismos, como, por exemplo, os mamíferos.

Em *T. cruzi*, a importância da via das pentoses fosfato na proteção contra o estresse oxidativo é evidenciada pelos níveis elevados de atividade das enzimas dessa via nas formas evolutivas presentes no hospedeiro vertebrado (tripomastigota e amastigota). Isso está relacionado a alta concentração de radicais livres produzidas pelas células do hospedeiro, principalmente macrófagos, para matar o parasita intracelular (50). Nessas situações, o parasita necessita de um grande fluxo de glicose através da via das pentoses fosfato a fim de produzir NADPH e dessa forma sobreviver ao estresse oxidativo. Por outro lado, a produção de ribose 5-fosfato é fundamental para as formas replicativas de *T. cruzi* (epimastigotas e amastigota). Com exceção da ribose 5-fosfato epimerase, as enzimas dessa via apresentam maior atividade na forma tripomastigota metacíclica (50,53).

### 1.7.1 Ribose 5-fosfato isomerase

A ribose 5-fosfato isomerase (Rpi, EC 5.3.1.6) catalisa a reação reversível entre a D-ribose-5-fosfato (R5P; **7**) e a D-ribulose-5-fosfato (Ru5P; **10**) desempenhando um papel central na via das pentoses fosfato. Essa catálise envolve duas etapas (Figura 1.6): uma de abertura do anel da R5P (forma fechada) e outra de isomerização em Ru5P por meio do intermediário de alta energia, o composto 1,2-cis-enediolato (54).



**Figura 1.6** Reação de isomerização da ribose 5-fosfato isomerase. Fonte: Adaptado de Capriles *et al.*, 2015 (55).

As enzimas Rpi podem ser divididas em duas famílias: tipo A e tipo B. Essas famílias, apesar de possuírem a mesma função enzimática, apresentam de maneira geral sequências e estruturas tridimensionais diferentes, provavelmente devido a origens evolutivas distintas, representando dessa forma um caso de analogia. As RpiAs são encontradas na maioria dos organismos eucariotos e em alguns

procariotos. Já as RpiBs existem principalmente em procariotos e em eucariotos inferiores, incluindo alguns tripanossomatídeos e outros protozoários. O tipo B também é encontrado em alguns fungos e em algumas plantas, embora existam diferenças marcantes entre as sequências destas e as sequências típicas de RpiBs. Alguns organismos, como *Escherichia coli*, *Shigella sp.* e *Enterobacter sp.* apresentam as duas famílias de enzimas. De qualquer forma, em genomas até então anotados, pelo menos um dos dois tipos de Rpis foi identificado. A Rpi de *T. cruzi* (TcRpiB) pertence à família das Rpis tipo B, enquanto a Rpi de humanos (HsRpiA) faz parte da família das Rpis tipo A (54).

A localização celular da RpiA em humanos e em outros eucariotos superiores acredita-se ser estritamente citoplasmática, assim como em outras enzimas da via das pentoses fostato (47). Já em tripanosomatídeos, a RpiB se localiza majoritariamente no citoplasma e minoritariamente nos glicossomos. Essa dupla localização da RpiB de tripanossomatídeos é demonstrada por estudos de permeabilização com digitonina e fracionamento celular por centrifugação diferencial em *T. cruzi* e estudos com digitonina e colocação por microscopia de fluorescência em *T. brucei* e em *L. infantum* (50,56). No entanto, a análise proteômica de glicossomos de *L. donovani* não conseguiu detectar a RpiB (57), podendo estar relacionado com a quantidade dessa enzima nos glicossomos (44).

Existem estruturas de Rpis resolvidas experimentalmente, tanto para RpiAs e RpiBs, disponíveis no *Protein Data Bank* (Tabela 1.1). Dois estados oligoméricos são encontrados em ambos os tipos de Rpis: o estado homodimérico e o homotetramérico. Em *E. coli* (58) a isoenzima RpiA (EcRpiA) se comporta como homodimérica, assim como a de *L. salivarius* (59), *V. vulnificus* (60), *Spinacia oleracea* (61) e *P. falciparum* (62). Já a forma homotetramérica de RpiA é encontrada em *P. horikoshii* (63) *M. jannaschii* (64) e *S. cerevisiae* (65). Diferentemente da sua RpiA dimérica, a RpiB em *E. coli* (EcRpiB) se encontra na forma homotetramérica (66). A forma tetramérica de RpiB também é encontrada em *C. immitis* (67) e, segundo dados cristalográficos, em *Clostridium thermocellum* (68). Homodímeros de RpiB estão presentes em *M. tuberculosis* (69) e em *T. cruzi*, embora, entre as estruturas de RpiB obtidas por cristalografia de raios-x para esse parasita, a forma tetramérica também tenha sido encontrada (54,70).

Tabela 1.1: Estruturas tridimensionais de Rpis resolvidas experimentalmente.

RpiA	
Organismo	PDB ID
<i>Pseudomonas aeruginosa</i>	4X84
<i>Thermus thermophilus</i>	1UJ4, 1UJ5, 1UJ6
<i>Escherichia coli</i>	1LKZ, 1KS2, 1O8B
<i>Streptococcus mutans</i>	3L7O
<i>Haemophilus influenzae</i>	1M0S
<i>Lactobacillus salivarius</i>	4GMK
<i>Neisseria gonorrhoeae</i>	5UF2
<i>Pyrococcus horikoshii</i>	1LK5, 1LK7
<i>Methanocaldococcus jannaschii</i>	3IXQ
<i>Vibrio vulnificus</i>	3ENQ, 3ENV, 3ENW
<i>Bartonella henselae</i>	3HHE
<i>Francisella tularensis</i>	4IO1
<i>Burkholderia thailandensis</i>	3UW1, 3U7J
<i>Plasmodium falciparum</i>	2F8M
<i>Toxoplasma gondii</i>	4NML
<i>Saccharomyces cerevisiae</i>	1XTZ

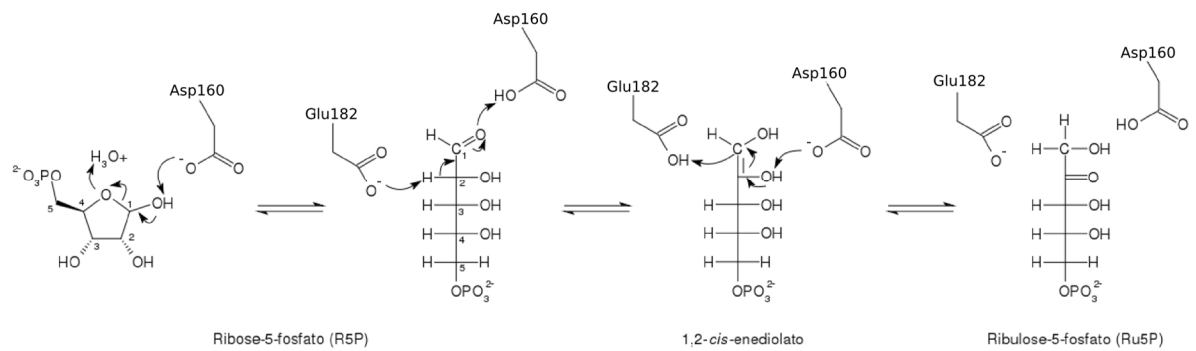
RpiB	
Organismo	PDB ID
<i>Pseudomonas aeruginosa</i>	4X84
<i>Thermus thermophilus</i>	1UJ4, 1UJ5, 1UJ6
<i>Escherichia coli</i>	1LKZ, 1KS2, 1O8B
<i>Streptococcus mutans</i>	3L7O
<i>Haemophilus influenzae</i>	1M0S
<i>Lactobacillus salivarius</i>	4GMK
<i>Neisseria gonorrhoeae</i>	5UF2
<i>Pyrococcus horikoshii</i>	1LK5, 1LK7
<i>Methanocaldococcus jannaschii</i>	3IXQ
<i>Vibrio vulnificus</i>	3ENQ, 3ENV, 3ENW
<i>Bartonella henselae</i>	3HHE
<i>Francisella tularensis</i>	4IO1
<i>Burkholderia thailandensis</i>	3UW1, 3U7J
<i>Bartonella henselae</i>	3HHE
<i>Francisella tularensis</i>	4IO1
<i>Burkholderia thailandensis</i>	3UW1, 3U7J
<i>Plasmodium falciparum</i>	2F8M
<i>Toxoplasma gondii</i>	4NML
<i>Saccharomyces cerevisiae</i>	1XTZ

O monômero das RpiAs em geral consiste de dois domínios  $\alpha/\beta$  de tamanhos distintos ligados por três a quatro fitas  $\beta$  curtas (55,62–65,74). Em HsRpiA, o domínio

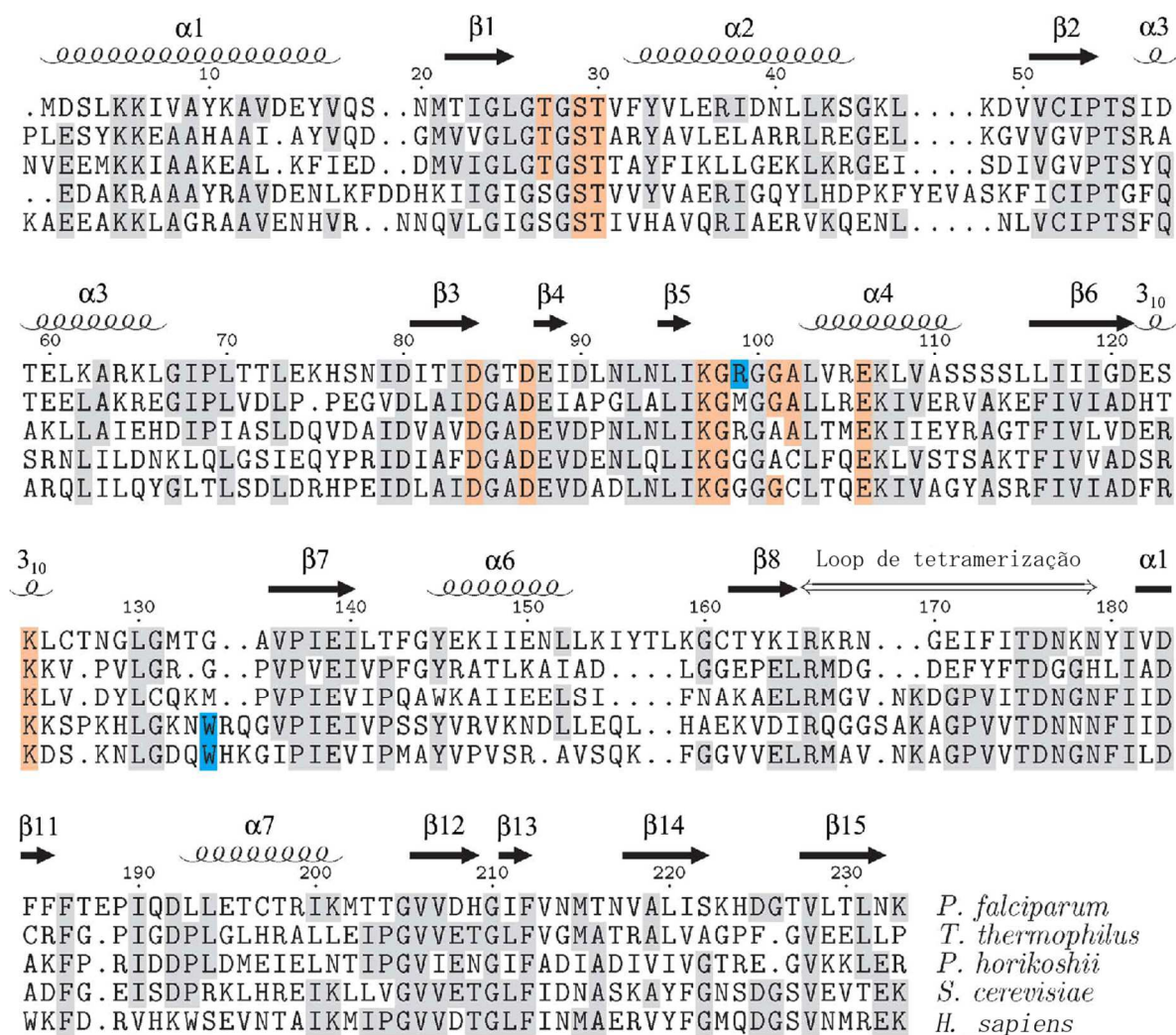
I é formado por seis folhas  $\beta$  paralelas e uma anti-paralela flanqueadas por quatro  $\alpha$ -hélices em uma extremidade e por duas na outra. O domínio II é menor e apresenta quatro folhas  $\beta$  anti-paralelas e duas  $\alpha$ -hélices (55). Essa arquitetura em ortólogos de RpiA é altamente conservada. A região de homodimerização (região de interação da cadeia A e B referente a HsRpiA e C e D) comum a todos os membros da família RpiA também é conservada, refletindo em um mesmo modo de dimerização em todas as RpiAs (65). RpiAs de archeas e eucariotos, com exceção de *P. falciparum* (62), apresentam um loop de tetramerização contendo de 16 a 17 resíduos (referente aos aminoácidos de 240 a 255 em HsRpiA) responsável pela estrutura tetramérica dessas RpiAs (55,62,63,65). Já em bactérias e em *P. falciparum*, o loop equivalente se encontra encurtado (possui 14 resíduos) o que poderia explicar a formação apenas de homodímeros nas RpiAs desses organismos e não de tetrâmeros (62) (Figura 1.8). A função do loop de tetramerização foi enfatizada por um estudo mostrando que a mutação de Arg189 (referente a Arg240 em HsRpiA) por Lys na RpiA de *S. cerevisiae* leva a perda de grande parte da atividade dessa enzima. Isso acontece pois a mutação provoca a dissociação do homotetrâmero em homodímeros e a destruição parcial da estrutura tridimensional do monômero (65,75).

Na estrutura das RpiAs, os resíduos Asp160, Lys173 e Glu182 (utilizando a numeração da HsRpiA como referência) representam os resíduos catalíticos (55,58,71). A função desses resíduos foi evidenciada através de estudos com RpiA de *S. oleracea* (SoRpiA) em que mutações em Asp87 e Lys100 (Asp160 e Lys173 em HsRpiA) diminuíram consideravelmente a constante catalítica da enzima (72). Já em *P. horikoshii*, a substituição de Glu107 (Glu182 de HsRpiA) para Gln causa a perda de atividade enzimática, enquanto a substituição do Asp85 (Asp160 de HsRpiA) por Asn causa diminuição dessa atividade a 2.5% (63). Quanto ao mecanismo de catálise (Figura 1.7), é proposto que a Asp160 juntamente com uma molécula de água atuem na abertura do anel da R5P (7) através da doação de um próton para O4 e a remoção do próton do O1, respectivamente, gerando a forma linear de D-ribose-5-fosfato. Em seguida, a etapa de isomerização envolvendo a formação do intermediário 1,2-cis-enediolato é possível através do Glu182 e novamente do Asp160. O Glu182 atua como uma base que remove um próton de C2 e então transfere-o para C1, enquanto a Lys173 ajuda na estabilização desse intermediário (55,58,71). Já a Asp160 devolve o próton para O1 e remove, em seguida, o próton do O2. Devido à grande conservação do sítio ativo das RpiAs, acredita-se que o mecanismo de catálise descrito seja

conservado nessas enzimas (65). O alinhamento entre as estruturas primárias de RpiAs de alguns dos organismos citados está representado na Figura 1.8.



**Figura 1.7:** Mecanismo de reação de HsRpiA. Figura adaptada de Baptista, 2011 (73).



**Figura 1.8:** Alinhamento múltiplo das estruturas primárias de RpiA. A sequência de HsRpiA está representada a partir do resíduo 77. Os resíduos conservados que contatam diretamente o 5-fosfato de ribose na estrutura de *T. thermophilus* (PDB ID 1UJ5) são indicados por sombreamento colorido. Fonte: Adaptado de Holmes *et al.*, 2006 (62).

No caso de RpiBs, cada subunidade da proteína possui geralmente cinco folhas  $\beta$  paralelas flanqueadas por três  $\alpha$ -hélices em um lado e duas no outro (54,66,68,69). Essa organização é denominada  $\alpha\beta\alpha$  sandwich fold e é comum em diversas proteínas, entre elas enzimas que ligam a nucleotídeos (66). Além disso, uma outra  $\alpha$ -hélice referente a região C-terminal se sobressai desse enovelado descrito e interage com o outro monômero do dímero/tetrâmero, estabilizando-o. Em cada dímero existem dois sítios ativos localizados entre as subunidades. Dessa forma, no caso de RpiBs tetramérica, são observados quatro sítios ativos.

Os resíduos do sítio ativo de RpiBs são altamente conservados. Apesar disso, a Cys69 (considerando a numeração de TcRpiB), um dos resíduos catalíticos, não está presente em RpiB de *M. tuberculosis* (MtRpiB), apresentando uma Gly em seu



lugar. Nesse caso, acredita-se que a função catalítica da Cys é substituída pelo Glu75 em MtRpiB (76). Outros aminoácidos catalíticos incluem a His102, Ser/Thr71 e em alguns organismos a His138 (na Figura 1.9 está representada o mecanismo de reação da TcRpiB). A His102 foi comprovada atuar em uma das etapas de abertura do anel da R5P (7), participando da doação de um próton para O4 (70). A etapa de remoção do próton do O1 gerando a forma linear de R5P é feita pela His138 (mesma numeração que TcRpiB), como foi demonstrado para RpiB de *L. donovani* (LdRpiB) (77). Já em TcRpiB essa função para a His138 foi descartada por estudos de mutação nesse resíduo (70), indicando que essa etapa de desprotonação do O1 do anel da R5P provavelmente é feita por uma molécula de água, como sugerido por Roos e colaboradores em 2008 (78) para a EcRpiB. A partir da forma linear de R5P, segue-se a etapa de isomerização a Ru5P pela Cys69 (Glu75 em MtRpiB) e pela Ser71, envolvendo a formação do intermediário 1,2-cis-enediolato. A Cys69 age como uma base que remove o próton do C2 e então transfere-o para C1, enquanto a Ser71 atua na transferência de um próton entre o O2 e O1 (54,70,76,77). Essa função da Cys69 em TcRpiB, LdRpiB e em RpiB de *Leishmania infantum* (LiRpiB) foi demonstrada por estudos de mutação sítio dirigida nos quais as RpiBs perderam sua capacidade de isomerização (44,70,77). O alinhamento entre as estruturas primárias de RpiBs de alguns dos organismos citados está representado na Figura 1.10.



Análises da cinética enzimática demonstraram que a EcRpiB é também uma aldose 6-fosfato isomerase, catalisando a isomerização reversível entre D-aldose 6-fosfato (All6P; **15**) e D-alulose 6-fosfato (Allu6P), ambos açúcares de seis carbonos (78). Ao contrário do que foi demonstrado para a EcRpiB, a TcRpiB e a MtRpiB não atuam na isomerização de D-aldose 6-fosfato; ao invés disso, esse açúcar atua na inibição de ambas as enzimas (54,78). Foi observado também que RpiBs de *Clostridium difficile*, *T. maritima*, *C. thermocellum*, *Streptococcus pneumoniae* e de *L. donovani* são capazes de catalisar a isomerização aldose-cetoses em vários outros açúcares (77,79–81). A RpiB de *Clostridium difficile* exibiu atividade apenas com aldoses como D-ribose (**17**), D-aldose (**18**), L-talose (**19**), L-lixose (**20**), D-gulose (**21**) e L-manose (**22**), enquanto a *T. maritima* apresentou atividade com aldoses como as formas D- e L de ribose, talose e aldose (79). A RpiB de *C. thermocellum*, possuiu atividade com substratos aldose como L-talose, D-ribose, D-aldose, L-aldose (**23**), L-ribose (**24**) e D-talose (**25**) (ordem decrescente de atividade). No caso de substratos cetoses a atividade seguiu a ordem: D-ribulose (**26**) > D-psicose (**27**) > L-tagatose (**28**) > L-ribulose(**29**) > L-psicose (**30**) > D-tagatose (**31**) (80). Em RpiB de *S. pneumoniae*, observou-se atividade particularmente alta com formas L- como L-lixose, L-xilulose (**32**), L-talose e L-tagatose (81). Já em LdRpiB, foi observada atividade com várias aldoses, na seguinte ordem, D-ribose > D-talose > D-aldose > D- arabinose (**33**) (77). Já no caso das RpiAs, foi demonstrado que a enzima de levedura está envolvida na biossíntese de piridoxina (**34**), uma forma de vitamina B6 e molécula precursora da coenzima piridoxal fosfato (**35**) (75). O piridoxal fosfato, por sua vez, desempenha importante função no metabolismo de aminoácidos.

#### **1.7.1.1 A essencialidade da ribose 5-fosfato isomerase na viabilidade celular**

As Rpis por terem um papel central na via das pentoses fosfato e, portanto, no metabolismo dos organismos, é essencial à sobrevivência de diversos deles. Essa essencialidade das Rpis A e B na viabilidade celular já foi evidenciada em várias espécies (44,56,82–85). Em leveduras, a deleção do gene de RpiA inviabiliza o crescimento desse organismo, enquanto que *E. coli* mutantes para a RpiA tornam-se incapazes de crescer em ausência de ribose (82,83). Nesse caso, a ribose (ou compostos contendo ribose) é sugerida como sendo um indutor da expressão de EcRpiB (84). Também em *E. coli*, duplos mutantes de RpiA-/RpiB- apresentaram crescimento prejudicado em todas as condições testadas (85). Em humanos foi

observado que paciente heterozigoto para alelo nulo do gene de HsRpiA e com mutação não sinônima de Ala61 para Val exibia HsRpiA com atividade enzimática parcial e manifestava leucoencefalopatia progressiva, neuropatia periférica e perfil metabólico alterado que incluía níveis aumentados de polióis como ribitol e arabitol (47,86). Somando-se a isso, a essencialidade demonstrada em outros organismos e o fato de nenhum caso de homozigose para alelos nulos de HsRpiA ter sido até hoje detectado levantam a hipótese de que indivíduos completamente nulos para HsRpiA sejam inviáveis (86).

Em tripanossomatídeos, estudos de Loureiro e colaboradores (56) mostraram que o silenciamento incompleto de RpiB de *T. brucei* (TbRpiB) na sua forma sanguínea reduzia o crescimento *in vitro* desses parasitas. Já camundongos infectados com os parasitas silenciados exibiam menor parasitemia e maior sobrevivência quando comparados ao grupo controle (56). Pelo fato de o silenciamento ter sido incompleto (com a presença de TbRpiB mesmo após a interferência), esse estudo foi capaz apenas de mostrar a importância dessa enzima para o parasita sem poder confirmar a sua essencialidade. Posteriormente, Faria e colaboradores (44) mostraram que mutantes nulos para um dos alelos de TbRpiB resultava em um parasita com fenótipo menos alterado do que aquele obtido com RNA de interferência, provavelmente devido a uma menor síntese da enzima naquele sistema. Todas as tentativas de gerar um mutante completamente nulo para TbRpiB falharam, indicando que TbRpiB é possivelmente essencial para a sobrevivência de *T. brucei*. Já em *L. infantum*, mutantes completamente nulos para LiRpiB só foram possíveis quando uma cópia epissomal do gene de RpiB era fornecida. Essa cópia epissomal foi retida mesmo com ausência de pressão seletiva por resistência ao fármaco, o que sugere que o gene de RpiB é essencial a sobrevivência do parasita nas formas amastigotas e promastigotas. Além disso, a incapacidade de gerar mutantes completamente nulos para LiRpiB em parasitas complementados com uma cópia epissomal de RpiB com mutação na Cys69, fundamental a atividade isomerase dessa enzima, sugere que a essencialidade descrita anteriormente é devido a essa função metabólica (44). Dessa forma, como a RpiB é essencial a sobrevivência de *L. infantum* e há indicações de que seja essencial também a *T. brucei*; numa perspectiva mais ampla essa enzima pode ser considerada um alvo terapêutico potencial em tripanossomatídeos de maneira geral (44).

### 1.7.1.2 Inibidores de ribose 5-fosfato isomerases

Entre os inibidores conhecidos de Rpis, o primeiro a ser descrito foi o iodoacetato (**37**, Apêndice B- Fig. 7.2), composto capaz de causar a perda total de atividade da EcRpiB sem afetar a RpiA (87). O mesmo efeito foi observado em MtRpiB e em TcRpiB, neste último caso iodoacetamida foi utilizado (69,70). Essa capacidade de inibição do iodoacetato e do iodoacetamida (**36**) está geralmente relacionada com o mecanismo de ação desses compostos que envolve a ligação covalente com resíduos de cisteína de uma proteína. Dessa forma, no caso da EcRpiB e TcRpiB, a interação covalente com a cisteína catalítica foi sugerida como causadora da perda de atividade enzimática (69,70). Já em MtRpiB, como a cisteína não está presente, a capacidade de inibição do iodoacetato supostamente está relacionada a reação com alguma das histidinas do sítio ativo (69). A Tabela 1.2 e Tabela 1.3 mostram todos os inibidores testados em RpiAs e RpiBs.

Baseado no intermediário de alta energia – 1,2-cis-enediolato - o ácido 4-fosfo-D-eritrônico, ou também chamado de e 4-fosfo-D-eritronato (4PEA, **38**), foi inicialmente demonstrado no ano de 1979 como um forte inibidor da SoRpiA ( $K_i$  0,0044 mM) (88). Além disso, outras moléculas se sobressaíram nesse estudo como a D-eritrose (**39**), que se liga fracamente a SoRpiA ( $K_i > 10$ mM), e a D-eritrose 4-fosfato (**40**), que é um bom inibidor dessa enzima ( $K_i$  0,32mM). Posteriormente, derivados de 4PEA (**38**) – ácido 4-fosfo-D-eritronohidroxâmico (4PEH, **41**), 4-fosfo-D-eritronohidrazida (4PEHz, **42**) e 4-fosfo-D-eritronamida (4PEAm, **43**) - representando novos análogos do intermediário da reação de isomerização da Rpi, foram testados em SoRpiA. Entre os compostos testados, além do próprio 4PEA, o 4PEH apresentou uma das melhores constantes de inibição ( $K_i$  de ~0,029 mM). Vale ressaltar que o 4PEA nesse estudo apresentou um valor diferente de  $K_i$  quando comparado ao estudo inicial ( $K_i$  de ~0,028 mM). Além disso, comparações do desempenho destas moléculas com outros análogos não fosforilados indicam a importância desse grupamento para se obter uma inibição efetiva no caso de RpiA (89). O grupamento fosfato, por si só, consegue inibir a SoRpiA ( $K_i$  7,9 mM), embora não fortemente (72). O 4-fosfonometil-D-eritronato (4PMEA, **44**), um análogo estável do 4PEA, foi sintetizado e testado em um outro estudo com SoRpiA, surgindo como outro potente inibidor dessa enzima ( $K_i$  0,074 mM) (90). Além disso, a arabinose 5-fosfato (**45**) também foi capaz de inibir a SoRpiA ( $K_i$  0.70 mM) e EcRpiA ( $K_i$  ~2,1 mM) (58,72). Alguns nucleotídeos, como o

AMP (**46**) e UMP (**47**), foram testados em Rpis e apresentaram capacidade de inibição dessa enzima (87,91).

Em um estudo de 2004, vários compostos foram selecionados baseados na similaridade com os substratos de Rpis, como a frutose 6-fosfato (**12**), D-eritrose (**39**), ácido D-4- eritrônico (**49**) e o ácido D-3-fosfoglicérico (**50**), e testados na MtRpiB. No entanto, nenhum desses compostos apresentou efeito na inibição da atividade mesmo em altas concentrações (até 10 mM). Nesse estudo, demonstrou-se também que o fosfato era um fraco inibidor ( $K_i$  130 mM)(69). Subsequentemente, ensaios com os inibidores já descritos para RpiAs como o 4PEA (**38**), 4PEAm (**43**), 4PEH (**41**) e o 4PEHz (**42**), compostos análogos ao intermediário de alta energia também das RpiBs, foram realizados em TcRpiB e MtRpiB. Desses ensaios, o 4PEH se destacou na inibição de ambas RpiBs com  $K_i$  de 0,057 mM e  $K_i$  de 1,2 mM para a MtRpiB e TcRpiB, respectivamente (70,76).

Considerando a capacidade de isomerização de All6P (**15**) a Allu6P (**16**) da enzima da EcRpiB, Mariano e colaboradores, em 2009 (92), sintetizaram análogos do intermediário de alta energia postulado para essa reação como o ácido 5-fosfo-D-ribonohidroxâmico (5PRH, **51**), 5-fosfo-D-ribonato (5PRA), 5-fosfo-D-ribonamida (5PRAm), N-(5-fosfo-D-ribonil)-hidrazina (5PRHz, **54**), N-(5-fosfo-D-ribonil)-metilamina (5PRMA), N-(5-fosfo-D-ribonil)-glicina (5PRGly, **56**) e o N-(5-fosfo-D-ribonil)-c-aminobutanoato (5PRGABA). Esses compostos foram testados em relação a atividade alose 6-fosfato isomerase de EcRpiB e a atividade Rpi de EcRpiB e de MtRpiB, os mais bem sucedidos para ambas atividades de EcRpiB foram 5PRH (**51**) ( $K_i \approx 0,09$  mM), 5PRAm ( $K_i \sim 0,07$  mM) e 5PRMA ( $K_i \sim 0,18$  mM), enquanto que para a atividade Rpi de MtRpiB, além dos compostos já descritos, 5PRA (**52**) e 5PRGly (**56**) também foram capazes de inibir essa enzima (92). Alguns desses mesmos análogos do intermediário da atividade alose 6-fosfato isomerase também foram testados em TcRpiB, mas nenhum deles inibiu significativamente essa enzima, apenas a All6P em si que apresentou uma inibição fraca ( $K_i$  de 15mM) (54)(70).

Tabela 1.2: Inibidores testados nas RpiBs

Atividade Ribose 5-fosfato isomerase				
Inibidor	K <sub>i</sub> (mM)	IC <sub>50</sub> (mM)	Organismo	Referência
Iodoacetato (37)	Inibição completa e irreversível			
Frutose 6-fosfato (12)	N.D.	-		
D-eritrose (39)	N.D.	-		
Ácido D-4-eritrônico (49)	N.D.	-		
Ácido D-3-fosfoglicérico (50)	N.D.	-		
Fosfato	130	-		
4PEH (41)	0,057	0,04		
4PEA (38)	1,7	6		
4PEAm (43)	N.D.	>30	<i>M. tuberculosis</i>	(69,76,92)
4PEHz (42)	N.D.	>35		
4PMEA (44)	~2	~12		
5PRH (51)	0,44 ± 0,02	0,26 ± 0,01		
5PRA (52)	0,031 ± 0,002	0,09 ± 0,001		
5PRAm (53)	0,22 ± 0,02	0,04 ± 0,01		
5PRHz (54)	1,0 ± 0,1	N.D.		
5PRMA (55)	0,19 ± 0,01	0,11 ± 0,02		
5PRGly (56)	0,57 ± 0,04	0,34 ± 0,01		
5PRGABA (57)	1,42 ± 0,03	N.D.		
5PRH (51)	0,17 ± 0,02	0,09 ± 0,02		
5PRA (52)	1,31 ± 0,06	N.D.		
5PRAm (53)	0,20 ± 0,02	0,07 ± 0,05		
5PRHz (54)	N.D.	N.D.		
5PRMA (55)	0,36 ± 0,02	0,18 ± 0,03	<i>E. coli</i>	(66,92)
5PRGly (56)	9 ± 1	N.D.		
5PRGABA (57)	2,0 ± 0,2	N.D.		
Iodoacetato (37)	Inibição completa e irreversível			
Aabinose 5-fosfato (45)	≥ 10	-		
4PEH (41)	1,2	0,7		
4PEA (38)	-	5mM	<i>T. cruzi</i>	(54,70)
4PEAm (43)	-	-		

4PEHz (42)	N.D.	> 10		
5PRH (51)	N.D.	> 10		
Iodoacetamida (36)	Inibição completa e irreversível			
All6P (15)	15mM	-		
5PRA (52)	N.I.S.	-		
5PRAm (53)	N.I.S.	-		
5PRMA (55)	N.I.S.	-		
5PRGly (56)	N.I.S.	-		
4PEH (41)	2,2	0,8	<i>T. brucei</i>	(56)
Atividade Alose 6-fosfato isomerase				
Inibidor	K <sub>i</sub> (mM)	IC <sub>50</sub> (mM)	Organismo	Referência
5PRH (51)	0,62 ± 0,05	0,43 ± 0,04		
5PRA (52)	1,33 ± 0,09	N.D.		
5PRAm (53)	1,8	N.D.		
5PRHz (54)	1,9 ± 0,3	N.D.	<i>E. coli</i>	(92)
5PRMA (55)	0,3	N.D.		
5PRGly (56)	N.I.	N.D.		
5PRGABA (57)	1,6 ± 0,3	N.D.		

N.D., não determinado; N.I., Não inibido; N.I.S., não inibido significativamente. Traços (-) representam que a informação não foi mencionada no estudo.



Tabela 1.3: Inibidores testados nas RpiAs

Atividade Ribose 6-fosfato isomerase				
Inibidor	Ki (mM)	IC50 (mM)	Organismo	Referência
Ácido D-malônico	1,3	-		
Ácido D-glutárico	> 10	-		
Ácido D-eritrônico (49)	0,32	-		
D-eritrose (48)	> 10	-		
Ácido D-3-fosfoglicérico (50)	1,3	-		
D-eritrose 4-fosfato (40)	0,67	-		
Ácido 5-fosfo-D-ribônico	0,119	-		
4PEA (1º estudo) (38)	0,0044	-		
4PEA (2º estudo) (38)	0,028 ± 0,005	0,010 ± 0,002	<i>S. oleracea</i>	(72,88–90)
4PEH (41)	0,029 ± 0,003	0,018 ± 0,003		
4PEHz (42)	1,8 ± 0,2	1,8 ± 0,2		
4PEAm (43)	2,5 ± 0,5	1,8 ± 0,2		
D-eritronato (49)	24 ± 2	-		
Ácido D-eritronohidroxâmico	32 ± 3	-		
4PMEA (44)	0,074	-		
Fosfato	7,9	-		
Xilulose 5-fosfato (14)	4	-		
Arabinose 5-fosfato (45)	0,7	-		
Arabinose 5-fosfato (45)	2,1 ± 0,9	-		
AMP (46)	-	-	<i>E. coli</i>	(58,87)
Glicose 6-fosfato (6)	-	-		

Traços (-) representam que a informação não foi mencionada no estudo.

### 1.8 Planejamento de novos fármacos auxiliado por computador

Duas estratégias têm sido utilizadas para a descoberta de novos candidatos a fármacos: a triagem fenotípica utilizando biblioteca de compostos e o desenho de racional de fármacos. Muitos fármacos antiparasitários disponíveis no mercado e a maioria dos candidatos a novos fármacos são tipicamente descobertos por meio de

triagens fenotípicas. Nesse contexto, as técnicas de triagem de alta vazão (*High Throughput Screening*; HTS) tem permitido a identificação de moléculas líderes através da realização de ensaios individuais com um grande número de compostos contra organismos inteiros (é importante notar que técnicas de HTS também podem ser utilizadas para a triagem de compostos contra macromoléculas alvo). Nas triagens fenotípicas, podem ser realizados inicialmente ensaios com sistemas de culturas *in vitro* seguido por testes *in vivo*. Pelo fato de ensaios desse tipo serem puramente fenotípicos, o alvo terapêutico e o mecanismo de ação do composto não são facilmente identificados, principalmente em casos em que o mesmo composto possui múltiplos alvos. Essa problemática torna a otimização de compostos líderes bastante desafiadora (93–95).

Já a identificação racional de novos fármacos envolve geralmente três etapas: (i) a identificação/busca de um alvo terapêutico; (ii) validação do alvo e (iii) desenho ou identificação de compostos capazes de interagir com o alvo e, dessa forma, promover o efeito biológico desejado (94). Alvos terapêuticos podem pertencer as diversas classes de macromoléculas biológicas como proteínas, lipídeos, ácidos nucléicos e carboidratos. No entanto, grande parte dos fármacos aprovados apresentam proteínas (sobretudo enzimas) como alvos terapêuticos, principalmente devido a problemas de toxicidade, especificidade e a inabilidade de se obter compostos capazes de inibir as outras classes de macromoléculas (96,97). Em doenças causadas por patógenos; para uma macromolécula, como uma enzima, ser considerada um potencial alvo terapêutico, ela deve ser preferencialmente exclusiva do parasita ou ter um baixo grau de similaridade estrutural com a macromolécula homóloga no organismo hospedeiro. Ou seja, esses alvos necessitam ser específicos do patógeno, homólogos distantes ou análogos entre o microrganismo patogênico em relação aos seus hospedeiros. Além disso, esses alvos devem ser essenciais para a sobrevivência do parasita de forma que, ao serem identificados inibidores contra esses alvos, haja a inativação da função destes, prejudicando assim o parasita, sem afetar o hospedeiro (94).

Para a identificação de alvos essenciais, abordagens *in vitro* como a superexpressão, silenciamento ou o *knockout* de genes podem ser utilizadas. Da mesma forma, ferramentas de bioinformática, que permitam fazer a análise e mineração de dados genômicos, proteômicos e transcriptômicos, são frequentemente empregadas para a identificação de alvos potenciais (98). A análise de sequências

genômicas do *T. cruzi*, em particular, permite a identificação dos genes ali codificados, suas proteínas correspondentes e suas possíveis funções a partir de análises comparativas de sequências entre diversas espécies. Isso possibilita também a identificação de genes codificantes de proteínas homólogas, análogas e específicas de relevância funcional em espécies próximas e/ou distantes, como um parasita e o seu hospedeiro, permitindo assim a inferência de potenciais alvos terapêuticos.

A validação desses alvos terapêuticos potenciais inclui a aplicação de métodos de eliminação seletiva da função de cada molécula alvo em particular e a subsequente observação da consequência desse processo em uma via bioquímica isolada ou em um organismo inteiro. Entre os métodos experimentais que permitem essa eliminação seletiva para a validação de proteínas alvos estão o silenciamento e o *knockout* de genes. Além disso, anticorpos específicos podem ser desenvolvidos visando o bloqueio da função do potencial alvo terapêutico. Alternativamente, um composto que modifique muito especificamente esse alvo pode servir para definir a função dele e validá-lo também, mesmo que esse composto por si só não seja adequado como um fármaco (35).

A terceira etapa do desenho racional de fármacos é a identificação de compostos capazes de interagir com a molécula alvo. Esses compostos podem ser identificados de diversas formas que vão de metodologias puramente experimentais a metodologias computacionais, conhecidas como o desenho de fármacos auxiliado por computador (*Computer-Aided Drug Design*; CADD). As abordagens baseadas em CADD, que serão o foco dessa tese, podem ser usadas uma vez que se tenha conhecimento sobre o sítio de ligação da molécula alvo ou sobre a estrutura de ligantes capazes de interagir com essa molécula (35). As abordagens CADD são amplamente utilizadas para triagem virtual de bibliotecas de compostos a fim de gerar um conjunto de compostos potencialmente ativos em relação a macromolécula alvo ou para otimizar compostos líderes, ou seja, transformar compostos biologicamente ativos em medicamentos adequados, melhorando as suas propriedades físico-químicas, farmacêuticas e farmacocinéticas (99). A principal vantagem desse processo de triagem *in silico* é a redução do número de compostos que precisam ser testados experimentalmente a fim de identificar compostos ativos e a consequente redução do tempo e dos custos envolvidos na descoberta de novos fármacos (100).

O CADD envolve diferentes metodologias que podem ser divididas principalmente em duas classes: desenho de fármacos baseado em ligante e desenho de fármacos baseado em estrutura do alvo terapêutico.

### 1.8.1 Desenho de fármacos baseado em ligante (DFBL)

O desenho de fármacos baseado em ligante envolve metodologias que precisam de conhecimento prévio sobre as moléculas que se ligam em uma macromolécula alvo a fim de guiar a identificação de compostos mais potentes. Esse conhecimento prévio pode incluir informações como a afinidade de ligação, estrutura química, propriedades físico-químicas dessas moléculas, entre outras (101). O DFBL se torna particularmente importante quando a estrutura do alvo terapêutico de interesse é desconhecida e a predição dessa estrutura não é possível por modelagem comparativa ou por abordagens *ab initio* (100). Em casos em que a estrutura da macromolécula alvo está disponível, o uso combinado do DFBL e do desenho de fármacos baseado em estrutura pode ser bastante vantajoso e eficiente uma vez que uma classe de metodologia compensa as limitações da outra. Além disso, as vantagens de ambas classes de metodologias são reforçadas (100).

As abordagens mais comuns do DFBL são relação quantitativa estrutura-atividade independente de receptor (*quantitative structure-activity relationship*; QSAR), a busca por similaridade molecular e a modelagem de farmacóforo baseada em ligante (100). O QSAR envolve a utilização de compostos muito e pouco potentes em relação à macromolécula alvo, seguido da produção de descritores 2D e 3D das propriedades físico-químicas e estruturais dessas moléculas, como peso molecular, volume, ligações com rotação, distâncias interatômicas, tipos de átomos, distribuição atômica, aromaticidade, propriedades de solvatação e etc. O  $IC_{50}$  (concentração necessária para inibir 50% da atividade de uma enzima) e a  $K_i$  (constante de inibição) são as principais medidas utilizadas em QSAR para a quantificação da atividade de compostos. A partir disso, o objetivo final é gerar um modelo capaz de descrever a correlação matemática entre estrutura/propriedade de um conjunto de moléculas e a sua atividade em relação a macromolécula alvo de forma que esse modelo possa ser usado para avaliar a atividade de novos compostos (100–102).

Nos métodos de busca por similaridade molecular, geralmente o *fingerprint* (representação molecular que permite comparação estrutural entre moléculas pequenas) de ligantes conhecidos da macromolécula alvo é usado na triagem de

bibliotecas de compostos visando encontrar moléculas similares. Esse método se baseia na teoria de que moléculas estruturalmente similares tendem a ter propriedades de ligação similares (100). Além disso, métodos baseados em *fingerprint* consideram todas as partes da molécula igualmente, sem diferenciar partes consideradas como importantes na atividade biológica (102).

Já o método de geração de farmacóforo baseado em ligante será descrito detalhadamente na seção 1.8.2.1, juntamente com modelagem de farmacóforo baseada em estrutura.

### **1.8.2 Desenho de fármacos baseado em estrutura (DFBE)**

O desenho de fármacos baseado em estrutura requer a existência da estrutura tridimensional da macromolécula alvo previamente determinada por metodologias experimentais, como cristalografia de raios-x, RMN (ressonância magnética nuclear) e criomicroscopia eletrônica ou por abordagens computacionais, como modelagem comparativa e *ab initio*. As abordagens de DFBE envolvem inicialmente a preparação e análise das estruturas de moléculas biológicas quanto a topologia do sítio de ligação, propriedade físico-químicas dos resíduos do sítio ativo, estados de protonação, entre outras. A partir disso, compostos são desenhados ou identificados baseados no conhecimento da estrutura alvo de forma a conterem características físico-químicas necessárias para interagir com a macromolécula alvo (103,104). Para este fim, o DFBE envolve metodologias, como a modelagem de farmacóforo baseado em estrutura, atracamento molecular (ou *docking* molecular) e a dinâmica molecular, descritas, respectivamente, nas seções 1.8.2.1, 1.8.2.2 e 1.8.2.3.

#### **1.8.2.1 Modelagem de farmacóforo**

A palavra farmacóforo (*pharmacophore* em inglês) é empregada na literatura desde o início do século XX. Inicialmente, essa palavra era utilizada para descrever grupos químicos específicos dentro de uma molécula, responsáveis pela sua atividade biológica (105). Atualmente, a definição, segundo a IUPAC (*International Union of Pure and Applied Chemistry*), descreve farmacóforo como sendo um conjunto de características estéricas e eletrônicas necessárias para garantir interações moleculares efetivas com um alvo biológico específico e, dessa forma, desencadear ou inibir uma resposta biológica. Ainda de acordo com a IUPAC, um farmacóforo não representa uma associação real com grupos químicos ou um composto em particular,

mas um conceito abstrato que ilustra as capacidades comuns de interação molecular de um conjunto de compostos em relação ao alvo biológico específico (106).

Nota-se que o conceito básico de farmacóforo mudou ao longo do tempo e a ênfase antes dada a grupos químicos como sendo representativos de interações essenciais expandiu para uma descrição química mais abrangente que frequentemente é denominada ponto farmacofórico (*pharmacophore feature* em inglês). Um farmacóforo compreende vários pontos farmacofóricos. Tipos comuns de pontos farmacofóricos incluem aceptores e doadores de ligação de hidrogênio, cátions, ânions, anéis aromáticos e centros hidrofóbicos. Além disso, alguns programas de modelagem de farmacóforo permitem adição de restrições estéricas denominados volumes de exclusão (107).

Com a introdução do CADD na química medicinal, diversas ferramentas de modelagem de farmacóforo surgiram representando novas perspectivas e aplicações do conceito de farmacóforo. A partir desses programas, um modelo de farmacóforo pode tanto ser gerado baseado na estrutura do ligante quanto baseado na estrutura da macromolécula alvo. A abordagem escolhida vai depender da disponibilidade de dados sobre os ligantes e as estruturas da proteína alvo em complexo com ligante ou sem ligante.

O farmacóforo baseado na estrutura do ligante é particularmente interessante na ausência de estrutura da macromolécula alvo, sendo necessário o conhecimento de ligantes ativos. Os modelos baseados em ligante são obtidos por sobreposição desses compostos seguido da extração de pontos farmacofóricos comuns entre eles e, teoricamente, essenciais a interação com a macromolécula alvo. Os programas que seguem esse tipo de abordagem geralmente realizam a construção do modelo de farmacóforo em duas etapas: (i) mapeamento do espaço conformacional de cada ligante, visando representar a flexibilidade conformacional deles e, (ii) sobreposição dos ligantes e determinação dos pontos farmacofóricos comuns para construir os modelos. Dessa forma, a representação da flexibilidade conformacional e a sobreposição são as etapas chaves dessa abordagem e os programas existentes diferem principalmente nos algoritmos que tratam dessas etapas. Os algoritmos que lidam com a flexibilidade conformacional devem ser principalmente capazes de gerar eficientemente todas as conformações possíveis que os ligantes assumem quando interagem com macromoléculas. Entre os algoritmos mais conhecidos de amostragem

do espaço conformacional de pequenas moléculas aplicados a programas de farmacóforo estão os Monte Carlo<sup>1</sup> e os algoritmos genéticos<sup>2</sup>.

Já os algoritmos de sobreposição podem ser classificados em duas categorias: baseados em ponto e baseados em propriedade. Na abordagem baseada em ponto, o ponto em si pode ser diferenciado em átomos, fragmentos ou características químicas e os pares de pontos são sobrepostos por minimização de distâncias. Já algoritmos baseados em propriedade usam descritores moleculares, representados geralmente por um conjunto de funções gaussianas, para gerar as sobreposições. É importante notar que, quando conjuntos de dados diferentes são utilizados para gerar um modelo de farmacóforo, diferentes modelos podem ser gerados para representar as interações essenciais com um alvo específico. Os tipos de ligantes, o tamanho do conjunto de dado e a diversidade química dos ligantes influenciam na qualidade do modelo e, portanto, na representatividade das interações essenciais (108,109).

No farmacóforo baseado na estrutura da macromolécula alvo, utiliza-se a estrutura tridimensional do alvo molecular complexado ou não com o ligante. De maneira geral, esse tipo de modelagem de farmacóforo envolve a análise de características químicas complementares do sítio ativo e do ligante, de forma que o modelo a ser gerado seja uma representação das interações intermoleculares observadas. Em casos em que o alvo molecular não está complexado a um ligante, apenas as informações físico-químicas dos aminoácidos do sítio ativo são utilizadas para a geração do modelo de farmacóforo. Esses casos podem ser problemáticos uma vez que a maioria dos programas não tem classificação de prioridade para auxiliar na seleção dos pontos farmacofóricos que constituirão o modelo de farmacóforo. Outra problemática comum aos modelos baseados na estrutura da macromolécula alvo é a

---

<sup>1</sup> O algoritmo de Monte Carlo é realizado visando a busca conformacional do ligante de interesse. De maneira geral, esse método envolve a geração de uma configuração inicial aleatória do ligante. A partir desta configuração inicial, uma nova conformação é gerada. Se esta configuração possuir menor energia que a inicial, a nova conformação é aceita e usada como referência para um próximo passo de iteração. Caso a energia seja maior, um critério é utilizado para decidir se a nova conformação será aceita ou não. As iterações são feitas até que se alcance o número desejado de conformações (52,61).

<sup>2</sup> Nos algoritmos genéticos utiliza-se a teoria da evolução e seleção natural de Darwin. Nesse método se trabalha com uma população de "indivíduos" representados por cromossomos (vetores contendo os parâmetros estruturais da conformação do ligante). A cada geração, novos cromossomos são formados através da troca de genes entre dois cromossomos progenitores (recombinação) e de mudanças aleatórias nos valores dos genes (mutação). Os cromossomos progenitores correspondem aqueles mais bem adaptados (isto é, com os menores valores de energia) da geração em questão. Este processo é recursivamente executado, até que um critério de parada seja encontrado, de forma que a população evolua para conformações (cromossomos) correspondendo a energia mínima global (52,61).

presença de muitos (mais que 7) pontos farmacofóricos, o que impossibilita, por exemplo, uma etapa posterior de triagem virtual. Dessa forma, é necessário selecionar um número limitado de pontos farmacofóricos para compor o modelo, o que não é uma tarefa fácil devido a já citada falta de prioridade explícita. Outra desvantagem é a falta de correlação do modelo com o QSAR, já que o modelo geralmente é derivado de um único complexo proteína-ligante. A utilização de mais de um complexo proteína-ligante e a seleção de pontos farmacofóricos em comum pode contribuir para uma maior representatividade do modelo final sobre as interações essenciais (108). Por outro lado, as vantagens de se utilizar abordagens baseadas em farmacóforo em comparação a outras metodologias, como o atracamento, são a "simplicidade" no processo de geração de modelos e a rapidez e acurácia em triar virtualmente bibliotecas contendo milhões de compostos (110).

De forma geral, ambas as abordagens de farmacóforo podem ser utilizadas para triagem virtual de bibliotecas de ligantes, desenho *de novo* de compostos e aplicações relacionadas a otimização de compostos líderes e desenho de fármacos multialvo. Na triagem virtual baseada em farmacóforo, em particular, objetiva-se identificar ligantes com características farmacofóricas similares ao modelo e, portanto, com alta probabilidade de serem biologicamente ativos. Alguns novos ligantes encontrados podem ter um arcabouço molecular semelhante aos ligantes conhecidos ou um arcabouço molecular completamente novo. Essa abordagem é uma vantagem sobre triagem virtual baseada em atracamento pois reduz problemas provenientes, por exemplo, de consideração inadequada de flexibilidade da proteína alvo, pois introduz uma tolerância de raio de cada ponto farmacofórico (107,108).

### **1.8.2.2 Atracamento molecular**

Uma das aplicações mais práticas do conceito de reconhecimento molecular é a estratégia computacional de atracamento molecular, seja de uma pequena molécula ou mesmo de proteínas. O atracamento molecular permite identificar ligantes, ou em alguns casos macromoléculas, que sejam potencialmente capazes de interagir e se encaixar – tanto geometricamente quanto energeticamente – no sítio de ligação de uma macromolécula receptora, geralmente uma proteína. Essa metodologia consegue avaliar complexos energeticamente mais favoráveis entre um receptor e uma outra molécula, baseando-se muitas vezes em estimativas de energia livre. Além disso, é possível também avaliar a conformação e a orientação do ligante, em



condições de equilíbrio, em uma região do receptor. Por auxiliar nessa compreensão de como um ligante se liga à macromolécula alvo, o atracamento molecular tem se tornado uma ferramenta muito utilizada na descoberta de novos fármacos (103,111). Os programas de atracamento molecular podem ser descritos como a combinação de dois componentes: (i) algoritmo de busca; (ii) função de pontuação ou avaliação.

Os algoritmos de busca são responsáveis pela geração dos possíveis modos de ligação entre a proteína e o ligante e geram inúmeras poses/conformações da molécula ligante no sítio de ligação. Embora durante o reconhecimento molecular proteína-ligante ocorram mudanças conformacionais (teoria do encaixe-induzido), em um programa de atracamento geralmente apenas a flexibilidade da estrutura do ligante é considerada (atracamento molecular semi-flexível). Nesse caso, são considerados os graus de liberdade rotacional, translacional e conformacional do ligante. Apesar da flexibilidade molecular do receptor ser uma característica importante para a representação da realidade em um programa de atracamento, a consideração do grande número de graus de liberdade da molécula receptora torna o processo muito custoso computacionalmente. Para contornar isso, alguns programas de atracamento permitem considerar a proteína parcialmente flexível além da consideração de todos os graus de liberdade do ligante. Outro método para tratar a flexibilidade da proteína é utilização de conjuntos de conformações proteicas (ao invés de uma única proteína) como alvo no atracamento – *ensemble docking* (111,112).

Os principais métodos de busca utilizados no atracamento molecular podem ser classificados basicamente em métodos de busca sistemática, métodos de simulação, também chamados de determinísticos, e métodos de busca estocástica.

No método de busca sistemática cada grau de liberdade apresenta um conjunto de valores e todos os graus de liberdade da molécula são explorados combinatoriamente durante a busca. O principal exemplo de busca sistemática é o algoritmo de construção incremental ou baseado em fragmento. No algoritmo de construção incremental, o ligante é dividido em fragmentos, levando em consideração regiões rígidas e flexíveis. Esses fragmentos são, então, ancorados/docados no sítio ativo do receptor de maneira incremental a partir de um fragmento base (fragmento âncora) até a reconstrução do ligante. A principal vantagem desse método é que a busca conformacional pode ser restrita somente as conformações que se ajustam no sítio de interesse (111,112).

Já o método de busca baseado em simulação ou determinístico envolve métodos de dinâmica molecular ou métodos de minimização de energia. As simulações de dinâmica molecular permitem explorar a influência de solvente explícito e os graus de liberdade do complexo proteína-ligante. Apesar disso, muitas vezes essas simulações são incapazes de atravessar barreiras de alta energia dentro de períodos de simulação viáveis e, portanto, só podem acomodar ligantes nos mínimos locais da superfície potencial. Para contornar esse problema, frequentemente é feita a tentativa de simular diferentes partes de um sistema de proteína-ligante a diferentes temperaturas ou dar início aos cálculos de dinâmica molecular a partir de diferentes posições do ligante. No caso da utilização de minimização de energia, visa-se encontrar um conjunto de coordenadas que minimizem a energia potencial do sistema de interesse. Esse método geralmente é usado em complementação a outros métodos, como o Monte Carlo, mas raramente é utilizado como única abordagem do algoritmo de busca já que apenas o mínimo local pode ser alcançado ele (111,112).

Em métodos estocásticos, a busca é regida por movimentos aleatórios, ocasionando em diferentes resultados finais para um mesmo estado inicial fornecido de entrada. Exemplos desses métodos são o Monte Carlo e os algoritmos genéticos (seção 1.8.2.1) (103,112).

Após a geração dos modos de ligação de ligantes contra uma proteína receptora, o programa de atracamento aplica uma função de pontuação ou avaliação. Uma boa função de pontuação é capaz de prever, de forma aproximada, a energia de ligação dentre as conformações geradas pelo algoritmo de busca para um ligante. Em ensaios de triagem virtual, a função de pontuação deve ser capaz também de ranquear corretamente um conjunto de ligantes de forma a refletir às suas energias livres de ligação/afinidade de ligação em relação à macromolécula alvo (111,112). Geralmente uma função mais simples costuma ser utilizada durante a avaliação das conformações geradas durante o algoritmo de busca. Enquanto que funções de avaliação utilizadas na predição do modo de ligação e no ranqueamento final dos ligantes em uma triagem virtual são mais complexas (111).

Existem três tipos de funções de pontuação: baseadas em campo de força, empíricas e baseadas no conhecimento. As funções baseadas em campo de força levam em conta as interações intermoleculares descritas por termos clássicos da mecânica molecular (descrito mais detalhadamente na seção 1.8.2.3 e Figura 1.11), permitindo, portanto, quantificar a energia proveniente das interações ligante-receptor.

Já as funções empíricas são ajustadas para serem correlacionadas com dados experimentais, tais como energia livre de ligação. Nessa função, cada termo descreve um tipo de evento físico envolvido na formação do complexo proteína-ligante, como ligação de hidrogênio, interação iônica e etc. No desenvolvimento inicial de uma função empírica, geralmente uma série de complexos proteína-ligante com energias livre de ligação ou  $K_i$  conhecidas são usadas como conjunto de treinamento para a geração de um modelo estatístico. As constantes geradas pelo modelo são utilizadas como coeficientes que ajustam os termos da equação e a soma desses termos é igual a pontuação. Essa dependência da precisão dos dados utilizados para desenvolver o modelo representa a principal desvantagem de funções de pontuação empíricas. No entanto, devido à simplicidade dos termos desse tipo de função, as funções empíricas são mais rápidas do que os métodos baseados em campo de força, por exemplo. Já as funções de pontuação baseadas no conhecimento utilizam análises estatísticas da frequência dos pares de átomos encontrados em complexos proteína-ligante conhecidos para obter uma função geral. Essa função de pontuação é, então, construída levando em consideração a frequência com que dois átomos diferentes são encontrados dentro de uma determinada distância no conjunto de dados estruturais utilizados. Diferentes tipos de interação observadas no conjunto de dados são classificados e pesados de acordo com as suas frequências de ocorrência e as pontuações são dadas pela soma das interações individuais observadas no complexo analisado. Esse método, portanto, visa “reproduzir” estruturas experimentais – características *drug-like*, por exemplo – ao invés de considerar as energias de ligação. Nessa abordagem, assume-se que apenas as interações moleculares que se “assemelham” as da base de dados favorecem a ligação e aumentam a afinidade (103,111,112).

Em termos de vantagens e desvantagens, a maioria dos programas de atracamento molecular é capaz de prever com êxito a conformação do ligante no sítio de ligação de uma proteína alvo, podendo isto ser confirmado pela comparação dos complexos preditos com os seus dados de cristalografia de raios-x correspondentes. No entanto, grande parte desses programas não consegue reproduzir a energia livre de ligação ou a afinidade de ligação de um complexo proteína-ligante em concordância satisfatória com dados experimentais. Questões como a consideração da dessolvatação e os efeitos entrópicos são exemplos dos desafios a serem superados pelos programas de atracamento molecular atuais. Apesar dessas

limitações, os programas de atracamento são historicamente bem sucedidos no processo de identificação de novos inibidores de alvos protéicos e no enriquecimento de bancos de dados de compostos em abordagens de triagem virtual baseada em estrutura (113).

### **1.8.2.3 Dinâmica molecular**

A teoria chave-fechadura, em que um receptor rígido era pensado acomodar uma pequena molécula sem sofrer qualquer rearranjo conformacional, foi substituída por modelos que consideram as mudanças conformacionais do complexo receptor-ligante – a já citada teoria do encaixe-induzido. Consequentemente, ferramentas complementares as técnicas experimentais de resolução de estrutura de proteínas foram necessárias para permitir uma visão dinâmica sobre alvos biológicos e ligação de pequenas moléculas. Nesse contexto, a dinâmica molecular (DM) usando aproximações baseadas na física Newtoniana tem sido aplicada na simulação do movimento dos átomos de um sistema em função do tempo (114). Dessa forma, a DM permite explorar quais conformações de uma macromolécula ou um complexo são termicamente acessíveis, além de permitir explorar também o espaço conformacional, por exemplo, em aplicações após o atracamento molecular de ligantes contra uma proteína receptora (115). Com isso, a DM se tornou um método importante no arsenal de ferramentas desenvolvidas para o desenho de novas moléculas bioativas.

A DM envolve primeiramente a preparação do modelo computacional de um sistema molecular a partir de dados experimentais ou modelados computacionalmente. A partir disso, os elementos essenciais para a simulação de DM são o conhecimento do potencial de interação para os átomos, dos quais as forças podem ser calculadas, e a aplicação da equação Newtoniana de movimento que governa a dinâmica desses átomos (116).

Para a resolução dos potenciais de interação e assim obtenção das forças atuando sobre cada átomo, cálculos baseados na mecânica molecular (MM) são aplicados nas simulações de DM. O termo MM se refere ao uso de funções de energia potencial (por exemplo, oscilador harmônico ou potenciais de Coulomb) para modelar sistemas moleculares. Dessa forma, nas simulações de DM clássica, a função de energia potencial, comumente denominada campo de forças, como a mostrada na Figura 1.11, descreve as interações entre os átomos. Um campo de força geralmente consiste nos chamados termos ligados referentes a átomos ligados covalentemente

(comprimento de ligações, ângulos e torções) e termos para átomos não ligados (interações de van der Waals e interações eletrostáticas) (117,118). Ligações químicas e os ângulos atômicos são modelados usando molas virtuais simples. Ângulos diedro (isto é, rotações sobre uma ligação) são modelados usando uma função senoidal que aproxima as diferenças de energia entre conformações eclipsadas e alternadas. Além disso, o potencial referente ao ângulo diedro impróprio é definido em alguns campos de força a fim de descrever as oscilações de um conjunto de quatro átomos (não demonstrado na Figura 1.11). Energias potenciais referentes a átomos não ligados surgem devido a interações de van der Waals, modeladas usando os potenciais Lennard-Jones, e interações eletrostáticas, modeladas usando a lei de Coulomb (114). Por fim, a energia total do sistema é obtida pela soma da energia potencial e a energia cinética dos átomos.

