

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Exatas e de Tecnologia
Departamento de Química
Programa de Pós-graduação

**Integração de Métodos *in silico* e *in vitro* para o Planejamento de
Inibidores da Enzima Cruzaína**

HELTON JOSÉ WIGGERS*

Tese apresentada ao Programa de Pós-
Graduação em Química para a obtenção
do título de doutor em Ciências com área
de concentração em Química Orgânica

Orientador: Prof. Dr. Carlos Alberto Montanari

*Bolsista FAPESP

São Carlos

2011

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária/UFSCar**

W655im

Wiggers, Helton José.

Integração de métodos *in silico* e *in vitro* para o planejamento de inibidores da enzima cruzaina / Helton José Wiggers. -- São Carlos : UFSCar, 2011.
150 f.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2011.

1. Química orgânica. 2. Chagas, Doença de. 3. Química medicinal. 4. *Trypanosoma cruzi*. 5. Planejamento de substâncias bioativas. I. Título.

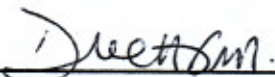
CDD: 547 (20^a)

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Exatas e de Tecnologia
Departamento de Química
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA
Curso de Doutorado

*Assinaturas dos membros da banca examinadora que avaliaram e aprovaram a defesa de tese de doutorado do candidato **Hélton José Wiggers**, realizada em 30 de junho de 2011:*



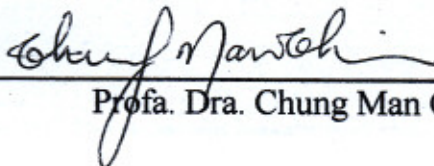
Prof. Dr. Carlos Alberto Montanari



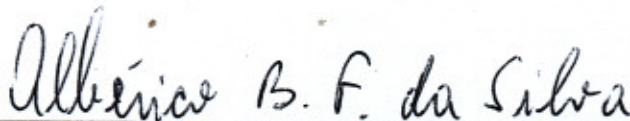
Profa. Dra. Dulce Helena Ferreira de Souza



Prof. Dr. Otavio Henrique Thiemann



Profa. Dra. Chung Man Chin



Prof. Dr. Albérico Borges Ferreira da Silva

DEDICATÓRIA

Aos meus pais Arni José e Santana da Silva Wiggers, aos meus irmãos e suas esposas Ivonei e Cleuni Fretta Wiggers; Helder Ivan e Thatiany Goularth Carneiro Wiggers; Heliton Augusto Wiggers e meus sobrinhos Lara Fretta Wiggers, Iuri Fretta Wiggers e Heloise Caroline Wiggers.

AGRADECIMENTOS

A Deus pela vida.

A minha família pela educação, incentivo aos estudos, orações, suporte em todos os momentos da vida e amor.

A Juliana Cheleski por aceitar trilharmos juntos os caminhos da vida.

Ao Prof. Dr. Carlos Montanari pela orientação, desafios propostos e confiança para o desenvolvimento do trabalho e sua esposa Dr. Maria Luíza Montanari (Malu) pela amizade e dividir suas experiências.

Aos amigos e colegas de grupo pela amizade, alegrias e descontrações durante o período de desenvolvimento da tese, discussões construtivas, colaborações nos artigos científicos, ajuda no planejamento e execução de experimentos, companhia nos congressos e inúmeras atividades alegres realizadas conjuntamente: Dr. Josmar Rocha e Adriane Pauluk; Geraldo Sartori; Renato e Mábia Freitas; Raphael Montanari; Dr. Peter W. Kenny; Vinicius Sciuti; Alessandra Figueiredo; Renato Bauab; Prof. Dr. Fabio e Patrícia Molfetta; Prof. Dr. Roberto Haiduke e Dr. Claudia Fileto (Guiga); Prof. Dr. Andrei e Patrícia Leitão, Igor Prokopczyk e Aline; Jean Francisco; Ricardo Goulart; Fabyana Soares; Emanuella Fonseca, Dr. Aderson Zottis; Juliane Sempionato; Prof. Dr. Maria Gotetti Vasconcelos Silva e Fabiana Rossini.

A todos meus amigos de Laranjeiras do Sul e de Guarapuava onde cursei a graduação (UNICENTRO).

Ao grupo de Enzimologia da Universidade Federal de São Paulo (UNIFESP) por fornecer o plasmídeo e orientações para a produção e purificação da enzima cruzaína. Prof. Dr. Luiz Juliano Neto e Dra. Maria Helena.

Ao grupo de Biologia Molecular e Bioquímica (IQSC) pela produção da enzima cruzaina e principalmente pela amizade. Prof. Dr. Júlio Borges, Sabrina Mattos, Natalia Rodrigues e Francisco Edvan.

Ao grupo de Biofísica Molecular Sérgio Mascarenhas: Andressa e Bel por sempre estarem dispostas a ajudar a resolver os problemas técnicos.

Ao Instituto de Química de São Carlos (IQSC) da Universidade de São Paulo (USP) pela autorização para o desenvolvimento do trabalho em suas dependências.

Aos funcionários da Universidade Federal de São Carlos pelo profissionalismo e competência.

Ao Grupo de Bioquímica de Proteínas, Fisiologia do Exercício (UFSCar) por estar de portas abertas para o uso dos equipamentos. Prof. Dra. Heloísa S. S. Araújo.

Grupo de Parasitologia, Faculdade de Medicina, USP, Ribeirão Preto pela oportunidade dos testes das substâncias contra o *T. cruzi*. Prof. Dr. João Santana e Dr. Renata Sesti.

As agências de fomento pelo financiamento dos projetos do NEQUIMED-PN e pela bolsa Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) e a Fundação de Amparo a pesquisa do Estado de São Paulo (FAPESP – Processo 2006/06291-3).

EPÍGRAFE

Tu és aquilo que o profundo desejo te impulsiona.

O que for teu desejo, assim será tua vontade.

O que for tua vontade, assim serão teus atos.

O que foram teus atos, assim será teu destino.

Brihadaranyaka Upanishad IV, 4.5.

RESUMO

A doença de Chagas, causada pelo parasito tripanossomatídeo *Trypanosoma cruzi*, é uma doença endêmica distribuída por toda América Latina. Os fármacos disponíveis são ineficientes e apresentam sérios efeitos colaterais. Portanto, são necessários novos fármacos para a quimioterapia da doença de Chagas. A enzima cruzaina (EC 3.4.22.51) de *T. cruzi* é expressa durante todo o ciclo de vida e representa um alvo validado contra a doença da Chagas. A busca de novos inibidores da enzima cruzaina foi realizada pelo desenvolvimento de uma estratégia integrada utilizando métodos *in silico* baseados na estrutura do ligante e do receptor e métodos *in vitro*. O consenso dos diferentes métodos de ensaios virtuais permitiu a seleção de 23 moléculas para os ensaios *in vitro*, a partir de uma coleção virtual com cerca de 8,5 milhões de estruturas provenientes do banco de moléculas comerciais ZINC e 12 apresentaram atividade inibitória frente à enzima cruzaina e sua homóloga catepsina-L de humanos. Os valores de IC_{50} variaram entre 5,6 e 73,9 μM para cruzaina e 8,6 a 89,1 para catepsina-L. As constantes de inibição aparentes (K_i^{app}) da série de compostos identificados variaram entre 3,7 e 64,5 μM para a cruzaina e 3,8 a 87,1 μM para a catepsina-L e apresentaram mecanismos competitivos de inibição. Novas classes moleculares inéditas de inibidores não covalentes da enzima cruzaina foram identificadas. Duas substâncias foram validadas como inibidores através de avaliação de análogos para o estabelecimento de relações entre a estrutura molecular e atividade biológica, uma substância com eficiência do ligante de $0,33 \text{ kcal mol}^{-1} \text{ NA}^{-1}$ foi identificada. Dois inibidores enzimáticos apresentaram atividade tripanossomicida frente à cepa Y do *T. cruzi* em sua forma tripomastigota com potência comparável ao Benzonidazol[®]. A inibição da enzima cruzaina com atividade confirmada frente ao *T. cruzi* oferece a oportunidade de otimização da estrutura molecular da substância matriz e valoriza a racionalização do processo de identificação de substâncias bioativas com modo de ação conhecido.

ABSTRACT

Chagas disease, caused by the flagellate protozoan of the family Trypanosomatidae *Trypanosoma cruzi*, is endemic in Latin America. The available drugs are ineffective and cause severe side effects. Therefore, it is necessary the discovery and development of new drugs for Chagas disease chemotherapy. The cruzain enzyme (EC 3.4.22.51) is expressed in all *T. cruzi* life cycle and represents a valid target against Chagas disease. The search for new cruzain enzyme inhibitors was carried out through the development of a new strategy using *in silico* methods, based on integrated ligand and target virtual screening. The consensus of different virtual screenings allowed the selection of 23 molecules for *in vitro* assays, from a virtual library containing approximately 8.5 million structures collected from the commercial database ZINC. The compounds were assayed against the cruzain and human homologous cathepsin-L and 12 presented inhibitory activity. The IC₅₀ values ranged from 5.6 to 73.9 μM for cruzain and 8.6 to 89.1 μM for cathepsin-L. The apparent inhibition constant (K_i^{app}) of the identified compounds ranged from 3.7 and 64.5 μM for cruzain and 3.8 to 87.1 μM for the cathepsin-L and showed competitive inhibition mechanisms. New molecular classes of non-covalent inhibitors of the enzyme cruzain were identified. Two substances were validated as inhibitors by evaluating compound analogs to establishing relationships between molecular structure and biological activity; a substance with the ligand efficiency of 0.33 kcal mol⁻¹ NA⁻¹ was identified. Two enzyme inhibitors showed trypanocidal activity against the Y strain of *T. cruzi* trypomastigotes with potency comparable to the drug benznidazole[®]. The micromolar activity of the compounds against the cruzain enzyme and the confirmed activity against the parasite provide the opportunity for molecular optimization and improve the bioactive compounds design with known mode of action.

LISTA DE FIGURAS

FIGURA 1.1: Distribuição Geográfica da Doença de Chagas	3
FIGURA 1.2: Ciclo de Vida do <i>Trypanosoma cruzi</i>	4
FIGURA 1.3: Representação das estruturas químicas do nifurtimox e benzonidazol .	6
FIGURA 1.4: Esquema de reação de clivagem de ligação peptídica por proteases .	11
FIGURA 1.5: A cisteína doa o próton para a histidina formando um par iônico correspondente a enzima ativada	12
FIGURA 1.6: Mecanismo de catálise de enzimas cisteíno proteases	13
FIGURA 1.7: Estrutura tridimensional da cruzaina (código PDB 2AIM)	14
FIGURA 1.8: Estrutura do sítio ativo representada com a superfície de Van der Waals, com a identificação das posições dos sub-sítios S1 a S4 (código PDB 3HD3).	14
FIGURA 1.9. Estruturas das bis-arilacil-hidrazidas identificadas por docagem molecular utilizando DOCK e derivado com alta potência frente à cruzaina	15
FIGURA 1.10: Estruturas (a) bidimensional e (b) tridimensional da cruzaina complexada com os inibidores Bz-Arg-Ala-CH ₂ F e Bz-Tyr-Ala-CH ₂ F	16
FIGURA 1.11: Representação da estrutura molecular do composto a) K11777 e seu análogo b) WRR-483.....	17
FIGURA 1.12: Estrutura química de dois potentes inibidores da cruzaina das classes a) 1,3-epoxi cetonas e b) epóxisucinil como grupos reativos a cisteína	18
FIGURA 1.13: Estruturas químicas dos primeiros inibidores de cruzaina não peptídicos identificados, classes a) aril tiouréia e b) aril uréia.....	18
FIGURA 1.14: Estruturas químicas dos compostos baseados em cetonas a) estrutura do composto com melhor seletividade frente as catepsinas B e L e b) composto co-cristalizado com a cruzaina (código PDB 1ME4)	19

FIGURA 1.15: Estruturas químicas representativas das classes a) Ceto amida, b) aldeído, c) ceto Ester e d) ceto ácido, potentes inibidores na cruzaina.....	19
FIGURA 1.16: Estruturas químicas representativas das classes a) isatinas e b) tiosemicarbazonas.....	20
FIGURA 1.17: Estruturas químicas representativas das classes a) vinilsulfonas não peptídicas, b) chalcona, c) tiosemicarbazonas, d) hidrazida e e) hidroximetilfurazona	21
FIGURA 1.18: Estruturas químicas representativas do composto a) 2,3,5,6-tetrafluorofenóximetil cetona e b) esqueleto molecular utilizado para SAR	22
FIGURA 1.19: Etapas do processo de descoberta e desenvolvimento de novos fármacos.....	23
FIGURA 1.20: Alvos dos fármacos na terapêutica distribuídos por classes bioquímicas.....	24
FIGURA 1.21: Etapas de identificação, seleção e otimização de substâncias bioativas durante o processo de descoberta de novos fármacos (fase pré-clínica)..	27
FIGURA 3.1: Ilustração do processo de interação entre enzima e inibidor	37
FIGURA 3.2: Ensaio virtual para identificação de novos ligantes	38
FIGURA 3.3: Fluxograma das etapas envolvidas no processo de seleção de moléculas para os ensaios bioquímicos	55
FIGURA 3.4: a) Superposição das estruturas cristalográficas dos 15 co-complexos provenientes do PDB e b) restrições farmacofóricas utilizadas demonstrando a complementaridade com os ligantes das estruturas 1ME3 e 1ME4.	60
FIGURA 3.5. Alinhamento das estruturas primárias de enzimas proteases de diferentes organismos com a cruzaina. As letras em branco com fundo vermelho representa conservação dos aminoácidos em todas as estruturas. Letras vermelhas conservação em pelo menos três estruturas. O alinhamento foi feito no servidor MultiAlign.	61

FIGURA 3.6: Estrutura do ligante proveniente da estruturas 1ME4 docado pelo programa FlexX na enzima cruzaína (código PDB 1ME4), verde representa a pose do co-cristal e cinza a pose da docagem	63
FIGURA 3.7: a) Curva ROC das funções de pontuação do programa FlexX e frequência de moléculas em função de b) D_SCORE, c) G_SCORE e d) CHEMSCORE	64
FIGURA 3.8: a) sobreposição de um ligante docado sobre a estrutura do inibidor da estrutura cristalográfica 1ME4 (em vermelho) e b) vários inibidores no sítio da enzima cruzaína	65
FIGURA 3.9: Superposição da conformação do inibidor hidroximetil cetona gerada pelo programa OMEGA2 (cinza) e a conformação cristalográfica extraída da estrutura da cruzaína (código PDB 1ME4) (verde).....	67
FIGURA 3.10: Estrutura do ligante proveniente da estruturas 1ME4 docado pelo programa FRED na enzima cruzaína (código PDB 1ME4), verde representa a pose do co-cristal e cinza a pose da docagem	68
FIGURA 3.11: a) Curva ROC das funções de pontuação do programa FRED e frequência de moléculas em função de b) Shapegauss.....	68
FIGURA 3.12: a) Estrutura tridimensional de um ligante bem pontuado pelo programa FRED superposto a estrutura cristalográfica (código PDB 1ME4) e b) vários ligantes bem pontuados pelo programa FRED demonstrando a ocupação do sítio ativo	69
FIGURA 3.13: (A) Curva ROC das métricas de pontuação do programa ROCS e frequência de moléculas em função da métrica Tversky; (a) inibidor_1ME4, (b) 92260 e (c) K11777.....	72
FIGURA 3.14: Frequência do número de estruturas em função da métrica Tversky para as estruturas dos inibidores de referência (a) inibidor_1ME4, (b) 92260 e (c) K11777.....	73

FIGURA 3.15: Frequência do número de estruturas em função da métrica Tversky para as estruturas dos inibidores de referência (a) inibidor_1ME4, (b) 92260 e (c) K11777; e das funções de pontuação dos programas FlexX CHEMSCORE (d), G_SCORE (e), D_SCORE (f) e FRED SHSPEGAUSS (g).....	75
FIGURA 3.16: a) Estrutura do ligante proveniente da estruturas 1ME4 docado pelo programa Glide na enzima cruzaina (código PDB 1ME4), verde representa a pose do co-cristal e cinza a pose da docagem e b) curva ROC da função de pontuação GlideScore utilizada pelo programa Glide.....	77
FIGURA 3.17: a) sobreposição de um ligante docado sobre a estrutura do inibidor da estrutura cristalográfica 1ME4 (em vermelho) e b) vários inibidores no sítio da enzima cruzaina	77
FIGURA 3.18: a) Construção do Holograma Molecular e b) Desenvolvimento do modelo de HQSAR	79
FIGURA 3.19: Análise de correlação dos valores de pK_i calculados pelos modelos de HQSAR em função dos valores experimentais; (a) modelo 2 ($r^2 = 0,958$) e (b) modelo 7 ($r^2 = 0,954$).....	81
FIGURA 3.20: Validação externa do modelo 7 de HQSAR ($r^2_{pred} = 0,970$).....	81
FIGURA 3.21: Mapa das contribuições dos fragmentos moleculares para a potência dos inibidores da enzima cruzaina: a) ligante da estrutura 1ME4 e b) 1ME3. O esquema de cores representa as contribuições dos fragmentos para a potência calculada pelo o modelo de HQSAR	82
FIGURA 3.22: Processo de autoescalamento das pontuações dos diferentes programas utilizados e análise consensual	84
FIGURA 3.23: Estrutura do inibidor da cruzaina K11777 co-cristalizado com a enzima a), estruturas de complexos preditos por docagem molecular com modo de interação semelhante ao inibidor cristalizado identificados na análise visual b), c) e d)	86

FIGURA 3.24. Representação das estruturas das substâncias selecionadas para os ensaios bioquímicos.....	87
FIGURA 4.1. Esquema ilustrativo do mecanismo de inibição reversível competitiva	93
FIGURA 4.2: Esquema ilustrativo do mecanismo de inibição reversível não competitiva	94
FIGURA 4.3: Esquema ilustrativo do mecanismo de inibição reversível incompetitiva	95
FIGURA 4.4: Diagrama de Jablonski modificado. (A) absorção de um fóton, (S_0) estado fundamental, (S_n) estado excitado singlete, (S_1) primeiro estado excitado singlete, (R_V) relaxamento vibracional, (CI) cruzamento interno, (CIS) cruzamento intersistemas, (T_n) estado excitado tripleto.....	98
FIGURA 4.5: Estratégia utilizada para gerar o sinal de fluorescência em reações catalisadas por cisteínas proteases	99
FIGURA 4.6: Estrutura do ZFRMCA (substrato sintético para a cruzaina)	99
FIGURA 4.7: Gel de Eletroforese da cruzaina. Banda 1) Marcador de peso molecular (MM), 2) amostra induzida com IPTG e 3) amostra não induzida da bactéria <i>E. coli</i> BL21-DE3; 4) Amostra induzida com IPTG e 5) Amostra não induzida da bactéria <i>E. coli</i> BL21-pLysS	108
FIGURA 4.8: Cromatograma da purificação por troca-iônica. A linha azul representa a absorbância em 280, a linha vermelha a razão de NaCl, e os números em vermelho os tubos coletados para análise em SDS-page.....	110
FIGURA 4.9: Gel de Eletroforese das frações coletadas na cromatografia de troca-iônica. Os números representam o número dos tubos onde se coletou as frações que apresentaram absorbância em 280 nm.....	111
FIGURA 4.10: Gel de Eletroforese das frações coletadas na cromatografia de exclusão por tamanho. Banda (MM) Marcador de peso molecular, 1 ao 7 frações que apresentaram absorbância em 280 nm.....	111

FIGURA 4.11: Curvas da velocidade em função da concentração de substrato ZFRMCA ajustadas na Equação de Michaelis-Menten a) cruzaina e b) catepsina-L	112
FIGURA 4.12: Efeito da razão $[S]/K_M$ no valor de IC_{50} em distintos mecanismos de inibição.....	114
FIGURA 4.13: Curva da inibição da enzima cruzaina (ajuste $R^2 = 0,99$; $Q^2/DoF = 0,71$).....	115
FIGURA 4.14: Curvas da velocidade em função da concentração de substrato ajustadas nas equações de mecanismo competitivo, não-competitivo e incompetitivo, o ensaio foi realizado na ausência dos inibidores e nas concentrações de 30, 50 e 70 μM	122
FIGURA 4.15: Estruturas dos compostos ativos contra a cruzaina que contém a eltildiamida, Ambedônio é o farmaco comercial e ao lado a subestrutura comum às três substâncias ativas, os compostos estão dispostos por ordem de afinidade pela cruzaina	125
FIGURA 4.16: Comparação entre estruturas e valores de IC_{50} entre compostos da literatura e o inibidor encontrado neste trabalho Nequimed38.....	126
FIGURA 4.17: Estruturas das moléculas análogas ao Nequimed24 testadas contra a enzima cruzaina, as unidades são IC_{50} em μM , LE em $kcal\ mol^{-1}\ NA^{-1}$ e MW $g\ mol^{-1}$	128
FIGURA 4.18: Estruturas dos compostos Nequimed162 e Nequimed161 docadas no sítio da cruzaina utilizando-se o programa Glide_XP. Estrutura cristalográfica demonstrando a posição de ligação do grupo oxadiazol na catepsina-L de Humanos	130
FIGURA 4.19: Estruturas das moléculas análogas ao Nequimed42 testadas contra a enzima cruzaina, as unidades são IC_{50} em μM , LE em $kcal\ mol^{-1}\ NA^{-1}$ e MW $g\ mol^{-1}$	132

FIGURA 4.20: Estruturas dos compostos Nequimed165 e Nequimed176 docadas no sítio da cruzaína utilizando-se o programa Glide_XP. Os sub-sítios da enzima estão enumerados como S1, S2 e S3..... 133

LISTA DE TABELAS

TABELA 3.1. Código PDB e parâmetros de resolução das estruturas da cruzaina..	57
TABELA 3.2: Valores da área sob a curva ROC, limites de energia estabelecidos para as funções de pontuação e número de moléculas encontradas no ZINC acima dos limites.....	65
TABELA 3.3. Valores da área sob a curva ROC, limites de energia estabelecidos para as funções de pontuação e número de moléculas encontradas no ZINC acima dos limites.....	69
TABELA 3.4. Valores da área sob a curva ROC, limites de energia estabelecidos para as funções de pontuação e número de moléculas encontradas no ZINC acima dos limites.....	73
TABELA 3.5. Resultados estatísticos para os modelos de HQSAR gerados para um conjunto de inibidores de cruzaina	80
TABELA 4.1: Comparação dos sistemas disponíveis para a produção de proteínas heterólogas.....	97
TABELA 4.2. Preparo do gel SDS-PAGE 15% para monitoramento da enzima.....	104
TABELA 4.3. Valores das constantes cinéticas e dados do ajuste na equação de Michaelis-Menten da reação catalisada pela enzima cruzaina.....	112
TABELA 4.4: Estruturas, valores de IC_{50} , LE e parâmetros de Lipinski das substâncias testadas frente às enzimas cruzaina e catepsina-L	117
TABELA 4.5: Valores dos parâmetros estatísticos dos ajustes das velocidades experimentais em função da concentração de substrato nas equações para inibidores competitivo, não-competitivo e incompetitivo	122
TABELA 4.6. Estrutura e valores de K_i dos compostos testados contra as enzimas cruzaina e catepsina-L de humanos, o mecanismo observado para as moléculas ativas foi o competitivo.....	123

LISTA DE ABREVIATURAS

ϵ	constante dielétrica de Coulomb
r^2	ajuste linear do modelo
v_b	velocidade inicial com inibidor
°C	grau Celsius
μL	microlitros
μM	micromolar
1D	uma dimensão ou unidimensional
2D	duas dimensões ou bidimensional
3D	três dimensões ou tridimensional
A	átomos
Å	Angstrom
ADME	absorção, distribuição, metabolismo e excreção (etapas da fase farmacocinética de um fármaco)
AIDS	síndrome da imunodeficiência adquirida
ANVISA	agência nacional de vigilância sanitária
ATP	adenosina trifosfato
AUC	área sob a curva ROC
B	ligações
C	conectividade
Ch	quiralidade
Cys25	aminoácido cisteína e o número correspondente a sua posição na cadeia primária
Da	Daltons (unidade de massa molecular)
DA	doador e aceitador de ligação hidrogênio
DDT	Ditiotreitol
ΔG	energia livre de Gibbs
DHODH	Diidroorotato desidrogenase
DMSO	dimetilsulfóxido
DNA	ácido desoxirribonucléico
DPG	1,3-difosfoglicerato
E	enzima
<i>E. coli</i>	<i>Escherichia coli</i>
E-64	[1-[N-[(L-3-trans-carboxioxirano-2-carbonila)-L-leucil]amino]-4-guanidinobutano] (inibidor de proteases)
EC	Comissão para classificação de enzimas
EDTA	ácido etilenodiamino tetra-acético

EI	complexo enzima-inibidor
ES	complexo enzima-substrato
ESI	complexo enzima/substrato/inibidor
FDA	Administração federal de alimentos e medicamentos dos Estados Unidos
FP _{rate}	falsos positivos recuperados
GAPDH	gliceraldeído-3-fosfato desidrogenase
Gln19 primária	aminoácido glutamina e o número correspondente a sua posição na cadeia primária
Glu205 primária	aminoácido glutamato e o número correspondente a sua posição na cadeia primária
Gly66 primária	aminoácido glicina e o número correspondente a sua posição na cadeia primária
H	hidrogênio
HGPRT	hipoxantina guanina fosforibosiltransferase
His159 primária	aminoácido histidina e o número correspondente a sua posição na cadeia primária
HQSAR Moleculares	relações quantitativas estrutura atividade baseado em hologramas Moleculares
HTS	ensaios em larga escala
I	inibidor
IC ₅₀	concentração para inibir 50% da atividade enzimática
IPTG	Isopropil β-D-1-tiogalactopiranosideo
IUPAC	união internacional de química pura e aplicada
K	graus Kelvin
K _a	constante de afinidade
kcal	kilocalorias
k _{cat}	constante catalítica
KCl	cloreto de potássio
kDa	kiloDalton
kg	kilograma
K _i	constante de inibição
K _M	constante de Michaelis Mentem
L	litros
LB	meio de cultura Luria Broth
LBVS	ensaio virtual baseado na estrutura do ligante
LE	eficiência do ligante

Leu67 primária	aminoácido leucina e o número correspondente a sua posição na cadeia primária
log P	logaritmo do coeficiente de partição octanol-água
mg	miligrama
mL	mililitros
MM	massa molecular
N	número de componentes
NA	número de átomos diferente de hidrogênio
NaCl	cloreto de sódio
NAD ⁺	nicotinamida adenina dinucleotídeo oxidado
NADPH	nicotinamida adenina dinucleótido fosfato reduzido
NCEs	novas substâncias químicas bioativas
ND	não determinado
ng	nanograma
nM	nanomolar
OMS	organização mundial de saúde
P	produto
PA	Para Análise (grau de pureza)
PBS	tampão fosfato salino
PDB	banco de dados de proteínas
PFT	farnesil transferase
PGGT	geranilgeranil transferase
pK_a	logaritmo da constante de associação
PLS	regressão multivariada por mínimos quadrados parciais
PMSF	fluoreto de fenilmetilsulfóxido
Q	carga
q^2	coeficiente de correção da validação cruzada
QSAR	relações quantitativas estrutura atividade
R	constante dos gases
r	distancia entre os átomos
R_{factor}	fator residual
R_{free}	fator residual livre
RMSD	raiz quadrada do desvio quadrático médio
RMSE	raiz quadrada do desvio quadrático médio do erro
RNA	ácido ribonucléico
ROC	curva característica operativa do receptor
rpm	rotações por minuto

S	substrato
S1 a S4	sub-sítios integrantes do sítio catalítico da enzima cruzaina
SAR	relação estrutura atividade
SDS	dodecil sulfato de sódio
SEE	desvio do ajuste linear
SEP	menor erro padrão da predição
SEP	desvio da preditividade do modelo
Sy.x.	Desvio padrão dos resíduos
<i>T. Brucei</i>	<i>Trypanossoma Brucei</i>
<i>T. cruzi</i>	<i>Trypanossoma cruzi</i>
TBVS	planejamento baseado na estrutura do receptor
TO	tripanotiona oxidase
TP _{rate}	verdadeiros positivos recuperados
TR	tripanotiona redutase
UV	ultravioleta
v/v	volume/volume
v _i	velocidade inicial sem inibidor
v _{max}	velocidade máxima
ZFRMCA	N-benzóxicarbonil-L-fenilalanila-L-arginina-7-amido-metil-cumarina (substrato sintético da enzima cruzaina)

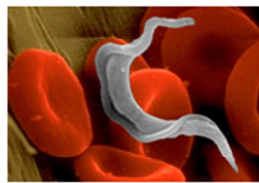
CONTEÚDO

RESUMO.....	i
ABSTRACT	ii
LISTA DE FIGURAS	vii
LISTA DE TABELAS	xiv
CAPÍTULO 1 – INTRODUÇÃO	1
1.1 Doença de Chagas.....	2
1.2 Enzima Cruzaína.....	10
1.3 A Química Medicinal e a gênese planejada de fármacos.....	22
1.3.1 O espaço químico-biológico	28
1.3.2 A Quiminformática na identificação e otimização de ligantes.....	29
CAPÍTULO 2 - OBJETIVOS	31
2.1 Objetivos Gerais.....	31
2.2 Objetivos Específicos	31
CAPÍTULO 3 – ENSAIOS VIRTUAIS	33
3.1 Coleções Virtuais de Moléculas e Filtros de seleção	34
3.2 Docagem Molecular	35
3.2.1 Algoritmos de busca e flexibilidade molecular.....	39
3.2.2 Funções de pontuação.....	41
3.3 Similaridade química	46
3.4 Relações Quantitativas Estrutura Atividade	47
3.5 Materiais.....	48
3.6 Métodos.....	49
3.6.1 Filtros moleculares	49
3.6.2 Docagem Molecular	49
3.6.3 Similaridade química.....	51

3.6.4	Construção da Coleção Multiconformacional.....	52
3.6.5	Relações Quantitativas Estrutura Atividade Baseada em Hogramas Moleculares	52
3.6.6	Avaliação dos modelos computacionais	53
3.6.7	Método de avaliação dos modelos computacionais.....	53
3.7	Resultados e Discussões.....	54
3.8	Filtros Moleculares.....	56
3.9	Docagem Molecular.....	57
3.9.1	Docagem com o programa FlexX	61
3.9.2	Docagem com o programa FRED.....	66
3.10	Similaridade Química.....	70
3.11	Escolha Consensual	74
3.12	Programa Glide.....	76
3.13	Relações Quantitativas Estrutura Atividade Baseado em Hogramas Moleculares	78
3.14	Análise consensual	83
CAPÍTULO 4 – ENSAIOS <i>IN VITRO</i>		89
4.1	Receptores Enzimáticos	90
4.2	Mecanismos Cinéticos de Catálise Enzimática.....	90
4.3	Inibições Enzimáticas	92
4.3.1	Tipos de Inibidores	93
4.3.2	Inibição reversível competitiva.....	93
4.3.3	Inibição reversível não competitiva.....	94
4.3.4	Inibição reversível incompetitiva	95
4.2	Obtenção de enzima por expressão heteróloga	96
4.3	Ensaio Cinético por Fluorimetria.....	97
4.1	Materiais	100
4.2	Métodos	101

4.2.1 Extração de DNA plasmidial.....	101
4.2.2 Transformação da Escherichia coli para expressão da proteína.	101
4.2.3 Pré-Inóculo.....	101
4.2.4 Inóculo e Indução.....	102
4.2.5 Lise das Células.....	102
4.2.6 Dissolução dos corpos de inclusão e renovelamento da enzima	103
4.2.8 Gel de eletroforese (SDS-page).....	103
4.2.9 Ativação e estocagem.....	104
4.2.10 Determinação da curva de Michaelis-Menten para a enzima cruzaina	105
4.2.5 Ensaios de inibição das enzimas cruzaina e catepsina-L para determinação de IC ₅₀	105
4.3 Determinação da constante de inibição aparente (K_i^{app}).....	106
4.4 Ensaio <i>in vitro</i> contra <i>T. cruzi</i> na forma Tripomastigota.....	106
4.6 Resultados e Discussões.....	107
4.7 Determinação da curva de Michaelis-Menten para as enzimas cruzaina e catepsina-L.....	112
4.8 Ensaios de inibição das enzimas cruzaina e catepsina-L.....	113
4.9 Determinação da constante de inibição.....	121
4.10 Relações Estrutura Atividade.....	126
4.11 Ensaio <i>in vitro</i> contra <i>T. cruzi</i> na forma Tripomastigota.....	133
CAPÍTULO 5 - CONCLUSÕES.....	135
PERSPECTIVAS.....	137
REFERÊNCIAS.....	139

CAPÍTULO 1 – INTRODUÇÃO



O Capítulo 1 apresenta uma revisão acerca da doença de Chagas, o alvo biológico estudado a enzima cruzaina e a contribuição da química medicinal como ferramenta para a descoberta de candidatos a fármacos.

1.1 Doença de Chagas

A doença de Chagas, também conhecida como tripanossomíase americana, juntamente com o seu agente causador *Trypanosoma (Schizotrypanum) cruzi*, já existia nas Américas como uma doença enzoótica (isto é, uma doença de animais silvestres) por milhões de anos. Vetores, principalmente dos gêneros *Triatoma*, *Panstrongylus* e *Rhodnius* (*Hemiptera, Reduviidae*), mas também de outros gêneros, transmitiram a doença entre os animais por quase 10 milhões de anos. Mais de 150 espécies de triatomíneos (vetores da doença) e mais de 100 espécies de mamíferos, principalmente as espécies selvagens, mantêm a infecção por *T. cruzi* na natureza[1]. A distribuição dos vetores e reservatórios silvestres de *T. cruzi* nas Américas se estende desde os Estados Unidos à Argentina e ao Chile. No total, 22 países são endêmicos para a doença enzoótica nas Américas. [2]

Estima-se que os seres humanos chegaram à América entre 26.000 e 12.000 anos atrás. As atividades agrícolas e domesticação de animais cerca de 10 mil anos atrás pode ter incentivado a infecção, apesar de há milhares de anos ser principalmente uma antropozoonose (uma infecção acidental do homem). No entanto, o ácido desoxirribonucléico (DNA) de *T. cruzi* foi encontrado em múmias do norte do Chile e sul do Peru, que têm cerca de 9.000 anos, e há evidências de *Triatoma infestans* nos domicílios em tempos pré-colombianos, nomeadamente Inca e das culturas Chinchorro, indicando a introdução gradual de transmissão doméstica. Nos últimos 200-300 anos, com o desmatamento para a criação de agricultura e pecuária, e a abertura de rotas terrestres de transporte (ferrovias e rodovias), os triatomíneos perderam gradualmente sua fonte primária de alimento de sangue animal selvagem e, ao mesmo tempo tiveram mais oportunidades para se espalhar. Eles se adaptaram para as zonas circundantes das habitações humanas, alimentando-se do sangue de animais domésticos e seres humanos. Desta forma, um novo ciclo de infecção foi estabelecida, e a doença de Chagas tornou-se uma zoonose (uma doença que transmite entre animais e seres humanos endemicamente). [1]

A doença de Chagas recebeu esse nome como homenagem ao pesquisador brasileiro e médico sanitário Doutor Carlos Chagas, que em 1909 descreveu toda sua etiologia. Apesar dos esforços ao longo do século de descoberta da doença e da redução da transmissão do parasito, a Organização Mundial de Saúde (OMS) estima

que haja cerca de 10 milhões de pessoas infectadas na América Latina e 25 milhões em risco de adoecer, apenas em 2008 ocorreram mais de 10 mil mortes por complicações desta doença. [3-4]

A doença de Chagas é endêmica na América Latina, entretanto é também encontrada no Canadá, Estados Unidos, Europa (principalmente Espanha e Portugal), Japão e Austrália, FIGURA 1.1. A presença do parasito nestas regiões tem sido atribuída aos imigrantes de regiões endêmicas. [1]

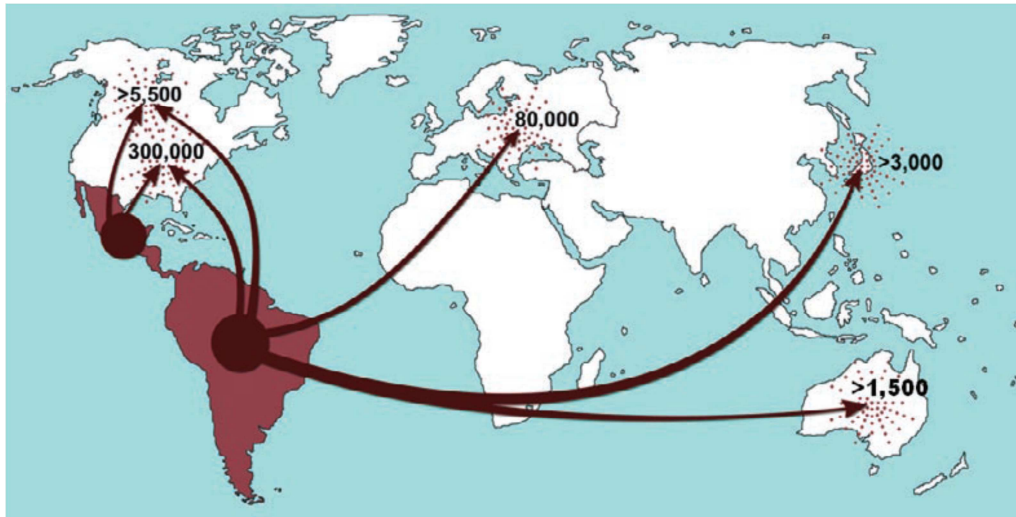


FIGURA 1.1: Distribuição Geográfica da Doença de Chagas [1]

O *T. cruzi* apresenta um complexo ciclo de vida do tipo heteroxênico, passando o parasito por uma fase de multiplicação intracelular no hospedeiro vertebrado (homem e mamíferos) e extracelular no inseto vetor (triatomíneos), FIGURA 1.2. O *T. cruzi* se apresenta em três principais estágios de desenvolvimento, distinguíveis morfologicamente: tripomastigota (eliminada pelo vetor invertebrado) ou sanguíneo (forma transmissível ao homem) e formas replicativas, amastigota e epimastigota. [5]

Considerando o mecanismo natural de infecção pelo *T. cruzi*, os tripomastigotas eliminados nas fezes e urina do vetor, durante ou logo após o hematofagismo, penetram pelo local da picada e causam a contaminação. Em seguida, ocorre a transformação dos tripomastigotas em amastigotas, que aí se multiplicam por divisão binária simples. A seguir ocorre a diferenciação dos amastigotas em tripomastigotas que são liberados da célula. Estes tripomastigotas seguem para a corrente circulatória, atingem outras células de qualquer tecido ou órgão para cumprir novo ciclo celular ou são destruídos por mecanismos

imunológicos do hospedeiro. Podem ainda ser ingeridos por triatomíneos, onde cumprirão seu ciclo extracelular, FIGURA 1.2. [6]

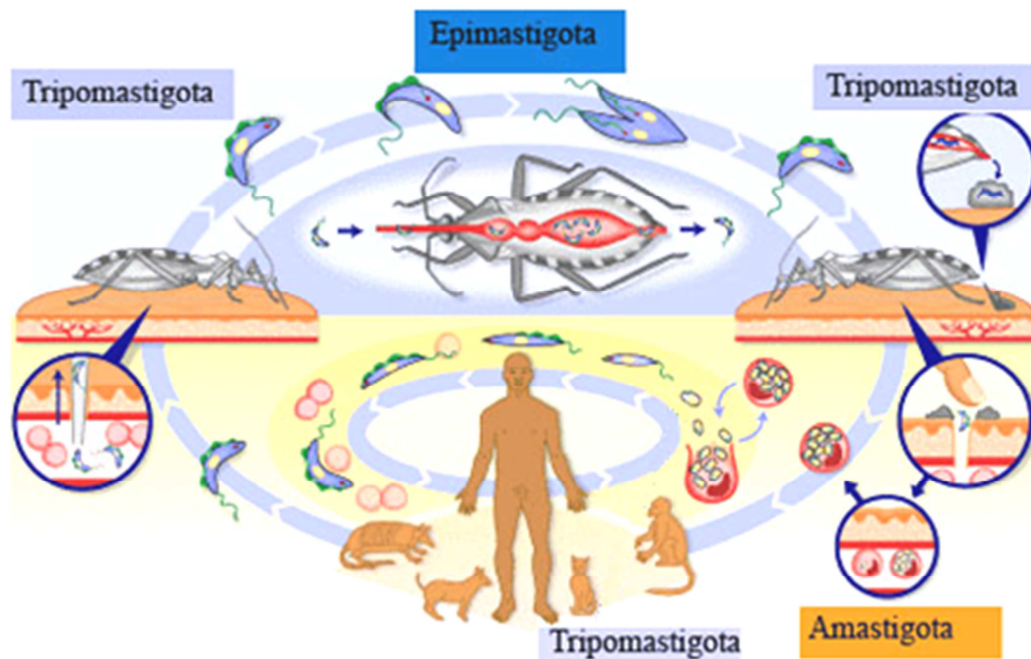


FIGURA 1.2: Ciclo de Vida do *Trypanosoma cruzi* [7]

No estômago do inseto eles se transformam na forma epimastigotas. No intestino médio, os epimastigotas se multiplicam por divisão binária simples, sendo, portanto, responsáveis pela manutenção da infecção no vetor. No reto, porção terminal do tubo digestivo, os epimastigotas se diferenciam em tripomastigotas (infectantes para os vertebrados), sendo eliminados nas fezes ou urina. Esta é a descrição clássica adotada para o ciclo do *T. cruzi*. [6]

A doença de Chagas apresenta uma fase aguda e uma fase crônica. A fase aguda pode ser sintomática ou assintomática (mais freqüente). Ambas estão relacionadas com o estado imunológico do hospedeiro. Há predomínio da forma aguda sintomática na primeira infância, levando à morte em alguns casos devido à meningoencefalite e mais raramente à falência cardíaca. A fase aguda inicia-se através das manifestações locais, quando o *T. cruzi* penetra na conjuntiva (sinal de Romanã) ou na pele (chagoma de inoculação). Estas lesões aparecem em casos agudos dentro de 4-10 dias após a picada do barbeiro, regredindo em um ou dois meses. As manifestações gerais não são exclusivas para a doença e são representadas por febre, edema localizado e generalizado, fadiga, dores no corpo, dor de cabeça, erupção cutânea, perda de apetite, diarreia, vômito e, às vezes, insuficiência cardíaca e perturbações neurológicas. O óbito, quando acontece, é

devido à meningoencefalite aguda ou devido à insuficiência cardíaca consequente à miocardite aguda difusa, uma das mais violentas de que se tem notícia. [8-10]

Após a fase aguda, os pacientes passam por um longo período assintomático (10 a 30 anos), ao fim do qual aparecem as primeiras manifestações clínicas relacionadas a fase crônica. O período assintomático, onde o paciente está contaminado mas não apresenta manifestações clínicas, é chamado de forma indeterminada da doença (latente) e caracterizada por exames sorológicos e parasitológicos, ausência de sintomas da doença, eletrocardiograma convencional normal e coração, esôfago e cólon radiologicamente normais. Cerca de 50 % dos pacientes chagásicos que tiveram a fase aguda pertencem a esta forma. [11]

Certos pacientes chagásicos após permanecerem assintomáticos por anos, com o correr do tempo apresentam sintomatologia relacionada com o sistema cardiocirculatório, digestivo, ou ambos. Isto devido ao fato de mudar inteiramente a fisionomia anatômica do miocárdio e do tubo digestivo (esôfago e cólon, principalmente). Observa-se reativação do processo inflamatório, com dano destes órgãos, nem sempre relacionada com o parasito, que se encontra extremamente escasso nesta fase.

Na cardiopatia chagásica crônica sintomática, a principal manifestação é a insuficiência cardíaca e isto se deve à diminuição da massa muscular que se encontra muito destruída. Além da insuficiência cardíaca são freqüentes os fenômenos tromboembólicos. Os trombos cardíacos são freqüentes (76 % dos casos desenvolvem insuficiência cardíaca), mas também podem se formar nas veias dos membros inferiores. A partir destes trombos, desprendem-se êmbolos que podem originar infartos no coração, pulmões, rins, baço e encéfalo. [12-13]

As manifestações digestivas são representadas principalmente no Brasil e na Argentina pelas megas dilatações dos órgãos do sistema digestório, megaesôfago, megacólon, megaduodeno, megabexiga etc. Os principais sintomas são disfagia, odinofagia, dor retroesternal, regurgitação, pirose, soluço, tosse e sialose. [14]

Os mecanismos de transmissão do parasito são:

Transmissão pelo vetor

Transfusão sanguínea

Transmissão congênita

Acidentes de laboratório

Transmissão oral

Transplante de órgãos

O tratamento contra a doença de Chagas ainda é ineficaz. Existem dois medicamentos na terapêutica atual: o Nifurtimox (Lampit[®]) e o Benzonidazol (Rochagan[®]), FIGURA 1.3, os quais são eficazes para a cura da infecção na fase aguda, com sucesso em até 80% dos casos agudos e nos casos crônicos até 20 %. Ambos os medicamentos exibem fortes efeitos colaterais, tais como anorexia, perda de peso, mudanças psicológicas, excitabilidade, tremores musculares, sonolência, alucinações, náusea, vômito, dor abdominal, diarreia e convulsões [15-16]. O Nifurtimox não é mais utilizado no Brasil devido a baixa suscetibilidade das cepas do parasito presente nos pacientes chagásicos brasileiros. A doença de Chagas é considerada uma das doenças “super-negligenciadas” segundo a organização mundial de saúde, estas são doenças tropicais infecciosas, as quais não possuem tratamentos específicos adequados. A busca de novos fármacos mais seguros e eficazes é de extrema necessidade. [16-17]

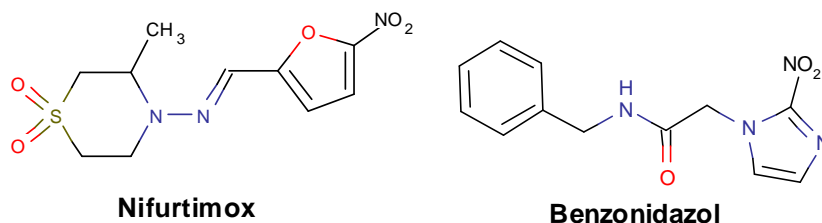


FIGURA 1.3: Representação das estruturas químicas do nifurtimox e benzonidazol

Devido à natureza empírica da descoberta e desenvolvimento do nifurtimox e do Benzonidazol, estes fármacos foram registrados e utilizados clinicamente por vários anos sem um entendimento claro dos seus mecanismos de ação. Hoje, se sabe que a ação antiparasitária do nifurtimox está associada com a produção de grandes quantidades de espécies reativas tóxicas de oxigênio (tais como anions superóxidos e peróxido de hidrogênio), resultantes da reação do oxigênio com metabólitos nitro-anions instáveis produzidos a partir grupo nitro presente na molécula. Entretanto, a citotoxicidade do benzonidazol não está associada com a geração de níveis significantes de espécies reativas tóxicas de oxigênio; ao invés, está provavelmente associada ao estresse redutivo, o qual envolve modificações covalentes de macromoléculas por intermediários de nitro-redução. [18-19]

A busca de novos fármacos mais seguros e eficazes é de extrema necessidade, há vários mecanismos de ação sendo estudados para a identificação

de alvos biomacromoleculares capazes de eliminar o *T. cruzi*. A seguir alguns exemplos:

Metabolismo de esteróis: Compostos que contêm azóis são a maior descoberta na terapia anti-fúngica. Foi demonstrado que os imidazóis e triazóis interferem na biossíntese de esterol e quando os compostos possuem algum heterociclo contendo nitrogênio pertencem à classe de inibidores da biossíntese de ergosterol [20]. O *T. cruzi* contém ergosterol e por esta razão os anti-fúngicos foram testados contra o parasito, os compostos miconazol e econazol apresentaram uma alta potência contra o crescimento do parasito, ao passo que diminuiu a quantidade do 5,7-dieno esterol [21]. Estudos posteriores demonstraram que o cetoconazol e outros potentes azóis antimicóticos foram também ativos na proteção de infecção letal de camundongos com *T. cruzi*, inibindo a multiplicação intra-celular do parasito e bloqueando a biossíntese de esterol. Os imidazóis, como cetoconazol e itraconazol, são inibidores da citocromo P450 dependente lanosterol 14 α -demetilase, a inibição desta enzima resulta no acúmulo do 14 α -metilesterol e diminuição na disponibilidade do ergosterol. Embora esta enzima esteja presente em células de mamíferos, é menos sensível a ação destes compostos comparativamente aos fungos e tripanosomatídeos. [22]

Os compostos que atuam nesta via são os mais avançados no desenvolvimento de um novo fármaco anti-chagásico. O anti-fúngico posaconazol (Merck) está anunciado entrar em fase clínica 2 na Espanha, o composto E12-24 (Esai, Japão) terminou a fase clínica 1 e está planejado a fase clínica 2 na Bolívia com potencial extensão para a Espanha. O composto Tak-187 (Takeda, Japão) foi avaliado em fase clínica 1. [23]

Prenilação de proteínas: foi demonstrada a ocorrência de prenilação de proteínas nos parasitos *T. cruzi* e *T. brucei* [24]. A prenilação de proteínas em mamíferos envolve a inserção de substituintes farnesil ou geranylgeranil em resíduos de aminoácidos cisteínas C-terminais uma série de proteínas celulares. Muitas destas proteínas são pequenas GTPases, que participam na sinalização celular de transdução e tráfego intracelular de vesículas [25-26]. Três enzimas foram identificadas em células eucarióticas incluindo as de mamíferos, plantas e leveduras que inserem substituintes prenil a proteínas, farnesil transferase (PFT), geranylgeranil transferase do tipo I (PGGT-I) e II (PGGT-II) [27]. Vários inibidores potentes da PFT foram descobertos com o objetivo inicial de desenvolver fármacos

para o tratamento de câncer. Alguns destes compostos inibiram uma PFT recombinante de *T. brucei* e *T. cruzi*, e são potenciais agentes tripanosomicidas [25, 28]. Uma etapa preliminar para a prenilação de proteínas é a reação catalisada pela enzima farnesil pirofosfato sintase, foi demonstrado que pode ser inibida por compostos bisfosfonatados análogos do pirofosfato [29]. Vários bisfosfonatos são potentes inibidores da reabsorção óssea e estão em uso clínico para o tratamento e prevenção da osteoporose, doença de Paget e hipercalcemia [30]. Bisfosfonatos apresentaram atividade *in vitro* e *in vivo* contra o *T. cruzi*, sem toxidez às células do hospedeiro [31]. Foi postulado que a atividade tripanosomicida preferencial é devida a acumulação dos bisfosfonatos na organela acidocalcisomos do parasito, a qual pode ser equivalente ao osso mineral. Acidocalcisomos são compartimentos ácidos ricos em cálcio, magnésio, sódio, zinco e polifosfatos de cadeias curtas e longas e estão presentes em vários tripanosomatídeos [32-33]. Milhões de pessoas são tratadas com medicamentos bisfosfonatados aprovados pelas agências reguladoras e por este motivo são atraentes para o uso do desenvolvimento de agentes tripanosomicidas.

Via glicolítica: Foi demonstrado que a forma intracelular amastigota do parasito *T. cruzi* é dependente da via glicolítica para a produção de energia na forma de adenosina trifosfato (ATP), desta forma a enzima gliceraldeído-3-fosfato desidrogenase (GAPDH) foi apontada como um alvo potencial [34]. A GAPDH catalisa a fosforilação oxidativa do gliceraldeído-3-fosfato em 1,3-difosfoglicerato (DPG) na presença de fosfato inorgânico e NAD^+ como coenzima. Vários produtos naturais foram identificados como inibidores da GAPDH, dentre eles flavonóides, cumarinas, ácidos anacárdicos, xantonas e biantronas. [35-36]

Via de salvação de purinas: tripanosomatídeos dependem exclusivamente da via de salvação de purinas, ao passo que mamíferos, além desta via, utilizam também a biossíntese *de novo* de bases purínicas. A enzima hipoxantina guanina fosforibosiltransferase (HGPRT) é responsável pela conversão das bases purínicas a ribonucleotídeos no *T. cruzi*, esta enzima pode também iniciar o metabolismo de alguns análogos das bases purínicas como o Allopurinol. Análogos de purinas foram testados contra HGPRT de *T. cruzi* e humanos e algumas mostraram afinidade pela enzima do parasito. [37]

Via de síntese da Tripanotiona: esta via é importante para a redução da quantidade de radicais livres e conseqüentemente para a manutenção do ambiente

intracelular do parasito. Duas enzimas são essenciais nesta via, a tripanotona oxidase (TO) e redutase (TR). A tripanotona redutase é uma flavoproteína dependente de NADPH responsável pela redução da tripanotona, esta enzima é alvo de estudos para o desenvolvimento racional de compostos contra o *T. cruzi* e leishmanioses, uma vez que foi demonstrado que a enzima de mamíferos difere da parasitária através de um estudo de especificidade pelo substrato [38]. Diferentes classes de compostos foram identificadas como inibidores da TR, como nitrofuranos, compostos triciclos derivados de tioridazina e clomipramina (antidepressivos), difenilsulfetos, poliaminas, alcalóides bis-benzil quinolínicos e complexos de platina II. [39-41]

DNA Topoisomerases: enzimas que alteram a topologia do DNA são denominadas DNA Topoisomerases. O gene codificante da DNA Topoisomerase do tipo II em *T. cruzi* foi isolado e a comparação da sequência de aminoácidos com as enzimas correspondentes em *T. brucei* e *Crithidia fasciculata* mostraram um alto grau de conservação [42]. Vários inibidores da DNA Topoisomerase II de bactérias foram ativos contra a enzima do *T. cruzi*, inibindo a forma proliferativa e do processo de diferenciação, causando danos ao Cinetoplasto e o Núcleo da forma epimastigota e sugerindo que ambas as organelas podem ser alvos para o desenvolvimento de fármacos. [43-44]

Via de síntese de novo de pirimidinas: A síntese de nucleotídeos é essencial para o desenvolvimento de todos e quaisquer organismos. Além de participarem como unidades fundamentais dos ácidos nucléicos, DNA e RNA, os nucleotídeos atuam também como ativadores e inibidores de várias vias do metabolismo intermediário das células.

