

UNIVERSIDADE FEDERAL DO AMAPÁ PRO-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS

PRISCILLA FARIAS OZELA

VALIDAÇÃO DE MÉTODO COMPUTACIONAL APLICADO À GALANTAMINA NO PLANEJAMENTO DE CANDIDATOS A FÁRMACOS PARA TRATAMENTO DA DOENÇA DE ALZHEIMER

Macapá 2017

PRISCILLA FARIAS OZELA

VALIDAÇÃO DE MÉTODO COMPUTACIONAL APLICADO À GALANTAMINA NO PLANEJAMENTO DE CANDIDATOS A FÁRMACOS PARA TRATAMENTO DA DOENÇA DE ALZHEIMER

Dissertação apresentada ao programa de Pós-graduação em Ciências Farmacêuticas da Universidade Federal do Amapá – UNIFAP, como requisito parcial à obtenção do Título de Mestre em Ciências Farmacêuticas, na área de concentração Tecnologia aplicada a fármacos.

Orientador: Prof^a. Dr^a. Lorane Izabel da Silva Hage Melim. Co-Orientador: Prof^o. Dr. Cleydson Breno Rodrigues dos Santos.

Macapá 2017

Dados Internacionais de Catalogação na Publicação (CIP)

Biblioteca Central da Universidade Federal do Amapá

615.321

O99v Ozela, Priscilla Farias.

Validação de método computacional aplicado à galantamina no planejamento de canditados a fármacos para tratamento da doença de alzheimer / Priscilla Farias Ozela; orientadora, Lorane Izabel da Silva Hage Melim. – Macapá, 2017.

73 f.

Dissertação (Mestrado) – Fundação Universidade Federal do Amapá, Programa de Pós-Graduação em Ciências Farmacêuticas.

PRISCILLA FARIAS OZELA

VALIDAÇÃO DE MÉTODO COMPUTACIONAL APLICADO À GALANTAMINA NO PLANEJAMENTO DE CANDIDATOS A FÁRMACOS PARA TRATAMENTO DA DOENÇA DE ALZHEIMER

Dissertação apresentada ao programa de Pós-graduação em Ciências Farmacêuticas da Universidade Federal do Amapá – UNIFAP, como requisito parcial à obtenção do Título de Mestre em Ciências Farmacêuticas, na área de concentração Tecnologia aplicada a fármacos.

Data de Aprovação: <u>31</u> / <u>03</u> / <u>2017</u>

Dra^a. Lorane Izabel da Silva Hage Melim – Orientadora

Lemande

Dr. Caio Pinho Fernandes – UNIFAP

Deimia

Dr. Daimio Chaves Brito - UEAP

À minha família, por ter sido sempre muito mais do que eu poderia ter sonhado. Amo vocês!

AGRADECIMENTOS

Primeiramente a Deus, por ter traçado sempre lindos planos para minha vida, me permitindo chegar até aqui sem jamais me abater.

Aos meus pais, por terem me ensinado desde cedo o valor do estudo e da perseverança.

À toda a minha família, por todo o auxílio, carinho e disponibilidade no decorrer de todo o meu desenvolvimento pessoal e intelectual.

Ao meu amor Robson Carvalho Barbosa, que no dia-a-dia me ajudou e incentivou a ser cada vez melhor me mostrando que ser Feliz é bem mais simples do que eu supunha.

À Universidade Federal do Amapá, pela oportunidade criada através do Programa de Pós-Graduação em Ciências Farmacêuticas.

À prof. Dra. Lorane Izabel da Silva Hage Melim, por sua orientação dedicada, apoio conselheiro e acolhimento paciente, sempre muito além do lado profissional.

Ao prof. Dr. Cleydson Breno Rodrigues dos Santos, pela co-orientação repleta de paciência, dedicação e companheirismo sem jamais perder a alegria da caminhada.

À equipe dos laboratórios PharMedChem e LMQC, por terem me recebido de braços abertos como uma verdadeira família.

Aos Amigos, que de perto ou de longe sempre acreditaram no meu potencial, me motivando a trilhar os melhores caminhos para chegar até aqui.

Aos colegas de turma, por todo apoio e escuta diária durante a batalha que travamos juntos no último ano.

A todos que colaboraram direta ou indiretamente para a concretização deste trabalho, compartilhando comigo seu tempo e seus conhecimentos.

"Inteligência é a capacidade de se adaptar às mudanças" (Stephen Hawking)

> "Somos o que repetidamente fazemos, portanto, a excelência não é um feito, mas um hábito". (Alberto Escudero)

RESUMO

A doença de Alzheimer (DA) é uma patologia neurodegenerativa causada por múltiplos fatores, tais como: deficiência de acetilcolina nas sinapses; agregação de peptídeos beta-amiloides para formação de placas senis no tecido cerebral e formação de emaranhados neurofibrilares no interior dos neurônios. É uma doença cada vez mais prevalente em idosos e que ganha importância crescente com o desenvolvimento social e o aumento da expectativa de vida no mundo. Atualmente, existem cinco fármacos aprovados e disponíveis no mercado brasileiro como drogas anti-Alzheimer, que são a tacrina, o donepezil, a rivastigmina, a galantamina e a memantina, atuando como fármacos alvo-específicos para redução dos sintomas, porém sem apresentar cura conhecida até hoje. Na tentativa de modificar esta realidade, a comunidade científica continua buscando fármacos mais eficazes, de forma que a modelagem molecular tem sido uma importante ferramenta para acelerar o processo de triagem de compostos potencialmente ativos e obter propriedades específicas desses potenciais fármacos. Sendo assim, este trabalho analisou os parâmetros geométricos da estrutura cristalográfica da galantamina, otimizando-a sob diferentes métodos computacionais objetivando a validação do melhor método computacional e conjunto de bases quando comparado ao modelo experimental, para ser usado em futuros estudos de planejamento de fármacos anti-Alzheimer. Para isso, este estudo utilizou vários métodos de química quântica, análise multivariada e estatística, a partir dos quais percebeu-se que, dentre os 15 métodos computacionais avaliados, o método Hartree-Fock associado ao conjunto de bases 6-31G** (HF6-31G**) foi o que melhor reproduziu os parâmetros geométricos experimentais, sendo, portanto, o método validado para galantamina neste trabalho, sugerindo que possa ser utilizado para outros alcaloides com regiões farmacofóricas semelhantes.

Palavras-chave: Doença de Alzheimer. Modelagem molecular. Validação. Análise multivariada. HF6-31G**.

ABSTRACT

Alzheimer's disease (AD) is a neurodegenerative disorder caused by multiple factors, such as: acetylcholine deficiency at synapses; Aggregation of beta-amyloid peptides for the formation of senile plaques in the brain tissue and the formation of neurofibrillary tangles within the neurons. It is an increasingly prevalent disease in the elderly and is gaining increasing importance with social development and increased life expectancy in the world. Currently, there are five drugs approved and available in Brazilian market as anti-Alzheimer drugs, which are tacrine, donepezil, rivastigmine, galantamine and memantine, acting as target-specific drugs to reduce symptoms, but AD remains without known cure until today. In an attempt to modify this reality, the scientific community continues to seek more effective drugs, and molecular modeling has been an important tool to accelerate the screening process for potentially active compounds and to obtain specific properties of these potential drugs. Thus, this work analyzed the geometric parameters of the crystal structure of galantamine, optimizing it under different computational methods aiming at the validation of the best computational method and basis sets when compared to the experimental model, and, therefore, be used in further studies of anti-Alzheimer drug planning. In order to do so, this study used several methods of quantum chemistry, multivariate and statistical analysis, from which it was observed that, among the 15 computational methods evaluated, the Hartree-Fock method associated to the 6-31G** basis set (HF6-31G**) was the one that best reproduced the experimental geometric parameters, being therefore the validated method for galantamine in this work, suggesting that this method can be applied to others alkaloids with similar phamacophore regions.

Keywords: Alzheimer Disease. Molecular Modelling. Validation. Multivariate Analysis. HF6-31G**.

LISTA DE ABREVIATURAS E SIGLAS

ACh	Acetilcolina						
AChE	Acetilcolinesterase						
AINEs	Anti-inflamatórios não esteroides						
AM1	Austin model 1						
ßA	Beta-amilóide						
BChE	Butirilcolinesterase						
CAT	Colina acetiltransferase						
CMAE	Erro de média absoluto corrigido						
DA	Doença de Alzheimer						
DFT	Teoria do funcional da densidade						
F	Fator de Fisher						
FDA	Food and Drug Administration						
Glu	Glutamato						
Gly	Glicina						
HCA	Análise de agrupamento hierárquico						
HF	Hartree-Fock						
HTS	"High-throughput screening" (Ensaios biológicos automatizados						
	em larga escala)						
IAChE	Inibidores de acetilcolinesterase						
IC ₅₀	Concentração inibitória mínima						
IChE	Inibidores das colinesterases						
IURCr	International Union of Crystallography						
MAE	Erro de média absoluto						
NMDA	N-metil-D-aspartato						
PC'S	Componentes principais						
PC1	Primeiro componente principal						
PC2	Segundo componente principal						
PC3	Terceiro componente principal						
PCA	Análise de componentes principais						
PDB	"Protein Data Bank"						
Phe	Fenilalanina						

Parametric method 3
Soma dos quadrados dos desvios
Validação cruzada
Relação estrutura atividade-quantitativa
Relação estrutura propriedade-quantitativa
Regressão linear
Regressão linear simples
Relação estrutura atividade
Erro padrão estimado
Serina
Desvio padrão da validação cruzada
Torpedo californica – acetilcolinesterase
Triptofano
Zerner's Intermediate Neglect of Differential Overlap

LISTA DE TABELAS

Tabela 1	-	Descrição dos alcaloides com atividade anti-						
		Alzheimer	25					
Tabela 2	-	Parâmetros geométricos teóricos e experimentais da região						
		farmacofórica da galantamina	46					
Tabela 3	Tabela 3 – Parâmetros geométricos selecionados pelo PCA incluin							
		desvio padrão dos métodos computacionais e matriz de						
		Correlação de Pearson	50					
Tabela 4	-	PCA do modelo selecionado para os métodos e para os						
		conjuntos de bases	51					
Tabela 5	-	Parâmetros geométricos selecionados pelo PCA incluindo os						
		métodos e conjuntos de bases com suas variações em relação						
		aos parâmetros geométricos experimentais	53					
Tabela 6	-	Principais parâmetros estatísticos calculados nos diferentes						
		métodos e conjuntos de bases	59					

LISTA DE FIGURAS

Figura 1	-	Aspectos macroscópicos da atrofia do hipocampo e do córtex	
		cerebral	18
Figura 2	-	Mecanismos de desenvolvimento da Doença de Alzheimer	19
Figura 3	-	Estruturas químicas de alcaloides isolados de plantas com	
		atividade anti-Alzheimer	23
Figura 4	-	Estrutura química da Galantamina	36
		a) numerada conforme utilizado neste estudo	36
		b) na forma de sal – bromidrato de galantamina	36
Figura 5	-	Principais resíduos de aminoácidos responsáveis pela	
		interação entre a galantamina e a AChE (PDB: 1QTI)	37
Figura 6	-	Plotagem dos parâmetros geométricos selecionados pelo	
		PCA	48
Figura 7	-	Plotagem dos scores obtidos na separação dos métodos	
		computacionais em relação aos dados experimentais após a	
		análise de PCA	49
Figura 8	-	Dendograma dos métodos computacionais estudados,	
		incluindo os dados dos parâmetros geométricos	
		experimentais da Galantamina	57
Figura 9	-	Gráfico de Regressão Linear Simples (RLS) obtido através da	
		comparação entre os dados experimentais e os dados teóricos	
		após a otimização pelo método HF6-31G**	62

SUMÁRIO

1	INTRODUÇÃO	14				
2	OBJETIVOS 1					
2.1	OBJETIVO GERAL 1					
2.2	OBJETIVOS ESPECÍFICOS 1					
3	REFERENCIAL TEÓRICO	17				
3.1	DOENÇA DE ALZHEIMER	17				
3.2	FISIOPATOLOGIA DA DOENÇA DE ALZHEIMER	17				
3.3	TRATAMENTOS DISPONÍVEIS	20				
3.4	PRODUTOS NATURAIS PARA TRATAMENTO DA DOENÇA DE					
	ALZHEIMER	22				
3.5	MODELAGEM MOLECULAR NO PLANEJAMENTO DE					
	FÁRMACOS	24				
3.6	MÉTODOS DE QUÍMICA QUÂNTICA APLICADOS À					
	MODELAGEM MOLECULAR DE COMPOSTOS					
	BIOATIVOS	27				
3.6.1	Métodos de química quântica	27				
3.6.2	Métodos de análise multivariada	29				
3.6.2.1	Análise de componentes principais	30				
3.6.2.2	Análise de agrupamento hierárquico	32				
3.6.3	Método de análise quantitativa	33				
4	MATERIAL E MÉTODOS	36				
4.1	SELEÇÃO DO COMPOSTO ESTUDADO – GALANTAMINA 36					
4.2	IDENTIFICAÇÃO DA REGIÃO FARMACOFÓRICA E SELEÇÃO					
	DOS PARÂMETROS GEOMÉTRICOS DA GALANTAMINA	37				
4.3	MODELAGEM MOLECULAR DA GALANTAMINA UTILIZANDO					
	DIFERENTES MÉTODOS COMPUTACIONAIS E CONJUNTOS					
	DE BASES	38				
4.4	APLICAÇÃO DA TÉCNICAS DE ANÁLISE MULTIVARIADA					
4.4.1	Análise de Componente Principais (PCA)	39				
4.4.2	Análise de Agrupamento Hierárquico (HCA)	40				
4.5	ANÁLISE ESTATÍSTICA	40				

4.5.1	Construção de modelos de Regressão Linear Simples (RLS)	40
5	RESULTADOS E DISCUSSÃO	43
5.1	DETERMINAÇÃO DOS PARÂMETROS GEOMÉTRICOS DA	
	GALANTAMINA	43
5.2	ANÁLISE MULTIVARIADA (PCA E HCA)	44
5.2.1	Análise de componentes principais (PCA)	48
5.2.2	Análise de agrupamento hierárquico (HCA)	57
5.3	ANÁLISE ESTATÍSTICA DAS VARIÁVEIS SELECIONADAS	58
6	CONCLUSÃO	64
	REFERÊNCIAS	66

1 INTRODUÇÃO

A Doença de Alzheimer (DA) é uma desordem neurodegenerativa que afeta prioritariamente a população acima de 65 anos e é considerada a demência mais prevalente no mundo, além de apresentar difícil diagnóstico. Esta doença tem na deficiência de acetilcolina (ACh) pela perda de neurônios colinérgicos; na formação de placas senis pela deposição de peptídeos beta-amiloides (ßA); e na hiperfosforilação da proteína tau formando emaranhados neurofibrilares, as principais causas conhecidas associadas a esta doença.

Vários fármacos têm sido estudados para o tratamento da DA a fim de prevenir e/ou controlar os avanços da doença e seus danos sobre a memória e sobre a função cognitiva dos seus portadores. Atualmente, poucas especialidades farmacêuticas circulam no mercado para a terapêutica da DA e historicamente o homem busca na observação da natureza a cura para diversos males. Assim, os produtos naturais podem representar uma importante contribuição para a descoberta de novos fármacos anti-Alzheimer.

A Galantamina (Nivalin[®]) e a Tacrina (Cognex[®]) são, até o momento, os únicos fármacos do mercado brasileiro usado no tratamento da DA, desenvolvidos a partir de um produto natural. Contudo, devido a elevada toxicidade associada ao uso da Tacrina, este fármaco tem tido seu uso cada vez mais restrito na prática clínica; enquanto que a Galantamina, que é um fármaco pertencente à classe dos alcaloides, têm sido cada vez mais utilizada, por demonstrar significativos resultados no alívio sintomático da DA, o que ocorre principalmente devido ao aumento da atividade dos receptores de ACh. Apesar disso, diante da gravidade do avanço do quadro clínico da DA, faz-se necessário o contínuo desenvolvimento de pesquisas para novos fármacos anti-Alzheimer, na tentativa de encontrar um fármaco que supere a capacidade de aliviar os sintomas da doença, sendo capaz de promover a melhora ou até mesmo a sua cura.

Para descoberta e desenvolvimento de novos fármacos podem ser aplicadas várias metodologias, dentre as quais o isolamento de compostos químicos, oriundos de plantas consagradas pelo uso popular, é um dos mais tradicionais no meio científico; enquanto que o uso da química combinatória e o planejamento racional de fármacos utilizando métodos de modelagem molecular, representam os grandes

avanços da química medicinal da atualidade, que tem no estudo das relações estrutura-atividade dos compostos químicos a sua principal ferramenta de trabalho.

Diante disso, um método computacional que reproduza adequadamente, *in silico*, os parâmetros experimentais relacionados à estrutura e atividade de alcaloides, seria de grande relevância para a comunidade científica, no intuito de facilitar futuros estudos para o desenvolvimento de moléculas promissoras destinadas ao tratamento da DA. Assim, este trabalho busca validar o método computacional que melhor reproduzir os parâmetros geométricos da estrutura cristalográfica protótipo, a ser aplicado na modelagem molecular de alcaloides como candidatos a fármacos para o tratamento da DA, utilizando como composto protótipo a Galantamina, um fármaco já aprovado e amplamente utilizado no tratamento da DA.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Validar o método computacional aplicado à galantamina para planejamento de fármacos anti-Alzheimer.

