



UNIVERSIDADE FEDERAL DE OURO PRETO
ESCOLA DE FARMÁCIA
DEPARTAMENTO DE FARMÁCIA



CAMILA VIANA DE SOUSA

**CONSTRUÇÃO E APLICAÇÃO DE UM PROTOCOLO DE TRIAGEM VIRTUAL NO
DESENVOLVIMENTO DE POTENCIAIS INIBIDORES DA CRUZAÍNA**

OURO PRETO - MINAS GERAIS - BRASIL

Agosto/2021



UNIVERSIDADE FEDERAL DE OURO PRETO
ESCOLA DE FARMÁCIA
DEPARTAMENTO DE FARMÁCIA



CAMILA VIANA DE SOUSA

**CONSTRUÇÃO E APLICAÇÃO DE UM PROTOCOLO DE TRIAGEM VIRTUAL NO
DESENVOLVIMENTO DE POTENCIAIS INIBIDORES DA CRUZAÍNA**

Trabalho de conclusão de curso apresentado
ao Curso de Farmácia da Universidade Federal de Ouro Preto,
como requisito parcial para aprovação na disciplina TCC 007.

Orientador: Saulo Fehelberg Pinto Braga.

OURO PRETO - MINAS GERAIS - BRASIL

Agosto/2021

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S725c Sousa, Camila Viana de .
Construção e aplicação de um protocolo de triagem virtual no desenvolvimento de potenciais inibidores da cruzaina . [manuscrito] / Camila Viana de Sousa. - 2021.
55 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Saulo Fehelberg Pinto Braga.
Monografia (Bacharelado). Universidade Federal de Ouro Preto.
Escola de Farmácia. Graduação em Farmácia .

1. Tripanossoma cruzi. 2. Chagas, Doença de. 3. Modelagem molecular. 4. Triagem virtual. I. Braga, Saulo Fehelberg Pinto . II. Universidade Federal de Ouro Preto. III. Título.

CDU 615.45

Bibliotecário(a) Responsável: Soraya Fernanda Ferreira e Souza - SIAPE: 1.763.787



FOLHA DE APROVAÇÃO

Camila Viana de Sousa

Construção e aplicação de um protocolo de triagem virtual no desenvolvimento de potenciais inibidores da cruzaina

Monografia apresentada ao Curso de Farmácia da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Farmácia.

Aprovada em 26 de agosto de 2021

Membros da banca

Prof. Dr. Saulo Fehelberg Pinto Braga - Orientador - Universidade Federal de Ouro Preto
Profa. Dra. Karina Taciana Santos Rúbio - Universidade Federal de Ouro Preto
Mestranda Rúbia Castro Fernandes Melo Reis - Universidade Federal de Ouro Preto

Prof. Dr. Saulo Fehelberg Pinto Braga, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 10/09/2021



Documento assinado eletronicamente por **Saulo Fehelberg Pinto Braga, PROFESSOR DE MAGISTERIO SUPERIOR**, em 10/09/2021, às 14:07, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0219471** e o código CRC **E1D3A4B3**.

AGRADECIMENTOS

Aos meus pais, Jorge e Cida por todo apoio, carinho e suporte. Eu amo vocês. Agradeço também as minhas irmãs, Renata e Gabriela pelo incentivo.

Ao orientador Saulo por ser um exemplo a ser seguido e à Universidade Federal De Ouro Preto pelo ensino público, gratuito e de qualidade. Aos professores da EFAR –UFOP por aceitarem o desafio de formar profissionais de saúde em meio à uma pandemia mundial.

A todos os meus familiares pelo amparo, em especial minha tia Heloísa por ser grande incentivadora. Dedico este trabalho à minha tia Helenice (*in memoriam*). Tenho certeza que a Senhora acompanhou todo este trabalho aí do céu.

E por fim, agradeço as eternas amigas que fiz na minha amada República Koxixo. Espero em breve poder comemorar com vocês o fim da pandemia e o início de um novo ciclo.

RESUMO

A doença de Chagas (DC), descoberta e descrita pelo pesquisador Carlos Chagas, é uma doença considerada negligenciada que acomete cerca de 6 a 7 milhões de pessoas na América Latina. Questões como desmatamento e desigualdade social afetam diretamente a epidemiologia desta doença, que além de ter diagnóstico difícil, tem tratamento baseado em apenas dois fármacos: benznidazol e nifurtimox. Ambos os fármacos apresentam baixa eficácia no tratamento da fase crônica da DC e a maioria dos pacientes apresentam efeitos adversos relacionados aos medicamentos, o que reduz a adesão ao tratamento. Neste contexto, faz-se necessário pesquisas com intuito de guiar o desenvolvimento de novos tratamentos para a DC. Das estratégias existentes para o planejamento de fármacos, destaca-se a triagem virtual que reduz os custos em pesquisas e acelera o processo de desenvolvimento de novas moléculas bioativas. O *docking* molecular é uma das estratégias de planejamento de novos fármacos e tem como intuito prever a interação de uma determinada molécula bioativa em um alvo molecular selecionado. A cruzaina, enzima envolvida em importantes processos de sobrevivência do parasita causador da DC, o *T. cruzi*, foi escolhida como alvo molecular desse trabalho. O presente trabalho teve como objetivo determinar o melhor protocolo de *docking* molecular para a busca de substâncias bioativas capazes de auxiliar no tratamento da DC. Para isso, questões como a seleção do alvo molecular e os parâmetros utilizados para o processo de *docking* foram analisados e discutidos. Após a seleção do melhor protocolo, este foi aplicado em um novo *set* de moléculas. Este novo set foi criado a partir do composto MB15, que já teve sua atividade tripanocida descrita anteriormente. Após o *docking*, as melhores moléculas passaram por uma inspeção visual com o intuito de determinar pontos importantes de interação entre alvo e ligantes.

Palavras chaves: Modelagem molecular, triagem virtual, *docking* molecular, cruzaina, doença de Chagas, *T. cruzi*.

LISTA DE ILUSTRAÇÕES

Figura 1 - Transmissão vetorial da doença de Chagas _____	12
Figura 2- Sinal de Romaña e chagoma de inoculação _____	13
Figura 3 - Estrutura nifurtimox (A) e benznidazol (B) _____	14
Figura 4 - Estrutura do MB15 _____	18
Figura 5 - Mecanismo de inibição irreversível da cruzaina por aceptor de Michael (vinilssufonas) _____	25
Figura 6 - Retrossíntese do MB15 _____	28
Figura 7 - Interações intermoleculares entre B95 e cruzaina _____	30
Figura 8 - Curvas ROC AUC dos protocolos aplicados _____	33
Figura 9 - Estrutura química e poses no sítio ativo da cruzaina encontradas pelo docking para os compostos selecionados (continua...) _____	37
Figura 10 - Subsítios da cruzaina _____	40
Figura 11- Interações entre resíduos de aminoácidos e o ligante 507 _____	41

LISTA DE ABREVIATURAS E SIGLAS

Ala	Alanina
Asp	Aspartato
CI ₅₀	Concentração necessária para se ter metade de um efeito
Cys	Cisteína
DC	Doença de Chagas
EF	Fator de enriquecimento (<i>enrichment factor</i>)
Gly	Glicina
His	Histidina
Leu	Leucina
Met	Metionina
ROC	<i>Receiver Operating Characteristic</i>
Trp	Triptofano

SUMÁRIO

1. INTRODUÇÃO	10
1.1 Doença De Chagas	10
1.1.1 Contexto epidemiológico	10
1.1.2 Transmissão	10
1.1.3 Vetores e ciclo do <i>T. cruzi</i>	11
1.1.4 Sintomas e diagnósticos	12
1.1.5 Tratamento atual	14
1.2 Desenvolvimento de novos fármacos para o tratamento da DC	15
1.2.1 Triagem virtual de compostos	15
1.2.2 A cruzaina como alvo molecular no tratamento da DC	17
1.2.3 Aplicação de um protocolo de triagem virtual na seleção de potenciais inibidores da cruzaina	17
2. OBJETIVOS	19
2.1 Objetivo Geral	19
2.2 Objetivos Específicos	19
3. METODOLOGIA	20
3.1. Obtenção e preparo da estrutura cristalográfica da cruzaina	20
3.2 Seleção de inibidores da cruzaina com atividade determinada experimentalmente	20
3.3 Obtenção dos <i>decoys</i>	21
3.4 Preparo dos ligantes	21
3.5 Criação de uma biblioteca virtual de compostos para triagem virtual	22
3.6 Desenvolvimento dos protocolos de ancoragem molecular	22
3.7 Normalização	23
3.8 Cálculo das métricas de validação	23
3.9 Virtual screening da biblioteca virtual e seleção dos compostos	23
4 RESULTADOS E DISCUSSÕES	24
4.1 Obtenção e preparo da estrutura cristalográfica da cruzaina	24
4.2 Seleção de inibidores de cruzaina de atividade determinada experimentalmente	24
4.3 Obtenção dos <i>decoys</i>	26
4.5 Criação de uma biblioteca virtual de compostos para triagem virtual	28
4.6 Desenvolvimento dos protocolos de ancoragem molecular	29
4.7 Normalização	31
4.8 Métricas e validação	31

4.9 Triagem virtual da biblioteca virtual e seleção dos compostos	36
5. CONCLUSÃO.....	42
REFERÊNCIAS BIBLIOGRÁFICAS.....	43
ANEXO A – Compostos utilizados para obtenção dos análogos do MB15	46

1. INTRODUÇÃO

1.1 Doença De Chagas

1.1.1 Contexto epidemiológico

Em Minas Gerais, no ano de 1909, o pesquisador, médico e sanitarista Carlos Chagas descreveu pela primeira vez a infecção parasitária que foi nomeada por ele de tripanossomíase americana e que subsequentemente ficou conhecida como doença de Chagas (DC). Tal descrição incluiu a identificação do *Trypanosoma cruzi*, agente etiológico da doença e um de seus vetores, o *Triatoma infestans* (COURA, 2003). Segundo dados da Organização Mundial de Saúde de 2019, estima-se que, na América Latina há cerca de 6 a 7 milhões de pessoas infectadas pelo *T. cruzi* e que outras 60 milhões de pessoas estão em situação de risco de infecção nos mais de 18 países endêmicos (WHO, 2019). Dentre esses países, o Brasil apresentou expressivo aumento nas últimas décadas, atingindo o número de cerca de 1,1 milhão de infectados (WHO, 2019). Aspectos sócio-econômicos e ambientais impactam no número de casos da doença de forma que residências localizadas em regiões interioranas e de famílias de baixa renda estão mais predispostas a alojar o inseto vetor. Já dentre os fatores ambientais podemos explicitar o desmatamento, que altera a disponibilidade de recursos de alimentação e proteção no habitat natural dos insetos vetores (DIAS, 2007). Desta maneira, estes insetos, na condição de hematófilos, migram para o ambiente doméstico e começam a se alimentar de humanos (DIAS, 2007). Apesar de acometer predominantemente as Américas, já foram diagnosticados casos de doença de Chagas em outros países não endêmicos como em alguns países da Europa, Japão e Austrália. Nestes países a disseminação da DC dá-se pela transmissão não vetorial (TZIZIK; BORCHARDT, 2018).

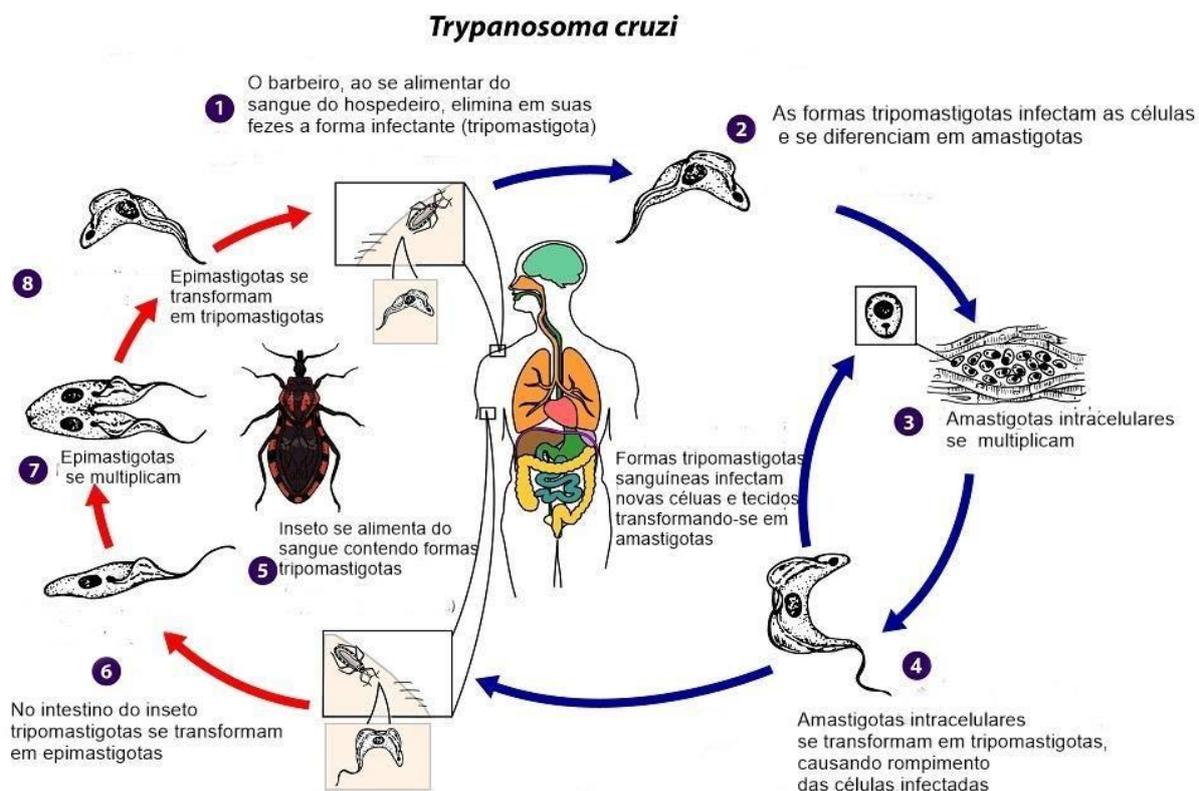
1.1.2 Transmissão

A transmissão da DC pode ocorrer de forma vetorial, vertical, oral, acidental e por meio de transfusões, sendo que somente a transmissão vetorial é intermediada por insetos. A transmissão vertical ocorre durante a gestação ou parto e é transmitida da gestante para o feto, principalmente quando a gestante está na fase aguda da DC. Também na fase aguda, a amamentação não é recomendada (LEAL; OLIVEIRA, 2009). A forma acidental de transmissão ocorre quando mucosas são expostas a materiais contaminados com as fezes de triatomíneos infectados pelo *T. cruzi* ou quando essas mucosas são expostas também de forma

acidental a sangue e material contaminado com este parasita. Na forma transfusional ocorre quando há transfusão de sangue ou transplante de órgãos de um portador da DC para um indivíduo saudável. Já a transmissão oral pode ocorrer de maneiras distintas que incluem a ingestão de alimentos contaminados com as fezes dos triatomíneos à ingestão de carne mal cozida de mamíferos contaminados com *T. cruzi* (ORGANIZAÇÃO PAN AMERICANA DE SAÚDE, 2009)