Nos últimos anos a enzima diidroorotato desidrogenase (DHODH), que catalisa a quarta etapa da síntese de novo de nucleotídeos de pirimidina, vem sendo estudada como um alvo promissor para o desenvolvimento de novos agentes terapêuticos contra parasitoses causadas por tripanossomatídeos. Experimentos recentes visando a validação da DHODH do *T. cruzi* mostraram que as epimastigotas, forma replicativa do parasito no inseto vetor, são totalmente dependentes da síntese de novo de pirimidinas para sobrevivência. [45-46]

Apesar do potencial da enzima DHODH como alvo terapêutico, apenas a enzima de *Plasmodium falciparum* (PfDHODH) e de *T. cruzi* (*TcDHODH*) têm sido objeto de estudos para o planejamento de inibidores capazes intervirem na via

metabólica de síntese de novo de nucleotídeos de pirimidinas e atuarem como agentes quimioterápicos contra a Malária e doença de Chagas

Via de digestão proteolítica: O *T. cruzi* possui uma enzima cisteíno protease responsável pela maior parte a atividade proteolítica, responsável pela nutrição do parasito, presente em todos os estágios do ciclo de vida do parasito chamada cruzipaína (cruzaína) [47-48]. Esta enzima foi o alvo de estudo deste projeto e será discutida com mais detalhes na seção 1.2.

Apesar dos esforços na erradicação da ploriferação do *T. cruzi* e dos esforços para a identificação de vias essenciais para a sobrevivência do parasito e descoberta de agentes tripanossomicidas, esta doença ainda assola milhões de pessoas na América Latina e grande parte da população deste continente está em áreas consideradas de risco de contaminação. Apesar de alguns compostos estarem atualmente fase clínica, não há registro de nenhum fármaco novo desde a introdução do Benzonidazol e Nifurtimox para o tratamento da doença. Diante deste cenário, este trabalho visou contribuir para a quimioterapia da doença de Chagas através da busca de compostos bioativos capazes de inibir a enzima cruzaina e erradicar o parasito *T. cruzi*.

1.2 Enzima Cruzaína

Proteases são enzimas chaves na patogenicidade de muitas doenças parasitárias, são alvos atraentes porque estão envolvidas na sobrevivência e replicação dos parasitos, elas induzem o dano no tecido do hospedeiro e facilitam a invasão do parasito ou auxiliam na utilização dos metabólitos de proteínas do hospedeiro. A elucidação do comportamento biológico das proteases e aquisição do conhecimento completo da bioquímica e estrutura são requisitos para o desígnio de fármacos antiparasitários que bloqueiem especificamente a ação proteolítica. [49]

Vários grupos de pesquisas independentes convergiram em uma protease tiol-dependente, no *T. cruzi*, como uma enzima metabólica chave [50-51] e alvo de resposta imune do hospedeiro [52]. A protease foi purificada inicialmente por e foi imunologicamente identificada nas três formas de vida do *T. cruzi*, epimastigotas, tripomastigotas e amastigotas. [53-54]

Análise da sequência de aminoácidos da proteína purificada revelou que o provável sítio ativo é correspondente ao da papaína (Cys25) [55]. Posteriormente, descobriu-se que a glicoproteína GP57/51 que tinha sido descrita previamente como o principal antígeno de *T. cruzi* em humanos infectados [56], era o mesmo da cisteína protease [52]. Foi sugerida a participação direta da protease no dano do tecido hospedeiro através de secreção da enzima dos parasitos que podem facilitar a ruptura de células ou incidentalmente por vazamento da protease após a lise e morte do parasito, desta forma há o estímulo à resposta imune do hospedeiro. [52]

A sequência completa do gene desta cisteína protease de *T. cruzi* organiza-se em tandem, repedindo seis vezes no genoma do parasito. A expressão heteróloga da enzima codificada por este gene em bactéria, caracterização cinética da cisteína protease purificada e sequência de aminoácidos classificou-a como membro da superfamília da papaína e o nome cruzipapaína foi sugerido. [52, 55, 57].

A cruzipapaína catalisa reações de hidrólise de ligações peptídicas, FIGURA 1.4, e seu número de classificação segundo a comissão de nomenclatura de enzimas (*Enzyme Commission*) é EC 3.4.22.51. Os números correspondem a: 3 - classe das hidrolases, 4 - atua em ligação peptídica, 22 - cisteína endopeptidase e 51 seu número específico.



FIGURA 1.4: Esquema de reação de clivagem de ligação peptídica por proteases

As proteases têm um papel importante em vários processos patológicos, pois estão implicadas em doenças tais como artrite reumatóide, doenças cardiovasculares, infecções bacterianas e virais, câncer e doença de Alzheimer. [58]

Todas as cisteína proteases têm em comum um sítio ativo que consiste dos resíduos de aminoácidos cisteína, histidina e glutamina. A cisteína e a histidina formam um par iônico que é estabilizado pela glutamina, há uma cavidade nesta região das cisteína proteases capaz de estabilizar intermediários aniônicos próximos ao resíduo de glutamina e por este motivo é chamado de cavidade do oxiânion (do inglês *oxyanion hole*).

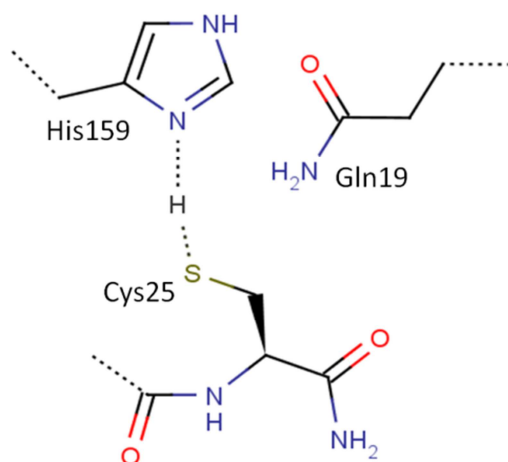


FIGURA 1.5: A cisteína doa o próton para a histidina formando um par iônico correspondente a enzima ativada [59-60]

O mecanismo de hidrólise de peptídeos desempenhado pelas cisteíno proteases envolve quatro etapas. A primeira é a desprotonação do grupo tiol da cisteína catalítica pela histidina adjacente. No próximo passo o enxofre aniônico nucleofílico se liga ao carbono carbonílico do substrato formando um intermediário tetraédrico, o qual é estabilizado pela cavidade do oxiânion. Em seguida o fragmento amino terminal do substrato é liberado e a histidina é restituída na sua forma desprotonada. Na sequência o intermediário acil-enzima é hidrolisado e um segundo intermediário tetraédrico é formado, novamente há a estabilização pela cavidade do oxiânion. Por fim ocorre o substituinte N-terminal é liberado restituindo a enzima no estado inicial. [59-61]

Um esquema representativo da sequência das etapas do mecanismo da reação é mostrado na FIGURA 1.6.

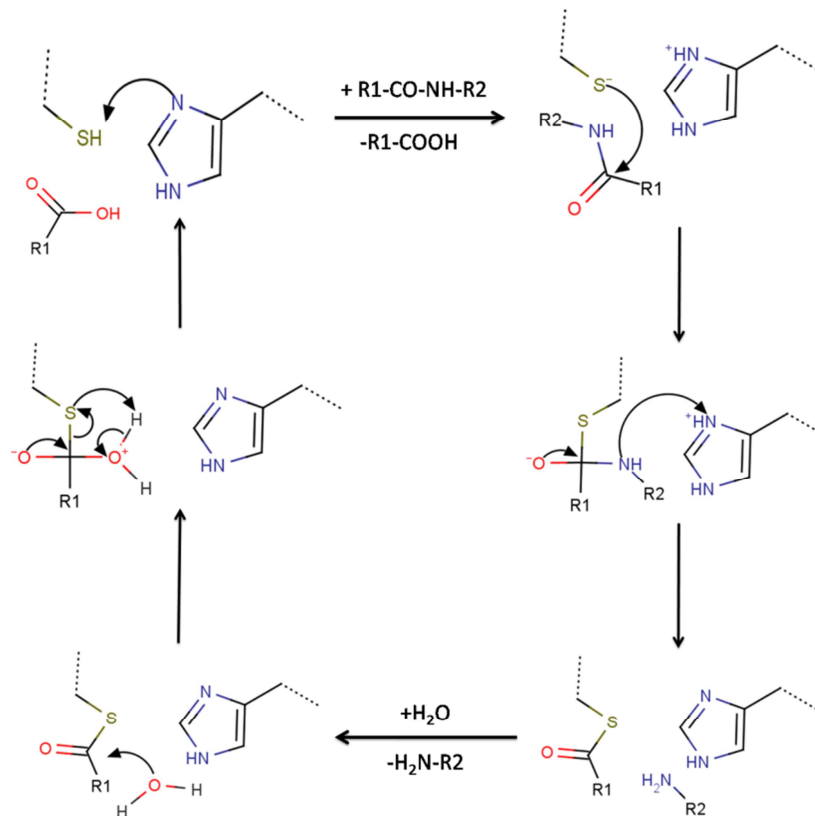


FIGURA 1.6: Mecanismo de catálise de enzimas cisteíno proteases [59-60]

A enzima cruzipaína expressa no sistema de expressão das bactérias passa por eventos proteolíticos e por isso foi desenvolvido uma enzima recombinante com a exclusão do domínio C-terminal com aproximadamente 130 resíduos de aminoácidos, cerca de 25 % da cruzipaína. A enzima recombinante foi denominada cruzaina e corresponde ao domínio catalítico (enzima “madura”) [57]. Foi desenvolvido um protocolo de cristalização da enzima [62], a resolução da estrutura da enzima revelou a cadeia polipeptídica com 215 resíduos de aminoácidos enovelados em dois domínios, um α -helicoidal e outro consistindo de extensivas interações de folhas- β . Os resíduos da tríade catalítica foram identificados Cys25, His159 e Gln19 (similar a papaína) e um extenso sítio de ligação do substrato foi encontrado em uma cavidade entre os dois domínios [63]. Uma representação da estrutura cristalográfica é mostrada na FIGURA 1.7.

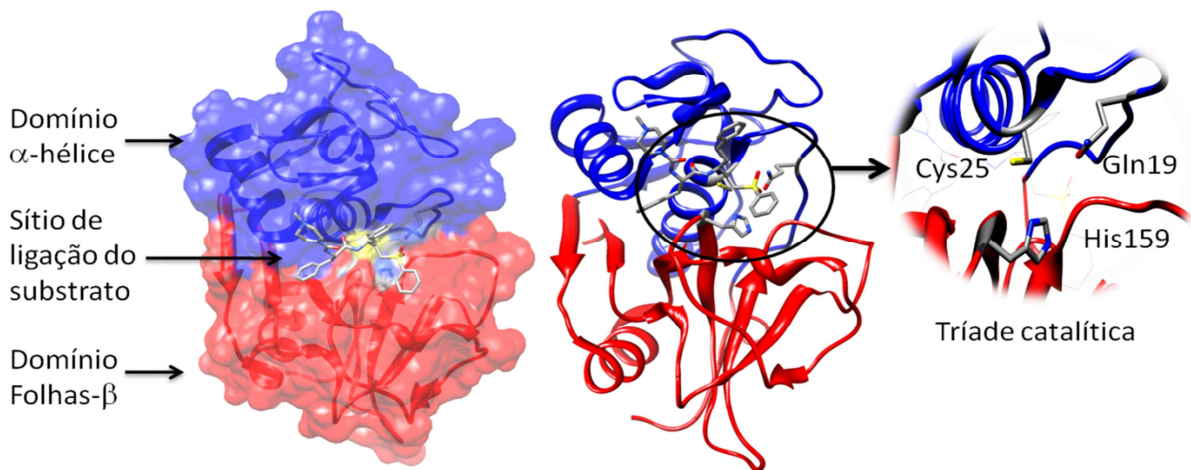


FIGURA 1.7: Estrutura tridimensional da cruzaina (código PDB 2AIM)

A cruzaina pertence à família de enzimas proteases e possui um alto grau de identidade com as enzimas humanas catepsina-L (EC 3.4.22.15) e B (EC 3.4.22.1) [64]. As catepsinas estão envolvidas em muitos processos fisiológicos como na reabsorção óssea, liberação de hormônios da tireoide e um papel fundamental no sistema imunológico. Elas também estão implicadas em um número de processos degenerativos que incluem a osteoporose, artrite reumatóide, enfisema e distrofia muscular. Há também forte evidência da participação das catepsinas no câncer por atuação sobre os processos de invasão de tumores e metástases. As catepsinas, em especial a catepsina-L, são alvos atrativos para o desenvolvimento de agentes anticânceres. [65]

A terminologia descrita por Schechter & Berger [66] para proteases foi adotada para a cruzaina, denominando sub-regiões do sítio ativo de S1 a S4, as posições dos sub-sítios estão mostradas na FIGURA 1.8.

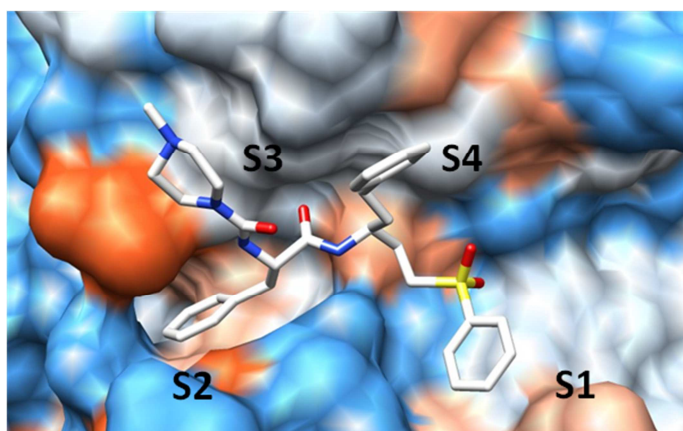


FIGURA 1.8: Estrutura do sítio ativo representada com a superfície de Van der Waals, com a identificação das posições dos sub-sítios S1 a S4 (código PDB 3HD3).

O sub-sítio S1 é onde se localiza a tríade catalítica da enzima, o sub-sítio S2 possui uma característica hidrofóbica localizado em uma região mais profunda, os sub-sítios S3 e S4 são mais superficiais e possuem também característica hidrofóbica. Apenas o centro do sítio possui característica hidrofílica, onde há o reconhecimento da ligação amídica.

Utilizando a estrutura tridimensional da enzima foram planejados os primeiros ligantes através do planejamento baseado na estrutura do receptor (TBVS). O programa de docagem molecular DOCK identificou uma classe de bis-arilacilhidrazidas, um derivado do ligante identificado apresentou alta potência contra a cruzaina ($IC_{50} \sim 600$ nM) [67], FIGURA 1.9. Antes deste trabalho os inibidores de cruzaina identificados eram os mesmos descritos previamente em outras proteases pertencentes à superfamília da papaína.

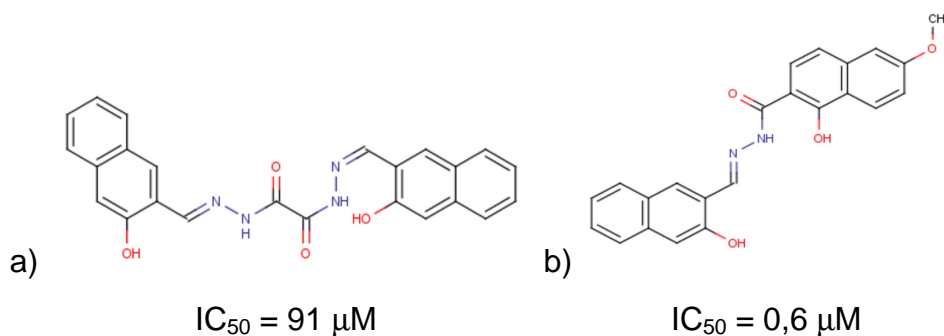


FIGURA 1.9. Estruturas das bis-arilacil-hidrazidas identificadas por docagem molecular utilizando DOCK e derivado com alta potência frente à cruzaina

A partir do protocolo de cristalização da cruzaina foram obtidas estruturas cristalográficas demonstrando os substituintes determinantes para a especificidade da interação. O resíduo Glu205 se ajusta para reestruturar parte do sítio ativo, denominado sítio S2, conferindo uma interação forte com inibidores básicos ou hidrofóbicos neste mesmo sítio de interação. A FIGURA 1.10 mostra a mudança conformacional no resíduo Glu205 que confere a especificidade de interação do sítio S2. [68]

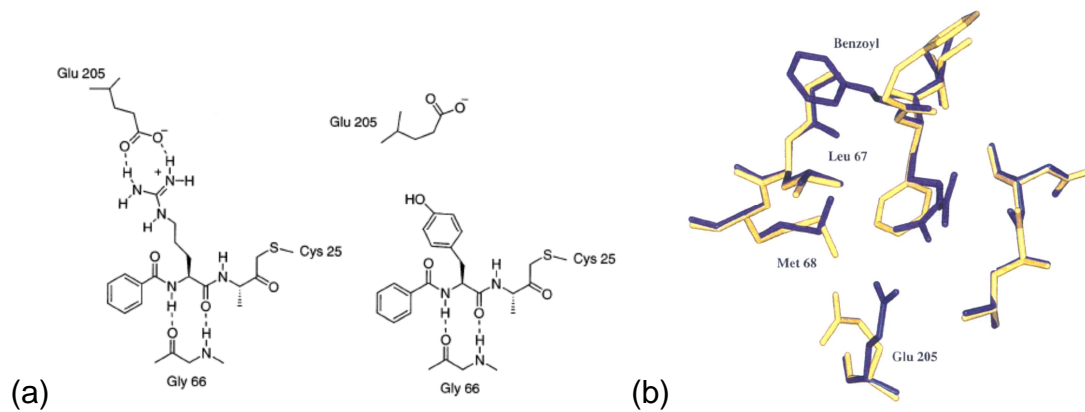


FIGURA 1.10: Estruturas (a) bidimensional e (b) tridimensional da cruzaina complexada com os inibidores Bz-Arg-Ala-CH₂F e Bz-Tyr-Ala-CH₂F

As funções da enzima cruzaina implicam em vários processos celulares como invasão da célula do hospedeiro e nutrição do parasito [69]. É também reconhecida como um antígeno imunodominante durante a infecção em humanos com uma possível participação da defesa do parasito durante a infecção. A enzima possui participação no processo de diferenciação entre os estágios de vida do parasito no ciclo de vida [56, 70]. Estudos demonstraram que a cruzaina é capaz de iniciar o programa de morte celular pela ativação direta de caspases [71]. A inibição seletiva desta enzima bloqueia a ploriferação das formas extracelular epimastigotas e intracelular amastigotas, indicando um papel fundamental na sobrevivência, diferenciação e crescimento do parasito. [72-73]

A enzima cruzaina transita na via de Golgi endosomal, localizada nos lisossomos (ou reservosomos) [74]. Tratamento destas culturas com inibidores de cruzaina alteram a estrutura e função do complexo de Golgi do parasito, resultando em sua morte. [75]

Um modelo animal da doença de Chagas em camundongos infectados foi tratado com inibidores projetados para inativar a enzima cruzaina. Os inibidores peptídios miméticos fluorometil cetonas evitaram a infecção letal dos animais. A melhor estrutura molecular continha a fenilalanina e homofenilalanina no esqueleto. Para obter a cura da infecção, foi inserido no esqueleto peptídico mimético um substituinte vinil sulfônico menos tóxico, resultando no N-metil piperazina-Phe-homoPhe-vinil sulfona (K777 ou K11777) [72]. A estrutura é mostrada na FIGURA 1.11(a).

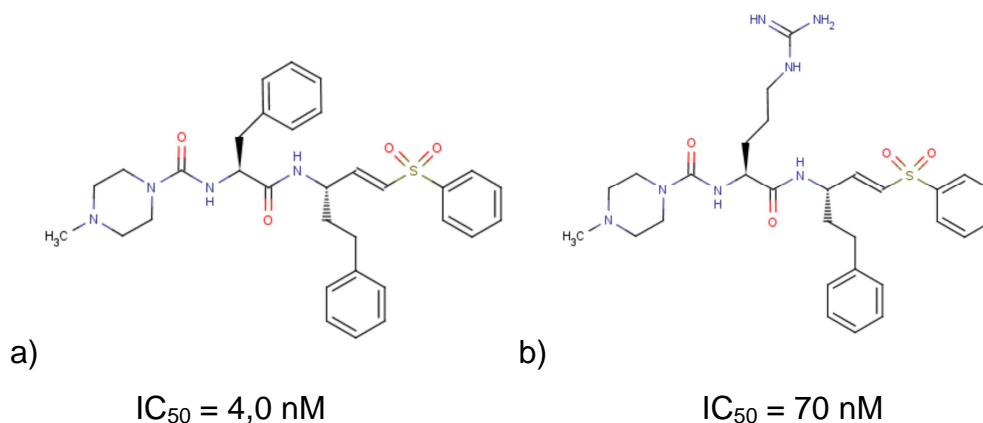


FIGURA 1.11: Representação da estrutura molecular do composto a) K11777 e seu análogo b) WRR-483

O K11777 apresentou boa eficácia no modelo animal da doença de Chagas em camundongos tanto em fase aguda como crônica com mínima toxidez. Estes estudos demonstraram pela primeira vez que a inibição da enzima cruzaina efetivamente interfere na infecção pelo parasito *T. cruzi* e os inibidores peptídicos vinil sulfônicos não são tóxicos em doses terapêuticas, colocando em questão a falta de seletividade deste inibidor em relação às catepsinas humanas. [76]

Um estudo sucessivo foi realizado em modelo animal da doença de Chagas em caninos, embora o tratamento com o K11777 não foi capaz de curar totalmente os animais infectados, houve uma redução significativa da infecção, reduzindo os danos no coração [77]. Posteriormente, este composto foi utilizado para o tratamento de um modelo animal imunodeficiente, uma vez que transplantes de coração são necessários em pacientes com doença de Chagas na fase crônica avançada que apresentam problemas cardíacos, demonstrando que o K11777 é capaz de curar a doença de camundongos imunodeficientes infectados [78]. Estes estudos contribuíram para a classe dos compostos vinil-sulfônicos emergirem como potenciais agentes para o tratamento de doenças parasitárias. [79]

Recentemente foi sintetizado um análogo do K11777 denominado WRR-483, FIGURA 1.11(b) o substituinte fenilalanina responsável pela interação no sub-sítio S2 foi trocado por um substituinte arginina baseado no trabalho de Guilmore ET AL. Apesar de o composto ser menos potente no ensaio enzimático em relação ao K11777 foi superior em ensaios celulares *in vitro* em culturas do *T. cruzi* e *in vivo* nos camundongos. [80]

Desde a validação da principal protease do *T. cruzi* [81] como um alvo promissor para o desenvolvimento de agentes tripanosomicidas houve intensificação

nos trabalhos que buscaram por novas classes de compostos e contribuíram para o descobrimento de vários inibidores da enzima cruzaina.

Uma entre as primeiras classes descobertas como inibidor da cruzaina foi a 1,3-epoxi cetona, o grupo epóxi age como grupo reativo à cisteína. Os compostos foram projetados a partir de um inibidor epóxido conhecido de proteases (E-64) e cetonas peptídicas (Cbz-Phe-Ala-FMK) [82]. Utilizando uma estratégia semelhante e guiada por métodos de modelagem molecular baseado na estrutura do receptor, as classes epoxisucinil e α,β -epóxiésteres dipeptídicos foram projetadas, as estruturas químicas dos compostos mais potentes são mostradas na FIGURA 1.12. [83]

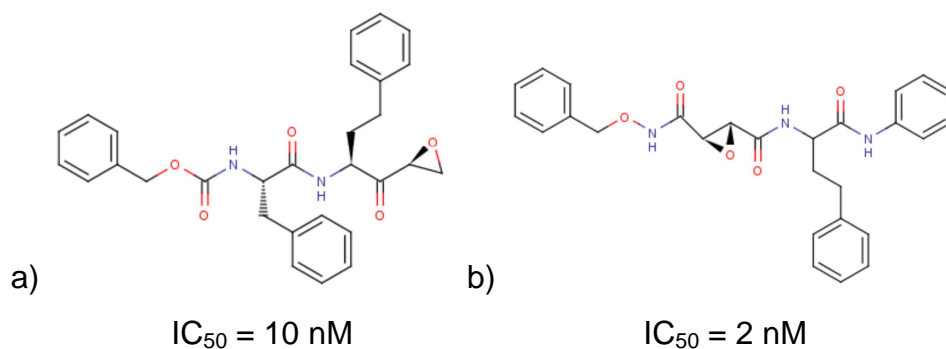


FIGURA 1.12: Estrutura química de dois potentes inibidores da cruzaina das classes a) 1,3-epoxi cetonas e b) epóxisucinil como grupos reativos a cisteína

Inibidores não covalentes de proteases utilizados na terapêutica para o tratamento da hipertensão e AIDS, atuando nas enzimas conversoras de angiotensina e HIV protease, respectivamente, inspiraram a busca por inibidores não covalentes para a cruzaina [84-85]. Baseado na relação bioisostérica entre ligações peptídicas e uréia, os primeiros inibidores não peptídicos reversíveis foram identificados como aril uréias e tiouréias, apresentando moderada potência contra a cruzaina e as formas amastigotas em culturas de *T. cruzi*. [86]

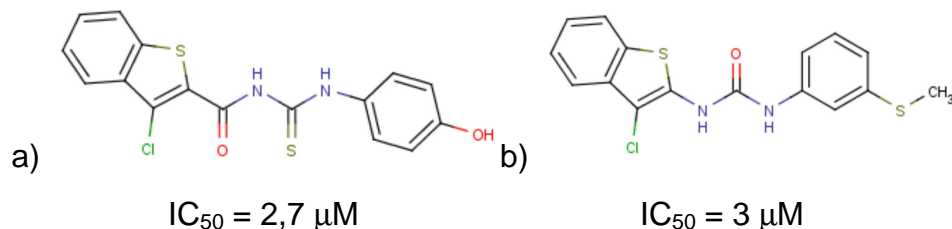


FIGURA 1.13: Estruturas químicas dos primeiros inibidores de cruzaina não peptídicos identificados, classes a) aril tiouréia e b) aril uréia

Inibidores mais potentes da enzima cruzaina foram obtidos através de uma coleção de compostos baseados em cetonas, a qual foi testada e otimizada contra a enzima e potentes inibidores, K_i entre 0,9 a 10 nM, foram identificados. Um

composto apresentou seletividade de 1000 e 100 vezes frente às catepsinas B e L, respectivamente, FIGURA 1.14.

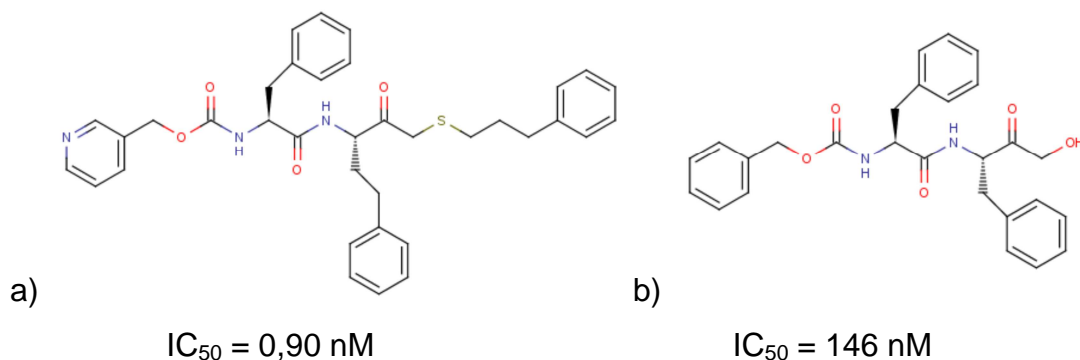


FIGURA 1.14: Estruturas químicas dos compostos baseados em cetonas a) estrutura do composto com melhor seletividade frente as catepsinas B e L e b) composto co-cristalizado com a cruzaina (código PDB 1ME4)

Dois compostos baseados em cetonas foram co-cristalizados com a cruzaina e estruturas de alta resolução foram obtidas, o modo de interação destes inibidores foram determinados e identificaram-se interações fundamentais para o reconhecimento molecular em regiões específicas no sítio da enzima. [87]

A partir destas informações, novas séries de ceto amidas, ceto ácidos, ceto éster e aldeídos foram desenvolvidos. Utilizando métodos de planejamento baseado na estrutura da enzima estes compostos foram otimizados e um total de 20 compostos exibiu potência picomolar nos ensaios de inibição enzimática, as estruturas representativas destas classes são apresentadas na FIGURA 1.15.

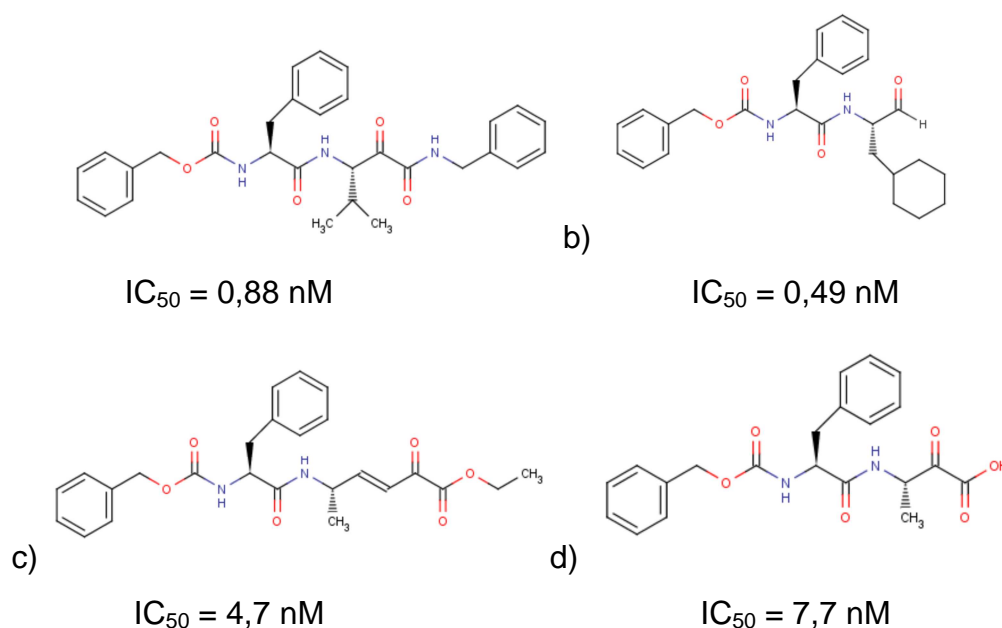


FIGURA 1.15: Estruturas químicas representativas das classes a) Ceto amida, b) aldeído, c) ceto Ester e d) ceto ácido, potentes inibidores na cruzaina

Os três inibidores mais potentes representando diferentes esqueletos moleculares foram testados em cultura de células de macrófagos infectados com *T. cruzi*. Os compostos apresentaram uma inibição dose-dependente da atividade tripanosomicida. O composto ceto amida apresentou melhor atividade que cetoéster e cetoácido, mas todos moderadamente potentes. A baixa potência no teste em macrófagos comparativamente aos ensaios enzimáticos foi atribuída a pobre permeabilidade a membrana do macrófago. Além disso, foi observado que a ligação éster de um composto cetoéster é rapidamente metabolizada *in vivo*. Portanto, apesar da alta potência apresentada por estes compostos, as pobres características físicas impediram os compostos avançarem nos estudos pré-clínicos. [88]

A estratégia utilizada para o desenvolvimento de inibidores da cruzaina baseado em peptídeos resultou na descoberta de vários compostos potentes contra a cruzaina, mas devido à suscetibilidade a hidrólise, metabolismo e baixa biodisponibilidade da maioria dos compostos há necessidade da busca de novos compostos não peptídicos.

A síntese e avaliação de isatinas e derivados de tiosemicarbazona demonstrou o potencial destas classes em inibir a cruzaina apesar de apresentarem somente afinidade moderada contra a enzima. [89]

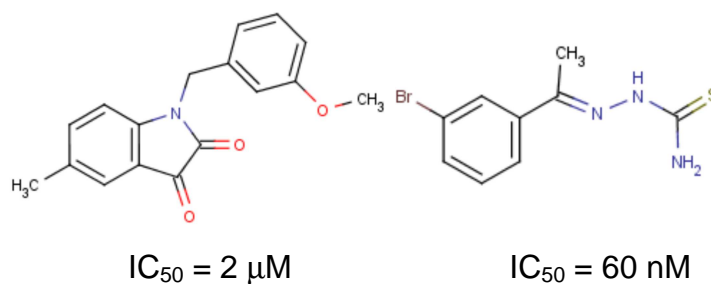


FIGURA 1.16: Estruturas químicas representativas das classes a) isatinas e b) tiosemicarbazonas

A síntese de uma série de tiosemicarbazonas e a determinação de relações estrutura atividade resultou na otimização da potência desta classe, os compostos identificados foram eficazes contra a cruzaina, apresentaram atividade tripanosomicida com potência moderada e baixa toxicidade em camundongos. Alguns compostos inativos frente à enzima apresentaram atividade nos ensaios tripanosomicidas e também outros protozoários, sugerindo a interação com alvos distintos das cisteíno proteases. [90-92]

Os trabalhos mais recentes na literatura com estruturas não peptídicas são vinilsulfonas não peptídicas [93], chalconas e hidrazonas [94], tiazol-hidrazonas [95], nitrofurazona e hidroximetilfurazona [96]. Algumas estruturas representativas destas classes são apresentadas na FIGURA 1.17.

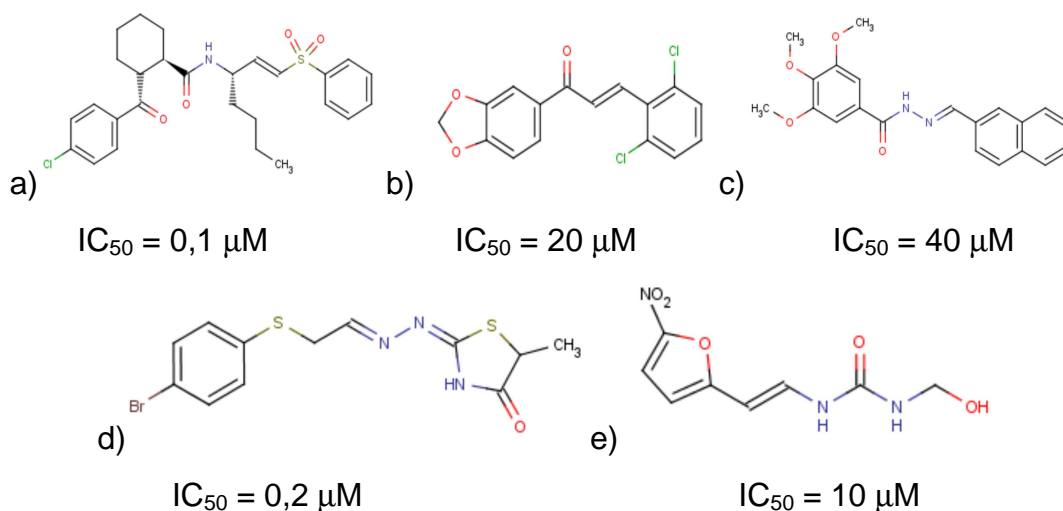


FIGURA 1.17: Estruturas químicas representativas das classes a) vinilsulfonas não peptídicas, b) chalcona, c) tiosemicarbazonas, d) hidrazida e e) hidroximetilfurazona

Os compostos mais potentes e promissores não peptídicos descobertos foram derivados de substratos triazóis, um potente inibidor irreversível 2,3,5,6-tetrafluorofenóximetil cetona foi identificado, o qual foi eficiente na eliminação de infecções do *T. cruzi* em culturas de células e mostrou estabilidade em plasma [97]. Os ensaios em camundongos foram promissores, foi observada baixa toxicidade e boa eficácia na eliminação da infecção pelo *T. cruzi*, doses diárias de 20 mg kg^{-1} foram necessárias duas vezes ao dia para obter resultados semelhantes ao composto vinil sulfônico K11777 com doses de 50 mg kg^{-1} . Entretanto, todos os camundongos tratados apresentaram inflamação no músculo cardíaco sugerindo a necessidade do desenvolvimento de inibidores mais potentes. Um trabalho de otimização foi realizado e obteu-se uma série de compostos promissores que estão sob investigação de propriedades farmacocinéticas em modelos animais [98]. As estruturas dos 2,3,5,6-tetrafluorofenóximetil cetona e do esqueleto molecular mais promissor para o desenvolvimento de um agente tripanosomicida desta classe são apresentados na FIGURA 1.18.

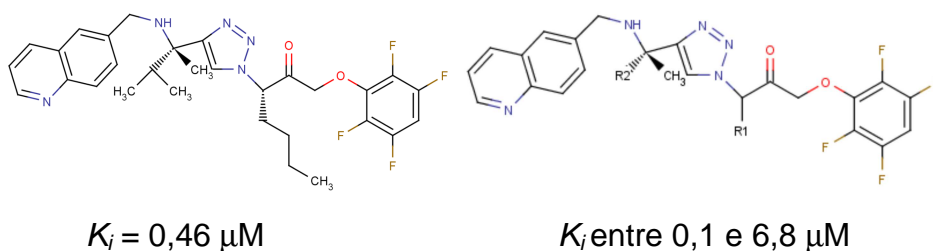


FIGURA 1.18: Estruturas químicas representativas do composto a) 2,3,5,6-tetrafluorofenóximetil cetona e b) esqueleto molecular utilizado para SAR

A enzima cruzaina é um excelente alvo promissor para o desenvolvimento de uma agente tripanosomicida, a validação desta enzima estimulou o desenvolvimento de vários inibidores ao longo dos últimos vinte anos. A descoberta de novos compostos, provenientes de classes químicas distintas, mais seguros e eficazes representa uma importante contribuição para o desenvolvimento da quimioterapia da doença de Chagas.

1.3 A Química Medicinal e a gênese planejada de fármacos

A definição oficial de Química Medicinal dada por uma comissão especializada da IUPAC é: [99]

"A Química Medicinal envolve a descoberta, o desenvolvimento e a identificação e interpretação do modo de ação de compostos biologicamente ativos ao nível molecular. A ênfase é colocada em fármacos, mas o interesse da Química Medicinal também está voltado para o estudo, identificação e síntese de seus produtos metabólicos bem como dos compostos relacionados."

Montanari & Pili deram uma definição mais ampla: "A Química Medicinal é uma ciência baseada na Química que engloba inovação; descoberta e desenvolvimento de novas substâncias químicas bioativas (NCEs); síntese ou modificação molecular; extração, isolamento, identificação e elucidação estrutural de princípios ativos naturais de plantas, animais ou minerais; descrição das moléculas desde a sua constituição atômica (passando por relações entre a estrutura e propriedades) até suas características estruturais quando da(s) interação(ões) com alvos biológicos de interesse terapêutico; compreensão, a nível molecular, de processos bioquímicos/farmacológicos, toxicológicos e farmacocinéticos e a criação de relações entre estrutura química e atividade farmacológica (SARs)". [100]

Desta forma, a Química Medicinal possui uma característica interdisciplinar e cobre um amplo domínio situado na interface da química orgânica com as ciências da vida, tais como a bioquímica, farmacologia, biologia molecular, imunologia, farmacocinética e a toxicologia, e também disciplinas baseadas na físicoquímica, como cristalografia, calorimetria, espectroscopia e métodos de modelagem molecular computacional. [101]

O processo de descoberta e desenvolvimento de novos fármacos é dividido em duas fases principais, como mostrado na FIGURA 1.19:: Fase pré-clínica (descoberta) e Fase clínica (desenvolvimento).

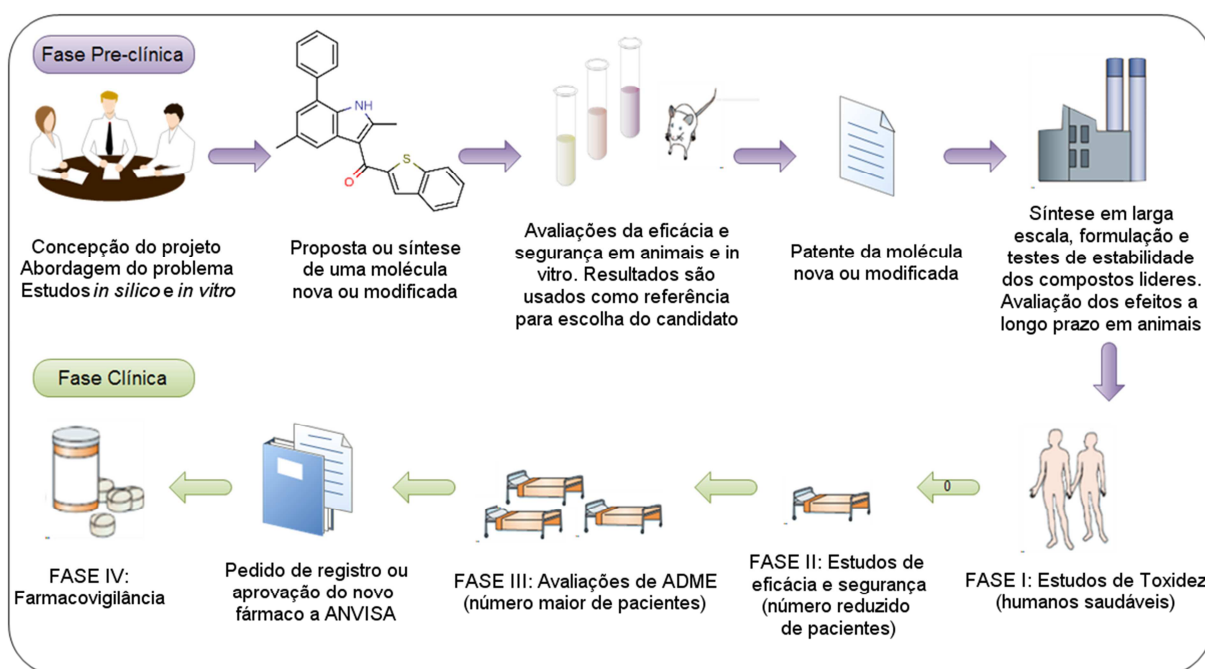


FIGURA 1.19: Etapas do processo de descoberta e desenvolvimento de novos fármacos [102]

A Fase pré-clínica inicia com identificação da necessidade de uma nova terapia e o julgamento sobre a adequação dos tratamentos existentes. Em seguida é necessário identificar e validar o alvo terapêutico, os mais importantes e estudados são enzimas, DNA, receptores hormonais, receptores nucleares e células. O número de alvos na terapêutica é 348 [103], a FIGURA 1.20 mostra uma distribuição dos alvos com medicamento aprovado na terapia [104], estabelecer modelos para estes alvos é essencial para o desenvolvimento do projeto.

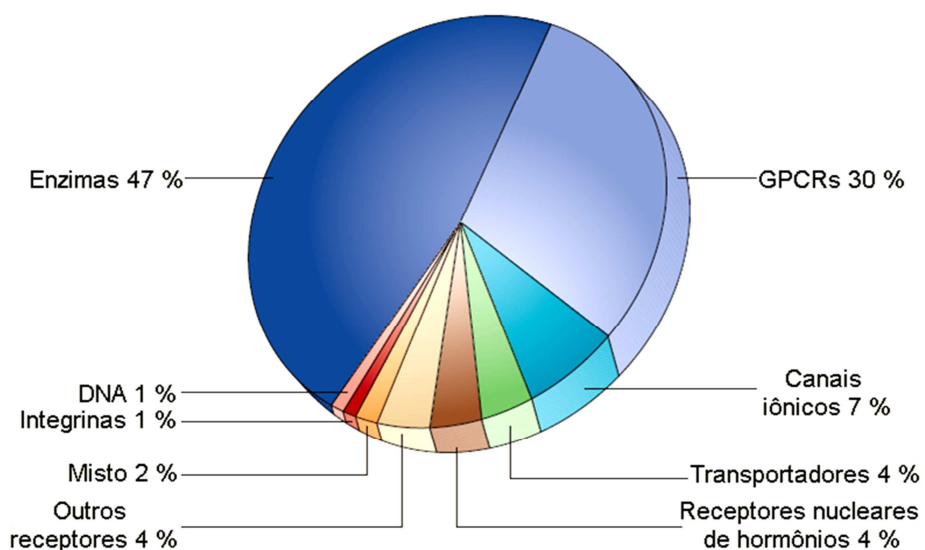


FIGURA 1.20: Alvos dos fármacos na terapêutica distribuídos por classes bioquímicas [104]

Os passos posteriores incluem a identificação de substâncias químicas que apresentam atividade biológica *in vitro*, chamados de ligantes, os quais precisam ser validados e otimizados para os testes nos modelos *in vivo* apropriados. Em seguida, é feita a síntese de estruturas análogas para a maximização das atividades e finalmente a seleção de um ou mais compostos, chamados compostos líderes, como candidatos para o desenvolvimento do fármaco. Estes candidatos passam por avaliações toxicológicas em animais e se forem aprovados em todos os testes podem ser submetidos à avaliação em uma agência reguladora que autorizará (ou não) os estudos em fases clínicas. Os testes clínicos necessitam de grandes quantidades da substância e normalmente há a necessidade da síntese em larga escala dos candidatos a fármacos para possibilitar os estudos de formulações, avaliações sobre o uso em longo prazo em animais e os estudos das fases clínicas. [102, 105]

Na fase clínica, a proposta é o aperfeiçoamento das propriedades farmacocinéticas e o ajuste fino das propriedades farmacêuticas (formulação química) das substâncias selecionadas para que estas se tornem aptas ao uso clínico. Para que isto seja possível são feitas avaliações sequenciais da tolerância em voluntários humanos saudáveis (Fase I), da eficácia e segurança em pacientes (Fase II) seguidas de uma ampla avaliação em milhares de pacientes para a compilação de dados acerca da eficácia e de propriedades ADME (Fase III). Para os poucos candidatos a fármacos que sobrevivem esta série de testes, todos os dados acumulados durante a pesquisa são submetidos à agência reguladora para na forma

de um pedido de Registro do novo Fármaco ou de Aplicação como novo Fármaco. Todos os dados são revisados por especialistas das próprias agências e somente após a aprovação o fármaco pode ser prescrito pelos médicos aos pacientes para tratar a doença para a qual foi designado. Na fase clínica IV, são realizados os monitoramentos pós-comercialização dos novos fármacos. [102, 105]

A ação de um dado fármaco depende de uma sequência de eventos físico-químicos que começa quando a molécula bioativa penetra no organismo vivo e termina quando a mesma é eliminada após ter alcançado seu alvo e provocado a resposta biológica apropriada. Classicamente é admitido que três fases características governam a atividade biológica de um fármaco dentro de um organismo vivo: as *fases farmacêutica, farmacocinética e farmacodinâmica*.

A *fase farmacêutica* é constituída pelo conjunto de fenômenos compreendidos entre a administração do medicamento e a absorção propriamente dita, que determinam a intensidade e a velocidade com que ocorre a entrada da substância ativa no organismo. Estes fenômenos compreendem basicamente a liberação e dissolução do fármaco contido no medicamento. Algumas possíveis rotas de administração são a oral, intravenosa, nasal, transdermal, etc. Na formulação é decidido se o composto irá ser administrado na forma de comprimidos, cápsulas, líquido, etc. Quando a formulação farmacêutica de um composto não é efetiva, modificações químicas sutis ou a formação de derivados bioreversíveis, ou seja, que podem ser facilmente convertidos pelo organismo aos compostos ativos (e.g. ésteres ou amidas), podem melhorar propriedades físico-químicas tais como lipofilia, pK_a , polaridade, e otimizar a taxa de dissolução nos fluidos biológicos e a passagem pelas primeiras membranas encontradas após sua administração (e.g. cutânea, intestinal, etc.). Deste modo, o resultado global é uma melhor absorção pelo organismo.

A *fase farmacocinética* corresponde ao estudo da evolução temporal do movimento do fármaco e seu destino *in vivo* após administrado em um organismo. Em ordem cronológica os eventos da fase farmacocinética ocorrem como segue: absorção, distribuição, metabolismo e excreção do fármaco, conhecidos simplesmente por ADME. Durante a absorção ele atravessa membranas biológicas e organelas. Esse processo é fortemente dependente das características físico-químicas do fármaco (ionizado, não ionizado, tamanho, coeficiente de partição, etc.) e pode ocorrer através de *difusão passiva* ou mecanismos fisiológicos mais

sofisticados. A corrente sanguínea se encarrega de distribuí-lo pelo organismo. A função do metabolismo é transformar quimicamente os fármacos ou qualquer outra substância estranha ao organismo (xenobióticos) em derivados solúveis em água para facilitar a sua eliminação através da urina. Esta mudança geralmente produz uma supressão da atividade farmacológica e da eventual toxidez. Entretanto, também pode acontecer de o metabolismo ativar um determinado composto ou mesmo produzir intermediários altamente reativos (a maioria eletrófilos) que induzem mecanismos de toxidez. Se for possível a ativação metabólica, fármacos inativos *in vitro* podem se mostrar ativos *in vivo* e este é o princípio de funcionamento dos chamados *pro-fármacos*. Uma vez que o efeito farmacológico é produzido, o fármaco e seus metabolitos devem ser eliminados do organismo com cinética apropriada. As principais rotas de eliminação são a renal (urina) e retal (fezes), mas ocasionalmente também pode ocorrer por outras vias (e.g. pulmonar, salivar, cutânea).

Os quatro processos de ADME estão relacionados com a biodisponibilidade dos fármacos. Como a biodisponibilidade é expressa como a porcentagem do composto que alcança a corrente sanguínea em um determinado tempo, a administração intravenosa representa então, por definição, 100% de biodisponibilidade. Depois de uma dosagem oral, por exemplo, uma biodisponibilidade de 100 % implicaria na completa absorção e nenhum metabolismo de primeira passagem.

A *fase farmacodinâmica* é a fase de maior interesse para o químico medicinal e trata diretamente da natureza e da qualidade das interações entre o fármaco e seu alvo biomacromolecular. Partindo de um composto pouco potente e não seletivo, o desafio é maximizar a potência e minimizar os efeitos adversos causados pelo composto. A resposta biológica é máxima quando o princípio ativo exibe uma perfeita complementaridade estereoquímica e eletrônica com a estrutura do alvo. Idealmente, o químico medicinal, baseado nos conhecimentos das características do alvo, tenta planejar compostos que possuam tamanho, forma, propriedades hidrofóbicas e hidrofílicas e a disposição dos grupos funcionais apropriadas. Quanto maior é o ajuste do composto com o sítio do alvo, mais seletivo será o fármaco em promover apenas a resposta biológica desejada. Por isso, a procura exaustiva dentro do universo infinito de combinações entre os átomos por aqueles compostos capazes de interagir eficientemente com as biomacromoléculas envolvidas em

processos importantes dentro do organismo é o ponto central do planejamento de fármacos. [101]

A Química Medicinal desempenha um importante papel na fase pré-clínica e as contribuições se concentram na sequência de passos ilustrados na FIGURA 1.21, que são: identificação, seleção e validação do alvo; identificação de ligantes; otimização dos ligantes a compostos-líderes otimização dos compostos-líderes. [106]

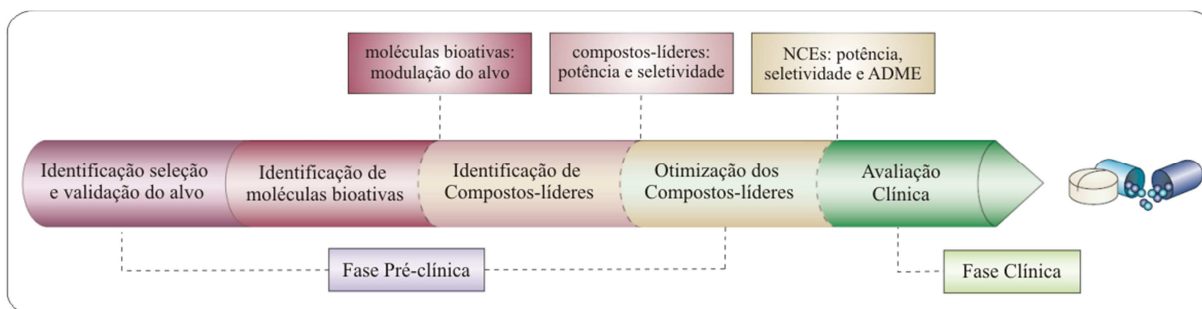


FIGURA 1.21: Etapas de identificação, seleção e otimização de substâncias bioativas durante o processo de descoberta de novos fármacos (fase pré-clínica) [106]

O conhecimento de alvos moleculares (e.g. enzimas, receptores, ácidos nucléicos) tem se beneficiado do progresso feito na biologia molecular, engenharia genética e na biologia estrutural. É crescente o número de alvos para os quais se conhece a estrutura 3D e a localização precisa do sitio ativo. Como consequência o planejamento de novas substâncias químicas bioativas é cada vez mais e realizado com base nos resultados obtidos de estudos de modelagem receptor-ligante. Além da identificação do alvo molecular, a validação também é fundamental, pois estabelece a sua relevância no processo fisiopatológico em estudo e evidencia se a sua modulação seletiva é capaz de gerar a resposta farmacológica esperada no controle da doença ou desordem metabólica em humanos. [106-107]

Os ligantes têm sua origem a partir de produtos naturais ou através de síntese orgânica e coleções combinatórias. Estas moléculas podem ser identificadas através de ensaios reais, virtuais (*in silico*) ou ainda através de *planejamento racional*, mas em todos os casos as suas propriedades biológicas devem ser determinadas experimentalmente para verificar suas capacidades de modulação do alvo de interesse. Nesta fase inicial geralmente são identificadas moléculas com baixa afinidade (e.g. 0,1-100 μM) que necessitam serem otimizadas em relação a

uma série de propriedades (e.g., potência, afinidade, seletividade, biodisponibilidade, toxidez) até se tornarem compostos-líderes promissores. [108]

Nem todos os ligantes possuem propriedades químicas e biológicas adequadas para serem considerados compostos-líderes. Primariamente, possuir alguma afinidade pelo alvo macromolecular confirmada é essencial, entretanto, desde a identificação dos ligantes pode-se priorizar compostos com características moleculares mais apropriadas do ponto de vista de acessibilidade sintética de análogos ou que possuam características físico-químicas compatíveis com o que já se conhece sobre o perfil necessário para aumentar a probabilidade de sucesso nas fases subsequentes da gênese do fármaco. Os ligantes com características adequadas são selecionados como compostos-líderes e otimizados até que possuam potência maior, em geral em concentrações inferiores 100 nM contra o alvo macromolecular e abaixo de 100 µM contra o modelo *in vivo*. [107]

Na última etapa da Fase Pré-clínica os compostos-líderes aperfeiçoados em relação às propriedades farmacocinéticas, e também é realizado um ajuste fino das propriedades farmacêuticas (formulação química) para torná-los *novas entidades químicas* aceitáveis ao uso na fase de desenvolvimento clínico até que alcancem a terapia.

