UNIVERSIDADE DE SÃO PAULO INSTITUTO DE FÍSICA DE SÃO CARLOS

MARIANA LAUREANO DE SOUZA

Identificação de novos inibidores da enzima cruzaína de *Trypanosoma cruzi* candidatos a fármacos contra a doença de Chagas

> São Carlos 2012

MARIANA LAUREANO DE SOUZA

Identificação de novos inibidores da enzima cruzaína de *Trypanosoma* cruzi candidatos a fármacos contra a doença de Chagas

> Dissertação apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Mestre em Ciências.

> Área de concentração: Física Aplicada Opção: Física Biomolecular Orientador: Prof. Dr. Adriano Defini Andricopulo

Versão Corrigida

(Versão original disponível na Unidade que aloja o Programa)

São Carlos 2012 ficha catalográfica

AGRADECIMENTOS

A Deus por iluminar meu caminho.

Aos meus pais, Sergio e Cidinha, por tudo que tenho hoje. Aos meus irmãos, Natália e Junior, com quem aprendi a compartilhar tudo em minha vida. A toda a minha família, que sempre esteve presente.

Ao Prof. Dr. Adriano D. Andricopulo, pela confiança, oportunidade, orientação, incentivo, amizade e acima de tudo pelo exemplo de compromisso e dedicação.

À Prof. Dra. Rafaela S. Ferreira, pelo total apoio, pela confiança, por toda ajuda e principalmente pela amizade.

Ao Prof. Dr. Rafael V. C. Guido, pela ajuda constante durante o desenvolvimento deste trabalho.

Às minhas amigas da sala 9, Wanessa, Bela, Renata, Simone e Ivani, que acompanharam o meu desenvolvimento e sempre estiveram prontas para ajudar. Também, pelos momentos inesquecíveis.

Aos amigos que sempre ajudaram com importantes discussões, em particular, Prof. Dr. Gustavo Trossini, Heline, Leonardo, Ricardo, Lívia e Napoleão.

Às minhas amigas queridas que estiveram sempre ao meu lado, mesmo quando não entendiam nada que eu dizia, Mari, Fer, Naná, Si e Régia.

A todos os professores, funcionários e colegas do laboratório de cristalografia do Instituto de Física de São Carlos.

Ao CNPq pelo auxílio financeiro durante este trabalho.

Por fim, a todos que direta ou indiretamente participaram de minha formação, meus sinceros agradecimentos.

RESUMO

DE SOUZA, M. L. **Identificação de novos inibidores da enzima cruzaína de** *Trypanosoma cruzi* **candidatos a fármacos contra a doença de Chagas. 2012. 84p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.**

A doença de Chagas, uma infecção parasitária amplamente distribuída na América Latina, é um problema grave de saúde pública com consequências devastadoras em termos de morbidade e mortalidade humana. O arsenal terapêutico contra a doença é bastante limitado e insuficiente em todos os aspectos clínicos. Visando o desenvolvimento de novos agentes antichagásicos, várias proteínas do parasita têm sido exploradas como alvos terapêuticos. Neste contexto, a enzima cruzaína, uma cisteíno protease envolvida nos estágios de desenvolvimento e diferenciação do Trypanosoma cruzi, foi selecionada para os nossos estudos, visando a identificação de inibidores através do uso do método de planejamento de fármacos baseado na estrutura do receptor (SBDD, do inglês, structure-based drug design). Esta metodologia engloba uma diversidade de estratégias, empregando estruturas cristalográficas de proteínas alvo, disponíveis usualmente no Protein Data Bank (PDB). Entre as técnicas modernas utilizadas no SBDD, destaca-se a triagem virtual baseada na estrutura do receptor (SBVS, do inglês, structure-based virtual screening), que possibilita a seleção de novos candidatos a ligantes de proteínas alvo, a partir de grandes bases de dados de compostos. No presente trabalho de dissertação, a seleção de 19 estruturas da enzima cruzaína, em complexo com ligantes, permitiu a aplicação de métodos de SBDD. Um conjunto com cerca de 3,4 milhões de compostos, com característica líder-similar (do inglês, lead-like), e outro conjunto com aproximadamente 450.000 compostos, com característica fragmento-similar (do inglês, fragment-like), foram coletados da base de dados ZINC. O programa DOCK 3.5.54 foi empregado na triagem virtual das bases de dados utilizando-se a estrutura cristalográfica PDB ID: 3KKU. Um subconjunto com 35.000 moléculas foi selecionado para estudos posteriores com os programas GOLD e Surflex. As 500 melhores moléculas selecionadas por cada um dos programas foram analisadas visualmente considerando-se diversas características estruturais dos subsítios da enzima cruzaína e dos ligantes (e.g., complementaridade molecular, flexibilidade, lipofilia do subsítio S2, presença de doadores e aceptores de hidrogênio entre os subsítios S2 e S1). Desta forma, um conjunto

final de 18 compostos foi priorizado para os ensaios bioquímicos frente a enzima cruzaína. Destes 18 compostos, 6 apresentaram atividade inibitória frente a cruzaína, com destaque para os 2 mais promissores, com valores de IC₅₀ (concentração de inibidor necessária para reduzir em 50% a atividade enzimática) de 20 μ M e 580 nM. O inibidor mais potente da série foi selecionado da base fragmento-similar e apresentou um valor de eficiência do ligante (EL) de 0,53 kcal/mol/átomo, considerado significativo para otimização em química medicinal. A integração de técnicas computacionais e experimentais permitiu a descoberta de ligantes com inovação estrutural, abrindo novas perspectivas para o planejamento de inibidores mais potentes e seletivos da enzima cruzaína de *T. cruzi*.

Palavras-chave: *Trypanosoma cruzi*. Cruzaína. Triagem virtual. Ensaio enzimático. Química medicinal.

ABSTRACT

DE SOUZA, M. L. **Discovery of novel inhibitors of the cruzain enzyme from** *Trypanosoma cruzi* as drug candidates against Chagas' disease. 2012. 84p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2012.

Chagas' disease, a parasitic infection widely distributed in Latin America, is a serious public health problem with devastating consequences in terms of human morbidity and mortality. The therapeutic arsenal against the disease is very limited and insufficient in all clinical aspects. This has led to a new paradigm for the discovery of new agents that act on specific enzymes or metabolic pathways. The enzyme cruzain, a cysteine protease essential for the survival of the parasite Trypanosoma cruzi, has been selected in this work as an attractive target for the development of new inhibitors through the use of structure-based drug design (SBDD). This approach brings together a diversity of strategies, which employs crystal structures of target proteins, usually available in the Protein Data Bank (PDB). Structurebased virtual screening (SBVS), one of the most important techniques used in SBDD, allows the selection of new ligands of target proteins from large libraries of compounds. In this work, 19 crystal structures of the cruzain enzyme, in complex with ligands, allowed the application of SBDD methods. A data set of about 3.4 million compounds, with *lead-like* characteristics, and a second data set, with approximately 450,000 compounds, with fragment-like characteristics, were collected from the ZINC data base. The docking program DOCK 3.5.54 was employed in the virtual screening of the data sets using the crystal structure PDB ID: 3KKU. A subset of 35,000 compounds was selected for further studies with the programs GOLD and Surflex. The 500 top ranked molecules for each of the programs were visually inspected considering a number of structural characteristics of the subsites of the cruzain enzyme, as well as of the ligands (e.g., molecular complementarity, flexibility, the hydrophobic S2 subsite, and the presence of hydrogen donors and acceptors between the subsites S2 and S1). Thus, a final subset of 18 compounds was prioritized for the biochemical assays against the cruzain enzyme. Six out of 18 compounds exhibited enzyme inhibition, with the most two promising inhibitors having IC_{50} values (IC_{50} refers to the concentration of compound required for 50% inhibition of cruzain) of 20 µM e 580 nM. The most potent inhibitor of the series was selected from the *fragment-like* data set and showed a ligand

efficiency of 0,53 kcal/mol/atom, which is considered significant in drug design. The integration of computational and experimental approaches allowed the discovery of compounds with innovative structures, providing new perspectives for the design of inhibitors of *T. cruzi* cruzain having increased potency and selectivity.

Keywords: Trypanosoma cruzi. Cruzain. Virtual screening. Enzymatic assay. Medicinal chemistry.

LISTA DE FIGURAS

Figura 1 – Distribuição mundial de casos de infecção por <i>Trypanosoma cruzi</i> e do vetor de transmissão, entre os anos de 2006 e 2009 (Adaptado de http://whqlibdoc.who.int/publications/2010/9789241564090_eng.pdf) 24
Figura 2 – Esquema geral do ciclo de vida do <i>Trypanosoma cruzi</i> . (Adaptado de http://www.dpd.cdc.gov/dpdx/HTML/ImageLibrary/TrypanosomiasisAmeri can_il.htm) 25
Figura 3 - Fármacos utilizados no tratamento da doença de Chagas 26
Figura 4 - Modelo de interação enzima-substrato para proteases. Nesse modelo os subsítios da enzima são representados pela letra "S", e os peptídeos que interagem com a enzima, pela letra "P". A clivagem do substrato está representada por uma seta12 28
Figura 5 - Estrutura da cruzaína formada por dois domínios, onde um deles é predominantemente hélice-α e o outro consiste de extensas interações de folhas-β antiparalelas 29
Figura 6 - Subsítios da cruzaína em destaque: S1' (verde), S1 (rosa), S2 (roxo) e S3 (amarelo) 30
 Figura 7 – Ao centro, a estrutura cristalográfica da cruzaína de <i>T. cruzi</i> (PDB ID: 3KKU). Ao redor, um conjunto de inibidores da enzima alvo de diferentes classes químicas: (1) vinilsulfonas, (2) fluorometilcetonas, (3) triazóis, (4) pirimidinas, (5) tiossemicarbazonas, (6) chalconas e (7) benzimidazóis 31
Figura 8 - Interações entre a cruzaína e os inibidores: A) Z-Tyr-Ala-FMK (PDB ID: 1AIM)

e B) Z-Arg-Ala-FMK (PDB ID: 2AIM). Os diagramas de interação 2D foram gerados com o auxílio do programa Poseview (http://poseview.zbh.unihamburg.de)26, desenvolvido pelo Centro de Bioinformática da Universidade de Hamburgo, Alemanha. As linhas pretas tracejadas representam ligações de hidrogênio e as linhas sólidas verdes representam interações hidrofóbicas27. As figuras de interação 3D foram geradas pelo programa Pymol e os subsítios estão sombreados conforme figura anterior. ----- 32

Figura 9 -	Interações entre a cruzaína e um inibidor reversível hidroximetilcetona (PDB ID: 1ME3) 32	3
Figura 10 -	Interações entre a cruzaína e um inibidor α-ceto éster (PDB ID: 1U9Q) 3-	4
Figura 11 -	Interações intermoleculares entre a cruzaína e o inibidor K777 3	5
Figura 12 -	Esquema geral da aplicação de métodos de SBVS 30	б

