



UNIVERSIDADE DE BRASÍLIA  
INSTITUTO DE FÍSICA

TESE DE DOUTORADO

SIMILARIDADE, DOCKING E DINÂMICA MOLECULAR: COMBINAÇÃO DE  
ESTRATÉGIAS NA BUSCA DE NOVOS INIBIDORES DA hAChE

Nádia Melo Borges

Orientador: Prof. Ricardo Gargano

Co-orientador: Prof. Carlos Alberto Montanari

Brasília, 15 de dezembro de 2017

SIMILARIDADE, DOCKING E DINÂMICA MOLECULAR: COMBINAÇÃO DE  
ESTRATÉGIAS NA BUSCA DE NOVOS INIBIDORES DA hACHE

Por

Nádia Melo Borges

Tese de Doutorado apresentada ao Instituto de Física da Universidade de Brasília  
como parte dos requisitos necessários para a obtenção do título de Doutor em Física.

Aprovada por:

---

Prof. Dr. Solemar Silva Oliveira  
Universidade Estadual de Goiás - UEG

---

Prof. Dr. João Batista Lopes Martins  
Instituto de Química - UnB

---

Dra. Alessandra Sofia Kiametis  
Instituto de Biologia - UnB

Brasília,

---

Prof. Dr. Geraldo Magela e Silva  
Coordenador de Pós-Graduação  
Instituto de Física  
Universidade de Brasília

Brasília, 15 de dezembro de 2017

# Agradecimentos

Agradeço, antes de tudo, a Deus por ser tão bom comigo, me amando mesmo eu não merecendo e me proporcionando momentos únicos vividos nesse período do doutorado.

Aos meus pais, Damião e Terezinha, pelo amor incondicional, incentivo aos estudos e apoio em todos os momentos pois eu não seria metade do que sou sem eles.

Ao meu esposo Jean que sempre esteve comigo me aconselhando, cuidando e incentivando a ser cada dia melhor no meio acadêmico. Ele que foi meu companheiro no dia a dia do doutorado, meu melhor amigo e confidente, também caminhou comigo nos momentos felizes do doutorado sanduíche e nos difíceis enfrentando as diferenças da cultura e da língua, tornando esse período o mais leve possível.

A minha irmã Magda, cunhado Gercy e adorável sobrinha Beatriz que são minha segunda família, cuidam de mim como uma filha, sempre me apoiando e incentivando a acreditar mais em mim, obrigada pelo amor que me concedem.

A todos os avós, tios e primos que sempre acreditaram em mim e mesmo de longe me fortaleceram com mensagens carinhosas.

Aos meus sogros, Ângelo e Oneida, cunhados e concunhados que se tornaram minha família e sempre me apoiaram nas decisões da pós-graduação com carinho e sabedoria.

Ao meu orientador e amigo, Professor Dr. Ricardo Gargano, pela orientação, dedicação, paciência e principalmente por se tornar um verdadeiro pai nesses últimos seis anos de pós-graduação, me apoiando em todas as decisões e escolhas e aconselhando com sabedoria nos momentos em que estive errada. Com certeza, um dos maiores ganhos durante todo o processo foi ter conhecido essa pessoa ímpar que contribuiu intensamente no meu crescimento científico e pessoal.

Ao co-orientador, Professor Dr. Carlos Alberto Montanari, a qual admiro e respeito, pelo apoio e oportunidade que me foram concedidos, por ter me recebido tão bem no grupo Nequimed contribuindo imensamente para o meu crescimento profissional, por todas as contribuições a mim e ao projeto e pela confiança em mim depositada ao assumir a co-orientação.

Ao meu supervisor de estágio, Professor Dr. Charles Loughton que me recebeu e tanto me ensinou no meu doutorado sanduíche na Escola de Farmácia da Universidade de Nottingham – Inglaterra. Meu agradecimento sincero por toda dedicação e amizade.

Ao Professor Dr. João Batista, do Instituto de Química da UnB, pela colaboração durante o desenvolvimento do projeto.

Aos professores da UnB, Geraldo Magela, Pedro Henrique, William Cunha, Alessandra Albernaz, Demétrio Filho e Luiz Roncaratti, pelos momentos e conhecimentos compartilhados.

Ao querido e amigo Professor Dr. Peter Kenny que conquistou minha eterna admiração com seu jeito singular e disponibilidade em ensinar, abraçando a todos, inclusive a mim que tive a honra de poder compartilhar espaço no laboratório Nequimed em São Carlos.

Ao Professor Dr. Andrei Leitão, que me acolheu muito bem no laboratório Nequimed, sempre disposto a ajudar.

A Queila, aquela amiga de todas as horas, que ficará guardada para sempre como uma pessoa especial, pois mesmo estando longe é companheira. Agradeço a todos os momentos que me apoiou sendo uma irmã em Brasília.

Ao amigo Alexandre Adriano, pois me hospedou humildemente e coração grande que só ele tem, quando mais precisei.

Aos amigos, Almir, Bruno, Dyorgge, Ísis, Jaqueline, Letícia Massa, Miguel, Philipp, Renata, Suzi, Verenna, Víncius, Virgílio e os não citados, da colina que tornaram os meus dias em Brasília mais divertidos e descontraídos. Serão sempre lembrados com carinho.

A todos os amigos do Instituto de Física da UnB em especial aqueles que estiveram mais presentes, Queila, Lauriane, Erinaldo, Igo, Rodolpho, Marcos, Fábio Moura, Jonathan, Vavá, Mirian, Bruno, Rhuiago, Gabriel, Henrique, Luciano, Alessandra Sofia, Liz, Marco, Mônica, Andriele, Tatiane, Mario, Cyntia, Thiago, Horácio, Gesiel, Aline, André Filho, Vanessa, Letti, Clauber, Clever, Felipe Luís, Pedro Dias, Keila, Ana Claudia, Sara, Vinicius Alves, Ana Virginia, Fernando Mendes, Oscar, Sandra e Thalles, meu muito obrigada por cada contribuição e luta do dia a dia.

Aos amigos, Geraldo e Josmar que me ajudaram imensamente no meu projeto de estágio em São Carlos, pois a eles devo tanto conhecimento adquirido e amizade verdadeira.

Aos amigos que conquistei no grupo Nequimed, Instituto de Química da USP e em São Carlos, em especial a Dr. Maria Luiza, Michelle, Dri, Cristian, Lina, Erika, Felipe, William, Daniel, Daniela, Lorenzo, Karen, Fran, Dani, Jussara, Leandro, Gabriella, Luciano Cordeiro, Ana Maria, Nati e Patiño,

Aos novos amigos que conquistei no tempo de doutorado sanduíche na Inglaterra, Ian, Marco Pasi, Khaled, Ioanna, Damiano, Xi Chen, Uijin Jeong, Jong Bong, Sebastian, Lori, Hannah, Zhi Yuan, Divneet, Max, Selva, Mohamed, Ismael, Christophe, Camila, Bruno, Angeliki e Davidae.

A CAPES, pelo apoio financeiro e bolsa concedida durante o projeto de doutorado sanduíche.

E a todos aqueles(as) que aqui não foram citados, mas que de alguma forma colaboraram na realização deste trabalho.

“Seu mundo muda quando você muda”  
*Bento Augusto*

## Resumo

A doença de Alzheimer é a forma mais comum de demência irreversível entre as pessoas idosas. A diminuição dos níveis de acetilcolina no cérebro de pacientes com a doença está relacionada com a fisiopatologia da doença. A hipótese colinérgica é baseada no aumento do nível de acetilcolina pela inibição reversível da enzima acetilcolinesterase. O objetivo principal deste trabalho é abordar o desempenho de métodos de triagem virtuais com base no ligante e na estrutura da proteína afim de encontrar novos candidatos a inibidores da acetilcolinesterase. Além disso, um protocolo foi desenvolvido para identificar e propor novos inibidores promissores da AChE a partir do banco de dados ZINC com 10 milhões de compostos disponíveis comercialmente. Neste sentido, busca por similaridade 3D usando similaridade por forma e eletrostática além do método de docagem para uma série de compostos foram utilizados para recuperar inibidores da AChE a partir de uma coleção de banco de dados. Simulação por dinâmica molecular foi realizada para estudar os melhores compostos docados. Alguns resíduos importantes foram identificados como indispensáveis para formação do duplo modo de interação entre os compostos selecionados e o sítio de interação da enzima. Todos os resultados indicam o relevante uso do método de similaridade eletrostática combinado com estratégias de docagem para identificar novos possíveis inibidores, e sete novas estruturas foram selecionadas como potentes candidatos anticolinesterásicos.

## Abstract

Alzheimer's disease is the most common cause of irreversible dementia among the elderly. The decrease of acetylcholine levels in the brain of patients with the disease is related to the pathophysiology of the disease. The cholinergic hypothesis is based on increasing the level of acetylcholine by the reversible inhibition of the enzyme acetylcholinesterase. The main purpose of this study is to address the performance of virtual screening methods based on ligands and the protein structure of acetylcholinesterase in order to retrieve novel hAChE inhibitors. In addition, a protocol was developed to identify novel hit compounds and propose new promising AChE inhibitors from ZINC database with 10 million commercially available compounds. In this sense, 3D similarity searches using Rapid Overlay of Chemical Structures and similarity analysis through comparison of electrostatic overlay of docked hits were used to retrieve AChE inhibitors from collected databases. Molecular dynamics simulation of 100 ns was carried out to study the best docked compounds from similarity search. Some key residues are identified as crucial for the dual binding mode of inhibitor with interaction site. All results indicated the relevant use of EON and docking strategy for identifying novel hit compounds as promising potential anticholinesterase candidates, and seven new structures were selected as potential hAChE inhibitors.

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
1.1	A Doença Alzheimer	15
1.2	Hipótese Colinérgica	16
1.2.1	Acetilcolinesterase	18
1.2.2	Butirilcolinesterase	20
<b>2</b>	<b>OBJETIVOS</b>	<b>21</b>
2.1	Objetivos Gerais	21
2.2	Objetivos Específicos	21
<b>3</b>	<b>METODOLOGIAS</b>	<b>22</b>
3.1	Conjunto de Dados	23
3.2	Métodos Baseados no Ligante	24
3.2.1	Similaridade Química	24
3.3	Métodos baseados na estrutura do alvo	26
3.3.1	Docagem Molecular	26
3.4	Métricas de avaliação dos métodos	28
3.5	Dinâmica Molecular	29
3.5.1	Análises de Triplet Clique	29
<b>4</b>	<b>RESULTADOS E DISCUSSÕES - VALIDAÇÃO DO PROTOCOLO DE SELEÇÃO</b>	<b>31</b>
4.1	ROCS e EON	31
4.2	Docagem Molecular	33
4.2.1	Procedimento de Redocagem Molecular	33
4.2.2	Análises da Docagem Molecular	34
<b>5</b>	<b>RESULTADOS E DISCUSSÕES – SELEÇÃO DOS COMPOSTOS</b>	<b>38</b>
5.1	Ensaio virtuais na busca por novos esqueletos base inibidores da hAChE	38
5.1.1	Filtros por ensaios virtuais	39
5.1.2	Dinâmica Molecular	44
5.1.3	Análise de triplet clique e caracterização do modo de ligação em termos de parâmetros geométricos	45
<b>6</b>	<b>CONSIDERAÇÕES FINAIS E PERSPECTIVAS</b>	<b>47</b>

<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>49</b>
<b>APÊNDICE A – INIBIDORES DA ENZIMA CPB DE LEISHMANIA MEXICANA: MODELAGEM POR HOMOLOGIA, DOCAJEM E DINÂMICA MOLECULAR .....</b>	<b>54</b>
<b>APÊNDICE B - SIMILARITY SEARCH COMBINED WITH DOCKING AND MOLECULAR DYNAMICS FOR NOVEL HACHE INHIBITOR SCAFFOLDS.....</b>	<b>68</b>

# Lista de Figuras

Figura 1: Primeiros fármacos ativos contra a enzima AChE. ....	16
Figura 2: Características funcionais do sistema colinérgico (Adaptado [1]). ....	17
Figura 3: Sítio ativo da AChE em complexo com Donepezil. ....	19
Figura 4: a) Neurotransmissor Acetilcolina b) Acetato e colina após a hidrólise da Acetilcolina. ....	19
Figura 5: Inibidores hAChE usados como estrutura de referência: (A) Tacrina IC <sub>50</sub> = 0.1 µM [46] (B) Galantamina IC <sub>50</sub> = 0.3 µM [46] (C) Rivastigmina IC <sub>50</sub> = 0.7 µM [46] (D) Donepezil IC <sub>50</sub> = 0.002 µM [46] (E) HuprinaW IC <sub>50</sub> = 0,001 µM [47]. ....	25
Figura 6: Estruturas cristalograficas da hAChE alinhadas em complexo com a HuprinaW (4BDT) em dourado, Dihydrotanshinone I (4M0E) em cinza e Donepezil (4EY7) em azul. ....	27
Figura 7: Curvas ROC dos programas ROCS e EON para os conjuntos de dados (a) DS1, (b) DS2 and (c) DS3. ....	32
Figura 8: Redocagem dos ligantes co-cristalizados. (a) Superposição entre os ligantes co-cristalizados (cinza) e as melhores poses previstas pela redocagem (verde) para (1) Huprina W (0.89Å), (2) Donepezil (1.34Å) e (3) Dihydrotanshinone I (0.82Å) e (b) suas estruturas moleculares. ....	34
Figura 9: Curvas ROC do programa de docagem molecular Glide para as estruturas cristalográficas (a) 4EY7, (b) 4BDT e (c) 4M0E. ....	35
Figura 10: Representação bidimensional das interações entre a estrutura da hAChE e o Donepezil. ....	36
Figura 11: Protocolo hierárquico de ensaios virtuais usados para identificar inibidores da hAChE. ....	38
Figura 12: Representação das estruturas propostas como novos potentes candidatos a inibidores hAChE. ....	39
Figura 13: Resultado de docagem do composto N01 na estrutura cristalográfica hAChE 4M0E. Ligações de hidrogênio estão representadas por linhas tracejadas amarelas; interações do tipo π-stacking estão representadas por linhas tracejadas azuis; interações do tipo Cátion-π estão representadas por linhas tracejadas verdes, ....	40
Figura 14: Estrutura dos complexos previstos por docagem molecular com modo de	

interação semelhante a inibidores co-cristalizados selecionados na análise visual (a) N01 (b) N02 (c) N03 (d) N04. ....	42
Figura 15: Estrutura dos complexos preditos por docagem molecular com modo de interação semelhante a inibidores co-cristalizados selecionados na análise visual (e) N05 (f) N06 (g) N07.....	43
Figura 16: Resultado da docagem molecular do composto N01 no sítio ativo da hAChE em tempos diferentes da simulação por dinâmica molecular: 0 ns, 25 ns, 50 ns, 75 ns e 100 ns. ....	44
Figura 17:RMSD dos compostos versus tempo de simulação de dinâmica molecular. ....	45
Figura 18: O complexo hAChE-N05 e os padrões geométricos da interação proteína-ligante. A união do ponto de origem de cada resíduo define um triângulo Trp83-Tyr334-N05 de lados 6,4, 4,7, e 7,5 Å e ângulos de $85^{\circ} \pm 4$ , $38^{\circ} \pm 3$ e $57^{\circ} \pm 3$ . ....	46

# Lista de Abreviaturas

1YL	inibidor Dihydrotanshinone I
3D	tridimensional
4BDT	estrutura tridimensional da acetilcolinesterase humana complexada com o inibidor HuprinaW
4EY7	estrutura tridimensional da acetilcolinesterase humana complexada com o inibidor Donepezil
4M0E	estrutura tridimensional da acetilcolinesterase humana complexada com o inibidor Dihydrotanshinone I
A $\beta$	placas $\beta$ amiloide
Acetil CoA	acetilcoenzima A
ACh	neurotransmissor acetilcolina
AChE	enzima acetilcolinesterase
Å	Angstrom
Ala	aminoácido alanina
AUC	área sob a curva ROC
BuChE	enzima butirilcolinesterase
CAS	sítio de interação catalítico
ChAT	enzima colina acetiltransferase
ChEs	enzimas colinesterases
ChEIs	inibidores colinesterásicos
DA	doença de Alzheimer
$\Delta G$	energia livre de Gibbs
DS	conjunto de dados
DUD-e	directory of useful decoys enhanced
E20	inibidor Donepezil
ENF	emaranhados neuro-fibrilares
EON	análise de similaridade através da comparação de sobreposição eletrostática
FDA	Administração federal de alimentos e medicamentos dos Estados Unidos

fs	femtossegundos
Glu292	aminoácido glutamato e o número correspondente a sua posição na cadeia primária
hAChE	enzima acetilcolinesterase humana
His447	aminoácido histidina e o número correspondente a sua posição na cadeia primária
HUW	inibidor HuprinaW
IC <sub>50</sub>	concentração para inibir 50% da atividade enzimática
Ile	aminoácido isoleucina
K	graus Kelvin
K <sup>+</sup>	potássio
Kcal	quilocalorias
Kcal/mol	uma quilocaloria de energia por mol de substância
K <sub>i</sub>	constante de inibição
LBVS	ensaio virtual baseado na estrutura do ligante
Leu	aminoácido leucina
MD	dinâmica molecular
m/s	metros por segundo
μM	micromolar
Na <sup>+</sup>	sódio
nM	nanomolar
ns	nanossegundos
PAS	sítio de interação periférico
PDB	Protein Data Bank: Banco de dados de proteínas
Phe205	aminoácido fenilalanina e o número correspondente a sua posição na cadeia primária
PPA	proteína precursora amiloide
PS	placas senis
ps	picossegundos
reISESA	área de superfície excluída do solvente
RMSD	raiz quadrada do desvio quadrático médio
RMN	ressonância magnética nuclear

ROC	curva característica operativa do receptor
ROCS	superposição rápida de estruturas químicas
s	segundo
SBVS	ensaio virtual baseado na estrutura do receptor
Ser203	aminoácido serina e o número correspondente a sua posição na cadeia primária
Trp86	aminoácido triptofano e o número correspondente a sua posição na cadeia primária
tcAChE	enzima acetilcolinesterase do <i>torpedo californica</i>
Tyr337	aminoácido tirosina e o número correspondente a sua posição na cadeia primária
Val	aminoácido valina
VS	ensaios virtuais
X	ângulo de torção

# 1 Introdução

## 1.1 A Doença Alzheimer

A doença de Alzheimer (DA) é dentre outras a forma mais comum de demência, entre pessoas com idade superior a 65 anos, causando um distúrbio neurodegenerativo, crônico e progressista, associado ao comprometimento de funções ligadas as atividades de vida diária [1-3]. As alterações apresentadas pelos pacientes dependem do estágio da DA. No estágio inicial os pacientes apresentam alterações cognitivas, comportamental e de humor, cujas principais características são esquecimento frequente e perguntas repetitivas. No estágio intermediário, há piora das funções cognitivas, perda de memória recente e problemas com linguagem. No último estágio há severas alterações comportamentais, perda de identidade e o paciente torna-se excessivamente dependente [1-7].

Estudos realizados para o tratamento da doença utilizam três fisiopatologias principais [6-15]. Os dois critérios histopatológicos para DA são, observações de depósitos extracelulares de peptídeos fibrilares chamadas placas senis (PS) e de emaranhados neuro-fibrilares (ENF) intraneuronais generalizados, ambos foram descritos por Alois Alzheimer em seu resultado de estudo de caso [12]. As PS são formadas a partir do acúmulo anormal extracelular e depósitos dos peptídeos  $\beta$  amiloide com 40 ou 42 resíduos de aminoácidos (conhecido como placas  $\beta$  amiloide ( $A\beta$ )) do metabolismo da proteína precursora amiloide (PPA) encontrada nas membranas das células e organelas como as mitocôndrias [11]. O acúmulo de proteínas  $\beta A$  que promove a sua precipitação e agregação em forma destas placas leva a atrofia e degeneração dos neurônios colinérgicos, impedindo as sinapses. Os ENF são formados devido as alterações na integridade da estrutura dos microtúbulos axônios, responsáveis pelos contatos Inter neuronais, causadas pela hiperfosforilação da proteína tau [16]. O acúmulo dos emaranhados no interior dos neurônios compromete o transporte axonal e a função neuronal. As PS e os ENF não são exclusivas da DA [6,13], podendo ser encontradas no envelhecimento normal e em outras doenças degenerativas, necessitando um melhor entendimento patogênico [17]. Bioquimicamente, a DA é decorrente da diminuição da neurotransmissão colinérgica, devida aos baixos níveis de acetilcolina (ACh) na fenda sináptica. Segundo a Hipótese Colinérgica, formulada a partir de observações de perda considerável de neurônios [18], o aumento do nível de acetilcolina por inibição reversível das enzimas colinesterases (ChEs) poderia melhorar o perfil cognitivo dos pacientes [19]. No segundo item deste capítulo tem uma discussão mais abrangente sobre a Hipótese Colinérgica, teoria bioquímica mais aceita mundialmente e a primeira forma racional de tratamento para esta neuropatologia.

Em todo o mundo, cerca de 36 milhões de pessoas tem a doença de Alzheimer ou alguma demência relacionada. A DA é mais comum na Europa Ocidental e menos prevalente na África Subsaariana. Apenas 1 em cada 4 pessoas com DA foram diagnosticados [20]. Uma pequena porcentagem dos casos de DA, provavelmente menos de 1%, são causadas por variações genéticas raras encontradas em um pequeno número de famílias em todo o mundo. Nestas pessoas, a doença tende a desenvolver antes dos 65 anos,

podendo ocorrer a partir dos 30 anos [21]. Trabalhos realizados em pacientes com demência revelam que quanto mais cedo o tratamento for iniciado melhor é a perspectiva ao longo prazo. Tratar a doença de Alzheimer em estágios iniciais tem como objetivo preservar as habilidades dos pacientes próximo do normal pelo maior tempo possível [22]. Entre os diferentes tipos de drogas, os inibidores da colinesterase (ChEIs), amplamente utilizados, representam a primeira classe de medicamentos aprovados para o tratamento sintomático específico da DA[23]. O primeiro inibidor aprovado pela FDA (Food and Drug Administration) foi a tacrina, entretanto sua aplicação tem sido limitada devido aos sérios efeitos colaterais, como hepatotoxicidade. Posteriormente donepezil, rivastigmina e galantamina se tornaram disponíveis (Figura 1) [23,24].

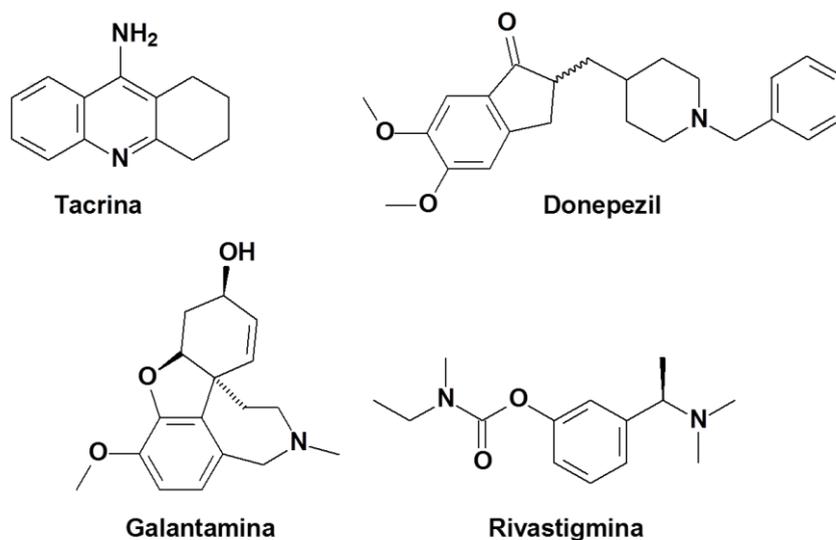


Figura 1: Primeiros fármacos ativos contra a enzima AChE.

Os efeitos terapêuticos dos fármacos em uso atualmente para o tratamento da DA limitam-se a diminuir a taxa de progressão natural da doença, permitindo apenas uma melhora temporária do estado funcional do paciente [25]. Um certo número de efeitos indesejáveis, tem sido associado a agentes terapêuticos colinérgicos convencionais. Assim, a busca por novas drogas que melhoram o perfil cognitivo dos pacientes e ao mesmo tempo reduz os efeitos secundários representa um importante avanço na luta contra doenças como o Alzheimer.

## 1.2 Hipótese Colinérgica

Estudos histológicos, complementados com técnicas avançadas de imagens, como a tomografia por emissão de pósitrons e a ressonância magnética, mostram que há perda da atividade colinérgica em pacientes com DA. Uma perda significativa ocorre em áreas do cérebro associados com os pensamentos, planos, lembranças e formação de novas lembranças, ou seja, do hipocampo e do córtex. A consequente diminuição da neurotransmissão colinérgica deu origem a Hipótese Colinérgica que consiste na restauração do nível do neurotransmissor ACh encontrado em baixas concentrações no cérebro de pacientes portadores da DA. Esta restauração pode ser feita através da inibição

reversível das colinesterases, AChE e butirilcolinesterase (BuChE), enzimas que naturalmente catabolizam a acetilcolina liberada no processo de neurotransmissão [26].

No processo de neurotransmissão (Figura 2), a membrana dos neurônios não estimulados apresenta uma diferença de potencial elétrico entre o seu interior e o seu exterior, denominado potencial membrana de repouso. Aplicando um estímulo a essa membrana, ocorre um desequilíbrio temporário entre as cargas elétricas de um lado e do outro da membrana, chamado de potencial local, e a membrana é dita polarizada. Sua polarização permite que uma grande quantidade de íons ( $\text{Na}^+$ ) entrem na célula e a quantidade relativa de carga positiva no seu interior aumenta, ocasionando sua despolarização. Posteriormente, uma grande quantidade de íons ( $\text{K}^+$ ) saem do interior da membrana que volta a ficar com excesso de cargas negativas, provocando sua repolarização. Essa onda de despolarizações e repolarizações que se propaga pela membrana plasmática do neurônio é o impulso nervoso, transmitido de uma célula a outra através das sinapses. O impulso nervoso é conhecido por potencial de ação, fenômeno de natureza eletroquímica que ocorre devido as modificações, descritas anteriormente, na permeabilidade da membrana do neurônio. Existem dois tipos de sinapses, química e elétrica. Nas sinapses químicas, quando o potencial de ação atinge as extremidades da célula pré-sináptica há liberação na fenda sináptica (espaço entre as membranas das células emissoras e receptoras) de neurotransmissores que interagem com os receptores presentes na membrana das células pós-sinápticas, desencadeando o potencial de ação. Na neurotransmissão colinérgica o neurotransmissor liberado é a acetilcolina. Para impedir que a ACh se ligue aos seus receptores e recomece o ciclo ela é destruída pela ação catabólica das enzimas AChE e BuChE, ou é absorvido, normalmente na terminação pré-sináptica. A taxa de propagação pode exceder 50 m/s em alguns neurônios, resultando na rápida propagação de informações de um neurônio para outro [27].

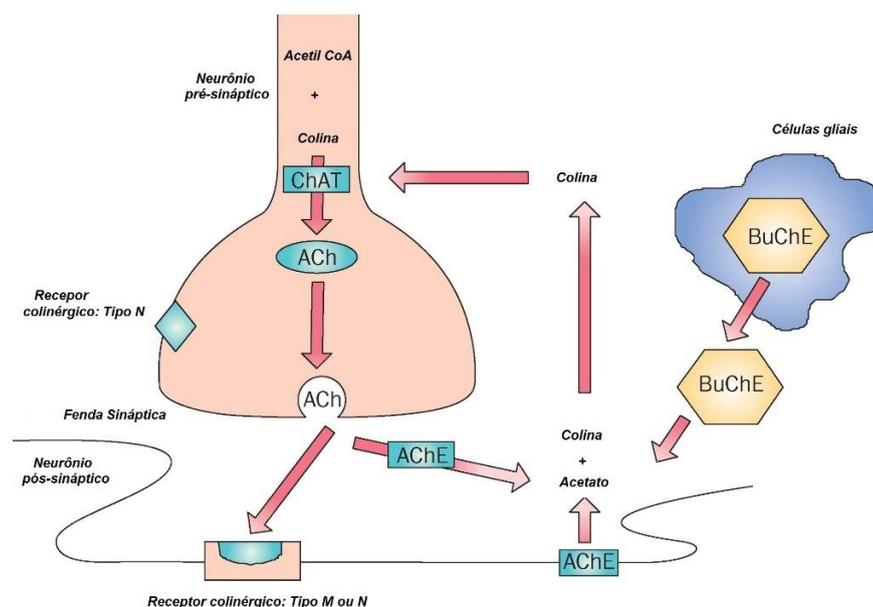


Figura 2: Características funcionais do sistema colinérgico (Adaptado [1]).