1.1.3 Vetores e ciclo do *T. cruzi*

A transmissão vetorial da DC é intermediada por insetos da família Reduviidae e são conhecidos como triatomíneos. Tais insetos se alimentam de qualquer tipo de sangue e para isso, são dotados de substâncias anestésicas e com potencial anticoagulante em sua saliva. Cinco das mais de quarenta e quatro espécies de triatomíneos identificadas no Brasil são consideradas domésticas e com importância epidemiológica primária. São elas: *Triatoma infestans*, *Panstrongylus megistus*, *T. brasiliensis* e *T. pseudomaculata* e *T. sordida*. Os vetores da DC, popularmente identificados como barbeiros, são infectados pela forma tripomastigota do *T. cruzi*, presente nas células sanguíneas dos hospedeiros vertebrados. A infecção ocorre quando estes insetos se alimentam de animais já infectados (TARTAROTTI, 2004). Com isso, os protozoários passam a habitar o intestino dos insetos e se diferenciam em epimastigotas. Já na porção final do sistema digestivo dos triatomíneos, as formas epimastigotas sofrem alterações e dão origem aos tripomastigotas, forma infectante eliminada nas fezes dos insetos. O ciclo se reinicia quando as formas infectantes tripomastigotas infectam um novo hospedeiro vertebrado e, após cerca de 30 horas, se diferenciam em amastigotas, a forma replicativa nas células de hospedeiros vertebrados (ECHEVERRIA; MORILLO, 2019). Os triatomíneos expelem *T. cruzi* em suas fezes e estes defecam concomitantemente com o momento da alimentação. Assim, ao se alimentar do sangue de um novo mamífero, o barbeiro evacua e as fezes contaminadas entram em contato com as mucosas deste novo animal, podendo atingir a circulação e infectando-o (Figura 1).

Figura 1: Transmissão vetorial da doença de Chagas

As setas vermelhas indicam o ciclo do *T. cruzi* no triatomíneo, enquanto as setas em azul indicam o ciclo no hospedeiro vertebrado.

FONTE: ADAPTADO CDC, 2017

1.1.4 Sintomas e diagnósticos

A DC apresenta sintomatologias distintas que variam de acordo com as fases da doença: aguda ou crônica. A fase aguda é o período inicial da doença e pode durar por até três meses e os primeiros sintomas iniciam entre oito e dez dias após a introdução do *T. cruzi* no organismo. Tais sintomas incluem febre, prostração, anorexia, diarreia, vômitos, cefaleia e mialgia. Em alguns, ainda na fase aguda, podem aparecer edemas que indicam a entrada do parasita no organismo, conhecidos como chagomas e sinais de Romana (Figura 2). A sintomatologia inespecífica na fase aguda após anos, ou até mesmo décadas depois do contágio, pode encaminhar para fase crônica ou para a fase aguda grave, podendo levar a óbito o indivíduo infectado. Na cronificação da doença, pode-se observar a DC cardíaca, digestiva, associada e indeterminada. A forma cardíaca é a maior responsável pelos casos de morte envolvendo a DC e causa dilatação do miocárdio e conseqüentemente insuficiência cardíaca congestiva. Na DC

digestiva observa-se uma frequente evolução para megacólon ou megaesôfago. A forma associada inclui os sintomas da DC cardíaca e digestiva enquanto a forma indeterminada é assintomática, podendo evoluir para qualquer uma das formas crônicas já citadas (ORGANIZAÇÃO PAN AMERICANA DE SAÚDE, 2009).

Figura 2: Sinal de Romaña e chagoma de inoculação



Fonte: DIAS; *et.al*, 2015.

O diagnóstico da DC é clínico e laboratorial, sendo o laboratorial, dividido em sorológicos e parasitológicos. Os exames parasitológicos por sua vez podem ser classificados em diretos e indiretos, sendo os métodos diretos mais limitados em relação aos métodos indiretos. No método diretos, apesar de mais rápidos e financeiramente mais viáveis, a baixa parasitemia pode gerar um resultado falso negativo (DIAS; *et.al*, 2015). Dentre os métodos indiretos temos a hemocultura e o xenodiagnóstico.

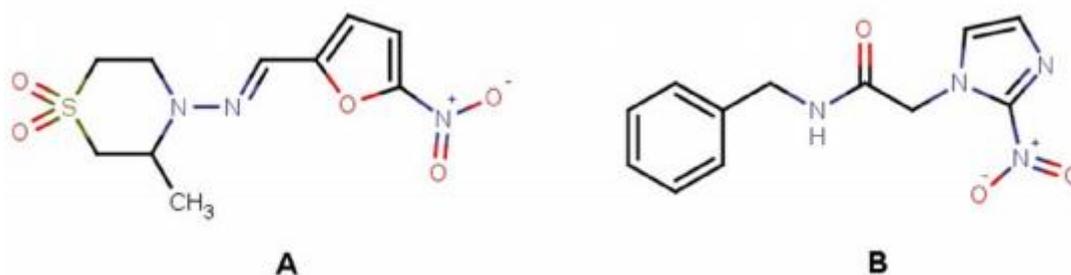
Já para os exames sorológicos, tem-se a pesquisa anticorpos que pode ser realizada por meio da imunofluorescência indireta, hemaglutinação direta e o método imunoenzimático (ELISA). Apesar de mais eficientes no diagnóstico da fase aguda da DC, os exames sorológicos também são usados no diagnóstico da fase crônica e recomenda-se a combinação de um teste com elevada sensibilidade com outro de elevada especificidade. Os exames parasitológicos não são indicados para descartar casos de DC na fase crônica por apresentarem baixa sensibilidade. Porém, caso o resultado seja positivo, este resultado representa diagnóstico absoluto (DIAS; *et.al*, 2015). Outros exames inespecíficos são importantes para o manejo clínico e diagnóstico da DC crônica, e buscam avaliar possíveis complicações causadas ao organismo humano pelo *T. cruzi*. Dentre estes exames temos: o hemograma completo com plaquetas, a urinálise, a eletrocardiografia, a tomografia de tórax e a prova de função hepática (ORGANIZAÇÃO PAN

AMERICANA DE SAÚDE, 2009). Devido à dificuldade no diagnóstico precoce da DC, acredita-se que tal doença seja subnotificada no país (EUZÉBIO, 2015).

1.1.5 Tratamento atual

O tratamento etiológico atual no Brasil é baseado no uso do benznidazol e, nos casos de intolerância a este medicamento, o nifurtimox é utilizado (Figura 3). Esses medicamentos são fornecidos gratuitamente pelo Ministério da Saúde do Brasil, porém são pouco eficazes na fase crônica da doença e apresentam alta toxicidade para o paciente (BRASIL, 2005; SCARIM, et al., 2018). Cerca de 65% dos pacientes em uso de benznidazol apresentam algum evento adverso relacionado a este medicamento, enquanto que entre os pacientes de nifurtimox essa taxa é de 85%. Parestesias, artralguas, intolerância gastrointestinal, alopecia, dermatites e rash cutâneo, são os principais eventos adversos relacionados ao uso do benznidazol. Os eventos observados com o uso de nifurtimox também estão relacionados a artralguas, intolerância gastrointestinal e danos dermatológicos (CONITEC; MINISTÉRIO DA SAÚDE, 2018).

Figura 3: Estrutura nifurtimox (A) e benznidazol (B)



FONTE: RANGEL, 2015

Além das limitações apresentadas, o tratamento etiológico atual é contraindicado para gestantes devido a toxicidade, sendo adotada apenas em casos de DC aguda grave na gestante (BRASIL, 2005).

O tratamento não etiológico da DC visa manejar os danos causados pela doença na fase crônica e é individualizado, levando em consideração o tipo de manifestação clínica apresentada pelo paciente. Tal tratamento farmacológico conta a utilização de medicamentos úteis no manejo da insuficiência cardíaca e miocardites agudas (quando o paciente evolui para forma da DC cardíaca). Para os casos de DC digestiva, as cirurgias são indicadas para o tratamento do megaesôfago e megacólon, além da adequação dos hábitos alimentares. Opções farmacológicas

para o tratamento do megaesôfago incluem o uso relaxantes das fibras musculares enquanto para o megacólon o tratamento farmacológico indicado é o uso de laxativos osmóticos (DIAS; et.al, 2015).

1.2 Desenvolvimento de novos fármacos para o tratamento da DC

Os tratamentos etiológicos para DC disponíveis atualmente, além de apresentarem baixa eficácia no tratamento da doença em sua fase crônica, apresentam alta toxicidade para os pacientes. Tal tratamento é baseado no uso prolongado de benznidazol e nifurtimox e provocam eventos adversos na maioria dos pacientes, o que reduz a adesão ao tratamento. (CONITEC;MINISTÉRIO DA SAÚDE, 2018). Além do tratamento pouco eficaz, a DC é uma doença negligenciada pois está presente predominantemente em ambientes sócio-economicamente desfavorecidos. (LINDOSO;LINDOSO, 2009).

Dado o contexto toxicológico, social e econômico em que a doença se encontra, faz-se necessário a utilização de metodologias eficientes e de baixo custo para a busca de novos tratamentos para a DC.

1.2.1 Triagem virtual de compostos

Entre metodologias que permitem a redução de tempo e custo no desenvolvimento de fármacos, pode-se destacar a triagem virtual de compostos. Esta tem como objetivo selecionar *in silico* estruturas com características desejáveis e potencialmente ativas no alvo molecular escolhido, reduzindo assim o número de moléculas necessárias em ensaios biológicos para que se encontre um composto ativo o que se traduz em uma redução de custos e agilidade na busca de moléculas de interesse farmacológico. Existem técnicas distintas para a execução da triagem virtual, dentre elas pode-se destacar a ancoragem molecular ou docking molecular. Nesta estratégia busca-se através de modelos matemáticos, prever o modo de interação e pontuar a afinidade dos ligantes com o alvo molecular (FERREIRA; et.al, 2015). Atualmente, vários programas de docking estão disponíveis, empregando diferentes algoritmos de busca e funções de pontuação, visando obter maior reprodutibilidade de dados experimentais (pose do ligante no complexo e/ou estimativa da energia de interação) com menor custo computacional possível (KIRCHMAIR; *et al.*, 2008). Dentre essas opções tem-se o GOLD, existente há mais 20 anos, e um dos programas de ancoragem molecular mais citados do mundo. Este programa oferece ao usuário quatro funções de pontuação distintas: GoldScore, ChemScore, ASP e ChemPLP

(LIEBESCHUETZ; COLE; KORB, 2012). Funções de pontuação, através de algoritmos, tem como objetivo atribuir valores para as posições dos ligantes no sítio ação e as possíveis interações que podem acontecer entre ligante-receptor. Desta maneira, a busca por compostos ativos ocorre de forma mais rápida e fácil (NEUDERT; KLEBE, 2011).

Em um *docking* que considera o ligante flexível, um programa de *docking* deve antes encontrar as conformações do ligante no sítio de ligação definido em um receptor. Para isso, diferentes algoritmos de busca podem ser utilizados. O GOLD, por exemplo, utiliza um algoritmo genético na sua busca. Algoritmos genéticos fazem uma alusão à teoria da evolução de Darwin pois armazena características estruturais de um ligante como os ângulos diedros e a posição do ligante em questão no sítio de ligação como genes. Essas informações combinadas formam o cromossomo para cada conformação de ligante. Para cada ligante de entrada, o algoritmo gera diversas conformações, estabelecendo assim uma população inicial cada conformação do ligante é avaliada pela função de pontuação do programa de docking. Aqueles mais aptos (melhor pontuação) tem maior chance de prosseguir no algoritmo, onde poderão sofrer mutações aleatórias em seus genes ou crossover com outro cromossomo. Após inúmeras iterações, o algoritmo tende a evoluir para a solução de maior pontuação (JONES; *et. al.*, 1997)

Muitos programas de docking permitem ao usuário alterar os parâmetros a serem utilizados, o que permite a elaboração de diferentes protocolos para a triagem virtual. Entre esses parâmetros destacam-se a área de busca, o número de corridas independentes e configurações específicas do algoritmo de busca.

Para avaliar o poder preditivo de protocolo de docking em um alvo molecular selecionado pode-se criar uma biblioteca virtual contendo compostos ativos e inativos, e aplicar métricas para validação dos resultados obtidos no docking, como o fator de enriquecimento e área sob curva ROC, de forma a auxiliar na avaliação da qualidade dos protocolos utilizados para a triagem virtual (KIRCHMAIR; *et al.*, 2008; RODRIGUES, *et al.*, 2012).

Essa biblioteca idealmente deve contar com compostos químicos comprovadamente ativos e que apresentem grande diversidade entre si. Tais compostos são obtidos através de bases de dados, como a ChEMBL, e a qualidade desta tem influência direta na qualidade do docking (RODRIGUES; *et,al*, 2012). A ChEMBL é uma plataforma de acesso gratuito, que fornece ao usuário informações importantes sobre a bioatividade dos cerca de dois milhões de compostos cadastrados (MENDEZ; *et,al*, 2018). Ainda, na avaliação de um protocolo de triagem virtual por docking, a biblioteca virtual a ser testada deve conter também compostos inativos. Estes

podem ser obtidos através da plataforma DUD-E (<http://dude.docking.org/>), capaz de gerar *decoys*, a partir de compostos comprovadamente ativos fornecidos à plataforma (MYSINGER *et al.*, 2012).