Todas essas etapas demandam o conhecimento multidisciplinar englobado pela Química Medicinal juntamente com a integração de diversas tecnologias para a investigação e o entendimento dos efeitos das propriedades moleculares nas diferentes fases de ação de pequenas moléculas dentro e fora do organismo humano.

1.3.1 O espaço químico-biológico

O processo de descoberta de uma substância bioativa é o ponto central da química medicinal. Tradicionalmente, isto era feito através de síntese orgânica sistemática (ou isolamento de produtos naturais) e ensaios biológicos de vários compostos. Todo este processo depende da exploração adequada do *espaço químico* para que seja possível selecionar aquelas moléculas com potencial para se tornarem bons compostos-matrizes.

Com a evolução da química medicinal, tornou-se claro que mesmo os métodos mais eficientes de química combinatória são capazes de gerar apenas uma fração diminuta da diversidade química que define o espaço químico a ser varrido na busca por substâncias químicas bioativas [109]. Hoje, estima-se que esse conjunto exceda 10^{20} compostos orgânicos, mesmo quando nos limitamos a aqueles com MM menores que 500 Da e que contêm os elementos tipicamente encontrados em fármacos, ou seja, C, O, H, N, P, S, F, Cl, Br, I. Devido à imensidão de moléculas contidas no espaço químico, o desafio na descoberta de fármacos é localizar as substâncias com atividade biológica desejada, contidas em uma região do espaço químico definida como espaço químico-biológico. [108, 110]

Com base nisso, os estudos focados em planejamento de substâncias bioativas tem como um de seus objetivos específicos a utilização de métodos computacionais ou experimentais que permitam “navegar” no *espaço-químico* e localizar aquelas regiões de *espaço químico-biológico* específicas para o alvo de interesse.

1.3.2 A Quiminformática na identificação e otimização de ligantes

A exploração sistemática do *espaço químico* e o delineamento das regiões estabelecidas pelo *espaço químico-biológico* só são possíveis através da criação de bancos de dados *in silico* para o armazenamento adequado das informações sobre as pequenas moléculas e os alvos macromoleculares. Como consequência disso, os métodos e ferramentas computacionais tem se tornando parte integral durante as etapas de identificação e otimização de ligantes para o armazenamento, gerenciamento, análise de dados e modelagem molecular de compostos, tanto dentro das indústrias farmacêuticas quanto na academia. O desenvolvimento de estratégias e a aplicação de métodos computacionais para a solução dos problemas associados à seleção e otimização de candidatos a fármacos fazem parte do escopo de uma disciplina da química medicinal que se chama *Quiminformática*.

O termo “*quiminformática*” foi definido pela primeira vez por Brown (1998) [111] como “*uma combinação de recursos para transformar dados em informação e informação em conhecimento, com o propósito de se tomar melhores decisões mais rapidamente dentro do processo de planejamento de fármacos*”. Uma definição mais

ampla foi dada por Gasteiger: “A *Quiminformática* consiste do uso dos métodos de *informática para resolver problemas químicos*”. [112]

Nos últimos anos, a *Quiminformática* tem emergido como um avanço tecnológico guiado pelas inovações na área de informática para alavancar as pesquisas pré-clínicas, uma vez que ela une todas as frentes de pesquisa, dos ensaios virtuais à toxicologia animal, por via de um único elemento central: a *estrutura química*. Os seguintes protocolos de química computacional fazem parte das atividades da *Quiminformática* no contexto da descoberta de fármacos: descrição, manipulação e análise de bancos de dados de estruturas químicas (virtuais ou reais); SAR (qualitativos ou quantitativos) para a identificação de séries de compostos-líderes; pós-processamento de dados obtidos em ensaios (reais ou virtuais) contra os alvos; planejamento *de novo* para geração e otimização de compostos-líderes; ensaios *in silico* para a predição de propriedades ADME, predição do metabolismo e toxicologia; mineração e visualização de dados; *ensaios virtuais*. [113]

CAPÍTULO 2 - OBJETIVOS

2.1 Objetivos Gerais

O objetivo geral deste trabalho foi identificar inibidores da enzima cruzaina com atividade tripanosomicida através da integração de métodos em quiminformática e bioquímicos.

2.2 Objetivos Específicos

(i) Selecionar uma coleção de compostos virtuais com grande diversidade química para a busca de novos ligantes da enzima cruzaina

(ii) Aplicar métodos de docagem molecular na sub-coleção de compostos para a identificação de novos ligantes da enzima cruzaina

(iii) Aplicar métodos de busca, baseados em similaridade química na sub-coleção de compostos para identificação de novos ligantes com características estruturais similares com inibidores da enzima cruzaina

(iv) Construir modelo de relações quantitativas estrutura atividade (HQSAR) baseado em hologramas moleculares utilizando inibidores conhecidos da cruzaina

(v) Estabelecer uma estratégia de uso consensual dos diferentes métodos computacionais para a seleção das moléculas a serem testadas nos ensaios de bioquímicos

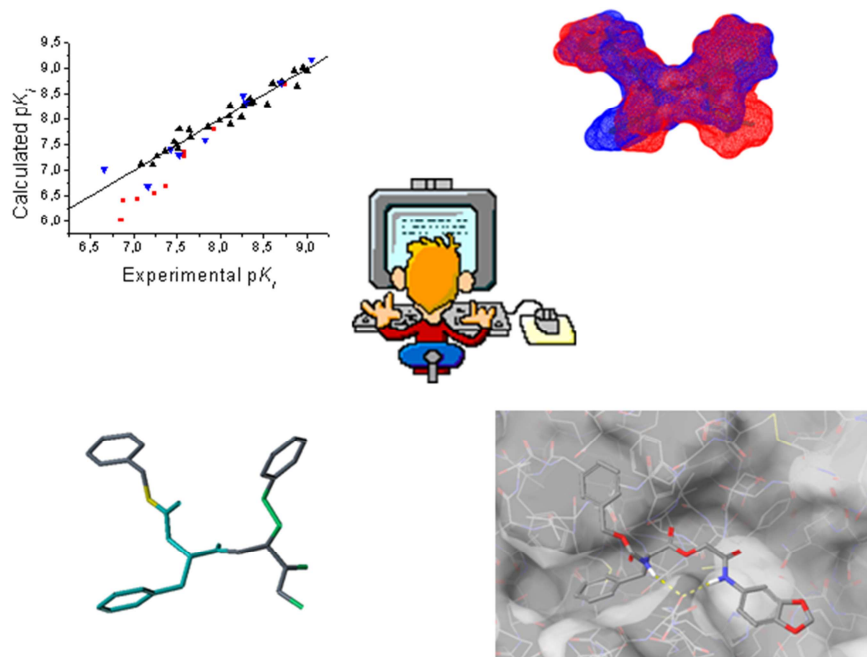
(vi) Determinar a atividade inibitória dos compostos selecionados pelos ensaios virtuais através de ensaios de inibição enzimática contra as enzimas cruzaina de *T. cruzi* e catepsina-L de humanos para verificar seletividade

(vii) Determinar o mecanismo de inibição dos inibidores identificados

(viii) Determinar a atividade tripanosomicida dos inibidores da enzima cruzaina

(ix) Estabelecer relações estrutura atividade para novas classes de compostos identificados para a validação da atividade inibitória através de ensaios de substâncias com estruturas análogas

CAPÍTULO 3 – ENSAIOS VIRTUAIS



O capítulo 3 apresenta os métodos em quiminformática utilizados nos ensaios virtuais para a busca de inibidores da enzima cruzaina

3.1 Coleções Virtuais de Moléculas e Filtros de seleção

A seleção da coleção virtual de moléculas para os ensaios virtuais é uma etapa fundamental na identificação de um ligante. A coleção deve conter a maior diversidade química disponível para haver a investigação em uma ampla região do espaço químico. A maior coleção de moléculas disponíveis comercialmente de domínio público é encontrada no banco de dados do ZINC (<http://zinc.docking.org>). [114]

O primeiro ensaio virtual a ser realizado é aplicação de filtros de seleção de moléculas com características moleculares desejáveis para o desenvolvimento de um novo fármaco. Para se tornar um fármaco, um composto necessita além de uma alta afinidade pelo alvo a incorporação de propriedades farmacocinéticas apropriadas [115]. As propriedades ADME derivam de uma vasta faixa de mecanismos *in vivo*, muitos dos quais são desconhecidos ou são multicomponentes. Durante a década de 90 as indústrias farmacêuticas começaram a notar que muitos compostos não passavam pela fase clínica em virtude de apresentar propriedades farmacocinéticas indesejáveis.

O desenvolvimento de coleções originadas de ensaios em larga escala (HTS), coleções combinatórias, assim como coleções virtuais, tem-se focado muito no conceito de fármaco-similar (drug-likeness) dos compostos. Lipinski define fármacos-similares como “aqueles compostos que apresentam propriedades ADME e toxidez suficientemente aceitáveis para sobreviverem até o final dos testes clínicos da fase I”. [116]

O conjunto dos parâmetros que é comumente utilizado é chamado de regra dos cinco de Lipinski [117], que leva em consideração a importância da lipofilia (coeficiente de partição óleo-água), expresso por $\log P$, massa molecular (MM) e o número de doadores e receptores de ligações de hidrogênio. Este primeiro filtro fármaco-similar prevê que a baixa absorção ou permeação é mais provável quando o composto apresenta mais de 5 doadores em ligações de hidrogênio, 10 receptores de ligações de hidrogênio, massa molecular de 500 Da e o coeficiente de partição 5 ($\log P$).

Outro conceito importante utilizado em filtros de seleção é o da molécula matriz-similar. Foi estabelecida a diferença entre fármaco e substância matriz e podem ser pontuados dois significados gerais para a substância matriz. Em uma

definição as substâncias têm dimensões reduzidas em relação ao intervalo das propriedades do fármaco (menor massa molecular, log P, aceitadores e doadores de ligação hidrogênio etc). Em outra definição, a descoberta de matriz similar se refere à análise das coleções de moléculas com baixa massa molecular, com detecção de afinidades na faixa de micromolar alto milimolar. Essas coleções são muitas vezes referidas como coleções de fragmentos moleculares. Baseado na regra dos cinco de Lipinski, uma regra de três foi cunhada para essas pequenas moléculas para a triagem de fragmentos; MM <300; log P < 3; doadores e aceitadores de ligação de hidrogênio < 3 e ligações livres < 3. O delineamento experimental em comparação com a diversidade molecular, teoricamente realizável é muito maior em uma biblioteca de fragmentos com baixa MM comparativamente a coleção convencional com MM maior. Matrizes possuem estrutura molecular menos complexa que fármacos e o processo de otimização a fármaco quase invariavelmente aumenta MM e log P. [118-119]

3.2 Docagem Molecular

A docagem molecular (do inglês *molecular docking*) é um método baseado na estrutura do receptor (TBVS), pode ser definida como uma tentativa de prever conformação e orientação da estrutura do complexo formado entre um ligante (uma pequena molécula ou até mesmo uma proteína) e um receptor (enzima, DNA, canais iônicos, receptores, etc.) [120]. Desta forma, um pré-requisito é a informação tridimensional detalhada da biomacromolécula em estudo, a qual é geralmente conseguida através da análise de estruturas obtidas pela resolução de estruturas de proteínas por raios-X, estudos de espectroscopia de ressonância magnética ou através de modelagem por homologia. [121]

Programas de docagem molecular permitem a busca automatizada de maneira mais objetiva, o modo de interação entre duas moléculas. O número de conformações que duas moléculas podem interagir é grande e, almejando uma busca eficiente, o problema da docagem molecular é simplificado permitindo-se apenas certos graus de liberdade a serem explorados. Em meados da década de 80 surgiu o primeiro programa de docagem molecular, as moléculas eram descritas como uma série de pontos espalhadas sobre a superfície de van der Waals; neste método a superfície molecular é descrita desde um ponto distante até a superfície de

van der Waals, onde uma esfera de hidratação é adicionada ao raio de van der Waals. O limite de superposição da superfície molecular ligante/receptor pode ser obtido sem problemas através desta abordagem. Um método analítico para a descrição tridimensional do contorno molecular foi desenvolvido por Connolly. Este programa foi crucial para o desenvolvimento dos algoritmos de docagem, bem como a deposição de um grande número de proteínas disponíveis no banco de dados de proteínas (PDB). Os primeiros algoritmos foram baseados principalmente em critérios geométricos, embora alguns poucos contivessem informações de energia.

O programa pioneiro de docagem molecular foi o DOCK, desenvolvido por Kuntz e colaboradores[122]. Este programa e seu modo atrativo da descrição da interação, pela interseção de esferas, inspirou o campo molecular de interação (GRID) que foi integrado em muitos algoritmos.

Há três principais características na docagem molecular:

- representação do sistema;
- busca conformacional no espaço;
- pontuação de potenciais candidatos;

A docagem simula a interação na superfície da proteína. Portanto, a primeira questão é como definir a superfície da proteína. Elas podem ser descritas por modelos matemáticos como, por exemplo, descritores de formas geométricas ou grades. Alternativamente, a descrição pode envolver um tratamento, estático ou dinâmico, das cadeias protéicas, como por exemplo, rígido vs flexível. A docagem envolve duas moléculas separadas, começa pelo enovelamento da proteína e as conformações dos ligantes. Entretanto, a predição computacional de estruturas complexas é relativamente limitada e é um problema enfrentado pelos programas de docagem. Uma abordagem é determinar a complementaridade entre fragmentos que provêm um estudo do enovelamento protéico consistente com a organização da interação intermolecular e mudanças conformacionais intramoleculares da proteína [123]. O processo de docagem exige um procedimento eficiente de busca e uma função de pontuação. Os dois elementos críticos para este fim são a velocidade e a eficiência em cobrir um espaço conformacional relevante; é essencial que a função de pontuação seja rápida o suficiente para permitir a obtenção de um grande número de potenciais soluções, e em princípio, discriminar entre as conformações docadas efetivamente e não efetivas. A função de pontuação deve incluir um peso apropriado para todos os componentes energéticos da interação. Desta forma, uma

solução apresentada por um programa de docagem em particular, não deve ser vista como uma representação completa, mas sim uma peça dentro de um vasto espaço conformacional. Desta forma, uma abordagem ideal, é combinar os melhores algoritmos e funções de pontuação de diferentes programas de docagem, de maneira a obter uma visão mais ampla do espaço conformacional do processo de interação. [124]

Para inibidor e enzima o objetivo da docagem é prever corretamente a estrutura do complexo $[E] + [I] \rightarrow [EI]$ sob condições de equilíbrio.

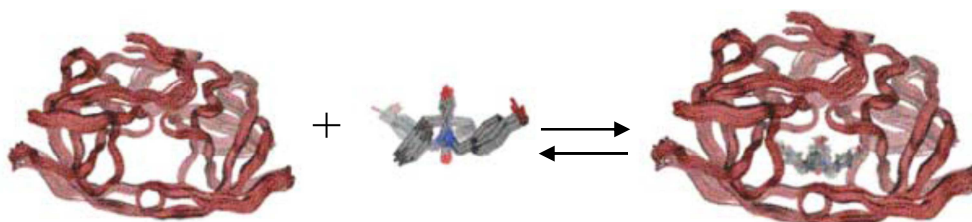


FIGURA 3.1: Ilustração do processo de interação entre enzima e inibidor [125]

A FIGURA 3.1 ilustra a interação do inibidor Dmp323 com a enzima HIV protease e é baseado em soluções das estruturas (código PDB: 1BVE). Múltiplas estruturas do complexo enzima-inibidor revelaram variações estruturais limitadas. A energia livre de interação ($\Delta G_{\text{interação}}$) é relacionada com a afinidade de interação pelas equações:

$$\Delta G_{\text{interação}} = -RT \ln K_a \quad K_a = K_i^{-1} = \frac{[EI]}{[E] + [I]} \quad (3.1)$$

onde R é a constante dos gases, K_a constante de afinidade, K_i constante de inibição, [EI] concentração do complexo enzima-inibidor, [E] concentração da enzima e [I] concentração do inibidor

A predição correta da estrutura do complexo (pose) não requer informações a respeito da constante de afinidade. Entretanto, a predição da afinidade (pontuação) requer esta informação; a função de pontuação deve considerar os seguintes fatores: estereo, eletrostático, ligação de hidrogênio, restrição conformacional do ligante (se flexível) e enzima. Considerando o equilíbrio de formação do complexo, há outros fatores importantes: dessolvatação, entropia rotacional e entropia translacional.

Esta técnica vem sendo empregada como um método de triagem inicial, onde bases de dados contendo de milhares a milhões de compostos reais ou virtuais são utilizados para identificar e classificar substâncias bioativas, inclusive propor uma possível conformação bioativa.

Essa técnica é aplicada em diferentes estágios no processo de desenvolvimento de fármacos com o objetivo de cumprir três propostas principais: prever o modo de ligação de uma molécula conhecida, identificar novos ligantes usando o ensaio virtual e prever a afinidade de ligação de compostos relacionados a partir de uma série de ligantes ativos. [126]

Coleções extensas de compostos podem ser “docadas”, em um programa computacional, com um grande número de conformações, na estrutura do receptor. Em seguida, como resultado, as moléculas são pontuadas com base na sua complementaridade com o receptor e as mais bem pontuadas podem ser subsequentemente avaliadas em um ensaio experimental, FIGURA 3.2.

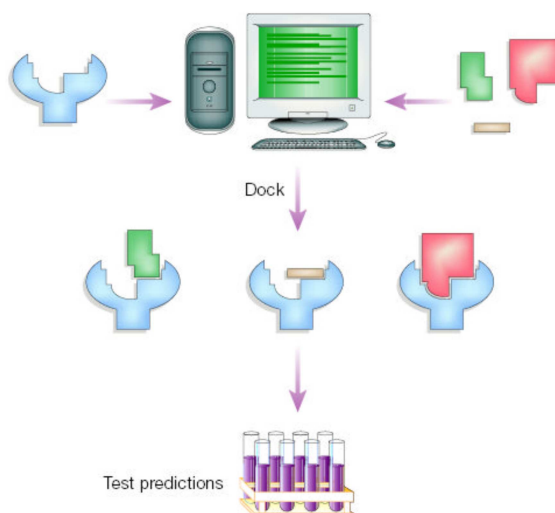


FIGURA 3.2: Ensaio virtual para identificação de novos ligantes

Grandes coleções são docadas na estrutura do receptor utilizando um programa computacional. Cada composto é testado milhares de possíveis configurações e pontuados com base na complementaridade com o receptor. Os compostos melhores pontuados são testados experimentalmente contra o alvo para avaliação da atividade biológica [126]

Os protocolos de docagem podem ser descritos como a combinação de duas componentes: um algoritmo de busca dos graus de liberdade configuracional e conformacional e uma função de pontuação ou avaliação. O algoritmo de busca deve

elucidar todos os possíveis modos de ligação entre o ligante e o receptor, explorados os graus de liberdade translacional e rotacional dos ligantes e do receptor. Usualmente, as funções de pontuação indicam a complementaridade estérea e química entre o ligante e o receptor. [125]

3.2.1 Algoritmos de busca e flexibilidade molecular

A flexibilidade do ligante é explorada por diferentes métodos e pode ser dividida basicamente em três categorias: algoritmos de biblioteca multiconformacional, algoritmos de construção incremental e algoritmos estocásticos.

3.2.1.1 Algoritmo de biblioteca multiconformacional

Este método realiza a docagem de corpos rígidos, a flexibilidade é considerada através da biblioteca multiconformacional de cada ligante. Os programas de docagem que utilizam este algoritmo utilizam complementaridade de forma ou interação, alguns exemplos de programas são FRED, FTDock, LIGIN e SANDOCK. No método de complementaridade de interação um farmacóforo inicial, o qual representa a proteína, é desenvolvido e utilizado para guiar a docagem. Uma conformação inicial do ligante é gerada e um farmacóforo do ligante é derivado desta conformação. As matrizes das distâncias (lista de todas as distâncias entre cada ponto farmacofórico) do ligante e a proteína são examinadas e vetores de posição são calculados para avaliar rotação e translação no sitio de interação da proteína. A velocidade é a maior vantagem deste método, ao passo que, sua acurácia na predição da pose do complexo ligante proteína é baixa e reflete na pontuação da energia de interação. Esta classe de programa de docagem é útil em coleções de compostos de médio a grande escala para o enriquecimento de moléculas ativas.

3.2.1.2 Algoritmo de busca sistemática ou crescimento incremental

Estes algoritmos exploram todos os graus de liberdade de uma molécula, mas enfrentam o problema da “explosão combinatória”. Para evitar uma busca exaustiva,

muitos algoritmos de busca usam um método de construção incremental, crescendo o ligante no sitio ativo, que consiste basicamente de três passos: seleção do fragmento principal, ajuste posicional do fragmento principal e crescimento incremental do ligante. No primeiro passo, o ligante é dividido em um fragmento rígido com cadeias laterais flexíveis; subsequentemente, essas cadeias laterais são divididas em cada nova ligação rotacionável, no segundo e terceiro passos, o fragmento é colocado e as cadeias laterais são ligadas incrementalmente, considerando-se os graus de liberdade rotacionais possíveis. O crescimento incremental pode ser realizado de diferentes modos. Um método é docar vários fragmentos moleculares na região do sitio ativo da enzima e em seguida conectá-los covalentemente (método de novo) ou, alternativamente dividir os ligantes docados em partes rígidas e flexíveis. Neste caso, uma vez que o fragmento rígido é definido, ele é docado do sitio ativo, em seguida as regiões flexíveis são adicionadas incrementalmente. Exemplos de programas que utilizam este algoritmo são DOCK, FlexX, FLOG, Hammerhead e Surfex). [127]

3.2.1.3 Algoritmos estocásticos

O algoritmo estocástico faz uma busca aleatória para um ligante, ou mesmo para uma população de ligantes, cada novo complexo de interação entre proteína e ligante obtido é avaliado baseado numa função de probabilidade pré-definida. Os dois métodos mais amplamente utilizados são Monte Carlo e algoritmo genético.

O algoritmo de Monte Carlo possui basicamente a seguinte forma: geração da conformação inicial do ligante no sitio ativo consistindo de uma forma aleatória de conformação, translação, rotação e pontuação da configuração inicial. Geração da segunda conformação e sua pontuação, uso do critério de Metrópolis¹ para determinar se a nova configuração é mantida e repetição de vários passos até um número desejado de configurações a serem obtidas. Exemplos de programas que

¹ Critério de Metrópolis: se uma nova solução é melhor que a prévia é imediatamente aceita - se a configuração não representa um novo mínimo de energia, uma função de probabilidade de Boltzman é aplicada; se a solução passa pelo teste da função de probabilidade é aceita, se não a configuração é descartada.

utilizam este algoritmo são DockVision, ICM, MCDOCK, ProDOCK e SLIDE.
Configuração

Os algoritmos genéticos são baseados na teoria da evolução de Darwin. A pose do ligante é representada por um cromossomo constituído de genes. Os genes codificam cada ângulo torsional, rotação do ligante e translação. As poses envolvem transmissão da informação dos genes (reprodução), alterada por operações genéticas como cruzamento e mutação. Existem vários métodos de seleção da próxima geração, mas, a mais utilizada é a sobrevivência da pose de melhor complementaridade, onde duas conformações com melhor energia passam para a próxima geração. O algoritmo possui basicamente os seguintes passos: fazem pequenas mudanças aleatórias na conformação inicial, pontua cada mudança de acordo com o valor da função de pontuação, determina quais mudanças são favoráveis à interação. Se a melhor modificação tem um valor menor de energia é aceita, e cada nova conformação é reportada com sua pontuação. Exemplos de programas são AutoDock, DARWIN, DIVALI, GOLD, EADock, FITTED e PSI-DOCK.
[128]

3.2.2 Funções de pontuação

As funções de pontuação possuem duas funções, direcionar a pose durante o procedimento de docagem e predizer a afinidade de interação da pose final. De uma forma geral a predição da afinidade é utilizada para listar de maneira ordenada as moléculas com maiores chances de apresentar atividade biológica. As funções de pontuação são classificadas em funções baseadas no campo de força, no conhecimento e empírica. As funções de pontuação têm limitações, muitas decorrentes de simplificações e exclusão de contribuições importantes para a energia de interação como solvatação e entropia. Para superar este obstáculo, uma alternativa tem sido utilizar mais de uma função de pontuação para selecionar a pose de melhor energia de interação nos programas de docagem. Como por exemplo, as funções ChemScore, ScreenScore e PLP implementadas no programa FRED. Além do consenso na docagem, a pontuação combina várias funções de pontuação, mostrando-se superior ao uso de uma única função em alguns casos.
[129-130]

3.2.2.1 Funções de pontuação baseadas em campos de força

Funções de pontuação baseadas em campos de força são tipicamente fáceis e rápidas, e por esta razão são recomendadas nos primeiros estágios do planejamento molecular para docagem em larga escala. Em geral, consistem de termos de campos de força de mecânica molecular (AMBER, CHARMM), onde se considera ligante e proteína não ligados e avaliação em fase gasosa das suas interações. [131]

A função de escore do DOCK utiliza parâmetros eletrostáticos e de van der Waals para o cálculo do potencial de AMBER:

$$E_{\text{interação}} = \sum_{ij} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \left(3320 \frac{Q_i Q_j}{\epsilon(r_{ij}) r_{ij}} \right) \quad (3.2)$$

O primeiro termo da equação é o potencial de Lennard-Jones para interação de van der Waals, A e B são os diâmetros de colisão para os respectivos átomos i e j . O termo exponencial 12 é responsável pela repulsão em pequenas distâncias, o termo exponencial 6 representa o termo de atração que se aproxima de zero à medida que a distância entre dois átomos aumenta.

O segundo termo da equação representa a energia potencial eletrostática pela soma da interação de Coulomb de dois átomos, onde Q é a carga, r a distância entre os átomos e ϵ constante dielétrica de Coulomb. [127]

Funções de pontuação padrões têm grandes limitações, principalmente devido a consideração do modelo entálpico em fase-gasosa que não inclui termos entrópicos e de solvatação. Extensões mais recentes introduziram entropia torcional e inclusão explícita de ligações de hidrogênio entre o receptor e o ligante, em programas como GOLD e AutoDock.

3.2.2.2 Funções de pontuação empíricas

As funções de pontuação empíricas têm o objetivo de reproduzir resultados experimentais. A afinidade de interação entre ligante e proteína é decomposta em

uma série de termos relevantes para a energia livre de interação e cada termo é multiplicado por um fator de peso determinado por uma função matemática derivada de dados experimentais. A primeira função de pontuação empírica foi proposta por Böhm, e neste caso é dada por:

$$\Delta G_{\text{interação}} = \Delta G_0 + \Delta G_{H\text{-bond}} \sum_{H\text{-bond}} f(\Delta R, \Delta \alpha) + \Delta G_{\text{ionica}} \sum_{\text{ionica}} f(\Delta R, \Delta \alpha) + \Delta G_{\text{lipo}} |A_{\text{lipo}}| + \Delta G_{\text{rot}} N_{\text{rot}} \quad (3.3)$$

onde $\Delta G_{\text{interação}}$ é a energia livre de interação, $\Delta G_{H\text{-bond}}$ a contribuição de uma interação de hidrogênio ideal, ΔG_{ionica} a contribuição de uma interação iônica e $f(\Delta R, \Delta \alpha)$ uma função de penalidade que considera os desvios das interações ideais.

Esta função foi originalmente calibrada pelo ajuste da constante de afinidade de 45 complexos ligante-proteína e foi capaz de reproduzir as constantes de afinidade deste conjunto de moléculas com um desvio padrão de $1,9 \text{ kcal mol}^{-1}$. [132]

Uma importante extensão da função de Böhm foi a implementação feita no programa FlexX. O termo da interação lipofílica foi substituído por:

$$\Delta G_{\text{lipo}} |A_{\text{lipo}}| = \Delta G_{\text{aromatic}} \sum_{\text{aromatic}} f(\Delta R, \Delta \alpha) + \Delta G_{\text{lipo}} \sum_{\text{lipo}} f(\Delta R) \quad (3.4)$$

onde a distinção é feita entre contatos aromáticos da proteína e do ligante e todos os contatos lipofílicos; as funções $f(\Delta R, \Delta \alpha)$ e $f(\Delta R)$ são funções de penalidade como citado anteriormente, entretanto, a determinação de ligações ideais e ângulos são modificados de maneira a refletir interações aromáticas e lipofílicas (a distância considerada ideal nestes contatos é o raio de van der Waals). Os coeficientes desta função foram originalmente calibrados utilizando 19 complexos proteína-ligante e este conjunto foi reproduzido com um desvio quadrático médio (RMSD) entre 0,5 e 1,2 Å quando utilizado o programa FlexX.

Dentre as funções de pontuação empíricas, a mais comumente utilizada é ChemScore, ela é utilizada por vários programas de docagem, como por exemplo GOLD e FRED. As funções de pontuação empíricas incluem contribuições não entálpicas assim como termos de rotação, as quais incluem penalidades entrópicas para a interação através da atribuição de pesos na soma do número de ligações rotacionáveis nos ligantes. [133]

3.2.2.3 Funções de pontuação baseadas no conhecimento

As funções de pontuação baseadas no conhecimento são derivadas pela análise de estrutura de complexos determinados experimentalmente. Nessa análise, um conjunto de ligantes e fragmentos de proteínas identificados é utilizado para listar as distâncias entre os fragmentos de todas as estruturas e compilá-los. A probabilidade de um determinado inter-fragmento ocorrer para cada par é calculada, então a distância dependente de dois fragmentos potenciais é descrita para cada par:

$$\Delta W_{ij}(R) = -\ln\left(\frac{g_{ij}(R)}{g(R)}\right) \quad (3.5)$$

onde $g_{ij}(R)$ é a probabilidade de distribuição de um par de fragmentos i,j , R a distância e $g(R)$ é a distribuição de referência selecionada; a medida que $\Delta W_{ij}(R)$ se aproxima de zero a distância R tende a infinito. Através da soma de todos os fragmentos potenciais obtém-se a energia livre total de interação;

$$\Delta G_{interacao} = \sum_{i,j} \Delta W_{ij} \quad (3.6)$$

Klebe et al. construíram a função DrugScore que foi derivada de dados provenientes de estruturas de 1376 complexos proteína-ligante depositados no banco de proteínas. A função foi usada com sucesso para reposicionar a estrutura de 158 complexos docados pelo programa FlexX. [134]

A maior vantagem das funções de pontuação baseadas no conhecimento é sua simplicidade computacional, isso permite de maneira eficiente testar grandes coleções de compostos, uma desvantagem é sua derivação ser essencialmente baseada em informações implícitas determinadas por um conjunto limitado de estruturas de complexos proteína-ligante.

3.2.2.4 Funções de pontuação Consensuais

Visto as limitações de cada função de pontuação, uma tendência tem sido introduzir esquemas de pontuação consensual. Este método combina informações de diferentes funções e balanceia em uma função única aumentando a probabilidade de encontrar um ligante. Um exemplo de implementação foi a função X-CSCORE

que combinam as funções GOLD-like, DOCK-like, ChemScore, PMF e FlexX. Entretanto o valor potencial da pontuação consensual pode ser limitado, em termos de diferentes funções de pontuação correlacionadas que podem amplificar erros de cálculos ao invés de balanceá-los. [135]

Tanja Schulz-Gasch e Martin Stahl [136] compararam programas de docagem com seus respectivos algoritmos de busca e suas funções de pontuação. Para isto, avaliaram sete complexos proteínas-ligantes com características de polaridade e tamanho da cavidade do sitio ativo distinta. O desempenho dos algoritmos é altamente dependente das características da estrutura do receptor, e como um guia geral tem-se:

Sítios lipofílicos onde ajustes estereoquímicos são de grande importância, o método recomendado é a docagem multiconformacional em combinação com uma função de pontuação que considera a orientação da ligação (ChemScore por exemplo).

Sítios predominantemente lipofílicos, mas com importantes grupos polares são melhores descritos por algoritmos de construção incremental.

Sítios polares onde os ligantes têm uma rede de interações dirigidas, um método de escolha é o de construção incremental, mas com uma função de pontuação que considera a orientação da ligação.

De um modo geral, a função ChemScore mostrou-se mais aplicável em combinação com um algoritmo multiconformacional.

A busca virtual de inibidores é mais consistente quando foca-se a busca com a inserção de restrições previamente definidas. Consequentemente, alguns detalhes como a definição das ligações no sitio ativo são críticas.

Os programas que correspondem satisfatoriamente a estes requisitos são o FRED e o FlexX. O programa FRED apresenta como restrição do sitio ativo coordenadas em uma caixa, ao passo que no FlexX pode-se definir um conjunto de átomos. A inclusão de restrições farmacofóricas em ambos os programas pode ser feita. Isto é essencial para incorporar informações prévias baseadas em ligantes posicionados em complexos como guia para a busca de novos ligantes. O sitio da enzima cruzaina possui característica lipofílica com interações de hidrogênio específica, por isso os algoritmos multiconformacionais e de construção incremental em combinação com as funções de pontuação Shapegauss, PLP, Chemgauss2,

Chemgauss3, CGO, CGT, Chemscore, OEChemscore e Screenscore disponíveis nestes dois programas foram empregadas neste trabalho.

3.3 Similaridade química

Os métodos de similaridade para a busca em bancos de dados foi introduzido cerca de vinte anos atrás. Estes métodos são normalmente aplicados nos estágios iniciais do processo de desenvolvimento de novos fármacos, quando há pouca informação acerca de um determinado alvo biológico de interesse. Uma vez que um composto matriz é identificado, métodos mais sofisticados podem ser usados para melhorar as propriedades do composto matriz inicial. [137]

O primeiro método de busca por similaridade química foi baseado em propriedades bidimensionais. Apesar de este método prover estruturas com alto grau de similaridade estrutural, a relação entre a estrutura e a atividade biológica é complexa, o que levou ao desenvolvimento de um vasto conjunto de métodos de similaridade. O primeiro método 2D foi baseado na descrição das moléculas ao nível atômico e das ligações químicas. Trabalhos subsequentes focaram a descrição das propriedades atômicas ao invés da descrição atômica elementar, o que levou ao desenvolvimento de diversos descritores moleculares.

Há também um considerável interesse em métodos de representação tridimensional (3D) da estrutura química. Os métodos 3D podem ser divididos em alinhamento independente, onde as moléculas são tipicamente representadas por vetores que usam técnicas similares aos métodos 2D, como o coeficiente de Tanimoto. Os métodos que requer maior demanda computacional são os que exigem um alinhamento prévio antes do cálculo da similaridade. A aplicação de análises de similaridade química no processo de desenvolvimento de fármacos é uma técnica amplamente difundida. Existem vários métodos de topologia bi e tridimensional (2D e 3D) para a determinação da similaridade química. Métodos que utilizam superposição 3D são normalmente mais lentos que o 2D. Isto é devido em grande parte ao fato de métodos 3D requererem um acesso as energias conformacionais do arranjo molecular, ao passo que os métodos 2D somente utilizam uma estrutura única. Métodos 3D, entretanto, têm a vantagem de encontrar quimicamente estruturas que têm a mesma forma e propriedades químicas. [138]

A busca por similaridade inicia com uma estrutura que tenha atividade conhecida ou o próprio substrato de uma enzima, a qual é usualmente denominada como estrutura de referência. Uma representação desta estrutura (2D ou 3D) é utilizada para comparar contra um banco de dados por meio de um coeficiente de similaridade. O banco de dados analisado é posicionado em ordem decrescente em relação aos valores do coeficiente de similaridade, o composto com maior similaridade ficará em primeiro lugar na lista.

3.4 Relações Quantitativas Estrutura Atividade

As pesquisas na área de Relações Quantitativas Estrutura Atividade (QSAR - *Quantitative Structure-Activity Relationships*) têm como principal objetivo a construção de modelos matemáticos que relacionem a estrutura química e a atividade biológica de uma série de compostos análogos. A ação terapêutica de fármacos resulta de interações destes com sistemas biológicos e é dependente de fatores relacionados com sua estrutura química e, conseqüentemente, de suas propriedades físico-químicas. Estes fatores sejam caráter eletrônico, hidrofóbico ou estérico, influenciam a interação do fármaco com a biofase e a sua distribuição nos compartimentos que compõem o sistema biológico. Assim, dois fármacos estruturalmente semelhantes, diferenciando-se apenas por um átomo ou posição que este ocupa na molécula, podem apresentar diferenças quanto às suas propriedades físico-químicas e, conseqüentemente, quanto a atividade biológica, tanto do ponto de vista quantitativo como qualitativo. [139]

Métodos de QSAR têm sido amplamente aplicados na indústria farmacêutica para a descoberta de substâncias biativas, otimização, e a previsão de toxicidade. O desafio é selecionar o grupo de descritores que descrevem as características estruturais e físico-químicas mais críticas associadas com a atividade. Descritores apropriados ou correta seleção de variáveis é parte integrante do processo de modelagem de QSAR. Obtenção de um modelo de QSAR de boa qualidade depende de muitos fatores, tais como a qualidade dos dados biológicos, a escolha de descritores e os métodos estatísticos adotados. [139-140]

O processo de desenvolvimento de modelos QSAR pode ser geralmente dividido em três etapas: preparação de dados, análise de dados e validação do

modelo. A primeira fase inclui a seleção de um conjunto de dados moleculares para estudos de QSAR, cálculo de descritores moleculares e seleção de um método de QSAR (análise estatística e correlação). Estes passos representam uma prática normal de qualquer modelagem de QSAR, e seus detalhes específicos são geralmente determinados pelos interesses dos pesquisadores e disponibilidade de software.

Uma etapa fundamental do desenvolvimento de um modelo QSAR é a validação. A maioria dos métodos de modelagem de QSAR aplica o método “deixe um de fora” (*leave-one-out*) como procedimento de validação cruzada. Neste procedimento, cada composto é sistematicamente excluído uma vez do conjunto, e em seguida a resposta biológica é predita com o restante dos compostos. Com as atividades preditas dos compostos “de fora” é possível obter o valor do coeficiente de correção da validação cruzada (q^2), utilizado para avaliar a capacidade de predição dos modelos.

Muitos autores consideram de alto q^2 (por exemplo, $q^2 > 0,5$) como um indicador ou mesmo como a prova definitiva de que o modelo é altamente preditivo. Uma abordagem amplamente utilizada para estabelecer a robustez do modelo é o chamado aleatorização. Trata-se de repetir o procedimento de cálculo com atividades aleatorizadas e posterior avaliação de probabilidade das estatísticas resultantes. Muitas vezes, é utilizado juntamente com a validação cruzada.

Além dos parâmetros estatísticos do modelo a validação da capacidade de predição é necessário utilizar estruturas com atividade conhecida que não foram utilizadas para construir o modelo para avaliar a capacidade de prever com precisão a atividade biológica. [141-142]

3.5 Materiais

Os ensaios virtuais e os estudos de modelagem molecular foram realizados no Laboratório de Quiminformática e Modelagem Molecular do Grupo de Estudos em Química Medicinal de Produtos Naturais – NEQUIMED-PN. Os trabalhos foram desenvolvidos em uma estação de trabalho Linux Fedora 10 utilizando os programas listados abaixo:

- FILTER2.0.2, MAKEFRAGLIB2.0.2, OMEGA2.0.2, FRED2.2.5, ROCS2.4.1 da empresa Openeye Scientific Software;
- FlexX2.0.2, HQSAR da empresa Trypos operando na plataforma Sybyl8.0;
- Glide4.5 da empresa Schrödinger operando na plataforma Maestro6.0

3.6 Métodos

3.6.1 Filtros moleculares

O filtro fármaco-similar utilizado com modificações para incluir também moléculas matrizes similares. Os parâmetros utilizados foram: massa molecular entre 200 e 600 Daltons, número de átomos pesados 15 a 35, número de anéis 0 a 5, número de átomos conectados em átomos não-cíclicos 0 a 15, número de grupos funcionais 0 a 18, átomos de carbono 7 a 35, átomos não ramificados 0 a 6, heretoátomos 2 a 20, percentagem de halogênios 0,5%, graus de liberdade da molécula 0 a 4, ligações rotacionais 0 a 20, ligações rígidas 0 a 35, número de doações de ligação de hidrogênio 0 a 6, número de receptores de ligação de hidrogênio 0 a 8, regra de Lipinski para ligação hidrogênio, carga formal -2 a +3, número de centro quirais 0 a 4, coeficiente de partição calculado (XlogP) 0 a 6, compostos solúveis e altamente solúveis, filtros de eliminação de compostos que formam agregados (conhecidos e preditos), eliminação de metais pesados e grupos funcionais indesejados como quinonas, alquil fosfatos, corantes etc. [143]

3.6.2 Docagem Molecular

A estrutura 3D da enzima utilizada para a docagem molecular foi preparada utilizando o módulo Preparação das Estruturas de Biopolímeros do programa Sybyl8. 0. O ligante e as moléculas de água foram extraídos e todos os hidrogênios foram adicionados à estrutura da enzima, considerando-se os estados de protonação dos resíduos de aminoácidos apropriados. Os resíduos Cys25 e His159, pertencentes à tríade catalítica, foram mantidos neutros e o hidrogênio azo foi posicionado no nitrogênio delta da histidina, uma vez que a análise estrutural demonstra uma ligação de hidrogênio entre este resíduo hidroxila do ligante α -hidróxi-cetona cristalizado. [144]

Com o intuito de explorar diversos algoritmos de busca, os programas de docagem molecular FRED (v2.2.5 – algoritmo de busca multiconformacional), FlexX (v2.0.2 - algoritmo de construção incremental), e Glide (v4.5 – algoritmo de busca estocástica com Monte Carlo) foram utilizados. Inicialmente foram realizadas etapas de redocagem e docagem cruzada para avaliar a capacidade de reprodução das estruturas cristalográficas dos complexos 1ME4 e 1ME3 com os diferentes programas. Os parâmetros com melhores resultados para cada programa são descritos a seguir:

FlexX: O procedimento de docagem foi o padrão do programa com algumas modificações; o sitio ativo definiu-se uma região esférica da proteína englobando todos os resíduos em um raio de 6,5 Å a partir das coordenadas do ligante. Três restrições farmacofóricas foram necessárias para obter um RMSD inferior a 2 Å, hidrofóbica no resíduo Leu67, aceitador de ligação de hidrogênio na Gly66 e doador de ligação hidrogênio em His159. Estas restrições foram inseridas após uma análise de todos os co-complexos presentes do PDB revelar que são interações presentes na maioria dos mesmos, e desta forma essenciais para o reconhecimento molecular. Somente a primeira molécula da lista e sua respectiva pose foram mantidas nesta análise. [145]

FRED: Os parâmetros padrões foram utilizados com as seguintes modificações; os mapas de contorno interno e externo foram desabilitados, desta forma o raio de van der Waals dos átomos do sitio ativo e a caixa foram os limites espaciais para a docagem. A caixa foi adicionada para cobrir todos os resíduos de aminoácidos a uma distância de 5 Å a partir das coordenadas do ligante de maneira a incluir todos os resíduos do sitio ativo da enzima. A seleção das poses realizou-se através de pontuação consensual utilizando as funções PLP, chemgauss3 e ochemscore. A pontuação final para a predição da afinidade de interação do complexo entre a enzima e as moléculas do banco de dados foi utilizada todas as funções disponíveis no programa (PLP, chemgauss3, ochemscore shapegauss, chemgauss2, chemscore, screenscore e zapbind). As mesmas restrições incluídas no programa FlexX foram utilizadas neste programa para obter RMSD < 2 Å na redocagem.

Os programas FlexX e FRED são mais rápidos em relação ao programa Glide, desta forma foram utilizados na sub-coleção proveniente da saída do filtro molecular contendo cerca de 3,5 milhões de estruturas químicas.

Glide: Os parâmetros utilizados foram os padrões de precisão extra (Glide XP), a caixa do receptor foi gerada com dimensões de 30×30×30 Å³, centrada nas coordenadas do ligante complexado. As mesmas restrições incluídas no programa FlexX foram utilizadas neste programa para obter RMSD < 2 Å na redocagem. O programa Glide foi aplicado somente na subcoleção da seleção consensual devido ao seu alto custo computacional.

3.6.3 Similaridade química

Uma das mais importantes decisões em busca por similaridade é a escolha da molécula de referência. Normalmente um composto conhecido de alta potência é utilizado para a busca, outra escolha é uma molécula seletiva entre alvos ortogonais ou ainda o próprio substrato da enzima quando não há ligantes identificados. Neste trabalho foram utilizadas três estruturas de potentes inibidores da enzima cruzaina como estruturas de referência para a busca nos bancos de dados. O primeiro deles foi o K11777 - um inibidor que se encontra na fase pré-clínica em estágio mais avançado na gênese; é um inibidor irreversível da cruzaina com baixa toxidez *in vivo* [77]. Um complexo deste inibidor com a enzima cruzaina está disponível no banco de dados de proteínas (código PDB 20Z2), de onde a estrutura foi extraída com a conformação que a molécula assume no sítio ativo da enzima (assumida como conformação “ativa”).

O segundo inibidor escolhido como molécula de referência foi o inibidor α -hidróxi-cetona, e vai ser referido aqui como inibidor_1ME4, que possui disponível um complexo da enzima com este composto (código PDB 1ME4) [87]. Este inibidor é reversível, o mecanismo de inibição desejável para obter substâncias com menor toxidez quando comparado com inibidores irreversíveis.

A terceira estrutura de referência escolhida é de um inibidor com potência picomolar (chamado aqui de 92260) e o mais seletivo em relação à catepsina-L, cerca de 100 vezes. [146]

Os três inibidores utilizados como referência foram testados contra este banco, onde a coleção multiconformacional foi gerada com o programa OMEGA2, como descrito na seção 3.6.4. Este teste foi utilizado para definir os valores das métricas de similaridade utilizadas pelo programa ROCS que definem um composto ativo contra a enzima utilizando como padrão a potência de inibidores conhecidos.

3.6.4 Construção da Coleção Multiconformacional

Os programas da Openeye ROCS (similaridade) e FRED (docagem) necessitam de uma biblioteca conformacional das moléculas provenientes do banco de dados. Os programas MAKEFRAGLIB e OMEGA são utilizados para a construção da Biblioteca Multiconformacional, o primeiro fragmenta a molécula em ligações livres (onde há rotação) e o segundo constrói a molécula em diferentes conformações.

Os parâmetros utilizados para gerar a coleção multiconformacional foram os padrões do OMEGA2, exceto os parâmetros de construção 3D: onde o campo de força utilizado foi o mmff94s (Merck Molecular Force Field); parâmetros de “torção” das ligações: janela de energia utilizada 25 kcal mol⁻¹, desvio quadrático médio 0,6, número máximo de confômeros 500. [147]

3.6.5 Relações Quantitativas Estrutura Atividade Baseada em Hologramas Moleculares

A série de moléculas utilizadas para a construção do modelo foi retirada da literatura e contém 41 inibidores, os quais foram divididos em conjunto de treinamento (32 moléculas) e teste (9 moléculas). Os valores da constante de afinidade K_i foram convertidos em pK_i ($-\log K_i$) antes de serem utilizados como variáveis dependentes na análise de HQSAR. Adicionalmente, uma série de 10 moléculas retiradas do trabalho de Huang et al. Foram utilizadas para a validação externa dos modelos.

Os hologramas moleculares foram gerados utilizando várias combinações de fragmentos e distinções de fragmentos, incluindo átomos (A), ligações (B), conectividade (C), hidrogênio (H), quiralidade (Ch) e doador e aceitador de ligação hidrogênio (DA). O tamanho dos fragmentos, foi utilizado o padrão (4-7) para gerar os modelos iniciais, os melhores modelos foram submetidos à variação do tamanho do fragmento (2-5, 3-6, 5-8, 6-9 e 7-10). Os comprimentos dos hologramas variaram entre 53 a 401 blocos. Todos os modelos foram produzidos usando a análise de regressão multivariada por mínimos quadrados parciais (PLS) correspondendo ao

menor erro padrão da predição (SEP) foram determinados usando procedimento de validação cruzada ,“deixe um de fora”.

3.6.6 Avaliação dos modelos computacionais

Os métodos computacionais foram avaliados através da capacidade de classificação de inibidores conhecidos de cruzaina em uma coleção de moléculas. Para isto, selecionou-se 27 inibidores com potência elevada e 326 moléculas retiradas do banco de dados ZINC com propriedades moleculares 1D semelhantes aos inibidores (massa molecular, número de ligações rotacionáveis, log P calculado, número de doadores e aceptores de ligação hidrogênio, etc), utilizando o programa FILTER. Estas moléculas foram definidas como inativas, apesar de, nunca serem testadas contra a enzima. A coleção teste totalizou 353 estruturas.

Os programas de docagem molecular, similaridade química e o modelo de HQSAR foram utilizados para discriminar os inibidores da cruzaina dentro da coleção de teste. Foi definido de valores limites e quais métricas de cada programa possuem melhor desempenho para a seleção de moléculas proveniente do banco de dados do ZINC a serem testadas nos ensaios bioquímicos.

3.6.7 Método de avaliação dos modelos computacionais

A análise utilizada para avaliação dos métodos foi a curva ROC (Receiver Operating Characteristcs). Ela é utilizada para organizar e classificar de forma a visualizar o desempenho do método. As curvas ROC são tradicionalmente utilizadas em decisões médicas, e nos últimos anos tem sido aplicada no descobrimento de novos fármacos, principalmente em ensaios virtuais de larga escala. Esta análise é realizada dividindo a coleção de compostos em quatro classes: compostos ativos recuperados pelo método (verdadeiros positivos), compostos inativos recuperados pelo método (falsos negativos), compostos ativos não recuperados (falsos negativos) e compostos inativos não recuperados (verdadeiros negativos). [148-149]

A taxa de verdadeiros positivos recuperados (TP_{rate}) pelo método é definida por:

$$TP_{rate} = \frac{N_{\text{verdadeiros_positivos}}}{N_{\text{ativos_total}}} \quad (3.7)$$

A taxa de falsos positivos recuperados (FP_{rate}) pelo método é definida por:

$$FP_{rate} = \frac{N_{falsos_positivos}}{N_{inativos_total}} \quad (3.8)$$

A sensibilidade do método é definida como a taxa de verdadeiros positivos recuperados, e a especificidade são definidas por:

$$Especificidade = \frac{Verdadeiros_negativos}{falsos_positivos + verdadeiros_negativos} \quad (3.9)$$

A curva ROC é a sensibilidade em função da especificidade do método, e pode ser expressa como a taxa de verdadeiros positivos encontrados em função da taxa de falso positivos. A área sob a curva ROC (AUC) é uma medida quantitativa que relaciona a métrica utilizada para a determinação do coeficiente de similaridade com a taxa de recuperação dos verdadeiros positivos.

3.7 Resultados e Discussões

A pesquisa de informações disponíveis na literatura possibilitou definir uma estratégia para a seleção de moléculas para os testes bioquímicos. Uma vez que a estrutura tridimensional da enzima e inibidores são conhecidas, foram utilizados métodos de ensaios virtuais baseados na estrutura do ligante (LBVS) e do alvo (TBVS). Um esquema geral das etapas envolvidas para a seleção das moléculas é apresentada na FIGURA 3.3.