A melhora da função cognitiva ocorre quando se aumenta os níveis de acetilcolina por inibição das enzimas colinesterases, principalmente a AChE. A inibição pode ser seletiva ou promíscua. Portanto, a inibição da BuChE deve ser reconsiderada, uma vez que, a atividade da BuChE no cérebro é apenas cerca de 10% da atividade da AChE e sua inibição parcial é suficiente para aumentar os níveis de acetilcolina [28].

## 1.2.1 Acetilcolinesterase

A acetilcolinesterase é uma enzima que hidrolisa o neurotransmissor acetilcolina e é extremamente importante para a regulação da neurotransmissão em todas as áreas do sistema nervoso. As estruturas 3D da AChE disponíveis no banco de dados Protein Data Bank (PDB) de espécies tais como arraia elétrica do Pacífico (*Torpedo californica*), camundongo (*Mus musculus*) e humana (*Homo sapiens*) são bastante semelhantes. Todos os casos revelam que o sítio ativo da enzima é formado por um grupo de cadeias laterais aromáticas e está localizado na parte inferior de um canal estreito com cerca de 20 Å, chamado “garganta”. Existe hoje cerca de 200 estruturas cristalográficas da AChE no PDB e 19 destas estruturas são acetilcolinesterase humana (hAChE). São elas, 4M0E, 4M0F, 4BDT, 4EY5, 4EY6, 4EY7, 5HF9, 2X8B, 4EY8, 4EY4, 3LII, 1B41, 1F8U, 5HF5, 5HF6, 5HF8, 5HFA, 5FPQ e 4PQE.

As sete primeiras estruturas descritas estão na forma holo, complexadas com ligante, como donepezil (4EY7, veja Figura 3) e galantamina (4EY6), fármacos ativos, entre outros e as demais estruturas estão na forma apo (sem ligante).

Analisando as interações entre ligante e estrutura alvo, nota-se que dentre os vários resíduos de aminoácidos, presentes no sítio ativo, os que interagem são os resíduos de aminoácidos apresentados na Figura 3. Sendo a Tyr337, o Trp286 e o Trp86, os resíduos de aminoácidos que interagem com a maioria dos ligantes. Comparando as estruturas, nota-se uma mudança conformacional considerável na cadeia lateral da Tyr337. A Figura 3 é o sítio ativo da AChE, da estrutura cristalográfica 4EY7, cujo ligante é o Donepezil.

### 1.2.1.1 Hidrólise da Acetilcolina

O sítio ativo da AChE é dividido em duas partes, o Sítio Aniônico Periférico [Trp286, Tyr341, Glu292] e o Sítio Catalítico [Trp86, Tyr133, Tyr337, Phe338] onde está a tríade catalítica [Ser203, His447, Glu202] [29]. No processo de hidrólise da ACh, o neurotransmissor é reconhecido pelos resíduos aromáticos situados entre o sítio aniônico periférico e o sítio catalítico, em seguida a ACh move-se para o fundo do sítio ativo ligando-se a tríade catalítica, onde a AChE hidrolisa a ACh em

acetato e colina (Figura 4).

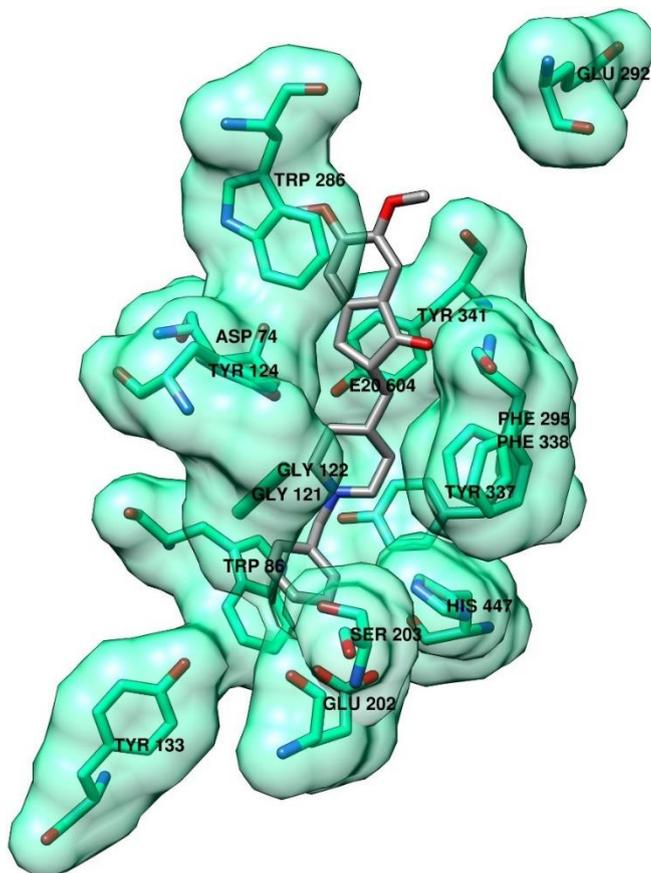


Figura 3: Sítio ativo da AChE em complexo com Donepezil.

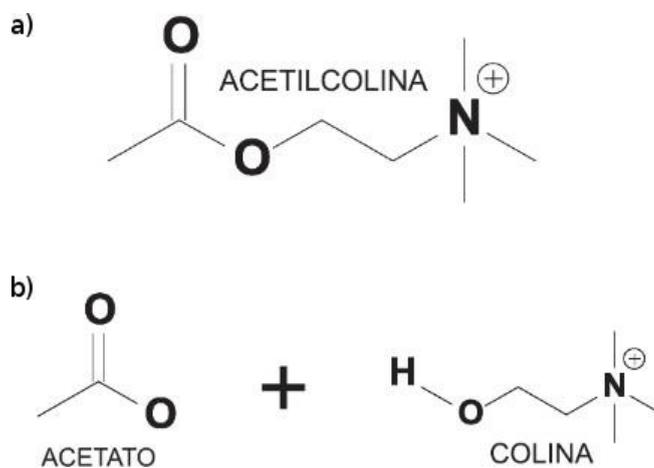


Figura 4: a) Neurotransmissor Acetilcolina b) Acetato e colina após a hidrólise da Acetilcolina.

## 1.2.2 Butirilcolinesterase

Butirilcolinesterase, também conhecida como pseudocolinesterase ou colinesterase não específica é uma enzima colinesterase. A primeira evidencia de sua existência foi por volta de 1932 [30].

Ambas AChE e BuChE mostram uma característica hidrofóbica na superfície da enzima. Contendo os dois sítios, periférico e catalítico. Porém, existem algumas diferenças estruturais em outras partes do sítio ativo, incluindo regiões hidrofóbicas presentes apenas na AChE. No sítio catalítico da AChE tem dois resíduos fenilalanina (PHE295 e PHE297) que na BuChE são substituídas por dois resíduos menores, como a valina e leucina. Esta mudança estrutural cria um espaço maior no fundo da “garganta” permitindo a interação com moléculas maiores. Estudos demonstram que em cérebros de camundongos, aproximadamente 80% da atividade das ChEs é devido a AChE, enquanto que 20% é a BuChE. No entanto, experiências de microdiálise mostram que a BuChE pode complementar a ação da AChE, regulando os níveis de acetilcolina no cérebro [31].

## **2 Objetivos**

### **2.1 Objetivos Gerais**

Este trabalho tem como objetivo a busca por novos agentes químicos com perfil anticolinérgico que possuam ação inibitória esperada, capazes de melhorar as funções cognitivas dos pacientes de DA e reduzir os efeitos colaterais provocados pelos fármacos convencionais.

### **2.2 Objetivos Específicos**

Comparar o desempenho de métodos de triagem virtual baseados na estrutura de ligantes ativos, previamente testados, e na estrutura da proteína acetilcolinesterase. Este processo visa desenvolver uma estratégia para identificação de novos compostos candidatos a inibidores da AChE e propor novos esqueletos base inéditos, com características estruturais apropriadas, a partir de bancos de dados comercialmente disponíveis.

### 3 Metodologias

A busca de novos agentes terapêuticos é um processo longo, de alto custo e envolve diversas áreas do conhecimento. A Química medicinal, subárea da Química, tem como objetivo planejar e produzir compostos químicos usados para a prevenção, o tratamento ou a cura de doenças humanas e animais. A aplicação de métodos computacionais, para a solução dos problemas associados a seleção e otimização de compostos químicos candidatos a fármacos, introduziu dentro da Química Medicinal a Quiminformática, que consiste do uso dos métodos de informática para resolver problemas químicos [32].

Os ensaios virtuais (VS, do inglês, Virtual Screening) são métodos amplamente utilizados, na Quiminformática, para identificar, selecionar e otimizar os ligantes candidatos a compostos líderes. Consistem da aplicação de modelos computacionais capazes de pontuar, classificar ou filtrar compostos, contidos em grandes bancos de dados, que possuam alguma afinidade por um determinado alvo, para realizar ensaios experimentais. Os ensaios virtuais podem ser feitos baseados na estrutura do ligante e são chamados de ensaios virtuais baseados no ligante (LBVS, do inglês, Ligand-Based Virtual Screening), é uma estratégia dependente do princípio da similaridade, a qual diz que moléculas similares devem exibir propriedades biológicas similares. Quando se conhece a estrutura do alvo, ensaios virtuais baseados na estrutura do receptor (SBVS, do inglês, Structure-Based Virtual Screening) são empregados, sendo a docagem e dinâmica molecular os mais utilizados [33].

A docagem molecular é um método para prever corretamente a conformação e orientação de um ligante de acordo com o sítio ativo do alvo, calcular sua energia de interação e classificar os ligantes de acordo com sua afinidade com o alvo macromolecular. Para utilização do método é necessário a estrutura tridimensional do alvo macromolecular escolhido. Neste trabalho utilizamos estruturas da enzima acetilcolinesterase, disponíveis no bando de dados PDB (Protein Data Bank), com parâmetros cristalográficos adequados.

A docagem molecular é dividida em três partes fundamentais, representação do alvo, busca no espaço conformacional do ligante e da proteína e pontuação das possíveis soluções. Na primeira parte, representação do alvo, o alvo e o ligante são tratados como rígidos ou flexíveis. Muitos programas de docagem tratam o alvo como rígido devido ao alto custo computacional relacionado ao cálculo de mudanças conformacionais de uma biomacromolécula, enquanto o ligante é tratado como flexível. Quando se deseja doar moléculas pequenas e macromoléculas, o método mais comumente utilizado é a representação por meio de grades (ou caixas). Onde as informações sobre as contribuições eletrostáticas e de van der Waals em cada ponto de uma grade 3D são armazenadas a fim de representar o sistema macromolecular durante as partes posteriores de docagem. A busca no espaço conformacional do ligante, interagindo na grade, usa diferentes algoritmos, que podem ser divididos em

sistemáticos, estocásticos e de simulação [34]. Todos esses algoritmos são efetivos na busca por conformações de interação de ligantes conhecidos. Cada pose receptor-ligante gerada pelo algoritmo de busca é avaliada de acordo com a energia de interação através da função de pontuação. Esta última parte do processo é dividida em três tipos disponíveis, funções de pontuação baseadas em campo de força, empíricas e baseadas no conhecimento [35].

Quando se busca informações mais detalhadas sobre o comportamento conformacional receptor-ligante a simulação por dinâmica molecular (MD, do inglês, Molecular Dynamics) é uma ótima opção, pois é um método de modelagem computacional que simula o comportamento de uma molécula ou um sistema molecular ao longo do tempo. Porém, o número de processos descritos pela mecânica quântica presentes em movimentos atômicos em um sistema complexo o torna inviável computacionalmente. Sendo assim, o método de MD clássico, usa aproximações simples baseadas na mecânica clássica newtoniana, para simular os movimentos atômicos reduzindo assim a complexidade computacional. O processo se inicia com o modelo atômico, contendo informação da posição de cada um dos átomos do sistema, preparado a partir de ressonância magnética nuclear (RMN), cristalografia, ou dados de modelagem por homologia. As forças que atuam sobre cada um dos átomos do sistema são, em seguida, calculadas a partir de parâmetros que coletivamente são chamados de campo de força [36]. O campo de força determina a força, atuante por todo o sistema em um único átomo, a posição e as velocidades dos átomos, que são calculadas para um intervalo de tempo de um a dois femtossegundos ( $10^{-15}$ s). O processo é repetido, tipicamente milhões de vezes até atingir o tempo de simulação desejado [37]. O campo de força molecular se define por parâmetros e funções potenciais, dependente apenas das posições relativas de cada átomo (energia potencial), que podem ser divididas em duas partes, uma equivalente a porção ligada e a outra a uma porção não ligada. A porção ligada descreve os átomos ligados por ligações covalentes descrevendo a energia associada ao movimento harmônico de distância, ângulo e diedro das ligações. A parte referente ao estado não ligado descreve a atração ou repulsão eletrostática e o potencial de Lennard-Jones calculando as forças de van der Waals [38].

### 3.1 Conjunto de Dados

Três conjuntos de dados (series de compostos) foram usadas para as análises de triagem virtual. No primeiro conjunto de dados, denominado de DS1, cerca de 1600 compostos previamente testados contra a enzima humana acetilcolinesterase (hAChE) foram coletados do banco de dados de bioatividade ChEMBL [39]. Como a finalidade de discriminar entre compostos ativos e inativos aqueles com valores de  $IC_{50}$  menores que  $1 \mu\text{M}$  foram considerados inativos. O segundo conjunto de dados, denominado DS2, foi coletado a partir de um diretório de iscas conhecido como DUD-e (directory of useful decoys enhanced) [40] contendo um conjunto de 453 ligantes conhecidos e testados contra o

alvo estudado e 26234 iscas para a hAChE. A ferramenta DUD-e é amplamente utilizada para validação de métodos de triagem virtual e está disponível em <http://dude.docking.org>. O terceiro e último conjunto de dados, denominado DS3, possui 243 compostos ativos contra a hAChE e 12115 iscas criadas usando o sistema online livre DUD-e. Todos os conjuntos de dados foram então preparados e usados para analisar a habilidade dos métodos de triagem virtual em discriminar entre os compostos ativos e inativos.

## 3.2 Métodos Baseados no Ligante

Com a finalidade de encontrar compostos ativos contra o alvo acetilcolinesterase, abordagens utilizadas para os métodos de triagem virtual baseados no ligante (Ligand Based Virtual Screening, LBVS) são altamente recomendadas [41,42]. Em casos em que a estrutura tridimensional do alvo é desconhecida e os métodos de docagem molecular não são aplicáveis os métodos LBVS realizam uma técnica ativa e promissora para reconhecer novos compostos com atividade biológica semelhante aos ligantes já conhecidos.

### 3.2.1 Similaridade Química

#### 3.2.1.1 Seleção do composto de referência

Buscando identificar compostos com estrutura semelhante comparado a medicamentos ativos e/ou aprovados contra a AChE alguns destes compostos foram usados como estrutura de referência para ambos os métodos de similaridade química, ROCS e EON. Os quatro medicamentos, Donepezil, Tacrina, Galantamina e Rivastigmina e o composto ativo HuprinaW foram considerados como referência no presente trabalho (Figura 5). As estruturas 2D dos compostos de referência foram desenhados usando o programa Marvin Sketch [43] e então as estruturas 3D foram geradas pelo programa OMEGA [44,45] da Openeye. O programa OMEGA é utilizado para gerar a coleção multiconformacional dos compostos moleculares. Inicialmente os compostos são fragmentados e então reconstruídos em diferentes conformações para gerar multiconformações. O algoritmo implementado tenta capturar todos os comprimentos e ângulos de ligação relevantes e conformações do anéis para a molécula de entrada. Os parâmetros utilizados para gerar a coleção multiconformacional foram os padrões, exceto os parâmetros de conformação máxima definido para gerar o número máximo de conformações possíveis ( $maxconfs=1$ ), de definição dos estereocentros não especificados

(*flipper=verdadeiro*) e de não exigir que todos os estereocentros sejam especificados (*stricstereo=falso*).

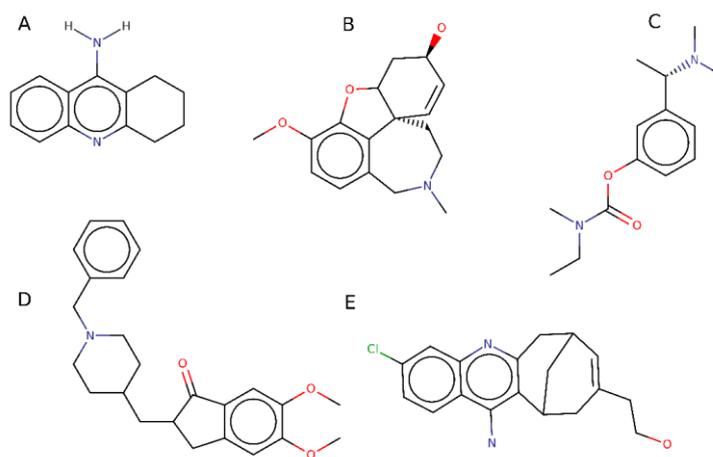


Figura 5: Inibidores hAChE usados como estrutura de referência: (A) Tacrina IC<sub>50</sub> = 0.1  $\mu$ M [46] (B) Galantamina IC<sub>50</sub> = 0.3  $\mu$ M [46] (C) Rivastigmina IC<sub>50</sub> = 0.7  $\mu$ M [46] (D) Donepezil IC<sub>50</sub> = 0.002  $\mu$ M [46] (E) HuprinaW IC<sub>50</sub> = 0,001  $\mu$ M [47].

Para todos os conjuntos de dados as bibliotecas conformacionais necessárias para os programas de similaridade química, ROCS e EON, foram construídas com o programa OMEGA. Os parâmetros padrões foram usados com as seguintes exceções: definição dos estereocentros não especificados (*flipper=verdadeiro*) e não exigir que todos os estereocentros sejam especificados (*stricstereo=falso*). Os arquivos com as todas as conformações possíveis foram usados como arquivos de entrada para o programa ROCS que gerou os arquivos de entrada para o programa EON.

### 3.2.1.1 ROCS e EON

As análises do programa ROCS foram realizadas para o método de superposição baseado na forma tridimensional. O ROCS é um programa bastante utilizado na busca por estruturas tridimensionais similares aos compostos dos bancos de dados. Os parâmetros padrões foram usados com as seguintes exceções: *besthits* (mostra no arquivo de saída os melhores resultados) definido como 10000 para a série DS1, 15000 para a série DS3 e 35000 para a série DS2. No programa ROCS cada conformação dos compostos é alinhada ao composto de referência. Desta forma, os compostos dos conjuntos de base foram ranqueados de acordo com os valores de coeficiente de Tanimoto. O coeficiente de Tanimoto varia entre os valores igual a um, se as duas formas são iguais, e igual a zero se forem totalmente diferentes.

Em complemento à similaridade por forma descrita para o programa ROCS foi utilizado o programa EON para calcular a similaridade eletrostática. O EON usa o coeficiente de Tanimoto Eletrostático baseado no campo eletrostático dos compostos para comparar o potencial eletrostático de

cada composto do conjunto de dados e o composto de referência. Neste caso em especial a equação eletrostática completa de Poisson-Boltzmann [48] é resolvida para calcular o potencial eletrostático dos compostos de referência e dos conjuntos de base. Com esta finalidade o EON calcula novas cargas parciais usando o campo de força MMFF94 e ajusta ambos, composto de referência e compostos dos conjuntos de dados, ao pH neutro. Para garantirmos uma boa qualidade de alinhamento entre os compostos, o programa ROCS gerou o melhor arquivo de entrada para o programa EON. Todos os cálculos feitos pelo programa EON foram realizados com valores padrões dos parâmetros em exceção do *besthits* que foi definido com o mesmo valor usado nos parâmetros do programa ROCS.

### 3.3 Métodos baseados na estrutura do alvo

O método de docagem molecular depende da disponibilidade da estrutura da proteína. Além disso, esta ferramenta implementada na química medicinal nos permite prever o modo de interação entre a proteína e ligante no sítio ativo nos permitindo encontrar a conformação mais favorável do complexo. A estrutura tridimensional do complexo proteína-ligante nos proporciona informações detalhadas a nível atômico, que é amplamente útil para explorar as interações entre o inibidor e a proteína auxiliando nas futuras otimizações do sistema [49-51].

#### 3.3.1 Docagem Molecular

##### 3.3.1.1 Preparação da Proteína

Todas as estruturas cristalográficas utilizadas neste trabalho foram coletadas do banco de dados Protein Data Bank (PDB) [52]. Três estruturas representativas da acetilcolinesterase humana foram selecionadas das 20 estruturas disponíveis no PDB, com os seguintes códigos, 4BDT, 4M0E e 4EY7. Nestas estruturas estão co-cristalizados inibidores conhecidos incluindo HuprinaW (HUW), Dihydratanshione I (1YL) e Donepezil (E20) respectivamente. Com a finalidade de explorar mudanças significativas no sítio ativo da proteína, as estruturas foram escolhidas devido a gama de conformações que o resíduo de aminoácido Tyr337 (na estrutura hAChE) (Figura 6) adota sendo caracterizado pelo RMSD do ângulo de torção  $\chi_1$  igual a 38°. As conformações podem ser atribuídas com base nos seus ângulos de torção  $\chi_1$ . Os valores dos ângulos de diedros  $\chi_1$  são 152°, 164° e 111° para as estruturas 4EY7, 4M0E e 4BDT respectivamente. As mesmas conformações foram encontradas na estrutura tcAChE (*torpedo californica* AChE) onde o resíduo de aminoácido Tyr337 é substituído pela Phe330 que é descrito como um “*swinging gate*” [53].

As estruturas 3D da hAChE usadas nos processos de docagem foram preparadas usando Protein Preparation Wizard [54]. Os resíduos de aminoácidos ionizáveis foram definidos no seu estado normal de ionização a pH 7.4 de acordo com o programa Propka [55]. Durante o processo de preparação, todos os átomos de hidrogênio foram adicionados, todas as moléculas de água e o ligante foram deletados e a energia minimizada usando o Macromodel [56] e campo de força OPLS-2005. Todas as outras configurações foram definidas no modo padrão. As caixas do receptor, região delimitada onde ocorrerá a docagem dos compostos, foram geradas usando o Glide com dimensões de 22x22x22 Å<sup>3</sup>, centrada nas coordenadas do ligante co-cristalizado.

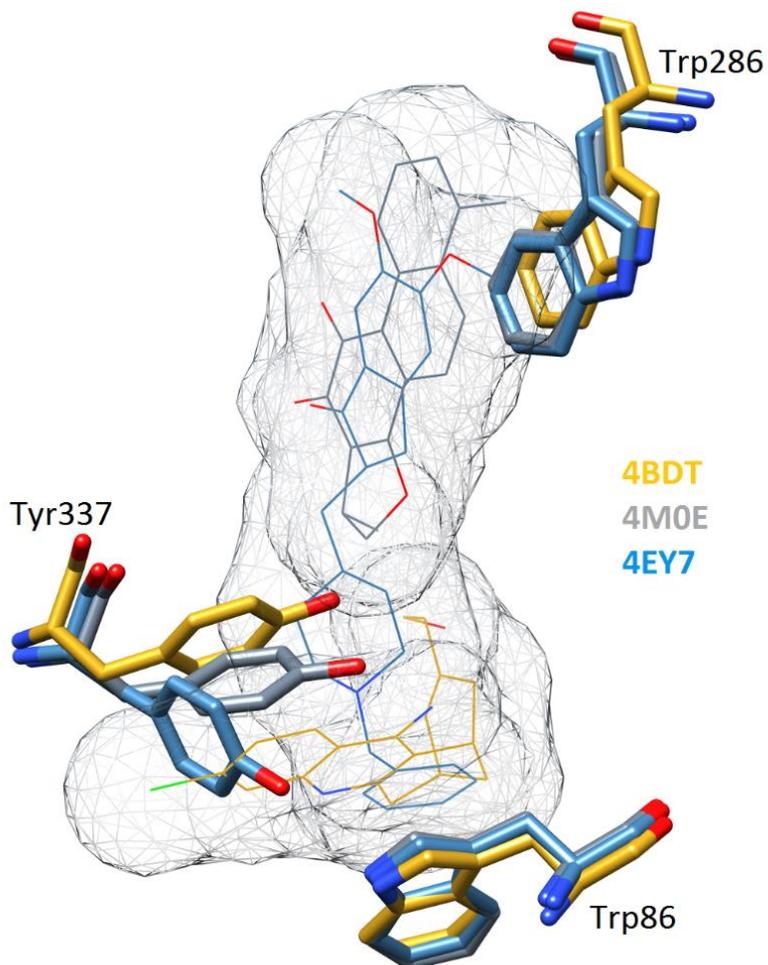


Figura 6: Estruturas cristalograficas da hAChE alinhadas em complexo com a HuprinaW (4BDT) em dourado, Dihydotanshinone I (4MOE) em cinza e Donepezil (4EY7) em azul.

### 3.3.1.2 Preparação do ligante

Para os estudos de docagem, todos os conjuntos de dados foram preparados usando LigPrep2.3 [57], Epik [58] e campo de força OPLS-2005 com IMPACT code [59], gerando todos os

possíveis tautômeros, enantiômeros e estados de protonação a pH  $7,4 \pm 0,5$ , bem como calculando todos os estados de ligação com metais. O mesmo protocolo foi usado para os ligantes co-cristalizados nos processos de redocagem descritos a seguir.

### 3.3.1.3 Processo de Docagem

O processo de docagem molecular foi realizado usando as estruturas cristalográficas da hAChE e os conjuntos de dados DS1, DS2 e DS3, através do programa Glide SP precisão padrão implementado no Schrodinger [60,61]. Os conjunto de dados foram docados considerando o ligante flexível pelo método de amostragem conformacional. Foram 50 poses geradas para cada ligante e apenas 5 foram reportadas. A melhor pose foi determinada pelo melhor valor do coeficiente GlideScore.

## 3.4 Métricas de avaliação dos métodos

A análise utilizada para determinar o desempenho dos diferentes métodos de triagem virtual presentes neste trabalho foi a curva ROC (Receiver Operating Characteristcs). A curva ROC foi gerada classificando os compostos, em ativos e inativos, e valores calculados pela função de pontuação. Se o composto é ativo (positivo) e o método o classifica como positivo (bem ranqueado), então este é contado com uma verdadeiro positivo; se o método o classifica como negativo (baixo ranqueado) este é contado com um falso negativo. Se o composto é inativo (negativo) e o método o classifica com negativo (baixo ranqueado) este é contado com uma verdadeiro negativo; se o método o classifica como positivo (bem ranqueado) este é contado com um falso positivo. A curva ROC descreve a sensibilidade em função da especificidade do método. A sensibilidade do método é definido como a taxa de verdadeiros positivos recuperados, e a especificidade é definida pelos verdadeiros negativos dividido pela soma dos falsos positivos e verdadeiros negativos [62], onde a área sob a curva ROC (AUC) é uma medida quantitativa. Segundo Anighoro e coautores [63], valores acima de 0,5 indicam uma distribuição aleatória, enquanto que os valores maiores indicam a recuperação de compostos ativos nas primeiras posições [64,65]. Esta métrica é uma ferramenta de validação amplamente utilizada para avaliar a qualidade dos protocolos de triagem virtual, monitorando a porcentagem de ligantes ativos recuperados em cada posição do ranking.

## 3.5 Dinâmica Molecular

Simulações de dinâmica molecular (DM) foram utilizadas para elucidar a importância da flexibilidade da proteína na interação com o ligante e explorar as interações entre a hAChE e os compostos que foram selecionados e apresentados no presente trabalho. Apenas as conformações proteína-ligante com melhores pontuações nos resultados de docagem foram submetidas as simulações. As simulações de DM descritas a seguir foram realizadas usando o pacote do Amber 14 e campo de força AMBER FF14sb. Para rodar as simulações, os arquivos de topologia e coordenadas estruturais iniciais foram gerados com o modulo tLEAP do Amber 14. Foi utilizado a técnica comum para inserir o complexo docado dentro de uma caixa octaédrica truncada com moléculas de água do modelo TIP3P. Íons de sódio ( $\text{Na}^+$ ) foram adicionados para neutralizar a carga líquida do sistema proteína-ligante. Antes de iniciar o processo de simulação DM, duas etapas de minimização foram realizadas para a estrutura inicial adquirida no processo de docagem. Primeiramente, as moléculas de água foram minimizadas durante 1000 ciclos iniciais utilizando o método *steepest descent* [66 e 67] seguidos por 1000 ciclos utilizando o método de gradiente conjugado [66 e 67]. Na segunda minimização, 2000 ciclos foram realizados utilizando o método *steepest descent* para todo o sistema, seguido de 1500 ciclos pelo método de gradiente conjugado. Após os ciclos de minimização de energia, o sistema foi aquecido de 0 a 300K durante 1,25 ns em *ensemble* NVT. Na etapa seguinte, após o sistema atingir a temperatura de 300K, realizou-se uma etapa de equilíbrio da densidade durante 7,5 ns. Então, foi realizada a etapa de equilíbrio final do sistema de 7,5 ns. Ambas em *ensemble* NTP. A etapa de simulação do sistema final foi então realizada a temperatura constante de 300 K, durante 100 ns e escalonamento isotrópico de pressão com tempo de relaxação de 1 ps. Em todas as etapas a partir dos ciclos de minimização de energia, o algoritmo SHAKE [67] foi utilizado para tratar todas as ligações com átomos de hidrogênio, um corte de 10 Å foi usado para todas as distâncias de interações entre átomos não ligados e a dinâmica de Langevin [67 e 68] com frequência de choque de 2 ps<sup>-1</sup> foi adotada para controle de temperatura. O tempo de simulação DM foi dividido em 10 corridas de 10 ns com intervalo de 2 fs e a cada 25000 passos os dados de trajetória eram salvos. Assim, 200 conformações (frames) foram obtidas a cada 10 ns de simulação.