Os parâmetros do campo de força são tipicamente ajustados para reproduzir dados experimentais (por exemplo, espectroscópicos) ou cálculos quântico-mecânicos. Esta parametrização inclui a identificação dos comprimentos ideais das molas que descrevem ligações químicas e ângulos atômicos, determinação da melhor carga atômica parcial utilizadas para calcular a interação eletrostática, identificação dos raios van der Waals atômicos apropriados e etc (114,118).

$$E = \sum_{\text{ligação}} K_r (1 - l_{eq})^2 + \sum_{\text{ângulo}} K_\theta (\theta - \theta_{eq})^2 + \sum_{\text{diedros}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

**Figura 1.11:** Exemplo de uma equação de campo de força. Fonte: Adaptado de Durrant *et al.*, 2011 (114).

A partir disso, a força resultante atuando sobre cada átomo ( $F_i$ ) em um instante de tempo  $t$  é calculada pela derivada primeira da energia potencial ( $E(r_i)$ ), obtida com o campo de força escolhido, na posição  $r_i$  – equivalente as coordenadas do átomo  $i$  (equação 1.1):

$$F_i(t) = -\frac{\partial E(r_i)}{\partial(r_i)} \quad 1.1$$

Com a  $F_i$  é possível calcular a aceleração do átomo  $i$  ( $a_i$ ) de massa  $m_i$  e, a partir dessa aceleração, pode-se obter a velocidade ( $v_i$ ) e a mudança de posição desse átomo ( $r_i$ ) (equação 1.2):

$$F_i(t) = m_i a_i = m_i \frac{d^2 r_i}{dt^2} = m_i \frac{dv_i}{dt} \quad 1.2$$

Dessa forma com as novas posições e velocidades de cada átomo do sistema, obtêm-se as energias potencial e cinética. A aplicação sucessiva das equações de movimento fornece a trajetória, que se refere ao conjunto de posições e velocidades de cada átomo do sistema ao longo do tempo. A integração numérica das equações Newtonianas de movimento na DM em si é realizada através de algoritmos, como o algoritmo de Verlet, algoritmo de Beeman e o algoritmo *leap-frog*, em que a integração é dividida em pequenos intervalos de tempo geralmente na escala de femtosegundos. No caso do algoritmo de Verlet, por exemplo, utiliza-se as posições e acelerações no tempo  $t$  assim como as posições do passo anterior,  $r(t - \Delta t)$ , a fim de determinar as novas posições em um tempo posterior ( $t + \Delta t$ ) (equação 1.3) (119):

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + a(t)\Delta t^2 \quad 1.3$$

Já no algoritmo *leap-frog* é usada a posição  $r$ , o tempo  $t$  e a velocidade  $v$  no tempo  $t - \frac{1}{2}\Delta t$ . O algoritmo atualiza as novas posições e velocidades usando as forças  $F(t)$  determinadas pelas posições no tempo  $t$  usando as seguintes relações (equações 1.4 e 1.5) (120):

$$v\left(t + \frac{1}{2}\Delta t\right) = v\left(t - \frac{1}{2}\Delta t\right) + \frac{\Delta t}{m}F(t) \quad 1.4$$

$$r(t + \Delta t) = r(t) + \Delta t v\left(t + \frac{1}{2}\Delta t\right) \quad 1.5$$

Tipicamente é importante simular acuradamente as condições experimentais a serem replicadas. Vários valores para condições físicas como para temperatura e pressão podem ser considerados na simulação. Esse conjunto de configurações e propriedades mantidas constantes que representam o estado do sistema durante a integração numérica das equações de movimento é chamado de *ensemble*. Somado a isso, para se ter uma representação razoável do ambiente em que a macromolécula de interesse está inserida, a solvatação do sistema também é importante de ser considerada e para isso existem modelos explícitos e implícitos<sup>3</sup> de solvatação. Uma vez que as simulações de DM são feitas em caixas virtuais (espaço tridimensional no qual o macromolécula e o solvente são colocados), para evitar efeitos de superfície indesejáveis e obter as propriedades macroscópicas a partir das trajetórias das simulações são utilizadas condições periódicas de contorno que envolvem a representação do sistema de forma periódica (119,121).

É importante destacar que antes da simulação da DM propriamente dita, o sistema com as coordenadas iniciais dos átomos do soluto (macromolécula) e solventes deve ser minimizado, principalmente por causa da solvatação a que o soluto foi submetido. A minimização é uma etapa importante para evitar impedimentos estéricos entre os átomos e visa encontrar um conjunto de coordenadas que minimizam a energia potencial do sistema fazendo com que ele alcance um mínimo de energia local próximo na superfície de potencial. Dessa forma, as posições dos átomos são ajustadas a fim de relaxar as distorções nas ligações químicas, nos ângulos entre átomos e nos contatos de van der Waals. Algoritmos como o método *steepest descent* e o método dos gradientes conjugados podem ser usados para esse processo de minimização. Após o sistema ter sido minimizado, ele pode ser submetido a elevação gradual de temperatura até a temperatura de interesse. A temperatura de um sistema é correlacionada com a média das velocidades das partículas, no caso átomos, assim as velocidades iniciais podem ser fornecidas pela distribuição de Maxwell-Boltzmann na temperatura desejada (119,121).

No início de uma simulação de DM, as condições encontradas diferem do equilíbrio termodinâmico, ou seja, enquanto o sistema não alcançar o equilíbrio ele provavelmente não será representativo do fenômeno de interesse. Esse período de

---

<sup>3</sup> Modelos implícitos de solvatação consideram, ao invés de um sistema contendo centenas ou milhares de moléculas de águas – modelo explícito de solvatação – apenas a influência média do solvente sobre o sistema. Em outras palavras, nesse modelo um meio contínuo infinito com propriedades dielétricas e hidrofóbicas da água é considerado (119,163).

equilíbrio é variável e depende do sistema em estudo. Quando se alcança o equilíbrio, pode-se então gerar as trajetórias da DM e calcular as propriedades para o sistema em questão (119,121).

A trajetória resultante da DM de um complexo macromolécula-ligante, por exemplo, é analisada para extrair informações como distâncias e interações entre átomos ou resíduos de interesse. A estabilidade global do complexo pode ser controlada por análises, por exemplo, do RMSD (sigla do inglês *root mean square deviation* ou desvio quadrático médio), a partir de uma conformação de referência, e das flutuações dos átomos por RMSF (sigla do inglês *root mean square fluctuation* ou flutuação quadrática média). Além disso, uma importante aplicação de simulações de DM – especialmente no contexto do desenho racional de fármacos – é a estimativa da variação de energia livre,  $\Delta G$ . A energia livre de ligação, que pode ser diretamente correlacionado a constante de ligação,  $K_i$  (117).

### **1.8.3 ADME-Tox (Absorção, Distribuição, Metabolismo, Excreção e Toxicidade)**

ADME ou farmacocinética é o termo utilizado para o estudo de como um fármaco é absorvido, distribuído, metabolizado e excretado pelo corpo. O estudo da toxicidade de um fármaco é outra propriedade importante de ser avaliada juntamente com as propriedades farmacocinéticas. Dessa forma, absorção é o processo pelo qual um fármaco alcança a corrente sanguínea a partir do local de administração, podendo a administração ser, por exemplo, oral, nasal, subcutânea, intramuscular. Em todos os tipos de administração, com exceção da administração intravenosa, o fármaco deve passar através de membranas celulares para alcançar a corrente sanguínea. No caso da administração oral, o fármaco ao entrar na corrente sanguínea é encaminhado para o fígado (também conhecido como efeito de primeira passagem), onde pode ser submetido a um extenso metabolismo antes de cair na circulação sistêmica. Uma vez que o fármaco tenha sido absorvido, ele é rapidamente distribuído na corrente sanguínea e então é distribuído mais vagarosamente para vários tecidos e órgãos. Nesse contexto, o termo distribuição se refere a quais “compartimentos” (tecidos e órgãos) o fármaco alcança. O metabolismo envolve a ação de enzimas específicas sobre o fármaco, convertendo-o a um ou mais metabólitos. Juntamente com a excreção do fármaco, o metabolismo é o principal meio pelo qual o corpo remove o fármaco do sistema. Um objetivo comum no processo de descoberta de fármacos é identificar um candidato a fármaco que apresente níveis terapêuticos no



sistema que possam ser mantidos seguindo um esquema de dosagem conveniente. Além disso, os metabólitos podem também ser biologicamente ativos, podendo a atividade biológica resultante causar efeitos desejáveis ou não. Já a excreção envolve processos pelos quais o corpo elimina um fármaco não metabolizado ou o seu metabólito resultante. A urina e as fezes são as principais rotas de excreção de fármacos, assim como a exalação no caso de metabólitos voláteis. Já a toxicidade se refere ao grau de dano que um composto pode causar a um organismo ou subestrutura do organismo, como células e órgãos, e continua a ser uma das razões mais importantes para o fracasso no desenvolvimento tardio de fármacos (35,122).

Apesar de ser possível a determinação de propriedades ADME-Tox através de triagens *in vitro*, a capacidade de predição *in silico* de algumas dessas propriedades nos primeiros estágios do desenvolvimento de novos fármacos tem se tornado cada vez mais importantes, uma vez que realizar experimentos ADME-Tox para um número vasto de compostos é muitas vezes inviável. Dessa forma, a combinação de métodos *in vitro* e *in silico* no processo de descoberta de novos fármacos auxilia na redução de problemas associados com a segurança desses compostos. De maneira geral, esses métodos *in silico* usam abordagens estatísticas, aprendizagem de máquina, descritores moleculares e dados experimentais para modelar processos como biodisponibilidade oral, absorção intestinal, permeabilidade e mutagenicidade (123).

Uma das primeiras etapas *in silico* para inferir a segurança de um composto é a avaliação de suas propriedades físico-químicas, como solubilidade, lipofilicidade, massa molecular, topologia e ionização. A análise de tais propriedades é importante uma vez que elas têm efeito significativo na farmacocinética, principalmente a relacionada a biodisponibilidade oral, a potência e a seletividade de um composto (122). Dessa forma, a abordagem mais comum envolve a predição dessas propriedades seguida pela aplicação de filtros ou regras que delimitam as propriedades físico-químicas adequadas de fármacos ou candidatos a fármacos. Entre as regras já criadas, a regra de cinco de Lipinsk é a mais famosa. Nessa regra quatro propriedades físico-químicas são consideradas para determinar se um composto é absorvido oralmente: massa molecular menor ou igual a 500 Da, coeficiente de partição octanol/água (LogP) menor ou igual a 5, número de doadores de ligação hidrogênio menor ou igual a 5 e o número de aceptores de ligação hidrogênio menor ou igual a 10. Se um composto viola uma ou mais regras, este composto pode não ser ativo oralmente (124). Em contraste com a regra de cinco de

Lipinsk, Verber e colaboradores mediram a biodisponibilidade oral de 1.100 candidatos a fármacos e descobriram que uma flexibilidade molecular reduzida é favorável para uma boa biodisponibilidade oral. Dessa forma, Veber determinou que um número de ligações rotacionáveis menor ou igual a 10, assim como uma área de superfície polar (PSA, do inglês *polar surface area*, que se refere a soma das superfícies dos átomos polares) menor ou igual a 140 Å<sup>2</sup> ou, alternativamente, uma quantidade total de ligações de hidrogênio menor ou igual a 12 são preditores importantes para uma boa biodisponibilidade oral, independente da massa molecular do composto em questão. Compostos que obedecem os critérios de Lipinsk ou Veber são frequentemente chamados de fármaco-similares (*drug-like*) (124,125). Além da predição da absorção oral baseadas em regras, a predição baseada em modelos de QSAR, obtidos a partir da correlação entre estrutura/propriedade molecular de um conjunto de compostos e os perfis de absorção oral deles, também tem sido largamente utilizados (122).

Já a distribuição de um composto pode ser predita, em parte, pela análise *in silico* de parâmetros como a interação do candidato a fármaco com proteínas do plasma sanguíneo, efluxo de glicoproteína P e a penetração na barreira hematoencefálica. Como, em geral, somente o fármaco livre é capaz de interagir com a proteína alvo, a análise da interação do candidato a fármaco com proteínas do plasma sanguíneo são importantes no estudo da distribuição (126). Os fármacos podem ligar a proteínas do plasma a taxas constantes, causando uma menor biodisponibilidade. Métodos computacionais que empregam atracamento e dinâmica molecular são muito utilizados, permitindo estimar a afinidade de ligação entre o fármaco e os sítios de ligação nas proteínas do plasma, como a albumina. Além disso, métodos baseados em QSAR também já foram descritos, permitindo assim a predição da interação com proteínas plasmáticas, utilizando-se as características estruturais dos candidatos a fármacos (122). Já a glicoproteína P é um transportador e membro importante da superfamília ABC (do inglês *ATP-Binding Cassette*), que contém proteínas que usam a ATP como fonte de energia para o bombeamento de substratos contra um gradiente de concentração. A glicoproteína P, em particular, é encontrada em células em todo o corpo, incluindo no intestino e na barreira hematoencefálica. Portanto, esse transportador é indicado como tendo uma função importante na distribuição de compostos, já que ele é capaz de transportar moléculas estruturalmente diversas de diferentes compartimentos no corpo (126). A fim de

predizer substratos, inibidores e moduladores de glicoproteína P, vários modelos *in silico* foram desenvolvidos usando abordagens como QSAR, bem como modelagem de farmacóforos e diversas técnicas de aprendizagem de máquina (122,127). Já a distribuição de candidatos a fármacos para o sistema nervoso central envolve a penetração da barreira hematoencefálica para que o composto possa assim alcançar seu alvo molecular. Em casos de alvos periféricos, pouca ou nenhuma penetração na barreira hematoencefálica pode ser requerida, visando evitar efeitos adversos. Abordagens de predição *in silico* da penetração de compostos pela barreira hematoencefálica envolvem, em sua grande maioria, modelos que partem do pressuposto de que os compostos são transportados através da barreira hematoencefálica por difusão passiva. Regras parecidas com a regra de cinco de Lipinsk envolvendo parâmetros que contribuem para a habilidade de moléculas atravessarem a barreira hematoencefálica tem sido empregadas. Por exemplo, a regra que determina que compostos com massa molecular menor que 450 Da ou com PSA menor que 100 Å<sup>2</sup> são mais prováveis de penetrar a barreira hematoencefálica. A maioria dos modelos preditivos são baseados em uma abordagem de regressão linear múltipla (gerando modelos de QSAR), no qual muitos usam propriedades físico-químicas (122,127).

O metabolismo de compostos é realizado principalmente pelas chamadas enzimas de fase I (oxidação, redução e hidrólise) e II (sulfo-conjugação, entre outros), que são, na sua maior parte, produzidas no fígado. A análise do metabolismo de um candidato a fármaco é importante a fim de identificar problema relacionados a baixa biodisponibilidade; efeitos tóxicos causados por metabolitos reativos e interações medicamentosas, incluindo inibição enzimática, indução e inativação baseada em mecanismos (126). Os modelos de predição relacionados ao metabolismo podem ser principalmente focados nos seguintes estudos: (i) modelos de interação de enzimas com xenobióticos, que são frequentemente usados para distinguir se um xenobiótico é um substrato ou inibidor do sistema monooxigenase do citocromo p450 (CYP450). Esses modelos são construídos utilizando principalmente abordagens de atracamento e dinâmica molecular; (ii) modelos de depuração (*clearance*) no fígado que podem prever quantitativamente a estabilidade dos xenobióticos, utilizando abordagens que vão de modelos de classificação baseados em características físico-químicas à aprendizagem de máquina; (iii) modelos de predição dos pontos na estrutura de um xenobiótico suscetíveis a metabolização. Abordagens de desenhos de fármacos

baseadas tanto em estrutura quanto em ligante são utilizados na construção desses modelos a fim de determinar a reatividade e acessibilidade dos átomos de um metabólito e, dessa forma, os pontos mais suscetíveis a metabolização; (iv) modelos de predição de potenciais metabólitos derivados de um xenobiótico. Esses modelos são determinados por sistemas e métodos estatísticos baseados em regras de reações metabólicas disponibilizadas na literatura (122,128).

Uma prioridade crítica no desenvolvimento de medicamentos é a identificação precoce de toxicidade grave em candidatos a fármacos a fim de evitar falhas desnecessárias nos testes pré-clínicos e clínicos. Portanto, a predição de toxicidade antes mesmo da síntese de fármacos é uma abordagem importante. Uma grande quantidade de modelos *in silico* aplicando diferentes métodos são utilizados para a predição de toxicidade, como abordagens envolvendo QSAR, aprendizagem de máquina e aplicação de modelagem de “farmacóforo” (a fim de descrever, nesse caso, grupos toxicóforicos) e métodos baseados na similaridade química e na identificação de fragmentos frequentemente representados em compostos tóxicos. Essas abordagens têm sido aplicadas em diversos estudos a fim de prever, por exemplo, toxicidade aguda, genotoxicidade e hepatotoxicidade (122,128).

## **1.9 Justificativa**

A doença de Chagas é um problema de saúde pública na América Latina, com ocorrência cada vez maior em países não endêmicos, como EUA, Canadá, Austrália, Japão e alguns países da Europa. Estimativas mais recentes mostram que cerca de sete milhões de pessoas estão infectadas com o protozoário *T. cruzi* em todo o mundo. Infelizmente, não há vacinas para a doença de Chagas, os tratamentos disponíveis têm efeito limitado sobre a fase crônica da doença e todos os recentes estudos clínicos com fármacos, antes promissores, foram malsucedidos. Por essa razão, o desenvolvimento de novos medicamentos mais eficazes contra o parasita é uma necessidade urgente.

Considerando, para este fim, uma abordagem de descoberta de novos fármacos baseada no alvo, a enzima TcRpiB é um alvo terapêutico com qualidades marcantes, principalmente por pertencer a uma via metabólica essencial ao *T. cruzi* e ser uma enzima análoga a Rpi do hospedeiro vertebrado humano. A partir dessa perspectiva, esse estudo visou a identificação de inibidores potenciais da atividade

enzimática TcRpiB através de metodologias de descoberta de fármacos assistidas por computador, como a modelagem de farmacóforo, busca por similaridade, atracamento molecular e dinâmica molecular. Desta forma pretendemos contribuir para estabelecer estratégias mais eficazes para o desenvolvimento de uma quimioterapia antiparasitária no controle da doença de Chagas.

## 2 OBJETIVOS

### 2.1 Objetivo Geral

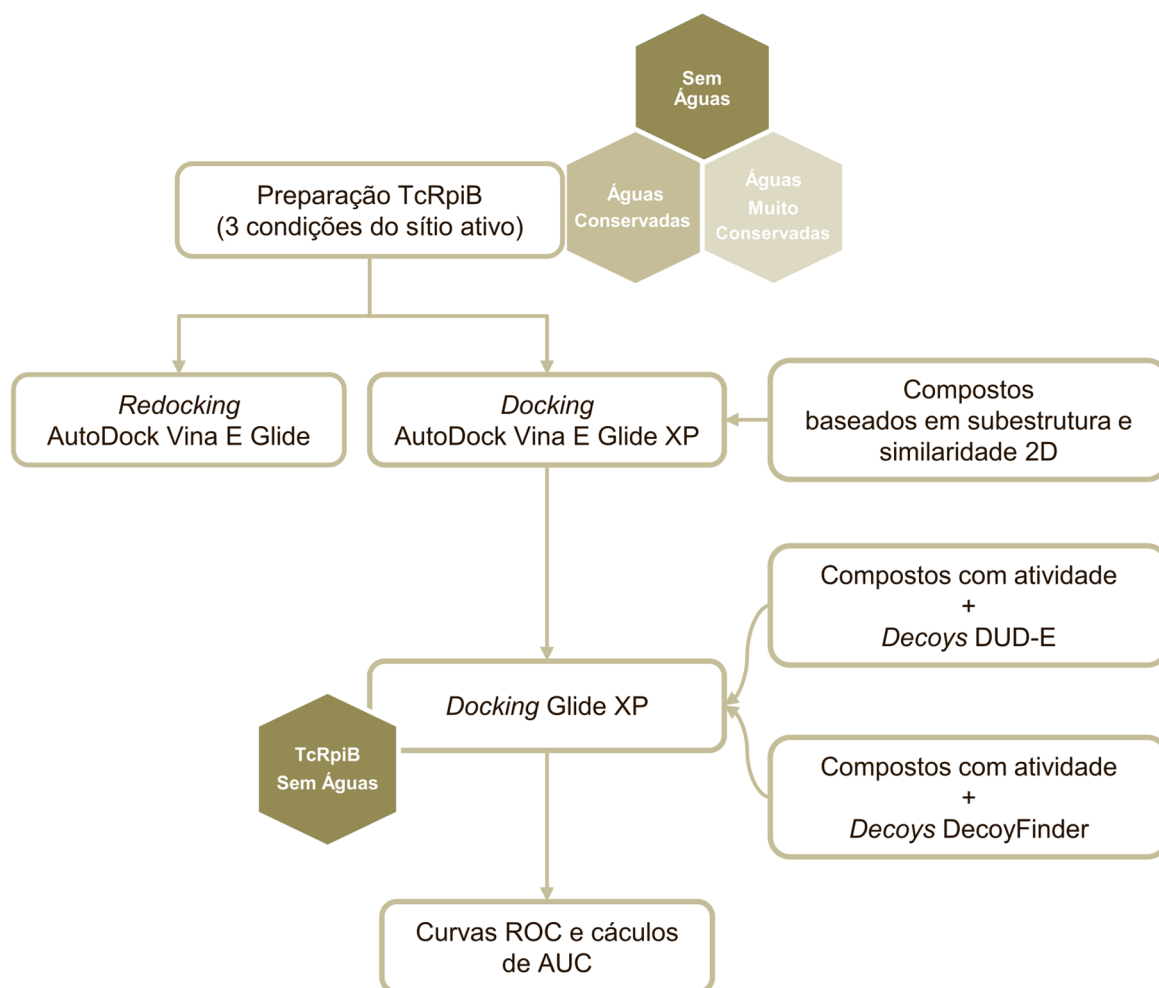
Identificação *in silico* de compostos com potencial de inibir a atividade enzimática da ribose 5-fosfato isomerase de *Trypanosoma cruzi* utilizando abordagens computacionais baseadas em ligante e estrutura.

### 2.2 Objetivos Específicos

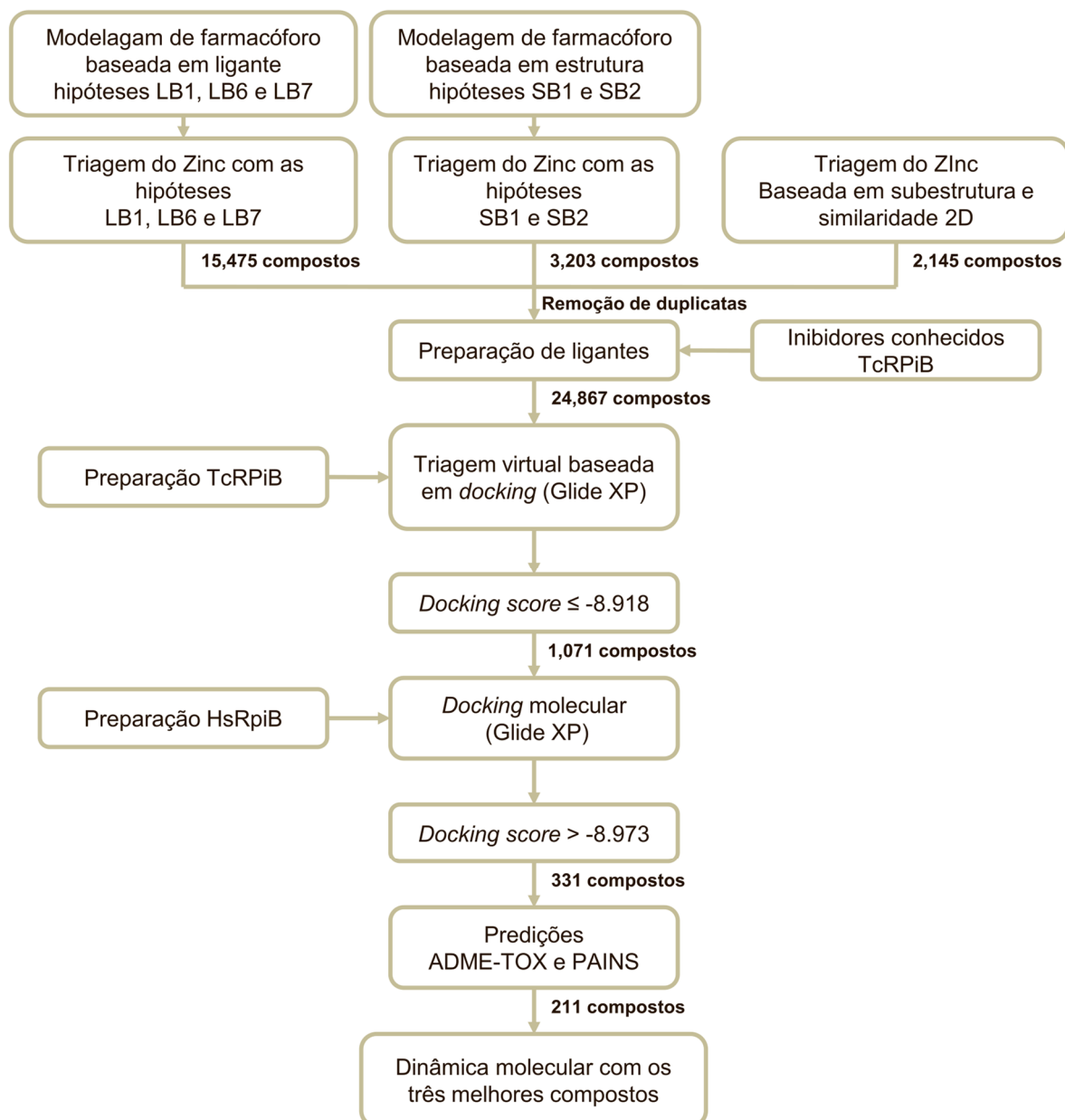
- i. Analisar e comparar as estruturas tridimensionais de HsRpiA e TcRpiB;
- ii. Avaliar a influência das moléculas de água conservadas no sítio ativo de TcRpiB em ensaios de atracamento molecular com os programas AutoDock Vina e Glide XP;
- iii. Realizar a modelagem de farmacóforo baseada em ligante e em estrutura considerando inibidores conhecidos de RpiB e estruturas cristalográficas disponíveis de TcRpiB;
- iv. Criar uma biblioteca de compostos enriquecida com potenciais inibidores de TcRpiB através da triagem virtual baseada em farmacóforo, similaridade 2D e subestrutura do banco de dados ZINC;
- v. Predizer a afinidade de ligação dos compostos da biblioteca de compostos no sítio ativo de TcRpiB e HsRpiA utilizando a abordagem de atracamento molecular;
- vi. Selecionar compostos com alta afinidade predita com TcRpiB e com baixa afinidade com HsRpiA;
- vii. Predizer propriedades farmacocinéticas e a toxicidade dos compostos selecionados;
- viii. Analisar a diversidade estrutural dos compostos selecionados;
- ix. Analisar a estabilidade de interação dos compostos mais bem sucedidos em complexo com TcRpiB utilizando a metodologia de dinâmica molecular.

### 3 MATERIAL E MÉTODOS

A descrição geral do processo de identificação de potenciais inibidores de TcRpiB está representada nas Figura 3.1 e Figura 3.2.



**Figura 3.1:** Visão geral dos procedimentos para a avaliação das melhores condições do sítio ativo quanto a presença de moléculas de águas conservadas e dos programas de atracamento molecular AutoDock Vina e Glide XP.



**Figura 3.2:** Visão geral das etapas desenvolvidas na identificação de potenciais inibidores de TcRpiB.

### 3.1 Conjunto de dados utilizados

Nesse estudo, as estruturas cristalográficas de TcRpiB complexadas ao substrato R5P (**7**; Apêndice B- Fig. 7.2) (PDB ID: 3K7S) e ao inibidor 4PEH (**41**) (ID PDB: 3K8C) foram obtidas a partir do RCSB *Protein Data Bank* (PDB) (<http://www.rcsb.org/>) e foram utilizadas nos ensaios de atracamento molecular e modelagem de farmacóforo. Além disso, para determinar as moléculas de águas



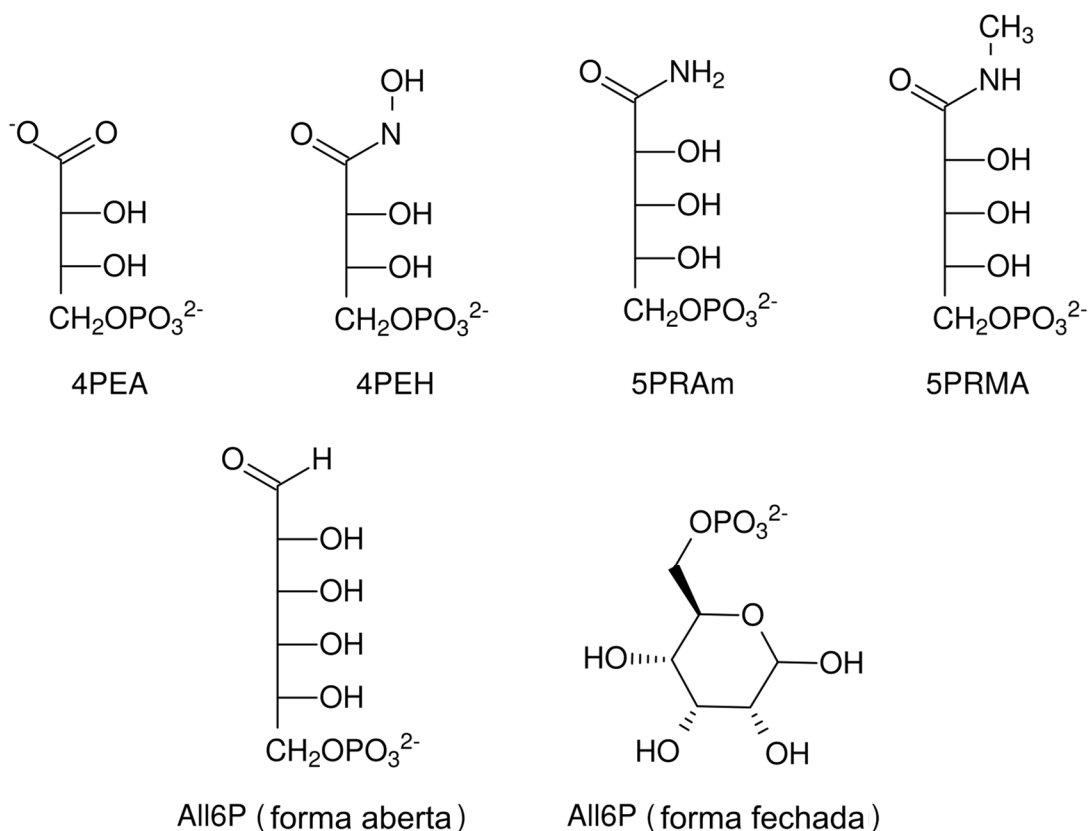
conservadas no sítio ativo da estrutura de TcRpiB (PDB 3K7S), as seguintes estruturas cristalográficas de RpiBs de diferentes organismos foram selecionadas: PDB Ids - 3K70, 3K8C, 3M1P, 3K7P (*T. cruzi*); 1USL, 2BES, 2BET, 2VVP, 2VVQ, 2VVO (*M. tuberculosis*); 2VVR, 1NN4 (*E. coli*); 3HE8, 3HEE, 3PH3, 3PH4 (*C. thermocellum*) e 4EM8 (*A. phagocytophilum*). O modelo estrutural de HsRpiA, também utilizado nos ensaios de atracamento molecular, foi obtido do trabalho previamente publicado por Capriles e colaboradores (55). Segundo os autores, o modelo foi construído através de modelagem comparativa utilizando o programa MODELLER (<https://salilab.org/modeller/>) (129) e abordagens de modelagem *ab initio* utilizando o programa GAPP (130). O modelo foi gerado como uma proteína tetramérica utilizando como *query* a sequência Hsa: 22934 do banco de dados KEGG (<http://www.genome.jp/kegg/>) e como moldes as proteínas de PDB 1XTZ (identidade de 46%), 1LK7 (identidade de 42%), 3HHE (identidade de 39%), 2F8M (identidade de 35%)(55). Os resíduos do sítio ativo dos moldes apresentam uma identidade compreendida entre 75% e 94% em relação a sequência de HsRpiA.

Um conjunto de três inibidores conhecidos de TcRpiB foi selecionado a partir da literatura para as etapas de modelagem de farmacóforo baseada em ligante e de ensaios de atracamento molecular, representando, nesse caso, os controles positivos. Os inibidores selecionados foram 4PEH (ácido 4-fosfo-D-eritrono-hidroxâmico, **41**), 4PEA (4-fosfo-D-eritronato, **38**) (70) e a forma aberta e fechada de All6P (D-Alose 6-fosfato, **15**) (54). Além disso, dois outros inibidores, 5PRAm (5-fosfo-D-ribonamida, **53**) e 5PRMA (N- (5-fosfo-D-ribonóil) –metilamina, **55**), que são descritos como bons inibidores de RpiBs de outros organismos (*Mycobacterium tuberculosis* e *E. Coli*) (92), foram selecionados para adicionar variedade à modelagem de farmacóforo baseada em ligante (Figura 3.3). As estruturas tridimensionais dos inibidores e substratos foram desenhadas no Maestro (versão 9.3), interface gráfica utilizada para todos os programas disponibilizadas pelo grupo Schrödinger. A biblioteca de compostos desse estudo foi obtida a partir do banco de dados ZINC (<http://ZINC.docking.org/>) (131) – subconjunto *purchasable* (versão de dezembro de 2014) contendo 22.724.825 compostos. Um conjunto de 700 compostos *decoys*<sup>4</sup> foi obtido no servidor DUD-E (<http://dude.docking.org/>) utilizando os substratos e inibidores conhecidos de TcRpiB

---

<sup>4</sup> *Decoys* são compostos com propriedades físicas similares aos compostos *query* (compostos ativos em relação a proteína alvo), mas com estruturas químicas distintas, representando, portanto, compostos sem atividade em potencial em relação a proteína alvo (164).

como *query*. Além disso, mais 700 *decoys* foram obtidos utilizando o programa DecoyFinder.



**Figura 3.3:** Inibidores de RpiBs utilizados na modelagem de farmacóforo baseado em ligante.

## 3.2 Análises das proteínas

### 3.2.1 Comparação das proteínas TcRpiB e HsRpiA

Foi realizada uma comparação entre as estruturas primária, secundária e terciária de HsRpiA e TcRpiB (PDB ID: 3K7S) com o objetivo de avaliar as diferenças e semelhanças entre essas proteínas. As estruturas primárias foram alinhadas umas contra as outras para acessar a identidade e similaridade entre elas usando o programa ClustalW através do módulo *Multiple Sequence Viewer* do Maestro (versão 9.3) (*Schrödinger Suite*). As estruturas secundárias também foram avaliadas com o módulo *Multiple Sequence Viewer* do Maestro. A sobreposição estrutural 3D foi realizada com o programa TM-align (<http://zhanglab.ccmb.med.umich.edu/TM-align/>) (132) que compara as estruturas usando uma abordagem independente de ordem da sequência e pode ser usada para identificar semelhanças locais entre as proteínas. Esse programa permite inferir também se as proteínas sobrepostas são

correlacionadas evolutivamente, valores acima de 0,5 indicam sobreposição de proteínas com o mesmo enovelamento, enquanto que valores abaixo de 0,3 se referem a sobreposições geradas por proteínas não correlacionadas (132). O sítio ativo de ambas as proteínas e os modos de ligação dos substratos foram avaliados utilizando o Pymol v1.8.2.3 (<http://www.pymol.org/>). Os potenciais da superfície eletrostática de HsRpiA e TcRpiB – com os estados de protonação adequados – foram calculados pelo software APBS com o objetivo de avaliar as diferenças eletrostáticas entre seus sítios ativos.

### **3.2.2 Análise das moléculas de águas conservadas no sítio ativo de TcRpiB**

A análise das moléculas de água conservadas no sítio ativo da proteína de interesse previamente a estudos de atracamento molecular é um processo crucial, já que a presença de água pode afetar a quantidade de ligações de hidrogênio que podem ocorrer entre o ligante e os aminoácidos no sítio ativo de uma enzima. Para determinar as moléculas de águas conservadas, foram avaliados os sítios ativos da proteína TcRpiB (PDB 3K7S) e de outras proteínas depositadas no PDB com atividade enzimática descrita de RpiB (*Enzyme Commission* (EC Number) 5.3.1.6) e com similaridade maior que 50% em relação à proteína TcRpiB. A partir das estruturas primárias referentes a essas proteínas, foi feito um alinhamento múltiplo utilizando o programa ClustalW. Para a avaliação das moléculas de água nos sítios ativos, todas as estruturas foram sobrepostas com auxílio do programa Pymol v1.8.2.3. As interações entre a enzima, ligante (quando presente) e as moléculas de água em cada complexo também foram avaliadas utilizando o programa Pymol. As moléculas de água das estruturas analisadas com RMSD menor que 0,5 Å em relação às moléculas de água da estrutura TcRpiB de referência foram consideradas como moléculas de águas conservadas. As moléculas de água que se mantiveram conservadas em mais que 60% das estruturas compuseram um grupo denominado de “águas conservadas”. Já as moléculas de água que se mantiveram conservadas em mais que 80% dos casos formaram um grupo denominado de “águas muito conservadas”.

## **3.3 Modelagem de farmacóforo**

### **3.3.1 Modelagem de farmacóforo baseado em ligante**

As hipóteses de farmacóforo baseadas em ligante foram geradas usando o servidor Pharmagist ([bioinfo3d.cs.tau.ac.il/PharmaGist/](http://bioinfo3d.cs.tau.ac.il/PharmaGist/)) (133) que é capaz de detectar grupos farmacofóricos a partir de um conjunto de compostos conhecidos por se ligarem ao alvo de interesse. O método utilizado pelo Pharmagist envolve quatro etapas principais: (i) representação do ligante; (ii) alinhamento par a par dos ligantes; (iii) alinhamento múltiplo de todos os ligantes e, (iv) agrupamento e geração da hipótese. Na primeira etapa, cada ligante de entrada é processado separadamente. O método detecta as ligações rotacionáveis de cada ligante, dividindo-o em grupos rígidos. Além disso, são atribuídas as características físico-químicas dos ligantes; isto é, os pontos farmacofóricos. Para o cálculo do alinhamento par-a-par, um ligante é tratado como rígido (ligante *pivot*) e um ligante é tratado como flexível (ligante alvo). A pontuação dos alinhamentos resultantes é uma soma ponderada das características do *pivot* que conseguem ser alinhadas ao alvo. Duas características, uma do *pivot* e uma de uma conformação do ligante alvo, são consideradas alinhadas se forem do mesmo tipo e a distância entre elas estiver abaixo de um limiar predefinido (1Å por padrão). Ao final dessa etapa são gerados vários alinhamentos par-a-par de alta pontuação entre o *pivot* e o ligante alvo. Os alinhamentos par-a-par entre o *pivot* e os ligantes alvos são combinados em alinhamentos múltiplos e novamente pontuações são fornecidas. O objetivo é encontrar subconjuntos de pontos farmacofóricos do *pivot* que sejam combinados por tantos alinhamentos par-a-par para diferentes ligantes alvo quanto possível. Por padrão, o algoritmo seleciona iterativamente cada um dos ligantes de entrada para servir como *pivot* por vez. Na última etapa, as hipóteses de farmacóforo candidatas derivadas das diferentes iterações com *pivot* são agrupadas e os de maior pontuação e não redundantes são relatados (133).

Nesse estudo, um conjunto de inibidores conhecidos para RpiBs, em geral (seção 3.1, Figura 3.3) no formato mol2, foi utilizado como arquivo de entrada para a modelagem de farmacóforo através do Pharmagist. Esse programa gera como arquivo de saída uma lista de hipóteses de farmacóforo, em que cada hipótese é gerada no formato jmol. Essas hipóteses, conforme descrito anteriormente, possuem uma pontuação associada a elas e contém um padrão de pontos farmacofóricos - teoricamente responsáveis pela ligação do ligante ao receptor – compartilhado por todos ou quase todos os ligantes que serviram de arquivos de entrada na etapa inicial. As hipóteses de farmacóforo geradas foram escolhidas com base nos seguintes critérios: (i) maiores pontuações; (ii) inclusão na construção das hipóteses de todos

os inibidores de RpiBs estudados; (iii) apresentarem pontos farmacofóricos conhecidamente importantes na interação proteína-ligante (iv) conseguirem retornar os substratos conhecidos de Rpis em busca no banco de dados ZINC, através do servidor ZINCPharmer (ZINCpharmer.csb.pitt.edu/) (134).

### **3.3.2 Modelagem de farmacóforo baseado em estrutura**

A hipótese de farmacóforo baseada em estrutura foi construída através do programa ZincPharmer (134) que, além de permitir a busca de compostos no banco de dados ZINC, possibilita também, através da tecnologia do programa Pharmer (135), a modelagem de farmacóforo a partir da estrutura da proteína de interesse. Essa modelagem envolve duas etapas: (i) a identificação de pontos farmacofóricos a partir do padrão detectado de interação proteína-ligante na estrutura 3D disponível e, (ii) seleção e edição pelo usuário dos pontos farmacofóricos mais relevantes que em conjunto representarão a hipótese de farmacóforo. A partir da hipótese de farmacóforo é possível fazer a triagem na biblioteca de compostos do próprio ZincPharmer, conforme abordado no item 3.3.3. A biblioteca do ZincPharmer contém compostos de vários *subsets* do ZINC, em que cada um desses compostos é representado por até dez confôrmeros gerados a partir de uma amostragem do espaço conformacional e seleção dos melhores confôrmeros. Cada um desses confôrmeros é representado por seus pontos farmacofóricos previamente identificados com o programa Pharmer. Esses pontos farmacofóricos são mantidos no banco de dados do ZincPharmer de maneira a otimizar o processo de triagem virtual de sua biblioteca de compostos (134).

Neste estudo, as estruturas cristalográficas 3K7S e 3K8C, contendo, respectivamente, o substrato ribose 5-fostato forma aberta e o inibidor 4PEH (41), foram sobrepostas e utilizadas para geração da hipótese de farmacóforo baseada em estrutura com o ZincPharmer. Essas estruturas cristalográficas foram selecionadas visto que seus ligantes possuem características farmacofóricas e grupos químicos em comum sabidamente relevantes para interação com a TcRpiB. Essas características farmacofóricas foram consideradas para a geração final das hipóteses de farmacóforo baseada em estrutura.

### **3.3.3 Triagem virtual baseada em farmacóforo e similaridade 2D**

Tanto as hipóteses selecionadas baseadas em ligante quanto as baseadas em estrutura foram utilizadas como entrada no servidor ZINCPharmer visando a

identificação de compostos no banco de dados ZINC que correspondessem às hipóteses em questão – isto é, triagem virtual baseada em farmacóforo. Esse processo envolveu a triagem de mais de 100 milhões de conformeros referentes aos ~22 milhões de compostos do ZINC (*subset purchasable*) a fim de identificar inibidores em potencial de RpiBs.

Adicionalmente, realizou-se uma busca baseada em similaridade estrutural também na base de dados ZINC, para obter compostos com mais de 50% de semelhança com substratos de TcRpiB (forma linear de R5P (**7**), forma de anel R5P e Ru5P (**10**)) ou compostos contendo grupos fosfato e/ou hidroxilas em sua subestrutura. As pesquisas incluíram apenas o subconjunto *purchasable* do banco de dados ZINC. Além disso, foram previamente eliminados dessa busca compostos que não se enquadraram na regra dos cinco de Lipinski (136), ou seja, compostos com mais que 5 doadores e 10 aceptores de ligações de hidrogênio, com peso molecular maior que 500 e logP maior que 5. Esse conjunto de compostos baseado em similaridade, em particular, foi utilizado na avaliação dos dois programas de atracamento (Glide e AutoDock Vina) empregados nesse estudo a fim de auxiliar na escolha do programa com melhor desempenho para a etapa de triagem virtual baseada em atracamento.

Os compostos recuperados, tanto pelas hipóteses de farmacóforo como também na busca de similaridade, foram utilizados para construir a biblioteca de compostos – enriquecida com inibidores em potencial de RpiBs – para a triagem virtual baseada em atracamento. Essa biblioteca foi submetida ao programa Knime (137) para remover as duplicatas e eliminar compostos obtidos pela modelagem de farmacóforo que não obedecessem a regra de Lipinski.

### **3.4 Atracamento molecular**

#### **3.4.1 Preparação dos ligantes**

Todas as estruturas 3D de inibidores de RpiB conhecidos, substratos e compostos de ZINC foram obtidos no formato mol2. Os *decoys* foram obtidos no formato smiles e convertidos a mol2 através do programa Maestro. Todos os ligantes foram preparados com o programa Ligprep da *Schrödinger Suite* e tiveram as suas energias minimizadas usando o campo de força OPLS-2005. Para cada ligante, todos

os possíveis estados de ionização em  $\text{pH } 7 \pm 1$  foram atribuídos, a quiralidade foi retida e tautômeros foram gerados. A quiralidade de todos os ligantes foi retida visando evitar a geração de estereoisômeros.

Além disso, para os ensaios de atracamento com o programa AutoDock Vina, os inibidores, os substratos e os compostos obtidos pela busca por similaridade 2D foram preparados e convertidos ao formato pdbqt utilizando o script `prepare_ligand4.py` disponível no pacote do programa MGLTools (138). No caso do Glide, todos os compostos preparados no formato mol2 com o Ligprep foram utilizados diretamente nos ensaios de atracamento com o programa.

### 3.4.2 Preparação das proteínas

A estrutura homotetramérica de TcRpiB (PDB 3K7S) complexada com a forma linear do substrato R5P (7) foi utilizada para os estudos de atracamento molecular com os programas AutoDock Vina (139) e Glide XP (140). A preparação do receptor envolveu inicialmente a retirada do ligante do arquivo PDB. Além disso, as subunidades C e D foram excluídos desse arquivo de modo que apenas as subunidades A e B foram mantidas. A exclusão de duas das quatro subunidades foi realizada uma vez que a TcRpiB é descrita como um dímero em solução e os sítios ativos se localizarem entre as subunidades idênticas “A e B” e “C e D” (54), sendo suficiente a representação de apenas um par de subunidades.

De acordo com as análises de água descritas no item 3.2, três arquivos PDBs referentes ao receptor TcRpiB foram criados quanto à presença e ausência de moléculas de água: (i) receptor com ausência de águas, (ii) receptor com presença do grupo de “águas conservadas” no sítio ativo e, (iii) receptor com presença do grupo de “águas muito conservadas” no sítio ativo. Para determinar os estados de protonação das cadeias laterais ionizáveis dos resíduos do sítio ativo de TcRpiB, foram feitas predições dos valores de pKas dessas cadeias utilizando o programa Propka (141). A partir dos resultados obtidos com o Propka e de acordo com o mecanismo de reação descrito para TcRpiB, MtRpiB e EcRpiB (70,78), as cadeias laterais ionizáveis referentes ao receptor foram ajustadas aos estados de protonação adequados em pH 7. Essa protonação foi realizada em cada um dos três PDBs de TcRpiB utilizando o programa *Protein Preparation Wizard* do Maestro, no caso do Glide, e o Chimera (142), no caso do AutoDock Vina.

Para o ensaio de atracamento com o programa AutoDock Vina, todos os arquivos PDB da TcRpiB, após terem seus estados de protonação assinalados, foram submetidos à adição de cargas Gasteiger e salvos no formato pdbqt (formato utilizado pelo AutoDock Vina) através do programa MGLtools (138).

Já a preparação dos receptores para o atracamento molecular com o programa Glide foi feita utilizando o *Protein Preparation Wizard* através do Maestro, no qual, além de terem sido gerados os estados de protonação adequados, foram adicionados ao receptor átomos de hidrogênio e assinaladas as ordens de ligação.

O modelo estrutural de HsRpiB foi utilizado em ensaios de atracamento com o programa Glide, sendo, portanto, preparado com *Protein Preparation Wizard* conforme descrito anteriormente. Além disso, as moléculas de água e ligantes não foram considerados na preparação de proteína. Os estados de protonação dos resíduos foram atribuídos em pH 7, usando a configuração de protonação descrita por Capriles e colaboradores (55) para os resíduos catalíticos com cadeias laterais ionizáveis.

### **3.4.3 Redocking e avaliação dos programas de atracamento AutoDock Vina e Glide XP**

Para os ensaios de atracamento molecular dois programas foram escolhidos AutoDock Vina (139) e Glide (precisão XP) (140). O programa AutoDock Vina utiliza como algoritmo de busca um método denominado *Iterated Local Search Global Optimizer* para busca estocástica global, em uma junção do método de Monte Carlo e algoritmos evolutivos (139,143). A função de pontuação do AutoDock Vina combina características de funções baseados no conhecimento e empíricas (139). No programa Glide, a busca conformacional se aproxima de uma busca sistemática completa dos graus de liberdade rotacional, translacional e conformacional do ligante, em um processo semelhante a construção incremental (1.8.2.2). Uma filtragem heurística visa eliminar rapidamente conformações que não são adequadas a ligação com o receptor, com eliminação também dos confôrmeros de alta energia. Ao final, as melhores conformações são refinadas usando o método de Monte Carlo. Já a função de pontuação do programa Glide (precisão XP), diferentemente do AutoDock Vina, é apenas empírica com os termos da função representando as propriedades físico-químicas teórica das interações proteína-ligante.

Nesse estudo, conforme representado na Figura 3.1, uma etapa de *redocking* com o receptor TcRpiB foi realizada a fim de avaliar o desempenho dos programas de



atracamento AutoDock Vina (139) e Glide (precisão XP) (140) e testar o efeito de diferentes preparações do receptor na reprodutibilidade do modo de ligação do ligante redocado em relação à pose do ligante cristalográfico. Esses procedimentos auxiliaram na escolha do programa de atracamento a ser utilizado na etapa de triagem virtual baseada em atracamento e também na determinação das melhores condições operacionais para o desenvolvimento dessa etapa.

Cada uma das três preparações do receptor TcRpiB, referentes à presença de água no sítio ativo, correspondeu a um procedimento separado de *redocking* em ambos os programas de atracamento molecular. Todos os procedimentos de *redocking* foram realizados com grid centralizada no ligante cristalográfico. Em todos os casos a grid incluiu, além de outros resíduos, todos os resíduos do sítio ativo de TcRpiB: Asp10, His11, Phe12, Tyr46, Cys69, Gly70, Ser71, Ile73, Gly74, His102, Asn103, Arg113, Arg141, Arg137 e His138.

A partir do conjunto de compostos baseado em similaridade (item 3.3.3) foi realizado um ensaio de atracamento molecular com os programas AutoDock Vina e Glide XP levando em consideração as três condições de solvatação do sítio ativo. Portanto, para cada um dos programas, três ensaios atracamento foram realizados equivalentes aos ensaios de *redocking*.

Além disso, uma outra etapa de avaliação do programa Glide XP foi realizada utilizando a proteína TcRpiB com ausência de moléculas de águas conservadas no sítio ativo. Nesse procedimento dois ensaios de atracamento separados foram realizados com os conjuntos de *decoys* (compostos inativos) obtidos no servidor DUD-E e DecoyFinder, juntamente, com o atracamento dos controles positivos e dos substratos de TcRpiB (compostos ativos). A partir dos resultados, foi possível construir as curvas ROC (*receiver operating characteristic*) – gráfico referente a taxa de verdadeiros positivos (sensibilidade) contra taxa de falsos positivos (especificidade) – desses ensaios e os cálculos da AUC (área sob curva) utilizando a biblioteca ROCR com o programa RStudio. Esse tipo de avaliação permite determinar se o programa de atracamento é capaz de distinguir compostos ativos de inativos. Quanto maior o valor da AUC (valor máximo 1), mais acurado o protocolo de atracamento é na determinação de verdadeiros positivos.

#### **3.4.4 Triagem virtual baseada em atracamento molecular**

Uma triagem virtual baseada em atracamento molecular com o programa Glide XP foi realizada utilizando a biblioteca de compostos gerada nesse estudo a fim de identificar potenciais inibidores de TcRpiB. Essa abordagem permitiu também a predição das interações intermoleculares de cada ligante da biblioteca em relação a proteína alvo de TcRpiB. Para realizar essa triagem, o protocolo XP foi escolhido por ter um maior desempenho de enriquecimento e predição de afinidade de ligação do que outros protocolos de ancoragem no Glide. Além disso, o Glide XP é descrito como sendo o melhor na eliminação de falsos positivos (140). Nessa triagem, apenas a proteína TcRpiB com ausência de moléculas de águas no sítio ativo foi utilizada. As configurações da grid seguiram como descrito na seção 3.4.3.

Nessa triagem virtual baseada em atracamento, os compostos mais bem ranqueados, isto é, com as melhores pontuações (*docking score* – DS), foram selecionados utilizando como valor de corte as pontuações dos controles positivos (4PEH (**41**), 4PEA (**38**) e a forma aberta e fechada de All6P (**15**)). Estes compostos mais bem ranqueados selecionados foram também atracados no sítio ativo de HsRpiA com o objetivo de eliminar compostos com elevada afinidade de ligação a esta proteína. Esse atracamento molecular foi realizado com a grid apresentando as seguintes coordenadas do centroide: X = 18,36; Y = 11,22 e Z = 89,57. A grid incluiu todos resíduos do sítio ativo de HsRpiA: Lys82, Thr107, Ser104, Gly105, Ser106, Asp160, Gly161, Ala162, Asp163, Lys173, Gly174, Gly175, Gly176, Gly177, Cys178, Glu182, Lys200 e Lys203.

### **3.5 Predição ADME-Tox e PAINS**

Os compostos mais bem ranqueados na triagem com a TcRpiB e com baixa afinidade de ligação com o sítio ativo de HsRpiB foram submetidos a predições de suas propriedades ADME-Tox (absorção, distribuição, metabolismo, excreção e toxicidade) e PAINS (*pan-assay interference compounds*) utilizando o servidor FAF-Drug3 ([fafdrugs3.mti.univ-paris-diderot.fr/](http://fafdrugs3.mti.univ-paris-diderot.fr/)) (144). O FAFDrug3 se baseia na aplicação de filtros ou regras, como as regras de cinco de Lipinsk, de Veber, de Egan<sup>5</sup>, GSK

---

<sup>5</sup> Regra de Egan:  $-1 \leq \log P \leq 5.8$  e  $PSA \leq 130 \text{ \AA}^2$ . Essa regra determina a biodisponibilidade oral de um candidato a fármaco (144).

4/400<sup>6</sup> e a Pfizer 3/75<sup>7</sup>, a fim de prever propriedades farmacocinéticas e toxicidade. Além disso, o FAFDrug3 realiza também a detecção de subestruturas, como potenciais grupos tóxicos, agregadores e PAINS. A análise de PAINS visa identificar compostos reativos que tendem a atuar como falsos positivos (144).

Dessa forma, compostos que apresentaram propriedades de ADME-Tox e PAINS insatisfatórios foram eliminados do grupo final de potenciais inibidores de TcRpiB. Essa etapa de filtragem ADME-Tox e PAINS é importante uma vez que muitos compostos apresentam bom desempenho na fase de validação experimental ou até mesmo nas fases com modelos animais, mas não conseguem ser aprovados nos testes clínicos. Isso se deve principalmente ao fato de que muitos compostos que chegam nessa etapa têm propriedades farmacocinéticas e toxicidade indesejáveis e a eliminação prévia deles é uma abordagem importante (145).

Além disso, para analisar se alguns dos compostos finais selecionados pertencem ao metabolismo humano foram feitas buscas, utilizando estes compostos como *query*, no *subset* do ZINC referente a moléculas endógenas de humanos.

### 3.6 Clusterização de compostos

O conjunto de compostos obtidos após a etapa de atracamento molecular e de predição de ADME-Tox e PAINS foi avaliado e clusterizado (subdividido) com o programa Scaffold Hunter 2.6.3 (146). Para cada composto foi calculado o *fingerprint* contendo 1024 *bits*. A partir disso, a clusterização hierárquica dos compostos foi realizada utilizando como critério de ligação de grupos o método Ward. Esse método agrupa dois grupos que resultam em um mínimo aumento de variância, dessa forma em cada etapa da clusterização dois grupos que apresentarem menor aumento global na variância dentro do grupo resultante são agrupados. A distância entre os compostos comparados foi obtida com o coeficiente de Tanimoto,  $d(a,b)$  (equação 3.1):

---

<sup>6</sup> Regra GSK 4/400:  $\log P > 4$  e peso molecular  $> 400$ . Candidatos a fármacos que se encaixam nessa regra apresentam maiores riscos de toxicidade, interações com *off-targets* ou dificuldades durante as etapas de desenvolvimento (144).

<sup>7</sup> Regra Pfizer 3/75:  $\text{ClogP} < 3$  e  $\text{PSA} > 75 \text{ \AA}^2$ . Essa regra afirma que candidatos a fármaco com os parâmetros físico-químicos descritos tem uma redução de seis vezes na toxicidade pré-clínica (ou uma redução de 24 vezes no caso de compostos básicos) (144).

$$d(a, b) = \frac{|a \wedge b|}{|a \vee b|} \quad 3.1$$

Em que  $d(a, b)$  é definido a partir dos *fingerprints*  $a = (\{0,1\}^1, \dots, \{0,1\}^n)$  e  $b = (\{0,1\}^1, \dots, \{0,1\}^n)$  representados como uma série de bits.

Além disso, para a avaliação da diversidade de arcabouços moleculares do conjunto de composto obtidos foi utilizado o algoritmo de *scaffold tree* do programa Scaffold Hunter. Esse algoritmo é capaz de computar, para um conjunto de compostos, as classificações hierárquicas baseadas nas características estruturais em comum desses compostos (146). Dessa forma, cada composto do conjunto final foi associado com um arcabouço molecular único que foi obtido pelo corte de todos as cadeias laterais terminais preservando-se apenas as duplas ligações diretamente ligadas a um anel. Em seguida, cada arcabouço molecular foi podado por um conjunto de regras determinísticas de uma forma gradual, removendo um único anel por vez. Estas regras determinísticas baseiam-se em considerações estruturais com o objetivo de preservar o “núcleo” estrutural mais característico, de forma que ao final o arcabouço molecular de cada composto consistiu de um único anel (146).

### **3.7 Simulação por dinâmica molecular**

Conforme já descrito em seções anteriores, o ensaio de atracamento molecular realizado nesse estudo consistiu em considerar os ligantes flexíveis e a estrutura da proteína alvo rígida. No entanto, a atividade de uma enzima é eventualmente associada com modificações tridimensionais no decorrer do tempo da reação. Para ser eficiente na sua função inibitória, um fármaco tem de interagir de forma estável com os aminoácidos do sítio de ligação ao longo do tempo (147). Para permitir esse tipo de avaliação mais robusta, ensaios de DM foram realizados em colaboração com o Dr. João Hermínio Martins da Silva (Grupo de Modelagem Computacional, FIOCRUZ- Ceará). Esses ensaios foram realizados com TcRpiB em complexo com os três compostos mais bem ranqueados no atracamento molecular (apenas compostos aprovados em todas as etapas de triagem foram considerados).

A partir disso, as estruturas dos compostos foram submetidas ao Automated Topology Builder (ATB) and Repository (148) para construir topologias adequadas para a DM. A DM foi realizada usando o programa Gromacs 5.1.2 (149), que utiliza

por padrão algoritmo *leap-frog* como integrador, e o campo de força Gromos 96. O seguinte tratamento foi feito para preparar os sistemas para as etapas de produção da DM: (i) minimização de energia no vácuo, usando o método *steepest descent* (2 ps); (ii) minimização de energia em água (10 ps); (iii) simulação DM em *ensemble* NVT (100 ps); (iv) simulação DM em *ensemble* NpT com restrições (100 ps); (v) simulações DM no *ensemble* NpT (1 ns). O sistema foi inicialmente submetido a minimização usando o método *steepest descent*. As minimizações foram completadas quando a tolerância de 1000 kJ / mol não foi excedida. Após estes passos, foram geradas trajetórias de 50 ns para cada complexo. As simulações foram realizadas utilizando condições periódicas de contorno, a 300 K e 1 atm de pressão. A eletrostática foi avaliada utilizando-se o algoritmo particle mesh Ewald (150). Todas as análises de  $\beta$ -fator, flutuação quadrática média (RMSF), ligações de hidrogênio foram realizadas usando o pacote de software Gromacs, Maestro e Visual Molecular Dynamics (VMD) v1.9.3.41.