- Figura 13 Esquema geral de docagem molecular de uma base de dados. As moléculas melhores classificadas podem ser selecionadas para avaliação biológica. ------ 37
- Figura 14 Estrutura da enzima cruzaína (PDB ID: 3KKU). A) Complexo proteína-ligante, onde a proteína está representada em modelo de fitas e o ligante em bastão. B) Representação do sítio de ligação da enzima (superfície) e as interações com o inibidor (bastão). As interações de hidrogênio estão representadas pelas linhas tracejadas. ------46
- Figura 15 Esquema geral dos processos de SBVS da Base 1. Inicialmente, o programa DOCK foi utilizado na docagem dos compostos. Um subconjunto foi selecionado para os estudos com os programas GOLD e Surflex. Os 500 melhores compostos de cada programa foram inspecionados visualmente para priorização de compostos para avaliação experimental em ensaios em in vitro. - 47
- Figura 16 Principais resíduos do sítio ativo da cruzaína que realizam interações com os compostos selecionados. ----- 48
- Figura 17 Esquema geral dos processos de SBVS da Base 2. Inicialmente, o programa DOCK foi utilizado na docagem dos compostos. Um subconjunto foi selecionado para os estudos com os programas GOLD e Surflex. Os 500 melhores compostos de cada programa foram inspecionados visualmente para priorização de compostos para avaliação experimental em ensaios em in vitro. - 49
- Figura 18 Análise por SDS-PAGE da expressão heteróloga e da purificação da cruzaína. Gel SDS-PAGE 15% corado com Coomassie-blue. Sendo M = padrão de massa molecular; 1 = fração pós coluna (flowthrough); 2-4 = frações pós purificação em coluna de níquel; 5 = enzima cruzaína após completada a ativação completa (alíquota em destaque em vermelho). -----51

- Figura 19 Classe de inibidores vilnilsulfona. Diagramas 2D gerados pelo Poseview para as estuturas: A) PDB ID: 1F29, B) PDB ID: 1F2B. C) Interações entre o inibidor A e a cruzaína (PDB ID: 1F29). D) Ocupação do sítio ativo da cruzaína pelas vinilsulfonas A e B.
- Figura 20 Interações entre a cruzaína e um inibidor purínico (PDB ID: 3IO6). ----- 57
- Figura 21 Interações entre a cruzaína e um inibidor não peptídico (PDB ID: 3IUT).----- 58
- Figura 22 Interações entre a cruzaína e o inibidor derivado de vinilsulfona (PDB ID: 3LXS).-----58
- Figura 23 Interações entre a cruzaína e o inibidor reversível benzimidazol (PDB ID: 3KKU)------59
- Figura 24 Estratégia de SBVS das Bases 1 e 2 com o programa DOCK para a identificação de candidatos a inibidores da enzima cruzaína.----- 60
- Figura 25 Estratégia de SBVS da Base 1 com os programas GOLD e Surflex para a identificação de candidatos a inibidores da enzima cruzaína.-----63
- Figura 26 Esquema da clivagem do substrato Z-Phe-Arg-AMC. ----- 69
- Figura 27 Gráfico concentração x resposta para a determinação de IC₅₀ do composto 16. 71
- Figura 28 Gráfico duplo-recíproco para o inibidor 16. Concentrações de inibidor: controle - zero (\bullet); 0,3 µM (O); 0,6 µM ($\mathbf{\nabla}$); 1,25 µM (\triangle) e 2,50 µM ($\mathbf{\Box}$).--- 72
- Figura 29 Determinação de K_i para o inibidor 16. Concentrações de substrato: 1,25 μ M (\bullet); 2,5 μ M (O); 20 μ M (\blacksquare) e 40 μ M (\Box). ------72

LISTA DE TABELAS

Tabela 1 - Alvos moleculares e rotas metabólicas e para o desenvolvimento de novos fármacos para o tratamento da doença de Chagas. 27
Tabela 2 – Numeração atual dos resíduos do sítio ativo da cruzaína 55
Tabela 3 – Compostos selecionados da Base 1 pelo programa DOCK para avaliação bioquímica frente a enzima cruzaína de T. cruzi.61
Tabela 4 – Compostos selecionados da Base 1 pelo programa GOLD para avaliação bioquímica frente a enzima cruzaína de T. cruzi.GOLD para avaliação63
Tabela 5 - Compostos selecionados da Base 1 pelo programa Surflex para avaliação bioquímica frente a enzima cruzaína de T. cruzi.65
Tabela 6 - Composto selecionado da Base 2 pelo programa DOCK para avaliação bioquímica frente a enzima cruzaína de <i>T. cruzi</i> .67
Tabela 7 – Compostos selecionados da Base 2 pelo programa GOLD para avaliação bioquímica frente a enzima cruzaína de T. cruzi.GOLD para avaliação67
Tabela 8 - Composto selecionado da Base 2 pelo programa Surflex para avaliação bioquímica frente a enzima cruzaína de T. cruzi.68
Tabela 9 – Avaliação bioquímica in vitro de uma série de compostos selecionados por SBVS, contra a enzima cruzaína70

LISTA DE ABREVIATURAS E SIGLAS

2D	Bidimensional
3D	Tridimensional
Å	Angstrom
AMC	7-Amino-4-metilcumarina
DMSO	Dimetilssulfóxido
D.O.	Densidade óptica
DTT	Ditiotreitol
E. coli	Escherichia coli
EDTA	Ácido etilenodiaminotetracético
EL	Eficiência do ligante
IC ₅₀	Concentração de inibidor para reduzir em 50% a atividade enzimática
IPTG	Isopropil-β-D-tiogalactopiranosídeo
Log P	Logaritmo do coeficiente de partição octanol-água
Ki	Constante de dissociação do complexo enzima-inibidor
$K_{ m M}$	Constante de Michaelis-Menten
OMS	Organização Mundial de Saúde
PDB	Protein Data Bank (Banco de dados de proteínas)
T. cruzi	Trypanosoma cruzi
Z	Benziloxicarbonil

SUMÁRIO

1	INT	INTRODUÇÃO				
	1.1	Doença de Chagas	23			
	1.2	Alvos Moleculares para a Doença de Chagas	26			
1.2.		1 Proteases	28			
	1	.2.1.1 Cruzaína	29			
	1.3	Planejamento de Fármacos Baseado na Estrutura do Receptor	35			
	1.3.	1 Docagem Molecular	36			
2	OB.	JETIVOS	41			
3	3 MATERIAIS E MÉTODOS					
	3.1	Estudos Estruturais	45			
	3.1.	1 Preparação do Alvo Molecular	45			
	3.2	Base de Dados	46			
	3.3	Triagem Virtual	47			
	3.4	Expressão e Purificação da Enzima Cruzaína	49			
	3.5	Ensaios Enzimáticos	52			
4	RES	SULTADOS E DISCUSSÃO	55			
	4.1	Estudos Estruturais	55			
	4.1	Triagem Virtual	59			
	4.2	Avaliação Bioquímica	69			
5	5 CONCLUSÕES77					
R	REFERÊNCIAS					



1 INTRODUÇÃO

1.1 Doença de Chagas

A doença de Chagas, também conhecida como tripanossomíase americana, é uma infecção parasitária causada pelo protozoário *Trypanosoma cruzi*. Estima-se atualmente que 10 milhões de pessoas estão infectadas, sendo a situação mais grave observada na América Latina, principalmente no Brasil, Argentina e México, onde a doença é endêmica¹. Mais recentemente, devido principalmente à intensa mobilidade da população mundial², a doença tem sido considerada também um problema relevante de saúde pública nos Estados Unidos, Canadá e em alguns países do leste europeu e da Ásia³. A Figura 1 apresenta o panorama global da distribuição geográfica e do número estimado de indivíduos infectados com o *T. cruzi*, de acordo com dados da Organização Mundial de Saúde (OMS), entre os anos de 2006-2009.



Figura 1 – Distribuição mundial de casos de infecção por *Trypanosoma cruzi* e do vetor de transmissão, entre os anos de 2006 e 2009 (Adaptado de http://whqlibdoc.who.int/publications/2010/9789241564090_eng.pdf).

A doença de Chagas é transmitida aos humanos predominantemente pelas fezes de triatomíneos, que são insetos hematófagos conhecidos como "barbeiros". Outras formas de contaminação menos comuns ocorrem em transfusões de sangue, transplante de órgãos, ingestão de alimentos infectados e transmissão congênita⁴. O *T. cruzi* apresenta um ciclo de vida complexo, que envolve um hospedeiro invertebrado e outro vertebrado, caracterizado por três formas distintas: (i) epimastigota, que se reproduz no intestino do inseto e não é infectante para os vertebrados; (ii) tripomastigota, fase extracelular, que circula no sangue, constituindo a forma infectante para os vertebrado; e (iii) amastigota, fase intracelular infectante, forma de replicação celular⁵.

De acordo com a Figura 2, o barbeiro ao se alimentar do sangue de um vertebrado libera em suas fezes tripomastigotas metacíclicos (1). Estas formas do parasita penetram no organismo do hospedeiro vertebrado através da ferida causada pela picada do inseto contaminado ou através do contato com as mucosas, como a conjuntiva. No hospedeiro, tripomastigotas invadem as células próximas ao local da inoculação, onde se diferenciam. Em sua fase intracelular, os tripomastigotas assumem uma forma ovoide e sem flagelo, chamada amastigota intracelular (2). As amastigotas se multiplicam rapidamente (3) causando o rompimento celular e a infestação do protozoário na corrente sanguínea. Neste estágio, o protozoário reassume a forma flagelada (tripomastigota sanguíneo), que se espalha pelo organismo infectando mais células e causando lesões, principalmente, em tecidos musculares cardíacos (4). As tripomastigotas sanguíneas se transformam em amastigotas intracelulares em locais de novas infecções. Manifestações clínicas podem ser originadas a partir deste ciclo infeccioso. A replicação é retomada quando os parasitas penetram em outras células ou são ingeridos por outro vetor. A perpetuação do ciclo se dá quando um inseto se alimenta de sangue humano ou animal que está contaminado com parasitas tripomastigotas (5). As tripomastigotas ingeridas pelo inseto se transformam em epimastigotas no intestino do vetor (6). As epimastigotas se multiplicam e se diferenciam no intestino delgado (7) e se diferenciam em tripomastigotas metacíclicos infectantes no intestino grosso (8), forma que é eliminada juntamente com as fezes do inseto quando este se alimenta, renovando, assim, o ciclo de transmissão.



