



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE FÍSICA

TESE DE DOUTORADO

**Reconhecimento Molecular dos Compostos Naturais
Cinamaldeído, Curcumina e Ácido Rosmarínico
pelo Peptídeo β -Amilóide(1-42)**

Camila D'Avila Braga Sette

Brasília
março/2020

Camila D'Avila Braga Sette

**Reconhecimento Molecular dos Compostos Naturais
Cinamaldeído, Curcumina e Ácido Rosmarínico pelo
Peptídeo β -Amilóide(1-42)**

Tese submetida ao Instituto de Física da
Universidade de Brasília como parte dos
requisitos para a obtenção do grau de
Doutora em Física.

Orientador: Prof. Dr. Ricardo Gargano

Co-orientadora: Dra. Alessandra Sofia Kiametis Wernik

Dedico este trabalho à minha família e ao meu marido, por serem minha fortaleza, e ao Luiz Augusto (*in memoriam*), para que seus sonhos possam ser realizados através dos sonhos de quem te ama.

"The block of granite which was an obstacle in the path of the weak, becomes a stepping stone in the path of the strong."

Thomas Carlyle

Agradecimentos

À Deus, pela centelha divina do aprendizado.

Aos meus pais, Luiza e Ricardo, por serem sempre e para sempre meu porto seguro.

Aos meus irmãos, Digo e Lô, pela parceria de todos os dias, mesmo de longe, não é mesmo Digo?

Ao meu marido, meu "princeso", Paulo Sette, por acreditar em mim, principalmente, nas horas que eu duvidei da miha capacidade, por ter olhos tão gentis e amorosos, que conseguem enxergar o melhor de mim todos os dias. Obrigada, meu amor, por todo apoio, paciência e por segurar minha mão e andar ao meu lado em todos os momentos, especialmente, naqueles que eu quis desistir de tudo.

À grande família que eu ganhei casando com o Vinícius. Essa família tão linda que torce por mim e me quer tanto bem.

A todos os meus amigos que participam da minha vida e contribuíram de alguma forma nessa caminhada longa e não muito fácil (para ser sincera, é difícil).

À família Ledbetter, por ser um lugar de paz, de muitas risadas e de muito amor.

Às BPN's, por serem o meu lugar sagrado, onde eu me encontro e onde as dificuldades da vida tornam-se pequenas.

Ao Camarim 1, vocês são o meu melhor presente dessa fase maluca da minha vida.

Aos meus orientadores, Ricardo Gargano e Alessandra Kiametis, por toda leveza e alegria. Quero dizer um muito obrigada especial para a Kia, porque você "ralou" junto comigo nessa tese e me ajudou a remar esse barco sem deixar que eu esmorecesse, me incentivando todos os dias para que eu não deixasse a "peteca cair".

Aos professores da pós-graduação pela minha formação acadêmica.

Aos funcionários da UnB pela ajuda com as burocracias.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), pois o presente trabalho foi realizado com seu apoio - Código de Financiamento 001.

Resumo

O envelhecimento é uma realidade da sociedade contemporânea. Fato comprovado pelo aumento da expectativa de vida, que hoje tem uma média mundial próxima dos 72 anos. Porém, o aumento da expectativa de vida tem apresentado alguns desafios para a ciência, como a demência senil, haja vista que 47 milhões de pessoas no mundo inteiro apresentam algum tipo de demência e a cada ano surgem cerca de 10 milhões de novos casos, sendo a doença de Alzheimer (DA) a causa mais comum dessa enfermidade, abrangendo 60% a 70% das ocorrências. A demência é um conceito bastante amplo que abrange os casos de doença mental caracterizados por perda da capacidade cognitiva, além de incluir diversas alterações como: perda gradativa da memória, distúrbios de linguagem e de habilidades visuoespaciais, dificuldade de raciocínio, de concentração e de aprendizado, entre outras. Embora a DA seja a causa mais comum de demência, ela é apenas um subtipo desta síndrome. Há outras causas possíveis, dentre elas: demência vascular, demência de corpos de Lewy (DCL), demência da doença de Parkinson (DDP), demência frontotemporal e demência da doença de Huntington. Desses exemplos de causas da demência, a DCL, a DDP e a DA compartilham um biomarcador conhecido como placas senis formadas pelo peptídeo β -Amilóide ($A\beta$).

Dentro desse contexto, neste trabalho foi feito um estudo do reconhecimento molecular dos compostos naturais cinamaldeído, curcumina e ácido rosmarínico (ligantes) pelo peptídeo $A\beta$, na sua forma de 42 aminoácidos ($A\beta_{1-42}$). Esse estudo foi viabilizado através do uso de um protocolo obtido da combinação da dinâmica molecular, *docking* e métodos computacionais de modelagem molecular. Os resultados obtidos indicam que os ligantes têm atividades sobre o peptídeo $A\beta_{1-42}$, principalmente, nos aminoácidos Phe4, Leu17, His13, Phe19, Phe 20, Ile31, Ile32 e Val39, que possivelmente têm um papel importante na agregação proteica do $A\beta_{1-42}$. Assim, quando esses aminoácidos interagem com os compostos estudados, tem-se o impedimento da agregação do peptídeo e, por conseguinte, da formação de placas senis. Esse impedimento está vinculado principalmente às interações do tipo ligações de hidrogênio e empilhamento de anéis aromáticos. A partir dessas interações, pode-se

também perceber que os grupos funcionais anéis aromáticos, carbonila e hidroxila são importantes para se ter ação no peptídeo $A\beta_{1-42}$. Esse mapeamento detalhado das interações entre o peptídeo e esses ligantes naturais pode proporcionar um vislumbre do que um fármaco precisa para alcançar o reconhecimento molecular pelo alvo molecular $A\beta_{1-42}$.

Abstract

Aging is a reality in contemporary society. This fact is confirmed by the increase in life expectancy, which today has an average of 72 years. However, the increase in life expectancy has presented some challenges for science, such as senile dementia, given that 47 million people worldwide have some type of dementia and each year around 10 million new cases arise, being Alzheimer's disease (AD) the most common cause of this disease, covering 60% to 70% of occurrences. Dementia is a very broad concept that covers cases of mental illness characterized by loss of cognitive skills, in addition to several changes such as: gradual loss of memory, language and visuospatial disorders, difficulty in reasoning, concentration and learning, among others. Although AD is the most common cause of dementia, it is only a subtype of this syndrome. There are other possible causes, including vascular dementia, Lewy body dementia (LBD), Parkinson's disease dementia (PDD), frontotemporal dementia and Huntington's disease dementia. Of these examples of causes of dementia, LBD, PDD and AD share a hallmark known as senile plaques formed by the β -Amyloid ($A\beta$) peptide.

In this context, this work was carried out in a molecular recognition study of natural compounds cinnamaldehyde, curcumin and rosmarinic acid (ligands) by the peptide $A\beta$, in the form of 42 amino acids ($A\beta_{1-42}$). This study was made possible through the use of a protocol applied by the combination of molecular dynamics, *docking* and computational methods of molecular modeling. The results obtained show that the ligands has activity on the $A\beta_{1-42}$ peptide, mainly in the amino acids Phe4, Leu17, His13, Phe19, Phe 20, Ile31, Ile32 and Val39, which are probably playing an important role on the aggregation of $A\beta_{1-42}$. Thus, when these amino acids interact with the studied compounds, they prevent the aggregation of peptide and, therefore, the formation of senile plaques. This impediment is mainly linked to the hydrogen bonding and aromatic ring stacking interactions. From these interactions, you can also see that the aromatic ring, carbonyl and hydroxyl groups are important for action on the $A\beta_{1-42}$ peptide. This detailed mapping of the interactions between peptides and these natural ligands can provide a glimpse of

what a drug needs to achieve molecular recognition by the molecular target $A\beta_{1-42}$.

Sumário

Sumário v

Lista de Tabelas viii

Lista de Figuras ix

1	Introdução	1
1.1	Síntese e Mecanismos dos Biomarcadores: α -sinucleína, proteína Tau e peptídeo A β	4
1.1.1	α -sinucleína	4
1.1.2	Proteína Tau	6
1.1.3	Peptídeo A β	7
1.2	Compostos Naturais com Ação em Proteínas Amiloidogênicas	13
1.2.1	Cinamaldeído	15
1.2.2	Curcumina	16
1.2.3	Ácido rosmarínico	17
1.3	Objetivos	17
2	Metodologias	19
2.1	Dinâmica Molecular Clássica	20
2.1.1	Funções de Energia Potencial	23

2.1.2	Simulação de Dinâmica Molecular	26
2.1.3	<i>Ensembles</i> Estatísticos	30
2.1.4	Trajетórias geradas na DM	34
2.2	Docking Molecular	37
2.2.1	Funções de <i>Scoring</i>	37
2.3	Métodos Computacionais voltados à Modelagem Molecular	40
2.3.1	Aproximação de Born-Oppenheimer	41
2.3.2	Equação de Schrödinger Eletrônica	44
2.3.3	Métodos Semi-empíricos	52
2.3.4	Teoria do Funcional de Densidade	53
2.3.5	O Método Integrado de N-camadas: ONIOM	62
3	Resultados e Discussões	66
3.1	Resultados da Dinâmica Molecular	66
3.1.1	Configurações Iniciais e Equilibração do Sistema	66
3.1.2	Análise de RMSD das Conformações do Peptídeo A β	68
3.2	Resultados de <i>Docking</i> Receptor-ligante	78
3.3	Resultados das Energias do Complexo Receptor-Ligante via ONIOM	83
3.3.1	Energias Eletrônicas das Interações do Cinamaldeído com o Peptídeo A β_{1-42}	85
3.3.2	Energias Eletrônicas das Interações da Curcumina com o Peptídeo A β_{1-42}	93
3.3.3	Energias Eletrônicas das Interações do Ácido Rosmarínico com o Peptídeo A β_{1-42}	104
4	Conclusões e Perspectivas	113
	Referências Bibliográficas	116
A	Estrutura Química dos Aminoácidos	129

B Enovelamento(<i>folding</i>) Proteico	133
C <i>Scripts</i> para Gerar os <i>Inputs</i> do ONIOM	136
D Artigo publicado	143

Lista de Tabelas

- 1.1 Tabela com os compostos cinamaldeído, curcumina e ácido rosmarínico. 14

Lista de Figuras

1.1	Diagrama das características patológicas presentes em alguns dos subtipos de demência.	3
1.2	Tecido de pacientes com doenças neurodegenerativas e sua provável evolução temporal.	5
1.3	Representação dos processos relacionados à proteína tau. Fonte: Adaptado de Giacobini, 2013 [17]	7
1.4	Representação das vias não-amiloidogênica e amiloidogênica. Fonte: Adaptado de Spies, Petra E.; et al. [27]	9
1.5	Esquema geral do mecanismo de agregação do peptídeo $A\beta$	10
1.6	Monômero do peptídeo $A\beta_{1-42}$ com duas α -hélices e distribuição dos aminoácidos no monômero de acordo com sua polaridade	12
1.7	Monômero do peptídeo $A\beta_{1-42}$ na sua forma de folha- β e um exemplo de trímero.	13
1.8	Isômeros trans (E) e cis(Z) do cinamaldeído, respectivamente.	15
1.9	Tautômeros da curcumina: formas cetônica(A) e enólica(B).	16
2.1	Tipos de interação no campo de força CHARMM.	26
2.2	Diagrama das caixas imagem, obtidas pela aplicação das condições de fronteiras periódicas.	28
2.3	Diagrama esquemático do método ONIOM para duas camadas. Fonte: Elaborado pela autora	64

3.1	Caixas das simulações para a conformação do $A\beta_{1-42}$ em α -hélice e em folha- β	68
3.2	RMSD da extensão total do peptídeo $A\beta_{1-42}$ com conformação inicial de α -hélice	70
3.3	Mudanças sofridas pelo peptídeo $A\beta_{1-42}$, na conformação α -hélice, a cada 20 ns da simulação	71
3.4	Medida de distância entre os aminoácidos Ile31(laranja) e Val34(cinza), a partir do centro massa de cada um deles, durante 200 ns de simulação de DM.	72
3.5	Medida de distância entre os aminoácidos Val24(branco) e Asn27(verde), a partir do centro massa de cada um deles, durante 200 ns de simulação de DM.	73
3.6	RMSD da região dos aminoácidos Lys28-Val39 do peptídeo $A\beta_{1-42}$ com conformação inicial de α -hélice durante 200ns de simulação de DM.	73
3.7	RMSD da região dos aminoácidos Ser8-Gly25 do peptídeo $A\beta_{1-42}$ com conformação inicial de α -hélice durante 200ns de simulação de DM.	74
3.8	RMSD da extensão total do peptídeo $A\beta_{1-42}$ com conformação inicial de folha- β durante 200 ns de simulação de DM	75
3.9	RMSD da região dos aminoácidos Asp1-Asp7 do peptídeo $A\beta_{1-42}$ com conformação inicial de folha- β durante 189 ns de simulação de DM. O segundo gráfico detalha o RMSD a partir do momento que a região analisada flutua na posição de equilíbrio	75
3.10	RMSD das regiões dos aminoácidos Gln15-Ala21 (à esquerda) e dos Val39-Ala42 (à direita) do peptídeo $A\beta_{1-42}$ com conformação inicial de folha- β durante 189 ns de simulação de DM.	76
3.11	Mudanças sofridas pelo peptídeo $A\beta_{1-42}$, na conformação de folha- β , a cada 20 ns da simulação	77

3.12	Medida de distância entre os aminoácidos His4(branco, folha- β magenta) e Val18(branco, folha- β azul), a partir do centro massa de cada um deles, durante 189 ns de simulação de DM.	78
3.13	Nuvem de <i>docking</i> do cinamaldeído na conformação de α -hélice do peptídeo A β_{1-42}	79
3.14	Nuvem de <i>docking</i> do cinamaldeído na conformação de folha- β do peptídeo A β_{1-42}	80
3.15	Nuvem de <i>docking</i> da curcumina na conformação de α -hélice do peptídeo A β_{1-42}	81
3.16	Nuvem de <i>docking</i> da curcumina na conformação de folha- β do peptídeo A β_{1-42}	81
3.17	Nuvem de <i>docking</i> do ácido rosmarínico na conformação de α -hélice do peptídeo A β_{1-42}	82
3.18	Nuvem de <i>docking</i> do ácido rosmarínico na conformação de folha- β do peptídeo A β_{1-42}	83
3.19	Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo A β_{1-42} mais o cinamaldeído (α -hélice)	86
3.20	Interações do aldeído cinâmico com os aminoácidos do A β_{1-42} no domínio hidrofílico (α -hélice)	87
3.21	Interações do aldeído cinâmico com os aminoácidos do A β_{1-42} no domínio hidrofóbico (α -hélice)	88
3.22	Representação dos orbitais HOMO e LUMO para o cinamaldeído isolado e interagindo com o peptídeo A β_{1-42} em sua conformação de α -hélice	89
3.23	Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo A β_{1-42} mais o cinamaldeído (folha- β)	90
3.24	Interações do aldeído cinâmico com os aminoácidos do A β_{1-42} no domínio hidrofílico (folha- β)	91

3.25	Interações do aldeído cinâmico com os aminoácidos do $A\beta_{1-42}$ no domínio hidrofóbico (folha- β)	92
3.26	Representação dos orbitais HOMO e LUMO para o cinamaldeído isolado e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de folha- β	93
3.27	Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo $A\beta_{1-42}$ mais a curcumina (α -hélice)	95
3.28	Interações da curcumina com os aminoácidos do $A\beta_{1-42}$ no domínio hidrofílico (α -hélice)	96
3.29	Interações da curcumina com os aminoácidos do $A\beta_{1-42}$ no domínio hidrofóbico (α -hélice)	97
3.30	Representação dos orbitais HOMO e LUMO para a curcumina isolada e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de α -hélice	98
3.31	Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo $A\beta_{1-42}$ mais a curcumina (folha- β)	99
3.32	Interações da curcumina com os aminoácidos do $A\beta_{1-42}$ no domínio hidrofílico (folha- β)	101
3.33	Interações da curcumina com os aminoácidos do $A\beta_{1-42}$ no domínio hidrofóbico (folha- β)	102
3.34	Representação dos orbitais HOMO e LUMO para a curcumina isolada e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de folha- β	103
3.35	Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo $A\beta_{1-42}$ mais o ácido rosmarínico (α -hélice)	105
3.36	Interações do ácido rosmarínico com os aminoácidos do $A\beta_{1-42}$ no domínio hidrofílico (α -hélice)	106
3.37	Interações do ácido rosmarínico com os aminoácidos do $A\beta_{1-42}$ no domínio hidrofóbico (α -hélice)	107

3.38	Representação dos orbitais HOMO e LUMO para o ácido rosmarínico isolado e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de α -hélice	108
3.39	Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo $A\beta_{1-42}$ mais ácido rosmarínico (folha- β)	109
3.40	Interações do ácido rosmarínico com os aminoácidos do $A\beta_{1-42}$ no domínio hidrofílico (folha- β)	110
3.41	Interações do ácido rosmarínico com os aminoácidos do $A\beta_{1-42}$ no domínio hidrofóbico (folha- β)	111
3.42	Representação dos orbitais HOMO e LUMO para o ácido rosmarínico isolado e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de folha- β	112
A.1	Fórmula geral dos aminoácidos.	129
B.1	Processo de agregação das proteínas	134
B.2	Níveis estruturais de uma proteína	135

Capítulo 1

Introdução

O comportamento dos seres vivos é o resultado da comunicação entre os genes e o meio ambiente. Nesse cenário, a aprendizagem e a memória são os mecanismos mais importantes pelos quais o ambiente altera o comportamento. É através da aprendizagem que se adquire conhecimento sobre o mundo, enquanto a memória é o processo pelo qual o conhecimento é codificado, armazenado e, depois, recuperado. No entanto, devido ao envelhecimento, o mecanismo da memória tem sofrido alterações significativas colocando o idoso sob o risco de uma nova epidemia: a demência, que é uma síndrome caracterizada pela perda de memória e prejuízo cognitivo de gravidade suficiente para interferir nas relações sociais e ocupacionais do ser humano [1].

O envelhecimento é uma realidade da sociedade contemporânea. Fato comprovado pelo aumento da expectativa de vida, que hoje tem uma média mundial de 72 anos [2]. Porém, o aumento da expectativa de vida tem apresentado alguns desafios para a ciência, como a demência senil, haja vista que 50 milhões de pessoas no mundo inteiro apresentam algum tipo de demência e a cada ano surgem cerca de 10 milhões de novos casos, sendo a doença de Alzheimer (DA) a causa mais comum dessa enfermidade, abrangendo 60% a 70% das ocorrências [2, 3].

A demência é um conceito bastante amplo que abrange os casos de doença mental caracterizados por perda da capacidade cognitiva, além de incluir diversas alterações como: perda gradativa da memória, distúrbios de linguagem e de habi-

lidades visuoespaciais, dificuldade de raciocínio, de concentração e de aprendizado, entre outras. Considerando-se a amplitude de consequências dessa doença, é necessário diferenciar a demência do envelhecimento natural, pois, com o chegar da idade mais avançada, ocorrem mudanças próprias dessa condição nas habilidades cognitivas como dificuldade de foco, de acesso a memórias específicas, por exemplo lembrar nomes, de solucionar problemas mais complexos, redução da capacidade psicomotora, etc. Essas mudanças não se adequam aos critérios do diagnóstico de demência, porque o prejuízo cognitivo provocado por essa síndrome é quase sempre acompanhado não só por distúrbios emocionais e neuropsiquiátricos (psicose, agitação, depressão, etc), mas também pelo declínio da capacidade de realizar as atividades diárias [4].

Embora a DA seja a causa mais comum de demência, ela é apenas um subtipo desta síndrome. Há outras causas possíveis, dentre elas: demência vascular, demência de corpos de Lewy (DCL), demência da doença de Parkinson (DDP), demência frontotemporal (DFT) e demência da doença de Huntington [4]. Entretanto, esse rol não se encerra nesses exemplos. Apesar disso, para o presente estudo há dois subtipos de demência que requerem uma maior atenção, além da própria DA: a DCL, que é responsável por 15%-20% dos casos dessa síndrome [3], e a DDP (3-4%) [5].

A DA, a DCL e a DDP podem ser confundidas entre si, pois além de compartilharem um quadro clínico semelhante, também possuem características patológicas em comum, a saber: as placas senis ou neuríticas (também chamadas de fibrilas amiloide), que são agregados proteicos de qualquer proteína com a característica comum de um núcleo de folhas- β , e os emaranhados neurofibrilares da proteína Tau (Figura 1.1). Além delas, um outro biomarcador conhecido como corpos de Lewy também está presente tanto na DCL, na DDP e na DA, sendo na última em menor quantidade [6, 7].

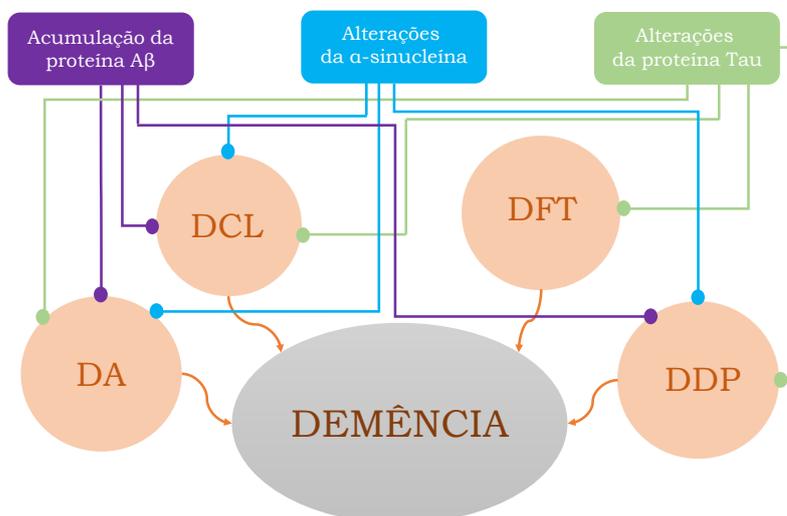


Figura 1.1: Diagrama das características patológicas presentes em alguns dos subtipos de demência.

Nessas três enfermidades, as placas senis são formadas pelo peptídeo β -Amilóide ($A\beta$), mas em pessoas com doença de Parkinson (DP) sem demência, a presença dessas placas é praticamente insignificante. Entretanto, apesar da DCL e da DDP terem como principal característica os corpos de Lewy, nas que possuem DDP também existe a presença do $A\beta$, que é bastante difundida e severa o que, provavelmente, contribui para a demência. Na DCL, esse peptídeo também é bem difundido e ultrapassa a quantidade observada na DDP [8].

Por conseguinte, é necessário ter melhor compreensão dos mecanismos moleculares presentes em cada doença. Não só por ser importante para o desenvolvimento de novas terapias, mas também para o entendimento mais substancial do que é a demência e de como funcionam as suas causas.

1.1 Síntese e Mecanismos dos Biomarcadores: α -sinucleína, proteína Tau e peptídeo $A\beta$

1.1.1 α -sinucleína

A doença de Parkinson e a demência de corpos de Lewy, ambas conhecidas também pelo termo genérico doença de corpos de Lewy, são α -sinucleinopatias, que são neuropatias caracterizadas pelo acúmulo anormal da proteína α -sinucleína hiperfosforilada nos neurônios [9, 10]. Essas duas neuropatias são caracterizadas pela presença de corpos de Lewy (intraneuronal) e neuritos de Lewy, que se alojam preferencialmente nos prolongamentos neuronais¹(Figura 1.2) [8, 13, 14].

¹Corpos de Lewy são inclusões intracitoplasmáticas com 5-25 μm de diâmetro, tem um núcleo eosinofílico (coloração obtida após usar o corante eosina) denso e um halo circundante mais claro. O seu núcleo é composto de um material filamentosos e granular que é cercado por filamentos orientados radialmente. Esses corpos são compostos pelas proteínas α -sinucleína (majoritariamente) e ubiquitina [11, 12].

Neurito é qualquer filamento ou crescimento em formato de ponta apresentado por células neurais em cultura de tecido ou células embrionárias. Estas projeções podem continuar-se para formarem uma árvore ramificada de dendritos ou um único axônio ou podem ser reabsorvidas num estágio adiantado de desenvolvimento. Os neuritos de Lewy são neuritos anormais que contém filamentos semelhantes aos encontrados em corpos de Lewy [13].

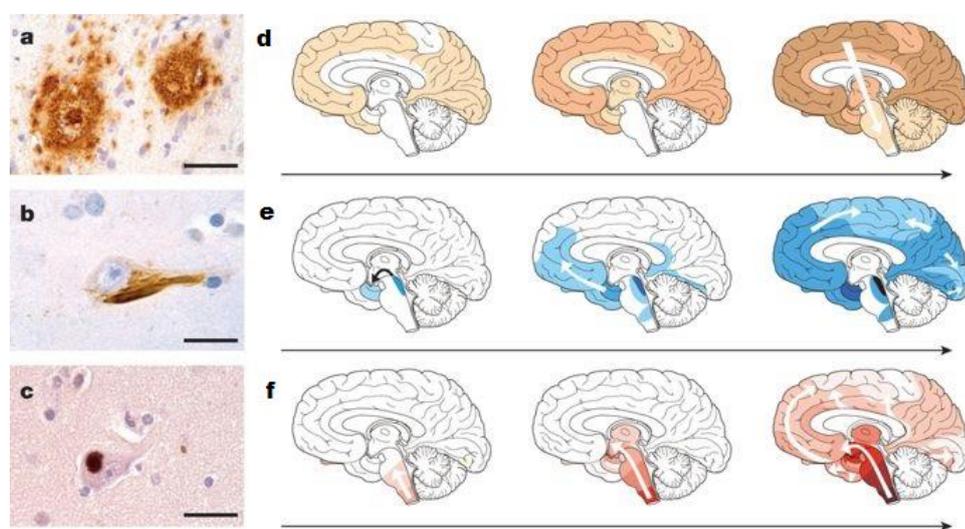


Figura 1.2: Tecido de pacientes com doenças neurodegenerativas e sua provável evolução temporal. (a) depósitos de proteína $A\beta$ (placas senis) no neocórtex de um paciente com DA. (b) inclusão emaranhado neurofibrilar da proteína Tau em um neurônio neocortical de um paciente com DA. (c) inclusão de α -sinucleína (corpo de Lewy) em um neurônio neocortical de um paciente com DP/DCL. Barras de escala - $50 \mu\text{m}$ em a, $20 \mu\text{m}$ em (b) e (c). (d-f), progressão de lesões proteicas específicas em doenças neurodegenerativas ao longo do tempo (t), inferidas a partir de análises *post-mortem* de cérebros; depósitos de $A\beta$ e inclusões de Tau em cérebros de pacientes com DA (d e e), inclusões de α -sinucleína no cérebro de pacientes com DP (f). Três estágios são mostrados para cada doença, com setas brancas indicando a possível propagação da lesões.

Fonte: Adaptado de Jucker & Walker, 2013 [15]

Na maioria dos pacientes, essas duas principais sinucleinopatias são fáceis de diferenciar clinicamente pela predominância de características motoras extrapiramidais² em pacientes com DP e demência em pacientes com DCL. Contudo, em alguns pacientes, demência e sinais extrapiramidais ocorrem em sucessão, levantando a questão sobre a nosologia³ de cada uma das doenças. Para fins de pesquisa, os casos são classificados como DCL se a demência ocorrer dentro de 12 meses antes das características parkinsonianas, ou parkinsonismo, e como demência da DP (DDP) se a demência ocorre um ano ou mais após um diagnóstico clínico de Parkin-

²O sistema extrapiramidal é formado pelo tálamo, cerebelo e gânglios da base. Estas estruturas, através de suas conexões, estão envolvidas em vários processos, inclusive na modulação do controle motor. A disfunção de estruturas do sistema extrapiramidal associa-se a transtornos dos movimentos [16].

³Ramo da medicina que estuda a classificação das doenças.

son [5]. Na prática clínica real, um diagnóstico é feito com base nas características clínicas preponderantes. A separação entre DDP e DCL é considerada por alguns como artificial, uma vez que tal separação implica que as duas síndromes clínicas têm diferentes substratos anatômicos, que podem, em última instância, ter diferentes desafios terapêuticos, e há pouca diferença na distribuição ou gravidade das características patológicas entre as duas [14].

1.1.2 Proteína Tau

A Tau é uma proteína altamente solúvel e associada à estabilidade neuronal. Ela está concentrada, principalmente, nos axônios e estabiliza a estrutura dos neurônios ao ligar-se nos microtúbulos. Seis isoformas de Tau são expressas no cérebro humano adulto por meio do *splicing* alternativo⁴ de mRNA do gene MAPT (do inglês *Microtubule-Associated Protein Tau*), todas elas foram detectadas em lesões associadas à DA. Nas doenças em que a proteína Tau está presente, ela está hiperfosforilada e incapaz de interagir com os microtúbulos, ocasionando a desestruturação do citoesqueleto e interrupção do fluxo no axônio [17].

A DA, em particular, caracteriza-se pela presença de altos níveis de inserções intraneuronais de Tau hiperfosforilada no cérebro. Esse defeito da proteína Tau parece promover sua auto-montagem em filamentos helicoidais emparelhados (do inglês, *Paired Helical Filaments* - PHFs) e filamentos retos, dessa forma esses eventos ocorrem antes da agregação da tau em emaranhados neurofibrilares (do inglês, *Neurofibrillary Tangles* - NFTs) (Figura 1.3). Após a morte de neurônios que possuem agregados solúveis da Tau, os filamentos dessa proteína podem estar presentes no espaço extracelular.

⁴É um processo regulado durante a expressão gênica que resulta na codificação de múltiplas proteínas por um mesmo gene.

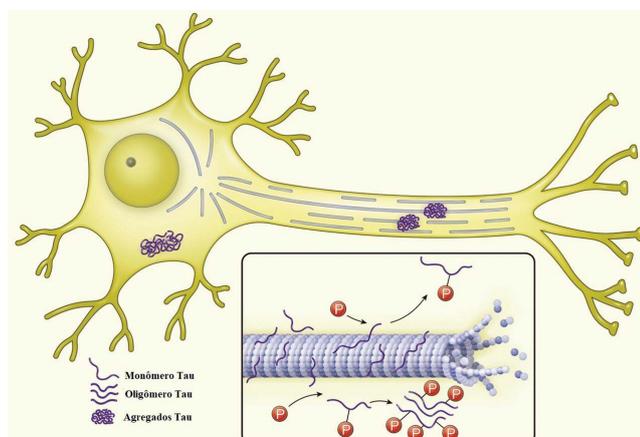


Figura 1.3: Representação dos processos relacionados à proteína tau.
 Fonte: Adaptado de Giacobini, 2013 [17]

1.1.3 Peptídeo $A\beta$

Dentro das vertentes de pesquisa voltadas para a DA, há uma hipótese chamada cascata amilóide, que no seu contexto originário dizia que a deposição de $A\beta$ no cérebro constitui a lesão patológica central da DA. Essa hipótese atribui a demência à morte neuronal causada pela toxicidade de fibrilas amilóides insolúveis. Recentemente, o foco dessa proposta tem passado das placas amilóides para o conceito de que a falha da memória é causada pela interrupção da função sináptica por pequenos oligômeros solúveis do peptídeo $A\beta$ [17].

Sendo assim, o foco dessa estratégia é a restauração da homeostase do $A\beta$ no cérebro, que é interrompida por alterações na produção e na remoção de $A\beta$ nessa enfermidade. Com base nessas premissas, foram desenvolvidas quatro estratégias principais para reduzir os níveis de $A\beta$ no cérebro: prevenção ou redução da formação de $A\beta$, visando, em particular, as enzimas proteolíticas que mediam o processamento da APP (β -secretase e γ -secretase); remoção de depósitos amilóides existentes por meio da imunoterapia (imunização passiva ou ativa); prevenção ou redução da agregação de $A\beta$, e aprimoramento da remoção $A\beta$ [17].

Ainda no âmbito da DA, em particular na DA familiar, um dos genes envolvidos nessa doença é o Precursor da Proteína Amilóide (PPA) que reside no

cromossomo 21, também responsável pela Síndrome de Down. Porém, só foi observada essa mutação no tipo precoce da DA. Outros genes mutantes identificados foram os da pressenilina 1 e pressenilina 2, proteínas transmembranares que modificam o processamento do PPA, resultando em quantidades aumentadas de uma forma particularmente tóxica do peptídeo $A\beta$, a $A\beta_{1-42}$ [18]. Na forma mais tardia da DA, os marcadores genéticos observados foram os da proteína apolipoproteína E (APOE), cujo gene que a codifica está no cromossomo 19, sendo o alelo $\epsilon 4$ que possui fator de risco para desenvolvimento da DA. Apesar disso, herdar esse alelo não é suficiente para causar a DA, apenas aumenta o risco de desenvolvê-la [19, 20].

Paralelamente, há evidências de que a atividade da isoforma da APOE4 aumenta a concentração do peptídeo $A\beta$, cujos agregados formam as placas neuríticas, ao diminuir a sua remoção do organismo. Além dessa proteína, há a enzima GSK-3 β que é essencial para a neurotoxicidade induzida pela $A\beta$. A $A\beta$ não só ativa a GSK-3 β , mas também ativa as enzimas quinase 5 dependente de ciclina (CDK5, do inglês *Cyclin-dependent Kinase 5*) e quinase 2 reguladora extracelular (ERK2, do inglês *Signal-regulated Extracellular Kinase 2*), as quais conduzem à hiperfosforilação da Tau e, por fim, à apoptose⁵. Além de todos esses acontecimentos, o peptídeo $A\beta$ é capaz de afetar a membrana neuronal ao ponto de desregular a concentração interna de íons, em especial, a do íon Ca^{2+} [21, 22].

No que tange a deposição da proteína $A\beta$, primeiramente, precisa ocorrer a sua formação que decorre da clivagem sequencial da APP (do inglês, *Amyloide Precursor Protein*), uma glicoproteína integral da membrana, através da via metabólica amiloidogênica. Há duas vias que competem pelo substrato APP, a que leva ao processamento amiloidogênico e a que conduz ao processo não amiloidogênico da proteína (Figura 1.4). Elas são mediadas por três enzimas proteolíticas: α , β e γ -secretases. Na via não-amiloidogênica, a clivagem da APP é feita pela α -secretase na superfície celular, liberando uma porção solúvel do N-terminal, sAPP α , e o fragmento C-terminal CTF α (do inglês, *C-Terminal Fragment*). Em sequência acontece

⁵Morte celular

a clivagem do $\text{CTF}\alpha$ por meio da γ -secretase, o que produz um peptídeo p3 extracelular solúvel e a APP de domínio intracelular (AICD - *Amyloid Precursor Protein Intracellular Domain*). Já a via amiloidogênica envolve a clivagem da APP pela β -secretase, conhecida como BACE1, no aminoterminal que libera a parte solúvel, $\text{sAPP}\beta$, no domínio extracelular e o fragmento $\text{CTF}\beta$. A clivagem do $\text{CTF}\beta$ através da γ -secretase, no carboxiterminal, produz peptídeos β -amilóide de comprimentos variados (39-43 aminoácidos), bem como o fragmento AICD [23, 24, 25].

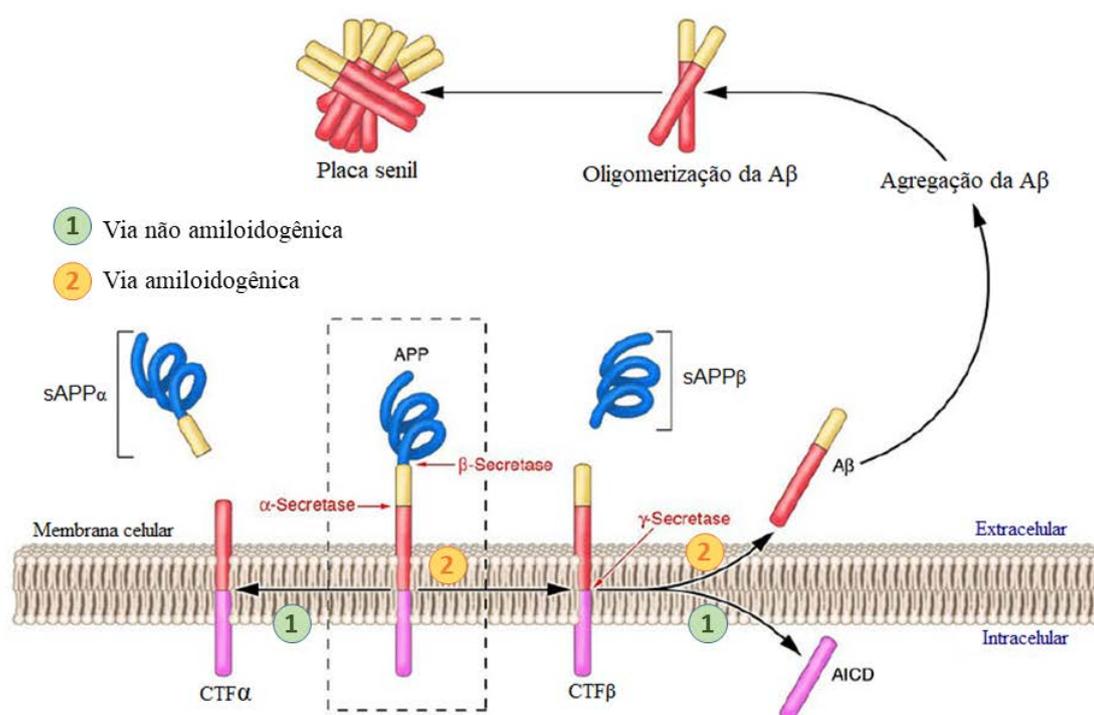


Figura 1.4: Representação das vias não-amiloidogênica e amiloidogênica.
Fonte: Adaptado de Spies, Petra E.; et al. [27]

O $\text{A}\beta$ produzido pode seguir essencialmente dois caminhos: continuar na forma solúvel no meio extracelular ou agregar-se (oligomerizar), ficando insolúvel e depositando-se como placas neuríticas. Para seguir a segunda via, quando o peptídeo é liberado, ele altera sua forma nativa principal de α -hélice para folha- β (detalhes no Apêndice B) permitindo sua autoassociação em oligômeros (dímeros, trímeros,

etc), protofibrilas e fibrilas (Figura 1.5) ou apenas em oligômeros [26].

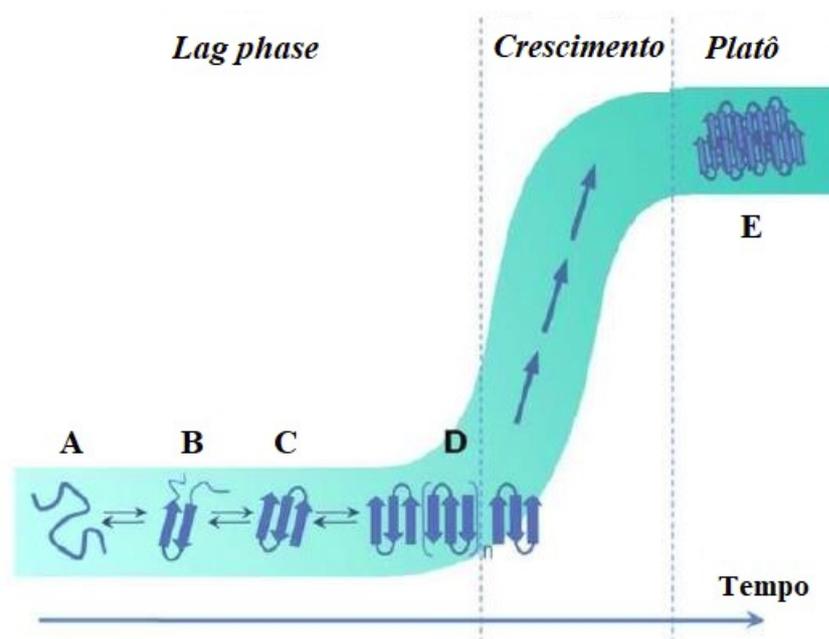


Figura 1.5: Esquema geral do mecanismo de agregação. (A) Fase lenta (*lag phase*) em que o peptídeo ou a proteína está na sua conformação nativa e se desdobra formando uma estrutura de folha- β (B) que agrega para formar o núcleo (C) e as protofibrilas (D) - até esse momento os agregados são solúveis. Em seguida, na fase de crescimento (fase mais curta) os monômeros e oligômeros associam-se ao núcleo de folha- β para formar as fibrilas maduras na fase platô - placas senis (E).

Fonte: Adaptado de Belo, 2014 [28]

Alguns indícios sugerem que a formação de $A\beta$ pode conduzir à patologia da Tau, mesmo que a deposição de Tau no cérebro anteceda a formação de placas senis. No entanto, se $A\beta$ é necessária para a neurotoxicidade da Tau ou se o contrário é verdadeiro continua a ser uma questão inconclusiva. Apesar disso, há crescentes evidências de que para ocorrer os eventos relacionados à DA, exige-se a presença dessas duas proteínas, ou seja, um único biomarcador não é suficiente para desencadear todos os processos da doença [29]. Somado a esse contexto, o compartilhamento da presença do $A\beta$ pelas causas apresentadas que levam à demência é uma das razões que motivou este trabalho a ter o foco voltado para esse peptídeo enquanto alvo molecular.

Peptídeo $A\beta_{1-42}$ como Alvo Molecular

O peptídeo $A\beta$ possui duas formas predominantes: o $A\beta$ de 40 aminoácidos ($A\beta_{1-40}$) e o de 42 aminoácidos ($A\beta_{1-42}$). Levando em consideração todos os possíveis comprimentos de $A\beta$, o peptídeo $A\beta_{1-40}$ é o de maior ocorrência no organismo como um todo e, também, um dos componentes principais das placas neuríticas juntamente com o $A\beta_{1-42}$ [30]. Apesar disso, mesmo o $A\beta_{1-40}$ participando da composição dessas placas, o componente predominante nestas estruturas é a forma $A\beta_{1-42}$, responsável por 90% da composição destes agregados [31].

Essas duas formas preponderantes de $A\beta$ possuem uma estrutura primária semelhante, dada pela seguinte sequência de aminoácidos: NH_2 - ¹DAEFRHDSGY ¹¹EVHHQKLVFF ²¹AEDVGSNKG ³¹IIGLMVGGVV ⁴¹IA - COOH (detalhes dos códigos e estruturas dos aminoácidos estão no Apêndice A). A diferença entre elas reside no dipeptídeo Ile41-Ala42, presente na forma $A\beta_{1-42}$ e ausente na $A\beta_{1-40}$. Esse dipeptídeo a mais no carboxiterminal tem um grande impacto no comportamento desses dois peptídeos de modo a alterar suas taxas e mecanismos de agregação. A forma $A\beta_{1-42}$ possui uma taxa de agregação muito maior que a da $A\beta_{1-40}$ [31]. Além disso, ela é a forma mais neurotóxica do peptídeo $A\beta$ [23, 32, 33]. Isso sugere que o impacto patogênico desse peptídeo varia com o seu tamanho.

Dentre as estruturas de $A\beta$, a sua forma monomérica (Figura 1.6 e Figura 1.7) é o primeiro estágio antes do início do processo de oligomerização. Esse processo ocorre por meio do rearranjo conformacional do peptídeo (Apêndice B), passando de α -hélice para folha- β e/ou de um *coil* para folha- β , levando à construção de agregados de vários tamanhos: dímeros, trímeros, tetramêros e outros agregados muito maiores até chegar na forma de placas senis, o produto final da agregação. Então, caracterizar essa estrutura inicial a nível atômico é essencial para elucidar seus mecanismos de agregação [34].