2.2 OBJETIVOS ESPECÍFICOS

- a) Realizar a modelagem molecular da Galantamina utilizando diferentes métodos computacionais e conjuntos de bases;
- b) Construir a matriz de dados (PCA e HCA) para selecionar os parâmetros geométricos mais importantes relacionados com a região farmacofórica;
- c) Construir modelos de regressão linear simples e avaliar os parâmetros estatísticos a fim de validar o método computacional;
- d) Determinar o método computacional e conjunto de bases que melhor reproduz os dados experimentais da Galantamina através das análises estatísticas;
- e) Propor a aplicação do método validado para futuras determinações de propriedades moleculares e atividade biológica de alcaloides.

3 REFERENCIAL TEÓRICO

3.1 DOENÇA DE ALZHEIMER

Os avanços na qualidade de vida e um maior acesso aos programas de saúde proporcionaram um aumento significativo na expectativa de vida das populações. Esses fatores também geram um impacto negativo nos indivíduos, tendo em vista que algumas doenças estão ligadas ao processo de envelhecimento, em especial as doenças neurodegenerativas. Elas são desordens progressivas do sistema nervoso que afetam a função e a manutenção de populações neuronais específicas (PINHO; POÇAS, 2009).

As doenças neurodegenerativas são caracterizadas pela morte neuronal acentuada e prematura de regiões focais do cérebro mantendo determinados grupos de neurônios intactos (DRAYER et al., 1986; MOSER, 2011). Uma classificação clara de neurodegeneração pode estar baseada na presença de componentes proteicos anormais que se acumulam no cérebro em classes determinadas de neurônios (FERNANDEZ et al., 2007; PINHO; POÇAS, 2009).

Dados recentes indicam que a população brasileira com idade maior que 60 anos deverá chegar à marca de 58,4 milhões no ano de 2060, o que tem atraído a atenção dos sistemas de saúde para várias doenças neurodegenerativas, dentre elas a DA é uma das mais focadas devido à sua grave e progressiva perda de funções psicológicas e motoras que podem levar a total incapacidade do portador, além de ainda não possuir cura conhecida (DIAS et al., 2015).

A DA foi descrita pela primeira vez por Alois Alzheimer em 1907 e, atualmente, é o tipo de demência mais comum entre os idosos, atingindo cerca 35 milhões de pessoas no mundo (DA ROCHA et al., 2011).

3.2 FISIOPATOLOGIA DA DOENÇA DE ALZHEIMER

A DA é a perda acentuada de neurônios na região do hipocampo e do córtex cerebral que provoca uma atrofia com consequente desordem neurodegenerativa (**Figura 1**), levando a prejuízos nas habilidades cognitivas e na memória (DA ROCHA et al., 2011; PICANÇO et al., 2017).



Figura 1 – Aspectos macroscópicos da atrofia do hipocampo e do córtex cerebral.

Fonte: Picanço et al. (2017).

A DA é uma doença multifatorial que ainda não tem suas causas plenamente elucidadas, mas sabe-se que a ação de ß-secretase, ɣ-secretase e a hiperfosforilação da proteína Tau são importantes mecanismos no desenvolvimento da doença por desencadearem a geração de fragmentos insolúveis de peptídeos ßA que se depositam no ambiente extracelular do tecido neuronal, e emaranhados neurofibrilares que se acumulam dentro das células neuronais (**Figura 2**), tudo isso provocando a perda massiva de neurônios e favorecendo assim o envelhecimento cerebral (DIAS et al., 2015; PICANÇO et al., 2017).

A DA é amplamente relacionada à diminuição das taxas de vários neurotransmissores e em especial à ACh. A produção de ACh está diretamente relacionada a estruturas chamadas núcleos basais de Meynert, que produzem a enzima colina acetiltransferase (CAT), a qual participa ativamente da síntese de ACh. Desta forma, em situações patológicas, quando os núcleo de Meynert encontram-se atrofiados há uma expressiva diminuição da produção de CAT e por conseguinte também da ACh (MOSER, 2011; DIAS et al., 2015; PICANÇO et al., 2017).



Figura 2 – Mecanismos de desenvolvimento da Doença de Alzheimer.

A inflamação, a oxidação e as excessivas concentrações de glutamato no cérebro, também, têm se mostrado como outros fatores relevantes na vulnerabilidade neuronal e na fisiopatologia da DA como um todo (DA ROCHA et al., 2011; DIAS et al., 2015). A inflamação é uma característica muito frequentemente encontrada nas regiões do cérebro acometidas pela deposição de ßA, o que se evidencia pela presença de células microgliais, ativadas quando o tecido cerebral é submetido a qualquer tipo de dano, necessitando eliminar e/ou remover células danificadas e/ou mortas deste tecido (DIAS et al., 2015). A oxidação induzida por ßA também se evidencia com a progressão da DA, o que se dá devido ao aumento da formação de espécies reativas de oxigênio, resultando no estresse oxidativo que compromete a função das células neuronais (BARNHAM; MASTERS; BUSH, 2004; SCHWAB; MCGEER, 2008; LIU et al., 2015; PICANÇO et al., 2017). Já a elevação da concentração de glutamato interfere na fisiopatologia da DA por provocar a morte

Fonte: Picanço et al. (2017).

excitotóxica de células nervosas, processo este, comumente mediado por receptores do tipo N-metil-D-aspartato (NMDA), que ao serem ativados permitem o influxo de Ca²⁺ para dentro dos neurônios, e assim ativam os processos neurodegenerativos (DA ROCHA et al., 2011).

O caráter multifatorial torna a DA uma patologia muito complexa e também muito estudada por diversos pesquisadores que buscam elucidar os componentes que podem perturbar as vias bioquímicas que causam a doença, além de entender se as placas amiloides e os emaranhados neurofibrilares induzem a doença ou são causados por ela (MOSER, 2011).

3.3 TRATAMENTOS DISPONÍVEIS

Atualmente, o tratamento para DA não é capaz de oferecer cura efetiva da doença, sendo meramente sintomático e buscando atenuar os sintomas e a evolução da doença através de estímulos colinérgicos como a inibição da AChE e a modulação de receptores nicotínicos pré-sinápticos; da intervenção na formação de fragmentos insolúveis de peptídeo ßA; e do antagonismo de receptores NMDA (DIAS et al., 2015).

Já no que tange o tratamento de estabilização da DA, várias são as abordagens utilizadas na atualidade, dentre as quais pode-se citar: as terapias antiamiloides, antioxidantes, drogas anti-inflamatórias, uso de agonistas muscarínicos, agentes capazes de alterar o processo de fosforilação da proteína tau, bloquadores de apoptose e moduladores da transmissão de glutamato (DA ROCHA et al., 2011; MOSER, 2011).

Uma pesquisa sistemática de Da Rocha e colaboradores (2011) revelou cerca de 70 entidades químicas sendo avaliadas para DA até Dezembro de 2009. Dentre os compostos relacionados estavam: o alfa-tocoferol; o resveratrol; o estrogênio; o valproato; o ácido docosa-hexanóico; a galantamina; a tacrina; a rivastigmina; a memantina e o donepezil. Desses, os cinco primeiros são derivados naturais em ensaio clínico de fase III, a galantamina é o único produto natural aprovado pelo "Food and Drug Administration" (FDA) e os demais são drogas anti-Alzheimer sintéticas também já aprovadas.

Os fármacos atualmente disponíveis no mercado brasileiro são: Tacrina, donepezil, rivastigmina, galantamina, e memantina (DIAS et al., 2015).

A tacrina é um inibidor das colinesterases (IChE) e foi o primeiro fármaco aprovado para o tratamento da DA. É considerada um inibidor reversível não competitivo e não seletivo para acetilcolinesterase (AChE), que apresenta uma eficácia dose dependente, meia vida curta, alta incidência de efeitos adversos e hepatotoxicidade (ALMEIDA, 1998; ENGELHARDT et al., 2005; NASCIMENTO, 2009).

O donepezil é uma droga da família das N-Benzilpiperidinas; foi desenvolvida, sintetizada e avaliada por uma indústria farmacêutica japonesa, sendo aprovado em 1996 para uso no tratamento contra a DA. É um inibidor reversível da AChE altamente seletivo e não competitivo que não apresenta efeitos adversos sérios e melhora significativamente os sintomas da DA (GREENBLATT et al., 1999; SUGIMOTO et al., 2002; ABIB JÚNIOR et al., 2013). Entretanto, por utilizar o sistema citocromo P450 na sua metabolização favorece o desenvolvimento de interações medicamentosas com outras drogas, de forma que a sua associação deve ser cautelosa (ENGELHARDT et al., 2005).

A rivastigmina é uma droga derivada da fisostigmina, é inibidora pseudoirreversível seletiva da AChE e butirilcolinesterase (BChE), o que prolongaria a atividade desta droga no organismo e também os benefícios desse tratamento. É uma droga que apresenta boa atividade e tolerância por parte dos portadores da DA e não envolve o sistema citocromo P450 na sua metabolização, reduzindo a possibilidade de ocorrerem interações medicamentosas com outras drogas, melhorando a cognição e gerando efeitos neuroprotetores nesses pacientes (ENGELHARDT et al., 2003; RACCHI et al., 2004; ENGELHARDT et al., 2005).

A galantamina é um alcaloide terciário de origem natural, extraído de várias espécies de Amarillidaceae, descoberto acidentalmente em 1950, representando um inibidor seletivo, competitivo e reversível da AChE, aprovado em 2001 para o tratamento da DA, já sendo usada ao longo dos tempos para diversas desordens neurológicas. Apresenta ainda uma ação moduladora de receptores nicotínicos e baixa hepatotoxidade, embora a sua associação com outras drogas exija cuidado já que a sua metabolização utiliza o sistema citocromo P450 (BARTOLUCCI et al., 2001; ENGELHARDT et al., 2005; NASCIMENTO, 2009).

A memantina é um antagonista de receptores NMDA, bloqueando-os de forma moderada e não competitiva, o que oferece uma ação protetora ao tecido cerebral em relação a toxicidade do glutamato (WENK et al., 2000). É o único fármaco aprovado

pelos órgãos reguladores internacionais que não atua na inibição da AChE e devido ao seu mecanismo de ação diferenciado, que evita o influxo excessivo de Ca²⁺ para dentro do citoplasma dos neurônios, ela tem sido aplicada nos estágios moderados a graves da DA (ENGELHARDT et al., 2005; GOODMAN; GILMAN, 2006). Além disso, seu metabolismo para excreção é renal e não envolve o citocromo P450, reduzindo o risco de interações medicamentosas e como não parece interferir nos aspectos farmacodinâmicos e farmacocinéticos dos inibidores de colinesterases, vários têm sido os benefícios da sua associação com o donepezil sobre os sintomas cognitivos e comportamentais mais avançados da DA (JARVIS; FIGGITT, 2003; ENGELHARDT et al., 2005).

3.4 PRODUTOS NATURAIS PARA TRATAMENTO DA DOENÇA DE ALZHEIMER

Produtos naturais são amplamente utilizados em diversas terapêuticas ao longo dos tempos e tem se destacado na DA por comumente apresentarem mais de um sítio ativo em suas moléculas, o que é uma vantagem em doenças multifatoriais como a DA (DE MELO; COSTA, 2005; DA ROCHA et al., 2011; DIAS et al., 2015).

Várias substâncias de origem natural têm sido estudadas como potenciais fármacos anti-Alzheimer entre alcaloides, flavonoides, terpenos, cumarinas, polifenóis e outros (DA ROCHA et al., 2011).

Segundo estudos de De Melo e Costa (2005), os alcaloides apresentaram uma boa atividade inibitória da AChE destacando-se a sanguinina, que se mostrou 10 vezes mais ativa que a galantamina, e a Huperzina A, que está em estudos clínicos de fase II e tem se mostrado altamente potente como neuroprotetor, antioxidante e inibidor da morte celular induzida por peptídeos ßA. Neste mesmo estudo, entre os terpenos, destacaram-se a α-onocerina que apresentou atividade anticolinesterásica melhor do que a do donepezil e os ginkgolidos A e B, isolados de *Ginkgo biloba*, que demonstraram capacidade protetora contra a morte neuronal causada por ßA; enquanto o decursinol foi a cumarina mais potente como Inibidores da Acetilcolinesterase (IAChE). Os autores verificaram ainda que os flavonoides são eficazes na inibição da agregação de ßA; e os polifenóis, além de inibir de forma dosedependente esta agregação, também foram capazes de desestabilizar as fibrilas de ßA já formadas.

Atualmente, existem aproximadamente 50 fármacos sendo avaliados quanto a sua eficácia e segurança como potenciais tratamentos na DA. Estes fármacos encontram-se em diferentes fases de testes clínicos e incluem drogas com potenciais tais como: imunoterápicos de anticorpos ßA, anti-inflamatórios não esteróides (AINEs), removedores de placas ßA, estatinas e antioxidantes (DA ROCHA et al., 2011).

Diversas espécies vegetais, incluindo Ginseng, Celastrus paniculatus, Centella asiatica, Curcuma longa, Ginkgo biloba, Huperzia serrata, Lycoris radiate, Galanthus nivalis, Magnolia officinalis, Polygala tenuifolia, Salvia lavandulaefolia, Salvia miltiorrhiza, Coptis chinensis, Crocus sativus, Evodia rutaecarpa, Sanguisorba officinalis, Veratrum grandiflorum e Picrorhiza kurvoa, são relatadas por Da Rocha e colaboradores (2011) como fontes potenciais de extratos ativos, além de mais de 60 metabólitos secundários sob avaliação quanto à eficácia no controle dos sintomas e no impedimento da progressão da DA (**Figura 3**).

Figura 3 – Estruturas químicas de alcaloides isolados de plantas com atividade anti-Azheimer.





Fonte: Adaptado de Da Rocha e colaboradores (2011).

As diversas espécies citadas apresentam moléculas promissoras para o desenvolvimento de fármacos anti-Alzheimer, entretanto o presente estudo focou apenas nas moléculas pertencentes à classe dos alcaloides (**Tabela 1**), visto que é a esta classe que pertence a Galantamina – único fármaco natural que até o momento comprovou sua eficácia no tratamento da DA e por este motivo foi escolhido para servir como padrão terapêutico nos estudos de validação proposto nesta pesquisa.

3.5 MODELAGEM MOLECULAR NO PLANEJAMENTO DE FÁRMACOS

Nas últimas décadas, um grande número de tecnologias tem sido criadas para geração de dados e informações, que quando aplicados ao estudo de sistemas biológicos e química medicinal, podem ser traduzidos em conhecimento e colaborar no processo de planejamento e desenvolvimento de fármacos desde a concepção do alvo biológico até a síntese do medicamento. Neste contexto, atualmente, um grande número de alvos moleculares vem sendo explorados de forma rápida e relativamente precisa através de técnicas como: docagem ("docking"), triagem virtual, ensaios biológicos automatizados em larga escala ("high-throughput screening" - HTS) e química combinatória. Em todos os casos, o primeiro passo é eleger um alvo biológico que seja comprovadamente relevante para a gênese e/ou progressão da doença, para só então selecionar potenciais ligantes capazes de interagir com o alvo biológico de escolha. Nessa seleção, devem ser consideradas as características químicas, estruturais e eletrônicas tanto do sítio de ligação do alvo molecular relacionado à doença, quanto dos seus potenciais ligantes, para que só então sejam propostas modificações moleculares que otimizem a molécula de forma a aumentar a sua atividade biológica (RODRIGUES et al., 2012).

N⁰	Substância	Classe Química	Espécie	Mecanismo de ação	IC50 (AChE)	Referência
(1)	Galantamina	Alcaloide	Lycoris radiata, Galanthus nivalis e Narcissus	Inibidor seletivo de AChE Agonistas de receptores nicotínicos	0,36µM	BARTOLUCCI et al., 2001
			spp			
(2)			Physostigma venenosum	Inibidor da AChE	22 nM	JACKISCH et al 2009
(-)	Fisostigmina	Alcaloide			0,43nM	KALOGIANNIS et al. 2011.
(3)	Huperzina A	Alcaloide	Huperzia serrata	Inibidor da AChE Antagonista de receptores NMDA	0,0104 µg/mL	ZHAO et al., 2013.
(4)	Licoramina	Alcaloide	Lycoris radiata	Inibidor da AChE	>17,815	SHI et al., 2014
					µg/me	2014.
(5)	Jatrorizina	Alcaloide	Coptis chinensis	Inibidor da MAO	2,10 µg/mL	SU et al., 2007.
(6)	Berberina	Alcaloide	Coptis chinensis	Inibidores da MAO	40 µg/mL	LEE et al., 2011.
(7)	Coptisina	Alcaloide	Coptis chinensis	Inibidor da AChE e potencializa a atividade do Fator de Crescimento Neural	0,80µM	JUNG et al., 2009.
(8)	Palmatina	Alcaloide	Coptis chinensis	Inibidor da AChE e potencializa a atividade do Fator de Crescimento Neural	0,52 µM	JUNG et al., 2009.

Tabela 1 – Descrição dos alcaloides com atividade anti-Alzheimer.

