

UNIVERSIDADE FEDERAL DE ALAGOAS
ESCOLA DE ENFERMAGEM E FARMÁCIA
PROGRAMA DE PÓS GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS

EDJAN CARLOS DANTAS DA SILVA

**Utilização do QSAR-3D no planejamento racional de novos inibidores HDAC8
do *Schistosoma mansoni***

Maceió

2020

EDJAN CARLOS DANTAS DA SILVA

**Utilização do QSAR-3D no planejamento racional de novos inibidores HDAC8
do *Schistosoma mansoni***

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em ciências farmacêuticas do instituto de ciências farmacêuticas da Universidade Federal de Alagoas, como requisito para obtenção do grau de Mestre em ciências farmacêuticas.

Orientador: Prof. Dr. Thiago Mendonça de Aquino

Maceió

2020

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central

Bibliotecária: Taciana Sousa dos Santos – CRB-4 – 2062

S586u Silva, Edjan Carlos Dantas da.
Utilização do QSAR-3D no planejamento racional de novos inibidores
HDAC8 do *Schistosoma mansoni* / Edjan Carlos Dantas da Silva. – 2020.
95 f. : il. ; figs. ; tabs. color.

Orientador: Thiago Mendonça de Aquino.
Dissertação (Mestrado em Farmácia) – Universidade Federal de Alagoas.
Instituto de Ciências Farmacêuticas. Maceió, 2020.

Bibliografia: f. 66-75.
Anexos: f. 76-95.

1. Esquistossomose. 2. *Schistosoma mansoni*. 3. Fármacos. 4. Enzima histona
deacetilase 8 (smHDAC8). 5. QSAR-3D. 6. Modelagem molecular. I. Título.

CDU: 615.28: 54



Ata de defesa de dissertação

Aos **vinte e dois dias** do mês de junho de **2020**, às dezenove horas, reuniu-se via teleconferência, a banca examinadora composta pelos docentes doutores Thiago Mendonça de Aquino, Edeildo Ferreira da Silva Júnior e Francisco Jaime Bezerra Mendonça Júnior, para o exame de defesa da dissertação intitulada: “**Utilização do QSAR-3D no planejamento racional de novos inibidores HDAC8 do Schistosoma mansoni**” elaborada pelo mestrando **Edjan Carlos Dantas da Silva**, regularmente matriculado no Programa de Pós-Graduação em Ciências Farmacêuticas, Nível Mestrado, sob a orientação do **Prof. Dr. Thiago Mendonça de Aquino**, que presidiu os trabalhos. Após apresentação por **46 minutos**, o mestrando foi arguido pela banca. Em seguida, reunidos em sessão secreta às **20 horas e 45 minutos**, os examinadores consideram a dissertação (**X**) **APROVADA OU () APROVADA COM RESTRIÇÃO**, emitindo o seguinte parecer:
A banca examinadora recomenda a realização das correções sugeridas.

Nada mais havendo a tratar, a sessão foi encerrada às **21 horas** e eu, Daniel de Brito Ricarte, Secretário do PPGCF, Mestrado, lavrei a presente ata que segue assinada pela Banca Examinadora e pelo Mestrando.
Maceió, 22 de junho de 2020

Presidente (Orientador): Prof. Dr. Thiago Mendonça de Aquino

Examinador interno: Prof. Dr. Edeildo Ferreira da Silva Júnior

Examinador Externo: Prof. Dr. Francisco Jaime Bezerra Mendonça Júnior

Mestrando: Edjan Carlos Dantas da Silva

AGRADECIMENTOS

Agradeço a CAPES e FAPESP pelo apoio financeiro cedido para bolsa de mestrado.

RESUMO

Há algum tempo, órgãos de saúde em todo mundo, tem demonstrado grande preocupação com a prevalência da esquistossomose, bem como a resistência às terapias atuais. A esquistossomose trata-se de uma enfermidade incluída no grupo de doenças tropicais negligenciadas e é causada pelo verme *S. mansoni*. Para o desenvolvimento de novos compostos os quais visam combater este parasita, é de grande interesse o estudo da inibição de alvos essenciais ao seu ciclo. Dentre os diversos alvos estudados, a enzima histona deacetilase 8 do *Schistosoma mansoni* (*smHDAC8*) é uma proteína atraente para o planejamento de drogas esquistossomicidas, pela justificativa de que a *smHDAC8* é essencial para o parasita e não é presente no hospedeiro humano. Os derivados do ácido hidroxinâmico são potentes inibidores enzimáticos da *smHDAC8*. Assim, visando o planejamento racional de novos inibidores *smHDAC8*, foi realizado um estudo de QSAR para produzir um modelo capaz de prever a atividade biológica para esta classe de compostos. Todas estruturas foram desenhadas e otimizadas para etapa de docking. As conformações obtidas no docking, foram utilizadas para cálculo dos descritores e conseqüentemente, construção do modelo de QSAR-3D. Um modelo com 3 descritores de Lennard-Jones e 1 de Coulomb foi capaz de prever a atividade biológica para o conjunto de treinamento ($R^2 = 0.92$, RMSEC = 0,42). Além disso, foi capaz de prever o valor de IC_{50} para um subconjunto externo ($Q^2_{ext} = 0.85$) de amostras. A interpretação dos descritores levou a vários *insights* sobre a relação entre atividade e estrutura dos inibidores *smHDAC8* baseados em hidroxamatos. Dessa forma, novos compostos podem ser planejados e ter suas atividades biológicas previstas *in silico* pelo modelo de QSAR-3D.

Palavras chave: Esquistossomose; *smHDAC8*; QSAR; modelagem molecular

ABSTRACT

For some time, health agencies worldwide have shown great concern about the prevalence of schistosomiasis, as well as resistance to current therapies. Schistosomiasis is a disease included in the group of neglected tropical diseases and is caused by the worm *S. mansoni*. For the development of new compounds which aim to combat this parasite, it is of great interest to study the inhibition of essential targets in its cycle. Among the various targets studied, the histone deacetylase 8 enzyme from *S. mansoni* (*smHDAC8*) is an attractive protein for planning schistosomicidal drugs, justifying the fact that *smHDAC8* is essential for the parasite. Hydroxamic acid derivatives are potent enzyme inhibitors of *smHDAC8*. Thus, aiming at the rational planning of new *smHDAC8* inhibitors, a QSAR study was carried out to produce a model capable of predicting the biological activity for this class of compounds. All structures were designed and optimized for the docking stage. The conformations obtained in the docking were used to calculate the descriptors and, consequently, construction of the QSAR-3D model. A model with 3 Lennard-Jones and 1 Coulomb descriptor was able to predict the biological activity for the training set ($R^2 = 0.92$, RMSEC = 0.42). In addition, it was able to predict the IC₅₀ value for an external subset ($Q^2_{ext} = 0.85$) of samples. The interpretation of the descriptors led to several insights on the relationship between activity and structure of hydroxamate-based *smHDAC8* inhibitors. In this way, new compounds can be planned and have their biological activities predicted in silico by the QSAR-3D model.

Keywords: Schistosomiasis; *smHDAC8*; QSAR; molecular modeling

LISTA DE FIGURAS

Figura 1 Distribuição da esquistossomose no mundo em 2012.

Figura 2 Regiões que necessitam de tratamento preventivo no mundo em 2016.

Figura 3 Estruturas químicas dos fármacos praziquantel e oxamniquina.

Figura 4 Ciclo de vida do *S. mansoni*.

Figura 5 Exemplos de inibidores que apresentaram atividade antiproliferativa frente a HDAC'S em diversos parasitas.

Figura 6 Estrutura 3D do *smHDAC8* (código pdb: 6HTT).

Figura 7 Estruturas químicas dos derivados de mercaptoacetoamida e Hidroxamato.

Figura 8 Características estruturais gerais de inibidores *smHDAC8*.

Figura 9. Fluxograma do processo de descoberta e desenvolvimento defármacos.

Figura 10 Estruturas químicas dos fármacos Imatinib, Nilotinib e Alisquireno.

Figura 11 Diferentes ferramentas *in silico* usadas no design defármacos.

Figura 12 Representação simplificada de uma molécula inserida na caixa virtual para cálculo dos descritores moleculares.

Figura 13 Sobreposição dos cristais da enzima *smHDAC8* escolhidos para realização do estudo de *docking*.

Figura 14 Tamanhos selecionados para caixas nas simulações de *docking*.

Figura 15 Alinhamento das conformações após *docking* inclusas no sítio ativo da enzima *smHDAC8*.

Figura 16 Dendograma das 73 moléculas utilizadas na construção do modelo de QSAR 3D, separadas por *clusters* de similaridade.

Figura 17 pIC₅₀ experimental *versus* pIC₅₀ calculado para o conjunto interno de 51 moléculas.

Figura 18 Valores de pIC₅₀ experimental *versus* valores de pIC₅₀ teóricos, calculados pelo modelo de QSAR 3D para o conjunto externo de 22 moléculas.

Figura 19 Intercepto da reta para os valores de R^2 do *y-randomization*.

Figura 20 Intercepto da reta para os valores de Q^2_{LOO} .

Figura 21 Amostragem para o teste de validação Q^2_{LNO} .

Figura 22 Representação espacial dos descritores moleculares.

Figura 23 Representação dos descritores moleculares no sítio ativo da enzima *smHDAC8*.

Figura 24 Representação dos descritores moleculares considerando a área superficial da enzima *smHDAC8*.

Figura 25 Disposição espacial em detalhe do descritor $QQ+$.

Figura 26 Representação espacial detalhada do descritor $LJ+1$.

Figura 27 Representação espacial detalhada do descritor $LJ+2$.

Figura 28 Representação espacial detalhada do descritor $LJ+3$.

LISTA DE QUADROS E TABELAS

Quadro 1 Descrição dos resíduos que compõe o sítio ativo da enzima *smHDAC8*.

Tabela 1 Cristais da enzima *smHDAC8* utilizados no estudo de *docking*.

Tabela 2 Descritores moleculares referentes ao melhor modelo de QSAR-3D calibrado.

LISTA DE ABREVIATURAS E SIGLAS

- 2D Bidimensional ou Segunda dimensão
- 3D Tridimensional ou Terceira dimensão
- 6D Hexadimensional ou Sexta dimensão
- CHARMM Nome do campo de força para o programa CHARMM
- CoMFA *Comparative Molecular Field Analysis*
- FDA *Food & Drug Administration*
- HDAC's Histonas deacetilases
- IC₅₀ Concentração necessária para promover 50% da inibição
- Ki Constante de inibição
- LJ Descritores Potencial de Lennard Jones
- LOO *Leave-one-out*
- LNO *Leave-N-Out*
- MIC Concentração inibitória mínima
- MIF *Molecular Interaction Field*
- OMS Organização Mundial da Saúde
- PDB *Protein Data Bank*
- Q² Coeficiente de Correlação de Validação Cruzada.
- QQ Descritores calculados com o potencial de Coulomb.
- QSAR *Quantitative Structure Activity Relationship*.
- QQP Pontencial de Coulomb positivo
- R² Coeficiente de Correlação de Determinação Múltipla.
- smHDAC8 Histona deacetylase 8 do *S. mansoni*
- VWP Potencial de Lennard Jones positivo

SUMÁRIO

1 INTRODUÇÃO.....	13
1.1 Esquistossomose.....	13
1.2 <i>Schistosoma mansoni</i>	16
1.3 Enzima HDAC8 do <i>S. mansoni</i> (<i>smHDAC8</i>)	18
1.4 Inibidores <i>smHDAC8</i>	21
1.5 Planejamento racional de novos Fármacos.....	22
1.6 <i>Computer-aided drug-design (CADD)</i>	24
1.6.1 <i>Structure-based drug design (SBDD)</i>	26
1.6.1.1 <i>Virtual screening (VS)</i>	26
1.6.1.2 <i>Molecular docking</i>	27
1.6.2 <i>Ligand-based drug design (LBDD)</i>	28
1.6.2.1 <i>Quantitative structure-activity relationship (QSAR)</i>	28
1.7 Importância do QSAR na química medicinal.....	32
2 OBJETIVOS.....	34
2.1 OBJETIVO GERAL.....	34
2.2 OBJETIVOS ESPECÍFICOS.....	34
3 MATERIAL E MÉTODOS.....	35
3.1 Material.....	35
3.1.1 Computador de alto desempenho.....	35
3.1.2 Conjunto de dados.....	35
3.2 Metodologia.....	35
3.2.1 Padronização do conjunto de dados e preparo das estruturas 3D dos ligantes.....	35
3.2.2 Preparo das estruturas 3D da enzima <i>smHDAC8</i>	35
3.2.3 <i>Docking</i>	36
3.2.4 Cálculo dos descritores moleculares e construção do modelo de QSAR- 3D.....	36
3.2.5 Validação do modelo de QSAR-3D.....	37
3.2.6 Interpretação mecanicista dos descritores moleculares.....	38
4 RESULTADOS E DISCUSSÕES.....	39
4.1 Construção do banco de dados.....	39
4.2 Estruturas 3D da <i>smHDAC8</i>	39
4.3 Otimização e alinhamento dos modelos moleculares.....	41

4.4 Cálculo dos descritores e filtro de variáveis.....	45
4.5 Distinção dos conjuntos de treinamento (interno) e de teste (externo).....	46
4.6 Criação e validação dos modelos de QSAR 3D.....	48
4.7 Interpretação mecanicista dos descritores moleculares.....	56
5 CONCLUSÃO.....	66
6 REFERÊNCIAS.....	67
7 ANEXOS.....	77
7.1 Anexo 1. Estruturas 2D das moléculas escolhidas para compor o estudo bem como suas respectivas atividades biológicas.....	81

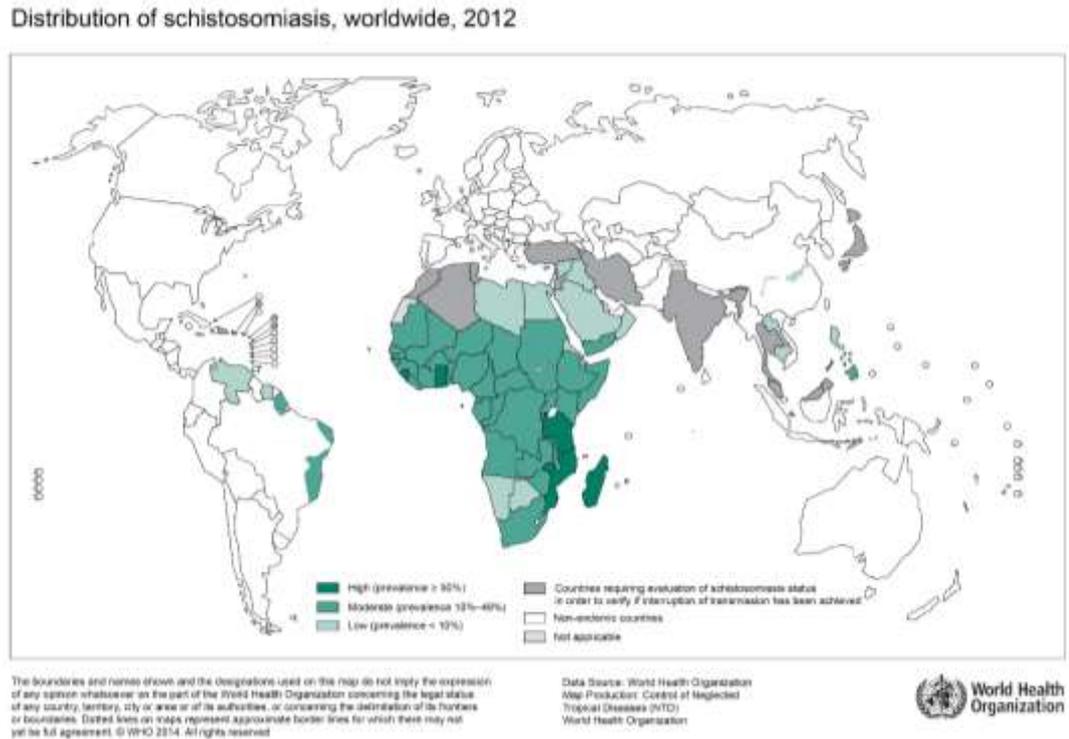
1 INTRODUÇÃO

1.1 Esquistossomose

As doenças tropicais negligenciadas caracterizam um grupo seletivo de enfermidades as quais, na sua grande maioria, possuem parasitas como agentes causadores. Estes parasitas possuem ciclos de vida variados, bem como fisiopatologias distintas, caracterizando mecanismos fisiopatológicos complexos que engajam os pesquisadores para elucidação de métodos capazes de interromper o ciclo de vida, preferencialmente, em seu principal hospedeiro definitivo: o ser humano (Brasil, 2019a).

Dentre as doenças negligenciadas, pode-se destacar a esquistossomose, a qual se trata de uma doença com ocorrência desde a África e leste mediterrâneo, até o sul da América, atingindo principalmente a região do Caribe, Venezuela e Brasil (Brasil, 2019b). A figura 1 apresenta a distribuição desta doença no mundo, na qual o Brasil exibe regiões com prevalência moderada.

Figura 1. Distribuição da esquistossomose no mundo em 2012.



Fonte: *World Health Organization* (Disponível em: https://www.who.int/schistosomiasis/Schistosomiasis_2012-01.png?ua=1).

A esquistossomose apresenta duzentos e vinte milhões de infectados no mundo e mais de três milhões de infectados no Brasil. Aproximadamente duzentos milhões de pessoas necessitaram de tratamento preventivo em 2016; dos quais mais de cem milhões eram crianças em idade escolar. No Brasil, quase dois milhões de pessoas necessitaram de tratamento preventivo no mesmo ano (WHO, 2016a). Apenas países em situações críticas recebem apoio da organização mundial de saúde (OMS) para tratamento preventivo; a Figura 2 apresenta as regiões do mundo que mais necessitam de um tratamento preventivo da esquistossomose.

Figura 2. Regiões que necessitam de tratamento preventivo no mundo em 2016.



Legenda: Na cor branca, países que não necessitam de tratamento preventivo para esquistossomose; na cor laranja, países que necessitam de tratamento preventivo para esquistossomose.

Fonte: *World Health Organization* (Disponível em: <http://apps.who.int/neglected_diseases/ntddata/sch/sch.html>)

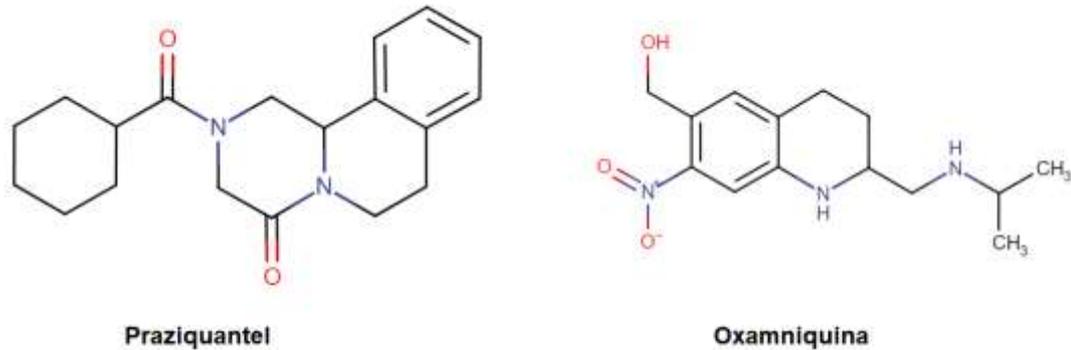
Dados do ministério da saúde apontam o estado de Alagoas com maior número de casos em área endêmica, no período de 2000 a 2016 (Brasil, 2016). O último levantamento quanto ao número de óbitos por esquistossomose no Brasil, por região e unidade federada, entres os anos de 1996 a 2015, explicita o estado de Alagoas com a terceira maior média de mortalidade (Brasil, 2017).

O alto número de óbitos por esquistossomose evidencia a ineficácia e toxicidade dos tratamentos atuais, onde a necessidade de desenvolvimento de novos fármacos se torna uma problemática eminente. No entanto, por se tratar de uma doença negligenciada, não atrai a indústria farmacêutica para investimento.

É interessante ressaltar que, atualmente, as pessoas infectadas são tratadas principalmente com dois fármacos esquistossomicidas: praziquantel e oxamniquina (figura 3). Existem relatos do aumento da resistência à principal linha de tratamento; o praziquantel (Whang, 2012). A OMS possui projetos de incentivo para erradicação

de parasitoses humanas. (WHO, 2016a). Acredita-se que este seja um ponto interessante e que daí possam surgir oportunidades de investimento em pesquisas voltadas para o desenvolvimento de novos fármacos nesta área.

Figura 3. Estruturas químicas dos fármacos praziquantel e oxamniquina.



Fonte: Elaborado pelo autor.

1.2 *Schistosoma mansoni*

O *S. mansoni*, parasita causador da esquistossomose, possui parte de seu ciclo no ser humano, causando diversas complicações clínicas. (Figura 4). Os ovos depositados pelos vermes adultos no trato gastrointestinal inferior do homem, eclodem em condições ideais, liberando os miracídeos, passando pelo hospedeiro intermediário até chegar à forma de cercárias, as quais infectam o homem. No ser humano o verme se desenvolve até a fase adulta para prosseguir o ciclo (Brasil, 2014).

Figura 4. Ciclo de vida do *S. mansoni*.



Fonte: (Brasil, 2014).

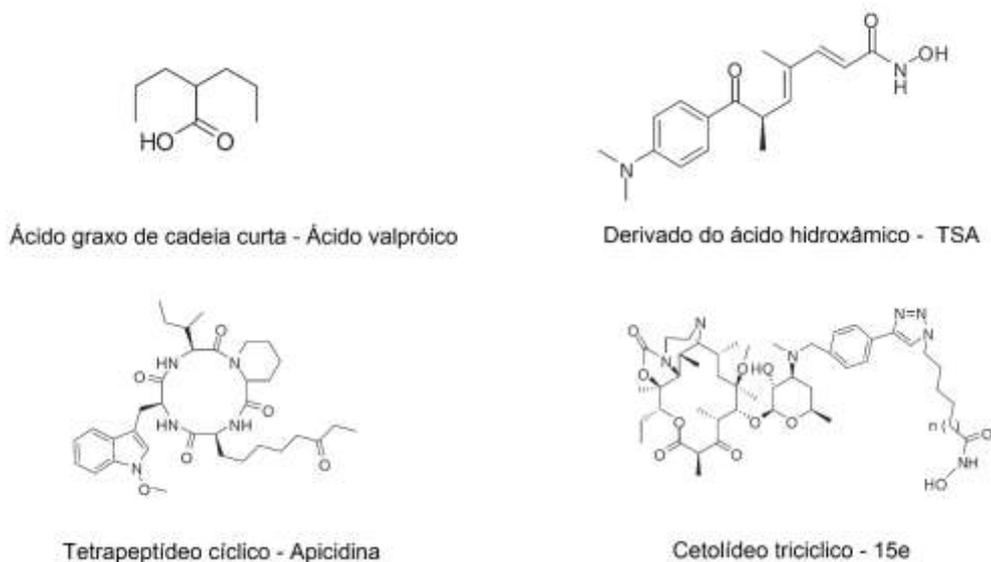
O parasita possui um ciclo extenso e diversas possibilidades de alvo para fármacos. As enzimas importantes para sobrevivência do parasita podem servir como alvo para um possível inibidor do ciclo parasitário. Dentre as enzimas presentes no *S. mansoni*, as mais estudadas são: histona deacetilase 8 (*smHDAC8*) (HEIMBURG *et al.*, 2016), arginase (HAI *et al.*, 2014), endopeptidase de aspariginil (MCKERROW, 2008), enzima de catabolização de NAD⁺ (MULLER-STEFFNER *et al.*, 2017) e tio-redoxina glutationa redutase (GUGLIELMO *et al.* 2014).

1.3 ENZIMA HDAC8 DO *S. mansoni* (*smHDAC8*)

Dentre os últimos estudo realizados com as enzimas histona deacetilases (HDAC's), tem-se observado que estas enzimas podem ser consideradas alvos atrativos, uma vez que estão diretamente ligadas com a replicação do parasita. Ainda assim, nota-se que estas enzimas possuem diversas isoformas, sendo a mais intensamente estudada pela sua importância para o ciclo do *S. Mansoni*.

Investigações recentes mostraram que as HDAC'S desempenham um papel fundamental na regulação gênica transcricional e na progressão do ciclo celular de parasitas eucariotos (ANDREWS, 2012). Diversos estudos demonstraram atividade antiproliferativa de inibidores de HDAC'S em importantes doenças parasitárias e alguns deles são apresentados na figura 5 (ANDREWS; PIERCE, 2012).

Figura 5. Exemplos de inibidores que apresentaram atividade antiproliferativa frente a HDAC'S em diversos parasitas.

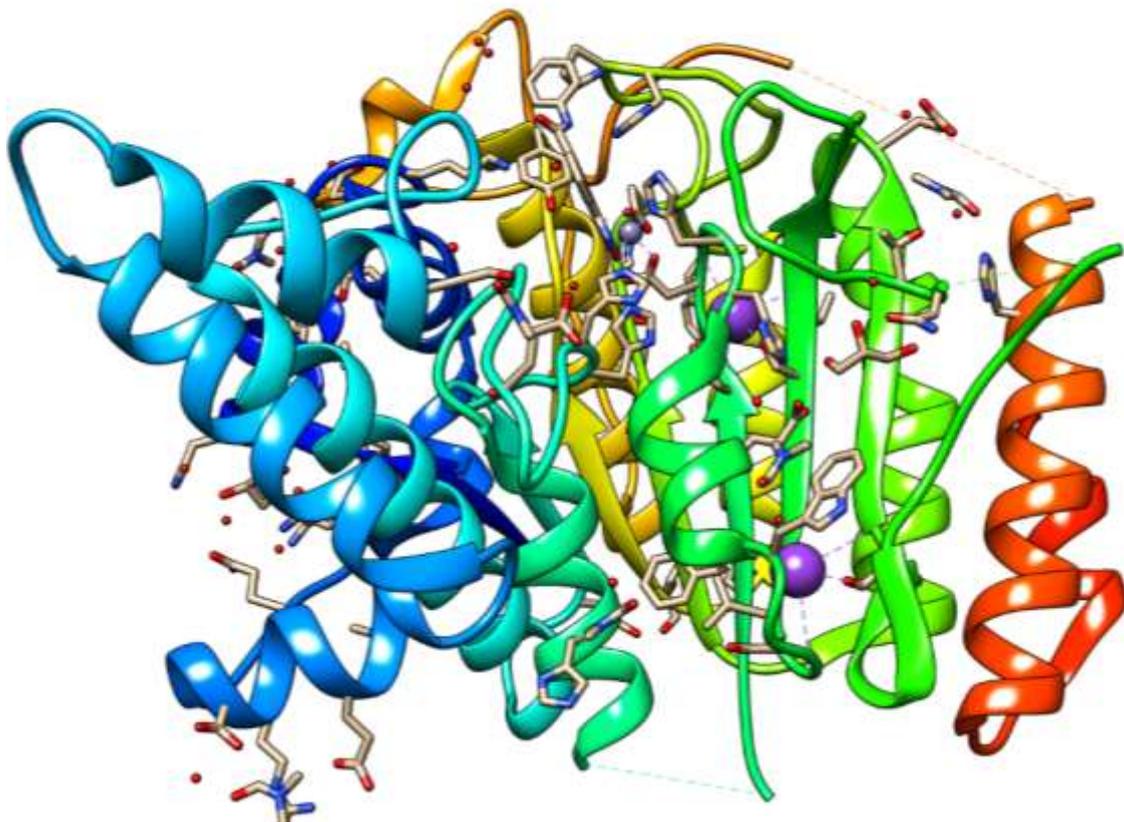


Fonte: Elaborado pelo autor.