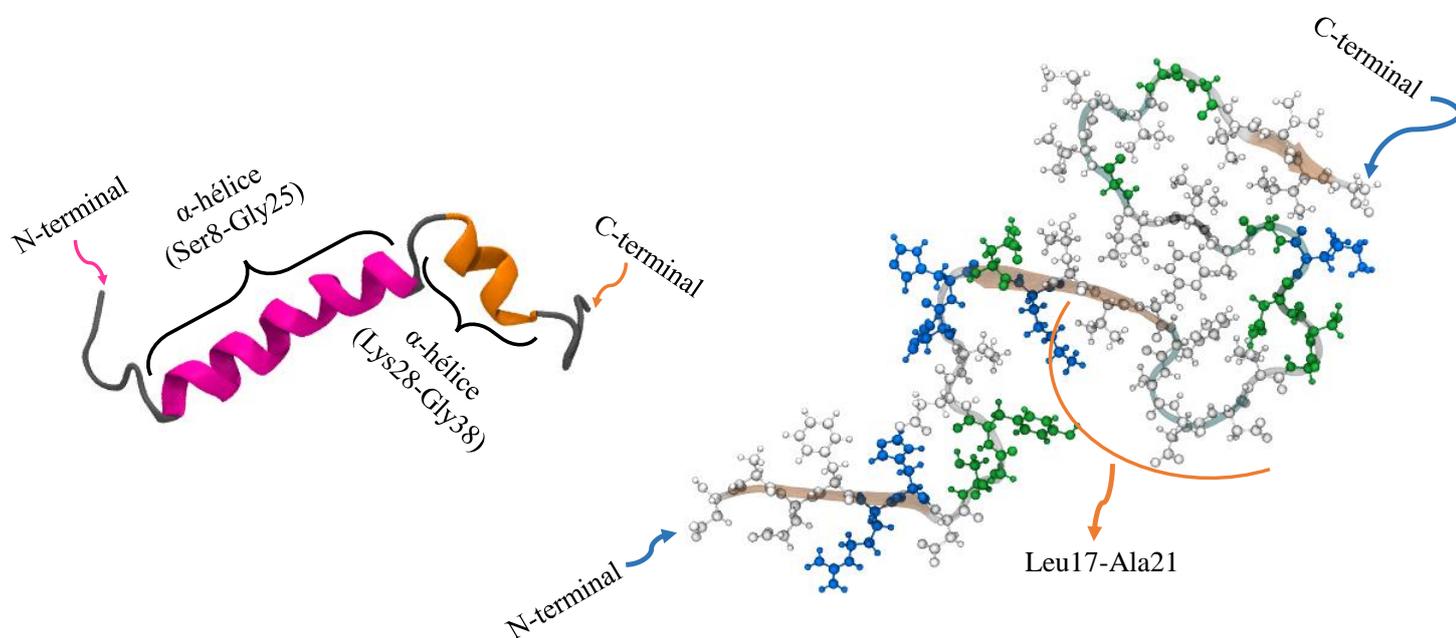


Figura 1.6: À esquerda, o monômero do peptídeo $A\beta_{1-42}$ com duas α -hélices nas regiões indicadas (código PDB:1Z0Q, modelo 25). À direita, a distribuição dos aminoácidos no monômero de acordo com sua polaridade: branco para apolares, verde para polares e azul para polares básicos. Em condições fisiológicas e pH neutro, os aminoácidos aspartato(D), glutamato(E), lisina(K) e arginina(R) possuem carga, enquanto a histidina(H) é considerada neutra [34]. A maioria dos aminoácidos hidrofóbicos estão na região Leu17-Ala21 e perto do C-terminal (Ala30-Ala42), com exceção das glicinas (Gly33,Gly37,Gly38) que são polares. Além disso, os primeiros 15 aminoácidos do N-terminal são predominantemente hidrofílicos [34].

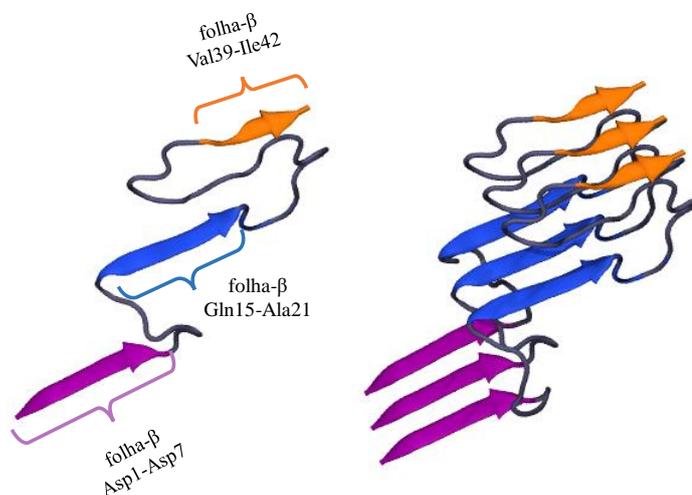


Figura 1.7: Na figura da esquerda, tem-se o monômero do peptídeo $A\beta_{1-42}$ na sua forma de folha- β (código PDB:2NAO, cadeia B); à direita, um exemplo de trímero (código PDB:2NAO, cadeias A,B e C), que é um agregado formado a partir de monômeros como o representado à esquerda.

Na atualidade, os meios pelos quais ocorre a oligomerização permanecem inconclusivos. Entretanto, há alguns achados de grande importância que contribuem para encontrar a estrutura desse processo de agregação. Entre essas descobertas está o fato de que as regiões compreendidas pelos aminoácidos Leu17-Ala21, Asp23-Lys28 e Gly29-Met35 têm papel essencial na agregação do peptídeo $A\beta$ [31, 35]. A importância de cada uma dessas regiões será devidamente abordada na seção de resultados e discussões deste trabalho.

1.2 Compostos Naturais com Ação em Proteínas Amiloidogênicas

Através dos séculos, plantas e temperos são usados tanto no âmbito nutricional como na esfera medicinal. Diante disso, uma dieta baseada nesses alimentos com ações terapêuticas vem ganhando espaço, pois cada vez mais estudos estão

comprovando que essas plantas vão além da adição de sabor à comida. Elas têm demonstrado grande utilidade para retardar, prevenir e até tratar doenças neurodegenerativas [36, 37].

Com base nesse cenário, o presente trabalho aborda três compostos naturais com ação no peptídeo $A\beta_{42}$. São eles: cinamaldeído, curcumina e ácido rosmarínico (Tabela 1.1). O primeiro possui uma estrutura química bem menor que a dos outros dois compostos, os quais compartilham semelhanças estruturais, sendo, portanto, um contraponto para observar o porquê do cinamaldeído também ter ação no peptídeo $A\beta$, mesmo com algumas literaturas demonstrando que compostos com similaridades estruturais às da curcumina são melhores candidatos para ter ação nessa estrutura proteica [38].

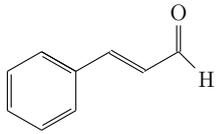
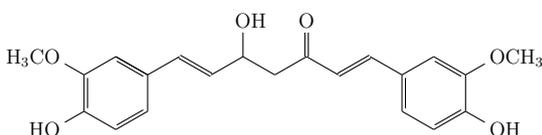
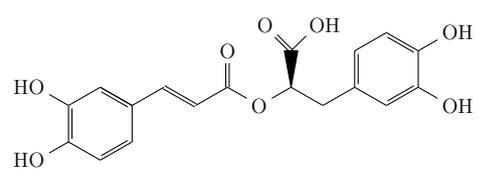
Composto	Fórmula molecular	Fórmula estrutural	IC ₅₀ (μ M)
Cinamaldeído	C ₉ H ₈ O		ND
Curcumina	C ₂₁ H ₂₀ O ₆		10 ^a
Ácido rosmarínico	C ₁₈ H ₁₆ O ₈		12 ^a

Tabela 1.1: Tabela com os compostos estudados, suas respectivas fórmulas moleculares, estruturais e o IC₅₀^b para inibição do peptídeo $A\beta_{42}$. ND é Não Determinado.
^a Reinke, Ashley A.; Gestwicki, Jason E.[38]

^b IC₅₀ (*Half maximal inhibitory concentration*) é o índice de concentração máxima necessária para um composto inibir 50% de um processo biológico. Assim quanto menor esse índice mais eficaz é um fármaco e quanto maior a porcentagem de inibição mais potente é o composto, ou seja, precisa-se de uma dose pequena para fazer efeito (inibir)

1.2.1 Cinamaldeído

A canela é um tempero comum usado por vários séculos em diferentes culturas pelo mundo. Ele é obtido da casca interna das árvores do gênero *Cinnamomum* da família Lauraceae, que possui aproximadamente 250 espécies. Essa planta tropical tem duas variedades principais: *Cinnamomum zeylanicum*, também conhecida como *Cinnamomum verum* e *Cinnamomum cassia* (ou *Cinnamomum aromaticum*/canela-chinesa) [39, 40]. O cinamaldeído ou aldeído cinâmico é o composto orgânico responsável pelo sabor e odor da canela, sendo o componente majoritário do óleo essencial dessa planta [41]. Esse aldeído possui isômeros cis(Z) e trans(E) (Figura 1.8), sendo o (E)-cinamaldeído a ocorrência predominante [42].

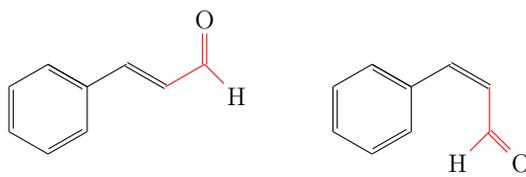


Figura 1.8: Isômeros trans (E) e cis(Z) do cinamaldeído, respectivamente.

Essa especiaria amplamente usada possui propriedades curativas que derivam dos seus vários compostos, quais sejam, cinamaldeído, eugenol, acetato de cinamila e uma ampla variedade de outras substâncias. Dentre os atributos medicinais da canela estão: controle glicêmico, anti-oxidante, anti-inflamatório e anti-microbiano [40, 41, 43]. Nesse contexto, o cinamaldeído é considerado o principal responsável por essas características [43, 44]. Além disso, foi demonstrado que essa especiaria tem um efeito inibitório na agregação de Tau relacionada a DA [41], bem como propriedades farmacológicas no tratamento de diabetes tipo II [37].

Ademais, estudos demonstram que o cinamaldeído possui atividade neuroprotetora contra o peptídeo $A\beta$ [40, 43], reduz sua oligomerização, diminui a agregação da Tau e provoca uma melhora cognitiva em modelo animal de DA [37, 45].

1.2.2 Curcumina

O açafrão-da-terra, também conhecido como cúrcuma, turmérico, entre outros nomes é uma planta herbácea da família do gengibre, Zingiberaceae, e espécie *Curcuma longa* originária da Ásia. Ele é diferente da especiaria extraída das flores de *Crocus sativus*, também conhecida como açafrão. É a partir da raiz da cúrcuma que se obtém uma especiaria homônima de cor dourada, usada como condimento ou corante. O composto responsável por essa cor é a curcumina, principal constituinte químico dessa planta. Além disso, é reconhecido como sendo responsável pela maioria dos efeitos terapêuticos como propriedades anti-inflamatórias, anticancerígenas, antibacteriana, antiviral e antifúngica [46, 47].

A curcumina é hidrofóbica por natureza e freqüentemente solúvel em dimetilsulfóxido, acetona, etanol e óleos. Quando exposta a condições ácidas, sua cor passa do amarelo para o vermelho escuro, e, nessa cor, é usada frequentemente em várias cerimônias religiosas [46]. Ela pode existir segundo duas formas tautoméricas distintas, a forma cetônica e a forma enólica (Figura 1.9), das quais a última é a mais estável, tanto em fase sólida quanto em solução [47].

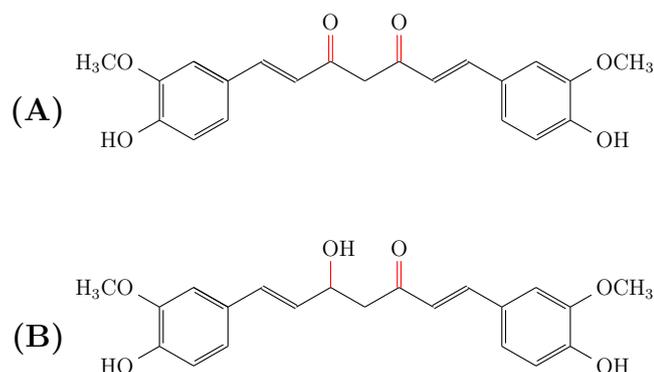


Figura 1.9: Tautômeros da curcumina: formas cetônica(A) e enólica(B).

Ademais, estudos recentes também mostraram que a curcumina pode ser útil no tratamento da DA devido à sua capacidade de impedir a agregação das proteínas amiloidogênicas. O mecanismo de sua propriedade protetiva em relação à DA é atribuído a vários fatores, incluindo sua capacidade de eliminar as espécies

reativas de oxigênio (ROs, do inglês *Reactive Oxygen species*) que são conhecidas por desempenhar um papel no dano cerebral. Principalmente, sabe-se que a curcumina impede a agregação das proteínas A β e Tau interagindo diretamente com elas para impedir sua agregação em espécies neurotóxicas no sistema nervoso central [30, 48].

1.2.3 Ácido rosmarínico

O ácido rosmarínico (estrutura na Tabela 1.1) é um éster proveniente do ácido cafeico e do ácido 3,4-dihidroxifenilacético, sendo um composto com caráter predominantemente lipofílico. Ele é encontrado em várias espécies vegetais como alecrim (*Rosmarinus officianalis L.*), manjericão (*Ocimum basilicum*), orégano (*Origanum vulgare*), salvia (*Salvia officianalis*) e muitas outras. Dentre as suas diversas atividades biológicas estão as antiséptica, antioxidante, anti-inflamatória, antiviral, hipoglicemiante, antitumoral e neuroprotetora [49]. Além de todas essas propriedades, o ácido rosmarínico previne a neurotoxicidade induzida pelo A β , protege contra as ROs e inibe a formação das placas amilóides vistas na DA agindo na agregação dos monômeros em oligômeros e destes em fibrilas [33, 50].

1.3 Objetivos

Haja vista a atividade inibitória desses três compostos no peptídeo A β_{1-42} , a presente pesquisa tem como objetivo compreender os detalhes moleculares dessas interações, por meio da busca de padrões que possam explicar a afinidade dessas três moléculas pelo mesmo alvo. Esse alvo molecular foi escolhido devido à sua participação no papel de componente majoritário das placas senis da DA e também ter um papel importante em outras causas de demência, por ser a estrutura mais neurotóxica desse peptídeo e, evidentemente, pela atividade que os compostos naturais citados neste capítulo possuem sobre ele.

O peptídeo A β_{1-42} será estudado nas suas diferentes conformações de monô-

mero, α -hélice e folha- β , pois essa estrutura é o estágio anterior ao início da agregação. Assim sendo, esse passo primário será investigado para observar se há possibilidades do impedimento da mudança de conformação desse peptídeo, de α -hélice para folha- β , pelos compostos naturais apresentados e, mesmo ocorrendo a mudança conformacional, se esses compostos conseguem interromper a continuidade do processo de agregação. Ao mapear essas relações, é possível direcionar a racionalização de novos fármacos com atividade nesse alvo molecular.

Capítulo 2

Metodologias

No presente capítulo, serão apresentadas as fundamentações teóricas pertinentes a cada um dos métodos usados para a análise do peptídeo $A\beta_{1-42}$ e das interações desse alvo molecular com os compostos naturais escolhidos, a saber: aldeído cinâmico (isômero *trans*), curcumina (forma enólica) e ácido rosmarínico. Dentre as metodologias aplicadas estão: dinâmica molecular clássica dos monômeros nas formas α -hélice e folha- β , *docking* molecular dos compostos naturais no peptídeo e análise energética dos possíveis sítios de interação através de método híbrido de Mecânica Quântica e Mecânica Molecular, que são métodos mais adequados para o tratamento de sistemas com uma grande quantidade de átomos como os estudados neste trabalho, pois proporcionam um menor custo computacional e um nível de cálculo mais apurado na região de interesse.

A dinâmica molecular foi empregada com o intuito de avaliar a flexibilidade do peptídeo, a estabilidade de suas estruturas secundárias, α -hélices e folhas- β , e para amostrar um grande número das possíveis conformações desse alvo molecular. Assim que o peptídeo atingiu o equilíbrio termodinâmico, um *ensemble* de várias conformações dessa fase equilibrada foi produzido. A partir deste *ensemble*, foram extraídas várias conformações do peptídeo para fazer o *docking* molecular sobre essa amostragem de equilíbrio, a fim de localizar os possíveis sítios de interação dos compostos naturais no alvo molecular. Como resultado do *docking*, foram obtidos

vários complexos dos ligantes mais o peptídeo em suas diversas conformações dadas pela dinâmica. Partindo desse cenário, a próxima etapa foi estimar as energias dos complexos através de cálculos híbridos de Mecânica Quântica e Mecânica Molecular objetivando apontar as configurações mais estáveis. Por fim, caracterizou-se as possíveis interações dos complexos com menor energia por meio de medidas de distância e análise dos orbitais de fronteira.

2.1 Dinâmica Molecular Clássica

A dinâmica molecular (DM) é um método amplamente usado para o estudo de uma diversidade de problemas desde fluidos simples a biopolímeros complexos e macromoléculas. Ela fornece informações sobre o comportamento dinâmico microscópico, dependente do tempo, dos átomos individuais que compõem o sistema. Já as propriedades macroscópicas como pressão, volume, temperatura, energia interna, entropia, energia livre são deduzidas por meio de expressões matemáticas da mecânica estatística que conectam as observáveis macroscópicas com o movimento e a distribuição dos átomos e moléculas [51, 52].

Nessa ferramenta, o comportamento dinâmico de um sistema molecular é simulado usando a integração numérica das equações de movimento regidas pelas leis de Newton. O resultado desses cálculos é a trajetória que especifica como as posições e velocidades das partículas no sistema variam com o tempo.

A trajetória é obtida ao ser resolvida a equação da segunda lei de Newton que para um sistema tridimensional para um átomo i com massa m_i e posição \vec{r}_i é dada por:

$$\vec{F}_i = \frac{d\vec{p}_i}{dt}, \quad (2.1)$$

a partir dessa expressão tem-se a relação entre velocidade e momento, dada pela equação:

$$\frac{d\vec{r}_i}{dt} = \frac{\vec{p}_i}{m_i}. \quad (2.2)$$

A força exercida sobre a partícula i , é dada pelo negativo do gradiente da função energia potencial em relação à posição de i , isto é:

$$\vec{F}_i = -\frac{dU}{d\vec{r}_i} = -\vec{\nabla}U(\vec{r}_i). \quad (2.3)$$

Nessa esfera, é possível diferenciar três tipos de problemas envolvendo as leis de Newton para o movimento: o primeiro caso e mais simples é a ausência de força atuante sobre cada partícula entre as colisões - de uma colisão para a outra, a posição da partícula muda pela equação $\vec{v}_i\delta t$, sendo \vec{v}_i a velocidade constante e δt o tempo entre as colisões; o segundo trata-se da atuação de uma força constante entre as colisões. Um exemplo disso é uma partícula carregada movendo-se em um campo elétrico uniforme; na terceira situação, a força que atua na partícula depende da sua posição em relação às outras partículas [53]. Nesse último caso, é bastante difícil descrever o movimento analiticamente e uma das alternativas para solucionar esses problemas são os métodos de diferenças finitas [54].

Dentro desses métodos, há vários algoritmos para integrar as equações de movimento. Em todos eles as posições e as propriedades dinâmicas podem ser aproximadas pelas expansões de séries de Taylor:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t\vec{v}(t) + \frac{1}{2}\delta t^2\vec{a}(t) + \frac{1}{6}\delta t^3\vec{b}(t) + \frac{1}{24}\delta t^4\vec{c}(t) + \dots \quad (2.4)$$

$$\vec{v}(t + \delta t) = \vec{v}(t) + \delta t\vec{a}(t) + \frac{1}{2}\delta t^2\vec{b}(t) + \frac{1}{6}\delta t^3\vec{c}(t) + \dots \quad (2.5)$$

$$\vec{a}(t + \delta t) = \vec{a}(t) + \delta t\vec{b}(t) + \frac{1}{2}\delta t^2\vec{c}(t) + \dots \quad (2.6)$$

sendo \vec{v} a velocidade, que é a primeira derivada da posição com respeito ao tempo, \vec{a} é a aceleração, que corresponde a segunda derivada, \vec{b} é a terceira derivada e assim por diante.

O algoritmo mais difundido, provavelmente, é o de Verlet [55]. Esse algoritmo faz o uso das posições, das acelerações no tempo t e das posições do passo anterior, $x(t-\delta t)$, para determinar as novas posições em $(t+\delta t)$. Então, as expansões de Taylor para as posições ficam descritas como:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) + \dots \quad (2.7)$$

e

$$\vec{r}(t - \delta t) = \vec{r}(t) - \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) + \dots \quad (2.8)$$

Adicionando as duas equações acima, tem-se:

$$\vec{r}(t + \delta t) = 2\vec{r}(t) - \vec{r}(t - \delta t) + \delta t^2 \vec{a}(t). \quad (2.9)$$

A velocidade não aparece explicitamente nas equações do algoritmo de Verlet, mas pode ser deduzida ao dividir a diferença das posições nos tempos $t + \delta t$ e $t - \delta t$ por $2\delta t$:

$$\vec{v} = \frac{\vec{r}(t + \delta t) - \vec{r}(t - \delta t)}{2\delta t}. \quad (2.10)$$

Uma das desvantagens desse cálculo é que as posições $\vec{r}(t + \delta t)$ são obtidas ao adicionar-se um termo pequeno, $\delta t^2 \vec{a}(t)$ à subtração de dois termos maiores, $2\vec{r}(t)$ e $\vec{r}(t - \delta t)$, o que pode ocasionar uma perda de precisão.

Há inúmeras variações desse algoritmo. Uma delas é o método da velocidade de Verlet [56] que dá as posições, velocidades e acelerações ao mesmo tempo sem comprometer a precisão. Nesse método, as equações 2.9 e 2.10 ficam escritas como:

$$\vec{r}(t + \delta t) = 2\vec{r}(t) + \vec{v}(t)\delta t + \frac{1}{2}\delta t^2\vec{a}(t) \quad (2.11)$$

e

$$\vec{v}(t + \delta t) = \vec{v}(t) + \frac{1}{2}\delta t[\vec{a}(t) + \vec{a}(t + \delta t)]. \quad (2.12)$$

Como a maior parte das metodologias de dinâmica são dependentes de uma função de energia adequada para reproduzir a paisagem de energia do sistema, faz-se necessário o uso de funções de energia convenientes que descrevam as interações inter e intramoleculares resultando na descrição apropriada da dinâmica do sistema [52].

2.1.1 Funções de Energia Potencial

Considerando o tamanho dos sistema modelado e o custo computacional, é bastante comum o uso da Mecânica Molecular (MM) para a realização da dinâmica molecular. Na MM, as moléculas ao serem descritas por forças newtonianas, podem ser tratadas como um grupo de partículas mantidas unidas por forças harmônicas ou elásticas. O conjunto completo dos potenciais de interação entre as partículas é referido como campo de força. O campo de força empírico, que é um caso especial de funções de energia ou potenciais interatômicos, permite que a energia potencial total do sistema, $V(\mathbf{r})$, seja calculada por meio da estrutura tridimensional (3D) do sistema. $V(\mathbf{r})$ é o somatório de termos da energia, incluindo os termos para átomos ligados como comprimentos e ângulos de ligação, ângulos diedrais e os termos para átomos não-ligados, como as interações de van der Waals e de Coulomb [51]. O campo de força é tipicamente representado pela equação:

$$V(\mathbf{r}) = \sum V_l + \sum V_\theta + \sum V_\phi + \sum V_{vdW} + \sum V_{elec}, \quad (2.13)$$

na qual V_l é a energia de estiramento da ligação em relação à sua posição de equilí-

brio, V_θ é a energia de deformação do ângulo de ligação, no que diz respeito ao seu valor de equilíbrio, V_ϕ é a energia de torção em torno de uma ligação, V_{vdW} representa a energia das interações de van der Waals e V_{elec} são as energias de atração ou repulsão eletrostática entre duas cargas.

Os campos de forças existentes foram desenvolvidos de maneira independente e com parâmetros específicos. Alguns incluem termos para descrever especificamente as ligações de hidrogênio ou as ligações para acoplar oscilações entre ângulos e comprimentos de ligação e, dessa forma, obter uma melhor concordância com os espectros vibracionais. A escolha do campo de força depende do sistema a ser estudado e das propriedades que serão investigadas. No caso de sistemas biomoleculares, os campos de força mais utilizados são CHARMM [57], AMBER [58], GROMOS [59], OPLS [60], entre outros.

No campo de força CHARMM, a função de energia potencial é baseada em cargas de ponto fixo e se divide em duas expressões, uma responsável pela energia intramolecular e outra pela energia intermolecular. Para os termos intramoleculares, tem-se a equação:

$$\begin{aligned}
 U(\vec{r}) = & \sum_{\text{ligações}} K_b(b - b_0)^2 + \sum_{\text{ângulos}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{diedros}} K_\phi(1 + \cos(n\phi - \delta)) \\
 & + \sum_{\substack{\text{diedros} \\ \text{impróprios}}} K_\psi(\psi - \psi_0)^2 + \sum_{\substack{\text{Urey-} \\ \text{Bradley}}} K_{UB}(r_{1,3} - r_{1,3;0})^2 + \\
 & + \sum_{\text{resíduos}} U_{CMAP}(\phi, \psi),
 \end{aligned} \tag{2.14}$$

e para os termos intermoleculares, termos não-ligados, tem-se:

$$U(\vec{r}) = \sum_{\text{n\~{a}o-ligados}} \frac{q_i q_j}{4\pi D r_{ij}} + \epsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right]. \quad (2.15)$$

Nas contribuiç~oes intramoleculares est~ao os termos de ligaç~oes(b), ângulos de valência(θ), torç~ao ou ângulos diedros(ϕ), diedros impróprios(ψ), o termo de Urey-Bradley e cor- reç~oes torcionais para o *backbone*, U_{CMAP} . Nesses termos, os coeficientes K s~ao as constantes de força e os parâmetros com sub-índice 0 s~ao os valores correspondentes ao equilíbrio.

O termo de Urey-Bradley é uma funç~ao quadrática de distância e representa a correlaç~ao entre o comprimento de ligaç~ao e o potencial angular. Já o termo de ângulo diedral impróprio é uma funç~ao quadrática usada para pontos como o C_α da proteína, em que átomos ligam-se a um átomo central e o diedro impróprio é um ângulo diedral fictício. Os termos de Urey-Bradley e de diedros impróprios s~ao usados para melhorar os ajustes na parte vibracional do espectro com relaç~ao aos movimentos fora do plano (Figura 2.1).

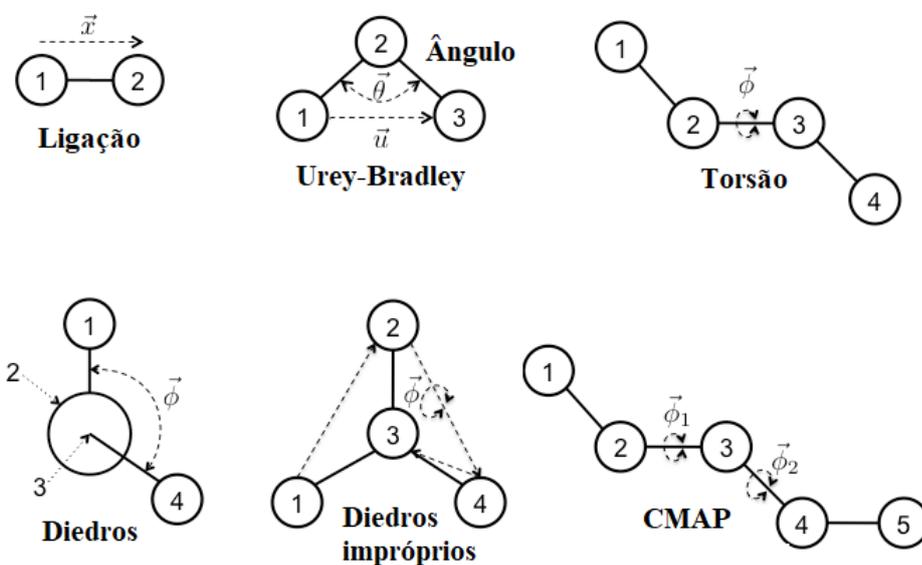


Figura 2.1: Tipos de interação no campo de força CHARMM.

Nas contribuições intermoleculares, estão inclusos os termos eletrostáticos e as interações de van der Waals, sendo q_i e q_j as cargas parciais dos átomos i e j , respectivamente, ϵ_{ij} a profundidade do poço, $R_{min,ij}$ é o raio no termo de Lennard-Jones para tratar as interações de van der Waals, e r_{ij} é a distância entre i e j .

2.1.2 Simulação de Dinâmica Molecular

Dinâmica molecular (DM) é um método de simulação computacional que estuda o movimento físico dos átomos e moléculas de um sistema, cujo potencial de interação, posições e velocidades iniciais entre suas partículas já são conhecidos. Essas simulações de DM podem ter seu processo dividido nas seguintes etapas: escolha das condições iniciais do sistema, o cálculo das forças exercidas sobre cada partícula devido às interações intermoleculares, equilíbrio, *ensembles* e controles de simulação e, por fim, a análise das trajetórias geradas.

Na configuração inicial do sistema, sem a proteína, é conveniente situar, inicialmente, as partículas nas posições de uma rede cristalina, que pode ser criada a

partir de um arranjo experimental disponível, para evitar uma sobreposição indesejada entre as partículas na caixa de simulação. A caixa cúbica é a mais usual e suas dimensões são determinadas de tal maneira que reproduzem a densidade numérica, $\frac{N}{V}$, do sistema real e de suas condições termodinâmicas. Para caracterizar-se o sistema, as condições fisiológicas são reproduzidas de modo a aproximá-lo do modelo real e, na maioria das simulações, o meio escolhido é o aquoso em que, na DM, a solvatação é explícita. No caso de solvente explícito, existem alguns modelos para a água, entre eles estão os da família TIP (*Transferable Intermolecular Potentials*) que são o TIP3P e TIP4P [60], SCP e SCP/E [52]. No entanto, os modelos mais conhecidos se utilizam da aproximação SHAKE [61] para manter-se a geometria molecular. Além disso, são adicionadas condições de contorno com o intuito de reduzir os efeitos de fronteira, pois as moléculas que estão na borda da caixa experimentam forças atrativas e/ou repulsivas diferentes das moléculas do centro.

As condições de contorno ou condições de fronteiras periódicas replicam a caixa de simulação em várias caixas adjacentes, nos três eixos cartesianos, conservando a configuração das moléculas da caixa original. Assim, durante a simulação da caixa central, as moléculas da superfície ficam cercadas por moléculas em todas as direções, evitando os efeitos indesejados nas bordas que gerariam descontinuidades na distribuição espacial das interações intermoleculares (Figura 2.2). Devido a aplicação das condições periódicas, as partículas da caixa central interagem com as partículas das caixas imagem, logo se um sistema contém N partículas e o potencial se expressa como a soma das interações entre pares de partículas, então as interações com as partículas das caixas imagem também tem de ser incluídas nessa soma. Com isso, a soma constará de infinitos termos e, na prática, isso requer a realização de aproximações.

Para o tratamento das interações de curto alcance, cujo potencial decai rapidamente com o aumento da distância entre as partículas, emprega-se o procedimento denominado convenção da imagem mínima, que é o truncamento do potencial em um raio de corte esférico, geralmente da ordem de $\frac{L}{2}$, em que L é o comprimento da

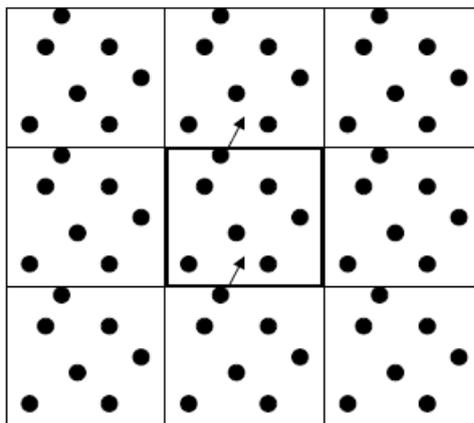


Figura 2.2: Diagrama das caixas imagem, obtidas pela aplicação das condições de fronteiras periódicas.

célula de simulação, para evitar a interação da partícula com sua própria imagem ou com a mesma partícula duas vezes. Contudo, no tratamento das interações de longo alcance, em que o potencial decai com a distância que é menor do que r^{-d} , sendo d a dimensionalidade do sistema, o truncamento abrupto do potencial produz efeitos que, em muitos casos, comprometem severamente os resultados das simulações. Esse problema tem sido contornado com a utilização de métodos baseados em somas infinitas como o PME (*Particle Mesh Ewald*) [62], uma das variantes do método de soma de Ewald adotada na maioria dos pacotes computacionais, que produzem trajetórias estáveis [51, 63].

A soma de Ewald é um método para calcular interações de longo alcance em sistemas periódicos. Ele foi desenvolvido pela primeira vez como um método para o cálculo de energias eletrostáticas de cristais iônicos e hoje é comumente usado para calcular interações de longo alcance em química computacional. A soma de Ewald é um caso especial da fórmula da soma de Poisson (Equação 2.16), que é uma equação que relaciona os coeficientes da série de Fourier da soma periódica de uma função com os valores da transformada de Fourier da função contínua. Na soma de Ewald, substitui-se da fórmula de Poisson o somatório das energias de interação no espaço real por uma soma equivalente no espaço de Fourier.

$$f_p(x) = \sum_{n=-\infty}^{\infty} f(x+n) = \sum_{n=-\infty}^{\infty} \hat{f}(2\pi n) e^{2\pi i n x}. \quad (2.16)$$

Na equação acima, f_p é uma função periódica de período 1, $f_p(x+1) = f_p(x)$. Se f for absolutamente integrável em \mathbb{R} , então segue do teorema de Fubini que f_p é absolutamente integrável em $[0, 1]$. \hat{f} é a transformada de Fourier da função f e os coeficientes $\hat{f}(2\pi n)$ vão para zero rapidamente.

A soma de Ewald é dada pela seguinte equação:

$$\phi(\vec{r}) = \phi_{sr}(\vec{r}) + \phi_{lr}(\vec{r}), \quad (2.17)$$

sendo o primeiro termo o de curto alcance e o segundo termo o de longo alcance. A interação de longo alcance é dividida em duas partes: uma contribuição de curto alcance e uma contribuição de longo alcance que não tem singularidade (a função é "bem comportada"). A contribuição de curto alcance é calculada no espaço real, enquanto a contribuição de longo alcance é calculada usando uma transformada de Fourier. A vantagem desse método é a rápida convergência da energia em comparação com a de uma soma direta, ou seja, o método possui alta precisão e velocidade razoável ao computar interações de longo alcance e, por isso, é um método bastante utilizado para o cálculo de interações de longo alcance em sistemas periódicos.

O *Particle Mesh* (PM) é um método computacional para determinar as forças em um sistema de partículas. Essas partículas podem ser átomos, estrelas ou componentes fluidos e, dessa forma, o método é aplicável a vários campos, incluindo dinâmica molecular e astrofísica. O princípio básico desse método é converter um sistema de partículas em uma grade de valores de densidade, a partir disso o potencial é, então, resolvido para essa grade de densidade e as forças são aplicadas a cada partícula com base na célula em que está e no local em que a célula está.

No PME, assim como na soma padrão de Ewald (Equação 2.17), o potencial de interação é dividido em dois termos. A idéia básica da soma de Ewald para

uma malha de partículas é substituir a soma direta de energias de interação entre partículas pontuais:

$$E_{Total} = \sum_{i,j} \phi(\vec{r}_j - \vec{r}_i) = E_{sr} + E_{lr}, \quad (2.18)$$

em que E_{Total} é a energia total, E_{sr} é a energia de curto alcance e E_{lr} é a energia de longo alcance. Sendo,

$$E_{sr} = \sum_{i,j} \phi_{sr}(\vec{r}_j - \vec{r}_i), \quad (2.19)$$

e

$$E_{lr} = \sum_k \hat{\Phi}_{lr}(\vec{k}) |\hat{\rho}(\vec{k})|, \quad (2.20)$$

na equação acima $\hat{\Phi}_{lr}(\vec{k})$ e $|\hat{\rho}(\vec{k})|$ são as transformadas de Fourier do potencial e da carga, respectivamente [62].

2.1.3 *Ensembles* Estatísticos

A Física Estatística, por meio da teoria de probabilidades, relaciona as leis e os resultados da Termodinâmica com o comportamento de um grande número de partículas presentes nos sistemas macroscópicos. No cenário das simulações, essa área é de grande importância, pois apesar do movimento das partículas microscópicas ser governado pelas leis da mecânica, a solução das equações do movimento para o grande número de partículas que compõem os sistemas, da ordem da constante de Avogadro ($A \approx 6,02 \times 10^{23}$), torna-se inviável. Esse ramo da ciência possibilita a construção de uma análise mecânico-estatística, na qual conecta-se os estados microscópicos (Mecânica) a um estado macroscópico (Termodinâmica). Com o objetivo de obter um formalismo mais adequado para o cálculo de grandezas termodinâmicas, surgiram os *ensembles* estatísticos, que são o conjunto dos estados microscópicos

aos quais se associam determinados pesos probabilísticos (pesos de Boltzmann). Em outras palavras, esses *ensembles* permitem considerar um grande número de cópias fictícias com a mesma natureza (macroestado) do sistema observado, cada uma representando um estado físico possível em que esse sistema pode estar [64]. A partir dos *ensembles* estatísticos é possível definir os *ensembles* termodinâmicos que, entre outras características, estão em equilíbrio estatístico, isso quer dizer que os valores médios de todas as quantidades físicas que caracterizam o estado são independentes do tempo, e através deles as propriedades dos sistemas termodinâmicos são alcançadas. Há vários *ensembles* termodinâmicos com características distintas, em que parâmetros diferentes permanecem constantes, como: o canônico (NVT), que o estado termodinâmico é caracterizado por um número fixo de átomos(N), volume(V) e temperatura(T) constantes; o isobárico-isoentálpico (NPH), no qual o número de átomos, a pressão e a entalpia são fixos, o gran-canônico (μVT), em que o potencial químico, o volume e a temperatura são constantes, o isobárico-isotérmico(NPT), cujo número de átomos, a pressão e a temperatura são constantes, etc [65].

O *ensemble* canônico se refere ao conjunto de estados possíveis que um sistema mecânico pode ser encontrado estando em equilíbrio térmico, mantido pelo contato com um banho de calor a uma temperatura fixa. O sistema pode trocar energia com o banho de calor para que os estados do sistema sejam diferentes em energia total [66]. Nesse *ensemble*, um sistema clássico com N partículas e temperatura constante T interagindo em equilíbrio é retratado pela distribuição de Boltzmann. Esse físico, em sua teoria que gerou a distribuição de Boltzmann, queria explicar o comportamento de um gás por meio da Mecânica Clássica, algo inviável pela quantidade de moléculas existentes num gás. Assim, para Boltzmann formular matematicamente sua teoria precisou fazer uma suposição conhecida como hipótese ergódica.

Para elucidar essa hipótese, considere a média de uma quantidade $G(\vec{r}^N)$, que é uma função da posição \vec{r} de todas as N partículas e é expressa na forma de uma integral de espaço de fase incluindo a energia potencial $U(\vec{r}^N)$:

$$\langle G \rangle = \frac{\int G(\vec{r}^N) \exp[-\beta U(\vec{r}^N)] d\vec{r}^N}{\int \exp[-\beta U(\vec{r}^N)] d\vec{r}^N} \quad (2.21)$$

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

Se o sistema em equilíbrio, depois de um tempo grande o suficiente e sob certas condições, ocupa todo o espaço de fase de microestados com a mesma energia, e a densidade de probabilidade é diferente de zero, o que quer dizer que todos os microestados acessíveis são igualmente prováveis nesse período de tempo, o sistema é dito ergódico. Essa é a chamada hipótese ergódica na qual a média de uma amostra ao longo do tempo é igual à média no *ensemble* definida por:

$$\langle G \rangle_t = \lim_{M \rightarrow \infty} \langle G \rangle_M = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{\mu=1}^M G[\vec{r}^N(t_\mu)] \quad (2.22)$$

nessa equação $\langle G \rangle_M$ é a média sobre as medições M, da quantidade G em todos os tempos, $(t_\mu)_{\mu=1}^M$. Para uma boa amostragem, que seria uma grande quantidade de medições, a média $\langle G \rangle_M$ é vista como uma aproximação conveniente para a média do *ensemble* equilibrado [67].

Na maioria das vezes, as simulações de DM fazem uso do *ensemble* microcanônico (NVE) para determinar suas condições iniciais, sendo a energia total E uma constante de movimento e sua média temporal igual a do *ensemble*. Todavia, os experimentos laboratoriais são realizados com volume e temperatura constantes, o que corresponde ao *ensemble* canônico, ou à pressão e temperatura constantes, que é o *ensemble* isobárico-isotérmico.

Mesmo no *ensemble* NVE, inicialmente à energia constante, a temperatura flutua devido a interconversão espontânea das componentes cinética e potencial da energia total. A energia interna total pode ser escrita como a soma das contribuições cinéticas (K) e potencial (U) envolvendo a soma sobre todos os pares de moléculas:

$$E = \langle K \rangle + \langle U \rangle. \quad (2.23)$$

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

$$\left\langle \frac{\sum_{i=1}^N |p_i|^2}{m_i} \right\rangle = 2 \langle k \rangle = 3Nk_B T, \quad (2.24)$$

na equação acima k_B é a constante de Boltzmann, m_i a massa e p_i o momento do átomo i , e N o número total de partículas. Essa equação está relacionada ao princípio da equipartição, assim a função de temperatura cinética instantânea pode ser escrita como:

$$T = \frac{2k}{3Nk_B} = \frac{1}{3Nk_B} \sum_{i=1}^N |p_i|^2. \quad (2.25)$$

Se as velocidades forem alteradas por um fator λ em uma dada temperatura $T(t)$ num instante t , as variações da temperatura podem ser calculadas por:

$$\begin{aligned} \Delta T &= \frac{1}{2} \sum_{i=1} 2 \frac{m_i (\lambda v_i)^2}{N_{df} k_B} - \frac{1}{2} \sum_{i=1} 2 \frac{m_i (v_i)^2}{N_{df} k_B} \\ \Delta T &= (\lambda^2 - 1) T(t) \\ \lambda &= \sqrt{\frac{T_0}{T(t)}} \end{aligned} \quad (2.26)$$

Este é o modo mais simples de controlar a temperatura, multiplicando as velocidades atômicas instantâneas pelo fator λ . Uma variante dessa abordagem é o termostato de Berendsen [68, 69], no qual o sistema é acoplado a um banho térmico com temperatura T_0 fixa, e durante as simulações as velocidades são reescaladas a cada passo de integração para ajustar a energia cinética do sistema até a temperatura desejada.

No *ensemble* NVT a energia total não é conservada e o início da simulação

é equiparado a um banho térmico, portanto esse *ensemble* é essencial para a termalização do sistema, quando tratamentos térmicos como o *Simulated Annealing*, que faz uma busca probabilística do mínimo global da energia termodinâmica, são aplicados [53, 70].

Já no *ensemble* NPT a variável que flutua é o volume do sistema podendo alterar as dimensões da caixa de simulação, além de poder flutuar na medida em que as posições forem reescaladas. O fator $\gamma = \sqrt[3]{\frac{p(t)}{p_0}}$ é o responsável por manter a taxa de flutuação equilibrada. A pressão do sistema pode ser calculada por meio do teorema de virial através da equação:

$$p(t) = \frac{Nk_B T}{V} - \frac{1}{3V} \sum_{i=1}^N \sum_{j>1}^N r_{ij}(t) \cdot F_{ij}(t), \quad (2.27)$$

$r_{ij} = r_j(t) - r_i(t)$ é o vetor posição das moléculas i e j , e F_{ij} é a força intermolecular entre essas duas moléculas no instante t . Existem alguns métodos que fazem a pressão ficar constante corretamente durante a simulação, um exemplo é o método de dinâmica de Langevin, que serve para descrever a dinâmica de uma partícula browniana [71].