(9)	Dehidroevodiamina	Alcaloide	Evodia ruecapa	Inibidor da AChE Reverte a injúria induzida pela escopolamina Aumenta o fluxo sanguíneo cerebral	37,8µM	PARK et al., 1996; JUNG; PARK et al., 2007.
-----	-------------------	-----------	-------------------	--	--------	--

Um dos mais importantes avanços no planejamento e descoberta de novos fármacos tem sido a utilização da modelagem molecular. Ela tem se consolidado como uma ferramenta indispensável não somente no processo de descoberta de novos fármacos, mas também na otimização de um protótipo já existente ou obtido pelo próprio estudo de modelagem molecular (TAYLOR; JEWSBURY; ESSEX, 2002), fornecendo informações importantes para este processo e permitindo a obtenção de propriedades específicas das moléculas que podem influenciar na interação com o receptor (CECHINEL FILHO; YUNES, 1998).

As interações de um fármaco com seu receptor biológico são determinadas por forças intermoleculares, ou seja, interações lipofílicas, polares, eletrostáticas e estéricas. Assim sendo, substâncias que apresentam propriedades terapêuticas, ao interagirem com um alvo específico (uma enzima, um receptor, um canal de íons, um ácido nucleico ou qualquer outra macromolécula biológica), devem possuir uma estrutura tridimensional de forma que as disposições de seus grupos funcionais favoreçam uma maior complementaridade ao sítio de ligação (ARROIO; HONÓRIO; DA SILVA, 2010).

Além disso, a química medicinal também estuda os fenômenos envolvidos no metabolismo e o estabelecimento de relações entre a estrutura química e atividade (SAR, do inglês, *structure activity relationships*). Nesse sentido, destacam-se também os métodos de química quântica para avaliação de QSAR (do inglês, *quantitative structure-activity relationshisps*), que têm como objetivos relacionar a informação em conjuntos de dados (estrutura química e atividade biológica) e criar modelos estatísticos com capacidade preditiva externa (ALONSO; BLIZNYUK; GREADY, 2006), o que torna os estudos de SAR e de QSAR muito importantes para guiar a síntese de novas moléculas com propriedades otimizadas, minimizando o universo de compostos a serem considerados nos programas de triagem biológica (CECHINEL FILHO; YUNES, 1998).

Dentre os atuais desafios que se impõem ao planejamento racional de novos fármacos, pode-se citar: as inflamações crônicas, as disfunções imunológicas e metabólicas, as doenças genéticas e várias outras que tem nos hábitos modernos e na longevidade seus maiores fatores de risco, como é o caso da DA e demais desordens neurodegenerativas (BOLOGNESI, 2013; DIAS et al., 2015).

3.6 MÉTODOS DE QUÍMICA QUÂNTICA APLICADOS À MODELAGEM MOLECULAR DE COMPOSTOS BIOATIVOS

A química quântica utiliza a equação de Schröndinger (1) para descrever propriedades atômicas e sistemas moleculares de interesse tecnológico e/ou biológico através de diferentes métodos teóricos e funções de bases que calculam diversas propriedades termodinâmicas; possíveis conformações de moléculas biológicas; entre outros, no intuito de facilitar e otimizar o processo de desenvolvimento de compostos bioativos, sendo utilizados assim como uma importante ferramenta computacional para orientar a criação de hipóteses sobre os possíveis e/ou prováveis mecanismos de ação desses compostos bioativos e consequentemente para o planejamento racional de fármacos (SANT'ANNA, 2009; SANTOS et al., 2014b; BATISTA; GRANATO; ANGELOTTI, 2016).

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + E_p\Psi = i\hbar\frac{\partial\Psi}{\partial t}$$
(1)

Onde, E_p é a energia potencial na região considerada; $\hbar = h/2\pi$; e m é massa da partícula associada a esta função de onda. $\partial \Psi / \partial t$ representa a derivada parcial, sendo f(x,y) uma função de duas variáveis, com derivada parcial de f em ordem de x calculado admitindo que y é uma constante. $\nabla^2 \Psi$ é o chamado laplaciano de Ψ , em coordenadas cartesianas descrito por $\nabla^2 \Psi = \partial^2 \Psi / \partial x^2 + \partial^2 \Psi / \partial y^2 + \partial^2 \Psi / \partial z^2$, onde $\partial^2 \Psi / \partial x^2$ é a segunda derivada parcial de Ψ em ordem de x, sendo Ψ dependente das variáveis x, y, z e t.

3.6.1 Métodos de química quântica

Há vários métodos computacionais, utilizados nos cálculos de propriedades moleculares, que realizam a derivação direta de descritores eletrônicos a partir da

função de onda molecular, os principais dentre eles são: semi-empírico; ab initio/ Hartree-Fock (HF) e a Teoria do Funcional de Densidade (DFT), cada um deles apresentando um diferente padrão de exatidão associados à utilização ou não de parâmetros empíricos, ao conjunto de base de valência selecionado e ao nível de correlação eletrônica entre os átomos envolvidos na formação da estrutura química do composto bioativo estudado (SANTOS et al., 2014d; SANTOS; CARVALHO; MACEDO, 2016).

Inicialmente, os métodos de química quântica utilizavam a dinâmica quântica para realizar cálculos teóricos que tentavam combinar termodinâmica, cinética química, mecânica quântica e a teoria da ligação de elétrons de valência para construir superfícies de energia potencial capazes de prever a combinação temporária de átomos no estado de transição (MIELKE et al., 2003; SANTOS; CARVALHO; MACEDO, 2016).

Os métodos semi-empíricos são considerados métodos rápidos, de baixo custo computacional e empregam parâmetros empíricos (derivados de dados experimentais) previamente calculados pela equação de Schröndinger, que neste caso (métodos semi-empíricos) considera apenas os elétrons da camada de valência do sistema, já que são eles os responsáveis pela ocorrência das reações químicas e demais interações intermoleculares. Nesses métodos ocorre a simplificação dos cálculos computacionais através da aproximação de valores matemáticos, visto que todos os orbitais atômicos envolvidos são considerados simétricos, desconsiderando assim as diferenças entre os diversos átomos, o que ocorre de forma semelhante nos métodos semi-empíricos mais utilizados como: Austin model 1 (AM1), Parametric method 3 (PM3) e Zerner's Intermediate Neglect of Differencial Overlap (ZINDO), que diferenciam-se entre si basicamente por suas parametrizações, que são as correções dos erros decorrentes do processamento de cada método computacional anteriormente utilizado (LEACH, 2001; SANT'ANNA, 2002; HEHRE, 2003; SANT'ANNA, 2009).

O método HF, desenvolvido por Douglas Hartree, considerava a função de onda como sendo uma combinação linear antissimétrica de produtos de spins-orbitais, baseando-se em cálculos realizados a partir de constantes físicas fundamentais em equações exatas, se diferenciando dos métodos semi-empíricos por não utilizar parâmetros experimentais e nem fazer aproximações nos cálculos realizados, além de considerar todos os elétrons das estruturas químicas analisadas durante a aplicação da equação de Schröndinger (SANTOS et al., 2014b).

O método HF foi o primeiro modelo que empregou a equação de Schröndinger completa e associou conjuntos de funções de bases do tipo Slater e/ou Gaussianas aos seus cálculos, conferindo maior qualidade à predição teórica de uma ampla gama de sistemas químicos; entretanto, por se tratarem de cálculos bem mais elaborados, seus resultados se mostraram mais lentos e desencadearam um custo computacional mais elevado (LEACH, 2001; SANT'ANNA, 2002).

O método DFT pode ser aplicado a estruturas que contenham de 50 a 100 átomos, como é o caso da galantamina que tem 42 átomos, e substitui a função de onda utilizada no método HF pela densidade eletrônica da estrutura química como um todo, considerando cada interação instantânea entre os pares de elétrons com seus spins opostos, diferentemente do que ocorre no método HF que considera uma densidade eletrônica média (LEACH, 1996; HEHRE, 2003).

Para avaliar todas as interações entre os pares de elétrons e seus spins opostos, o método DFT utiliza o teorema de Honeberg e Kohn, que permite descrever de forma aproximada a energia eletrônica total em função da densidade eletrônica através da seguinte equação: $E(p)=E_{KE}(p)+E_{C}(p)+E_{H}(p)+E_{XC}(p)$, onde $E_{KE}(p)$ é a energia cinética; $E_{c}(p)$ é o termo de interação núcleo-elétron; $E_{H}(p)$ é a energia de Coulomb; e $E_{XC}(p)$ contém as contribuições de troca e correlação (BATISTA; GRANATO; ANGELOTTI, 2016; SANTOS; CARVALHO; MACÊDO, 2016).

No método DFT, os cálculos também são expressos em termos de função de bases do tipo Gaussianas e orbitais de Slater, tal qual no método HF; e como, mesmo nos modelos DFT, o funcional de densidade exato não é conhecido, foram desenvolvidos diversos métodos para identificar o conjunto de funcionais de densidade utilizado para expressar cada resultado, sendo que o método B3LYP é um dos mais utilizados por se tratar de um método híbrido que calcula uma parte do funcional através da mecânica quântica e a outra parte através da parametrização pela adição de funcionais de correlação (FORESMAN; FRISCH, 1993; SANTOS; CARVALHO; MACÊDO, 2016).

3.6.2 Métodos de Análise Multivariada

Os métodos de análise multivariada (PCA e HCA) utilizam a avaliação de descritores moleculares para a classificação de compostos químicos e predição de correlação com o potencial de atividade biológica com o mínimo de perdas possível em relação às informações totais analisadas. Para isso, comumente os resultados são organizados em uma matriz de dados que correlaciona os descritores em colunas, com os objetos de estudo em linhas, facilitando a avaliação de dados multidimensionais (BARBOSA et al., 2011; SANTOS et al., 2013; SANTOS et al., 2014a).

Assim, quando os métodos de química quântica são aplicados a estudos de modelagem molecular, torna-se possível obter informações sobre as propriedades moleculares das estruturas químicas analisadas, sendo possível, a partir disto, estabelecer seus padrões de SAR; QSAR e Relação Estrutura Propriedade-Quantitativa (QSPR), tendo em vista que qualquer pequena mudança na estrutura química pode gerar uma maior ou menor alteração na sua atividade biológica, de forma que, quanto maior for a similaridade ou semelhança de uma estrutura química em relação a um modelo padrão, maior deverá ser a probabilidade dessa estrutura garantir e/ou apresentar resultados semelhantes também em termos de atividade biológica (JOHNSON; MAGGIORA, 1990; TODESCHINI; GRAMATICA, 1998; LEACH, 2001; MARTIN, 2010; GUHA, 2012; SANTOS; CARVALHO; MACÊDO, 2016).

3.6.2.1 Análise de componentes principais

A Análise de Componentes Principais (PCA) é considerado um método exploratório que ajuda na elaboração de hipóteses gerais a partir de dados experimentais coletados, separando as informações importantes das informações redundantes (SENA et al., 2000). Neste tipo de análise é feita uma manipulação matemática da matriz de dados fundamentada na correlação entre as variáveis, de forma que as variáveis com alta colinearidade entre si são consideradas redundantes, residindo em um subespaço da matriz de dados total, podendo portanto, ser eliminadas e substituídas por variáveis independentes, capazes de serem melhor descritas através de uma combinação linear das antigas variáveis, possibilitando assim a projeção de amostras muito mais informativas e a visualização das mesmas amostras em uma matriz menos densa e/ou em um gráfico de menor dimensionalidade (PARREIRA, 2003)

Assim, a principal finalidade da PCA é a redução dos dados a partir de combinações lineares das variáveis originais, decompondo a matriz de dados em produtos de vetores que podem ser calculados através do ajuste de mínimos quadrados, dado pela equação: $X = t_1 p'_1 + t_2 p'_2 + \dots + t_h p'_h$, onde t_h são os "scores" e p'_h são os "loadings"; de forma que os "scores" fornecem a composição dos componentes principais em relação às amostras ou objetos e os "loadings" fornecem a composição dos componentes principais em relações entre os objetos, quanto as relações entre as variáveis, permitindo ainda avaliar as suas formas de agrupamento, estimar a influência de cada uma das variáveis em relação a cada objeto analisado e identificar as amostras que não se encaixam no modelo, chamadas de "outliers" (SENA et al., 2000).

Deste modo, a PCA baseia-se na correlação entre o menor conjunto de variáveis, denominadas componentes principais, garantindo o mínimo de perda de informações relevantes em relação às variáveis originais, descritas na matriz de dados inicialmente completa. Este tipo de análise, diferentemente da regressão linear, não tem sua aplicação limitada pela presença de variáveis altamente correlacionadas, demonstrando, portanto, de forma qualitativa as características dos compostos químicos através de "scores", bem como o tipo de correlação estabelecida entre as variáveis analisadas através de "loadings" (FRANKE; GRUSKA, 1996; FERREIRA, 2002; FERREIRA; MONTANARI; GAUDIO, 2002; BARBOSA et al., 2011).

Do ponto de vista matemático, PCA é descrito como:

$$y_{ij} = \sum_{k=1}^{p} P_{ik} a_{kj} + \sum_{k=p+1}^{m} P_{ik}^{(0)} a_{kj}^{(0)}$$

Onde, P_{ik} são vetores ortogonais chamados de componentes principais (PC's) que correspondem às características do composto analisado e são determinados de maneira que a matriz de dados seja adequadamente reproduzida; a_{kj} representa o peso ou a contribuição (importância) do componente principal em relação à variável (FRANKE; GRUSKA, 1996).

Entretanto, segundo Wold (1987), sendo PCA um método ajustado por mínimos quadrados, torna-se essencial realizar o escalonamento da matriz de dados antes de aplicar a técnica aos dados analisados para que todas as variáveis sejam padronizadas e tratadas com a mesma importância, o que comumente é feito através da aplicação da equação:

$$Y_{jm}' = \frac{(y_{jm} - \bar{y}_j)}{S_j}$$

Onde, y_{jm} é o valor da variável *j* para o objeto *m*; \bar{y}_j é a média dos valores das variáveis *j*; e S_j é o desvio padrão dos valores da variável *j* (RIBEIRO, 2001; PARREIRA, 2003; SANTOS; CARVALHO; MACÊDO, 2016).

3.6.2.2 Análise de agrupamento hierárquico

A Análise de Agrupamento Hierárquico (HCA) é uma técnica que examina as distancias entre os pontos de todas as amostras do conjunto de dados e representa essa informação na forma de um gráfico bidimensional conhecido como dendograma, onde é possível visualizar os agrupamentos entre os objetos e as similaridades entre as variáveis. O gráfico em questão é feito calculando-se a distância entre todas as amostras do conjunto em pares, para em seguida definir uma matriz de similaridade *S* onde os elementos (índices de similaridade) variam seus índices entre 0 e 1, podendo ser calculados pela equação a seguir (PARREIRA, 2003).

$$S_{mn} = 1 - \frac{d_{mn}}{d_{max}}$$

Onde, S_{mn} é um elemento de S; d_{max} é a distância máxima para um par de amostras; e d_{mn} é a distância euclidiana entre as amostras m e n, calculadas por: $d_{mn} = [(X_{m1} - X_{n1})^2 + (X_{m2} - X_{n2})^2 + \dots + (X_{mh} - X_{nh})^2]^{1/2}$, sendo X_{mh} um elemento da matriz de dados original; e d_{max} é a maior distância encontrada entre todas as amostras do conjunto (BEEBE; PELL; SEASHOLTZ, 1998; FERREIRA; MONTANARI; GAUDIO, 2002).

Na matriz de similaridade, quanto maior o índice de similaridade, menor será a distância entre os objetos analisados, aumentando a similaridade entre eles, o que costuma ser avaliado através da medida das distâncias Euclidianas entre as variáveis do conjunto de dados (PARREIRA, 2003).

A técnica de HCA é comumente utilizada para investigar as relações existentes no interior de um grupo de variáveis multivariadas aonde ainda não foi realizada nenhum tipo de caracterização, podendo ser do tipo divisiva, quando as amostras de um mesmo grupo são divididas em subgrupos de acordo com suas similiaridades; ou aglomerativa, quando as amostras de grupos inicialmente diferentes são agrupadas formando um grupo final único (BRO, 1997).

Outras aplicações da técnica de HCA são: a identificação de grupos dentro de um conjunto de dados; o teste de hipóteses envolvendo agrupamentos; a identificação de membros de um determinado grupo; e a formação mais conveniente de agrupamentos quando as variáveis apresentam características diversas (BRATCHELL, 1989; PARREIRA, 2003).

Diante do exposto, a técnica de HCA utiliza a matriz de dados, demonstrandoos de forma a acentuar os seus agrupamentos naturais e padrões, classificando-os em suas respectivas categorias e permitindo a comparação de cada amostra e/ou variáveis em relação ao modelo padrão. Este tipo de análise baseia-se em cálculos que comparam as distâncias entre pares de amostras, transformando os dados em uma matriz de similaridade, na qual as distâncias entre as amostras são inversamente proporcionais a sua similaridade, bem como ocorre entre esta similaridade e os tamanhos dos ramos do dendograma que expressam os resultados (FERREIRA, 2002; FERREIRA; MONTANARI; GAUDIO, 2002; SANTOS et al., 2014a; SANTOS et al., 2014c; SANTOS; CARVALHO; MACÊDO, 2016).

3.6.3 Método de Análise Quantitativa

A construção de um modelo estatístico relaciona de forma quantitativa um ou mais objetos analisados com as respectivas propriedades de interesse das variáveis utilizadas, tornando possível a quantificação dessas propriedades onde elas são desconhecidas (PARREIRA, 2003). A quantificação em questão pode ocorrer de forma univariada, quando relaciona um único bloco de respostas à propriedade de interesse; ou multivariada, quando relaciona vários blocos respostas a esta mesma propriedade de interesse (BEEBE; PELL; SEASHOLTZ, 1998; FERREIRA et al., 1999).