## 4 RESULTADOS E DISCUSSÃO

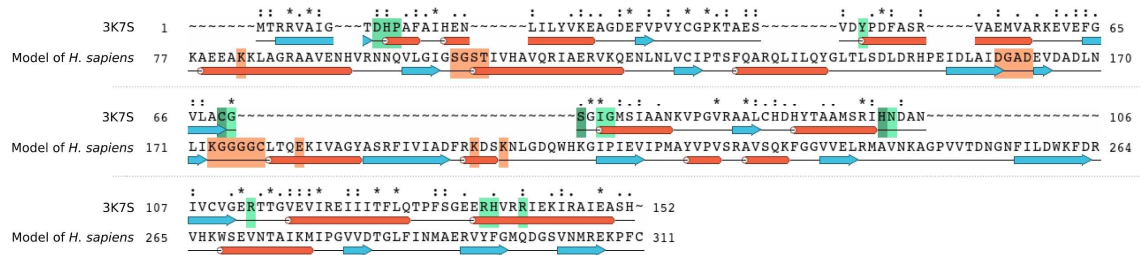
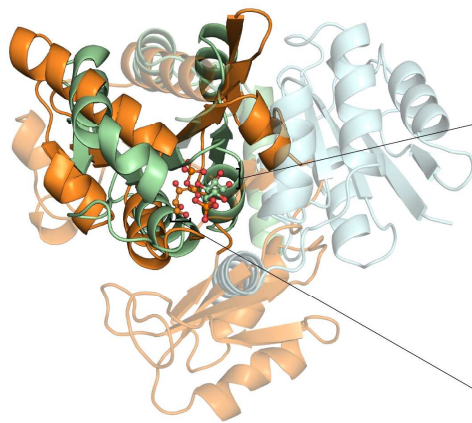
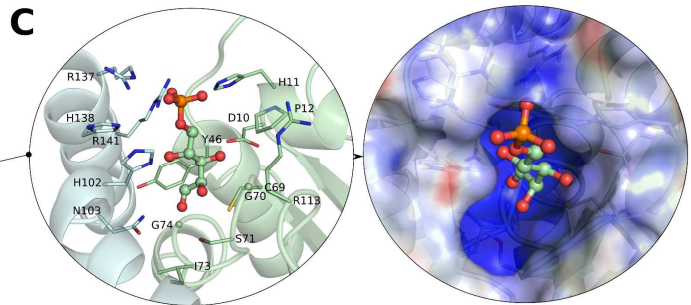
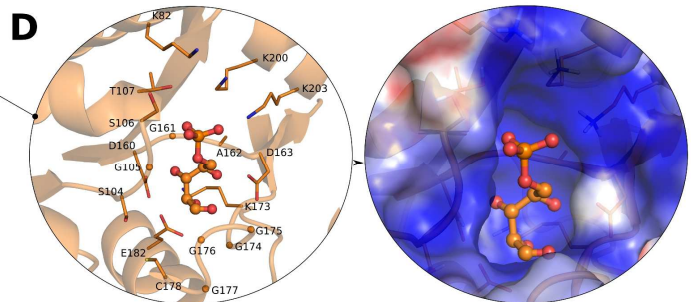
### 4.1 Análises de proteínas

#### 4.1.1 Comparação das proteínas TcRpiB e HsRpiA

Uma análise comparativa estrutural da proteína do patógeno, no caso o *T. cruzi*, em relação à proteína humana correspondente é uma etapa extremamente relevante na descoberta de fármacos. Esta análise é necessária para avaliar a possibilidade de um inibidor em potencial atuar também na proteína do hospedeiro, desencadeando efeitos colaterais. Embora, no caso de enzimas análogas, essa possibilidade de interação seja reduzida (devido a estrutura tridimensional distinta ou pouco similar), uma avaliação das semelhanças e diferenças entre essas proteínas ainda é necessária, já que ambas enzimas interagem com os mesmos substratos e catalisam a mesma atividade enzimática podendo, portanto, apresentar características em comum.

A comparação das estruturas primária, secundária e terciária de HsRpiA e TcRpiB revelou diferenças entre essas proteínas, como esperado devido à analogia (Figura 4.1). O alinhamento global das sequências das proteínas HsRpiA e TcRpiB mostrou uma identidade e similaridade de 8% e 21% (Figura 4.1A), respectivamente. É importante notar que o alinhamento da sequência não conseguiu alinhar os resíduos do sítio ativo de HsRpiA e TcRpiB, incluindo os resíduos catalíticos. Além disso, a sobreposição estrutural realizada com o TM-align entre a estrutura cristalográfica de TcRpiB (PDB 3K7S) e o modelo construído para HsRpiA não foi bem sucedida, sugerindo uma possível sobreposição de proteínas não relacionadas (Figura 4.1B). Isto é evidenciado pelo valor obtido de TM-score de 0,36065. Segundo essa métrica, apenas as sobreposições estruturais entre proteínas com valores de pontuação superiores a 0,5 são indicativas de serem sobreposições de proteínas evolutivamente relacionadas, contendo geralmente o mesmo enovelamento no SCOP/CATH (132). Estes resultados corroboram a hipótese de analogia entre HsRpiA e TcRpiB, mostrando que, apesar de apresentarem a mesma função enzimática (provavelmente devido a um fenômeno de convergência evolutiva), suas sequências e estruturas tridimensionais globais são diferentes.

Uma comparação visual da arquitetura de ambos os sítios ativos revelou a presença de muitos resíduos aromáticos (Tyr46, His11, His102 e His138) em TcRpiB, enquanto que HsRpiA não apresenta nenhum (Figura 4.1C e D) (55). Além disso, a distribuição de resíduos com potencial de apresentarem cargas negativas (a presença da carga vai depender do pH do meio e do pKa da cadeia lateral do resíduo em questão) é diferente; TcRpiB possui a Cys69 e Asp10, e a HsRpiA possui o Asp160, Asp163 e Glu182. Apesar destas diferenças, considerando o substrato R5P (**7**) como ponto de referências entre TcRpiB e HsRpiA (Figura 4.1C e D), ambos os sítios ativos dessas enzimas apresentam resíduos básicos próximo do grupamento fosfato de R5P, altamente eletronegativo (His11, His102, Arg113, His138, Arg137 e Arg141 em TcRpiB, e Lys82, Lys200 e Lys203 em HsRpiA). Esta característica comum a TcRpiB e HsRpiA reflete nas semelhanças de superfície eletrostática potencial observadas nos sítios ativos dessas proteínas, caracterizados por superfícies predominantemente eletropositivas (Figura 4.1C e D). Mesmo com a analogia descrita, estas semelhanças convergentes, em particular nos sítios ativos de HsRpiA e TcRpiB, eram esperadas, uma vez que essas enzimas catalisam a mesma reação e, portanto, interagem com o mesmo substrato. Estes resultados suportam a demanda por estudos de atracamento molecular não apenas com TcRpiB, mas também com HsRpiA, visando a identificação de compostos capazes de inibir a enzima do parasita, mas com pouca ou nenhuma atividade em relação a HsRpiA.

**A****B****C****D**

**Figura 4.1:** Comparação entre as estruturas primárias, secundárias e terciárias de HsRpiA e TcRpiB. (A) Alinhamento de seqüências das proteínas HsRpiA (Hsa: 22934) e TcRpiB (Tcr: 509199.24) utilizando o programa Maestro / Clustalw. Os resíduos do sítio ativo de TcRpiB estão realçados em verde e os resíduos catalíticos estão representados em verde escuro. Os resíduos do sítio ativo de HsRpiB estão representados em laranja. As estruturas secundárias estão representadas por setas azuis (folha  $\beta$ ), cilindros vermelhos ( $\alpha$ -hélice) e linhas pretas (loops); (B) Sobreposição estrutural baseada nas similaridades estruturais entre a estrutura cristalográfica de TcRpiB (PDB 3K7S, cadeias A/B em verde/azul) e o modelo de HsRpiA (cadeia A em laranja). A região de sobreposição entre estas estruturas está realçada em laranja escuro e verde escuro; (C) R5P no sítio ativo de TcRpiB; (D) R5P no sítio ativo de HsRpiA. Os substratos R5P de TcRpiB e de HsRpiA estão representados em *ball and stick* com carbonos coloridos em verde e laranja, respectivamente. As cadeias laterais de resíduos para ambos os sítios ativos estão representadas como *sticks*, exceto as glicinas que estão representadas como bolas pequenas. Os potenciais de superfície eletrostática dos sítios ativos também estão representados (vermelho é -25 kBT e azul é 25 kBT) com uma superfície transparente. As figuras das estruturas das proteínas foram geradas usando Pymol v1.8.2.3.

#### 4.1.2 Análise das moléculas de águas conservadas no sítio ativo da estrutura cristalográfica TcRpiB (PDB 3K7S)

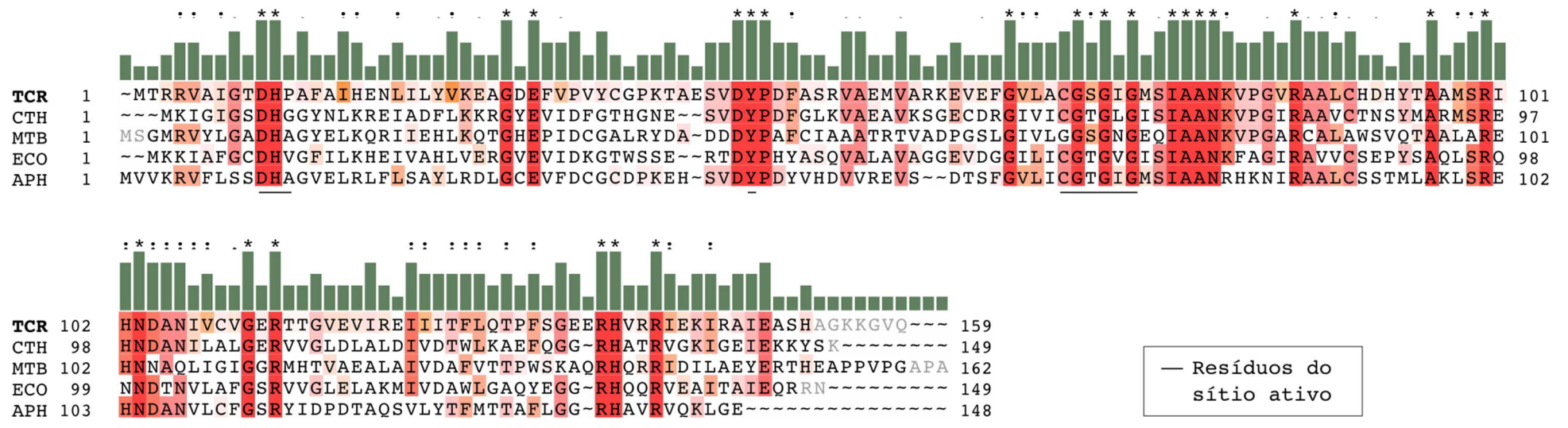


As moléculas de água conservadas no sítio ativo da proteína alvo TcRpiB (PDB 3K7S) foram determinadas pela comparação com os sítios ativos de outras estruturas cristalográficas de enzimas RpiBs disponíveis no PDB e com similaridade maior que 50% a TcRpiB. Entre as 18 estruturas cristalográficas selecionadas, cinco pertenciam ao *T. cruzi* (3K7S, 3K70, 3K8C, 3M1P e 3K7P), seis a *Mycobacterium tuberculosis* (1USL, 2BES, 2BET, 2VVP, 2VVQ e 2VVO), duas a *Escherichia coli* (2VVR e 1NN4), quatro ao *Clostridium thermocellum* (3HE8, 3HEE, 3PH3 e 3PH4) e uma ao *Anaplasma phagocytophilum* (4EM8) (Tabela 4.1). O alinhamento entre as sequências primárias, referentes a essas enzimas nos diferentes organismos analisados, evidenciou a conservação entre os resíduos pertencentes ao sítio ativo dessas enzimas (Figura 4.2). Dessa forma, a conservação entre esses sítios ativos torna mais precisa as inferências sobre as moléculas de águas conservadas e com funções importantes na interação entre ligante e enzima. Moléculas de água com essas características são normalmente denominadas moléculas de águas estruturais.

**Tabela 4.1:** Estruturas cristalográficas de RpiBs utilizadas nas análises de moléculas de águas estruturais no sítio ativo de TcRpiB

Organismo	PDB ID	Resolução (Å)	Referência
<i>T. cruzi</i>	3K7S	1,90	Stern <i>et al.</i> , 2011 (54)
	3K70	2,00	
	3K8C	2,10	
	3M1P	2,20	
<i>M. tuberculosis</i>	3K7P	1,40	Roos <i>et al.</i> , 2004 (69) Roos <i>et al.</i> , 2005 (76) Roos <i>et al.</i> , 2008 (78)
	1USL	1,88	
	2BES	2,10	
	2BET	2,20	
	2VVP	1,65	
	2VVQ	2,00	
<i>E. coli</i>	2VVR	2,10	Roos <i>et al.</i> , 2008 (78)
	1NN4	2,20	Zhang, <i>et al.</i> , 2009 (151)
<i>C. thermocellum</i>	3HE8	1,90	Jung, <i>et al.</i> , 2011 (68)
	3HEE	2,00	
	3PH3	2,07	
	3PH4	2,07	
<i>A. phagocytophilum</i>	4EM8	1,95	-

A análise da estrutura tridimensional dos sítios ativos sobrepostos (Tabela 4.2) indicou que as moléculas de água 176A, 175A, 180A, 188A, 204B e 222B (nomenclatura referentes à estrutura de PDB 3K7S) se mantiveram conservadas em mais que 60% das estruturas e formaram um grupo denominado de águas conservadas (AC). Dentro do grupo de águas conservadas, as moléculas de água 176A, 175A, 180A e 188A foram as mais frequentes, aparecendo em mais que 80% dos casos, formando um subgrupo denominado águas muito conservadas (AMC). A seleção do grupo e subgrupo de águas conservadas tem especial importância nos estudos de atracamento molecular, podendo otimizar os seus resultados, já que essas moléculas de águas conservadas podem mediar às interações entre o ligante e a proteína. Observando-se apenas as distâncias em Å entre as moléculas, dentre as moléculas de águas selecionadas, as moléculas de águas 176A, 180A e 222B provavelmente realizam ligações de hidrogênio tanto com o substrato, R5P (7; Apêndice B- Fig. 7.2), quanto com os resíduos do sítio ativo da proteína, representando moléculas de águas estruturais. Dessa forma, o oxigênio da água 176A possui uma distância de 2,8Å e 3,1Å, respectivamente, com um dos oxigênios de R5P e com um dos nitrogênios da cadeia lateral da Arg141. Já o oxigênio da água 180A apresenta 2,6Å de distância em relação ao nitrogênio da cadeia lateral da Asn103 e 3,1Å de distância em relação a um dos oxigênios do substrato. No caso da água 222B, é observada a distância de 3,0Å entre seu oxigênio e um dos oxigênios de R5P e a distância de 3,1Å desse oxigênio da água em relação a um dos nitrogênios da cadeia lateral da His102. É interessante ressaltar que ligações de hidrogênio podem apresentar distâncias entre o acceptor e o doador de ligação de hidrogênio de 2.5Å até 3.7Å, mas o que determina que a interação seja de fato efetiva é o ângulo formado entre o acceptor, o hidrogênio e o doador de ligação de hidrogênio (que pode ir de 90° até 180°) (125).



**Figura 4.2:** Alinhamento múltiplo entre as sequências de aminoácidos da ribose 5-fosfato isomerase de *T. cruzi* (TCR), de *M. tuberculosis* (MTB), de *E. coli* (ECO), de *C. thermocellum* (CTH) e de *A. phagocytophilum* (APH). Foi utilizado para o alinhamento o ClustalW2 através do programa Maestro. Os resíduos foram realçados em variações de vermelho indicando a porcentagem de identidade no alinhamento. O gráfico consenso está representado em barras verdes. Os resíduos pertencentes ao sítio ativo também foram identificados.

**Tabela 4.2:** Análise das moléculas de águas conservadas no sítio ativo das estruturas cristalográficas com similaridade maior que 50% com a proteína alvo TcRpiB (PDB 3K7S).

Org.	PDB ID	Ligante	Ident.	Simil.	RMSD em relação à 3K7S (Å)	Moléculas de água da TcRpiB (PDB 3K7S)								
						427A	344A	195A	176A	175A	180A	188A	204B	222B
TCR	3K7S	R5P	100	100	0	+	+	+	+	+	+	+	+	+
	3K70	-	100	100	0,17	-	-	-	+	+	+	+	+	-
	3K8C	4PEH	100	100	0,17	-	-	-	+	+	+	+	+	+
	3M1P	ALL6P	100	100	0,18	-	-	-	+	+	+	+	+	-
	3K7P	Pi	100	100	0,19	-	-	-	+	+	-	+	+	-
MTB	1USL	Pi	37	53	0,82	-	+	-	+	-	+	+	-	+
	2BES	4PEH	37	53	0,8	-	+	+	+	+	+	+	-	+
	2BET	4PEA	37	53	0,8	-	+	+	+	+	+	+	-	+
	2VVP	R5P	37	53	0,78	-	+	+	+	+	+	+	+	+
	2VVQ	5PRA	37	53	0,81	-	+	+	+	+	+	+	+	+
	2VVO	ALL6P	37	53	0,84	-	+	+	+	+	+	+	-	-
ECO	2VVR	-	40	63	0,89	-	-	-	+	+	-	+	-	-
	1NN4	Pi	41	63	0,96	-	-	-	+	+	+	+	+	-
CTH	3HE8	Glicerol	44	65	0,56	-	-	+	+	+	+	+	+	+
	3HEE	R5P	44	65	0,54	-	+	+	+	+	+	+	+	+
	3PH3	D-ribose	44	65	0,57	-	-	-	+	+	+	+	+	+
	3PH4	D-alose	44	65	0,56	-	-	+	+	+	+	+	+	+
APH	4EM8	-	41	58	1,35	-	-	-	+	+	-	+	-	
Frequência						1/18 (5,5%)	8/18 (44%)	9/18 (50%)	18/18 (100%)	17/18 (94%)	15/18 (83%)	18/18 (100%)	12/18 (66%)	11/18 (61%)

Abreviações: Org, organism; Ident, identidade; Simil, similaridade; Pi, fosfato; TCR, *T. cruzi*; MTB, *M. tuberculosis*; ECO, *E. coli*; CTH, *C. thermocellum*; APH, *A. phagocytophilum*.

## 4.2 Modelagem de farmacóforo

### 4.2.1 Modelagem de farmacóforo baseado em ligante

Para prever as características químicas essenciais - e a organização espacial destas - que um composto deve possuir para interagir com a proteína TcRpiB, foi realizada a modelagem de farmacóforo baseada em ligante e baseada em estrutura. Na modelagem de farmacóforo baseada em ligante realizada com o programa Pharmagist, um total de 40 hipóteses de farmacóforo (conforme mostrado na Tabela 4.3) foram geradas utilizando o conjunto de inibidores descritos para RpiB – 4PEH (**41**), All6P (aberta, **15**), All6P (fechada), 4PEA (**38**), 5PRAm (**53**) e 5PRMA (**55**). A partir disso, três hipóteses foram selecionadas – LB1, LB6 e LB7 (representadas na Figura 4.3). Estas hipóteses foram selecionadas seguindo os seguintes critérios: i) utilizarem na sua construção todos os inibidores indicados anteriormente; ii) possuírem pelo menos um ponto farmacofórico com carga negativa entre os seus pontos farmacofóricos. Esse critério é importante de ser considerado, já que o sítio ativo de RpiBs é caracterizado por possuir uma natureza eletropositiva; e iii) recuperar no banco de dados ZINC, através do servidor ZINCPHARMER, os substratos de Rpi - R5P (**7**) (cadeia aberta e forma de anel) e Ru5P (**10**).

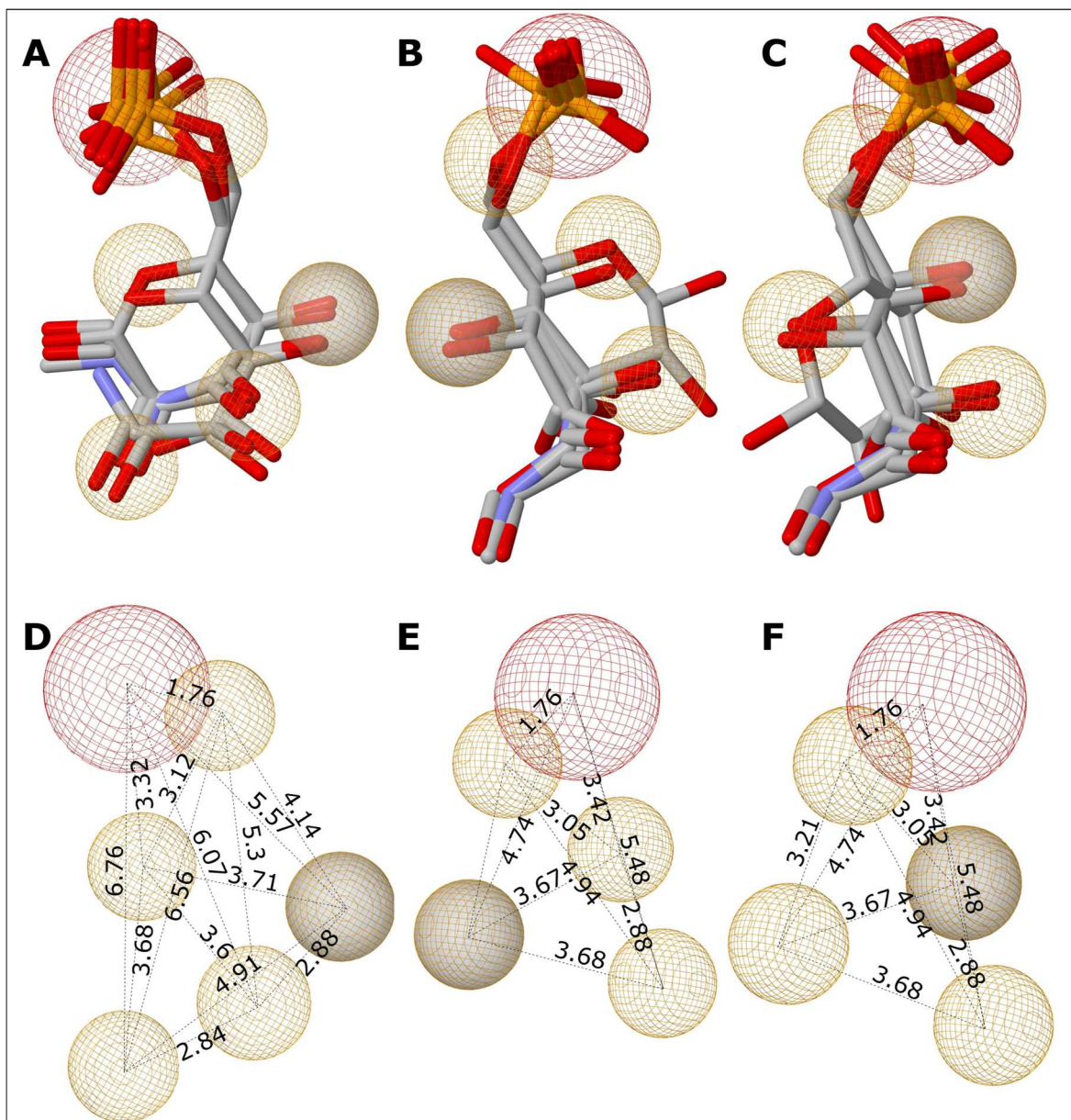
A hipótese LB1 se caracterizou por usar como ligante de referência (*pivot*) o inibidor ALL6P (**15**) em sua forma fechada e por possuir, além do ponto farmacofórico carregado negativamente (N), cinco aceptores de ligação de hidrogênio (HA) e um doador de ligação de hidrogênio (HD) (Figura 4.3A e D). Já as hipóteses LB6 e LB7 se caracterizaram por terem como ligante de referência o inibidor 5PRMA (**55**) e por possuírem um ponto farmacofórico N, quatro HA e um HD. Ambas as hipóteses diferem entre si apenas pela posição do ponto farmacofórico doador de ligação de hidrogênio, como pode ser visto nas Figura 4.3B, C, E e F.

**Tabela 4.3:** Hipóteses de farmacóforo baseadas em ligante geradas a partir do conjunto de inibidores selecionados de RpiB

Hipót.	Pont.	A	H	HD	HA	N	P	Compostos
LB1	26.062	0	0	1	5	1	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB2	23.625	0	0	2	5	0	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB3	22.5	0	0	0	5	0	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB4	22.274	0	0	2	5	0	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB5	21.744	0	0	1	5	0	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB6	21.562	0	0	1	4	1	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB7	21.562	0	0	1	4	1	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB8	21.562	0	0	1	4	1	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB9	21	0	0	0	4	1	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB10	21	0	0	0	4	1	0	oAll6P, rAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB11	25.102	0	0	2	5	1	0	oAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB12	25.102	0	0	2	5	1	0	oAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB13	25.102	0	0	2	5	1	0	oAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB14	24.572	0	0	2	4	1	0	oAll6P, rAll6P, 4PEH, 5PRAm e 5PRMA
LB15	20.86	0	0	2	4	1	0	oAll6P, rAll6P, 4PEH, 5PRAm e 5PRMA
LB16	20.86	0	0	2	4	1	0	oAll6P, rAll6P, 4PEH, 5PRAm e 5PRMA
LB17	20.86	0	0	2	4	1	0	oAll6P, rAll6P, 4PEH, 5PRAm e 5PRMA
LB18	20.86	0	0	2	4	1	0	oAll6P, 4PEH, 4PEA, 5PRAm e 5PRMA
LB19	20.329	0	0	2	3	1	0	oAll6P, rAll6P, 4PEH, 5PRAm e 5PRMA
LB20	19.512	0	0	3	3	1	0	oAll6P, rAll6P, 4PEH, 5PRAm e 5PRMA
LB21	27.45	0	0	3	5	1	0	oAll6P, 4PEH, 5PRAm e 5PRMA
LB22	23.481	0	0	3	4	1	0	oAll6P, rAll6P, 5PRAm e 5PRMA
LB23	23.481	0	0	3	4	1	0	oAll6P, 4PEH, 5PRAm e 5PRMA
LB24	23.481	0	0	3	4	1	0	oAll6P, 4PEH, 5PRAm e 5PRMA
LB25	23.481	0	0	2	5	1	0	oAll6P, rAll6P, 5PRAm e 5PRMA
LB26	21.739	0	0	3	4	1	0	oAll6P, 4PEH, 5PRAm e 5PRMA
LB27	21.331	0	0	3	5	0	0	oAll6P, rAll6P, 5PRAm e 5PRMA
LB28	20.835	0	0	3	4	0	0	rAll6P, 4PEH, 5PRAm e 5PRMA

Hipót.	Pont.	A	H	HD	HA	N	P	Compostos
LB29	20.339	0	0	2	4	0	0	oAll6P, 4PEH, 5PRAm e 5PRMA
LB30	20.008	0	0	3	4	1	0	oAll6P, 4PEH, 5PRAm e 5PRMA
LB31	22.199	0	0	4	4	1	0	oAll6P, 5PRAm e 5PRMA
LB32	22.199	0	0	3	5	1	0	oAll6P, 5PRAm e 5PRMA
LB33	22.199	0	0	3	5	1	0	oAll6P, 5PRAm e 5PRMA
LB34	21.739	0	0	3	4	1	0	oAll6P, rAll6P e 5PRAm
LB35	21.739	0	0	2	5	1	0	oAll6P, rAll6P e 4PEH
LB36	19.749	0	0	4	4	0	0	rAll6P, 5PRAm e 5PRMA
LB37	19.29	0	0	3	4	0	0	oAll6P, 4PEH e 5PRAm
LB38	19.29	0	0	3	4	0	0	rAll6P, 4PEH e 5PRAm
LB39	19.29	0	0	3	4	0	0	oAll6P, rAll6P e 4PEH
LB40	18.83	0	0	1	5	0	0	oAll6P, rAll6P e 4PEA

Abreviaturas: Hipót., hipóteses; Pont, pontuação LB, hipótese de farmacóforo baseado em ligante; A, aromático; H, hidrofóbico; HD, doador de ligação de hidrogênio; HA, aceptor de ligação de hidrogênio; N, carga negativa, P, carga positiva; oALL6P, forma de cadeia aberta de D-alose-6-fosfato; rALL6P, forma em anel de D-alose-6-fosfato; 4PEH, ácido 4-fosfo-D-eritrono-hidroxâmico; 4PEA, 4-fosfo-D-eritronato; 5PRAm, 5-fosfo-D-ribonamida; 5PRMA, N- (5-fosfo-D-ribonoil)-metilamina.



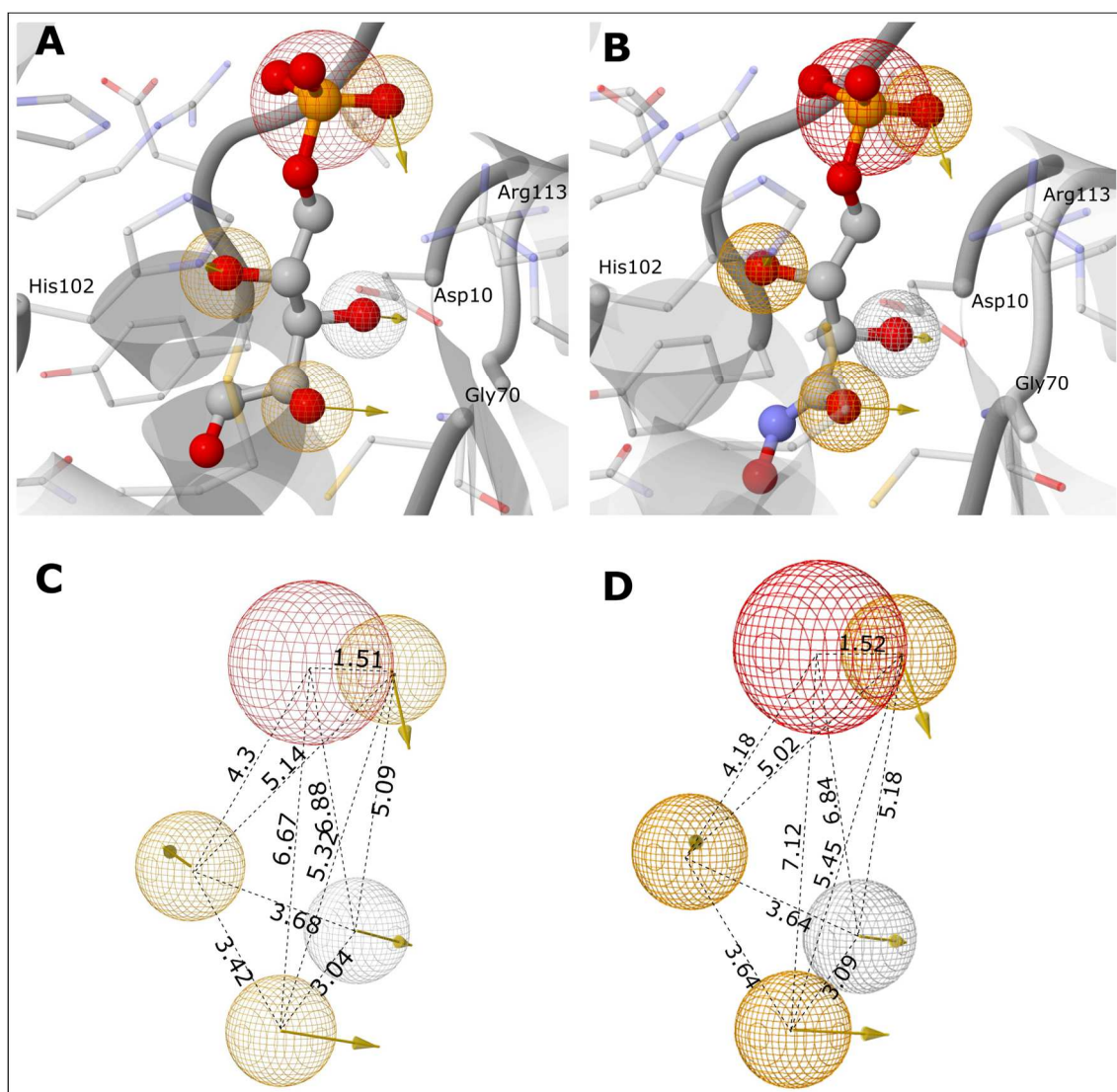
**Figura 4.3:** Hipóteses de farmacóforo baseadas em ligante. (A), (B) e (C) hipóteses de farmacóforo LB1, LB6 e LB7; (D), (E) e (F) representações esquemáticas das hipóteses de farmacóforo com as distâncias entre os pontos farmacofóricos em Å. As esferas em vermelho representam os pontos farmacofóricos de carga negativa, em laranja representam os pontos farmacofóricos aceptores de ligação de hidrogênio e em cinza representam as sobreposições dos pontos farmacofóricos aceptores e doadores de ligação de hidrogênio.

#### 4.2.2 Modelagem de farmacóforo baseado em estrutura

As modelagens de farmacóforo baseadas nas estruturas de PDB 3K7S e 3K8C geradas com o programa ZINCPHARMER retornaram os pontos farmacofóricos existentes entre os complexos proteína-ligante. A partir disso, foram selecionados os pontos farmacofóricos considerados significativos e essenciais na interação molecular dos ligantes com a TcRpiB, sendo, portanto, avaliadas as características químicas em comum dos ligantes e o ambiente químico do sítio ativo. Duas hipóteses semelhantes



de farmacóforo baseadas em estrutura foram geradas, ambas contendo os mesmos pontos farmacofóricos (SB1 e SB2). Como se pode observar na Figura 4.4, os pontos farmacofóricos selecionadas foram: (i) um ponto farmacofórico carregado negativamente referente ao grupamento fosfato dos ligantes 4PEH (**41**) e R5P (**7**) que interagem eletrostaticamente com os resíduos do sítio ativo com carga positiva em potencial (His11, His102, Arg113, Arg137, His138 e Arg141); (ii) três aceptores de ligação de hidrogênio referentes a ligação de hidrogênio com a Arg113 (o oxigênio do grupamento fosfato), a His102 e a cadeia principal da Gly70; (iii) um doador de ligação de hidrogênio, tornando possível uma ligação de hidrogênio com o Asp10. Estes pontos farmacofóricos refletem características químicas relevantes que devem existir no ligante para interagir com o receptor TcRpiB e, conseqüentemente, são importantes para identificar compostos biologicamente ativos nas etapas de triagem virtual.



**Figura 4.4:** Hipóteses de farmacóforo baseadas em estrutura. (A) e (B) hipóteses de farmacóforo SB1 e SB2, respectivamente; (C) e (D) representações esquemáticas das hipóteses de farmacóforo com as distâncias entre os pontos farmacofóricos em Å. As esferas em vermelho representam os pontos farmacofóricos de carga negativa, em laranja representam os pontos farmacofóricos aceptores de ligação de hidrogênio, em branco representam os pontos farmacofóricos doadores de hidrogênio e em cinza representam as sobreposições dos pontos farmacofóricos aceptores e doadores de ligação de hidrogênio.

#### 4.2.3 Triagem virtual baseada em farmacóforo e similaridade 2D

Com base nas hipóteses de farmacóforo geradas (três hipóteses de farmacóforo baseadas em ligante e duas baseadas em estrutura), foi possível triar o banco de dados ZINC, através do servidor ZINCPHARMER, com o objetivo de recuperar moléculas apresentando os mesmos pontos farmacofóricos e arquitetura dos modelos de farmacóforo. Em todas as hipóteses selecionadas, os raios dos pontos farmacofóricos aceptores e doadores de ligação de hidrogênio foram de 1Å e

os de carga negativa foram de 1,5Å. Esse raio representa a o quão flexível a hipótese gerada será. Como resultado, foram obtidos 18.038 compostos (sem considerar duplicatas). Além disso, uma triagem baseada em similaridade estrutural (a fim de triar compostos com mais de 50% de semelhança com substratos de TcRpiB) foi realizada diretamente no banco de dados ZINC e recuperou 2.145 compostos, esse conjunto de compostos em particular foi utilizado na avaliação dos programas de atracamento AutoDock Vina e Glide XP, descrito no próximo item. Todos estes compostos juntos, obtidos tanto pela triagem baseada em farmacóforos quanto em similaridade estrutural, foram submetidos à geração de tautômeros e estado de ionização adequado. As 24.863 estruturas finais representaram a biblioteca de compostos deste estudo *in silico*.

### 4.3 Atracamento molecular

#### 4.3.1 Avaliação dos programas de atracamento AutoDock Vina e Glide XP

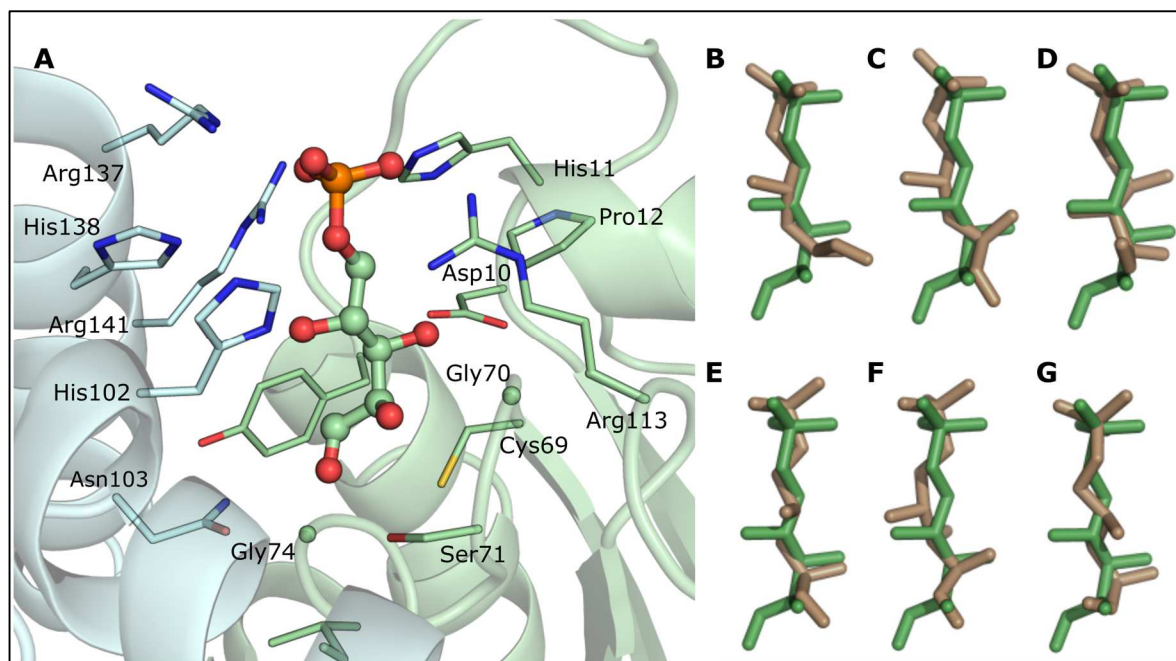
##### 4.3.1.1 Estudos de *Redocking* com TcRpiB utilizando os programas AutoDock Vina e Glide XP

Estudos de *redocking* com o receptor TcRpiB (PDB ID: 3K7S) foram realizados visando validar a precisão dos programas de atracamento AutoDock Vina e Glide XP e testar também as diferentes preparações do receptor na reprodutibilidade do modo de ligação do ligante redocado em relação à pose cristalográfica. Nesses estudos foram levadas em consideração as três condições de solvatação do sítio ativo: (i) ausência de moléculas de águas; (ii) presença de moléculas de águas conservadas (AC) e (iii) presença de moléculas de águas muito conservadas (AMC). Portanto, para cada um dos programas de atracamento, três ensaios *redocking* foram realizados.

O programa AutoDock Vina apresentou, de maneira geral, uma boa eficiência em reproduzir a pose cristalográfica do ligante, com o valor de RMSD referentes às poses de menor RMSD com valor médio de  $1,59 \pm 0,066$  Å. As diferentes preparações do receptor quanto à presença e à ausência de moléculas de águas cristalográficas conservadas no sítio ativo aparentemente apresentaram pouca influência nos resultados dos *redockings*, o que é evidenciado pelas variações menores que 0,1Å entre as médias dos RMSDs (3º coluna da Tabela 4.4). Apesar disso, pode-se observar que *redockings* (*redockings* 2 e 3) que envolveram a preparação do receptor com a presença de grupos de AC e AMC apresentaram resultados ligeiramente melhores, uma vez que, comparativamente, obtiveram os menores valores de energia de ligação, representando uma diminuição de até 0.9 kcal/mol de energia. Da mesma forma, as poses de menor energia desses *redockings* alcançaram os menores valores de RMSD, equivalentes a 1,57 Å e 1,84 Å para *redockings* envolvendo respectivamente AMC e AC contra o RMSD de 5,48Å para a pose de menor energia referente aos *redockings* em que não se considerou a presença das moléculas de água no sítio ativo (8º coluna da Tabela 4.4). É interessante ressaltar que no *redocking* envolvendo AMC a pose de menor energia corresponde também à pose de menor RMSD.

Já os resultados do *redocking* obtidos com o Glide XP indicam que este programa foi muito bem sucedido em alcançar poses similares a pose do ligante cristalográfico em todas as variações das condições do receptor quanto a presença de moléculas de águas cristalográficas no sítio ativo. Isso é evidenciado pelo baixo valor médio de RMSD das poses de menor RMSD ( $0,69 \pm 0,265\text{\AA}$ ). Além disso, em dois dos três *redockings* (moléculas de águas ausentes e AMC) a pose de menor pontuação também foi a pose de menor RMSD obtida. As três condições de *redocking* apresentaram poucas variações entre as médias dos RMSDs (3º coluna da Tabela 4.4) indicando a pouca influência das condições de solvatação do sítio ativo nos ensaios de *redocking* com o programa Glide XP. No entanto, o ensaio envolvendo a ausência de água no sítio ativo resultou na pose de menor RMSD ( $0,39\text{\AA}$ ) e nas poses de menor pontuação entre todas as poses geradas pelos ensaios de *redocking* com o Glide XP.

Quanto a comparação dos dois programas de atracamento molecular, o *redocking* com o programa Glide XP se mostrou muito mais acurado em reproduzir a pose cristalográfica do ligante do que o programa AutoDock Vina, como pode ser observado comparando os valores de RMSD entre todos os seis *redockings* realizados. Isso fica evidente, por exemplo, quando se compara a média total do RMSD das poses obtidas nos *redockings* com o Glide XP apresentando valor de  $1,37\text{\AA}$  contra  $3,69\text{\AA}$  obtido com o programa AutoDock Vina (3º coluna da Tabela 4.4). Na Figura 4.5 estão representadas as poses de menor RMSD (obtidas nos ensaios de *redocking* referentes as três condições diferentes com os dois programas de atracamento) em comparação com o ligante cristalográfico, assim como estão representadas também as ligações de hidrogênio estabelecidas entre o ligante cristalográfico R5P (7) e a proteína alvo TcRpiB.



**Figura 4.5:** (A) Pose do ligante cristalográfico R5P em relação à proteína alvo TcRpIB (PDB ID: 3K7S). Em (B), (C) e (D), observa-se em rosa a pose de menor RMSD para os *redockings* com ausência de moléculas de águas, com AMC e AC, respectivamente, obtidas com o programa AutoDock Vina. Em (E), (F) e (G), observa-se em rosa a pose de menor RMSD para os *redockings* com ausência de moléculas de águas, com AMC e AC, respectivamente, obtidas com o programa Glide XP. Em verde se destaca a estrutura do ligante cristalográfico R5P.

**Tabela 4.4:** Resultados dos *redockings* realizados utilizando diferentes preparações do receptor.

<b>Redocking AutoDock Vina</b>									
Ensaio	Águas cristalográficas no sítio ativo	Média dos RMSDs (Å)	Desvio Padrão (Å)	Pose de menor RMSD			Pose de menor energia de ligação		
				RMSD (Å)	EL (kcal/mol)	Posição	RMSD (Å)	EL (kcal/mol)	Posição
1	Ausente	3,63	1,28	1,66	-6,70	9°	5,48	-6,80	1°
2	AMC	3,70	1,69	1,57	-7,40	1°	1,57	-7,40	1°
3	AC	3,73	1,81	1,53	-7,10	6°	1,84	-7,70	1°
Média total		3,69	-	1,59	-7,07	-	2,96	-7,30	-

<b>Redocking Glide XP</b>									
Ensaio	Águas cristalográficas no sítio ativo	Média dos RMSDs (Å)	Desvio Padrão (Å)	Pose de menor RMSD			Pose de menor pontuação		
				RMSD (Å)	Docking Score	Posição	RMSD (Å)	Docking Score	Posição
1	Ausente	1,41	0,40	0,39	-10,24	1°	0,39	-10,24	1°
2	AMC	1,33	0,20	0,85	-7,424	1°	0,85	-7,424	1°
3	AC	1,38	0,24	0,85	-7,154	5°	1,49	-8,43	1°
Média total		1,37	-	0,69	-8,27	-	1,15	-8,69	-

**Abreviações:** EL, energia de ligação; DS- *docking score*; AMC, águas muito conservadas; AC, águas conservadas.

#### 4.3.1.2 Ensaios de atracamento com AutoDock Vina e Glide XP utilizando a TcRpiB e o conjunto de inibidores baseados em similaridade 2D

A fim de complementar os resultados dos estudos do *redocking* na avaliação das melhores condições do sítio ativo quanto à presença de moléculas de água cristalográficas e dos programas de atracamento, foram realizados ensaios de atracamento considerando as três condições do receptor TcRpiB: (i) com ausência de água, (ii) com presença de moléculas de águas conservadas (AC), e (iii) com presença de moléculas de águas muito conservadas (AMC). Nesse ensaio, o conjunto de compostos utilizado foi obtido por busca por similaridade 2D no banco de dados ZINC, conforme descrito no item 4.2.3, e apresentou ao todo 2.145 compostos. Nesse conjunto de compostos também foram incluídos como controles positivos os substratos da proteína alvo TcRpiB (R5P forma aberta e fechada (**7**) e a Ru5P(**10**)) e os seus inibidores conhecidos – ácido D-eritronohidroxiâmico (4PEH, **41**), alose 6-fosfato forma aberta e fechada (All6P; **15**) e 4-fosfo-D-eritronato (4PEA, **38**). Controles negativos não foram considerados.

Ao final de cada ensaio, os compostos foram classificados empregando como critério a função de pontuação própria de cada programa utilizado. As Tabela 4.5 e Tabela 4.6 correspondem, respectivamente, aos resultados desses ensaios de atracamento com os programas AutoDock Vina e Glide XP, nas quais são mostrados os três compostos de menor pontuação e os controles positivos referentes a cada um dos ensaios realizados. Já na Figura 4.6, é possível observar as interações moleculares entre o receptor TcRpiB e os mais bem classificados em cada um dos seis ensaios de atracamento (três referentes ao programa AutoDock Vina e três referentes ao Glide XP).

Comparando os resultados dos três ensaios de atracamento realizados com o programa AutoDock Vina, podemos observar que a presença das moléculas de águas conservadas e muito conservadas no sítio ativo influenciou na identificação de compostos com menores energias de ligação com a enzima receptora. Além disso, os controles positivos foram mais bem ranqueados e possuíram menores energias de ligação em ambos os ensaios envolvendo a presença de moléculas de águas. Isso fica bastante evidente no ensaio envolvendo AC, no qual o inibidor alose 6-fosfato ficou ranqueado entre os três compostos de menor energia e seis dos sete controles



positivos (incluindo os substratos) ficaram em posições menores ou iguais a 100. Tal fato mostra a importância da presença das moléculas de água selecionadas nesses ensaios realizados com o programa AutoDock Vina para um bom ranqueamento dos inibidores testados. Outro fato interessante a ser observado é a semelhança estrutural dos compostos melhores ranqueados em relação aos substratos naturais da enzima TcRpiB nos ensaios envolvendo moléculas de águas no sítio ativo. Nos casos de atracamento com ausência de água no sítio ativo, os compostos melhores colocados não apresentam tal semelhança, além de possuírem maior volume. Entre os três compostos mais bem ranqueados de cada um dos dois ensaios envolvendo a presença de água, foi identificado um composto em comum a ambos: ZINC03869779. Enquanto que no ensaio em que não envolveu a presença das moléculas de águas, não houve compostos mais bem ranqueados similares aos outros dois ensaios.

Mas é importante considerar que a escolha de uma condição envolvendo presença de moléculas de águas no sítio ativo no caso do AutoDock Vina pode representar um viés para compostos mais parecidos com os substratos, com menores volumes, conforme observado nos ensaios, (devido a maior restrição espacial do sítio ativo) ou que precisem das moléculas de água para realizar a interação em detrimento de compostos que não precisariam de tais moléculas de água ou que não caberiam no sítio por causa da presença destas moléculas. No entanto, com o AutoDock Vina as condições envolvendo a presença de água parece dar ao experimento um ambiente mais fidedigno a realidade do que ocorre no sítio ativo da TcRpiB e pode auxiliar na diminuição de falsos positivos. Isso fica evidente no caso de ensaios realizados por Lenselink e colaboradores (2014)(152) com o receptor adenosina A2A em que é mostrado que a presença de certas de moléculas de águas cristalográficas explícitas no sítio ativo em ensaios de triagem virtual baseado em estrutura aumenta significativamente o enriquecimento de compostos ativos em detrimento de *decoys*, além de mostrar que a otimização das posições dos hidrogênios das moléculas de águas são importantes a fim de se obter melhores resultados (Lenselink *et al.*, 2014)(152). Da mesma forma, Huang e Shoichet (2008) já haviam demonstrado que, em pelo menos metade dos alvos do conjunto de dados do DUD (banco de dados de *decoys*, antecessor do DUD-E), a presença de moléculas de águas cristalográficas em triagens virtuais é capaz de melhorar o enriquecimento do conjunto final de compostos (153). Além disso, estudo envolvendo a análise de aproximadamente 2300 estruturas cristalográficas de alta resolução de complexos, hidratados e não

hidratados, com compostos com  $K_i$  ou  $K_d$  relatados, mostraram que os compostos interagindo com moléculas de águas estruturais são tão potentes quanto aqueles interagindo somente com resíduos do sítio ativo (154). Isso realça a importância da consideração das moléculas de água em estudos de atracamento molecular.

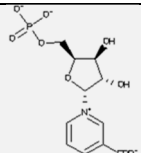
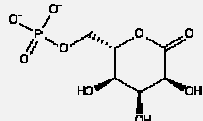
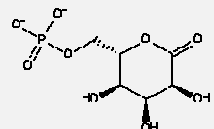
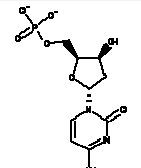
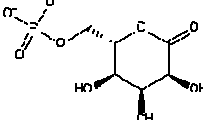
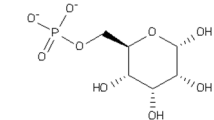

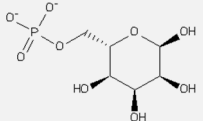
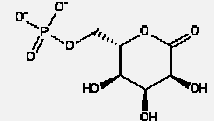
Entre os ensaios de atracamento realizados com o programa Glide XP, ao contrário daqueles com o AutoDock Vina, o ensaio envolvendo a ausência de moléculas de águas cristalográficas no sítio ativo obteve os melhores resultados, o que já era refletido pelos resultados de *redocking*. Nesse ensaio, em comparação com aqueles envolvendo a presença de moléculas de águas cristalográficas, os compostos mais bem ranqueados possuíam *docking scores* (DS) relativamente menores e, apesar de os substratos não estarem tão bem ranqueados, a maioria dos inibidores (controles positivos) foram comparativamente melhor posicionados e apresentaram em todos casos menores DS que nos outros ensaios. Comparando-se apenas os dois ensaios envolvendo a presença de moléculas de águas cristalográficas no sítio ativo, o que mais se destaca é aquele envolvendo a presença de moléculas de águas conservadas, em que além de os substratos e os inibidores estarem mais bem ranqueados, eles apresentam também DS mais baixos. Observando os compostos mais bem ranqueados de todos os ensaios com o Glide XP e comparando com os ensaios envolvendo o programa AutoDock Vina, a presença de compostos iguais que estejam similarmente bem ranqueados não é observada.

Os resultados distintos entre os programas AutoDock Vina e Glide XP, com relação a necessidade da presença de moléculas de águas conservadas no sítio ativo, evidenciam as abordagens distintas seguidas por cada programa e a importância de se preparar a proteína receptora levando essas diferenças em consideração. No programa AutoDock Vina, não é descrita nenhuma abordagem para lidar com as moléculas de água explícitas e com a solvatação do sítio ativo, de forma que se as moléculas de águas não forem incluídas explicitamente no ensaio, elas não serão consideradas. Isso pode explicar por que os ensaios com AutoDock Vina envolvendo a presença de moléculas de águas estruturais se destacam em relação aqueles com ausência. Já no caso do Glide XP, a sua função de pontuação envolve termos que levam em consideração os efeitos da dessolvatação e deslocamento de moléculas de água de regiões hidrofóbicas. Esse programa inclui para cada pose energeticamente competitiva do ligante uma etapa de atracamento de moléculas de águas explícitas no complexo proteína-ligante na região do sítio ativo. Nesse processo, descritores

baseados na interação dessas moléculas de água com vários grupos químicos do ligante e da proteína são usados como medida para saber se o complexo proteína-ligante e moléculas de água é fisicamente realístico. Dessa forma, penalidades advindas da dessolvatação podem ser contabilizadas. Essa estatística da pontuação de moléculas de águas também é utilizada para determinar se pontuações favoráveis para ligações de hidrogênio devem ser contabilizadas (140). Esse tratamento das moléculas de água permite ao programa rejeitar uma grande fração de falsos positivos, separar ligantes ativos de inativos e pode ser também a explicação do porque não houve a necessidade da presença de moléculas das águas cristalográficas estruturais no sítio ativo para se obter bons resultados nos ensaios da TcRpiB envolvendo o programa Glide XP (140).

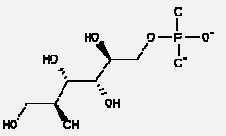
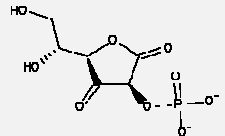
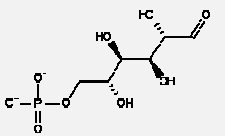
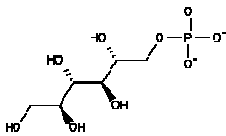
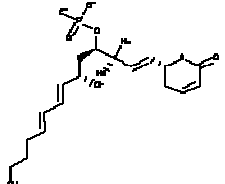
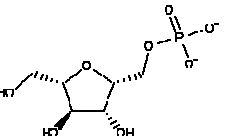
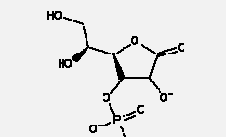
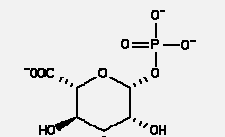
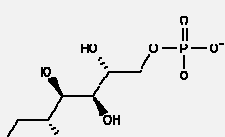
Os resultados obtidos e as abordagens descritas para cada um dos programas de atracamento molecular, quanto as moléculas de água, permitiram a escolha do programa Glide XP para as próximas etapas de triagem virtual. Ao ser comparado com outros programas de atracamento molecular quanto a acurácia de predição da pose de ligação e ao poder de estimar a afinidade de ligação, o Glide XP apresenta um dos melhores desempenhos na predição de poses, perdendo apenas para o programa Gold (155). No entanto, as funções de pontuação do AutoDock Vina e do Gold, ao serem correlacionadas as afinidades de ligação experimentais, foram melhores sucedidas. Contudo, mesmo nos melhores casos, a correlação entre os resultados das funções de pontuação e as afinidades de ligação experimentais foram baixas, indicando que as funções de pontuação dos programas atuais não são confiáveis o suficientemente para esse tipo de predição (155). Nesse estudo foi utilizado um conjunto diverso de estruturas de proteínas complexadas a ligantes. Como em estudos de atracamento molecular o tipo de proteína e as condições do estudo influenciam muito na performance do ensaio, uma análise do desempenho da função de pontuação do Glide XP em relação a TcRpiB foi realizada.

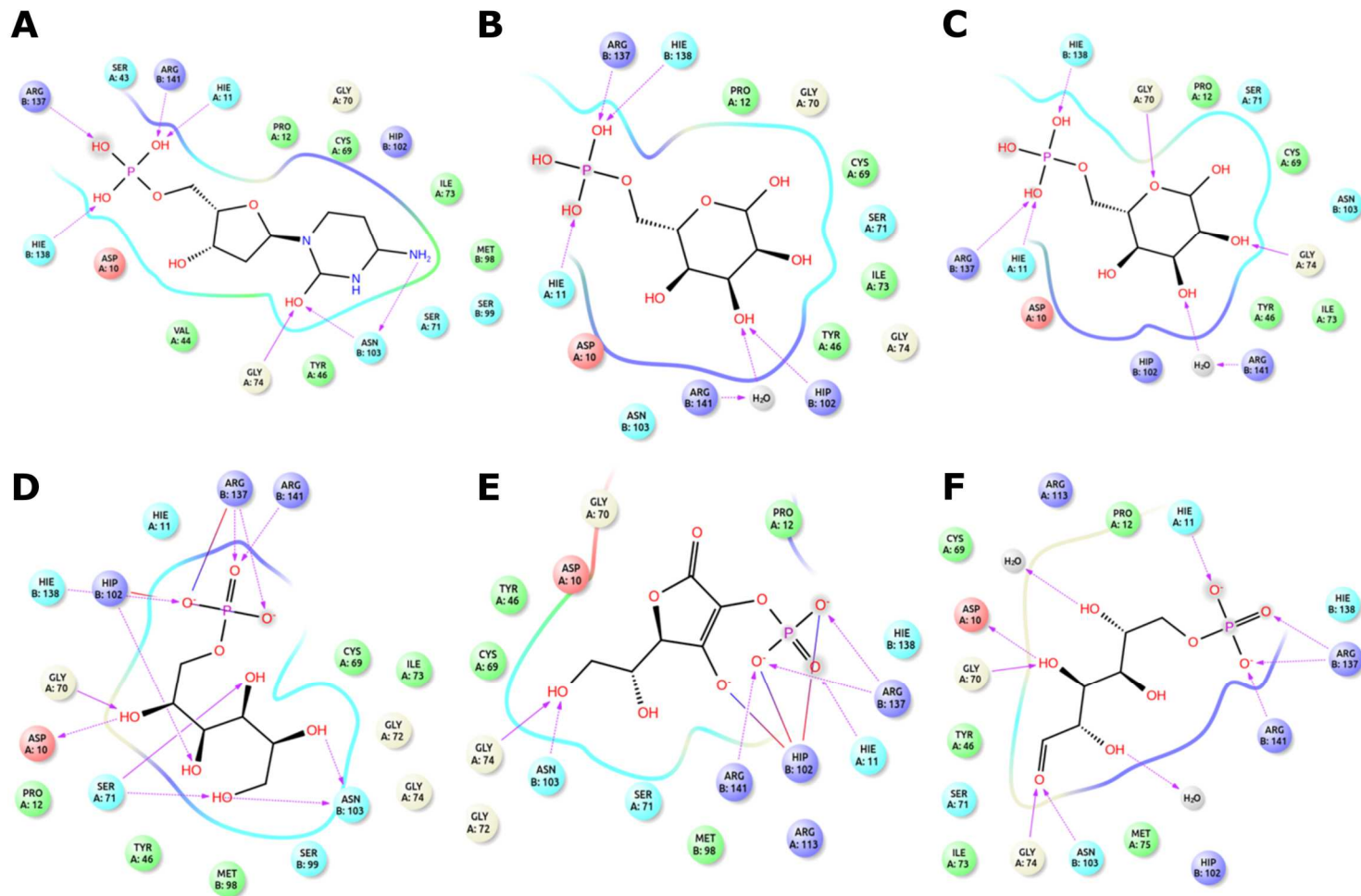
**Tabela 4.5:** Resultado do ensaio de atracamento com o programa AutoDock VINA utilizando o receptor em três condições em relação à presença de moléculas de água no sítio ativo.

Ensaio de atracamento <i>com</i> o AutoDock VINA usando diferentes condições do receptor								
Sem águas conservadas			Com águas muito conservadas			Com águas conservadas		
Composto	Posição	EL (kcal/mol)	Composto	Posição	EL (kcal/mol)	Composto	Posição	EL (kcal/mol)
 ZINC03870117	1°	-8,6	 ZINC03869779	1°	-8,9	 ZINC03869779	1°	-8,8
 ZINC03869817	2°	-8,6	 ZINC03869779_1	2°	-8,7	 All6P (fechada)	2°	-8,7
 ZINC00001812	3°	-8,4	 ZINC03869396	3°	-8,6	 ZINC03869779_1	3°	-8,6
R5P (fechada)	170°	-7,3	R5P (fechada)	111°	-7,6	R5P (fechada)	28°	-8,1
Ru5P	238°	-7,1	Ru5P	188°	-7,3	R5P (aberta)	93°	-7,7
R5P (aberta)	364°	-6,9	R5P (aberta)	314°	-7,0	Ru5P	103°	-7,6
4PEH	60°	-7,7	4PEH	8°	-8,2	4PEH	6°	-8,4
All6P (aberta)	218°	-7,2	All6P (aberta)	236°	-7,2	All6P (aberta)	71°	-7,8
All6P (fechada)	52°	-7,8	All6P (fechada)	7°	-8,3	All6P (fechada)	2°	-8,7
4PEA	505°	-6,7	4PEA	122°	-7,6	4PEA	100°	-7,7

**Abreviações:** EL, energia de ligação; AMC, águas muito conservadas; AC, águas conservadas.

**Tabela 4.6:** Resultado do ensaio de atracamento com o programa Glide XP utilizando o receptor em três condições em relação à presença de moléculas de água no sítio ativo

Ensaio de atracamento com o Glide XP usando diferentes condições do receptor								
Sem águas conservadas			Com águas muito conservadas			Com águas conservadas		
Composto	Posição	DS	Composto	Posição	DS	Composto	Posição	DS
 ZINC02522704	1°	-14,142	 ZINC18020695	1°	-11.806	 ZINC02508229	1°	-11,622
 ZINC03869810	2°	-13,054	 ZINC27308265	2°	-10,44	 ZINC44699073	2°	-10,386
 ZINC31460124	3°	-12,621	 ZINC25722529	3°	-10,25	 ZINC03869812	3°	-10,304
R5P (fechada)	62°	-10,974	R5P (fechada)	91°	-8,525	R5P (fechada)	27°	-9,367
Ru5P	266°	-9,407	Ru5P	137°	-8,016	Ru5P	148°	-8,172
R5P (aberta)	173°	-9,892	R5P (aberta)	275°	-7,050	R5P (aberta)	70°	-8,803
4PEH	139°	-10,152	4PEH	240°	-7,256	4PEH	249°	-7,414
All6P (aberta)	9°	-12,283	All6P (aberta)	31°	-9,235	All6P (aberta)	132°	-8,280
All6P (fechada)	248°	-9,484	All6P (fechada)	249°	-7,165	All6P (fechada)	19°	-9,658
4PEA	387°	-8,918	4PEA	335°	-6,655	4PEA	385°	-6,554



**Figura 4.6:** Interações moleculares 2D indicadas pelo MAESTRO entre o receptor TcRpiB e os mais bem classificados nos ensaios de atracamento com AutoDock Vina e Glide XP envolvendo ausência de água (A, D), com presença de moléculas de águas muito conservadas (B,E) e com presença de moléculas de águas conservadas (C, F). Os resíduos estão distinguidos em carga negativa (rosa), carregado positivamente (roxo), polar (azul) e hidrofóbico (verde). As ligações de hidrogênio estão representadas como setas roxas e as interações eletrostáticas estão representadas como linhas azuis / vermelhas.