Os *decoys* são estruturas químicas a priori inativas com características estruturais muito semelhantes às dos compostos ativos e tem como objetivo avaliar a capacidade de um protocolo de *docking* de discernir ativos de inativos (MYSINGER *et al.*, 2012).

1.2.2 A cruzaina como alvo molecular no tratamento da DC

A cruzaina é a principal cisteína protease do *T. cruzi* e está presente nas formas amastigotas, epimastigotas e tripomastigotas deste parasita. Tal enzima tornou-se importante alvo molecular nas pesquisas relacionadas a novos tratamentos da DC, pois tem participação fundamental na aquisição de nutrientes, degradação de proteínas do hospedeiro e evasão do sistema imune (CAFFREY; *et al.*, 2001; McKERROW, 2009).

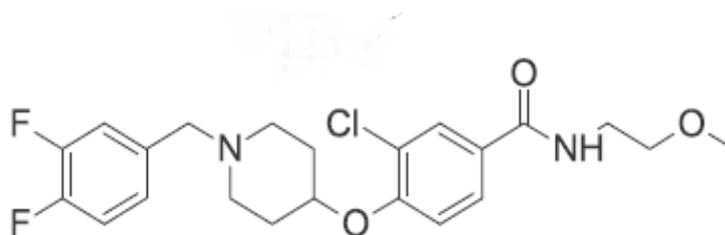
Algumas classes de inibidores da cruzaina já foram descritas em literatura. Tal inibição pode ocorrer de maneira covalente e irreversível ou não-covalente e reversível (SILVA; *et al.*, 2017, ROCHA; *et al.*, 2018). Grande parte dos inibidores já descritos em literatura são peptídicos e apresentam em sua estrutura uma porção eletrofílica, o que confere à essas moléculas características de inibidores irreversíveis. Características essas que não são interessantes quando se analisa biodisponibilidade e fatores sintéticos (RISHTON, 2003; BRAK; *et al.*, 2008).

1.2.3 Aplicação de um protocolo de triagem virtual na seleção de potenciais inibidores da cruzaina

Uma vez estabelecido um protocolo de *docking* em um alvo molecular, como a cruzaina, este pode ser aplicado em novas bibliotecas virtuais com o objetivo de selecionar os compostos mais promissores para atuarem naquele alvo. Neste contexto, análogos do composto MB15 (3-cloro-4-((1-(3,4-difluorobenzil)piperidin-4-il)oxi)-N-(2-metoxietil)benzamida) representam uma classe promissora com potencial tripanocida (Figura 4). O MB15 foi caracterizado como um inibidor de atividade potente frente à enzima cruzaina, com valor de CI_{50} de 5 μ M (PEREIRA; *et al.*, 2019). Além de potente frente a cruzaina, este composto apresentou também boa atividade frente a rodesaina presente no *T. brucei*, parasita responsável pela doença do sono (PEREIRA; *et al.*, 2019). Consistente com esses resultados, o MB15 demonstrou, ainda,

atividade *in vitro* contra diferentes tripanossomatídeos, incluindo o *T. cruzi* (CI₅₀ 32 μM) (KAISER, *et al.*, 2015).

Figura 4 – Estrutura do MB15



FONTE: PEREIRA *et.al*, 2019

Além desses resultados apresentados, o MB15 possui baixa similaridade estrutural com outros compostos já estudados (PEREIRA; *et al*, 2019). Outra vantagem é a presença de diversos pontos passíveis de modificação, o que permite o desenvolvimento de estruturas análogas que, além de aumentar a gama de opções de potenciais inibidores da cruzaina, nos permite explorar pontos chaves de interação entre enzima e ligante.

Assim, a partir de estruturas análogas do MB15 possíveis de serem obtidas utilizando apenas reagentes de partida de fácil acesso, como os disponíveis no laboratório de Química Medicinal e Bioensaios (EFAR-UFOP), pode-se criar uma nova biblioteca virtual com essa classe promissora e, utilizando a triagem virtual por docking molecular, selecionar as moléculas mais promissoras para síntese e avaliação biológica.

2. OBJETIVOS

2.1 Objetivo Geral

O presente trabalho teve como intuito desenvolver um protocolo de triagem virtual com ancoragem molecular para a seleção de potenciais inibidores da cruzafina, base para o desenvolvimento de fármacos para o tratamento da doença de Chagas.

2.2 Objetivos Específicos

Para atingir esse objetivo, objetivou-se, especificamente:

- Analisar e selecionar uma estrutura tridimensional da enzima cruzafina;
- Selecionar, a partir de dados experimentais, estruturas ativas contra a cruzafina;
- Obter uma série de *decoys* baseados nas estruturas ativas selecionadas;
- Realizar a ancoragem molecular dos ativos e *decoys* com diferentes protocolos;
- Analisar a capacidade preditiva dos protocolos desenvolvidos a partir do fator de enriquecimento e curva ROC (AUC);
- Criar uma biblioteca virtual de moléculas passíveis de serem obtidas a partir de reagentes disponíveis no laboratório de Química Medicinal e Bioensaios da Escola de Farmácia da Universidade Federal de Ouro Preto, para obtenção de análogos dessa estrutura;
- Aplicar o melhor protocolo de triagem virtual ao conjunto de moléculas criadas;
- Selecionar as moléculas melhores ranqueadas para priorização na síntese

3. METODOLOGIA

3.1. Obtenção e preparo da estrutura cristalográfica da cruzaina

Utilizando a base de dados do *Protein Data Bank* (PDB - <https://www.rcsb.org/>) foi obtida a estrutura tridimensional da cruzaina em complexo com um ligante de caráter não-peptídico e modo de ligação reversível (código PDB 3KKU).

Após a obtenção do complexo cristalográfico, o mesmo foi preparado utilizando o programa Hermes (CCDC). O preparo consistiu inicialmente em análise prévia dos aminoácidos constituintes do sítio de ligação da 3KKU, observando a presença de grupos ionizáveis, tautômeros da His, inversões de Asn e Gln, além da ocupância dos resíduos no sítio de ligação. A conformação da Cys25 de maior ocupância foi escolhida. As moléculas de água presentes no sítio de ligação da estrutura cristalográfica, o artefato de cristalização Z222 e o ligante B95 foram removidos. Como os aminoácidos His162 e Cys25 podem estar na forma neutra ou ionizada, foram estabelecidos protocolos de dockings distintos, cada um considerando a formação do par iônico entre eles ou a forma neutra.

3.2 Seleção de inibidores da cruzaina com atividade determinada experimentalmente

Utilizando o programa DataWarrior para importar informações da base de dados ChEMBL, obteve-se 32.273 compostos com atividade inibitória determinada experimentalmente contra a enzima cruzaina. Estruturas contendo os grupos nitrilas (MOTT; et al., 2010) epóxidos (SILVA, 2017), vinilssufonas (ENGEL; et al., 1998) e aceptores de Michael (RANGEL, 2015) foram removidas por meio de filtros do DataWarrior por apresentarem potencial efeito inibitório irreversível da cruzaina devido a presença desses grupos eletrofílicos. Após tal remoção, estruturas contendo elementos como selênio, boro, silício e arsênio também foram removidas devido por não serem usualmente encontrados na estrutura de fármacos e não reconhecidos por alguns métodos de modelagem molecular. A atividade da molécula bem como suas características físico-químicas foram calculadas e utilizadas como critérios para seleção do conjunto de moléculas. Foram mantidas as moléculas com as seguintes características:

- Massa molar entre 150 e 500 daltons
- LogP de 0 a 5
- Máximo de 10 aceptores de hidrogênio por estrutura

- Máximos de 5 doadores de hidrogênio por estrutura
- CI_{50} contra cruzaína $< 10 \mu\text{M}$

No final, para obter um conjunto com estruturas mais diversificadas, utilizou-se o DataWarrior para selecionar 250 moléculas mais diferentes entre si dos 21.416 compostos ativos restantes. À esta lista foi adicionado o MB15, protótipo utilizado na construção da biblioteca virtual, totalizando 251 compostos inibidores da enzima cruzaína.

3.3 Obtenção dos *decoys*

Para a obtenção de *decoys*, foi utilizada a plataforma online DUD-E (<http://dude.docking.org/>), utilizando como entrada o código SMILES dos 251 inibidores selecionados na etapa anterior. Para cada inibidor, foram obtidos 50 *decoys*.

3.4 Preparo dos ligantes

As estruturas dos inibidores de cruzaína (item 3.2) e dos *decoys* (item 3.3), foram submetidos ao módulo cxcalc da Marvin Suite (<https://chemaxon.com/products/marvin>), para determinar o estado de protonação majoritário dos ligantes em pH de 5,5. Os inibidores e *decoys* tiveram, então, suas estruturas tridimensionais geradas e minimizadas utilizando o programa Open Babel com auxílio do *script* POAP. Para esta etapa foram utilizadas as seguintes configurações:

- Campo de força MMFF94
- Método de minimização: *steepest descent*
- Número de etapas: 5000
- Critério de convergência: 1×10^{-6}
- Método de geração das coordenadas 3D: mais lento
- Solicitamos que a ferramenta gerasse 10 confôrmeros para cada estrutura pelo método *obconformer* seguido de 250 passos de minimização e selecionasse a de menor energia dentre essas 10 opções.
-

3.5 Criação de uma biblioteca virtual de compostos para triagem virtual

Para criação de uma biblioteca virtual, utilizou-se o MB15 como protótipo e levou-se em consideração os reagentes disponíveis no Laboratório de Química Medicinal e Bioensaios da Escola de Farmácia da Universidade Federal de Ouro Preto, para obtenção de análogos dessa estrutura. Após a análise retrossintética do MB15, estabeleceu-se que aminas, aldeídos e ácido carboxílicos seriam necessários para síntese dos análogos, sendo encontrados 3 aldeídos, 10 ácidos carboxílicos e 19 aminas disponíveis no laboratório. Com o auxílio do programa DataWarrior, foram realizadas análises combinatórias com esses reagentes, fornecendo todos possíveis produtos que poderiam ser formados e posteriormente utilizados como reagentes na obtenção dos análogos do MB15.

3.6 Desenvolvimento dos protocolos de ancoragem molecular

Para realizar os cálculos de docking foi utilizado o programa GOLD (CCDC). Diferentes protocolos foram desenvolvidos abordando três variáveis: tamanho da área de busca (distância a partir do ligante cristalográfico), função de pontuação e estado de ionização dos aminoácidos His162 e Cys25 presentes no sítio de ligação. Levando em consideração tais variáveis, os seguintes protocolos descritos na Tabela 1 foram estabelecidos:

Tabela 1: Descrição dos protocolos de ancoragem molecular analisados para a validação

Código do protocolo	Estado de ionização do par His162/Cys25	Tamanho da área de busca	Função de pontuação
8iongold	Ionizado	8 Å	Goldscore
8neugold	Não ionizado	8 Å	Goldscore
8ionplp	Ionizado	8 Å	ChemPLP
8neuplp	Não ionizado	8 Å	ChemPLP
5iongold	Ionizado	5 Å	Goldscore
5neugold	Não Ionizado	5 Å	Goldscore
5ionplp	Ionizado	5 Å	ChemPLP
5neuplp	Não Ionizado	5 Å	ChemPLP

FONTE: AUTORIA PRÓPRIA

Os demais parâmetros foram padronizados para todos dockings, sendo a eficiência do algoritmo de busca definida como 30% e o número de cálculos independentes realizados para cada ligante

em dez. A opção *early termination* foi ativada para que os cálculos fossem encerrados caso as três melhores soluções tivessem um RMSD menor do que 1,5 Å entre elas. Para cada ligante, apenas a melhor solução foi salva.

3.7 Normalização

Ao score final de cada ligante (para ambas funções de pontuação) foram aplicadas a seguinte fórmula de normalização:

$$Score\ normalizado = \frac{Score}{N^{\frac{1}{3}}}$$

Onde N é igual ao número de átomos pesados na estrutura (todos, exceto os hidrogênios), previamente calculados (item 3.2).

3.8 Cálculo das métricas de validação

Como métricas de validação foram analisados os fatores de enriquecimento e a área sob a curva ROC (AUC). Esta foi obtida através da plataforma *Screening explorer* (<http://stats.drugdesign.fr/>) e levou em consideração todos os compostos ranqueados. Para o cálculo do fator de enriquecimento (EF), utilizou-se as frações 1%, 2% e 10% dos compostos mais bem ranqueados e foi utilizada a seguinte fórmula:

$$EF = \frac{\left(\frac{a}{b}\right)}{\left(\frac{A}{B}\right)}$$

Onde:

a corresponde ao número de compostos ativos presentes na fração analisada; b o número de total de compostos na mesma fração analisada; A é número total de compostos com atividade conhecida e B o número total de compostos submetidos ao docking (PAN, 2003).

3.9 Virtual screening da biblioteca virtual e seleção dos compostos

O protocolo com os melhores resultados de fator de enriquecimento 1% e AUC ROC foi escolhido e aplicado aos compostos da biblioteca virtual criada (item 3.5).

4 RESULTADOS E DISCUSSÕES

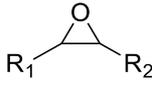
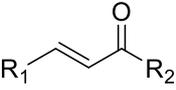
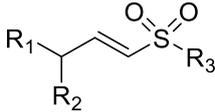
4.1 Obtenção e preparo da estrutura cristalográfica da cruzaina

Utilizando a base de dados do *Protein Data Bank* e buscando pelo termo “*cruzain*” encontrou-se 31 resultados. Dentre estes, a estrutura cristalográfica escolhida de código 3KKU por representar um complexo com ligante não covalente (seguindo o padrão dos ligantes utilizados neste trabalho) e boa resolução (1,28 Å). Durante o preparo da proteína foram analisadas a presença de moléculas de água no sítio de ligação dos ligantes. Foi observado que tais moléculas não possuem interações importantes no sítio de ação e por isso foram removidas. Observou-se que o par His162/Cys25 presente no sítio de ligação pode apresentar dois estados de ionização diferentes e por isso os dois estados de ionização foram considerados, obtendo ao final desta etapa a 3KKU com o par His162/Cys25 não ionizado e ionizado.