2.1.4 Trajetórias geradas na DM

As simulações de DM consistem em cinco fases, quando já estão definidas as estruturas tridimensionais e as topologias dos sistemas de interesse, a saber: etapa inicial, minimização (ou relaxação), termalização, equilibração e produção [72]. Na etapa inicial, é criada a caixa de solvatação contendo receptor e/ou ligante, solvente e íons, caso sejam necessários para neutralizar as cargas do sistema. Na etapa de minimização, o sistema é relaxado para corrigir possíveis distorções nas ligações, ângulos e interações, reduzindo consideravelmente as forças atuantes sobre os átomos em valores próximo de zero e, por conseguinte, a energia potencial do sistema [53]. Dessa forma, a superfície do potencial é percorrida na direção do decréscimo da

energia levando o sistema a um mínimo de energia local próximo ou ao mínimo global. Os algoritmos de minimização mais empregados na DM são o método *Steepest Descent*, que partindo da derivada primeira consegue determinar a direção do mínimo tendo como ponto inicial uma estrutura de alta energia potencial, e o método dos gradientes conjugados, no qual retira-se informações da derivada primeira para levar o sistema o mais próximo possível do mínimo global, determinando a direção e o passo para acelerar a busca pelo mínimo da superfície[51, 53]. Após a relaxação, a força atuante sobre os átomos do sistema minimizado é extremamente pequena, de maneira que os átomos deixam de ter uma aceleração significativa e, conseqüentemente, não adquirem velocidade (equação 2.2), logo não ocupam uma nova posição no sistema. Nesse contexto, ao começar a etapa de termalização, as velocidades iniciais são atribuídas aos átomos aleatoriamente em conformidade com a distribuição de Maxwell-Boltzmann. Dado que esse processo de fixar velocidades para cada temperatura pode ocasionar perturbações no sistema, incluindo deformações nas macromoléculas, o procedimento mais usual é aquecer o sistema gradualmente iniciando em uma temperatura mais baixa até a temperatura de equilíbrio, com o fim de uniformizar as energias do sistema [51]. Por último, na etapa de produção, são obtidas as trajetórias que serão utilizadas para análise.

Na produção, o sistema evolui temporalmente, o que gera as trajetórias. Nessa etapa, as posições e outras configurações são gravadas para a realização das análises das propriedades do sistema.

Das inúmeras propriedades passíveis de serem calculadas a partir das trajetórias, estão propriedades estruturais, como a função de distribuição de pares, $g(r)$, que fornece a probabilidade de se encontrar uma partícula β em uma camada esférica de espessura dr , a uma distância r da partícula α , posicionada na origem. Levando em consideração as coordenadas cartesianas retiradas das trajetórias de simulação, a função radial de pares pode ser dada por:

$$g(r) = \rho^{-2} \left\langle \sum_i \sum_{i \neq j} \delta(r_i) \delta(r_j - r) \right\rangle = \frac{V}{N^2} \left\langle \sum_i \sum_{i \neq j} \delta(r - r_{ij}) \right\rangle. \quad (2.28)$$

Por conseguinte, para o número de partículas $N(r)$, localizado na região Δr a uma distância r da origem, a distribuição radial fica na forma:

$$g(r) = \frac{N(r)}{4\pi r^2 N}. \quad (2.29)$$

Além da distribuição radial, as trajetórias possibilitam o cálculo de médias como o desvio quadrático médio ("*Root Mean Square Deviation*- RMSD), que fornece a diferença estrutural do sistema num determinado passo da simulação em relação às configurações iniciais. Na DM, esse desvio é usado para comparar as coordenadas atômicas das estruturas moleculares e, através disso, observar as flutuações em torno da posição de equilíbrio e as regiões mais instáveis da estrutura. A equação que rege as flutuações ao redor da posição de equilíbrio é:

$$\langle RMSD \rangle = \sqrt{\frac{\sum_{j=1}^N [r_j(t_i) - r_0]^2}{N}}, \quad (2.30)$$

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

Na situação que se deseja observar, os átomos ou as regiões da estrutura que flutuam mais ou menos, tem-se a equação:

$$\langle RMSD_j \rangle = \sqrt{\frac{\sum_{j=1}^M [r_j(t_i - r_0)]^2}{M \cdot \Delta t}}, \quad (2.31)$$

sendo M o número de configurações e Δt o intervalo de tempo entre os passos da simulação, o resultado da multiplicação dessas duas variáveis é o tempo total da simulação.

2.2 Docking Molecular

A técnica de *docking* molecular é uma abordagem *in silico* amplamente utilizada para auxiliar na descoberta e na concepção de novos compostos bioativos.

No *docking* ligante-receptor, busca-se a conformação e a orientação de um ligante com relação a um determinado receptor quando a estrutura deste é conhecida ou pode ser modelada. Em geral, existem dois objetivos a serem alcançados no estudo de *docking*: uma modelagem estrutural e a previsão dos modos de ligação. No entanto, a identificação de características moleculares que são responsáveis por reconhecimento biológico específico, ou a previsão de modificações estruturais que melhorem a potência do composto, são questões complexas que muitas vezes são difíceis de entender e, ainda mais, de simular em um computador [73]. Em vista desses desafios, o processo do *docking* é dividido em etapas no qual cada passo introduz um ou mais graus adicionais de complexidade, mas pode ser resumido em dois momentos: no primeiro é necessário um algoritmo que explore a superfície de energia e preveja a conformação preferencial do ligante no sítio de ação biológica do receptor, que é a pose; e no segundo, coordenar as conformações por intermédio de uma função de avaliação (*scoring function*), o que significa, aplicar um modelo de ponderação da energia livre de ligação, *scorings*, que separe os distintos modos de ligação do mesmo ligante, ou que, sendo os ligantes diferentes, selecione aquele com maior afinidade pelo sítio ativo [74].

2.2.1 Funções de *Scoring*

Ao longo da execução do *docking*, os algoritmos de busca investigam um número excessivo de conformações de ligantes, usualmente milhares. As funções de *scoring* são usadas para avaliar a qualidade desses modos, direcionando a busca para conformações de ligante relevantes e, dessa forma, reduzindo custo computacional.

Geralmente, o método de docagem computacional é realizado com o emprego de um campo de força simples e explorando-se uma região ampla do espaço

conformacional. A depender do programa os campos de força podem variar. A função de *scoring* do programa AutoDock [75] é baseada no campo de força molecular AMBER.

No AutoDock, a energia livre de ligação é a diferença entre a energia do ligante e receptor no estado não-ligado e a energia do ligante e do receptor no complexo. Isso acontece em duas fases: análise das energias intramoleculares da transição do estado não-ligado para o ligado, para cada uma das moléculas separadamente, e a avaliação energética intermolecular das moléculas no complexo.

O campo de força abrange seis termos de potenciais de interação que envolve contribuições entre átomos ligados e não-ligados e um termo de contribuição entrópica conformacional de ligação (ΔS_{conf}) descritos na equação abaixo:

$$\Delta G = (V_{ligado}^{L-L} - V_{n\tilde{a}o-ligado}^{L-L}) + (V_{ligado}^{R-R} - V_{n\tilde{a}o-ligado}^{R-R}) + (V_{ligado}^{R-L} - V_{n\tilde{a}o-ligado}^{R-L} + \Delta S_{conf}), \quad (2.32)$$

L está relacionado ao ligante e R ao receptor. Os primeiros dois termos são referentes às energias intramoleculares para os ligantes nos estados ligado e não-ligado, e os dois subsequentes são correspondentes as energias intramoleculares dos receptores nos mesmos estados. O último parêntese constitui as mudanças de energia intermolecular entre os estados ligado e não-ligado.

Os termos atômicos por pares incorporam a dispersão/repulsão, as ligações de hidrogênio, a parte eletrostática e de dessolvatação de acordo com a equação a seguir:

$$V = W_{vdW} \sum_{i,j} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + W_{H-ligação} \sum \frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \quad (2.33)$$

$$+ W_{elec} \sum \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}} + W_{sol} \sum_{i,j} S_i V_j + S_j V_i e^{\left(\frac{-r_{ij}^2}{2\sigma^2}\right)}.$$

As constantes de ponderação (W) são otimizadas para ajustar a energia livre empírica com base em um conjunto de complexos caracterizados experimentalmente. O primeiro termo é o potencial de Lennard-Jones para interações de dispersão/repulsão. As variáveis A e B foram extraídas do campo de força AMBER. O segundo termo é um potencial para ligações de hidrogênio. Os parâmetros C e D foram atribuídos para melhorar a descrição do potencial de ligação para O-H, N-H e S-H. As interações eletrostáticas são avaliadas segundo um potencial similar ao coulombiano e o último termo representa o potencial de dessolvatação baseado no volume (V) dos demais átomos em torno de um átomo definido, ponderado por um parâmetro de solvatação (S) e um termo exponencial devido a distância. O termo para entropia torsional é proporcional ao número de ângulos de rotação da molécula (N_{tors}):

$$\Delta S_{conf} = W_{conf} N_{tors}. \quad (2.34)$$

O número de ângulos de rotações abarca todos os graus de liberdade torsionais, incluindo rotações de átomos de hidrogênio polar em grupos hidroxila e similares [76].

O AutoDock Vina é uma ferramenta de *docking* molecular que possui vários algoritmos de busca implementados com base no Algoritmo Genético Lamarckiano (do inglês, *Lamarckian Genetic Algorithms* - LGA) [77], que procura um mínimo global de energia de ligação do complexo. Durante o processo de docagem, a proteína é mantida rígida e as mudanças conformacionais são feitas pelo ligante. A inclusão da flexibilidade é de grande importância para os resultados serem compatíveis ao sistema real, sendo fundamental para a compreensão das interações farmacobiológicas. Apesar disso, a flexibilidade do sistema ainda é um desafio para os métodos de docagem.

2.3 Métodos Computacionais voltados à Modelagem Molecular

Nas últimas décadas, ocorreram avanços significativos no desempenho computacional. Esses avanços impactaram positivamente o ramo da Química computacional levando a resultados mais precisos de diversas propriedades moleculares, em especial às relacionadas a grandezas termoquímicas. Mesmo com o desenvolvimento de métodos cada vez mais sofisticados e funções de base extensas para descrever sistemas eletrônicos, a realização de cálculos envolvendo um grande número de partículas ainda é uma tarefa com elevado custo computacional. Algumas estratégias têm tido sucesso e tornado possível a abordagem mais precisa de sistemas moleculares relativamente grandes. Dentre elas destacam-se a Teoria do Funcional de Densidade (DFT, do inglês *Density Functional Theory*), o método ONIOM (do inglês *Own N-layered Integrated Molecular Orbital and Molecular Mechanics*) e também os métodos compostos [78].

O desafio central para os cálculos estruturais em modelagem molecular é a solução da equação de Schrödinger não relativística e independente do tempo:

$$\hat{H}\psi(\mathbf{r},\mathbf{R}) = E\psi(\mathbf{r},\mathbf{R}) \quad (2.35)$$

sendo $\psi(\mathbf{r},\mathbf{R})$ a autofunção do sistema, \mathbf{r} as coordenadas eletrônicas, \mathbf{R} as coordenadas nucleares e \hat{H} o operador Hamiltoniano molecular, escrito em unidades atômicas, é dado por:

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{A=1}^M \sum_{i=1}^N \frac{Z_A}{r_{Ai}} + \sum_{A=1}^{M-1} \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} + \sum_{i=1}^{N-1} \sum_{j>i}^N \frac{1}{r_{ij}} \quad (2.36)$$

em que o primeiro termo é operador energia cinética dos elétrons, o segundo o operador energia cinética dos núcleos, o terceiro a interação coulombiana entre os

núcleos e os elétrons, o quarto a interação coulombiana entre os núcleos e, por fim, o quinto a interação coulombiana entre os elétrons. M_A e Z_A é a massa e o número atômico do núcleo A, respectivamente.

Substituindo a equação (2.36) na equação (2.35), temos a seguinte expressão:

$$\left[\frac{-1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{A=1}^M \sum_{i=1}^N \frac{Z_A}{r_{Ai}} + \sum_{A=1}^{M-1} \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} + \sum_{i=1}^{N-1} \sum_{j>i}^N \frac{1}{r_{ij}} \right] \psi(\mathbf{r}, \mathbf{R}) = E\psi(\mathbf{r}, \mathbf{R}) \quad (2.37)$$

Essa equação (2.37) é extremamente complexa e solucioná-la, mesmo empregando técnicas computacionais, se torna uma tarefa quase impossível para a maioria dos sistemas. Dessa forma, precisamos fazer o uso de aproximações para obtermos a solução desejada. No caso desse trabalho, a aproximação usada é a aproximação de Born-Oppenheimer (ABO).

2.3.1 Aproximação de Born-Oppenheimer

Pelo fato do núcleo ser mais massivo que os elétrons, ele se move muito devagar. A ABO trata desse movimento, isto é, qualitativamente essa aproximação considera o núcleo fixo com os elétrons se movimentando ao seu redor. Fazendo o uso do teorema adiabático, que afirma que se a perturbação num sistema for lenta o suficiente, o sistema possui a capacidade de ajustar-se à nova configuração e seu autoestado é conservado [79].

A partir desse teorema é possível aplicar a expansão adiabática na autofunção do sistema molecular:

$$\psi(\mathbf{r}, \mathbf{R}) = \phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R}) \quad (2.38)$$

na qual $\phi(\mathbf{r}; \mathbf{R})$ é a autofunção eletrônica que depende de forma explícita das coordenadas eletrônicas e parametricamente das coordenadas nucleares e $\chi(\mathbf{R})$ é a autofunção nuclear que descreve o movimento dos núcleos. Desse modo, substituindo essa expansão na função de onda da equação (2.38), encontra-se

$$\left[-\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{A=1}^M \sum_{i=1}^N \frac{Z_A}{r_{Ai}} + \sum_{A=1}^{M-1} \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} + \sum_{i=1}^{N-1} \sum_{j>i}^N \frac{1}{r_{ij}} \right] \phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R}) = E\phi(\mathbf{r}; \mathbf{R})\chi(\mathbf{R}). \quad (2.39)$$

A partir da equação acima, outras duas são estabelecidas

$$\left[-\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 \chi(\mathbf{R}) + \sum_{A=1}^{M-1} \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} - E \right] \chi(\mathbf{R}) = -\epsilon(\mathbf{R})\chi(\mathbf{R}), \quad (2.40)$$

e

$$\left[-\frac{1}{2} \sum_{i=1}^N \nabla_i^2 \phi(\mathbf{r}; \mathbf{R}) - \sum_{A=1}^M \sum_{i=1}^N \frac{Z_A}{r_{Ai}} + \sum_{i=1}^{N-1} \sum_{j>i}^N \frac{1}{r_{ij}} \right] \phi(\mathbf{r}; \mathbf{R}) = \epsilon(\mathbf{R})\phi(\mathbf{r}; \mathbf{R}). \quad (2.41)$$

O desacoplamento da equação 2.39 nas equações 2.40 e 2.41 é conhecido como a aproximação Born-Oppenheimer. Sendo a localização dos núcleos constante, pode escrever-se para uma configuração \mathbf{R}_a qualquer

$$\left[-\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{A=1}^M \sum_{i=1}^N \frac{Z_A}{r_{Ai}} + \sum_{i=1}^{N-1} \sum_{j>i}^N \frac{1}{r_{ij}} \right] \phi(\mathbf{r}; \mathbf{R}_a) = \epsilon(\mathbf{R}_a) \phi(\mathbf{r}; \mathbf{R}_a) \quad . \quad (2.42)$$

O termo em colchetes é o hamiltoniano eletrônico, e a equação (2.42) é a equação de Schrödinger eletrônica, em que $\epsilon(\mathbf{R}_a)$ representa a energia eletrônica e $\phi(\mathbf{r}; \mathbf{R}_a)$ é a autofunção eletrônica para uma molécula de configuração fixa (\mathbf{R}_a).

Como a energia eletrônica é calculada para várias configurações nucleares, há a possibilidade de substituir esse conjunto de energias na equação (2.40), o que leva a

$$\left[-\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + \sum_{A=1}^{M-1} \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} + \epsilon(\mathbf{R}) \right] \chi(\mathbf{R}) = E(\mathbf{R}) \chi(\mathbf{R}). \quad (2.43)$$

O valor do termo

$$\sum_{A=1}^{M-1} \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} + \epsilon(\mathbf{R})$$

depende da configuração nuclear e, além disso, ele representa a curva de energia potencial, sistema diatômico, ou superfície de energia potencial, sistema com mais de dois núcleos, para o movimento nuclear.

Já as soluções da equação (2.43) descrevem a translação, rotação e vibração das moléculas, ou seja, a dinâmica molecular, e ela é conhecida como equação de Schrödinger nuclear, e pode ser escrita como

$$\left[-\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + V(\mathbf{R}) \right] \chi(\mathbf{R}) = E(\mathbf{R}) \chi(\mathbf{R}). \quad (2.44)$$

em que

$$V(\mathbf{R}) = \sum_{A=1}^{M-1} \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} + \epsilon(\mathbf{R}) \quad (2.45)$$

é o potencial efetivo ao qual os núcleos são submetidos.

2.3.2 Equação de Schrödinger Eletrônica

A equação de Schrödinger eletrônica, que deve ser resolvida para cada configuração nuclear \mathbf{R} é:

$$\hat{H}_e \phi(\mathbf{r}; \mathbf{R}) = \epsilon(\mathbf{R}) \phi(\mathbf{r}; \mathbf{R}), \quad (2.46)$$

em que \hat{H}_e descreve o movimento dos elétrons para uma certa configuração nuclear e $\epsilon(\mathbf{R})$ é o espectro eletrônico, ou seja, corresponde às energias eletrônicas do sistema.

Para resolver a equação de Schrödinger eletrônica, precisa-se de métodos numéricos eficazes, pois resolvê-la analiticamente não é factível mesmo para uma molécula simples como o H_2 . Hoje, há uma quantidade razoável de métodos computacionais capazes de solucionar essa equação. Esses métodos podem ser classificados como semi-empíricos (SE), *ab initio* e DFT.

Teoria do Orbital Molecular

A teoria dos orbitais moleculares, uma ampliação das ideias de Bohr, considera que não só todos os elétrons de valência têm uma influência na estabilidade da molécula, mas também que os orbitais atômicos do nível de valência deixam de existir quando a molécula é formada, sendo substituídos por um novo conjunto de níveis energéticos correspondentes às novas distribuições da nuvem eletrônica, que é a densidade de probabilidade. Esses novos níveis de energia constituem uma propriedade da molécula como um todo e são chamados de orbitais moleculares (OM). De forma mais direta, há uma função de estado para cada elétron e sua composição

para formar a função de estado total dos N elétrons foi desenvolvida em primeira instância por Hartree, Roothan, Löwdin, dentre outros [76, 80].

A partir disso, tem-se o OM $\phi_i(r^\mu) = \phi_i(x^\mu, y^\mu, z^\mu)$ em que x^μ, y^μ, z^μ são as coordenadas do elétron μ e i indica o estado eletrônico. Sendo $\xi(S^\mu)$ a função de spin, o produto entre $\phi_i(r^\mu)$ e essa função resulta no spin-orbital molecular (SOM), definido por:

$$\psi(q^\mu) = \phi_i(x^\mu, y^\mu, z^\mu)\xi(S^\mu) = \phi_i(\mu)\xi(\mu) = \phi_i^\mu\xi^\mu, \quad (2.47)$$

$$\text{em que } \xi(S^\mu) = \begin{cases} \alpha(\mu) \uparrow (\text{spin up}) \\ \beta(\mu) \downarrow (\text{spin down}) \end{cases}.$$

Para satisfazer o princípio de exclusão de Pauli, a função de estado total do sistema de N-elétrons (Φ) precisa ser escrita como um produto antissimétrico dos spin-orbitais:

$$\Phi = (N!)^{\frac{1}{2}} \hat{A} \left\{ \psi_1^{(1)} \psi_2^{(2)} \dots \psi_N^{(N)} \right\}, \quad (2.48)$$

na expressão acima \hat{A} é o operador antissimetrizador, $\Phi_{a,b,\dots,k}(q^1, q^2, \dots, q^N)$ é representado por Φ e $\psi_j(q^\mu)$ por ψ_j^μ . A relação 2.48 pode ser reescrita como um determinante de Slater [81]:

$$\Phi = \frac{1}{\sqrt{N!}} \begin{bmatrix} \psi_1^{(1)} & \psi_2^{(1)} & \dots & \psi_N^{(1)} \\ \psi_1^{(2)} & \psi_2^{(2)} & \dots & \psi_N^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1^{(N)} & \psi_2^{(N)} & \dots & \psi_N^{(N)} \end{bmatrix} \quad (2.49)$$

A função de estado acima possui duas propriedades: os SOM's são linearmente independentes e o princípio de exclusão de Pauli é satisfeito reflexamente. Além disso, a constante $\frac{1}{\sqrt{N!}}$ é um fator de normalização e o determinante de Slater inverte o foco do problema de achar a autofunção para N-elétrons, ou seja, agora precisa-se resolver N autofunções para um único elétron [80].

Como os spin-orbitais ψ_k são linearmente independentes, é possível escrevê-los, formando um conjunto ortonormal, como:

$$\int \psi_k^*(\mu)\psi_l(\mu)d\tau_\mu = \delta_{kl}. \quad (2.50)$$

Assim sendo,

$$\int \Phi^*\Phi d\tau = 1. \quad (2.51)$$

Expandindo a equação 2.51

$$\int \Phi^*\Phi d\tau = \int \psi_1^{(1)\dagger}\psi_1^{(1)} d\tau_1 \int \dots \psi_N^{(N)\dagger}\psi_N^{(N)} d\tau_N = 1. \quad (2.52)$$

Constituindo-se a energia do sistema na forma da seguinte equação:

$$E = \int \Phi^*\hat{H}\Phi d\tau \equiv E[\Phi], \quad (2.53)$$

na qual o operador Hamiltoniano em unidades atômicas é dado por:

$$\hat{H} = \sum_{\mu} \hat{h}(\mu) + \frac{1}{2} \sum_{\mu \neq \nu} \frac{1}{r^{\mu\nu}}, \quad (2.54)$$

e na equação 2.54 $\hat{h}(\mu)$ é o operador hamiltoniano de uma partícula para o μ -ésimo elétron movendo-se no campo dos núcleos, além de ser um operador hermitiano e linear representado matematicamente, em u.a, por

$$\hat{h}(\mu) = -\frac{\nabla_{\mu}^2}{2} - \sum_A \frac{1}{r^{\mu A}}, \quad (2.55)$$

$\frac{\nabla_{\mu}^2}{2}$ é a energia cinética do elétron μ , e o segundo termo a energia coulombiana de interação elétron-núcleo.

O método Hartree-Fock busca uma solução aproximada para o estado fun-

damental de um sistema de elétrons em um átomo, uma molécula ou um sólido considerando apenas um determinante de Slater. A partir do determinante (2.49), a parte espacial e a dependência relacionada aos spins podem ser escritas separadamente:

$$\psi_a^1 = \phi_{r_1} \alpha(1) \quad (2.56)$$

ou

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

nessas equações acima α e β são, respectivamente, spin *up* \uparrow e *down* \downarrow . Se não houver restrição para os orbitais moleculares, o método é chamado de Hartree-Fock não Restrito(UHF), geralmente usado em sistemas de camada aberta. Sistemas de camada aberta também podem ser descritos por uma autofunção do tipo restrita, na qual as partes espaciais dos spins-orbitais duplamente ocupados, sejam necessariamente as mesmas. Nesse cenário, tem-se o método Hartree-Fock Restrito de Camada Aberta (ROHF). Já em sistemas que o estado eletrônico for camada fechada, há a restrição que cada orbital espacial seja associado a dois elétrons, um com spin α e outro com spin β , ou seja, sem elétrons desemparelhados, nesse caso o método é chamado Hartree-Fock Restrito(RHF) [79, 82].

Método Hartree-Fock

O teorema variacional diz que a autofunção mais próxima da solução exata, dentre todos os determinantes possíveis de Slater, é aquela para a qual a energia é mínima [82]:

$$\mathcal{E}(\psi) = \sum_a \langle \psi_a | h | \psi_a \rangle + \frac{1}{2} \sum_{a,b} (\langle \psi_a \psi_b | \psi_a \psi_b \rangle - \langle \psi_a \psi_b | \psi_b \psi_a \rangle), \quad (2.57)$$

Isso significa que, para encontrar a autofunção, a energia é minimizada em relação às alterações nos orbitais de um elétron, a partir do qual o determinante Slater foi construído. Para tanto, é necessário que as funções de *spin* sejam ortonormais, isto é:

$$\langle \psi_a | \psi_b \rangle - \delta_{ab} = 0 \quad (2.58)$$

Como o objetivo é encontrar os melhores orbitais que minimizem a energia, ou que pelo menos permaneçam estacionários em relação às mudanças em ψ_i , mantendo a propriedade de serem ortonormais. Pelo princípio variacional, a energia calculada será sempre maior que a energia real do estado fundamental. Nesse caso, é necessário achar o conjunto de orbitais moleculares que minimizam o valor de \mathcal{E} . Então, usando o "método dos multiplicadores indeterminados de Lagrange", no que concerne aos orbitais moleculares, a função Lagrange pode ser escrita como:

$$\mathcal{L}[\psi] = \mathcal{E}[\psi] - \epsilon_{ba} (\langle \psi_a | \psi_b \rangle - \delta_{ab}) \quad (2.59)$$

nessa equação, ϵ_{ba} são os multiplicadores de Lagrange. Como \mathcal{L} é real e $\langle \psi_a | \psi_b \rangle = \langle \psi_b | \psi_a \rangle^*$, os multiplicadores de Lagrange precisam ser elementos de uma matriz Hermitiana, $\epsilon_{ba} = \epsilon_{ab}^*$. Haja vista essas condições, quando cada orbital molecular é modificado por uma quantidade infinitesimal $\delta\psi$, \mathcal{L} é um mínimo e por isso $\delta\mathcal{L} = 0$, tem-se:

$$\delta\mathcal{L} = \delta\mathcal{E} - \sum_{a,b=1}^N \epsilon_{ba} (\langle \delta\psi_a | \psi_b \rangle + \langle \psi_a | \delta\psi_b \rangle) = 0 \quad (2.60)$$

Fixando \mathcal{J}_b e \mathcal{K}_b como os operadores de Coulomb e de Troca, respectivamente, por meio das equações a seguir

$$\mathcal{J}_b(1)\psi_a(1) = \left\langle \psi_b(2) \left| \frac{1}{r_{12}} \right| \psi_b(2) \right\rangle \psi_a(1) \quad (2.61)$$

e

$$\mathcal{K}_b(1)\psi_a(1) = \left\langle \psi_b(2) \left| \frac{1}{r_{12}} \right| \psi_b(2) \right\rangle \psi_b(1), \quad (2.62)$$

a energia com relação às mudanças nos orbitais é escrita da seguinte forma:

$$\delta\mathcal{E} = \sum_{a=1}^N (\langle \delta\psi_a | h | \psi_a \rangle + \langle \psi_a | h | \delta\psi_a \rangle) + \sum_{a,b=1}^N (\langle \delta\psi_a | \mathcal{J}_b - \mathcal{K}_b | \psi_a \rangle + \langle \psi_a | \mathcal{J}_b - \mathcal{K}_b | \delta\psi_a \rangle), \quad (2.63)$$

A partir da equação acima tem-se um novo operador, $\mathcal{F}(1) = h(1) + \sum_b (\mathcal{J}_b(1) - \mathcal{K}_b(1))$, conhecido como operador de Fock. Ele é um operador efetivo de um-elétron, associado a variação na energia. Assim, colocando a expressão energética em termos do operador de Fock:

$$\delta\mathcal{E} = \sum_{a=1}^N (\langle \delta\psi_a | \mathcal{F}(1) | \psi_a \rangle + \langle \psi_a | \mathcal{F}(1) | \delta\psi_a \rangle), \quad (2.64)$$

e

$$\delta\mathcal{L} = \sum_{a=1}^N (\langle \delta\psi_a | \mathcal{F}(1) | \psi_a \rangle + \langle \psi_a | \mathcal{F}(1) | \delta\psi_a \rangle) + \sum_{a,b=1}^N \epsilon_{ab} (\langle \delta\psi_a | \psi_b \rangle + \langle \psi_a | \delta\psi_b \rangle) = 0, \quad (2.65)$$

reorganizando os termos da equação 2.65, obtém-se a seguinte expressão, que é a equação Hartree-Fock(HF):

$$\mathcal{F}(1)\psi_a(1) = \sum_{b=1}^N \epsilon_{ba}\psi_b(1) \quad (2.66)$$

As equações de HF expressas em sua forma canônica, $\mathcal{F}(1)\psi'_a(1) = \epsilon_a\psi'_a(1)$, formam um conjunto de equações de autovalor. Um orbital de Fock específico só pode ser determinado uma vez que todos os outros orbitais ocupados são conhecidos e, por isso, métodos iterativos devem ser empregados para determinar os orbitais. Um conjunto de orbitais, que é uma solução para as equações de HF, são chamados

orbitais de campo autoconsistente (SCF, do inglês "*Self-Consistent Field*") [82].

Método de Hartree-Fock-Roothan

O método HF encontra uma solução aproximada para o estado fundamental de um sistema de elétrons num átomo considerando apenas um determinante de Slater. Como isso é feito de forma numérica, para sistemas poliatômicos o cálculo torna-se inviável. Essa barreira foi transposta pelas equações de Roothan-Hall [82], que são obtidas estendendo-se os conceitos do princípio variacional e da combinação linear de orbitais atômicos (LCAOs, do inglês "*Linear Combination of Atomic Orbitals*") às equações de HF, que serão obtidas através de um conjunto de funções de base conhecidas (tipo Gaussianas ou Slater). Com um conjunto adequado de funções de base, tem-se:

$$\phi_p(r) = \sum_{\nu=1}^k C_{\nu p} g_{\nu}(r), \quad (2.67)$$

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

As funções mais usadas em cálculos moleculares são as gaussianas. Ao escrever esse tipo de função em coordenadas cartesianas, obtém-se a expressão a seguir:

$$g(x, y, z) = N x^l y^m z^n e^{-\alpha r^2}, \quad (2.68)$$

em que l , m e n são números inteiros, $r = \sqrt{x^2 + y^2 + z^2}$ e N uma constante de normalização. Para as funções gaussianas cartesianas, é possível classificá-las de acordo com a soma dos expoentes l , m e n . Se essa soma for 0, a função é do tipo s, se for 1, do tipo p, 2, do tipo d e 3, do tipo f.

Substituindo 2.67 na equação de HF (2.66), chega-se à seguinte expressão:

$$\mathcal{F}(\mathbf{r}_1) \sum_{\nu} C_{\nu p} g_{\nu}(\mathbf{r}_1) = \epsilon_p \sum_{\nu} C_{\nu p} g_{\nu}(\mathbf{r}_1). \quad (2.69)$$

Multiplicando a equação 2.69 por $g_{\mu}^*(\mathbf{r}_1)$ e integrando, a equação integro-diferencial se transforma em uma equação matricial,

$$\sum_{\nu} C_{\nu p} \langle g_{\nu}(1) | \mathcal{F}(1) | g_{\nu} \rangle = \epsilon_p \sum_{\nu} C_{\nu p} \langle g_{\mu}(1) | g_{\nu} \rangle, \quad (2.70)$$

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

A equação acima possibilita a definição de duas matrizes, são elas: a matriz de superposição com os elementos determinados por

$$S_{\mu\nu} = \langle g_{\mu}(1) | g_{\nu} \rangle \quad (2.71)$$

e a matriz de Fock com os elementos descritos por

$$F_{\mu\nu} = \langle g_{\mu}(1) | \mathcal{F}(1) | g_{\nu} \rangle. \quad (2.72)$$

Por meio dessas definições, pode-se escrever

$$\sum_{\nu} F_{\mu\nu} C_{\nu p} = \epsilon_p \sum_{\nu} S_{\mu\nu} C_{\nu p}, \quad (2.73)$$

que em sua forma matricial fica

$$\mathbf{FC} = \mathbf{SC}\epsilon. \quad (2.74)$$

Essa expressão é a equação de Hartree-Fock-Roothan e ϵ é a matriz diagonal das energias orbitais $\epsilon_{pq} = \epsilon_p \delta_{pq}$.

2.3.3 Métodos Semi-empíricos

As equações de Hartree-Fock-Roothaan, no caso restrito ou no irrestrito, ficam limitadas à sistemas mais simples, pois o tempo computacional necessário cresce com a ordem de M^4 , em que M é o número de funções atômicas de base [83]. Outra dificuldade é o fato desse método não descrever de forma satisfatória a energia associada ao sistema, pois não abrange a energia de correlação eletrônica. Com o objetivo de se obter soluções para problemas moleculares com um grande número de elétrons, surgiu a linha de pesquisa dos métodos semi-empíricos. Nele o formalismo autoconsistente da teoria de orbitais moleculares usando a expansão em orbitais atômicos (SCF-LCAO) é mantido e o foco fica em torno de resolver as dificuldades próprias do método HF com a introdução de parâmetros. Há inúmeras aproximações usadas para implementar os métodos semi-empíricos e algumas delas são o CNDO (*Complete Neglect Differential Overlap*), o INDO (*Intermediate Neglect Differential Overlap*) e o NDDO (*Neglect of Diatomic Differential Overlap*) todos baseados na aproximação ZDO (*Zero Differential Overlap*) [84, 85].

No presente trabalho, o método semi-empírico usado para fazer uma primeira otimização, foi o *Parametric Model 6* (PM6). Esse método é um aprimoramento dos modelos AM1 (*Austin Model 1*) e PM3 (*Parametric Model 3*), pertencentes ao método MNDO (*Modified Neglect of Diatomic Overlap*) [86], feita por Stewart [87]. Além de usar um conjunto muito maior de dados de referência para os parâmetros, o PM6 também introduziu várias melhorias nos termos núcleo-núcleo. Ele substituiu as correções gaussianas para o termo núcleo-núcleo do AM1 e PM3 usando um termo de correção de núcleo-núcleo, originalmente introduzido por Voityuk e Rösch, que emprega parâmetros emparelhados em vez de parâmetros específicos de elemento. O PM6 ainda usa diferentes potenciais de repulsão núcleo-núcleo para N-H, O-H, C-C e Si-O com o intuito de corrigir falhas específicas na parametrização. Por fim, o método PM6 também adiciona d-orbitais na base atômica para certos elementos, assim como o método MNDO/d. O mais recente dos modelos PMx de Stewart é o

PM7, que inclui ainda uma correção de dispersão e correção de ligações de hidrogênio com base na correção H^+ . Adicionalmente, várias integrais semi-empíricas são modificadas por funções de comutação para impor convergência ao resultado clássico em distâncias maiores [88].

2.3.4 Teoria do Funcional de Densidade

A teoria do funcional de densidade (DFT) permite calcular todas as propriedades dos sistemas pela densidade de elétrons $\rho(r)$ que é uma função de três variáveis: $\rho(r) = f(x, y, z)$. Como a densidade é uma função da função de onda, refere-se a ela como um funcional. Para um sistema de N elétrons, a densidade eletrônica depende de $3N$ variáveis (ou $4N$, se levar o *spin* em consideração). Existem dois tipos de densidades de elétrons para sistemas polarizados por *spin*, um para *spin up*, $\rho(r) \uparrow$, e outro para elétrons *spin down*, $\rho(r) \downarrow$. O fato de que as propriedades do estado fundamental são funcionais da densidade eletrônica foi introduzido por Hohenberg e Kohn, em 1964, e é a estrutura básica dos métodos modernos de funcionais de densidade. Em 1965, os cientistas Kohn e Sham desenvolveram o método KS, que dava resultados mais acurados na aplicação da DFT para a energia cinética [78, 82]. Hohenberg e Kohn (HK) propuseram os seguintes teoremas:

- O potencial externo sentido pelos elétrons é um funcional da densidade eletrônica $\rho(r)$.
- A energia do estado fundamental é mínima para a densidade exata.

Dessa forma, partindo da aproximação de Born-Oppenheimer, o estado fundamental do sistema de elétrons é resultado das posições dos núcleos. No Hamiltoniano, a energia cinética de elétrons e a interação elétron-elétron se ajustam ao potencial externo V_{ext} , proveniente dos núcleos. Uma vez que o potencial externo começa a funcionar em um sistema, todo o resto, incluindo a densidade eletrônica, se organiza para fornecer a menor energia total possível do sistema.

A partir desse cenário, Kohn e Sham propuseram um funcional de energia para um sistema de partículas não-interagentes dado pela equação:

$$E_{KS}[\rho] = T_0[\rho] + U_H[\rho] + V_{ext}[\rho] + V_{XC}[\rho], \quad (2.75)$$

em que $T_0[\rho]$ é a energia cinética média, $U_H[\rho]$ é a energia coulombiana média entre os elétrons, $V_{ext}[\rho]$ é o potencial efetivo externo e $V_{XC}[\rho]$ é o termo de troca e correlação, com correções para a energia cinética e para a energia de Hartree.

O método KS descreve a repulsão elétron-elétron de Coulomb e um funcional universal $G[\rho]$:

$$E_\nu[\rho] = G[\rho] + \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} + \int \rho(r)\nu(r)dr, \quad (2.76)$$

sendo $G[\rho] = T_s[\rho] + E_{xc}[\rho]$. $T_s[\rho]$ é o funcional de energia cinética de um sistema de elétrons não-interagentes, que possui a mesma densidade eletrônica do sistema interagente. $E_{xc}[\rho]$ abrange o termo de troca e correlação, interação elétron-elétron, e a parte residual da energia cinética, $T[\rho] - T_s[\rho]$, sendo $T[\rho]$ a energia cinética exata para o sistema interagente.

Para partículas não-interagentes, é possível escrever um Hamiltoniano com um potencial local efetivo, $\nu_{ef}(r)$:

$$H^{KS} = -\frac{1}{2}\nabla^2 + \nu_{ef}(r). \quad (2.77)$$

A autofunção ψ^{KS} do estado fundamental para partículas que não interagem é aproximada por um produto anti-simétrico de N funções de onda de um elétron, $\psi_i(r_i)$, podendo ser escrita como um determinante de Slater:

$$\psi^{KS} = \frac{1}{\sqrt{N!}} \begin{bmatrix} \psi_1^{KS}(r_1) & \psi_2^{KS}(r_1) & \cdots & \psi_N^{KS}(r_1) \\ \psi_1^{KS}(r_2) & \psi_2^{KS}(r_2) & \cdots & \psi_N^{KS}(r_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1^{KS}(r_N) & \psi_2^{KS}(r_N) & \cdots & \psi_N^{KS}(r_N) \end{bmatrix} \quad (2.78)$$

Assim, os orbitais de Kohn-Sham(KS), ψ_i^{KS} , são determinados através da equação de Schrödinger de um elétron:

$$\left(-\frac{1}{2}\nabla^2 + \nu_{ef}(r) \right) \psi_i^{KS} = \varepsilon_i \psi_i^{KS}, \quad (2.79)$$

em que ε_i são os autovalores e ψ_i são as autofunções.

O operador KS depende apenas de r , e não do índice (natureza) dos elétrons, ou seja, é o mesmo para todos os elétrons. Os orbitais KS, ψ_i^{KS} , podem ser usados de forma imediata para calcular a densidade, e a partir dela pode-se calcular um potencial efetivo otimizado. Por conseguinte, a ligação entre o sistema não interagente e o sistema real a partir da escolha desse potencial otimizado, de modo que a densidade eletrônica resultante seja igual à do estado fundamental, é dada por

$$\rho_s(r) = \sum_i^N 2|\psi_i^{KS}(r)|^2 = \rho_0(r). \quad (2.80)$$

A densidade também pode ser utilizada para o cálculo da energia total, equação 2.75, e também para calcular a energia cinética $T_s[\rho]$, que é obtida dos orbitais correspondentes:

$$T_s[\rho] = \sum_i^N \left\langle \psi_i^{KS} \left| -\frac{1}{2}\nabla_i^2 \right| \psi_i^{KS} \right\rangle. \quad (2.81)$$

Por meio da minimização da equação 2.76 e considerando as funções de um elétron ortonormais, obtém-se o potencial efetivo da seguinte forma:

$$\nu_{ef}(r) = \nu(r) + \int \frac{\rho(r_1)}{|r - r_1|} dr_1 + \nu_{xc}(r). \quad (2.82)$$

Nessa equação, $\nu_{xc} = \frac{\delta E_{XC}[\rho]}{\delta \rho(r)}$. O termo E_{XC} é o potencial de troca e correlação, que inclui todos os termos de interação entre partículas. O tratamento desse termo é complexo e um obstáculo para a DFT, mas existem aproximações que definem esse potencial e, conseqüentemente, trazem confiabilidade para a teoria.

Funcional de Troca e Correlação

Uma aproximação muito usada no formalismo da DFT é expressar o funcional de troca e correlação em termos do gradiente da densidade de carga total. Ela é conhecida como expansão generalizada em termos do gradiente (do inglês, *Generalized Gradient Approximation* - GGA), e o seu funcional é escrito como a soma de dois termos, um para a troca e outro para a correlação:

$$E_{XC}^{GGA} = E_X^{GGA} + E_C^{GGA}. \quad (2.83)$$

Para a energia de troca, um dos funcionais mais usados é o de Becke, o B88, dado por

$$E_X^{B88} = E_X^{LSDA} - b \sum_{\sigma=\alpha,\beta} \int \frac{(\rho^\sigma)^{\frac{4}{3}} \chi_\sigma^2}{1 + 6b \chi_\sigma \sinh^{-1} \chi_\sigma} d\mathbf{r}, \quad (2.84)$$

nessa equação $\chi_\sigma \equiv |\nabla \rho^\sigma| / (\rho^\sigma)^{4/3}$, $\sinh^{-1} = \ln[x + (x^2 + 1)^{1/2}]$ e b é um fator semiempírico igual a 0,0042u.a. auferido após o cálculo de energia de troca pelo método HF para vários átomos. E_X^{LSDA} é

$$E_X^{LSDA} = -\frac{3}{4} \left(\frac{6^{\frac{1}{3}}}{\pi} \right) \int [(\rho^\alpha)^{\frac{4}{3}} + (\rho^\beta)^{\frac{4}{3}}] d\mathbf{r}. \quad (2.85)$$

Outros funcionais de correlação corrigidos pelo gradiente muito usados são

o de Lee-Yang-Parr (LYP) e os de Becke. Além desses, os funcionais híbridos são bastante empregados. Eles são uma combinação linear dos funcionais DFT puros e o termo de troca obtido pelo método HF(E_X^{HF}). Para sistemas de camada fechada, o termo de troca é dado pela seguinte expressão:

$$E_X = -\frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n \left\langle \phi_i^{KS}(1) \phi_j^{KS}(2) \left| \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \phi_j^{KS}(1) \phi_i^{KS}(2) \right\rangle. \quad (2.86)$$

Nesta pesquisa, para a otimização dos sistemas moleculares após o uso do método semi-empírico PM6, foi usado o funcional híbrido B3LYP, que tem três parâmetros de Becke mais os funcionais de correlação LYP. O B3LYP é definido pela expressão:

$$E_{XC}^{B3LYP} = (1 - a_0 - a_x) E_X^{LSDA} + a_0 E_X^{HF} + (1 - a_c) E_C^{VWN} + a_c E_C^{LYP}, \quad (2.87)$$

sendo o termo de correlação E_C^{VWN} proveniente da aproximação da densidade local de *spin* - LSDA (do inglês, *Local-Spin-Density Approximation*), desenvolvida por Vosko-Wilk-Nusair [89].

