

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GENÉTICA E BIOQUÍMICA
CURSO DE BIOTECNOLOGIA

Virtual screening aplicado na busca de potenciais inibidores da proteína lanosterol 14-alfa demetilase (CYP51) em *Trypanosoma cruzi*

Asaph Souza Diniz

Monografia apresentada à Coordenação do Curso de Biotecnologia, da Universidade Federal de Uberlândia, para obtenção do grau de Bacharel em Biotecnologia.

Uberlândia – MG

Dezembro – 2017

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GENÉTICA E BIOQUÍMICA
CURSO DE BIOTECNOLOGIA

Virtual screening aplicado na busca de potenciais inibidores da proteína lanosterol 14-alfa demetilase (CYP51) em *Trypanosoma cruzi*

Asaph Souza Diniz

Nilson Nicolau Junior

Monografia apresentada à Coordenação do Curso de Biotecnologia, da Universidade Federal de Uberlândia, para obtenção do grau de Bacharel em Biotecnologia.

Uberlândia – MG

Dezembro – 2017

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GENÉTICA E BIOQUÍMICA
CURSO DE BIOTECNOLOGIA

Virtual screening aplicado na busca de potenciais inibidores da proteína lanosterol 14-alfa demetilase (CYP51) em *Trypanosoma cruzi*

Asaph Souza Diniz

Nilson Nicolau Junior

Homologado pela coordenação do Curso
de Biotecnologia em __/__/__

Edgar Silveira Campos

Uberlândia – MG

Dezembro - 2017
UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GENÉTICA E BIOQUÍMICA
CURSO DE BIOTECNOLOGIA

Virtual screening aplicado na busca de potenciais inibidores da proteína lanosterol 14-alfa demetilase (CYP51) em *Trypanosoma cruzi*

Asaph Souza Diniz

Aprovado pela Banca Examinadora em: / / Nota: _____

Uberlândia, 20 de dezembro de 2017

RESUMO

A Doença de Chagas é, atualmente, um problema de saúde pública que afeta de 6 a 7 milhões de pessoas em 21 países apenas na América Latina, segundo a Organização Mundial da Saúde (OMS). Os tratamentos existentes não são específicos para o parasita, de modo que os efeitos colaterais gerados por tais medicamentos trazem um grande desconforto aos pacientes, e aliado ao fato de que o tratamento da fase crônica não é efetivo, se faz necessário a busca por novos compostos com maior especificidade e melhores características farmacocinéticas. Uma das formas de atingir tais resultados se dá pela técnica de *Virtual Screening*, utilizando a enzima lanosterol 14-alfa demetilase (CYP51) de *Trypanosoma cruzi*, um alvo validado na literatura. Utilizando o pacote de programas da OpenEye Scientific Software, Inc, e o banco de dados Drugs Now do ZINC, foi gerado uma *Hit list* dos 500 melhores compostos obtidos via *virtual screening*, sendo os melhores compostos selecionados de acordo com suas características de ADMET (absorção, distribuição, metabolismo, excreção e toxicidade) geradas a partir do *web server* pkCSM, e analisadas pelo *software* Orange3.

LISTA DE SIGLAS E ABREVIATURAS

ADMET	Absorção, Distribuição, Metabolismo, Excreção E Toxicidade
AUC	Área Sobre a Curva
CADD	Design de Drogas Auxiliado por Computadores
CYP	Superfamília P450
CYP51	Lanosterol 14-Alfa Demetilase
DC	Doença De Chagas
DM	Docagem Molecular
HTS	High Throughput Screening
NMR	Ressonância Magnética Nuclear
OMS	Organização Mundial Da Saúde
PCA	Principal Component Analysis
PDB	Protein Data Bank
PGP	Glicoproteína-P
ROC	Receiver Operating Characteristic
SNC	Sistema Nervoso Central
VS	Virtual Screening

SUMÁRIO

1. INTRODUÇÃO	1
2. OBJETIVOS	6
3. MATERIAIS E MÉTODO	7
4. RESULTADOS E DISCUSSÃO	11
5. CONCLUSÃO	23
6. REFERÊNCIAS	24

1. INTRODUÇÃO

A Doença de Chagas (DC) é uma parasitose caracterizada pela ação do agente etiológico *Trypanosoma cruzi*. Está na lista da Organização Mundial da Saúde como uma das doenças tropicais negligenciadas, afetando cerca de 21 países, e 8 milhões de pessoas só na América Latina. Mais de 10.000 pessoas morrem por ano por causa da doença, e cerca de 25 milhões de pessoas estão em risco de adquirirem-na. É considerado um problema de saúde mundial, com casos relatados em outros continentes (WHO , 2017).

Um dos seus vetores é o *Triatoma infestans*, inseto da família *Reduviidae*. É comumente chamado de barbeiro, e por ser um hematófago, a transmissão para o vetor ocorre quando ele se alimenta de sangue de outros mamíferos contaminados. Neste caso, o mamífero contaminado contém em sua corrente sanguínea formas tripomastigotas do parasito, e quando o *T. infestans* ingere o sangue contaminado, as células se diferenciam em amastigotas no intestino médio do vetor. Após isso, as células amastigotas se diferenciam, novamente, em células epimastigotas, células essas, que juntamente com as amastigotas, são replicativas. As células epimastigotas se ligam à parede cerosa do intestino médio do inseto por meio de interações hidrofóbicas, e se diferenciam em formas metacíclicas, estas por sua vez, são excretadas pelo *T. infestans* enquanto se alimenta, assim, por meio da ferida, as células metacíclicas caem na corrente sanguínea da presa, e reiniciam o ciclo de vida do protozoário (TYLER; ENGMAN , 2001).

Atualmente, a DC é tratada com quimioterapia por meio da utilização dos medicamentos nifurtimox e benznidazol, porém, ambos apresentam diversos efeitos colaterais, incluindo efeitos gastrointestinais e neuronais (CASTRO et al. , 2006). A doença pode ser curada durante a fase inicial da infecção, sendo de extrema importância o diagnóstico precoce para que o tratamento seja efetivo, mas já que a fase aguda da doença apresenta sintomas muito parecidos com infecções comuns, grande parte dos hospedeiros não sabe que contraíram a doença,

diminuindo a eficiência dos tratamentos existentes. Tais medicamentos atingiram cerca de 76% (16 de 21 pacientes) de cura durante a fase aguda da doença, porém ela é uma fase muito breve em relação à crônica, e a eficácia do tratamento é prejudicada nesse último período, apresentando cerca de 8% (9 de 113) de pacientes com testes sorológicos constantemente negativos (CANÇADO , 2002).

A Busca de potenciais alvos para doenças negligenciadas já é uma realidade, e no caso da proteína lanosterol 14-alfa demetilase (CYP51), diversos tratamentos estão sendo desenvolvidos contra os organismos dos mais diversos reinos, gêneros e espécies. (CHOI et al. , 2014). Essa proteína é da superfamília P450 (CYP), família essa responsável pelas proteínas mais antigas, estando presente em quase todos os reinos biológicos. No tocante aos protozoários, a CYP51 está atrelada à rota metabólica do esqualeno, sendo crucial para a formação de ergosterol, o equivalente metabólico para o colesterol em vertebrados. É um importante componente da membrana celular, e sua falta induz a formação de poros na mesma, causando o desbalanço das concentrações dos íons, e consequentemente a morte da célula (KINSKY , 1970). Desta forma, dada sua importância na sobrevivência do protozoário, pode-se afirmar que a proteína CYP51 é um importante e promissor alvo biológico.

