



**UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ
CENTRO DE ENGENHARIAS E CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**DOENÇA DE ALZHEIMER: ESTUDOS IN SILICO DE
INIBIDORES DA PROTEINA CASPASE 3**

Paula Beatriz Jesus Santos

**Toledo – PR
2023**



**UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ
CENTRO DE ENGENHARIAS E CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**DOENÇA DE ALZHEIMER: ESTUDOS IN SILICO DE
INIBIDORES DA PROTEINA CASPASE 3**

Paula Beatriz Jesus Santos

Dissertação apresentada ao Programa de Pós-Graduação em Química da Universidade Estadual do Oeste do Paraná, Unioeste/*Campus* de Toledo, como parte dos requisitos para a obtenção do Título de Mestre em Química.

Orientador(a): Prof. Dr. Eduardo Borges de Melo

**Toledo – PR
2023**

Ficha de identificação da obra elaborada através do Formulário de Geração Automática do Sistema de Bibliotecas da Unioeste.

Jesus Santos, Paula Beatriz
DOENÇA DE ALZHEIMER: ESTUDOS IN SILICO DE INIBIDORES DA
PROTEÍNA CASPASE 3 / Paula Beatriz Jesus Santos; orientador
Eduardo Borges de Melo. -- Toledo, 2023.
60 p.

Dissertação (Mestrado Acadêmico Campus de Toledo) --
Universidade Estadual do Oeste do Paraná, Centro de
Engenharias e Ciências Exatas, Programa de Pós-Graduação em
Química, 2023.

1. Relação estrutura-atividade quantitativa. 2. Caspase 3.
3. Doença de Alzheimer (DA). I. Borges de Melo, Eduardo,
orient. II. Título.

Espaço dedicado a folha de aprovação

DEDICATÓRIA

Dedico este trabalho a pessoas muito especiais em minha vida, cujo apoio e encorajamento foram inestimáveis ao longo desta jornada.

.

AGRADECIMENTOS

A Deus, agradeço pelo dom da vida e pela oportunidade de trilhar esta jornada acadêmica. Sou grata por todas as experiências, desafios e conquistas que moldaram meu percurso e me tornaram a pessoa que sou hoje.

Ao meu esposo, que sempre esteve ao meu lado, me apoiando e sendo meu maior incentivador, mesmo quando eu mesma não acreditava em mim. Seu apoio me deu coragem para enfrentar todos os desafios e alcançar meus objetivos.

Aos meus queridos pais e ao meu irmão, sou grata por todo o suporte e confiança que depositaram em mim desde o início da minha trajetória. Vocês sempre acreditaram em meu potencial e me encorajaram a ir além. A presença de vocês, mesmo que de longe, é um lembrete diário de que nunca estou sozinha.

Ao professor Dr. Eduardo Borges de Melo, cujo conhecimento foi fundamental para o meu crescimento acadêmico. Sua orientação, paciência e compromisso em compartilhar seus conhecimentos foram essenciais para o desenvolvimento desta dissertação.

SUMÁRIO

RESUMO	11
ABSTRACT	12
1. INTRODUÇÃO	13
2. OBJETIVOS	14
3. REVISÃO BIBLIOGRÁFICA.....	15
3.1. Doença de Alzheimer	15
3.1.1. Caracterização	15
3.1.2. Fisiopatologia	16
3.2. Potenciais alvos para o tratamento da doença de Alzheimer	18
3.2.1. Caspase 3.....	19
3.3 Planejamento de fármacos	20
3.3.1 Planejamento de fármacos auxiliados por computador (CADD)	21
3.3.2 Método QSAR e estudo de Docking	22
3.3.3 Triagem virtual	22
4. PROCEDIMENTO EXPERIMENTAL.....	23
4.1. Modelo QSAR	23
4.1.1. Obtenção e preparação do conjunto de dados:.....	23
4.1.2. Descritores moleculares.....	26
4.1.3. Seleção de variáveis e construção do modelo	26
4.1.4. Validação do modelo.....	27
4.2. Triagem virtual baseada por similaridade 2D	28
4.2.1. Toxicidade in silico e domínio de aplicabilidade	28
4.3. Estudo de Docking.....	29
4.3.1. Modelagem farmacofórica	29
4.3.2. Triagem virtual 3D	29

4.3.3.	Avaliação da toxicidade in silico	30
4.3.4.	Avaliação farmacocinética.....	30
4.3.5.	Docking.....	30
4.3.6.	Aplicação do modelo QSAR 2D.....	32
5.	RESULTADOS E DISCUSSÃO	33
5.1.	Estudo QSAR.....	33
5.2.	Triagem virtual 2D.....	39
5.3.	Triagem virtual 3D.....	40
5.4.	Estudo de Docking.....	43
6.	CONCLUSÃO	48
7.	REFERÊNCIAS.....	49
	APÊNDICE	53

LISTA DE TABELAS

Tabela 4.1: Compostos derivados da estrutura base A e B com respectivas denotações SMILES e pIC_{50}	24
Tabela 5.1: Descritores utilizados para o modelo	34
Tabela 5.2: Parâmetros estatísticos obtidos para o conjunto de treinamento após PLS	35
Tabela 5.3: Resultados gerados utilizando a metodologia γ -randomization	36
Tabela 5.4: Resultados gerados utilizando a metodologia Leave-N-out	38
Tabela 5.5: Resultados da validação externa	39
Tabela 5.6: Novos compostos hits obtidos pelo estudo de triagem virtual	41
Tabela 5.7: Dados de energia de ligação obtidos no estudo de docking	46
Tabela 5.8: Código smiles e atividade prevista dos compostos com os melhores resultados de energia de ligação	47

LISTA DE FIGURAS

Figura 4.1: Gráfico de dispersão dos valores de pIC_{50} demonstrando o gap observado entre os valores de 1,66 a 4,04. Os valores destacados na cor amarela foram retirados do conjunto amostral.	24
Figura 4.2. Configurações dos filtros adotados dentro programa Pharmit. Hit Reduction utilizado as bibliotecas Nubbe, Coconut e Life Chemicals (A), Hit Screening utilizado nas bibliotecas Nubbe e Life Chemicals (B) e, Hit Screening utilizado na biblioteca Coconut (C).	29
Figura 4.3. Arquivos gerados dentro programa Discovery Studio. Proteína CASPASE 3 (A), estrutura 3D (B) e 2D (C) do inibidor co-cristalizado.	31
Figura 4.4: Representação gráfica do cubo com as dimensões definidas no software PyRx/AutoDock Vina.	31
Figura 5.1: Contribuição dos descritores moleculares para o vetor de regressão.	34
Figura 5.2: Representação gráfica da sobreposição dos dados de R2 e Q2LOO e seus interceptos utilizando a metodologia y-randomization	37
Figura 5.3: Representação gráfica do resultado Leave-14-Out	38
Figura 5.4: Gráfico de leverage x resíduos de Student para avaliação da presença de outlier	39
Figura 5.5: Representação dos sítios farmacofóricos selecionados no servidor Pharmit	43
Figura 5.6: Disposição das moléculas no modelo boiled-egg	44
Figura 5.7: Sobreposição do ligante na sua forma nativa e da best pose obtida no redocking	45
Figura 5.8: A) Representação do ancoramento da molécula A com o proteína 1GFW (superfície colorida pelo grau de hidrofobicidade); B) Diagrama 2D exibindo os principais tipos de interação entre o ligante e os resíduos de aminoácido da proteína.	47
Figura 5.9: A) Representação do ancoramento da molécula B com o proteína 1GFW (superfície colorida pelo grau de hidrofobicidade); B) Diagrama 2D exibindo os principais tipos de interação entre o ligante e os resíduos de aminoácido da proteína	48
Figura 5.10: A) Representação do ancoramento da molécula C com o proteína 1GFW (superfície colorida pelo grau de hidrofobicidade); B) Diagrama 2D exibindo os principais tipos de interação entre o ligante e os resíduos de aminoácido da proteína	48
Figura 5.11: A) Representação do ancoramento da molécula D com o proteína 1GFW (superfície colorida pelo grau de hidrofobicidade); B) Diagrama 2D exibindo os principais tipos de interação entre o ligante e os resíduos de aminoácido da proteína.	49

Doença de alzheimer: estudos in silico de inibidores da proteina caspase 3

Paula Beatriz Jesus Santos

RESUMO

A caspase 3 é uma enzima que desempenha importante papel na morte celular programada (apoptose). Sua ativação desenfreada em células neurais, cada vez mais estudada, é atrelada a uma das causas da perda de neurônios na Doença de Alzheimer (DA). Assim, este alvo é considerado um alvo promissor para o desenvolvimento de novos fármacos. Posto isto, nesse trabalho foi realizado um estudo de relação estrutura-atividade quantitativa 2D (QSAR-2D) baseado em 71 compostos descritos na literatura como inibidores da caspase 3. O modelo obtido, utilizando regressão por PLS, foi aprovado nos quesitos de validação interna e externa, e foi utilizado para realizar previsão de novas moléculas oriundas de um estudo de triagem virtual baseado em similaridade 2D e posteriormente em 3D. Após avaliação de toxicidade e propriedades farmacocinéticas, foram identificados 23 compostos promissores como potenciais inibidores da caspase 3. Esses resultados fornecem importantes subsídios para futuros estudos *in vitro*, visando o desenvolvimento de novos tratamentos para a DA. A abordagem computacional utilizada neste estudo demonstrou sua eficácia na identificação de compostos com potencial atividade biológica, contribuindo para otimizar o processo de descoberta de novos fármacos direcionados ao tratamento da DA.

Palavras-Chave: Relação estrutura-atividade quantitativa; Caspase 3; Doença de Alzheimer (DA).

Alzheimer's disease: in silico studies of caspase 3 protein inhibitors

Paula Beatriz Jesus Santos

ABSTRACT

Caspase 3 is an enzyme that plays an important role in programmed cell death (apoptosis). Its unbridled activation in neural cells, increasingly studied, is linked to one of the causes of neuron loss in Alzheimer's Disease (AD). Thus, this target is considered a promising target for the development of new drugs. That said, in this work a 2D quantitative structure-activity relationship (QSAR-2D) study was carried out based on 71 compounds described in the literature as caspase 3 inhibitors. The model obtained, using regression by PLS, was approved in terms of internal validation and external, and was used to predict new molecules from a virtual screening study based on 2D and later 3D similarity. After evaluating the toxicity and pharmacokinetic properties, 23 promising compounds were identified as potential caspase 3 inhibitors. These results provide important subsidies for future in vitro studies, aiming at the development of new treatments for AD. The computational approach used in this study demonstrated its effectiveness in identifying compounds with potential biological activity, contributing to optimize the discovery process of new drugs aimed at the treatment of AD.

Keywords: Quantitative structure-activity relationship; CASPASE 3; Alzheimer's Disease (AD).

1. INTRODUÇÃO

A doença de Alzheimer (DA) é um tipo de demência que afeta mais de 50 milhões de pessoas em todo o mundo [1]. Embora a medicina tenha avançado no sentido de melhorar a qualidade de vida, aliviar os sintomas e prolongar a sobrevivência dos pacientes, atualmente não existem medicamentos capazes de reverter a progressão da doença. Apesar da abundância de informações disponíveis na literatura sobre os mecanismos envolvidos na progressão da DA, os avanços em estudos *in vivo* têm sido limitados.

Diante dessa realidade, a necessidade de disponibilizar mais opções de fármacos no mercado é eminente, não apenas devido ao crescimento da população mundial, mas também pela ausência de um tratamento curativo. Isso justifica a busca constante por compostos que possam atuar como protótipos para o desenvolvimento de novos medicamentos.

Nessa perspectiva, compostos que modulam a atividade enzimática da caspase 3, inibindo-a, podem desempenhar um papel fundamental na regulação da morte neuronal. Além da doença de Alzheimer, a caspase 3 é alvo de estudo na doença de Parkinson, a segunda patologia neurodegenerativa mais comum em todo o mundo [2]. Portanto, a descoberta de um fármaco seletivo que atue nesse alvo biológico é de grande interesse tanto para a comunidade científica quanto para a indústria farmacêutica, uma vez que atualmente não existem fármacos aprovados para suprimir essa proteína.

A criação de um modelo quantitativo de relação estrutura-atividade auxiliará na previsão de compostos com potencial de inibição da enzima caspase 3, além de fornecer subsídios para reduzir os custos gerais relacionados à síntese desses compostos. Essa abordagem pode impulsionar a descoberta de novos candidatos a fármacos e acelerar o desenvolvimento de tratamentos mais eficazes para doenças neurodegenerativas, como a doença de Alzheimer.

2. OBJETIVOS

Considerando o cenário apresentado, esse trabalho tem como objetivo propor um modelo QSAR a partir de um conjunto de moléculas que apresentam capacidade inibitória frente à caspase 3 selecionados na literatura científica.

Ademais, este trabalho também tem como objetivo a busca por novos compostos *hits* com potencial capacidade de inibição da protease estudada, utilizando abordagens de triagem virtual por similaridade 2D e por triagem utilizando uma abordagem de planejamento baseado na estrutura do alvo terapêutico.

3. REVISÃO BIBLIOGRÁFICA

3.1. Doença de Alzheimer

3.1.1. Caracterização

A Doença de Alzheimer é a denominação de uma enfermidade neurodegenerativa caracterizada pela perda progressiva dos neurônios, resultando em danos nas funções cognitivas, as quais se associam às alterações do comportamento e incapacidade funcional em seu portador [1,2]. A principal manifestação da doença de Alzheimer é a perda de memória, porém logo surgem outros sintomas como alterações de humor, agressividade, dificuldades para realizar atividades do dia a dia e até o desligamento total da realidade em que se vive o paciente [1-3].

Os primeiros sinais que indiciam essa doença é a deficiência da memória recente, sendo que as lembranças remotas são preservadas até certo estágio da doença. Além das dificuldades de atenção e fluência verbal, outras funções cognitivas deterioram à medida que a patologia evolui, entre elas a capacidade de fazer cálculos, as habilidades visuais e espaciais e a capacidade de utilizar objetos comuns e ferramentas. Todavia, até a doença atingir um estado avançado, não é observado perda da capacidade de vigília, lucidez e fraqueza motora [3,4]. Além do mais, é importante ressaltar que esses sintomas são frequentemente acompanhados por distúrbios comportamentais, como agressividade, alucinações, hiperatividade, irritabilidade e depressão [3].

A designação “Doença de Alzheimer” se deve ao neurologista alemão Alois Alzheimer, que foi o primeiro a descrever sua ocorrência quando diagnosticou o caso de um paciente de cinquenta e um anos de idade que apresentava sintomas psiquiátricos, como delírios, associados à perda progressiva da memória. O quadro do paciente foi se agravando ao longo de quatro anos, quando ao fim desse período já havia perdido completamente suas faculdades mentais. Após realização de necropsia, foi possível verificar atrofia cerebral generalizada, degeneração neurofibrilar, e deposição de material metabólico desconhecido [5,6].

De acordo com dados da Organização Mundial de Saúde, a doença de Alzheimer é a forma mais comum de demência, e atualmente e pode contribuir para 60-70% dos casos. Dados do Ministério da Saúde informam que no Brasil, cerca de 1,2 milhão pessoas vivem com alguma forma de demência e 100 mil novos casos são diagnosticados por ano. Em todo o mundo, o número chega a 50 milhões de pessoas. Segundo estimativas da

Alzheimer's Disease International (<https://www.alzint.org>), os números poderão chegar a 74,7 milhões em 2030 e 131,5 milhões em 2050, devido ao envelhecimento da população. Esse cenário mostra que a doença caracteriza uma crise global de saúde que deve ser enfrentada [7].

O principal fator de risco para a doença de Alzheimer é a idade, sendo mais comum a manifestação dos sintomas após os 65 anos de idade. Todavia, ela pode acometer indivíduos mais jovens, sendo denominada de doença pré-senil, ocorrendo normalmente em torno de 40 anos de idade. Devido ao fato de ser uma doença de evolução progressiva, classifica-se em três fases principais de acordo com a gravidade do comprometimento cognitivo e com o grau de independência do indivíduo, sendo as fases leve, moderada ou grave. Normalmente, ela costuma evoluir pelos estágios de agravamento de forma lenta [3-7].

- Fase inicial: caracteriza-se como um quadro leve da doença, onde o paciente costuma apresentar alterações na memória, na personalidade e nas habilidades visuais e espaciais. Verifica-se nesse estágio, transtornos de memória recente e diminuição na capacidade de aprender novas informações;
- Fase moderada: o paciente começa a apresentar distúrbios de memória mais acentuados, além de dificuldade para falar, realizar tarefas simples e coordenar movimentos. Também é comum a presença de agitação e insônia;
- Estágio avançado: o paciente necessita de assistência integral, sendo comum apresentar resistência à execução de tarefas diárias, bem como incontinência urinária e fecal, e dificuldade para se alimentar.