A *smHDAC8* é a enzima (da família das HDAC'S) mais abundante no parasita e, está presente em todo o ciclo de vida deste verme (DUBOIS, 2009). Existem diversos depósitos no banco de dados de proteínas (PDB) referentes a enzima *smHDAC8*.

Estas estruturas 3D possuem diversas resoluções e diferentes moléculas complexadas junto a proteína. Como exemplo, uma estrutura 3D da *smHDAC8* pode ser visualizada na Figura 6.

Figura 6. Estrutura 3D do *smHDAC8* (código PDB: 6HTT)



Fonte: Elaborado pelo autor.

Os dados cristalográficos de alvos proteicos são essenciais para estudos de QSAR – método que será mais bem explanado ao longo do texto. A *smHDAC8* possui cadeias em alfa-hélice e algumas folhas beta, que dão mobilidade conformacional às cadeias peptídicas. O principal ponto a se destacar é a presença do elemento Zinco⁺² no centro da proteína. Este metal, junto a alguns resíduos de aminoácidos, são extremamente importantes para o modo de interação e atividade biológica dos inibidores que serão discutidos em seguida. Os resíduos de aminoácidos que compõe o sítio catalítico da enzima *smHDAC8* estão dispostos no quadro 1.

Quadro 1. Descrição dos resíduos que compõe o sítio ativo da enzima *smHDAC8*.

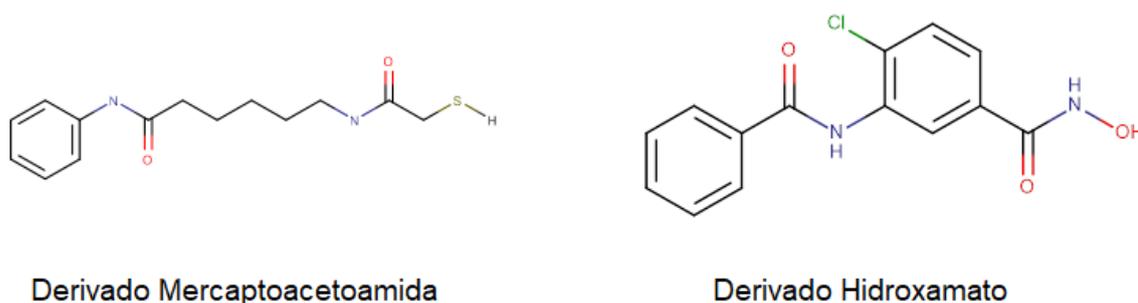
Resíduo de aminoácido	Código no cristal do PDB	Tipo de interação intermolecular	Resíduos essenciais
Aspartato	ASP186	Coordenação com Zinco	Essencial
Aspartato	ASP285	Coordenação com Zinco	Essencial
Glicina	GLY150	Ligação de hidrogênio	Não essencial
Histidina	HIS141	Empilhamento Pi	Não essencial
Histidina	HIS142	Ligação de hidrogênio	Não essencial
Histidina	HIS188	Coordenação com Zinco	Essencial
Histidina	HIS292	Ligação de hidrogênio / empilhamento Pi	Não essencial
Lisina	LYS20	Ligação de hidrogênio	Não essencial
Fenilalanina	PHE216	Hidrofóbica	Não essencial
Prolina	PRO 291	Hidrofóbica	Não essencial
Tirosina	TYR341	Ligação de hidrogênio	Não essencial
Zinco ⁺²	ZN	Interação coordenada	Essencial

Fonte: elaborado pelo autor.

1.4 Inibidores *smHDAC8*

Diversos estudos têm relatado moléculas com atividade inibitória frente à *smHDAC8* (BAYER *et al.*, 2018; CABY *et al.*, 2017; STOLFA *et al.*, 2014). Como exemplos podemos citar os derivados de mercaptoacetamidas e derivados do hidroxamato (figura 7). Assim como outras classes de inibidores, estes compostos possuem, a capacidade de interagir com a zinc no sítio catalítico no qual está contido o elemento Zn^{+2} , desempenhando um papel inibitório diante da *smHDAC8*.

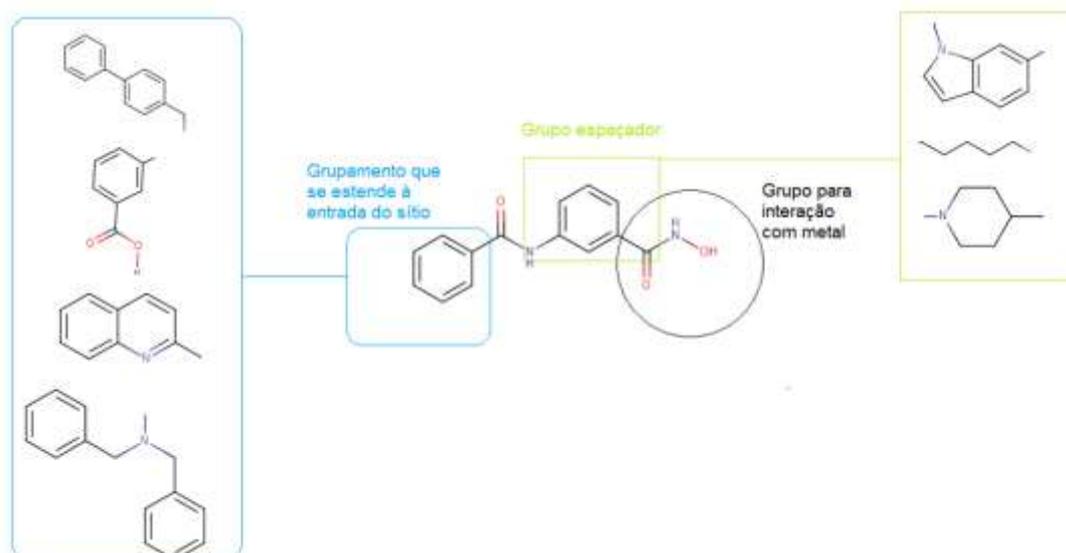
Figura 7. Estruturas químicas dos derivados de mercaptoacetamida e Hidroxamato.



Fonte: Elaborado pelo autor.

Até o momento, foi verificado que inibidores de *smHDAC8* devem possuir três elementos essenciais para interação no sítio catalítico: (i) um grupo que possua extrema afinidade pelo elemento Zn^{+2} e aos resíduos de aminoácidos que o circundam (ver quadro 1); (ii) um grupamento espaçador e (iii) um grupo volumoso que se estende até a entrada da cavidade hidrofóbica do sítio (MAREK *et al.*, 2018). Um exemplo de estrutura básica de uma série de inibidores *smHDAC8* pode ser visualizado na figura 8. O desenho mostra um inibidor com alguns exemplos de grupos substituintes para o grupo espaçador e para o grupamento que se estende à entrada do sítio.

Figura 8. Características estruturais gerais de inibidores *smHDAC8*.

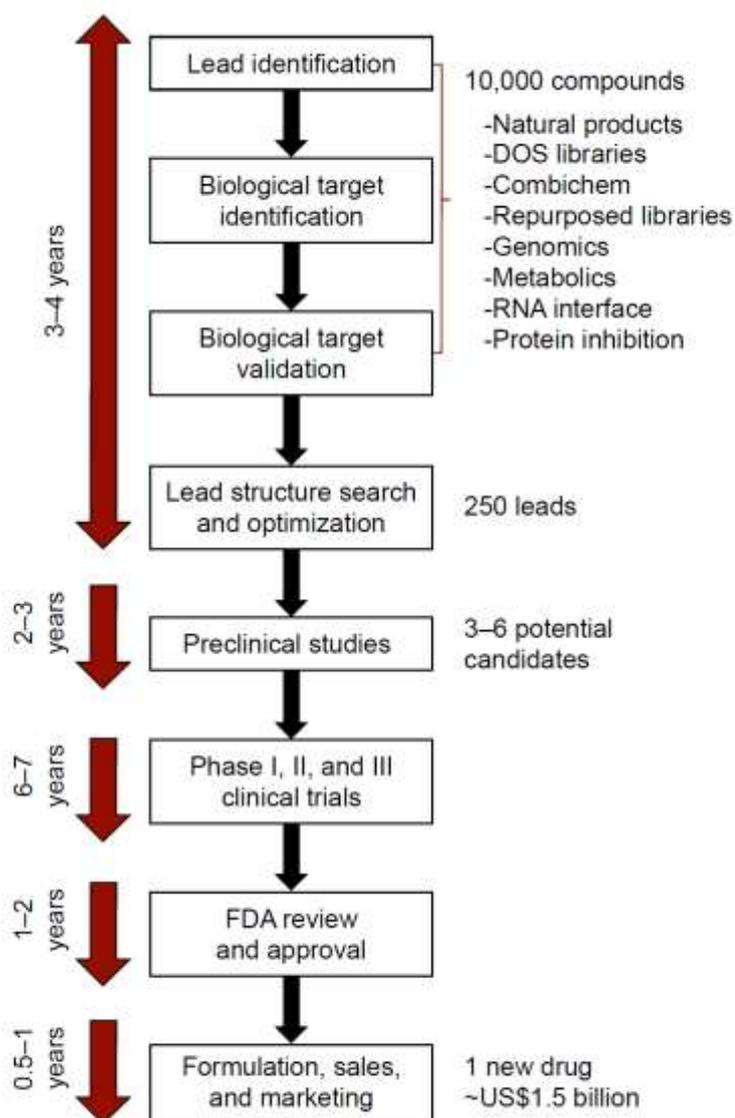


Fonte: Elaborado pelo autor.

1.5 Planejamento racional de novos Fármacos

É sabido que para ocorrer o desenvolvimento de um novo fármaco, desde a descoberta e síntese de uma estrutura, até o produto que chega ao paciente, se trata de um processo demorado que demanda empenho de equipes multiprofissionais e especializadas. O custo para tudo isso é alto. Portanto, se faz necessário o desenvolvimento de técnicas que racionem a relação tempo/custo de forma positiva. É dentro desta visão que se opta por realizar o planejamento racional de novos compostos. Tal tipo de planejamento é acoplado a diversas técnicas, nas mais diversas etapas de desenvolvimento de um novo fármaco. Neste sentido, o presente trabalho se intensifica na etapa inicial e primordial, que se trata da elaboração e descoberta de possíveis moléculas que possam chegar ao status de *hit* ou composto líder. A figura 9 apresenta um fluxograma que abrange o tempo e o valor gasto para se desenvolver um fármaco.

Figura 9. Fluxograma do processo de descoberta e desenvolvimento de fármacos.



Fonte: (ARODOLA; SOLIMAN, 2017).

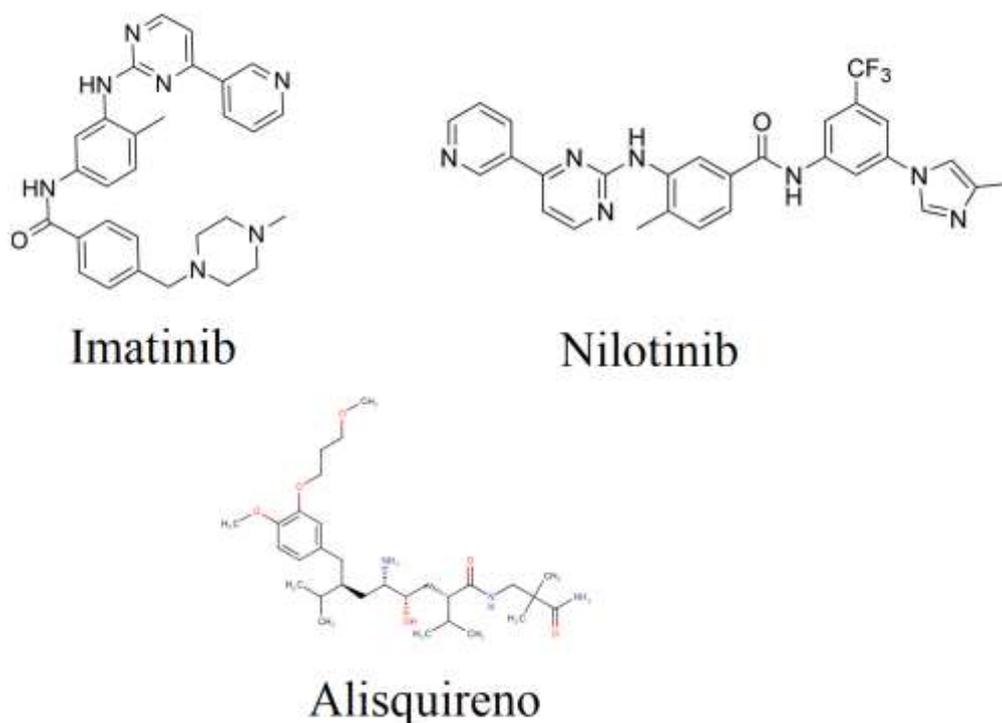
Os estudos de modelagem molecular são as principais ferramentas teóricas no campo da química medicinal, as quais têm sido largamente usados para descoberta de novos fármacos. A modelagem molecular abrange todos os métodos usados para

modelar ou imitar o comportamento das moléculas. Os métodos de modelagem molecular são utilizados nos campos da química computacional, *design* de fármacos, bioinformática e ciência dos materiais, para estudar sistemas moleculares que variam de pequenos sistemas químicos à macromoléculas biológicas e conjuntos de materiais. A característica comum dos métodos de modelagem molecular é a descrição do nível atômico dos sistemas moleculares. Isso pode incluir tratar os átomos como a menor unidade individual (uma abordagem de mecânica molecular) ou modelar explicitamente prótons e nêutrons com seus quarks, anti-quarks e elétrons com seus fótons (uma abordagem química quântica) (RAMACHANDRAN et al., 2008; ANDREW, 2001).

1.6 *Computer-aided drug-design* (CADD)

As técnicas de modelagem molecular são diversas, onde cada uma possui uma aplicação diferente, que deve ser apreciada para cada caso. Dentro do campo da modelagem molecular, existem os estudos de CADD, o qual tem como objetivo o desenvolvimento e melhora da eficácia de fármacos usando ferramentas computacionais modernas, as quais se encaixem como rápidas e econômicas em comparação aos métodos convencionais. A abordagem de CADD já foi aplicada em vários fármacos de sucesso, como na descoberta dos quimioterápicos imatinib e nilotinib (DRUKER; LYDON, 2000; WEISBERG et al., 2005). Outro exemplo, é o fármaco alisquireno, usado no tratamento de hipertensão arterial, o qual foi descoberto através de estudos de modelagem molecular (WOOD *et. al.*, 2003). A estrutura 2D dos medicamentos citados anteriormente estão apresentados na figura 10.

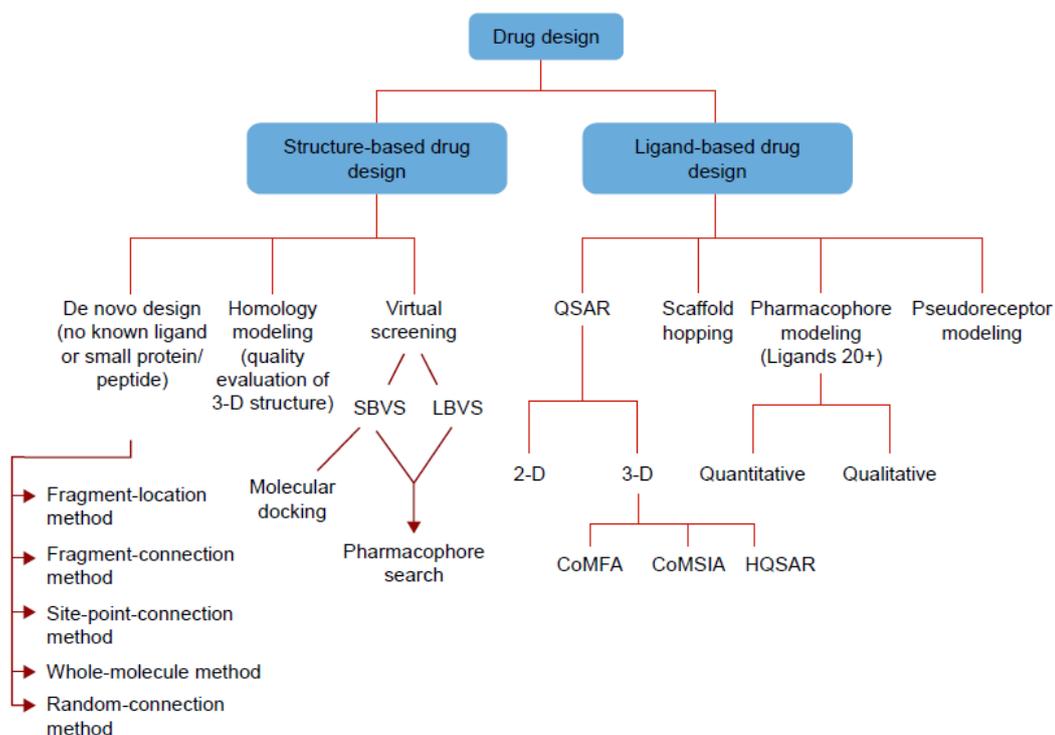
Figura 10. Estrutura químicas dos fármacos Imatinib, Nilotinib e Alisquireno.



Fonte: Elaborado pelo autor.

Em termos gerais, o CADD é subdividido em duas áreas: *Structure-based Drug Design* (SBDD) e *Ligand-based Drug Design* (LBDD). O SBDD abrange ferramentas utilizadas para o desenvolvimento de fármacos baseadas na estrutura, ou seja, no alvo de ação o qual o fármaco planejado irá interagir, que pode ser uma proteína ou uma sequência de DNA, por exemplo. O LBDD consiste em técnicas baseado na estrutura de ligantes, os quais podem ser fármacos já usados clinicamente, ou *hits* em fase de triagem, por exemplo (ARODOLA; SOLIMAN, 2017; AUGEN, 2002). A figura 11 apresenta um fluxograma com as diferentes ferramentas *in silico* utilizadas no planejamento de fármacos.

Figura 11. Diferentes ferramentas in silico usadas no design de fármacos.



Fonte: (ARODOLA; SOLIMAN, 2017).

1.6.1 Structure-based Drug Design (SBDD)

1.6.1.1 Virtual screening (VS)

Uma das principais ferramentas de SBDD é o *virtual screening* (VS). No VS, é possível realizar uma triagem com diversos alvos macromoleculares, ou pequenos ligantes já conhecidos, a fim de se construir um modelo farmacofórico e, a partir disto, realizar o desenho de novas moléculas. No VS existem dois métodos: o *structure-based virtual screening* (SBVS) e o *ligand-based virtual screening* (LBVS) (GIMENO et al., 2019).

No SBVS é realizada a ancoragem de uma biblioteca de moléculas para um, ou vários alvos, a fim de selecionar um subconjunto desses compostos com base no

score de ligação através do *docking* molecular (esta ferramenta será explanada mais adiante neste manuscrito). No LBVS, são utilizadas as informações presentes em ligantes ativos conhecidos, sem a necessidade de conhecimento do alvo molecular. Assim, abordagens virtuais baseadas em ligantes podem ser empregadas no intuito de se encontrar novas moléculas, avaliando-se a semelhança entre os ligantes candidatos a fármaco e os compostos ativos conhecidos. Em ambos os casos, SBVS e LBVS, o intuito principal é conhecer o modelo farmacofórico ideal para interação (BANEGAS-LUNA; CERÓN-CARRASCO; PÉREZ-SÁNCHEZ, 2018; WU; ARIGHI; ROSS, 2017).

1.6.1.2 *Molecular docking*

No campo da modelagem molecular, o *docking* é um método que prediz a orientação preferida de uma molécula em um sítio de ligação. O conhecimento desta orientação pode ser usado para prever a força da associação ou afinidade de ligação entre duas moléculas usando, por exemplo, funções de pontuação. O encaixe molecular (*docking*) é um dos métodos mais freqüentemente usados no projeto de fármacos com base na estrutura, devido à sua capacidade de prever a conformação de ligação de pequenas moléculas no sítio de ligação de um alvo apropriado. A caracterização do comportamento de ligação desempenha um papel importante no planejamento racional de fármacos, bem como na elucidação de processos bioquímicos fundamentais. Este método consiste em calcular a energia livre de Gibbs entre duas espécies químicas, geralmente pequenas moléculas e macromoléculas biológicas. Estes valores são expressos em Kcal/mol e quanto mais negativo for esse valor, significa que maior é a interação entre uma determinada molécula com os aminoácidos de uma proteína, por exemplo (JOHN R.GIUDICESSI, BA.MICHAEL J.ACKERMAN., 2008).

1.6.2 Ligand-based Drug Design (LBDD)

1.6.2.1 *Quantitative structure-activity relationship (QSAR)*

De forma geral, o QSAR é um método empregado para correlacionar a atividade biológica de uma gama de compostos com suas respectivas estruturas químicas, de forma quantitativa. Para tanto, realiza-se a construção de um modelo estatístico capaz de prever a atividade biológica de moléculas para um determinado alvo. Os estudos de QSAR são extremamente confiáveis, uma vez que podem ser validados quanto a robustez estatística, servindo como métrica de confiabilidade aos resultados obtidos (VERMA, 2010).

Neste tipo de estudo, inicialmente, é importante a realização de uma revisão bibliográfica apurada, na qual algumas características devem ser levadas em conta como os critérios de inclusão para compor o banco de dados do estudo. São eles: (i) todas as moléculas usadas devem possuir estrutura química completamente elucidada; (ii) atividade biológica definida para o alvo em questão; e (iii) que o protocolo usado para os testes de atividade biológica seja o mesmo em todas as séries inclusas no banco de dados. (FOURCHES, 2010).

Os estudos QSAR que codificam apenas propriedades físico-químicas são chamadas de QSAR-2D. Ainda existem modelos de QSAR-3D a 6D, que também utilizam os métodos de regressão citados anteriormente para criar modelos de predição. A diferença entre eles são os tipos de descritores utilizados (XUE; BAJORATH, 2000). Em um QSAR 3D são consideradas propriedades estéricas e eletrônicas das moléculas. Quando se usa simulações de dinâmica molecular para gerar as conformações das moléculas, considera-se uma variante a mais no estudo, e, portanto, o modelo de QSAR gerado é denominado 4D.

Desta forma, o QSAR baseia-se na ideia de que moléculas análogas pertencentes a uma mesma classe e um mesmo sistema biológico possam ter suas propriedades de interação quantificadas através de modelos matemáticos. O modelo de QSAR deve clarear as variáveis independentes que são os descritores os quais

codificam aspectos estruturais e físico químicos de interesse, e as variáveis dependentes, que são as atividades biológicas de uma classe de compostos.

O termo variável dependente em QSAR é designado às atividades biológicas obtidas de estudos experimentais *in vivo* ou *in vitro*. Normalmente são utilizados valores de IC₅₀ (dose necessária para inibir 50% da atividade biológica de um determinado alvo), DL₅₀ (dose necessária para matar 50% da população de um determinado organismo vivo), Ki (coeficiente de inibição enzimática), e MIC (concentração inibitória mínima) (NANTASENAMAT et al., 2009).

Desta forma, tem-se em um estudo de QSAR a construção de uma matriz, na qual as variáveis independentes (propriedades da molécula) estão dispostas no eixo X, e as variáveis dependentes (valores de atividade biológica) no eixo Y. Os valores dispostos nessa matriz são correlacionados através da regressão linear multivariada através de diversas abordagens, incluindo o *Comparative Molecular Field Analysis* (CoMFA), *Comparative Molecular Similarity Indices Analysis* (COMSIA) e o *Hologram Quantitative Structure Activity Relationship* (HQSAR).

A abordagem mais citada de QSAR-3D é conhecida como CoMFA (CRAMMER; PATTERSON; BUNCE, 1988). No QSAR-3D, a estrutura 3D de uma molécula é colocada dentro de uma caixa virtual, representando um sítio ativo hipotético rígido, onde são avaliadas as propriedades eletrostáticas e estéricas dos compostos. Estas propriedades são traduzidas em energias de interações em cada ponto de uma grade na caixa virtual. Primordialmente, são calculados os potenciais de Coulomb (QQ) e de Lennard-Jones (LJ), sendo chamados de descritores moleculares. Para o potencial QQ temos descritores eletrostáticos e para o potencial LJ, descritores de Vander Waals. As equações 1 e 2, demonstram como são calculados os potenciais QQ e LJ respectivamente.

$$E_{QQ} = \sum_{K=1}^K \frac{q_{sonda}q_k}{Dr_{sondak}} \quad (1)$$

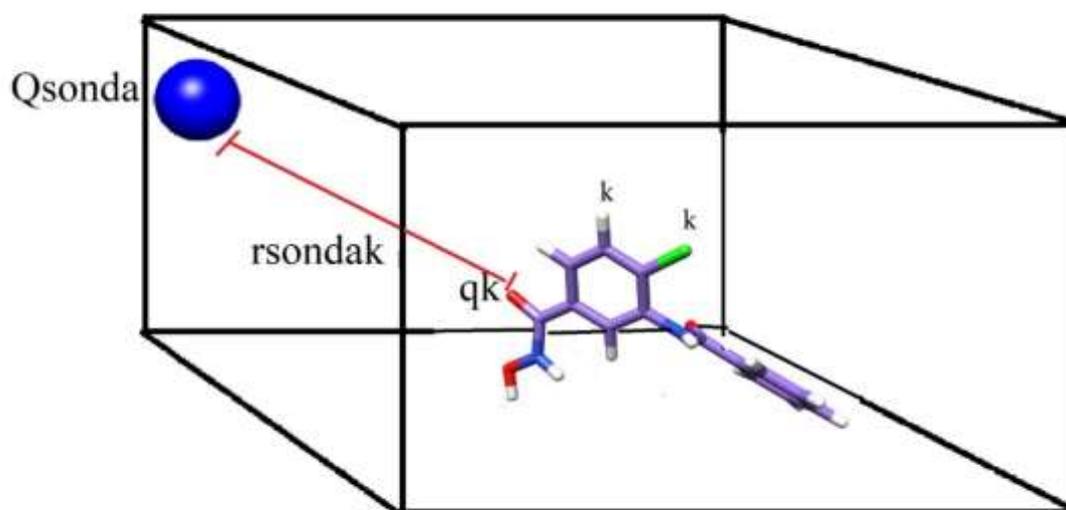
$$E_{vdw} = \sum_{k=1}^k (Ar_{sondak}^2 - Cr_{sondak}^{-6}) \quad (2)$$

Fonte: (PAULO et al.,2009).