Além disso, uma vez construído o modelo, neste tipo de análise é possível realizar determinações simultâneas de vários objetos e/ou variáveis mesmo que

existam interferentes, sendo possível, portanto, diminuir o erro estimado (principalmente quando são utilizadas múltiplas variáveis); estabelecer a relação entre o sinal medido e a propriedade que se deseja quantificar no conjunto de dados; e assegurar a validação do modelo que refletir de forma mais adequada o comportamento do objeto de estudo (PARREIRA, 2003).

No caso da utilização de métodos quantitativos objetivando a validação, como é de interesse para este trabalho, este processo de validação poderá ser feito através de um conjunto de validação externa ou através de validação cruzada; no primeiro caso se utiliza o conjunto de amostras retirados da matriz de dados; enquanto que no segundo caso a matriz de dados é dividida em pequenos grupos, sendo que um dos grupos é removido da matriz original para geração de uma matriz reduzida, composta apenas pelas variáveis que se mostraram mais relevantes para o conjunto total de dados, a qual é decomposta em scores e pesos que, por sua vez, geram uma previsão de resíduos entre os valores reais (experimentais) e os valores estimados (calculados) para cada objeto estudado, o que é repetido para cada um dos pequenos grupos do conjunto de dados (PARREIRA, 2003; MARTINS; FERREIRA, 2013).

Segundo Sena e colaboradores (2000), neste processo de validação a regressão linear é o método mais simples a ser a ser empregado para correlação dos dados, podendo ser utilizado para isso tanto a regressão linear múltipla (RLM), quanto a regressão linear simples (RLS), embora esta última seja o método ideal, visto que Ferreira, Montanari e Gaudio (2002) afirmam que quanto mais variáveis são consideradas no processo, maior a probabilidade dos ajustes matemáticos divergirem dos ajustes naturais das variáveis em relação aos valores observados experimentalmente para a atividade da estrutura química analisada (CRUZ, 2012).

O desafio, entretanto, para o uso da RLS dentro do processo de validação é a seleção de um único tipo de variável relevante ao processo estudado, o que não ocorreu neste trabalho devido a prévia utilização de métodos qualitativos de análise multivariada, os quais têm sido amplamente empregados em diversos estudos de química computacional, demonstrando comprovada eficácia na seleção de variáveis significativas empregadas nos diversos processos de estudos que envolvem a relação estrutura química e atividade biológica (BEEBE; PELL; SEASHOLTZ, 1998; SENA et al., 2000; FREITAS; PAZ; CASTILHO, 2009; MARTINS; FERREIRA, 2013).

Segundo Santos e colaboradores (2013), outras formas de avaliar quantitativamente a eficácia do modelo construído é através da avaliação de: desvio
padrão *S*; coeficiente de correlação *R*; coeficiente de correlação ajustado (R²_A); erro padrão estimado (SEE – do inglês "Standard Error of Estimation"); Soma dos quadrados do erro de predição (PRESS – do inglês "Prediction Residual Error Sum of Squares"); desvio padrão da validação cruzada (S_{PRESS}); coeficiente de correlação de validação cruzada (Q²); média de erro absoluto (MAE – do inglês "Mean Absolute Error"); e média de erro absoluto corrigido (CMAE - do inglês "Corrected Mean Absolute Error"), todos descritos nas equações a seguir.

$$PRESS = \sum_{i=1}^{n} (y - y_i)^2$$
$$S_{PRESS} = \frac{\sqrt{PRESS}}{n - k - 1}$$

$$Q^2 = 1.0 - \frac{\text{PRESS}}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Onde $\bar{y} = y_{média}$

$$MAE = \frac{\sum n |\delta calcul - \delta experim|}{n}$$
$$CMAE = \frac{\sum n |\delta corrigido - \delta experim|}{n}$$
Onde $\delta_{corrigido} = (\delta_{calcul} - a)/b$

Enquanto que Birck, Campos e Melo (2016), além das técnicas matemáticas já citadas acima, consideram ainda o erro quadrático médio (RMSE do inglês "Root Mean Square Error"), descrito na equação a seguir, para as avaliações do modelo preditivo descrito em seus estudos.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y - y_i)^2}$$

4 MATERIAL E MÉTODOS

4.1 SELEÇÃO DO COMPOSTO ESTUDADO - GALANTAMINA

A galantamina (**Figura 4a**) é um produto natural que apresenta atividade inibitória sobre a AChE, disponível no mercado desde o ano de 2000 como fármaco indicado para o tratamento da DA. Este produto tem sido amplamente estudado como protótipo no desenvolvimento de novos fármacos anti-Alzheimer por produzir efeitos benéficos no tratamento da DA através do aumento da liberação de ACh nas sinapses com consequente melhoria da função cognitiva (DE MELO; COSTA, 2005; FORLENZA, 2005; DA ROCHA et al., 2011; DIAS et al., 2015).

O bromidrato de galantamina (**Figura 4b**) foi selecionado para a validação do método computacional, pois é um sal de alcaloide da galantamina que apresenta dados cristalográficos disponíveis (PEETERS; BLATON; RANTER, 1997).

A estrutura cristalográfica do bromidrato de galantamina foi retirada a partir do banco de dados eletrônico International Union of crystallography (IUCr NA1303) de acordo com estudos realizados por Peeters, Blaton e Ranter (1997).

Figura 4 – Estrutura química da Galantamina.

a) numerada conforme utilizada nesteb) na forma de sal – bromidato de galantamina.





4.2 IDENTIFICAÇÃO DA REGIÃO FARMACOFÓRICA E SELEÇÃO DOS PARÂMETROS GEOMÉTRICOS DA GALANTAMINA

Estudos baseados em Bartolucci (2001) utilizaram a acetilcolinesterase de *Torpedo californica* (TcAChE) depositada no Protein Data Bank (PDB) sob código PDB ID: 2ACE (RAVES et al., 1997) e avaliaram um complexo entre a estrutura 3D do fármaco anti-Alzheimer, galantamina, com TcAChE (PDB ID: 3I6M), depositada por Bartolucci e colaboradores (2010), demonstrando assim as interações entre o fármaco e os resíduos de aminoácidos do sítio ativo no complexo galantamina-TcAChE, e exibindo dessa forma o modo de ligação da região farmacofórica da galantamina com os principais resíduos de aminoácidos (Trp84, Phe330, Glu199, Ser200, Gly118) no sítio ativo da TcAChE, possibilitando a identificação da região farmacofórica do fármacofórica do fármaco usado neste estudo (**Figura 5**).

Figura 5 – Principais resíduos de aminoácidos responsáveis pela interação entre a galantamina e a AChE (PDB ID: 1QTI).



Por conseguinte, os dados cristalográficos da galantamina descrito no estudo de Peeters, Blaton e Ranter (1997) foram utilizados para a seleção dos parâmetros geométricos na região farmacofórica identificada anteriormente, tornando possível a validação do método computacional que melhor representa os parâmetros geométricos teóricos e modelagem molecular, conforme estudos anteriores realizados por Santos e colaboradores (2014a), que validaram o método computacional HF/6-31G** propondo uma combinação de métodos de química quântica e análise estatística para estudar os parâmetros geométricos na região do grupo farmacofórico.

4.3 MODELAGEM MOLECULAR DA GALANTAMINA UTILIZANDO DIFERENTES MÉTODOS COMPUTACIONAIS E CONJUNTOS DE BASES.

A estrutura cristalográfica da galantamina foi inicialmente otimizada completamente com os diferentes métodos e conjuntos de base – Semi-empíricos (AM1, PM3 e ZINDO), ab initio/Hartree-Fock (HF/3-21G, HF/3-21G*, HF/3-21G*, HF/6-31G, E HF/6-31G**) e Teoria do funcional da densidade (DFT/3-21G, DFT/3-21G*, DFT/3-21G**, DFT/6-31G, DFT/6-31G*, DFT/6-31G**) – implementados no programa Gaussian 03 (FRISCH et al., 2003), para encontrar o método e conjunto de base com o melhor ajuste entre o tempo computacional e a precisão da informação relativa aos dados experimentais (SANTOS et al., 2013).

Depois da completa otimização, foi possível determinar os parâmetros geométricos teóricos (comprimentos de ligação, ângulos de ligação e ângulos de torção ou diedros entre os átomos) da galantamina na região farmacofórica a fim de avaliar a qualidade da função de onda molecular, comparando os parâmetros geométricos teóricos nos diferentes métodos e conjuntos de bases com os dados experimentais, objetivando validar aquele que melhor reproduzisse os parâmetros geométricos experimentais.

4.4 APLICAÇÃO DAS TÉCNICAS DE ANALISE MULTIVARIADA

A comparação entre os dados experimentais e os parâmetros teóricos obtidos na etapa anterior foi inicialmente realizada com a utilização do programa Pirouette (INFOMETRIX, 2001), que permite a organização dos dados em uma matriz e realiza a análise multivariada através de técnicas conhecidas como Análise de Componentes Principais (PCA) e Análise de Agrupamento Hierárquico (HCA). Estas técnicas são capazes de avaliar os parâmetros organizados na matriz de dados mesmo quando as variáveis analisadas mantêm altos níveis de correlação entre si (SANTOS et al., 2013), permitindo assim uma melhor avaliação dos parâmetros geométricos obtidos da região farmacofórica da galantamina.

Assim, este tipo de análise otimiza a compreensão da relação estabelecida entre os parâmetros geométricos da estrutura cristalográfica com os diferentes métodos computacionais utilizados, de forma a demonstrar tais dados com o mínimo de perda de informação total, ou seja, permitindo a melhor seleção dos principais parâmetros geométricos analisados, bem como o autoescalonamento dos parâmetros geométricos a fim de tratar todas as variáveis com o mesmo peso matemático através da técnica de PCA; além de permitir, através da técnica de HCA, a organização dos métodos e conjuntos de bases em um ranking de similaridades que avalia a reprodutibilidade dos dados experimentais para validação do melhor método computacional, conforme sugerido por Santos e colaboradores (2014a) em estudos anteriores.

4.4.1 Análise de Componentes Principais (PCA)

A matriz de dados foi construída colocando cada parâmetro geométrico retirado da região farmacofórica em sua própria linha (variáveis) até que todos estivessem expressos na matriz em construção. Já nas colunas foram colocados os diferentes métodos e conjuntos de bases utilizados para otimizar a estrutura estudada (descritores), incluindo os valores experimentais, gerando uma matriz de dados medindo 16 x 31.

Posteriormente à construção da matriz de dados, mostrou-se necessário um refinamento da seleção de parâmetros geométricos para otimizar a reprodutibilidade dos dados experimentais; para isso foram feitas várias combinações aleatórias entre as variáveis (parâmetros geométricos – comprimentos de ligações, ângulos de ligação e ângulos de torção) até encontrar o grupo de variáveis que conseguissem separar de forma satisfatória os diferentes métodos computacionais descritos nas linhas da matriz de dados, diminuindo o número de variáveis independentes da matriz e mantendo apenas os parâmetros que demonstrassem as relações mais fortes com as variáveis dependentes ou descritores (métodos computacionais).

Ademais, os parâmetros geométricos selecionados foram correlacionados com os seus desvios padrão, como uma forma de avaliar a importância de cada uma dessas variáveis para a técnica de PCA, a qual foi realizada, com as variáveis autoescalonadas, para construção e avaliação dos 3 componentes principais (PC's) – chamados de PC1, PC2 e PC3, com a prévia avaliação da variância cumulativa dos 3PC's para compreensão do nível de reprodutibilidade da variáveis selecionadas em relação ao total de informações da matriz original.

4.4.2 Análise de Agrupamento Hierárquico (HCA)

Com os parâmetros geométricos da matriz selecionados de forma satisfatória e os métodos computacionais devidamente separados dentro da matriz de dados, seguiu-se para a análise de HCA onde avaliou-se o padrão de similaridade entre os dados experimentais e cada um dos métodos computacionais estudados.

Com este objetivo, os parâmetros geométricos anteriormente selecionados precisaram ser continuamente triados, otimizando cada vez mais os elementos da matriz de dados original, até que os métodos computacionais utilizados para otimizar a estrutura da galantamina estivessem adequadamente agrupados e apenas um deles demonstrasse acentuada similaridade com as medidas geométricas entre os átomos da estrutura química em questão, quando comparados aos dados experimentais da estrutura cristalográfica do modelo padrão descrito (galantamina), favorecendo o processo de validação.

4.5 ANÁLISE ESTATÍSTICA

4.5.1 Construção dos modelos de Regressão Linear Simples (RLS)

Para construção dos modelos de RLS, foi necessário excluir os parâmetros geométricos do tipo comprimento de ligação (Å) visto que este parâmetro apresenta unidade de medida diferente dos demais parâmetros geométricos avaliados, tais como: ângulos de ligação e ângulos de torção ou diedros entre os átomos. O objetivo dessa exclusão é garantir que todas as variáveis tenham um peso igual em termos matemáticos (CARDOSO et al., 2007; CARDOSO et al., 2008; SANTOS et al., 2013). Assim, os parâmetros geométricos selecionados e usados para a análise de RLS, utilizando o programa Statistica 6.1 (STATSOFT, 2004), foram os ângulos de ligação e os ângulos de torção ou diedros.

A análise estatística via RLS foi o método estatístico de escolha devido a sua capacidade de correlacionar os dados dos parâmetros mais relevantes para predizer valores teóricos através da construção de um modelo matemático utilizando um conjunto completo de dados e valores reais como: y_i(r), R²_A, fator de Fisher (F) e erro padrão estimado (SEE). Assim as equações, obtidas após a construção do modelo matemático, foram testadas sobre o seu poder preditivo através do procedimento de validação cruzada, que, segundo Santos e colaboradores (2013), é um método prático e confiável para verificar o poder preditivo de equações de regressão linear (RL).

As equações de RL, por sua vez, foram feitas através da abordagem "leaveone-out", que omite apenas uma das amostras em cada etapa, gerando diferentes modelos de RLS, onde os dados omitidos são chamados de valores preditos (y_i); sendo que, quando associa-se y_i aos valores experimentais ou reais (y), ao número de amostras usada para construir o modelo (n) e ao número de parâmetros geométricos (k), nas fórmulas 1 (PRESS) e 2 (S_{PRESS}) descritas a seguir, torna-se possível avaliar a performance do modelo construído.

$$PRESS = \sum_{i=1}^{n} (y - y_i)^2$$
(1)

$$S_{PRESS} = \frac{\sqrt{PRESS}}{n-k-1}$$
(2)

Outra forma de quantificar a habilidade preditiva do modelo é através dos termos do coeficiente de correlação de validação cruzada (Q²), descrito na equação 3 a seguir, onde $\bar{y} = y_{média}$.

$$Q^2 = 1.0 - \frac{\text{PRESS}}{\sum_{i=1}^n (y_i - \bar{y})^2}$$
 (3)

Assim, considerando que outros parâmetros estatísticos também estão disponíveis para avaliar a precisão do cálculo, torna-se importante ressaltar que nenhum dos parâmetros avaliados é completamente satisfatório se utilizado sozinho, por isso, cada modelo de RLS construído considerou também os valores de *a* e *b*, sendo gerados, portanto, pela seguinte equação:

$$\delta$$
calcul = a + b δ experim (4)

A partir da construção dos modelos de RLS, foram calculados ainda o erro de média absoluto (MAE), o erro de média absoluto corrigido (CMAE), e o erro médio quadrático (RMSE), para correção de erros sistemáticos, ambas descritas, respectivamente, pelas fórmulas 5 e 6 a seguir, onde $\delta_{corrigido} = (\delta_{calcul} - a)/b$.

$$MAE = \frac{\sum n \mid \delta calcul - \delta experim \mid}{n}$$
(5)

$$CMAE = \frac{\sum n \mid \delta \text{corrigido} - \delta \text{experim} \mid}{n} \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y - y_i)^2}$$
(7)

Todos esses cálculos foram realizados objetivando, primeiramente, a seleção do melhor modelo de RLS e, por fim, a validação do melhor método computacional. Assim, o método computacional e conjunto de bases que demonstrou a maior similaridade estatística com os parâmetros geométricos experimentais da galantamina foi o método validado e, por conseguinte, poderá ser utilizado para futuras avaliações de substâncias químicas de origem vegetal da classe dos alcaloides como recurso terapêutico para DA.

5 RESULTADOS E DISCUSSÃO

5.1 DETERMINAÇÃO DOS PARÂMETROS GEOMÉTRICOS DA GALANTAMINA

Os parâmetros geométricos da galantamina foram determinados avaliando os comprimentos de ligação, ângulos de ligação e ângulos de torção entre os átomos da região farmacofórica da estrutura cristalográfica da galantamina, obtida pelo modelo de Peeters, Blaton e Ranter (1997), ver **Tabela 2**.

Os parâmetros geométricos obtidos após a otimização da galantamina sob os diversos métodos e conjuntos de bases foram relacionados na **Tabela 2** e comparados com os parâmetros geométricos do modelo experimental de Peeters, Blaton e Ranter (1997), para obtenção do desvio padrão, o que representou a primeira forma de avaliação dos métodos computacionais no sentido de validar um deles para posteriores estudos de obtenção de fármacos oriundos de alcaloides com atividade anti-Alzheimer, conforme foi realizado em estudos anteriores de Santos e colaboradores (2013) que também utilizaram a medição dos parâmetros geométricos para avaliar a estrutura da artemisinina.

