

UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA - DOUTORADO

**Efeito de interações de longo alcance no sistema
macromolecular Ibuprofeno-Albumina do soro humano**

TAMIRES CRISTINA DA SILVA RIBEIRO

Maceió

2017

TAMIRES CRISTINA DA SILVA RIBEIRO

**Efeito de interações de longo alcance no sistema
macromolecular Ibuprofeno-Albumina do soro humano**

Tese de doutorado apresentado ao Instituto de Física da Universidade Federal de Alagoas como parte dos requisitos para obtenção do título de Doutora em Física.

Orientador: Prof. Dr. Marcelo Leite Lyra

Co-orientador: Prof. Dr. Vinícius Manzoni Vieira

Maceió - Alagoas - Brasil

Dezembro

2017

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central

Bibliotecária Responsável: Janaina Xisto de Barros Lima

R482e Ribeiro, Tamires Cristina da Silva.
Efeito de interações de longo alcance no sistema macromolecular Ibuprofeno-Albumina do soro humano / Tamires Cristina da Silva Ribeiro. – 2017.
172 f. : il.

Orientador: Marcelo Leite Lyra.
Coorientador: Vinícius Manzoni Vieira.
Tese (Doutorado em Física) – Universidade Federal de Alagoas. Instituto de Física, Maceió, 2017.

Bibliografia: f. 157-172.

1. Ibuprofeno. 2. Albumina do Soro Humano. 3. Interação fármaco-proteína. 4. Métodos QM/MM. 5. MFCC I. Título.

CDU: 53:547.962.3



PARECER DA BANCA EXAMINADORA DE DEFESA DE
TESE DE DOUTORADO

***“Efeito de interações de longo alcance no sistema
macromolecular Ibuprofeno-Albumina do soro humano”***


por

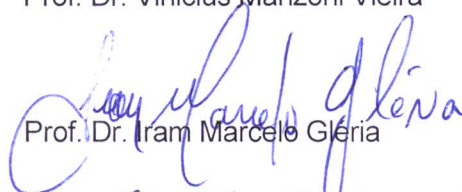
Tamires Cristina da Silva Ribeiro

A Banca Examinadora composta pelos professores Marcelo Leite Lyra (orientador), do Instituto de Física da Universidade Federal de Alagoas, Vinicius Manzoni Vieira (coorientador), do Instituto de Física da Universidade Federal de Alagoas, Iram Marcelo Gléria, do Instituto de Física da Universidade Federal de Alagoas, Alexandre Manoel de Moraes Carvalho, do Instituto de Física da Universidade Federal de Alagoas, Umberto Laino Fulco, da Universidade Federal do Rio Grande do Norte e Eudenilson Lins de Albuquerque, da Universidade Federal do Rio Grande do Norte, consideram a candidata **aprovada com grau “B”**.


Maceió, 1º de dezembro de 2017


Prof. Dr. Marcelo Leite Lyra


Prof. Dr. Vinicius Manzoni Vieira


Prof. Dr. Iram Marcelo Gléria


Prof. Dr. Alexandre Manoel de Moraes Carvalho


Prof. Dr. Umberto Laino Fulco


Prof. Dr. Eudenilson Lins de Albuquerque

Ao Supremo Criador, a Isaac e Carlos.

AGRADECIMENTOS

Primeiramente agradeço ao Criador pela minha vida e dos meus familiares.

Em segundo lugar, agradeço à minha família pela assistência, em especial meu esposo Carlos Alberto, pelo apoio incondicional, e ao meu filho Isaac por proporcionar tantos momentos especiais. Ao meu irmão Bruno, à minha mãe Noêmia Ramos e minha avó Raimunda Ramos. Enfim, a todos meus familiares que contribuíram diretamente para que eu alcançasse meus objetivos.

Aos professores Marcelo Lyra e Vinícius Manzoni, pela orientação e auxílio necessários para tornar este trabalho possível. E finalmente, à CAPES pelo suporte financeiro fornecido por meio das bolsas concedidas.

Tudo posso Naquele que me fortalece.

Apóstolo Paulo

Resumo

A interação entre fármacos e proteínas tem sido amplamente estudada desde a década de 1970, quando a cristalização dos primeiros complexos macromoleculares permitiu a investigação, em nível molecular, dos tipos de forças envolvidas nestes sistemas. A determinação de estruturas tridimensionais por meio de técnicas experimentais de raios X e espectroscopia de RNM proporcionaram um significativo avanço no estudo da função e estrutura de proteínas, bem como da interação proteína-ligante, auxiliando de maneira decisiva no design de novas drogas e no desenvolvimento da modelagem molecular. Neste escopo, diversas aproximações computacionais têm se mostrado importantes na otimização do processo de concepção de novos fármacos. Desta maneira, muitas destas metodologias computacionais estão empenhadas em prever detalhadamente as interações específicas entre proteína e ligante. Dentre estas, destacam-se as abordagens baseadas em Mecânica Molecular (MM) e Mecânica Quântica (MQ), que têm sido muito utilizadas para uma descrição mais acurada das interações envolvidas no reconhecimento molecular entre o alvo receptor e a molécula ligante. Neste trabalho, serão calculadas as energias de interação entre a molécula do Anti-Inflamatório Não-Esteróide (AINE) Ibuprofeno (IBU) com a proteína Albumina do Soro Humano (ASH) através dos métodos quânticos Hartree-Fock (HF), Teoria do Funcional da Densidade (DFT) e o Método Perturbativo de Møller-Plesset de segunda ordem (MP2), onde são estudados os efeitos das interações de longo alcance, como as interações eletrostáticas, que foram inseridas nos cálculos como um fundo de cargas pontuais remanescentes e por meio do ajuste das interações Coulombianas proporcionadas pelo funcional híbrido CAM-B3LYP.

Palavras-chave: Ibuprofeno, Albumina do Soro Humano, Interação fármaco-proteína, MFCC, Métodos QM/MM.

Abstract

The interaction between proteins and ligands has been extensively studied since the 1970s when the crystallization of the first macromolecular complexes allowed molecular investigation of the types of forces involved in these systems. The determination of three-dimensional structures by means of experimental X-ray techniques and NMR spectroscopy provided a significant advance in the study of protein function and their structure, as well as protein-binding interaction, contributing to the development for design of new drugs and for molecular modelling. In this scope, several computational approaches have been shown to be important in the optimization of the design process of new drugs. In this way, many of these computational methodologies are engaged in predicting in detail the specific interactions between protein and ligand. Among these, we highlight the approaches based on Molecular Mechanics (MM) and Quantum Mechanics (MQ), which have been widely used for a more accurate description of the interactions involved in molecular recognition between the receptor target and the binding molecule. In this work, the interaction energies between the molecule of the Nonsteroidal anti-inflammatory Drug (NSAID) Ibuprofen (IBU) and the Human Serum Albumin Protein (HSA) (HF), Density Functional Theory (DFT) and second-order Møller-Plesset Perturbative Method (MP2), where the effects of long-term interactions are studied, such as the electrostatic interactions, which were inserted into the calculations as a background of remaining point charges and for using the Coulomb-attenuating provided by the CAM-B3LYP hybrid functional.

Keywords: Ibuprofen, Human Serum Albumin, Protein–Ligand Interactions, MFCC, QM/MM Methods.

Lista de Figuras

1.1	A estrutura molecular dos aminoácidos consiste em um carbono central C_{α} ligado a um hidrogênio, um grupo funcional carboxila, a um grupo amina e uma cadeia lateral R. Ressaltando-se que os grupos carboxila e amina são ionizados em meios com pH fisiológico.	4
1.2	A figura representa os diferentes níveis estruturais de uma proteína. A estrutura primária consiste na sequência de aminoácidos ligados covalentemente via ligações peptídicas. A Estrutura secundária configura cadeias polipeptídicas dispostas repetidamente, formando assim estruturas protéicas características, como por exemplo uma Alfa-hélice, que por sua vez pode integrar a estrutura terciária de um polipeptídeo dobrado. As várias subunidades de polipeptídeos dobrados formam uma estrutura quaternária, como por exemplo o arranjo protéico da Hemoglobina.	6
1.3	Representação do sítio ativo de uma macromolécula.	8
1.4	Modelos de interação do substrato com o sítio ativo.	9
1.5	Ligação iônica entre o íon carboxilato do resíduo ácido aspártico e o íon amônio da lisina.	10
1.6	Representação de uma ponte de hidrogênio. O átomo doador (denotado por D) está ligado ao hidrogênio que é fornecido para o átomo aceitador (denotado por A).	12
1.7	Interação de van der Waals entre duas Fenilalaninas (PHE), que são dois resíduos aromáticos e hidrofóbicos.	13

1.8	A energia de interação de van der Waals como uma função da distância entre o núcleo e um dímero de argônio. Nota-se claramente a atração eletrostática (interação de longo alcance) e a interação repulsiva devido ao princípio de exclusão de Pauling (interação de curto alcance).	15
3.1	Estrutura da proteína HSA composta por três domínios, subdivididos em subdomínios Ia (Amarelo), Ib (verde), IIa (vermelho), IIb (magenta), IIIa (azul) e IIIb (ciano).	58
3.2	Sítios de Sudlow I e II, os quais encontram-se nos subdomínios IIa e IIIa (também denominados FA7 e FA3/FA4, respectivamente). Na figura são mostrados os principais sítios da ASH revelados pelo estudo cristalográfico da ASH complexada com ácidos graxos de cadeia longa e média. . . .	59
3.3	Estrutura molecular do ligante Ibuprofeno (IBU). A estrutura molecular é composta por um grupo químico carboxílico na região do propionato, onde ocorre a presença do carbono quiral do composto (na posição- α). A outra porção da molécula é constituída pelo anel aromático do benzeno juntamente com o grupo metil propil.	61
3.4	Estrutura cristalina do IBU complexado com a proteína ASH. Os dados cristalográficos, encontrados no arquivo PDB identificado por 2BXG, foram obtidos através de difração de raios x e possuem uma resolução de 2,7 Å.	66
3.5	Nesta figura é possível visualizar a região do sítio ativo FA3/FA4 (área sombreada) e o raio de 13 angstroms (Å), traçado em torno do centro do fármaco, que delimita a região que contém os resíduos inclusos no cálculo das energias de interação.	67
3.6	Estrutura molecular do ligante Ibuprofeno (IBU). A fim de facilitar a análise das interações da molécula com os resíduos da proteína, o IBU é dividido nas regiões i, ii e iii. A região i compreende o grupo carboxílico, enquanto que a região ii contempla o anel aromático e a região iii os 4 carbonos do grupo metilpropil ligado ao benzeno da molécula.	69
3.7	Painel gráfico BIRD obtido para o método quântico HF. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas).	86

- 3.8 Painel gráfico BIRD obtido para o método quântico HF na presença de cargas pontuais. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas). 87
- 3.9 Painel gráfico BIRD obtido para o método quântico DFT. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas). 88
- 3.10 Painel gráfico BIRD obtido para o método quântico DFT na presença de cargas pontuais. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas). 89
- 3.11 Painel gráfico BIRD obtido para o método quântico DFT incluindo correção de longo alcance através do funcional CAM-B3LYP. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas). 90
- 3.12 Painel gráfico BIRD obtido para o método quântico DFT incluindo correção de longo alcance através do funcional CAM-B3LYP e na presença de cargas pontuais. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas). 91
- 3.13 Comparação em relação à inclusão de cargas pontuais para as energias calculadas através do método HF com conjunto de base 6-311+G(d). As barras horizontais pretas descrevem o caso em que não são consideradas as cargas pontuais, à medida que as barras horizontais cinzas indicam que a inclusão do background de cargas nos cálculos da energia de interação IBU-HSA. 92
- 3.14 Comparação em relação à inclusão de cargas pontuais para as energias calculadas através do método DFT com conjunto de base 6-311+G(d). As barras horizontais pretas descrevem o caso em que não são consideradas as cargas pontuais, à medida que as barras horizontais cinzas indicam que a inclusão do background de cargas nos cálculos da energia de interação IBU-HSA. 93

3.15	Comparação em relação à inclusão de cargas pontuais para as energias calculadas através do método DFT, com funcional para interações de longo alcance CAM-B3LYP, e conjunto de base 6-311+G(d). As barras horizontais pretas descrevem o caso em que não são consideradas as cargas pontuais, à medida que as barras horizontais cinzas indicam que a inclusão do background de cargas nos cálculos da energia de interação IBU-HSA.	94
3.16	O painel gráfico BIRD (com 17 aminoácidos) obtido para as energias calculadas com conjunto de base 6-311+G(d), sem cargas pontuais, para os métodos HF (barras horizontais cinzas com listras diagonais cruzadas), MP2 (barras horizontais cinzas com listras diagonais), CAM (barras horizontais cinzas) e B3LYP (barras horizontais pretas).	95
3.17	O painel gráfico BIRD (com 12 aminoácidos) obtido para as energias calculadas com conjunto de base 6-311+G(d), sem cargas pontuais, para os métodos HF (barras horizontais cinzas com listras diagonais cruzadas), MP2 (barras horizontais cinzas com listras diagonais), CAM (barras horizontais cinzas) e B3LYP (barras horizontais pretas).	96
3.18	O painel gráfico BIRD obtido para as energias calculadas com conjunto de base 6-311+G(d), na presença de cargas pontuais, para os métodos HF (barras horizontais cinzas com listras diagonais cruzadas), MP2 (barras horizontais cinzas com listras diagonais), CAM (barras horizontais cinzas) e B3LYP (barras horizontais pretas).	97
3.19	Disposição estrutural de 4 dos 13 aminoácidos mais importantes nas interações do complexo IBU-HSA.	99
3.20	Disposição estrutural de 4 dos 13 aminoácidos mais importantes nas interações do complexo IBU-HSA.	100
3.21	Disposição estrutural de 5 dos 13 aminoácidos mais importantes nas interações do complexo IBU-HSA.	101

- 3.22 Desvio Absoluto tomando como referência o uso do conjunto de base 6-311+G(d). A maioria dos desvios flutua em valores muito próximos de zero, onde somente os resíduos LEU387, PH403, ARG445 e ASN483 obtiveram DA superior a 1kcal/mol (indicado pela linha tracejada vermelha). O eixo das abscissas dispõe as mínimas distâncias em ordem crescente. Os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias. 104
- 3.23 Desvio Relativo tomando como referência o uso do conjunto de base 6-311+G(d). Excetuando os resíduos CYS438, LEU453, LEU387, ALA490, SER454 e ALA406 que mostram desvios maiores que 1, a maioria apresenta DR abaixo desse valor. Os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias. 105
- 3.24 Desvio Absoluto tomando como referência o uso do funcional CAM-B3LYP. Para melhor visualização da região delimitada pelo retângulo com borda preta, o gráfico ao lado direito se trata do aumento da referida área do gráfico. Os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias. 107
- 3.25 Desvio Relativo tomando como referência o uso do funcional CAM-B3LYP. Assim como no gráfico do DA, área delimitada pelo retângulo está expandida ao lado direito. Os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias. 108
- 3.26 Desvio Absoluto tomando como referência os métodos onde as cargas pontuais foram inclusas. Os gráficos A, B e C correspondem às expansões das regiões delimitadas pelo retângulos com borda preta A, B e C. Como os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias, o gráfico A é compatível com regiões próximas do sítio ativo e concentra a maior parte dos resíduos com DA maior que 1 kcal/mol (indicado pela linha tracejada vermelha). 110
- 3.27 Desvio Relativo tomando como referência os métodos onde as cargas pontuais foram inclusas. As energias calculadas via DFT sofreram os maiores desvios em relação à presença das cargas pontuais. 111

- 3.28 Desvio Absoluto tomando como referência as energias calculadas pelo método HF. Visto que os resíduos estão dispostos em ordem crescente, nota-se que muitos resíduos com DA maior que 1 kcal/mol (indicado pela linha tracejada vermelha) se concentram nas regiões próximas do sítio ativo. A maioria destes desvios se referem às energias calculadas com MP2. 113
- 3.29 Desvio Relativo tomando como referência as energias calculadas pelo método HF. Assim como em casos anteriores, as regiões próximas do sítio ativo concentram a maior quantidade de aminoácidos cujos desvios ultrapassam 1 (indicado pela linha tracejada vermelha). A maioria destes resíduos tiveram suas energias calculadas via MP2. 114
- 3.30 Desvio Absoluto tomando como referência as energias calculadas pelo método MP2. Assim como em casos anteriores, as regiões próximas do sítio ativo concentram a maior quantidade de aminoácidos cujos desvios ultrapassam 1 kcal/mol (indicado pela linha tracejada vermelha). A maioria dos resíduos com DA significativa tiveram suas energias calculadas via DFT. 116
- 3.31 Desvio Relativo tomando como referência as energias calculadas pelo método MP2. Assim como em casos anteriores, as regiões próximas do sítio ativo concentram a maior quantidade de aminoácidos cujos desvios relativos ultrapassam 1 (indicado pela linha tracejada vermelha). Os métodos são proporcionais nos quesitos intensidade dos DRs e quantidade de resíduos para os quais o DR seja extenso. 117
- 3.32 Convergência da energia em função da distância radial do resíduo em torno do ligante. Cada curva, denominada CBPR, representa um método empregado no cálculo das energias de interação entre o IBU e a ASH. Aqui são comparadas as energias obtidas pelos métodos HF, DFT, DFT (com funcional CAM-B3LYP) e MP2, ao método DFT aliado à simulação em meio contínuo PCM, com constantes dielétricas $\epsilon = 0$ (curva na cor ciano) e $\epsilon = 40$ (curva no cor magenta) [85]. 120

- 3.33 Convergência da energia em função da distância radial do resíduo em torno do ligante. Cada curva, denominada CBPR, representa um método empregado no cálculo das energias de interação entre o IBU e a ASH. Aqui são comparadas as energias obtidas pelos métodos, na presença das cargas pontuais; HF, DFT e DFT (com funcional CAM-B3LYP), ao método DFT aliado à simulação em meio contínuo PCM, com constantes dielétricas $\varepsilon = 0$ (curva na cor ciano) e $\varepsilon = 40$ (curva no cor magenta) [85]. 121

Lista de Tabelas

1.1	Os 20 aminoácidos (AA) normalmente encontrados como resíduos em proteínas. Outros menos comuns, decorrentes de modificações em aminoácidos comuns após a síntese protéica, também ocorrem em proteínas.	6
3.1	Nesta tabela, temos a relação de quais métodos quânticos foram empregados no cálculo da energia de interação resíduo-ligante, assim como quais conjuntos de base e em quais casos foi incluído o background de cargas pontuais.	64
3.2	A tabela lista os aminoácidos do sítio ativo (marcados com *) e circunvizinhança para os quais foram calculadas as energias de interação com o ligante IBU01. No total, foram escolhidos 73 resíduos, distribuídos num raio de 13 Å em torno do centróide do ligante.	68
3.3	A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.	71
3.4	A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.	72

- 3.5 A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 73
- 3.6 A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 74
- 3.7 A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 75
- 3.8 A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 76
- 3.9 A tabela mostra as energias obtidas pelos métodos HF, B3LYP, CAM e MP2 utilizando o conjunto de bases 6-311+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 77

- 3.10 A tabela mostra as energias obtidas pelos métodos HF, B3LYP, CAM e MP2 utilizando o conjunto de bases 6-311+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 78
- 3.11 A tabela mostra as energias obtidas pelos métodos HF, B3LYP, CAM e MP2 utilizando o conjunto de bases 6-311+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 79
- 3.12 A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-311+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 80
- 3.13 A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-311+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 81
- 3.14 A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-311+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP. 82
- 3.15 Esta tabela mostra a média (em porcentagem) dos desvios relativos da energia, considerando-se a mudança de base, isto é, quando com as energias calculadas via base 6-311+G(d) são definidas como os valores de referência. 103

-
- 3.16 A tabela mostra a média (em porcentagem) dos desvios relativos da energia, considerando-se como os valores de referência as energias calculadas com a inclusão das interações de longo alcance via funcional CAM. 107
- 3.17 A tabela mostra a média (em porcentagem) dos DRs em relação às energias obtidas na presença de cargas pontuais. 110
- 3.18 A tabela mostra as médias dos desvios relativos (DR) em relação ao método HF. As médias mais altas correspondem à comparação entre as energias obtidas pelos métodos HF e MP2. Outrossim, a média dos DRs são menores para as energias calculadas com métodos que incluem a contribuição eletrostática das cargas pontuais. 112
- 3.19 A tabela mostra os valores das médias do DR das energias em relação o método MP2. 115

SUMÁRIO

Folha de rosto	i
Resumo	ix
Abstract	xi
Lista de Figuras	xiii
Lista de Tabelas	xx
1 Introdução	1
1.1 Proteínas	3
1.1.1 Sítio de ligação	7
1.1.2 Interações iônicas (ou eletrostáticas)	10
1.1.3 Interações de hidrogênio	11
1.1.4 Interações de Van der Waals	13
1.2 Estudo de proteínas via Métodos de Modelagem Molecular	17
1.3 Estudo de proteínas através de métodos baseados em Mecânica Quântica	20
2 Métodos Quânticos	23
2.1 O método Hartree-Fock	25

2.1.1	A equação canônica de Hartree-Fock	27
2.2	O Método Perturbativo de Møller-Plesset (MPPT)	34
2.2.1	A correção de Segunda Ordem	37
2.3	A Teoria do Funcional de Densidade (DFT)	38
2.3.1	Teoremas de Hohenberg e Kohn	40
2.3.2	Funcionais de Correlação e Troca	47
2.4	O método MFCC (Molecular Fractionation with Conjugate Caps)	51
2.4.1	Estratégias para o refinamento do MFCC	52
2.4.2	Aproximação MFCC-CPCM	53
2.4.3	Aproximação EE-GMFCC	55
2.4.4	Aproximação EE-GMFCC-CPCM	56
2.4.5	Aproximação EE-AMFCC	57
3	Aplicação: Interação do Ibuprofeno (IBU) com a Albumina do Soro Humano	55
3.1	A proteína Albumina do Soro Humano (ASH)	56
3.1.1	Características estruturais da ASH	56
3.2	O ligante Ibuprofeno (IBU)	60
3.3	Detalhes da metodologia	62
3.4	Resultados	69
3.5	Gráficos BIRDs (Binding site, Interaction energy and Residues Domain)	83
3.5.1	Convergência de Bases	83
3.5.2	O efeito das cargas pontuais nas energias de interação	84

3.5.3	Comparação entre os métodos	85
3.6	Características estruturais das interações IBU-HSA	98
3.7	Desvio Absoluto (DA) e Relativo (DR)	101
3.7.1	Desvios Absoluto (DA) e Relativo (DR) em relação à base 6-311+G(d)	102
3.7.2	Desvios Absoluto (DA) e Relativo (DR) em relação ao uso do funcional CAM-B3LYP	105
3.7.3	Desvios Absoluto (DA) e Relativo (DR) em relação à inclusão das cargas pontuais	108
3.7.4	Desvios Absoluto (DA) e Relativo (DR) em relação ao método HF	111
3.7.5	Desvios Absoluto (DA) e Relativo (DR) em relação ao método MP2	114
3.8	As Curvas CBPR (Converged Binding Pocket Radius)	117
4	Conclusão	121
	Referências Bibliográficas	124

Capítulo 1

Introdução

A elucidação da estrutura molecular e atômica de macromoléculas biológicas, assim como de seu funcionamento e desempenho nos mais variados tipos de organismos, é crucial para a compreensão dos vários processos de vida básicos, tais como os mecanismos de reação de uma enzima fármaco-conversora, a transdução de sinal de uma proteína para outra, os modos de ação de uma droga e as consequências de uma mutação sobre uma enzima [1].

Os estudos de sistemas biológicos a nível microscópico, como por exemplo a análise detalhada de uma molécula individual e sua interação com pequenas estruturas que formam organismos vivos complexos, incorporam conhecimentos de biologia molecular, bioquímica e biofísica [2]. Por outro lado, o avanço desta área se deu principalmente a partir do desenvolvimento da modelagem computacional, por meio da elaboração de modelos tridimensionais de moléculas biológicas via computação gráfica sofisticada, assim como por meio de informações estruturais detalhadas fornecidas por cristalografia de raios-X e espectroscopia de ressonância magnética nuclear [1, 3].

As proteínas são moléculas complexas e é esta mesma complexidade que transmite sua atividade biológica. Deste modo, o estudo de proteínas não é um trabalho trivial, embora sua função e estrutura estejam fortemente relacionadas. Assim sendo, existe uma grande variedade de métodos

dedicados ao estudo de proteínas. Sejam no sentido de compreender a relação entre estrutura e atividade, ou no entendimento das interações enredadas no processamento da ação biológica destas macromoléculas, são diversas as técnicas experimentais e teóricas que se complementam e buscam desenvolver e ampliar o conhecimento sobre estas estruturas moleculares fundamentais para a desenvolvimento, regulação e funcionamento de organismos vivos [3].

Dentro do amplo espectro de interesses que envolvem o estudo de proteínas, o entendimento de como as drogas interagem com estas macromoléculas é um dos principais objetivos da interdisciplinar química medicinal, além disso a compreensão dos mecanismos envolvidos na interação ligante-proteína possui um papel essencial na descoberta e planejamento de novas drogas [4, 5]. O design racional de fármacos integra uma importante área baseada no desenvolvimento de drogas a partir do conhecimento dos alvos moleculares biologicamente ativos [6]. Neste ensejo, a modelagem molecular computacional e as técnicas experimentais são duas correntes paralelas que contribuem para o impulso deste ramo, pois enquanto a modelagem molecular (*In silico*) se baseia principalmente em estimativas e probabilidades, a biologia experimental (*In vitro* e *In vivo*) é considerada mais precisa e confiável. No entanto, ambas atuam em uma espécie de simbiose, uma vez que informações geradas experimentalmente podem aperfeiçoar a capacidade de cálculo e estimativa da modelagem molecular, por sua vez, esta última pode revelar problemas de resultados experimentais e auxiliar no entendimento dos mesmos [7].

As técnicas experimentais são indispensáveis e mantêm sua importância em todas as fases da concepção de novos fármacos, porém são dispendiosas e carecem de longos períodos para alcançar resultados bem sucedidos [8]. Em contrapartida, o uso da modelagem molecular, ou ainda, a utilização de simulações *In Silico* auxilia na investigação das interações entre um dado ligante com o seu alvo receptor, permitindo a exploração dos fatores estruturais relacionados ao efeito biológico, de forma rápida e satisfatória. O emprego de ferramentas computacionais, principalmente nas etapas iniciais do design racional de drogas, possibilitam uma rápida análise da descrição da geometria e das propriedades dos fármacos, viabilizando o entendimento das bases químicas e moleculares envolvidas na interação fármaco-receptor. A alta disponibilidade de bancos de dados e de softwa-

res de química computacional são ferramentas poderosas na geração de informações, facilitando assim o emprego e utilização de técnicas de modelagem computacional na pesquisa de fármacos e proteínas [2, 9].

Utilizando uma metodologia baseada em modelagem computacional, este estudo pretende investigar a influência das correções de longo alcance na energia de interação entre o fármaco Ibuprofeno (IBU) e a proteína Albumina do Soro Humano (ASH). Por meio de métodos de mecânica molecular e mecânica quântica (MM/MQ), foram calculadas as energias de interação do IBU com a região bioativa da ASH, onde foram inseridas as contribuições das forças de longo alcance, assim como foram averiguados os efeitos decorrentes da presença de cargas pontuais na energia entre ligante e receptor. Neste capítulo serão introduzidas as características gerais de proteínas, quais os métodos tradicionais são mais utilizados no estudo da sua estrutura, atividade e interação com drogas. No segundo capítulo serão apresentados os modelos teóricos de mecânica molecular e mecânica quântica aplicados ao estudo da interação entre fármacos e macromoléculas, os quais foram usados nesta investigação. Já o terceiro capítulo mostrará a aplicação computacional dos métodos teóricos expostos no capítulo anterior, no sistema de interesse IBU-ASH, apresentando detalhadamente os métodos e os resultados obtidos mediante a metodologia proposta. Finalmente, no quarto capítulo será apresentada a conclusão e suas devidas discussões acerca dos resultados decorrentes do referido estudo.

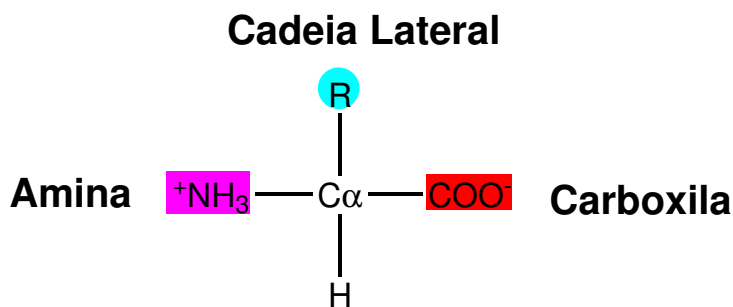
1.1 Proteínas

Devido a importância das várias funções que desempenham nos diversos tipos de organismos, as proteínas são as moléculas cujas formas são mais exploradas, visto que, além do esclarecimento do funcionamento e estrutura dos seres vivos, o conhecimento sobre estas macromoléculas pode fornecer a chave para o desenvolvimento de novos medicamentos, materiais e procedimentos diagnósticos. Além disso, o estudo da estrutura e dinâmica molecular de macromoléculas biológicas, especialmente de proteínas e ácidos nucleicos, viabiliza a caracterização da base molecular de

muitas doenças, podendo também fornecer terapias específicas para estas [2].

Proteínas são grandes sistemas macromoleculares formados por aminoácidos que unem-se covalentemente por meio de ligações peptídicas. Tipicamente, as proteínas contêm entre 50 e 2000 aminoácidos, de modo que cada proteína possui uma sequência particular de aminoácidos. Dito isto, a estrutura básica de cada aminoácido é composta por um átomo de carbono central (carbono α) ligado a um grupo carboxílico (COO^-), um grupo amínico (NH_3^+), um átomo de hidrogênio (H) e uma cadeia lateral característica (denotado por R), onde as diferentes propriedades químicas das cadeias laterais designam o papel de cada aminoácido na estrutura e função das proteínas [10]. A figura 1.1 mostra a estrutura molecular geral dos aminoácidos, notando-se que em meios com pH fisiológico, o funcional carboxílico e o grupo amínico são ionizados.

Figura 1.1: A estrutura molecular dos aminoácidos consiste em um carbono central C_α ligado a um hidrogênio, um grupo funcional carboxila, a um grupo amina e uma cadeia lateral R. Ressaltando-se que os grupos carboxila e amina são ionizados em meios com pH fisiológico.



Fonte: Cooper, 2007

A tabela 1.1 lista os 20 principais aminoácidos, os quais podem ser classificados de acordo com as propriedades de suas respectivas cadeias laterais. Desta forma, são considerados apolares os aminoácidos Glicina (GLY), Alanina (ALA), Valina (VAL), Leucina (LEU), Isoleucina (ILE), Prolina (PRO), Cisteína (CYS), Metionina (MET), Fenilalanina (PHE) e Triptofano (TRP). Tais

aminoácidos possuem cadeias laterais hidrofóbicas e carga líquida nula. Por outro lado, os aminoácidos Serina (SER), Treonina (THR), Tirosina (TYR), Asparagina (ASN) e Glutamina (GLN) apresentam cadeias laterais com carga líquida nula, no entanto são polares uma vez que possuem grupos químicos suscetíveis para a formação de pontes de hidrogênio, como a carboxila e a amida. No caso dos aminoácidos com cadeias laterais carregadas, a Lisina (LYS), Arginina (ARG) e a Histidina (HIS) são aminoácidos básicos cujas cadeias laterais detêm carga líquida positiva. No caso do Ácido Aspártico (ASP) e do Ácido Glutâmico (GLU), também chamados de Aspartato e Glutamato respectivamente, suas cadeias laterais são negativamente carregadas. Os aminoácidos carregados positivamente ou negativamente são hidrofílicos e tendem a se localizar na superfície da proteína, enquanto que os aminoácidos hidrofóbicos, como os de cadeia lateral apolar, geralmente encontram-se no interior protéico [8, 10].

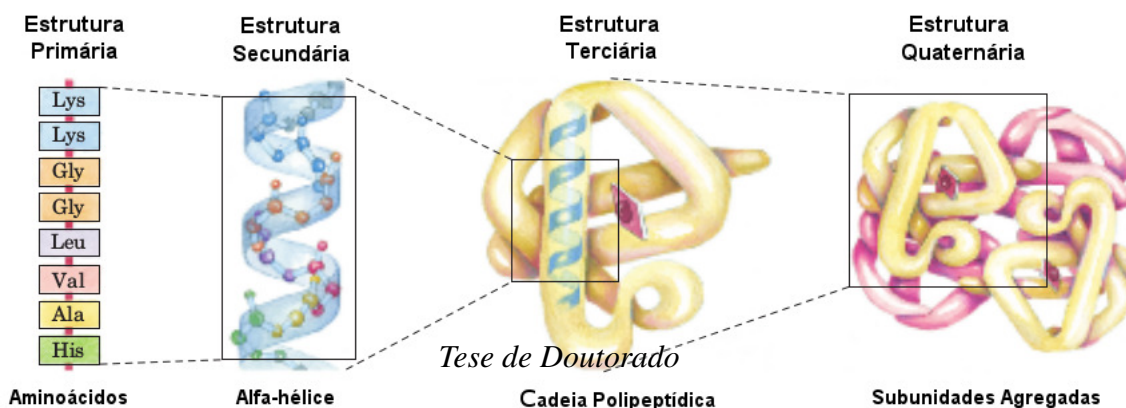
Fonte: Patrick, 1995

A estrutura protéica pode organizar-se em quatro níveis diferentes: Primária, Secundária, Terciária e Quaternária. A estrutura primária consiste no nível mais simples de conformação da proteína, visto que é a sequência de resíduos aminoácidos unidos entre si por ligações covalentes (incluindo ligações dissulfídicas) em uma cadeia polipeptídica. Estruturas secundárias referem-se a arranjos estáveis de aminoácidos que formam padrões estruturais recorrentes, como no caso das estruturas de Alfa-Hélice e Folha-Beta. Já as estruturas terciárias se caracterizam pelo dobramento tridimensional de um polipeptídeo. O arranjo espacial de duas ou mais subunidades polipeptídicas configuram uma estrutura quaternária [11]. A figura 1.2 mostra os quatro diferentes níveis organizacionais de uma proteína, onde a Hemoglobina é utilizada como modelo para exemplificá-los.

Tabela 1.1: Os 20 aminoácidos (AA) normalmente encontrados como resíduos em proteínas. Outros menos comuns, decorrentes de modificações em aminoácidos comuns após a síntese protéica, também ocorrem em proteínas.

Aminoácidos sintetizados no corpo humano	Carga líquida da cadeia lateral (Polaridade)
Alanina (ALA)	neutra (apolar)
Arginina (ARG)	positiva
Asparagina (ASN)	neutra (polar)
Ácido Aspártico (ASP)	negativa
Cisteína (CYS)	neutra (apolar)
Ácido Glutâmico (GLU)	negativa
Glutamina (GLN)	neutra (polar)
Glicina (GLY)	neutra (apolar)
Prolina (PRO)	neutra (apolar)
Serina (SER)	neutra (polar)
Tirosina (TYR)	neutra (polar)
Aminoácidos essenciais à dieta	
Histidina (HIS)	positiva
Isoleucina (ILE)	neutra (apolar)
Leucina (LEU)	neutra (apolar)
Lisina (LYS)	positiva
Metionina (MET)	neutra (apolar)
Fenilalanina (PHE)	neutra (apolar)
Treonina (THR)	neutra (polar)
Triptofano (TRP)	neutra (apolar)
Valina (VAL)	neutra (apolar)

Figura 1.2: A figura representa os diferentes níveis estruturais de uma proteína. A estrutura primária consiste na sequência de aminoácidos ligados covalentemente via ligações peptídicas. A Estrutura secundária configura cadeias polipeptídicas dispostas repetidamente, formando assim estruturas protéicas características, como por exemplo uma Alfa-hélice, que por sua vez pode integrar a estrutura terciária de um polipeptídeo dobrado. As várias subunidades de polipeptídeos dobrados formam uma estrutura quaternária, como por exemplo o arranjo protéico da Hemoglobina.



De um modo geral, o entendimento da estrutura da proteína precede ao entendimento do funcionamento dos processos envolvidos na sua interação com os fármacos. Assim sendo, a conformação dos níveis de estrutura, principalmente da estrutura terciária em uma proteína pode definir sua função e interação com drogas. Os dobramentos ocorrem automaticamente e dependem da forma como as proteínas são sintetizadas nas células. No entanto existem outros fatores determinantes da forma que uma estrutura protéica assumirá. Dentre eles, pode-se citar as forças envolvidas entre as cadeias laterais dos resíduos que constituem uma cadeia peptídica. Os resíduos interagem entre si, atraindo ou repelindo um ao outro, fazendo com que a proteína atinja a conformação mais favorável, isto é, a estrutura terciária [8].

As ligações intramoleculares envolvidas entre os resíduos desempenham importante papel no que diz respeito à formação e estabilidade da estrutura protéica. Sendo assim, os tipos de ligação que se destacam nestes processos são as de natureza atrativa, tais como: as ligações covalentes, as ligações iônicas e as ligações de hidrogênio. Em compensação, as ligações de Van der Waals e as interações hidrofóbicas também são relevantes na formação da estrutura e função das proteínas.

1.1.1 Sítio de ligação

A interação, ou ligação, de uma molécula farmacológica com seu alvo macromolecular ocorre em regiões específicas conhecidas como sítio de ligação (figura 1.3). Desta forma, os aminoácidos presentes no sítio ativo desempenham um importante papel na função da macromolécula, de modo que tais aminoácidos podem estar diretamente envolvidos na ligação do substrato com o sítio ativo, ou estar envolvido no mecanismo de reação.

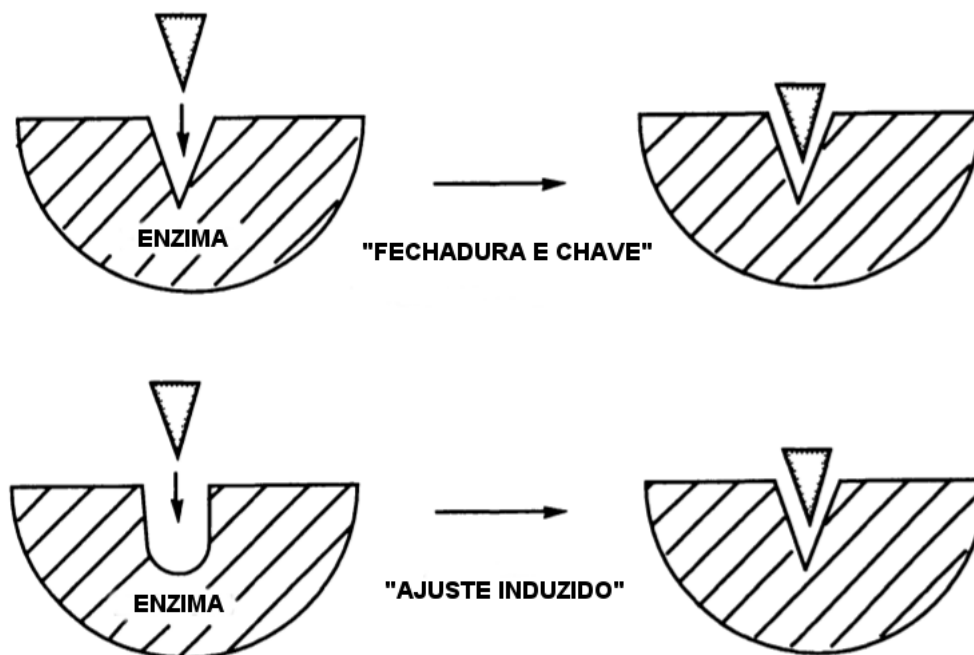
Figura 1.3: Representação do sítio ativo de uma macromolécula.



Fonte: Patrick, 1995

Era usual comparar a ligação da droga no sítio ativo com o encaixe perfeito de uma chave (que seria o substrato) em sua fechadura (o alvo macromolecular), como podemos ver na figura 1.4. Porém existem casos em que a mesma proteína pode ligar-se a diferentes substratos, com maior ou menor eficiência. Logo, nestes cenários não cabe a analogia da chave e fechadura. Assim, a Teoria do Ajuste Induzido de Koshland assume que a ligação do substrato no sítio ativo não se dá a partir de um encaixe perfeito, mas ao entrar na região ativa ocorre um processo de ajuste do sítio em relação ao ligante, como mostra de modo (figura 1.4). Todavia, o substrato deve induzir a sítio de ligação ao melhor ajuste, não necessariamente a forma mais estável, de maneira que o sítio de ligação não tenha sua forma excessivamente modificada, preservando assim o mecanismo de reação [8].

Figura 1.4: Modelos de interação do substrato com o sítio ativo.



Fonte: Patrick, 1995

É comum os sítios de ligação estarem localizados próximos ou na própria superfície das proteínas. Por esse motivo, as forças iônicas assumem importante papel na ligação do substrato com o sítio ativo. Ademais, as forças intermoleculares envolvidas na ligação do substrato com a enzima são as mesmas forças intramoleculares que controlam a formação e estabilidade da estrutura protéica. Assim, embora algumas drogas reajam de modo irreversível ligando-se covalentemente¹ com o sítio ativo de uma macromolécula, na maioria dos casos o substrato se liga ao sítio ativo por meio de interações intermoleculares menos fortes, como por exemplo as ligações iônicas (ou eletrostáticas), as interações de Van der Waals, as pontes de hidrogênio e as interações repulsivas. Portanto, vamos agora considerar os possíveis tipos de interação intermolecular que ocorrem no

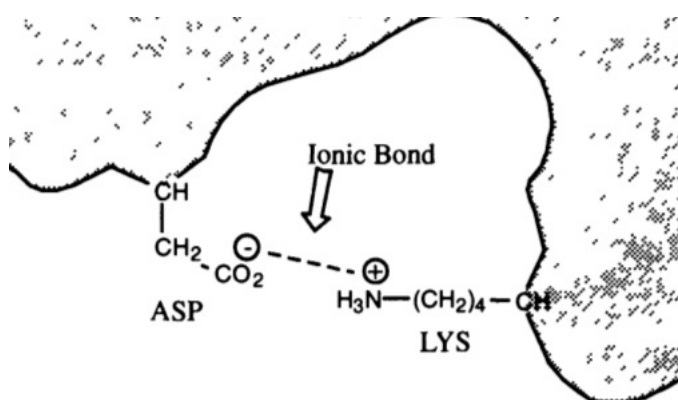
¹As ligações covalentes apresentam altas energias com cerca de 47,80 a 95,60 kcal/mol

processo de ligação de uma droga com a região bioativa de um alvo macromolecular.

1.1.2 Interações iônicas (ou eletrostáticas)

As interações eletrostáticas ocorrem entre grupos atômicos eletricamente carregados e envolvem forças atrativas e repulsivas. A natureza destas interações depende dos sinais das cargas elétricas dos átomos, isto é, para cargas de mesmo sinal ocorrerá interação atrativa, enquanto que para cargas opostas haverá repulsão elétrica. De um modo geral, dois tipos de força eletrostática são dominantes em proteínas. São elas, a interação iônica e as interações de hidrogênio [12].

Figura 1.5: Ligação iônica entre o íon carboxilato do resíduo ácido aspártico e o íon amônio da lisina.



Fonte: Patrick, 1995

A interação iônica (1.5) envolve a atração de íons ou moléculas com cargas permanentemente completas de sinais opostos e é a mais forte entre as forças intermoleculares, com energia em torno de 4,78 a 9,56 kcal/mol. Este tipo de interação é muito comum em macromoléculas e desempenham funções como estabilizar a estrutura de proteínas e sua interação com outras moléculas, além de direcionar as catálises enzimáticas [13]. Este tipo de interação é descrita pela lei de Coulomb para

duas cargas pontuais em um meio uniforme, assim, a energia desta interação é inversamente proporcional à distância entre os dois átomos carregados. Portanto as interações iônicas (ou eletrostáticas) são consideradas de longo alcance.

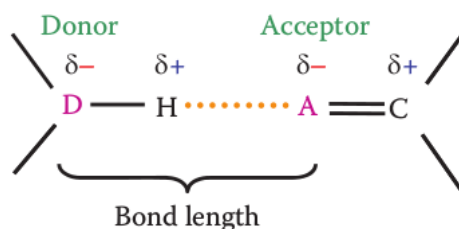
De acordo com a lei de Coulomb, esta interação possui também a dependência com a constante dielétrica, isto é, o meio é um aspecto relevante na magnitude da mesma. A constante dielétrica representa a habilidade de mascarar a interação eletrostática, destarte no meio cuja constante dielétrica seja alta, deve ocorrer uma diminuição na energia de interação entre as cargas. No caso das regiões internas da proteína, tais como suas regiões bioativas, a constante dielétrica varia, sendo necessária a inclusão da energia de polarização do meio protéico. Já no caso dos sítios de ligação das proteínas, os quais são hidrofóbicos, o efeito das interações iônicas são mais evidentes [14].