4.2 Seleção de inibidores de cruzaina de atividade determinada experimentalmente

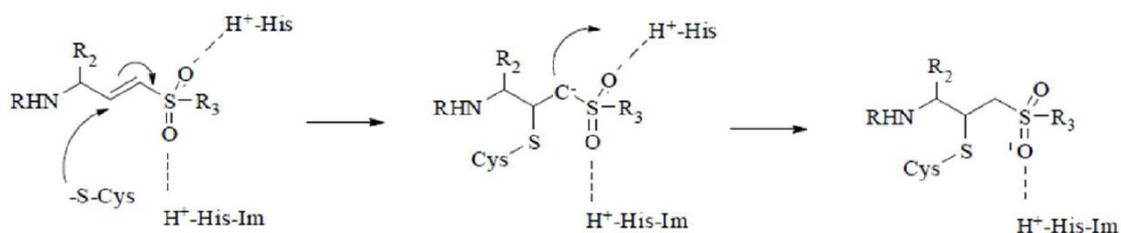
A enzima cruzaina presente no *T. cruzi* é um importante alvo de pesquisa para desenvolvimentos de fármacos que possam atuar de maneira a inibir sua atividade (CAFFREY; *et al.*, 2001; McKERROW, 2009). Tal inibição pode ser reversível ou irreversível, de acordo com as características da interação entre o ligante e o sítio de ação. Ligações do tipo covalente estão relacionadas a uma inibição irreversível da cruzaina. Efeitos *off-target* e toxicidade são eventos mais frequentes com inibidores irreversíveis em relação aos reversíveis (RANGEL, 2015). Com isso, dentre os ligantes obtidos no item 3.2, optou-se pela seleção de compostos que não apresentaram grupos químicos relacionados à ligação irreversível. Entre esses grupos, destacam-se as nitrilas (MOTT; *et al.*, 2010), epóxidos (SILVA, 2017), vinilssufonas (ENGEL; *et al.*, 1998) e aceptores de Michael (RANGEL, 2015) que inclui, além das vinilssufonas, as carbonilas $\alpha\beta$ -insaturada (Tabela 2).

Tabela 2 - Grupos químicos relacionados à ligação covalente com a cruzaina.

Grupo	Estrutura química
Epóxido	
Nitrila	$R-C\equiv N$
Carbonilas $\alpha\beta$ -insaturada	
Vinissulfonas	

FONTE: AUTORIA PRÓPRIA

O mecanismo de inibição da cruzaina envolvendo aceptores de Michael (Figura 5) pode ser descrito da seguinte maneira: inicialmente o resíduo de aminoácido Cys25 presente no sítio catalítico atua como nucleófilo e ataca o carbono eletrofílico da insaturação conjugada do inibidor, em uma adição de Michael, fazendo com que a enzima fique covalente ao inibidor e com isso inibida de forma irreversível; em seguida o carbânion formado capta um próton do meio reacional, no caso, do resíduo de histidina (RANGEL, 2015).

Figura 5 - Mecanismo de inibição irreversível da cruzaina por aceptor de Michael (vinilssulfonas).

FONTE: MEHRTENS,2007.

Ainda na seleção dos ligantes, estruturas contendo semimetais foram excluídas do presente trabalho por não serem comuns em fármacos, estarem fora do escopo de síntese apresentado e apresentarem potencial tóxico. Dentre as estruturas restantes e ainda utilizando o programa DataWarrior, aplicou-se a regra de Lipinski para determinar o conjunto de moléculas utilizadas

no docking. A regra de Lipinski direciona o planejamento de fármacos de tal forma que a molécula tenha propriedades físico-químicas adequadas para uma boa biodisponibilidade oral (SANTOS, 2018). Essa regra é composta pelos seguintes parâmetros:

- A estrutura deverá ter 5 ou menos doadores de ligação de hidrogênio;
- A estrutura deverá ter 10 ou menos aceptores de ligação de hidrogênio;
- O peso molecular deverá ser menor do que 500 daltons;
- O valor de logP deverá ser inferior a 5 (SANTOS, 2018).

Além desses critérios, outro parâmetro importante utilizado na seleção dos ligantes foi o valor de CI_{50} . Este valor determinado experimentalmente determina a dose necessária para inibir metade da atividade enzimática das enzimas-alvo, que neste caso é a cruzaina. Sendo assim, é desejável valores mais baixos de CI_{50} (BAG, 2016). Com isso, compostos que apresentaram valores de CI_{50} maiores que $10\mu\text{M}$ não foram selecionados para compor o conjunto de moléculas ativas deste trabalho.

Após aplicação destes filtros, restaram 21.416 moléculas. Para adequar o trabalho à capacidade de processamento disponível foi solicitado ao DataWarrior a seleção dos 250 compostos mais distintos entre si, com intuito de utilizar moléculas com ampla diversidade estrutural na validação do protocolo. Por ser um inibidor com propriedade *drug-like*, CI_{50} igual a $5\mu\text{M}$ (PEREIRA; et al., 2019) e ser o protótipo investigado neste trabalho, o MB15 foi adicionado manualmente a esta lista de 250 compostos, já que seleção aleatória do DataWarrior não o havia selecionado, totalizando então, 251 compostos ativos contra a cruzaina selecionados.

4.3 Obtenção dos *decoys*

Decoys são estruturas químicas a priori inativas com características estruturais muito semelhantes às do composto comprovadamente ativo. O uso de *decoys* na ancoragem molecular é de suma importância, pois, analisa a capacidade do *docking* de preterir os inativos em relação aos ativos. Quanto melhor a capacidade do *docking* em diferenciar ativos dos inativos, menores serão as respostas falso-positivas (GRAVES, 2005). Utilizou-se a plataforma online DUD-E (<http://dude.docking.org/>) para gerar 50 *decoys* para cada ativo, sendo que cada *decoy* foi gerado de forma que suas características físico-químicas fossem semelhantes à do ativo. A plataforma DUD-E conta com um banco de dados com 102 proteínas catalogadas e 22 mil moléculas que são ativas em tais proteínas, mas, para o presente trabalho utilizou-se a função

de geração de decoys a partir de uma biblioteca de ativos previamente preparada. Desta maneira, a partir dos ativos, a plataforma gerou decoys estruturalmente distintos dos ativos, mas, preservando características físico-químicas similares aos dos ativos de origem (MYSINGER; et.al, 2012).

4.4 Preparo dos ligantes

Os ligantes obtidos nos itens 3.2 e 3.3 foram submetidos ao módulo cxcalc da Marvin Suite (<https://chemaxon.com/products/marvin>), para determinar o estado de protonação majoritário dos ligantes em pH de 5,5. Esta etapa foi realizada, pois o pH ótimo de atuação da cruzafina é de 5,5 (GILMOR; CRAIK; FLETTERICK, 1997) e o estado de protonação do ligante é importante para análise das interações da enzima – ligante (PEREIRA, 2017). Após esta etapa, os protômeros majoritários foram minimizados por meio do programa OpenBabel. Quando desenhadas, mesmo em três dimensões, as estruturas não consideram aspectos estéricos, eletrostáticos e torcionais da molécula, o que torna esta representação diferente do que se espera encontrar na molécula real (CARVALHO, 2003). Com intuito de tornar essas representações mais similares às moléculas reais, faz-se a minimização das estruturas. Esta minimização pode ser feita através de modelos matemáticos baseados em mecânica molecular que buscam minimizar a energia estérica da molécula, dada pelo desvio dos ângulos, comprimentos de ligação e ângulos torcionais de valores considerados ideais (CARVALHO, 2003). Neste contexto, o programa Open Babel oferece três opções predefinidas de campo de força:

- MMFF94: ideal para moléculas orgânicas e *druglike*
- UFF: ideal para estruturas com geometria mais complexa e não suportadas pelo campo MMFF94 (O'BOYEL, 2011)
- GAFF: aplicável para a maioria dos fármacos.

O campo de força mais recomendado para otimizações é o MMFF94, por ser mais eficiente e amplamente utilizado. Por isso, os demais campos só deverão ser utilizados quando não for possível a otimização pelo MMFF94 (RABELLO, 2016). Durante a minimização do conjunto de moléculas do presente trabalho, três ativos não tiveram suas estruturas minimizadas pelo programa. Analisando suas estruturas, observou-se que estas apresentavam muitos centros quirais e conseqüentemente geometria complexa, o que pode explicar a falha do programa para minimizar essas estruturas pelo campo de força MMFF94. Com isso, optou-se pela exclusão dessas moléculas e seguiu-se com 248 no conjunto de estruturas ativas contra a cruzafina.

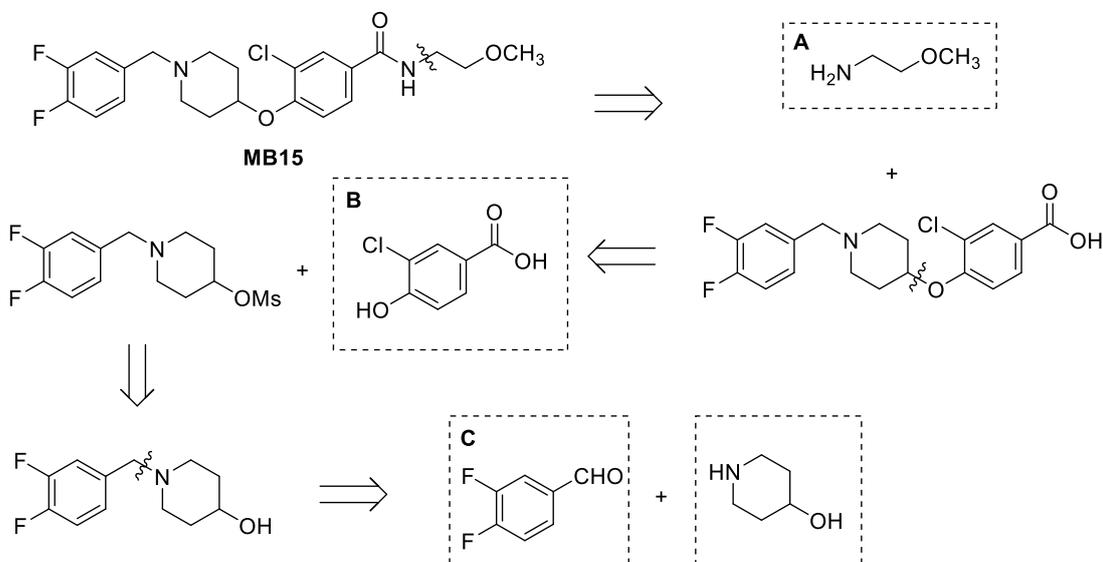
O mesmo protocolo foi aplicado aos *decoys*. Da mesma maneira, o OpenBabel falhou em obter estruturas tridimensionais minimizadas para alguns desses ligantes e o total de *decoys* minimizados obtidos foi de 13.133.

4.5 Criação de uma biblioteca virtual de compostos para triagem virtual

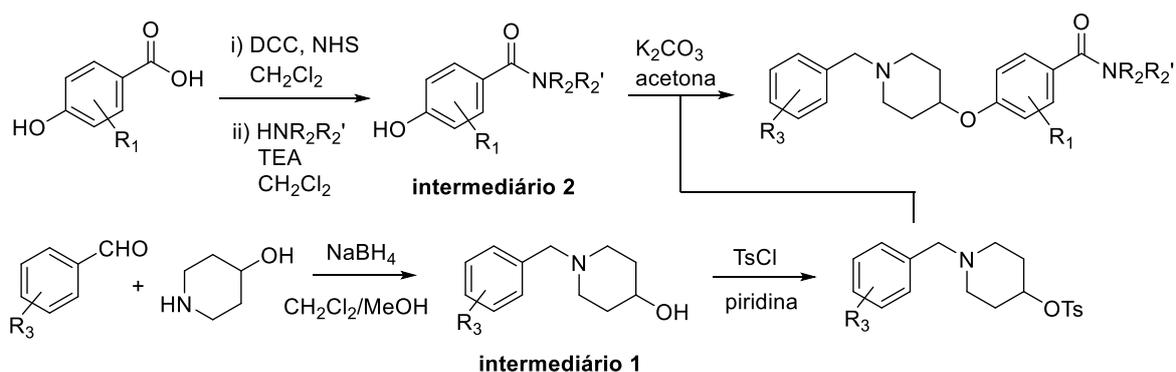
No presente trabalho se propôs avaliar com o protocolo de docking estabelecido análogos do MB15 por meio de uma criação de uma biblioteca virtual. Para tal, inicialmente foi realizada a análise retrossintética deste composto seguida de uma proposta de síntese (Figura 6).

Figura 6: Análise retrossintética e proposta de rota de síntese

A) Análise retrossintética do MB15



B) Rota de síntese proposta



A) Análise retrossintética do MB15 e B) Proposta de rota sintética para obtenção de análogos do MB15

FONTE:AUTORIA PRÓPRIA

Após análise retrossintética, definiu-se que seriam necessários como reagentes de partida para obtenção dos análogos, aldeídos e a 4-hidroxipiperidina (obtenção do intermediário 1) e ácidos carboxílicos e aminas (obtenção do intermediário 2). Após acoplamento desses intermediários por substituição nucleofílica, uma grande variedade estrutural do MB15 poderia ser obtida. A biblioteca de análogos do MB15 foi então construída a partir dos reagentes disponíveis no laboratório de Química Medicinal e Bioensaios da Escola de Farmácia da Universidade Federal de Ouro Preto, sendo: 19 aminas, 3 aldeídos e 10 ácidos carboxílicos (Tabelas 1, 2 e 3 do ANEXO A). As estruturas destes análogos foram obtidas por análise combinatória desses reagentes utilizando a função *Enumerate Combinatorial Library* do DataWarrior, totalizando 570 análogos do MB15.

Para as aminas contendo mais de um nitrogênio (ex.: *N*-(1-naftil)etilenodiamina), o produto foi limitado para a reação no nitrogênio mais nucleofílico, sendo essa seleção feita de forma manual. Após a obtenção dos 570 análogos, estes foram preparados de acordo com o protocolo estabelecido no item 4.4 e posteriormente submetidos ao *docking*.