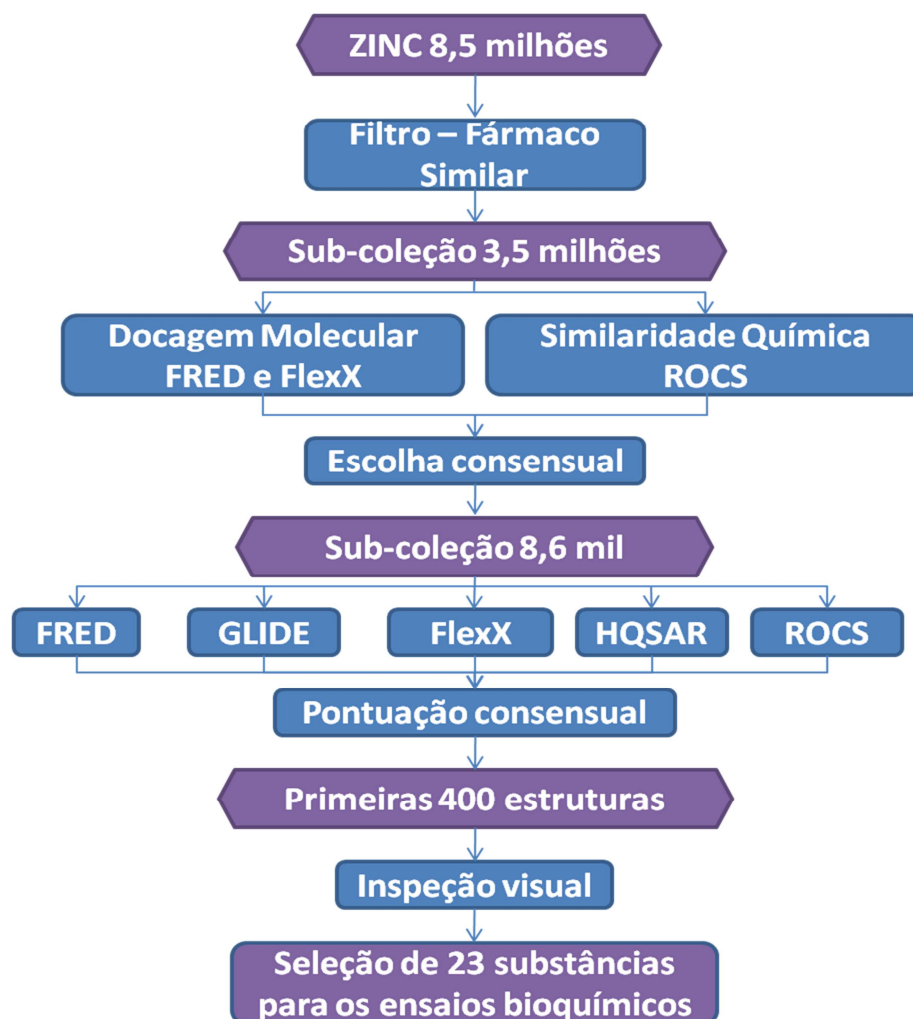


FIGURA 3.3: Fluxograma das etapas envolvidas no processo de seleção de moléculas para os ensaios bioquímicos

Resumidamente, as etapas desenvolvidas para a seleção de moléculas, como ilustrada na FIGURA 3.3, foram: estruturar uma coleção virtual de compostos disponíveis comercialmente (ZINC), a qual continha cerca de 8,5 milhões de estruturas. Em seguida foi aplicado um filtro molecular com características fármaco-similar, e a coleção decresceu para uma sub-coleção 3,5 milhões de estruturas. Esta sub-coleção foi submetida à docagem molecular com os programas FRED e FlexX e similaridade química com o programa ROCS; em seguida uma escolha consensual destes métodos foi utilizada para definir a sub-coleção final com 8 635 moléculas, o critério utilizado foi o composto ter sido recuperado nos três programas ROCS, FRED e FlexX.

A sub-coleção final foi listada através de pontuação consensual das métricas de similaridade do ROCS, funções de pontuação dos programas de docagem e valores de pK_i calculados por um modelo de HQSAR construído com inibidores

conhecidos da enzima cruzaina. As primeiras 400 estruturas foram inspecionadas visualmente para seleção dos compostos a serem testados nos ensaios bioquímicos. Os detalhes de cada etapa serão discutidos ao longo deste capítulo.

3.8 Filtros Moleculares

O programa utilizado para aplicar os filtros de seleção neste trabalho foi o FILTER da Openeye [143]. Este programa tem disponível um filtro padrão que envolve muito dos princípios de filtragem tradicionais, tais como remover substituintes tóxicos e reativos. Os critérios usados pelo FILTER para que uma dada molécula seja selecionada ou descartada dividem-se em três categorias: propriedades físicas, conteúdo atômico, grupos funcionais e propriedades topológicas.

As propriedades físicas simples incluem massa molecular, área superficial, $\log P$ e solubilidade em água. Os filtros também incluem o conteúdo relativo e absoluto de heteroátomos, assim como, os limites do número de uma ampla variedade de grupos funcionais. Os filtros topológicos apontam questões a respeito do número e tamanho dos anéis, a flexibilidade da molécula, o tamanho e a forma de cadeias alifáticas. [40]

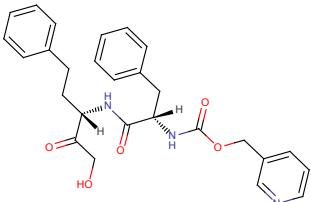
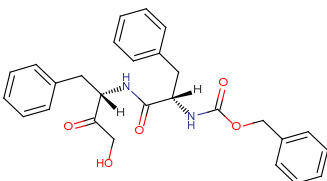
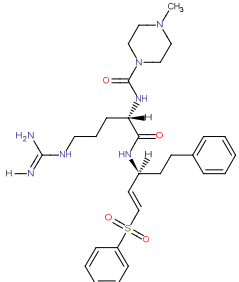
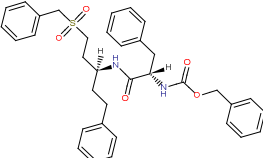
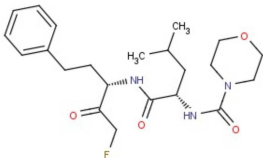
O filtro fármaco-similar foi utilizado para filtrar a coleção de moléculas, pois contempla as substâncias que se adequam as características moleculares definidas por Lipinski e Veber e possuem maior chance de apresentarem características farmacocinéticas apropriadas. Além disto, foram incluídas características 1D de inibidores conhecidos da enzima cruzaina, como o aumento das ligações com rotação livre. A solubilidade é outro fator crítico, principalmente nos ensaios iniciais e por isso o filtro para moléculas com alta solubilidade predita foi selecionado.

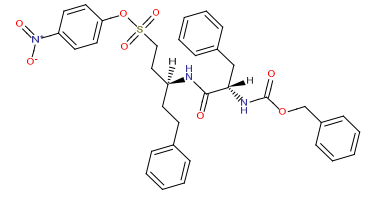
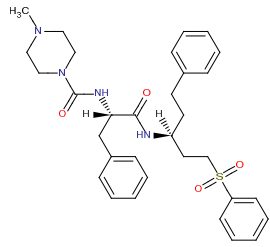
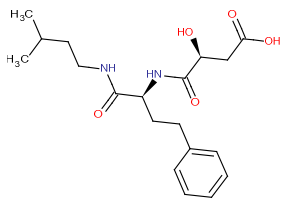
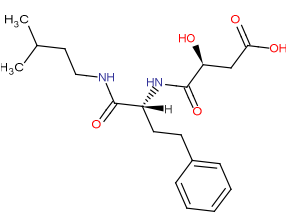
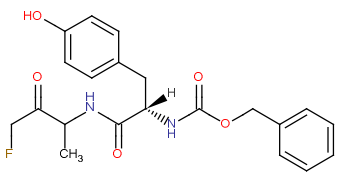
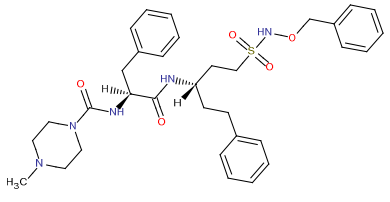
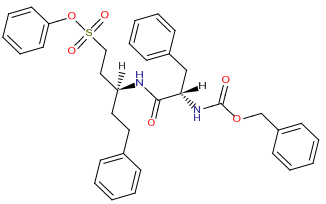
O número de compostos remanescentes após a aplicação do FILTER decresceu de 8,5 para 3,5 milhões de estruturas moleculares. Esta sub-coleção foi submetida aos métodos de busca de similaridade química e docagem molecular para seleção das moléculas para os ensaios bioquímicos.

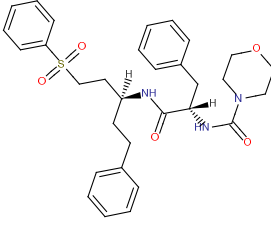
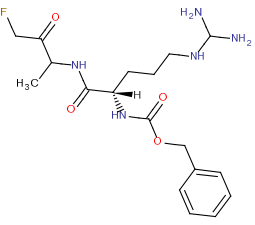
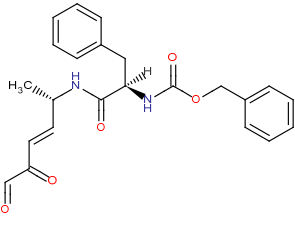
3.9 Docagem Molecular

Inicialmente buscou-se no banco de dados de proteínas (PDB) por estruturas da enzima cruzaina complexadas com inibidores, um total de 15 estruturas depositadas foi encontrado e os dados da qualidade da resolução estão apresentados na TABELA 3.1. [150]

TABELA 3.1. Código PDB e parâmetros de resolução das estruturas da cruzaina

Código PDB	Resolução (Å)	R _{factor}	R _{free}	(R _{free} - R _{factor})	Ligante co-cristalizado
1ME3	1,2	0,1	0,14	0,032	
1ME4	1,2	0,09	0,12	0,033	
2EFM	1,5	0,15	0,19	0,042	
1F2A	1,6	0,18	0,22	0,035	
1EWP	1,8	0,18	0,21	0,023	

1F2B	1,8	0,17	0,2	0,030	
2OZ2	2	0,16	0,21	0,048	
1EWM	2	0,17	0,21	0,033	
1EWL	2	0,17	0,23	0,062	
1AIM	2	0,19	0,25	0,059	
1F2C	2	0,18	0,23	0,050	
1WEO	2,1	0,22	0,26	0,041	

1F29	2,2	0,19	0,21	0,023	
2AIM	2,2	0,19	0,25	0,063	
1U9Q	2,3	0,14	0,22	0,082	

Um importante parâmetro nos estudos cristalográficos é a resolução dos dados, que é expressa em Å, onde números menores indicam resoluções mais altas. Em altas resoluções ($< 1,5 \text{ \AA}$) o modelo é provavelmente mais que 95% uma consequência dos dados observados. Contudo, em baixas resoluções ($> 2,5 \text{ \AA}$), a modelagem dos detalhes nas estruturas das proteínas é muito mais subjetiva do que baseada em informações experimentais. [151]

A habilidade do modelo em explicar os dados experimentais é usualmente identificada por meio do parâmetro chamado R_{factor} . Contudo, como esse parâmetro está sujeito ao ajuste forçado (*over-fitting*) dos dados, Brünger introduziu o conceito de validação cruzada no refinamento cristalográfico com a introdução do parâmetro R_{free} , cuja definição é idêntica à do R_{factor} , exceto que o R_{free} é calculado para um pequeno conjunto de reflexões que nunca são usadas no refinamento do modelo. Portanto, o parâmetro R_{free} mede a capacidade do modelo em prever observações experimentais que não são usadas para ajustar o modelo. Os valores desejáveis são $< 0,25$ para o R_{factor} e $< 0,40$ para R_{free} . A diferença entre os parâmetros R_{free} e R_{factor} é uma medida parcial da extensão na qual o modelo superestima os dados. Esta diferença ($R_{\text{free}} - R_{\text{factor}}$) deve ser baixa para o modelo final, idealmente $< 0,05$. [152]

Baseado nesses fatores, várias estruturas têm características apropriadas para os ensaios virtuais. As estruturas com melhor resolução são as PDB 1ME3 e 1ME4. Outra característica a ser ressaltada é que os inibidores co-cristalizados destas estruturas são reversíveis, contrariamente a todas outras estruturas que apresentam inibidores irreversíveis. A superposição das 15 estruturas cristalográficas permitiu identificar interações intermoleculares essenciais para o reconhecimento molecular e são mostrados na FIGURA 3.4.

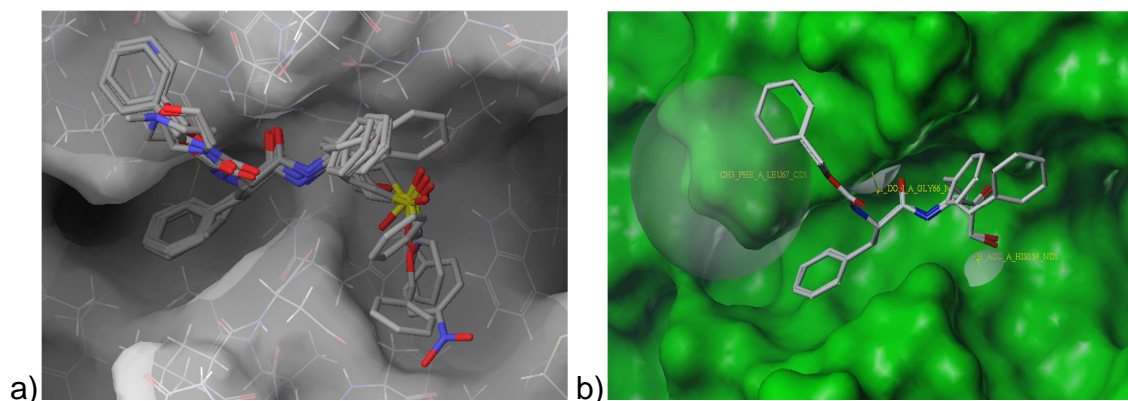


FIGURA 3.4: a) Superposição das estruturas cristalográficas dos 15 co-complexos provenientes do PDB e b) restrições farmacofóricas utilizadas demonstrando a complementaridade com os ligantes das estruturas 1ME3 e 1ME4.

Os sub-sítios S2 e S3 possuem característica hidrofóbica e a Leu67 na interface entre eles, este aminoácido possui complementaridade com todos os ligantes co-cristalizados com a enzima cruzaina. Na região central do sítio ocorre o reconhecimento da amida dos ligantes, e a Gly66 interage via ligação de hidrogênio da cadeia principal com o oxigênio da amida dos ligantes. No sítio S1 estão localizados os aminoácidos da tríade catalítica e fazem importantes ligações de hidrogênio. Desta forma, os aminoácidos Leu67, Gly66 e His159 foram selecionados como restrições farmacofóricas nos programas de docagem, a redocagem e docagem cruzada apresentaram melhores resultados com as restrições e, além disso, as informações disponíveis de ligantes conhecidos podem ser utilizadas para a seleção de novos ligantes.

Além disso, uma análise comparativa da identidade da cadeia principal de aminoácidos foi realizada com enzimas pertencentes à família das proteases de diferentes organismos utilizando o servidor BLASTp. As cinco enzimas com maior identidade são apresentadas na FIGURA 3.5.

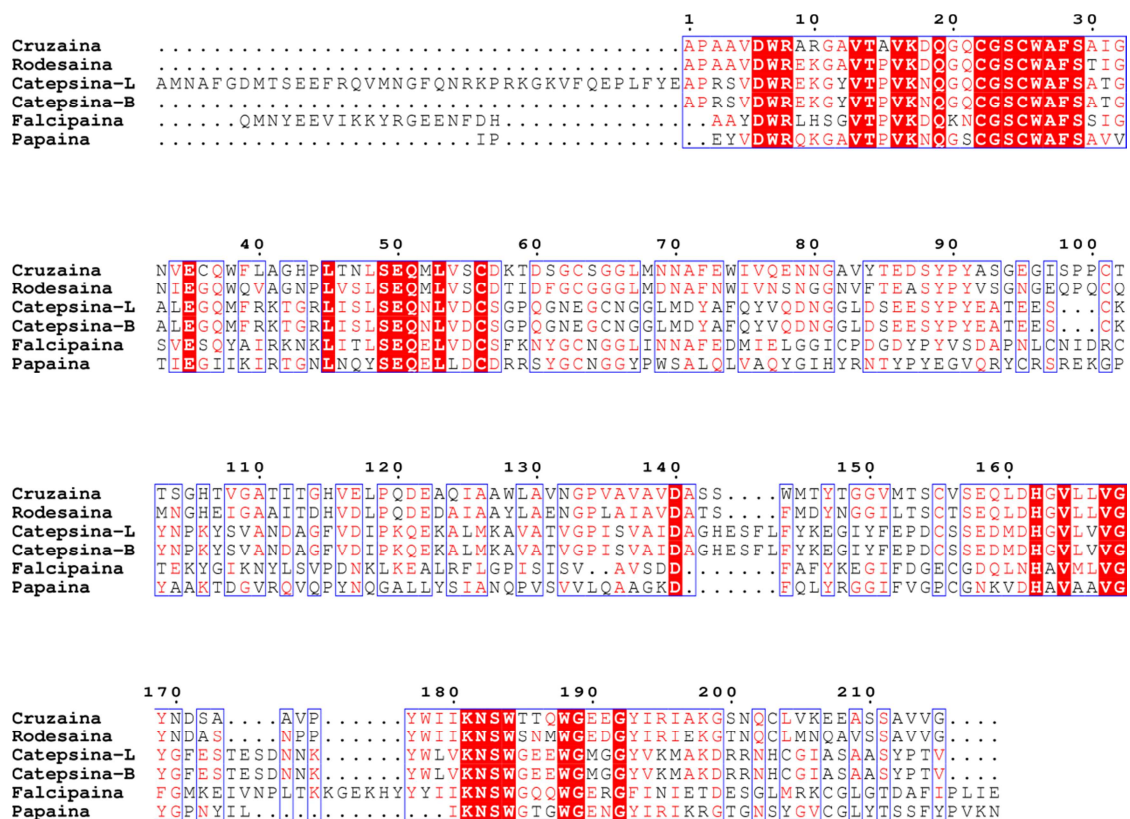


FIGURA 3.5. Alinhamento das estruturas primárias de enzimas proteases de diferentes organismos com a cruzaina. As letras em branco com fundo vermelho representam conservação dos aminoácidos em todas as estruturas; as letras vermelhas representam conservação em pelo menos três estruturas. O alinhamento foi feito no servidor MultiAlign.

A cruzaina apresenta as seguintes idêntidades com proteases: rodessaína pertencente ao parasito *Trypanosoma brucei* 70 %, falcipaína pertencente ao parasito *Plasmodium falciparum* 40 %, papaína pertencente a *Carica papaya* 39 % e as enzimas de humanos catepsinas L e B 43 % em ambas.

O elevado grau de identidade entre estas estruturas sugere que inibidores da enzima cruzaina potencialmente podem apresentar atividade nestas enzimas. Hipoteticamente, um fármaco para a doença de Chagas, atuando pelo mecanismo de ação da inibição da enzima cruzaina, poderia apresentar atividade biológica também para a doença do sono, malária, câncer ou osteoporose, uma vez que estas proteases são alvos para estas doenças.

3.9.1 Docagem com o programa FlexX

O programa FlexX, trata o ligante flexível e o receptor rígido. O programa divide o ligante ao longo de suas ligações simples em fragmentos, ajusta primeiro um fragmento base no sitio ativo e então conecta os outros fragmentos. Ao invés de

identificar pontos no receptor onde o ligante pode se localizar, o programa FlexX define os chamados pontos de interação para cada possível grupo de interação do ligante e do receptor. Tendo designado os sítios de interação, um tipo de interação (receptor de ligação de hidrogênio, doador de ligação de hidrogênio, etc.) é modelado através de uma geometria de interação consistindo de um centro de interação e uma superfície esférica. Uma interação entre dois grupos A e B ocorre se: i) os tipos A e B são compatíveis; ii) o centro de interação de A se posiciona sobre a superfície de interação de B, e vice versa. Desse modo o programa FlexX distingue entre os seguintes tipos de interação: ponte salina, ligações de hidrogênio, complexos metálicos, interações específicas entre o centro de anéis aromáticos com alguns grupos e contatos hidrofóbicos não específicos. [153]

A primeira etapa foi redocar o inibidor da estrutura PDB 1ME4 na proteína para avaliar a capacidade do programa reproduzir a estrutura cristalográfica. Em seguida fez-se uma docagem cruzada do inibidor da estrutura 1ME3 na enzima da estrutura 1ME4, estas duas estruturas foram selecionadas por serem os únicos inibidores reversíveis, que não apresentam ligação covalente com a enzima. Ambos os inibidores apresentam vários graus de liberdade na molécula. Isto representa a principal limitação para o programa em gerar a conformação e o modo de interação correspondente à estrutura cristalográfica: o número de confôrmeros aumenta exponencialmente com o aumento do número de rotações livres da molécula.

O programa FlexX foi capaz de reproduzir as estruturas cristalográficas satisfatoriamente. O valor recomendado para desvio quadrático médio (RMSD) é inferior a 2 Å [154]. Este mede o desvio da posição dos átomos do inibidor da estrutura cristalográfica com relação à posição predita pelo programa de docagem. A redocagem apresentou um valor de RMSD de 1,247 Å, FIGURA 3.6.

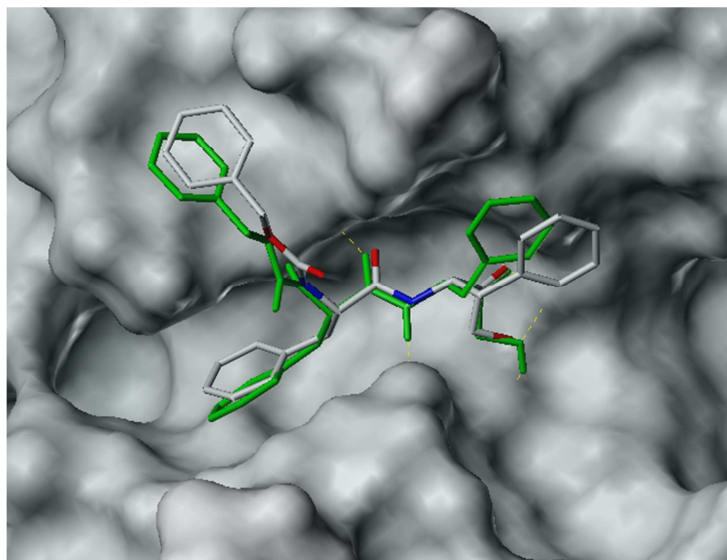


FIGURA 3.6: Estrutura do ligante proveniente da estruturas 1ME4 docado pelo programa FlexX na enzima cruzaina (código PDB 1ME4), verde representa a pose do co-cristal e cinza a pose da docagem

A redocagem satisfatória é o requisito mínimo exigido de um programa de docagem para que possa ser empregado nos ensaios virtuais. Adicionalmente foi construída uma coleção de teste e o programa foi submetido a uma avaliação da capacidade de discriminar inibidores conhecidos em um banco de moléculas, a coleção de teste foi construída com inibidores conhecidos e moléculas provenientes do banco ZINC com propriedades 1D similares, como massa molecular, log P, número de ligações livres etc., que foram definidas como controle negativo. A capacidade de discriminação entre as moléculas ativas e inativas foram avaliadas com o uso da curva ROC. A FIGURA 3.7 mostra a avaliação da recuperação das moléculas ativas pelo programa FlexX com as quatro funções de pontuação empregadas pelo programa e as frequências do número de moléculas em função do valor da pontuação com bom desempenho.

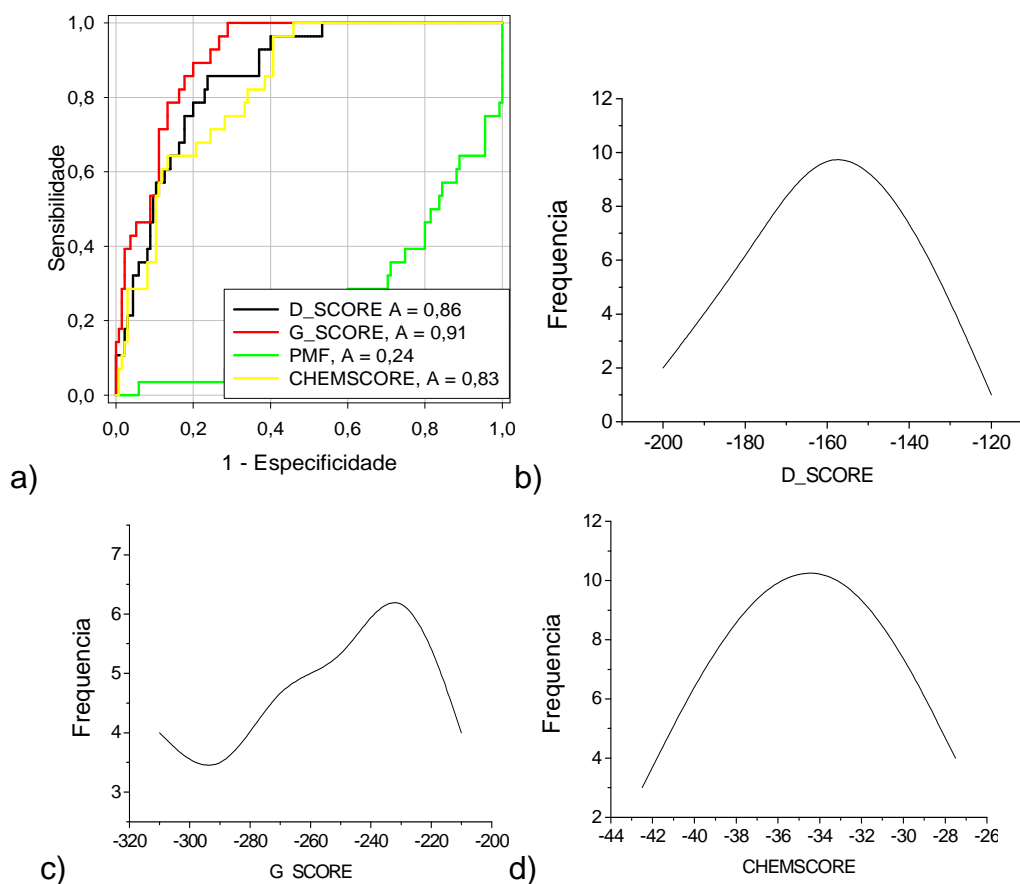


FIGURA 3.7: a) Curva ROC das funções de pontuação do programa FlexX e freqüência de moléculas em função de b) D_SCORE, c) G_SCORE e d) CHEMSCORE

Através da curva ROC [155] foi possível identificar as funções de pontuação que tiveram uma boa taxa de recuperação, quando as moléculas ativas são recuperadas nas primeiras posições a curva possui formato vertical e área sob a curva com valor igual a 1 significa recuperação de todas as moléculas ativas, valores igual ou menor que 0,5 significa que não há recuperação das moléculas ativas. Analisando a FIGURA 3.7(a) é possível identificar que somente a função PMF não é capaz de pontuar corretamente os inibidores da cruzaina, pois possui área sob a curva ROC 0,24, e por esta razão foi excluída da análise consensual.

Esta análise permitiu estabelecer os valores das funções de pontuações para aqueles inibidores conhecidos da cruzaina, e estabelecer os limites mínimos para que as moléculas provenientes do banco ZINC tivessem a possibilidade de serem testadas *in vitro* no ensaio enzimático. Os valores limite para as funções de docagem foram definidos como a menor energia (interação mais favorável) determinada para o conjunto de inibidores conhecidos.

A sub-coleção de 3,5 milhões de moléculas foi docada utilizando o programa FlexX e somente as moléculas com energia de interação calculada pela docagem superior aos limites estabelecidos foram mantidas para a análise consensual posterior. A TABELA 3.2 sumariza a capacidade de recuperação e moléculas ativas no banco conhecido, o valor de energia estabelecido como corte e o número de moléculas remanescentes após a docagem da sub-coleção do ZINC proveniente da saída do FILTER.

TABELA 3.2: Valores da área sob a curva ROC, limites de energia estabelecidos para as funções de pontuação e número de moléculas encontradas no ZINC acima dos limites

Função de Pontuação	Área sob a curva ROC	Valor limite estabelecido	Nº de moléculas do ZINC acima do limite
D_SCORE	0,86	- 200	3359
G_SCORE	0,91	- 300	1642
CHEMSCORE	0,83	- 042	4286
PMF	0,24	ND	ND

As funções de pontuação do programa FlexX apresentaram uma capacidade de recuperação das moléculas ativas superior a 80 % no banco construído com moléculas ativas e inativas utilizadas para a calibração do método. Os valores de energia estabelecidos foram elevados e a docagem de uma coleção contendo cerca de 3,5 milhões de estruturas docadas na enzima cruzaina resultaram em um número extremamente reduzido de compostos com valores de energia da docagem melhores que os limites estabelecidos. Entretanto, as estruturas mantidas possuem, segundo o programa FlexX, uma alta complementaridade com o sitio da enzima cruzaina com promissor potencial de apresentar atividade biológica nos ensaios *in vitro*. Algumas estruturas selecionadas pelo programa FlexX são apresentadas na FIGURA 3.8.

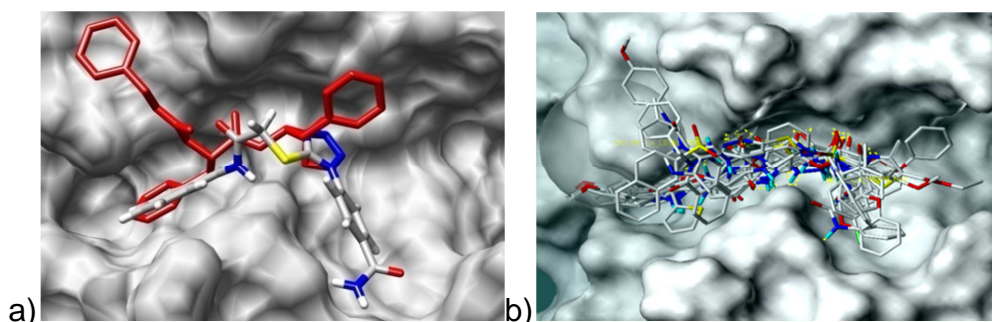


FIGURA 3.8: a) sobreposição de um ligante docado sobre a estrutura do inibidor da estrutura cristalográfica 1ME4 (em vermelho) e b) vários inibidores no sítio da enzima cruzaina

Pode ser observado na FIGURA 3.8(b) que os ligantes bem pontuados pelo programa FlexX ocuparam parcialmente a mesma posição do inibidor da estrutura cristalográfica em que foram docados. Os sub-sítios S1 e S2 foram ocupados pelos ligantes docados de maneira satisfatória, entretanto os sub-sítios S3 e S4 não mostram ocupação pelos ligantes com melhor energia de interação predita. Apesar da ocupação do sítio ser parcial a energia de docagem predita foi superior aos ligantes conhecidos, além disso, as classes moleculares selecionadas são distintas das previamente identificadas como inibidores contribuindo para aumentar a diversidade molecular de inibidores de cruzaína.

3.9.2 Docagem com o programa FRED

FRED (*Fast Rigid Exhaustive Docking*) é um programa de docagem rígida de pequenas moléculas em proteínas. A avaliação da flexibilidade do ligante requer uma coleção multiconformacional de cada molécula como arquivo de entrada. O processo de docagem do programa FRED é constituído de duas etapas: ajuste da forma e otimização. Na primeira etapa o ligante é colocado em uma caixa com uma resolução de 0.5 Å que engloba todos os átomos do sítio ativo utilizando um potencial Gaussiano. Em seguida uma série de três otimizações é processada: (1) refinamento da posição de átomos de hidrogênios de hidroxilas do ligante, (2) otimização do corpo rígido e (3) otimização da pose do ligante com relação ao ângulo diedro no espaço. Para a etapa de otimização, quatro funções de pontuação são disponíveis: forma de Gaussian, ChemScore, PLP e ChemScore para os três filtros de otimização. [156]

Os confôrmeros da cada molécula são gerados através do programa OMEGA2 (OpenEye Scientific Software). Esse programa é composto de dois componentes principais: modelo de construção e torção dirigida. Os componentes são independentes e podem ser usados dessa forma. O OMEGA2 constrói modelos iniciais das estruturas com base em padrões de ligações sigma. As conformações para cada fragmento molecular são pré-gerados usando o programa Makefraglib (OpenEye Scientific Software). A construção das moléculas é acompanhada por um vetor simples de alinhamento, uma vez que todos os fragmentos são conectados ao longo das ligações sigma. A partir do modelo inicial o OMEGA2 gera modelos

adicionais enumerando conformações de anéis e átomos de nitrogênio. O programa define todos os substituintes exocíclicos, alinha-os e os conecta com relação a cada conformação do ciclo. O OMEGA2 tenta gerar cada combinação possível de conformações prováveis para uma dada estrutura [147]. A FIGURA 3.9 ilustra um exemplo da conformação do inibidor hidroximetil cetona gerada pelo OMEGA2 superposta à conformação cristalográfica extraída do complexo da enzima cruzaina 1ME4.

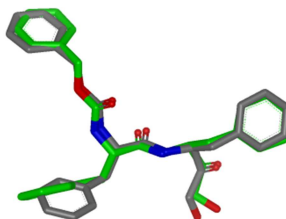


FIGURA 3.9: Superposição da conformação do inibidor hidroximetil cetona gerada pelo programa OMEGA2 (cinza) e a conformação cristalográfica extraída da estrutura da cruzaina (código PDB 1ME4) (verde)

O programa OMEGA2 foi capaz de reproduzir conformação correspondente à estrutura cristalográfica com um desvio quadrado médio dos átomos igual a 0,903. O inibidor hidroximetil cetona possui 13 ligações rotacionáveis e o número de possíveis conformações é extremamente elevado, mesmo assim, o programa OMEGA2 gerou uma conformação muito próxima da conformação cristalográfica, atestando a capacidade de reproduzir a possível conformação reconhecida pela enzima cruzaina. Desta forma, foi construída a biblioteca multiconformacional da subcoleção de 3,5 milhões de moléculas para a docagem com o programa FRED.

O resultado para a redocagem do inibidor da estrutura cristalográfica 1ME4 na cruzaina é apresentado na FIGURA 3.10.

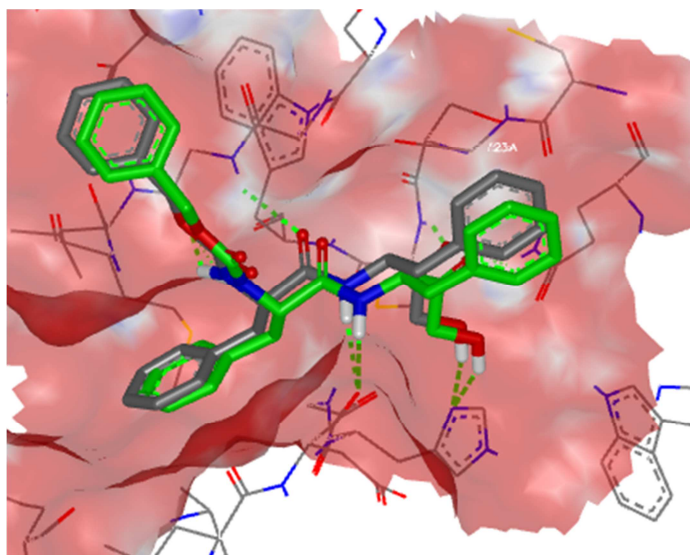


FIGURA 3.10: Estrutura do ligante proveniente da estruturas 1ME4 docado pelo programa FRED na enzima cruzaina (código PDB 1ME4), verde representa a pose do co-cristal e cinza a pose da docagem

O programa FRED foi capaz de reproduzir a estrutura cristalográfica satisfatoriamente com o de RMSD igual a 0,798 Å. A coleção de teste construída com inibidores conhecidos foi utilizada para avaliar a capacidade de recuperação das moléculas ativas e as funções de pontuação do programa FRED. A curva ROC é mostrada na FIGURA 3.11.

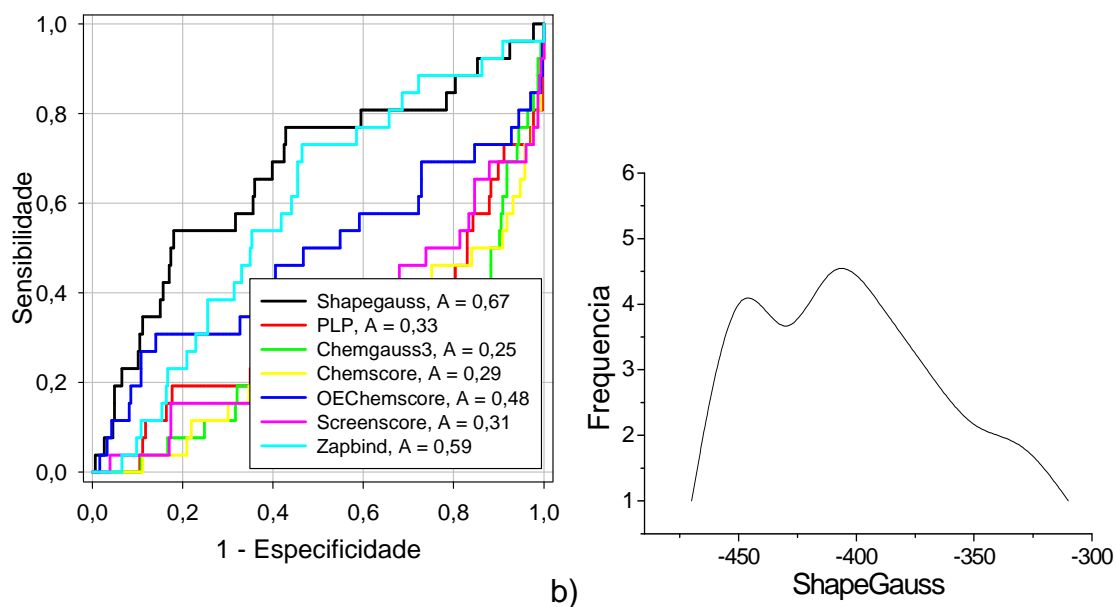


FIGURA 3.11: a) Curva ROC das funções de pontuação do programa FRED e freqüência de moléculas em função de b) Shapegauss

As funções de pontuação do programa FRED não apresentaram um bom desempenho, somente a função Shapegauss foi utilizada na busca por moléculas

bioativas na sub-coleção de 3,5 milhões de moléculas provenientes do ZINC. A curva de frequência de moléculas em função de Shapegauss mostra uma ampla distribuição de valores de energia calculada de interação, a TABELA 3.3 mostra os valores das áreas sob a curva ROC de todas as funções de pontuação o limite de energia estabelecido para a função Shapegauss e o número de moléculas encontradas com energia de interação calculada superior ao limite estabelecido.

TABELA 3.3. Valores da área sob a curva ROC, limites de energia estabelecidos para as funções de pontuação e número de moléculas encontradas no ZINC acima dos limites

Função de Pontuação	Área sob a curva ROC	Valor limite estabelecido	Nº de moléculas do ZINC acima do limite
SHAPEGAUSS	0,67	- 470	16151
PLP	0,33	ND	ND
Chemgauss3	0,25	ND	ND
Chemscore	0,29	ND	ND
OEChemscore	0,48	ND	ND
Screenscore	0,31	ND	ND
Zapbind	0,59	ND	ND

ND – não determinado

A FIGURA 3.12 mostra alguns exemplos de ligantes bem pontuados docados na estrutura da enzima cruzaina para análise da ocupação do sítio.

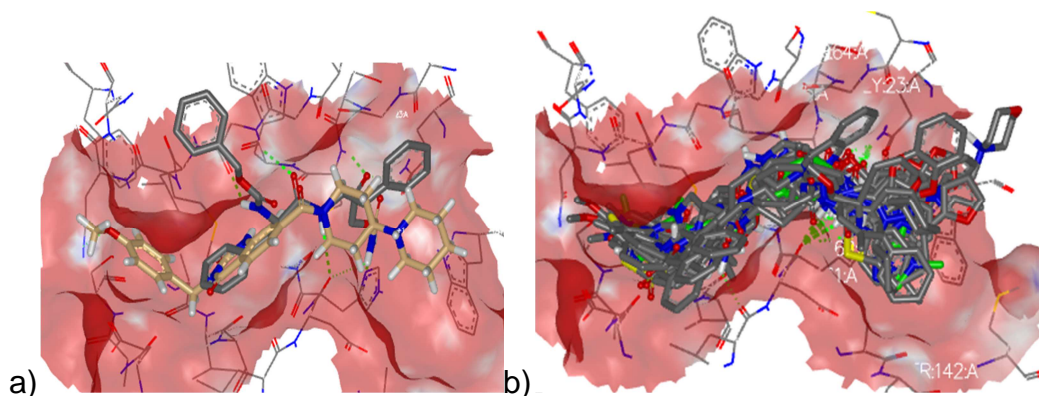


FIGURA 3.12: a) Estrutura tridimensional de um ligante bem pontuado pelo programa FRED superposto a estrutura cristalográfica (código PDB 1ME4) e b) vários ligantes bem pontuados pelo programa FRED demonstrando a ocupação do sítio ativo

As posições de docagem obtidas com o programa FRED foram similares àquelas obtidas com o programa FlexX, onde não houve ocupação dos sub-sítios S3 e S4 na estrutura da enzima, entretanto, alguns ligantes como mostrado na FIGURA 3.12(a) ocuparam uma pequena cavidade não explorada localizada em S2. Os programas FlexX e FRED foram aplicadas a sub-coleção 3,5 milhões de moléculas

para gerar uma nova sub-coleção reduzida com estruturas capazes de serem reconhecidas pelo sítio da enzima cruzaina e obedecerem as restrições farmacofóricas impostas para a seleção.

3.10 Similaridade Química

O método de similaridade empregado neste trabalho foi o 3D, através do uso do programa ROCS (Rapid Overlay of Chemical Structures). O ROCS é utilizado para identificar similaridade química entre moléculas baseado na sua forma tridimensional, e o recobrimento do volume.

Esta abordagem foi introduzida originalmente por Masek et al. que definiu o volume da molécula considerando-a como esfera fundida onde o raio era o atômico [157]. Na prática, o procedimento era lento e insuficiente para a determinação do mínimo global de energia devido ao uso de funções descontínuas para a descrição do volume molecular. Um passo crucial foi a introdução da descrição da forma molecular através de funções Gaussianas contínuas, centralizadas nos átomos. Isto foi utilizado para definir uma função com um comportamento muito similar a uma função de volume de uma esfera rígida, com uma predição total do volume de pequenas moléculas com erro inferior a 1%. [158]

Uma métrica bastante difundida para o cálculo da similaridade molecular é o coeficiente de Tanimoto, definido por:

$$Tanimoto_{A,B} = \frac{O_{A,B}}{O_A + O_B - O_{A,B}} \quad (3.10)$$

O coeficiente de Tanimoto tem reconhecidamente valor igual a um, se as duas formas são iguais, e zero se forem completamente diferentes. Duas moléculas nunca são completamente diferentes, entretanto podem apresentar formas iguais para duas moléculas diferentes.

Uma forma alternativa de medir a similaridade molecular é através do coeficiente de Tversky, ele é similar ao coeficiente de Tanimoto e definido por:

$$Tversky_{A,B} = \frac{O_{A,B}}{\alpha O_A + \beta O_B} \quad (3.11)$$

A principal diferença entre Tanimoto e Tversky é que o primeiro mede a diferença de similaridade considerando a unidade ($\neq 1$), o segundo não assume uma simples transformação, se A difere de B, e A e B são ambos diferente de zero, Tversky mede a representação da similaridade parcial da forma molecular. [159]

Normalmente $\alpha + \beta = 1$, e para o programa ROCS é assumido o valor 0,95; isto introduz uma assimetria. O coeficiente de Tversky depende de como a molécula se auto-recobre para definir um pré-valor de α . O ROCS calcula um pré-valor de α para a molécula de referência e um segundo valor para a molécula proveniente do banco de dados.

Outra métrica utilizada pelo programa ROCS é o campo de força *Color*. Cada composto assume um valor calculado pelo campo de força Color relacionado à natureza das ligações e átomos constituintes. A cada átomo diferente de carbono e hidrogênio é calculado um valor para a interação complementar entre os compostos, e ocorre um refinamento na superposição baseado na similaridade química.

Existe ainda uma métrica da junção de Tanimoto e Color, denominada *Combo*, pode assumir valores entre 0 e 2, sendo a métrica padrão para a listagem por similaridade.

O ROCS é um programa utilizado para a busca em grande escala de estruturas 3D em bancos de dados. Primeiramente é feito um alinhamento; as funções Gaussianas que simulam as moléculas como esferas rígidas são utilizadas para maximizar o recobrimento do volume molecular. O ROCS usa somente átomos pesados do ligante; os hidrogênios são ignorados. Uma vez que forma e volume molecular são relacionados, a maximização do recobrimento do volume é um excelente método para obter a similaridade da forma. Este programa tem uma taxa de recobrimento de 600-800 confôrmeros por segundo, este rápido desempenho do programa torna-o atrativo para a busca de novas substâncias bioativas para a enzima cruzaina. [160]

O programa ROCS [161] também foi submetido à análise da curva ROC, a mesma coleção de teste construída para a análise dos programas de docagem molecular foi empregada. As três estruturas de referência foram avaliadas, bem como as métricas de similaridade e a determinação dos valores limites de similaridade, FIGURA 3.13.

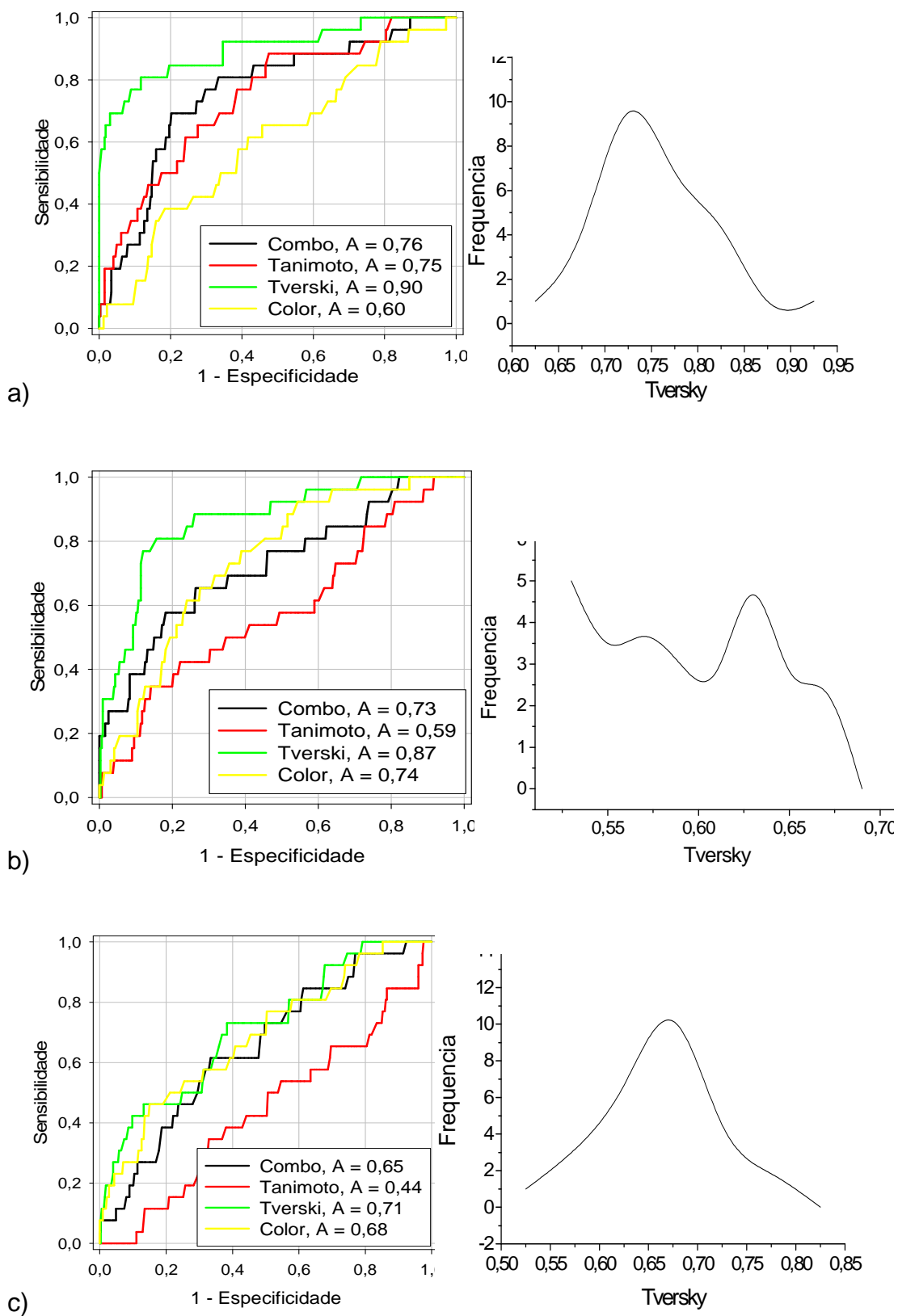


FIGURA 3.13: (A) Curva ROC das métricas de pontuação do programa ROCS e frequência de moléculas em função da métrica Tversky; (a) inibidor_1ME4, (b) 92260 e (c) K11777

A métrica que apresentou a melhor taxa de recuperação das moléculas ativas foi a de Tversky para as três estruturas de referências. Tradicionalmente a métrica utilizada pra medir a similaridade é a de Tanimoto, a principal diferença entre Tanimoto e Tversky é que o primeiro mede a diferença de similaridade total da molécula considerando a unidade, enquanto Tversky mede a representação da similaridade parcial da forma molecular, sendo mais eficiente na busca por subestrutura. [162]

Os limites impostos para as três estruturas de referência foram definidos como o valor da métrica Tversky em que apenas os inibidores da cruzaina apareceram nas primeiras posições da lista de saída do programa ROCS. Os valores foram: inibidor_1ME4 = 0,81, 92260 = 0,66 e K11777 = 0,68. Na sequência a sub-coleção proveniente do ZINC foi submetida à busca por similaridade, a frequência de moléculas em função da métrica de Tverski é mostrada na FIGURA 3.14.

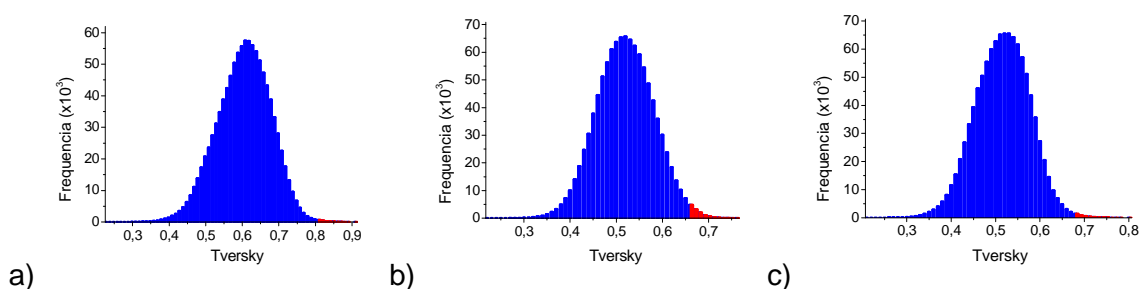


FIGURA 3.14: Frequência do número de estruturas em função da métrica Tversky para as estruturas dos inibidores de referência (a) inibidor_1ME4, (b) 92260 e (c) K11777

Pode ser observado que o número de compostos acima do limite estabelecido é extremamente pequeno quando comparado ao número total de estruturas na sub-coleção, áreas vermelha e azul das curvas, respectivamente. Os valores das áreas sob a curva ROC, similaridade de Tverski estabelecido como limite e número de moléculas acima do limite estabelecido na coleção do ZINC para as estrutura de referência são mostradas na TABELA 3.4

TABELA 3.4. Valores da área sob a curva ROC, limites de energia estabelecidos para as funções de pontuação e número de moléculas encontradas no ZINC acima dos limites

Estrutura de referência	Área sob a curva ROC	Limite estabelecido	Nº de moléculas do ZINC acima do limite
1ME4	0,90	0,81	1490
92260	0,87	0,66	12303
K11777	0,71	0,68	3159

Forma molecular é uma característica fundamental para o reconhecimento molecular entre ligantes e alvos biológicos. Freitas et al. utilizando o programa ROCS foi capaz de demonstrar seletividade entre inibidores de cruzaína e catepsina-L, este programa apresentou uma excelente performance na classificação correta de 37 ligantes de cruzaína nas primeiras posições de um banco contendo inibidores de catepsina-L e cruzaína misturados. A estrutura da cruzaína possui uma forma característica, com seus sub-sítios (S1-S4, capítulo 1) presentes em várias estruturas cristalográficas com diferentes ligantes, os quais poderiam induzir mudanças conformacionais na enzima. A conservação da estrutura do sítio é uma demonstração de que a forma molecular atua de maneira fundamental no reconhecimento de ligantes pela enzima cruzaína.

3.11 Escolha Consensual

Os programas de docagem molecular e similaridade química empregados de forma paralela para seleção de compostos identificaram moléculas com informações baseadas na estrutura dos ligantes conhecidos e da estrutura tridimensional da enzima cruzaína. A coleção utilizada para esta análise continham cerca de 3,5 milhões de estruturas provenientes da saída do programa FILTER, após esta etapa de análises os compostos acima dos valores limite estabelecidos para os programas FlexX, FRED e ROCS totalizaram cerca de 30 000 (trinta mil) estruturas. Esta foi uma das etapas cruciais na seleção das moléculas onde houve uma grande redução no número de estruturas, as moléculas remanescentes foram selecionadas por pelo menos uma função de pontuação ou métrica de similaridade acima dos limites estabelecidos. Uma observação foi haver moléculas selecionadas por um programa serem reprovadas por outro, a falta de concordância entre os programas foi utilizada como um critério de exclusão das estruturas. Somente aqueles compostos que foram recuperados por todos os programas e tiveram pelo menos um valor de energia de docagem ou similaridade acima dos limites estabelecidos foram selecionados para compor uma sub-coleção final com 8635 (oito mil e seiscentos) estruturas. A FIGURA 3.15 mostra a frequência de compostos remanescentes em função das funções de pontuação dos programas ou valores de similaridade.