3.1.2. Fisiopatologia

Múltiplos fatores são considerados possíveis para a incidência da doença de Alzheimer, sendo os principais a idade, o fator genético, além do estilo de vida. Além do mais, a ausência de um tratamento definitivo está relacionada com a falta de conhecimento sobre os eventos por trás da doença [5].

Embora o grau e o tipo de sintomas possam variar muito de paciente para paciente, observações *post mortem* no sistema nervoso central de indivíduos acometidos pela doença evidenciaram algumas características em comum. Do ponto de vista macroscópico, o cérebro do paciente doente apresenta atrofia difusa e não uniforme, sendo mais proeminente nas regiões frontais, temporais, e parietais, afetando sobretudo

áreas corticais associativas. Microscopicamente, observa-se a maciça perda sináptica e morte neuronal observada nas regiões cerebrais responsáveis pelas funções cognitivas, incluindo o córtex cerebral, o hipocampo, o córtex entorrinal e o estriado ventral. Além disso, também é observado acúmulo de placas neuríticas anormais e presença de emaranhados neurofibrilares [3-5, 7].

Na fisiopatologia da doença existem duas proteínas principais envolvidas: a beta-amiloide e a Tau. A beta-amiloide atinge níveis anormais no cérebro dos indivíduos afetados, resultando na formação de placas neuríticas que se acumulam entre os neurônios e interferem na função celular. Por sua vez, a proteína Tau também apresenta níveis desregulados e forma emaranhados neurofibrilares dentro dos neurônios, causando o bloqueio do sistema de transporte intracelular. Esses eventos contribuem para a progressão da doença, resultando em danos neuronais e comprometimento cognitivo [5].

As placas neuríticas também chamadas de placas senis, são formadas em um processo externo ao neurônio, no qual a proteína beta amiloide envolve a célula neural e provoca a sua morte. Já a formação dos emaranhados neurofibrilares também provoca a morte dos neurônios, porém é um processo interno ao neurônio. A formação desses emaranhados ocorre devido à hiperfosforilação da proteína Tau, formando filamentos helicoidais pareados dentro do neurônio, com desorganização da citoarquitetura dos túbulos neurais [5]. Estas alterações ocorrem em estruturas do lobo temporal medial, incluindo o hipocampo e o giro para-hipocampal, consideradas estruturas essenciais para os processos de memória. Com a evolução da doença, o processo degenerativo se espalha para o neocórtex de associação, atingindo áreas cerebrais responsáveis por outros processos cognitivos [6].

É relevante ressaltar que o fator genético exerce uma influência significativa na maioria dos casos da doença, com mais de vinte genes identificados como sendo relacionados a essa condição. Entre esses genes, o mais amplamente associado ao aumento do risco de desenvolvimento da doença de Alzheimer é o gene apolipoproteína E tipo 4 (APOE4), que está presente em cerca de 50% de todos os casos. Outros genes envolvidos na demência incluem a Proteína Precursora Amiloide β (*amyloid precursor protein*, APP), bem como os genes Presenilina 1 (PSEN1) e Presenilina 2 (PSEN2). No entanto, esses genes são responsáveis por uma proporção pequena, estimada em menos de 1%, dos casos da doença. É importante notar que mutações genéticas envolvendo esses genes estão associadas a casos de herança dominante e início precoce da doença, respondendo por 30% a 50% dos casos nessa categoria. Dados da literatura científica

demonstram que indivíduos que herdam mutações nos genes APP ou PSEN1 têm garantia de desenvolver a doença, enquanto aqueles com mutações hereditárias no gene PSEN2 têm aproximadamente 95% de chance de desenvolvê-la [5,8].

3.2. Potenciais alvos para o tratamento da Doença de Alzheimer

O tratamento da Doença de Alzheimer está diretamente relacionado ao seu estágio, e é essencialmente paliativo, visando retardar a progressão da doença e melhorar a qualidade de vida do paciente. Este processo geralmente envolve uma abordagem combinada, utilizando tanto estratégias farmacológicas como intervenções psicossociais. Os tratamentos não farmacológicos consistem em atividades que visam estimular a cognição do paciente, com um papel fundamental desempenhado pelo apoio familiar. Por outro lado, os tratamentos farmacológicos se baseiam na inibição das colinesterases, aumentando a disponibilidade da acetilcolina, e na diminuição dos níveis de glutamato nos neurônios. Atualmente, inúmeras substâncias psicoativas têm sido propostas para preservar ou restabelecer a cognição, o comportamento e as habilidades funcionais do indivíduo com doença de Alzheimer. Nesse sentido, o tratamento farmacológico para a doença de Alzheimer pode ser categorizado em quatro níveis: terapia específica, abordagem profilática, tratamento sintomático e terapia complementar [6,9].

De acordo com o *Food and Drug Administration* (FDA), órgão regulatório de medicamentos dos Estados Unidos da América (EUA), dois grupos de medicamentos são aprovados para o tratamento sintomático da doença de Alzheimer. O primeiro grupo inclui fármacos a donepezila, rivastigmina e galantamina. Esses medicamentos aumentam a disponibilidade de acetilcolina na fenda sináptica graças a inibição da enzima acetilcolinesterase, o que permite uma melhora dos sintomas. O segundo grupo é composto pelo inibidor do receptor NMDA memantina, que reduz a entrada excessiva de íons de cálcio nos neurônios pós-sinápticos. O influxo excessivo desse íon nos receptores, devido à alta despolarização causada pelo glutamato, promove alterações de memória e destruição neuronal [10-11]. Entretanto, é válido ressaltar que ainda não há cura definitiva para esta doença, mesmo com a condição reversível que pode ser obtida em alguns casos, de acordo com o tratamento empregado [12]. Assim, continua sendo necessário à pesquisa e o desenvolvimento de novos medicamentos para o tratamento desta doença.

Recentemente, o FDA, através do programa de aprovação acelerada, liberou o uso do anticorpo monoclonal Aducanumabe para o tratamento da DA. Apesar dos benefícios

ainda serem incertos, esse fármaco promove a redução placas senis do cérebro dos pacientes portadores da doença [13-20].

3.2.1. Caspase 3

A evolução da doença de Alzheimer está associada à perda de conexões entre as células cerebrais, resultante da morte celular, o que leva ao agravamento dos sintomas cognitivos. Estudos recentes sugerem que a caspase-3 ativada desempenha um papel significativo na perda progressiva de neurônios associada à doença. As caspases são proteases aspárticas de cisteína que compõem uma família de endoproteases importantes na regulação do equilíbrio celular, através da regulação da apoptose (morte celular programada) e inflamação. Elas possuem um resíduo de cisteína em seu sítio ativo e clivam seus substratos quase que invariavelmente após resíduos de ácido aspártico. Cada tipo de caspase desempenha funções específicas na coordenação da morte celular programada, ou apoptose [3].

Este termo refere-se a morte celular que ocorre de forma não acidental. Esse tipo de morte celular pode ocorrer em diversas situações, como por exemplo, na organogênese e hematopoese normal e patológica, na reposição fisiológica de certos tecidos maduros, na atrofia dos órgãos, na resposta inflamatória e na eliminação de células após dano celular por agentes genotóxicos. As características morfológicas da apoptose são bastante coordenadas, sendo esse fenômeno considerado bastante rápido, onde inicialmente corre uma retração da célula, ocasionando a perda da aderência com a matriz extracelular e células vizinhas. Tais alterações morfológicas resultam de uma sequência de eventos moleculares e bioquímicos específicos e geneticamente regulados. Nesse sentido, a apoptose é considerada um programa de morte celular extremamente regulado e de grande eficiência, sendo crucial para a manutenção da homeostase [13,15-16].

De acordo com a literatura, existem 14 caspases humanas, agrupadas com base em sua especificidade e atividade biológica. Das caspases conhecidas, seis (caspases 3, 6, 7, 8, 9, 10) desempenham um papel na apoptose. Por outro lado, as caspases 1, 4, 5, 11, 12, 13 e 14 estão envolvidas na maturação de citocinas, e sua contribuição na apoptose permanece não esclarecida [13,15-16].

Estudos têm demonstrado a presença das caspases 1 a 3 e 5 a 9 nos cérebros de indivíduos afetados pela doença de Alzheimer. A caspase 3, em particular, é uma caspase classificada como efetora, e possui um papel importante na degeneração sináptica. Ela

atua na clivagem da proteína tau e APP, contribuindo para a formação de placas β -amiloides no cérebro dos pacientes. Assim, medicamentos direcionados para prevenir a ativação e execução da apoptose pela caspase 3 representam uma abordagem promissora no tratamento da doença de Alzheimer [13,15-16].

3.3 Planejamento de fármacos

A indústria farmacêutica é a principal responsável pela produção, distribuição, e comercialização de medicamentos, sendo um dos setores industriais de maior relevância no cenário mundial. Com o objetivo de promover a saúde e atender às necessidades da população mundial, há um alto investimento financeiro em pesquisa, desenvolvimento e inovação de produtos. A indústria farmacêutica utiliza tecnologia de ponta e conta com mão de obra altamente qualificada para garantir a qualidade, segurança e eficácia dos medicamentos destinados ao tratamento e controle de doenças [1,12]. No Brasil o setor farmacêutico representa um faturamento anual em torno de 10,3 bilhões de dólares devido ao consumo desses produtos, colocando o país entre os dez maiores mercados farmacêuticos do mundo [1].

Várias etapas antecedem o emprego e a comercialização de qualquer medicamento no tratamento de enfermidades. Estas incluem a pesquisa de moléculas biologicamente ativas, produção em escala industrial e o processamento final. Todo este processo pode levar muitos anos, e necessitar de altíssimos investimentos, incluindo aí o desenvolvimento da produção em escala industrial. Segundo a literatura, estima-se que a cada 10.000 moléculas sintetizadas nas fases iniciais de pesquisa, apenas uma chega a fase final com a função de princípio ativo [1,13].

O desenvolvimento de um fármaco requer investimentos significativos, muitas vezes na faixa dos milhões ou bilhões de dólares, por parte da indústria farmacêutica responsável. Esses investimentos são destinados à pesquisa, ensaios clínicos e outros processos necessários para garantir a eficácia e a segurança do medicamento [12]. Assim, do processo de descoberta e desenvolvimento de um novo fármaco até ele chegar ao mercado demora em torno de 10 anos, com custos em torno de 2,6 bilhões de dólares [5, 18].

Antes da evolução da síntese orgânica, os compostos químicos com propriedades farmacológicas eram majoritariamente obtidos a partir de fontes naturais. No entanto, com o progresso da tecnologia e a busca por melhor custo-benefício, os compostos

naturais bioativos foram gradualmente substituídos por compostos sintéticos. Como resultado, a indústria farmacêutica desempenha um papel fundamental na fabricação de substâncias químicas para a produção de medicamentos, abrangendo diversas classes terapêuticas e especialidades farmacêuticas, tanto alopáticas quanto homeopáticas [5,12,14].

3.3.1 Planejamento de fármacos auxiliados por computador (CADD)

Buscando a otimização do processo de desenvolvimento de novos medicamentos, a modelagem molecular se tornou uma ferramenta essencial no planejamento de fármacos. Essa abordagem permite reduzir o número de compostos a serem sintetizados, uma vez que fornece informações sobre propriedades estruturais e eletrônicas. A modelagem molecular se baseia em um conjunto de técnicas computacionais que possibilitam uma compreensão quantitativa e qualitativa das interações moleculares. Desse modo, é possível compreender o funcionamento dos sistemas biológicos, com base em informações relacionadas a aspectos eletrônicos, moleculares, atômicos e forças hidrofóbicas [5,18].

Dentre as técnicas mais utilizadas, encontram-se os métodos de Planejamento de Fármacos Auxiliado por Computador (*Computer-Aided Drug Design, CADD*). A aplicação dessas técnicas computacionais contribui de forma significativa ao processo de desenvolvimento de novos fármacos, reduzindo a necessidade de síntese em larga escala, experimentos extensivos e a utilização de animais de laboratório [21-22].

O CADD abrange uma ampla gama de técnicas computacionais e recursos para o armazenamento, gestão, análise e modelagem de compostos. Essa metodologia tem sido importante na descoberta e desenvolvimento de novos fármacos nos últimos anos, especialmente na identificação de novos candidatos a fármacos e seleção de compostos para estudos mais aprofundados. Além disso, o CADD também auxilia na avaliação precoce da potência, seletividade e potencial desvantagens tóxicas ou farmacocinéticas [21-22].

Por meio do CADD, é possível obter informações importantes sobre a natureza estérica, eletrônica e lipofílica da provável interação do fármaco (ligante) com o alvo biológico (receptor), além de estabelecer as relações quantitativas entre a estrutura química e a atividade biológica de uma série de compostos, proporcionando uma melhor compreensão do mecanismo de ação responsável pelo efeito biológico [21-22]. Assim,

com as técnicas de CADD é possível obter resultados experimentais antes mesmo de serem realizados experimentalmente em laboratório, economizando tempo e recurso.

3.3.2 Método QSAR e estudo de Docking

Dois dos métodos de CADD mais empregados atualmente são o ancoramento molecular (*molecular docking*) e os métodos de relação estrutura-atividade quantitativa (*Quantitative Structure–Activity Relationship*, QSAR).

O ancoramento molecular visa estudar a interação entre um composto candidato a fármaco e o sítio ativo de uma macromolécula alvo, buscando compreender a formação de um novo complexo molecular. Por outro lado, o QSAR estima as propriedades moleculares dos compostos com o objetivo de explicar sua atividade biológica com base em sua estrutura. De acordo com Ribeiro, o QSAR é um método que considera a variação da atividade biológica observada em um conjunto de compostos com base em suas diferenças estruturais gerando um modelo matemático formado por parâmetros ou descritores, avaliado por índices estatísticos, que tem o objetivo de estimar/prever a atividade de novas moléculas. Por meio dessas relações, é possível compreender como as forças de interação estão envolvidas no mecanismo de ação dos fármacos, além de como se fundamentam no desenvolvimento de novos compostos com as propriedades biológicas almejadas. Estes modelos são construídos utilizando processos matemáticos comuns, como regressão linear múltipla (*Multiple Linear Regression*, MLR), análise de componentes principais (*Principal Components Analysis*, PCA), etc [18,25-27].

Os descritores moleculares desempenham um papel fundamental na construção dos modelos QSAR. Eles são informações químicas que descrevem a estrutura molecular dos compostos e são transformados em números que podem ser analisados por métodos de análise multivariada. Esses descritores fornecem informações sobre o tamanho, forma, distribuição de cargas, interações intramoleculares, entre outros aspectos da molécula.

3.3.3 Triagem virtual

A triagem virtual é poderosa ferramenta computacional utilizada com o objetivo de selecionar as melhores moléculas a partir das técnicas computacionais. Ela consiste em duas estratégias principais: a triagem baseada na estrutura do alvo molecular (*structure-based virtual screening*, SBVS) e a triagem baseada em ligantes (*ligand-based*

virtual screening, LBVS). Embora essas técnicas ainda estejam em constante aprimoramento, elas têm sido amplamente utilizadas na área [27-32].

Na SBVS utiliza-se o conhecimento sobre a estrutura tridimensional do alvo biológico para selecionar ligantes que apresentem afinidade e complementaridade com o sítio de ligação. Já a abordagem LBVS pode ser utilizada sem o alvo biológico, estando sua estrutura tridimensional disponível ou não. Nesse caso, a triagem é realizada com base nas características dos ligantes conhecidos [27-32]. Ambas as técnicas de triagem virtual permitem extrair informações estruturais e físico-químicas tanto dos ligantes quanto do receptor. Essas informações possibilitam a elaboração de um conjunto de regras para averiguar a afinidade de formação do complexo receptor-ligante [32-33].

Portanto, a triagem virtual tem se estabelecido como uma ferramenta indispensável na fase inicial do desenvolvimento de medicamentos, proporcionando uma abordagem mais eficiente, econômica e sustentável para a descoberta de novos compostos com potencial farmacológico.

4. PROCEDIMENTO EXPERIMENTAL

4.1. Modelo QSAR

4.1.1. Obtenção e preparação do conjunto de dados:

Um conjunto de dados formado por 78 compostos descritos como inibidores da caspase 3 foram selecionados da literatura [27-29]. É importante destacar que os estudos mencionados utilizaram a mesma metodologia para mensuração dos valores de atividade biológica, por meio de um ensaio de doseamento fluorimétrico. Nesse ensaio, foi mensurado o acúmulo de um produto fluorogênico, a 7-amino-4-metilcumarina (7-AMC). Os valores de IC_{50} (concentração necessária para inibir a atividade enzimática em 50%), apresentados em μM , foram convertidos para pIC_{50} ($-\log IC_{50}$, ou $\log 1/IC_{50}$) para normalização dos dados. A faixa de variação de atividade obtida foi de 1,66 a 9,39 sendo observado um *gap* de 1,11 unidades logarítmicas de 2,93 a 4,04. Uma vez que a faixa de atividade é ampla e o número de compostos é extenso, foram excluídos os compostos com as atividades de 1,66 a 2,93. Dessa forma, a faixa final de variação de atividade foi de 5,35 unidades logarítmicas (pIC_{50} 4,04 a 9,39), compreendendo um total de 71 compostos.