Figura 2 – Esquema geral do ciclo de vida do *Trypanosoma cruzi*. (Adaptado de http://www.dpd.cdc.gov/dpdx/HTML/ImageLibrary/TrypanosomiasisAmerican_il.htm).

A doença de Chagas possui duas fases clínicas, uma aguda e uma crônica. A fase aguda inicia imediatamente após a infecção. Em adultos é normalmente oligossintomática, não sendo valorizada pelo paciente ou pelo agente de saúde. Os sintomas persistem por até dois meses, sendo possível o tratamento. Contudo, a doença é dificilmente identificada neste estágio, passando para a fase crônica que tipicamente se prolonga por toda a vida do hospedeiro. Por outro lado, a fase aguda sintomática pode ocorrer principalmente em crianças, na primeira década de vida, podendo levar à morte devido a complicações decorrentes de insuficiência cardíaca e processos inflamatórios⁶. Em 2009, celebrou-se o primeiro centenário da descoberta da doença de Chagas, realizada pelo médico brasileiro Carlos Chagas. Após 103 anos, não há qualquer medicamento eficaz disponível para o tratamento dessa grave enfermidade⁴.

Os fármacos disponíveis (Figura 3), benzonidazol (Rochagan[®], da Roche) e nifurtimox (Lampit[®], da Bayer), desenvolvidos na década de 1970, são extremamente

limitados e apresentam sérios problemas, como baixa eficácia e elevada toxicidade⁷. A necessidade de novas alternativas terapêuticas é clara, evidenciando a importância de investimentos em programas de pesquisa e desenvolvimento (P&D) de novos fármacos.





1.2 Alvos Moleculares para a Doença de Chagas

A busca por novas alternativas quimioterápicas pode envolver a seleção de um alvo molecular do parasita. Com o sequenciamento do genoma e os recentes avanços relacionados à proteômica e biologia do *T. cruzi*, foram identificados diversos alvos macromoleculares atrativos para o desenvolvimento de novos candidatos a fármacos (Tabela 1), sendo a maioria de enzimas^{4,6}.

Alvo molecular, via ou classe metabólica	Enzima, organela ou rota
(função)	metabólica
Durotogog	Císteino proteases
rioleases	Serino proteases
(multiplas funções que envolvem desde invasao celular, diferenciação,	Metaloproteases
desenvolvimento ate evasao do parasita do sistema imune)	Treonino proteases
	Esterol 14-demetilase
Biossíntese de esteróis	Lanoesterol sintase
(essencial para a composição estrutural de membranas, mitocôndrias e	3-Hidroxi-3-metilglutaril
plasma)	coenzima A redutase
	Farnesiltransferase
	Gliceraldeído-3-fosfato
Via glicolítica	desidrogenase
(produção de energia)	Hexoquinase
	Fosfofrutoquinase
Biossíntese de lipídeos	Aquil-lisofosfolipídeos
(essencial em diversos constituintes celulares)	Glicosfingolipídeos
Matabolismo dependente de grupos tióis	Tripanotiona redutase
(mecanismo de defesa contra estresse oxidativo)	Tripanotiona sintetase
	Triparedoxina peroxidase
Metabolismo de pentose fosfato	6-Fosfogluconato
(mecanismo de defesa contra estresse oxidativo)	desidrogenase
Super família de proteínas quinases	Arginina quinase
(produção de energia)	Fosfatidilinositol-3-quinase
Transporte e metabolismo de poliaminas	Arginina descarboxilase
(defesa contra estresse oxidativo)	Glutationil espermidina
	sintetase
	Purina fosforibosil transferases
Síntese de nucleotídeos	Di-hidrolato redutase
(precursores da síntese de matérias genéticos)	Pteridina redutase
	Di-hidro-orato desidrogenase
Transferência de ácido siálico	
(transferência do ácido a partir de glicoconjugados do hospedeiro e	Trans-sialidase
incorporação de moléculas de mucina ligadas à membrana parasitária)	
DNA topoisomerase	DNA topoisomerase I
(replicação do DNA do T. cruzi)	DNA topoisomerase II

 Tabela 1 - Alvos moleculares e rotas metabólicas e para o desenvolvimento de novos fármacos para o tratamento da doença de Chagas⁷.

1.2.1 Proteases

Proteases, peptidases ou enzimas proteolíticas possuem a função de catalisar a degradação de ligações peptídicas em proteínas e peptídeos. Em protozoários como o *T. cruzi*, as proteases possuem múltiplas funções que envolvem desde a invasão celular até a evasão do parasita do sistema imune do hospedeiro⁸. As proteases podem ser divididas de acordo com o tipo de reação que promovem, sendo classificadas como exopeptidases, quando removem resíduos de aminoácidos das extremidades N-terminal (aminopeptidases) ou C-terminal (carboxipeptidases), e como endopeptidases, quando são responsáveis por catalisar a hidrólise de ligações peptídicas internas em proteínas e peptídeos⁹. Outro critério utilizado para classificar proteases está relacionado com o mecanismo catalítico de clivagem, agrupando-as em seis classes principais: serino proteases, cisteíno proteases, aspartil proteases, metaloproteases, treonino proteases e glutamato proteases¹⁰.

As proteases possuem cavidades em suas estruturas denominadas subsítios. Um modelo foi desenvolvido para descrevê-los, onde os subsítios da enzima, representados pela letra "S", são os locais onde se ligam os resíduos de aminoácidos do substrato, representados pela letra "P". Esses subsítios são numerados de S1 a Sn seguindo a direção N-terminal do substrato e numerados de S1' a Sn' seguindo a direção C-terminal do substrato. Os resíduos do substrato acomodados nesses subsítios são numerados de P1 a Pn se direcionados ao N-terminal e de P1' a Pn' se direcionados ao C-terminal da ligação peptídica do substrato¹¹ (Figura 4).



Figura 4 - Modelo de interação enzima-substrato para proteases. Nesse modelo os subsítios da enzima são representados pela letra "S", e os peptídeos que interagem com a enzima, pela letra "P". A clivagem do substrato está representada por uma seta¹².

1.2.1.1 Cruzaína

A enzima cruzaína (EC 3.4.22.51), principal cisteíno protease do *T. cruzi*, é considerada um alvo molecular validado para a doença de Chagas¹³ e foi selecionada para os estudos nesta dissertação de mestrado. Expressa durante todo o ciclo de vida, a enzima é fundamental para a nutrição e desenvolvimento do parasita, evasão do sistema imune e invasão celular do hospedeiro¹⁴. A cruzaína, uma cisteíno protease da família da papaína, tem estrutura formada por dois domínios, onde um deles é predominantemente hélice- α e o outro consiste de extensas interações de folhas- β antiparalelas (Figura 5).



Figura 5 - Estrutura da cruzaína formada por dois domínios, onde um deles é predominantemente héliceα e o outro consiste de extensas interações de folhas-β antiparalelas.

O sítio ativo se encontra na interface entre os dois domínios, onde estão presentes os resíduos de aminoácidos Cys25, His159 e Asn175, que formam a chamada tríade catalítica¹⁵. Este sítio catalítico possui 4 subsítios de ligação, denominados S1, S2, S3 e S1' (Figura 6). O subsítio S2 (roxo) é o principal responsável pela especificidade da enzima. Pouco exposto ao

solvente, é delimitado pelas cadeias laterais dos resíduos hidrofóbicos Met68, Ala133, Leu157 e Gly160, o que confere especificidade por grupos hidrofóbicos. Por outro lado, a presença e a flexibilidade do resíduo de Glu205 na extremidade deste subsítio atribui uma característica negativa ao mesmo, o que possibilita também interações com grupos carregados positivamente¹⁶. Os subsítios S1'(verde), S1 (vermelho) e S3 (amarelo) são menos definidos e mais expostos ao solvente.



Figura 6 - Subsítios da cruzaína em destaque: S1' (verde), S1 (rosa), S2 (roxo) e S3 (amarelo).

A Figura 7 apresenta as principais classes de inibidores da enzima cruzaína⁷. Merecem destaque os derivados peptídicos, como as vinilsulfonas $(1)^{17, 18, 19}$ e as fluorometilcetonas $(2)^{16}$; e os derivados não peptídicos, como os triazóis $(3)^{20}$, as pirimidinas $(4)^{21}$, as tiossemicarbazonas $(5)^{22}$, as chalconas $(6)^{23}$ e os benzimidazóis $(7)^{24}$.



Figura 7 – Ao centro, a estrutura cristalográfica da cruzaína de *T. cruzi* (PDB ID: 3KKU). Ao redor, um conjunto de inibidores da enzima alvo de diferentes classes químicas: (1) vinilsulfonas, (2) fluorometilcetonas, (3) triazóis, (4) pirimidinas, (5) tiossemicarbazonas, (6) chalconas e (7) benzimidazóis.

A primeira estrutura cristalográfica da cruzaína foi determinada em complexo com um inibidor irreversível do tipo fluorometilcetona²⁵, servindo de guia para uma variedade de estudos da enzima. Outras estruturas de inibidores peptídicos em complexo com a cruzaína estão depositadas no PDB¹⁶.

Uma análise das estruturas destes complexos indica a especificidade do subsítio S2 por resíduos hidrofóbicos e básicos, que interagem principalmente com a Leu67. O resíduo Glu205 pode assumir duas conformações, de acordo com os grupos do inibidor presentes na posição P2. Como pode ser observada na Figura 8, a flexibilidade do Glu205 possibilita tanto a interação com a hidroxila do inibidor benziloxicarbonil-tirosina-alanina-fluorometilcetona (Z-Tyr-Ala-FMK) (Figura 8, A), quanto a formação de uma ponte salina com a guanidina da arginina do inibidor benziloxicarbonil-arginina-alanina-fluorometilcetona (Z-Arg-Ala-FMK) (Figura 8, B). A ligação de hidrogênio observada com a Gly66 é normalmente encontrada em complexos cruzaína-inibidor.



Figura 8 - Interações entre a cruzaína e os inibidores: A) Z-Tyr-Ala-FMK (PDB ID: 1AIM) e B) Z-Arg-Ala-FMK (PDB ID: 2AIM). Os diagramas de interação 2D foram gerados com o auxílio do *Poseview* (http://poseview.zbh.uni-hamburg.de)²⁶, desenvolvido pelo Centro de Bioinformática da Universidade de Hamburgo, Alemanha. As linhas pretas tracejadas representam ligações de hidrogênio e as linhas sólidas verdes representam interações hidrofóbicas²⁷. As figuras de interação 3D foram geradas pelo programa *Pymol* e os subsítios estão sombreados conforme figura anterior.