A parte eletrostática da energia de interação intermolecular total é bastante proeminente. Assim, muitos programas de mecânica molecular incluem o termo energético correspondente, os quais dependem das cargas parciais atômicas. Deste modo, a energia eletrostática de macromoléculas pode ser calculada por meio de diversas aproximações, tais como procedimentos topológicos, ou através de métodos de funções de onda na MQ, onde utiliza-se a análise populacional de cargas [15].

1.1.3 Interações de hidrogênio

Interações de hidrogênio são muito comuns em proteínas, de maneira que são consideradas fatores importantes para a formação destas estruturas supramoleculares [16, 17, 18].

Figura 1.6: Representação de uma ponte de hidrogênio. O átomo doador (denotado por D) está ligado ao hidrogênio que é fornecido para o átomo aceitador (denotado por A).



Fonte: Kessel, 2011

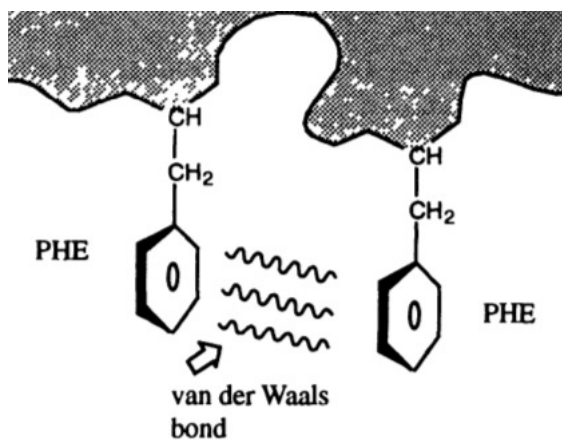
Com uma energia que varia entre 3,82 e 14,34 kcal/mol, a ponte de hidrogênio (ver figura 1.6) se dá entre um heteroátomo Y eletronegativo, isto é, que possui um par de elétrons desemparelhados (geralmente um Oxigênio (O) ou Nitrogênio(N)) e um hidrogênio que normalmente está ligado covalentemente a um átomo X também eletronegativo (O ou N). Assim, o átomo protonado X-H tende a se aproximar do átomo eletronegativo e o hidrogênio passa a compartilhar cargas com o mesmo, criando desta maneira a ponte de hidrogênio. Neste caso, o grupo que fornece o átomo de hidrogênio para a ligação é conhecido como doador (D) e o grupo que fornece o átomo com elétrons excedentes são denominados aceitadores (A). Visto que os átomos envolvidos neste processo de interação são parcialmente carregados, então a ponte de hidrogênio é considerada uma interação eletrostática fraca. As distâncias características para a formação de pontes de hidrogênio ficam em torno de 1,5-2,2 Å [8].

Moléculas de água no interior de proteínas também podem formar pontes de hidrogênios entre si e com as cadeias laterais da proteína. De outra maneira, átomos de carbono (do tipo C-H) podem fornecer hidrogênios mais fracos do que pontes de hidrogênio convencionais, pois a eletronegatividade de C é mais baixa que a do oxigênio e do nitrogênio [19].

1.1.4 Interações de Van der Waals

A interação de van der Waals resulta da interação entre moléculas hidrofóbicas, como por exemplo a interação entre anéis aromáticos² (figura 1.7). Consideradas interações fracas com energias típicas de 0,48 a 0,96 kcal/mol, a força destas interações diminuem rapidamente com a distância entre as moléculas, com um decaimento da ordem da sétima potência da distância de separação. Desta forma, as interações de van der Waals são cruciais na descrição das forças envolvidas em regiões apolares da proteína [8].

Figura 1.7: Interação de van der Waals entre duas Fenilalaninas (PHE), que são dois resíduos aromáticos e hidrofóbicos.



Fonte: Patrick, 1995

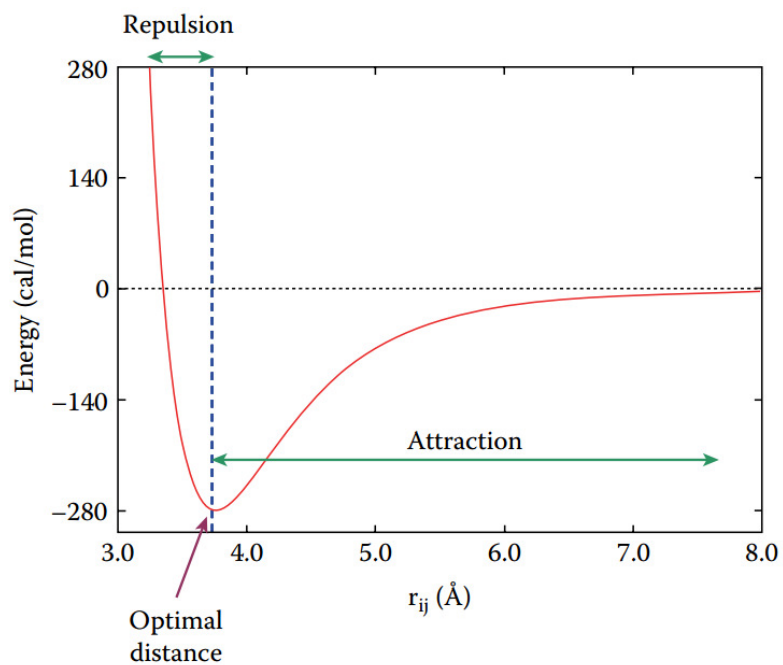
Devidos às flutuações quânticas transitórias nos elétrons, dipolos atômicos são gerados e suas mudanças constantes de direção, levam à variações no campo elétrico do átomo, o que significa que dois elétrons a uma distância em torno de 7 Å induzem mudanças opostas no campo elétrico um do outro. Deste modo, os dois dipolos opostos interagem eletrostaticamente e como resultado,

²Aproximadamente 60 % das cadeias laterais aromáticas estão envolvidas em interações do tipo π - π e próton- π

se atraem. Porém, na medida em que os átomos se aproximam, suas nuvens eletrônicas começam a se repelir, obedecendo assim o princípio da exclusão de Pauli ³. Estas forças atrativas e repulsivas constituem as interações de van der Waals, e ocorrem em átomos adjacentes de maneira que as distâncias interatômicas são suficientemente curtas para permitir interações atrativas, mas distantes o bastante para não produzir repulsão (figura 1.8). Conseqüentemente, cada átomo recebe um raio de van der Waals, que é igual à melhor distância do núcleo do átomo ao invólucro exterior dos átomos vizinhos. Nas macromoléculas, como as proteínas, os átomos são frequentemente embalados uns contra os outros de modo que a distância entre eles produz considerável energia de van der Waals [12, 20].

³De acordo com o princípio da exclusão de Pauli, os elétrons não podem se sobrepor espacialmente.

Figura 1.8: A energia de interação de van der Waals como uma função da distância entre o núcleo e um dímero de argônio. Nota-se claramente a atração eletrostática (interação de longo alcance) e a interação repulsiva devido ao princípio de exclusão de Pauling (interação de curto alcance).



Fonte: Kessel, 2011

A equação 1.1 descreve o potencial de Lennard-Jones, onde a interação de van der Waals é descrita empiricamente, já que para a mesma não possui uma expressão exata. Neste caso, r_{ij} são as distâncias entre os núcleos atômicos, enquanto que as constantes A_{ij} e B_{ij} representam propriedades físico-químicas específicas de cada par de átomos [21].

$$E_{vdw} \cong \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \quad (1.1)$$

As interações de van der Waals são constituídas por três tipos diferentes de forças, conhecidas como: interações de Keeson, interações de Debye e interações de dispersão de London.

Interações de Keeson ou interações dipolo permanente-dipolo permanente

As interações de Keeson são as mais fortes dentre as interações de van der Waals. Este tipo de interação ocorre entre dois dipolos permanentes, em que a parte negativa do dipolo de uma molécula interage com a parte positiva do dipolo de outra molécula. A energia de interação dipolo-dipolo depende da orientação entre os dois dipolos e varia com $1/r^3$ [22].

Interações de Debye ou interações dipolo permanente–dipolo induzido

As interações de Debye ocorrem entre moléculas polares e apolares. Assim, quando uma molécula com dipolo permanente se aproxima de outra molécula com alta polarizabilidade, de modo que a primeira molécula possa induzir um dipolo nesta última, ocorre uma interação dipolo permanente-dipolo induzido. Este tipo de interação varia com o inverso da quarta potência da distância intermolecular $1/r^4$ [22].

Interações de dispersão de London ou interações de dipolo induzido-dipolo induzido

A interação de dispersão de London decorre em todas as interações moleculares, de maneira que para moléculas apolares, tal interação é a única força presente. Quando duas moléculas apolares se aproximam, ambas perturbam a densidade eletrônica uma da outra, originando assim dipolos induzidos. Esta é a força mais fraca entre as interações de van der Waals e varia com $1/r^6$ [22].

1.2 Estudo de proteínas via Métodos de Modelagem Molecular

Inicialmente, aproximadamente até a década de 80, os medicamentos eram desenvolvidos através de metodologias baseadas em triagem cega, onde eram realizados diversos ensaios biológicos para várias moléculas, que eram testadas aleatoriamente, sem conhecimento dos mecanismos de ação e interação dos ligantes com o alvo receptor [23, 24]. No entanto, apesar de muitos fármacos utilizados nos dias atuais terem sido obtidos através destes métodos, eram necessários tempos cada vez maiores e altos custos para a descoberta e desenvolvimento de novas drogas, por esses processos. Neste cenário, o avanço da biologia molecular e das técnicas experimentais para isolamento e preparação de proteínas, a identificação de diversos alvos moleculares relacionados a doenças e o crescimento de informações estruturais, proporcionados principalmente por projetos como GENOMAS e PROTEOMAS, conduziram ao desenvolvimento de ferramentas e métodos computacionais para o auxílio no planejamento de novos medicamentos [25]. Por conseguinte, muitos bancos de dados disponibilizam estruturas moleculares, tanto de complexos macromoleculares (como o Protein Data Bank), quanto de moléculas ligantes (Cambridge Structural Database) permitindo desta maneira um grande avanço nas estratégias computacionais de planejamento de fármacos [26]

Diversas técnicas computacionais, como visualização, *docking*, simulações moleculares, química quântica e design *de novo*, foram concebidos com o propósito de identificar os tipos de interação envolvidas em sistemas ligante-proteína, assim como tentar estimar a afinidade com novas entidades moleculares bioativas. Dentro deste contexto foi desenvolvida a área de Desenho Racional de Fármacos Baseados em Estruturas (DRBE), que é um conjunto de metodologias sistemáticas que tem como objetivo encontrar compostos protótipos. Para isto, é realizado o estudo de estruturas tridimensionais das macromoléculas receptoras, levando em consideração as interações intermoleculares envolvidas no processo de ligação do fármaco com o sítio ativo do alvo macromolecular. Dentro do DRBE estão diversas metodologias amplamente utilizadas tanto para a descoberta, quanto para a otimização de compostos já conhecidos [27, 28].

A partir da identificação de alvos macromoleculares, conseguida através de técnicas experimen-

tais como raios x e NMR, é feita uma triagem *In Silico* dos ligantes que podem desencadear o efeito biológico esperado. Para isso os métodos teóricos computacionais como QSAR ((*sigla para Quantitative Structure–Activity Relationship*)) e *Docking* são amplamente utilizados, principalmente nas fases de otimização do composto líder, onde modificações mais precisas são essenciais para que a molécula ligante se ajuste ao perfil de ADME-tox⁴ sem perder a afinidade com o alvo [4, 28].

O método QSAR (Quantitative Structure–Activity Relationship) estuda a relação quantitativa entre estrutura e atividade, podendo ser descrito como uma aplicação de métodos de análise de dados estatísticos para o desenvolvimento de modelos que possam prever a atividade ou propriedade biológica de compostos baseados em suas estruturas moleculares, ou seja, um modelo QSAR correlaciona as propriedades estruturais, físico-químicas e conformacionais com suas atividades biológicas, utilizando-se de aproximações estatísticas [29]. Existem diferentes abordagens para este método, das quais podemos citar os métodos QSAR-2D (que faz uso de descritores⁵ obtidos experimentalmente e calculados a partir da estrutura molecular bidimensional), o QSAR-3D (que utiliza descritores moleculares obtidos de estruturas tridimensionais pode ser classificado como um método dependente ou independente do receptor) e o QSAR-4D (que inclui flexibilidade conformacional e liberdade de alinhamento como quarta dimensão) [30, 31, 32, 33]. O método CoMFA (Comparative molecular field analysis) é um tipo de método QSAR-3D, que inclui Análise comparativa de campo molecular e considera que as interações intermoleculares entre o ligante e o sítio ativo do alvo macromolecular, especialmente as estéricas e eletrostáticas contribuem fortemente para ação biológica [34].

Aproximações computacionais de reconhecimento molecular, chamadas de metodologias de *docking*, "ancoram" pequenas estruturas moleculares nas regiões bioativas de alvos macromoleculares (biorreceptor que pode ser uma proteína), a fim de ativar ou bloquear sua função biológica. Estas metodologias são extensamente aplicadas na identificação de compostos protótipos novos,

⁴ADME-tox é o acrônimo utilizado na farmacocinética e na farmacologia para as características que definem a performance e a atividade farmacológica dos compostos moleculares, como absorção, distribuição, metabolismo e excreção, levando em consideração a sua toxicidade

⁵Descritores moleculares são propriedades estruturais calculadas ou medidas experimentalmente

assim como para a otimização de moléculas bioativas já conhecidas. A partir do conhecimento do alvo terapêutico e a determinação de sua estrutura molecular tridimensional (obtidos por meio de técnicas experimentais e disponíveis em bancos de dados como o PDB), de modo semelhante às metodologias tradicionais, é realizada uma triagem *in silico* nos bancos de estruturas moleculares, com o objetivo de selecionar os candidatos mais promissores a composto protótipo [4]. Através de metodologias *docking* mais acuradas, os candidatos mais promissores passam por otimizações e refinamentos em suas estruturas moleculares, de modo que suas funções esperadas sejam potencializadas e suas propriedades farmacocinéticas sejam condizentes com as de um fármaco apropriado para consumo.

As metodologias de *docking* mais específicas se propõem a descrever as interações intra e intermoleculares envolvidas no reconhecimento molecular entre a micro molécula ligante e o alvo macromolecular receptor. Todavia, prever a energia de ligação entre ligante-receptor de forma precisa exige um alto custo computacional. Assim, para a fase de busca conformacional, muitos programas de *docking* utilizam modelos de funções simples para a energia potencial, baseadas em campos de força clássicos via mecânica molecular. Já as funções *scoring*, consideradas mais sofisticadas, incluem termos entrópicos oriundos da solvatação do sistema e são utilizadas para avaliar as conformações do complexo geradas a partir das primeiras. Em síntese, as funções utilizadas pelos programas de *docking* se dividem em três classes [35], baseadas em campos de força (como os campos GROMOS [36], AMBER [37], CHARMM [38] e MMFF94 [39]), em funções empíricas (estas reproduzem dados experimentais da energia livre de ligação, e são implementadas em programas como GOLD [40], LUDI [41], FlexX [42]) e em funções baseadas em conhecimento (utilizam informações estruturais experimentais, porém avaliam a mudança de energia livre em função de coordenadas intratômicas, através da estatística de Boltzmann, como o por exemplo programa DrugScore [43]). Além disso, muitas metodologias *docking* ligante-receptor incluem cada vez mais variantes, como maior grau de flexibilidade por exemplo, na tentativa de buscar o maior número de soluções e encontrar o resultado mais acurado possível. Para isto, programas de *docking* utilizam métodos de Dinâmica Molecular [44] e Monte Carlo [45] em seus algoritmos, na tentativa

de encontrar a energia mínima destes tipos de sistema [27, 4].

1.3 Estudo de proteínas através de métodos baseados em Mecânica Quântica

Os sistemas ligante-receptores são complexos e possuem considerável flexibilidade, assim como funções ácidas e básicas que podem induzir ao equilíbrio tautomérico e conferir diferentes conformações e características estéricas e eletrônicas [46, 47, 48]. Neste ensejo, o uso de métodos de química quântica tem se tornado um importante tipo de abordagem no estudo das interações ligante-proteína. Esta via de estudo constitui modelos que visam um entendimento mais aprofundado das propriedades fundamentais de fármacos [49, 50]. O uso desde métodos *ab initio*⁶, como Hartree-Fock (HF), até métodos pós HF, aliados às metodologias mais tradicionais como QSAR e mecânica molecular, tem permitido sistematicamente o refinamento das descrições da interações intra e intermoleculares em complexos ligante-proteína [50]. Aproximações do tipo QM/MM (do inglês Quantum Mechanics/ Molecular Mechanics) têm se mostrado metodologias promissoras na predição da afinidade e energia de ligação entre ligante e proteínas, porém na maioria dos casos, devido ao tamanho extenso da proteína, esta é calculada via mecânica molecular, enquanto que somente as moléculas dos ligantes são tratados via mecânica quântica, afetando assim a confiabilidade dos cálculos realizados mediante QM/MM [51, 52, 53].

O problema dos altos custos computacionais, devido à larga escala dos sistemas protéicos, foram em certa medida solucionadas a partir do desenvolvimento dos métodos de escala linear baseados no fracionamento molecular, como o MFCC (Molecular Fractional with Caps Conjugated), onde cálculos mais sofisticados, a nível de DFT (Density Functional Theory) por exemplo, podem ser aplicados em sistemas de milhares de átomos [54, 55, 56]. Os detalhes teóricos em que se baseiam tais métodos serão apresentados no próximo capítulo, ressaltando que no presente trabalho,

⁶O termo *ab initio* (palavra originária do latim, significa "desde o início") está relacionado aos cálculos que buscam a resolução das equações sem referência a outros resultados

além de aplicarmos os métodos HF e DFT, realizaremos também o estudo das energias de interação no sistema proposto a nível perturbativo via MP2 (Método perturbativo de Møller-Plesset de segunda ordem).

Capítulo 2

Métodos Quânticos

Métodos computacionais baseados em mecânica quântica (MQ) são importantes ferramentas para o entendimento e/ou predição de interações químicas-biológicas, auxiliando no desenho racional de fármacos [57]. No entanto, resultados computacionais dependem tanto dos parâmetros, como da qualidade dos conjuntos de bases utilizados nas modelagens de sistemas biológicos via MQ. Deste modo, novos métodos teóricos de química quântica foram e têm sido desenvolvidos, com o objetivo de tratar sistemas macromoleculares, como ácidos nucleicos e proteínas. A descrição de aminoácidos e peptídeos têm sido frequentemente realizadas via métodos *ab initio* ou semiempírico e suas energias são calculadas com o auxílio de cálculos que envolvem o uso de campos de força e métodos de mecânica molecular [58].

Os métodos quânticos podem ser classificados em duas classes de acordo com suas funcionalidades: uma classe que inclui os métodos usados para quantificar energias e otimizar estruturas, enquanto que outra classe abrange aqueles que são voltados para o cálculo de propriedades moleculares, tais como cargas parciais, comprimentos de ligações e ângulos de torção [59]. Outrossim, um aspecto importante em relação a cada método quântico é o seu custo computacional. Desta maneira, como alguns métodos são consideravelmente mais dispendiosos que outros, a escolha do

método deve levar em consideração, dentre outras questões, o tamanho do sistema molecular. Assim, o custo computacional do método Hartree-Fock é proporcional a N^3 , com N sendo o número de funções base, isto é, se o número de funções base dobrar, o custo do cálculo aumentará em 8 vezes. No caso de métodos mais acurados esse custo é maior, como no caso do método perturbativo de Møller-Plesset de segunda ordem, o qual escala nominalmente como N^5 . O método da Teoria do Funcional da densidade apresenta escala nominal semelhante ao método de Hartree-Fock (cerca de N^3), porém devido a termos de proporcionalidade maiores, na prática exibem cálculos computacionais mais onerosos [60].

A aplicação direta de métodos quânticos em sistemas macromoleculares depende do tamanho da proteína, uma vez que o poder computacional atualmente disponível não é suficiente para gerar resultados com nível de segurança similar aos experimentais. Desta forma, para aumentar a precisão dos cálculos e possibilitar o estudo e descrições de macromoléculas, são propostas metodologias híbridas que mesclam mecânica quântica e mecânica molecular (MQ/MM), assim como são empregadas diferentes abordagens, como por exemplo as estratégias de fragmentação molecular, denominadas MFCC (sigla do termo inglês Molecular Fractional with Caps Conjugated) [59].

Neste capítulo serão abordadas as bases teóricas dos métodos quânticos que sustentam a metodologia proposta neste trabalho, isto é, os métodos que foram usados na descrição e investigação do efeito de interações de longo alcance na energia entre fármaco e proteína. A primeira seção trata sobre os métodos *ab initio*, enquanto que a segunda apresenta os fundamentos da Teoria do Funcional da Densidade (DFT). Serão mostrados os conceitos e fundamentações teóricas sobre as cargas pontuais e a correção de longo alcance para os efeitos eletrostáticos. O método perturbativo de Møller-Plesset de segunda ordem (MP2) também será discutido, assim como a implementação das

estratégia de fragmentação molecular (MFCC) para o estudo das interações entre ligante e proteína.

2.1 O método Hartree-Fock

Obter soluções exatas da equação de Schrödinger para sistemas poliatômicos não é uma tarefa factível. No entanto, o uso de algumas técnicas e metodologias torna possível a obtenção de soluções aproximadas bastante precisas. Desta forma, propriedades de sistemas atômicos e moleculares podem ser obtidas por meio de métodos aproximativos como por exemplo o método de Hartree-Fock (HF) [4].

Embora o método de Hartree-Fock forneça solução aproximada satisfatória para o problema de muitos elétrons, ele é considerado como um ponto de partida para outros métodos, tais quais fornecem resultados mais próximos da solução exata. A partir da aproximação de Born-Oppenheimer (BO) o método de HF busca a solução para a equação de Schrödinger eletrônica 2.1,

$$\hat{H}\psi(\mathbf{r}; \mathbf{R}) = E(\mathbf{R})\psi(\mathbf{r}; \mathbf{R}) \quad (2.1)$$

Neste caso, as dependências temporais (o interesse central é obter soluções para estados estacionários) e os efeitos relativísticos são desprezados. \hat{H} é a parte eletrônica do operador hamiltoniano do sistema e E os autovalores da energia eletrônica. A função de onda, representada por ψ descreve os estados eletrônicos nas posições \mathbf{r} com parametrização em \mathbf{R} , que consiste no conjunto fixo das posições dos núcleos atômicos.

O hamiltoniano para a equação 2.1 é dado por:

$$\hat{H} = -\frac{\hbar^2}{2m_e} \sum_i^n \nabla_i^2 - \sum_i^n \sum_I^N \frac{Z_I e^2}{4\pi\epsilon_0 r_{Ii}} + \frac{1}{2} \sum_{ij}^n \frac{e^2}{4\pi\epsilon_0 r_{ij}} \quad (2.2)$$

onde m_e é a massa do elétron, Z_I o número atômico do núcleo I , R_{Ii} a distância entre o núcleo I e o elétron i e r_{ij} a distância entre os elétrons i e j . Os índices n e N se referem ao número de elétrons e núcleos, respectivamente.

O hamiltoniano eletrônico 2.2 pode ser reescrito utilizando o sistema de unidades atômicas de Hartree (a.u.)¹, tomando a seguinte forma:

$$\hat{H} = -\frac{1}{2} \sum_i^n \nabla_i^2 - \sum_i^n \sum_I^N \frac{Z_I}{r_{Ii}} + \sum_{ij}^n \frac{1}{r_{ij}} \quad (2.3)$$

O primeiro termo descreve a energia cinética dos elétrons movimentando-se no campo dos núcleos atômicos estacionários. O segundo termo se refere à interação elétron-núcleo e o último à interação elétron-elétron dos sistema. Em cálculos de estruturas moleculares é convencional não incluir o termo de repulsão núcleo-núcleo no hamiltoniano, uma vez que o mesmo depende somente do parâmetro \mathbf{R} e contribui como uma constante. Além disso, tal interação pode ser adicionada como um termo clássico no final dos cálculos [61].

Para obter a melhor aproximação possível na descrição de um sistema de muitos elétrons, o método de HF baseia-se no uso de funções de estado expandidas em formas de determinantes, visto que estas funções de ondas devem ser antissimétricas mediante uma permutação entre dois elétrons. A expansão da função de onda dos elétrons em um conjunto completo de funções de estado para um sistema de N-elétrons é conhecida como o determinante de Slater (2.4),

$$\Phi_0 = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \quad (2.4)$$

Assim, a introdução de um único determinante de Slater simplifica o problema de se encontrar uma única função de onda para N elétrons de um sistema, que agora centra-se em calcular N funções de onda de um elétron, chamadas de spin-orbitais moleculares. Por outro lado, a função de onda de

¹No sistema de unidades atômicas: $m_e = \hbar = e = 1/4\pi\epsilon_0 = a_0 = 1$, com energia expressa em Hartree e a distância em termo do raio de Bohr a_0 .

HF, dada pelo determinante de Slater, pode ser reescrita da seguinte forma:

$$\Phi_0 = \frac{1}{N!} \sum_{i=1}^{N!} (-1)^{p_i} P_i \{ \chi_1(\mathbf{x}_1) \chi_2(\mathbf{x}_2) \dots \chi_N(\mathbf{x}_N) \} \quad (2.5)$$

onde P_i é o operador permutação de N elementos e p_i é o número de transposições de P_i .

As funções χ 's denotam os spin-orbitais moleculares (SOM) de cada elétron e podem ser separadas em termos das coordenadas espaciais e do spin, isto é,

$$\chi_a(\mathbf{x}_1) = \phi_p(\mathbf{r}_1)\alpha(1) \quad (2.6)$$

ou

$$\chi_a(\mathbf{x}_1) = \phi_p(\mathbf{r}_1)\beta(1) \quad (2.7)$$

onde o spin "up" é representado por α e o "down" por β . Notada por ϕ , a parte espacial da função de estado eletrônica é conhecida como os orbitais moleculares (OMs).

A depender das restrições impostas sobre os orbitais moleculares, o método HF pode ser não restrito (UHF) e restrito (RHF)². Geralmente, o método UHF é aplicado ao caso de sistemas com camadas abertas, onde não são feitas restrições sobre os SOM. Já no caso de sistemas com camadas fechadas, onde os orbitais moleculares são ocupados por pares de elétrons, um com spin "up" (α) e "down" (β), é recomendado o uso do método RHF[62].

2.1.1 A equação canônica de Hartree-Fock

Os spin-orbitais moleculares (SOM) que compõem o determinante de Slater dado pela equação 2.4 são obtidos por meio da equação de autovalor de Hartree-Fock (HF). No entanto é necessário desenvolver um critério de escolha destas funções de estado, o qual é fornecido através do método

²UHF é a sigla em inglês de Unrestricted Hartree-Fock e RHF vêm do termo Restricted Hartree-Fock.

variacional [63, 64, 65].

O teorema variacional pode ser útil para a dedução da equação de Schrödinger independente do tempo, a partir do argumento de que a mesma obedece à condição de extremo do funcional de energia $E[\phi]$, que é justamente o valor médio do operador Hamiltoniano. Por outro lado, o uso do teorema variacional consiste na obtenção de funções de estado dependentes de parâmetros que minimizem $E[\phi]$, de modo que,

$$E[\phi] \geq E_0 \quad (2.8)$$

onde $E[\phi] = \langle \Phi | \hat{H} | \Phi \rangle$.

A equação 2.8 mostra uma das definições do teorema variacional, onde E deve ser sempre maior ou igual á energia do estado fundamental E_0 . Assim quanto menor E , mais próxima estará da solução procurada. Em outras palavras, a função de onda mais próxima da solução exata é aquela que minimiza E , e esta última é o valor médio do operador hamiltoniano \hat{H} . Logo, como \hat{H} é um funcional dos spin-orbitais moleculares a energia também o será. Assim sendo, temos:

$$E = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle = E[\chi] \quad (2.9)$$

Da equação 2.9, têm-se que a minimização da energia implica na minimização do valor médio do operador \hat{H} ,

$$\langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \langle \Phi_0 | O_1 + O_2 | \Phi_0 \rangle \quad (2.10)$$

Com,

$$\hat{H} = O_1 + O_2 \quad (2.11)$$

onde,

$$O_1 = \sum_{i=1}^N h(i) = \sum_{i=1}^N \left\{ -\frac{1}{2} \nabla_i^2 - \sum_{\alpha} \frac{Z_{\alpha}}{r_{i\alpha}} \right\} \quad (2.12)$$

e

$$O_2 = \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (2.13)$$

O primeiro termo da equação 2.11 é definido como o operador de um elétron 2.12, enquanto que o segundo termo como o operador de dois elétrons (2.13).

Os valores esperados dos operadores de 1 e 2 elétrons são

$$\langle \Phi_0 | O_1 | \Phi_0 \rangle = \sum_{a=1}^N \langle a | h | a \rangle \quad (2.14)$$

e

$$\langle \Phi_0 | O_2 | \Phi_0 \rangle = \frac{1}{2} \sum_{a=1}^N \sum_{b=1}^N \langle ab || ab \rangle \quad (2.15)$$

respectivamente.

A partir das equações 2.14 e 2.15, o valor esperado do hamiltoniano pode ser escrito como:

$$\langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \sum_{a=1}^N \langle a | h | a \rangle + \frac{1}{2} \sum_{a=1}^N \sum_{b=1}^N \langle ab || ab \rangle \quad (2.16)$$

Finalmente, utilizando o funcional 2.16 a equação 2.9 se torna,

$$E[\chi] = \sum_a \langle \chi_a | h | \chi_a \rangle + \frac{1}{2} \sum_{a,b} (\langle \chi_a \chi_b | \chi_a \chi_b \rangle - \langle \chi_a \chi_b | \chi_b \chi_a \rangle) \quad (2.17)$$

onde o operador de 2 elétrons é notado a partir da definição

$$\langle ab||ab \rangle = \langle ab|ab \rangle - \langle ab|ba \rangle \quad (2.18)$$

O funcional dado em 2.17 deve ser minimizado sob a condição de ortonormalidade dos SOM's, isto é,

$$\langle \chi_a | \chi_b \rangle - \delta_{ab} = 0 \quad (2.19)$$

Devido ao vínculo 2.19, a técnica de multiplicadores de lagrange pode ser aplicada a minimização do funcional 2.17, ou seja,

$$L[\chi] = E[\chi] - \sum_{a,b} \epsilon_{ba} (\langle \chi_a | \chi_b \rangle - \delta_{ab}) \quad (2.20)$$

onde os multiplicadores de Lagrange são os coeficientes ϵ_{ab} e L é real. Desta forma, como E é real, então o segundo termo da equação 2.20 também deve ser, logo

$$\langle \chi_a | \chi_b \rangle^* = \langle \chi_b | \chi_a \rangle \quad (2.21)$$

Da expressão 2.21, pode-se concluir que:

$$\epsilon_{ab}^* = \epsilon_{ba} \quad (2.22)$$

Portanto, os multiplicadores de lagrange são os elementos que formam uma matriz hermitiana.

Prosseguindo com o procedimento de minimização de L , uma pequena variação nos SOM's produzirá uma variação em L dada por,

$$\delta L = \sum_a \langle \delta \chi_a | h | \chi_a \rangle + \sum_{a,b} \langle \delta \chi_a \chi_b | \chi_a \chi_b \rangle - \langle \delta \chi_a \chi_b | \chi_b \chi_a \rangle - \sum_{a,b} \epsilon_{ab} \langle \delta \chi_a | \chi_b \rangle + c.c. \quad (2.23)$$

com c.c. sendo o complexo conjugado dos demais termos.

Ao utilizar a seguinte notação,

$$\begin{aligned}\langle \chi_a \delta \chi_b | \chi_a \chi_b \rangle &= \left\langle \chi_a(1) \delta \chi_b(2) \left| \frac{1}{r_{12}} \right| \chi_a(1) \chi_b(2) \right\rangle \\ \langle \delta \chi_b \chi_a | \chi_b \chi_a \rangle &= \left\langle \delta \chi_b(1) \chi_a(2) \left| \frac{1}{r_{12}} \right| \chi_b(1) \chi_a(2) \right\rangle\end{aligned}\tag{2.24}$$

são definidos os operadores de Coulomb (\mathcal{J}_b) e de troca (\mathcal{K}_b), expressos nas respectivas equações

$$\mathcal{J}_b(1) \chi_a(1) = \left\langle \chi_b(2) \left| \frac{1}{r_{12}} \right| \chi_b(2) \right\rangle \chi_a(1)\tag{2.25}$$

$$\mathcal{K}_b(1) \chi_a(1) = \left\langle \chi_b(2) \left| \frac{1}{r_{12}} \right| \chi_a(2) \right\rangle \chi_b(1)\tag{2.26}$$

Em termos dos operadores de Coulomb (2.25) e de troca 2.26, a equação 2.23 pode ser reescrita como,

$$\delta L = \sum_a \langle \delta \chi_a(1) | \left\{ \left[h(1) + \sum_b [\mathcal{J}_b(1) - \mathcal{K}_b(1)] \right] | \chi_a(1) \rangle - \sum_b \epsilon_{ba} | \chi_b(1) \rangle \right\} + c.c. \tag{2.27}$$

Para minimizar L faz-se $\delta L = 0$, e como $\delta \chi_a$ é arbitrário, temos

$$\left\{ h(1) + \sum_b [\mathcal{J}_b(1) - \mathcal{K}_b(1)] \right\} \chi_a(1) = \sum_b \epsilon_{ba} \chi_b(1)\tag{2.28}$$

O termo da expressão 2.28 que se encontra entre chaves é chamado de operador de Fock,

$$\mathcal{F}(1) = h(1) + \sum_b [\mathcal{J}_b(1) - \mathcal{K}_b(1)]\tag{2.29}$$

Reescrevendo a equação 2.28 em termos de 2.29, obtemos

$$\mathcal{F}(1)\chi_a(1) = \sum_b \epsilon_{ba}\chi_b(1) \quad (2.30)$$

Como os multiplicadores de lagrange sob uma transformação unitária podem ser escritos na forma matricial,

$$\epsilon' = \mathbf{U}^\dagger \epsilon \mathbf{U} \quad (2.31)$$

onde ϵ é uma matriz hermitiana. Então, para uma transformação unitária que diagonalize ϵ' , os elementos desta última matriz serão dados por,

$$\epsilon'_{ab} = \epsilon'_a \delta_{ab} \quad (2.32)$$

Utilizando 2.32, equação 2.30 se torna,

$$\mathcal{F}(\mathbf{x}_1)\chi'_a(1) = \sum_b \epsilon'_a \delta_{ab}\chi'_b(1) = \epsilon'_a \chi'_a(1) \quad (2.33)$$

que pode ser reescrita, sem as linhas, como

$$\mathcal{F}(\mathbf{x}_1)\chi'_a(1) = \epsilon_a \chi_a(1) \quad (2.34)$$

A equação 2.34 é denominada como equação canônica de Hartree-Fock, enquanto que os SOM's obtidos por meio desta são conhecidos como spin-orbitais canônicos. Além disso, substituindo o operador de Fock na mesma, multiplicando pela função $\chi_a^*(1)$ e integrando a expressão resultante, têm-se:

$$\epsilon_a = \langle \chi_a(1) | h(1) + \sum_b [\mathcal{J}_b(1) - \mathcal{K}_b(1)] | \chi_a(1) \rangle \quad (2.35)$$

que expressa as energias orbitais.

Ao utilizar as expressões dos operadores $\mathcal{J}_b(1)$ e $\mathcal{K}_b(1)$ em 2.35, chegamos a,

$$\epsilon_a = \langle a|h|a\rangle + \sum_b \langle ab||ab\rangle \quad (2.36)$$

Comparando 2.36 com a expressão para o funcional 2.17, temos a equação para a energia eletrônica total:

$$E = \sum_a \epsilon_a - \frac{1}{2} \sum_{a,b} \langle ab||ab\rangle \quad (2.37)$$

Assim, a energia resulta do somatório das energia orbitais, que são autovalores do operador de Fock. Com a replusão elétron-elétron é contada duas vezes, a mesma é subtraída.

Visto que o operador de Fock depende da solução dos spin-orbitais calculados pela equação de HF, e cada SOM existe uma equação que depende de outros orbitais por meio do operador de Fock, as equações são resolvidas iterativamente através de aproximações. Devido a isto, o método HF é dito autoconsistente. O método de HF se aplica melhor a sistemas de poucos elétrons, uma vez que para sistemas moleculares maiores, a resolução numérica da equação de HF não é factível. Neste último caso, são utilizadas as equações de Hartree-Fock-Roothaan ou de Pople-Nesbet, cujas resoluções são baseadas em métodos que expandem a parte espacial dos SOM's a partir de um conjunto de funções-base conhecidas. Não obstante, o método HF utiliza uma metodologia variacional baseada em um único determinante. Dessa maneira, este método perde alguns detalhes relacionados às interações eletrônicas, pois o movimento correlacionado entre as partículas passa a ser tratado por uma interação média autoconsistente [65]. Por esta razão, se faz necessário definir uma energia de correlação como sendo a diferença entre a energia calculada via método HF restrito e a energia exata E_0 , isto é,

$$E_{corr} = E_0 - E_{HF}. \quad (2.38)$$

Deste modo, diversos métodos teóricos foram desenvolvidos visando obter a energia de correlação eletrônica. Dentre eles, destaca-se o Método Perturbativo de Møller-Plesset (MPPT), que será apresentado adiante.

2.2 O Método Perturbativo de Møller-Plesset (MPPT)

A Teoria de Perturbação de Muito Corpos (MBPT)³ é um dos métodos pós-HF que incluem a correlação eletrônica na descrição da estrutura eletrônica de átomos e moléculas [66].

Na teoria de perturbação, o hamiltoniano do sistema é separado em duas partes distintas,

$$\hat{H} = \hat{H}_0 + \hat{V} \quad (2.39)$$

onde \hat{H}_0 é o hamiltoniano não-perturbado, com autovalores e autoestados conhecidos. Já o termo \hat{V} é definido como a perturbação no sistema. Assim, a energia exata E é dada pela resolução da equação independente do tempo de Schrödinger:

$$\hat{H}|\psi\rangle = E|\psi\rangle, \quad (2.40)$$

expressa como uma soma de contribuições dependentes da perturbação \hat{V} .

A teoria de perturbação de Møller-Plesset (MPPT do inglês *Møller-Plesset Perturbation Theory*) define o hamiltoniano não perturbado, H_0 , como a soma dos operadores de Fock, isto é,

$$\hat{H}_0 = \sum_{\mu=1}^N \hat{F}(\mu) = \sum_{\mu=1}^N \left\{ \hat{h}(\mu) + \sum_{j=1}^N [\hat{J}_j(\mu) - \hat{K}_j(\mu)] \right\}. \quad (2.41)$$

³MBPT é a sigla proveniente do termo em inglês *Many Body Perturbation Theory*.

A perturbação \hat{V} pode ser obtida através de:

$$\hat{V} = \hat{H} - \hat{H}_0 \quad (2.42)$$

onde, abrindo os termos do hamiltoniano chegamos em,

$$\hat{V} = \sum_{\mu=1}^N \left[\sum_{\nu>\mu} \frac{1}{r_{\mu\nu}} - V^{HF}(\mu) \right]. \quad (2.43)$$

O último termo se refere ao operador de Hartree-Fock, dado por:

$$V^{HF}(\mu) = \sum_j \left[\hat{J}_j(\mu) - \hat{K}_j(\mu) \right]. \quad (2.44)$$

Os autovalores de energia $E_0^{(0)}$ correspondentes ao sistema não-perturbado descrito pelo hamiltoniano 2.41, são calculados pela soma das energias de todos os spin-orbitais ocupados $|\psi_a^\mu\rangle^A$, ou seja,

$$\hat{H}_0|\Phi_0\rangle = E_0^{(0)} = (\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_N)|\psi_1^{(1)}, \psi_2^{(2)}, \dots, \psi_N^{(N)}\rangle^A \quad (2.45)$$

onde,

$$E_0^{(0)} = \sum_{a=1}^N \varepsilon_a. \quad (2.46)$$

As funções de onda de HF (Φ^{HF}) constituem o espaço \mathcal{P} , enquanto que as funções Φ^s formam o espaço \mathcal{Q} e são geradas pelas excitações simples (S), duplas (D), triplas (T), etc. dos elétrons de spins orbitais ocupados ψ_i, ψ_j, \dots para orbitais virtuais ψ_a, ψ_b, \dots . As funções Φ^{HF} e Φ^s são ortogonais entre si, o que simplifica o cálculo das diferentes correções de energia. Assim, dado o

operador perturbação (equação 2.43),

$$\hat{V} = \sum_{\mu=1}^N \left\{ \sum_{v>\mu}^N \frac{1}{r_{\mu v}} - \sum_{j=1}^N [\hat{J}_j(\mu) - \hat{K}_j(\mu)] \right\} \quad (2.47)$$

então, ao aplicar o operador perturbação no estado de referência de HF, teremos a correção de primeira ordem na energia, $E_0^{(1)}$, dada por:

$$E_0^{(1)} = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle. \quad (2.48)$$

De onde, usando as regras de Condon-Slater para o termo de duas e uma partícula, nesta mesma ordem teremos,

$$\mathcal{O}_2 = \sum_{\mu} \sum_{v<\mu} \frac{1}{r_{\mu v}} \quad (2.49)$$

e

$$\mathcal{O}_1 = \sum_{\mu} V^{HF}(\mu). \quad (2.50)$$

Portanto, a aplicação destes operadores no estado de referência resulta em:

$$\langle \Phi_0 | \mathcal{O}_2 | \Phi_0 \rangle = \frac{1}{2} \sum_{i,j=1}^N \langle \psi_i \psi_j | | \psi_i \psi_j \rangle \quad (2.51)$$

$$\langle \Phi_0 | \mathcal{O}_1 | \Phi_0 \rangle = \sum_{i,j=1}^N \langle \psi_i \psi_j | | \psi_i \psi_j \rangle \quad (2.52)$$

Assim, a correção de primeira ordem é dada como,

$$E_0^{(1)} = -\frac{1}{2} \sum_{i,j=1}^N \langle \psi_i \psi_j | | \psi_i \psi_j \rangle. \quad (2.53)$$

Logo, a equação 2.53 revela que a energia de HF é idêntica à energia calculada através de MP1,

$$E_0 = E_0^{(0)} + E_0^{(1)} = \sum_{i=1}^N \varepsilon_i - \frac{1}{2} \sum_{i,j=1}^N \langle \psi_i \psi_j | \psi_i \psi_j \rangle = E_{HF} \quad (2.54)$$

que é a própria energia de HF para o estado fundamental. Deste modo, o método de HF é correto até a primeira ordem do MPPT e a energia de correlação começa a aparecer da segunda ordem em diante. Isso ocorre também para o momento de dipolo, a densidade eletrônica e outras propriedades eletrônicas [67].

2.2.1 A correção de Segunda Ordem

A correção na energia de correlação, MP2, é dada por:

$$E_0^{(2)} = \sum_{i \neq 0} \frac{|\langle \phi_0 | \hat{V} | \phi_i \rangle|^2}{\varepsilon_0 - \varepsilon_i}, \quad (2.55)$$

onde o somatório é sobre todos os orbitais moleculares ocupados. As autofunções $|\phi_0\rangle$ correspondem ao determinante de Slater do estado de referência de HF, enquanto que os autoestados ϕ_i correspondem aos determinantes de Slater para as excitações simples, duplas e de ordens superiores.

Para as excitações simples, $|\phi_a^r\rangle$, os orbitais ocupados a são substituídos pelos orbitais virtuais r . Assim, temos:

$$\langle \phi_0 | \hat{V} | \phi_a^r \rangle = \langle \phi_0 | \hat{H} | \phi_a^r \rangle - \langle \phi_0 | \hat{H}_0 | \phi_a^r \rangle. \quad (2.56)$$

Como consequência do Teorema de Brillouin o termo $\langle \phi_0 | \hat{H} | \phi_a^r \rangle$ se anula. Já o segundo termo nos dá:

$$\langle \phi_0 | \hat{H}_0 | \phi_a^r \rangle = \sum_j \varepsilon_j \langle \phi_0 | \phi_a^r \rangle = 0. \quad (2.57)$$

Dessa maneira, chegamos a,

$$\langle \phi_0 | \hat{V} | \phi_a^r \rangle = 0. \quad (2.58)$$

Portanto, o resultado da equação 2.58 revela que excitações simples não alteram a energia do estado de referência de HF.