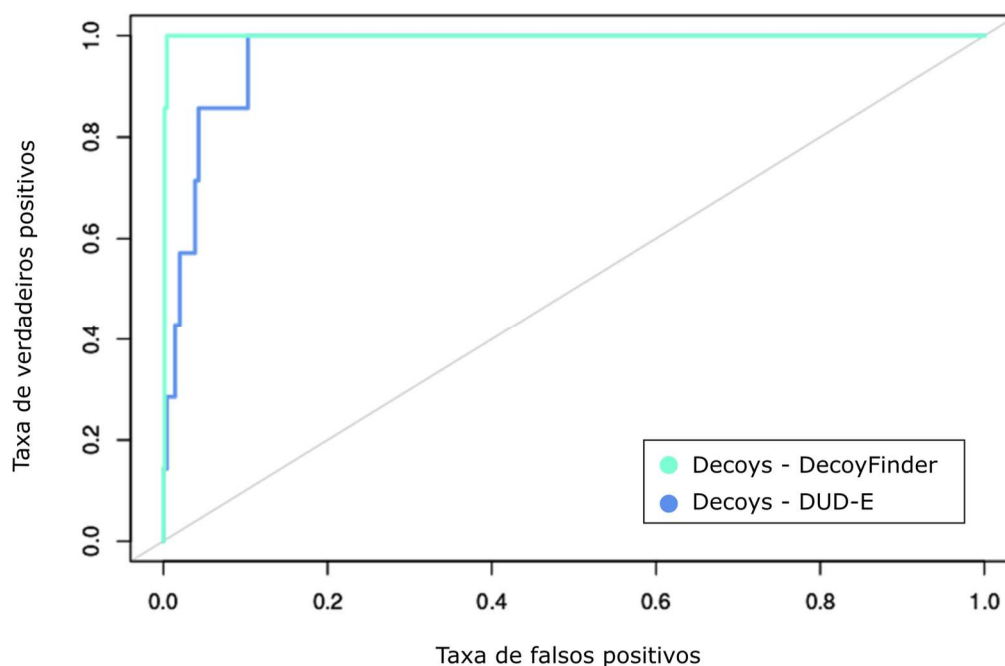
#### 4.3.2 Análise de desempenho do programa Glide XP com base na curva ROC

A partir da escolha do programa de atracamento molecular a ser utilizado nas etapas de triagem virtual e a condição do sítio ativo de TcRpiB mais adequada (condição com ausência de moléculas de águas cristalográficas no sítio ativo), uma análise do desempenho da função de pontuação do Glide XP foi realizada. O objetivo era determinar se essa função na condição escolhida era capaz de classificar melhor compostos ativos do que os inativos. Para isso foram feitas análises da curva ROC (*Receiver Operating Characteristic*) gerada utilizando um conjunto de compostos sabidamente ativos e dois conjuntos de *decoys*, compostos teoricamente inativos. A escolha desses dois conjuntos distintos de *decoys* se baseou no fato de que o resultado desse tipo de análise pode também ser influenciado pelos *decoys* utilizados, uma vez que a atividade real dos *decoys* geralmente é desconhecida.

A curva ROC permitiu mostrar a sensibilidade em relação a especificidade do método, ou seja, a taxa de verdadeiros positivos (ativos) versus a taxa de falsos positivos. A partir disso, o cálculo da área sob a curva (AUC) foi capaz de quantificar o desempenho da função de pontuação. Quanto maior o valor de AUC (valor máximo de AUC é 1), mais eficaz é a função de pontuação em discriminar compostos ativos de inativos. Valores de AUC acima de 0,8 são considerados bons e entre 0,9 a 1 são considerados excelentes. Uma AUC de 0,9, por exemplo, indica que, 9 em cada 10 compostos ativos selecionados aleatoriamente tem uma pontuação maior que um composto inativo randomicamente selecionado (156).

As duas curvas ROC (Figura 4.7), referentes aos compostos ativos em relação aos *decoys* do DUD-E e DecoyFinder, obtiveram valores de AUC de 0,96 e 0,99, respectivamente. Esse resultado indica que a função de pontuação do Glide XP utilizando a TcRpiB na condição com ausência de moléculas de águas cristalográficas é acurado em identificar compostos ativos como verdadeiros positivos, o que indica que essa abordagem tem grande potencial de identificação de compostos ativos em uma triagem virtual, conseguindo discriminar satisfatoriamente positivos de falsos positivos. A análise mais detalhada da curva ROC mostra que, no caso do teste com os *decoys* do DecoyFinder, 100% de compostos ativos haviam sido identificados com apenas 1% de falsos positivos. No caso do teste com os *decoys* do DUD-E, 85% dos compostos ativos já haviam sido identificados em 5% de falsos positivos e 100% com 20% de falsos positivos. Essa diferença entre resultados usando os dois testes

provavelmente se deve aos *decoys* do DUD-E terem propriedades físico químicas (por exemplo, peso molecular, LogP calculado, número de aceptores e doadores de ligação de hidrogênio) muito parecidas com os compostos ativos, apesar da topologia bidimensional dissimilar, enquanto que no caso do DecoyFinder apenas o peso molecular foi considerado como ponto de similaridade entre *decoys* e compostos ativos. Dessa forma, os *decoys* do DUD-E apresentam maiores “fatores de complicação” para a função de pontuação, fazendo com a chance de um composto inativo do DUD-E ser identificado como positivo (i.e falsos positivos) seja provavelmente maior do que os compostos obtidos com o DecoyFinder. No entanto, é importante destacar que mesmo com esse fator complicador, o desempenho da função de pontuação para o estudo em questão foi excelente (valor de AUC maior que 0,96).



**Figura 4.7:** Curva ROC do desempenho da função de pontuação do Glide XP em relação a proteína TcRpiB com ausência de moléculas de águas cristalográficas no sítio ativo.

### 4.3.3 Triagem virtual baseada em atracamento molecular

Estudos de triagem virtual baseada em atracamento molecular foram realizados permitindo prever a afinidade de ligação e as interações intermoleculares dos compostos da biblioteca *in silico* gerada (biblioteca com os ligantes baseados em farmacóforo e similaridade 2D) em relação ao sítio ativo, visando propor potenciais inibidores de TcRpiB. O programa Glide XP foi escolhido para a realização dos



ensaios de triagem virtual baseada em atracamento, uma vez que esse programa foi bem sucedido nos estudos de *redocking*, nos ensaios de atracamento utilizando a condição de TcRpiB com ausência de moléculas de águas e nas análises de curva ROC. A condição com a ausência de moléculas de águas nesse caso é interessante, pois o Glide XP leva em consideração as penalidades de dessolvatação e as interações ligante-solvente através de uma abordagem de adição de água baseada em grid (“atracamento” de moléculas de águas explícitas), sem a necessidade de que as moléculas de águas sejam mantidas (140). Além disso, essa condição aumenta a probabilidade de se obter uma maior diversidade estrutural no processo de triagem virtual de ligantes.

Dessa forma, todos os compostos obtidos a partir dos modelos de farmacóforo e pela busca de similaridade na base de dados de ZINC foram então atracados no sítio ativo de TcRpiB. O conjunto de três inibidores conhecidos de TcRpiB utilizados como controles positivos neste atracamento molecular – 4PEH (**41**) (obteve pontuação – Docking Score de -10,152), All6P de cadeia aberta (DS de -12,283), anel All6P (DS de -9,484) e 4PEA (**38**) (DS de -8,918) – foram importantes para determinar o valor de corte (*cut-off*) do DS de -8,918. Esse valor de corte baseado nos atracamentos com os controles positivos foi utilizado para determinar quais compostos (a partir dos valores de DS obtidos por esses compostos na triagem virtual) seriam classificados como potenciais inibidores de TcRpiB. A partir disso, os compostos que apresentam valores de DS iguais ou inferiores ao valor de DS do inibidor de 4PEA foram selecionados e considerados como potenciais inibidores para TcRpiB. Dessa forma, um total de 1.071 compostos, com pontuações menores ou iguais a pontuação do inibidor 4PEA (valor de corte), foram selecionados.

Estes compostos selecionados foram também atracados no sítio ativo de HsRpiA com o objetivo de identificar e eliminar compostos com potencial de inibir esta enzima também. Embora HsRpiA seja análogo ao TcRpiB, os sítios ativos destas proteínas têm algumas semelhanças (item 4.1.1), como a presença de vários resíduos com características básicas e as, conseqüentes, superfícies eletrostáticas predominantemente eletropositivas. Por esta razão, é importante assegurar a escolha de inibidores seletivos para o alvo de *T. cruzi*, com pouco ou nenhum potencial de inibir HsRpiA. Os controles positivos desse ensaio foram a forma linear de R5P (**7**), a forma de anel de R5P e o substrato Ru5P (**10**) que apresentaram os valores de DS de -8,973, -9,823 e -10,070, respectivamente. Nessa segunda triagem foram

selecionados os compostos que tinham valores de DS superiores a -8.973, uma vez que isto resultaria em um conjunto de compostos com baixo potencial de interagir com o sítio ativo de HsRpiA e, conseqüentemente, de inibir esta enzima. A partir disso, um total de 331 compostos foram selecionados como potenciais inibidores de TcRpiB e, provavelmente, com baixa ou nenhuma atividade inibitória em relação à enzima análoga humana.

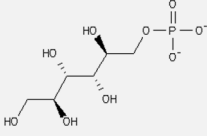
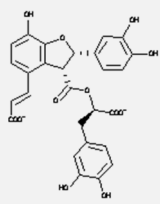
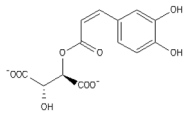
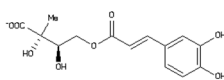
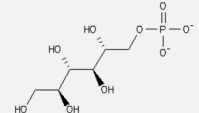
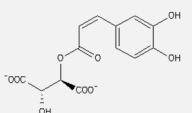
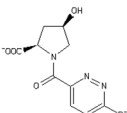
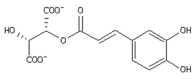
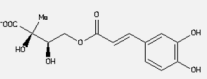
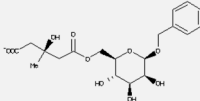
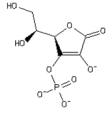
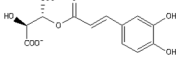
Na Tabela 4.7, são mostrados os seis primeiros colocados e os controles positivos da triagem virtual com a TcRpiB e HsRpiA, respectivamente. Como é possível observar, existiram compostos, principalmente com arcabouço molecular básico contendo catecol e grupamentos carboxila, com alta afinidade a HsRpiA, além da afinidade com TcRpiB. Inclusive um desses compostos, o composto ZINC33832844, ficou entre os mais bem ranqueados para ambas as enzimas. Outros casos também foram observados analisando-se os 20 primeiros colocados para essas enzimas e considerando também compostos com arcabouço molecular idêntico com mudança apenas na quiralidade, como é o caso dos compostos ZINC31168253 (5º colocado para TcRpiB) e ZINC31168249 (2º colocado para HsRpiA). Provavelmente algumas das características similares já descritas de ambos os sítios (item 4.1.1) foram importantes para que esses compostos tivessem alta afinidade predita por essas duas enzimas e, provavelmente, atuem como inibidores delas. É interessante ressaltar que nenhum composto até hoje foi indicado como inibidor de HsRpiA, apenas para outras RpiAs (58,72,87–90).

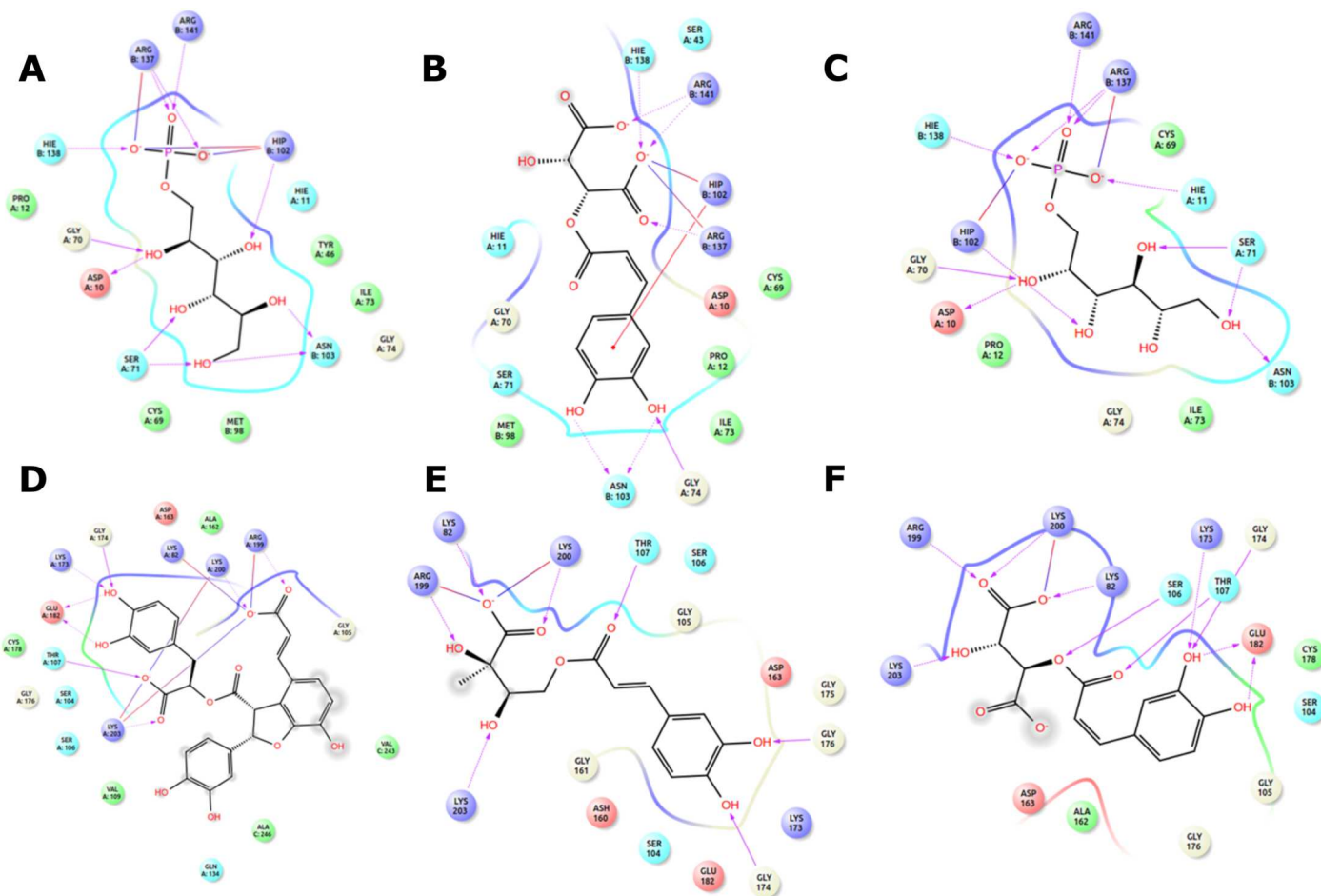
Todos os três primeiros colocados das duas triagens e suas interações com os sítios ativos das respectivas enzimas estão representados na Figura 4.8. Analisando, em particular, o composto ZINC33832844 (Figura 4.8B, Figura 4.8F e Figura 4.9), bem ranqueado em ambas as triagens, é possível observar padrões de interação semelhantes nos sítios de TcRpiB e HsRpiA, como por exemplo, as interação do tipo ligação hidrogênio e eletrostática de um dos grupos carboxilato com os aminoácidos de carga positiva e interação de uma das duas hidroxilas do catecol com glicina. Além dessas interações em comum, outras interações em particular em HsRpiA são observadas, como a interação do catecol também com Glu182 e Lys173, além das interações observadas com Gly174 e Gly176. Esse padrão de interação é frequente entre compostos contendo catecol e açúcares na forma de anel com a enzima de humano. Enquanto que os grupamentos de carga negativa dos compostos, como os grupamentos carboxila já citados, interagem com o núcleo positivo do sítio ativo

contendo principalmente lisinas. Já em TcRpiB, compostos com catecol ou açúcares na forma de anel interagem normalmente, além da Gly74, com a Asn103, enquanto os grupamentos negativos dos compostos interagem com as histidinas e argininas presentes nesse sítio (núcleo positivo).

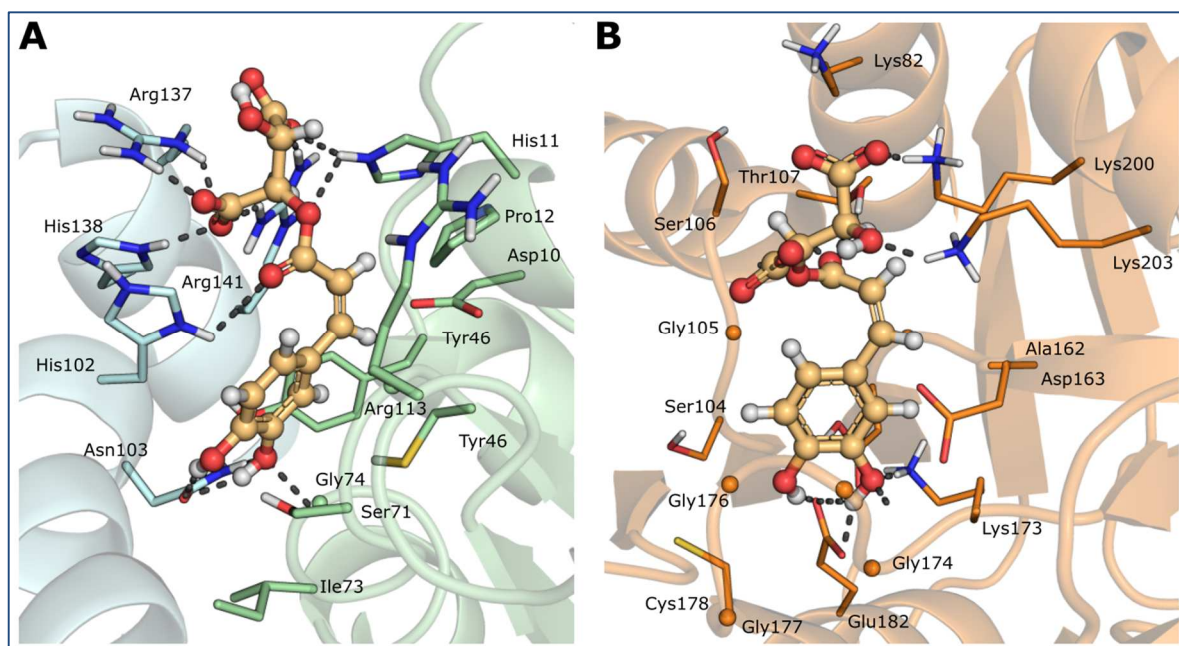
A diferença mais destoante entre as enzimas HsRpiA e TcRpiB é a possibilidade de a enzima do parasita ser capaz de fazer interações do tipo *stacking* devido a presença de resíduos com anéis aromáticos no sítio ativo (Tyr46, His11, His102 e His138), conforme já descrito anteriormente, além de possuir uma quantidade menor de aminoácidos carregados negativamente. Já o sítio ativo da enzima humana não possui anéis aromáticos (item 4.1.1). A triagem virtual com inclusão apenas de compostos que interagiram significativamente com TcRpiB, mas não com HsRpiA, aparentemente contribuíram para a presença de compostos com anéis aromáticos e sem o favorecimento daqueles com grupamentos positivos. A maioria desses compostos com anéis aromáticos realizou interações *pi-stacking* e/ou interações cation-*pi* com a histidina protonada (His102) presente no sítio ativo de TcRpiB. Nas interações *pi-stacking*, predominaram as configurações T (*T-shaped* ou *edge-to-face*). Estudos indicam que a conformação T (dependendo das condições do solvente) juntamente com a interação *pi-stacking* de anéis aromáticos com histidinas protonadas são energeticamente mais favoráveis (157,158). Além da contribuição dos resíduos aromáticos para a seletividade, a arquitetura e a disposição distinta dos aminoácidos nos sítios ativos de TcRpiB e HsRpiA parecem influenciar a seletividade de muitos compostos, que diferem, por exemplo, quanto as suas configurações de ligação de hidrogênio. Outro ponto importante é o sítio ativo de TcRpiB ser mais profundo, enquanto o de HsRpiA ser mais superficial e descrito como tendo um volume menor (55). Essa característica do sítio de HsRpiA contribui para uma exposição desfavorável ao solvente de alguns compostos potencialmente seletivos para a enzima do parasita. Esses resultados da triagem virtual com TcRpiB e HsRpiA reforçam que em estudos para o desenho racional de fármacos, mesmo em casos de analogia, a análise conjunta das duas enzimas deve ser realizada a fim de favorecer compostos que interajam somente com a enzima alvo.

**Tabela 4.7:** Resultado da triagem virtual baseada em atracamento molecular com o programa Glide XP para as proteínas TcRpiB e HsRpiA

Triagem virtual - TcRpiB			Triagem virtual - HsRpiA		
Composto	Posição	DS	Composto	Posição	DS
	1 <sup>o</sup>	-14,142		1 <sup>o</sup>	-13,282
ZINC02522704			ZINC67913694		
	2 <sup>o</sup>	-13,153		2 <sup>o</sup>	-13,186
ZINC33832844			ZINC31168249		
	3 <sup>o</sup>	-13,054		3 <sup>o</sup>	-13,036
ZINC03869810			ZINC33832844		
	4 <sup>o</sup>	-12,735		4 <sup>o</sup>	-12,935
ZINC36975961			ZINC02389524		
	5 <sup>o</sup>	-12,633		5 <sup>o</sup>	-12,902
ZINC31168253			ZINC31155215		
	6 <sup>o</sup>	-12,621		6 <sup>o</sup>	-12,889
ZINC31460124			ZINC14720637		
R5P (fechada)	-	-10,974	Ru5P	-	-10,070
Ru5P	-	-9,407	R5P (fechada)	-	-9,823
R5P (aberta)	-	-9,892	R5P (aberta)	-	-8,973
oALL6P	-	-12,283	oALL6P	-	-11,507
4PEH	-	-10,152	4PEH	-	-9,898
rALL6P	-	-9,484	rALL6P	-	-9,449
4PEA	-	-8,918	4PEA	-	-9,818



**Figura 4.8:** Interação 2D dos primeiros colocados nas triagens baseadas em atracamento molecular com o programa Glide XP para as proteínas TcRpiB e HsRpiA. (A), (B) e (C) são os três primeiros colocados para a triagem virtual baseada em atracamento com o receptor TcRpiB. (D), (E) e (F) são os três primeiros colocados para a triagem virtual baseada em atracamento com o receptor HsRpiA. Os resíduos estão distinguidos em carga negativa (rosa), carregado positivamente (roxo), polar (azul) e hidrofóbico (verde). As ligações de hidrogênio estão representadas como setas roxas e as interações eletrostáticas estão representadas como linhas azuis / vermelhas.



**Figura 4.9:** Modo de ligação do composto ZINC33832844, 2<sup>o</sup> e 3<sup>o</sup> colocado das triagens com TcRpiB e HsRpiA, respectivamente, no sítio ativo dessas enzimas. (A) Interação do composto ZINC33832844 no sítio ativo de TcRpiB; (B) (A) Interação do composto ZINC33832844 no sítio ativo de HsRpiA. O ZINC33832844 está representado em *ball and stick* em laranja na imagem principal, as subunidades A e B da proteína TcRpiB estão representadas em cores verde-clara e ciano-claro, respectivamente, a subunidade A de HsRpiA está representado em laranja e as interações de ligações de hidrogênio estão representadas como traços cinzas.

#### 4.4 Predição ADME-Tox e PAINS

Para determinar as propriedades farmacocinéticas, toxicidade potencial e para identificar também a presença de grupamentos indesejáveis ou compostos potencialmente promíscuos entre os 331 compostos previamente selecionados, predições de ADME-Tox e PAINS foram realizadas. Sendo assim, 211 compostos foram aprovados nessa triagem (Apêndice A- Tabela 7.1) e um total de 120 compostos foram classificados como intermediários ou rejeitados, representando moléculas que incluem alertas estruturais de baixo ou alto risco e/ou pertencem à classe PAINS.

A Tabela 4.8 resume os resultados das predições de ADME-Tox e PAINS obtidos para inibidores de TcRpiB conhecidos e para os dez compostos mais bem ranqueados que passaram por todas as triagens. Além disso, o conjunto de inibidores conhecidos para TcRpiB também foi avaliado quanto as suas propriedades farmacocinéticas e toxicidade. Desse conjunto, a forma aberta do inibidor All6P foi o único composto rejeitado na análise ADME-Tox e PAINS devido à presença de um

grupo aldeído classificado como sendo de alto risco. Como se pode observar na Tabela 4.8, os compostos identificados neste estudo apresentam melhores pontuações do que a maioria dos inibidores conhecidos de TcRpiB (valores de DS  $\leq$  -11,109) – principalmente 4PEH (**41**), 4PEA (**38**) e a forma em anel de All6P, sugerindo que estes novos compostos poderiam ser inibidores competitivos mais interessantes para a enzima do parasita. Além disso, os inibidores conhecidos de TcRpiB têm comparativamente valores menores de pontuação para HsRpiA (valores DS  $\leq$  -9,449) podendo, portanto, ser considerados como potenciais inibidores também para RpiAs de maneira geral, como já demonstrado para 4PEH e 4PEA em RpiA de espinafre (89). Por outro lado, os novos compostos propostos parecem ser mais específicos para RpiB, representando uma característica desejável em compostos líderes para o desenvolvimento de novos fármacos antichagásicos menos tóxicos para os seres humanos.

Como os compostos iniciais utilizados nas etapas de triagem foram provenientes do subconjunto *purchasable* do banco de dados ZINC, compostos encontrados no metabolismo humano também estão presentes nesse subconjunto. Dessa forma, entre os compostos finais selecionados como potenciais inibidores de TcRpiB, alguns compostos presentes no metabolismo humano também foram selecionados. Do total de 211 compostos finais selecionados, 42 compostos são encontrados no metabolismo humano ou apresentam arcabouço molecular idêntico a compostos encontrados nesse metabolismo, variando apenas quanto a quiralidade (Apêndice A- Tabela 7.1). Entre os compostos identificados como potenciais inibidores de TcRpiB e que se encaixam na categoria de metabolitos humanos estão nucleotídeos como o UDP (difosfato de uridina), GMP (monofosfato de guanosina, **58**) e CMP (monofosfato de citidina, **59**) e açúcares como a ribose (**17**), alose (**18**), manose (**22**), entre outros. Isso é interessante visto que alguns desses compostos já haviam sido indicados na literatura como sendo substratos ou inibidores de Rpis em geral. Esse é o caso das aldoses, alose e ribose, que atuam como substratos da RpiB de *L. donovani*, *Clostridium difficile*, *T. maritima* e *C. thermocellum* (77,79,80). Além disso, o nucleotídeo AMP (monofosfato de adenosina, **46**) já havia sido demonstrado ter atividade sobre EcRpA (87) e RpiB de *Rhodospirillum rubrum* como indicado no banco de dados Brenda (<https://www.brenda-enzymes.org/>). No caso de RpiA de *Cyberlindnera jadinii* (anteriormente chamada de *Candida utilis*), além do nucleotídeo AMP (**46**), os nucleotídeos UMP (**47**), GMP (**58**) e CMP (**59**) também foram capazes

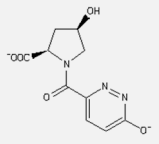
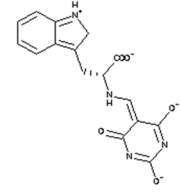
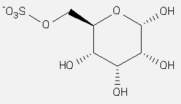
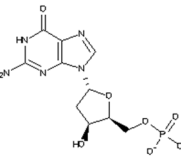
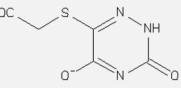
inibir essa enzima (91). Esses resultados são um bom indicativo de que o conjunto final de compostos selecionados está de fato enriquecido com inibidores potenciais de TcRpiB.

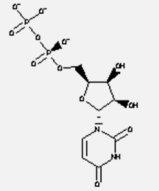
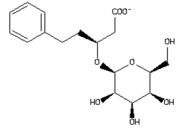
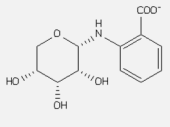
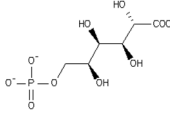
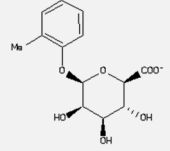
Entre os compostos finais selecionados para TcRpiB e ausentes no metabolismo de humanos, nenhum deles está entre os fármacos já aprovados por órgãos internacionais regulatórios, como o FDA. No entanto, o composto ZINC36521966 é análogo ao fármaco mizoribine (**60**), aprovado neste órgão. O mizoribine é um agente imunossupressor descrito por ter capacidade de inibir a síntese de DNA e apresentar a capacidade de inibição da inosina monofosfato desidrogenase e guanosine monofosfato sintetase (159).

Quanto aos compostos com propriedades farmacológicas conhecidas ou com mesmo arcabouço molecular de tais compostos (somente quiralidade distinta), estão, por exemplo, o ZINC00002053 e ZINC05085274 referentes a salicina (**61**) e o pirazofurin (**62**), respectivamente, entre outros (indicados no Apêndice A- Tabela 7.2). O pirazofurin, por exemplo, é um análogo de nucleosídeos que apresenta atividades antivirais, antibiótico e antitumorais e é descrito por potencialmente inibir a orotidine 5'-monophosphate (OMP) decarboxilase, impedindo a síntese de novo de uridina (160). Já a salicina, que é extraída do salgueiro, possui atividade anti-inflamatória e por ter estrutura química semelhante à da aspirina, possui ação parecida em humanos. Quando consumida, a salicina é metabolizada em ácido salicílico (161).



**Tabela 4.8:** Propriedades ADME previstas, filtros PAINS e DS dos controles positivos e dos dez compostos mais bem ranqueados para TcRpiB em relação a HsRpiA

Compostos	MW	logP	HD	HA	LV	OB VEBER	OB EGAN	PAINS Filter	DS TcRpiB	DS HsRpiA	Ligações H TcRpiB
 ZINC36975961	254,219	-2,03	2	8	0	Bom	Bom	Aceito	-12,735	-8,877	6
 ZINC63480117	341,298	-0,05	2	9	0	Bom	Bom	Aceito	-11,927	-8,503	8
 ZINC43763931	249,216	-3,9	4	9	0	Bom	Bom	Aceito	-11,849	-8,571	8
 ZINC03869847	347,221	-3,4	5	12	1	Bom	Bom	Aceito	-11,659	-8,465	6
 ZINC19300946	204,184	-1,08	3	7	0	Bom	Bom	Aceito	-11,601	-8,795	8

Compostos	MW	logP	HD	HA	LV	OB VEBER	OB EGAN	PAINS Filter	DS TcRpiB	DS HsRpiA	Ligações H TcRpiB
 ZINC08585028	402,145	-5,21	4	14	1	Bom	Bom	Aceito	-11,429	-8,625	8
 ZINC35456138	356,368	-0,18	4	8	0	Bom	Bom	Aceito	-11,41	-8,723	7
 ZINC34924542	269,251	-0,08	4	7	0	Bom	Bom	Aceito	-11,383	-8,671	7
 ZINC02047359	276,135	-4,46	4	10	0	Bom	Bom	Aceito	-11,314	-6,828	8
 ZINC35643607	284,262	0,37	3	7	0	Bom	Bom	Aceito	-11,109	-6,913	5
oALL6P	260,14	-4,53	6	9	1	Bom	Bom	Rejeitado	-12,283	-11,507	10
4PEH	231,100	-3,97	6	9	1	Bom	Bom	Aceito	-10,152	-9,898	8

Compostos	MW	logP	HD	HA	LV	OB VEBER	OB EGAN	PAINS Filter	DS TcRpiB	DS HsRpiA	Ligações H TcRpiB
rALL6P	260,14	-4,89	6	9	1	Bom	Bom	Aceito	-9,484	-9,449	7
4PEA	216,08	-3,23	5	8	0	Bom	Bom	Aceito	-8,918	-9,818	9

**Abreviações:** MW, peso molecular; LogP, coeficiente de partição octanol/água; HD, doador de ligação hidrogênio; HA, acceptor de ligação hidrogênio; LV, Violação de Lipinski; OB, biodisponibilidade oral; PAINS, *pan-assay interference compounds*; DS, *docking score*; ligações H, ligações de hidrogênio; oALL6P, forma de cadeia aberta de D-alose-6-fosfato; rALL6P, forma em anel de D-alose-6-fosfato; 4PEH, ácido 4-fosfo-D-eritrono-hidroxâmico; 4PEA, 4-fosfo-D-eritronato.

#### 4.5 Clusterização dos compostos

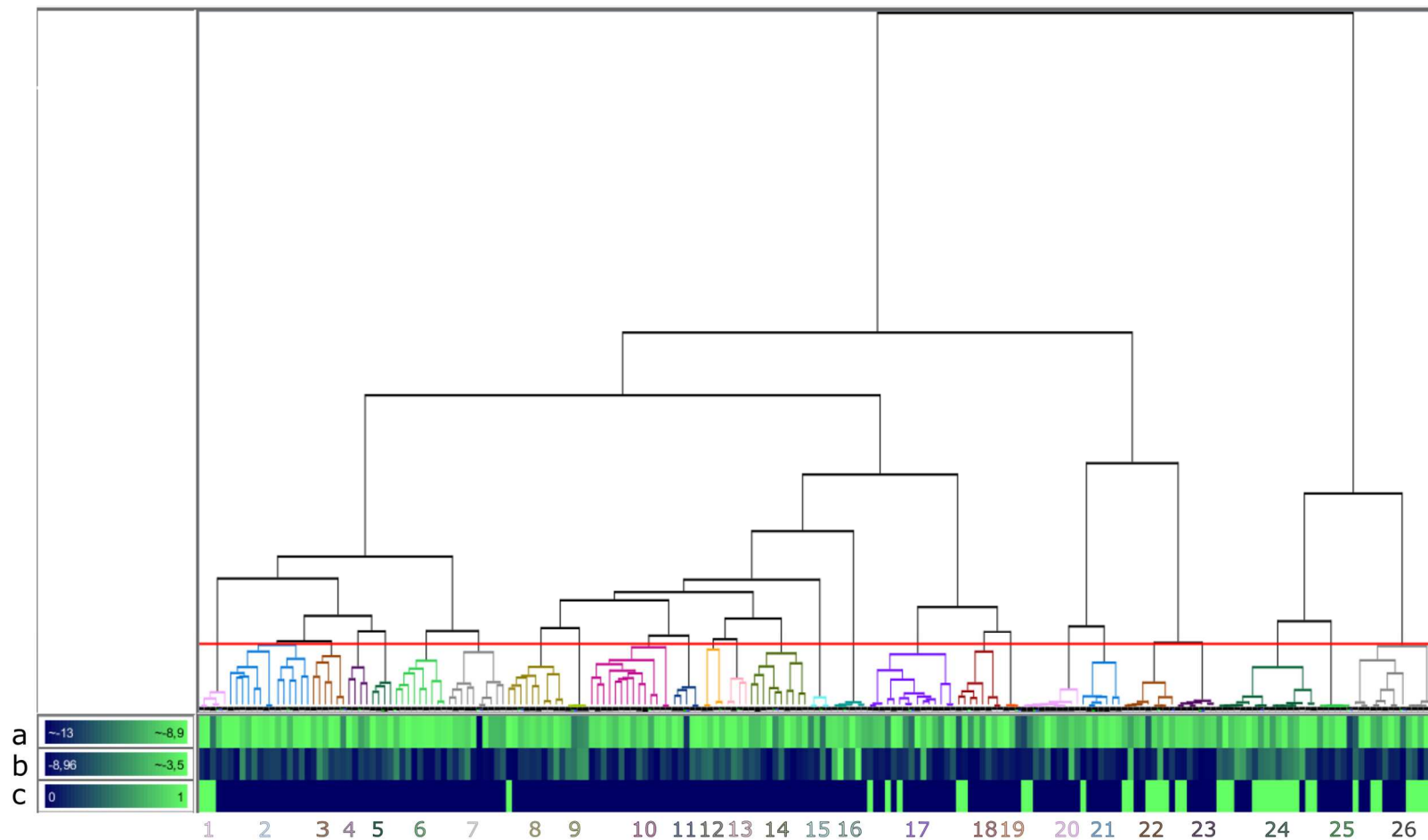
Geralmente, no final de uma triagem virtual, existem muitos compostos identificados que se assemelham em certo grau uns aos outros, de forma que não há a necessidade de seguir com todos os compostos para as próximas etapas mais custosas de avaliação, como as etapas de testes *in vitro*. Nestes casos, é interessante agrupar (clusterizar) a lista de resultados em grupos de compostos semelhantes e selecionar apenas um composto representativo de cada grupo para as etapas posteriores. Dessa forma, o conjunto de 211 compostos, selecionados (Apêndice A-Tabela 7.1) após a etapa de atracamento molecular e predição de ADME-Tox e PAINS, foi avaliado quanto a sua composição de arcabouços moleculares. Além disso, os compostos foram hierarquicamente clusterizados a partir das similaridades calculadas entre esses compostos, usando o coeficiente de Tanimoto. A clusterização, nesse estudo, é interessante para observar a diversidade de arcabouço molecular e auxiliar na escolha de compostos para etapas posteriores, como as etapas de dinâmica molecular e os testes *in vitro*. Dessa forma, será possível a seleção dos compostos mais diversos entre si e que sejam representativos de seus grupos.

O dendograma referente a clusterização dos 211 compostos está representado na Figura 4.10. Ao todo 26 grupos foram considerados, contendo de três a dezessete compostos cada. O *heat map*, juntamente com o dendograma, permitiu avaliar a distribuição dos valores de DS para TcRpiB e HsRpiA dos compostos participantes de cada grupo e se estes compostos são ou não metabólitos humanos. Os representantes de cada grupo estão indicados na Figura 4.11 e Figura 4.12. De todos os grupos, o grupo 24 apresentou o maior número de compostos (17 compostos), representando 9% do conjunto total. Nesse grupo, carboidratos foram os mais frequentes, entre eles as já citadas manose, ribose e alose. Grande parte desse grupo foi composto de metabólitos humanos. O grupo 17 foi o segundo maior grupo em número de compostos. Nesse grupo, compostos contendo ácido glicarônico (**63**) formando ligações glicosídicas com diversos grupos químicos (constituindo os chamados glicoranídeos) foram comuns, entre eles o próprio composto representativo desse grupo (Figura 4.12) e os metabólitos humanos, paracetamol glicoronídeo (**64**) e indoxil glicoronídeo (**65**). Além desse grupo, outro grupo menor e hierarquicamente próximo, o grupo 18, também apresentou glicoranídeos. Outros grupos com muitos

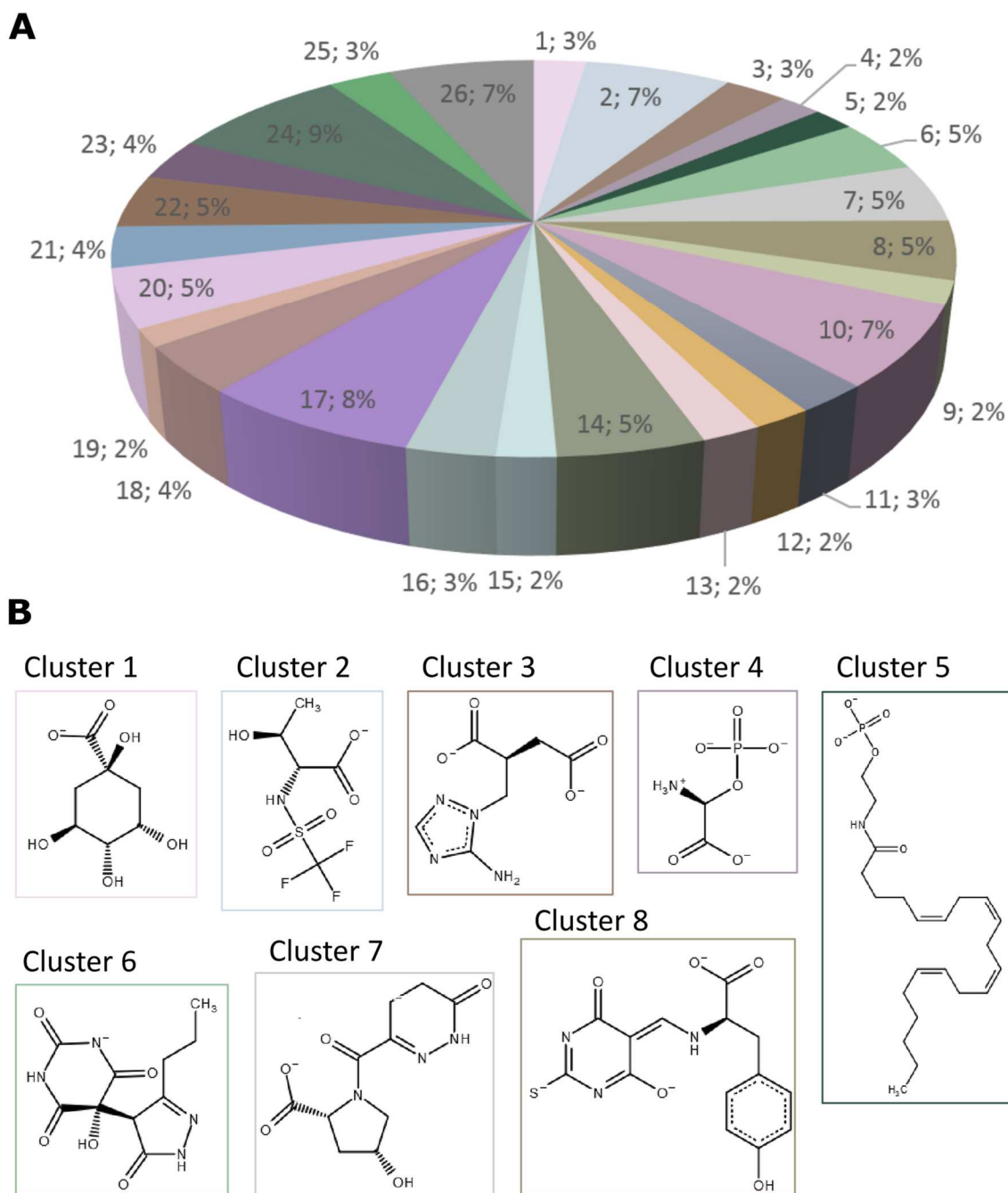
compostos foram os grupos 2, 10 e 26. Destes, os grupos 2 e 10 apresentaram uma composição aparentemente mais heterogênea, o que poderá ocasionar na escolha de mais representantes destes grupos para as etapas futuras de teste *in vitro*. Já o grupo 26 contém a presença majoritária de açúcares ligados a grupamentos fosfato. O grupo 25 também apresentou açúcares, nesse caso de 6 carbonos, porém ligados a grupamentos sulfato. Além de muitos grupos contendo carboidratos semelhantes aos próprios substratos de Rpi, vários grupos contendo nucleotídeos e seus análogos foram formados, como o grupo 22, contendo nucleotídeos derivados de pirimidinas como UDP (UMP, **47**) e TMP (**66**), assim como o grupo 23 contendo CMP (**59**). Já o grupo 20 apresentou nucleotídeos derivados de purina, como o GMP (**58**), e alguns análogos como o composto representante do grupo. Além disso, o grupo 21 apresentou compostos análogos a AMP (**46**), um dos nucleotídeos descritos por apresentar atividade de inibição em Rpis (91) (87). Um outro grupo interessante é o 14, que contém basicamente compostos contendo sulfonamidas. Além dele, o grupo 2, apesar de ser bastante heterogêneo, também apresenta alguns compostos com sulfonamidas (**67**). Sulfonamidas, de maneira geral, são amplamente usadas em química medicinal e estão presentes em diversos fármacos antimicrobianos. Diversos compostos com sulfonamida são inibidores competitivos da enzima ácido tetraidropteróico sintetase de bactérias (162), mas nenhuma atividade específica em RpiB foi reportada ainda.

Para a avaliação da diversidade de arcabouços moleculares do conjunto final de compostos foi utilizado o algoritmo *scaffold tree*, um método baseado em presença de anéis, de forma que cada composto foi associado com um arcabouço molecular único. A diversidade de arcabouços moleculares do conjunto de compostos está representada na Figura 4.13 e Figura 4.14. Nessa representação, arcabouços moleculares compartilhados por mais de um composto foram representados uma única vez, correspondendo a base da ramificação. As extremidades da ramificação representam um segundo nível hierárquico de arcabouços moleculares associados aos arcabouços moleculares base, referentes, portanto, a presença de mais de um anel em certos compostos. Como se pode observar, a diversidade de arcabouços moleculares entre os compostos identificados é bastante vasta, em torno de 44 arcabouços moleculares básicos, com grande parte deles sendo de anéis aromáticos. A presença de muitos arcabouços moleculares contendo anéis aromáticos pode ter sido favorecida por interações do tipo *pi-stacking* com o sítio ativo de TcRpiB, tipo de

interação que não acontece com HsRpiA. A Figura 4.15 é uma segunda representação dos arcabouços moleculares encontrados no conjunto final de compostos, associando ainda esses arcabouços moleculares com o número de moléculas correspondentes e com a média dos valores de DS em relação a TcRpiB. O arcabouço molecular mais frequente entre os compostos foi o oxano ou também chamado de tetraidropirano, presente em 28 compostos, principalmente entre açúcares referentes a diversos grupos, principalmente o grupo 24. Já o arcabouço molecular com maior afinidade predita a TcRpiB foi aquele contendo os anéis indol e derivado de pirimidina, referente apenas ao composto ZINC63480117 (grupo 11).

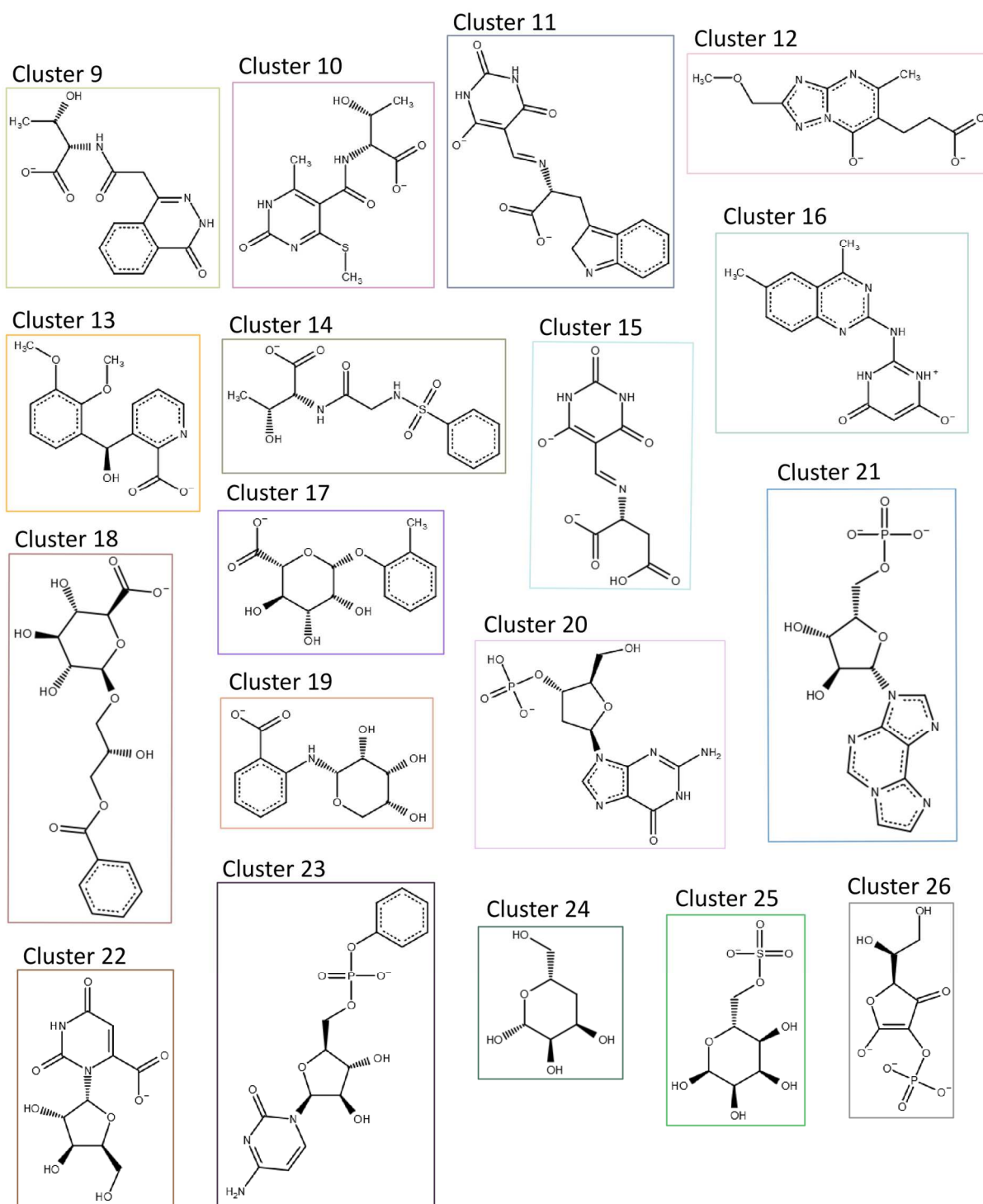


**Figura 4.10:** *Heat map* e dendograma referente a clusterização dos 211 potenciais inibidores de TcRpiB. O ponto de corte para a consideração de grupos individuais é representado pela linha vermelha. Ao todo 26 grupos foram gerados, cada um representado por um número e cor específica. Em a é representado o *heat map* referente a DS de cada composto no sítio ativo de TcRpiB. Em b é representado o *heat map* referente a DS de cada composto no sítio ativo de HsRpiA. Em c é representado o *heat map* referente a ausência (0) ou presença (1) do composto no metabolismo humano.

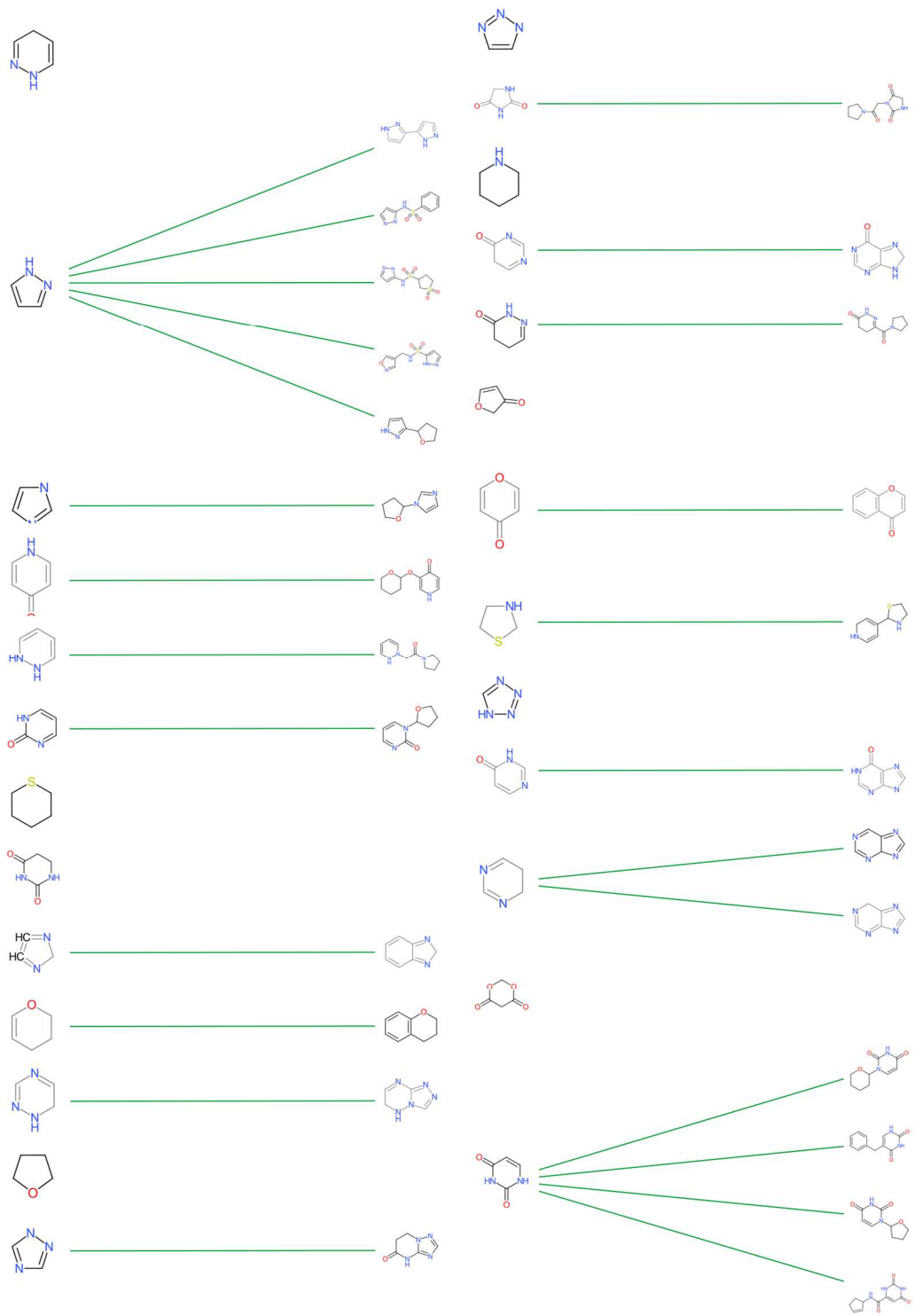


**Figura 4.11:** Resultado da clusterização dos 211 potenciais inibidores de TcRpiB. (A) Gráfico de pizza representando as porcentagens de compostos em cada grupo. Em a; b%, a é o número do grupo e b% é a porcentagem de compostos naquele grupo. (B) Compostos representativos de cada grupo. Foram escolhidos como representantes de cada grupo compostos de maior afinidade de ligação predita para TcRpiB e que não fossem metabólitos humanos.

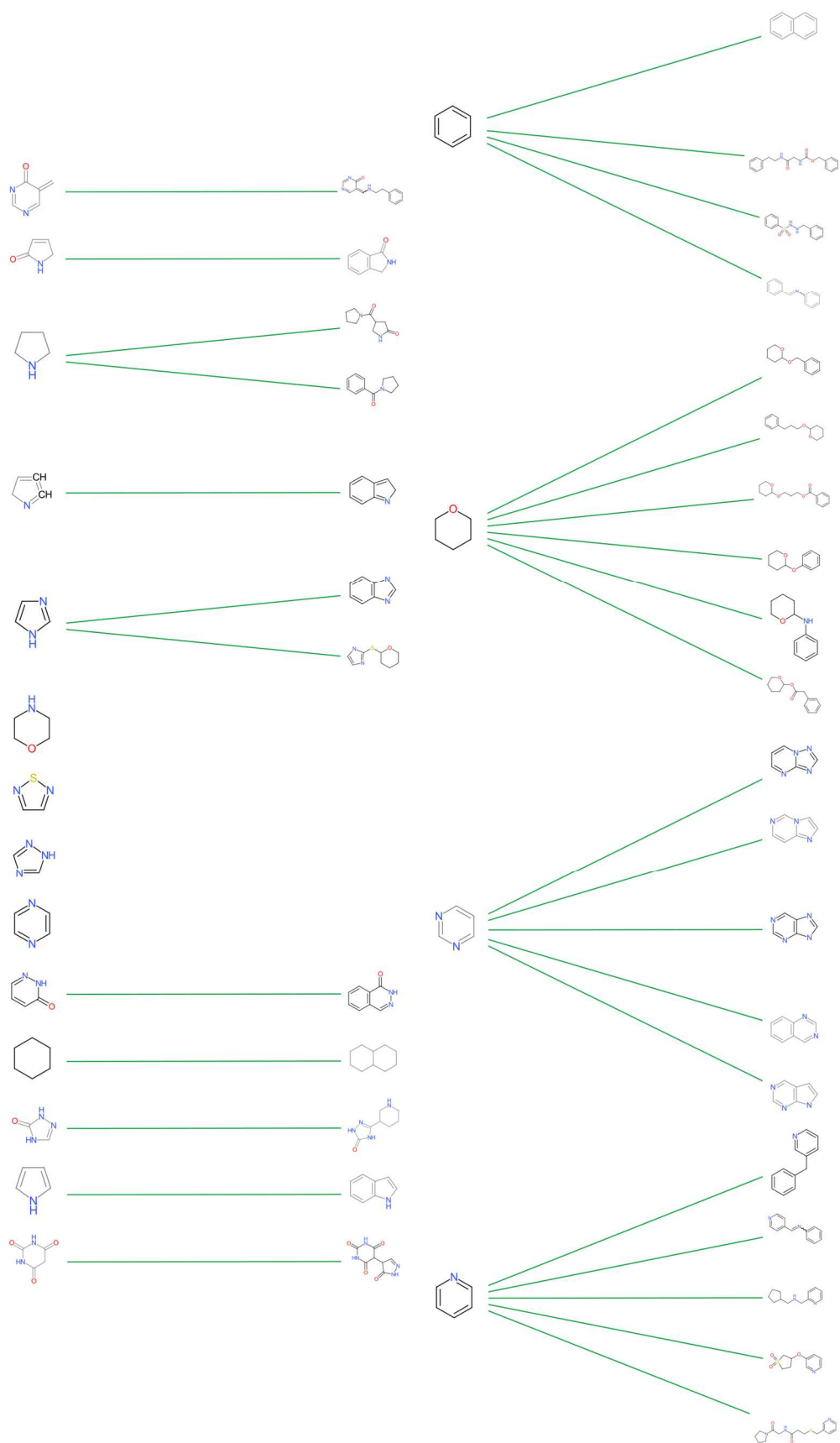




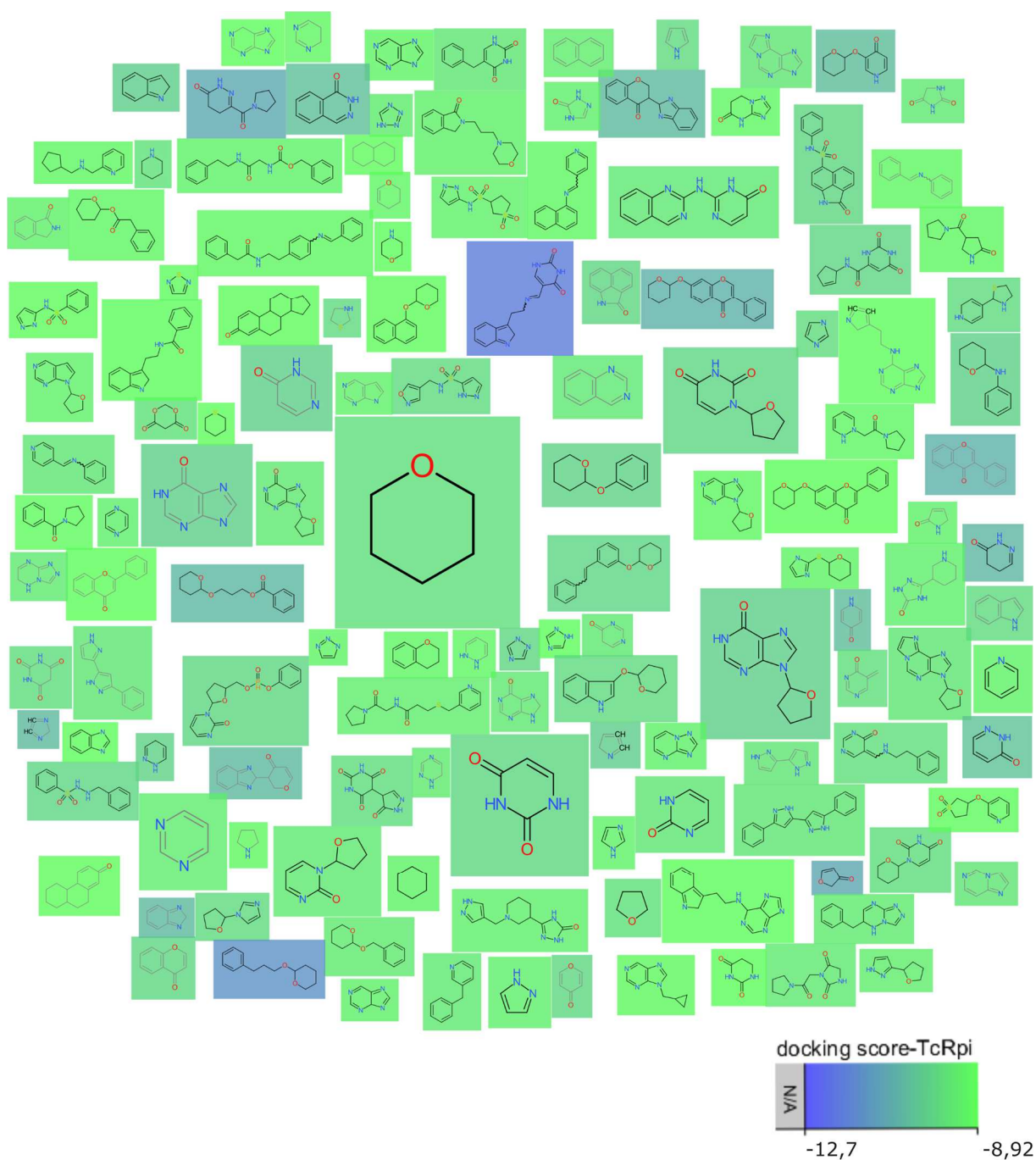
**Figura 4.12:** Continuação da Figura 4.11 com compostos representativos de cada grupo. Foram escolhidos como representantes de cada grupo compostos de maior afinidade de ligação predita para TcRpiB e que não fossem metabólitos humanos.