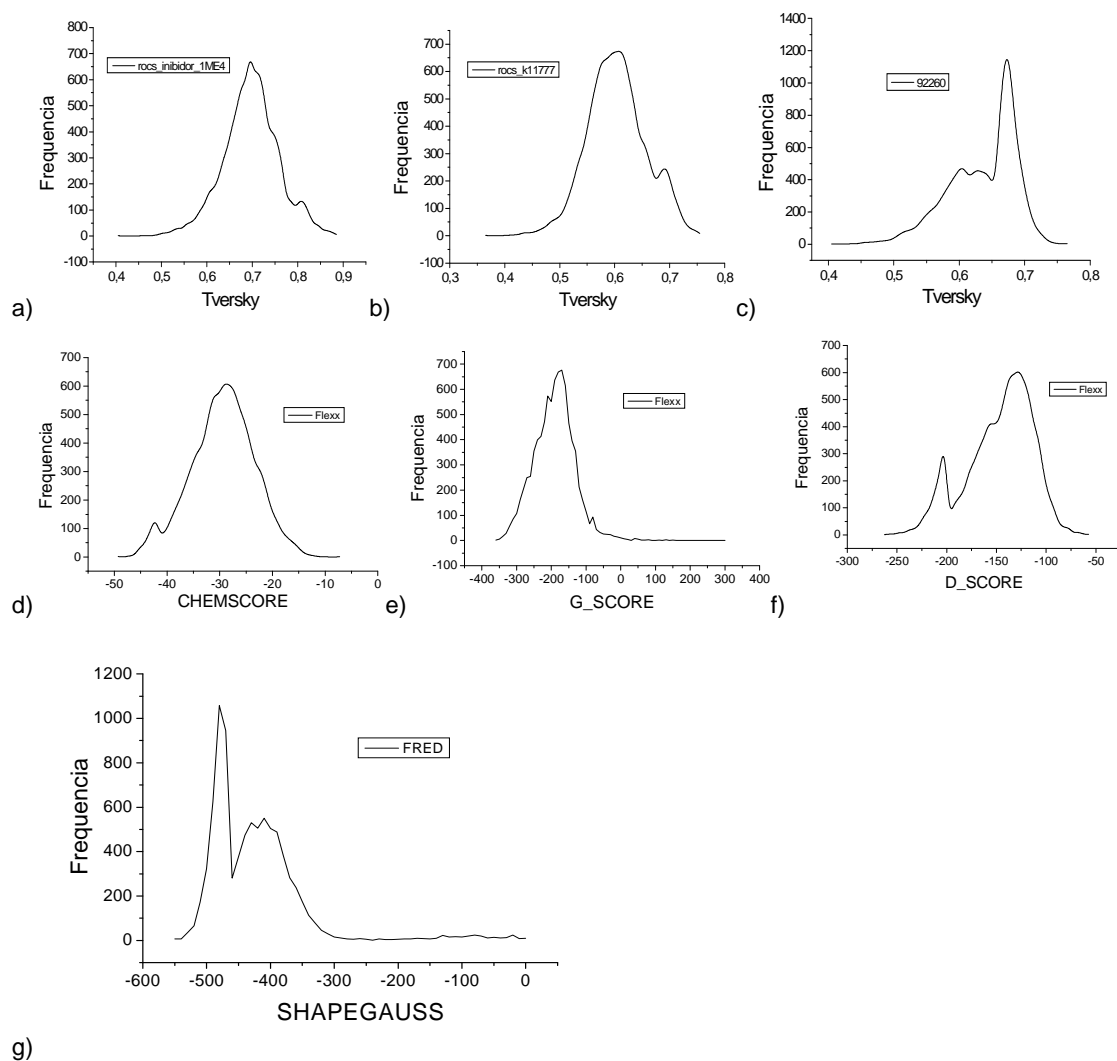


FIGURA 3.15: Frequência do número de estruturas em função da métrica Tversky para as estruturas dos inibidores de referência (a) inibidor_1ME4, (b) 92260 e (c) K11777; e das funções de pontuação dos programas FlexX CHEMScore (d), G_SCORE (e), D_SCORE (f) e FRED SHSPEGAUSS (g)

A análise mostrada na FIGURA 3.15 mostra que os critérios utilizados para a seleção dos compostos resultaram em um enriquecimento da coleção com estruturas que continham valores acima dos limites estabelecidos em todos os programas. A integração de métodos baseados na estrutura do ligante e do alvo resultaram na seleção de uma série de compostos com alto potencial de inibir a enzima cruzaina e com grande diversidade química, características fundamentais para a identificação de novas classes químicas para este alvo enzimático.

A sub-coleção de 8635 estruturas foi submetida à análise consensual dos valores das métricas, nesta etapa foram inclusos o cálculo da afinidade utilizando o modelo de HQSAR e também o programa de docagem GLIDE com precisão extra, o qual não foi utilizado anteriormente devido ao seu alto custo computacional.

3.12 Programa Glide

O programa Glide faz a busca do espaço conformacional do ligante mantendo a estrutura do receptor rígido. Glide usa uma série de filtros hierárquicos para procurar possíveis localizações do ligante na região do sítio ativo do receptor e produz um conjunto de conformações iniciais dos ligantes. Dadas estas conformações do ligante, as buscas iniciais são realizadas em todo o espaço disponível para o ligante para localizar poses promissoras. Esta busca prévia reduz drasticamente a fase de busca da região do espaço e cálculo de energia, a qual possui um alto custo computacional.

A partir das poses selecionadas pela triagem inicial, o ligante é minimizado no domínio do receptor usando uma função de padrão de energia mecânica molecular (neste caso, o campo de força OPLS-AA) em conjunto com um modelo dependente da distância dielétrica. Finalmente, de três a seis poses de menor energia são submetidos a um procedimento de Monte Carlo, que examina os mínimos nas proximidades conformacional. Este procedimento é necessário em alguns casos para orientar adequadamente os grupos periféricos e, ocasionalmente, altera ângulos de torção interna. O programa Glide utiliza uma função empírica de pontuação chamada GlideScore e possui um módulo de precisão extra. Este módulo possui penalidades para poses que violam princípios físico-químicos estabelecidos como grupos polares ou com carga ser expostos ao solvente. Esta versão com precisão extra minimiza a taxa de falsos positivos e é recomendada principalmente na fase de otimização de um ligante ou outros estudos onde uma coleção pequena de moléculas deve ser investigada com alta precisão, pois possui um alto custo computacional [163-164]. Por esta razão, utilizou-se o programa Glide com o modo de precisão extra somente na coleção final de compostos com 8 635 estruturas.

O resultado para a redocagem e a análise da curva ROC para verificar o desempenho do programa Glide em discriminar as moléculas ativas na coleção de teste é mostrada na FIGURA 3.16.

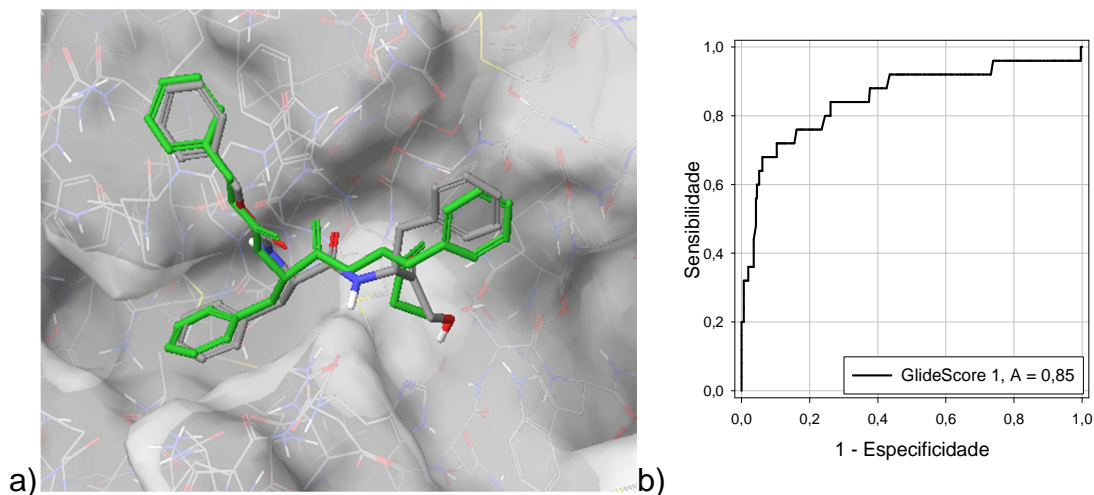


FIGURA 3.16: a) Estrutura do ligante proveniente da estruturas 1ME4 docado pelo programa Glide na enzima cruzaina (código PDB 1ME4), verde representa a pose do co-cristal e cinza a pose da docagem e b) curva ROC da função de pontuação GlideScore utilizada pelo programa Glide

O programa Glide foi capaz de reproduzir a estrutura cristalográfica satisfatoriamente com o de RMSD igual a 1,335 Å e apresentou um excelente desempenho, com uma área sob a curva ROC de 0,85 com as moléculas ativas recuperadas nas primeiras posições da lista de energia. A coleção final foi submetida à docagem com o programa Glide e nenhuma molécula foi descartada pelo programa. Isto significa uma concordância entre os programas de docagem quanto ao tamanho e forma molecular dos ligantes identificados, apesar de suas funções de pontuação atribuir diferentes valores de energia e poses para os complexos identificados. Por isso a necessidade de uma análise consensual entre os programas para a seleção dos compostos a serem ensaiados nos testes bioquímicos. A FIGURA 3.17 mostra a pose de uma série de ligantes bem pontuados pelo programa Glide.

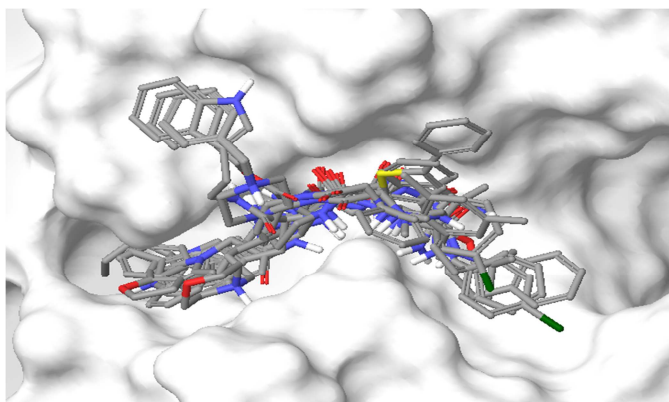


FIGURA 3.17: a) sobreposição de um ligante docado sobre a estrutura do inibidor da estrutura cristalográfica 1ME4 (em vermelho) e b) vários inibidores no sítio da enzima cruzaina

Pode ser observado na FIGURA 3.17 que alguns dos ligantes bem pontuados pelo programa Glide ocuparam totalmente o sítio da enzima, inclusive os sub-sítios S3 e S4, diferentemente dos outros programas utilizados previamente. A complementaridade dos ligantes pela enzima foi analisada através destas estruturas de docagem na análise visual e seleção dos compostos, após a análise de pontuação consensual.

3.13 Relações Quantitativas Estrutura Atividade Baseado em Hogramas Moleculares

A relação quantitativa estrutura atividade (QSAR) é um modelo matemático que relaciona a estrutura química com um processo bem definido, como a atividade biológica ou reatividade química. Na química medicinal ela é empregada em várias etapas da gênese, durante os processos de identificação e otimização de substâncias bioativas. [163]

O holograma QSAR (HQSAR – *hologram quantitative structure-activity relationships*) é um método de QSAR bidimensional baseado no conceito de impressão digital molecular. Os hologramas moleculares são originários da fragmentação molecular bidimensional, cada molécula do conjunto de dados é dividida em fragmentos estruturais que são identificados e colocados nos blocos de um vetor com comprimento fixo para formar o holograma molecular. Os blocos do holograma molecular são ocupados pelos descritores estruturais (variáveis independentes), com a variação na resposta biológica é obtida usando o método de regressão linear PLS. Ao contrário de outros métodos baseados em fragmentos, o método de HQSAR considera e codifica todos os possíveis fragmentos (linear, ramificado, sobrepostos, etc.) no holograma. Opcionalmente, informações acerca das propriedades tridimensionais, como a hibridização e a quiralidade também podem ser consideradas na construção do holograma, FIGURA 3.18. [164]

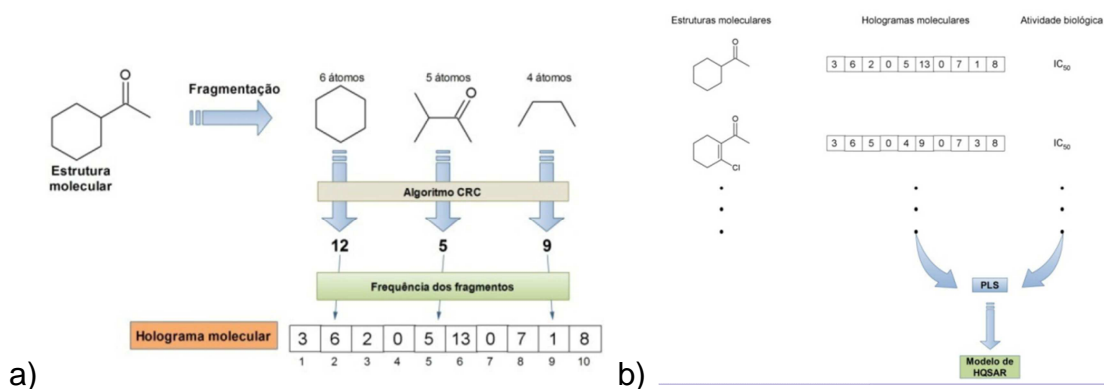


FIGURA 3.18: a) Construção do Holograma Molecular e b) Desenvolvimento do modelo de HQSAR

Os parâmetros importantes para o desenvolvimento do modelo são: (i) comprimento do holograma, (ii) tamanho dos fragmentos e (iii) distinção dos fragmentos. O comprimento do holograma define o número de posições no holograma disponível para que diferentes fragmentos sejam incorporados. O tamanho de fragmentos é definido pelo número máximo e mínimo de átomos que é imposto aos fragmentos a serem gerados, enquanto a distinção de fragmentos definem as características moleculares que distinguem um fragmento do outro. Os parâmetros utilizados para distinguir os fragmentos são variáveis ajustáveis e define o tipo de informação bidimensional que será incluída na criação dos descritores moleculares, estes são definidos por: (i) Átomos (A), Ligações (B), Conectividade (C), Hidrogênio (H), Quiralidade (Ch) e Doador e Aceitador (DA); A combinação destes parâmetros é utilizada para determinar a melhor correlação entre as estruturas e atividade.

A geração dos hologramas moleculares foi realizada utilizando-se várias combinações de distinção de fragmentos, comprimento dos blocos dos hologramas e tamanhos dos fragmentos moleculares [144]. Os parâmetros estatísticos dos melhores modelos são apresentados na TABELA 3.5.

TABELA 3.5. Resultados estatísticos para os modelos de HQSAR gerados para um conjunto de inibidores de cruzaina

Modelo	Distinção dos fragmentos	HL	N	Tamanho dos Fragmentos	r^2	q^2	SEE	SEP	RMSE
1	A/B	53	6	4-7	0,935	0,722	0,162	0,334	0,537
2	A/B/C	151	6	4-7	0,960	0,809	0,128	0,277	0,519
3	A/B/H	307	6	4-7	0,875	0,582	0,224	0,410	0,531
4	A/B/Ch	353	6	4-7	0,941	0,661	0,154	0,369	0,597
5	A/B/D	71	6	4-7	0,934	0,662	0,163	0,369	0,505
6	A/B/C/H	307	6	4-7	0,912	0,703	0,188	0,345	0,430
7	A/B/C/Ch	151	6	4-7	0,954	0,794	0,135	0,288	0,394
8	A/B/C/D	401	6	4-7	0,911	0,717	0,189	0,337	0,572
9	A/B/H/Ch	59	6	4-7	0,884	0,634	0,216	0,384	0,654
10	A/B/H/D	307	6	4-7	0,892	0,628	0,208	0,387	0,580
11	A/B/Ch/D	307	6	4-7	0,889	0,573	0,211	0,414	0,710
12	A/B/C/H/Ch	307	6	4-7	0,952	0,661	0,139	0,369	0,434
13	A/B/C/H/D	353	6	4-7	0,901	0,647	0,200	0,377	0,426
14	A/B/C/Ch/D	257	6	4-7	0,900	0,639	0,200	0,381	0,545
15	A/B/C/H/Ch/D	53	6	4-7	0,889	0,589	0,211	0,406	0,776

A – átomos, B – ligações, C – conectividade, H – hidrogênios, Ch – chiralidade, D – doador e aceitador de ligação de hidrogênio, HL – comprimentos dos blocos dos hologramas, N – número de componentes, r^2 ajuste linear do modelo, q^2 preditividade do modelo por validação cruzada, SEE – desvio do ajuste linear, SEP – desvio da preditividade do modelo e RMSE - raiz quadrada do desvio quadrático médio do erro da predição das séries de teste e validação externa

Os parâmetros estatísticos a serem observados são a preditividade do modelo indicada pelo valor de q^2 , juntamente com o parâmetro r^2 que indica a linearidade do modelo. Os valores de q^2 variaram entre 0,573 e 0,809 (TABELA 3.5) e os dois melhores modelos obtidos foram 2 e 7, a variação no tamanho dos fragmentos não melhorou os modelos. Embora um baixo valor de q^2 para o conjunto de treinamento seja um indicativo de baixa preditividade do modelo, alto valor de q^2 não significa necessariamente em alta preditividade. Isto quer dizer que, alto valor de q^2 é necessário, mas não suficiente para modelos com alta capacidade de predição. Portanto, a validação do modelo precisa ser feita através de séries de teste. Neste trabalho o conjunto de moléculas retirado da literatura continha 41 estruturas divididas em dois conjuntos, um para o treinamento do modelo (32) e outro para testar a preditividade (9 moléculas). Uma análise para avaliar a preditividade do modelo é o ajuste linear dos valores de pK_a calculados pelos modelos de HQSAR em

função dos valores experimentais (r^2_{pred}) A FIGURA 3.19 mostra esta análise para os dos melhores modelos de HQSAR obtidos.

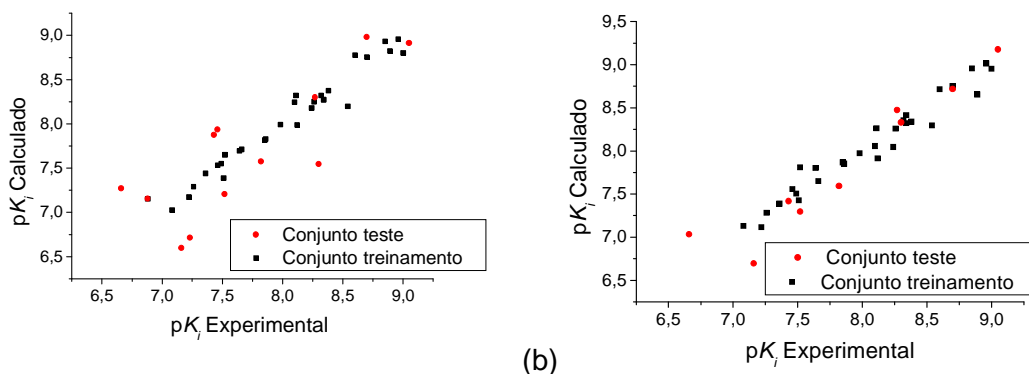


FIGURA 3.19: Análise de correlação dos valores de pK_i calculados pelos modelos de HQSAR em função dos valores experimentais; (a) modelo 2 ($r^2 = 0,958$) e (b) modelo 7 ($r^2 = 0,954$).

Os resultados mostram que apesar do modelo 7 apresentar menor valor de q^2 em relação ao modelo 2 é mais preditivo para a série de teste. Os valores de r^2_{pred} foram 0,970 para o modelo 7 e 0,942 para o modelo 2, além disso o valor de RMSE para o modelo 7 é menor, a FIGURA 3.19 mostra que em baixos valores de pK_i os valores calculados possuem um maior desvio em relação aos experimentais.

Adicionalmente, a capacidade de predição do modelo 7 foi avaliada frente a um conjunto [165] extra para o teste que não foi utilizado para a construção do modelo de HQSAR (validação externa), FIGURA 3.20.

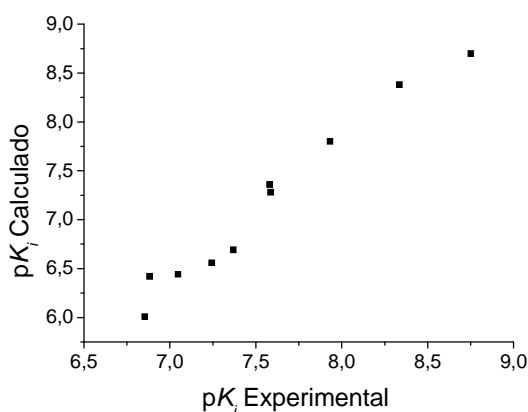


FIGURA 3.20: Validação externa do modelo 7 de HQSAR ($r^2_{pred} = 0,970$).

O modelo 7 apresentou também uma ótima preditividade frente a este conjunto de teste. A única diferença entre o modelo 2 e 7 é a utilização da quiralidade como distinção de fragmento no melhor modelo. Este é um indicativo da robustez e da capacidade de predição do modelo 7, uma vez que os inibidores

utilizados para a construção e validação do modelo são peptídios miméticos e possuem átomos de carbono quirais.

O método de HQSAR permite identificar os átomos que contribuem positivamente ou negativamente para a atividade biológica segundo o modelo construído. Um esquema de cores chamado mapa das contribuições indicam quais fragmentos moleculares são os principais responsáveis para o reconhecimento molecular. A análise das estruturas cristalográficas dos inibidores co-cristalizados com a enzima cruzaina permitiu identificar fragmentos moleculares frequentes nas moléculas, a análise do mapa das contribuições dos fragmentos permitiu avaliar a capacidade de identificação dos átomos importantes para o reconhecimento molecular. Os ligantes das estruturas 1ME4 e 1ME3 foram submetidos a esta análise e os mapas das contribuições dos fragmentos para a atividade biológica são mostrados na FIGURA 3.21.

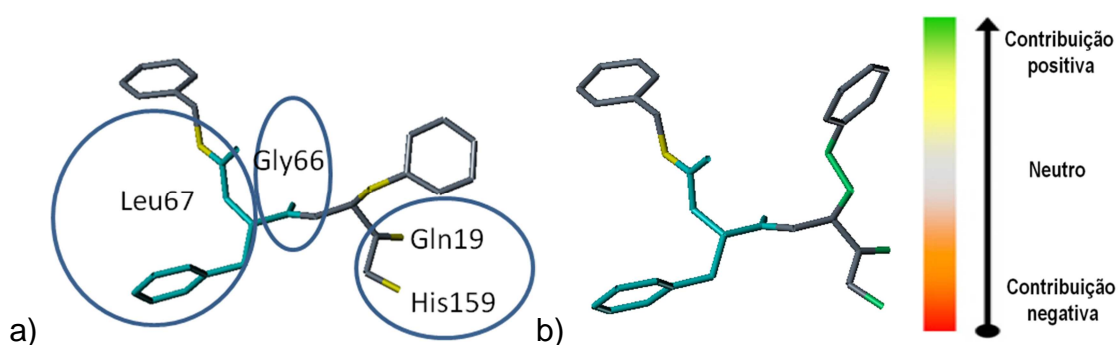


FIGURA 3.21: Mapa das contribuições dos fragmentos moleculares para a potência dos inibidores da enzima cruzaina: a) ligante da estrutura 1ME4 e b) 1ME3. O esquema de cores representa as contribuições dos fragmentos para a potência calculada pelo o modelo de HQSAR

Os ligantes das estruturas 1ME4 e 1ME3 ocupam a mesma posição no sitio da enzima e os mapas de contribuições para atividade biológica são semelhantes. Três regiões das estruturas moleculares foram indicadas como importantes para a atividade biológica pelo modelo de HQSAR e correspondem às interações com a estrutura da enzima cruzaina nos resíduos de aminoácidos Leu67, Gly66 e em dois aminoácidos da tríade catalítica His159 e Asp19. Desta forma, é possível atestar uma concordância entre o modelo de HQSAR e as restrições farmacofóricas utilizadas nos programas de docagem, apesar de serem modelos totalmente distintos onde HQSAR é baseado apenas na estrutura de ligantes conhecidos, ao passo que, a docagem molecular na estrutura tridimensional da enzima. A

integração dos métodos baseados na estrutura do ligante e do alvo é o ponto central para prover uma taxa de acerto significativa de inibidores da enzima no processo de seleção de moléculas a serem ensaiadas nos testes bioquímicos.

O modelo de HQSAR é consistente e de acordo com resultados publicados pelo grupo NEQUIMED-PN previamente [166] e a principal vantagem do método de HQSAR em relação a métodos de QSAR 3D, como campos moleculares de interação, é a necessidade de apenas estruturas 2D como entrada do programa, isso elimina a necessidade do alinhamento molecular e predição da conformação de interação. Como resultado o método de HQSAR é mais rápido que métodos de QSAR 3D e pode ser utilizado em grandes coleções virtuais no procedimento de ensaio virtual. Apesar do método ser 2D, opcionalmente, informações 3D como hibridação e quiralidade podem ser codificadas nos hologramas moleculares.

A sub coleção final com 8 635 estruturas moleculares foi submetida ao cálculo de pK_i utilizando o modelo de HQSAR construído e foi utilizado na análise consensual para a seleção dos compostos para os ensaios bioquímicos.

3.14 Análise consensual

A análise consensual foi proposta originalmente para os programas de docagem (*consensus docking*), a utilização de diferentes funções de pontuação poderiam balancear os erros de apenas uma função e aperfeiçoar a probabilidade de identificar ligantes conhecidos, reduzindo a taxa de falsos positivos [167-168]. O sucesso da pontuação consensual tem sido descrito em vários trabalhos, onde o seu uso é geralmente superior ao método da função de pontuação única. Neste trabalho fizemos uma extensão deste conceito incluindo na análise consensual não apenas dos métodos de docagem, mas também os valores de similaridade e afinidade calculados.

Os programas apresentam distintas maneiras de pontuação, por isso, efetuou-se um autoescalamento dos valores para que se tornassem comparáveis [169]. Nesse método os valores de pontuação são autoescalados para um número entre 0 e 1 usando a Equação 3.12:

$$X_{scaled} = (x - x_{min}) / (x_{max} - x_{min}) \quad (3.12)$$

Onde X_{scaled} é o valor escalonado, x é a pontuação atribuída por um programa, x_{min} o pior e x_{max} o melhor valor calculado para uma série de moléculas. Este procedimento foi aplicado porque cada programa pontua os compostos com diferentes parâmetros, por exemplo, energia de docagem ou pK_i calculado pelo HQSAR, que não podem ser comparados diretamente. Quando se escalonam os valores, independentemente do programa e parâmetro utilizado, ficam compreendidos em um intervalo que vai de 0 (para o composto pior pontuado) a 1 (para o composto melhor pontuado). Deste modo os valores podem ser somados e dar origem a uma pontuação e classificação consensual, denominada “listagem pelo número escalonado” (*scaled rank-by-number*). A FIGURA 3.22 ilustra este procedimento.

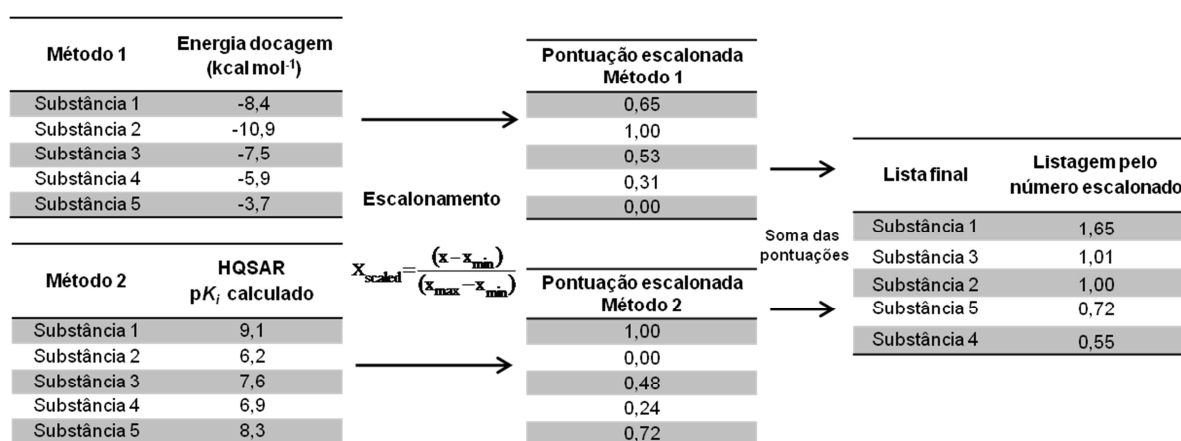


FIGURA 3.22: Processo de autoescalonamento das pontuações dos diferentes programas utilizados e análise consensual

Apesar de haver vários exemplos de enriquecimento de coleções por pontuação consensual, a utilização deste método para docagem gera opiniões divergentes. Um problema comum é a utilização de várias funções de pontuação em apenas um método de docagem, se a pose do complexo é predita erroneamente, as várias funções irão pontuar este complexo e inevitavelmente irão falhar. Além disso, ainda não existem métodos de docagem e pontuação que demonstrem um desempenho consistente e acurado em vários alvos distintos [170]. Por isso, neste trabalho utilizaram-se três métodos de docagem distintos para a docagem, e as funções de pontuação foram avaliadas previamente para a identificação de quais poderiam efetivamente selecionar moléculas com potencial inibitório da enzima cruzaina.

Alguns trabalhos na literatura apontam os fatores que podem contribuir para o desempenho superior de métodos de análises consensuais. Os fatores incluem (i)

melhora estatística devido a média da repetição de amostras serem mais próximas do valor real em relação às medidas individuais, (ii) moléculas ativas são acumuladas de forma mais efetiva que as inativas e, portanto, repetição de amostragem enriquece mais o conjunto de ativas em relação as inativas, (iii) a concordância na classificação de ativas é maior que inativas quando diferentes métodos são utilizados. [171-172]

A sub-coleção final contendo 8635 compostos foi submetida à pontuação de cinco diferentes programas: FlexX, FRED, GLIDE, ROCS e HQSAR. A combinação destas análises de forma consensual permitiram a classificação desta sub-coleção, as 400 (quatrocentas) moléculas melhores classificadas na análise consensual foram submetidas a uma inspeção visual e busca em bancos de dados como o SciFinder, Colaborative Drug Discovery, PubChem e ZINC, para assegurar o ineditismo das moléculas selecionadas como inibidores da cruzaina. Os compostos selecionados nunca foram testados contra as enzimas cruzaina e catepsina-L. Um dos 23 compostos selecionados foi predito ser ativo contra a catepsina-L no site do ZINC (SEA predictions), esta observação legitima a estratégia dos ensaios virtuais empregados para a seleção das moléculas para ensaios bioquímicos.

Os critérios utilizados para seleção dos compostos baseados na análise visual foram: a ocupação do sitio, principalmente nos sub-sítios S1, S2 e S3, e ligações de hidrogênio, com ênfase no aminoácido Gly66 inserido como restrição farmacofórica na docagem molecular, e diversidade química onde foram selecionados compostos similares aos inibidores conhecidos e também novas classes moleculares. A FIGURAS 3.23(a) mostra a estrutura cristalográfica do inibidor K11777, utilizado como referência no método de similaridade química, as FIGURAS 3.23(b-d) são as posições da docagem de inibidores selecionados para os testes *in vitro*.

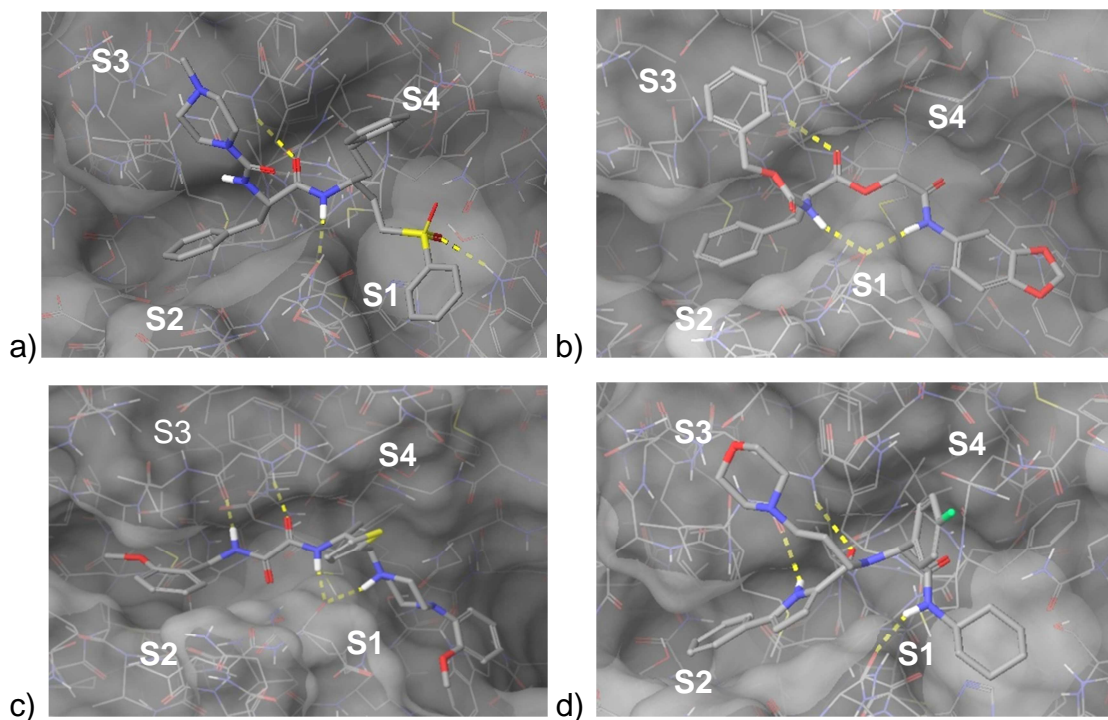


FIGURA 3.23: Estrutura do inibidor da cruzaina K11777 co-cristalizado com a enzima a), estruturas de complexos preditos por docagem molecular com modo de interação semelhante ao inibidor cristalizado identificados na análise visual b), c) e d)

Os sub-sítios S1, S2, S3 e S4 possuem características hidrofóbicas, a ocupação por substituintes com interações complementares nos compostos analisados foi considerada como aspecto positivo na inspeção visual. A Gly66 faz uma ligação de hidrogênio importante no reconhecimento molecular, e está na região central do sítio em uma porção hidrofílica do mesmo [173], substituintes capazes de receber ligação de hidrogênio nesta posição foram priorizados para os ensaios *in vitro*. Além disso, foram selecionadas substâncias pertencentes a diferentes classes químicas, algumas semelhantes aos inibidores conhecidos para avaliar a capacidade de identificação de inibidores pelos modelos computacionais desenvolvidos e outras pertencentes a classes inéditas.

Outro aspecto importante a ressaltar é que o modo de interação predito pelas poses da docagem dos compostos selecionados na inspeção visual foi semelhante às poses das estruturas cristalográficas de vários inibidores. Entretanto, os compostos das estruturas cristalográficas se ligam covalentemente à enzima ao passo que os compostos selecionados para os ensaios não possuem características de aceptores de adição eletrofílica, o que em teoria seria o perfil ideal de inibidores enzimáticos: reversíveis e competitivos. Uma representação das estruturas das substâncias selecionadas para os ensaios bioquímicos é mostrada na FIGURA 3.24.

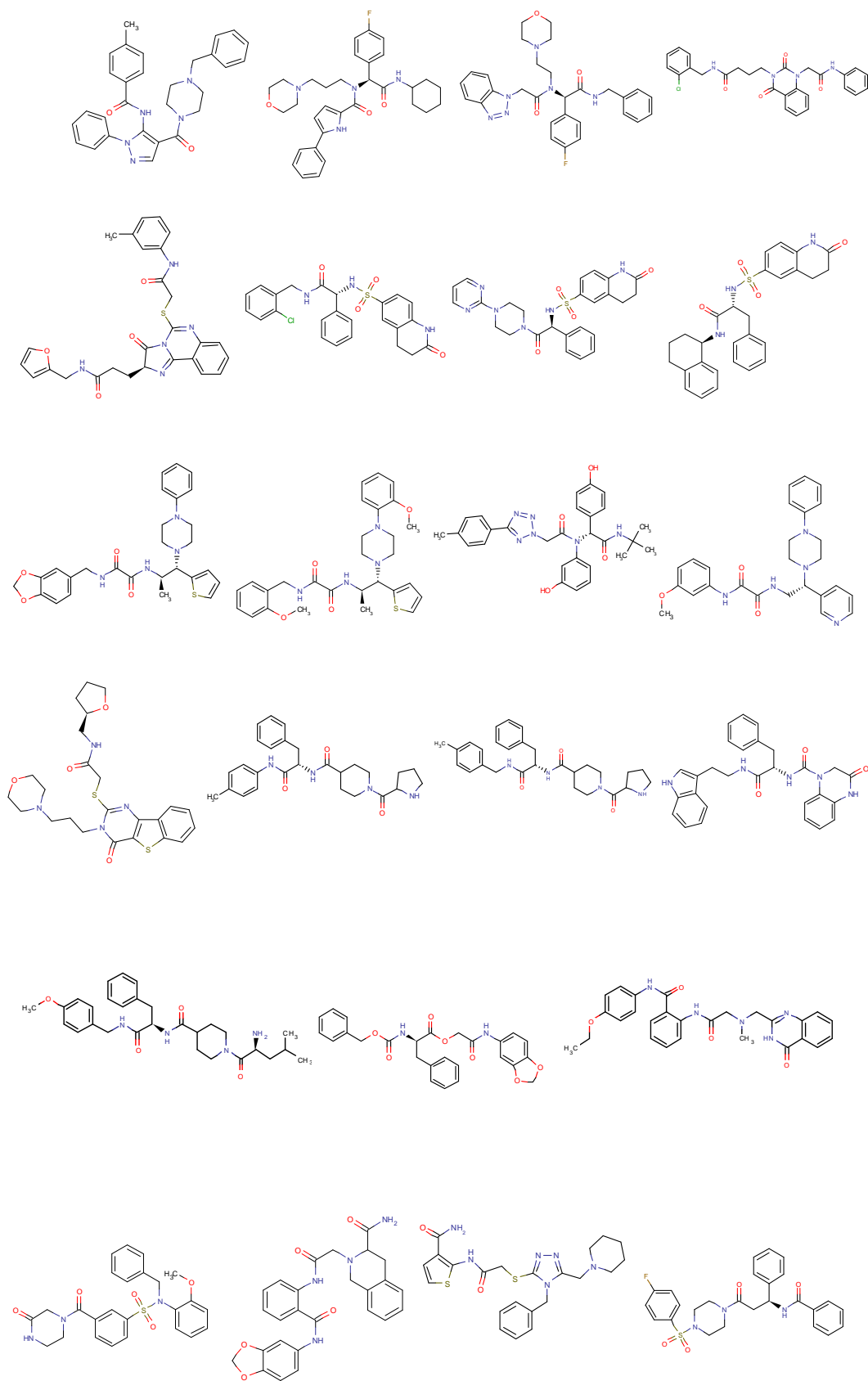
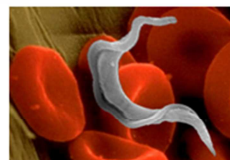
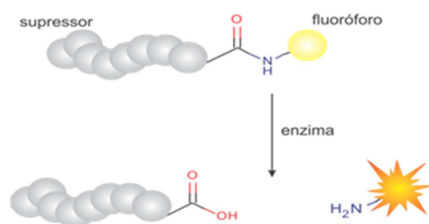
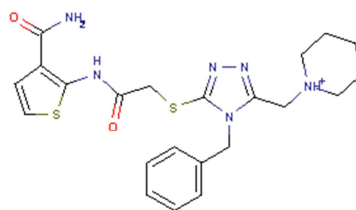
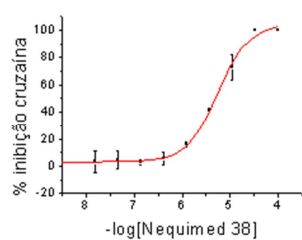


FIGURA 3.24. Representação das estruturas das substâncias seleccionadas para os ensaios bioquímicos

CAPÍTULO 4 – ENSAIOS *IN VITRO*



O capítulo 4 apresenta os métodos experimentais utilizados para obtenção da enzima cruzaina por expressão heteróloga e avaliação da atividade biológica das substâncias selecionadas pelos ensaios virtuais frente às enzimas cruzaina e catepsina-L e ao parasito *Trypanosoma cruzi* na forma tripomastigota

4.1 Receptores Enzimáticos

Os receptores podem ser proteínas de membrana, receptores hormonais, ácido desoxirribonucléico (DNA), enzimas, etc. As enzimas são um tipo especial de receptor; são proteínas especializadas em catalisar reações químicas. Elas estão associadas a biomoléculas, devido a sua extraordinária especificidade e poder catalítico.

Muitas enzimas são compostas de uma cadeia polipeptídica simples. Este é o caso de enzimas como a ribonuclease, lisozima, tripsina, pepsina e algumas alfa-amilase. Por outro lado, existe um grande número de enzimas compostas por mais de uma cadeia peptídica. A enzima cruzaina de *Trypanosoma cruzi* é composta por uma cadeia polipeptídica. A repetição das cadeias polipeptídicas na construção de uma macromolécula de proteína caracteriza a estruturação que esta pode assumir. [174-175]

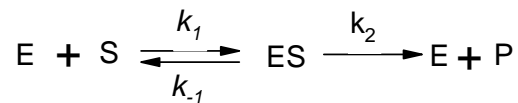
Para a realização de uma análise quantitativa com o sistema enzimático é importante conhecer: a estequiometria da reação catalítica; se a enzima requer adição de cofatores; a dependência da enzima sob a concentração do cofator ou substrato; pH ótimo de catálise e a temperatura pela qual a enzima é estável e apresenta uma elevada atividade. [176]

4.2 Mecanismos Cinéticos de Catálise Enzimática

Em 1894 Emil Fischer propôs o modelo chave – fechadura para explicar a ação enzimática, um modelo em que a estrutura da enzima e do substrato são rígidas [177]. Já no modelo proposto por Koshland em 1958, tanto a enzima quanto o substrato passam por mudanças de conformação para propiciar o encaixe [178]. A enzima não aceita simplesmente o substrato, o substrato é distorcido para conformação exata do estado de transição, denominado encaixe flexível.

Ao completar a reação catalítica a enzima libera o produto e retorna a forma original. O processo (em sua forma simplificada) ocorre em duas etapas, a primeira reversível e a outra irreversível, a formação do produto (P) a partir de um único

substrato (S) na presença da enzima (E) forma o complexo de enzima-substrato (ES):



Os termos k_1 , k_{-1} e k_2 são as constantes, para a associação da enzima com o substrato, sua dissociação e a formação do produto.

Este estudo é muito útil para a compreensão do mecanismo da atividade catalítica que cada enzima segue no processo metabólico, de como sua atividade é controlada na célula e a atuação dos medicamentos moduladores da atividade enzimática.

A velocidade da reação (v) é limitada pela etapa de conversão do complexo $[ES]$ em P , e esta dependerá da concentração da enzima que interage com o substrato. Assim:

$$v = k_2[ES] \quad (4.1)$$

onde, k_2 é denominado k_{cat} quando a enzima encontra-se saturada, sob condições de estado estacionário e a reação atinge a velocidade máxima (v_{max}).

Com excesso de substrato, tem-se $[S] \gg [E]$, estando o sistema no estado estacionário. O complexo $[ES]$ está em equilíbrio com $[E]$ e $[S]$, assim a $[ES]$ é constante. A formação do complexo $[ES]$ dependerá da constante k_1 e da disponibilidade da enzima e do substrato.

A dissociação do complexo $[ES]$ pode ocorrer de duas maneiras, seja pela formação do produto ou pela não reatividade do substrato pela enzima. Assim, o estado estacionário pode ser escrito:

$$k_1[E][S] = (k_{-1} + k_2)[ES] \quad (4.2)$$

Dividindo ambos os lados por $(k_{-1} + k_2)$:

$$\frac{k_1[E][S]}{(k_{-1} + k_2)} = [ES] \quad (4.3)$$

A partir da Equação 4.4 obtém-se a constante de Michaelis Mentem (K_M), que pode ser usada para descrever a afinidade da enzima pelo substrato. Quanto menor o valor de K_M , maior a afinidade da enzima pelo substrato.

$$\frac{(k_{-1} + k_2)}{k_1} = K_M \quad (4.4)$$

Sendo: $v_{\max} = k_2[E_0]$

$$[E_0] = [E] + [ES]$$

onde $[E_0]$ é a concentração total de enzima, $[E]$ a concentração de enzima livre, $[ES]$ a enzima na forma complexada e a velocidade máxima (v_{\max}).

Rearranjando as equações 1, 2, 3 e 4 chega-se na Equação de Michaelis-Menten:

$$v = \frac{v_{\max}[S]}{K_M + [S]} \quad (4.5)$$

Na determinação da velocidade máxima da catálise enzimática, a concentração do substrato tende a aumentar segundo uma cinética de primeira ordem, até se atingir uma taxa constante da velocidade (ordem zero). [179]

A velocidade máxima de atividade da enzima é obtida quando seu sitio ativo está saturado pelo substrato. Há outros fatores que influenciam na velocidade da reação, como por exemplo, o pH e a temperatura. [180]

4.3 Inibições Enzimáticas

Quando uma enzima é inibida, o substrato é acumulado, e os produtos metabólicos não são gerados (a menos que haja outra enzima capaz de catalisar a mesma reação, ou que haja outras rotas metabólicas para os mesmos produtos). Qualquer composto que diminua ou bloqueie a catálise de uma enzima é chamado de inibidor enzimático. Se uma célula tem deficiência de um substrato para a enzima alvo, e essa deficiência resulta em um estado doentio para a célula, então, a inibição dessa enzima pode impedir a degradação desse substrato aumentando sua concentração.

Um exemplo é a inibição da enzima xantina oxidase, que catalisa a conversão de xantina em ácido úrico. A inibição dessa enzima diminui a concentração de ácido úrico no organismo resultando em um efeito anti-hiperuricêmico. [181-182]

Além disso, um inibidor enzimático pode possuir uma estrutura química similar a estrutura do substrato ou do produto da enzima alvo; desta forma, a escolha da enzima alvo prontamente disponibiliza um direcionamento na busca dos inibidores. [183]

4.3.1 Tipos de Inibidores

Muitos tipos de moléculas inibem as enzimas e interagem de forma diferente com o alvo biomacromolecular. A principal distinção entre os inibidores é o mecanismo de inibição, reversível e irreversível.

A forma de interação entre alvo e ligante que envolve interações não-covalentes é dita reversível e pode ser classificada como inibição reversível competitiva, não competitiva e incompetitiva. Na inibição irreversível há a formação de ligações químicas entre o alvo e o ligante (ligações covalentes) e inativação do alvo biológico. [184]

4.3.2 Inibição reversível competitiva

Na inibição reversível competitiva a molécula de inibidor pode apresentar estrutura semelhante ao substrato da enzima. Neste caso, o inibidor compete com o substrato pelo mesmo sitio ativo. Se o sitio ativo estiver ocupado pelo inibidor, o substrato deixa de interagir nesta região e então o processo catalítico de conversão em produto não se realiza. Portanto, na inibição reversível competitiva, o inibidor compete pelo mesmo sitio catalítico do substrato, como no esquema abaixo:

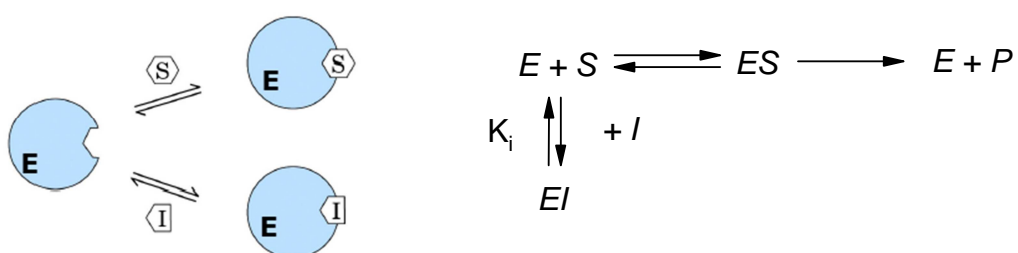


FIGURA 4.1. Esquema ilustrativo do mecanismo de inibição reversível competitiva

Sendo E (enzima), I (inibidor), (S) substrato, (EI) complexo enzima/inibidor, (ES) enzima/substrato e (P) produto. A equação de Michaelis-Menten, na presença de inibidor competitivo torna-se:

$$v = \frac{V_{\max} [S]}{K_M \left(1 + \frac{[I]}{K_i} \right) + [S]} \quad (4.6)$$

O efeito do inibidor competitivo modifica o valor da constante de Michaelis-Menten, mas não altera a velocidade máxima da reação. [185]

4.3.3 Inibição reversível não competitiva

A inibição reversível não competitiva ocorre quando uma molécula pode ligar-se com o complexo enzima/substrato formando o complexo enzima/substrato/inibidor [ESI] ou interagir com a enzima livre formando o complexo enzima/inibidor [EI]. Neste processo há modificação da conformação da enzima e tornando o processo catalítico ineficiente: [186]

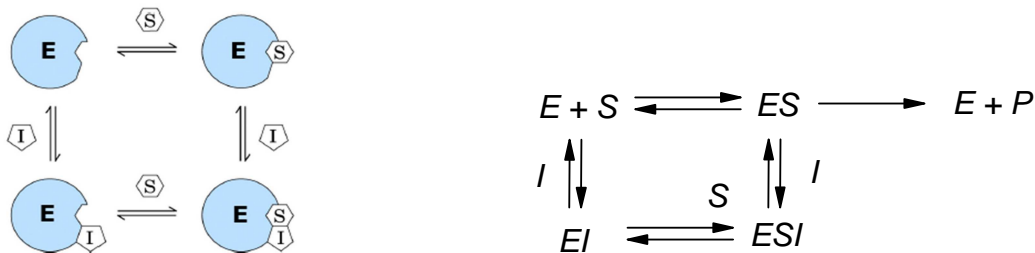


FIGURA 4.2: Esquema ilustrativo do mecanismo de inibição reversível não competitiva

O inibidor não competitivo apresenta uma grande afinidade pela enzima, seja similar ou não com a estrutura do substrato. A equação de Michaelis-Menten fica:

$$v = \frac{v_{\max} [S]}{K_M \left(1 + \frac{[I]}{K_i} \right) + (S) \left(1 + \frac{[I]}{K_i} \right)} \quad (4.7)$$

O efeito desse inibidor é a diminuição da velocidade máxima da reação, enquanto K_M permanece constante. [187]

4.3.4 Inibição reversível incompetitiva

A inibição reversível incompetitiva se diferencia da inibição reversível não competitiva pelo fato do inibidor ter afinidade apenas pelo complexo enzima/substrato $[ES]$, formando o complexo inativo enzima/substrato/inibidor $[ESI]$ e não interage com a enzima livre. É raro este comportamento de inibição. O esquema abaixo demonstra o processo de inibição:

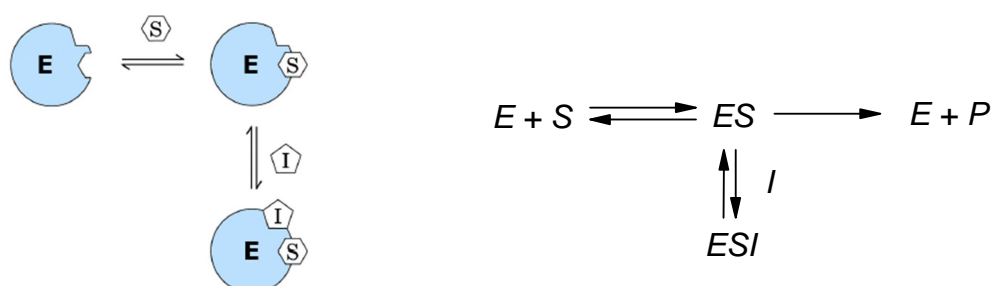


FIGURA 4.3: Esquema ilustrativo do mecanismo de inibição reversível incompetitiva

A redução efetiva da concentração do complexo $[ES]$ aumenta a afinidade aparente da enzima pelo substrato, devido ao princípio de Le Chatelier, há o deslocamento do equilíbrio para a formação do complexo $[ES]$. A equação de Michaelis-Menten na presença de inibidor incompetitivo fica:

$$v = \frac{v_{\max} [S]}{\left(1 + \frac{[I]}{K_i}\right) \left(\frac{K_M}{\left(1 + \frac{[I]}{K_i}\right)} + [S]\right)} \quad (4.8)$$

O efeito desse inibidor modifica a velocidade da reação, diminuindo o valor de v_{\max} e o valor de K_M . [188]

A eficácia de um inibidor enzimático como agente terapêutico dependerá da (i) potência do inibidor, (ii) sua seletividade ao alvo enzimático, (iii) a escolha da rota biossintética na qual a enzima alvo está inserida e (iv) as características farmacocinéticas apropriadas do inibidor ou seus derivados. Maior potência implica

no uso de menor quantidade do fármaco e, portanto, redução de efeitos colaterais, para obter-se a resposta farmacológica desejada. [189-191]

4.2 Obtenção de enzima por expressão heteróloga

A obtenção das enzimas de interesse pode ser realizada através da técnica de DNA recombinante. Para isto, um fragmento de DNA de qualquer origem é incorporado em um plasmídeo pequeno, com o qual o hospedeiro (geralmente a bactéria *Escherichia coli*) é transformado. Este plasmídeo, incluindo o fragmento de DNA nele incorporado, se reproduz no hospedeiro, originando cópias do DNA “estranho”. [192]

A fragmentação do DNA é feita com o uso de enzimas de restrição. Com elas o preparado de DNA em questão é cortado em fragmentos precisamente definidos. Quando junta-se um pedaço qualquer de DNA a um plasmídeo, mantendo a possibilidade de se propagar indefinidamente este DNA, dá-se o nome de “clone” (cDNA). Assim, após a etapa de purificação da enzima, é possível obter esta em grande quantidade, para que se possam realizar os estudos *in vitro*. [193]

A produção de uma proteína *in vivo* é um processo extremamente complexo, o qual envolve modificações pós-transducionais, necessárias para estabilidade e atividade biológica, glicosilação, fosforilação e correto enovelamento, são exemplos. A produção de uma proteína fora de seu hospedeiro natural é chamada de produção de proteínas heterólogas, e a proteína expressa é chamada de proteína heteróloga [194]. O desenvolvimento de ferramentas em biologia molecular permitiu a expressão de várias proteínas heterólogas de interesse terapêutico em espécies simples como bactérias. As ferramentas essenciais para a produção de uma determinada proteína são o cDNA que codifica a proteína desejada, um vetor e um sistema biológico capaz de transcrever e traduzir a proteína codificada pelo cDNA (sistema de expressão). O sistema de expressão deve ser capaz de produzir a proteína na conformação correta, ter boa reprodutibilidade, ser de fácil manipulação e manutenção, ser seguro, econômico e possuir um sistema simples de produção. Entretanto, não há um único sistema biológico que atenda este amplo número de critérios.