De modo análogo, as excitações duplas, $|\phi_{ab}^{rs}\rangle$, onde os orbitais ocupados a e b são substituídos pelos orbitais virtuais r e s , apresentam contribuição nula para os operadores de um elétron, contribuindo somente para o operador de dois elétrons, o que resulta em

$$\langle \phi_0 | \hat{V} | \phi_{ab}^{rs} \rangle = \langle ab || rs \rangle. \quad (2.59)$$

Finalmente, usando o resultado 2.59, a expressão 2.55 pode ser escrita como,

$$E_0^{(2)} = \sum_{a \neq r, b \neq s} \frac{|\langle ab || rs \rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s}. \quad (2.60)$$

que expressa a energia de segunda ordem e inclui somente substituições duplas. Além disso, como a correção de primeira ordem recupera apenas a energia de HF, a correção de segunda ordem (ou método MP2) é o tratamento de mais baixa ordem em MBPT.

2.3 A Teoria do Funcional de Densidade (DFT)

A Teoria do Funcional de Densidade (DFT)⁴ é uma alternativa ao método de HF para a resolução do problema de N corpos. Desenvolvida em 1964 por Hohenberg, Kohn and Sham, a DFT se baseia na idéia de que as propriedades do estado fundamental de um sistema de muitas partículas podem ser caracterizadas em termos da densidade eletrônica total do sistema [68]. Deste modo,

⁴DFT é a sigla para o termo em inglês Density Functional Theory

este método converte o problema de encontrar uma função de onda com $3N$ variáveis (se o spin não for considerado) em um problema de calcular a densidade eletrônica do sistema com somente três variáveis. O uso da DFT tornou-se popular devido a fatores como: o baixo custo computacional (quando comparada com métodos pós HF), pode ser usada nos cálculos de moléculas com $N_{\text{átomos}} \geq 20$ ⁵ em um tempo significativamente menor que os métodos baseados na metodologia de HF, além de reproduzir frequentemente resultados muito próximos de resultados experimentais [61].

O próprio nome da teoria sucede do fato de que a energia é um funcional da densidade eletrônica $\rho(\mathbf{r})$, isto é, $E[\rho(\mathbf{r})]$. Porém, o uso da densidade eletrônica como variável básica tem sido amplamente explorado em modelos de condução térmica e elétrica, baseados na aplicação da teoria dos gases a metais. Dentre estes modelos, pode-se apontar o modelo de Tomas-Fermi para o gás de elétrons, que posteriormente foi aperfeiçoado por Dirac, que incluiu a energia de troca. Assim, o funcional de energia de Thomas-Fermi-Dirac (TFD) é expresso da seguinte forma:

$$E_{TFD}[\rho] = \lambda \int \rho(r)^{5/3} dr + \int \rho(r)v(r)dr + \frac{1}{2} \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2 - \frac{3}{4}e^2 \left(\frac{3}{\pi}\right)^{1/3} \int \rho^{4/3} dr \quad (2.61)$$

com,

$$\lambda = \frac{3}{10}(3\pi^2)^{2/3}.$$

O primeiro termo do lado direito, representa a energia cinética dos elétrons, enquanto que o segundo se refere à energia decorrente da interação elétron-núcleo. O terceiro termo corresponde ao potencial de Coulomb, isto é, a energia de interação elétron-elétron e o último termo é a energia de troca ou *exchange* de um gás de elétrons livres. A densidade eletrônica é representada por ρ , o potencial devido aos núcleos é denotado por $v(r)$ e as coordenadas são notadas por r .

A introdução do conceito de energia como funcional da densidade ρ , juntamente com a interpre-

⁵ $N_{\text{átomos}}$ designa o número de átomos que formam um sistema molecular. No caso em que $N_{\text{átomos}} \geq 20$, os sistemas são considerados de moderados a grandes.

tação de Slater para o termo de *exchange*, o qual depende principalmente da densidade eletrônica do sistema, foram essenciais para a formulação da DFT, uma vez que Hohenberg e Kohn se basearam em tais idéias para desenvolver os fundamentos da mesma. Sendo assim, em 1964 e 1965 os teoremas de Hohenberg, Kohn e Sham foram propostos em dois artigos, onde as concepções básicas da DFT foram enunciadas. Esta teoria causou um grande impacto nos cálculos de propriedades de moléculas e sólidos e apresenta alta aplicabilidade (que continua em expansão) em diversos problemas. Assim, devido à importância da DFT, em 1998 Walter Kohn dividiu o prêmio Nobel de Química com John Pople⁶ pelo desenvolvimento desta teoria.

2.3.1 Teoremas de Hohenberg e Kohn

A DFT apoia-se sobre dois teoremas propostos por Hohenberg e Kohn [69] que estabelecem a existência de um funcional único da densidade e que a energia do estado fundamental é mínima para esta densidade. Entretanto, antes de enunciar e analisar tais teoremas, deve-se lembrar que o hamiltoniano de um sistema eletrônico molecular, considerando a aproximação de BO e desprezando os efeitos relativísticos é,

$$\hat{H}_{BO} = - \sum_i^N \frac{1}{2} \nabla_i^2 - \sum_A^M \sum_i^N \frac{Z_A}{|R_A - r_i|} + \sum_{i < j}^N \sum_j^N \frac{1}{|r_i - r_j|} \quad (2.62)$$

onde os índices i e j se referem aos elétrons, enquanto que A e B representa os núcleos. Já as coordenadas dos elétrons e dos núcleos são dadas por r_i e R_A , respectivamente⁷.

Da equação 2.62, o operador energia cinética \hat{T} corresponde ao primeiro termo, ao passo que o operador da energia potencial \hat{U} é representado pelo segundo termo. O terceiro termo configura os operadores de interação repulsiva elétron-elétron \hat{V}_e . Dessa forma, reescrevendo o hamiltoniano

⁶John Pople recebeu o prêmio Nobel de química pelo desenvolvimento de métodos computacionais em química quântica.

⁷A equação 2.62 está expressa em unidades naturais.

eletrônico em termos dos operadores, obtemos

$$\hat{H} = \hat{T} + \hat{U} + \hat{V}_e \quad (2.63)$$

de onde, pode-se escrever a equação de Schrödinger eletrônica 2.1 como,

$$\hat{H}\Psi = (\hat{T} + \hat{U} + \hat{V}_e)\Psi = E\Psi \quad (2.64)$$

Por sua vez, a densidade eletrônica é definida pela equação:

$$\rho(r) = \int \cdots \int \Psi(r_1, r_2, \cdots, r_N)^* \Psi(r_1, r_2, \cdots, r_N) dr_2 dr_3 \dots dr_N \quad (2.65)$$

onde $\Psi(r_1, r_2, \cdots, r_N)$ deve ser a solução para o estado fundamental. Assim, a energia total do sistema é determinada por,

$$E = \int \Psi(r_1, r_2, \cdots, r_N)^* \hat{H}_{BO} \Psi(r_1, r_2, \cdots, r_N) dr_1 dr_2 \dots dr_N = \langle \Psi | \hat{H}_{BO} | \Psi \rangle. \quad (2.66)$$

Separando o potencial externo como um funcional da densidade, a energia total fica:

$$E_0 = \langle \Psi | (\hat{T} + \hat{V}_e) | \Psi \rangle + \int \rho(r) v(r) dr. \quad (2.67)$$

Assim o hamiltoniano do sistema é definido pelo número de elétrons e pelo potencial externo em que estes se movimentam.

Teorema 2.3.1 (Primeiro Teorema de Hohenberg e Kohn). *O potencial externo $v(\mathbf{r})$ sentido pelos elétrons é um funcional único da densidade eletrônica $\rho(\mathbf{r})$.*

Assim sendo, considerando que para um sistema cujo estado fundamental seja Ψ , o hamiltoniano \hat{H} que descreve o sistema possui um potencial externo $v(r)$. Para o caso em que há outro potencial externo dado por $v'(r)$, o hamiltoniano correspondente será \hat{H}' e o estado fundamental

Ψ' . Por outro lado, consideremos que a densidade $\rho(r)$ deve ser igual para ambos os casos, isto é, para $v(r)$ e $v'(r)$. Desse modo, de acordo com o teorema variacional, temos,

$$\langle \Psi | \hat{H} | \Psi \rangle < \langle \Psi' | \hat{H} | \Psi' \rangle = \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | (\hat{U} - \hat{U}') | \Psi' \rangle \quad (2.68)$$

Usando a expressão para o potencial externo,

$$\hat{U} = \sum_i^N v(r_i) \quad (2.69)$$

e integrando seu valor esperado em relação às posições eletrônicas, a equação 2.68 se torna,

$$E < E' + \int [v(r) - v'(r)] \rho(r) d^3r \quad (2.70)$$

Procedendo do mesmo modo, para $\langle \Psi' | \hat{H}' | \Psi' \rangle$, então

$$E' < E + \int [v'(r) - v(r)] \rho(r) d^3r \quad (2.71)$$

Somando 2.70 e 2.71, obtem-se:

$$E + E' < E' + E \quad (2.72)$$

Porém, como a densidade eletrônica deve ser a mesma para $v(r)$ e $v'(r)$, a equação não é viável, pois $\Psi \neq \Psi'$. Assim, para evitar este conflito, assumir a unicidade da densidade $\rho(r)$ implica em considerar $\Psi = \Psi'$. Portanto, a densidade do estado fundamental $\rho(r)$ deve conter todas as informações da função de onda deste estado. Dito isto, como o valor esperado de um dado observável físico (designado pelo operador \hat{O}) é determinado do seguinte modo:

$$O = \langle \Psi | \hat{O} | \Psi \rangle = O[\rho(r)] \quad (2.73)$$

Então este operador deverá ser um funcional único da densidade.

Teorema 2.3.2 (Segundo Teorema de Hohenberg e Kohn). *A energia do estado fundamental $E_0[\rho]$ é mínima para a densidade $\rho(r)$ exata,*

$$E[\rho] = \langle \Psi | (\hat{T} + \hat{U} + \hat{V}_e) | \Psi \rangle \quad (2.74)$$

O segundo teorema define a energia como sendo um funcional da densidade eletrônica ρ , cujo valor mínimo deve ser fornecido pela densidade do estado fundamental. Para isto, considerando ρ' a densidade para um considerado estado Φ e ρ a densidade para o estado fundamental Ψ , então a equação 2.74 pode ser escrita como:

$$E[\rho'] = \langle \Psi' | (\hat{T} + \hat{U} + \hat{V}_e) | \Psi' \rangle E_0[\rho] = \langle \Psi | (\hat{T} + \hat{U} + \hat{V}_e) | \Psi \rangle \quad (2.75)$$

De onde, através do teorema variacional é possível demonstrar que,

$$E_0[\rho] \leq E[\rho'] \quad (2.76)$$

onde a energia de um estado qualquer deverá ser igual à do estado fundamental se as densidades eletrônicas forem iguais.

De acordo com os teoremas de HK, o estado fundamental de um sistema para um dado potencial externo pode ser obtido utilizando-se a densidade eletrônica tridimensional como variável básica. Assim, o funcional da energia pode ser calculada a partir da definição de um *funcional universal*, dado por:

$$F[\rho] = \langle \Psi | (\hat{T} + \hat{V}_e) | \Psi \rangle \quad (2.77)$$

com \hat{T} e \hat{V}_e válidos para qualquer sistema coulombiano. Logo, a equação 2.74 pode ser escrita

como,

$$E[\rho] = F[\rho] + \langle \Psi | \hat{U} | \Psi \rangle \quad (2.78)$$

onde o potencial externo \hat{U} depende do sistema eletrônico em questão. Por outro lado, visto que as interações de Coulomb são de longo alcance, convém separar a contribuição coulombiana clássica do funcional universal, de modo que,

$$F[\rho] = \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{r_1 - r_2} dr_1 dr_2 + G[\rho] \quad (2.79)$$

Assim, o funcional da energia se torna,

$$E[\rho] = G[\rho] + \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} + \int \rho(r)v(r)dr \quad (2.80)$$

onde $G[\rho]$ é um novo funcional universal da densidade, definido como,

$$G[\rho] \equiv T_0[\rho] + E_{xc}[\rho] \quad (2.81)$$

para $T_0[\rho]$ sendo o funcional de energia cinética de um sistema eletrônico sem interação entre os elétrons e com densidade ρ . O termo $E_{xc}[\rho]$ é a energia de *exchange* e correlação de um sistema eletrônico interagente com densidade ρ e inclui também a correção da energia cinética do referido sistema. Dessa forma, ao definir esta primeira aproximação para a forma de $G[\rho]$, Kohn e Sham introduziram a estratégia de usar um sistema de partículas independentes como referência, para o cálculo da estrutura eletrônica de sistemas de muitas partículas. Por fim, vale ressaltar que o funcional $E_{xc}[\rho]$ não possui uma fórmula analítica, sendo necessário buscar formas aproximadas para este termo.

Considerando o funcional de energia,

$$E[\rho] = T_0[\rho] + \int \rho(r) E_{xc}(\rho(r)) dr + \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} + \int \rho(r)v(r) dr, \quad (2.82)$$

estabelecendo que a condição de que a carga eletrônica total seja fixa, isto é,

$$\int \rho(r) dr = N \quad (2.83)$$

então pelo princípio variacional, o funcional de energia deve satisfazer a condição de extremo,

$$\delta \left(E[\rho] - \mu \left[\int \rho(r) dr - N \right] \right) = 0. \quad (2.84)$$

Assim, utilizando a definição de derivada funcional a um funcional da densidade $F[\rho]$ em relação a uma variação $\delta\rho(r)$:

$$\frac{\delta F[\rho]}{\delta\rho(r)} = F[\rho + \delta\rho(r)] - F[\rho(r)] = \int \frac{\delta F[\rho]}{\delta\rho(r)} \delta\rho(r) dr \quad (2.85)$$

sendo que o funcional a ser derivado é o próprio funcional da densidade, dado pela equação 2.82, chegamos à seguinte expressão:

$$\int \delta\rho(r) \left\{ \frac{\delta T_0}{\delta\rho} + v(r) + \int \frac{\rho(r_1)}{|r - r_1|} dr + v_{xc}[\rho] - \mu \right\} dr = 0, \quad (2.86)$$

em que o termo conhecido como potencial de exchange-correlação, $v_{exc}[\rho]$, é dado por:

$$v_{exc}[\rho] = \frac{\delta E_{xc}}{\delta\rho}. \quad (2.87)$$

Dada a densidade de carga auxiliar,

$$\rho(r) = \sum_{i=1}^N |\psi_i(r)|^2, \quad (2.88)$$

a equação 2.86 pode ser resolvida através da equação de Schrödinger de uma partícula,

$$\left(-\frac{1}{2}\nabla^2 + v^{KS}[\rho]\right)\psi_i^{KS} = \varepsilon_i\psi_i^{KS} \quad (2.89)$$

onde, a energia cinética é calculada de forma autoconsistente, por meio da seguinte equação:

$$T_s[\rho] = \sum_i^N \left\langle \psi_i^{KS} \left| -\frac{1}{2}\nabla_i^2 \right| \psi_i^{KS} \right\rangle. \quad (2.90)$$

O termo,

$$v^{KS} = v(r) + \int \frac{\rho(r_1)}{|r - r_1|} dr_1 + v_{xc}(\rho) \quad (2.91)$$

é o potencial efetivo de Kohh-Sham, o qual é uma função de da variável espacial r e deve ser local, isto é, não depende dos valores de $v_{ef}(r)$ em outros pontos do espaço, r_1 .

A equação 2.89 pode ser reescrita em termos do hamiltoniano de Kohn-Sham, \hat{h}^{KS} , como

$$\hat{h}^{KS}\psi_i^{KS} = \varepsilon_i\psi_i^{KS} \quad (2.92)$$

que é conhecida como equação de Kohn-Sham (KS) e é resolvida iterativamente, posto que o potencial efetivo v^{KS} depende da densidade, que por sua vez depende de v^{KS} . Portanto, os orbitais de Konh-Sham, ψ_i^{KS} , são obtidos através de um cálculo autoconsistente.

Por sua vez, a energia eletrônica total é determinada por:

$$E = \sum \varepsilon_i - \frac{1}{2} \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2 + E_{xc}[\rho] - \int \rho(r)v_{xc}dr \quad (2.93)$$

onde ε_i são os autovalores da equação de KS. Assim, o esquema autoconsistente de KS fornece a densidade eletrônica do estado fundamental, a partir da qual todas as outras propriedades do sistema podem ser obtidas.

2.3.2 Funcionais de Correlação e Troca

No método DFT não é conhecida uma forma exata para os funcionais de correlação e troca $E_{xc}[\rho]$. Logo, algumas aproximações são realizadas para este funcional, de modo que a dependência deste com a densidade eletrônica é dada por:

$$E_{xc}[\rho] = \int \rho(\mathbf{r})\varepsilon_{xc}[\rho(\mathbf{r})]d^3\mathbf{r} \quad (2.94)$$

A quantidade $\varepsilon_{xc}[\rho(\mathbf{r})]$ representa a densidade de energia e pode ser expressa como,

$$\varepsilon_{xc}[\rho(\mathbf{r})] = \varepsilon_x[\rho(\mathbf{r})] + \varepsilon_c[\rho(\mathbf{r})], \quad (2.95)$$

ou seja, o termo de troca é dado por $\varepsilon_x[\rho(\mathbf{r})]$, enquanto que $\varepsilon_c[\rho(\mathbf{r})]$ representa o termo de correlação eletrônica.

Entre as aproximações para o termo de correlação e troca mais utilizados, estão a Aproximação da Densidade Local (LDA)⁸ e Aproximação do Gradiente Generalizado (GGA)⁹, sendo esta último uma versão mais refinada da LDA.

Aproximação LDA

A aproximação LDA, foi proposta por Kohn e Sham, podendo descrever adequadamente sistemas cuja densidade eletrônica seja aproximadamente uniforme [70]. Para este funcional é considerado que em cada ponto \mathbf{r} a densidade eletrônica $\rho(\mathbf{r})$ é constante. Isto posto, a energia de correlação e troca é dada por:

⁸LDA é a abreviação para o termo em inglês Local Density Approximation

⁹GGA é o acrônimo do termo em inglês Generalized Gradient Approximations

$$E_{xc}^{LDA} = \int \rho(\mathbf{r}) \varepsilon_{xc}^h[\mathbf{r}] d^3\mathbf{r}, \quad (2.96)$$

com $\varepsilon_{xc}^h[\rho(\mathbf{r})]$ sendo a energia de correlação e troca por elétron, em um gás homogêneo de elétrons interagentes, enquanto que $\rho(\mathbf{r})$ a densidade eletrônica dada no ponto \mathbf{r} .

A energia por elétron pode ser dividida em dois termos correspondentes ao termo de troca e ao de correlação, isto é:

$$\varepsilon_{xc}^h = \varepsilon_x^h + \varepsilon_c^h \quad (2.97)$$

onde ε_x^h é calculado analiticamente através de:

$$\varepsilon_x^h(\mathbf{r}) = -\frac{3}{4} \left(\frac{3\rho}{\pi} \right)^{1/3} \quad (2.98)$$

ao passo que obtém-se ε_c^h de forma aproximada¹⁰ através da expressão,

$$\varepsilon_c^h = -\frac{0,44}{r_s + 7,8}, \quad (2.99)$$

sendo r_s (o raio da esfera) dado por:

¹⁰A melhor aproximação para o termo de correlação foi obtida por Ceperley e Adler em 1980.

$$r_s = \left(\frac{3}{4\pi\rho} \right)^{1/3} \quad (2.100)$$

Aproximação GGA

A aproximação GGA [71] deve ser utilizada para sistemas cuja densidade eletrônica não seja uniforme. Os erros devido à heterogeneidade do meio são corrigidos pela inclusão da dependência do funcional da densidade de energia ε_{xc} com o gradiente da densidade eletrônica. Assim, o funcional assume a seguinte forma:

$$E_{xc}^{GGA} = \int \varepsilon_{xc}^{GGA}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})] d^3\mathbf{r},$$

to

(2.102)

onde $\varepsilon_{xc}^{GGA}[\rho(\mathbf{r})]$ possui dependência com $\rho(\mathbf{r})$ (a densidade de energia) e com $\nabla\rho(\mathbf{r})$ (o gradiente da densidade).

Funcionais Híbridos

Os funcionais híbridos são tipos de funcionais que mesclam parte do funcional de troca do método HF com partes de outros funcionais de correlação e troca do método DFT. Assim, estes tipos de funcionais são escritos da seguinte forma:

$$E_{xc} = a_0 E_{xc}^{HF} + \sum_{i>0} a_i E_{xc}^i, \quad (2.103)$$

Com a_0 e a_i representando os parâmetros da fração do termo de troca para o método HF e correlação e troca para o DFT, nesta ordem. Tais parâmetros podem ser determinados semi-empiricamente por meio de métodos experimentais.

Nos cálculos de energia realizados no presente trabalho, foram utilizados os funcionais híbridos B3LYP e CAM-B3LYP. O primeiro (B3LYP) é dado pela seguinte expressão:

$$E_{xc}^{B3LYP} = (1 - a) E_x^{LDA} + a E_x^{HF} + b \Delta E_x^{B88} + (1 - c) E_c^{VWN} + c E_c^{LYP}, \quad (2.104)$$

onde $a = 0,20$, $b = 0,72$ e $c = 0,81$ são parâmetros semi-empíricos do funcional B3PW91 proposto por Becke [72].

No caso do funcional híbrido CAM-B3LYP, são incluídas correções de longo alcance por meio atenuação dos efeitos de Coulomb [73]. E a correção de longo alcance na energia (E_x^{lr}) é dado por:

$$E_x^{lr} = -\frac{1}{2} \sum_{\sigma} \sum_i^{occ} \sum_j^{occ} \int \int \phi_{i\sigma}^*(\mathbf{r}_1) \phi_{j\sigma}^*(\mathbf{r}_1) \times \frac{erf(\mu)r_{12}}{r_{12}} \phi_{i\sigma}(\mathbf{r}_2) \phi_{j\sigma}(\mathbf{r}_2) \quad (2.105)$$

onde $\phi_{i\sigma}$ é o i -ésimo spin-orbital molecular e o parâmetro μ determina o balanço entre os funcionais de troca DFT e HF na distância r_{12} .

2.4 O método MFCC (Molecular Fractionation with Conjugate Caps)

O método MFCC (Molecular Fragmentation with Conjugate Caps) é uma aproximação de escala linear capaz de fornecer cálculos quantitativos e eficientes da energia de interação de proteínas com compostos moleculares através de métodos *ab initio*, DFT ou com nível mais alto, como métodos perturbativos por exemplo, os quais exigiriam alto custo computacional para sistemas macromoleculares (da ordem de milhares de átomos). Como regra geral, o custo computacional de um método mecânico quântico promissor para sistemas moleculares grandes deve ser linearmente escalar em relação ao tamanho do sistema a ser tratado. Uma abordagem de escala linear baseia-se principalmente na propriedade local da interação de elétrons, porque o efeito da perturbação de energia em uma área é geralmente localizado em sua vizinhança e decai rapidamente se afastando dela. Nessa abordagem, os métodos de divisão e conquista (DAC)¹¹ e similares são comumente empregados em cálculos teóricos [54, 74, 75, 76, 77]. Isto é, a idéia principal é dividir a energia de interação de pequenas moléculas com macromoléculas (como uma proteína, por exemplo), em somas individuais de interação que possam ser calculadas por métodos baseados em mecânica quântica [56].

Proposto originalmente por Zhang *et al* [78], o método MFCC consiste na fragmentação da proteína a partir da quebra das ligações peptídicas, e no lugar em que a quebra das ligações é feita são introduzidos os caps¹² conjugados, com o intuito de preservar as propriedades da ligação de valência e simular o efeito da parte molecular cortada na parte eletrônica do fragmento, isto é, os caps conjugados devem na medida do possível, mimetizar o ambiente protéico em torno da porção fragmentada.

Na metodologia MFCC de Zhang, a energia de interação entre uma molécula M e a proteína

¹¹DAC é a abreviatura do termo inglês Divide and Conquer, que se refere ao paradigma em que algoritmos são construídos para resolver problemas maiores, dividindo-os em problemas menores.

¹²O termo em inglês caps pode ser traduzido como "capas" e se refere a um grupo de átomos ou radicais inseridos nas extremidades do fragmento.

P , denotada por $V(M - P)$ é calculada através da seguinte expressão:

$$V(M - P) = \sum_i^N V(M - C_{ap}^{i-1} * A_i C_{ap}^i) - \sum_i^{N-1} V(M - C_{ap}^i * C_{ap}^i) \quad (2.106)$$

onde o primeiro termo representa a interação entre a molécula M e o fragmento encapsado $V(M - C_{ap}^{i-1} * A_i C_{ap}^i)$, isto é, um dado aminoácido A_i unido covalentemente aos caps conjugados $C_{ap}^{i-1} * C_{ap}^i$. Já o segundo termo expressa a energia de interação entre a molécula M (ligante) e os caps conjugados. Desta forma, através da equação 2.106, é possível obter a energia de interação entre a proteína e uma molécula ligante por meio da soma das energias individuais dos fragmentos, as quais podem ser calculadas via métodos baseados em química quântica, tais como métodos *ab initio*, DFT e até níveis mais sofisticados.

2.4.1 Estratégias para o refinamento do MFCC

De um modo geral, através do MFCC o problema de se calcular a energia entre o ligante e a proteína se resume a cálculos individuais da energia de interação do ligante com cada aminoácido componente da proteína em questão, acoplado com seus caps, diminuindo consideravelmente o tamanho do sistema a ser estudado. Desta forma, normalmente cada cálculo possui um número reduzido de átomos, possibilitando assim o uso de métodos quânticos [79].

Interações em sistemas macromoleculares tem sido amplamente estudados através do esquema de fragmentação molecular MFCC [80, 81, 82, 83]. Por outro lado, a necessidade de incluir efeitos não locais na interação dos aminoácidos com as moléculas ligantes e tornar a descrição dos sistemas mais realísticas levaram ao desenvolvimento de outras metodologias, que embora mantenham a idéia principal do MFCC original, incorporam modelos e técnicas que pretendem aumentar a eficiência dos cálculos da energia de interação de sistemas complexos como proteínas-ligantes.

As variações do método de Zhang *et al* [78] incluem efeitos negligenciados nesta primeira aproximação. Tais metodologias procuram melhorar a descrição da interação entre proteínas e ligantes, citando-se o aprimoramento na mimetização do meio protéico e o aumento do efeito da blindagem eletrostática nos sítios ativos através da inserção de vários caps conjugados [84], assim como o uso de modelos de solvatação contínua com constantes dielétricas típicas para meios protéicos [85] e a incorporação de efeitos eletrostáticos [86]. Em seguida, alguns métodos de fragmentação modificados serão sucintamente apresentados.

2.4.2 Aproximação MFCC-CPCM

A aproximação GMFCC-CPCM acopla o uso da estratégia MFCC com o método de solvatação eletrostática, conhecido como *conductor-like polarizable continuum model* (CPCM). Neste tipo de cálculo, a proteína é tratada como soluto, que por sua vez é decomposta em fragmentos via MFCC. O potencial eletrostático na interface soluto-solvente é calculado através da soma das contribuições de cada fragmento, subtraindo-se a interação dos caps conjugados. Dito isto, as cargas induzidas na interface soluto-solvente é resolvida iterativamente através do modelo da cavidade condutora COSMO ¹³. Finalmente, a carga induzida é utilizada como um background de cargas no qual o soluto terá sua estrutura eletrônica calculada via MQ [87].

Considerando que o hamiltoniano de uma molécula de soluto em solução seja,

$$H = H_0 + H', \quad (2.107)$$

com H_0 representando a fase gasosa do soluto, enquanto que H' representa a perturbação causada pelo solvente, dada por:

$$H' = \sum_{\mu,\alpha} \frac{q_\mu Z_\alpha}{|r_\mu - R_\alpha|} - \sum_{\mu,i} \frac{q_\mu}{|r_\mu - r_i|}. \quad (2.108)$$

¹³COSMO é a sigla para o termo em inglês *conductor-like screening model*

onde as cargas superficiais são representadas por q_μ e suas respectivas distâncias por r_μ . Já as cargas nucleares são denotadas por Z_α e suas posições são representadas por R_α . As coordenadas do i -ésimo elétron é dada por r_i . Assim sendo, a energia de solvatação eletrostática ($G(ele)$) é dada por:

$$G(ele) = E(wfd) + G(es). \quad (2.109)$$

Na equação 2.109, o primeiro termo do lado direito se refere á energia de polarização do soluto, onde a sigla "wfd" é a sigla para wave function distortion, isto é, a distorção da função de onda e o último termo é a energia de interação eletrostática entre soluto e solvente.

O método CPCM define uma cavidade que envolve esferas centradas nos átomos do soluto. A superfície da cavidade é formada por porções caracterizadas pela sua posição, área e vetor normal [88]. Assim, o meio solvente contínuo é polarizado pelo soluto ao passo que ocorre uma indução de cargas na superfície. Esta superfície discreta de cargas induzidas q é calculada através da resolução da equação linear,

$$B + Aq = 0, \quad (2.110)$$

onde a carga q_μ é o produto da densidade de carga σ_μ e a área da superfície S_μ , isto é,

$$q_\mu = \sigma_\mu S_\mu. \quad (2.111)$$

Os elementos da matriz A (eq. 2.110) são dados por:

$$A_{\mu\nu} = \frac{1}{|r_\mu - r_\nu|} (1 - \delta_{\mu\nu}) + 1.07 \sqrt{\frac{4\pi}{S_\mu}} \delta_{\mu,\nu} \quad (2.112)$$

Já os elementos da matriz B são dados por:

$$B_{\mu} = \sum_{\alpha} \frac{Z_{\alpha}}{|r_{\mu} - R_{\alpha}|} - \phi(r_{\mu}), \quad (2.113)$$

em que o potencial eletrostático na superfície do sítio μ é representado por $\phi(r_{\mu})$. O termo r_{μ} nos dá a posição desta superfície do sítio μ criada pelos elétrons do soluto.

Vale ressaltar que a equação 2.110 é válida somente para condutores com constante dielétrica infinita. Porém, a superfície de cargas é corrigida pela constante dielétrica finita através da equação,

$$\tilde{q} = \frac{\varepsilon - 1}{\varepsilon} q_{\mu}, \quad (2.114)$$

com ε sendo a constante dielétrica do solvente. No caso de macromoléculas, em que são definidas largas superfícies, é necessário resolver a equação linear 2.110 via métodos iterativos [88].

2.4.3 Aproximação EE-GMFCC

Uma das estratégias de melhoramento do MFCC é a aproximação EE-GMFCC (Electrostatically Embedded Generalized Molecular Fractionation with Conjugate Caps), que combina a metodologia de fragmentação molecular (MFCC) generalizada com a incorporação dos efeitos eletrostáticos [89, 86]. A energia total da proteína é calculada por meio de uma combinação linear da energia obtida via MQ dos resíduos vizinhos com a energia de interação (MQ) entre dois-corpos de resíduos não-vizinhos que estejam espacialmente em contato [90]. Todos os fragmentos dos cálculos são introduzidos em um campo de cargas pontuais representando o meio protéico restante. Essa aproximação baseia-se em um método proposto por He *et al.* [90] denominado GMFCC/MM, em que os campos de forças moleculares são inseridos para representar as interações de longo alcance entre os fragmentos não-vizinhos mais distantes, enquanto que as interações de curto alcance são calculadas via MQ. No entanto, a aproximação GMFCC/MM não inclui o efeito do meio eletrostático nos cálculos dos fragmentos.

Diversos estudos têm mostrado que a polarização eletrônica advinda do meio é importante para incluir o efeito da interação de muitos corpos em modelos de fragmentação [89, 91, 92, 93, 94]. Por sua vez, o método EE-GMFCC adota estes efeitos a partir da incorporação do campo de cargas pontuais no cálculo da energia de cada fragmento, com o objetivo de mimetizar o meio eletrostático do sistema remanescente. Este método é capaz de descrever, satisfatoriamente, a estrutura eletrônica de grandes sistemas de proteínas nos níveis de cálculos HF, DFT e MP2 [86].

No EE-GMFCC, a energia total da proteína é dada pela expressão,

$$E_{EE-GMFCC} = \sum_{i=2}^{N-1} (C_{i-1}^* R_i C_{i+1}) - \sum_{i=2}^{N-2} (C_i^* C_{i+1}) + \sum_{i,j < i+2} (\tilde{E}_{ij} - \tilde{E}_i - \tilde{E}_j)_{QM} - \left\{ \sum_{k,l} \sum_{m,n} \frac{q_{m(k)} q_{n(l)}}{R_{m(k)n(l)}} - \sum_{i,j > i+2} \sum_{m',n'} \frac{q_{m'(i)} q_{n'(j)}}{R_{m'(i)n'(j)}} \right\}, \quad (2.115)$$

onde \tilde{E} representa a soma sobre a autoenergia do fragmento e a energia de interação entre o fragmento e background de cargas dos sistema restante. As energias dos resíduos i e j , cujas ligações peptídicas foram substituídas por hidrogênios, são denotadas por \tilde{E}_i e \tilde{E}_j , respectivamente. Já o termo \tilde{E}_{ij} representa a energia do dímero formado pelos resíduos i e j . As cargas pontuais referente ao m -ésimo átomo do fragmento k são denotadas pelos termos $q_{m(k)}$.

2.4.4 Aproximação EE-GMFCC-CPCM

Na aproximação EE-GMFCC-CPCM a distorção da função de onda é obtida através da expressão,

$$E_{(wfd)} = \sum_{k=1}^N \delta \tilde{E}_k - \sum_{k=1}^{N_c} \delta \tilde{E}_k^c + \sum_{k=1}^{N_{Gc}} \delta (\tilde{E}_k^{ij} - \tilde{E}_k^i - \tilde{E}_k^j) \quad (2.116)$$

onde \tilde{E} representa a soma das autoenergias de cada fragmento e seu correspondente background de cargas. Os termos N , N_c e N_{Gc} denota o número de resíduos individuais anexados aos seus

respectivos caps, concaps e Gconcaps, nessa ordem. Já o termo $\delta(\tilde{E}_k$ se refere à energia das funções de onda distorcidas dos fragmentos individuais, à medida que a energia da distorção da função de onda dos caps conjugados é denotada por $\delta(\tilde{E}_k^c$. A interação de dois-corpos com resíduos não-vizinhos colocados em contato próximo é denotada por $\delta(\tilde{E}_k^{ij} - \tilde{E}_k^i - \tilde{E}_k^j)$.

No que diz respeito à energia da interação eletrostática soluto-solvente $G(es)$, a mesma é expressa por,

$$G(s) = \frac{1}{2} \sum_{\mu} q_{\mu} \phi(r_{\mu}), \quad (2.117)$$

em que o potencial eletrostático ϕ_{μ} da superfície generalizada, em decorrência dos elétrons e núcleos do soluto, é agora corrigido pela adição da interação de dois-corpos. Assim,

$$\phi(r_{\mu}) = \sum_{k=1}^N \phi_k(r_{\mu}) - \sum_{k=1}^{N_c} \phi_k^c(r_{\mu}) + \sum_{k=1}^{N_{Gc}} [\phi_k^{ij}(r_{\mu}) - \phi_k^i(r_{\mu}) - \phi_k^j(r_{\mu})]. \quad (2.118)$$

Portanto, na abordagem EE-GMFCC-CPCM a correção de dois-corpos é introduzida na expressão da energia total e no potencial eletrostático da superfície da cavidade. Desta maneira, esta sistematização aumenta o nível de precisão da energia de distorção da função de onda, assim como da energia de interação eletrostática entre soluto e solvente. A partir de métodos da MQ, cada fragmento tem sua energia calculada na presença das cargas superficiais e a nova função de onda obtida é utilizada para determinar as cargas superficiais por meio da equação 2.110. Logo, o processo se repete iterativamente até a carga superficial da cavidade convergir e a energia de interação eletrostática soluto-solvente final ser encontrada [89].

2.4.5 Aproximação EE-AMFCC

No presente trabalho, as energias de interação foram calculadas usando uma adaptação do esquema MFCC que inclui o efeito das interações eletrostáticas, EE-GMFCC [89, 86]. Neste caso,

a diferença entre a presente metodologia e a metodologia EE-GMFCC, reside no fato de que no corrente método (EE-AMFCC)¹⁴, as interações são inseridas por meio das cargas pontuais dos átomos do sistema remanescente em cada termo dos cálculos de energia. Isto é, em todos os termos de energia os átomos dos aminoácidos implícitos são inseridos através de suas respectivas cargas atômicas, considerando-se assim um background de cargas pontuais e sua contribuição eletrostática na energia de interação. Desta forma, a expressão para a energia de interação ($E_{(L-R_i)}$) entre um dado resíduo de interesse (R_i) e o ligante (L) é dada por:

$$E(L - R_i) = E(L - C_{i-1}R_iC_{i+1} - C_j^{cp}) - E(C_{i-1}R_iC_{i+1} - C_j^{cp}) - E(L - C_{i-1}C_{i+1} - C_j^{cp}) + E(C_{i-1}C_{i+1} - C_j^{cp}), \quad (2.119)$$

com C_{i+1} e C_{i-1} denotando os caps sucessor e antecessor, respectivamente. O primeiro termo do lado direito da equação 2.119 se refere à energia de interação entre o ligante (L) e o resíduo de interesse (R_i) acoplado com seus respectivos caps (C_{i-1} e C_{i+1}), o segundo termo representa somente a interação do resíduo de interesse com seus caps conjugados, o terceiro termo representa a energia do ligante com os caps e o quarto termo representa apenas a energia de interação dos caps. O índice cp representa as cargas pontuais, enquanto que j caracteriza o conjunto de resíduos em torno do resíduo referência R_i , sem considerar os Caps, onde cada átomo é substituído pela sua carga elétrica. Assim sendo, o conjunto de cargas pontuais, denotado por C_j^{cp} é introduzido em todos o termos da equação 2.119.

¹⁴EE-AMFCC é a sigla para o termo inglês Electrostatically Embedded Adaptation of MFCC.

Capítulo 3

Aplicação: Interação do Ibuprofeno (IBU) com a Albumina do Soro Humano

No presente capítulo, será apresentado o sistema macromolecular adotado para a aplicação da metodologia proposta neste trabalho. Tal metodologia, composta por métodos quânticos utilizados em conjunto com uma estratégia de fragmentação molecular MFCC, será exposta em detalhes neste capítulo. Todavia, primeiramente serão apresentados os interesses biológicos e as principais características estruturais da proteína Albumina do Soro Humano (ASH). Do mesmo modo, será exposto um breve levantamento das propriedades farmacológicas do ligante Ibuprofeno (IBU). Em seguida, será feita uma descrição minuciosa da metodologia empregada para obter as energias de interação entre a molécula ligante e os aminoácidos do sítio ativo (vizinhança) da proteína. Ao final do capítulo, os resultados serão exibidos, e oportunamente explanados, levando-se em consideração

o escopo e os métodos explorados neste estudo.

3.1 A proteína Albumina do Soro Humano (ASH)

A Albumina do Soro Humano (ASH) é a proteína mais abundante no plasma sanguíneo, representando cerca de 50% a 60% das proteínas constituintes do sangue humano [95]. Ela é sintetizada no fígado e possui uma concentração típica de 5 g/100ml [96, 97]. A ASH é uma proteína versátil, e além de servir como uma proteína entreposto, ela também possui vasta capacidade de transportar diferentes tipos de moléculas insolúveis em água, como os ácidos graxos por exemplo. A ASH pode se ligar a vários tipos de compostos endógenos e exógenos, influenciando fortemente como eles são transportados pelo corpo humano [98].

Uma das principais funções da ASH é manter o pH e a pressão osmótica do plasma sanguíneo. Porém, podemos destacar que sua capacidade de ligar-se reversivelmente a uma grande variedade de ligantes e fazer o transporte de fármacos aos órgãos alvos, pode ser pensado como um dos papéis fisiológicos mais interessantes que a ASH desempenha [96]. Além disso, a presença abundante da ASH no sangue é um fator importante na farmacocinética de muitas drogas, afetando diretamente a taxa de transporte do ligante, tal como sua eficácia [99]. A importância fisiológica e uma relativa facilidade para isolar e purificar a ASH em larga escala levaram a um grande número de estudos envolvidos na tentativa de compreender suas propriedades fisiológicas e farmacológicas, bem como sua alta afinidade com diferentes estruturas moleculares, incluindo metais, ácidos graxos e uma grande variedade de drogas [95, 96, 98, 99, 100, 101].

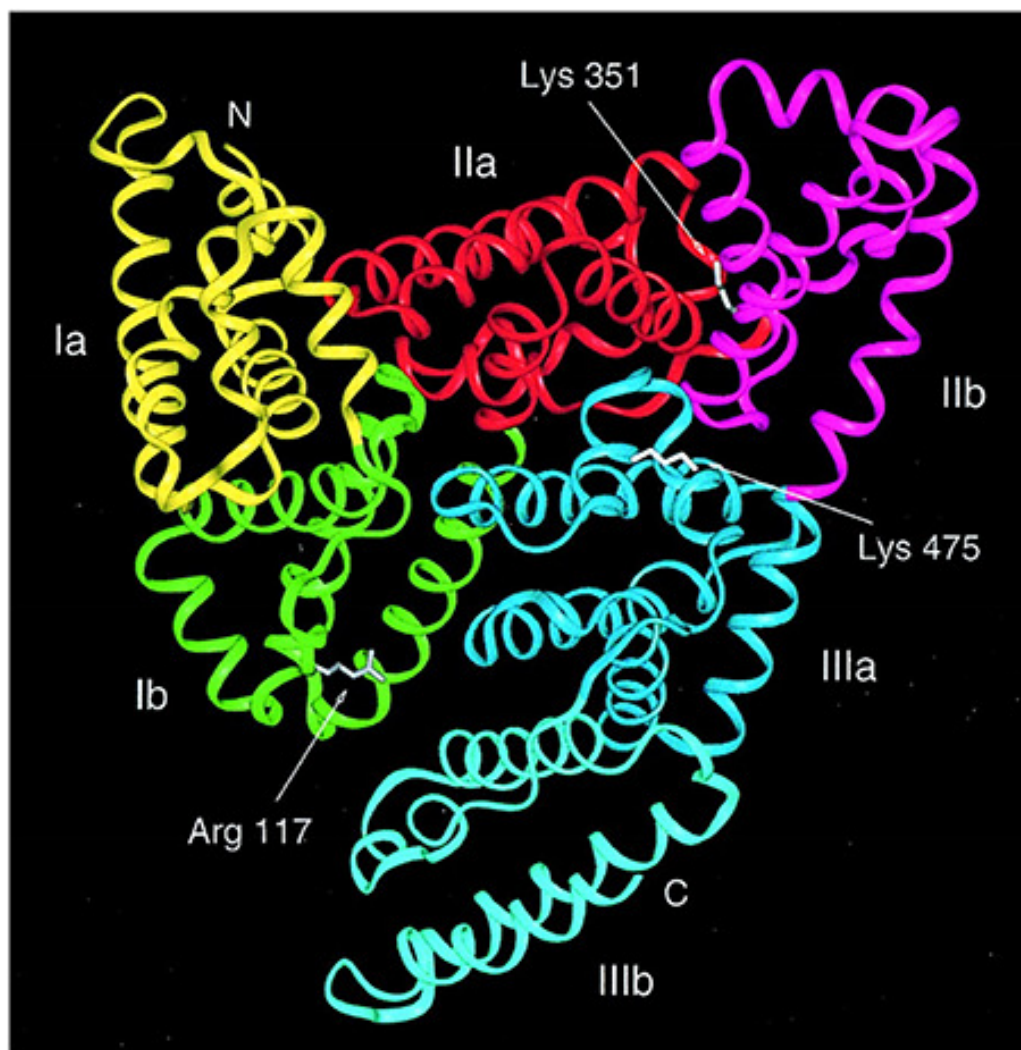
3.1.1 Características estruturais da ASH

Classificada como uma proteína com baixo peso molecular, a albumina possui uma massa molecular em torno de 66, 348 Da [102]. Esta proteína apresenta uma estrutura primária simples que é composta por três domínios homólogos numerados em I, II e III, os quais são agrupados em sub-

domínios A e B (conforme Figura 3.1). Cada tipo de subdomínio (A ou B) apresenta características estruturais similares entre si [96, 99].