O alinhamento global de proteínas humana (Uniprot ID: Q16850) e do parasito (Uniprot ID: M1KK91), utilizando o software EMBOSS needle (RICE et al. , 2000), mostra uma similaridade de 30,7% dos aminoácidos, apontando uma baixa chance de que haja inibidores comuns, corroborando com o fato de que há diferença significativa entre as duas proteínas para que possa haver o desenvolvimento de drogas seletivas.

Hoje sabe-se que para a comercialização de uma nova droga no mercado, centenas de milhares de outras moléculas foram analisadas e descartadas durante o processo de desenvolvimento, e isso se converte em tempo e dinheiro desperdiçados, seja no isolamento ou

sintetização das moléculas, ou nos próprios testes de bancada. Entre os anos 1960 e 1980, o custo para produzir uma nova droga passou de 4 milhões, para 350 milhões de dólares, desta forma, fez-se necessário o desenvolvimento de novas técnicas e ferramentas para atingir os resultados esperados (OOMS , 2000).

O design de drogas auxiliado por computadores (CADD) foi uma dessas técnicas muito utilizada no design racional de drogas, com o intuito de se reduzir o tempo requerido, e aumentar as taxas de eficiência e eficácia dos testes pré-clínicos. Tal método tira proveito do poder computacional para calcular a similaridade de ligantes, simular interações entre estruturas moleculares, entre outras possibilidades. Para tal, foi necessário o desenvolvimento de novas formas de se analisar os dados obtidos através de técnicas empíricas, como a Ressonância Magnética Nuclear (NMR) e a Cristalografia de Raios X, com o intuito de caracterizar e agrupar as moléculas conhecidas.

De acordo com Wermuth et al. (1998) "Texto traduzido", tem-se que farmacóforo, ou modelo farmacofórico, é "O conjunto de características estéricas e eletrônicas que são necessárias para garantir a interação supramolecular ótima com um alvo biológico, e acionar - ou bloquear - sua resposta biológica.", ou em outras palavras, um modelo farmacofórico é uma entidade virtual, que contém informações a respeito de uma determinada molécula, ou conjunto de moléculas, de forma que interações hidrofóbicas, doação e acepção de elétrons, grupos formadores de ligação de hidrogênio, e sítios de interação eletrostática têm suas coordenadas tridimensionais anotadas, permitindo o reconhecimento molecular do alvo por esse pequeno conjunto de informações.

A modelagem de farmacóforos se dá por duas vias principais, a modelagem baseada no ligante e a baseada na estrutura. Na primeira, as informações farmacofóricas dos ligantes conhecidos na literatura são extraídas e usadas como um molde para comparar as informações

com outras moléculas, e determinar sua similaridade. Já a segunda se baseia nas informações da estrutura do alvo biológico conhecidas, de modo que um molde negativo é formado de acordo com as características farmacofóricas do sítio de ligação do alvo. Neste caso, o conhecimento tanto da estrutura, como da exata localização do sítio de ligação tem papéis fundamentais no processo de CADD.

É dado o nome *Virtual Screening* (VS), ou Triagem Virtual, ao processo de comparação de modelos farmacofóricos distintos. Tal técnica recebeu sua denominação por ser uma derivação da técnica de *High Throughput Screening* (HTS), onde diversos compostos diferentes são testados empiricamente contra o mesmo alvo biológico afim de se predizer o composto com o maior coeficiente de inibição, ou ativação dependendo do resultado esperado. A VS se utiliza do poder computacional para calcular as distâncias tridimensionais entre os pontos farmacofóricos do modelo e compara-los com os das moléculas analisadas, de modo que as mais variadas funções de pontuação podem ser usadas para gerar o valor final de cada análise.

A docagem molecular (DM) pode ser usada como uma forma de triagem, porém, apenas baseada no volume do ligante. De acordo com KITCHEN et al. , (2004) o processo da DM envolve a predição da conformação – caracterizada pela liberdade do ligante das ligações da molécula – e orientação do ligante – chamada de pose – dentro da caixa de ligação selecionada do alvo, e através de uma função de pontuação que calcula as interações dos ligantes com o sítio da proteína, os ligantes são ranqueados, variando para cada software. No caso do software FRED presente no pacote OEDOCKING 3.2.0.2: OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>, os métodos são: busca exaustiva e otimização. Na primeira parte, o software calcula todas as poses da molécula dentro da cavidade do alvo, e elimina aquelas que colidem com as paredes dela, calcula a pontuação em relação às ligações de hidrogênio formadas entre o ligante e a proteína, entre o ligante e solventes implícitos, como a água, entre interações metal-quelante, e a forma do ligante, que calcula as distancias dos raios

de van der Waals dos átomos pesados do ligante em relação aos da proteína, considerando apenas os que a soma dos raios dos átomos da proteína e do ligante estiverem entre a proporção de 1,25 e 2,5 vezes. Já a segunda utiliza uma versão alterada da função de pontuação, de modo que os termos de ligação de hidrogênio e interação metal-quelante são melhorados.

2. OBJETIVOS

2.1. Objetivos gerais

Construir um modelo farmacofórico visando a busca por potenciais fármacos, via triagem virtual, contra a CYP51 de *T. cruzi*.

2.2. Objetivos específicos

- Gerar e validar um modelo farmacofórico baseado na forma de estruturas conhecidas extraídas do banco de dados do PDB em complexo com a proteína CYP51 de *T. cruzi*;
- Realizar uma busca baseada na forma e nas cores do modelo, utilizando uma biblioteca de ligantes;
- Estudar a interação entre o receptor, CYP51, e os compostos selecionados na triagem, por meio da técnica de ancoragem molecular;
- Avaliar as propriedades farmacocinéticas *in silico* dos compostos selecionados.
- Sugerir ligantes com potencial de inibição *in vivo*.

3. MATERIAIS E MÉTODO

3.1. Geração do modelo farmacofórico

O modelo farmacofórico foi gerado a partir de inibidores conhecidos na literatura, que compartilhavam características favoráveis em relação à inibição do crescimento do parasita. Tais inibidores foram extraídos do banco de dados do Protein Data Bank (PDB) (BERMAN et al. , 2000), de modo que foram escolhidas 5 moléculas, cujos códigos de identificação são J5Y, T9H, WVH, 26N e TW5. As moléculas foram alinhadas de acordo com seus pontos farmacofóricos com a utilização do *web server* PharmaGist (PIRES et al. , 2015), e o melhor resultado baseado no score final gerado foi submetido ao software vROCS 3.2.0.4 , OpenEye Scientific Software, Santa Fe, NM (HAWKINS et al. , 2007), onde os pontos farmacofóricos desnecessários foram excluídos, restando apenas 8.

3.2. Validação do modelo

A validação do modelo foi feita utilizando o software vROCS 3.2.1.4 Software, Santa Fe, NM (HAWKINS et al. , 2010), de modo que se utilizou a área sobre a curva (AUC) da curva *Receiver Operating Characteristic* (ROC) como característica avaliada, onde valores $\geq 0,8$ foram considerados aceitáveis.