Os primeiros inibidores não covalentes descritos para a enzima foram derivados hidroximetilcetona²⁸. Um desses, mostrado na Figura 9, contém um grupo fenilalanina na posição P2. De maneira similar ao observado para outras estruturas, o subsítio S2 apresentou interações hidrofóbicas com o ligante através dos resíduos Leu67, Ala133 e Leu157. Como esperado, devido a sua flexibilidade, o resíduo Glu205 no fundo do bolsão S2 se afastou do

grupo fenilalanina do inibidor. Nesse complexo, o inibidor é estabilizado por ligações de hidrogênio com os resíduos Gln19, Gly66 e Ser61, apresentando também uma ligação de hidrogênio com a His159, resíduo pertencente à tríade catalítica.



Figura 9 - Interações entre a cruzaína e um inibidor reversível hidroximetilcetona (PDB ID: 1ME3).

Uma série de inibidores α -cetamidas, α -ceto ácidos e α -ceto ésteres foram identificados por processos de triagens bioquímicas e posteriormente otimizados utilizando métodos de SBDD aplicados à modificação de resíduos específicos nas posições P1, P2, P3 e P1'. Um destes inibidores, mostrado na Figura 10, em complexo com a cruzaína, deu origem a uma nova estrutura cristalográfica com resolução de 2,3 Å²⁹. Esse inibidor está ligado a enzima através de ligação covalente com o resíduo Cys25 . O fragmento benziloxicarbonil (Z) ocupa o subsítio S3, interagindo com os resíduos Gly66, Ala133 e Leu67. O anel aromático do grupo Z faz interações hidrofóbicas fracas com Leu 67, enquanto que o grupo fenilalanina na posição P2 está bem acomodado no subsítio hidrofóbico S2. O Glu205, localizado na base do S2, está na conformação aberta, não interferindo nas interações deste fragmento com o S2. Além disso, ligações de hidrogênio são observadas entre o inibidor e os resíduos Gly66, Asp158 e Gln19.



Figura 10 - Interações entre a cruzaína e um inibidor a-ceto éster (PDB ID: 1U9Q).

Em particular, entre os inibidores de cruzaína descritos na literatura, destaca-se a vinilsulfona denominada "K777", que além de elevada afinidade e seletividade, apresentou potente atividade em ensaios *in vitro* e *in vivo* contra o *T. cruzi*, contribuindo para a comprovação da cruzaína como alvo terapêutico e para a consolidação do papel crucial do desenvolvimento de inibidores candidatos a fármacos. A estrutura cristalográfica desse inibidor irreversível, que está atualmente em testes pré-clínicos, em complexo com a cruzaína (PDB ID: 20Z2) foi determinada com resolução de 1,95 Å ¹⁷. Como mostrado na Figura 11, o inibidor K777 ocupa todos os subsítios da proteína, de S3 a S1'. As interações polares, observadas com os resíduos Gln19, Gly66, Asp161, His162 e Trp184, são responsáveis pelo ancoramento do inibidor no sítio ativo da enzima. Esse padrão de interação é consistente em outros complexos com inibidores desta classe química. O Glu208, em conformação aberta, aponta para a superfície da enzima a fim de evitar interações desfavoráveis com o grupo fenilalanina na posição P2. As interações hidrofóbicas ocorrem principalmente com os resíduos Gly23, Ser64, Gly65, Gly66, Leu67, Ala138, Leu160, Asp161, Ala141 e Met145.



Figura 11 - Interações intermoleculares entre a cruzaína e o inibidor K777.

1.3 Planejamento de Fármacos Baseado na Estrutura do Receptor

Um método bastante útil para o desenvolvimento de inibidores enzimáticos é o planejamento de fármacos baseado na estrutura do receptor (SBDD, do inglês *structure-based drug design*)³⁰. Esta estratégia engloba diversas técnicas computacionais que geralmente utilizam estruturas cristalográficas de proteínas alvo para a identificação, planejamento e otimização de novos ligantes. Aspectos fundamentais, como mecanismo e modo de ligação, são explorados para a caracterização das interações intermoleculares predominantes na formação de complexos entre o receptor alvo e os compostos bioativos. Entre as técnicas modernas utilizadas em SBDD, destaca-se a triagem virtual baseada na estrutura do receptor (SBVS, do inglês, *structure-based virtual screening*). Normalmente empregada nos estágios iniciais de projetos de descoberta e desenvolvimento de fármacos, a SBVS possibilita a seleção de moléculas candidatas a inibidores a partir de grandes bases de dados de compostos³¹.

Um esquema geral de SBVS é apresentado na Figura 12. A primeira etapa envolve a escolha de um alvo molecular para a doença em estudo. Após a identificação do alvo, uma estrutura tridimensional (3D) deve ser selecionada para a realização da etapa de triagem virtual de bases de dados de compostos, através de docagem molecular (do inglês, *molecular docking*). Os compostos selecionados da triagem virtual são avaliados em ensaios *in vitro* para
a determinação da potência, afinidade, seletividade e mecanismo de ação. Possíveis modificações estruturais ocorrem nessa fase de desenvolvimento. Os compostos mais promissores são empregados em estudos de química medicinal para a otimização de propriedades farmacodinâmicas e farmacocinéticas até a descoberta de um candidato promissor para a fase de desenvolvimento clínico³².



Figura 12 - Esquema geral da aplicação de métodos de SBVS.

1.3.1 Docagem Molecular

A docagem molecular é uma das principais estratégias de SBDD aplicadas na identificação de novas moléculas bioativas. A docagem envolve a predição das conformações dos ligantes no sítio de ligação da proteína alvo³³, seguida da avaliação (pontuação) dos modos de ligação. Neste contexto, ocorre inicialmente o processo de acoplamento (do inglês, *pose*) dos ligantes das bases de dados na cavidade apropriada da proteína alvo. Diversas

conformações e orientações de cada molécula são avaliadas de acordo com suas interações intermoleculares com os resíduos de aminoácido do sítio de ligação, sendo classificadas pelo algoritmo de docagem de cada programa com base em uma função de pontuação (do inglês, *scoring*). No final do processo, uma classificação é gerada para os ligantes da base de dados, sendo possível priorizar os candidatos mais promissores para avaliação experimental *in vitro* (Figura 13)³⁴.



Figura 13 - Esquema geral de docagem molecular de uma base de dados. As moléculas melhores classificadas podem ser selecionadas para avaliação biológica.

Os programas de docagem molecular possuem algoritmos de busca e funções de pontuação. Os algoritmos são responsáveis pela identificação dos possíveis modos de ligação entre a proteína e as moléculas. As funções de pontuação são aplicadas para tentar predizer a afinidade de ligação das moléculas acopladas. De acordo com o programa de docagem utilizado, as funções podem ser físico-químicas, experimentais ou baseadas no conhecimento. Existem atualmente cerca de 30 programas de docagem molecular³⁵, entre os quais se destacam o DOCK³⁶, GOLD³⁷ e Surflex³⁸, que foram utilizados nesta dissertação de mestrado.

O programa DOCK 3.5.54 utiliza um algoritmo de construção incremental que divide o ligante em diversos fragmentos e os posiciona na cavidade do sítio. Neste processo, tem-se uma construção parte a parte, com minimizações de energia após a adição de cada fragmento, até que toda a estrutura do ligante esteja no sítio. As moléculas são pontuadas através de uma função baseada em mecânica molecular, levando-se em consideração as interações de van der Waals, as interações eletrostáticas e as penalidades pela dessolvatação do ligante. O programa GOLD 4.0.1 (do inglês, *Genetic Optimization for Ligand Docking*) utiliza um algoritmo genético que se baseia em biologia evolutiva, como hereditariedade, mutação, seleção e recombinação. A função Goldscore, uma das mais utilizadas, emprega métodos de mecânica molecular baseados em campos de força. Esta função quantifica a energia das interações entre receptor e ligante, bem como a energia interna do ligante. O programa Surflex utiliza um algoritmo incremental e a função empírica de Hammerhead. O ligante é idealizado no sítio ativo, gerando um protomol, onde as poses são construídas. Funções como esta utilizam ajustes teóricos a partir de dados experimentais, incluindo energias de ligação e conformações derivadas de cristalografia de raios X.

As interações intermoleculares mais comumente envolvidas no processo de reconhecimento molecular^{39,40} são as ligações de hidrogênio⁴¹, de van der Waals, iônicas, hidrofóbicas, do tipo cátions- π envolvendo grupamentos positivamente carregados e anéis aromáticos⁴², e do tipo π - π . Desta forma, após concluído o trabalho de docagem, as melhores moléculas selecionadas devem ser inspecionadas visualmente para garantir a maior complementaridade possível com a proteína alvo.



2 **OBJETIVOS**

O objetivo geral desta dissertação de mestrado é a identificação e caracterização cinética de novos inibidores da enzima cruzaína de *T. cruzi*, candidatos a fármacos para o tratamento da doença de Chagas, empregando uma combinação de métodos computacionais (SBVS) e experimentais (avaliação bioquímica *in vitro*). Os objetivos específicos deste trabalho incluem:

- Realização de estudos estruturais de complexos cruzaína-inibidor para a caracterização de aspectos moleculares fundamentais para os estudos de SBDD;
- Planejamento e execução do processo de SBVS de bases de dados líder-similar (*lead-like*) e fragmento-similar (*fragment-like*) da coleção de compostos ZINC;
- Utilização dos programas DOCK, GOLD e Surflex nas etapas de SBVS;
- Seleção de compostos candidatos a novos inibidores da enzima-alvo;
- Expressão e purificação da enzima cruzaína de *T. cruzi*;
- Realização dos ensaios bioquímicos para a determinação de parâmetros como potência (IC₅₀) e afinidade (*K*_i), bem como para a elucidação do mecanismo de inibição;



3 MATERIAIS E MÉTODOS

3.1 Estudos Estruturais

Estudos estruturais da enzima cruzaína de *T. cruzi* foram realizados utilizando informações de estruturas cristalográficas disponíveis no PDB, com o auxílio dos programas de visualização molecular *Pymol*⁴³ e *Chimera*⁴⁴. As interações entre os complexos receptor-ligante também foram analisadas pela ferramenta "*Ligand Explorer*", disponível no PDB. Esta ferramenta permite observar as principais interações entre proteína e inibidor (ligações de hidrogênio, interações hidrofóbicas, entre outras). Desta forma, foi possível identificar os resíduos do sítio ativo da enzima que interagem com inibidores, entre outros, principalmente, os resíduos Gly66, Asp161, His162, e Glu208.

