



UnB

UNIVERSIDADE DE BRASÍLIA

INSTITUTO DE FÍSICA

TESE DE DOUTORADO

MÔNICA DE ABREU SILVA

**Modelos Preditivos Baseados em Descritores
Moleculares e Modos de Interação Receptor-Ligante
para Inibidores da Acetilcolinesterase**

Brasília

2017

MÔNICA DE ABREU SILVA

**Modelos Preditivos Baseados em Descritores
Moleculares e Modos de Interação Receptor-Ligante
para Inibidores da Acetilcolinesterase**

Orientador: Prof. Dr. Ricardo Gargano

Coorientador: Dra. Alessandra Sofia Kiametis

Brasília

2017

AGRADECIMENTOS

Agradeço primeiramente à Deus pelo dom da vida, por ser meu refúgio e fortaleza.

A minha família, Painho e Mainha, por todo amor, generosidade, respeito, ensinamentos e pelo exemplo, por serem o meu poema de Drummond (**Pra sempre...***Fosse eu Rei do Mundo, baixava uma lei: Mãe e Pai não morrem nunca, mãe e pai ficarão sempre juntos do seu filho e ele, velho embora, será pequenino feito grão de milho.*) À Lôra, Nandinha, Tony e Pê, pelo incentivo e por me mostrarem, todos os dias, que a distância é ínfima diante de todo o amor que sentimos.

Aos meus “velhinhos” amados, Lita, Gerson e Dete, por tudo que representam para nossa família, por dividir comigo as histórias de uma vida sertaneja de luta e perseverança.

A Robson, por todo amor, companheirismo e respeito, seu apoio foi essencial para a conclusão desse trabalho e é essencial também na minha vida.

A família que a gente escolhe e que não precisa ter conta sanguínea, Glorinha, May, Mary, Fábio, Cláudia, Ni, Nádia e Gobira, Rose, Rosa; por me estimularem, pela horas de descontração, por todos os momentos de pura alegria que me proporcionaram.

Aos meus amigos, Dri, Taty, Lauri, Thiago, Lucélia pelos momentos compartilhados, por tornarem a vida longe de casa mais bonita e pelo carinho que sempre tiveram, vocês são muito importantes para mim. A todos os amigos que fiz, e foram muitos, muito obrigada!

Ao meu orientador, Professor Dr. Ricardo Gargano, uma das melhores pessoas que tive a oportunidade de conhecer. Muito obrigada pela confiança, pela excelente orientação, paciência, amizade, disponibilidade e motivação, toda a minha gratidão!

A minha coorientadora, Dr. Alessandra Kiametis, que esteve presente em todas as etapas deste trabalho, com muita boa vontade, paciência, competência e comprometimento, sem a qual esse trabalho não seria possível. Obrigada pelo carinho e amizade, toda a minha admiração pessoal e profissional.

Ao professor, Dr. Werner Treptow, pela confiança e colaboração, pela oportunidade que nos deu de conhecermos “um mundo novo”, que ajudou a direcionar esse trabalho! E aos alunos do Laboratório de Biologia Teórica e Computacional, em especial ao Caio, por todo auxílio computacional e disposição em ajudar nas diversas tarefas.

Aos professores Dr. Luiz Romeiro, pela colaboração e entusiasmo, Dr. João Martins e Dr. Roberto Politi, pelas discussões muito proveitosas que motivaram esse trabalho.

Ao CNPq, pelo apoio financeiro.

A todos que de forma direta ou indireta contribuíram para a realização deste trabalho, meus sinceros agradecimentos ... ninguém faz nada sozinho!

“O que sabemos é uma gota; o que ignoramos é um oceano.”

(Isaac Newton)

RESUMO

O Alzheimer é uma doença neurodegenerativa progressiva que compromete o sistema cognitivo e causa demência. Em geral, a doença de Alzheimer (DA) afeta pessoas com mais de 65 anos de idade e projeções futuras indicam que a população com DA irá aumentar significativamente, causando forte impacto social. Os fármacos atualmente comercializadas somente amenizam a progressão da doença. Nesse sentido, a busca por novos fármacos é um tema relevante na química medicinal. O objetivo principal deste trabalho é a modelagem molecular de novos agentes terapêuticos para o tratamento da DA. A estratégia terapêutica aqui adotada é a hipótese colinérgica. Um conjunto de candidatos a inibidores da acetilcolinesterase foi desenhado a partir de um composto natural obtido a partir do líquido da casca da castanha de caju (cardanol) com o intuito de modular positivamente a ação colinérgica através da inibição da acetilcolinesterase. Uma triagem *in silico* foi realizado baseado em descritores eletrônicos e relação quantitativa de estrutura-atividade para melhor compreender o perfil inibitório desses candidatos. A análise de componentes principais foi capaz de agrupar ligantes com perfil inibitório similar mostrando que esses descritores poderiam guiar o planejamento de novos agentes anticolinérgicos. O conhecimento dos modos de interação entre o ligante o sítio ativo biológico é essencial para o planejamento de fármacos. Com o intuito de investigar essas interações, simulações de dinâmica molecular e *docking* foram empregadas. A caracterização detalhada com movimento coordenado intrínseco do sítio ativo, o qual desempenha importante papel para o mecanismo de interação receptor-ligante, é apresentada aqui. Por meio do tratamento estatístico, os modos de interação mais favoráveis revelaram como o ambiente proteico foi modificado na presença dos ligantes.

ABSTRACT

Alzheimer is a neurodegenerative progressive disease which compromises the cognitive system and causes dementia. In general, the Alzheimer disease (AD) affects elderly people over 65 years and the future projections indicate that population with AD will increase significantly, causing a huge social impact factor. The drugs currently marketed only smooth the progression of this disease and present adverse effects. In this sense, the search of the new drugs is a relevant theme to the medicinal chemistry. The main goal of this work is the molecular modeling of new therapeutic agents for treating AD. The therapeutic strategy adopted here is the cholinergic hypothesis. A set of acetylcholinesterase inhibitor candidates was designed from a natural compound obtained from the cashew nutshell liquid (cardanol) in order to modulate positively the cholinergic function through the inhibition of the acetylcholinesterase. An *in silico* screening was performed based on electronic descriptors and quantitative structure- relationship activity to better understand the inhibitory profile of these candidates. Principal component analysis was able to cluster ligands with similar inhibitory profile showing that these descriptors could guide new anticholinergic drug design. The knowledge of the interaction modes between the ligand and biological active site is essential for the drug design. In order to investigate these interactions, molecular dynamics combined with molecular docking simulations were employed. A detailed characterization of the intrinsic coordinated movement of the active site, which plays an important role for the receptor-ligand interaction mechanism, is presented here. From a statistical treatment, the most favourable interaction modes revealed how the protein environment was modified in the presence of the ligands.

LISTA DE FIGURAS

Figura 1 – Agregação do peptídeo β -amiloide.	17
Figura 2 – Hiperfosforilação da proteína tau.	17
Figura 3 – Sinapse colinérgica-Liberação da ACh.	18
Figura 4 – Estrutura tridimensional da AChE humana (PDB:4EY7, cadeia A), as 12 folhas- β representadas na cor amarela, e as 14 α -hélices, na cor roxa.	20
Figura 5 – Gorge da AChE humana (código PDB:4EY7) com os sítios aromático (laranja), aniônico (amarelo), oxianiônico (púrpura), acil (verde), periférico (azul) e catalítico (vermelho). A região transparente, refere-se ao Omega loop. Em destaque, figura superior à direita, as estruturas dos resíduos de cada sítio, com sua respectiva cor.	21
Figura 6 – Portas alternativas existentes na AChE, “ <i>main door</i> ”, “ <i>back door</i> ” e “ <i>side door</i> ”, a superfície em vermelho é a representação para a Ser234, aminoácido pertencente à tríade catalítica, para ilustrar que todas as portas levam ao sítio catalítico da Gorge.	23
Figura 7 – Fármacos utilizados no tratamento da DA.	24
Figura 8 – Interações realizadas pelo Donepezil no sítio da AChE humana (PDB:4EY7).	25
Figura 9 – Estrutura química dos lipídeos não isoprenóides presentes no LCC.	27
Figura 10 – Estrutura química do cardanol e pontos de alterações para o desenho racional de derivados do cardanol.	27
Figura 11 – Estrutura molecular do grupo de derivados aminoderivados heterocíclicos.	29
Figura 12 – Estrutura molecular do grupo de derivados piperazínicos e arilpiperazínicos.	30
Figura 13 – Estrutura molecular do grupo de derivados etanolamínicos N-Alquil ou N-benzilssubstituídos.. . . .	31
Figura 14 – Estrutura molecular do grupo de derivados benzilamínicos.	32
Figura 15 – Visualização gráfica da variância das PCs	55
Figura 16 – Caixa de Simulação de Dinâmica Molecular contendo o solvente e a proteína (AChE: 4EY7), com 57480 átomos no total, dos quais 49149 correspondem aos átomos das moléculas de água, 100 são átomos das moléculas de NaCl e 8231 é a quantidade de átomos da proteína.	69

Figura 17 – Gráfico mostrando a disposição dos compostos LDT140(1), LDT141(2), LDT142(3) e LDT144(4) nas duas primeiras componentes principais-PC1 e PC2.	71
Figura 18 – Superfície de densidade eletrônica dos orbitais HOMO e LUMO para os aminoderivados do cardanol.	74
Figura 19 – Superfície de densidade eletrônica dos orbitais HOMO-1 e LUMO+1 para os aminoderivados do cardanol.	74
Figura 20 – Mapa de potencial eletrostático (do inglês , <i>molecular electrostatic potential</i> -MEP) para os aminoderivados do cardanol.	76
Figura 21 – Gráfico mostrando a disposição dos compostos LDT155(1), LDT156(2), LDT157(3), LDT158(4), LDT475(5) e LDT489(6) nas duas primeiras componentes principais-PC1 e PC2.	78
Figura 22 – Superfície de densidade eletrônica dos orbitais HOMO e LUMO para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos.	81
Figura 23 – Superfície de densidade eletrônica dos orbitais HOMO-1 e LUMO+1 para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos.	82
Figura 24 – Mapa de potencial eletrostático (do inglês , <i>molecular electrostatic potential</i> -MEP), para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos.	83
Figura 25 – Gráfico mostrando a disposição dos compostos LDT159(1), LDT160(2), LDT161(3), LDT167(4) e LDT169(5) nas duas primeiras componentes principais-PC1 e PC2.	84
Figura 26 – Superfície de densidade eletrônica dos orbitais HOMO e LUMO, para os derivados benzilamínicos do cardanol.	87
Figura 27 – Superfície de densidade eletrônica dos orbitais HOMO-1 e LUMO+1, para os derivados benzilamínicos do cardanol.	88
Figura 28 – Mapa de potencial eletrostático (do inglês , <i>molecular electrostatic potential</i> -MEP), para os derivados benzilamínicos do cardanol.	89
Figura 29 – Desvio quadrático médio (do inglês <i>root mean square deviation RMSD</i>) da AChE (4EY7) humana durante 206 ns de simulação de DM.	92

Figura 30 – RMSD dos aminoácidos da Gorge da AChE (4EY7) humana (Tyr103, Asp105, Trp117, Gly152, Gly153, Tyr155, Tyr164, Ser234, Ala235, Trp267, Glu316, Trp317, Phe326, Phe328, Glu365, Tyr368, Tyr372 e His478) durante 200 ns de simulação de DM.	93
Figura 31 – Medida de distância entre os aminoácidos Phe328 pertencente ao sub sitio Acil e Tyr372 pertencente ao PAS, a partir do centro massa de cada um dos aminoácidos, durante 200 ns de simulação de DM.	95
Figura 32 – Posicionamento dos aminoácidos Phe328 e Tyr372 na Gorge (superfície molecular na cor verde) e distâncias assumidas entre dois átomos específicos (Phe328:CE1, Tyr372:CE1) destes dois aminoácidos em um <i>frame</i> da trajetória. Os átomos de carbono e hidrogênio estão coloridos de ciano e branco, respectivamente, enquanto os átomos de nitrogênio estão na cor azul e os de oxigênios em vermelho.	96
Figura 33 – Medida de distância entre os aminoácidos Trp317 e Tyr372 pertencentes ao PAS, a partir do centro massa de cada um dos aminoácidos, durante 200 ns de simulação de DM.	97
Figura 34 – Posicionamento dos aminoácidos Trp317 e Tyr372 na Gorge (superfície molecular na cor verde) e distâncias assumidas entre dois átomos específicos (Phe328:CE1, Tyr372:CE1) destes dois aminoácidos em um <i>frame</i> da trajetória. Os átomos de carbono e hidrogênio estão coloridos de ciano e branco, respectivamente, enquanto os átomos de nitrogênio estão na cor azul e os de oxigênios em vermelho.	98
Figura 35 – Medida de distância entre os aminoácidos Tyr155 e Tyr372 pertencentes ao PAS, a partir do centro massa de cada um dos aminoácidos, durante 200 ns de simulação de DM.	99
Figura 36 – Posicionamento dos aminoácidos Tyr155 e Tyr372 na Gorge (superfície molecular na cor verde) e distâncias assumidas entre dois átomos específicos (Tyr155:OH, Tyr372:OH) destes dois aminoácidos em um <i>frame</i> da trajetória. Os átomos de carbono e hidrogênio estão coloridos de ciano e branco, respectivamente, enquanto os átomos de nitrogênio estão na cor azul e os de oxigênios em vermelho.	100

- Figura 37 – Representação da construção do perfil da Gorge traçado pelo programa HOLE: A Gorge é mostrada na superfície na cor vermelha, e no centro da Gorge um vetor na direção do eixo Z do sistema. A superfície na cor verde, criada no programa VMD, é o resultado gerado pelo programa HOLE, a partir do raio da Gorge em cada ponto do vetor no eixo Z, ao lado direito dessa superfície a visualização gráfica dos resultados do perfil da Gorge fornecidos pelo HOLE. 102
- Figura 38 – Curvas para os perfis (raio da Gorge ao longo do eixo Z, onde o ponto 0Å representa a região de estreitamento da Gorge e nas proximidades de 10Å a *main door*) dos estados: *CloseOpen*, curva na cor turquesa; *OpenClosed*, curva na azul; *Closed*, curva na cor verde; *Open*, curva na cor preta. Curva para o perfil do cristal da hAChE 4EY7, linha tracejada na cor vermelha. 103
- Figura 39 – Mapa volumétrico da AChE, nos estados *Closed*, (lado esquerdo) e *Open* (lado direito). Os pontos mais azuis indicam maior volume de conjunto de dados e pontos mais vermelhos menor volume no conjunto de dados 105
- Figura 40 – Superposição das estruturas do Donepezil no cristal, representada nas cores ciano (carbonos), vermelha (oxigênio) e azul (nitrogênio); e no *ensemble* representada na cor amarela. 106
- Figura 41 – Nuvem de soluções para o *Docking* do Donepezil no sítio ativo da AChE (código PDB:4EY7) contendo 842 conformações do Donepezil e duas orientações preferenciais. A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde. 107
- Figura 42 – Nuvem de soluções para o *Docking* do ligante LDT160 no sítio ativo da AChE (código PDB:4EY7) contendo 987 conformações do Donepezil e duas orientações preferenciais. A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde. 108
- Figura 43 – Nuvem de soluções para o *Docking* do ligante LDT161 no sítio ativo da AChE (código PDB:4EY7) contendo 994 conformações do Donepezil e duas orientações preferenciais. A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde. 109

Figura 44 – Nuvem de soluções para o <i>Docking</i> do ligante LDT167 no sítio ativo da AChE (código PDB:4EY7 contendo 960 conformações do Donepezil e duas orientações preferenciais). A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde.	110
Figura 45 – Nuvem de soluções para o <i>Docking</i> do ligante LDT169 no sítio ativo da AChE (código PDB:4EY7) contendo 997 conformações do Donepezil e duas orientações preferenciais. A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde.	111
Figura 46 – Gráfico do RMSD versus Energia, para os resultados de <i>Docking</i> do Donepezil, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas pelo Donepezil.	113
Figura 47 – Gráfico do RMSD versus Energia, para os resultados de <i>Docking</i> do ligante LDT160, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas por esse ligante.	114
Figura 48 – Gráfico do RMSD versus Energia, para os resultados de <i>Docking</i> do ligante LDT161, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas por esse ligante.	115
Figura 49 – Gráfico do RMSD versus Energia, para os resultados de <i>Docking</i> do ligante LDT167, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas por esse ligante.	116
Figura 50 – Gráfico do RMSD versus Energia, para os resultados de <i>Docking</i> do ligante LDT169, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas por esse ligante.	117
Figura 51 – Sobreposição do ligante LDT161, como exemplo de parte do procedimento adotado para a investigação, a partir da construção das densidades, de grupos moleculares da proteína favoráveis a interações com o ligante.	118
Figura 52 – Densidade de probabilidade de encontrar resíduos da proteína que contenham partes aromáticas próximo dos ligantes LDT160, LDT161, LDT167, LDT169 e para o Donepezil, para a primeira população de soluções, representada pela isosuperfície amarela.	121

Figura 53 – Densidade de probabilidade de encontrar átomos de hidrogênio da proteína próximo dos ligantes LDT160, LDT161, LDT167, LDT169 e para o Donepezil, representada pela isosuperfície cinza.	122
Figura 54 – Densidade de probabilidade de encontrar átomos de nitrogênio da proteína próximo dos ligantes LDT160, LDT161, LDT167, LDT169 e para o Donepezil, representada pela isosuperfície azul.	124
Figura 55 – Densidade de probabilidade de encontrar átomos de oxigênio da proteína próximo dos ligantes LDT160, LDT161, LDT167, LDT169 e para o Donepezil, representada pela isosuperfície vermelha.	125
Figura 56 – Conformações assumidas pela molécula LDT167 nas soluções de <i>Docking</i> . No centro, a proteína mostrando a orientação preferencial dessa molécula, grupo amina voltado para o fundo da Gorge, e em volta, as posições encontradas com o sistema no mesmo plano.	127
Figura 57 – Exemplo de conformação da AChE, equilibrada em solução. A Gorge é mostrada na isosuperfície verde, com seus aminoácidos. Representado como esferas de <i>van der Waals</i> , no interior da Gorge, as três moléculas de água, que permanecem na Gorge durante a dinâmica.	133
Figura 58 – Separação das regiões da proteína em solução, complexada com o ligante LDT161. A região menor, delimitada pelo retângulo vermelho é a região QM, e a região maior, delimitada pelo retângulo azul é a região MM.	134

LISTA DE TABELAS

Tabela 1 – Tabela com os valores do IC_{50} e inibição para a classificação da atividade biológica dos aminoderivados heterocíclicos	28
Tabela 2 – Tabela com os valores do IC_{50} e inibição para a classificação da atividade biológica dos derivados piperazínicos e arilpiperazínicos.	29
Tabela 3 – Tabela com os valores do IC_{50} e inibição para a classificação da atividade biológica dos derivados benzilamínicos.	31
Tabela 4 – Peso das variáveis mais significativas para três primeiras componentes principais, para os aminoderivados heterocíclicos.	72
Tabela 5 – Variâncias acumuladas pelas três primeiras componentes principais, para os aminoderivados heterocíclicos.	72
Tabela 6 – Valores das propriedades Volume (Å^3), HOMO-1(eV), HOMO(eV) e LUMO(eV), utilizadas na PCA para os aminoderivados heterocíclicos.	73
Tabela 7 – Peso das variáveis mais significativas nas três primeiras componentes principais, para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos.	78
Tabela 8 – Variâncias acumuladas pelas três primeiras componentes principais, para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos.	79
Tabela 9 – Valores das propriedades LUMO (eV), LUMO+1 (eV), pKa e logD, utilizadas na PCA dos derivados etanolamínicos N-benzila ou N-alquilassubstituídos.	80
Tabela 10 – Peso das variáveis mais significativas para as três primeiras componentes principais para os derivados benzilamínicos.	85
Tabela 11 – Variâncias acumuladas pelas três primeiras componentes principais, para os derivados benzilamínicos.	85
Tabela 12 – Valores das propriedades HOMO-1(eV), HOMO(eV), LUMO(eV) e LUMO+1(eV), utilizadas na PCA para os derivados benzilamínicos.	86

SUMÁRIO

1	Introdução	16
1.1	<i>Estrutura da Acetilcolinesterase Humana</i>	19
1.2	<i>Inibidores da Acetilcolinesterase</i>	23
1.3	<i>Lipídeos Fenólicos presentes no Líquido da Castanha do Caju</i>	26
2	Metodologias	33
2.1	<i>A Equação de Schrödinger Eletrônica</i>	33
2.1.1	Solução da Equação de Schrödinger Eletrônica	36
2.1.2	Equação de Hartree-Fock	39
2.1.3	A equação de Hartree-Fock-Roothan	41
2.2	<i>Teoria do Funcional da Densidade</i>	43
2.2.1	Funcional de Troca e Correlação	46
2.2.2	Conjunto de Funções de Bases	47
2.3	<i>Descritores Moleculares</i>	48
2.4	<i>Análise de Componentes Principais</i>	52
2.5	<i>Dinâmica Molecular Clássica</i>	56
2.5.1	Funções de Energia Potencial	58
2.5.2	Simulação de Dinâmica Molecular (DM)	60
2.6	<i>Docking Molecular</i>	65
2.6.1	Funções de <i>Scoring</i>	66
3	Resultados e Discussões	70
3.1	<i>Propriedades Estruturais e Eletrônicas para os derivados do Cardanol</i>	70
3.1.1	Resultados PCA para os aminoderivados heterocíclicos	71
3.1.2	Resultados PCA para os derivados piperazínicos	76
3.1.3	Resultados de PCA para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos	77
3.1.4	Resultados de PCA para os derivados benzilamínicos	83
3.2	<i>Resultados da Dinâmica Molecular</i>	90
3.2.1	Equilibração do Sistema	91

3.2.2	Dinâmica dos resíduos da Gorge	94
3.2.3	Resultados de Docking receptor-ligante	105
4	Conclusões	129
5	Perspectivas	132
5.1	<i>Docking molecular na presença de água</i>	132
5.2	<i>Cálculos com método híbrido QM/MM</i>	133
	Referências	135
	Apêndice A – Potential acetylcholinesterase inhibitors: molecular docking, molecular dynamics, and <i>in silico</i> prediction	143

1 INTRODUÇÃO

A expectativa de vida tem aumentado significativamente nos últimos anos, o que torna a população idosa demograficamente maior em todo o mundo. Segundo o relatório da ADI (do inglês, "*Alzheimer's Disease International* ") de 2013 confeccionado com dados coletados em 2010, apontam que nos próximos 40 anos o número de idosos que necessitam de cuidados irá quase triplicar, de 101 milhões em 2010 para 277 milhões em 2050, dos quais quase metade estarão acometidos por um tipo de demência, aumentando expressivamente o número de pessoas com a doença de Alzheimer (DA). De acordo com a Organização Mundial de Saúde (OMS), até 2050 serão cerca de 115,4 milhões de pessoas com DA (PRINCE et al., 2013).

A DA é uma alteração neurodegenerativa progressiva e de natureza crônica que compromete o sistema cognitivo, sendo o tipo de demência mais comum com incidência de 50% a 75% dos casos. A demência é caracterizada pelo declínio na memória, na linguagem e em outras habilidades cognitivas que afetam a capacidade de realizar atividades cotidianas. Seus primeiros sintomas aparecem habitualmente depois dos 60 anos de idade, atingindo 5% dos indivíduos com mais de 65 anos e 20% dos indivíduos com mais de 80 anos. Atualmente, a DA é considerada uma doença multifatorial da qual o fator hereditário contribui com aproximadamente 15 % dos casos diagnosticados (MOSCONI et al., 2008; ASSOCIATION et al., 2016).

Essa doença causa mudanças neuropatológicas e bioquímicas por alterações estruturais, ocasionando distúrbios de cognição e memória relacionadas ao estresse neural. As características do cérebro de pacientes com Alzheimer são: a agregação do peptídeo β -Amiloide ($A\beta$)¹ e formação de emaranhados neurofibrilares, como ilustrado na Figura 1. O acúmulo de $A\beta$ é considerado uma das principais causas da DA, e essa hipótese propõe o aumento da produção e acumulação de peptídeo β -amiloide que leva a agregados provocando a morte de células neuronais (VALASANI et al., 2014; FLÔRES; QUARANTINI, 2011).

¹ O peptídeo β -amilóide possui uma cadeia de tamanho variável, de 39-43 aminoácidos é proveniente da Proteína Precursora Amiloide (do inglês, "Amyloid precursor protein-APP"), que está presente em várias células e tecidos do organismo, incluindo os neurônios (HOWLETT et al., 1995).

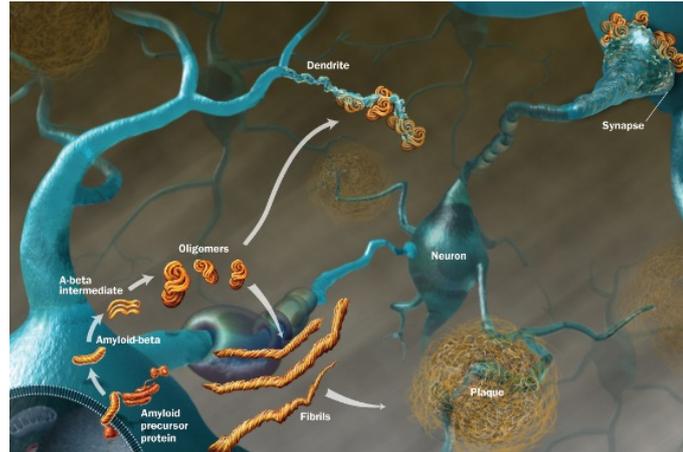


Figura 1 – Agregação do peptídeo β -amiloide.

Adaptado de CNSforum:Image Bank

Outra hipótese para o desenvolvimento da doença de Alzheimer envolve a proteína Tau. No cérebro com Alzheimer, o defeito na proteína Tau leva a emaranhados neurofibrilares intraneuronal de filamentos helicoidais emparelhados (do inglês, “*paired helical filaments*”-PHF) que podem apresentarem-se misturados com filamentos retos (do inglês, *straight filaments*-SF). Esses emaranhados de filamentos hiperfosforilados (Figura 2) também são observados em regiões em torno do núcleo das placas β -amiloide. A degeneração neurofibrilar causada pela hiperfosforilação da Tau é visualizada em exames clínicos da DA (IQBAL et al., 2010).

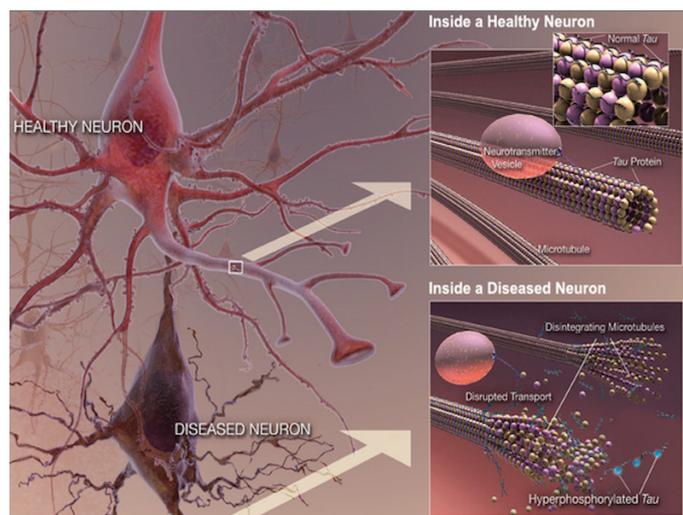


Figura 2 – Hiperfosforilação da proteína tau.

Adaptado de CNSforum:Image Bank

Uma das hipóteses mais antigas que descrevem o Alzheimer é a hipótese colinérgica. Investigações bioquímicas sistemáticas em cérebros de pacientes com Alzheimer, entre os anos 1960 e 1970, procuravam explicações neuroquímicas que identificassem a DA. A resposta para essas investigações foi dada em meados de 1970 quando informações neuroquímicas para grandes deficiências neocorticais de proteína colina acetiltransferase (do inglês, *choline acetyltransferase* -ChAT), foram claramente identificadas em relatórios médicos de pacientes com DA (FRANCIS et al., 1999).

De maneira mais detalhada, a hipótese neuroquímica da DA consiste no efeito colinérgico no qual, na região *nucleus basalis* de Meynert, ocorre a produção da enzima ChAT, que cataliza a síntese da acetilcolina (ACh) a partir da colina (Ch) e da coenzima A (CoA). Assim que formada, a ACh é liberada na fenda sináptica atuando como agonista nos receptores nicotínico ou muscarínico. A ACh é hidrolisada em colina e acetato pela enzima acetilcolinesterase (AChE) de acordo com a Figura 3. A DA altera a síntese da ChAT, conseqüentemente diminuindo a síntese da acetilcolina em face da atrofia no *nucleus basalis* de Meynert. Adicionalmente a formação da proteína beta-amiloide, hiperfosforilação da proteína Tau e causando também a restrição da captação de colina e aumento da atividade da AChE (HACHISU et al., 2015).

A relação entre a disfunção colinérgica e o progressivo agravamento da DA fornece uma justificativa para o uso terapêutico de inibidores de AChE (do inglês, “*Acetylcholinesterase inhibitors*”-AChEIs). A inibição da AChE foi comprovada como método bem sucedido para amenizar os sintomas cognitivos e comportamentais de pacientes com Alzheimer.

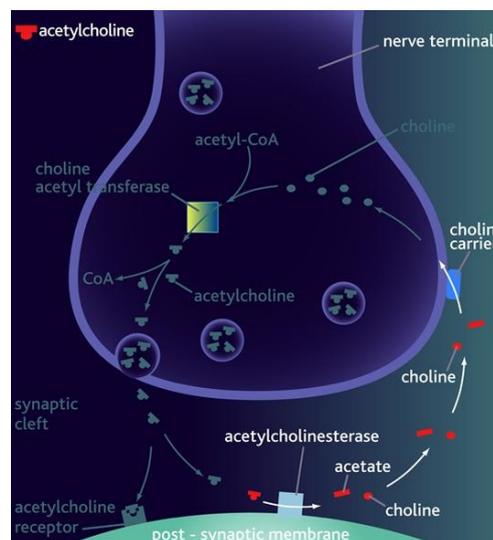


Figura 3 – Sinapse colinérgica-Liberação da ACh.

A AChE (acetilcolinesterase ou acetilcolina acetil-hidrolase),—enzima multifuncional presente nos sistemas nervosos central (SNC) e periférico (SNP) podendo ser encontrada na forma solúvel ou ligada à membrana —é caracterizada pela sua atividade de esterase na conclusão da transmissão de impulsos nervosos no processo de sinapse colinérgica pela hidrólise da ACh (BUTINI et al., 2008).

1.1 *Estrutura da Acetilcolinesterase Humana*

Várias estruturas tridimensionais para a AChE, determinadas por cristalografia de raios-X, podem ser encontradas no *Protein Data Bank* (BERMAN et al., 2000). No total são 165 estruturas de vários organismos diferentes (*Mus musculus*, *Torpedo californica*, *Homo sapiens*, *Drosophila*), das quais 20 são humanas. No entanto, muitas estruturas não estão completas por alguma falha na cristalização. Até o momento existem 5 estruturas da AChE humana com o núcleo catalítico bem resolvido, porém nem todas foram complexadas com fármacos.

Na Figura 4 apresentamos uma estrutura de AChE humana recombinante (rhAChE) de código PDB:4EY7 cadeia A, com resolução de 2,15 Å, complexada com Donepezil. Nessa estrutura (rhAChE), a conformação com a qual o Donepezil é reconhecido, é significativamente diferente se comparada ao complexo com a proteína do *Torpedo californica* AChE (TcAChE), evidenciando que a enzima humana é mais precisa para o estudo da interação do fármaco (CHEUNG et al., 2012).

A AChE (PDB:4EY7) é composta por 537 aminoácidos (8231 átomos), com 14 α -hélices e 12 folhas- β , representadas Figura 4 abaixo.

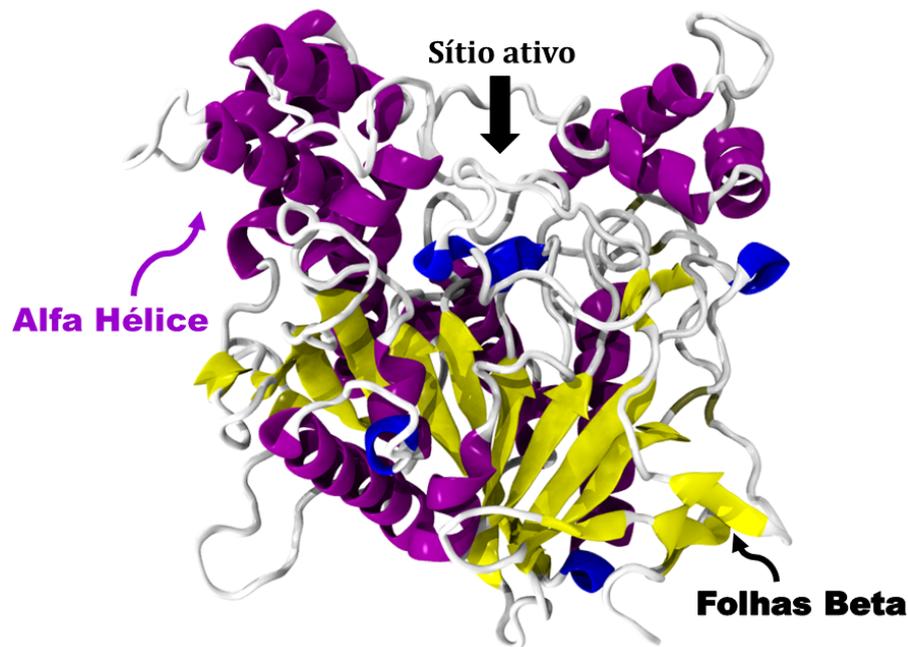


Figura 4 – Estrutura tridimensional da AChE humana (PDB:4EY7, cadeia A), as 12 folhas- β representadas na cor amarela, e as 14 α -hélices, na cor roxa.

O sítio ativo da AChE consiste numa tríade catalítica (Ser234, Glu365 e His478, sequência da 4EY7), localizada no fundo de um “gargalo”, chamado de Gorge que tem cerca de 20 Å de profundidade. O acesso à tríade catalítica é regulada pela “abertura” do subconjunto de 14 resíduos aromáticos presentes na Gorge que criam um estreitamento (do inglês, “*bottleneck*”) a 10 Å da superfície da proteína. Além disso sua secção transversal no ponto mais estreito é menor do que diâmetro do grupo quaternário de colina, indicando que a AChE deve sofrer movimentos dominantes de “respiração” *do inglês*, “*breathing*”, para realizar sua função catalítica (XU et al., 2008).

No fundo da Gorge também encontra-se o sítio aniônico (Trp117 e proximidades) que é o sítio de ligação da colina que possibilita contatos importantes para a parte estrutural da colina positivamente carregada. O bolso de ligação acila (Acil, Trp267, Phe326, Phe328) é fundamental para a estabilização do grupo funcional acetila ($-COCH_3$) da ACh. O sítio oxianiônico (átomo de nitrogênio da cadeia principal, Gly152, Gly153 e Ala235) exerce função relevante na estabilização para estados de transição de acilação por meio da preservação de carga negativa sobre o oxigênio da carbonila. Esses locais são cruciais para catálise da enzima e suas funções captadas nas informações estruturais são obtidas a partir da cristalografia de raios-X da enzima (BENNION et al., 2015).

Como meio para atrair acetilcolina, os aminoácidos da proteína são reposicionados para criar um campo elétrico que atrai o substrato carregado positivamente para a entrada “gargalo” do sítio ativo. Localizado na entrada da Gorge, que chamamos de *main door*, encontra-se o sítio de ligação aniônico periférico (do inglês, *peripheral anionic binding site-PAS*), (Tyr103, Asp105, Tyr155, Glu316, Trp317 e Tyr372) que promove o controle alostérico do sítio ativo. A última, e não menos importante região, é o sítio aromático (Tyr164 e Tyr368) que, juntamente com a Tyr155 (pertencente ao PAS) controlam o fluxo de substratos no sítio catalítico (RADIĆ; REINER; TAYLOR, 1991; ARIEL et al., 1998).

A aproximadamente 90° da *main door* encontra-se o *omega loop* que também está envolvido nos movimentos da proteína contribuindo para o alargamento momentâneo do estreitamento da Gorge, o que possivelmente aumenta o acesso do substrato e saída do produto. O *omega loop* é composto pelos resíduos Cys100 ao Cys127 (SHI et al., 2003).

Os sítios aromático, aniônico, oxianiônico, acil, periférico e catalítico, e seus respectivos resíduos aromáticos, são mostrados na Figura 5, a seguir. Nesta Figura também pode ser vista a superfície que corresponde ao *omega loop*.

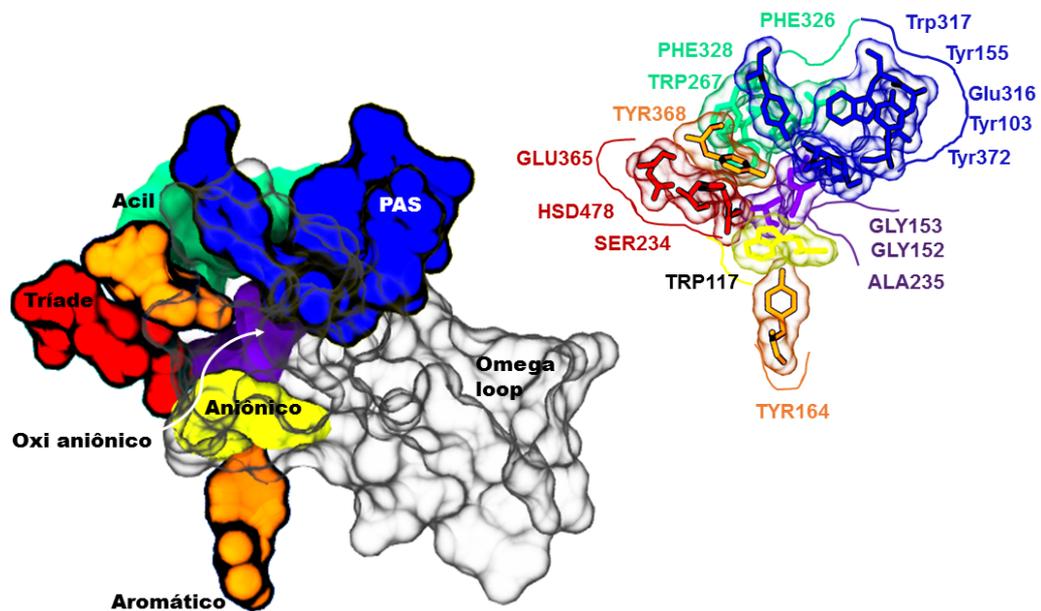


Figura 5 – Gorge da AChE humana (código PDB:4EY7) com os sítios aromático (laranja), aniônico (amarelo), oxianiônico (púrpura), acil (verde), periférico (azul) e catalítico (vermelho). A região transparente, refere-se ao Omega loop. Em destaque, figura superior à direita, as estruturas dos resíduos de cada sítio, com sua respectiva cor.

Muitos estudos tem revelado a alta eficiência catalítica da AChE, a causa mais provável pode ser a existência de portas alternativas, além da porta principal que pode levar o fármaco até o alvo catalítico. Essas portas alternativas, podem ser canais para moléculas pequenas, como a água, de maneira que a entrada de reagentes e saída de produtos pode ser feita de forma mais equilibrada. Até o momento, três portas alternativas foram sugeridas, a *main door*, *back door* e *side door* (FANG et al., 2011; SANSON et al., 2011; TARA; STRAATSMA; MCCAMMON, 1999).

A primeira alternativa é a *back door*, foi revelada em 1994 por *Sussam et al.* (GILSON et al., 1994). Neste trabalho, *Sussam* realiza uma simulação de dinâmica molecular em água e encontra um canal temporário, próximo ao aminoácido TrP117, grande o suficiente para passar uma molécula de água. A *back door* justificaria o fato da enzima apresentar uma alta taxa catalítica, mesmo com o local ativo sendo muito estreito. As simulações realizadas revelaram que além da água (solvente), substratos e produtos poderiam mover-se através da *back door* e ainda, os cálculos eletrostáticos expressaram um forte campo eletrostático na *back door*, que promovem sua abertura e atrai substratos.

O primeiro estudo computacional, para mostrar que produtos hidrolisados na AChE podem sair pela *back door*, foi realizado por *Xu et al.* (XU et al., 2010). Esta investigação é baseada num estudo de dinâmica molecular para compreender o tráfego e remoção da tiocolina, através do sítio ativo da AChE do organismo *Torpedo Californica*-(TcAChE). Como conclusão, foi descoberto que produtos de hidrólises saem não somente pelo sítio ativo, mas também pela *back door* e por uma terceira alternativa, a *side door*.

Na Figura 6, está representada a *back door* da AChE humana (PDB:4EY7), formada pelos aminoácidos Val163, Tyr480 e Glu483.

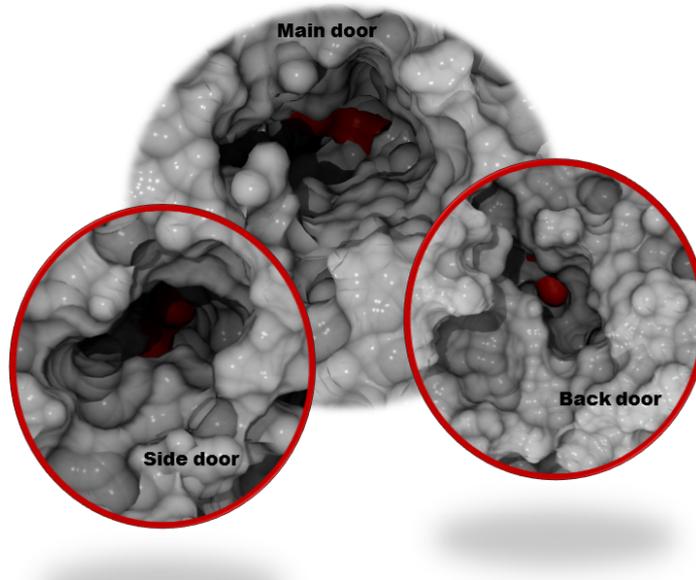


Figura 6 – Portas alternativas existentes na AChE, “*main door*”, “*back door*” e “*side door*”, a superfície em vermelho é a representação para a Ser234, aminoácido pertencente à tríade catalítica, para ilustrar que todas as portas levam ao sítio catalítico da Gorge.

1.2 *Inibidores da Acetilcolinesterase*

Neste trabalho será considerada a hipótese colinérgica, estratégia que norteou a maioria dos fármacos desenvolvidos atualmente. Embora essa abordagem terapêutica seja viável e atraente para amplificar a ação da ACh no cérebro de pacientes com DA, as drogas estudadas como potenciais para o tratamento da DA, limitam-se a reduzir a taxa de progressão da doença, são tratamentos sintomáticos (SUGIMOTO et al., 2000; BUTINI et al., 2008).

O primeiro AChEI a ser comercializado foi a Tacrina (THA) aprovada em 1993 e comercializada como Cognex. No entanto, sua administração apresentou toxicidade hepática e alta incidência de efeitos adversos, além disso ensaios laboratoriais comprovam eventos de falta de benefício para a DA (EAGGER; LEVY; SAHAKIAN, 1991).

O segundo foi o donepezil (cloridrato de donepezila) aprovado em 1996 e comercializado como Aricept. O donepezil é considerado um medicamento de alta seletividade, reversível e o único AChEI *dual-binding-site*, aprovado para o tratamento da DA (SAVINI et al., 2003).

Em 1998, a Rivastigmina (RIVA) foi o terceiro fármaco aprovado, esse agente melhorou significativamente a cognição dos pacientes durante os testes clínicos em comparação com aos agentes já existentes. A RIVA é comercializada como Exelon, é classificada como um agente de ação intermediária ou pseudo-irreversível, podendo inibir a AChE por até 10 horas, e bioquimicamente, apresenta seletividade do sistema nervoso central com inibição periférica (JANN, 2000).

A GALA (Galantamina) aprovada em 2001, comercializada com nome de Reminyl. É considerada como um fármaco reversível e também exerce modulação alostérica dos receptores nicotínicos colinérgicos. Ela se diferencia pela boa tolerância e melhorias na cognição dos pacientes, no entanto, sua administração exige cuidados no que se refere a seu metabolismo hepático, já que pode ocorrer interação deste medicamento com outros fármacos (CARAMELLI et al., 2004).

A estrutura química dos fármacos tacrina, donepezil, rivastigmina e galantamina, é apresentada na Figura 7.

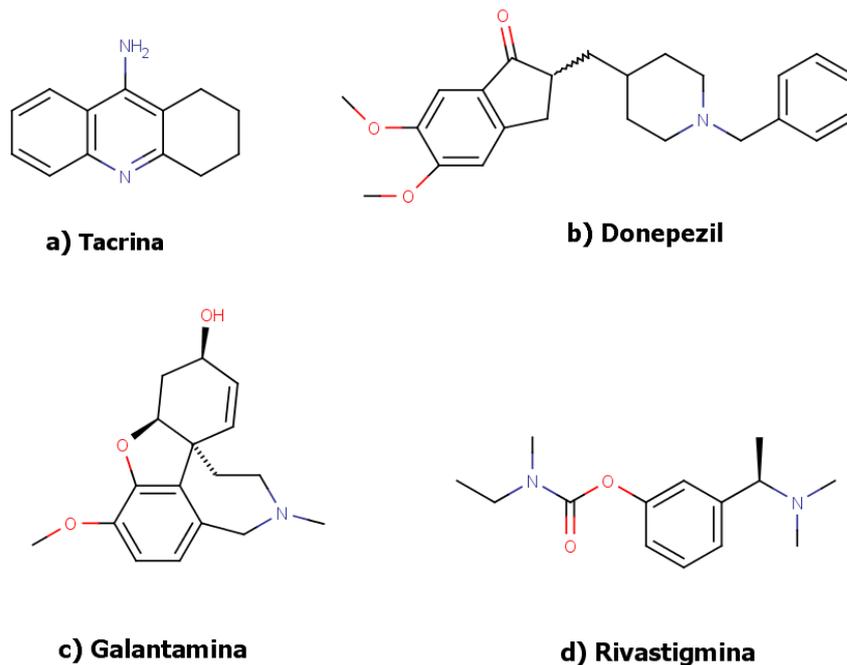


Figura 7 – Fármacos utilizados no tratamento da DA.

Dos fármacos citados acima, o Donepezil é o único que liga-se simultaneamente em dois sítios da Gorge. O fato da Tacrina interagir somente no sítio catalítico, e apresentar algumas falhas na eficiência, tem levado cientistas da química farmacêutica a investir

na busca de fármacos simultaneamente se ligue as PAS e ao sítio catalítico da AChE, chamados inibidores duplos de sítios de ligação que possuem eficácia melhorada. Além disso, esses agentes podem inibir a ação da AChE em relação ao péptido β -amilóide o que pode frear a doença (BOLOGNESI et al., 2007).

O cristal da AChE (PDB:4EY7) está complexada com o Donepezil, e a cristalização revela que o Donepezil estende-se por todo o comprimento da Gorge e possui uma orientação preferencial. Na Figura 8, interações feitas pelo Donepezil (E2020) no sítio da Gorge, mostram que esse fármaco não interage com o sítio catalítico. Os dois grupos aromáticos no final da estrutura do Donepezil (grupo indanona) forma uma interação de empilhamento π com o anel do Trp317 do PAS, o grupo carbonila (aceitador) forma uma interação de ligação de hidrogênio com o NH da Phe326 do sítio acil, sua extensão pode realizar ainda interações hidrofóbicas com resíduos do sítio ativo e no fundo da Gorge, faz uma interação de empilhamento π com anel do Trp117 do sítio aniônico (CHEUNG et al., 2012; FANG et al., 2014).

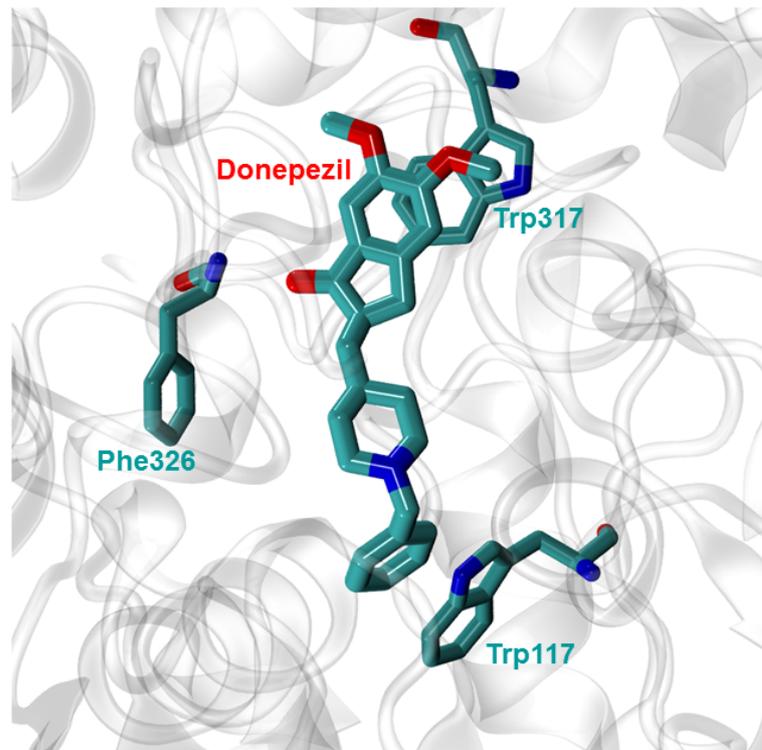


Figura 8 – Interações realizadas pelo Donepezil no sítio da AChE humana (PDB:4EY7).

Estudos anteriores indicam que derivados do cardanol (subproduto de uma mistura fenólica) apresentam boas propriedades anticolinérgicas. Motivados pelo interesse encontrar

propriedades em moléculas sejam potenciais inibidores de AChE, e guie o estudo para a descoberta de novos AChEIs com habilidades *dual binding*, faremos um estudo de uma série de derivados planejados a partir do cardanol, obtidos do líquido da casca da castanha do caju (LCC), como produto natural base para planejamento racional fármacos, a partir de seus lipídeos fenólicos não-isoprenóides do cajueiro (PAULA et al., 2009).

1.3 *Lipídeos Fenólicos presentes no Líquido da Castanha do Caju*

O cajueiro é uma planta nativa do Brasil, da família botânica Anacardiaceae, de nome científico *Anacardium occidentale* é original da região nordeste do Brasil. Segundo o IBGE, a área plantada no país até 2012 é de 779.351 hectares, sendo os principais estados produtores Ceará, Piauí, Rio Grande do Norte e Bahia. Essa planta também é descrita como de utilidade medicinal, com potencial analgésico, diurético, bem como utilizada em problemas respiratórios, gripe, bronquite, tosse, eczema, doenças de pele, verrugas e agente no processo de cicatrização (MAZZETTO; LOMONACO; MELE, 2009).