A sequência primária da HSA apresenta uma cadeia única contendo 585 aminoácidos, incluindo 17 pontes de enxofre (ou ligações dissulfeto) e uma cisteína livre (Cys34)[96, 103]. As ligações dissulfetos proporcionam estabilidade à estrutura dobrada de proteínas plasmáticas, ajudando assim a manter suas funções durante a frequente movimentação das proteínas no sistema circulatório [104]. A ASH contém apenas uma molécula de triptofano, uma das características responsáveis pela fluorescência ultravioleta da proteína, o que favorece o uso de métodos espectroscópicos para o estudo dos seus mecanismos moleculares de ligação [97].

Figura 3.1: Estrutura da proteína HSA composta por três domínios, subdivididos em subdomínios Ia (Amarelo), Ib (verde), IIa (vermelho), IIb (magenta), IIIa (azul) e IIIb (ciano).



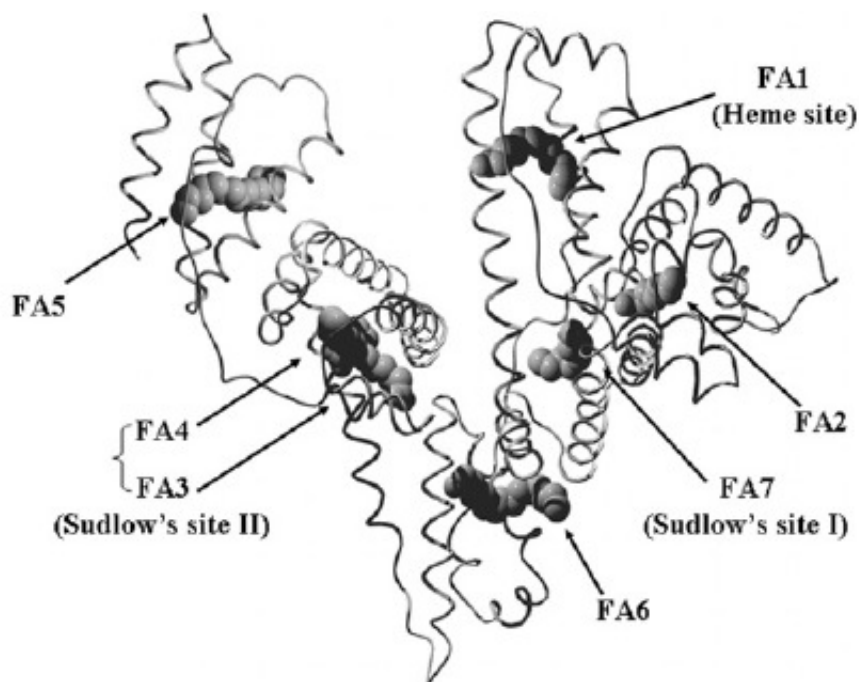
Fonte: Sugio et al., 1999

Embora a ASH seja composta por uma única cadeia polipeptídica, ela se comporta como se fosse composta por várias subunidades e apresenta vários sítios de ligação, dos quais os que possuem maior afinidade com moléculas orgânicas se localizam em aberturas hidrofóbicas [97, 105]. Contudo, estruturas cristalinas com alta resolução revelam que existem dois principais sítios em

que a maioria das moléculas se ligam, tais quais possuem formas e potenciais eletrostáticos diferentes. Estes sítios ficam localizados nos subdomínios IIa e IIIa (ver Figura 3.2) e são conhecidos, respectivamente, como sítios de Sudlow I e II.

O sítio de Sudlow I (localizado numa área sobreposta à região também conhecida como FA7) se caracteriza por uma cavidade formada pelas cadeias laterais de aminoácidos hidrofóbicos. Por outro lado, a entrada desta cavidade é cercada por resíduos carregados positivamente, como Argininas, Lisinas e Histidinas. Já no sítio de Sudlow II (formado por sítios de ligação também denominados FA3 e FA4), além das cadeias laterais hidrofóbicas que revestem o interior da cavidade, é marcada a presença de uma ponte dissulfeto [96, 99].

Figura 3.2: Sítios de Sudlow I e II, os quais encontram-se nos subdomínios IIa e IIIa (também denominados FA7 e FA3/FA4, respectivamente). Na figura são mostrados os principais sítios da ASH revelados pelo estudo cristalográfico da ASH complexada com ácidos graxos de cadeia longa e média.



Fonte: Fanali et al., 2012

Estudos teóricos de relação quantitativa estrutura-atividade, o QSAR (sigla para Quantitative Structure–Activity Relationship), assim como simulações de docking molecular e observações da estrutura cristalina, evidenciam a contribuição de forças hidrofóbicas, eletrostáticas e de ligações de hidrogênio na formação do complexo macromolecular HSA-ligante [1]. Desta forma, os cálculos de energia intermolecular entre o ligante e os resíduos da ASH são fundamentais para um entendimento mais preciso das interações que envolvem a ligação entre esta proteína e a molécula de interesse. Em contrapartida, há a necessidade de aferir quais métodos descrevem melhor a energia e a contribuição dos diferentes tipos de forças envolvidas na interação entre o ligante e os resíduos do sítio ativo do sistema macromolecular.

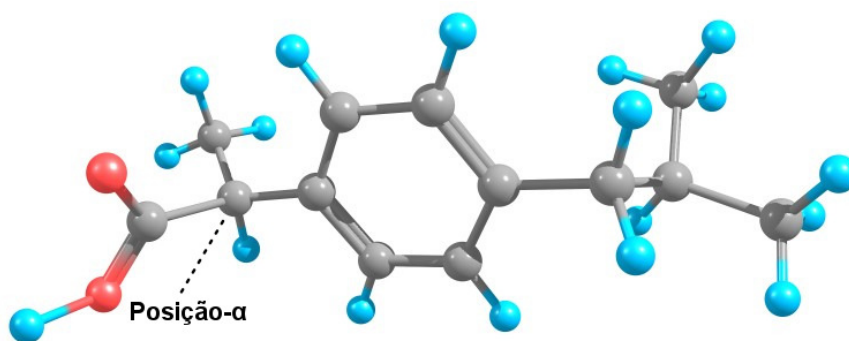
3.2 O ligante Ibuprofeno (IBU)

O Ibuprofeno (IBU) é um fármaco anti-inflamatório não-esteróide (AINE). Os AINE's estão entre os medicamentos mais prescritos em todo o mundo, visto que são utilizados em muitos cuidados paliativos, assim como no tratamento de dor aguda, febre e inflamações [85]. O IBU encontra-se entre os medicamentos mais importantes e essenciais em um sistema básico de saúde, segundo lista publicada pela OMS (Organização Mundial de Saúde) [106]. Este fármaco foi produzido como uma alternativa segura à aspirina. Inicialmente, a partir da década de 70, seu uso medicinal era aplicado ao tratamento de artrite reumatóide e somente anos mais tarde, mais precisamente em 1983, passou a ser o primeiro AINE (além da aspirina) licenciado e acessível para a venda sem receita médica nos países do Reino Unido [107].

A molécula do IBU possui fórmula molecular $C_{13}H_{18}O_2$ e dispõe das seguintes nomenclaturas, segundo normas da IUPAC: ácido (RS)-2-(4-(2-metilpropil)fenil)propanóico, ácido 2-(4-isobutilfenil)propanóico e ácido 2-(4-isobutilfenil)propiônico. Sua estrutura molecular é derivada do ácido propiônico, apresentando assim um grupo carboxílico na porção propionato da molécula, como mostra a figura 3.3. Esta molécula também é constituída por uma região não polar, composta pelo anel aromático do benzeno ligado ao grupo metilpropil [108]. Por possuir um elevado peso

molecular e uma baixa polaridade, o IBU possui solubilidade praticamente nula (muito baixa) em água, mas apresenta alta solubilidade em alguns solventes orgânicos como metanol, etanol, clorofórmio e acetona [109]. Devido à quiralidade do carbono na posição- α da cadeia lateral proveniente do ácido propiônico, o IBU manifesta isomerismo óptico, apresentando as formas S-Ibuprofeno e R-Ibuprofeno. Normalmente existe uma mistura racêmica na produção do IBU, porém o isômero dextrógiro (S-Ibuprofeno) é a forma que desempenha maior atividade biológica. Neste estudo usamos os dados cristalográficos do isômero S-Ibuprofeno co-cristalizado com a proteína ASH.

Figura 3.3: Estrutura molecular do ligante Ibuprofeno (IBU). A estrutura molecular é composta por um grupo químico carboxílico na região do propionato, onde ocorre a presença do carbono quiral do composto (na posição- α). A outra porção da molécula é constituída pelo anel aromático do benzeno juntamente com o grupo metil propil.



Fonte: Autora, 2017

Os AINE's inibem as duas isoformas da enzima ciclo-oxigenase (COX), conhecidas como COX-1 e COX-2. Estas enzimas convertem o ácido aracdônico em prostaglandinas H₂(PGH₂), que por sua vez, são as proteínas responsáveis por regular atividades inflamatórias no organismo [110]. Portanto, as ações farmacológicas antiinflamatória, analgésica e antipirética dos AINE's ocorrem devido a sua capacidade de inibir a COX-1 e COX-2, reduzindo assim a síntese das prostaglandinas mediadoras da dor febre e inchaço. O IBU é designado como um AINE tradicional (ou convencional) e apesar de não ser seletivo (inibe tanto a COX-1 quanto a COX-2) esta molécula

exibe um alto poder de inibição da COX-1 [111].

A função individual de cada isoforma da COX no efeitos analgésicos e antiinflamatórios, assim como nos efeitos colaterais provocados por cada tipo de AINE ainda são incertos e amplamente discutidos, de modo que diferentes compostos moleculares apresentam diferentes níveis de analgesia e reações adversas [112]. Embora o IBU apresente atividade anti-inflamatória menor que outros AINE's, ele oferece menos efeitos colaterais, como o risco de infarto do miocárdio e diminuição de hemorragias gastrointestinais, com seu uso contínuo [113].

Estudos epidemiológicos prévios afirmam que o uso de AINE's, especialmente do IBU, está relacionado ao baixo risco de desenvolvimento de doenças neurodegenerativas como Alzheimer e Parkinson, uma vez que o IBU apresenta potencialmente o efeito neuroprotetor que outros analgésicos não possuem [114, 115]. Por outro lado, o uso muito prolongado do IBU poderia desencadear diversos danos, como por exemplo, problemas nos sistemas vasculares, urinário e digestivo [111, 116].

3.3 Detalhes da metodologia

A metodologia aplicada para calcular a energia de interação entre o IBU e os resíduos da ASH se baseia no uso de métodos quânticos de primeiros princípios, como o Hartree-Fock (HF), além de métodos ab initio como o método da Teoria do Funcional de Densidade, DFT (Density Functional Theory), e do método perturbativo de segunda ordem MP2 (Møller–Plesset perturbation) ¹.

Cálculos quânticos de sistemas com muitos átomos demandam alto custo computacional, impondo assim uma restrição no tempo gasto para obter a energia de interação em sistemas macromoleculares. Por outro lado, métodos de fragmentação molecular podem auxiliar na obtenção das energias de interação fármaco-proteína, uma vez que o sistema total de milhares de átomos se reduz a vários sistemas menores da ordem de algumas dezenas de átomos, minimizando desta forma o

¹Todos os métodos foram calculados através do software G09 [117]

custo computacional e, conseqüentemente, tornando viável o uso de métodos quânticos nestes sistemas. Deste modo, para o sistema proposto, utilizamos o método de fragmentação molecular com caps conjugados, o MFCC (sigla para o termo em inglês Molecular Fractionation with Conjugate Caps). Através do uso da estratégia MFCC no sistema IBU-ASH, a energia de interação entre o ligante IBU e cada aminoácido, $E(\text{IBU} - R_i)$, foi calculada considerando-se a seguinte relação:

$$\begin{aligned} E(\text{IBU} - R_i) = & E(\text{IBU} - C_{i-1}R_iC_{i+1}) - E(\text{IBU} - C_{i-1}C_{i+1}) \\ & - E(C_{i-1}R_iC_{i+1}) + E(C_{i-1}C_{i+1}), \end{aligned} \quad (3.1)$$

Onde C_{i-1} e C_{i+1} representam os caps conjugados, isto é, os aminoácidos precedentes e subsequentes ao resíduo de interesse R_i na cadeia peptídica. O termo $E(\text{IBU} - C_{i-1}R_iC_{i+1})$ corresponde à energia total do sistema composto pela molécula do IBU e o resíduo acoplado aos caps, ao passo que $E(\text{IBU} - C_{i-1}C_{i+1})$ é a energia total do sistema formado pelo ligante e os caps. A energia total dos caps em conjunto com o resíduo R_i é dada pelo termo $E(C_{i-1}R_iC_{i+1})$, e finalmente, o termo $E(C_{i-1}C_{i+1})$ reproduz a energia total do sistema formado somente pelos caps conjugados. Vale mencionar que os átomos de nitrogênio (N) e carbono (C) ligados aos terminais de cada cap conjugado, onde a cadeia peptídica é segmentada no processo de fragmentação molecular, são substituídos por átomos de hidrogênio.

A fim de verificar a contribuição eletrostática na energia de interação ligante-resíduo, foi adicionada a contribuição da polarização induzida pelo campo gerado pelas cargas atômicas da proteína. Para isso, os demais átomos da proteína, aqueles que não entram explicitamente nos cálculos de energia são considerados como cargas pontuais. Onde as cargas pontuais atômicas são calculadas, via métodos auto consistentes (via HF e DFT), a fim de fornecer o potencial eletrostático molecular apropriado [118]. Neste caso, as cargas pontuais representam a própria carga atômica efetiva, reproduzindo assim o potencial eletrostático molecular [119, 120].

Ao acrescentar a contribuição eletrostática nos termos da energia, a equação 3.1 se torna:

$$\begin{aligned}
 E(\text{IBU} - R_i) = & E(\text{IBU} - C_{i-1}R_iC_{i+1} - C_j^{cp}) - E(\text{IBU} - C_{i-1}C_{i+1} - C_j^{cp}) \\
 & - E(C_{i-1}R_iC_{i+1} - C_j^{cp}) + E(C_{i-1}C_{i+1} - C_j^{cp}),
 \end{aligned}
 \tag{3.2}$$

Onde o índice cp representa as cargas pontuais, enquanto que j caracteriza o conjunto de resíduos em torno do resíduo referência R_i , sem considerar os Caps, e onde cada átomo é substituído pela sua carga.

Com o objetivo de refinar a descrição da interação entre o IBU e os resíduos da ASH, foi incluída a correção de longo alcance no método DFT, por meio do funcional híbrido de correlação e troca chamado CAM-B3LYP. Por conseguinte, cada termo contribuinte na expressão da energia de interação foi calculado por meio dos métodos quânticos HF, DFT, DFT com funcional CAM-B3LYP (onde CAM se refere à correção de longo alcance) e MP2, mediante o uso dos conjuntos de bases gaussianas 6-31+G(d) e 6-311+G(d), onde o sinal + e o (d) indicam a presença de uma função difusa e uma polarizada, respectivamente.

Tabela 3.1: Nesta tabela, temos a relação de quais métodos quânticos foram empregados no cálculo da energia de interação resíduo-ligante, assim como quais conjuntos de base e em quais casos foi incluído o background de cargas pontuais.

Métodos Quânticos	Conjunto de bases	Cargas pontuais
Hartree-Fock (HF)	6-31+G(d), 6-311+G(d)	Ausência/Presença
DFT	6-31+G(d), 6-311+G(d)	Ausência/Presença
DFT com funcional CAM-B3LYP	6-31+G(d), 6-311+G(d)	Ausência/Presença
MP2	6-311+G(d)	Ausência

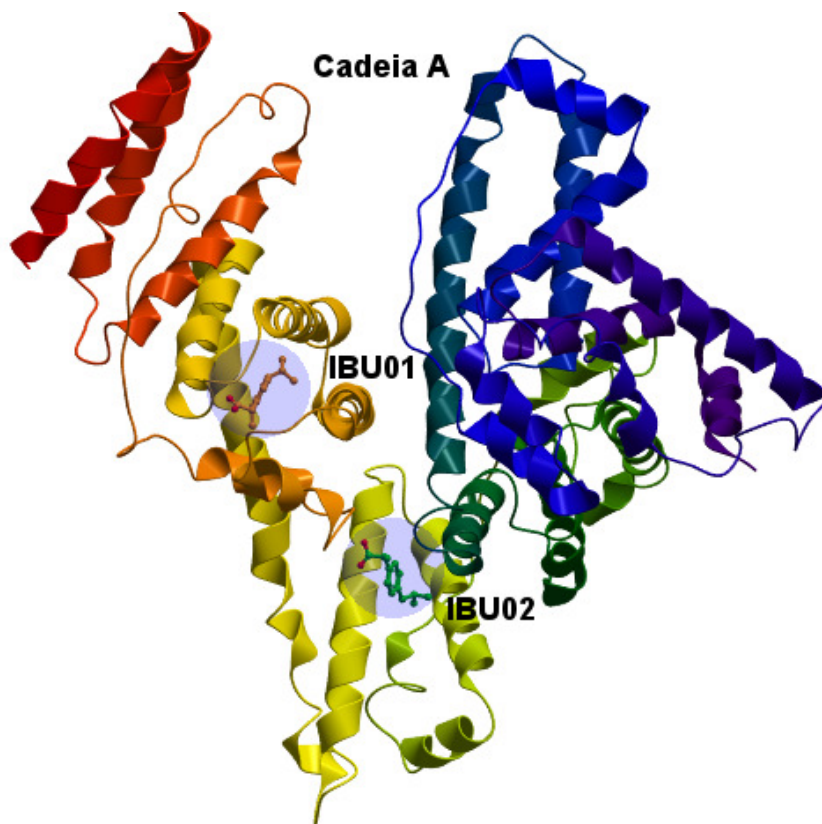
Fonte: Autora, 2017

A tabela 3.1 ordena os métodos quânticos usados para os cálculos de energia, com as devidas especificações do conjunto de bases e a consideração do efeito eletrostático por inclusão de cargas pontuais. Ressalta-se que, devido ao elevado custo computacional, o método perturbativo MP2

foi aplicado somente com o conjunto de bases 6-31+G(d) e na ausência do background de cargas pontuais. Não obstante, foram utilizadas as bases 6-31+G(d) e 6-311+G(d) nos seguintes métodos quânticos: HF, DFT e DFT com funcional CAM-B3LYP. Nestes mesmos métodos os cálculos de energia foram realizados tanto na ausência, quanto na presença das cargas pontuais.

Os dados cristalográficos do IBU complexado com a ASH foram retirados do arquivo PDB (Protein Data Bank) identificado por 2BXG. A co-cristalização do complexo IBU-ASH foi obtida através de métodos de difração de raios X, apresentando assim uma resolução de 2.7 Å. O cristal 2BXG apresenta estrutura com duas cadeias monoméricas, a cadeia A e a cadeia B. Cada cadeia possui dois sítios ativos ocupados pelas moléculas do ligante IBU. Os ligantes localizados no sítio FA3/FA4 (também denominado sítio 2 de Sundlow) foram nomeados por IBU01, enquanto que os ligantes no sítio FA6 são chamados IBU02. A figura 3.4 mostra a cadeia A e os sítios ocupados pelas moléculas correspondentes IBU01 e IBU02.

Figura 3.4: Estrutura cristalina do IBU complexado com a proteína ASH. Os dados cristalográficos, encontrados no arquivo PDB identificado por 2BXG, foram obtidos através de difração de raios x e possuem uma resolução de 2,7 Å.

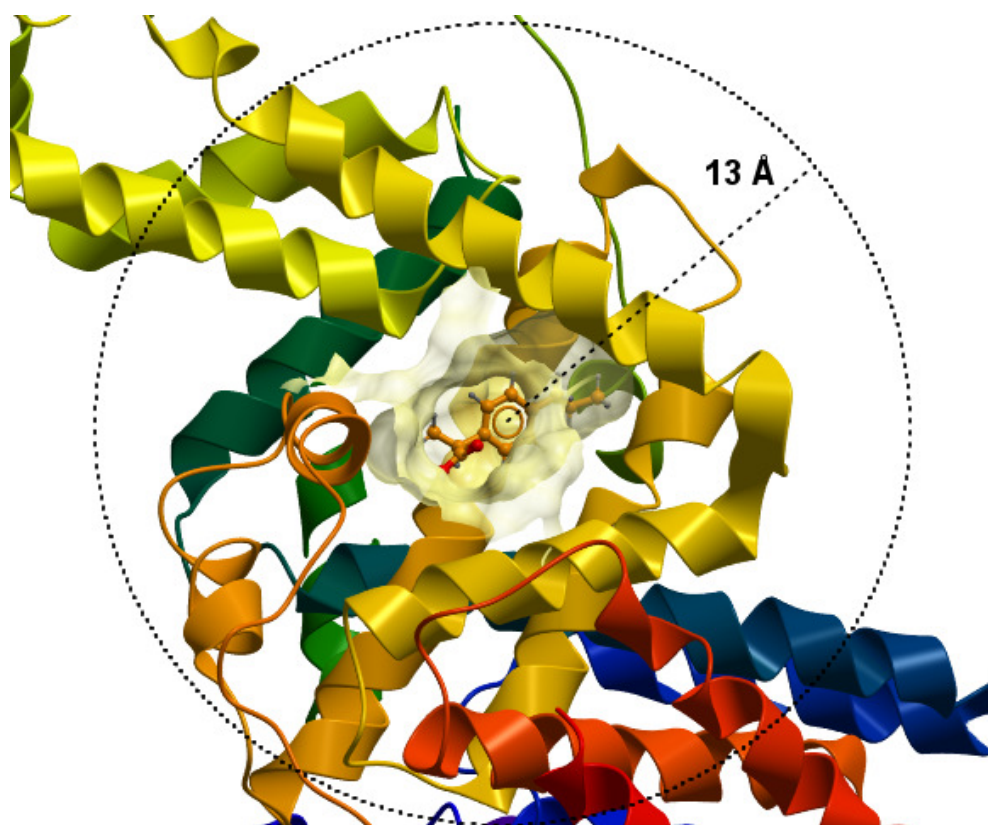


Fonte: Autora, 2017

Os átomos de hidrogênio espalham os raios x muito fracamente. Desta forma, as posições dos hidrogênios não podem ser definidas por meio da técnica de difração de raios x, motivo pelo qual a estrutura cristalográfica não fornece as posições atômicas destes átomos, sendo estes inseridos e em seguida otimizados via campo de força CHARMM, enquanto que os átomos pesados como Carbono, Nitrogênio, Oxigênio e Enxofre são mantidos fixos. Feita a otimização dos átomos de hidrogênio, partimos para a fragmentação via MFCC, ou seja, a separação dos aminoácidos (juntamente com seus caps conjugados) cujas energias devem ser calculadas. Para os cálculos de energia, foram escolhidos os aminoácidos do sítio ativo do ligante IBU01 da cadeia A, onde também fo-

ram considerados os aminoácidos na circunvizinhança do sítio. Ao todo, foram selecionados 73 aminoácidos, distribuídos num raio de 13 angstroms (\AA) em relação ao centróide da molécula do ligante. A figura 3.12 destaca a região onde se encontram os aminoácidos selecionados para os cálculos de energia, enquanto que a tabela 3.2 aponta tais aminoácidos e suas respectivas posições na distribuição radial feita em torno do centro de massa do IBU01.

Figura 3.5: Nesta figura é possível visualizar a região do sítio ativo FA3/FA4 (área sombreada) e o raio de 13 angstroms (\AA), traçado em torno do centro do fármaco, que delimita a região que contém os resíduos inclusos no cálculo das energias de interação.



Fonte: Autora, 2017

Tabela 3.2: A tabela lista os aminoácidos do sítio ativo (marcados com *) e circunvizinhança para os quais foram calculadas as energias de interação com o ligante IBU01. No total, foram escolhidos 73 resíduos, distribuídos num raio de 13 Å em torno do centróide do ligante.

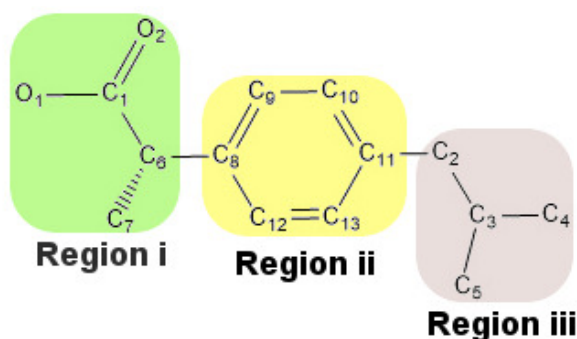
Raio (Å)	
3	LEU453*
4,5	LEU387* ASN391*
5	ILE388* PHE403* LEU407* ARG410* LEU430*
5,5	VAL433* ARG485*
6	ALA449*
6,5	TYR411* SER489*
7	GLY431 GLU450
7,5	CYS392* GLY434* LEU457
8	PHE395 TYR452
8,5	PRO384 GLN390 LEU394 LYS414* CYS437* CYS438* VAL456
9	LYS389 ASN429 LYS432 PHE488*
9,5	ASN386 GLN404 VAL426 SER427 MET446 SER454
10	GLU383 GLU393 ALA406 SER435 ARG445 CYS448 ASP451 PRO486
10,5	GLN385 LYS436 VAL455
11	ARG348 LEU408 ARG428 ARG484 ALA490
11,5	VAL344 GLU396 LEU398 PRO447
12	GLU400 ASN458 LEU460 CYS487 LEU491
12,5	SER342 GLY399 LYS402 ASN405 LYS439 ASN483
13	GLU382 GLN397 VAL409 GLU425 HIS440

Fonte: Autora, 2017

3.4 Resultados

A estrutura do ligante IBU01, conforme figura 3.6, foi dividida em três regiões de interesse com as quais os aminoácidos interagem, são elas a região i, ii e iii. Essa divisão foi feita a fim de proporcionar uma análise mais detalhada e constatar a localidade do ligante em que os resíduos apresentam uma maior afinidade. A região i compreende o grupo carboxílico e com maior eletro-negatividade em relação ao restante do fármaco, onde esse grupo químico favorece a formação de interações mais fortes, tais como as que são mediadas por interações de hidrogênio. Já a região ii, representa o grupo Arila e contempla o anel aromático, enquanto que a região iii compreende aos carbonos que compõem o grupo metilpropil (grupo Isobutilo) anexado ao benzeno da molécula.

Figura 3.6: Estrutura molecular do ligante Ibuprofeno (IBU). A fim de facilitar a análise das interações da molécula com os re'íduos da proteína, o IBU é dividido nas regiões ii, ii e iii. A região i compreende o grupo carboxílico, enquanto que a região ii contempla o anel aromático e a região iii os 4 carbonos do grupo metilpropil ligado ao benzeno da molécula.



Fonte: Autora, 2017

Conforme os cálculos realizados por meio dos métodos quânticos DFT, DFT com funcional CAM-B3LYP e MP2, mediante o uso do background de cargas pontuais, foram detectados pequenas flutuações na energia de interação entre o IBU01 e os dados resíduos. Estas diferenças são identificadas em relação ao uso dos diferentes conjuntos de bases, 6-31g e 6-311g, bem como em

relação ao uso das cargas pontuais. Por outro lado, o nível de cálculo assumido pelos diferentes métodos, também proporciona essas flutuações na energia de interação. Nas tabelas de 3.3 a 3.14, são mostradas todas as energias obtidas segundo cada método empregado. Nota-se que para a maioria dos resíduos, de acordo com a devida metodologia aplicada, as energias convergem para um dado valor, não sofrendo grandes variações de um método para outro. No entanto, alguns resíduos carregados, sofrem consideráveis variações conforme alguns refinamentos são introduzidos nos cálculos. No caso de alguns aminoácidos polarizados e com carga oposta a do ligante, como por exemplo a LYS414, a LYS432, a LYS402 e a LYS389, os quais são positivamente carregados, estes apresentaram energias com diferenças superiores a 10 kcal/mol quando se compara os métodos que incluem a correção de longo alcance via funcional CAM-B3LYP. Na tabelas a seguir são mostradas as energias calculadas para os 73 aminoácidos em acordo com cada um dos métodos e correções de longo alcance empregados. Os resultados são comparados com as energias dos resultados mostrados na referência [85]), onde as energias foram calculadas usando o método DFT em conjunto com funcional de Perdew, Burke e Ernzerhof (PBE) adotando conjunto de bases duplamente polarizada (DNP), juntamente com a correção de longo alcance de Grimme (GGA + D). Para aumentar a precisão dos cálculos da energia de interação intermolecular entre o IBU e os resíduos da ASH, a metodologia utilizada em [85] se baseou no uso do modelo de solvatação contínua COSMO (Conductor-like Screening Model), a fim de garantir uma eficácia maior nos resultados e incluir a contribuição de forças eletrostáticas. Neste caso, para a constante dielétrica foram adotados os casos em que $\epsilon = 0$ e $\epsilon = 40$.

Tabela 3.3: A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)				
				Conjunto de bases 6-31+G(d)			Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	$\epsilon = 0$	$\epsilon = 40$
3,0	LEU453	1,801	HD11 - iii(H9)C5	2,577	1,300	0,116	-6,700	-4,540
4,5	LEU387	1,891	HD21 - ii(H17)C12	5,203	1,444	0,744	-0,800	-0,740
4,5	ASN391	2,170	HB2 - iii(H1)C2	7,089	5,378	4,545	-1,580	-0,960
5,0	ARG410	1,983	HH21 - i(O2)COO(-)	-85,543	-85,110	-86,584	-84,850	-16,500
5,0	PHE403	2,612	HZ - iii(H1)C2	1,353	0,868	0,712	-3,530	-2,810
5,0	ILE388	2,668	HD13 - iii(H8)C5	1,017	1,190	0,886	-0,220	-1,340
5,0	LEU430	2,828	HD23 - ii(H14)C9	-1,610	-1,441	-1,835	-4,620	-3,160
5,0	LEU407	3,592	HD11 - ii(H15)C10	0,670	0,508	0,379	-0,580	-2,720
5,5	VAL433	1,900	HG11 - iii(H7)C5	3,806	3,292	1,364	-2,930	-2,700
5,5	ARG485	2,419	HD2 - i(H11)C7	-37,506	-39,019	-38,908	-37,440	-1,570
6,0	ALA449	2,087	HB1 - iii(H8)C5	1,703	1,074	0,594	0,240	-0,040
6,5	SER489	1,411	HG - i(H12)C7	10,290	7,071	6,001	-15,810	-7,720
6,5	TYR411	1,878	HH - i(O1)COO(-)	-8,947	-9,944	-11,919	-5,450	-1,380
7,0	GLY431	4,790	HA1 - ii(H15)C10	-0,081	-1,027	-0,216	-0,520	0,140
7,0	GLU450	5,041	HA - iii(H9)C5	29,937	30,190	30,126	30,510	1,020
7,5	GLY434	2,176	HA1 - iii(H4)C4	0,433	0,272	-0,308	-0,420	-0,170
7,5	CYS392	3,641	SG - iii(H6)C4	-0,081	0,338	0,091	0,790	-0,030
7,5	LEU457	4,156	HD11 - i(H11)C7	-1,216	-1,158	-1,195	-0,480	-1,250
8,0	PHE395	4,061	HB1 - iii(H5)C4	-0,674	-0,518	-0,578	-0,960	-1,410
8,0	TYR452	4,445	HB2 - iii(H7)C5	-1,539	-1,673	-1,626	-2,150	-0,300
8,5	LYS414	2,475	HZ3 - i(O1)COO(-)	-93,335	-82,541	-93,485	-81,410	-9,160
8,5	CYS438	3,412	SG - iii(H6)C4	0,655	0,676	0,379	0,370	-0,030
8,5	CYS437	3,704	HB1 - iii(H4)C4	-1,178	-1,148	-1,198	-0,780	-0,050

¹ Energias calculadas em [85].

Tabela 3.4: A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)				
				Conjunto de bases 6-31+G(d)			Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	$\epsilon = 0$	$\epsilon = 40$
8,5	GLN390	6,154	HB2 - ii(H16)C12	-2,600	-1,243	-2,241	-0,250	-0,070
8,5	LEU394	6,224	HB2 - iii(H1)C2	-1,604	-1,535	-1,559	-0,490	-0,090
8,5	VAL456	6,374	HG11 - ii(H14)C9	-1,722	-1,787	-1,784	-2,200	-0,200
8,5	PRO384	6,467	O - ii(H16)C12	1,223	1,170	1,191	0,940	0,050
9,0	PHE488	3,650	HB2 - i(H13)C7	0,303	-0,349	-0,220	0,350	-0,450
9,0	LYS432	5,402	C - iii(H4)C4	-20,970	-33,335	-21,830	-39,630	-0,990
9,0	LYS389	6,452	N - iii(H2)C2	-19,896	-32,722	-20,878	-23,940	-0,640
9,0	ASN429	6,457	O -iii(H3)C3 / O	-1,859	-1,865	-1,860	1,290	-0,100
9,5	MET446	6,077	HA - iii(H8)C5	1,143	0,951	1,160	-0,310	-0,010
9,5	SER427	6,532	HA - ii(H14)C9	0,166	-0,578	0,193	1,250	0,040
9,5	VAL426	6,865	HG13 - ii(H14)C9	0,887	0,750	0,781	0,660	-0,030
9,5	ASN386	6,984	O - ii(H16)C12	-1,044	-1,153	-1,124	0,120	-0,040
9,5	SER454	7,252	HN - iii(H9)C5	-0,246	0,020	-0,038	-0,260	-0,020
9,5	GLN404	7,667	HA - ii(H15)C10	0,973	1,005	1,037	-2,120	0,040
10,0	SER435	4,889	HN - iii(H4)C4	-0,536	-0,422	-0,454	-1,370	-0,020
10,0	PRO486	5,606	HA - i(H11)C7	2,848	2,496	2,944	0,080	0,040
10,0	ARG445	5,714	O - iii(H8)C5	-20,458	-20,918	-20,766	-23,970	-0,670
10,0	CYS448	5,814	C - iii(H7)C5	-0,830	-0,988	-0,941	-0,650	-0,010
10,0	GLU383	6,565	OE2 - i(H12)C7	31,517	31,185	31,298	32,790	1,070
10,0	GLU393	7,067	HN - iii(H6)C4	20,269	20,635	20,492	28,690	0,660
10,0	ASP451	7,196	HN - iii(H7)C5	20,602	21,022	20,881	25,030	0,690
10,0	ALA406	7,870	O - i(O2)COO(-)	-0,547	-0,876	-0,800	0,190	0,010
10,5	LYS436	5,316	HN - iii(H4)C4	-1,344	-1,332	-1,340	-0,850	-0,040
10,5	GLN385	8,059	O - ii(H16)C12	0,842	0,876	0,882	0,240	0,050
10,5	VAL455	8,675	HN - iii(H9)C5	-0,737	-0,744	-0,751	-1,230	-0,100
11,0	ALA490	5,610	HN - i(H13)C7	-0,819	-0,354	-0,469	0,350	0,180

¹ Energias calculadas em [85].

Tabela 3.5: A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)				
				Conjunto de bases 6-31+G(d)			Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	$\epsilon = 0$	$\epsilon = 40$
11,0	ARG484	6,335	O - i(H13)C7	-21,492	-22,219	-21,889	-27,590	-0,810
11,0	ARG348	8,313	H22 - i(H11)C7	-24,253	-24,313	-24,290	-36,300	-0,880
11,0	ARG428	8,355	O - ii(H15)C10	-19,419	-19,514	-19,472	-35,470	-0,770
11,0	LEU408	8,722	N - i(O2)COO(-)	1,115	1,313	1,294	3,500	0,130
11,5	GLU396	7,408	HN - iii(H5)C4	18,845	19,319	19,159	20,960	0,640
11,5	PRO447	8,223	O - iii(H7)C5	0,704	0,696	0,715	-0,260	-0,010
11,5	LEU398	9,160	HD13 - iii(H1)C2	-1,091	-1,075	-1,086	-1,860	-1,000
11,5	VAL344	9,485	HG21 - i(H11)C7	-0,670	-0,692	-0,693	0,210	-0,040
12,0	LEU491	6,370	HN - i(H13)C7	3,352	2,891	2,997	1,720	-0,030
12,0	CYS487	6,374	C - i(H13)C7	-1,204	-1,381	-1,337	1,050	0,180
12,0	GLU400	7,945	HA - iii(H5)C4	21,785	21,853	21,820	30,000	0,840
12,0	LEU460	8,307	HD23 - i(H10)C6	-1,263	-1,200	-1,228	-0,970	-0,110
12,0	ASN458	10,220	HN - i(H10)C6	-0,466	-0,204	-0,261	-1,280	-0,070
12,5	LYS439	7,159	HN - iii(H4)C4	-0,568	-0,406	-0,455	-	-
12,5	ASN483	8,328	O - i(H11)C7 / HN	-0,015	-0,121	-0,094	2,290	0,320
12,5	GLY399	8,977	O - iii(H5)C4	0,771	0,648	0,676	-0,400	0,000
12,5	SER342	9,563	HB2 - iii(H8)C5	-0,147	-0,191	-0,175	0,110	-0,070
12,5	ASN405	10,545	O - i(O2)COO(-)	0,774	0,689	0,724	-	-
12,5	LYS402	10,728	C - iii(H5)C4	-17,536	-32,930	-19,441	-	-
13,0	HIS440	8,240	HN - iii(H6)C4	0,595	0,552	0,561	-	-
13,0	VAL409	9,279	C - i(O1)COO(-)	-1,283	-1,339	-1,348	-	-
13,0	GLN397	9,362	HN - iii(H5)C4	-1,125	-1,085	-1,102	0,260	-0,010
13,0	GLU425	10,476	O - ii(H14)C9	19,819	19,930	19,887	-	-
13,0	GLU382	10,689	O - ii(H16)C12	16,845	16,887	16,860	-	-

¹ Energias calculadas em [85].

Tabela 3.6: A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)				
				Conjunto de bases 6-31+G(d) com cargas pontuais			Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	$\epsilon = 0$	$\epsilon = 40$
3,0	LEU453	1,801	HD11 - iii(H9)C5	3,964	2,773	1,578	-6,700	-0,070
4,5	LEU387	1,891	HD21 - ii(H17)C12	5,203	4,247	3,484	-0,800	-0,090
4,5	ASN391	2,170	HB2 - iii(H1)C2	12,267	10,748	9,846	-1,580	-0,200
5,0	ARG410	1,983	HH21 - i(O2)COO(-)	-81,171	-79,972	-81,617	-84,850	0,050
5,0	PHE403	2,612	HZ - iii(H1)C2	3,695	3,255	3,087	-3,530	-0,450
5,0	ILE388	2,668	HD13 - iii(H8)C5	1,445	1,617	1,342	-0,220	-0,990
5,0	LEU430	2,828	HD23 - ii(H14)C9	-1,036	-0,839	-1,236	-4,620	-0,640
5,0	LEU407	3,592	HD11 - ii(H15)C10	5,995	6,023	5,844	-0,580	-0,100
5,5	VAL433	1,900	HG11 - iii(H7)C5	4,313	2,958	1,655	-2,930	-0,010
5,5	ARG485	2,419	HD2 - i(H11)C7	-31,578	-32,173	-32,433	-37,440	0,040
6,0	ALA449	2,087	HB1 - iii(H8)C5	2,382	1,779	1,279	0,240	-0,030
6,5	SER489	1,411	HG - i(H12)C7	13,368	10,396	9,268	-15,810	-0,040
6,5	TYR411	1,878	HH - i(O1)COO(-)	-5,612	-7,302	-9,197	-5,450	-0,020
7,0	GLY431	4,790	HA1 - ii(H15)C10	-0,200	-0,043	-0,073	-0,520	0,040
7,0	GLU450	5,041	HA - iii(H9)C5	30,373	30,388	30,431	30,510	-0,020
7,5	GLY434	2,176	HA1 - iii(H4)C4	0,383	0,194	-0,378	-0,420	0,040
7,5	CYS392	3,641	SG - iii(H6)C4	-0,058	0,315	0,096	0,790	-0,670
7,5	LEU457	4,156	HD11 - i(H11)C7	-1,142	-1,166	-1,176	-0,480	-0,010
8,0	PHE395	4,061	HB1 - iii(H5)C4	-0,920	-0,785	-0,837	-0,960	1,070
8,0	TYR452	4,445	HB2 - iii(H7)C5	-0,911	-1,189	-0,878	-2,150	0,660
8,5	LYS414	2,475	HZ3 - i(O1)COO(-)	-89,549	-88,060	-89,269	-81,410	0,690
8,5	CYS438	3,412	SG - iii(H6)C4	0,372	0,352	0,074	0,370	0,010
8,5	CYS437	3,704	HB1 - iii(H4)C4	-0,714	-0,820	-0,649	-0,780	-0,040

¹ Energias calculadas em [85].

Tabela 3.7: A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)				
				Conjunto de bases 6-31+G(d) com cargas pontuais			Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	$\epsilon = 0$	$\epsilon = 40$
8,5	GLN390	6,154	HB2 - ii(H16)C12	-1,699	-1,610	-1,667	-0,250	-0,070
8,5	LEU394	6,224	HB2 - iii(H1)C2	-1,473	-1,370	-1,353	-0,490	-0,090
8,5	VAL456	6,374	HG11 - ii(H14)C9	-1,401	-1,339	-1,381	-2,200	-0,200
8,5	PRO384	6,467	O - ii(H16)C12	1,859	1,744	1,815	0,940	0,050
9,0	PHE488	3,650	HB2 - i(H13)C7	5,761	5,924	5,835	0,350	-0,450
9,0	LYS432	5,402	C - iii(H4)C4	-19,579	-19,447	-19,485	-39,630	-0,990
9,0	LYS389	6,452	N - iii(H2)C2	-18,905	-18,974	-18,947	-23,940	-0,640
9,0	ASN429	6,457	O - iii(H3)C3 / O	-1,409	-1,119	-1,222	1,290	-0,100
9,5	MET446	6,077	HA - iii(H8)C5	1,263	1,193	1,233	-0,310	-0,010
9,5	SER427	6,532	HA - ii(H14)C9	0,617	0,574	0,622	1,250	0,040
9,5	VAL426	6,865	HG13 - ii(H14)C9	1,284	0,985	1,206	0,660	-0,030
9,5	ASN386	6,984	O - ii(H16)C12	-0,717	-0,782	-0,717	0,120	-0,040
9,5	SER454	7,252	HN - iii(H9)C5	0,019	0,268	0,220	-0,260	-0,020
9,5	GLN404	7,667	HA - ii(H15)C10	1,490	1,454	1,526	-2,120	0,040
10,0	SER435	4,889	HN - iii(H4)C4	-0,414	-0,346	-0,364	-1,370	-0,020
10,0	PRO486	5,606	HA - i(H11)C7	3,257	3,115	3,258	0,080	0,040
10,0	ARG445	5,714	O - iii(H8)C5	-18,232	-18,318	-18,295	-23,970	-0,670
10,0	CYS448	5,814	C - iii(H7)C5	-0,383	-0,514	-0,461	-0,650	-0,010
10,0	GLU383	6,565	OE2 - i(H12)C7	34,032	33,308	33,872	32,790	1,070
10,0	GLU393	7,067	HN - iii(H6)C4	19,881	20,519	20,039	28,690	0,660
10,0	ASP451	7,196	HN - iii(H7)C5	20,077	20,322	20,232	25,030	0,690
10,0	ALA406	7,870	O - i(O2)COO(-)	0,265	0,006	0,072	0,190	0,010
10,5	LYS436	5,316	HN - iii(H4)C4	-0,941	-0,861	-0,886	-0,850	-0,040
10,5	GLN385	8,059	O - ii(H16)C12	0,979	0,954	0,993	0,240	0,050
10,5	VAL455	8,675	HN - iii(H9)C5	-0,707	-0,697	-0,717	-1,230	-0,100
11,0	ALA490	5,610	HN - i(H13)C7	-0,528	0,141	-0,041	0,350	0,180

¹ Energias calculadas em [85].