Devido à sua eficiência computacional e precisão razoável, a DFT com aproximações semilocais à energia de troca e correlação como, por exemplo, a LSDA e o padrão Perdew-Burke-Ernzerhof (PBE) da GGA, é um dos métodos de estrutura eletrônica mais amplamente usado em várias áreas. As aproximações semilocais exibem um cancelamento de erro bem compreendido entre troca e correlação em regiões de ligação. Assim, alguns efeitos de correlação de médio alcance, importantes para ligações fortes e fracas, são realizados pela parte da troca da aproximação. No entanto, é sabido que essas aproximações não podem produzir corretamente forças de dispersão assintóticas de longo alcance. Isso levanta dúvidas sobre a adequação de

aproximações semi-locais para a descrição de interações fracas, incluindo ligações de hidrogênio e interações de van der Waals, mesmo perto do equilíbrio. No entanto, essa deficiência é solucionada por funcionais conhecidos como meta-GGAs (MGGA). Comparados aos GGAs, que usam a densidade $\rho(\mathbf{r})$ e seu gradiente $\nabla\rho$, os MGGA's adicionam a densidade de energia cinética positiva, $\tau = \sum_i \frac{|\nabla\psi_i|^2}{2}$, dos orbitais ocupados ψ_i [90]. O funcional MGGA usado neste trabalho para o cálculo de energia de *single-point* foi o funcional M06-L [91]. A energia de troca desse funcional, com σ referente ao momento angular de *spin* do elétron que pode assumir dois valores, $\alpha(+1/2)$ e $\beta(-1/2)$, é escrita da seguinte forma:

$$E_X^{M06-L} = \sum_{\sigma} \int [F_{X\sigma}^{PBE}(\rho_{\sigma}, \nabla\rho_{\sigma})f(w_{\sigma}) + \varepsilon_{X\sigma}^{LSDA}h_X(x_{\sigma}, z_{\sigma})]dr, \quad (2.88)$$

sendo $h_X(x_{\sigma}, z_{\sigma})$, com $d_5 = 0$, definido por

$$h(x_{\sigma}, z_{\sigma}) = \left(\frac{d_0}{\gamma(x_{\sigma}, z_{\sigma})} + \frac{d_1x_{\sigma}^2 + d_2z_{\sigma}}{\gamma_{\sigma}^2(x_{\sigma}, z_{\sigma})} + \frac{d_3x_{\sigma}^4 + d_4x_{\sigma}^2z_{\sigma} + d_5z_{\sigma}^2}{\gamma_{\sigma}^3(x_{\sigma}, z_{\sigma})} \right), \quad (2.89)$$

$F_{X\sigma}^{PBE}(\rho_{\sigma}, \nabla\rho_{\sigma})$ é a densidade da energia de troca do modelo de troca de Perdew-Burte-Ernzerhof (PBE), $\varepsilon_{X\sigma}^{LSDA}$ é a LSDA para a troca

$$\varepsilon_{X\sigma}^{LSDA} = \frac{3}{2} \left(\frac{3}{4\pi} \right)^{\frac{1}{3}} \rho_{\sigma}^{\frac{4}{3}} \quad (2.90)$$

e $f(w_{\sigma})$ é o fator de aprimoramento da densidade de energia cinética de *spin*

$$f(w_{\sigma}) = \sum_{i=0}^m a_i w_{\sigma}^i, \quad (2.91)$$

nessa expressão w_{σ} é uma função de t_{σ} , que é uma função da densidade de energia cinética de *spin* (τ_{σ}) e da densidade de *spin* ρ_{σ} ,

$$w_\sigma = \frac{t_\sigma - 1}{t_\sigma + 1}, \quad (2.92)$$

$$t_\sigma = \frac{\tau_\sigma^{LSDA}}{\tau_\sigma}, \quad (2.93)$$

e

$$\tau_\sigma \equiv \frac{3}{10}(6\pi^2)^{\frac{2}{3}}\rho_\sigma^{\frac{5}{3}}. \quad (2.94)$$

A parte da energia de correlação do funcional M06-L para spins opostos é dada por:

$$E_C^{\alpha\beta} = \int e_{\alpha\beta}^{UEG} [g_{\alpha\beta}(x_\alpha, x_\beta) + h_{\alpha\beta}(x_{\alpha\beta}, z_{\alpha\beta})] dr, \quad (2.95)$$

sendo, na equação acima,

$$g_{\alpha\beta}(x_\alpha, x_\beta) = \sum_{i=0}^n c_{C\alpha\beta,i} \left(\frac{\gamma_{C\alpha\beta}(x_\alpha^2 + x_\beta^2)}{1 + \gamma_{C\alpha\beta}(x_\alpha^2)} \right)^i \quad (2.96)$$

e $h_{\alpha\beta}(x_{\alpha\beta}, z_{\alpha\beta})$ definido pela equação 2.89, com $x_{\alpha\beta}^2 \equiv x_\alpha^2 + x_\beta^2$ e $z_{\alpha\beta} \equiv z_\alpha + z_\beta$.

Para spins paralelos tem-se:

$$E_C^{\sigma\sigma} = \int e_{\sigma\sigma}^{UEG} [g_\sigma(x_\sigma + h_{\sigma\sigma}(x_\sigma, z_\sigma))] D_\sigma dr, \quad (2.97)$$

em que

$$g_{\sigma\sigma}(x_\sigma) = \sum_{i=0}^n c_{C\sigma\sigma,i} \left(\frac{\gamma_{C\sigma\sigma}(x_\sigma^2)}{1 + \gamma_{C\sigma\sigma}(x_\sigma^2)} \right)^i \quad (2.98)$$

e $h_{\sigma\sigma}(x_\sigma, z_\sigma)$ é definido pela equação 2.89. D_σ é o fator de correção de auto-interação, que para sistemas de 1-elétron ele desaparece,

$$D_\sigma = 1 - \frac{x_\sigma^2}{4(z_\sigma + C_F)}. \quad (2.99)$$

Nas equações 2.95 e 2.97, $e_{\alpha\beta}^{UEG}$ e $e_{\sigma\sigma}^{UEG}$ são a densidade da energia de correlação no limite do gás de elétrons uniforme (UEG, do inglês *uniform-electron-gas*) para os casos de spins anti-paralelos e spins paralelos, respectivamente, e podem ser extraídas da densidade total da energia de correlação do UEG. A energia total de correlação é dada pela seguinte equação:

$$E_C = E_C^{\alpha\beta} + E_C^{\alpha\alpha} + E_C^{\beta\beta}. \quad (2.100)$$

Os parâmetros não lineares $\gamma_{C\alpha\beta} = 0,0031$ e $\gamma_{C\sigma\sigma} = 0,006$ são dados pelo trabalho de Zhao e Truhlar [91].

Conjuntos de Funções de Base

Os primeiros tipos de funções de base utilizados foram os orbitais do tipo Slater (STO - do inglês, "*Slater Type Orbital*"), que apresentam semelhanças com as funções hidrogenóides. As funções de Slater não permitem uma resolução analítica rápida das integrais de dois elétrons de mais de um centro ou ainda integrais de atração nuclear de três centros. Para resolver esse óbice da solução analítica de integrais de energia, funções do tipo gaussiana (GTO - do inglês, "*Gaussian Type Orbitals*") foram introduzidas [92].

Tendo em vista a melhora do desempenho computacional, usa-se a combinação linear de gaussianas, em que duas ou mais gaussianas primitivas (PGTOs, do inglês "*Primitive Gaussian-Type Orbitals*") são combinadas linearmente, formando gaussianas contraídas. Essas contrações são chamadas de contração de orbitais do tipo gaussiana (CGTO, do inglês "*Contracted Gaussian-Type Orbitals*") e a eficiência computacional está no fato que apenas as CGTOs são calculadas. Há incontáveis conjuntos de bases voltados para o cálculo de sistemas poliatômicos, por isso será

descrito nessa seção apenas as bases usadas neste trabalho.

O conjunto de base Pople aplicado nesta pesquisa corresponde ao conjunto de bases K-nlmG ou k-nlm++G, conhecido também como conjunto de bases de valência separada. Nesse conjunto, k é a quantidade de PGTOs contraídas para os orbitais do núcleo, nlm indica a valência separada, ou seja, refere-se à quantidade de PGTOs utilizadas para descrever os orbitais de valência, mais externos. Um único + denota funções de difusão, (s, p_x, p_y, p_z) , para os átomos pesados e ++ indica funções de difusão para os átomos pesados e para os hidrogênios [93]. As bases usadas no presente trabalho foram a 6-311G(d,p), para otimização do sistema molecular, e a 6-311+G(d,p), para o cálculo de *single point*. Nessas bases, 6 funções gaussianas contraídas descrevem cada orbital do núcleo, 3 GTOs contraídas referem-se à parte mais interna da camada de valência, 1 GTO diz respeito à parte mais externa da camada de valência e mais 1 gaussiana para descrever a parte externa da camada de valência mais distante. Dessa forma, a função de base escolhida trata o orbital interno por meio de três gaussianas, e os orbitais médios e externos como gaussianas únicas (6-311 G).

Além disso, essas bases permitem acrescentar funções de polarização e funções difusas. Nesta pesquisa, foi acrescentada uma função de polarização de simetria d aos átomos pesados, e funções de polarização de simetria p aos átomos mais leves, hidrogênio (d,p). As funções de polarização retêm informações sobre as deformações ocorridas nos orbitais que participam de interações. Já as funções difusas são dotadas de informações sobre uma região maior do espaço dos orbitais ocupados e são necessárias sempre que haja elétrons fracamente ligados. O uso de funções difusas melhora a descrição de sistemas com densidade eletrônica relevantes, que estejam distantes do núcleo [79].

2.3.5 O Método Integrado de N-camadas: ONIOM

A partir de meados da década de 70, surgiram os métodos híbridos de mecânica quântica (QM) e mecânica molecular (QM/MM). Esses métodos consistem na combinação de diferentes tipos de aproximações, *ab initio*, semiempírico e mecânica molecular, tentando assim aproveitar as vantagens de cada uma e contornando algumas de suas limitações.

A aproximação QM/MM envolve a separação do sistema em regiões QM e MM, onde o potencial molecular é determinado parcialmente pelo cálculo de estrutura eletrônica por mecânica quântica e parcialmente pelo campo de força da MM. O motivo dessa divisão são vários fatores: tamanho dos fragmentos e natureza dos átomos, moléculas presentes no sistema, tipo de propriedade de interesse, etc. Isso significa que, nas partes consideradas mais difíceis (sistema "modelo"), nas quais há átomos não parametrizados, interações incomuns, ou átomos diretamente participantes do estado de transição (ET) de uma reação, são aplicados métodos precisos e mais demandantes computacionalmente, enquanto o restante da molécula (sistema "real"), constituída por interações consideradas mais simples, é tratado por métodos rápidos, baseados na mecânica molecular [78, 94]. Para os métodos QM/MM, a energia total para o sistema completo, $E_{QM/MM}$, é escrita de forma aditiva, fazendo uma soma da energia do sistema modelo pelo método QM (E_{QM}), a energia do sistema real pelo método MM (E_{MM}), e as interações (E_{QM-MM}) entre o sistema modelo QM e o sistema real MM:

$$E_{QM/MM} = E_{QM} + E_{MM} + E_{QM-MM}. \quad (2.101)$$

As energias dos dois sistemas e as interações entre eles são adicionadas para obter a energia do sistema como um todo. O acoplamento Hamiltoniano de QM/MM (E_{QM-MM}), as interações entre os sistemas "modelo" e "real", geralmente inclui: interações ligadas para ligações covalentes que cortam o limite QM/MM (alongamento, contribuições flexíveis e torcionais) e interações não-ligadas (van der Waals e inte-

rações eletrostáticas). Todas as interações QM-MM são classicamente avaliadas, embora as interações eletrostáticas possam ser computadas "semiclassicamente".

A primeira implementação dessa nova geração de métodos ficou conhecida como *Integrated Molecular Orbital and Molecular Mechanics* (IMOMM), esse método pode ser constituído de métodos MO (*Molecular Orbital*) em diferentes níveis de aproximação, a essa integração de dois métodos MO deu-se o nome de método IMOMO. Como esses métodos foram bem sucedidos, novos estudos foram realizados para atingir uma maior versatilidade nos cálculos de estrutura eletrônica. Assim, apareceu o método ONIOM [95], que pode ser descrito como uma superposição de cálculos como numa "cebola", além disso, ele abrange os métodos IMOMM e IMOMO. De acordo com o esquema teórico desse método, é possível dividir qualquer sistema molecular em diferentes níveis ligados seguindo a ordem mais conveniente do problema estudado. Cada parte pode ser tratada sob qualquer nível de cálculo e ao integrar-se os resultados obtidos nessas partes, obtém-se uma extrapolação atingindo valores de energia mais precisos sobre todo o sistema [96]. Quando dois níveis diferentes de aproximação são integrados, que foi o modelo usado para este trabalho (ONIOM2), o método ONIOM segue o diagrama apresentado na Figura 2.3.

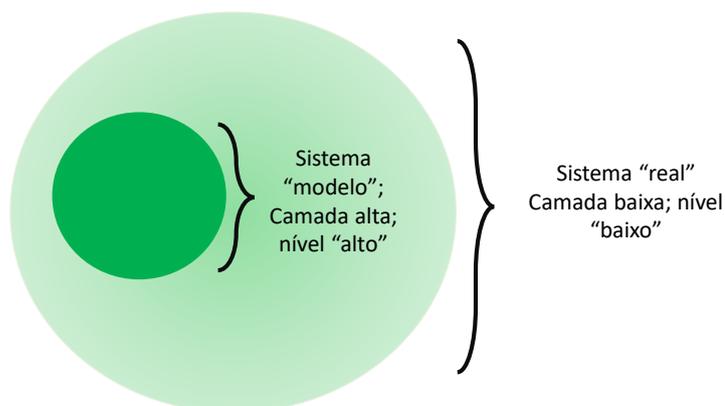


Figura 2.3: Diagrama esquemático do método ONIOM para duas camadas.
 Fonte: Elaborado pela autora

Ao contrário dos esquemas aditivos de QM/MM que estimam a E_{QM-MM} (2.101), o IMOMM é um método "subtrativo" ou "extrapolativo": a energia total do sistema completo ("real") é calculada como a energia MO (ou QM) do sistema modelo ($E_{(QM,model)}$) mais a energia MM do sistema real ($E_{(MM,real)}$) e menos a energia MM do sistema modelo ($E_{(MM,model)}$), que remove a "contagem dupla" das contribuições MM:

$$E_{IMOMM} = E_{ONIOM2(QM:MM)} = E_{(QM,model)} + E_{(MM,real)} - E_{(MM,model)}, \quad (2.102)$$

Embora o ONIOM possa ser usado como um método QM/MM de duas camadas, ele também pode, exclusivamente, combinar diferentes métodos de QM e QM e, dessa forma, ser estendido para várias camadas [96].

Dessa forma, tendo o diagrama como base, a equação para a energia final de um sistema com duas camadas é definida pela seguinte expressão:

$$E = E_{(high, model)} + E_{(low, real)} - E_{(low, model)}, \quad (2.103)$$

sendo *high* e *low* referente aos níveis de teoria "alto" e "baixo" de aproximação, enquanto *model* e *real* descrevem os sistemas de tamanho "pequeno" ("modelo") e "real", respectivamente. No caso desta pesquisa, os sistemas moleculares estudados passaram por três etapas: duas de otimização e uma de *single point*. Em todas elas, no nível "baixo" foi implementado o campo de força AMBER como método de MM. Para nível "alto" (QM), a aproximação usada foi o PM6, na primeira otimização, e B3LYP para as últimas duas etapas.

Capítulo 3

Resultados e Discussões

No presente capítulo será apresentado o comportamento das estruturas secundárias do peptídeo $A\beta_{1-42}$ ao longo da DM. Esse comportamento será avaliado via RMSD. Logo após, serão mostradas as nuvens de solução advindas do *docking*, as quais apresentam as poses preferenciais do ligante. Em seguida, serão descritas as energias alcançadas via QM/MM em função das distâncias encontradas entre peptídeo e os ligantes para as configurações escolhidas. Por fim, a partir dessas conformações de interesse, serão exibidas as prováveis interações entre o alvo molecular e os ligantes que levam a um mínimo de energia dentro dos resultados encontrados nesta pesquisa. Juntamente com essas possíveis interações, serão mostrados os orbitais de fronteira relativos aos compostos naturais. Neste capítulo, também são descritos os detalhamentos da aplicação de cada um dos métodos empregados.

3.1 Resultados da Dinâmica Molecular

3.1.1 Configurações Iniciais e Equilibração do Sistema

A simulação de dinâmica molecular para o receptor, tanto a partir da sua forma cristalográfica de α -hélice (Código PDB:1Z0Q, modelo 25) quanto de folha- β (Código PDB:2NAO, cadeia A), foi realizada no pacote computacional NAMD [97],

com o campo de força CHARMM36 [98]. As moléculas de água foram adicionadas usando o modelo TIP3P [60] e as caixas cúbicas de simulação com dimensões 67 x 51 x 40(Å)³, para a α -hélice, e 36 x 58 x 77(Å)³, para a folha- β (Figura 3.1). O sistema foi neutralizado e moléculas de NaCl foram adicionadas para reproduzir a solução fisiológica de concentração 0,15 μ M. A distância entre pares de interações não-ligados teve o *cut off* de 12 Å. O PME foi utilizado para as interações eletrostáticas de longo alcance.

Nessa simulação, considerou-se constante o número de moléculas, a pressão e a temperatura (*ensemble* NPT), usando a dinâmica de Langevin para manter constantes a temperatura e a pressão, 300 K e 1 atm, respectivamente. As trajetórias foram acumuladas com um passo de tempo, *timestep*, de 1 fs. O tempo de amostragem total, incluindo a termalização e a relaxação, foi de 211ns para a conformação em α -hélice e 200 ns para a estrutura em folha- β , a evolução temporal e a mobilidade do sistema foram computados pelo cálculo do RMSD. O tempo de simulação para estágio de termalização e relaxação do sistema foi de 11 ns para as duas conformações iniciais do peptídeo: α -hélice e folha- β . Os demais 200 ns, para o primeiro caso, e 189 ns, para o segundo, consistem na amostragem da dinâmica do sistema. Cada arquivo de trajetória corresponde a 1 ns composto por 100 frames, totalizando 21042 frames, para o formato α -hélice, e 19938 frames, para o formato folha- β , na simulação completa.

As imagens tridimensionais, assim como a construção e visualização do sistema para a simulação de DM, foram produzidas usando o programa gráfico *Visual Molecular Dynamics* (VMD) [99].

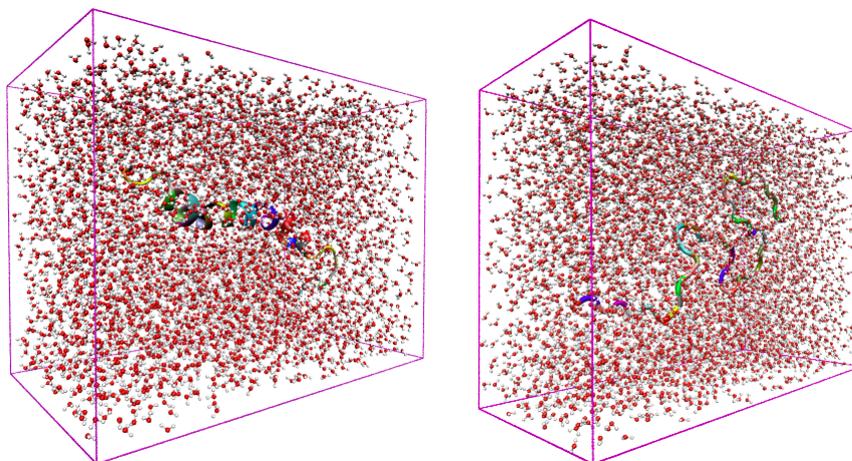


Figura 3.1: Caixas das simulações com 13287 átomos, para a conformação em α -hélice (à esquerda), e 14489, para a forma de folha- β (à direita), sendo 627 átomos do peptídeo e o restante dividido em moléculas de água e íons do NaCl.

3.1.2 Análise de RMSD das Conformações do Peptídeo A β

O conhecimento das flutuações conformacionais das proteínas é importante para compreender as funções biológicas desenvolvidas pelos sistemas, pois a relação entre movimento e função proteica não é totalmente conhecida [74], ainda mais em doenças neurodegenerativas ocasionadas por erro do *fold*ing proteico (Apêndice B), como as causas de demência apresentadas neste trabalho, em que uma proteína específica modifica sua forma nativa. Assim, é de suma importância conhecer essas variações e mapeá-las.

A dinâmica conformacional do peptídeo A β nas suas várias extensões, como a A β_{1-42} e a A β_{1-40} , e de seus segmentos vem sendo amplamente estudada. Uma vez que, é necessário compreender melhor o mecanismo pelo qual esse peptídeo altera sua forma nativa (α -hélice ou *coil*) para a forma em folha- β levando à formação de placas senis. Nesse sentido, foi realizada a análise das flutuações dos aminoácidos da A β_{1-42} , para ambas as conformações, a partir de 11 ns de simulação, pois para controle da simulação de DM, nos primeiros 11 nanosegundos o *backbone* do peptídeo foi liberado aos poucos até atingir uma dinâmica em que o peptídeo pôde movimentar-se livremente, sem qualquer rigidez, até alcançar o equilíbrio termodinâmico.

Dessa forma, visando uma melhor descrição dessas flutuações para o alvo molecular estudado neste trabalho, a verificação da estabilidade do peptídeo $A\beta$, tanto no formato α -hélice quanto no de folha- β , foi feita utilizando o RMSD das trajetórias no decorrer da simulação, calculado sobre o *backbone* da proteína para todos os frames gerados. Por conseguinte, todas as estruturas foram sobrepostas e as flutuações em torno da posição média, tendo como valor de referência o primeiro conjunto de coordenadas do sistema, são mostradas nos gráficos da presente seção.

Assim sendo, o peptídeo $A\beta_{1-42}$ será descrito considerando toda sua extensão, que tem como estrutura primária os aminoácidos ¹DAEFRHDSGY ¹¹EVHHQKLVFF ²¹AEDVGSNKG ³¹IIGLMVGGVV ⁴¹IA (detalhes dos códigos e estruturas dos aminoácidos estão no Apêndice A). A partir dessa divisão, também serão analisadas algumas regiões específicas.

Conformação Inicial de α -hélice

Nesta parte da seção, serão feitas discussões sobre os resultados obtidos considerando a divisão do peptídeo nas suas duas regiões de α -hélice, segmentos compreendidos pelos aminoácidos Ser8-Gly25 e Lys28-Gly38, e nas regiões que tenham alguma particularidade a ser mencionada.

No gráfico da Figura 3.2, tem-se o RMSD de toda extensão do peptídeo para sua conformação em α -hélice desde o princípio da DM. Nota-se que a partir da marca de 126 nanosegundos de simulação as variações em torno da posição de equilíbrio, até os 211 nanosegundos finais, alcançam o equilíbrio, flutuando em torno de 7Å indicando alta mobilidade do peptídeo. Mesmo com um RMSD aparentemente alto, o valor alcançado nessa simulação é menor do que o descrito na literatura [25, 31, 100].

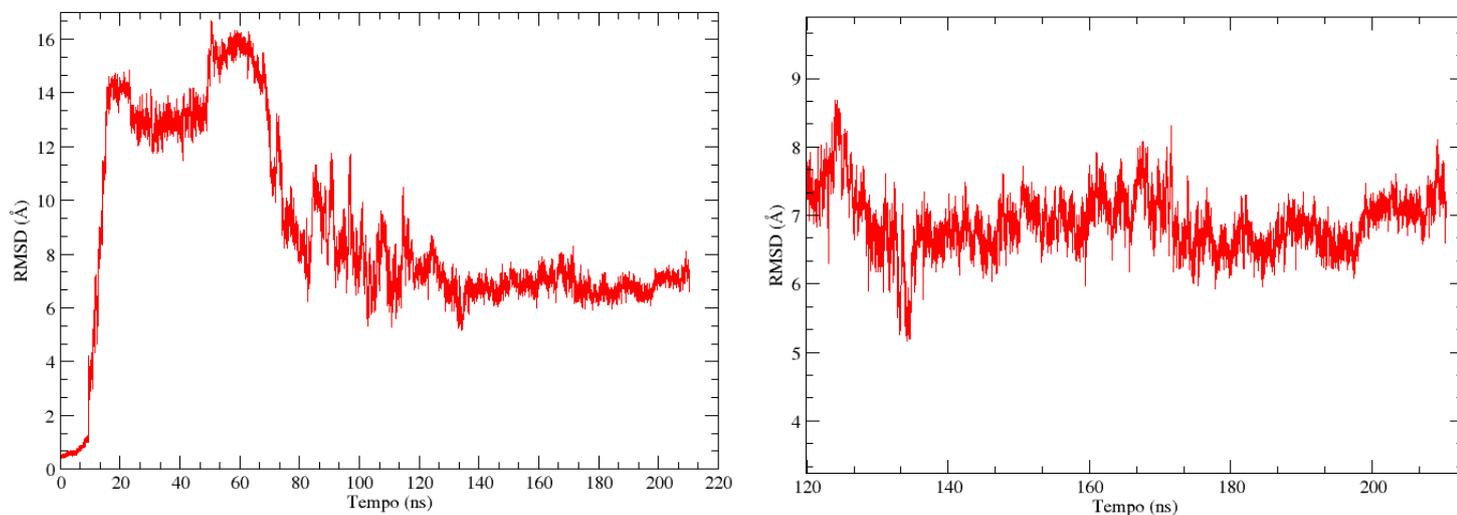


Figura 3.2: RMSD da extensão total do peptídeo $A\beta_{1-42}$ com conformação inicial de α -hélice durante 211ns de simulação de DM. O segundo gráfico detalha o RMSD a partir do momento que o peptídeo flutua na posição de equilíbrio.

A Figura 3.3 ilustra as mudanças do peptídeo a cada 20 ns da simulação. Aos 20 ns, observa-se que o peptídeo começa a perder a estrutura de α -hélice na região dos aminoácidos Lys28-Val39 e ela é completamente desfeita aos 40 ns. A partir de 80 ns em diante, a região que compreende os aminoácidos Ser8-Asp23 fixa sua forma de α -hélice, ao passo que a região dos aminoácidos Gly29-Val36 assume a forma de folha- β e, aos 120 ns, a região compreendida pelos aminoácidos Ile31-Leu34 formam uma dobra (essa aproximação pode ser vista através da Figura 3.4). A desestruturação da α -hélice próxima ao C-terminal do peptídeo, coloca os resíduos hidrofóbicos presentes nessa região, que são quase todos com exceção das glicinas (aminoácidos 33, 37 e 38), em contato com o solvente, o que provavelmente contribui na oligomerização da $A\beta_{1-42}$ para minimizar o acesso do solvente a essa área [31, 35, 101]. A instabilidade da α -hélice dessa região se deve, possivelmente, à presença de uma única ponte de hidrogênio entre os *backbones* dos aminoácidos Ile32 e Met35 e de interações hidrofóbicas, que são interações mais fracas e mais fáceis de serem desfeitas.

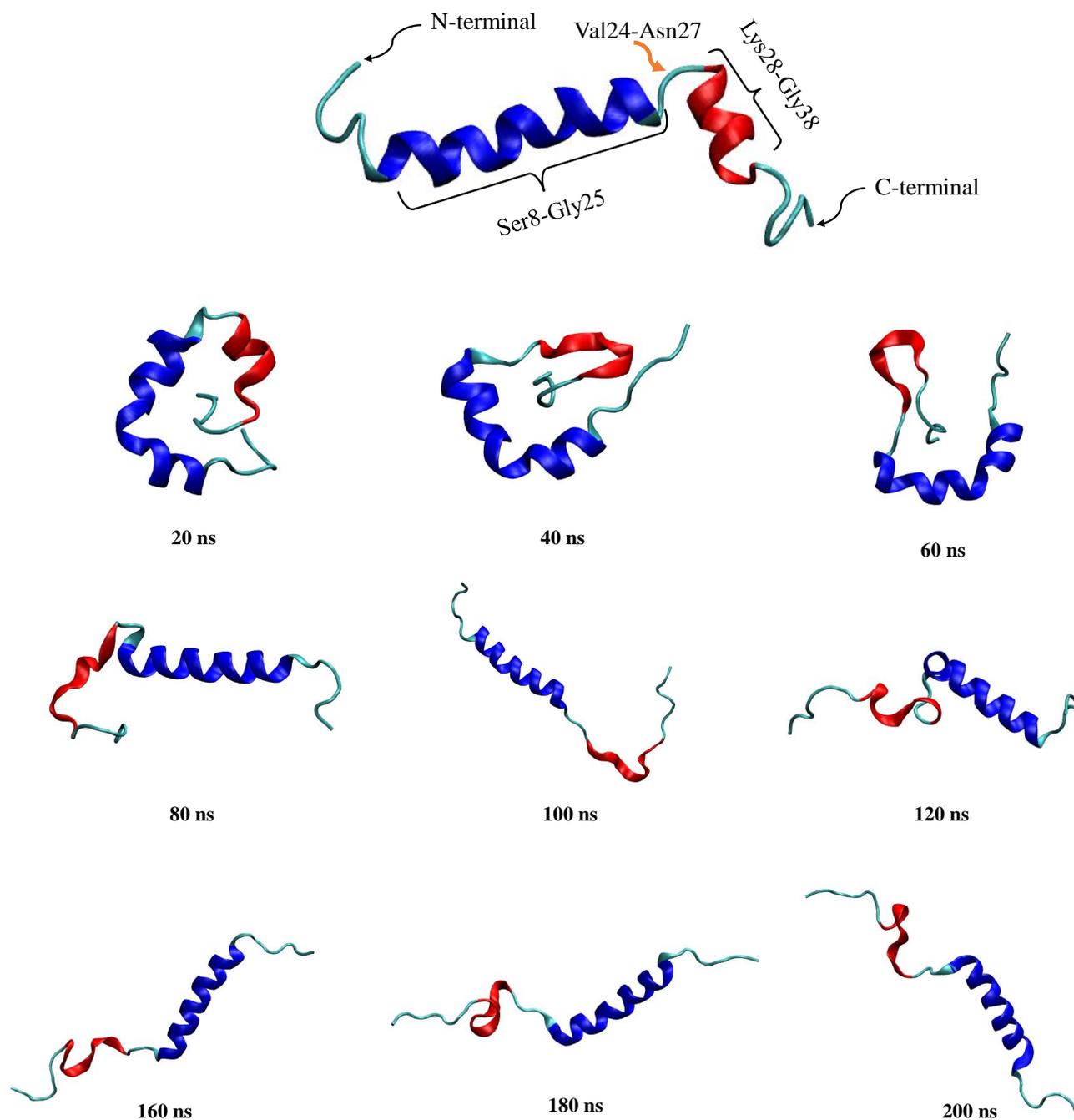


Figura 3.3: Mudanças sofridas pelo peptídeo A β_{1-42} a cada 20 ns da simulação. A primeira figura mostra a conformação α -hélice do cristal(PDB:1Z0Q) e suas principais estruturas. Na sequência estão dispostas representações das modificações do peptídeo ao longo dos 200 ns da simulação.

A dobra da região Ile31-Leu34 é composta por uma maioria de aminoácidos apolares com exceção da Gly33. Esses aminoácidos são reportados na literatura como componentes do núcleo de pentâmeros, forma preferencial de agregação do peptídeo $A\beta_{1-42}$ [35]. Logo, pode ser que essa seja uma formação que favoreça o processo de amiloidose.

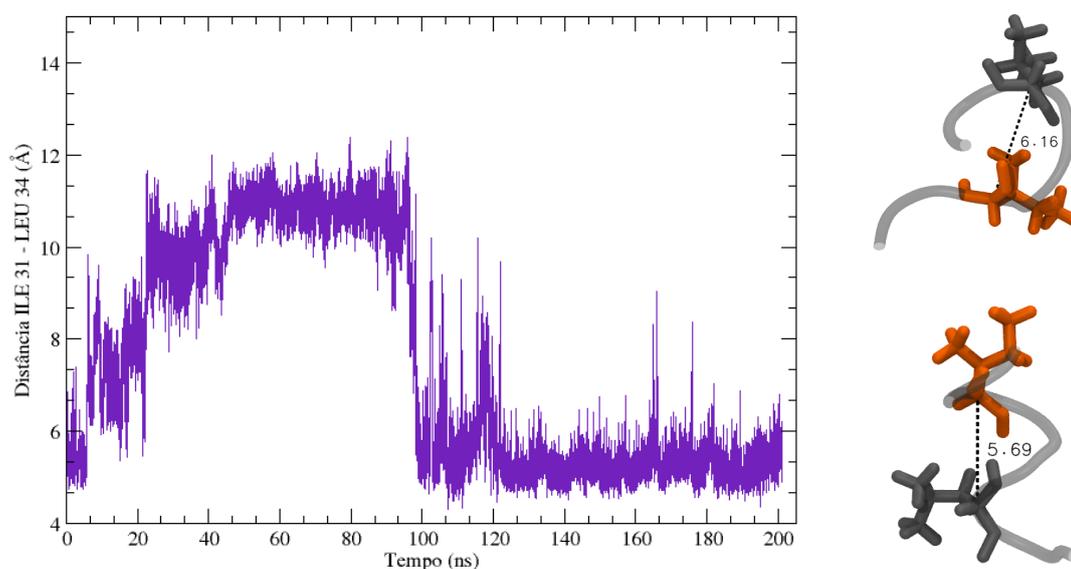


Figura 3.4: Medida de distância entre os aminoácidos Ile31(laranja) e Val34(cinza), a partir do centro massa de cada um deles, durante 200 ns de simulação de DM.

Na região Val24-Asn27 há um afastamento de seus aminoácidos mostrados na Figura 3.5. Esse aumento de distância impede a formação de uma ponte salina pelos aminoácidos aspartato(23) e lisina(28). Essa ponte está presente no peptídeo $A\beta_{1-40}$, que possui uma taxa de agregação muito menor que o $A\beta_{1-42}$ [31]. Assim, a ausência dessa estrutura pode ser um dos fatores que corroboram para a taxa de agregação do monômero $A\beta_{1-42}$ ser maior que a do $A\beta_{1-40}$, pois é menos uma interação para manter a conformação do monômero em α -hélice estável.

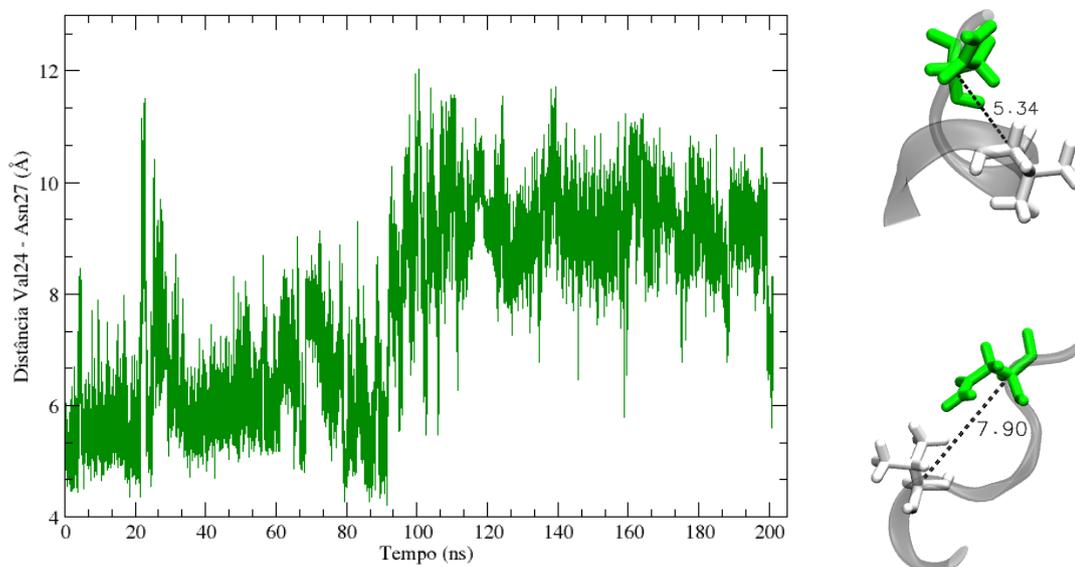


Figura 3.5: Medida de distância entre os aminoácidos Val24(branco) e Asn27(verde), a partir do centro massa de cada um deles, durante 200 ns de simulação de DM.

A seguir, os gráficos das Figuras 3.6 e 3.7 mostram, respectivamente, uma alta flutuação da região dos aminoácidos Lys28-Gly38, com RMSD em torno de 6Å , provocada pela mudança no seu formato, de α -hélice para folha- β , e uma baixa flutuação da região dos aminoácidos Ser8-Asp23, com RMSD em torno de $2,5\text{Å}$, demonstrando sua tendência a permanecer em α -hélice.

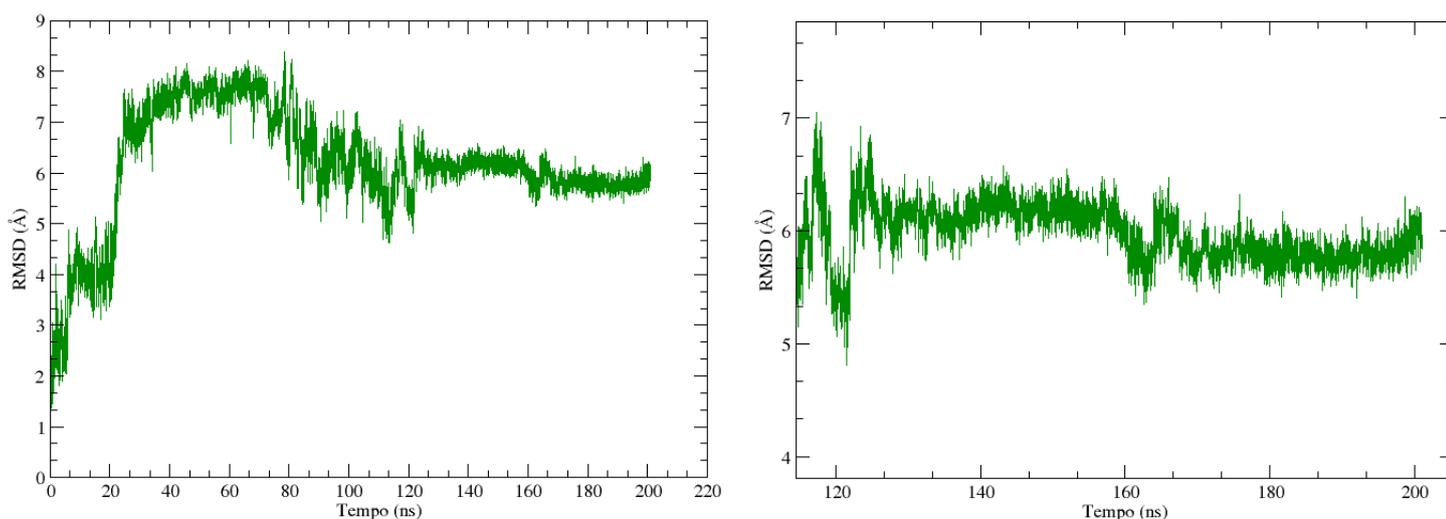


Figura 3.6: RMSD da região dos aminoácidos Lys28-Val39 do peptídeo $A\beta_{1-42}$ com conformação inicial de α -hélice durante 200ns de simulação de DM.

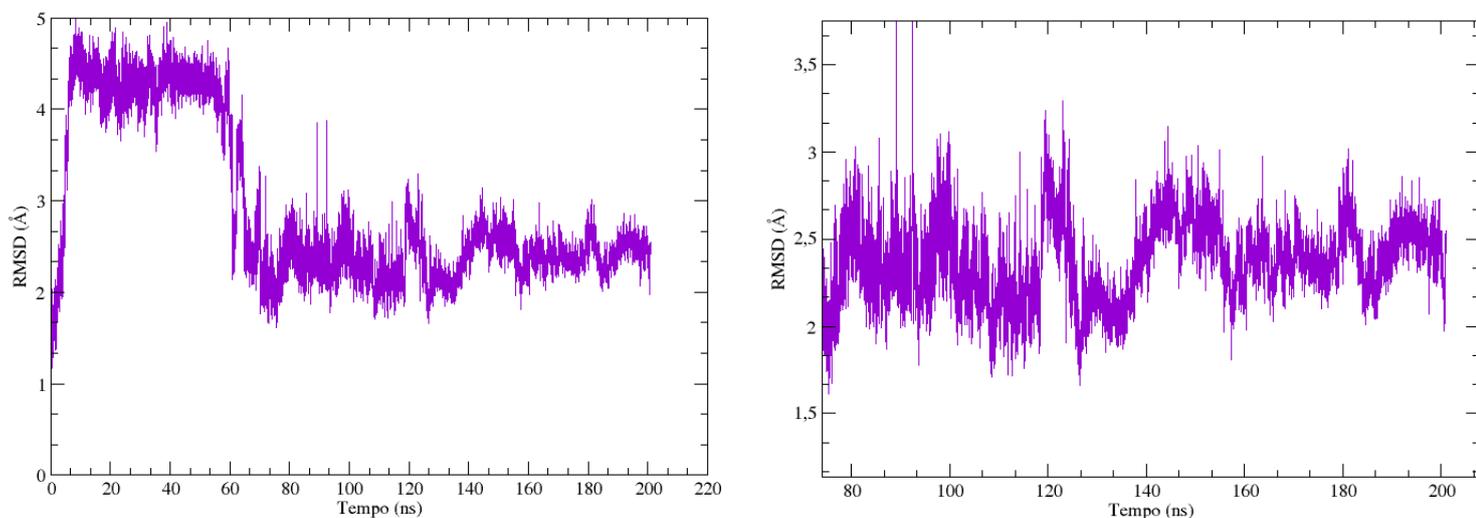


Figura 3.7: RMSD da região dos aminoácidos Ser8-Gly25 do peptídeo $A\beta_{1-42}$ com conformação inicial de α -hélice durante 200ns de simulação de DM.

Conformação Inicial de folha- β

Nesta parte da seção, serão feitas discussões sobre os resultados obtidos considerando o peptídeo como um todo e sua divisão nas regiões de folha- β - segmentos compreendidos pelos aminoácidos Asp1-Asp7, Gln15-Ala21 e Val39-Ala42.

Para o RMSD do peptídeo completo, apresentado no gráfico da Figura 3.8, o momento em que o alvo molecular mostra uma menor flutuação e um possível equilíbrio é observado a partir dos 85 ns até os 200 ns flutuando em torno do valor de 13,5 Å. Esse valor elevado mostra uma grande mobilidade do peptídeo, que também foi constatada na sua forma de α -hélice.