Para tal, foram preparados dois bancos de dados, o primeiro, chamado de ativos, foi preparado contendo moléculas descritas na literatura com atividade inibitória contra o alvo, utilizando 18 inibidores da CYP51 de *T. cruzi*, extraídos de complexos proteína-ligante depositados no PDB, cujos códigos de identificação estão presentes na tabela 1, e 18 moléculas descritas na literatura que pertencem ao grupo dos azoles, cujos códigos de identificação do ZINC estão presentes na tabela 1, totalizando 36 moléculas distintas; já o segundo, chamado de *decoys*, ou seja, moléculas inativas, foram obtidas ao se submeter o banco de ativos no *web*

server DUD-E (MYSINGER et al. , 2012), de modo que 2600 moléculas distintas foram geradas.

3.3. Preparação do banco de dados

O banco de dados utilizado foi o ZINC versão 12 (IRWIN; SHOICHET , 2005), *subset* Drugs-Now de 24/11/2014, baixado em 20/09/2017, que contem 10.639.555 moléculas distintas. O banco foi tratado pelos softwares FILTER 2.5.1.4, QUACPAC 1.7.0.2, e OMEGA 2.5.1.4 (HAWKINS et al. , 2010) Software, Santa Fe, NM.<http://www.eyesopen.com>. que visam remover compostos que não estão de acordo com a regra dos 5 de Lipinski, fixar o estado de protonação no pH fisiológico (7,4) e gerar isômeros por tautomeria, e gerar no máximo 200 conformações biologicamente ativas, quando possível, para cada composto respectivamente.

3.4. Triagem virtual

A triagem foi feita utilizando o software ROCS 3.2.1.4 Software, Santa Fe, NM (HAWKINS et al. , 2010), e o modelo validado, empregando os parâmetros *Tanimoto combo*, *color force field*: Implicit Mills-Dean, e máximo de 500 compostos na *Hit list*. Os dados foram visualizados utilizando o software VIDA 4.3.0 Software, Santa Fe, NM.

3.5. Ancoragem molecular

Primeiramente foi confeccionado o sítio ativo do alvo. Para isso, a estrutura cristalina da proteína, cujo código de identificação no PDB era 4C0C, foi submetida no módulo MAKE_RECEPTOR do pacote OEDOCKING 3.2.0.2: OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>. Para a ancoragem molecular foi utilizado o módulo FRED (MCGANN , 2011) do mesmo pacote, de modo que a classificação se deu pelo algoritmo Chemgauss4. Os 500 melhores com a adição das moléculas utilizadas para a confecção do modelo foram submetidos à ancoragem, totalizando 505 moléculas. Para a visualização e seleção foi utilizado o módulo DOCKING_REPORT do mesmo pacote. Os ligantes que

obtiveram até 99% de similaridade com o ligante J5Y, que obtivera o melhor resultado, foram selecionados para a etapa seguinte, totalizando 6 moléculas.

3.6. Análise das propriedades farmacocinéticas

As moléculas selecionadas, em conjunto com as que compunham o modelo foram submetidas ao *web server* PkCSM (PIRES et al. , 2015), onde as características de absorção, distribuição, metabolismo, excreção e toxicidade foram preditas.

Para a análise dos dados, foi utilizado o software Orange3 (DEMŠAR et al. , 2013). Os dados qualitativos foram discretizados, assumindo os valores 1 para “No”, e 2 para “Yes”. Após isso, os valores foram normalizados pela média, subsequentemente foram utilizados os modelos estatísticos *Principal Component Analysis* (PCA) e K-Means, onde no primeiro, foram utilizados apenas os dois principais componentes, e no segundo foram utilizados, baseado na pontuação da silhueta, três grupos.

4. RESULTADOS E DISCUSSÃO

Conforme descrito por HARGROVE et al., (2013), tem-se que a inibição de citocromos P450 é governada por dois fatores: primeiro e mais importante, coordenação de um átomo básico com o ferro heme, que é responsável pela diminuição do potencial da enzima em ser reduzida; e segundo, múltiplas interações fracas da parcela não coordenada com a cavidade da enzima. Desta forma, os pontos farmacofóricos escolhidos no modelo seguiram tais características, como pode ser visto na figura 1

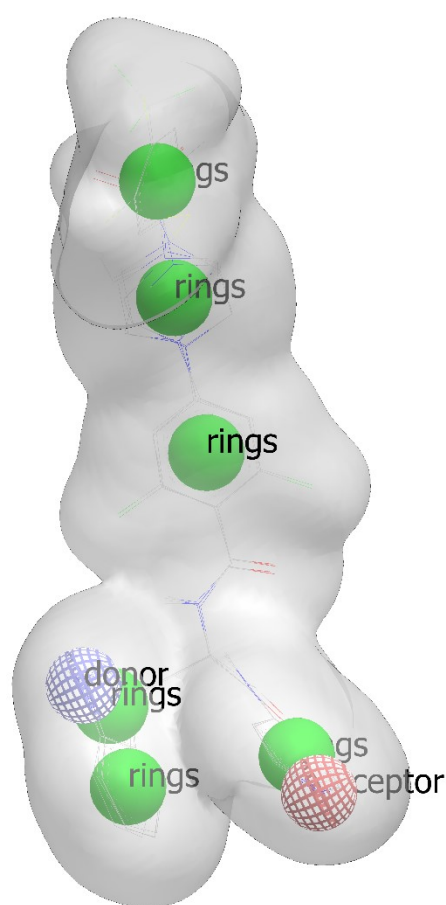


Figura 1. Modelo farmacofórico. Figura representando o modelo gerado. Os pontos em verde indicam pontos farmacofóricos de anéis, em vermelho, os pontos farmacofóricos de aceptores de elétrons, e em azul, pontos farmacofóricos de doadores de elétrons. Em cinza, está representado a forma (volume) do modelo.

A validação do modelo deu a certeza da qualidade do modelo utilizado, com uma pontuação de AUC de 0,842, superior ao limite estabelecido de 0,8. Isso demonstra a probabilidade de o modelo ser realmente uma molécula ativa, ao invés de ser um falso positivo, categorizando-o como um resultado aleatório. A figura 2 mostra a curva obtida.