Tabela 3.8: A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-31+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)				
				Conjunto de bases 6-31+G(d) com cargas pontuais			Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	$\epsilon = 0$	$\epsilon = 40$
11,0	ARG484	6,335	O - i(H13)C7	-19,560	-19,901	-19,797	-27,590	-0,810
11,0	ARG348	8,313	H22 - i(H11)C7	-23,766	-23,724	-23,732	-36,300	-0,880
11,0	ARG428	8,355	O - ii(H15)C10	-19,339	-19,426	-19,378	-35,470	-0,770
11,0	LEU408	8,722	N - i(O2)COO(-)	1,963	2,263	2,243	3,500	0,130
11,5	GLU396	7,408	HN - iii(H5)C4	17,872	18,260	18,075	20,960	0,640
11,5	PRO447	8,223	O - iii(H7)C5	0,929	0,920	0,943	-0,260	-0,010
11,5	LEU398	9,160	HD13 - iii(H1)C2	-1,036	-1,008	-1,016	-1,860	-1,000
11,5	VAL344	9,485	HG21 - i(H11)C7	-0,689	-0,697	-0,702	0,210	-0,040
12,0	LEU491	6,370	HN - i(H13)C7	4,498	4,115	4,193	1,720	-0,030
12,0	CYS487	6,374	C - i(H13)C7	-0,339	-0,850	-0,494	1,050	0,180
12,0	GLU400	7,945	HA - iii(H5)C4	21,416	22,022	21,376	30,000	0,840
12,0	LEU460	8,307	HD23 - i(H10)C6	-1,313	-1,230	-1,268	-0,970	-0,110
12,0	ASN458	10,220	HN - i(H10)C6	-0,312	-0,035	-0,098	-1,280	-0,070
12,5	LYS439	7,159	HN - iii(H4)C4	-0,447	-0,300	-0,348	-	-
12,5	ASN483	8,328	O - i(H11)C7 / HN	0,700	-0,300	0,578	2,290	0,320
12,5	GLY399	8,977	O - iii(H5)C4	0,985	0,834	0,893	-0,400	0,000
12,5	SER342	9,563	HB2 - iii(H8)C5	-0,137	-0,261	-0,195	0,110	-0,070
12,5	ASN405	10,545	O - i(O2)COO(-)	1,110	0,903	1,002	-	-
12,5	LYS402	10,728	C - iii(H5)C4	-16,862	-17,165	-17,066	-	-
13,0	HIS440	8,240	HN - iii(H6)C4	0,350	0,236	0,291	-	-
13,0	VAL409	9,279	C - i(O1)COO(-)	-1,025	-1,011	-1,031	-	-
13,0	GLN397	9,362	HN - iii(H5)C4	-1,186	-1,094	-1,151	0,260	-0,010
13,0	GLU425	10,476	O - ii(H14)C9	20,158	20,558	20,265	-	-
13,0	GLU382	10,689	O - ii(H16)C12	17,117	17,084	17,086	-	-

¹ Energias calculadas em [85].

Tabela 3.9: A tabela mostra as energias obtidas pelos métodos HF, B3LYP, CAM e MP2 utilizando o conjunto de bases 6-311+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)					
				Conjunto de bases 6-311+G(d)				Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	MP2	$\epsilon = 0$	$\epsilon = 40$
3,0	LEU453	1,801	HD11 - iii(H9)C5	2,238	1,122	-0,097	-5,530	-6,700	-4,540
4,5	LEU387	1,891	HD21 - ii(H17)C12	2,375	1,382	0,659	-2,955	-0,800	-0,740
4,5	ASN391	2,170	HB2 - iii(H1)C2	6,868	5,397	4,513	-1,099	-1,580	-0,960
5,0	ARG410	1,983	HH21 - i(O2)COO(-)	-85,908	-85,246	-86,729	-88,771	-84,850	-16,500
5,0	PHE403	2,612	HZ - iii(H1)C2	1,268	0,910	2,164	-1,696	-3,530	-2,810
5,0	ILE388	2,668	HD13 - iii(H8)C5	0,972	1,168	1,542	-1,618	-0,220	-1,340
5,0	LEU430	2,828	HD23 - ii(H14)C9	-1,715	-1,483	-1,868	-4,758	-4,620	-3,160
5,0	LEU407	3,592	HD11 - ii(H15)C10	0,513	0,433	0,304	-1,961	-0,580	-2,720
5,5	VAL433	1,900	HG11 - iii(H7)C5	3,638	3,250	1,230	-3,286	-2,930	-2,700
5,5	ARG485	2,419	HD2 - i(H11)C7	-37,596	-38,996	-38,930	-41,881	-37,440	-1,570
6,0	ALA449	2,087	HB1 - iii(H8)C5	1,523	0,981	0,464	-1,135	0,240	-0,040
6,5	SER489	1,411	HG - i(H12)C7	10,094	6,785	5,681	4,706	-15,810	-7,720
6,5	TYR411	1,878	HH - i(O1)COO(-)	-9,309	-10,279	-12,307	-14,379	-5,450	-1,380
7,0	GLY431	4,790	HA1 - ii(H15)C10	-0,074	-0,986	-0,170	-0,343	-0,520	0,140
7,0	GLU450	5,041	HA - iii(H9)C5	29,953	30,240	30,172	29,788	30,510	1,020
7,5	GLY434	2,176	HA1 - iii(H4)C4	0,373	0,290	-0,321	-1,878	-0,420	-0,170
7,5	CYS392	3,641	SG - iii(H6)C4	-0,100	0,319	0,077	-0,764	0,790	-0,030
7,5	LEU457	4,156	HD11 - i(H11)C7	-1,280	-1,159	-1,198	-1,829	-0,480	-1,250
8,0	PHE395	4,061	HB1 - iii(H5)C4	-0,688	-0,493	-0,550	-1,048	-0,960	-1,410
8,0	TYR452	4,445	HB2 - iii(H7)C5	-1,556	-1,696	-1,645	-2,038	-2,150	-0,300
8,5	LYS414	2,475	HZ3 - i(O1)COO(-)	-93,456	-82,774	-93,575	-93,629	-81,410	-9,160
8,5	CYS438	3,412	SG - iii(H6)C4	0,531	0,590	0,292	-0,685	0,370	-0,030
8,5	CYS437	3,704	HB1 - iii(H4)C4	-1,246	-1,187	-1,236	-1,972	-0,780	-0,050

¹ Energias calculadas em [85].

Tabela 3.10: A tabela mostra as energias obtidas pelos métodos HF, B3LYP, CAM e MP2 utilizando o conjunto de bases 6-311+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\varepsilon = 0$ e $\varepsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)					
				Conjunto de bases 6-311+G(d)				Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	MP2	$\varepsilon = 0$	$\varepsilon = 40$
8,5	GLN390	6,154	HB2 - ii(H16)C12	-2,603	-1,260	-2,274	-2,986	-0,250	-0,070
8,5	LEU394	6,224	HB2 - iii(H1)C2	-1,605	-1,528	-1,550	-1,700	-0,490	-0,090
8,5	VAL456	6,374	HG11 - ii(H14)C9	-1,744	-1,795	-1,792	-1,956	-2,200	-0,200
8,5	PRO384	6,467	O - ii(H16)C12	1,218	1,161	1,180	1,026	0,940	0,050
9,0	PHE488	3,650	HB2 - i(H13)C7	0,331	-0,353	-0,214	-1,422	0,350	-0,450
9,0	LYS432	5,402	C - iii(H4)C4	-20,944	-33,102	-21,768	-21,255	-39,630	-0,990
9,0	LYS389	6,452	N - iii(H2)C2	-19,879	-32,487	-20,813	-20,131	-23,940	-0,640
9,0	ASN429	6,457	O -iii(H3)C3 / O	-1,904	-1,865	-1,864	-2,027	1,290	-0,100
9,5	MET446	6,077	HA - iii(H8)C5	1,145	0,962	1,155	1,037	-0,310	-0,010
9,5	SER427	6,532	HA - ii(H14)C9	0,159	-0,590	0,179	0,071	1,250	0,040
9,5	VAL426	6,865	HG13 - ii(H14)C9	0,882	0,756	0,785	0,670	0,660	-0,030
9,5	ASN386	6,984	O - ii(H16)C12	-1,154	-1,215	-1,188	-1,319	0,120	-0,040
9,5	SER454	7,252	HN - iii(H9)C5	-0,241	0,009	-0,047	-0,044	-0,260	-0,020
9,5	GLN404	7,667	HA - ii(H15)C10	1,028	1,032	1,063	1,103	-2,120	0,040
10,0	SER435	4,889	HN - iii(H4)C4	-0,501	-0,378	-0,403	-0,550	-1,370	-0,020
10,0	PRO486	5,606	HA - i(H11)C7	2,903	2,559	2,958	2,794	0,080	0,040
10,0	ARG445	5,714	O - iii(H8)C5	-20,517	-20,935	-18,922	-21,024	-23,970	-0,670
10,0	CYS448	5,814	C - iii(H7)C5	-0,827	-1,004	-0,959	-1,163	-0,650	-0,010
10,0	GLU383	6,565	OE2 - i(H12)C7	31,486	31,165	31,274	31,137	32,790	1,070
10,0	GLU393	7,067	HN - iii(H6)C4	20,275	20,656	20,513	20,420	28,690	0,660
10,0	ASP451	7,196	HN - iii(H7)C5	20,606	21,044	20,904	20,886	25,030	0,690
10,0	ALA406	7,870	O - i(O2)COO(-)	-0,680	-0,900	-0,825	-1,121	0,190	0,010
10,5	LYS436	5,316	HN - iii(H4)C4	-1,345	-1,324	-1,334	-1,448	-0,850	-0,040
10,5	GLN385	8,059	O - ii(H16)C12	0,850	0,873	0,878	0,841	0,240	0,050
10,5	VAL455	8,675	HN - iii(H9)C5	-0,740	-0,747	-0,753	-0,780	-1,230	-0,100
11,0	ALA490	5,610	HN - i(H13)C7	-0,314	-0,383	-0,490	-0,028	0,350	0,180

¹ Energias calculadas em [85].

Tabela 3.11: A tabela mostra as energias obtidas pelos métodos HF, B3LYP, CAM e MP2 utilizando o conjunto de bases 6-311+G(d), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)					
				Conjunto de bases 6-311+G(d)				Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	MP2	$\epsilon = 0$	$\epsilon = 40$
11,0	ARG484	6,335	O - i(H13)C7	-21,552	-22,232	-21,917	-22,063	-27,590	-0,810
11,0	ARG348	8,313	H22 - i(H11)C7	-24,290	-24,319	-24,293	-24,405	-36,300	-0,880
11,0	ARG428	8,355	O - ii(H15)C10	-19,238	-19,515	-19,473	-19,292	-35,470	-0,770
11,0	LEU408	8,722	N - i(O2)COO(-)	1,142	1,308	1,290	1,260	3,500	0,130
11,5	GLU396	7,408	HN - iii(H5)C4	18,866	19,336	19,177	19,152	20,960	0,640
11,5	PRO447	8,223	O - iii(H7)C5	0,698	0,676	0,694	0,650	-0,260	-0,010
11,5	LEU398	9,160	HD13 - iii(H1)C2	-1,096	-1,073	-1,083	-1,036	-1,860	-1,000
11,5	VAL344	9,485	HG21 - i(H11)C7	-0,674	-0,686	-0,687	-0,705	0,210	-0,040
12,0	LEU491	6,370	HN - i(H13)C7	3,349	2,876	2,981	2,627	1,720	-0,030
12,0	CYS487	6,374	C - i(H13)C7	-1,233	-1,425	-1,385	-1,466	1,050	0,180
12,0	GLU400	7,945	HA - iii(H5)C4	21,795	21,859	21,824	21,778	30,000	0,840
12,0	LEU460	8,307	HD23 - i(H10)C6	-1,263	-1,186	-1,214	-1,220	-0,970	-0,110
12,0	ASN458	10,220	HN - i(H10)C6	-0,296	-0,205	-0,261	-0,074	-1,280	-0,070
12,5	LYS439	7,159	HN - iii(H4)C4	-0,569	-0,407	-0,456	-0,485	-	-
12,5	ASN483	8,328	O - i(H11)C7 / HN	-0,055	-1,187	-0,092	1,833	2,290	0,320
12,5	GLY399	8,977	O - iii(H5)C4	0,770	0,646	0,675	0,605	-0,400	0,000
12,5	SER342	9,563	HB2 - iii(H8)C5	-0,145	-0,186	-0,169	-0,174	0,110	-0,070
12,5	ASN405	10,545	O - i(O2)COO(-)	0,850	0,683	0,717	0,810	-	-
12,5	LYS402	10,728	C - iii(H5)C4	-17,514	-32,644	-19,317	-17,710	-	-
13,0	HIS440	8,240	HN - iii(H6)C4	0,597	0,552	0,562	0,508	-	-
13,0	VAL409	9,279	C - i(O1)COO(-)	-1,311	-1,364	-1,373	-1,390	-	-
13,0	GLN397	9,362	HN - iii(H5)C4	-1,130	-1,083	-1,099	-1,095	0,260	-0,010
13,0	GLU425	10,476	O - ii(H14)C9	19,815	19,926	19,883	19,886	-	-
13,0	GLU382	10,689	O - ii(H16)C12	16,843	16,892	16,863	16,855	-	-

¹ Energias calculadas em [85].

Tabela 3.12: A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-311+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)				
				Conjunto de bases 6-311+G(d) com cargas pontuais			Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	$\epsilon = 0$	$\epsilon = 40$
3,0	LEU453	1,801	HD11 - iii(H9)C5	3,691	2,650	1,423	-6,700	-4,540
4,5	LEU387	1,891	HD21 - ii(H17)C12	5,059	4,246	3,464	-0,800	-0,740
4,5	ASN391	2,170	HB2 - iii(H1)C2	12,219	10,778	9,831	-1,580	-0,960
5,0	ARG410	1,983	HH21 - i(O2)COO(-)	-81,268	-80,057	-81,710	-84,850	-16,500
5,0	PHE403	2,612	HZ - iii(H1)C2	3,623	3,292	3,128	-3,530	-2,810
5,0	ILE388	2,668	HD13 - iii(H8)C5	1,397	1,567	1,296	-0,220	-1,340
5,0	LEU430	2,828	HD23 - ii(H14)C9	-1,096	-0,831	-1,215	-4,620	-3,160
5,0	LEU407	3,592	HD11 - ii(H15)C10	5,976	6,100	5,923	-0,580	-2,720
5,5	VAL433	1,900	HG11 - iii(H7)C5	4,127	2,934	1,543	-2,930	-2,700
5,5	ARG485	2,419	HD2 - i(H11)C7	-31,544	-32,117	-32,397	-37,440	-1,570
6,0	ALA449	2,087	HB1 - iii(H8)C5	2,280	1,690	1,155	0,240	-0,040
6,5	SER489	1,411	HG - i(H12)C7	13,184	10,089	8,920	-15,810	-7,720
6,5	TYR411	1,878	HH - i(O1)COO(-)	-5,938	-7,608	-9,565	-5,450	-1,380
7,0	GLY431	4,790	HA1 - ii(H15)C10	-0,235	-0,054	-0,090	-0,520	0,140
7,0	GLU450	5,041	HA - iii(H9)C5	30,390	30,408	30,452	30,510	1,020
7,5	GLY434	2,176	HA1 - iii(H4)C4	0,306	0,184	-0,416	-0,420	-0,170
7,5	CYS392	3,641	SG - iii(H6)C4	-0,080	0,283	0,069	0,790	-0,030
7,5	LEU457	4,156	HD11 - i(H11)C7	-1,202	-1,151	-1,166	-0,480	-1,250
8,0	PHE395	4,061	HB1 - iii(H5)C4	-0,931	-0,755	-0,805	-0,960	-1,410
8,0	TYR452	4,445	HB2 - iii(H7)C5	-0,895	-1,189	-0,851	-2,150	-0,300
8,5	LYS414	2,475	HZ3 - i(O1)COO(-)	-89,591	-87,983	-89,225	-81,410	-9,160
8,5	CYS438	3,412	SG - iii(H6)C4	0,269	0,262	-0,017	0,370	-0,030
8,5	CYS437	3,704	HB1 - iii(H4)C4	-0,768	-0,608	-0,680	-0,780	-0,050

¹ Energias calculadas em [85].

Tabela 3.13: A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-311+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)				
				Conjunto de bases 6-311+G(d) com cargas pontuais			Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	$\epsilon = 0$	$\epsilon = 40$
8,5	GLN390	6,154	HB2 - ii(H16)C12	-1,691	-1,602	-1,659	-0,250	-0,070
8,5	LEU394	6,224	HB2 - iii(H1)C2	-1,445	-1,338	-1,315	-0,490	-0,090
8,5	VAL456	6,374	HG11 - ii(H14)C9	-1,411	-1,335	-1,375	-2,200	-0,200
8,5	PRO384	6,467	O - ii(H16)C12	1,855	1,732	1,800	0,940	0,050
9,0	PHE488	3,650	HB2 - i(H13)C7	5,786	5,909	5,832	0,350	-0,450
9,0	LYS432	5,402	C - iii(H4)C4	-19,519	-19,408	-19,445	-39,630	-0,990
9,0	LYS389	6,452	N - iii(H2)C2	-18,905	-19,010	-18,979	-23,940	-0,640
9,0	ASN429	6,457	O -iii(H3)C3 / O	-1,428	-1,136	-1,244	1,290	-0,100
9,5	MET446	6,077	HA - iii(H8)C5	1,260	1,190	1,228	-0,310	-0,010
9,5	SER427	6,532	HA - ii(H14)C9	0,607	0,557	0,601	1,250	0,040
9,5	VAL426	6,865	HG13 - ii(H14)C9	1,284	1,014	1,212	0,660	-0,030
9,5	ASN386	6,984	O - ii(H16)C12	-0,762	-0,827	-0,769	0,120	-0,040
9,5	SER454	7,252	HN - iii(H9)C5	0,020	0,253	0,208	-0,260	-0,020
9,5	GLN404	7,667	HA - ii(H15)C10	1,544	1,476	1,545	-2,120	0,040
10,0	SER435	4,889	HN - iii(H4)C4	-0,388	-0,315	-0,325	-1,370	-0,020
10,0	PRO486	5,606	HA - i(H11)C7	3,281	3,071	3,212	0,080	0,040
10,0	ARG445	5,714	O - iii(H8)C5	-18,220	-18,310	-18,286	-23,970	-0,670
10,0	CYS448	5,814	C - iii(H7)C5	-0,371	-0,525	-0,471	-0,650	-0,010
10,0	GLU383	6,565	OE2 - i(H12)C7	34,014	33,278	33,856	32,790	1,070
10,0	GLU393	7,067	HN - iii(H6)C4	19,867	20,530	20,043	28,690	0,660
10,0	ASP451	7,196	HN - iii(H7)C5	20,079	20,340	20,250	25,030	0,690
10,0	ALA406	7,870	O - i(O2)COO(-)	0,248	-0,016	0,047	0,190	0,010
10,5	LYS436	5,316	HN - iii(H4)C4	-0,922	-0,833	-0,862	-0,850	-0,040
10,5	GLN385	8,059	O - ii(H16)C12	0,978	0,940	0,980	0,240	0,050
10,5	VAL455	8,675	HN - iii(H9)C5	-0,714	-0,705	-0,724	-1,230	-0,100
11,0	ALA490	5,610	HN - i(H13)C7	-0,528	0,117	-0,067	0,350	0,180

¹ Energias calculadas em [85].

Tabela 3.14: A tabela mostra as energias obtidas pelos métodos HF, B3LYP e CAM utilizando o conjunto de bases 6-311+G(d) (na presença do background de cargas pontuais), comparando-as com as energias calculadas no resultado [85] através do método de solvatação COSMO, com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$, via funcional e base GGA + D/PBE/DNP.

Raio (Å)	Resíduos	Mínima Distância (Å)	HSA-IBU01	Energia (kcal/mol)				
				Conjunto de bases 6-311+G(d) com cargas pontuais			Conjunto de bases DNP ¹	
				HF	B3LYP	CAM	$\epsilon = 0$	$\epsilon = 40$
11,0	ARG484	6,335	O - i(H13)C7	-19,532	-19,901	-19,799	-27,590	-0,810
11,0	ARG348	8,313	H22 - i(H11)C7	-23,755	-23,720	-23,727	-36,300	-0,880
11,0	ARG428	8,355	O - ii(H15)C10	-19,336	-19,431	-19,383	-35,470	-0,770
11,0	LEU408	8,722	N - i(O2)COO(-)	1,980	2,238	2,219	3,500	0,130
11,5	GLU396	7,408	HN - iii(H5)C4	17,880	18,261	18,078	20,960	0,640
11,5	PRO447	8,223	O - iii(H7)C5	0,919	0,899	0,919	-0,260	-0,010
11,5	LEU398	9,160	HD13 - iii(H1)C2	-1,037	-1,002	-1,007	-1,860	-1,000
11,5	VAL344	9,485	HG21 - i(H11)C7	-0,693	-0,693	-0,699	0,210	-0,040
12,0	LEU491	6,370	HN - i(H13)C7	4,501	4,106	4,181	1,720	-0,030
12,0	CYS487	6,374	C - i(H13)C7	-0,386	-0,910	-0,555	1,050	0,180
12,0	GLU400	7,945	HA - iii(H5)C4	21,418	22,032	21,369	30,000	0,840
12,0	LEU460	8,307	HD23 - i(H10)C6	-1,305	-1,206	-1,244	-0,970	-0,110
12,0	ASN458	10,220	HN - i(H10)C6	-0,313	-0,037	-0,098	-1,280	-0,070
12,5	LYS439	7,159	HN - iii(H4)C4	-0,447	-0,301	-0,349	-	-
12,5	ASN483	8,328	O - i(H11)C7 / HN	0,704	-0,226	0,576	2,290	0,320
12,5	GLY399	8,977	O - iii(H5)C4	0,983	0,833	0,891	-0,400	0,000
12,5	SER342	9,563	HB2 - iii(H8)C5	-0,131	-0,251	-0,185	0,110	-0,070
12,5	ASN405	10,545	O - i(O2)COO(-)	1,099	0,883	0,983	-	-
12,5	LYS402	10,728	C - iii(H5)C4	-16,840	-17,162	-17,065	-	-
13,0	HIS440	8,240	HN - iii(H6)C4	0,350	0,237	0,291	-	-
13,0	VAL409	9,279	C - i(O1)COO(-)	-1,063	-1,048	-1,068	-	-
13,0	GLN397	9,362	HN - iii(H5)C4	-1,192	-1,092	-1,148	0,260	-0,010
13,0	GLU425	10,476	O - ii(H14)C9	20,146	20,556	20,252	-	-
13,0	GLU382	10,689	O - ii(H16)C12	17,113	17,083	17,083	-	-

¹ Energias calculadas em [85].

3.5 Gráficos BIRDs (Binding site, Interaction energy and Residues Domain)

Os resultados dispostos nas subseções subsequentes mostram o comportamento do perfil energético das interações entre os aminoácidos selecionados e a molécula do ligante. Para isto, são analisados os gráficos BIRDs (Binding site, Interaction energy and Residues Domain), os quais nos permitem inferir a intensidade da interação intermolecular entre o ligante e os resíduos de interesse. Cada gráfico BIRD abrange os aminoácidos com energias maiores que 3 kcal/mol, os quais se localizam no sítio ativo e num raio de 13 Å a partir do centróide do ligante IBU01. Os resíduos estão dispostos em ordem crescente da mínima distância entre aminoácido e ligante.

Em várias situações, um dado resíduo apresenta energia dentro do valor de corte quando calculado através de um dado método, porém para outro método este mesmo resíduo pode apresentar energia inferior a 3 kcal/mol, pois métodos com correções de longo alcance e inclusão de efeitos eletrostáticos na descrição tendem a diminuir os valores superestimados de energia, de modo que as energias de interação se aproximam de resultados mais realistas. Assim sendo, ocorreram casos em que foram inseridos aminoácidos com energias inferiores à energia de corte a fim de comparação entre as energias individuais. Finalmente, com exceção da análise da convergência de bases, nas outras subseções as energias usadas para as aferições foram calculadas com o conjunto de bases 6-311+G(d).

3.5.1 Convergência de Bases

Os gráficos BIRDs mostram os aminoácidos que interagem mais fortemente com a proteína HSA, de modo que as energias de interação entre ligante e proteína devem ser maiores que a energia de corte aqui adotada (3 kcal/mol). Os gráficos de 3.7 a 3.12 mostram os BIRDs comparando as energias calculadas via métodos HF, DFT e DFT (com funcional CAM-B3LYP) por meio dos conjuntos de base 6-31+G(d) (barras horizontais pretas) e 6-311+G(d) (barras horizontais cinzas).

O perfil energético de interação IBU-ASH para todos os métodos aqui utilizados, mostram que os resíduos polares como as Argininas, Lisinas, Ácido Aspártico e Ácido Glutâmico exibem expressivas energias de interação, com valores maiores que 20 kcal/mol. Por outro lado, aminoácidos apolares como TYR411, ASN391 e SER489 também apresentam energias de interação intensas, de modo que é notável que todos os aminoácidos pertencentes ao sítio ativo revelaram interações significativas, e embora a maioria exiba energias repulsivas, as energias atrativas possuem valores maiores em módulo.

No geral, não foram detectadas variações relevantes entre as energias calculadas por meio do conjunto de bases 6-31+G(d) e 6-311+G(d). Assim, os resíduos com interações mais intensas mantiveram o padrão energético em todos os métodos e as energias de interação convergem em relação às bases adotadas neste estudo.

3.5.2 O efeito das cargas pontuais nas energias de interação

Nos gráficos BIRDS 3.13, 3.14 e 3.15 (para os métodos HF, DFT e DFT com funcional CAM-B3LYP, respectivamente) as interações energéticas mais importantes são desvendadas, de maneira que estas são comparadas mediante a inclusão do background de cargas nos cálculos. Logo, as barras horizontais pretas assinalam as energias computadas sem o background de cargas, enquanto as barras horizontais cinzas representam as energias obtidas na condição em que as cargas pontuais são inseridas em cada cálculo individual. As energias dispostas nos referidos BIRDS foram computadas por meio do conjunto de bases 6-311+G(d).

No caso em que a contribuição das cargas pontuais é adicionada o módulo das energias atrativas diminui, enquanto que o das energias repulsivas aumenta. Assim, os resíduos que na ausência do background de cargas pontuais possuem energias inferiores a 3 kcal/mol, como o LEU407, LEU453, PHE403 e PHE488 também passam a compor o domínio de interações com energia maior que o critério de corte.

É visível que os aminoácidos localizados no sítio de ligação da proteína, como SER489, TYR411, ARG410, ASN391, ARG485, LYS414, LEU407 e PHE488, foram mais sensíveis às interações eletrostáticas proporcionadas pelas cargas pontuais dos resíduos remanescentes. De outra maneira, o método DFT com cargas pontuais superestima as interações entre ligante e proteína fora do sítio ativo (LYS432, ARG484, ARG445 e GLU383), de modo que tais aminoácidos também são bastante suscetíveis às alterações provocadas pela inclusão das cargas pontuais neste método. As diferenças energéticas para as interações obtidas via DFT (com e sem cargas pontuais) superam 15 kcal/mol.

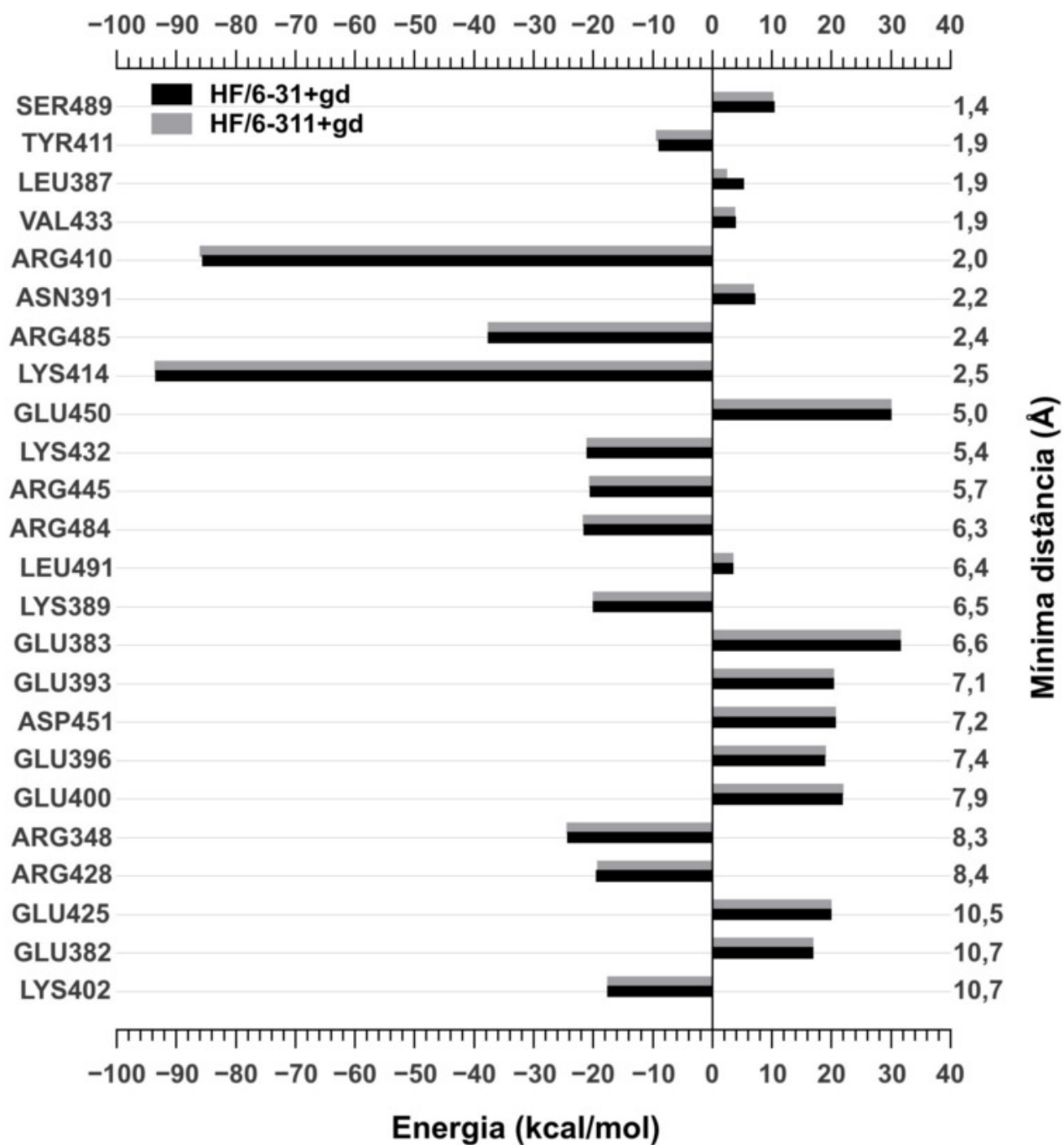
3.5.3 Comparação entre os métodos

Os gráficos BIRDs de 3.16 a 3.17 mostram as interações energéticas mais relevantes, obtidas por intermédio dos métodos HF (barras horizontais cinzas com listras diagonais cruzadas), MP2 (barras horizontais cinzas com listras diagonais), DFT (barras horizontais pretas) e DFT com funcional CAM-B3LYP (barras horizontais cinzas) executados com o conjunto de base 6-311+G(d).

Para uma visualização mais clara, as energias dos 25 resíduos computadas com os métodos sem cargas pontuais estão dispostas nos gráficos 3.16 e 3.17. Com exceção dos aminoácidos LYS432, LYS389 e LYS402, os resíduos mais sensíveis aos efeitos eletrostáticos atribuídos às cargas pontuais estão localizados no sítio ativo da HSA. Destaca-se a interação do LEU453 obtida via MP2, uma vez que a mesma torna-se invertida o sentido, tornando-se atrativa quando calculada com tal método.

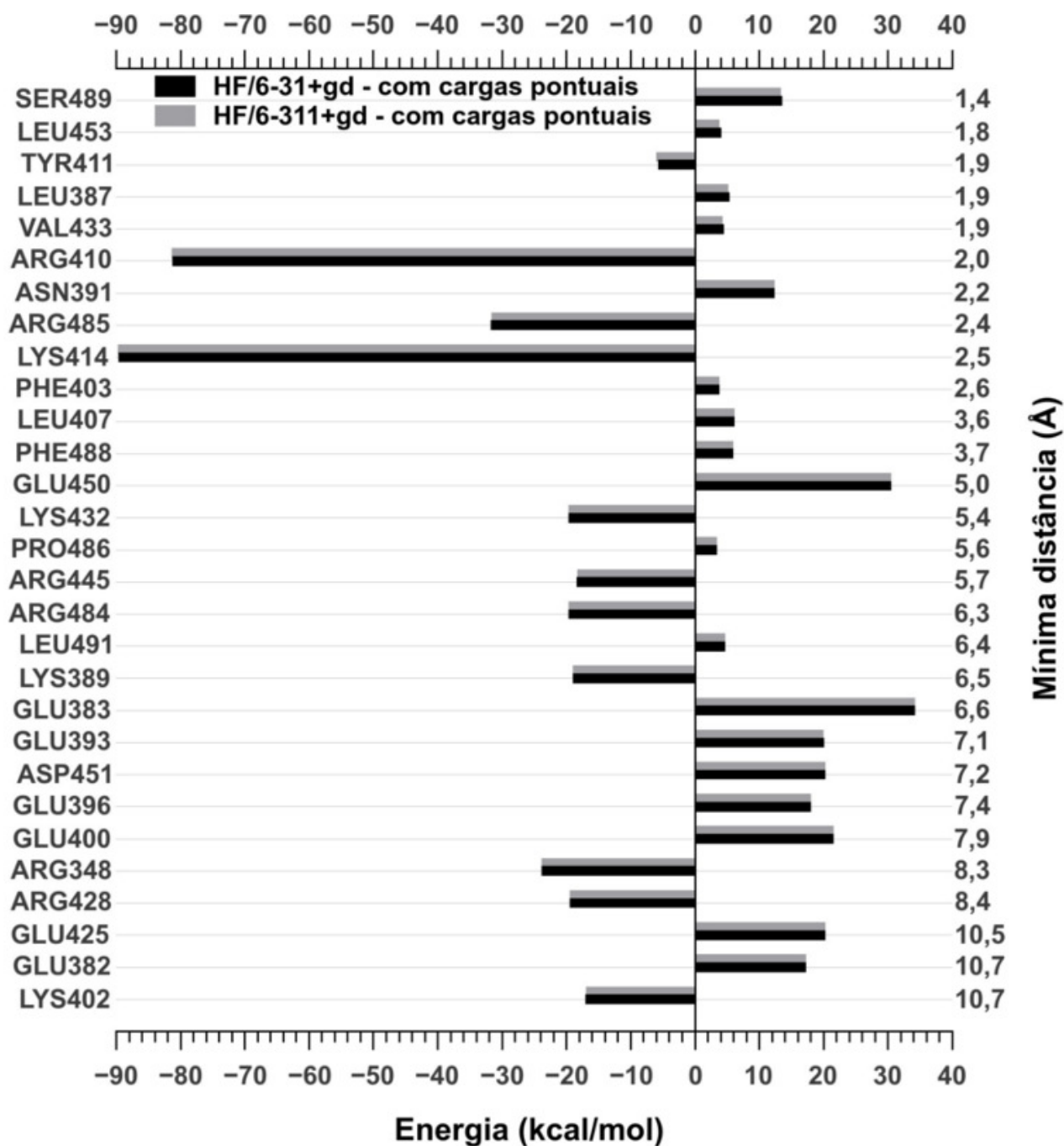
Embora as energias de interação tenham convergido, em relação aos diferentes métodos aqui utilizados, o método DFT descreveu energias de interação sobrestimadas para os resíduos LYS432, LYS389 e LYS402. Todavia, para os aminoácidos do sítio ativo, tal método foi o que calculou as menores energias de interação atrativa. Por outro lado, conforme o gráfico 3.17, a inclusão dos efeitos eletrostáticos gerados pelas cargas pontuais inibiu as variações mais intensas nas energias de interação IBU-ASH.

Figura 3.7: Painel gráfico BIRD obtido para o método quântico HF. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas).



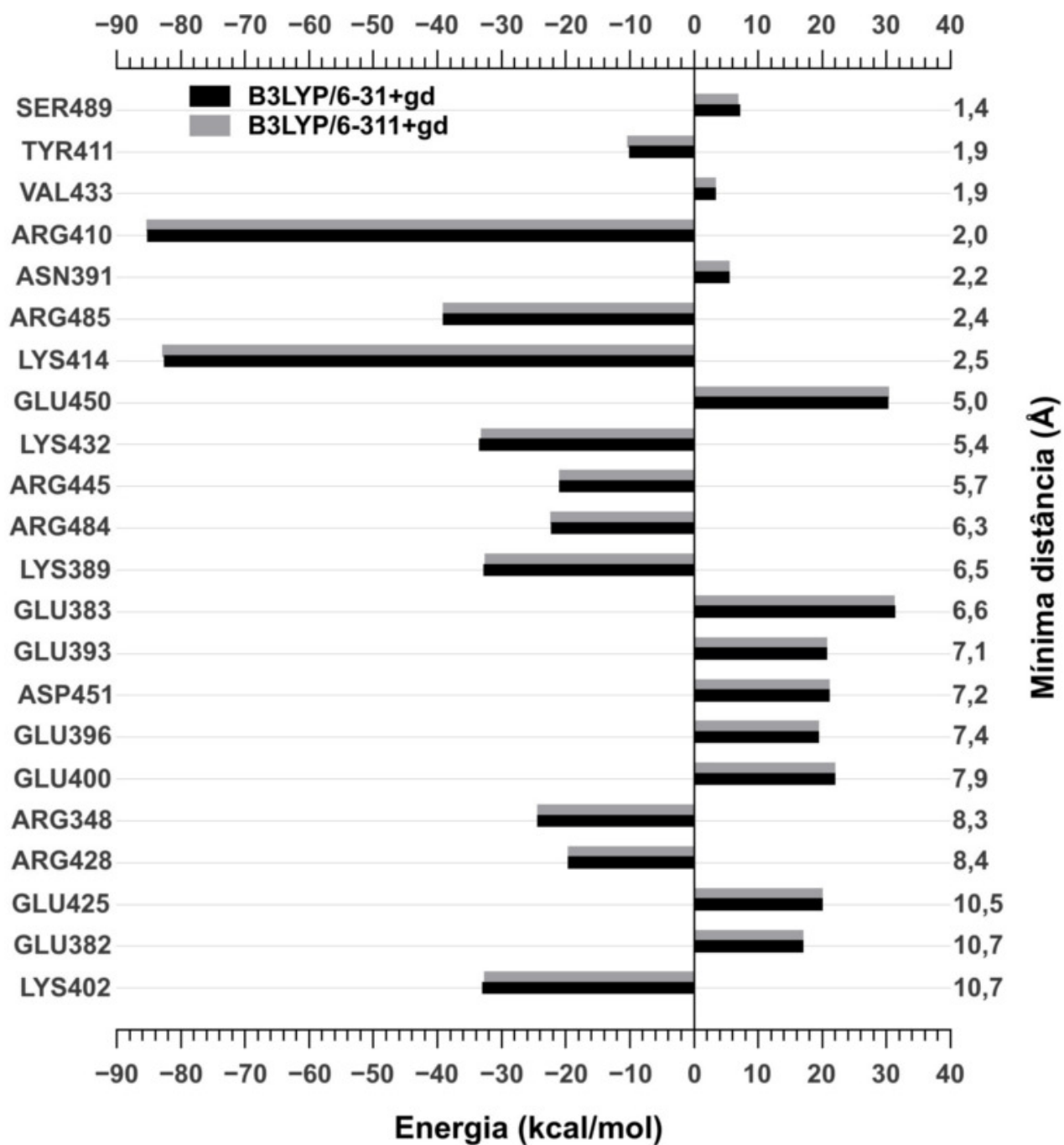
Fonte: Autora, 2017

Figura 3.8: Painel gráfico BIRD obtido para o método quântico HF na presença de cargas pontuais. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas).



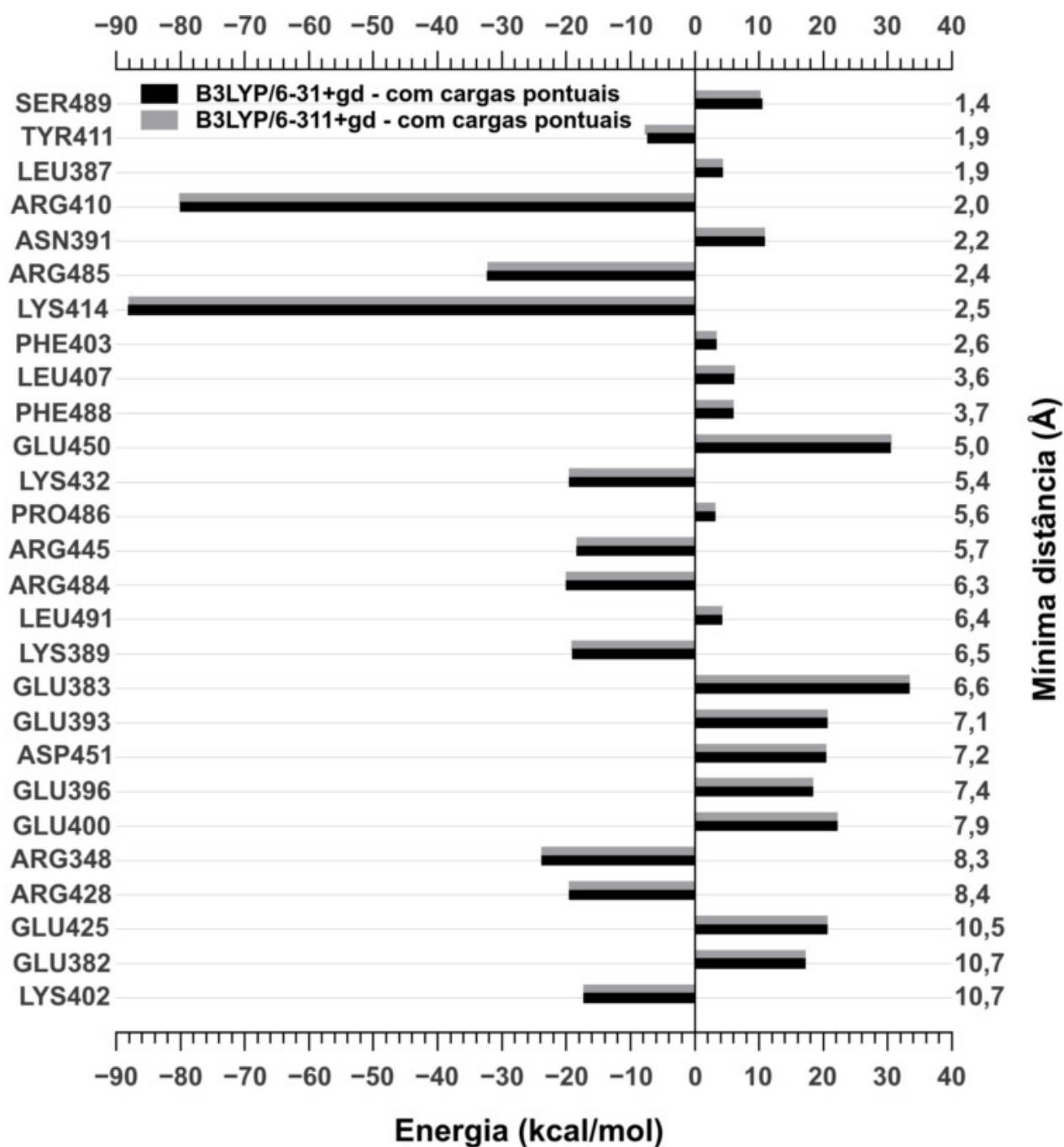
Fonte: Autora, 2017

Figura 3.9: Painel gráfico BIRD obtido para o método quântico DFT. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas).



Fonte: Autora, 2017

Figura 3.10: Painel gráfico BIRD obtido para o método quântico DFT na presença de cargas pontuais. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas).



Fonte: Autora, 2017

Instituto de Física - UFAL

Figura 3.11: Painel gráfico BIRD obtido para o método quântico DFT incluindo correção de longo alcance através do funcional CAM-B3LYP. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas).