3.1.1 Preparação do Alvo Molecular

As 19 estruturas da cruzaína em complexo com ligantes, depositadas no PDB (PDB IDs: 3HD3, 2OZ2, 1U9Q, 1EWL, 1EWM, 1EWO, 1ME3, 1ME4, 1F29, 1F2A, 1F2B, 1F2C, 1EWP, 1AIM, 2AIM, 3IO6, 3IUT, 3KKU e 3LXS), foram utilizadas para caracterizar as principais interações intermoleculares entre a proteína e os ligantes. Após análise detalhada de todos os complexos, a estrutura de código 3KKU²⁴ (Figura 14) foi selecionada para os estudos de SBVS. Esta estrutura apresenta alta resolução (1,28Å) e o ligante em complexo faz apenas interações não covalentes com a proteína. Com o auxílio do programa Sybyl 8.0, a estrutura foi preparada removendo-se as moléculas de água, o ligante e as duplas conformações dos resíduos da estrutura. Foram adicionados átomos de hidrogênios e checado o estado de protonação dos resíduos His162 (forma protonada) e Cys25 (forma desprotonada)^{24,34,45}.



Figura 14 - Estrutura da enzima cruzaína (PDB ID: 3KKU). A) Complexo proteína-ligante, onde a proteína está representada em modelo de fitas e o ligante em bastão. B) Representação do sítio de ligação da enzima (superfície) e as interações com o inibidor (bastão). As interações de hidrogênio estão representadas pelas linhas tracejadas.

3.2 Base de Dados

A base de dados ZINC⁴⁶ contem uma ampla variedade de compostos disponíveis comercialmente. Esta base, que foi empregada nos estudos de SBVS deste trabalho de dissertação, reúne, atualmente, mais de 21 milhões de compostos (na época do desenvolvimento de nosso estudo reunia cerca de 14 milhões) que estão divididos em diversos subconjuntos de acordo com características físico-químicas específicas, entre outras. Para o processo de triagem virtual foram utilizadas duas bases de compostos da ZINC. A primeira (denominada Base 1), com 3.409.091 milhões de compostos, apresenta exclusivamente característica líder-similar (*lead-like*)⁴⁷: clog P < 3,5; massa molar < 350 g/mol; e ligações rotacionáveis ≤ 7 . A segunda (denominada Base 2), por sua vez, com cerca de 450.000 compostos, apresenta característica fragmento-similar (*fragment-like*)⁴⁸: clog $P \leq 2,5$; massa molar ≤ 250 g/mol; e ligações rotacionáveis ≤ 5 .

3.3 Triagem Virtual

A base de dados líder-similar (Base 1) foi primeiramente utilizada nos estudos de SBVS. A docagem dos 3.409.091 compostos foi realizada com o programa DOCK 3.5.54 (Figura 15). Os 500 melhores compostos (pontuação) passaram por criteriosa inspeção visual. Um subconjunto do programa DOCK, com 35.000 compostos (aproximadamente 1% do número de compostos da base de dados inicial), foi utilizado para os processos de SBVS com os programas GOLD 4.0.1 e Surflex, levando a seleção dos 500 melhores compostos (pontuação).



Figura 15 - Esquema geral dos processos de SBVS da Base 1. Inicialmente, o programa DOCK foi utilizado na docagem dos compostos. Um subconjunto foi selecionado para os estudos com os programas GOLD e Surflex. Os 500 melhores compostos de cada programa foram inspecionados visualmente para priorização de compostos para avaliação experimental em ensaios em *in vitro*.

De acordo com a Figura 15, as 500 melhores moléculas de cada processo individual de SBVS (DOCK, GOLD e Surflex) foram inspecionadas visualmente, considerando a pontuação de cada molécula, as características dos subsítios da enzima alvo e as interações comumente observadas em complexos cristalográficos entre cruzaína e inibidores. Neste contexto, foram favorecidos grupos aromáticos e hidrofóbicos ou grupos carregados positivamente no subsítio S2. Entre os subsítios S2 e S1, observou-se a presença de doadores e aceptores de ligação de hidrogênio, envolvidos principalmente em interações com os resíduos Gly66, Asp161 e Gln19, e em alguns casos, com o Glu208. A presença dessas interações foi considerada devido a conservação nas estruturas cristalográficas analisadas. Além disso, foram observadas possíveis ligações de hidrogênio e interações com grupos aromáticos entre os ligantes e os resíduos His162 e Trp184. Na Figura 16 podem ser observados os principais resíduos que realizam as interações com os compostos candidatos selecionados.



Figura 16 - Principais resíduos do sítio ativo da cruzaína que realizam interações com os compostos selecionados.

Um esquema similar, apresentado na Figura 17, caracteriza o trabalho de SBVS realizado com a base fragmento-similar (Base 2), com cerca de 450.000 compostos.



Figura 17 - Esquema geral dos processos de SBVS da Base 2. Inicialmente, o programa DOCK foi utilizado na docagem dos compostos. Um subconjunto foi selecionado para os estudos com os programas GOLD e Surflex. Os 500 melhores compostos de cada programa foram inspecionados visualmente para priorização de compostos para avaliação experimental em ensaios em *in vitro*.

3.4 Expressão e Purificação da Enzima Cruzaína

Para a realização dos experimentos de cinética enzimática planejados neste trabalho de dissertação, foi indispensável a obtenção da proteína pura e ativa. Algumas alterações foram aplicadas a um protocolo de expressão e purificação previamente descrito⁴⁹.

O vetor pQE30 contendo a construção gênica que codifica a cruzaína, gentilmente cedido pela Prof. Dra. Ana Paula Cabral de Araujo Lima, da Universidade Federal do Rio de Janeiro – UFRJ, foi transformado em células de *Escherichia coli* linhagem SG13009. Essa construção contém as regiões que codificam o pró-domínio e a região central da enzima, mas é truncada na extensão C-terminal, pois este possui resíduos que não são essenciais para a atividade enzimática⁵⁰. Além disso, esse plasmídeo permite a síntese de uma cauda de histidina, o que possibilita a purificação da enzima através de uma coluna de afinidade.

Para a expressão da cruzaína, células de *E. coli* linhagem SG13009 contendo o vetor pQE30 foram cultivadas em 500 mL de meio Luria Bertani líquido (LB-líquido) suplementado com 100 µg/mL kanamicina e 100 µg/mL¹ ampicilina, sob agitação constante (200 rpm) a 37° C, até atingir uma densidade óptica (D.O., a 600 nm) de 0,6. Após, 1 mM de IPTG (Isopropil- β -D-tiogalactopiranosídeo) foi adicionado ao meio para indução da expressão da proteína durante 4 h a 30° C. Na sequência, o meio foi centrifugado a 4.500 rpm durante 60 min a 4° C, o sobrenadante foi descartado, o precipitado (*pellet*) de bactérias foi ressuspendido em tampão de lise (100 mM fosfato de sódio, 10 mM Tris/HCl, 8 M uréia, pH 8) e incubado sob agitação por 1 h a 4° C. Após promover a lise, a solução foi centrifugada a 9.000 rpm durante 40 min a 4° C e o sobrenadante contendo a proteína foi utilizado na sequência para a purificação.

Na primeira etapa do processo de purificação foram utilizadas colunas de níquel (Ni-NTA Superflow - QUIAGEN). A coluna (aproximadamente 5 mL) foi equilibrada em tampão de lise e em seguida a solução contendo a proteína foi eluida. A proteína foi lavada com excesso de tampão de lise (Figura 18, fração 1). A eluição se deu por meio de gradientes de tampão (100 mM fosfato de sódio, 10 mM Tris/HCl, 8 M uréia, pH 6,3, 200 mM imidazol) até que toda a proteína fosse retirada da coluna e distribuída em alíquotas de 5 mL. Esta etapa de purificação foi monitorada por eletroforese em gel de poliacrilamida 15% (SDS-PAGE) (Figura 18, frações 2, 3 e 4). A análise do gel proporcionou o agrupamento das alíquotas da proteína, que foram concentradas até 1 mg/mL, utilizando concentradores Amicon com uma membrana de corte de 10 kDa.

A proteína purificada foi incubada em 10 mM de DTT a 37° C por 45 min e na sequencia diluída (1:100) em tampão Tris/HCl 100 mM pH 8, contendo EDTA 1 mM, L-arginina 250 mM e glicerol 20%, para o renovelamento. Após incubação de 20 h a 4° C, a proteína foi novamente concentrada até atingir uma concentração de 0,5 mg/mL.

Para ativação da proteína, foi adicionado o tampão de ativação (acetato de sódio 100 mM pH 5,5, EDTA 10 mM, DTT 5 mM e NaCl 1 M). Esta mistura foi ajustada para um pH

5,3 e a ativação da proteína ocorreu a 37º C. Durante a ativação, a atividade da enzima foi monitorada através de ensaios enzimáticos a cada 30 min, por um período de 4 a 7 h, até que a atividade atingisse um valor máximo. A ativação completa da enzima foi confirmada SDS-PAGE (Figura 18, fração 5).



Figura 18 - Análise por SDS-PAGE da expressão heteróloga e da purificação da cruzaína. Gel SDS-PAGE 15% corado com Coomassie-blue. Sendo M = padrão de massa molecular; 1 = fração pós coluna (*flowthrough*); 2-4 = frações pós purificação em coluna de níquel; 5 = enzima cruzaína após completada a ativação completa (alíquota em destaque em vermelho).

A última etapa da purificação foi realizada em coluna de resina tiopropil sefarose 6B (Thiopropyl Sepharose® 6B – GE Healthcare Life Sciences), previamente equilibrada com tampão de ligação pH 7,2 (20 mM fosfato de sódio, 150 mM NaCl). Para tanto, a proteína foi adicionada a resina equilibrada juntamente com o tampão de ligação e incubada durante a noite a 4° C, sob agitação lenta. Posteriormente, esta mistura foi adicionada a uma coluna para eluição da proteína através de tampão de ligação suplementado com 20 mM de DTT. A enzima pura foi lavada em sistema de ultra filtração Amicon para a troca de tampão, sendo armazenada em alíquotas de 30 μ L a -80° C, com o tampão de 0,1 M NaAc e 0,000005% Triton X-100.

3.5 Ensaios Enzimáticos

Ensaios de fluorescência foram padronizados e validados para a avaliação bioquímica dos compostos selecionados neste trabalho de dissertação. Para o monitoramento da fluorescência em fluorímetro de placa VICTOR3TM (PerkinElmer), utilizando-se placas pretas de 96 poços, foram utilizados os comprimentos de onda de 355 nm para excitação e de 460 nm para emissão⁵¹. Os ensaios enzimáticos foram realizados com uma mistura reacional contendo 0,5-1,5 nM de cruzaína e 5,0 μ M de substrato Z-Phe-Arg-AMC (benzoil-fenilalanina-arginina-aminometilcumarina) ($K_{\rm M} = 1,6 \,\mu$ M), em tampão acetato de sódio 100 mM pH 5,5, na presença de 5 mM de DTT e 0,01% de Triton X-100.