O fruto do caju é a castanha, da qual se aproveita a amêndoa, a casca e o líquido da casca da castanha (LCC). A amêndoa é uma parte muito apreciada e exportada, especialmente para Europa e América do Norte. Segundo o Ministério de Desenvolvimento, Indústria e Comércio Exterior (MDIC) as exportações de castanha de caju somaram US\$ 185,7 milhões em 2012. O LCC compreende 25% do peso do fruto, sendo uma das fontes mais abundantes de lipídeos fenólicos não isoprenóides (MAZZETTO; LOMONACO; MELE, 2009).

O LCC é um líquido viscoso, vermelho, cáustico, acrimonioso e inflamável encontrado nos alvéolos do mesocarpo da castanha do caju, constituído por quatro lipídeos fenólicos não isoprenóides o ácido anacárdico, cardol, cardanol e metilcardol (ver figura 9) (JULIS et al., 2014).

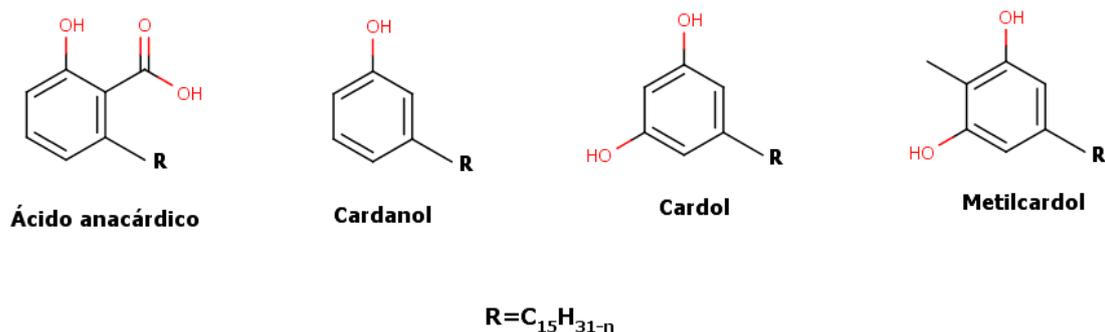


Figura 9 – Estrutura química dos lipídeos não isoprenóides presentes no LCC.

Foram realizadas modificações na estrutura do cardanol nos pontos de acessos mostrados na Figura 10 cujo esqueleto molecular possui características químicas e físico-químicas particulares devido à sua característica estrutural composta por cadeias alquílica e anel fenólico. Informações a respeito das estratégias de síntese e do planejamento estrutural são melhor explanados na referência (LEMES et al., 2016).

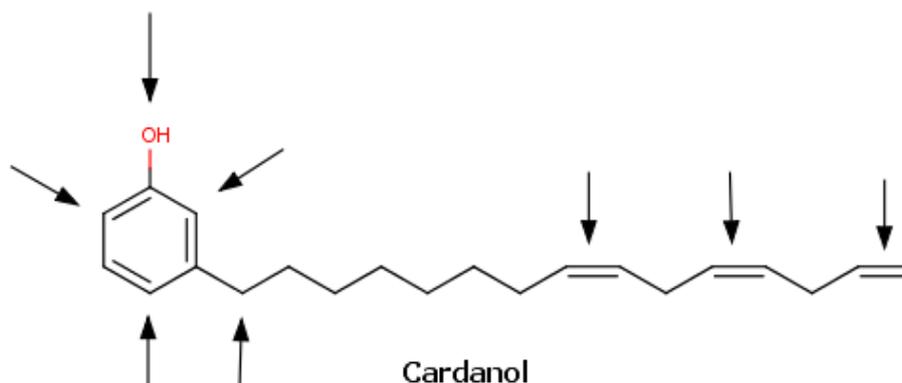


Figura 10 – Estrutura química do cardanol e pontos de alterações para o desenho racional de derivados do cardanol.

Depois de síntese dos candidatos derivados do cardanol, realizou-se ensaios farmacológicos *in vitro* de inibição da enzima acetilcolinesterase purificada do peixe elétrico (do inglês, *Electrophorus electricus*-EeAChE), buscando identificar quais funcionalidades teriam cada molécula a partir das modificações estruturais. Este estudo experimental *in vitro*, mostrou que algumas modificações levou a derivados que inibiam a enzima EeAChE a concentrações micromolares baixas. E em particular, um composto LDT161, o mais

potente da série (valor de $IC_{50} = 6.6\mu M$), foi testado na isoforma humana da AChE e apresentou resulta de $IC_{50} = 5.7\mu M$. Foram realizados também teste de citotoxicidade, e foi constatado a ausência de citotoxicidade significativa na célula humana HT-29, até a concentração mais elevada de $100\mu M$ (LEMES et al., 2016).

A atividade inibitória foi classificada com base nos valores de IC_{50} e % de inibição, à concentração de $100\mu M$. O IC_{50} (do inglês, *half maximal inhibitory concentration*), é o índice de concentração máxima necessária para um fármaco inibir 50% de um processo biológico. Essa medida é muito utilizada para mensurar a potência de fármacos (YUNG-CHI; PRUSOFF, 1973).

Para que o estudo fique mais didático, os candidatos foram divididos em grupos de acordo com as similaridades dos grupos substituintes (ver figuras 11, 12, 13 e 14). Os resultados experimentais a respeito da atividade biológica serão apresentados também por grupo.

Os grupos de compostos candidatos a inibidores da acetilcolinesterase e suas respectivas respostas biológicas a partir do ensaio experimental, são apresentados nas Tabelas 1, 2 e 3 e nas Figuras 11, 12, 13 e 14.

Tabela 1 – Tabela com os valores do IC_{50} e inibição para a classificação da atividade biológica dos aminoderivados heterocíclicos

Molécula	%inibitório($100\mu M$)	$IC_{50}(\mu M)$
LDT140(1)	87.5	26.4
LDT141(2)	41.8	–
LDT142(3)	42.1	–
LDT144(4)	–	–

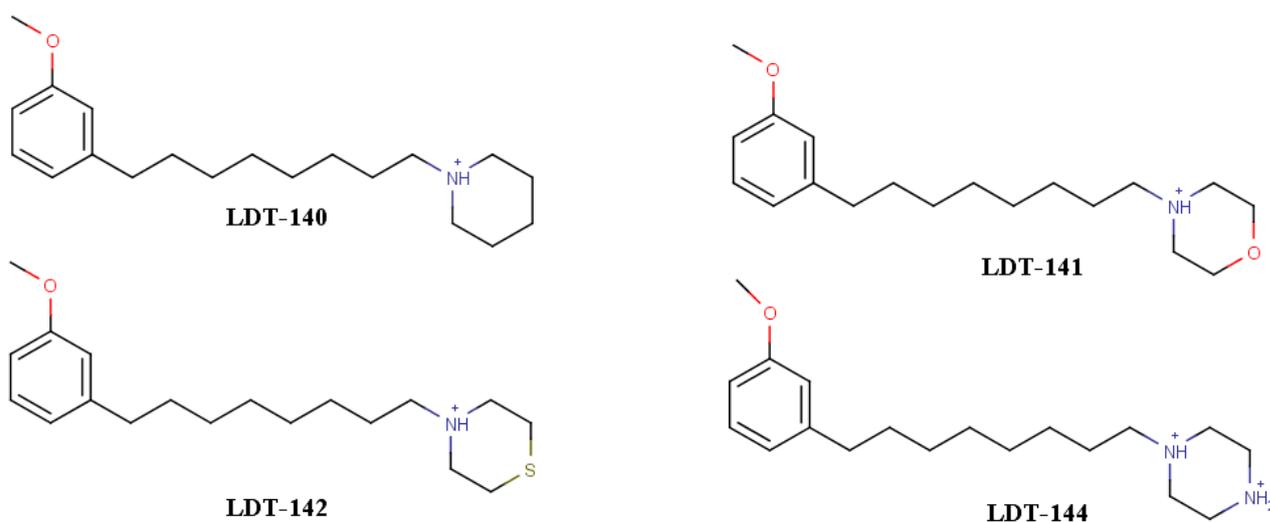


Figura 11 – Estrutura molecular do grupo de derivados aminoderivados heterocíclicos.

Tabela 2 – Tabela com os valores do IC_{50} e inibição para a classificação da atividade biológica dos derivados piperazínicos e arilpiperazínicos.

Molécula	%inibitório(100 μ M)	IC_{50} (μ M)
LDT7(1)	24.7	–
LDT143(2)	63.3	59.8
LDT145(3)	49.8	–
LDT146(4)	–	–
LDT147(5)	28.4	–
LDT166(6)	–	–
LDT168(7)	55.5	–
LDT170(8)	–	–
LDT242(9)	39.4	56.8
LDT473(10)	–	–
LDT474(11)	–	–

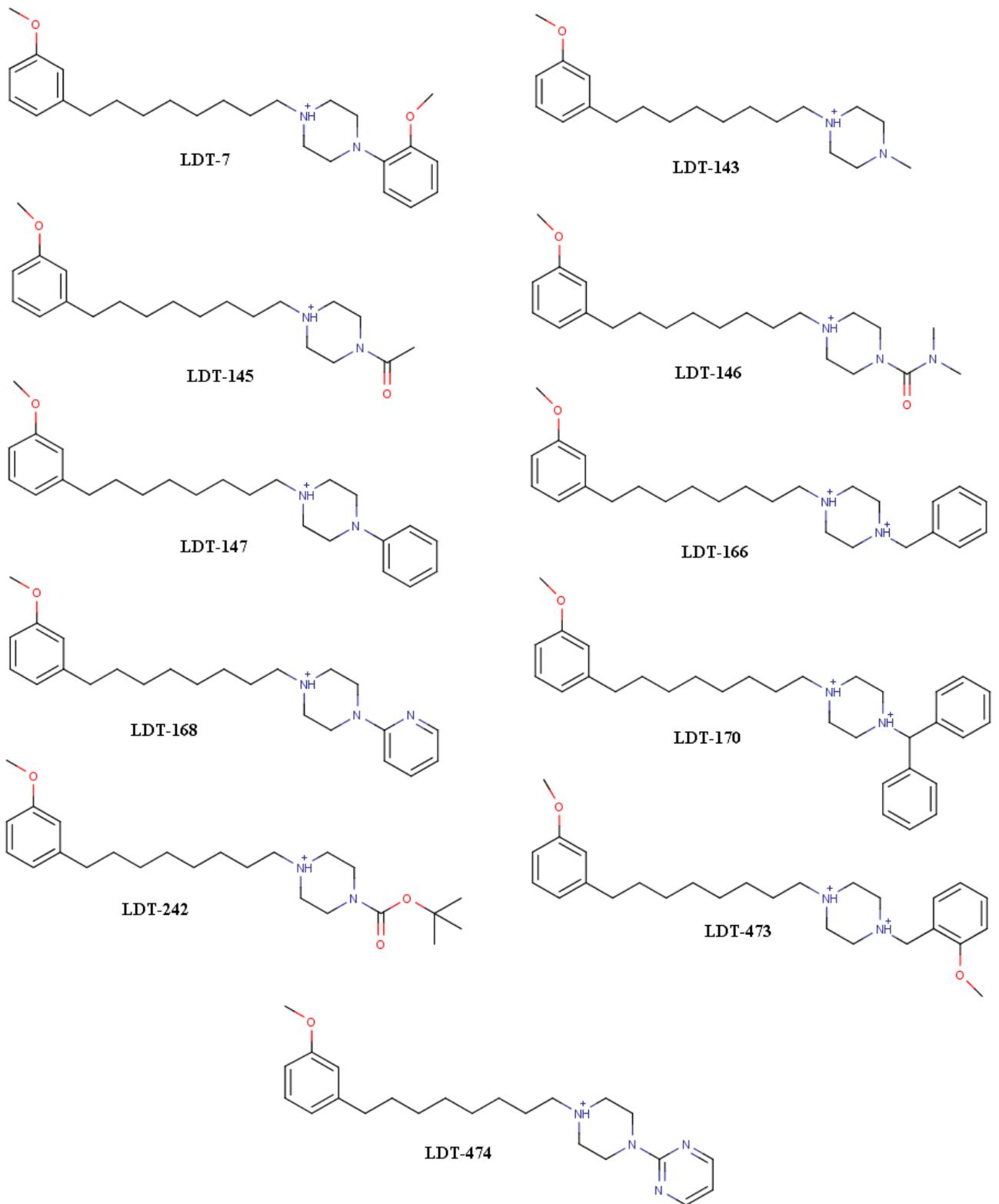


Figura 12 – Estrutura molecular do grupo de derivados piperazínicos e arilpiperazínicos.

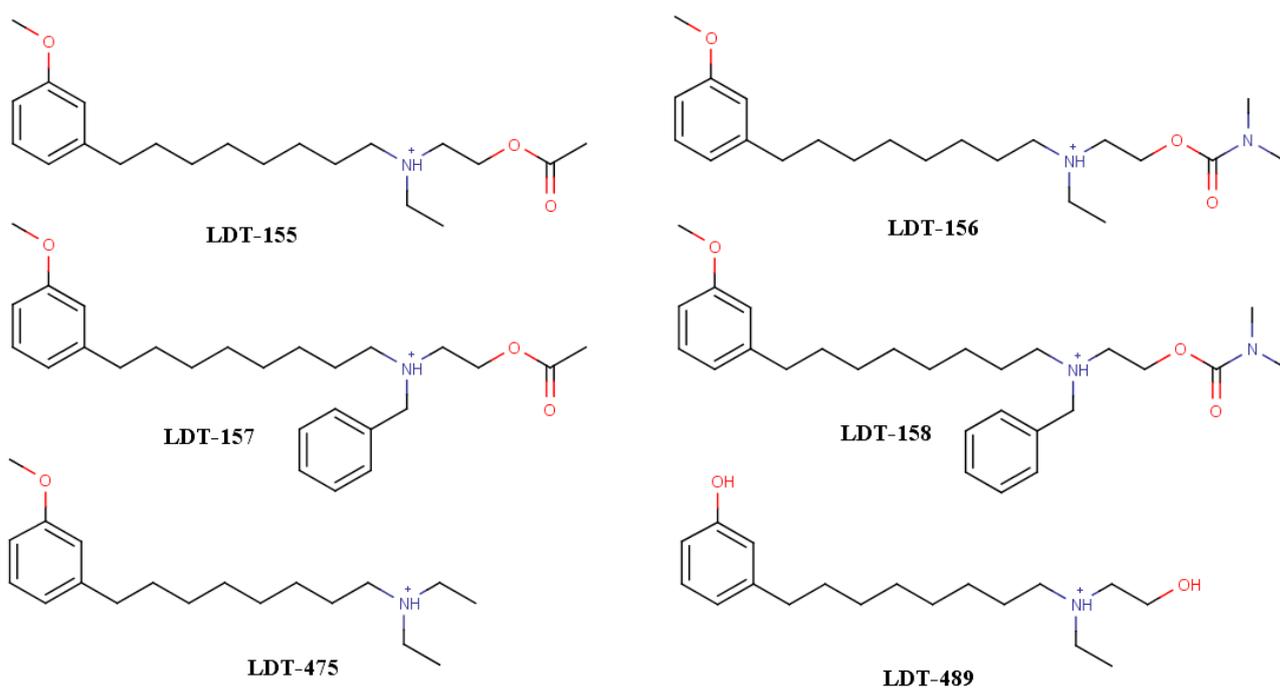


Figura 13 – Estrutura molecular do grupo de derivados etanolamínicos N-Alquil ou N-benzilsubstituídos..

Tabela 3 – Tabela com os valores do IC_{50} e inibição para a classificação da atividade biológica dos derivados benzilamínicos.

Molécula	%inibitório(100 μ M)	IC_{50} (μ M)
LDT159(1)	–	–
LDT160(2)	81.5	16.1
LDT161(3)	91.6	6.6
LDT167(4)	95.5	17.2
LDT169(5)	24.9	–

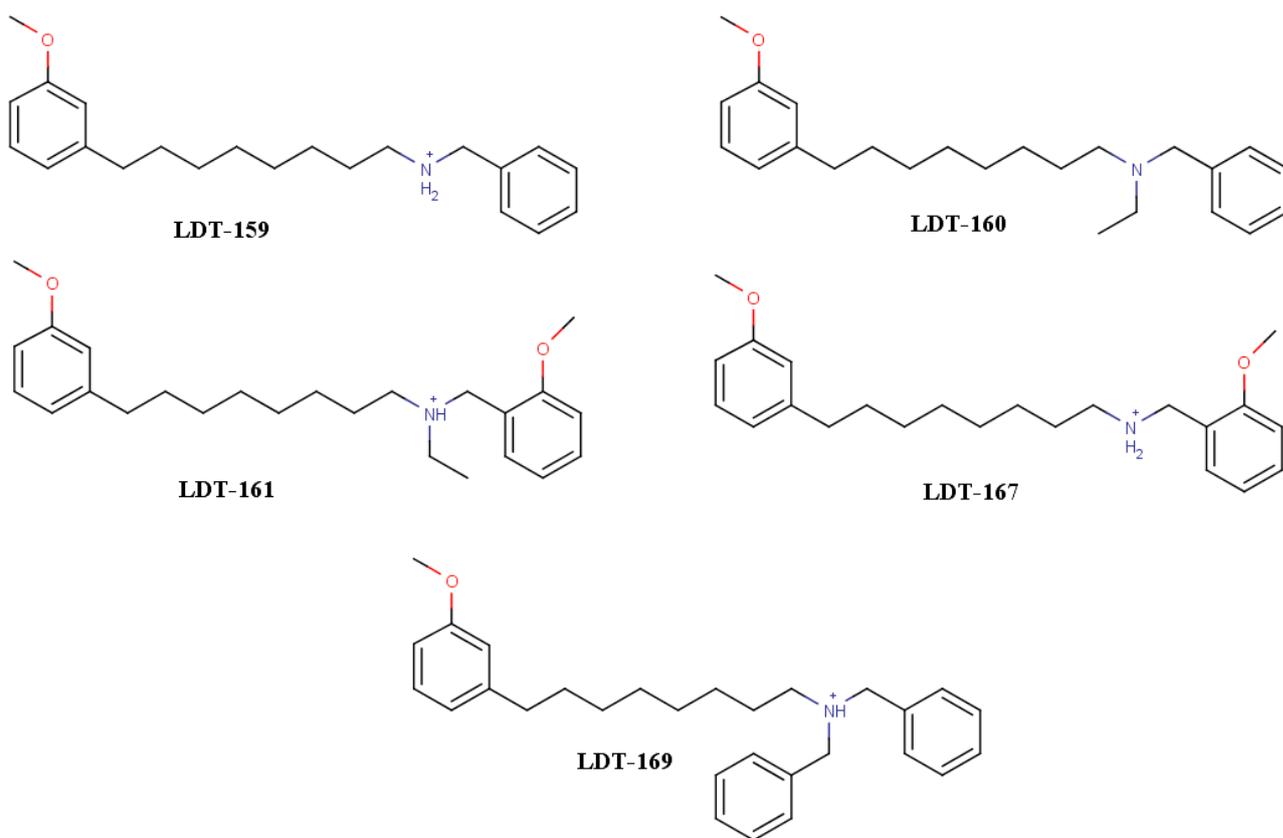


Figura 14 – Estrutura molecular do grupo de derivados benzilamínicos.

É de grande interesse da comunidade científica, principalmente das ciências química e física voltada para área medicinal, a busca por agentes terapêuticos mais eficientes que melhorem o sistema cognitivo de indivíduos com DA e diminua os efeitos adversos, produzidos pelos fármacos atualmente comercializados. Propor novos candidatos à fármacos para a DA, utilizando matéria-prima renovável e abundante encontrada, corroboram para a importância social e econômica para o país.

Pensando nisso, esta pesquisa propõe o estudo teórico, baseado no formalismo da mecânica quântica, simulações de dinâmica molecular clássica e *docking* receptor-ligante, buscando o modo de reconhecimento molecular dos ligantes no sítio da enzima AChE (PDB:4EY7), de candidatos à novos agentes terapêuticos para a DA, planejados e desenvolvidos para a DA no Laboratório de Desenvolvimento de Estratégias Terapêuticas (LADETER/UCB), pelo grupo de pesquisadores da UCB e pelo grupo da Faculdade de Ciências da Saúde da UnB. O capítulo seguinte, explora as metodologias aplicadas neste estudo.

2 METODOLOGIAS

O planejamento racional de fármacos, bem como o estudo teórico de candidatos à fármacos, trabalham com estratégias de investigação de estruturas moleculares. Entre as técnicas empregadas, a modelagem molecular é a mais utilizada. O estudo computacional de candidatos à fármacos é uma ferramenta muito promissora no planejamento de novos agentes terapêuticos, além da confiabilidade que essa ferramenta possui, diminui de maneira considerável os gastos financeiros empregados nas investigações experimentais (LEACH; KIER, 1997).

A análise conformacional é o estudo do modelo molecular tridimensional e das variações de energias associadas a rotações de diedros significativos nas ligações das moléculas. As rotações em torno dessas ligações formam arranjos moleculares ou conformações, que são formas temporárias e apresentam estabilidades diferentes. As preferências conformacionais assumidas por candidatos à fármacos é de grande interesse, uma vez que conhecendo as conformações energeticamente mais apropriadas podemos obter informações sobre a estereoquímica das interações com o sítio ativo. Durante a construção dessas estruturas e na geração desses conformeros podem haver deformações nos comprimentos de ligação e ângulos associados, por isso, as conformações mais estáveis são obtidas por meio da otimização de geometria via cálculos de mecânica quântica e molecular (BARREIRO et al., 1997).

O objetivo do presente capítulo consiste em apresentar ideias e conceitos essenciais à compreensão das ferramentas utilizadas na realização desse trabalho. Métodos e teorias bem estabelecidas como cálculos de estrutura eletrônica, simulações de dinâmica molecular, *Docking* molecular e seus detalhes técnicos, podem ser consideradas com maiores detalhes em diversas referências ((SZABO; OSTLUND, 1989) ,(LEACH, 2001), (LEVINE, 2009)). Além disso, as referências pertinentes também são indicadas no texto.

2.1 A Equação de Schrödinger Eletrônica

As moléculas candidatas à fármacos estudadas neste trabalho possuem em média 80 átomos, configurando assim um problema eletrônico de muitos corpos. Dentro da mecânica

quântica um sistema molecular composto por N elétrons e M núcleos pode ser descrito pela equação de Schrödinger independente do tempo e desprezando quaisquer efeitos relativísticos:

$$\mathcal{H}|\Phi\rangle = \mathcal{E}|\Phi\rangle \quad (1)$$

onde \mathbf{r} destina-se às coordenadas eletrônicas e \mathbf{R} às nucleares; \hat{H} é o operador Hamiltoniano molecular, dado por (em unidades atômicas)

$$\mathcal{H} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{A=1}^M \sum_{i=1}^N \frac{Z_A}{r_{iA}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (2)$$

onde i e j são índices para os elétrons e r_{ij} a distância entre eles, M_A é a massa do núcleo, r_{iA} a distância elétron-núcleo e R_{AB} a distância entre os núcleos. O primeiro termo de (2) é o operador energia cinética dos elétrons. Enquanto o segundo é o operador energia cinética dos núcleos. O terceiro termo é a energia potencial de atração entre os elétrons e o núcleo, com r_{iA} a distância entre os elétrons i e o núcleo A . O quarto termo é a energia potencial de repulsão entre os núcleos. O último termo é a energia potencial de repulsão entre os elétrons (LEVINE, 2009).

Resolver a equação de Schrödinger para um sistema molecular, mesmo para os mais simples, não é fácil, tornando imprescindível o uso de aproximações, dentre algumas das mais utilizadas pela comunidade científica encontra-se a Aproximação de Born-Oppenheimer (ABO).

2.1.0.1 A aproximação de Born-Oppenheimer

A aproximação de Born-Oppenheimer é fundamental para a química quântica. Considerando o fato dos núcleos serem mais pesados que os elétrons, eles movem-se mais lentamente. Sendo assim, uma boa aproximação é considerar que os elétrons em uma molécula movem-se em um campo de núcleos fixos (SZABO; OSTLUND, 1989).

Essa é uma consideração proveniente do teorema adiabático, que sustenta que se a perturbação imposta ao sistema for lenta o bastante para que o sistema se adapte à sua nova condição, o autoestado é conservado. Dessa maneira é possível desacoplar os movimentos dos elétrons ao dos núcleos (LEACH, 2001).

Dentro dessa aproximação, o segundo termo da equação 2, a energia cinética dos núcleos, pode ser negligenciado e o penúltimo termo, a repulsão entre os núcleos, pode ser considerada constante. Os termos restantes em 2 são conhecidos como a Hamiltoniana Eletrônica que descreve o movimento de N elétrons em um campo de M cargas pontuais,

$$\mathcal{H} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{A=1}^M \sum_{i=1}^N \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}}. \quad (3)$$

A solução para a equação de Schrödinger abrange a Hamiltoniana eletrônica,

$$\mathcal{H}_{el} \Phi_{el} = \mathcal{E}_{el} \Phi_{el} \quad (4)$$

e a função de onda eletrônica,

$$\Phi_{el} = \Phi_{el}(\{\mathbf{r}_i\}; \{\mathbf{R}_A\}) \quad (5)$$

que retrata o movimento dos elétrons e depende explicitamente das coordenadas eletrônicas e parametricamente das coordenadas nucleares, como na energia eletrônica,

$$\mathcal{E}_{el} = \mathcal{E}_{el}(\{\mathbf{R}_A\}). \quad (6)$$

A dependência paramétrica significa que, para diferentes arranjos nucleares, Φ_{el} é uma função de diferentes coordenadas eletrônicas. De maneira que, coordenadas nucleares não aparecem explicitamente em Φ_{el} . E a energia total para núcleos fixos compreende também uma constante de repulsão nuclear.

$$\mathcal{E}_{tot} = \mathcal{E}_{el} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}, \quad (7)$$

as equações 3 e 7 consistem o problema eletrônico.

Resolvido o movimento para os elétrons, é possível solucionar o movimento dos núcleos utilizando a mesma aproximação. Como os elétrons movem-se mais rapidamente que os núcleos, quando o núcleo muda lentamente a sua configuração, os elétrons imediatamente se ajustam a essa mudança. Essa é uma aproximação coerente para substituir as coordenadas eletrônicas por seus valores médios, calculada sobre a função de onda eletrônica. Isso produz uma Hamiltoniana nuclear para o movimento dos núcleos em um campo médio de elétrons,

$$\begin{aligned}
\mathcal{H}_{nuc} &= -\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + \left\langle -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{A=1}^M \sum_{i=1}^N \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \right\rangle \quad (8) \\
&+ \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \\
&= -\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + \mathcal{E}_{el}(\{\mathbf{R}_A\}) + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \\
&= -\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + \mathcal{E}_{tot}(\{\mathbf{R}_A\})
\end{aligned}$$

A energia total $\mathcal{E}_{tot}(\{\mathbf{R}_A\})$ gera um potencial para o movimento nuclear. Então na aproximação de Born-Oppenheimer move uma energia potencial de superfície obtida pela solução do problema eletrônico (SZABO; OSTLUND, 1989). E a solução para a equação de Schrödinger nuclear,

$$\mathcal{H}_{nuc} \Phi_{nuc} = \mathcal{E}_{nuc} \Phi_{nuc} \quad (9)$$

explica a vibração, rotação e translação da molécula,

$$\Phi_{nuc} = \Phi_{nuc}(\{\mathbf{R}_A\}) \quad (10)$$

logo, \mathcal{E} (energia total para a aproximação de Born-Oppenheimer 1) inclui a energia eletrônica, vibracional, rotacional e translacional (BUNKER; MOSS, 1977).

2.1.1 Solução da Equação de Schrödinger Eletrônica

A teoria dos orbitais moleculares foi desenvolvida depois do advento da mecânica quântica, ela é uma extensão das ideias de Bohr. Segundo ela, existe uma função de estado para cada elétron e a composição destas forma uma função de estado total de N elétrons. A formulação matemática para o problema de muitos corpos, a partir teoria dos orbitais, foi exposta por Hartree, Slater, Roothaan (que na época era um aluno de Mulliken), Löwdin e outros colaboradores, levou ao prêmio Nobel em 1966, concedido à Robert S. Mulliken (VIANNA, 2004).

Temos inicialmente o orbital molecular (O.M.) $\varphi(r^\mu) = \varphi_i(x^\mu, y^\mu, z^\mu)$ onde x^μ, y^μ, z^μ são as coordenadas do elétron μ e i diz respeito ao estado eletrônico. Se o spin for considerado, tem-se o spin-orbital molecular (S.O.M.):

$$\psi_i(q^\mu) = \varphi_i(x^\mu, y^\mu, z^\mu)\xi(S^\mu) = \varphi_i(\mu)\xi(\mu) = \varphi_i^\mu\xi^\mu, \quad (11)$$

onde

$$\xi(S^\mu) = \begin{cases} \alpha(\mu) & \uparrow \\ \beta(\mu) & \downarrow. \end{cases} \quad (12)$$

A função de estado total para o sistema de N-elétrons que satisfaz o princípio da exclusão de Pauli, é um produto anti-simétrico dos spin-orbitais:

$$\Phi = (N!)^{\frac{1}{2}} \hat{A} \left\{ \psi_1^{(1)} \psi_2^{(2)} \dots \psi_N^{(N)} \right\}, \quad (13)$$

em que Φ é a função de estado total para N elétrons e \hat{A} é o operador antisimetrizador (operador de projeção). Nessa expressão $\Phi_{a,b,\dots,k}(q^1, q^2, \dots, q^N)$ por Φ e cada estado $\psi_j(q^\mu)$ por ψ_j^μ . A relação 13 pode ser escrita na forma de determinante, o determinante de Slater:

$$\Phi = (N!)^{-\frac{1}{2}} \begin{bmatrix} \psi_1^{(1)} & \psi_2^{(1)} & \dots & \psi_N^{(1)} \\ \psi_1^{(2)} & \psi_2^{(2)} & \dots & \psi_N^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1^{(N)} & \psi_2^{(N)} & \dots & \psi_N^{(N)} \end{bmatrix}. \quad (14)$$

em que

ψ^μ 's são funções das coordenadas espaciais e spin de um único elétron, os S.O.M.

e tem como característica,

(a) Os S.O.M são linearmente independentes,

(b) O princípio da exclusão de Pauli é satisfeito automaticamente.

Admitindo que os ψ^μ 's sejam ortonormais, o fator $(N!)^{-\frac{1}{2}}$ é uma constante de normalização para Φ_0 . O determinante único faz uma simplificação no sentido de substituir o problema de encontrar uma função de onda que depende das coordenadas de N elétrons, pelo problema de encontrar N funções de onda de um elétron (S.O.M.). Trocar coordenadas de dois elétrons significa trocar duas linhas do determinante, o que implica numa inversão de sinais. A função de onda Hartree-Fock (HF), definida em termos de um único determinante de Slater, é antissimétrica diante da troca das coordenadas de dois elétrons. Como um

determinante com dois S.O.M. iguais possui duas colunas iguais, e portanto, é nulo. O que satisfaz o princípio da exclusão de Pauli, que afirma que dois elétrons não pode ocupar o mesmo estado quântico (mesmo S.O.) (COUTINHO, 2007).

Os S.O.M. ψ_k são linearmente independentes, podemos escrevê-los como:

$$\int \psi_k^*(\mu)\psi_\ell(\mu)d\tau_\mu = \delta_{k\ell}. \quad (15)$$

Conseqüentemente, teremos

$$\int \Phi^*\Phi d\tau = 1. \quad (16)$$

Então a equação 16 torna-se

$$\int \Phi^*\Phi d\tau = \int \psi_1^{(1)\dagger}\psi_1^{(1)}d\tau_1 \int \dots \int \psi_N^{(N)\dagger}\psi_N^{(N)}d\tau_N = 1.$$

A energia do sistema é dada por:

$$E = \int \Phi^*\hat{H}\Phi d\tau \equiv E[\Phi], \quad (17)$$

onde o operador Hamiltoniano em unidades atômicas é definido por:

$$\hat{H} = \sum_{\mu} \hat{h}(\mu) + \frac{1}{2} \sum_{\mu \neq \nu} \frac{1}{r^{\mu\nu}}. \quad (18)$$

Em 18, $\hat{h}(\mu)$ é o operador hamiltoniano de uma partícula para o μ -ésimo elétron movendo-se no campo dos núcleos; $\hat{h}(\mu)$ é um operador hermitiano e linear expresso por

$$\hat{h}(\mu) = -\frac{\nabla_{\mu}^2}{2} - \sum_A \frac{1}{r^{\mu A}} \quad (19)$$

com $\frac{\nabla_{\mu}^2}{2}$ sendo a energia cinética do elétron μ , e $\sum_A \frac{1}{r^{\mu A}}$ a energia coulombiana de interação elétron-núcleo.

Os determinantes de Slater são a essência do método Hartree-Fock (COUTINHO, 2007). Tomando o único determinante de Slater 14, pode-se escrever as dependências com relação a parte espacial e de spin dos S.O.M separadamente,

$$\psi_a^1 = \phi_{r_1}\alpha(1)$$

ou

$$\psi_a^1 = \phi_{r_1} \beta(1) \quad (20)$$

onde α e β representam, respectivamente, spin “para cima” e “para baixo”. Se nenhuma restrição for feita aos orbitais moleculares, o método é denominado Hartree-Fock não Restrito(UHF), sendo vantajoso para sistemas de camada aberta. Sistemas de camada aberta podem ser descrito também por uma função de onda do tipo restrita, onde as partes espaciais dos spins-orbitais, duplamente ocupados, sejam obrigatoriamente as mesmas. Nestas condições, tem-se o método Hartree-Fock Restrito de Camada Aberta (ROHF). Para sistemas que o estado eletrônico for camada fechada, é necessário que faça-se uma restrição, na qual cada orbital espacial seja associado a dois elétrons, um com spin α e outro com spin β , para esse sistema o método é chamado Hartree-Fock Restrito (COUTINHO, 2007; SZABO; OSTLUND, 1989).

2.1.2 Equação de Hartree-Fock

Como a função de onda que mais se aproxima da solução exata é a que conduz a um mínimo da energia,

$$\mathcal{E}(\psi) = \sum_a \langle \psi_a | h | \psi_a \rangle + \frac{1}{2} \sum_{a,b} (\langle \psi_a \psi_b | \psi_a \psi_b \rangle - \langle \psi_a \psi_b | \psi_b \psi_a \rangle) \quad (21)$$

pode ser minimizada sujeita à restrição que os S.O.M. sejam ortonormais, ou seja,

$$\langle \psi_a | \psi_b \rangle - \delta_{ab} = 0 \quad (22)$$

Usando a técnica de multiplicadores de Lagrange para a condição de vínculo (SLATER, 1951), temos

$$\mathcal{L}[\psi] = \mathcal{E}[\psi] - \sum_{a,b} \epsilon_{ba} (\langle \psi_a | \psi_b \rangle - \delta_{ab}) \quad (23)$$

onde os ϵ_{ab} são os multiplicadores de Lagrange. Dada as condições:

- \mathcal{L} é real,

- \mathcal{E} é real,

- a e b assumem os mesmos valores, podem ser trocados no duplo somatório.

Se cada S.O.M. sofrer uma variação $\delta\psi$, então L também sofrerá. Podemos escrevê-lo como:

$$\begin{aligned} \delta\mathcal{L} = & \sum_a \langle \delta\psi_a | h | \psi_a \rangle + \sum_{a,b} \langle \delta\psi_a \psi_b | \psi_a \psi_b - \delta\psi_a \psi_b | \psi_b \psi_a \rangle \\ & - \sum_{a,b} \epsilon_{ba} \langle \delta\psi_a | \psi_b \rangle + \text{complexo conjugado.} \end{aligned} \quad (24)$$

Definindo os operadores de Coulomb, \mathcal{J}_b , e de Troca, \mathcal{K}_b através das expressões

$$\mathcal{J}_b(1)\psi_a(1) = \left\langle \psi_b(2) \left| \frac{1}{r_{12}} \right| \psi_b(2) \right\rangle \psi_a(1)$$

e

$$\mathcal{K}_b(1)\psi_a(1) = \left\langle \psi_b(2) \left| \frac{1}{r_{12}} \right| \psi_b(2) \right\rangle \psi_a(1) \quad (25)$$

Para que \mathcal{L} seja um mínimo devemos ter $\delta\mathcal{L}=0$. Como $\delta\psi_a$ é arbitrário, tem-se

$$\{h(1) + \sum_b [\mathcal{J}_b(1) - \mathcal{K}_b(1)]\} \psi_a(1) = \sum_b \epsilon_{ba} \psi_b(1), \quad (26)$$

a expressão entre chaves

$$\mathcal{F}(1) = h(1) + \sum_b [\mathcal{J}_b(1) - \mathcal{K}_b(1)] \quad (27)$$

é chamada de operador de Fock. Este operador depende das soluções da equação S.O.M. Existe uma equação para cada orbital atômico que depende dos outros orbitais através do operador de Fock. Portanto, as equações devem ser resolvidas de maneira acoplada, por meio de aproximações sucessivas. Devido a isso o método HF é um método autoconsistente, no final do processo os ψ_s que são soluções da equação de Hartree-Fock, devem ser os mesmos ψ_s a partir dos quais o operador de Fock foi obtido.

Em termos deste operador, temos

$$\mathcal{F}(1)\psi_a(1) = \sum_b \epsilon_{ba} \psi_b(1) \quad (28)$$

usando as definições dos operadores de Coulomb e de troca, temos

$$\epsilon_a = \langle a | h | a \rangle + \sum_b \langle ab || ab \rangle \quad (29)$$

E a energia é escrita como

$$\mathcal{E} = \sum_a \epsilon_a - \frac{1}{2} \sum_{a,b} \langle ab || ab \rangle \quad (30)$$

ou seja, a energia eletrônica total não é igual a soma das energias orbitais. As energias dos orbitais são autovalores do operador de Fock que inclui a energia de repulsão para um determinado elétron com os outros. A energia de repulsão elétron-elétron deve ser descontada, já que essa energia de repulsão elétron-elétron total é contada duas vezes na soma das energias dos orbitais (LEVINE, 2009; COUTINHO, 2007; SZABO; OSTLUND, 1989).

2.1.3 A equação de Hartree-Fock-Roothan

O método de Hartree-Fock é bem resolvido numericamente para sistemas de átomos, tornando-se inviável para sistemas grandes (moléculas poliatômicas). Uma solução para esse problema foi apresentado por Slater, mas formalmente apresentada por Roothaan (1951), com a expansão da parte espacial dos spins-orbitais moleculares em termos de um conjunto de funções conhecidas, chamadas de funções base. Transformando a solução das equações integro-diferenciais acopladas em um problema de álgebra matricial. A formulação matricial do método de Hartree-Fock é a equação de Hartree-Fock-Roothaan (HFR) (LEVINE, 2009; COUTINHO, 2007; SZABO; OSTLUND, 1989).

A ideia é escrever os orbitais moleculares como:

$$\phi_p(r) = \sum_{\nu=1}^k C_{\nu p} g_{\nu}(r) \quad (31)$$

onde

- $g_{\nu}(r)$ são as funções base conhecidas,
- k é o número de funções do conjunto,
- $C_{\nu p}$ são os coeficientes a serem determinados.

As funções tipo gaussianas são as mais utilizadas nos cálculos moleculares, por causa da facilidade de se calcular integrais usando este tipo de função. Uma função gaussiana escrita em coordenadas cartesianas é determinada como

$$g(x, y, z) = N x^l y^m z^n e^{(-\alpha r^2)} \quad (32)$$

onde l, m e n são números inteiros, $r = \sqrt{x^2 + y^2 + z^2}$ e N é uma constante de normalização.

Por convenção, funções gaussianas cartesianas podem ser classificadas como

$$l + m + n = 0 \rightarrow \text{função tipo } s$$

$$l + m + n = 1 \rightarrow \text{função tipo } p$$

$$l + m + n = 2 \rightarrow \text{função tipo } d$$

$$l + m + n = 3 \rightarrow \text{função tipo } f$$

Substituindo a expressão 31 na equação de Hartree-Fock, temos

$$f(\mathbf{r}_1) \sum_{\nu} C_{\nu p} g_{\nu}(\mathbf{r}_1) = \epsilon_p \sum_{\nu} C_{\nu p} g_{\nu}(\mathbf{r}_1) \quad (33)$$

Multiplicando por $g_{\mu}^*(\mathbf{r}_1)$ e integrando, tem-se

$$\sum_{\nu} C_{\nu p} \langle g_{\nu}(1) | f(1) | g_{\nu} \rangle = \epsilon_p \sum_{\nu} C_{\nu p} \langle g_{\mu}(1) | g_{\nu} \rangle \quad (34)$$

onde $g_{\nu}(1) \equiv g_{\nu}(\mathbf{r}_1)$.

Definindo os elementos da matriz de superposição como

$$S_{\mu\nu} = \langle g_{\mu}(1) | g_{\nu} \rangle \quad (35)$$

e da matriz de Fock como

$$F_{\mu\nu} = \langle g_{\mu}(1) | f(1) | g_{\nu} \rangle \quad (36)$$

podemos escrever

$$\sum_{\nu} F_{\mu\nu} C_{\nu p} = \epsilon_p \sum_{\nu} S_{\mu\nu} C_{\nu p} \quad (37)$$

a forma matricial de 37 é

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (38)$$

onde ϵ é a matriz diagonal que contém as energias orbitais $\epsilon_{pq} = \epsilon_p \delta_{pq}$. Essa equação matricial é conhecida como equação de Hartree-Fock-Roothaan (COHEN; ROOTHAAN, 1965).

2.2 Teoria do Funcional da Densidade

A teoria do funcional da densidade (DFT, do inglês "Density Functional Theory") é uma forma de solucionar o problema para sistemas de muitos elétrons, utilizando a densidade eletrônica total $\rho(r)$. A equação de Schrödinger eletrônica de N elétrons com uma função de onda com 3N variáveis pode ser escrita como uma equação da densidade eletrônica com três variáveis.

Historicamente a utilização da densidade eletrônica $\rho(r)$ como variável fundamental para a descrição de um sistema eletrônico data de 1927 quando Llewellyn H. Thomas e Enrico Fermi propuseram a Aproximação de Thomas-Fermi, uma alternativa para a resolução da equação de Schrödinger, usando a densidade eletrônica $\rho(r)$ ao invés função de onda.

Segundo Coutinho *et al.* (COUTINHO, 2007) a densidade eletrônica foi usada por Drude, desde o início do século XX, para desenvolver a teoria sob condução térmica e elétrica, quando aplicou a teoria dos gases a um metal, considerando como um gás homogêneo de elétrons. E a partir de então, outros modelos como os de Sommerfeld, Thomas, Fermi e Dirac foram propostos e aperfeiçoados.

A solução exata, com o uso da densidade eletrônica, foi reconhecida em 1964 através da publicação de dois teoremas por Hohenberg e Kohn, e ficou conhecida como Teoria do Funcional da Densidade-DFT. Este trabalho resultou num prêmio Nobel de Química dado a Walter Kohn em 1998. A aplicação da Teoria do Funcional da Densidade em átomos, moléculas e sólidos aumentou consideravelmente, e vem se mostrando o método mais eficiente para o cálculo de propriedades estruturais e eletrônicas, para sistema no estado fundamental. A partir do formalismo de Hohenberg e Kohn elementos importantes para a descrição química de compostos são descritos bem melhor pela Teoria do Funcional da Densidade (COUTINHO, 2007; BURKE, 2012; VESSECCHI *et al.*, 2008).

A densidade eletrônica ($\rho(\mathbf{r})$) como variável básica na descrição do sistema eletrônico, foi usado primeiramente por Thomas e Fermi, conhecido como modelo de Thomas-Fermi-Dirac (TFD). Em 1964 foi legitimada por Hohenberg e Kohn. Em 1965 Kohn e Sham apresentaram melhores resultados para a energia cinética, o método KS (VIANNA, 2004; PERDEW *et al.*, 1982).

O primeiro teorema de Honhenber-Kohn (HK) estabelece que o potencial externo é um funcional único de $\rho(r)$ além de uma constante aditiva. A energia do sistema é calculada mediante resolução da equação de Schrodinger eletrônica, $\hat{H}_{BO}\Psi = E\Psi$, e a energia é determinada pela densidade eletrônica $\rho(\mathbf{r})$.

$$E = E_v[\rho] \quad (39)$$

O segundo teorema estabelece que, havendo qualquer aproximação da densidade eletrônica, $\tilde{\rho}(r)$, de modo que $\tilde{\rho}(r) \geq 0$ e $\int \tilde{\rho}(r)dr = N$, a energia total será sempre maior ou igual a energia exata do sistema, ou seja, $E\tilde{\rho}(r) \geq E[\rho] = E_0$.

Kohn e Sham propuseram um modelo de partículas não interagentes, no qual o funcional energia é dada por:

$$E_{KS}[\rho] = T_0[\rho] + U_H[\rho] + V_{ext}[\rho] + V_{XC}[\rho] \quad (40)$$

onde $T_0[\rho]$ é a energia cinética média de um sistema não interagente de elétrons cuja densidade é igual a de um sistema interagente, $U_H[\rho]$ é a energia coulombiana média entre os elétrons, $V_{ext}[\rho]$ é o potencial efetivo externo e $V_{XC}[\rho]$ é o termo de troca e correlação, contendo correções para energia cinética e energia de Hartree (KOHN; SHAM, 1965).

O método KS estabelece a repulsão elétron-elétron de Coulomb e uma nova função universal $G[\rho]$:

$$E_v[\rho] = G[\rho] + \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2 + \int \rho(r)v(r)dr, \quad (41)$$

em que

$$G[\rho] = T_s[\rho] + E_{xc}[\rho]$$

e $T_s\rho$ é o funcional de energia cinética de um sistema de elétrons que não interagem e possuem a mesma densidade eletrônica do sistema de elétrons que interagem. De forma que, $E_{xc}[\rho]$ inclui não somente o termo de interação elétron-elétron (troca e correlação), mas também a parte residual da energia cinética, $T[\rho] - T_s[\rho]$, em que $T[\rho]$ é a energia cinética exata para o sistema de elétrons que interagem.

Assim, é possível empregar em um sistema de elétrons que não interagem um Hamiltoniano que tenha potencial local efetivo, $v_{ef}(r)$.

$$H^{KS} = -\frac{1}{2}\nabla^2 + v_{ef}(r). \quad (42)$$

A função de onda Ψ^{KS} , do estado fundamental para elétrons não interagentes é aproximada por um produto anti-simétrico de N funções de onda de um elétron, $\psi_i(r_i)$, representado pelo determinante de Slater:

$$\Psi^{KS} = \frac{1}{\sqrt{N!}} \begin{bmatrix} \psi_1^{KS}(r_1) & \psi_2^{KS}(r_1) & \dots & \psi_N^{KS}(r_1) \\ \psi_1^{KS}(r_2) & \psi_2^{KS}(r_2) & \dots & \psi_N^{KS}(r_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1^{KS}(r_N) & \psi_2^{KS}(r_N) & \dots & \psi_N^{KS}(r_N) \end{bmatrix}. \quad (43)$$

Deste modo, os orbitais Kohn-Sham (KS), ψ_i^{KS} , são obtidos a partir da equação de Schrödinger de um elétron:

$$\left(-\frac{1}{2}\nabla^2 + v_{ef} \right) \psi_i^{KS} = \varepsilon_i \psi_i^{KS} \quad (44)$$

onde ε_i são os autovalores e ψ_i são as autofunções.

A ligação entre esse sistema e o sistema real pode ser feita pela escolha do potencial efetivo, de maneira que a densidade eletrônica resultante seja equivalente à densidade no estado fundamental,

$$\rho_s(r) = \sum_i^N 2|\psi_i^{KS}(r)|^2 = \rho_0(r). \quad (45)$$

A energia cinética é então, calculada com precisão e de maneira autoconsistente, de acordo com a equação 46

$$T_s[\rho] = \sum_i^N \left\langle \psi_i^{KS} \left| -\frac{1}{2}\nabla_i^2 \right| \psi_i^{KS} \right\rangle \quad (46)$$

O potencial efetivo é obtido através da minimização da equação da energia 41, desde que as funções de um elétron sejam ortonormais ($\langle \psi_i^{KS} | \psi_j^{KS} \rangle = \delta_{ij}$), que pode ser vista com mais detalhe na referencia (PARR, 1980). E a equação para o potencial efetivo é,

$$v_{ef}(r) = v(r) + \int \frac{\rho(r_1)}{|r - r_1|} dr_1 + v_{xc}(r), \quad (47)$$

em que

$$v_{xc}(r) = \frac{\delta E_{XC}[\rho]}{\delta \rho(r)}. \quad (48)$$

nesta equação o termo E_{XC} é conhecido como potencial de troca e correlação, este inclui todos os termos de interação entre as partículas. O grande impasse para a DFT é a dificuldade em encontrar boas aproximações para este termo, já que funcionais exatos para troca e correlação não são conhecidos, com restrição do gás de elétrons livres. Contudo existem aproximações que definem o potencial de troca e correlação, uma vez calculado este termo, a teoria torna-se consistente. Com isso o cálculo de propriedades do sistema é feito de maneira primorosa, aproximando-se dos dados obtidos experimentalmente, conferindo confiabilidade ao método.

2.2.1 Funcional de Troca e Correlação

O método de Aproximação do Gradiente Generalizado (do inglês, *Generalized-Gradiente Approximation-GGA*) utiliza aproximação não local. O funcional de troca e correlação E_{XC}^{GGA} é escrito como a soma de duas partes, uma para a troca e outra para a correlação,

$$E_{XC}^{GGA} = E_X^{GGA} + E_C^{GGA}. \quad (49)$$

Um dos funcionais de troca, com densidade corrigida pelo gradiente, mais utilizados é o de Becke de 1988 conhecido como B88, escrito como segue

$$E_X^{B88} = E_X^{LSDA} - b \sum_{\sigma=\alpha,\beta} \int \frac{(\rho^\sigma)^{\frac{4}{3}} \chi_\sigma^2}{1 + 6b\chi_\sigma \sinh^{-1} \chi_\sigma} d\mathbf{r}, \quad (50)$$

em que $\chi_\sigma \equiv |\nabla \rho^\sigma| / (\rho^\sigma)^{\frac{4}{3}}$, $\sinh^{-1} = \ln[x + (x^2 + 1)^{\frac{1}{2}}]$ e b é um parâmetro semi-empírico que vale $0,0042u.a.$ obtido após o cálculo de energia de troca pelo método HF realizado e ajustado para vários átomos, onde E_X^{LSDA} é escrito como,

$$E_X^{LSDA} = -\frac{3}{4} \left(\frac{6}{\pi} \right)^{\frac{1}{3}} \int \left[(\rho^\alpha)^{\frac{4}{3}} + (\rho^\beta)^{\frac{4}{3}} \right] d\mathbf{r} \quad (51)$$

Outros funcionais de correlação corrigidos pelo gradiente bastante utilizados atualmente são o de Lee-Yang-Parr (LYP) e os funcionais de Becke. Contudo, os funcionais híbridos são muito empregado, esses funcionais são uma combinação linear dos funcionais

DFT puros e o termo de troca calculado pelo método Hartree-Fock (E_X^{HF}). Para sistemas de camada fechada, o termo de troca é escrito como:

$$E_X = -\frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n \left\langle \phi_i^{KS}(1) \phi_j^{KS}(2) \left| \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \phi_j^{KS}(1) \phi_i^{KS}(2) \right\rangle. \quad (52)$$

O funcional híbrido empregado nesta pesquisa foi o B3LYP, que possui 3 parâmetros de Becke e mais os funcionais de correlação LYP. O B3LYP é definido como:

$$E_{XC}^{B3LYP} = (1 - a_0 - a_x) E_X^{LSDA} + a_0 E_X^{HF} + a_x E_X^{B88} + (1 - a_C) E_C^{VWN} + a_C E_C^{LYP}, \quad (53)$$

onde o termo de correlação E_C^{VWN} provém do processo de aproximação da densidade local-LSDA (*Local-Spin-Density Approximation*) desenvolvido por Vosko-Wilk-Nusair. Os demais parâmetros $a_0 = 0.20$, $a_x = 0.72$ e $a_C = 0.81$ foram escolhidos de maneira a imprimir os resultados para a energia de otimização e energias atômicas totais.