### 3.5.1 Análises de Triplet Clique

Para cada resíduo de aminoácido da proteína a área de superfície excluída do solvente (do inglês, solvent-excluded surface area (relSESA)) foi calculada usando o UCSF Chimera, programa para visualização e análise interativa de estruturas moleculares [69], e comparada a relSESA de cada resíduo de tripeptídeo padrão em Gly-X-Gly. Apenas resíduos hidrofóbicos (Ala, Val, Ile, Leu, Tyr, Phe e

Trp) com relSESA abaixo de 0,3 foram considerados para as próximas análises. Nas análises de simulação, uma conformação a cada 2 ns foi considerada para o cálculo dos valores de relSESA e os resíduos hidrofóbicos que ficaram com o valor médio abaixo de 0,3 durante toda a simulação foram submetidos a análise de Clique.

Um conjunto de três resíduos em contato um com o outro é definido como um Clique. Foi considerado como contato os resíduos que apresentaram as menores distâncias interatômicas abaixo de 3,8 Å para interação de átomos pesados na cadeia principal, como está descrito por Basu e coautores [70,71]. Vetores internos para cada resíduo, assim como vetores de clique, foram definidos como descrito por Basu e coautores [71] para as conformações obtidas a partir dos resultados de simulação DM do presente trabalho. No caso dos cliques formados pelo ligante, foi considerado, para definir os vetores, o anel aromático localizado mais próximo dos outros dois resíduos presentes no clique. Os parâmetros, *tilt* e *swivel angles*, usados para analisar a estabilidade dos cliques, descrevem respectivamente a orientação de cada resíduo no plano do clique e a rotação do resíduo em relação ao seu eixo interno Z [70,71].

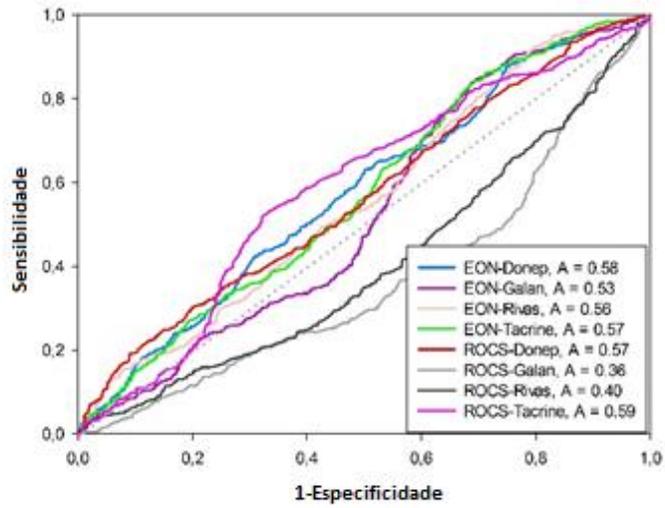
## 4 Resultados e Discussões - Validação do protocolo de seleção

A partir de métodos em Quiminformática busca-se identificar inibidores da enzima acetilcolinesterase. Uma vez que a estrutura tridimensional da enzima acetilcolinesterase e seus inibidores são conhecidos, utilizou-se métodos de ensaios virtuais baseados nos ligantes e na estrutura do alvo (LBVS e SBVS).

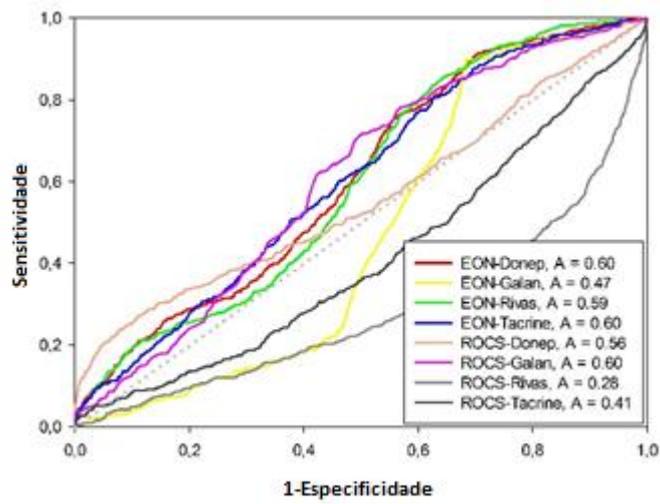
### 4.1 ROCS e EON

Duas estratégias computacionais foram analisadas nesta etapa: análises dos métodos de similaridade baseado na forma denominado de ROCS e similaridade eletrostática denominado EON. Mais especificamente, ROCS é usado para identificar a semelhança entre moléculas baseando-se em suas formas tridimensionais, enquanto que o método EON compara o mapa de potencial eletrostático das moléculas pré-alinhadas. A métrica de Tanimoto baseado na forma e potencial eletrostático dos pacotes ROCS e EON foram testadas para todos os conjuntos de dados e compostos de referência descritos no capítulo de metodologia (veja Figura 5). Uma vez que os resultados dos métodos ROCS e EON são altamente dependentes da molécula de referência adotada, análises com quatro diferentes moléculas de referência foram realizadas para testar a consistência do desempenho dos programas. A área sob a curva ROC (AUC) variou consideravelmente em relação as moléculas de referência e os programas usados. Entretanto, o melhor desempenho em ambos os métodos mostrou uma ligeira diferença de AUC (0,60 e 0,70, para ROCS e EON, respectivamente). Além disso, considerando ambos os programas usados e todos os conjuntos de dados o melhor resultado foi alcançado pelo composto de referência Donepezil com AUC superior a 0,57 em todas as curvas ROC.

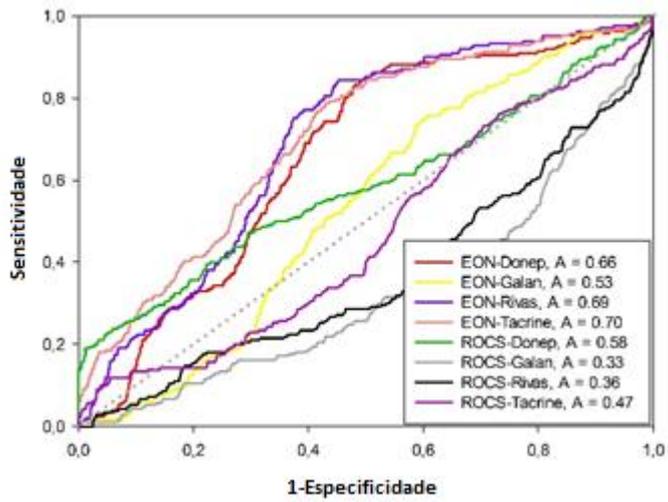
A busca por similaridade baseado no potencial eletrostático, método EON, mostrou desempenho adequado para todas as moléculas de referência, principalmente para o Donepezil e Tacrina. O resultado mostrado pelo uso de busca por similaridade 3D suporta o uso do método EON em projetos baseados em métodos de ensaio virtual que visam a busca por inibidores da acetilcolinesterase.



(a)



(b)



(c)

Figura 7: Curvas ROC dos programas ROCS e EON para os conjuntos de dados (a) DS1, (b) DS2 and (c) DS3.

## 4.2 Docagem Molecular

### 4.2.1 Procedimento de Redocagem Molecular

Com o intuito de avaliar a capacidade de reprodução das estruturas cristalográficas, foram realizadas etapas de redocagem para as três estruturas utilizadas neste trabalho. No processo de redocagem, cada ligante é redocado em sua estrutura cristalográfica, a fim de reproduzir o modo de interação dos ligantes na estrutura cristalográfica e validar o tipo de docagem. O modo de interação foi validado para os seguintes ligantes co-cristalizados em suas respectivas estruturas cristalográficas: Donepezil (E20), HuprinaW (HUW) e Dihydratanshinone I (1YL), nas estruturas 4EY7, 4BDT e 4M0E.

Nas simulações de redocagem é esperado que os resultados gerem desvio quadrático médio (do inglês, root mean square deviation, RMSD) melhores que 2 Å quando comparado às conformações de interação dos ligantes originalmente co-cristalizados [60].

Verificou-se que os ligantes ligaram ao sítio ativo da proteína alvo hAChE semelhantemente aos ligantes co-cristalizados (veja Figura 8). Dessa forma, o resultado das redocagens mostraram que o protocolo usado pelo programa Glide foi capaz de reproduzir as poses adotadas pelas estruturas 4BDT, 4EY7 e 4M0E com RMSD dos átomos iguais a 0, 89 Å, 1, 34 Å e 0,82 Å. A função de pontuação do programa Glide para os complexos 4EY7 e 4M0E ficaram bem próximas com valores de -11,65 kcal/mol e -8,52 kcal/mol comparadas aos valores de afinidade experimentais de  $\Delta G_{4EY7} = -11,86$  kcal/mol (usando o valor de  $IC_{50} = 0,0002 \mu M$ ) e  $\Delta G_{4M0E} = -8,52$  kcal/mol (usando o valor de  $IC_{50} = 1 \mu M$ ). Para a estrutura cristalográfica 4BDT o valor de afinidade experimental é  $\Delta G_{4BDT} = -12,27$  kcal/mol (usando o valor de  $IC_{50} = 0,001 \mu M$ ) que é ligeiramente maior comparado ao valor de -10,73 kcal/mol alcançado pelo programa Glide. Os valores de  $\Delta G$  foram calculados usando os valores de afinidade  $IC_{50}$  encontrados experimentalmente em  $\Delta G = RT \ln (IC_{50})$ . Portanto, os valores da função de pontuação dos ligantes redocadas estão em perfeito acordo como os valores de afinidade experimental. Este resultado valida a ideia de que a predição dos modos de ligação das estruturas cristalográficas pelo protocolo de docagem aplicado é adequado para estudos de ensaios virtuais. A superposição entre os ligantes co-cristalizados e as conformações preditas pela redocagem são mostrados na Figura 8.

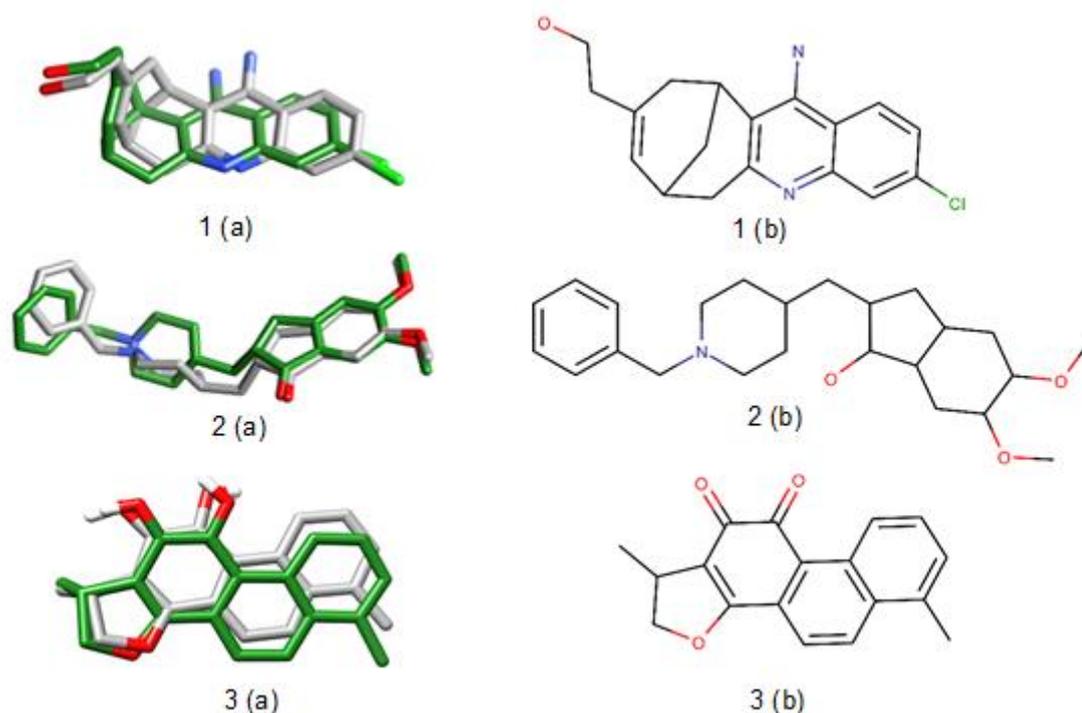
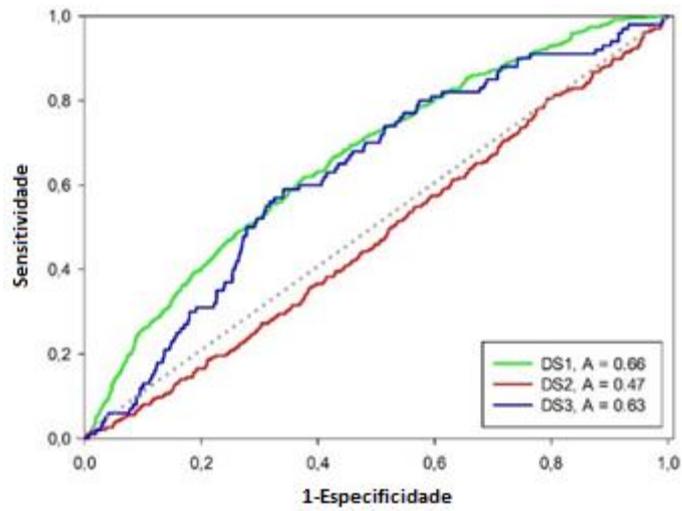


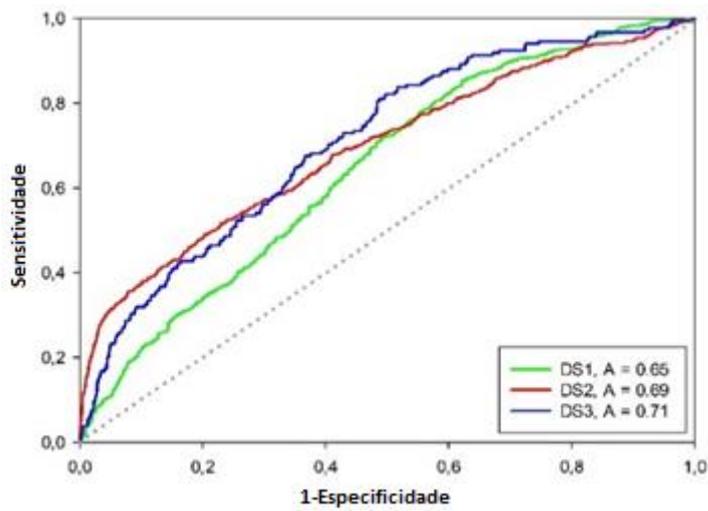
Figura 6. Redocagem dos ligantes co-cristalizados. (a) Superposição entre os ligantes co-cristalizados (cinza) e as melhores poses previstas pela redocagem (verde) para (1) HuPrina W (0.89Å), (2) Donepezil (1.34Å) e (3) Dihydratanthone I (0.82Å) e (b) suas estruturas moleculares.

#### 4.2.2 Análises da Docagem Molecular

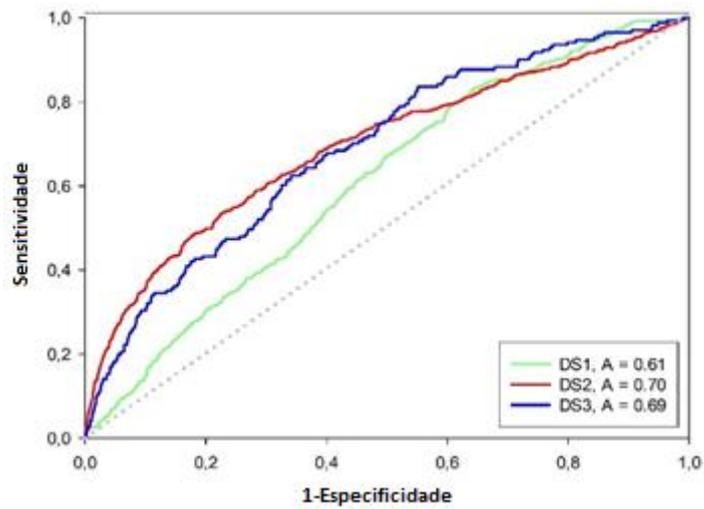
Nos estudos de docagem molecular, o mesmo protocolo, previamente validado na etapa de redocagem, foi definido para discriminar entre compostos ativos e inativos de todos os conjuntos de dados (DS1, DS2 e DS3) em três estruturas diferentes (4EY7, 4BDT e 4M0E). Após obter os modos de interação dos compostos no sítio ativo estes foram ranqueados de acordo com os valores da função de pontuação do Glide, apenas os melhores valores foram considerados para cada composto. As curvas ROC correspondentes podem ser verificadas na Figura 9. O programa de docagem Glide teve um bom desempenho alcançando o valor de AUC de 0,71. A estrutura 4EY7 foi incapaz de recuperar os compostos ativos para o conjunto de dados DS2, alcançando o valor de AUC igual a 0,47. Para os outros casos, os resultados mostraram consistência de desempenho apropriada, com as estruturas 4BDT e 4M0E que apresentaram recuperação esperada dos compostos ativos acima da recuperação aleatória.



(a)



(b)



(c)

Figura 9: Curvas ROC do programa de docagem molecular Glide para as estruturas cristalográficas (a) 4EY7, (b) 4BDT e (c) 4MOE.

Os resultados desta primeira etapa de docagem sugerem que as conformações e modos de ligação dos inibidores da hAChE foram satisfatórias. Uma mesma análise com restrição de *smarts* e dimensão da caixa de docagem de 6X6X6 Å<sup>3</sup> foram feitas, entretanto não houve melhora nos dados de recuperação.

Após esta validação exaustiva das diferentes metodologias de ensaios virtuais, sendo elas, busca por similaridade química, ROCS e EON, e docagem molecular, encontramos que busca por similaridade eletrostática, EON, sobressai quando comparado com os resultados de busca por similaridade por forma. Os resultados também revelaram que as estruturas 4BDT e 4M0E alcançaram melhor desempenho no método de docagem, baseado na estrutura do receptor, para inibidores da hAChE. É bastante conhecido que o Donepezil é um inibidor dual, isto é, interage simultaneamente nos sítios de interação catalítico (CAS), periférico (PAS) e no meio da gorge (entre o CAS e o PAS) (veja Figura 10).

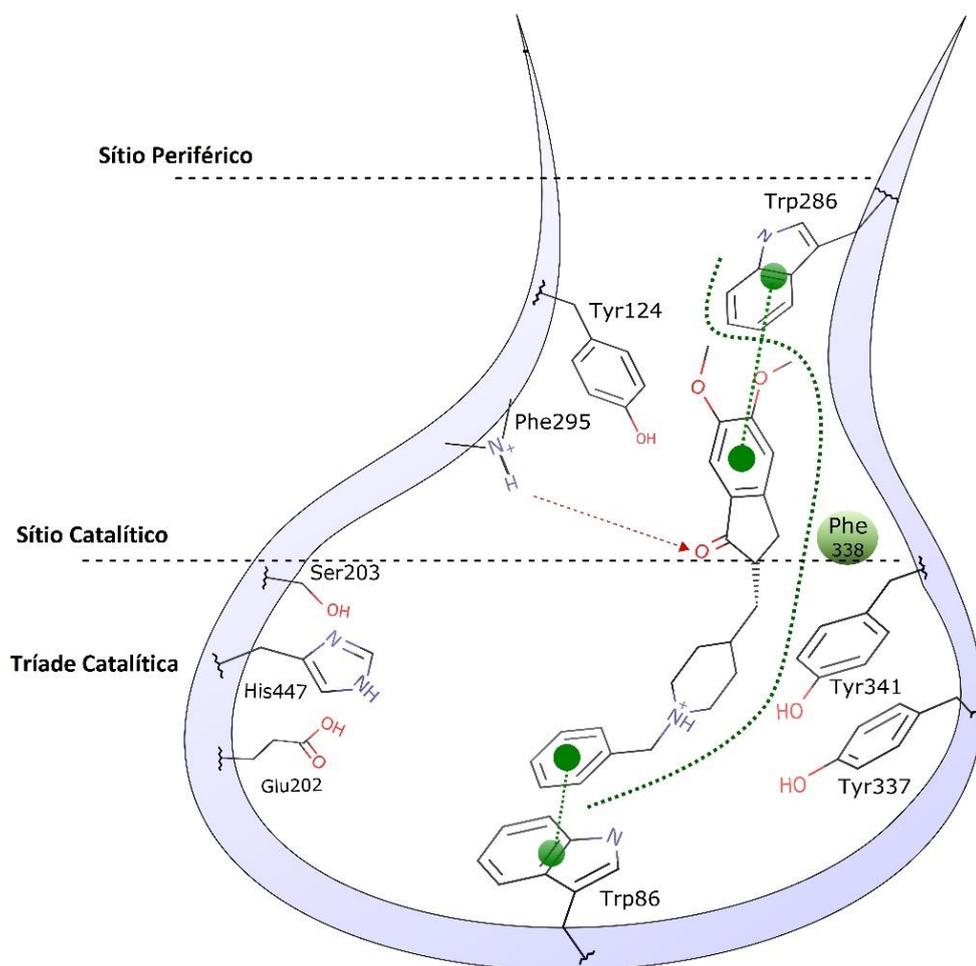


Figura 10: Representação bidimensional das interações entre a estrutura da hAChE e o Donepezil.

Em mais detalhes, o fármaco Donepezil tem algumas principais interações, sendo elas,  $\pi$ -stacking com os resíduos de aminoácidos Trp286 e Trp86, representadas por linhas pontilhadas verde com círculos na extremidade indicando o anel da interação, hidrofóbicas com os resíduos Tyr341,

Tyr337 e Phe338, representado por linhas pontilhadas verdes, e ligação de hidrogênio com o resíduo de aminoácido Phe295 representado por linhas vermelhas pontilhadas. Portanto, o Donepezil permite identificar candidatos a inibidores da doença de Alzheimer mais efetivos e promissores. Sendo assim, sua estrutura foi usada como referência para as próximas etapas. Além disso, a estrutura cristalográfica 4M0E, com melhor resolução e resultados no processo de docagem, foi selecionada para a busca por novos inibidores da hAChE.

## 5 Resultados e Discussões – Seleção dos compostos

### 5.1 Ensaios virtuais na busca por novos esqueletos base inibidores da hAChE

Procedimentos hierárquicos de ensaios virtuais foram aplicados para identificar esqueletos base inibidores da hAChE. Resumidamente, o procedimento consiste de: Filtro dos ligantes usando busca por similaridade eletrostática; predição das poses de ligação usando o programa Glide; e ranqueamento dos compostos docados usando a função pontuação GlideScore. O procedimento dos ensaios virtuais está sumarizado na Figura 11.



Figura 11: Protocolo hierárquico de ensaios virtuais usados para identificar inibidores da hAChE.

## 5.1.1 Filtros por ensaios virtuais

Primeiramente, as conformações da molécula de referência Donepezil da estrutura cristalográfica 4EY7 e dos compostos do banco de dados ZINC foram geradas para a busca de similaridade eletrostática como foi descrito anteriormente. O método EON foi realizado para os 10 milhões de compostos disponíveis comercialmente do banco de dados ZINC. Os compostos foram ranqueados de acordo com o valor de Tanimoto eletrostático, onde 1 milhão de compostos foram selecionados para as análises seguintes com pontuação entre 0,60 e 0,94.

Os compostos remanescentes após a aplicação do filtro de similaridade eletrostática, foram submetidos ao método de docagem molecular usando o programa Glide, com a finalidade de seleção das moléculas mais promissoras para as etapas de dinâmica molecular e ensaios bioquímicos. O ligante co-cristalizado foi retirado da estrutura 4M0E e a sub-coleção foi docada no sítio ativo do alvo. As moléculas melhores classificadas pelo valor de GlideScore foram submetidas a uma inspeção visual e busca em bancos de dados para assegurar o ineditismo das moléculas selecionadas como inibidores hAChE. Os critérios usados para seleção dos melhores compostos na inspeção visual foram: a ocupação do sítio ativo, interação com alguns dos principais resíduos de aminoácidos, Trp286, Trp86, Tyr337, Phe295 e com os resíduos de aminoácidos da tríade catalítica, Ser203, His447 e Glu 202, preferência para compostos com interação dual do sítio e quantidade de ligação de hidrogênio. Sete compostos foram selecionados e foram ainda avaliados usando simulações MD. Uma representação bidimensional das estruturas é mostrada na Figura 12.

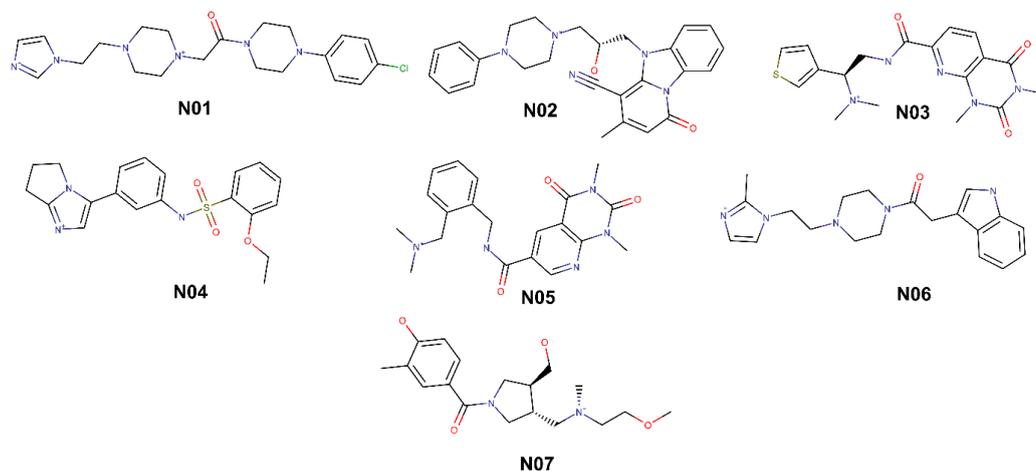


Figura 12: Representação das estruturas propostas como novos potentes candidatos a inibidores hAChE.

A Figura 13 apresenta o resultado de docagem molecular do composto N01 em 4M0E para explorar as interações entre o ligante e o sítio ativo da enzima hAChE. O modo de interação sugere que alguns resíduos de aminoácidos principais do sítio de ligação interage com o composto fazendo ligações de hidrogênio e interações do tipo  $\pi$ -stacking e cation- $\pi$ . Os seguintes átomos fazem ligação de hidrogênio: o átomo de oxigênio com o resíduo de aminoácido Phe295 a uma distância de 2,68 Å, o

átomo de nitrogênio do imidazol com o resíduo de aminoácido His447 a uma distância de 1,92 Å e o átomo de nitrogênio da piperazina com o resíduo de aminoácido Tyr124 a uma distância de 2,15 Å. Os anéis aromáticos da cadeia lateral do Trp286 e Trp86 formaram interação do tipo  $\pi$ - $\pi$  como o imidazol e benzeno do ligante. Pode-se observar que as duas interações  $\pi$ -stacking e a ligação de hidrogênio que a molécula de referência Donepezil faz com o sítio ativo da hAChE, representado pela Figura 10, também ocorre entre o composto N01 e os mesmos resíduos de aminoácidos.