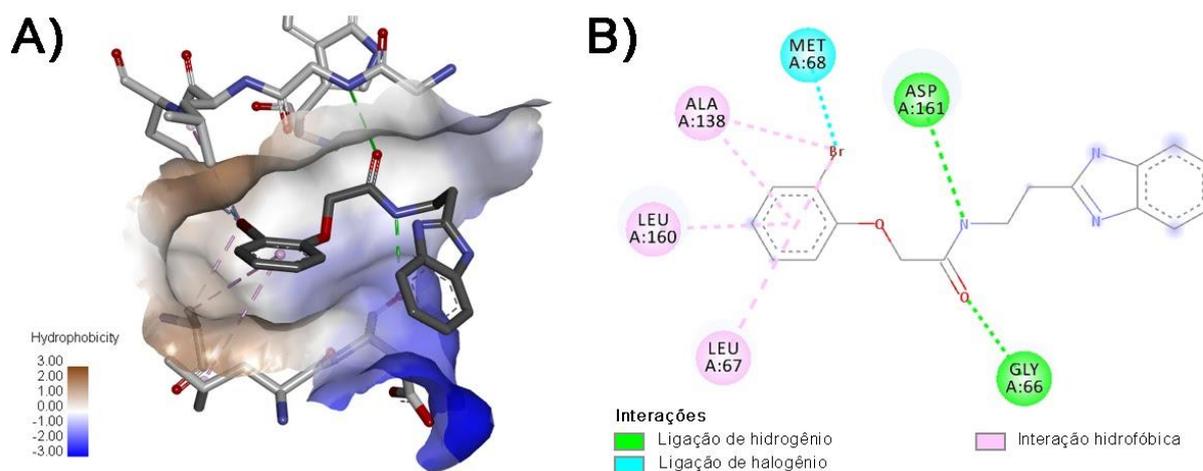
4.6 Desenvolvimento dos protocolos de ancoragem molecular

O programa GOLD (CCDC) existe há mais 20 anos e é um dos programas de ancoragem molecular mais citados do mundo (LIEBESCHUETZ; COLE; KORB, 2012). Funções de pontuação, por meio de algoritmos, atribuem valores para as possíveis interações entre ligante e para o posicionamento destes ligantes no sítio de ação. Cada função de pontuação tem uma maneira particular de atribuir tais pontuações. O programa utilizado neste trabalho oferece ao seu usuário quatro funções de pontuação distintas: GoldScore, ChemScore, ASP e ChemPLP. (VERDONK; *et.al*, 2003) e optou-se por desenvolver protocolos com duas, das quatro funções disponíveis: a pioneira GoldScore e a introduzida mais recentemente, ChemPLP. Todas as funções de pontuação possuem diferenças significativas entre si. Das funções utilizadas neste trabalho e em relação as interações ligante-proteína, a GoldScore pontua as interações entre ligante e proteína por uma função utilizando mecânica molecular, considerando a energia das interações do tipo van der Waals e ligações de hidrogênio (LIEBESCHUETZ; COLE; KORB 2012). Já a função ChemPLP, é uma função empírica que pode ser customizada conforme o tipo da proteína alvo, e pontua a interação conforme naturezas que os átomos do ligante e os átomos da proteína podem apresentar. Tais naturezas são descritas como aceptoras, doadoras,

não polar e aceitadores e doadores concomitantemente. A presença de íons metálicos também é pontuada conforme a natureza do metal (CCDC, 2016).

Além da escolha da função de pontuação adequada, a área de busca também é um parâmetro fundamental no desenvolvimento de protocolos de triagem virtual. Uma área de busca pequena pode restringir que os ligantes mais volumosos façam um encaixe efetivo no sítio de ação, enquanto que uma área de busca muito grande pode aumentar o tempo de busca de forma significativa com encaixes não efetivos. Partindo de um sítio de ligação já conhecido, pode-se restringir a área de busca para esta região do complexo cristalográfico utilizado (FEINSTEIN; BRYLINSK, 2015). No presente trabalho, partiu-se de um sítio de ligação já bastante estudado e, baseando-se na presença de um ligante, estabeleceu-se como área de busca os resíduos distantes até 5 Å e 8 Å do mesmo. Na figura 7 podemos observar as interações entre o ligante B95 e complexo cristalográfico 3KKU, que envolve tanto interações hidrofóbicas com o anel halogenado quanto ligações de hidrogênio com a amida. Quando se amplia a área de busca para resíduos a 8 Å desse ligante, além de abranger as interações já registradas, permite o encaixe de moléculas maiores e que eventualmente possam apresentar pontos adicionais de interação.

Figura 7: Interações intermoleculares entre B95 e cruzaina.



Interações intermoleculares entre B95 e cruzaina em 3D (A) e diagrama 2D (B). Em (A) uma superfície acessível ao solvente foi construída em torno dos resíduos do sítio e colorida de acordo com a hidrofobicidade do resíduo. As linhas tracejadas representam as interações intermoleculares observadas. Em verde, ligações de hidrogênio, em rosa as interações hidrofóbicas do tipo Carbono–Carbono e Sistema Pi–Carbono e em azul as ligações de halogênios. Átomos expostos ao solvente estão demarcados com um círculo azulado.

FONTE: AUTORIA PRÓPRIA

Já em relação ao estado de ionização, o complexo cristalográfico 3KKU foi obtido através de difração de raios-X e com isso não é possível determinar com certeza o estado de ionização do par His162/Cys25, uma vez que a densidade eletrônica dos hidrogênios não é suficiente para ser determinar sua posição nas resoluções comumente obtidas por essa técnica (PEREIRA 2017). Baseado nisto, optou-se por utilizar ambos estados de ionização (forma neutra e ionizada), uma vez que o pKa da cadeia lateral desses resíduos (6,0 e 8,18, respectivamente) gera incerteza quanto ao seu estado de ionização (LEHNINGER; NELSON; COX,2014).

4.7 Normalização

Os modelos matemáticos utilizados no cálculo da interação entre ligantes e sítio de ação tendem a pontuar melhor as moléculas com maior número de átomos, uma vez que estes podem estabelecer inúmeras interações de van der Waals, de caráter sempre aditivo na função de pontuação (PAN, 2003). Desta forma, os compostos com as maiores massas moleculares geralmente acabam sendo favorecidos durante os cálculos. Com isso, faz-se necessário a normalização para ajustar a pontuação considerando os átomos pesados, ou seja, aqueles que são diferentes de hidrogênio e contabilizados nessa interação (PAN, 2003). Para o presente trabalho, foi feita a normalização considerando todos os diferentes de hidrogênio, utilizando a seguinte fórmula (PAN, 2003):

$$Score\ normalizado = \frac{Score}{N^{\frac{1}{3}}}$$

Optou-se pelo expoente 1/3 no número de átomos pesados (N) como forma de reduzir a tendência do *docking* em pontuar melhor as moléculas mais pesadas. Este expoente foi selecionado baseado na observação experimental de Pan e colaboradores em que concluíram que o expoente 1/3 foi capaz de reduzir a análise tendenciosa do *docking* sem se tornar discrepante em relação aos dados da biblioteca original (não normalizada). Durante a seleção virtual dos compostos, optou-se por não utilizar moléculas com massa superior à 500 daltons, o que caracteriza uma estratégia adicional para reduzir a tendência do *docking* em favorecer moléculas mais pesadas.

4.8 Métricas e validação

O fator de enriquecimento nos auxilia a mensurar o desempenho do algoritmo de busca na triagem molecular, já que informa quantas vezes a mais o protocolo aplicado é capaz de

selecionar os ligantes ativos em relação a uma seleção ao acaso, de ativos em meio aos *decoys*. Para melhorar análise de desempenho das funções de pontuação, preteriu-se analisar frações dos compostos com as melhores pontuações, uma vez que, quando se utiliza bibliotecas de compostos extensa, a tendência é que as funções de pontuação apresentem desempenho similares (LIEBESCHUETZ; COLE; KORB, 2012). Já a curva ROC (Receiver Operating Characteristic Curve - Característica de Operação do Receptor) é obtida ordenando os resultados de uma triagem virtual em ordem decrescente e plotando o percentual de ativos verdadeiros encontrados na triagem virtual pelo percentual da biblioteca explorada. Assim, em um cenário ideal, todos os compostos ativos seriam ranqueados na fração inicial da triagem virtual e curva adquiriria uma inclinação inicial elevada, até que todos os compostos ativos fossem encontrados. Normalmente utiliza-se a área sob a curva ROC (AUC) para mensurar quanto o protocolo aplicado foi capaz de ranquear melhor os ativos em relação aos inativos, no caso, *decoys*. Tal métrica apresenta valores entre 0 e 1, sendo que o valor de 0,5 representa uma seleção aleatória da triagem e valores inferiores à 0,5 correspondem à protocolos em que os ativos foram preteridos em relação aos *decoys* (BROZELL, 2012). Para o presente trabalho, observou-se melhora nos valores de fator de enriquecimento conforme as normalizações foram aplicadas. Isso ocorreu, pois, a tendência da triagem em preferir moléculas com alto peso molecular foi atenuada (Tabela 3).

Tabela 3 – Fator de enriquecimento considerando a fração 1%, 2% e 10% dos compostos melhores ranqueados considerando o score antes e após normalização pelo número de átomos pesados. (continua...)

Fator de enriquecimento top 1%

Protocolo	5iongold	5neugold	5ionplp	5neuplp	8iongold	8neugold	8ionplp	8neuplp
EF	1,42	1,21	2,82	2,37	1,21	2,01	2,82	2,82
EF (normalizado)	2,84	2,81	4,03	3,62	1,62	1,21	4,03	3,22

Fator de enriquecimento top 2%

Protocolo	5iongold	5neugold	5ionplp	5neuplp	8iongold	8neugold	8ionplp	8neuplp
EF	1,65	1,41	2,42	2,01	1,41	2,21	2,21	2,62
EF (normalizado)	2,12	2,22	2,41	2,82	1,61	2,01	2,82	3,02

Tabela 3 – Fator de enriquecimento considerando a fração 1%, 2% e 10% dos compostos melhores ranqueados considerando o score antes e após normalização pelo número de átomos pesados. (...conclusão)

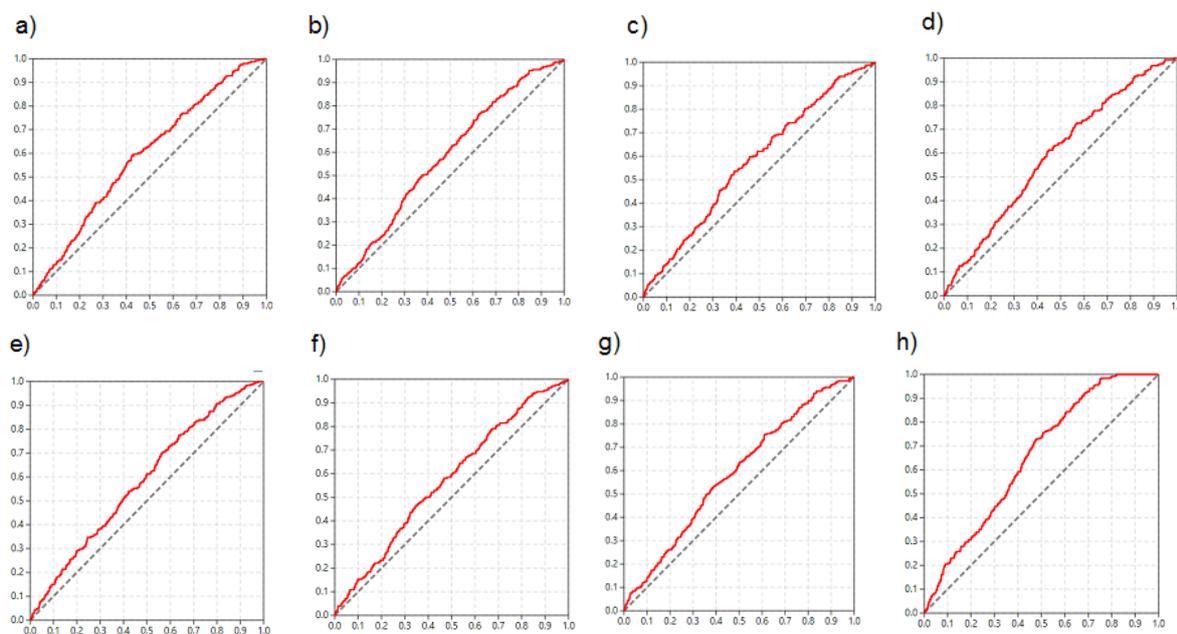
Fator de enriquecimento top 10%

Protocolo	5iongold	5neugold	5ionplp	5neuplp	8iongold	8neugold	8ionplp	8neuplp
EF	1,60	1,37	1,21	1,25	1,81	1,41	1,33	1,37
EF (normalizado)	1,51	1,65	1,49	1,53	1,69	1,89	1,65	1,61

FONTE: AUTORIA PROPRIA

Os valores de AUC da curva ROC, mesmo após as normalizações, não apresentaram melhoras importantes (Figura 8) Nenhum valor de AUC foi inferior a 0,50, logo, em todos os protocolos, mesmo não alçando o valor desejável da área sobre a curva, a seleção feita pelo algoritmo foi melhor do que o aleatório (BROZELL, 2012). No entanto, observou-se que, no geral, o valor da AUC é maior nas porções iniciais do ranking, fato que foi refletido no fator de enriquecimento considerando diferentes porções do ranking. Assim, os protocolos utilizados possuem maior probabilidade em selecionar compostos ativos nas frações iniciais do ranking (top1% e top 2%).

Figura 8 - Curvas ROC AUC dos protocolos aplicados.