é computacionalmente mais vantajoso. Funcionais híbridos, tal como B3LYP, escolhido como funcional de troca e correlação para calcular a energia eletrônica mínima, melhoram a precisão dos parâmetros fornecidos pelo método (BECKE, 1993; LEE; YANG; PARR, 1988).

2.2.2 Conjunto de Funções de Bases

As funções do tipo Slater, anteriormente descritas, descrevem com grande eficiência sistemas monoatômicos e diatômicos. No entanto, mesmo integrais para orbitais monoelétrônicos são de difíceis resolução, limitando o emprego dos orbitais do tipo Slater (STOs) para sistemas pequenos. Para moléculas maiores é comum a utilização de funções do tipo gaussiana (GTO do inglês "Gaussian Type Orbital") conhecidas. Uma maneira de facilitar os cálculos e, conseqüentemente, tornar ágil o processo computacional é a combinação linear de gaussianas, onde duas ou mais gaussianas primitivas (PGTOs do inglês "Primitive Gaussian-Type Orbitals") são combinadas linearmente, gerando gaussianas contraídas, conhecidas como contração de orbitais do tipo gaussiana (CGTO do inglês "Contracted Gaussian-Type Orbitals"). A vantagem computacional dar-se justamente pelo fato de que somente as gaussianas contraídas serão calculadas. Na literatura há varios conjuntos de

bases aplicados à cálculos de sistemas poliatômicos, segue uma pequena descrição da base empregada neste trabalho.

O conjunto de base Pople utilizada neste trabalho corresponde ao conjunto de bases K-nlmG também conhecidas como valência desdobrada, onde k é a quantidade de PGTOs contraídas para representar os orbitais mais internos, ditos orbitais de caroço e nlm refere-se a quantidade de PGTOs utilizadas para descrever os orbitais mais externos, de valência. A base utilizada no presente trabalho foi a 6-311+G(2d,p), nesta 6 GTOs contraídas descreve cada orbital do caroço, 3 GTOs contraídas descrevem a parte mais interna da camada de valência, 1 GTO descreve a parte mais externa da camada de valência e mais 1 GTO para descrever a parte externa da camada de valência mais distante. Assim, a função de base utilizada foi escolhida de modo que, o orbital interno é representado por três gaussianas, e os orbitais médios e externos representados como gaussianas únicas (6-311 G).

Uma vantagem de utilizar essas bases é que pode-se acrescentar a elas funções de polarização e funções difusas. No nosso caso, foram acrescentadas duas funções de polarização de simetria d aos átomos pesados, e funções de polarização de simetria p aos átomos mais leves, hidrogênio (2d,p). As funções de polarização são muito importantes, elas carregam informações sobre deformações ocorridas nos orbitais que participaram de interações, alguns estudos revelam que essas funções mostram-se fundamentais para reprodução de resultados experimentais (FURET et al., 1985).

As funções difusas nos fornece informações sobre uma região maior do espaço dos orbitais ocupados. Estas fazem-se essenciais sempre que se possui elétrons fracamente ligados. A utilização de funções difusas melhoram a representação de sistemas com densidade eletrônica relevantes, contudo se encontram distantes do núcleo. Estas funções (s, p_x, p_y, p_z) são adicionadas aos átomos pesados e simbolizadas pelo sinal positivo (+) (SZABO; OSTLUND, 1989).

2.3 *Descritores Moleculares*

Estudos quantitativos de relação estrutura-atividade QSAR (do inglês, “*Quantitative structure-activity relationship*”) são de inquestionáveis importância para a química e

bioquímica. Esse modelo buscar compreender os prováveis efeitos que uma estrutura química de um composto pode produzir enquanto interage com o receptor biológico.

A interação de substâncias que apresentam propriedades terapêuticas com o alvo específico é governada por um conjunto extenso de descritores moleculares, que de maneira direta influenciam nestas interações. Portanto descritores moleculares são poderosas ferramentas para prever propriedades de moléculas, investigar possíveis interações entre sistemas, apontando a viabilidade de realizar-se uma investigação experimental, que por sua vez é financeiramente mais custosa. Sendo assim, a procura por diferentes descritores vem aumentando, pois acredita-se que eles sejam a chave para solucionar problemas em estudos sobre estrutura-atividade (KARELSON; LOBANOV; KATRITZKY, 1996).

No estudo com candidatos à fármacos derivados do cardanol um conjunto foi utilizado um conjunto de descritores estruturais, físico-químicos e eletrônicos na intenção de que esses descritores nos forneça maior detalhamento sobre a atividade biológica de moléculas com estruturas semelhantes, porém que apresentem respostas biológicas diferentes. A identificação adequada desses parâmetros físico-químicos, eletrônicos ou químico-quânticos, e ainda aqueles que as distingue geométrico e topologicamente, é de fundamental importância para alicerçar o planejamento racional de novos compostos que apresentem perfil terapêutico (ARROIO et al., 2010; TAVARES, 2004).

No que se refere a conformação da molécula e informações sobre sua orientação tridimensional determina-se duas categorias diferentes de descritores, os bidimensionais e os tridimensionais. A abordagem realizada nesta pesquisa é referente a descritores bidimensionais como medidas de elementos da estrutura química, estrutura geométrica e topológica e parâmetros eletrônicos calculados por métodos quânticos. Dentro da categoria dos bidimensionais, os descritores fundamentais podem ainda ser subdivididos em eletrônicos, geométricos, constitucionais, topológico e lipofilia.

Eletrônicos: Obtidos por meio de cálculos computacionais químico-quântico, envolvem estudo sobre a densidade de distribuição eletrônica sob a perspectiva de ação de substituintes. Fazem parte destes descritores as energias dos orbitais de fronteira, cargas atômicas, potenciais de ionização, energia eletrônica, momento de dipolo, afinidade eletrônica, calor de formação, polarizabilidade, refratividade molar, entre outros.

Os orbitais de fronteira HOMO (*High Occupied Molecular Orbital*-Último Orbital Molecular Ocupado) e LUMO (*Lower Unoccupied Molecular Orbital*-Primeiro Orbital

Molecular Não-ocupado) são parâmetros importantes, eles representam, respectivamente, a capacidade doadora e receptora de elétrons.

A polarizabilidade molecular é uma medida da disposição que a molécula tem de modificar sua distribuição eletrônica em resposta a um campo elétrico externo, fazendo com que os elétrons da nuvem fiquem mais deslocalizados.

Constitucionais: Referentes aos elementos que compõem a estrutura molecular e às ligações, como massa molecular, número total de elétrons, número e tipo de ligação.

A constante de dissociação (pKa) é um parâmetro muito importante para este estudo especificamente, pois esta constante diz respeito quanto o meio de interesse encontra-se protonado, visto que todas as moléculas de interesse foram protonadas para realização dos cálculos e da investigação, afim de reproduzir fielmente as condições iniciais que serviram de ponto de partida para a investigação experimental. Se pKa do composto for igual ao pH do meio, 50% deste encontra-se em estado ionizado e 50% não ionizado. Compostos ácidos em pH com duas unidades logarítmicas menor encontram-se 99% não ionizados. Compostos básicos em pH com duas unidades logarítmicas menor encontram-se 99% ionizados. Portanto, o grau de ionização do fármaco depende do pKa (GROSS; SEYBOLD; HADAD, 2002).

Geométricos: A geometria molecular traça o arranjo espacial dos átomos em uma molécula, exemplos de descritores moleculares geométricos são volume molecular, a superfície de área polar (PSA do inglês "Polar Surface Area"), informações a respeito da superfície molecular.

Descritor de Lipofilia: O coeficiente de Partição (log P) é o logaritmo da razão entre a solubilidade de uma substância em um solvente orgânico e a solubilidade desta em água (octanol/água). Esse parâmetro refere-se a disposição do fármaco no organismo, sua permeabilidade através da membrana plasmática e caracteriza a hidrofobicidade de uma molécula (KARELSON; LOBANOV; KATRITZKY, 1996). O coeficiente de partição é dado por:

$$P = \frac{[droga]_{octanol}}{[droga]_{agua}} \quad (54)$$

onde $[droga]_{octanol}$ é a concentração da droga em solução no octanol e $[droga]_{agua}$ é a concentração da droga na água. Experimentalmente, uma amostra de fármaco é agitada

com uma mistura de octanol e água e sua concentração em cada camada é determinada pelo logP.

No entanto, a maioria dos fármacos são ácidos ou bases fracas que ionizam parcialmente quando dissolvidos em água. Então o pH da solução aquosa interfere na proporção entre as formas molecular e ionizadas da droga. E a constante de distribuição da droga é dada por:

$$D = \frac{[droga_{molecula}]_{octanol}}{[droga_{molecula}]_{agua} [droga_{ion}]_{agua}} \quad (55)$$

na qual $[droga_{molecula}]_{octanol}$ é a concentração da droga (forma molecular) em octanol, $[droga_{molecula}]_{agua}$ é a concentração da droga (forma molecular) em água e $[droga_{ion}]_{agua}$ é a concentração da droga (forma ionizada) em água. A constante de distribuição é dependente do pH e o termo logD é utilizado para refletir a lipofilicidade de uma droga. Para o estudo que segue utilizou-se o descritor lipofílico coeficiente de distribuição (logD), pois os compostos estão ionizados em pH fisiológico (BHARATE; KUMAR; VISHWAKARMA, 2016).

Os descritores calculados para este trabalho, com base nos descritores relevantes para o estudo de novos fármacos apresentados na literatura, foram o volume molecular, a polarizabilidade, os orbitais de fronteira HOMO, HOMO-1, LUMO e LUMO+1, a carga sobre o átomo de nitrogênio protonado, o logD, a constante de dissociação (pKa), a área de superfície polar (PSA) e momento de dipolo. O que esse trabalho busca, justamente é com base nos resultados experimentais de atividade biológica e concentração, identificar quais parâmetros moleculares são relevantes para que estas moléculas candidatas a novo fármaco, apresentem perfil terapêutico na inibição da acetilcolinesterase. Tomar conhecimento dos descritores mais relevantes que possam levar à interação ligante e receptor biológico, não é uma tarefa fácil, uma vez que o conjunto de descritores gera muitas possibilidades, sendo assim, fez-se necessária a utilização de um método estatístico como ferramenta de análise. Para a análise dos resultados escolhemos o método de análise multivariada Análise de Componentes Principais, que faz parte de umas das técnicas mais utilizadas que procura estabelecer uma possível correlação entre descritores teóricos e a atividade biológica (ARROIO et al., 2010).

2.4 *Análise de Componentes Principais*

A análise de componentes principais (PCA) (*Principal Components Analysis*) é um método estatístico multivariado que por meio da redução de dados, busca o reconhecimento de padrões. Outros métodos multivariados como a Modelagem Independente de Analogia de Classes (SIMCA) (do inglês, “*Soft Independent Modeling by Class Analogy*”) que é um método de classificação, e Regressão em Componentes Principais (PCR) (do inglês “*Principal Component Regression*”) que é um método de calibração, ou regressão por mínimos quadrados (PLS do inglês “*Partial Least Squares*”), derivam do PCA. Estes métodos estatísticos de análise de dados, quando aplicados em problemas químicos, são chamados de métodos quimiométricos (ZHOU et al., 2008; LOZANO et al., 2012).

A PCA foi criada por Pearson(1901) e Hotelling(1933), e provavelmente é o método multivariado mais utilizado para redução de dados (PEARSON, 1901). E consiste em fatorar a matriz de dados iniciais X, escalonando-a de maneira a transformar os descritores para serem comparados numa escala igual, mesmo que suas unidades sejam diferentes. Então, temos uma matriz com as variáveis de dados e uma segunda matriz com as amostras.

Os descritores escolhidos para trabalhar com cada grupo de moléculas derivadas do cardanol foram volume molar (V), polarizabilidade (Pol), HOMO (H), HOMO-1 (H-1), LUMO (L), LUMO+1 (L+1), carga do nitrogênio (qN), pKa, PSA e momento de dipolo (μ). Assim a matriz de dados inicial é uma matriz moléculas *versus* descritores.

Os elementos da matriz de dados iniciais possuem unidades de medida diferentes, e a comparação entre elementos de natureza diferente não é praticável. Para tornar essa comparação possível é preciso fazer com que estes elementos (variáveis) sejam adimensionais, com média nula e variância unitária. E um termo que represente todas as variáveis modificadas é escrito da seguinte maneira:

$$y_{mn} = \frac{x_{nm} - \bar{x}_m}{\sigma_m} \quad (56)$$

em que os índices m e n, representam, respectivamente, m-ésimo descritor e n-ésima molécula. E a média aritmética dos valores calculados para o m-ésimo descritor,

$$\bar{x}_m = \frac{\sum_{n=1}^N x_{nm}}{N} \quad (57)$$

é o desvio padrão do m-ésimo descritor é:

$$\sigma_m = \sqrt{\frac{\sum_{n=1}^N (x_{nm} - \bar{x}_m)^2}{N - 1}} \quad (58)$$

Tem-se agora uma nova matriz de variáveis simétrica de ordem $m \times m$, a matriz de covariância C , dada por:

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mm} \end{pmatrix}$$

A covariância mede o grau de relação linear entre duas variáveis. Um valor "grande" positivo indica dados correlacionados positivamente. Da mesma maneira que um "grande" valor negativo, indica que essas variáveis estão pouco correlacionadas. Em magnitude, a covariância mede o grau de redundância (SHLENS, 2014).

Com o propósito de diminuir a correlação entre as variáveis, os elementos fora da diagonal principal na matriz C , precisam ser minimizados, enquanto os termos da diagonal principal maximizados, isso significa maximizar o sinal medido pela variância (GROTH et al., 2013).

Para que isso ocorra utiliza-se a diagonalização da matriz de covariância C , de maneira a obter uma matriz P que equivale aos autovetores de C , cujo autovalor associado é a variância de cada componente principal. Assim, determinar a componente principal, aquela que tem maior relevância, resume-se em encontrar o autovetor ao qual esteja associado o maior autovalor, então

$$CP = VP \quad (59)$$

onde a matriz de autovalores é chamada de V .

De forma geral, PC1 é formada de maneira que tenha a maior parte da variância. PC2 é retirada da variância restante, e da mesma maneira gera-se a PC3, até que a PCN seja gerada. PC2 é então uma combinação linear de variáveis observadas não correlacionadas com a primeira coluna (PC1) que tenha máximo de resto da variação total. A matriz de variância pode então ser vista da seguinte maneira:

$$Var = \begin{pmatrix} \text{Variância da PC1} & 0 & \cdots & 0 \\ 0 & \text{Variância da PC2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \text{Variância da PCn} \end{pmatrix}$$

Geometricamente, os objetos são representados por uma nuvem de N pontos (amostras) em um espaço multidimensional, com um eixo para cada um dos descritores (variáveis que dizem respeito às amostras), o centro dos pontos é definido pela média de cada amostra, e a variância é a média dos quadrados da diferença dos N pontos com relação à média de cada amostra.

$$V_i = \frac{1}{n-1} \sum_{m=1}^n (X_{im} - \bar{X}_i)^2. \quad (60)$$

Os autovetores podem ser representados matricialmente da seguinte forma:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}$$

As componentes principais, obtidas em ordem decrescente de relevância, são escritas como uma combinação linear de variáveis modificadas

$$PC_j = a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n. \quad (61)$$

O grau com que cada variável é linearmente correlacionada é representado pela sua covariância:

$$C_{ij} = \frac{1}{n-1} \sum_{m=1}^n (X_{im} - \bar{X}_i)(X_{jm} - \bar{X}_j) \quad (62)$$

onde,

– C_{ij} = covariância de variáveis i e j ;

– \sum = soma sobre todos os n objetos;

– X_{im} = valor da variável i no objeto m ;

– \bar{X}_i = média de variável i .

O objetivo do PCA é rotacionar rigidamente os eixos desse espaço p-dimensional para novas posições, que são os eixos principais, de forma que o eixo principal 1 tenha maior variância que o eixo 2, e assim por diante, até o último eixo que tenha menor variância. Os eixos principais não são correlacionados, o PCA realiza rotações dos eixos cartesianos de forma que o antigo eixo horizontal contenha a maior informação sobre as distribuições das amostras, e o PC2 a segundo maior. Numa visualização gráfica desses pontos no espaço multidimensional pode ser visto na figura 15 que mostra a variância das PCs.

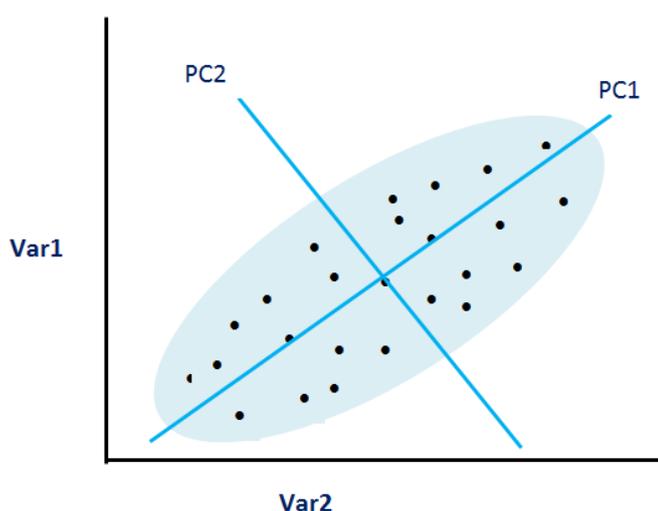


Figura 15 – Visualização gráfica da variância das PCs

Inicialmente os candidatos foram desenhados na forma tridimensional no *software* GaussView (DENNINGTON; KEITH; MILLAM, 2009). Para determinação da conformação mais estável pelo cálculo de otimização de geometria, foram escolhidos o método da Teoria do Funcional de Densidade (DFT) (B3LYP), com funções de base 6-311+G(2d,p), instaladas no pacote computacional Gaussian 09 (FRISCH et al., 2009) com o *Polarizable Continuum Model* (PCM)¹. Essa escolha foi realizada com base em trabalhos anteriores feitos com derivados obtidos a partir do LCC, para os quais esses métodos resultou em bons resultados para a obtenção de descritores utilizados no estudo de QSAR (PAULA et al., 2009; KIAMETIS et al., 2013)

¹ O PCM é um método *ab initio* para cálculos de efeito de solvente, que trata o solvente como um contínuo polarizável ao invés de tratá-lo como moléculas individuais, tornando possível o tratamento *ab initio* do solvente, este método encontra-se implementado em vários pacotes computacionais e detalhes a cerca do mesmo podem ser vistos nas referências (COSSI et al., 2002; TOMASI; MENNUCCI; CAMMI, 2005)

Como o problema é bidimensional, os padrões podem ser visualizados plotando as duas primeiras componentes principais, que contêm a maior parte da variância, caso essas não sejam suficientes para exibir as correlações, a terceira componente pode ser plotada auxiliando na identificação de similaridades. No nosso estudo as PCs foram construídas utilizando o *software* STATISTICA (OKLAHOMA, 2008). A tabela de dados iniciais foi construída para cada grupo de moléculas contendo os descritores volume[Å^3], polarizabilidade [Å^3], as energias dos orbitais de fronteira (HOMO, HOMO-1, LUMO e LUMO+1) [eV], carga sobre o nitrogênio [C], PSA [Å^2], momento de dipolo [Debye], logD e pKa, com suas respectivas unidades.

2.5 *Dinâmica Molecular Clássica*

A Modelagem Molecular Clássica é um método amplamente aplicado para investigar sistemas moleculares pela aplicação de funções de energia potencial da mecânica molecular². Usando modelos simples, da abordagem da mecânica molecular, as simulações de Dinâmica Molecular (DM) consistem na integração numérica das equações do movimento de Newton. Considerando um sistema tridimensional, para um átomo i de massa m_i e posição \vec{r}_i , a relação entre sua velocidade e momento, p_i , é dada por:

$$\frac{dr_i}{dt} = \frac{p_i}{m_i}. \quad (63)$$

A força que o sistema exerce sobre o átomo i , é dada pelo negativo do gradiente da função energia potencial com referência a posição do átomo i , e tem a forma:

$$F_i = -\frac{dU}{dr_i} = -\nabla U(\vec{r}_i), \quad (64)$$

e a equação de movimento Newtoniana para o átomo i , é dada pela expressão:

$$\frac{dp_i}{dt} = F_i. \quad (65)$$

² Dentre as funções utilizadas pela mecânica molecular clássica, estão o oscilador harmônico e potencial de Coulomb.

Considerando uma determinada componente do vetor posição \vec{r}_i , ou seja, a posição ao longo de uma única dimensão x , em um tempo particular t , após um pequeno intervalo de tempo Δt , a posição pode ser escrita como uma série de Taylor:

$$x(t + \Delta t) = x(t) + \frac{dx(t)}{dt} \Delta t + \frac{d^2x(t)}{dt^2} \frac{\Delta t^2}{2} + \dots \quad (66)$$

A posição, velocidade e aceleração, respectivamente, $x(t)$, $dx(t)/dt$ e $d^2x(t)/dt^2$, são suficientes para resolver numericamente as equações de movimento. Para a dimensão x , a segunda Lei de Newton, escrita de acordo com a equação abaixo, descreve a aceleração

$$\frac{d^2x(t)}{dt^2} = \frac{F_x}{m} \quad (67)$$

onde F_x é a componente da força que atua no átomo, paralela à direção x .

Essas equações diferenciais podem ser resolvidas por métodos de diferenciais finitas. O Algoritmo de Verlet é derivado das expansões da série de Taylor para as posições, utiliza as posições e acelerações atômicas no tempo t e as posições do passo anterior, $x(t - \Delta t)$, para determinar as novas posições em $(t + \Delta t)$. De maneira que as expansões de Taylor para as posições são:

$$x(t + \Delta t) = x(t) + \frac{dx(t)}{dt} \Delta t + \frac{d^2x(t)}{dt^2} \frac{\Delta t^2}{2} + \frac{d^3x(t)}{dt^3} \frac{\Delta t^3}{6} + O\Delta t^4 \quad (68)$$

$$x(t - \Delta t) = x(t) - \frac{dx(t)}{dt} \Delta t + \frac{d^2x(t)}{dt^2} \frac{\Delta t^2}{2} - \frac{d^3x(t)}{dt^3} \frac{\Delta t^3}{6} + O\Delta t^4. \quad (69)$$

Adicionando as equações 68 e 69, para o algoritmo de Verlet, tem-se:

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + \frac{d^2x(t)}{dt^2} \Delta t^2 + O\Delta t^4. \quad (70)$$

O termo $+O\Delta t^4$ é o erro de truncamento, e a velocidade das partículas é dada pela equação 71 abaixo:

$$v(t) = \frac{x(t + \Delta t) - x(t - \Delta t)}{2\Delta t} + O\Delta t^2. \quad (71)$$

A velocidade não aparece explicitamente nas equações do algoritmo para a determinação das trajetórias, mas seu valor é indispensável para o cálculo da energia cinética. Como no início da simulação as posições e velocidades iniciais são aquelas referentes as configurações iniciais do sistema, e portanto não se conhece essas variáveis em passos

de simulação anteriores ou superiores, as equações para posição e velocidade podem ser truncadas no termo de segunda e retirada também o erro de truncamento, de maneira que as equações 70 e 71, são escritas respectivamente:

$$x(t + \Delta t) = x(t) + \frac{dx(t)}{dt} \Delta t + \frac{d^2x(t)}{dt^2} \frac{\Delta t^2}{2} \quad (72)$$

$$v(t + \Delta t) = v(t) + \frac{a(t + \Delta t) + a(t)}{2} \Delta t \quad (73)$$

onde $a = \frac{d^2x(t)}{dt^2}$ (GRUBMÜLLER et al., 1991; ALLEN et al., 2004).

Os métodos de simulação de dinâmica são utilizados para obter informações relativas à evolução temporal das conformações do sistema (proteínas e outras macromoléculas biológicas), bem como informações cinéticas e termodinâmicas tempo-dependentes, permitindo o conhecimento de inúmeros aspectos biomoleculares como estruturais, de reconhecimento e função. Os dados fornecidos pelas trajetórias na DM são geralmente posições atômicas, velocidades e energias de *single-point*, em nível de comportamento microscópico. As propriedades macroscópicas como pressão, capacidade de calor, energia interna, volume, temperatura, entropia, dentre outras, são deduzidas através de expressões matemáticas da mecânica estatística, que estabelece uma conexão entre as distribuições e movimentos dos átomos e moléculas e as observáveis macroscópicas.

Em geral, os métodos de DM dependem extremamente de uma função de energia apropriada para reproduzir a energia de *landscape* do sistema. A utilização de funções de energia que descrevam convenientemente as interações intermoleculares e intramoleculares é essencial para que a simulação de DM faça-se satisfatória (ADCOCK; MCCAMMON, 2006).

2.5.1 Funções de Energia Potencial

Na DM a energia total do sistema é dada pelo potencial nuclear em função das coordenadas atômicas. O conjunto completo de potenciais clássicos de energia e os parâmetros ajustáveis a eles associados é denominado campo de força (JORGENSEN; TIRADORIVES, 2005). Os campos de força, em sua maioria, são puramente aditivos, e a função de energia contém termos para interações internas e externas, como para os campos de força CHARMM (BROOKS et al., 1983), AMBER (WEINER; KOLLMAN, 1981; WANG

et al., 2004), OPLS (JORGENSEN; MAXWELL; TIRADO-RIVES, 1996), GROMOS (FRATERNALI; GUNSTEREN, 1996), etc. A função de energia potencial utilizada normalmente pelo programa CHARMM está fundamentado em cargas de ponto fixo, e é descrito de acordo com a equação 74 abaixo:

$$\begin{aligned}
 U(\vec{R}) = & \sum_{\text{ligações}} K_b(b - b_0)^2 + \sum_{\text{ângulos}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{Urey-Bradley}} K_{UB}(S - S_0)^2 \quad (74) \\
 & + \sum_{\text{diedrais}} k_\varphi(1 + \cos(n\varphi - \delta)) + \sum_{\text{impróprios}} K_\omega(\omega - \omega_0)^2 \\
 & + \sum_{\text{não-ligados}} \left\{ \varepsilon_{ij}^{\text{min}} \left[\left(\frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\} + \sum_{\text{resíduos}} U_{CMAP}(\varphi, \psi)
 \end{aligned}$$

As contribuições intramoleculares contém termos de ligações (índices b , do inglês *bond*), ângulos de valência (θ), Urey-Bradley (UB, S), ângulos diedrais (ε), ângulos impróprios (ω) e contribuições de correções torcionais para o *backbone* ($CMAP, \varepsilon, \psi$). Os parâmetros K_b , K_ε , K_{UB} , K_θ e K_ω são as específicas constantes de força e as variáveis com sub-índice 0 são valores referentes ao equilíbrio. Considerando a ligação entre três átomos A–B–C o termo de Urey-Bradley é uma função quadrática da distância ,S, entre os átomos A e C. O termo de ângulo diedral impróprio é uma função quadrática usada para pontos de ramificação, como no caso de um C_α da proteína, onde átomos A–B–D ligam-se a um átomo central, o ângulo impróprio é um ângulo diedral fictício. Os termos de Urey-Bradley e os diedros impróprios são usados para aperfeiçoar os ajustes no espectro vibracional para movimentos fora do plano.

As contribuições intermoleculares ou não-ligadas envolvem termos de interações Coulombianas entre cargas pontuais (q_i e q_j), interações do tipo Lennard-Jones (LJ), utilizadas para tratar a repulsão caroço-caroço, e termos de van der Waals para interações atrativas. A constante dielétrica relativa, ε , frequentemente, é definida para simulações que incluem representações de solvente explícito e ε_0 é a permissividade elétrica no vácuo, sendo o termo eletrostático inversamente proporcional à r_{ij}^2 , o quadrado da distância de interação entre os átomos. No campo de força CHARMM, os valores para $\varepsilon_{ij}^{\text{min}}$ são calculados mediante média geométrica ($\varepsilon_{ij}^{\text{min}} = \sqrt{\varepsilon_{ii}^{\text{min}} \varepsilon_{jj}^{\text{min}}}$) e R_{ij}^{min} mediante média aritmética, $R_{ij}^{\text{min}} = (R_i^{\text{min}} + R_j^{\text{min}})/2$, onde i e j são os índices dos átomos que estão interagindo e R_{ij}^{min} é a distância na qual o termo de LJ tem seu mínimo (BROOKS et al., 2009; HALGREN, 1992; ADCOCK; MCCAMMON, 2006).

2.5.2 Simulação de Dinâmica Molecular (DM)

As simulações de DM são técnicas para determinar o movimento das partículas de um sistemas, para o qual o potencial de interação entre as partículas (função de energia potencial) já é conhecido. Esse de descrição das posições instantâneas e velocidades instantâneas das partículas, pode ser dividido nas seguintes estágios: estabelecer a configuração inicial para o sistema, a computação das forças exercidas sobre cada partícula em razão das interações intermoleculares, movimentação das partículas ou equilibração de fase, *ensembles* e controles de simulação, e finalmente, análise das configurações (trajetórias) obtidas.

A configuração inicial do sistema, tem como frequente metodologia a criação de uma "caixa de simulação", que geralmente é uma caixa cúbica, onde todas as moléculas do sistema se encontram. Uma boa prática das condições iniciais é que o sistema não contenha interações com altas energias, que podem por sua vez, causar instabilidades na simulação. Se o sistema possuir um arranjo experimental disponível (por exemplo, estrutura de raio X) esta será a configuração inicial suficiente para o início da simulação. A estrutura cristalográfica mais comum é a rede cúbica de face centrada, mas podem ser escolhidas outras geometrias adequadas. As dimensões da caixa de simulação devem ser escolhidas de maneira que reproduza a densidade numérica, N/V do sistema real e suas condições termodinâmicas. A caracterização do sistema considera a reprodução das condições fisiológicas que o aproximam do modelo real, e os sistemas são tratados, de forma que, para a maior parte das simulações o ambiente escolhido é um solvente aquoso, podendo ser uma solvatação implícita ou explícita. Para o caso de solvente explícito, existem alguns modelos explícitos para a água, que incluem os modelos TIP3P, TIP4P (JORGENSEN et al., 1983), SCP e SCPC/E (ADCOCK; MCCAMMON, 2006), e os modelos mais conhecidos são consistentes com a aproximação SHAKE (ANDERSEN, 1983) para manutenção da geometria molecular, e ainda para diminuir os efeitos de fronteiras, condições de contorno são utilizadas (LEACH, 2001).

Para iniciar a simulação de DM, a solução da equação de movimento, é necessário fornecer não somente as posições das partículas, mas também as suas velocidades iniciais, já que as equações diferenciais são de segunda ordem.

Durante a simulação as forças entre as partículas decaem com a distância entre estas, interações Coulombianas, a medida mais aplicada é na caixa de simulação, substituir a força de interação entre duas partículas por uma força efetiva, soma das interações entre todas as imagens periódicas dessas partículas, pelo método de soma infinita. O método de soma de Edwald (YORK; YANG, 1994) oferece uma abordagem rigorosa para a avaliação das interações eletrostáticas, além do Edwald, existem métodos que são variantes deste, como o método *Particle Mesh Ewald* (PME), que analisa as forças diferenciando analiticamente, diminuindo significativamente a demanda de memória computacional. Pela eficiência em velocidade de processamento e rigor, o PME está implementado em muitos pacotes de simulação (TOUKMAJI; BOARD, 1996; WANG; GAO; FANG, 2016).

2.5.2.1 *Ensemble* Estatístico

Um sistema clássico com N partículas interagindo em equilíbrio, a uma temperatura constante T, no *ensemble* canônico é reproduzido pela distribuição de Boltzmann. A média da quantidade $G(\vec{r}^N)$, é unicamente uma função da posição \vec{r} de todas as N partículas e é expressada como uma integral de espaço de fase incluindo a energia potencial $U(\vec{r}^N)$:

$$\langle G \rangle = \frac{\int G(\vec{r}^N) \exp(-\beta U(\vec{r}^N)) d\vec{r}^N}{\int \exp(-\beta U(\vec{r}^N)) d\vec{r}^N} \quad (75)$$

na equação 75 acima $\beta = \frac{1}{k_B T}$, onde k_B é a constante de Boltzmann.

Se, depois de um tempo suficientemente grande, sob certas condições, o sistema em equilíbrio explora todo o espaço de fase de microestados com a mesma energia, onde a densidade de probabilidade é não nula, ou seja todos os microestados acessíveis são igualmente prováveis nesse longo período de tempo, esse sistema é *ergódico*. Essa é a hipótese ergódica no qual a média de uma amostragem ao longo do tempo, equivale a média no *ensemble* que é definida como:

$$\langle G \rangle_t = \lim_{M \rightarrow \infty} \langle G \rangle_M = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{\mu=1}^M G(\vec{r}^N(t_\mu)) \quad (76)$$

onde $\langle G \rangle_M$ é a média sobre as medições M, da quantidade G em todos os tempos $\{t_\mu\}_{\mu=1}^M$. Para uma boa amostragem, ou seja, um grande número de medições M, a média $\langle G \rangle_M$ é considerada uma aproximação conveniente para a média do *ensemble* em equilíbrio (SALINAS, 1997).

No decorrer da simulação, integração das equações de Newton, são mantidas constantes uma coleção de configurações e propriedades que caracteriza o estado do sistema. Essa coleção de todos os prováveis sistemas que possuem estados microscópicos diferentes, mas participam de um único estado macroscópico (termodinâmico), chamamos de *ensemble* estatístico. Existem vários *ensemble* com características diferentes, nos quais distintos parâmetros permanecem constantes simultaneamente como o *ensemble* canônico (NVT): conjunto de sistemas nos quais o estado termodinâmico é caracterizado por um número fixo de átomos, N, fixado o volume, V, e fixada a temperatura, T; o *ensemble* isobárico-isoentálpico (NPH- número de átomos, pressão e entalpia constantes), o *ensemble* gran-canônico (μ VT- potencial químico, volume e temperatura constantes), o *ensemble* isobárico-isotérmico (NPT- número de átomos, pressão e temperatura constantes), dentre outros (CHANDLER, 1987).

Grande parte das simulações de DM tem suas condições iniciais representadas pelo *ensemble* microcanônico (NVE), onde a energia total E é uma constante de movimento, e sua média temporal são iguais a sua média no *ensemble*. Contudo as experiências laboratoriais (por exemplo, cristalografia de raios X) são realizadas com volume e temperatura constantes (*ensemble* canônico) ou à pressão e temperatura constantes (*ensemble* isobárico-isotérmico), sendo uma prática na DM reproduzir, sob condições fisiológicas, esses estados (SCHLICK, 2010). Inicialmente à energia constante, a temperatura flutua como consequência da interconversão espontânea das componentes cinética e potencial da energia total. A energia interna total pode ser escrita como a soma das contribuições cinéticas (\mathcal{K}) e potencial (\mathcal{U}) envolvendo a soma sobre todos os pares de moléculas:

$$E = \langle \mathcal{K} \rangle + \langle \mathcal{U} \rangle \quad (77)$$

A temperatura e pressão podem ser calculados via teorema do virial. Para o caso atômico, somando sobre o $3N$ termos, a energia cinética tem a forma:

$$\left\langle \sum_{i=1}^N |p_i|^2 / m_i \right\rangle = 2 \langle \kappa \rangle = 3Nk_B T, \quad (78)$$

onde k_B é a constante de Boltzmann, m_i e p_i são, respectivamente, massa e momento do átomo i e N é o número total de átomos. A equação 78 acima fundamenta-se no princípio

da equipartição e a função de temperatura cinética instantânea pode ser determinada como:

$$\mathcal{T} = 2\mathcal{K}/3Nk_B = \frac{1}{3Nk_B} \sum_{i=1}^N |p_i|^2/m_i. \quad (79)$$

Sendo a temperatura $\mathcal{T}(t)$ no instante t, e as velocidades alteradas por um fator λ , as variações de temperatura podem ser calculadas da seguinte maneira:

$$\Delta\mathcal{T} = \frac{1}{2} \sum_{i=1}^N 2 \frac{m_i(\lambda v_i)^2}{N_{df}k_B} - \frac{1}{2} \sum_{i=1}^N 2 \frac{m_i v_i^2}{N_{df}k_B} \quad (80)$$

$$\Delta\mathcal{T} = (\lambda^2 - 1)\mathcal{T}(t)$$

$$\lambda = \sqrt{\mathcal{T}_0/\mathcal{T}(t)}$$

a temperatura tem como controle simples, multiplicar as velocidade atômicas instantâneas pelo fator λ . Este método tem como variante o termostato de Berendsen (HÜNENBERGER, 2005), que modera a razão com a qual a temperatura do sistema chega ao valor determinado. Assumindo o ensemble NVT, para o qual a energia total não é conservada, o início da simulação equipara-se de fato a um banho térmico, fazendo desse *ensemble* indispensável para o estagio de termalização, quando tratamentos térmicos como o *Simulated Annealing* são utilizados (KARPLUS; PETSKO et al., 1990; ERPENBECK; WOOD, 1984).

No *ensemble* NPT o volume do sistema pode flutuar alterando as dimensões da caixa de simulação e também pelo reescalonamento das posições, o fator $\gamma = \sqrt[3]{p(t)/p_0}$ é assim, responsável por manter a taxa de flutuação equilibrada, de maneira similar como feito para a temperatura. A pressão pode ser calculada com a contribuição do teorema do virial de acordo com a equação:

$$p(t) = \frac{Nk_B T}{V} - \frac{1}{3V} \sum_{i=1}^N \sum_{j>1}^N r_{ij}(t) \cdot F_{ij}(t) \quad (81)$$

onde $r_{ij}(t) = r_j(t) - r_i(t)$ é o vetor posição das moléculas i e j, e $F_{ij}(t)$ a força intermolecular entre i e j, no instante t. Existem alguns métodos em simulações de DM que fazem com que a pressão fique corretamente constante durante a simulação, como por exemplo, o método dinâmica de Langevin (FELLER et al., 1995).

2.5.2.2 Análise e armazenamento trajetórias geradas na DM

O estágio inicial da dinâmica a partir da configuração inicial é a termalização, e deve ser inspecionado com muito cuidado até que a configuração inicial não mais persistir. Neste estágio, como mencionado acima, a temperatura e a pressão devem ser escalonadas, e o banho térmico sistematicamente observado, até que o sistema se encontre no equilíbrio. A credibilidade dos resultados no estágio seguinte, produção, depende do monitoramento adequado do estágio de equilibração.

O estágio de produção é a reunião das trajetórias (coordenadas ou velocidades atômicas) no tempo de simulação que sucede aquele do equilíbrio. A partir dessas trajetórias é que serão calculadas propriedades físico-químicas do sistema e médias sobre os *ensembles* das trajetórias (SCHLICK, 2010).

Dentre várias propriedades que podem ser calculadas a partir das trajetórias estão as propriedades estruturais, como a função de distribuição de pares, $g(r)$, que produz a probabilidade de se encontrar uma partícula β em uma camada esférica de espessura dr , e sua distância até a partícula α vale r posicionada na origem. Considerando as coordenadas cartesianas retiradas das trajetórias de simulação, a função radial de pares pode ser dada por:

$$g(r) = \rho^{-2} \left\langle \sum_i \sum_{j \neq i} \delta(r_i) \delta(r_j - r) \right\rangle = \frac{V}{N^2} \left\langle \sum_i \sum_{j \neq i} \delta(r - r_{ij}) \right\rangle \quad (82)$$

Assim sendo, o número de partículas $N(r)$ que localiza-se na faixa estreita Δr a uma distância r da origem, remete a distribuição radial dada pela seguinte equação:

$$g(r) = \frac{N(r)}{4\pi r^2 N} \quad (83)$$

O deslocamento quadrático médio RMSD (do inglês, *root mean square displacement*) calcula a diferença estrutural do sistema em um determinado passo de simulação com relação as configurações iniciais, ou seja, realiza uma comparações entre as coordenadas atômicas de estruturas proteicas, por exemplo, adquiridos das trajetórias de DM, que tem como dois principais interesses:

Observar a evolução temporal, dado um conjunto de átomos específicos é realizada uma média de RMSD a cada passo de simulação, aferindo as mudanças ocorridas durante a simulação, que são indicadas por flutuações em torno de uma posição de equilíbrio. A

evolução temporal de RMSD é geralmente visualizada através de um gráfico de RMSD médio versus tempo, e o RMSD médio é calculado segundo a equação:

$$\langle RMSD \rangle = \sqrt{\frac{\sum_{j=1}^N [r_j(t_i) - r_0]^2}{N}} \quad (84)$$

onde $r_j(t_i)$ é a distância do átomo j no passo de simulação i , r_0 é a distância de referência e N é o número total de átomos.

Analisar a média do RMSD por átomo, que é encontrado gerando as médias dos RMSD, para cada átomo, em todos os passos da simulação, e tem como finalidade principal, por comparação, encontrar quais átomos ou regiões do sistema flutuam mais ou menos. A média do RMSD é encontrado utilizando a equação:

$$\langle RMSD_j \rangle = \sqrt{\frac{\sum_{i=1}^M [r_j(t_i) - r_0]^2}{M \cdot \Delta t}} \quad (85)$$

na equação acima M é o número de configurações e Δt o intervalo de tempo entre os passos de simulação, de maneira que multiplicando M por Δt , tem-se o tempo de simulação total.

2.6 Docking Molecular

Muitas estratégias computacionais permeiam os conceitos da descoberta de novos fármacos, como técnicas de *Virtual screening* (VS), usada na descoberta de novas drogas, que busca numa biblioteca de pequenas moléculas, quais destas mais provavelmente se ligariam no alvo. Ainda quando sabe-se pouco a respeito do ligante e informações estruturais não estão disponíveis para o alvo, abordagens como QSAR podem ser aplicadas. Com relação ao Desenho Racional de Fármacos Baseado em Estrutura (DRBE), o *docking* molecular é o método *in silico* mais comum e tem sido amplamente utilizado na descoberta e no desenho de novos compostos bioativos (BOHLOOLI; SEPEHRI; RAZZAGHI-ASL, 2017).

O *Docking* molecular ou ancoragem molecular, tem como objetivo básico a predição da estrutura do complexo receptor-ligante, ou seja, a predição do modo de ligação de uma molécula menor, o ligante, que pode ser um inibidor, na região de ligação ou sítio biológico do alvo molecular (uma molécula receptora vinculada a alguma doença ou processo fisiopatológico), pretendendo ainda prever a afinidade entre o ligante e o receptor. A ancoragem utiliza métodos computacionais e pode ser resumida em duas fases: primeiro,

a amostragem das conformações do ligante no sítio de ação biológica do receptor, que necessita de um algoritmo robusto para de forma eficiente explorar a hipersuperfície de energia e prever a conformação e a orientação preferencial do ligante, usualmente referida como pose; e segundo, coordenar as conformações por intermédio de uma função de avaliação (do inglês *scoring function*), o que significa, empregar um modelo de ponderação da energia livre de ligação, *scorings*, que diferencie os distintos modos de ligação do mesmo ligante, ou que, sendo desiguais os ligantes, selecione aquele com maior afinidade pelo sítio ativo (CHEUNG et al., 2012; MENG et al., 2011).

No processo de reconhecimento molecular, receptor e ligante experimentam mudanças conformacionais gerando centenas de milhares de graus de liberdade devido ao grande número de possíveis modos de ligação, e gerar todas as possibilidades conformacionais é computacionalmente muito caro. Vários algoritmos de amostragem foram desenvolvidos e são extensivamente utilizados em *softwares* de *docking* molecular. Dentre os *softwares* de *docking* mais usados estão, FlexAID (GAUDREAUULT; NAJMANOVICH, 2015), BetaDock (KIM et al., 2011), AutoDock Vina (TROTT; OLSON, 2010), Blaster (IRWIN et al., 2009), GOLD (VERDONK et al., 2003), DOCK (EWING et al., 2001), DARWIN (TAYLOR; BURNETT, 2000), etc.

2.6.1 Funções de *Scoring*

Para que os modos de ligação gerados e selecionados sejam eficazes na descrição da interação receptor-ligante, o modelo para classificar cada conformação precisa ser confiável e o custo computacional viável, já que, como dito anteriormente, a estimativa da energia livre de ligação do complexo receptor-ligante é computacionalmente dispendiosa. A necessidade de funções de avaliação mais ágeis e úteis aos programas de *docking*, direciona estudos para o desenvolvimento de funções que adotem aproximações para a avaliação do complexo. O *docking* computacional é realizado, frequentemente, empregando um campo de força simples da mecânica molecular e explorando uma região mais ampla do espaço conformacional, dependendo do programa utilizado os campos de força podem variar entre GROMOS, AMBER, OPLS ou CHARMM. A função de avaliação implementada no programa AutoDock (NAMASIVAYAM; GÜNTHER, 2007) é fundamentado no campo de força molecular AMBER.

Na abordagem adotada pelo AutoDock a energia livre de ligação resulta da diferença entre a energia do ligante e receptor no estado não-ligado, e a energia do ligante e do receptor no complexo. Isso em consequência do funcionamento da avaliação que ocorre em duas fases: análise das energias intramoleculares da transição do estado não-ligado para a conformação ligada, para cada uma das moléculas separadamente, e a análise energética intermolecular das moléculas no complexo.

O campo de força engloba seis termos de potenciais de interação que envolve contribuições entre átomos ligados e não-ligados e um termo de contribuição entrópica conformacional de ligação (ΔS_{conf}) de acordo com a seguinte equação:

$$\Delta G = (V_{ligado}^{L-L} - V_{n\tilde{a}o-ligado}^{L-L}) + (V_{ligado}^{R-R} - V_{n\tilde{a}o-ligado}^{R-R}) + (V_{ligado}^{R-L} - V_{n\tilde{a}o-ligado}^{R-L} + \Delta S_{conf}) \quad (86)$$

Na equação 86 acima, L diz respeito ao ligante e R ao receptor no complexo receptor-ligante, sendo válida para toda e qualquer tipo de molécula no complexo. Os primeiros dois termos de energia intramoleculares para o ligante nos estados ligados e não-ligado, e os dois termos posteriores as energias intramoleculares para o receptor nos estados ligado e não-ligado. O terceiro parentese contém as mudanças de energia intermolecular entre os estados ligado e não-ligado.

Os termos atômicos por pares incluem a dispersão/repulsão, ligações de hidrogênio, eletrostático e dessolvatação e conforme a equação abaixo:

$$\begin{aligned} V = & W_{vdW} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + W_{hligac\tilde{a}o} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{elec} \sum_{i,j} \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} \\ & + W_{sol} \sum_{i,j} (S_i V_j + S_j V_i) e^{(-r_{ij}^2/2\sigma^2)} \end{aligned} \quad (87)$$

As constantes de ponderação W são otimizadas para calibrar a energia livre empírica com base em um conjunto de complexos que foram caracterizados experimentalmente. O primeiro termo é o potencial 6-12 de Lennard-Jones para interações de dispersão/repulsão. Os parâmetros A e B foram extraídos do campo de força AMBER. O segundo termo é um potencial para ligações de hidrogênio. Os parâmetros C e D foram atribuídos para melhorar a descrição do potencial de ligação para O-H e N-H e S-H. As interações eletrostáticas são avaliadas segundo um potencial idêntico ao de Coulomb e o último termo representa o potencial de dessolvatação baseado no volume (V) dos demais átomos em torno de um

átomo definido, pesado por um parâmetro de solvatação (S) e um termo exponencial devido a distância.

O termo para entropia torsional é proporcional ao número de ângulos de rotação da molécula (N_{tors}):

$$\Delta S_{conf} = W_{conf} N_{tors} \quad (88)$$

O número de ângulos de rotações compreendem todos os graus de liberdade torcionais, abrangendo rotações de átomos de hidrogênio polar em grupos hidroxilo e similares (BROOIJMANS; KUNTZ, 2003; HUEY et al., 2007).

O AutoDock Vina (TROTT; OLSON, 2010) é uma ferramenta de *docking* molecular que apresenta excelentes resultados na estimativa dos modos de ligação preferenciais do receptor-ligante. Esse *software* possui implementado vários algoritmos de busca baseado no Algoritmo Genéticos Lamarckiano (do inglês *Lamarckian Generic Algorithms - LGA*) (MORRIS et al., 1998) que procura um mínimo global de energia de ligação do complexo. Durante o processo de ancoramento molecular a proteína é mantido rígida e as mudanças conformacionais são devidas ao ligante, incluir a flexibilidade é grandemente importante para que os resultados se aproximem do modelo realístico, sendo fundamental para a compreensão das interações farmacobiológicas, mas a flexibilidade do sistema ainda é um desafio para os métodos de ancoragem (JONES et al., 1997).

2.6.1.1 Detalhes da Simulação

A simulação de dinâmica molecular para o receptor foi realizada no pacote computacional NAMD (PHILLIPS et al., 2002), com o campo de força CHARMM36 (KLAUDA et al., 2010). As moléculas de água foram adicionadas usando o modelo TIP3P (JORGENSEN et al., 1983) e a caixa cúbica de simulação com dimensões 76 x 88 x 92 Å que pode ser visualizada na Figura 16. O sistema foi neutralizado e moléculas de NaCl foram adicionadas para reproduzir a solução fisiológica de concentração 0.15 μM. A distância entre pares de interações não-ligados foi truncado em 12 Å, *cut off*. O *particle mesh Ewald - PME* foi utilizado para as interações eletrostáticas de longo alcance. A simulação de DM foi executada considerando constante o número de moléculas, a pressão e a temperatura (*ensemble* NPT), usando a dinâmica de Langevin para manter constantes a temperatura e

pressão, 300 K e 1 atm, respectivamente. As trajetórias foram acumuladas com um passo de tempo, *timestep*, de 2fs. O tempo de amostragem total, incluindo a termalização e a relaxação, foi de 200 ns, a evolução temporal e a mobilidade do sistema foram computados pelo cálculo do deslocamento quadrático médio (RMSD).

A ancoragem molecular (*docking*) foi realizada no programa AutoDock Vina, usando o método LGA. As imagens tridimensionais, bem como a construção e visualização do sistema para a simulação de DM, foram produzidos usando o programa gráfico *Visual Molecular Dynamics* (VMD)(HUMPHREY; DALKE; SCHULTEN, 1996).