Ao longo do tempo vários sistemas de expressão foram desenvolvidos, e são ilustrados na TABELA 4.1.

TABELA 4.1: Comparação dos sistemas disponíveis para a produção de proteínas heterólogas

Sistema	Custo de produção	Tempo de produção	Produção em larga escala	Nível de expressão	Risco de contaminação
Bactéria	Baixo	Curto	Alto	Alto	Endotoxinas
Levedura	Médio	Médio	Alto	Baixo-alto	Baixo risco
Célula de inseto	Alto	Médio	Médio	Baixo-alto	Alto
Célula de mamíferos	Alto	Longo	Muito baixo	Baixo-moderado	Vírus e príons
Animal transgênico	Alto	Muito Longo	Baixo	Moderado-alto	Vírus e príons
Célula de plantas	Baixo	Curto	Alto	Moderado-alto	Baixo risco
Plantas transgênicas	Muito baixo	Longo	Muito alto	Moderado-alto	Baixo risco

A bactéria *Escherichia coli* (*E. coli*) foi o sistema de expressão pioneiro o qual produziu a primeira proteína, a somatostatina. A insulina foi a primeira proteína aprovada para a farmacoterapia pela agência reguladora FDA, produzida em um sistema de *E. coli* [195]. Gradualmente, várias modificações foram introduzidas nas bactérias de maneira que pudessem produzir proteínas cada vez mais complexas [196]. Da mesma forma, outros sistemas de expressões foram desenvolvidos para produção heteróloga de proteínas, como células de mamíferos [197], leveduras [198], insetos [199], plantas transgênicas [200] e animais transgênicos. [201]

Apesar de haverem vários sistemas de expressão disponíveis nos dias de hoje, o principal ainda é sistema bacteriológico utilizando a *E. coli*. Comparativamente aos outros sistemas, listados na TABELA 4.1, o sistema de expressão em bactéria possui baixo custo e tempo de produção, possui altos níveis de expressão e possibilidade de produção em larga escala.

4.3 Ensaio Cinético por Fluorimetria

O ensaio cinético padrão utilizado para a determinação da atividade das enzimas cruzaína e catepsina-L é fluorescência [202]. O fenômeno de fluorescência é definido pela capacidade de uma molécula emitir radiação eletromagnética na região do ultravioleta próximo ao visível, quando excitado por uma fonte de luz. Os estados eletrônicos da maior parte das moléculas orgânicas podem agrupar-se em dois tipos principais: singleto e tripleto. No estado singleto, todos os elétrons da

molécula têm os spins emparelhados, enquanto, no estado tripleto, há um conjunto de spins eletrônicos não emparelhados. [203]

O mecanismo fundamental de fluorescência é a emissão de fótons quando a molécula é excitada no estado singleto, onde o spin do elétron no orbital excitado mantém sua orientação original. O processo de emissão e excitação no fenômeno de fluorescência é muito rápido com tempo de vida na ordem de nano segundo. Desta forma, é possível observar facilmente fluorescência à temperatura ambiente e diretamente em solução, o que torna o procedimento experimental fluorimétrico bastante simples.

A FIGURA 4.4 representa o um diagrama de Jablonski [204] e mostra as transições eletrônicas permitidas para uma molécula.

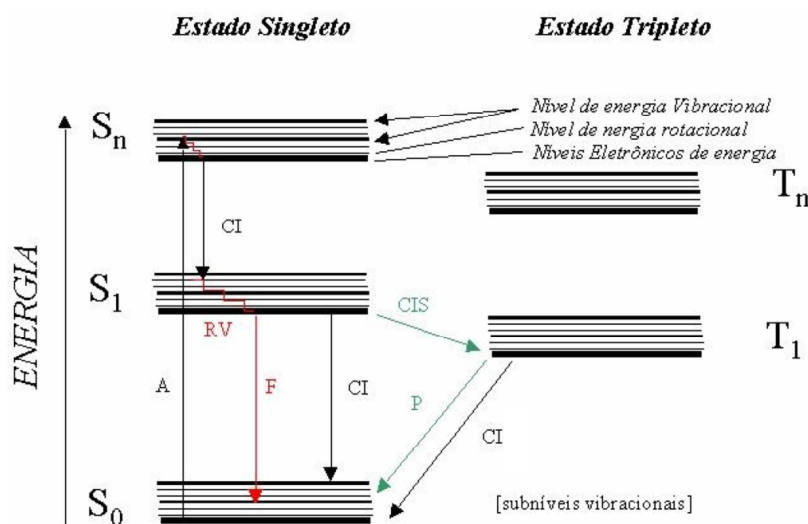


FIGURA 4.4: Diagrama de Jablonski modificado. (A) absorção de um fóton, (S_0) estado fundamental, (S_n) estado excitado singleto, (S_1) primeiro estado excitado singleto, (R_V) relaxamento vibracional, (CI) cruzamento interno, (CIS) cruzamento intersistemas, (T_n) estado excitado tripleto [204]

Nesse esquema o eixo das ordenadas representa a energia, ao passo que, o eixo das abscissas os níveis energéticos ocupados. O estado S_0 representa o estado fundamental, S_1 e S_n estados excitados. A partir de S_1 , ela pode seguir dois caminhos para retornar ao estado fundamental. Primeiro, se a diferença de energia entre S_1 e S_0 for pequena e existir possibilidade de sobreposição de níveis vibracionais, a molécula pode ser levada ao mais baixo nível vibracional de S_1 por relaxamento vibracional sem emissão de radiação eletromagnética, ou seja, ocorre um cruzamento interno de níveis de energia. Se, no entanto, a diferença energética entre S_1 e S_0 for relativamente grande, a desativação para o estado fundamental dá-

se com emissão de radiação na forma de fluorescência. Pode ocorrer ainda um cruzamento intersistemas e os elétrons ocuparem um estado tripleto possibilitando o fenômeno de fosforescência. [204]

O método de fluorescência tem sido amplamente utilizado para a determinação da cinética das enzimas do tipo cisteíno-proteases. Zimmerman et al. sintetizaram a primeira classe de substratos fluorogênicos para as proteases tripsina, elastase e quimotripsina [205]. Esses substratos eram amidas de acil aminoácidos ou peptídeos com o substituinte 7-amino-4-metilcumarina (MCA). A FIGURA 4.5 ilustra a estratégia utilizada para a determinação da formação de produto catalisada pelas enzimas.

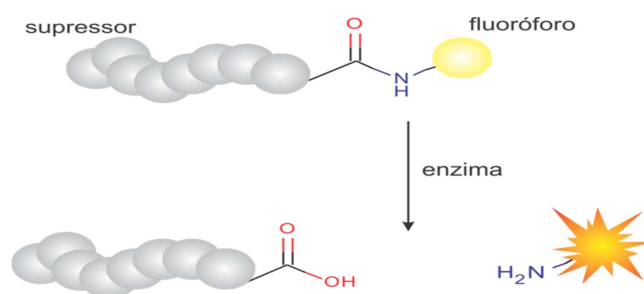


FIGURA 4.5: Estratégia utilizada para gerar o sinal de fluorescência em reações catalisadas por cisteíno proteases

Aos grupos terminais carbóxi e amino de um dipeptídeo são ligados um grupo fluorescente e supressor, respectivamente. Após a quebra da ligação peptídica pela enzima, o grupo fluorescente é liberado em solução permitindo a detecção da formação do produto e determinação da cinética de reação catalisada pela enzima.

Diversos substratos fluorogênicos foram utilizados como substratos para a enzima cruzaina [206-207]. O substrato estabelecido como padrão para as análises cinéticas de novos inibidores da enzima foi o N-benzóxicarbonil-L-fenilalanil-L-arginina-7-amido-metil-cumarina (Z-Phe-Arg-MCA ou ZFRMCA), FIGURA 4.6.

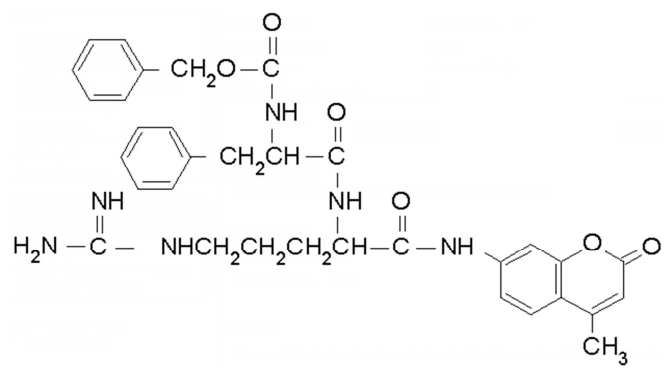


FIGURA 4.6: Estrutura do ZFRMCA (substrato sintético para a cruzaina)

Os valores das constantes cinéticas da enzima cruzaina reportadas na literatura para este substrato são: K_M entre 0,96 e 1,4 μM , k_{cat} entre 3,9 e 11,6 s^{-1} e k_{cat}/K_M entre 4677 e 9667 $\text{mM}^{-1} \text{s}^{-1}$. [57, 206-207] Os valores das constantes cinéticas para a enzima catepsina-L são: K_M entre 1,2 e 2,4 μM , k_{cat} entre 10 e 25,8 s^{-1} e k_{cat}/K_M entre 4166 e 11700 $\text{mM}^{-1} \text{s}^{-1}$. [208-210]

4.1 Materiais

Os reagentes fosfatos de sódio monobásico e dibásico, ácido etilenodiamino tetra-acético (EDTA), dimetilsulfóxido (DMSO), Triton-x100, acetato de sódio, glicerol, cloreto de sódio, azul de bromofenol, isopropanol, ZFRMCA, Isopropil β -D-1-tiogalactopiranosideo (IPTG), enzima catepsina-L de fígado humano, ampicilina, Tris-HCl, uréia, Lisozima, Dnase e fluoreto de fenilmetilsulfóxido (PMSF) foram adquiridos da empresa Sigma-Aldrich.

Os reagentes *comassie brilliant blue*, acrilamida bis-acrilamida, temed e persulfato de amônia foram adquiridos da empresa Bio-Rad.

O reagente Ditiotretol (DDT) foi adquirido da empresa USB Corporation com grau de pureza “ultra-puro”. O kit de extração de DNA plasmidial, QIAprep Spin Miniprep Kit foi utilizado da empresa Qiagen. O meio de cultura Luria Broth (LB) foi adquirido da empresa Fermentas.

As substâncias adquiridas para os ensaios de inibição foram adquiridas das empresas Enamine, Chemdiv e Interbioscreen.

Todos os reagentes têm certificação do fabricante de pureza > 99 %. Nenhuma purificação/avaliação da integridade adicional foi realizada.

A expressão da enzima foi realizada no Laboratório de Química Medicinal de Produtos Naturais (NEQUIMED-PN) do Instituto de Química de São Carlos (IQSC), Universidade de São Paulo (USP), equipado com Shaker da Marconi, Centrífuga Eppendorf e Desmembrador Fischer. A purificação foi feita no Laboratório de Bioquímica e Biologia Molecular (BMB) do IQSC utilizando o sistema FPLC Akta da GE e colunas da Bio-Rad.

Os experimentos fluorimétricos foram conduzidos no Laboratório de Farmacologia/Bioquímica no departamento de ciências fisiológicas da Universidade Federal de São Carlos com um fluorímetro da empresa Molecular-Devices (SpectraMax Gemini XPS). [211]

4.2 Métodos

4.2.1 Extração de DNA plasmidial

A obtenção do DNA plasmidial contendo o gene da cruzaina foi realizada através do método de lise alcalina em pequena escala, centrifugaram-se as células durante 5 minutos em 10 mL de cultura a 13,2 rpm. Em seguida o precipitado foi solubilizado e lisado com solução tampão alcalina, precipitando as demais proteínas presente no meio celular, restando apenas o DNA plasmidial em solução aquosa. Purificou-se e armazenou-se a -20 °C. Para esta etapa utilizou-se o Kit Miniprep promega (Wizard Plus SV Miniprep DNA purification System).

4.2.2 Transformação da *Escherichia coli* para expressão da proteína

O plasmídeo utilizado para expressão da enzima foi o pUC18 com o gene da cruzaina [57]. Foram adicionados 20 ng do plasmídeo às bactérias competentes *E. coli* BL21-DE3 e *E. coli* BL21-pLysS e mantido em gelo por 5 minutos seguido de choque térmico à 42 °C por 120 segundos e incubadas em gelo por mais 5 minutos. As bactérias foram fixadas em placas contendo o meio de cultura LB Agar, esterilizadas previamente com o antibiótico ampicilina 100 mg L⁻¹ e mantidas em estufa à 37 °C por 16 horas.

4.2.3 Pré-Inóculo

Uma colônia da bactéria transformada em células de expressão *E. Coli* BL21 DE3 e pLysS foi selecionada e adicionada em 5 mL do meio de cultura LB em solução aquosa, previamente esterilizado com ampicilina 100 mg L⁻¹. Deixou-se 16 h a 37°C, 250 rpm.

4.2.4 Inóculo e Indução

Preparou-se 500 mL do meio de cultura LB com ampicilina 100 mg L⁻¹. Adicionaram-se os 5 mL do pré-inóculo nesta solução e levou-se ao misturador (*shaker*). Deixou-se por aproximadamente 2:30 h à 37 °C, 250 rpm e mediu-se a densidade óptica da solução a 600 nm, quando atingiu-se um valor entre 0,5 – 0,6 adicionou-se o agente indutor IPTG 1,0 mM. Deixou-se 16 h a 37 °C, 250 rpm. Retirou-se uma amostra do induzido de cada linhagem de bactéria BL21 DE3 e pLysS para verificar qual possui melhor nível de expressão.

4.2.5 Lise das Células

Colocou-se as células nos tubos adequados pra centrifugação a 5.000 rpm, 4 °C, por 10 minutos . O sobrenadante foi descartado e o precipitado foi estocado à -20 °C (etapa de estocagem podendo ficar por até um mês se não for prosseguir no mesmo dia).

A lise das bactérias foi realizada descongelando-se o precipitado (sem aquecimento) em seguida suspendeu-o em 20 mL do tampão Tris (50 mM Tris- HCl, 1,0 mM EDTA, 50 mM NaCl pH 8) . Manteve-se sempre resfriado em banho de gelo moído. Adicionou-se ao tampão Tris:

- 10 µL de PMSF (0,5 M em isopropanol – solução estoque) para dar uma concentração final de 100 µM da solução;
- Lisozima 0,2 mg mL⁻¹ (Concentração final).
- Triton-X100, a uma concentração de 2,0 % v/v.
- 5 µL de Dnase (10 unidades µL⁻¹).

Deixou-se incubando em banho de gelo moído, mexendo de vez em quando, durante 1 hora. Após este período sonicou-se a suspensão através de 10 ciclos de 30 segundos com intervalos de 60 s.

A suspensão foi centrifugada e o sobrenadante foi descartado, a proteína é expressa em corpos de inclusão e fica no precipitado. O precipitado foi lavado com o tampão Tris contendo TritonX-100 (50 mM Tris.HCl, 1,0 mM EDTA, 50 mM NaCl pH 8 e Triton-X100 2,0 % v/v) através de três ciclos de ressuspensão e centrifugação.

4.2.6 Dissolução dos corpos de inclusão e renovelamento da enzima

Uma solução de tampão Tris com 8,0 M de uréia foi preparada (50 mM Tris-HCl, 1,0 mM EDTA, 50 mM NaCl pH 10,7 e 8,0 M de uréia). Suspendeu-se o precipitado contendo a enzima com 40 mL desta solução e deixou agitando-a com um agitador magnético por 4 horas, a solução foi mantida resfriada em banho de gelo moído.

Esta solução foi dialisada em tampão Tris (50 mM Tris-HCl, 1,0 mM EDTA, 50 mM NaCl pH 8,0). A solução da proteína em solução com uréia foi colocada em membrana de diálise com poros de 10 kDa e imersa em um béquer contendo o tampão de diálise, sem a uréia. A solução do béquer foi trocada a cada 2 horas por 4 vezes.

Após a diálise a solução contendo a proteína foi centrifugada a 12000 rpm, 4 °C, durante 15 minutos. Nesta etapa manteve-se o sobrenadante e descartou-se o precipitado. Em seguida a solução foi filtrada; primeiro em papel plegueado duplo e depois usando membrana de poro 0,22 – 0,45 micrometros utilizando-se bomba de vácuo. O filtrado foi coletado e mantido a 4 °C.

4.2.7 Purificação

A purificação foi realizada em um sistema Akta da GE Healthcare utilizando uma coluna de troca-iônica (High Q, Bio-Rad). Tampões: A) 50 mM Tris-HCl, 50 mM NaCl, 1,0 mM EDTA pH 8 e B) 50 mM Tris-HCl, 1,0 M NaCl, 1,0 mM EDTA pH 8,0.

Procedeu-se à purificação através de gradiente salino 50 mM a 1,0 M de NaCl, com um passo na concentração de 0,3 M. A proteína foi monitorada com detector de UV em um comprimento de onda a 280 nM. As frações que apresentaram absorvância em 280 nM foram coletadas e realizaram-se testes de atividade catalítica com o substrato ZFRMCA.

As frações contendo atividade foram concentradas por centrifugação em membranas do tipo Amicon para um volume de 2 mL e uma nova etapa de purificação por cromatografia por exclusão de tamanho foi utilizada com uma coluna Sephadex-200 (Bio-Rad). O tampão utilizado foi 50 mM Tris-HCl, 150 mM NaCl, 1,0 mM EDTA pH 8,0. A proteína foi monitorada com detector de UV em um comprimento de onda a 280 nM. As frações que apresentaram absorvância em 280

nM foram coletadas e realizaram-se testes de atividade catalítica com o substrato ZFRMCA.

4.2.8 Gel de eletroforese (SDS-page)

As amostras coletadas em cada etapa dos processos de expressão e purificação foram diluídas em tampão de amostra para eletroforese (Tris-HCl 50 mM, pH 6,8, DTT 100 mM, SDS 2,0 % m/v, azul de bromofenol 0,1% e glicerol 10 %), fervidas durante 5 minutos e centrifugadas por 2 minutos em uma microcentrífuga *Eppendorff* em temperatura ambiente. Os géis de empilhamento e resolução foram preparados como descrito na TABELA 4.2.

TABELA 4.2. Preparo do gel SDS-PAGE 15% para monitoramento da enzima.

Gel de empilhamento	Tris-HCl 2,0 M pH 6,8	125 µL
	SDS 10% m/v	20 µL
	Acrilamida bis-acrilamida (30%)	300 µL
	Água ultra pura (Mili-Q)	1,52 mL
	Temed (PA)	5,0 µL
	Persulfato de amônia (PA)	18,7 µL
Gel de resolução	Tris-HCl 2,0 M pH 8,8	1,5 mL
	SDS 10% m/v	80 µL
	Acrilamida bis-acrilamida (30%)	4,0 mL
	Água ultra pura	2,3 mL
	Temed (PA)	10 µL
	Persulfato de ammonia (PA)	50 µL

O experimento foi realizado em voltagem constante de 175 Volts por aproximadamente 45 minutos. O gel foi corado por 20 minutos em etanol: ácido acético: água 5:1:15 (v:v:v) e *Comassie Brilliant Blue R* 0,25 % v/v e descorado em ácido acético:etanol:água 3:2:35 (v:v:v).

4.2.9 Ativação e estocagem

A enzima purificada foi ativada utilizando o tampão 100 mM fosfato de sódio, 0,9 M de NaCl, 10 mM de DTT, 10 mM de EDTA e pH 6,3. Deixou a solução em temperatura de 37 °C por 16 h. Em seguida adicionou-se 20 % v/v de glicerol e

congelou-se a proteína gradativamente, passando por 4 °C, -6 °C e -20 °C. E enzima foi mantida em estoque congelada a -20 °C.

4.2.10 Determinação da curva de Michaelis-Menten para a enzima cruzaina

O tampão utilizado foi fosfato de sódio 100 mM em pH 6,3, com 100 mM de NaCl, 10 mM EDTA, 5,0 % v/v de DMSO, 5,0 mM de DTT. A concentração da enzima cruzaina foi 2,0 nM e de substrato entre 0,1 a 10 µM.

A enzima foi incubada com a solução tampão contendo DTT por 5 minutos antes de adicionar o substrato, a hidrólise do substrato foi acompanhada por fluorescência com a excitação em 355 nm e emissão 460 nm durante 6 minutos. A velocidade foi determinada pela lei da velocidade inicial. Todos os experimentos foram realizados no mínimo em triplicatas para avaliação do erro experimental.

4.2.5 Ensaios de inibição das enzimas cruzaina e catepsina-L para determinação de IC₅₀

As condições experimentais para as enzimas foram:

Cruzaina - tampão fosfato de sódio 100 mM em pH 6,3, 100 mM de NaCl, 10 mM EDTA, 5,0 % v/v de DMSO, 5,0 mM de DTT e triton X-100 0,01 % v/v.

Catepsina-L - tampão acetato de sódio 100 mM em pH 5,5, 100 mM de NaCl, 10 mM EDTA, 5,0 % v/v de DMSO, 5,0 mM de DTT e triton X-100 0,01 % v/v.

As concentrações das enzimas utilizadas foram 2,0 nM e do substrato ZFRMCA 2,0 µM.

Os valores de concentração para inibir 50% da atividade enzimática (IC₅₀) foram determinados através da curva dose-resposta. As concentrações dos inibidores nos ensaios variaram de 0,1 a 150 µM. Estas concentrações, do mais concentrado para o menos concentrado, foram realizadas segundo um esquema de diluição serial na proporção de diluição 2 vezes, partindo da concentração mais alta de 150 µM [188]. O procedimento utilizado para ambas as enzimas foi incubar as enzimas na solução tampão contendo DTT por 5 minutos, em seguida o inibidor foi incubado também por 5 minutos antes de adicionar o substrato, a hidrólise do substrato foi acompanhada por fluorescência com a excitação em 355 nm e emissão

460 nm durante 6 minutos. A velocidade foi determinada pela lei da velocidade inicial. Todos os experimentos foram feitos no mínimo em duplicata para avaliação do erro experimental. As curvas de percentual de inibição em função da concentração das substâncias testadas foram ajustadas na curva dose-resposta utilizando o programa Origin 7.0. [212]

4.3 Determinação da constante de inibição aparente (K_i^{app})

As curvas de Michaelis-Menten para a enzima cruzaina foram determinadas utilizando o tampão fosfato de sódio 100 mM em pH 6,3, 100 mM de NaCl, 10 mM EDTA, 5,0 % v/v de DMSO, 5,0 mM de DTT. A concentração da enzima cruzaina foi 2 nM e de substrato variou entre 0,1 a 10 μ M.

As curvas de Michaelis-Menten para a enzima catepsina-L de humanos foram determinadas utilizando o tampão acetato de sódio 100 mM em pH 5,5, 100 mM de NaCl, 10 mM EDTA, 5,0 % v/v de DMSO, 5,0 mM de DTT. A concentração da enzima catepsina-L foi 2,0 nM e de substrato variou entre 0,1 a 10 μ M. O procedimento utilizado para ambas as enzimas foi incubar as enzimas na solução tampão contendo DTT por 5 minutos, em seguida o inibidor foi incubado também por 5 minutos antes de adicionar o substrato, a hidrólise do substrato foi acompanhada por fluorescência com a excitação em 355 nm e emissão 460 nm durante 6 minutos. A velocidade foi determinada pela lei da velocidade inicial. Todos os experimentos foram realizados no mínimo em triplicatas para avaliação do erro experimental.

Os inibidores foram testados nas concentrações de 30, 50 e 70 μ M. As curvas foram ajustadas nos modelos de inibição competitivo, incompetitivo e não-competitivo utilizando o programa *SigmaPlot* 10.0, módulo cinético. [213]

4.4 Ensaio *in vitro* contra *T. cruzi* na forma Tripomastigota

O ensaio foi realizado misturando-se 90 μ L de sangue de camundongo swiss fêmea infectado com *T. cruzi* e 10 μ L da solução da substância a ser testada.

Inicialmente dissolveu-se a substância em DMSO puro em uma concentração igual a 50 mM, para o estoque inicial. Em seguida diluiu-se esta solução em um tampão fosfato salino (PBS) NaCl 137 mM, Fosfato 10 mM, KCl 2,7 mM, e um pH

de 7,4, a concentração da substância foi igual a 5 mM em uma solução PBS contendo DMSO 10 %. Esta solução estoque foi utilizada para a mistura com o sangue, de modo que a concentração de DMSO no ensaio foi de 1%. A substância teve sua concentração superior em 500 μ M e foi diluída serialmente (2X) por cinco vezes até a concentração inferior de 15,6 μ M.

As substâncias foram incubadas com o sangue do parasito em uma estufa com temperatura em 37° C por 24 horas. O resultado foi obtido através da contagem do número de parasitos, adicionando-se 5 μ L do sangue incubado com a substância em uma lamínula e contando-se o número de parasitos em 20 campos. O controle negativo foi PBS com DMSO 1,0 % v/v e controle positivo Benzonidazol[®] nas mesmas concentrações das substâncias.

4.6 Resultados e Discussões

A obtenção do alvo macromolecular para ensaios *in vitro* é uma etapa fundamental no processo de descoberta de substâncias bioativas. Este projeto foi desenvolvido em colaboração com o laboratório de Biofísica da Universidade Federal de São Paulo (UNIFESP) coordenado pelo Prof. Dr. Luiz Juliano Neto, forneceu o DNA plasmidial estocado em bactéria *E. coli* linhagem DH5 α . O plasmídeo foi originalmente construído utilizando o vetor pUC18 de forma que a expressão fosse realizada nesta linhagem de bactéria [57]. Entretanto, visando um aumento do rendimento da enzima, foram utilizadas novas linhagens de *E. coli*, para que altos níveis de expressão pudessem ser alcançados. [214]

As linhagens de *E. coli* BL21-DE3 e BL21-pLysS são apropriadas para a expressão da cruzaina. Elas possuem a vantagem de serem deficientes das proteases lon e ompT, além da compatibilidade com o promotor operon Lac T7. O antibiótico utilizado para a seleção das colônias transformadas é ampicilina, o mesmo utilizado na construção do plasmídeo da cruzaina. [215]

O plasmídeo, mantido intacto durante o processo, foi extraído da bactéria *E. coli* DH5 α utilizando o método padrão de extração em meio alcalino. O DNA plasmidial é solúvel na solução alcalina ao passo que o material celular da bactéria é precipitado nesta solução, em seguida com uma única etapa de purificação por exclusão de tamanho em coluna aberta (gravidade) é obtido o plasmídeo puro. Este

plasmídeo contendo o gene da cruzaina foi utilizado para transformar as bactérias de *E. coli* BL21-DE3 e BL21-pLysS, uma das grandes vantagens da utilização das bactérias é a facilidade de transformação através de choque térmico. Em seguida foi realizado um teste de indução para verificar qual linhagem de bactéria produz o melhor nível de expressão, alíquotas dos meios de cultura induzidos de ambas as linhagens de bactérias foram utilizados para análise em SDS-PAGE 15%, FIGURA 4.7.

As alíquotas utilizadas para a análise continham as mesmas quantidades de meio de cultura para permitir a análise comparativa; a enzima tem massa molecular de 27 kDa.

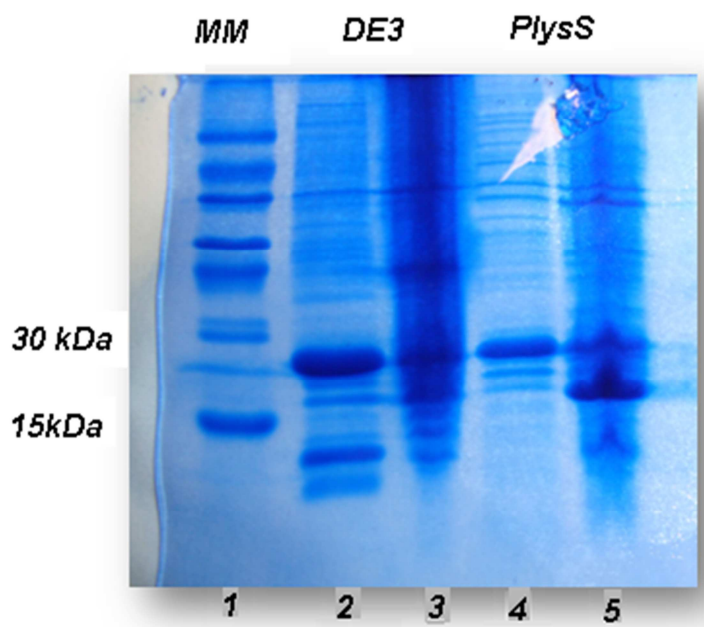


FIGURA 4.7: Gel de Eletroforese da cruzaina. Banda 1) Marcador de peso molecular (MM), 2) amostra induzida com IPTG e 3) amostra não induzida da bactéria *E. coli* BL21-DE3; 4) Amostra induzida com IPTG e 5) Amostra não induzida da bactéria *E. coli* BL21-pLysS

Após a indução é observada uma banda pronunciada entre 15 kDa e 30 kDa, a qual corresponde a enzima cruzaina em ambas as linhagens de bactérias. Entretanto, pode ser observada uma banda mais larga no meio induzido contendo a bactéria *E. coli* BL21-DE3, isto indica que há um maior nível de expressão neste sistema.

Após o sucesso da troca do sistema de expressão, substituição da *E. coli* DH5 α pela *E. coli* BL21-DE3 foi adaptado o protocolo de expressão e purificação da cruzaina baseado no descrito na literatura Eakin et al [57].

Inicialmente, utilizou-se o meio comercial chamado de Luria broth ou LB (Tryptona, Extrato de Levedura e NaCl) para o meio de cultura de crescimento das bactérias *E. coli* BL21-DE3 contendo o DNA plasmidial com o gene da cruzaína. O antibiótico ampicilina foi utilizado para selecionar somente as bactérias transformadas, as quais possuem resistência a este antibiótico quando ocorre inserção do plasmídeo. O agente indutor deste sistema de expressão é o *isopropil beta-D-galactosídeo* (IPTG), sua função ligar-se ao centro alostérico da proteína repressora, inativando-a; isto permite a transcrição dos genes do operon da lactose e, conseqüentemente, a produção das enzimas codificada no DNA plasmidial. Como o IPTG não é degradado pelas proteínas presentes na bactéria, a enzima vai sendo continuamente produzida e conseqüentemente a sua concentração aumenta na célula.

As bactérias induzidas super expressam a enzima de interesse, e após esta etapa são submetidas ao processo de lise celular para extração e purificação da enzima. Para a lise é utilizado lisozima que atua no rompimento da parede celular, PMSF é um inibidor de protease, o qual garante a integridade da enzima de interesse, Dnase hidrolisa DNA para facilitar o processo de purificação da enzima por troca-iônica e TritonX-100 é utilizado para solubilizar as proteínas da bactéria separando-as do material insolúvel. A principal limitação da produção da enzima é o fato dela ser expressa em corpos de inclusão, os quais são depositados no precipitado após a lise celular. A etapa de lavagem do precipitado com TritonX-100 é utilizada para a separação de proteínas solúveis o que facilita a purificação da cruzaína.

Para a cruzaína ser solubilizada é necessário a dissolução dos corpos de inclusão, isto exige uma solução de guanidina ou uréia 8,0 M e um pH alcalino de 10,7, esta alta concentração de uréia resulta no desenovelamento da proteína. Originalmente o renovelamento era realizado através de uma etapa de diluição 10 vezes da solução de enzima contendo a uréia e em seguida uma diminuição lenta do pH de 10,7 para 8,0. O protocolo foi modificado e o renovelamento foi realizado através de diálise, desta forma a uréia ao invés de ter sua concentração reduzida em 10 vezes, passou a ter 10^4 vezes, pois foi trocado a solução de diálise quatro vezes. Além disso, é essencial a redução lenta do pH para permitir as interações intramoleculares durante o processo de renovelamento, a diálise demora cerca de pelo menos 8 h ao passo que anteriormente era feito através de um gotejamento

lento de cerca de 1 h. Essa modificação do protocolo além de visar uma melhora no renovelamento melhora também o aspecto técnico. Ao final desta etapa a solução dialisada possui um volume de cerca de 40 mL que pode ser colocada em um sistema de super-alça de 50 mL e purificada em uma única etapa, anteriormente o volume era de 400 mL e exigia pelo menos 8 etapas de preenchimento do super-alça.

A enzima cruzaina possui ponto isoeletrônico de 5,5, em pH básico adquire cargas negativas, pois doa prótons para o tampão e adquire afinidade pela coluna de troca aniônica. Desta forma a enzima é purificada por cromatografia de troca aniônica com gradiente salino, o cromatograma da FIGURA 4.8 ilustra o perfil de eluição da enzima.

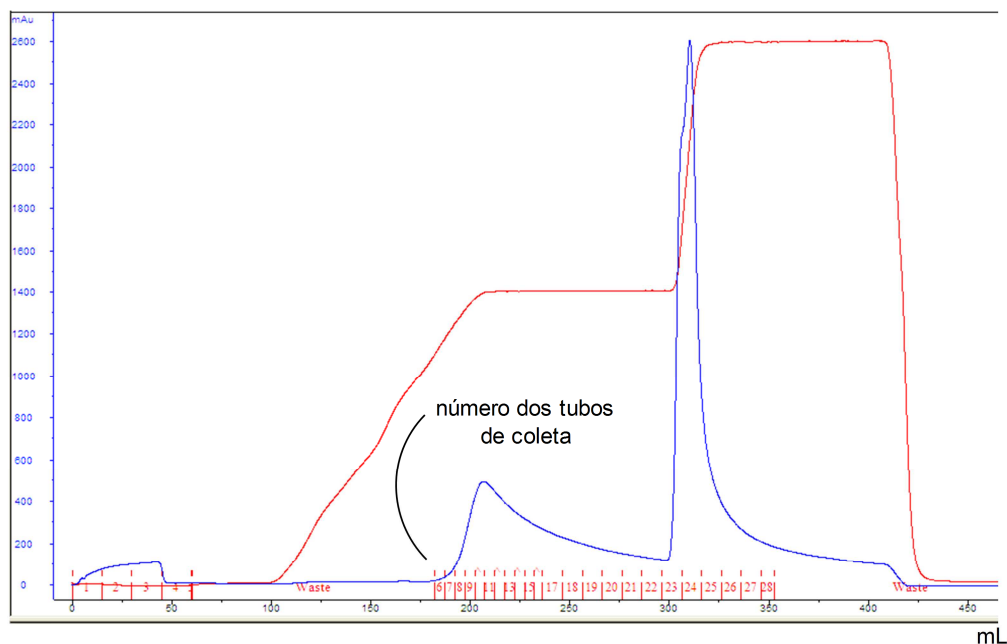


FIGURA 4.8: Cromatograma da purificação por troca-iônica. A linha azul representa a absorbância em 280, a linha vermelha a razão de NaCl, e os números em vermelho os tubos coletados para análise em SDS-page

A presença de dois picos e uma banda larga no cromatograma indica que a enzima não eluiu totalmente pura, foi utilizado SDS-page para verificação da pureza das frações que apresentaram absorbância em 280 nM, FIGURA 4.9.

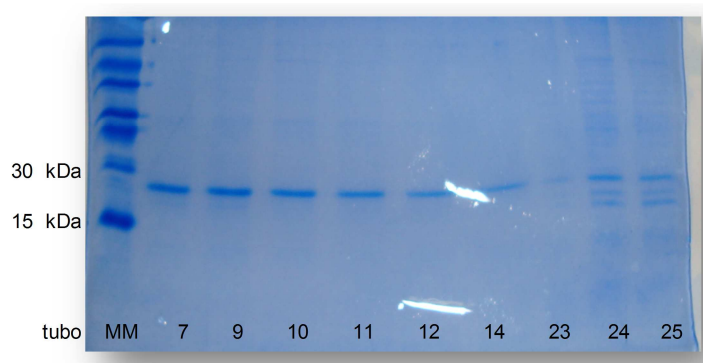


FIGURA 4.9: Gel de Eletroforese das frações coletadas na cromatografia de troca-iônica. Os números representam o número dos tubos onde se coletou as frações que apresentaram absorvância em 280 nm

A análise do gel da FIGURA 4.9 revela que a enzima cruzaina está presente em todas as frações, entretanto há também impurezas, demonstradas pelas bandas mais fracas em regiões de maior massa molecular. Adicionalmente foi realizada uma etapa de cromatografia por exclusão de tamanho (gel filtração), visando uma maior pureza da amostra da enzima. As frações que apresentaram absorvância em 280 nm foram analisadas utilizando SDS-page, FIGURA 4.10.

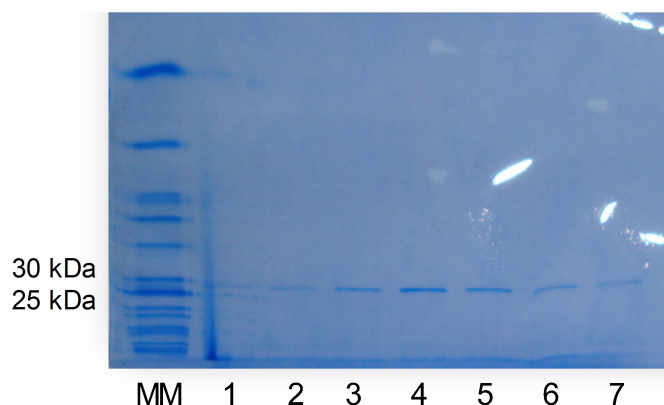


FIGURA 4.10: Gel de Eletroforese das frações coletadas na cromatografia de exclusão por tamanho. Banda (MM) Marcador de peso molecular, 1 ao 7 frações que apresentaram absorvância em 280 nm

As frações 3 a 6 apresentaram bandas limpas e correspondem a enzima com um alto grau de pureza. Estas frações foram utilizadas nos ensaios cinéticos, após a verificação da atividade catalítica da hidrólise do substrato ZFRMCA e determinação dos parâmetros cinéticos de Michaelis-Menten.

4.7 Determinação da curva de Michaelis-Menten para as enzimas cruzaina e catepsina-L

Previamente aos ensaios de inibição enzimática os parâmetros de Michaelis-Menten das enzimas cruzaina e catepsina-L foram determinados pelo método padrão da fluorimetria para a verificação do funcionamento apropriado das condições experimentais e dos lotes de enzima utilizados neste trabalho.

O substrato ZFRMCA possui uma baixa solubilidade em água e foi necessário utilizar DMSO 5,0 % v/v nos ensaios. As curvas cinéticas das enzimas cruzaina e catepsina-L são apresentadas na FIGURA 4.11

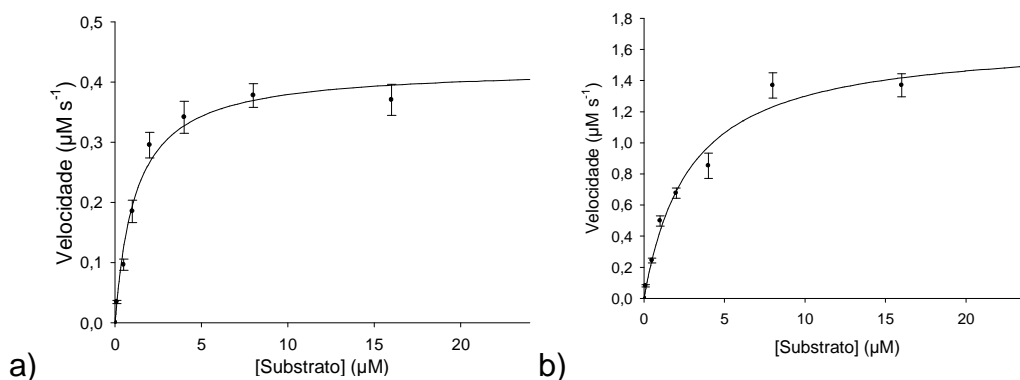


FIGURA 4.11: Curvas da velocidade em função da concentração de substrato ZFRMCA ajustadas na Equação de Michaelis-Menten a) cruzaina e b) catepsina-L

Os valores das constantes cinéticas da cruzaina, juntamente com os dados do ajuste, são apresentados na Tabela 4.3.

TABELA 4.3. Valores das constantes cinéticas e dados do ajuste na equação de Michaelis-Menten da reação catalisada pela enzima cruzaina

Parâmetros	Cruzaina		Catepsina-L	
	Valor com desvio padrão	95% Intervalo de Confiança	Valor com desvio padrão	95% Intervalo de Confiança
V_{max} ($\mu\text{M s}^{-1}$)	$0,42 \pm 0,02$	0,38 a 0,46	$1,67 \pm 0,09$	1,47 a 1,89
K_M (μM)	$1,18 \pm 0,20$	0,76 a 1,60	$2,32 \pm 0,28$	1,74 a 2,90
k_{cat} (s^{-1})	$6,87 \pm 1,23$	5,24 a 7,98	$10,1 \pm 1,21$	9,58 a 12,7
k_{cat}/K_M ($\times 10^6 \text{ s}^{-1} \text{ M}^{-1}$)	5,82		4,35	
R^2	0,96		0,97	
$Sy.x^*$	0,05		0,01	

Sy.x. Desvio padrão dos resíduos

Os valores obtidos das constantes cinéticas corroboraram com os valores da literatura para ambas as enzimas. A enzima catepsina-L, de origem comercial, possui maiores valores de K_M e k_{cat} , entretanto a eficiência catalítica é muito próxima. A reprodução dos valores cinéticos da cruzaina atesta a atividade cinética da enzima, e conseqüentemente pode-se inferir que a enzima possui sua conformação e estado de enovelamento corretos, o que significa sucesso no protocolo de expressão e purificação da cruzaina no sistema de expressão da bactéria *E. coli* BL21-DE3.

Além da atividade da enzima outro fator importante a ser mencionado é o rendimento da proteína, o qual foi quantificado pelo método do coeficiente de extinção molar. O protocolo da literatura relata o rendimento de 1,0 mg de proteína para 1,0 litro de indução, no protocolo deste trabalho 0,5 L de indução rendeu esta mesma quantia. A enzima cruzaina obtida foi utilizada para os estudos de inibição da enzima. Paralelamente, os mesmos estudos foram realizados com a enzima catepsina-L de humanos para verificação da seletividade dos inibidores identificados.

4.8 Ensaio de inibição das enzimas cruzaina e catepsina-L

A determinação da inibição das enzimas pelos compostos selecionados pelos métodos computacionais foi realizada por curvas dose-resposta para medir a concentração do composto que inibe 50 % da atividade enzimática (IC_{50}). O IC_{50} é uma medida de potência entre os inibidores ensaiados na mesma condição experimental. Fatores como pH, força iônica e temperatura podem causar mudanças no valor de IC_{50} , e por isso devem ser mantidos constantes a fim de gerar dados comparáveis para uma série de compostos. Outro fator importante que influencia o valor de IC_{50} é a concentração do substrato, a FIGURA 4.12 ilustra a influência da concentração do substrato em inibidores que apresentam distintos mecanismos de inibição.

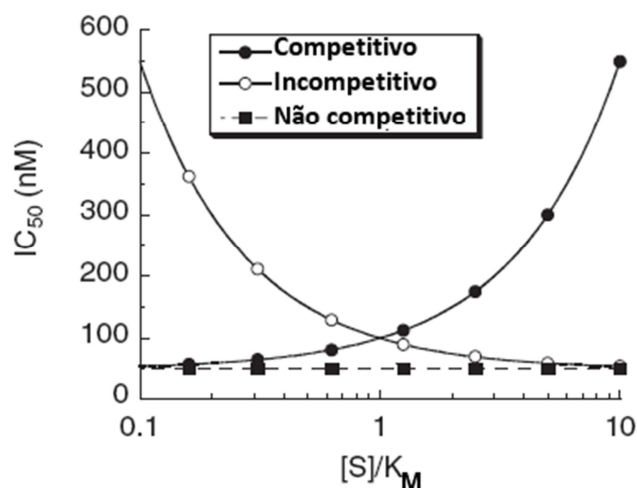


FIGURA 4.12: Efeito da razão $[S]/K_M$ no valor de IC_{50} em distintos mecanismos de inibição [216]

Este exemplo é uma simulação do efeito da razão $[S]/K_M$ no valor de IC_{50} para um inibidor que possui alta afinidade pela enzima ($K_d = 50 \text{ nM}$) [216]. A concentração de substrato que fornece a análise mais abrangente e o valor mais próximo do K_d é a de K_M ($[S] = K_M$), por esta razão a concentração de ZFRMCA utilizada para os ensaios foi $2,0 \mu\text{M}$. Para que a enzima opere em estado estacionário é necessário que a concentração de substrato seja pelo menos 1000 vezes maior que a de enzima, conseqüentemente a concentração das enzimas utilizadas nos ensaios foi $2,0 \text{ nM}$.

A qualquer concentração de inibidor, a concentração total de enzima é a soma da enzima livre e o complexo enzima-inibidor. A atividade fracional da enzima v_i/v_b (v_i atividade sem inibidor e v_b com inibidor) está diretamente relacionada com a razão entre a concentração de enzima livre e a concentração total de enzima. A fração ocupada pelo inibidor, segundo o balanço de massas, será $1 - v_i/v_b$ a uma concentração fixa de enzima e substrato; a velocidade da reação diminui com o aumento da concentração do inibidor até não haver atividade residual (exceto em caso de inibição parcial). Uma curva da fração da velocidade (ou percentual de inibição da enzima) em função da concentração do inibidor terá um formato sigmoidal (perfil de curva dose-resposta); o ponto médio da curva, onde a fração da velocidade é igual a 0,5 corresponde a 50 % de inibição da enzima (IC_{50}).

As curvas do percentual de inibição em função da concentração dos inibidores foram determinadas para todos os compostos selecionados para os testes

contra as enzimas cruzaina e catepsina-L. O percentual de inibição em cada ponto de concentração foi calculado utilizando-se a Equação 4.9. [217]

$$\% \text{ inibição} = 100 \left(1 - \frac{v_i}{v_b} \right) \quad (4.9)$$

onde v_i é a velocidade na presença de inibidor e v_b a velocidade do experimento controle. A FIGURA 4.13 mostra o perfil das curvas com o ajuste matemático da equação da curva dose-resposta dos pontos determinados experimentalmente para as substâncias que apresentaram atividade inibitória frente à enzima cruzaina.

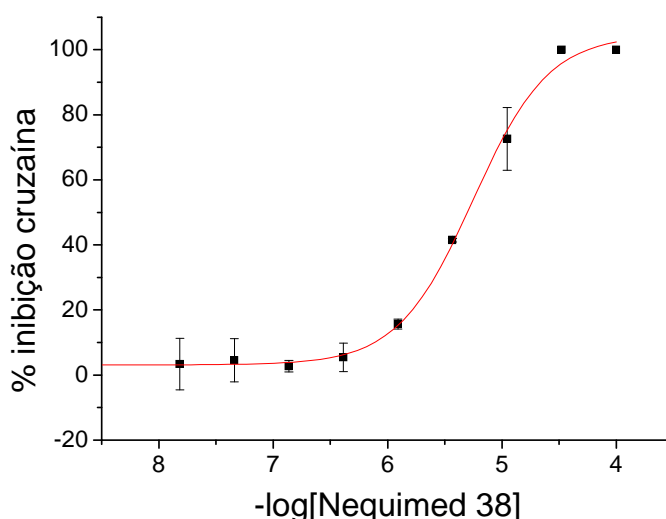


FIGURA 4.13: Curva da inibição da enzima cruzaina (ajuste $R^2 = 0,99$; $Q^2/\text{DoF} = 0,71$)

Os valores do percentual de inibição das enzimas foram considerados, a variável dependente (eixo Y) enquanto os valores da concentração dos inibidores foram calculados os logaritmos e considerados as variáveis independentes (eixo X).

Os valores de IC_{50} foram calculados utilizando a Equação 4.10, o método de ajuste foi não-linear dos mínimos quadrados.

$$\% \text{ inibição} = A_0 + \frac{(A_1 - A_0)}{1 + 10^{(\text{Log}IC_{50})h}} \quad (4.10)$$

onde A_0 é o patamar inferior da curva, A_1 é o patamar superior, $\text{Log}IC_{50}$ é o logaritmo da concentração de IC_{50} e h o coeficiente de Hill. Os valores de A_1 e A_0

representam a resposta máxima e ausência de resposta, respectivamente, da enzima na presença do inibidor. O coeficiente de Hill está relacionado à estequiometria da interação, a cooperatividade entre monômeros de enzimas alostéricas (no caso de enzimas multiméricas) e também ao grau da relação dose-resposta: coeficiente igual a 1 indica interação completamente independente entre monômeros, maior que 1 cooperatividade positiva e menor que um negativa. A enzima cruzaina é monomérica e, portanto não é esperado o fenômeno de cooperatividade, idealmente o perfil de inibição dos compostos deve apresentar o coeficiente de Hill próximo a 1. Além do grau de cooperatividade o coeficiente de Hill pode ser um diagnóstico de comportamento de inibição não-ideal. Compostos que causem uma variação abrupta ($h \gg 1$), usualmente refletem um mecanismo não específico de inibição, como agregação por exemplo. [216]

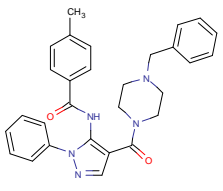
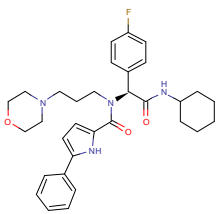
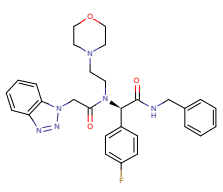
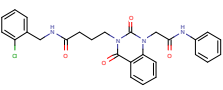
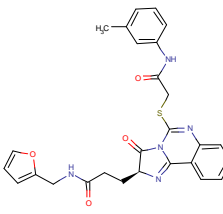
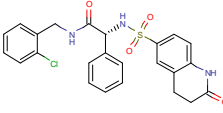
Além da potência existe outro parâmetro amplamente difundido para comparação de ligantes provenientes de diferentes séries e classes de compostos nas etapas de identificação e otimização do ligante chamada Eficiência do Ligante (LE). A definição de Eficiência do Ligante é a energia livre de interação entre um ligante e o alvo por cada átomo diferente de hidrogênio (comumente chamados de átomos pesados). LE pode ser calculada por:

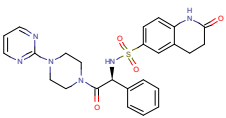
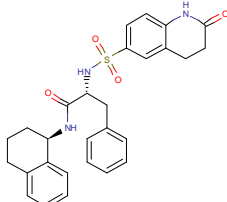
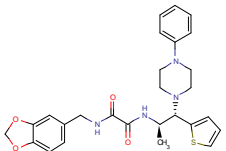
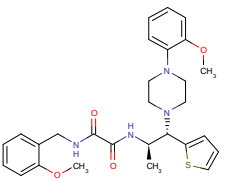
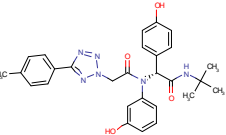
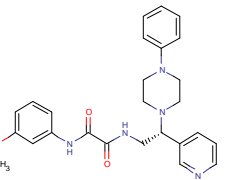
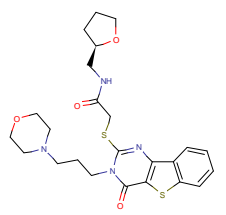
$$LE = \frac{-\Delta G}{NA} = \frac{-RT \ln K_i}{NA} \approx \frac{-RT \ln IC_{50}}{NA} \quad (4.11)$$

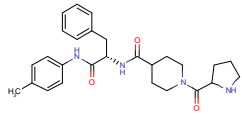
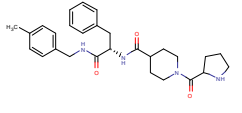
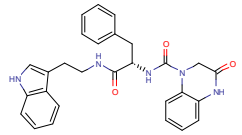
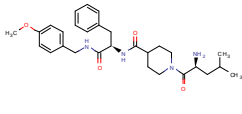
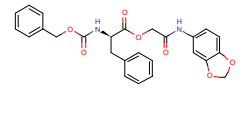
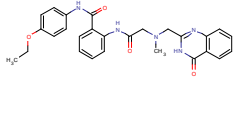
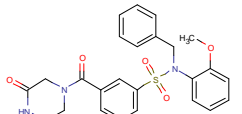
onde NA é o número de átomos diferente de hidrogênio, R a constante dos gases, T a temperatura em Kelvin, K_i a constante de inibição e IC_{50} a concentração necessária para inibir 50 % da atividade enzimática. A unidade de LE é $\text{kcal mol}^{-1} \text{NA}^{-1}$. Se consideramos um fármaco com alta potência $IC_{50} < 10 \text{ nM}$ e massa molar de 500 Daltons (para satisfazer a regra de Lipinski) o valor mínimo a ser atingido de LE é $0,3 \text{ kcal mol}^{-1} \text{NA}^{-1}$. [218-219]

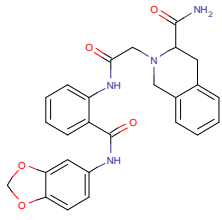
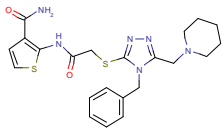
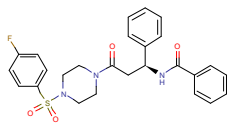
Os resultados da inibição das substâncias selecionadas para os ensaios, estruturas, parâmetros de Lipinski e LE são apresentados na TABELA 4.4. As substâncias adquiridas comercialmente, provenientes de produtos naturais ou sintetizadas pelo NEQUIMED-PN são catalogadas na forma de um banco de dados, os códigos utilizados para se referir às substâncias nesse trabalho são correspondentes a este banco.