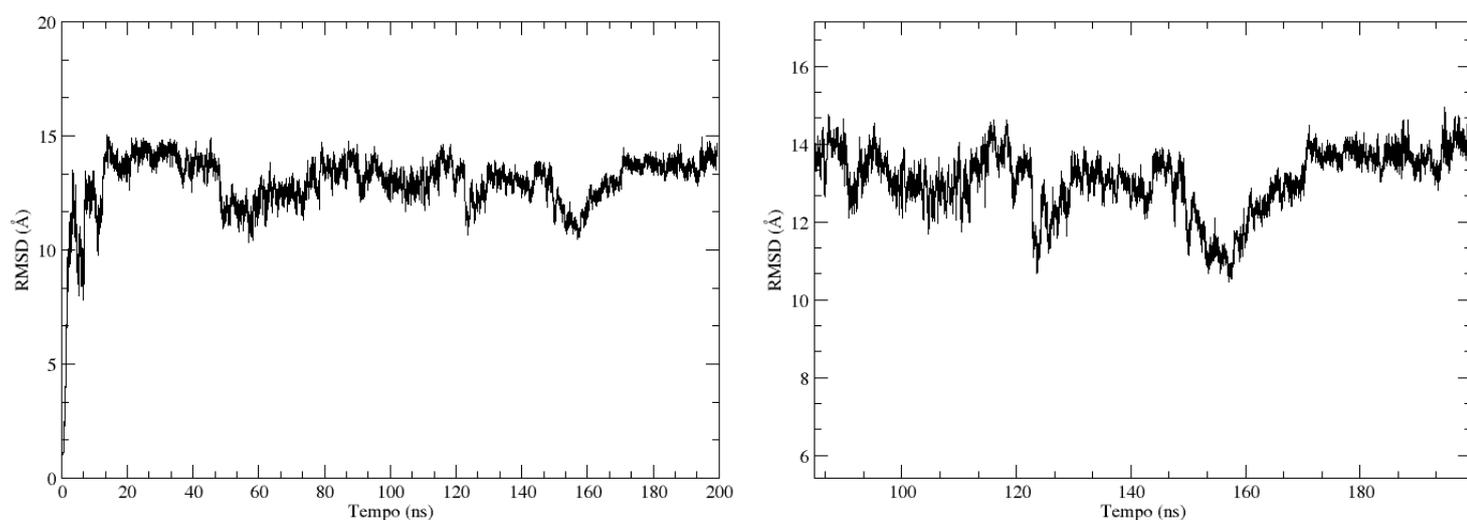


Figura 3.8: RMSD da extensão total do peptídeo $A\beta_{1-42}$ com conformação inicial de folha- β durante 200ns de simulação de DM. O segundo gráfico detalha o RMSD a partir do momento que o peptídeo flutua numa posição com menores variações de RMSD.

Nas regiões de folha- β , nota-se uma maior estabilidade, visto que os RMSDs encontrados são de valores bem menores. Para a região entre os aminoácidos Asp1 e Asp7, o valor calculado foi de 3,5 Å (Figura 3.9), para Gln15-Ala21 de 4,5 Å (Figura 3.10) e para Val39-Ala42 de 2,5 Å (Figura 3.10).

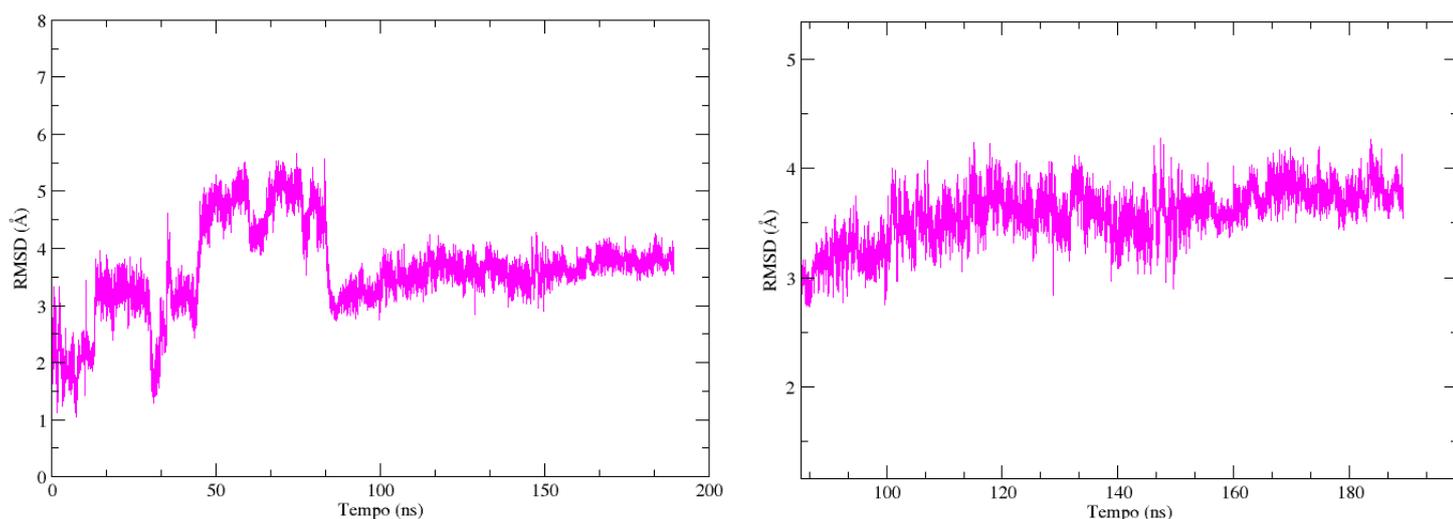


Figura 3.9: RMSD da região dos aminoácidos Asp1-Asp7 do peptídeo $A\beta_{1-42}$ com conformação inicial de folha- β durante 189 ns de simulação de DM. O segundo gráfico detalha o RMSD a partir do momento que a região analisada flutua na posição de equilíbrio

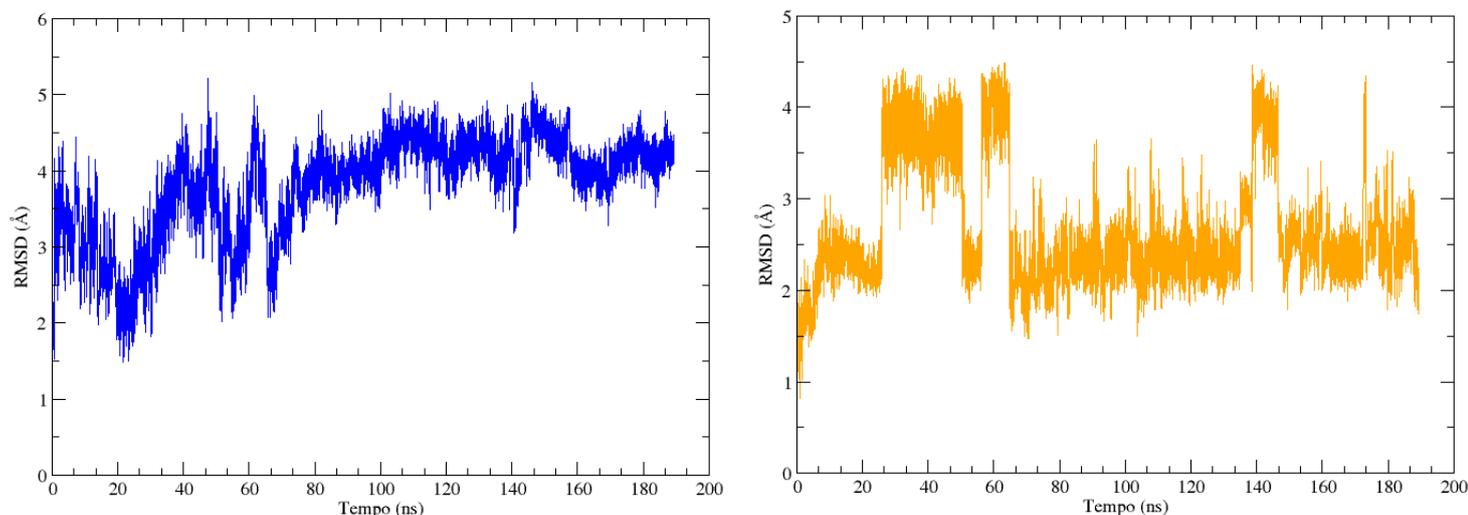


Figura 3.10: RMSD das regiões dos aminoácidos Gln15-Ala21 (à esquerda) e dos Val39-Ala42 (à direita) do peptídeo $A\beta_{1-42}$ com conformação inicial de folha- β durante 189 ns de simulação de DM.

Observando os baixos valores dos RMSDs das regiões de folha- β , pode-se dizer que não há grandes mudanças nessa estrutura secundária. Esse fato é confirmado pela Figura 3.11, que mostra as várias poses do peptídeo durante os 189 ns de simulação de DM que ele está completamente livre, sem rigidez. Nessa figura, as regiões que apresentam folha- β permanecem na mesma conformação. A mudança mais significativa observada é a aproximação das regiões Asp1-Asp7 e Gln15-Ala21 (Figura 3.12), que sugere uma afinidade desses aminoácidos e indicam dois segmentos com prováveis parcelas de responsabilidade pela agregação desse peptídeo. Uma possível explicação seria que quando essas regiões interagem devem conferir estabilidade aos oligômeros formados pelo $A\beta_{1-42}$. Essa aproximação durante a simulação de DM é mostrada na Figura 3.12.

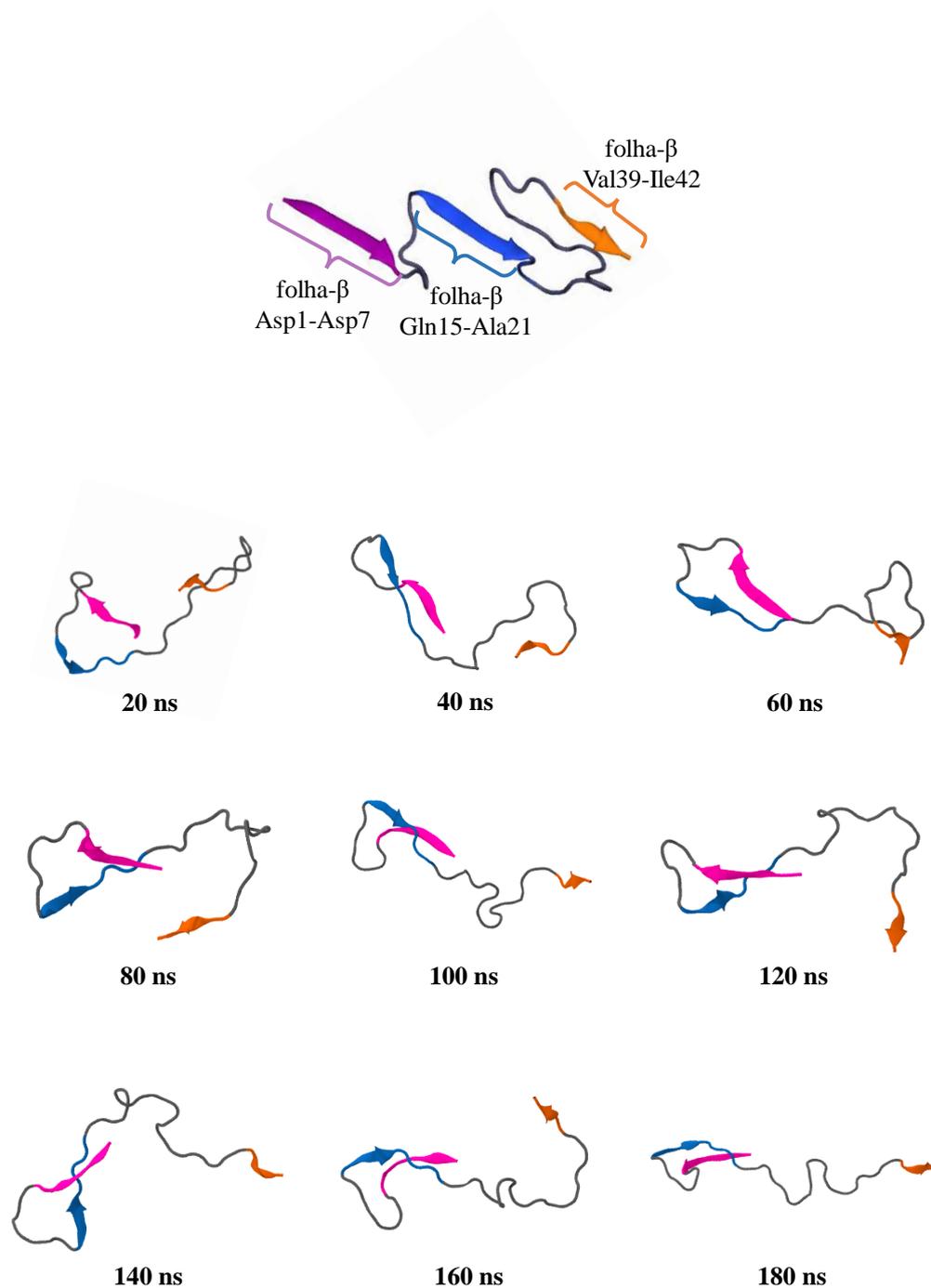


Figura 3.11: Mudanças sofridas pelo peptídeo $A\beta_{1-42}$ a cada 20 ns da simulação. A primeira figura mostra a conformação folha- β do cristal(PDB:2NAO) e suas principais estruturas. Na sequência estão dispostas representações das modificações do peptídeo ao longo dos 189 ns da simulação.

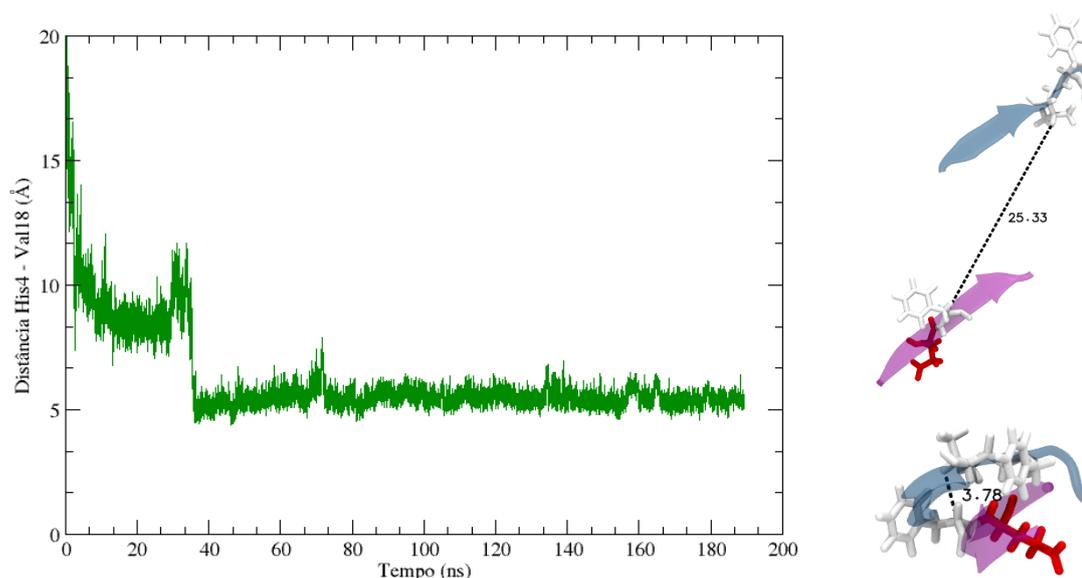


Figura 3.12: Medida de distância entre os aminoácidos His4(branco, folha- β magenta) e Val18(branco, folha- β azul), a partir do centro massa de cada um deles, durante 189 ns de simulação de DM.

3.2 Resultados de *Docking* Receptor-ligante

Os estudos de *docking* foram realizados usando as conformações do peptídeo encontradas na DM. Essas nuvens de soluções são uma clusterização das poses fornecidas pelo programa Autodock Vina e mostram a densidade dessas soluções distribuídas ao longo do peptídeo.

O Docking realizado para os compostos aqui estudados e suas nuvens de soluções obtidas para as conformações α -hélice e folha- β do peptídeo $A\beta_{1-42}$ serão apresentadas a seguir. Na forma de α -hélice, foram geradas 8600 soluções para os três ligantes. Na folha- β , foram 9400 soluções para os três compostos naturais.

As especificações de cores usadas nas figuras são as mesma para todos os ligantes. As cores vermelho e azul estão relacionadas à orientação dos átomos dos ligantes. Nesta pesquisa, essa questão não será analisada, mas na Figura 3.13 tem-se um exemplo dessas orientações para o cinamaldeído na α -hélice. No que tange a essas orientações dos átomos do ligante, observa-se um espectro de cores entre

o vermelho e o azul [102]. Conforme o ligante muda sua orientação entre as duas posições mostradas na Figura 3.13, tem-se um vermelho mais claro, passando por um azul claro até chegar no azul escuro da conformação apresentada na figura. A cor branca é a ausência do ligante naquela posição.

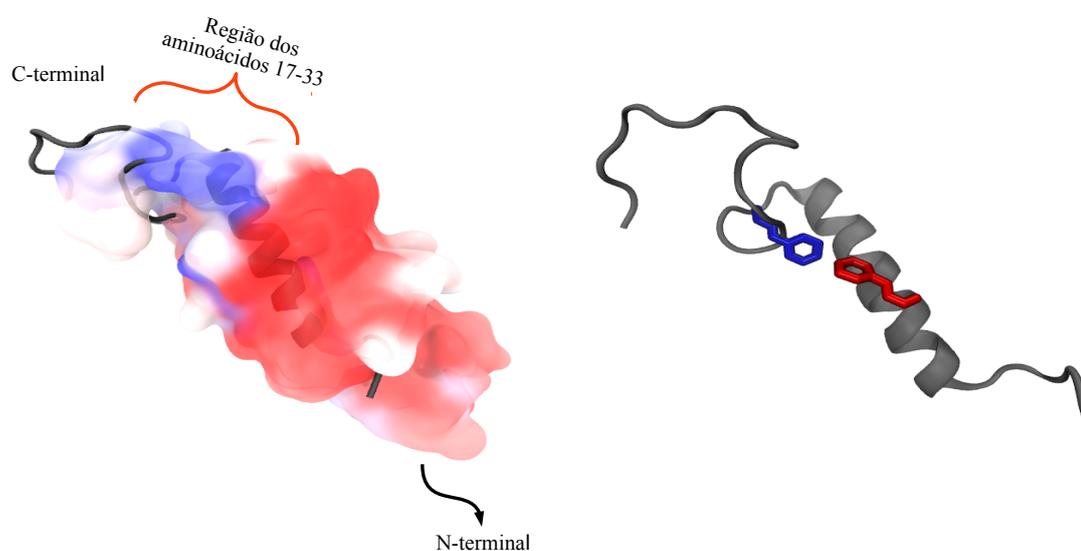


Figura 3.13: Nuvem de *docking* do cinamaldeído na conformação de α -hélice do peptídeo $A\beta_{1-42}$. A segunda figura mostra as orientações do ligante determinadas pelas cores azul e vermelha, mesmas cores das nuvens nas regiões que essas conformações são encontradas.

Na figura acima, nota-se uma preferência de conformação, dada pela cor azul, na região dos aminoácidos Leu17-Gly33 e na área próxima ao N-terminal há tendência para a outra conformação possível do cinamaldeído, dada pela cor vermelha.

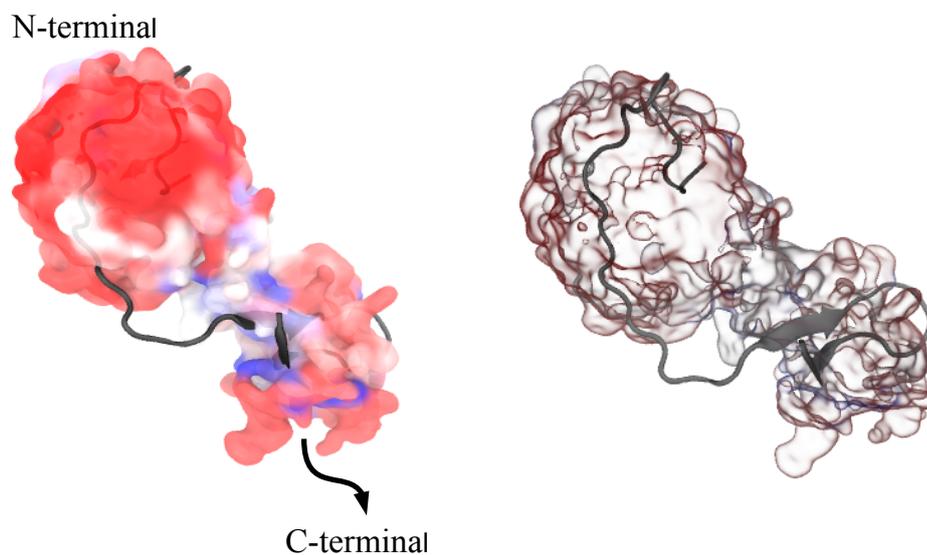


Figura 3.14: Nuvem de *docking* do cinamaldeído na conformação de folha- β do peptídeo A β_{1-42} . A segunda figura detalha a posição do peptídeo da imagem à esquerda.

Na nuvem representativa das soluções do aldeído cinâmico para a folha- β , observa-se uma preferência da conformação do peptídeo no N-terminal, que apresenta uma cor vermelha mais intensa do que nas outras regiões do peptídeo.

Para a curcumina, tem-se as nuvens de *docking* em ambas conformações do A β_{1-42} mostradas nas figuras 3.15 e 3.16. Na α -hélice, observa-se a mesma tendência de separação das conformações preferenciais vista no cinamaldeído. Na folha- β , não existe uma divisão tão clara, mostrando uma nuvem voltada para a cor vermelha distribuída de maneira mais homogênea por todo o peptídeo.

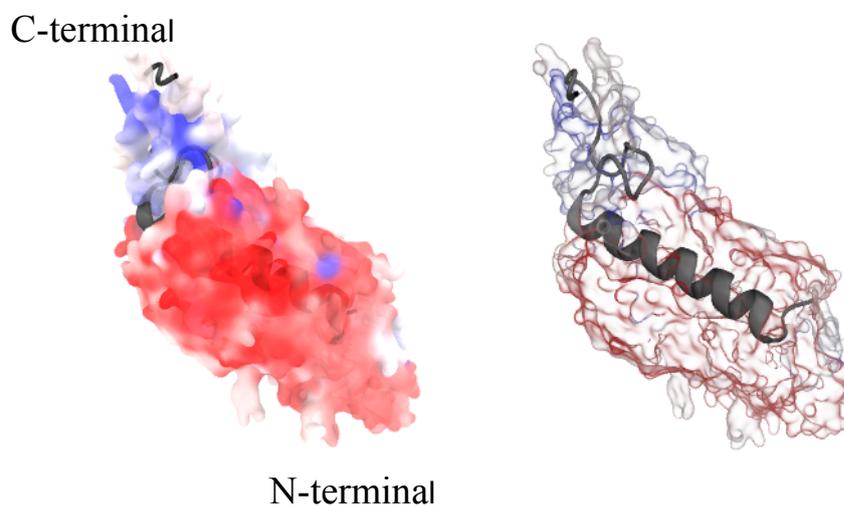


Figura 3.15: Nuvem de *docking* da curcumina na conformação de α -hélice do peptídeo $A\beta_{1-42}$. A segunda figura detalha a posição do peptídeo da imagem à esquerda.

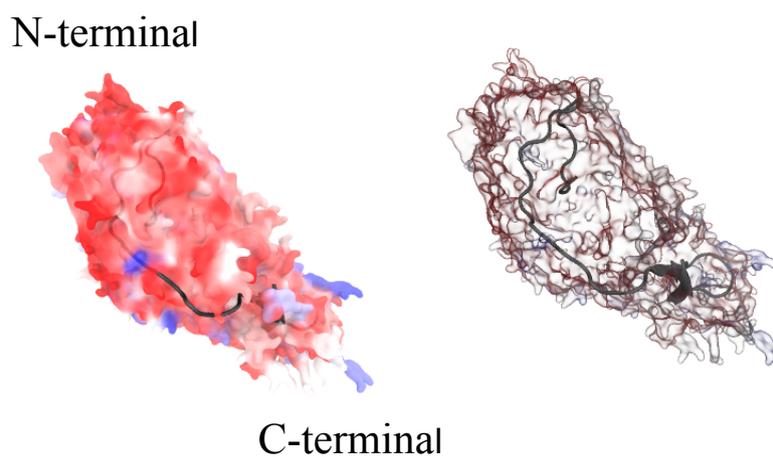


Figura 3.16: Nuvem de *docking* da curcumina na conformação de folha- β do peptídeo $A\beta_{1-42}$. A segunda figura detalha a posição do peptídeo da imagem à esquerda.

No âmbito do ácido rosmarínico, suas nuvens estão desenhadas nas figuras 3.17 e 3.18. Observa-se um comportamento bastante parecido com o da curcumina. No entanto, na conformação de α -hélice do peptídeo não existe uma divisão para a preferência de orientação do composto entre os aminoácidos dos dois terminais, há uma preferência clara por uma orientação do ácido rosmarínico pelos aminoácidos do N-terminal.

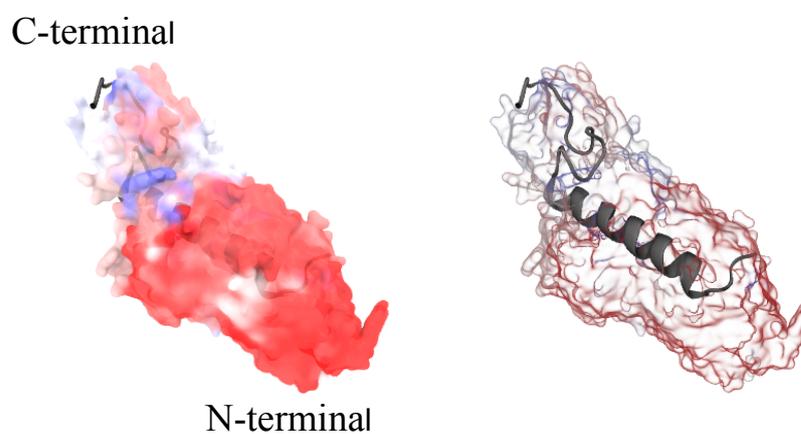


Figura 3.17: Nuvem de *docking* do ácido rosmarínico na conformação de α -hélice do peptídeo A β ₁₋₄₂. A segunda figura detalha a posição do peptídeo da imagem à esquerda.

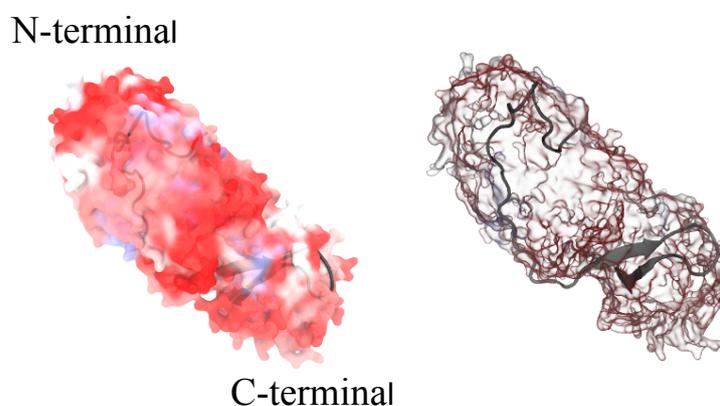


Figura 3.18: Nuvem de *docking* do ácido rosmarínico na conformação de folha- β do peptídeo $A\beta_{1-42}$. A segunda figura detalha a posição do peptídeo da imagem à esquerda.

3.3 Resultados das Energias do Complexo Receptor-Ligante via ONIOM

A fim de obter as energias dos complexos formados pelos compostos naturais mais o peptídeo $A\beta_{1-42}$ via método ONIOM2, neste trabalho, após a docagem dos compostos no peptídeo, foram escolhidas 50 poses de *docking* para cada ligante considerando as duas conformações do peptídeo, totalizando 300 poses.

Para encontrar as poses em que as duas estruturas estão mais próximas, o peptídeo foi dividido em dois domínios: um hidrofílico (aminoácidos Asp1-Lys28) e o outro hidrofóbico (aminoácidos Gly29-Ala42) [30]. Esses domínios são determinados com base nas características das cadeias laterais da maioria dos aminoácidos presentes em cada região. Assim, essas conformações foram escolhidas com base na distância do centro geométrico desses dois domínios do peptídeo até o centro geométrico do ligante por meio de um *script* em linguagem Python. A distância limite fixada foi de até 3Å.

A partir dessa triagem, os aminoácidos selecionados para ficarem na camada alta do *input* do ONIOM juntamente com o ligante foram escolhidos de acordo com o seu posicionamento dentro das poses adotadas. Desse modo, para todos os ligantes, foram selecionados 3 aminoácidos, além do ligante, para constituírem a camada alta.

Tendo esse cenário como ponto de partida, os *inputs* para realizar os cálculos de energia eletrônica via método ONIOM2 foram construídos com o auxílio do pacote *Toolkit to Assist ONIOM Calculations* (TAO) [103] e de *scripts* para automatizar os comandos disponibilizados por esse pacote (alguns dos *scripts* usados estão disponíveis no Apêndice C). Os detalhes desse procedimento para cada ligante serão apresentados nas seções oportunas.

Assim sendo, para encontrar as energias eletrônicas das conformações com maior precisão, fez-se duas otimizações usando o programa Gaussian 09 [104]: uma com o método PM6 e outra com o nível de cálculo B3LYP e a base de Pople 6-311G(d,p), ambos para o tratamento da camada alta. O PM6 foi usado como uma "pré-otimização". Com as estruturas otimizadas, foi feito um cálculo de *single-point* usando o nível de cálculo M06-L e a base de Pople 6-311+G(d,p) na camada alta. Esse nível de cálculo foi escolhido para contemplar possíveis contribuições dispersivas nas interações envolvendo o peptídeo e os ligantes. Tanto nas otimizações como no *single-point*, a camada baixa foi tratada pelo campo de força AMBER.

Após o cálculo da energia do complexo, fez-se um *single-point* apenas para o peptídeo, preservando as divisões das camadas alta e baixa, e outro apenas para o ligante, com o intuito de observar a contribuição energética das estruturas separadas e ficar apenas com a energia eletrônica envolvida nas interações (Equação 3.1) [105].

$$\Delta E = E_{complexo} - E_{peptideo} - E_{ligante}. \quad (3.1)$$

Com o ΔE determinado, observou-se as poses de maior e menor energia dos complexos formados pelo peptídeo mais cada um dos ligantes, com o intuito de mapear as prováveis interações presentes que contribuem para uma maior estabi-

lidade dessas conformações. A compreensão dos fatores estruturais que permitem um alcance de mínimos de energia é imprescindível para o planejamento de fármacos. Assim sendo, serão apresentados os resultados alcançados para cada composto estudado em seções separadas.

3.3.1 Energias Eletrônicas das Interações do Cinamaldeído com o Peptídeo $A_{\beta_{1-42}}$

Após todo o processo de otimização, nem todas as conformações convergiram. Por conseguinte, para o composto cinamaldeído, das 25 conformações iniciais de cada domínio (hidrofílico e hidrofóbico) obteve-se 18 estruturas otimizadas para cada um dos dois domínios do peptídeo na forma de α -hélice. Na forma de folha- β , foram 16 para o domínio hidrofílico e 8 para o hidrofóbico.

A camada alta, na conformação de α -hélice do peptídeo, foi formada pelo aldeído cinâmico mais os aminoácidos His13, His14 e Leu17 - domínio hidrofílico. Para o domínio hidrofóbico, os aminoácidos selecionados foram Ile32, Gly33 e Met35. Na folha- β , os aminoácidos da camada alta foram Phe4, His6, Phe19 - domínio hidrofílico -, Ile31, Leu34 e Met35 - domínio hidrofóbico.

Os resultados obtidos para cada uma das conformações do peptídeo são apresentados a seguir em seções separadas.

$A_{\beta_{1-42}}$ em α -hélice

A Figura 3.19 apresenta a distribuição da diferença de energia para os dois domínios do peptídeo em α -hélice. Nota-se que as menores energias, em sua maioria, estão localizadas no domínio hidrofóbico.

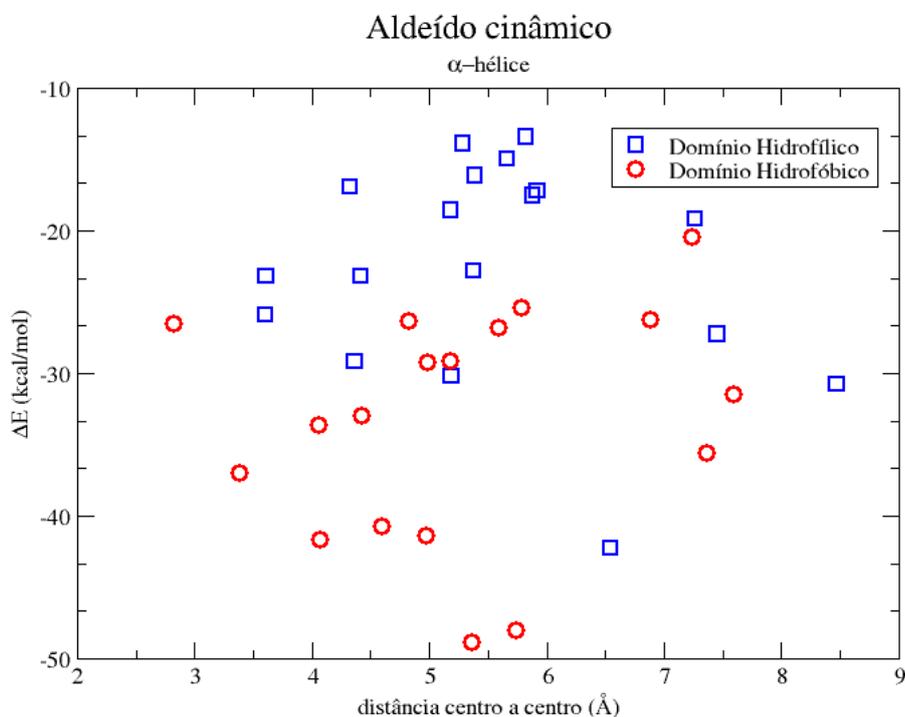


Figura 3.19: Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo $A\beta_{1-42}$ mais o cinamaldeído, levando em consideração a divisão do peptídeo nos domínios hidrofílico e hidrofóbico. A distância apresentada é do centro geométrico do ligante para o centro geométrico do grupo de aminoácidos escolhidos para a camada alta.

As Figuras 3.20 e 3.21 mostram, para cada domínio considerado, as conformações de menor e de maior energia. Por meio dessas figuras, é possível observar que os tipos de interações mais importantes, em termos energéticos, são:

- a) Interação por ponte de hidrogênio entre o grupo amina da cadeia lateral da Lisina 16 (LYS 16) e a carbonila do ligante;
- b) Interação hidrofóbica da cadeia carbônica da Lys16 com a cadeia do ligante.

Além dessas, existem interações hidrofóbicas que possuem uma contribuição energética menor, mas também levam a um mínimo de energia, que são o π -stacking da Fenilalanina 20 (PHE 20) com o anel aromático do cinamaldeído e a interação do anel da Histidina 13 (HIS 13) com a cadeia carbônica do ligante. Em outras con-

formações, é bastante provável que o π -stacking ocorra com a Fenilalanina 19 (PHE 19). No domínio hidrofóbico, a interação mais relevante é a ponte de hidrogênio do grupo amina da cadeia lateral da Asparagina 27 (ASN 27) com a carbonila do aldeído cinâmico. O restante das interações nesse domínio são puramente hidrofóbicas, a saber: a cadeia lateral da Ile31 com o anel do cinamaldeído e a cadeia lateral da Met35 com a cadeia carbônica do ligante.

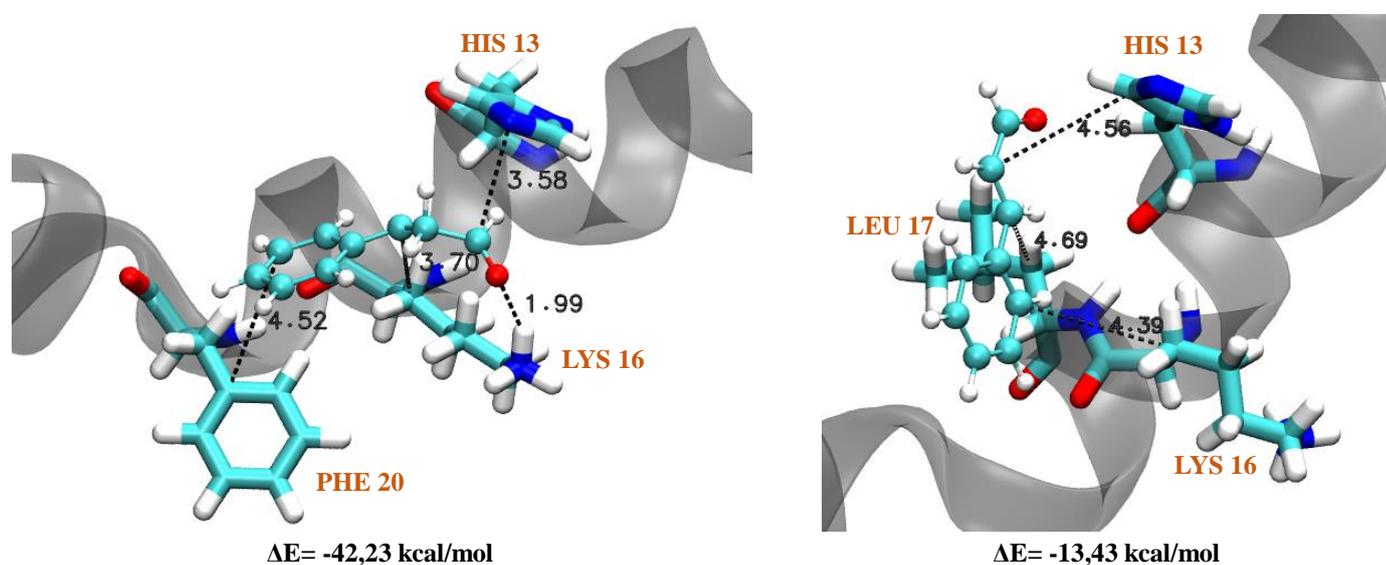


Figura 3.20: Interações do aldeído cinâmico (representação de esfera e varetas) com os aminoácidos (representação em varetas) do $A\beta_{1-42}$ no domínio hidrofílico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são π -stacking com a Phe20 (4,52Å), uma ligação de hidrogênio com a Lys16 (1,99Å) e duas interações hidrofóbicas, uma com o anel da His13 (3,58Å) e outra com a cadeia carbônica da Lys16 (3,70Å). Na figura da direita (maior energia), observa-se apenas interações hidrofóbicas com os aminoácidos mostrados: cadeia carbônica da Lys16 mais o anel do ligante (4,39Å), cadeias carbônicas do ligante mais da Leu17 (4,69Å) e cadeia carbônica do ligante com o anel da His13 (4,56Å).

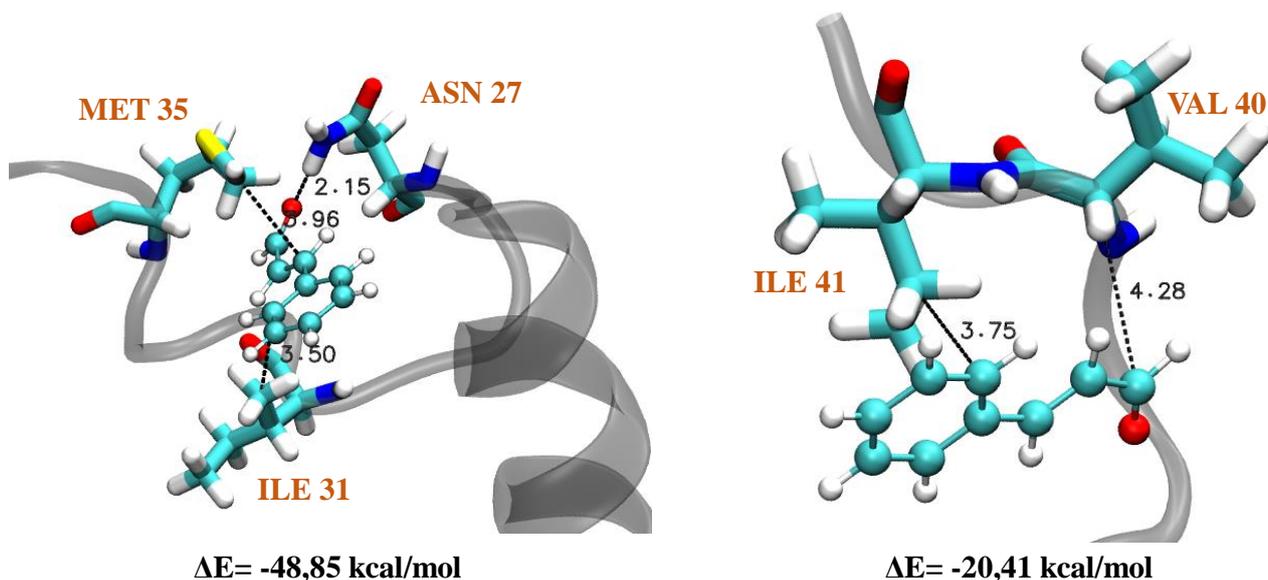


Figura 3.21: Interações do aldeído cinâmico (representação de esfera e varetas) com os aminoácidos (representação em varetas) do Aβ₁₋₄₂ no domínio hidrofóbico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são uma ligação de hidrogênio com a Asn27 (2,15Å) e duas interações hidrofóbicas, uma do anel do ligante com a cadeia lateral da Ile31 (3,50Å) e outra da cadeia carbônica do ligante com a cadeia lateral da Met35 (3,96Å). Na figura da direita (maior energia), observa-se apenas interações hidrofóbicas com os aminoácidos mostrados: cadeia carbônica do ligante com a cadeia lateral da Val40 (4,28Å) e anel do ligante mais a cadeia lateral da Ile41 (3,75Å).

Partindo das interações apresentadas nas Figuras 3.20, 3.21 e da Figura 3.22, que ilustra os orbitais de fronteira do cinamaldeído (HOMO - *Highest Occupied Molecular Orbital*, e LUMO - *Lowest Unoccupied Molecular Orbital*) nas configurações de menor energia e também do ligante isolado, sem estar interagindo com o peptídeo, é possível ver a alteração das nuvens eletrônicas de acordo com as interações do ligante com os aminoácidos ao seu redor. As principais alterações são observadas nos orbitais HOMO, que sofrem um deslocamento na região do ligante e possuem uma maior superfície de densidade quando há interações do peptídeo com o ligante. Além disso, quanto menor a energia, maior a superfície de densidade dos orbitais

HOMO na região da carbonila do aldeído cinâmico.