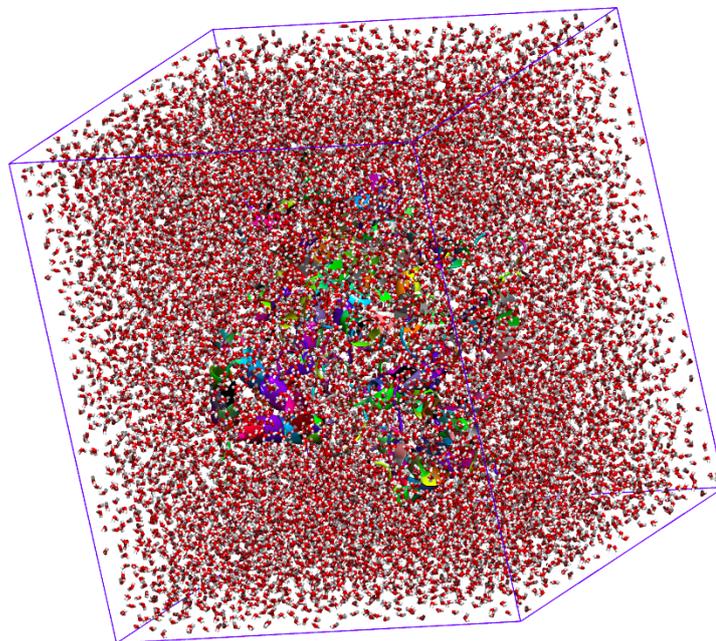


Figura 16 – Caixa de Simulação de Dinâmica Molecular contendo o solvente e a proteína (AChE: 4EY7), com 57480 átomos no total, dos quais 49149 correspondem aos átomos das moléculas de água, 100 são átomos das moléculas de NaCl e 8231 é a quantidade de átomos da proteína.

3 RESULTADOS E DISCUSSÕES

3.1 *Propriedades Estruturais e Eletrônicas para os derivados do Cardanol*

Nesta seção apresentaremos os resultados dos cálculos das propriedades estruturais e eletrônicas, e a análise PCA para cada conjunto de moléculas estudadas. Como apresentado nos capítulos que antecedem, as moléculas foram divididas em grupos, cujo critério de seleção foi baseado nos substituintes mais parecidos; a análise então será realizada para cada grupo separadamente. Algumas moléculas já possuem testes de atividade, percentual inibitório e IC_{50} , apresentados nas tabelas 1, 2 e 3, e portanto as correlações obtidas pela PCA vislumbrarão os descritores relevantes para a formação de *cluster* (agrupamento) dentro de um mesmo conjunto de moléculas, afim de elaborar suposições sobre a relevância de determinado descritor para o resultado da atividade.

Os descritores calculados foram volume, polarizabilidade, energias dos orbitais de fronteira HOMO, HOMO-1, LUMO e LUMO+1, a carga sobre o átomo de nitrogênio protonado (qN), logD, pKa, PSA e o momento de dipolo(μ). Como o número de propriedades calculadas, ao todo 11, é considerável, a análise baseada nas combinações lineares desses descritores dá origem a uma extensa quantidade de componentes principais. A fim de garantir que a primeira componente principal tivesse a maior variância possível, tornando esses resultados mais seguros, elegemos apenas as combinações em que a variância da primeira componente principal (PC1) representasse mais de 70% da variância total, embora esse não seja o único parâmetro de análise, mas também a disposição das moléculas nas duas primeiras componentes principais, para a possível formação de *cluster*.

As análises que se seguem no decorrer do capítulo, para os diferentes grupos, foram selecionadas de um grande conjunto de possibilidades de PCs. Dentro de todas as possibilidades, as combinações lineares, para diferentes descritores demonstram um padrão quanto a disposição das moléculas nas duas primeiras componentes principais. Dentro desse conjunto de possibilidades que apresentam um padrão, escolhemos aquela na qual a PC1 apresentasse a maior variância, e portanto, dentro desses critérios de escolha, as análises eleitas são as que melhores denotam o padrão encontrado.

3.1.1 Resultados PCA para os aminoderivados heterocíclicos

Para o grupo formado pelas moléculas LDT140(1), LDT141(2), LDT142(3) e LDT144(4), na maioria das análises as estruturas moleculares 1 e 2 são aproximadas nas duas componentes. A estrutura 4 se aproxima da 2 somente na PC1 e encontra-se afastada das demais nas duas componentes. A estrutura 1 encontra-se afastada de 2, 3 e 4 nas PC1 e PC2. Essa distribuição pode ser visualizada na Figura 17 abaixo:

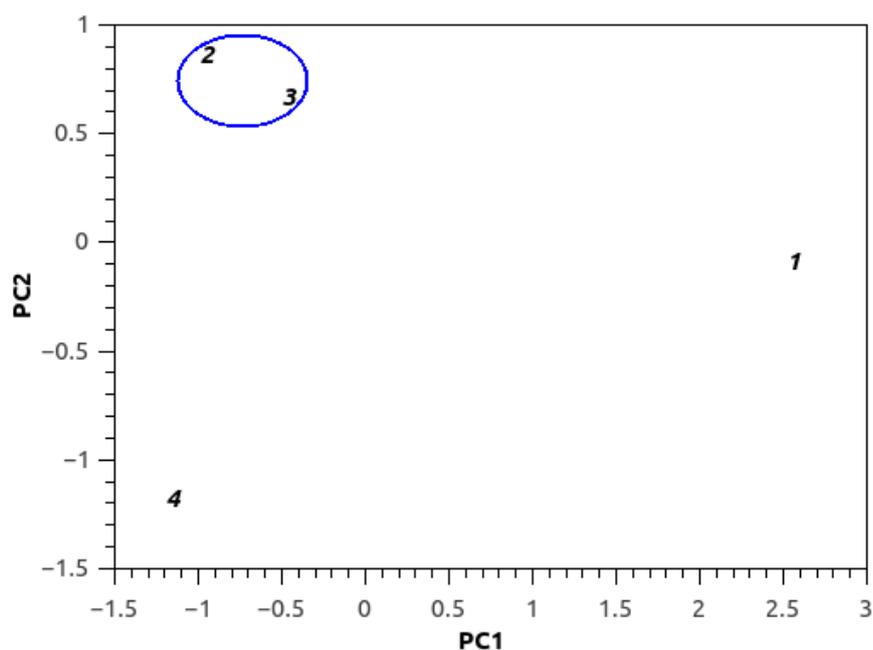


Figura 17 – Gráfico mostrando a disposição dos compostos LDT140(1), LDT141(2), LDT142(3) e LDT144(4) nas duas primeiras componentes principais-PC1 e PC2.

Os descritores reconhecidos como os mais adequados para este resultado foram Volume, HOMO, HOMO-1 e LUMO, dos quais as energias dos orbitais HOMO e HOMO-1 apresentam maior contribuição para a PC1 (ver tabela 4) responsável, aproximadamente, por 75% da informação da correlação (ver tabela 5). A PC2 contendo cerca de 21% da variância total, é dominada pelo descritor LUMO (ver tabela 4).

Tabela 4 – Peso das variáveis mais significativas para três primeiras componentes principais, para os aminoderivados heterocíclicos.

Variáveis	PC1	PC2
Volume	0.534001	-0.256962
HOMO	0.566629	-0.094022
HOMO-1	0.568048	-0.113003
LUMO	0.266637	0.955176

Tabela 5 – Variâncias acumuladas pelas três primeiras componentes principais, para os aminoderivados heterocíclicos.

Componentes	Autovalor	% Variância Total	Acumulativo
PC1	3.010431	75.26078	75.26078
PC2	0.855336	21.38339	96.64417

Para as estruturas 2 e 3, a PCA mostra correlação entre estas, o que está de acordo com o resultado experimental para o percentual inibitório destas duas estruturas, que corresponde a 41.8 e 42.1, respectivamente (ver tabela 1). A diferença básica na estrutura molecular das moléculas representadas pelos números 2 e 3, é a retirada do átomo de oxigênio na molécula 2, e a substituição do enxofre para molécula 3 (ver figura 11). O resultado experimental mostrou que a diferença estrutural entre essas duas moléculas, não refletiu mudança significativa no percentual de inibição da enzima acetilcolinesterase.

A PCA para este grupo possui o volume molecular como um dos descritores mais significativos para a primeira componente principal (ver tabela 4), comparando os valores para os volumes das respectivas moléculas (ver tabela 6), há uma diferença nos valores dessa propriedade, enquanto os valores para as energias dos orbitais de fronteira HOMO, HOMO-1 e LUMO, para as moléculas 2 e 3 são parecidas, o que refletiu na proximidade dessas moléculas na PC1 e na PC2.

A molécula 1, que possui a melhor atividade com percentual de inibição de 87.5 e IC_{50} igual a 26.4 micromolar, encontra-se afastada das demais moléculas nas duas componentes principais, que é um reflexo da diferença entre os valores dos descritores volume, HOMO e HOMO-1 dessa molécula para com as demais (ver tabela 6). A diferença estrutural na molécula 1 é o grupo metileno (CH_2), cuja substituição por oxigênio (molécula

2) ou enxofre (molécula 3) gerou o aumento da densidade de carga negativa e levou à existência de aceptores de ligação de hidrogênio, que possivelmente ocasionou a piora no reconhecimento molecular dessa região. A substituição do NH_2 (molécula 4) tornou a molécula um dicátion que por sua vez exibiu atividade inibitória inferior.

Tabela 6 – Valores das propriedades Volume (\AA^3), HOMO-1(eV), HOMO(eV) e LUMO(eV), utilizadas na PCA para os aminoderivados heterocíclicos.

Molécula	Volume	HOMO	HOMO-1	LUMO
LDT140	1129.15	-6.080162	-6.282345	-0.51566
LDT141	1066.39	-6.284522	-6.915287	-0.521919
LDT142	1093.21	-6.295678	-6.926716	-0.517021
LDT144	1085.4	-6.295406	-6.926687	-0.68546

A molécula 4 manteve-se distante das moléculas 1 e 3 nas duas componentes e se aproximou da molécula 2 na primeira componente principal (ver Figura 17), explicado também pelos valores HOMO e HOMO-1 serem bem parecidos, no entanto, o LUMO as distanciam na PC2. Como essa molécula ainda não possui dados experimentais, os resultados obtidos pelo estudo de PCA para essa molécula nos leva a supor que, por ela se manter distante da molécula 1, que possui a melhor atividade para esse grupo, os resultados para atividade podem ser inferiores, tornando desnecessária a investigação experimental dessa estrutura.

As moléculas pertencentes a esse grupo, possui seu orbital HOMO concentrado no grupo aromático, de acordo com a Figura 18, compreendendo a região com maior carácter elétron doador. Esse grupo funcional pode fazer interações do tipo dipolo-dipolo, aromático-hidrofóbico, íon-dipolo, interagindo possivelmente com aminoácidos presentes no sítio periférico.

A contribuição do orbital LUMO, mostrado na Figura 18, representa a região com maior carácter eletro-receptor e encontra-se concentrada principalmente na subunidade farmacofórica para todas as moléculas, indicando uma possível interação dessa região com o sítio catalítico da AChE. Essa região é a parte onde observa-se mais diferenças nas superfícies desse orbital em relação as diferenças para o orbital HOMO, já que as substituições moleculares foram realizadas nesse ponto.

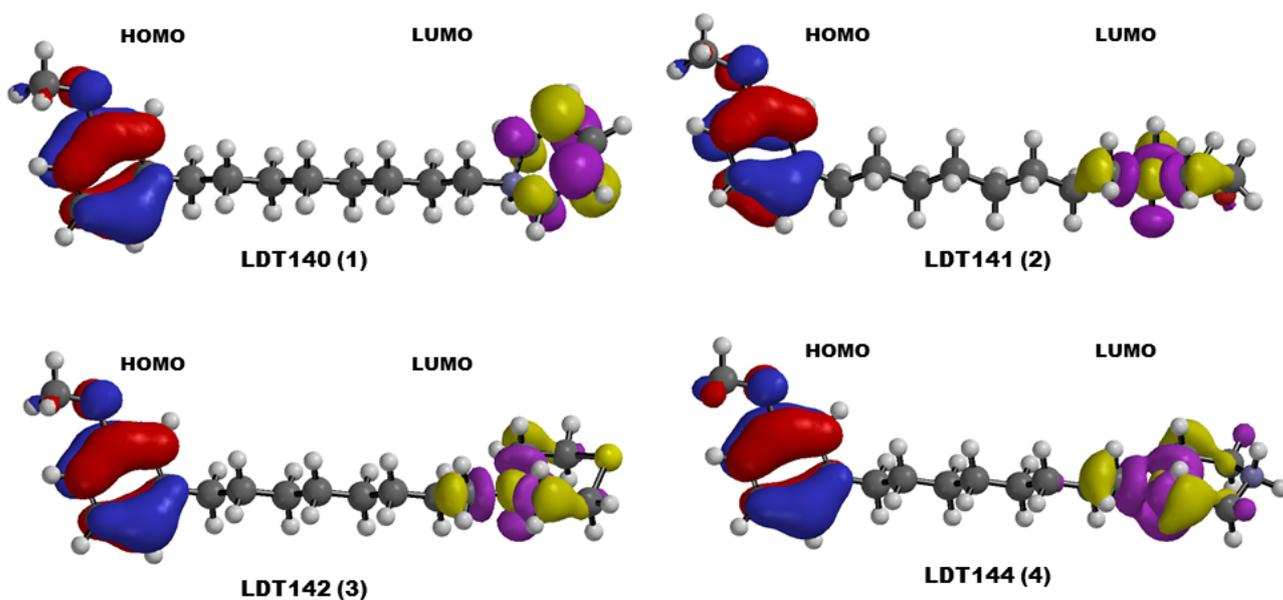


Figura 18 – Superfície de densidade eletrônica dos orbitais HOMO e LUMO para os aminoderivados do cardanol.

Na maior parte das moléculas a contribuição do orbital HOMO-1 (Figura 19), encontra-se no anel aromático do cardanol e na ligação entre os dois primeiros carbonos ligados ao anel, expressando que essa região pode ser um plausível sítio doador de elétrons.

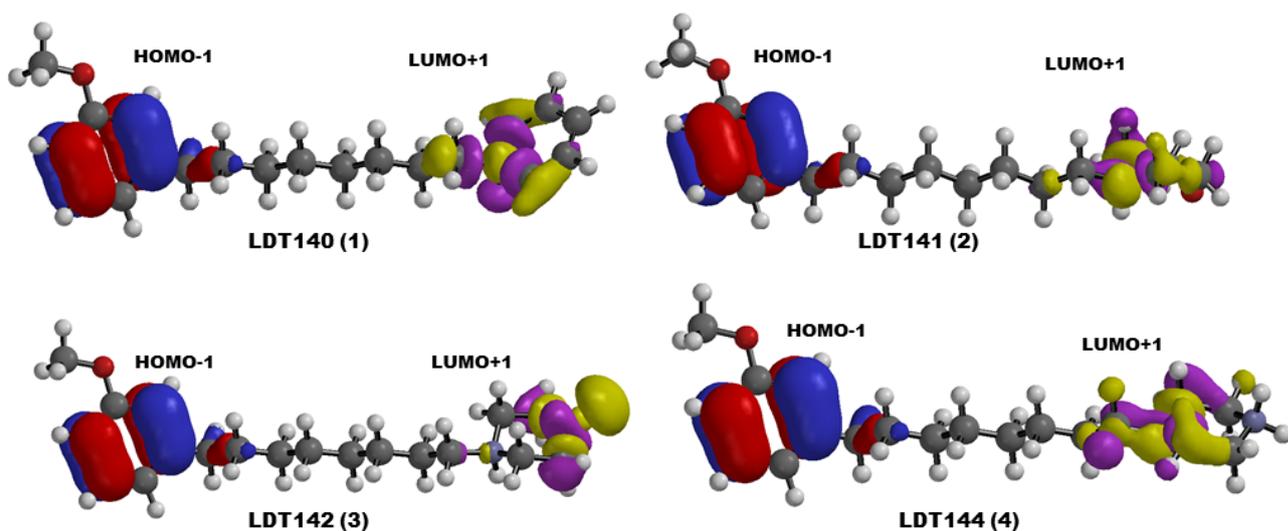


Figura 19 – Superfície de densidade eletrônica dos orbitais HOMO-1 e LUMO+1 para os aminoderivados do cardanol.

O mapa de potencial eletrostático (do inglês, *molecular electrostatic potential*-MEP) é uma maneira de visualizar a distribuição de carga de uma molécula. A densidade de elétrons é obtida nos cálculos computacionais quânticos de otimização e energia, e depois é gerada a superfície de potencial.

O potencial eletrostático para os grupos de derivados do cardanol foram calculados no Gaussian (FRISCH et al., 2009) e gerados no GaussView (DENNINGTON; KEITH; MILLAM, 2009), onde o potencial eletrostático em um ponto qualquer (x,y,z) é dado pela energia potencial eletrostática entre um íon imaginário carregado positivamente, situado nesse ponto, e a molécula. Se o íon é repelido pela molécula então naquele ponto o potencial é positivo, e se for atraído esse potencial é negativo. Considerando que o íon possui carga +1, ele é repelido por regiões pobres em elétrons e atraído em regiões ricas em elétrons; portanto, regiões pobres em elétrons tem potenciais positivos, e potenciais ricos em elétrons tem potenciais negativos.

Nos MEPs gerados para os derivados do cardanol, regiões onde o potencial eletrostático é positivo foram coloridas de azul, e onde o potencial eletrostático é negativo foram coloridas de vermelho. Regiões de potencial intermediário, seguem a ordem de cores do espectro: vermelho < laranja < amarelo < verde < azul.

A Figura 20 mostra o mapa de potencial eletrostático para os aminoderivados do cardanol, no qual observamos a concentração de densidade negativa sobre o grupo auxofórico aromático, mais especificamente sobre o elemento oxigênio, mostrando que essa região da molécula possui potencial atrativo. E a densidade de carga positiva localizada na subunidade farmacofórica, especialmente sobre o nitrogênio, exibindo o caráter repulsivo (em relação ao íon com carga +1).

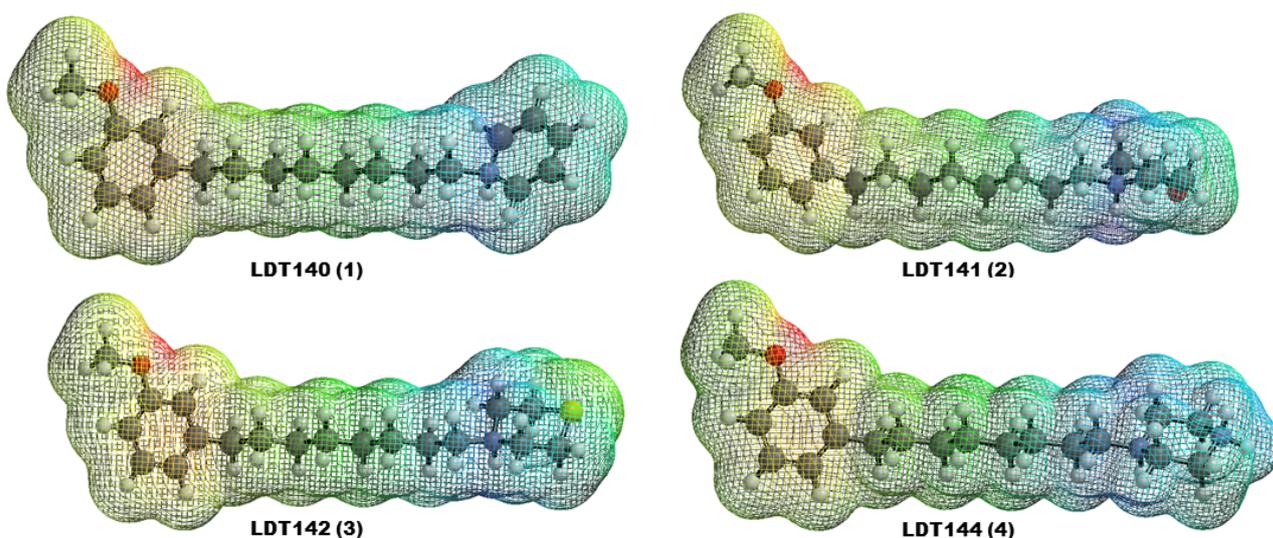


Figura 20 – Mapa de potencial eletrostático (do inglês , *molecular electrostatic potential*-MEP) para os aminoderivados do cardanol.

A obtenção de MEPs contribui no sentido de guiar a avaliação das interações moleculares dos candidatos estudados, em relação a interações destes com aminoácidos no sítio ativo da AChE.

3.1.2 Resultados PCA para os derivados piperazínicos

No segundo grupo de moléculas estudadas, LDT7(1), LDT143(2), LDT145(3), LDT146(4), LDT147(5), LDT166(6), LDT168(7), LDT170(8), LDT242(9), LDT473(10) e LDT474(11), somente as moléculas representadas pelos números 1, 2, 3, 5, 7 e 9, apresentam percentual de inibição da acetilcolinesterase, e destas somente as moléculas 2 e 9, possuem dados de IC_{50} (ver tabela 2). Das informações experimentais conhecidas para estas moléculas, em geral não apresentam resultados satisfatórios como candidatos a agentes AChEI, no entanto, conhecer informações estruturais e eletrônicas do ponto de vista da mecânica quântica podem ser de grande valia para o planejamento de novos candidatos que de alguma maneira se distanciem das obtidas para esse grupo específico de moléculas derivadas do cardanol.

O estudo baseado nas componentes principais não revelou uma tendência para esse grupo, já que não houve formação de agrupamentos de moléculas com atividades semelhantes e não encontramos padrão nas análises das combinações lineares entre os

descritores físico-químicos. Estruturalmente esses compostos são bem diferentes, figura 12, essa diferença estrutural pode ter levado a não correlação entre estes compostos. Podemos ainda supor que algumas amostras (moléculas), por apresentarem muitas diferenças, sejam estruturais ou eletrônicas, das demais moléculas do grupo, influenciaram o resultado de maneira negativa. Devido a insuficiência dos resultados de PCA, não foi possível identificar com clareza quais descritores possuem caráter preditivo para esse grupo, mesmo que a identificação destes partisse do princípio de que estas moléculas não compreendem potentes candidatos a AChEI.

3.1.3 Resultados de PCA para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos

O grupo de derivados etanolamínicos N-benzila é formado pelos compostos LDT155(1), LDT156(2), LDT157(3), LDT158(4), LDT475(5) e LDT489(6) e até o momento, nenhum destes tem resultados experimentais de atividade anticolinesterásica.

Na Figura 21, para as primeiras componentes principais, é possível visualizar a formação de *clusters* com as moléculas 1 e 2, e com as moléculas 3 e 4. As moléculas 1 e 2 possuem diferença estrutural molecular no grupo amina, no entanto, os resultados obtidos mostram que esta diferença não interfere na correlação entre estas moléculas, já que a PCA encontra padrões para estes agrupamentos, e os descritores selecionados (LUMO, LUMO+1, pKa e logD) apresentam-se com maior frequência nestes padrões.

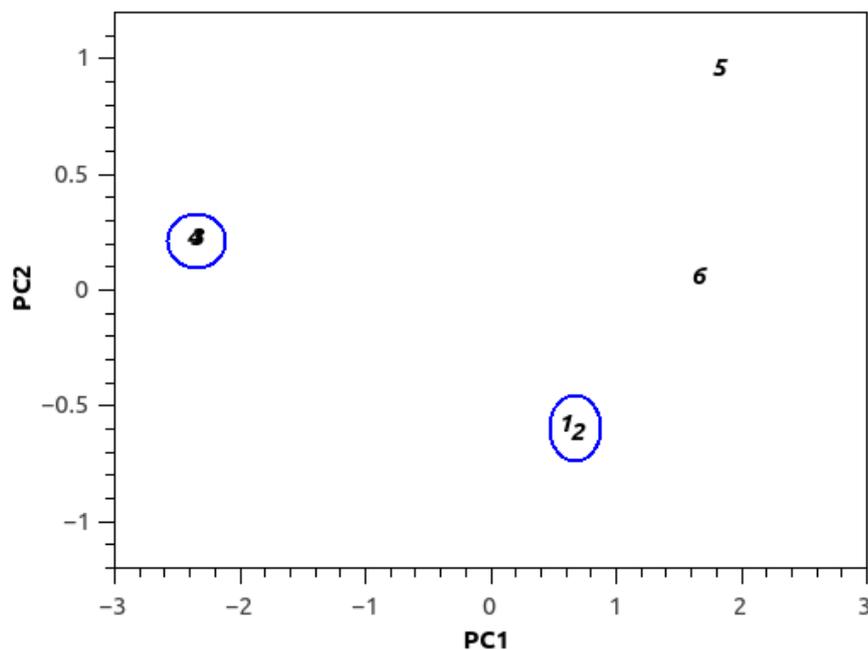


Figura 21 – Gráfico mostrando a disposição dos compostos LDT155(1), LDT156(2), LDT157(3), LDT158(4), LDT475(5) e LDT489(6) nas duas primeiras componentes principais-PC1 e PC2.

Na segunda componente principal o descritor de maior relevância foi o pKa, como pode ser visto na Tabela 7, que também apresenta-se como um importante descritor para a PC1. O pKa identifica o grau de ionização, neste caso para todas as moléculas o pKa assume valores entre 8.8 e 10.4 (ver Tabela 9), mostrando que essas moléculas podem ser vistas como bases fracas, e sofrem ionização significativa em água.

Tabela 7 – Peso das variáveis mais significativas nas três primeiras componentes principais, para os derivados etanolamínicos N-benzila ou N-alkilassubstituídos.

Variáveis	PC1	PC2	PC3
LUMO	0.501491	-0.504602	0.327901
LUMO+1	0.522013	-0.157564	0.377726
pKa	0.456426	0.843411	0.170141
logD	-0.517382	0.095965	0.849033

As variâncias acumuladas pela três primeiras componentes principais mostrando que para esse grupo a PC1 possui 90.24% de informação a respeito da correlação, enquanto

a PC2 apenas 8.59% (ver Tabela 8), e os descritores mais significativos para a PC1 foram LUMO e LUMO+1, com LUMO+1 contribuindo um pouco mais que o LUMO, o que pode ser verificado na Tabela 7.

Tabela 8 – Variâncias acumuladas pelas três primeiras componentes principais, para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos.

Componentes	Autovalor	% Variância Total	Acumulativo
PC1	3.609724	90.24309	90.24309
PC2	0.343601	8.59002	98.83311
PC3	0.042811	1.07027	99.90337

O descritor LUMO+1 domina a PC1, analisando-o para o agrupamento formado pelas moléculas 1 e 2, este descritor mostra-se responsável pela agrupamento não só das moléculas 1 e 2 como também das moléculas 3 e 4, de maneira que estes valores, mostrados na Tabela 9, discriminam muito os dois *clusters* observados no gráfico de PCs. Essa similaridade, entres os valores obtidos para as propriedades e os agrupamentos formados, é observado também para o descritor LUMO.

A molécula 6 possui valor de LUMO equivalente à molécula 1, aproximadamente -0.52 eV, contudo os valores do descritor LUMO+1 para estas moléculas são bem diferentes, o que as distanciou, principalmente, na PC1.

O pKa uniu as moléculas 1 e 2, e, 3 e 4 na PC2, mas certamente foi o responsável por separar as moléculas 5 e 6, pois seus valores para estas são distintos. As moléculas 5 e 6 estão correlacionadas na primeira componente principal pelo descritor LUMO+1, que para essas moléculas são, -0.26 eV e -0.29 eV, respectivamente. Possivelmente, 5 e 6 reproduzam a influência do logD, que governa a terceira componente principal, apesar de representar pouca relevância para as demais componentes, e possui valores próximos para estas duas moléculas (ver tabela 9).

Tabela 9 – Valores das propriedades LUMO (eV), LUMO+1 (eV), pKa e logD, utilizadas na PCA dos derivados etanolamínicos N-benzila ou N-alquilassubstituídos.

Molécula	LUMO	LUMO+1	pKa	logD
LDT155	-0.519742	-0.42559	9.2	3.02
LDT156	-0.517293	-0.343683	9.2	3.14
LDT157	-1.206291	-0.925462	8.77	4.8
LDT158	-1.19976	-0.922201	8.77	4.92
LDT475	-0.57743	-0.262864	10.33	2.6
LDT489	-0.518381	-0.296879	9.76	2.05

Os resultados obtidos com a análise de componentes principais leva-nos a levantar a hipótese de que a atividade anticolinesterásica das moléculas que formaram *clusters* sejam semelhantes, e ainda as modificações no grupo farmacofórico, entre as moléculas 1 e 2, e 3 e 4, podem não fazer diferença para a atividade. Já as alterações estruturais da molécula 5 em relação a 6, diferenciaram estas moléculas com relação ao caráter acceptor (orbitais de fronteira LUMO e LUMO+1) da região farmacofórica.

As superfícies de densidade eletrônica para os orbitais LUMO e LUMO+1 dos derivados etanolamínicos N-benzila ou N-alquilassubstituídos, Figura 22 mostra a igualdade dessas nuvens para as moléculas agrupadas na PCA, e uma singela diferença entre as moléculas 5 e 6. A visualização dessas superfícies, bem como os resultados de PCA, evidenciam que modificações, ainda que pequenas (a diferença entre as moléculas 5 e 6 é um grupo carboxila), levaram ao deslocamento da densidade do orbital LUMO, podendo alterar significamente a atividade biológica dessas moléculas.

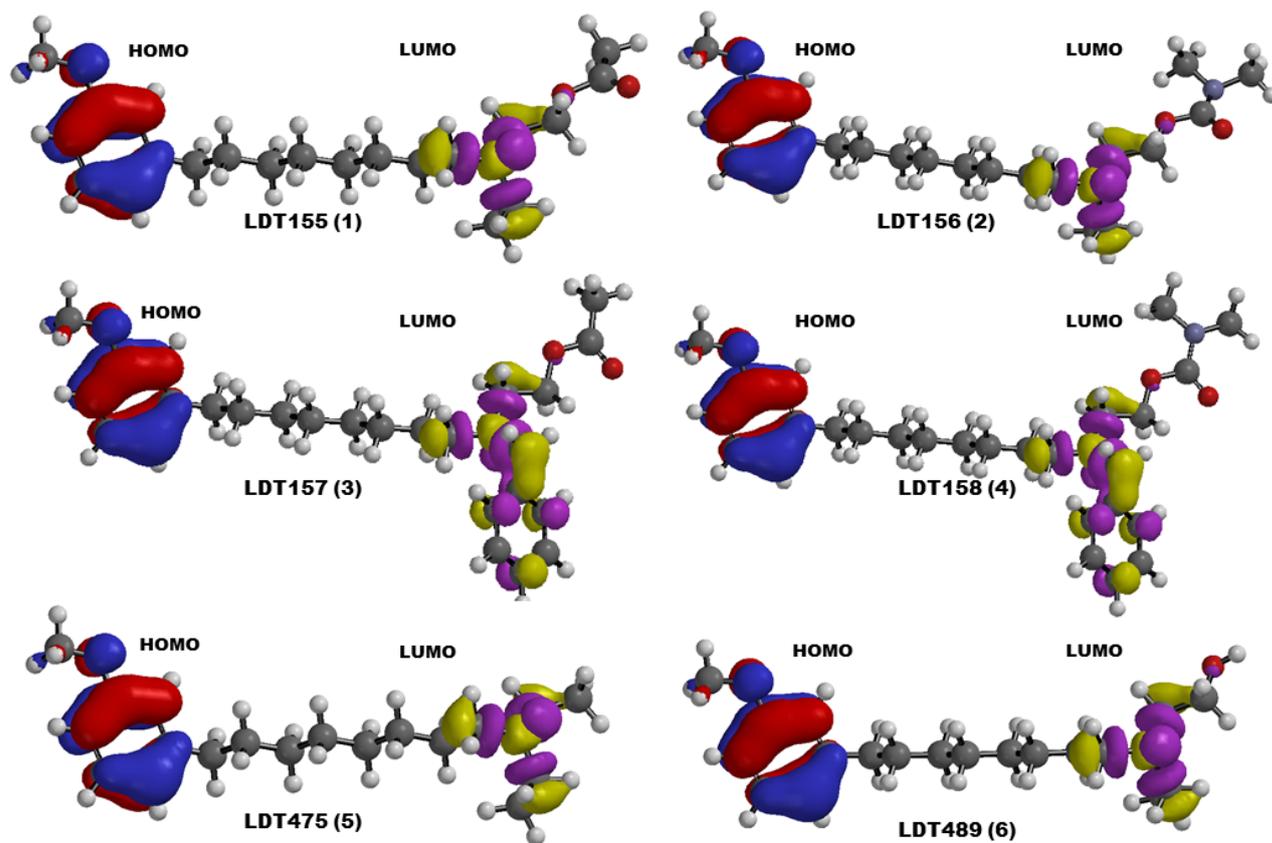


Figura 22 – Superfície de densidade eletrônica dos orbitais HOMO e LUMO para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos.

As densidades para o orbital HOMO-1 é igual para todas as moléculas, mostrando que esse orbital não reflete as modificação realizadas na região farmacofórica, e podem ser vistas na Figura 23. Já o orbital LUMO+1 é diferente para todas as moléculas, e na Figura 23 é possível verificar que há uma concentração de densidade na região da amina, para as moléculas 1 e 2, mas a adição de um anel aromático modifica a posição dessa densidade. Se fizermos uma comparação entre as moléculas 2 e 4, cuja diferença é justamente o anel aromático, vemos que a PCA separa essas moléculas nas duas componentes principais, e a PC1, com maior peso, que é definida por esses orbitais.

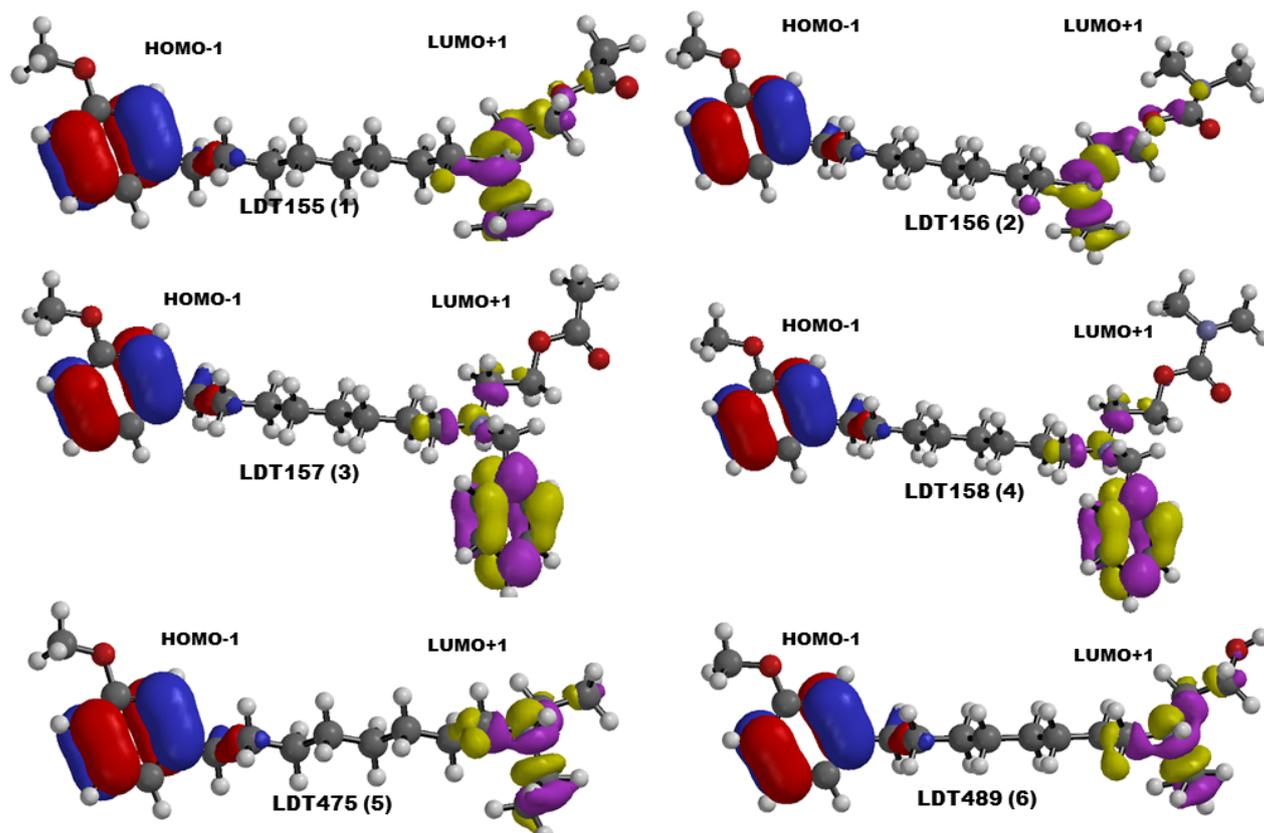


Figura 23 – Superfície de densidade eletrônica dos orbitais HOMO-1 e LUMO+1 para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos.

Os MEPs para os derivados etanolamínicos N-benzila ou N-alquilassubstituídos, Figura 24, mostra que a diferença estrutural entre as moléculas 1 e 2 não produz diferença nas densidades eletrostáticas dessas moléculas, acontecendo o mesmo para as moléculas 3 e 4. E para as moléculas 5 e 6, há uma sutil diferença no superfície de potencial positivo.

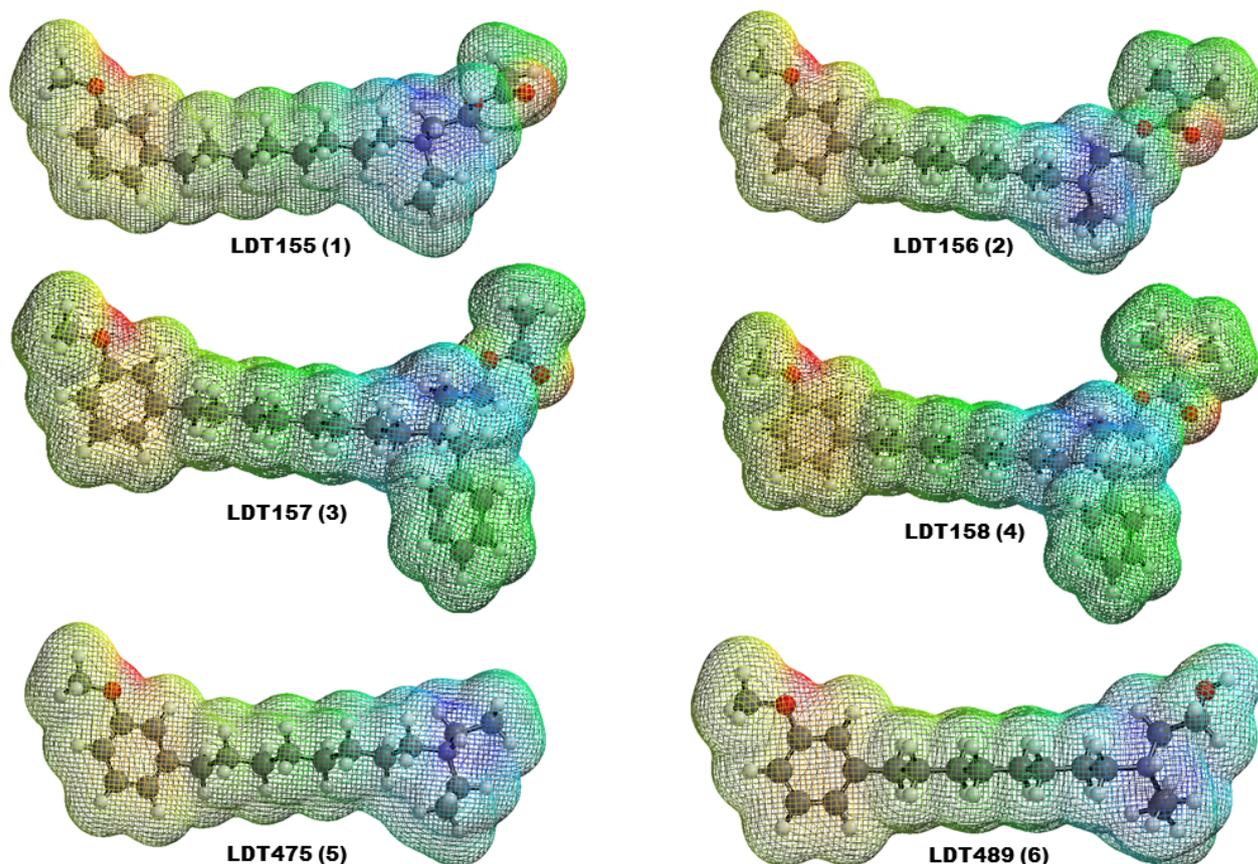


Figura 24 – Mapa de potencial eletrostático (do inglês , *molecular electrostatic potential*-MEP), para os derivados etanolamínicos N-benzila ou N-alkilassubstituídos.

3.1.4 Resultados de PCA para os derivados benzilamínicos

O grupo com derivados cuja parte farmacofórica possui substituintes benzilamínicos, formado pelas moléculas LDT159(1), LDT160(2), LDT161(3), LDT167(4) e LDT169(5), foi o que apresentou melhores resultados experimentais quanto ao percentual de inibição e IC_{50} (ver tabela 3), com exceção da molécula 5, com % de inibição de $24.9 \mu M$ considerado insatisfatório para um potente candidato à fármaco, e a molécula 1 que não possui resultados experimentais.

Na figura 25, podemos verificar a formação de *cluster* entre as moléculas 1 e 2, e entre as moléculas 3 e 4, e a molécula 5 encontra-se afastada dos *clusters* formados.

As moléculas de 1 a 5 tem em comum o orbital de fronteira HOMO, mas se diferenciam nos demais descritores, inclusive no HOMO-1 que juntamente com o HOMO são os que mais contribuem para a PC2, responsável apenas por 5.25% da correlação.

A PC1, responsável por 90% da correlação, tem os orbitais de fronteira LUMO e LUMO+1, como os descritores que mais expressivos (ver tabela 10).

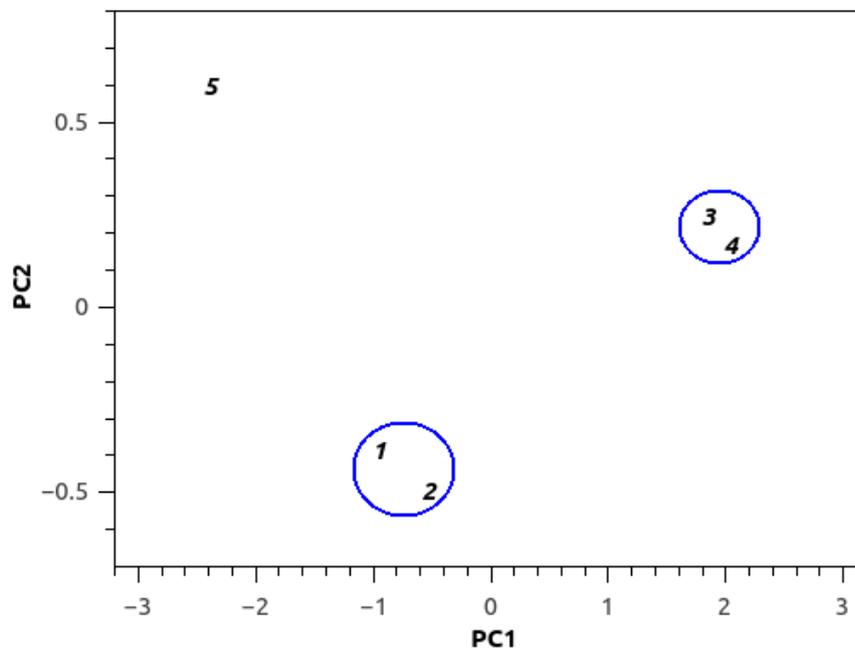


Figura 25 – Gráfico mostrando a disposição dos compostos LDT159(1), LDT160(2), LDT161(3), LDT167(4) e LDT169(5) nas duas primeiras componentes principais-PC1 e PC2.

As moléculas 1 e 2 contam estruturalmente com uma distinção com relação ao grupo etila, como pode ser visto na figura 14. No que diz respeito aos descritores utilizados para a análise de componentes principais, essas moléculas são correspondentes, uma vez que para todos os descritores os valores são muito próximos (ver tabela 10), diferenciando-se minimamente no LUMO+1, o que possivelmente é responsável pela separação entre elas na PC1.

O descritor LUMO+1, é mais significativo do que o LUMO para a primeira componente principal. A molécula 1, como foi dito anteriormente, não possui dado experimental, porém a molécula 2 apresenta inibição da enzima acetilcolinesterase de 81.5 % e IC_{50} igual a 16.1 μM , considerados valores muito bons. A proximidade dessas duas moléculas no estudo de PCs pode indicar uma semelhança nos resultados de atividade anticolinesterásica da molécula 1 com a molécula 2.

Tabela 10 – Peso das variáveis mais significativas para as três primeiras componentes principais para os derivados benzilamínicos.

Variáveis	PC1	PC2	PC2
HOMO	0.489188	-0.170044	0.855097
HOMO-1	0.487821	0.813949	-0.108809
LUMO	0.501726	-0.550794	-0.381022
LUMO+1	0.520574	-0.072094	-0.334352

As moléculas 3 e 4 são as que apresentam os melhores valores experimentais para o percentual de inibição e IC_{50} (ver tabela 3). A diferença entre 3 e 4 é basicamente o grupo etila, a molécula 4 possui % inibitório de 95.5 e IC_{50} 17.2 μM , que são resultados bons.

A substituição de um hidrogênio próximo ao grupo farmacofórico por um etil na molécula 3, se comparada com a molécula 4, melhorou a atividade biológica, uma vez que o IC_{50} foi de 17.2 μM para 6.6 μM , mostrando a relevância desse grupo para a atividade.

Os valores dos descritores obtidos para as moléculas 3 e 4, mostram resultados muito próximos, com uma diferença pequena no LUMO+1, resultando numa proximidade entre essas duas moléculas nas duas primeiras componentes principais. A molécula 5, que possui resultado experimental muito inferior ao das moléculas 3 e 4, apresenta valores diferentes aos dessas moléculas nos descritores HOMO-1, LUMO e principalmente no LUMO+1, o que a distanciou dessas moléculas no gráfico de PCA.

A primeira componente principal contribui com 90.31% de informação como pode ser visto na Tabela 11, é a componente responsável pela diferenciação dos moléculas. Nessa componente, as variáveis com maiores pesos foram LUMO e LUMO+1, que contribuem igualmente na PC1, de acordo com a Tabela 10. A segunda componente principal contribui somente com 5.25% da informação e é dominada pela variável HOMO-1.

Tabela 11 – Variâncias acumuladas pelas três primeiras componentes principais, para os derivados benzilamínicos.

Componentes	Autovalor	% Variância Total	Acumulativo
PC1	3.612464	90.31161	90.31161
PC2	0.210344	5.25860	95.57022
PC3	0.177163	4.42908	99.99930

Os valores para as variáveis mais significativas, selecionadas como aquelas que melhor representam o padrão encontrado, são mostradas na Tabela 12. Os orbitais moleculares HOMO-1, HOMO, LUMO e LUMO+1, mostraram-se presentes na maior parte dos conjuntos de variáveis que apresentavam esse padrão, e juntos conseguiram discriminar moléculas com bons resultados experimentais das demais moléculas, que possuem atividades razoáveis.

Tabela 12 – Valores das propriedades HOMO-1(eV), HOMO(eV), LUMO(eV) e LUMO+1(eV), utilizadas na PCA para os derivados benzilamínicos.

Molécula	HOMO	HOMO-1	LUMO	LUMO+1
LDT159	-6.283705	-6.918552	-1.134452	-0.8822
LDT160	-6.282889	-6.916919	-1.142343	-0.904242
LDT161	-6.282617	-6.729976	-1.04411	-0.580152
LDT167	-6.282073	-6.735146	-1.056355	-0.607091
LDT169	-6.283977	-6.914742	-1.268605	-1.11105

Na Figura 26, estão representados as superfícies de densidades dos orbitais HOMO e LUMO dos derivados benzilamínicos. Nessa figura, o orbital HOMO não se diferencia para todos os derivados, permanecendo na região auxofórica. As superfícies de densidade do orbital LUMO, por sua vez, sofrem variações devido as modificações estruturais ocorridas na parte farmacofórica, onde os mesmos se localizam.

Comparando as moléculas 1 e 2 (a molécula 1 tem o H_2 ligada ao nitrogênio, enquanto, a molécula 2 tem nessa mesma posição o grupo C_2H_5), o orbital LUMO possui maior concentração próximo ao nitrogênio e ao H_2 da molécula 1. Quando esse grupo é substituído por etila, molécula 2, a superfície de densidade LUMO é mais distribuída, provavelmente aumentando a superfície com caráter receptor, o que pode significar alterações importantes na maneira com que essas moléculas interagem no sítio biológico.

O mesmo acontece entre as moléculas 3 e 4, como a alteração estrutural entre 3 e 4 é a mesma entre 1 e 2, mais uma vez podemos verificar as diferenças ocorridas nas superfícies de densidade LUMO quanto a substituição de H_2 por C_2H_5 . A densidade LUMO é mais distribuída na molécula LDT161(3), e como essas moléculas apresentam IC_{50} diferentes (LDT161($IC_{50} = 6.6\mu M$), LDT167($IC_{50} = 17.2\mu M$)), as diferenças nesses orbitais podem ter contribuído para a melhora da atividade.