Legenda. *E_{QQ}*: Energia dos descritores eletrostáticos; *E_{vdw}*: energias dos descritores de Van der Waals (WDV); *Q_{sonda}*: carga do átomo sonda; *q_k*: carga parcial do átomo *k* da molécula; *D*: constante dielétrica; *r_{sondak}*: distância da sonda dentro da grade virtual até o átomo *k* da molécula; *A* e *C*: constantes dependentes do raio de van der Waals parametrizados para o campo de força usado no cálculo da energia.

Para se entender como estas equações são aplicadas no campo molecular, é necessário compreender que as moléculas se encontram dispostas em um espaço tridimensional (uma grade virtual). A figura 12 demonstra uma representação simplificada de como uma molécula é inserida dentro de uma caixa virtual, para a realização dos cálculos dos descritores moleculares. A partir de cada átomo, as equações são aplicadas para calcular a interação entre o mesmo e um átomo sonda, no caso um carbono sp³. O resultado de cada equação gera um descritor eletrostático e um descritor de Vander Walls para cada ponto na grade. A partir disto, espera-se a geração de algumas dezenas de milhares de descritores. São estes valores que são postos na matriz, e correlacionados aos valores de atividade através da regressão linear multivariada.

Figura 12. Representação simplificada de uma molécula inserida na caixa virtual para cálculo dos descritores moleculares.



Fonte: Elaborado pelo autor.

Note que, analisando em conjunto, as equações 1 e 2 com a figura 11, é perceptível que uma molécula está inserida em uma grade virtual 3D. Um átomo de oxigênio (qk) da molécula, está a uma distância "X" (rsondak) do átomo sonda (Qsonda). Alguns valores de constantes são aplicados na equação, como a constante dielétrica (D). Outros valores são constantes, porém dependentes do raio de Van der Waals utilizado pelo campo de força escolhido (A e C).

No entanto, para alcançar a primazia na construção de modelos de QSAR, é necessário que todas as etapas que antecedem a construção dele, sejam feitas de forma criteriosa, obedecendo padrões e métodos consistentes já citados na literatura. Todo modelo de QSAR deve seguir os padrões e normas técnicas exigidos pela *Organisation for Economic Co-operation and Development* (OECD). A OECD é uma

organização econômica intergovernamental com 36 países membros, fundada em 1961 para estimular o progresso econômico e o comércio mundial. É necessário que um bom modelo de QSAR possua um processo de validação transparente com determinação objetiva da confiabilidade. Em novembro de 2004, os países membros da OCDE estabeleceram os princípios de validação dos modelos QSAR visando uma avaliação regulamentar destes modelos. Os princípios acordados fornecem a base para avaliar a aplicabilidade regulatória dos modelos de QSAR e contribuem para seu uso aprimorado e uma avaliação eficiente (OECD, 2004, 2007).

Segundo os padrões da OECD, um modelo de QSAR adequado deve possuir: (i) Um ponto final definido; (ii) um algoritmo inequívoco; (iii) um domínio de aplicabilidade definido; (iv) medidas apropriadas de qualidade de ajuste, robustez, e previsibilidade; esse possível (v) uma interpretação mecanicista.

1.7 Importância do QSAR na química medicinal

Os métodos atuais de busca de novos fármacos contam com forte contribuição dos estudos de QSAR, os quais visam a identificação de compostos líderes e dão embasamento físico-químico para o planejamento de análogos mais específicos, com maior atividade intrínseca ou com melhor perfil farmacológico. Desta forma, aumentam-se as chances de sucesso de inserção de novos fármacos no arsenal terapêutico.

Diversos trabalhos têm demonstrado a alta aplicabilidade do QSAR para o descobrimento de novos *hits*. Esta ferramenta se adequa a proposta do planejamento racional de fármacos, possibilitando ao pesquisador, prever qual será a atividade de uma determinada molécula antes de sua síntese. Através do QSAR, pesquisadores conseguiram encontrar novos *hits* com atividade esquistossomicida semelhante ao fármaco praziquantel (NEVES et al., 2016). Uma análise de QSAR bem sucedida foi relatada na descoberta de inibidores potentes da tirosina fosfatase humana, sendo usados como ponto de partida para o desenvolvimento de quimioterápicos que combatem células cancerígenas (TAHA et al., 2007). As análises de QSAR também foram citadas em estudos de sucesso que relataram o desenvolvimento de novas moléculas capazes de inibir enzimas bacterianas e células do parasita *Trypanosoma*

cruzi.(MELO-FILHO et al., 2019; TAHA et al., 2008; TUFFAHA; HATMAL; TAHA, 2019).

Em muitos casos, o uso integrado de estratégias de SBDD e LBDD pode gerar informações úteis no planejamento de fármacos, por meio da sinergia e complementaridade de conhecimentos entre as estratégias. Diversos estudos apresentam resultados satisfatórios utilizando a combinação de técnicas de SBDD e LBDD (GUIDO; ANDRICOPULO; OLIVA, 2010) (COMB 0). As diversas técnicas combinadas, podem maximizar os resultados obtidos. Pesquisadores obtiveram sucesso ao combinar as ferramentas de *docking* e QSAR-3D no desenvolvimento de inibidores seletivos da desidrogenase presente no parasita *T. cruzi*, (GUIDO et al., 2008). A integração dos métodos também é citada na identificação de potentes fármacos desenvolvidos para o tratamento de fibrodisplasia óssea (SATO et al., 2020).

Neste sentido, o presente trabalho visou a aplicação de técnicas de modelagem molecular, para o desenvolvimento de um modelo de QSAR-3D capaz de prever a atividade biológica (*in silico*) de moléculas, frente a enzima *smHDAC8*.

2 OBJETIVOS

2.1 Objetivo geral

Desenvolver um modelo de QSAR-3D para o planejamento racional de novos inibidores HDAC8 do *S. mansoni*.

2.2 Objetivos específicos

- Padronizar um conjunto de moléculas contidas na literatura com atividade frente a *smHDAC8*;
- Preparar e otimizar estruturas 3D da enzima *smHDAC8* para realização dos estudos de docking;
- Realizar o *docking* de todo o conjunto de moléculas com a enzima *smHDAC8*;
- Calcular descritores moleculares 3D para todo o conjunto de moléculas;
- Criar e validar estatisticamente o modelo de QSAR-3D, segundo os padrões da OECD;
- Entender a nível molecular a relação existente entre as estruturas químicas das classes de moléculas em estudo e suas respectivas atividades biológicas;

3 MATERIAL E MÉTODOS

3.1 Material

3.1.1 Estação de trabalho

Todos os cálculos foram realizados usando *workstation*, Dell *Precision Workstation T7810*, com processador Intel® Xeon® E5-2623 v3 (4C, 3.0 GHz, 10 Mb Cache, 105 w), com duas placas de vídeo NVIDIA® Quadro® NVS510 de 2GB (4DP), memória de 16 GB (4x4GB) 2400 MHz DDR4, e disco rígido SATA de 1 TB (7200 RPM).

3.1.2 Conjunto de dados

Para compor o conjunto de moléculas, foram utilizados 73 derivados de hidroxamatos ou mercaptoacetoamidas (KANNAN *et al.* 2014; STOLFA *et al.*, 2014; HEIMBURG *et al.*, 2016; BAYER *et al.*, 2018). As moléculas estão dispostas no anexo 1.

3.2.1 Padronização do conjunto de dados e preparo das estruturas 3D dos ligantes

Inicialmente, foram inseridos ao banco de dados todos os inibidores da enzima *smHDAC8*, descritos na literatura com estrutura química definida, atividade biológica determinada com mesma ordem de grandeza, e mesmo protocolo de ensaios de atividade biológica (as estruturas utilizadas neste estudo estão dispostas no **ANEXO 1**, ao final deste manuscrito).

Foi calculado o estado de protonação de cada molécula, de acordo com o pH fisiológico. Foram atribuídas as cargas positiva ou negativa, nos casos em que a molécula se encontrava na sua forma ionizada no pH fisiológico. No software Marvin foi calculada a conformação energeticamente mais estável, de acordo com as seguintes especificações: geometria Hessian, QM DFT, Cálculo Hessian, nível de teoria DFT (B3LYP / 6 31G *) e cargas derivados de Merz-Singh-Kollman.

3.2.2 Preparo das estruturas 3D da enzima *smHDAC8*

As estruturas cristalinas da *smHDAC8* foram pesquisadas no banco de dados *PDB*. Foram escolhidos os cristais com melhor resolução (resolução < 3.0 Ångström).

e que possuíam ligantes complexados. Dentre as estruturas selecionadas, foi pesquisado quais ligantes complexados (no sítio catalítico da enzima *smHDAC8*) se assemelhavam mais com o esqueleto principal das 73 moléculas que compõe o banco de dados. Foi verificado o grau de similaridade dos cristais da enzima *smHDAC8*, utilizando-se o módulo *match maker* do USCF Chimera (PETTERSEN, 2004). As estruturas das proteínas foram otimizadas quanto ao estado de protonação no pH fisiológico e, também foram adicionadas cargas formais AM1BCC. As moléculas otimizadas, junto aos cristais aprimorados, foram usadas para os estudos de *docking* molecular.

3.2.3 Docking

Para as simulações de docking, foram utilizados 3 softwares diferentes: Autodock vina (TROTT, 2010), Autodock 4.2 e o ICM (Molsoft). Para a escolha do tamanho ideal da caixa, todas as moléculas foram submetidas a ensaios com 3 dimensões: 9 x 12 x 6 Å, 14 x 17 x 11 Å e 19 x 22 x 16 Å. As conformações geradas para cada ensaio de *docking* foram inspecionadas visualmente a fim de se encontrar qual conformação mais se assemelhava a conformação do ligante complexado ao cristal. Após análise, foi escolhida 1 conformação preferível para cada molécula, segundo os parâmetros estabelecidos. Todas as estruturas foram salvas em um único arquivo, para realização da etapa de cálculo de descritores.

3.2.4 Cálculo dos descritores moleculares e construção do modelo de QSAR-3D

Após obtenção da conformação ideal para todas as moléculas da série, foram utilizados os potenciais de QQ e LJ para cálculo dos descritores moleculares. O Open3DQSAR foi utilizado para calcular dezenas de milhares de descritores dispersos em um espaço 3D. Foram utilizados filtros de variáveis para exclusão de descritores redundantes.

Nas etapas seguintes, os inibidores foram separados em dois conjuntos: interno (ou treinamento) — o qual contém as moléculas utilizadas para construção do modelo — e conjunto externo (ou validação) — composto por moléculas usadas para validação do modelo de QSAR (todas estruturas utilizadas nos conjuntos interno e externo,

estão dispostas no **ANEXO 1** ao final deste manuscrito). Para a seleção das moléculas de cada grupo, foi utilizado o uso de um dendograma gerado pelo software ICM, baseado na similaridade das estruturas e suas respectivas atividades biológicas, separando as moléculas em clusters.

Os descritores calculados de LJ e QQ, para o conjunto interno foram submetidos ao software *Nanobridges* (AMBURA *et. al.*, 2009) Para geração dos modelos de QSAR. os melhores foram submetidos aos testes de validação para escolha daquele que melhor descreveu a atividade biológica.

3.2.5 Validação do modelo de QSAR-3D

Inicialmente, foi analisado o valor de R^2 para o conjunto interno. No R^2 , os valores de atividade experimental do conjunto interno foram plotados com os valores de atividade calculados pelo modelo de QSAR.

A validação externa do modelo foi avaliada pelo desvio entre os valores de atividade biológica experimentais e os preditos pelo modelo.

Os valores preditos pelo modelo foram calculados da seguinte forma: para cada molécula, foram selecionados os valores dos descritores que melhor descrevem a atividade biológica, segundo o modelo de QSAR 3D. O vetor de regressão (VR) de cada descritor foi multiplicado pelo valor de cada descritor molecular, e foi realizada então a somatória destes valores. Esta somatória foi multiplicada pela constante do modelo. O resultado desta operação foi descrito como a atividade biológica predita pelo modelo, para cada molécula.

Para averiguar a qualidade da previsão externa, foi analisada a inclinação da reta entre os valores preditos e experimentais (Q^2 externo). Nesta etapa da validação, as moléculas que compõem o conjunto externo, ou seja, que não foram usadas para construção do modelo, foram submetidas ao modelo de QSAR-3D.

Para verificar se os descritores obtidos no modelo de QSAR estavam realmente relacionados com a farmacodinâmica dos compostos, foi avaliado o intercepto do R^2 para o teste denominado *y-randomization*. No *y-randomization* foram construídos

novos modelos de QSAR usando os mesmos descritores calculados anteriormente, porém, com os valores de atividade biológica embaralhados. Estes modelos foram reconstruídos N vezes o número de moléculas do conjunto interno. Com isso, para cada modelo reconstruído com atividades embaralhadas, foi gerado um valor de R^2 . Com estes valores, foi realizado o intercepto da reta em relação ao valor de R^2 do modelo original.

Para avaliar o grau de previsibilidade do modelo foi realizada uma validação cruzada utilizando o teste *Leave-One-Out* (Q^2_{LOO}). Foi excluída uma amostra de cada vez do conjunto de treinamento. O procedimento foi realizado quantas vezes for o número das amostras do conjunto de treinamento. A diferença entre o valor experimental e o estimado para as amostras retiradas foram usados para calcular o erro padrão da validação cruzada (*SEV, standart error of validation*) e o coeficiente de correlação da validação LOO (Q^2).

O Processo denominado *Leave-N-out* foi usado para avaliar a robustez do modelo de QSAR-3D. Foram retiradas do conjunto de treinamento N amostras em bloco. Um novo modelo foi construído sem estas amostras, e os valores de atividade biológica são previstos para o bloco de amostras separadas para validação.

Em seguida, foi realizada uma interpretação mecanicista dos descritores contidos no melhor modelo avaliado.

3.2.6 Interpretação mecanicista dos descritores moleculares

Foi realizada uma interpretação minuciosa dos descritores no espaço 3D. Foram explanadas quais interações são favoráveis ou desfavoráveis, no ponto o qual se encontra cada descritor dentro do sítio ativo. Esta interpretação foi realizada da seguinte forma: foram comparados, para cada descritor, quais moléculas possuíam o menor e o maior valor para o referido descritor. Em seguida foi verificado qual grupo funcional, nas duas moléculas, se encontrava mais próximo do descritor analisado. A partir disto, foi distinguindo grupos favoráveis ou desfavoráveis para interação a partir da natureza dos descritores. Esta análise foi realizada para cada descritor obtido no modelo de QSAR-3D.

4 RESULTADOS E DISCUSSÕES

4.1 Construção do banco de dados

Os 73 inibidores selecionados a partir da literatura, possuíam estrutura química definida, sem presença de moléculas com centros quirais indefinidos. Todas estruturas possuíam valores de atividade biológica na mesma grandeza, com valores de IC₅₀ e boa variabilidade de atividade biológica, desde a escala micromolar até a nanomolar (nM). Esta variabilidade foi importante para que as equações de regressão linear utilizadas fossem capazes de correlacionar valores de atividade baixos, porém não inativos, até valores altos, com as diversas propriedades moleculares das estruturas. Todas as moléculas escolhidas apresentavam o mesmo tipo de ensaio enzimático à enzima *smHDAC8*. O desenho 2D de todos inibidores utilizados neste estudo e suas respectivas atividades, bem como os artigos científicos os quais estas moléculas foram retiradas, estão dispostos no **ANEXO 1**, ao final deste manuscrito.

4.2 Estruturas 3D da *smHDAC8*

Ao todo, foram encontrados 28 depósitos referentes a cristais de *smHDAC8*. As estruturas cristalográficas que melhor outorgaram os padrões estabelecidos no item 3.2.2 para escolha dos cristais estão dispostos na tabela 1.

Tabela 1. Cristais da enzima *smHDAC8* utilizados no estudo de *docking*.

Código PDB	Resolução (Ångström)	Série de ligante cristalografado
4BZ6	2.0	Hidroxamatos alifáticos
6HTT	1.748	Hidroxamatos Benzílicos
4CQF	2.3	Mercaptoacetoamidas

Fonte: *Protein data bank* (<https://www.rcsb.org/>)

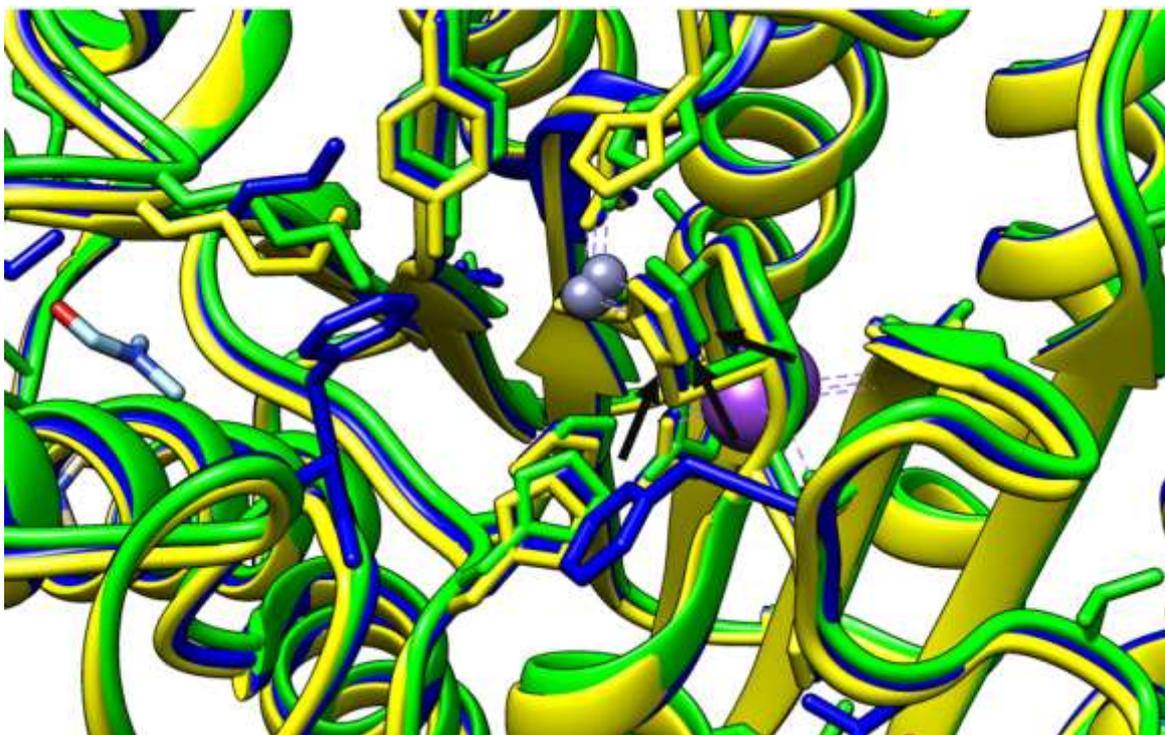
As determinações experimentais são bases magnificentes nos estudos de modelagem molecular. Portanto, é de grande interesse a escolha de estruturas cristalográficas,

nas quais contenham ligantes, com o maior grau de paridade possível com as moléculas da série. Com as enzimas retiradas do PDB, foi possível ter referências primorosas, as quais abrangem todas classes de inibidores presentes no banco de dados. Neste sentido, os hidroxamatos alifáticos foram submetidos aos estudos de *docking* com o PDB de código 4BZ6, os hidroxamatos benzílicos com o PDB 6HTT, e as poucas mercaptoacetoamidas, com o PDB 4CQF.

No entanto, em um estudo de docking, as conformações de todas as moléculas devem possuir um exímio alinhamento, sendo necessário atenção ao usar estruturas enzimáticas de códigos diferentes. É interessante ressaltar que, por mais que estas estruturas escolhidas possuam ligantes complexados, categoricamente distintos, existe grande similaridade em suas estruturas peptídicas, como pode ser visualizada na sobreposição dos cristais apresentada na figura 13. As 3 estruturas da *smHDAC8* estão representadas nas cores amarelo, azul e verde.

Não necessariamente estas estruturas necessitam de extrema similaridade. Porém, é de grande valia que o grau de similaridade seja alto para que os resultados das simulações de *docking* sejam precisos o suficiente para gerar conformações de ligantes que se sobreponham. Na figura 13, é perceptível um notável alinhamento entre as estruturas, principalmente em relação aos aminoácidos que cercam o sítio ativo. Podem ser vistos as conformações de folha beta e alfa hélice sobrepostas. A título de exemplo, pode-se observar 3 resíduos de histidina, pertencentes a cada uma das estruturas, sobrepostos, assinalados por setas de cor preta. Desta forma, foi possível realizar os estudos de *docking* com cristais diferentes sem alterações significativas dos resíduos de aminoácidos que circundam o sítio ativo.

Figura 13. Sobreposição dos cristais da enzima *smHDAC8* escolhidos para realização do estudo de *docking*.

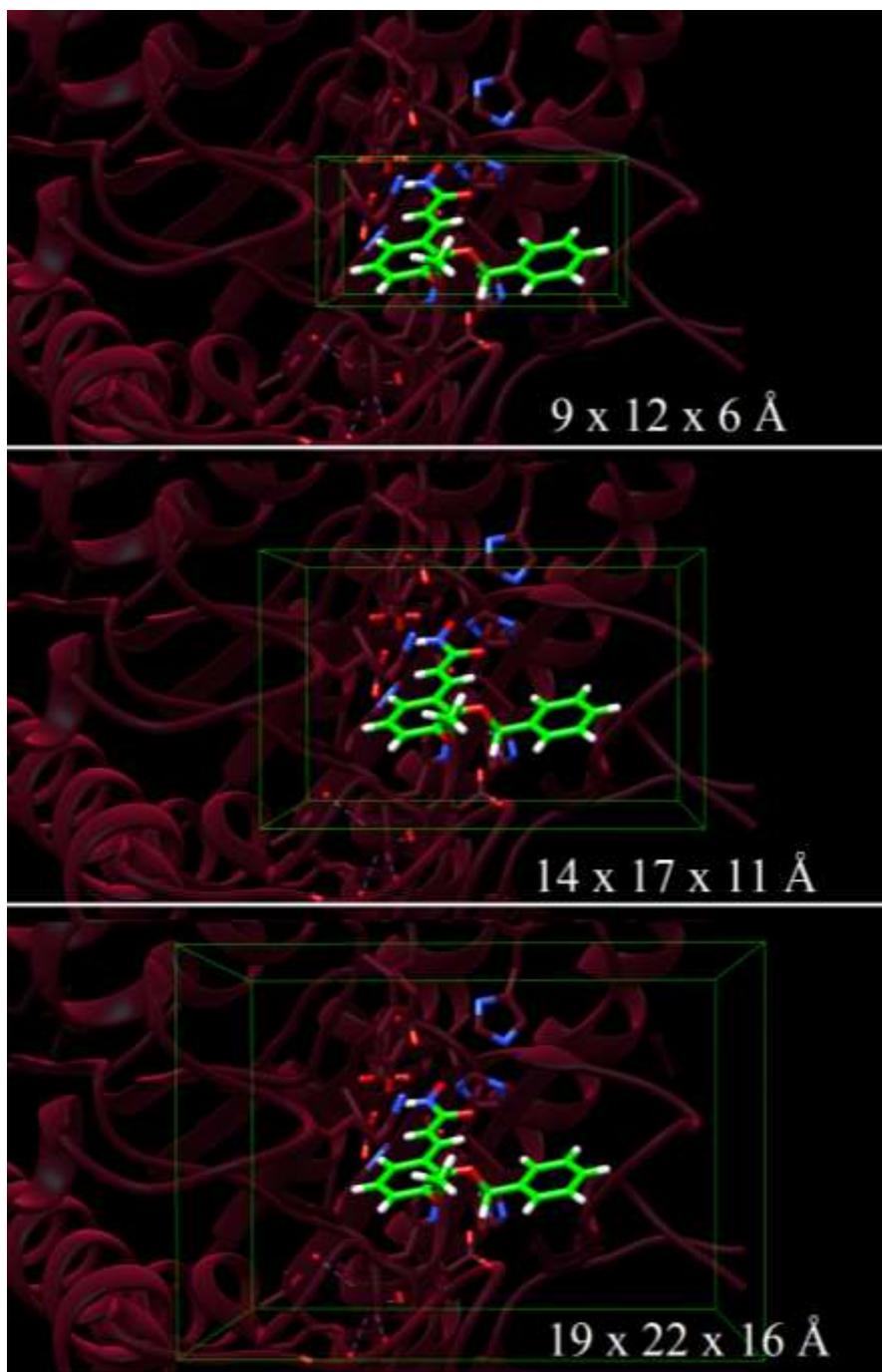


Fonte: elaborado pelo autor.

4.3 Otimização e alinhamento dos modelos moleculares

Após o cálculo das conformações 3D de todos compostos e padronização das estruturas proteicas, foi realizado o estudo de *docking*. Para tanto, foram utilizados 3 tamanhos de caixa diferentes. Por mais semelhantes que as moléculas sejam em seu esqueleto principal, existem moléculas pequenas formadas por 10 átomos, até moléculas mais extensas contendo 25 átomos ou mais. Desta forma, em uma caixa muito pequena moléculas grandes poderiam não ter espaço suficiente para que fosse verificada de forma correta as conformações obtidas por *docking*. Da mesma forma, moléculas menores, em caixas grandes, poderiam fugir ao sítio ativo, devido às próprias inconsistências dos algoritmos usados. De qualquer forma, foi primordial para escolha das conformações a semelhança com o ligante co-cristalizado com a enzima *smHDAC8*. Os tamanhos de caixa escolhidos podem ser visualizados na figura 14

Figura 14. Tamanhos selecionados para caixas nas simulações de *docking*.