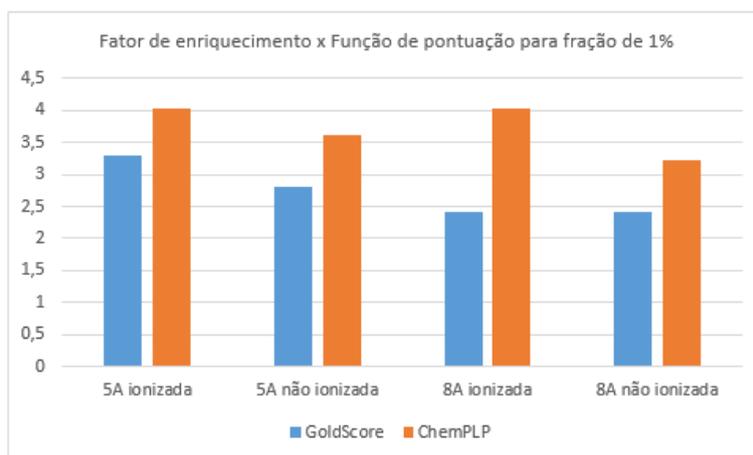


Fracção de falsos positivos x Fracção de verdadeiramente positivos. Em **a)** protocolo 5neuGold ROC AUC: 0.591, **b)** protocolo 5neuPlp ROC AUC: 0.581, **c)** protocolo 8neuPlp ROC AUC: 0.580, **d)** protocolo 5ionPlp ROC AUC:0.573, **e)** protocolo 5ionGold ROC AUC: 0.584, **f)** protocolo 8neuGold ROC AUC: 0.59, **g)** protocolo 8ionPlp ROC AUC:0.587, **h)** protocolo 8ionGold. ROC AUC: 0.653.

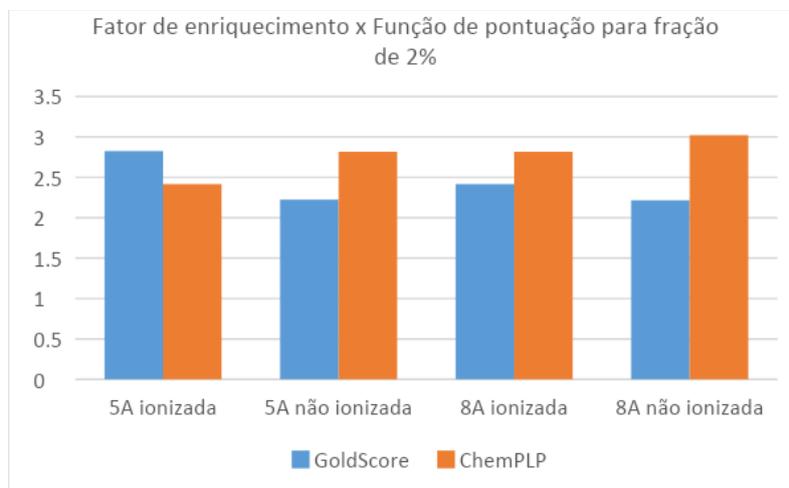
Pode-se levantar algumas hipóteses que justificam o valor de AUC abaixo do esperado, são elas:

- A proporção entre *decoys* e ativos: levando em consideração todos os ativos e *decoys* descartados na minimização, a triagem foi feita com uma proporção de cerca de 54 *decoys* para cada ativo. Isso aumenta em cerca de 54 vezes a chance de o ativo ser preterido em relação ao *decoy* o que se traduz em uma dificuldade adicional para o protocolo de triagem
- A possibilidade dos *decoys* serem ativos: Os *decoys* foram gerados a partir de estruturas dos compostos verdadeiramente ativos. Tais estruturas foram tidas como inativas, por não apresentarem estudos que citem sua atividade. Porém, existe a possibilidade de algumas estruturas dos *decoys* serem compostos verdadeiramente ativos e ainda não estudados.
- Limitações da função do programa: O GOLD, apesar de considerado um dos melhores programas de docking existentes, não apresenta um bom sistema de classificação para as melhores poses (WANG; *et al*,2016)
- Limitações da estrutura cristalográfica: Mesmo utilizando a 3KKU, cristal que apresenta uma boa resolução, não é possível determinar com precisão o estado de ionização dos resíduos, como o que ocorre com o par His162/Cys25 (PEREIRA, 2017). Tal incerteza pode direcionar a resultados e interpretações incorretas.
- Limitações da metodologia do docking: como a amostragem limitada dos ligantes e não consideração da água como solvente explícito e a flexibilidade da proteína (FERREIRA, 2015).

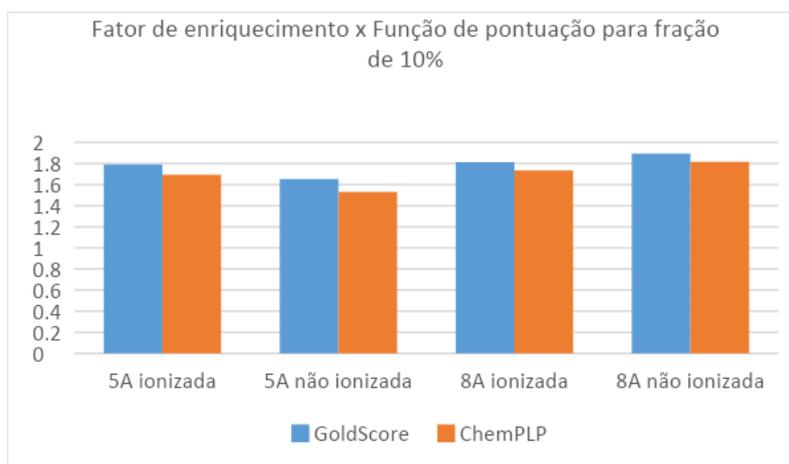
Utilizando o fator de enriquecimento nas frações 1%, 2% e 10%, comparou-se os parâmetros variados nos protocolos. Os gráficos a seguir (Gráfico 1, Gráfico 2 e Gráfico 3) comparam o desempenho das funções GoldScore e ChemPLP, o tamanho da área de busca e o estado de ionização do par His162/Cys25 quando aplicadas a biblioteca formada pelos ativos e *decoys*.

Gráfico 1: Fator de enriquecimento x Função de pontuação para fração de 1%

FONTE: AUTORIA PROPRIA

Gráfico 2: Fator de enriquecimento x Função de pontuação para fração de 2%

FONTE: AUTORIA PROPRIA

Gráfico 3: Fator de enriquecimento x Função de pontuação para fração de 10%

FONTE: AUTORIA PROPRIA

Para a fração de 1% das moléculas melhores ranqueadas, o estado ionizado Cys/His apresentou pequena vantagem sobre o par não ionizado. A função de pontuação ChemPLP apresentou melhor desempenho nas frações 1% e 2%. Já na fração de 10%, as funções de pontuação não apresentaram grande diferença sobre o enriquecimento. Este mesmo padrão foi observado por Liebeschuetz, Cole, e Korb, (2012) que concluíram, também de forma experimental, que quando se trabalha com um extenso set de moléculas, as funções de pontuação apresentam desempenhos similares, enquanto para set menores o uso da função ChemPLP é mais vantajoso e isso acontece, pois, quando diminui o tamanho da biblioteca ocorre também uma redução na porcentagem de falsos positivos. Desta maneira, pode-se observar com mais clareza a atuação da função de pontuação e do protocolo escolhido. Além de discriminar de forma mais eficiente os *decoys* de ativos em set reduzidos, a função ChemPLP realiza os cálculos de forma mais rápida (cerca de 190 encaixes em 23 segundos). A função GoldScore demora 90 segundos para realizar a mesma quantidade de cálculos (LIEBESCHUETZ, 2012). Assim, é preferível utilizar a função ChemPLP, dado o ganho de desempenho nos cálculos.

Já em relação ao tamanho da área de busca, observou-se uma pequena vantagem para os protocolos com a maior área de busca 8 Å, principalmente considerando a função ChemPLP. Tal fato pode ser justificado pela maior liberdade de encaixe no sítio, buscando pontos adicionais de interação, como já discutido no item 4.6.

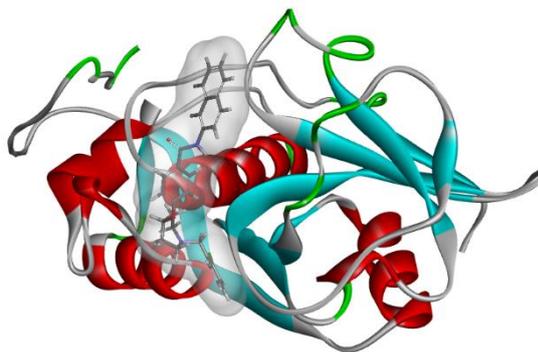
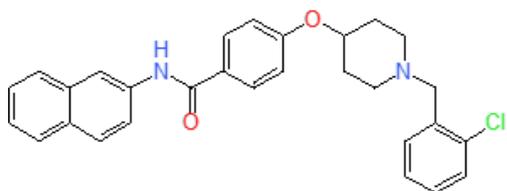
4.9 Triagem virtual da biblioteca virtual e seleção dos compostos

A partir do protocolo escolhido (função de pontuação ChemPLP, par His162/Cys25 ionizado e área de busca 8 Å) foi realizada a triagem virtual da biblioteca de análogos do MB15 criada. Os valores de score obtidos neste docking, também foram normalizados pelo mesmo critério utilizado na validação (fórmula no item 4.7). Apesar do melhor fator de enriquecimento corresponder a fração de 1%, optou-se por selecionar 2% do total de análogos devido a quantidade reduzida de compostos nesta lista (570 compostos no total) a fim de se selecionar ao menos 12 compostos.

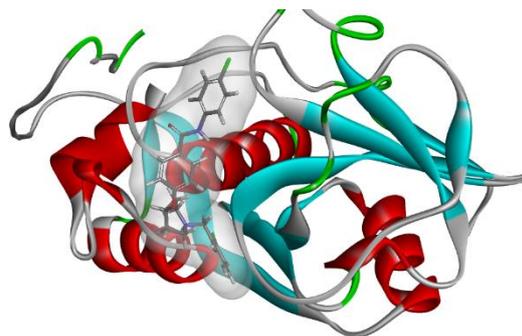
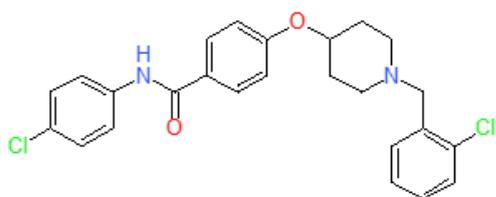
Após a aplicação do protocolo, os compostos com os melhores ranqueados foram analisados. Tal análise baseou-se na posição dos análogos no sítio de ligação da estrutura cristalográfica da 3KKU. Observou-se que as moléculas da biblioteca que apresentaram os melhores scores ocupam posições similares no sítio de ligação (Figura 9B e Anexo B).

Figura 9 – Estrutura química e poses no sítio ativo da cruzaína encontradas pelo docking para os compostos selecionados (continua...)

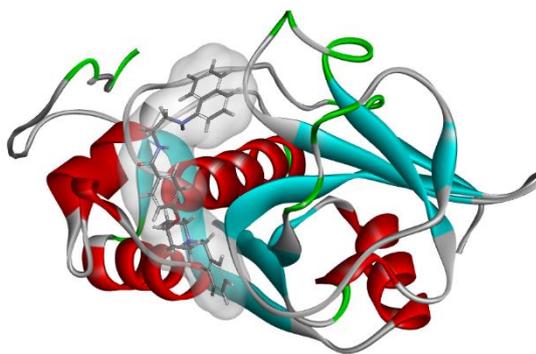
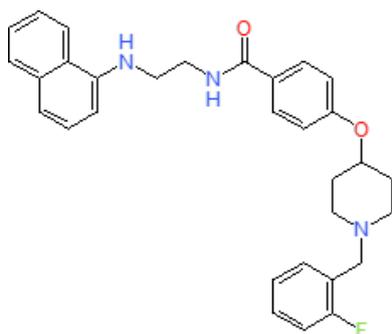
Composto 507



Composto 417



Composto 565



Composto 504

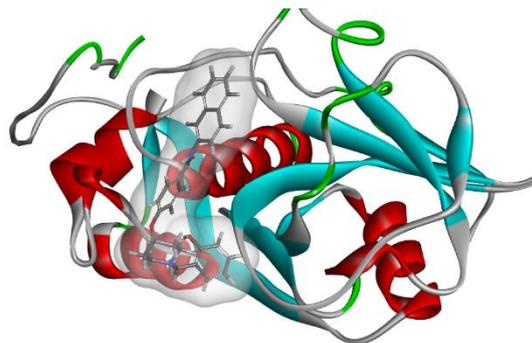
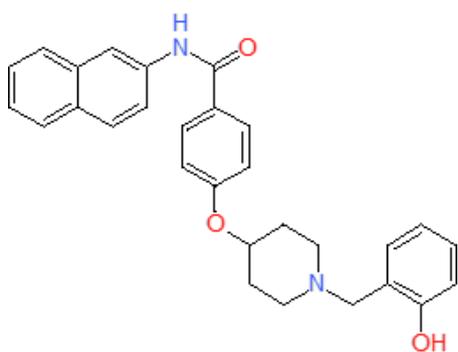
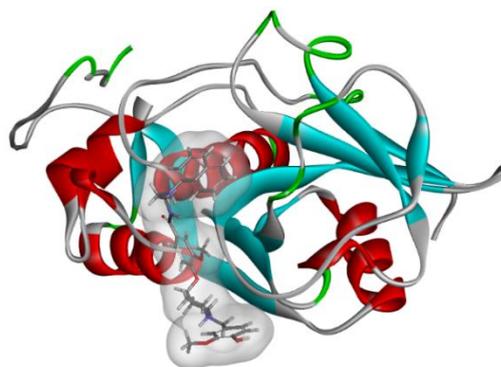
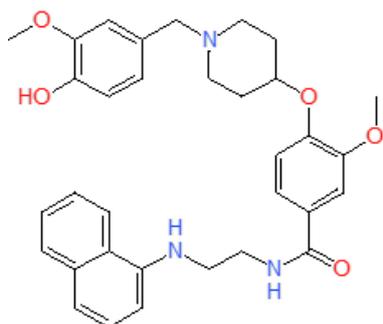
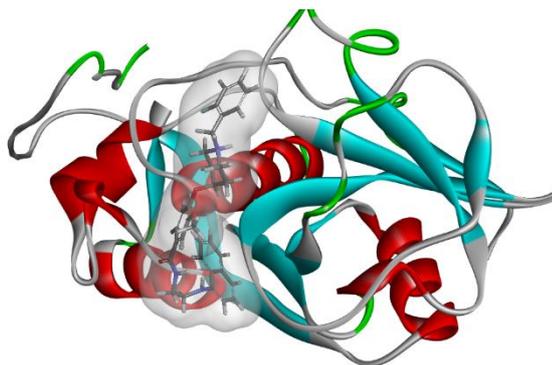
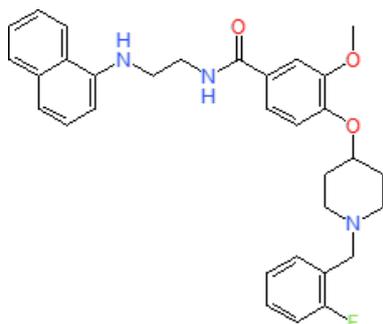


Figura 9 – Estrutura química e poses no sítio ativo da cruzaina encontradas pelo docking para os compostos selecionados (continua...)