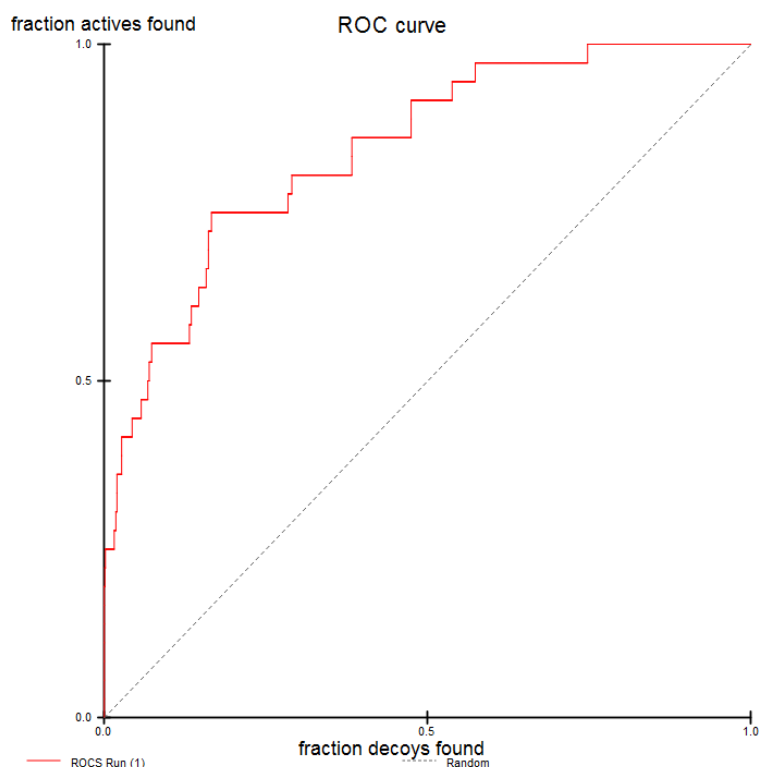


Figura 2. Curva AUC. Gráfico da curva gerada a partir da análise de compostos ativos (actives), e inativos (decoys). Quanto mais afastada da linha pontilhada de “Random”, menor a probabilidade de ser um modelo aleatório, e quanto maior a área sob a curva (AUC), maior a probabilidade de ser um modelo similar a moléculas ativas.

A triagem resultou nos 500 compostos que obtiveram maior pontuação tanto quanto à similaridade estrutural (volume), como à proximidade espacial dos pontos farmacofóricos (cores) em relação ao modelo.

É possível observar a presença do grupo indol, caracterizado pela presença de um anel benzênico fundido a um anel pirrol, em todas as moléculas mostradas, fruto dos pontos

farmacofóricos escolhidos no modelo. A figura 3 mostra os 6 primeiros compostos.

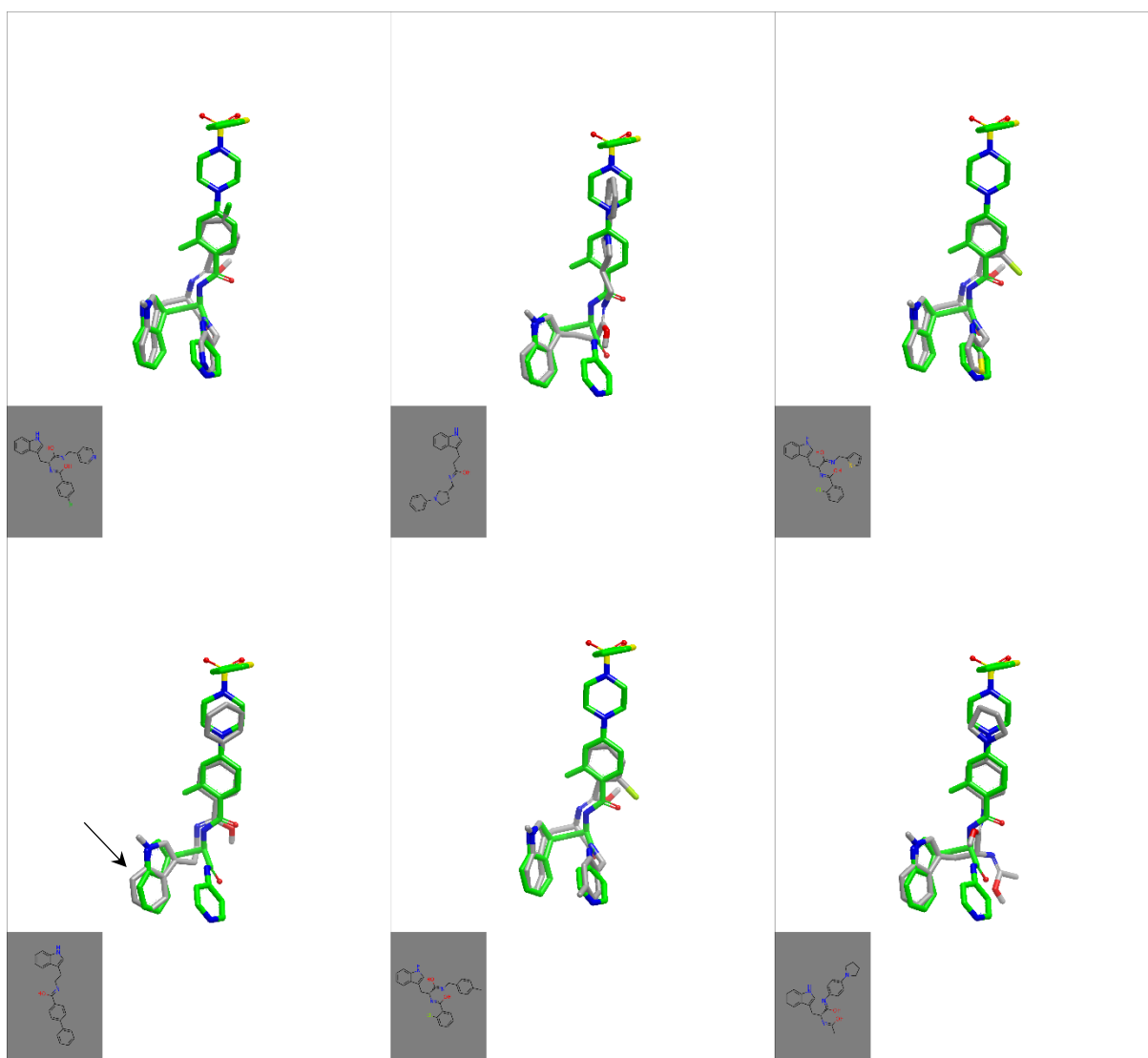


Figura 3. Triagem virtual. A figura mostra a sobreposição das 6 melhores moléculas, representadas pela cor cinza, com a T9H, representada em verde, uma das moléculas que compunha o modelo. Em cinza a representação bidimensional da molécula. A seta indica o grupo indol.

A ancoragem molecular não apresentou falhas, sendo que todas as 505 moléculas foram ancoradas com sucesso. Porém foi observado que, em geral, os compostos interagem com o sítio de ligação da enzima de maneira aleatória, não alinhando o grupo indol como no caso de todas as moléculas componentes do modelo. Isso se deve principalmente ao fato de não se utilizar o ferro heme como restrição, de modo que ele seria, como descrito por CHOI; PODUST;

ROUSH, 2014; HARGROVE et al., 2013, um elemento essencial na coordenação e correto alinhamento das moléculas na enzima.