**Figura 4.13:** Ramificação de arcabouços moleculares, com até dois níveis, representando o conjunto final de potenciais inibidores de TcRpiB. As duas colunas de arcabouços moleculares são independentes uma da outra.



**Figura 4.14:** Continuação da Figura 4.13 contendo ramificação de arcabouços moleculares, com até dois níveis, representando o conjunto final de potenciais inibidores de TcRpiB. As duas colunas de arcabouços moleculares são independentes uma da outra.

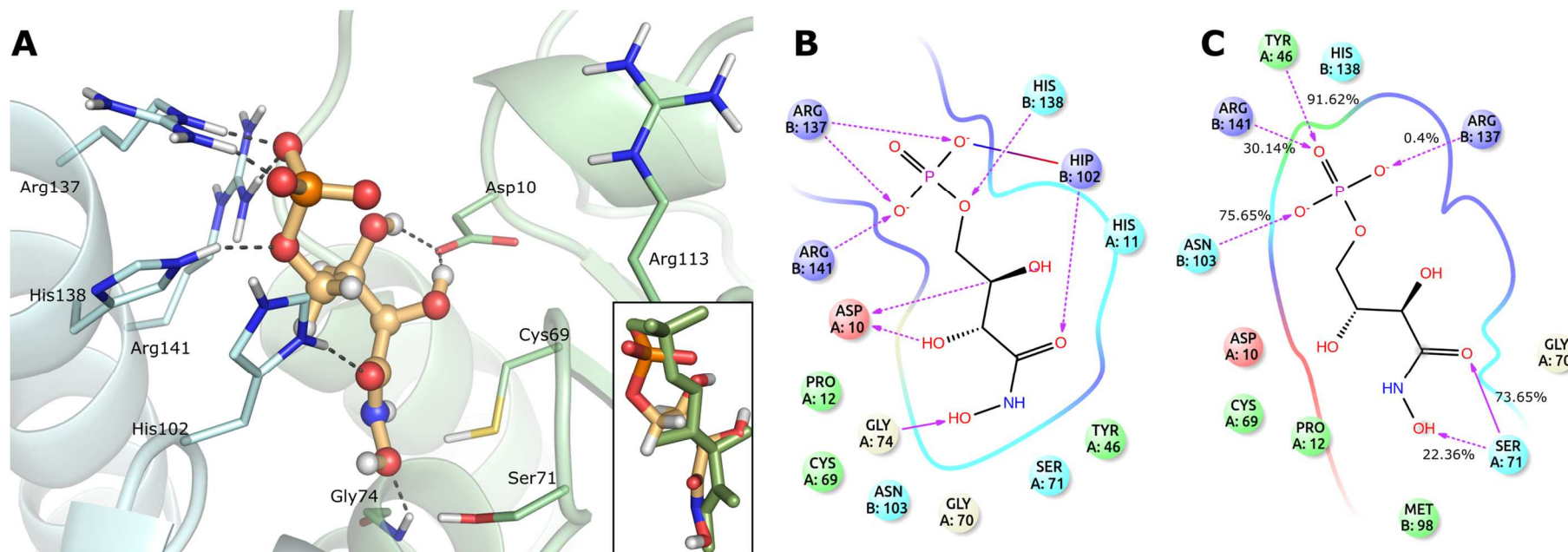


**Figura 4.15:** Arcabouços moleculares do conjunto final de potenciais inibidores de TcRpiB. O tamanho dos arcabouços moleculares é proporcional ao número de compostos que ele representa. As cores de cada arcabouço molecular estão relacionados ao DS dos compostos que o arcabouço molecular representa.

#### **4.6 Análises de interação molecular dos compostos mais bem ranqueados em relação a TcRpiB**

Um controle positivo e três compostos pertencentes a três grupos distintos, com as melhores pontuações de DS e com as propriedades ADME-Tox e PAINS adequadas foram selecionados para realização de análises mais detalhadas de seus modos de ligação no sítio ativo de TcRpiB: ZINC36975961 (grupo 7), ZINC63480117 (grupo 11) e ZINC43763931 (grupo 25) (Figura 4.17, Figura 4.18 e Figura 4.19), e o controle positivo (4PEH (**41**), Figura 4.16) no sítio ativo de TcRpiB. Todos esses três compostos selecionados foram identificados com a triagem de farmacóforo baseada em ligante, com exceção do composto ZINC43763931 que também foi identificado na triagem virtual baseada em estrutura com o PDB 3K7S.

O controle positivo 4PEH, que já possuía dados cristalográficos para TcRpiB (PDB 3K8C), foi redocado no sítio ativo dessa proteína, permitindo obter o valor de DS da pose redocada e comparar com o DS de outros candidatos a inibidores (Tabela 4.8). Essa pose possuiu um modo de ligação ligeiramente diferente da pose cristalográfica, com um RMSD de 1,33 Å (54) (Figura 4.16A). No modo de ligação obtido com o *redocking*, o grupamento fosfato interagiu através de quatro ligações de hidrogênio com o grupo imidazol da His138 e com os grupos guanidino dos resíduos Arg137 e Arg141 (Figura 4.16B). Além disso, dependendo da conformação da cadeia lateral da Arg113 é possível também uma ligação de hidrogênio do grupamento fosfato com o grupo guanidino dessa arginina. Foi também observada uma interação eletrostática do grupamento fosfato com o anel imidazol da His102. O grupo carbonila da 4PEH interagiu com o anel imidazol da His102 através de uma ligação de hidrogênio. Os três grupos hidroxila do inibidor efetuaram interações através de três ligações de hidrogênio com o grupo carboxilato do Asp10 e com a cadeia principal da Gly74.

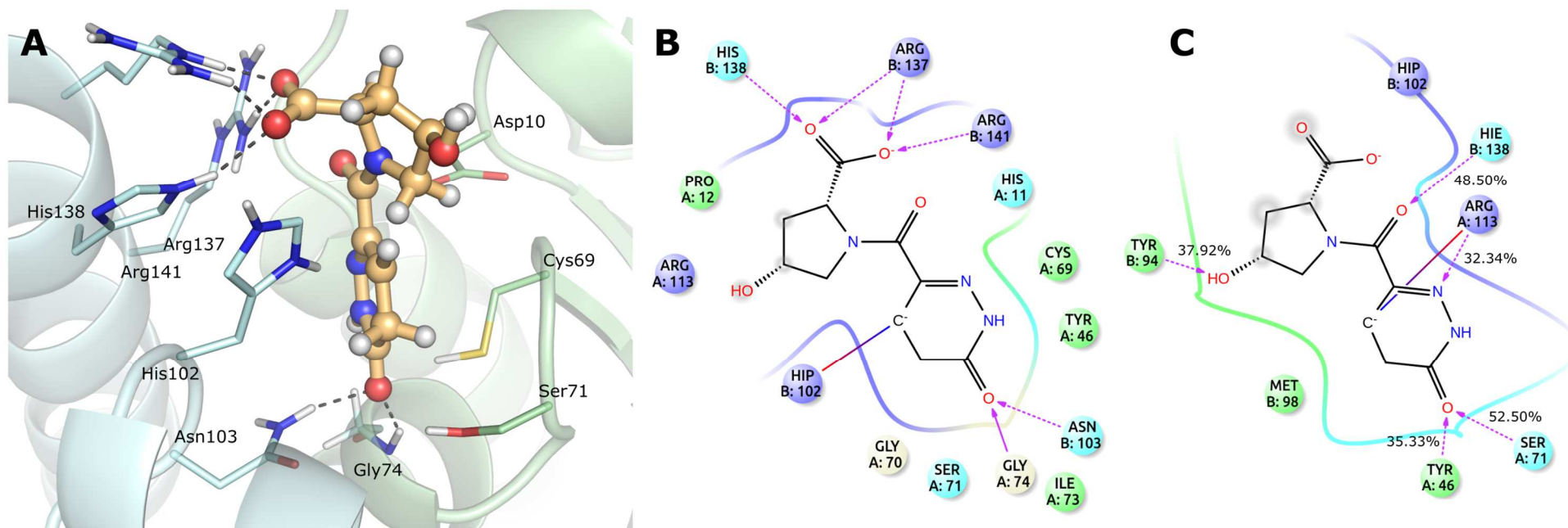


**Figura 4.16:** Modo de ligação predito entre 4PEH e proteína TcRpiB. (A) O modo de ligação de 4PEH no sítio ativo de TcRpiB predito na abordagem de atracamento molecular. A caixa quadrada mostra a sobreposição da pose de 4PEH obtida com atracamento (laranja) e da pose cristalográfica (verde escuro). O 4PEH está representado em *ball and stick* em laranja na imagem principal, as subunidades A e B da proteína estão representadas em cores verde-clara e ciano-claro, respectivamente, e as interações de ligações de hidrogênio estão representadas como traços cinzas; (B) Diagrama de interação 2D do 4PEH no sítio ativo de TcRpiB, de acordo com os resultados de atracamento molecular. (C) diagrama de interação 2D do modo de ligação do 4PEH após as simulações de dinâmica molecular e a ocupância das ligações de hidrogênio. Os resíduos estão distinguidos em carga negativa (rosa), carregado positivamente (roxo), polar (azul) e hidrofóbico (verde). As ligações de hidrogênio estão representadas como setas roxas e as interações eletrostáticas estão representadas como linhas azuis / vermelhas. As imagens 3D e 2D foram geradas usando Pymol v1.8.2.3 e Maestro, respectivamente.

O composto ZINC36975961 apresentou o melhor valor de DS (DS -12,735) e exibiu interações com vários resíduos do sítio ativo de TcRpiB, formando no total seis ligações de hidrogênio e uma interação eletrostática (Figura 4.17A e B). O grupo carboxilato desse possível inibidor fez três ligações de hidrogênio com os grupos guanidino da Arg137 e Arg141 e uma ligação de hidrogênio com o NE2 do grupo imidazol da His138. O grupo carbonila do anel derivado de piridazina apresentou uma ligação de hidrogênio com a cadeia principal da Gly74 e outra ligação de hidrogênio com o grupo amida da cadeia lateral da Asn103. Além disso, observou-se uma interação eletrostática entre o carbanion da ZINC36975961 e o ND1 do grupo imidazol da His102.

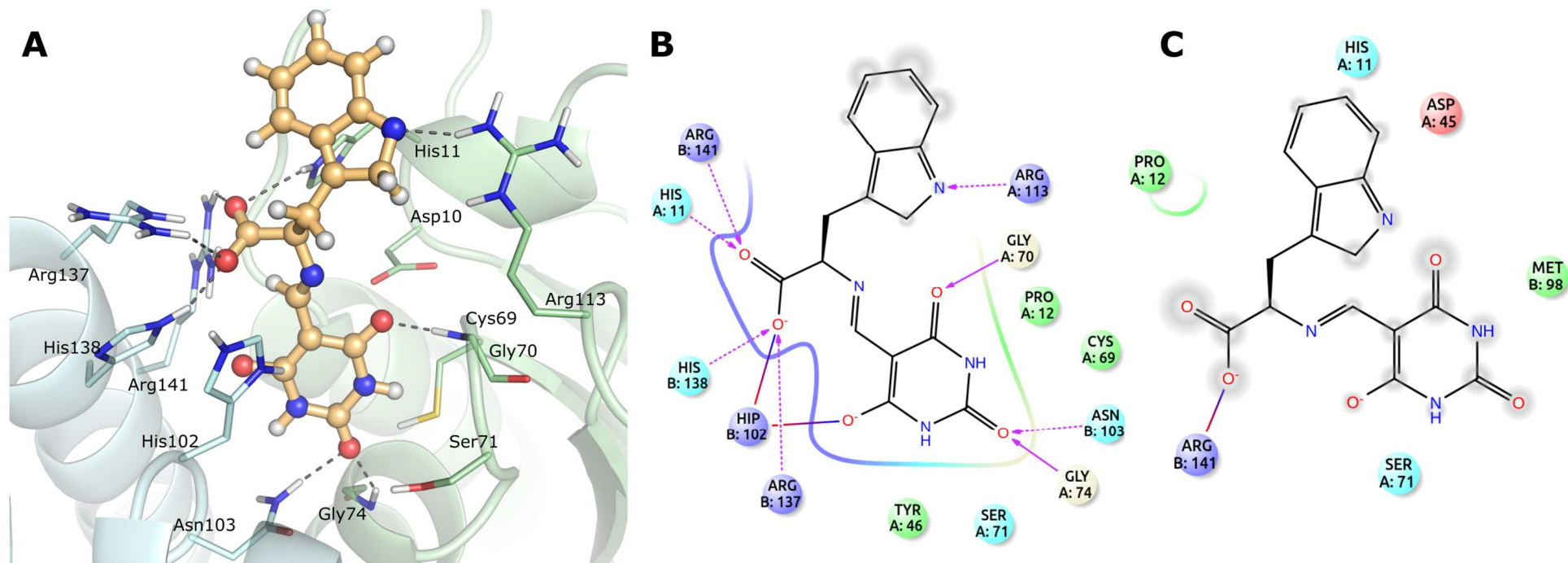
O segundo mais bem ranqueado, o composto ZINC63480117, apresentou uma pontuação de DS de -11,927 e de acordo com o seu modo de ligação predito, oito ligações de hidrogênio, duas interações eletrostáticas e uma possível interação pi-pi (*pi-stacking*) foram observadas (Figura 4.18A e B). Neste modo de ligação, o nitrogênio do grupo indol apresentou ligações de hidrogênio com o grupo guanidino da Arg113. De um modo semelhante ao descrito para ZINC36975961, o grupo carboxilato do ligante foi estabilizado através de ligações de hidrogênio com as cadeias laterais da Arg137, Arg141 e His138 e de uma ligação de hidrogênio com o grupo imidazol da His11 que não foi observada na interação com ZINC36975961. Além disso, este grupo carboxilato teve uma interação eletrostática com o grupo imidazol da His102 e esse mesmo resíduo interagiu com o oxigênio negativamente carregado da porção do derivado de pirimidina. Os dois grupos carbonila da porção do derivado de pirimidina do ligante formaram três ligações de hidrogênio no total: uma com o grupo amida da cadeia lateral da Asn103 e duas com as cadeias laterais da Gly70 e Gly74. Além disso, houve uma possível interação pi-pi entre o anel benzeno da porção indol do ligante com o grupo imidazol da His11.

O modo de ligação do composto ZINC43763931 no sítio ativo de TcRpiB apresentou oito ligações de hidrogênio e uma interação eletrostática, representando o valor de DS de -11,849 (Figura 4.19A e B). Do mesmo modo que o grupo carboxilato do composto ZINC63480117, o íon sulfato do ZINC43763931 fez interações através de quatro ligações de hidrogênio com as cadeias laterais dos resíduos Arg137, Arg141, His138 e His11 e uma interação eletrostática com a His102. Três das quatro hidroxilas do ligante interagiram através de quatro ligações de hidrogênio com o grupo carboxilato do Asp10 e com as cadeias laterais dos resíduos Gly70, Ser71 e Gly74.

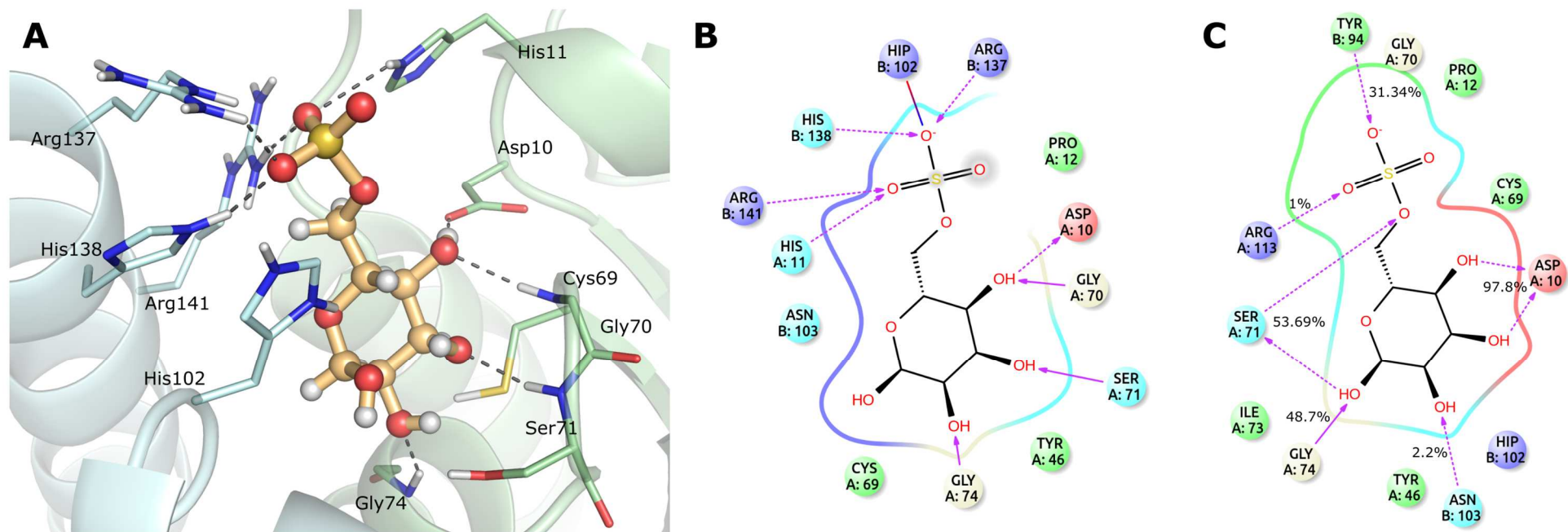


**Figura 4.17:** Modo de ligação predito entre ZINC36975961 e a proteína TcRpiB. (A) modo de ligação obtido com o atracamento molecular de ZINC36975961 no sítio ativo de TcRpiB. O inibidor está representado em *ball and stick* em laranja, as subunidades A e B da proteína estão representadas em cores verde-clara e ciano-claro, respectivamente, e as interações referentes a ligações de hidrogênio estão representadas como traços cinzas; (B) diagrama de interação 2D do resultado do atracamento molecular de ZINC36975961. (C) diagrama de interação 2D obtido após as simulações de dinâmica molecular do composto ZINC36975961 em complexo com a proteína TcRpiB e as ocupâncias das ligações de hidrogênio. Os resíduos no diagrama 2D estão distinguidos em carga negativa (rosa), positivamente carregados (roxa), polar (azul), hidrofóbicos (verde). As ligações de hidrogênio estão representadas como setas roxas e as interações eletrostáticas estão representadas como linhas azuis/vermelhas. As imagens 3D e 2D foram geradas usando Pymol v1.8.2.3 e Maestro, respectivamente.





**Figura 4.18:** Modo de ligação predito entre ZINC63480117 e a proteína TcRpiB. (A) modo de ligação obtido com o atracamento molecular de ZINC63480117 no sítio ativo de TcRpiB. O inibidor está representado em ball and stick em laranja, as subunidades A e B da proteína estão representadas em cores verde-clara e ciano-claro, respectivamente, e as interações referentes a ligações de hidrogênio estão representadas como traços cinzas; (B) diagrama de interação 2D do resultado do atracamento molecular de ZINC63480117. (C) diagrama de interação 2D obtido após as simulações de dinâmica molecular do composto ZINC63480117 em complexo com a proteína TcRpiB e as ocupâncias das ligações de hidrogênio. Os resíduos no diagrama 2D estão distinguidos em carga negativa (rosa), positivamente carregados (roxa), polar (azul), hidrofóbicos (verde). As ligações de hidrogênio estão representadas como setas roxas e as interações eletrostáticas estão representadas como linhas azuis/vermelhas. As imagens 3D e 2D foram geradas usando Pymol v1.8.2.3 e Maestro, respectivamente.



**Figura 4.19:** Modo de ligação predito entre ZINC43763931 e a proteína TcRpiB. (A) modo de ligação obtido com o atracamento molecular de ZINC43763931 no sítio ativo de TcRpiB. O inibidor está representado em ball and stick em laranja, as subunidades A e B da proteína estão representadas em cores verde-clara e ciano-claro, respectivamente, e as interações referentes a ligações de hidrogênio estão representadas como traços cinzas; (B) diagrama de interação 2D do resultado do atracamento molecular de ZINC43763931. (C) diagrama de interação 2D obtido após as simulações de dinâmica molecular do composto ZINC43763931 em complexo com a proteína TcRpiB e as ocupâncias das ligações de hidrogênio. Os resíduos no diagrama 2D estão distinguidos em carga negativa (rosa), positivamente carregados (roxa), polar (azul), hidrofóbicos (verde). As ligações de hidrogênio estão representadas como setas roxas e as interações eletrostáticas estão representadas como linhas azuis/vermelhas. As imagens 3D e 2D foram geradas usando Pymol v1.8.2.3 e Maestro, respectivamente.

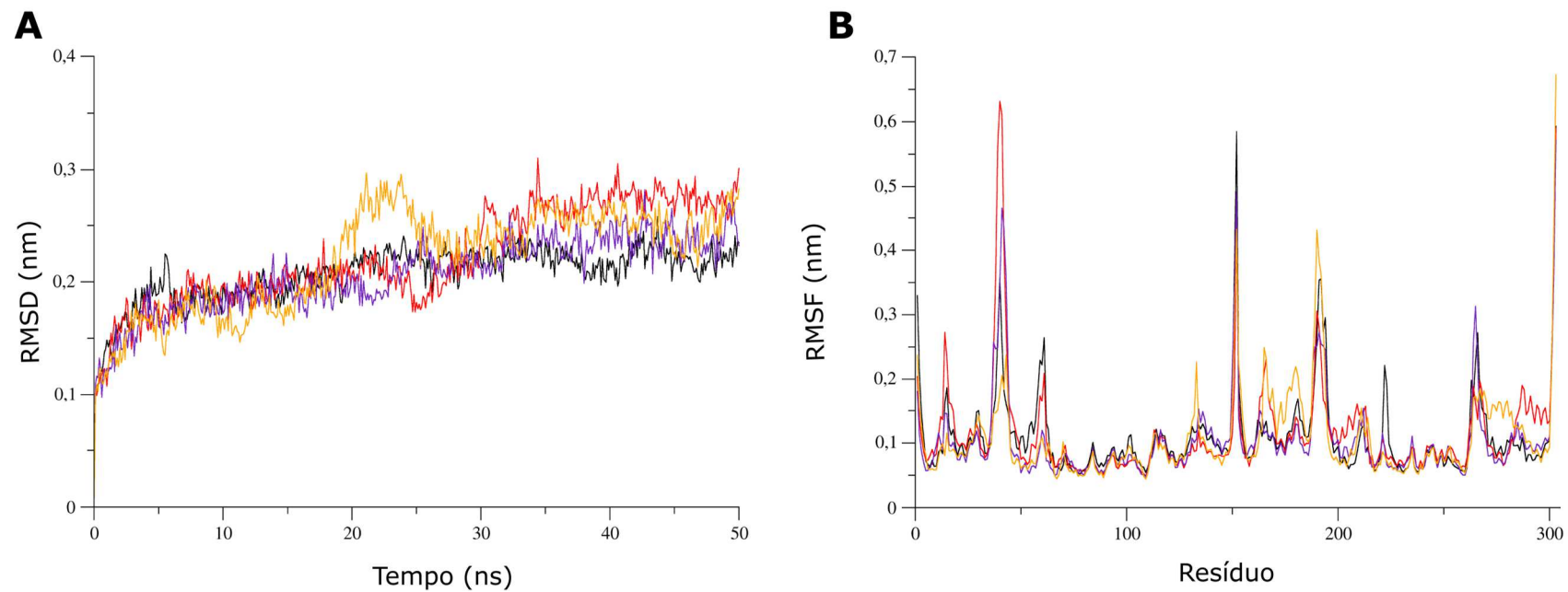
Estes compostos mostraram resultados de atracamento molecular promissores, uma vez que eles formaram interações-chave com alguns resíduos catalíticos da proteína alvo, incluindo His102 e Ser71, e com resíduos descritos como importantes na ligação ao substrato como a His11, Asp10, Gly70, Gly74, Asn103, Arg113, Arg137, His138 e Arg141 (54,70). De modo semelhante, algumas destas interações são também encontradas no inibidor 4PEH co-cristalizado e também na pose obtida com o *redocking*. De fato, o grupamento com carga negativa, sugerida como uma característica farmacofórica essencial a inibição de Rpi do tipo B devido à natureza altamente eletropositiva do sítio ativo dessa proteína (77), está presente em todos os compostos selecionados, tornando possível algumas destas interações-chave.

#### **4.7 Simulações de dinâmica molecular**

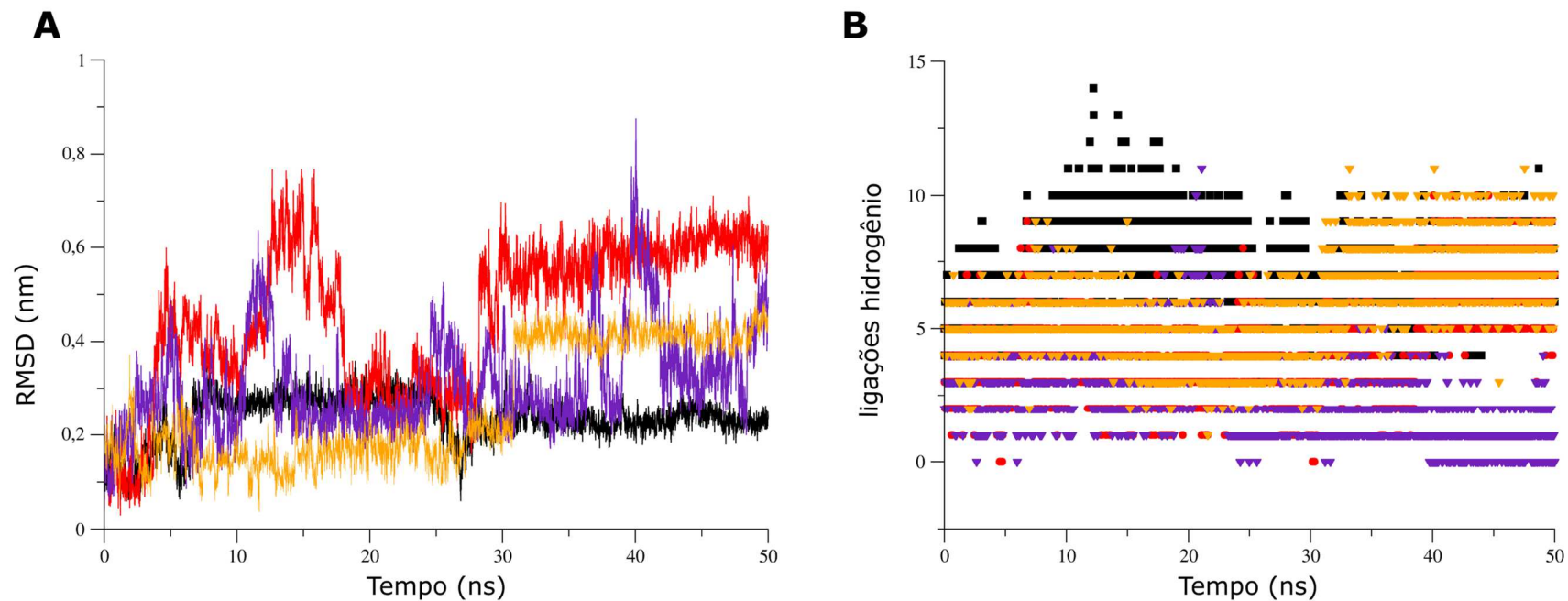
Simulações de DM de 50 ns foram realizadas para analisar a estabilidade dos complexos proteína-ligante obtidos nos ensaios de atracamento molecular (Figuras 4.16 à 4.21). A partir destas simulações de DM, os RMSD dos C $\alpha$  das proteínas nos complexos foram calculados como mostrado na Figura 4.20A. O complexo TcRpiB-controle positivo (4PEH, 41) foi o primeiro a atingir o equilíbrio a 20ns com valor médio de RMSD de 0,22 nm. Os complexos TcRpiB-ZINC63480117 e TcRpiB-ZINC43763931 atingiram o *plateau* após 30 ns com RMSD médio de 0,24 nm e 0,25 nm, respectivamente. O último a atingir o equilíbrio foi o complexo TcRpiB-ZINC36975961 com RMSD de 0,27 nm. As regiões de maior flexibilidade na proteína dos complexos foram avaliadas através do RMSF dos C $\alpha$  como mostrado na Figura 4.20B. Flutuações acima de 0,25 nm foram observadas nos complexos para Met1A, Ala15A, Ser71B, Thr40A / Thr40B-Ser43A / Ser43B, Glu61A / Glu161B, Ser151A / Ser151B e His152A / His152B, mas nenhum deles são resíduos do sítio ativo com os potenciais inibidores. A estabilidade de ligação destes inibidores foi assegurada através dos seus valores de RMSD (Figura 4.21A). Todos os inibidores se desviaram de suas estruturas iniciais atingindo conformações estáveis após 30 ns, com exceção do ligante ZINC63480117 que apresentou flutuações nos valores de RMSD ao longo de toda a simulação. Os ligantes 4PEH, ZINC36975961 e

ZINC43763931 estabilizaram com RMSD médio de 0,23 nm, 0,58 nm e 0,4 nm, respectivamente. É importante notar que todos os inibidores permaneceram no sítio de ligação ao longo de toda a simulação.

As interações entre os ligantes e a proteína TcRpiB foram avaliadas mais detalhadamente através da análise das ligações de hidrogênio formadas durante as trajetórias (Figura 4.21B). No caso do controle positivo 4PEH uma média de 7 interações de ligações de hidrogênio com o sítio ativo de TcRpiB foi observada ao longo da trajetória. Em relação as interações observadas com o atracamento molecular, durante a trajetória este ligante manteve ligações de hidrogênio com o Asp10 (98,8% de ocupância), Gly74 (25%), His138 (84,83%) e Arg141 (30%). Além disso, novas ligações de hidrogênio foram formadas com a Tyr46, Ser71 e Asn103 com ocupância superior a 70%. O inibidor potencial ZINC36975961 obteve uma média de 4 ligações de hidrogênio ao longo da simulação e manteve ligações de hidrogênio com a Asn103 (37,92% de ocupância), His138 (48,5%) e Arg141 (7%) daquelas anteriormente observada no complexo obtido com o atracamento. Durante a simulação foram formadas novas interações com Ser71 (52,5% de ocupância), Arg113 (32,34%), His102 (49,7%), Tyr94 (37,52%) e Tyr46 (35,33%). No caso do composto ZINC63480117, durante a dinâmica ocorreu a formação de duas ligações de hidrogênio com ocupâncias superiores a 20% com a Asn103, His 11, Ser43, Arg141 e His102. No final da simulação, o ligante perdeu todas as ligações de hidrogênio e foi mantido no sítio catalítico através da interação eletrostática do grupo carboxilato do ligante com Arg141 (Figura 4.18F). Já entre o composto ZINC43763931 e TcRpiB, uma média de 6 ligações de hidrogênio se formaram durante a simulação. Em comparação com as interações observadas na pose do atracamento, este ligante conservou ligações de hidrogênio com o Asp10 (97,8% de ocupância), a cadeia principal (40,12%) e cadeia lateral (53,69%) da Ser71 e Gly74 (48,7%), porém baixas ocupâncias foram obtidas com a Arg141 (4,19%) e His138 (8,58%). Foram formadas novas ligações de hidrogênio com ocupâncias consideráveis com Asn103 (40,92%), His102 (28,54%) e Tyr94 (31,34%). As interações observadas para cada inibidor – 4PEH, ZINC36975961, ZINC63480117 e ZINC43763931 – no final dos 50 ns de simulação são mostradas na Figura 4.16C e nas Figura 4.17C, Figura 4.18C e Figura 4.19C, respectivamente.



**Figura 4.20:** Resultados dos estudos de dinâmica molecular dos novos inibidores potenciais e o 4PEH no sítio ativo de TcRpiB. (A) Resultados da trajetória da simulação de dinâmica molecular ilustrando a RMSD dos C $\alpha$  de proteína TcRpiB em complexo com ZINC36975961 (vermelho), ZINC63480117 (índigo), ZINC43763931 (laranja) e o controle positivo 4PEH (preto); (B) RMSF dos C $\alpha$  da proteína TcRpiB nos complexos.



**Figura 4.21:** Resultados dos estudos de dinâmica molecular dos novos inibidores potenciais e o 4PEH no sítio ativo de TcRpiB. (A) RMSD dos inibidores potenciais ZINC36975961 (vermelho), ZINC63480117 (índigo), ZINC43763931 (laranja) e o controle positivo 4PEH (preto); (B) Ligações de hidrogênio entre a proteína TcRpiB e ZINC36975961, ZINC63480117, ZINC43763931 e o 4PEH ao longo das simulações.

Estes resultados indicam que, conforme esperado para um controle positivo, a 4PEH apresentou interações moleculares estáveis com o sítio ativo de TcRpiB ao longo de todos os 50 ns de simulação, sendo estas interações consideravelmente mais estáveis do que nos outros inibidores em potencial. As simulações de DM também revelaram que os ligantes ZINC36975961 e ZINC43763931 fizeram ao longo da simulação interações importantes com resíduos de TcRpiB envolvidos na catálise e na ligação ao substrato, como também já havia sido observado nos estudos de atracamento (item 4.6). Além disso, algumas dessas interações observadas nos inibidores em potencial também são encontradas no inibidor 4PEH (controle positivo) ao longo da simulação de DM, no ensaio de atracamento e na pose cristalográfica. Este padrão de interações e os valores de *docking score* de ligação obtidos nos estudos de atracamento molecular sugerem que os compostos ZINC36975961 e ZINC43763931 são candidatos promissores a inibidores de TcRpiB. Já o composto ZINC63480117 se manteve instável ao longo dos 50 ns de simulação e perdeu algumas interações importantes no final desse tempo de simulação. No entanto, a pose obtida no atracamento molecular predita para este composto apresentou interações notáveis; somado a isso, o composto ZINC63480117 permaneceu no sítio ativo durante toda simulação DM. Dessa forma, a pose do atracamento do composto ZINC63480117 pode ser interessante para guiar na proposta de novas estruturas químicas para estudos de desenho de fármacos.

## 5 CONCLUSÕES

Após a realização desse estudo, podemos concluir que:

- ✓ As estruturas primárias, secundárias e terciárias de HsRpiA e TcRpiB apresentam diferenças marcantes que reforçam a hipótese de analogia entre essas enzimas, tornando a TcRpiB um alvo interessante para o desenvolvimento de fármacos;
- ✓ Apesar da analogia, os sítios ativos de HsRpiA e TcRpiB possuem algumas características similares, como a presença de resíduos básicos próximos a entrada dos sítios e a, conseqüente, superfície eletrostática majoritariamente positiva;
- ✓ A análise de estruturas cristalográficas de RpiBs permitiu identificar a presença de moléculas de águas conservadas (presentes em mais que 60% das estruturas) e de moléculas de águas muito conservadas (presentes em mais que 80% das estruturas) no sítio ativo de TcRpiB;
- ✓ Ensaios de atracamento molecular com o programa Glide XP mostraram que a ausência de moléculas de água conservadas ou muito conservadas no sítio ativo de TcRpiB geram melhores resultados;
- ✓ Três modelos de farmacóforo baseados em ligante foram gerados considerando inibidores conhecidos de RpiB;
- ✓ Dois modelos de farmacóforo baseados em estrutura foram construídos a partir das estruturas cristalográficas disponíveis de TcRpiB;
- ✓ Uma biblioteca contendo 20.183 compostos, enriquecida com potenciais inibidores de TcRpiB, foi construída através da triagem virtual do banco de dados ZINC baseada nos modelos de farmacóforo, em similaridade 2D e subestrutura;



- ✓ Estudos de atracamento molecular com TcRpiB e, em seguida, com HsRpiA permitiram a identificação de 331 potenciais inibidores seletivos para a enzima do parasita;
- ✓ Dos 331 compostos selecionados, 211 apresentaram propriedades farmacocinéticas e toxicidade preditas adequadas;
- ✓ A clusterização desses 211 compostos resultou na formação de 26 grupos, contendo de três a dezessete compostos cada, determinados pelo grau de similaridades (coeficiente de Tanimoto) entre os compostos;
- ✓ Análises mais detalhadas do modo de ligação dos três compostos pertencentes a três grupos distintos (ZINC36975961 do grupo 7, ZINC63480117 do grupo 11 e ZINC43763931 do grupo 25) e com as melhores pontuações de DS mostraram padrões de interação promissores;
- ✓ O composto ZINC63480117, embora tenha permanecido no sítio ativo durante toda a trajetória de dinâmica molecular, mostrou instabilidade durante a simulação e perdeu algumas interações importantes;
- ✓ Simulações de dinâmica molecular revelaram que os compostos ZINC36975961 e ZINC43763931, além de formarem novas interações, mantiveram, com o sítio ativo de TcRpiB, algumas interações chaves observadas nos estudos de atracamento;
- ✓ Estes resultados indicam que os compostos identificados ZINC36975961 e ZINC43763931 são inibidores potenciais para TcRpiB e, após suas atividades serem confirmadas experimentalmente, podem também ser utilizados como condutores para o desenho de novos inibidores de TcRpiB.

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

1. Schmunis G a. Epidemiology of Chagas disease in non-endemic countries: the role of international migration. Mem Inst Oswaldo Cruz [Internet]. 2007 Oct 30;102 Suppl(PAHO 2006):75–85.
2. Liu Q, Zhou X. Preventing the transmission of American trypanosomiasis and its spread into non-endemic countries. Infect Dis poverty [Internet]. 2015;4:60.
3. WHO | Epidemiology [Internet]. WHO. World Health Organization; 2016 [cited 2016 Dec 21]. Available from: <http://www.who.int/chagas/epidemiology/en/>
4. WHO | Chagas disease (American trypanosomiasis) [Internet]. WHO. World Health Organization; 2017. Available from: <http://www.who.int/mediacentre/factsheets/fs340/en/>
5. Coura JR. The main sceneries of chagas disease transmission. The vectors, blood and oral transmissions - A comprehensive review. Mem Inst Oswaldo Cruz. 2015;110(3):277–82.
6. Bermudez J, Davies C, Simonazzi A, Pablo Real J, Palma S. Current drug therapy and pharmaceutical challenges for Chagas disease. Acta Trop [Internet]. 2016;156:1–16.
7. Neglected tropical diseases: A statistical update [Internet]. World Health Organization (WHO); 2014 [cited 2016 Dec 22].
8. Stanaway JD, Roth G. The Burden of Chagas Disease Estimates and Challenges [Internet]. Vol. 10, Global Heart. World Heart Federation (Geneva); 2015. p. 139–44.

9. Schofield CJ, Jannin J, Salvatella R. The future of Chagas disease control. *Trends Parasitol.* 2006;22(12):583–8.
10. Dias JCP, Silveira AC, Schofield CJ. The impact of Chagas disease control in Latin America: a review. *Mem Inst Oswaldo Cruz* [Internet]. 2002 Jul;97(5):603–12.
11. Benziger CP. Chagas Cardiomyopathy: Clinical Presentation and Management in the Americas Chagas disease Neglected diseases Cardiomyopathy Epidemiology Trypanocidal agents. *Cardiol Clin* [Internet]. 2016;35(1):31–47.
12. Pérez-Molina J a, Norman F, López-Vélez R. Chagas disease in non-endemic countries: epidemiology, clinical presentation and treatment. *Curr Infect Dis Rep* [Internet]. 2012 Jun [cited 2013 Mar 6];14(3):263–74.
13. Rassi A, Marcondes de Rezende J. American trypanosomiasis (Chagas disease). *Infect Dis Clin North Am* [Internet]. 2012 Jun [cited 2013 Mar 9];26(2):275–91.
14. Chatelain E. Chagas Disease Drug Discovery: Toward a New Era. *J Biomol Screen.* 2015;20(1):22–35.
15. Rassi A, Marin-Neto JA. Chagas disease. *Lancet* [Internet]. 2010 Apr 17 [cited 2013 Mar 4];375(9723):1388–402.
16. Muñoz-Saravia SG, Haberland A, Wallukat G, Schimke I. Chronic Chagas' heart disease: a disease on its way to becoming a worldwide health problem: epidemiology, etiopathology, treatment, pathogenesis and laboratory medicine. *Heart Fail Rev* [Internet]. 2012 Jan [cited 2013 Mar 21];17(1):45–64.

17. Freitas HFG, Chizzola PR, Paes ??ngela T., Lima ACP, Mansur AJ. Risk stratification in a Brazilian hospital-based cohort of 1220 outpatients with heart failure: Role of Chagas' heart disease. *Int J Cardiol.* 2005 Jul 10;102(2):239–47.
18. Matsuda NM, Miller SM, Evora PRB. The chronic gastrointestinal manifestations of Chagas disease. *Clinics (Sao Paulo).* 2009;64(12):1219–24.
19. Zingales B, Miles MA, Moraes CB, Luquetti A, Guhl F, Schijman AG, et al. Drug discovery for chagas disease should consider *Trypanosoma cruzi* strain diversity. *Mem Inst Oswaldo Cruz.* 2014;109(6):828–33.
20. Zingales B, Andrade SG, Briones MRS, Campbell DA, Chiari E, Fernandes O, et al. A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: Second revision meeting recommends TcI to TcVI. *Mem Inst Oswaldo Cruz.* 2009;104(7):1051–4.
21. MARCILI A, LIMA L, CAVAZZANA M, JUNQUEIRA AC V., VELUDO HH, MAIA DA SILVA F, et al. A new genotype of *Trypanosoma cruzi* associated with bats evidenced by phylogenetic analyses using SSU rDNA, cytochrome b and Histone H2B genes and genotyping based on ITS1 rDNA. *Parasitology [Internet].* 2009 May 16 [cited 2017 Aug 9];136(6):641.
22. Brenière SF, Waleckx E, Barnabé C. Over Six Thousand *Trypanosoma cruzi* Strains Classified into Discrete Typing Units (DTUs): Attempt at an Inventory. *PLoS Negl Trop Dis.* 2016;10(8):1–19.
23. Souza W de. Structural organization of *Trypanosoma cruzi*. *Mem Inst Oswaldo Cruz [Internet].* 2009 Jul [cited 2016 Dec 26];104:89–100.
24. De Souza W. Basic cell biology of *Trypanosoma cruzi*. *Curr Pharm Des [Internet].* 2002;8(4):269–85.

25. Vianna Martins A, Patrícia Gomes A, de Mendonça EG, Rangel Fietto JL, Alberto Santana L, de Almeida Oliveira MG, et al. Biology of *Trypanosoma cruzi*: An update. *Infectio*. 2012;16(1):45–58.
26. Teixeira DE, Benchimol M, Crepaldi PH, de Souza W. Interactive multimedia to teach the life cycle of *Trypanosoma cruzi*, the causative agent of Chagas disease. *PLoS Negl Trop Dis* [Internet]. 2012 Aug [cited 2013 Mar 10];6(8):e1749.
27. Teixeira DE, Benchimol M, Crepaldi PH DSW. Atlas didático: Ciclo de vida do *Trypanosoma cruzi*. 1ª. Rio de Janeiro: Fundação CECIERJ, Consórcio CEDERJ; 2011.
28. Lima FM, Oliveira P, Mortara R a, Silveira JF, Bahia D. The challenge of Chagas' disease: has the human pathogen, *Trypanosoma cruzi*, learned how to modulate signaling events to subvert host cells? *N Biotechnol* [Internet]. 2010 Dec 31 [cited 2013 Mar 13];27(6):837–43.
29. Nagajyothi F, Machado FS, Burleigh BA, Jelicks LA, Scherer PE, Mukherjee S, et al. Mechanisms of *Trypanosoma cruzi* persistence in Chagas disease. *Cell Microbiol*. 2012;14(5):634–43.
30. Bern C. Antitrypanosomal therapy for chronic Chagas' disease. *N Engl J Med* [Internet]. 2011 Jun 30;364(26):2527–34.
31. Vanessa de Vasconcelos Sinatti Castilho. Clonagem, expressão heteróloga e caracterização molecular das enzimas gliconato cinase e uracil fosforribosiltransferase de *Tripanosoma cruzi*. Instituto Oswaldo Cruz- Fiocruz; 2014.
32. Wilkinson SR, Kelly JM. Trypanocidal drugs: mechanisms, resistance and new

- targets. *Expert Rev Mol Med*. 2009;11(October 2009):e31.
33. WHO. Control of Chagas disease : second report of the WHO expert committee [Internet]. Geneva; 2002.
  34. Morillo CA, Avezum A, Rassi A, Rosas F, Villena E, Quiroz R, et al. Randomized Trial of Benznidazole for Chronic Chagas' Cardiomyopathy. *N Engl J Med*. 2015;373:1295–306.
  35. Silverman RB, Holladay MW. *The Organic Chemistry of Drug Design and Drug Action*. Third Edit. Elsevier Inc.; 2014. 10, 13-14, 19-20 p.
  36. FDA. The Drug Development Process - Step 3: Clinical Research [Internet]. Office of the Commissioner; [cited 2017 Jan 5]. Available from: <http://www.fda.gov/ForPatients/Approvals/Drugs/ucm405622.htm>
  37. Molina I, Gómez i Prat J, Salvador F, Treviño B, Sulleiro E, Serre N, et al. Randomized Trial of Posaconazole and Benznidazole for Chronic Chagas' Disease. *N Engl J Med* [Internet]. 2014 May 15 [cited 2016 Dec 30];370(20):1899–908.
  38. A Study of the Use of Oral Posaconazole (POS) in the Treatment of Asymptomatic Chronic Chagas Disease (P05267) - Full Text View - ClinicalTrials.gov [Internet]. [cited 2016 Dec 30]. Available from: <https://clinicaltrials.gov/ct2/show/study/NCT01377480>
  39. Study of use of Oral Posaconazole on the Treatment of asymptomatic chronic CHAGAS disease - American College of Cardiology [Internet]. [cited 2016 Dec 30]. Available from: <http://www.acc.org/latest-in-cardiology/clinical-trials/2016/04/03/15/46/stop-chagas>

40. Azoles E1224 – DNDi [Internet]. [cited 2016 Dec 30]. Available from: <http://www.dndi.org/diseases-projects/portfolio/completed-projects/azoles-e1224/>
41. K777 (Chagas) – DNDi [Internet]. [cited 2016 Dec 30]. Available from: <http://www.dndi.org/diseases-projects/portfolio/completed-projects/k777/>
42. Otto TD, Guimarães ACR, Degraive WM, de Miranda AB. AnEnPi: identification and annotation of analogous enzymes. *BMC Bioinformatics* [Internet]. 2008 Jan [cited 2013 Mar 21];9:544.
43. Alves-Ferreira M, Guimarães ACR, Capriles PVSZDSZ, Dardenne LE, Degraive WM. A new approach for potential drug target discovery through in silico metabolic pathway analysis using *Trypanosoma cruzi* genome information. *Mem Inst Oswaldo Cruz* [Internet]. 2009 Dec;104(8):1100–10.
44. Faria J, Loureiro II, Santarem N, Cecilio P, Macedo-Ribeiro S, Tavares J, et al. Disclosing the essentiality of ribose-5-phosphate isomerase B in *Trypanosomatids*. *Sci Rep* [Internet]. 2016 May 27 [cited 2016 Jun 2];6(January):26937.
45. Nelson DL, Cox MM. *Princípios de bioquímica de Lehninger*. Princípios de bioquímica de Lehninger. 2011. p. 1274.
46. Kovářová J, Barrett MP. The Pentose Phosphate Pathway in Parasitic *Trypanosomatids*. *Trends Parasitol*. 2016;32(8):622–34.
47. Wamelink MMC, Struys EA, Jakobs C. The biochemistry, metabolism and inherited defects of the pentose phosphate pathway: A review. *J Inherit Metab Dis* [Internet]. 2008;31(6):703–17.

48. Eicks M, Maurino V, Knappe S, Flügge U-I, Fischer K. The Plastidic Pentose Phosphate Translocator Represents a Link between the Cytosolic and the Plastidic Pentose Phosphate Pathways in Plants<sup>1</sup>. *Plant Physiol* [Internet]. 2002;128(2):512–22.
49. Moyersoen J, Choe J, Fan E, Hol WGJ, Michels PAM. Biogenesis of peroxisomes and glycosomes: trypanosomatid glycosome assembly is a promising new drug target. *FEMS Microbiol Rev* [Internet]. 2004 Nov 1 [cited 2017 Aug 12];28(5):603–43.
50. Maugeri DA, Cazzulo JJ. The pentose phosphate pathway in *Trypanosoma cruzi*. *FEMS Microbiol Lett* [Internet]. 2004;234(1):117–23.
51. Maugeri DA, Cazzulo JJ, Burchmore RJS, Barrett MP, Ogbunode POJ. Pentose phosphate metabolism in *Leishmania mexicana*. *Mol Biochem Parasitol* [Internet]. 2003 Aug [cited 2013 Mar 13];130(2):117–25.
52. Augustyns K, Amssoms K, Yamani A, Rajan PK, Haemers A. Trypanothione as a target in the design of antitrypanosomal and antileishmanial agents. [Review] [134 refs]. *Curr Pharm Des*. 2001;7(Netherlands PT-Journal Article PT-Review LG-English):1117–41.
53. Maugeri DA, Cannata JJB, Cazzulo JJ. Glucose metabolism in *Trypanosoma cruzi*. *Essays Biochem* [Internet]. 2011;51:15–30.
54. Stern AL, Naworyta A, Cazzulo JJ, Mowbray SL. Structures of type B ribose 5-phosphate isomerase from *Trypanosoma cruzi* shed light on the determinants of sugar specificity in the structural family. *FEBS J*. 2011;278(5):793–808.
55. Capriles PVSZ, Baptista LPR, Guedes IA, Guimarães ACR, Custodio FL, Alves-Ferreira M, et al. Structural modeling and docking studies of ribose 5-phosphate



- isomerase from *Leishmania major* and *Homo sapiens*: A comparative analysis for Leishmaniasis treatment. *J Mol Graph Model* [Internet]. 2015;55:134–47.
56. Loureiro I, Faria J, Clayton C, Macedo-Ribeiro S, Santarém N, Roy N, et al. Ribose 5-Phosphate Isomerase B Knockdown Compromises *Trypanosoma brucei* Bloodstream Form Infectivity. *PLoS Negl Trop Dis*. 2015;9(1):1–11.
  57. Jamdhade MD, Pawar H, Chavan S, Sathe G, Umasankar PK, Mahale KN, et al. Comprehensive Proteomics Analysis of Glycosomes from *Leishmania donovani*. *Omi A J Integr Biol* [Internet]. 2015 Mar [cited 2017 Mar 27];19(3):157–70.
  58. Zhang RG, Andersson CE, Savchenko A, Skarina T, Evdokimova E, Beasley S, et al. Structure of *Escherichia coli* ribose-5-phosphate isomerase: A ubiquitous enzyme of the pentose phosphate pathway and the Calvin cycle. *Structure*. 2003;11(1):31–42.
  59. Lobley CMC, Aller P, Douangamath A, Reddivari Y, Bumann M, Bird LE, et al. Structure of ribose 5-phosphate isomerase from the probiotic bacterium *Lactobacillus salivarius* UCC118. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2012;68(12):1427–33.
  60. Kim TG, Kwon TH, Min K, Dong M-S, Park YI, Ban C. Crystal structures of substrate and inhibitor complexes of ribose 5-phosphate isomerase A from *Vibrio vulnificus* YJ016. *Mol Cells* [Internet]. 2009 Jan 5 [cited 2017 Jan 27];27(1):99–103.
  61. Rault M, Giudici-Orticoni MT, Gontero B, Ricard J. Structural and functional properties of a multi-enzyme complex from spinach chloroplasts. 1. Stoichiometry of the polypeptide chains. *Eur J Biochem* [Internet]. 1993;217(3):1065–73.

62. Holmes MA, Buckner FS, Van Voorhis WC, Verlinde CLMJ, Mehlin C, Boni E, et al. Structure of ribose 5-phosphate isomerase from *Plasmodium falciparum*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* [Internet]. 2006 May 1 [cited 2016 Aug 11];62(5):427–31.
63. Ishikawa K, Matsui I, Payan F, Cambillau C, Ishida H, Kawarabayasi Y, et al. A hyperthermostable D-ribose-5-phosphate isomerase from *Pyrococcus horikoshii* characterization and three-dimensional structure. *Structure*. 2002;10(6):877–86.
64. Strange RW, Antonyuk S V., Ellis MJ, Bessho Y, Kuramitsu S, Yokoyama S, et al. The structure of an archaeal ribose-5-phosphate isomerase from *Methanocaldococcus jannaschii* (MJ1603). *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2009;65(12):1214–7.
65. Graille M, Meyer P, Leulliot N, Sorel I, Janin J, Van Tilbeurgh H, et al. Crystal structure of the *S. cerevisiae* D-ribose-5-phosphate isomerase: Comparison with the archaeal and bacterial enzymes. *Biochimie* [Internet]. 2005 [cited 2016 Aug 11];87(8):763–9.
66. Zhang RG, Andersson CE, Skarina T, Evdokimova E, Edwards AM, Joachimiak A, et al. The 2.2 Å resolution structure of RpiB/AlsB from *Escherichia coli* illustrates a new approach to the ribose-5-phosphate isomerase reaction. *J Mol Biol*. 2003;332(5):1083–94.
67. Edwards TE, Abramov AB, Smith ER, Baydo RO, Leonard JT, Leibly DJ, et al. Structural characterization of a ribose-5-phosphate isomerase B from the pathogenic fungus *Coccidioides immitis*. *BMC Struct Biol* [Internet]. 2011;11(1):39.
68. Jung J, Kim JK, Yeom SJ, Ahn YJ, Oh DK, Kang LW. Crystal structure of *Clostridium thermocellum* ribose-5-phosphate isomerase B reveals properties

- critical for fast enzyme kinetics. *Appl Microbiol Biotechnol*. 2011;90(2):517–27.
69. Roos AK, Andersson CE, Bergfors T, Jacobsson M, Karlén A, Unge T, et al. *Mycobacterium tuberculosis* ribose-5-phosphate isomerase has a known fold, but a novel active site. *J Mol Biol*. 2004;335(3):799–809.
  70. Stern AL, Burgos E, Salmon L, Cazzulo JJ. Ribose 5-phosphate isomerase type B from *Trypanosoma cruzi*: kinetic properties and site-directed mutagenesis reveal information about the reaction mechanism. *Biochem J* [Internet]. 2007;401(1):279–85.
  71. Hamada K, Ago H, Sugahara M, Nodake Y, Kuramitsu S, Miyano M. Oxyanion hole-stabilized stereospecific isomerization in ribose-5-phosphate isomerase (Rpi). *J Biol Chem* [Internet]. 2003 [cited 2016 Aug 15];278(49):49183–90.
  72. Jung CH, Hartman FC, Lu TY, Larimer FW. D-Ribose-5-Phosphate Isomerase from Spinach: Heterologous Overexpression, Purification, Characterization, and Site-Directed Mutagenesis of the Recombinant Enzyme. *Arch Biochem Biophys* [Internet]. 2000;373(2):409–17.
  73. Baptista LP. Modelagem Molecular da Ribose-5-fosfato isomerase de *Leishmania major* e de *Homo sapiens*: Predição de Estruturas e Estudos de Atracamento Molecular. Laboratório Nacional de Computação Científica; 2011.
  74. Rangarajan ES, Sivaraman J, Matte A, Cygler M. Crystal structure of D-ribose-5-phosphate isomerase (RpiA) from *Escherichia coli*. *Proteins Struct Funct Genet*. 2002;48(4):737–40.
  75. Kondo H, Nakamura Y, Dong Y-X, Nikawa J, Sueda S. Pyridoxine biosynthesis in yeast: participation of ribose 5-phosphate ketol-isomerase. *Biochem J* [Internet]. 2004 Apr 1 [cited 2017 Feb 4];379(Pt 1):65–70.

76. Roos AK, Burgos E, Ericsson DJ, Salmon L, Mowbray SL. Competitive inhibitors of *Mycobacterium tuberculosis* ribose-5-phosphate isomerase B reveal new information about the reaction mechanism. *J Biol Chem*. 2005;280(8):6416–22.
77. Kaur PK, Tripathi N, Desale J, Neelagiri S, Yadav S, Bharatam P V., et al. Mutational and structural analysis of conserved residues in ribose-5-phosphate isomerase B from *Leishmania donovani*: Role in substrate recognition and conformational stability. *PLoS One* [Internet]. 2016;11(3):e0150764.
78. Roos AK, Mariano S, Kowalinski E, Salmon L, Mowbray SL. d-Ribose-5-Phosphate Isomerase B from *Escherichia coli* is Also a Functional d-Allose-6-Phosphate Isomerase, While the *Mycobacterium tuberculosis* Enzyme is Not. *J Mol Biol*. 2008;382(3):667–79.
79. Yeom SJ, Kim BN, Park CS, Oh DK. Substrate specificity of ribose-5-phosphate isomerases from *Clostridium difficile* and *Thermotoga maritima*. *Biotechnol Lett*. 2010;32(6):829–35.
80. Yoon RY, Yeom SJ, Kim HJ, Oh DK. Novel substrates of a ribose-5-phosphate isomerase from *Clostridium thermocellum*. *J Biotechnol*. 2009;139(1):26–32.
81. Park C-S, Yeom S-J, Lim Y-R, Kim Y-S, Oh D-K. Substrate specificity of a recombinant ribose-5-phosphate isomerase from *Streptococcus pneumoniae* and its application in the production of l-lyxose and l-tagatose. *World J Microbiol Biotechnol* [Internet]. 2011 Apr 22 [cited 2017 Jan 27];27(4):743–50.
82. Skinner AJ, Cooper RA. The regulation of ribose-5-phosphate isomerisation in *Escherichia coli* K12. *FEBS Lett* [Internet]. 1971;12(5):293–6.
83. Miosga T, Zimmermann FK. Cloning and characterization of the first two genes of the non-oxidative part of the *Saccharomyces cerevisiae* pentose-phosphate

- pathway. *Curr Genet.* 1996;30(5):404–9.
84. Skinner AJ, Cooper RA. Genetic Studies on Ribose Mutants of *Escherichia coli* Ribose 5-Phosphate Isomerase Mutants of *Escherichia coli* K-12. *J Bacteriol.* 1974;118(3):1183–5.
  85. SØRENSEN KI, Hove-Jensen B. Ribose catabolism of *Escherichia coli*: Characterization of the *rpiB* gene encoding ribose phosphate isomerase B and of the *rpiR* gene, which is involved in regulation of *rpiB* expression. *J Bacteriol.* 1996;178(4):1003–11.
  86. Wamelink MMC, Grüning NM, Jansen EEW, Bluemlein K, Lehrach H, Jakobs C, et al. The difference between rare and exceptionally rare: Molecular characterization of ribose 5-phosphate isomerase deficiency. *J Mol Med.* 2010;88(9):931–9.
  87. Essenberg MK, Cooper R a. Two ribose-5-phosphate isomerases from *Escherichia coli* K12: partial characterisation of the enzymes and consideration of their possible physiological roles. *Eur J Biochem.* 1975;55(2):323–32.
  88. Woodruff W. Inhibition of Ribose-5-phosphate 4-Phosphoerythronate\* Isomerase. 1979;254(13):5866–7.
  89. Burgos E, Salmon L. Synthesis and evaluation of new 4-phospho-D-erythronic acid derivatives as competitive inhibitors of spinach ribose-5-phosphate isomerase. *Tetrahedron Lett.* 2004;45(4):753–6.
  90. Burgos E, Salmon L. Synthesis and kinetic evaluation of 4-deoxy-4-phosphonomethyl-D- erythronate, the first hydrolytically stable and potent competitive inhibitor of ribose-5-phosphate isomerase. *Tetrahedron Lett.* 2004;45(17):3465–9.