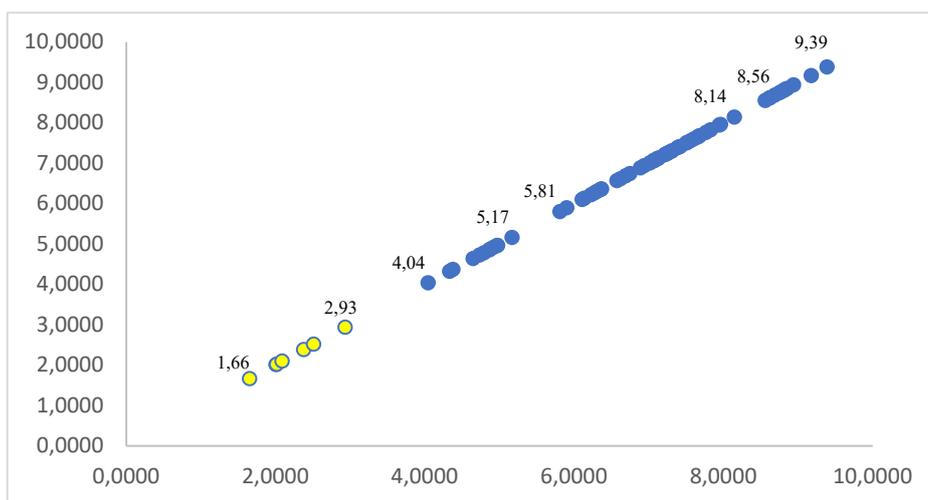
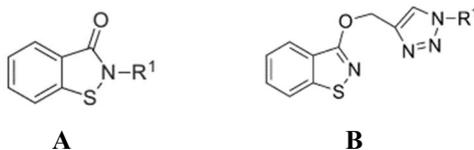


Figura 4.1: Gráfico de dispersão dos valores de pIC_{50} demonstrando o *gap* observado entre os valores de 1,66 a 4,04. Os valores destacados na cor amarela foram retirados do conjunto amostral.

Utilizando o editor *ChemSketch* (<https://www.acdlabs.com>), foram obtidas as notações *SMILES* (do inglês, *Simplified Molecular Input Line Entry Specification*) de todos os compostos (Tabela 4.1).

Tabela 4.1: Compostos derivados da estrutura base A e B com respectivas denotações *SMILES* e pIC_{50}



Composto	Derivado	SMILES	pIC_{50}
1	A	<chem>O=C1NSc2ccccc21</chem>	4,330
2	A	<chem>O=C1c2ccccc2SN1C</chem>	4,373
3	A	<chem>O=C1c2ccccc2SN1CC</chem>	4,040
4	A	<chem>O=C1c2ccccc2SN1CC=C</chem>	4,644
5	A	<chem>O=C1c2ccccc2SN1c1cccc1</chem>	6,273
6	A	<chem>Cc1ccc(cc1)N1Sc2ccccc2C1=O</chem>	6,332
7	A	<chem>O=C1c2ccccc2SN1Cc1cccc1</chem>	4,967
8	A	<chem>COc1ccc(cc1)CN1Sc2ccccc2C1=O</chem>	4,918
9	A	<chem>Fc1ccc(cc1)CN1Sc2ccccc2C1=O</chem>	4,865

10	A	O=C1c2cccc2SN1CCc1cccc1	4,732
11	A	O=C1c2cccc2SN1C(=O)Nc1cccc1	7,119
12	A	COc1ccc(cc1)NC(=O)N1Sc2cccc2C1=O	4,970
13	A	Fc1cccc1NC(=O)N1Sc2cccc2C1=O	4,789
14	A	Fc1cccc(c1)NC(=O)N1Sc2cccc2C1=O	6,359
15	A	Fc1ccc(cc1)NC(=O)N1Sc2cccc2C1=O	6,227
16	A	O=C1c2cccc2SN1C(=O)NCc1cccc1	7,066
17	A	CC(NC(=O)N1Sc2cccc2C1=O)c1cccc1	7,071
18	A	COc1cccc1CNC(=O)N1Sc2cccc2C1=O	7,056
19	A	COc1ccc(c1)CNC(=O)N1Sc2cccc2C1=O	7,131
20	A	COc1ccc(cc1)CNC(=O)N1Sc2cccc2C1=O	7,398
21	A	Fc1cccc1CNC(=O)N1Sc2cccc2C1=O	6,686
22	A	Fc1cccc(c1)CNC(=O)N1Sc2cccc2C1=O	6,618
23	A	Fc1ccc(cc1)CNC(=O)N1Sc2cccc2C1=O	6,939
24	A	O=C1c2cccc2SN1C(=O)NCCc1cccc1	7,509
25	A	Cc1ccc(nc1)NC(=O)N1Sc2cccc2C1=O	7,419
26	A	O=C1c2cccc2SN1C(=O)NCc1cccn1	7,382
27	A	O=C1c2cccc2SN1C(=O)NCCc1cccn1	7,306
28	A	O=C1c2cccc2SN1C(=O)NCc1cco1	7,094
29	A	O=C1c2cccc2SN1C(=O)NCCc1cccs1	7,508
30	A	O=C1c2cccc2SN1C(=O)NCCCc1cccc1	7,309
31	A	O=C1c2cccc2SN1C(=O)Nc1cccc1N1CCOCC1	6,569
32	A	O=C1c2cccc2SN1C(=O)Nc1cccc(c1)N1CCOCC1	7,256
33	A	O=C1c2cccc2SN1C(=O)Nc1ccc(cc1)N1CCOCC1	8,939
34	A	CN1CCN(CC1)c1cccc1NC(=O)N1Sc2cccc2C1=O	4,868
35	A	CN1CCN(CC1)c1cccc(c1)NC(=O)N1Sc2cccc2C1=O	6,743
36	A	CN1CCN(CC1)c1ccc(cc1)NC(=O)N1Sc2cccc2C1=O	7,208
37	A	O=C1c2cccc2SN1C(=O)Nc1ccc(cc1)N1CCCC1	7,009
38	A	O=C1c2cccc2SN1C(=O)Nc1ccc(cc1)c1cccn1	7,656
39	A	O=C1c2cccc2SN1C(=O)NCc1ccc(cc1)N1CCOCC1	7,245
40	A	O=C(c1cccc1)c1cc(ccc1)NC(=O)N1Sc2cccc2C1=O	6,366
41	A	O=S(=O)(N1CCOCC1)c1ccc(cc1)NC(=O)N1Sc2cccc2C1=O	6,106
42	A	CN1CCN(CC1)S(=O)(=O)c1ccc(cc1)NC(=O)N1Sc2cccc2C1=O	7,218
43	A	O=C1c2cccc2SN1CC(=O)Nc1ccc(cc1)N1CCOCC1	5,896
44	A	O=C1c2cccc2SN1CC(=O)Nc1ccc(cc1)N1CCCC1	5,807
45	A	O=C1c2cccc2SN1CC(=O)Nc1ccc(cc1)c1cccn1	6,300
46	A	O=C1c2cccc2SN1CC(=O)Nc1ccc(cc1)n1cccc1	6,137
47	B	Fc1ccc(cc1)n1cc(nn1)COc1nsc2cccc21	7,667
48	A	Fc1cccc1n1cc(nn1)CN1Sc2cccc2C1=O	8,770
49	A	Fc1ccc(cc1)n1cc(nn1)CN1Sc2cccc2C1=O	8,836
50	A	COc1ccc(c1)n1cc(nn1)CN1Sc2cccc2C1=O	9,387
51	A	[O-][N+](=O)c1cccc1n1cc(nn1)CN1Sc2cccc2C1=O	8,807
52	A	[O-][N+](=O)c1cccc(c1)n1cc(nn1)CN1Sc2cccc2C1=O	8,815
53	A	O=C(O)c1ccc(cc1)n1cc(nn1)CN1Sc2cccc2C1=O	8,629
54	A	Oc1ccc(cc1)n1cc(nn1)CN1Sc2cccc2C1=O	8,854
55	A	O=C(OCC)c1ccc(cc1)n1cc(nn1)CN1Sc2cccc2C1=O	8,143

56	A	<chem>O=C1c2cccc2SN1Cc1cn(nn1)c1ccc2OCCOc21</chem>	8,747
57	A	<chem>O=C1c2cccc2SN1Cc1cn(nn1)c1cccc1c1cccc1</chem>	8,686
58	A	<chem>O=C1c2cccc2SN1Cc1cn(nn1)c1ccc(cc1)N1CCOCC1</chem>	8,611
59	A	<chem>O=C1c2cccc2SN1Cc1cn(nn1)c1ccc(cc1)N1CCCCC1</chem>	8,558
60	A	<chem>O=C(c1cccc1)c1ccc(cc1)n1cc(nn1)CN1Sc2cccc2C1=O</chem>	9,174
61	A	<chem>Fc1cccc1n1cc(nn1)CNC(=O)N1Sc2cccc2C1=O</chem>	7,943
62	A	<chem>Fc1cccc(c1)n1cc(nn1)CNC(=O)N1Sc2cccc2C1=O</chem>	7,824
63	A	<chem>Fc1ccc(cc1)n1cc(nn1)CNC(=O)N1Sc2cccc2C1=O</chem>	7,959
64	A	<chem>COc1cccc1n1cc(nn1)CNC(=O)N1Sc2cccc2C1=O</chem>	7,251
65	A	<chem>[O-][N+](=O)c1cccc1n1cc(nn1)CNC(=O)N1Sc2cccc2C1=O</chem>	7,684
66	A	<chem>O=C(OCC)c1ccc(cc1)n1cc(nn1)CNC(=O)N1Sc2cccc2C1=O</chem>	7,764
67	A	<chem>O=C1c2cccc2SN1C(=O)NCc1cn(nn1)c1ccc(cc1)N1CCOCC1</chem>	6,885
68	A	<chem>O=C1c2cccc2SN1C(=O)NCc1cn(nn1)c1cccc1N1CCOCC1</chem>	7,556
69	A	<chem>CCN(CC)c1ccc(cc1)n1cc(nn1)CNC(=O)N1Sc2cccc2C1=O</chem>	7,606
70	A	<chem>O=C1c2cccc2SN1C(=O)NCc1cn(nn1)c1ccc(cc1)C1CCCN1</chem>	5,166
71	A	<chem>O=C1c2cccc2SN1C(=O)NCc1cn(nn1)c1cccc2CCc21</chem>	7,115

4.1.2. Descritores moleculares

Os descritores moleculares, variáveis independentes deste modelo, contendo apenas as informações que não dependem da estrutura tridimensional, foram obtidos por meio do programa *Dragon 6.0* utilizando as notações *SMILES*. Os descritores gerados passaram por um processo de redução no próprio programa, com a intenção de excluir aqueles sem informações relevantes. Assim, foram removidos descritores com valores constantes e quase constantes; com desvio padrão inferior a 0,001; descritores com ao menos um valor ausente; e descritores que possuísem correlação maior ou igual a 0,9 com outro descritor do conjunto (nesse processo, um dos descritores do par é removido de modo aleatório).

Uma última etapa de redução de variáveis foi realizada no programa *QSAR Modeling* (*download*: <http://lqta.iqm.unicamp.br>), onde foram excluídos os descritores que apresentaram valores absolutos do coeficiente de correlação de Pearson ($|r|$) com a atividade biológica inferior a 0,2. Ao final da etapa de redução, foi obtida uma matriz composta por 213 descritores.

4.1.3. Seleção de variáveis e construção do modelo

A matriz reduzida passou, a seguir, pelo processo de seleção de variáveis. O intuito desse processo é selecionar o subconjunto de descritores que melhor explica a variável dependente y . Para isso, foi utilizado o algoritmo Seleção de Preditores

Ordenados (*Ordered Predictors Selection*, OPS), disponível no programa *QSAR Modeling*. Nessa metodologia, os descritores são ordenados pelo seu grau de importância que é feito com base em um vetor informativo [26].

A construção do modelo foi feita utilizando o método de regressão por quadrados mínimos parciais (*Partial Least Squares*, PLS) com os 71 compostos que fazem parte do conjunto de dados [30], obtendo-se assim um modelo auxiliar. A escolha dessa metodologia está atrelada a capacidade de se trabalhar com um elevado número de variáveis correlacionadas entre si [26,31].

A partir desse modelo, o conjunto de dados foi dividido de modo aleatório em conjunto de treinamento (57 compostos, 80% do conjunto de dados) e conjunto de teste (14 compostos, 20% do conjunto de dados). Estes subconjuntos foram utilizados para a validação da qualidade de predição dos modelos obtidos pela abordagem de validação interna (com o conjunto de treinamento) e externa (com o conjunto teste).

4.1.4. Validação do modelo

A validação de um modelo QSAR permite testar, por meio de parâmetros estatísticos, a qualidade de predição da atividade para novos compostos em estudo [32-33]. Na validação interna, uma série de parâmetros recomendados pela literatura foram utilizados (Apêndice 1), sendo eles, o coeficiente de determinação (R^2); teste de *Fisher* com 95% de significância; erro quadrático médio de calibração (*RMSEC*); capacidade de previsão interna do modelo por meio da validação cruzada *Leave-One-Out* (LOO); robustez através da validação cruzada *leave-N-out* (LNO); e a presença de correlações espúrias utilizando o teste de *y-randomization*. Adicionalmente, fez-se a avaliação dos compostos que se comportam como pontos extremos (*outliers*) e valores influentes [26].

Para viabilizar a validação externa, os compostos do conjunto teste tiveram sua atividade biológica predita pelo modelo obtido. Nesse parâmetro, foram avaliados o coeficiente de determinação da validação externa (R^2_{pred}); as inclinações das linhas de regressões lineares de predição (k e k'); a média de r^2 modificado (*average r_m^2*); a diferença entre os valores dos coeficientes de determinação ($|R^2_0 - R'^2_0|$); e a diferença entre os valores do r^2 modificado (Δr_m^2). Esta etapa foi realizada no programa *Xternal validation 1.0*. (<http://dtclab.webs.com/software-tools>).

4.2. Triagem virtual baseada por similaridade 2D

Dentre as várias abordagens de CADD, encontram-se aquelas relacionadas à triagem em bibliotecas virtuais de compostos químicos, ou simplesmente triagem virtual [34]. Deste modo, foi realizada uma busca baseada em similaridade estrutural 2D, utilizando como padrão de busca a estrutura molecular mais simples (composto **1**) e a estrutura que representa o composto mais ativo (composto **50**) (Tabela 4.1). Este estudo exploratório foi conduzido utilizando a ferramenta *Swiss Similarity* (<http://www.swissimilarity.ch>), empregando como abordagem de busca o método combinado e utilizando os SMILES como dados de entrada.

4.2.1. Toxicidade in silico e domínio de aplicabilidade

As moléculas identificadas na etapa de triagem virtual tiveram seus descritores moleculares calculados e sua atividade predita utilizando o modelo QSAR obtido. Nesse contexto, foi avaliado se todos os compostos apresentavam atividade predita dentro da faixa linear (pIC_{50} 4,04 – 9,387). Os compostos com atividade fora dessa faixa foram removidos do conjunto de dados, uma vez que suas previsões seriam oriundas de extrapolação.

Devido à grande quantidade de moléculas obtidas, foi necessário adotar abordagens de redução para se chegar a um número gerenciável desses compostos. Para isso, utilizou-se primeiramente a abordagem de redução baseada na toxicologia, por meio dos filtros de carcinogenicidade, mutagenicidade e hepatotoxicidade disponíveis no programa *VEGA QSAR 1.1.4* (<http://www.vega-qsar.eu/download.html>). Todos os compostos que apresentaram efeitos positivos para esses filtros, também foram removidos do conjunto de dados [35].

Na sequência, foi avaliado se os compostos restantes se enquadravam no domínio da aplicabilidade (DA) do modelo. Para isso, utilizou-se a medida da distância euclidiana entre os descritores moleculares obtidos para essas novas moléculas e os descritores do conjunto teste, com o auxílio do programa *Euclidian Applicability Domain 1.0* (download: http://teqip.jdvu.ac.in/QSAR_Tools). O composto foi mantido no estudo se apresentasse distância média normalizada entre 0-1, uma vez a confiabilidade de previsão é aumentada dentro desse intervalo [36].

4.3. Estudo de Docking

4.3.1. Modelagem farmacofórica

A proteína cristalizada utilizada como farmacóforo do estudo foi obtida a partir do banco de dados do PDB (<https://www.rcsb.org/>), sendo intitulada como 1GFW. O arquivo original contendo o complexo da proteína e seu inibidor, passou por um processo de preparação no qual foi realizado desacoplamento do complexo cristalográfico e retirada de estruturas que não eram de interesse. Ao final desse processo, foram gerados arquivos individuais no formato .pdb para cada molécula envolvida, e para isso, utilizou-se o software *Discover Studio 2021 Client*.

