UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

ADRIANO MORAES AMARANTE

Estudo do recobrimento biológico de nanossuperfícies por modelagem
computacional: aplicação no desenvolvimento de nanoimunossensores

São Carlos
2019

ADRIANO MORAES AMARANTE

Estudo do recobrimento biológico de nanossuperfícies por modelagem computacional: aplicação no desenvolvimento de nanoimunossensores

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para realização da Defesa, como parte dos requisitos para obtenção do título de Doutor em Ciências.

Área de concentração: Física Aplicada
Opção: Física Biomolecular
Orientador: Prof. Dr. Fábio de Lima Leite

Versão Corrigida
(versão original disponível na Unidade que aloja o Programa)

São Carlos
2019

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Amarante, Adriano Moraes

Estudo do recobrimento biológico de nanossuperfícies por modelagem computacional: aplicação no desenvolvimento de nanoimunossensores / Adriano Moraes Amarante; orientador Fabio de Lima Leite - versão corrigida -- São Carlos, 2019.

172 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2019.

1. Modelagem molecular computacional. 2. Atomic force microscope. 3. Dinâmica molecular. 4. Dinâmica molecular direcionada. 5. Nanoimunossensor. I. Leite, Fabio de Lima, orient. II. Título.

Aos meus pais Herculano e Sonia.

AGRADECIMENTOS

Agradeço aos meus pais, Herculano e Sonia, e aos meus irmãos Juliano e Natália pelo apoio incondicional e presença na minha vida a todo o momento.

Agradeço aos meus mentores e amigos Fábio, Eduardo, Gomide e Osvaldo por toda confiança e ensinamentos de vida e profissionais.

Agradeço aos meus amigos e companheiros de grupo Jéssica, Pâmela, Ariana, Guilherme e Guedmiller por todo trabalho e conhecimento compartilhado.

Agradeço a todo grupo de NanoNeuroBioFísica por todo apoio e tutoria.

Agradeço aos meus professores do IFSC/USP e da UFSCar-SOR pelos conhecimentos transmitidos, apoio e confiança depositada.

Agradeço aos meus amigos e companheiros de estudos Jorge e Tiago por todo trabalho e conhecimento compartilhado.

Agradeço a todos os meus amigos que estiveram comigo nesse período tão importante da minha vida.

Agradeço especialmente à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela confiança depositada e o fomento da pesquisa desenvolvida sob o número de processo FAPESP nº 2014/12466-7.

Agradeço, enfim, a todos aqueles que sempre estiveram por perto e que me ajudaram a chegar aonde cheguei.

RESUMO

AMARANTE, A. M. **Estudo do recobrimento biológico de nanossuperfícies por modelagem computacional:** aplicação no desenvolvimento de nanoimunossensores. 2019. 172 p. Tese (Doutorado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2019.

Neste trabalho foram utilizadas técnicas de modelagem molecular computacional para descrever nanossuperfícies funcionalizadas com biomoléculas do sistema imunológico correlacionando resultados experimentais obtidos com o microscópio de força atômica, simulações de dinâmica molecular e dinâmica molecular direcionada. O objetivo principal proposto é avaliar as forças intermoleculares provenientes das interações antígeno-anticorpo (funcionalizados em nanossuperfícies) para aplicação no desenvolvimento de nanoimunossensores e detecção de doenças desmielinizantes, como a Neuromielite Óptica. A Neuromielite Óptica é uma doença inflamatória autoimune na qual o próprio sistema imunológico reage contra os nervos ópticos e a medula espinhal, causando lesão desmielinizante. Estudos na literatura estabeleceram anticorpo anti-aquaporina4 como um importante biomarcador da doença. Neste contexto, um nanoimunossensor vem sendo desenvolvido com a técnica de Microscopia de Força Atômica, o qual visa detectar o anticorpo anti-aquaporina4 no soro de portadores da doença. Tal estudo necessitou de uma nova abordagem computacional para a descrição de estruturas tridimensionais de anticorpos. Essa nova aproximação consistiu na aplicação de técnicas de computacionais de modelagem e engenharia molecular para a geração de modelos de anticorpos com base em sucessivas substituições dos resíduos componentes do sítio de interação com o antígeno. Testes realizados envolvendo modelos de anticorpos disponíveis em bancos de dados especializados indicaram $(48 \pm 18) \%$ e $(65 \pm 14) \%$ de identidade das cadeias leve e pesada, respectivamente, entre os modelos gerados computacionalmente e as estruturas 3D reais de anticorpos. Por fim, para comprovar o funcionamento dos nanoimunossensores, foi desenvolvido um modelo estatístico para tratar e interpretar os dados experimentais. Este modelo foi eficiente para distinguir os pacientes soropositivos de sujeitos soronegativos para determinados biomarcadores relacionados à Neuromielite Óptica e a Esclerose Múltipla, fornecendo assim um novo e mais preciso processo para diagnóstico de doenças desmielinizantes.