Legenda: Na cor vinho: enzima *smHDAC8*. Linhas finas verdes: limites dimensionais da caixa.
Bastões em verde limão: ligante.

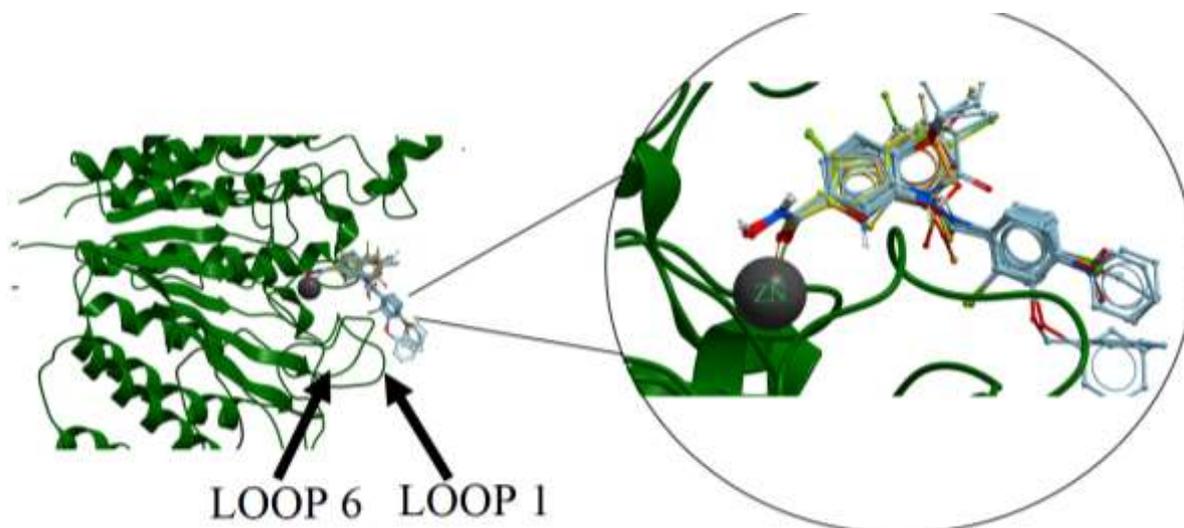
Fonte: elaborado pelo autor.

Outro fator que aprimorou a metodologia de *docking* utilizada neste trabalho foi o emprego de mais de um *software* para geração das conformações. Ao todo, foram 219 simulações (3 tipos de caixa para cada molécula) em 3 softwares diferentes, resultando num total de 657 simulações de *docking*. Cada simulação gerou 10 conformações. Cada conformação foi analisada minuciosamente, sem exceção. No total, foram realizadas 6.570 análises, inspecionadas visualmente, para se chegar a 73 conformações (1 para cada inibidor) prediletas.

Como já mencionado, as moléculas utilizadas no estudo foram analisadas quanto a conformação adquirida e não conforme ao *score* obtido por *docking*. Isto se dá pelo fato de que para um estudo de QSAR, a escolha de uma conformação análoga à conformação do ligante natural da série – molécula contida no depósito cristalográfico da proteína no site PDB – ser mais importante do que uma conformação com alto *score* de *docking* (THILAGAVATHI, 2003). Portanto, selecionando cada conformação mais relevante obteve-se um *pool* de conformações, que pode ser visualizado na Figura 15.

A conformação escolhida se torna um dos pontos chave para obtenção de um modelo preditivo. Pois é a partir das coordenadas espaciais de cada ligante que são calculados os descritores moleculares que se traduzem em pontos de interação e por conseguinte serão visualizados no espaço 3D do sítio ativo da enzima.

Figura 15. Alinhamento das conformações após *docking* inclusas no sítio ativo da enzima *smHDAC8*.



Legenda: ligantes em formato de bastão. Enzima *smHDAC8* em forma de bastão e folhas alfa e beta. Elemento zinco em forma de bola na cor cinza. Em verde escuro: cadeias da enzima *smHDAC8*. Em vermelho: átomos de oxigênio. Em azul: átomos de nitrogênio. Em verde: átomos de cloro. Nesta visualização, todas outras cores não citadas, representam cadeias de carbono.

Fonte: Elaborado pelo autor.

Os inibidores se encaixaram no sítio ativo com semelhança aos seus respectivos ligantes cristalografados, bem como apresentaram um modo de interação já relatado em estudos anteriores (MAREK *et. al.*, 2018). Devido à similaridade conformacional entre os ligantes submetidos ao *docking* e os ligantes contidos nos cristais (do site PDB) da enzima *smHDAC8*, é sugestivo que ambos apresentem interações intermoleculares similares. Dentre estas interações, pode ser citada a interação com o elemento zinco, essencial à atividade biológica. Outra interação, ocorre na porção contrária da molécula que interage com o elemento zinco; esta região da enzima, é subdividida em dois *loops*, os quais são denominados de *loop 1* e *loop 6*. A presença de grupos hidrofóbicos nestas regiões possibilitam estabilidade

conformacional à molécula, de tal forma que ela permaneça próxima ao elemento zinco.

Estes dados sustentam relevante resultado para os estudos de *docking*, pois denotam que as conformações obtidas, coincidem com as conformações obtidas por cristalografia de raios-x, dos cristais da enzima *smHDAC8* depositados no PDB. Conseqüentemente, foram observadas analisando a distância entre os resíduos de aminoácidos e os grupos funcionais de cada molécula, interações intermoleculares semelhantes às citadas nos artigos publicados anteriormente. Assim, as conformações obtidas podem ser usadas com segurança para cálculo dos descritores.

4.4 Cálculo dos descritores e filtro de variáveis

As conformações de *docking* obtidas foram submetidas ao processo de cálculo de descritores moleculares com o software Open3DQSAR. Foram obtidos 32.330 descritores moleculares para cada potencial de LJ e QQ. Foram observados diversos descritores redundantes, dispostos em pontos muito próximos um do outro, ou muito distantes das estruturas moleculares, por exemplo. Para tais inconsistências, se fez necessário o uso de filtros de variáveis. Algoritmos em *shell script* já previamente validados foram aplicados.

Como estes descritores moleculares se traduzem em pontos de interações intermoleculares, sabe-se que, para ocorrer uma interação intermolecular entre um ligante e um resíduo de aminoácido, é necessária uma distância ideal entre as duas espécies químicas. Além disto, descritores dispostos no mesmo raio, muito próximos um do outro, acabam por descrever propriedades idênticas, e, portanto, precisam ser excluídos antes da etapa de regressão linear múltipla. Através destes filtros, foi possível eliminar descritores com baixa variância (ponto de corte de 0,02), ou seja, descritores excessivamente distantes das estruturas dentro da caixa virtual. Descritores com alta correlação (ponto de corte 0,03), ou seja, muito próximos um do outro, também foram descartados. Para os descritores de LJ foram descartados os que possuíam valores absurdos, que não condizem com a realidade (> 30 kcal/mol). Após os filtros, o número de descritores diminuiu para 11.300 para LJ e 8.580 para

QQ. Assim estes valores foram correlacionados com os dados de atividade biológica através da regressão linear multivariada.

4.5 Distinção dos conjuntos de treinamento (interno) e de teste (externo)

Com o dendograma gerado pelo software ICM (figura 16), foi possível escolher um conjunto de 22 inibidores para compor o conjunto externo (validação) e, conseqüentemente, 51 moléculas compuseram o conjunto interno (treinamento). No dendograma, as moléculas estão separadas por clusters de similaridade estrutural e de atividade biológica. A ideia nesta etapa foi incluir no conjunto externo a maior variabilidade estrutural possível. Cada ramo terminal presente no dendograma, representa um inibidor. Quanto mais próximo está um ramo do outro, mais similares são as estruturas. As cores indicam os diferentes clusters, ou melhor, os grandes ramos. Então, dentro de cada grande ramo, foi escolhido um número arbitrário de moléculas, de modo que o conjunto de 22 inibidores compreendesse todos os ramos. Vale ressaltar que, durante a escolha das estruturas, não se sabia qual molécula estava sendo escolhida, apenas que se tratava de moléculas distintas por estrutura e atividade. Deste modo, evitando-se uma escolha tendenciosa de amostras. As moléculas escolhidas para compor o conjunto externo, segundo a numeração ID estabelecida no **ANEXO1**, foram: 1, 6, 7, 10, 13, 17, 22, 24, 27, 30, 32, 38, 43, 45, 47, 53,54, 57, 62, 64, 68 e 72. Em consequência, as 51 moléculas restantes compreenderam o conjunto de treinamento. Desta forma, as moléculas do conjunto de treinamento, bem como seus respectivos descritores moleculares e atividades biológicas, foram utilizados para construção do modelo de QSAR-3D. O conjunto externo foi utilizado na etapa de validação do modelo.

4.6 Criação e validação dos modelos de QSAR 3D

Através da ferramenta *Nanobridges* foi possível gerar vários modelos de QSAR, sendo necessário apenas inserir a matriz na qual estão dispostos - os dados dos descritores moleculares no eixo X e, as atividades biológicas no eixo Y. É importante enfatizar que, para alimentar esta matriz com valores de atividade biológica, os quais estão na razão de IC_{50} (quanto menor o IC_{50} , mais potente é a molécula), é preciso calcular o inverso do logaritmo (antilog) do valor de IC_{50} , sendo denominado como pIC_{50} . Desta forma, os valores de atividades potentes (concentrações baixas) se tornaram valores altos e, os valores de atividades fracas (concentrações altas) se tornaram valores baixos. De forma prática, foi construída uma matriz na qual possuiu 51 linhas (valores de pIC_{50} das 51 moléculas do conjunto interno) e milhares de colunas (descritores calculados para cada molécula). A partir disto, exaustivos cálculos de regressão linear múltipla correlacionaram os dados dos descritores com as atividades biológicas. Foram gerados diversos modelos. Porém, foi preciso averiguar qual destes obedeciam aos padrões da OECD para serem utilizados no estudo. O modelo foi escolhido de forma categórica através dos testes de validação. O melhor modelo elegido apresentou 4 descritores moleculares, os quais estão dispostos na Tabela 2.

Tabela 2. Descritores moleculares referentes ao melhor modelo de QSAR-3D calibrado.

Descritor	Coefficiente de regressão	Interpretação rápida
LJ+ 1 (VWP)	0,03702	Interação favorável com Aspartato 100
QQ+ (QQP)	0,87106	Hidroxamida essencial para atividade
LJ+2 (VWP)	0,01796	Interação de empilhamento Pi favorável
LJ+ 3 (VWP)	0,01749	Presença de Anéis aromáticos favorável à interação

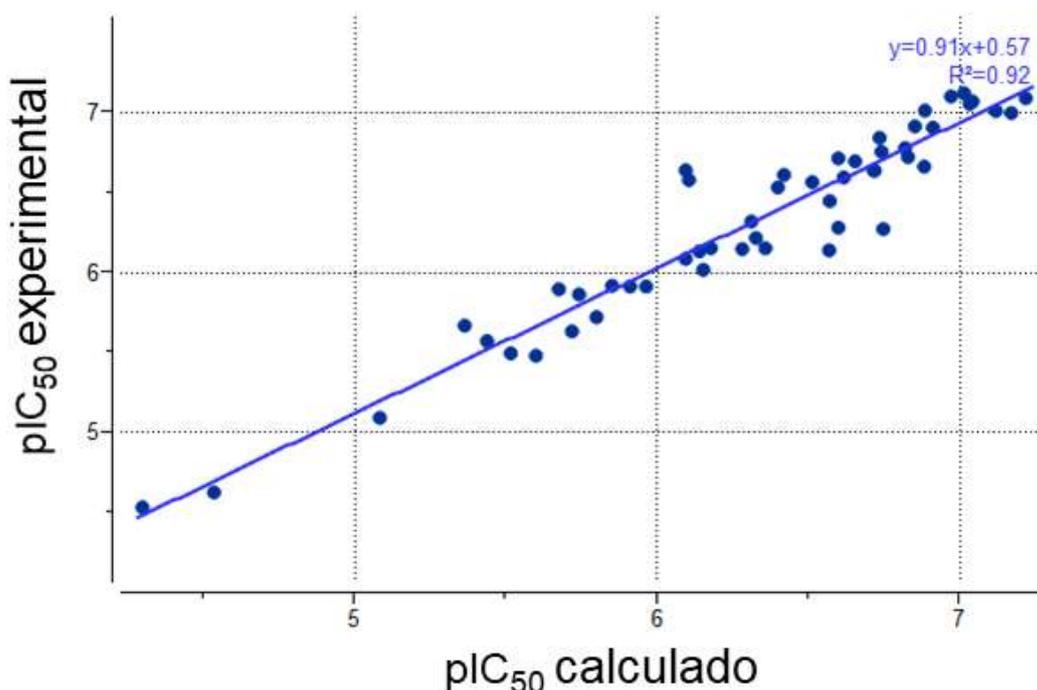
Legenda: VWP, descritor de Vander Walls positivo; QQP, descritor de Coulomb positivo.

Fonte: elaborado pelo autor.

Em seguida, foi avaliado se o modelo apresentava medidas aceitáveis de qualidade de ajuste, robustez e previsibilidade, através de testes estatísticos (R^2 , Q^2 externo, *y-randomization*, *Leave-One-Out* (LOO) e *leave-N-Out* (LNO)).

A figura 17 apresenta os resultados para o teste de R^2 . O gráfico demonstra as atividades experimentais do conjunto interno, plotadas às atividades calculadas pelo modelo de QSAR 3D.

Figura 17. pIC₅₀ experimental *versus* pIC₅₀ calculado para o conjunto interno de 51 moléculas.



Legenda: Pontos em azul, valores de pIC₅₀ do conjunto interno, calculados pelo modelo de QSAR 3D.

Fonte: Elaborado pelo autor.

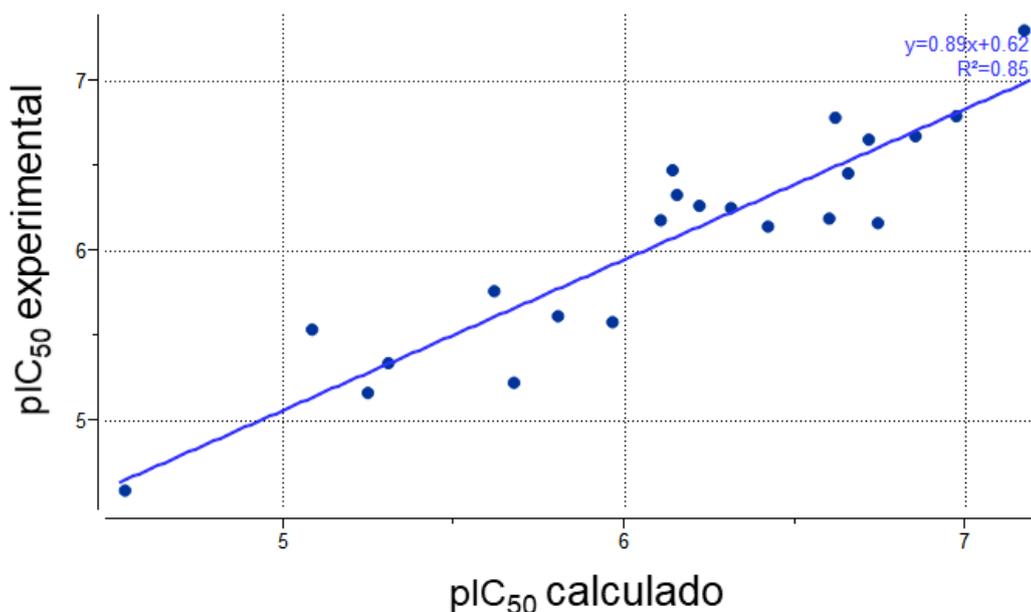
Cada ponto em azul representa a atividade biológica (pIC₅₀) calculada pelo modelo para cada molécula, as quais compõem o conjunto interno. Quanto mais esse valor calculado se aproxima do valor experimental, maior será a proximidade em relação a reta, plotada no gráfico e, conseqüentemente, mais próximo de 1,00 será o valor de R².

Um modelo no qual apresente valor de R² inferior a 0,7, é considerado inadequado para estudos de QSAR. Portanto, não pode ser utilizado para etapas

seguintes do estudo. Vale ressaltar que, como as estruturas submetidas a esta previsão (R^2), são as mesmas utilizadas para construção do modelo, existe grande possibilidade de ocorrer sobre-ajuste do modelo. Sobre-ajuste (do inglês: *overfitting*), é um termo usado em estatística para descrever quando um modelo estatístico se ajusta muito bem ao conjunto de dados anteriormente observado, mas se mostra ineficaz para prever novos resultados.

O modelo apresentou um valor R^2 de 0,92, considerado extremamente satisfatório. No entanto, como estas moléculas (51 moléculas do conjunto interno) foram utilizadas para construção do modelo, pode ocorrer que este modelo de QSAR 3D esteja sobre-ajustando os dados. Em outras palavras, é fácil que um modelo preveja dados para amostras nas quais ele foi treinado. portanto, mesmo o modelo de QSAR tendo apresentado valores adequados de R^2 para o conjunto interno, foi necessário uma validação cruzado com o conjunto externo, no qual é mensurada pelo valor de Q^2 externo (previsão de atividade biológica para o conjunto externo: 22 moléculas que não foram utilizadas para construção do modelo). A figura 18 apresenta o resultado para o teste de Q^2 externo.

Figura 18. Valores de pIC_{50} experimental *versus* valores de pIC_{50} teóricos, calculados pelo modelo de QSAR 3D para o conjunto externo de 22 moléculas.

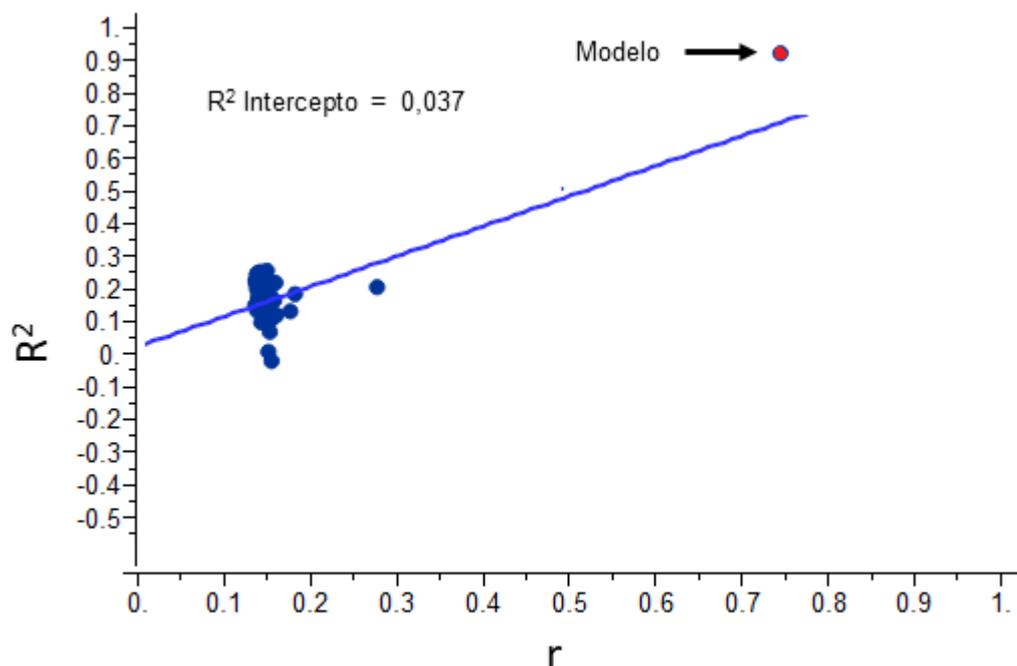


Legenda: Pontos em azul, valores de pIC_{50} calculados (conjunto externo) pelo modelo de QSAR 3D.

Fonte: elaborado pelo autor.

O modelo obtido apresentou um valor de Q^2 externo de 0,85, significando que ele é capaz de prever a atividade biológica para um conjunto desconhecido de moléculas, ou seja, é qualificado a realizar previsões com dados, os quais ele não foi treinado. No entanto, ainda é necessário buscar padrões mais rebuscados para valorização do modelo de QSAR. Para isto, se fez necessário realizar mais testes para que o modelo se enquadre aos padrões exigidos nas normas técnicas da OECD. Sendo assim, mesmo obtendo um modelo com valores de Q^2 externo satisfatórios, ainda pode-se explorar mais a qualidade do modelo, como por exemplo nos testes de *Y-randomization*, *Leave-One-Out (LOO)* e *leave-N-Out (LNO)*. O intercepto da reta para o teste de *y-randomization* está disposto na figura 19.

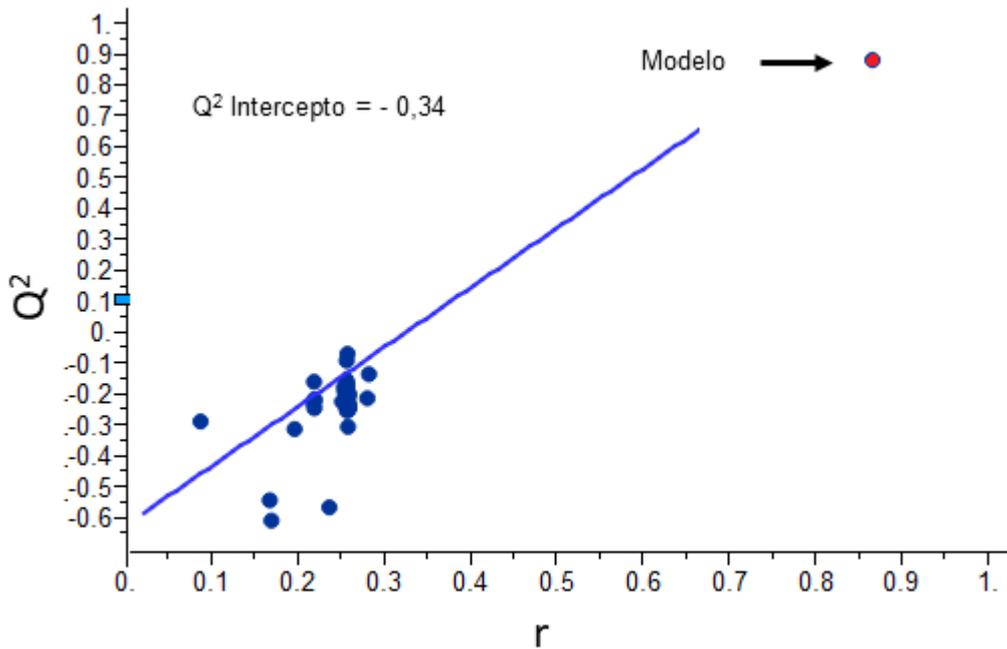
Figura 19. Intercepto da reta para os valores de R^2 do *y-randomization*.



Fonte: Elaborado pelo autor.

No gráfico exposto acima, cada ponto em azul representa o valor de R^2 , para cada modelo reconstruído, com as atividades biológicas embaralhadas N vezes o número de amostras do conjunto interno. O modelo original, o qual apresentou valor de R^2 de 0,92, está representado por um ponto de cor vermelha. É notório que os valores calculados foram ruins ($<0,30$), como era de se esperar em um modelo estatístico, no qual não há correlação com o acaso. Foi observado um valor de 0,037 para o intercepto da reta. Pode-se considerar o resultado satisfatório, pois a literatura enfatiza que este valor não deve ultrapassar 0,3. Desta maneira, o modelo de QSAR 3D obtido não prevê a atividade biológica ao acaso. Outro teste que denota as chances do modelo de QSAR apresentar *overfitting*, é o teste de LOO, mensurado pelo valor de Q^2 LOO. O resultado para o teste de LOO, está disposto na figura 20.

Figura 20. Intercepto da reta para os valores de Q^2_{Loo} .

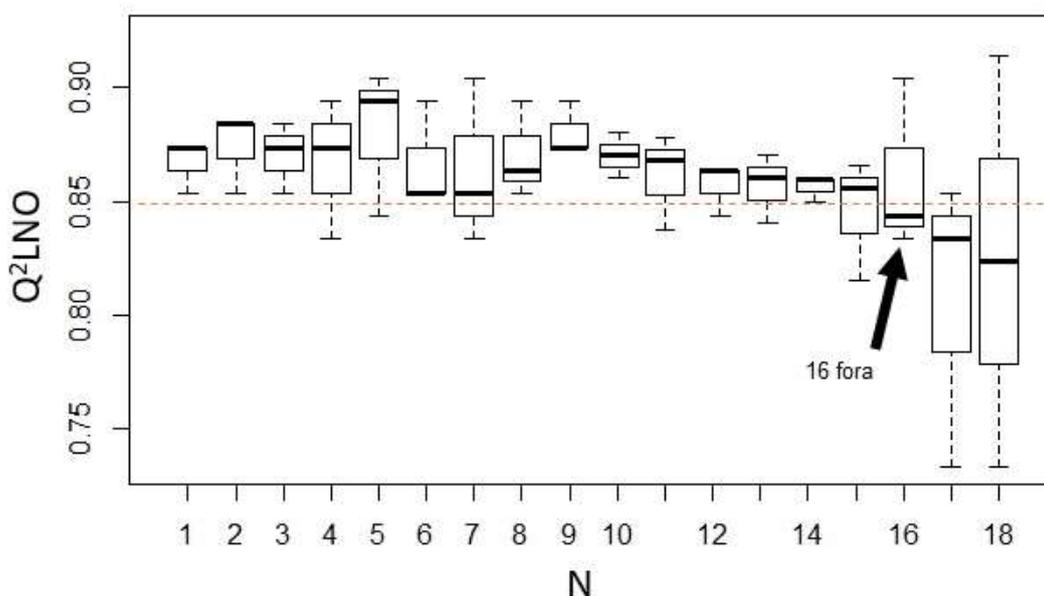


Fonte: Elaborado pelo autor.

De forma semelhante, o valor de -0,34 foi satisfatório. Este teste enfatiza ainda mais que o modelo não possuiu sobre-ajuste.

A OECD explica que os modelos de QSAR devem ser robustos. Para analisar tal fato, foi realizado o teste LNO, mensurado pelo valor de Q^2_{LNO} . Na figura 21 está apresentado os resultados para o teste de Q^2_{LNO} .

Figura 21. Amostragem para o teste de validação Q^2_{LNO} .



Fonte: elaborado pelo autor.