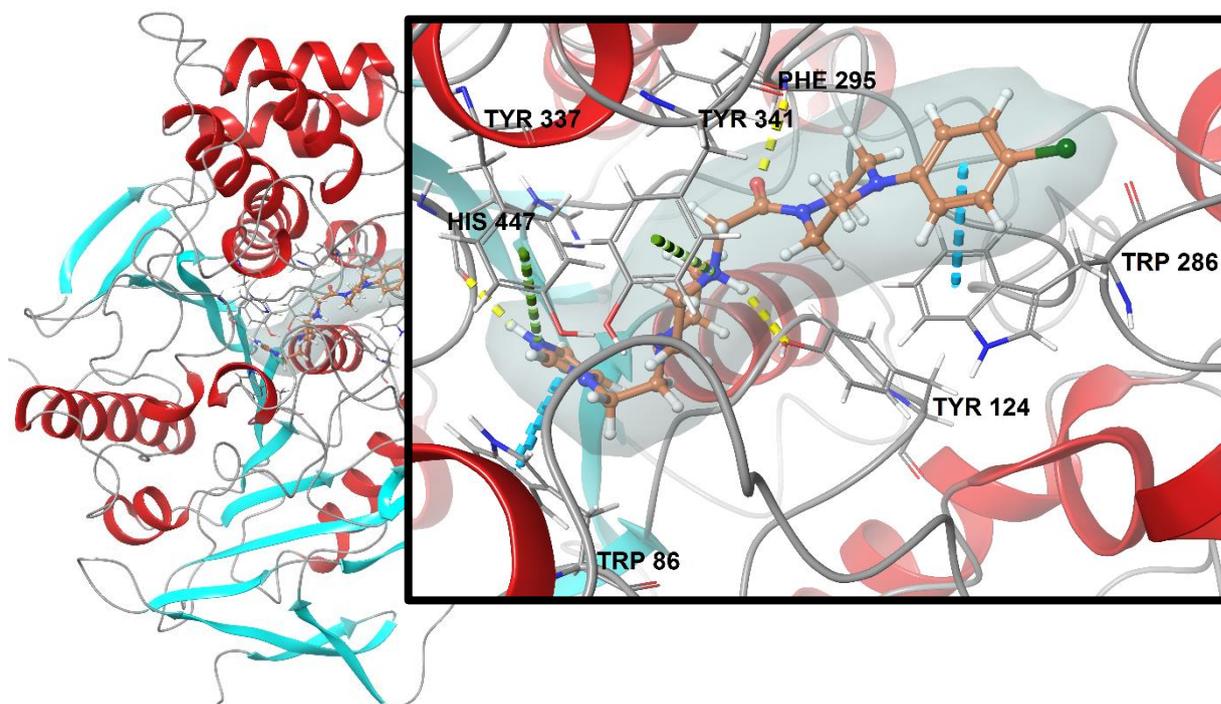


Figura 13: Resultado de docagem do composto N01 na estrutura cristalográfica hAChE 4M0E. Ligações de hidrogênio estão representadas por linhas tracejadas amarelas; interações do tipo  $\pi$ -stacking estão representadas por linhas tracejadas azuis; interações do tipo Cátion- $\pi$  estão representadas por linhas tracejadas verdes,

Analisando as interações a partir dos resultados de docagem molecular (ver Figuras 14 e 15) observa-se que o modo de interação predito pelas poses de docagem dos compostos selecionados na inspeção visual foi semelhante as poses das estruturas cristalográficas dos sete inibidores disponíveis no banco de dados PDB. Alguns resíduos de aminoácidos foram destacados como principais, como Ser293, His 447, Tyr 124 e Tyr 133 que formam ligação de hidrogênio. Interações do tipo cátion- $\pi$  e  $\pi$ -stacking foram observadas entre os ligantes e os resíduos Tyr337, His447, Trp286, Tyr341, Phe338, Trp86 e Tyr124. Além disto, foi observado que os resíduos Phe295, Tyr337, Tyr124, Trp286, trp86 e Try 341 desempenham um papel crucial interagindo com a maioria destes compostos. Outro aspecto importante a ressaltar é que o sítio catalítico da proteína hAChE possui características hidrofóbicas, esse tipo de interação está presente em todos os compostos analisados, embora não estejam representadas nas figuras, foi considerado como aspecto positivo na inspeção visual. O modo de ligação predito pela simulação de docagem molecular, para estes sete compostos selecionados, sugere-os como potenciais candidatos a inibidores da hAChE.

Os critérios geométricos por trás da atribuição das interações relatadas são: Para as ligações de hidrogênio a distância máxima é 2,8 Å, o ângulo mínimo do doador é 120° enquanto que o ângulo mínimo de aceitador é 90°. Para as interações do tipo  $\pi$ -stacking o ângulo entre os planos dos anéis deve ser menor que 30° e a distância entre o centroide dos anéis deve ser menor que 4,4 Å, para o caso *face to face*, ou o ângulo entre os planos dos anéis deve estar entre 60° a 120° e a distância entre o centroide dos anéis deve ser menor que 5,5 Å, para o caso, *edge to face*. Para as interações do tipo cátion- $\pi$  a distância máxima entre o centro do cátion e o centro do anel é de 6,6 Å e o ângulo entre o plano do anel e a linha entre o centro do cátion e o centro do anel não pode desviar em mais de 30° na perpendicular.

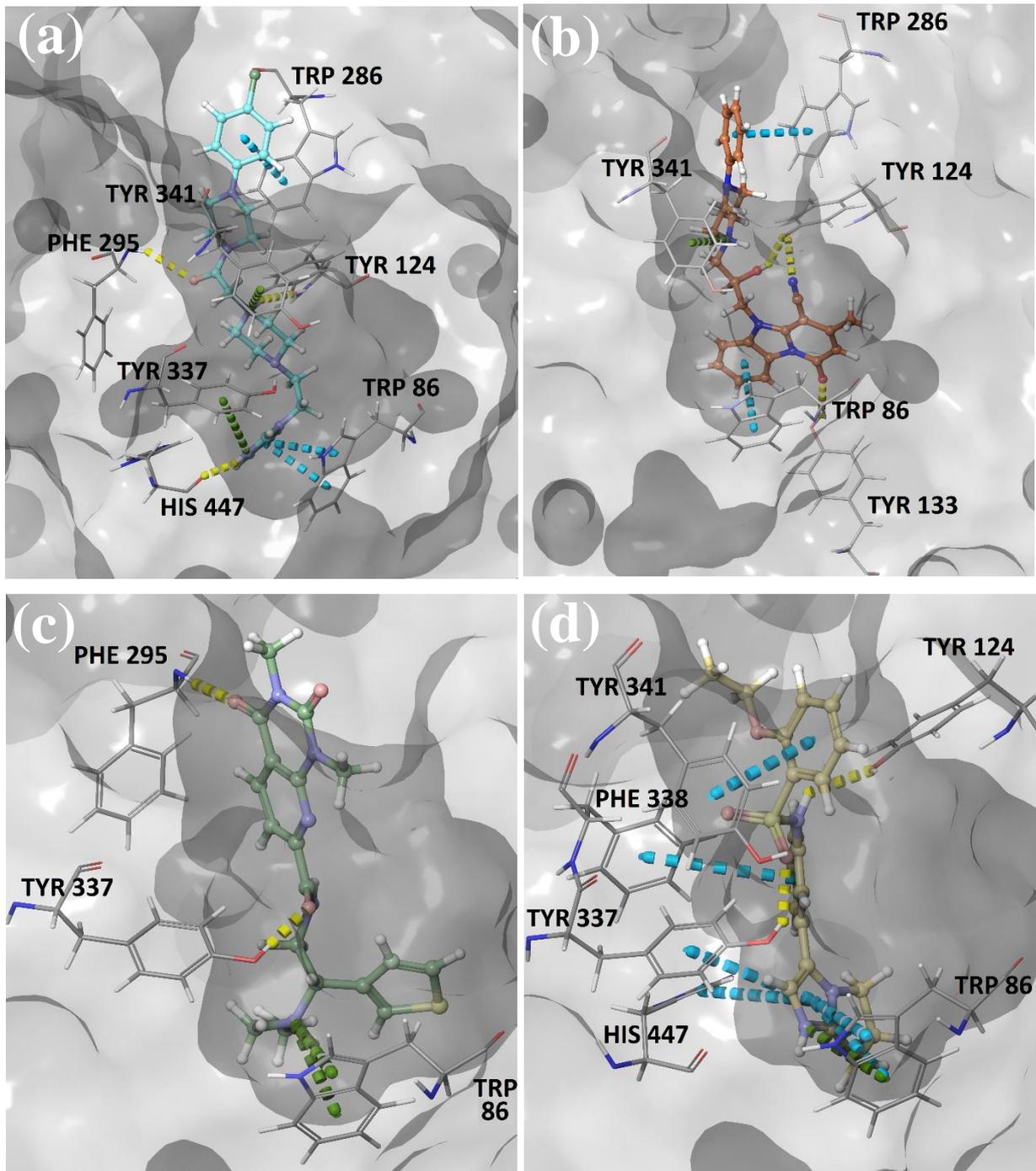


Figura 14: Estrutura dos complexos preditos por docagem molecular com modo de interação semelhante a inibidores co-cristalizados selecionados na análise visual (a) N01 (b) N02 (c) N03 (d) N04.

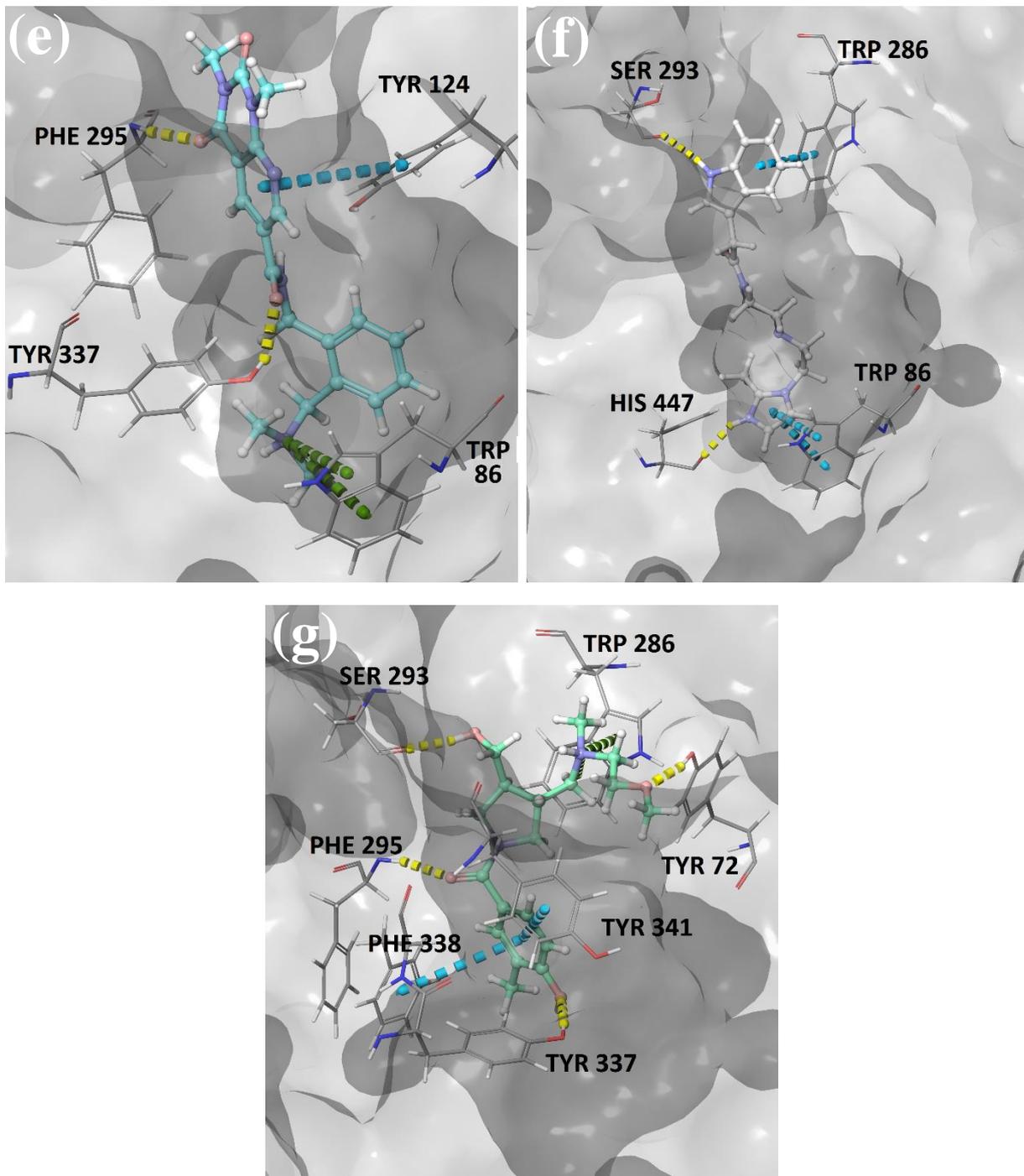


Figura 15: Estrutura dos complexos preditos por docagem molecular com modo de interação semelhante a inibidores co-cristalizados selecionados na análise visual (e) N05 (f) N06 (g) N07.

## 5.1.2 Dinâmica Molecular

Com a finalidade de investigar ainda mais esses compostos filtrados nas etapas anteriores, 100 ns de simulação por dinâmica molecular foi realizada para cada um dos sete ligantes complexados com a estrutura da hAChE. Quanto ao comportamento dinâmico do composto N01, mais promissor da série, foi observado que o ligante permanece estável dentro do sítio ativo da enzima. Foi também observado as interações do ligante com os resíduos Trp286, Trp86, Tyr337, Tyr341, His447, Tyr124 e Phe295, como pode ser analisado na sobreposição das conformações na Figura 16. Estas observações vão de encontro com os resultados de docagem molecular.

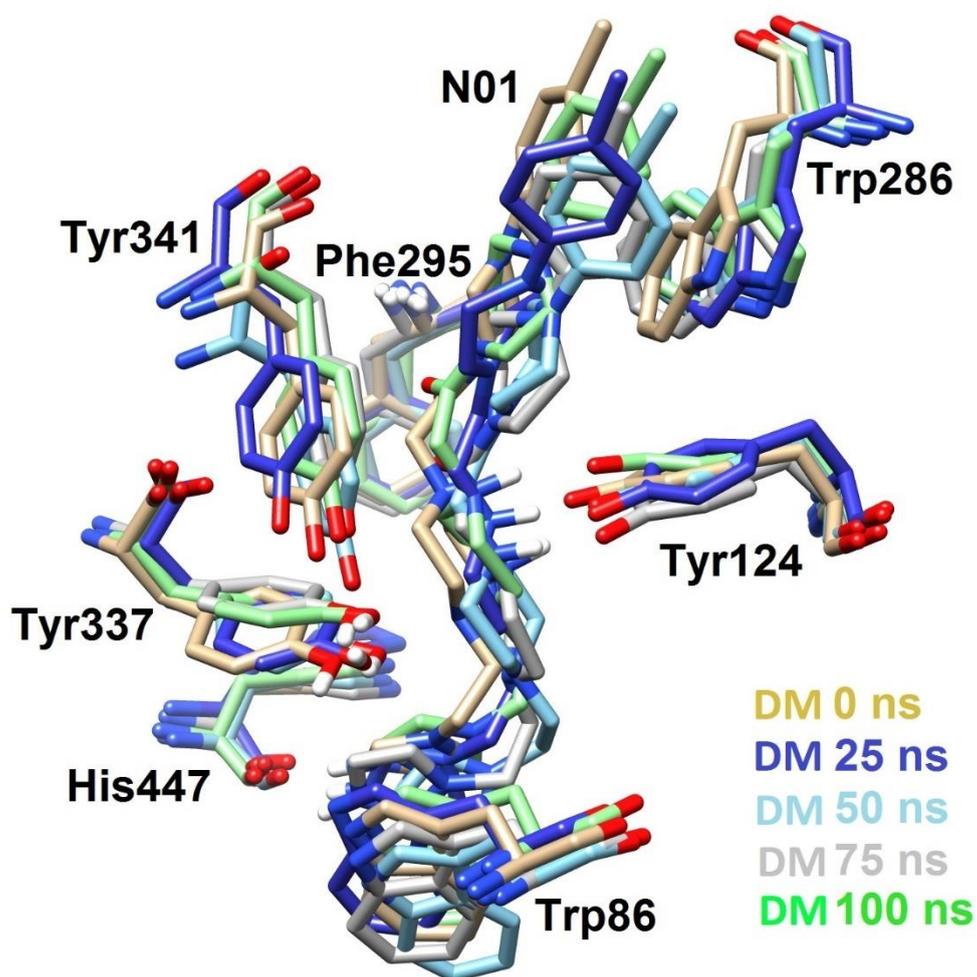


Figura 16: Resultado da docagem molecular do composto N01 no sítio ativo da hAChE em tempos diferentes da simulação por dinâmica molecular: 0 ns, 25 ns, 50 ns, 75 ns e 100 ns.

Para explorar a estabilidade dos ligantes no decorrer da dinâmica do sistema, os valores de RMSD dependentes do tempo dos carbonos  $\alpha$  foram calculados durante as trajetórias da DM nos 100

ns. Os resultados de RMSD mostraram que os compostos ficaram estáveis após 55 ns de simulação para todos os casos. No gráfico do RMSD versus o tempo, apresentado na Figura 17, também é possível observar a conformação estável dos ligantes que não mostraram valor de RMSD superior a 6 Å durante todo o tempo de simulação. Os resultados observados para o sistema indicam instabilidade durante o tempo de simulação DM com valores médios de RMSD de 1,8 Å a 2,1 Å. Outro fato importante a se destacar foi a otimização das interações entre o alvo hACHe e os compostos através de algumas mudanças principais que ocorreram com a conformação relatando estabilidade.

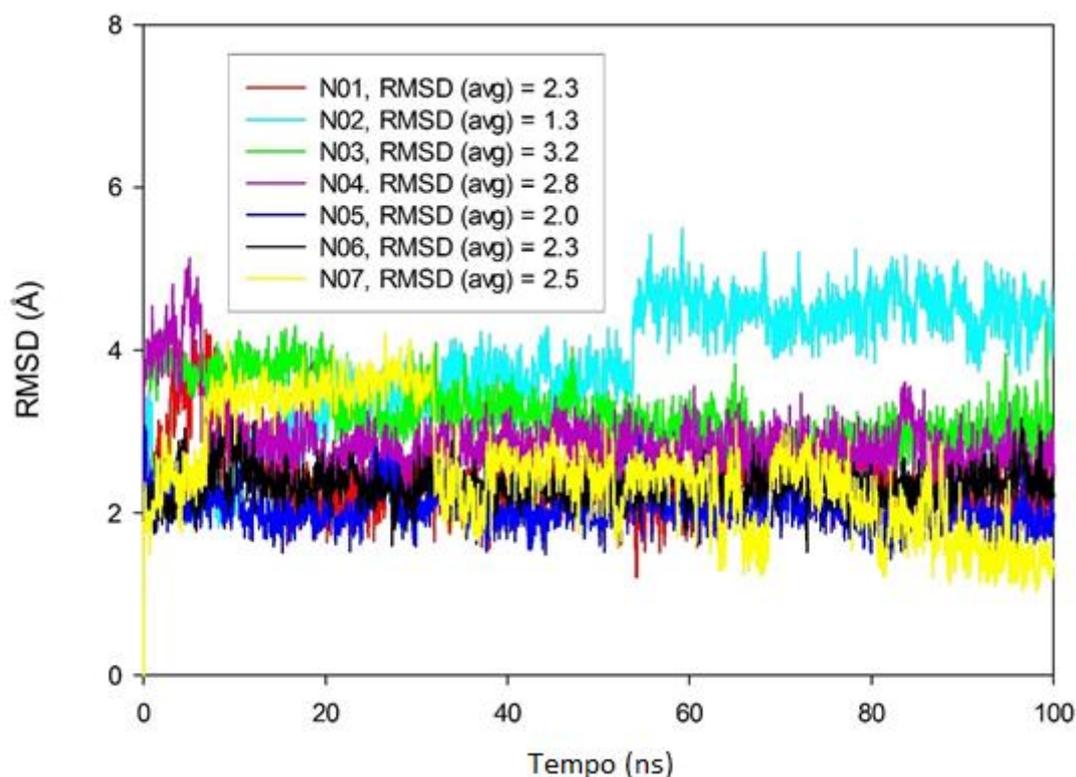


Figura 17:RMSD dos compostos versus tempo de simulação de dinâmica molecular.

### 5.1.3 Análise de triplet clique e caracterização do modo de ligação em termos de parâmetros geométricos

A presença dos cliques foi cuidadosamente investigada para caracterizar a associação entre os resíduos aromáticos enterrados presentes no sítio ativo enzimático e os anéis dos ligantes em termos de restrições geométricas impostas durante a formação do complexo binário. Todos os sete ligantes propostos no presente trabalho possuem um “clique” com resíduos enterrados da proteína, pelo menos em alguma parte da simulação. Enquanto a maioria dos ligantes fizeram parte de cliques efêmeros, em dois casos eles apresentaram boa estabilidade. O primeiro deles, molécula N05, mostrou a presença do clique com os resíduos Trp83 e Tyr334 em 80% dos passos da DM, evidenciando uma boa

interação aromática do ligante com estes resíduos. No segundo ligante, uma mudança conformacional da molécula N02 leva a formação do clique com os mesmos resíduos citados para o composto N05. Neste caso, o clique aparece durante a simulação e está presente em 68% da última metade dos passos da simulação.

Uma análise mais detalhada foi feita para o complexo entre hAChE e o composto N05 para descrever os padrões geométricos da interação entre a proteína e o ligante, veja Figura 18. A união dos pontos de origem de cada resíduo define um triângulo Trp83-Tyr334-N05 de lados 6,4, 4,7, e 7,5 Å e ângulos de  $85^\circ \pm 4$ ,  $38^\circ \pm 3$  e  $57^\circ \pm 3^\circ$ .

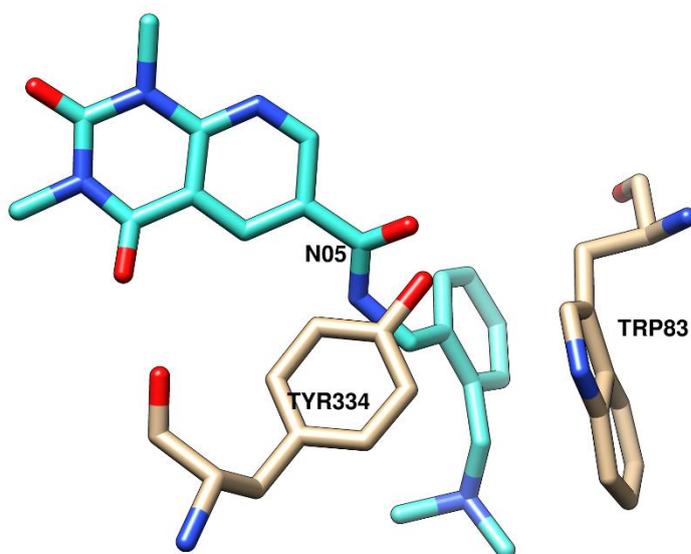


Figura 18: O complexo hAChE-N05 e os padrões geométricos da interação proteína-ligante. A união do ponto de origem de cada resíduo define um triângulo Trp83-Tyr334-N05 de lados 6,4, 4,7, e 7,5 Å e ângulos de  $85^\circ \pm 4$ ,  $38^\circ \pm 3$  e  $57^\circ \pm 3^\circ$ .

Foi observado durante a simulação uma mudança conformacional da Tyr334 caracterizada por modificações de ângulo de inclinação (*tilt angle*) de  $60^\circ$  a  $125^\circ$ . Este mesmo parâmetro foi usado para mostrar um pequeno desvio de  $6^\circ$  da cadeia lateral do resíduo Trp83 cuja conformação mais estável apresentou ângulo de  $66,1^\circ$ . Por outro lado, observa-se através deste parâmetro uma orientação aleatória do anel aromático do ligante N05 localizado no fundo do sítio ativo da hAChE durante a simulação. Este fato pode indicar uma falta de forte interação aromáticas entre o ligante N05 e os resíduos tirosina e triptofano. Um comportamento semelhante é observado para o ângulo de rotação, que permaneceu estável a  $63^\circ$  e  $170^\circ$  durante a simulação para os resíduos Trp83 e Tyr334 e apresentou maior flutuação para o ligante N05.

## 6 Considerações Finais e Perspectivas

Os resultados aqui apresentados são decorrentes do estudo teórico-computacional envolvendo a enzima acetilcolinesterase, que foi escolhida por ser um dos principais alvos contra a doença de Alzheimer descrito pela hipótese colinérgica.

A busca de novos esqueletos base inibidores da enzima acetilcolinesterase foi realizada utilizando métodos de ensaios virtuais baseados tanto na estrutura do ligante quanto na estrutura do receptor. Foram feitas buscas por similaridade química por forma e eletrostática, utilizando diferentes inibidores como moléculas de referência e o método de docagem molecular para predizer e pontuar o modo de interação dos complexos gerados. Entre os métodos estudados, os compostos de referência e as estruturas usadas, os que mais se destacaram na recuperação de compostos ativos a partir de um banco de moléculas ativas e inativas foram o método de similaridade eletrostática, Donepezil como estrutura de referência e a estrutura hAChE 4MOE.

A integração dos métodos de similaridade eletrostática e docagem molecular foi abordada através da elaboração de um protocolo de busca que se iniciou filtrando os compostos mais semelhantes ao Donepezil a partir do banco de dados comerciais ZINC com cerca de 10 milhões de estruturas resultando em 1 milhão de compostos. Subsequentemente, o processo de docagem molecular desta sub-coleção permitiu uma inspeção visual, das moléculas melhores listadas pela função pontuação, focada em moléculas com características apropriadas para uma boa interação molecular com a enzima acetilcolinesterase.

A inspeção visual das moléculas melhores pontuadas permitiu a seleção de 7 moléculas que foram avaliadas nas etapas de docagem e dinâmica molecular. As simulações de docagem e dinâmica molecular foram usadas para explorar as interações entre a hAChE e estes compostos. Alguns resíduos principais tal como, Phe295, Tyr 124, Trp286, Trp86 e Tyr341 presentes no sítio ativo da enzima foram identificados como cruciais na recuperação de inibidores duais promissores a candidatos anti-Alzheimer. Adicionalmente, os estudos mostraram que todos os sete candidatos aqui propostos apresentam no mínimo três das principais interações presentes na maioria dos inibidores já conhecidos contra a hAChE tornando os fortes candidatos a inibidores.

Os resultados sobre dinâmica molecular descritos acima sugerem que as trajetórias das interações, presentes no sítio ativo, no decorrer do tempo validaram os resultados apresentados na docagem molecular, cuja análise nos permitiu selecionar os sete compostos propostos. Embora que a abordagem de docagem molecular considere os resíduos da proteína não flexíveis os resultados de DM (resíduos flexíveis) mostraram que as interações identificadas previamente e utilizadas na etapa de inspeção visual se mantiveram estáveis. Além disso, um clique estável formado entre os anéis do ligante e os resíduos Trp 83 e Tyr 334 destaca-se como uma possibilidade valiosa no desenvolvimento de novos inibidores AChE com melhor interação hidrofóbica e perfil inibitório.

Em resumo, todos os resultados indicaram o uso relevante dos métodos EON e docagem molecular como estratégia para identificação de novos compostos promissores a inibidores da acetilcolinesterase. Nesse contexto, foi possível apresentar sete esqueletos base inéditos e diferentes estruturalmente, como potentes candidatos anticolinesterásicos.

O desenvolvimento de uma estratégia integrada de métodos de ensaios virtuais baseados nas estruturas do ligante e do receptor, juntamente com a inspeção visual das moléculas melhores pontuadas resultou na seleção de 7 moléculas a serem avaliadas nos ensaios bioquímicos. Inicialmente as substâncias foram adquiridas e estão na fase de teste para determinação de  $IC_{50}$  (concentração para inibir 50% da atividade enzimática) e  $K_i$  (constante de inibição) contra as enzimas AChE e BuChE. Os resultados permitirão validar a estratégia de busca descrita no presente trabalho. Além disso, a partir dos compostos com boa inibição enzimática, estruturas análogas serão encontradas, a estratégia computacional desenvolvida será utilizada novamente para seleção de novas séries de compostos que serão submetidos aos ensaios bioquímicos a fim de estabelecer relações entre a estrutura química e a atividade biológica para novas classes de compostos inibidores da enzima alvo, etapa imprescindível no processo de descoberta de substâncias bioativas. Sendo assim, o uso dos protocolos computacionais validados serão úteis para a otimização molecular, novos compostos serão propostos e a partir de ensaios bioquímicos, novas informações poderão ser adicionadas aos modelos virtuais desenvolvidos para que esta se torne uma ferramenta mais abrangente. Por fim, uma vez identificado compostos seguros aos ensaios *in vitro* e celulares contra a enzima AChE, o modo de interação poderá ser confirmado por cristalografia de Raios-X, método experimental capaz de resolver a estrutura tridimensional da enzima em complexo com os ligantes.