Palavras-chave: Modelagem molecular computacional. *Atomic force microscope*. Dinâmica molecular. Dinâmica molecular direcionada. Nanoimunossensor.

ABSTRACT

AMARANTE, A. M. **Study of the biological coverage of nanosurfaces by computational modeling:** application in the development of nanoimunosensores. 2019. 172 p. Tese (Doutorado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2019.

Study of the biological coverage of nanosurfaces by computational modeling: application in the development of nanoimunosensores. In this work, computational molecular modeling techniques were applied to describe nanosurfaces functionalized with immune system biomolecules, correlating data from atomic force microscope experiments, molecular dynamics, and steered molecular dynamics simulations. The main goal of this research was to evaluate intermolecular forces involved in the antigen-antibody interaction on the nanosurfaces during the development of nanoimunosensores for demyelinating diseases detection, especially neuromyelitisoptica. The neuromyelitisoptica is an autoimmune inflammation in which components of the immune system respond against optical nerves and spinal cord, resulting in demyelinating lesions. In the literature, studies have established anti-aquaporin 4 as an important biomarker for neuromyelitisoptica. Then, a nanoimunosenzor for anti-aquaporin 4 antibodies detection in neuromyelitisoptica patients' serum via Atomic Force Microscopy is in development. This study requested a computational approach for describing the tridimensional structure of antibodies. The novel approach consisted of computer molecular modeling and engineering to perform successive substitutions in residues of the antigen interaction site. Tests carried out using antibody structures available in specialized data banks demonstrated the similarity of $(48 \pm 18) \%$ and $(65 \pm 14) \%$ for light and heavy chains, respectively, of the computationally generated models and experimental 3D structures of antibodies. Additionally, a statistical model was developed to prove the nanoimunosenzor sensing activity, which was useful to treat and interpret the experimental data. This statistical model was efficient to distinguish seropositive patients from seronegative subjects considering specific biomarkers related to neuromyelitisoptica and multiple sclerosis, providing a novel and more precise process for demyelinating disease diagnosis.

Keywords: Computational molecular modeling. Atomic force microscope. Molecular dynamics. Steered molecular dynamics. Nanoimunosenzor.