4.3.2. Triagem virtual 3D

Os arquivos gerados na etapa anterior serviram como dados de entrada na ferramenta online Pharmit (Pharmit Search Engine - pitt.edu). Com o auxílio dessa ferramenta, definiu-se o farmacóforo com base nas principais interações realizadas da proteína (1GFW) com o seu ligante (Molécula 33).

Este servidor aplica em suas buscas o método definido como Formas Moleculares Alinhadas (VAMS), que procura identificar compostos moleculares correspondentes aos utilizados como base da busca a partir de critérios de inclusão e de exclusão. Os critérios adotados neste trabalho como efeito de inclusão a forma do ligante, e de exclusão, a forma do receptor.

Dentro da ferramenta de busca, utilizou-se bibliotecas oriundas de contribuições, intituladas como Nubbe, Coconut e Life Chemicals. Os filtros utilizados para redução dos compostos são apresentados na (Figura 4.2).

Panel	Filter	Value
A: Hit Reduction	Max Hits per Conf:	1
	Max Hits per Mol:	1
	Max Total Hits:	1
B: Hit Screening (Nubbe, Coconut)	MolWeight	0 ≤ ... ≤ 500
	RotBonds	0 ≤ ... ≤ 7
	LogP	0 ≤ ... ≤ 5
	PSA	0 ≤ ... ≤ 140
	Aromatics	0 ≤ ... ≤ 10
	HBA	0 ≤ ... ≤ 10
	HBD	0 ≤ ... ≤ 5
C: Hit Screening (Coconut)	MolWeight	200 ≤ ... ≤ 400
	RotBonds	0 ≤ ... ≤ 7
	LogP	0 ≤ ... ≤ 5
	PSA	0 ≤ ... ≤ 140
	Aromatics	0 ≤ ... ≤ 10
	HBA	0 ≤ ... ≤ 10
	HBD	0 ≤ ... ≤ 5

Figura 4.2. Configurações dos filtros adotados dentro programa Pharmit. Hit Reduction utilizado as bibliotecas Nubbe, Coconut e Life Chemicals (A), Hit Screening utilizado nas bibliotecas Nubbe e Life Chemicals (B) e, Hit Screening utilizado na biblioteca Coconut (C).

Após a triagem, foi realizado *download* dos resultados de busca referente a cada biblioteca pesquisada. Para que os dados fossem compatíveis com os demais programas que seriam utilizados, foi necessário realizar alteração do formato original (.sdf) para o formato .mol2 e para isso, utilizou-se o *software Open Babel*.

4.3.3. Avaliação da toxicidade in silico

Para as moléculas resultantes da busca utilizando a biblioteca Coconut, realizou-se avaliação da toxicidade *in silico*, utilizando o programa *VEGA QSAR 1.1.4*, com o objetivo de redução do número de moléculas. Durante essa etapa, foram aplicados filtros para verificar possíveis efeitos mutagênicos, carcinogênicos e hepatotóxicos. Todas as moléculas que demonstraram resultados positivos nessas avaliações, foram removidas do conjunto total de dados.

4.3.4. Avaliação farmacocinética

O conjunto de dados final foi avaliado utilizando o modelo denominado *Boiled-Egg*. Nesse modelo, disponível como uma ferramenta na plataforma *SwissADME* (<http://www.swissadme.ch/index.php>), é possível observar a capacidade das moléculas em apresentarem absorção gastrointestinal e permeação pela barreira hematoencefálica. Além disso, é possível visualizar se as moléculas avaliadas possuem características para serem substrato da glicoproteína P.

As moléculas que não demonstraram susceptibilidade *in silico* de absorção gastrointestinal e permeação na barreira hematoencefálica, foram retiradas do conjunto de dados.

4.3.5. Docking

Visando analisar a capacidade do conjunto de dados em conseguir realizar interação com o sítio da caspase 3, foi feito o estudo de ancoramento molecular (*docking*). Para este fim, o software utilizado foi *PyRx/AutoDock Vina*.

Neste programa, a abordagem utilizada para ancoramento é de um sítio fixo com um ligante flexível. Os arquivos utilizados para o estudo foram a proteína 1GFW e seu

inibidor co-cristalizado, previamente preparados conforme descrito no item de modelagem farmacofórica.

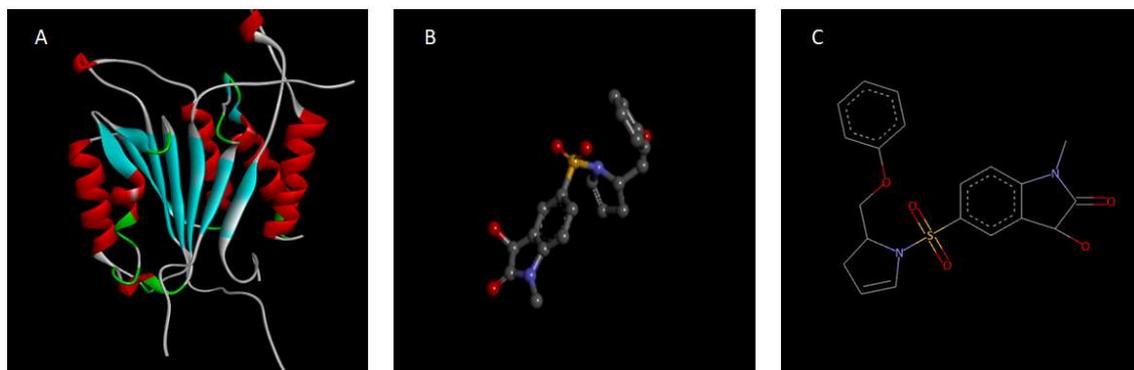


Figura 4.3. Arquivos gerados dentro programa Discovery Studio. Proteína CASPASE 3 (A), estrutura 3D (B) e 2D (C) do inibidor co-cristalizado.

O átomo de enxofre do inibidor foi estabelecido Figura 4.4 como ponto de referência para a criação de um espaço cúbico no qual o algoritmo do programa executaria as buscas necessárias. Como coordenadas foram utilizado os seguintes valores: center_x: 36,91 Å, center_y: 33,02 Å, center_z: 27,98 Å, size_x: 15,10, size_y: 15,01 e size_z: 15,03.

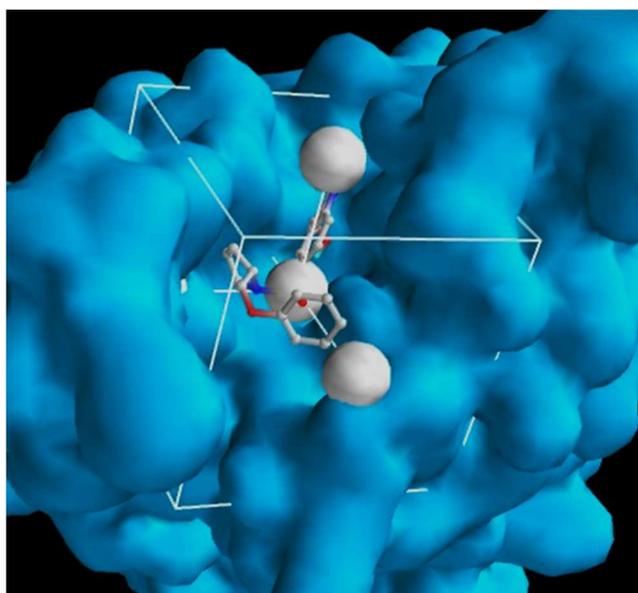


Figura 4.4: Representação gráfica do cubo com as dimensões definidas no *software PyRx/AutoDock Vina*.

O tempo gasto na busca para realização do cálculo de ancoragem (exhaustiveness) foi definido como 100, sendo este eficiente para obtenção de poses similares da estrutura do ligante cristalizado.

Após a validação da metodologia, com os parâmetros mencionados anteriormente, as moléculas do banco de dados foram utilizadas como input no software *PyRx/AutoDock Vina* para o cálculo de ancoragem com o receptor 1GFW. As energias de ligação obtidas no *docking* foram compiladas em um arquivo do Excel e foram selecionadas apenas as moléculas que apresentaram ΔG igual ou superior ao ligante de referência.

4.3.6. Aplicação do modelo QSAR 2D

O modelo de predição da atividade biológica, desenvolvido e validado durante a execução deste trabalho, foi utilizado para previsão de atividade das moléculas cujo ΔG apresentou-se igual ou superior ao ligante de referência.

Todas as moléculas tiveram seus descritores calculados com auxílio do programa Dragon 6.0, e apenas os descritores de interesse foram utilizados para realização do cálculo.

5. RESULTADOS E DISCUSSÃO

5.1. Estudo QSAR

Ao término do processo de seleção de variáveis, o modelo obtido (apresentado na equação 1) foi construído com 6 descritores moleculares utilizando o método de regressão por PLS. Estes descritores originaram 2 variáveis latentes (VLs).

A regressão obtida mostrou-se significativa, uma vez que o valor do teste F (63,60) foi muito maior ao valor tabelado (3,18), para um $p=2$ e $n-p-1=54$, com um nível de significância de 95% ($\alpha=0,05$). A definição dos descritores selecionados se encontram na (Tabela 5.1).

$$pIC_{50} = 1,342 + 0,498(CATS2D_04_AA) + 0,179(CATS2D_06_AL) + 0,015(P_VSA_LogP_4) + 1,825(GATS2i) + 0,452(VE1_A) - 0,39(CATS2D_03_DL) \quad (1)$$

Tabela 5.1: Descritores utilizados para o modelo

Nome	Descrição ¹	Classe
CATS2D_04_AA	Acceptor-Acceptor de lag 4.	CATS 2D
CATS2D_06_AL	Acceptor-Lipofílico de lag 6.	CATS 2D
P_VSA_LogP_4	P_VSA-like no LogP.	P_VSA-like
GATS2i	Autocorrelação Geary de lag 2 ponderado pelo potencial de ionização.	Autocorrelação 2D
VE1_A	Soma do coeficiente do último autovetor da matriz de adjacência.	Descritores baseados em matriz 2D
CATS2D_03_DL	Doador-Lipofílico de lag 3.	CATS 2D

¹ – Dados retirados da lista de descritores calculados pelo programa *Dragon 6.0*

A Figura 5.1 ilustra a contribuição individual dos descritores moleculares com a atividade biológica. Observa-se que o descritor CATS2D_04_AA apresenta a contribuição positiva de maior relevância. Por outro lado, o descritor VE1_A é identificado como o de menor relevância para a atividade biológica.

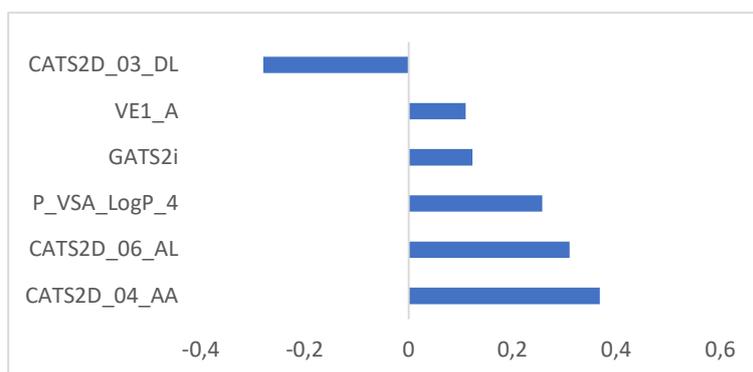


Figura 5.1: Contribuição dos descritores moleculares para o vetor de regressão.

A escolha do número ideal de variáveis latentes para a construção do modelo levou em consideração a quantidade de informação acumulada e os valores de RMSECV. Como demonstrado na Tabela 5.2, duas variáveis latentes acumulam aproximadamente 50% da variância dos dados e possuem o menor erro associado.

O modelo escolhido possui uma significativa capacidade de explicar os dados, uma vez que ele consegue explicar 70,2% da variabilidade presente nos dados da atividade biológica ($R^2 = 0,702 \geq 0,6$), fornecendo uma boa compreensão da relação entre os descritores moleculares selecionados com os dados da atividade. Com o coeficiente de determinação da validação cruzada de 0,635 ($Q^2_{LOO} \geq 0,5$), o modelo é capaz de prever 63,5% da variabilidade observada nos valores de atividade biológica, sendo um importante indicador da capacidade de previsão [37].

A pequena diferença de apenas 0,067 entre os valores de R^2 e Q^2_{LOO} indica que o modelo possui uma baixa probabilidade de se ajustar em excesso aos dados do conjunto de treinamento. Isso significa que o modelo não está apenas se ajustando bem aos dados, mas também tem uma boa capacidade de prever com precisão e confiabilidade novos conjuntos de dados [37].

Tabela 5.2: Parâmetros estatísticos obtidos para o conjunto de treinamento após PLS.

VL	% Inf. Acumuladas	RMSECV	RMSEC	R^2	Q^2_{LOO}
1	32,303	0,807	0,754	0,678	0,631
2	47,274	0,803	0,724	0,702	0,635
3	70,583	0,805	0,720	0,706	0,633
4	85,352	0,804	0,719	0,707	0,633
5	91,484	0,805	0,719	0,707	0,633
6	100	0,805	0,719	0,707	0,633

A presença de correlações ao acaso entre os descritores moleculares e a atividade biológica foi realizada utilizando a metodologia *y-randomization*, com 40 aleatorizações. Os resultados foram tabelados e podem ser visualizados na (Tabela 5.3).

Tabela 5.3: Resultados gerados utilizando a metodologia *y-randomization*.

Aleatorização	R ²	Q ²	R (yal,y)	Aleatorização	R ²	Q ²	R (yal,y)
1	0,051	-0,198	0,204	21	0,028	-0,336	0,041
2	0,147	-0,096	0,206	22	0,107	-0,121	0,239
3	0,192	-0,086	0,017	23	0,157	-0,075	0,314
4	0,097	-0,190	0,138	24	0,120	-0,094	0,095
5	0,046	-0,404	0,134	25	0,126	-0,111	0,095
6	0,117	-0,227	0,048	26	0,040	-0,206	0,134
7	0,111	-0,111	0,012	27	0,092	-0,135	0,075
8	0,072	-0,184	0,073	28	0,077	-0,203	0,151
9	0,096	-0,199	0,113	29	0,138	-0,077	0,008
10	0,082	-0,165	0,059	30	0,058	-0,211	0,014
11	0,074	-0,286	0,083	31	0,063	-0,217	0,131
12	0,051	-0,208	0,149	32	0,138	-0,157	0,136
13	0,109	-0,159	0,141	33	0,127	-0,148	0,214
14	0,159	-0,016	0,108	34	0,076	-0,206	0,008
15	0,108	-0,163	0,114	35	0,169	-0,051	0,045
16	0,106	-0,142	0,000	36	0,130	-0,155	0,010
17	0,063	-0,252	0,157	37	0,084	-0,169	0,269
18	0,112	-0,099	0,017	38	0,102	-0,208	0,074
19	0,021	-0,226	0,065	39	0,107	-0,091	0,232
20	0,164	-0,104	0,223	40	0,211	0,024	0,011
Modelo sem aleatorização: R ² = 0,702 , Q ² = 0,635 , R (yal, y) = 1							

Os valores obtidos dos interceptos na metodologia *y-randomization* foram de 0,06 para o R² e -0,21 para o Q²_{LOO}, os quais estão de acordo com o recomendado pela literatura ($\alpha_Q < 0,05$; $\alpha_R < 0,3$) [37], podendo ser visualizados na (Figura .

Realizando a permuta aleatória da variável resposta (*y* = atividade biológica) e mantendo as variáveis explicativas inalteradas (*x* = descritores moleculares) quebra-se a relação real que existe entre elas e espera-se que os resultados encontrados sejam inferiores aos obtidos no modelo original. Analisando os dados encontrados, verifica-se que as informações explicadas e preditas pelo modelo são reais e não são resultantes de correlações espúrias devido a diferença nos resultados obtidos de R² e Q²_{LOO}.

Em relação à robustez do modelo, um outro parâmetro que foi considerado na validação interna é a metodologia *Leave-N-Out* (LNO). O valor de N escolhido foi 14, correspondendo a aproximadamente 25% do conjunto de treinamento.

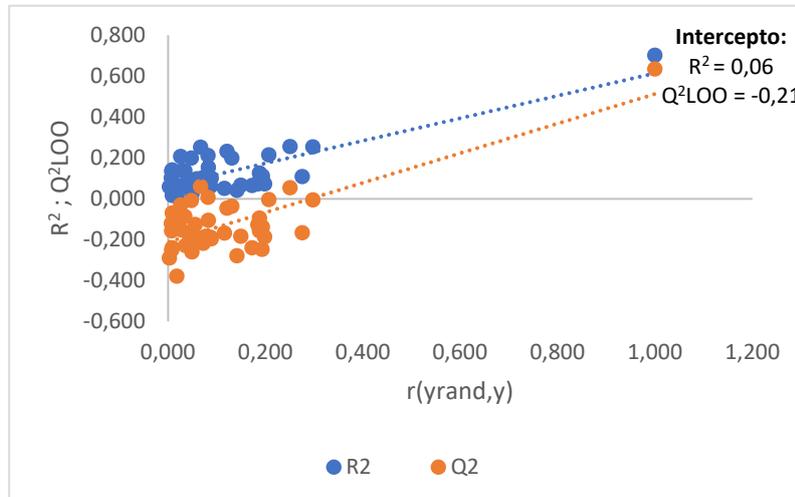


Figura 5.2: Representação gráfica da sobreposição dos dados de R^2 e Q^2_{LOO} e seus interceptos utilizando a metodologia *y-randomization*

Os resultados dos testes revelaram que o modelo apresenta uma sólida robustez, uma vez que pequenas flutuações em torno da média do valor de Q^2_{LOO} foram observados em sua sextuplicata (FIG. 5.3 e TAB. 5.4). As mínimas flutuações obtidas indicam a capacidade do modelo em lidar com pequenas e intencionais variações dos dados, podendo-se confiar na sua capacidade de previsões para além do conjunto de treinamento utilizado.