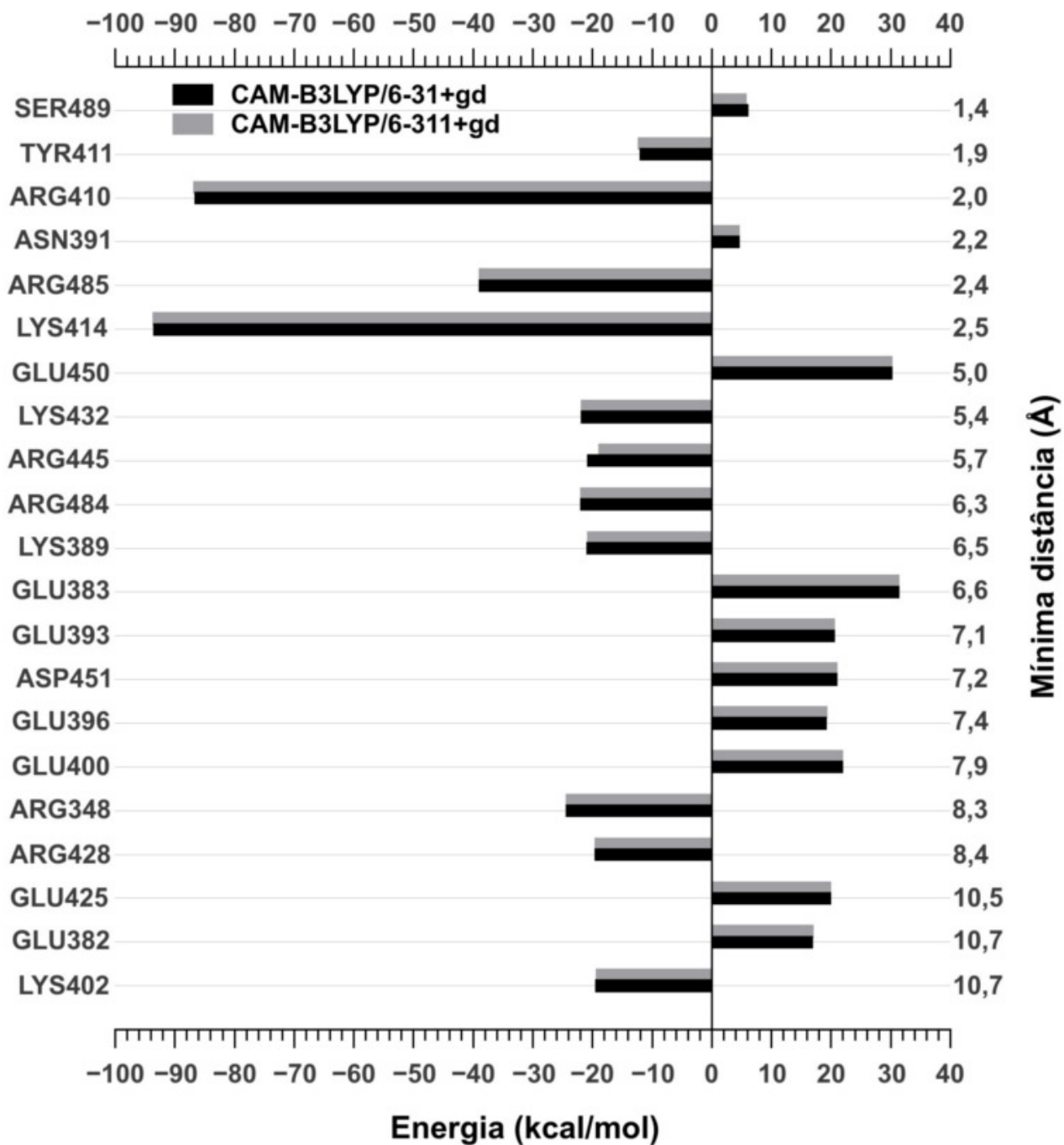
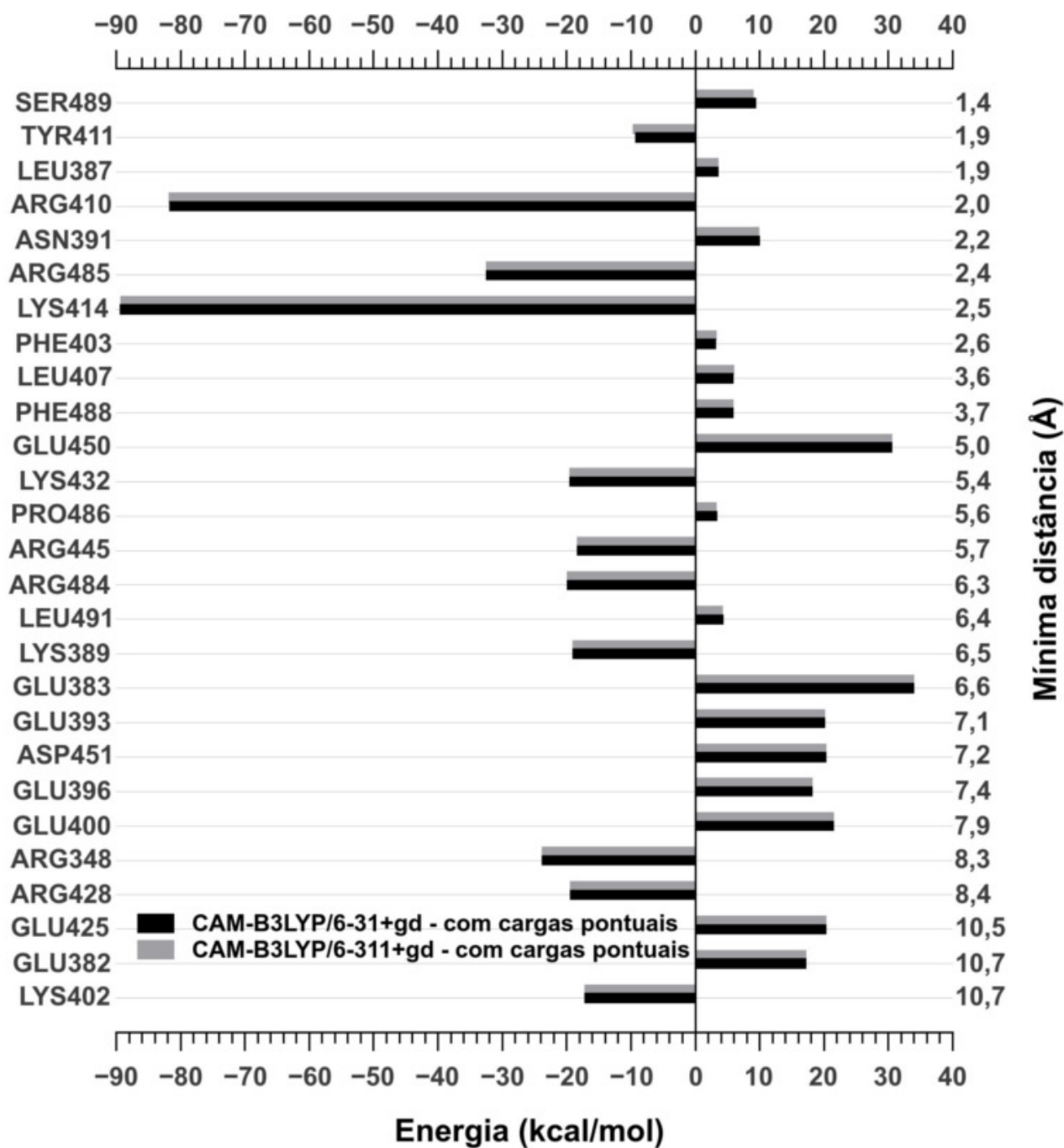


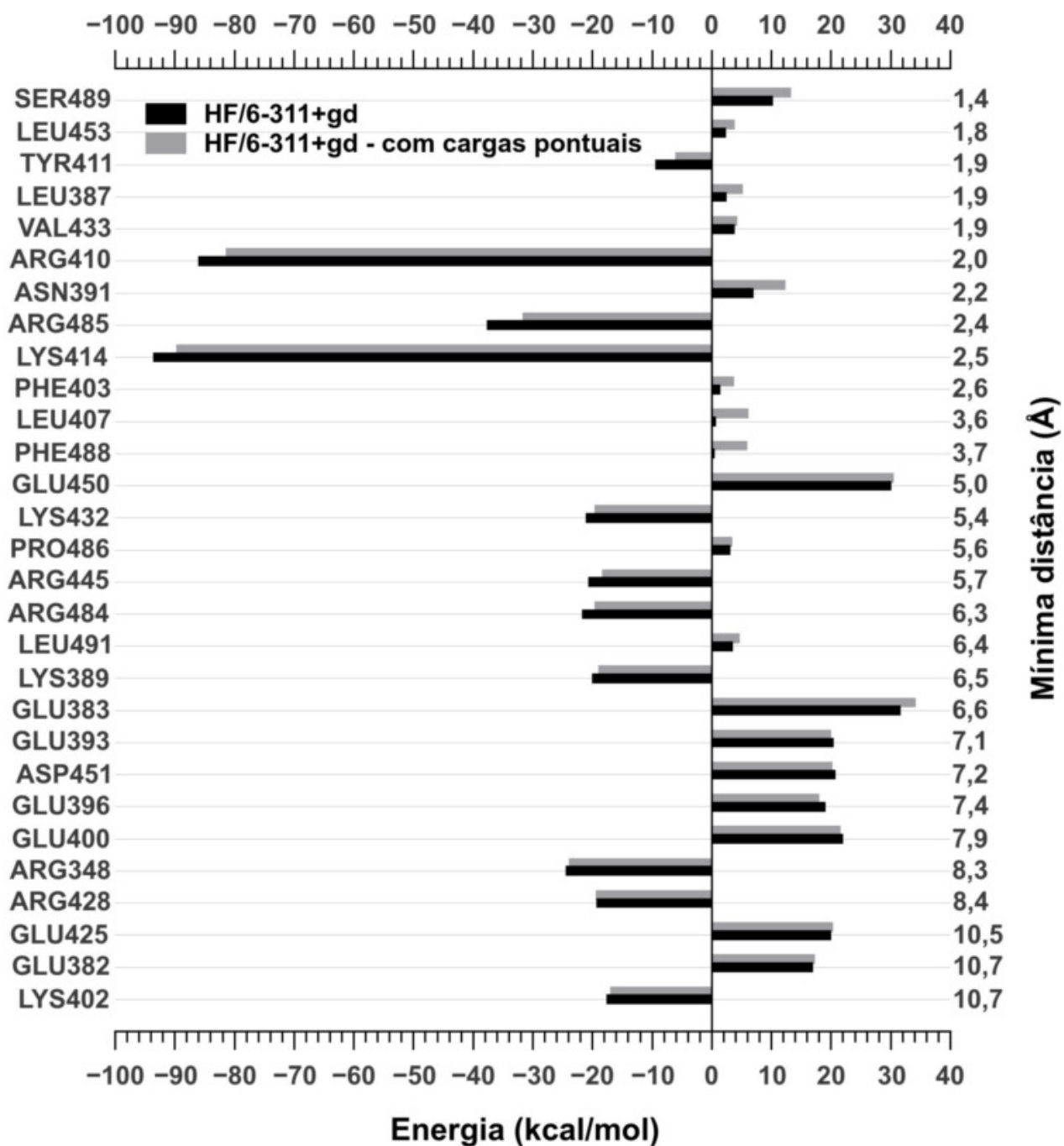
Figura 3.12: Painel gráfico BIRD obtido para o método quântico DFT incluindo correção de longo alcance através do funcional CAM-B3LYP e na presença de cargas pontuais. São comparadas as energias obtidas por meio dos conjuntos de bases 6-31+G(d) (barras horizontais pretas) e 6-311+g(d) (barras horizontais cinzas).



Fonte: Autora, 2017

Instituto de Física - UFAL

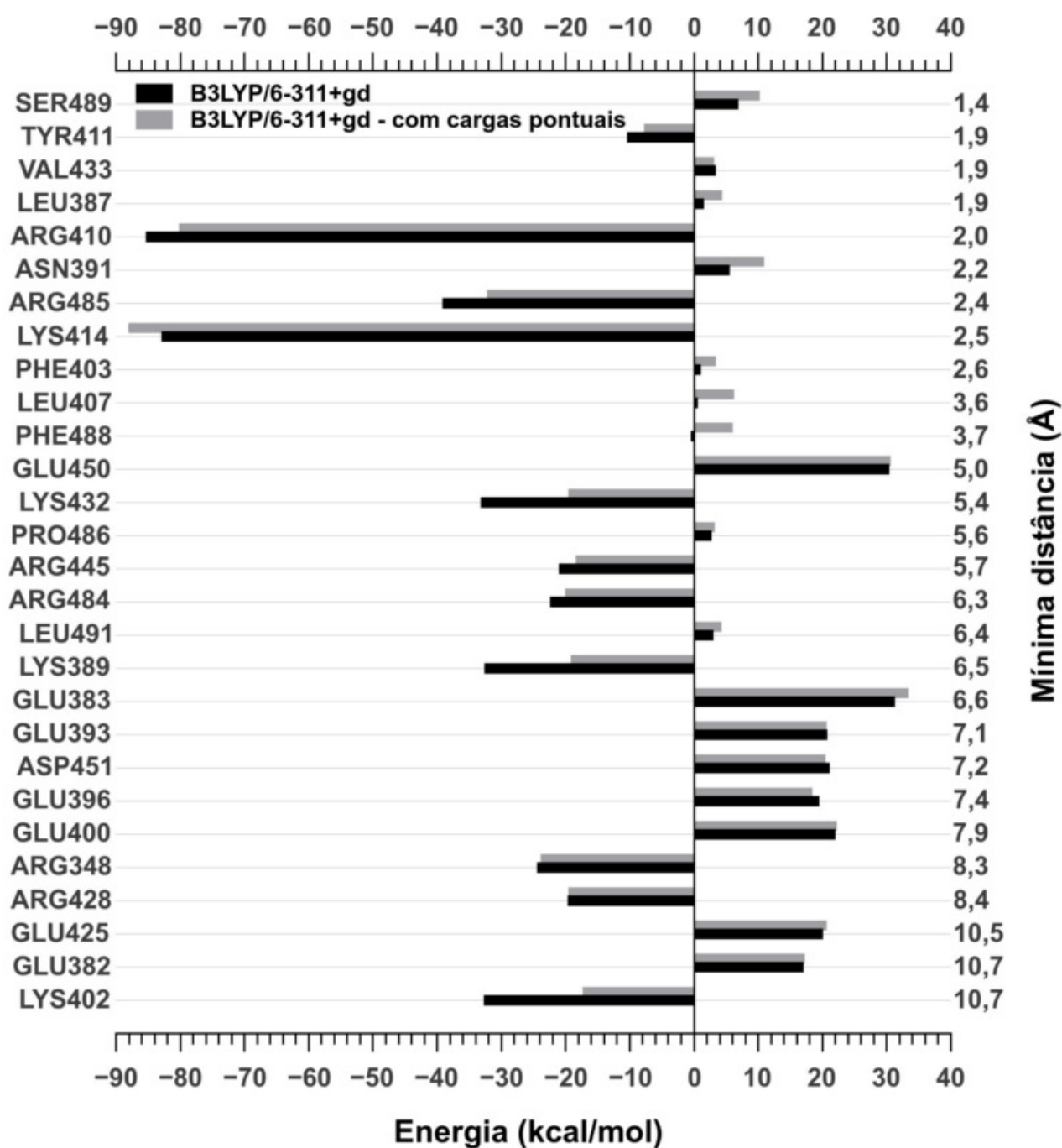
Figura 3.13: Comparação em relação à inclusão de cargas pontuais para as energias calculadas através do método HF com conjunto de base 6-311+G(d). As barras horizontais pretas descrevem o caso em que não são consideradas as cargas pontuais, à medida que as barras horizontais cinzas indicam que a inclusão do background de cargas nos cálculos da energia de interação IBU-HSA.



Fonte: Autora, 2017

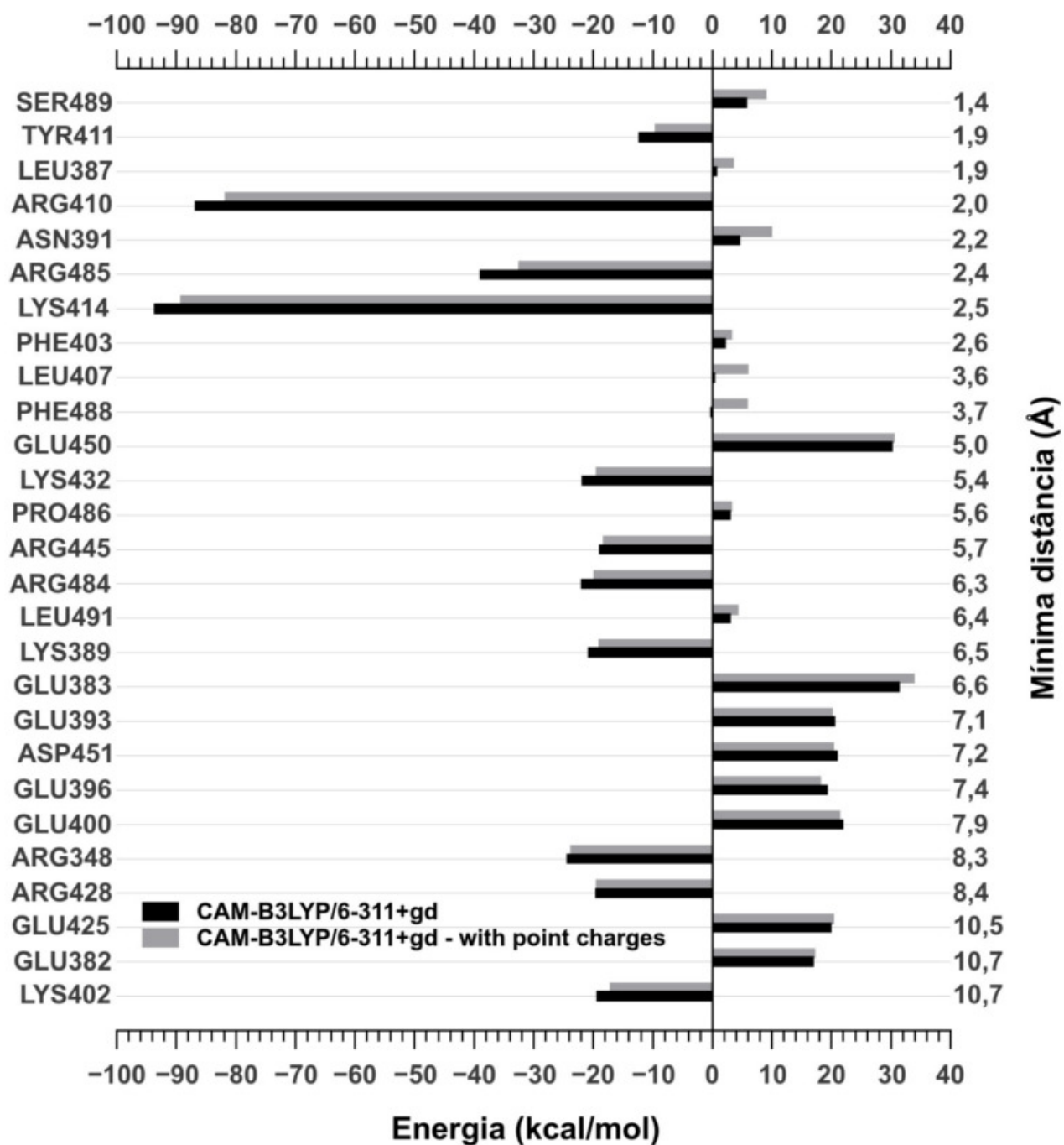
Tese de Doutorado

Figura 3.14: Comparação em relação à inclusão de cargas pontuais para as energias calculadas através do método DFT com conjunto de base 6-311+G(d). As barras horizontais pretas descrevem o caso em que não são consideradas as cargas pontuais, à medida que as barras horizontais cinzas indicam que a inclusão do background de cargas nos cálculos da energia de interação IBU-HSA.



Fonte: Autora, 2017

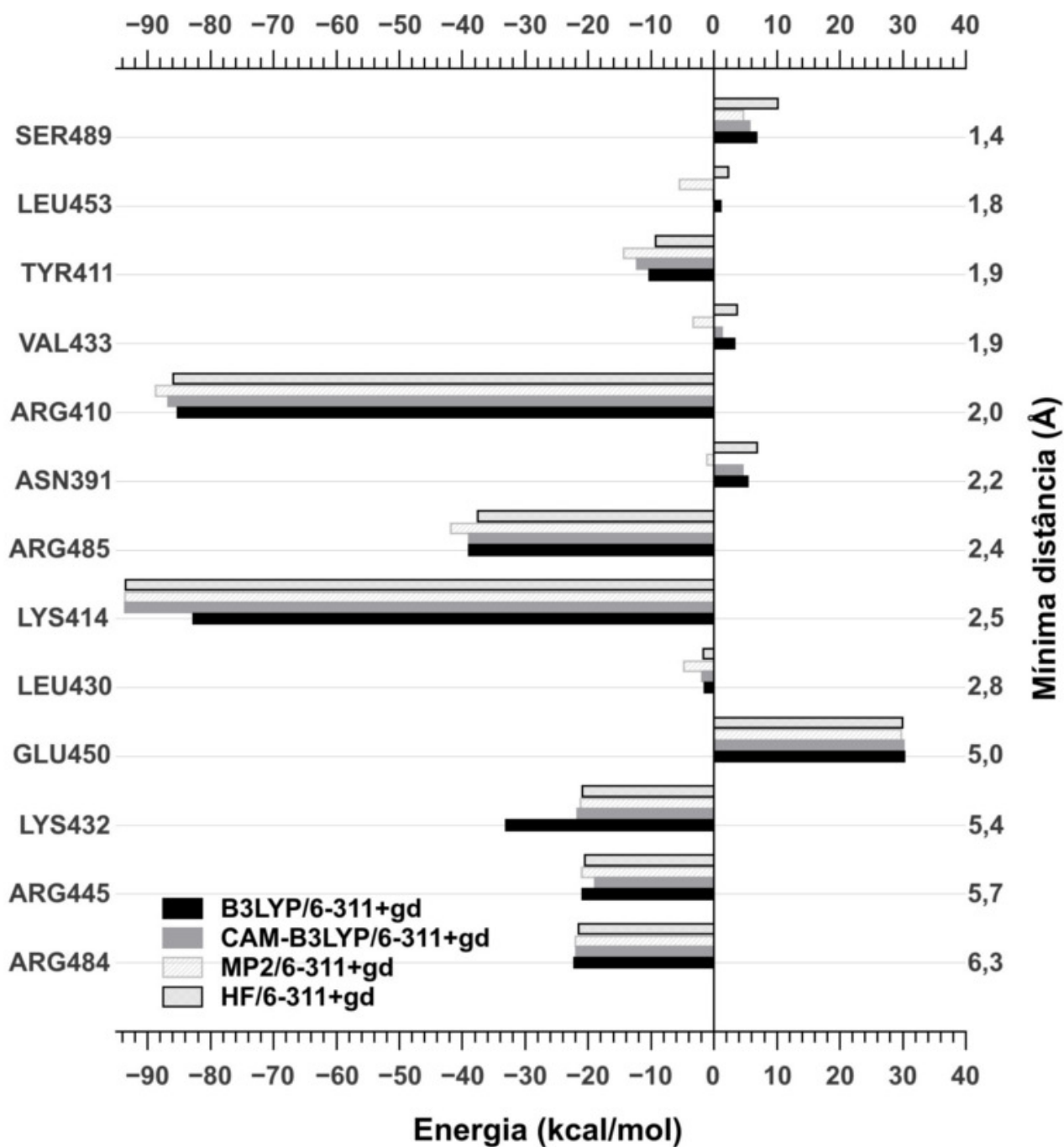
Figura 3.15: Comparação em relação à inclusão de cargas pontuais para as energias calculadas através do método DFT, com funcional para interações de longo alcance CAM-B3LYP, e conjunto de base 6-311+G(d). As barras horizontais pretas descrevem o caso em que não são consideradas as cargas pontuais, à medida que as barras horizontais cinzas indicam que a inclusão do background de cargas nos cálculos da energia de interação IBU-HSA.



Fonte: Autora, 2017

Tese de Doutorado

Figura 3.16: O painel gráfico BIRD (com 17 aminoácidos) obtido para as energias calculadas com conjunto de base 6-311+G(d), sem cargas pontuais, para os métodos HF (barras horizontais cinzas com listras diagonais cruzadas), MP2 (barras horizontais cinzas com listras diagonais), CAM (barras horizontais cinzas) e B3LYP (barras horizontais pretas).



Fonte: Autora, 2017

Instituto de Física - UFAL

Figura 3.17: O painel gráfico BIRD (com 12 aminoácidos) obtido para as energias calculadas com conjunto de base 6-311+G(d), sem cargas pontuais, para os métodos HF (barras horizontais cinzas com listras diagonais cruzadas), MP2 (barras horizontais cinzas com listras diagonais), CAM (barras horizontais cinzas) e B3LYP (barras horizontais pretas).

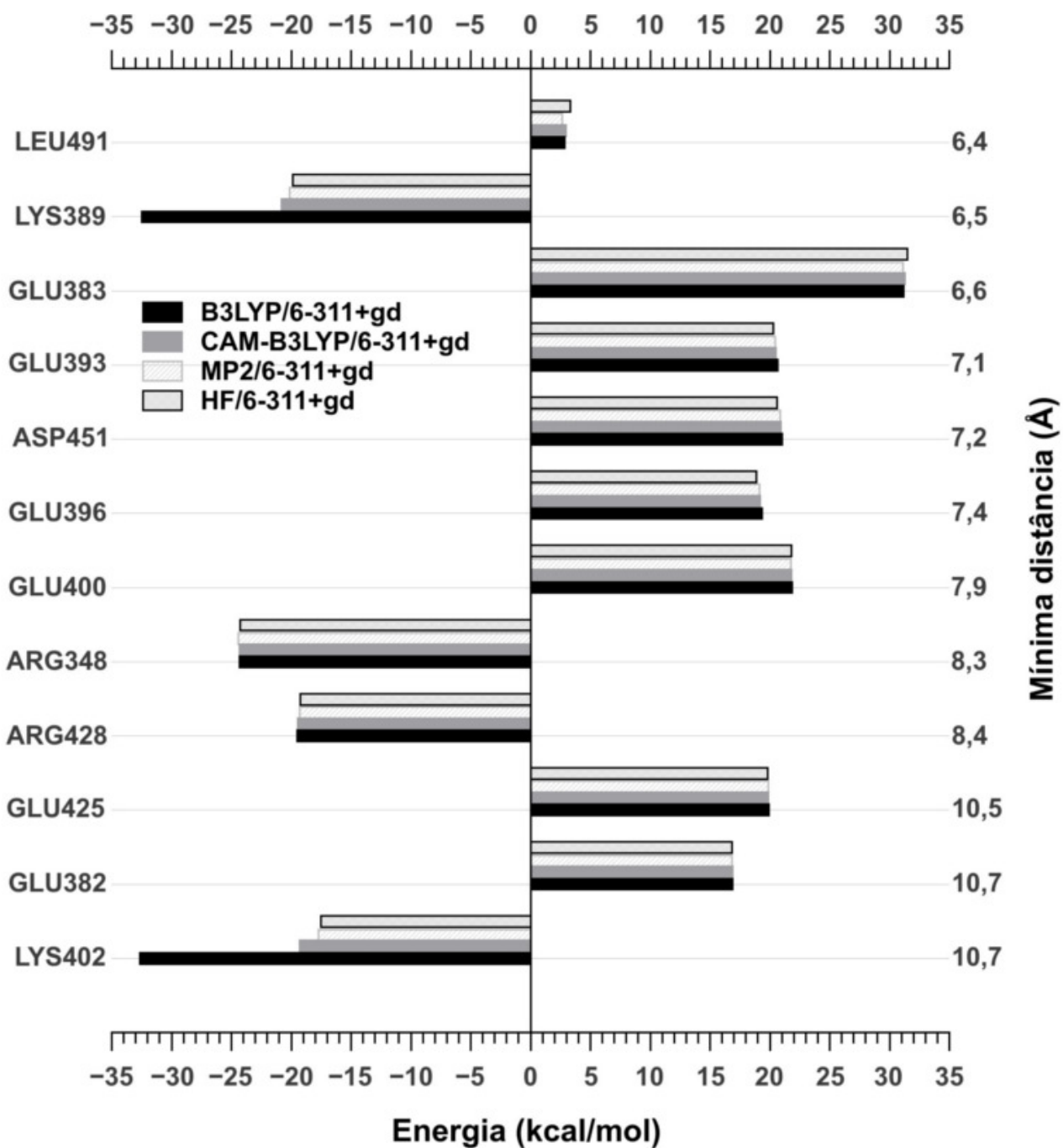
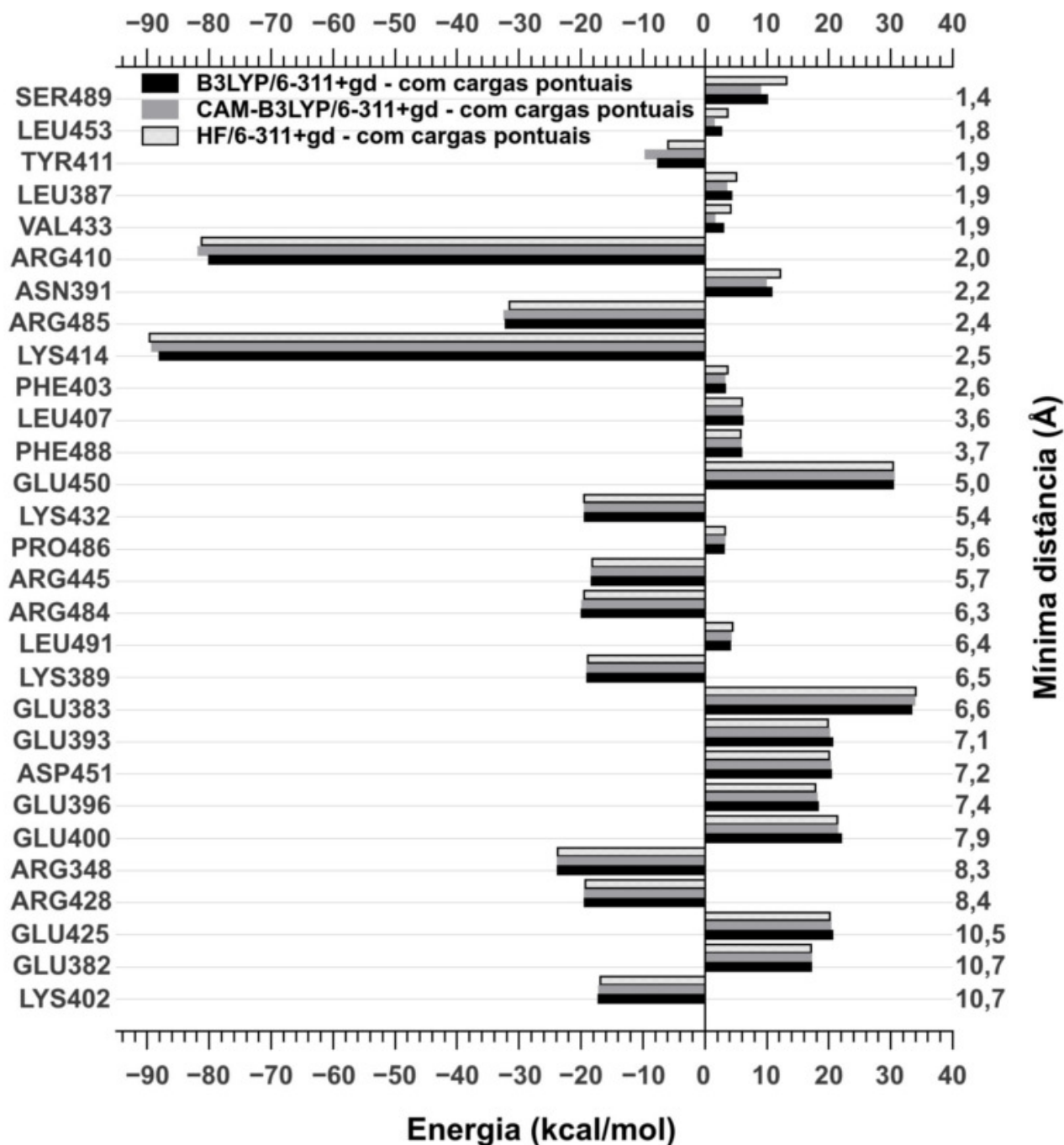


Figura 3.18: O painel gráfico BIRD obtido para as energias calculadas com conjunto de base 6-311+G(d), na presença de cargas pontuais, para os métodos HF (barras horizontais cinzas com listras diagonais cruzadas), MP2 (barras horizontais cinzas com listras diagonais), CAM (barras horizontais cinzas) e B3LYP (barras horizontais pretas).



Fonte: Autora, 2017

Instituto de Física - UFAL

3.6 Características estruturais das interações IBU-HSA

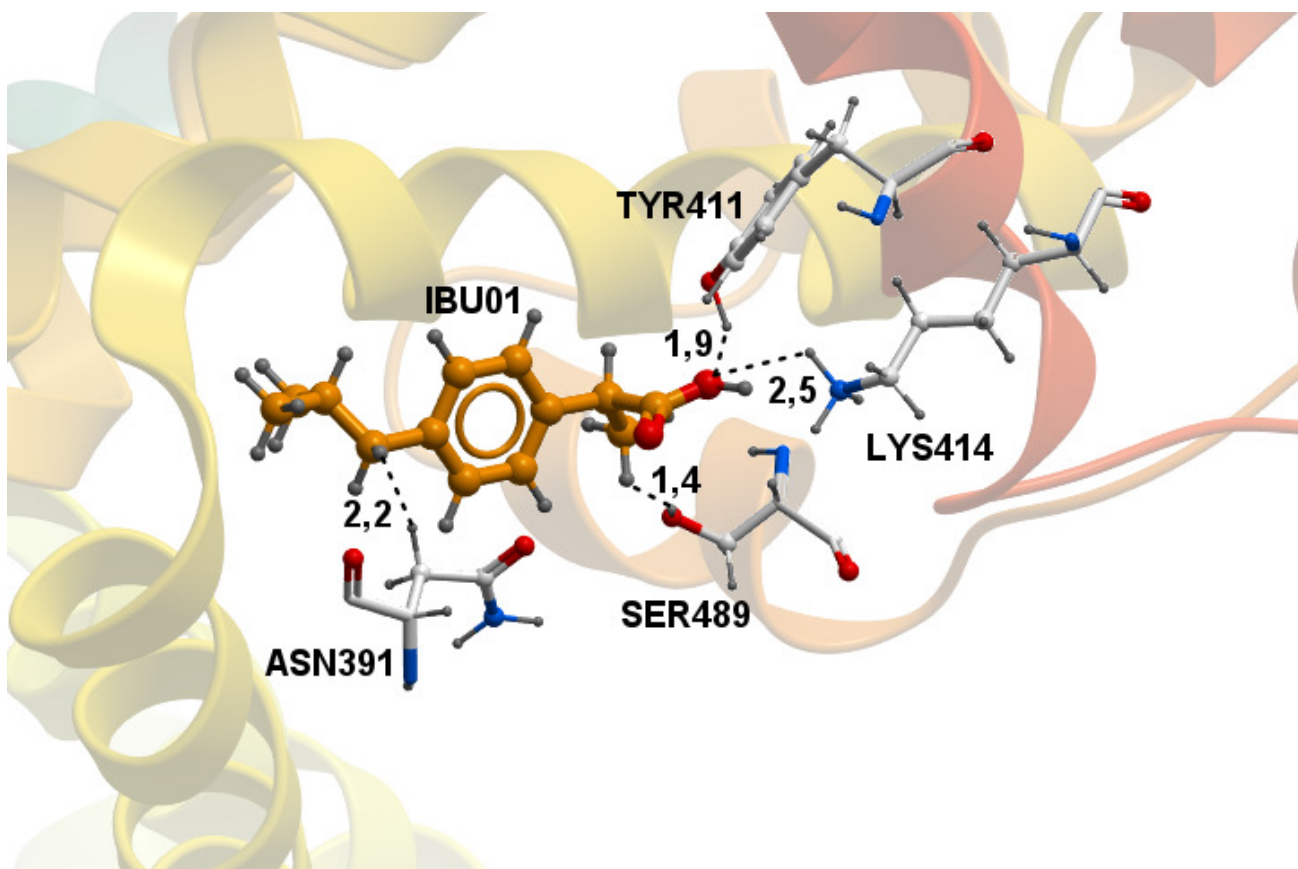
O perfil de interação entre o IBU-ASH (mostrados nos gráficos BIRDS) evidencia que os resíduos cujas energias de interação são mais significativas, situam-se no sítio ativo. Enquanto que os outros resíduos com energias de interação relevantes localizam-se em regiões distantes do sítio ativo e possuem cadeias laterais carregadas. Desta forma, a fim de verificar-se quais características estruturais estão envolvidas e contribuem com a energia de interação entre a micromolécula IBU01 e os resíduos da proteína ASH, as figuras 3.19, 3.20 e 3.21 exibem os resíduos que pertencem à região do sítio de ligação da proteína (com distância mínima de até 5,5 Å) e possuem energia de interação igual ou maior que 5 kcal/mol (em módulo).

Os critérios designados para distinguir quais os aminoácidos mais importantes para a interação ligante-proteína, baseiam-se no fato de que o sítio FA3/FA4 (também denominado sítio de Sudlow II) abriga uma cavidade hidrofóbica formada principalmente por resíduos com cadeias laterais neutras e apolares, como a VAL433, LEU387, LEU407 e PHE403, que interagem diretamente com os grupos Arila e Isobutilo correspondentes às regiões ii e iii da micromolécula IBU, respectivamente. Nestas regiões, dominam interações de van der Waals, mediadas por forças dispersivas do tipo dipolo induzido-dipolo induzido (interações de dispersão de London). Por outro lado, nas proximidades da entrada da cavidade hidrofóbica localiza-se o fragmento centrado nos resíduos ARG410 e TYR411, que também inclui os aminoácidos LYS414 e SER489. Assim, as cadeias laterais positivamente carregadas do ARG410 e LYS414, assim como as cadeias polares neutras da TYR411 e LYS414, interagem com o grupo carboxílico, correspondente à região i do IBU01 (porção polar negativa da micromolécula). Assim, nesta região prevalecem as interações eletrostáticas (para os resíduos ARG410 e LYS414), que proporcionam as energias de interação mais fortes no sistema IBU-ASH, e as interações do tipo dipolo permanente-dipolo induzido (interações de dispersão de Debye) para os aminoácidos TYR411 e SER489.

Finalmente, o arranjo do sítio de Sudlow II estabelece um equilíbrio entre as interações eletrostáticas e de van der Waals, contribuindo assim para a estabilização do complexo macromolecular

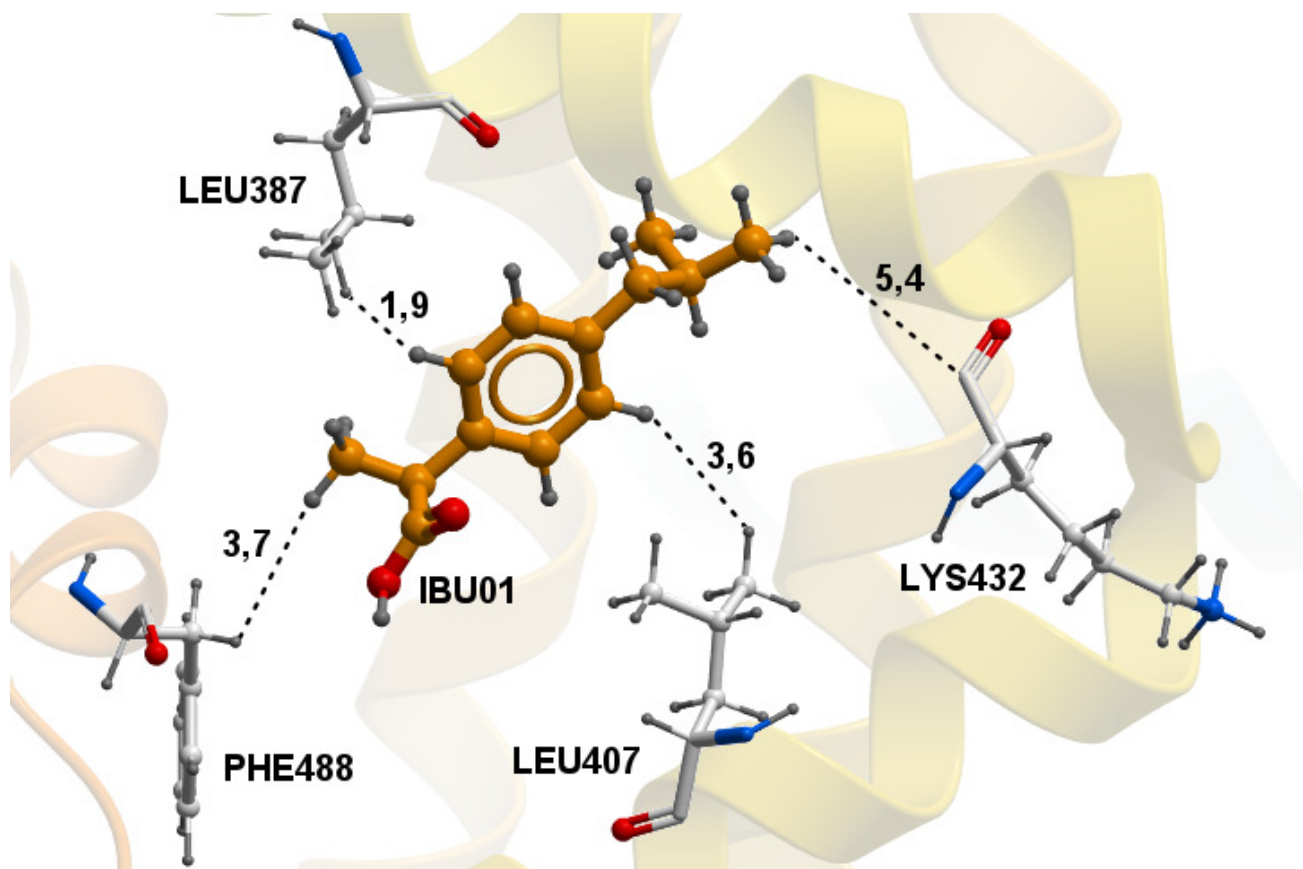
IBU-ASH e conferindo ao ligante maior especificidade e afinidade nos modos de interação com a proteína.

Figura 3.19: Disposição estrutural de 4 dos 13 aminoácidos mais importantes nas interações do complexo IBU-HSA.



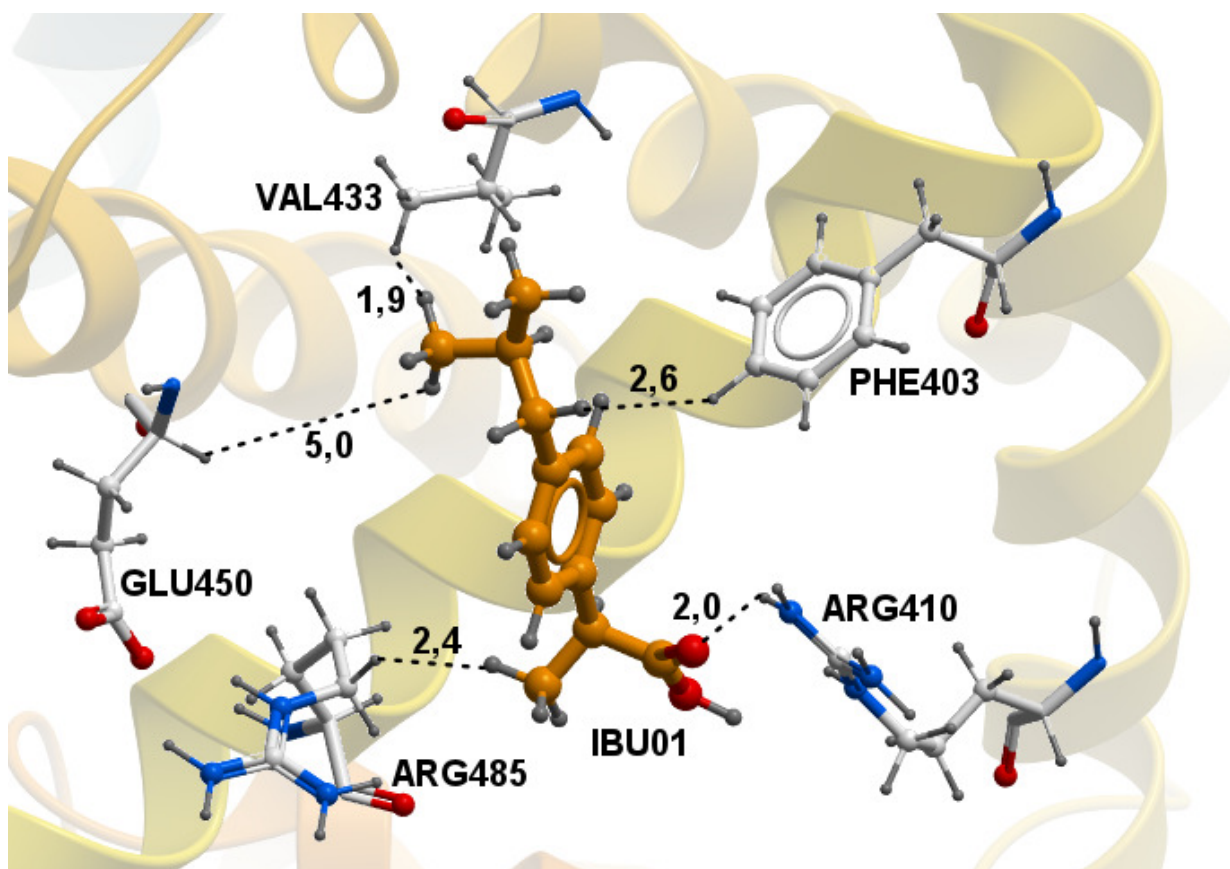
Fonte: Autora, 2017

Figura 3.20: Disposição estrutural de 4 dos 13 aminoácidos mais importantes nas interações do complexo IBU-HSA.