LISTA DE FIGURAS

Figura 1 – Diagrama esquemático do sistema de Microscopia de Força Atômica	28
Figura 2 – (A) Curva de força típica e (B) Diagrama esquemático do movimento vertical da ponta do AFM durante a aproximação e retração de uma curva de força	30
Figura 3 – Esquema representativo do mecanismo de ação da neuromielite óptica	34
Figura 4 – Estados de compressão, equilíbrio e alongamento para moléculas ligadas	37
Figura 5 – RMSD de duas diferentes estruturas, para uma DM de 10000 ps	38
Figura 6 – (A) Representação esquemática sem escala do <i>cantilever</i> de AFM exercendo uma força em uma molécula alvo. (B) Esquema do sistema para investigações das interações enzima-inibidores e antígeno-anticorpos	39
Figura 7 – Remoção de uma molécula ligante a partir da cavidade de ligação de uma proteína	40
Figura 8 – Ciclo de energia livre de ligação	42
Figura 9 – Descrição do processo de funcionalização da superfície de Si_3N_4 da ponta do AFM. (A) Molécula de APTES. (B) Superfície de Si_3N_4 funcionalizada com APTES	48
Figura 10 – (A) Estrutura inicial da porção Fab da IgG, destacando as alças mutadas. (B) Complexo inicial com epítipo alvo nposicionado por cálculos de <i>docking</i> molecular	51
Figura 11 – Modelagem molecular da proteína AQP4, destacando um possível epítipo-alvo do anticorpo anti-AQP4	53
Figura 12 – Curvas características da distribuição gama, dependentes dos parâmetros α e β	56
Figura 13 – Ponta de AFM funcionalizada com APTES e peptídeos da AQP4	57
Figura 14 – MMC da IgG tipo 1. (A) Interação entre monômeros de proteína A e IgG segundo dados cristalográficos. (B) Interação entre moléculas de proteína A íntegras e IgG segundo MMC	59
Figura 15 – Nanossuperfície de mica funcionalizada com IgGs tipo 1	60
Figura 16 – Homogeneidade da imobilização orientada e não orientada	61
Figura 17 – Estrutura dos peptídeos AKFRHD, HQLDPAFG e EPANPSEKNSPSTQYC sintetizados <i>in silico</i>	63

Figura 18 – Estrutura prevista do complexo antígeno anticorpo com epítomos AKFRHD.	65
Figura 19 – Estrutura prevista do complexo antígeno anticorpo com epítomos HQLDPAFG.	68
Figura 20 – Estrutura prevista do complexo antígeno anticorpo com epítomos EPANPSEKNSPSTQYC.	68
Figura 21 – MMC dos epítomos da proteína AQP4. Peptídeo linear sintetizado <i>in silico</i>	69
Figura 22 – (A) Estrutura prevista do complexo antígeno anticorpo anti-AQP4. (B) Potencial eletrostático do epítopo GTEKPLPVDM e do anticorpo anti-AQP4. (C) RMSD do complexo formado.	71
Figura 23 - Modelo esquemático da aplicação de uma força externa para remoção do peptídeo imunogênico do complexo da porção Fv do anticorpo específico por SMD mutada.	72
Figura 24 – Curva de força de SMD obtida para a interação peptídeo imunogênico da AQP4 e anticorpo modelado, com força de desprendimento foi 224 pN.	73
Figura 25 – Moda dos valores de força de adesão para o soro purificado de pacientes soro positivos para anti-AQP4	74

LISTA DE TABELAS

Tabela 1 – Etapas de desenvolvimento do protocolo de engenharia molecular	55
---	----

LISTA DE SIGLAS E ABREVIATURAS

AFM	Microscópio de Força Atômica
anti-AQP4	Anti-aquaporina 4
AQP4	Aquaporina4
DM	Dinâmica Molecular
EM	Esclerose Múltipla
Fab	Do inglês: <i>Fragment with specific antigen binding</i>
GNN	Grupo de Pesquisa em Nanoneurobiofísica
IgG	Imunoglobulina G
LCR	líquido cefalorraquidiano
MMC	Modelagem molecular computacional
MOG	Do inglês <i>myelinoligodendrocyteglycoprotein</i>
NMO	Neuromielite óptica
RMN	Ressonância magnética nuclear
RMSD	Do inglês: <i>Root Mean Square Deviation</i>
SMD	Do inglês: Steered Molecular Dynamics
SNC	Sistema nervoso central
PDB	Do inglês: Protein Data Bank