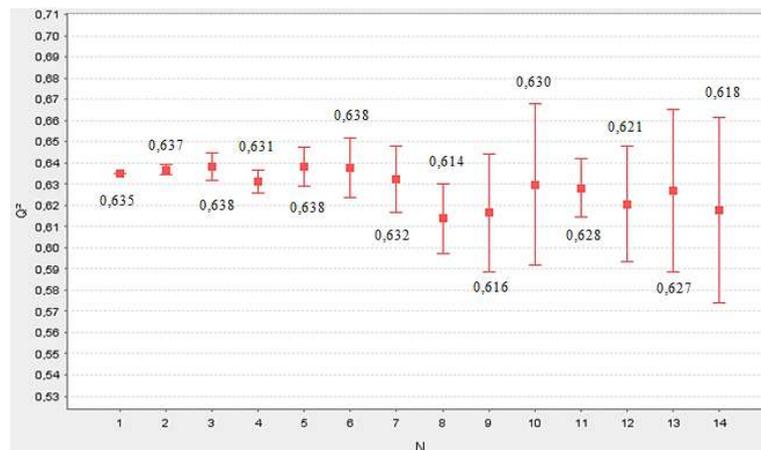


Figura 5.3: Representação gráfica do resultado *Leave-14-Out*.

Tabela 5.4: Resultados gerados utilizando a metodologia *Leave-N-out*.

Identificação/ Réplicas	1	2	3	4	5	6	Média	Desvio Padrão
<i>Leave-1-out</i>	0,635	0,635	0,635	0,635	0,635	0,635	0,635	0,000
<i>Leave-2-out</i>	0,636	0,638	0,639	0,632	0,637	0,640	0,637	0,003
<i>Leave-3-out</i>	0,642	0,635	0,648	0,629	0,635	0,640	0,638	0,007
<i>Leave-4-out</i>	0,623	0,637	0,633	0,633	0,634	0,628	0,631	0,005
<i>Leave-5-out</i>	0,636	0,651	0,634	0,627	0,647	0,635	0,638	0,009
<i>Leave-6-out</i>	0,634	0,657	0,621	0,646	0,624	0,646	0,638	0,014
<i>Leave-7-out</i>	0,625	0,616	0,654	0,635	0,617	0,647	0,632	0,016
<i>Leave-8-out</i>	0,587	0,612	0,617	0,630	0,631	0,605	0,614	0,017
<i>Leave-9-out</i>	0,654	0,588	0,624	0,586	0,641	0,605	0,616	0,028
<i>Leave-10-out</i>	0,626	0,647	0,637	0,622	0,681	0,566	0,630	0,038
<i>Leave-11-out</i>	0,606	0,635	0,620	0,640	0,642	0,625	0,628	0,014
<i>Leave-12-out</i>	0,647	0,623	0,569	0,631	0,636	0,617	0,621	0,027
<i>Leave-13-out</i>	0,663	0,563	0,617	0,609	0,655	0,654	0,627	0,038
<i>Leave-14-out</i>	0,545	0,641	0,643	0,668	0,593	0,618	0,618	0,044

A avaliação de *outliers* e pontos influentes foi realizada por meio do gráfico de resíduos studentizados (σ) pelo *leverage* (FIG. 5.4), gerado no programa *QSAR modeling*. No gráfico, se observa a presença de um *outlier* ($\sigma=2,8$), representado pelo composto **23**, e dos pontos influentes com os seguintes valores: 0,111, 0,133 e 0,145, correspondentes aos compostos **1**, **14** e **29**, respectivamente.

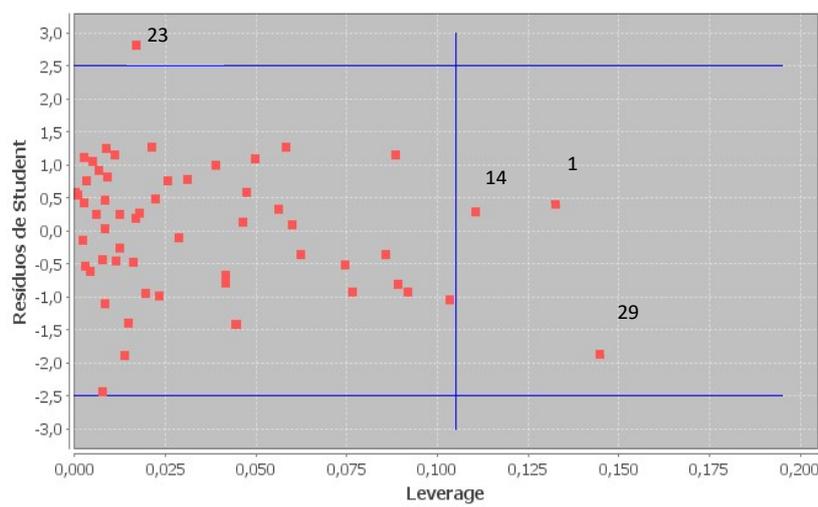


Figura 5.4: Gráfico de *leverage* x resíduos de Student para avaliação da presença de *outliers*.

Considerando que o modelo QSAR é um modelo de previsão, optou-se por manter essas amostras no conjunto de dados, visando a manutenção da quantidade amostral. Além disso, outro fator que corrobora para manutenção dos pontos, tanto os influentes

quanto o outlier, é o fato do modelo atender a todos os critérios estatísticos e apresentar adequada capacidade preditiva.

Um fator preocupante seria o fato dos pontos influentes fossem também considerados outliers, pois isso indicaria que o coeficiente de determinação estaria sendo influenciado por um ponto cujo resíduo é alto, o que não é observado nesta análise.

Na etapa de validação externa, o modelo teve sua capacidade preditiva avaliada por meio do conjunto teste (N= 14 compostos). Nessa fase, foram calculados e avaliados alguns parâmetros estatísticos para determinar a qualidade de previsão do modelo criado. Os dados podem ser visualizados na (TAB. 5.5).

Tabela 5.5: Resultados da validação externa.

Compostos	pIC_{50} real	pIC_{50} predito
2	4,373	4,397
7	4,967	5,584
11	7,119	6,507
16	7,066	6,551
21	6,686	6,995
30	7,309	6,584
39	7,245	7,252
41	6,106	7,242
45	6,300	6,612
49	8,836	7,999
56	8,747	8,916
58	8,611	8,644
62	7,824	7,369
68	7,556	8,363
Parâmetro	Resultado	Valor de referência
R^2_{pred}	0,796	> 0,6
k	0,997	0,85 < k < 1,15
k'	0,997	0,85 < k' < 1,15
$R^2_0 - R'^2_0$	0,039	< 0,3
$\Delta r_m^2(\text{pred})\text{-scaled}$	0,135	< 0,2
Average r_m^2 (pred)-scaled	0,715	> 0,5

Com base nos dados de atividade obtidos através do modelo proposto, pode-se inferir que o modelo possui uma boa capacidade preditiva. Uma evidência disso é a

pequena diferença observada entre os valores reais e os valores preditos, com exceção do composto **41**, cuja diferença foi de 1,14. Embora essa diferença seja alta, ainda está muito próxima de 1, o que confirma a acurácia do modelo.

Os dados estatísticos gerados pelo programa *Xternal validation 1.0* reforçam a confiabilidade do modelo, uma vez que os resultados obtidos se apresentaram dentro dos critérios utilizados como referência [38-39].

5.2. Triagem virtual 2D

Na busca por novos fármacos, a estrutura 2D dos compostos 1 e 50 foram utilizados como base. Ao final da busca, um total de 1.600 moléculas foram obtidos por meio da plataforma *SwissSimilarity*. Esses compostos tiveram seus descritores calculados e sua atividade predita conforme o modelo disposto na equação 1.

Para garantir a significância da atividade predita do conjunto amostral, foram excluídos os compostos cujas atividades ficaram fora da faixa linear (pIC_{50} 4,04 – 9,387). Isso resultou em um conjunto final de 1516 moléculas para a análise toxicológica *in silico*.

Dentro desse conjunto, foram realizadas avaliações de carcinogenicidade, mutagenicidade e hepatotoxicidade. Os compostos que apresentaram perfis indesejáveis nessas análises foram excluídos, resultado em um conjunto final de 20 compostos promissores (*hits*), conforme apresentado na (TAB. 5.6).

O domínio da aplicabilidade, medido pela distância média normalizada (DMN), é usado para avaliar o grau de similaridade estrutural entre as amostras resultantes da triagem virtual e as amostras utilizadas para construir o modelo proposto. Os descritores moleculares calculados foram utilizados para obtenção dos resultados apresentados na (TAB. 5.6).

Durante a avaliação, foi observado que o composto nomeado como “r” apresentou uma DMN de 1,017, o que está acima do limite estabelecido de 1. Isso indica que o modelo tem confiabilidade de previsão para os 19 compostos restantes, visto que as previsões realizadas estão dentro da faixa de interpolação dos dados.

Tabela 5.6: Novos compostos *hits* obtidos pelo estudo de triagem virtual.

ID	SMILES	pIC_{50} predito	DMN ¹
a	<chem>O=C(Nc3c(c1nc(ccc1)C2CC2)c(ns3)c4ccc5nn(cc5(c4))C)C6CC6(C)</chem>	8,363	0,109

b	<chem>O=C3N(N=C(c2ccc(OC)c(OCc1ncccc1)c2)C3(C)C)C(C)C</chem>	7,382	0,004
c	<chem>n1cnn2cc(ccc12)c4c(nc(NCc3ccccc3C)s4)c5nc(ccc5)C</chem>	6,868	0,345
d	<chem>O=C(N)c3enc1ccc(cc1c3(Nc2ccccc(OC)c2))S(=O)(=O)C4CCCC4</chem>	8,253	0,427
e	<chem>n4nc3n(N=C(c2ccc(OC)c(OC1CCCC1)c2)CS3)c4c5cc(OC)ccc5(OC)</chem>	8,912	0,271
f	<chem>O=C(N)c4c(OC(c1cccc1Cl)C)cc(n3cnc2ccc(OC)cc23)s4</chem>	8,529	0,018
g	<chem>O=C(Nc2c(c1nc(ccc1)C(C)C)c(ns2)c3ccc4nn(cc4(c3))C)C5CC5(C)</chem>	8,299	0,109
h	<chem>n3nc2n(N=C(c1ccc(OC)c(OC)c1)CS2)c3c4cc(OC)ccc4(OC)</chem>	8,254	0,262
i	<chem>n4nc3n(N=C(c2ccc(OC)c(OCC1CC1)c2)CS3)c4c5cc(OC)ccc5(OC)</chem>	8,848	0,271
j	<chem>O=C(N)c1cccn1CC</chem>	5,832	0,037
k	<chem>O=C(NC)c1cnccc1</chem>	5,534	0,056
l	<chem>O=C(Nc1cnccc1)Nc2ccc(F)cc2</chem>	5,337	0,038
m	<chem>n3c(nc(NCC1CCCCC1)c4NC(c2ccc(cc2)Cl)CNc34)N</chem>	6,004	0,191
n	<chem>O=C(NCC#C)c1ccc[nH]1</chem>	4,673	0,398
o	<chem>O=C(N)C(N1C(=O)CCC1)CC</chem>	5,734	0,036
p	<chem>O=C(OCC=C)c1cccc1(N)</chem>	4,543	0,357
q	<chem>O=C(OC1(C(=O)C)(CCCC1))C</chem>	6,250	0,026
r	<chem>N1C(C)SC(C)SC1CC(C)C</chem>	4,643	1,017
s	<chem>O=C(OCC)CC(C(=O)OCC)C</chem>	5,891	0,037
t	<chem>O=C(Nc1ncc(Cl)s1)C(c2ccc(cc2)Cl)C(C)(C)C</chem>	6,544	0,030

l-Distância média normalizada

Além disso, é importante destacar que a maioria dos 19 compostos apresentam altos valores de pIC_{50} , o que os torna candidatos promissores para investigações futuras.

5.3. Triagem virtual 3D

Dando continuidade à busca por novos compostos com potencial de inibição da proteína alvo caspase 3, os estudos de triagem virtual foram ampliados para levar em consideração a estrutura tridimensional da proteína em estudo. Esse tipo de triagem permite uma abordagem mais precisa na identificação de possíveis compostos que possam interagir com o sítio ativo da proteína, levando a efeito desejáveis.

O completo cristalográfico 1GFW foi selecionado para esse estudo devido a sua interação com um inibidor pertencente a classe isatina-sulfonamina, por ser derivado do organismo *Homo sapiens* e por possuir uma resolução adequada 2.80 Å, o que permite uma visualização clara e detalhada da estrutura.

O complexo obtido no *Protein Data Bank* passou por algumas alterações a fim de remover estruturas irrelevantes para o estudo, como moléculas de água e co-fatores. O resultado desse processo resultou em dois arquivos independentes, um contendo apenas a estrutura da proteína e outro contendo a estrutura do inibidor em questão.

O ligante selecionado para o estudo de triagem foi a molécula de número 33 (TAB.4.1). Essa escolha foi baseada em seu alto valor de pIC_{50} e na sua estrutura relativamente mais simples em comparação com as moléculas 50 e 60, que também apresentam valores elevados de pIC_{50} . Essa decisão teve como objetivo aumentar as chances de encontrar mais compostos durante a busca.

Os arquivos da proteína cristalografada e o da molécula **33** foram utilizados como dados de input no servidor *Pharmit* e nele, foram identificados pontos de interação do ligante com a proteína. Para identificação dos farmacóforos essenciais ao desempenho da atividade de interesse, utilizou-se o artigo de referência da proteína 1GFW e da síntese da molécula mencionada [28,40].

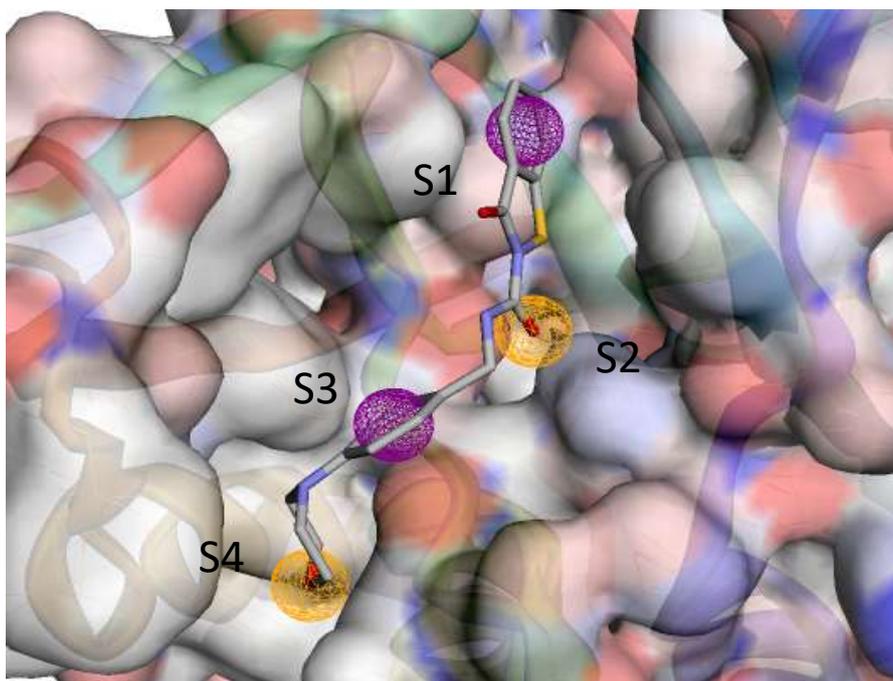


Figura 5.5. Representação dos sítios farmacofóricos selecionados no servidor *Pharmit*.

Foram selecionados um total de 4 pontos farmacofóricos, como pode ser visualizado na (FIG 5.5). Cada um desses pontos selecionados abrangem as 4 cavidades essenciais para a atividade inibitória da proteína (S1, S2, S3 e S4) [28]. A seleção

compreende 2 grupamentos aromáticos, representados pela esfera roxa e, dois grupamentos aceptores de hidrogênio, representados pela esfera amarela.