Os compostos foram testados inicialmente na concentração de 100 μ M, sendo que apenas os que apresentaram inibição enzimática $\geq 30\%$ foram selecionados para a determinação da potência biológica. Os valores de IC₅₀ (que se referem à concentração necessária de inibidor para reduzir em 50% a atividade da enzima cruzaína) foram determinados com pelo menos seis concentrações diferentes de inibidor. Todos os ensaios foram realizados em triplicata e os resultados reportados representam a média e o respectivo erro padrão. Os dados cinéticos foram analisados com o auxílio dos programas GraphPad Prism 5 e SigmaPlot 10.



4.1 Estudos Estruturais

Por ser uma cisteíno protease da família da papaína, as cadeias de aminoácidos das estruturas da cruzaína seguiam a mesma numeração das estruturas da papaína. No entanto, mais recentemente, as estruturas da cruzaína passaram a apresentar uma numeração diferente, seguindo a ordem dos resíduos de aminoácidos que formam as cadeias. Como a estrutura utilizada nos estudos de triagem virtual deste trabalho adota esta nova numeração, os resultados apresentados seguem a numeração dos resíduos encontrados no sítio ativo da enzima de acordo com a Tabela 2.

Resíduo	Numeração anterior	Numeração atual	
Asp	158	161	
His	159	162	
Trp	177	184	
Glu	205	208	

Tabela 2 – Numeração atual dos resíduos do sítio ativo da cruzaína.

Os estudos estruturais dos 19 complexos cruzaína-ligante levaram a compreensão das interações intermoleculares predominantes, como, por exemplo, a capacidade de clivar peptídeos contendo tanto resíduos hidrofóbicos quanto positivos em S2, o subsítio de maior importância para a especificidade da cruzaína. Para a análise das interações entre proteína e inibidor (presentes nos complexos) foi utilizada a ferramenta do PDB *Ligand Explorer*, que permitiu observar as ligações de hidrogênio e interações hidrofóbicas, entre outras. Algumas dessas interações foram representadas nos formatos 2D e 3D, com o auxílio do *Poseview* e do *Pymol*, respectivamente. Entre estas, encontram-se as que são majoritariamente conservadas nos complexos depositados, como as ligações de hidrogênio entre os ligantes e os resíduos Gln19, Ser61, Gly66, Asp161, His162 e Trp184, e as interações hidrofóbicas com os resíduos Gly23, Cys25, Ser64, Gly65, Gly66, Leu67, Ala138, Asp161, His162, Ala133 e Leu157.

Na Figura 19 (A e B), duas vinilsulfonas (inibidores irreversíveis) foram utilizadas para ilustrar algumas dessas interações com a cruzaína. Pode ser observado que os 2 inibidores possuem na posição P1 grupos homofenilalanina e na posição P2 grupos fenilalanina. Os grupos na posição P3 variam entre anéis aromáticos e heterociclos não-aromáticos como a morfolina. Nesses complexos estão presentes ligações de hidrogênio entre os inibidores e os resíduos Gly66, Asp158, Gln19 e Trp177. As interações hidrofóbicas ocorrem com Leu67, Ala133 e Leu157. A Figura 19 (C e D) permitiu estudar o modo de interação 3D dos inibidores desta classe química com os diferentes subsítios da proteína. O resíduo Glu205 (correspondente ao Glu208 na numeração nova) está presente em todas as estruturas analisadas na conformação aberta, não interagindo com os grupos P2 dos ligantes.



Figura 19 – Classe de inibidores vilnilsulfona. Diagramas 2D gerados pelo Poseview para as estuturas: A)
PDB ID: 1F29, B) PDB ID: 1F2B. C) Interações entre o inibidor A e a cruzaína (PDB ID: 1F29). D) Ocupação do sítio ativo da cruzaína pelas vinilsulfonas A e B.

O aduto covalente formado com a Cys25 da cruzaína pode ser observado na estrutura cristalográfica (PDB ID: 3I06; resolução de 1,1 Å) da Figura 20, na presença de um inibidor purínico⁵². As interações hidrofóbicas entre o inibidor covalente reversível e a enzima ocorrem principalmente com os resíduos Gly23, Ser64, Leu67, Ala138. Essas interações acomodam o inibidor entre os subsítios S1 e S2. As ligações de hidrogênio entre inibidor e enzima são formadas pelos resíduos Gln19, Cys25 e Asp161. Em S2, o Glu208 não faz nenhuma interação com o ligante.



Figura 20 - Interações entre a cruzaína e um inibidor purínico (PDB ID: 3IO6).

A flexibilidade do resíduo Glu208 pode também ser verificada no complexo entre a enzima e o inibidor irreversível da Figura 21. O derivado triazólico não apresenta na posição P2 volume estrutural suficiente para ocupar totalmente S2, o que permite a orientação do resíduo Glu208 em direção ao interior do subsítio⁵³. O grupo quinolina na posição P3 interage com o subsítio S3, sendo estabilizado por interações hidrofóbicas e por uma ligação de hidrogênio com o resíduo Ser61. Outras ligações de hidrogênio ocorrem com os resíduos Gln19 e Gly66.



Figura 21 - Interações entre a cruzaína e um inibidor não peptídico (PDB ID: 3IUT).

O análogo do inibidor K777, denominado WRR-483 (Figura 22), mostrou-se eficaz em testes *in vitro* e *in vivo*⁵¹. Esta vinilsulfona se encontra covalentemente ligada a cruzaína. Em contraste com o grupo fenilalanina do K777 presente em P2 (Figura 11), o inibidor WRR-483 possui uma cadeia lateral arginina que ocupa o subsítio S2, sendo capaz de formar uma ligação de hidrogênio com o Glu208. Outras interações observadas anteriormente também auxiliam na estabilização do complexo proteína-ligante.



Figura 22 - Interações entre a cruzaína e um inibidor derivado vinilsulfona (PDB ID: 3LXS).

Merece especial destaque o derivado benzimidazol do complexo cristalográfico da Figura 23 (PDB ID: 3KKU; resolução 1,28 Å)²⁴. Este inibidor reversível competitivo,

descoberto através de triagem biológica automatizada em larga escala (HTS, do inglês, *high-throughput screening*), possui valor de K_i de 2 μ M, caracterizando-se como um importante composto líder para estudos de química medicinal. O sítio ativo apresenta alto grau de conservação em comparação com outras estruturas da cruzaína. Contudo, uma mudança significativa é observada em relação ao resíduo Gln159, que assume uma conformação para interagir com o ligante. Pequenas diferenças são observadas na orientação dos resíduos que formam o S2, onde Glu208, Leu67 e Leu160 se orientam na direção do ligante. Assim como observado em outras estruturas, ligações de hidrogênio são encontradas entre ligante e os resíduos Gly66 e Asp161. O grupo *m*-bromofenil ocupa o S2. Por várias razões, incluindo a alta qualidade da estrutura, a orientação dos resíduos no sítio ativo, as interações observadas no complexo proteína-ligante e o modo de interação não covalente do inibidor, esta estrutura foi escolhida para a realização dos estudos de SBVS neste trabalho de dissertação.



Figura 23 - Interações entre a cruzaína e um inibidor reversível benzimidazol (PDB ID: 3KKU).

4.2 Triagem Virtual

O processo de SBVS da Base 1 foi realizado primeiramente utilizando-se o programa DOCK (Figura 24). Os programas CHEMGRID⁵⁴ e Delphi⁵⁵ foram utilizados para calcular os mapas de energia potencial. Múltiplas conformações e orientações de cada ligante foram pontuadas pelo programa, através de uma função baseada em mecânica molecular, que levou

em consideração as interações de van der Waals, as interações eletrostáticas e as penalidades pela dessolvatação do ligante⁵⁶. A melhor orientação de cada molécula foi usada para classificar a base de dados levando a seleção de 35.000 compostos, que foram posteriormente alvo de estudos de SBVS com os programas GOLD e Surflex. Os 500 melhores compostos pontuados pelo DOCK passaram por uma cuidadosa inspeção visual e apenas aqueles com melhor complementaridade multifuncional em relação ao sítio ativo foram selecionados entre os 10 que formaram o subconjunto final. Entre estes, 6 eram comercialmente disponíveis e foram adquiridos. De maneira similar, o processo de SBVS da Base 2 levou a seleção de 10 compostos, sendo que apenas 1 foi adquirido (Figura 24).



Figura 24 – Estratégia de SBVS das Bases 1 e 2 com o programa DOCK para a identificação de candidatos a inibidores da enzima cruzaína.

As estruturas 2D dos 6 compostos selecionados da Base 1 são mostrados na Tabela 3. A classificação dos compostos no programa DOCK, os respectivos códigos ZINC e as principais interações intermoleculares entre proteína e ligante são apresentados também na Tabela 3. As interações foram geradas com o auxilio do programa *Chimera* e as figuras com o programa *Pymol*.



Tabela 3 - Compostos selecionados da Base 1 pelo programa DOCK para avaliação bioquímica frente a enzima cruzaína de T. cruzi.

continua



O subconjunto com 35.000 compostos selecionados pelo DOCK foi empregado em estudos de SBVS com o programa GOLD (Figura 25). Os parâmetros padrões do GOLD foram utilizados e o sítio de ligação foi definido com um raio de 10 Å a partir da coordenada do átomo de carbono carbonílico da cadeia central amídica do ligante da estrutura utilizada. A avaliação das orientações das moléculas foi realizada pela função Goldscore. As 500 melhores

moléculas classificadas pelo programa GOLD passaram por uma cuidadosa inspeção visual levando a seleção de um subconjunto final de 10 compostos, sendo que 3 destes foram adquiridos para a avaliação *in vitro* (Tabela 4).



- Figura 25 Estratégia de SBVS da Base 1 com os programas GOLD e Surflex para a identificação de candidatos a inibidores da enzima cruzaína.
- Tabela 4 Compostos selecionados da Base 1 pelo programa GOLD para avaliação bioquímica frente a enzima cruzaína de T. cruzi.



continua



De forma semelhante, o subconjunto de 35.000 compostos foi alvo de estudos de SBVS com o programa Surflex (Figura 25). A estrutura foi preparada pelo módulo Biopolymer (Sybyl), com a adição dos hidrogênios e a consideração das cargas. Este programa utiliza um algoritmo que gera uma representação computacional do sítio de ligação (um protomol), onde os ligantes são idealizados e as poses construídas, sendo possível a minimização de energia após a docagem. As moléculas foram pontuadas por uma função empírica que inclui diversos termos como lipofilia, polaridade, entropia e solvatação. Ao final do processo, 5 compostos foram adquiridos para avaliação bioquímica (Tabela 5).



 Tabela 5 - Compostos selecionados da Base 1 pelo programa Surflex para avaliação bioquímica frente a enzima cruzaína de T. cruzi.

continua



Para a SBVS da Base 2 com o programa DOCK foi aplicada a estratégia esquematizada na Figura 24, levando a seleção final de 1 composto que foi adquirido para avaliação frente a enzima cruzaína. A Tabela 6 apresenta as informações correspondentes a esta molécula.