CÓDIGO DOS AMINOÁCIDOS

Nome	Símbolo	Abreviação
Ácido aspártico	ASP	D
Ácido glutâmico	GLU	E
Alanina	ALA	A
Arginina	ARG	R
Asparagina	ASN	N
Cisteína	CYS	C
Fenilalanina	PHE	F
Glicina	GLY	G
Glutamina	GLN	Q
Histidina	HIS	H
Isoleucina	ILE	I
Lisina	LYS	K
Leucina	LEU	L
Metionina	MET	M
Prolina	PRO	P
Tirosina	TYR	Y
Treonina	THR	T
Triptofano	TRP	W
Serina	SER	S
Valina	VAL	V

SUMÁRIO

1 INTRODUÇÃO	23
2 FUNDAMENTAÇÃO TEÓRICA.....	27
3 OBJETIVOS	45
4 METODOLOGIA.....	47
5 RESULTADOS E DISCUSSÃO	57
6 CONCLUSÕES E PERSPECTIVAS	77
7 PRODUÇÕES DECORRENTES DESTE TRABALHO.....	79
REFERÊNCIAS	81
APÊNDICE A	93
APÊNDICE B.....	105
ANEXO A.....	109
ANEXO B	115
ANEXO C.....	123

1 INTRODUÇÃO

O desenvolvimento teórico de nanobiossensores baseados em pontas do microscópio de força atômica (AFM, do inglês *Atomic Force Microscope*) foi abordado neste trabalho pelo uso de técnicas de modelagem molecular computacional (MMC). Nanobiossensores são sensores biológicos com a sensibilidade e a seletividade na nanoescala (compreendidos na escala de 1-100 nm).(1) O nosso grupo de pesquisa (Grupo de Pesquisa em Nanoneurobiofísica, GNN) desenvolve diversas linhas de pesquisas que estudam interações proteína-ligantes e o desenvolvimento de “superfícies inteligentes” para a construção de nanobiossensores(2–6), baseados na detecção de forças moleculares por pontas AFM funcionalizadas.(7-8) Os nanobiossensores começaram a ser desenvolvidos para a detecção do defensor químico Diclofop.(9) A modelagem do nanobiossensor foi realizada de maneira que a ponta do AFM, elemento sensor, foi funcionalizada com enzimas Acetil-coA Carboxilase (ACC) (10) para a detecção do substrato funcionalizado com o Diclofop. Tal força foi calculada teoricamente por MMC e comparada a força de adesão medida experimentalmente, respondendo a questões, envolvendo nanobiossensores de ponta de AFM, não respondidas até o momento, como: (i) número de enzimas úteis na ponta do AFM que podem interagir com o substrato; (ii) o número de sítios ativos disponíveis; (iii) consequências da imobilização das enzimas e do substrato e; (iv) força de adesão teórica entre a ponta do AFM e o substrato.(11) Os modelos desenvolvidos nesta etapa foram cruciais para o desenvolvimento dos modelos de funcionalização propostos e que resultaram na publicação dos artigos intitulados: *A new rigidez model and its application in nanobiosensor development* e *A Computational Protein Structure Refinement of the Yeast Acetohydroxyacid Synthase*, publicados nos periódicos *Journal of Molecular Graphics & Modelling* e *Journal of the Brazilian Chemical Society*, respectivamente, disponíveis nos **ANEXOS A e B**.