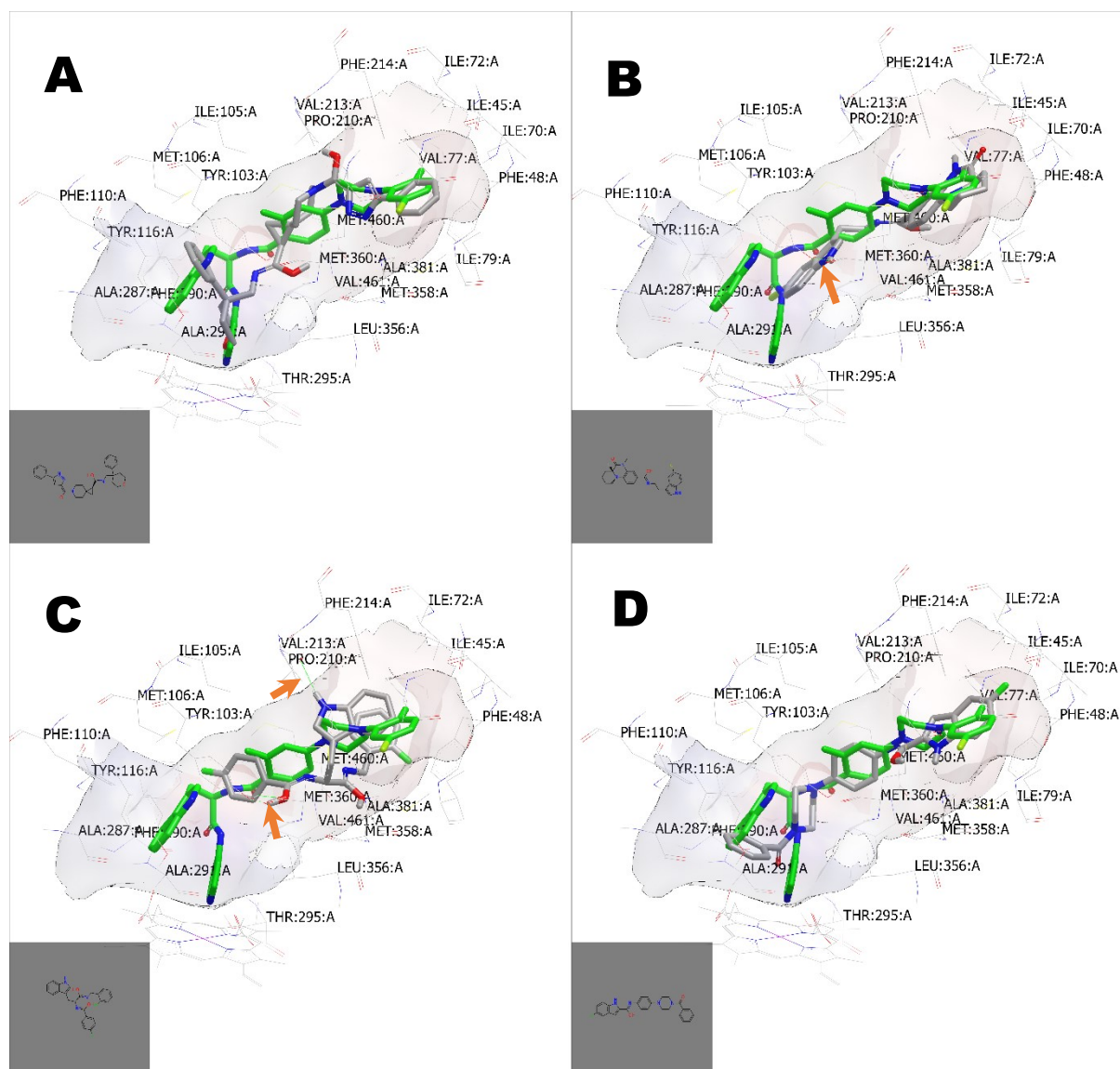


Figura 4. Ancoragem molecular. A figura mostra a sobreposição da ancoragem da J5Y (uma das moléculas que compunha o modelo) em verde, e as quatro primeiras moléculas. Em “A” está representada a molécula ZINC11784865_6_22, e pode-se observar que apenas ela não contém o grupo indol. Em “B” está representada a molécula ZINC21938563_2_168, e observa-se que o grupo indol está no local esperado. Em “C” está representada a molécula ZINC14202273_4_134, e pode-se observar um exemplo de ancoragem diferente do esperado, onde os dois braços menores da molécula triada estão no local onde braço maior da molécula modelo está. Em “D” está representada a molécula ZINC09429860_3_125, e é possível notar a presença do grupo indol no local oposto ao esperado, partindo da suposição de que o modelo

estaria no local correto. As setas indicam ligações de hidrogênio do ligante com resíduos da proteína.

Os resultados obtidos pelo *web server* pkCSM foram divididos em seis tabelas. Na tabela 1, estão presentes as predições das características físicas de cada molécula. Pôde-se observar que há uma divisão bem clara entre as moléculas que compunham o modelo e as triadas, em relação ao peso molecular. Isso se deve ao fato de que o subset DRUGS NOW contém moléculas selecionadas a partir da regra dos cinco de Lipinski, onde o peso molecular não deve exceder 500 kDa. Além disso, observa-se que a área superficial e o peso molecular têm uma forte correlação entre si (cerca de 0,995).

Uma vez que o parasita, durante a fase crônica da doença, se encontra no interior das células, faz-se necessário alta lipofilicidade por parte do composto, para que este possa penetrar pela membrana celular humana, e atingir seu alvo. Desta forma, valores de solubilidade em água pequenos são essenciais. Para avaliar tal característica, é utilizada o coeficiente de partição, que é a razão da concentração do composto diluído em fase aquosa, e da concentração do composto diluído em fase oleosa, geralmente octanol. Valores maiores que 0 indicam compostos mais lipofílicos, e valores menores que 0 indicam compostos mais hidrofílicos. Os resultados mostram que todos os compostos testados apresentam alta lipofilicidade, maior inclusive que a do exano que é em torno de 4.

Tabela 1: Propriedades moleculares

Nomes	MW	LP	RB	AP	DN	SA
<i>26N</i>	630,646	6,027	8,000	5,000	3,000	261,596
<i>J5Y</i>	611,121	5,970	8,000	5,000	3,000	259,403
<i>T9H</i>	632,743	4,254	9,000	7,000	3,000	257,136
<i>TW5</i>	597,094	5,662	8,000	5,000	3,000	253,038
<i>WVH</i>	598,629	5,287	8,000	5,000	3,000	251,066
<i>ZINC09429860_3_125</i>	442,494	4,906	4,000	3,000	2,000	190,157
<i>ZINC11784865_6_22</i>	498,627	6,027	6,000	6,000	2,000	217,459
<i>ZINC13723309_2_199</i>	498,970	5,262	5,000	3,000	2,000	212,180
<i>ZINC14202273_4_134</i>	433,458	5,519	7,000	2,000	3,000	182,971
<i>ZINC21938563_2_168</i>	436,943	4,704	4,000	3,000	2,000	185,691
<i>ZINC34928371_1_14</i>	352,437	5,033	2,000	1,000	1,000	157,812

**WM significa peso molecular, LP coeficiente de partição, RB ligações com rotação livre, AP aceptores, DN doadores, SA área superficial.*

Na tabela 2 estão expressos os valores referentes à predição das características de absorção das moléculas. Após a normalização, é possível observar uma forte correlação negativa entre o coeficiente de partição e a solubilidade em água (cerca de -0,912), corroborando com a afirmação anterior. Os resultados da permeabilidade em células de linhagem Caco-2, um modelo para absorção intestinal, não foram positivos para todas as substâncias. Pôde-se observar que para as moléculas T9H, TW5, ZINC09429860_3_125, ZINC11784865_6_22, e ZINC34928371_1_14 foram obtidos valores de alta permeabilidade (maior que 0,9), mas o restante permaneceu na zona de permeabilidade mediana. A absorção intestinal foi predita positiva para todas as moléculas, ficando todas acima dos 30%.