Seguindo a mesma estratégia geral da Figura 25 aplicada para a Base 1, realizou-se o processo de SBVS da Base 2 com os programas GOLD e Surflex. Ao final do trabalho, 2 compostos foram selecionadas pelo programa GOLD (Tabela 7) e 1 pelo programa Surflex (Tabela 8).



 Tabela 6 - Composto selecionado da Base 2 pelo programa DOCK para avaliação bioquímica frente a enzima cruzaína de T. cruzi.

Tabela 7 – Compostos selecionados da Base 2 pelo programa GOLD para avaliação bioquímica frente a enzima cruzaína de *T. cruzi*.



continua



 Tabela 8 - Composto selecionado da Base 2 pelo programa Surflex para avaliação bioquímica frente a enzima cruzaína de T. cruzi.



Ao final de todo o processo de SBVS, 18 compostos comercialmente disponíveis foram priorizados e adquiridos, sendo 14 compostos provenientes da Base 1 (líder-similar) e 4 da Base 2 (fragmento-similar). Todos os compostos foram adquiridos da empresa ENAMINE[©] (http://www.enamine.net).

4.3 Avaliação Bioquímica

A atividade catalítica da enzima cruzaína é monitorada com base na clivagem do substrato Z-Phe-Arg-AMC (benziloxicarbonil-fenilalanina-arginina-aminometilcumarina), que tem baixa fluorescência, mas libera o grupo 7-amino-4-metilcumarina, que apresenta alta fluorescência. O esquema da hidrólise do substrato pode ser observado na Figura 26.



Figura 26 - Esquema da clivagem do substrato Z-Phe-Arg-AMC.

Os 18 compostos selecionados do SBVS foram avaliados experimentalmente em concentração única de 100 μ M, sem e com incubação prévia com a enzima. Os compostos que apresentaram inibição superior a 30% foram selecionados para a determinação da potencia biológica (IC₅₀). Os resultados são apresentados na Tabela 9.

Composto	% Inibição	% Inibição contra a cruzaína	
	sem incubação	10 minutos de incubação	IC ₅₀ (μM)
1	0	6±3	ND
2	0	0	ND
3	7±5	15±3	ND
4	22±4	53±12	>100
5	8±1	41±5	>100
6	0	5±3	ND
7	21±2	38±4	>100
8	0	7±1	ND
9	0	6±1	ND
10	7±6	24±3	ND
11	5±3	16±7	ND
12	8±5	33±8	>100
13	3±1	66±7	20±2
14	6±4	7±2	ND
15	9±6	6±1	ND
16	100±0	100±0	0,58±8
17	25±13	4±2	ND
18	0	0	ND

Tabela 9 – Avaliação bioquímica *in vitro* de uma série de compostos selecionados por SBVS, contra a enzima cruzaína.

Todos os ensaios foram realizados em triplicada. Os valores de IC_{50} foram obtidos com 5-7 concentrações de inibidor com base em 2 experimentos independentes.

Os compostos 1-3, 6, 8-11, 14, 15, 17 e 18 foram inativos ou apresentaram modesta inibição, não sendo possível determinar valores de IC_{50} . Os compostos 4, 5, 7 e 12 apresentaram baixa potência, com valores de $IC_{50} > 100 \mu$ M. Verificou-se ainda uma relação tempo-dependente na inibição apresentada por estes compostos. O composto 13, embora também apresentando um perfil tempo-dependente, foi mais potente, com um valor de IC_{50} de 20 μ M. Por outro lado, o composto 16 mostrou uma inibição de 100% da enzima na concentração de 100 μ M, não apresentando relação tempo-dependente. Surpreendentemente, um valor de IC_{50} de 580 nM foi determinado para este composto 16 (Figura 27).



Figura 27 – Gráfico concentração x resposta para a determinação de IC₅₀ do composto 16.

Os resultados promissores obtidos levaram a realização de experimentos para a determinação do mecanismo de inibição. Claramente, a estratégia de SBVS empregada neste trabalho visava à identificação de inibidores do tipo reversível competitivo. O perfil característico de um inibidor competitivo é observado através de um gráfico denominado Lineweaver-Burk (ou duplo-recíproco), que é obtido com medidas de v_0 (velocidade inicial de reação) para diferentes concentrações de substrato ([S]), variando-se as concentrações de inibidor ([I]). Neste caso, as interseções com o eixo y (valores recíprocos de v_0) fornecem o valor de velocidade máxima $(1/V_{max})$, que não se altera na presença de um inibidor competitivo. Os interceptos com o eixo x (valores recíprocos de [S]) fornecem os valores de $K_{\rm M}^{\rm app}$ ($K_{\rm M}$ aparente, -1/ $K_{\rm M}$), que aumentam com o aumento de [I] de um inibidor competitivo. O resultado da avaliação do composto 16 é mostrado na Figura 28, confirmando um mecanismo do tipo competitivo. A curva de menor inclinação refere-se à ausência de inibidor, sendo possível determinar o valor de K_M no intercepto x. Com o aumento de [I], as inclinações se tornam mais acentuadas, obtendo-se valores de $K_{\rm M}^{\rm app}$ cada vez maiores em x. Os valores de V_{max} não se alteram com diferentes [I]. Este resultado é extremamente importante porque confirma o sucesso da estratégia de SBVS empregada neste trabalho. O inibidor 16 se liga a enzima de maneira mútua e exclusiva em relação ao substrato utilizado.


Figura 28 – Gráfico duplo-recíproco para o inibidor 16. Concentrações de inibidor: controle - zero (●); 0,3 μM
(O); 0,6 μM (▼); 1,25 μM (△) e 2,50 μM (■).

O valor da constante de afinidade do inibidor K_i (constante de equilíbrio para a dissociação do complexo E•I), de 100 nM, foi determinado através do gráfico de Dixon, que é obtido com medidas de v_o (velocidade inicial de reação) para diferentes concentrações de inibidor ([I]), variando-se as concentrações de substrato ([S]). No caso, para o inibidor competitivo as linhas convergem acima do eixo x, sendo que o valor de x referente à intersecção das retas correspondente ao valor de K_i (- K_i).



Figura 29 – Determinação de K_i para o inibidor 16. Concentrações de substrato: 1,25 μ M (\bullet); 2,5 μ M (O); 20 μ M (\blacksquare) e 40 μ M (\Box).

É importante mencionar que o composto **16** foi classificado na posição 131 pelo programa GOLD, sendo assim considerado em nossos estudos de SBVS. Por outro lado, este composto não apresentou uma classificação razoável nos programas DOCK e Surflex. O programa DOCK classificou o composto **16** na posição 10.822, ao passo que o programa Surflex, somente na posição 20.951. Desta maneira fica claro que existem diferenças significativas entre os métodos empregados (algoritmos de busca e funções de pontuação), que levaram o inibidor **16** a ser classificado como um falso negativo em dois dos programas. Portanto, a estratégia de usar múltiplos programas pode ser útil em estudos de SBVS. Outra observação pertinente é em relação à seleção de moléculas por consenso. Embora seja uma estratégia interessante, em muitos casos, como o do presente trabalho, não parece ser ideal. Uma análise por consenso não levaria a identificação e seleção do inibidor nanomolar **16**.

No planejamento de inibidores enzimáticos utilizando triagens virtuais de bases fragmentos-similares, um parâmetro que vem sendo atualmente muito utilizado é a eficiência do ligante $(EL)^{57}$. EL é a medida da energia livre de ligação de um ligante para uma proteína específica. Matematicamente, a EL pode ser definida pela razão entre a energia livre de Gibbs (ΔG) e o número de átomos diferentes de hidrogênio presentes no ligante (Equação 1).

$$EL = -\Delta G/N \qquad (1)$$

sendo ΔG a energia livre de Gibbs e *N* o número de átomos diferentes de hidrogênio.

Em geral, um ligante pode ser considerado promissor para valores de $EL \ge 0,30$ kcal/mol/átomo⁵⁷, pois esse valor tende a diminuir durante o processo de otimização com o aumento do peso molecular do composto. Como a partir da triagem da base fragmento-similar obtivemos apenas um composto potente, calculamos o valor de EL para que possa nortear os futuros estudos na otimização deste composto líder.

De acordo com a equação 1, um valor de EL de 0,53 kcal/mol/átomo foi obtido para o composto **16**, o que é bastante expressivo para um fragmento molecular que poderá ser alvo futuro de otimização molecular.



5 CONCLUSÕES

Nessa dissertação de mestrado foram apresentados os resultados obtidos da aplicação da estratégia de SBVS de bases de dados de compostos para a enzima cruzaína de *T. cruzi*, alvo terapêutico para a doença de Chagas.

A integração de métodos computacionais e experimentais é uma estratégia moderna de planejamento de fármacos, que se mostrou extremamente útil neste trabalho de dissertação. A estratégia computacional desenvolvida, explorando duas bases de dados de compostos e três métodos de docagem molecular, permitiu explorar uma impressionante diversidade química, levando a seleção de um subconjunto final promissor de candidatos a inibidores da enzima alvo. Em particular, devido a suas características moleculares, o uso da base fragmento-similar (Base 2) caracterizou um importante desafio para a identificação de novos ligantes.

A realização dos ensaios experimentais permitiu a caracterização da potência (IC₅₀) dos melhores inibidores, bem como do mecanismo de ação, da constante de afinidade (K_i) e da eficiência do ligante (EL) do composto mais promissor da série (**16**), que apresentou um valor de IC₅₀ de 580 nM, de K_i de 100 nM e de EL de 0,53 kcal/mol/átomo. Vale destacar que este potente composto, selecionado de uma base com características similar a de fragmentos, é um excelente composto-líder para o desenvolvimento de um candidato a novo fármaco para o tratamento doença de Chagas.

REFERÊNCIAS

1 WORLD HEALTH ORGANIZATION - WHO. Working to overcome the global impact of neglected tropical diseases. Geneva, 2010. 184 p. Disponível em: < http://whqlibdoc.who.int/publications/2010/9789241564090_eng.pdf >. Acesso em: 15 Jun. 2012.

2 SCHMUNIS, G. The globalization of Chagas disease. **Isbt Science Series**, v. 2, n. 1, p. 6-11, 2007.

3 COURA, J. R.; VINAS, P. A. Chagas disease: a new worldwide challenge. **Nature,** v. 465, n. 7301, p. S6-S7, 2010. Supplement.

4 COURA, J. R.; CASTRO, S. L. D. A critical review on Chagas disease chemotherapy. **Memórias do Instituto Oswaldo Cruz,** v. 97, p. 3-24, 2002.

5 HOARE, C. A.; WALLACE, F. G. Developmental stages of trypanosomatid flagellates - a new terminology. **Nature**, v. 212, n. 5068, p. 1385-1386, 1966.