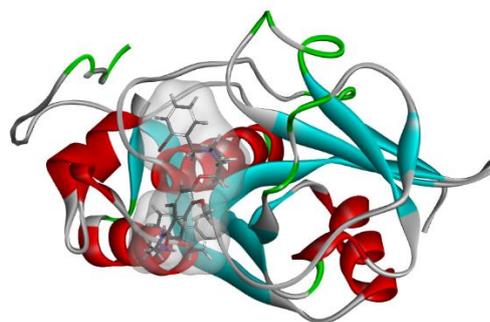
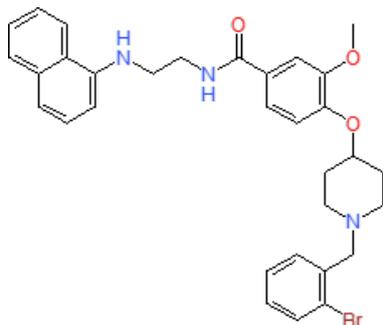
Composto 541



Composto 546



Composto 548



Composto 549

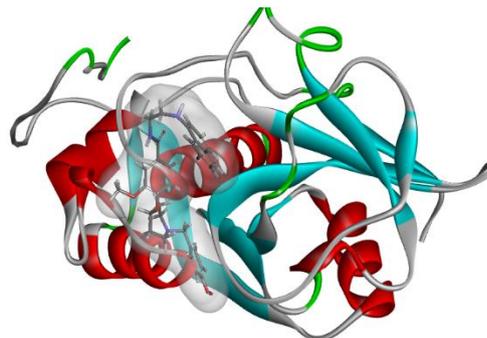
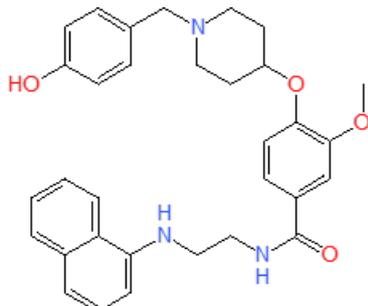
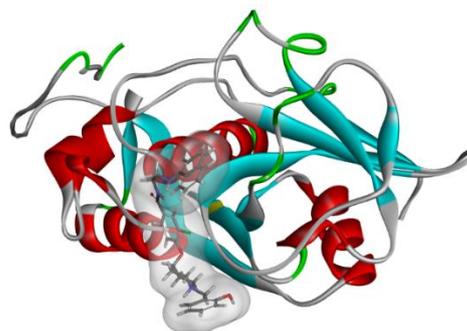
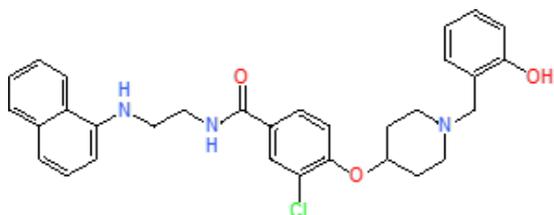
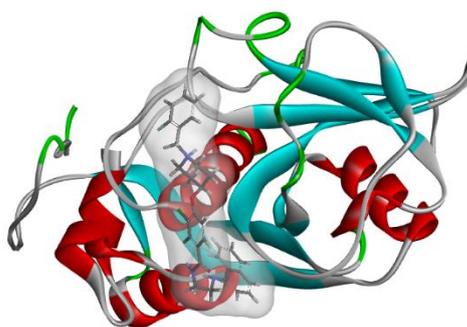
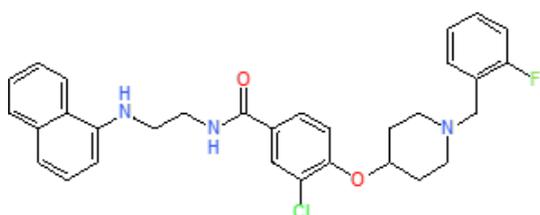


Figura 9 – Estrutura química e poses no sítio ativo da cruzaina encontradas pelo docking para os compostos selecionados (...conclusão)

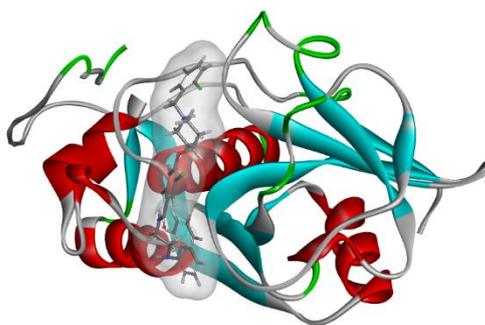
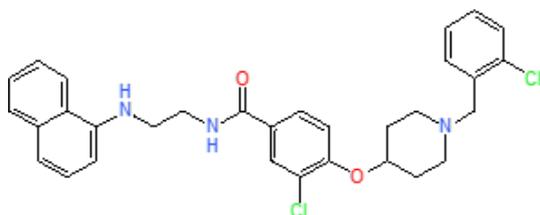
Composto 554



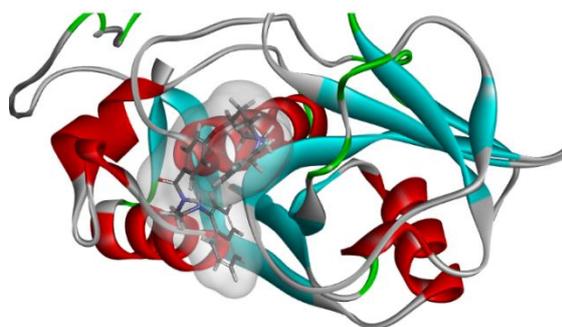
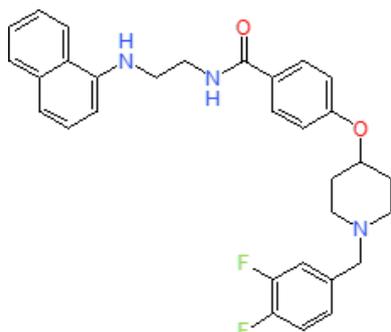
Composto 556



Composto 557



Composto 570

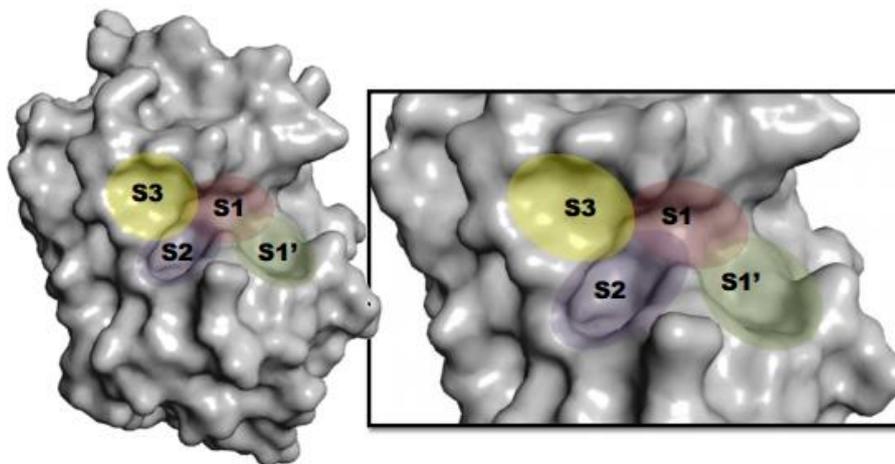


A cruzaina é mostrada em cartoon e os ligantes como sticks com a superfície destacando a posição dos ligantes no sítio de ligação

FONTE: AUTORIA PROPRIA

O sítio catalítico da cruzaina possui uma tríade catalítica composta pelos resíduos de Cys25, Hys162 e Asn175. Tal sítio pode ser subdividido ainda em subsítios, responsáveis por reconhecer determinados aminoácidos do substrato peptídico a ser clivado. São eles: S1, S1', S2 e S3 sendo que em cada um deles observa-se resíduos de aminoácidos, distintos entre si, que definem a especificidade da cruzaina (SOUZA, 2012).

Figura 10 – Subsítios da cruzaina



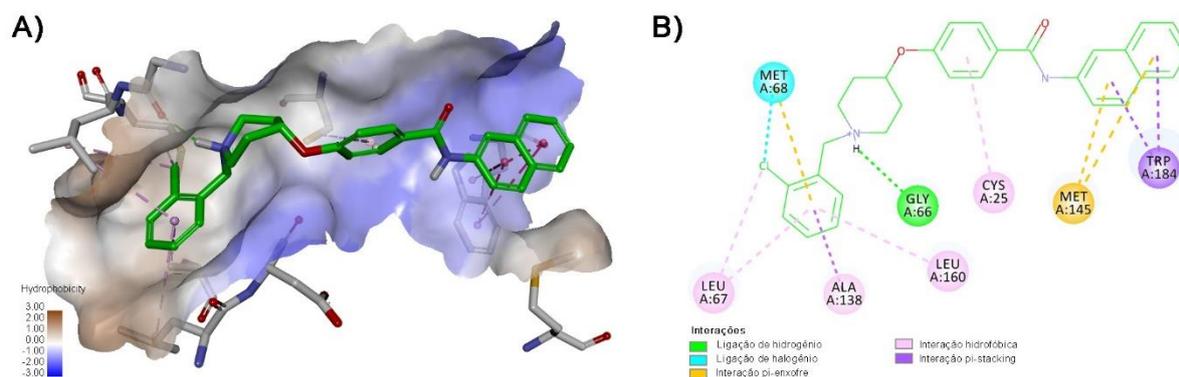
FONTE: SOUZA. (2012, p. 31)

Neste trabalho observou-se que as moléculas mais bem ranqueadas na triagem virtual são aquelas capazes de ocupar eficientemente estes subsítios e estabelecer interações com os resíduos de aminoácidos ali presentes (SOUZA, 2012). Dentre estes subsítios, é de grande relevância a interação no subsítio S2, principal responsável pela especificidade da cruzaina (MARTINS, 2016). Esse subsítio é constituído pelos resíduos de aminoácido Met68, Ala133, Leu157 e Gly160, conferindo seletividade para compostos hidrofóbicos (SOUZA, 2012). Observou-se que os anéis aromáticos dos compostos selecionados ocupam esse subsítio, destacando-se a importante de interação neste resíduo entre halogênios presentes e o resíduo Met68 (Figura 11).

Baseando-se nas informações prévias em relação aos subsítios presentes na cruzaina e na característica do MB15 de apresentar vários pontos passíveis de modificações estruturais, realizar a triagem virtual por *docking* molecular com os análogos deste composto é uma estratégia interessante, já que inúmeros compostos podem ser avaliados simultaneamente e, a partir dos resultados, interações com pontos importantes da enzima, como ilustrado na figura

abaixo com o análogo 507 no sítio de ligação podem ser observados, direcionando o desenho de potenciais inibidores.

Figura 11: Interações entre resíduos de aminoácidos e o ligante 507



Interações intermoleculares entre composto 507 e cruzaina em 3D (A) e diagrama 2D (B). Em (A) uma superfície acessível ao solvente foi construída em torno dos resíduos do sítio e colorida de acordo com a hidrofobicidade do resíduo. As linhas tracejadas representam as interações intermoleculares observadas. Em verde, ligações de hidrogênio, em rosa as interações hidrofóbicas do tipo Carbono–Carbono e Sistema Pi–Carbono em azul as ligações de halogênio, em laranja interações entre Sistema Pi– Enxofre e em roxo Pi-*stacking*. Átomos expostos ao solvente estão demarcados com um círculo azulado.

FONTE: AUTORIA PROPRIA

Na Figura 11 podem ser observados pontos importantes de interação do ligante análogo do MB15 507 e a cruzaina. Essas interações foram observadas em grande parte dos compostos selecionados da biblioteca criada. Destaca-se a presença de um halogênio em posição favorável no subsítio S2 para realizar interações com resíduo Met68 e nitrogênio protonado da amina interagindo por ligação de hidrogênio com Gly66. Além de serem pontos em comum entre os análogos com as melhores pontuações, essas interações também são observadas no ligante cristalográfico B95 (Figura 7).

5. CONCLUSÃO

Neste trabalho, após a seleção de compostos ativos contra a cruzaina na literatura, criou-se uma biblioteca virtual com 248 ativos e 13.133 *decoys* derivados desses. Diferentes protocolos de triagem foram analisados com as métricas fator de enriquecimento e área sob a curva ROC (AUC) e, apesar dos valores da AUC estarem fora do esperado, o protocolo que utilizou a função de pontuação ChemPLP, área de busca de 8 Å e o par His162/Cys25 ionizado, apresentou capacidade preditiva quatro vezes superior em relação a uma seleção aleatória de compostos, quando considera-se o Top 1% das moléculas ranqueadas. Com isso, este protocolo foi aplicado a uma nova biblioteca, gerada a partir de análogos do MB15, estrutura promissora na busca por novos fármacos com atividade tripanocida e gerada baseando-se nos reagentes presentes no laboratório de química medicinal da UFOP. Após aplicação deste protocolo nesta nova biblioteca, os 12 compostos melhores ranqueados foram selecionados para priorização de síntese em uma etapa futura. Entre esses compostos, observou-se que:

- Os melhores compostos ranqueados apresentaram conformações similares no sítio de ação, similar a uma letra “J”.
- A presença de halogênios no anel aromático, ocupando o subsítio S2 demonstrou-se importante, pois, está presente em alguns dos melhores compostos ranqueados. Tal importância pode ser justificada pela interação estabelecida entre halogênio do ligante e resíduo Met68, observada também no ligante cristalográfico B95.
- A presença de uma amina ionizada também se demonstrou importante, pela possibilidade de interação do tipo ligação de hidrogênio com resíduo Gly66
- Os anéis aromáticos são importantes para seletividade dos compostos, já que são capazes de ocupar o subsítio hidrofóbico S2, de grande importância.