Todas as moléculas foram classificadas como potenciais substratos da Glicoproteína-P (PGP), um importante transportador de membrana que expulsa toxinas e xenobióticos da célula, porém, todas as moléculas também foram preditas como inibidoras das PGP I e PGP II, de modo que a interferência na entrada e permanência na célula é provavelmente pouco significativa.

Tabela 2: Propriedades de absorção

Nomes	WS	Caco2 P	IA	SP	PgpS	Pgp1I	Pgp2I
26N	-3,841	0,865	94,872	-2,735	Yes	Yes	Yes
J5Y	-4,329	0,842	94,229	-2,735	Yes	Yes	Yes
T9H	-3,879	0,955	85,048	-2,735	Yes	Yes	Yes
TW5	-3,921	0,903	93,744	-2,735	Yes	Yes	Yes
WVH	-3,886	0,839	90,961	-2,735	Yes	Yes	Yes
ZINC09429860_3_125	-3,815	1,109	89,232	-2,737	Yes	Yes	Yes
ZINC11784865_6_22	-5,455	0,909	90,423	-2,866	Yes	Yes	Yes
ZINC13723309_2_199	-4,418	0,767	90,056	-2,737	Yes	Yes	Yes
ZINC14202273_4_134	-4,103	0,770	91,630	-2,735	Yes	Yes	Yes
ZINC21938563_2_168	-4,851	0,710	92,303	-2,835	Yes	Yes	Yes
ZINC34928371_1_14	-5,191	0,921	92,224	-2,735	Yes	Yes	Yes

**WS significa solubilidade em água, Caco2 P permeabilidade em células de linhagem Caco-2, IA absorção intestinal, SP permeabilidade pela pele, PgpS substrato da Glicoproteína P(PGP), Pgp1I inibidor da PGP I, Pgp2I inibidor da PGP II.*

À tabela 3 estão associados os valores preditos para as propriedades de distribuição dos

	Property 1	Property 2	Property 3	Property 4	Property 5
26N	Green	Yellow	Green	Green	Green
J5Y	Orange	Orange	Green	Green	Green
T9H	Green	Yellow	Red	Green	Green
TW5	Green	Yellow	Green	Green	Green
WVH	Green	Orange	Orange	Yellow	Green
ZINC09429860_3_125	Red	Yellow	Orange	Red	Green
ZINC11784865_6_22	Orange	Orange	Orange	Yellow	Green
ZINC13723309_2_199	Yellow	Orange	Yellow	Green	Green
ZINC14202273_4_134	Orange	Red	Yellow	Red	Green
ZINC21938563_2_168	Red	Yellow	Yellow	Green	Green
ZINC34928371_1_14	Orange	Yellow	Yellow	Green	Green

potenciais fármacos. Observa-se que a molécula ZINC21938563_2_168 apresentou um valor de volume de distribuição alto (maior que 0,45), o que se considera ideal para que a molécula se encontre dentro das células; as moléculas ZINC34928371_1_14, e WVH, apresentaram um valor abaixo de -0,15, valor esse considerado baixo, demonstrando a tendência dessas moléculas em permanecer nos vasos sanguíneos; as demais moléculas ficaram em condições

intermediárias. A fração desligada de todas as moléculas foi baixa, demonstrando uma possível dificuldade de atravessar a membrana celular, porém, um maior tempo de meia-vida no organismo, por permanecerem ligadas às proteínas do sangue.

Uma vez que o parasita não se multiplica no cérebro, medicamentos que atravessam a membrana hematoencefálica são desnecessários por uma série de questões que envolvem efeitos indesejados. Os valores da permeabilidade à membrana hematoencefálica mostram que todas as moléculas que compunham o modelo são pobremente distribuídas no cérebro, mostrando que poucos efeitos indesejados são esperados, e evidenciando seu avanço no seu desenvolvimento como droga utilizável. Já as moléculas triadas apresentam valores intermediários, ficando entre -1 e 0,3 que são os limites entre pobremente distribuído e altamente distribuído no cérebro respectivamente. Apenas as moléculas ZINC34928371_1_14 e ZINC09429860_3_125 apresentaram um valor acima do normal, considerado facilmente distribuído no cérebro. Já para a permeabilidade ao SNC observa-se uma média menor das moléculas componentes do modelo em relação às triadas, porém todas com exceção da WVH e T9H são capazes de penetrar no sistema nervoso central (SNC), o que não é o ideal

Tabela 3: Propriedades de distribuição

<i>Nomes</i>	<i>VDss</i>	<i>FU</i>	<i>BBBp</i>	<i>CNSp</i>
<i>26N</i>	0,019	0,000	-1,051	-1,916
<i>J5Y</i>	0,096	0,000	-1,002	-2,069
<i>T9H</i>	-0,103	0,000	-1,300	-3,160
<i>TW5</i>	0,015	0,000	-0,912	-1,959
<i>WVH</i>	-0,162	0,000	-1,153	-2,821
<i>ZINC09429860_3_125</i>	0,318	0,000	-0,829	-1,761
<i>ZINC11784865_6_22</i>	0,168	0,000	-0,927	-1,877
<i>ZINC13723309_2_199</i>	-0,097	0,000	-0,838	-1,800
<i>ZINC14202273_4_134</i>	-0,052	0,018	-1,212	-1,803
<i>ZINC21938563_2_168</i>	0,464	0,001	-0,068	-1,815
<i>ZINC34928371_1_14</i>	-0,408	0,000	0,426	-1,517

**VDss significa volume de distribuição aparente, FU fração livre, BBBp permeabilidade à barreira hematoencefálica, CNSp permeabilidade ao SNC.*

Pode -se observar que na tabela 4, referente às predições sobre as propriedades de metabolização, todas as moléculas que compunham o modelo, além da ZINC11784865_6_22 não se mostraram serem substratos da CYP2D6, uma importante enzima da superfamília P450 presente em altas quantidades no fígado, mas também na massa cinzenta. Todas as moléculas se mostraram substratos da enzima CYP3A4, também pertencente à superfamília P450, muito envolvida na ativação de pró-fármacos. Sua ação é extremamente necessária, uma vez que o tempo de meia-vida das moléculas é afetado pela sua ação, e sua toxicidade está diretamente atrelada à atuação das duas enzimas.

Desta forma, a inibição dessas enzimas é característica de fundamental entendimento para se determinar a farmacocinética e possíveis interações medicamentosas. Os resultados obtidos mostram um padrão seguido pelas moléculas componentes do modelo, que não foi completamente traduzido nas moléculas triadas.