Na figura 21, no eixo X encontra-se o número de moléculas retiradas do modelo, e no eixo Y o valor de Q^2 externo calculado. A linha tracejada representa o ponto de corte, para que o modelo seja considerado aceitável. É possível observar que mesmo retirando 15 amostras do modelo, ele conseguiu calibrar dentro do desvio permitido. Após a retirada de 16 moléculas, o valor decaiu. A literatura exige que, para que um modelo de QSAR seja considerado robusto, é necessário que o mesmo apresente um desvio máximo de 0,01, retirando ao menos, 20% das amostras (KIRALJ; FERREIRA, 2009). Diante dos resultados apresentados, pode-se considerar o modelo obtido robusto.

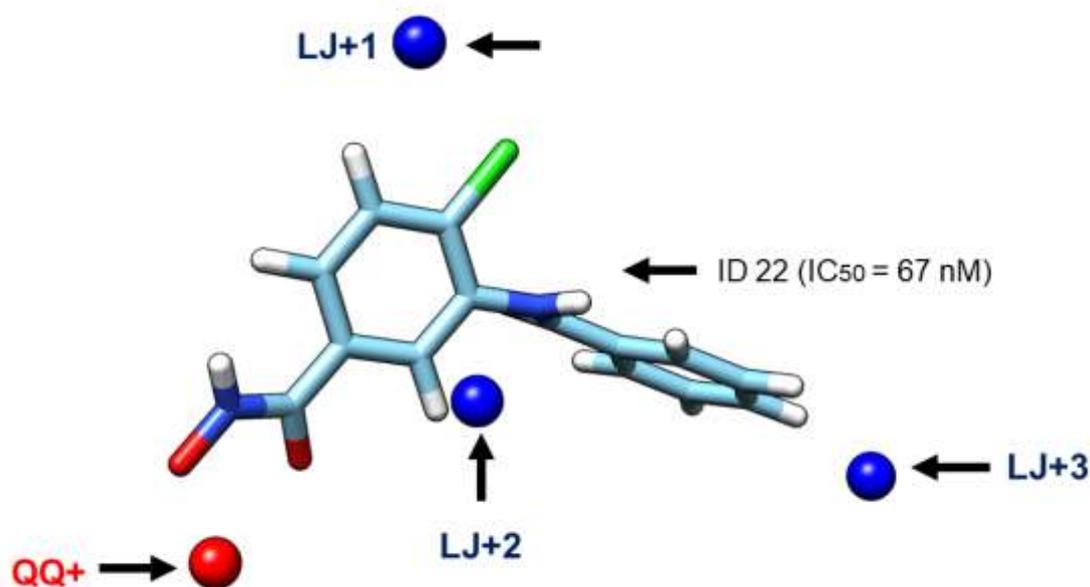
Ainda citando a OECD como padrão ouro de referência para validação do modelo de QSAR-3D, outras abordagens foram verificadas, no intuito de se verificar a adequação do modelo gerado.

A partir do momento que o modelo de QSAR-3D foi validado, é confiável utilizar de seus dados para prever atividade biológica dos compostos e usá-lo para o design racional de novos fármacos. Uma forma adequada para auxiliar o design é a interpretação dos descritores moleculares obtidos com o modelo de QSAR 3D.

4.7 Interpretação mecanicista dos descritores moleculares

Ao analisar a disposição dos descritores no plano 3D, foi possível observar inicialmente como estes descritores estavam dispostos ao redor de uma molécula da série (figura 22). Para fins de melhor visualização, nesta imagem em específico, foi colocada apenas uma molécula com boa potência (ID 22, do **ANEXO 1**), a qual possui grupos moleculares que facilitam a exposição dos resultados.

Figura 22. Representação espacial dos descritores moleculares.



Legenda. Descritor obtido através da equação do potencial de Coulomb está representando por uma esfera na cor vermelha. Os descritores obtidos através da equação de potencial de Lennard-Jones, estão representados por esferas na cor azul.

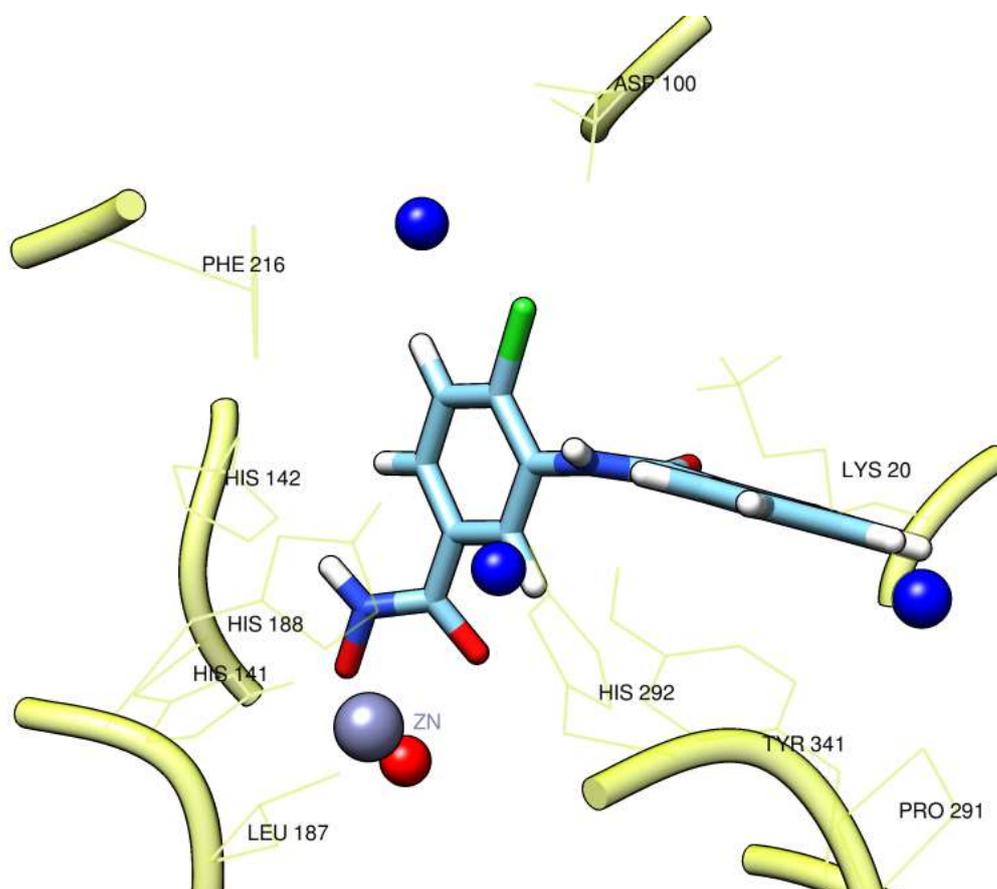
Fonte: elaborado pelo autor.

Assim, partir desta análise foi possível observar quais grupos moleculares ficaram próximos aos descritores obtidos no modelo de QSAR-3D. E melhor, sabendo

que cada molécula apresenta um valor calculado diferente, para cada um destes descritores, é possível analisar, para cada descritor em específico, qual molécula apresentou o melhor ou pior valor para tal descritor.

Também foi possível visualizar que os descritores obtidos estavam próximos de resíduos de aminoácidos essenciais para atividade biológica de inibidores *smHDAC8* (figura 23).

Figura 23. Representação dos descritores moleculares no sítio ativo da enzima.



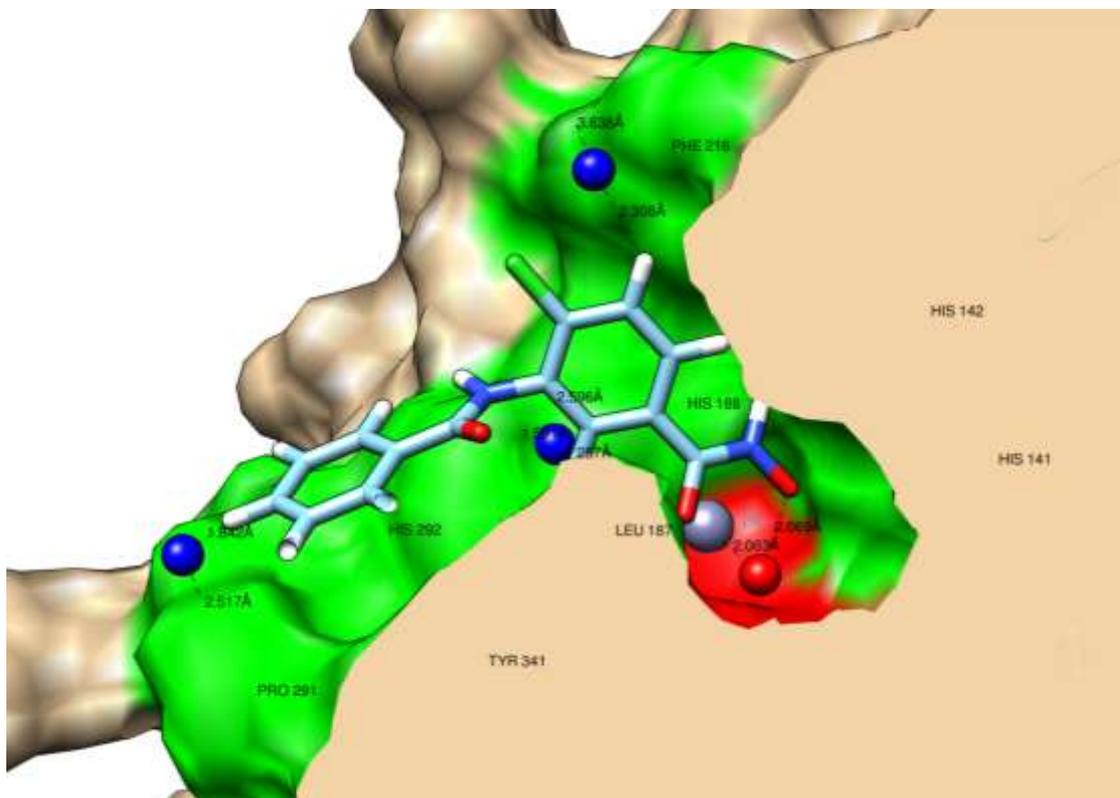
Legenda: Em azul claro: átomos de carbono do composto **ID 22**. Em branco: átomos de hidrogênio. Em vermelho: átomos de oxigênio. Em azul escuro: átomos de nitrogênio. Em verde incandescente: átomo de cloro. Em cinza: elemento zinco. Em amarelo claro: resíduos de aminoácido da *smHDAC8*.

Fonte: elaborado pelo autor.

É possível observar (figura 23), de uma forma geral, a posição dos descritores, relacionando-os não só à molécula ligante (azul), mas também aos resíduos de aminoácidos (por exemplo: Leucina 187, Histidina 292 e Tirosina 341), os quais fazem interação com esta série de inibidores. Em uma visão geral, é notório que, quando plotados no plano espacial os descritores se ajustaram entre ligante e sítio ativo da enzima, estabelecendo uma interpretação mecanicista destes dados.

Ao aplicar a área superficial nos resíduos de aminoácidos, foi possível constatar que alguns descritores apresentaram disposições espaciais, as quais coincidem com cavidades específicas do sítio ativo, estas extremamente importantes para atividade biológica desta série de compostos (Figura 24).

Figura 24. Representação dos descritores moleculares considerando a área superficial da enzima.



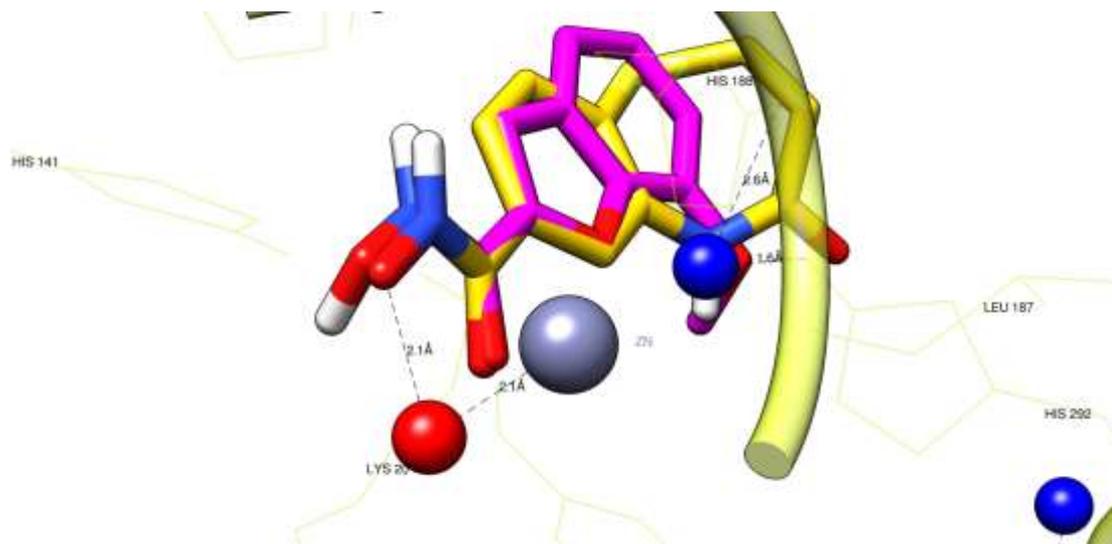
Legenda: Em forma de bastão: ligante. Em área superficial: enzima *smHDAC8*. Em azul claro: átomos de carbono do composto **ID 22**. Em branco: átomos de hidrogênio. Em vermelho: átomos de oxigênio. Em azul escuro: átomos de nitrogênio. Em verde incandescente: átomo de cloro. Em cinza: elemento zinco. Área superficial verde: área dos aminoácidos importantes para interação. Área superficial vermelha: cavidade na qual está contida o elemento zinco. Esferas em azul: descritores de LJ. Esferas em vermelho: descritores QQ.

Fonte: elaborado pelo autor.

O sítio ativo apresenta uma cavidade (em vermelho) na qual se encontra o elemento zinco, que possui significativa relevância à atividade biológica destes compostos e, nesta mesma cavidade, está presente um descritor QQ positivo. Já é descrito na literatura que os ligantes com atividade inibitória à *smHDAC8* necessitam interagir neste bolso da enzima.

O descritor QQ+, representado por uma esfera vermelha, o qual no plano espacial ficou próximo ao elemento de zinco, revelou que o grupo hidroxamato, na posição indicada, está diretamente relacionada com a variação da atividade biológica (figura 25).

Figura 25. Disposição espacial em detalhe do descritor QQ+.



Legenda: Em forma de bastão: ligantes. Em linhas finas: resíduos de aminoácidos da enzima *smHDAC8*. Em rosa: átomos de carbono do composto **ID 54**. Em amarelo ouro: átomos de carbono do composto **ID 49**. Em branco: átomos de hidrogênio. Em vermelho: átomos de oxigênio. Em azul escuro: átomos de nitrogênio. Em cinza: elemento zinco. Esferas em azul: descritores de LJ. Esferas em vermelho: descritores QQ.

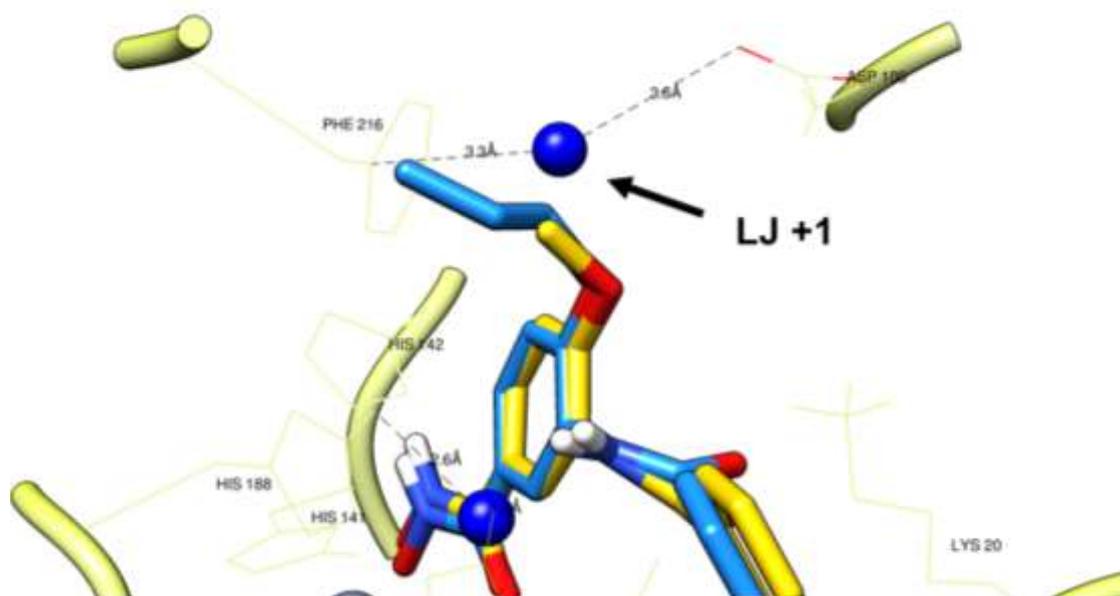
Fonte: elaborado pelo autor.

Na figura 25, estão representadas duas moléculas: a molécula **ID 49** (amarelo), a qual apresentou valores altos para o descritor QQ+ (esfera em vermelho), e a molécula **ID 54** (rosa), a qual apresentou valores baixos para este mesmo descritor. Analisando a conformação das duas estruturas neste ponto (próximo ao descritor QQ+), percebe-se que o **ID 49** ($IC_{50} = 1220$ nM) possui o átomo de oxigênio terminal do grupo hidroxamato ionizado. Já o **ID 54** ($IC_{50} = 2100$ nM), apresenta este grupo na forma molecular. Neste sentido, foi observado que grupos ionizados com cargas negativas favorecem seguramente a atividade biológica no referido ponto da enzima

smHDAC8. Portanto, este resultado ficou em concordância com os dados contidos na literatura, os quais, correlacionam a atividade dos inibidores com o grupo hidroxamato ionizado, devido sua avidez em coordenar com o metal zinco, como também pela capacidade de realizar interações intermoleculares do tipo ligação de hidrogênio com os aminoácidos adjacentes. Isto denota que a natureza descritiva deste descritor está estritamente relacionada com os dados de atividade biológica.

Na figura 26 está representado detalhadamente o descritor LJ+1 (esfera azul) que está próximo aos resíduos de Aspartato (ASP 100) e fenilalanina (PHE 216). Para este descritor, foi observado uma diminuição da atividade biológica em moléculas as quais apresentavam grupos volumosos hidrofóbicos voltados ao resíduo de fenilalanina 216.

Figura 26. Representação espacial detalhada do descritor LJ+1.



Legenda: Em forma de bastão: ligantes. Em linhas finas: resíduos de aminoácidos da enzima *smHDAC8*. Em azul turquesa: átomos de carbono do composto **ID 26**. Em amarelo ouro: átomos de carbono do composto **ID 20**. Em branco: átomos de hidrogênio. Em vermelho: átomos de oxigênio. Em azul escuro: átomos de nitrogênio. Em cinza: elemento zinco. Esferas em azul: descritores de LJ.

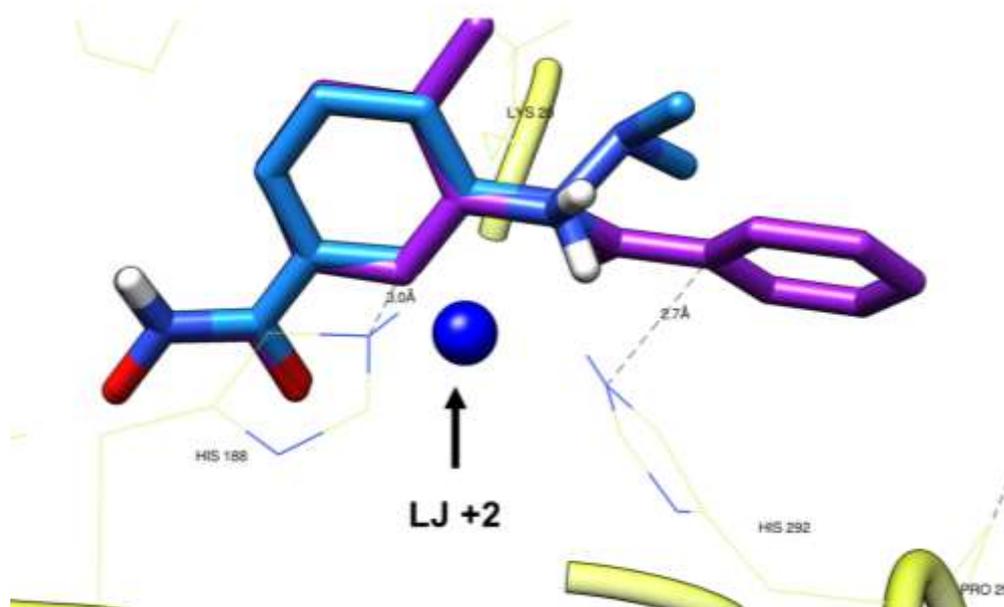
Fonte: Elaborado pelo autor.

Os derivados que detinham grupos volumosos hidrofóbicos na posição “*para*” em relação ao grupo hidroxamato apresentaram valores baixos no referido descritor (LJ+1). Já os compostos menos volumosos, apresentavam valores mais altos. A molécula **ID 20** em amarelo, ($IC_{50} = 190$ nM) apresentou valores altos para o referido descritor e, a molécula **ID 26** em azul, ($IC_{50} = 267$ nM) apresentou valores baixos. Desta forma, sugere-se que neste ponto, os grupos volumosos se voltam ao resíduo hidrofóbico da fenilalanina (PHE 216). Neste caso, pode ser levantada a hipótese de que a interação com a fenilalanina, diminui a atividade das moléculas por dificultar o encaixe dos inibidores no sítio ativo. É interessante observar também que o descritor se encontra próximo a um resíduo de aspartato (ASP 100), que é um aminoácido capaz de realizar interações intermoleculares do tipo ligação de hidrogênio. Desta forma, esta informação pode ser considerada no design racional, inserindo grupos neste ponto, que possam contemplar este tipo de interação.

Ainda vale ressaltar que substituições com o átomo de cloro na posição “*para*” resultam em uma melhora nos valores de atividade biológica. Esta substituição não diz respeito a nenhum tipo de interação entre o heteroátomo e resíduos de aminoácidos próximos, mas sim em relação ao efeito indutivo causado no anel de benzeno ligado diretamente ao grupo hidroxamato, favorecendo o encaixe da molécula como um todo no sítio. Desta vez, diferentemente do descritor abordado anteriormente (QQ+), é algo inédito para esta série de inibidores, uma vez que não se têm descrito na literatura este tipo de abordagem neste ponto de interação na enzima *smHDAC8*.

Para o descritor de LJ+2, existem dois resíduos de histidina que formam uma camada superficial e facilita interações de empilhamento Pi (figura 27).

Figura 27. Representação espacial detalhada do descritor LJ+2.



Legenda: Em forma de bastão: ligantes. Em linhas finas: resíduos de aminoácidos da enzima *smHDAC8*. Em roxo: átomos de carbono do composto **ID 18**. Em azul turquesa: átomos de carbono do composto **ID 14**. Em branco: átomos de hidrogênio. Em vermelho: átomos de oxigênio. Em azul escuro: átomos de nitrogênio. Esferas em azul: descritores de LJ.

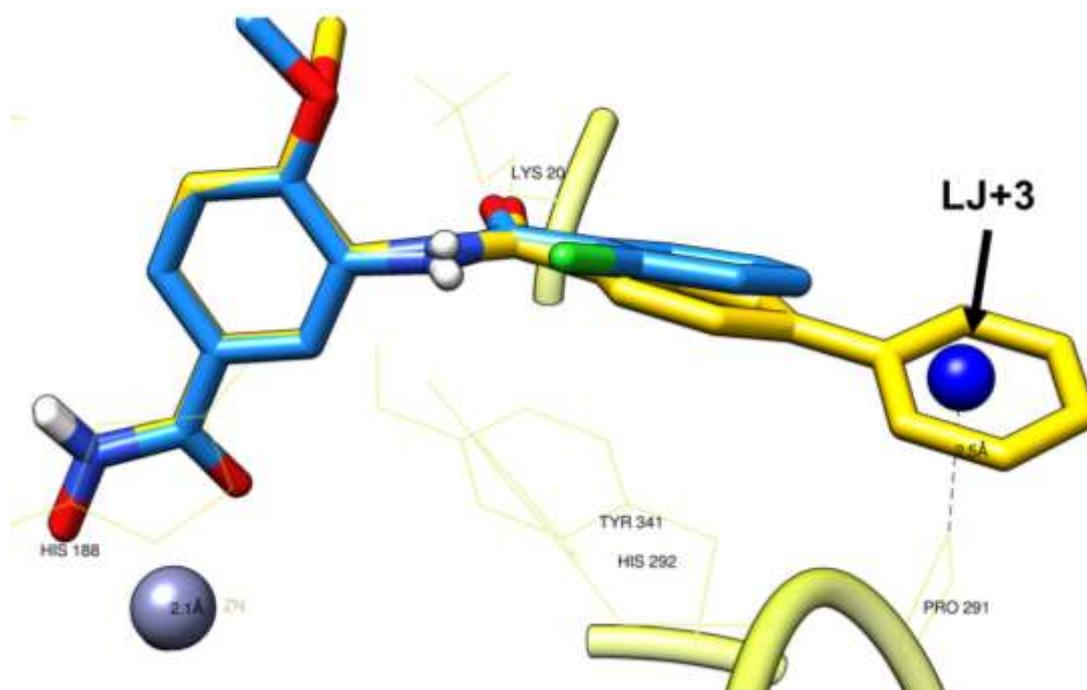
Fonte: elaborado pelo autor.

A molécula **ID 18** ($IC_{50} = 468$ nM), representada na cor roxa, apresentou maiores valores para o descritor LJ+2 em comparação a molécula **ID 14** ($IC_{50} = 3630$ nM) representada na cor azul turquesa. É perceptível que o descritor se encontra entre dois resíduos de histidina (HIS 188 e HIS 292). A histidina é um aminoácido que interage com ligantes, preferencialmente, através de interações intermoleculares de empilhamento Pi. Sabendo que anéis de benzeno apresentam capacidade de realizar tal tipo de interação, notou-se que a molécula a qual possui um anel benzeno adicional (**ID 18**), apresenta não só melhores valores para o descritor relatado, como também melhores valores de atividade biológica, comparando-a com outra molécula que não

contém este anel aromático adicionalmente (**ID 14**). A partir desta análise, supõe-se que interações de empilhamento *Pi* nesta parte da molécula, favorecem o aumento da atividade biológica de inibidores *smHDAC8*.