Tais conclusões podem ser utilizadas, ainda, no desenho de novas estruturas visando otimizar a interação com a cruzaina e com o intuito de acelerar a busca por novos fármacos para o tratamento da DC.

REFERÊNCIAS BIBLIOGRÁFICAS

- BAG, A; GHORAI, P.K (2016). Development of Quantum Chemical Method to Calculate Half Maximal Inhibitory Concentration (IC50). *Molecular Informatics*, 35(5), 2016.
- BRASIL, Ministério da Saúde. Consenso Brasileiro em doença de Chagas. **Revista da Sociedade Brasileira de Medicina Tropical**, 38, SUPPL. 3, 2005.
- BROZELL, et.al. Evaluation of DOCK 6 as a pose generation and database enrichment tool. **J Comput Aided Mol Des**. 26(6), 749-73, 2012.
- CAFFREY, C. R., *et al.* Active site mapping, biochemical properties and subcellular localization of rhodesain, the major cysteine protease of *Trypanosoma brucei rhodesiense*. **Molecular and Biochemical Parasitology**. 118, 1, 61-73, 2001.
- CARVALHO. I. *et.al.* Introdução a modelagem molecular de fármacos no curso experimental de química farmacêutica. **Quím. Nova, São Paulo**. 26(3), 428-438, 2003.
- CONITEC; MINISTÉRIO DA SAÚDE. **Protocolo Clínico e Diretrizes Terapêuticas Doença de Chagas**. p. 1-144, 2018. Disponível em: <http://conitec.gov.br/images/Protocolos/Relatorio_PCDT_Doenca_de_Chagas.pdf>
- COURA. J.C. Tripanosomose, doença de chagas. **Cienc. Cult. vol.55 no.1 São Paulo**, Jan./Mar 2003. Disponível em: <http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252003000100022&lng=en&nrm=iso>. ISSN 0009-6725.
- DIAS, J.C.P; et al. II Consenso Brasileiro em Doença de Chagas, 2015 **Epidemiol. Serv. Saúde** . 25, n.esp, 2016.
- DIAS, J.C.P. Globalização, iniquidade e doença de Chagas. **Cad. Saúde Pública, Rio de Janeiro**, 23 Sup 1:S13-S22, 2007.
- ECHEVERRIA, L.E; MORILLO, C.A. American Trypanosomiasis (Chagas Disease). **Infectious Disease Clinics of North America**, 33(1), 2019.
- ENGEL, J. C. et al. Cysteine protease inhibitors cure an experimental *Trypanosoma cruzi* infection. **Journal of Experimental Medicine**. 188(4), 725-734, 1998.
- EUZÉBIO. D.M. **Novos casos da doença de chagas em uma área rural no município de Tobias Barreto – Sergipe**. Tese – mestrado. Universidade Federal de Sergipe – Núcleo de pós-graduação em biologia parasitária, 2015. Disponível em: <https://ri.ufs.br/jspui/handle/riufs/3249>
- FEDERICO B.L **Planejamento e validação anti-proliferativa e anti-leishmania, de novos híbridos tri-funcionalizados unidos através do anel 1,2,3-triazol e compostos similares**. Tese- doutorado. Universidade de São Paulo, 2016.
- FEINSTEIN W.P, BRYLINSK M. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. **Journal of ChemInformatics**. 15, 7-18, 2015.
- FERREIRA, *et.al.* Molecular docking and structure-based drug design strategies, **Molecules**. 20, 13384-13421, 2015.

- FORLI, S, Charting a path to success in virtual screening, **Molecules**. 20, 18732–18758, 2015.
- GILMOR S. A; CRAICK C.S; FILETTERICK R.J; Structural determinants of specificity in the cysteine protease cruzain. **Protein Science**. 6(8), 1603–1611, 1997.
- GRAVES, A.P; BRENK, R; SHOICHET, B.K. Decoys for Docking. **Journal of Medicinal Chemistry**. 48(11), 2005.
- JONES.G, *et al.* Development and Validation of a Genetic Algorithm for Flexible Docking, **J. Mol. Biol.** 267, 727–748, 1997.
- KIRCHMAIR, J. *et al.* Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection - What can we learn from earlier mistakes? **Journal of Computer-Aided Molecular Design**. 22(3–4), 213–228, 2008.
- LEAL, P.S; OLIVEIRA, F.N. Transmissão vertical da doença de Chagas em gestantes. **Enfermagem Brasil**, novembro/dezembro, 2019.
- LIEBESCHUETZ, J.W; COLE,J.C; KORB,O. Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. **Journal of Computer-Aided Molecular Design**. 26(6), 737-748, 2012.
- LINDOSO, J.A.L.; LINDOSO, A. A. B. P. Doenças tropicais negligenciadas no Brasil. **Revista do Instituto de Medicina Tropical de São Paulo**. 51(5), 247-253, 2009.
- MARTINS, L.C. **Síntese, ensaios *in vitro* e estudos computacionais de aminoquinolinas possíveis inibidoras da cruzaina e rodesaina**. Tese (mestrado em Ciências Farmacêuticas) – Faculdade de Farmácia, Universidade Federal de Minas Gerais, 2016.
- MEHRTENS, J. M. **The design, synthesis and biological assay of cysteine protease specific inhibitors**, Tese (Doutorado): University of Canterbury, 2007.
- MENDEZ; *et.al.* ChEMBL: towards direct deposition of bioassay data, **Nucleic Acids Research**, 47(D1), 930–940, 2018.
- MOTT.B. T, *et al.* Identification and optimization of inhibitors of trypanosomal cysteine proteases: Cruzain, rhodesain, and TbCatB. **Journal of Medicinal Chemistry**. 53(1), 52–60, 2010.
- MYSINGER *et.al.* Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. **J. Med. Chem.** 55(14), 6582–6594, 2012.
- NEUDERT G; KLEBE G. A knowledge-bases scoring function assement of protein-ligant complexes. **Journal of Chemical and Modeling**. 51(10), 2731-2745, 2011.
- O’BOYEL, N.M *et.al.* Open Babel: An open chemical toolbox. **Journal of Cheminformatics**. 3(33), 2011.
- PEREIRA, G. A. N. *et al.* Discovery and characterization of trypanocidal cysteine protease inhibitors from the ‘malaria box’. **European Journal of Medicinal Chemistry**, 179, 765-778 2019.

PEREIRA, J.C. **Melhoramento de docking-based virtual screening usando abordagem de deep learning**. Tese (Doutorado)-Fundação Oswaldo Cruz, Instituto Oswaldo Cruz, Rio de Janeiro, 2017.

RABELLO, M.M. **Desenvolvimento e automação de metodologias in silico para o estudo de complexos de inclusão utilizados na inovação terapêutica**. Tese (Pós-graduação em Inovação terapêutica) – Centro de biociências, Universidade Federal de Pernambuco, 2016.

RANGEL, K.C. **Planejamento e Avaliação de fragmentos moleculares como inibidores da enzima cruzaina**. Tese (Mestrado) – Centro de Ciências Exatas e de Tecnologia, Departamento De Química, Universidade Federal de São Carlos, 2015.

RODRIGUES, *et.al.* Estratégias de Triagem Virtual no Planejamento de Fármacos. **Revista Virtual de Química**. 4(6), 2012.

SANTOS, V.L.A; GONSALVES, A.A; ARAUJO, C.R.M. Abordagem didática para o desenvolvimento de moléculas bioativas: regra dos cinco de Lipinski e preparação de heterocíclico 1,3,4-oxadiazol em forno de micro-ondas doméstico. **Química Nova**. 41(1), 110-115, 2018.

SILVA, D.G. **Planejamento, síntese e avaliação de inibidores da enzima cruzaina e de agentes tripanossomidas derivados de imidazopiridina**. Tese (doutorado) – Instituto de química de São Carlos, Universidade de São Paulo, 2017.

SOUZA, M.L. Identificação de novos inibidores da enzima cruzaina de *Trypanossoma cruzi* candidatos a fármaco contra doença de chagas. **Tese (Mestrado) – Instituto de Física de São Carlos – USP, 2012**.

SOUZA, T. B. DE *et al.* New Eugenol Glucoside-based Derivative Shows Fungistatic and Fungicidal Activity against Opportunistic *Candida glabrata*. **Chemical Biology and Drug Design**. 87(1), 83–90, 2016.

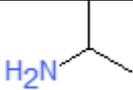
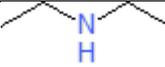
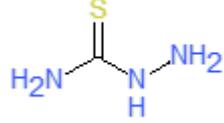
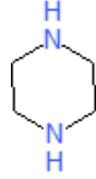
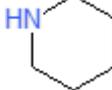
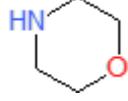
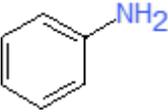
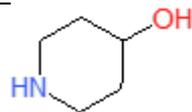
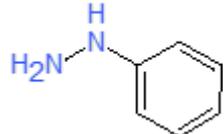
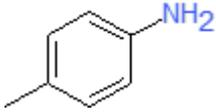
SCARIM, C. B. *et al.* Current advances in drug discovery for Chagas disease. **European Journal of Medicinal Chemistry**, 155, 824–838, 2018.

TARTAROTTI, E; *et.al.* Problemática vetorial da Doença de Chagas. **Arq. Ciênc. Saúde**. 11(1), 44-47, 2004.

VERDONK, *et.al.* Improved Protein–Ligand Docking Using GOLD, **Proteins Struct. Funct. Genet**. 52, 609–623, 2003.

ANEXO A – Compostos utilizados para obtenção dos análogos do MB15

Tabela A.1: Aminas disponíveis no almoxarifado do laboratório de Química Medicinal e Bioensaios da Escola de Farmácia da Universidade Federal de Ouro Preto.

NOME	ESTRUTURA
Dimetilamina	
Isopropilamina	
Dietilamina	
2-metoxietilamina	
Tiossemicarbazida	
Piperazina	
Piperidina	
Morfolina	
Anilina	
4-hidroxipiperidina	
Fenilhidrazinha	
<i>p</i> -toluidina	

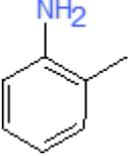
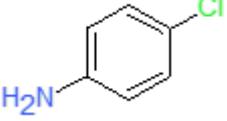
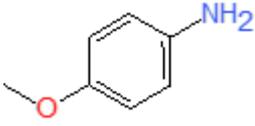
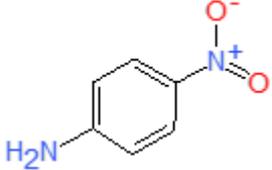
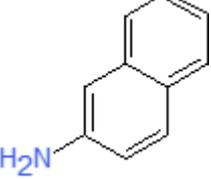
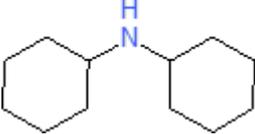
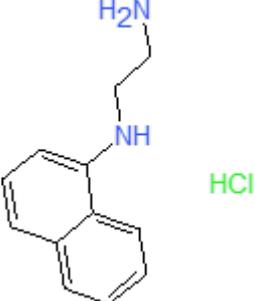
2-metilanilina	
4-cloroanilina	
<i>p</i> -anisidina	
4-nitroanilina	
2-naftilamina	
diciclohexilamina	
Cloridrato de <i>N</i> -(1-naftil)etilenoidiamina	

Tabela A.2: Aldeídos disponíveis no almoxarifado do laboratório de Química Medicinal e Bioensaios da Escola de Farmácia da Universidade Federal de Ouro Preto.

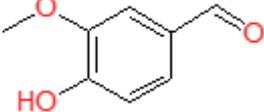
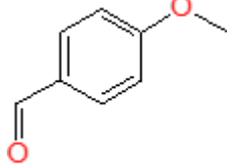
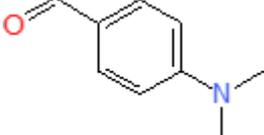
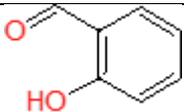
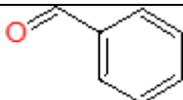
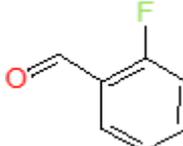
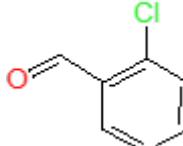
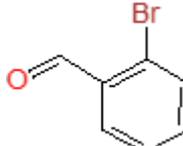
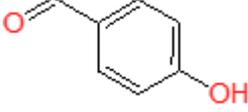
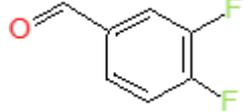
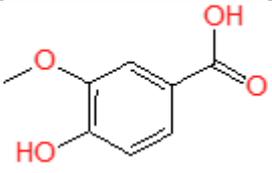
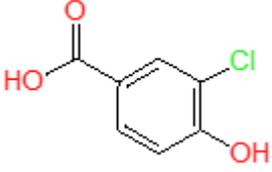
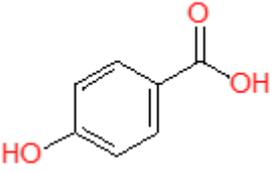
NOME	ESTRUTURA
Vanilina	
4-metoxibenzaldeído (anisaldeído)	
4-dimetilamino-benzaldeído	
Salicilaldeído	
Benzaldeído	
2-fluorobenzaldeído	
2-clorobenzaldeído	
2-bromobenzaldeído	
4-hidroxibenzaldeído	
3,4-difluorobenzaldeído	

Tabela A.3: Ácidos carboxílicos disponíveis no almoxarifado do laboratório de Química Medicinal e Bioensaios da Escola de Farmácia da Universidade Federal de Ouro Preto.

NOME	ESTRUTURA
Ácido da vanilina	 <chem>COc1ccc(O)cc1C(=O)O</chem>
Ácido 3-cloro-4-hidroxibenzóico	 <chem>O=C(O)c1cc(O)c(Cl)cc1</chem>
Ácido 4-hidroxibenzóico	 <chem>O=C(O)c1ccc(O)cc1</chem>