## Referências Bibliográficas

1. Scarpini E, Schelterns P, Feldman H (2003). Treatment of Alzheimer's disease; current status and new perspectives. *Lancet Neurol* 2:539–547
2. Mckhann GM, Knopman DS, Chertkow H, et al (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging and the Alzheimer's Association workgroup. *Alzheimer's Dement* 1–7
3. Herrup K (2011). Commentary on “Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.” Addressing the challenge of Alzheimer's disease in the 21st century. *Alzheimer's Dement* 7:335–337
4. Citron M (2002). Alzheimer's disease: treatments in discovery and development. *Nat Neurosci* 5:1055–1057
5. Yiannopoulou KG, Papageorgiou SG (2013). Current and future treatments for Alzheimer's disease. *Ther Adv Neurol Disord* 6:19–33
6. Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT (2011). Neuropathological alterations in Alzheimer disease. *Cold Spring Harb Perspect Med*
7. Reitz C, Mayeux R (2014). Alzheimer disease: Epidemiology, diagnostic criteria, risk factors and biomarkers. *Biochem Pharmacol* 88:640–651
8. Francis PT, Palmer AM, Snape M, Wilcock GK (1999). The cholinergic hypothesis of Alzheimer's disease: a review of progress. *J Neurol Neurosurg Psychiatry* 66:137–147
9. Perry EK (1986). The cholinergic hypothesis ten years on. *Br Med Bull* 42:63–9
10. Dumas JA, Newhouse PA (2011). The cholinergic hypothesis of cognitive aging revisited again: Cholinergic functional compensation. *Pharmacol Biochem Behav* 99:254–261
11. Miyashita N, Straub JE, Thirumalai D (2009). Structures of  $\beta$ -amyloid peptide 1-40, 1-42, and 1-55-the 672-726 fragment of APP-in a membrane environment with implications for interactions with  $\beta$ -secretase. *J Am Chem Soc* 131:17843–17852
12. Gallagher D, Coen RF, Lawlor BA (2016). Alzheimer disease. *Neurodegener. Disord. A Clin. Guid.*
13. Kepp KP (2012). Bioinorganic chemistry of Alzheimer's disease. *Chem Rev* 112:5193–5239
14. Small G, Bullock R (2011). Defining optimal treatment with cholinesterase inhibitors in Alzheimer's disease. *Alzheimer's Dement* 7:177–184

15. Anand P, Singh B (2013). A review on cholinesterase inhibitors for Alzheimer's disease. *Arch Pharm Res* 36:375–399
16. Qureshi HY, Li T, MacDonald R, et al (2013). Interaction of 14-3-3 $\beta$  with microtubule associated protein tau within alzheimer's disease neurofibrillary tangles. *Biochemistry* 52:6445–6455
17. Kasper PK (2012). Bioinorganic chemistry of alzheimer's disease. *Chemical reviews* 112:5193–5239
18. Elaine KP (1986). The cholinergic hypothesis ten years on. *British Medical Bulletin* 42:63–69
19. Babic T, et al (1999). The cholinergic hypothesis of alzheimer's disease: a review of progress. *Journal of Neurology, Neurosurgery and Psychiatry* 67:558
20. Nori Graham (1997). Alzheimer's disease international. *International Psychogeriatrics* 9:5–6
21. Alzheimer's Association et al (2010). 2010 Alzheimer's disease facts and figures. *Alzheimer's & dementia: the journal of the Alzheimer's Association* 6:158
22. Gary S, Roger B (2011). Defining optimal treatment with cholinesterase inhibitors in alzheimer's disease. *Alzheimer's & Dementia* 7:177–184
23. Misik J, Korabecny J, Nepovimova E, et al (2016). Effects of novel tacrine-related cholinesterase inhibitors in the reversal of 3-quinuclidinyl benzilate-induced cognitive deficit in rats-Is there a potential for Alzheimer's disease treatment? *Neurosci Lett* 612:261–268
24. Colovic MB, Krstic DZ, Lazarevic-Pasti TD, et al (2013). Acetylcholinesterase Inhibitors: Pharmacology and Toxicology. *Curr Neuropharmacol* 11:315–335
25. G Waldemar, B Dubois, M Emre, J Georges, IG McKeith, M Rossor, P Scheltens, P Tariska, and B Winblad (2007). Recommendations for the diagnosis and management of alzheimer's disease and other disorders associated with dementia: Efn's guideline. *European Journal of Neurology* 14:e1–26
26. M Allgaier and C Allgaier (2014). An update on drug treatment options of alzheimer's disease. *Frontiers in bioscience (Landmark edition)* 19:1345
27. R Mark Wightman, Leslie J May, and Adrian C Michael (1998). Detection of dopamine dynamics in the brain. *Analytical Chemistry* 60:769A–793A
28. Cerbai F, Giovannini MG, Melani C, et al (2007). N1phenethyl-norcymserine, a selective butyrylcholinesterase inhibitor, increases acetylcholine release in rat cerebral cortex: A comparison with donepezil and rivastigmine. *Eur J Pharmacol* 572:142–150
29. Maria LB, Vincenza A, Manuela B, Andrea C, Anna M, Maurizio R, Michela R, Vincenzo T, Carlo M (2005). Heterocyclic inhibitors of ache acylation and peripheral sites. *Il Farmaco* 60:465–473

30. Edgar S, Ellen S, Leslie HE (1932). Choline-esterase. An enzyme present in the blood-serum of the horse. *Biochemical Journal* 26:2056
31. Ezio G (2001). Selective inhibitors of butyrylcholinesterase. *Drugs & aging* 18:891–898
32. Johann G (2006). The central role of chemoinformatics. *Chemometrics and intelligent laboratory systems* 82:200–209
33. Lavecchia A, Di Giovanni C (2013). Virtual screening strategies in drug discovery: A critical review. *Current medicinal chemistry* 20:2839–2860
34. Vincent Z, Aurélien G, Olivier M (2009). Docking, virtual high throughput screening and in silico fragment-based drug design. *Journal of cellular and molecular medicine* 13:238–248
35. Coupez B, Lewis RA (2006). Docking and scoring-theoretically easy, practically impossible? *Current medicinal chemistry* 13:2995–3003
36. Wendy D C, Piotr C, Christopher I B, Ian R G, Kenneth M M, David M F, David C S, Thomas F, James W C, Peter A K (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society* 117:5179–5197
37. Jacob D D, McCammon J A (2011). Molecular dynamics simulations and drug discovery. *BMC biology* 9:71
38. Kenno V, Olgun G, Alexander D. M Jr (2014). Molecular mechanics. *Curr Pharm Des.* 20: 3281–3292
39. Gaulton A, Bellis LJ, Bento AP, et al (2012). ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Res* 40:D1100-D1107
40. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK (2012). Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *J Med Chem* 55:6582–6594
41. Lavecchia A (2015). Machine-learning approaches in drug discovery: Methods and applications. *Drug Discov Today* 20:318–331
42. Lavecchia A, Di Giovanni C (2013). Virtual screening strategies in drug discovery: A critical review. *Curr Med Chem* 20:2839–2860
43. ChemAxon (2013). Marvin Sketch. In: <https://www.chemaxon.com/products/marvin/>.
44. OMEGA (2013). 2.5.1.4: OpenEye Scientific Software.

45. Hawkins PC, Nicholls A (2012). Conformer Generation with OMEGA Learning from the Data Set and the Analysis of Failures. *J Chem Inf Model* 52:2919–2936
46. Catto M, Pisani L, Leonetti F, et al (2013). Design, synthesis and biological evaluation of coumarin alkylamines as potent and selective dual binding site inhibitors of acetylcholinesterase. *Bioorganic Med Chem* 21:146–152
47. Ronco C, Foucault R, Gillon E, et al (2011). New huprine derivatives functionalized at position 9 as highly potent acetylcholinesterase inhibitors. *Chem Med Chem* 6:876–888
48. Leach A R (2007). Ligand-Based Approaches: Core Molecular Modeling. In Taylor J B & Triggle D J. *Comprehensive Medicinal Chemistry II* (Vol. 4, Chap. 4.05, pp. 87-118). Stevenage: Elsevier.
49. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004). Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 3:935–949
50. Wang G, Zhu W (2016). Molecular docking for drug discovery and development: a widely used approach but far from perfect. *Future Med Chem* 8:1707–1710
51. Meng X-Y, Zhang H-X, Mezei M, Cui M (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des* 7:146–57
52. Berman HM, Westbrook J, Feng Z, et al (2000). The protein data bank. *Nucleic Acids Res* 28:235–242
53. Kryger G, Silman I, Sussman JL (1999). Structure of acetylcholinesterase complexed with E2020 (Aricept): Implications for the design of new anti-Alzheimer drugs. *Structure* 7:297–307
54. Schrödinger (2009). Protein Preparation Wizard.
55. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004). PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res*
56. Macromodel (2009). Schrödinger, LLC.
57. LigPrep (2009). Schrödinger, LLC.
58. Epik (2009). Schrödinger, LLC,.
59. Banks JL, Beard HS, Cao Y, et al (2005). Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J Comput Chem* 26:1752–1780
60. Glide (2009). Schrodinger, LLC, NY.

61. Friesner RA, Banks JL, Murphy RB, et al (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem* 47:1739–1749
62. Tom Fawcett (2006). An introduction to roc analysis. *Pattern recognition letters* 27:861– 874
63. Anighoro A, Bajorath J (2016). Three-Dimensional Similarity in Molecular Docking: Prioritizing Ligand Poses on the Basis of Experimental Binding Modes. *J Chem Inf Model* 56:580–587
64. Triballeau N, Acher F, Brabet I, et al (2005). Virtual screening workflow development guided by the “receiver operating characteristic” curve approach. Application to highthroughput docking on metabotropic glutamate receptor subtype 4. *J Med Chem* 48:2534–2547
65. Fawcett T (2006). An introduction to ROC analysis. *Pattern Recognit Lett* 27:861–874
66. Zhang J (2015). The Hybrid Method of Steepest Descent: Conjugate Gradient with Simulated Annealing. In: *Molecular Structures and Structural Dynamics of Prion Proteins and Prions. Focus on Structural Biology*, vol 9. Springer, Dordrecht
67. Stewart A. Adcock and J. Andrew McCammon (2006). Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev* 106:1589-1615
68. Feller S E, Zhang Y, Pastor R W (1995). Constant pressure molecular dynamics simulation: The Langevin piston. *J. Chem. Phys* 103: 4613-4621
69. Eric F. P, Thomas D. G, Conrad C. H, Gregory S. Co, Daniel M. G, Elaine C. M, Thomas E. F (2004). UCSF Chimera – A Visualization System for Exploratory Research and Analysis. *Journal of Computational Chemistry* 25:1605-1612
70. Basu S, Bhattacharyya D, Banerjee R (2011). Mapping the Distribution of Packing Topologies Within Protein Interiors Shows Predominant Preference for Specific Packing Motifs. *BMC Bioinformatics* 12:195
71. Banerjee R, Sen M, Bhattacharyya D, et al (2003). The Jigsaw Puzzle Model: Search for Conformational Specificity in Protein Interiors. *J Mol Biol Volume* 333: 211-226

# Apêndice A – Inibidores da enzima CPB de *Leishmania Mexicana*: Modelagem por homologia, docagem e dinâmica molecular

O trabalho abaixo descrito foi desenvolvido durante o estágio de doutorado sanduíche por um período de seis meses na Escola de Farmácia da Universidade de Nottingham – Inglaterra sob supervisão do professor Dr. Charles Laughton.

## Resumo

Leishmaniose é uma doença causada por protozoários parasitas do gênero *Leishmania*. Atualmente, devido à toxicidade e aos efeitos colaterais das terapias existentes, e ao desenvolvimento de resistência, a busca de novas drogas é de extrema importância. Os parasitas de *Leishmania* contêm altos níveis de enzimas cisteína proteases. Neste estudo a estrutura 3D de uma das mais significativas destas, CPB (cisteína protease B), foi predita usando a ferramenta de modelagem por homologia MODELLER e a estrutura cristalográfica da enzima Rodesaína como referência. Para testar a exatidão e utilidade do modelo por homologia, utilizou-se estudos de docagem molecular de 24 moléculas *in-home* como atividade inibitória contra a CPB determinadas experimentalmente. Embora se acredite que as moléculas são inibidores covalentes reversíveis da enzima, baseou-se na hipótese de que as afinidades relativas dos complexos não covalentes presente antes da formação da ligação covalente entre a proteína e o ligante deveriam se correlacionar com a atividade inibitória. O resultado esperado foi obtido, no entanto, as atividades de alguns dos inibidores mais potentes (contendo o substituinte terc-butil-N-metilperazol) foram significativamente subestimadas pela função pontuação do programa de docagem para o modelo por homologia inicial. Portanto, o modelo por homologia da CPB foi refinado através de simulação de dinâmica molecular, realizada na presença de um ligante ligado covalentemente. O processo de docagem molecular não covalente foi então realizado utilizando a estrutura obtida a partir da DM e os compostos, que por sua vez correlacionaram muito bem com suas respectivas atividades inibitória enzimática. Dessa forma, temos um modelo de homologia da CPB validado, que pode ser utilizado prospectivamente para orientar na seleção e síntese de novos inibidores altamente potentes e seletivos.

## Abstract

Leishmaniasis is a disease induced by protozoan parasites of the genus *Leishmania*.

Nowadays, because of the toxicity and side effects of existing therapies, and developing resistance, the search for new improved drugs is of utmost importance. Leishmania parasites contain high levels of cysteine proteases. In this study, the 3D structure of one of the most significant of these, CPB (cysteine protease B), was predicted using the homology modelling tool MODELLER, using the known crystal structure of Rhodesain as the template. To test the correctness and utility of the model, it has been used for molecular docking (Glide) studies on 24 in-house molecules with experimentally-determined CPB inhibitory activity. Though the molecules are believed to be reversible covalent inhibitors of the enzyme, we hypothesise that the relative affinities of the non-covalent encounter complexes formed prior to drug-protein cross linking should correlate with inhibitory activity. Fair agreement was indeed obtained, however the activities of some of the most potent inhibitors (containing a tert-butyl-N-methylpyrazole moiety) were significantly underestimated by the docking to this initial model. Therefore, the model for the protein has been refined through extensive molecular dynamics simulation, performed in the presence of a covalently-bound ligand. Repeating the non-covalent docking calculations with models for the protein structure obtained from the MD results in calculated binding affinities of a range of dipeptidyl nitrile ligands that correlate very well with their established enzyme inhibitory activity. Thus, we now have a well-validated model for CPB, and a ligand analysis methodology, that can be used prospectively to guide the selection and synthesis of new highly potent and selective inhibitors.

## Introdução

Leishmaniose é uma doença causada por parasitas protozoários do gênero Leishmania e afeta mais de 12 milhões de pessoas ao redor do mundo. Os casos oficiais totalizam mais de 58 mil de leishmaniose visceral e 220 mil casos gerais por ano [1]. Existem três formas principais da doença, leishmaniose cutânea (CL), leishmaniose visceral (VL) ou Kala-azar e leishmaniose mucocutânea (MCL) [2]. Atualmente, devido à toxicidade e aos efeitos colaterais das terapias existentes, e ao desenvolvimento de resistência, a busca de novas drogas é de extrema importância. Os parasitas de Leishmania contêm altos níveis de enzimas da classe cisteína protease sendo a CPB (cisteína protease B) uma das mais significativas. Nenhuma estrutura cristalográfica para a CPB encontra-se disponível na literatura. Sendo assim, o objetivo deste estudo foi construir um modelo por homologia validado capaz de orientar na seleção e síntese de novos inibidores altamente potentes e seletivos.

## Metodologia e Resultados

### Modelagem molecular por homologia: *LmCPB2.8*

A sequência completa de aminoácidos da enzima cisteína protease B de *Leishmania Mexicana* foi adquirida no banco de dados de sequência de proteínas NCBI [3]. O programa BLAST [4] foi usado para pesquisar uma estrutura de referência disponível no PDB. A estrutura cristalográfica da Rodesaína do organismo *T. Brucei*, código do PDB 2P&U, com resolução de 1,65 Å e 61% de identidade com a sequência da CPB foi usada como referência para construir o modelo da CPB. O modelo por homologia para a CPB foi construído pelo programa MODELLER [5].

Inicialmente, a estrutura 3D de referência foi alinhada com a sequência da CPB. Segundo, as características espaciais, tais como, distância de carbono alfa ( $C\alpha - C\alpha$ ), ligações de hidrogênio e ângulos diedros da cadeia principal e lateral, foram transferidas da estrutura de referência para o modelo. Assim, foram obtidas várias restrições espaciais em sua estrutura. Por último, 100 modelos 3D da CPB foram obtidos satisfazendo o máximo de restrições possível. O melhor modelo foi selecionado pelo mais baixo valor da função pontuação DOPE (Discrete Optimized Protein Energy). Uma vez que o melhor modelo foi escolhido, a região da alça, resíduos 57 – 61, foi otimizada com minimização de energia gerando 100 novos modelos. O modelo final foi finalmente escolhido pelo menor valor de DOPE.

A estrutura secundária da sequência de aminoácidos da CPB foi predita por PSIPRED v3.3 [6]. O alinhamento 2D entre as sequências da CPB e da Rodesaína estão representados na Figura 1. O modelo obtido foi avaliado por inspeção do gráfico de Ramachandran (Figura 2) obtido a partir das análises PROCHECK [7]. As análises do modelo da CPB mostraram que 97% dos resíduos encontraram regiões favoráveis e apenas 3% dos resíduos foram observados fora das regiões permitidas, sugerindo boa qualidade do modelo.

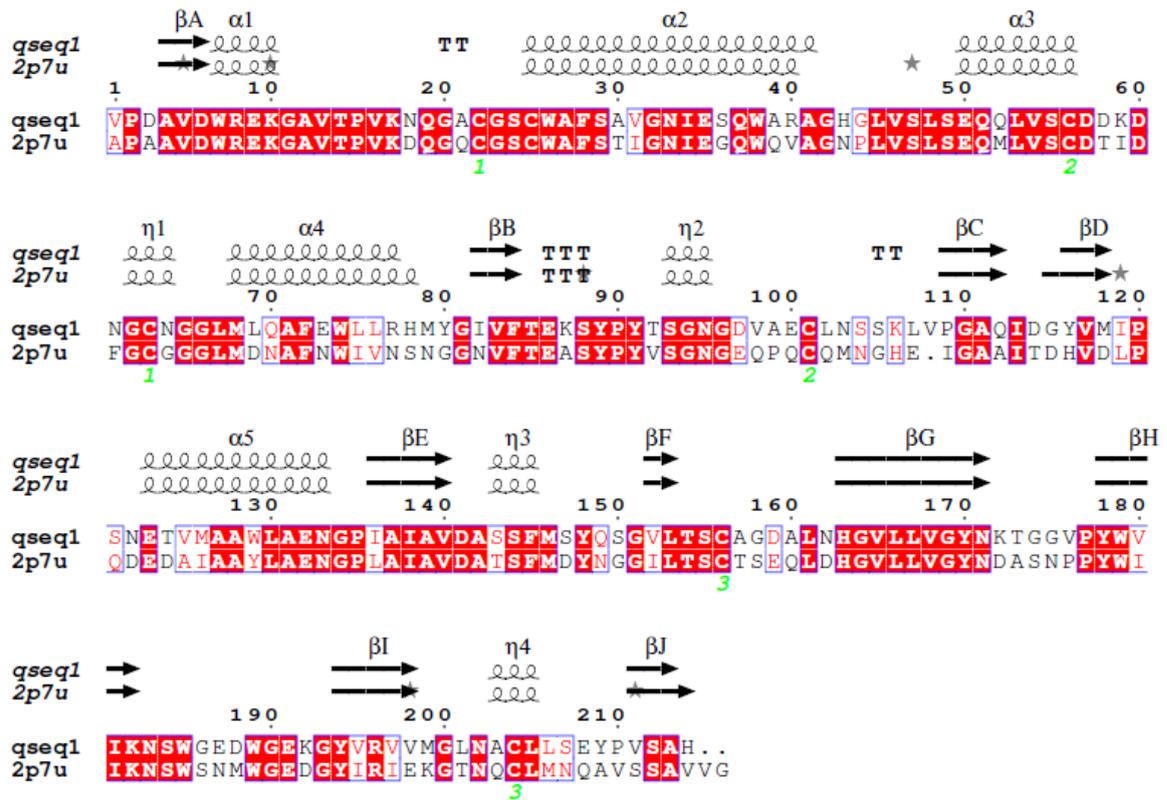


Figura 1. Alinhamento bidimensional entre a sequência de aminoácido da enzima CPB (qseq1) e da estrutura de referência (2p7u).

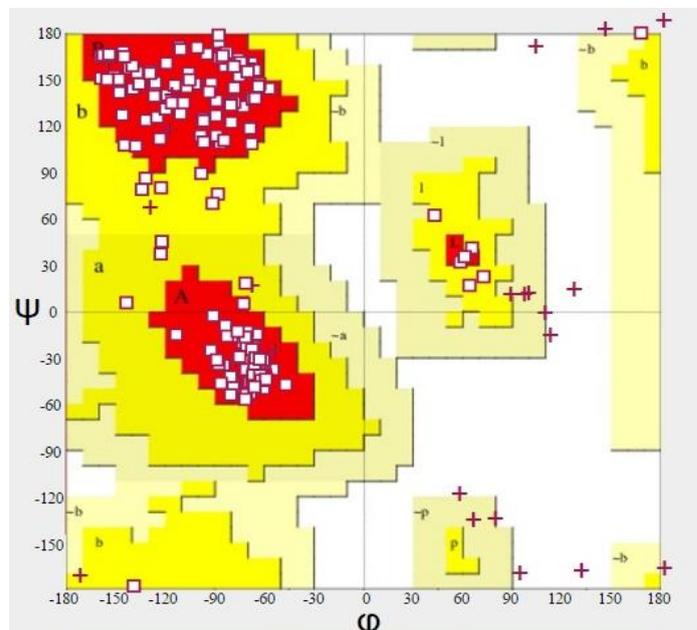


Figura 2. Gráfico de Ramachandran do modelo da CPB.

O programa VERIFY-3D [8] foi usado para analisar a compatibilidade atômica entre o modelo 3D criado e sua própria sequência de aminoácido 1D. O resultado mostrou que pelo menos

80% dos aminoácidos tiveram pontuação maior ou igual a 0,2 o que corresponde localização da cadeia lateral aceitável.

O modelo por homologia construído foi então comparado com o a estrutura de referência (PDB ID: 2P7U) usando o programa Chimera [9]. A Figura 3 mostra o alinhamento entre as estruturas. Os resíduos de aminoácido conservados entre as estruturas estão representados em azul e os não conservados estão em vermelho. Foi possível observar que no sub-sítio S2 da enzima Rodesaína o resíduo Ala208 foi substituído por uma Tyr209 na estrutura modelo da CPB. Além disso, no sub-sítio S3 da Rodesaína contém uma Phe61 enquanto que na enzima CPB contém uma Asp61. Contudo, os principais resíduos dos sub-sítios S1 e S1' foram conservados.

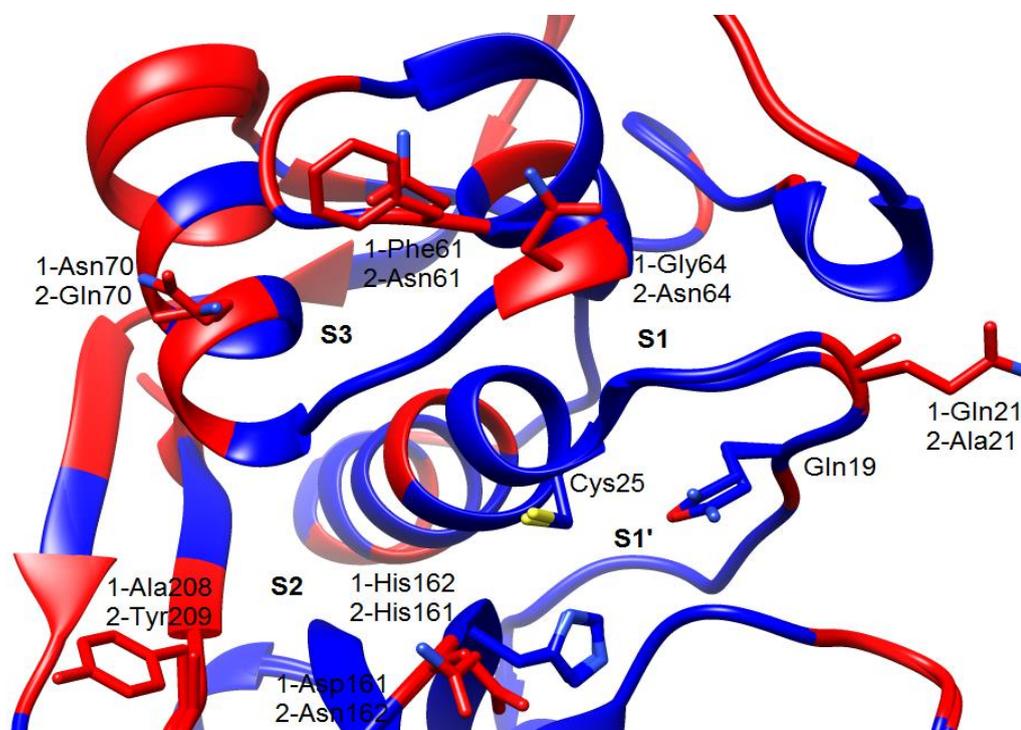


Figura 3. Alinhamento entre a estrutura de referência 2P7U representada pelo número 1 e o modelo por homologia construído para a CPB representada pelo número 2.

Modelagem por homologia é uma ferramenta poderosa na predição de estruturas de proteína. No presente estudo, o modelo por homologia da enzima CPB foi com sucesso construído usando a estrutura de referência da enzima Rodesaína e os métodos mencionados acima. O modelo da CPB foi então, nos próximos passos, validado por métodos de ensaio virtual baseados na estrutura do receptor, são eles docagem e dinâmica molecular. Uma série de 24 compostos sintetizados e testados *in-home* contra a CPB foi utilizada.

## Docagem Molecular

Com o modelo por homologia, os compostos do banco de dados *in-home* foram analisados por ensaio virtual. O método de docagem molecular foi realizado para testar se era possível prever a afinidade de ligação com alguma precisão e encontrar uma correlação entre os valores de  $p_{ki}$  e a pontuação do programa de docagem.

Os cálculos de docagem desenvolvidos neste projeto foram realizados usando o programa Glide [10]. Os compostos foram preparados usando LigPrep [11], Epik[12] e campo de força OPLS-2005, todos os possíveis estados tautoméricos e de protonação foram determinados a pH 5,5, bem como calculou os estados de ligação de metal. Os estereoisômeros foram determinados a partir da estrutura 3D previamente definida. A estrutura da proteína foi preparada usando Protein Preparation Wizard implementado no programa Maestro (Schrodinger 2016). Para ajuste dos estados de protonação da proteína Propka [13] a pH 5,5 foi utilizado. Todas as Histidinas e o resíduo Asp57 foram considerados protonados, o resíduo Glu208 foi considerado carregado negativamente. Os resíduos Cys25 e His161 formaram o par de tiolato/imidazol necessário para o mecanismo catalítico. Após ajustes dos estados de protonação foi realizada a etapa de minimização de energia usando campo de força OPLS-2005 e os parâmetros padrões. A região do sítio ativo foi definido pelo Receptor Grid Generation [14]. O tamanho da grade foi ajustado para 20x20x20Å centrado nos resíduos selecionados, Asn162, Cys25 e Gly66. O processo de docagem molecular foi então realizado no programa Glide com precisão padrão SP. O número de 100 poses por ligante foi escolhido.

Inicialmente, o composto Neq0409 da estrutura PDB 4QH6 foi escolhido como referência para as análises seguintes. O resultado de docagem, não covalente e sem restrição, do composto Neq0409 na estrutura modelo da CPB está representado na Figura 4. A Figura 4 mostra a superposição do composto Neq0409 na sua estrutura cristalográfica 4QH6, em azul, com o resultado de docagem em cinza. O valor de RMSD entre eles foi de 2.6 Å. O modo de interação mostra que as posições P1, P2 e P3 ocupam os subsítios S1, S2 e S3, respectivamente.

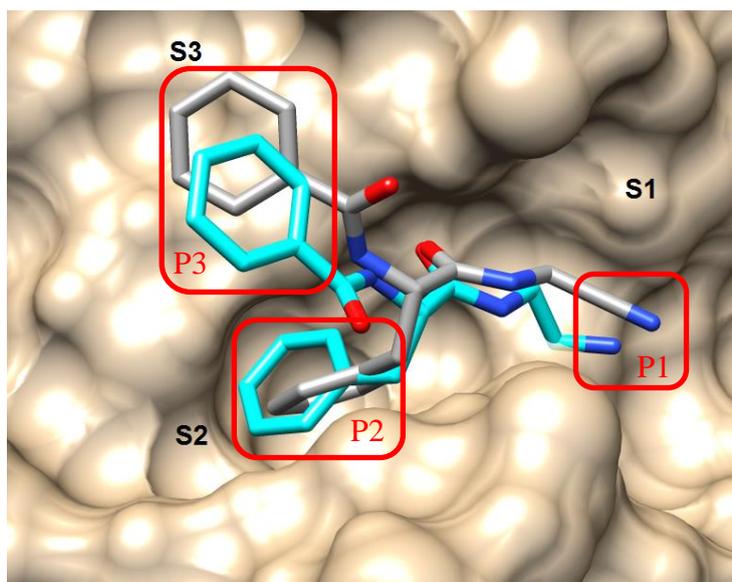


Figura 4. Superposição dos modos de ligação do composto Neq0409 a partir do resultado de docagem (cinza) e sua estrutura cristalográfica a partir do PDB com código 4QH6 (azul).