O descritor de LJ+3 localizado próximo ao resíduo de prolina 291 (figura 28), foi relacionado com a de anéis conjugados com capacidade de interagir com este resíduo de aminoácido de característica hidrofóbica. Derivados com grupos que não alcançavam o resíduo de prolina, apresentaram valores baixos para este descritor.

Figura 28. Representação espacial detalhada do descritor LJ+3.



Legenda: Em forma de bastão: ligantes. Em linhas finas: resíduos de aminoácidos da enzima *smHDAC8*. Em amarelo ouro: átomos de carbono do composto **ID 29**. Em azul turquesa: átomos de carbono do composto **ID 32**. Em branco: átomos de hidrogênio. Em vermelho: átomos de oxigênio. Em azul escuro: átomos de nitrogênio. Em cinza: elemento zinco. Esfera em azul: descritor de LJ.

Fonte: elaborado pelo autor.

Na figura 28 o descritor LJ+3 está evidenciado por uma seta. A molécula **ID 29** ($IC_{50} = 75,4$ nM) está representada na cor amarela. O inibidor **ID 32** ($IC_{50} = 699$ nM) está evidenciado na cor azul. É perceptível a diferença entre as duas moléculas citadas, no que diz respeito a adição de um anel benzeno. Os valores encontrados para o descritor LJ+3 foram altos para o **ID 29**, e baixos para o **ID 32**. É perceptível que o referido descritor se encontra entre um resíduo de prolina e um anel de benzeno do ligante **ID 29**. A prolina possui capacidade de realizar interações de van der Waals com ligantes que possuem grupos hidrofóbicos. Desta forma, supõe-se que neste ponto, a presença de bifenilas está relacionada com o aumento da atividade inibitória frente a enzima *smHDAC8*.

A interpretação mecanicista dos descritores forneceu dados valiosos para o desenho de novas estruturas. Foi possível identificar não só grupos funcionais que se relacionam melhor com a atividade biológica, como também grupos que não favorecem o aumento da atividade.

5 CONCLUSÃO

Todos os objetivos almejados foram alcançados com êxito e, a construção de um modelo de QSAR-3D robusto, foi realizada com vultoso exímio. É importante ressaltar a importância das ferramentas de modelagem molecular, como o QSAR, utilizadas no planejamento racional de novos compostos. Pois foi possível desenvolver um modelo para previsão de atividade biológica, de maneira célere e com baixo custo. Com a disponibilidade desta ferramenta, será possível testar a atividade biológica *in silico* de novos compostos projetados, com segurança. A perspectiva atrelada ao resultado deste trabalho é norteadada pelo desenvolvimento racional de novos inibidores, com previsão de atividade *in silico* (pelo modelo de QSAR-3D), síntese e testes de atividade biológica *in vitro*, para que os protótipos desenvolvidos em trabalhos posteriores, possam chegar às fases subsequentes do desenvolvimento de novos fármacos.

6 REFERÊNCIAS

A. R. Leach, *Molecular Modelling: Principles and Applications*, 2001, [ISBN 0-582-38210-6](#)

ANDREWS, K. T.; Haque, A.; Jones, M. K. HDAC inhibitors in parasitic diseases *Immunol. Cell Biol.* **2012**, 90,66– 77.

AMBURE, Pravin; AHER, Rahul Balasaheb; GAJEWICZ, Agnieszka; *et al.* “NanoBRIDGES” software: Open access tools to perform QSAR and nano-QSAR modeling. *Chemometrics and Intelligent Laboratory Systems*, v. 147, p. 1–13, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.chemolab.2015.07.007>>.

ARODOLA, Olayide A.; SOLIMAN, Mahmoud E.S. Quantum mechanics implementation in drug-design workflows: Does it really help? *Drug Design, Development and Therapy*, v. 11, p. 2551–2564, 2017.

AUGEN, Jeffrey. The evolving role of information technology in the drug discovery process. *Drug Discovery Today*, v. 7, n. 5, p. 315–323, 2002.

BAK, A. *et al.* RSC Advances guided pharmacophore mapping : investigation of activity profile for a series of drug absorption promoters †. *RSC Advances*, v. 6, p. 76183–76205, 2016.

BAN, Tomohiro; OHUE, Masahito; AKIYAMA, Yutaka. Multiple grid arrangement improves ligand docking with unknown binding sites: Application to the inverse docking problem. *Computational Biology and Chemistry*, v. 73, p. 139–146, 2018. Disponível em: <<https://doi.org/10.1016/j.compbiolchem.2018.02.008>>.

BANEGAS-LUNA, Antonio Jesús; CERÓN-CARRASCO, José P.; PÉREZ-SÁNCHEZ, Horacio. A review of ligand-based virtual screening web tools and screening algorithms in large molecular databases in the age of big data. *Future Medicinal Chemistry*, v. 10, n. 22, p. 2641–2658, 2018.

BAYER, Theresa; CHAKRABARTI, Aloкта; LANCELOT, Julien; *et al.* Synthesis, Crystallization Studies, and in vitro Characterization of Cinnamic Acid Derivatives as

SmHDAC8 Inhibitors for the Treatment of Schistosomiasis. **ChemMedChem**, v. 13, n. 15, p. 1517–1529, 2018.

Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância Epidemiológica. **Leishmaniose visceral : recomendações clínicas para redução da letalidade** – Brasília : Ministério da Saúde, 2011. 78 p. : il. – (Série A. Normas e Manuais Técnicos).

BRASIL. **Ministério da Saúde**. VIGILÂNCIA DA ESQUISTOSSOMOSE MANSONI Diretrizes Técnicas . Brasil, por Região e Unidade Federada. 1998 a 2015. Brasília, 2014,

Brasil. **Ministério da Saúde**. Secretaria de Vigilância em Saúde. Departamento de Vigilância Epidemiológica. Vigilância da Esquistossomose Mansonii: diretrizes técnicas / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância das Doenças Transmissíveis. – 4. ed. – Brasília: Ministério da Saúde, 2014.

BRASIL. **Ministério da Saúde**. Número de Óbitos por esquistossomose. Brasil, por Região e Unidade Federada. 1998 a 2015. Brasília, 2016.

BRASIL. **Ministério da Saúde**. Número de casos na área endêmica - Brasil, por Região e Unidade Federada. 2000 a 2016. Brasília, 2017.

BRASIL. Ministério da Saúde. **Saúde de A-Z. Esquistossomose: descrição da doença**. 2019a. Disponível em: < <http://portalms.saude.gov.br/saude-de-a-z/esquistossomose/11240-descricao-da-doenca>. Acesso em: 08 abr. 2019.

BRASIL. Ministério da Saúde. **Esquistossomose: causas, sintomas, tratamento e prevenção**. 2019b. Disponível em: <<https://http://portalms.saude.gov.br/saude-de-a-z/esquistossomose/>>. Acesso em: 18 fev. 2019.

CABY, Stéphanie; PAGLIAZZO, Lucile; LANCELOT, Julien; *et al.* Analysis of the interactome of *Schistosoma mansoni* histone deacetylase 8. **PLoS Neglected Tropical Diseases**, v. 11, n. 11, p. 1–15, 2017.

CALIXTO, João B.; JR., Jarbas M. Siqueira. Desenvolvimento de Medicamentos no Brasil: Desafios. **Gaz. méd. Bahia;78 (Suplemento 1):98-106**, v. 78, n. 1, p. 98–106, 2008. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.86.011108>>.

CORREIA, N.; GIRA, M. Construction of 4D-QSAR models for use in the design of novel p38-MAPK inhibitors. **Journal of Computer-Aided Molecular Design**, p. 385–400, 2005.

CRAMMER, R.D.; PATTERSON, D.E.; BUNCE, J.D. Comparative molecular-field analysis (CoMFA).1. Effect of shape on binding of steroids to carrier proteins. **Journal of the American Chemical Society**, v.110, p. 5959-5967, 1988.

DEBNATH, Sudhan; DEBNATH, Tanusree; BHAUMIK, Samhita; *et al.* Discovery of novel potential selective HDAC8 inhibitors by combine ligand-based, structurebased virtual screening and in-vitro biological evaluation. p. 1–14, 2019.

DRUKER, Brian J.; LYDON, Nicholas B. Lessons learned from the development of an Abl tyrosine kinase inhibitor for chronic myelogenous leukemia. **Journal of Clinical Investigation**, v. 105, n. 1, p. 3–7, 2000.

DUBOIS, F.; Caby, S.; Oger, F.; Cosseau, C.; Capron, M.; Grunau, C.; Dissous, C.; Pierce, R. J. Histone deacetylase inhibitors induce apoptosis, histone hyperacetylation and up-regulation of gene transcription in *Schistosoma mansoni* **Mol. Biochem. Parasitol.** 2009, 168, 7– 15.

ERLANSON, Daniel A. Introduction to Fragment-Based Drug Discovery. **Topics in Current Chemistry**, v. 310, n. June 2011, p. 1–32, 2011.

FEINSTEIN, Wei P.; BRYLINSKI, Michal. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. **Journal of Cheminformatics**, v. 7, n. 1, 2015. Disponível em: <<http://dx.doi.org/10.1186/s13321-015-0067-5>>.

FOURCHES, Denis; MURATOV, Eugene; TROPSHA, Alexander. Trust, But Verify: On the Importance of Chemical Structure Curation in Cheminformatics and QSAR

Modeling Research. **Journal of Chemical Information and Modeling**, v. 50, n. 7, p. 1189–1204, 2010.

GAUDIO, A. C.; ZANDONADE, E. Proposition, Validation and Analysis of Qsar Models. **Quim Nova**. SBQ, v. 24, n. 5, p. 658–671, 2001.

GIMENO, Aleix; OJEDA-MONTES, María José; TOMÁS-HERNÁNDEZ, Sarah; *et al.* The light and dark sides of virtual screening: What is there to know? **International Journal of Molecular Sciences**, v. 20, n. 6, 2019.

GUGLIELMO, Stefano; CORTESE, Daniela; VOTTERO, Francesca; *et al.* New praziquantel derivatives containing NO-donor furoxans and related furazans as active agents against *Schistosoma mansoni*. **European Journal of Medicinal Chemistry**, v. 84, p. 135–145, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.ejmech.2014.07.007>>.

GUIDO, Rafael V.C.; ANDRICOPULO, Adriano D.; OLIVA, Glaucius. Planejamento de fármacos, biotecnologia e química medicinal: Aplicações em doenças infecciosas. **Estudos Avancados**, v. 24, n. 70, p. 81–98, 2010.

GURRATH, Marion; MÜLLER, Gerhard; HÖLTJE, Hans Dieter. Pseudoreceptor Modelling in Drug Design: Applications of Yak and PrGen. **Perspectives in Drug Discovery and Design**, v. 12--14, p. 135–157, 1998.

HAI, Yang; EDWARDS, Jennifer E.; VAN ZANDT, Michael C.; *et al.* Crystal structure of *Schistosoma mansoni* arginase, a potential drug target for the treatment of schistosomiasis. **Biochemistry**, v. 53, n. 28, p. 4671–4684, 2014.

HEIMBURG, Tino; CHAKRABARTI, Alokta; LANCELOT, Julien; *et al.* Structure-Based Design and Synthesis of Novel Inhibitors Targeting HDAC8 from *Schistosoma mansoni* for the Treatment of Schistosomiasis. **Journal of Medicinal Chemistry**, v. 59, n. 6, p. 2423–2435, 2016.

JOHN R.GIUDICESSI, BA.MICHAEL J.ACKERMAN., 2013. Molecular Docking: A powerful approach for structure-based drug discovery. **current computer aided drug**

design, v. 23, n. 1, p. 1–7, 2008. Disponível em:

<<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf>>.

K, Ojha Lokendra; RACHANA, Sharma; RANI, Bhawsar Mukta. Modern drug design with advancement in QSAR : A review. **Int. J. Res. BioSciences 1 International Journal of Research in BioSciences**, v. 2, n. 1, p. 1–12, 2013. Disponível em: <<http://www.ijrbs.in>>.

KANNAN, Srinivasaraghavan; MELESINA, Jelena; HAUSER, Alexander Thomas; *et al.* Discovery of inhibitors of schistosoma mansoni hdac8 by combining homology modeling, virtual screening, and in vitro validation. **Journal of Chemical Information and Modeling**, v. 54, n. 10, p. 3005–3019, 2014.

K.I.Ramachandran, G Deepa and Krishnan Namboori. P.K. *Computational Chemistry and Molecular Modeling Principles and Applications* 2008 [3] [ISBN 978-3-540-77302-3](#) Springer-Verlag GmbH.

KIRALJ, Rudolf; FERREIRA, Márcia M C. Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application. **Journal of the Brazilian Chemical Society**, v. 20, n. 4, p. 770–787, 2009.

Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer III, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, **J. Chem. Theory Comput.**, 13(7) 3185–3197 (2017). (doi: [10.1021/acs.jctc.7b00174](https://doi.org/10.1021/acs.jctc.7b00174)).

MAREK, Martin; SHAIK, Tajith B.; HEIMBURG, Tino; *et al.* Characterization of Histone Deacetylase 8 (HDAC8) Selective Inhibition Reveals Specific Active Site Structural and Functional Determinants. **Journal of Medicinal Chemistry**, v. 61, n. 22, p. 10000–10016, 2018.

MCKERROW, J; CAFFREY, C; POWERS, J. Aza-peptidyl Michael Acceptors. A New Class of Potent and Selective Inhibitors of {...}. **Jmc**, p. 2816–2832, 2008. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/jm701311r>>.

MELO-FILHO, Cleber C.; BRAGA, Rodolpho C.; MURATOV, Eugene N.; *et al.* Discovery of new potent hits against intracellular *Trypanosoma cruzi* by QSAR-based virtual screening. **European Journal of Medicinal Chemistry**, v. 163, p. 649–659, 2019. Disponível em: <<https://doi.org/10.1016/j.ejmech.2018.11.062>>.

MULLER-STEFFNER, Hélène; JACQUES, Sylvain A.; KUHN, Isabelle; *et al.* Efficient Inhibition of SmNACE by Coordination Complexes Is Abolished by *S. mansoni* Sequestration of Metal. **ACS Chemical Biology**, v. 12, n. 7, p. 1787–1795, 2017.

NANTASENAMAT, C. *et al.* A practical overview of quantitative structure-activity relationship. **EXCLI Journal**, v. 8, p. 74–88, 2009.

NEVES, Bruno J.; DANTAS, Rafael F.; SENGER, Mario R.; *et al.* Discovery of New Anti-Schistosomal Hits by Integration of QSAR-Based Virtual Screening and High Content Screening. **Journal of Medicinal Chemistry**, v. 59, n. 15, p. 7075–7088, 2016.

NOGUEIRA, Luciano José; MONTANARI, Carlos Alberto; DONNICI, Cláudio Luis. Histórico da evolução da química medicinal e a importância da lipofilia : de Hipócrates e Galeno a Paracelsus e as contribuições de Overton e de Hansch Histórico da evolução da química medicinal e a importância da lipofilia : de Hipócrates e Galeno a Para. **Rev. Virtual Quim.**, v. 1, n. 3, p. 227–240, 2009.

OECD. Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)Sar] Models. **Transport**, v. 2, n. February, p. 1–154, 2007.

OECD. PRINCIPLES FOR THE VALIDATION, FOR REGULATORY PURPOSES, OF (QUANTITATIVE) STRUCTURE-ACTIVITY RELATIONSHIP MODELS These principles were agreed by OECD member countries at the 37. **Biotechnology**, n. November, p. 3–4, 2004.

ORTIZ, Christopher Llynard D.; COMPLETO, Gladys C.; NACARIO, Ruel C.; *et al.* Potential Inhibitors of Galactofuranosyltransferase 2 (GlfT2): Molecular Docking, 3D-QSAR, and In Silico ADMETox Studies. **Scientific Reports**, v. 9, n. 1, p. 1–28, 2019.

PAULO, João; MARTINS, A; BARBOSA, Euzébio G; *et al.* LQTA-QSAR: A New 4D-QSAR Methodology. p. 1428–1436, 2009. Disponível em: <<https://pubs.acs.org/doi/pdf/10.1021/ci900014f>>.

PETTERSEN, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. **Journal of computational chemistry**, v. 25, n. 13, p. 1605–12, out. 2004.

PETUKH, Marharyta; STEFL, Shannon; ALEXOV, Emil. The Role of Protonation States in Ligand-Receptor Recognition and Binding. **Current Pharmaceutical Design**, v. 19, n. 23, p. 4182–4190, 2013.

PIERCE, R. J.; Dubois

Abdeselem, F.; Lancelot, J.; Andrade, L.; Oliveira, G. Targeting schistosome histone modifying enzymes for drug development **Curr. Pharm. Des.** 2012, 18, 3567–78.

Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability”, R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince III, E. G. Hohenstein, U. Bozkaya, A. Yu. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A.

RAVI, M. *et al.* 4D-QSAR Analysis of a Set of Ecdysteroids and a Comparison to CoMFA Modeling. **Journal Chemical Information Computer**, p. 1587–1604, 2001.

ROY, Kunal; KAR, Supratik; DAS, Rudra Narayan. **Newer QSAR Techniques**. [s.l.: s.n.], 2015.

SATO, Tomohiro; SEKIMATA, Katsuhiko; SAKAI, Naoki; *et al.* Structural Basis of Activin Receptor-Like Kinase 2 (R206H) Inhibition by Bis-heteroaryl Pyrazole-Based Inhibitors for the Treatment of Fibrodysplasia Ossificans Progressiva Identified by the Integration of Ligand-Based and Structure-Based Drug Design Approaches. **ACS Omega**, v. 2, 2020.

SAFAVI-SOHI, J. B. G. R.; BARBOSA, E. G. 4D-LQTA-QSAR and docking study on potent gram-negative specific LpxC inhibitors : a comparison to CoMFA modeling. **Molecular Diversity**, p. 203–213, 2012.

SAXENA, Akansha; SANGWAN, Rajender Singh; MISHRA, Sanjay. Fundamentals of Homology Modeling Steps and Comparison among Important Bioinformatics Tools: An Overview. **Science International**, v. 1, n. 7, p. 237–252, 2013.

SCHNEIDER, Gisbert; FECHNER, Uli. Computer-based de novo design of drug-like molecules. **Nature Reviews Drug Discovery**, v. 4, n. 8, p. 649–663, 2005.

STOLFA, Diana A.; MAREK, Martin; LANCELOT, Julien; *et al.* Molecular basis for the antiparasitic activity of a mercaptoacetamide derivative that inhibits histone deacetylase 8 (HDAC8) from the human pathogen schistosoma mansoni. **Journal of Molecular Biology**, v. 426, n. 20, p. 3442–3453, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.jmb.2014.03.007>>.

SUN, Hongmao; TAWA, Gregory; WALLQVIST, Anders. Classification of scaffold-hopping approaches. **Drug Discovery Today**, v. 17, n. 7–8, p. 310–324, 2012.

TAHA, Mutasem O.; BUSTANJI, Yasser; AL-BAKRI, Amal G.; *et al.* Discovery of new potent human protein tyrosine phosphatase inhibitors via pharmacophore and QSAR analysis followed by in silico screening. **Journal of Molecular Graphics and Modelling**, v. 25, n. 6, p. 870–884, 2007.

TAHA, Mutasem O.; ATALLAH, Najji; AL-BAKRI, Amal G.; *et al.* Discovery of new MurF inhibitors via pharmacophore modeling and QSAR analysis followed by in-silico screening. **Bioorganic and Medicinal Chemistry**, v. 16, n. 3, p. 1218–1235, 2008.

TANRIKULU, Yusuf; SCHNEIDER, Gisbert. Pseudoreceptor models in drug design: Bridging ligand- and receptor-based virtual screening. **Nature Reviews Drug Discovery**, v. 7, n. 8, p. 667–677, 2008.

TAVARES, Leoberto Costa. QSAR: A abordagem de Hansch. **Quimica Nova**, v. 27, n. 4, p. 631–639, 2004.

THILAGAVATHI, Ramasamy; CHAKRABORTI, Asit K. Importance of Alignment in Developing 3-D QSAR Models of 1,5-Diaryl Pyrazoles for Prediction of COX-2 Inhibitory Activity. **Internet Electronic Journal of Molecular Design**, v. 2, p. 0–11, 2003.

TOSCO, P.; BALLE, T. Open3DQSAR: a new open-source software aimed at highthroughput chemometric analysis of molecular interaction fields. **Journal of molecular modeling**, v. 17, n. 1, p. 201–8, jan. 2011.

TROTT, O.; OLSON, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of computational chemistry**, v. 31, n. 2, p. 455–61, 30 jan. 2010.

TUFFAHA, Ghada Omar; HATMAL, Ma'mon M.; TAHA, Mutasem O. Discovery of new JNK3 inhibitory chemotypes via QSAR-Guided selection of docking-based pharmacophores and comparison with other structure-based pharmacophore modeling methods. **Journal of Molecular Graphics and Modelling**, v. 91, p. 30–51, 2019. Disponível em: <<https://doi.org/10.1016/j.jmgm.2019.05.015>>.

VEERASAMY, Ravichandran; RAJAK, Harish; JAIN, Abhishek; *et al.* Ravichandran Veerasamy, et al : Validation of QSAR Models-Strategies and Importance Validation of QSAR Models -Strategies and Importance. **International Journal of Drug Design and Discovery** , v. 2, n. 3, p. 511–519, 2011. Disponível em: <http://ijddonline.com/issues/511_full.pdf>.

VERMA, Jitender; KHEDKAR, Vijay; COUTINHO, Evans. 3D-QSAR in Drug Design - A Review. **Current Topics in Medicinal Chemistry**, v. 10, n. 1, p. 95–115, 2010. Disponível em: <<http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1568-0266&volume=10&issue=1&spage=95>>.

Wang W, Wang L, Liang Y. Susceptibility or resistance of praziquantel in human schistosomiasis: a review. **Parasitol Research**. 2012;111:1871–7.

WEISBERG, Ellen; MANLEY, Paul W.; BREITENSTEIN, Werner; *et al.* Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl. **Cancer Cell**, v. 7, n. 2, p. 129–141, 2005.

WHO. Global Health Estimates 2015: Deaths by Cause, Age, Sex, by Country and by Region, 2000-

2015. http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html . **Geneva, World Health Organization**; 2016a.

WHO. Global Health Observatory (GHO) data 2016.

http://www.who.int/gho/neglected_diseases/schistosomiasis/en/. **Geneva, World Health Organization**; 2016b.

WHO. Schistosomiasis . Disponível em: <

<http://www.who.int/mediacentre/factsheets/fs310/en/>>. Acesso em: 21 de JUL. 2018.

WOOD, Jeanette M; MAIBAUM, Jürgen; RAHUEL, Joseph; *et al.* Structure-based design of aliskiren, a novel orally effective renin inhibitor. **Biochemical and Biophysical Research Communications**, v. 308, n. 4, p. 698–705, 2003.

WU, Cathy H; ARIGHI, Cecilia; ROSS, Karen. Protein Bioinformatics: From Protein Modifications and Networks to Proteomics. **Methods in Molecular Biology**, v. 1558, n. 1558, p. 472, 2017. Disponível em:

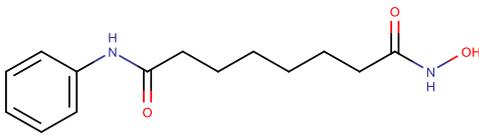
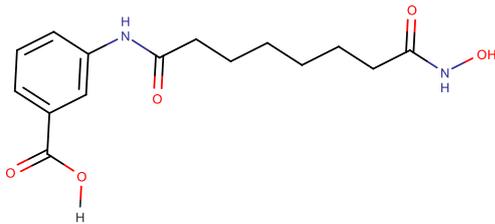
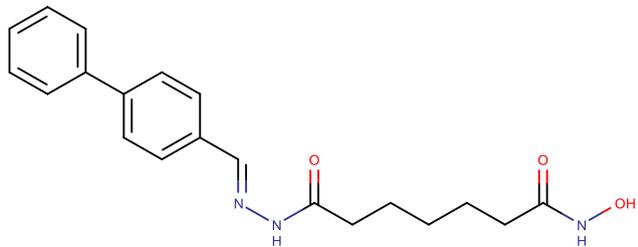
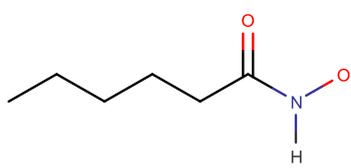
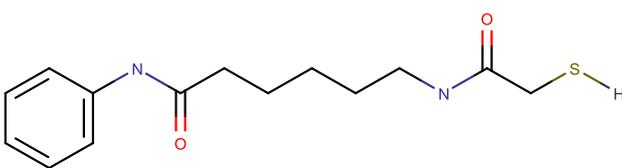
<<https://www.springer.com/us/book/9781493967810>>.

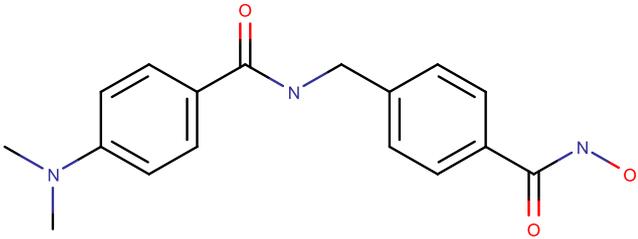
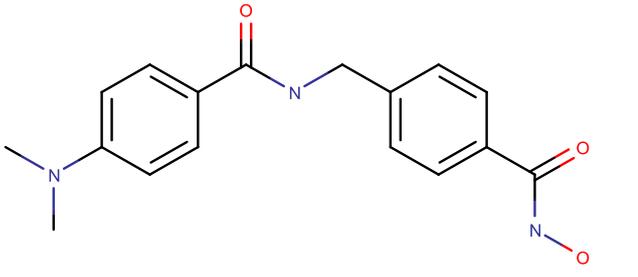
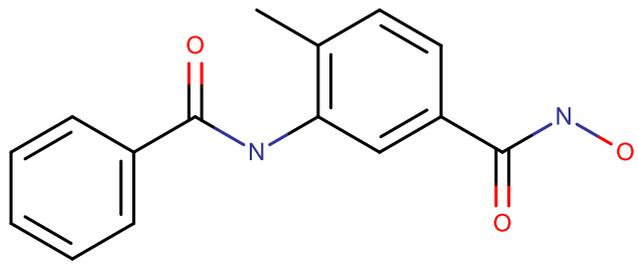
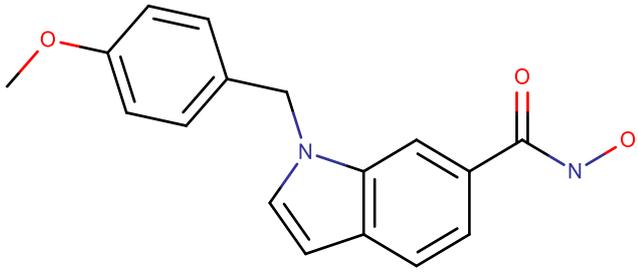
XUE, L.; BAJORATH, J. Molecular descriptors for effective classification of biologically active compounds based on principal component analysis identified by a genetic algorithm. **Journal of chemical information and computer sciences**, v. 40, n. 3, p. 801–9, jan. 2000.