Os nanobiossensores empregados em nosso grupo de pesquisa, originalmente na detecção de contaminantes ambientais, deram origem ao estudo de doenças desmielinizantes¹ e (2–5,11–16), como a Neuromielite Óptica (NMO). A NMO é uma doença inflamatória autoimune na qual o próprio sistema imunológico reage contra os nervos ópticos e a medula espinhal do paciente. (17–18) Dados da literatura apontam para um papel central de anticorpos anti-aquaporina4 (anti-AQP4) no diagnóstico de tal patologia. (18–19) O anti-AQP4 reconhece uma proteína dos canais de água nas terminações dos axônios dos astrócitos, a aquaporina4 (AQP4) e, em conjunto com as moléculas do sistema complemento, causa lesão desmielinizante e axonal, necrose, espessamento vascular. (20) Para o diagnóstico de doenças autoimunes desmielinizantes, *e.g.* a NMO e a Esclerose Múltipla (EM), as ferramentas mais utilizadas são a neuroimagem (21) e a análise do líquido cefalorraquidiano (LCR). (21) Em casos clássicos da NMO e EM, ambas as doenças têm distinção devido à diferença de sintomas, como neurite óptica aguda e mielite transversa que atingem pacientes portadores de NMO, mas raramente atingem os pacientes de EM, nos estágios iniciais da doença. (20) No entanto, quando se trata do progresso das doenças, há uma linha tênue para a distinção entre NMO e EM, uma vez que a síndrome de NMO pode ocorrer em ambas, ao ponto de a NMO ter sido considerada uma variante da EM por décadas. (22) Atualmente, para diagnóstico da NMO, além do quadro clínico, são realizados testes laboratoriais para análise da presença do anticorpo anti aquaporina-4 (anti-AQP4) no soro dos pacientes. (23)

A obtenção de modelos 3D de estruturas de anticorpos, como as estruturas cristalográficas do banco de dados de proteínas (24–27), é fundamental para os cálculos de MMC, Docking Molecular, Dinâmica Molecular (DM) e a dinâmica molecular direcionada

¹Projetos aprovados:

Desenvolvimento de Nanobiossensores para Estudo de Doenças Neurodegenerativas e Contaminantes Ambientais (Processo: 302666/2012-5, Financiamento: CNPq) (2013-2016)

Estudo de Doenças Neurodegenerativas usando Técnicas de Nanobiotecnologia: Mecanismos de Ação e Interpretação Neuropatológica (Processo: 23038006985201116, Financiamento: CAPES) (2013-2014)

Desenvolvimento de Nanobiossensores para Estudo de Doenças Neurodegenerativas (Processo: 471632/2012-0 (Edital Universal) (2013)

Desenvolvimento de Cantiléveres Inteligentes para Estudo de Doenças Neurodegenerativas (Processo: 483303/2011-9, Financiamento: CNPq) (2013)

(SMD, do inglês *Steered Molecular Dynamics*). (2-5,11-16) Normalmente, a estrutura é obtida por modelos de cristalografia de raios-X e de ressonância magnética nuclear (RMN) (28-29) Porém, apesar dos métodos experimentais em questão serem capazes de gerar um modelo tridimensional detalhado, preciso e total de biomoléculas, uma série de empecilhos pode dificultar e até inviabilizar a obtenção da estrutura tridimensional de interesse, sendo estes com relação ao tempo de análise (30-31), obtenção e purificação da macromolécula em estudo, cristalização e obtenção de cristais de qualidade para análise por cristalografia de raios-X, tamanho da molécula ou complexo molecular no caso de análise por RMN (32) e descrição das regiões de alças protéicas. (33-34) Dessa forma, ferramentas computacionais surgem como alternativas úteis e rápidas para descrição estrutural de biomoléculas, a considerar a grande evolução dos computadores ao longo das últimas décadas. (35) Desta forma, optou-se pela aplicação de técnicas de MMC para a obtenção da estrutura dos anticorpos, especialmente pela aplicação de métodos de engenharia molecular computacional. Neste protocolo, a modelagem da região fragmento ligador de antígeno (Fab, do inglês *Fragment with specific antigen binding*) da Imunoglobulina G (IgG) constitui a mais difícil de se prever, já que variam muito dentre os anticorpos e são específicas para prover a ligação com o antígeno. (30-31)