Após docagem não covalente sem restrição da série com 24 compostos *in-home*, cujo resultado será mostrado a seguir, o processo de docagem foi refeito com o uso de restrições de *smart* nas posições P1, P2 e P3. Para gerar as restrições, a grade de docagem foi criada com restrições de *smart* nas três posições como está mostrado na Figura 5.

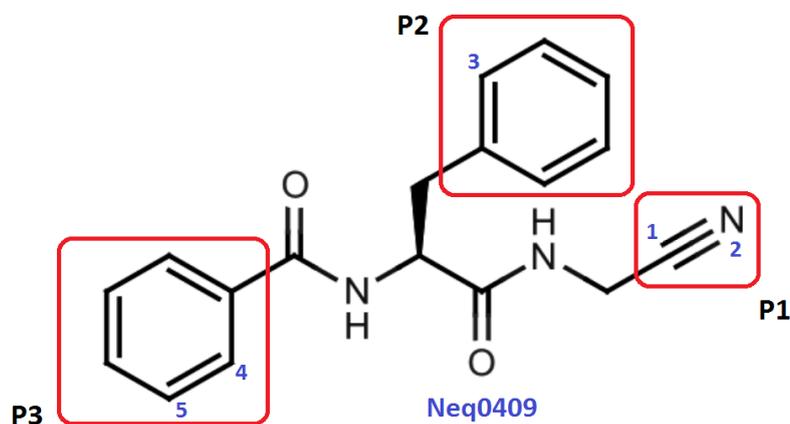


Figura 5. Estrutura 2D do inibidor Neq0409. Átomos 1, 2, 3, 4 e 5 (Átomos 1605, 1606, 1597, 1612 e 1611 na estrutura PDB ID: 4QH6)

A fim de usar os átomos do composto Neq0409 como referência de posição sua estrutura cristalográfica, 4QH6, foi alinhada a estrutura do modelo da CPB. Para a posição P1 o *smart* C#N a uma distância de até 5,0Å do centroide dos átomos 1 e 2 foi usado. Na posição P2 e P3 os *smart*

específicos de cada composto foi usado a uma distância de até 4,0 Å e 4,6 Å do átomo 3 e centroide dos átomos 4 e 5 respectivamente.

A Figura 6 apresenta as diferenças entre os valores da função pontuação do programa Glide a partir da docagem não covalente sem uso de restrição para as melhores poses selecionadas e a partir da docagem não covalente com uso das restrições de SMARTS em P1, P2 e P3. Para seleção das melhores poses a partir do resultado de docagem sem restrição considerou-se as melhores conformações em que P1, P2 e P3 ocupassem os sub-sítios S1, S2 e S3 alcançando o mais próximo possível da conformação do composto de referência Neq0409. Para a docagem com restrição das três posições foi considerado os melhores valores da função de pontuação.

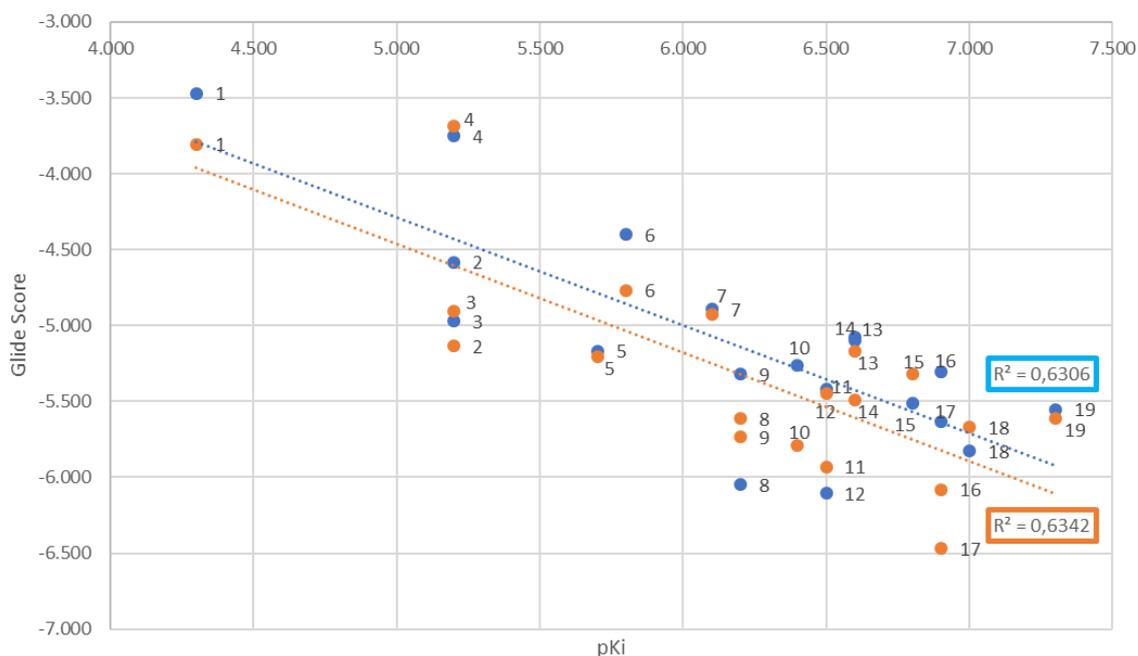


Figura 6. Função pontuação do programa Glide (Glide Score) versus os valores de pKi. Os resultados da docagem molecular não covalente sem uso de restrição representado em azul e os resultados da docagem molecular não covalente com uso de restrições de posição de *smart* em laranja.

Os coeficientes de determinação  $R^2$  foram equivalentes em ambos os processos de docagem, mostrando correlação suficiente para recuperar os inibidores da CPB mais potentes do bando de dados *in-home*. Apesar da correlação ser suficiente, através da docagem não covalente, as poses de docagem mostrou uma resistência em posicionar o grupo tert-butyl-N-metil pirazol no sub-sítio S3. Sendo assim, este fato nos levou ao uso de simulação por dinâmica molecular cujo objetivo foi analisar o comportamento deste grupo no sub-sítio S3 e refinar o modelo por homologia da CPB na presença de um ligante ligado covalentemente.

Além dos métodos de docagem molecular descritos anteriormente também foi realizado processos de docagem molecular não covalente com restrição de *smart* apenas na posição P1, outro

processo com restrições apenas nas posições P2 e P3 e por fim docagem molecular covalente, entretanto nenhuma melhora na correlação foi encontrada.

## Dinâmica Molecular

Nos estudos de docagem molecular foi observado diferentes conformações do grupo Tert-Butil-N-metil pirazol em P3. Por essa razão, simulações de DM foi utilizada para explorar as interações entre o modelo da CPB e o composto Neq0569 usando o pacote Amber 16 [15] com o campo de força AMBER. O composto Neq0569 foi o escolhido para representar todos os compostos da série que possuem este mesmo grupo na posição P3 além de apresentarem os melhores valores de pKi. Para realizar a simulação, o complexo foi solvatado em uma caixa truncada octaédrica usando o modelo de água do tipo TIP3P. Antes de rodar a simulação, duas etapas de minimização foram feitas para estrutura inicial adquirida dos resultados de docagem. Primeiramente, as moléculas de água foram minimizadas durante 1000 ciclos iniciais utilizando o método *steepest descendente* [16] seguidos por 1000 ciclos utilizando o método de gradiente conjugado [17]. Na segunda minimização, 2000 ciclos foram realizados utilizando o método *steepest descent* para todo o sistema, seguido de 1500 ciclos pelo método de gradiente conjugado. Após os ciclos de minimização de energia, o sistema foi aquecido de 0K a 300K durante 1,25 ns em *ensemble* NVT. Na etapa seguinte, após o sistema atingir a temperatura de 300K, realizou-se uma etapa de equilíbrio da densidade durante 1,25 ns. Então, foi realizada a etapa de equilíbrio final do sistema de 7,5 ns. Ambas em *ensemble* NTP. A etapa de simulação do sistema final foi então realizada a temperatura constante de 300 K, durante 100 ns e escalonamento isotrópico de pressão com tempo de relaxação de 1 ps. Em todas as etapas a partir dos ciclos de minimização de energia, o algoritmo SHAKE [17] foi utilizado para tratar todas as ligações com átomos de hidrogênio, um corte de 10 Å foi usado para todas as distâncias de interações entre átomos não ligados e a dinâmica de Langevin [18] com frequência de choque de 2 ps<sup>-1</sup> foi adotada para controle de temperatura. O tempo de simulação DM foi dividido em 10 corridas de 10 ns com intervalo de 2 fs e a cada 25000 passos os dados de trajetória eram salvos. Assim, 200 conformações (frames) foram obtidas a cada 10 ns de simulação.

Após a análise da simulação por DM, o mesmo processo de docagem molecular não covalente com restrições nas posições P1, P2 e P3 foi reproduzido entre a série com 24 compostos *in-home* e duas conformações diferentes obtidas através da DM. Ambas conformações estão representadas na Figura 7.

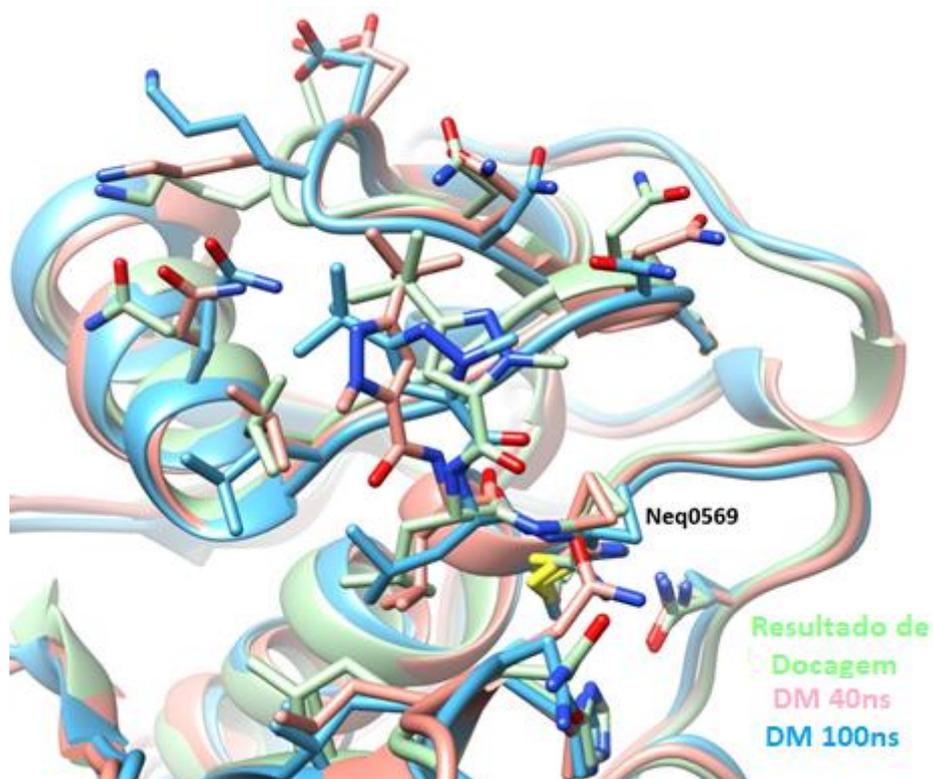


Figura 7. Superposição do modo de ligação do composto Neq0569 em complexo com a CPB a partir dos resultados de DM e Docagem Molecular não covalente com restrições de posição em P1, P2 e P3.

Os valores de RMSD da proteína permaneceram estáveis durante o tempo de simulação (ver Figura 8). Duas conformações principais são possíveis para o composto Neq0569 (Figura 9). Por essa razão, no primeiro processo de docagem usou-se a estrutura referente a conformações de 40 ns de simulação. No segundo processo, a estrutura adquirida aos 100 ns de simulação foi usada. Sendo o último caso o que conseguiu melhor correlação entre os resultados da docagem e os valores de  $p_{ki}$ , com  $R^2$  igual a 0,62 (Figura 10), desconsiderando o composto Neq0413 que foi um valor atípico.

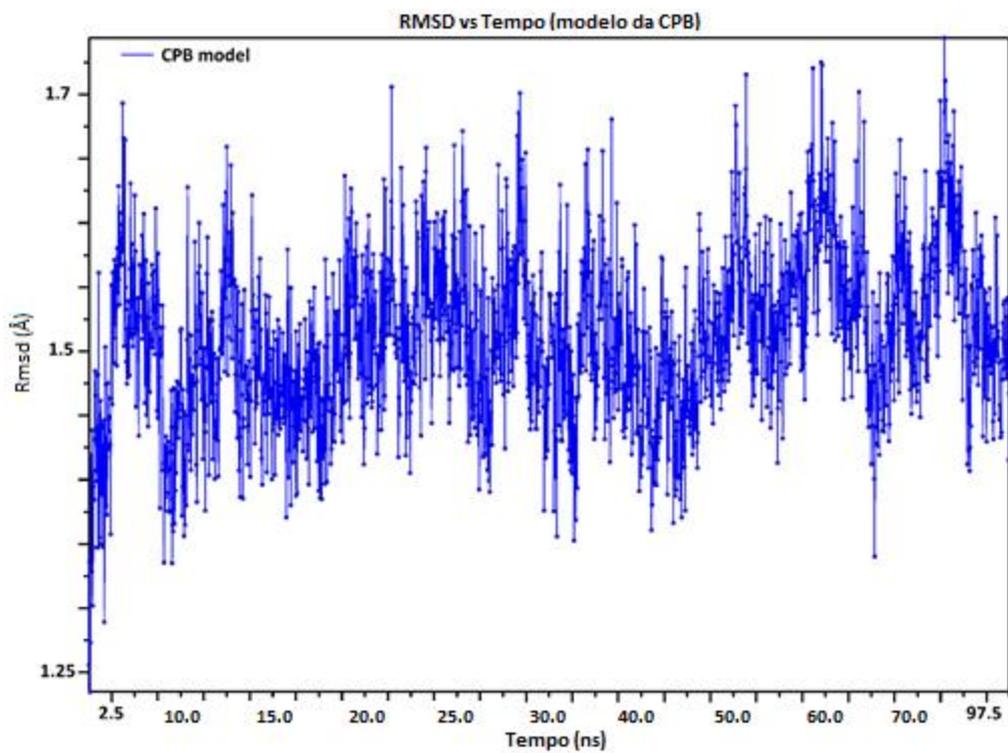


Figura 8. Evolução temporal da DM versus valores de RMSD da proteína para o modelo da CPB.

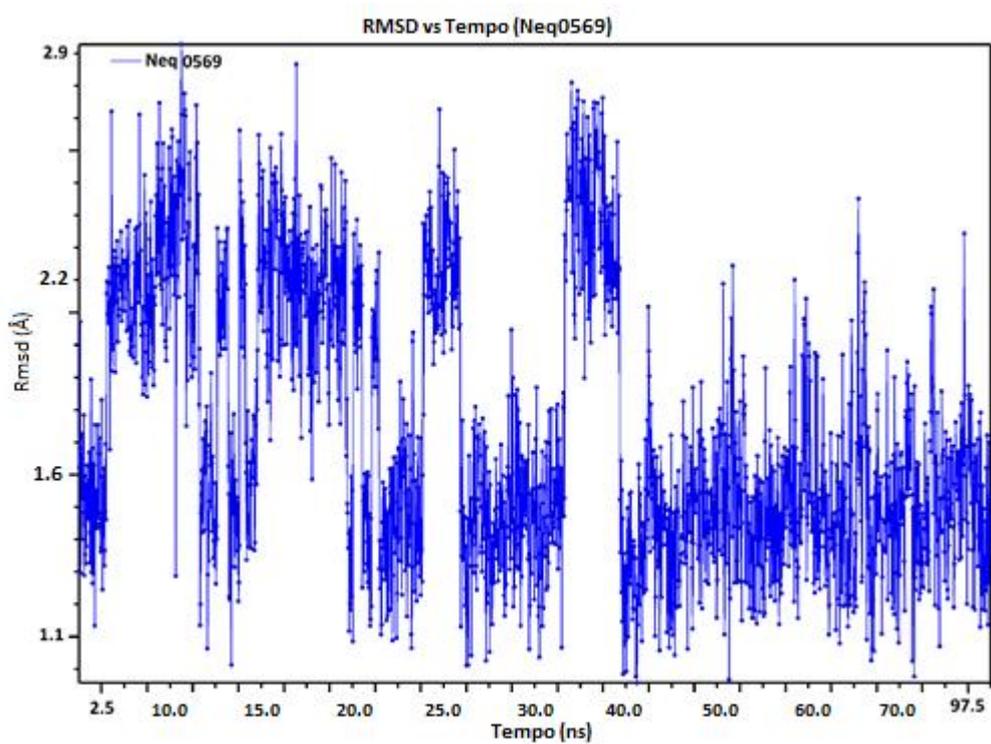


Figura 9. Evolução temporal da DM versus valores de RMSD do ligante Neq0569.

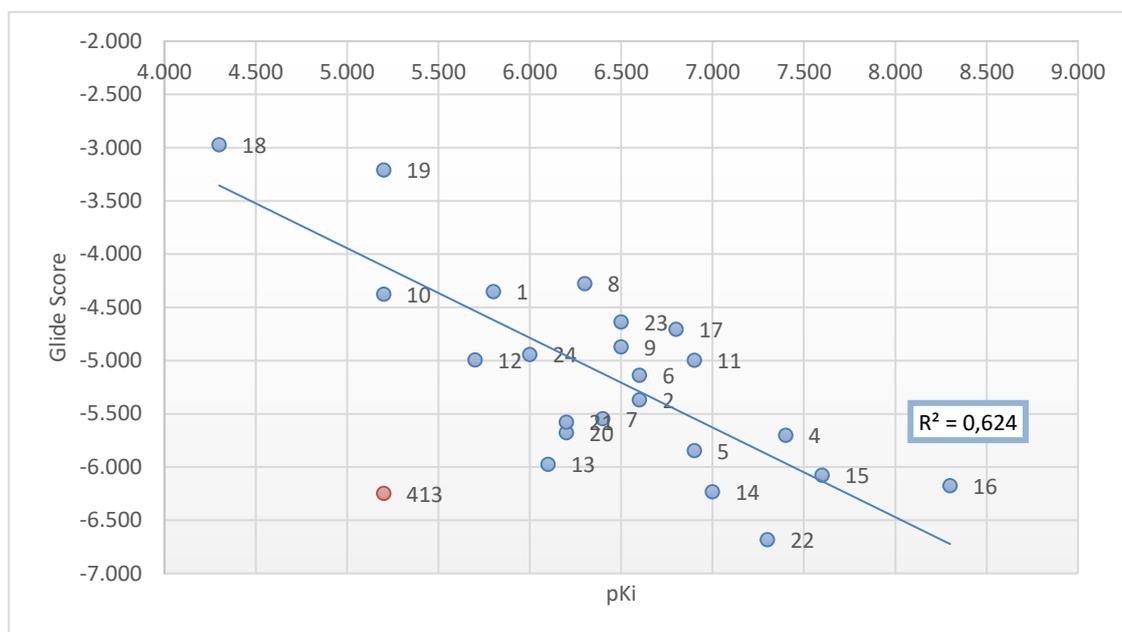


Figura 10. Melhores valores da função pontuação do Glide (GlideScore) versus valores de pki. Processo de docagem não covalente com restrição das posições P1, P2 e P3 no modelo por homologia da CPB. Composto 413 um valor atípico.

## Conclusões

No presente projeto, modelagem por homologia, docagem e dinâmica molecular foram com sucesso usados para estudar a série *in-home* com 24 compostos inibidores da CPB. O modelo tridimensional da cisteína protease B de *Leishmania mexicana* 2.8 foi muito bem construído. Docagem e dinâmica molecular foram usados para explorar as interações entre a CPB e seus inibidores. Os resultados de docagem molecular usando as três restrições de posição e a estrutura obtida a partir da DM leva a uma melhor análise das características estruturais que podem ser usadas para orientar a seleção e síntese de novos inibidores potentes e seletivos da CPB. A correlação positiva encontrada após as etapas de dinâmica e docagem molecular não covalente valida a qualidade do modelo proposto. Além disso, trabalhos experimentais que estão sendo desenvolvidos no grupo de pesquisa mostrou a alta semelhança entre o modelo desenvolvido neste projeto e a estrutura obtida por cristalografia de raios-x.

## Referências

1. Alvar J, Velez ID, Bern C, Herrero M, Desjeux P, Cano J, et al (2012). Leishmaniasis Worldwide and Global Estimates of Its Incidence. *PLoS ONE* 7:e35671
2. Fernández-Figueroa, E A, Rangel-Escareño, C, Espinosa-Mateos, V, Carrillo-Sánchez, K, Salaiza-Suazo, N, et al (2012). Disease Severity in Patients Infected with *Leishmania mexicana* Relates to IL-1 $\beta$ . *PLoS Negl. Trop. Dis.* 6:e1533
3. National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information (1988) – [cited 2017]. Available from: <https://www.ncbi.nlm.nih.gov/>
4. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215:403-410
5. Sali and T. L. Blundell (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779-815
6. Jones DT. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292: 195-202
7. Laskowski R A, MacArthur M W, Moss D S, Thornton J M (1993). PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.* 26, 283-291
8. Bowie JU, Lüthy R, Eisenberg D (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253:164-70
9. Eric F. P, Thomas D. G, Conrad C. H, Gregory S. Co, Daniel M. G, Elaine C. M, Thomas E. F (2004). UCSF Chimera – A Visualization System for Exploratory Research and Analysis. *Journal of Computational Chemistry* 25:1605-1612
10. Glide (2009) Schrodinger, LLC, NY.
11. LigPrep (2009) Schrödinger, LLC.
12. Epik (2009) Schrödinger, LLC.
13. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* doi:10.1093/nar/gkh381
14. J.L.; Beard, H.S.; Cao, Y.; Cho, A.E.; Damm, W.; Farid, R.; Felts, A.K.; Halgren, T.A.; Mainz, D.T.; Maple, J.R.; Murphy, R.; Philipp, D.M.; Repasky, M.P.; Zhang, L.Y.; Berne, B.J.; Friesner, R.A.;

Gallicchio, E.; Levy, R.M (2005). Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comp. Chem.* 26:1752

15. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman (2016), AMBER 2016, University of California, San Francisco.

16. Zhang J (2015). The Hybrid Method of Steepest Descent: Conjugate Gradient with Simulated Annealing. In: *Molecular Structures and Structural Dynamics of Prion Proteins and Prions. Focus on Structural Biology*, vol 9. Springer, Dordrecht

17. Stewart A. Adcock and J. Andrew McCammon (2006). Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev* 106:1589-1615

18. Feller S E, Zhang Y, Pastor R W (1995). Constant pressure molecular dynamics simulation: The Langevin piston. *J. Chem. Phys* 103: 4613-4621

## **Apêndice B - Similarity search combined with Docking and Molecular Dynamics for novel hAChE Inhibitor Scaffolds.**



# Similarity search combined with docking and molecular dynamics for novel hAChE inhibitor scaffolds

Nadia Melo Borges<sup>1</sup> · Geraldo Rodrigues Sartori<sup>2</sup> · Jean F. R. Ribeiro<sup>2</sup> · Josmar R. Rocha<sup>2</sup> · João B. L. Martins<sup>3</sup> · Carlos A. Montanari<sup>2</sup> · Ricardo Gargano<sup>1</sup>

Received: 21 June 2017 / Accepted: 27 November 2017  
© Springer-Verlag GmbH Germany, part of Springer Nature 2017

## Abstract

The main purpose of this study was to address the performance of virtual screening methods based on ligands and the protein structure of acetylcholinesterase (AChE) in order to retrieve novel human AChE (hAChE) inhibitors. In addition, a protocol was developed to identify novel hit compounds and propose new promising AChE inhibitors from the ZINC database with 10 million commercially available compounds. In this sense, 3D similarity searches using rapid overlay of chemical structures and similarity analysis through comparison of electrostatic overlay of docked hits were used to retrieve AChE inhibitors from collected databases. Molecular dynamics simulation of 100 ns was carried out to study the best docked compounds from similarity searches. Some key residues were identified as crucial for the dual binding mode of inhibitor with the interaction site. All results indicated the relevant use of EON and docking strategy for identifying novel hit compounds as promising potential anticholinesterase candidates, and seven new structures were selected as potential hAChE inhibitors.

**Keywords** Virtual screening · hAChE inhibitors · ROCS and EON methods · Docking and molecular dynamics

## Introduction

Alzheimer's disease (AD) is the most common cause of dementia among people over 65 years old, causing an irreversible and progressive brain disorder associated with their daily life and activities [1–3]. The alterations presented by patients depend on the stage of AD, and the following stages are found during evolution of this disease [1–7]:

- In the initial stage patients present cognitive, behavioral and mood alterations, whose main characteristics are frequent forgetting and repetitive questions.

- In moderate stage dementia patients, there is a cognitive decline, loss of recent memory and problems with language.
- In the final stage of dementia, there are severe behavioral changes, loss of identity and the patient becomes excessively dependent.

Recent studies on treatment of AD suggest three major pathophysiological [6–15]. The two main histopathological criteria for AD are observations of extracellular deposits of fibrillar peptides called senile plaques (SP) and of neurofibrillary tangles (NFTs), both described by Alois Alzheimer in his original case report result [12]. The SP are formed from the abnormal extracellular accumulation and deposition of the amyloid- $\beta$  peptide with 40 or 42 amino acids [known as  $\beta$ -amyloids ( $A\beta$ )] of the metabolism of the amyloid precursor protein (APP) found in the membranes of cells and organelles such as mitochondria [11]. The NFTs consist of twisted strands of hyperphosphorylated tau protein, a protein important for the structural integrity of microtubules, structural tubulin polymers of the cytoskeleton of the neurons [16]. These plaques and tangles are commonly observed upon neurodegeneration, thus they are also not unique to AD [6, 13]. The biochemical

---

This paper belongs to Topical Collection VI Symposium on Electronic Structure and Molecular Dynamics – VI SeedMol

---

✉ Nadia Melo Borges  
nadiameloborges@unb.br; nadiameloborges@gmail.com

<sup>1</sup> Institute of Physics, University of Brasilia, Brasilia, DF, Brazil

<sup>2</sup> Institute of Chemistry of São Carlos, University of São Paulo, São Carlos, SP, Brazil

<sup>3</sup> Institute of Chemistry, University of Brasilia, Brasilia, DF, Brazil

investigation of patients with AD led to the “cholinergic hypothesis” that proposed the degeneration of cholinergic neurons in the synaptic cleft and the associated loss of cholinergic neurotransmission in the cerebral cortex that contributed to the cognitive function impairments [8, 17–22]. This hypothesis suggests that an increase in the neurotransmitter acetylcholine (ACh) through the inhibition of the enzymes cholinesterases (ChEs) i.e., acetylcholinesterase (AChE) and butyrylcholinesterase (BuChE), that hydrolyze ACh into acetic and choline, would improve the cognitive profile of patients [9].

Among different types of drugs, the widely used cholinesterase inhibitors (ChEIs) represent the first class of drugs approved for the symptomatic treatment of AD [23]. The first inhibitor approved by the FDA (Food and Drug Administration) was tacrine; however, its application has been limited due to side effects. Subsequently, donepezil, rivastigmine and galantamine became available [23, 24]. ChE inhibition can be selective (donepezil, a selective AChE inhibitor) or promiscuous (rivastigmine and galantamine, inhibitors of AChE and BuChE). However, experiments show that a partial inhibition of BuChE activity is enough to increase ACh levels, and may also reduce the number of plaques SP [25]. In this sense, the aim of this study was to compare the performance of virtual screening (VS) methods based on ligands and the protein structure of AChE. In addition, we also wanted to develop a strategy for identifying novel hit compounds and to propose new and superior AChE inhibitors from a free database of commercially available compounds (ZINC [26]). Therefore, 3D similarity searches using ROCS (rapid overlay of chemical structures) [27, 28] and EON (similarity analysis through comparison of electrostatic overlay) [29], and docking (a structure-based virtual screening) using Glide program [30] were used to retrieve AChE inhibitors from collected databases. Molecular dynamics (MD) simulation using Amber14 program [31] was carried out to study the selected compounds.

## Methods

### Dataset

Three datasets were used for VS analysis. For the first dataset (DS1), ca. 1600 compounds assayed against the *Homo sapiens* AChE enzyme (hAChE), were collected from the ChEMBL bioactivity database [32]. Compounds with  $IC_{50}$  less than 1  $\mu$ M were considered inactive. The second dataset (DS2) was collected from a directory of useful decoys enhanced (DUD-e) [33], a benchmarking set of 453 known ligands and 26,234 decoy sets for hAChE. DUD-e has been used widely for the validation of virtual screening, and is available at <http://dude.docking.org>. The last dataset (DS3) was generated using DUD-e free on-line system with 243 active compounds against hAChE. In total 12,115 decoys

were created. These datasets were then prepared and used to assess the ability of the virtual screening in order to discriminate between active and inactive compounds.