6 SOEIRO, M. N. C.; DE CASTRO, S. L. *Trypanosoma cruzi* targets for new chemotherapeutic approaches. **Expert Opinion on Therapeutic Targets**, v. 13, n. 1, p. 105-121, 2008.

7 DIAS, L. C. et al. Chemotherapy of chagas' disease: state of the art and perspectives for the development of new drugs. **Quimica Nova**, v. 32, n. 9, p. 2444-2457, 2009.

8 SAJID, M.; MCKERROW, J. H. Cysteine proteases of parasitic organisms. **Molecular Biochemical Parasitolology**, v. 120, n. 1, p. 1-21, 2002.

9 KENNEDY, J.; TURAN, N. Handbook of proteolytic enzymes. **Bioseparation**, v. 8, n. 6, p. 340-341, 1999.

10 RAO, M. B. et al. Molecular and biotechnological aspects of microbial proteases. **Microbiology Molecular Biology Review**, v. 62, n. 3, p. 597-635, 1998.

11 SCHECHTER, I.; BERGER, A. On the size of the active site in proteases. I. papain. **Biochemical and Biophysical Research Communications**, v. 27, n. 2, p. 157-162, 1967.

12 TURK, B. Targeting proteases: successes, failures and future prospects. **Nature Reviews Drug Discovery**, v. 5, n. 9, p. 785-99, 2006.

13 CAZZULO, J. J.; STOKA, V.; TURK, V. The major cysteine proteinase of *Trypanosoma cruzi*: a valid target for chemotherapy of Chagas disease. **Current Pharmaceutical Design**, v. 7, n. 12, p. 1143-1156, 2001.

14 MCKERROW J. H. Development of cysteine protease inhibitors as chemotherapy for parasitic diseases: insights on safety, target validation, and mechanism of action. **International Journal of Parasitology**, v. 29, n. 6, p. 833-837, 1999.

15 SAJID, M. et al. Cruzain : the path from target validation to the clinic. Advances in Experimental Medicine and Biology, v. 712, p. 100-115, 2011.

16 GILLMOR, S. A.; CRAIK, C. S.; FLETTERICK, R. J. Structural determinants of specificity in the cysteine protease cruzain. **Protein Science**, v. 6, n. 8, p. 1603-1611, 1997.

17 KERR, I. D. et al. Vinyl sulfones as antiparasitic agents and a structural basis for drug design. **Journal of Biological Chemistry**, v. 284, n. 38, p. 25697-25703, 2009.

18 BRINEN, L. S. et al. A target within the target: probing Cruzain's P1 ' site to define structural determinants for the Chagas' disease protease. **Structure with Folding & Design**, v. 8, n. 8, p. 831-840, 2000.

19 BRYANT, C. et al. Novel non-peptidic vinylsulfones targeting the S2 and S3 subsites of parasite cysteine proteases. **Bioorganic & Medicinal Chemistry Letters,** v. 19, n. 21, p. 6218-6221, 2009.

20 BRAK, K. et al. Identification of a new class of nonpeptidic inhibitors of cruzain. **Journal of American Chemical Society**, v. 130, n. 20, p. 6404-10, 2008.

21 ZANATTA, N. et al. Convergent synthesis and cruzain inhibitory activity of novel 2-(N 'benzylidenehydrazino)-4-trifluoromethyl-pyrimidines. **Bioorganic & Medicinal Chemistry**, v. 16, n. 24, p. 10236-10243, 2008.

22 FUJII, N. et al. Discovery of potent thiosemicarbazone inhibitors of rhodesain and cruzain. **Bioorganic & Medicinal Chemistry Letters,** v. 15, n. 1, p. 121-123, 2005.

23 BORCHHARDT, D. M. et al. Biochemical evaluation of a series of synthetic chalcone and hydrazide derivatives as novel inhibitors of cruzain from *Trypanosoma cruzi*. Journal of the Brazilian Chemical Society, v. 21, n. 1, p. 142-150, 2010.

24 FERREIRA, R. S. et al. Complementarity between a docking and a high-throughput screen in discovering new cruzain inhibitors. **Journal of Medicinal Chemistry**, v. 53, n. 13, p. 4891-4905, 2010.

25 M MCGRATH, M. E. et al. The crystal structure of cruzain: a therapeutic target for Chagas' disease. **Journal of Molecular Biology**, v. 247, n. 2, p. 251-9, 1995.

26 STIERAND, K.; RAREY, M. From modeling to medicinal chemistry: automatic generation of two-dimensional complex diagrams. **ChemMedChem**, v. 2, n. 6, p. 853-60, 2007.

27 STIERAND, K.; MAAB, P. C.; RAREY, M. Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. **Bioinformatics**, v. 22, n. 14, p. 1710-1716, 2006.

28 HUANG, L.; BRINEN, L. S.; ELLMAN, J. A. Crystal structures of reversible ketonebased inhibitors of the cysteine protease cruzain. **Bioorganic & Medinal Chemistry**, v. 11, n. 1, p. 21-9, 2003.

29 CHOE, Y. et al. Development of alpha-keto-based inhibitors of cruzain, a cysteine protease implicated in Chagas disease. **Bioorganic & Medicinal Chemistry,** v. 13, n. 6, p. 2141-2156, 2005.

30 ANDRICOPULO, A. D.; SALUM, L. B.; ABRAHAM, D. J. Structure-based drug design strategies in medicinal chemistry. **Current Topics in Medicinal Chemistry**, v. 9, n. 9, p. 771-790, 2009.

31 FERREIRA, R. S. et al. In silico screening strategies for novel inhibitors of parasitic diseases. **Expert Opinion on Drug Discovery**, v. 6, n. 5, p. 481-489, 2011.

32 BARAS, A. I.; BARAS, A. S.; SCHULMAN, K. A. Drug development risk and the cost of capital. **Nature Reviews on Drug Discovery**, v. 11, n. 5, p. 347-8, 2012.

33 KITCHEN, D. B. et al. Docking and scoring in virtual screening for drug discovery: methods and applications. **Nature Reviews on Drug Discovery**, v. 3, n. 11, p. 935-49, 2004.

34 FERREIRA, R. S.; GLAUCIUS, O.; ANDRICOPULO, A. D. Integração das técnicas de triagem virtual e triagem biológica automatizada em alta escala: oportunidades e desafios em P&D de fármacos. **Química Nova,** v. 34, p. 1770-1778, 2011.

35 LEACH, A. R.; SHOICHET, B. K.; PEISHOFF, C. E. Prediction of protein-ligand interactions. docking and scoring: successes and gaps. **Journal of Medicinal Chemistry**, v. 49, n. 20, p. 5851-5855, 2006.

36 SHOICHET, B. K.; KUNTZ, I. D. Protein docking and complementarity. Journal of Molecular Biology, v. 221, n. 1, p. 327-346, 1991.

37 JONES, G. et al. Development and validation of a genetic algorithm for flexible docking. **Journal of Molecular Biology,** v. 267, n. 3, p. 727-748, 1997.

38 JAIN, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. **Journal of Medicinal Chemistry**, v. 46, n. 4, p. 499-511, 2003.

39 DAVIS, A.; TEAGUE, S. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. **Angewandte Chemie International Edition**, v. 38, n. 6, p. 736-749, 1999.

40 BISSANTZ, C.; KUHN, B.; STAHL, M. A medicinal chemist's guide to molecular interactions. Journal of Medicinal Chemistry, v. 53, n. 14, p. 5061-5084, 2010.

41 PANIGRAHI, S. K.; DESIRAJU, G. R. Strong and weak hydrogen bonds in the proteinligand interface. **Proteins**, v. 67, n. 1, p. 128-141, 2007.

42 MINOUX, H.; CHIPOT, C. Cation $-\pi$ interactions in proteins: can simple models provide an accurate description? Journal of the American Chemical Society, v. 121, n. 44, p. 10366-10372, 1999.

43 **The PyMOL Molecular Graphics System.** Disponível em < http://www.pymol.org/>. Acesso em: 20 Dez. 2009.

44 PETTERSEN, E. F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1605-1612, 2004.

45 MLADENOVIC, M. et al. On the origin of the stabilization of the zwitterionic resting state of cysteine proteases: a theoretical study. **Journal of American Chemical Society**, v. 130, n. 27, p. 8696-705, 2008.

46 IRWIN, J.; SHOICHET, B. ZINC - a free database of commercially available compounds for virtual screening. Journal of Chemical Information and Modeling, v. 45, n. 1, p. 177-182, 2005.

47 TEAGUE, S. J. et al. The design of leadlike combinatorial libraries. **Angewandte Chemie International Edition**, v. 38, n. 24, p. 3743-3748, 1999.

48 CARR, R. A. E. et al. Fragmente-based lead discovery: leads by design. **Drug Discovery Today**, v. 10, n. 14, 2005.

49 EAKIN, A. E. et al. The sequence, organization, and expression of the major cysteine protease (cruzain) from *Trypanosoma cruzi*. **Journal of Biological Chemistry**, v. 267, n. 11, p. 7411-20, Apr 1992.

50 EAKIN, A. E. et al. Production of crystallizable cruzain, the major cysteine protease from *Trypanosoma cruzi*. Journal of Biological Chemistry, v. 268, n. 9, p. 6115-8, 1993.

51 CHEN, Y. T. et al. In vitro and in vivo studies of the trypanocidal properties of WRR-483 against *Trypanosoma cruzi*. **PLoS Neglected Tropical Diseases**, v. 4, n. 9, 2010. DOI: 10.1371/journal.pntd.0000825.

52 MOTT, B. T. et al. Identification and optimization of inhibitors of trypanosomal cysteine proteases: cruzain, rhodesain, and TbCatB. **Journal of Medicinal Chemistry**, v. 53, n. 1, p. 52-60, 2010.

53 BRAK, K. et al. Nonpeptidic tetrafluorophenoxymethyl ketone cruzain inhibitors as promising new leads for Chagas disease chemotherapy. **Journal of Medicinal Chemistry**, v. 53, n. 4, p. 1763-1773, 2010.

54 MENG, E. C.; SHOICHET, B. K.; KUNTZ, I. D. Automated docking with grid-based energy evaluation. Journal of Computational Chemistry, v. 13, n. 4, p. 505-524, 1992.

55 GILSON, M. K.; HONIG, B. H. Calculation of electrostatic potentials in an enzyme active site. **Nature**, v. 330, n. 6143, p. 84-86, 1987.

56 SHOICHET, B.; LEACH, A.; KUNTZ, I. Ligand solvation in molecular docking. **Proteins-Structure Function and Genetics**, v. 34, n. 1, p. 4-16, 1999.

57 HOPKINS, A. L.; GROOM, C. R.; ALEX, A. Ligand efficiency: a useful metric for lead selection. **Drug Discovery Today**, v. 9, n. 10, p. 430-431, 2004.