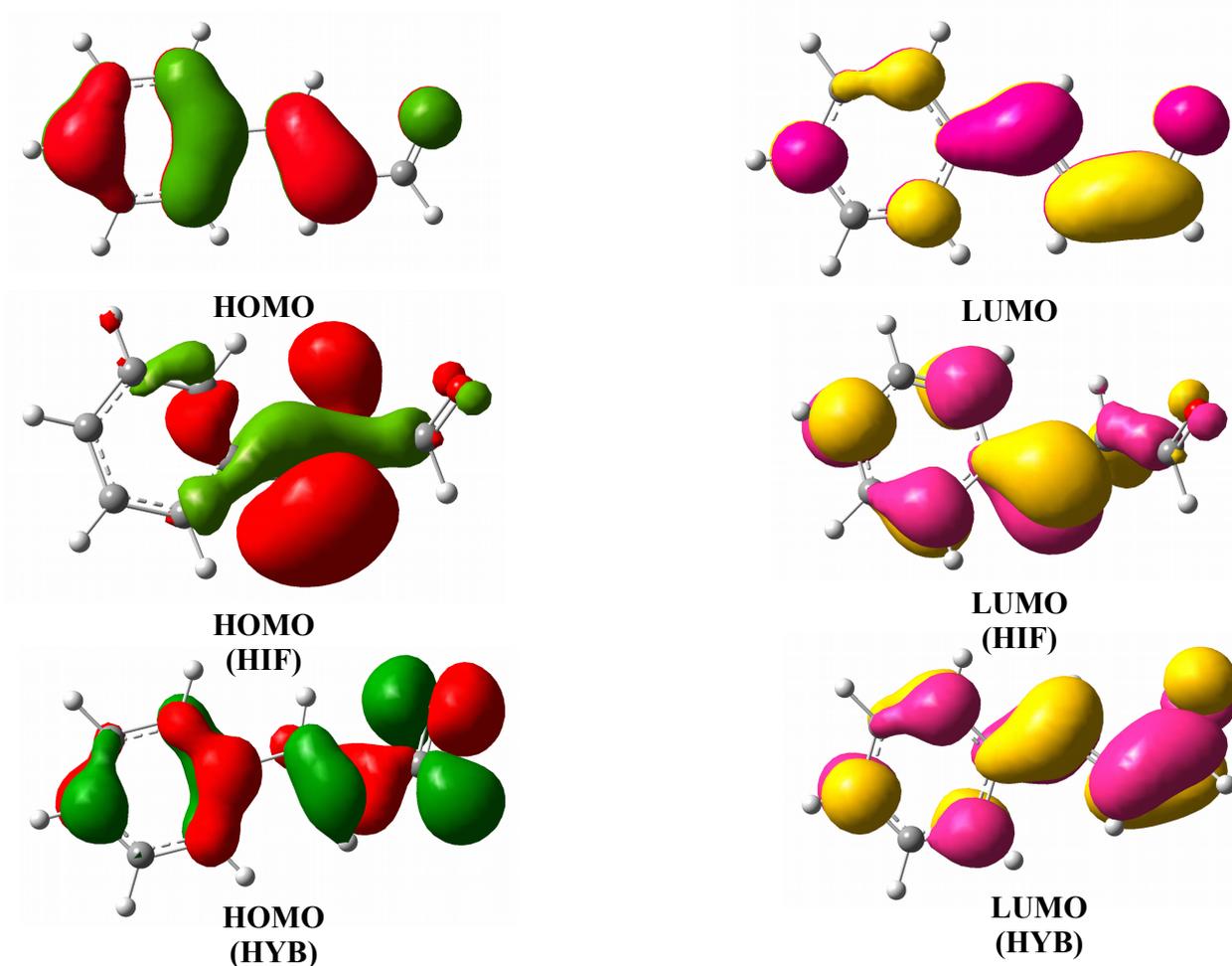


Figura 3.22: Representação dos orbitais HOMO e LUMO para o cinamaldeído isolado e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de α -hélice. De cima para baixo, a representação do HOMO e LUMO do ligante isolado (sem interação), depois a representação do HOMO e LUMO da interação do cinamaldeído no domínio hidrofílico (HIF) do peptídeo - configuração de menor energia - e, por fim, a representação do HOMO e LUMO da interação do cinamaldeído no domínio hidrofóbico (HYB) - configuração de menor energia

$A\beta_{1-42}$ em Folha- β

Para a conformação em folha- β , a Figura 3.23 mostra a distribuição da diferença de energia para os dois domínios do peptídeo. Nessa Figura, é possível ob-

servar que, diferentemente dos resultados para a conformação α -hélice do peptídeo, os dois domínios possuem energias bem similares.

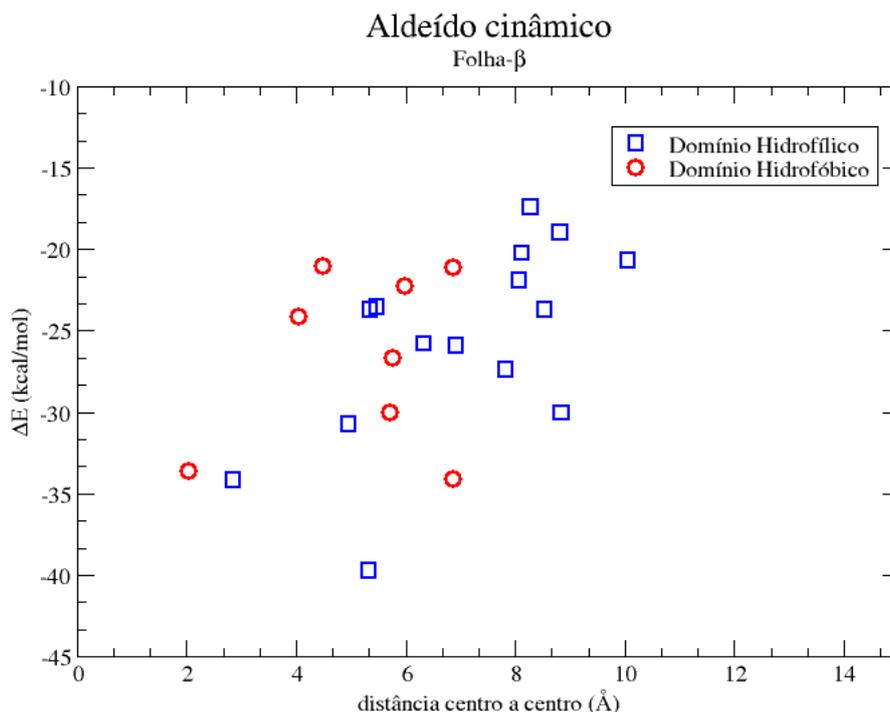


Figura 3.23: Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo $A\beta_{1-42}$ mais o cinamaldeído, levando em consideração a divisão do peptídeo nos domínios hidrofílico e hidrofóbico. A distância apresentada é do centro geométrico do ligante para o centro geométrico do grupo de aminoácidos escolhidos para a camada alta.

Com base na Figura 3.23 e analisando as conformações de menor e maior energia, dentro de cada um dos domínios (Figuras 3.24 e 3.25), é possível observar as interações mais relevantes energeticamente. De fato, no domínio hidrofílico, tem-se a ponte de hidrogênio entre o grupo amina do *backbone* da Lisina 16 (LYS 16) e a carbonila do ligante, seguida das interações próprias de anéis aromáticos, um empilhamento do anel do ligante com a Fenilalanina 19 (PHE 19) fazendo um π -stacking ou um T-stacking com a Fenilalanina 4 (PHE 4). Além dessas, há a interação hidrofóbica da cadeia carbônica do ligante com o anel da Histidina 13 (HIS 13). No

domínio hidrofóbico, a interação mais substancial observada é uma ligação de hidrogênio do grupo amina do *backbone* da Glicina 29 (GLY 29) com a carbonila do cinamaldeído e a cadeia lateral desse mesmo aminoácido com a cadeia carbônica do ligante. O restante das interações prováveis nesse domínio são hidrofóbicas: cadeias laterais da Gly33 e Leu34 com o anel do cinamaldeído.

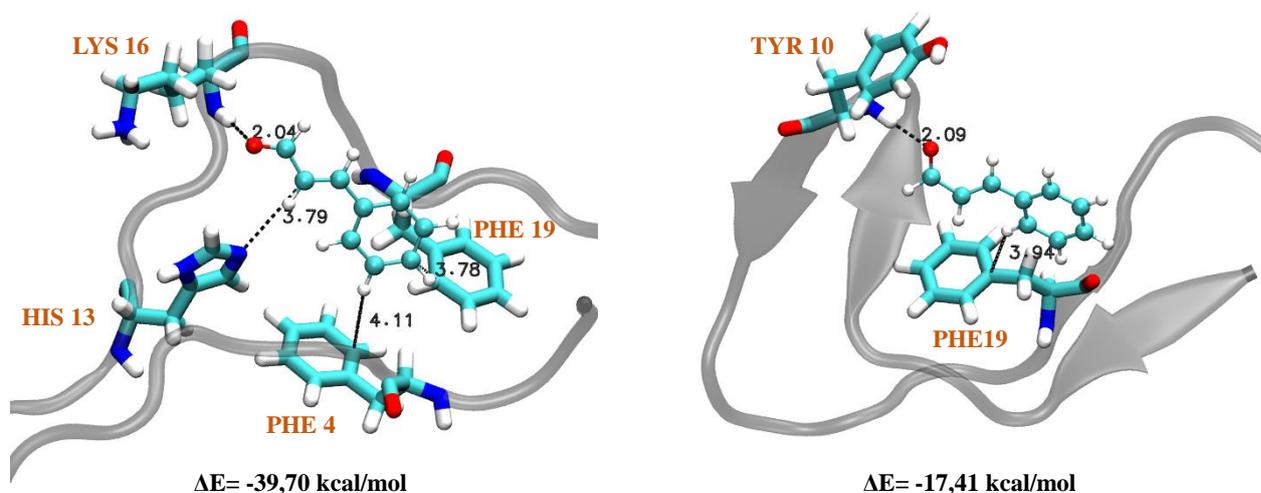


Figura 3.24: Interações do cinamaldeído (representação de esfera e varetas) com os aminoácidos (representação em varetas) do $A\beta_{1-42}$ no domínio hidrofílico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são π -stacking com a Phe19 (3,78Å), T-stacking com a Phe4 (4,11Å), uma ligação de hidrogênio com a amina do *backbone* da Lys16 (2,04Å) e a interação do anel da His13 (3,79Å) com a cadeia do ligante. Na figura da direita (maior energia), observa-se: um T-stacking com a Phe19(3,94Å) e uma ponte de hidrogênio da carbonila do ligante com a amina do backbone da Tyr10(2,09Å).

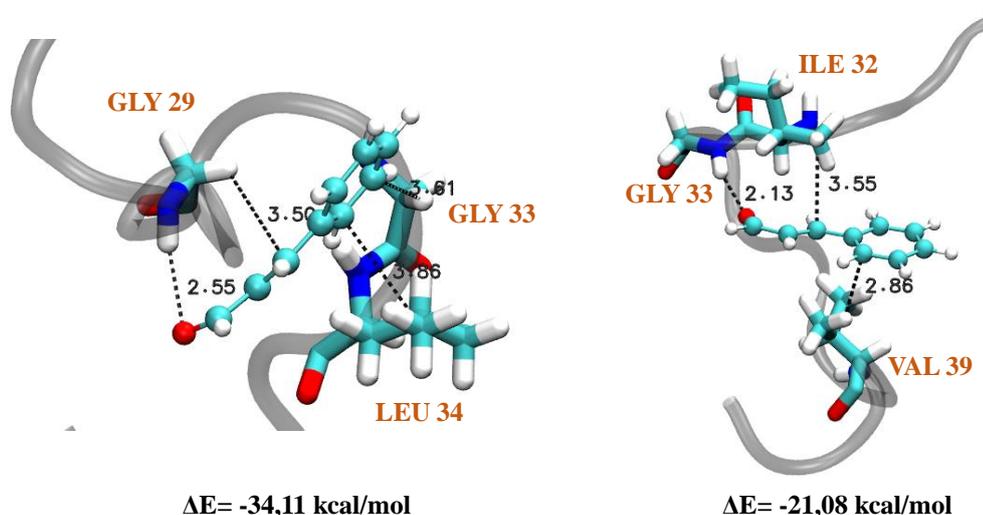


Figura 3.25: Interações do aldeído cinâmico (representação de esfera e varetas) com os aminoácidos (representação em varetas) do $A\beta_{1-42}$ no domínio hidrofóbico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são uma ligação de hidrogênio da amina do *backbone* da Gly29 com a carbonila do composto (2,55Å) e três interações hidrofóbicas, duas interações possíveis com o anel do ligante, uma com a cadeia lateral da Gly33 (3,61Å) e outra com a cadeia lateral da Leu34(3,86Å), e a cadeia lateral da Gly29 com a cadeia do composto (3,50Å). Na figura da direita (maior energia), observa-se as interações: cadeia carbônica do ligante com a cadeia lateral da Ile32 (3,55Å), cadeia lateral da Val39 e anel do ligante (2,86Å) e uma ligação de hidrogênio da amina do *backbone* da Gly33 com a carbonila do cinamaldeído (2,13Å).

Haja vista as interações apresentadas nas Figuras 3.24 e 3.25 e da Figura 3.26, que ilustra os orbitais de fronteira do cinamaldeído (HOMO e LUMO) nas configurações de menor energia e também do ligante isolado, sem estar interagindo com o peptídeo, é possível ver a alteração das nuvens eletrônicas de acordo com as interações do ligante com os aminoácidos ao seu redor. As mesmas alterações vistas nas interações com o peptídeo em α -hélice são observadas para as interações com o peptídeo em folha- β .

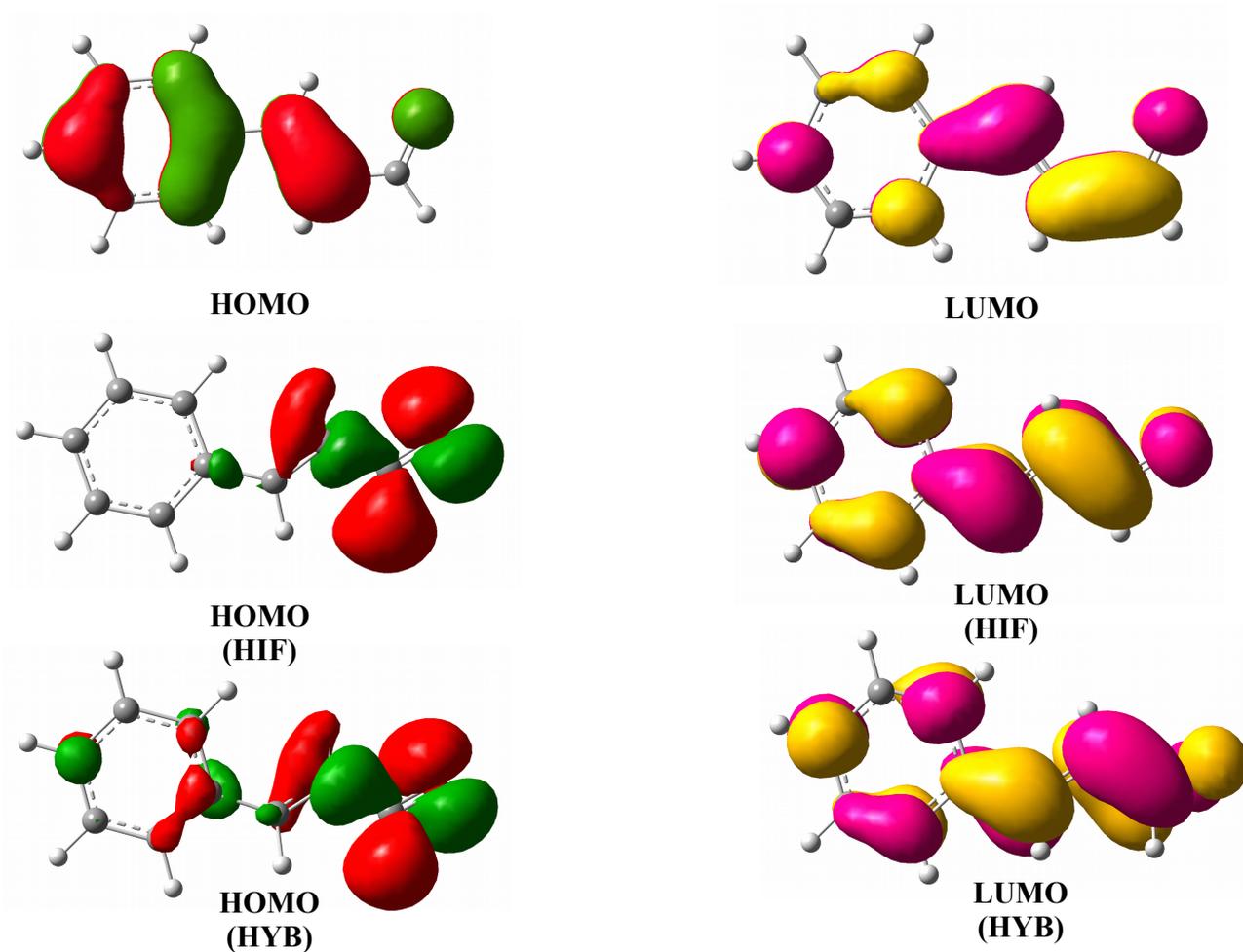


Figura 3.26: Representação dos orbitais HOMO e LUMO para o cinamaldeído isolado e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de folha- β . De cima para baixo, a representação do HOMO e LUMO do ligante isolado (sem interação), depois a representação do HOMO e LUMO da interação do cinamaldeído no domínio hidrofílico (HIF) do peptídeo - configuração de menor energia - e, por fim, a representação do HOMO e LUMO da interação do cinamaldeído no domínio hidrofóbico (HYB) - configuração de menor energia

3.3.2 Energias Eletrônicas das Interações da Curcumina com o Peptídeo

$A\beta_{1-42}$

Para a curcumina das 25 conformações iniciais de cada domínio obteve-se 6 estruturas otimizadas para o domínio hidrofílico e 13 para o hidrofóbico, na forma de

α -hélice do peptídeo. Na folha- β , 7 para o domínio hidrofílico e 5 para o hidrofóbico.

A camada alta, na conformação de α -hélice do peptídeo, foi formada pela curcumina mais os aminoácidos Tyr10, His13 e His14 - domínio hidrofílico. Para o domínio hidrofóbico, os aminoácidos selecionados foram Ala30, Met35 e Val39. Na folha-*beta*, os aminoácidos da camada alta foram Phe4, His6, Phe19 - domínio hidrofílico -, Ile31, Leu34 e Met35 - domínio hidrofóbico.

$A_{\beta_{1-42}}$ em α -hélice

A Figura 3.27 mostra a distribuição da diferença de energia para os dois domínios do peptídeo em α -hélice. Nota-se que, assim como no cinamaldeído, as menores energias estão localizadas no domínio hidrofóbico.

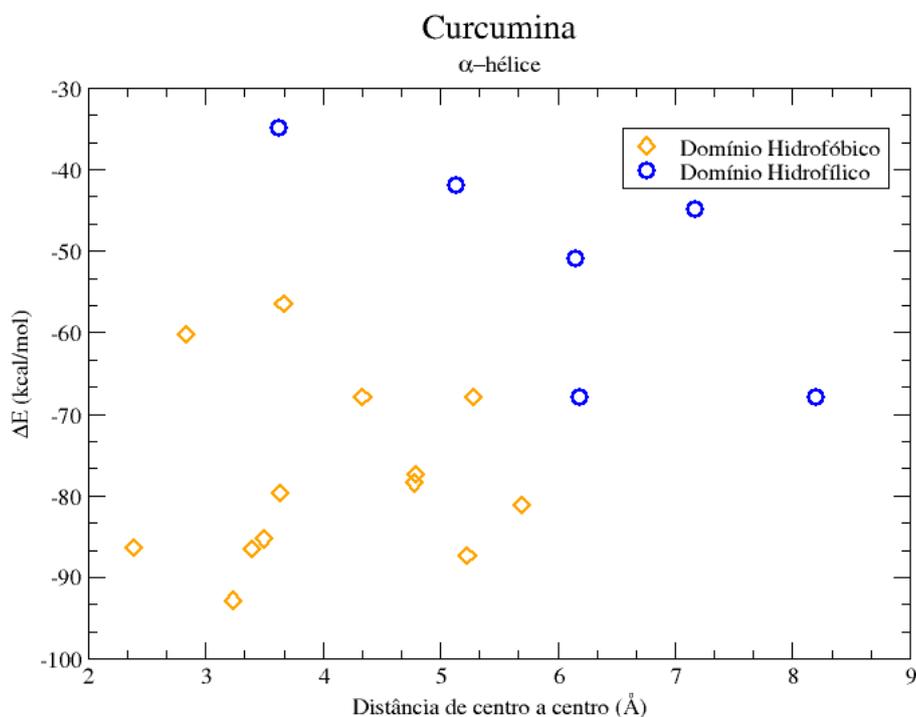


Figura 3.27: Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo $A\beta_{1-42}$ mais a curcumina, levando em consideração a divisão do peptídeo nos domínios hidrofílico e hidrofóbico. A distância apresentada é do centro geométrico do ligante para o centro geométrico do grupo de aminoácidos escolhidos para a camada alta.

Tendo a Figura 3.27 como ponto de partida e, também, as conformações de menor e maior energia dentro de cada um dos domínios (Figuras 3.28 e 3.29) é possível observar os tipos de interações mais relevantes. No domínio hidrofílico, são uma possível ligação de hidrogênio com a cadeia lateral da Glutamina 15 (GLN 15) com a carbonila do ligante e outra ligação de hidrogênio com a carbonila do Glutamato 22 (GLU 22) mais uma das hidroxilas do anel da curcumina. A interação hidrofóbica observada foi com a cadeia lateral da Valina 18 mais um dos anéis aromáticos do ligante. Para o domínio hidrofóbico, a interação mais relevante é a ponte de hidrogênio do grupo amina da cadeia lateral da Asparagina 27 (ASN 27) com o oxigênio entre o anel aromático e o grupo metila da curcumina. As interações

hidrofóbicas nesse domínio foram com as cadeias laterais dos aminoácidos Ile32 e Ile41 mais os anéis aromáticos da curcumina.

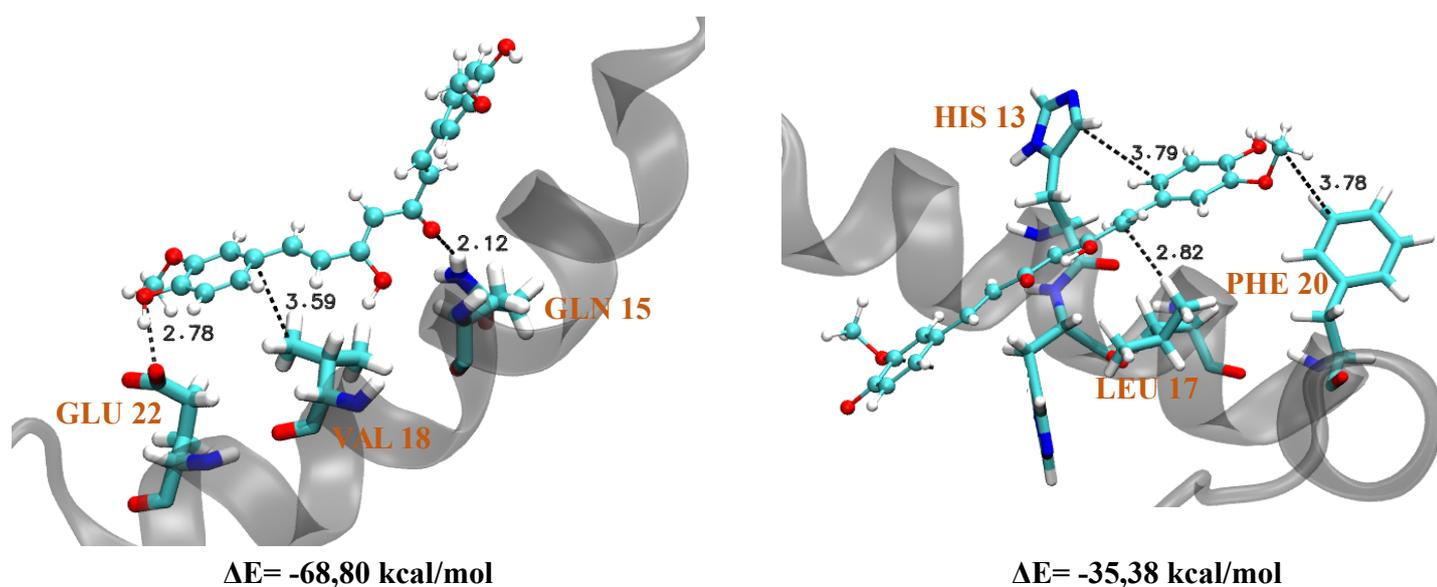


Figura 3.28: Interações da curcumina (representação de esfera e varetas) com os aminoácidos (representação em varetas) do $A\beta_{1-42}$ no domínio hidrofílico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são uma ligação de hidrogênio com a Gln15 (2,78Å) e outra com o Glu22(2,12Å), interação hidrofóbica com a Val18(3,59Å). Na figura da direita (maior energia), observa-se apenas interações hidrofóbicas com os aminoácidos mostrados: um π -stacking com a His13(3,79Å), uma do anel da Phe20 com o grupo metila do ligante (3,78Å) e uma da cadeia lateral da Leu17 com a parte central da curcumina (2,82Å).

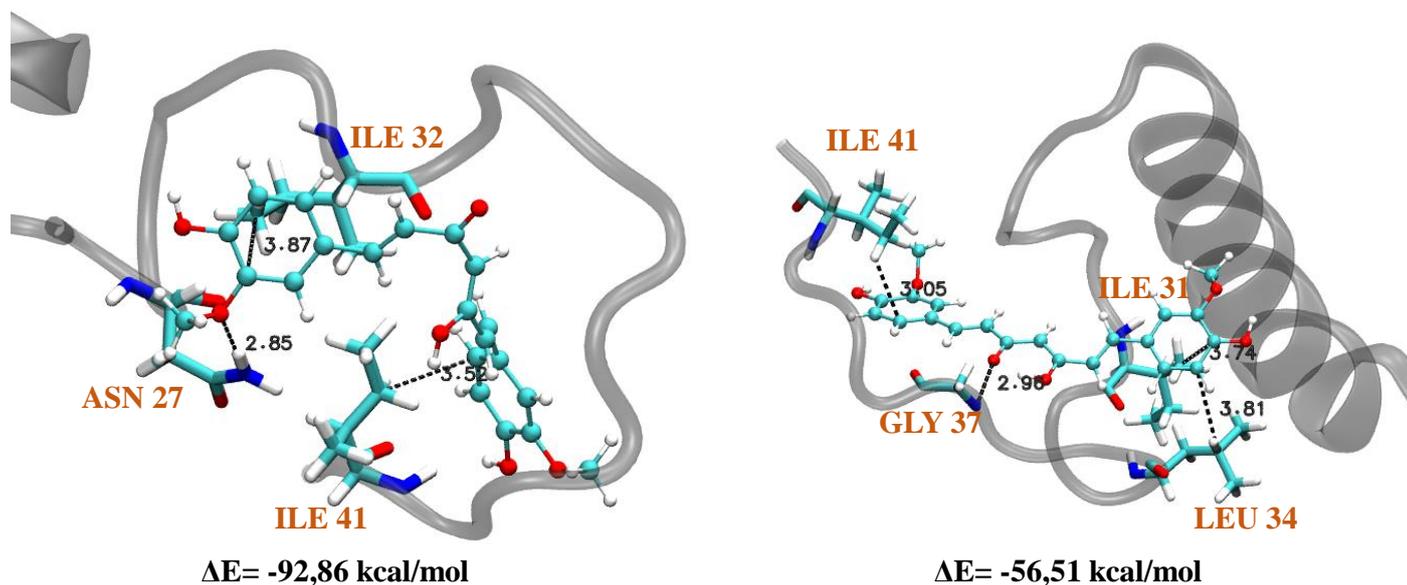


Figura 3.29: Interações da curcumina (representação de esfera e varetas) com os aminoácidos (representação em varetas) do $A\beta_{1-42}$ no domínio hidrofóbico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são uma ligação de hidrogênio com a Asn27 (2,85Å) e duas interações hidrofóbicas, uma do anel do ligante com a cadeia lateral da Ile32 (3,87Å) e outra da cadeia lateral da Ile41 com um dos anéis da curcumina (3,52Å). Na figura da direita (maior energia), observa-se as seguintes interações: cadeia lateral da Ile31 com um dos anéis (3,74Å), cadeia lateral da Leu34 no mesmo anel da Ile31 (3,81Å), ligação de hidrogênio da amina da Gly37 com a carbonila da curcumina (2,96Å) e a cadeia lateral da Ile41 com um dos anéis aromáticos do composto (3,05Å).

Partindo das interações apresentadas nas Figuras 3.28 e 3.29 e da figura 3.30, que ilustra os orbitais de fronteira da curcumina nas configurações de menor energia e também do ligante isolado, sem estar interagindo com o peptídeo, é possível ver a alteração das nuvens eletrônicas de acordo com as interações do ligante com os aminoácidos ao seu redor. As principais alterações são observadas nos orbitais HOMO, que apesar de continuar mais concentrado em uma região da curcumina, ele também se desloca para o outro anel do ligante, apresentando pequenas porções

da sua superfície, e também para a região central, onde estão os grupos hidroxila e carbonila.

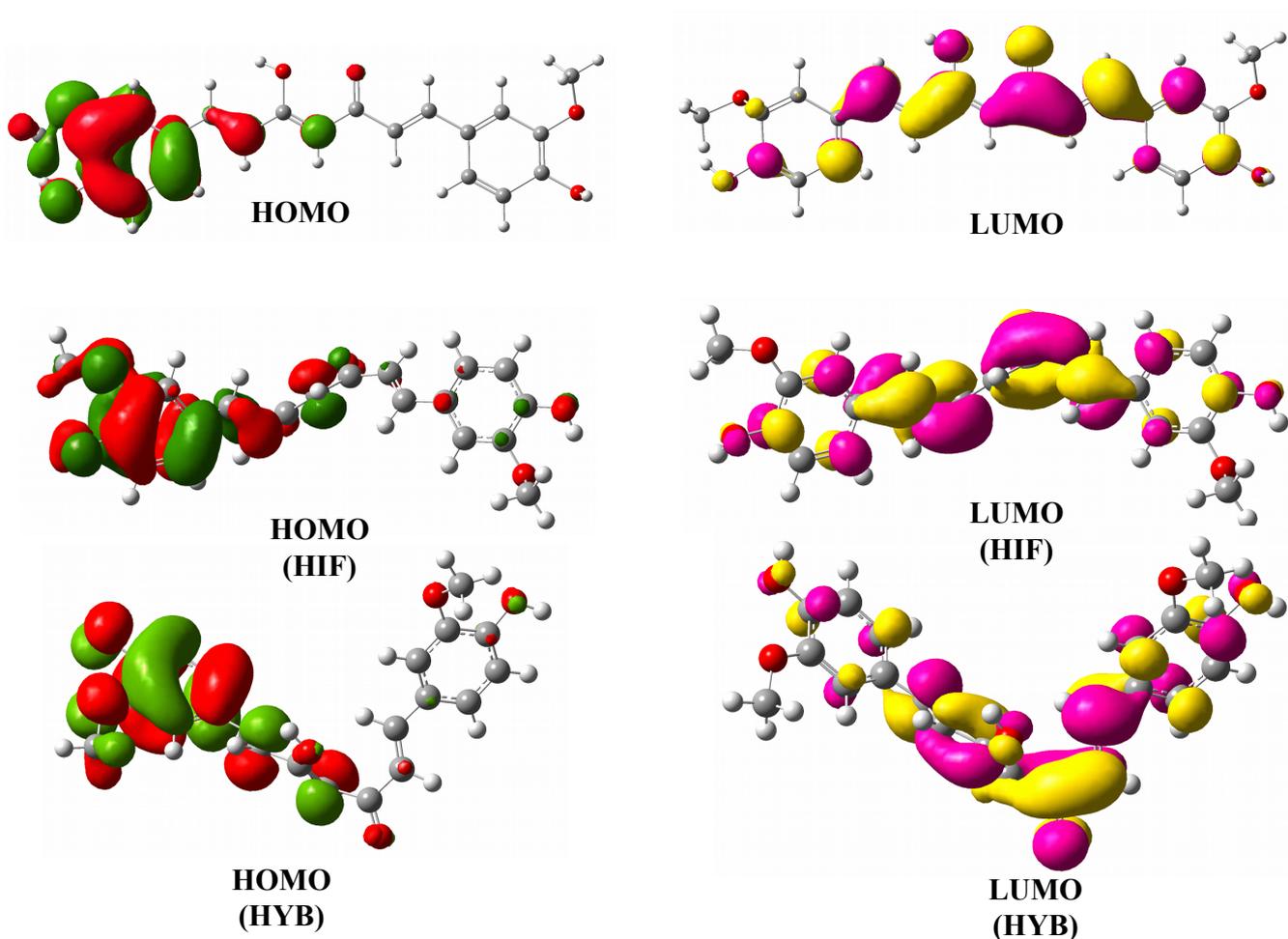


Figura 3.30: Representação dos orbitais HOMO e LUMO para a curcumina isolada e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de α -hélice. De cima para baixo, a representação do HOMO e LUMO do ligante isolado (sem interação), depois a representação do HOMO e LUMO da interação da curcumina no domínio hidrofílico (HIF) do peptídeo - configuração de menor energia - e, por fim, a representação do HOMO e LUMO da interação da curcumina no domínio hidrofóbico (HYB) - configuração de menor energia

A β_{1-42} em Folha- β

Em se tratando da folha- β , ao observar a distribuição da diferença de energia para os dois domínios do peptídeo (Figura 3.31), percebe-se uma faixa de energia menor para o domínio hidrofílico.

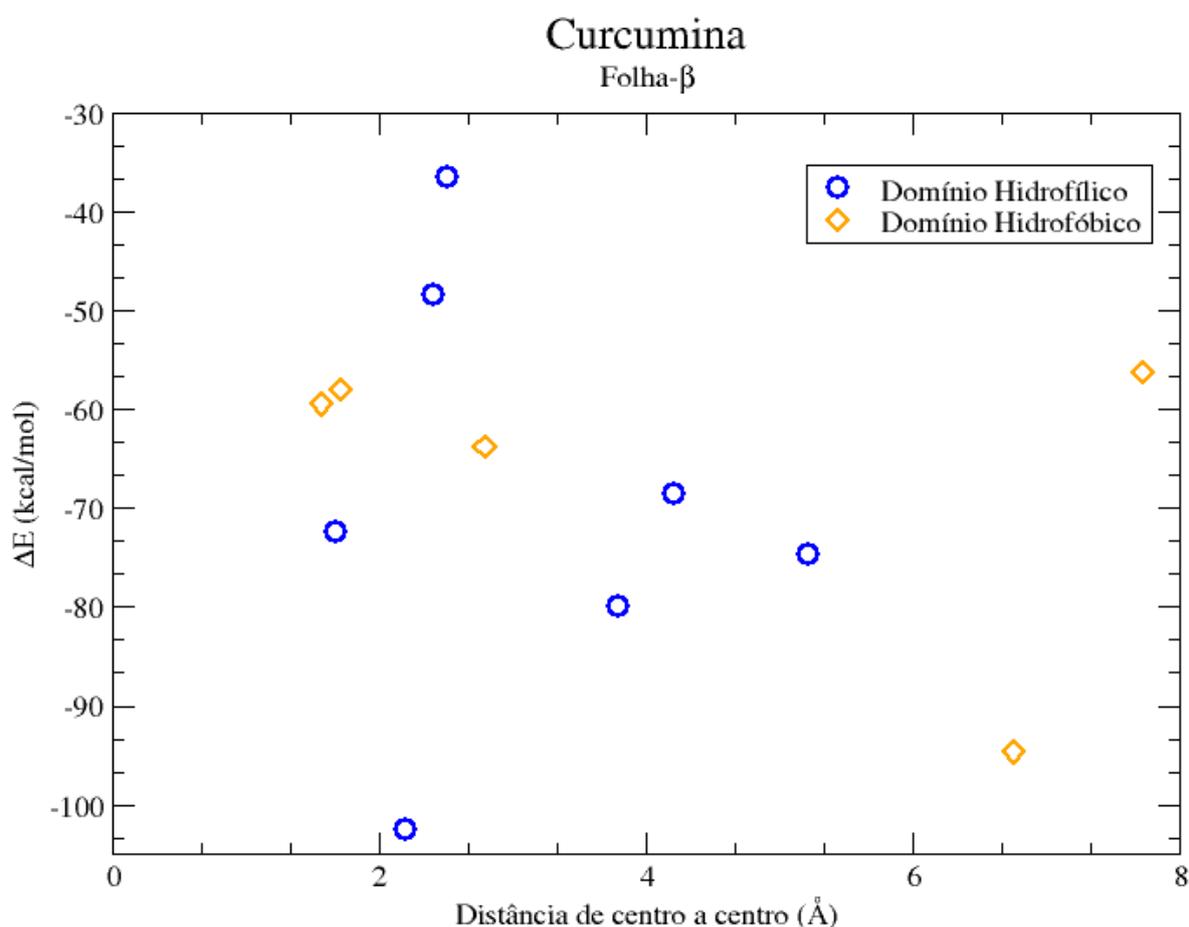


Figura 3.31: Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo A β_{1-42} mais a curcumina, levando em consideração a divisão do peptídeo nos domínios hidrofílico e hidrofóbico. A distância apresentada é do centro geométrico do ligante para o centro geométrico do grupo de aminoácidos escolhidos para a camada alta.

Tendo a Figura 3.31 como ponto de partida e, também, as conformações de menor e maior energia dentro de cada um dos domínios, mostradas nas Figuras

3.32 e 3.33, é possível observar que os tipos de interações mais relevantes, no domínio hidrofílico, dadas por uma possível interação cátion- π com a cadeia lateral da Arginina 5(ARG 5) e a ponte de hidrogênio entre o grupo amina da cadeia lateral da Lisina 16 (LYS 16) e a carbonila do ligante. A interação hidrofóbica vista é o π -*stacking* da Fenilalanina 4 (PHE 4) com um dos anéis aromáticos da curcumina. No domínio hidrofóbico, as interações de maior relevância energética são:

- a) Uma ponte de hidrogênio do grupo carbonila do *backbone* da Valina 36 (VAL 36) com a hidroxila central da curcumina mais interação da cadeia lateral desse aminoácido com o anel aromático;
- b) Outra ligação de hidrogênio da amina do backbone da Valina 39 (VAL 39) com a carbonila central do composto, além da interação da cadeia lateral desse aminoácido com o anel aromático e
- c) Interação da cadeia lateral da Isoleucina 41 (ILE 41) com um dos anéis aromáticos.

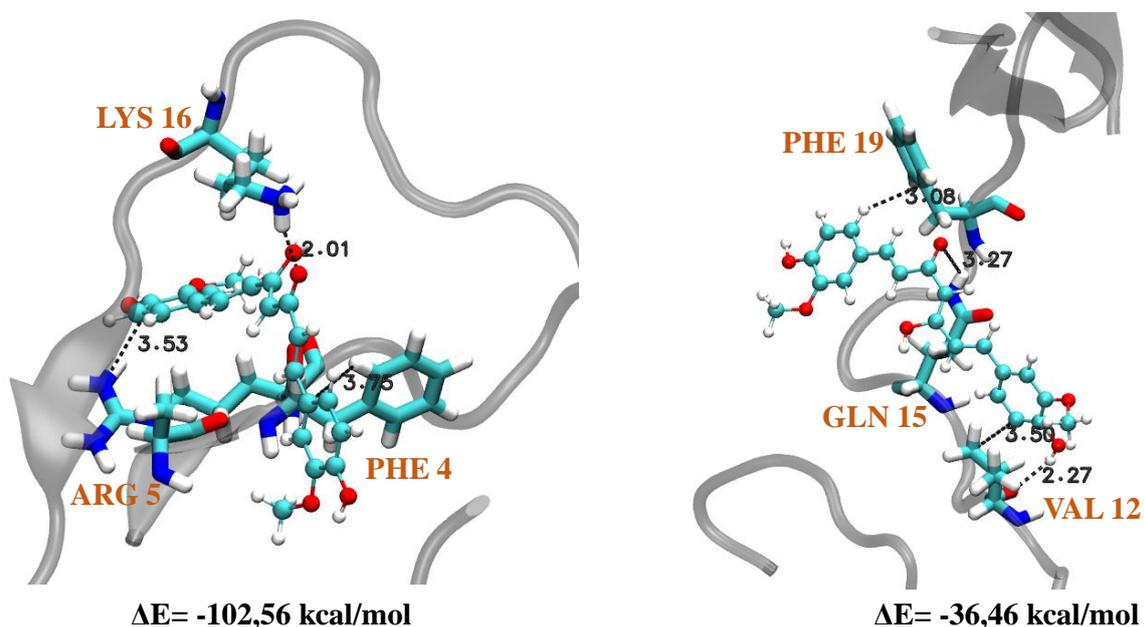


Figura 3.32: Interações da curcumina (representação de esfera e varetas) com os aminoácidos (representação em varetas) do $A\beta_{1-42}$ no domínio hidrofílico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são uma interação cátion- π com a Arg5 (3,53Å), uma ligação de hidrogênio com a Lys16 (2,01Å) e um π -stacking com a Phe4(3,75Å). Na figura da direita (maior energia), observa-se as interações: um T-stacking com a Phe19(3,09Å), uma ligação de hidrogênio da cadeia lateral da Gln15 com a carbonila central da curcumina (3,27Å) e duas interações com a Val12, uma da cadeia lateral com um dos anéis da curcumina (3,50Å) e uma ligação de hidrogênio da carbonila do *backbone* do aminoácido com uma das hidroxilas do ligante (2,27Å).

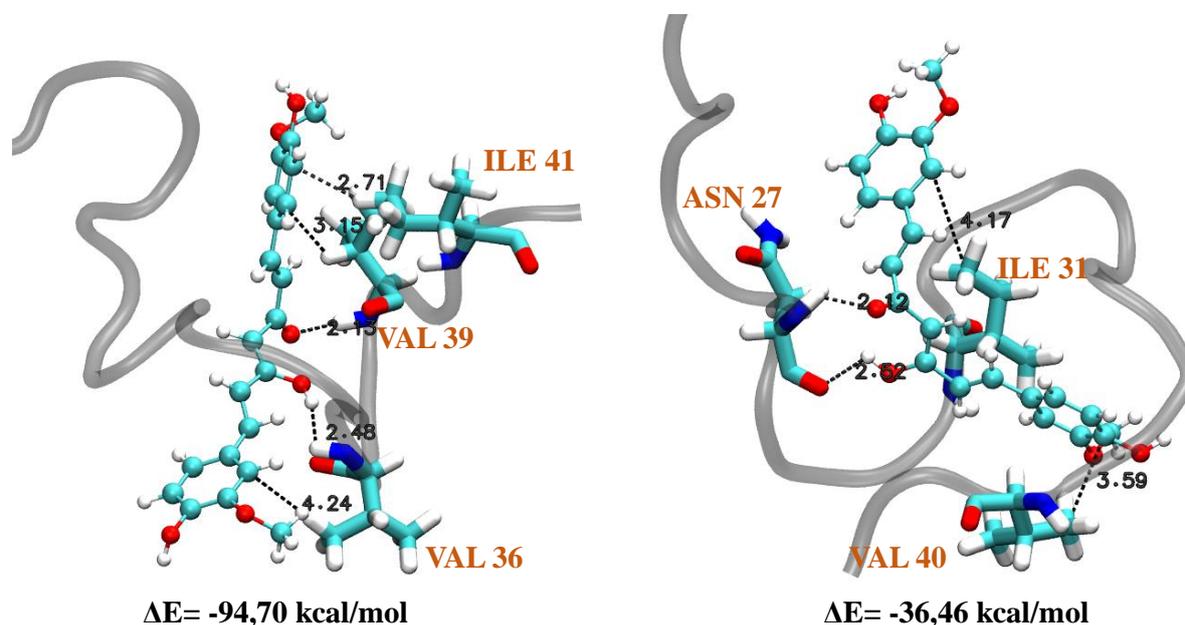


Figura 3.33: Interações da curcumina (representação de esfera e varetas) com os aminoácidos (representação em varetas) do $A\beta_{1-42}$ no domínio hidrofóbico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são duas com a Val36, uma ligação de hidrogênio (2,48Å) e outra interação hidrofóbica com um anel da curcumina (4,24Å), duas com a Val39, uma ligação de hidrogênio (2,13Å) e outra interação hidrofóbica com um anel da curcumina (3,15Å), e uma interação hidrofóbica com a Ile41 (2,71Å). Na figura da direita (maior energia), observa-se as seguintes interações: duas com a Asn27, uma ligação de hidrogênio da carbonila central da curcumina com a amina do *backbone* do aminoácido (2,12Å) e a carbonila do *backbone* do aminoácido com a hidroxila central do ligante (2,52Å), e outras duas interações hidrofóbicas, a cadeia lateral da Ile31 com um anel aromático (4,17Å) e a cadeia lateral da Val40 com o outro anel aromático da curcumina (3,59Å).

Tendo em vista as interações apresentadas nas Figuras 3.32 e 3.33 e da figura 3.34, que ilustra os orbitais de fronteira da curcumina nas configurações de menor energia e também do ligante isolado, é possível ver a alteração das nuvens eletrônicas de acordo com as interações do ligante com os aminoácidos ao seu redor. Para as interações da curcumina na folha- β , também há um deslocamento do orbital HOMO, mas isso ocorre de forma mais igualitária pela superfície do ligante do que

nos complexos com o $A\beta_{1-42}$ em conformação de α -hélice, o que provavelmente ajuda a alcançar uma menor energia e uma maior estabilidade do complexo, pois as energias dessas configurações são menores do que as energias dos complexos da curcumina com o peptídeo em α -hélice.