Neste contexto, uma das propostas principais deste trabalho envolve a mutação e modelagem das seis alças de interação da IgG com o antígeno específico. Para isso foi desenvolvido uma metodologia que engloba técnicas de engenharia molecular computacional para a geração de modelos para anticorpos. Na literatura, autores como Loyauet al. 2015 (32), Choi e Deane 2011 (36), Kuroda et al. 2012 (30), Marks e Deane 2017 (37), Sevy e Meiler 2014 (31), e Zhao 2018 (38), realizaram cálculos de engenharia molecular na estrutura de anticorpos já descritos, variando aminoácidos em posições específicas, principalmente quando se tratando da alça H3, das alças de interação com o antígenos, com a finalidade de obtenção

de uma nova estrutura de um anticorpo de interesse. Porém nenhum desses estudos fazem a previsão da estrutura de anticorpos partindo apenas da estrutura de epítomos de interesse, nem mesmo programas e servidores online como PredictionofImmunoGlobulinStructure(39-40), Rosetta(41), Biovia Discovery Studio (42) e StructuralAntibodyPrediction Server (SAbPred).(43) Dessa forma, a metodologia desenvolvida neste trabalho consiste em mutar as alças de interação da região Fab da IgG de acordo com a estrutura do antígeno de interesse. As mutações foram realizadas e analisadas de forma a selecionar complexos antígeno-anticorpo cada vez mais específicos. O modelo foi capaz de prever a estrutura do anticorpo anti-AQP4, algo inédito na literatura. Dispondo de tal estrutura, as técnicas computacionais como a MMC, docking Molecular, DM e a SMD, permitiram abordar questões fenomenológicas que até o momento não eram solucionáveis de maneira geral, tais como: (i) MMC das nanossuperfícies (ponta de AFM e substrato) funcionalizadas (recobertas); (ii) os protocolos iniciais para os cálculos de forças de adesão entre as nanossuperfícies funcionalizadas. Os resultados provenientes dos protocolos desenvolvidos direcionaram e corroboraram os resultados experimentais, em parceria com as alunas de doutorado Ariana de Souza Moraes (Projeto FAPESP nº 2015/05283-6) e Pâmela Soto Garcia (Projeto FAPESP nº 2014/21540-0).

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Desenvolvimento de nanoimunossensores de AFM

Os nanobiossensores desenvolvidos em nosso grupo de pesquisa são construídos utilizando o AFM.(2,11,13,15,44-45) O AFM é composto basicamente de uma sonda (ponta - *tip*), cerâmicas piezelétricas (*scanners*) para posicionar a amostra e realizar varreduras, circuitos elétricos de realimentação para controlar a posição vertical da sonda e um computador para mover os scanners de varredura, armazenar dados e convertê-los em imagens com *software* específico. No AFM, a amostra é varrida por uma ponta montada em uma mola. Os átomos da ponta interagem com aqueles da superfície, produzindo deflexão na mola devido às alterações causadas pelas forças de atração ou repulsão.

Os componentes essenciais de um AFM, e da maioria dos microscópios de varredura por sonda, são uma cerâmica piezoelétrica (*scanner*), um *cantilever* com a ponta, detector, que mede as variações de posição e de intensidade da luz produzidas pelas deflexões do *cantilever*, sistema de realimentação (*feedback*) e um computador, que armazena e processa as diferentes deflexões da sonda, transformando-as em imagens topográficas da superfície (2D ou 3D), como mostra a FIG. 1.(46)

Figura1 – Diagrama esquemático do sistema de Microscopia de Força Atômica.