No estudo de triagem virtual tridimensional, foram identificados um total de 819 compostos, dos quais 31 foram obtidos na biblioteca Life Chemicals, 2 na biblioteca Nubbe e 786 na biblioteca Coconut. Devido ao grande número de compostos na biblioteca Coconut, foi necessário aplicar filtros para obter uma quantidade gerenciável de moléculas para análises posteriores.

Utilizando o programa *VEGA QSAR 1.1.4*, foram selecionados filtros específicos para reduzir o número de compostos, com foco em critérios relacionados à mutagenicidade, carcinogenicidade e hepatotoxicidade. Após a aplicação desses filtros, o número de compostos foi reduzido para um total de 9 moléculas.

É importante mencionar que, devido ao baixo número de compostos obtidos nas bibliotecas Life Chemicals e Nubbe, optou-se por não avaliá-los pelos filtros de toxicidade. É válido ressaltar que o objetivo deste trabalho não é selecionar compostos com características toxicológicas ideais, mas sim usar as moléculas triadas como um guia para futuras otimizações.

O pool final composto por 42 moléculas passou por uma avaliação de suas características farmacocinéticas utilizando a modelo *Boiled-Egg* disponível no servidor online *SwissADME*.

O modelo *boiled-egg* fornece informações farmacocinéticas relevantes, como a previsão da capacidade das moléculas de atravessar a barreira hematoencefálica (representada pela parte mais escura do ovo) e a previsão da facilidade de absorção pelo trato gastrointestinal (parte mais clara do ovo). Essas características são altamente desejáveis no desenvolvimento de fármacos e, portanto, todas as moléculas que apresentaram essas características foram mantidas no estudo. Apenas duas moléculas ficaram fora do intervalo esperado e são representadas por um círculo tracejado na (FIG. 5.6).

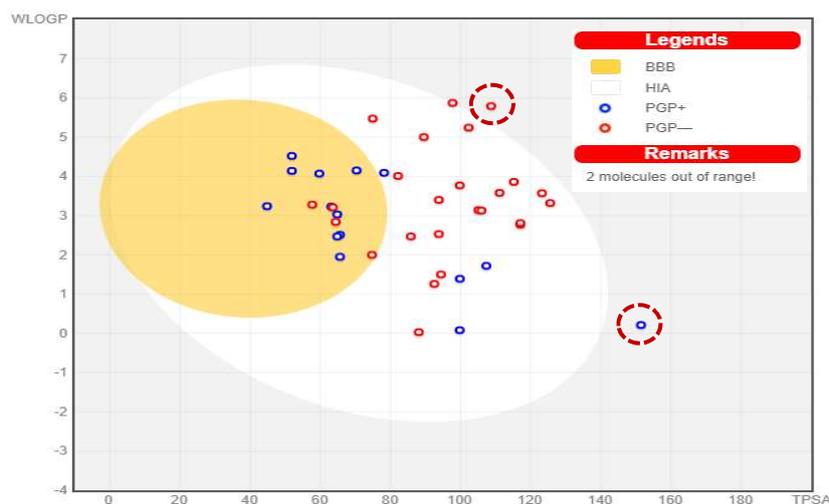


Figura 5.6: Disposição das moléculas no modelo *boiled-egg*.

Os pontos em azul no gráfico representam as moléculas que foram previstas como potenciais substratos da glicoproteína P (PGP+), enquanto os pontos em vermelho indicam as moléculas que foram previstas como não substratos (PGP-). Levando em consideração que o alvo biológico atua no sistema nervoso central, os inibidores eficazes da caspase 3 seriam aqueles que apresentam boa permeação através da barreira hematoencefálica e baixa afinidade com a proteína PGP. Isso permitiria um acúmulo adequado no local de ação, potencializando a resposta terapêutica. No entanto, o número de moléculas para as análises subsequentes ficaria consideravelmente reduzido.

Dessa forma, para à análise de ancoramento molecular, foram baixadas as 40 estruturas no formato tridimensional.

5.4. Estudo de Docking

A fim de investigar o modo de interação entre os potenciais inibidores da caspase 3 e o receptor, foram realizados cálculos de ancoramento molecular.

Para obtenção de resultados confiáveis, foi necessário à validação de uma metodologia de *redocking* do ligante co-cristalizado com a proteína 1GFW. Essa metodologia tem como objetivo inserir o ligante novamente na cavidade da proteína e posicioná-lo de forma que se tenha uma boa sobreposição com sua conformação nativa.

Na etapa de validação da metodologia as coordenadas utilizadas para o ancoramento molecular foram de center_x: 36,91 Å, center_y: 33,02 Å, center_z: 27,98 Å, size_x: 15,10, size_y:15,01 e size_z: 15,03 utilizando o exhaustiveness de 100. A

energia de ligação obtida (ΔG) foi de $-8,7 = \text{Kcal.mol}^{-1}$ e este resultado foi utilizado para fins de comparação.

A FIG. 5.7 **Figura** revela que a melhor posição obtida é semelhante à pose apresentada na estrutura co-cristalizada, o que indica que o método é capaz de reproduzir de maneira confiável a interação dos ligantes em teste com a proteína 1GFW.

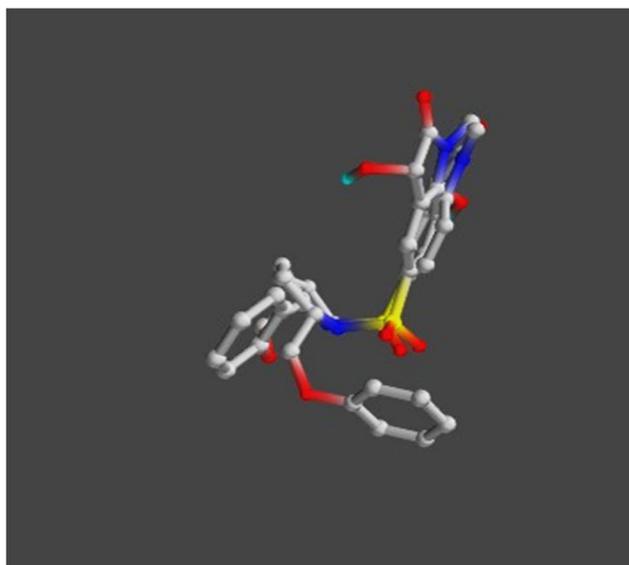


Figura 5.7: Sobreposição do ligante na sua forma nativa e da best pose obtida no *redocking*.

Os 40 compostos foram submetidos ao estudo de ancoramento molecular, e apenas 4 moléculas apresentaram uma energia de ligação igual ou superior a $-8,7 \text{ Kcal.mol}^{-1}$. Essas moléculas apresentaram resultados promissores em termos de energia, indicando uma alta probabilidade de se ligarem ao receptor de interesse e desempenharem a atividade biológica desejada. Os resultados estão apresentados na tabela a seguir.

Os quatro compostos selecionados foram avaliados quanto às suas atividades biológicas utilizando o modelo descrito na equação 1. Abaixo estão apresentados os resultados de predição dessas atividades, juntamente com seus respectivos códigos *smiles*.

Observa-se que todas as atividades preditas estão dentro da faixa linear do modelo (4,04 a 9,39) e a distância média normalizada apresentou valores inferiores a 1, o que indica confiabilidade dos dados obtidos.

Tabela 5.7: Dados de energia de ligação obtidos no estudo de *docking*

Composto	Energia (ΔG) Kcal.mol^{-1}	Composto	Energia (ΔG) Kcal.mol^{-1}
A	-9.2	U	-7.7

B	-9.2	V	-7.6
C	-9.1	W	-7.6
D	-8.8	X	-7.6
E	-8.6	Y	-7.5
F	-8.5	Z	-7.4
G	-8.5	A'	-7.4
H	-8.4	B'	-7.4
I	-8.4	C'	-7.4
J	-8.2	D'	-7.3
K	-8.2	E'	-7.2
L	-8.0	F'	-7.2
M	-8.0	G'	-7.2
N	-7.9	H'	-7.2
O	-7.8	I'	-7.1
P	-7.7	J'	-7.1
Q	-7.7	K'	-6.9
R	-7.7	L'	-6.8
S	-7.7	M'	-6.8
T	-7.7	N'	-4.3

Tabela 5.8: Código *smiles* e atividade prevista dos compostos com os melhores resultados de energia de ligação.

ID	Smiles	Atividade (<i>pIC</i>₅₀)	DMN¹
A	<chem>O=c1c(c2c(OC)cccc2)coc2c1ccc1OCN(Cc21)[C@@H](c1cccc1)C</chem>	8,600	0,399
B	<chem>O=C1c2c(C(=O)c3c1oc1c3ccc3OCN(Cc13)CCc1cccc1)cccc2</chem>	8,425	0,003
C	<chem>S(=O)(=O)(Oc1cc2c(oc3c2cccc3)cc1)c1c2c3c(cc1)NC(=O)c3ccc2</chem>	6,228	0,064
D	<chem>N#Cc2nc1cccc1n2CC(=O)NCc3ccc4NC(=O)CCc4(c3)</chem>	6,316	0,311

¹-Distância média normalizada

Na FIG. 5.8 a 5.11 são apresentadas as principais interações intermoleculares entre os compostos selecionados no estudo de ancoramento molecular e os resíduos de aminoácidos que revestem a cavidade interna do sítio ativo da proteína 1GFW.

Dados da literatura demonstram que a caspase 3 possui um sítio ativo com quatro cavidades distintas, sendo denominadas de S1, S2, S3 e S4. Ao lado direito das cavidades é visível a predominância de resíduos com propriedades hidrofílicas, ao ponto que ao lado esquerdo é visto resíduos com propriedades hidrofóbicas.

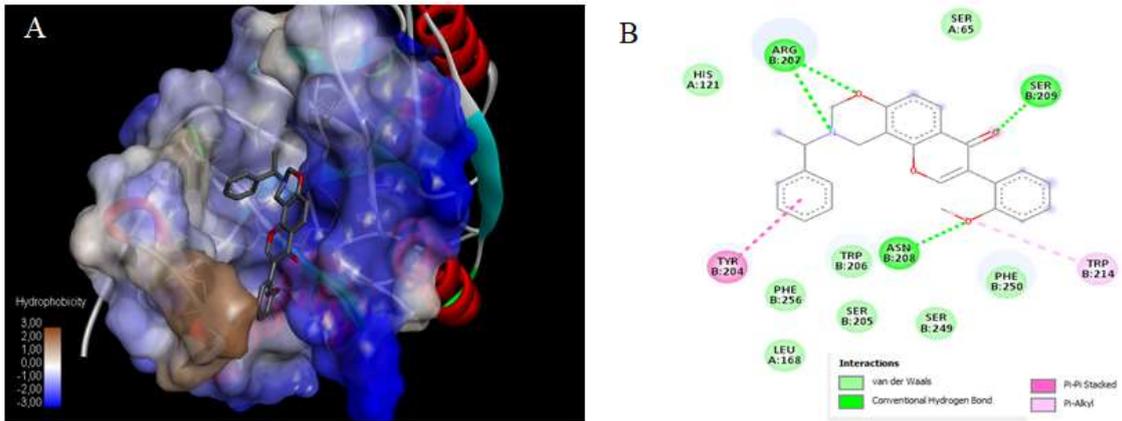


Figura 5.8: A) Representação do ancoramento da molécula A com o proteína 1GFW (superfície colorida pelo grau de hidrofobicidade); B) Diagrama 2D exibindo os principais tipos de interação entre o ligante e os resíduos de aminoácido da proteína.

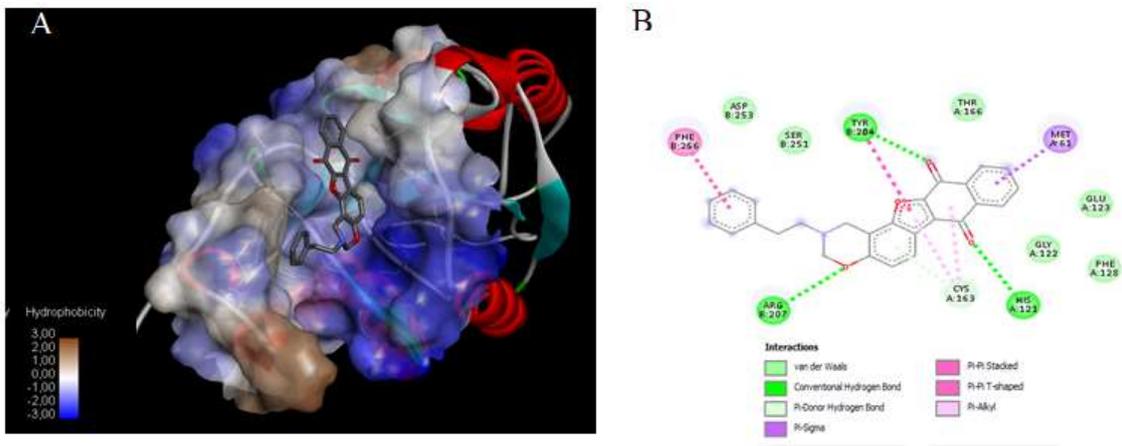


Figura 5.9: A) Representação do ancoramento da molécula B com o proteína 1GFW (superfície colorida pelo grau de hidrofobicidade); B) Diagrama 2D exibindo os principais tipos de interação entre o ligante e os resíduos de aminoácido da proteína.

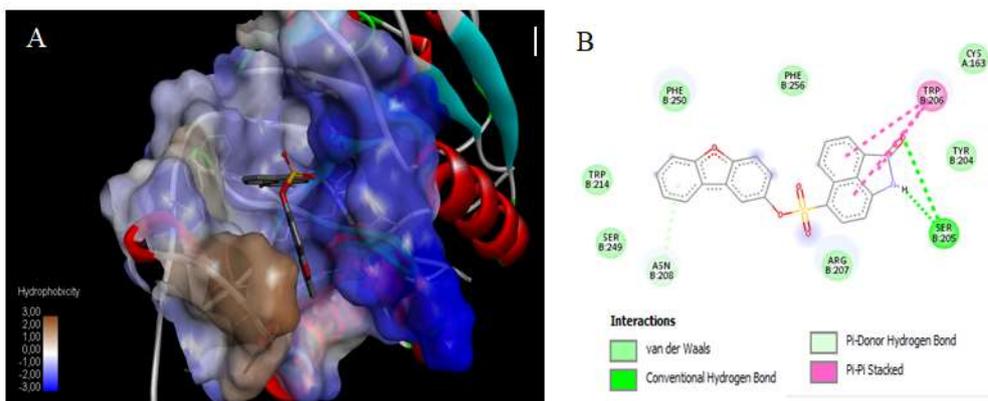


Figura 5.10: A) Representação do ancoramento da molécula C com o proteína 1GFW (superfície colorida pelo grau de hidrofobicidade); B) Diagrama 2D exibindo os principais tipos de interação entre o ligante e os resíduos de aminoácido da proteína.

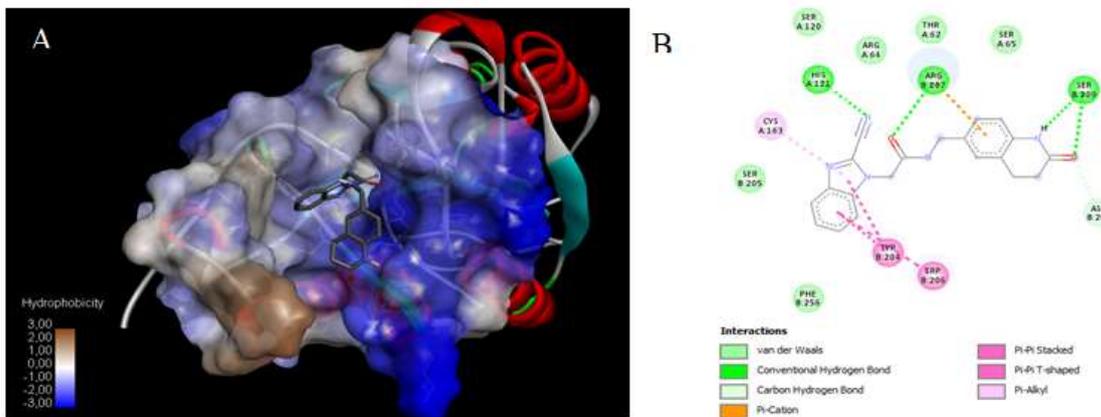


Figura 5.11: A) Representação do ancoramento da molécula D com o proteína 1GFW (superfície colorida pelo grau de hidrofobicidade); B) Diagrama 2D exibindo os principais tipos de interação entre o ligante e os resíduos de aminoácido da proteína.

O composto A apresentou o maior valor de atividade em comparação às outras três moléculas avaliadas. Esse resultado pode ser atribuído às fortes interações estabelecidas pelo composto com os resíduos presentes nas quatro cavidades da proteína. Essas interações são particularmente relevantes, considerando que as cavidades S2 e S3 estão relacionadas à inibição e seletividade da enzima, respectivamente. A extensão da cadeia até a cavidade S4 contribui com o aumento da atividade de inibição das moléculas [29,42].