Neste estudo, os resultados de desvio padrão obtidos na comparação dos métodos e conjuntos de bases avaliados em relação ao modelo experimental demonstraram valores de 64,992; 65,089; 89,623; 1,839; 1,839; 2,186; 2,762; 8,145; 2,835; 9,090; 1,819; 1,877; 5,296; 1,460; 1,160; respectivamente para os métodos AM1; PM3; ZINDO; HF3-21G; HF3-21G*; HF3-21G**; HF6-31G; HF6-31G*; HF6-31G**; DFT3-21G; DFT3-21G*; DFT3-21G**; DFT6-31G; DFT6-31G*; DFT6-31G**; exibindo uma grande variação (1,160 - 89,623), sendo que os métodos semiempíricos foram os que demonstraram a maior variação entre si (64,992 – 89,623), imediatamente seguidos pelos métodos DFT (1,160 - 9,090) e HF (1,839 - 8,145); demonstrando que o método HF apresentava a menor variância, em termos de desvio padrão, para os seus diferentes conjuntos de bases. Dessa forma, esta primeira análise demonstra que, no que diz respeito à semelhança dos parâmetros geométricos de cada método computacional em relação ao modelo experimental, os métodos semi-empíricos se mostraram muito inferiores quando comparados aos resultados obtidos pelos métodos DFT e HF; sendo que, em termos de desvio padrão, destacavam-se de forma clara, respectivamente, os métodos DFT6-31G**, DFT631G*, DFT3-21G*, HF3-21G/HF3-21G*, DFT3-21G**, HF3-21G**, HF6-31G e HF6-31G**.

Resultados diferentes foram obtidos por Batista, Granato e Angelotti (2016) que avaliaram a variabilidade de diversas propriedades da artemisinina através da avaliação do desvio padrão, observando baixa variabilidade dos parâmetros geométricos em sua amostra, muito embora tenham chegado a semelhante conclusão de que o método DFT era o método mais adequado a ser utilizado em suas análises.

Apesar do exposto e do desvio padrão ser considerado uma boa medida de dispersão de dados, possibilitando verificar consideráveis diferenças entre os métodos estudados, nota-se que este tipo de avaliação estatística (desvio padrão) sofre muita interferência da variância entre as amostras; além disso, como neste caso foram avaliados parâmetros geométricos de grandezas diferentes (comprimentos de ligação – expressos em centímetros; ângulos de ligação e ângulos de torção ou diedros – expressos em Ângstrom), os resultados podem não fornecer adequada segurança estatística no sentido prático, sugerindo a necessidade de continuidade do processo de avaliação dos parâmetros geométricos (variáveis) em relação aos descritores (métodos computacionais e dados cristalográficos da galantamina), utilizando outras técnicas mais robustas de validação (VIEIRA, 2011).

Assim, diante das importantes limitações expostas, do advento e evolução da química computacional e do crescente desuso de análises do tipo univariadas, técnicas de análise de dados matemáticas e estatísticas mais complexas tornam-se necessárias, principalmente tendo-se em vista o caráter essencial da análise de dados para a obtenção de resultados fidedignos que possibilitem uma melhor visualização espacial das amostras, a exemplo do que ocorre nas análises de dados multivaridadas, capazes de avaliar mais de um tipo de variável, bem como a correlação entre elas, permitindo a extração de uma quantidade muito maior de informações acerca da estrutura química em estudo, que neste caso foi a galantamina (SENA et al., 2000; PARREIRA, 2003; SILVA, 2012).

5.2 ANÁLISE MULTIVARIADA (PCA E HCA)

A análise de PCA e HCA foram as técnicas de análise multivariada utilizadas no processo de validação do método computacional que melhor reproduzisse os parâmetros geométricos da estrutura cristalográfica da galantamina, com o objetivo de facilitar e conferir maior credibilidade a futuros estudos *in silico* de alcaloides oriundos de produtos naturais para o desenvolvimento de fármacos com atividade anti-Alzheimer. Segundo Santos, Carvalho e Macêdo (2016) estes tipos de análises são consideradas mais robustas do que a simples avaliação do desvio padrão entre os métodos, visto que as análises de PCA e HCA são capazes de avaliar de forma mais ampla a matriz de dados construída, considerando simultaneamente todos os parâmetros estruturais da molécula em estudo, bem como as correlações decorrentes da interação entre as variáveis (parâmetros geométricos).

As técnicas de análise multivariada são amplamente utilizadas em diversos tipos de estudos que objetivam correlacionar fatores em matrizes de dados separando amostras significativas, a exemplo dos estudos Sena e Poppi (2000), que utilizaram estas técnicas para avaliar propriedades de solos submetidas a diferentes tipos de manejo; de estudos Stenberg, Pell e Torstensson (1998) que mediram as variáveis químicas, físicas e biológicas em diferentes solos cultivados na Suécia, ambas possibilitando a classificação dos solos quanto a sua fertilidade; e de estudos de Freitas, Paz e Castilho (2009), que avaliaram modelos QSAR 3D em um conjunto de inibidores de butirilcolinesterase humana.

	Semi-empírico					Hartree-	Fock/HF					DFT/E	33LYP			EVD
Parametros	AM1	PM3	ZINDO	3-21G	3-21G*	3-21G**	6-31G	6-31G*	6-31G**	3-21G	3-21G*	3-21G**	6-31G	6-31G*	6-31G**	EXP
Comprimento																
de ligações																
C5O4	1,424	1,406	1,385	1,436	1,436	1,438	1,428	1,398	1,323	1,459	1,459	1,460	1,451	1,421	1,419	1,257
O4C3	1,380	1,380	1,374	1,366	1,366	1,367	1,367	1,347	1,279	1,382	1,382	1,382	1,386	1,386	1,363	1,209
H31O12	0,967	0,950	1,048	0,970	0,970	0,945	0,955	0,950	0,862	1,003	1,003	0,975	0,984	0,984	0,970	0,723
O12C11	1,418	1,409	1,399	1,440	1,440	1,443	1,436	1,405	1,334	1,463	1,463	1,464	1,465	1,465	1,429	1,273
O8C9	1,464	1,454	1,410	1,473	1,473	1,473	1,464	1,430	1,352	1,511	1,511	1,511	1,505	1,505	1,463	1,294
O8C6	1,387	1,382	1,386	1,387	1,387	1,387	1,380	1,355	1,281	1,404	1,404	1,403	1,399	1,399	1,373	1,218
C14C13	1,333	1,330	1,367	1,317	1,317	1,319	1,325	1,321	1,255	1,335	1,335	1,336	1,342	1,342	1,337	1,161
C13C11	1,495	1,502	1,475	1,512	1,512	1,514	1,507	0,510	1,399	1,516	1,516	1,518	1,512	1,512	1,512	1,331
N18C20	1,450	1,485	1,433	1,458	1,458	1,460	1,453	1,447	1,383	1,472	1,472	1,474	1,468	1,468	1,460	1,329
N18C19	1,445	1,479	1,425	1,461	1,461	1,464	1,452	1,445	1,374	1,473	1,473	1,475	1,464	1,464	1,454	1,316
Ângulo de																
ligações																
C5O4C3	115,578	116,810	113,889	120,429	120,429	120,605	121,150	119,205	119,196	117,564	117,564	117,630	118,246	118,246	117,636	116,850
O4C3C2	125,220	126,056	124,042	125,703	125,703	125,577	125,324	126,171	126,826	126,253	126,253	126,096	125,758	125,758	126,321	127,011
H31O12C11	107,384	107,506	107,130	107,686	107,686	106,899	111,426	108,402	110,363	103,857	103,857	103,785	107,289	107,289	105,829	109,226
O12C11C10	111,547	113,315	111,900	111,253	111,253	111,172	111,624	112,384	112,845	110,976	110,976	111,005	111,187	111,187	112,034	112,421
C6O8C9	106,839	107,167	107,030	106,240	106,240	106,323	107,119	106,626	106,394	104,790	104,790	104,865	105,554	105,554	105,761	105,437
O8C6C3	125,155	124,722	125,648	124,404	124,404	124,391	124,866	124,149	124,414	123,774	123,774	123,782	124,141	124,141	123,789	123,524
C15C14C13	125,074	125,176	124,755	124,841	124,841	124,872	124,993	125,048	124,593	124,541	124,541	124,572	124,772	124,772	124,872	124,484

Tabela 2 – Parâmetros geométricos teóricos e experimentais da região farmacofórica da galantamina.

C14C15C16	111,705	110,253	118,851	113,255	113,255	113,147	112,500	112,381	112,213	113,401	113,401	113,277	112,516	112,516	112,523	112,505
C19N18C17	114,154	114,387	113,559	116,453	116,453	116,562	117,117	115,448	113,908	115,450	115,450	115,460	116,243	116,243	114,989	113,740
N18C19H39	109,005	109,391	110,504	109,224	109,224	109,364	109,424	109,393	109,673	108,840	108,840	109,031	109,191	109,191	109,288	109,577
Ângulos de																
torção																
C5O4C3C2	15,178	11,933	6,110	2,472	2,472	2,537	1,026	0,750	2,341	1,611	1,611	1,796	0,839	0,839	0,646	2,287
O4C3C2C1	177,767	178,064	172,817	-179,63	-179,63	-179,65	-179,71	-179,93	-177,85	-179,64	-179,64	-179,67	-179,76	-179,76	-179,97	-177,75
O12C11C10H28	40,035	42,656	55,851	43,423	43,423	43,599	41,406	40,452	41,614	44,403	44,403	44,362	42,812	42,812	41,886	41,740
H31O12C11C10	53,685	59,110	64,090	48,068	48,068	47,881	48,326	53,667	54,558	48,450	48,450	48,645	49,287	49,287	53,374	54,654
C6O8C9C10	-142,45	-137,93	-102,61	-150,72	-150,72	-150,50	-151,22	-152,38	-151,18	-150,08	-150,08	-149,83	-150,64	-150,64	-151,05	-151,12
C9O8C6C3	-171,16	-173,71	171,549	-164,06	-164,06	-164,15	-164,99	-164,13	-167,08	-165,39	-165,39	-165,54	-166,17	-166,17	-165,84	-166,96
O8C6C3O4	-0,220	-1,172	-8,579	0,244	0,244	0,256	0,294	0,184	-1,889	0,317	0,317	0,346	0,411	0,411	0,245	-1,929
C16C15C14C13	125,081	128,967	143,222	118,977	118,977	119,061	118,857	117,898	118,167	119,320	119,320	119,364	119,060	119,060	118,722	118,310
C15C16C17H37	-55,657	-50,828	-62,349	-56,805	-56,805	-57,021	-57,258	-57,891	-54,666	-56,198	-56,198	-56,420	-56,672	-56,672	-57,395	-54,698
C17N18C19H39	49,540	48,540	47,211	46,864	46,864	46,229	45,740	49,199	49,013	49,688	49,688	49,196	48,274	48,274	50,432	49,002
N18C17C16C15	68,077	11,933	6,110	2,472	2,472	2,537	1,026	0,750	2,341	1,611	1,611	1,796	0,839	0,839	0,646	67,105
Desvio Padrão	64,992	65,089	89,623	1,839	1,839	2,186	2,762	8,145	2,835	9,090	1,819	1,877	5,296	1,460	1,160	0,000

5.2.1 Análise de componentes principais (PCA)

Objetivando a compressão da extensa matriz de dados descrita na **Tabela 2**, a análise de PCA foi realizada, demonstrando que vários dos parâmetros geométricos da região farmacofórica se mostravam pouco representativos para o processo de validação, podendo portanto, serem retirados da matriz de dados com o intuito de otimizá-la, permanecendo assim, ao final da otimização, apenas os parâmetros geométricos que mostravam-se representativos (**Figura 6**), sendo capazes de alcançar o objetivo de separar, de forma adequada, os diferentes métodos computacionais estudados em três classes: semi-empíricos (AM1, PM3 e ZINDO); HF (3-21G; 3-21G*; 3-21G*; 6-31G; 6-31G* e 6-31G**) e DFT (3-21G; 3-21G*; 3-21G*; 3-21G**).



Figura 6 – Plotagem dos parâmetros geométricos selecionados pelo PCA.

Nota-se, então, que os resultados de PCA expressos na **Figura 6** denotam as variáveis N18C19; H31O12; C6O8C9 e O4C3C2C1, como sendo as mais relevantes para a separação dos métodos computacionais (**Figura 7**) com o mínimo de perdas em relação às informações totais da matriz de dados. Assim, após a compressão da matriz de dados pela aplicação da técnica de PCA, exibida na **Tabela 2**, tem-se 2 variáveis do tipo comprimento de ligação (N18C19 e H31O12) medindo a distância entre dois átomos; uma variável do tipo ângulo de ligação (C6O8C9) medindo o ângulo

formado entre 3 átomos; e uma variável do tipo ângulo de torção ou diedro (O4C3C2C1) medindo a torção entre 4 átomos.

Figura 7 – Plotagem dos scores obtidos na separação dos métodos computacionais em relação aos dados experimentais após a análise de PCA.



Após a análise de PCA exibida na **Figura 7**, tornou-se possível observar a divisão dos métodos em 3 grupos (semi-empíricos; HF e DFT) de acordo com PC2, tendo em vista que os métodos semi-empíricos se reuniram no quadrante superior; os conjuntos de bases do método DFT se reuniram no quadrante inferior; e os conjuntos de bases do método HF permaneceram na parte central do gráfico, se dividindo entre valores positivos e negativos de PC2, dando destaque ao método e conjunto de base HF6-31G^{**} que se manteve consideravelmente mais próximo do modelo experimental.

Ao analisar simultaneamente a **Figuras 6** e a **Figura 7**, pode-se dizer ainda que, o método DFT e três dos conjuntos de bases do método HF (HF3-21G, HF3-21G* e HF3-21G**) sofrem maior influência das variáveis N18C19 e H31O12, deslocando os métodos para regiões de escore negativo em PC2; enquanto que todos os métodos semi-empíricos e os outros três conjuntos de base do método HF (HF6-31G, HF6-31G*, HF6-31G**) sofrem maior influência das variáveis C6O8C9 e O4C3C2C1, deslocando os métodos para regiões de escore positivo em PC2; além disso, observa-se que a contribuição de cada uma das variáveis para PC1 é inversamente proporcional à eficiência do método, pois o aumento do escore de PC1 resulta no distanciamento do método em relação ao modelo experimental.

A **Tabela 3**, por sua vez, além de exibir a matriz de dados comprimida, após o refinamento feito pela aplicação da técnica de PCA, exibi também uma matriz de correlação entre os parâmetros geométricos e os desvios padrão dos métodos computacionais e experimental estudados, a fim de confirmar a representatividade das variáveis (parâmetros geométricos) selecionadas pela técnica de PCA.

		DESVIO			
MÉTODOS	Comprimento	Comprimento	Ângulo	Torção	
	H31O12	N18C19	C6C8C9	O4C3C2C1	FADRAU
AM1	0,967	1,444	106,839	177,767	64,992
PM3	0,950	1,479	107,167	178,064	65,089
ZINDO	1,048	1,425	107,030	172,817	89,623
HF 3-21G	0,955	1,452	107,119	-179,705	1,839
HF 3-21G*	0,950	1,445	106,626	-179,931	1,839
HF 3-21G**	0,862	1,473	106,394	-177,849	2,186
HF 6-31G	0,970	1,461	106,240	-179,626	2,762
HF 6-31G*	0,970	1,461	106,240	-179,626	8,145
HF 6-31G**	0,945	1,464	106,323	-179,648	2,835
DFT 3-21G	0,984	1,464	105,554	-179,764	9,090
DFT 3-21G*	0,984	1,464	105,554	-179,764	1,819
DFT 3-21G**	0,970	1,454	105,761	-179,968	1,877
DFT 6-31G	1,003	1,473	104,790	-179,639	5,296
DFT 6-31G*	1,003	1,473	104,790	-179,639	1,460
DFT 6-31G**	0,975	1,475	104,865	-179,665	1,160
EXPERIM.	0,723	1,316	105,437	-177,749	0,000
H31O12		0,719	0,032	0,232	0,311
N18C19			0,0002	-0,026	-0,032
C6O8C9				0,584	0,574
O4C3C2C1					0,979

Tabela 3 – Parâmetros geométricos selecionados pelo PCA incluindo desvio padrão dos métodos computacionais e matriz de Correlação de Pearson.

Ao analisar a matriz de correlação exibida na **Tabela 3**, nota-se que as correlações entre os parâmetros geométricos apresentam valores menores ou iguais a 0,584; enquanto que a correlação entre os parâmetros geométricos e os desvios

padrão são menores ou iguais a 0,979. Assim, os resultados obtidos e expressos na matriz de correlação da **Tabela 3** confirmam que os parâmetros geométricos selecionados pela técnica de PCA apresentam as características necessárias para validar o melhor dentre os métodos computacionais avaliados neste estudo, tendo em vista que, em sua maioria, apresentaram valores de correlação maiores que 0,30; considerados como valores de moderado a excelente por vários autores, tais como Callegari-Jacques (2003); Appolinário (2006) e Moura e colaboradores (2015), que também utilizaram a correlação de Pearson nas avaliações realizadas em seus estudos.