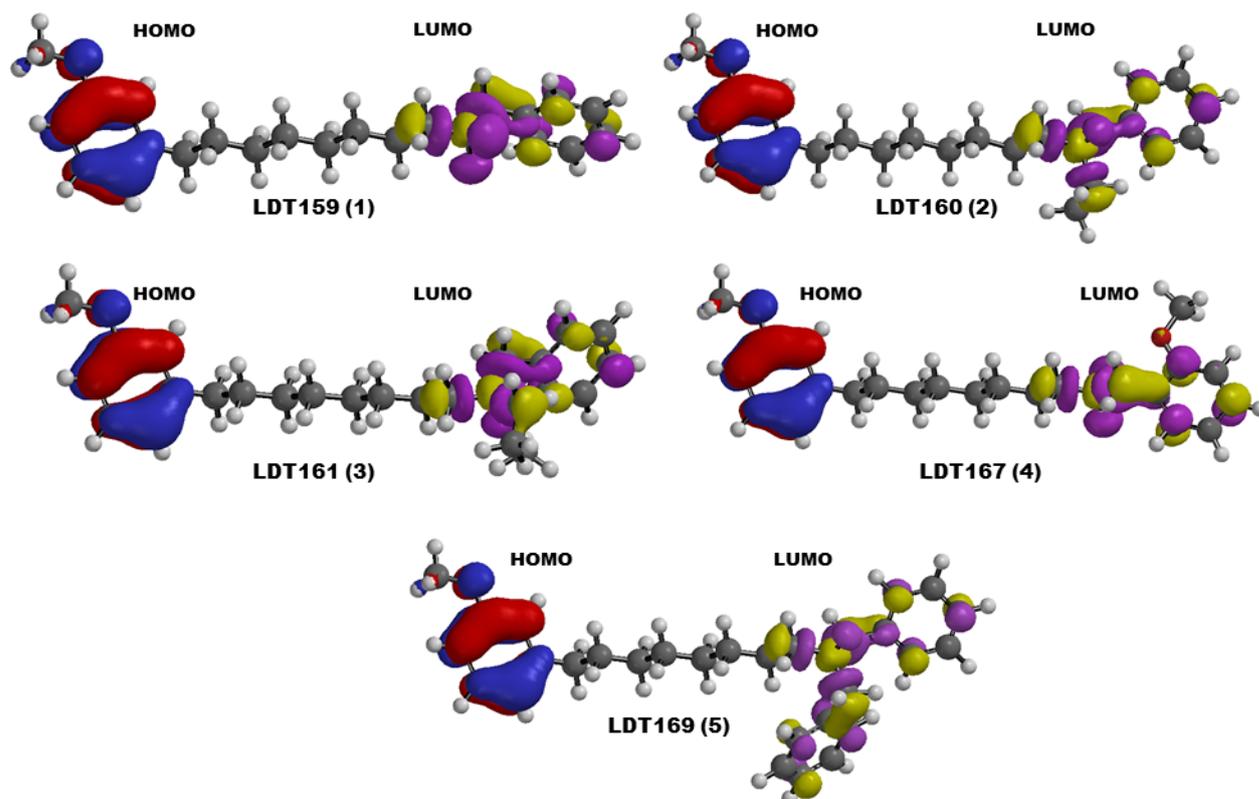


Figura 26 – Superfície de densidade eletrônica dos orbitais HOMO e LUMO, para os derivados benzilamínicos do cardanol.

Para superfícies de densidade para orbitais moleculares HOMO-1 e LUMO+1, apresentadas na Figura 27, as observações realizadas são as mesmas feitas para os orbitais HOMO e LUMO. Lembrando que os orbitais LUMO e LUMO+1, são aqueles que mais contribuíram na PC1 e diferenciaram as moléculas com relação as superfícies de densidade.

Como esses orbitais possuem caráter eletro-receptor, o fato de estarem localizados no final da subunidade farmacofórica de todos os grupos de moléculas analisados (amino, etanolamínicos e benzilamínicos), e mostraram-se como variáveis significantes nos resultados para as componentes principais, pode indicar que ao encaixar-se no sítio de interesse, supondo que as moléculas alcancem o fundo da Gorge, essa porção interaja com aminoácidos da tríade catalítica. E modificações, que produzam valores menores para IC_{50} s, supostamente devem ser realizadas em regiões de concentração destes orbitais.

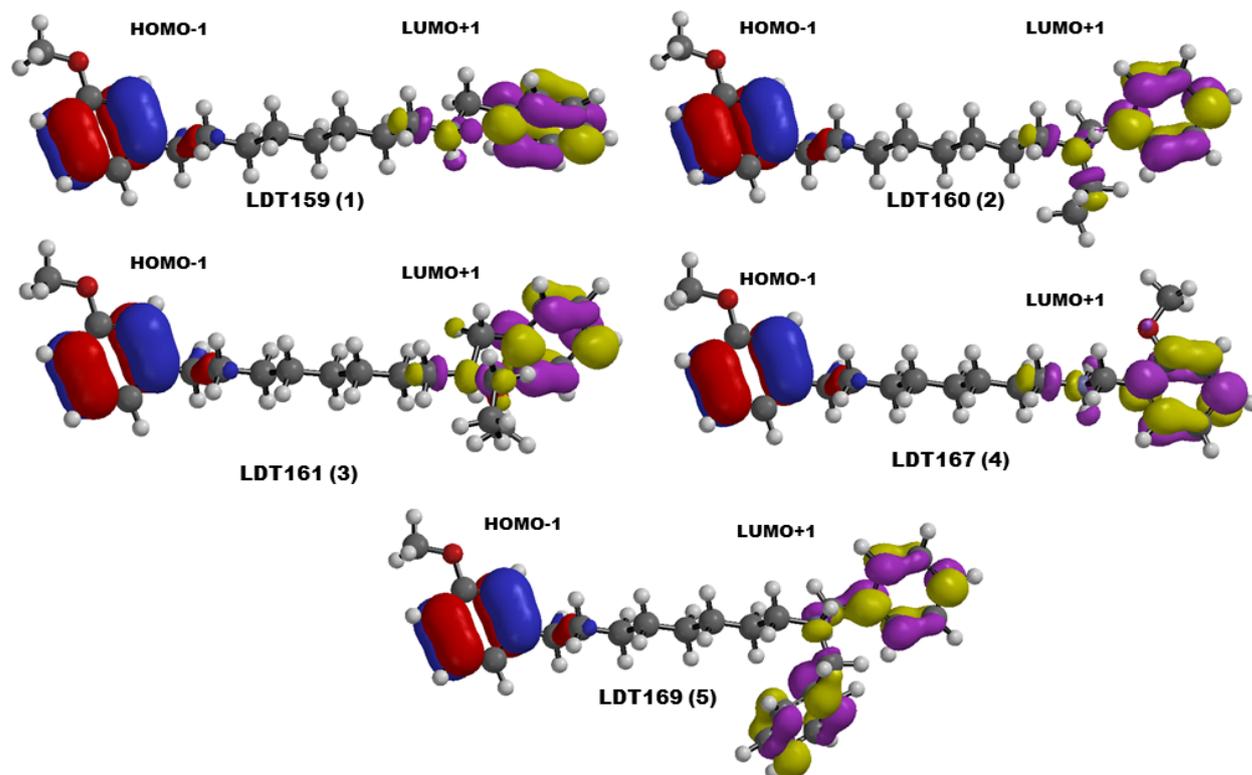


Figura 27 – Superfície de densidade eletrônica dos orbitais HOMO-1 e LUMO+1, para os derivados benzilamínicos do cardanol.

Os MEPs para este grupo, representado na Figura 28, exibem, assim como nos outros grupos, que as alterações realizadas na estrutura das moléculas interferem na distribuição de das mesmas.

A região de potencial eletrostático intermediário, próximo a região de potencial atrativo (com relação a carga pontual positiva, localizada no oxigênio da parte auxofórica), da molécula LDT161(3) em especial, apresenta densidade de potencial mais positiva do que a mesma região para as demais moléculas. As modificações realizadas nessa molécula, levaram ao aumento da região de potencial positivo (repulsiva), isso pode ter relação com o fato dessa molécula apresentar os melhores resultados experimentais (%inibitório=91.6 e $IC_{50} = 6.6\mu M$).

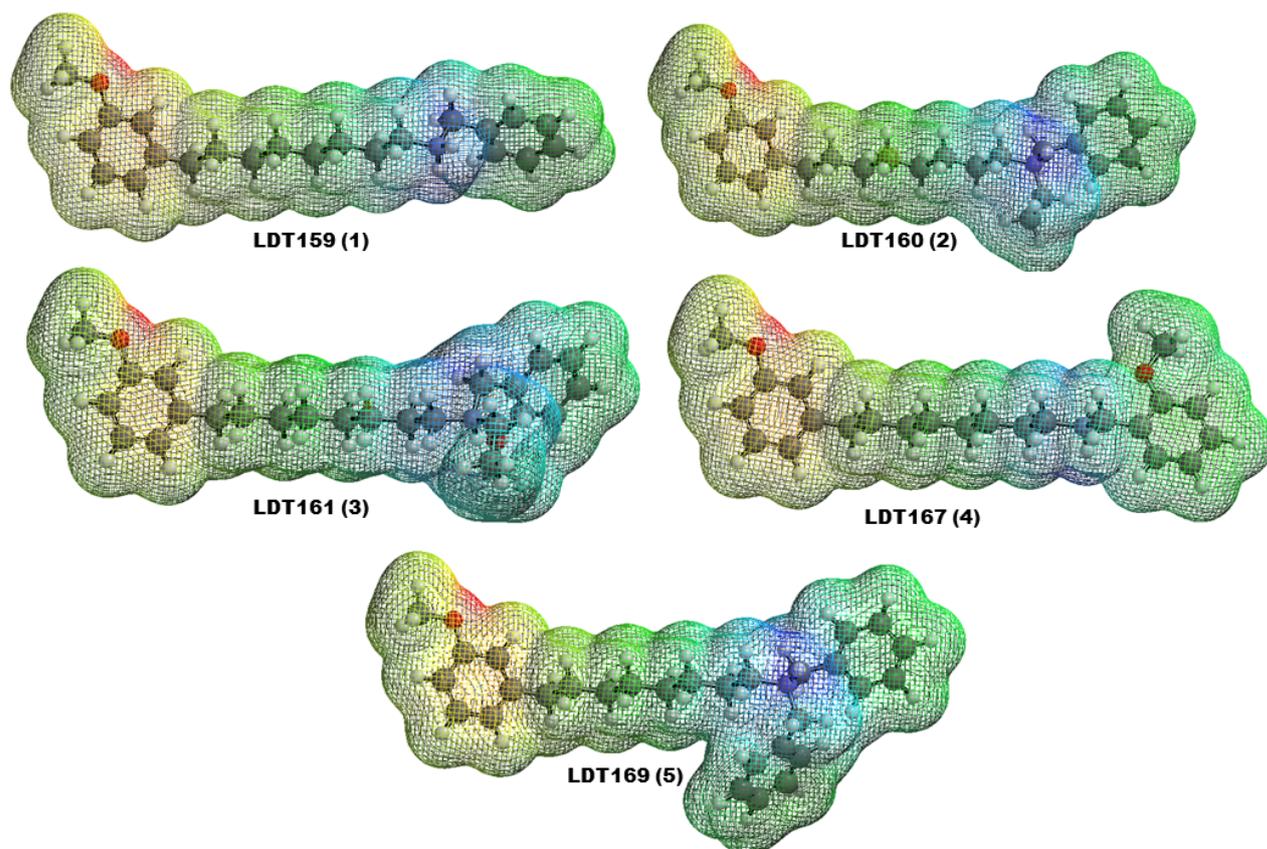


Figura 28 – Mapa de potencial eletrostático (do inglês , *molecular electrostatic potential*-MEP), para os derivados benzilamínicos do cardanol.

Apesar de todas as moléculas apresentarem uma protonação na amina, a inserção de grupos na região farmacofórica mudaram a distribuição de potencial positivo na amina e a disposição deste, na molécula como um todo. O que podemos pressupor, baseado nos MEPs, é que moléculas que apresentam mais regiões de potencial positivo, podem ter mais afinidade de interação com o sítio biológico.

As análises das Componentes Principais nos leva a previsões e interpretações de descritores significativos que estabeleçam padrão de similaridade entre moléculas, e ainda, descritores que mostrem-se importantes para o entendimento das interações que a droga realiza nas regiões ativas do receptor, podendo também auxiliar no desenho de novas drogas.

Respalda no conhecimento estrutural e eletrônico para o conjunto de moléculas, e na correlação entre suas atividades e os descritores inerentes a cada uma apresentadas pelo PCA, selecionamos o grupo dos benzilamínicos (que apresenta os melhores resultados experimentais) para realizar o estudo de Docking, no intuito de compreendermos orientações

moleculares preferenciais desses ligantes com o sítio ativo da AChE, e estabelecer uma relação com os descritores eleitos na PCA.

O processo de *Drug Design* envolve a concepção de moléculas que se complementam com o alvo biológico de forma que realiza interações com o mesmo. O conhecimento das moléculas candidatas a novas drogas é de fundamental interesse para o desenho racional de fármacos, no entanto, isso não é suficiente. Encontrar novos medicamentos exige o conhecimento detalhado do alvo biológico, que no nosso caso é a AChE humana (código PDB:4EY7), e só então é possível prever a forma pela qual a proteína (receptor) reconhece o fármaco (ligante). Por isso a próxima seção é dedicada ao estudo da biomolécula receptora.

3.2 *Resultados da Dinâmica Molecular*

Para várias estruturas cristalográficas da AChE de múltiplos organismos, diversos grupos tem realizado estudos de dinâmica molecular que revelam a complexa natureza das flutuações do sítio ativo gorge da AChE, devido principalmente ao estreitamento entre os resíduos aromáticos Tyr155 e Phe369 da gorge na região do gargalo. Esses estudos apontam um estado “*open*” desse estreitamento, e essa população de conformações “*open*” é acessada somente durante cerca de 2% da simulação (BUI; HENCHMAN; MCCAMMON, 2003). Esses resultados de DM, em sua maioria para a água e para o substrato natural ACh, revelam ainda a entrada e ou saída de ambos ao sítio catalítico, por outros acessos, “*back door*” e “*side door*” (SANSON et al., 2011). Grande parte desses estudos apresentam tempo de simulação de DM em torno de 10 ns (TAI et al., 2002; FANG et al., 2014).

O estudo realizado por Brian Bennion (BENNION et al., 2015), no qual o tempo simulação de DM para a AChE foi de 300 ns, foi um tanto esclarecedor, no que se refere as flutuações no volume da gorge, foi observado o movimento de respiração no comportamento do volume. O tempo demandado para essa observação foi de 100 ns.

Afim de compreender a interação entre as moléculas pertencentes ao grupo dos benzilamínicos e a hAChE escolhida (PDB: 4EY7), sabendo que o modo como estas moléculas interagem com os resíduos da Gorge, inibindo a AChE não são conhecidos, e investigar o que leva moléculas estruturalmente próximas apresentarem potenciais de inibição diferentes, é de grande relevância para descobrir as bases moleculares que levaram a esses resultados. Desta forma, os mecanismos de interação deste grupo de inibidores foram

desenvolvidos pelas abordagens de simulação de dinâmica molecular, para o receptor, e acoplamento molecular receptor-ligante. Nessa pesquisa, portanto, simulações de DM foram aplicadas ao estudo de mecanismo da AChE e os resultados gerados serão apresentados a seguir.

3.2.1 Equilibração do Sistema

O tempo de simulação total da AChE para as etapas de termalização e relaxação e amostragem foi de 206 ns, de acordo com as configurações iniciais descritas na seção de metodologia de DM. O tempo de simulação para estágio de termalização e relaxação do sistema foi de 6 ns, e os demais 200 ns consiste na amostragem de dinâmica do sistema. Cada arquivo trajetória corresponde a 1 ns composto por 500 *frames*, totalizando 100000 *frames* em toda a simulação.

O conhecimento das flutuações conformacionais das proteínas são importantes para compreender as funções biológicas desenvolvidas pelos sistemas já que a relação entre movimento e função proteica não é totalmente conhecida (BUTINI et al., 2008). A verificação da estabilidade da proteína AChE (4EY7) foi feita utilizando o RMSD (do inglês *root mean square deviation RMSD*) das trajetórias no decorrer da simulação, calculado sobre o *backbone* da proteína para todos os *frames*.

Todas as estruturas foram sobrepostas e as flutuações em torno da posição média (tomando como valor de referência o primeiro conjunto de coordenadas do sistema) são mostradas no gráfico da Figura 29, onde pode-se notar que por volta dos 30 primeiros nanosegundos de simulação as variações em torno da posição de equilíbrio, até os 206 nanosegundos finais de simulação, alcançam o equilíbrio, flutuando na média em torno de 2,25Å indicando baixa mobilidade da proteína.

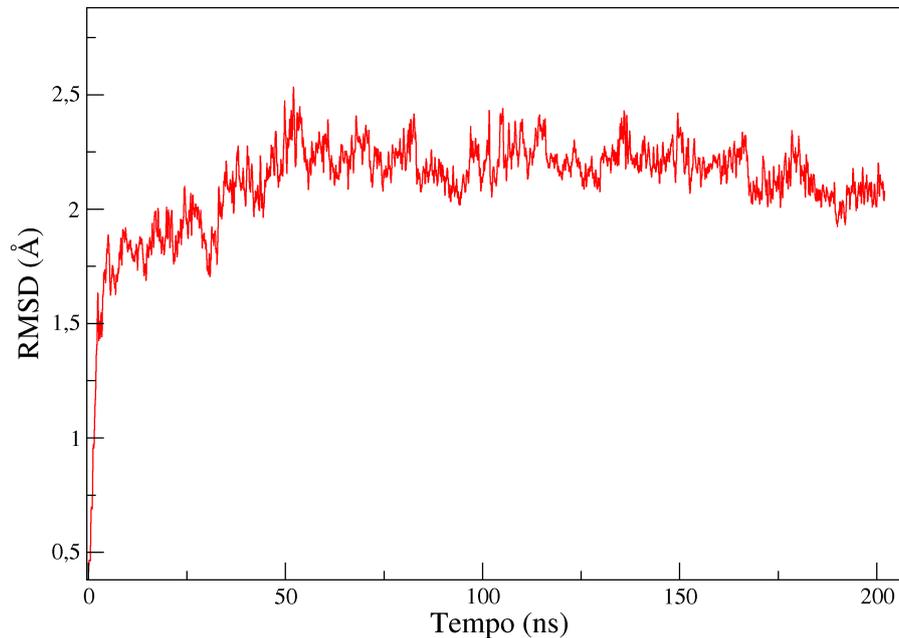


Figura 29 – Desvio quadrático médio (do inglês *root mean square deviation RMSD*) da AChE (4EY7) humana durante 206 ns de simulação de DM.

A dinâmica conformacional da Gorge vem sendo largamente estudada, uma vez que esse sítio ativo apresenta alta flexibilidade, e tal flexibilidade é apontada como fundamental para a atividade da enzima (HAQ et al., 2009). Nesse sentido, foi realizada a análise das flutuações dos aminoácidos da Gorge a partir do sexto nanossegundo de simulação, devido ao fato de que, para controle da simulação de DM, nos primeiros seis nanossegundos as porções da proteína foram liberadas aos poucos para uma dinâmica "livre" obedecendo uma ordem de prioridade, na qual primeiramente o *backbone*, depois as folhas beta e por último os aminoácidos pertencentes a Gorge flutuaram livremente.

Os resultados obtidos na análise a partir da evolução temporal do RMSD dos movimentos dos aminoácidos da Gorge é apresentado no gráfico abaixo, Figura 30. No RMSD calculado para esses aminoácidos, no início da simulação até por volta dos 30 primeiros nanossegundos, a Gorge apresenta alta flexibilidade, no entanto, desse ponto até o final da amostragem, na média o valor de RMSD fica em torno de 2,5Å, sugerindo uma estabilidade da Gorge.

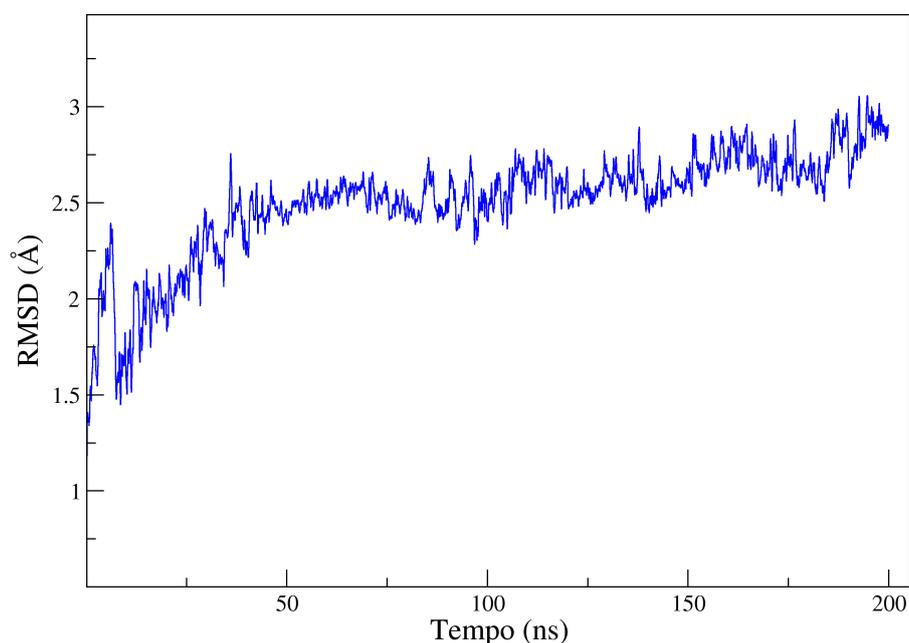


Figura 30 – RMSD dos aminoácidos da Gorge da AChE (4EY7) humana (Tyr103, Asp105, Trp117, Gly152, Gly153, Tyr155, Tyr164, Ser234, Ala235, Trp267, Glu316, Trp317, Phe326, Phe328, Glu365, Tyr368, Tyr372 e His478) durante 200 ns de simulação de DM.

Pode-se concluir pelos gráficos, figuras 29 e 30, que o RMSD para ambas as seleções (*backbone* e Gorge) mostram que o sistema está equilibrado e assim, na média, a mobilidade conformacional é baixa. Com o sistema devidamente equilibrado, realizamos o estudo de *Docking* para os benzilamínicos, no entanto, as primeiras análises desses resultados revelaram que o comportamento dinâmico da AChE ocorre de maneira muito complicada, de modo que para *frames* consecutivos em uma trajetória, a Gorge subitamente era pouco acessível ou até mesmo, inacessível. Nesse caso, para que os resultados de *Docking* sejam seguros, é imprescindível o conhecimento das particularidades conformacionais envolvidas nesse comportamento complicado. O estudo foi, nesse momento, direcionado para a investigação da dinâmica dos resíduos da Gorge, afim de encontrar a melhor conformação da local ativo para a interação com o ligante.

3.2.2 Dinâmica dos resíduos da Gorge

Estudos revelam que pequenas flutuações no *backbone* são importantes para o mecanismo de *gating* constatado para muitas enzimas, inclusive para a AChE. Esse mecanismo é responsável por controlar o acesso do substrato e de ligantes vinculados ao processo de catalise enzimática, pela imposição de restrições ao ligante, como constrições geométricas, especificidade de interações moleculares como ligações de hidrogênio, interações hidrofóbicas, entre outras; e ainda pode diferenciar pontos mais estreitos ao longo do caminho do sitio, formando um "gargalo" ou estreitamento (do inglês, "*bottleneck*"), como no caso da AChE. Os "portões" (do inglês "*gate*") constituem regimes dinâmicos que fazem transições reversíveis entre estados "*open*" e "*closed*", dada várias possibilidades de *gating*, os mecanismos mais complexos são aqueles que envolvem o movimento de dois os mais resíduos específicos do sítio ativo (GORA; BREZOVSKY; DAMBORSKY, 2013).

A AChE possui um mecanismo de *gating* bem complexo, que modula o acesso do substrato e de outros ligantes no sitio ativo. O movimento de "respiração" (*breathing motions*) dos resíduos da Gorge, que opera uma dinâmica coletiva com a proteína gera os estados "*open*" e "*closed*" na região de estreitamento da Gorge. Esse mecanismo foi avaliado por Yechun Xu *et al.* (XU *et al.*, 2008) para quarenta e sete estruturas de *Torpedo Californica* AChE (TcAChE) e outros estudo, para estruturas cristalográficas de outros organismos também comprovam esse movimento característico na AChE e ainda, que o movimento coletivo da enzima está acoplado com o movimento do ligante através da região do estreitamento (BUI; HENCHMAN; MCCAMMON, 2003).

O estudo da dinâmica dos resíduos da proteína que contribuem para o mecanismo de *gating*, presentes na região de estreitamento da Gorge, é fundamental para identificar as conformações da proteína, já que mudanças conformacionais dos aminoácidos do sítio ativo estão diretamente relacionadas com a interação de ligantes no mesmo. Os aminoácidos com alta flexibilidade no estreitamento da Gorge observados nesse estudo foram: Trp317, Tyr155 e Tyr372 localizado no PAS e Phe328 localizada sub sítio Acil e os resultados obtidos serão exibidos a seguir.

O primeiro resultado é a medida de distância entre os aminoácidos Phe328 presente no subsítio Acil e Tyr372 localizada no sítio aniônico periférico. De acordo com o gráfico da Figura 31 a distância entre os centros de massa da Phe328 e da Tyr372 tem valor

mínimo de aproximadamente 7\AA e máximo de aproximadamente 14\AA , mostrando que há uma flutuação considerável nas diferenças de distâncias. Isso mostra que estes dois aminoácidos tem papel importante no mecanismo de *gating* da Gorge.

Nos primeiros 10 ns os aminoácidos estão afastados, depois disso conservam uma distância média em torno de 8\AA até os 50 ns, quando essa distância média passa para 9\AA até os 100 ns. Entre 100 e 200 ns observamos maior frequência de grandes diferenças nas distâncias, indicando que, a utilização desse parâmetro como método para o resgate de conformações da proteína, onde possivelmente o estreitamento encontraria-se no estado "open", pode ser realizado para várias trajetórias durante a dinâmica.

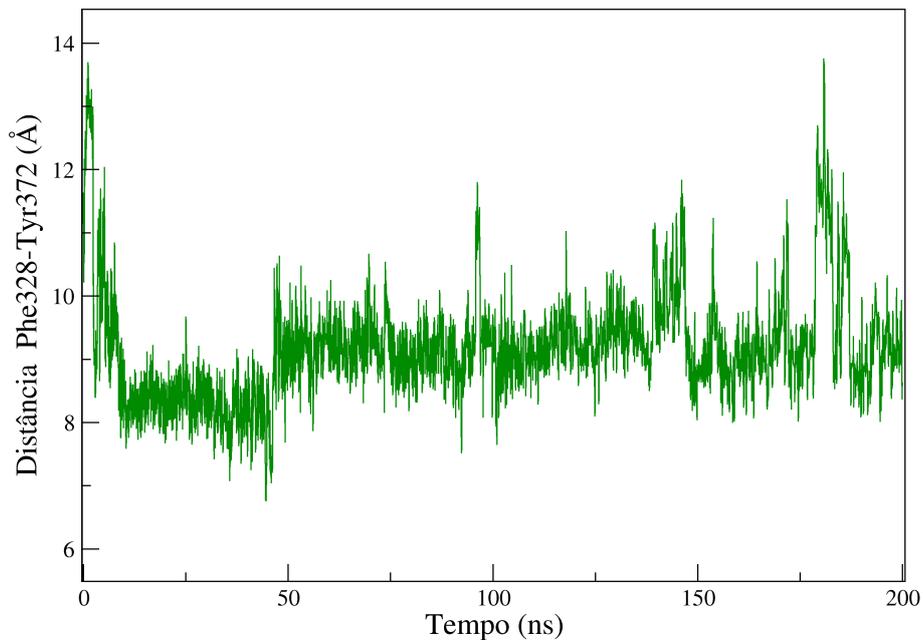


Figura 31 – Medida de distância entre os aminoácidos Phe328 pertencente ao sub sitio Acil e Tyr372 pertencente ao PAS, a partir do centro massa de cada um dos aminoácidos, durante 200 ns de simulação de DM.

Na Figura 32 é possível ver a localização dos aminoácidos Phe328 e Tyr372 na Gorge, e o afastamento e aproximação que esses aminoácidos sofrem durante a dinâmica. A medida de distância apresentada nessa figura foi feita considerando um átomo do anel aromático de cada aminoácido, e pode-se perceber que esses átomos em diferentes *frames*

de uma trajetória, podem assumir distâncias razoavelmente pequenas como 5.61\AA , até distâncias que ultrapassam o dobro desse valor, como 12.79\AA .

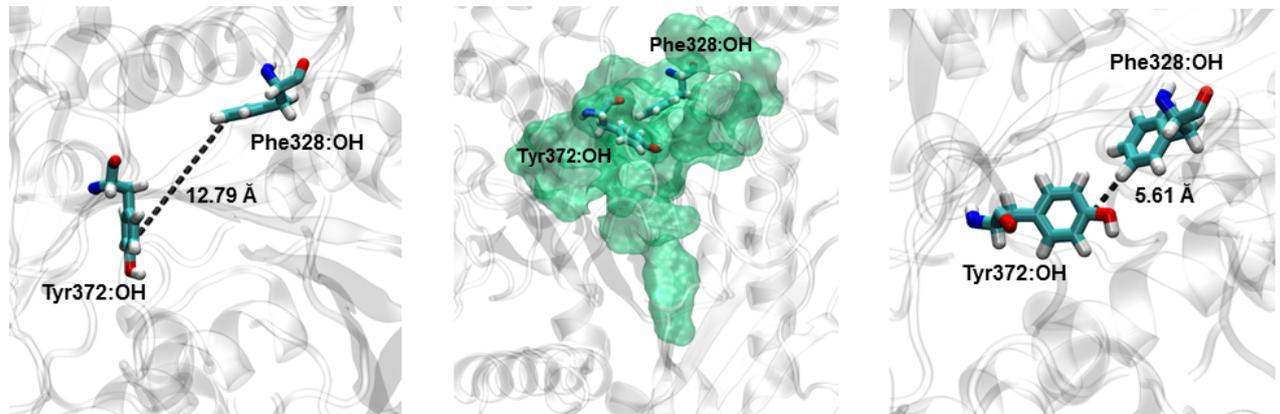


Figura 32 – Posicionamento dos aminoácidos Phe328 e Tyr372 na Gorge (superfície molecular na cor verde) e distâncias assumidas entre dois átomos específicos (Phe328:CE1, Tyr372:CE1) destes dois aminoácidos em um *frame* da trajetória. Os átomos de carbono e hidrogênio estão coloridos de ciano e branco, respectivamente, enquanto os átomos de nitrogênio estão na cor azul e os de oxigênios em vermelho.

O Trp317 e a Tyr372 são aminoácidos pertencentes ao PAS que durante as observações da amostragem da dinâmica mostraram-se importantes para o estreitamento, principalmente por estarem próximos à *main door*. As distâncias entre os centros de massa desses dois aminoácidos, apresentadas no gráfico da Figura 33, possui variação máxima de aproximadamente 4\AA entre as maiores e menores distâncias computadas ao longo dos 200 ns de amostragem, e obedecem a um padrão de frequência de flutuação.

Na média a distância entre os centros de massa mantém-se em 8\AA , havendo mudança suaves e crescentes em alguns pontos da amostragem como logo após os 50 ns, onde as distâncias gradualmente aumentam chegando valores acima de 10\AA e depois caem novamente voltando à média. A medida de distância entre Trp317 e Tyr372 como estratégia para o reconhecimento do estado "open" mostra que há poucas trajetórias, assumindo que esse estado exista para as maiores distâncias encontradas (maiores ou iguais a 10\AA), que apresentam distâncias características do estado "open".

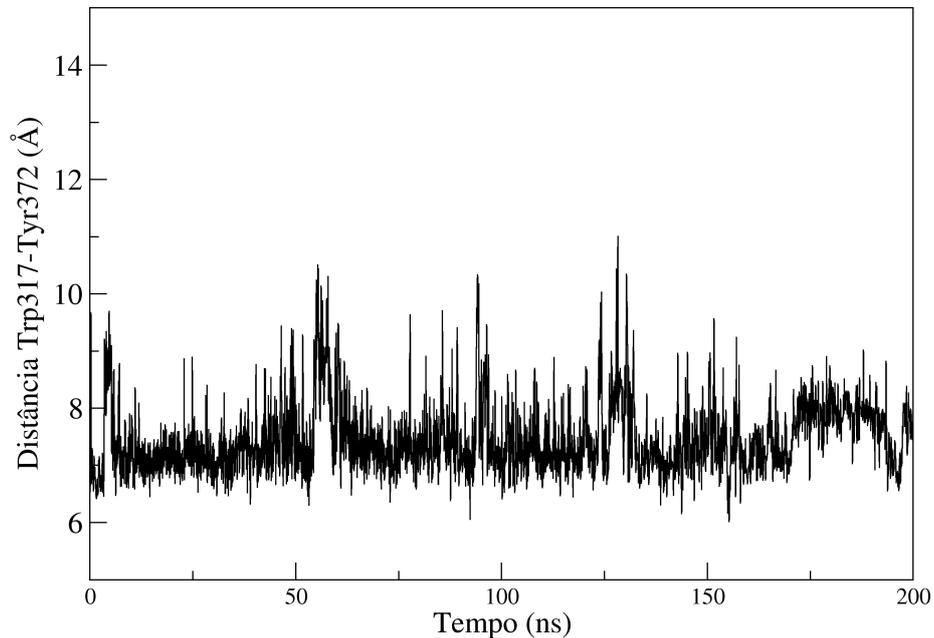


Figura 33 – Medida de distância entre os aminoácidos Trp317 e Tyr372 pertencentes ao PAS, a partir do centro massa de cada um dos aminoácidos, durante 200 ns de simulação de DM.

Na Figura 34 pode-se notar que tanto o Trp317 quanto a Tyr372 estão rente a *main door*, comprovando que a aproximação e afastamento desses aminoácidos pode bloquear ou liberar a *main door*, interferindo no acesso de ligantes e até mesmo do substrato na Gorge. A medida de distância entre dois átomos específicos dos aminoácidos, (Trp317:CZ2, Tyr372:OH) também contidos na figura, deixa claro que esses átomos podem assumir distâncias com diferenças próximas de 10 Å, ainda percebe-se a mobilidade da Tyr372 e do Trp317 que ao se aproximarem ou se afastarem realizam movimentos conhecidos como *swinging doors*¹.

¹ O movimento dos resíduos: *Swinging Doors*, representa aquele em que dois resíduos podem girar, na mesma direção ou em direção oposta, estabilizando conformação *open* ou *closed* a um *gate*, como no caso da Acetilcolinesterase (AChE), esses portões também são conhecidos como inclináveis (GORA; BREZOVSKY; DAMBORSKY, 2013).

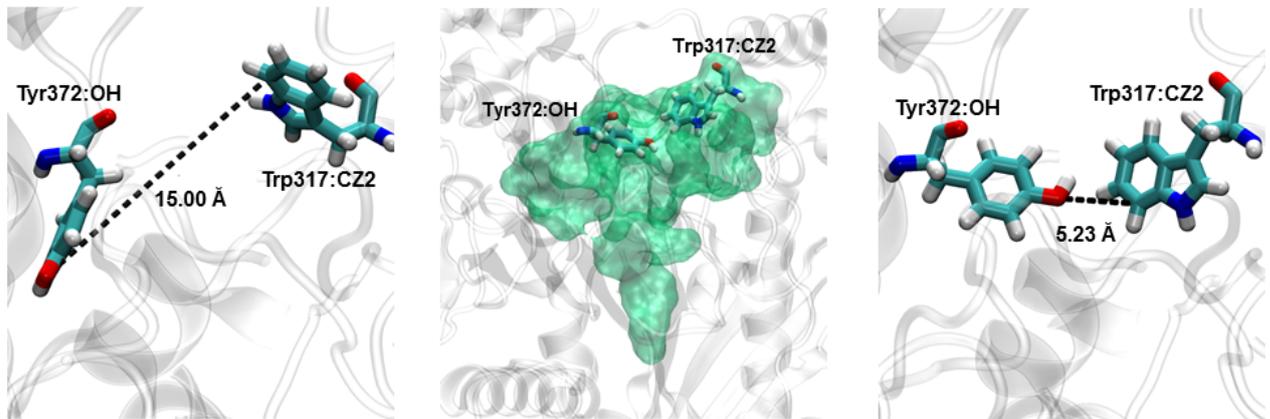


Figura 34 – Posicionamento dos aminoácidos Trp317 e Tyr372 na Gorge (superfície molecular na cor verde) e distâncias assumidas entre dois átomos específicos (Phe328:CE1, Tyr372:CE1) destes dois aminoácidos em um *frame* da trajetória. Os átomos de carbono e hidrogênio estão coloridos de ciano e branco, respectivamente, enquanto os átomos de nitrogênio estão na cor azul e os de oxigênios em vermelho.

O último par de aminoácidos analisados, com relação à distância, para a busca na distinção entre os estados *open* e *closed* foram as tirosinas Tyr155 e Tyr372, também pertencentes ao PAS. Durante a visualização do sistema no tempo de amostragem, esse par chamou muita atenção pela forma como se aproximavam, mantendo distancias muito pequenas por um período relativamente curto de tempo, e então se afastavam e as distâncias de afastamento eram, comparadas com os outros pares de resíduos analisados, as menores.

No gráfico da Figura 35 a medida de distância entre o centro de massa dos resíduos Tyr155 e Tyr372 ao longo dos 200 ns de amostragem traça nos primeiros 10 ns um perfil onde a distância média é cerca de 11Å, depois cai bruscamente para 8Å, permanecendo assim até os 50 ns, onde volta a apresentar distâncias maiores, chegando acima de 13Å no final da amostragem. Tomando esse resultado como artifício para localizar estados *open* e *closed*, os primeiros 50 ns de simulação seriam negligenciados, e somente a partir desse ponto da amostragem, o sistema mantém conformações onde o estado *open* é mais provável.

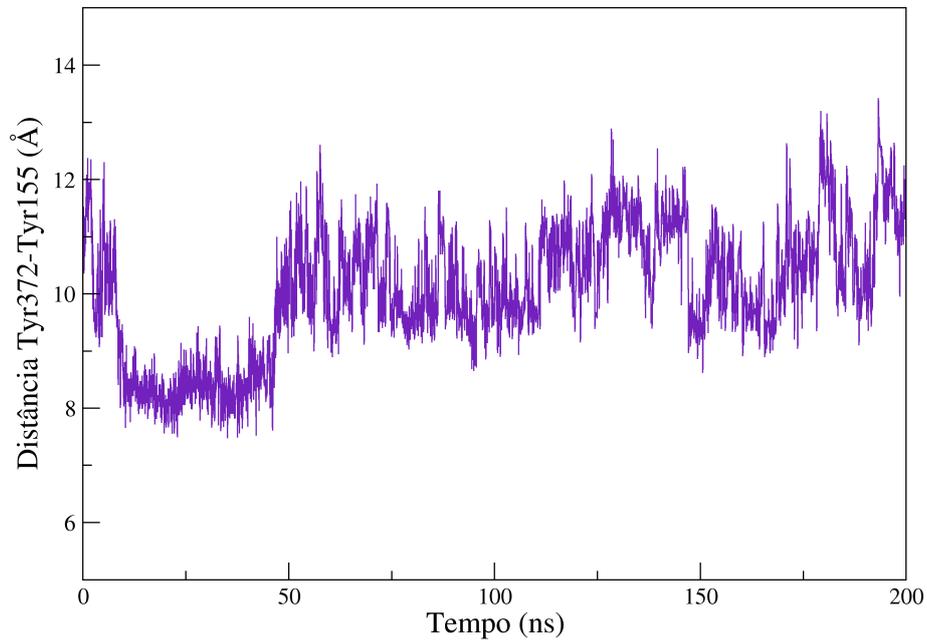


Figura 35 – Medida de distância entre os aminoácidos Tyr155 e Tyr372 pertencentes ao PAS, a partir do centro massa de cada um dos aminoácidos, durante 200 ns de simulação de DM.

Apesar da medida de distância do centro de massa para Tyr155 e Tyr372 indicarem que estes resíduos acham-se afastados na maior parte do tempo de amostragem, a medida de distância entre os dois átomos específicos para estes aminoácidos (Tyr155:OH, Tyr372:OH) evidenciam que eles podem ser encontrados muito próximos e contribuir imensamente para a constrição do Gorge. Uma medida de distância entre esses dois átomos para um *frame* de determinada trajetória, é apresentada na Figura 36, e exibe uma distância muito pequena de 2.84Å, podendo no momento de afastamento alcançar 9.47Å.

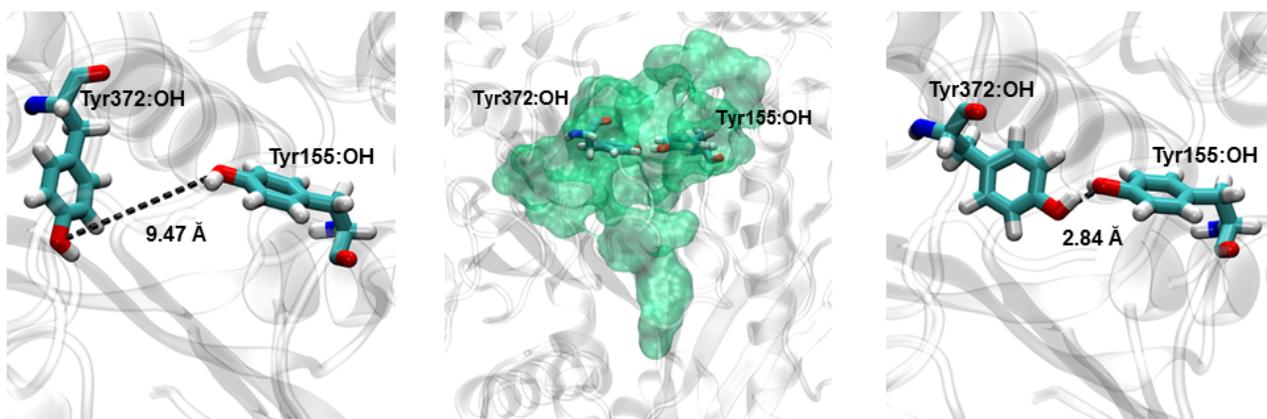


Figura 36 – Posicionamento dos aminoácidos Tyr155 e Tyr372 na Gorge (superfície molecular na cor verde) e distâncias assumidas entre dois átomos específicos (Tyr155:OH, Tyr372:OH) destes dois aminoácidos em um *frame* da trajetória. Os átomos de carbono e hidrogênio estão coloridos de ciano e branco, respectivamente, enquanto os átomos de nitrogênio estão na cor azul e os de oxigênios em vermelho.

Como demonstrado nas análises dos resíduos que certamente estão envolvidos na caracterização dos estados do estreitamento que definem a Gorge ("open" e "closed"), selecionar estados no qual estamos interessados ("open") não é uma tarefa fácil, uma vez que a dinâmica da AChE apresenta uma "respiração" e o mecanismo de *gating* devido aos resíduos do estreitamento parece estar sincronizado com a "respiração", de maneira que essa "respiração" pode deslocar os resíduos.

Para discriminar os estados da proteína "open" e "closed", alguns parâmetros podem ser avaliados, como por exemplo, o comportamento dos resíduos constituintes (todos os resíduos do sistema, ou uma parte específica), os aminoácidos que tornam a estrutura flexível e possibilitam a mudança conformacional e o diâmetro ou raio do estreitamento, buscando mudanças ocorridas neste ao longo do tempo.

Os dois primeiros parâmetros foram utilizados nas análises anteriores, e a partir de agora, apresenta-se o terceiro parâmetro que é a seleção dos estados "open" e "closed" por meio da análise da variação do raio da região de estreitamento, para cada conformação assumida pela proteína nas trajetórias, durante os 200 ns de simulação.

A análise das dimensões do estreitamento e da Gorge como um todo, foi realizada empregando o programa HOLE (SMART; GOODFELLOW; WALLACE, 1993; SMART et

al., 1996) implementado no VMD (HUMPHREY; DALKE; SCHULTEN, 1996), que simula o percurso percorrido por uma esfera de raio variável, ao longo de um poro ou cavidade do sítio ativo, por exemplo. Neste programa as coordenadas dos átomos do sistema são fornecidos e a definição de um ponto inicial p que passe por qualquer local da cavidade é necessário. Também é preciso definir um vetor aproximadamente na direção da cavidade, assim o programa faz a leitura dos átomos no arquivo de coordenadas e associa um raio de *van der Waals* para cada um. A esfera é então centrada no ponto p de forma que seu raio é maximizado sem sobreposição com a superfície de *van der Waals* em todas as direções, e o raio da esfera é contabilizado num plano normal ao vetor da direção da cavidade.

A Figura 37 mostra a definição de um vetor que atravessa a Gorge na direção da coordenada Z , para a aplicação do programa HOLE. O ponto inicial p foi definido pelas coordenadas de um aminoácido da AChE que se encontra além do fundo da Gorge (próximo à tríade catalítica). Os resultados do perfil da Gorge gerado a partir do HOLE contém o raio da Gorge para cada ponto do vetor na coordenada Z , esses resultados podem ser visualizados na forma gráfica ou na superfície gerada no VMD. Essas duas maneiras de visualizar o perfil encontram-se na figura 37, onde a *main door* encontra-se localizada no ponto 10Å do eixo z , a região de constricção da Gorge na vizinhança do ponto 0Å no eixo Z e na posição -10Å no eixo Z situa-se fundo da Gorge (ponto p).

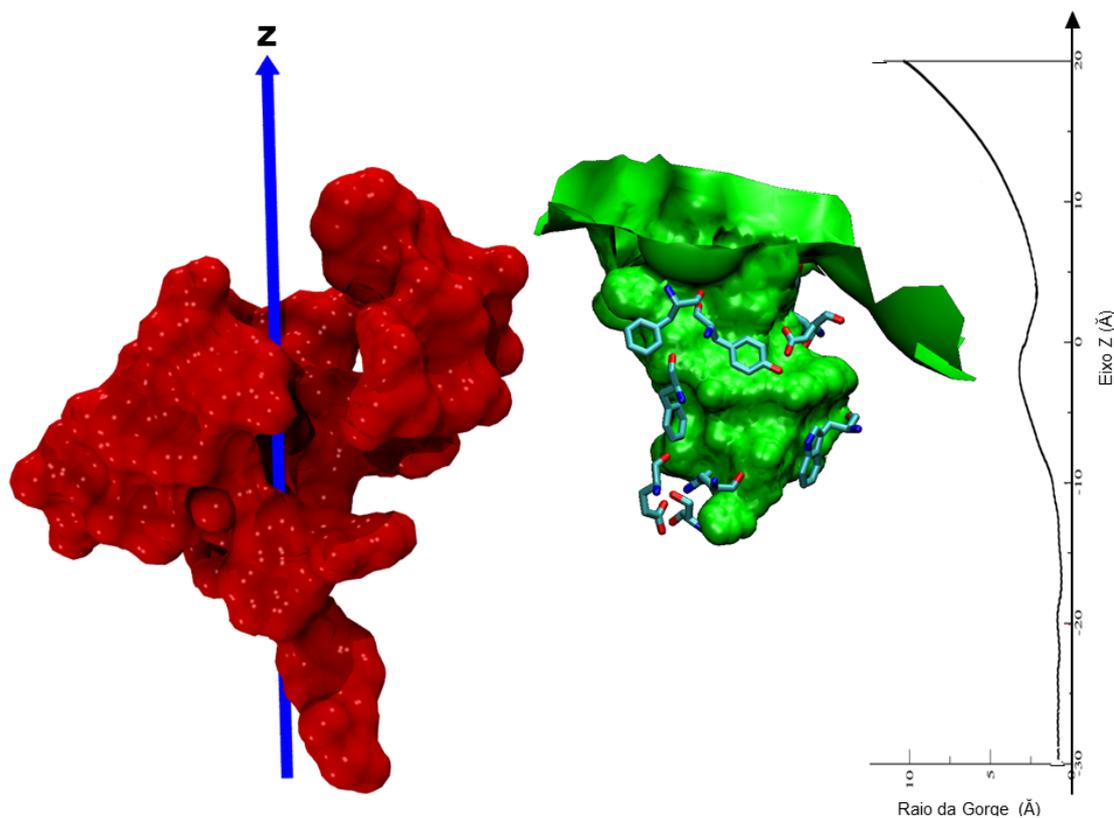


Figura 37 – Representação da construção do perfil da Gorge traçado pelo programa HOLE: A Gorge é mostrada na superfície na cor vermelha, e no centro da Gorge um vetor na direção do eixo Z do sistema. A superfície na cor verde, criada no programa VMD, é o resultado gerado pelo programa HOLE, a partir do raio da Gorge em cada ponto do vetor no eixo Z, ao lado direito dessa superfície a visualização gráfica dos resultados do perfil da Gorge fornecidos pelo HOLE.

O número total de *frames* no tempo de amostragem de 200 ns foi de 100000 (cem mil) *frames*. Realizamos o cálculo do perfil da Gorge para todos os *frames*, gerando cem mil arquivos de perfis da Gorge. A avaliação desses perfis foi realizada de maneira sistemática e minuciosa e interpretamos distribuições do sistema em quatro estados para a Gorge.

Aos estados atribuímos dois aspectos que os definem: o raio encontrado na região da constrição da Gorge e o raio encontrado nas redondezas da *main door*. Para o perfil que tem raio pequeno (em torno de 1\AA) na região da constrição, mas apresenta raio maior que 3\AA na região da *main door*, denominamos estado *ClosedOpen*. O perfil que apresenta raio em torno de 1\AA na região da constrição e possui raio menor que 3\AA na região da *main door*, denominamos estado *Closed*. O perfil no qual a região da constrição possui raio maior que 3\AA , porém nas redondezas da *main door* esse raio tem valor de 1\AA , chamamos

de estado *OpenClosed*. E o último e mais relevante estado é aquele onde as duas regiões descritas (*main door* e *constricção*) possuem raio maior que 3\AA , e titulamos de estado *Open*.

Com base nas definições acima para os perfis, separamos os quatro estados de maneira que o estado *Closed* corresponde a 14% dos perfis para o sistema, sendo que 34,6% dos perfis representam o estado *ClosedOpen*, 49% diz respeito ao estado *OpenClosed* e somente 2,4% dos perfis configuram o estado *Open*. O gráfico da Figura 38 contém as curvas médias do perfil de cada estado, e uma curva referente ao perfil encontrado para a estrutura cristalográfica da AChE-4EY7 humana. Ainda na Figura 38, a curva tracejada na cor vermelha representa o perfil para o cristal, que é similar àquela encontrada para o estado *Open*, transparecendo a confiabilidade dos procedimentos adotados, considerando que esse cristal está complexado com o Donepezil.

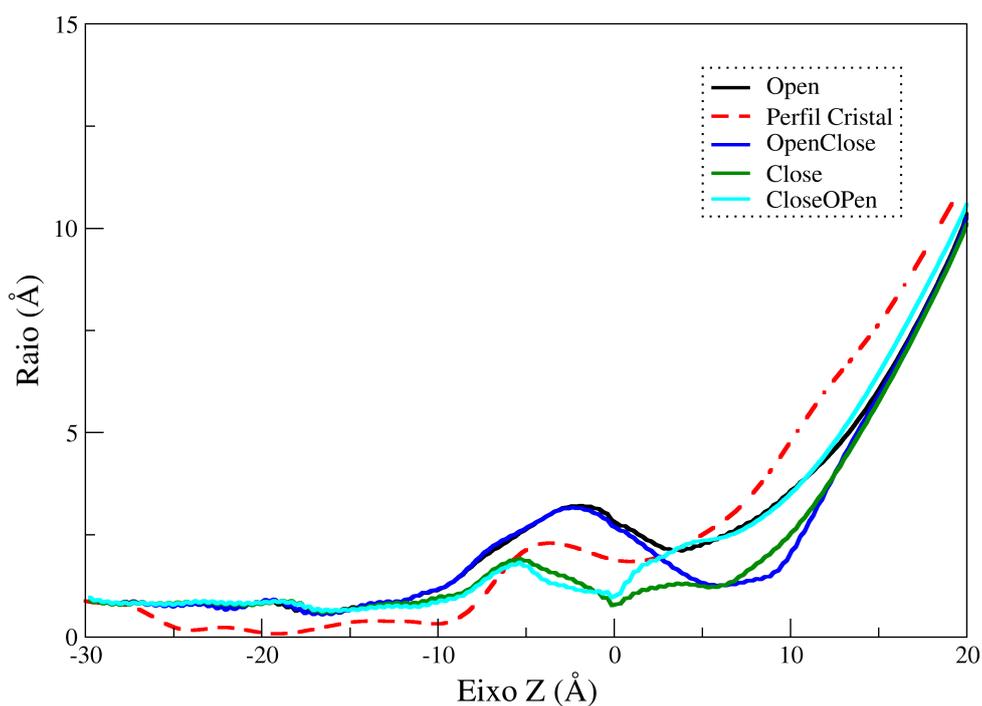


Figura 38 – Curvas para os perfis (raio da Gorge ao longo do eixo Z, onde o ponto 0\AA representa a região de estreitamento da Gorge e nas proximidades de 10\AA a *main door*) dos estados: *CloseOpen*, curva na cor turquesa; *OpenClosed*, curva na azul; *Closed*, curva na cor verde; *Open*, curva na cor preta. Curva para o perfil do cristal da hAChE 4EY7, linha tracejada na cor vermelha.