Tabela 4: Propriedades de metabolização

<i>Nomes</i>	<i>2D6s</i>	<i>3A4s</i>	<i>1A2i</i>	<i>2C19i</i>	<i>2C9i</i>	<i>2D6i</i>	<i>3A4i</i>
<i>26N</i>	No	Yes	No	Yes	Yes	No	Yes
<i>J5Y</i>	No	Yes	No	Yes	Yes	No	Yes
<i>T9H</i>	No	Yes	No	Yes	Yes	No	Yes
<i>TW5</i>	No	Yes	No	Yes	Yes	No	Yes
<i>WVH</i>	No	Yes	No	Yes	Yes	No	Yes
<i>ZINC09429860_3_125</i>	Yes	Yes	Yes	Yes	Yes	No	Yes
<i>ZINC11784865_6_22</i>	No	Yes	No	No	No	No	Yes
<i>ZINC13723309_2_199</i>	Yes	Yes	No	Yes	Yes	No	Yes
<i>ZINC14202273_4_134</i>	Yes	Yes	Yes	Yes	Yes	No	Yes
<i>ZINC21938563_2_168</i>	Yes	Yes	Yes	Yes	Yes	No	Yes
<i>ZINC34928371_1_14</i>	Yes	Yes	Yes	Yes	Yes	No	No

**2D6s significa substrato da CYP2D6, 3A4s substrato da CYP3A4, 1A2i inibidor da CYP1A2, 2C19i inibidor da CYP2C19, 2C9i inibidor da CYP2C9, 2D6i inibidor da CYP2D6, 3A4i inibidor da CYP3A4.*

A tabela 5 contém os resultados referentes às predições sobre as propriedades de excreção de cada molécula. A proteína Transportadora de Cátions Orgânicos II (OCT2 do inglês) desempenha um importante papel na eliminação de cátions exógenos e endógenos pela

urina, desta forma, ser potencial substrato deste transportador interfere no tempo de meia-vida do fármaco, e indiretamente na sua toxicidade. Nenhum dos fármacos se mostrou provável substrato da OCT2 renal.

A depuração total expressa em log(ml/min/kg) demonstra a rapidez em que a droga é completamente eliminada do corpo, desta forma, valores maiores significam eliminação mais

Tabela 5. Propriedades de excreção

<i>Nomes</i>	<i>TC</i>	<i>R-OCT2s</i>
<i>26N</i>	0,548	No
<i>J5Y</i>	-0,215	No
<i>T9H</i>	-0,088	No
<i>TW5</i>	-0,226	No
<i>WVH</i>	0,003	No
<i>ZINC09429860_3_125</i>	0,274	No
<i>ZINC11784865_6_22</i>	0,535	No
<i>ZINC13723309_2_199</i>	-0,172	No
<i>ZINC14202273_4_134</i>	0,686	No
<i>ZINC21938563_2_168</i>	0,021	No
<i>ZINC34928371_1_14</i>	0,599	No

**TC significa depuração total, R-OCT2s substrato da OCT2 renal.*

rápida.

Por fim, a tabela 6 contém os dados referentes à predição das propriedades tóxicas dos compostos. O teste AMES resumidamente dá, para um certo composto, sua probabilidade de causar mutações no DNA de bactérias, podendo atuar como um possível carcinógeno. Apenas as moléculas ZINC14202273_4_134, e ZINC34928371_1_14 foram preditas como possíveis agentes mutagênicos.

A máxima dose tolerada é uma estimativa do limite da dose máxima. Expressa em Log(mg/kg/dia), valores iguais ou abaixo de 0,477 são considerados baixos, ou seja, a dose necessária para causar toxicidade é baixa; já o restante dos valores é considerado alto, ou seja, a dose necessária para causar toxicidade é alta.

Os canais de potássio codificados pelo gene hERG estão presentes no miocárdio, e desempenham um papel fundamental no funcionamento do mesmo. Sua inativação, ou por inibidores, ou por mutações genéticas, pode gerar uma desordem potencialmente fatal chamada Síndrome do QT longo. Logo, o “*anti-targeting*” de drogas que causem a inibição de tais canais é essencial para no desenvolvimento de drogas. Todas as moléculas obtiveram resultados idênticos, sendo classificadas como inibidoras da hERG II, porém não inibidoras da hERG I.

A toxicidade aguda oral em ratos é o provável valor de LD₅₀ em ratos de um dado composto, ou seja, a quantidade, em mol/kg da molécula, necessária para matar 50% dos animais testados. Todas as moléculas apresentaram valores altos, onde, a título de conhecimento, o menor valor ZINC14202273_4_134, de 904,63 g/kg é muito superior ao da aspirina, de 200 mg/kg. A toxicidade crônica oral em ratos consiste na determinação do LOEL, ou seja, a menor dose de um dado composto que cause um efeito adverso observável. Novamente os compostos apresentaram valores altos.

A hepatotoxicidade é um fator muito considerado no desenvolvimento de drogas, uma vez que injúrias no tecido hepático causadas pelo uso crônico de fármacos são intensificadas com a duração do tratamento, desta forma, compostos com pouca ou nenhuma hepatotoxicidade são ideais no tratamento a longo prazo. Os únicos compostos com hepatotoxicidade predita como negativa foram o ZINC14202273_4_134, e ZINC34928371_1_14.

A sensibilização da pele é um efeito adverso de análise muito importante ao se considerar o desenvolvimento de drogas administradas pela pele, porém, não é um parâmetro analisado para este caso, uma vez que esta via de administração não é visada.

A toxicidade em *T. pyriformis* é um teste de IGC₅₀, ou seja, inibição de 50% do crescimento do protozoário, onde valores maiores que -0,5 são considerados tóxicos, desta forma, todos os compostos são considerados tóxicos neste parâmetro.

Tabela 6: Propriedades de toxicidade

<i>Nomes</i>	<i>AMESt</i>	<i>MTD</i>	<i>hERG1I</i>	<i>hERG2I</i>	<i>ORAT</i>	<i>ORCT</i>	<i>Hpt</i>	<i>SS</i>	<i>TPT</i>	<i>MT</i>
<i>26N</i>	No	0,117	No	Yes	2,792	0,949	Yes	No	0,285	0,305
<i>J5Y</i>	No	0,437	No	Yes	3,290	2,505	Yes	No	0,285	-1,442
<i>T9H</i>	No	0,185	No	Yes	2,414	2,049	Yes	No	0,285	0,481
<i>TW5</i>	No	0,087	No	Yes	2,812	1,020	Yes	No	0,285	0,104
<i>WVH</i>	No	0,103	No	Yes	2,769	2,061	Yes	No	0,285	0,664
<i>ZINC09429860_3_125</i>	No	0,015	No	Yes	2,686	2,021	Yes	No	0,300	1,063
<i>ZINC11784865_6_22</i>	No	-0,556	No	Yes	3,668	0,844	Yes	No	0,319	2,617
<i>ZINC13723309_2_199</i>	No	-0,160	No	Yes	2,345	2,376	Yes	No	0,289	-1,922
<i>ZINC14202273_4_134</i>	Yes	0,834	No	Yes	2,087	1,330	No	No	0,292	-0,923
<i>ZINC21938563_2_168</i>	No	-0,429	No	Yes	2,583	1,069	Yes	No	0,595	-1,541
<i>ZINC34928371_1_14</i>	Yes	0,262	No	Yes	2,265	0,131	No	No	0,308	0,836

**AMESt significa teste de AMES, MTD dose máxima tolerada, hERG1I inibidor da hERG I, hERG2I inibidor da hERG II, ORAT toxicidade aguda oral em ratos, ORCT toxicidade crônica oral em ratos, Hpt hepatotoxicidade, SS sensibilização da pele, TPT toxicidade em T. pyriformis, MT toxicidade em peixes.*

A toxicidade em peixes é um parâmetro que avalia a LC₅₀, ou seja, a menor dose capaz de matar 50% dos animais. Para ser considerado de toxicidade aguda, o valor deve ser inferior a -0,3, o que ocorreu para os compostos J5Y, ZINC21938563_2_168, ZINC14202273_4_134, e ZINC13723309_2_199.