Na FIG 5.8 é possível observar que o composto apresenta duas ligações de hidrogênio com Arg207, principal resíduo da cavidade S2. O composto também apresenta ligações do mesmo tipo com os resíduos de Ser209 (S3) e Asn208 (S4). Além disso, é observado ligações do tipo pi-pi stacked com o resíduo Tyr204 (S1) e pi-akyl com o resíduo Trp214 (S4).

Apesar dos outros três compostos não estabelecerem interações específicas com os diferentes resíduos das 4 cavidades da proteína, eles exibem boa energia de ligação e valores elevados de atividade.

6. CONCLUSÃO

Com base em todos os resultados obtidos, podemos concluir que o modelo QSAR 2D demonstrou ser significativo, robusto, livre de correlações ao acaso e com boa capacidade preditiva. O estudo de triagem virtual por similaridade 2D permitiu a identificação de 19 compostos hits com baixo risco de carcinogenicidade, mutagenicidade e hepatotoxicidade, o que indica sua relevância para futuros estudos como potenciais candidatos a fármacos.

Além disso, o segundo estudo de triagem virtual, baseado na similaridade 3D, possibilitou a identificação de moléculas com padrões farmacofóricos semelhantes as moléculas que serviram como base para esse estudo. Esses compostos passaram por filtros de toxicidade e propriedades farmacocinéticas *in silico*, garantindo maior segurança no processo de seleção de candidatos a fármacos.

O estudo de ancoramento molecular foi realizado para observar as interações entre esses compostos e a estrutura cristalográfica da proteína caspase 3. Nesse estudo, quatro compostos mostraram perfis promissores como inibidores da caspase, e o composto A se destacou por apresentar interações específicas com os resíduos das diferentes cavidades da proteína.

Esses resultados indicam que a abordagem de triagem virtual combinada com a modelagem molecular foi eficaz na identificação de compostos potenciais para atuarem como inibidores da caspase 3. Essas descobertas podem impulsionar futuras pesquisas e estudos para o desenvolvimento de novos fármacos que visem combater doenças relacionadas a essa proteína, como a doença de Alzheimer e outras enfermidades associadas a processos apoptóticos. No entanto, é importante ressaltar que mais experimentos e ensaios biológicos são necessários para validar esses candidatos em potencial antes que eles possam ser efetivamente utilizados na prática clínica.

7. REFERÊNCIAS

- [1] Yin L, Zhou Y, Liu H, Li Y. LYECs: lysosome-enhancing compounds as potential therapeutic approaches for Alzheimer disease. *Autophagy*. 2022; 6:1-2.
- [2] Da Silva Correia MG, Paixão AO, Jesus AVF et al. Doença de Parkinson: uma desordem neurodegenerativa. *Caderno de Graduação-Ciências Biológicas e da Saúde-UNIT-SERGIPE*. 2013; 1: 57-65.
- [3] Sereniki A, Vital MABF. A doença de Alzheimer: aspectos fisiopatológicos e farmacológicos. *Revista de Psiquiatria do Rio Grande do Sul*. 2008; 30.
- [4] Smith MAC. A doença de Alzheimer. *Revista Brasileira de Neurologia e Psiquiatria*. 1999; 20; 3.
- [5] Calabrò M, Rinaldi C, Santoro G, Crisafulli C. The biological pathways of Alzheimer disease: a review. *AIMS Neuroscience*. 2021; 8: 86–132.
- [6] Forlenza OV. Tratamento farmacológico da doença de Alzheimer. *Archives of Clinical Psychiatry*. 2005; 32:137–148.
- [7] Alzheimer's Association. Demência do Alzheimer no Brasil. Disponível em: <https://www.alz.org/br/demencia-alzheimer-brasil.asp>. Acesso em: 11/07/2023.
- [8] Ribeiro HF, Santos JSF, Souza JN. Doença de Alzheimer de início precoce (DAIP): características neuropatológicas e variantes genéticas associadas. *Revista de Neuro-Psiquiatria*. 2021; 84:113-127.
- [9] Pinto SR, Neto EMR. Doença de alzheimer: abordagem farmacoterapêutica. *Mostra Científica da Farmácia*. 2016; 2:1.
- [10] Medeiros NVP. Estudos *in silico* de inibidores alostéricos do domínio amino terminal de receptores NMDA que contenham a subunidade GluN2B [dissertação]. Natal-RN. Universidade Federal do Rio Grande do Norte, Programa de pós-graduação em ciências farmacêuticas; 2022.
- [11] Silva YJ, Lessa RT, Araujo GN. Advances in early diagnosis of Alzheimer's disease and new treatment perspectives: a systematic review of the literature. *Brazilian Journal of Health Review*. 2021; 4:10121–10135.
- [12] Carraro MF. Estudos da seletividade e afinidade da interação de anticorpos de cadeia única NUsc a oligômeros do peptídeo β^2 -amiloide. [Tese]. São Paulo- SP de Doutorado. Universidade de São Paulo, Departamento de Bioquímica e Imunologia; 2021.
- [13] Porter AG, Jänicke RU. Emerging roles of caspase-3 in apoptosis. *Cell Death and Differentiation*. 1999; 6: 99–104.
- [14] Troy CM, Akpan N, Jean YY. Regulation of caspases in the nervous system: Implications for functions in health and disease. *Progress in Molecular Biology and Translational Science*. 2011; 99.

- [15] Monteiro PAM. Estratégias Terapêuticas Baseadas na Modulação da Atividade Enzimática das Caspases. [Dissertação]. Porto: Universidade Fernando Pessoa, Faculdade Ciências da Saúde; 2014.
- [16] Su JH, Zhao M, Anderson AJ, Srinivasan A, Cotman CW. Activated caspase-3 expression in Alzheimer's and aged control brain: Correlation with Alzheimer pathology. *Brain Research*. 2001; 898:350–357.
- [17] Akkari ACS, Munhoz IP, Tomioka J, Santos NMBF, Santos RF. Inovação tecnológica na indústria farmacêutica: diferenças entre a Europa, os EUA e os países farmaemergentes. *Gest. Prod.* 2016; 23: 365-380.
- [18] Ribeiro FF. Uso de ferramentas do planejamento racional de fármacos auxiliado por computador aplicado a uma série de tiossemicarbazonas e seus bioisómeros frente à doença de chagas. [Tese]. Recife-PE: Universidade Federal de Pernambuco, Centro de Biociências, Programa de Pós-graduação em inovação terapêutica; 2018.
- [19] Calixto JB. Biodiversidade como fonte de medicamentos. *Ciência e cultura*. 2003; 55: 37-39.
- [20] Sangi DP. Estratégias de síntese na descoberta de fármacos: o emprego da síntese orientada pela diversidade estrutural. *Quim. Nova*. 2016; 39:995-1006.
- [21] SANT'ANNA, CMR. Glossário de termos usados no planejamento de fármacos (recomendações da IUPAC para 1997). *Química Nova*. 2002; 25: 505-512.
- [22] Alves VM, Braga RC, Andrade CH. Aplicação de metodologias do CADD a inibidores da 14 α -esterol desmetilase: triagem virtual e busca por novos candidatos a fármacos antichagásicos. *Relatório Pibic*. 2010; 2009:48–60.
- [23] Myint KZ, Xie XQ. Recent advances in fragment-based QSAR and multi-dimensional QSAR methods. *International journal of molecular sciences*. 2010; 11:3846-3866.
- [24] SILVA BAVG. Aplicação de metodologias do CADD (Computer-Aided Drug Design) a um conjunto de pirrolidina carboxamidas: mapeamento do farmacóforo e planejamento de novos protótipos tuberculostáticos potenciais. [Dissertação]. São Paulo-SP: Universidade de São Paulo, Faculdade de Ciências Farmacêuticas, Programa de Pós-Graduação em Fármaco e Medicamentos, Área de Insumos Farmacêuticos; 2018.
- [25] Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M. et al. QSAR modeling: Where have you been? Where are you going to?. *Journal of Medicinal Chemistry*. 2014; 57: 4977–5010.
- [26] Martins JPA, Ferreira MMC. QSAR modeling: um novo pacote computacional open source para gerar e validar modelos QSAR. *Química Nova*. 2013; 36: 554–560.

- [27] Liu D, Tian Z, Yan Z, Wu L, Ma Y, Wang Q, Liu W, Zhou H, Yang C. Design, synthesis and evaluation of 1, 2-benzisothiazol-3-one derivatives as potent caspase-3 inhibitors. *Bioorganic & medicinal chemistry*. 2013; 21:2960-2967.
- [28] Wu L, Lu M, Yan Z, Tang X, Sun B, Liu W, et al. 1, 2-benzisothiazol-3-one derivatives as a novel class of small-molecule caspase-3 inhibitors. *Bioorganic & Medicinal Chemistry*. 2014; 22: 2416-2426.
- [29] Guo Z, Yan Z, Zhou X, Wang Q, Lu M, Liu W, et al. Synthesis and biological evaluation of novel 1, 2-benzisothiazol-3-one-derived 1, 2, 3-triazoles as caspase-3 inhibitors. *Medicinal Chemistry Research*. 2015; 24:1814-1829.
- [30] Veerasamy R, Rajak H, Jain A, Sivadasan S, Varghese C, Agrawal R. Validation of QSAR models-strategies and importance. *Int. J. Drug Des. Discov*. 2011; 3: 511-519.
- [31] Haenlein M, Kaplan AM. A beginner's guide to partial least squares analysis. *Understanding statistics*. 2004; 3: 283-297.
- [32] Tropsha A. Best practices for QSAR model development, validation, and exploitation. *Molecular informatics*. 2010; 29: 476-488.
- [33] Gramatica P. Principles of QSAR models validation: internal and external. *QSAR & combinatorial science*. 2007; 26: 694-701.
- [34] Rodrigues RP, Mantoani SP, Almeida JR, Pinsetta FR, Semighini EP, Silva VB, Silva CHTP. Estratégias de triagem virtual no planejamento de fármacos. *Revista Virtual de Química*. 2012; 4:739-776.
- [35] Benfenati E, Manganaro A, Gini GC. VEGA-QSAR: AI inside a platform for predictive toxicology. *Health & Environmental Research Online*. 2013;1: 21-28.
- [36] Minovski N, Zuperl S, Drgan V, Novic M. Assessment of applicability domain for multivariate counter-propagation artificial neural network predictive models by minimum Euclidean distance space analysis: a case study. *Analytica chimica acta*. 2013; 759: 28-42.
- [37] Kiralj R, Ferreira M. Basic validation procedures for regression models in QSAR and QSPR studies: theory and application. *Journal of the Brazilian Chemical Society*. 2009; 20:770-787.
- [38] Roy K, Kar S, Das RN. Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment. 1ª edição. Academic Press; 2015.
- [39] Golbraikh A, Tropsha A. Beware of q²!. *Journal of molecular graphics and modelling*. 2002; 20: 269-276.
- [40] Lee D, Long SA, Adams JL, Chan G, Vaidya KS, Francis TA, et al. Potent and selective nonpeptide inhibitors of caspases 3 and 7 inhibit apoptosis and maintain cell functionality. *Journal of Biological Chemistry*. 2000; 275:16007-16014.

[41] Costa AS, Martins JPA, Melo EB. SMILES-based 2D-QSAR and similarity search for identification of potential new scaffolds for development of SARS-CoV-2 MPRO inhibitors. *Structural Chemistry*. 2022; 33:1691-1706.

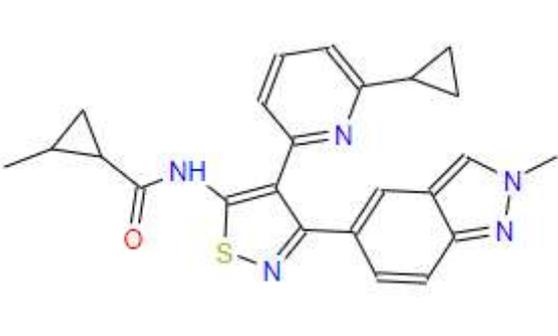
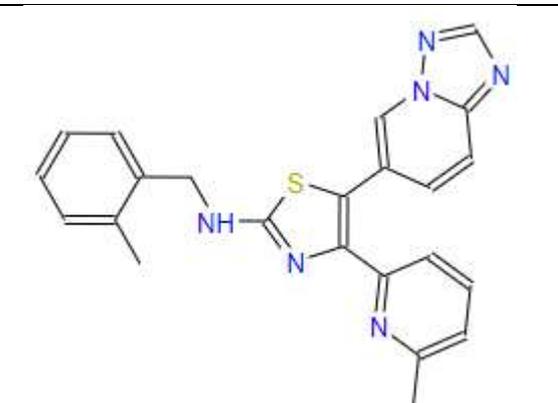
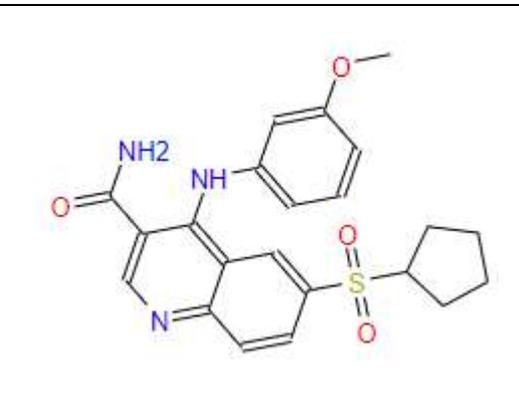
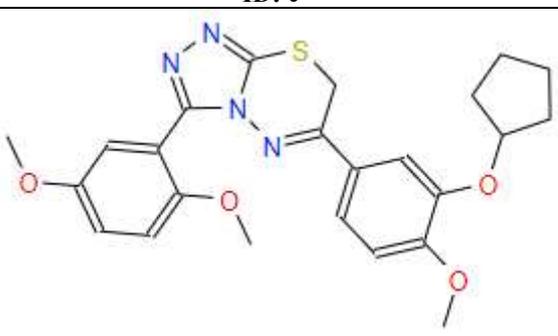
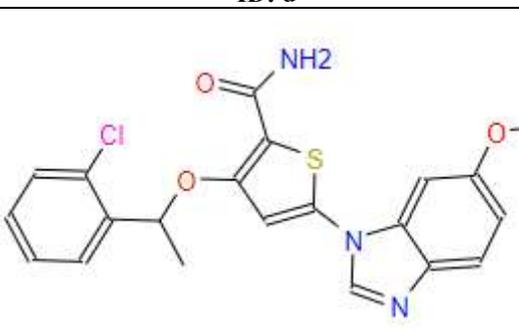
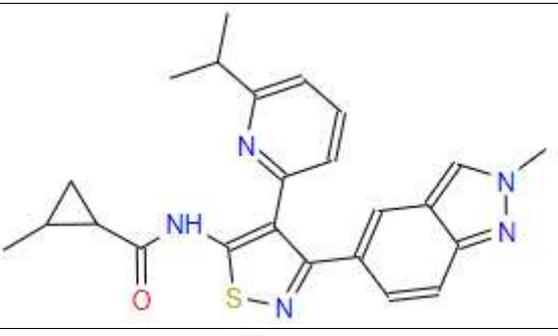
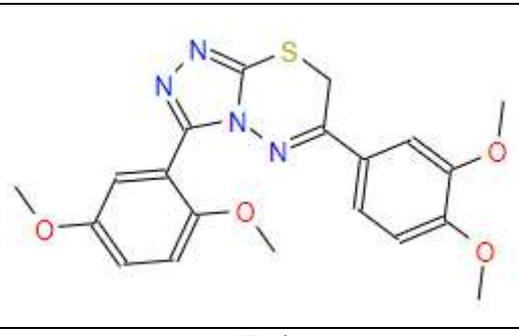
[42] Firoozpour L, Gao L, Moghimi S, Pasalar P, Davoodi J, Wang MW, et al. Efficient synthesis, biological evaluation, and docking study of isatin based derivatives as caspase inhibitors. *J Enzyme Inhib Med Chem*. 2020;35:1674–84.

Tabela 1: com os parâmetros estatísticos empregados na validação interna.

Etapa	Parâmetro	Definição	Especificação
Avaliação inicial do modelo	R^2	Coefficiente de determinação	$\geq 0,6$
	F	Teste de Fisher ($F_{p,n-p-1}$) com 95% de significância. Onde p corresponde ao nº de variáveis latentes e n o número de compostos do modelo	Maior que o valor tabelado
	$RMSEC$	Raiz quadrada do erro da calibração	Menor possível
Capacidade de previsão interna	Q^2_{Loo}	Coefficiente de determinação da validação cruzada	$\geq 0,5$
	$RMSECV$	Raiz quadrada do erro da validação cruzada	Menor possível
Sobreajuste	$R^2 - Q^2_{Loo}$	Diferença entre $R^2 - Q^2_{Loo}$	$< 0,2$
<i>y-randomization</i>	α_Q e α_R	Intercepto do R^2 e Q^2_{Loo}	$\alpha_Q < 0,05$ $\alpha_R < 0,3$

Fonte: Adaptado de Costa et al (2022) [40].