Com a finalidade visualizar, de maneira mais clara, os estados descritos através do estudo dos perfis, selecionamos um *frame* representativo para os estados *Open* e *Closed*, e por meio do Volmap, que é um *plugin* do VMD (HUMPHREY; DALKE; SCHULTEN, 1996) que cria mapas volumétricos tridimensionais com base nos dados moleculares, geramos um mapa volumétrico para cada um destes *frames*, optando pelo tipo de mapa baseado na "ocupação".

No mapa por "ocupação", considera-se os raios atômicos dos átomos que são tratados como esferas, e todo o sistema é construído de pequenas grades cubicas. A cada ponto dessa grade é associado o valor 0, se dentro da grade não contém átomos, e 1 se dentro da grade conter átomos. Encontrando átomo num ponto da grade, esse então é considerado como "ocupado". O mapa foi gerado usando a representação *VolumeSlice*, e as cores compreendem o intervalo de valores escalares do conjunto de dados, com vermelho especificando baixos valores e azul indicando alto valores nos dados.

A Figura 39 expõe os mapas encontrados para os estados *open* e *Closed*, nela é possível identificar a "ocupado" de átomos nas regiões do estreitamento e na *main door* no lado esquerdo da figura, e a Gorge sem esses estreitamentos, ou seja, vazia ou sem "ocupação". Ainda fica evidente as diferenças nas distribuições das "ocupações" da proteína como um todo, se comparamos a distribuição dos lados direito e esquerdo, que torna mais forte as convicções de que os movimentos dos átomos nas regiões de constrição estão acoplados aos movimentos da proteína, quando a mesma "respira".

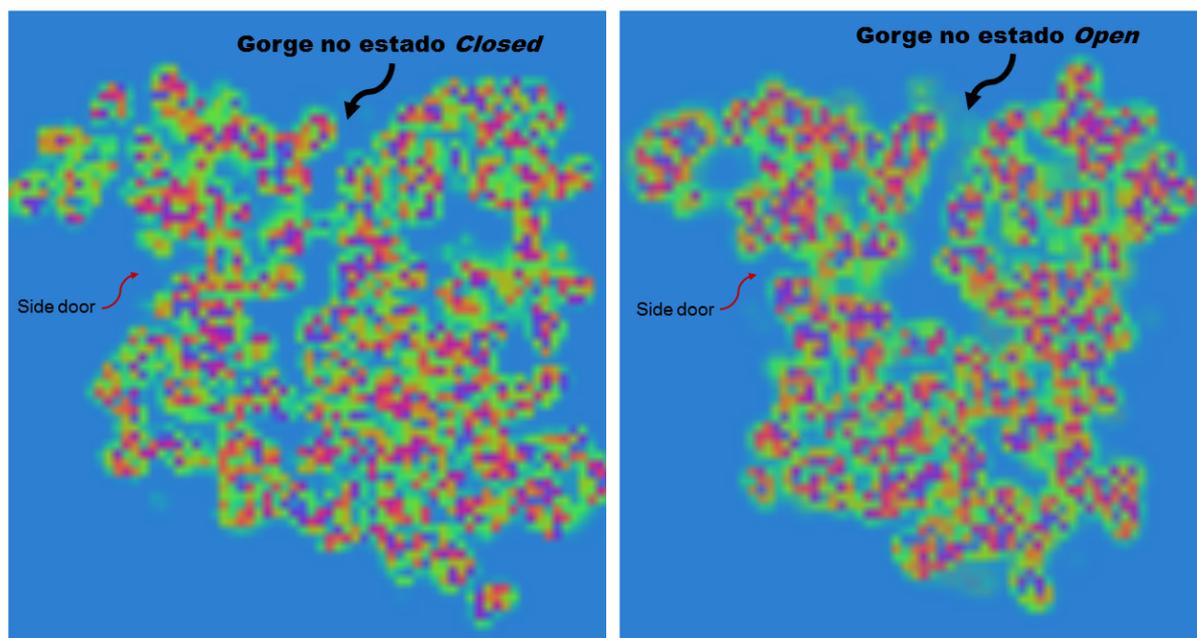


Figura 39 – Mapa volumétrico da AChE, nos estados *Closed*, (lado esquerdo) e *Open* (lado direito). Os pontos mais azuis indicam maior volume de conjunto de dados e pontos mais vermelhos menor volume no conjunto de dados

Uma vez encontrado o estado *Open*, estamos interessados em compreender as prováveis orientações preferencias com as quais os ligantes interagem com os resíduos do sitio biológico da proteína. Os próximos resultados a serem apresentados são os resultados para o atracamento molecular dos ligantes LDT160, LDT161, LDT167, LDT169 e do Donepezil com a Gorge, no estado *Open*.

3.2.3 Resultados de Docking receptor-ligante

Os estudos de atracamento molecular (*docking* receptor-ligante) foram realizados considerando as conformações da proteína no estado *Open* de acordo com o perfil apresentado na Figura 38, e a geometria inicial utilizada para os ligantes foram as de menor energia, otimizadas anteriormente para a análise de PCA. Uma vez que a proteína hAChE (Código PDB: 4EY7) está complexada com o Donepezil, com finalidade de validar se o *Docking* realizado pelo Autodock Vina era fidedigno para esse sistema, realizamos o *Docking* do Donepezil no conjunto de soluções (*ensemble*) conformacionais da hAChE no estado *Open*.

As 842 conformações encontradas para o acoplamento do Donepezil no *ensemble* foram comparados com a pose do cristal, através do cálculo de RMSD para todas as 842 conformações, e o resultado foi muito satisfatório, já que o valor do RMSD entre uma das soluções de *Docking* e a conformação do Donepezil no cristal foi de 1.34\AA ². Isto significa que o *Docking* realizado no programa AutoDock Vina foi capaz de reproduzir a orientação do Donepezil no cristal (modelo), assegurando que é adequado para o sistema da hAChE. Para termos uma visão melhor a respeito das conformações do Donepezil comparadas e mencionadas acima, a Figura 40 mostra a sobreposição estrutural do Donepezil no cristal, com a pose que o representa no *ensemble*, constatando que as conformações são muito próximas, o que justifica o valor encontrado para o RMSD.

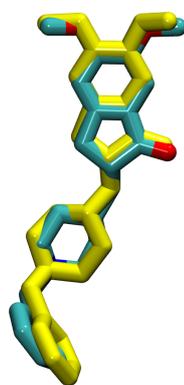


Figura 40 – Superposição das estruturas do Donepezil no cristal, representada nas cores ciano (carbonos), vermelha (oxigênio) e azul (nitrogênio); e no *ensemble* representada na cor amarela.

O Donepezil é uma mistura racêmica de anantiômeros R e S, possuindo a mesma forma molecular, exceto para o anel piperidina. Com relação a interação com a AChE, ligam-se de maneira semelhante e possuem afinidades de ligação equivalentes (SUGIMOTO et al., 2000). A ligação do Donepezil com a Gorge compreende resíduos ao longo da Gorge: o anel benzil empilha com o Trp117 (no sub sítio aromático), e a outra extremidade empilha com o Trp317 no sítio aniônico periférico, podendo a parte central, piperidina, interagir com Tyr372. Para Cheung *et al.*, mudanças conformacionais na Gorge podem influenciar nos modos de ligação de inibidores mais longos que possuam variabilidade

² Para comparações de *Docking*, este é considerado conveniente se o valor médio de RMSD entre o encaixe e a pose de referência (cristal) obedecer a um limite de 2\AA (FANG et al., 2014).

conformacional, como duplos inibidores de AChE que simultaneamente se ligam em dois sítios (CHEUNG et al., 2012).

Como dito previamente, a docagem do Donepezil no *ensemble* de conformações da AChE, gerou 842 soluções de posições para o Donepezil, as quais vamos referir como uma nuvem de soluções. Essa nuvem de soluções está representada na Figura 41, onde as partes em azul nessa nuvem são referentes aos nitrogênios protonados da piperidina, e a vermelha aos oxigênios presentes na outra extremidade do Donepezil. Essas duas "manchas", azul e vermelha, se misturam na *main door* e no fundo da Gorge indicando que o *Docking* encontrou dois modos de ligação preferenciais para o Donepezil.

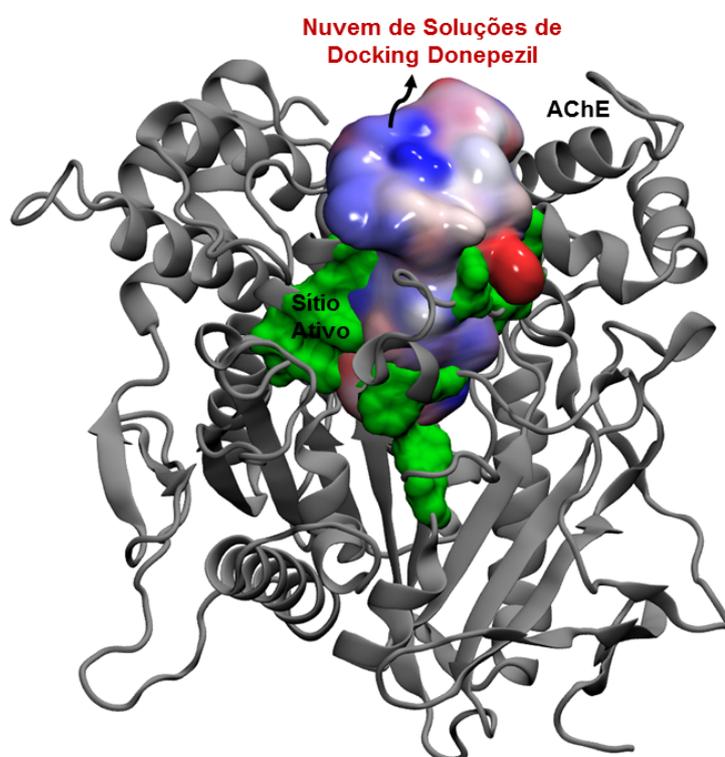


Figura 41 – Nuvem de soluções para o *Docking* do Donepezil no sítio ativo da AChE (código PDB:4EY7) contendo 842 conformações do Donepezil e duas orientações preferenciais. A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde.

O *Docking* foi realizado para todos os ligantes, LDT160, LDT161, LDT167 e LDT169 e as nuvens de soluções de *Docking* obtidas para cada um deles estão apresentadas a seguir, e foram construídas obedecendo as mesmas especificações de cores da nuvem para o Donepezil.

O *Docking* para o ligante LDT160 gerou 987 soluções e também apresentou uma mistura de duas principais orientações preferências, de modo que a sua nuvem apresenta tons vermelhos e azuis na *main door* e no fundo da Gorge e pode ser visto na Figura 42.

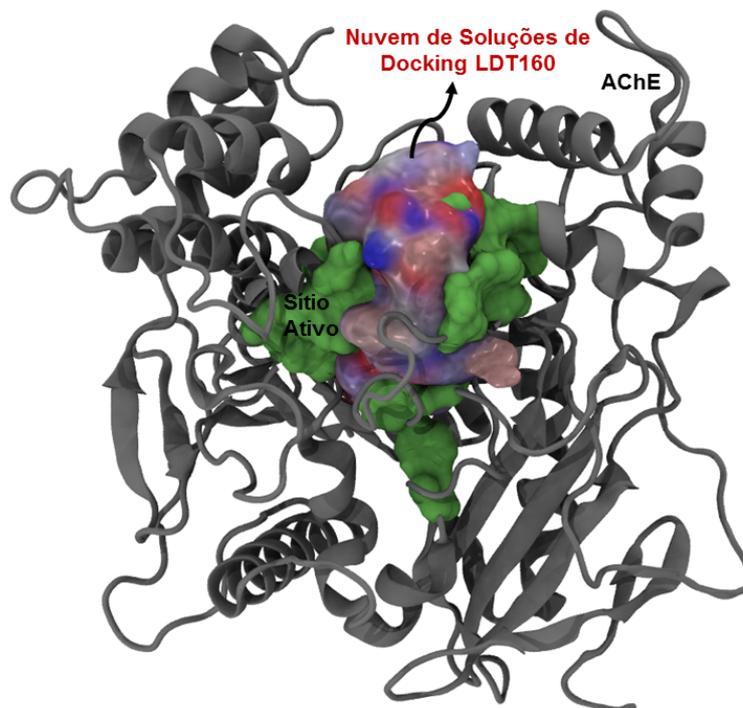


Figura 42 – Nuvem de soluções para o *Docking* do ligante LDT160 no sítio ativo da AChE (código PDB:4EY7) contendo 987 conformações do Donepezil e duas orientações preferenciais. A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde.

As soluções encontradas para o ligante LDT161 foram ao todo 994 e, da mesma forma que para o LDT160, essas soluções apresentam duas orientações preferenciais. No entanto, a nuvem para o LDT161, ver Figura 43, apresenta uma diferença significativa em relação a nuvem do LDT160, na parte superior da nuvem (região da *main door*) nota-se a dominância de tons vermelhos, indicando que apesar de existirem dois modos de ligação, há um que prevalece.

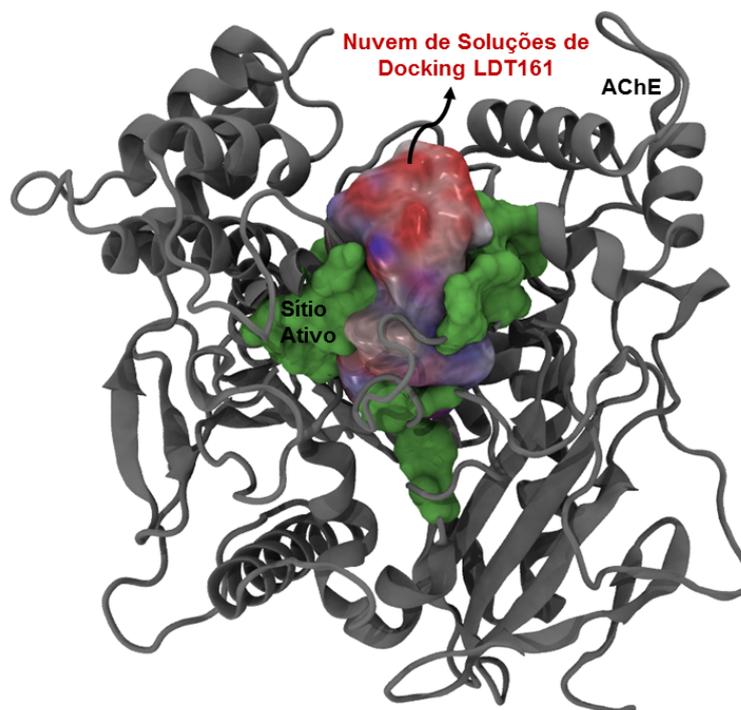


Figura 43 – Nuvem de soluções para o *Docking* do ligante LDT161 no sítio ativo da AChE (código PDB:4EY7) contendo 994 conformações do Donepezil e duas orientações preferenciais. A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde.

Para o ligante LDT167 o número de soluções de *Docking* foi 960, com duas conformações preferenciais, e a análise para essa molécula é bem parecida com a feita pra LDT161. Levando em consideração a nuvem de soluções da Figura 44, na qual pode-se ver a predominância da cor vermelha na parte superior da nuvem, e comparando essa nuvem com aquela encontrada para o ligante LDT161, existe um ponto nessa parte superior da nuvem do LDT167 em que a cor azul é bem acentuada, sugerindo que o número de soluções, cuja conformação do ligante indica o nitrogênio próximo à *main door*, é maior para o LDT167 do que para LDT161.

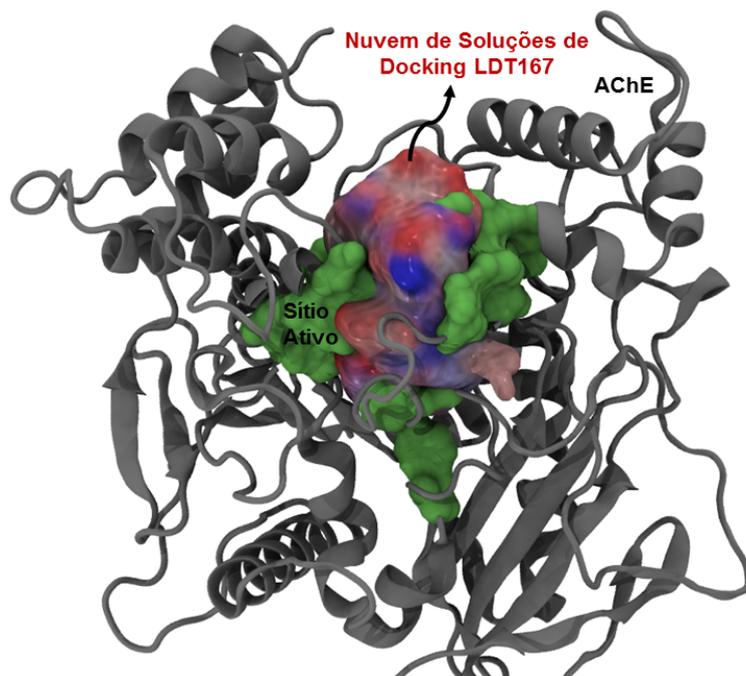


Figura 44 – Nuvem de soluções para o *Docking* do ligante LDT167 no sítio ativo da AChE (código PDB:4EY7 contendo 960 conformações do Donepezil e duas orientações preferenciais). A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde.

Já para as soluções do ligante LDT169, para qual foram encontradas 997 soluções em duas orientações preferenciais, a nuvem de soluções, Figura 45, não deixa muito claro se existe alguma predominância de modo de ligação. Nessa nuvem, presume-se que quantidade das duas orientações sejam muito próximas.

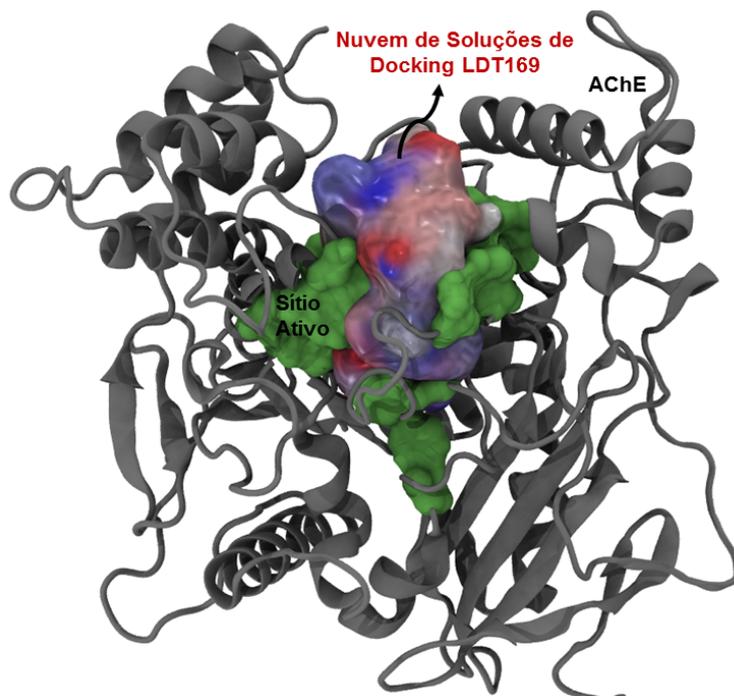


Figura 45 – Nuvem de soluções para o *Docking* do ligante LDT169 no sítio ativo da AChE (código PDB:4EY7) contendo 997 conformações do Donepezil e duas orientações preferenciais. A enzima AChE está representada pelo modelo de fitas na cor cinza grafite e o sítio ativo é indicado na superfície verde.

3.2.3.1 Orientações Preferenciais apontadas pelo *Docking*

Com o propósito de interpretar as soluções de *Docking* para os dois modos de ligação revelados, de maneira mais esclarecedora, e portanto, aumentar o grau de informação a respeito das interações dos ligantes com o sítio biológico da AChE, adotamos o RMSD e as energia de ligação (dadas pelo *Docking*) entre as duas conformações como parâmetro para diferenciá-las. O algoritmo criado seleciona a solução de *Docking* de menor energia de ligação (escolhida como referência), para cada um dos ligantes e para o Donepezil separadamente, e compara com todas as outras soluções por meio do cálculo do RMSD.

A distinção entre os dois modos de ligação encontrados no *Docking* para o Donepezil, de acordo com o gráfico da Figura 46, retrata as duas populações que a metodologia foi capaz de discriminar. A população na parte inferior do gráfico representa aquelas para as quais o Donepezil ancora no sítio da AChE com a piperidina protonada voltada para o

fundo da Gorge. Nessa população inferior, localizamos a solução para o cristal, identificada no gráfico com uma estrela vermelha.

A população superior no gráfico é aquela em que o Donepezil acessa o sítio com sentido contrário ao que é encontrado nas estruturas cristalográficas, nela a piperidina protonada coloca-se próxima à *main door*. A análise desses resultados refletem ainda que além de encontrarmos duas populações com modos de ligação distintos, para as quais o Donepezil é capaz de ligar-se em duas orientações, as energias de ligação para essas duas populações podem ter o mesmo valor.

Segundo Jiansong Fang (FANG et al., 2014), várias estruturas cristalográficas da AChE recombinada com a o Donepezil, em diferentes organismos, revelam que o Donepezil interage com a AChE simultaneamente em dois sítios e esses cristais revelam que existe uma orientação com a qual o Donepezil interage.

Considerando os dois modos de ligação encontrados no *Docking*, resolvemos nesse primeiro momento considerá-las, já que o mesmo aconteceu com os outros ligantes, e também para estabelecer critérios equivalentes para a divisão das duas orientações de todos os ligantes.

Na parte superior e inferior do gráfico 46 as soluções com energia menor ou igual a -9 kcal/mol, em torno de 1,5 kcal/mol a menos do que a energia encontrada para a solução do cristal, que foi de -10,4 kcal/mol, serão divididas em duas populações. Portanto, a população de soluções com a piperidina protonada perto da *main door* corresponde 21 % do número total de soluções, e a população inferior no gráfico (piperidina protonada no fundo da Gorge) corresponde a 18 % do número total de soluções e são aquelas em que o Donepezil encontra-se na mesma orientação do cristal.

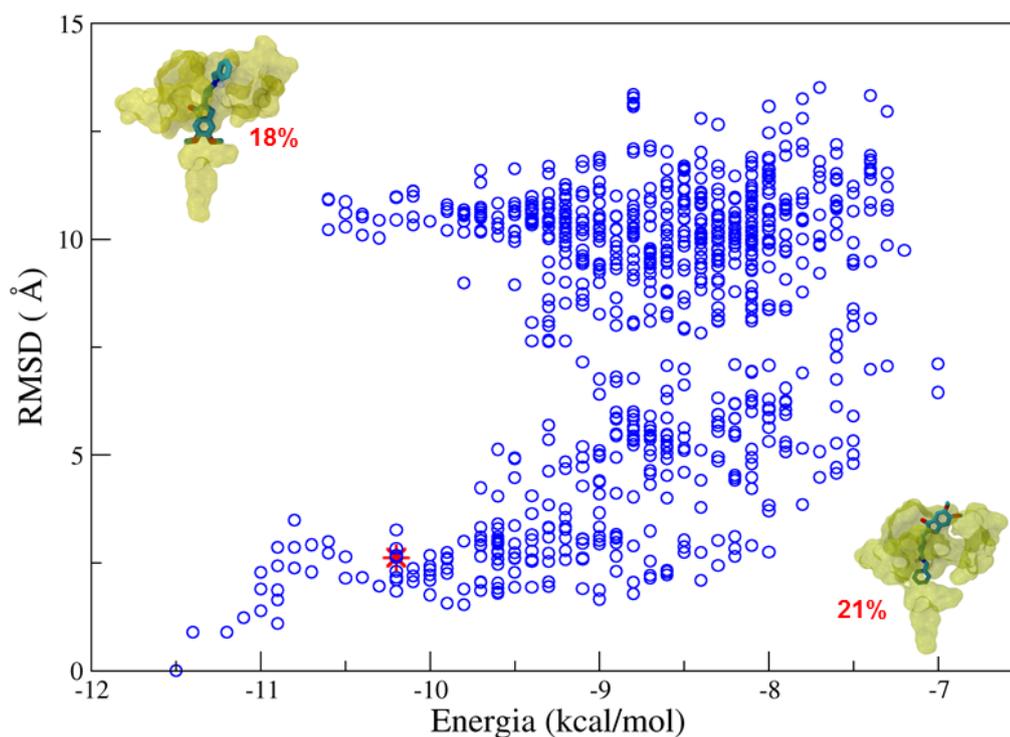


Figura 46 – Gráfico do RMSD versus Energia, para os resultados de *Docking* do Donepezil, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas pelo Donepezil.

A análise baseada no RMSD e energia para as soluções do *Docking* do ligante LDT160, mostrada no gráfico da Figura 47, evidencia a separação entre as soluções de *Docking* em duas populações. Como na análise para o Donepezil a estrutura de referência foi aquela de menor energia, e todas as outras são comparadas por meio do RMSD. Para esse ligante, diferentemente do Donepezil, a solução de menor energia encontrava-se com a “parte da amina protonada voltado para a *main door* e representa a população na parte inferior do gráfico 47. A população na parte superior do gráfico, diz respeito aquela em que a amina protonada encontra-se no fundo da Gorge e substancialmente contém o maior número de soluções.

Obedecendo aos critérios de seleção determinados para o Donepezil, para este e para os demais ligantes as duas populações serão contabilizadas com uma diferença de energia de 1,5 kcal/mol a menos daquela encontrada para a solução de menor energia. E para o ligante LDT160 a solução de menor energia no gráfico 47 encontra-se em cerca de -9,4 kcal/mol, sendo assim, próximas análises serão realizadas para populações na

parte superior e inferior do gráfico que tenham energia de ligação igual ou menor que -8.0 kcal/mol. De maneira que soluções da população com a amina protonada no fundo da Gorge equivale a 18 % do número total de soluções, e a população com a amina protonada na *main door* representa 6% do número total de soluções.

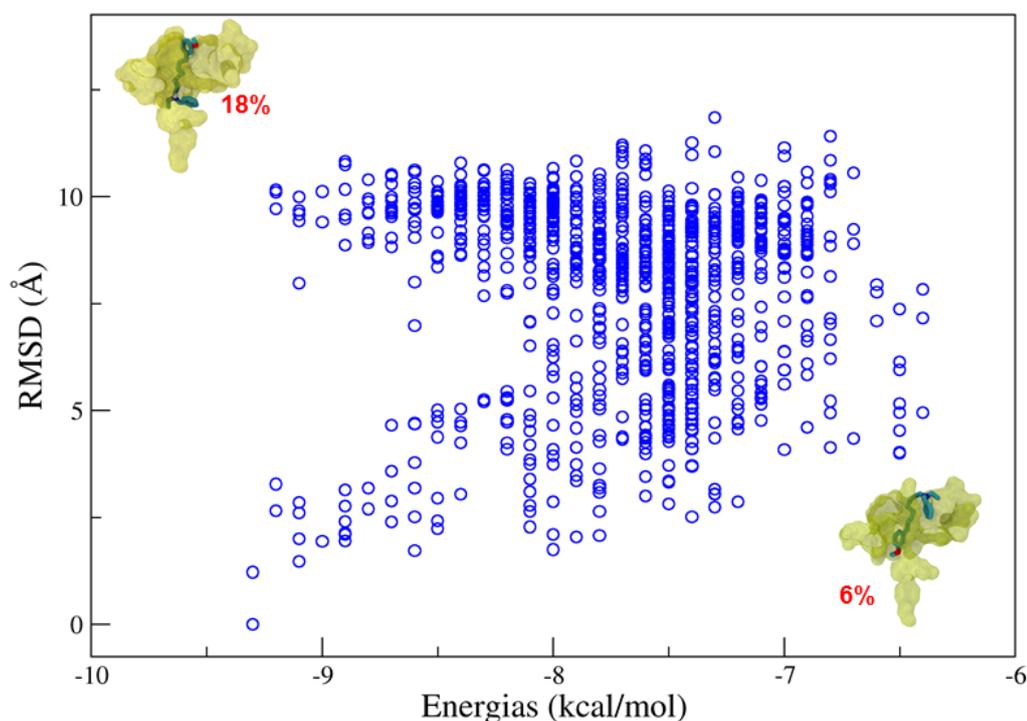


Figura 47 – Gráfico do RMSD versus Energia, para os resultados de *Docking* do ligante LDT160, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas por esse ligante.

As diferentes populações que o *Docking* encontrou para o ligante LDT161 são apresentadas no gráfico da Figura 48. A população de soluções localizada na parte superior do gráfico representa as soluções onde a amina protonada encontra-se na região da *main door*, e a parte inferior diz respeito a população onde a amina protonada está próxima ao fundo da Gorge. No gráfico é fácil observar que a maior parte das soluções pertencem a população para a amina no fundo da Gorge, apesar de existir dois modos de ligação bem distintos, esse é mais plausível para esse ligante. Para as investigações futuras dessas populações, como a solução de menor energia para essa população tem energia de ligação igual a -9,4 kcal/mol, soluções que representarão ambas as populações são aquelas com energia de ligação menor ou igual a -8 kcal/mol, assim a população para a amina protonada

no fundo da Gorge do gráfico equivale a 20% do número total de soluções e a população com a amina protonada na *main door* corresponde a 4,2% do total de soluções.

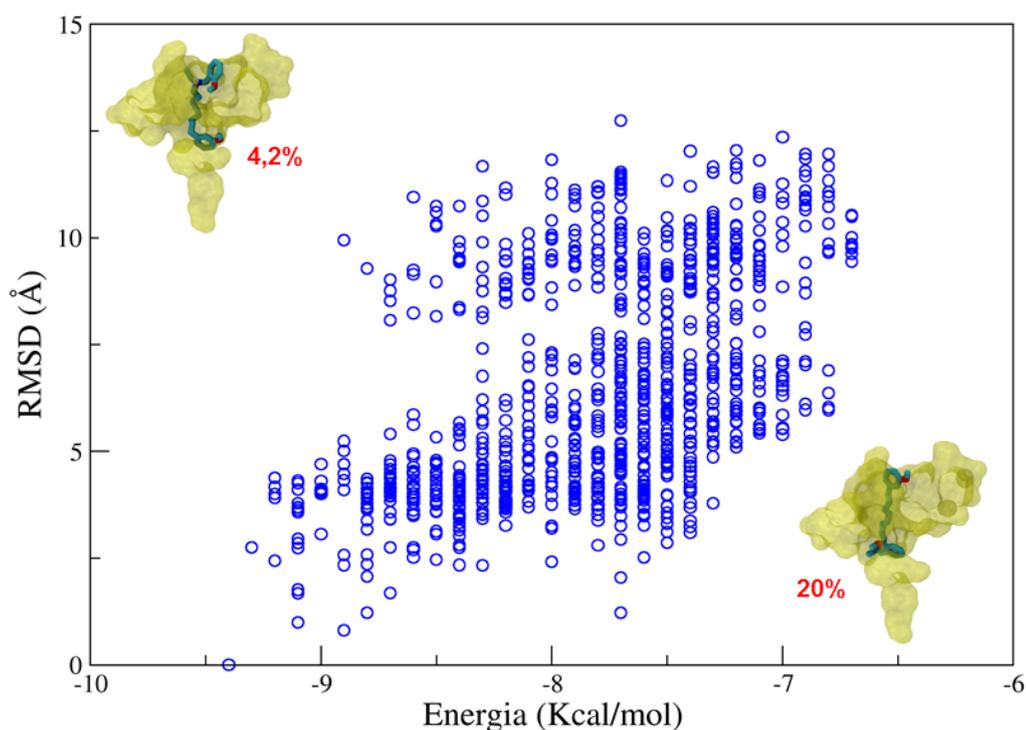


Figura 48 – Gráfico do RMSD versus Energia, para os resultados de *Docking* do ligante LDT161, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas por esse ligante.

Para as soluções de *Docking* do ligante LDT167, o gráfico da Figura 49 mostra a disposição das 960 soluções, como resposta da discriminação por meio do RMSD e energia para duas populações de modos de ligação diferentes, com a parte superior do gráfico correspondendo a população com a amina protonada localizada na região da *main door* e soluções com a amina protonada no fundo da Gorge pertencentes a população da parte inferior do gráfico. As duas populações tem um comportamento muito próximo as do ligante LDT161, contudo o número de soluções em que a amina protonada encontra-se mais próxima a *main door* é maior para esse ligante, concordando com as observações feitas para esses ligantes na análise para a nuvem de soluções. A solução de menor energia para esse ligante possui energia de ligação igual a -9,5 kcal/mol e, portanto, soluções que possuam energia de ligação igual ou menor que -8,0 kcal/mol para os dois modos de ligação.

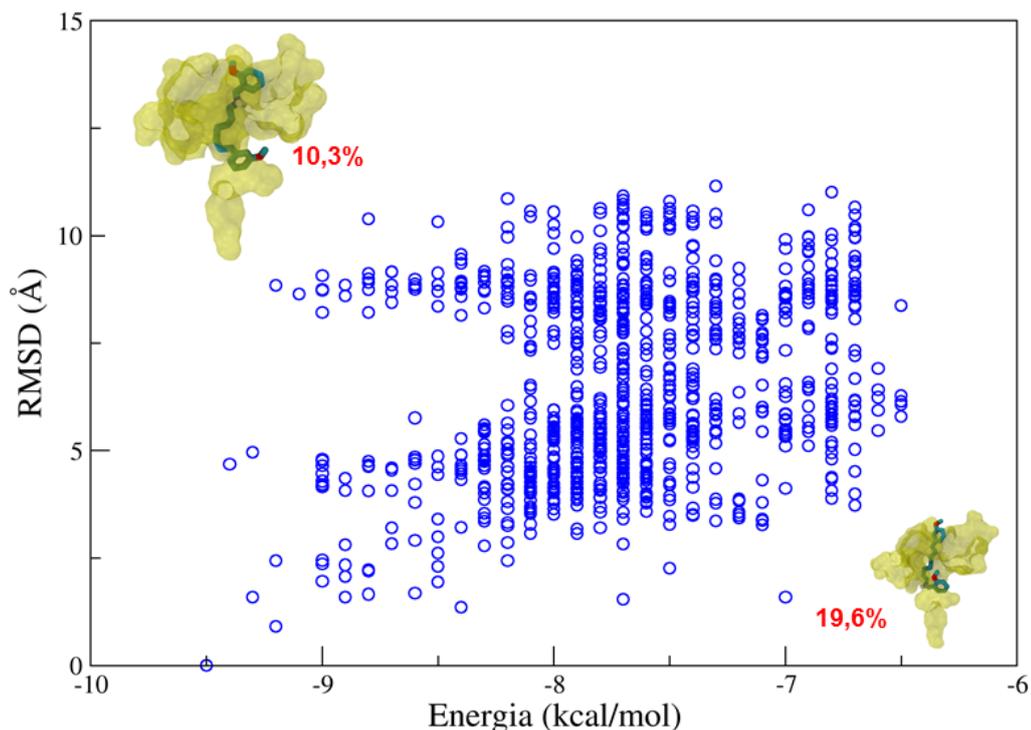


Figura 49 – Gráfico do RMSD versus Energia, para os resultados de *Docking* do ligante LDT167, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas por esse ligante.

O ligante LDT169 também apresenta duas orientações preferenciais e no gráfico da Figura 50 as duas populações, grupo amina protonado próximo a *main door* concentrando-se na parte superior do gráfico e grupo amina protonado no fundo da Gorge reunidas na parte inferior do gráfico, são expostas. A quantidade de soluções encontradas com a amina protonada voltada para o fundo da Gorge é bem maior do que o número de soluções para a outra população em que a mesma está próxima a *main door*. As análises seguintes para esse ligante serão realizadas com soluções que apresentem energia de ligação igual ou menor que -9,0 kcal/mol, um pouco menos de 1,5 kcal/mol de diferença em relação a solução com menor energia de ligação que é de -10,3 kcal/mol. Com base nessas condições, a população correspondente a aquela em que a amina protonada encontra-se na fundo da Gorge responde por 13% do número total de soluções, enquanto a outra população representa 6%.

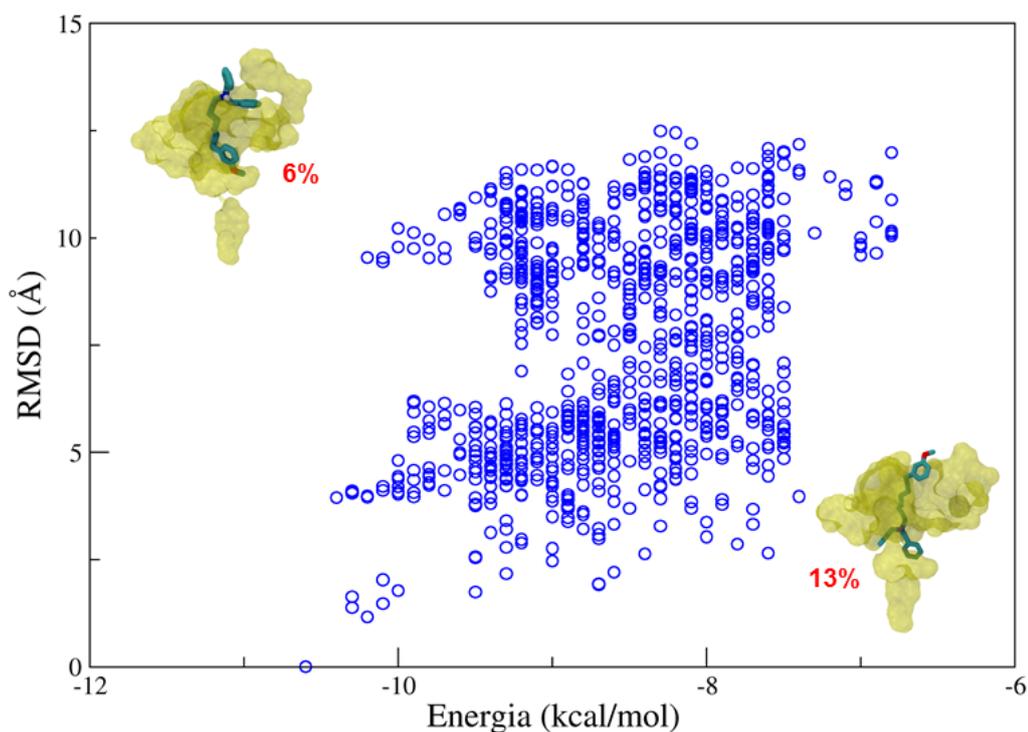


Figura 50 – Gráfico do RMSD versus Energia, para os resultados de *Docking* do ligante LDT169, explicitando as duas populações encontradas referentes as duas orientações preferencias apresentadas por esse ligante.

As orientações de melhor ajuste dos ligantes no sitio ativo da proteína foram reconhecidas e divididas em duas populações nas análises anteriores. Entender o processo de reconhecimento molecular para ligantes que encerram como características a alta flexibilidade não é uma tarefa fácil. Os resultados baseados na metodologia de *Docking*, que vem continuamente sofrendo ajustes e melhoramentos na *scoring functions* e tem retornado maior precisão nas previsões do modo de ligação do complexo receptor-ligante, mesmo que tenham sido realizados com rigor, deve ficar claro que o *Docking* é uma premissa energética para descrever as afinidades dos ligantes com a proteína.

Tendo como base os resultados para as populações dos dois modos de ligações encontrados para cada um dos ligantes e para o Donepezil, foi adotada uma estratégia fundamentada na amostragem estatística do espaço de soluções dos modos de ligação, voltados para o comportamento espacial dos grupos moleculares da proteína que interagem com os ligantes, uma vez que a análise da distribuição dos grupos moleculares da proteína satisfazem a frustração energética do ligante.

Primeiramente foram separadas os arquivos de soluções de *Docking* de acordo com as populações discriminadas na última análise, depois sobrepomos as estruturas dos ligantes nas conformações diversas dadas pelo ancoramento, de todos os ligantes e para o Donepezil com suas respectivas populações. A Figura 51 ilustra a sobreposição das conformações do ligante LDT161 (nessa figura contém somente algumas conformações do LDT161, não correspondendo a sobreposição todas as soluções de posições do ligante em uma das populações). A sobreposição dos ligantes levou a distribuição do *ensemble* da proteína em torno da sobreposição.

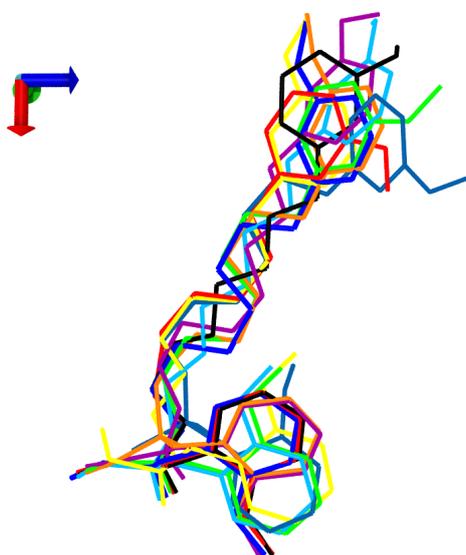


Figura 51 – Sobreposição do ligante LDT161, como exemplo de parte do procedimento adotado para a investigação, a partir da construção das densidades, de grupos moleculares da proteína favoráveis a interações com o ligante.

Dada a distribuição da proteína em torno do ligante, geramos um mapa de densidade buscando a densidade de probabilidade de encontrar um tipo particular de grupo funcional ou átomos envolvidos em interações nessa proteína descritos na literatura, próximos ao ligante. Para isso, utilizamos a ferramenta Volmap do VMD que cria um mapa de densidade atômica ponderada em cada ponto da grade, na qual cada átomo é representado por uma distribuição gaussiana normalizada de largura (desvio padrão) igual ao seu raio atômico, que são aditivamente distribuídas na grade (HUMPHREY; DALKE; SCHULTEN, 1996). Os mapas de densidade foram gerados para a média da combinação de todos os arquivos de *Docking* (coordenadas do ligante e da proteína) de cada população para determinado

ligante, onde buscamos a densidade de probabilidade de encontrar grupos aromáticos, oxigênio, nitrogênio e hidrogênio a uma distância limite de 4Å do ligante. Essas densidades médias foram visualizadas no VMD com o método de visualização padrão para o conjunto de dados volumétricos Isosurface, com o controle de representação de valor de isosuperfície (isovalue) iguais para as isosuperfícies comparadas, e o ligante representa a estrutura média da sobreposição. Esses resultados serão mostradas e discutidas a seguir.

3.2.3.2 Densidade Volumétrica da AChE para os ligantes e o Donepezil

Porções aromáticas estão presentes em grande quantidade desde a parede até a base do sítio da AChE, e estudos bioquímicos destacam uma variedade de sub-sítios de ligações hidrofóbicas e aniônicas no sítio da AChE. Devido ao fato de que o substrato natural da AChE, a ACh (acetilcolina) é carregada positivamente, uma série de estudos designam o sítio ativo como aniônico, já que o local de ligação da ACh no organismo *Electrophorus* contém de 6 a 9 cargas negativas. Posteriormente viu-se que não necessariamente existe uma localização de cargas negativas invariantes na AChE, estudos teóricos demonstraram um gradiente de potencial ao longo do comprimento da Gorge que deve mover a ACh até o sítio catalítico, e ainda que grandes contribuições são dadas por aminoácidos pertencentes a *main door* que colaboram vigorosamente com o momento de dipolo da proteína. Como existem aminoácidos na Gorge responsáveis pelo mecanismo de *gating* as interações entre eles também afetam o gradiente de potencial da Gorge. No que se refere as interações de inibidores no sítio, interações preferenciais entre nitrogênios quartenários e elétrons π dos anéis aromáticos são mais frequentes, uma vez que, devido a polarizabilidade do íon, grupos aromáticos interagem mais fortemente com amônios quartenários (DVIR et al., 2010).

Considerando a importância das interações dos tipos hidrofóbica e cátion- π dos aminoácidos que compõem a Gorge com ligantes dentro do sitio ativo, os mapas de densidade volumétrica foram realizados procurando identificar a densidade de grupos (aromáticos, oxigênio, nitrogênio e hidrogênio) em em torno do ligante, que participam dessas interações.

As análises seguintes serão feitas somente para a população em que amina protonada dos ligantes encontra-se próxima ao fundo da Gorge, uma vez que esta é a orientação preferencial do Donepezil.

O primeiro mapa diz respeito a densidade de probabilidade de encontrar os aromáticos próximos aos ligantes. A Figura 52 contém os mapas de densidade aromática para as moléculas LDT160, LDT161, LDT167, LDT169 e para o Donepezil, que foi incluído nas análises a nível de comparação, já que este apesar de ser estruturalmente diferente dos derivados do cardanol, é um fármaco AChEI e o comportamento da proteína em relação a ele pode nos guiar com relação às interações, em razão da AChE ser seletiva.

A densidade de aromáticos em torno do Donepezil está basicamente concentrada na piperidina, o que mostra que a ferramenta utilizada está reproduzindo corretamente as interações para esse ligante, logo que é sabido que o anel benzênico perto da piperidina realiza interações do tipo cátion- π e π - π e hidrofóbicas com o Trp117 da AChE. Para o ligante LDT160 a densidade de aromáticos é acumulada na região do grupo fenol, mostrando que esse grupo pode realizar interações hidrofóbicas ou aromáticas com aminoácidos da proteína, mas que no benzeno praticamente não existe densidade aromática.

Dentre os ligantes, a maior densidade de aromático foi encontrada em volta do LDT161, que de certa maneira conserva a densidade do fenol encontrada para o LDT160. Além disso, próximo à amina tem uma densidade parecida com a que encontramos para o Donepezil, e a densidade de aromáticos no fenol próximo à amina é muito similar à encontrada para o LDT167 na mesma porção do ligante. O LDT169 é o que mais se diferencia para densidades aromáticas da figura 52, próximo ao fenol tem uma densidade parecida com a do LDT160, porém um pouco menor, e no aromático (próximo à amina) foi encontrada uma nuvem de aromáticos, mostrando que essa molécula pode realizar interações aromáticas ou hidrofóbicas nessa região. Com exceção da molécula LDT161, todas apresentam densidade nesse aromático.

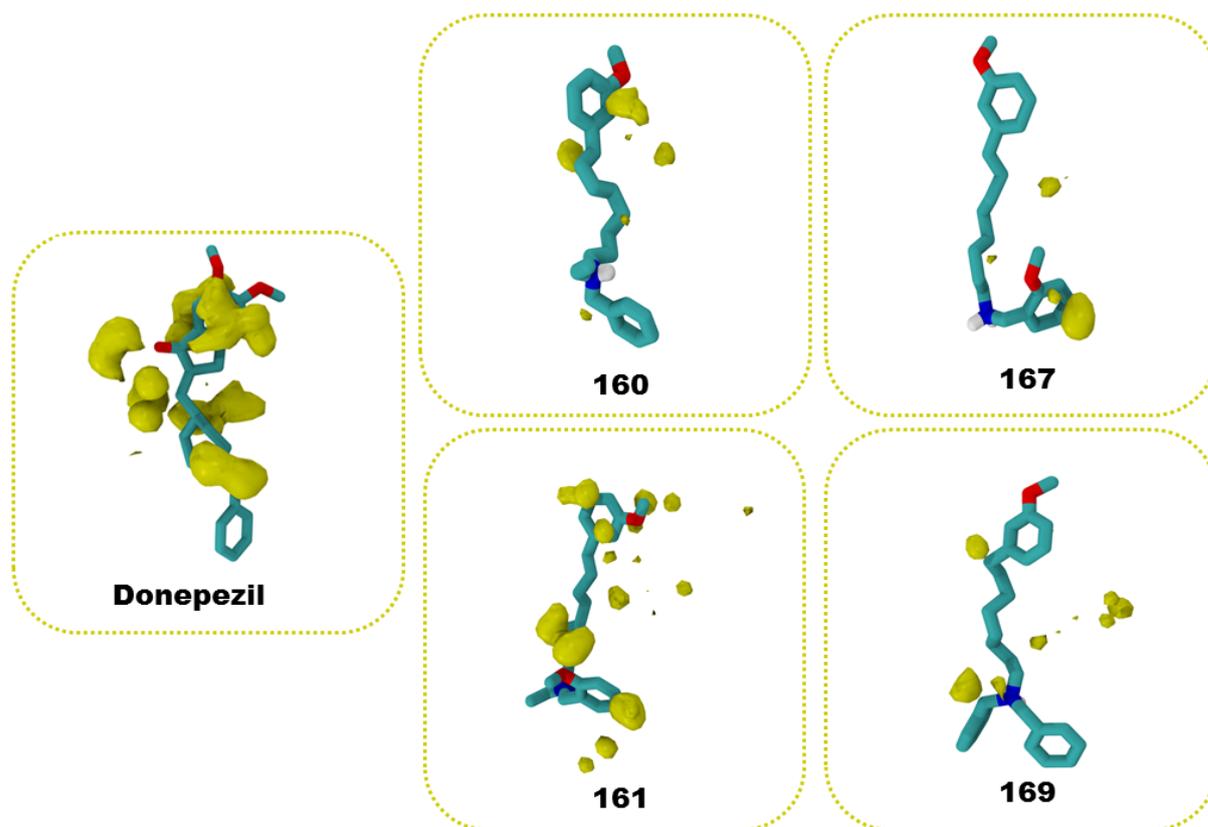


Figura 52 – Densidade de probabilidade de encontrar resíduos da proteína que contenham partes aromáticas próximo dos ligantes LDT160, LDT161, LDT167, LDT169 e para o Donepezil, para a primeira população de soluções, representada pela isosuperfície amarela.

Identificar densidades de hidrogênio é muito importante, já que os ligantes, assim como o Donepezil, possuem grupos funcionais deceptor (heteroátomo rico em elétrons- N ou O) e doador (heteroátomo deficiente em elétrons- N-H ou O-H) de ligação de hidrogênio. Mais especificamente, os grupos funcionais aceptores de ligação de hidrogênio que estes compostos apresentam são os anéis aromáticos e as regiões hidrofóbicas que podem atrair densidades de hidrogênio por meio, por exemplo, de forças de dispersão; já as hidroxilas são doadores de ligação de hidrogênio que podem melhorar a interação no receptor.