Os resultados descritos na **Tabela 4**, por sua vez, também confirmam a análise feita para **Tabela 3** ao exibir os valores de variância e variância cumulativa dos 3 PC's, bem como os valores de contribuição de cada uma das variáveis selecionadas para PC1 e PC2, que segundo Santos e colaboradores (2014a), podem ser descritos como a combinação linear das variáveis selecionadas por PCA e são expressos matematicamente pelas equações a seguir:

$$PC_1 = 0.677 (H_{31}O_{12}) + 0.633 (N_{18}C_{19}) + 0.187 (C_6O_8C_9) + 0.325 (O_4C_3C_2C_1)$$

$$PC_2 = -0.190 (H_{31}O_{12}) - 0.321 (N_{18}C_{19}) + 0.688 (C_6O_8C_9) + 0.623 (O_4C_3C_2C_1)$$

Tabela 4 – PCA do modelo selecionado para os métodos e para os conjuntos de bases.

	Comp	oonentes Princip	oais
—	PC ₁	PC ₂	PC ₃
Variância %	28,654	23,260	6,096
Variância Acumulativa (%)	47,757	86,523	96,683
	Contr	ribuição	
Parâmetros geomé	tricos	PC ₁	PC ₂
H ₃₁ O ₁₂		0,677	-0,190
N18C19		0,633	-0,321
C6O8C9	0,187	0,688	
$O_4C_3C_2C_1$	0,325	0,623	

Podendo-se perceber, então, ao analisar de forma mais completa a **Tabela 4**, que o modelo de PCA foi construído utilizando 3 PC's, sendo que o primeiro componente (PC1) é capaz de descrever 28,654%; o segundo componente (PC2) corresponde a 23,260%; e o terceiro componente (PC3) a 6,096% do total de informações da matriz de dados original, mantendo a capacidade de explicar 96,683% da variância dos dados da matriz original ao se considerar os valores dos 3 PC's. Apesar disso e diante da pequena capacidade de PC3, que apresentou o menor valor de variância, decidiu-se manter apenas PC1 e PC2 na análise de contribuição dos parâmetros geométricos, o que, por sua vez, permite a visualização bidimensional dos PC's, mantendo ainda uma variância acumulativa de 86,523% para explicar a matriz de dados original.

Ainda na **Tabela 4** é possível verificar que os parâmetros geométricos $H_{31}O_{12}$ (0,677) e $N_{18}C_{19}$ (0,633) são os que mais contribuem para PC1; enquanto que os parâmetros geométricos $C_6O_8C_9$ (0,688) e $O_4C_3C_2C_1$ (0,623) são os que mais contribuem para PC2.

Os parâmetros geométricos selecionados por PCA (H31O12, N18C19, C6O8C9 e O4C3C2C1) também foram avaliados neste estudo através da sua variação em relação aos dados experimentais, expressas pelos valores de análise de variação (Δ) e Δ %, exibidos na **Tabela 5**, tal qual o que foi feito por Santos e colaboradores (2013), que também utilizaram este tipo de análise para realizar a avaliação químico-quântica da estrutura química da artemisinina.

Parâmatras	Semi-empírico												
	AM1	Δ_{AM1}	Δ%	PM3	Δ _{PM3}	Δ%	ZINDO	Δ _{ZINDO}	Δ%				
Comprimentos de ligação													
H31O12	0,967	-0,245	-33,849	0,950	-0,227	-31,431	1,048	-0,326	-45,053	0,723			
N18C19	1,445	-0,129	-9,826	1,479	-0,163	-12,418	1,425	-0,109	-8,311	1,316			
Ângulo de ligação													
C6O8C9	106,839	-1,402	-1,330	107,167	-1,730	-1,641	107,030	-1,593	-1,511	105,437			
Ângulo de torção													
O4C3C2C1	177,767	-355,516	200,010	178,064	-355,813	200,177	172,817	-350,566	197,225	-177,749			
				Ab initio	o / Hartree-Fo	ck (HF)				F			
Parametros _	3-21G	Δ _{3-21G}	Δ%	3-21G*	Δ _{3-21G*}	Δ%	3-21G**	Δ _{3-21G**}	Δ%	_ Ехр			
Comprimentos de ligação													
H31O12	0,970	-0,247	-34,226	0,970	-0,247	-34,226	0,945	-0,223	-30,826	0,723			
N18C19	1,461	-0,146	-11,087	1,461	-0,146	-11,087	1,464	-0,149	-11,294	1,316			
Ângulo de ligação C6O8C9	106 240	-0.803	-0 762	106 240	-0.803	-0 762	106 323	-0.886	-0 840	105 437			
Ângulo de torção	100,210	0,000	0,102	100,210	0,000	0,102	100,020	0,000	0,010	100,101			
O4C3C2C1	-179,626	1,877	-1,056	-179,626	1,877	-1,056	-179,648	1,899	-1,068	-177,749			

Tabela 5 – Parâmetros geométricos selecionados pelo PCA incluindo os métodos e conjuntos de bases com suas variações em relação aos parâmetros geométricos experimentais.

Parâmatras	Ab initio / Hartree-Fock (HF)												
	6-31G	Δ _{6-31G}	Δ%	6-31G*	Δ _{6-31G*}	Δ%	6-31G**	Δ _{6-31G**}	Δ%	_ схр			
Comprimentos de ligação H31O12 N18C19	0,955 1,452	-0,232 -0,137	-32,149 -10,398	0,950 1,445	-0,227 -0,130	-31,433 -9,847	0,862 1,374	-0,139 -0,059	-19,283 -4,455	0,723 1,316			
Ângulo de ligação C6O8C9	107,119	-1,682	-1,595	106,626	-1,189	-1,128	106,394	-0,957	-0,908	105,437			
Ângulo de torção O4C3C2C1	-179,705	1,956	-1,100	-179,931	2,182	-1,228	-177,849	0,100	-0,056	-177,749			
Parâmetros	DFT / B3LYP												
	3-21G	Δ _{3-21G}	Δ%	3-21G*	Δ _{3-21G*}	Δ%	3-21G**	Δ _{3-21G**}	Δ%				
Comprimentos de ligação H31O12 N18C19	1,003 1,473	-0,280 -0,157	-38,735 -11,968	1,003 1,473	-0,280 -0,157	-38,735 -11,968	0,975 1,475	-0,252 -0,159	-34,897 -12,087	0,723 1,316			
Ângulo de ligação C6O8C9	104,790	0,647	0,614	104,790	0,647	0,614	104,865	0,572	0,543	105,437			
Ângulo de torção O4C3C2C1	-179,639	1,890	-1,063	-179,639	1,890	-1,063	-179,665	1,916	-1,078	-177,749			

Donômotroo	DFT / B3LYP												
Parametros	6-31G	Δ _{6-31G}	Δ%	6-31G*	Δ _{6-31G*}	Δ%	6-31G**	Δ _{6-31G**}	Δ%	_ Ехр			
Comprimentos de ligação													
H31012	0,984	-0,262	-36,197	0,984	-0,262	-36,197	0,970	-0,248	-34,253	0,723			
N18C19	1,464	-0,148	-11,251	1,464	-0,148	-11,251	1,454	-0,139	-10,558	1,316			
Ângulo de ligação C6O8C9	105,554	-0,117	-0,111	105,554	-0,117	-0,111	105,761	-0,324	-0,307	105,437			
Ângulo de torção O4C3C2C1	-179,764	2,015	-1,134	-179,764	2,015	-1,134	-179,968	2,219	-1,248	-177,749			

Notas: Δ = experimental – teórico; Δ % = (Δ *100)/experimental

Através da Δ exibida na **Tabela 5**, torna-se possível observar que os comprimentos de ligação pouco contribuem para a diferenciação dos modelos, pois em todos os métodos computacionais avaliados os valores de Δ demonstraram boa concordância com os dados experimentais ao se manterem bem próximos de zero, entre -0,059 (Δ HF6-31G**) e -0,326 (Δ ZINDO); já os ângulos de ligação e os ângulos de torção demonstraram ser bem representativos para a diferenciação dos modelos em relação aos dados experimentais, visto que geraram valores de Δ muito diferenciados entre si.

Assim, analisando os métodos semi-empíricos observa-se que eles apresentaram os maiores desvios para o ângulo de ligação (C6O8C9) e principalmente para o ângulo de torção (O4C3C2C1), exibindo, respectivamente, valores de Δ =-1,402 (Δ %=-1,330) e Δ =-355,516 (Δ %=200,010) para AM1; Δ =-1,730 (Δ %=-1,641) e Δ =-355,813 (Δ %=200,177) para PM3; Δ =-1,593 (Δ %=-1,511) e Δ =-350,566 (Δ %=197,225) para ZINDO.

Já analisando o ângulo de ligação (C6O8C9), nota-se que os conjuntos de base do método DFT se mostraram ligeiramente superiores aos conjuntos de base do método HF, exibindo valores de Δ =-0,803 (Δ %=-0,762) para HF3-21G; Δ =-0,803 (Δ %=-0,762) para HF3-21G*; Δ =-0,886 (Δ %=-0,840) para HF3-21G**; Δ =-1,682 (Δ %=-1,595) para HF6-31G; Δ =-1,189 (Δ %=-1,128) para HF6-31G*; Δ =-0,957 (Δ %=-0,908) para HF6-31G**; Δ =0,647 (Δ %=0,614) para DFT3-21G; Δ =0,647 (Δ %=0,614) para DFT3-21G*; Δ =0,572 (Δ %=0,543) para DFT3-21G**; Δ =-0,117 (Δ %=-0,111) para DFT6-31G; Δ =-0,117 (Δ %=-0,111) para DFT6-31G* e Δ =-0,324 (Δ %=-0,307) para DFT6-31G**, o que ocorre, provavelmente, pela considerável similaridade, em termos de Δ e Δ %, de todos os métodos DFT, especialmente nos conjuntos de bases 6-31G, 6-31G* e 6-31G**.

O ângulo de torção (O4C3C2C1), por sua vez, demonstrou relativa equivalência entre os métodos DFT e HF, exibindo valores de Δ=1,877 (Δ%=-1,056) para HF3-21G; Δ=1,877 (Δ%=-1,056) para HF3-21G*; Δ=1,899 (Δ%=-1,068) para HF3-21G**; Δ=1,956 (Δ%=-1,100) para HF6-31G; Δ=2,182 (Δ%=-1,228) para HF6-31G*; Δ=0,100 (Δ%=-0,056) para HF6-31G**; Δ=1,890 (Δ%=-1,063) para DFT3-21G; Δ=1,890 (Δ%=-1,063) para DFT3-21G*; Δ=1,916 (Δ%=-1,078) para DFT3-21G**; Δ=2,015 (Δ%=-1,134) para DFT6-31G; Δ=2,015 (Δ%=-1,134) para DFT6-31G* e Δ=2,219 (Δ%=- 1,248) para DFT6-31G^{**}, dentre os quais se evidencia a superioridade de HF6-31G^{**} por ter demonstrado ótimos valores tanto de Δ , quanto de Δ %.

5.2.2 Análise de agrupamento hierárquico (HCA)

O HCA foi a técnica de análise multivariada utilizada para classificar e validar os métodos computacionais de acordo com a sua similaridade em relação ao modelo experimental; para isso foi realizado um processo de autoescalonamento baseado no método incremental e nas distâncias Euclideanas, gerando um gráfico no formato de um dendograma, demonstrado na **Figura 8**, onde percebe-se que os métodos semiempíricos foram os que exibiram as menores similaridades em relação ao modelo experimental, o que é caracterizado na figura pelo delineamento de ramos longos. Aplicando esta mesma análise para os demais métodos e conjuntos de bases avaliados neste estudo, nota-se ainda que todos os conjuntos de bases dos métodos DFT e HF apresentam ramos curtos, o que vários autores afirmam representar as suas altas similaridades em relação ao modelo experimental (BARBOSA et al., 2011; SANTOS et al., 2013; SANTOS et al., 2014a; SANTOS; CARVALHO; MACÊDO, 2016).

Figura 8 – Dendograma dos métodos computacionais estudados, incluindo os dados dos parâmetros geométricos experimentais da Galantamina.



Além disso, a **Figura 8** também demonstra a divisão dos métodos e conjuntos de bases em 2 grupos, onde um grupo engloba a grande maioria dos métodos e conjuntos de bases avaliados neste estudo (AM1, PM3, ZINDO, HF3-21G, HF3-21G*, HF3-21G**, HF6-31G, HF6-31G*, DFT3-21G, DFT3-21G*, DFT3-21G**, DFT6-31G, DFT6-31G* e DFT6-31G**), enquanto o outro grupo é formado pelo modelo experimental juntamente com o método HF6-31G**, evidenciando a similaridade deste último método em relação aos dados experimentais, visto que ele é o único método computacional que se manteve no grupo do modelo experimental, o que reforça e confirma os resultados expressos pela técnica de PCA.

5.3 ANÁLISE ESTATÍSTICA DAS VARIÁVEIS SELECIONADAS

Após a aplicação das técnicas de analise multivariada, o modelo construído foi avaliado também sobre o aspecto estatístico para avaliar o grau de significância e erro do modelo construído na etapa anterior. Para isso os principais parâmetros estatísticos foram descritos na **Tabela 6**.

Na **Tabela 6**, os valores de R%; R²_A%; F e Q² são responsáveis pela avaliação estatística da capacidade preditiva do modelo construído pelos parâmetros geométricos selecionados no processo de análise multivariada; enquanto que os valores de SEE, MAE, CMAE, RMSE, PRESS e S_{PRESS}, são responsáveis pela avaliação dos erros estatísticos associados ao mesmo modelo construído (SOUZA, 2012; SANTOS, 2013; LEAL, 2015).

Métodos e		а	h	R%	R ² .%	F	SEE	MAF	CMAE	RMSF	PRESS	Spree	O ² %	
COI	njunto	o de bases	u	5	1170	IT A/O	•	ULL		OIII/(E		11200	OFRE35	Q /0
	0	AM1	-2,611	0,755	67,630	42,883	16,016	76,160	19,338	51,720	77,681	126.720,437	18,736	25,322
-ime	oírice	PM3	-3,454	0,757	67,693	42,972	16,071	76,100	20,041	52,797	77,796	127.096,847	18,764	25,188
Ň	eml	ZINDO	7,669	0,454	32,368	5,765	2,224	97,824	39,580	140,451	107,120	240.966,998	25,836	-83,033
_		3-21G	0,087	0,999	99,975	99,948	38221,685	2,305	1,669	1,669	2,195	101,163	0,529	99,950
:)/ At		3-21G*	0,087	0,999	99,975	99,948	38221,685	2,305	1,669	1,669	2,195	101,163	0,529	99,950
K (HF	<u>.o</u>	3-21G**	0,142	0,999	99,972	99,941	33951,064	2,445	1,802	1,796	2,330	113,979	0,562	99,944
-Focl	init	6-31G	0,288	0,995	99,973	99,942	34738,209	2,417	1,795	1,979	2,353	116,255	0,568	99,943
Irtee-		6-31G*	0,268	0,997	99,989	99,977	86395,787	1,533	1,216	1,304	1,503	47,419	0,362	99,977
На		6-31G**	-0,157	0,998	99,998	99,997	580691,157	0,591	0,360	0,601	0,654	8,967	0,158	99,996
		3-21G	0,262	1,001	99,976	99,950	40074,747	2,251	1,548	1,598	2,170	98,874	0,523	99,951
_		3-21G*	0,262	1,001	99,976	99,950	40074,747	2,251	1,548	1,598	2,170	98,874	0,523	99,951
ЗГҮР		3-21G**	0,274	1,001	99,977	99,951	40744,359	2,232	1,525	1,588	2,154	97,448	0,520	99,952
-T/B3		6-31G	0,401	0,998	99,985	99,969	63512,675	1,788	1,304	1,364	1,740	63,591	0,420	99,969
D		6-31G*	0,401	0,998	99,985	99,969	63512,675	1,788	1,304	1,364	1,740	63,591	0,420	99,969
		6-31G**	0,329	0,998	99,991	99,980	101378,762	1,415	1,036	1,101	1,382	40,106	0,333	99,980

Tabela 6 – Principais parâmetros estatísticos calculados nos diferentes métodos e conjuntos de bases.

Notas: As equações de Regressão Linear Simples podem ser construídas por: $\delta_{experimental} = a + b\delta_{calculado}$, com n = 21. Média de erro absoluto: MAE = $\sum_{n} |\delta_{experimental} - \delta_{calculado}|/n$. Média de erro absoluto corrigido: CMAE = $\sum_{n} |\delta_{experimental} - \delta_{corrigido}|/n$, onde $\delta_{corrigido} = (\delta_{calculado} - a)/b$, para correção de erros sistemáticos. Erro médio quadrático estimado: RMSE = $\sqrt{|1/n\sum_{i=1}^{n}(y - y_i)^2|}$.