Tabela 2: Estrutura química dos compostos *hits* obtidos pelo estudo de triagem virtual 2D.

 <p>Chemical structure of compound ID: a, featuring a thiazole ring substituted with a cyclopropylcarbamoyl group, a cyclopropyl group, and a fused benzimidazole system.</p>	 <p>Chemical structure of compound ID: b, featuring a pyridine ring substituted with a methoxy group and a side chain containing a 1,3,4-oxadiazole ring and a quaternary carbon.</p>
ID: a	ID: b
 <p>Chemical structure of compound ID: c, featuring a thiazole ring substituted with a benzyl group, a pyridine ring, and a fused benzimidazole system.</p>	 <p>Chemical structure of compound ID: d, featuring a benzimidazole ring system substituted with an amino group, a methoxy group, and a sulfonamide group attached to a cyclopentane ring.</p>
ID: c	ID: d
 <p>Chemical structure of compound ID: e, featuring a thiazole ring substituted with a benzimidazole system, a methoxy group, and a cyclopentane ring.</p>	 <p>Chemical structure of compound ID: f, featuring a thiazole ring substituted with a benzimidazole system, a chlorine atom, and a methoxy group.</p>
ID: e	ID: f
 <p>Chemical structure of compound ID: g, featuring a thiazole ring substituted with a cyclopropylcarbamoyl group, a pyridine ring, and a fused benzimidazole system.</p>	 <p>Chemical structure of compound ID: h, featuring a thiazole ring substituted with a benzimidazole system and a methoxy group.</p>
ID: g	ID: h

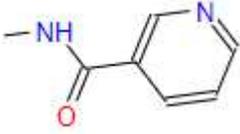
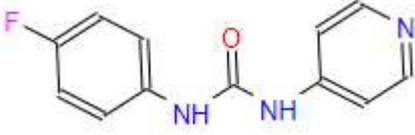
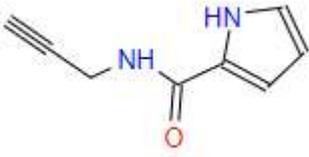
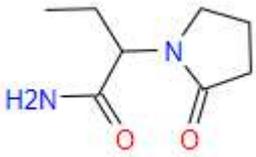
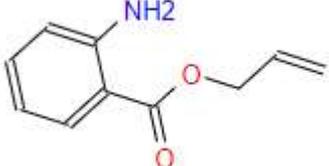
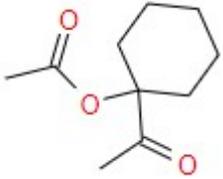
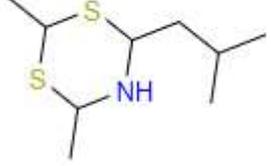
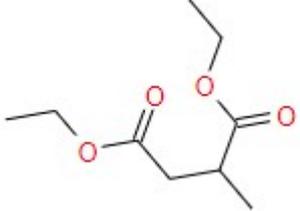
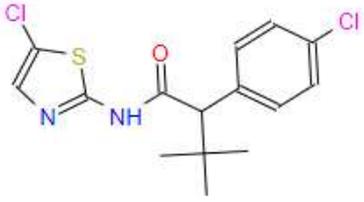
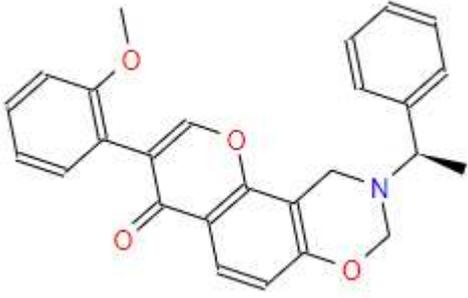
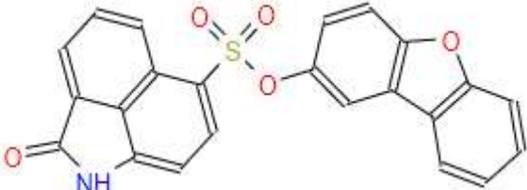
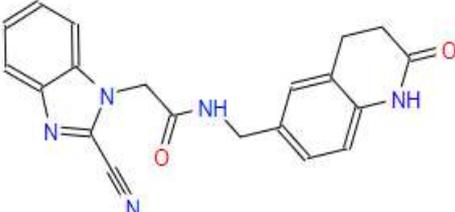
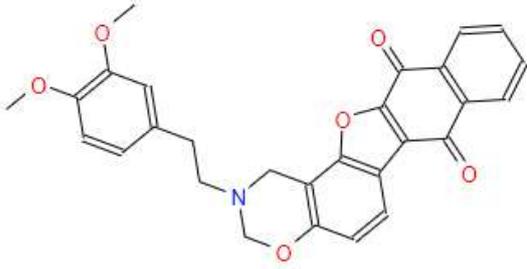
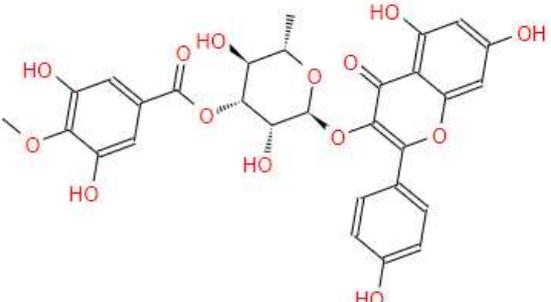
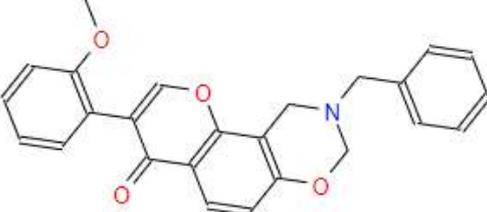
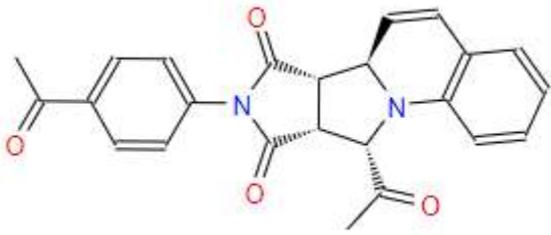
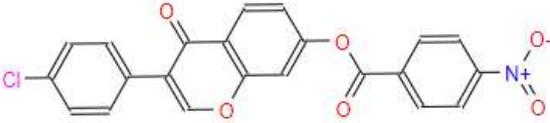
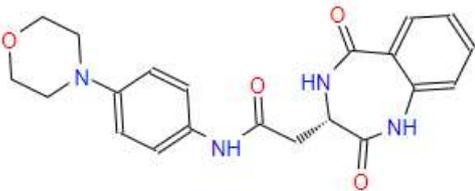
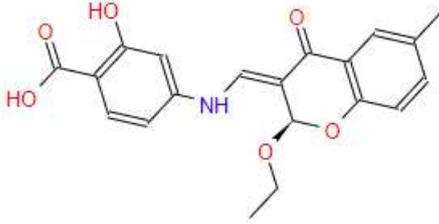
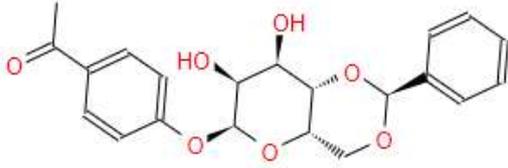
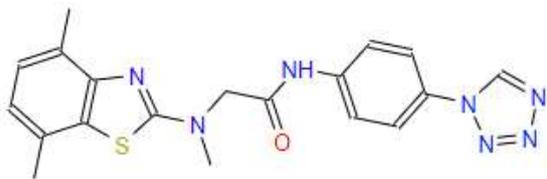
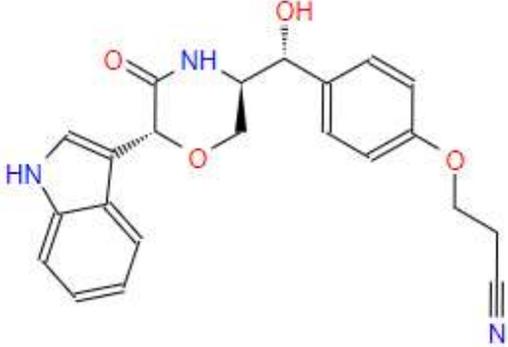
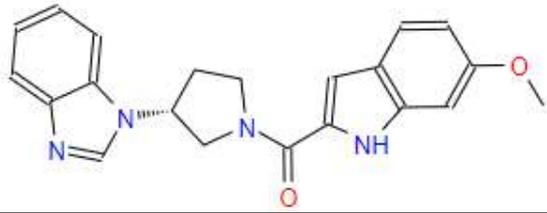
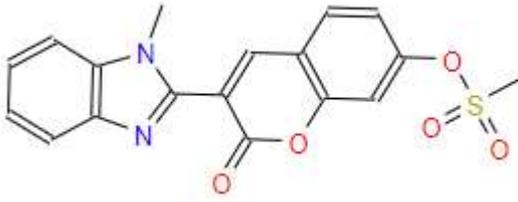
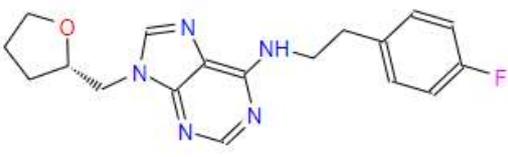
	
<p>ID: i</p>	<p>ID: j</p>
	
<p>ID: k</p>	<p>ID: l</p>
	
<p>ID: m</p>	<p>ID: n</p>
	
<p>ID: o</p>	<p>ID: p</p>
	
<p>ID: q</p>	<p>ID: r</p>
	
<p>ID: s</p>	<p>ID: t</p>

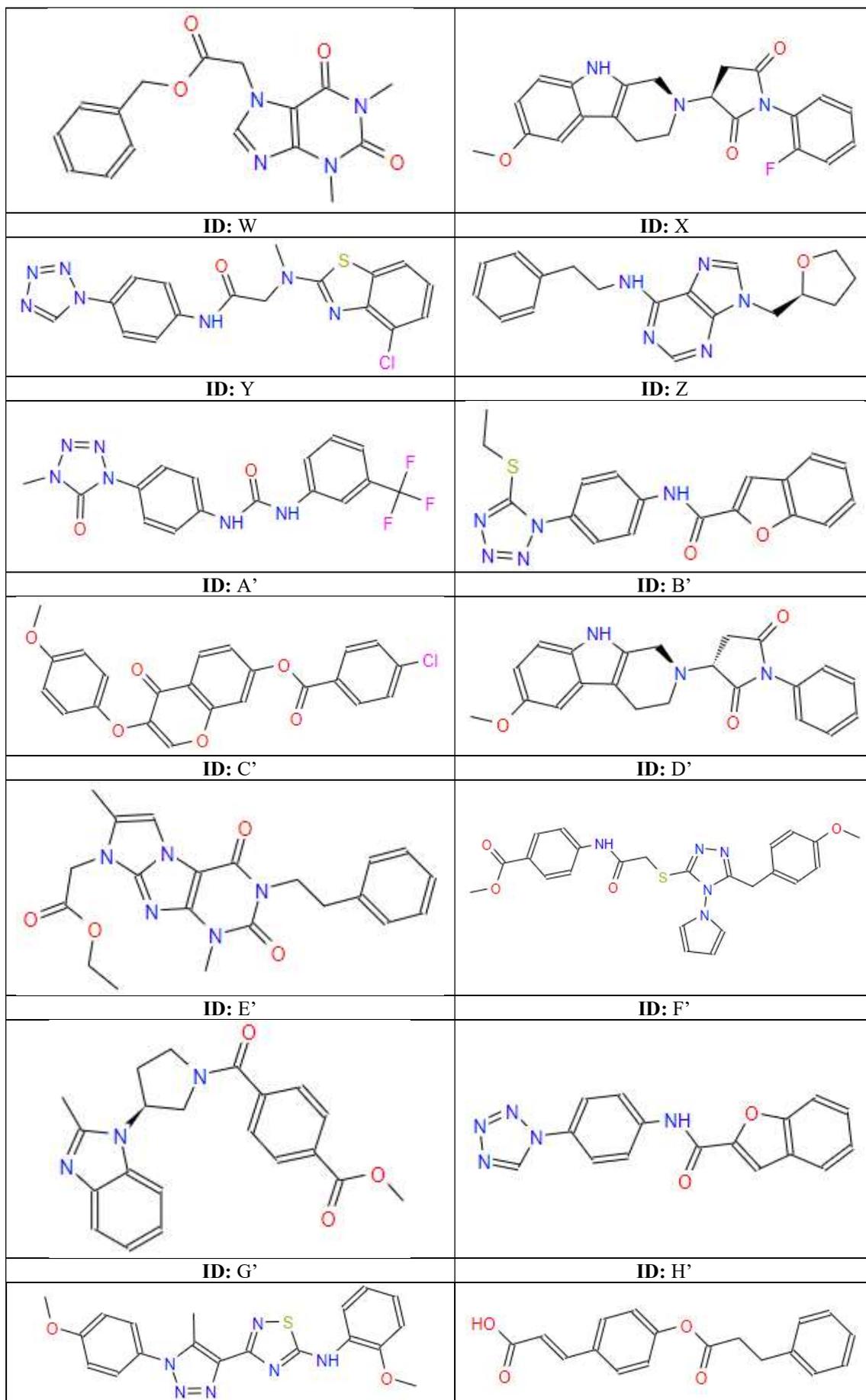
Tabela 3: Compostos obtidos na triagem virtual 3D com seus respectivos códigos de busca nas bases de dados triadas.

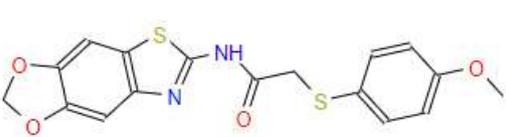
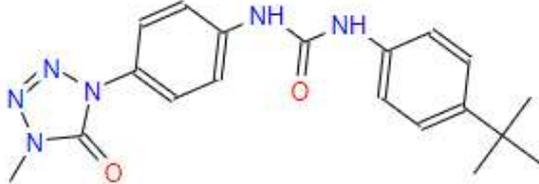
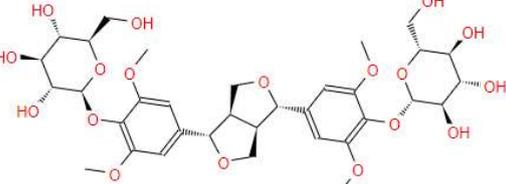
Composto	Base de dados	Código de identificação na base de dados
A	Life Chemicals	F3385-6017
B	Life Chemicals	F1217-0041
C	Life Chemicals	F1414-1154
D	Coconut	CNP0009925
E	Life Chemicals	F1094-0206
F	Life Chemicals	F6672-1375
G	Nubbe	ZINC000148010066
H	Life Chemicals	F3385-6005
I	Coconut	CNP0219174
J	Life Chemicals	F2629-0109
K	Life Chemicals	F3139-1054
L	Life Chemicals	F6671-3190
M	Coconut	CNP0409192
N	Coconut	CNP0466814
O	Life Chemicals	F2077-0993
P	Coconut	CNP0183629
Q	Life Chemicals	F6225-0779
R	Coconut	CNP0192370
S	Life Chemicals	F6559-1117
T	Life Chemicals	F3385-5554
U	Life Chemicals	F3228-0004
V	Life Chemicals	F6796-0764
W	Life Chemicals	F3259-0272
X	Coconut	CNP0127577
Y	Life Chemicals	F6225-1036
Z	Life Chemicals	F6796-0689
A'	Life Chemicals	F6233-0235
B'	Life Chemicals	F3375-0039
C'	Life Chemicals	F0916-2511
D'	Coconut	CNP0396021
E'	Life Chemicals	F3234-0454
F'	Life Chemicals	F2563-0707
G'	Life Chemicals	F6559-1265
H'	Life Chemicals	F3375-0040
I'	Life Chemicals	F3406-4687
J'	Coconut	CNP0181015
K'	Life Chemicals	F6796-0450
L'	Life Chemicals	F2816-0473
M'	Life Chemicals	F6233-0224
N'	Nubbe	ZINC000085486993

Tabela 4: Estrutura química dos compostos obtidos na triagem virtual 3D.

	
ID: A	ID: B
	
ID: C	ID: D
	
ID: E	ID: F
	
ID: G	ID: H
	
ID: I	ID: J

	
<p>ID: K</p>	<p>ID: L</p>
	
<p>ID: M</p>	<p>ID: N</p>
	
<p>ID: O</p>	<p>ID: P</p>
	
<p>ID: Q</p>	<p>ID: R</p>
	
<p>ID: S</p>	<p>ID: T</p>
	
<p>ID: U</p>	<p>ID: V</p>



<p style="text-align: center;">ID: I'</p> 	<p style="text-align: center;">ID: J'</p> 
<p style="text-align: center;">ID: K'</p> 	<p style="text-align: center;">ID: L'</p> 
<p style="text-align: center;">ID: M'</p>	<p style="text-align: center;">ID: N'</p>