Na Figura 53, para o Donepezil, vemos que há densidade de hidrogênio mais concentrada no grupo doador O, no anel aromático e piperidina e o anel benzênico abaixo da amina não acumula densidade de hidrogênio. Os mapas de densidade de hidrogênio para a LDT160 e LDT161 são muito parecidos, para a região do ligante em que o aromático se encontra distante da amina, apresentando densidades similares; contudo a densidade no

fenol próximo a amina, no composto LDT161 encerra uma superfície de densidade maior que o benzeno na LDT160, indicando que a hidroxila é importante para interações com a proteína das quais o hidrogênio faça parte e nesse composto a densidade perto da amina é parecida com aquela para a amina no Donepezil. O composto LDT167, apesar de se diferenciar dos compostos LDT160, LDT161 por não exibir densidade próxima ao fenol, se assemelha e a LDT161 quando retrata densidade na amina e no fenol próximo a esta. O LDT169 com relação as densidades de hidrogênio é comparável somente ao LDT167, ainda assim apresenta pouca densidade na amina e no benzeno, apontando que a diferença na estrutura, para esse composto em relação as demais, fez muita diferença nas concentrações de grupos funcionais da AChE em que o hidrogênio está presente, em torno desse ligante.

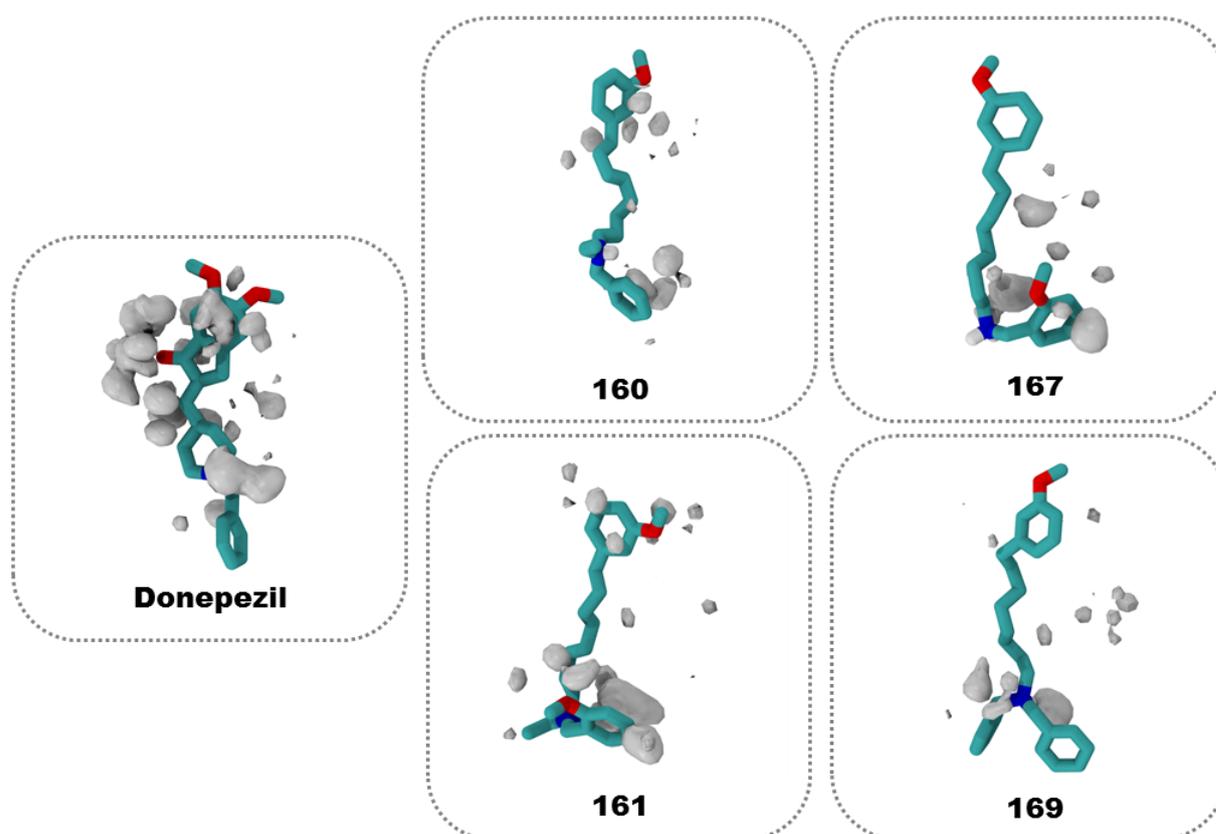


Figura 53 – Densidade de probabilidade de encontrar átomos de hidrogênio da proteína próximo dos ligantes LDT160, LDT161, LDT167, LDT169 e para o Donepezil, representada pela isosuperfície cinza.

Interações intermoleculares envolvem além das interações descritas nas análises anteriores para as densidade de hidrogênio e de aromáticos, interações entre grupos de cargas opostas, do tipo dipolo-dipolo, íon-dipolo, cátion- π e íon-íon; e também interações

iônicas são muito importantes quando um fármaco entra no sítio ativo, como previamente dito, tudo indica que o substrato natural ACh (que tem uma amina protonada) seja fortemente levado ao sítio catalítico da AChE por um forte gradiente eletrostático existente na Gorge. Sendo assim, as próximas densidades serão feitas para oxigênios e nitrogênios pertencentes a aminoácidos da proteína que se distribuem em torno dos ligantes.

As densidade de nitrogênio foram construídas para os ligantes LDT160, LDT161, LDT167, LDT169 e para o Donepezil como apresentado na Figura 54. Os resultados obtidos para as densidades de nitrogênio em torno do Donepezil deixam claro que grupos doadores de ligação de hidrogênio, piperidina e hidroxila, e o acceptor de ligação de hidrogênio ($=O$) atraem densidade de nitrogênio.

A distribuição dessas superfícies de densidade em torno do LDT161 estão de acordo com o resultado para o Donepezil, e se assemelha ao do LDT167, com relação a localização dessas densidades em partes específicas da estrutura molecular dos ligantes, no entanto, o LDT167 possui mais superfícies de densidade do que o LDT161. O LDT169, também apresenta uma grande densidade basicamente na amina e outra menor na hidroxila do fenol, e o LDT160 é aquele que apresenta disparidade em relação aos demais compostos, no que diz respeito a distribuição de densidade de nitrogênio próximo a amina, contudo, conserva a densidade perto do fenol. Nessa análise, os compostos que de maneira similar possuem distribuição de densidade de nitrogênio ao seu redor são LDT161 e LDT167.

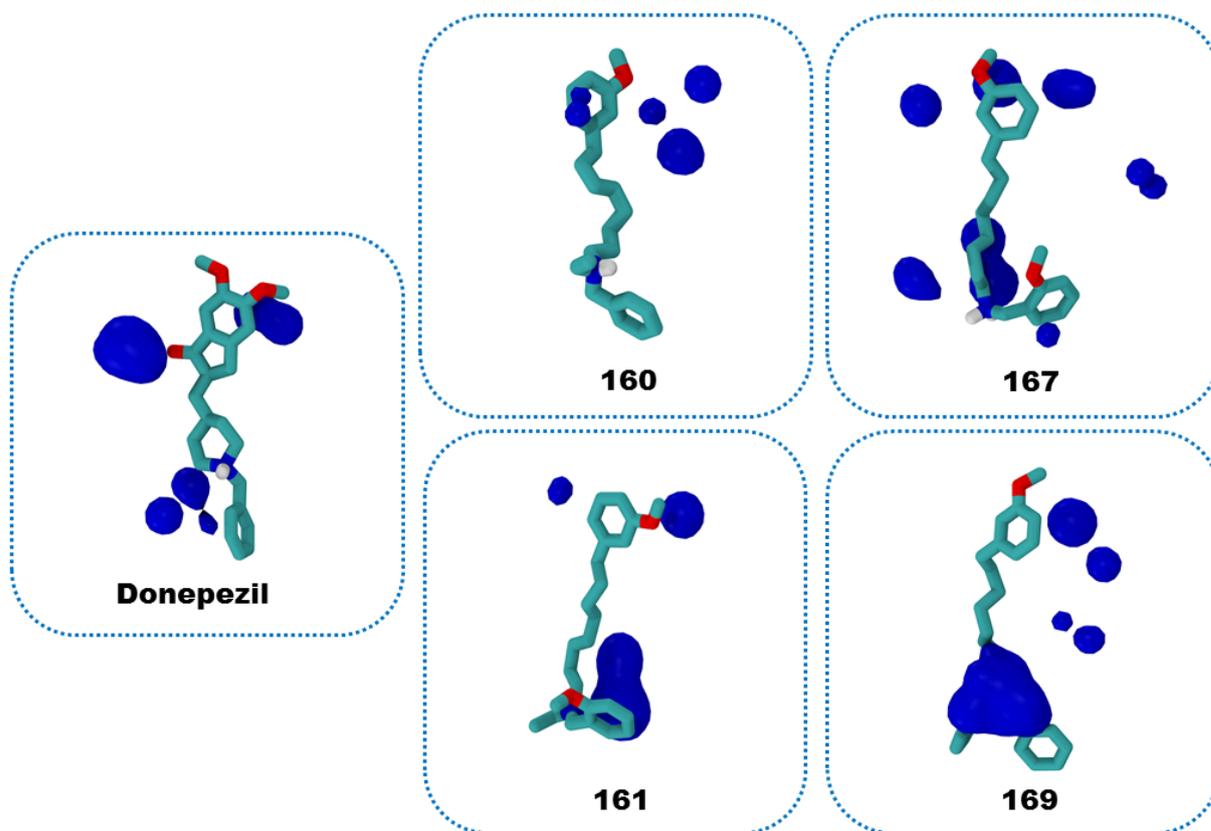


Figura 54 – Densidade de probabilidade de encontrar átomos de nitrogênio da proteína próximo dos ligantes LDT160, LDT161, LDT167, LDT169 e para o Donepezil, representada pela isosuperfície azul.

As superfícies de distribuições de densidades de oxigênio em torno dos ligantes são apresentadas na Figura 55. A densidade de oxigênio para o Donepezil localiza-se sobretudo na região da piperidina mostrando a identidade da densidade com doadores de ligação de hidrogênio e ou grupos funcionais carregados positivamente. Nos compostos, as densidades de oxigênio perto do grupo amina foram comum a todos, mas com superfícies diferentes, para o LDT167 a probabilidade de encontrar oxigênios próximo ao grupo amina é maior que para os demais e ele ainda apresenta densidade no grupo fenol.

O composto LDT161 possui densidade na amina e no fenol nas duas extremidades de sua estrutura. O LDT160 apresenta densidade próxima ao grupo amina, no entanto é a menor superfície de densidade para esse grupo funcional comparada com os demais compostos, e se comparado estruturalmente com o LDT161, pode-se dizer que a hidroxila presente no LDT161 melhorou a identificação desse composto por oxigênios presentes na proteína. Já o composto LDT169 exibe densidade de oxigênio na região da amina, que

é comum aos outros compostos, e no fenol, mas para esse composto próximo ao fenol a densidade é muito maior que para os demais. Dentre os compostos aqueles que mais se aproximaram com relação as densidades de oxigênio foram LDT161 e LDT167.

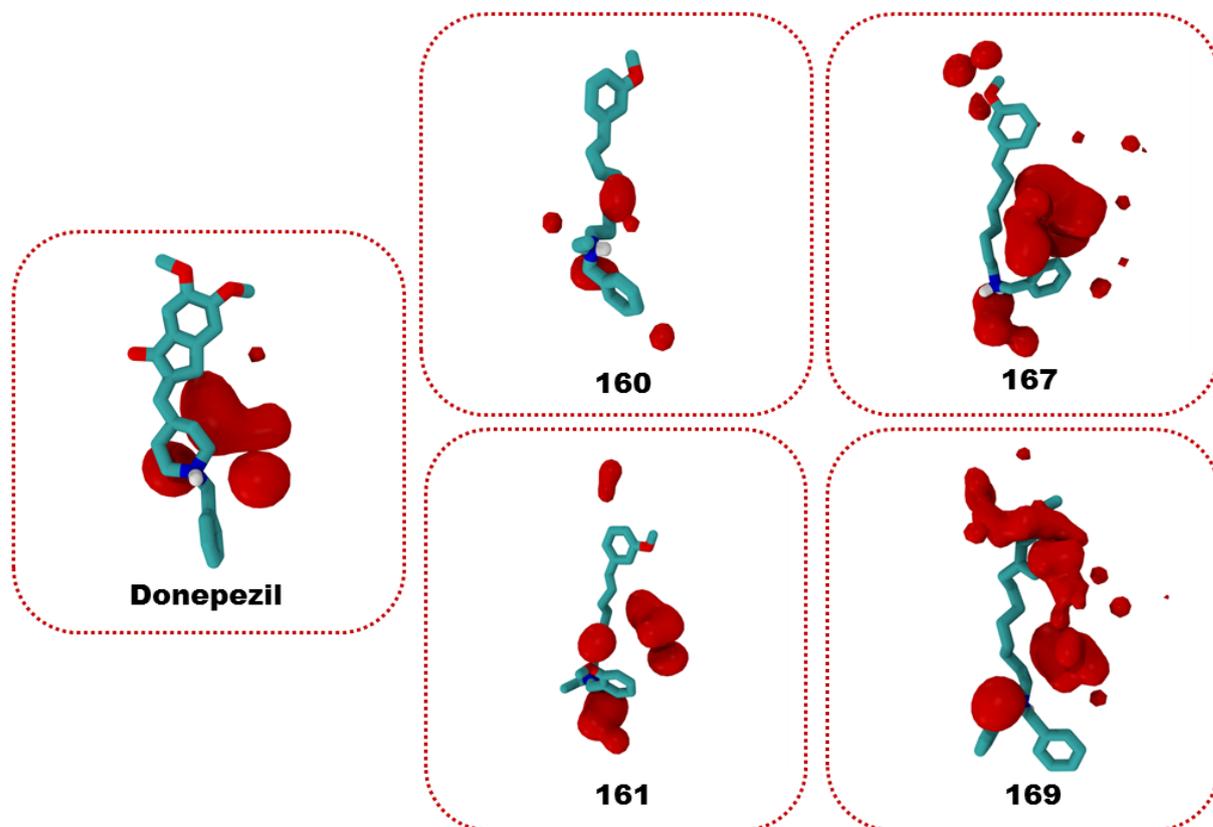


Figura 55 – Densidade de probabilidade de encontrar átomos de oxigênio da proteína próximo dos ligantes LDT160, LDT161, LDT167, LDT169 e para o Donepezil, representada pela isosuperfície vermelha.

A união das informações obtidas a partir das densidades dos grupos de oxigênio, nitrogênio, aromáticos e hidrogênio, para a molécula LDT161 indica que esta possui grande afinidade com o sitio ativo da proteína. Além de encontrarmos uma densidade relevantes para todos os grupos, as densidades em torno desse ligante mostram uma correspondência espacial com aquelas encontradas para o Donepezil.

Uma equivalência pode ser notada entre as densidades e a nuvem de soluções de *Docking* para o LDT161 representada na Figura 43. Na região da *main door* a nuvem é praticamente vermelha, enquanto na região do fundo da Gorge, a cor que mais prevalece é azul, especificando que a na maior parte das soluções encontradas, a LDT161 encontra-se com a amina ancorada no fundo da Gorge e o fenol interage com o PAS.

Em relação aos mapas de densidade do aromático, hidrogênio e nitrogênio para o ligante LDT167 mostram uma deficiência de superfície de densidade nas mediações do grupo fenol da molécula. Uma comparação entre as densidades e a nuvem de soluções de *Docking*, Figura 44, pode ser usada como alternativa para esclarecer a ausência dessas densidades.

Na nuvem de soluções foi encontrada uma mancha azul mais intensa próxima a região da *main door* da Gorge, essa mancha refere-se a posições assumidas pelo nitrogênio nas conformações da molécula. Essa concentração de nitrogênio próximo a *main door*, deve-se não somente pelo fato do *Docking* ter encontrado soluções em que o fenol ligava-se no fundo da Gorge e o nitrogênio localizado no PAS, mas também pelo fato de que, para uma grande parte das soluções, a LDT167 estava ancorava-se somente no PAS.

Considerando a alta flexibilidade desse ligante, as soluções que estavam somente no PAS apresentaram conformações variadas, e essa molécula foi encontrada "dobrada" (os seus dois fenóis empilhados) de acordo com a Figura 56 Uma vez que testes experimentais revelaram-a como um potencial inibidor da AChE, para estes estudos realizados até o momento, uma provável sugestão para a interpretação das interações com o receptor, seria que essa molécula inibe a ação da AChE, preferencialmente, por interações com os aminoácidos do PAS.

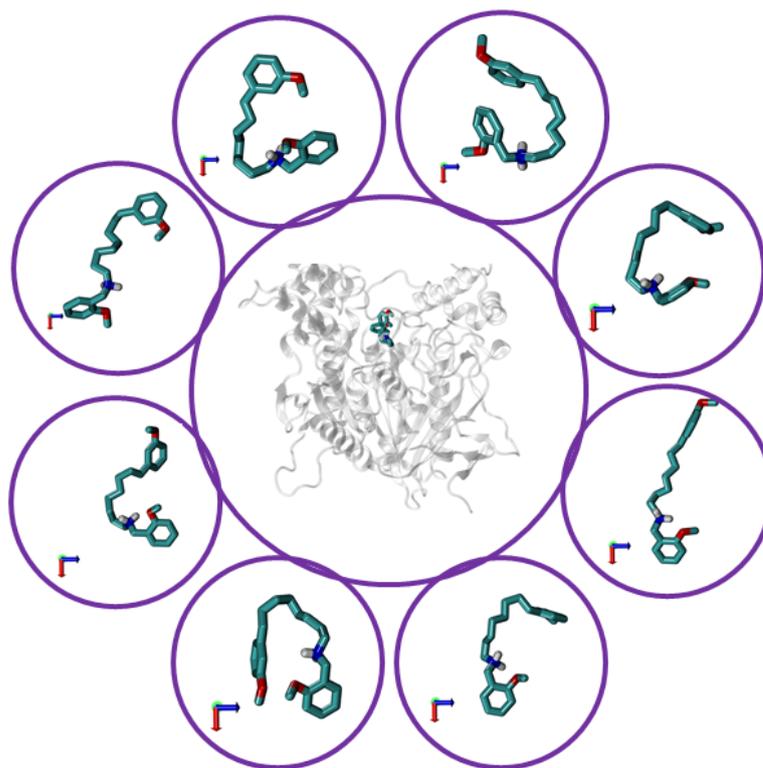


Figura 56 – Conformações assumidas pela molécula LDT167 nas soluções de *Docking*. No centro, a proteína mostrando a orientação preferencial dessa molécula, grupo amina voltado para o fundo da Gorge, e em volta, as posições encontradas com o sistema no mesmo plano.

A molécula LDT160, no que diz respeito as conformações das soluções de *Docking*, se comporta de maneira parecida com a LDT167. Na Figura 42 é possível percebermos a mesma característica da nuvem para a molécula LDT167, porém a intensidade de manchas azuis é menor. Observando as conformações dessa molécula, existem soluções em que esta molécula também encontra-se "dobrada", localizando-se totalmente na *main door*, mais em menor número do que a molécula LDT167.

Já a molécula LDT169, em todas as soluções de *Docking* desse *ensemble*, encontramos a molécula "estendida" ao longo da Gorge. E portanto, as manchas azuis que aparecem na nuvem de soluções, na Figura 45, diz respeito a outra orientação que o *Docking* encontrou.

De modo geral, as densidades encontradas para as moléculas LDT161 e LDT167 sinalizam que o comportamento da proteína em relação à molécula LDT161 é diferente do comportamento da mesma para LDT167, no entanto, experimentalmente foi

verificado que elas possuem % inibitório muito parecidos, LDT161(%inibitório=91.6) e LDT167(%inibitório=95.5). Assim sendo, o que faz com que a molécula LDT161 ancore no fundo da Gorge e a molécula LDT167 interaja preferencialmente no PAS?

No máximo grau de detalhamento que as metodologias adotadas e as análises realizadas possibilitaram compreender o sistema, a comparação entre as densidades, as nuvens de *Docking* e os MEPs apontam que a diferença na distribuição de potencial eletrostático positivo, que é maior para a LDT161 do que para LDT167, pode influenciar na afinidade do ligante com o sitio ativo do receptor.

Como a Gorge é considerada um ambiente aniônico, moléculas como a LDT161, supostamente chegam ao fundo da Gorge atraídas por meio de um gradiente potencial negativo. Com base nos resultados pode-se ainda supor que, se por um lado, moléculas com maiores superfícies de potencial eletrostático positivo interagem mais facilmente com o sítio, moléculas em que essa superfície positiva são mais sucintas podem interagir basicamente com o PAS e inibir, ainda assim, a ação da AChE.

4 CONCLUSÕES

Neste trabalho realizamos o estudo para 26 moléculas candidatas a inibidores da acetilcolinesterase, derivadas do cardanol, divididas em grupos conformes suas semelhanças estruturais. Determinamos um conjunto de descritores moleculares com capacidade de predição de atividade biológica que diferencie moléculas com grande potencial de atividade na biofase, de moléculas que não apresente esse resultado. Para o cálculo das propriedades eletrônicas e de geometria realizamos o cálculo mecânico quântico de otimização de geometria bem como a busca de propriedades SAR, baseadas na estrutura espacial das moléculas que para o nosso estudo foi considerado a série congênere (similaridade estrutural).

A análise de componentes principais sinalizou a predisposição de correlação entre moléculas congêneres, cuja atividade biológica anticolinesterásica, se aproximasse. Para o primeiro grupo estudado ficou clara a importância do grupo metileno para a atividade, sua substituição refletiu nos valores da energia dos orbitais de fronteira. Ainda no que se refere ao primeiro grupo de moléculas, a análise de componentes principais obteve resultados, visualizados bidimensionalmente através do gráfico de PCs, que concordam com os resultados experimentais. Essa concordância entre resultados experimentais e teóricos obtidos por meio da PCA pode ser verificado para moléculas que possuem dados experimentais até então. Para moléculas que ainda não possuem testes experimentais, presentes em todas as análises, este estudo revela tendências destas quando encontram-se próximas ou afastadas de moléculas que possuem informações sobre atividade biológica. A PCA apontou, por meio das correlações encontradas, a possibilidade de semelhanças entre valores de percentual de inibição e IC_{50} , de moléculas que participaram da correlação, como no caso das moléculas LDT141(2) e LDT142(3) no primeiro grupo (que possuem resultados experimentais semelhantes).

Para o grupo de derivados piperazínicos a análise de PCA não contribuiu na identificação e discriminação de descritores relevantes para a atividade. Essa análise não deixou evidente a possível correlação entre compostos, mesmo entre aqueles que não apresentam bons dados de atividade.

Verificamos a importância de descritores que correspondem principalmente as energias dos orbitais de fronteira HOMO, HOMO-1, LUMO e LUMO+1, principalmente

a energia do LUMO que esteve presente em todos os padrões selecionados. Este orbital se concentra na parte farmacofórica das moléculas e possíveis interações aconteceram preferencialmente com grupos farmacofóricos localizados na parte de concentração desse orbital e aminoácidos presentes na enzima. Os mapas de potencial eletrostático das moléculas expressaram o perfil de distribuição de cargas, colaborando no reconhecimento de prováveis interações ligante-receptor.

No estudo de modelagem molecular para a predição do reconhecimento molecular do complexo (receptor-ligante), duas metodologias - dinâmica molecular considerando campos de força clássicos e *Docking* molecular - foram empregadas na investigação do mecanismo dinâmico da proteína AChE (código PDB:4EY7) e modos de interação preferenciais com as moléculas pertencentes ao grupo dos derivados benzilamínicos do cardanol. Em sua totalidade, os resultados gerados nos permitem extrair informações a respeito das interações dessas moléculas com o sítio ativo da AChE.

Primeiramente foram identificados os perfis para as posições os resíduos da Gorge assumiam durante a dinâmica. Esses perfis configuram os quatro estados encontrados para a proteína, que foram nomeados de *Closed*, *OpenClosed*, *ClosedOpen* e *Open*, determinados através dos raios das esferas variáveis ao longo do eixo de simetria (Z) que atravessa a Gorge. A identificação do perfil *Open* foi indispensável para seleção de conformações da proteína para o estudo de ancoramento.

O *Docking* das moléculas LDT160, LDT161, LDT167, LDT169 e o Donepezil (adicionado no estudo como parâmetro para comparação das análises) no *ensemble* de conformações da proteína, no estado *Open*, revelaram que todos os ligantes, inclusive o Donepezil, assumem duas orientações preferenciais. Devido ao fato de que, estudos anteriores garantem que o Donepezil tem uma orientação de interação preferencial (FANG et al., 2014), as análises para os derivados benzilamínicos foram realizadas para um variedade de conformações desse ligante, cuja orientação condiz com a do Donepezil.

Dada a relevância da compreensão dos modos de interação entre ligante e receptor, como forma de minimizar o erro associado as predições de energia livre dadas pelo processo de *Docking*, foi realizado o estudo estatístico do comportamento do ambiente da proteína em relação ao ligante, com base na densidade de probabilidade de encontrar átomos ou grupos funcionais da proteína em torno do ligante. Esse estudo foi muito satisfatório, uma vez que conseguiu reproduzir com fidelidade as interações para o Donepezil, o que guiou as análises dos derivados do cardanol.

Nos resultados obtidos para as densidades, podemos concluir que o composto LDT161 apresenta densidades de aromáticos que remetem as interações realizadas pelo Donepezil com o TRP117, no sítio Aromático da Gorge, e densidades aromáticas no fenol do LDT161 indicam ainda, que este pode interagir com aminoácidos no sítio aniônico periférico, apontando a possibilidade desta molécula interagir, simultaneamente, no sítio aromático e no PAS. As densidades para a LDT161, no geral, sinalizam que esta molécula tem muita afinidade com a Gorge.

Os resultados encontrados para as densidades do ligante LDT167, principalmente na região do fenol, mostrou grande diferença das demais moléculas. Essa diferença retrata as conformações assumidas por essa molécula nas soluções de *Docking*. Na média, essa molécula interage unicamente com o PAS, na região da *main door*, e em muitas soluções, encontra-se "dobrada". Algo parecido acontece para a molécula LDT160, no entanto, essa molécula apresenta uma medida próxima entre conformações "dobrada" e "estendida".

Uma relação entre as densidades encontradas, a nuvem de soluções de *Docking* e os MEPs para estas moléculas, foi estabelecida para compreender os diferentes resultados para os testes experimentais de % de inibição dessas moléculas. Esta relação, leva a suposição de que moléculas que apresentam maior superfície de densidade eletrostática positiva, é possivelmente mais capaz de ancorar no fundo da Gorge, apresentando maior afinidade com o sítio. Moléculas, cuja superfície de densidade eletrostática positiva é mais uniforme, e portanto menos intensa, parece não ser capaz de chegar ao fundo da Gorge, mas pode interagir com o PAS, e apresentar um bom % inibitório.

Como conclusão, o presente trabalho mostra que o mapeamento do ligante através do estudo quântico de estrutura eletrônica, juntamente com a modelagem molecular clássica, podem desempenhar um papel relevante para a busca e planejamento de novos fármacos com alto poder de combate a DA. Dessa maneira, os estudos realizados neste trabalho para os derivados benzilamínicos do cardanol, revelam aspectos preditivos inéditos para os modos de interação destes com o receptor AChE (código PDB:4EY7).

5 PERSPECTIVAS

Apesar deste estudo realizar uma caracterização dos estados da proteína e a predição estatística, por meio do *Docking* molecular, das interações dos derivados benzilamínicos do cardanol com a AChE humana (código PDB:4EY7), existem uma série de aspectos que envolvem a compreensão dos modos de ligação que ainda não foram explorados, ou que não podem ser observados com essas abordagens. Nesse contexto, para o avanço dos estudos utilizando esse e outros métodos, os resultados obtidos oferecem uma base fundamental, evidenciando que nossas pesquisas são necessárias para o esclarecimento dessas e de outras questões que envolvam a interação desses ligantes com a Gorge.

5.1 *Docking* molecular na presença de água

O *Docking* molecular tem sido frequentemente utilizado como uma ferramenta valiosa para estudos em química medicinal e devido ao sucesso dos resultados, tornando-se significativa para a descoberta de novas drogas. Ainda assim, a reprodução de conformações bioativa de ligantes na AChE aplicando programas de ancoragem molecular, mostrou tem mostrado alguns problemas em consequência de alguns fatores, como a conformação da proteína e a presença de água (SENIYA; KHAN; UCHADIA, 2014).

A conformação da proteína já foi devidamente avaliada neste trabalho, resultando no *ensemble* de conformações *Open*. A verificação da interação dos ligantes, através da metodologia de *Docking*, considerando a presença de água no sítio ativo, é uma das perspectivas de continuidade desse estudo.

Durante a amostragem das trajetórias nos cálculos de dinâmica da AChE, percebemos que no interior da Gorge permaneciam em média três moléculas de água para conformações da proteína no estado *Open*, como pode ser visto na Figura 57 O *Docking* dos ligantes no *ensemble* de conformações *Open* na presença de água, pode nos informar se essas moléculas de água, de alguma maneira interferem, auxiliando ou dificultando as interações dos ligantes com os resíduos do sítio ativo.

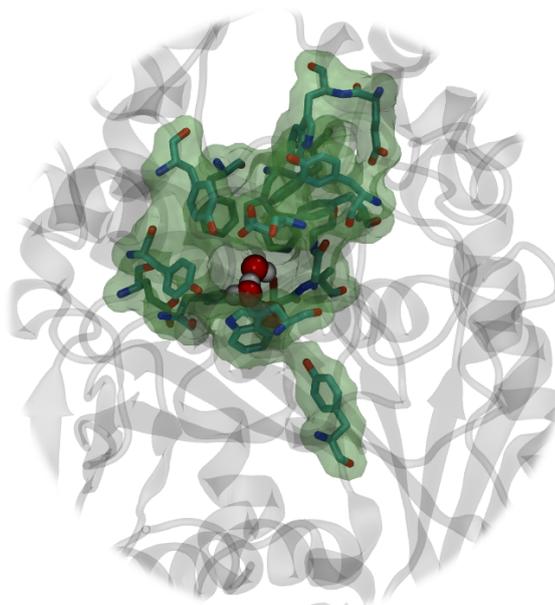


Figura 57 – Exemplo de conformação da AChE, equilibrada em solução. A Gorge é mostrada na isosuperfície verde, com seus aminoácidos. Representado como esferas de *van der Waals*, no interior da Gorge, as três moléculas de água, que permanecem na Gorge durante a dinâmica.

5.2 Cálculos com método híbrido QM/MM

As simulações de dinâmica molecular baseadas na mecânica molecular (MM) apresentam limitações ao reproduzir fenômenos, nos quais não há significativas mudanças estruturais das ligações dos átomos do sistema. Ou seja, os potenciais clássicos não conseguem descrever fenômenos ligados as modificações das distribuições eletrônicas. Assim, tratar o sistema quanticamente é necessário para calcular propriedades eletrônicas em processos reativos ou de interação, como mecanismos enzimáticos. No entanto, tratar via mecânica quântica, sistemas muito grandes embebidos em solução, é algo muito dispendioso do ponto de vista computacional.

Os métodos chamados híbridos combinam mecânica quântica e mecânica clássica para tratar sistemas de macromoléculas, buscam a modelagem de eventos eletrônicos locais em grandes sistemas. O método híbrido QM/MM (do inglês, *quantum mechanics/molecular mechanics*) tem sido muito utilizado em estudo de mecanismos enzimáticos e relatam aspectos das dinâmicas da proteína (SENN; THIEL, 2007).

A ideia básica desse método é descrever o local ativo (região onde ocorrem reações químicas, interações químicas ou excitações eletrônicas) por meio da mecânica quântica, com exatidão necessária, e compreender o ambiente molecular (que inclui o solvente e a macromolécula) com a mecânica molecular, isto é, no nível dos campos de força clássicos (SCHLICK, 2010).

Para compreendermos melhor as interações e os efeitos do ambiente sobre os derivados benzilamínicos do cardanol, e até mesmo o efeito desses ligantes sobre o ambiente da proteína, pretendemos realizar cálculos de QM/MM, cujo particionamento das regiões de tratamento mecânico quântico e mecânico molecular podem ser definidas de acordo com a Figura 58, que tem como exemplo uma solução de *Docking* para a molécula LDT161.

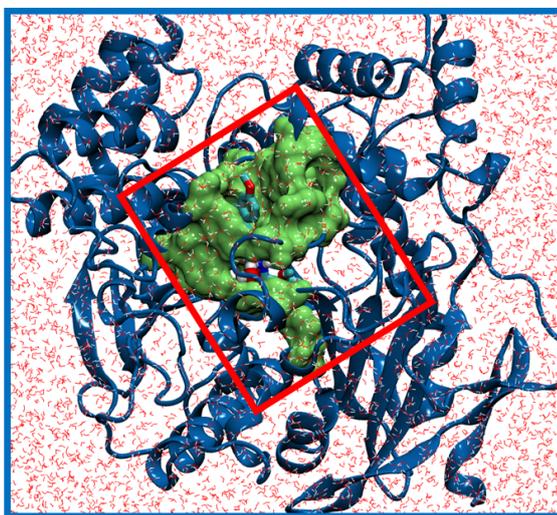


Figura 58 – Separação das regiões da proteína em solução, complexada com o ligante LDT161. A região menor, delimitada pelo retângulo vermelho é a região QM, e a região maior, delimitada pelo retângulo azul é a região MM.

Pretendemos a princípio obter os mapas de densidades eletrostáticas para a interação do complexo e confrontá-las com os mapas para os ligantes obtidos por meio dos cálculos quânticos mostrados nesse trabalho. Além disso, pode-se ainda compreender os efeitos de polarização e solvente para essas interações. Acreditamos que esse resultado terá grande contribuição nos desafios que envolvem o entendimento a cerca das modificações químicas, caracterizações bioquímicas e biofísicas que podem ser realizados nos derivados do cardanol, afim de melhor a afinidade deste com a AChE, e ainda fornecer uma rota para o desenho de novos inibidores da AChE.

REFERÊNCIAS

- ADCOCK, S. A.; MCCAMMON, J. A. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical reviews*, ACS Publications, v. 106, n. 5, p. 1589–1615, 2006.
- ALLEN, M. P. et al. Introduction to molecular dynamics simulation. *Computational soft matter: from synthetic polymers to proteins*, NIC Series, v. 23, p. 1–28, 2004.
- ANDERSEN, H. C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, Elsevier, v. 52, n. 1, p. 24–34, 1983.
- ARIEL, N. et al. The ‘aromatic patch’ of three proximal residues in the human acetylcholinesterase active centre allows for versatile interaction modes with inhibitors. *Biochemical Journal*, Portland Press Limited, v. 335, n. 1, p. 95–102, 1998.
- ARROIO, A. et al. Propriedades químico-quânticas empregadas em estudos das relações estrutura-atividade. *Química Nova*, Sociedade Brasileira de Química, v. 33, n. 3, p. 694–699, 2010.
- ASSOCIATION, A. et al. 2016 alzheimer’s disease facts and figures. *Alzheimer’s & Dementia*, Elsevier, v. 12, n. 4, p. 459–509, 2016.
- BARREIRO, E. J. et al. Molecular modeling: a tool for rational drug design in medicinal chemistry. *Química nova*, SciELO Brasil, v. 20, n. 3, p. 300–310, 1997.
- BECKE, A. D. Density-functional thermochemistry. iii. the role of exact exchange. *The Journal of chemical physics*, AIP Publishing, v. 98, n. 7, p. 5648–5652, 1993.
- BENNION, B. J. et al. A wrench in the works of human acetylcholinesterase: Soman induced conformational changes revealed by molecular dynamics simulations. *PLoS one*, Public Library of Science, v. 10, n. 4, p. e0121092, 2015.
- BERMAN, H. M. et al. The protein data bank. *Nucleic acids research*, Oxford Univ Press, v. 28, n. 1, p. 235–242, 2000.
- BHARATE, S. S.; KUMAR, V.; VISHWAKARMA, R. A. Determining partition coefficient (log p), distribution coefficient (log d) and ionization constant (pka) in early drug discovery. *Combinatorial chemistry & high throughput screening*, Bentham Science Publishers, v. 19, n. 6, p. 461–469, 2016.
- BOHLOOLI, F.; SEPEHRI, S.; RAZZAGHI-ASL, N. Response surface methodology in drug design: A case study on docking analysis of a potent antifungal fluconazole. *Computational Biology and Chemistry*, Elsevier, 2017.
- BOLOGNESI, M. L. et al. Multi-target-directed drug design strategy: from a dual binding site acetylcholinesterase inhibitor to a trifunctional compound against alzheimer’s disease. *Journal of medicinal chemistry*, ACS Publications, v. 50, n. 26, p. 6446–6449, 2007.
- BROOIJMANS, N.; KUNTZ, I. D. Molecular recognition and docking algorithms. *Annual review of biophysics and biomolecular structure*, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 32, n. 1, p. 335–373, 2003.

- BROOKS, B. R. et al. Charmm: the biomolecular simulation program. *Journal of computational chemistry*, Wiley Online Library, v. 30, n. 10, p. 1545–1614, 2009.
- BROOKS, B. R. et al. Charmm: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry*, Wiley Online Library, v. 4, n. 2, p. 187–217, 1983.
- BUI, J. M.; HENCHMAN, R. H.; MCCAMMON, J. A. The dynamics of ligand barrier crossing inside the acetylcholinesterase gorge. *Biophysical journal*, Elsevier, v. 85, n. 4, p. 2267–2272, 2003.
- BUNKER, P.; MOSS, R. The breakdown of the born-oppenheimer approximation: the effective vibration-rotation hamiltonian for a diatomic molecule. *Molecular Physics*, Taylor & Francis, v. 33, n. 2, p. 417–424, 1977.
- BURKE, K. Perspective on density functional theory. *The Journal of chemical physics*, AIP Publishing, v. 136, n. 15, p. 150901, 2012.
- BUTINI, S. et al. Exploiting protein fluctuations at the active-site gorge of human cholinesterases: further optimization of the design strategy to develop extremely potent inhibitors. *Journal of medicinal chemistry*, ACS Publications, v. 51, n. 11, p. 3154–3170, 2008.
- CARAMELLI, P. et al. Effects of galantamine on attention and memory in alzheimer's disease measured by computerized neuropsychological tests: results of the brazilian multi-center galantamine study (gal-bra-01). *Arquivos de neuro-psiquiatria*, SciELO Brasil, v. 62, n. 2B, p. 379–384, 2004.
- CHANDLER, D. *Introduction to modern statistical mechanics*. New York, Oxford: Oxford University Press, 1987. ISBN 0-19-504276-X.
- CHEUNG, J. et al. Structures of human acetylcholinesterase in complex with pharmacologically important ligands. *Journal of medicinal chemistry*, ACS Publications, v. 55, n. 22, p. 10282–10286, 2012.
- COHEN, H. D.; ROOHTAAN, C. Electric dipole polarizability of atoms by the hartree—fock method. i. theory for closed-shell systems. *The Journal of Chemical Physics*, AIP, v. 43, n. 10, p. S34–S39, 1965.
- COSSI, M. et al. New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. *The Journal of Chemical Physics*, AIP, v. 117, n. 1, p. 43–54, 2002.
- COUTINHO, N. *Métodos de Química Teórica E Modelagem Molecular*. [S.l.]: Editora Livraria da Física, 2007. ISBN 9788588325876.
- DENNINGTON, R.; KEITH, T.; MILLAM, J. *GaussView Version 5*. 2009. Semichem Inc. Shawnee Mission KS.
- DVIR, H. et al. Acetylcholinesterase: from 3d structure to function. *Chemico-biological interactions*, Elsevier, v. 187, n. 1, p. 10–22, 2010.
- EAGGER, S. A.; LEVY, R.; SAHAKIAN, B. J. Tacrine in alzheimer's disease. *The Lancet*, Elsevier, v. 337, n. 8748, p. 989–992, 1991.

ERPENBECK, J. J.; WOOD, W. W. Molecular dynamics calculations of the hard-sphere equation of state. *Journal of statistical physics*, Springer, v. 35, n. 3-4, p. 321–340, 1984.

EWING, T. J. et al. Dock 4.0: search strategies for automated molecular docking of flexible molecule databases. *Journal of computer-aided molecular design*, Springer, v. 15, n. 5, p. 411–428, 2001.

FANG, J. et al. Inhibition of acetylcholinesterase by two genistein derivatives: kinetic analysis, molecular docking and molecular dynamics simulation. *Acta Pharmaceutica Sinica B*, Elsevier, v. 4, n. 6, p. 430–437, 2014.

FANG, L. et al. Active site gating and substrate specificity of butyrylcholinesterase and acetylcholinesterase: insights from molecular dynamics simulations. *The Journal of Physical Chemistry B*, ACS Publications, v. 115, n. 27, p. 8797–8805, 2011.

FELLER, S. E. et al. Constant pressure molecular dynamics simulation: the langevin piston method. *The Journal of chemical physics*, AIP, v. 103, n. 11, p. 4613–4621, 1995.

FLÔRES, D. G.; QUARANTINI, L. d. C. Delirium e deficiência colinérgica. 2011.

FRANCIS, P. T. et al. The cholinergic hypothesis of alzheimer's disease: a review of progress. *Journal of Neurology, Neurosurgery & Psychiatry*, BMJ Publishing Group Ltd, v. 66, n. 2, p. 137–147, 1999.

FRATERNALI, F.; GUNSTEREN, W. V. An efficient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution. *Journal of molecular biology*, Elsevier, v. 256, n. 5, p. 939–948, 1996.

FRISCH, M. et al. *Gaussian 09, Revision A. 02, Gaussian*. 2009.

FURET, P. et al. Substituent effects on acetylene stability. a comparison of sto-3g, 6-31g, 6-31g**, and 6-311g** calculations. *Canadian journal of chemistry*, NRC Research Press, v. 63, n. 11, p. 2990–2994, 1985.

GAUDREAU, F.; NAJMANOVICH, R. J. Flexaid: revisiting docking on non-native-complex structures. *Journal of chemical information and modeling*, ACS Publications, v. 55, n. 7, p. 1323–1336, 2015.

GILSON, M. et al. Open"back door" in a molecular dynamics simulation of acetylcholinesterase. *SCIENCE-NEW YORK THEN WASHINGTON-*, American Association for the Advancement of Science, p. 1276–1276, 1994.

GORA, A.; BREZOVSKY, J.; DAMBORSKY, J. Gates of enzymes. *Chemical reviews*, ACS Publications, v. 113, n. 8, p. 5871–5923, 2013.

GROSS, K. C.; SEYBOLD, P. G.; HADAD, C. M. Comparison of different atomic charge schemes for predicting pka variations in substituted anilines and phenols. *International journal of quantum chemistry*, Wiley Online Library, v. 90, n. 1, p. 445–458, 2002.

GROTH, D. et al. Principal components analysis. *Computational Toxicology: Volume II*, Springer, p. 527–547, 2013.

GRUBMÜLLER, H. et al. Generalized verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Molecular Simulation*, Taylor & Francis, v. 6, n. 1-3, p. 121–142, 1991.

- HACHISU, M. et al. Beyond the hypothesis of serum anticholinergic activity in alzheimer's disease: Acetylcholine neuronal activity modulates brain-derived neurotrophic factor production and inflammation in the brain. *Neurodegenerative Diseases*, Karger Publishers, v. 15, n. 3, p. 182–187, 2015.
- HALGREN, T. A. The representation of van der waals (vdw) interactions in molecular mechanics force fields: potential form, combination rules, and vdw parameters. *Journal of the American Chemical Society*, ACS Publications, v. 114, n. 20, p. 7827–7843, 1992.
- HAQ, Z. ul et al. Molecular dynamics simulation of axillaridine-a: A potent natural cholinesterase inhibitor. *Journal of enzyme inhibition and medicinal chemistry*, Taylor & Francis, v. 24, n. 5, p. 1101–1105, 2009.
- HOWLETT, D. R. et al. Aggregation state and neurotoxic properties of alzheimer beta-amyloid peptide. *Neurodegeneration*, Elsevier, v. 4, n. 1, p. 23–32, 1995.
- HUEY, R. et al. A semiempirical free energy force field with charge-based desolvation. *Journal of computational chemistry*, Wiley Online Library, v. 28, n. 6, p. 1145–1152, 2007.
- HUMPHREY, W.; DALKE, A.; SCHULTEN, K. Vmd: visual molecular dynamics. *Journal of molecular graphics*, Elsevier, v. 14, n. 1, p. 33–38, 1996.
- HÜNENBERGER, P. H. Thermostat algorithms for molecular dynamics simulations. In: *Advanced computer simulation*. [S.l.]: Springer, 2005. p. 105–149.
- IQBAL, K. et al. Tau in alzheimer disease and related tauopathies. *Current Alzheimer Research*, Bentham Science Publishers, v. 7, n. 8, p. 656–664, 2010.
- IRWIN, J. J. et al. Automated docking screens: a feasibility study. *Journal of medicinal chemistry*, ACS Publications, v. 52, n. 18, p. 5712–5720, 2009.
- JANN, M. W. Rivastigmine, a new-generation cholinesterase inhibitor for the treatment of alzheimer's disease. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, Wiley Online Library, v. 20, n. 1, p. 1–12, 2000.
- JONES, G. et al. Development and validation of a genetic algorithm for flexible docking. *Journal of molecular biology*, Elsevier, v. 267, n. 3, p. 727–748, 1997.
- JORGENSEN, W. L. et al. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, AIP, v. 79, n. 2, p. 926–935, 1983.
- JORGENSEN, W. L.; MAXWELL, D. S.; TIRADO-RIVES, J. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, ACS Publications, v. 118, n. 45, p. 11225–11236, 1996.
- JORGENSEN, W. L.; TIRADO-RIVES, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proceedings of the National Academy of Sciences of the United States of America*, National Acad Sciences, v. 102, n. 19, p. 6665–6670, 2005.
- JULIS, J. et al. Selective ethenolysis and oestrogenicity of compounds from cashew nut shell liquid. *Green Chemistry*, Royal Society of Chemistry, v. 16, n. 5, p. 2846–2856, 2014.

- KARELSON, M.; LOBANOV, V. S.; KATRITZKY, A. R. Quantum-chemical descriptors in qsar/qspr studies. *Chemical reviews*, ACS Publications, v. 96, n. 3, p. 1027–1044, 1996.
- KARPLUS, M.; PETSKO, G. A. et al. Molecular dynamics simulations in biology. *Nature*, v. 347, n. 6294, p. 631–639, 1990.
- KIAMETIS, A. S. et al. Acetylcholinesterase inhibitors: Modeling potential candidates. *International Journal of Quantum Chemistry*, Wiley Online Library, v. 113, n. 10, p. 1461–1466, 2013.
- KIM, D.-S. et al. Betadock: shape-priority docking method based on beta-complex. *Journal of Biomolecular Structure and Dynamics*, Taylor & Francis, v. 29, n. 1, p. 219–242, 2011.
- KLAUDA, J. B. et al. Update of the charmm all-atom additive force field for lipids: validation on six lipid types. *The journal of physical chemistry B*, ACS Publications, v. 114, n. 23, p. 7830–7843, 2010.
- KOHN, W.; SHAM, L. J. Self-consistent equations including exchange and correlation effects. *Physical review*, APS, v. 140, n. 4A, p. A1133, 1965.
- LEACH, A. *Molecular Modelling: Principles and Applications (2nd Edition)*. 2. ed. [S.l.]: Prentice Hall, 2001. Paperback. ISBN 0582382106.
- LEACH, A.; KIER, L. B. Molecular modeling: Principles and applications. *Journal of Medicinal Chemistry*, [Easton, Pa.]: American Chemical Society, [c1963-, v. 40, n. 18, p. 2969, 1997.
- LEE, C.; YANG, W.; PARR, R. G. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, APS, v. 37, n. 2, p. 785, 1988.
- LEMES, L. F. N. et al. Cardanol-derived ache inhibitors: Towards the development of dual binding derivatives for alzheimer's disease. *European journal of medicinal chemistry*, Elsevier, v. 108, p. 687–700, 2016.
- LEVINE, I. N. *Quantum chemistry*. [S.l.]: Pearson Prentice Hall Upper Saddle River, NJ, 2009.
- LOZANO, N. B. H. et al. Theoretical models for the antitrypanosomal activity of thiosemicarbazone derivatives. *International Journal of Quantum Chemistry*, Wiley Online Library, v. 112, n. 20, p. 3364–3370, 2012.
- MAZZETTO, S. E.; LOMONACO, D.; MELE, G. Cashew nut oil: opportunities and challenges in the context of sustainable industrial development. *Quimica Nova*, SciELO Brasil, v. 32, n. 3, p. 732–741, 2009.
- MENG, X.-Y. et al. Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design*, Bentham Science Publishers, v. 7, n. 2, p. 146–157, 2011.
- MORRIS, G. M. et al. Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry*, v. 19, n. 14, p. 1639–1662, 1998.

- MOSCONI, L. et al. Multicenter standardized 18f-fdg pet diagnosis of mild cognitive impairment, alzheimer's disease, and other dementias. *Journal of Nuclear Medicine*, Soc Nuclear Med, v. 49, n. 3, p. 390–398, 2008.
- NAMASIVAYAM, V.; GÜNTHER, R. Pso@ autodock: A fast flexible molecular docking program based on swarm intelligence. *Chemical biology & drug design*, Wiley Online Library, v. 70, n. 6, p. 475–484, 2007.
- OKLAHOMA. Statsoft. 2008.
- PARR, R. G. Density functional theory of atoms and molecules. In: *Horizons of Quantum Chemistry*. [S.l.]: Springer, 1980. p. 5–15.
- PAULA, A. D. et al. New potential ache inhibitor candidates. *European journal of medicinal chemistry*, Elsevier, v. 44, n. 9, p. 3754–3759, 2009.
- PEARSON, K. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Taylor & Francis, v. 2, n. 11, p. 559–572, 1901.
- PERDEW, J. P. et al. Density-functional theory for fractional particle number: derivative discontinuities of the energy. *Physical Review Letters*, APS, v. 49, n. 23, p. 1691, 1982.
- PHILLIPS, J. C. et al. Namd: Biomolecular simulation on thousands of processors. In: IEEE. *Supercomputing, ACM/IEEE 2002 Conference*. [S.l.], 2002. p. 36–36.
- PRINCE, M. et al. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimer's & Dementia*, Elsevier, v. 9, n. 1, p. 63–75, 2013.
- RADIĆ, Z.; REINER, E.; TAYLOR, P. Role of the peripheral anionic site on acetylcholinesterase: inhibition by substrates and coumarin derivatives. *Molecular pharmacology*, ASPET, v. 39, n. 1, p. 98–104, 1991.
- SALINAS, S. R. *Introdução a Física Estatística Vol. 09*. [S.l.]: Edusp, 1997.
- SANSON, B. et al. Backdoor opening mechanism in acetylcholinesterase based on x-ray crystallography and molecular dynamics simulations. *Protein Science*, Wiley Online Library, v. 20, n. 7, p. 1114–1118, 2011.
- SAVINI, L. et al. Specific targeting of acetylcholinesterase and butyrylcholinesterase recognition sites. rational design of novel, selective, and highly potent cholinesterase inhibitors. *Journal of medicinal chemistry*, ACS Publications, v. 46, n. 1, p. 1–4, 2003.
- SCHLICK, T. *Molecular modeling and simulation: an interdisciplinary guide: an interdisciplinary guide*. [S.l.]: Springer Science & Business Media, 2010.
- SENIYA, C.; KHAN, G. J.; UCHADIA, K. Identification of potential herbal inhibitor of acetylcholinesterase associated alzheimer's disorders using molecular docking and molecular dynamics simulation. *Biochemistry research international*, Hindawi Publishing Corporation, v. 2014, 2014.
- SENN, H. M.; THIEL, W. Qm/mm studies of enzymes. *Current opinion in chemical biology*, Elsevier, v. 11, n. 2, p. 182–187, 2007.