Na análise estatística dos métodos computacionais, dentre os métodos semiempíricos, o método PM3 foi o que demonstrou os melhores resultados (R%=67,693; R²_A%=42,972; F=16,071; Q²%=25,188; SEE=76,100; MAE=20,041; CMAE=52,797; PRESS=127.096,847; Spress=18,764 e RMSE=77,796), enquanto ZINDO demonstrou os piores resultados (R%=32,368; R²_A%=5,765; F=2,224; Q²%=-83,033; SEE=97,824; MAE=39,580; CMAE=140,451; PRESS=240.966,998; Spress=25,836 e RMSE=107,120). Dentre os conjuntos de bases do método DFT, os piores resultados DFT3-21G/DFT3-21G* (R%=99,976; R²_A%=99.950: foram expressos por Q²%=99,951; F=40.074,747; SEE=2,251; MAE=1,548; CMAE=1,598; PRESS=98,874; SPRESS=0,523 e RMSE=2,170), enquanto que os melhores resultados foram descritos pelo DFT6-31G** (R%=99,991; R²_A%=99,980; F=101.378,762; Q²%=99,980; SEE=1,415; MAE=1,036; CMAE=1,101; PRESS=40,106; S_{PRESS}=0,333 e RMSE=1,382). Dentre os conjuntos de bases do método HF, o que apresentou os piores resultados foi o HF6-31G (R%=99,973; R²_A%=99,942; F=34.738,209; MAE=1,795; Q²%=99,943; SEE=2,417; CMAE=1,979; PRESS=116,255; SPRESS=0,568 e RMSE=2,353), enquanto HF6-31G** demonstrou os melhores resultados (R%=99,998; R²_A%=99,997; F=580.691,157; Q²%=99,996; SEE=0,591; MAE=0,360; CMAE=0,601; PRESS=8,967; Spress=0,158 e RMSE=0,654).

Diante dos dados expostos, nota-se que a análise estatística confirma a grande deficiência preditiva dos métodos semi-empíricos para o modelo construído, tal como já havia sido sugerido pelos resultados das análises multivariadas da etapa anterior deste estudo, visto que, nesta etapa, todos os parâmetros estatísticos dos três métodos semi-empíricos avaliados geraram baixos resultados de capacidade preditiva, como Q²=25,322% (AM1); Q²=25,188% (PM3) e Q²=-83,033% (ZINDO), associados a altos índices de probabilidade de erros, a exemplo dos valores de RMSE=77,681 (AM1); RMSE= 77,796 (PM3) e RMSE=107,120 (ZINDO).

Resultados diferentes foram obtidos em estudos realizados por Albuquerque (2008), que considerou os métodos semi-empíricos como os mais adequados para a modelagem molecular de sistemas nanoscópicos bioativos que avaliam a hemocompatibilidade do dendrímero poliamido amina (PAMAM). Tal diferença na escolha do método mais adequado é associada pelo próprio autor à existência de um bom banco de dados com as adequadas parametrizações dos compostos estudados por ele.

Já nos resultados obtidos no presente estudo, torna-se importante enfatizar que, segundo Martins e Ferreira (2013), o valor tão inferior de Q² exibido na análise de ZINDO, em relação aos demais métodos, confirma que o modelo não foi construído ao acaso, enquanto que Morais (2016) afirma que valores negativos deste parâmetro estatístico significam que o modelo fez previsões com menor probabilidade do que a possibilidade aleatória, demonstrando que tal avaliação apresenta resultados imprecisos com a metodologia utilizada, tornando ZINDO um método inadequado para a realização de futuras predições a respeito da atividade biológica da galantamina e seus derivados utilizando processos de validação cruzada, bem como denota a baixa qualidade dos resultados expressos através dos métodos AM1 e PM3 avaliados sob a mesma metodologia.

Os métodos DFT e HF, por sua vez, apresentaram resultados bastante satisfatórios em termos estatísticos, tanto no que diz respeito à capacidade preditiva do modelo construído, gerando resultados de Q² superiores a 95%, quanto no que tange a probabilidade de erros do processo metodológico, gerando baixos valores de RMSE, MAE, CMAE e S_{PRESS} para todos os conjuntos de bases dos métodos em questão (DFT e HF), dando destaque ao método HF6-31G**, que foi o método que demonstrou os melhores resultados para todos os parâmetros estatísticos avaliados neste estudo, apresentando assim uma considerável superioridade em relação aos demais métodos computacionais estudados (HF3-21G; HF3-21G*; HF3-21G**; HF6-31G; HF6-31G*; DFT3-21G; DFT3-21G*; DFT3-21G**; DFT6-31G; DFT6-31G**).

Diante do exposto, devido a pequena variação entre os diferentes conjuntos de bases dos métodos DFT e HF no que se refere aos valores de Q² [99,943% – 99,996%], os resultados das análises estatísticas, quando isolados das demais análises de validação aqui realizadas, sugerem que qualquer um dos conjuntos de bases dos métodos DFT e/ou HF poderiam ser utilizados em futuros estudos de modelagem molecular de alcaloides análogos da galantamina, visto que apresentariam resultados estatisticamente significativos sobre suas propriedades moleculares, embora respeitando cada qual o seu nível de capacidade preditiva, corroborando com os estudos de Batista, Granato e Angelotti (2016), que também sugerem a descrição satisfatória de mais de um funcional de densidade no estudo das propriedades da artemisinina como potencial antimalárico e anticâncer.

Sabendo disso, e após ter todos os parâmetros estatísticos tabelados (**Tabela 6**), torna-se possível construir as equações de RLS para cada método computacional avaliado, o que segundo Peternelli (2005); Cruz (2012); e Martins e Ferreira (2013), é um método estatístico muito eficaz para verificar a existência de uma relação funcional entre uma variável dependente e uma ou mais variáveis independentes através da construção de um modelo matemático, explicando de forma adequada como se comportam os valores da variável dependente (método computacional) em função das variáveis independentes (parâmetros geométricos da estrutura cristalográfica). Desta forma, após a construção do modelo de RLS para cada método computacional, o melhor resultado foi exibido na **Figura 9**.

Figura 9 – Gráfico de Regressão Linear Simples (RLS) obtido através da comparação entre os dados experimentais e os dados teóricos após otimização pelo método HF6-31G**.



Assim, nota-se que novamente o método HF6-31G** se destacou dentre os demais métodos computacionais avaliados, o que, de acordo com Santos e

colaboradores (2013), se deu devido às maiores similaridades entre os valores teóricos dos parâmetros geométricos selecionados e os valores experimentais da estrutura cristalográfica, mantendo todos os pontos do gráfico, em maior ou menor proporção, sobre a linha da equação ideal, oferecendo assim os melhores resultados através da análise por RLS.

6 CONCLUSÃO

Realizando uma avaliação conjunta de todas as análises realizadas neste estudo (PCA, HCA e análise estatística) para validação do melhor método computacional a ser aplicado na modelagem molecular de alcaloides, tornou-se possível verificar que a técnica de PCA obteve sucesso selecionando os parâmetros geométricos N18C19, H31O12, C6O8C9 e O4C3C2C1 como sendo essenciais para a separação dos métodos computacionais em três grupos: semi-empírico, HF e DFT, dos quais o primeiro era formado de três métodos diferentes; enquanto o segundo e o terceiro eram formados, cada um, por 6 conjuntos de bases diferentes, totalizando os 15 métodos computacionais avaliados neste estudo.

Dentro da análise realizada neste trabalho, o método HF associado ao conjunto de bases 6-31G** foi o que demonstrou os melhores resultados para a maioria das técnicas aplicadas, isto porque, apesar de não ter sido o melhor método na análise isolada do desvio padrão, foi o método que apresentou os melhores resultados na análise de PCA, aproximando-se dos resultados dos dados experimentais, obtidos da estrutura cristalográfica da galantamina; bem como na análise de variância dos desvios padrão, que desde o início já demonstravam que os diferentes conjuntos de bases do método HF eram os que exibiam a menor variância entre si, revelando posteriormente que o método HF6-31G** era o único que apresentava bons resultados de variâncias para todos os parâmetros geométricos essenciais selecionados, além de ser o método que melhor reproduzia, pela técnica de PCA, os parâmetros geométricos experimentais da galantamina.

Semelhante a isso, o método HF6-31G** também mostrou sua superioridade após a análise de HCA, onde revelou ser o método mais próximo dos dados experimentais, sendo o único a permanecer no mesmo grupo do modelo experimental. Desta forma, diante da clara tendência analítica a favor da validação do método HF6-31G**, a análise estatística final corroborou com os resultados das demais análises deste estudo ao validar este método (HF6-31G**) como sendo o melhor método computacional, dentre os avaliados nesta pesquisa, para determinar as propriedades químicas da galantamina, podendo, portanto, ser aplicado a futuros estudos de desenvolvimento de potenciais fármacos análogos da galantamina que apresentem atividade anti-Alzheimer. Dessa forma, os resultados satisfatórios obtidos durante a avaliação da galantamina, validando um método computacional capaz de reproduzir parâmetros experimentais, tende a acelerar o processo de seleção de potenciais fármacos análogos da galantamina.

REFERÊNCIAS

ABIB JÚNIOR, E.; DUARTE, L.F.; PEREIRA, R.; LEMES, A.B.; MORAIS, D.C.; LIMA, L.G.; SAVIO, D.; PENGO, S. Estudo de biodisponibilidade relativa/ bioequivalencia de duas formulações de cloridrato de donepezila em voluntários sadios de ambos os sexos. **RMB**, v. 70, n.1/2, p. 31-35, 2013.

ALBUQUERQUE, C.A. Modelagem molecular aplicada ao desenvolvimento de sistemas nanoscópicos bioativos. **Dissertação (Mestrado)**. Universidade Federal de Itajubá. Programa de Pós-graduação em materiais para engenharia. Minas Gerais, 2008.

ALMEIDA, O.P. Tratamento da Doença de Alzheimer: avaliação crítica sobre o uso de anticolinesterásicos. **Arq. Neuropsiquiatr.**, v. 53, n. 3-B, p. 688-696, 1998.

ALONSO, H.; BLIZNYUK, A.A.; GREADY, J.E. Combining docking and molecular dynamic simulations in drug design. Med. Res. Rev. 2006, 26, 531. **Med. Res. Rev.** p. 531–568, 2006.

APPOLINÁRIO, F. **Metodologia da Ciência: filosofia e prática da pesquisa**. 1ª ed. São Paulo: Editora Thomson, 2006.

ARROIO, A.; HONÓRIO, K.M.; DA SILVA, A.B.F. Propriedades químico-quânticas empregadas em estudos das relações estrutura-atividade. **Quim. Nova**, v. 33, n. 3, p. 694-699, 2010.

BARBOSA, J.P.; FERREIRA, J.E.V.; FIGUEIREDO, A.F.; ALMEIDA, R.C.O.; SILVA, O.P.P.; CARVALHO, J.R.C.; CRISTINO, M.G.G.; CIRÍACO-PINHEIRO, J.; VIEIRA, J.L.F.; SERRA, R.T.A. Molecular modeling and chemometric study of anticâncer derivatives of artemisinin. **J. Serb. Chem. Soc.**, v. 76, n. 9, p. 1263-1282, 2011.

BARNHAM, K.J.; MASTERS, C.L.; BUSH, A.I. Neurodegenerative diseases and oxidative stress. **Nat. Rev. Drug Discov.**, v. 3, p. 205-214, 2004.

BARTOLUCCI, C.; PEROLA, E.; PILGER, C.; FELS, G.; LAMBA, D. Threedimensional Structures of a Complexo f Galanthamine (Nivalin[®]) with Acetylcholinesterase From Torpedo californica: Implication for the Design of New Anti-Alzheimer Drugs. **Protein Struct. Funct. Genet.**, v.42, p. 182-191, 2001.

BARTOLUCCI, C.; HALLER, L.A.; JORDIS, U.; FELS, G.; LAMBA, D. Probing Torpedo californica acetylcholinesterase catalytic gorge with two novel bis-functional galanthamine derivatives. **J. Med. Chem.**, v. 53, n. 2, p. 745-751, 2010.

BATISTA, V.H.S.C.; GRANATO, A.C.; ANGELOTTI, W.F.D. Comparação entre funcionais de densidade no estudo de propriedades eletrônicas de derivados da artemisinina. **Quím. Nova**, v. 39, n. 3, p.298-304, 2016.

BEEBE, K.R.; PELL, R.J.; SEASHOLTZ, M.B. **Chemometrics: a Pratical Guide**. John Wiley & Sons, Nova Iorque, 1998.

BOLOGNESI, M.L. Polypharmacology in a single drug: multitarget drug. **Curr. Med. Chem.**, v.20, p. 1623, 2013.

BRATCHELL, N. Cluster Analysis. Chemom. Intell. Lab. Sys, v. 6, p. 105-125, 1989.

BRO, R. PARAFAC: tutorial and applications. **Chemom. Intell. Lab. Syst.**, v. 38, p.149-171, 1997.

CALLEGARI-JACQUES, S.M. **Bioestatística: princípios e aplicações**. Porto Alegre: Artmed; 2003.

CARDOSO, F.J.B.; COSTA, R.B.; FIGUEIREDO, A.F; BARBOSA, J.P.; NAVA-JUNIOR, I.; PINHEIRO, J.C.; ROMERO, O.A.S. Modeling artemisinin derivatives with potent activity against *P. falciparum* malaria with *ab initio* and PLS methods. *Int. Electon.* **J. Mol. Des**., *v.* 6, n.5, p. 122–134, 2007.

CARDOSO, F.J.B.; FIGUEIREDO, A.F.; LOBATO, M.S.; MIRANDA, R.M.; ALMEIDA, R.C.O.; PINHEIRO, J.C. A study on antimalarial artemisinin derivatives using MEP maps and multivariate QSAR. **J. Mol. Model.** *v. 14*, n.1, p. 39–48, 2008.

CHEMPLUS: Modular Extensions for HyperChem Release 6.02, Molecular Modeling for Windows, HyperClub, Inc., Gainesville, 2000.

CECHINEL FILHO, V.; YUNES, R. A. Estrategies for obtaining pharmacologically active compounds from medicinal plants: concepts about structural modification for improve the activity. **Quím. nova**, v. 21, n. 1, p. 99–105, 1998.

CRUZ, D.S. Estudos de QSAR 2D e 3D para derivados de aminoimidazóis, aminohidantoínas e aminipiridinas com atividade inibitória sobre a enzima betasecretase humana. **Dissertação (Mestrado). Universidade Estadual de Feira de Santana**. Programa de Pós Graduação em Biotecnologia. Bahia, 2012.

DA ROCHA, M.D. Viegas, F.P.D., Campos, H.C.; Nicastro, P.C.; Fossaluzza, P.C.; Fraga, C.A.M.; Barreiro, E.J.; Viegas Jr, C. The role of natural products in the discovery of new drug candidates for the treatment of neurodegenerative disorders II: Alzheimer's disease. **CNS Neurol. Disord. Drug Targets**, v. 10, n. 2, p. 251–270, 2011.

DE MELO, G.O.; COSTA, S.S. Produtos naturais para o tratamento da Doença de Alzheimer: Promessa e Desafio. **Rev. Fitos**, v. 1, n. 2, p. 41-47, 2005.

DENNINGTON, R.; KEITH, T.; MILLAM, J. GaussView, Version 5, Semichem Inc., Shawnee Mission KS, 2009.

DIAS, K.S.T.; DE PAULA, C.T.; RIQUIEL, M.M.; LAGO, S.T; COSTA, K.C.M; VAZ, S.M.; MACHADO, R.P.;LIMA, L.M.S.; VIEGAS JUNIOR, C. Recent Applications of the Multi-Target Directed Ligands Approach for the Treatment of Alzheimer's Disease. **Rev. Virtual Quím.**, v. 7, n. 2, 2015.

DRAYER, B.P; OLANOW, W.; BURGER, P.; JOHNSON, G.A; HERFKENS, R; RIEDERER S. Parkinson plus syndrome: diagnosis using high field MR imaging of brain iron. **Radiology**, v. 159, n. 2, p. 493-498, 1986.

ENGELHARDT, E.; BERTOLUCCI, P.; BRITO-MARQUES, P. Eficácia da rivastigmina no desempenho cognitivo de pacientes com doença de Alheimer provável leve a moderada: resultados do braço brasileiro de um estudo aberto multicêntrico. **Arq. Neuropsiquiatr.**, v. 61(Suppl), p. S54-S55, 2003.

ENGELHARDT, E.; BRUCKI, S.M.T.; CAVALCANTI, J.L.S.; FORLENZA, O.V.; LAKS, J.; VALE, F.A.C. Tratamento da Doença de Alzheimer: Recomendações e sugestões do Departamento Científico de Neurologia Cognitiva e do Envelhecimento da Academia Brasileira de Neurologia. **Arq. Neuropsiquiatr.**, v. 63, n.4, p.1104-1112, 2005.

FERNANDEZ, L.L; FORNARI, L. H. T; BARBOSA, M. V; SCHRODER, N. Ferro e neurodegeneração. **Sci. Med.**, v. 17, p. 218-224, 2007.

FERREIRA, M.M.C. Multivariate QSAR. J. Braz. Chem. Soc., v. 13, n. 6, p. 742-753, 2002.

FERREIRA, M.M.C.; ANTUNES, A.M.; MELGO, M.S.; Volpe, P. L. O. Quimiometria I: calibração multivariada, um tutorial. **Quím. Nova**, v. 22, p. 724-731, 1999.

FERREIRA, M.M.C.; MONTANARI, C.A.; GAUDIO, A.C. Seleção de Variáveis em QSAR. **Quím. Nova**, v. 25, n. 3, p. 439-448, 2002.

FORESMAN, J.B.; FRISCH, A.E. **Exploring Chemistry with Electronic Structure Methods: A Guide to Using Gaussian**. Gaussian, Inc., USA, 1993.