### Ligand-based method

In order to find compounds for the target, the approaches used for the Ligand-Based Virtual Screening (LBVS) methods are recommended [34, 35]. In the case of a 3D structure, where docking is not highly applicable, LBVS methods carry out an active and promising technique to recognize new compounds with similar biological activity.

### Similarity search

**Query selection** In order to identify compounds with a common substructure, compared to the approved drugs with AChE-inhibitory activity, some drugs were used as queries for both ROCS and EON similarity methods. The following four drugs and an active compound were considered in the present study: Donepezil, Tacrine, Galantamine, Rivastigmine and HuprineW (Fig. 1). Their structures were drawn using the program Marvin Sketch [36] and then converted into a three-dimensional (3D) structure using the OMEGA [37, 38] package from OpenEye. OMEGA is a conformation generator for molecules. OMEGA fragments the molecules and reassembles them to generate multi-conformers. The algorithm implemented in OMEGA attempts to capture all of the relevant bond length and angles, and ring conformations for an input molecule. For OMEGA one, default parameters were used except, *maxconfs* = 1 (sets the maximum number of conformation to be generated), *flipper* set to true (flips all unspecified stereocenters), and *strictstereo* set to false (not requires all stereocenters to be specified).

**Dataset preparation** For all datasets, the conformer libraries required by ROCS and EON programs were built with OMEGA for use in the calculations. Default parameters were used with the following exceptions: *flipper* set to true (flips all unspecified stereocenters), and *strictstereo* set to false (not requires all stereocenters to be specified). The program fragments the compounds and reassembles them to generate all possible conformations. These conformer files were used as input database for performing the ROCS, and then to generate input files for EON program.

**ROCS and EON details** ROCS analysis was carried out for the shape-based superposition method. Default parameters were used with the following exception: *besthits* set to 10,000 (for DS1), 15,000 (for DS3) and 35,000 (for DS2). In ROCS every conformer is aligned to each query molecule. Therefore, the database compounds were ranked according to their Shape Tanimoto (ST) score, considering shape search score of

ligands, ranging from 0 to 1, where 1 represents the best match between the query and the compound.

The EON program was employed to calculate the electrostatic similarity that uses a field-based measure of Electrostatic Tanimoto (ET) score, using full Poisson-Boltzmann electrostatics, to compare the electrostatic potential of each database compound with the query molecule. For this purpose, EON calculates new partial charges using MMFF94 force field, and has the ability to adjust both query and database molecule to a neutral pH model. To ensure a good quality alignment between the molecules, the ROCS program provides the best input to EON. All EON calculations were performed with default values except the *besthits* (set to same values used in the ROCS code).

### Structure-based method

Docking depends on the available protein structure. Furthermore, it involves the prediction of the complex between protein and ligand into the active site, and enables the most favorable conformation to be determined. The 3D structure of protein–ligand provides detailed information at atomic level, which is very useful for exploring the interactions between an inhibitor and protein, and to lead optimization [41–43].

### Molecular docking

**Protein preparation** All protein X-ray crystal structures were downloaded from the protein data bank (PDB) [44]. Three representative hAChE targets were selected from the 20 3D structures available in the PDB, with access codes 4BDT, 4M0E and 4EY7. These structures contain known co-crystallized inhibitors including Huprine W (HUW), Dihydropantanshine I (1YL) and

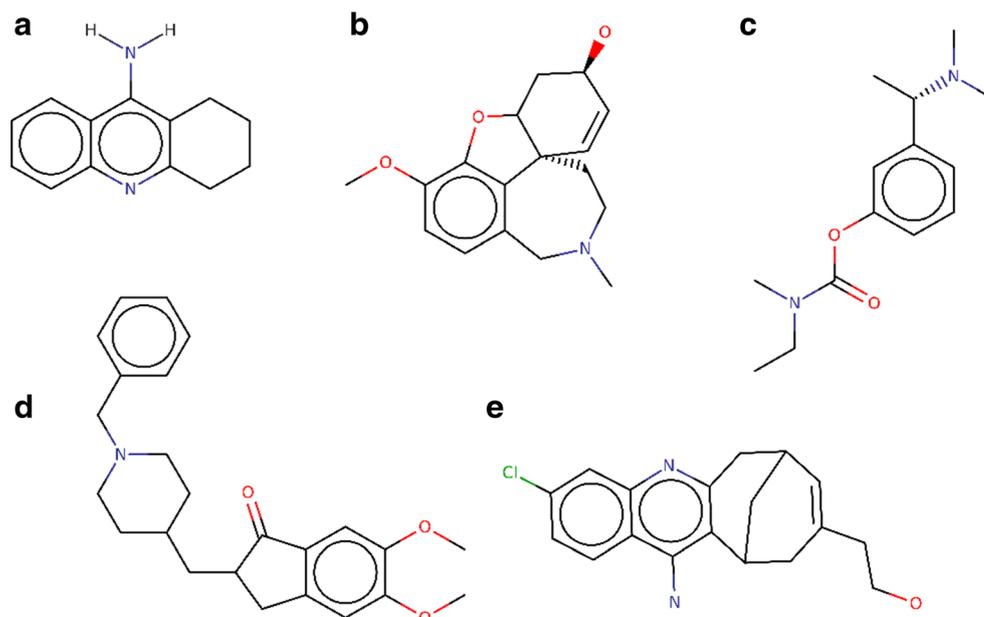
Donepezil (E20), respectively. In order to explore significant changes in the active site, the structures were chosen because their Tyr337 in hAChE (Fig. 2) adopts a wide range of conformation characterized by a RMSD of  $\chi_1$  torsion angle of  $38^\circ$ . The conformations can be assigned based on their  $\chi_1$ . The  $\chi_1$  dihedral values are  $-152^\circ$ ,  $164^\circ$  and  $111^\circ$  for 4EY7, 4M0E and 4BDT, respectively. The same conformations were found in the tcAChE (*Torpedo californica* AChE) where the Tyr337 is replaced by Phe330, which has been described as a “swinging gate” [45].

The 3D structures of hAChE used for docking was prepared using Protein Preparation Wizard [46]. The ionizable residues were set to their normal ionization states at pH 7.4 according the program Propka [47]. During the preparation, all missing hydrogen atoms were added, all water molecules and ligand deleted, and the energy minimized in the Macromodel [48] using the OPLS-2005 force field. All other settings were kept as default setting. The receptor grids were generated using the grid generation module in Glide within a grid box of  $22\text{\AA} \times 22\text{\AA} \times 22\text{\AA}$ , centered on co-crystallized ligands.

**Ligand preparation** For docking studies, all the datasets were prepared using LigPrep2.3 [49], Epik [50] and OPLS-2005 force field with IMPACT code [51], generating all possible tautomeric, enantiomeric, and protonation states at target pH  $7.4 \pm 0.5$ , as well as calculating metal binding states. The same protocol was used to prepare co-crystallized ligands for redocking calculations described below.

**Docking details** Molecular docking was carried out using the hAChE crystallographic structures and the databases, DS1, DS2 and DS3, by using the program Glide SP as implemented in Schrödinger Suite [30, 52]. The datasets were docked

**Fig. 1** *Homo sapiens* acetylcholinesterase (hAChE) inhibitors used as queries structures. **a** Tacrine  $IC_{50} = 0.1 \mu\text{M}$  [39]. **b** Galantamine  $IC_{50} = 0.3 \mu\text{M}$  [39]. **c** Rivastigmine  $IC_{50} = 0.7 \mu\text{M}$  [39]. **d** Donepezil  $IC_{50} = 0.002 \mu\text{M}$  [39]. **e** HuprineW  $IC_{50} = 0.001 \mu\text{M}$  [40]



considering ligand flexibility by ligand conformational sampling method implemented in Glide program. Fifty poses per ligand was included, and five poses were reported. The best pose for a given ligand was determined by the GlideScore score.

### Evaluation metrics

To determine the performance of the different VS, a ROCS curve was generated. The receiver operating characteristic (ROC) curve was generated by classifying the molecules (active and inactive) and docking energy values calculated by the scoring function. If the molecule is active (positive) and the docking classifies it as positive (well ranked), then this occurrence is counted as a true positive; if it is classified as negative (low ranked) it is counted as a false negative. If the molecule is inactive (negative) and the docking classifies it as negative (low ranked) it is counted as a true negative; if it is classified as positive (well ranked) it is counted as a false positive. The ROC curve describes the sensitivity as a function of the specificity of the method (docking). The sensitivity of the docking method is defined as the true recovered positive rate, and specificity is defined by true negatives divided by the sum of false positives and true negatives [53], where this area under the ROC curve (AUC) has an information value. In this case as stated by Anighoro and co-workers [54], values up to 0.5 indicate a random compound distribution, while further increasing values are indicative of active compound enrichment at high rank positions [55, 56]. It is a widely used validation tool for assessing the quality of VS protocol by monitoring the percentage of active ligands retrieved at each position of the ranking. ROC analysis can also be applied to studies of protein–ligand docking scoring [56].

### Molecular dynamics

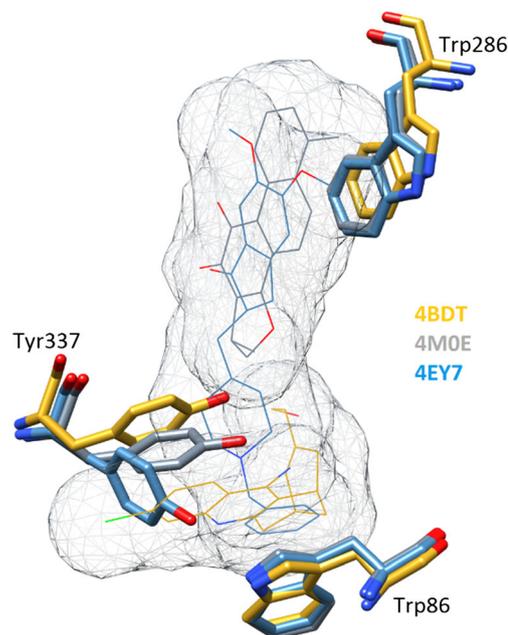
MD simulations were used to elucidate the importance of protein flexibility in ligand binding, and to explore the interactions between hAChE and the selected compounds. Only the highest scoring pose from docking results was subjected to simulations. The MD simulations described below were performed using the Amber 14 package with AMBER FF14sb force field. To carry out the simulations, the topologic and coordinate files of initial structures were built with tLEAP module of Amber 14. We have used the common technique to impose the docked complex into a truncated octahedron box using TIP3D water model. Sodium ( $\text{Na}^+$ ) counter ions were added to neutralize the net charge of protein–ligand system. Before performing the MD simulations, two minimization stages were carried out for the initial structures, which come from docking results. First, water molecules were minimized for 1000 minimization cycles with the steepest descent method, followed by 1000 cycles with the conjugate gradient method. In the second minimization, 2000 cycles were

performed using the steepest descent method for the whole system, followed by 1500 cycles of the conjugate gradient method. MD simulations steps were implemented. First, we slowly heated the system from 0 to 300 K for 1.25 ns with constant volume and no pressure control. Subsequently, density equilibration was executed with constant pressure for 1.25 ns. After that, a final equilibration at 300 K was performed for 7.5 ns. The production run was carried out at 300 K for 100 ns. In those simulations, the Shake algorithm was applied to constraint all bonds involving hydrogen, and a cutoff of 10 Å was used for all non-bonded interactions. The MD time step was taken to be 2 fs, and one snapshot was sampled every 25,000 steps. Thus, 200 conformations (frames) were obtained every 10 ns during the MD runs.

### Triplet clique search

The relative solvent-excluded surface area (relSESA) for each residue of proteins was calculated using UCSF Chimera and compared to the solvent-excluded surface area for each standard residue in Gly-X-Gly tripeptides [57]. Only hydrophobic residues (Ala, Val, Ile, Leu, Tyr, Phe, Trp) with relSESA below 0.3 was considered for further analysis. For simulation analysis, one frame to each 2 ns was considered for relSESA calculation, and those hydrophobic residues with an average below 0.3 during simulation were submitted to clique search.

A set of three residues in contact one with each other is defined as a clique. We considered as contact residues those which presented the lowest inter residue atomic distance



**Fig. 2** Aligned hAChE crystallographic structures in complex with Huprine W (4BDT) in *goldenrod*, Dihydrotanshinone I (4M0E) in *slate gray* and Donepezil (4EY7) in *steel blue*

smaller than 3.8 Å for side-chain heavy atoms, as described by Basu et al. [58, 59]. Internal vectors for each residue as well as the clique vectors were defined as described by Basu et al. [59] for simulation frames. In the case of ligand, we considered to define the vectors the nearest aromatic ring to the other two residues part of clique. The parameters tilt and swivel angles used to analyze the stability of cliques, describes respectively the orientation of each residue to the plane of clique and the rotation of residue about their internal Z axis [58, 59].

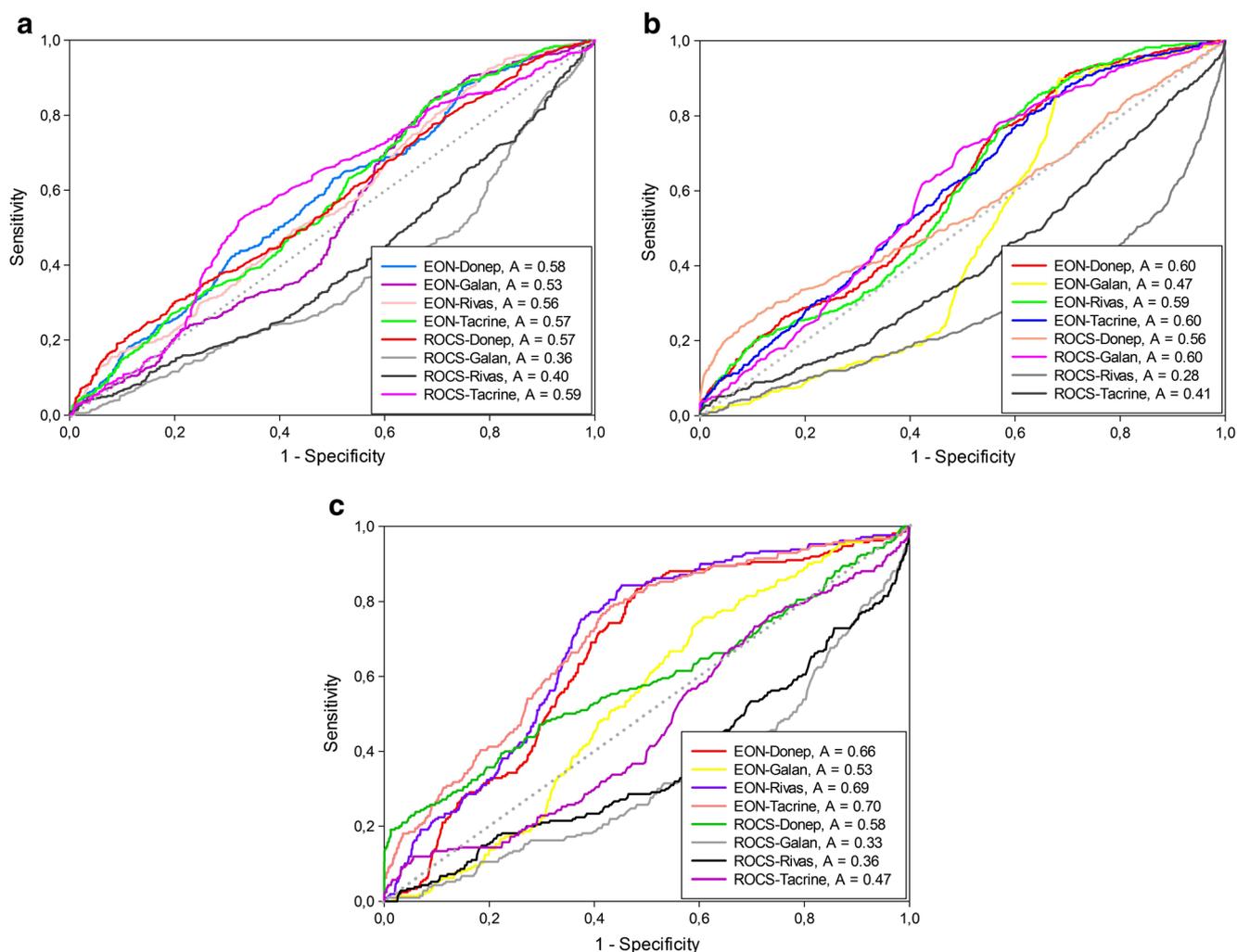
## Results and discussion

### ROCS and EON

Two computational strategies were employed: ROCS shape-based similarity analysis and EON electrostatic comparison. More specifically, ROCS is used to identify similarity between

molecules based on their 3D shape, while EON compares electrostatic potential maps of pre-aligned molecules. ST and ET metrics of the ROCS and EON packages were tested for all databases and queries (see Fig. 3). Given that the ROCS and EON results are highly dependent on query molecule, the analysis was carried out with four different queries to test the program performance consistency. The AUC varies considerably among the queries and the programs. However, the best performance of both codes showed slight difference of AUC (0.60 and 0.70, for ROCS and EON, respectively). In addition, considering both packages and the complete dataset the best result was achieved by Donepezil query, with AUC higher than 0.5 in all the ROCS curves.

Electrostatics-based screening showed suitable performance for all queries, mainly for Donepezil and Tacrine. This result of application of 3D similarity supports the use of EON code in a virtual screening project for AChE inhibitor discovery.



**Fig. 3** Area under the curve (AUC) plots using rapid overlay of chemical structures (ROCS) and EON (similarity analysis through comparison of electrostatic overlay) packages for **a** DS1, **b** DS2 and **c** DS3 datasets

## Docking

### Redocking procedure

Initially, we performed redocking calculations to reproduce the binding mode of the co-crystal structures. The ligand binding process was validated by redocking of the co-crystallized Donepezil (E20), Huprine W (HUW) and Dihydrotanshinone I (1YL) with 4EY7, 4BDT and 4M0E structures respectively. In redocking simulations, it is expected that the results generate root mean square deviation (RMSD) greater than 2 Å when compared to co-crystallized ligands [52].

We verified that the ligands bound to the hAChE active site, analogous to the co-crystallized ligand (see Fig. 4). Therefore, the results from the redocking procedure showed that the docking protocol could reproduce the binding mode of the co-crystal structures 4BDT, 4EY7 and 4M0E, with RMSD values equal to 0.89 Å, 1.34 Å and 0.82 Å. The Glide score of the complex in redocking results for 4EY7 and 4M0E is similar with values of  $-11.65 \text{ kcal mol}^{-1}$  and  $-8.52 \text{ kcal mol}^{-1}$  compared to experimental affinity of  $\Delta G_{4EY7} = -11.86 \text{ kcal mol}^{-1}$  ( $IC_{50} 0.002 \text{ }\mu\text{M}$ ) and  $\Delta G_{4M0E} = -8.50 \text{ kcal mol}^{-1}$  ( $IC_{50} 1 \text{ }\mu\text{M}$ ). For the 4BDT crystal structure the experimental affinity is  $\Delta G_{4BDT} = -12.27 \text{ kcal mol}^{-1}$  ( $IC_{50} 0.001 \text{ }\mu\text{M}$ ) which is slightly lower compared to the  $-10.73$  score achieved by Glide. Here, the scores for the redocked ligand is in perfect agreement with experimental affinity. These validate the idea that the prediction of binding modes of co-crystal structures is enough to define a

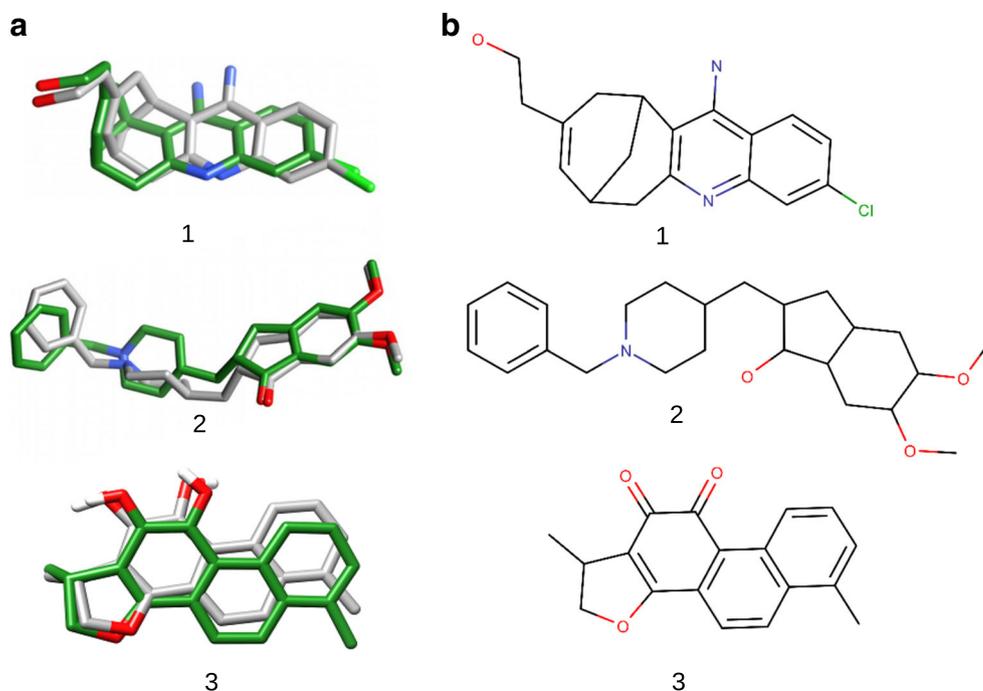
docking protocol for virtual screening. The superposition between co-crystallized and the predicted compounds is shown in the Fig. 4.

### Docking analysis

In our docking studies, the same protocol (validated using redocking) was set up to subsequently discriminate between active and inactive compounds for all datasets (DS1, DS2 and DS3) into three different structures (4EY7, 4BDT and 4M0E). After obtaining the binding mode, the compounds were ranked according to the Glidescore, considering only the best docked scores. The corresponding ROC curves are shown in Fig. 5. The program Glide achieved an AUC of 0.71. The structure 4EY7 was unable to retrieve active compounds from the DS2 database, achieving an AUC of 0.47. In the other cases, the results showed appropriate consistency performance, with 4BDT and 4M0E structures, which yielded expected enrichment of active compounds beyond random selection. This suggests that the binding conformation and binding modes of the hAChE inhibitors are satisfactory. The same analysis using SMARTS patterns of the Glide platform for the constraints and grid dimensions of  $6\text{Å} \times 6\text{Å} \times 6\text{Å}$  were carried out; however, no improvement in the enrichment data was observed.

After an exhaustive validation of different virtual screening methodologies, we have found that electrostatic-based search performed better as compared to shape-based analysis. The study also revealed that 4BDT and 4M0E structures

**Fig. 4** Redocking of co-crystallized ligands for (1) Huprine W (0.89 Å) (2) Donepezil (1.34 Å) (3) Dihydrotanshinone I (0.82 Å). **a** Superposition between co-crystallized (*gray*) and the best docking pose predicted (*green*) and **b** molecular structure



achieved the best results for structure-based method of hAChE inhibitors. It is well known that Donepezil is a dual binding site inhibitor, i.e., interacting simultaneously with the catalytic (CAS), peripheral (PAS) and mid gorge (between the CAS and the PAS) binding sites (see Fig. 6). Therefore, Donepezil should identify more effective and promising candidates for AD [60], and it was used as the query for the next steps. In addition, the 4M0E crystallography structure (with the best resolution and results from redocking process) was chosen for the search of new inhibitors of AD.

### Virtual screening for novel hAChE inhibitor scaffolds

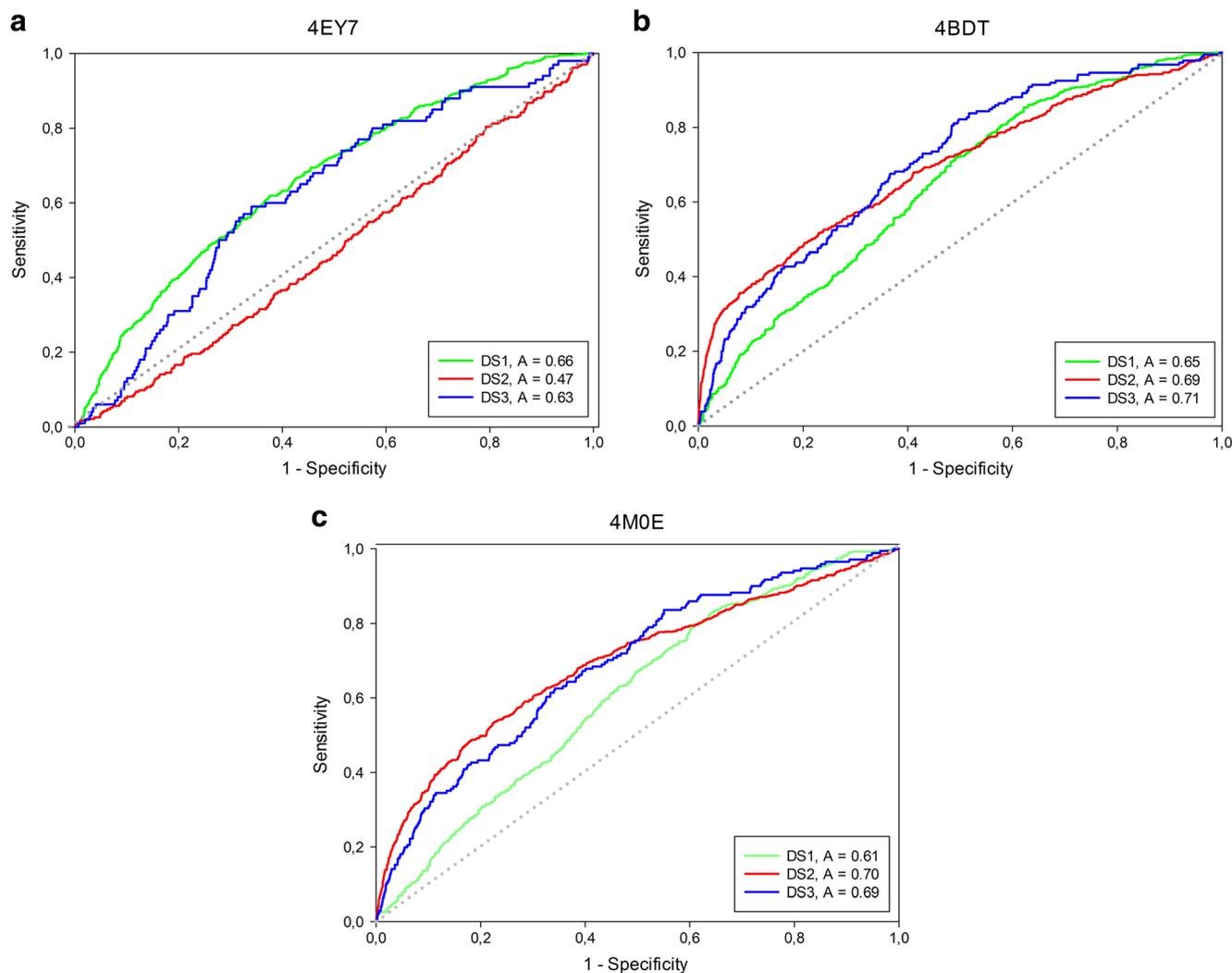
Hierarchical virtual screening procedure was employed to identify hAChE inhibitor scaffolds targeting the active site. Shortly, our hierarchical procedure consists of: Filtering ligands using Electrostatic Similarity Search;

predicting the binding poses of ligands using the Glide package, and ranking the docked ligands using Glidescore. The VS procedure was summarized in Fig. 7.

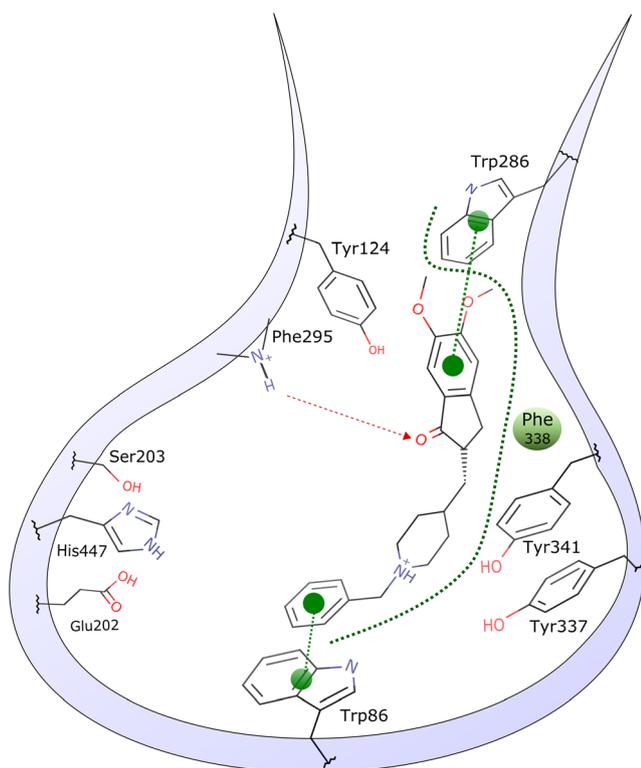
### In silico virtual screening filter

Firstly, the donepezil query from 4EY7 and ZINC database conformation was generated for electrostatic similarity search as described previously. EON was carried out against the ZINC database with 10 million commercially available compounds. The top 1 million hits, which had ET scores between 0.60 and 0.94, were saved for further analysis.