91. Horitsu H, Sasaki I, Kikuchi T, Suzuki H, Sumida M, Tomoyeda M. Purification, Properties, and Structure of Ribose 5-Phosphate Ketol Isomerase from *Candida utilis*. *Agric Biol Chem*. 1976;40(2):257–64.
92. Mariano S, Roos AK, Mowbray SL, Salmon L. Competitive inhibitors of type B ribose 5-phosphate isomerases: design, synthesis and kinetic evaluation of new d-allose and d-allulose 6-phosphate derivatives. *Carbohydr Res* [Internet]. 2009;344(7):869–80.
93. Bielska E, Lucas X, Czerwoniec A, Kasprzak JM, Kaminska KH, Bujnicki JM. Virtual screening strategies in drug design - methods and applications. *Biotechnologia*. 2011;92(3):249–64.
94. Muller J, Hemphill A. Drug target identification in protozoan parasites. *Expert Opin Drug Discov* [Internet]. 2016;11(June):2029–38.
95. Bleicher KH, Böhm H-J, Müller K, Alanine AI. A guide to drug discovery: Hit and lead generation: beyond high-throughput screening. *Nat Rev Drug Discov* [Internet]. 2003 May [cited 2016 Jun 7];2(5):369–78.
96. Hopkins AL, Groom CR. The druggable genome. *Nat Rev Drug Discov* [Internet]. 2002;1(9):727–30.
97. Bull SC, Doig AJ. Properties of protein drug target classes. *PLoS One* [Internet]. 2015 [cited 2017 Aug 10];10(3):e0117955.
98. Schenone M, Dančík V, Wagner BK, Clemons P a. Target identification and mechanism of action in chemical biology and drug discovery. *Nat Chem Biol* [Internet]. 2013;9(4):232–40.

99. Cerqueira NMFSA, Gesto D, Oliveira EF, Santos-Martins D, Brás NF, Sousa SF, et al. Receptor-based virtual screening protocol for drug discovery. *Arch Biochem Biophys* [Internet]. 2015;582:56–67.
100. Leelananda SP, Lindert S. Computational methods in drug discovery. *Beilstein J Org Chem* [Internet]. 2016 [cited 2017 Apr 1];12:2694–718.
101. Katsila T, Spyroulias GA, Patrinos GP, Matsoukas M-T. Computational approaches in target identification and drug discovery. *Comput Struct Biotechnol J* [Internet]. 2016 [cited 2017 Apr 1];14:177–84.
102. Sliwoski G, Kothiwale S, Meiler J, Lowe E. Computational methods in drug discovery. *Pharmacol Rev* [Internet]. 2014;66(1):334–95.
103. Ferreira L, dos Santos R, Oliva G, Andricopulo A. Molecular Docking and Structure-Based Drug Design Strategies. *Molecules* [Internet]. 2015;20(7):13384–421.
104. Lionta E, Spyrou G, Vassilatis DK, Cournia Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem* [Internet]. 2014;14(16):1923–38.
105. Güner OF, Bowen JP. Setting the record straight: The origin of the pharmacophore concept. *J Chem Inf Model*. 2014;54(5):1269–83.
106. Wermuth CG, Ganellin CR, Lindberg P, Mitscher LA. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl Chem* [Internet]. 1998 Jan 1 [cited 2017 Apr 4];70(5):1129–43.
107. Kaserer T, Beck KR, Akram M, Odermatt A, Schuster D, Willett P.

- Pharmacophore models and pharmacophore-based virtual screening: Concepts and applications exemplified on hydroxysteroid dehydrogenases. *Molecules*. 2015;20(12):22799–832.
108. Yang S-Y. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov Today* [Internet]. 2010;15(11–12):444–50.
  109. Gaurav A, Gautam V. Structure-based three-dimensional pharmacophores as an alternative to traditional methodologies. *J Receptor Ligand Channel Res*. 2014;7:27–38.
  110. Gao Q, Yang L, Zhu Y. Pharmacophore based drug design approach as a practical process in drug discovery. *Curr Comput Aided Drug Des*. 2010;6(1):37–49.
  111. Kitchen D, Decornez H, Furr J, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov*. 2004;3(11):935–49.
  112. Guedes IA, Magalhães CS de, Dardenne LE. *Bioinformática Da Biologia À Flexibilidade Molecular*. São Paulo: Sociedade Brasileira de Bioquímica e Biologia Molecular; 2014. 189-208 p.
  113. Grinter SZ, Zou X. Challenges, applications, and recent advances of protein-ligand docking in structure-based drug design. *Molecules*. 2014;19(7):10150–76.
  114. Durrant JD, McCammon JA. Molecular dynamics simulations and drug discovery. *BMC Biol* [Internet]. 2011 Oct 28 [cited 2017 Apr 7];9:71.



115. Hansson T, Oostenbrink C, van Gunsteren WF. Molecular dynamics simulations. *Curr Opin Struct Biol* [Internet]. 2002;12:190.
116. Perilla JR, Goh BC, Cassidy CK, Liu B, Bernardi RC, Rudack T, et al. Molecular dynamics simulations of large macromolecular complexes. *Curr Opin Struct Biol* [Internet]. 2015 Apr [cited 2017 Apr 7];31:64–74.
117. Mortier J, Rakers C, Bermudez M, Murgueitio MS, Riniker S, Wolber G. The impact of molecular dynamics on drug design: Applications for the characterization of ligand-macromolecule complexes. *Drug Discov Today*. 2015;20(6):686–702.
118. Adcock SA, McCammon JA. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem Rev* [Internet]. 2006 May [cited 2017 Apr 7];106(5):1589–615.
119. Namba AM, Da Silva VB, Da Silva CHTP. Dinâmica molecular: Teoria e aplicações em planejamento de fármacos. *Eclét Quim*. 2008;33(4):13–24.
120. Abraham M, Hess B, Spoel D van der, Lindahl E. GROMACS User Manual version 5.0.7. [www.gromacs.org](http://www.gromacs.org). 2015;
121. Verli H. *Bioinformática da Biologia à Flexibilidade Molecular*. São Paulo: Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq; 2014. 173-187 p.
122. Wang Y, Xing J, Xu Y, Zhou N, Peng J, Xiong Z, et al. In silico ADME/T modelling for rational drug design. *Q Rev Biophys* [Internet]. 2015;48(4):488–515.
123. Moroy G, Martiny VY, Vayer P, Villoutreix BO, Miteva MA. Toward in silico

- structure-based ADMET prediction in drug discovery. *Drug Discov Today* [Internet]. 2012;17(1–2):44–55.
124. Tian S, Wang J, Li Y, Li D, Xu L, Hou T. The application of in silico drug-likeness predictions in pharmaceutical research. *Adv Drug Deliv Rev* [Internet]. 2015;86:2–10.
  125. Patrick GL. An Introduction to Medicinal Chemistry. UCL Lect notes. 2009;40:752.
  126. Lagorce D, Douguet D, Miteva MA, Villoutreix BO. Computational analysis of calculated physicochemical and ADMET properties of protein-protein interaction inhibitors. *Sci Rep* [Internet]. 2017;7(December 2016):46277.
  127. van de Waterbeemd H, Gifford E. ADMET in silico modelling: Towards prediction paradise? *Nat Rev Drug Discov*. 2003;2(3):192–204.
  128. Alqahtani S. *In silico* ADME-Tox modeling: progress and prospects. *Expert Opin Drug Metab Toxicol* [Internet]. 2017;0(0):1–12.
  129. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* [Internet]. 1993 Dec 5;234(3):779–815.
  130. Custodio FL, Barbosa HJC, Dardenne LE. A multiple minima genetic algorithm for protein structure prediction. *Appl Soft Comput J* [Internet]. 2014 [cited 2017 May 6];15:88–99.
  131. Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. *J Chem Inf Model* [Internet]. [cited 2016 May 17];45(1):177–82.

132. Zhang Y, Skolnick J. TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* 2005;33(7):2302–9.
133. Schneidman-Duhovny D, Dror O, Inbar Y, Nussinov R, Wolfson HJ. PharmaGist: a webserver for ligand-based pharmacophore detection. *Nucleic Acids Res [Internet]*. 2008 Jul 1 [cited 2016 May 17];36(Web Server issue):223–8.
134. Koes DR, Camacho CJ. ZINCPharmer: pharmacophore search of the ZINC database. *Nucleic Acids Res [Internet]*. 2012 Jul [cited 2016 May 17];40(W1):W409–14.
135. Koes DR, Camacho CJ. Pharmer: Efficient and exact pharmacophore search. *J Chem Inf Model.* 2011;51:1307–14.
136. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* 2001;46(1):3–26.
137. Berthold MR, Cebon N, Dill F, Gabriel TR, Kötter T, Meinel T, et al. KNIME - The Konstanz Information Miner. *SIGKDD Explor [Internet]*. 2009;11(1):26–31.
138. Morris GM, Ruth H, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem.* 2009;30(16):2785–91.
139. Trott O, Olson AJ. AutoDock Vina. *J Comput Chem.* 2010;31:445–61.
140. Friesner R a., Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren T a., et al. Extra precision glide: Docking and scoring incorporating a model of

- hydrophobic enclosure for protein-ligand complexes. *J Med Chem.* 2006;49(21):6177–96.
141. Li H, Robertson AD, Jensen JH. Very fast empirical prediction and rationalization of protein pK<sub>a</sub> values. *Proteins Struct Funct Genet.* 2005;61(4):704–21.
142. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004;25(13):1605–12.
143. Handoko SD, Ouyang X, Tran C, Su T, Kwoh CK, Ong YS. QuickVina: Accelerating AutoDock Vina Using Gradient-Based Heuristics for Global Optimization. [cited 2017 May 7];
144. Lagorce D, Sperandio O, Baell JB, Miteva MA, Villoutreix BO. FAF-Drugs3: a web server for compound property calculation and chemical library design. *Nucleic Acids Res.* 2015;43(W1):W200-7.
145. Kennedy T. Managing the drug discovery/development interface. *Drug Discov Today* [Internet]. 1997 Oct [cited 2017 May 7];2(10):436–44.
146. Schäfer T, Kriege N, Humbeck L, Klein K, Koch O, Mutzel P. Scaffold Hunter: a comprehensive visual analytics framework for drug discovery. *J Cheminform* [Internet]. 2017;9(1):28.
147. Proctor EA, Yin S, Tropsha A, Dokholyan N V. Discrete molecular dynamics distinguishes natively-like binding poses from decoys in difficult targets. *Biophys J.* 2012;102(1):144–51.
148. Koziara KB, Stroet M, Malde AK, Mark AE. Testing and validation of the

Automated Topology Builder (ATB) version 2.0: Prediction of hydration free enthalpies. *J Comput Aided Mol Des*. 2014;28(3):221–33.

149. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. GROMACS: Fast, flexible, and free. Vol. 26, *Journal of Computational Chemistry*. 2005. p. 1701–18.
150. Darden T, York D, Pedersen L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J Chem Phys [Internet]*. 1993;98(12):10089.
151. Zhao FJ, Zhang S, Yu ZM, Xia SJ, Li H. Specific targeting of prostate cancer cells in vitro by the suicide gene/prodrug system, uracil phosphoribosyltransferase/5-fluorouracil, under the control of prostate-specific membrane antigen promoter/enhancer. *Prostate Cancer Prostatic Dis [Internet]*. 2009;12(2):166–71.
152. Lenselink EB, Beuming T, Sherman W, Van Vlijmen HWT, Ijzerman AP. Selecting an optimal number of binding site waters to improve virtual screening enrichments against the adenosine A2A receptor. *J Chem Inf Model*. 2014;54(6):1737–46.
153. Huang N, Shoichet BK. Exploiting Ordered Waters in Molecular Docking. *J Med Chem [Internet]*. 2008 Aug [cited 2017 May 31];51(16):4862–5.
154. García-Sosa AT. Hydration Properties of Ligands and Drugs in Protein Binding Sites: Tightly-Bound, Bridging Water Molecules and Their Effects and Consequences on Molecular Design Strategies. *J Chem Inf Model [Internet]*. 2013 Jun 24 [cited 2017 May 31];53(6):1388–405.
155. Wang Z, Sun H, Yao X, Li D, Xu L, Li Y, et al. Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction

accuracy of sampling power and scoring power. *Phys Chem Chem Phys* [Internet]. 2016;18(18):12964–75.

156. Triballeau N, Acher F, Brabet I, Pin J-P, Bertrand H-O. Virtual Screening Workflow Development Guided by the “Receiver Operating Characteristic” Curve Approach. Application to High-Throughput Docking on Metabotropic Glutamate Receptor Subtype 4. *J Med Chem* [Internet]. 2005 Apr [cited 2017 Apr 14];48(7):2534–47.
157. Cauët E, Rooman M, Wintjens R, Liévin J, Biot C. Histidine-aromatic interactions in proteins and protein-ligand complexes: Quantum chemical study of X-ray and model structures. *J Chem Theory Comput.* 2005;1(3):472–83.
158. Churchill CDM, Wetmore SD. Noncovalent interactions involving histidine: The effect of charge on  $\pi$ - $\pi$  stacking and T-shaped interactions with the DNA nucleobases. *J Phys Chem B.* 2009;113(49):16046–58.
159. Kawasaki Y. Mizoribine: a new approach in the treatment of renal disease. *Clin Dev Immunol* [Internet]. 2009;2009:681482.
160. Cadman EC, Dix DE, Handschumacher RE. Clinical, biological and biochemical effects of pyrazofurin. *Cancer Res* [Internet]. 1978;38(3):682–8.
161. Shara M, Stohs SJ. Efficacy and Safety of White Willow Bark (*Salix alba*) Extracts. Vol. 29, *Phytotherapy Research*. 2015. p. 1112–6.
162. Kalgutkar AS, Jones R, Sawant A. Sulfonamide as an essential functional group in drug design. Metabolism, pharmacokinetics and toxicity of functional groups. RSC drug discovery series. 2010.

163. Onufriev A. Implicit solvent models in molecular dynamics simulations: A brief overview. *Annu Rep Comput Chem*. 2008;
  
164. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J Med Chem* [Internet]. 2012 Jul 26 [cited 2017 Jun 29];55(14):6582–94.

## 7 APÊNDICES

### Apêndice A - DESCRIÇÕES DAS PROPRIEDADES DOS 211 POTENCIAIS INIBIDORES DE TCRPIB

**Tabela 7.1:** Propriedades ADME previstas, filtros PAINS e DS dos 211 compostos finais para TcRpiB em relação a HsRpiA e seus respectivos grupos (clusters). Destacados em negrito estão os compostos representativos de cada cluster

COMPOSTOS	DS TCRPI	DS HSRPI	LOGP	HD	HA	LV	OB VEBER	OB EGAN	PAINS	CLUSTER	ENDÓGENO HUMANOS	NOMES DE ENDÓGENOS
<b>ZINC00895164</b>	-9.615	-8.368	-2.37	5.0	6.0	0.0	Bom	Bom	Aceito	1	Não	Não se aplica
ZINC00901628	-8.99	-7.626	-4.49	5.0	6.0	0.0	Bom	Bom	Aceito	1	Não	Não se aplica
ZINC03100658	-9.012	-8.004	-1.64	4.0	8.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC04115754	-9.597	-7.568	0.02	2.0	8.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC05641501	-9.081	-6.978	-3.81	5.0	9.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC20255238	-9.177	-7.725	1.3	3.0	7.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC35085609	-9.017	-7.605	1.4	3.0	7.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
<b>ZINC36983314</b>	-9.923	-7.08	-0.08	3.0	6.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC36983325	-9.096	-8.237	-1.87	5.0	10.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC64874461	-8.969	-8.829	0.5	3.0	6.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC77301044	-8.947	-8.709	-0.18	3.0	5.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC77313266	-9.081	-8.377	-0.76	3.0	6.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC85971925	-8.927	-6.676	-2.11	4.0	5.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC91996777	-9.078	-8.201	-0.61	3.0	5.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC92890194	-9.691	-7.903	-0.1	2.0	7.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC94610447	-9.392	-5.934	0.66	2.0	5.0	0.0	Bom	Bom	Aceito	2	Não	Não se aplica
ZINC04476560	-9.518	-8.301	-0.09	5.0	7.0	0.0	Bom	Bom	Aceito	3	Não	Não se aplica
ZINC20534732	-9.071	-7.507	-1.31	2.0	7.0	0.0	Bom	Bom	Aceito	3	Não	Não se aplica
<b>ZINC31776002</b>	-10.631	-7.653	-1.38	4.0	8.0	0.0	Bom	Bom	Aceito	3	Não	Não se aplica



COMPOSTOS	DS TCRPI	DS HSRPI	LOGP	HD	HA	LV	OB VEBER	OB EGAN	PAINS	CLUSTER	ENDÓGENO HUMANOS	NOMES DE ENDÓGENOS
ZINC87440332	-9.662	-6.664	-1.71	3.0	9.0	0.0	Bom	Bom	Aceito	3	Não	Não se aplica
ZINC87445535	-8.966	-7.794	0.18	1.0	7.0	0.0	Bom	Bom	Aceito	3	Não	Não se aplica
ZINC94221752	-9.136	-5.884	0.57	4.0	7.0	0.0	Bom	Bom	Aceito	3	Não	Não se aplica
ZINC05831103	-8.919	-7.485	-1.36	3.0	5.0	0.0	Bom	Bom	Aceito	4	Não	Não se aplica
<b>ZINC09970438</b>	-10.281	-7.602	-5.05	5.0	7.0	0.0	Bom	Bom	Aceito	4	Não	Não se aplica
ZINC21985229	-9.387	-8.722	-4.84	3.0	8.0	0.0	Bom	Bom	Aceito	4	Não	Não se aplica
ZINC83407208	-9.909	-7.937	-4.35	5.0	10.0	0.0	Bom	Bom	Aceito	4	Não	Não se aplica
ZINC11592857	-8.957	-7.613	0.18	4.0	8.0	0.0	Bom	Bom	Aceito	5	Não	Não se aplica
ZINC13507425	-9.413	-8.129	3.82	3.0	7.0	0.0	Bom	Bom	Aceito	5	Não	Não se aplica
<b>ZINC13761105</b>	-9.487	-7.482	4.26	3.0	6.0	0.0	Bom	Bom	Aceito	5	Não	Não se aplica
ZINC72319957	-9.009	-7.146	2.03	3.0	5.0	0.0	Bom	Bom	Aceito	5	Não	Não se aplica
<b>ZINC09694454</b>	-9.709	-6.013	-1.73	3.0	9.0	0.0	Bom	Bom	Aceito	6	Não	Não se aplica
ZINC20160154	-8.931	-6.688	-4.05	4.0	7.0	0.0	Bom	Bom	Aceito	6	Não	Não se aplica
ZINC32628321	-9.083	-7.603	0.51	4.0	7.0	0.0	Bom	Bom	Aceito	6	Não	Não se aplica
ZINC37449495	-9.01	-8.727	-0.87	5.0	8.0	0.0	Bom	Bom	Aceito	6	Não	Não se aplica
ZINC86433775	-8.981	-8.757	-4.03	3.0	7.0	0.0	Bom	Bom	Aceito	6	Não	Não se aplica
ZINC87311798	-9.205	-7.882	-1.64	4.0	7.0	0.0	Bom	Bom	Aceito	6	Não	Não se aplica
ZINC92984751	-9.152	-7.896	1.65	3.0	6.0	0.0	Bom	Bom	Aceito	6	Não	Não se aplica
ZINC92984768	-9.143	-8.882	-1.26	3.0	9.0	0.0	Bom	Bom	Aceito	6	Não	Não se aplica
ZINC93229808	-9.137	-8.882	-1.77	3.0	8.0	0.0	Bom	Bom	Aceito	6	Não	Não se aplica
ZINC36971700	-9.072	-7.384	-1.91	6.0	8.0	1.0	Bom	Bom	Aceito	7	Não	Não se aplica
<b>ZINC36975961</b>	-12.735	-8.877	-2.03	3.0	8.0	0.0	Bom	Bom	Aceito	7	Não	Não se aplica
ZINC37431651	-9.95	-7.59	-0.2	4.0	8.0	0.0	Bom	Bom	Aceito	7	Não	Não se aplica
ZINC37437566	-9.568	-6.185	0.92	2.0	7.0	0.0	Bom	Bom	Aceito	7	Não	Não se aplica
ZINC37449060	-9.432	-8.134	-2.19	3.0	9.0	0.0	Bom	Bom	Aceito	7	Não	Não se aplica
ZINC37449313	-9.086	-8.516	-1.96	2.0	7.0	0.0	Bom	Bom	Aceito	7	Não	Não se aplica
ZINC37449519	-9.11	-8.761	-1.84	2.0	8.0	0.0	Bom	Bom	Aceito	7	Não	Não se aplica

COMPOSTOS	DS TCRPI	DS HSRPI	LOGP	HD	HA	LV	OB VEBER	OB EGAN	PAINS	CLUSTER	ENDÓGENO HUMANOS	NOMES DE ENDÓGENOS
ZINC37451246	-10.031	-7.024	-0.64	3.0	8.0	0.0	Bom	Bom	Aceito	7	Não	Não se aplica
ZINC92435320	-9.957	-8.843	-0.46	2.0	7.0	0.0	Bom	Bom	Aceito	7	Não	Não se aplica
ZINC94236295	-9.404	-7.15	-1.21	3.0	5.0	0.0	Bom	Bom	Aceito	7	Não	Não se aplica
ZINC01576100	-9.206	-8.764	0.42	5.0	9.0	0.0	Low	Bom	Aceito	8	Não	Não se aplica
ZINC16478605	-9.075	-8.956	2.55	4.0	7.0	0.0	Bom	Bom	Aceito	8	Não	Não se aplica
ZINC19799225	-9.394	-7.847	0.99	5.0	9.0	0.0	Bom	Bom	Aceito	8	Não	Não se aplica
<b>ZINC20572839</b>	-9.666	-7.912	2.69	3.0	8.0	0.0	Bom	Bom	Aceito	8	Não	Não se aplica
ZINC23187566	-9.262	-8.051	-1.99	2.0	7.0	0.0	Bom	Bom	Aceito	8	Não	Não se aplica
ZINC37451146	-9.024	-5.921	-1.05	6.0	8.0	1.0	Bom	Bom	Aceito	8	Não	Não se aplica
ZINC64219433	-9.535	-7.359	-0.92	5.0	9.0	0.0	Bom	Bom	Aceito	8	Não	Não se aplica
ZINC75278199	-9.487	-7.902	0.73	2.0	10.0	0.0	Bom	Bom	Aceito	8	Não	Não se aplica
ZINC92493836	-9.343	-7.575	-2.32	3.0	6.0	0.0	Bom	Bom	Aceito	8	Não	Não se aplica
ZINC09575936	-9.47	-6.373	-0.16	4.0	8.0	0.0	Bom	Bom	Aceito	9	Não	Não se aplica
ZINC09577104	-10.521	-5.693	-0.59	4.0	8.0	0.0	Bom	Bom	Aceito	9	Não	Não se aplica
<b>ZINC13552852</b>	-10.906	-6.5	-0.16	4.0	8.0	0.0	Bom	Bom	Aceito	9	Não	Não se aplica
ZINC13654785	-10.652	-5.747	-0.16	4.0	8.0	0.0	Bom	Bom	Aceito	9	Não	Não se aplica
ZINC05085274	-9.328	-8.801	-1.54	6.0	9.0	1.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC05479266	-9.506	-8.451	2.46	4.0	8.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC13213453	-9.79	-8.71	-4.73	6.0	9.0	1.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC13549482	-9.385	-8.246	-2.14	5.0	12.0	1.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC20611195	-9.292	-6.932	-1.96	6.0	10.0	1.0	Bom	Bom	Aceito	10	Não	Não se aplica
<b>ZINC36972227</b>	-10.182	-8.596	-0.81	4.0	8.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC36972227	-9.153	-8.96	-0.81	4.0	8.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC37448985	-9.117	-8.76	1.27	3.0	9.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC38549580	-9.414	-6.829	-1.81	3.0	9.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC39275476	-9.246	-7.656	-1.41	2.0	9.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC44963783	-9.15	-7.002	-1.03	4.0	8.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica

COMPOSTOS	DS TCRPI	DS HSRPI	LOGP	HD	HA	LV	OB VEBER	OB EGAN	PAINS	CLUSTER	ENDÓGENO HUMANOS	NOMES DE ENDÓGENOS
ZINC67695651	-9.015	-7.692	-2.39	3.0	9.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC67695654	-9.534	-8.037	-2.39	3.0	9.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC92177362	-9.87	-8.55	-1.17	4.0	8.0	0.0	Bom	Bom	Aceito	10	Não	Não se aplica
ZINC38617785	-8.98	-8.728	1.56	3.0	7.0	0.0	Bom	Bom	Aceito	11	Não	Não se aplica
ZINC63314483	-9.02	-6.822	-2.75	2.0	8.0	0.0	Bom	Bom	Aceito	11	Não	Não se aplica
ZINC63314487	-8.924	-7.258	-2.75	2.0	8.0	0.0	Bom	Bom	Aceito	11	Não	Não se aplica
<b>ZINC63480117</b>	<b>-11.927</b>	<b>-8.503</b>	<b>-0.05</b>	<b>3.0</b>	<b>9.0</b>	<b>0.0</b>	<b>Bom</b>	<b>Bom</b>	<b>Aceito</b>	<b>11</b>	<b>Não</b>	<b>Não se aplica</b>
ZINC63761405	-9.261	-7.981	-1.67	5.0	9.0	0.0	Bom	Bom	Aceito	11	Não	Não se aplica
ZINC04616624	-9.096	-6.724	1.47	3.0	7.0	0.0	Bom	Bom	Aceito	12	Não	Não se aplica
ZINC09167236	-8.976	-7.564	0.88	1.0	7.0	0.0	Bom	Bom	Aceito	12	Não	Não se aplica
ZINC09969825	-9.513	-6.622	0.58	3.0	7.0	0.0	Bom	Bom	Aceito	12	Não	Não se aplica
<b>ZINC20153858</b>	<b>-9.533</b>	<b>-7.131</b>	<b>0.57</b>	<b>1.0</b>	<b>8.0</b>	<b>0.0</b>	<b>Bom</b>	<b>Bom</b>	<b>Aceito</b>	<b>12</b>	<b>Não</b>	<b>Não se aplica</b>
ZINC87131155	-9.009	-7.563	1.26	2.0	5.0	0.0	Bom	Bom	Aceito	13	Não	Não se aplica
<b>ZINC91306114</b>	<b>-9.437</b>	<b>-6.815</b>	<b>1.46</b>	<b>2.0</b>	<b>6.0</b>	<b>0.0</b>	<b>Bom</b>	<b>Bom</b>	<b>Aceito</b>	<b>13</b>	<b>Não</b>	<b>Não se aplica</b>
ZINC91998368	-9.012	-6.407	-0.32	2.0	7.0	0.0	Bom	Bom	Aceito	13	Não	Não se aplica
ZINC94654182	-9.083	-7.632	1.94	2.0	5.0	0.0	Bom	Bom	Aceito	13	Não	Não se aplica
ZINC06223189	-9.864	-7.255	-0.61	3.0	6.0	0.0	Bom	Bom	Aceito	14	Não	Não se aplica
ZINC11804506	-8.921	-8.936	0.37	3.0	7.0	0.0	Bom	Bom	Aceito	14	Não	Não se aplica
ZINC34957425	-8.952	-8.514	1.86	1.0	7.0	0.0	Bom	Bom	Aceito	14	Não	Não se aplica
ZINC36983462	-9.239	-8.929	-0.82	1.0	9.0	0.0	Bom	Bom	Aceito	14	Não	Não se aplica
<b>ZINC37431881</b>	<b>-10.58</b>	<b>-7.878</b>	<b>-0.39</b>	<b>4.0</b>	<b>8.0</b>	<b>0.0</b>	<b>Bom</b>	<b>Bom</b>	<b>Aceito</b>	<b>14</b>	<b>Não</b>	<b>Não se aplica</b>
ZINC60360593	-9.858	-7.536	3.42	3.0	7.0	0.0	Bom	Bom	Aceito	14	Não	Não se aplica
ZINC83323577	-9.821	-6.613	-2.67	4.0	6.0	0.0	Bom	Bom	Aceito	14	Não	Não se aplica
ZINC86860377	-9.496	-8.439	-0.28	5.0	7.0	0.0	Bom	Bom	Aceito	14	Não	Não se aplica
ZINC92325932	-9.769	-8.613	0.18	3.0	9.0	0.0	Bom	Bom	Aceito	14	Não	Não se aplica
ZINC92391774	-9.196	-8.418	-0.25	4.0	8.0	0.0	Bom	Bom	Aceito	14	Não	Não se aplica
ZINC06261710	-10.712	-8.111	-0.7	3.0	8.0	0.0	Bom	Bom	Aceito	15	Não	Não se aplica

COMPOSTOS	DS TCRPI	DS HSRPI	LOGP	HD	HA	LV	OB VEBER	OB EGAN	PAINS	CLUSTER	ENDÓGENO HUMANOS	NOMES DE ENDÓGENOS
ZINC06261710	-9.53	-8.438	-0.7	3.0	8.0	0.0	Bom	Bom	Aceito	15	Não	Não se aplica
ZINC08996391	-9.322	-8.616	-1.39	4.0	10.0	0.0	Bom	Bom	Aceito	15	Não	Não se aplica
ZINC08996391	-11.001	-6.47	-1.39	4.0	10.0	0.0	Bom	Bom	Aceito	15	Não	Não se aplica
<b>ZINC04855421</b>	-9.242	-5.171	1.88	3.0	7.0	0.0	Bom	Bom	Aceito	16	Não	Não se aplica
ZINC04855423	-9.57	-3.491	1.49	3.0	8.0	0.0	Bom	Bom	Aceito	16	Não	Não se aplica
ZINC04855425	-9.892	-8.068	1.49	3.0	8.0	0.0	Bom	Bom	Aceito	16	Não	Não se aplica
ZINC04957420	-8.962	-6.546	1.88	3.0	7.0	0.0	Bom	Bom	Aceito	16	Não	Não se aplica
ZINC05042523	-9.407	-3.509	2.25	3.0	7.0	0.0	Bom	Bom	Aceito	16	Não	Não se aplica
<b>ZINC09860934</b>	-9.746	-7.001	1.88	3.0	7.0	0.0	Bom	Bom	Aceito	16	Não	Não se aplica
ZINC00002053	-10.549	-7.927	-1.41	5.0	7.0	0.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC05225545	-9.732	-8.554	0.76	5.0	8.0	0.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC15206550	-9.64	-8.309	-1.05	5.0	10.0	0.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC26977303	-8.948	-8.277	1.33	4.0	9.0	0.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC31160567	-8.936	-7.559	1.08	6.0	12.0	2.0	Bom	Bom	Aceito	17	Não	Não se aplica
<b>ZINC35643607</b>	-11.109	-6.913	0.37	4.0	7.0	0.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC35644601	-9.653	-4.859	1.91	5.0	8.0	0.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC39937400	-9.147	-7.923	-3.55	6.0	9.0	1.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC63698176	-10.52	-8.742	1.0	2.0	7.0	0.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC65748606	-9.894	-8.776	1.57	6.0	9.0	1.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC93120380	-9.013	-8.053	1.24	2.0	5.0	0.0	Bom	Bom	Aceito	17	Não	Não se aplica
ZINC93120429	-9.232	-7.109	1.57	2.0	4.0	0.0	Bom	Bom	Aceito	17	Não	Não se aplica
<b>ZINC35457690</b>	-10.558	-8.915	-0.87	5.0	10.0	0.0	Bom	Bom	Aceito	18	Não	Não se aplica
ZINC35644136	-10.428	-8.176	-1.05	4.0	9.0	0.0	Bom	Bom	Aceito	18	Não	Não se aplica
ZINC62722512	-9.012	-7.36	-1.69	3.0	5.0	0.0	Bom	Bom	Aceito	18	Não	Não se aplica
ZINC77285380	-9.389	-8.899	-0.33	6.0	10.0	1.0	Bom	Bom	Aceito	18	Não	Não se aplica
ZINC93114422	-9.645	-8.143	2.53	2.0	4.0	0.0	Bom	Bom	Aceito	18	Não	Não se aplica
ZINC93114518	-9.098	-8.532	1.74	2.0	4.0	0.0	Bom	Bom	Aceito	18	Não	Não se aplica

COMPOSTOS	DS TCRPI	DS HSRPI	LOGP	HD	HA	LV	OB VEBER	OB EGAN	PAINS	CLUSTER	ENDÓGENO HUMANOS	NOMES DE ENDÓGENOS
ZINC04580783	-8.941	-8.193	0.98	5.0	7.0	0.0	Bom	Bom	Aceito	19	Não	Não se aplica
ZINC05356198	-9.603	-8.404	0.35	5.0	7.0	0.0	Bom	Bom	Aceito	19	Não	Não se aplica
<b>ZINC34924542</b>	-11.383	-8.671	-0.08	5.0	7.0	0.0	Bom	Bom	Aceito	19	Não	Não se aplica
ZINC00005848	-8.932	-6.534	-0.87	5.0	9.0	0.0	Bom	Bom	Aceito	20	Não	Não se aplica
ZINC03872749	-9.33	-8.917	-3.0	7.0	13.0	2.0	Bom	Bom	Aceito	20	Não	Não se aplica
ZINC03872751	-9.909	-8.306	-3.0	7.0	13.0	2.0	Bom	Bom	Aceito	20	Não	Não se aplica
ZINC05224737	-9.978	-8.8	-3.0	7.0	13.0	2.0	Bom	Bom	Aceito	20	Não	Não se aplica
<b>ZINC14963248</b>	-10.079	-7.083	-2.78	6.0	12.0	2.0	Bom	Bom	Aceito	20	Não	Não se aplica
ZINC14963248	-9.739	-7.767	-2.78	6.0	12.0	2.0	Bom	Bom	Aceito	20	Não	Não se aplica
ZINC36521966	-9.816	-7.602	-1.21	5.0	9.0	0.0	Bom	Bom	Aceito	20	Não	Não se aplica
ZINC38600444	-9.544	-4.882	-0.88	6.0	10.0	1.0	Bom	Bom	Aceito	20	Não	Não se aplica
ZINC04533489	-9.179	-8.316	-1.32	3.0	11.0	1.0	Bom	Bom	Aceito	21	Não	Não se aplica
<b>ZINC04533499</b>	-9.824	-8.957	-2.29	4.0	12.0	1.0	Bom	Bom	Aceito	21	Não	Não se aplica
ZINC16939884	-9.226	-8.249	-1.56	4.0	10.0	0.0	Bom	Bom	Aceito	21	Não	Não se aplica
ZINC22013293	-9.054	-8.259	-2.01	2.0	9.0	0.0	Bom	Bom	Aceito	21	Não	Não se aplica
ZINC79272843	-9.08	-7.719	-2.22	5.0	11.0	1.0	Bom	Bom	Aceito	21	Não	Não se aplica
ZINC85210532	-9.252	-8.723	-0.58	2.0	9.0	0.0	Bom	Bom	Aceito	21	Não	Não se aplica
ZINC01610270	-9.445	-8.077	-1.7	3.0	9.0	0.0	Bom	Bom	Aceito	22	Não	Não se aplica
ZINC12503903	-9.691	-8.65	-2.81	4.0	10.0	0.0	Bom	Bom	Aceito	22	Não	Não se aplica
<b>ZINC24980922</b>	-9.698	-7.75	-2.12	5.0	11.0	1.0	Bom	Bom	Aceito	22	Não	Não se aplica
ZINC01611080	-9.502	-8.869	-2.4	4.0	9.0	0.0	Bom	Bom	Aceito	23	Não	Não se aplica
ZINC01736046	-9.211	-7.581	-4.16	6.0	11.0	2.0	Bom	Bom	Aceito	23	Não	Não se aplica
ZINC04096130	-9.435	-8.846	-2.93	5.0	10.0	0.0	Bom	Bom	Aceito	23	Não	Não se aplica
ZINC04535606	-9.638	-8.599	-4.16	6.0	11.0	2.0	Bom	Bom	Aceito	23	Não	Não se aplica
<b>ZINC17027397</b>	-9.856	-8.954	-1.89	4.0	11.0	1.0	Bom	Bom	Aceito	23	Não	Não se aplica
ZINC04557020	-9.36	-5.637	-2.4	4.0	6.0	0.0	Bom	Bom	Aceito	24	Não	Não se aplica
ZINC22063668	-9.013	-5.345	-2.03	3.0	5.0	0.0	Bom	Bom	Aceito	24	Não	Não se aplica

COMPOSTOS	DS TCRPI	DS HSRPI	LOGP	HD	HA	LV	OB VEBER	OB EGAN	PAINIS	CLUSTER	ENDÓGENO HUMANOS	NOMES DE ENDÓGENOS
<b>ZINC33930908</b>	-9.466	-5.341	-2.17	4.0	5.0	0.0	Bom	Bom	Aceito	24	Não	Não se aplica
ZINC88466005	-8.968	-6.859	-2.15	5.0	7.0	0.0	Bom	Bom	Aceito	24	Não	Não se aplica
ZINC04095527	-10.23	-8.613	-4.17	5.0	9.0	0.0	Bom	Bom	Aceito	25	Não	Não se aplica
ZINC06753359	-9.227	-8.238	-4.17	5.0	9.0	0.0	Bom	Bom	Aceito	25	Não	Não se aplica
ZINC06753370	-9.061	-7.527	-4.17	5.0	9.0	0.0	Bom	Bom	Aceito	25	Não	Não se aplica
ZINC13507434	-9.501	-8.561	-4.17	5.0	9.0	0.0	Bom	Bom	Aceito	25	Não	Não se aplica
<b>ZINC43763931</b>	-11.849	-8.571	-4.17	5.0	9.0	0.0	Bom	Bom	Aceito	25	Não	Não se aplica
ZINC88466498	-9.046	-7.985	-4.17	5.0	9.0	0.0	Bom	Bom	Aceito	25	Não	Não se aplica
ZINC09970474	-9.404	-7.094	-3.5	6.0	9.0	1.0	Bom	Bom	Aceito	26	Não	Não se aplica
<b>ZINC13373392</b>	-10.947	-8.936	-1.61	4.0	9.0	0.0	Bom	Bom	Aceito	26	Não	Não se aplica
ZINC22062810	-9.015	-8.312	-3.34	5.0	10.0	0.0	Bom	Bom	Aceito	26	Não	Não se aplica
ZINC22063317	-9.434	-8.22	-3.01	4.0	10.0	0.0	Bom	Bom	Aceito	26	Não	Não se aplica
ZINC27643761	-9.212	-7.426	-1.98	5.0	7.0	0.0	Bom	Bom	Aceito	26	Não	Não se aplica
ZINC27643761	-8.965	-8.608	-1.98	5.0	7.0	0.0	Bom	Bom	Aceito	26	Não	Não se aplica
ZINC19300946	-11.601	-8.795	-1.08	4.0	7.0	0.0	Bom	Bom	Aceito	não alocado	Não	Não se aplica
ZINC00895179	-9.057	-8.009	-1.8	4.0	7.0	0.0	Bom	Bom	Aceito	1	Sim	isocitrate
ZINC01656422	-8.973	-8.576	-2.56	5.0	8.0	0.0	Bom	Bom	Aceito	1	Sim	acid garcinia
ZINC03581021	-10.96	-7.496	-2.51	6.0	8.0	1.0	Bom	Bom	Aceito	1	Sim	d-glucarate
ZINC12405009	-10.72	-7.505	-2.51	6.0	8.0	1.0	Bom	Bom	Aceito	1	Sim	d-glucarate
ZINC02530801	-8.921	-8.89	-3.69	6.0	7.0	1.0	Bom	Bom	Aceito	8	Sim	tyr-thr
ZINC04556592	-9.993	-7.814	0.14	5.0	8.0	0.0	Bom	Bom	Aceito	17	Sim	indoxylglucuronide
ZINC14811803	-10.727	-8.636	0.58	6.0	13.0	3.0	Bom	Bom	Aceito	17	Sim	6-malonylgenistin
ZINC76935860	-9.104	-8.084	-0.81	5.0	9.0	0.0	Bom	Bom	Aceito	17	Sim	paracetamol glucuronide
ZINC35456138	-11.41	-8.723	-0.18	5.0	8.0	0.0	Bom	Bom	Aceito	18	Sim	não informado
ZINC35644664	-9.12	-8.525	1.99	4.0	8.0	0.0	Bom	Bom	Aceito	18	Sim	ibuprofen glucuronide
ZINC03869847	-11.659	-8.455	-2.78	6.0	12.0	2.0	Bom	Bom	Aceito	20	Sim	dgmp
ZINC03869847	-10.703	-8.465	-2.78	6.0	12.0	2.0	Bom	Bom	Aceito	20	Sim	dgmp

COMPOSTOS	DS TCRPI	DS HSRPI	LOGP	HD	HA	LV	OB VEBER	OB EGAN	PAINS	CLUSTER	ENDÓGENO HUMANOS	NOMES DE ENDÓGENOS
ZINC31259283	-9.723	-8.946	-3.59	6.0	10.0	1.0	Bom	Bom	Aceito	21	Sim	air
ZINC03870253	-9.747	-8.847	-2.58	4.0	10.0	0.0	Bom	Bom	Aceito	22	Sim	dtmp
ZINC08585028	-11.429	-8.625	-4.8	5.0	14.0	1.0	Bom	Bom	Aceito	22	Sim	UDP
ZINC13523519	-10.268	-4.69	-2.58	4.0	10.0	0.0	Bom	Bom	Aceito	22	Sim	dtmp
ZINC14618594	-9.838	-5.199	-2.57	5.0	10.0	0.0	Bom	Bom	Aceito	22	Sim	igual oritidine
ZINC14618595	-9.124	-8.577	-2.57	5.0	10.0	0.0	Bom	Bom	Aceito	22	Sim	igual oritidine
ZINC14618596	-9.422	-8.19	-2.57	5.0	10.0	0.0	Bom	Bom	Aceito	22	Sim	orotidine
ZINC04536469	-9.242	-8.656	-3.61	6.0	11.0	2.0	Bom	Bom	Aceito	23	Sim	cmp
ZINC04536485	-9.25	-8.614	-3.61	6.0	11.0	2.0	Bom	Bom	Aceito	23	Sim	2cmp
ZINC00895055	-9.741	-6.888	-3.24	5.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	d-galactose cerelose alose galactose manose, vadex
ZINC00895350	-10.256	-6.799	-3.24	5.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	alose
ZINC00967344	-10.311	-6.31	-2.65	4.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	methyl mannoside
ZINC02042990	-9.074	-6.387	-2.65	4.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	methyl mannoside
ZINC03830874	-10.227	-6.609	-2.29	5.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	fru
ZINC03830875	-9.375	-7.688	-2.29	5.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	fru
ZINC03860714	-9.369	-6.482	-2.32	4.0	5.0	0.0	Bom	Bom	Aceito	24	Sim	ribose
ZINC04097152	-8.922	-7.145	-3.24	5.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	Cerelos D-galactose allose galactose manose vadex
ZINC05226983	-10.037	-6.239	-2.65	4.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	metil manoside

COMPOSTOS	DS TCRPI	DS HSRPI	LOGP	HD	HA	LV	OB VEBER	OB EGAN	PAINS	CLUSTER	ENDÓGENO HUMANOS	NOMES DE ENDÓGENOS
ZINC08579428	-9.734	-5.939	-2.65	4.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	metil manoside
ZINC13527852	-9.541	-7.253	-2.65	4.0	6.0	0.0	Bom	Bom	Aceito	24	Sim	metil manoside
ZINC44963311	-9.059	-7.687	-1.44	4.0	7.0	0.0	Bom	Bom	Aceito	24	Sim	ethylglucuronide
ZINC44964245	-9.067	-7.537	1.07	4.0	8.0	0.0	Bom	Bom	Aceito	24	Sim	ypag
ZINC02047359	-11.314	-6.714	-4.46	7.0	10.0	1.0	Bom	Bom	Aceito	26	Sim	L-Guluronic Acid 6-Phosphate
ZINC03869945	-8.94	-7.542	-6.22	7.0	9.0	1.0	Bom	Bom	Aceito	26	Sim	glucosamine-1p
ZINC03869945	-9.616	-6.775	-6.22	7.0	9.0	1.0	Bom	Bom	Aceito	26	Sim	glucosamine-1p
ZINC04095567	-9.481	-8.502	-4.28	6.0	9.0	1.0	Bom	Bom	Aceito	26	Sim	galactose 1-phosphate dpm mannose-1-p
ZINC04228241	-9.953	-8.665	-3.97	5.0	8.0	0.0	Bom	Bom	Aceito	26	Sim	ribose-5p
ZINC04228296	-10.248	-8.241	-2.64	4.0	7.0	0.0	Bom	Bom	Aceito	26	Sim	deoxy-ribose-5p
ZINC35924159	-9.08	-7.279	-6.22	7.0	9.0	1.0	Bom	Bom	Aceito	26	Sim	glucosamine-1p
ZINC35924159	-9.809	-8.761	-6.22	7.0	9.0	1.0	Bom	Bom	Aceito	26	Sim	glucosamine-1p



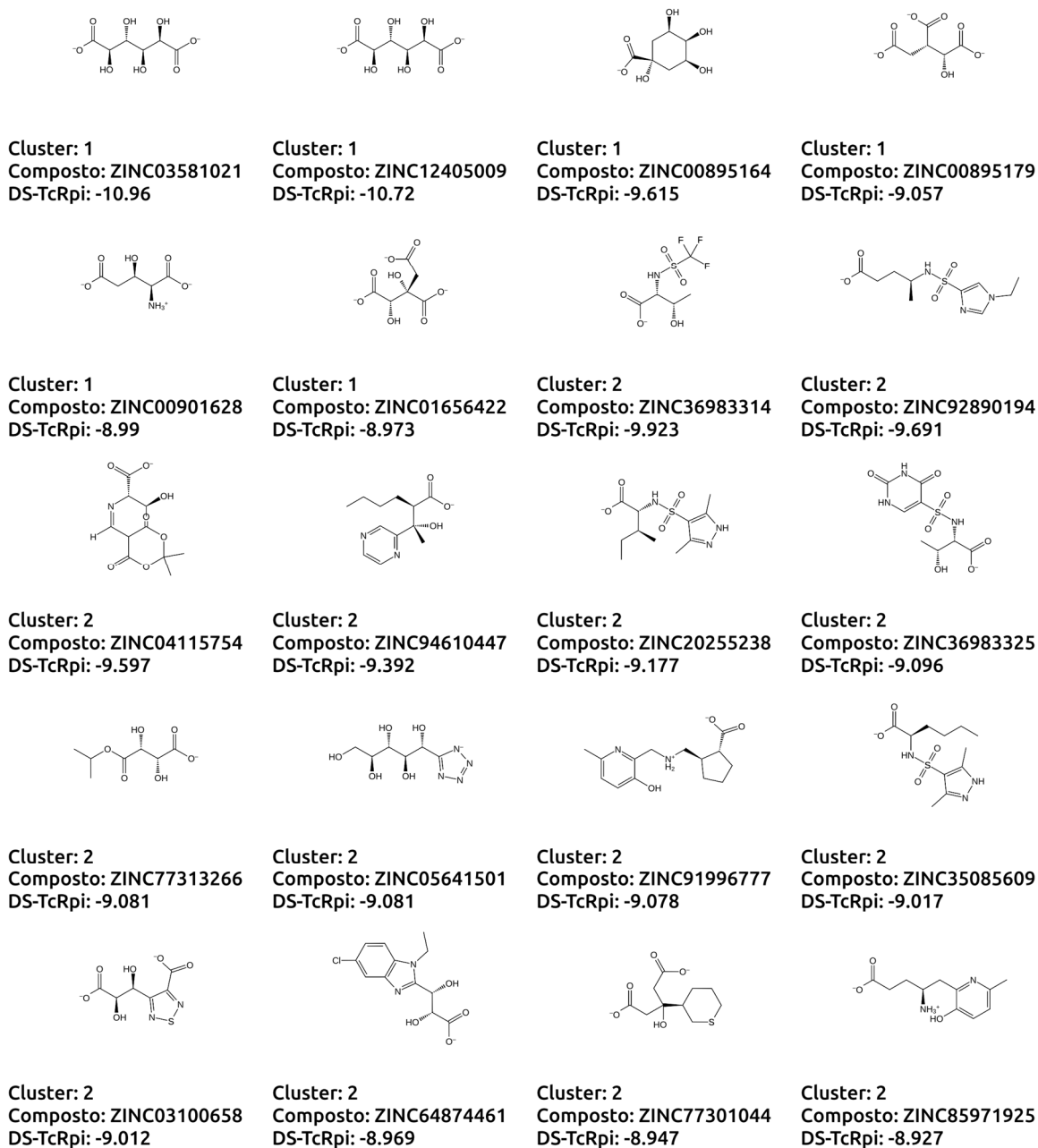
**Tabela 7.2:** Compostos com propriedades farmacológicas conhecidas, ou com mesmo arcabouço molecular de tais compostos (somente quiralidade distinta) entre o conjunto final após filtro ADME-Tox e PAINS (exceto metabólitos humanos). Fonte: PubChem e Zinc

Composto	Compostos de mesmo arcabouço molecular com propriedades farmacológicas descritas	Atividades Farmacológicas	Mecanismo de ação	OBS
ZINC05085274	ZINC4217381 (Pyrazofurin)	Diversas: Antibiótica Antiviral Antineoplásica Antiparasitário	É um análogo de nucleosídeos* descrito por potencialmente inibir a orotidine 5'-monophosphate (OMP) decarboxylase, impedindo síntese de novo de uridina	-
ZINC00895164	ZINC100009542 (Kinate) CID 6508 (Quinic acid)	Diversas	É um produto natural já cristalografado com 3-dehydroquinatase dehydratase de <i>Salmonella Enterica</i> e <i>Corynebacterium Glutamicum</i>	-
ZINC11592857	ZINC3876071 (Dexamethasone fosfato) CID 940 (dexamethasone phosphate)	Diversos	Pertence a classe de corticosteroides que afetam o metabolismo de carboidratos, inibem a secreção do hormônio adrenocorticotropico e possuem atividade antiinflamatória pronunciada.	-

<b>Composto</b>	<b>Compostos de mesmo arcabouço molecular com propriedades farmacológicas descritas</b>	<b>Atividades Farmacológicas</b>	<b>Mecanismo de ação</b>	<b>OBS</b>
ZINC13549482	CID 9548643 (Glucocheirolin)	-	-	Composto de origem vegetal que pode ser encontrado em alimentos
ZINC26977303	CID 11972479 (Torachryson 8-O-Glucoside)	Atividade antimicobacteriana	-	-
ZINC36521966	ZINC3812887 CID 104762 (Mizoribine)	Diversos: Antibiótica Antiinflamatória Antineoplásica Antiparasitário Agente imunossupressivo	Inibição seletiva da selectiveda inosine monophosphate synthetase e guanosine monophosphate synthetase	Aprovado pelo FDA
ZINC93120380	CID 29986850 (cis-resveratrol 3-O-glucuronide)	-	-	Encontrado em alimentos
ZINC01610270	CID 165151 (ddTMP)	Demonstrou atividade inibitória, em bioensaio, levando a acúmulo de UDP-N-acetylglucosamine no lúmen das membranas do complexo de Golgi	Ligação ao transportador UDP-N-acetylglucosamine (UDP-GlcNAc)	-
ZINC04535606	ZINC4096919 (Cytidine-3'-Monophosphate)	-	Inibição em bioensaio da RNAase A e inibição da ribonuclease A bovina	-

Composto	Compostos de mesmo arcaibouço molecular com propriedades farmacológicas descritas	Atividades Farmacológicas	Mecanismo de ação	OBS
ZINC33930908	CID 444914 (4-deoxy-alpha-d-glucosyl)	-	É um produto natural já cristalografado com enzimas como chondroitinase B de <i>Flavobacterium heparinum</i>	-
ZINC00002053	CID 439503 (Salicin)	Atividade anti-inflamatória. Por ter estrutura química semelhante a da aspirina, possui ação parecida em humanos	Inibição da cyclooxygenase (prostaglandin-endoperoxide synthases)	-
ZINC13761105 CID 5283388 (Anandamide 0-phosphate)	-	-	Afinidade de ligação <i>in vitro</i> contra o receptor de Cannabinoide 1 de rato	-

## Apêndice B - ESTRUTURAS QUÍMICAS DOS COMPOSTOS CITADOS NA TESE



**Figura 7.1** Estruturas químicas dos 211 compostos finais de TcRpiB. Estão indicados o nome de cada composto, seus respectivos grupos (clusters) e valores de *docking score* (DS) referente a TcRpi.

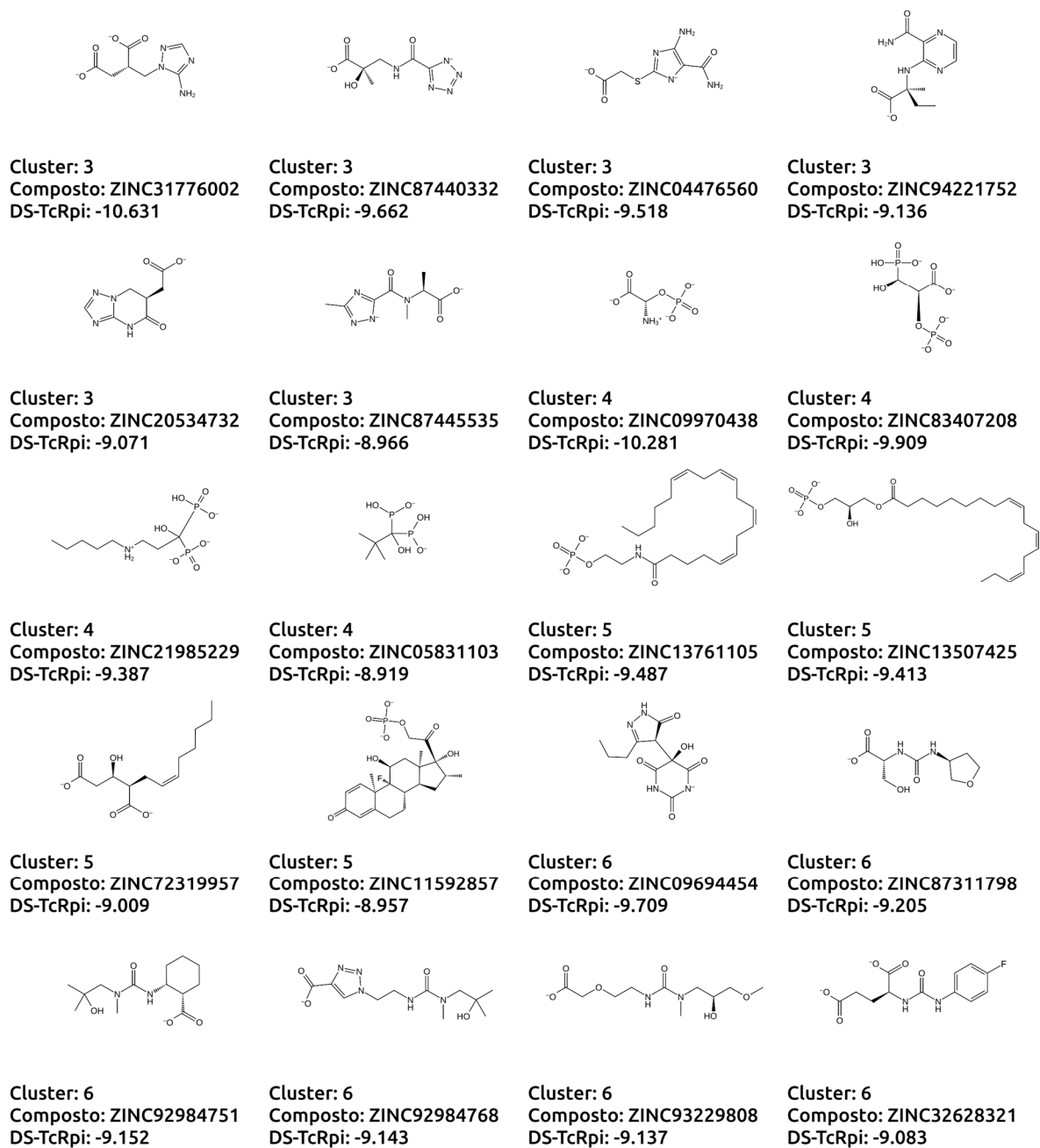
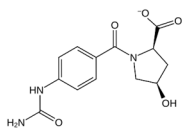
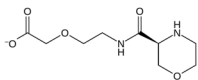


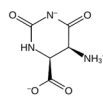
Figura 7.1. **Continuação**



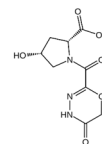
Cluster: 6  
Composto: ZINC37449495  
DS-TcRpi: -9.01



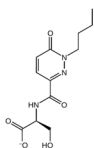
Cluster: 6  
Composto: ZINC86433775  
DS-TcRpi: -8.981



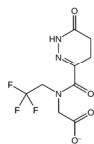
Cluster: 6  
Composto: ZINC20160154  
DS-TcRpi: -8.931



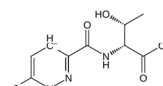
Cluster: 7  
Composto: ZINC36975961  
DS-TcRpi: -12.735



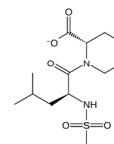
Cluster: 7  
Composto: ZINC37451246  
DS-TcRpi: -10.031



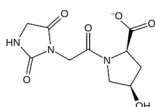
Cluster: 7  
Composto: ZINC92435320  
DS-TcRpi: -9.957



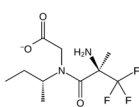
Cluster: 7  
Composto: ZINC37431651  
DS-TcRpi: -9.95



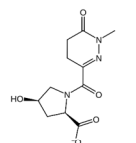
Cluster: 7  
Composto: ZINC37437566  
DS-TcRpi: -9.568



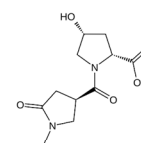
Cluster: 7  
Composto: ZINC37449060  
DS-TcRpi: -9.432



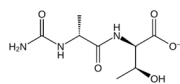
Cluster: 7  
Composto: ZINC94236295  
DS-TcRpi: -9.404



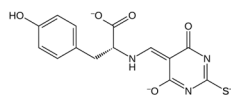
Cluster: 7  
Composto: ZINC37449519  
DS-TcRpi: -9.11



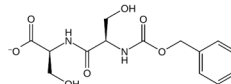
Cluster: 7  
Composto: ZINC37449313  
DS-TcRpi: -9.086



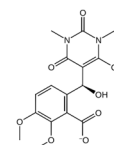
Cluster: 7  
Composto: ZINC36971700  
DS-TcRpi: -9.072



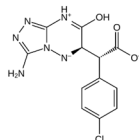
Cluster: 8  
Composto: ZINC20572839  
DS-TcRpi: -9.666



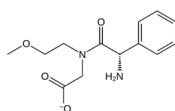
Cluster: 8  
Composto: ZINC64219433  
DS-TcRpi: -9.535



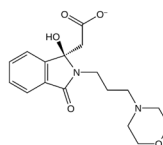
Cluster: 8  
Composto: ZINC75278199  
DS-TcRpi: -9.487



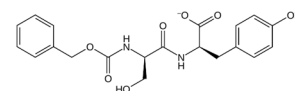
Cluster: 8  
Composto: ZINC19799225  
DS-TcRpi: -9.394



Cluster: 8  
Composto: ZINC92493836  
DS-TcRpi: -9.343

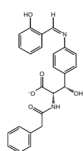


Cluster: 8  
Composto: ZINC23187566  
DS-TcRpi: -9.262

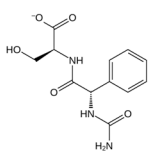


Cluster: 8  
Composto: ZINC01576100  
DS-TcRpi: -9.206

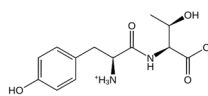
Figura 7.1. Continuação



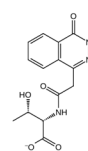
Cluster: 8  
Composto: ZINC16478605  
DS-TcRpi: -9.075



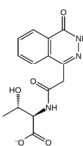
Cluster: 8  
Composto: ZINC37451146  
DS-TcRpi: -9.024



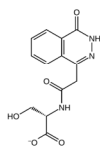
Cluster: 8  
Composto: ZINC02530801  
DS-TcRpi: -8.921



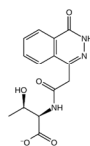
Cluster: 9  
Composto: ZINC13552852  
DS-TcRpi: -10.906



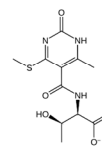
Cluster: 9  
Composto: ZINC13654785  
DS-TcRpi: -10.652



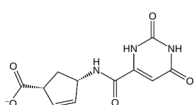
Cluster: 9  
Composto: ZINC09577104  
DS-TcRpi: -10.521



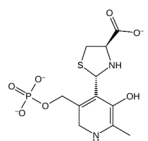
Cluster: 9  
Composto: ZINC09575936  
DS-TcRpi: -9.47



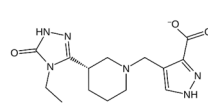
Cluster: 10  
Composto: ZINC36972227  
DS-TcRpi: -10.182



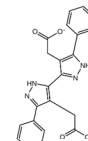
Cluster: 10  
Composto: ZINC92177362  
DS-TcRpi: -9.87



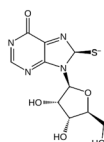
Cluster: 10  
Composto: ZINC13213453  
DS-TcRpi: -9.79



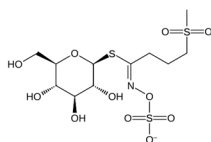
Cluster: 10  
Composto: ZINC67695654  
DS-TcRpi: -9.534



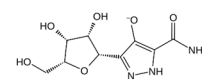
Cluster: 10  
Composto: ZINC05479266  
DS-TcRpi: -9.506



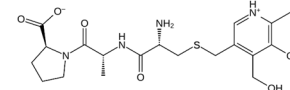
Cluster: 10  
Composto: ZINC38549580  
DS-TcRpi: -9.414