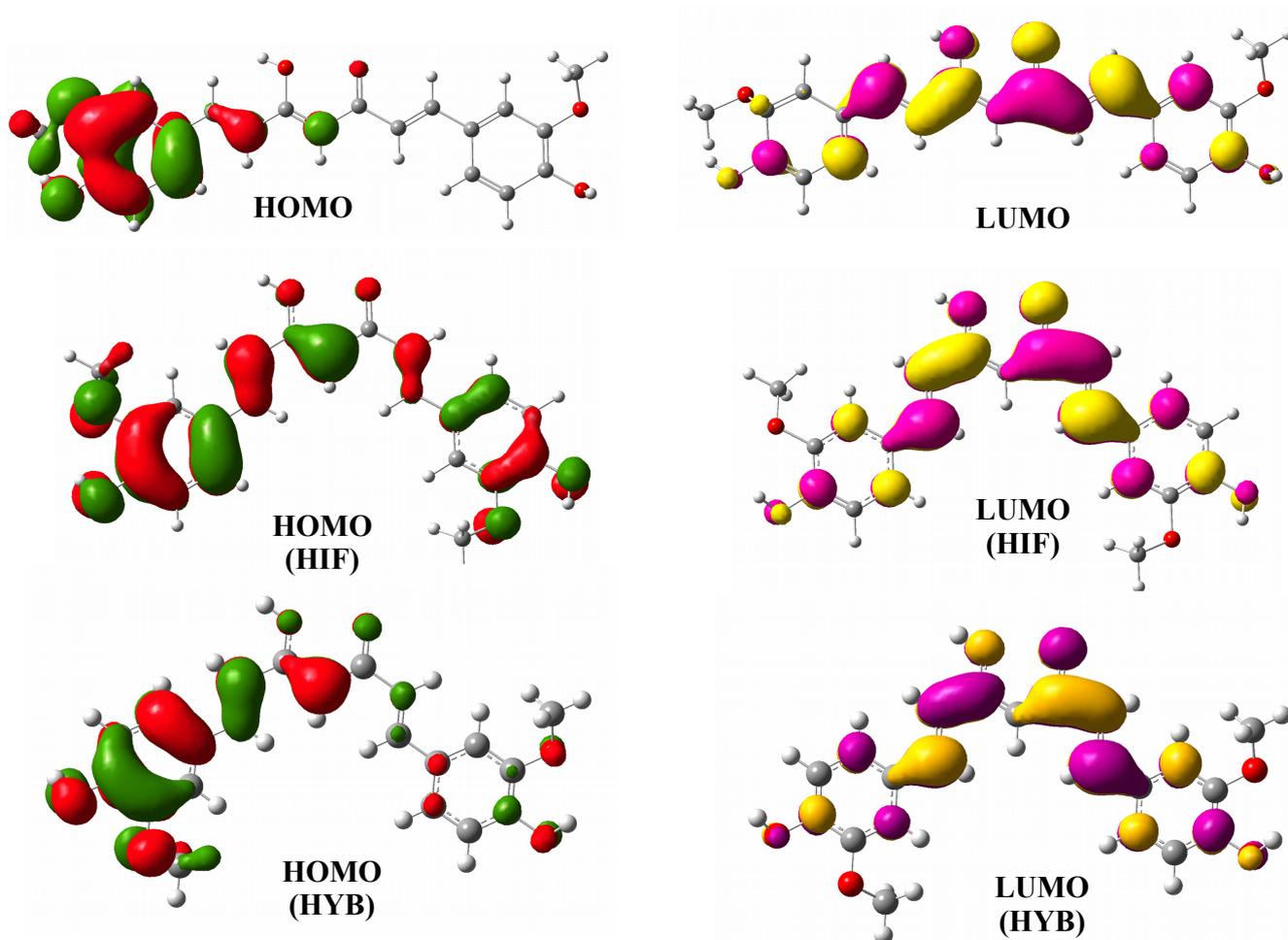


Figura 3.34: Representação dos orbitais HOMO e LUMO para a curcumina isolada e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de folha- β . De cima para baixo, a representação do HOMO e LUMO do ligante isolado (sem interação), depois a representação do HOMO e LUMO da interação da curcumina no domínio hidrofílico (HIF) do peptídeo - configuração de menor energia - e, por fim, a representação do HOMO e LUMO da interação do ligante no domínio hidrofóbico (HYB) - configuração de menor energia

3.3.3 Energias Eletrônicas das Interações do Ácido Rosmarínico com o Peptídeo $A_{\beta_{1-42}}$

Para o ácido rosmarínico das 25 conformações iniciais de cada domínio, obteve-se 8 estruturas otimizadas para o domínio hidrofílico e 7 para o hidrofóbico, na forma de α -hélice do peptídeo. Na folha- β , 12 para o domínio hidrofílico e 7 para o hidrofóbico.

A camada alta, na conformação de α -hélice do peptídeo, foi formada pelo composto mais os aminoácidos His13, Leu17 e Phe20 - domínio hidrofílico. Para o domínio hidrofóbico, os aminoácidos selecionados foram Ile32, Gly33 e Met35. Na folha- β , os aminoácidos da camada alta foram Phe4, Arg5 e Phe19 - domínio hidrofílico -, Ile31, Leu34 e Met35 - domínio hidrofóbico.

$A_{\beta_{1-42}}$ em α -hélice

Na Figura 3.35 é mostrada a distribuição da diferença de energia para os dois domínios do peptídeo em α -hélice. Nota-se que, assim como no cinamaldeído e na curcumina, as menores energias estão localizadas no domínio hidrofóbico.

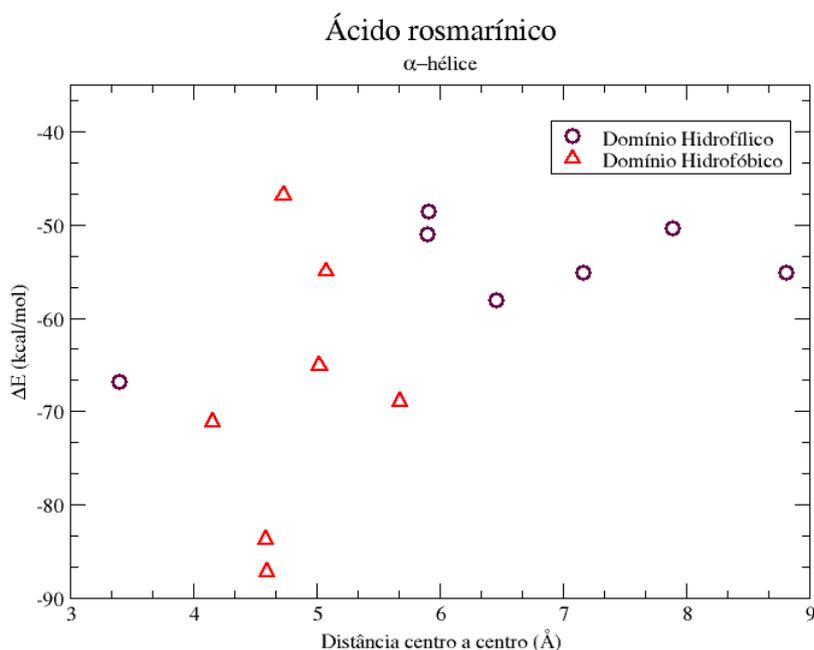


Figura 3.35: Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo $A\beta_{1-42}$ mais o ácido rosmarínico, levando em consideração a divisão do peptídeo nos domínios hidrofílico e hidrofóbico. A distância apresentada é do centro geométrico do ligante para o centro geométrico do grupo de aminoácidos escolhidos para a camada alta.

A partir da Figura 3.35 e, também, as conformações de menor e maior energia dentro de cada um dos domínios, mostradas nas Figuras 3.36 e 3.37, é possível observar que os tipos de interações mais relevantes, no domínio hidrofílico, são dadas por uma possível ligação de hidrogênio com a carbonila da cadeia lateral da Aspartato 23 (ASP 23) com a hidroxila do ligante e outra ligação de hidrogênio com a carbonila do *backbone* da Leucina 17 (Leu 17) mais uma das hidroxila do ácido rosmarínico. A interação hidrofóbica observada foi um π -stacking com o anel aromático da Fenilalanina 20 (PHE 20). Para o domínio hidrofóbico, as interações mais relevantes foram as pontes de hidrogênio formadas pelos *backbones* dos aminoácidos Glutamato 22 (GLU 22), Leucina 34 (LEU 34), Metionina 35 (MET 35) e Valina 39 (VAL 39) com o ácido rosmarínico, além da ponte de hidrogênio da carbonila do Glu22 com a hidroxila do ligante.

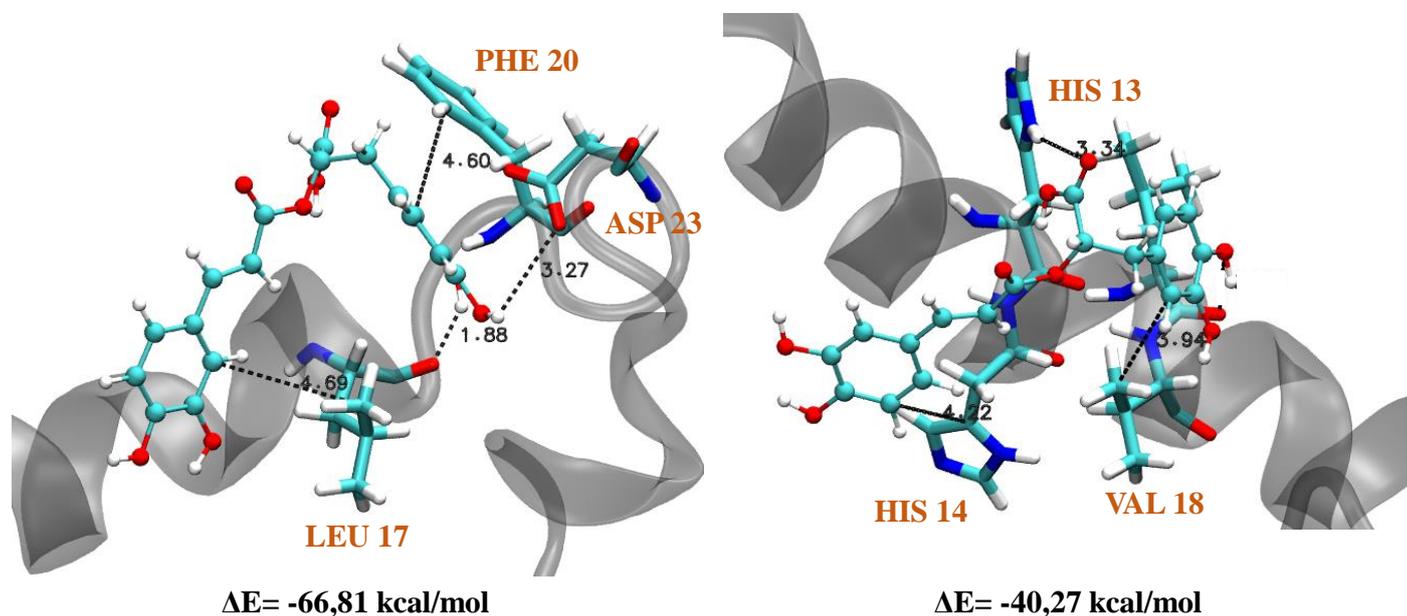


Figura 3.36: Interações do ácido rosmarínico (representação de esfera e varetas) com os aminoácidos (representação em varetas) do A β_{1-42} no domínio hidrofílico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são um π -stacking com a Phe20 (4,60Å), uma ligação de hidrogênio com Asp23(3,27Å) e outra com a Leu17 (1,88Å) e, por fim, um anel do composto com a cadeia lateral da Leu17 (4,69Å). Na figura da direita (maior energia), percebe-se as interações: um π -stacking com a His14 (4,22Å), uma ligação de hidrogênio com a His13 (3,34Å) e a cadeia lateral da Val18 com um anel aromático do ácido rosmarínico (3,94Å).

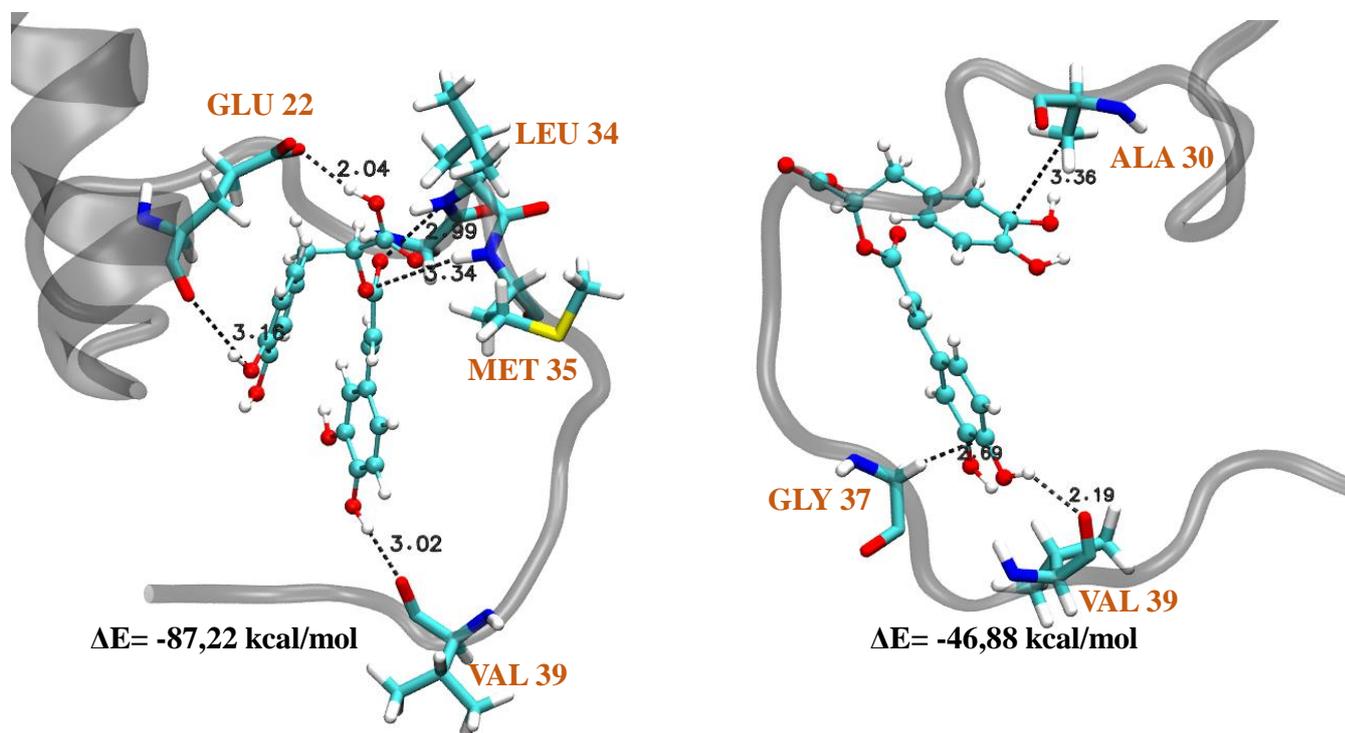


Figura 3.37: Interações do ácido rosmarínico (representação de esfera e varetas) com os aminoácidos (representação em varetas) do A β_{1-42} no domínio hidrofóbico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são todas ligações de hidrogênio, a saber: Glu 22, carbonila do *backbone* do aminoácido com hidroxila do ligante (3,16Å) e outra da carbonila da cadeia lateral do Glu22 com hidroxila do composto (2,04Å), Leu 34, grupo amina do *backbone* do aminoácido com carbonila do ligante (2,99Å), Met35, grupo amina do *backbone* do aminoácido com carbonila do ligante (3,34Å), e Val39, carbonila do *backbone* do aminoácido com hidroxila do ligante (3,02Å). Na figura da direita (maior energia), observa-se as seguintes interações: uma ligação de hidrogênio da carbonila do *backbone* do aminoácido Val39 com hidroxila de um dos anéis do ácido rosmarínico (2,19Å), e outras duas interações hidrofóbicas, a cadeia lateral da Ala30 com um anel aromático (3,36Å) e a cadeia lateral da Gly37 com o outro anel aromático da curcumina (2,69Å).

Tendo em vista as interações apresentadas acima e da figura 3.38, que ilustra os orbitais de fronteira do ácido rosmarínico nas configurações de menor energia e também do ligante isolado, é possível ver a alteração das nuvens eletrônicas de acordo com as interações do ligante com os aminoácidos ao seu redor. Ao contrário

dos outros dois compostos, o ácido rosmarínico apresenta um deslocamento tanto do orbital HOMO, quanto do orbital LUMO em relação aos orbitais do ligante isolado.

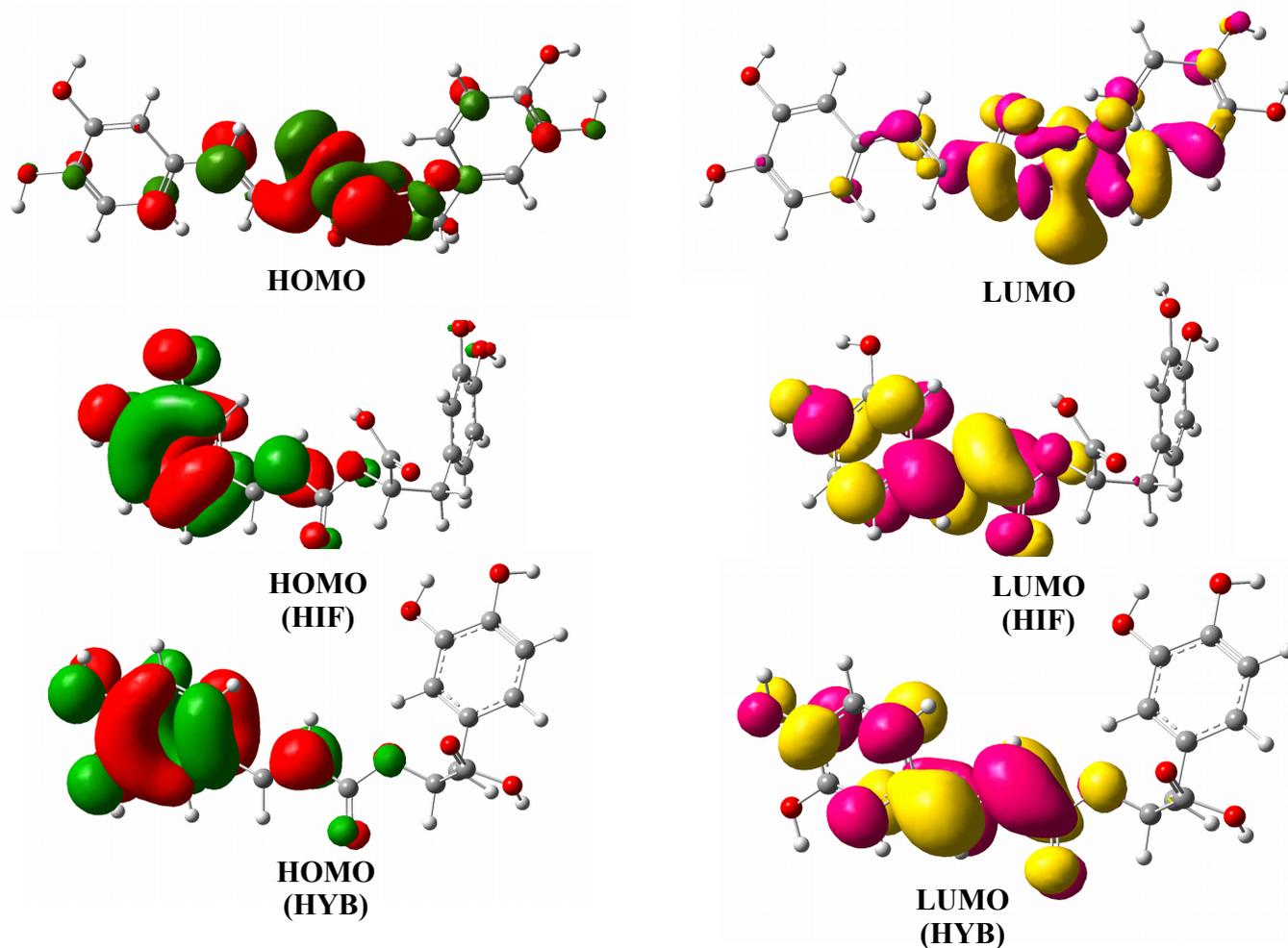


Figura 3.38: Representação dos orbitais HOMO e LUMO para o ácido rosmarínico isolado e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de α -hélice. De cima para baixo, a representação do HOMO e LUMO do ligante isolado (sem interação), depois a representação do HOMO e LUMO da interação do ácido rosmarínico no domínio hidrofílico (HIF) do peptídeo - configuração de menor energia - e, por fim, a representação do HOMO e LUMO da interação do ligante no domínio hidrofóbico (HYB) - configuração de menor energia

$A_{\beta_{1-42}}$ em Folha- β

Em se tratando da folha- β , ao observar a distribuição da diferença de energia para os dois domínios do peptídeo (Figura 3.39), percebe-se uma faixa de energia menor para o domínio hidrofílico.

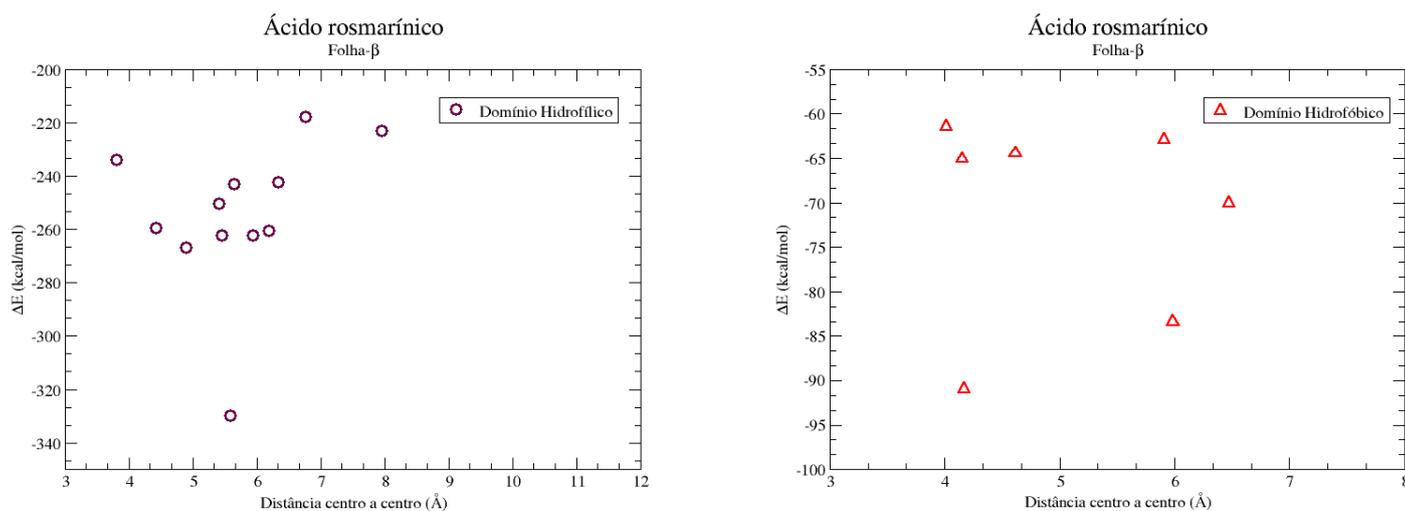


Figura 3.39: Distribuição da diferença de energia eletrônica do complexo formado pelo peptídeo $A_{\beta_{1-42}}$ mais ácido rosmarínico, levando em consideração a divisão do peptídeo nos domínios hidrofílico e hidrofóbico. A distância apresentada é do centro geométrico do ligante para o centro geométrico do grupo de aminoácidos escolhidos para a camada alta.

Analisando a Figura 3.39 e, também, as conformações de menor e maior energia dentro de cada um dos domínios (Figuras 3.40 e 3.41) é possível observar que os tipos de interações mais relevantes, no domínio hidrofílico, que possui nas duas conformações energéticas valores bem menores comparados a todos os outros encontrados nos resultados deste trabalho, são dadas por ligações de hidrogênio observadas com os aminoácidos Arginina 5 (ARG 5), amina do *backbone* do aminoácido com carbonila central do ligante, Glutamina 15 (GLN 15), carbonila do *backbone* do aminoácido com hidroxila do ligante - configuração de menor energia -, Glicina 9 (GLY 9), carbonila do *backbone* do aminoácido com hidroxila do ligante, Histidina 13 (HIS 13), carbonila do *backbone* do aminoácido com hidroxila central do ligante,

e Arg5, carbonila do *backbone* do aminoácido com hidroxila central do ligante - configuração de maior energia. As interações hidrofóbicas são todas por empilhamento, são elas: dois T-*stacking* com os aminoácidos Phe4 e Phe20, para a configuração de menor energia, e um π -*stacking* com a His6 e um T-*stacking* com com a Phe19, na configuração de maior energia.

No que tange ao domínio hidrofóbico, as interações de maior relevância energética são:

- As ligações de hidrogênio do grupo carbonila do *backbone* dos aminoácidos Ile32, Gly33, Met35 e Gly37 com as hidroxilas dos anéis do ácido rosmarínico;
- A ligação de hidrogênio do grupo amina do *backbone* da Val39 com a carbonila central do composto, além da interação da cadeia lateral deste aminoácido com o anel aromático e
- A interação hidrofóbica da cadeia lateral da Isoleucina 31 (ILE 31) com um dos anéis aromáticos do ligante.

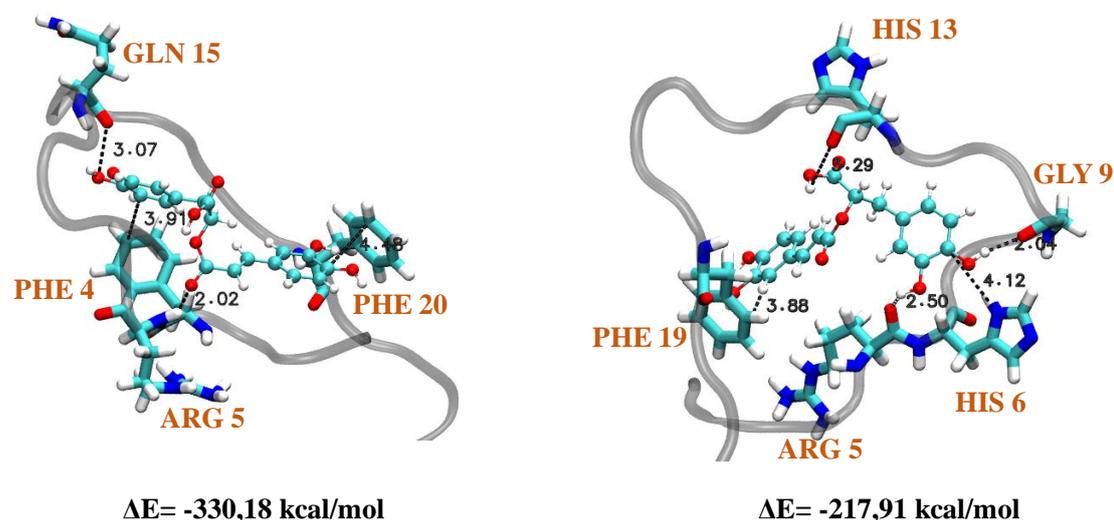


Figura 3.40: Interações do ácido rosmarínico (representação de esfera e varetas) com os aminoácidos (representação em varetas) do $A\beta_{1-42}$ no domínio hidrofílico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) e na figura da direita (maior energia) estão descritas no texto.

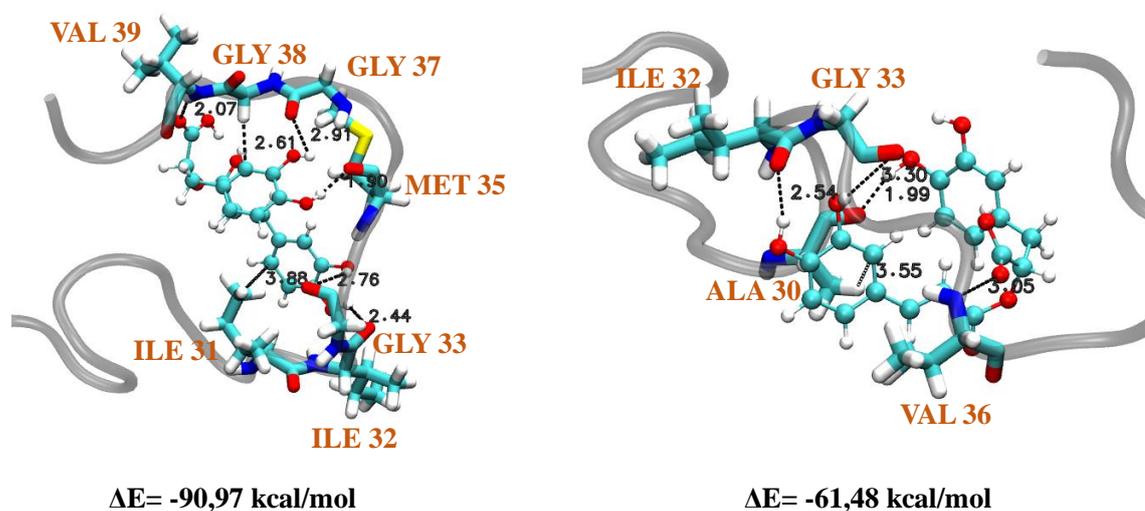


Figura 3.41: Interações do ácido rosmarínico (representação de esfera e varetas) com os aminoácidos (representação em varetas) do $A\beta_{1-42}$ no domínio hidrofóbico. A cor vermelha mostra os oxigênios, a azul os nitrogênios e a branca os hidrogênios. As distâncias são mostradas em Angströms. As interações mostradas na figura da esquerda (menor energia) são todas ligações de hidrogênio e estão descritas no texto. Na figura da direita (maior energia), observa-se as seguintes interações: ligações de hidrogênio do grupo carbonila do *backbone* dos aminoácidos Ala30 (1,99Å), Ile32 (2,54Å) e Gly33 (3,30Å) com as hidroxilas dos anéis do ácido rosmarínico e, também do grupo amina do *backbone* da Val36 com a carbonila do ácido rosmarínico (3,05Å). A interação hidrofóbica presente é da cadeia lateral da Ala30 com um dos anéis do ligante (3,55Å).

Considerando as interações apresentadas nas Figuras 3.40 e 3.41 e os orbitais de fronteira do ácido rosmarínico nas configurações de menor energia e também do ligante isolado (Figura 3.42), é possível ver a alteração das nuvens eletrônicas de acordo com as interações do ligante com os aminoácidos ao seu redor. É observado as mesmas alterações vistas com os complexos do peptídeo em α -hélice.

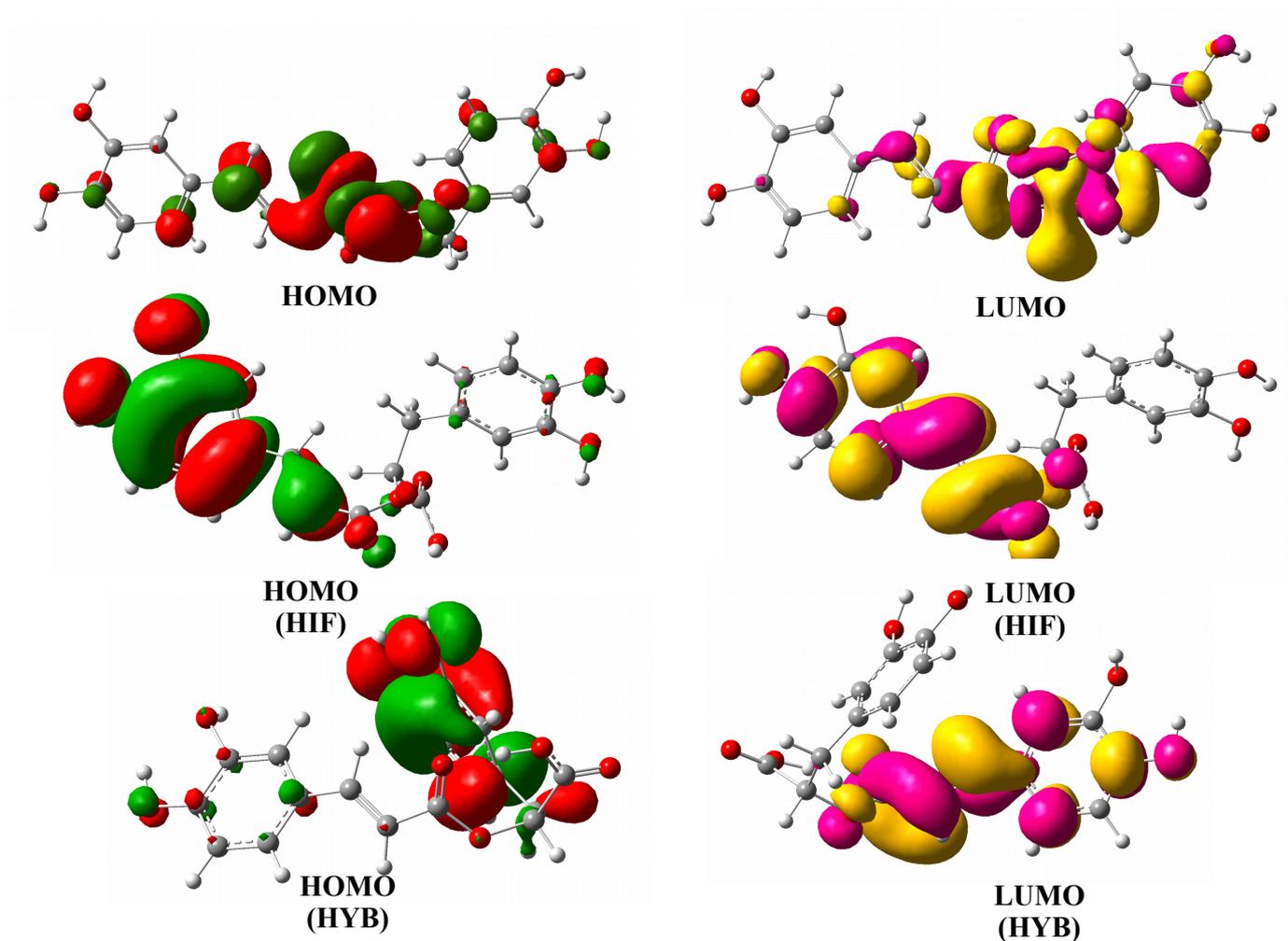


Figura 3.42: Representação dos orbitais HOMO e LUMO para o ácido rosmarínico isolado e interagindo com o peptídeo $A\beta_{1-42}$ em sua conformação de folha- β . De cima para baixo, a representação do HOMO e LUMO do ligante isolado (sem interação), depois a representação do HOMO e LUMO da interação do ácido rosmarínico no domínio hidrofílico (HIF) do peptídeo - configuração de menor energia - e, por fim, a representação do HOMO e LUMO da interação do ligante no domínio hidrofóbico (HYB) - configuração de menor energia

Capítulo 4

Conclusões e Perspectivas

Neste trabalho foi feito um estudo de reconhecimento molecular dos compostos naturais cinamaldeído, curcumina e ácido rosmarínico pelo peptídeo $A\beta_{1-42}$. A fim de viabilizar esse estudo, fez-se o uso de dinâmica molecular para avaliação da flexibilidade do peptídeo, da estabilidade de suas estruturas secundárias, α -hélices e folhas- β , e para amostrar um grande número das possíveis conformações desse alvo molecular. A partir do *ensemble* de conformações adquirido por meio da DM fez-se o *docking* molecular para localizar os possíveis sítios de interação dos compostos naturais no alvo molecular. Partindo desse cenário, as energias dos complexos obtidos via *docking* foram estimadas através de cálculos híbridos de QM/MM, usando o método ONIOM2, para serem observadas as configurações mais estáveis. Por fim, caracterizou-se as possíveis interações dos complexos com menor energia usando medidas de distância e análise dos orbitais de fronteira.

Diante dos resultados obtidos, pode-se sugerir que os aminoácidos possivelmente dotados de um papel importante para impedir a agregação do peptídeo $A\beta_{1-42}$ ou minimamente manter sua estrutura de α -hélice próxima do N-terminal ao interagir com os compostos estudados são Phe4, Leu17, His13, Phe19, Phe 20, Ile31, Ile32 e Val39. Esses aminoácidos tiveram interação com os três compostos. Há também os aminoácidos que compartilharam interações com pelo menos dois dos compostos e, provavelmente, também têm uma participação, a ser considerada,

para evitar o processo de agregação, são eles: Arg5, His14, Gln15, Lys16, Glu22, Asn27, Gly33, Leu34, Met35, Val36, Gly37 e Ile41. A importância desses aminoácidos está vinculada às interações do tipo ligações de hidrogênio e empilhamento de anéis aromáticos. Sendo as ligações de hidrogênio feitas com o *backbone* dos aminoácidos, as que levam à menor energia. Portanto, é bastante provável que, quando o $A\beta_{1-42}$ sofre a mudança da sua estrutura secundária de α -hélice para folhas- β , o que faz esses compostos naturais impedirem a agregação do peptídeo para formar placas neuríticas sejam as interações dessa natureza, pois, assim, impede-se que um núcleo de estruturas do tipo folha- β seja formado. Lembrando que este núcleo é o que caracteriza as proteínas como amiloidogênicas, ou seja, formadoras de placas neuríticas.

A partir dessas interações, também pode-se perceber que grupos funcionais como anéis aromáticos, carbonila e hidroxila são personagens importantes para se ter ação no peptídeo $A\beta_{1-42}$. Embora, a curcumina e o ácido rosmarínico tenham mostrado energias menores em seus complexos em comparação com o cinamaldeído, entende-se, por meio dos resultados obtidos nesta pesquisa, que o tamanho da estrutura química não é o mais importante para obter-se um fármaco com ação no $A\beta_{1-42}$ e sim estruturas que tenham seus anéis aromáticos acoplados a hidroxilas (grupo catecol) e grupos carbonila ao longo da cadeia carbônica do ligante, mesmo que não seja uma cadeia extensa. Além disso, a flexibilidade do ligante parece ter influência na obtenção de mínimos de energia, pois o ácido rosmarínico, que apresentou as menores energias dos resultados obtidos, mostrou uma flexibilidade na sua região central, que contribui para o alcance de aminoácidos com um papel importante na agregação do peptídeo.

Perspectivas

Esse estudo abre várias perspectivas futuras, entre os quais mencionamos:

- 1) Realizar um estudo de decomposição da energia total de interação dos complexos formados pelo peptídeo $A\beta_{1-42}$ e os compostos naturais cinamaldeído, curcumina e ácido rosmarínico. Este estudo visa determinar as principais contribuições das interações não-covalentes (troca, coulomb, dispersão e difusão) para entender com mais detalhes como os compostos naturais cinamaldeído, curcumina e ácido rosmarínico podem inibir o processo de oligomerização do peptídeo $A\beta_{1-42}$. A realização deste estudo pode ser realizada utilizando o método SAPT (do inglês *Symmetry-adapted perturbation theory*);
- 2) Determinar a transferências de cargas durante as interações entre o peptídeo $A\beta_{1-42}$ e os compostos naturais cinamaldeído, curcumina e ácido rosmarínico. Esse estudo torna-se importante pois complementa o estudo da decomposição das principais componentes das interações dos complexos aqui estudados;
- 3) Estimar, através da teoria das estruturas de transição, a taxa de reação em função da temperatura (com e sem efeitos de tunelamento), do primeiro estágio antes do início do processo de oligomerização, ou seja, o rearranjo conformacional do peptídeo $A\beta_{1-42}$ ao passar de α -hélice para folha- β . A viabilização deste estudo pode ser importante para conhecer como se dá a construção de agregados de vários tamanhos (dímeros, trímeros, tetramêros, etc) até chegar na forma de placas senis, o produto final da agregação. Esse estudo pode ser estendido para os complexos formados pelo peptídeo $A\beta_{1-42}$ e compostos naturais cinamaldeído, curcumina e ácido rosmarínico;
- 4) Utilizar o mesmo protocolo do presente trabalho para realizar um novo estudo com outros ligantes. A escolha dessas novas moléculas pode ser feita através do método virtual *screening*. Os critérios de seleção dos possíveis novos ligantes podem ser baseados nos resultados obtidos nesta tese de doutorado.

Referências Bibliográficas

- [1] KANDEL, Eric; SCHWARTZ, James; JESSELL, Thomas. *Princípios da Neurociência*. Editora Manole, 2003.
- [2] Who. https://www.who.int/gho/mortality_burden_disease/life_tables/situation_trends/en/. acessado em novembro de 2019.
- [3] BREITVE, Monica H.; CHWISZCZUK, Luiza J.; HYNNINEN, Minna J. et al. A systematic review of cognitive decline in dementia with lewy bodies versus alzheimer's disease. *Alzheimer's research & therapy*, 6(5):53, 2014.
- [4] CUMMINGS, Jeffrey L. Dementia: the falling brain. *The Lancet*, 345, 1995.
- [5] KALAITZAKIS, M.E.; PEARCE, R.K.B. The morbid anatomy of dementia in parkinson's disease. *Acta Neuropathologica*, 118(587), 2009.
- [6] TAVARES, Almir; AZEREDO, Camilo. Demência com corpos de lewy: uma revisão para o psiquiatra. *Revista de Psiquiatria Clínica*, 30(1):29–34, 2003.
- [7] GUERREIRO, Rita; ESCOTT-PRICE, Valentina; DARWENT, Lee. et al. Genome-wide analysis of genetic correlation in dementia with lewy bodies, parkinson's and alzheimer's diseases. *Neurobiology of aging*, 38:214–e7, 2016.
- [8] HEPP, Dagmar H.; VERGOOSSEN, Dana L. E.; HUISMAN, Evelien. et al. Distribution and load of amyloid- β pathology in parkinson disease and dementia with lewy bodies. *Journal of Neuropathology & Experimental Neurology*, 75(10):936–945, 2016.

-
- [9] MASUDA, Masami; SUZUKI, Nobuyuki; TANIGUCHI, Sayuri. et al. Small molecule inhibitors of α -synuclein filament assembly. *Biochemistry*, 45(19):6085–6094, 2006.
- [10] FUJIWARA, Hideo; HASEGAWA, Masato; DOHMAE, Naoshi. et al. α -synuclein is phosphorylated in synucleinopathy lesions. *Nature cell biology*, 4(2):160–164, 2002.
- [11] SPILLANTINI, Maria Grazia; SCHMIDT, Marie Luise; LEE, Virginia M-Y. et al. α -synuclein in lewy bodies. *Nature*, 388(6645):839–840, 1997.
- [12] BABA, Minami; NAKAJO, Shigeo; TU, Pang-Hsien. et al. Aggregation of alpha-synuclein in lewy bodies of sporadic parkinson's disease and dementia with lewy bodies. *The American journal of pathology*, 152(4):879, 1998.
- [13] SPILLANTINI, Maria Grazia; CROWTHER, R. Anthony; JAKES, Ross; HASEGAWA, Masato; GOEDERT, Michel. α -synuclein in filamentous inclusions of lewy bodies from parkinson's disease and dementia with lewy bodies. *Proceedings of the National Academy of Sciences*, 95(11):6469–6473, 1998.
- [14] HALLIDAY, Glenda M.; HOLTON, Janice L.; REVESZ, Tamas; DICKSON, Dennis W. Neuropathology underlying clinical variability in patients with synucleinopathies. *Acta neuropathologica*, 122(2):187–204, 2011.
- [15] JUCKER, Mathias & WALKER, Lary C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature*, 501(7465):45–51, September 2013.
- [16] JÚNIOR, C. O. G.; FELÍCIO, A. C.; PRADO, G. F. Sistema extrapiramidal: Anatomia e síndromes clínicas. *Rev. Neurociencia*, 14(1):48–51, 2006.
- [17] GIACOBINI, Ezio; GOLD, Gabriel. Alzheimer disease therapy - moving from amyloid- β to Tau. *Nature Reviews Neurology*, 9(12):677–686, 2013.