Fonte: LEITE et al.(15)

O AFM também é aplicável ao estudo de interações moleculares com resolução na ordem de 10^{-12} N sugerindo a possibilidade de medida de interações individuais usando a técnica de curva de força ou espectroscopia de força atômica.(45) Quando se determina a curva de força em função da distância em um local determinado da superfície da amostra, denomina-se espectroscopia de força local. A curva de força depende de vários parâmetros, como o material que constitui a ponta da agulha de prova e sua funcionalização, o tipo de superfície em estudo, as moléculas adsorvidas na superfície do material e o ambiente que envolve a superfície e a ponta. Mizesetal.(47) foram os primeiros a medir diretamente a variação espacial da adesão. Desde então, duas linhas principais têm caracterizado a pesquisa de curvas de força: (1) o estudo de diferentes interações em vários meios e (2) a “cartografia” de tais interações, traçado em curvas de força, a fim de distinguir materiais com diferentes propriedades físico-químicas (uma espécie de “espectroscopia” de superfície). A FIG. 2 mostra uma curva de força típica, identificando suas regiões: A-B: a ponta do *cantilever* é induzida ao contato com a amostra; B-C: a ponta é pressionada contra a superfície da amostra

até a região C e o cantilever se curva para cima; C-D: o *cantilever* começa a entrar em equilíbrio com as forças de superfície; D-E: o *cantilever* sofre uma deflexão devido às forças de atração (adesão); E-F: no ponto E tem-se a ruptura da ponta com o substrato. No ponto G, o *cantilever* volta à sua posição inicial. A distância vertical F-G corresponde ao valor da força de adesão entre a ponta e a amostra (pontual) e; G-H: a ponta e a amostra já não estão mais em contato.(48-49)

A detecção de forças interfaciais com resolução nanométrica é importante para entender, em nível molecular, fenômenos como fricção, lubrificação, adesão e fratura em interfaces, forças estruturais, estabilidade coloidal e sensoriamento. A resolução em força no AFM pode ser menor do que a magnitude da ligação química mais fraca (7), o que torna possível investigar interações moleculares individuais. Entretanto, grupos químicos específicos adsorvidos na ponta do AFM e interagindo com outras superfícies são ainda pouco definidos ou estudados e devem ser exploradas.

Figura2 – (A) Curva de força típica e (B) Diagrama esquemático do movimento vertical da ponta do AFM durante a aproximação e retração de uma curva de força.
 Fonte: Adaptada de BUTT(48);CAPPELLA.(49)

O conceito de modificação química de sondas (8) torna o AFM sensível a interações moleculares de interesse.(50) Empregam-se, nesse caso, pontas funcionalizadas no estudo de forças específicas por meio de curva de força. Tais técnicas já são usadas para: (a) detectar forças entre diferentes grupos moleculares, (b) medir energia de superfície em escala nanométrica e (c) mapear a distribuição espacial de grupos funcionais e seu estado de ionização.(48) Na microscopia de força química, pontas funcionalizadas são usadas em espectroscopia de força e obtenção de imagem (contraste) de domínios químicos.(3,51-52) A funcionalização pode ser feita com monocamadas auto-organizadas de grupos tióis (-SH) orgânicos (53) funcionalizados sobre pontas recobertas com ouro (Au) ou de grupos silanos (-SiOX) (54) depositados sobre pontas de silício.

Os atuais nanobiossensores divulgados na literatura são usados no monitoramento de analitos e contaminantes ambientais de forma eficiente (1-4,11-12,55-57), porém nenhum trabalho faz menção à detecção de doenças autoimunes e neurodegenerativas(58-60), muito provavelmente devido a vários problemas experimentais associados com a perda de atividade biológica de anticorpos pós imobilização em superfícies sólidas. Existem diversas razões para

que isto ocorra e uma delas pode ser atribuída à orientação randômica de moléculas assimétricas nos suportes das superfícies utilizadas.