TABELA 4.4: Estruturas, valores de IC₅₀, LE e parâmetros de Lipinski das substâncias testadas frente às enzimas cruzaina e catepsina-L

Estrutura	Códigos Nequimed	IC ₅₀ (μM) cruzaina	IC ₅₀ (μM) catepsina-L	NA	LE	MM	clogP	HA	HD
	Nequimed21	Inativa	Inativa	36	-	480,6	4,2	7	2
	Nequimed22	Inativa	Inativa	40	-	547,7	4,8	7	3
	Nequimed23	Inativa	Inativa	39	-	531,6	2,5	9	2
	Nequimed24	9,4 ± 1,2	8,6 ± 1,1	36	0,20	504,9	3,3	8	2
	Nequimed25	40,5 ± 1,3	51,0 ± 3,8	38	0,16	517,6	3,3	9	4
	Nequimed26	63,2 ± 1,1	50,9 ± 2,2	33	0,18	483,9	3,2	7	3

	Nequimed27	Inativa	Inativa	36	-	506,6	0,5	10	2
	Nequimed28	Inativa	Inativa	36	-	503,6	3,5	7	3
	Nequimed29	$17,4 \pm 6,6$	$6,6 \pm 1,1$	36	0,19	507,6	3,2	8	3
	Nequimed30	$7,5 \pm 2,3$	$9,9 \pm 1,3$	37	0,20	523,7	3,1	8	3
	Nequimed31	Inativa	Inativa	36	-	514,6	3,3	10	3
	Nequimed32	$58,7 \pm 8,3$	$51,8 \pm 9,8$	34	0,18	460,5	1,9	8	3
	Nequimed33	$73,9 \pm 6,3$	$81,3 \pm 8,8$	34	0,17	503,6	2,1	8	3

	Nequimed34	Inativa	Inativa	34	-	463,6	2,5	7	3
	Nequimed35	Inativa	Inativa	35	-	477,6	2,6	7	3
	Nequimed36	48,3 ± 9,2	20,4 ± 4,9	36	0,17	481,5	1,7	8	3
	Nequimed37	73,9 ± 3,2	81,5 ± 9,3	37	0,16	509,7	4,1	9	5
	Nequimed38	5,6 ± 1,1	31,1 ± 1,1	35	0,21	476,5	3,5	9	2
	Nequimed39	Inativa	Inativa	36	-	486,5	3,0	9	4
	Nequimed40	Inativa	Inativa	34	-	479,5	2,6	8	1

	Nequimed41	63,1 ± 14	89,1 ± 21	35	0,17	473,5	1,9	9	4
	Nequimed42	21,6 ± 5,4	35,1 ± 11	32	0,21	471,6	3,3	8	3
	Nequimed43	Inativa	Inativa	35	-	495,6	3,0	7	1

NA – número de átomos diferentes de hidrogênio

LE – eficiência do ligante para a enzima cruzaina, unidade kcal mol⁻¹ NA⁻¹

MM – massa molecular, unidade g mol⁻¹

clog P – coeficiente de partição calculado

HA – número de aceitadores de ligação hidrogênio

HD – número de doadores de ligação hidrogênio

A potência das substâncias inibidoras da enzima cruzaina variou entre 5,6 ± 1,1 µM e 73,9 ± 3,2 µM. As substâncias inibiram a enzima catepsina-L com potência muito próxima à cruzaina, a substância Nequimed38 foi 6 vezes mais potente frente a cruzaina em relação a catepsina-L, um índice baixo para definir seletividade. A eficiência média do ligante para as substâncias ativas foi 0,18 kcal mol⁻¹ NA⁻¹, este parâmetro é um indicativo das interações intermoleculares efetivas entre os átomos do ligante e a enzima, um valor baixo como foi observado mostra que há átomos que não contribuem positivamente para a interação. Uma vez que a massa molecular das substâncias é de aproximadamente 500 Da a otimização da potência não pode ser feita adicionando-se novos grupos para a interação, a melhor estratégia é identificar os átomos que não contribuem positivamente para a interação e substituí-los por átomos com interações favoráveis.

Trabalhos recentes mostram a busca de inibidores da cruzaina no banco ZINC utilizando métodos virtuais. Ferreira *et al.* através de docagem molecular

encontraram um composto ativo dentre 17 testados com potência $IC_{50} = 77 \mu M$ [220] e Malvezzi *et al.* encontraram um composto com potência baixo micromolar dentre 19 selecionados utilizando um modelo farmacofórico. [221]

Comparativamente, o número de inibidores encontrados neste trabalho empregando os métodos *in silico* é superior. Isto pode ser atribuído à estratégia de utilizar o consenso entre diferentes métodos baseados na estrutura do receptor e do ligante. O espaço químico disponível para a busca de inibidores da cruzaina consistiu de um banco com cerca de 8,5 milhões de estruturas, a busca por compostos com atividade biológica, que podem ser definidos como moléculas que pertencem ao espaço químico-biológico não é uma tarefa trivial. Os modelos desenvolvidos computacionalmente permitiram encontrar 12 ligantes em apenas 23 testados e desta maneira pode-se concluir que o espaço químico-biológico dos compostos capazes de inibir a cruzaina foram descritos eficientemente. Esta etapa de identificação do ligante foi realizada com sucesso, pois o resultado é muito superior ao ensaio aleatório.

Um aspecto importante a ressaltar é que os compostos ativos contra a cruzaina possuem características estruturais que satisfazem as regras de Lipinski, A área superficial polar dos compostos também foi calculada com os valores variando entre 98 e 137 \AA^2 e estão dentro dos valores desejados para que um composto seja passível de atravessar a membrana biológica ($<140 \text{\AA}^2$). [222]

4.9 Determinação da constante de inibição

Uma das etapas cruciais para a identificação do ligante é a determinação do mecanismo de inibição de constante de inibição. O mecanismo de inibição é um indício do modo de interação do ligante com a enzima e a constante de inibição é um parâmetro relacionado com a constante de associação dos complexos enzima-inibidor e independente da concentração da substância. Desta forma, é o parâmetro mais apropriado para estudos de estrutura atividade, onde é necessária a comparação da atividade entre diferentes substâncias.

As constantes de inibição foram determinadas através da curva de Michaelis-Menten na ausência de inibidor, e em seguida nas concentrações de 30, 50 e 70 μM , valores de concentração próximos ao IC_{50} . A FIGURA 4.14 ilustra o perfil das curvas

cinéticas obtidas ajustadas no gráfico de Lineweaver-Burk para melhor visualização do mecanismo de inibição, entretanto para os cálculos foram utilizados o método não-linear.

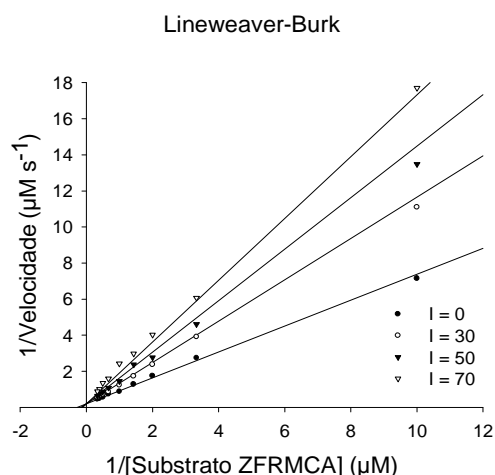


FIGURA 4.14: Curvas da velocidade em função da concentração de substrato ajustadas nas equações de mecanismo competitivo, não-competitivo e incompetitivo, o ensaio foi realizado na ausência dos inibidores e nas concentrações de 30, 50 e 70 μM .

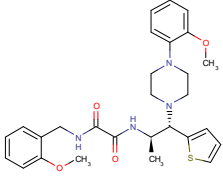
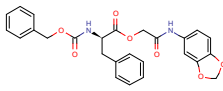
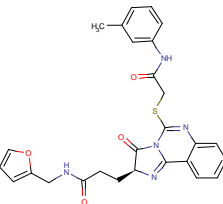
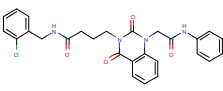
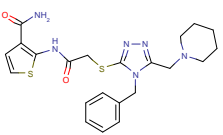
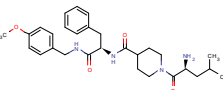
O mecanismo de inibição e K_i^{app} foram determinados ajustando os valores experimentais das velocidades em diferentes concentrações do substrato ZFRMCA nas equações para inibição reversível competitiva, não-competitiva e incompetitiva, analisaram-se os ajustes através dos parâmetros estatísticos, como exemplificado na TABELA 4.5.

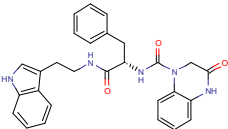
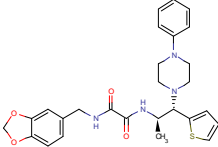
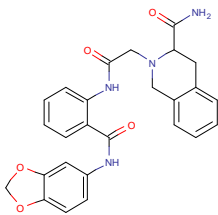
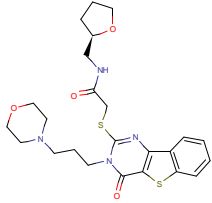
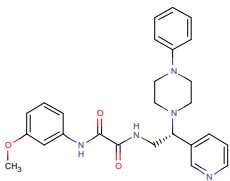
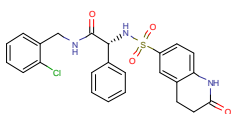
TABELA 4.5: Valores dos parâmetros estatísticos dos ajustes das velocidades experimentais em função da concentração de substrato nas equações para inibidores competitivo, não-competitivo e incompetitivo

R ²	Equação	R ²	AICc	Sy.x	Teste
1	Competitivo	0.982	-172	0.00838	Aprovado
2	Não-competitivo	0.979	-168	0.00888	Reprovado
3	Incompetitivo	0.958	-143	0.12575	Reprovado

O mecanismo de inibição observado para todos os inibidores testados foi o competitivo, as constantes de inibição foram determinadas para as enzimas cruzaina e catepsina-L, as estruturas dos compostos e seus valores estão listados na TABELA 4.6.

TABELA 4.6. Estrutura e valores de K_i dos compostos testados contra as enzimas cruzaina e catepsina-L de humanos, o mecanismo observado para as moléculas ativas foi o competitivo

Estrutura	Códigos Nequimed	K_i^{app} (μM) Cruzaina	LE Cruzaina	K_i^{app} (μM) Catepsina-L	LE Catepsina-L
	Nequimed30	$3,7 \pm 1,4$	0,21	$16,7 \pm 3,4$	0,18
	Nequimed38	$5,8 \pm 0,6$	0,21	$31,1 \pm 1,1$	0,18
	Nequimed25	$6,92 \pm 0,8$	0,19	$52,8 \pm 14$	0,16
	Nequimed24	$16,5 \pm 1,7$	0,19	$3,8 \pm 0,7$	0,21
	Nequimed42	$18,5 \pm 5,5$	0,21	$35,6 \pm 8,2$	0,20
	Nequimed37	$26,3 \pm 3,1$	0,18	$41,1 \pm 14$	0,17

	Nequimed36	30,2 ± 3,7	0,18	19,1 ± 1,1	0,19
	Nequimed29	34,3 ± 3,4	0,18	15,1 ± 4,3	0,19
	Nequimed41	38,8 ± 4,1	0,18	87,1 ± 30	0,16
	Nequimed33	40,1 ± 1,2	0,18	67,6 ± 6,6	0,17
	Nequimed32	50,6 ± 8,3	0,18	44,4 ± 1,8	0,18
	Nequimed26	64,5 ± 14	0,18	63,4 ± 14	0,18

Os valores de K_i^{app} para as moléculas ativas variaram entre 3,7 e 64,5 μM para a enzima cruzaina, os três compostos mais potentes apresentaram uma afinidade pela cruzaina um pouco maior que em relação à catepsina-L, cerca de 6 vezes.

Um dos objetivos deste projeto foi encontrar novos esqueletos moleculares ativos contra a enzima cruzaina, o que foi alcançado com as substâncias

Nequimed30, Nequimed25, Nequimed24 e Nequimed42. Foi realizada uma pesquisa por subestruturas no banco de dados de fármacos, que os compostos Nequimed30, Nequimed29 e Nequimed32 contêm a diamida comum ao fármaco Ambenônio, utilizado para o tratamento de miastênia gravis atuando como um inibidor da colinesterase, as estruturas são apresentadas na FIGURA 4.15.

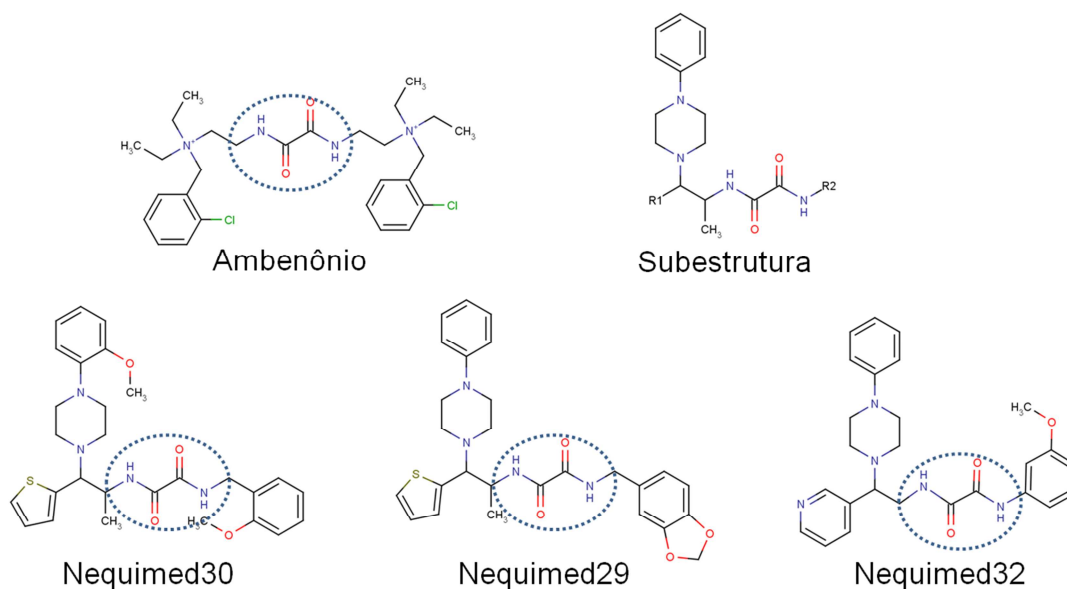


FIGURA 4.15: Estruturas dos compostos ativos contra a cruzaina que contém a eltidiamida, Ambedônio é o farmaco comercial e ao lado a subestrutura comum às três substâncias ativas, os compostos estão dispostos por ordem de afinidade pela cruzaina

Os compostos apresentados na FIGURA 4.15 possuem uma subestrutura comum e indicam a primeira relação entre a estrutura e atividade entre eles, a substituição do grupo tiofeno pela pirimidina resultou no composto menos potente. Os grupos fenila com metoxilas nas posições orto (Nequimed30) resultaram em melhor potência em relação ao grupo 1,3-benzodioxolano em R2 (Nequimed29).

A maioria dos inibidores reportados na literatura são compostos que se ligam covalentemente a enzima, entretanto é conhecido que compostos irreversíveis possuem maiores chances de apresentarem efeitos indesejados devido a problemas farmacocinéticos. Um dos compostos identificados pela estratégia de ensaios virtuais aplicados neste trabalho possui uma grande similaridade com substâncias identificadas anteriormente, porém, com uma importante modificação, a substituição dos grupos reativos com o grupo tiol da cisteína catalítica por um grupo incapaz de ligar-se covalentemente, FIGURA 4.16.

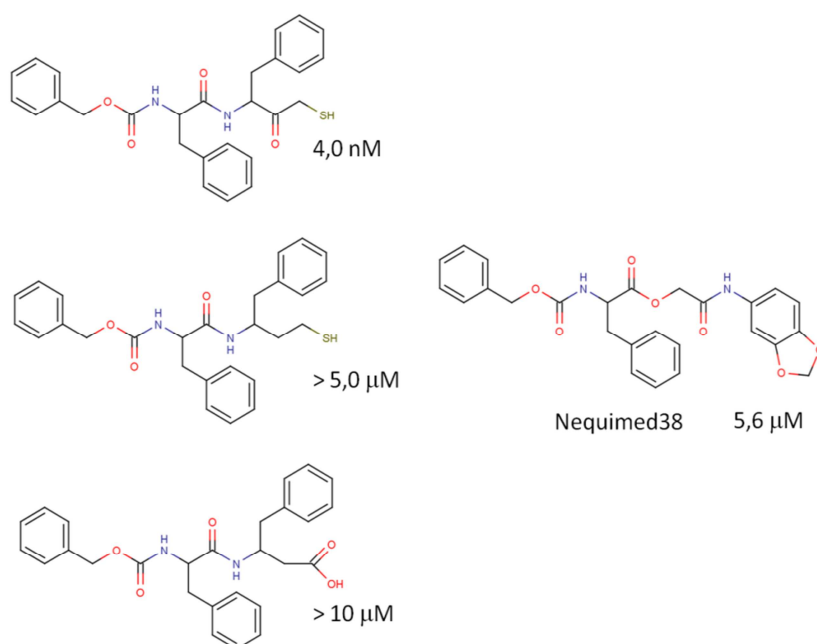


FIGURA 4.16: Comparação entre estruturas e valores de IC₅₀ entre compostos da literatura e o inibidor encontrado neste trabalho Nequimed38

As duas principais diferenças entre os inibidores da literatura [87] e o composto Nequimed38 são, a ausência de um grupo metil fenila e a inserção do grupo 1,3-benzodioxolano. A variação no valor de K_i^{app} dentro da série da literatura entre 4 nM e > 10 μM indica que a variação do grupo reativo possui uma influência mais pronunciada que o grupo metil-fenila, uma vez que este faz apenas interações hidrofóbicas com a enzima no sítio denominado como S4. Os compostos 1 e 4 possuem como diferença de um oxigênio cetônico na posição β em relação ao grupo tiol, que resulta em uma diferença na afinidade do composto pela enzima de cerca de mil vezes devido a uma interação de hidrogênio com histidina da tríade catalítica (His159). O oxigênio cetônico é mantido na mesma posição no composto Nequimed38, mesmo assim, a potência é a ordem de micro molar. Estas observações indicam que a ligação covalente com a cisteína é imprescindível para maior potência nesta classe de compostos.

4.10 Relações Estrutura Atividade

Uma etapa imprescindível no processo de descoberta de substâncias bioativas são a confirmação da atividade inibitória de ligantes identificados e o estabelecimento de relações entre a estrutura química e a atividade biológica.

O mecanismo de inibição competitivo e a constante de inibição de substâncias pertencentes a várias classes moleculares foram identificadas, através dos métodos *in silico* e confirmação nos ensaios bioquímicos via fluorimetria. Este importante resultado motiva a confirmação e validação da atividade inibitória através do uso de substâncias similares aos compostos identificados previamente. Foi realizado uma busca por análogos das substâncias pertencentes as novas classes identificadas neste trabalho Nequimed24, Nequimed25, Nequimed30 e Nequimed42 nos bancos de moléculas comerciais ZINC e *eMolecules*. Foram encontradas análogos das substâncias Nequimed24 e Nequimed42. Os modelos computacionais desenvolvidos previamente foram utilizados novamente para a seleção das novas séries de moléculas.

Os parâmetros utilizados para avaliar a atividade das substâncias foram potência (IC_{50}) e eficiência do ligante (LE) e revelaram que existem átomos que não contribuem efetivamente para a atividade biológica, desta forma, as novas séries de substâncias testadas possuíam estruturas simplificadas com remoção de grupos em diferentes posições e variações de substituintes de forma a examinar o modo de interação e o esqueleto responsável pelo reconhecimento molecular.

Uma vez que, o mecanismo de inibição das substâncias ativas foi identificado como competitivo e há somente o sítio catalítico da enzima disponível para interação, assumiu-se que seus análogos possuem o mesmo mecanismo de inibição, desconsiderando possíveis efeitos alostéricos.

Foram adquiridas e testadas dez substâncias análogas ao Nequimed24. O anel dipirimidinona com a amida na posição β do nitrogênio 2, circulado com a elipse em azul na FIGURA 4.17, foi o esqueleto mantido na série, uma vez que o modo de interação predito por docagem molecular indicam que este esqueleto interage na região da tríade catalítica com a amida sendo reconhecida pela Gly66, como é observado nos complexos co-cristalizados. Dentre as dez substâncias testadas, cinco foram ativas; no entanto, menos potentes que a substância de partida. Por outro lado, a eficiência do ligante apresentada por estas novas substâncias foi mantida. Isto significa que os átomos preservados no esqueleto podem contribuir efetivamente no processo de reconhecimento molecular.

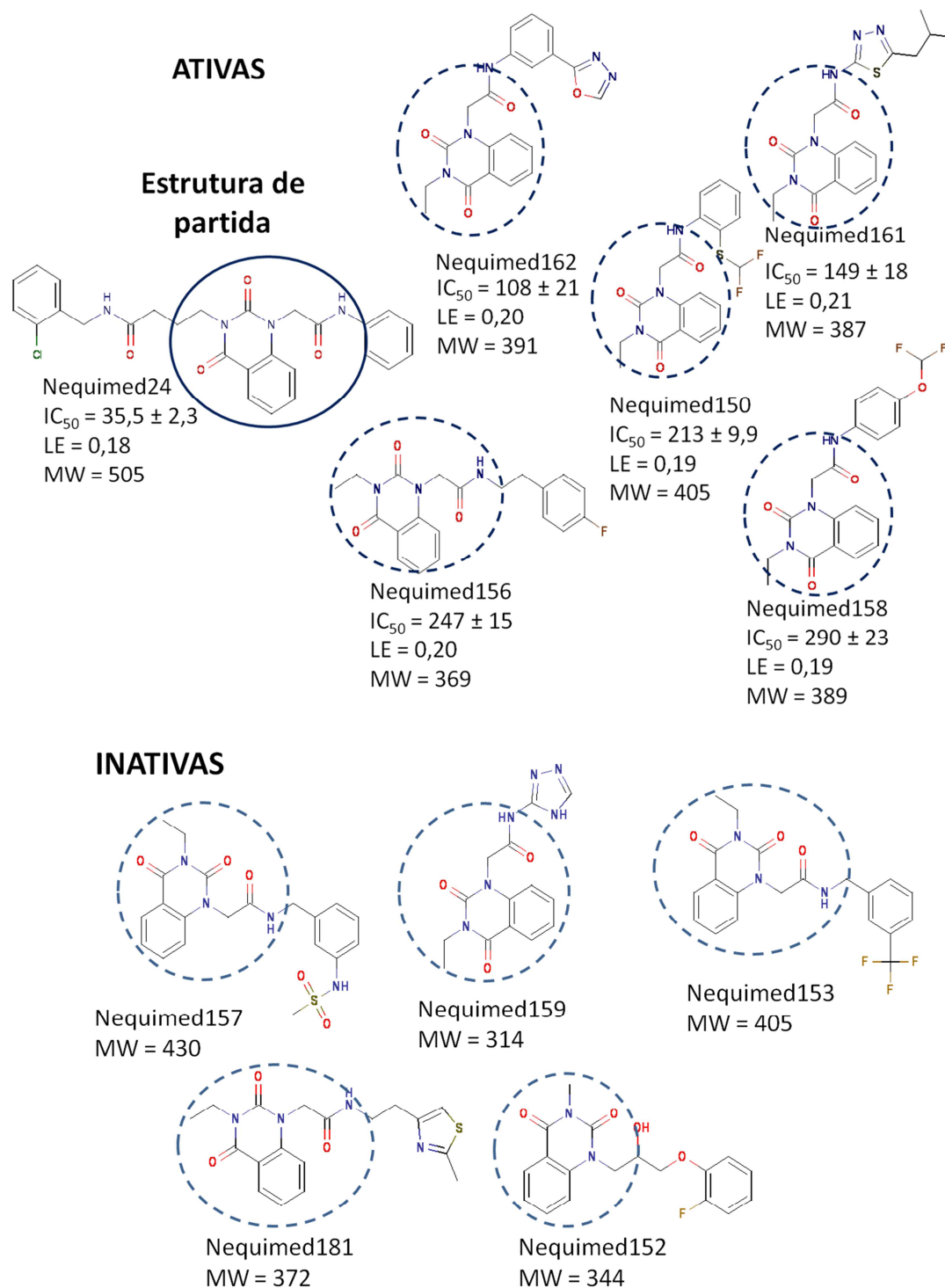


FIGURA 4.17: Estruturas das moléculas análogas ao Nequimed24 testadas contra a enzima cruzaina, as unidades são IC_{50} em μM , LE em $kcal\ mol^{-1}\ NA^{-1}$ e MW g mol^{-1}

Uma hipótese abordada é o anel dipirimidinônico ocupar o sub-sítio S1 da enzima, segundo o modo de interação predito por docagem molecular, FIGURA 4.18. Assumindo este modo de interação a amida ligada ao anel dipirimidinônico interage na porção central do sítio, onde há uma característica hidrofílica e o grupo onde houve as variações estruturais podem ocupar o sub-sítio S2 da proteína, onde

há o maior interesse por ser o sítio com maior chance de apresentar seletividade frente à enzima catepsina-L de humanos.

Os grupos selecionados para o estudo de SAR variaram a distância do anel subsequente a amida, a polaridade e tamanho dos grupos ligados a estes anéis. Existe um pequeno bolsão na proteína em uma região próxima ao sub-sítio S2 (próximo à Glu205) que foi identificada recentemente por métodos de dinâmica molecular [223], estes substituintes permitem sondar a orientação e tamanho dos grupos passíveis de serem acomodados nesta região da proteína.

A principal característica estrutural observada que diferencia as substâncias inativas das ativas foram a distância de ligações entre o nitrogênio da amida e o anel subsequente, com exceção da substância Nequimed159, todos os outros compostos possuem pelo menos um CH₂ alongando a cadeia. Uma hipótese é o alongamento causar impedimento estérico e não permitir a acomodação no sub-sítio S2.

A substância Nequimed159 difere dos demais compostos pela presença do triazol ligado ao nitrogênio da amida, conferindo uma maior polaridade ao grupo, o que pode ser detrimental para atividade nesta posição, uma vez que o sítio S2 tem característica hidrofóbica. O composto Nequimed152 não mantém a amida e é inativo, evidenciando que este grupo é fundamental para atividade.

Entre as substâncias ativas, a que apresentou melhor potência foi a Nequimed162, com o anel fenila ligado ao nitrogênio da amida e com o grupo oxadiazol na posição meta. A melhor eficiência do ligante foi Nequimed161 com o grupo tiadiazol ligado diretamente ao nitrogênio da amida. As substituições em posições orto e para do anel fenila, nos compostos Nequimed150 e Nequimed158 apresentaram potência e eficiência dos ligantes mais baixas, demonstrando assim, que a melhor posição para os substituintes aumentarem a potência é a meta. O modo de interação proposto por docagem molecular é apresentado na FIGURA 4.18.

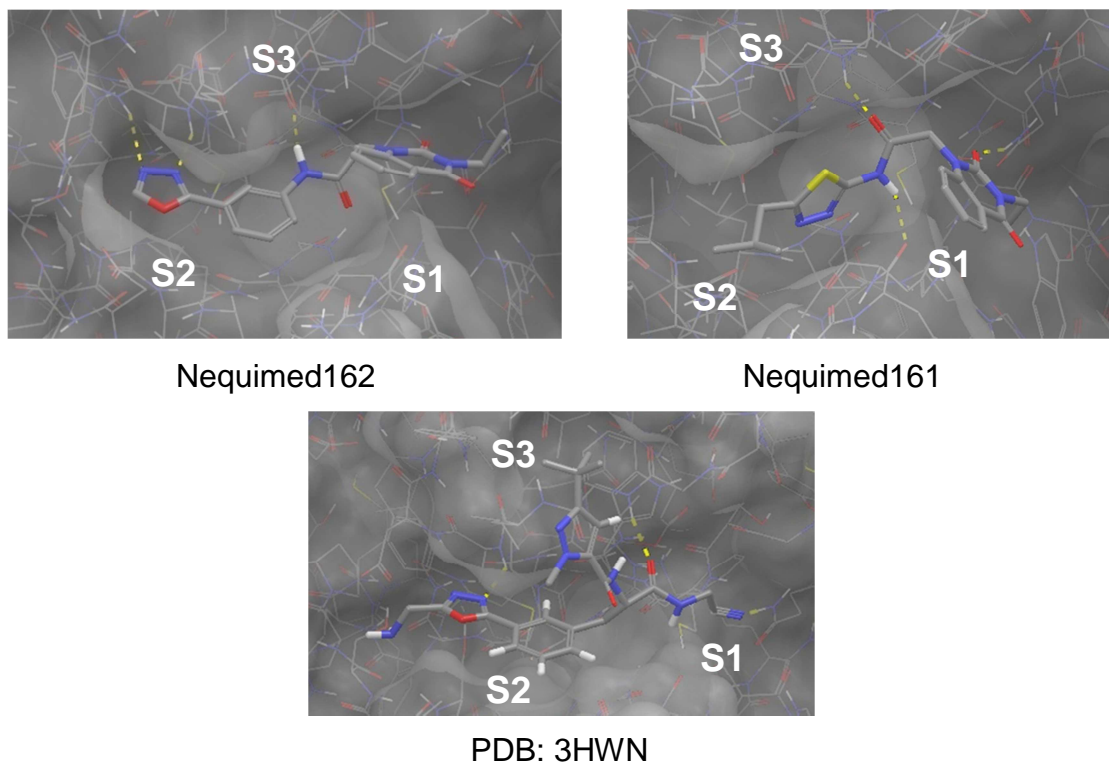


FIGURA 4.18: Estruturas dos compostos Nequimed162 e Nequimed161 docadas no sítio da cruzaína utilizando-se o programa Glide_XP. Estrutura cristalográfica demonstrando a posição de ligação do grupo oxadiazol na catepsina-L de Humanos [224]. Os subsítios da enzima estão enumerados como S1, S2 e S3

O substituinte dipirimidinona fica alocado no sub-sítio S1, onde está a tríade catalítica. O sub-sítio S2 é hidrofóbico e acomoda a fenila ou a isobutila, nas substâncias Nequimed162 e Nequimed161, respectivamente. Observando-se o complexo com a substância Nequimed162 fica claro que o oxadiazol na posição meta do anel fenila fica direcionado na posição do pequeno bolsão da extensão do sub-sítio S2, em uma pose semelhante aos inibidores da catepsina-L cristalizados com a enzima, o que demonstra uma possível razão para as menores potências das substituições em orto e para. O sub-sítio S3, mais superficial que os demais, são inexplorados por estas substâncias e oferece a oportunidade de inserção de substituintes nesta posição para o aumento de potência.

A outra série de substâncias adquiridas foi análoga a Nequimed42, a estratégia de simplificação molecular foi abordada devido às substâncias identificadas possuírem massa molecular próximo de 500 Da e potência na faixa de baixo micromolar, o aumento da potência é obtido com a inserção de novos grupos capazes de realizarem interações específicas, o que normalmente resulta em um crescimento molecular. Os aumentos da massa das substâncias identificadas neste projeto resultariam em moléculas que violariam as regras de Lipinski, diminuindo as

chances de um bom perfil farmacocinético. Desta forma, a identificação das porções moleculares conferem atividade para estas moléculas é essencial para iniciar a etapa de otimização de potência.

A complexidade molecular de Nequimed42 desta molécula gerou a proposição de dois esqueletos moleculares distintos, um com o grupo tiofeno com as duas amidas, e outro com o triazol substituído com nas posições 1 e 2 com a fenila e piperazina. Foram adquiridas oito substâncias a FIGURA 4.19 apresenta os valores de potência, eficiência do ligante e estruturas químicas.

Quando o grupo tiofeno com as duas amidas foi mantido, a variação da estrutura foi remover os grupos fenila e piperazina para avaliar sua contribuição para potência. Por outro lado, quando o grupo triazol foi mantido foram testadas três substâncias, uma com grupo fenila com características hidrofóbicas, uma nitrila que poderia ligar-se a cisteína catalítica covalentemente e uma sulfona cíclica com características hidrofílicas.

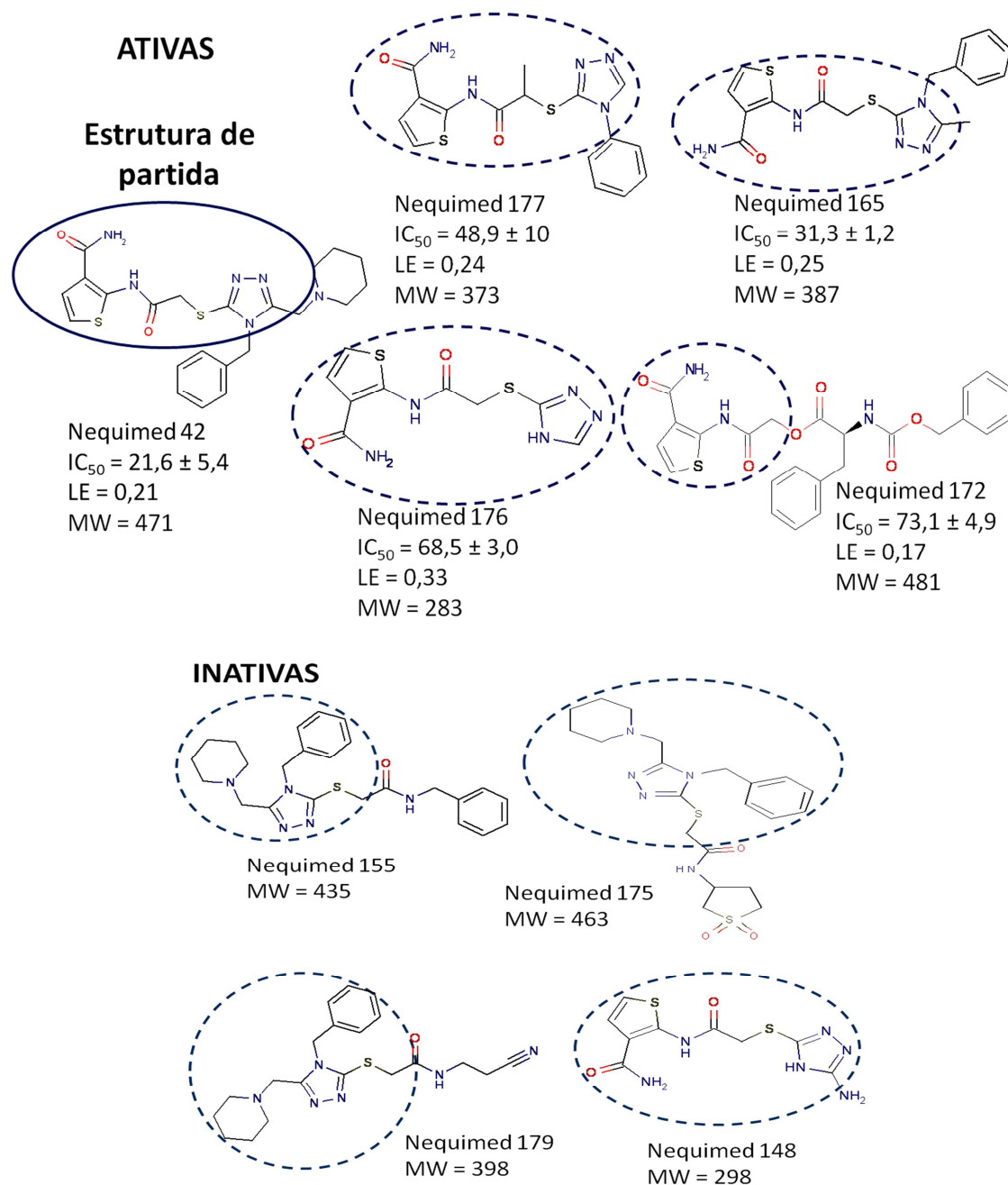


FIGURA 4.19: Estruturas das moléculas análogas ao Nequimed42 testadas contra a enzima cruzaina, as unidades são IC_{50} em μM , LE em $kcal\ mol^{-1}\ NA^{-1}$ e MW $g\ mol^{-1}$

Através desta série foi identificado que o grupo responsável pela atividade é o tiofeno com as amidas, pois foram as substâncias que apresentaram atividade, quando este grupo foi removido as substâncias foram inativas. Comparando-se as estruturas das moléculas Nequimed165 e Nequimed177 com a molécula de partida, observa-se que a potência mantém-se praticamente inalterada, entretanto a massa molecular das novas substâncias é menor, o que resulta em um aumento da eficiência do ligante e a demonstração que o grupo piperazina não contribui para a

atividade. A remoção da fenila e piperazina (Nequimed176) resultou em uma pequena diminuição da atividade, entretanto a eficiência do ligante aumentou significativamente, alcançando níveis de substância matriz ($LE > 0,30$). A substância Nequimed172 é uma hibridação entre Nequimed42 e Nequimed38, apesar de inibir a enzima esta substância apresentou potência e LE menores que os respectivos precursores. Uma proposta para o modo de interação destas substâncias é apresentada na FIGURA 4.20.

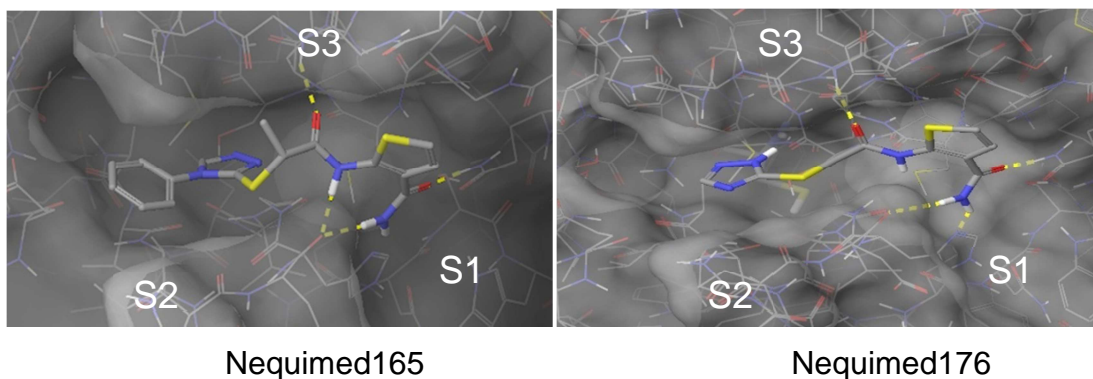


FIGURA 4.20: Estruturas dos compostos Nequimed165 e Nequimed176 docadas no sítio da cruzaina utilizando-se o programa Glide_XP. Os sub-sítios da enzima estão enumerados como S1, S2 e S3

O grupo amido-tiofeno é posicionado no sub-sítio S1, e possui boa complementaridade com a tríade catalítica. Uma característica extremamente importante a ressaltar nestas substâncias é a inibição ser reversível e não covalente, uma vez que não há grupos capazes de reagir covalentemente com a cisteína catalítica. A validação destes ligantes, juntamente com a classe inovadora para esta enzima qualificam estas substâncias como promissoras para o desenvolvimento de novas substâncias, que podem ser obtidas através da otimização da potência destas substâncias explorando o sub-sítio S3.

4.11 Ensaio *in vitro* contra *T. cruzi* na forma Tripomastigota

As seis substâncias mais potentes encontradas frente a enzima cruzaina foram submetidas ao ensaio *in vitro* contra a forma tripomastigota do parasito *T. cruzi* em sangue infectado de camundongo. Este ensaio foi realizado na Escola de Medicina da USP de Ribeirão Preto, sob supervisão do professor doutor João Santana. O resultado é apresentado na FIGURA 4.21.

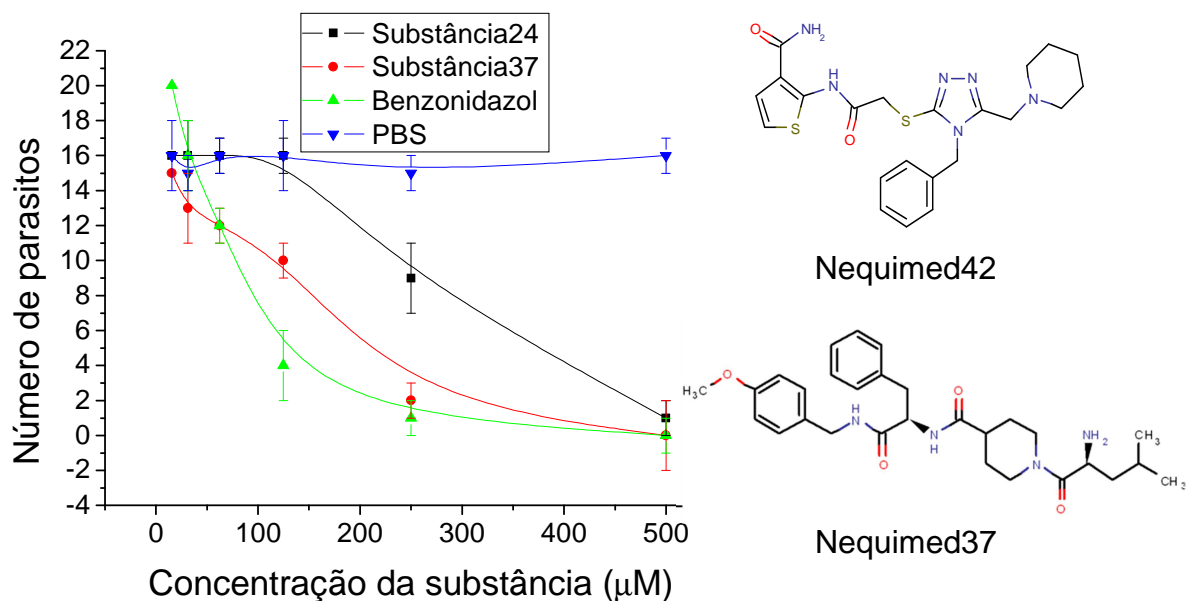


FIGURA 4.21: Contagem do número de parasitos em sangue infectado de camundongo em diferentes concentrações das substâncias avaliadas que apresentaram atividade

Este ensaio semi-quantitativo permite a triagem das substâncias a serem avaliadas posteriormente nos ensaios de toxicidade e quantitativos contra o *T. cruzi*. O controle negativo utilizado nos ensaios foi PBS com DMSO e controle positivo foi o fármaco comercial Benzonidazol. Dentre as seis substâncias testadas, as substâncias Nequimed42 e Nequimed37 apresentaram atividade tripanossomicida em níveis comparáveis ao benzonidazol. As duas substâncias pertencem a classes moleculares inéditas de inibidores da cruzaina, sendo que a classe tiofeno diamida (Nequimed42) foi validada como inibidor da enzima através do estudo de relação estrutura atividade.

Estes resultados provêm uma ótima validação das estratégias abordadas neste projeto que integram tecnologias, iniciando-se com ensaios virtuais aliados com biologia molecular e ensaios bioquímicos para a descoberta de novas substâncias bioativas promissoras no desenvolvimento de novos agentes quimioterápicos para o tratamento da doença de Chagas.

CAPÍTULO 5 - CONCLUSÕES

A busca de novos inibidores da enzima cruzaina foi realizada utilizando métodos virtuais baseados na estrutura do ligante e do receptor, foram construídos modelos de QSAR, busca por similaridade química utilizando descritores moleculares 3D e docagem molecular com distintos programas que contém algoritmos de busca conformacional e funções de pontuação diferentes.

A integração dos distintos métodos *in silico* foi abordada através do desenvolvimento de uma estratégia de busca por moléculas em banco de dados que iniciou no emprego de filtros moleculares no banco de moléculas comerciais ZINC com cerca de 8,5 milhões de estruturas, utilização de programas de docagem de alto desempenho que permitiu a avaliação de cerca de 3,5 milhões de moléculas com concomitante busca por similaridade química para gerar uma sub-coleção focada em moléculas com características apropriadas para o reconhecimento molecular pela enzima cruzaina que contém 8 635 estruturas.

A integração dos métodos *in silico*, juntamente com a inspeção visual das moléculas melhores listadas pela pontuação consensual permitiu a seleção de 23 moléculas pertencentes a diferentes classes químicas para serem avaliados nos ensaios bioquímicos. Inicialmente as substâncias foram testadas para a determinação de IC₅₀ via fluorimetria e identificaram-se 12 substâncias ativas com potência entre 5,6 e 73,9 µM para cruzaina e 8,6 e 89,1 µM para a catepsina-L, validando a estratégia de busca e os modelos virtuais estabelecidos previamente. Os compostos ativos possuem os valores calculados das propriedades de Lipinski satisfatórios, o que os capacitam a ter boas propriedades farmacocinéticas.

O mecanismo e a constante de inibição das substâncias ativas foram determinados e obteve-se K_i^{app} entre 3,7 e 64,5 µM para a cruzaina e 3,8 e 87,1 µM para a catepsina-L. O mecanismo de inibição observado para todos os compostos foi o reversível competitivo e substâncias pertencentes a novas classes de inibidores com afinidade baixo-micromolar foram identificadas: Nequimed24, Nequimed25, Nequimed30 e Nequimed42. Os processos integrados de descoberta de novos ligantes ativos contra a cruzaina levaram à identificação de inibidores em concentrações micromolar sem que houvesse clara identificação de seletividade frente à catepsina-L.

A obtenção do alvo biomacromolecular para estudos *in vitro* é essencial para a identificação de substâncias bioativas, desta forma foi implementado o protocolo de expressão e purificação da enzima cruzaina no grupo de Química Medicinal NEQUIMED-PN. O sistema de expressão foi modificado, o DNA plasmidial contendo o gene da cruzaina foi retirado da bactéria *E. coli* DH5 α e inserido em *E. coli* BL21-DE3, para a etapa de renovelamento da proteína foi utilizado o método de diálise em substituição ao método da diluição, os parâmetros cinéticos da enzima foram determinados e corroboraram com os valores descritos na literatura confirmando o sucesso na obtenção da enzima ativa.

Uma etapa fundamental na descoberta de uma substância bioativa é a validação do ligante, realizada através de teste de compostos análogos com o objetivo de estabelecer uma relação entre a estrutura química e a atividade biológica. Foram testados análogos de duas substâncias identificadas previamente por métodos computacionais e com o mecanismo e afinidade pela enzima cruzaina elucidados. As duas substâncias de partida foram simplificadas e tiveram suas massas moleculares diminuídas, os análogos ativos análogos à Nequimed24 apresentaram coeficientes de eficiência do ligante semelhantes à substância de partida e os análogos à Nequimed42 apresentaram um aumento da eficiência do ligante, sendo que um deles (Nequimed176) atingiu o valor 0,33 kcal mol⁻¹ NA⁻¹ superior ao limite inferior de uma substância matriz. O estudo de relações estrutura atividade, proveu informações a respeito de grupos potenciais para interação na cavidade localizada na extensão do sub-sítio S2 identificado recentemente, o grupo oxadiazol na posição meta da fenila apresentou um aumento da potência através de uma possível interação nesta extensão do sub-sítio S2.

A avaliação dos seis inibidores da enzima cruzaina identificados com maior afinidade contra o *T. cruzi* cepa Y resultou na descoberta de duas substâncias tripanossomicidas com potência comparável ao fármaco utilizado na terapêutica, o Benzonidazol[®], com alto valor agregado de informações. Isto porque foi proposto um projeto com alvo biológico promissor, descobertos os inibidores deste alvo, demonstrado a atividade nos ensaios bioquímicos cinéticos enzimáticos e ensaios contra a forma infectiva do parasito. A integração das tecnologias em química medicinal permitiu uma significativa contribuição através da identificação de substâncias promissoras para a gênese de um novo fármaco contra a doença de Chagas.

PERSPECTIVAS

O desenvolvimento de uma estratégia integrada de métodos de ensaios virtuais baseados na estrutura do receptor e do ligante resultou na descoberta de uma nova classe molecular com atividade frente à enzima cruzaina e ao parasito *T. cruzi*. Desta forma, a utilização dos modelos computacionais desenvolvidos e validados será útil nas etapas de otimização molecular, através da retro-alimentação dos modelos; propondo novas estruturas, aplicando nos modelos, realizando os testes *in vitro* e adicionando as novas informações para torná-los mais abrangentes e preditivos, principalmente no desenvolvimento de novos modelos de relações quantitativas entre as estruturas e a atividade biológica.

Em relação à nova classe molecular, os análogos do composto Nequimed42 ativos contra a enzima serão submetidos aos ensaios *in vitro* contra o parasito e também de toxidez em linhas celulares de mamíferos. Uma vez que os compostos se mostrem seguros, deverão ser submetidos à otimização da potência contra a enzima cruzaina. Para isto, o modo de interação deverá ser confirmado por um método experimental como cristalografia de proteínas para resolução da estrutura tridimensional. De acordo com o modo de interação predito por docagem molecular as substâncias ocupam os sub-sítios S1 e S2, a inserção de substituintes capazes de ocuparem os sub-sítios S3 e S4 são a oportunidade de explorar novas relações estrutura atividade desta classe molecular.

Por fim, as informações dos testes de atividade contra o parasito é uma oportunidade do desenvolvimento de modelos para a predição da atividade tripanossomicida de novas substâncias. Baseado na estrutura dos inibidores enzimáticos ativos contra o parasito é possível a construção de um banco de dados com as informações físico-químicas e estruturais das substâncias e também da atividade biológica. Alguns exemplos de modelos a serem desenvolvidos são métodos quimiométricos, para discriminar entre ativos e inativos, métodos de similaridade química e relações estrutura atividade.

REFERÊNCIAS

- [1] COURA, J. R.; VÍAS, P. A. "Chagas disease: A new worldwide challenge". *Nature* **465**: S6; 2010.
- [2] SCHOFIELD, C. J.; GORLA, D. E. "Who, how, what and where?". *Nature* **465**: S8; 2010.
- [3] OMS. Disponível em: <http://www.who.int/mediacentre/factsheets/fs340/en/> Acesso em: 26/03/2011
- [4] CLAYTON, J. "Chagas disease 101". *Nature* **465**: S4; 2010.
- [5] PESSOA, S. B. M., A. V. . "Parasitologia médica". *Guanabara Koogan* **11**: 872.
- [6] TYLER, K. M.; ENGMAN, D. M. "The life cycle of *Trypanosoma cruzi* revisited". *International Journal for Parasitology* **31**: 472; 2001.
- [7] OMS. Disponível em: <http://apps.who.int/tdr/svc/diseases/chagas> Acesso em: 25/03/2011
- [8] CHAGAS, C. R. J. D. "Tripanosomiase americana: forma aguda da molestia". *Mem. Inst. Oswaldo Cruz* **8**: 23; 1916.
- [9] PRATA, A. "Clinical and epidemiological aspects of Chagas disease". *The Lancet Infectious Diseases* **1**: 92; 2001.
- [10] MONCAYO, A.; YANINE, M. I. O. "An update on Chagas disease (human American trypanosomiasis)". *Ann. Trop. Med. Parasitol.* **100**: 663; 2006.
- [11] DE OLIVEIRA FERREIRA, H. "Treatment of the indeterminate form of Chagas' disease with nifurtimox and benznidazole". *Revista da Sociedade Brasileira de Medicina Tropical* **23**: 209; 1991.
- [12] AÑEZ, N.; CRISANTE, G.; CARABALLO, F.; DELGADO, W.; PARADA, H. "*Trypanosoma cruzi* persistence at oral inflammatory foci in chronic chagasic patients". *Acta Tropica* **117**: 207; 2011.
- [13] DIAS, J. C. P.; SILVEIRA, A. C.; SCHOFIELD, C. J. "The impact of Chagas disease control in Latin America - A review". *Memorias do Instituto Oswaldo Cruz* **97**: 603; 2002.
- [14] REZENDE FILHO, J. "Diagnosis of gastrointestinal forms of Chagas disease". *Diagnóstico das manifestações digestivas da doença de Chagas* **9**: 22; 2007.
- [15] PINK, R.; HUDSON, A.; MOURIES, M. A.; BENDIG, M. "Opportunities and challenges in antiparasitic drug discovery". *Nature Reviews Drug Discovery* **4**: 727; 2005.
- [16] COURA, J. R. "Present situation and new strategies for Chagas disease chemotherapy - a proposal". *Memorias do Instituto Oswaldo Cruz* **104**: 549; 2009.
- [17] URBINA, J. A.; DOCAMPO, R. "Specific chemotherapy of Chagas disease: Controversies and advances". *Trends in Parasitology* **19**: 495; 2003.
- [18] MORENO, S. N. J.; DOCAMPO, R.; MASON, R. P. "Different behaviors of benznidazole as free radical generator with mammalian and *Trypanosoma cruzi* microsomal preparations". *Archives of Biochemistry and Biophysics* **218**: 585; 1982.
- [19] DIAS, L. C.; DESSOY, M. A.; SILVA, J. J. N.; THIEMANN, O. H.; OLIVA, G.; ANDRICOPULO, A. D. "Chemotherapy of Chagas' disease: State of the art and perspectives for the development of new drugs". *Quimioterapia da doença de Chagas: Estado da arte e perspectivas no desenvolvimento de novos fármacos* **32**: 2444; 2009.
- [20] VANDEN BOSSCHE, H. "Biochemical targets for antifungal azole derivatives: hypothesis on the mode of action". *Current topics in medical mycology* **1**: 313; 1985.
- [21] DIXON, H.; GINGER, C. D.; WILLIAMSON, J. "Trypanosome sterols and their metabolic origins". *Comparative Biochemistry and Physiology -- Part B: Biochemistry and* **41**: 1; 1972.
- [22] BUCKNER, F.; YOKOYAMA, K.; LOCKMAN, J.; AIKENHEAD, K.; OHKANDA, J.; SADILEK, M.; SEBTI, S.; VAN VOORHIS, W.; HAMILTON, A.; GELB, M. H. "A class of sterol 14-demethylase inhibitors as anti-*Trypanosoma cruzi* agents". *Proceedings of the National Academy of Sciences of the United States of America* **100**: 15149; 2003.
- [23] CLAYTON, J. "Chagas disease: Pushing through the pipeline". *Nature* **465**: S12; 2010.