- SHI, J. et al. Nanosecond dynamics of the mouse acetylcholinesterase cys69–cys96 omega loop. *Journal of Biological Chemistry*, ASBMB, v. 278, n. 33, p. 30905–30911, 2003.
- SHLENS, J. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.
- SLATER, J. C. A simplification of the hartree-fock method. *Physical Review*, APS, v. 81, n. 3, p. 385, 1951.
- SMART, O. S.; GOODFELLOW, J. M.; WALLACE, B. The pore dimensions of gramicidin a. *Biophysical journal*, Elsevier, v. 65, n. 6, p. 2455–2460, 1993.
- SMART, O. S. et al. Hole: a program for the analysis of the pore dimensions of ion channel structural models. *Journal of molecular graphics*, Elsevier, v. 14, n. 6, p. 354–360, 1996.
- SUGIMOTO, H. et al. Donepezil hydrochloride (e2020) and other acetylcholinesterase inhibitors. *Current medicinal chemistry*, Bentham Science Publishers, v. 7, n. 3, p. 303–339, 2000.
- SZABO, A.; OSTLUND, N. S. *Modern quantum chemistry: introduction to advanced electronic structure theory*. [S.l.]: Courier Corporation, 1989.
- TAI, K. et al. Mechanism of acetylcholinesterase inhibition by fasciculin: a 5-ns molecular dynamics simulation. *Journal of the American Chemical Society*, ACS Publications, v. 124, n. 21, p. 6153–6161, 2002.
- TARA, S.; STRAATSMA, T.; MCCAMMON, J. A. Mouse acetylcholinesterase unliganded and in complex with huperzine a: a comparison of molecular dynamics simulations. *Biopolymers*, Wiley Online Library, v. 50, n. 1, p. 35–43, 1999.
- TAVARES, L. C. Qsar: a abordagem de hansch. *Química Nova*, SciELO Brasil, v. 27, n. 4, p. 631–639, 2004.
- TAYLOR, J. S.; BURNETT, R. M. Darwin: a program for docking flexible molecules. *Proteins: Structure, Function, and Bioinformatics*, Wiley Online Library, v. 41, n. 2, p. 173–191, 2000.
- TOMASI, J.; MENNUCCI, B.; CAMMI, R. Quantum mechanical continuum solvation models. *Chemical reviews*, ACS Publications, v. 105, n. 8, p. 2999–3094, 2005.
- TOUKMAJI, A. Y.; BOARD, J. A. Ewald summation techniques in perspective: a survey. *Computer physics communications*, Elsevier, v. 95, n. 2-3, p. 73–92, 1996.
- TROTT, O.; OLSON, A. J. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, Wiley Online Library, v. 31, n. 2, p. 455–461, 2010.
- VALASANI, K. R. et al. Structure based design, synthesis, pharmacophore modeling, virtual screening, and molecular docking studies for identification of novel cyclophilin d inhibitors. *Journal of chemical information and modeling*, ACS Publications, v. 54, n. 3, p. 902–912, 2014.

- VERDONK, M. L. et al. Improved protein–ligand docking using gold. *Proteins: Structure, Function, and Bioinformatics*, Wiley Online Library, v. 52, n. 4, p. 609–623, 2003.
- VESSECCHI, R. et al. Application of computational quantum chemistry to chemical processes involved in mass spectrometry. *Química Nova*, SciELO Brasil, v. 31, n. 4, p. 840–853, 2008.
- VIANNA, J. D. M. *Teoria Quântica de Moléculas e Sólidos: simulação computacional*. [S.l.]: Livraria da Física, 2004.
- WANG, H.; GAO, X.; FANG, J. Multiple staggered mesh ewald: Boosting the accuracy of the smooth particle mesh ewald method. *Journal of Chemical Theory and Computation*, ACS Publications, v. 12, n. 11, p. 5596–5608, 2016.
- WANG, J. et al. Development and testing of a general amber force field. *Journal of computational chemistry*, Wiley Online Library, v. 25, n. 9, p. 1157–1174, 2004.
- WEINER, P. K.; KOLLMAN, P. A. Amber: Assisted model building with energy refinement. a general program for modeling molecules and their interactions. *Journal of Computational Chemistry*, Wiley Online Library, v. 2, n. 3, p. 287–303, 1981.
- XU, Y. et al. Flexibility of aromatic residues in the active-site gorge of acetylcholinesterase: X-ray versus molecular dynamics. *Biophysical journal*, Elsevier, v. 95, n. 5, p. 2500–2511, 2008.
- XU, Y. et al. Long route or shortcut? a molecular dynamics study of traffic of thiocholine within the active-site gorge of acetylcholinesterase. *Biophysical journal*, Elsevier, v. 99, n. 12, p. 4003–4011, 2010.
- YORK, D.; YANG, W. The fast fourier poisson method for calculating ewald sums. *The Journal of Chemical Physics*, AIP, v. 101, n. 4, p. 3298–3300, 1994.
- YUNG-CHI, C.; PRUSOFF, W. H. Relationship between the inhibition constant (k_i) and the concentration of inhibitor which causes 50 per cent inhibition (i_{50}) of an enzymatic reaction. *Biochemical pharmacology*, Elsevier, v. 22, n. 23, p. 3099–3108, 1973.
- ZHOU, J. et al. Application of multivariate statistical approach to identify heavy metal sources in sediment and waters: a case study in yangzhong, china. *Environmental Geology*, Springer, v. 54, n. 2, p. 373–380, 2008.

**Apêndice A – POTENTIAL ACETYLCHOLINESTERASE INHIBITORS:
MOLECULAR DOCKING, MOLECULAR DYNAMICS, AND *IN SILICO*
PREDICTION**

Potential acetylcholinesterase inhibitors: molecular docking, molecular dynamics, and in silico prediction

Alessandra S. Kiametis¹ · Mônica A. Silva² · Luiz A. S. Romeiro³ · João B. L. Martins⁴ · Ricardo Gargano²

Received: 6 September 2016 / Accepted: 12 January 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract This paper deals with molecular modeling of new therapeutic agents for treating the Alzheimer's disease. The therapeutic line adopted for this study is the *cholinergic hypothesis*. To modulate positively the cholinergic function through the inhibition of the acetylcholinesterase, a set of candidates was designed from a natural compound extracted from the cashew nutshell liquid, anacardic acid. In silico screening of this chemical library revealed a ligand that is more promising once it is correlated with an active drug through specific topological and electronic descriptors. The protein–ligand docking showed stable binding modes and the binding free energy computed for the active site of the receptor suggests that our ligand presents a potential biological response.

Keywords AChE inhibitors design · Molecular docking · Molecular dynamics · In silico prediction

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

Electronic supplementary material The online version of this article (doi:10.1007/s00894-017-3228-9) contains supplementary material, which is available to authorized users.

✉ Mônica A. Silva
abreu_monica@yahoo.com.br

¹ Institute of Biology, University of Brasilia, Brasilia, Brazil

² Institute of Physics, University of Brasilia, Brasilia, Brazil

³ Department of Tropical Medicine, University of Brasilia, Brasilia, Brazil

⁴ Institute of Chemistry, University of Brasilia, Brasilia, Brazil

Introduction

First described in 1906 [1], Alzheimer's disease (AD) is the most common form of irreversible dementia among the elderly, although it can also affect young people. It is a kind of neurodegenerative disorder related to progressive cognitive impairments such as memory loss, relative decline in language skill and guidance [2]. The disease severely compromises the quality of life of the patients, causing a huge impact for the social, economic, and political areas [3, 4]. Considered a multi-factorial disease [5], AD affects about 30 million people worldwide and, as a modern epidemic, the number of patients grows significantly every year. Estimates suggest that in 2050 about 107 million people will be affected by AD [6], considering the increased longevity of the population and the age-dependent character of the disease.

Despite that the etiology of AD is not completely known, diverse factors including β -amyloid plaques [7, 8], τ -protein aggregation [9], oxidative stress [10, 11] and decreased levels of acetylcholine (ACh) [12] have been associated with the disease's pathophysiology. From a neuropathological point of view, AD is characterized by premature degeneration of the cholinergic subcortical neurons, especially those located in the basal forebrain (basal nucleus of Meynert). This cholinergic deficit is due to acetylcholinesterase (AChE) enzyme action, which hydrolyses the ACh in the synaptic process. The decreased cholinergic neurotransmission gave rise to the *cholinergic hypothesis*. According to this hypothesis, the increase in the ACh levels by a reversible inhibition of AChE would improve the cognitive profile of the patients [13]. Conventional therapeutic agents are used to inhibit the action of the AChE, preventing the hydrolysis of ACh and thus enabling the

maintenance of the neurotransmitter in acceptable levels but present undesirable side effects. In this sense, the search for a new generation of drugs with cholinergic profile able to improve cognitive functions (and at the same time reduce side effects) has aroused great interest for medicinal chemistry. We propose here a new AChE inhibitor with potential biological response derived from the anacardic acid natural compound. This work is organized as follows: in Section “Methods” we describe the AChE inhibitor design, the ligand–receptor mechanism, and the theoretical methods used to predict the inhibitory action of this new drug. In Section “Results and discussion” we present the main results and discussions. In Section “Conclusions” we outline the conclusions and perspectives.

Methods

Acetylcholinesterase

The crystal structure used as receptor was the human AChE (PDB entry: 4EY7, chain A) complexed with donepezil, a drug with a non-competitive and reversible cholinergic action (Fig. 1). AChE cavity looks like a deep narrow Gorge (Fig. 2), with about 20 Å length, and its active site, where AChE hydrolyzes ACh into acetate and choline, is located at the *catalytic triad site* [SER234, GLU365, HSD478]. AChE also presents other important binding sites for the ligand–receptor mechanism richly described in the literature [14–16]: *peripheral anionic site* (PAS) [TYR103, ASP105, TYR155, GLU316, TRP317, TYR372], *acyl pocket* [TRP117], *anionic site* [TRP267, PHE326, PHE328], *oxyanionic hole* [GLY152, GLY153, ALA235] and *aromatic patch* [TYR164, TYR368].

A molecular dynamics was simulated for the receptor in the NPT ensemble, using Langevin dynamics and the Langevin piston algorithm as implemented in NAMD computational package [17]. Temperature was set at 300K and pressure at 1atm. The system, previously neutralized,

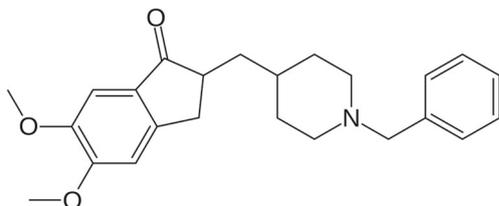


Fig. 1 Donepezil, a known medication for AD treatment with a non-competitive and reversible cholinergic action

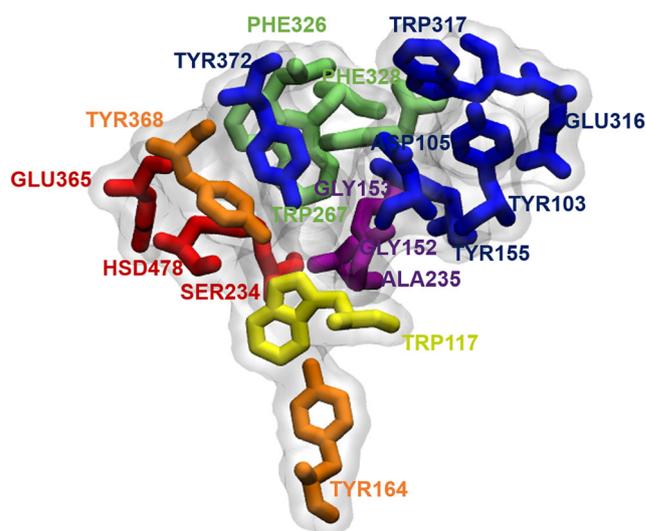


Fig. 2 Important binding sites for ligand–receptor mechanism represented by colors: Red catalytic triad, Blue peripheral anionic site, Yellow the acyl pocket, Green anionic site, Magenta oxyanionic hole, Orange aromatic patch. The sequence of residues were aligned according to the UniProt database (www.uniprot.org)

was immersed in a physiological concentration of NaCl (0.15 μM). Classical non-polarizable force field CHARMM36 and the solvation model TIP3P for water were used during the simulation. Pairwise non-bonded interactions were truncated at a distance of 12 Å, with a smooth switching function above 8 Å. Full electrostatic terms were included to calculations through the particle mesh Ewald method, using a grid spacing of 1.2 Å. A base time step of 2 fs was used to perform trajectories. After the relaxation of the molecular system, the root-mean-square deviation (RMSD) remains constant (about 2.5 Å). The sampling time of our simulation to compute average properties was almost 200 ns, which is large enough to reveal AChE relevant conformational changes.

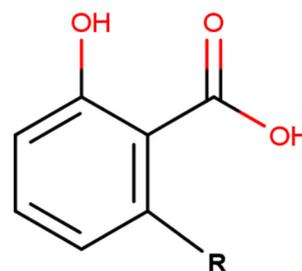


Fig. 3 Anacardic acid, a phenolic derivative extracted from the cashew nutshell (*Anacardium occidentale*). AChE inhibitors were designed adding several pharmacophore groups to the radical (R) linked by a hydrophobic spacer

AChE inhibitory design

The AChE inhibitors were designed from *anacardic acid* (Fig. 3), a phenolic lipid derivative extracted from the cashew nutshell liquid, which is a natural resin found in the honeycomb structure of the cashew (*Anacardium occidentale*) nutshell. The reason to choose anacardic acid as a platform for the design of new compounds is related to previous studies involving ligands based on cardanol, another type of phenolic lipid derivative of the cashew with similar molecular structure to anacardic acid [18–22]. Enzyme inhibition assays were realized for compounds synthesized from cardanol as hydrochlorides, using standard methods [19]. The most potent one presented a IC_{50} of $50\mu M$, the same order of magnitude as the IC_{50} of the galantamine, for example, which is a known medication for AD treatment [23]. Our expectation is to reach the same or even better results in *in vitro* experiments for ligands designed from anacardic acid. A set of pharmacophore groups was added to the radical of the anacardic acid and linked to it through a hydrophobic spacer constituted by eight methylene units giving rise to our library of ligands (Fig. 4). The function of the spacer is first to enable large conformational flexibility to the ligands and second to allow both subunits (the anacardic acid without the alkyl chain and the pharmacophore groups) to interact simultaneously at the PAS and the bottom of the cavity. The benzyl group of the

anacardic acid is expected to interact with the PAS residues through π -stacking interactions forming a kind of “anchor” at the entrance of the enzyme so the inhibition would be reversible. Also, this aromatic portion of the anacardic acid is recognized by AChE as an acceptor of hydrogen bonds. The following pharmacophoric groups are considered for this study: *O*-acylestere and *O*-carbamates of the 3-hydroxypyrrolidine, 3-hydroxypiperidine, and 4-hydroxypiperidine. It is a prospective study to explore the ACh hydrolysis in the catalytic triad, and also using quinoline and tacrine functional groups to simulate the same pattern of the tacrine dimer. All ligands have in common an amino group that is able to interact with the residue TRP117 (acyl pocket), mimicking the substrate of AChE. The addition of a benzyl substituent to the amino group also enables an interaction with the PHE326 residue (anionic site).

In silico screening and molecular docking

Although activity studies have not been made for our ligands yet, principal component analysis (PCA) is a powerful tool for discriminating molecules with respect to biological response when they present a correlation to an active drug for a suitable set of descriptors. In fact, previous studies showed that PCA was really able to predict, for molecules designed from other types of phenolic lipid derivatives of the cashew, the most efficient from the experimental point

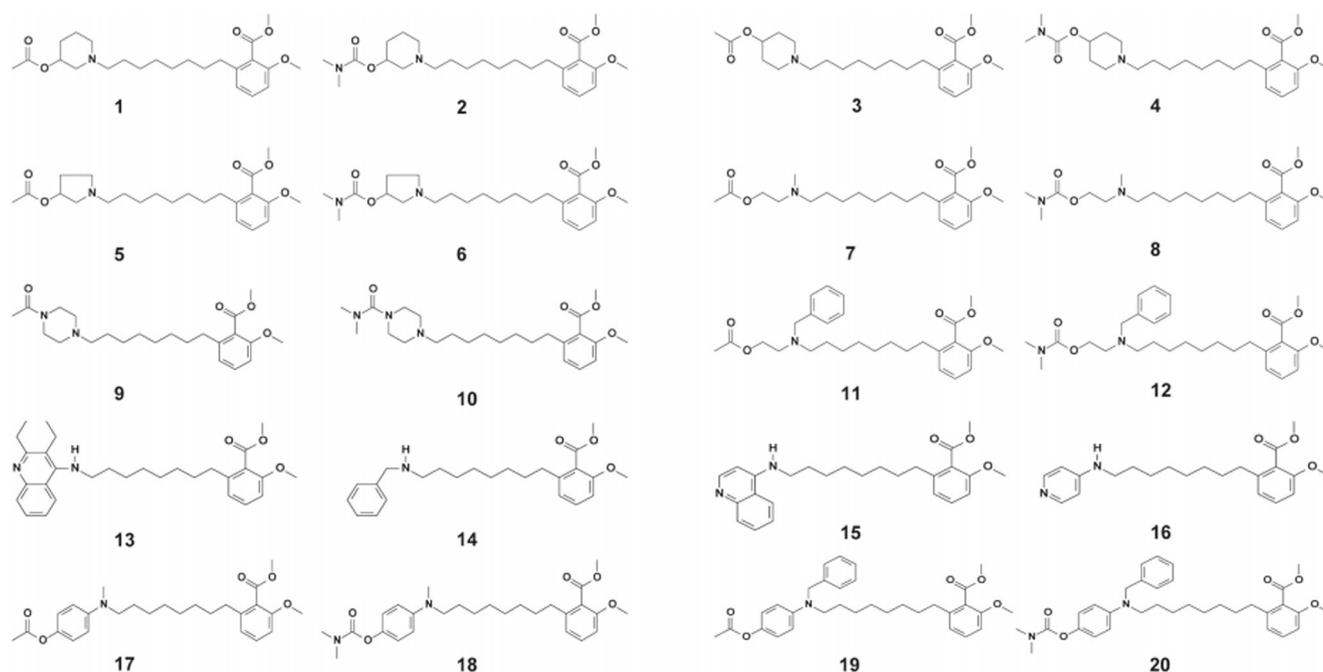


Fig. 4 Set of ligands derived from the anacardic acid. The figure shows the molecular structures and their respective IDs (1–20)

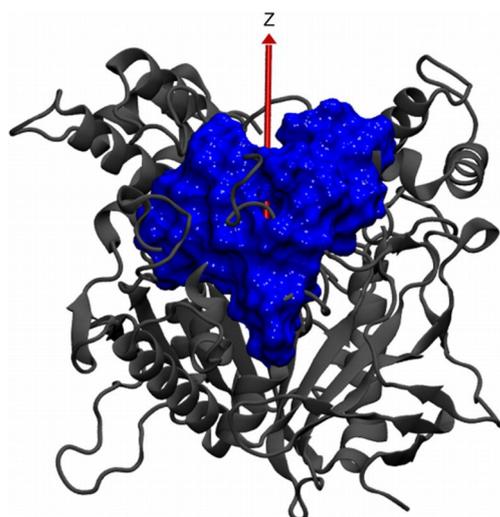


Fig. 5 Acetylcholinesterase structure (PDB entry: 4EY7), chain A. Backbone (represented by grey surface) and the blue surface corresponds to the cavity where the entrance is located at the XY plane

of view [19]. So our goal here is to use the PCA method to point out the most promising molecule (ligand), assuming that this screening protocol leads to reliable results. For this purpose, we used donepezil as a target [24], since we know its binding mode inside the cavity; such structure was solved by X-ray and it can be a parameter for comparison. To make the *in silico* screening of the ligands, we used two kinds of descriptors: electronic and topologic properties [25, 26]. Geometries were optimized using B3LYP/6-311+G(2d,p) density functional through computational Gaussian09 package [27]. The polarizable continuum model (PCM) was

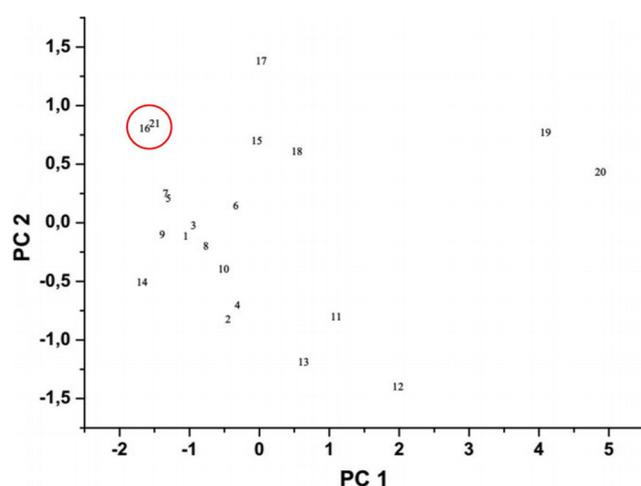


Fig. 6 Score plot of PC1 versus PC2: two-dimensional clustering involving the ID 16 and donepezil molecules

Table 1 Eigenvalues, percentage, and cumulative of total variance for the first three principal components

	Eigenvalue	Percentage	Cumulative
PC1	4.04	80.92	80.92
PC2	0.57	11.55	92.48
PC3	0.28	5.77	98.26

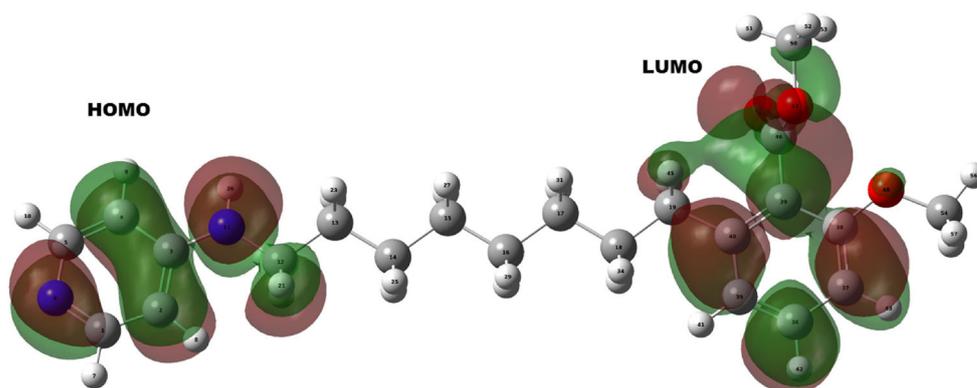
included in the *ab initio* calculations to simulate solvation effects. The following properties were calculated from optimized geometries with the same level: the two highest occupied molecular orbital (HOMO) energies, $E(\text{HOMO})$ and $E(\text{HOMO}-1)$, and the two lowest unoccupied molecular orbital (LUMO) energies, $E(\text{LUMO})$ and $E(\text{LUMO}+1)$, for the most stable conformers. Once the values for the frontier orbital energies are computed, we derived the LUMO-HOMO energy difference (GAP). Partial atomic charges were fit to the electrostatic potential according to CHELPG scheme. Topological descriptors such as molecular weight, mass, volume, polarizability, polar surface area, hydration energy, and LogP (partition coefficient) were computed through quantitative structure activity relationship (QSAR) using HyperChem package [28].

Molecular docking was performed, for the molecule discriminated in PCA as the most promising one, using AutoDock Vina software [29]. Flexible ligand conformations were used for the docking and dihedral scans were performed using *ab initio* HF/6-31G. The conformers were docked into the statistical ensemble of equilibrium for the AChE *open state*. The protein displays a kind of breathing during the simulation, which has already been described in the literature [16]. On average, the atomic displacements of the whole protein are basically due to thermal fluctuations, but the Gorge presents distinct conformations along the dynamics. Most of the time, residues TYR155 and TYR372 are sufficiently close to cause the *stricture* of the Gorge. In order to characterize an open conformation, we computed AChE profile along the z' -axis (Fig. 5) for all trajectories

Table 2 The most significant variables for the observed correlation and their respective weights for the first three principal components

DESCRIPTORS	PC1	PC2	PC3
HOMO	0.367352	0.852401	-0.358881
Nitrogen	0.433330	0.179907	0.873089
Energy	-0.479416	0.246280	0.260668
Volume	0.470093	-0.370413	-0.191164
Polarizability	0.475868	-0.207813	-0.066545

Fig. 7 HOMO and LUMO orbitals plotted for molecule ID **16**. HOMO orbital is concentrated at the ring bonded to the amino group while the LUMO orbital is concentrated at the dimethoxy-phenyl ring



using HOLE computational package [30] to analyze the pore dimensions. For this experiment, we constructed a R' referential in which the origin matches the center of the Gorge and the z' -axis is parallel to the canonical z -axis. So the probe starts at the origin and runs through the cavity in both directions (positive and negative) measuring the local radius.

Results and discussion

The PCA procedure was performed taking into account all possible linear combinations of the initial data set. The components were classified in order of significance, so the first one (PC1) corresponds to the maximum dispersion axis. Only those combinations in which the PC1 variance showed equal or greater than 75 % have been selected for appreciation. The reader can find the full table of descriptors (for all ligands and the donepezil) in the Supporting Information (SI). After systematic analysis, among several patterns

of distribution of the ligands around the target (ID **21**) we observed a two-dimensional clustering involving molecule ID **16** and the active drug, for a particular set of descriptors. The score plot of PC1 versus PC2 (Fig. 6) clearly shows this. Table 1 shows the eigenvalues and percentage of total variance for the first three principal components. Note that PC1 is responsible for more than 80 % of the information, so a two-dimensional model is sufficient to describe molecular distribution. Table 2 shows the most relevant descriptors for the observed correlation - the HOMO energy, the atomic charge of the nitrogen belonging to the amino group (N), the global minimum energy for the most stable conformer (*single-point* calculations), the polarizability, the molecular volume, and their respective contributions for the principal components, being all equally relevant for AChE action.

Molecule ID **16** has a benzene heterocyclic ring bonded to the amino group, where the HOMO orbital (Fig. 7) is located, showing the electron donor character of the ligand in this region (HOMO energy is -6.12 hartree). The LUMO (Fig. 7) orbital is concentrated at the dimethoxy-phenyl ring,

Fig. 8 Molecular electrostatic potential for molecule ID **16**. The *reddish* surface represents the attractive region of the ligand while the *bluish* surface represents the repulsive region of the ligand

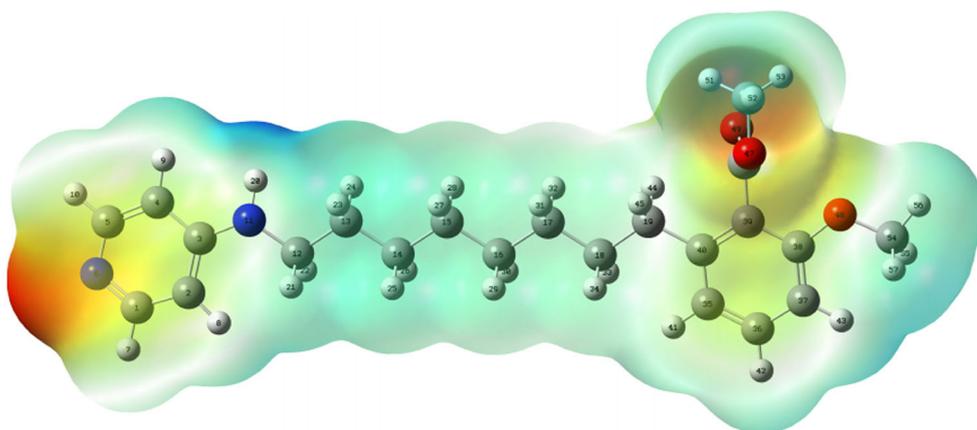
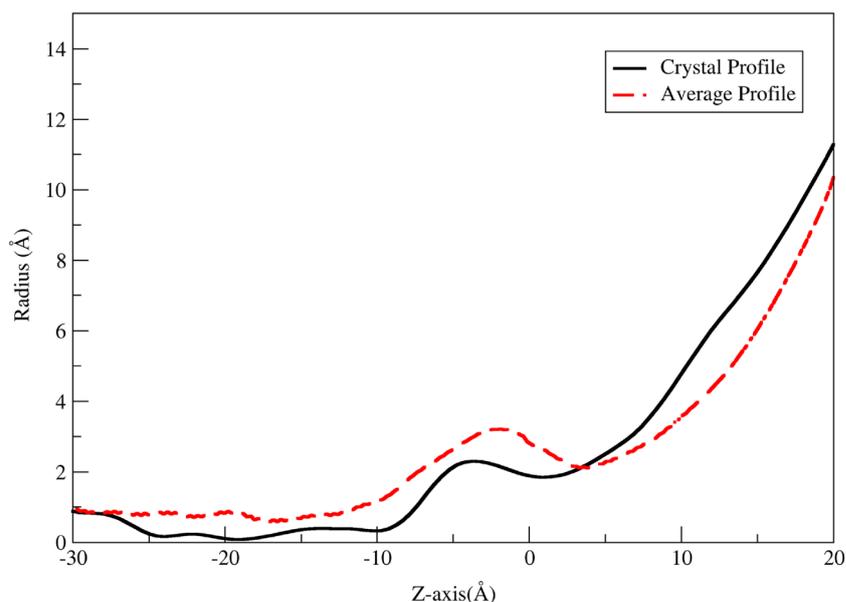


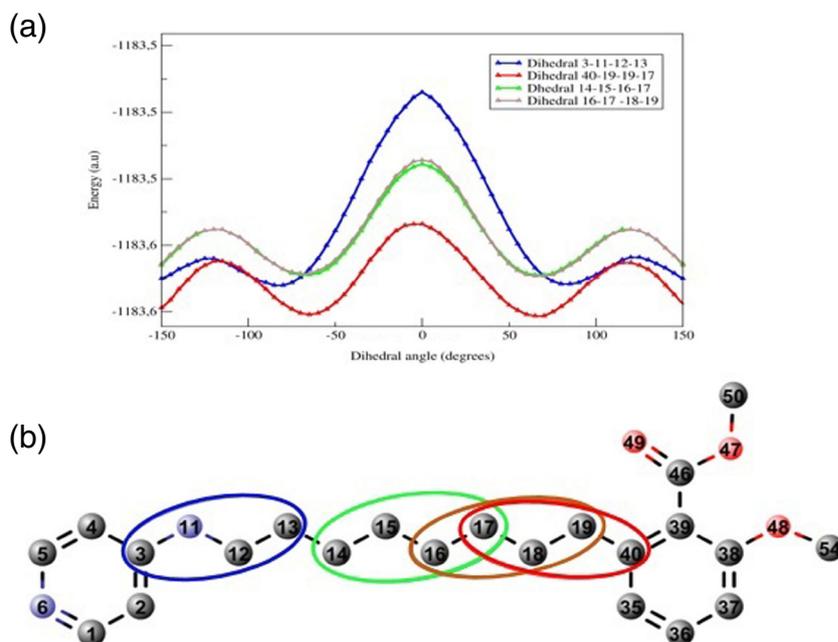
Fig. 9 Profile along the z-axis. The black and red curves correspond to the crystal profile and average of open microstates, respectively



showing the electron acceptor character of the ligand in this region (LUMO energy is -0.75 hartree). Charge density distribution over the ligand can be analyzed through the molecular electrostatic potential (Fig. 8). Two main donor regions are found, the amine group and the benzyl groups of the anacardic acid, which are capable of interacting with the PAS residues. Figure 9 shows the computed profile (black curve) for the receptor crystal, according to the procedure

described on methodology, which is in an open-state conformation. This opening radius is sufficiently large to enable the ligand to access the entire cavity. The open microstate profiles were selected by RMSD criteria (less than 2.5 \AA) in relation to the crystal profile. The ensemble of equilibrium for the open state was generated based on profiles with similar behavior, represented by the average profile (red curve). The molecule ID **16** was docked into the statistical ensemble

Fig. 10 Angular scan of the molecule ID **16**: **a** Associated energy to each dihedral angle for generated conformations. **b** The atom notation for the respective dihedral angle highlighted in circles



of equilibrium. All receptor structures were centered at the origin and superimposed. The docking method tried to find solutions in a search space box (60ÅX60ÅX60Å) centered at the origin. It also generated the ligands' conformations according to the number of rotate angles; 13 active torsions were detected (Fig. 10). The sampling has found 1000 docking solutions, less than 1 % corresponds to distinct regions of the cavity and were discarded. Docking has identified two binding sites. Binding site 1 corresponds to a buried cavity population and binding site 2 corresponds to an entrance one (Fig. 11). According to Table 3, the binding site 1 presents a higher solution density (number of solutions of the site per total number of solutions) than the binding site 2. The binding site free energies ($\Delta G_{binding}$) due to the non-bonded interactions were estimated by AutoDock Vina through the scoring functions:

$$\Delta G_{binding} = \Delta G_{gauss} + \Delta G_{repulsion} + \Delta G_{hbond} + \Delta G_{hydrophobic} + \Delta G_{torsion}$$

where ΔG_{gauss} is the attractive term for dispersion, $\Delta G_{repulsion}$ is the square of the distance (if closer the threshold value), ΔG_{hbond} and $\Delta G_{hydrophobic}$ are ramp functions and $\Delta G_{torsion}$ is proportional to the number of rotate bonds. The free binding energy here is an estimate (AutoDock standard error is about 2.85 kcal/mol) and more accurate methods as free energy perturbation (FEP), for example, could be performed. Binding site 1 proved to be more energetically stable and the complex constant dissociation (K_d) is also greater for this binding site (Table 3).

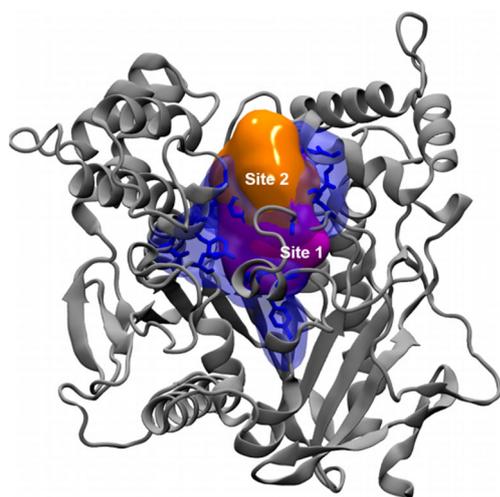


Fig. 11 Density solution representations: *bluish* surface corresponds to the residues of the Gorge. *Magenta* surface corresponds to the solutions at the entrance of the Gorge. *Orange* surface corresponds to the solutions inside the Gorge

The best binding modes for both binding sites were ranked according to energy criteria and number of contacts. We have found three most favorable energetic poses, two for the site 1 (notation pose 1.1 and 1.2), and one for the site 2 (notation pose 2.1). The particular poses 1.1 and 1.2 (more energetically favorable) for binding site 1 are shown in Figs. 12 and 13. In pose 1.1 (-8.5 kcal/mol), we can see the benzyl group binding to residue TRP117 (acyl pocket) through π -stacking interactions (3.79Å) while the dimethoxy-phenyl ring (which is a hydrogen and carboxylate bond donor) binds to residue TYR155 (PAS) through a dipole-dipole interaction (3.34Å). It is interesting to note that the amino group of the ligand (a hydrogen bond donor) is faced to the residue HSD478 (catalytic triad) making a hydrogen bond (2.49Å). In pose 1.2 (-8.4 kcal/mol) the benzyl group anchors at the PAS through a π -stacking interaction with TYR372 (PAS) residue (3.54Å) while the dimethoxy-phenyl ring binds to TRP117 (acyl pocket), which is a hydrogen bond acceptor, through dipole-dipole interactions (3.03Å). In terms of energy, pose 1.1 practically does not differ from pose 1.2. For the binding site 2, the most stable 3D structure for the complex (-7.3 kcal/mol) is shown in Fig. 14. In pose 2.1, although the benzyl group is anchored at the PAS, interacting with residue TYR372 (PAS) through a π -stacking interaction (3.66Å), the another part of the ligand is outside the cavity. Comparing the binding site poses, their respective energies and solution densities, the most probable situation is the ligand at the interior of the cavity. In this case, the hydrophobic spacer can be stretched, so the ligand anchors at the PAS, or it can be bended, so the ligand interacts with the anionic site and the acyl pocket at the same time. In any way, the ligand would be acting as reversible and non competitive AChE inhibitor. The interaction analysis of the crystal structure (protein-ligand complexes) of human AChE (PDB entry: 4EY7, chain A) shows two main interactions along the Gorge. In the bottom of the active site, the donepezil binds to residue TPR117 (acyl pocket) through π -stacking interaction, measured distance is 3.60Å. Figure 15 depicts the active site entrance where the ligand binds to TRP317

Table 3 The binding sites and their respective values for solution densities, binding free energy ($\Delta G_{binding}$) and dissociation constant (K_d), estimated through scoring functions of AutoDock Vina

Binding sites	Solution density	$\Delta G_{binding}$ (kcal/mol)	K_d
Site 1	0.82	-4.64	0.540
Site 2	0.13	-3.28	0.086

Fig. 12 Pose 1.1 (Docking of molecule ID 16 in active site of the AChE): **a** Representation of the three dimensional structure of the AChE, showing the important binding sites of the Gorge and the conformation of the ligand with binding energy of -8.5 kcal/mol. **b** Three-dimensional analysis of the interactions: π -stacking interactions with TRP117 (acyl pocket), dipole-dipole interaction with TYR155 (PAS) and hydrogen bond interaction with HSD478 (catalytic triad)

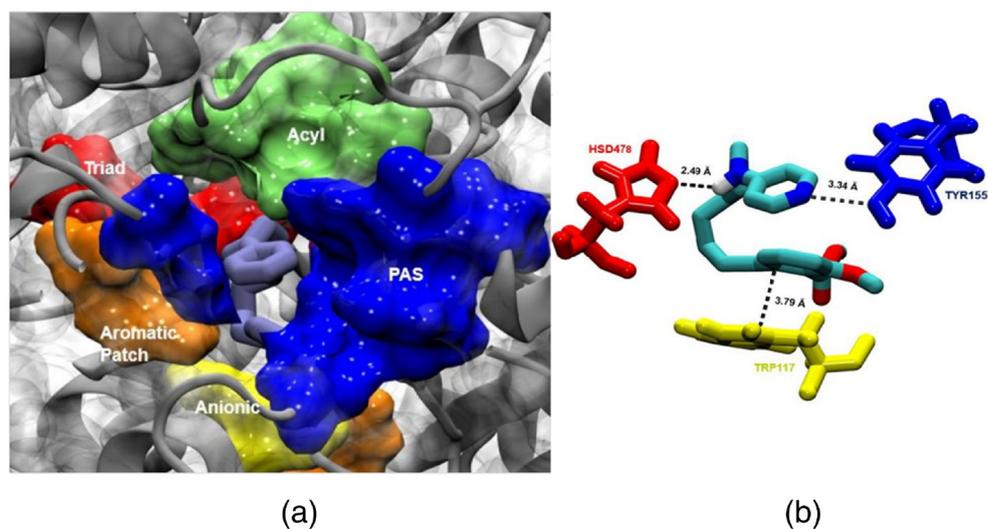


Fig. 13 Pose 1.2 (Docking of molecule ID 16 in active site of the AChE): **a** Representation of the three-dimensional structure of the AChE showing the important binding sites of the Gorge and the conformation of the ligand with binding energy of -8.4 kcal/mol. **b** Three-dimensional analysis of the interactions: ligand anchors at the PAS through a π -stacking interaction with TYR372 residue (with a distance of 3.54\AA). The ligand also binds to TRP117 (acyl pocket) through cation- π interaction (with a distance of 3.03\AA)

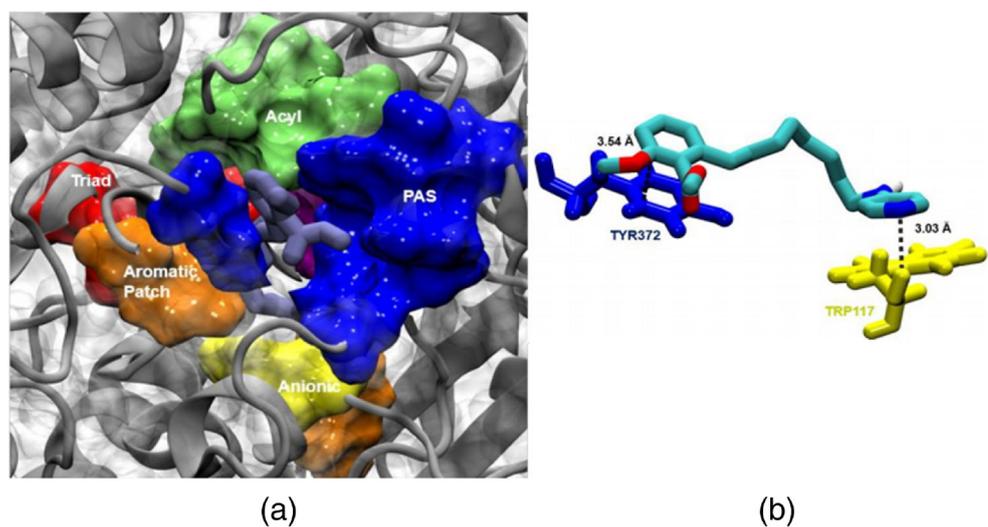
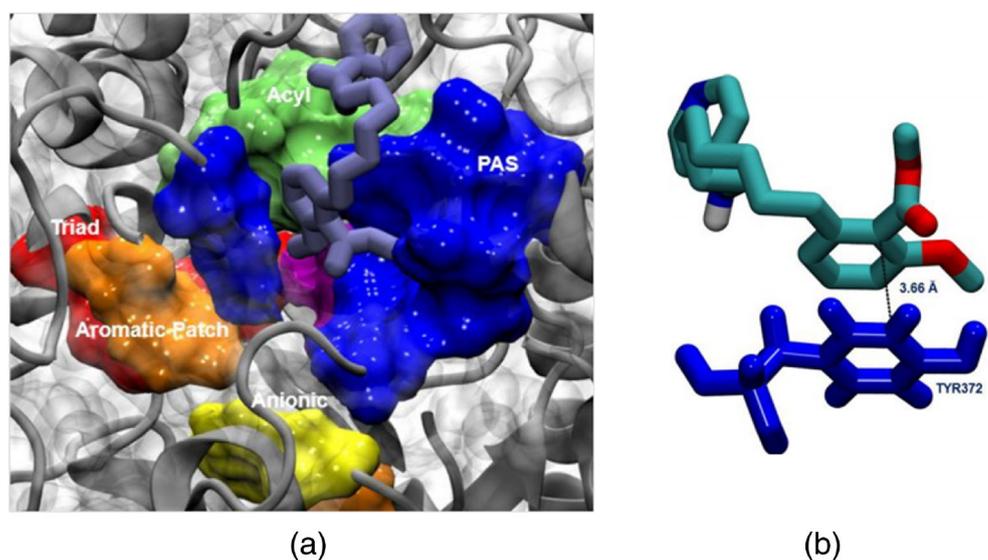


Fig. 14 Pose 2.1 (Docking of molecule ID 16 in active site of the AChE): **a** Representation of the three-dimensional structure of the AChE showing the important binding sites of the Gorge and the conformation of the ligand with binding energy of -7.3 kcal/mol. **b** Three-dimensional analysis of the interactions: ligand anchors at the PAS through a π -stacking interaction with TYR372 residue (with a distance of 3.66\AA)



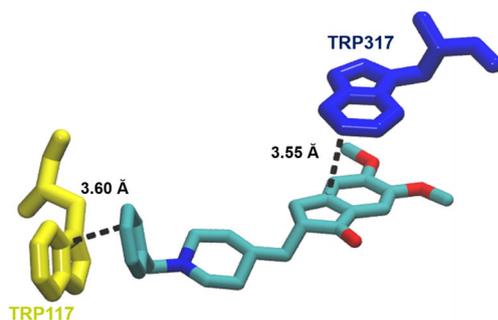


Fig. 15 Representation of the interactions between the crystal structure human AChE (PDB entry: 4EY7, chain A) and donepezil. The donepezil binds with TRP317 residue in the PAS, via a π -stacking interaction (with a distance of 3.55 Å) and also binds with TRP117 (acyl pocket) through a π -stacking interaction (with distance of 3.60 Å)

(PAS) through π -stacking interaction (3.55 Å). Comparison between donepezil and our ligand binding modes (poses of the site **1**) shows a similarity of their structures. The analysis of the interactions from molecular docking for the poses 1.1 and 1.2 show that, as well as the donepezil, the ligand (ID **16**) interacts with residue TRP117 at the bottom of the Gorge. Furthermore, they also interact with PAS residues. The measured distances are 3.60 Å and 3.79 Å, for donepezil and the ligand, respectively.

Conclusions

A new therapeutic agent for the treatment of Alzheimer disease is presented here as an alternative non-competitive AChE inhibitor. Its advantages over other known AChE inhibitors is the fact that it was designed from natural raw material with a large spacer that provides the ligand to anchor simultaneously to the PAS and the bottom of the cavity. Our modeling, based on *in silico* molecular screening and molecular docking, is able to predict the inhibitor binding modes, and that the ligand docked into the statistical ensemble of equilibrium for the receptor open state is energetically stable. Our results suggest that the selected ligand (ID **16**) presents a potential biological response.

Acknowledgments The authors gratefully acknowledge the financial support from the Brazilian Research Councils CAPES, CNPq, FINATEC, FAPDF, and the computational support of LBTC-UnB.

References

- Alzheimer A (1906) Über einen eigenartigen schweren Erkrankungsprozess der Hirnrinde. *Neurologisches Centralblatt* 23:1129–1136
- 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
- Goedert M, Spillantini MGA (2006) Century of Alzheimer's disease. *Science* 314:777–781
- Morley JE (2011) Anticholinergic medications and cognition. *J Am Med Dir Assoc* 12:543–543
- Igbal K, Grundke-Igbal I (2000) Alzheimer disease is multifactorial and heterogeneous. *Neurobiol Aging* 21:901–902
- Jellinger KA (2006) Alzheimer 100—highlights in the history of Alzheimer research. *J Neural Transm* 113:1603–1623
- Castro A, Martinez A (2006) Targeting beta-amyloid pathogenesis through acetylcholinesterase inhibitors. *Curr Pharm Des* 12:4377–4387
- Karran E, Mercken M, Strooper BD (2011) The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nature Rev Drug Discov* 10:698–712
- Alonso AC, Grundke-Igbal I, Igbal K (1996) Alzheimer's disease hyperphosphorylated tau sequesters normal tau into tangles of filaments and disassembles microtubules. *Nat Med* 2:783–787
- Andersen JK (2004) Oxidative stress in neurodegeneration: cause or consequence *Nat Med* 10:S18–25
- Dumont M, Beal MF (2011) Neuroprotective strategies involving ROS in Alzheimer disease. *Free Rad Biol Med* 51:1014–1026
- Cummings JL (2004) Treatment of Alzheimer's disease: current and future therapeutic approaches. *Rev Neurol Dis* 1:60–69
- Mangialasche F, Solomon A, Winblad B (2010) Alzheimer's disease: clinical trials and drug development. *Lancet Neurol* 9:702–716
- Shaffer A, Kronman C, Flashner Y, Leitner M, Grosfeld H, Ordentlich A, Gozes Y, Cohen S, Ariel N, Barak D (1992) Mutagenesis of human acetylcholinesterase. *Identif Resid Invol Catalyt Actv Polypep Fold* 267:17640–17648
- Dvira H, Silman I, Harela M, Rosenberry TL, Sussman JL (2010) Acetylcholinesterase: from 3D structure to function. *Chemico-Biol Interact* 187:10–22
- Bennion BJ, Essiz SG, Lau EY, Fattbert J-L, Emigh A, Lightstone FC (2015) A wrench in the works of human acetylcholinesterase: soman induced conformational changes revealed by molecular dynamics simulations. *PLoS ONE* 10:1–31
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802
- Martins JBL, Neves AA, Gargano R, Santos ML, Romeiro LAS (2007) Electronic structure calculations toward new potentially AChE inhibitors. *Chem Phys Lett* 446:304–308
- De Paula AAN, Martins JBL, Santos ML, Nascente LC, Romeiro LAS, Areas TFMA, Vieira KST, Gambo NF, Castro NG, Gargano R (2009) New potential AChE inhibitor candidates. *Eur J Med Chem* 44:3754–3759
- Kiametis AS, Martins JBL, Romeiro LAS, Gargano R (2013) Acetylcholinesterase inhibitors: modeling potential candidates. *Int J Quantum Chem* 113:1461–1466
- Nascimento ECM, Martins JBL (2011) Electronic structure and PCA analysis of covalent and non-covalent acetylcholinesterase inhibitors. *J Mol Model* 17:1371–1379
- Nascimento ECM, Martins JBL, Gargano R, Santos ML (2008) Theoretical study of classical acetylcholinesterase inhibitors. *Chem Phys Lett* 458:285–289
- Samochochi M, Hoffle A, Fehrenbacher A, Jostock R, Ludwig J, Christner C, Radina M, Zerlim M, Ullmer C, Pereira E, Lubbert H, Albuquerque E (2003) Galantamine is an allosterically potentiating ligand of neuronal nicotinic but not of muscarinic acetylcholine receptors. *J Pharmacol Exp Ther* 305:1024–1036
- Cheung J, Rudolph M, Burshteyn F, Cassidy M, Gary E, Love J, Height J, Franklin M (2012) Structures of human acetylcholinesterase in complex with pharmacologically important ligands. *J Med Chem* 55:10282–10288

25. Raevsky OA (2004) Physicochemical descriptors in property-based drug design. *Mini Rev Med Chem* 4:1041–1052
26. Sakkiah S, Lee KW (2012) Pharmacophore-based virtual screening and density functional theory approach to identifying novel butyrylcholinesterase inhibitors. *Acta Pharmacol Sin* 33:964–978
27. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JAJr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2009). Gaussian09, Inc., Wallingford, CT, USA, Gaussian, Inc., Wallingford, CT
28. HyperChem Computational Chemistry (1996). Hypercube Inc., New York
29. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem* 31:455–461
30. Smart OS, Neduelil JG, Wang X, Wallace BA, Sansom MSP (1996) HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* 14:354–360