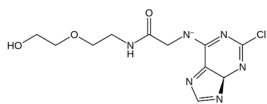
Cluster: 10  
Composto: ZINC13549482  
DS-TcRpi: -9.385



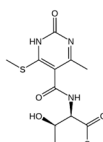
Cluster: 10  
Composto: ZINC05085274  
DS-TcRpi: -9.328



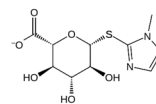
Cluster: 10  
Composto: ZINC20611195  
DS-TcRpi: -9.292



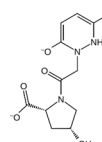
Cluster: 10  
Composto: ZINC39275476  
DS-TcRpi: -9.246



Cluster: 10  
Composto: ZINC36972227  
DS-TcRpi: -9.153



Cluster: 10  
Composto: ZINC44963783  
DS-TcRpi: -9.15



Cluster: 10  
Composto: ZINC37448985  
DS-TcRpi: -9.117

Figura 7.1. Continuação

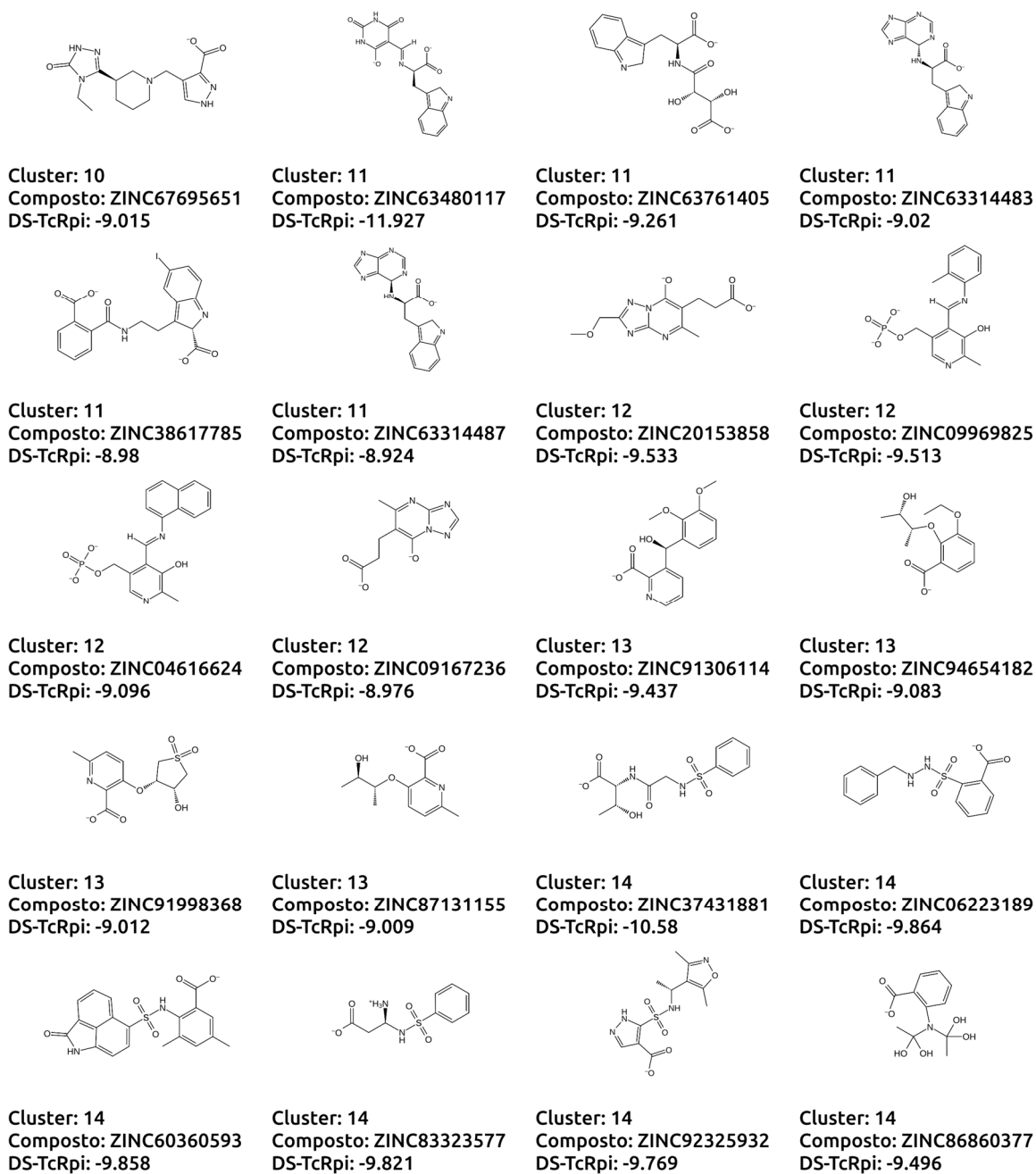
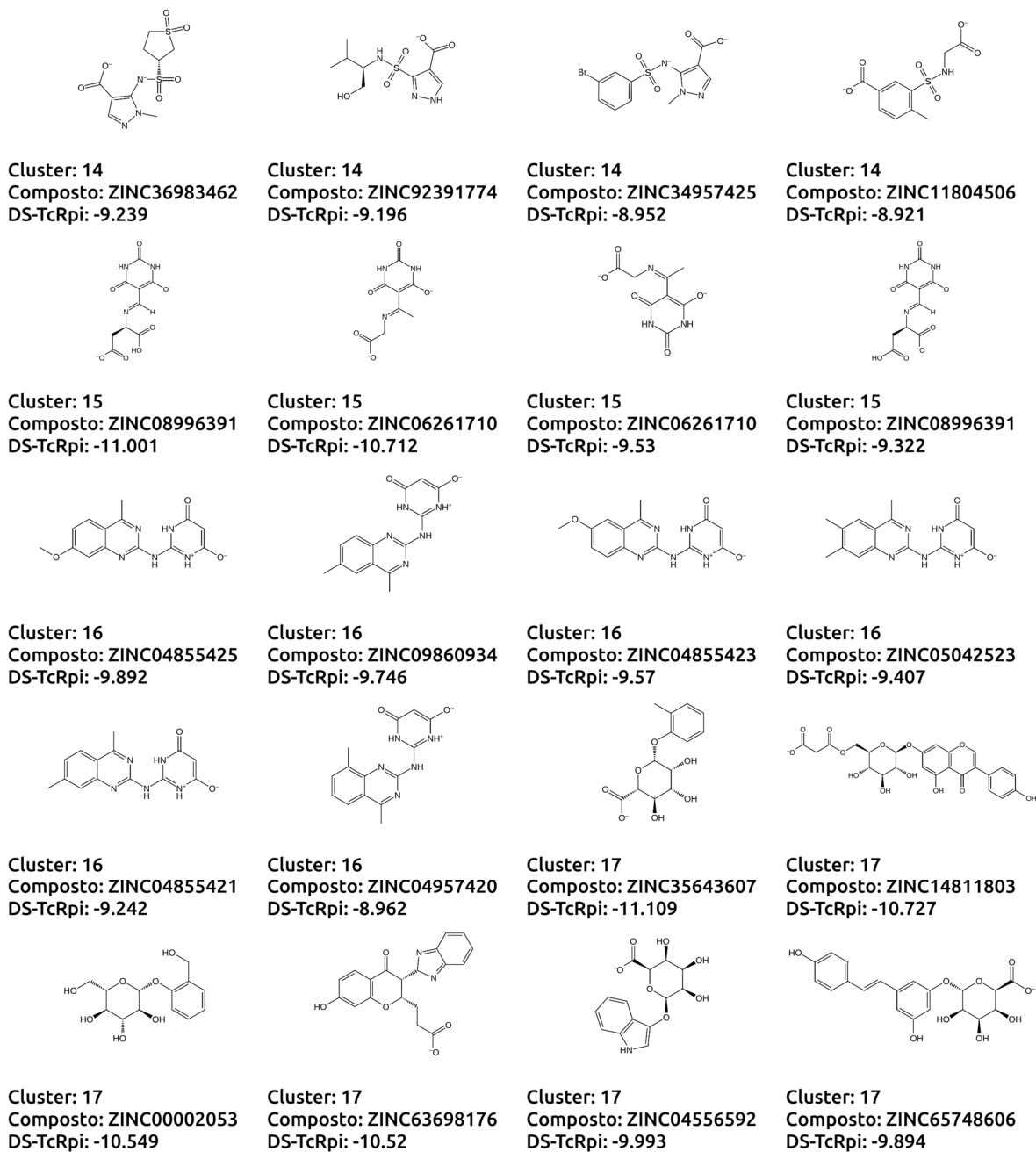


Figura 7.1. Continuação





**Figura 7.1.** Continuação

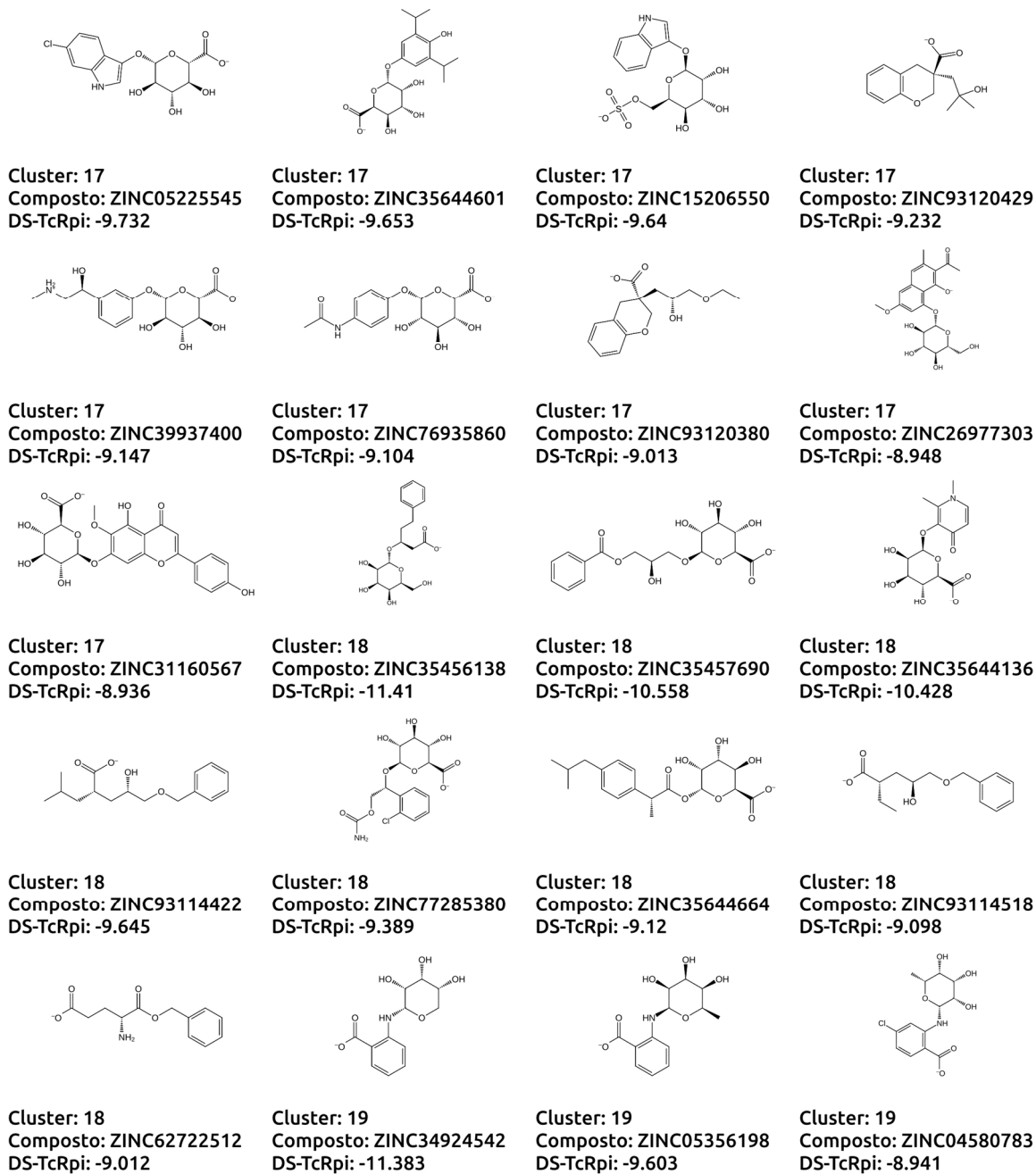


Figura 7.1. Continuação

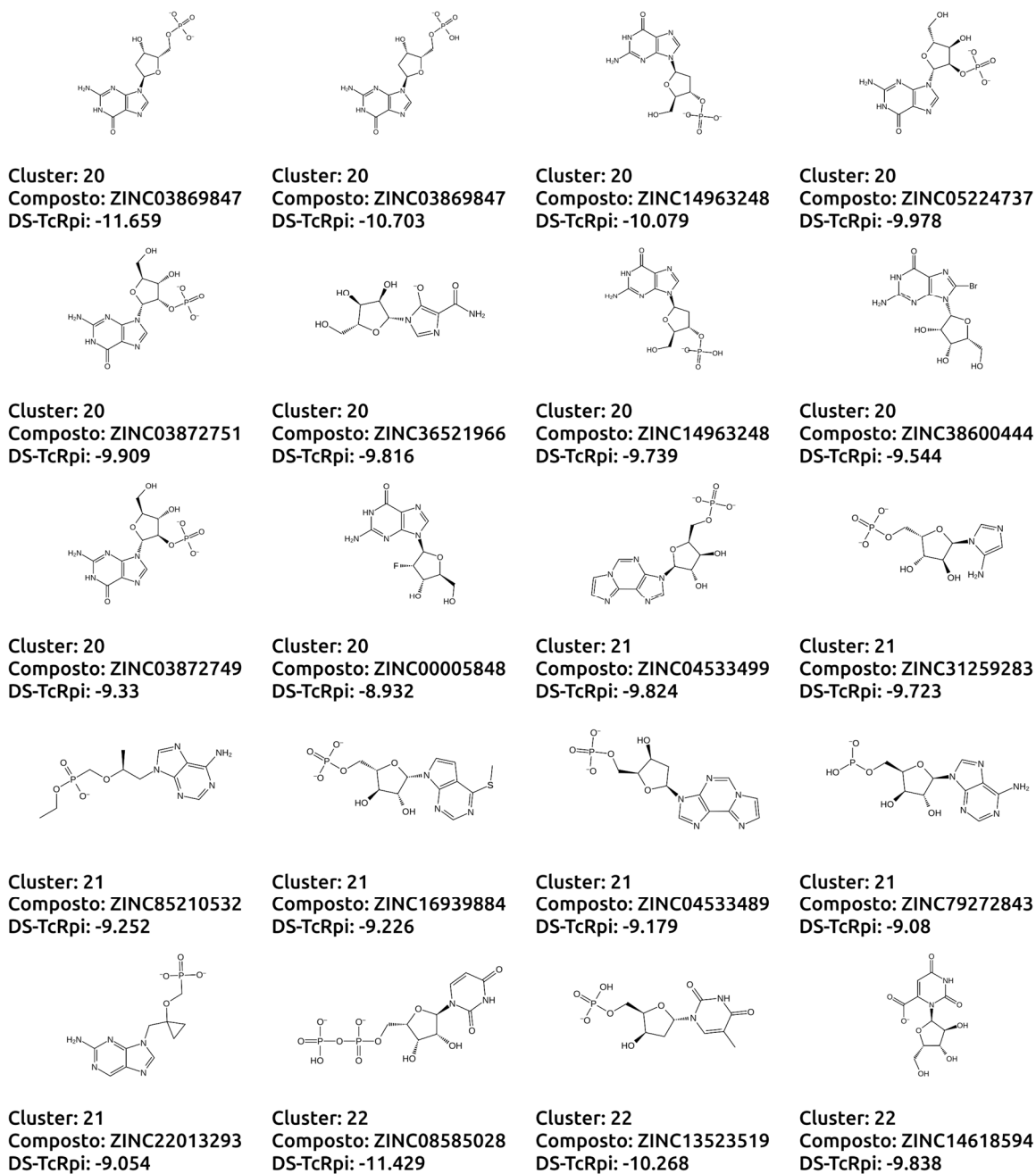


Figura 7.1. Continuação

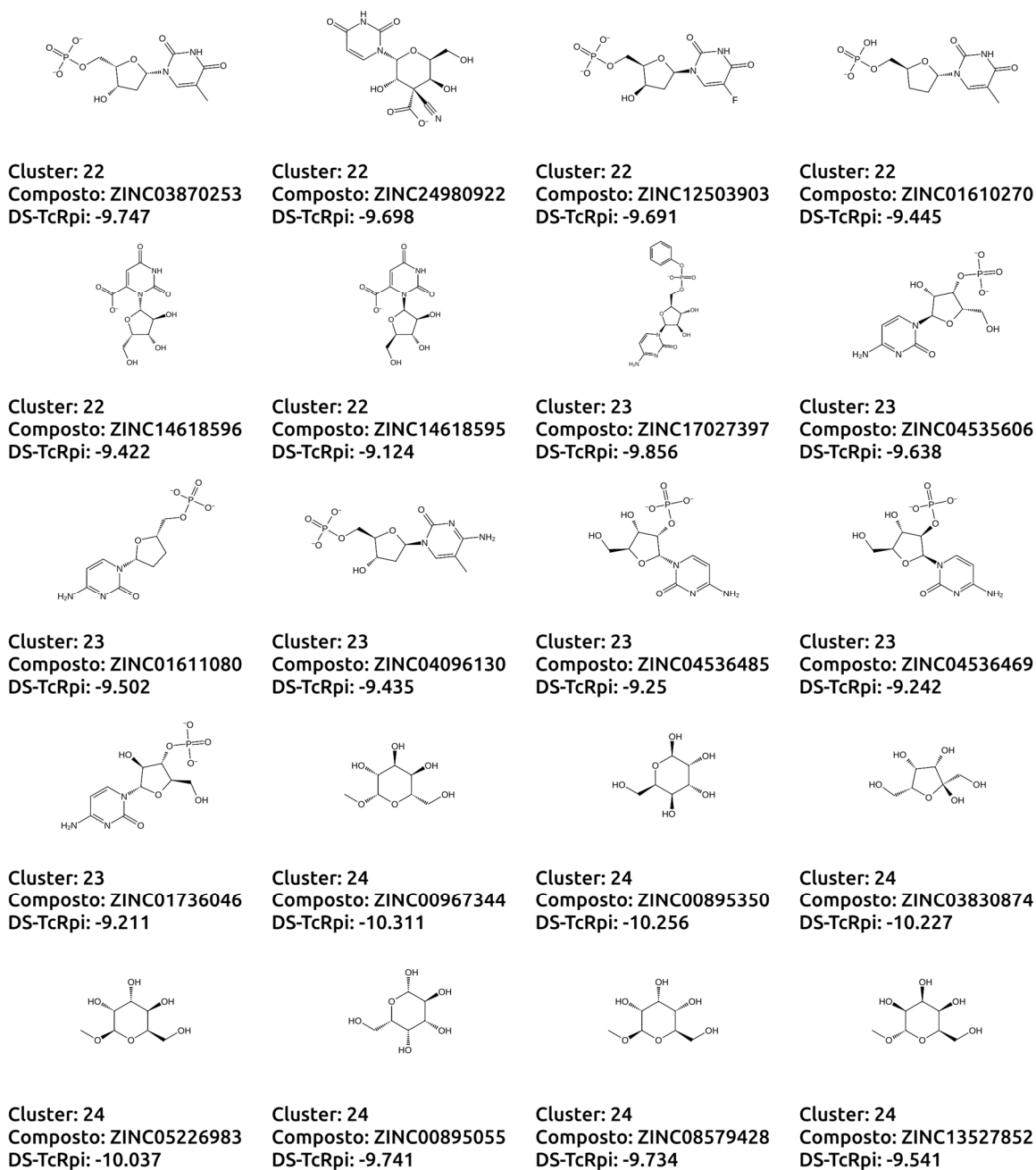
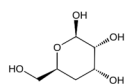
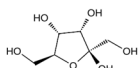


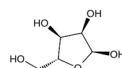
Figura 7.1. Continuação



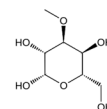
Cluster: 24  
Composto: ZINC33930908  
DS-TcRpi: -9.466



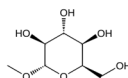
Cluster: 24  
Composto: ZINC03830875  
DS-TcRpi: -9.375



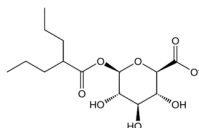
Cluster: 24  
Composto: ZINC03860714  
DS-TcRpi: -9.369



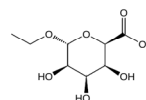
Cluster: 24  
Composto: ZINC04557020  
DS-TcRpi: -9.36



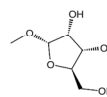
Cluster: 24  
Composto: ZINC02042990  
DS-TcRpi: -9.074



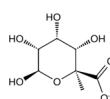
Cluster: 24  
Composto: ZINC44964245  
DS-TcRpi: -9.067



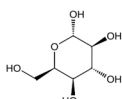
Cluster: 24  
Composto: ZINC44963311  
DS-TcRpi: -9.059



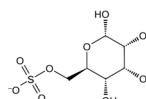
Cluster: 24  
Composto: ZINC22063668  
DS-TcRpi: -9.013



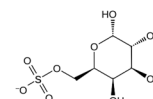
Cluster: 24  
Composto: ZINC88466005  
DS-TcRpi: -8.968



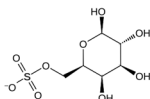
Cluster: 24  
Composto: ZINC04097152  
DS-TcRpi: -8.922



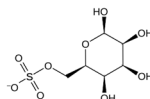
Cluster: 25  
Composto: ZINC43763931  
DS-TcRpi: -11.849



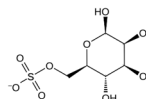
Cluster: 25  
Composto: ZINC04095527  
DS-TcRpi: -10.23



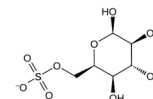
Cluster: 25  
Composto: ZINC13507434  
DS-TcRpi: -9.501



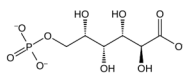
Cluster: 25  
Composto: ZINC06753359  
DS-TcRpi: -9.227



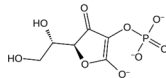
Cluster: 25  
Composto: ZINC06753370  
DS-TcRpi: -9.061



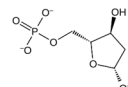
Cluster: 25  
Composto: ZINC88466498  
DS-TcRpi: -9.046



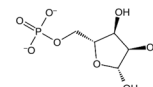
Cluster: 26  
Composto: ZINC02047359  
DS-TcRpi: -11.314



Cluster: 26  
Composto: ZINC13373392  
DS-TcRpi: -10.947

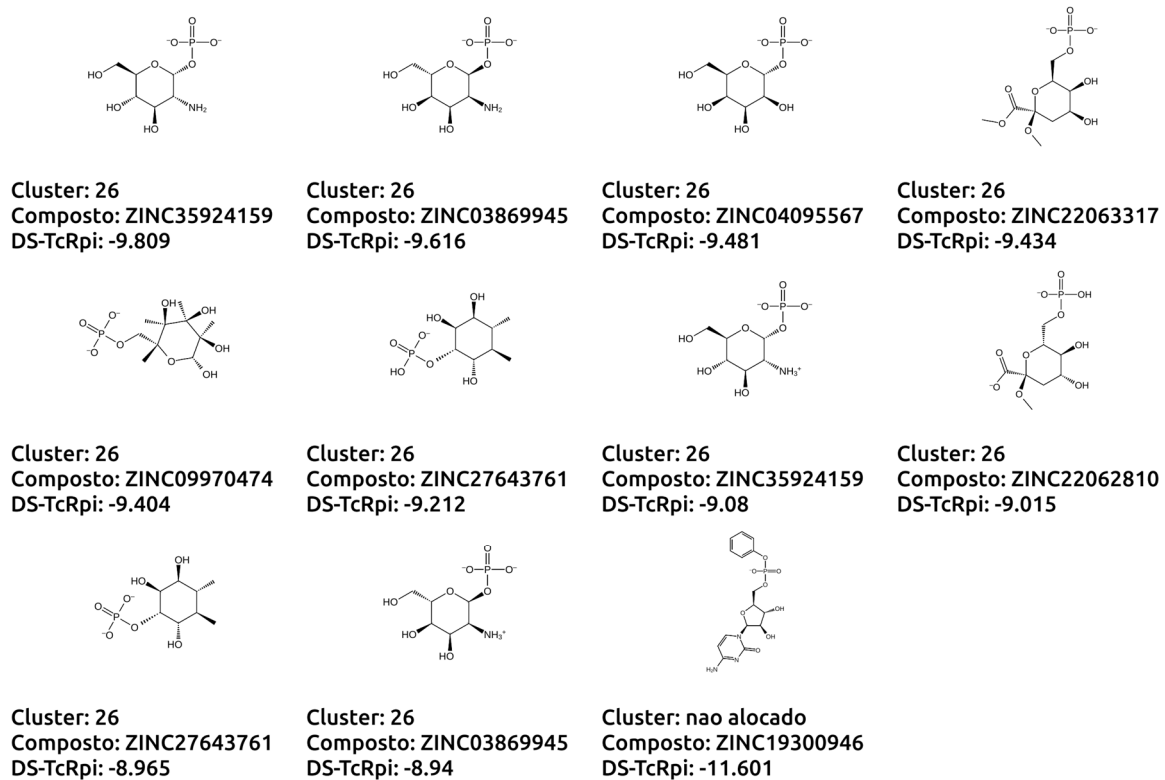


Cluster: 26  
Composto: ZINC04228296  
DS-TcRpi: -10.248

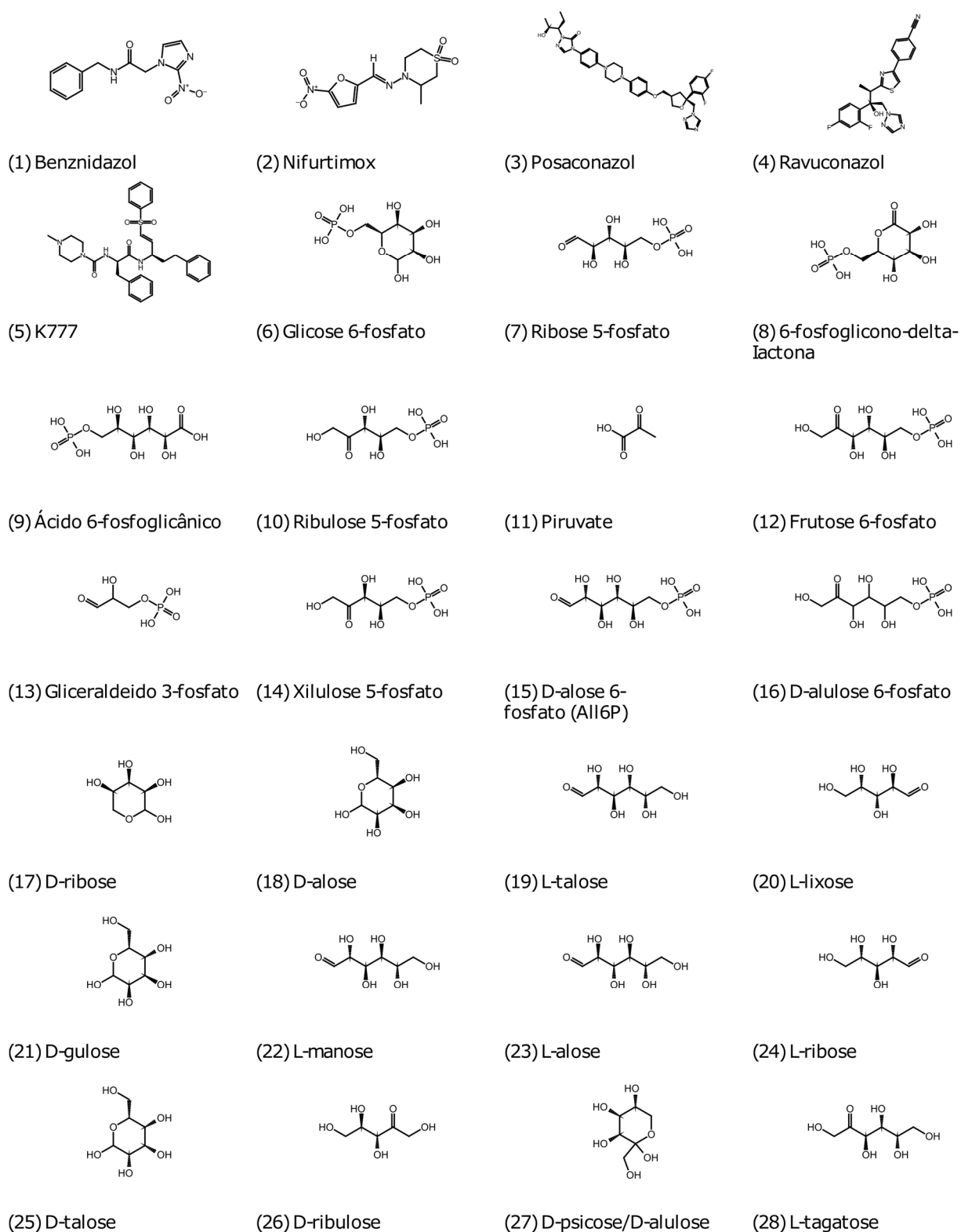


Cluster: 26  
Composto: ZINC04228241  
DS-TcRpi: -9.953

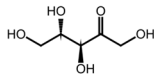
Figura 7.1. Continuação



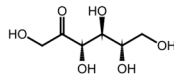
**Figura 7.1.** Continuação



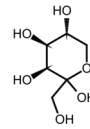
**Figura 7.2** Estruturas químicas de alguns dos compostos químicos citados na tese. Não foram inclusos compostos já representados na figura 7.1. A cada composto foi atribuído um número equivalente a ordem de citação na tese.



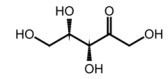
(29) L-ribulose



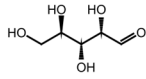
(30) L-psicose



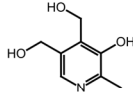
(31) D-tagatose



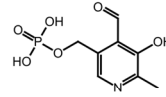
(32) L-xilulose



(33) D-arabinose



(34) Piridoxina



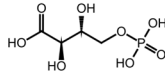
(35) Piridoxal fosfato



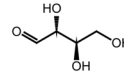
(36) Iodoacetamida



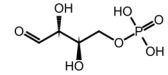
(37) Ácido iodoacético



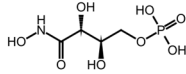
(38) Ácido 4-fosfo-D-eritrônico (4PEA)



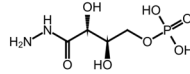
(39) D-eritrose



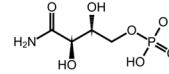
(40) D-eritrose 4-fosfato



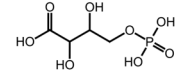
(41) Ácido 4-fosfo-D-eritronohidroxâmico (4PEH)



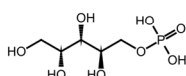
(42) 4-fosfo-D-eritronohidrazida (4PEHz)



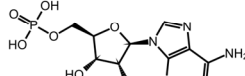
(43) 4-fosfo-D-eritronamida (4PEAm)



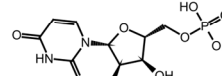
(44) 4-fosfonometil-D-eritronato (4PMEA)



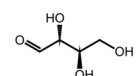
(45) Arabinose 5-fosfato



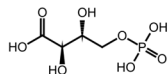
(46) AMP



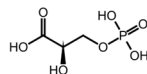
(47) UMP



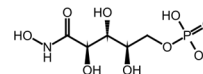
(48) D-eritrose



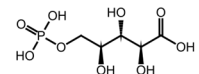
(49) Ácido D-4-eritrônico



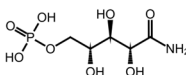
(50) Ácido D-3-fosfoglicérico



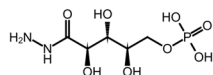
(51) Ácido 5-fosfo-D-ribonohidroxâmico (5PRH)



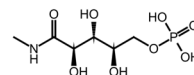
(52) 5-fosfo-D-ribonato (5PRA)



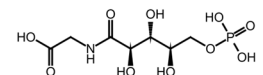
(53) 5-fosfo-D-ribonamida (5PRAm)



(54) N-(5-fosfo-D-ribonil)-hidrazina (5PRHz)



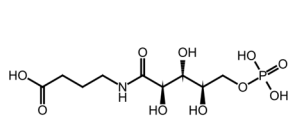
(55) N-(5-fosfo-D-ribonil)-metilamina (5PRMA)



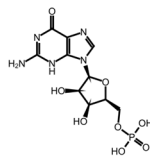
(56) N-(5-fosfo-D-ribonil)-glicina (5PRGly)

**Figura 7.2** Continuação

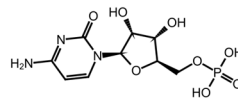




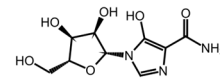
(57) N-(5-fosfo-D-ribonil)-c-aminobutanoato (5PRGABA)



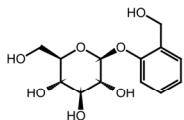
(58) GMP



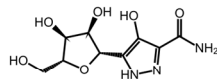
(59) CMP



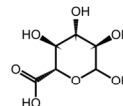
(60) Mizoribine



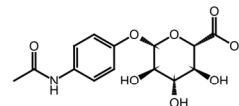
(61) Salicina



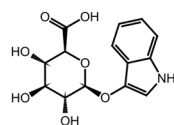
(62) Pirazofurin



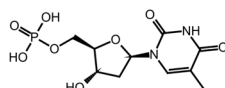
(63) Ácido glucarônico (Geral)



(64) Paracetamol



(65) Indoxil glucoronídeo



(66) TMP



(67) Sulfonamidas (Geral)

**Figura 7.2** Continuação

## Apêndice C - PRODUÇÃO ACADÊMICA NO PERÍODO REFERENTE AO DOUTORADO

### ***Publicação em periódicos***

1. de V. C. Sinatti, Vanessa, Ribeiro Baptista, Luiz Phillippe, Alves-Ferreira, Marcelo, Dardenne, Laurent, Hermínio Martins da Silva, João, Guimarães, Ana Carolina In silico identification of inhibitors of ribose 5-phosphate isomerase from *Trypanosoma cruzi* using ligand and structure based approaches. Journal of Molecular Graphics and Modelling, 77 (2017) 168–180.

### ***Participação em projetos***

1. Nome do projeto de mestrado: Identificação *in silico* de possíveis inibidores da fumarato hidratase de *Leishmania major*  
Aluna: Aline Beatriz Rodrigues (BCS)  
Orientador: Ana Carolina Ramos Guimarães
2. Nome do projeto de doutorado: Identificação e avaliação computacional de possíveis inibidores alostéricos e competitivos de PEPCK-M humana: uma alternativa para a terapia contra o carcinoma de pulmão  
Aluno: Luiz Phillippe Ribeiro Baptista (BCS)  
Orientador: Ana Carolina Ramos Guimarães
3. Nome do projeto de doutorado: Aspártico peptidases de *Trypanosoma cruzi*: atração molecular com Inibidores de peptidase do HIV, análise de expressão gênica e geração de proteína heteróloga.  
Aluna: Keyla Cristiny da Silva Gonçalves (BCM)  
Orientador: Claudia Masini d'Avila-Levy e Rubem Menna-Barreto

### ***Congressos e Simpósios***

1. X-meeting/BSB 2013  
Nome do trabalho apresentado: In silico and in vitro characterization of the enzymes gluconate kinase and uracil phosphoribosyltransferase from *Trypanosoma cruzi*.

- Cidade do Evento: Recife, PE  
Ano do evento: 2013  
Meio de divulgação: Apresentação poster
2. Ciclo Carlos Chagas de Palestras  
Nome do trabalho apresentado: Identificação e caracterização das enzimas uracil fosforribosiltransferase e gliconato cinase de *Trypanosoma cruzi*.  
Cidade do Evento: Rio de Janeiro, RJ  
Ano do evento: 2014  
Meio de divulgação: Apresentação oral
3. ISCB-LA/X-MEETING/BSB/SoiBio  
Nome do trabalho apresentado: Virtual screening aiming to the identification of inhibitors against ribose 5-phosphate isomerase from *Trypanosoma cruzi*.  
Cidade do Evento: Belo Horizonte, MG  
Ano do evento: 2014  
Meio de divulgação: Apresentação de poster
4. 7º Simpósio Brasileiro em Química Medicinal - BrazMedChem  
Nome do trabalho apresentado: Identification of cytotoxic substances to *Trypanosoma cruzi* in Chagas disease chemotherapy  
Cidade do Evento: Campos do Jordão, SP  
Ano do evento: 2014  
Meio de divulgação: Apresentação de poster
5. X-Meeting 2015 - 11th International Conference of the AB3C + Brazilian Symposium of Bioinformatics  
Nome do trabalho: In Silico identification of inhibitors against ribose 5-phosphate isomerase from *Trypanosoma cruzi*  
Cidade do Evento: São Paul, SP  
Ano do evento: 2015  
Meio de divulgação: Apresentação de poster
6. 8th Brazilian Symposium on Medicinal Chemistry- BrazMedChem

Cidade do evento: Búzios, RJ

Ano do evento: 2016

Nome do trabalho: In silico identification of potential inhibitors of Ribose 5-phosphate isomerase of *Trypanosoma cruzi*

Meio de divulgação: Apresentação de poster

### **Workshops e cursos**

1. 1º Workshop em Métodos Computacionais aplicados ao desenvolvimento de fármacos

Cidade do Evento: UFRJ- Faculdade Farmácia, Rio de Janeiro

Ano do evento: 2013

#### **Minicursos assistidos:**

Tutoriais de programas de ADMET. (Carga horária: 5h).

Tutoriais dos programas Pymol e AutoDock4. (Carga horária: 5h).

2. VII Escola De Modelagem Molecular Em Sistemas Biológico

Cidade do evento: LNCC - Petrópolis, RJ

Ano do evento: 2014

#### **Minicursos assistidos:**

Dinâmica Molecular Básica (Programa NAMD). (Carga horária:6).

Cálculo de Modos Normais em Proteínas. (Carga horária: 5h)

3. International Workshop on Target-based Drug Design

Local evento: Fiocruz- Rio de Janeiro, RJ

Ano do evento: 2015

4. Curso de Bioimageamento por Microscopia Automatizada

Local evento: Fiocruz- Rio de Janeiro, RJ

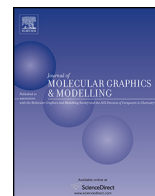
Ano do evento: 2015

5. Workshop Writing a Scientific Paper

Local evento: Fiocruz- Rio de Janeiro, RJ

Ano do evento: 2016

## **Apêndice D - PUBLICAÇÃO CIENTÍFICA RELACIONADA À TESE**



## Topical Perspectives

## *In silico* identification of inhibitors of ribose 5-phosphate isomerase from *Trypanosoma cruzi* using ligand and structure based approaches



Vanessa de V. C. Sinatti<sup>a,\*</sup>, Luiz Phillippe R. Baptista<sup>a</sup>, Marcelo Alves-Ferreira<sup>b,c</sup>, Laurent Dardenne<sup>d</sup>, João Hermínio Martins da Silva<sup>e</sup>, Ana Carolina Guimarães<sup>a</sup>

<sup>a</sup> Fiocruz, Instituto Oswaldo Cruz, Laboratório de Genômica Funcional e Bioinformática, Av. Brasil 4365, Manguinhos, 21040-900, Rio de Janeiro, RJ, Brazil

<sup>b</sup> Fiocruz, Laboratório de Modelagem de Sistemas Biológicos, Centro de Desenvolvimento Tecnológico em Saúde, Av. Brasil 4036, Manguinhos, 21040-361, Rio de Janeiro, Brazil

<sup>c</sup> Instituto Nacional de Ciência e Tecnologia em Inovação em Doenças de Populações Negligenciadas, INCT-IDPN, CNPq, Brazil

<sup>d</sup> Laboratório Nacional de Computação Científica, Grupo de Modelagem Molecular de Sistemas Biológicos, Av. Getúlio Vargas, 333, Quitandinha, 25651-075, Petrópolis, RJ, Brazil

<sup>e</sup> Fiocruz-Ceará, Computational Modeling Group, Avenida Santos Dumont, 5753, Papicu, 60175-047, Fortaleza, CE, Brazil

## ARTICLE INFO

## Article history:

Received 10 May 2017

Received in revised form 11 July 2017

Accepted 7 August 2017

Available online 12 August 2017

## Keywords:

Ribose 5-phosphate isomerase

*Trypanosoma cruzi*

Chagas disease

Virtual screening

Pharmacophore modelling

Molecular docking

Molecular dynamics

ADMET

## ABSTRACT

Chagas disease, caused by the protozoan *Trypanosoma cruzi*, affects approximately seven million people, mainly in Latin America, and causes about 7000 deaths annually. The available treatments are unsatisfactory and search for more effective drugs against this pathogen is critical. In this context, the ribose 5-phosphate isomerase (Rpi) enzyme is a potential drug target mainly due to its function in the pentose phosphate pathway and its essentiality (previously shown in other trypanosomatids). In this study, we propose novel potential inhibitors for the Rpi of *T. cruzi* (TcRpi) based on a computer-aided approach, including structure-based and ligand-based pharmacophore modeling. Along with a substructural and similarity search, the selected pharmacophore hypotheses were used to screen the purchasable subset of the ZINC Database, yielding 20,183 candidate compounds. These compounds were submitted to molecular docking studies in the TcRpi and Human Rpi (HsRpi) active sites in order to identify potential selective inhibitors for the *T. cruzi* enzyme. After the molecular docking and ADME-T (absorption, distribution, metabolism, excretion and toxicity)/PAINS (pan-assay interference compounds) screenings, 211 molecules were selected as potential TcRpi inhibitors. Out of these, three compounds – ZINC36975961, ZINC63480117, and ZINC43763931 – were submitted to molecular dynamics simulations and two of them – ZINC36975961 and ZINC43763931 – had good performance and made interactions with important active site residues over all the simulation time. These compounds could be considered potential TcRpi inhibitors candidates and also may be used as leads for developing new TcRpi inhibitors.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

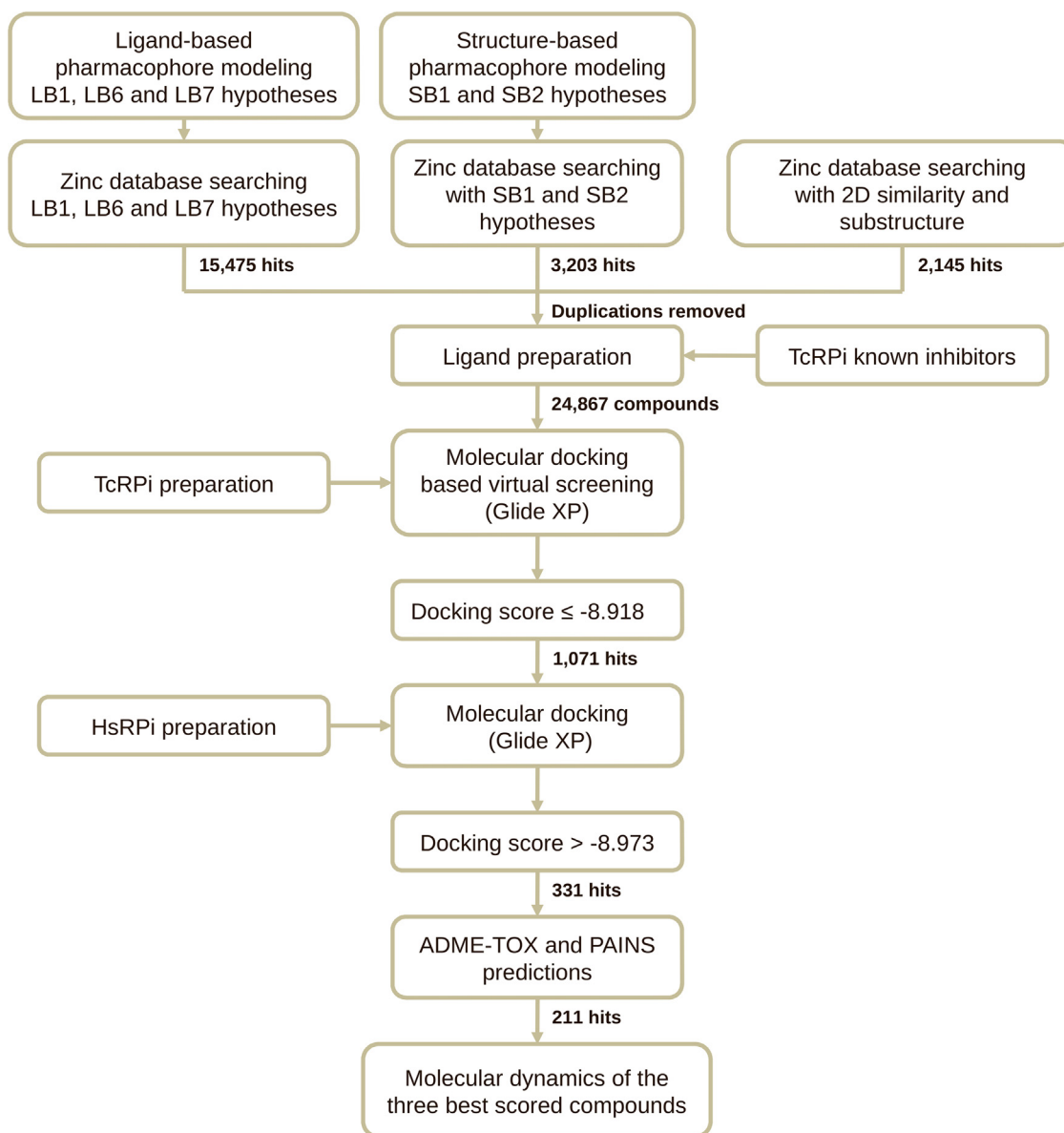
Chagas disease is a public health problem in Latin America and has also spread to non-endemic countries, as EUA, Canada, Australia, Japan and some countries in Europe, where the frequent underdiagnoses increase the risk of development of the symptomatic chronic form of the disease mainly cardiomyopathy and heart failure [1]. The latest estimates show that around seven million people in the world are infected by the kinetoplastid protozoan *Trypanosoma cruzi*, the etiological agent of Chagas disease [2]. Despite this, there are no vaccines for Chagas disease and the cur-

rent drugs available for the treatment are effective only in the acute phase, with limited efficacy in the chronic stage of the disease and many toxic side effects [3–5]. Therefore, the development of new drugs against the parasite is an urgent necessity. For this purpose, new therapeutic targets of critical importance in the metabolism of the parasite and the absence of structurally similar macromolecules in the host organism are essential points to be considered in the rational drug design against the Chagas disease.

Considering the rational drug design approach, analogous and specific enzymes of the parasite in relation to its host are interesting molecular targets. Analogous enzymes differ in their primary and tertiary structures but have the same enzymatic functions. The analogy happens due to different evolutionary origins of the genes encoding the same enzymatic activity by convergent evolution [6–8]. AnEnPi (<http://www.dbbm.fiocruz.br/AnEnPi/>) is a bioinfor-

\* Corresponding author.

E-mail address: [nessasinatti@gmail.com](mailto:nessasinatti@gmail.com) (V. de V. C. Sinatti).



**Fig. 1.** Schematic overview of the study approach used in the identification of TcRpi potential inhibitors. Abbreviations: LB, Ligand-Based pharmacophore hypothesis; SB, Structure-Based pharmacophore hypothesis.

matics tool that, besides other functionalities, predicts analogy and specificity among enzymes of distinct organisms [9].

In this context, the enzyme ribose 5-phosphate isomerase (Rpi; EC 5.3.1.6) is a promising target for the Chagas disease chemotherapy, since the Rpi of *T. cruzi* (TcRpi) and *H. sapiens* (HsRpi) are considered analogous enzymes [8,10]. The Rpi enzymes can be divided into two protein families: type A and type B. The TcRpi belongs to type B family (RpiB), generally found in prokaryotes and lower eukaryotes. On the other hand, the HsRpi belongs to type A family, which has representatives in most eukaryotes, as well as in prokaryotes. Type A Rpis (RpiA) are not present in trypanosomatids [11,12]. This enzyme (Rpi) plays a central role in the pentose phosphate pathway (PPP) that is responsible for the production of nucleotide precursors and NADPH. NADPH is critical for the protection of trypanosomatids against oxidative stress and for biosynthetic reactions. In the PPP, the Rpi enzyme catalyzes the reversible isomerization reaction between D-ribose-5-phosphate (R5P) and D-ribulose-5-phosphate (Ru5P) [13–17]. The TcRpi enzyme activity is possible mainly through the catalytic

residues Cys69, Ser71 and His102 which are present in the active site of the protein [11].

The importance of Rpis in cell viability was first shown by Skinner and Cooper in 1971, demonstrating that *Escherichia coli* mutants for RpiA were unable to grow in the absence of ribose [18]. Subsequently, Sorensen and Hove-Jensen showed that RpiA-/RpiB- double mutants of *E. coli* exhibited impaired growth in any tested conditions [19]. The deletion of RpiA gene of *Saccharomyces cerevisiae* was lethal [20]. RpiB silencing performed in *Trypanosoma brucei* using RNAi approach reduced the *in vitro* growth of these parasites, and the infected mice presented a lower parasitemia and higher survival [21]. Moreover, a recent study revealed that efforts to generate a null RpiB mutant in *T. brucei* failed, probably because of the essentiality of this enzyme. In *Leishmania infantum* the null mutants of RpiB were only possible when an episomal copy of RpiB was present, indicating the essentiality of RpiB for trypanosomatids in general [22]. For these reasons and due to the pivotal role of TcRpi in PPP, this enzyme is likely to be essential to *T. cruzi* viability, and it is an interesting rational target for antichagasic drug

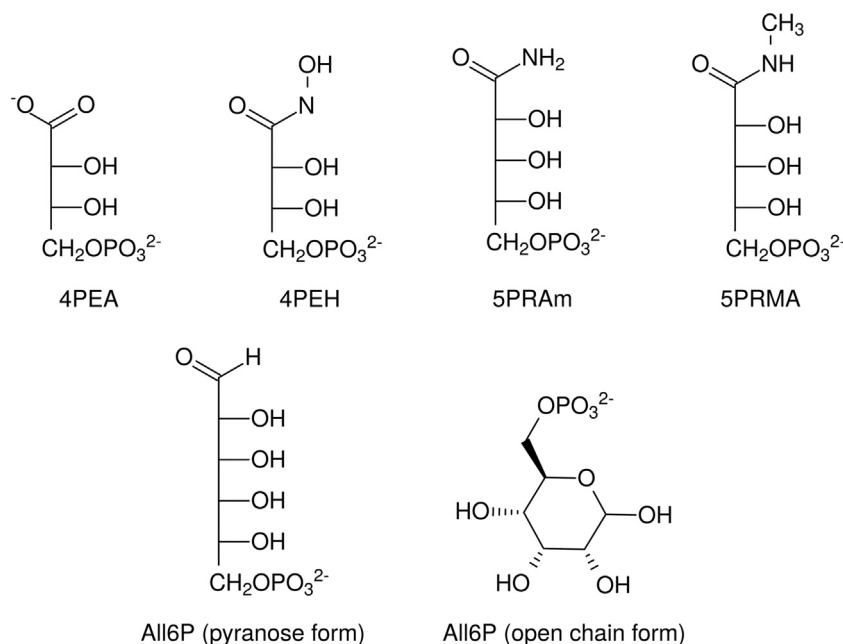


Fig. 2. Set of RpiB inhibitors used in ligand-based pharmacophore modeling.

design. This is particularly important since all proposed inhibitors for TcRpi are weak ( $K_i$  higher than 1.2 mM) and the more promisor, 4-phospho-D-erythrono-hydroxamic acid (4PEH), an analog to the reaction intermediate cis-enediolate, is described to be a potent competitive inhibitor of RpiA of spinach ( $\sim 30 \mu\text{M}$  at pH 7.5) [11,23].

The aim of this study is to propose new potential TcRpi inhibitors through a computer-aided drug discovery approach represented in Fig. 1. The first step included a structure-based and ligand-based pharmacophore modeling. The selected pharmacophore hypotheses along with a substructural and similarity search were used to screen the purchasable compounds from the ZINC database. The retrieved compounds were submitted to molecular docking in the TcRpi and HsRpi active sites, aiming to select molecules with potential to inhibit the parasite enzyme but not the human one. After an *in silico* ADME-T (absorption, distribution, metabolism, excretion and toxicity)/PAINS (pan-assay interference compounds) prediction, 211 molecules were proposed as potential TcRpi inhibitors. Further binding mode analyses and molecular dynamics simulations were performed with the three high ranked compounds and one positive control to access the stability of docked complexes. Two of the three compounds had good performance and can be considered as potential TcRpi's inhibitors.

## 2. Materials and methods

### 2.1. Dataset and protein analyses

In the present study, the crystal structures TcRpi bound to R5P substrate (PDB ID: 3K7S) and bound to the 4PEH (PDB ID: 3K8C) [12] inhibitor were obtained from the RCSB Protein Data Bank (<http://www.rcsb.org/>). The structural model of HsRpi was constructed with MODELLER (<https://salilab.org/modeller/>) [24] and *ab initio* modeling approaches as described in previous work [25]. This model was built as a tetrameric protein using the Hsa: 22934 sequence from Kegg database as query and the 1XTZ (identity of 46%), 1LK7 (identity of 42%), 3HHE (identity of 39%), 2F8M (identity of 35%) crystallographic proteins from PDB as templates. The identity of the active site residues of the templates in relation to the HsRpi sequence ranges from 75% to 94% (Supplementary Material Fig. S1).

A comparison between the primary, secondary and tertiary structures of HsRpi and TcRpi (PDB ID: 3K7S) was carried out aiming to evaluate the differences between these proteins. The primary structures were aligned against each other to access their identity and similarity using ClustalW through Multiple Sequence Viewer module of Maestro (version 9.3) of Schrödinger Suite. The secondary structures were also evaluated with Multiple Sequence Viewer module of Maestro. The 3D structural superposition was performed with TM-align (<http://zhanglab.ccmb.med.umich.edu/TM-align/>) [26] that compares the folds using an independent sequence-order approach and can be used to identify local similarities between the proteins. The active site of both proteins and the binding modes of the substrates were evaluated using Pymol v1.8.2.3 (<http://www.pymol.org/>). The electrostatic surface potentials of HsRpi and TcRpi were calculated by APBS software with the purpose to evaluate the electrostatic differences between their active sites.

A set of three known inhibitors within  $\text{IC}_{50} \leq 5 \text{ mM}$  against TcRpi was selected from literature to perform ligand-based pharmacophore modeling and as positive controls in molecular docking. The selected inhibitors were 4PEH (4-phospho-D-erythrono-hydroxamic acid), 4PEA (4-phospho-D-erythronate) [11] and the open and ring form of All6P (D-Allose-6-Phosphate) [12]. Besides that, two another inhibitors, 5PRAm (5-phospho-D-ribo-5-aminopentamide) and 5PRMA (N-(5-phospho-D-ribo-5-onyl)-methylamine), which are strong inhibitors to RpiBs of other organisms like *Mycobacterium tuberculosis* and *E. coli* [27], were selected to add variety to the ligand-based pharmacophore modeling (Fig. 2). Substrates and other inhibitors described in the literature for RpiBs from different organisms were used as a test set to validate the generated pharmacophore hypotheses [27–29]. These inhibitors and substrates structures were designed in Maestro. The compound library to be screened in the molecular docking was obtained from ZINC Database (<http://ZINC.docking.org/>) [30] – purchasable subset (12/2014) with 22,724,825 compounds – through search based on pharmacophore hypotheses and 2D similarity as described below. A second set of 1000 compounds was randomly obtained from the ZINC database, using the tool from the Bioinformatics and Chemical Informatics Research Center (<http://bcirc.docking.org/random.shtml>). In addition, a decoy set contain-



ing 700 compounds was generated using the database Directory of Useful Decoys, Enhanced (DUD-E; <http://dude.docking.org>) [31].

## 2.2. Ligand-based pharmacophore modeling

Ligand-based pharmacophore model was built with the Pharmagist web server ([bioinfo3d.cs.tau.ac.il/PharmaGist/](http://bioinfo3d.cs.tau.ac.il/PharmaGist/)) [32], using the set of known RpiB inhibitors selected from literature as input. The pharmacophore hypotheses were chosen based on some criteria: (i) high scores; (ii) the inclusion of all known RpiB inhibitors in the construction of the hypotheses; (iii) having at least one negative ionizable pharmacophoric feature, previously known as important to the protein-ligand interaction and (iv) the return of the known Rpi substrates to ZINC Database through the ZINCPharmer ([ZINCpharmer.csb.pitt.edu/](http://ZINCpharmer.csb.pitt.edu/)) [33].

## 2.3. Structure-based pharmacophore modeling

ZINCPharmer web server was used to generate a structure-based pharmacophore model. For this model, the PDB crystallographic structures 3K7S and 3K8C containing, respectively, the open-chain form of the ribose 5-phosphate and the 4PEH inhibitor were used since they have pharmacophoric features and chemical groups in common, known as relevant to the ligand-receptor interactions.

## 2.4. Pharmacophore validation

The performance of the generated pharmacophore models in discriminating between known active and inactive compounds was validated by calculating enrichment factor and AUC values (area under curve) of the ROC (receiver operating characteristic) curves. In order to perform the validation approach, for each pharmacophore model, a data set containing the test set (active compounds not used in the pharmacophore generation) and the DUD-E decoys was prepared, and conformers were generated. The pharmacophore models were used to screening the data sets through the Pharmer software (pharmacophore software used by the ZincPharmer) [34]. The hit compounds were scored and ranked using the RMSD of the best fit conformation to the pharmacophore query and based on this the metrics for performance assessment were calculated.

## 2.5. ZINC database screening

Both the ligand-based and structure-based pharmacophore models were used as input in the ZINCPharmer web server aiming to search compounds from the ZINC Database which match the hypotheses of the models. Additionally, a structural similarity search was carried out in the ZINC Database in order to get compounds with more than 50% similarity to TcRpi substrates (R5P open chain form, R5P ring form, and Ru5P), or compounds containing phosphate and/or hydroxyl groups in their substructure. The searches only included the purchasable subset of the ZINC Database. The retrieved compounds, both from the pharmacophore models and also from the similarity search, were used to construct the compound library of this study. After that, the library was submitted to Knime software [35] to remove duplicates and was filtered with Lipinski's rule of five [36].

## 2.6. Protein preparation

The crystal structure of TcRpi bound to R5P substrate with 1.9 Å resolution (PDB ID: 3K7S) was used for molecular docking with Glide software (version 5.8) [37]. The ligand, and C and D subunits were excluded from the PDB file. Just A and B subunits were used.

The exclusion of two from four subunits was carried out since TcRpi acts as a dimer in solution [12] and the active sites are between each pair of subunits (A-B and C-D). Moreover, we performed an analysis of the conserved waters in the TcRpi protein (PDB ID: 3K7S). Other proteins from PDB database annotated with Rpi enzymatic activity (Enzyme Commission (EC) Number 5.3.1.6) and with more than 50% of similarity in relation to TcRpi were used in this analysis (Supplementary Material Table S1). The water molecules in the active site of PDB 3K7S that are also present in more than 80% of the analyzed structures were considered highly conserved waters. In order to evaluate if these conserved waters would be maintained or excluded, redocking studies and performance analysis (using the active compounds – substrates and TcRpi inhibitors – and DUD-E decoys) were performed before the molecular docking steps.

The protonation states of TcRpi residues were assigned at pH 7 considering previous studies of the enzyme reaction mechanism and using PROPKA program via Protein Preparation Wizard [11,12,38]. In the reaction mechanism, it was proposed for RpiB of *Clostridium thermocellum* that the catalytic Cys69, despite the need of its deprotonated form for the isomerization, remains protonated in the active side and it is deprotonated only in the beginning of the reaction through the Asp10 [38]. Besides that, the His102 needs to be protonated for the opening of the R5P ring form [28]. For this reasons, the catalytic residues Cys69 and His102 were set to the neutral and protonated forms, respectively. In the structural model of HsRpi [25], water molecules and ligands were not considered in protein preparation. The protonation states of the residues were assigned at pH 7, and we also followed the protonation set up of Capriles and collaborators [25] for the catalytic residues with Glu182 considered negatively charged.