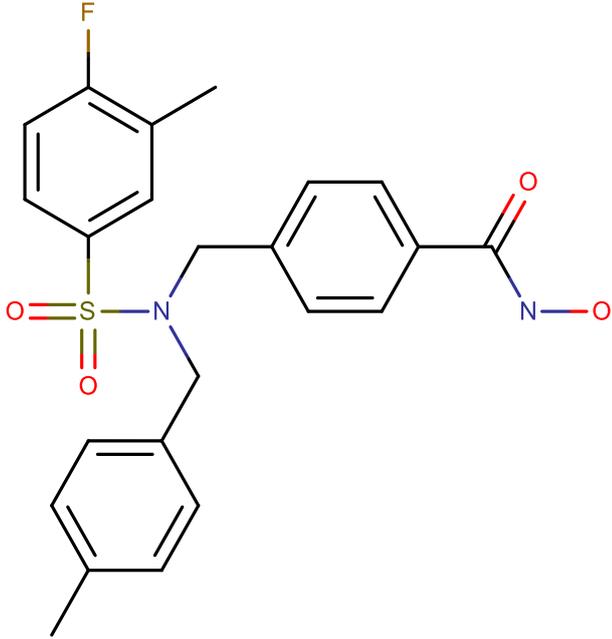
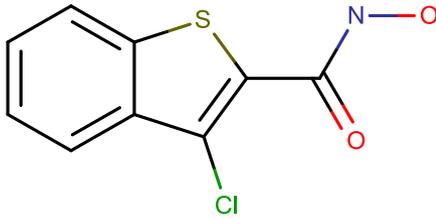
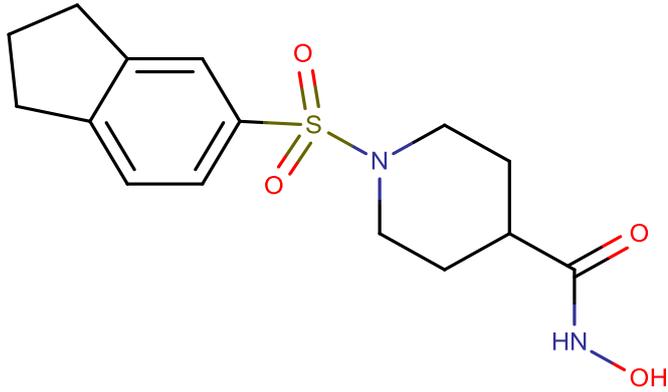
YANG, Sheng Yong. Pharmacophore modeling and applications in drug discovery: Challenges and recent advances. **Drug Discovery Today**, v. 15, n. 11–12, p. 444–450, 2010. Disponível em: <<http://dx.doi.org/10.1016/j.drudis.2010.03.013>>.

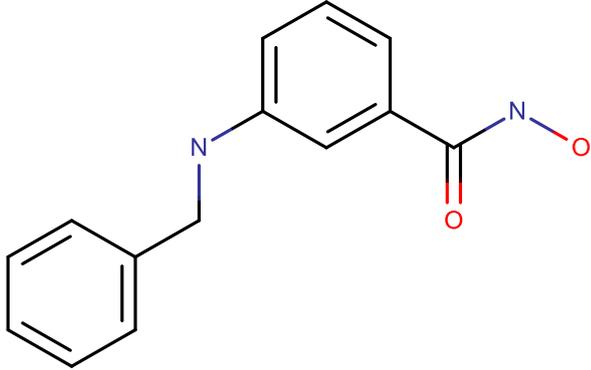
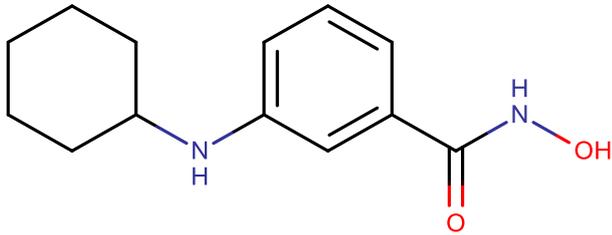
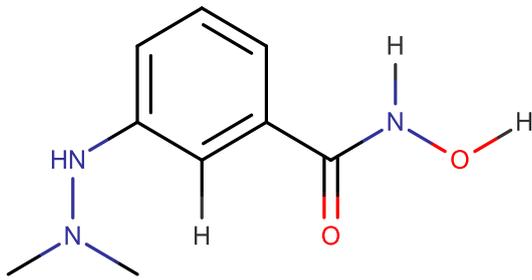
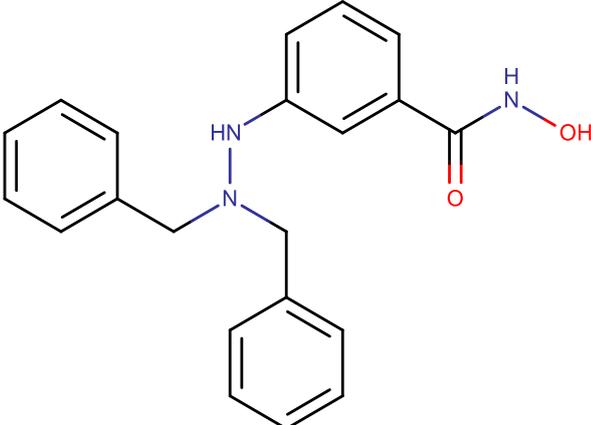
7 ANEXOS

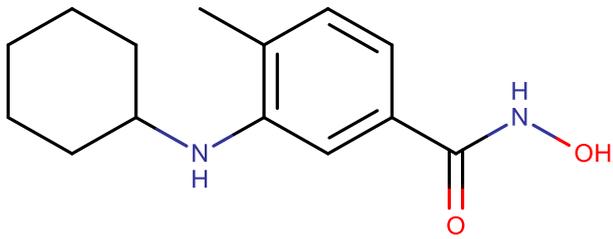
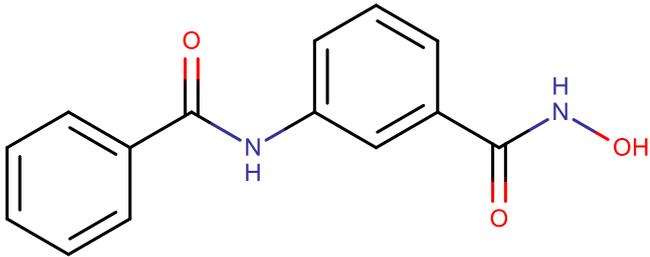
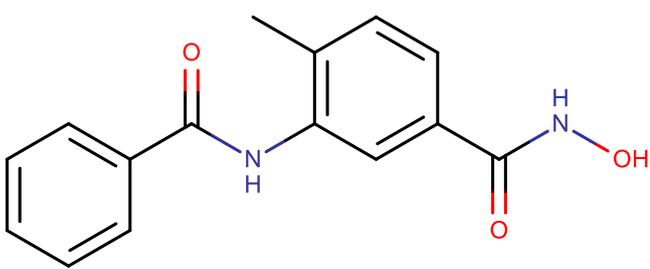
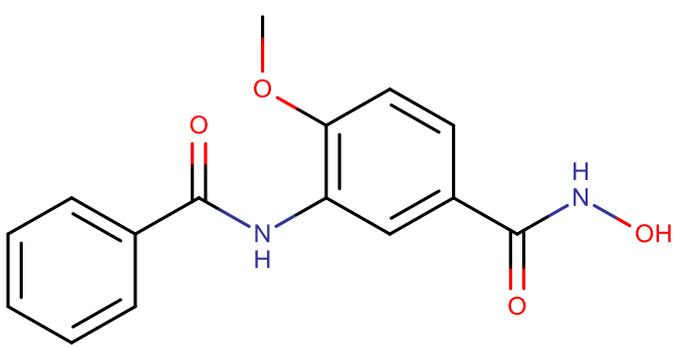
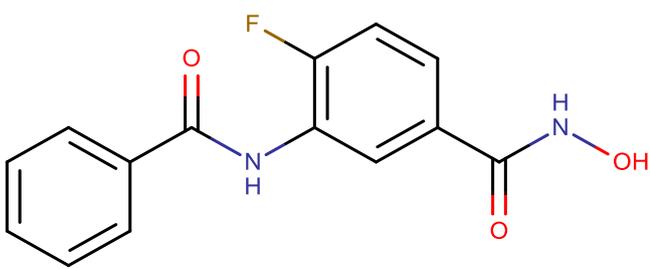
7.1 Anexo 1. Estruturas 2D das moléculas escolhidas para compor o estudo bem como suas respectivas atividades biológicas.

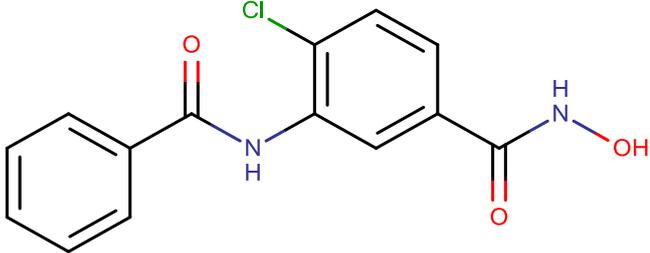
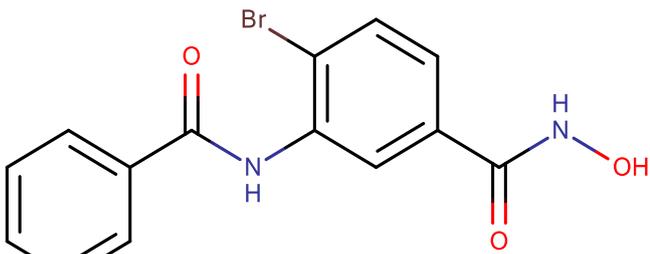
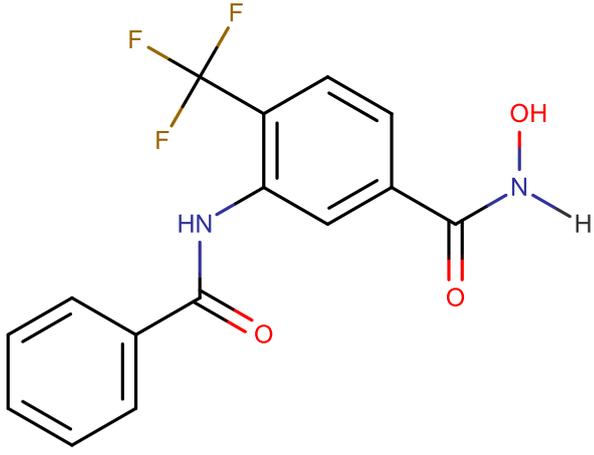
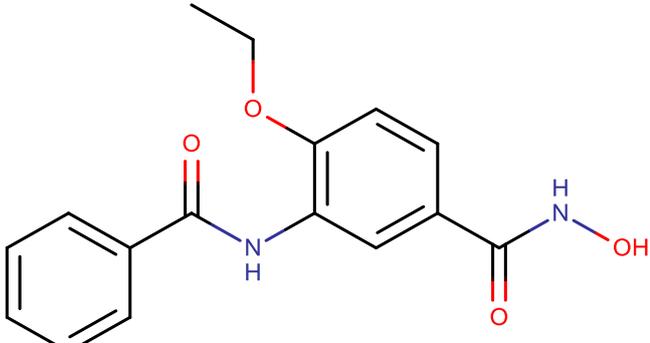
ID	Estrutura 2D	Artigo de referência	Ic50 (Nanomolar -nM)
1		STOLFA <i>et al.</i> , 2014	1900
2		STOLFA <i>et al.</i> , 2014	800
3		STOLFA <i>et al.</i> , 2014	900
4		STOLFA <i>et al.</i> , 2014	2500
5		STOLFA <i>et al.</i> , 2014	50000

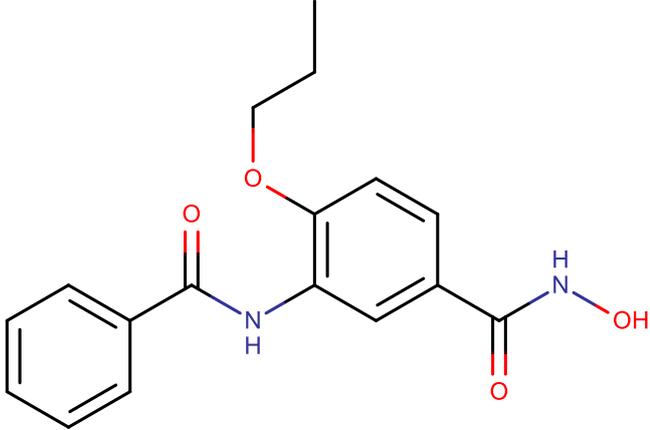
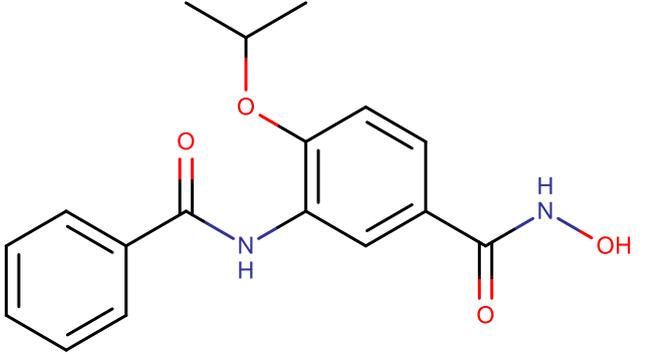
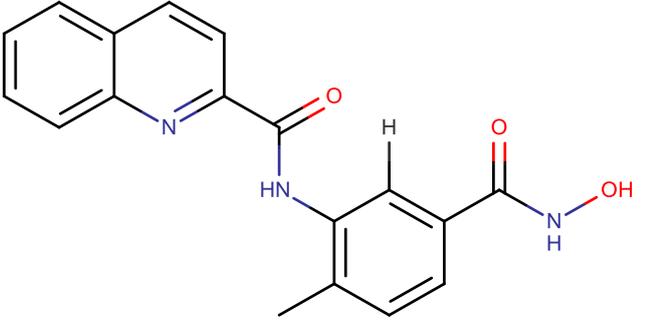
6		STOLFA <i>et al.</i> , 2014	2400
7		STOLFA <i>et al.</i> , 2014	29000
8		STOLFA <i>et al.</i> , 2014	1400
9		STOLFA <i>et al.</i> , 2014	400

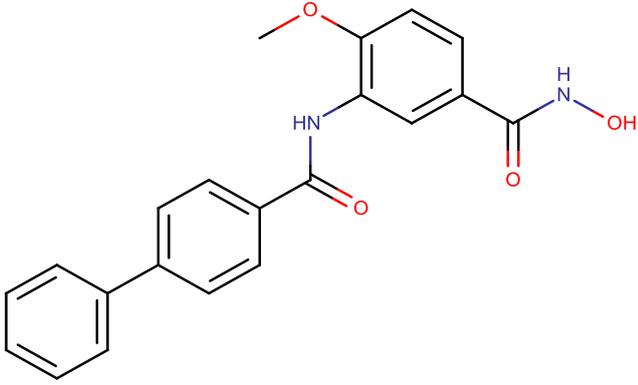
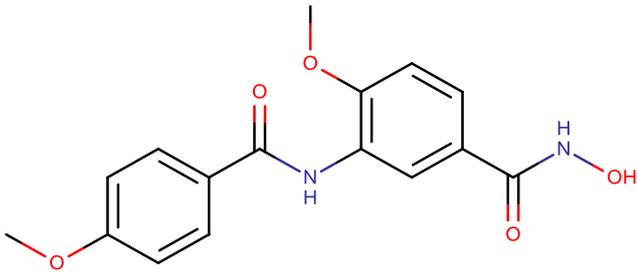
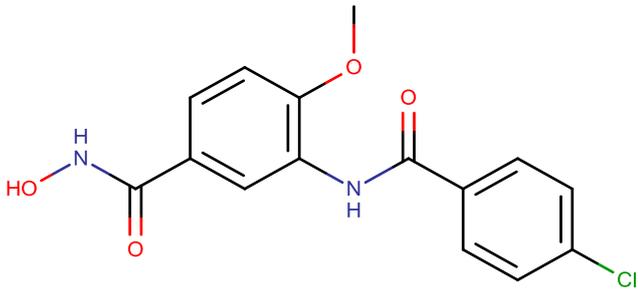
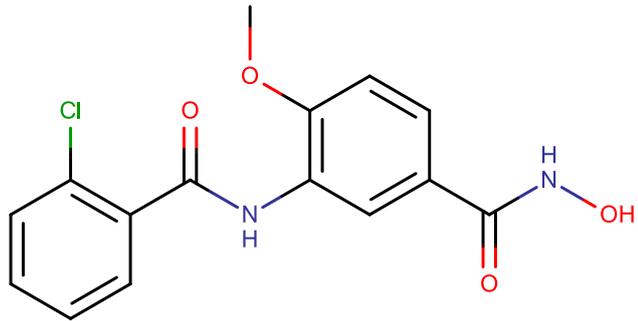
10	 <chem>Cc1ccc(cc1)N(S(=O)(=O)N)Cc2ccc(cc2)C(=O)N=O</chem>	KANNAN <i>et al.</i> 2014	5630
11	 <chem>O=[N+]([O-])C(=O)c1c(Cl)sc2ccccc12</chem>	KANNAN <i>et al.</i> 2014	4320
12	 <chem>OC(=O)N1CCCCC1NS(=O)(=O)c2ccc(cc2)C3CCCC3</chem>	KANNAN <i>et al.</i> 2014	3030

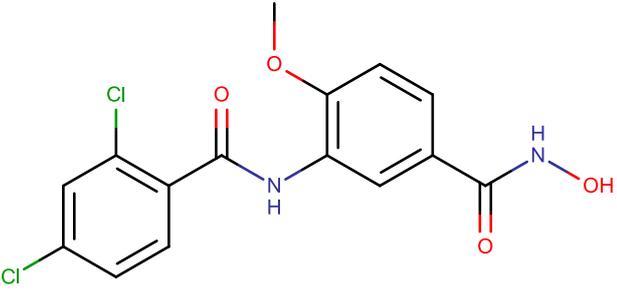
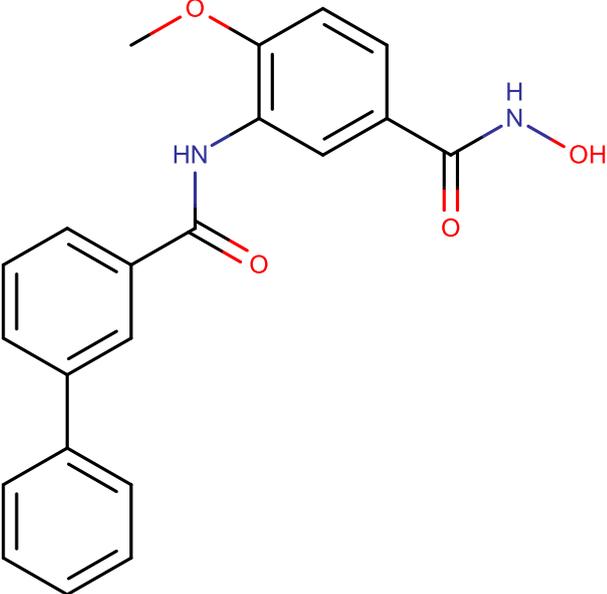
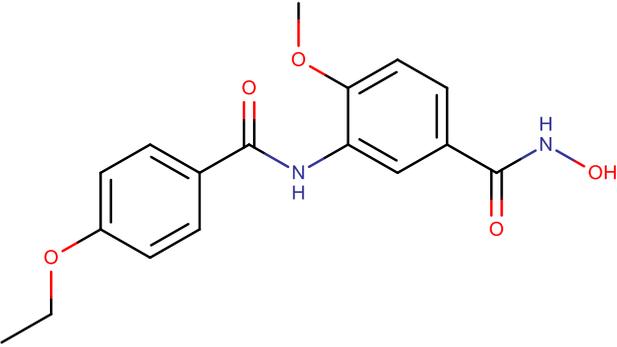
13	 <p>Chemical structure of N-(benzylamino)-3-nitrobenzamide. It features a benzamide core with a nitro group at the 3-position and a benzylamino group at the 1-position.</p>	HEIMBUR G <i>et al.</i> , 2016	1080
14	 <p>Chemical structure of N-(cyclohexylamino)-3-hydroxybenzamide. It features a benzamide core with a hydroxyl group at the 3-position and a cyclohexylamino group at the 1-position.</p>	HEIMBUR G <i>et al.</i> , 2016	3630
15	 <p>Chemical structure of N-(dimethylamino)-3-hydroxybenzamide. It features a benzamide core with a hydroxyl group at the 3-position and a dimethylamino group at the 1-position.</p>	HEIMBUR G <i>et al.</i> , 2016	1580
16	 <p>Chemical structure of N-(benzyl(phenyl)amino)-3-hydroxybenzamide. It features a benzamide core with a hydroxyl group at the 3-position and a benzyl(phenyl)amino group at the 1-position.</p>	HEIMBUR G <i>et al.</i> , 2016	9290

17	 <p>Chemical structure of N-(cyclohexyl)N-(3-hydroxyimino-4-methylphenyl)acetamide. It features a cyclohexane ring attached to the nitrogen of an acetamide group, which is further substituted with a 3-hydroxyimino-4-methylphenyl group.</p>	HEIMBUR G <i>et al.</i> , 2016	600
18	 <p>Chemical structure of N-(benzoylamino)-3-hydroxyiminoacetamide. It consists of a benzamide group attached to the nitrogen of a 3-hydroxyiminoacetamide group.</p>	HEIMBUR G <i>et al.</i> , 2016	468
19	 <p>Chemical structure of N-(benzoylamino)-3-hydroxyimino-4-methylacetamide. It features a benzamide group attached to the nitrogen of a 3-hydroxyiminoacetamide group, with a methyl group at the 4-position of the phenyl ring.</p>	HEIMBUR G <i>et al.</i> , 2016	116
20	 <p>Chemical structure of N-(benzoylamino)-3-hydroxyimino-4-methoxyacetamide. It features a benzamide group attached to the nitrogen of a 3-hydroxyiminoacetamide group, with a methoxy group at the 4-position of the phenyl ring.</p>	HEIMBUR G <i>et al.</i> , 2016	190
21	 <p>Chemical structure of N-(benzoylamino)-3-hydroxyimino-4-fluoroacetamide. It features a benzamide group attached to the nitrogen of a 3-hydroxyiminoacetamide group, with a fluorine atom at the 4-position of the phenyl ring.</p>	HEIMBUR G <i>et al.</i> , 2016	178

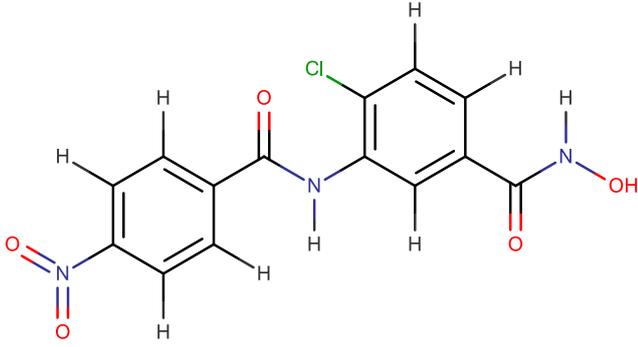
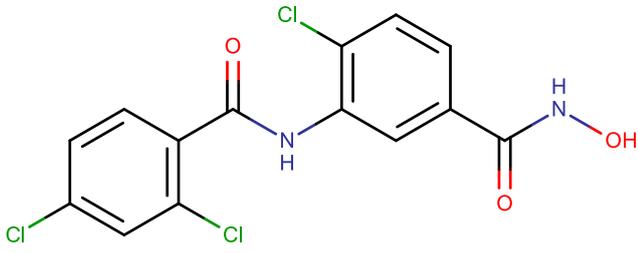
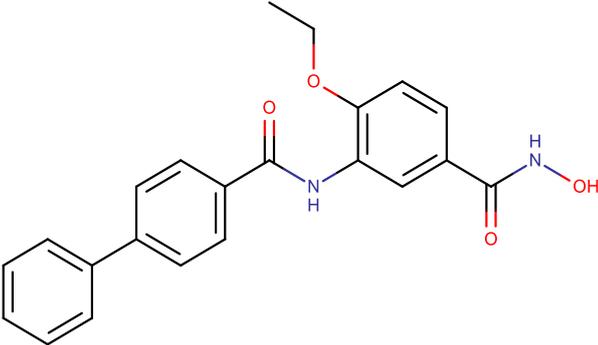
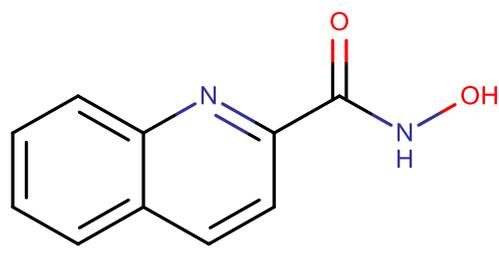
22	 <p>Chemical structure of N-(3-chloro-4-(benzoylamino)phenyl)acetamide hydroxylamine. The central benzene ring has a chlorine atom at the 3-position, a benzoylamino group (-NH-C(=O)-C₆H₅) at the 4-position, and an acetamide hydroxylamine group (-C(=O)-NH-OH) at the 1-position.</p>	HEIMBUR G <i>et al.</i> , 2016	67
23	 <p>Chemical structure of N-(3-bromo-4-(benzoylamino)phenyl)acetamide hydroxylamine. The central benzene ring has a bromine atom at the 3-position, a benzoylamino group (-NH-C(=O)-C₆H₅) at the 4-position, and an acetamide hydroxylamine group (-C(=O)-NH-OH) at the 1-position.</p>	HEIMBUR G <i>et al.</i> , 2016	150
24	 <p>Chemical structure of N-(3-(difluoromethyl)-4-(benzoylamino)phenyl)acetamide hydroxylamine. The central benzene ring has a difluoromethyl group (-CH₂F₂) at the 3-position, a benzoylamino group (-NH-C(=O)-C₆H₅) at the 4-position, and an acetamide hydroxylamine group (-C(=O)-NH-OH) at the 1-position.</p>	HEIMBUR G <i>et al.</i> , 2016	140
25	 <p>Chemical structure of N-(3-ethoxy-4-(benzoylamino)phenyl)acetamide hydroxylamine. The central benzene ring has an ethoxy group (-OCH₂CH₃) at the 3-position, a benzoylamino group (-NH-C(=O)-C₆H₅) at the 4-position, and an acetamide hydroxylamine group (-C(=O)-NH-OH) at the 1-position.</p>	HEIMBUR G <i>et al.</i> , 2016	129