Fonte: Autora, 2017

Figura 3.21: Disposição estrutural de 5 dos 13 aminoácidos mais importantes nas interações do complexo IBU-HSA.



Fonte: Autora, 2017

3.7 Desvio Absoluto (DA) e Relativo (DR)

Uma vez que através dos gráficos BIRDS são identificados os aminoácidos com as interações atrativas ou repulsivas mais intensas, os resultados das próximas subseções servem para aferirmos acerca das flutuações nas energias, devido a inclusão da contribuição das interações de longo alcance. Assim, para avaliar a convergência dos cálculos mediante os variados métodos e determinar as variações decorrentes da inclusão das interações de longo alcance nos cálculos da energia,

foram calculados os Desvios Absolutos (DA) e Relativos (DR) das energias em relação à base 6-311+G(d), ao uso do funcional CAM-B3LYP, à inclusão das cargas pontuais e em relação aos métodos HF e MP2. Por conseguinte, o Desvio Absoluto (DA) é a diferença absoluta entre um dado valor e outro de referência, enquanto que o Desvio Relativo (DR) refere-se à concordância do valor calculado com um nível de referência. Deste modo, o DA (3.3) especifica a variação absoluta entre os valores individuais ao passo que o DR (3.4) está associado à confiabilidade das medidas. Logo, quanto menor o DR, maior a exatidão do resultado, em relação ao dado utilizado como referência.

$$DA = E_{referência} - E_x \quad (3.3)$$

$$DR = \frac{E_{referência} - E_x}{E_{referência}} \quad (3.4)$$

onde x representa os valores calculados por meio dos diferentes métodos e os valores de *referência* correspondem aos métodos em relação aos quais são calculados os Desvios Absoluto e relativo.

3.7.1 Desvios Absoluto (DA) e Relativo (DR) em relação à base 6-311+G(d)

O Desvio Absoluto (DA) tomando como referência os valores de energia calculados com a maior base usada neste estudo, 6-311+G(d), aponta a convergência da energia para os diferentes conjuntos de base. Portanto, foi calculada a diferença entre as energias calculadas com a base menor 6-31+G(d) e com a base maior 6-311+G(d), os valores de energia convergiram e não apresentaram grandes discrepâncias. Assim, com exceção dos aminoácidos LEU387, PHE403, ARG445 e ASN483, que apresentaram DA maior que 1 kcal/mol, as energias de interação entre o IBU e os

resíduos da HSA não exibiram grande variação em relação ao uso do conjunto de base 6-311+G(d), principalmente para os métodos DFT, DFT com correção de longo alcance CAM-B3LYP e métodos em que as cargas pontuais foram incluídas.

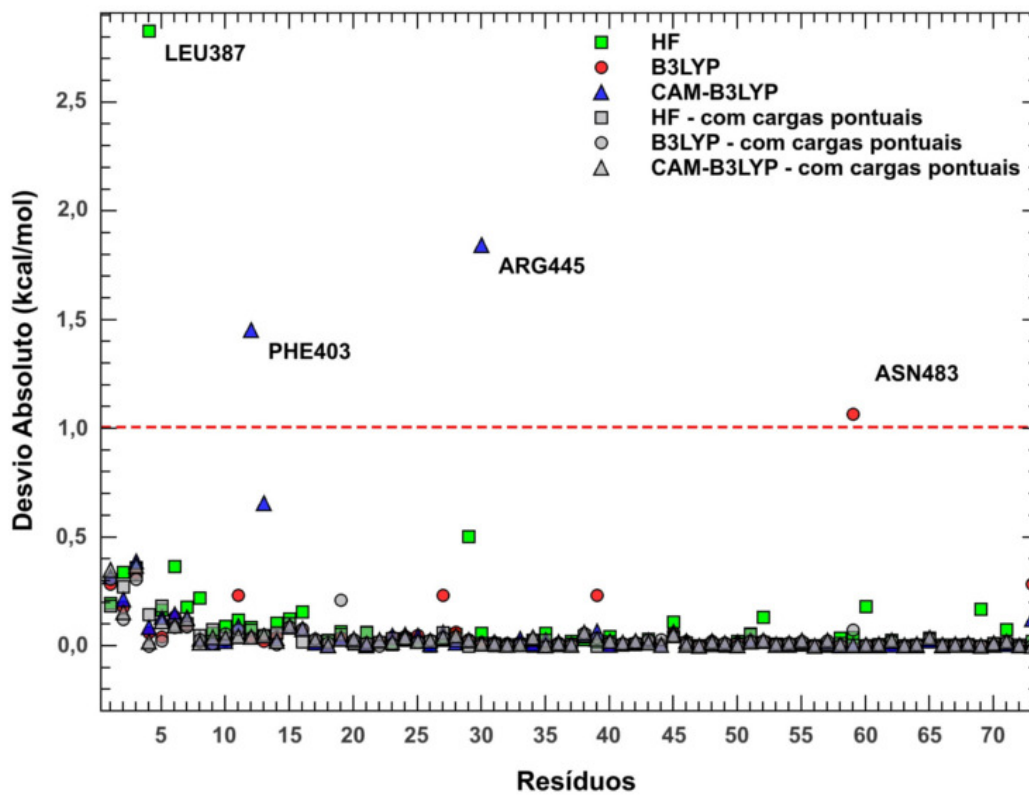
Os gráficos 3.22 e 3.23 expressam o DA e o DR, respectivamente, das energias obtidas com o método HF, DFT e DFT com correção de Coulomb (via funcional CAM-B3LYP) na ausência e presença das cargas pontuais. Destarte, nota-se que para os métodos onde foi incluída a descrição eletrostática do background de cargas pontuais, o DA da energia flutua em torno de valores próximos de zero. Quanto ao DR demonstrado no gráfico 3.23, dentre os seis resíduos que exibiram DR maior que 1, o CYS438 (5,26) e o LEU453 (2,20) expressaram DR mais significativos, evidenciando-se assim a convergência das energias em relação ao uso da base 6-311+G(d). A tabela 3.15 lista a média dos DRs da energia mediante a mudança de base, onde é constatado que as energias de interação obtidas via método HF (sem cargas pontuais) apresentam maior dispersão em relação à mudança do conjunto de bases, enquanto que o método DFT (CAM-B3LYP) com cargas pontuais apresenta a maior média do DR (6,13%), expressando assim maior dispersão nos valores de energia.

Tabela 3.15: Esta tabela mostra a média (em porcentagem) dos desvios relativos da energia, considerando-se a mudança de base, isto é, quando com as energias calculadas via base 6-311+G(d) são definidas como os valores de referência.

Média dos Desvios Relativos		
		Com Carga Pontuais
HF	4,07%	0,32%
DFT	1,30%	0,13%
DFT (CAM-B3LYP)	2,39%	6,13%

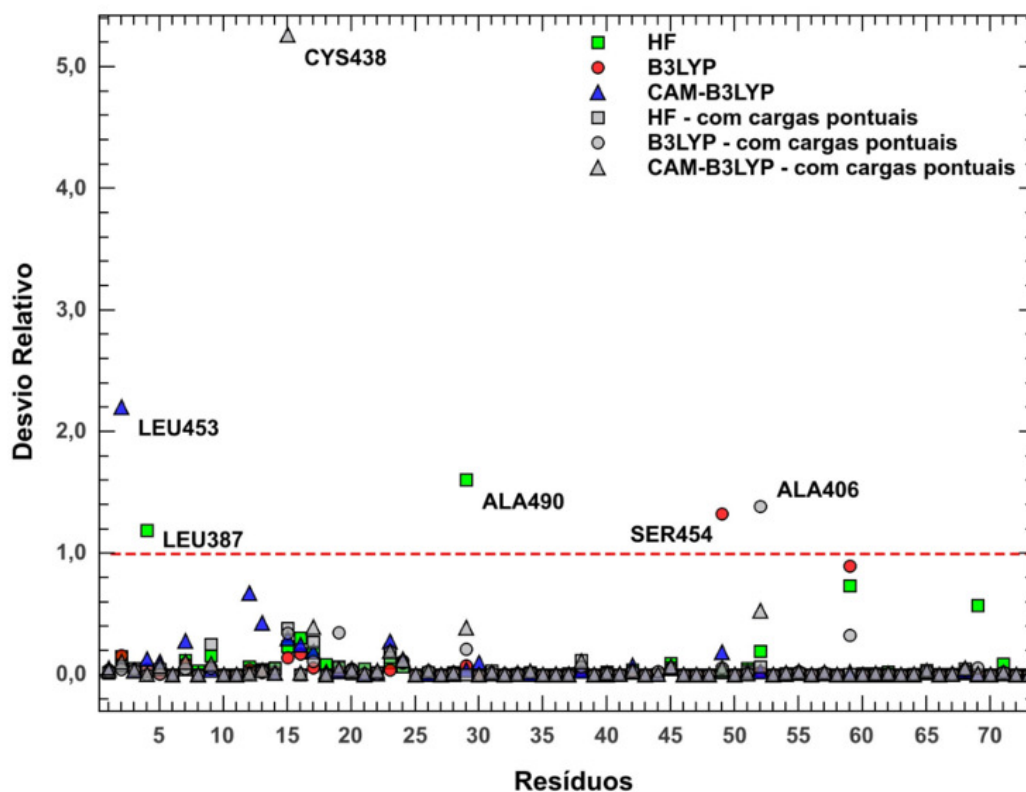
Fonte: Autora, 2017

Figura 3.22: Desvio Absoluto tomando como referência o uso do conjunto de base 6-311+G(d). A maioria dos desvios flutua em valores muito próximos de zero, onde somente os resíduos LEU387, PH403, ARG445 e ASN483 obtiveram DA superior a 1kcal/mol (indicado pela linha tracejada vermelha). O eixo das abscissas dispõe as mínimas distâncias em ordem crescente. Os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias.



Fonte: Autora, 2017

Figura 3.23: Desvio Relativo tomando como referência o uso do conjunto de base 6-311+G(d). Excetuando os resíduos CYS438, LEU453, LEU387, ALA490, SER454 e ALA406 que mostram desvios maiores que 1, a maioria apresenta DR abaixo desse valor. Os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias.



Fonte: Autora, 2017

3.7.2 Desvios Absoluto (DA) e Relativo (DR) em relação ao uso do funcional CAM-B3LYP

Na presente subseção, verificamos o DA (gráfico 3.24) e DR (gráfico 3.25) usando como referência o método DFT com correção de longo alcance CAM-B3LYP na ausência e presença de cargas pontuais, sendo os dois casos logrados com base 6-311+G(d). Para melhor visualização da

região delimitada pelo retângulo com borda preta, o gráfico ao lado direito (Figura 3.24) se trata do aumento da referida área do gráfico. Sendo assim, a maioria dos aminoácidos que sofreram maiores desvios absolutos em relação ao uso do funcional de longo alcance CAM-B3LYP foram aqueles cujas energias foram efetuadas com os métodos DFT e MP2, na ausência das cargas pontuais.

Uma importante constatação sobre o DA em relação ao uso do funcional CAM-B3LYP é a de que os resíduos polares como LYS414 (10,80 kcal/mol), LYS432 (11,33 kcal/mol), LYS389 (11,67 kcal/mol) e LYS402 (13,33 kcal/mol), calculados via DFT sem cargas pontuais, demonstram os maiores DAs. Por conseguinte, o maior número de resíduos (um total de 17) com DA maior que 1 kcal/mol tiveram suas energias obtidas pelo método MP2 na ausência de cargas pontuais. Da mesma forma, o DR indica que 14 aminoácidos calculados via MP2 (sem carga pontual) apresentaram DR acima de 1 kcal/mol.

Como os resíduos estão dispostos segundo a ordem crescente das mínimas distâncias, a área destacada pelo retângulo nos gráficos 3.24 e 3.25 representa as regiões mais próximas do sítio ativo. Assim, esta região concentra a maior quantidade de resíduos com DA e DR relevantes, mostrando que os aminoácidos do sítio ativo são os que mais sofrem variações nos valores de energia em detrimento do uso do funcional CAM-B3LYP.

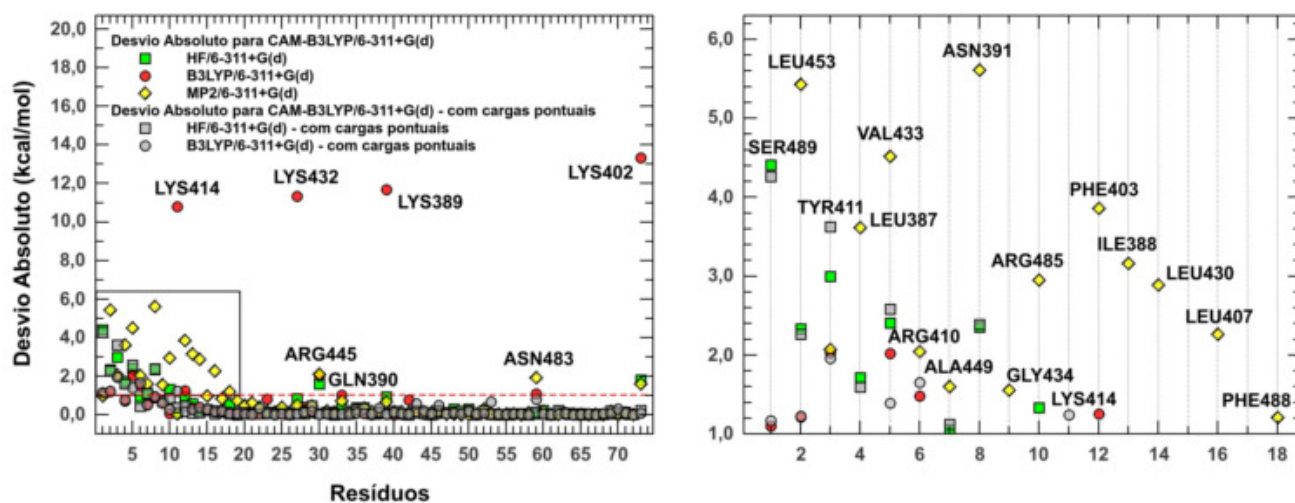
Na tabela 3.16 são notórios os altos valores nas médias dos desvios relativos em relação ao uso do funcional CAM-B3LYP, principalmente para os métodos onde não é inclusa a contribuição eletrostática das cargas pontuais. Desta forma, o uso do funcional CAM-B3LYP implica em mudanças significativas nas energias de interação, quando compara-se as energias calculadas através de métodos na ausência de cargas pontuais.

Fonte: Autora, 2017

Tabela 3.16: A tabela mostra a média (em porcentagem) dos desvios relativos da energia, considerando-se como os valores de referência as energias calculadas com a inclusão das interações de longo alcance via funcional CAM.

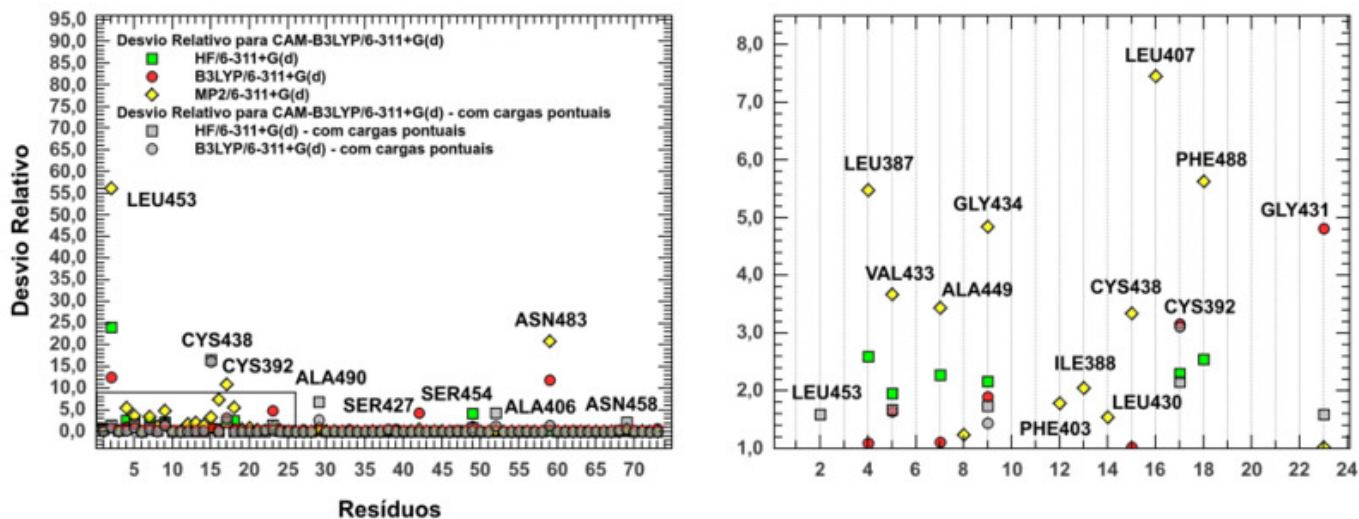
Média dos Desvios Relativos		
		Com Carga Pontuais
HF	69,59%	62,67%
DFT	70,75%	46,87%
MP2	188,98%	-

Figura 3.24: Desvio Absoluto tomando como referência o uso do funcional CAM-B3LYP. Para melhor visualização da região delimitada pelo retângulo com borda preta, o gráfico ao lado direito se trata do aumento da referida área do gráfico. Os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias.



Fonte: Autora, 2017

Figura 3.25: Desvio Relativo tomando como referência o uso do funcional CAM-B3LYP. Assim como no gráfico do DA, área delimitada pelo retângulo está expandida ao lado direito. Os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias.



Fonte: Autora, 2017

3.7.3 Desvios Absoluto (DA) e Relativo (DR) em relação à inclusão das cargas pontuais

Os gráficos 3.26 e 3.27 mostram os DA e DR da energias calculadas com os métodos HF, DFT e DFT com funcional CAM-B3LYP, na situação em que a contribuição eletrostáticas das cargas pontuais foram inclusas, onde suas respectivas energias são utilizados como os valores referência.

A maioria dos resíduos possuem DA da energia, em relação ao uso das cargas pontuais, variando em torno de zero, porém três regiões de interesse podem ser destacadas. Os retângulos A, B e C delimitam estas regiões onde se encontram os resíduos com DA maior que 1 kcal/mol. O gráfico 3.26 mostra a expansão destas áreas (A, B e C), onde a maioria dos resíduos (num total

de 11) que são sensíveis à inclusão das cargas pontuais estão localizados na região A, isto é, nas regiões próximas do sítio ativo. Entretanto, a área C também se sobressai com os 9 aminoácidos cujos valores do DA ultrapassam 1 kcal/mol. Deste modo, tais aminoácidos demonstraram sensibilidade em relação à adição de cargas pontuais em cada método confrontado, ou seja, nenhum dos métodos (HF, DFT e DFT com funcional CAM-B3LYP) sobrelevou-se. Portanto, a inserção das cargas pontuais nos cálculos de energia conferiu importante variação no DA, independente dos métodos adotados. Interessante notar que, assim como no caso em que as interações de longo alcance adicionadas via funcional CAM-B3LYP, a inclusão das interações eletrostáticas através do background de cargas proporcionou uma variação significativa nas energias de interação dos resíduos carregados positivamente, como a LYS432, LYS389 E LYS402 (área B) calculados através do método DFT.

O DR (gráfico 3.27) indica que as energias calculadas via DFT sofreram os maiores desvios em relação à presença das cargas pontuais. No entanto, além de revelar desvios relativos significativos (até 18,60 para a energia do ALA406), o método DFT com a correção das interações de longo alcance (funcional CAM-B3LYP) apresentou a maior quantidade de resíduos cujas energias tiveram DR maior que 1. As energias calculadas através do método HF foram as que exibiram menores DRs em relação às obtidas via métodos com cargas pontuais.

Na tabela 3.17 são mostradas as médias dos desvios relativos da energia, em relação ao uso do background de cargas pontuais. O método DFT exibe um visível afastamento da energia quando confrontada com a metodologia em que as contribuições eletrostáticas estão presentes. Dentre os métodos aqui analisados, o HF apresentou o menor desvio relativo médio de 20%, demonstrando assim que este tipo de método possui menor percepção das correções eletrostáticas originadas das cargas pontuais.

Fonte: Autora, 2017

Tabela 3.17: A tabela mostra a média (em porcentagem) dos DRs em relação às energias obtidas na presença de cargas pontuais.

Média dos Desvios Relativos	
HF	21,80%
DFT	105,94%
DFT (CAM-B3LYP)	40,35%

Figura 3.26: Desvio Absoluto tomando como referência os métodos onde as cargas pontuais foram inclusas. Os gráficos A, B e C correspondem às expansões das regiões delimitadas pelo retângulos com borda preta A, B e C. Como os resíduos estão dispostos de acordo com a ordem crescente das mínimas distâncias, o gráfico A é compatível com regiões próximas do sítio ativo e concentra a maior parte dos resíduos com DA maior que 1 kcal/mol (indicado pela linha tracejada vermelha).

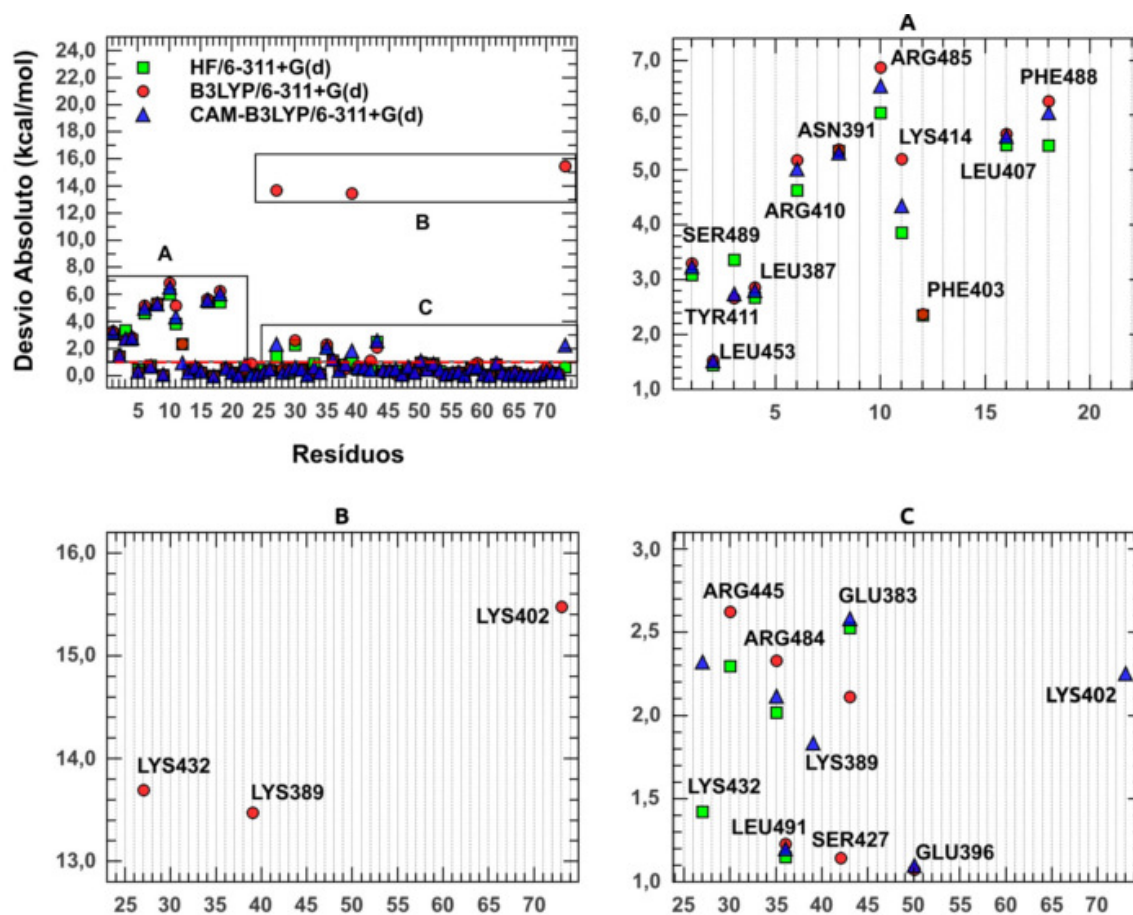
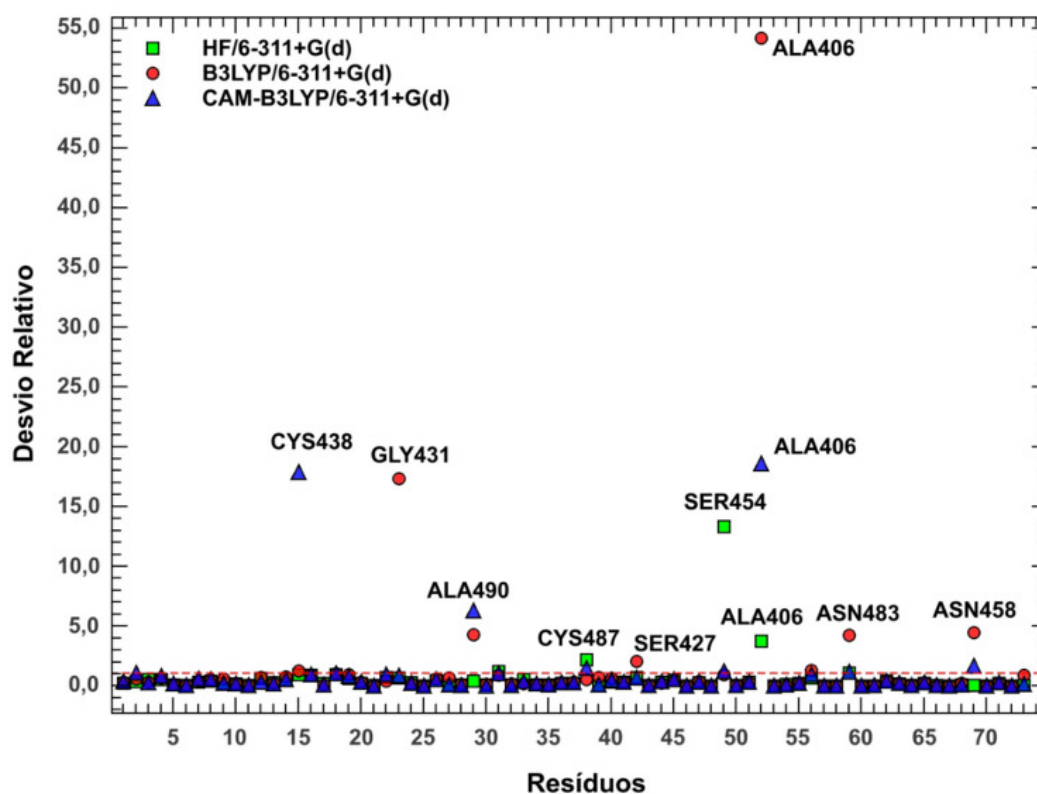


Figura 3.27: Desvio Relativo tomando como referência os métodos onde as cargas pontuais foram incluídas. As energias calculadas via DFT sofreram os maiores desvios em relação à presença das cargas pontuais.



Fonte: Autora, 2017

3.7.4 Desvios Absoluto (DA) e Relativo (DR) em relação ao método HF

Na presente análise, para o cálculo dos desvios, as energias obtidas pelo método HF foram usadas como valores de referência. Assim, os gráficos 3.28 e 3.29 mostram os desvios absoluto (DA) e relativo (DR), respectivamente.

Assim como no caso dos desvios em relação aos métodos em que as correções de longo alcance foram inseridas (por meio do funcional CAM-B3LYP), os desvios das energias em relação

ao método HF são mais evidentes para os resíduos do sítio ativo e suas proximidades. Da mesma maneira, conforme o gráfico 3.28, o método que culminou com a maior quantidade de resíduos com desvios significativos na energia foi o MP2 (17 resíduos). Todavia, evidencia-se que para os aminoácidos polares, como LYS414 (10,68 kcal/mol), LYS432 (12,16 kcal/mol), LYS389 (12,60 kcal/mol) e LYS402 (15,13 kcal/mol), as energias de interação calculadas pelo método DFT sem cargas pontuais sofrem os maiores DAs em relação àquelas calculadas via HF.

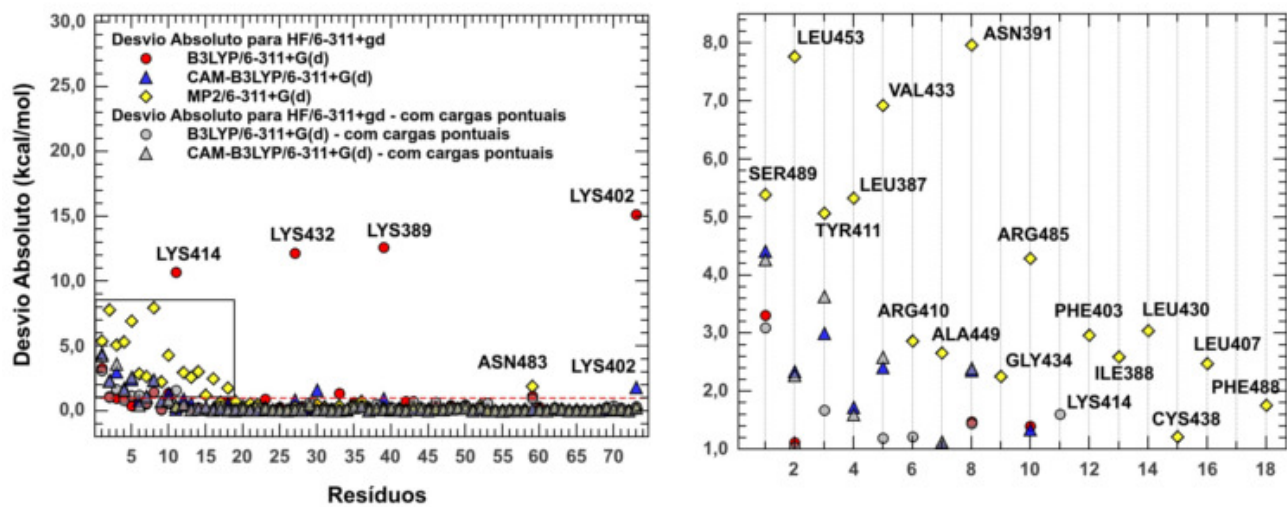
Com exceção dos aminoácidos distantes da região do sítio de ligação como GLY431, CYS487, SER427, SER454 e ASN483, a maioria dos desvios na energia foram constatados em resíduos localizados no sítio ativo da proteína. De modo que, As energias calculadas com o método MP2 são as que mais apresentaram DR (gráfico 3.29) superior a 1. As médias dos valores do DR estão expostas na tabela 3.18, segundo a qual os desvios em relação aos métodos onde foi inserida a contribuição eletrostática apresentaram menor dispersão, ou seja, apresentaram os menores valores da média. Por outro lado, as energias calculadas por meio de métodos na ausência de cargas pontuais mostraram desvios maiores, evidenciando maior convergência para métodos em que é inclusa a contribuição eletrostática fornecida por meio do background de cargas pontuais. Desta maneira, as energias de interação calculadas pelo método HF com cargas pontuais, convergem com aquelas obtidas por meio dos métodos DFT e DFT com funcional CAM-B3LYP.

Tabela 3.18: A tabela mostra as médias dos desvios relativos (DR) em relação ao método HF. As médias mais altas correspondem à comparação entre as energias obtidas pelos métodos HF e MP2. Outrossim, a média dos DRs são menores para as energias calculadas com métodos que incluem a contribuição eletrostática das cargas pontuais.

Média dos Desvios Relativos		
		Com Carga Pontuais
DFT	26,47%	1,13%
DFT (CAM-B3LYP)	9,34%	1,71%
MP2	76,91%	-

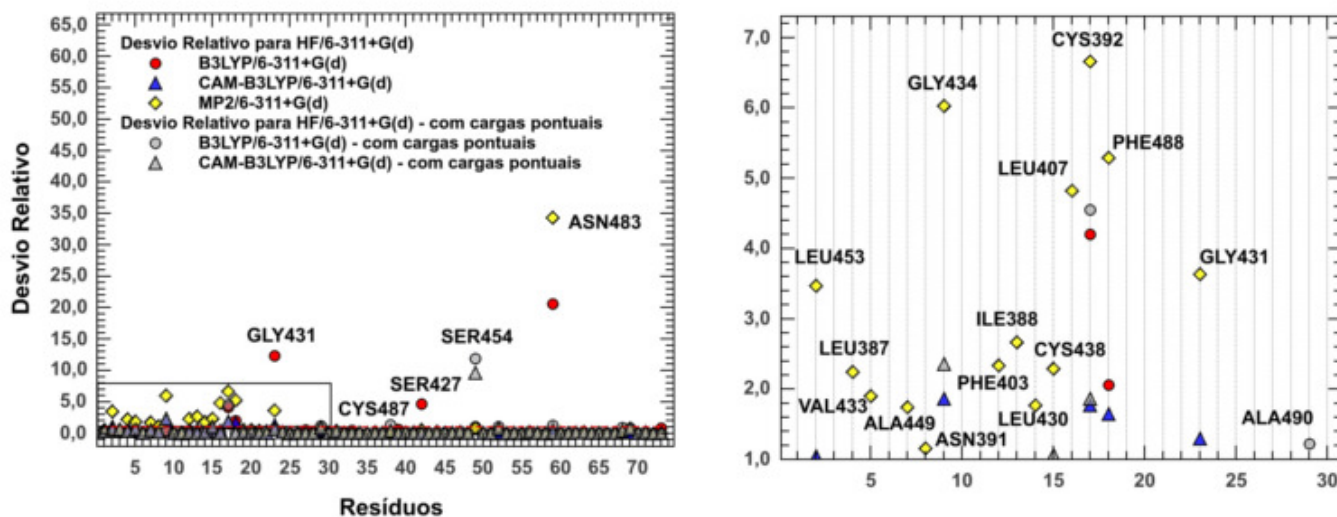
Fonte: Autora, 2017

Figura 3.28: Desvio Absoluto tomando como referência as energias calculadas pelo método HF. Visto que os resíduos estão dispostos em ordem crescente, nota-se que muitos resíduos com DA maior que 1 kcal/mol (indicado pela linha tracejada vermelha) se concentram nas regiões próximas do sítio ativo. A maioria destes desvios se referem às energias calculadas com MP2.



Fonte: Autora, 2017

Figura 3.29: Desvio Relativo tomando como referência as energias calculadas pelo método HF. Assim como em casos anteriores, as regiões próximas do sítio ativo concentram a maior quantidade de aminoácidos cujos desvios ultrapassam 1 (indicado pela linha tracejada vermelha). A maioria destes resíduos tiveram suas energias calculadas via MP2.



Fonte: Autora, 2017

3.7.5 Desvios Absoluto (DA) e Relativo (DR) em relação ao método MP2

Os desvios absoluto e relativo expostos nos gráficos 3.30 e 3.31 utilizam as energias calculadas via MP2 como valores de referência. Deste modo, prosseguimos avaliando estes desvios na energia de interação calculadas por meio dos diferentes métodos.

Conforme o gráfico 3.30, assim como em casos anteriores, os resíduos cujas as energias que mais apresentaram DA maior que 1 kcal/mol estão localizados no sítio ativo e sua vizinhança. Na área expandida ao lado direito no gráfico podemos visualizar melhor quais resíduos e para quais métodos os desvios absolutos maiores que 1 kcal/mol são detectados. Dessa forma, mesmo que

o DA em relação ao método MP2 acima de 1 kcal/mol seja constatado com bastante frequência para os métodos HF e DFT (CAM-B3LYP) (ambos com 17 resíduos), o método DFT lidera com 23 resíduos, para os quais o DA possui valor significativo. Visto que os valores do DA variam entre 1 e 15 kcal/mol, os resíduos para os quais o DA superou os 5 kcal/mol são LEU387 (5,33 kcal/mol), SER489 (5,39 kcal/mol), VAL433 (6,92 kcal/mol), LEU453 (7,77 kcal/mol), ASN391 (7,97 kcal/mol), todos com energias obtidas via HF. Outrossim, as energias dos resíduos LYS414 (10,85 kcal/mol), LYS432 (11,85 kcal/mol), LYS389 (12,36 kcal/mol), LYS402 (14,93 kcal/mol), as quais foram calculadas através do método DFT, apresentaram os maiores desvios absolutos. Portanto, o método DFT demonstrou maior percepção em relação às energias calculadas com o método perturbativo MP2.

Apesar dos resíduos com DR superior a 1 estarem dipostos nas proximidades da região ativa da proteína, o gráfico 3.31 mostra que os valores dos desvios em sua maioria não ultrapassa 3. Dentre os resíduos com os DRs mais expressivos, podemos citar: SER454 (4,44), ASN391 (7,25) e ALA490 (10,41), para o método HF; SER427 (9,36) para o DFT e ALA490 (16,77) para o DFT (com funcional CAM-B3LYP). No geral, nenhum método se sobressai em relação à intensidade dos desvios, nem em comparação com a quantidade de resíduos para os quais o DR seja extenso.

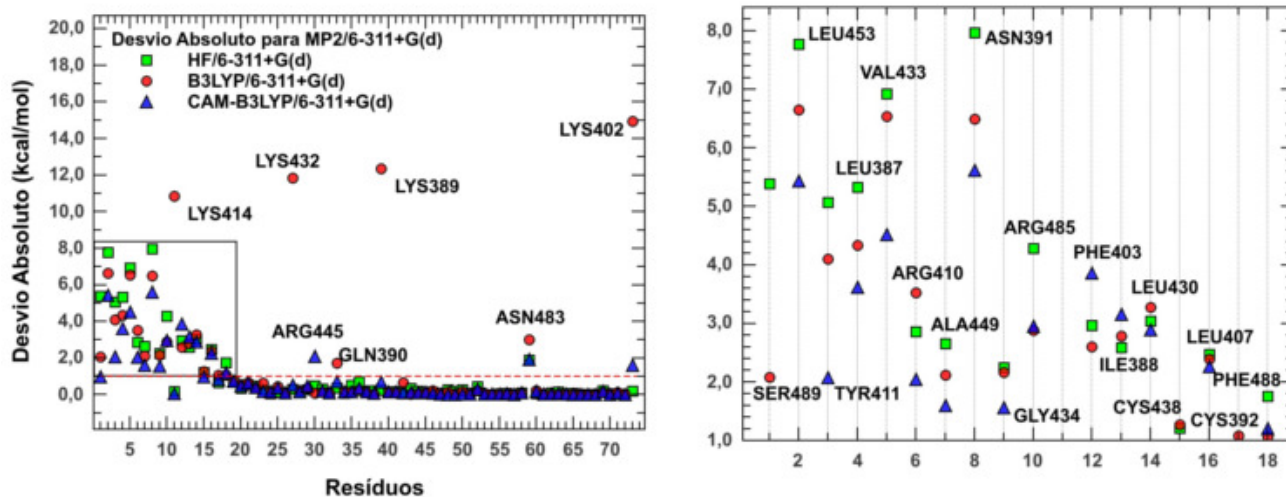
As informações da tabela 3.19 indicam o valor médio dos desvios relativos. Destarte, pode-se inferir que, a despeito de o método HF ser considerado o menos refinado dentre os demais, o mesmo demonstrou energias de interação com dispersão de 12,01% em relação às energias calculadas através do método perturbativo MP2. De outro modo, a média dos DRs das energias logradas com o método DFT (com funcional CAM-B3LYP) revelou maior convergência em relação aos valores calculados por meio do método MP2.

Tabela 3.19: A tabela mostra os valores das médias do DR das energias em relação o método MP2.

Média dos Desvios Relativos	
HF	12,01%
DFT	26,71%
DFT (CAM-B3LYP)	4,10%

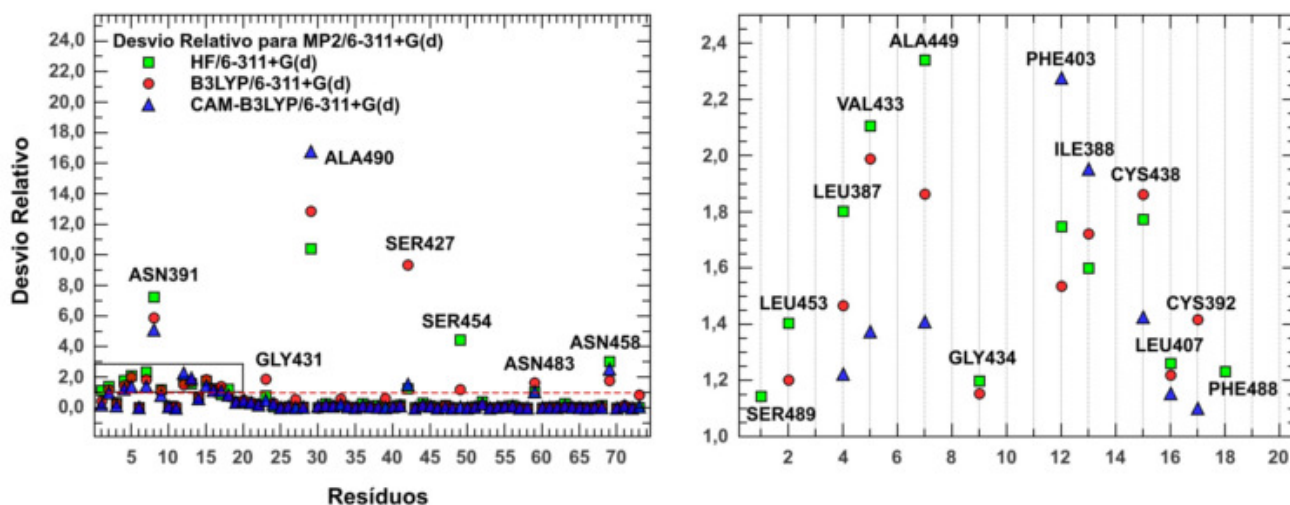
Fonte: Autora, 2017

Figura 3.30: Desvio Absoluto tomando como referência as energias calculadas pelo método MP2. Assim como em casos anteriores, as regiões próximas do sítio ativo concentram a maior quantidade de aminoácidos cujos desvios ultrapassam 1 kcal/mol (indicado pela linha tracejada vermelha). A maioria dos resíduos com DA significativa tiveram suas energias calculadas via DFT.



Fonte: Autora, 2017

Figura 3.31: Desvio Relativo tomando como referência as energias calculadas pelo método MP2. Assim como em casos anteriores, as regiões próximas do sítio ativo concentram a maior quantidade de aminoácidos cujos desvios relativos ultrapassam 1 (indicado pela linha tracejada vermelha). Os métodos são proporcionais nos quesitos intensidade dos DRs e quantidade de resíduos para os quais o DR seja extenso.



Fonte: Autora, 2017

3.8 As Curvas CBPR (Converged Binding Pocket Radius)

Originalmente, o sítio ativo do complexo IBU-HSA abrange 19 aminoácidos. No entanto, para uma avaliação mais abrangente da energia de interação entre fármaco e proteína, deve-se incluir resíduos distantes da molécula do ligante, isto é, os aminoácidos externos à camada mais próxima da região bioativa do complexo macromolecular [80, 82, 83]. Desta forma, expandimos a área de interesse e fizemos um estudo das energias de interação dos resíduos num raio de 13 Å, permitindo assim uma análise da convergência da energia total de interação entre fármaco e proteína. As curvas de convergência da energia total de interação com o raio do sítio ativo, também chamada

CBPR (Converged Binding Pocket Radius) tendem a se estabilizar conforme aumenta o raio das camadas do sítio de ligação, isto é, os aminoácidos mais distantes desta região passam a exprimir contribuições menos relevantes no valor da energia total de interação com o fármaco.

Os gráficos 3.32 e 3.33 mostram a convergência da energia total de interação em função da mínima distância entre os resíduos da proteína HSA e o ligante IBU. Todavia, até o raio (13 Å) em que foram selecionados os aminoácidos para o cálculo das energias de interação, o sistema IBU-HSA não apresentou uma tendência de estabilização da curva de convergência.