In order to filter the highlighted hits for a biological assay, the compounds were analyzed by molecular docking using Glide package. The co-crystallized ligand was extracted from 4M0E structure and the filtered

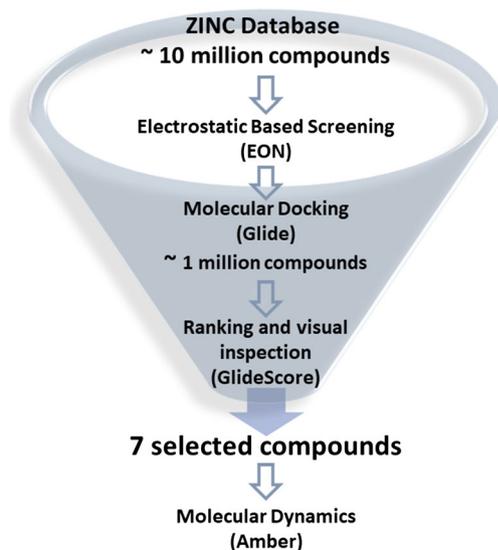


**Fig. 5** AUC plots of Glide package for **a** 4EY7, **b** 4BDT and **c** 4M0E crystallography structure



**Fig. 6** Two-dimensional representation of the interactions between the hAChE structure and donepezil. Red dashed lines Hydrogen bonds, green dashed lines with circles at the end  $\pi$ -stacking interactions, green dashed lines hydrophobic interactions

database was docked. The compounds were ranked by Glidescore. Additionally, the visual inspection helped to identify compounds that exhibited the best interactions with the binding site in hAChE. Seven hits were selected



**Fig. 7** In silico screening protocol used to identify hAChE inhibitors

after molecular docking procedures (see Fig. 8), and they were further evaluated using MD simulations.

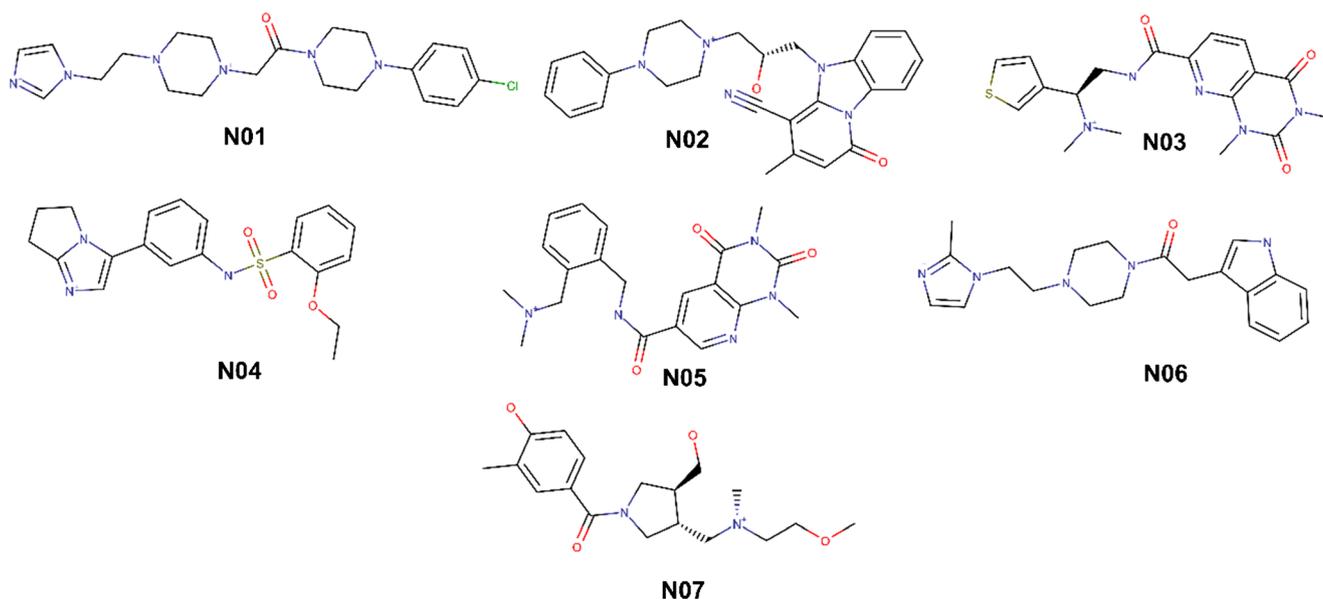
The most promising compound, N01, was selected to explore the interaction between the inhibitors and hAChE. Figure 9 presents the docking results for compound N01 in 4M0E. It suggests that some key amino acids in the binding pocket interact with the hAChE inhibitor through hydrogen bonding,  $\pi$ -stacking and cation- $\pi$ . The following interactions form hydrogen bonds: O atom of N01 (2.68 Å) with Phe295, the N atom of imidazole (1.92 Å) with His447, and the N atom of piperazine (2.15 Å) with Tyr124. The aromatic-ring side chains of Trp86 and Trp286 participate in  $\pi$ - $\pi$  interactions with imidazole and benzene, respectively.

By analyzing the interactions from the docking results (see Fig. 10) formed between all the compounds and hAChE, it was observed that some key amino acids, such as Ser293, Phe295, His447, Tyr 124, Tyr 133, form hydrogen bonds. Cation- $\pi$  and  $\pi$ -stacking interactions were observed between the ligands and Tyr337, His447, Trp86, Tyr341, Phe338, Trp286 and Tyr124. It is observed that Phe295, Tyr337, Tyr124, Trp286, Trp86 and Tyr341 play a crucial role interacting with the most of these compounds. The binding mode predicted by docking simulations, for these seven selected compounds, suggests them as potential candidates to hAChE inhibitors.

The geometric criteria behind assigning cation- $\pi$  and  $\pi$ -stacking interactions related are as follows: hydrogen bonds, the maximum distance is 2.8 Å, the minimum donor angle is 120.0°, and the minimum acceptor angle is 90.0°.  $\pi$  stacking interaction, the angle between the ring planes is less than 30° and the distance between the ring centroids is less than 4.4 Å (face-to-face), or the angle between the ring planes is between 60° and 120° and the distance between the ring centroids is less than 5.5 Å (edge-to-face). Cation  $\pi$  interaction, the maximum distance between the cation center and the ring center is 6.6 Å, and the angle between the ring plane and the line between the cation center and the ring center does not deviate from the perpendicular by more than 30°.

### Molecular dynamics

To further investigate these filtered compounds, 100 ns of MD analysis was performed for our seven selected ligands complexed with the structure of hAChE. Regarding the dynamic behavior of the more promising compound, it was observed that this ligand remained stable inside the hAChE active site. It was also observed the interaction of ligand with Trp286, Trp86, Tyr337, Tyr341, His447, Tyr124 and Phe295 as can be seen in the superposition of



**Fig. 8** Proposed structures of new potential candidate inhibitors

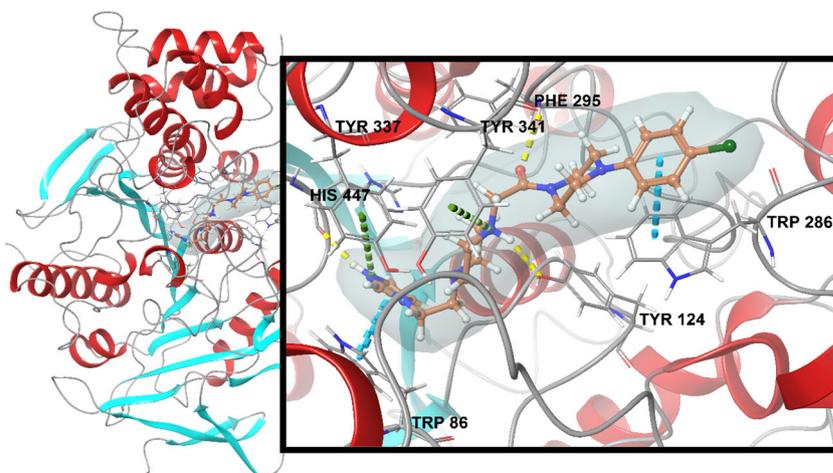
frames in Fig. 11. These observations corroborate the docking results.

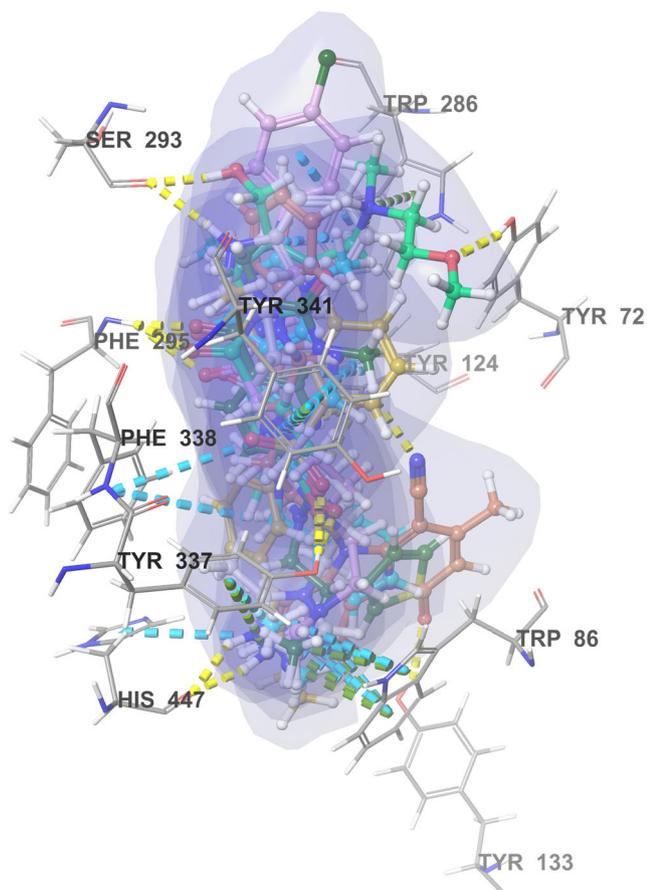
To explore the dynamic stability of the ligands, the time dependent RMSD values of the  $\alpha$ -carbon were calculated during MD trajectories of 100 ns. The RMSD results for the simulation showed that the compounds are stable after 55 ns of simulation for all cases. In the plot of temporal RMSD, presented in Fig. 12, it is also possible to observe the stable conformation of the ligands, with RMSD values never surpassing 6 Å during the simulation. The observed results for the system indicated the stability during the MD simulations, with RMSD average values from 1.8 to 2.1 Å. MD simulations optimized the interactions between hAChE and the compounds through some main change that occurred to the conformation reporting the stability.

#### Clique analysis and binding mode characterization in terms of geometrical parameters

The presence of cliques (closed subgraphs) has been thoroughly investigated to characterize the associations between buried aromatic residues present on the enzymatic active site and ligands rings in terms of geometric constraints imposed during the binary complex formation. All seven proposed ligands constituted a 'clique' with buried protein residues at least in part of simulation. While most of them were part of ephemeral cliques, in two of cases they present a good stability. The first one, molecule N05, showed the clique with residues Trp83 and Tyr334 in 80% of analyzed frames, evidencing a good aromatic interaction of ligand with these residues. In the second case, a conformational change in molecule N02 led to the formation of clique with the same residues of N05

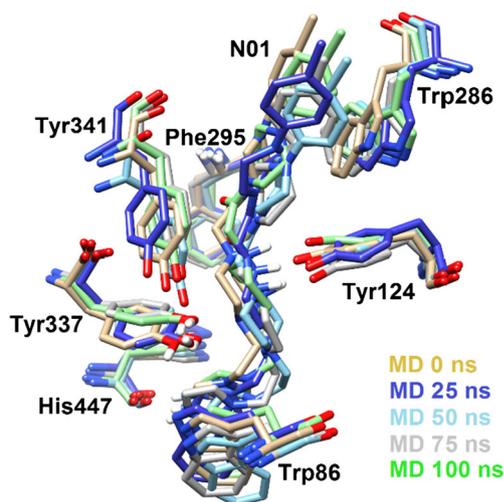
**Fig. 9** Docking results for compound N01 in the 4M0E hAChE crystallography structure. *Yellow dashed lines* Hydrogen bonds, *blue dashed lines*  $\pi$ -stacking interactions, *green dashed lines* cation- $\pi$  interactions



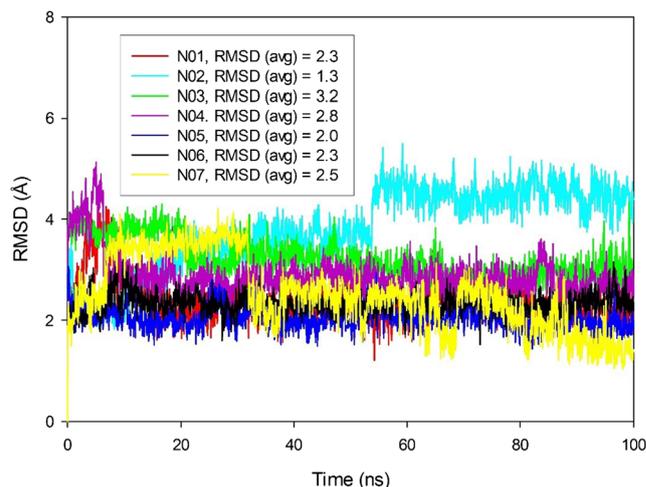


**Fig. 10** Superposition of docking results corresponding to all selected compounds inside the hAChE active site. *Yellow dashed lines* Hydrogen bonds, *blue dashed lines*  $\pi$ -stacking interactions, *green dashed lines* cation- $\pi$  interactions

one. In this case, the clique appears to be formed during the simulation and it is present in 68% of last half of analyzed frames.



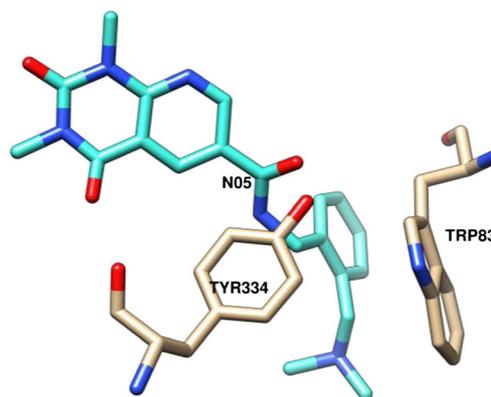
**Fig. 11** Docking results for compound N01 inside hAChE at various molecular dynamics (MD) simulation time points: 0 ns, 25 ns, 50 ns, 75 ns and 100 ns



**Fig. 12** Root mean square deviation (RMSD) of compounds versus MD simulation time

An extended analysis was carried out for AChE–N05 complex to describe the geometric patterns of the protein–ligand interaction, Fig. 13. The joining of origin point from each residue define a triangle Trp83–Tyr334–N05 of side 6.4, 4.7 and 7.5 Å and angles  $85^\circ \pm 4$ ,  $38^\circ \pm 3$  and  $57^\circ \pm 3$ .

We observed during the simulation a conformational change of Tyr334 characterized by modification of tilt angle from  $60^\circ$  to  $125^\circ$ . This same parameter was useful to show a small deviation of  $6^\circ$  for Trp83 side-chain from the preferred angle of  $66.1^\circ$ . On the other hand, through this parameter, a random orientation of aromatic ring of N05 located in the bottom of active site of AChE was observed during simulation. This may be an indicative of a lack of strong aromatic interaction between N05 and the residues Tyr and Trp. A similar behavior is observed for swivel angle, which remained stable at  $63^\circ$  and  $170^\circ$  during the simulation for the residues Trp83 and Tyr334, and had a higher fluctuation for N05 ligand.



**Fig. 13** The AChE–N05 complex and the geometric patterns of the protein–ligand interaction. The joining of origin point from each residue define a triangle Trp83–Tyr334–N05 of side 6.4, 4.7 and 7.5 Å and angles  $85.8^\circ \pm 4$ ,  $38.1^\circ \pm 3$  and  $57.1^\circ \pm 3$

## Conclusions

3D similarity search and molecular docking were successfully used to select seven new hAChE inhibitors as potential candidate inhibitors used in AD. Molecular docking and MD simulations were used to explore the interactions between hAChE and these compounds. Some key residues, such as Phe295, Tyr337, Tyr124, Trp286, Trp86 and Tyr341 at the binding site of hAChE were identified as crucial interaction to fulfill dual binding site inhibitor. In summary, all the results indicated the relevant use of EON and docking strategy for identifying novel hit compounds as promising potential anti-cholinesterase candidates. Furthermore, a stable clique among the ligands rings, Trp83 and Tyr334 stands out as a valuable possibility for the development of novel AChE inhibitors with better hydrophobic interaction and inhibition profile. In that context, in this work it was possible to suggest seven structures as potential hAChE inhibitors.

**Acknowledgments** The authors thank the Brazilian agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Apoio a Pesquisa do Distrito Federal (FAPDF) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for funding this work. We thank Professor Carlos Alberto Montanari group for providing the High-Performance Computing resource at University of São Paulo.

## References

- Scarpini E, Schelterns P, Feldman H (2003) Treatment of Alzheimer's disease; current status and new perspectives. *Lancet Neurol* 2:539–547. [https://doi.org/10.1016/S1474-4422\(03\)00502-7](https://doi.org/10.1016/S1474-4422(03)00502-7)
- Mckhann GM, Knopman DS, Chertkow H et al (2011) The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging and the Alzheimer's Association workgroup. *Alzheimers Dement* 7:263–269. <https://doi.org/10.1016/j.jalz.2011.03.005>
- Herrup K (2011) Commentary on “recommendations from the National Institute on Aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease.” addressing the challenge of Alzheimer's disease in the 21st century. *Alzheimers Dement* 7:335–337. <https://doi.org/10.1016/j.jalz.2011.04.002>
- Citron M (2002) Alzheimer's disease: treatments in discovery and development. *Nat Neurosci* 5:1055–1057. <https://doi.org/10.1038/nn940>
- Yiannopoulou KG, Papageorgiou SG (2013) Current and future treatments for Alzheimer's disease. *Ther Adv Neurol Disord* 6:19–33. <https://doi.org/10.1177/1756285612461679>
- Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT (2011) Neuropathological alterations in Alzheimer disease. *Cold Spring Harb Perspect Med* 1(1):a006189. <https://doi.org/10.1101/cshperspect.a006189>
- Reitz C, Mayeux R (2014) Alzheimer disease: epidemiology, diagnostic criteria, risk factors and biomarkers. *Biochem Pharmacol* 88:640–651. <https://doi.org/10.1016/j.bcp.2013.12.024>
- Francis PT, Palmer AM, Snape M, Wilcock GK (1999) The cholinergic hypothesis of Alzheimer's disease: a review of progress. *J Neurol Neurosurg Psychiatry* 66:137–147. <https://doi.org/10.1136/jnnp.66.2.137>
- Perry EK (1986) The cholinergic hypothesis—ten years on. *Br Med Bull* 42:63–69. <https://doi.org/10.1093/oxfordjournals.bmb.a072100>
- Dumas JA, Newhouse PA (2011) The cholinergic hypothesis of cognitive aging revisited again: cholinergic functional compensation. *Pharmacol Biochem Behav* 99:254–261. <https://doi.org/10.1016/j.pbb.2011.02.022>
- Miyashita N, Straub JE, Thirumalai D (2009) Structures of beta-amyloid peptide 1–40, 1–42, and 1–55-the 672–726 fragment of APP-in a membrane environment with implications for interactions with gamma-secretase. *J Am Chem Soc* 131:17843–17852. <https://doi.org/10.1021/ja905457d>
- Gallagher D, Coen RF, Lawlor BA (2016) Alzheimer disease. In: Hardiman O, Doherty CP, Elamin M, Bede P (eds) *Neurodegenerative disorders: a clinical guide*. Springer, Berlin, pp 57–74
- Kepp KP (2012) Bioinorganic chemistry of Alzheimer's disease. *Chem Rev* 112:5193–5239. <https://doi.org/10.1021/cr300009x>
- Small G, Bullock R (2011) Defining optimal treatment with cholinesterase inhibitors in Alzheimer's disease. *Alzheimers Dement* 7:177–184. <https://doi.org/10.1016/j.jalz.2010.03.016>
- Anand P, Singh B (2013) A review on cholinesterase inhibitors for Alzheimer's disease. *Arch Pharm Res* 36:375–399. <https://doi.org/10.1007/s12272-013-0036-3>
- Qureshi HY, Li T, MacDonald R et al (2013) Interaction of 14-3-3 $\beta$  with microtubule-associated protein tau within alzheimer's disease neurofibrillary tangles. *Biochemistry* 52:6445–6455. <https://doi.org/10.1021/bi400442d>
- Kiametis AS, Martins JBL, Romeiro LAS, Gargano R (2013) Acetylcholinesterase inhibitors: modeling potential candidates. *Int J Quantum Chem* 113:1461–1466. <https://doi.org/10.1002/qua.24341>
- Kiametis AS, Silva MA, Romeiro LAS et al (2017) Potential acetylcholinesterase inhibitors: molecular docking, molecular dynamics, and in silico prediction. *J Mol Model* 23(2):67. <https://doi.org/10.1007/s00894-017-3228-9>
- Nascimento ECM, Martins JBL (2011) Electronic structure and PCA analysis of covalent and non-covalent acetylcholinesterase inhibitors. *J Mol Model* 17:1371–1379
- De Paula AAN, Martins JBL, dos Santos ML et al (2009) New potential AChE inhibitor candidates. *Eur J Med Chem* 44:3754–3759
- Nascimento ECM, Martins JBL, dos Santos ML, Gargano R (2008) Theoretical study of classical acetylcholinesterase inhibitors. *Chem Phys Lett* 458:285–289
- De Paula AAN, Martins JBL, Gargano R et al (2007) Electronic structure calculations toward new potentially AChE inhibitors. *Chem Phys Lett* 446:304–308. <https://doi.org/10.1016/j.cplett.2007.08.055>
- Misik J, Korabecny J, Nepovimova E et al (2016) Effects of novel tacrine-related cholinesterase inhibitors in the reversal of 3-quinuclidinyl benzilate-induced cognitive deficit in rats—is there a potential for Alzheimer's disease treatment? *Neurosci Lett* 612:261–268. <https://doi.org/10.1016/j.neulet.2015.12.021>
- Colovic MB, Krstic DZ, Lazarevic-Pasti TD et al (2013) Acetylcholinesterase inhibitors: pharmacology and toxicology. *Curr Neuropharmacol* 11:315–335. <https://doi.org/10.2174/1570159X11311030006>
- Cerbai F, Giovannini MG, Melani C et al (2007) N1phenethyl-noracetylserine, a selective butyrylcholinesterase inhibitor, increases acetylcholine release in rat cerebral cortex: a comparison with donepezil and rivastigmine. *Eur J Pharmacol* 572:142–150. <https://doi.org/10.1016/j.ejphar.2007.06.053>
- Irwin JJ, Sterling T, Mysinger MM et al (2012) ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model* 52:1757–1768. <https://doi.org/10.1021/ci3001277>

27. Hawkins PCD, Skillman AG, Nicholls A (2007) Comparison of shape-matching and docking as virtual screening tools. *J Med Chem* 50:74–82. <https://doi.org/10.1021/jm0603365>
28. ROCS (2015) 3.2.0.4: Openeye Scientific Software, Santa Fe, NM
29. EON (2013) 2.2.0.5: OpenEye Scientific Software, Santa Fe, NM
30. Glide (2009) Schrödinger, LLC, New York, NY
31. Case DA, Babin V, Berryman JT et al (2014) AMBER14. San Francisco, CA
32. Gaulton A, Bellis LJ, Bento AP et al (2012) ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res* 40(Database issue):D1100–D1107 <https://doi.org/10.1093/nar/gkr777>
33. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK (2012) Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J Med Chem* 55:6582–6594. <https://doi.org/10.1021/jm300687e>
34. Lavecchia A (2015) Machine-learning approaches in drug discovery: methods and applications. *Drug Discov Today* 20:318–331. <https://doi.org/10.1016/j.drudis.2014.10.012>
35. Lavecchia A, Di Giovanni C (2013) Virtual screening strategies in drug discovery: a critical review. *Curr Med Chem* 20:2839–2860. <https://doi.org/10.2174/09298673113209990001>
36. ChemAxon (2013) Marvin sketch. In: <https://www.chemaxon.com/products/marvin/>
37. OMEGA (2013) 2.5.1.4: OpenEye Scientific Software, Santa Fe, NM
38. Hawkins PCD, Nicholls A (2012) Unknown\_Conformer generation with OMEGA learning from the data set and the analysis of failures(3). *J Chem Inf Model* 52:2919–2936. <https://doi.org/10.1021/ci300314k>
39. Catto M, Pisani L, Leonetti F et al (2013) Design, synthesis and biological evaluation of coumarin alkylamines as potent and selective dual binding site inhibitors of acetylcholinesterase. *Bioorg Med Chem* 21:146–152. <https://doi.org/10.1016/j.bmc.2012.10.045>
40. Ronco C, Foucault R, Gillon E et al (2011) New huprine derivatives functionalized at position9 as highly potent acetylcholinesterase inhibitors. *ChemMedChem* 6:876–888. <https://doi.org/10.1002/cmdc.201000523>
41. Kitchen DB, Decomez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 3:935–949. <https://doi.org/10.1038/nrd1549>
42. Wang G, Zhu W (2016) Molecular docking for drug discovery and development: a widely used approach but far from perfect. *Future Med Chem* 8:1707–1710. <https://doi.org/10.4155/fmc-2016-0143>
43. Meng X-Y, Zhang H-X, Mezei M, Cui M (2011) Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des* 7:146–157. <https://doi.org/10.2174/157340911795677602>
44. Berman HM, Westbrook J, Feng Z et al (2000) The protein data bank. *Nucleic Acids Res* 28:235–242. <https://doi.org/10.1093/nar/28.1.235>
45. Kryger G, Silman I, Sussman JL (1999) Structure of acetylcholinesterase complexed with E2020 (Aricept<sup>®</sup>): implications for the design of new anti-Alzheimer drugs. *Structure* 7:297–307. [https://doi.org/10.1016/S0969-2126\(99\)80040-9](https://doi.org/10.1016/S0969-2126(99)80040-9)
46. Schrödinger (2009) Protein preparation wizard. Schrödinger, LLC, New York, NY
47. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkh381>
48. Macromodel (2009) Schrödinger, LLC, New York, NY
49. LigPrep (2009) Schrödinger, LLC, New York, NY
50. Epik (2009) Schrödinger, LLC, New York, NY
51. Banks JL, Beard HS, Cao Y et al (2005) Integrated modeling program, applied chemical theory (IMPACT). *J Comput Chem* 26:1752–1780. <https://doi.org/10.1002/jcc.20292>
52. Friesner RA, Banks JL, Murphy RB et al (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 47:1739–1749. <https://doi.org/10.1021/jm0306430>
53. Basu S, Wallner B (2016) Finding correct protein-protein docking models using ProQDock. *Bioinformatics* 32:i262–i270. <https://doi.org/10.1093/bioinformatics/btw257>
54. Anighoro A, Bajorath J (2016) Three-dimensional similarity in molecular docking: prioritizing Ligand poses on the basis of experimental binding modes. *J Chem Inf Model* 56:580–587. <https://doi.org/10.1021/acs.jcim.5b00745>
55. Triballeau N, Acher F, Brabet I et al (2005) Virtual screening workflow development guided by the “receiver operating characteristic” curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J Med Chem* 48:2534–2547. <https://doi.org/10.1021/jm049092j>
56. Fawcett T (2006) An introduction to ROC analysis. *Pattern Recogn Lett* 27:861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
57. Bendell CJ, Liu S, Aumentado-Armstrong T et al (2014) Transient protein-protein Interface prediction: datasets, features, algorithms, and the RAD-T predictor. *BMC Bioinf* 15:82. <https://doi.org/10.1186/1471-2105-15-82>
58. Basu S, Bhattacharyya D, Banerjee R (2011) Mapping the distribution of packing topologies within protein interiors shows predominant preference for specific packing motifs. *BMC Bioinf* 12:195. <https://doi.org/10.1186/1471-2105-12-195>
59. Banerjee R, Sen M, Bhattacharyya D et al (2003) The jigsaw puzzle model: search for conformational specificity in protein interiors. *J Mol Biol* Vol 333:211–226. <https://doi.org/10.1016/j.jmb.2003.08.013>
60. Galdeano C, Viayna E, Arroyo P et al (2010) Structural determinants of the multifunctional profile of dual binding site acetylcholinesterase inhibitors as anti-Alzheimer agents. *Curr Pharm Des* 16:2818–2836