- [24] FIELD, H.; BLENCHE, I.; CROFT, S.; FIELD, M. C. "Characterisation of protein isoprenylation in procyclic form *Trypanosoma brucei*". *Molecular and Biochemical Parasitology* **82**: 67; 1996.
- [25] YOKOYAMA, K.; TROBRIDGE, P.; BUCKNER, F. S.; SCHOLTEN, J.; STUART, K. D.; VAN VOORHIS, W. C.; GELB, M. H. "The effects of protein farnesyltransferase inhibitors on trypanosomatids: Inhibition of protein farnesylation and cell growth". *Molecular and Biochemical Parasitology* **94**: 87; 1998.
- [26] YOKOYAMA, K.; TROBRIDGE, P.; BUCKNER, F. S.; VAN VOORHIS, W. C.; STUART, K. D.; GELB, M. H. "Protein farnesyltransferase from *Trypanosoma brucei*. A heterodimer of 61- and 65-kDa subunits as a new target for antiparasite therapeutics". *Journal of Biological Chemistry* **273**: 26497; 1998.
- [27] CASEY, P. J.; SEABRA, M. C. "Protein prenyltransferases". *Journal of Biological Chemistry* **271**: 5289; 1996.
- [28] EASTMAN, R. T.; BUCKNER, F. S.; YOKOYAMA, K.; GELB, M. H.; VAN VOORHIS, W. C. "Fighting parasitic disease by blocking protein farnesylation". *Journal of Lipid Research* **47**: 233; 2006.
- [29] FISHER, J. E.; ROGERS, M. J.; HALASY, J. M.; LUCKMAN, S. P.; HUGHES, D. E.; MASARACHIA, P. J.; WESOLOWSKI, G.; RUSSELL, R. G. G.; RODAN, G. A.; RESZKA, A. A. "Alendronate mechanism of action: Geranylgeraniol, an intermediate in the mevalonate pathway, prevents inhibition of osteoclast formation, bone resorption, and kinase activation in vitro". *Proceedings of the National Academy of Sciences of the United States of America* **96**: 133; 1999.
- [30] RODAN, G. A. Mechanisms of action of bisphosphonates. *volume 38*; 1998: 375.
- [31] URBINA, J. A.; MORENO, B.; VIERKOTTER, S.; OLDFIELD, E.; PAYARES, G.; SANOJA, C.; BAILEY, B. N.; YAN, W.; SCOTT, D. A.; MORENO, S. N. J.; DOCAMPO, R. "*Trypanosoma cruzi* contains major pyrophosphate stores, and its growth in vitro and in vivo is blocked by pyrophosphate analogs". *Journal of Biological Chemistry* **274**: 33609; 1999.
- [32] DOCAMPO, R.; MORENO, S. N. J. "Acidocalcisome: A novel Ca²⁺ storage compartment in Trypanosomatids and apicomplexan parasites". *Parasitology Today* **15**: 443; 1999.
- [33] MORENO, S. N. J.; DOCAMPO, R. "The role of acidocalcisomes in parasitic protists". *Journal of Eukaryotic Microbiology* **56**: 208; 2009.
- [34] ENGEL, J. C.; FRANKE DE CAZZULO, B. M.; STOPPANI, A. O. M.; CANNATA, J. J. B.; CAZZULO, J. J. "Aerobic glucose fermentation by *Trypanosoma cruzi* axenic culture amastigote-like forms during growth and differentiation to epimastigotes". *Molecular and Biochemical Parasitology* **26**: 1; 1987.
- [35] FREITAS, R. F.; PROKOPCZYK, I. M.; ZOTTIS, A.; OLIVA, G.; ANDRICOPULO, A. D.; TREVISAN, M. T. S.; VILEGAS, W.; SILVA, M. G. V.; MONTANARI, C. A. "Discovery of novel *Trypanosoma cruzi* glyceraldehyde-3-phosphate dehydrogenase inhibitors". *Bioorganic and Medicinal Chemistry* **17**: 2476; 2009.
- [36] DE MACEDO, E. M. S.; WIGGERS, H. J.; SILVA, M. G. V.; BRAZ-FILHO, R.; ANDRICOPULO, A. D.; MONTANARI, C. A. "A new bianthron glycoside as inhibitor of *Trypanosoma cruzi* glyceraldehyde 3-phosphate dehydrogenase activity". *Journal of the Brazilian Chemical Society* **20**: 947; 2009.
- [37] EAKIN, A. E.; GUERRA, A.; FOCIA, P. J.; TORRES-MARTINEZ, J.; CRAIG III, S. P. "Hypoxanthine phosphoribosyltransferase from *Trypanosoma cruzi* as a target for structure-based inhibitor design: Crystallization and inhibition studies with purine analogs". *Antimicrobial Agents and Chemotherapy* **41**: 1686; 1997.
- [38] AUGUSTYNS, K.; AMSSOMS, K.; YAMANI, A.; RAJAN, P. K.; HAEMERS, A. "Trypanothione as a target in the design of antitrypanosomal and antileishmanial agents". *Current Pharmaceutical Design* **7**: 1117; 2001.
- [39] HOLLOWAY, G. A.; CHARMAN, W. N.; FAIRLAMB, A. H.; BRUN, R.; KAISER, M.; KOSTEWICZ, E.; NOVELLO, P. M.; PARISOT, J. P.; RICHARDSON, J.; STREET, I. P.; WATSON, K. G.; BAELL, J. B. "Trypanothione reductase high-throughput screening

- campaign identifies novel classes of inhibitors with antiparasitic activity". *Antimicrobial Agents and Chemotherapy* **53**: 2824; 2009.
- [40] HORVATH, D. "A virtual screening approach applied to the search for trypanothione reductase inhibitors". *Journal of Medicinal Chemistry* **40**: 2412; 1997.
- [41] RIVAROLA, H. W.; PAGLINI-OLIVA, P. A. "Trypanosoma cruzi trypanothione reductase inhibitors: Phenothiazines and related compounds modify experimental Chagas' disease evolution". *Current Drug Targets - Cardiovascular and Haematological Disorders* **2**: 43; 2002.
- [42] FRAGOSO, S. P.; GOLDENBERG, S. "Cloning and characterization of the gene encoding *Trypanosoma cruzi* DNA topoisomerase II". *Molecular and Biochemical Parasitology* **55**: 127; 1992.
- [43] GOLDENBERG, S.; FRAGOSO, S. P.; GONZALES-PERDOMO, M. "Trypanosoma cruzi topoisomerase II: A target enzyme for the chemotherapy of Chagas' disease? Cloning and characterization of its encoding gene". *Biological Research* **26**: 77; 1993.
- [44] GONZALES-PERDOMO, M.; DE CASTRO, S. L.; MEIRELLES, M. N. S. L.; GOLDENBERG, S. "Trypanosoma cruzi proliferation and differentiation are blocked by topoisomerase II inhibitors". *Antimicrobial Agents and Chemotherapy* **34**: 1707; 1990.
- [45] CHRISTOPHERSON, R. I.; LYONS, S. D.; WILSON, P. K. "Inhibitors of *de novo* nucleotide biosynthesis as drugs". *Accounts of Chemical Research* **35**: 961; 2002.
- [46] ARAKAKI, T. L.; BUCKNER, F. S.; GILLESPIE, J. R.; MALMQUIST, N. A.; PHILLIPS, M. A.; KALYUZHNIY, O.; LUFT, J. R.; DETITTA, G. T.; VERLINDE, C. L. M. J.; VAN VOORHIS, W. C.; HOL, W. G. J.; MERRITT, E. A. "Characterization of *Trypanosoma brucei* dihydroorotate dehydrogenase as a possible drug target; structural, kinetic and RNAi studies". *Molecular Microbiology* **68**: 37; 2008.
- [47] URBINA, J. A. "Chemotherapy of Chagas disease". *Current Pharmaceutical Design* **8**: 287; 2002.
- [48] URBINA, J. A. "New chemotherapeutic approaches for the treatment of Chagas disease (American Trypanosomiasis)". *Expert Opinion on Therapeutic Patents* **13**: 661; 2003.
- [49] MCKERROW, J. H. "Parasite proteases". *Experimental Parasitology* **68**: 111; 1989.
- [50] BONTEMPI, E.; DECAZZULO, B. M.; RUIZ, A. M.; CAZZULO, J. J. "Purification and some properties of an acidic protease from epimastigotes of *Trypanosoma cruzi*". *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology* **77**: 599; 1984.
- [51] BONTEMPI, E.; MARTINEZ, J.; CAZZULO, J. J. "Subcellular localization of a cysteine proteinase from *Trypanosoma cruzi*". *Molecular and Biochemical Parasitology* **33**: 43; 1989.
- [52] MURTA, A. C. M.; PERSECHINI, P. M.; PADRON, T. D.; DESOUZA, W.; GUIMARAES, J. A.; SCHARFSTEIN, J. "Structural and functional identification of GP57/51 antigen of *Trypanosoma cruzi* as a cysteine proteinase". *Molecular and Biochemical Parasitology* **43**: 27; 1990.
- [53] RANGEL, H. A.; ARAUJO, P. M. F.; CAMARGO, I. J. B.; BONFITTO, M.; REPKA, D.; SAKURADA, J. K.; ATTA, A. M. "Detection of a proteinase common to epimastigote, trypomastigote and amastigote of different strains of *Trypanosoma cruzi*". *Tropenmedizin Und Parasitologie* **32**: 87; 1981.
- [54] RANGEL, H. A.; ARAUJO, P. M. F.; REPKA, D.; COSTA, M. G. "Trypanosoma cruzi isolation and characterization of a proteinase". *Experimental Parasitology* **52**: 199; 1981.
- [55] CAZZULO, J. J.; COUSO, R.; RAIMONDI, A.; WERNSTEDT, C.; HELLMAN, U. "Further characterization and partial amino-acid sequence of a cysteine proteinase from *Trypanosoma cruzi*". *Molecular and Biochemical Parasitology* **33**: 33; 1989.
- [56] SCHARFSTEIN, J.; SCHECHTER, M.; SENNA, M.; PERALTA, J. M.; MENDONCAPREVIATO, L.; MILES, M. A. "Trypanosoma cruzi characterization and isolation of a 57/51,000 MW surface glycoprotein (GP57/51) expressed by epimastigotes and blood stream trypomastigotes". *Journal of Immunology* **137**: 1336; 1986.

- [57] EAKIN, A. E.; MILLS, A. A.; HARTH, G.; MCKERROW, J. H.; CRAIK, C. S. "The sequence, organization, and expression of the major cysteine protease (cruzain) from *Trypanosoma cruzi*". *Journal of Biological Chemistry* **267**: 7411; 1992.
- [58] SALISBURY, C. M.; ELLMAN, J. A. "Rapid identification of potent nonpeptidic serine protease inhibitors". *ChemBioChem* **7**: 1034; 2006.
- [59] LECAILLE, F.; KALETA, J.; BROMME, D. "Human and parasitic papain-like cysteine proteases: Their role in physiology and pathology and recent developments in inhibitor design". *Chemical Reviews* **102**: 4459; 2002.
- [60] POWERS, J. C.; ASGIAN, J. L.; EKICI, O. D.; JAMES, K. E. "Irreversible inhibitors of serine, cysteine, and threonine proteases". *Chemical Reviews* **102**: 4639; 2002.
- [61] LEUNG, D.; ABBENANTE, G.; FAIRLIE, D. P. "Protease inhibitors: Current status and future prospects". *J. Med. Chem.* **43**: 305; 2000.
- [62] EAKIN, A. E.; MCGRATH, M. E.; MCKERROW, J. H.; FLETTERICK, R. J.; CRAIK, C. S. "Production of crystallizable cruzain, the major cysteine protease from *Trypanosoma cruzi*". *Journal of Biological Chemistry* **268**: 6115; 1993.
- [63] MCGRATH, M. E.; EAKIN, A. E.; ENGEL, J. C.; MCKERROW, J. H.; CRAIK, C. S.; FLETTERICK, R. J. "The crystal structure of cruzain - a therapeutic target for Chagas disease". *Journal of Molecular Biology* **247**: 251; 1995.
- [64] DARDONVILLE, C. "Recent advances in antitrypanosomal chemotherapy: patent literature 2002-2004". *Expert Opinion on Therapeutic Patents* **15**: 1241; 2005.
- [65] CHOWDHURY, S. F.; SIVARAMAN, J.; WANG, J.; DEVANATHAN, G.; LACHANCE, P.; QI, H. T.; MENARD, R.; LEFEBVRE, J.; KONISHI, Y.; CYGLER, M.; SULEA, T.; PURISIMA, E. O. "Design of noncovalent inhibitors of human cathepsin L. From the 96-residue proregion to optimized tripeptides". *Journal of Medicinal Chemistry* **45**: 5321; 2002.
- [66] SCHECHTE, I.; BERGER, A. "On size of active site in proteases .I. papain". *Biochemical and Biophysical Research Communications* **27**: 157; 1967.
- [67] LI, R. S.; CHEN, X. W.; GONG, B. Q.; SELZER, P. M.; LI, Z.; DAVIDSON, E.; KURZBAN, G.; MILLER, R. E.; NUZUM, E. O.; MCKERROW, J. H.; FLETTERICK, R. J.; GILLMOR, S. A.; CRAIK, C. S.; KUNTZ, I. D.; COHEN, F. E.; KENYON, G. L. "Structure-based design of parasitic protease inhibitors". *Bioorganic & Medicinal Chemistry* **4**: 1421; 1996.
- [68] GILLMOR, S. A.; CRAIK, C. S.; FLETTERICK, R. J. "Structural determinants of specificity in the cysteine protease cruzain". *Protein Science* **6**: 1603; 1997.
- [69] VERMELHO, A. B.; DE SIMONE, S. G.; D'AVILA-LEVY, C. M.; DO SANTOS, A. L. S.; DE MELO, A. C. N.; SILVA JR, F. P.; DA SILVA BON, E. P.; BRANQUINHA, M. H. "Trypanosomatidae peptidases: A target for drugs development". *Current Enzyme Inhibition* **3**: 19; 2007.
- [70] KLEMPA, M.; GOLDBERG, D. E. "Biological roles of proteases in parasitic protozoa". *Annual Review of Biochemistry* **71**: 275; 2002.
- [71] STOKA, V.; TURK, B.; MCKERROW, J. H.; BJORK, I.; CAZZULO, J. J.; TURK, V. "The high stability of cruzipain against pH-induced inactivation is not dependent on its C-terminal domain". *Febs Letters* **469**: 29; 2000.
- [72] ENGEL, J. C.; DOYLE, P. S.; HSIEH, I.; MCKERROW, J. H. "Cysteine protease inhibitors cure an experimental *Trypanosoma cruzi* infection". *Journal of Experimental Medicine* **188**: 725; 1998.
- [73] CAFFREY, C. R.; SCORY, S.; STEVERDING, D. "Cysteine Proteinases of Trypanosome Parasites: Novel Targets for Chemotherapy". *Current Drug Targets* **1**: 155; 2000.
- [74] SOARES, M. J.; SOUTO-PADRON, T.; DE SOUZA, W. "Identification of a large pre-lysosomal compartment in the pathogenic protozoon *Trypanosoma cruzi*". *Journal of Cell Science* **102**: 157; 1992.
- [75] ENGEL, J. C.; DOYLE, P. S.; PALMER, J.; HSLEH, I.; BAINTON, D. F.; MCKERROW, J. H. "Cysteine protease inhibitors alter Golgi complex ultrastructure and function in *Trypanosoma cruzi*". *Journal of Cell Science* **111**: 597; 1998.

- [76] JACOBSEN, W.; CHRISTIANS, U.; BENET, L. Z. "In vitro evaluation of the disposition of a novel cysteine protease inhibitor". *Drug Metabolism and Disposition* **28**: 1343; 2000.
- [77] BARR, S. C.; WARNER, K. L.; KORNREIC, B. G.; PISCITELLI, J.; WOLFE, A.; BENET, L.; MCKERROW, J. H. "A cysteine protease inhibitor protects dogs from cardiac damage during infection by *Trypanosoma cruzi*". *Antimicrobial Agents and Chemotherapy* **49**: 5160; 2005.
- [78] DOYLE, P. S.; ZHOU, Y. M.; ENGEL, J. C.; MCKERROW, J. H. "A cysteine protease inhibitor cures Chagas' disease in an immunodeficient-mouse model of infection". *Antimicrobial Agents and Chemotherapy* **51**: 3932; 2007.
- [79] KERR, I. D.; LEE, J. H.; FARADY, C. J.; MARION, R.; RICKERT, M.; SAJID, M.; PANDEY, K. C.; CAFFREY, C. R.; LEGAC, J.; HANSELL, E.; MCKERROW, J. H.; CRAIK, C. S.; ROSENTHAL, P. J.; BRINEN, L. S. "Vinyl sulfones as antiparasitic agents and a structural basis for drug design". *Journal of Biological Chemistry* **284**: 25697; 2009.
- [80] CHEN, Y. T.; BRINEN, L. S.; KERR, I. D.; HANSELL, E.; DOYLE, P. S.; MCKERROW, J. H.; ROUSH, W. R. "In vitro and in vivo studies of the trypanocidal properties of WRR-483 against *Trypanosoma cruzi*". *PLoS Neglected Tropical Diseases* **4**; 2010.
- [81] 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* **7**: 1143; 2001.
- [82] ROUSH, W. R.; GONZÁLEZ, F. V.; MCKERROW, J. H.; HANSELL, E. "Design and synthesis of dipeptidyl α,β -epoxy ketones, potent irreversible inhibitors of the cysteine protease cruzain". *Bioorganic and Medicinal Chemistry Letters* **8**: 2809; 1998.
- [83] GONZÁLEZ, F. V.; IZQUIERDO, J.; RODRÍGUEZ, S.; MCKERROW, J. H.; HANSELL, E. "Dipeptidyl- α,β -epoxyesters as potent irreversible inhibitors of the cysteine proteases cruzain and rhodesain". *Bioorganic and Medicinal Chemistry Letters* **17**: 6697; 2007.
- [84] WILLIAMS, G. H. "Converting-enzyme inhibitors in the treatment of hypertension". *New England Journal of Medicine* **319**: 1517; 1988.
- [85] LAM, P. Y. S.; JADHAV, P. K.; EYERMANN, C. J.; HODGE, C. N.; RU, Y.; BACHELER, L. T.; MEEK, J. L.; OTTO, M. J.; RAYNER, M. M.; WONG, Y. N.; CHANG, C. H.; WEBER, P. C.; JACKSON, D. A.; SHARPE, T. R.; ERICKSON-VIITANEN, S. "Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors". *Science* **263**: 380; 1994.
- [86] DU, X.; HANSELL, E.; ENGEL, J. C.; CAFFREY, C. R.; COHEN, F. E.; MCKERROW, J. H. "Aryl ureas represent a new class of anti-trypanosomal agents". *Chemistry and Biology* **7**: 733; 2000.
- [87] HUANG, L.; BRINEN, L. S.; ELLMAN, J. A. "Crystal structures of reversible ketone-Based inhibitors of the cysteine protease cruzain". *Bioorganic and Medicinal Chemistry* **11**: 21; 2003.
- [88] CHOE, Y.; BRINEN, L. S.; PRICE, M. S.; ENGEL, J. C.; LANGE, M.; GRISOSTOMI, C.; WESTON, S. G.; PALLAI, P. V.; CHENG, H.; HARDY, L. W.; HARTSOUGH, D. S.; MCKIN, M.; TILTON, R. F.; BALDINO, C. M.; CRAIK, C. S. "Development of α -keto-based inhibitors of cruzain, a cysteine protease implicated in Chagas disease". *Bioorganic and Medicinal Chemistry* **13**: 2141; 2005.
- [89] CHIYANZU, I.; HANSELL, E.; GUT, J.; ROSENTHAL, P. J.; MCKERROW, J. H.; CHIBALE, K. "Synthesis and evaluation of isatins and thiosemicarbazone derivatives against cruzain, falcipain-2 and rhodesain". *Bioorganic and Medicinal Chemistry Letters* **13**: 3527; 2003.
- [90] GREENBAUM, D. C.; MACKAY, Z.; HANSELL, E.; DOYLE, P.; GUT, J.; CAFFREY, C. R.; LEHRMAN, J.; ROSENTHAL, P. J.; MCKERROW, J. H.; CHIBALE, K. "Synthesis and structure-activity relationships of parasiticidal thiosemicarbazone cysteine protease inhibitors against *Plasmodium falciparum*, *Trypanosoma brucei*, and *Trypanosoma cruzi*". *Journal of Medicinal Chemistry* **47**: 3212; 2004.

- [91] FUJII, N.; MALLARI, J. P.; HANSELL, E. J.; MACKEY, Z.; DOYLE, P.; ZHOU, Y. M.; GUT, J.; ROSENTHAL, P. J.; MCKERROW, J. H.; GUY, R. K. "Discovery of potent thiosemicarbazone inhibitors of rhodesain and cruzain". *Bioorganic and Medicinal Chemistry Letters* **15**: 121; 2005.
- [92] SILES, R.; CHEN, S. E.; ZHOU, M.; PINNEY, K. G.; TRAWICK, M. L. "Design, synthesis, and biochemical evaluation of novel cruzain inhibitors with potential application in the treatment of Chagas' disease". *Bioorganic and Medicinal Chemistry Letters* **16**: 4405; 2006.
- [93] BRYANT, C.; KERR, I. D.; DEBNATH, M.; ANG, K. K. H.; RATNAM, J.; FERREIRA, R. S.; JAISHANKAR, P.; ZHAO, D.; ARKIN, M. R.; MCKERROW, J. H.; BRINEN, L. S.; RENSLO, A. R. "Novel non-peptidic vinylsulfones targeting the S2 and S3 subsites of parasite cysteine proteases". *Bioorganic and Medicinal Chemistry Letters* **19**: 6218; 2009.
- [94] BORCHHARDT, D. M.; MASCARELLO, A.; CHIARADIA, L. D.; NUNES, R. J.; OLIVA, G.; YUNES, R. A.; ANDRICOPULO, A. D. "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* **21**: 142; 2010.
- [95] HERNANDES, M. Z.; RABELLO, M. M.; LEITE, A. C. L.; CARDOSO, M. V. O.; MOREIRA, D. R. M.; BRONDANI, D. J.; SIMONE, C. A.; REIS, L. C.; SOUZA, M. A.; PEREIRA, V. R. A.; FERREIRA, R. S.; MCKERROW, J. H. "Studies toward the structural optimization of novel thiazolylhydrazone- based potent antitrypanosomal agents". *Bioorganic and Medicinal Chemistry* **18**: 7826; 2010.
- [96] TROSSINI, G. H. G.; MALVEZZI, A.; AMARAL, A. T. D.; RANGEL-YAGUI, C. O.; IZIDORO, M. A.; CEZARI, M. H. S.; JULIANO, L.; CHIN, C. M.; MENEZES, C. M. S.; FERREIRA, E. I. "Cruzain inhibition by hydroxymethylnitrofurazone and nitrofurazone: Investigation of a new target in *Trypanosoma cruzi*". *Journal of Enzyme Inhibition and Medicinal Chemistry* **25**: 62; 2010.
- [97] BRAK, K.; DOYLE, P. S.; MCKERROW, J. H.; ELLMAN, J. A. "Identification of a new class of nonpeptidic inhibitors of cruzain". *Journal of the American Chemical Society* **130**: 6404; 2008.
- [98] BRAK, K.; KERR, I. D.; BARRETT, K. T.; FUCHI, N.; DEBNATH, M.; ANG, K.; ENGEL, J. C.; MCKERROW, J. H.; DOYLE, P. S.; BRINEN, L. S.; ELLMAN, J. A. "Nonpeptidic tetrafluorophenoxymethyl ketone cruzain inhibitors as promising new leads for chagas disease chemotherapy". *Journal of Medicinal Chemistry* **53**: 1763; 2010.
- [99] WERMUTH, C. G.; GANELLIN, C. R.; LINDBERG, P.; MITSCHER, L. A. Glossary of terms used in medicinal chemistry. *volume 33: Pure Applied Chemistry*; 1998: 385.
- [100] MONTANARI, C. A.; PILLI, R. A. "Journal of the Brazilian Chemical Society: Editorial". *Journal of the Brazilian Chemical Society* **13**: 713; 2002.
- [101] WERMUTH, C. G. "Medicinal chemistry: definition and objectives, the three main phases of drug activity, drug and disease classifications". *The practice of medicinal chemistry*: 29; 2003.
- [102] LOMBARDINO, J. G.; LOWE III, J. A. "The role of the medicinal chemist in drug discovery - Then and now". *Nature Reviews Drug Discovery* **3**: 853; 2004.
- [103] ZHU, F.; HAN, B.; KUMAR, P.; LIU, X.; MA, X.; WEI, X.; HUANG, L.; GUO, Y.; HAN, L.; ZHENG, C.; CHEN, Y. "Update of TTD: Therapeutic Target Database". *Nucleic Acids Research* **38**: D787; 2009.
- [104] HOPKINS, A. L.; GROOM, C. R. "The druggable genome". *Nature Reviews Drug Discovery* **1**: 727; 2002.
- [105] TAMIMI, N. A. M.; ELLIS, P. "Drug development: From concept to marketing!". *Nephron - Clinical Practice* **113**: c125; 2009.
- [106] WESS, G.; URMANN, M.; SICKENBERGER, B. "Medicinal chemistry: Challenges and opportunities". *Angewandte Chemie - International Edition* **40**: 3341; 2001.
- [107] BLEICHER, K. H.; BÖHM, H. J.; MÜLLER, K.; ALANINE, A. I. "Hit and lead generation: Beyond high-throughput screening". *Nature Reviews Drug Discovery* **2**: 369; 2003.

- [108] KESERU, G. M.; MAKARA, G. M. "Hit discovery and hit-to-lead approaches". *Drug Discovery Today* **11**: 741; 2006.
- [109] LIPINSKI, C.; HOPKINS, A. "Navigating chemical space for biology and medicine". *Nature* **432**: 855; 2004.
- [110] VAN DEURSEN, R.; REYMOND, J. L. "Chemical space travel". *ChemMedChem* **2**: 636; 2007.
- [111] BROWN, F. K. Chemoinformatics: What is it and How does it Impact Drug Discovery., *volume 33*. Annual Reports in Medicinal Chemistry; 1998: 375.
- [112] GASTEIGER, J. "The central role of chemoinformatics". *Chemometrics and Intelligent Laboratory Systems* **82**: 200; 2006.
- [113] OPREA, T. I. *Chemoinformatics and the Quest for Leads in Drug Discovery*. Wiley-VCH Verlag GmbH; 2008. p.
- [114] IRWIN, J. J.; SHOICHET, B. K. "ZINC - A free database of commercially available compounds for virtual screening". *Journal of Chemical Information and Modeling* **45**: 177; 2005.
- [115] BAURIN, N.; BAKER, R.; RICHARDSON, C.; CHEN, I.; FOLOPPE, N.; POTTER, A.; JORDAN, A.; ROUGHLEY, S.; PARRATT, M.; GREANEY, P.; MORLEY, D.; HUBBARD, R. E. "Drug-like annotation and duplicate analysis of a 23-supplier chemical database totalling 2.7 million compounds". *Journal of Chemical Information and Computer Sciences* **44**: 643; 2004.
- [116] MUEGGE, I. "Selection criteria for drug-like compounds". *Medicinal Research Reviews* **23**: 302; 2003.
- [117] LIPINSKI, C. A.; LOMBARDO, F.; DOMINY, B. W.; FEENEY, P. J. "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings". *Advanced Drug Delivery Reviews* **23**: 3; 1997.
- [118] LIPINSKI, C. A. "Lead- and drug-like compounds: The rule-of-five revolution". *Drug Discovery Today: Technologies* **1**: 337; 2004.
- [119] OPREA, T. I.; ALLU, T. K.; FARA, D. C.; RAD, R. F.; OSTOPOVICI, L.; BOLOGA, C. G. "Lead-like, drug-like or "Pub-like": How different are they?". *Journal of Computer-Aided Molecular Design* **21**: 113; 2007.
- [120] BROOIJMANS, N.; KUNTZ, I. D., *volume 32*; 2003: 335.
- [121] ROSINI, M.; ANDRISANO, V.; BARTOLINI, M.; BOLOGNESI, M. L.; HRELIA, P.; MINARINI, A.; TAROZZI, A.; MELCHIORRE, C. "Rational approach to discover multipotent anti-Alzheimer drugs". *Journal of Medicinal Chemistry* **48**: 360; 2005.
- [122] TANIUCHI, H.; PARR, G. R.; JUILLERAT, M. A. "Complementation in folding and fragment exchange". *Methods in Enzymology* **Vol. 131**: 185; 1986.
- [123] BISSANTZ, C.; FOLKERS, G.; ROGNAN, D. "Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations". *Journal of Medicinal Chemistry* **43**: 4759; 2000.
- [124] LEACH, A. R.; SHOICHET, B. K.; PEISHOFF, C. E. "Prediction of protein-ligand interactions. Docking and scoring: Successes and gaps". *Journal of Medicinal Chemistry* **49**: 5851; 2006.
- [125] KITCHEN, D. B.; DECORNEZ, H.; FURR, J. R.; BAJORATH, J. "Docking and scoring in virtual screening for drug discovery: Methods and applications". *Nature Reviews Drug Discovery* **3**: 935; 2004.
- [126] SHOICHET, B. K. "Virtual screening of chemical libraries". *Nature* **432**: 862; 2004.
- [127] EWING, T. J. A.; MAKINO, S.; SKILLMAN, A. G.; KUNTZ, I. D. "DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases". *Journal of Computer-Aided Molecular Design* **15**: 411; 2001.
- [128] MOITESSIER, N.; ENGLEBIENNE, P.; LEE, D.; LAWANDI, J.; CORBEIL, C. R. "Towards the development of universal, fast and highly accurate docking/scoring methods: A long way to go". *British Journal of Pharmacology* **153**: S7; 2008.
- [129] HUANG, S. Y.; GRINTER, S. Z.; ZOU, X. "Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions". *Physical Chemistry Chemical Physics* **12**: 12899; 2010.

- [130] PENICHEVA, T.; SOUMANA, O. S.; PAJEVA, I.; MITEVA, M. A. "Post-docking virtual screening of diverse binding pockets: Comparative study using DOCK, AMMOS, X-Score and FRED scoring functions". *European Journal of Medicinal Chemistry* **45**: 2622; 2010.
- [131] BÖHM, H. J. "On the use of LUDI to search the Fine Chemicals Directory for ligands of proteins of known three-dimensional structure". *Journal of Computer-Aided Molecular Design* **8**: 623; 1994.
- [132] ROGNAN, D.; LAUEMOLLER, S. L.; HOLM, A.; BUUS, S.; TSCHINKE, V. "Predicting binding affinities of protein ligands from three-dimensional models: Application to peptide binding to class I major histocompatibility proteins". *Journal of Medicinal Chemistry* **42**: 4650; 1999.
- [133] GOHLKE, H.; HENDLICH, M.; KLEBE, G. "Predicting binding modes, binding affinities and 'hot spots' for protein-ligand complexes using a knowledge-based scoring function". *Perspectives in Drug Discovery and Design* **20**: 115; 2000.
- [134] GOHLKE, H.; HENDLICH, M.; KLEBE, G. "Knowledge-based scoring function to predict protein-ligand interactions". *Journal of Molecular Biology* **295**: 337; 2000.
- [135] WANG, R.; LAI, L.; WANG, S. "Further development and validation of empirical scoring functions for structure-based binding affinity prediction". *Journal of Computer-Aided Molecular Design* **16**: 11; 2002.
- [136] SCHULZ-GASCH, T.; STAHL, M. "Binding site characteristics in structure-based virtual screening: Evaluation of current docking tools". *Journal of Molecular Modeling* **9**: 47; 2003.
- [137] SHERIDAN, R. P. "Chemical similarity searches: When is complexity justified?". *Expert Opinion on Drug Discovery* **2**: 423; 2007.
- [138] MOFFAT, K.; GILLET, V. J.; WHITTLE, M.; BRAVI, G.; LEACH, A. R. "A comparison of field-based similarity searching methods: CatShape, FBSS, and ROCS". *Journal of Chemical Information and Modeling* **48**: 719; 2008.
- [139] FERREIRA, M. M. C.; MONTANARI, C. A.; GAUDIO, A. C. "Variable selection in QSAR". *Seleção de variáveis em QSAR* **25**: 439; 2002.
- [140] TAVARES, L. C. "QSAR: The Hansch's approach". *QSAR: A abordagem de Hansch* **27**: 631; 2004.
- [141] GOLBRAIKH, A.; SHEN, M.; XIAO, Z.; XIAO, Y. D.; LEE, K. H.; TROPSHA, A. "Rational selection of training and test sets for the development of validated QSAR models". *Journal of Computer-Aided Molecular Design* **17**: 241; 2003.
- [142] GOLBRAIKH, A.; TROPSHA, A. "Beware of q²!". *Journal of Molecular Graphics and Modelling* **20**: 269; 2002.
- [143] FILTER. version 2.0.2. Santa Fe: Openeye Scientific Software Inc; 2007.
- [144] SYBYL. version 8.0. Saint Louis: Tripos; 2007.
- [145] MALVEZZI, A.; DE REZENDE, L.; AMARAL, A. D. "Pharmacophore Model of Cruzain Inhibitors". *QSAR & Combinatorial Science* **28**: 781; 2009.
- [146] WOOD, W. J. L.; HUANG, L.; ELLMAN, J. A. "Synthesis of a Diverse Library of Mechanism-Based Cysteine Protease Inhibitors". *Journal of Combinatorial Chemistry* **5**: 869; 2003.
- [147] BOSTRÖM, J.; GREENWOOD, J. R.; GOTTFRIES, J. "Assessing the performance of OMEGA with respect to retrieving bioactive conformations". *Journal of Molecular Graphics and Modelling* **21**: 449; 2003.
- [148] KIRCHMAIR, J.; RISTIC, S.; EDER, K.; MARKT, P.; WOLBER, G.; LAGGNER, C.; LANGER, T. "Fast and efficient in silico 3D screening: Toward maximum computational efficiency of pharmacophore-based and shape-based approaches". *Journal of Chemical Information and Modeling* **47**: 2182; 2007.
- [149] FAWCETT, T. "An introduction to ROC analysis". *Pattern Recognition Letters* **27**: 861; 2006.
- [150] PDB. Disponível em: <http://www.pdb.org> Acesso em: 10/12/2007
- [151] DAVIS, A. M.; TEAGUE, S. J.; KLEYWEGT, G. J. "Application and limitations of x-ray crystallographic data in structure-based ligand and drug design". *Angewandte Chemie - International Edition* **42**: 2718; 2003.

- [152] BRUNGER, A. T. "Free R value: A novel statistical quantity for assessing the accuracy of crystal structures". *Nature* **355**: 472; 1992.
- [153] RAREY, M.; KRAMER, B.; LENGAUER, T.; KLEBE, G. "A Fast Flexible Docking Method using an Incremental Construction Algorithm". *Journal of Molecular Biology* **261**: 470; 1996.
- [154] KELLENBERGER, E.; RODRIGO, J.; MULLER, P.; ROGNAN, D. "Comparative evaluation of eight docking tools for docking and virtual screening accuracy". *Proteins: Structure, Function and Genetics* **57**: 225; 2004.
- [155] FAWCETT, T. "An introduction to ROC analysis". *Pat. Recog. Let.* **27**: 861; 2006.
- [156] OPENEYE. FRED. *volume 2.2.5*. Santa Fe, NM; 2010.
- [157] MASEK, B. B.; MERCHANT, A.; MATTHEW, J. B. "Molecular shape comparison of angiotensin II receptor antagonists". *Journal of Medicinal Chemistry* **36**: 1230; 1993.
- [158] RUSH III, T. S.; GRANT, J. A.; MOSYAK, L.; NICHOLLS, A. "A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction". *Journal of Medicinal Chemistry* **48**: 1489; 2005.
- [159] SENGER, S. "Using tversky similarity searches for core hopping: Finding the needles in the haystack". *Journal of Chemical Information and Modeling* **49**: 1514; 2009.
- [160] ROCS. *version 2.4.1*. Santa Fe: Openeye Scientific Software Inc; 2007.
- [161] GRANT, J. A.; GALLARDO, M. A.; PICKUP, B. T. "A fast method of molecular shape comparison: A simple application of a Gaussian description of molecular shape". *J. Comput. Chem.* **17**: 1653; 1996.
- [162] MEEK, P. J.; LIU, Z. W.; TIAN, L. F.; WANG, C. Y.; WELSH, W. J.; ZAUHAR, R. J. "Shape signatures: speeding up computer aided drug discovery". *Drug Discov. Today* **11**: 895; 2006.
- [163] ANDRICOPULO, A. D.; MONTANARI, C. A. "Structure-activity relationships for the design of small-molecule inhibitors". *Mini-Reviews in Medicinal Chemistry* **5**: 585; 2005.
- [164] MYINT, K. Z.; XIE, X. Q. "Recent advances in fragment-based QSAR and multi-dimensional QSAR methods". *International Journal of Molecular Sciences* **11**: 3846; 2010.
- [165] HUANG, L.; ELLMAN, J. A. "General solid-phase method to prepare novel cyclic ketone inhibitors of the cysteine protease cruzain". *Bioorganic and Medicinal Chemistry Letters* **12**: 2993; 2002.
- [166] FREITAS, R. F.; OPREA, T. I.; MONTANARI, C. A. "2D QSAR and similarity studies on cruzain inhibitors aimed at improving selectivity over cathepsin L". *Bioorganic and Medicinal Chemistry* **16**: 838; 2008.
- [167] CHARIFSON, P. S.; CORKERY, J. J.; MURCKO, M. A.; WALTERS, W. P. "Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins". *J. Med. Chem.* **42**: 5100; 1999.
- [168] FEHER, M. "Consensus scoring for protein-ligand interactions". *Drug Discov. Today* **11**: 421; 2006.
- [169] ODA, A.; TSUCHIDA, K.; TAKAKURA, T.; YAMAOTSU, N.; HIRONO, S. Comparison of consensus scoring strategies for evaluating computational models of protein-ligand complexes. *4th Indo-US Workshop on Mathematical Chemistry, volume 46*. Pune, INDIA: Amer Chemical Soc; 2005: 380.
- [170] MITEVA, M. A.; LEE, W. H.; MONTES, M. O.; VILLOUTREIX, B. O. "Fast Structure-Based Virtual Ligand Screening Combining FRED, DOCK, and Surflex". *Journal of Medicinal Chemistry* **48**: 6012; 2005.
- [171] WANG, R.; WANG, S. "How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment". *Journal of Chemical Information and Computer Sciences* **41**: 1422; 2001.
- [172] BABER, J. C.; SHIRLEY, W. A.; GAO, Y.; FEHER, M. "The Use of Consensus Scoring in Ligand-Based Virtual Screening". *Journal of Chemical Information and Modeling* **46**: 277; 2005.
- [173] HUANG, L.; BRINEN, L. S.; ELLMAN, J. A. "Crystal structures of reversible ketone-based inhibitors of the cysteine protease cruzain". *Bioorg. Med. Chem.* **11**: 21; 2003.

- [174] LEHNINGER, A.; NELSON, D. L.; COX, M. M. *Princípios de Bioquímica*. São Paulo: Sarvier; 1995. 839 p.
- [175] STRYER, L.; BERG, J. M.; TYMOCZKO, J. L. *Bioquímica*. São Paulo: Reverte; 1995. p.
- [176] VOET, D.; J.G.VOET. *Fundamentos de Bioquímica*. São Paulo: Artmed; 2006. 1596 p.
- [177] VERLI, H.; BARREIRO, E. J. "A medicinal chemistry paradigm: Ligands and receptor flexibility". *Um paradigma da química medicinal: A flexibilidade dos ligantes e receptores* **28**: 95; 2005.
- [178] KOSHLAND, D. E. "Application of a theory of enzyme specificity to protein synthesis". *Proceedings of the National Academy of Sciences of the United States of America* **44**: 98; 1958.
- [179] NELSON, J. M.; VOSBURGH, W. C. "Kinetics of invertase action". *Journal of the American Chemical Society* **39**: 790; 1917.
- [180] SCHNELL, S.; MAINI, P. K. "Enzyme kinetics at high enzyme concentration". *Bulletin of Mathematical Biology* **62**: 483; 2000.
- [181] ELION, G. B. "The purine path to chemotherapy". *Science* **244**: 41; 1989.
- [182] ERNST, M. E.; FRAVEL, M. A. "Febuxostat: A selective xanthine-oxidase/xanthine-dehydrogenase inhibitor for the management of hyperuricemia in adults with gout". *Clinical Therapeutics* **31**: 2503; 2009.
- [183] EGNER, U.; KRÄTZSCHMAR, J.; KREFT, B.; POHLENZ, H. D.; SCHNEIDER, M. "The target discovery process". *ChemBioChem* **6**: 468; 2005.
- [184] LUDEWIG, S.; KOSSNER, M.; SCHILLER, M.; BAUMANN, K.; SCHIRMEISTER, T. "Enzyme kinetics and hit validation in fluorimetric protease assays". *Current Topics in Medicinal Chemistry* **10**: 368; 2010.
- [185] SCHNELL, S.; MENDOZA, C. "Time-dependent closed form solutions for fully competitive enzyme reactions". *Bulletin of Mathematical Biology* **62**: 321; 2000.
- [186] SEGEL, I. H. *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems* New York: Wiley-Interscience; 2006. 957 p.
- [187] HOH, C. Y.; CORD-RUWISCH, R. "A practical kinetic model that considers endproduct inhibition in anaerobic digestion processes by including the equilibrium constant". *Biotechnology and Bioengineering* **51**: 597; 1996.
- [188] COPELAND, R. A. *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*. New Jersey: Willey-Interscience; 2005. 265 p.
- [189] SHITARA, Y.; SUGIYAMA, Y. "Pharmacokinetic and pharmacodynamic alterations of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors: Drug-drug interactions and interindividual differences in transporter and metabolic enzyme functions". *Pharmacology and Therapeutics* **112**: 71; 2006.
- [190] YAO, C.; LEVY, R. H. "Inhibition-based metabolic drug-drug interactions: Predictions from in vitro data". *Journal of Pharmaceutical Sciences* **91**: 1923; 2002.
- [191] SCHMIDER, J.; VON MOLTKE, L. L.; SHADER, R. I.; HARMATZ, J. S.; GREENBLATT, D. J. "Extrapolating in vitro data on drug metabolism to in vivo pharmacokinetics: Evaluation of the pharmacokinetic interaction between amitriptyline and fluoxetine". *Drug Metabolism Reviews* **31**: 545; 1999.
- [192] WATSON, J. D.; MYERS, R. M.; WITKOWSKI, J. A. *Recombinant DNA: Genes and Genomes - A Short Course*. New York: Scientific American Books; 1992. 929 p.
- [193] LEWIN, B. *Genes VII*. New York: Oxford University Press; 2002. 990 p.
- [194] BANEYX, F. "Recombinant protein expression in Escherichia coli". *Current Opinion in Biotechnology* **10**: 411; 1999.
- [195] PILLAI, O.; PANCHAGNULA, R. "Insulin therapies - Past, present and future". *Drug Discovery Today* **6**: 1056; 2001.
- [196] SØRENSEN, H. P.; MORTENSEN, K. K. "Advanced genetic strategies for recombinant protein expression in Escherichia coli". *Journal of Biotechnology* **115**: 113; 2005.

- [197] MOREIRA, A. R. "The evolution of protein expression and cell culture". *BioPharm International* **20**: 56; 2007.
- [198] WILDT, S.; GERNGROSS, T. U. "The humanization of N-glycosylation pathways in yeast". *Nature Reviews Microbiology* **3**: 119; 2005.
- [199] JARVIS, D. L. "Developing baculovirus-insect cell expression systems for humanized recombinant glycoprotein production". *Virology* **310**: 1; 2003.
- [200] STRASSER, R.; ALTMANN, F.; MACH, L.; GLÖSSL, J.; STEINKELLNER, H. "Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking β 1,2-linked xylose and core α 1,3-linked fucose". *Febs Letters* **561**: 132; 2004.
- [201] SIMONS, J. P.; WILMUT, I.; CLARK, A. J.; ARCHIBALD, A. L.; BISHOP, J. C.; LATHE, R. "Gene transfer into sheep". *Bio/Technology* **6**: 179; 1988.
- [202] HARRIS, J. L.; BACKES, B. J.; LEONETTI, F.; MAHRUS, S.; ELLMAN, J. A.; CRAIK, C. S. "Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries". *Proceedings of the National Academy of Sciences of the United States of America* **97**: 7754; 2000.
- [203] SKOOG, D. A.; HOLLER, F. J.; NIEMAN, T. A. *Princípios de Análise Instrumental*. Porto Alegre: Bookman; 2002. 1536 p.
- [204] NOOMNARM, U.; CLEGG, R. M. "Fluorescence lifetimes: Fundamentals and interpretations". *Photosynthesis Research* **101**: 181; 2009.
- [205] ZIMMERMAN, M.; ASHE, B.; YUREWICZ, E. C.; PATEL, G. "Sensitive assays for trypsin, elastase, and chymotrypsin using new fluorogenic substrates". *Analytical Biochemistry* **78**: 47; 1977.
- [206] LIMA, A. P. C. A.; DOS REIS, F. C. G.; SERVEAU, C.; LALMANACH, G.; JULIANO, L.; MÉNARD, R.; VERNET, T.; THOMAS, D. Y.; STORER, A. C.; SCHARFSTEIN, J. "Cysteine protease isoforms from *Trypanosoma cruzi*, cruzipain 2 and cruzain, present different substrate preference and susceptibility to inhibitors". *Molecular and Biochemical Parasitology* **114**: 41; 2001.
- [207] ALVES, L. C.; MELO, R. L.; CEZARI, M. H. S.; SANDERSON, S. J.; MOTTRAM, J. C.; COOMBS, G. H.; JULIANO, L.; JULIANO, M. A. "Analysis of the S2 subsite specificities of the recombinant cysteine proteinases CPB of *Leishmania mexicana*, and cruzain of *Trypanosoma cruzi*, using fluorescent substrates containing non-natural basic amino acids". *Molecular and Biochemical Parasitology* **117**: 137; 2001.
- [208] BARICOS, W. H.; ZHOU, Y.; MASON, R. W.; BARRETT, A. J. "Human kidney cathepsins B and L. Characterization and potential role in degradation of glomerular basement membrane". *Biochemical Journal* **252**: 301; 1988.
- [209] MASON, R. W. "Species variants of cathepsin L and their immunological identification". *Biochemical Journal* **240**: 285; 1986.
- [210] FAIRHEAD, M.; VAN DER WALLE, C. F. "The heavy-light chain loop of human cathepsin-L modulates its activity and stability". *Protein and Peptide Letters* **15**: 47; 2008.
- [211] DEVICES, M. Disponível em: <http://www.moleculardevices.com> Acesso em: 05/10/2010
- [212] ORIGIN. *version 7.0*. USA: Microcal; 2002.
- [213] SIGMAPLOT. *version 10.0*. USA: SigmaPlot; 2007.
- [214] PHUE, J. N.; SANG, J. L.; TRINH, L.; SHILOACH, J. "Modified *Escherichia coli* B (BL21), a superior producer of plasmid DNA compared with *Escherichia coli* K (DH5 α)". *Biotechnology and Bioengineering* **101**: 831; 2008.
- [215] GRÄSLUND, S.; NORDLUND, P.; WEIGELT, J.; HALLBERG, B. M.; BRAY, J.; GILEADI, O.; KNAPP, S.; OPPERMAN, U.; ARROWSMITH, C.; HUI, R.; MING, J.; DHE-PAGANON, S.; PARK, H. W.; SAVCHENKO, A.; YEE, A.; EDWARDS, A.; VINCENTELLI, R.; CAMBILLAU, C.; KIM, R.; KIM, S. H.; RAO, Z.; SHI, Y.; TERWILLIGER, T. C.; KIM, C. Y.; HUNG, L. W.; WALDO, G. S.; PELEG, Y.; ALBECK, S.; UNGER, T.; DYM, O.; PRILUSKY, J.; SUSSMAN, J. L.; STEVENS, R. C.; LESLEY, S. A.; WILSON, I. A.; JOACHIMIAK, A.; COLLART, F.; DEMENTIEVA, I.; DONNELLY, M. I.; ESCHENFELDT, W. H.; KIM, Y.; STOLS, L.; WU, R.; ZHOU, M.; BURLEY, S. K.; EMTAGE, J. S.; SAUDER, J. M.; THOMPSON, D.; BAIN, K.; LUZ, J.; GHEYI, T.; ZHANG, F.; ATWELL, S.; ALMO, S. C.;

- BONANNO, J. B.; FISER, A.; SWAMINATHAN, S.; STUDIER, F. W.; CHANCE, M. R.; SALI, A.; ACTON, T. B.; XIAO, R.; ZHAO, L.; MA, L. C.; HUNT, J. F.; TONG, L.; CUNNINGHAM, K.; INOUE, M.; ANDERSON, S.; JANJUA, H.; SHASTRY, R.; HO, C. K.; WANG, D.; WANG, H.; JIANG, M.; MONTELIONE, G. T.; STUART, D. I.; OWENS, R. J.; DAENKE, S.; SCHÜTZ, A.; HEINEMANN, U.; YOKOYAMA, S.; BÜSSOW, K.; GUNSALUS, K. C. "Protein production and purification". *Nature Methods* **5**: 135; 2008.
- [216] COPELAND, R. A., ed. *Evaluation of enzyme inhibitors in drug discovery. A guide for medicinal chemists and pharmacologists*; 2005.
- [217] ZHANG, S.; ZHAO, H.; JOHN, R. "Development of a quantitative relationship between inhibition percentage and both incubation time and inhibitor concentration for inhibition biosensors - Theoretical and practical considerations". *Biosensors and Bioelectronics* **16**: 1119; 2001.
- [218] HOPKINS, A. L.; GROOM, C. R.; ALEX, A. "Ligand efficiency: A useful metric for lead selection". *Drug Discov. Today* **9**: 430; 2004.
- [219] LEESON, P. D.; SPRINGTHORPE, B. "The influence of drug-like concepts on decision-making in medicinal chemistry". *Nat. Rev. Drug Discov.* **6**: 881; 2007.
- [220] FERREIRA, R. S.; BRYANT, C.; ANG, K. K. H.; MCKERROW, J. H.; SHOICHET, B. K.; RENSLO, A. R. "Divergent modes of enzyme inhibition in a homologous structure-activity series". *Journal of Medicinal Chemistry* **52**: 5005; 2009.
- [221] MALVEZZI, A.; DE REZENDE, L.; DO AMARAL, A. T. "Pharmacopore model of cruzain inhibitors". *QSAR and Combinatorial Science* **28**: 781; 2009.
- [222] VEBER, D. F.; JOHNSON, S. R.; CHENG, H. Y.; SMITH, B. R.; WARD, K. W.; KOPPLE, K. D. "Molecular properties that influence the oral bioavailability of drug candidates". *Journal of Medicinal Chemistry* **45**: 2615; 2002.
- [223] DURRANT, J. D.; KERÄNEN, H.; WILSON, B. A.; MCCAMMON, J. A. "Computational identification of uncharacterized cruzain binding sites". *PLoS Neglected Tropical Diseases* **4**; 2010.
- [224] BETHEL, P. A.; GERHARDT, S.; JONES, E. V.; KENNY, P. W.; KAROUTCHI, G. I.; MORLEY, A. D.; OLDHAM, K.; RANKINE, N.; AUGUSTIN, M.; KRAPP, S.; SIMADER, H.; STEINBACHER, S. "Design of selective Cathepsin inhibitors". *Bioorganic and Medicinal Chemistry Letters* **19**: 4622; 2009.