No gráfico 3.32, são mostradas as curvas CBPR para as energias obtidas pelos métodos quânticos na ausência do background eletrostático das cargas pontuais, ao passo que o gráfico 3.33 exhibe as curvas CBPR para as energias calculadas por meio de método na presença do background de cargas pontuais. Além do que, as curvas CBPR calculadas via métodos de solvatação contínua (PCM) com constantes dielétricas $\epsilon = 0$ e $\epsilon = 40$ [85], também são inclusas em ambos os gráficos.

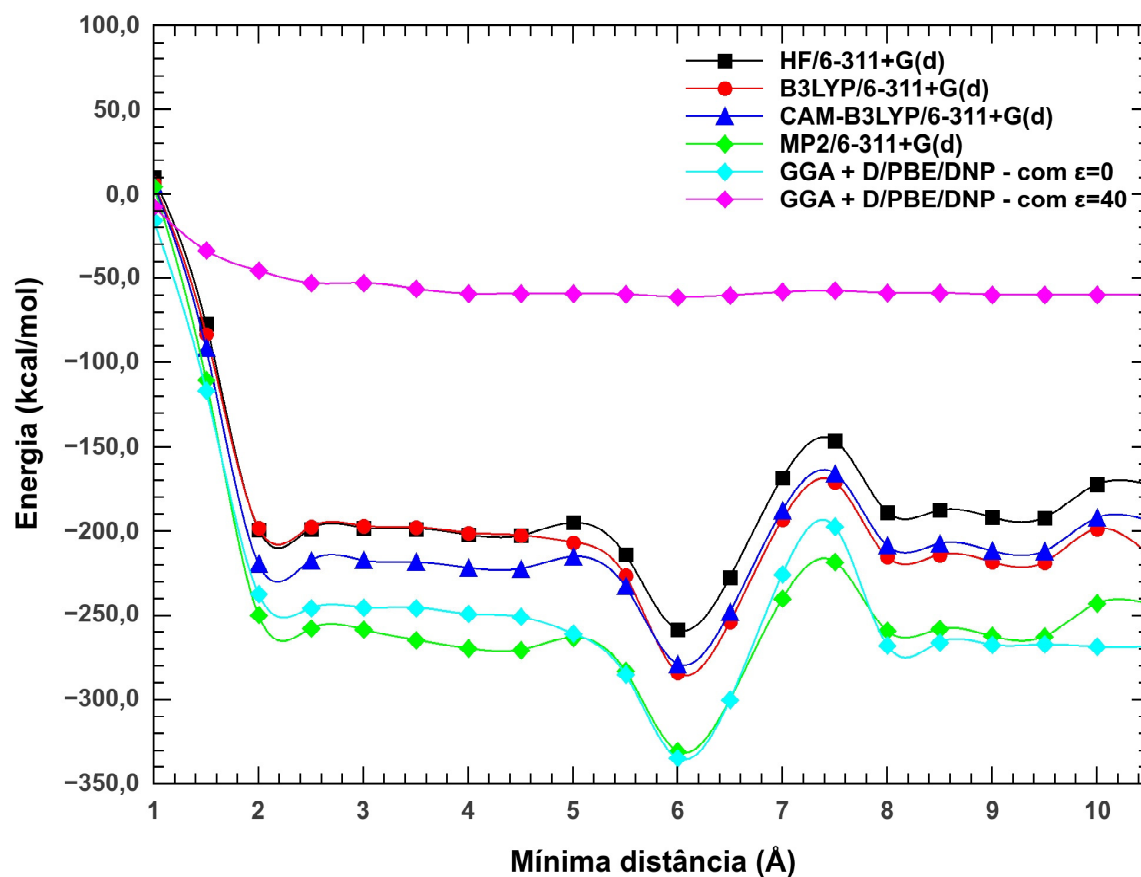
As curvas do gráfico 3.32 apresentaram comportamento análogo à metodologia com nível de cálculo GGA + D/PBE/DNP e constante dielétrica $\epsilon = 0$ (curva na cor ciano). Para o nível GGA + D/PBE/DNP com constante dielétrica $\epsilon = 40$ (curva na cor magenta) a energia total de interação converge com o raio em torno de 4 Å, enquanto que para $\epsilon = 0$ em torno de 9 Å. Para os demais métodos, a energia total não converge, mostrando assim que até o raio considerado neste trabalho as contribuições energéticas de outros aminoácidos são bastante significativas nas interações da proteína com o fármaco, ou ainda, os métodos aqui utilizados não reproduziram uma blindagem adequada das interações dos resíduos eletricamente carregados, cujas interações de longo alcance interferem fortemente com os resíduos do sítio ativo da proteína e do ligante IBU. Inclinações negativas nestas curvas (em 2, 6 e 8 Å) se devem à inclusão dos aminoácidos com cadeias laterais eletricamente positivas como ARG485, LYS414, ARG484, LYS389, ARG348 e ARG428. Por outro lado, as variações positivas (5, 7,5 e 10 Å) ocorrem devido à contribuição dos aminoácidos polares negativos como GLU450, GLU400 e GLU425. Por conseguinte, na região do sítio ativo (de 1 a 4,5 Å), os métodos HF e DFT reproduzem resultados equivalentes. As curvas das interações

totais calculadas via MP2 são muito próximas das obtidas no método PCM, porém no sítio de ligação o mesmo apresentou energias totais mais baixas que as obtidas via método PCM com $\varepsilon = 0$.

As curvas CBPR para os métodos nos quais as cargas pontuais foram inclusas (gráfico 3.33) mantêm a mesma forma das curvas que foram calculadas na ausência das cargas pontuais. Porém, as curvas para os métodos HF, DFT e CAM DFT sofreram um deslocamento em relação aos valores totais sem a contribuição eletrostática das cargas pontuais. Assim, na presença das cargas pontuais, os valores das energias aumentaram, expressando assim interações totais menos atrativas. Logo, para tais métodos, pode-se sugerir que a presença do background de cargas favoreceu uma blindagem mais expressiva das interações entre o fármaco e resíduos eletricamente carregados que se encontram em regiões distantes do sítio de ligação. Dentre tais métodos, deve-se sublinhar que as energias mais baixas foram calculadas via DFT com funcional CAM-B3LYP.

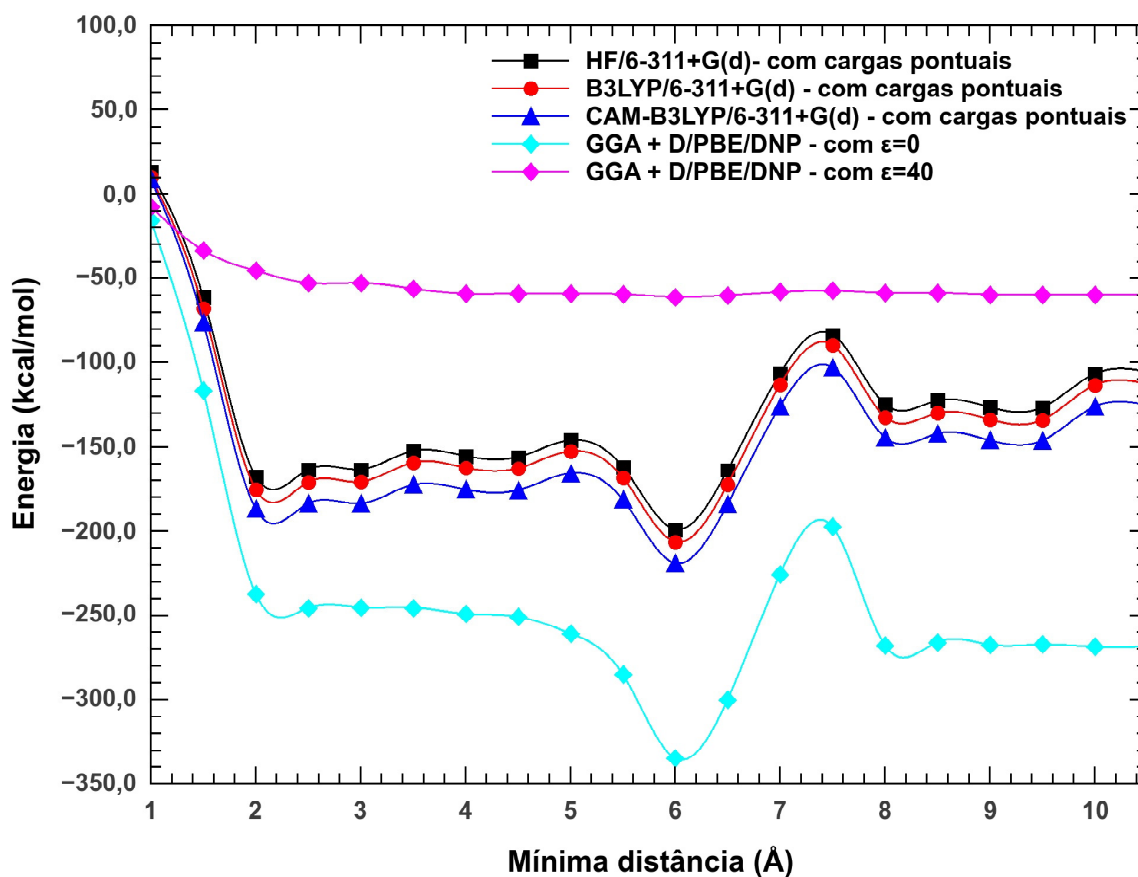
Em linhas gerais, percebe-se que os métodos que indicaram menores valores de energia foram aqueles em que não se incluiu a contribuição das interações eletrostáticas providas pelo background de cargas pontuais. Vale ressaltar que o método DFT (CAM-B3LYP) resultou nos menores valores da energia quando comparado com DFT e HF. Quanto aos métodos PCM, na descrição do meio com constante dielétrica nula, a heterogeneidade protéica não é evidenciada, de modo que há uma forte blindagem das interações de longo alcance dos resíduos da proteína. Já para o caso em que o meio é descrito por $\varepsilon = 40$, a forma da CBPR segue o mesmo padrão das curvas geradas pelos métodos HF, DFT e CAM-B3LYP, tanto na presença quanto na ausência do background de cargas.

Figura 3.32: Convergência da energia em função da distância radial do resíduo em torno do ligante. Cada curva, denominada CBPR, representa um método empregado no cálculo das energias de interação entre o IBU e a ASH. Aqui são comparadas as energias obtidas pelos métodos HF, DFT, DFT (com funcional CAM-B3LYP) e MP2, ao método DFT aliado à simulação em meio contínuo PCM, com constantes dielétricas $\epsilon = 0$ (curva na cor ciano) e $\epsilon = 40$ (curva no cor magenta) [85].



Fonte: Autora, 2017

Figura 3.33: Convergência da energia em função da distância radial do resíduo em torno do ligante. Cada curva, denominada CBPR, representa um método empregado no cálculo das energias de interação entre o IBU e a ASH. Aqui são comparadas as energias obtidas pelos métodos, na presença das cargas pontuais; HF, DFT e DFT (com funcional CAM-B3LYP), ao método DFT aliado à simulação em meio contínuo PCM, com constantes dielétricas $\epsilon = 0$ (curva na cor ciano) e $\epsilon = 40$ (curva no cor magenta) [85].



Fonte: Autora, 2017

Capítulo 4

Conclusão

Neste trabalho foram investigados os efeitos das interações de longo alcance nas energias de interação entre os resíduos do sítio ativo da proteína ASH e o ligante IBU. Para isto foram adicionadas nos cálculos de energia as contribuições das interações de longo alcance eletrostáticas, através do background de cargas pontuais e da contribuição Coulombiana inclusa no funcional CAM-B3LYP. Neste ensaio, utilizou-se uma metodologia onde foram utilizados cálculos de mecânica molecular (MM) e de mecânica quântica (MQ). Os cálculos de MM foram aplicados na otimização das posições atômicas dos átomos de hidrogênio no sistema cocrystalizado proteína-ligante, enquanto que os cálculos da energia de interação foram realizados por meio dos métodos quânticos HF, DFT e MP2. Para a realização dos cálculos individuais de energia, ou seja, da interação de cada resíduo de interesse com o fármaco, utilizamos o método de fragmentação molecular MFCC (Molecular Fractionation with Conjugate Caps).

Inicialmente as posições atômicas dos átomos pesados como Carbono, Oxigênio, Nitrogênio e Enxofre foram mantidos fixos, ao mesmo tempo que os átomos de Hidrogênio tiveram suas posições atômicas otimizadas através da aplicação do campo de força clássico CHARMM. Após a referida otimização, foram realizados os cálculos da energia de interação entre o ligante IBU e

os resíduos do sítio ativo e vizinhança, dentro de um raio de 13 Å traçado a partir do centróide da micromolécula.

Os cálculos de energia foram efetuados por meio dos métodos HF, DFT, DFT com funcional CAM-B3LYP e MP2. Com exceção do método MP2 que foi efetuado somente na ausência das cargas pontuais, todos os outros métodos foram implementados tanto na ausência quanto na presença do background de cargas pontuais. Além disso, para os diferentes níveis de cálculos via MQ, foram utilizados dois conjuntos de bases gaussianas, 6-31+G(d) e 6-311+G(d), com uma função polarizada e uma função difusa. A partir das energias calculadas foram traçados os perfis energéticos da interação do IBU com o sítio de ligação da ASH. Os perfis energéticos são esquematizados no painel gráfico BIRD (Binding site, Interaction energy an Residues Domain). No total, foram calculadas as energias de 73 aminoácidos, que incluem os resíduos do sítio ativo e a vizinhança dentro do raio estipulado.

Comparando os diferentes BIRDs foi possível inferir a influência das interações eletrostáticas proporcionadas pela presença das cargas pontuais remanescentes e as interações de longo alcance descritas pelo funcional CAM-B3LYP. Desta maneira, com base nas verificações dos resultados sistematizados nos BIRDs, ao considerar a inclusão das contribuições eletrostáticas no cálculos da energia de interação por meio do background das cargas pontuais, pode-se concluir que além de apresentarem significativa convergência em relação ao uso das bases 6-31+G(d) e 6-311+G(d), as energias de interação diminuem em módulo quando são consideradas as contribuições das cargas pontuais, mudando assim o perfil energético das interações entre o IBU e a ASH. Esta mudança é observada principalmente nos resíduos com cadeia lateral carregada, como as Lisinas, Argininas e Ácidos glutâmicos. Entretanto, resíduos neutros do sítio ativo também exibiram variações relevantes na energia, como o caso dos aminoácidos ASN391, SER489, TYR411, PHE403, PHE488 e LEU407.

Em relação aos Desvios Absolutos (DA) e Relativos (DR), a média dos DRs para o método DFT com correção de longo alcance (via funcional CAM-B3LYP) revelou maior precisão quando

comparado com o método MP2, exibindo o menor valor médio do DR, quando o método MP2 é utilizado como referência. Deste modo, embora o método perturbativo MP2 forneça um nível mais preciso, o método CAM-DFT apresentou resultados muito próximos a este e com um custo computacional substancialmente menor. De um modo geral, a inclusão da contribuição eletrostática gerada tanto pelo background de cargas pontuais, quanto pela correção coulombiana via funcional CAM-B3LYP, inibiu variações mais intensas nas energias de interação, de modo que os aminoácidos mais sensíveis aos efeitos eletrostáticos estão localizados no sítio ativo da HSA, com exceção dos aminoácidos LYS432, LYS389 e LYS402, os quais se encontram distantes do região bioativa da proteína.

As curvas CBPR evidenciaram os efeitos eletrostáticos providos pela inclusão das cargas pontuais, uma vez que para os métodos em que houve a inclusão do background eletrostático, a energia total diminuiu. Ademais, a convergência das energias em função da mínima distância calculadas por tais curvas, mostrou que na região do sítio ativo o método MP2 descreve melhor as energias de interação dos aminoácidos da IBU-HSA, de outra forma o método DFT (CAM-B3LYP) mostrou os menores valores de energia se comparado com método HF e DFT, na ausência e presença da contribuição eletrostática. Finalmente, a introdução de contribuições eletrostáticas fornece uma modesta blindagem para as interações do IBU com aminoácidos com cadeias laterais polares distantes do sítio ativo.

Em suma, o presente estudo demonstrou que efeitos de longo-alcance oriundos de interações eletrostáticas precisam ser incorporados de forma sistemática em cálculos baseados em MQ com o intuito de obter perfis mais precisos da interação entre fármacos e proteínas. Esperamos que a metodologia aqui introduzida venha a ser aplicada em outros sistemas, tornando-se padrão entre as técnicas teórico-computacionais voltadas ao estudo de sistemas macromoleculares.

Referências Bibliográficas

- 1 MORIN, A.; SLIZ, P. Structural biology computing: Lessons for the biomedical research sciences. *Biopolymers*, v. 99, n. 11, p. 809–816, 2013. ISSN 00063525.
- 2 National Institute of General Medical Sciences; NIH. The Structures of Life. n. July, p. 3–12, 2007.
- 3 BANASZAK, L. *Foundations of Structural Biology*. [S.l.]: Elsevier Science, 2000.
- 4 COUTINHO, K.; MORGON, N. *Métodos de Química Teórica E Modelagem Molecular*. [S.l.]: Editora Livraria da Física, 2007.
- 5 HAGE, D. S.; TWEED, S. A. Recent advances in chromatographic and electrophoretic methods for the study of drug-protein interactions. *Journal of Chromatography B*, v. 699, p. 499–525, 1997.
- 6 Krogsgaard-Larsen, Povl, T. L.; MADSEN, U. *Textbook of Drug Design and Discovery*. 3. ed. [S.l.: s.n.], 2002. 1–7 p.
- 7 NAYARISSERI, A. *Mechanistics in Drug Design - Experimental Molecular Biology vs. Molecular Modeling*. Nayarisseri2015: [s.n.], 2015. 3–4 p.
- 8 PATRICK, G. L. An Introduction to Medicinal Chemistry. *UCL lecture notes*, v. 40, p. 752, 2009. ISSN 00116513.

- 9 CARVALHO, I.; PUPO, M. T.; BORGES, Á. D. L. INTRODUÇÃO A MODELAGEM MOLECULAR DE FÁRMACOS NO CURSO EXPERIMENTAL DE QUÍMICA FARMACÊUTICA. *Quim. Nova.*, v. 26, n. 3, p. 428–438, 2003.
- 10 COOPER, R. E. H. G. M. *The Cell: A Molecular Approach, Fourth Edition*. 4. ed. [S.l.]: Sinauer Associates, Inc., 2007.
- 11 Nelson D.L., C. M. *Lehninger Principles of Biochemistry*. 4. ed. [S.l.: s.n.], 2004. 85–89 p. ISBN 1292021748.
- 12 KESSEL, A.; BEN-TAL, N. *Introduction to proteins*. [S.l.: s.n.], 2011. 1–12 p. ISBN 978-1-4398-1072-9.
- 13 TSAI C.J., X. D.; NUSSINOV, R. Structural motifs at protein–protein interfaces: protein cores versus two-state and three-state model complexes. *Protein Sci.*, v. 6, p. 1793–805, 1997.
- 14 SHEINERMAN F.B., N. R.; HONIG, B. Electrostatic aspects of protein–protein interactions. *Curr Opin Struct Biol*, v. 10, p. 153–9, 2000.
- 15 HOLTJE, H.-D.; FOLKERS, G. *Molecular Modeling - Basic Principles and Applications*. Weinheim ;New York ;Basel ;Cambridge ;Tokyo: Wiley - VCH, 1996. (Methods and principles in medicinal chemistry, v. 5).
- 16 WEDEMAYER G.J., P. P. W. L. S. P.; STEVENS, R. Structural insights into the evolution of an antibody combining site. *Science.*, v. 276, p. 1665–9, 1997.
- 17 PATTEN P.A., G. N. Y. P. M. C. W. G. B. J. S. R.; SCHULTZ, P. The immunological evolution of catalysis. *Science.*, v. 271, p. 1086–91, 1996.
- 18 SCHREIBER, G.; FERSHT, A. Energetics of protein–protein interactions: analysis of the barnase–barstar interface by single mutations and double mutant cycles. *J Mol Biol.*, v. 248, p. 478–86, 1995.

- 19 CLACKSON, T.; WELLS, J. A hot spot of binding energy in a hormone–receptor interface. *Science*, v. 267, p. 383–6, 1995.
- 20 KUMAR, S.; NUSSINOV, R. Close-range electrostatic interactions in proteins. *Chembiochem*, v. 3, n. 7, p. 604–617, July 2002.
- 21 BALDWIN, R. L. Energetics of Protein Folding. *Journal of Molecular Biology*, v. 371, n. 2, p. 283–301, 2007. ISSN 0022-2836. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022283607007371>>.
- 22 ROCHA, W. R. Interações intermoleculares. *Cadernos Temáticos de Química Nova na Escola*, p. 31–36, 2001.
- 23 LOMBARDINO, J. G.; III, J. A. L. A guide to drug discovery: The role of the medicinal chemist in drug discovery — then and now. *Nature Reviews Drug Discovery*, v. 3, p. 853–62, 2004.
- 24 SUN, E.; COHEN, F. E. A hot spot of binding energy in a hormone–receptor interface. *Gene*, v. 137, p. 127–32, 1993.
- 25 KUNTZ, I. D. Structure-Based Strategies for Drug Design and Discovery. *Science*, American Association for the Advancement of Science, v. 257, n. 5073, p. 1078–1082, 1992. ISSN 0036-8075. Disponível em: <<http://science.sciencemag.org/content/257/5073/1078>>.
- 26 Ou-Yang Si-sheng et al. Computational drug discovery. *Acta Pharmacologica Sinica*, CPS and SIMM, v. 33, p. 1131, aug 2012.
- 27 JORGENSEN, W. L. The many roles of computation in drug discovery. *SCIENCE*, v. 303, n. 5665, p. 813–1818, MARCH 2004.
- 28 Lounnas Valère et al. Current progress in Structure-Based Rational Drug Design marks a new mindset in drug discovery. *Computational and Structural Biotechnology Journal*, Research Network of Computational and Structural Biotechnology (RNCSB) Organization, v. 5, p.

- e201302011, feb 2013. ISSN 2001-0370. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3962124/>>.
- 29 MARTINS, J. P. A.; C., F. M. M. QSAR modeling: um novo pacote computacional open source para gerar e validar modelos QSAR. *Química Nova*, scielo, v. 36, p. 554–560, 00 2013. ISSN 0100-4042. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422013000400013&nrm=iso>.
- 30 NASCIMENTO, M. N. do. *Desenvolvimento de modelos de QSAR e planejamento de novos inibidores da enzima dUTPase de Plasmodium falciparum*. Dissertação (Mestrado) — Programa de Pós-Graduação em Ciências Farmacêuticas, Goiânia, 2015.
- 31 HANSCH, C.; FUJITA, T. p - σ - π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *Journal of the American Chemical Society*, v. 86, n. 8, p. 1616–1626, 1964. Disponível em: <<http://dx.doi.org/10.1021/ja01062a035>>.
- 32 HOPFINGER, A. J. et al. Construction of 3D-QSAR Models Using the 4D-QSAR Analysis Formalism. *Journal of the American Chemical Society*, v. 119, n. 43, p. 10509–10524, 1997. Disponível em: <<http://dx.doi.org/10.1021/ja9718937>>.
- 33 ANDRADE KERLY F. M. PASQUALOTO, E. I. F. C. H.; HOPFINGER, A. J. 4d-qsar: Perspectives in drug design. *Molecules*, v. 15, n. 5, p. 3281–3294, 2010.
- 34 CRAMER, R. D.; PATTERSON, D. E.; BUNCE, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society*, v. 110, n. 18, p. 5959–5967, 1988. PMID: 22148765. Disponível em: <<http://dx.doi.org/10.1021/ja00226a005>>.
- 35 Kitchen Douglas B. et al. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery*, Nature Publishing Group, v. 3, p. 935, nov 2004. Disponível em: <<https://www.nature.com/articles/nrd1549\supplementary-information>>.

- 36 SCOTT, W. R. P. et al. The GROMOS Biomolecular Simulation Program Package. *The Journal of Physical Chemistry A*, v. 103, n. 19, p. 3596–3607, 1999. Disponível em: <<http://dx.doi.org/10.1021/jp984217f>>.
- 37 WEINER, S. J. et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, v. 106, n. 3, p. 765–784, 1984. Disponível em: <<http://dx.doi.org/10.1021/ja00315a051>>.
- 38 BROOKS, B. R. et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, John Wiley & Sons, Inc., v. 4, n. 2, p. 187–217, 1983. ISSN 1096-987X. Disponível em: <<http://dx.doi.org/10.1002/jcc.540040211>>.
- 39 HALGREN, T. A. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *Journal of Computational Chemistry*, John Wiley Sons, Inc., v. 17, n. 5-6, p. 490–519, 1996. ISSN 1096-987X. Disponível em: <[http://dx.doi.org/10.1002-\(SICI\)1096-987X\(199604\)17:5/6<490::AID-JCC1>3.0.CO;2-P](http://dx.doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P)>.
- 40 Jones Gareth; Willett Peter; Glen Robert C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *Journal of Molecular Biology*, v. 245, n. 1, p. 43–53, 1995. ISSN 0022-2836. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022283695800379>>.
- 41 BöHM, H.-J. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3d database search programs. *Journal of Computer-Aided Molecular Design*, v. 12, n. 4, p. 309–323, July 1998.
- 42 RAREY, M. et al. A Fast Flexible Docking Method using an Incremental Construction Algorithm. *Journal of Molecular Biology*, v. 261, n. 3, p. 470–489, 1996. ISSN 0022-2836. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022283696904775>>.
- 43 GOHLKE, H.; HENDLICH, M.; KLEBE, G. Knowledge-based scoring function to predict protein-ligand interactions¹¹Edited by R. Huber. *Journal of Molecular Biology*, v. 295, n. 2, p.

337–356, 2000. ISSN 0022-2836. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022283699933715>>.

44 WU, G. et al. Detailed analysis of grid-based molecular docking: A case study of CDOCKER—A CHARMM-based MD docking algorithm. *Journal of Computational Chemistry*, Wiley Subscription Services, Inc., A Wiley Company, v. 24, n. 13, p. 1549–1562, 2003. ISSN 1096-987X. Disponível em: <<http://dx.doi.org/10.1002/jcc.10306>>.

45 LIU, M.; WANG, S. Mcdock: A monte carlo simulation approach to the molecular docking problem. *Journal of Computer-Aided Molecular Design*, v. 13, p. 435–451, 1999.

46 De Benedetti Pier G.; Fanelli Francesca. Computational quantum chemistry and adaptive ligand modeling in mechanistic QSAR. *Drug Discovery Today*, v. 15, n. 19, p. 859–866, 2010. ISSN 1359-6446. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1359644610002631>>.

47 GAO, J.; TRUHLAR, D. G. QUANTUM MECHANICAL METHODS FOR ENZYME KINETICS. *Annual Review of Physical Chemistry*, v. 53, n. 1, p. 467–505, 2002. PMID: 11972016. Disponível em: <<https://doi.org/10.1146/annurev.physchem.53.091301.150114>>.

48 FANELLI, F.; De Benedetti, P. G. Computational Modeling Approaches to StructureFunction Analysis of G Protein-Coupled Receptors. *Chemical Reviews*, v. 105, n. 9, p. 3297–3351, 2005. PMID: 16159154. Disponível em: <<http://dx.doi.org/10.1021/cr000095n>>.

49 KARELSON, M.; LOBANOV, V. S.; KATRITZKY, A. R. Quantum-Chemical Descriptors in QSAR/QSPR Studies. *Chemical Reviews*, v. 96, n. 3, p. 1027–1044, 1996. PMID: 11848779. Disponível em: <<http://dx.doi.org/10.1021/cr950202r>>.

50 RAHA, K. et al. The role of quantum mechanics in structure-based drug design. *Drug Discovery Today*, v. 12, n. 17, p. 725–731, 2007. ISSN 1359-6446. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1359644607002723>>.

- 51 Khandelwal Akash et al. A Combination of Docking, QM/MM Methods, and MD Simulation for Binding Affinity Estimation of Metalloprotein Ligands. *Journal of medicinal chemistry*, v. 48, n. 17, p. 5437–5447, aug 2005. ISSN 0022-2623 1520-4804. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2896055/>>.
- 52 MERZ, K. M.; BANCI, L. Binding of Bicarbonate to Human Carbonic Anhydrase II: A Continuum of Binding States. *Journal of the American Chemical Society*, v. 119, n. 5, p. 863–871, 1997. Disponível em: <<http://dx.doi.org/10.1021/ja963296a>>.
- 53 Klähn Marco et al. On Possible Pitfalls in ab initio QM/MM Minimization Approaches For Studies of Enzymatic Reactions. *The journal of physical chemistry. B*, v. 109, n. 32, p. 15645–15650, aug 2005.
- 54 DIXON, S. L.; JR., K. M. M. Semiempirical molecular orbital calculations with linear system size scaling. *The Journal of Chemical Physics*, v. 104, n. 17, p. 6643–6649, 1996. Disponível em: <<https://doi.org/10.1063/1.471382>>.
- 55 DIXON, S. L.; JR., K. M. M. Fast, accurate semiempirical molecular orbital calculations for macromolecules. *The Journal of Chemical Physics*, v. 107, n. 3, p. 879–893, 1997. Disponível em: <<https://doi.org/10.1063/1.474386>>.
- 56 GAO, A. M. et al. An efficient linear scaling method for ab initio calculation of electron density of proteins. *Chemical Physics Letters*, v. 394, n. 4, p. 293–297, 2004. ISSN 0009-2614. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S000926140400990X>>.
- 57 BENEDETTI, P. G. D.; FANELLI, F. Computational quantum chemistry and adaptive ligand modeling in mechanistic QSAR. *Drug Discovery Today*, v. 15, n. 19, p. 859–866, 2010.
- 58 SAPSE, A.-M. *Molecular Orbital Calculations for Amino Acids and Peptides*. Sapse1999: [s.n.], 1999. 8–10 p. ISBN 9781461271093.
- 59 ZHOU, T.; HUANG, D.; CAFLISCH, A. Quantum mechanical methods for drug design. *Current topics in medicinal chemistry*, v. 10, n. 1, p. 33–45, 2010. ISSN 15680266.

- 60 SAAD, Y. *Numerical Methods for Large Eigenvalue Problems: Revised Edition*. Society for Industrial and Applied Mathematics, 2011. (Classics in Applied Mathematics). ISBN 9781611970739. Disponível em: <<https://books.google.com.br/books?id=gViDLbUDjZ8C>>.
- 61 ATKINS, P. W.; FRIEDMAN, R. S. *Molecular Quantum Mechanics*. 4th. ed. [S.l.]: Oxford University Press Inc., 2005.
- 62 SLATER, J. C. The Theory of Complex Spectra. *Physical Review*, v. 34, p. 1293, 1929.
- 63 LEVINE, I. *Quantum Chemistry*. [S.l.]: Prentice-Hall Of India Pvt. Limited, 1999.
- 64 SZABO, A.; OSTLUND, N. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. [S.l.]: Dover Publications, 2012. (Dover Books on Chemistry).
- 65 VIANNA, J.; FAZZIO, A.; CANUTO, S. *Teoria quântica de moléculas e sólidos: simulação computacional*. [S.l.]: Livraria da Física, 2004.
- 66 CREMER, D. Møller–plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, John Wiley Sons, Inc., v. 1, n. 4, p. 509–530, 2011. ISSN 1759-0884. Disponível em: <<http://dx.doi.org/10.1002/wcms.58>>.
- 67 MØLLER, C.; PLESSET, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.*, American Physical Society, v. 46, p. 618–622, Oct 1934. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRev.46.618>>.
- 68 DREIZLER, R.; GROSS, E. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. [S.l.]: Springer Berlin Heidelberg, 2012.
- 69 HOHENBERG, P.; KOHN, W. Inhomogeneous electron gas. *Phys. Rev.*, American Physical Society, v. 136, p. B864–B871, Nov 1964. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRev.136.B864>>.

- 70 KOHN, W.; SHAM, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, American Physical Society, v. 140, p. A1133–A1138, Nov 1965. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRev.140.A1133>>.
- 71 PERDEW, J. P. et al. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B*, American Physical Society, v. 46, p. 6671–6687, Sep 1992. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevB.46.6671>>.
- 72 COHEN, A. J.; HANDY, N. C. Dynamic correlation. *Molecular Physics*, Taylor & Francis, v. 99, n. 7, p. 607–615, 2001. Disponível em: <<http://dx.doi.org/10.1080/00268970010023435>>.
- 73 YANAI, T.; TEW, D. P.; HANDY, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chemical Physics Letters*, v. 393, n. 1, p. 51–57, 2004. ISSN 0009-2614. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0009261404008620>>.
- 74 YANG, W. Direct calculation of electron density in density-functional theory. *Phys. Rev. Lett.*, v. 66, p. 1438–1441, 1991.
- 75 KOHN, W. Density functional and density matrix method scaling linearly with the number of atoms. *Phys. Rev. Lett.*, American Physical Society, v. 76, p. 3168–3171, Apr 1996. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.76.3168>>.
- 76 HE, X.; ZHANG, J. Z. H. The generalized molecular fractionation with conjugate caps/molecular mechanics method for direct calculation of protein energy. *The Journal of Chemical Physics*, v. 124, n. 18, p. 184703, 2006. Disponível em: <<https://doi.org/10.1063/1.2194535>>.
- 77 YASUDA, K.; YAMAKI, D. Simple minimum principle to derive a quantum-mechanical/molecular-mechanical method. *The Journal of Chemical Physics*, v. 121, n. 9, p. 3964–3972, 2004. Disponível em: <<https://doi.org/10.1063/1.1772354>>.

- 78 ZHANG, D. W.; ZHANG, J. Z. H. Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein–molecule interaction energy. *The Journal of Chemical Physics*.
- 79 ZHANG, D. W.; CHEN, X. H.; ZHANG, J. Z. H. Molecular caps for full quantum mechanical computation of peptide–water interaction energy. *Journal of Computational Chemistry*, v. 24, n. 15, 2003.
- 80 COSTA, R. F. da et al. Explaining statin inhibition effectiveness of HMG-CoA reductase by quantum biochemistry computations. *Phys. Chem. Chem. Phys.*, v. 14, p. 1389–1398, 2012.
- 81 RODRIGUES, C. et al. Quantum biochemistry study of the T3-785 tropocollagen triple-helical structure. *Chemical Physics Letters*, v. 559, p. 88–93, 2013. ISSN 0009-2614. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0009261413000134>.
- 82 RIBEIRO, T. C. da S. et al. The quantum biophysics of the isoniazid adduct NADH binding to its InhA reductase target. *New J. Chem.*, v. 38, p. 2946–2957, 2014.
- 83 ZANATTA, G. et al. Antipsychotic Haloperidol Binding to the Human Dopamine D3 Receptor: Beyond Docking Through QM/MM Refinement Toward the Design of Improved Schizophrenia Medicines. *ACS Chemical Neuroscience*, v. 5, n. 10, p. 1041–1054, 2014. PMID: 25181639. Disponível em: <http://dx.doi.org/10.1021/cn500111e>.
- 84 JIN, X.; ZHANG, J. Z. H.; HE, X. Full QM Calculation of RNA Energy Using Electrostatically Embedded Generalized Molecular Fractionation with Conjugate Caps Method. *The Journal of Physical Chemistry A*, v. 121, n. 12, p. 2503–2514, 2017. PMID: 28264557. Disponível em: <https://doi.org/10.1021/acs.jpca.7b00859>.
- 85 DANTAS, D. S. et al. Quantum molecular modelling of ibuprofen bound to human serum albumin. *RSC Adv.*, The Royal Society of Chemistry, v. 5, p. 49439–49450, 2015. Disponível em: <http://dx.doi.org/10.1039/C5RA04395F>.

- 86 WANG, X. et al. Electrostatically Embedded Generalized Molecular Fractionation with Conjugate Caps Method for Full Quantum Mechanical Calculation of Protein Energy. *The Journal of Physical Chemistry A*, v. 117, n. 32, p. 7149–7161, 2013.
- 87 TONG, Y. et al. Quantum Calculation of Protein Solvation and Protein-ligand binding free energy for HIV-1 Protease/Water complex. *Journal of Theoretical and Computational Chemistry*, v. 08, n. 06, p. 1265–1279, 2009.
- 88 MEI, Y.; JI, C.; ZHANG, J. Z. H. A new quantum method for electrostatic solvation energy of protein. *The Journal of Chemical Physics*, v. 125, n. 9, p. 094906, 2006. Disponível em: <<https://doi.org/10.1063/1.2345201>>.
- 89 JIA, X. et al. An improved fragment-based quantum mechanical method for calculation of electrostatic solvation energy of proteins. *The Journal of Chemical Physics*, v. 139, n. 21, p. 214104, 2013.
- 90 HE, X.; ZHANG, J. Z. H. The generalized molecular fractionation with conjugate caps/molecular mechanics method for direct calculation of protein energy. *The Journal of Chemical Physics*, v. 124, n. 18, p. 184703, 2006.
- 91 EXNER, T. E.; MEZEY, P. G. Ab Initio-Quality Electrostatic Potentials for Proteins: An Application of the ADMA Approach. *The Journal of Physical Chemistry A*, v. 106, n. 48, p. 11791–11800, 2002. Disponível em: <<http://dx.doi.org/10.1021/jp0263166>>.
- 92 LI, W.; LI, S.; JIANG, Y. Generalized Energy-Based Fragmentation Approach for Computing the Ground-State Energies and Properties of Large Molecules. *The Journal of Physical Chemistry A*, v. 111, n. 11, p. 2193–2199, 2007. PMID: 17388268. Disponível em: <<http://dx.doi.org/10.1021/jp067721q>>.
- 93 DAHLKE, E. E.; TRUHLAR, D. G. Electrostatically Embedded Many-Body Expansion for Large Systems, with Applications to Water Clusters. *Journal of Chemical Theory and*

- Computation*, v. 3, n. 1, p. 46–53, 2007. PMID: 26627150. Disponível em: <<http://dx.doi.org/10.1021/ct600253j>>.
- 94 JIANG, N.; MA, J.; JIANG, Y. Electrostatic field-adapted molecular fractionation with conjugated caps for energy calculations of charged biomolecules. *The Journal of Chemical Physics*, v. 124, n. 11, p. 114112, 2006.
- 95 ULRICH, K. H. Molecular aspects of ligand binding to serum albumin. *Pharmacol Rev*, v. 33, n. 1, p. 17–53, 1981.
- 96 SUGIO, S. et al. Crystal Structure of Human Serum Albumin at 2.5 Å. *Protein Eng*, v. 12, n. 6, p. 439–446, 1999.
- 97 SPECTOR, A. A. Fatty acid binding to plasma albumin. *Journal of Lipid Research*, v. 16, n. 2, p. 165–179, 1975.
- 98 PETERS, T. Serum Albumin. *Advances in protein chemistry*, v. 37, p. 161–245, 1985. ISSN 0065-3233.
- 99 FASANO, M. et al. The extraordinary ligand binding properties of human serum albumin. *IUBMB life*, v. 57, n. 12, p. 787–96, 2005. ISSN 1521-6543. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/16393781>>.
- 100 Klaus J. Fehske, Uwe Schalafr, Uwe Wollert; MULLER, W. E. Characterization of an Important Drug Binding Area on Human Serum Albumin Including the High-Affinity Binding Sites of Earfarin and Azapropazone. *Molecular Pharmacology*, v. 21, p. 387–393, 1982.
- 101 FREI, E. Albumin binding ligands and albumin conjugate uptake by cancer cells. *Diabetology {&} metabolic syndrome*, BioMed Central Ltd, v. 3, n. 1, p. 11, 2011. ISSN 1758-5996.

- 102 MERLOT, A. M.; KALINOWSKI, D. S.; RICHARDSON, D. R. Unraveling the mysteries of serum albumin-more than just a serum protein. *Frontiers in Physiology*, v. 5, n. August, p. 1–7, 2014. ISSN 1664042X.
- 103 LAW, S. W. et al. Nucleotide sequence and the encoded amino acids of human serum albumin mrna. *Proceedings of the National Academy of Sciences of the United States of America*, v. 79, p. 71–75, 1982.
- 104 JR., T. P. Intracellular PrecursorForms of PlasmaProteins: TheirFunctions andPossibleOccurrencein Plasma. *Clinical Chemistry*, v. 33, n. 8, p. 1317–1325, 1987.
- 105 PEDERSONT, D. M.; FOSTER, J. F. Subtilisin Cleavage of Bovine Plasma Albumin. Reversible Association of the Two Primary Fragments and Their Relation to the Structure of the Parent Protein“. *Biochemistry*, v. 8, n. 6, p. 2357–2365, 1969.
- 106 WHO. *WHO Model List of Essential Medicines - 19th List*.
<http://www.who.int/medicines/publications/essentialmedicines/en/>, April 2015.
Disponível em: <http://www.who.int/medicines/publications/essentialmedicines-/EML_2015_FINAL_amended_NOV2015.pdf?ua=1>.
- 107 HALFORD, G. M.; LORDKIPANIDZÉ, M.; WATSON, S. P. 50th anniversary of the discovery of ibuprofen: an interview with Dr Stewart Adams. *Platelets*, Taylor & Francis, v. 23, n. 6, p. 415–422, 2012. Disponível em: <<http://dx.doi.org/10.3109/09537104.2011.632032>>.
- 108 KD, R. Discovery, mechanisms of action and safety of ibuprofen. *Int J Clin Pract Suppl*, v. 135, p. 3–8, 2003.
- 109 ADAMS, S. S. The propionic acids: a personal perspective. *Journal of clinical pharmacology*, v. 32, n. 4, p. 317–23, 1992. ISSN 0091-2700.
- 110 GRACIN, S.; RASMUSON, Å. C. Solubility of phenylacetic acid, p-hydroxyphenylacetic acid, p-aminophenylacetic acid, p-hydroxybenzoic acid, and ibuprofen in pure solvents. *Journal of Chemical and Engineering Data*, v. 47, n. 6, p. 1379–1383, 2002. ISSN 00219568.

- 111 RICCIOTTI, E.; FITZGERALD, G. A. Prostaglandins and inflammation. *Arteriosclerosis, Thrombosis, and Vascular Biology*, v. 31, n. 5, p. 986–1000, 2011. ISSN 10795642.
- 112 BATLOUNI, M. Anti-Inflamatórios Não Esteroides: Efeitos Cardiovasculares, Cérebro-Vasculares e Renais. *Arquivos Brasileiros de Cardiologia*, v. 94, n. 4, p. 556–563, 2010. ISSN 0066-782X.
- 113 KAKUTA, H. et al. Cyclooxygenase-1-Selective Inhibitors Are Attractive Candidates for Analgesics That Do Not Cause Gastric Damage. Design and in Vitro/in Vivo Evaluation of a Benzamide-Type Cyclooxygenase-1 Selective Inhibitor. *Journal of Medicinal Chemistry*, v. 51, n. 8, p. 2400–2411, 2008.
- 114 SIMON, K. C.; CHEN, H.; SCHWARZSCHILD, M. diabetes , and risk of Parkinson disease. p. 863–869, 2011.
- 115 TOWNSEND, K. P.; PRATICÒ, D. Novel therapeutic opportunities for Alzheimer’s disease: focus on nonsteroidal anti-inflammatory drugs. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, v. 19, n. 12, p. 1592–1601, 2005. ISSN 1530-6860.
- 116 PEPE, V. L. E.; CASTRO, C. G. S. Osorio-de. *Prescrição de Medicamentos*. [s.n.], 2010. 79–103 p. ISSN 1098-6596. ISBN 9788533417366. Disponível em:
- 117 FRISCH, M. J. et al. *Gaussian09 Revision E.01*. Gaussian Inc. Wallingford CT 2009.
- 118 SIMMONETT, A. C.; ANDREW, T. B. G.; GILL, P. M. W. An optimal point-charge model for molecular electrostatic potentials. *Molecular Physics*, Taylor & Francis, v. 103, n. 20, p. 2789–2793, 2005. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/00268970500187910>>.
- 119 SMITH, C. M.; HALL, G. G. The approximation of electron densities. *Theoretica chimica acta*, v. 69, n. 1, p. 63–69, Jan 1986. ISSN 1432-2234. Disponível em: <<https://doi.org/10.1007/BF00526293>>.

120 HALL, G. G.; SMITH, C. M. Fitting electron densities of molecules. *International Journal of Quantum Chemistry*, John Wiley & Sons, Inc., v. 25, n. 5, p. 881–890, 1984. ISSN 1097-461X.
Disponível em: <<http://dx.doi.org/10.1002/qua.560250510>>.