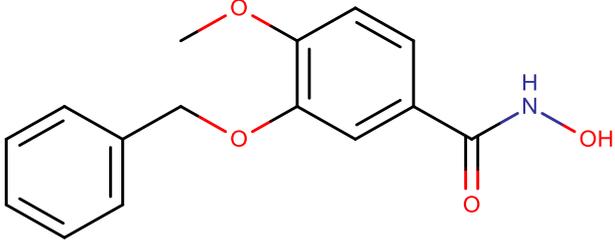
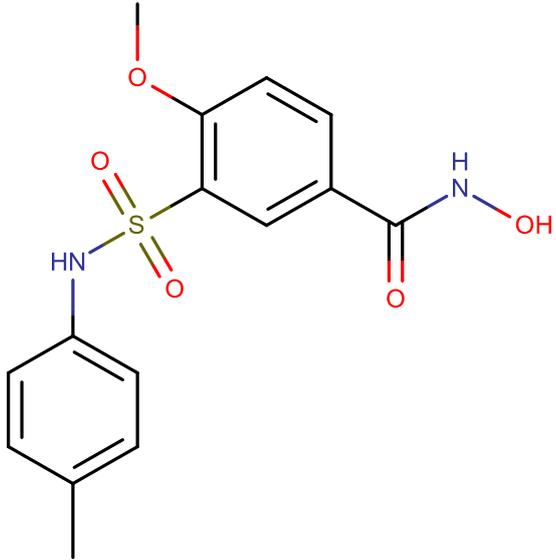
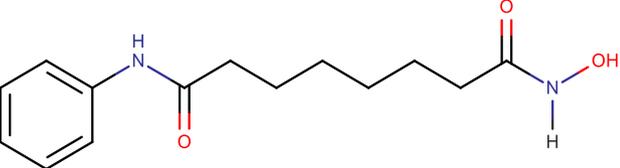
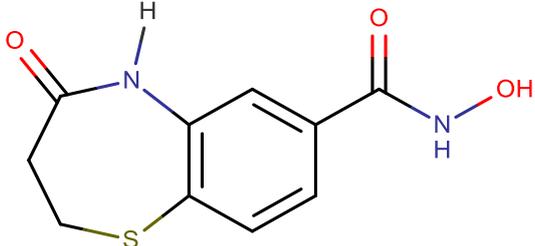
26	 <p>The structure shows a benzamide group (a benzene ring attached to a carbonyl group, which is further attached to an NH group) connected to a central benzene ring. This central benzene ring also has a hydroxamic acid group (-C(=O)NHOH) and an ethoxy group (-OCH2CH2CH3) attached to it.</p>	HEIMBUR G <i>et al.</i> , 2016	267
27	 <p>The structure is similar to the one in row 26, but the ethoxy group is replaced by an isopropoxy group (-OCH2CH(CH3)2).</p>	HEIMBUR G <i>et al.</i> , 2016	220
28	 <p>The structure features a quinoline ring system (two fused benzene rings with a nitrogen atom) attached to a carbonyl group, which is further attached to an NH group. This NH group is connected to a central benzene ring. This central benzene ring also has a hydroxamic acid group (-C(=O)NHOH) and a methyl group (-CH3) attached to it.</p>	HEIMBUR G <i>et al.</i> , 2016	96,1

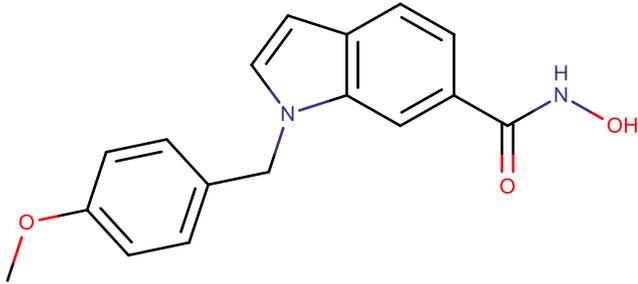
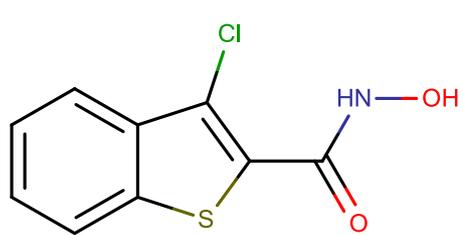
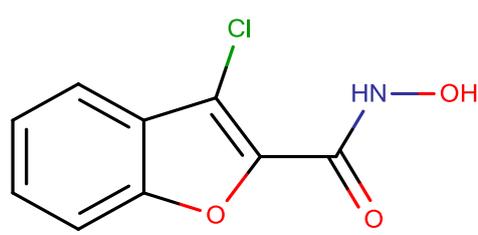
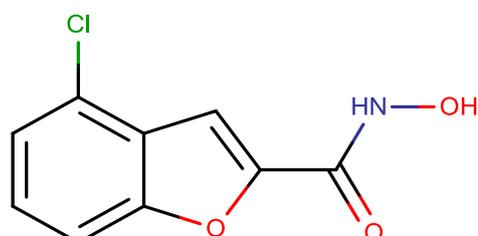
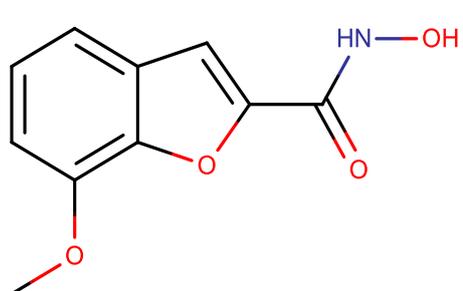
29		HEIMBUR G <i>et al.</i> , 2016	75,4
30		HEIMBUR G <i>et al.</i> , 2016	106
31		HEIMBUR G <i>et al.</i> , 2016	146
32		HEIMBUR G <i>et al.</i> , 2016	699

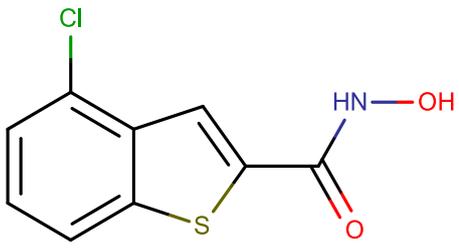
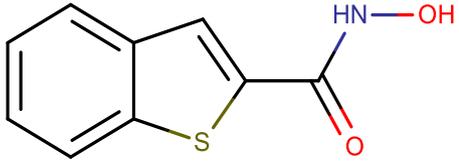
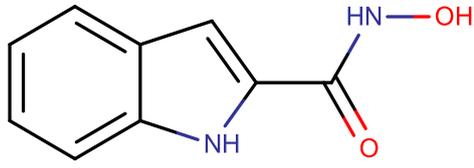
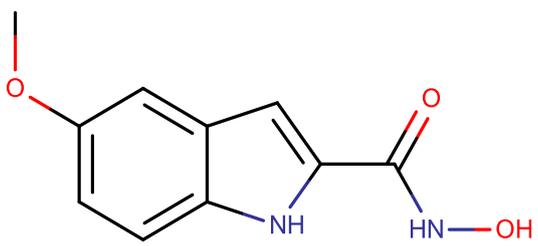
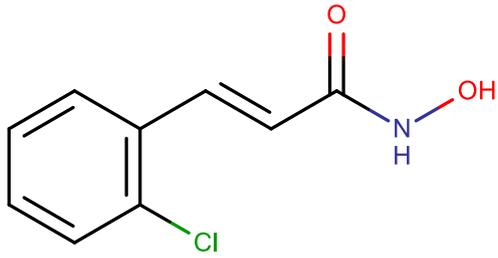
33		HEIMBUR G <i>et al.</i> , 2016	122
34		HEIMBUR G <i>et al.</i> , 2016	290
35		HEIMBUR G <i>et al.</i> , 2016	305

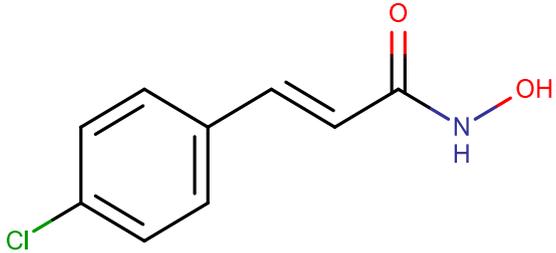
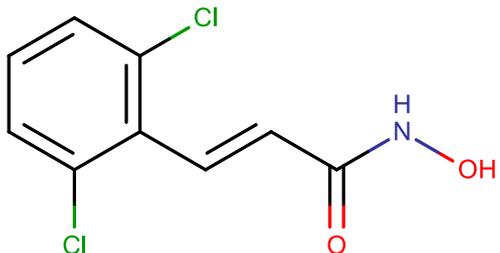
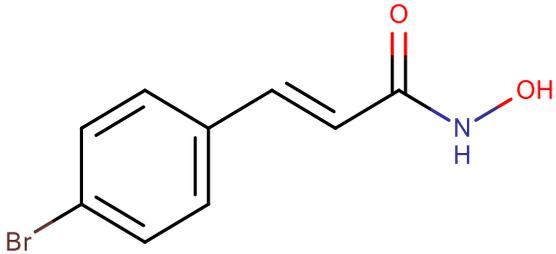
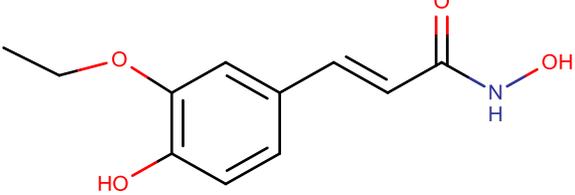
36	<chem>COc1ccc(cc1NC(=O)Cc2ccccc2)C(=O)NO</chem>	HEIMBUR G <i>et al.</i> , 2016	183
37	<chem>COc1ccc(cc1NC(=O)c2cc(Cl)ccc2)C(=O)NO</chem>	HEIMBUR G <i>et al.</i> , 2016	147
38	<chem>COc1ccc(cc1OCc2ccccc2)NC(=O)c3cc(Cl)ccc3C(=O)NO</chem>	HEIMBUR G <i>et al.</i> , 2016	378
39	<chem>Oc1ccc(cc1)Oc2ccc(cc2)NC(=O)c3cc(Cl)ccc3C(=O)NO</chem>	HEIMBUR G <i>et al.</i> , 2016	396
40	<chem>Oc1ccc(cc1)Oc2ccc(cc2)Oc3ccc(cc3)NC(=O)c4cc(Cl)ccc4C(=O)NO</chem>	HEIMBUR G <i>et al.</i> , 2016	979
41	<chem>Oc1ccc(cc1)C(=O)Nc2cc(Cl)ccc2C(=O)c3ccc(Cl)cc3</chem>	HEIMBUR G <i>et al.</i> , 2016	235

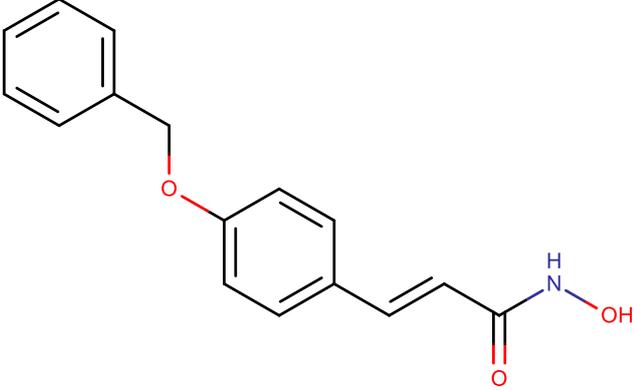
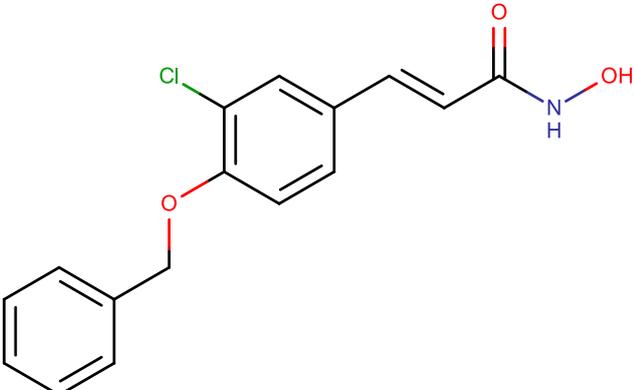
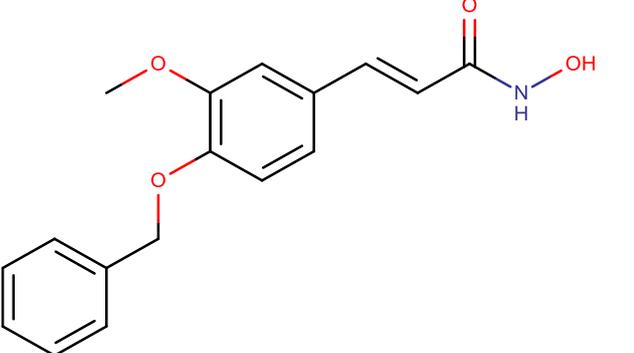
42		HEIMBUR G <i>et al.</i> , 2016	394
43		HEIMBUR G <i>et al.</i> , 2016	191
44		HEIMBUR G <i>et al.</i> , 2016	92
45		HEIMBUR G <i>et al.</i> , 2016	8210

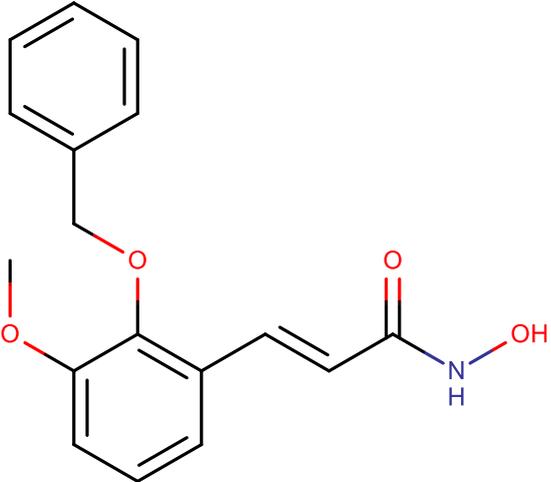
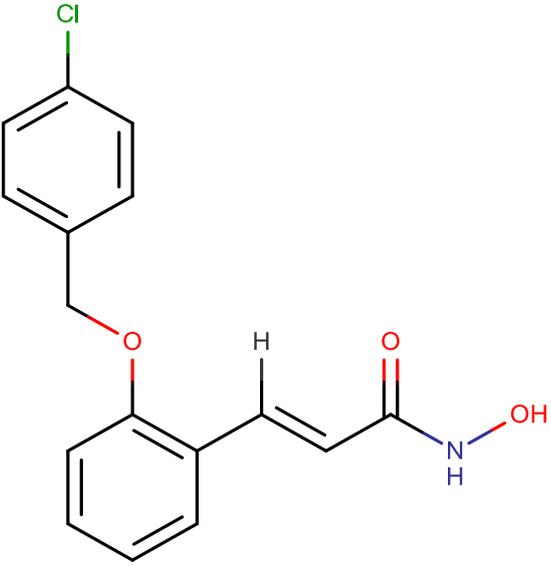
46	 <p>The structure shows a benzene ring with a methoxy group (-OCH₃) at the 3-position and a benzyloxy group (-OCH₂Ph) at the 4-position. At the 1-position, there is a hydroxamic acid group (-C(=O)NHOH).</p>	HEIMBUR G <i>et al.</i> , 2016	268
47	 <p>The structure shows a benzene ring with a methoxy group (-OCH₃) at the 3-position and a sulfamoyl group (-SO₂NH-Ph-4-Me) at the 4-position. At the 1-position, there is a hydroxamic acid group (-C(=O)NHOH).</p>	HEIMBUR G <i>et al.</i> , 2016	485
48	 <p>The structure shows a heptanedioic acid chain with a hydroxamic acid group (-C(=O)NHOH) at one end and a benzamide group (-NH-C(=O)-Ph) at the other end.</p>	HEIMBUR G <i>et al.</i> , 2016	1560
49	 <p>The structure shows a benzene ring with a 1,2,3,4-tetrahydrothiazolo[5,4-c]pyridin-5-yl group at the 4-position and a hydroxamic acid group (-C(=O)NHOH) at the 1-position.</p>	HEIMBUR G <i>et al.</i> , 2016	1220

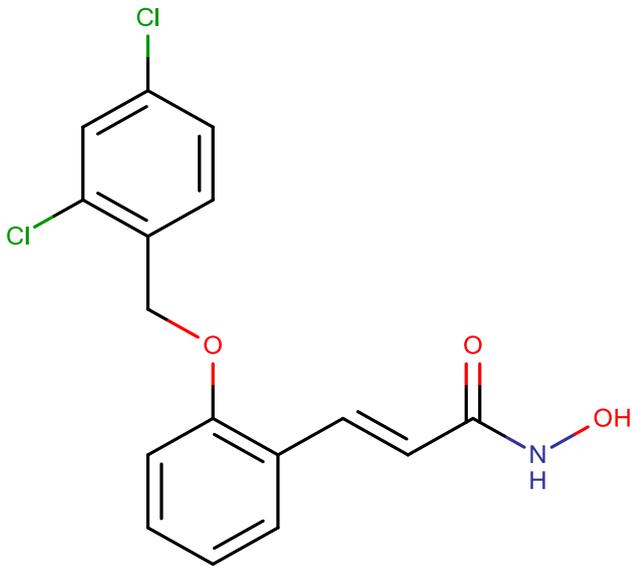
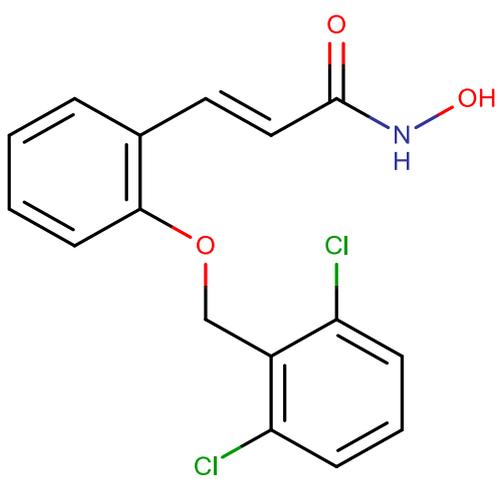
50	 <p>Chemical structure of a benzimidazole derivative. It features a benzimidazole ring system with a 4-methoxybenzyl group attached to the nitrogen atom and a hydroxamic acid group (-C(=O)NHOH) attached to the benzene ring.</p>	HEIMBUR G <i>et al.</i> , 2016	436
51	 <p>Chemical structure of a benzothiazole derivative. It features a benzothiazole ring system with a chlorine atom at the 5-position and a hydroxamic acid group (-C(=O)NHOH) at the 2-position.</p>	BAYER <i>et al.</i> , 2018	4300
52	 <p>Chemical structure of a benzofuran derivative. It features a benzofuran ring system with a chlorine atom at the 3-position and a hydroxamic acid group (-C(=O)NHOH) at the 2-position.</p>	BAYER <i>et al.</i> , 2018	240
53	 <p>Chemical structure of a benzofuran derivative. It features a benzofuran ring system with a chlorine atom at the 6-position and a hydroxamic acid group (-C(=O)NHOH) at the 2-position.</p>	BAYER <i>et al.</i> , 2018	780
54	 <p>Chemical structure of a benzofuran derivative. It features a benzofuran ring system with a methoxy group (-OCH₃) at the 6-position and a hydroxamic acid group (-C(=O)NHOH) at the 2-position.</p>	BAYER <i>et al.</i> , 2018	2100

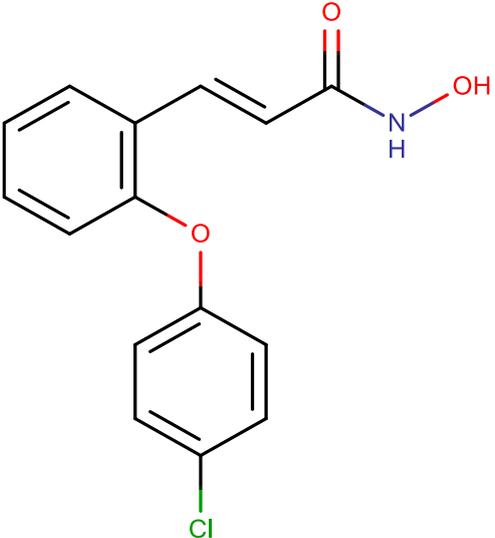
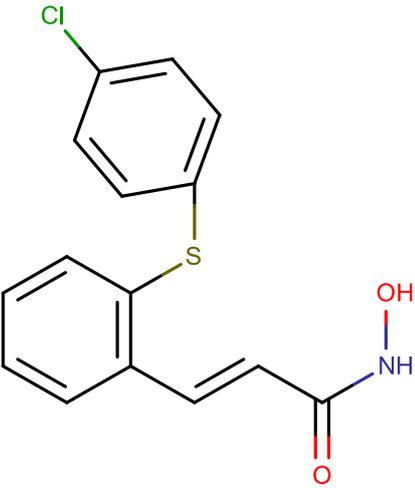
55	 <chem>O=C(O)C1=Cc2cc(Cl)ccc2S1</chem>	BAYER <i>et al.</i> , 2018	250
56	 <chem>O=C(O)C1=Cc2ccccc2S1</chem>	BAYER <i>et al.</i> , 2018	630
57	 <chem>O=C(O)C1=Cc2ccccc2N1</chem>	BAYER <i>et al.</i> , 2018	4900
58	 <chem>O=C(O)C1=Cc2cc(OC)ccc2N1</chem>	BAYER <i>et al.</i> , 2018	1800
59	 <chem>O=C(O)C=Cc1ccccc1Cl</chem>	BAYER <i>et al.</i> , 2018	60

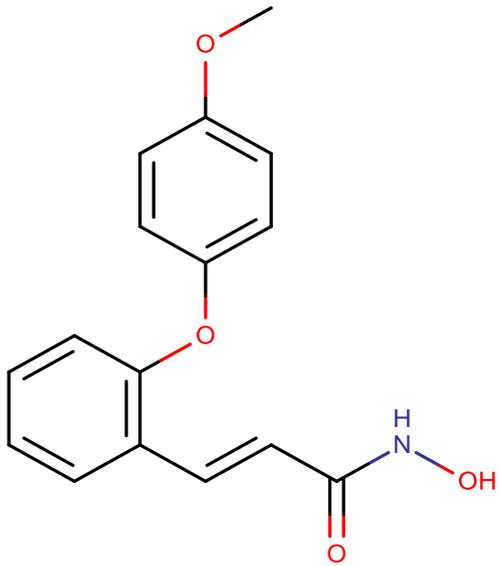
60	 <p>Chemical structure of 4-chlorobenzylideneacetone oxime. It features a benzene ring with a chlorine atom at the para position, connected to a propenal chain which is in its oxime form.</p>	BAYER <i>et al.</i> , 2018	800
61	 <p>Chemical structure of 2,6-dichlorobenzylideneacetone oxime. It features a benzene ring with chlorine atoms at the 2 and 6 positions, connected to a propenal chain which is in its oxime form.</p>	BAYER <i>et al.</i> , 2018	210
62	 <p>Chemical structure of 4-bromobenzylideneacetone oxime. It features a benzene ring with a bromine atom at the para position, connected to a propenal chain which is in its oxime form.</p>	BAYER <i>et al.</i> , 2018	240
63	 <p>Chemical structure of 4-ethoxyphenylideneacetone oxime. It features a benzene ring with an ethoxy group and a hydroxyl group at the para position, connected to a propenal chain which is in its oxime form.</p>	BAYER <i>et al.</i> , 2018	660

64	 <chem>O=C(O)C=Cc1ccc(OCC2=CC=CC=C2)cc1</chem>	BAYER <i>et al.</i> , 2018	250
65	 <chem>O=C(O)C=Cc1ccc(Cl)c(OCC2=CC=CC=C2)c1</chem>	BAYER <i>et al.</i> , 2018	3300
66	 <chem>O=C(O)C=Cc1ccc(OC)c(OCC2=CC=CC=C2)c1</chem>	BAYER <i>et al.</i> , 2018	1500

67	 <p>The structure shows a benzene ring with a methoxy group (-OCH₃) at the 2-position and a benzyl ether group (-CH₂-C₆H₅) at the 1-position. This benzene ring is connected at the 3-position to a propenoic acid derivative chain: -CH₂-CH=CH-C(=O)-NH-OH. The oxygen atoms are red, the nitrogen is blue, and the hydroxyl group is red.</p>	BAYER <i>et al.</i> , 2018	130
68	 <p>The structure shows a benzene ring with a chlorine atom (-Cl) at the 4-position and a benzyl ether group (-CH₂-C₆H₅) at the 1-position. This benzene ring is connected at the 3-position to a propenoic acid derivative chain: -CH₂-CH=CH-C(=O)-NH-OH. The oxygen atoms are red, the nitrogen is blue, and the hydroxyl group is red.</p>	BAYER <i>et al.</i> , 2018	720

69		BAYER <i>et al.</i> , 2018	520
70		BAYER <i>et al.</i> , 2018	440

71	 <p>The structure shows a benzene ring with a hydroxamic acid group (-CH=CH-C(=O)NH-OH) at the 1-position and a 4-chlorophenoxy group (-O-C₆H₄-Cl) at the 2-position. The chlorine atom is highlighted in green.</p>	BAYER <i>et al.</i> , 2018	220
72	 <p>The structure shows a benzene ring with a hydroxamic acid group (-CH=CH-C(=O)NH-OH) at the 1-position and a 4-chlorophenylsulfanyl group (-S-C₆H₄-Cl) at the 2-position. The chlorine atom is highlighted in green.</p>	BAYER <i>et al.</i> , 2018	180

73		BAYER <i>et al.</i> , 2018	90
----	---	----------------------------	----