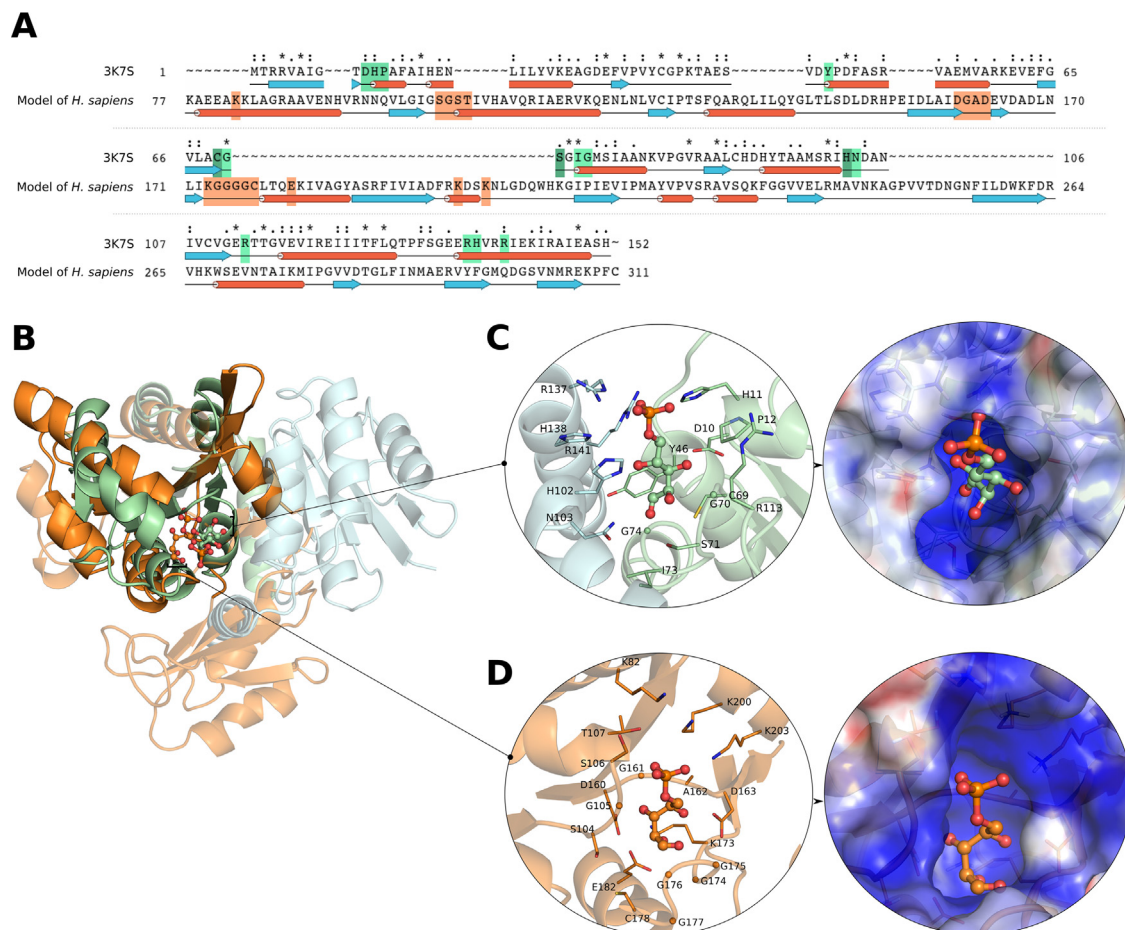
## 2.7. Ligand preparation

All 3D structures from known Rpi inhibitors, substrates, and compounds from ZINC were retrieved in the mol2 format. All ligands were prepared with Ligprep software from the Schrödinger Suite and energy minimized using OPLS-2005 force field. For each ligand, all possible ionization states were assigned at pH  $7 \pm 1$ , the chirality was retained, and the tautomers were generated.

## 2.8. Molecular docking

Binding pose predictions and scoring of compounds from the generated *in silico* library and the random set, against the TcRpi target, were performed through a molecular docking approach. To accomplish the molecular docking, the extra precision protocol of Glide software (Glide XP) [39] was chosen – Glide XP has a higher enrichment performance and binding affinity prediction than other docking protocols within Glide. Besides that, the Glide XP is better at eliminating false positives [39]. The prepared TcRpi crystal structure was used for grid generation on Glide XP, where the center is the active site of the receptor with the co-crystallized ligand. The grid dimensions were  $12 \times 12 \times 12$  Å of inner box and  $22 \times 22 \times 22$  Å of outer box. That grid included, besides other residues, all the active site residues of TcRpi: Asp10, His11, Phe12, Tyr46, Cys69, Gly70, Ser71, Ile73, Gly74, His102, Asn103, Arg113, Arg141, Arg137, and His138. Aiming to validate the molecular docking approach, the co-crystallized R5P substrate was submitted to a redocking in the target active site.

In this TcRpi structure-based virtual screening, the top-scored compounds were selected using the determined cut-off with the docking score (DS) values of the positive controls (4PEH, 4PEA and the open and ring form of All6P). These top-scored compounds were also docked into the HsRpi active site aiming to eliminate compounds with the potential to inhibit this human enzyme. That molecular docking was performed with the grid of



**Fig. 3.** Comparison between primary, secondary and tertiary structures of HsRpi and TcRpi. (A) Sequence alignment of HsRpi (Hsa:22934) and TcRpi (Tcr:509199.24) proteins using Maestro/Clustaw. The residues of the active site of TcRpi are highlighted in green and the catalytic residues are represented in dark green. The residues of the active site of HsRpi are depicted in orange. Secondary structures are represented by blue arrows ( $\beta$  sheet), red cylinders ( $\alpha$ -helix) and black lines (loops). (B) Three-dimensional structure alignment based on the structural similarity between TcRpi crystallographic structure (PDB 3K7S, green/blue – chains A/B) and HsRpi model (orange – chain A). The superposition between these structures is highlighted with dark orange and dark green. (C) R5P pose at the active site of TcRpi; (D) R5P pose at the active site of HsRpi. R5P substrate of TcRpi and HsRpi are represented as thick ball and stick with carbons colored in green and orange, respectively. Side chains of residues for both active sites are represented as thin sticks, except glycines that are represented as small balls. Electrostatic surface potentials of the active sites are also shown (red is  $-25$  kBT and blue is  $25$  kBT) with a transparent surface. Protein structural figures were generated using Pymol v1.8.2.3.

$12 \times 12 \times 12 \text{ \AA}$  and centroid coordinates of  $X = 18.36$ ;  $Y = 11.22$  and  $Z = 89.57$ . The grid included the active sites residues of HsRpi: Lys82, Thr107, Ser104, Gly105, Ser106, Asp160, Gly161, Ala162, Asp163, Lys173, Gly174, Gly175, Gly176, Gly177, Cys178, Glu182, Lys200, and Lys203.

### 2.9. ADME-T and PAINs predictions

The set of potential selective inhibitors to TcRpi (top-rated compounds in the molecular docking against TcRpi and with high docking scores for the HsRpi active site) was submitted to ADME-Tox (absorption, distribution, metabolism, excretion and toxicity) and PAINS (pan-assay interference compounds) analyses using FAF-Drug3 ([fafdrugs3.mti.univ-paris-diderot.fr/](http://fafdrugs3.mti.univ-paris-diderot.fr/)) [40].

### 2.10. Molecular dynamic simulations

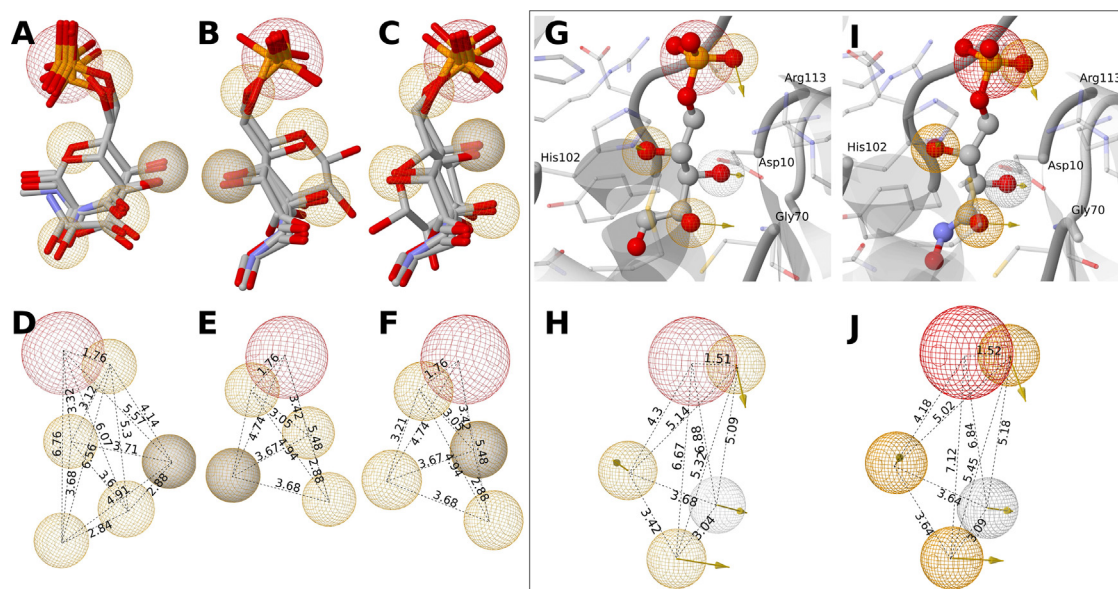
Compound structures were submitted to Automated Topology Builder (ATB) and Repository [41] in order to build suitable topologies to Molecular Dynamic (MD). MD was performed using Gromacs 5.1.2 [42] program and Gromos 96 force-field. The following treatment was done to prepare the systems for the production runs: (i) energy minimization in vacuum, using the steepest descent energy

(2 ps); (ii) energy minimization in water (10 ps); (iii) MD simulation in isochoric-isothermal ensemble (NVT) (100 ps); (iv) MD simulation in isothermal-isobaric ensemble (NpT) with restraints (100 ps); (v) MD simulations in ensemble NpT (1 ns). The system was initially subjected to minimization using the steepest descent energy. The minimizations were completed when the tolerance of  $1000 \text{ kJ/mol}$  was no longer exceeded. After these thermalization steps, production trajectories of 50 ns for each complex were generated. The simulations were performed using periodic boundary conditions, at 300 K and 1 atm pressure. Electrostatics were evaluated using the particle mesh Ewald algorithm [43]. All  $\beta$ -factor, root mean square fluctuation (RMSF), hydrogen bonds analyses were performed using Gromacs software package, Maestro and Visual Molecular Dynamics (VMD) v1.9.3 [44].

## 3. Results and discussion

### 3.1. Analyses of HsRpi and TcRpi proteins

A structural comparative analysis of the pathogen protein toward the correspondent human protein is a relevant step in the rational drug discovery. This analysis is necessary to evaluate the chance of potential off-target behaviors even if these proteins are described as analogous enzymes. The comparison of the pri-



**Fig. 4.** Ligand-based and structure-based pharmacophore hypotheses. (A) LB1, (B) LB6 and (C) LB7 ligand-based pharmacophore hypotheses; (D) LB1, (E) LB6 and (F) LB7 schematic representation with the distances between the features in angstrom (Å). (G) and (I) Selected pharmacophore features of SB1 and SB2 structure-based pharmacophore hypotheses in TcRpi active site (PDB code: 3K7S). (H) and (J) Schematic representation of SB1 and SB2 structure-based pharmacophore hypothesis with distances between the features in angstroms (Å); Pharmacophore features are labeled with red (negative ionic), orange (H-bond acceptor), white (H-bond donor) and gray (superposition of H-bond acceptor and donor features) colors. The figures were generated through jmol in ZINCpharmer interface.

**Table 1**

Results of the pharmacophore validation.

Pharmacophore	TA	TI	TP	TN	FP	FN	Se	Sp	EF	AUC
LB6.1	10	700	7	700	0	3	0.7	1	71	0.818
LB6.6	10	700	10	667	33	0	1	0.952	16.511	0.994
LB6.7	10	700	10	674	26	0	1	0.962	19.722	0.992
SB1	15	700	15	411	289	0	1	0.587	2.351	0.886
SB2	15	700	15	679	21	0	1	0.97	19.861	0.998

Abbreviations: TA total actives; TI total inactives; TP true positives; TN true negatives, FP, false positives; FN false negatives; Se; Sensitivity; Sp Specificity; EF Enrichment factor; AUC area under the curve.

mary, secondary and tertiary structures of HsRpi and TcRpi (Fig. 3) revealed the difference between these proteins, as expected due to the analogy.

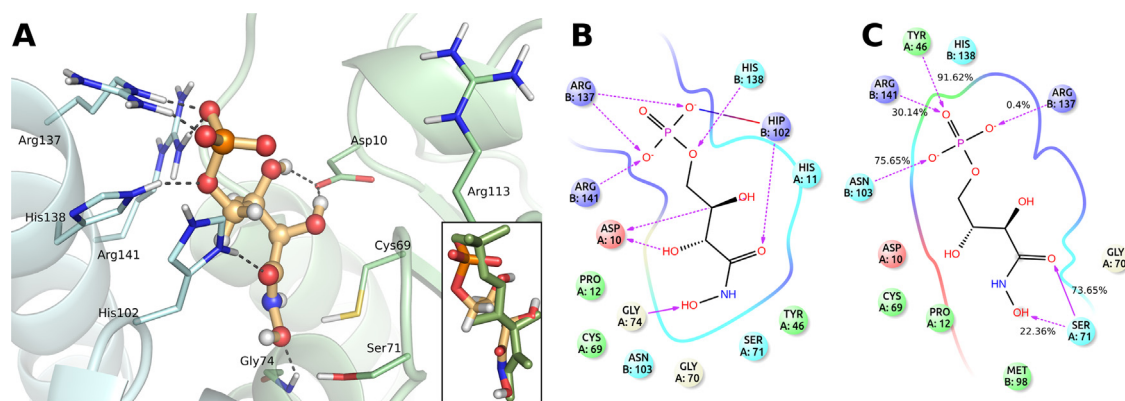
The global alignment of HsRpi and TcRpi protein sequences showed an identity and similarity of 8% and 21% (Fig. 3A), respectively. It is important to note that sequence alignment failed to align residues from the active site of HsRpi and TcRpi, including catalytic ones. Three-dimensional structure alignment between TcRpi crystallographic structure (PDB 3K7S) and the model constructed for HsRpi showed a poor overall structural fitting, suggesting a possible superposition of unrelated proteins (Fig. 3B). This is evidenced by the obtained TM-score value of 0.36065 since only superimposed proteins with score values higher than 0.5 are likely to be related proteins with the same fold [26]. These results emphasize the analogy between these two proteins, showing that despite the same function their sequences and overall structures are different.

A visual comparison of the architecture of both active sites reveals the presence of many aromatic residues (Tyr46, His11, His102 and His138) in TcRpi, while HsRpi shows none (Fig. 3C and D) [25]. Furthermore, the distribution of (potentially) negatively charged residues are different; TcRpi has the Cys69 and Asp10 and the HsRpiA has the Asp160, Asp163, and Glu182. Despite these differences, both active sites present basic residues near the highly electronegative phosphate group of R5P: His11, His102, Arg113, His138, Arg137 and Arg141 on TcRpi, and Lys82, Lys200 and Lys203 on HsRpi. This feature reflects the electrostatic surface potential similarities observed in the active sites, characterized by predom-

inantly electropositive surfaces (Fig. 3C and D). Even with the described analogy, these particular convergent similarities in HsRpi and TcRpi active sites could be supposed, since they are enzymes that catalyze the same reaction and therefore interacts with the same substrate. It supports the demand for molecular docking studies not only with TcRpi, but also with HsRpi as well, aiming to select compounds able to inhibit the parasite enzyme, but with little/none activity to HsRpi.

### 3.2. Pharmacophore modeling

To predict the essential chemical features – and their spatial organization – necessary in a compound to interact with the target protein TcRpi, ligand-based and structure-based pharmacophore modeling was carried out. A total of 40 ligand-based pharmacophore hypotheses was generated from the set of selected RpiB inhibitors (Supplementary Material Table S2) and three hypotheses were selected. These hypotheses (LB1, LB6 and LB7, represented in Fig. 4A–F) were chosen since they make use of all RpiB inhibitors, have at least one negative ionizable pharmacophoric feature (important because Rpi active site is characterized by having an electropositive nature) and could retrieve in the ZINC Database (using ZINCPharmer) the substrates of Rpi – R5P (open-chain and ring form) and Ru5P. The LB1 hypothesis has besides the negative ionizable feature (N), five hydrogen bond acceptors (HA) and one hydrogen bond donor (HD) features (Fig. 4A and D). In addition, the hypotheses LB6 and LB7 have one N, four HA and one HD features.



**Fig. 5.** The predicted binding interaction between 4PEH and TcRpi protein. (A) The binding mode of 4PEH in the active site of TcRpi predicted in the molecular docking approach. The square box shows the superimposition of the 4PEH docking pose (orange) and the crystallographic pose (dark green). The 4PEH is represented as orange ball and sticks in the main image, the protein subunits A and B are represented in light green and light cyan colors, respectively, and the hydrogen bonds interactions are represented as gray dashes; (B) The 2D interaction diagram of 4PEH in the active site of TcRpi, in accordance with the docking results. (C) 2D interaction diagram of 4PEH binding mode after the molecular dynamics simulations and the occupancy of the hydrogen bonds. Residues are distinguished in negatively charged (pink), positively charged (purple), polar (blue) and hydrophobic (green). The hydrogen bonds are represented as purple arrows and salt bridges are represented as blue/red lines. The 3D and 2D images were generated using Pymol v1.8.2.3 and Maestro, respectively.

**Table 2**  
Redocking studies and performance analyses using the TcRpi with and without conserved waters in the active site.

Conditions	Redocking		AUC	EF1%	EF2%
	RMSD	DS			
TcRpi without waters	0.39	−10.240	0.96	29	22
TcRpi with conserved waters	0.85	−7.424	0.85	14	14

Abbreviations: RMSD, root mean standard deviation; DS; docking score; AUC, area under the curve; EF1%, enrichment factor at 1% of the database; EF2%, enrichment factor at 2% of the database.

Both hypotheses differ only by the position of HD feature, as can be seen in Fig. 4B, C, E, and F.

The structure-based pharmacophore hypotheses were carried out with 3K7S and 3K8C PDB structures. We selected significant and essential features in the receptor-ligand molecular interaction considering the chemical characteristics of the ligand and the chemical environment of the active site. Two similar structure-based pharmacophore hypotheses were generated with the same pharmacophoric features (SB1 and SB2). As can be seen in Fig. 4G, H, I, and J, the selected features were: (i) one negative ionizable feature for the ligand phosphate group that electrostatically interacts with the potentially positively charged residues in the active site (His11, His102, Arg113, Arg137, His138 and Arg141); (ii) three hydrogen acceptors, two related to the hydrogen bond with Arg113 (the oxygen of phosphate group), His102 and Gly70, respectively; and, (iii) one hydrogen donor making possible a hydrogen bond with the Asp10. These features reflect relevant chemical characteristics that must exist in the ligand to interact with TcRpi receptor and, therefore, they are important to retrieve biologically active compounds in the virtual screening steps.

All these ligand-based and structure-based pharmacophore hypotheses were validated for their performance in differentiating between active compounds and decoys (supposed to be inactive compounds). For the performance assessment, the enrichment factor, sensitivity, specificity and AUC of the hypotheses were calculated (Table 1), and the ROC curve was also generated (Supplementary Material Fig. S2). The enrichment factors ranged from 2.35 to 71 and these values indicate the number of times that pharmacophore models performed better results than a random selection. The high values of sensitivity (Se) (0.7–1) reflected the good ability

of the models to retrieve the active compounds. The specificities were also very high, except for the SB1 model ( $Sp = 0.58$ ), indicating the ability in discarding inactive compounds. Furthermore, all the five hypotheses presented an AUC of over 0.81, suggesting that these pharmacophore models have good quality.

### 3.3. Database searching based on pharmacophore hypothesis and 2D similarity search

Based on the generated pharmacophore hypotheses (three ligand-based and two structure-based pharmacophore hypotheses), it was possible to screen the ZINC Database aiming to retrieve molecules presenting the same required features and architecture of the pharmacophore models. In all selected hypotheses, the radius of hydrogen bond donor/acceptor pharmacophore feature and negatively charged pharmacophore feature were 1 Å and 1.5 Å, respectively. As a result, 18,038 hit compounds (without considering duplications) were obtained. Moreover, a structural similarity search carried out in ZINC Database retrieved 2,145 compounds. All these compounds together, both from pharmacophore models and similarity search, were submitted to the generation of tautomers and suitable ionization state. The 24,863 final structures represented the compounds library of this *in silico* study.

### 3.4. Molecular docking

Redocking and performance analyses were carried out using TcRpi (PDB 3K7S) with absence and presence of the conserved water molecules in the active site. These analyses were performed to validate the molecular docking approach and to assess the ability of the TcRpi with and without waters in distinguishing active compounds from inactive ones. The redocking considered the co-crystallized R5P substrate in the active site of TcRpi with and without conserved waters, resulting in a docked pose with a root mean square deviation (RMSD) of 0.39 Å and 0.85 Å, respectively (Table 2). Moreover, the structure without waters showed a higher AUC and enrichment factor values (Table 2; Supplementary Material Fig. S3) than the structure with conserved waters, suggesting that the non-consideration of conserved waters in the active site of TcRpi increases the discrimination power of molecular docking. These results indicate that the best choice for TcRpi is to perform molecular docking studies without considering conserved waters.

All compounds retrieved from the pharmacophore models and the similarity screening in the ZINC Database were then docked

**Table 3**

ADME properties, PAINS filters and Docking Scores of the positive controls and the ten best scoring compounds for TcRpi in relation HsRpi.

Compounds	MW	logP	HBD	HBA	LV	OB VEBER	OB EGAN	PAINS Filter	DS TcRpi	DS HsRpi	H Bonds TcRpi
ZINC36975961	254.219	-2.03	2	8	0	Good	Good	Accepted	-12.735	-8.877	6
ZINC63480117	341.298	-0.05	2	9	0	Good	Good	Accepted	-11.927	-8.503	8
ZINC43763931	249.216	-3.9	4	9	0	Good	Good	Accepted	-11.849	-8.571	8
ZINC03869847	347.221	-3.4	5	12	1	Good	Good	Accepted	-11.659	-8.465	6
ZINC19300946	204.184	-1.08	3	7	0	Good	Good	Accepted	-11.601	-8.795	8
ZINC08585028	402.145	-5.21	4	14	1	Good	Good	Accepted	-11.429	-8.625	8
ZINC35456138	356.368	-0.18	4	8	0	Good	Good	Accepted	-11.41	-8.723	7
ZINC34924542	269.251	-0.08	4	7	0	Good	Good	Accepted	-11.383	-8.671	7
ZINC02047359	276.135	-4.46	4	10	0	Good	Good	Accepted	-11.314	-6.828	8
ZINC35643607	284.262	0.37	3	7	0	Good	Good	Accepted	-11.109	-6.913	5
oALL6P	260.14	-4.53	6	9	1	Good	Good	Rejected	-12.283	-11.507	10
4PEH	231.100	-3.97	6	9	1	Good	Good	Accepted	-10.152	-9.898	8
rALL6P	260.14	-4.89	6	9	1	Good	Good	Accepted	-9.484	-9.449	7
4PEA	216.08	-3.23	5	8	0	Good	Good	Accepted	-8.918	-8.818	9

Abbreviations: MW, molecular weight; LogP, octanol/water partition coefficient; HBD, hydrogen bond donor; HBA, hydrogen bond acceptor; LV, Lipinski Violation; OB, oral bioavailability; PAINS, pan-assay interference compounds; DS, docking score; H bonds, hydrogen bonds; oALL6P, open-chain form of D-allose-6-phosphate; rALL6P, ring form of D-allose-6-phosphate; 4PEH, 4-phospho-D-erythronhydroxamic acid; 4PEA, 4-phospho-D-erythronate.

**Table 4**

Analyses of the interactions observed in the molecular docking studies with the three final hits and 4PEH.

Compounds	DS	Asp10	His11	Gly70	Ser71	Gly74	His102	Asn103	Arg137	His138	Arg141
ZINC36975961	-12.735	-	-	-	-	H	H; SB	H	H	H	H
ZINC63480117	-11.927	-	H; pi-pi	H	-	H	SB	H	H	H	H
ZINC43763931	-11.849	H	H	H	H	H	SB	-	H	H	H
4PEH		H	-	-	-	H	H; SB	-	H	H	H

Abbreviations: DS, docking score; H, hydrogen bond; SB; salt bridge; pi-pi; pi-stacking.

at the TcRpi active site without crystallographic waters. These molecular docking studies were performed enabling to predict the binding pose of the selected compounds. In addition, the compounds were ranked based on their docking scores aiming to propose potential inhibitors to TcRpi. The set of three known TcRpi inhibitors used as positive controls in this molecular docking, 4PEH (DS of -10.152), open-chain All6P (DS of -12.283), ring All6P (DS of -9.484) and 4PEA (DS of -8.918), was important to determine the docking score cut-off of -8.918. To evaluate how stringent was this cutoff, a new molecular docking study was performed using a set of random compounds from ZINC database (likely non-inhibitors). The DS of these compounds were compared to the results obtained with the previously generated library – that is enriched with potential TcRpi inhibitors. The distributions (Supplementary Material Fig. S4) of the DS from these two sets were evaluated using Welch's *t*-test. A *p*-value of 2.2e-16 was obtained, indicating that the two distributions are significantly different. The cutoff of -8.918 was stringent enough to exclude all the compounds from the random set, indicating it is a good value to separate non-inhibitors from inhibitors in the enriched library of this study. Consequently, the compounds which present DS values equal or less than the DS value of 4PEA inhibitor was considered as best-scored compounds. A total of 1071 top-scored compounds were selected (Supplementary Material Table S3).

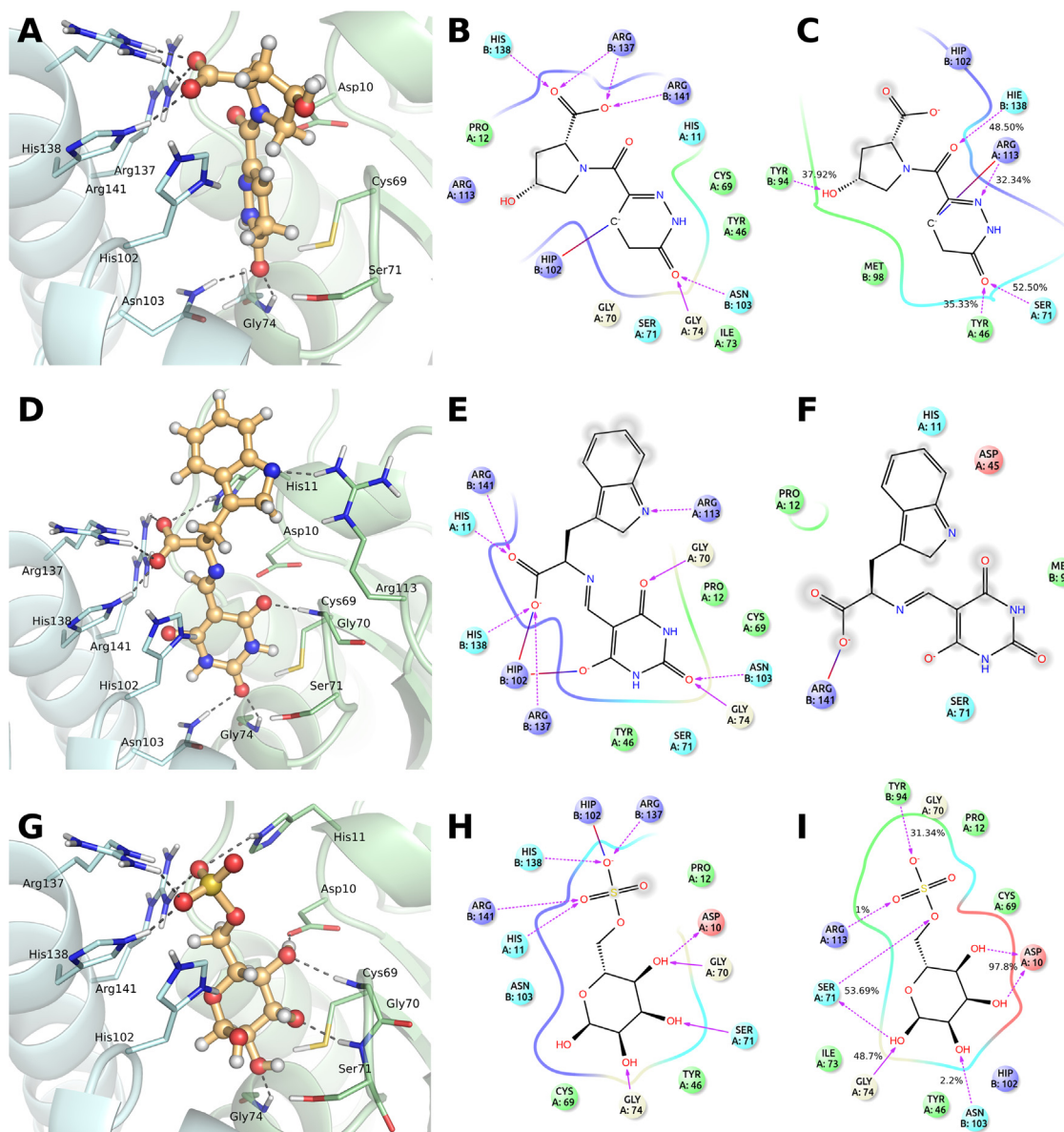
These top-scored compounds were also docked into the HsRpi active site aiming to eliminate compounds with the potential to inhibit this human enzyme. Although HsRpi is analogous to TcRpi, the active sites of these proteins have some similarities. For this reason, it is important to ensure the selection of inhibitors with a high potential affinity for the *T. cruzi* target and low or none interaction with HsRpi. Positive controls of these assay were the open chain form and ring form of R5P and Ru5P substrate which yielded the DS values -8.973, -9.823 and -10.070, respectively. We selected hits that had DS values greater than -8.973 since this would result in a set of compounds with lower inhibition potential towards HsRpi. After this second screening, we selected a total of

331 compounds with potential high affinity to TcRpi and likely low or none inhibitory activity towards the human analogous enzyme (Supplementary Material Table S4).

The potential selectivity observed in some of these 331 compounds appears to be mainly due to the aromatic residues – Tyr46, His11, His102 and His138 – present in the active site of the TcRpi and absent in the active site of the human protein. As consequence of the screening based on the selectivity, a frequency of more than 70% of these final retrieved compounds contained aromatic rings. Most of these aromatic rings performed pi-stacking interactions and/or cation-pi interactions with the protonated histidine (His102) present in the active site. In the pi-stacking interactions, the T-shaped (edge-to-face) arrangement predominate. Studies indicate that T-shaped conformation (depending upon the solvent conditions) and pi-stacking interaction of aromatic rings with protonated histidines are energetically more favorable than in the opposite cases [45,46]. Although the lysines present in the active site of HsRpi could also make cation-pi interactions with the selective compounds, this was observed at a much lower frequency – 2.5 times less – than in TcRpi. In fact, lysines are known to make much less this kind of interaction than arginines, for example [47]. Besides the contribution of the aromatic residues to the selectivity, the architecture and the distinct arrangement of the amino acids at the TcRpi and HsRpi active sites appeared to influence the selectivity of many compounds by changing, for example, their hydrogen bonding configuration. The cavity of the TcRpi is deeper, while the HsRpi is shallower and described to have a smaller volume [25] (Supplementary Material Fig. S5), which contributes to an unfavorable solvent exposure of some potential selective compounds.

### 3.5. ADME-T and PAINS predictions

ADME-T and PAINS predictions were important in determining the pharmacokinetic properties and potential toxicity, and to identify the presence of undesirable moieties or potentially promiscuous compounds for the 331 previously selected compounds.



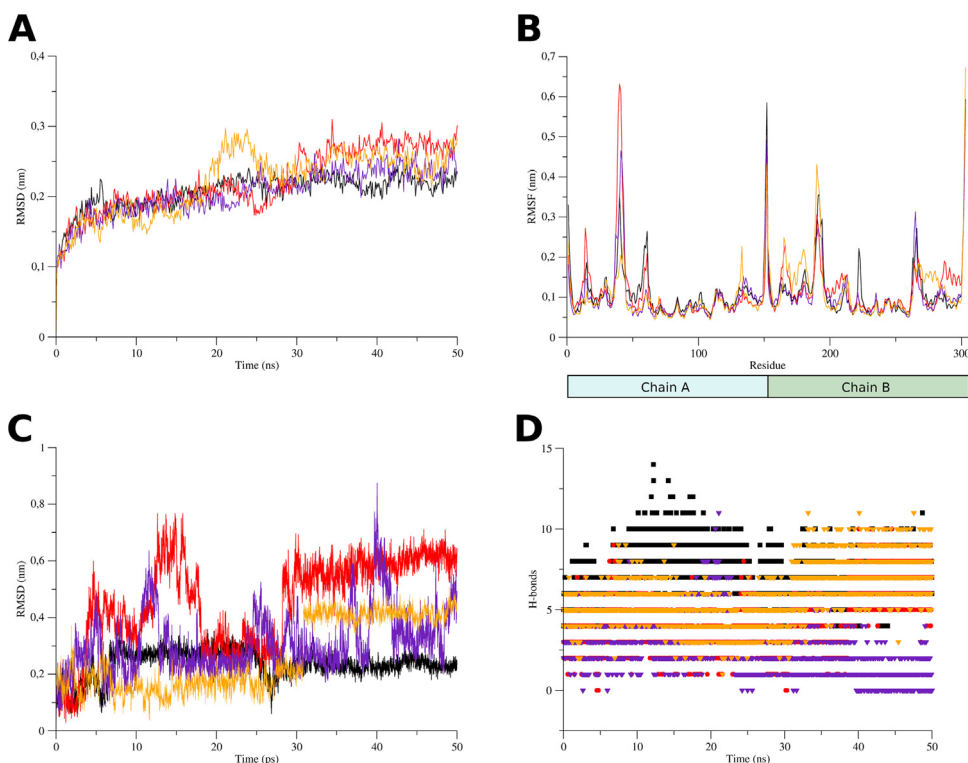
**Fig. 6.** The predicted binding interaction between the ZINC36975961, ZINC63480117 and ZINC43763931 potential inhibitors and TcRpi protein. (A), (D) and (G) represent the binding mode obtained with the molecular docking of ZINC36975961, ZINC63480117, and ZINC43763931, respectively, in the active site of TcRpi. The inhibitors are represented as orange ball and sticks, the protein subunits A and B are represented in light green and light cyan colors, respectively, and the hydrogen bonds interactions are represented as gray dashes; (B), (E) and (H) are the 2D interactions diagram of molecular docking results of ZINC36975961, ZINC63480117 and ZINC43763931, respectively. The 2D interactions obtained after the molecular dynamics simulations of ZINC36975961, ZINC63480117 and ZINC43763931 and the occupancy of the hydrogen bonds are represented in (C), (F) and (I), respectively. The 2D Residues are distinguished in negatively charged (pink), positively charged (purple), polar (blue), hydrophobic (green). The hydrogen bonds are represented as purple arrows and salt bridges are represented as blue/red lines. The 3D and 2D images were generated using Pymol v1.8.2.3 and Maestro, respectively.

Among these, 211 compounds passed the screening of ADME-T and PAINS predictions. A total of 120 were classified as intermediates or rejected representing molecules that include low or high-risk structural alerts and/or belong to the PAINS class. Table 3 summarizes the ADME-T and PAINS results obtained for the ten best scoring compounds identified in the TcRpi molecular docking assay and for known TcRpi inhibitors. From these set of known TcRpi inhibitors, the All6P open form was the only compound rejected in the ADME-T and PAINS analysis because of the presence of a high-risk aldehyde group. The ten compounds had better docking scores for TcRpi (DS values  $\leq -11.109$ ) than most of the known TcRpi inhibitors- mainly 4PEH, 4PEA and the ring form of All6P-, suggesting that they could be more interesting competitive inhibitors for the parasite enzyme. Furthermore, the known TcRpi inhibitors were also well scored for

HsRpi (DS values  $\leq -9.449$ ) and appear as potential inhibitors for RpiA too, besides RpiB, as already demonstrated for 4PEH and 4PEA in RpiA of spinach [23]. On the other hand, the new compounds seem to be more specific for RpiB, which is a desirable characteristic in lead compounds for the development of new antichagasic drugs less toxic to humans.

### 3.6. Molecular interaction analyses of the best scoring compounds for TcRpi

Three compounds with the best docking scores, ZINC36975961, ZINC63480117 and ZINC43763931 (Fig. 6), which have no reported activity in RpiBs yet, and one positive control (4PEH, Fig. 5) had further analyses of their binding poses in TcRpi's active site



**Fig. 7.** Molecular dynamics studies of the novel potential inhibitors and a 4PEH in TcRpi active site. (A) Molecular dynamics simulation trajectory results illustrating the RMSD of C $\alpha$  atoms of TcRpi protein in complex with ZINC36975961 (red), ZINC63480117 (indigo), ZINC43763931 (orange) and 4PEH positive control (black); (B) RMSF of C $\alpha$  atoms of TcRpi protein in the complexes; (C) RMSD of the potential inhibitors and the positive control; (D) Hydrogen-bonds between TcRpi protein and ZINC36975961, ZINC63480117, ZINC43763931 and 4PEH throughout the simulations.

and molecular interactions (Table 4). The positive control 4PEH inhibitor that already had crystallographic data to TcRpi was redocked into TcRpi active site permitting to access the docking score of the obtained pose and compare with the docking score of another inhibitors candidates. The redocked pose had a slightly different binding mode in relation to the crystallized pose with an RMSD of 1.33 Å (11) (Fig. 5A). In the redocked pose, the phosphate group interacted through four hydrogen bonds with the imidazole group of His138 and with the guanidino groups of Arg137 and Arg141 residues (Fig. 5B). Besides that, there was one possible hydrogen bond with the guanidino group of Arg113 dependently of its side chain conformation. One electrostatic interaction of the phosphate group with the imidazole ring of His102 was also observed. Carbonyl group of 4PEH interacted with the imidazole ring of His102 through one hydrogen bond. The three hydroxyl groups made interactions through three hydrogen bonds with the carboxylate group of Asp10 and with the backbone of Gly74.

The ZINC36975961 compound exhibited the best docking score of  $-12.735$  and had interactions with various amino acid residues of TcRpi active site by forming six hydrogen bonds and one salt bridge (Fig. 6A and B). The carboxylate group of the candidate inhibitor made three hydrogen bonds with the guanidino groups of Arg137 and Arg141 and one hydrogen bond with the NE2 of the imidazole group of His138. The carbonyl of the amide group presented one hydrogen bond with the backbone of Gly74 and another hydrogen bond with the amide group of Asn103 side chain. Moreover, one salt bridge interaction was observed between the carbanion of ZINC36975961 and the ND1 of the imidazole group of His102.

The second best hit, ZINC63480117 compound, had a docking score of  $-11.927$  and was predict to make eight hydrogen bonds, two electrostatic interactions and one possible pi-pi interaction (Fig. 6D and E). In this binding mode, the nitrogen of the

indole moiety made interactions through hydrogen bonding with the guanidino group of Arg113. In a similar way that described for ZINC36975961, the ligand carboxylate group was stabilized through hydrogen bonds with the side chains of Arg137, Arg141, and His138 and one hydrogen bond with the imidazole group of His11 that was not in ZINC36975961 interaction. Besides that, this carboxylate group had an electrostatic interaction with the imidazole group of His102, and that same residue interacted with the oxygen anion of the pyrimidine derivative moiety. The two carbonyl groups of the pyrimidine derivative moiety of the ligand made three hydrogen bonds in total: one with the amide group of Asn103 side chain and two with the backbones of Gly70 and Gly74. Besides, there was a possible pi-pi interaction between the benzene ring of the ligand's indole moiety with the imidazole group of His11.

The binding mode of the ZINC43763931 compound at the active site of TcRpi showed eight hydrogen bonds and one salt bridge, representing a docking score of  $-11.849$  (Fig. 6G and H). In the same way of the carboxylate group of ZINC63480117, the sulfate ion of ZINC43763931 made interactions through four hydrogen bonds with the side chains of the residues Arg137, Arg141, His138 and His11 and one electrostatic interaction with His102. Three of four ligand hydroxyls interacted through four hydrogen bonds with the carboxylate group of Asp10 and with the backbones of the Gly70, Ser71 and Gly74 residues.

These compounds showed promising molecular docking results once they were found to form key interactions with some catalytic residues of the target protein including His102 and Ser71 and with substrate-binding residues like His11, Asp10, Gly70, Gly74, Asn103, Arg113, Arg137, His138, and Arg141 [11,12]. Similarly, some of these interactions are also found in the cocrystallized 4PEH known inhibitor and their docking pose. In fact, the negative group sug-

gested as a pharmacophoric feature to inhibit type B Rpi due to the highly electropositive nature of the Rpi cavity is present in all selected compounds making possible some of these key interactions [48].

### 3.7. Molecular dynamics simulations

Molecular dynamics simulations of 50 ns were performed to analyze the stability of the docked protein–ligand complexes (Fig. 5–7). The protein in the complex was maintained in the protonation states considered in molecular docking with the His102 and the Cys69 set to protonated and neutral, respectively. From these molecular dynamic simulations, RMSDs of the C $\alpha$  atoms of the proteins in the complexes were calculated as shown in Fig. 7A. The TcRpi-positive control complex was the first to reach the equilibrium at 20 ns with RMSD mean of 0.22 nm. The TcRpi-ZINC63480117 and TcRpi-ZINC43763931 complexes reached the plateau after 30 ns with RMSD means of 0.24 nm and 0.25 nm, respectively. The last one to attain equilibrium was TcRpi-ZINC36975961 complex with RMSD of 0.27 nm. The regions of higher flexibility on the proteins were evaluated through RMSF of the C $\alpha$  atoms as shown in Fig. 7B. Fluctuations above 0.25 nm were observed for Met1A, Ala15A, Ser71B, Thr40A/Thr40B-Ser43A/Ser43B, Glu61A/Glu161B, Ser151A/Ser151B and His152A/His152B on the complexes, but none of them are residues of the active site with the potential inhibitors. The binding stability of these inhibitors was ensured through their RMSD values (Fig. 7C). All inhibitors deviated from their starting structures reaching stable conformations after 30 ns, with the exception of ZINC63480117 that showed continued fluctuations. The 4PEH, ZINC36975961, and ZINC43763931 stabilized with RMSD means of 0.23 ns, 0.58 ns and 0.4 ns, respectively. It is important to notice that all the inhibitors remained in the binding pocket throughout the simulation.

The interactions between the ligands and TcRpi were further evaluated through analysis of the hydrogen bonds during the trajectories (Fig. 7D). The positive control 4PEH presented an average of 7 hydrogen bonds interactions with the TcRpi active site. In relation to the docked complex, this ligand maintained hydrogen bonds with Asp10 (98.8% of frequency), Gly74 (25%), His138 (84.83%) and Arg141 (30%). Besides that, new hydrogen bonds were formed with Tyr46, Ser71 and Asn103 with occupancy higher than 70%. The ZINC36975961 potential inhibitor had an average of 4 hydrogen bonds throughout the simulation and maintained hydrogen bonds with Asn103 (37.92% occupancy), His138 (48.5%) and Arg141 (7%) from those observed in the docking complex. During the simulation new interactions were formed with Ser71 (52.5% of occupancy), Arg113 (32.34%), His102 (49.7%), Tyr94 (37.52%) and Tyr46 (35.33%). The ZINC63480117, during the dynamics, formed around two hydrogen bonds with occupancies higher than 20% at Asn103, His 11, Ser43, Arg141, and His102. At the end of the simulation, the ligand lost all hydrogen bonds and was maintained in the pocket through electrostatic interaction of the ligand carboxylate group with Arg141 (Fig. 6F). An average of 6 hydrogen bonds was formed between ZINC43763931 and TcRpi during the simulation. From those interactions observed in the docking pose, this ligand conserved hydrogen bonds with Asp10 (97.8% of occupancy), main (40.12%) and side chain (53.69%) of Ser71 and Gly74 (48.7%), however low occupancy was obtained at Arg141 (4.19%) and His138 (8.58%). New hydrogen bonds were formed with considerable occupancy at Asn103 (40.92%), His102 (28.54%), and Tyr94 (31.34%). The interactions observed for each inhibitor, 4PEH, ZINC36975961, ZINC63480117, and ZINC43763931, at the end of the 50 ns of simulation is shown in Fig. 5C and Fig. 6C, F, and I, respectively. The comparison between the pose of the most populated cluster obtained for each molecular dynamics simulations and the respec-

tive molecular docking pose are represented in the supplementary material Fig. S6.

These results indicate that, as expected for a positive control, 4PEH had stable molecular interactions with the TcRpi active site over all the 50 ns and was considerably more stable than the potential inhibitors. The molecular dynamics simulations also revealed that ZINC36975961 and ZINC43763931 made important interactions along the simulation with TcRpi residues involved in catalysis and substrate binding as previously observed in docking studies. Besides that, some of these observed interactions are also found in the 4PEH known inhibitor over the simulation time, in the docking, and the crystallographic pose. This pattern of interactions and the docking scores obtained in the docking studies suggest that ZINC36975961 and ZINC43763931 compounds are promising TcRpi inhibitor candidates. The ZINC63480117 had noticeable fluctuations throughout the 50 ns and lost some important interactions at the end of the simulation time. Nevertheless, the docked pose predicted for this compound presented remarkable interactions and remained attached to the binding pocket during the dynamic simulation indicating that this pose could guide the proposal of new chemical structures for further drug design studies.

## 4. Conclusion

The drugs currently available for the Chagas disease are inadequate for their inefficacy in the treatment of the chronic phase, for their toxicity and the resistance developed by some strains of the parasite [5,49,50]. For this reason, the search for more effective drugs against this pathogen is critical and the Rpi appears as a potential drug target mainly due to its essential role in PPP and, consequently, in its function against parasite oxidative stress [11,12]. The essentiality of the Rpi was already shown for some organisms like *S. cerevisiae*, *E. coli*, and *L. infantum* [18–21]. In *Trypanosoma brucei*, even though the essentiality was not completely proven, all the efforts to create null Rpi mutants failed, and silencing of Rpi led to the reduction of the parasite growth and less parasitemia of infected mice [21].

In the present study for the identification of new potential TcRpi inhibitors, firstly, it was performed a ligand-based and structure-based pharmacophore modeling considering known RpiB inhibitors and available TcRpi crystallographic structures. A substructure and 2D similarity search were also carried out. The aim was to screen a large virtual database of purchasable compounds. Subsequently, the retrieved compounds were analyzed through molecular docking against the active site of TcRpi and HsRpi and through ADMETOX/PAINS predictions. Out of a total of 211 compounds positively selected as good TcRpi potential inhibitors, the top three hit compounds – ZINC36975961, ZINC63480117, and ZINC43763931 – had further analyses of their binding modes in TcRpi's active site and were found to form interactions with key residues involved in catalysis and substrate binding. When compared to the known TcRpi inhibitors, these compounds had significantly higher docking scores to the human enzyme (*i.e.*, they were among the worst scored compounds) indicating that they are best choices as leads to more selective inhibitors for the *T. cruzi* enzyme. Besides that, molecular dynamics simulations revealed that ZINC36975961 and ZINC43763931 maintained with the TcRpi's active site some key interactions observed in the docking studies and also formed new ones. The ZINC63480117, even though remained in the binding pocket, showed instability throughout the simulation and lost some important interactions. These results indicate that the identified compounds ZINC36975961 and ZINC43763931 are potential inhibitors for TcRpi and may be used as leads for the drug design of novel TcRpi inhibitors.



## Funding

This work was supported by Plataforma Tecnológica Fiocruz – Bioinformática RPT04A with computational support; and the Brazilian National Council of Research (CNPq) Contract grants: MCTI/CNPq N° 14/2013 – Universal (480866/2013-9) and PAPES VII/CNPq (401808/2015-7)

## Acknowledgments

The authors are grateful to Plataforma Tecnológica Fiocruz – Bioinformática RPT04A for providing us the computational support. We thank Brazilian National Council of Research (CNPq) for supporting this work.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jmngm.2017.08.007>.

## References

- [1] M. Traina, S. Meymandi, J.S. Bradfield, Heart failure secondary to chagas disease: an emerging problem in non-endemic areas, *Curr. Heart Fail. Rep.* (2016) 1–7, <http://dx.doi.org/10.1007/s11897-016-0305-9>.
- [2] WHO, Chagas Disease (American Trypanosomiasis), WHO, 2016, <http://www.who.int/mediacentre/factsheets/fs340/en/>. Published 2016. (Accessed 30 December 2016).
- [3] C.M. Beaumier, P.M. Gillespie, U. Strych, T. Hayward, P.J. Hotez, M.E. Bottazzi, Status of vaccine research and development of vaccines for Chagas disease, *Vaccine* 34 (26) (2016) 2996–3000, <http://dx.doi.org/10.1016/j.vaccine.2016.03.074>.
- [4] P.V.M. Pereira, E.C. Navarro, Challenges and perspectives of Chagas disease: a review, *J. Venom. Anim. Toxins Incl. Trop. Dis.* 19 (1) (2013) 34, <http://dx.doi.org/10.1186/1678-9199-19-34>.
- [5] J. Bermudez, C. Davies, A. Simonazzi, J. Pablo Real, S. Palma, Current drug therapy and pharmaceutical challenges for Chagas disease, *Acta Trop.* 156 (2016) 1–16, <http://dx.doi.org/10.1016/j.actatropica.2015.12.017>.
- [6] W.M. Fitch, Distinguishing homologous from analogous proteins, *Syst. Biol.* 19 (2) (1970) 99–113, <http://dx.doi.org/10.2307/2412448>.
- [7] M.Y. Galperin, D.R. Walker, E.V. Koonin, Analogous enzymes: independent inventions in enzyme evolution, *Genome Res.* 8 (8) (1998) 779–790, <http://dx.doi.org/10.1101/gr.8.8.779>.
- [8] Capriles PVSZ, A.C.R. Guimarães, T.D. Otto, A.B. Miranda, L.E. Dardenne, W.M. Degraive, Structural modelling and comparative analysis of homologous, analogous and specific proteins from *Trypanosoma cruzi* versus *Homo sapiens*: putative drug targets for chagas' disease treatment, *BMC Genom.* 11 (2010) 610, <http://dx.doi.org/10.1186/1471-2164-11-610>.
- [9] T.D. Otto, A.C.R. Guimarães, W.M. Degraive, A.B. de Miranda, AnEnPi: identification and annotation of analogous enzymes, *BMC Bioinf.* 9 (2008) 544, <http://dx.doi.org/10.1186/1471-2105-9-544>.
- [10] M. Alves-Ferreira, A.C.R. Guimarães, Capriles PVSZDSZ, L.E. Dardenne, W.M. Degraive, A new approach for potential drug target discovery through in silico metabolic pathway analysis using *Trypanosoma cruzi* genome information, *Mem. Inst. Oswaldo Cruz* 104 (8) (2009) 1100–1110, <http://dx.doi.org/10.1590/S0074-02762009000800006>.
- [11] A.L. Stern, E. Burgos, L. Salmon, J.J. Cazzulo, Ribose 5-phosphate isomerase type B from *Trypanosoma cruzi*: kinetic properties and site-directed mutagenesis reveal information about the reaction mechanism, *Biochem. J.* 401 (1) (2007) 279–285, <http://dx.doi.org/10.1042/BJ20061049>.
- [12] A.L. Stern, A. Naworyta, J.J. Cazzulo, S.L. Mowbray, Structures of type B ribose 5-phosphate isomerase from *Trypanosoma cruzi* shed light on the determinants of sugar specificity in the structural family, *FEBS J.* 278 (5) (2011) 793–808, <http://dx.doi.org/10.1111/j.1742-4658.2010.07999.x>.
- [13] M.P. Barrett, The pentose phosphate pathway and parasitic protozoa, *Parasitol. Today Pers. Ed.* 13 (1) (1997) 11–16, [http://dx.doi.org/10.1016/S0169-4758\(96\)10075-2](http://dx.doi.org/10.1016/S0169-4758(96)10075-2).
- [14] J. Kovářová, M.P. Barrett, The pentose phosphate pathway in parasitic trypanosomatids, *Trends Parasitol.* 32 (8) (2016) 622–634, <http://dx.doi.org/10.1016/j.pt.2016.04.010>.
- [15] D.A. Maugeri, J.J. Cazzulo, R.J.S. Burchmore, M.P. Barrett, P.O.J. Ogbunude, Pentose phosphate metabolism in *Leishmania mexicana*, *Mol. Biochem. Parasitol.* 130 (2) (2003) 117–125, [http://dx.doi.org/10.1016/S0166-6851\(03\)00173-7](http://dx.doi.org/10.1016/S0166-6851(03)00173-7).
- [16] D.A. Maugeri, J.J. Cazzulo, The pentose phosphate pathway in *Trypanosoma cruzi*, *FEMS Microbiol. Lett.* 234 (1) (2004) 117–123, <http://dx.doi.org/10.1016/j.femsle.2004.03.018>.
- [17] M. Igoillo-Esteve, D. Maugeri, L. Stern, A. P. Beluardi, J.J. Cazzulo, The pentose phosphate pathway in *Trypanosoma cruzi*: a potential target for the chemotherapy of Chagas disease, *An. Acad. Bras. Cienc.* 79 (4) (2007) 649–663, S0001-37652007000400007 [pii].
- [18] A.J. Skinner, R.A. Cooper, The regulation of ribose-5-phosphate isomerisation in *Escherichia coli* K12, *FEBS Lett.* 12 (5) (1971) 293–296, [http://dx.doi.org/10.1016/0014-5793\(71\)80202-8](http://dx.doi.org/10.1016/0014-5793(71)80202-8).
- [19] K.I. SØRENSEN, B. Hove-Jensen, Ribose catabolism of *Escherichia coli*: characterization of the rpiB gene encoding ribose phosphate isomerase B and of the rpiR gene, which is involved in regulation of rpiB expression, *J. Bacteriol.* 178 (4) (1996) 1003–1011, <http://dx.doi.org/10.1128/jb.178.4.1003-1011.1996>.
- [20] T. Miosga, F.K. Zimmermann, Cloning and characterization of the first two genes of the non-oxidative part of the *Saccharomyces cerevisiae* pentose-phosphate pathway, *Curr. Genet.* 30 (5) (1996) 404–409, <http://dx.doi.org/10.1007/s002940050149>.
- [21] I. Loureiro, J. Faria, C. Clayton, et al., Ribose 5-Phosphate isomerase B knockdown compromises *trypanosoma brucei* bloodstream form infectivity, *PLoS Negl. Trop. Dis.* 9 (1) (2015) 1–11, <http://dx.doi.org/10.1371/journal.pntd.0003430>.
- [22] J. Faria, I.I. Loureiro, N. Santarem, et al., Disclosing the essentiality of ribose-5-phosphate isomerase B in *Trypanosomatids*, *Sci. Rep.* 6 (January) (2016) 26937, <http://dx.doi.org/10.1038/srep26937>.
- [23] E. Burgos, L. Salmon, Synthesis and evaluation of new 4-phospho-D-erythronic acid derivatives as competitive inhibitors of spinach ribose-5-phosphate isomerase, *Tetrahedron Lett.* 45 (4) (2004) 753–756, <http://dx.doi.org/10.1016/j.tetlet.2003.11.038>.
- [24] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (3) (1993) 779–815, <http://dx.doi.org/10.1006/jmbi.1993.1626>.
- [25] Capriles PVSZ, L.P.R. Baptista, I.A. Guedes, et al., Structural modeling and docking studies of ribose 5-phosphate isomerase from *Leishmania major* and *Homo sapiens*: a comparative analysis for Leishmaniasis treatment, *J. Mol. Graph. Model.* 55 (2015) 134–147, <http://dx.doi.org/10.1016/j.jmngm.2014.11.002>.
- [26] Y. Zhang, J. Skolnick, TM-align: a protein structure alignment algorithm based on the TM-score, *Nucleic Acids Res.* 33 (7) (2005) 2302–2309, <http://dx.doi.org/10.1093/nar/gki524>.
- [27] S. Mariano, A.K. Roos, S.L. Mowbray, L. Salmon, Competitive inhibitors of type B ribose 5-phosphate isomerases: design, synthesis and kinetic evaluation of new d-allose and d-allulose 6-phosphate derivatives, *Carbohydr. Res.* 344 (7) (2009) 869–880, <http://dx.doi.org/10.1016/j.carres.2009.02.024>.
- [28] A.K. Roos, E. Burgos, D.J. Ericsson, L. Salmon, S.L. Mowbray, Competitive inhibitors of *Mycobacterium tuberculosis* ribose-5-phosphate isomerase B reveal new information about the reaction mechanism, *J. Biol. Chem.* 280 (8) (2005) 6416–6422, <http://dx.doi.org/10.1074/jbc.M412018200>.
- [29] R.G. Zhang, C.E. Andersson, T. Skarina, et al., The 2.2 Å resolution structure of RpiB/AlsB from *Escherichia coli* illustrates a new approach to the ribose-5-phosphate isomerase reaction, *J. Mol. Biol.* 332 (5) (2003) 1083–1094, <http://dx.doi.org/10.1016/j.jmb.2003.08.009>.
- [30] J.J. Irwin, B.K. Shoichet, ZINC—a free database of commercially available compounds for virtual screening, *J. Chem. Inf. Model.* 45 (1) (2005) 177–182, <http://dx.doi.org/10.1021/ci049714+>.
- [31] M.M. Mysinger, M. Carchia, J.J. Irwin, B.K. Shoichet, Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking, *J. Med. Chem.* 55 (14) (2012) 6582–6594, <http://dx.doi.org/10.1021/jm300687e>.
- [32] D. Schneidman-Duhovny, O. Dror, Y. Inbar, R. Nussinov, H.J. Wolfson, PharmaGist: a webserver for ligand-based pharmacophore detection, *Nucleic Acids Res.* 36 (2008) 223–228, <http://dx.doi.org/10.1093/nar/gkn187> (Web Server issue).
- [33] D.R. Koes, C.J. Camacho, ZINCPharmer: pharmacophore search of the ZINC database, *Nucleic Acids Res.* 40 (W1) (2012) W409–W414, <http://dx.doi.org/10.1093/nar/gks378>.
- [34] D.R. Koes, C.J. Camacho, Pharmer Efficient and exact pharmacophore search, *J. Chem. Inf. Model.* 51 (2011) 1307–1314, <http://dx.doi.org/10.1021/ci200097m>.
- [35] M.R. Berthold, N. Cebon, F. Dill, et al., KNIME – the Konstanz information miner, *SIGKDD Explor.* 11 (1) (2009) 26–31, [http://dx.doi.org/10.1007/978-3-540-78246-9\\_38](http://dx.doi.org/10.1007/978-3-540-78246-9_38).
- [36] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.* 46 (1) (2001) 3–26, [http://dx.doi.org/10.1016/S0169-409X\(00\)00129-0](http://dx.doi.org/10.1016/S0169-409X(00)00129-0).
- [37] R.A. Friesner, J.L. Banks, R.B. Murphy, et al., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J. Med. Chem.* 47 (7) (2004) 1739–1749, <http://dx.doi.org/10.1021/jm0306430>.
- [38] J. Wang, W. Yang, Concerted proton transfer mechanism of clostridium thermocellum ribose-5-phosphate isomerase, *J. Phys. Chem. B* 117 (32) (2013) 9354–9361, <http://dx.doi.org/10.1021/jp404948c>.
- [39] R. A. Friesner, R.B. Murphy, M.P. Repasky, et al., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for

- protein-ligand complexes, *J. Med. Chem.* 49 (21) (2006) 6177–6196, <http://dx.doi.org/10.1021/jm051256o>.
- [40] D. Lagorce, O. Sperandio, J.B. Baell, M.A. Miteva, B.O. Villoutreix, FAF-Drugs3: a web server for compound property calculation and chemical library design, *Nucleic Acids Res.* 43 (W1) (2015) W200–W207, <http://dx.doi.org/10.1093/nar/gkv353>.
- [41] K.B. Koziara, M. Stroet, A.K. Malde, A.E. Mark, Testing and validation of the Automated Topology Builder (ATB) version 2.0: prediction of hydration free enthalpies, *J. Comput. Aided Mol. Des.* 28 (3) (2014) 221–233, <http://dx.doi.org/10.1007/s10822-014-9713-7>.
- [42] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible, and free, *J. Comput. Chem.* 26 (16) (2005) 1701–1718, <http://dx.doi.org/10.1002/jcc.20291>.
- [43] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems, *J. Chem. Phys.* 98 (12) (1993) 10089, <http://dx.doi.org/10.1063/1.464397>.
- [44] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph.* 14 (1) (1996) 33–38, [http://dx.doi.org/10.1016/0263-7855\(96\)00018-5](http://dx.doi.org/10.1016/0263-7855(96)00018-5).
- [45] E. Cauët, M. Rومان, R. Wintjens, J. Liévin, C. Biot, Histidine-aromatic interactions in proteins and protein-ligand complexes: quantum chemical study of X-ray and model structures, *J. Chem. Theory Comput.* 1 (3) (2005) 472–483, <http://dx.doi.org/10.1021/ct049875k>.
- [46] C.D.M. Churchill, S.D. Wetmore, Noncovalent interactions involving histidine: the effect of charge on  $\pi$ - $\pi$  stacking and T-shaped interactions with the DNA nucleobases, *J. Phys. Chem. B* 113 (49) (2009) 16046–16058, <http://dx.doi.org/10.1021/jp907887y>.
- [47] D.A. Dougherty, Cation- $\pi$  interactions involving aromatic amino acids, *J. Nutr.* 137 (6) (2007) 1504–1508, 0022-3166/07.
- [48] P.K. Kaur, N. Tripathi, J. Desale, et al., Mutational and structural analysis of conserved residues in ribose-5-phosphate isomerase B from *Leishmania donovani*: role in substrate recognition and conformational stability, *PLoS One* 11 (3) (2016) e0150764, <http://dx.doi.org/10.1371/journal.pone.0150764>.
- [49] L.S. Filardi, Z. Brener, Susceptibility and natural resistance of *Trypanosoma cruzi* strains to drugs used clinically in Chagas disease, *Trans. R. Soc. Trop. Med. Hyg.* 81 (5) (1987) 755–759, [http://dx.doi.org/10.1016/0035-9203\(87\)90020-4](http://dx.doi.org/10.1016/0035-9203(87)90020-4).
- [50] J. Rodrigues Coura, S.L. de Castro, A critical review on Chagas disease chemotherapy, *Mem. Inst. Oswaldo Cruz* 97 (1) (2002) 3–24, <http://dx.doi.org/10.1590/S0074-02762002000100001>.