FORLENZA, O. V. Tratamento farmacológico da doença de Alzheimer. **Rev. Psiq. Clín.**, v.32, n.3, p.137-148, 2005.

FRANKE, R.; GRUSKA, A. Multivariate data analysis of chemical and biological data in: van de WATERBEEMD, H. (Eds.). **Chemometrics Methods in Molecular Design**. New York: VHC, p.115, 1996.

FREITAS, H.F.; PAZ, O.S; CASTILHO, M.S. Estudos de QSAR 3D para um conjunto de inibidores butirilcolinesterase humana. **Quím. Nova**, v. 32, n. 8, p. 2114-2121, 2009.

FRISCH, M. J. G.; TRUCKS, W.; SCHLEGEL, H. B.; SCUSERIA, G. E.; ROBB, M. A.; CHEESEMAN, J. R.; MONTGOMERY, J. A.; VREVEN, JR. T.; KUDIN, K. N.; BURANT, J. C.; MILLAM, J. M.; IYENGAR, S. S.; TOMASI, J.; BARONE, V.; MENNUCCI, B.; COSSI, M.; SCALMANI, G.; REGA, N.; PETERSSON, G. A.; 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, J. E.; HRATCHIAN, H. P.; CROSS, J. B.; ADAMO, C.; JARAMILLO, J.; GOMPERTS, R.; STRATMANN, R. E.; YAZYEV, O.; AUSTIN, A. J.; CAMMI, R.; POMELLI, C.; OCHTERSKI, J. W.; AYALA, P. Y.; MOROKUMA, K.; VOTH, G. A.;

SALVADOR, P.; DANNENBERG, J. J.; ZAKRZEWSKI, V. G.; DAPPRICH, S.; DANIELS, A. D.; STRAIN, M. C.; FARKAS, O.; MALICK, D. K.; RABUCK, A. D.; RAGHAVACHARI, K.; FORESMAN, J. B.; ORTIZ, J. V.; CUI, Q.; BABOUL, A. G.; CLIFFORD, S.; CIOSLOWSKI, J.; STEFANOV, B. B.; LIU, G.; LIASHENKO, A.; PISKORZ, P.; KOMAROMI, I.; MARTIN, R. L.; FOX, D. J.; KEITH, T.; AL-LAHAM, M. A.; PENG, C. Y.; NANAYAKKARA, A.; CHALLACOMBE, M.; GILL, P. M. W.; JOHNSON, B.; CHEN, W.; M. WONG, W.; GONZALEZ, C.; POPLE, J. A. Gaussian 03, Revision A.1; Gaussian, Inc.: Pittsburgh PA, 2003.

GOODMAN, L.S.; GILMAN, A. **Godman e Gilman: As bases farmacológicas da terapêutica**. 11^a ed. Rio de Janeiro: MacGraw-Hill Interamericana do Brasil; 2006.

GREENBLATT, H.M.; KRYGER, G.; LEWIS, T.; SILMAN, I.; SUSSMAN, J.L. Structure of acetylcholinesterase complexed with (3)-galanthamine at 2.3A resolution. **FEBS Letters**, v.463, p. 321-326, 1999.

GUHA, R. Exploring structure-activity data using the landscape paradigma. **WIRES Comput. Mol. Sci.**, v. 2, n. 6, p. 829-841, 2012.

HEHRE, W.J. A Guide to Molecular Mechanics and Quantum Chemical Calculations: Wavefunction, Inc. 2003.

INFOMETRIX, Inc. Pirouette Software. Version 3.01. [S.I.]: Infometrix Incorporated, 2001.

JACKISCH, R.; FÖRSTER, S.; KAMMERER, M.; ROTHMAIER, A.K.; EHRET, A.; ZENTNER, J.; FEUERSTEIN, T.J. Inhibitory potency of choline esterase inhibitors on acetylcholine release and choline esterase activity in fresh specimens of human and rat neocortex. **J. Alzheimers Dis.**, v. 16, n. 3, p. 635-647, 2009.

JARVIS, B; FIGGITT, D.P. Memantine. Drugs Aging, v. 20, p. 465-476, 2003.

JOHNSON, M.; MAGGIORA, G. Concepts and Applications of Molecular Similarity. **John Wiley & Sons**. New York, 1990.

JUNG, H.A.; MIN, B.S.; YOKOZAWA, T.; LEE, J.H.; KIM, Y.S.; CHOI, J.S. Anti-Alzheimer and antioxidant activities of Coptidis Rhizoma alkaloids. **Biol. Pharm. Bull.**, v. 32, n. 8, p. 1433-1438, 2009.

JUNG, M.; PARK, M. Acetylcholinesterase inhibition by flavonoids from *Agrimonia pilosa*. **Molecules**, v. 12, n. 9, p. 2130–2139, 2007.

KALOGIANNIS, M.; HSU, E.; WILLIE, J.T.; CHEMELLI, R.M.; KISANUKI, Y.Y.; YANAGISAWA, M.; LEONARD, C.S. Cholinergic Modulation of Narcoleptic Attacks in Double Orexin Receptor Knockout Mice. **PLoS One**, v. 6, n.4, 2011.

LEACH, A.R. **Molecular Modelling Principles and Applications**. Longman, Essex, 1996.

LEACH, A.R. Molecular Modelling: Principles and Applications. Prentice Hall.

Englewood Cliffs. NJ, 2001.

LEAL, F.D. QSAR 2D/3D e docagem de inibidores da enzima DYRK1A como potenciais fármacos da doença de Alzheimer. **Tese (Doutorado). Universidade Federal do Rio de Janeiro** (UFRJ)/Faculdade de Farmácia, 159p., 2015.

LEE, I.S.; AHN, B.R.; CHOI, J.S.; HATTORI, M.; MIN, B.S.; BAE, K.H. Selective cholinesterase inhibition by lanostane triterpenes from fruiting bodies of Ganoderma lucidum. **Bioorg. Med. Chem. Lett.**, v. 21, p. 6603–6607, 2011.

LIU, J.; LIU, Z.; ZHANG, Y.; YIN, F. A Novel Antagonistic Role of Natural Compound Icariin on Neurotoxicity of Amyloid β Peptide. **Indian J. Med. Res.**, v. 142, p. 190-195, 2015.

MARTIN, Y. **Quantitative Drug Design: A Critical Introduction**. Boca Raton. 2^a Ed. FL: CRC Press, 2010.

MARTINS, J.P.A; FERREIRA, M.M.C. QSAR modeling: um novo pacote computacional Open Souce para gerar e validar modelos QSAR. **Quím. Nova**, v.36, n. 4, p. 554-560, 2013.

MIELKE, S.L.; PETERSON, K.A.; SCHWENKE, D.W.; GARRET, B.C.; TRUHLAR, D.G.; MICHAEL, J.V.; SU, M.C.; SUTHERLAND, J.W. H+H₂ Thermal Reaction: A Convergence of Theory and Experiment. **Phys. Rev. Lett.**, v. 91, n. 6, p. 63201-63204, 2003.

MORAIS, A.L. **SQL Server 2014 para iniciantes: Guia de referência prático para aprender a utilizar o SQL Server 2014**. Ed. Ciencia Moderna. 1ª Edição, 354p., 2016.

MOSER, D.A. Mal de Alzheimer na primeira fase: contribuições da psicolinguística. **Tese (Doutorado). UNIVERSIDADE FEDERAL DE SANTA CATARINA.** PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA. Florianópolis, 2011.

MOURA, F.A.; JACINTO, C.A.C.; FELICISSIMO, C.T.; PRUDÊNCIO, M.V.; MERCADANTE, L.A.; CUNHA, S.A. Concordância e correlação entre três métodos distintos para quantificação da altura do salto vertical. **Rev. Bras. Educ. Fís. Esp.**, v. 29, n.1, p. 25-34, 2015.

NASCIMENTO E.C.M. Determinação teórica de propriedades relevantes para a atividade de inibidor da acetilcolinesterase. **Dissertação (Mestrado em Química). UNIVERSIDADE DE BRASÍLIA.** PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA. Brasília - DF, 2009

PARK, C.H.; KIM, S.H.; CHOI, W.; LEE, Y.J.; KIM, J.S.; KANG, S.S.; SUH, Y.H. Novel anticholinesterase and antiamnesic activities of dehydroevodiamine, a constituent of Evodia rutaecarpa. **Planta Med.**, v. 62, n. 5, p.405–409, 1996.

PARREIRA, T.F. Utilização de métodos quimiométricos em dados de natureza multivariada. **Dissertação (Mestrado). Universidade Estadual de Campinas**
(UNICAMP), Instituto de Química. Campinas, SP, [s.n.], 248p., 2003.

PEETERS, O.M.; BLATON, N.M.; DE RANTER, C.J. (-)-Galanthaminium Bromide. Acta Crystallogr. C, v. 53, n. 9, p. 1284-1286, 1997.

PETERNELLI, L.A. **INF 162 - Estatística I**. Material didático institucional, Cap. 9, Universidade Federal de Viçosa, 2005. Disponível em: http://www.dpi.ufv.br/~peternelli/inf162.www.16032004/materiais/CAPITULO9.pdf

PICANÇO, L.C.S.; OZELA, P.F.; BRITO, M.F.B.; PINHEIRO, A.A.; PADILHA, E.C.; BRAGA, F.S.; SILVA, C.H.T.P.; SANTOS, C.B.R.; ROSA, J.M.C.; HAGE-MELIM, L.I.S. Alzheimer's Disease: A Review from the Pathophysiology to Diagnosis, New Perspectives for Pharmacological Treatment. **Curr. Med. Chem.**, v. 24, p. 1-19, 2017.

PINHO, G.M.; POÇAS, R.O. Modelos de Neurodegeneração e Stress do Retículo Endoplasmático em Drosophila. [s.l.]. **Dissertação (Mestrado em Biologia Molecular e Celular). UNIVERSIDADE DE AVEIRO**, 2009.

RACCHI, M.; MAZZUCCHELLI, M.; PORRELLO, E.; LANNI, C.; GOVONI, S. Acetylcholinesterase inhibitors: novel activities of old molecules. **Pharm. Res.**, v. 50, n. 4, p. 441-451, 2004.

RAVES, M.L.; HAREL, M.; PANG, Y.-P.; SILMAN, I.; KOZIKOWSKI, A.P.; SUSSMAN, J.L. Structure of acetylcholinesterase complexed with the nootropic alkaloid, (-)-huperzine A. **Nat. Struct. Biol.**, v. 4, n.1, p. 57-63, 1997.

Ribeiro, F.A.L. Aplicação de métodos de análise multivariada no estudo de hidrocarbonetos policíclicos aromáticos. **Dissertação (Mestrado). UNICAMP**, 2001.

RODRIGUES, R.P. MANTOANI, S.P.; DE ALMEIDA, J.R.; PINSETTA, F.R.; SEMIGHINI, E.P.; DA SILVA, V.B.; DA SILVA, C.H.P. Estratégias de Triagem Virtual no Planejamento de Fármacos. **Rev. Virt. Quim.**, v. 4, n. 6, p. 739–776, 2012.

SANT'ANNA, C.M.R. Glossário de Termos Usados no Planejamento de Fármacos (Recomendações da IUPAC para 1997). **Quím. Nova**, v. 25, p. 505, 2002.

SANT'ANNA, C.M.R. Métodos de modelagem molecular para estudo e planejamento de compostos bioativos: uma introdução. **Rev. Virt. Quim.**, v. 1, n. 1, p. 49-57, 2009.

SANTOS, C.B.R.; LOBATO, C.C.; VIEIRA, J.B.; BRASIL, D.S.B; BRITO, A.U.; MACÊDO, W.J.C.; CARVALHO, J.C.T.; PINHEIRO, J.C. Evaluation of quantum chemical methods and basis sets applied in the molecular modeling of artemisinin. **Comput. Mol. Biosci.**, v. 3, p. 66-79, 2013.

SANTOS, C.B.R.; VIEIRA, J.B.; FORMIGOSA, A.S.; COSTA, E.V.M.; PINHEIRO, M.T.; SILVA, J.O.; MACÊDO, W.J.C.; CARVALHO, J.C.T. Validation of computational methods applied in molecular modeling of artemisinin with antimalarial activity. **J. Comput. Theor. Nanosci.**, v.11, n. 3, p. 1-9, 2014a.

SANTOS, C.B.R.; LOBATO, C.C.; BRAGA, F.S.; MORAIS, S.S.S.; FERNANDES, C.P.; BRASIL, D.S.B.; HAGE-MELIM, L.I.S.; MACÊDO, W.J.C.; CARVALHO, J.C.T. Application of Hartree-fock method for modeling of bioactive molecules using SAR and QSPR. **Comput. Mol. Biosci.**, v. 4, p. 1-24, 2014b.

SANTOS, C.B.R.; LOBATO, C.C.; SOUSA, M.A.C.; MACÊDO, W.J.C.; CARVALHO, J.C.T. Molecular Modeling: Origin, Fundamental Concepts and Applications Using Structure-Activity Relationship and Quantitative Structure-Activity Relationship. **Rev. Theor. Sci.**, v. 2, n. 2, p. 91-115, 2014c.

SANTOS, C.B.R.; VIEIRA, J.B.; LOBATO, C.C. ; HAGE-MELIM, L.I.S.; SOUTO, R.N.P.; LIMA, C.S.; COSTA, E.V.M.; BRASIL, D.S.B.; MACÊDO, W.J.C.; CARVALHO, J.C.T. A SAR and QSAR Study of New Artemisinin Compounds with Antimalarial Activity. **Molecules**, v. 19, p. 367–399, 2014d.

SANTOS, C.B.R.; CARVALHO, J.C.T.; MACÊDO, W.J.C. **Planejamento Racional de Fármacos Antimaláricos**. Química Medicinal Aplicada. Novas Edições Acadêmicas. Deutschland/Niemcy: OmniScriptum GmbH & Co. KG. Saarbrucken; 2016.

SCHWAB, C.; MCGEER, P. L. Inflamatory aspects of Alzheimer Disease and other neurodegenerative desorders. **J. Alzheimer Dis.**, v. 13, p. 359-369, 2008.

SENA, M.M.; POPPI, R.J.; FRIGHETTO, R.T.S.; VALARINI, P.J. Avaliação do uso quimiométrico em análise de solos. **Quím. Nova**, v. 23, n.4, 2000.

SHI, Z.Q.; SONG, D.F.; LI, R.Q.; YANG, H.; QI, L.W.; XIN, G.Z.; WANG, D.Q.; SONG, H.P.; CHEN, J.; HAO, H.; LI, P. Identification of effective combinatorial markers for qualitystandardization of herbal medicines. **J. Chromatogr. A**, v. 1345, p. 78-85, 2014.

SILVA, C.P. Aplicação da quimiometria para caracterização química de combustíveis tipo MTR por fluorescência de Raios X. **Dissertação (Mestrado). Instituto de Pesquisas Energéticas e Nucleares**, Autarquia associada à Universidade de São Paulo (USP). São Paulo, SP, 52p., 2012.

SOUZA, S.D. Estudo de inibidores da colinesterase aplicando técnicas de QSAR-2D (HQSAR) e docking molecular. **Tese (Doutorado). Universidade Federal do Rio de Janeiro** (UFRJ)/Faculdade de Farmácia, 114p., 2012.

STASOFT, Inc. STATISTICA (Data Analysis Software System). Version 6.1. [S.I.]: Stasoft Incorporated, 2004. Disponível em: http://www.stasoft.com.

STENBERG, B.; PELL, M.; TORSTENSSON, L. Integrated evaluation of variations in biological, chemical and physical soil properties. **Ambio**, v. 27, p. 9-15, 1998.

SU, C.R.; Ueng, Y.F.; Dung, N.X.; Reddy, M.V.B.; Wu, T.S. Cytochrome P3A4 Inhibitors and Other Constituents of Fibraurea tinctorial. **J. Nat. Prod.**, v. 70, n. 12, p. 1930–1933, 2007. SUGIMOTO, H.; OGURA, H.; ARAI, Y.; IIMURA, Y.; YAMANISHI, Y. Research and development of donepezil hydrochloride, a new type of acetylcholinesterase inhibitor. **JPN J. Pharmacol.**, v.89, p. 7-20, 2002.

TAYLOR, R.D.; JEWSBURY, P.J.; ESSEX, J. W. A review of protein-small molecule docking methods, **J. comput. Aided Mol. Des.**, v. 16, n. 3, p. 151-66, 2002.

TODESCHINI, R.; GRAMATICA, P. New 3D molecular descriptors: the WHIM and QSAR applications. **Perspct. Drug Discov.**, v. 9-11, n.26, p. 355-380, 1998.

VIEIRA, S. Introdução à bioestatística. Elsevier, RJ, recurso digital: il., 4^a ed., 345 p., 2011.

WENK, G.L; QUACK, G.; MO EBIUS, H.J.; et al. No interaction of memantine with acethylcholinesterase inhibitors approved for clinical use. **Life Sci.**,v. 66, p. 1079-1083, 2000.

Wold, S. Principal Component Analysis. **Chemom. Intell. Lab. Sys.**, v. 2, p. 37-52, 1987.

ZHAO, X.M.; WANG, Z.Q.; SHU, S.H.; WANG, W.J.; XU, H.J.; AHN, Y.J.; WANG, M.; HU, X. Ethanol and Methanol Can Improve Huperzine a Production From Endophytic Colletotrichum Gloeosporioides ES026. **PLoS One**, v. 8, n. 4, 2013.