Os resultados obtidos pelo método de PCA e K-Means estão representados a figura 5, e mostram o agrupamento dos dados. Para o PCA, foram escolhidos os quatro componentes principais que juntos representam 79% da variância dos dados, onde PC1 e PC2 foram escolhidos para representar os eixos do gráfico. Como pode-se observar, as moléculas ZINC13723309_2_199, ZINC09429860_3_125 e ZINC21938563_2_168 se enquadraram no mesmo grupo das moléculas componentes do modelo, evidenciando sua provável ação superior às moléculas que não faziam parte do grupo, por serem mais similares às que compunham o modelo. Porém, como observado, nenhuma das moléculas triadas se situou no 3º quadrante, local esse que provavelmente está relacionado a características favoráveis ao efeito



Figura 5. Gráfico do agrupamento. A figura mostra o resultado da redução do número de componentes, e posterior agrupamento. Observa-se que as moléculas que compunham o modelo ficaram bem próximas entre si, demonstrando sua similaridade farmacocinética, e além disso, se situaram no quadrante 3, indicando uma região comum de características favoráveis à ação antichagásica. Estão ocultos os dados da molécula T9H, cujas coordenadas eram -3,51 e -2,12.

antichagásico pelo fato de todas as moléculas componentes do modelo estarem presentes nesta região.

5. CONCLUSÃO

Como foi mostrado anteriormente, conclui-se que o desenvolvimento de drogas com a utilização da VS é, sem sombra de dúvidas, a melhor maneira de se alcançar resultados mais confiáveis, de maneira mais rápida e com menor custo.

O resultado da técnica de PCA mostrou três prováveis compostos com potencial antichagásico baseadas nas moléculas encontradas na literatura, ZINC13723309_2_199, ZINC09429860_3_125 e ZINC21938563_2_168, porém, tais compostos apresentam características farmacocinéticas que ainda precisam ser melhoradas, afim de que possam prosseguir com testes pré-clínicos.

Desta forma, como futuras perspectivas, espera-se realizar o método de dinâmica molecular afim de se calcular e avaliar a energia de interação entre os presentes ligantes, e realizar alterações estruturais afim de alcançar melhores características farmacocinéticas.

REFÊRENCIAS

- BERMAN, H. M.; WESTBROOK, J.; FENG, Z.; et al. **The protein data bank**. Nucleic acids research, v. 28, n. 1, p. 235–242, 2000.
- CANÇADO, J. R. **Long term evaluation of etiological treatment of Chagas disease with benznidazole**. Revista do Instituto de Medicina Tropical de Sao Paulo, v. 44, n. 1, p. 29–37, 2002.
- CASTRO, J. A.; DEMECCA, M. M.; BARTEL, L. C. **Toxic side effects of drugs used to treat Chagas' disease (American trypanosomiasis)**. Human & experimental toxicology, v. 25, n. 8, p. 471–479, 2006. Sage Publications Sage CA: Thousand Oaks, CA.
- CHOI, J. Y.; PODUST, L. M.; ROUSH, W. R. **Drug strategies targeting CYP51 in neglected tropical diseases**. Chemical Reviews, v. 114, n. 22, p. 11242–11271, 2014.
- DEMŠAR, J.; CURK, T.; ERJAVEC, A.; et al. **Orange: Data Mining Toolbox in Python**. Journal of Machine Learning Research, v. 14, p. 23492353, 2013.
- HARGROVE, T. Y.; WAWRZAK, Z.; ALEXANDER, P. W.; et al. **Complexes of trypanosoma cruzi Sterol 14 α -Demethylase (CYP51) with Two Pyridine-based Drug Candidates for Chagas Disease: Structural basis for pathogen selectivity**. Journal of Biological Chemistry, v. 288, n. 44, p. 31602–31615, 2013.
- HAWKINS, P. C. D.; SKILLMAN, A. G.; NICHOLLS, A. **Comparison of Shape-Matching and Docking as Virtual Screening Tools**. Journal of Medicinal Chemistry, v. 50, n. 1, p. 74–82, 2007. American Chemical Society.
- HAWKINS, P. C. D.; SKILLMAN, A. G.; WARREN, G. L.; et al. **Conformer generation with OMEGA: Algorithm and validation using high quality structures from the protein databank and cambridge structural database**. Journal of Chemical Information and Modeling, v. 50, n. 4, p. 572–584, 2010.
- IRWIN, J. J.; SHOICHET, B. K. **ZINC – A Free Database of Commercially Available Compounds for Virtual Screening**. J. Chem. Inf. Model, v. 45, n. December 2004, p. 177–182, 2005.
- KINSKY, S. C. **Antibiotic interaction with model membranes**. Annual review of pharmacology, v. 10, p. 119–42, 1970. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/9090114>>. .
- KITCHEN, D. B.; DECORNEZ, H.; FURR, J. R.; et al. **Docking and scoring in virtual screening for drug discovery: methods and applications**. Nature reviews. Drug discovery, v. 3, n. 11, p. 935–49, 2004.
- MCGANN, M. **FRED Pose Prediction and Virtual Screening Accuracy**. Journal of Chemical Information and Modeling, v. 51, n. 3, p. 578–596, 2011. American Chemical Society.
- MYSINGER, M. M.; CARCHIA, M.; IRWIN, J. J.; et al. **Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking**. Journal of Medicinal Chemistry, v. 55, n. 14, p. 6582–6594, 2012. American Chemical Society.
- OOMS, F. **Molecular Modeling and Computer Aided Drug Design. Examples of their Applications in Medicinal Chemistry**. Current Medicinal Chemistry, v. 7, n. 2, p. 141–158, 2000. Disponível em: <<http://www.eurekaselect.com/openurl/content.php?genre=article&issn=0929->

8673&volume=7&issue=2&spage=141>. .

PIRES, D. E. V.; BLUNDELL, T. L.; ASCHER, D. B. **pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures**. Journal of Medicinal Chemistry, v. 58, n. 9, p. 4066–4072, 2015.

RICE, P.; LONGDEN, I.; BLEASBY, A. **EMBOSS: the European Molecular Biology Open Software Suite**. Trends in genetics : TIG, v. 16, n. 6, p. 276—277, 2000. Disponível em: <[https://doi.org/10.1016/S0168-9525\(00\)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2)>. .

TYLER, K. M.; ENGMAN, D. M. **The life cycle of Trypanosoma cruzi revisited**. International Journal for Parasitology, v. 31, n. 5–6, p. 472–481, 2001.

WERMUTH, C.-G.; GANELLIN, C. R.; LINDBERG, P.; et al. **Chapter 36. Glossary of Terms Used in Medicinal Chemistry (IUPAC Recommendations 1997)**. In: J. A. B. B. T.-A. R. in M. Chemistry (Org.); . v. Volume 33, p.385–395, 1998. Academic Press.

WHO. **Chagas disease (American trypanosomiasis)**. Disponível em: <<http://www.who.int/mediacentre/factsheets/fs340/en/>>. Acesso em: 25/9/2017.