-
- [18] CITRON, Martin; OLTERS DORF, Tilman; HAASS, Christian. et al. Mutation of the β -amyloid precursor protein in familial alzheimer's disease increases β -protein production. *Nature*, 360(6405):672–674, 1992.
- [19] PURVES, Dale; AUGUSTINE, George J.; FITZPATRICK, David. et al. *Neurociências-4*. Artmed editora, 2010.
- [20] CORDER, E. H.; SAUDERS, A. M.; STRITTMATTER, W. J. et al. Gene dose of apolipoprotein e type 4 allele and the risk of alzheimer's disease in late onset families. *Science*, 261(5123):921–923, 1993.
- [21] LEON, Rafael; GARCIA, Antonio G.; MARCO-CONTELLES, José. Recent advances in the multitarget-directed ligands approach for the treatment of alzheimer's disease. *Medicinal research reviews*, 33(1):139–189, 2013.
- [22] DIAS, Kris Simone T.; VIEGAS, Claudio. Multi-target directed drugs: a modern approach for design of new drugs for the treatment of alzheimer's disease. *Current neuropharmacology*, 12(3):239–255, 2014.
- [23] CANTER, Rebecca G.; PENNEY, Jay; TSAI, Li-Huei. The road to restoring neural circuits for the treatment of alzheimer's disease. *Nature*, 539(7628):187–196, 2016.
- [24] TORRES, Carolina Akkari. Glicogênio sintase quinase3b e proteína precursora do amilóide em plaquetas de indivíduos com comprometimento cognitivo leve e doença de alzheimer. Tese de doutorado, Universidade de São Paulo, 2009.
- [25] YANG, Cao; LI, Jinyu; LI, Yu; ZHU, Xiaolei. The effect of solvents on the conformations of amyloid β -peptide(1-42) studied by molecular dynamics simulation. *Journal of Molecular Structure: THEOCHEM*, 895(1):1–8, 2009.
- [26] GLABE, Charles C. Amyloid accumulation and pathogenesis of alzheimer's disease: significance of monomeric, oligomeric and fibrillar $\alpha\beta$. *Sub-cellular biochemistry*, 38:167–177, 2005.

-
- [27] SPIES, Petra E.; VERBEEK, Marcel M.; van GROEN, Thomas; CLAASSEN, Jurgen A. H. R. Reviewing reasons for the decreased csf $a\beta_{42}$ concentration in alzheimer disease. *Frontiers in bioscience (Landmark edition)*, 17:2024–2034, 2012.
- [28] BELO, Paulo Ricardo Neiva. Formação de placas amilóides em doenças neurodegenerativas. Dissertação de mestrado, Universidade de Coimbra, 2014.
- [29] FORNER, Stefania; BAGLIETTO-VARGAS, David; MARTINI, Alessandra C. et al. Synaptic impairment in alzheimer's disease: A dysregulated symphony. *Trends in Neurosciences*, 40(6):347–357, 2017.
- [30] BU, Xian-Le; RAO, Praveen P. N.; WANG, Yan-Jiang. Anti-amyloid aggregation activity of natural compounds: Implications for alzheimer's drug discovery. *Molecular Neurobiology*, 53(6):3565–3575, 2016.
- [31] TRIGUERO, Luciano; SINGH, Rajiv; PRABHAKAR, Rajeev. Comparative molecular dynamics studies of wild-type and oxidized forms of full-length alzheimer amyloid β -peptides $a\beta(1-40)$ and $a\beta(1-42)$. *The journal of physical chemistry. B*, 112(23):7123–7131, June 2008.
- [32] XIAO, Y.; MA, B.; MCELHENY, D. et al. $A\beta(1-42)$ fibril structure illuminates self-recognition and replication of amyloid in alzheimer's disease. *Nature Structural & Molecular Biology volume*, 22:499–505, 2015.
- [33] NGOUNGOURE, Viviane L. Ndam; SCHLUESENER, Jan; MOUNDIPA, Paul F.; SCHLUESENER, Hermann. Natural polyphenols binding to amyloid: A broad class of compounds to treat different human amyloid diseases. *Molecular Nutrition & Food Research*, 59(1):8–20, 2015.
- [34] URBANC, Brigita; CRUZ, Luis; TELOW, David B.; STANLEY, H. Eugene. Computer simulations of alzheimer's amyloid β -protein folding and assembly. *Current Alzheimer research*, 3(5):493–504, December 2006.

-
- [35] URBANC, B.; CRUZ, L.; YUN, S. et al. In silico study of amyloid β -protein folding and oligomerization. *Proceedings of the National Academy of Sciences of the United States of America*, 101(50):17345â17350, December 2004.
- [36] MIRMOSAYYEB, Omid; TANHAEI, Amirpouya; SOHRABI, Hamid. et al. Possible role of common spices as a preventive and therapeutic agent for alzheimer's disease. *International Journal of Preventive Medicine*, 8(1):5, 2017.
- [37] FRYDMAN-MAROM, Anat; LEVIN, Aviad; FARFARA, Dorit. et al. Orally administrated cinnamon extract reduces β -amyloid oligomerization and corrects cognitive impairment in alzheimer's disease animal models. *Plos One*, 6:1-11, 01 2011.
- [38] REINKE, Ashley A.; GESTWICKI, Jason E. Structure-activity relationships of amyloid beta-aggregation inhibitors based on curcumin: Influence of linker length and flexibility. *Chemical Biology & Drug Design*, 70(3):206-215, 2007.
- [39] RANASINGHE, Priyanga; PIGERA, Shehani; PREMAKUMARA, G. A. et al. Medicinal properties of "true" cinnamon (*cinnamomum zeylanicum*): a systematic review. *BMC complementary and alternative medicine*, 13:275, October 2013.
- [40] MOMTAZ, Saeideh; HASSANI, Shokoufeh; KHAN, Fazlullah; ZIAEE, Mojtaba; ABDOLLAHI, Mohammad. Cinnamon, a promising prospect towards alzheimer's disease. *Pharmacological research*, 130:241-258, 2018.
- [41] PETERSON, Dylan W.; GEORGE, Roshni C.; SCARAMOZZINO, Francesca. et al. Cinnamon extract inhibits tau aggregation associated with alzheimer's disease in vitro. *Journal of Alzheimer's disease: JAD*, 17(3):585â597, 2009.
- [42] EILERMAN, Robert G. *Cinnamic Acid, Cinnamaldehyde, and Cinnamyl Alcohol*, pages 1-11. American Cancer Society, 2014.

-
- [43] EMAMGHOREISHI, Masoumeh; FARROKHI, Majid; AMIRI, Atena; KESHAVARZ, Mojtaba. The neuroprotective mechanism of cinnamaldehyde against amyloid- β in neuronal SHSY5Y cell line: The role of N-methyl-D-aspartate, ryanodine, and adenosine receptors and glycogen synthase kinase-3 β . *Avicenna Journal of Phytomedicine*, 9(3):271–280, 2019.
- [44] MEHRAEIN, Fereshteh; ZAMANI, Maryam; NEGAHDAR, Feraidoon; SHOJAEE, Asieh. Cinnamaldehyde attenuates dopaminergic neuronal loss in substantia nigra and induces midbrain catalase activity in a mouse model of parkinson's disease. *Journal of Basic and Clinical Pathophysiology*, 6(1):9–16, 2018.
- [45] MODI, Khushbu K.; ROY, Avik; BRAHMACHARI, Saurabh. et al. Cinnamon and its metabolite sodium benzoate attenuate the activation of p21rac and protect memory and learning in an animal model of alzheimer's disease. *PLOS ONE*, 10(6):1–22, 06 2015.
- [46] AGGARWAL, Bharat B.; SUNDARAM, Chitra; MALANI, Nikita; ICHIKAWA, Haruyo. *CURCUMIN: THE INDIAN SOLID GOLD*, pages 1–75. Springer US, Boston, MA, 2007.
- [47] SHARMA, R.A.; GESCHER, A.J.; STEWARD, W.P. Curcumin: the story so far. *European journal of cancer (Oxford, England: 1990)*, 41(13):1955–1968, 2005.
- [48] RAO, Praveen P. N.; MOHAMED, Tarek; TECKWANI, Karan; TIN, Gary. Curcumin binding to beta amyloid: A computational study. *Chemical Biology & Drug Design*, 86(4):813–820, 2015.
- [49] ALAGAWANY, Mahmoud; ABD EL-HACK, Mohamed; FARAG, Mayada Ragab. et al. Rosmarinic acid: modes of action, medicinal values and health benefits. *Animal Health Research Reviews*, 18(2):167–176, 2017.

-
- [50] AIROLDI, Cristina; SIRONI, Erika; DIAS, Catarina. et al. Natural Compounds against Alzheimer's Disease: Molecular Recognition of A β 1-42 Peptide by *Salvia sclareoides* Extract and its Major Component, Rosmarinic Acid, as Investigated by NMR. *Chemistry - An Asian Journal*, 8(3):596–602, 2013.
- [51] NAMBA, A. M.; da SILVA, V. B.; da SILVA, C. H. T. P. Dinâmica molecular: teoria e aplicações em planejamento de fármacos. *Eclética Química*, 33(4), 2008.
- [52] ADCOCK, Stewart A.; MCCAMMON, J. Andrew. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chemical reviews*, 106(5):1589–1615, 2006.
- [53] ANDREW, R. Leach. *Molecular modeling principles and applications*. Pearson Education Limited, 2001.
- [54] LISZKA, Tadeusz; ORKISZ, Janusz. The finite difference method at arbitrary irregular grids and its application in applied mechanics. *Computers & Structures*, 11(1-2):83–95, 1980.
- [55] GRUBMÜLLER, Helmut; HELLER, Helmut; WINDEMUTH, Andreas; SCHULTEN, Klaus. Generalized verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Molecular Simulation*, 6(1-3):121–142, 1991.
- [56] MARTYS, Nicos S.; MOUNTAIN, Raymond D. Velocity verlet algorithm for dissipative-particle-dynamics-based models of suspensions. *Physical Review E*, 59(3):3733, 1999.
- [57] VANOMMESLAEGHE, Kenno; HATCHER, Elizabeth; ACHARYA, Chayan. et al. Charmm general force field: A force field for drug-like molecules compatible with the charmm all-atom additive biological force fields. *Journal of computational chemistry*, 31(4):671–690, 2010.

-
- [58] WANG, Junmei; WOLF, Romain M.; CALDWELL, James W. et al. Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9):1157–1174, 2004.
- [59] FRATERNALI, F.; VAN GUNSTEREN, W.F. An efficient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution. *Journal of molecular biology*, 256(5):939–948, 1996.
- [60] JORGENSEN, William L.; MAXWELL, David S.; TIRADO-RIVES, Julian. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc*, 118(45):11225–11236, 1996.
- [61] ANDERSEN, Hans C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52(1):24–34, 1983.
- [62] DARDEN, Tom; YORK, Darrin; PEDERSEN, Lee. Particle mesh ewald: An $n \log(n)$ method for ewald sums in large systems. *The Journal of chemical physics*, 98(12):10089–10092, 1993.
- [63] ANJOS, Luiz Carlos Araújo dos. Simulação molecular e predição de propriedades físico-químicas de lubrificantes biodegradáveis. Tese de doutorado, Universidade Federal de Pernambuco, 2009.
- [64] LARANJEIRAS, Cássio C.; CHIAPPIN, José R.N. A construção de uma teoria de ensembles: antecedentes em Maxwell e Boltzmann. *Revista Brasileira de Ensino de Física*, 30, 2008.
- [65] BARROS, Emília Pécora de. Estudo computacional da solvatação e anisotropia de fluorescência resolvida no tempo de subtilisina carlsberg em misturas de água e glicerol. Dissertação de mestrado, Universidade Estadual de Campinas-Unicamp, 2015.

-
- [66] SADUS, Richard J. *Molecular simulation of fluids: theory, algorithms and object-orientation*. Elsevier Science, 2002.
- [67] OLIVEIRA, Krerley; VIANA, Marcelo. Fundamentos da teoria ergódica. *Endereço eletrônico* <http://www.impa.br/viana/out/fte.pdf>, 2014.
- [68] MUDI, Anirban; CHAKRAVARTY, Charusita. Effect of the berendsen thermostat on the dynamical properties of water. *Molecular Physics*, 102(7):681–685, 2004.
- [69] BERENDSEN, Herman J. C.; POSTMA, J. P. M. van; van GUNSTEREN, Wilfred F. et al. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.
- [70] TSALLIS, Constantino; STARIOLO, Daniel A. Generalized simulated annealing. *Physica A: Statistical Mechanics and its Applications*, 233(1):395–406, 1996.
- [71] FELLER, Scott E.; ZHANG, Yuhong; PASTOR, Richard W.; BROOKS, Bernard R. Constant pressure molecular dynamics simulation: the langevin piston method. *The Journal of chemical physics*, 103(11):4613–4621, 1995.
- [72] KARPLUS, M.; PETSKO, G. A. Molecular dynamics simulations in biology. *Nature*, 347(6294):631–639, October 1990.
- [73] KITCHEN, Douglas B.; DECORNEZ, Hélène; FURR, John R.; BAJORATH, Jürgen. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews Drug discovery*, 3(11):935–949, 2004.
- [74] MENG, Xuan-Yu; ZHANG, Hong-Xing; MEZEI, Mihaly; CUI, Meng. Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design*, 7(2):146–157, 2011.

-
- [75] MORRI, Garrett M.; HUEY, Ruth; OLSON, Arthur J. Using autodock for ligand-receptor docking. *Current protocols in bioinformatics*, pages 8–14, 2008.
- [76] SILVA, Mônica de Abreu. Modelos preditivos baseados em descritores moleculares e modos de interação receptor-ligante para inibidores de acetilcolinesterase. *Tese de doutorado. Universidade de Brasilia-UnB*, 2017.
- [77] FUHRMANN, Jan; RURAINSKI, Alexander; LENHOF, Hans-Peter; NEUMANN, Dirk. A new lamarckian genetic algorithm for flexible ligand-receptor docking. *Journal of computational chemistry*, 31(9):1911–1918, 2010.
- [78] HEERDT, Gabriel; MORGON, Nelson H. Validação computacional de métodos compostos no estudo de propriedades moleculares. *Química Nova*, 34:868–873, 2011.
- [79] SZABO, Attila; OSTLUND, Neil S. *Modern quantum chemistry: Introduction to advanced electronic structure theory*. Dover Publications, 1996.
- [80] FARIA, Cristiano Geraldo de. Estudo da geometria e estrutura eletrônica de nanotubos de carbono quirais puros e dopados. Dissertação de mestrado, Universidade Federal de Itajubá, 2009.
- [81] SLATER, J. C. A simplification of the hartree-fock method. *Phys. Rev.*, 81:385–390, Feb 1951.
- [82] RAMACHADRAN, K.I.; DEEPA, G.; NAMBOORI, K. *Computational Chemistry and Molecular Modeling*. Springer-Verlag Berlin Heidelberg, 2008.
- [83] CLEMENTI, Enrico. Computation of large molecules with the hartree-fock model. *Proceedings of the National Academy of Sciences*, 69(10):2942–2944, 1972.
- [84] DAHL, J. P. On the application of zero differential overlap methods to complex molecules. *Acta Chemica Scandinavica*, 21:1244–1250, 1967.

-
- [85] KOLLMAR, Christian; BÖHM, Michael C. An analysis of the zero differential overlap approximation. towards an improved semiempirical mo method beyond it. *Theoretica chimica acta*, 92(1):13–47, Jul 1995.
- [86] DEWAR, Michael J. S.; THIEL, Walter. Ground states of molecules. 38. the mndo method. approximations and parameters. *Journal of the American Chemical Society*, 99(15):4899–4907, 1977.
- [87] STEWART, James J. P. Optimization of parameters for semiempirical methods v: Modification of nndo approximations and application to 70 elements. *Journal of Molecular Modeling*, 13(12):1173–1213, 2007.
- [88] CHRISTENSEN, A. S.; KUBAŘ, T.; CUI, Q.; ELSTNER, M. Semiempirical quantum mechanical methods for noncovalent interactions for chemical and biochemical applications. *Chemical reviews*, 116(9):5301–5337, 2016.
- [89] VOSKO, S. H.; WILK, L.; NUSAIR, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.
- [90] SUN, Jianwei; XIAO, Bing; FANG, Yuan. et al. Density functionals that recognize covalent, metallic, and weak bonds. *Phys. Rev. Lett.*, 111, Sep 2013.
- [91] ZHAO, Yan; TURHLAR, Donald G. . A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *The Journal of Chemical Physics*, 125(19), 2006.
- [92] MORGON, Nelson Henrique; CUSTÓDIO, Rogério. Funções de base: o ajuste variacional. *Revista Chemkeys*, pages 1–11, set 2018.
- [93] BINKLEY, J. Stephen; POPLE, John A.; HEHRE, Warren J. Self-consistent molecular orbital methods. 21. small split-valence basis sets for first-row elements. *Journal of the American Chemical Society*, 102(3):939–947, 1980.

-
- [94] BRAGA, Ataulpa A. C.; MORGON, Nelson H. Cálculos teóricos de afinidades por prÃade n-alquilaminas usando o método ONIOM. *Química Nova*, 29:187–193, 04 2006.
- [95] MOROKUMA, Keiji. Oniom and its applications to material chemistry and catalyses. *Bulletin of the Korean Chemical Society*, 24(6):797–801, 06 2003.
- [96] CHUNG, Lung Wa; SAMEERA, W. M. C.; RAMOZZI, Romain. et al. The oniom method and its applications. *Chemical reviews*, 115(12):5678–5796, June 2015.
- [97] PHILLIPS, James C.; ZHENG, Gengbin.; KUMAR, Sameer; KALÉ, Laxmikant V. Namd: Biomolecular simulation on thousands of processors. In *Supercomputing, ACM/IEEE 2002 Conference*, pages 36–36. IEEE, 2002.
- [98] KLAUDA, Jeffery B.; VENABLE, Richard M.; FREITES, J. Alfredo. et al. Update of the charmm all-atom additive force field for lipids: validation on six lipid types. *The journal of physical chemistry B*, 114(23):7830–7843, 2010.
- [99] HUMPHREY, William; DALKE, Andrew; SCHULTEN, Klaus. Vmd: visual molecular dynamics. *Journal of molecular graphics*, 14(1):33–38, 1996.
- [100] FLÖCK, Dagmar; COLACINO, Stefano; COLOMBO, Giorgio; DI NOLA, Alfredo. Misfolding of the amyloid β -protein: A molecular dynamics study. *Proteins: Structure, Function, and Bioinformatics*, 62(1):183–192, 2006.
- [101] MA, Buyong; NUSSINOV, Ruth. Simulations as analytical tools to understand protein aggregation and predict amyloid conformation. *Current Opinion in Chemical Biology*, 10(5):445 – 452, 2006.
- [102] Rscb:pdb-101. <https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/dealing-with-coordinates>. acessado em março de 2020.

-
- [103] TAO, Peng; SCHLEGEL, H. Bernhard. A toolkit to assist oniom calculations. *Journal of computational chemistry*, 31(12):2363–2369, September 2010.
- [104] FRISCH, M. J.; TRUCKS, G. W.; SCHLEGEL, H. B. et al. Gaussian¹/₄09 Revision E.01. Gaussian Inc. Wallingford CT 2009.
- [105] LUKAC, Iva; ABDELHAKIM, Hend; WARD, Richard A. et al. Predicting protein-ligand binding affinity and correcting crystal structures with quantum mechanical calculations: lactate dehydrogenase a. *Chemical science*, 10(7):2218–2227, February 2019.
- [106] BARREIRO, Eliezer J.; FRAGA, Carlos Alberto Manssour. *Química Medicinal: As bases moleculares da ação dos fármacos*. Artmed Editora, 2008.
- [107] CAMPBELL, Mary K.; FARRELL, Shawn O. *Bioquímica*. Cengage Learning, 2016.
- [108] CORNEJO, Victor Hugo; HETZ, Claudio. The unfolded protein response in alzheimer’s disease. *Seminars in Immunopathology*, pages 277–292, 2013.
- [109] ELLIS, R. John; PINHEIRO, Teresa J. T. Danger-misfolding proteins. *Nature*, 416(6880):483–484, 2002.
- [110] FAÍSCA, Patrícia F. N. O mistério da forma das proteínas. *Gazeta de Física*, 29:34–39, 2006.
- [111] GODOI CONTESSOTO, Vinícius de; OLIVEIRA JUNIOR, Antonio Bento de; CHAHINE, Jorge. et al. Introdução ao problema de enovelamento de proteínas: uma abordagem utilizando modelos computacionais simplificados. *Revista Brasileira de Ensino de Física*, 40, 2018.
- [112] Openstax. <https://openstax.org/books/biology-ap-courses/pages/3-4-proteins>. acessado em dezembro de 2019.

Apêndice A

Estrutura Química dos Aminoácidos

Os aminoácidos são moléculas orgânicas que participam na formação de proteínas como unidade fundamental. Em sua estrutura molecular (Figura A.1) há sempre um carbono central (C_{α}) ligado a um hidrogênio (H), a um grupo carboxila (COOH), a um grupo amina (NH_2) e a um radical "R", que varia de aminoácido para aminoácido. Esse radical é que determina as características do aminoácido [106, 107].

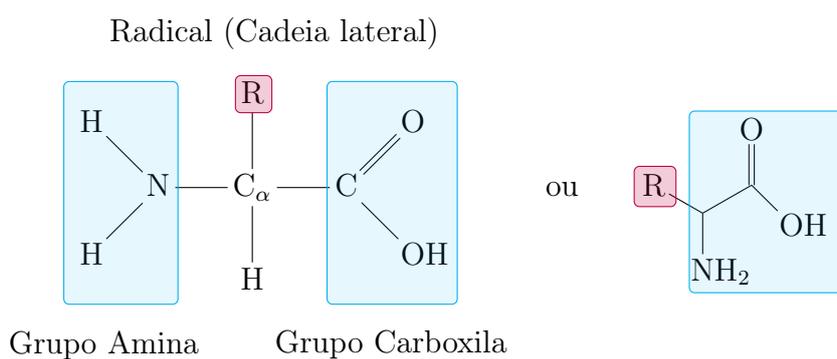


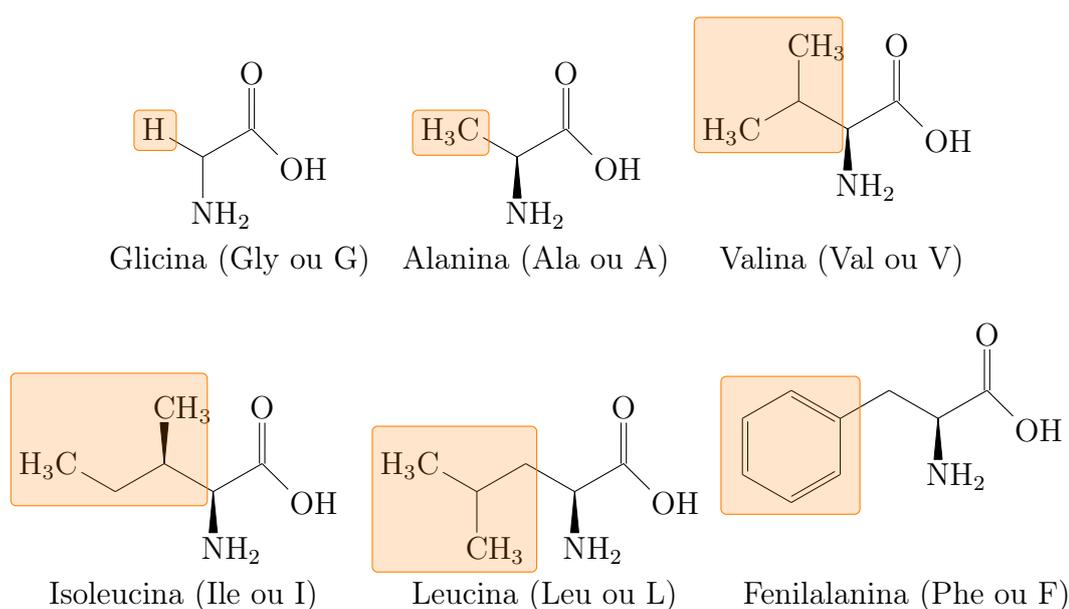
Figura A.1: Fórmula geral dos aminoácidos.

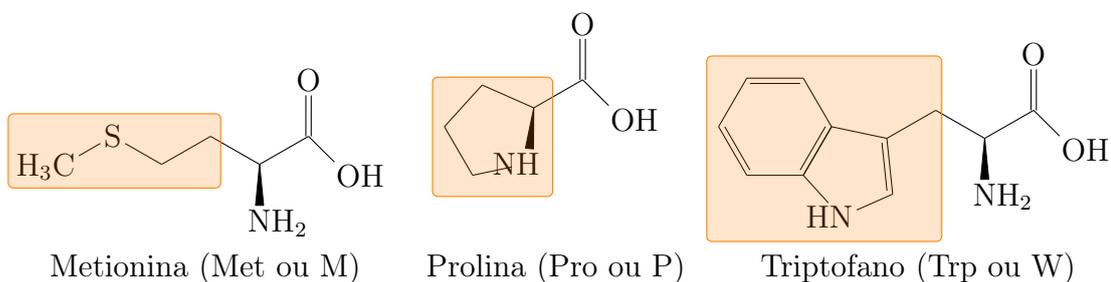
Quando dois aminoácidos se unem, eles são chamados de dipeptídeo, se são três forma-se um tripeptídeo e assim sucessivamente, sendo que a união de

vários aminoácidos constrói uma cadeia polipeptídica. São conhecidos 20 aminoácidos - Alanina, Arginina, Aspartato, Asparagina, Cisteína, Fenilalanina, Glicina, Glutamato, Glutamina, Histidina, Isoleucina, Leucina, Lisina, Metionina, Prolina, Serina, Tirosina, Treonina, Triptofano e Valina - encontrados nas proteínas, com sua síntese controlada por mecanismos genéticos, envolvendo a replicação do DNA e transcrição do RNA. A metade dos aminoácidos é sintetizada pelo organismo e é destinada a suprir as necessidades celulares. Os aminoácidos que não são produzidos pelo próprio organismo precisam estar presentes na dieta e são chamados de aminoácidos essenciais, os que são sintetizados são conhecidos como não-essenciais [107].

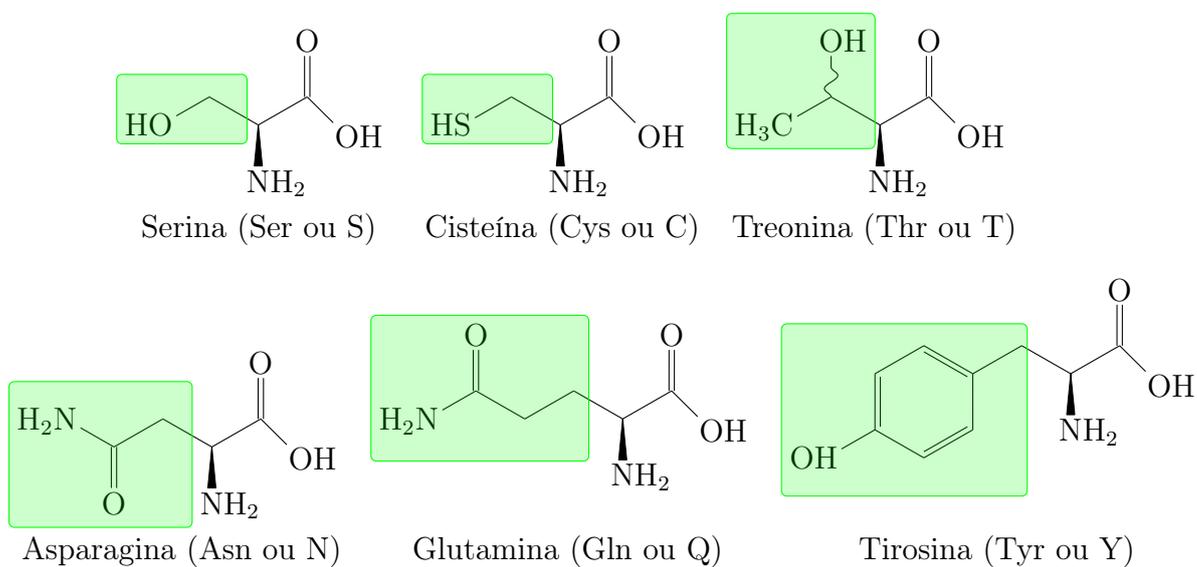
Os aminoácidos são frequentemente divididos pelas características das suas cadeias laterais, pois o *backbone* é o mesmo para todos. Abaixo segue a lista dos 20 aminoácidos presentes nas proteínas separados por sua classificação em apolares, polares, básicos (grupo amina como cadeia lateral) e ácidos (grupo carboxila como cadeia lateral).

Apolares

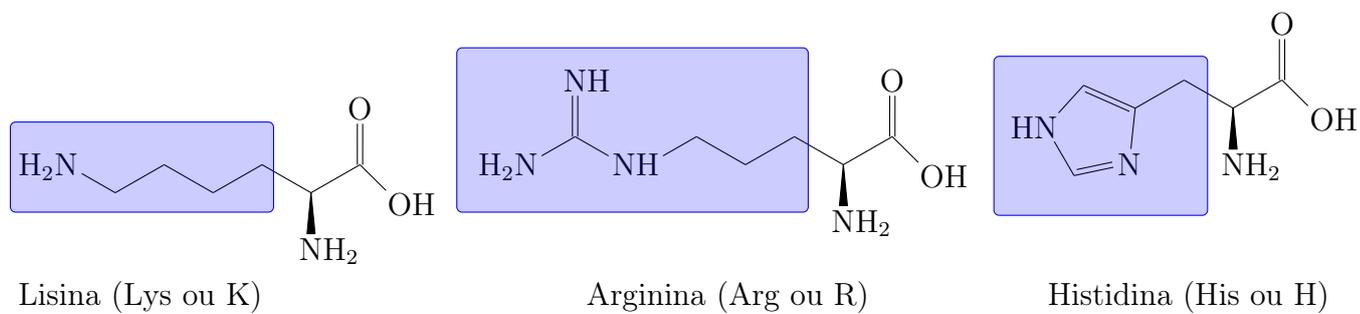




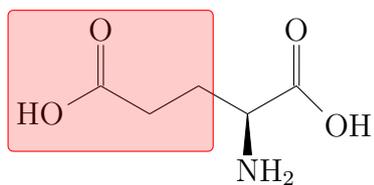
Polares



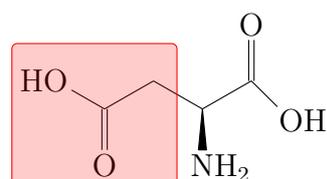
Básicos



Ácidos



Ácido glutâmico ou Glutamato (Glu ou E)



Ácido aspártico ou Aspartato (Asp ou D)

Apêndice B

Enovelamento(*fold*ing) Proteico

Quando aborda-se os biomarcadores presentes na maioria das doenças neurodegenerativas, o principal achado histopatológico compartilhado entre elas são os agregados e as inclusões anormais de proteína compostos por proteínas "mal enoveladas" como, por exemplo, as placas neuríticas presentes na DA. Esses acúmulos ocorrem devido ao erro de enovelamento proteico (*misfold*ing, em inglês) (Figura B.1) resultando em agregados que não possuem a função da proteína enovelada corretamente (*fold*ing, em inglês) e, algumas vezes, são grandes o suficiente para serem insolúveis [108, 109].

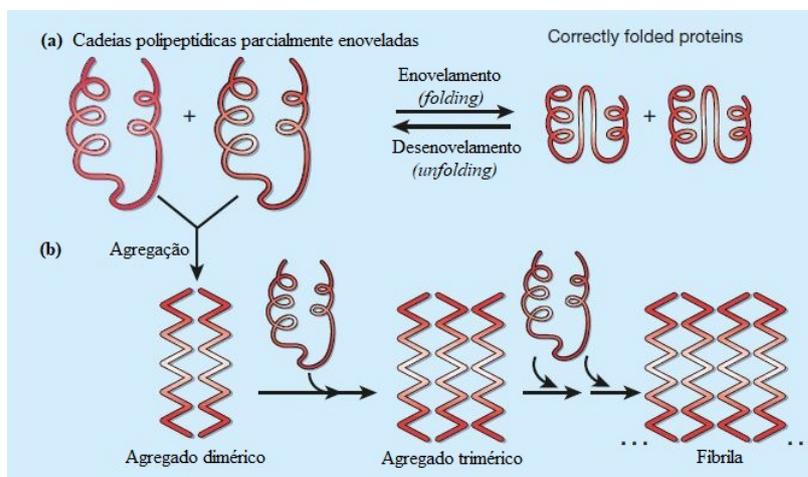


Figura B.1: Processo de agregação das proteínas. **(a)** cadeias polipeptídicas parcialmente dobradas, liberadas pelos ribossomos (estruturas que fazem a síntese das proteínas), normalmente tornam-se proteínas na sua forma nativa e, portanto, funcional. **(b)** no entanto, polipeptídeos parcialmente dobrados podem às vezes associar-se a cadeias semelhantes para formar agregados. Os agregados variam em tamanho, desde dímeros e trímeros solúveis até estruturas fibrilares insolúveis. Ao contrário da maioria das proteínas dobradas corretamente, agregados solúveis e insolúveis podem ser tóxicos para as células através de mecanismos ainda desconhecidos

Fonte: Adaptado de Ellis & Pinheiro[109]

Para que uma proteína possa funcionar corretamente, ela precisa de uma estrutura tridimensional única que determina sua função biológica. Essa estrutura é chamada de nativa, sendo ela o produto final do *folding* proteico da cadeia de aminoácidos que compõe a proteína. Basta uma falha, por menor que seja, na estrutura nativa para que esta não funcione da maneira correta [110].

Há quatro níveis de organização para a estrutura de uma proteína: a estrutura primária, que é a sequência dos aminoácidos formadores da proteína; a secundária, que são as conformações locais como folhas- β e α -hélices; a terciária (nativa), que é seu arranjo tridimensional; e a quaternária, que é um arranjo tridimensional de várias estruturas terciárias em um complexo multi-proteico (Figura B.2). O processo pelo qual uma cadeia linear de aminoácidos (estrutura primária) chega ao seu estado tridimensional nativo é chamado de enovelamento de proteínas [111]. Assim, quando há uma falha nesse procedimento surgem os agregados a partir

de estruturas intermediárias que expõem temporariamente suas regiões hidrofóbicas na sua superfície, permitindo que interajam com outras moléculas ao invés de se voltarem para a parte interna da estrutura final [109].

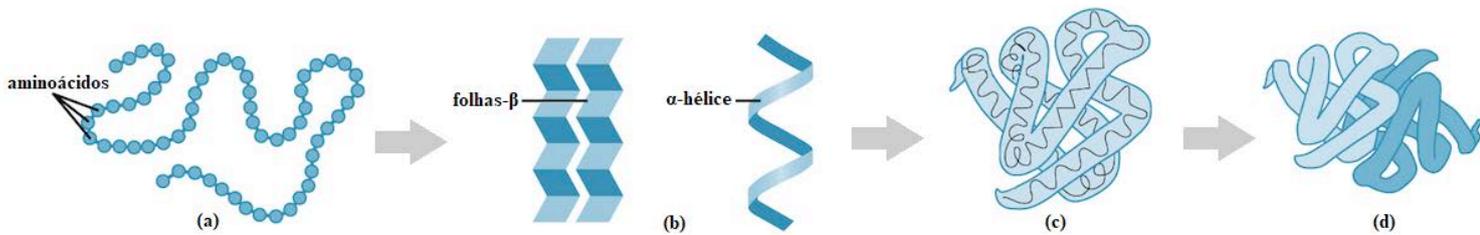


Figura B.2: Níveis estruturais de uma proteína. (a) Estrutura primária, (b) Estrutura secundária, (c) Estrutura terciária e (d) Estrutura quaternária.

Fonte: Adaptado de Openstax biology [112]

Nesse cenário, percebe-se que o processo defeituoso do *fold*ing é a porta de entrada para as doenças neurodegenerativas como a DA.

Apêndice C

Scripts para Gerar os *Inputs* do ONIOM

Nesta seção, serão apresentados alguns dos scrips usados para gerar os *inputs* para o cálculo com o ONIOM, que foram preparados a partir dos arquivos pdb produzidos no *docking*.

Dicionário dos tipos atômicos do $A\beta_{1-42}$ da topologia CHARMM para AMBER

```
#!/bin/bash
# Ajusta os tipos atômicos apenas da proteína beta amilóide para o gaussian
reconhecê-los
dir=/diretorio/dos/pdbs/do/docking
for i in $dir
do sed -i '.bck' -e "s/HT1/H /g;
s/HT2/H /g;
s/HT3/H /g;
s/HN/H /g;
s/HSD/HID/g;
s/HB2 ASP/HB3 ASP/g;
```

s/HB1 ASP/HB2 ASP/g;
s/HB2 ASN/HB3 ASN/g;
s/HB1 ASN/HB2 ASN/g;
s/HB2 ARG/HB3 ARG/g;
s/HB1 ARG/HB2 ARG/g;
s/HG2 ARG/HG3 ARG/g;
s/HG1 ARG/HG2 ARG/g;
s/HD2 ARG/HD3 ARG/g;
s/HD1 ARG/HD2 ARG/g;
s/HB2 GLN/HB3 GLN/g;
s/HB1 GLN/HB2 GLN/g;
s/HG2 GLN/HG3 GLN/g;
s/HG1 GLN/HG2 GLN/g;
s/HB2 GLU/HB3 GLU/g;
s/HB1 GLU/HB2 GLU/g;
s/HG2 GLU/HG3 GLU/g;
s/HG1 GLU/HG2 GLU/g;
s/HA2 GLY/HA3 GLY/g;
s/HA1 GLY/HA2 GLY/g;
s/HB2 HID/HB3 HID/g;
s/HB1 HID/HB2 HID/g;
s/HG12 ILE/HG13 ILE/g;
s/HG11 ILE/HG12 ILE/g;
s/ HD1 ILE/HD11 ILE/g;
s/ HD2 ILE/HD12 ILE/g;
s/ HD3 ILE/HD13 ILE/g;
s/CD ILE/CD1 ILE/g;
s/HB2 LEU/HB3 LEU/g;
s/HB1 LEU/HB2 LEU/g;

s/HB2 LYS/HB3 LYS/g;
s/HB1 LYS/HB2 LYS/g;
s/HG2 LYS/HG3 LYS/g;
s/HG1 LYS/HG2 LYS/g;
s/HD2 LYS/HD3 LYS/g;
s/HD1 LYS/HD2 LYS/g;
s/HE2 LYS/HE3 LYS/g;
s/HE1 LYS/HE2 LYS/g;
s/HB2 MET/HB3 MET/g;
s/HB1 MET/HB2 MET/g;
s/HG2 MET/HG3 MET/g;
s/HG1 MET/HG2 MET/g;
s/HB2 PHE/HB3 PHE/g;
s/HB1 PHE/HB2 PHE/g;
s/HB2 SER/HB3 SER/g;
s/HB1 SER/HB2 SER/g;
s/HG1 SER/HG SER/g;
s/HB2 TYR/HB3 TYR/g;
s/HB1 TYR/HB2 TYR/g;
s/OT1 ALA/O ALA/g;
s/OT2 ALA/O ALA/g;
s/ A / /g;
s/0.00 AB/0.00 /g;"\$i
done

Observação: também foi feito um script igual a esse para cada ligante.

Comando pdb2oniom do Pacote TAO

Esse comando transforma o arquivo pdb em um *input* para o ONIOM.

```
#!/bin/bash
b="input"
mkdir -p ../$b
cp /home/./corelist.txt /home/./pdb2oniom/
dir=/home/./pdb2oniom
name=nome.

cd $dir
for ((i=0;i<=24;i++))
do /home/./taopackage/pdb2oniom -i $name$i.pdb -resid corelist.txt -near 5 -o
$name$i.gjf
mv $dir/$name$i.gjf $name$i.gjf.onb /home/./input
done
```

Observação: Para esse comando reconhecer o ligante é preciso gerar um arquivo `.prep`, que é próprio do pacote AmberTools.

Arquivo `.prep` do Ambertools

O comando abaixo do Ambertools gera o arquivo `prep` a partir de um arquivo `pdb` com cargas RESP (comando `bcc`), típico do AMBER, para que o Gaussian reconheça os átomos do ligante na hora de gerar o *input* do ONIOM.

```
antechamber -fi pdb -fo prep1 -i arquivo.pdb -o arquivo.prep -c bcc -pf y
```

Observação: Pode usar outras cargas. O manual do Ambertools mostra cada um dos códigos.

Procedimento após a escolha das camadas

Após usar os *scripts* descritos acima, as camadas foram escolhidas para apenas um arquivo. De forma a transformar esse arquivo em um modelo e facilitar os outros processos, que serão descritos a seguir. A escolha das camadas do ONIOM foi feita no *software* Gaussview.

Logo depois de escolher as camadas, usou-se os seguintes comandos do pacote TAO:

```
chargesum -g arquivo.gjf  
parmlook -g arquivo.log -o arquivo.txt
```

O primeiro comando calcula as cargas de cada camada e o segundo é usado para definir os parâmetros que o programa Gaussian não encontrar.

Com todos esses parâmetros definidos para um único arquivo, foi usado o *script* abaixo nos outros arquivos para trocar apenas as colunas das coordenadas, que são três, pois a tabela de conectividade e a camada alta foi a mesma dentro do contexto de cada ligante.

```
#!/bin/bash  
c="inputopt"  
mkdir -p ../$c  
dir=../input  
name=nome.  
name2=nomedois.  
  
cd $dir for ((i=1;i<=24;i++)) # não colocar o número do arquivo que usar de  
modelo
```

```
do awk 'FNR==NR{a[NR]=$3;b[NR]=$4;d[NR]=$5;next}FNR>=13FNR<=656{a[FNR];
$4=b[FNR];$5=d[FNR]}' $name$i.opt.gjf arquivo.0.gjf > $name2$i.gjf
mv $dir/$name2$i.gjf /home/./inputopt
```

Os números 3, 4 e 5 correspondem as colunas de coordenadas. Os números 13 a 656 são o intervalo de linhas que as coordenadas ocupam.

Observação: Quanto à tabela de conectividade, o próprio Gaussview gera quando o arquivo é salvo.

Comando do pacote TAO para pegar a última geometria da otimização

```
#!/bin/bash
b="inputopt2"
mkdir -p ../$b
dir=/home/./inputopt
name=nome
name2=nome2

cd $dir
for ((i=0;i<=24;i++))
do
/home/./taopackage/oniomlog -oi -t $name.$i.gjf -fo $name2.$i.gjf -i $name.$i.log
mv $dir/$name2.$i.gjf /home/./inputopt2
done
```

Exemplo do cabeçalho do input do ONIOM

```
%chk=nome.chk
%mem=3700MB
```

```
%nprocshared=9  
#p oniom(b3lyp/6-311g(d,p):amber=hardfirst) nosymm geom=connectivity  
iop(2/15=3) test opt=quadmac
```

A chave b3lyp/6-311g(d,p) é para a camada alta e a chave amber=hardfirst é para a camada baixa.

Apêndice D

Artigo publicado

Molecular modeling of cardanol-derived AChE inhibitors

Autores: Mônica de Abreu Silva , Camila D'Avila Braga Sette , Alessandra Sofia Kiametis, Luiz Antônio Soares Romeiro, Ricardo Gargano

DOI:10.1016/j.cplett.2019.07.019