2.2 Sistema nervoso central e doenças desmielinizantes

O sistema nervoso é um tecido que tem a função de regular e responder aos estímulos internos e externos do corpo, sendo responsável também por comandar todos os movimentos, voluntários ou involuntários.(61) O sistema nervoso é dividido em sistema nervoso periférico e sistema nervoso central (SNC). Por sua vez, o SNC é composto pela medula espinal e encéfalo que são envoltos pelo LCR na coluna vertebral e caixa craniana.(62) No tecido nervoso encontram-se os neurônios, que são células especializadas responsáveis por transmitir impulsos nervosos pelo corpo. Os neurônios são compostos basicamente por um corpo celular e axônios, estes revestidos pela bainha condutora formada pela mielina.(62) A mielina, portanto, está presente em grande quantidade em regiões do encéfalo com grande quantidade de neurônios. Essas regiões são conhecidas como massa branca, sendo as regiões compostas predominantemente dos corpos celulares dos neurônios chamadas de massa cinzenta.(62-63)

A desmielinização ocorre quando as proteínas constituintes são reconhecidas como autoantígenos por anticorpos do sistema imunológico, em um processo inflamatório.(64) A perda da mielina leva a deterioração dos neurônios, células formadoras e ao redor, podendo ocorrer no SNC e periférico.(65) Esse estado está relacionado à diversas doenças autoimunes e neurológicas, como a NMO, EM e a encefalomielite aguda.(62,64-72)

2.3 Neuromielite Óptica (NMO) e o marcador anti-AQP4

A neuromielite óptica NMO é uma doença autoimune que afeta o Sistema Nervoso Central (SNC), causando lesão na medula espinhal e nervo óptico e, assim, cegueira e paraplegia. Hipóteses sugerem que a NMO seja mediada por anticorpos específicos contra aquaporina4 (AQP4), uma biomolécula específica que funciona como canal de entrada de água na estrutura dos astrócitos.(17-18) O anti-AQP4, desta forma, pode ser considerado um biomarcador específico para NMO (18-19) e sua detecção no soro de pacientes propicia o diagnóstico efetivo e tratamento precoce para evitar ou retardar possíveis reincidências. Os testes diagnósticos empregam diferentes técnicas, como a Imunofluorescência indireta (73,74) e ELISA (do inglês *Enzyme-linked Immunosorbent Assay*) (75), possuem especificidade na faixa de 91 a 100%, no entanto, a sensibilidade de apenas 30% a 70% (18,23,76-78) para a detecção do anti-AQP4, o que dificulta o diagnóstico preciso da NMO.

A identificação precoce da NMO é importante sob vários aspectos, dentre estes, pode-se destacar: (i) A NMO é uma doença com prognóstico diretamente relacionado à gravidade das lesões e potencial de óbito (17,19,79); e (ii) a definição diagnóstica com um único episódio e detecção da IgG que atua como anticorpo anti-AQP4 é fundamental para proteger pacientes de próximos eventos por possibilitar tratamento adequado. Atualmente, apesar de avanços com os critérios de Wingerchuck(20) e dos revisados (23), bem como sobre a distinção clínica, laboratorial e de ressonância magnética da NMO, há certa confusão quanto ao diagnóstico diferencial entre NMO e EM, particularmente nas formas não típicas.(80) O tratamento destas doenças é distinto, sendo o da NMO com imunossupressores e anticorpos monoclonais (81–84), enquanto o da EM com imunomoduladores.(85) Pacientes com NMO confundidos com EM e submetidos a tratamento com imunomoduladores podem ter o agravamento da doença.(86) Os mecanismos moleculares que descrevem a NMO e a EM ainda

não estão bem estabelecidos e as ferramentas utilizadas no diagnóstico e tratamento não vêm fornecendo resultados promissores.(21,87-88)

A fisiopatologia da NMO envolve o reconhecimento da AQP4, uma proteína dos canais de água (89-90) localizada nos astrócitos, por anticorpos específicos do próprio sistema imunológico, o anti-AQP4, e a ação das moléculas do complemento como ilustrado na FIG. 3. O reconhecimento do antígeno e a ação do complemento causam lesão desmielinizante, axonal e necrose, espessamento e hialinização vascular, aspectos que caracterizam a lesão na NMO.(22) Neste contexto, esta pesquisa fornece modelos descritivos dos sistemas em questão, bem como irá auxiliar no desenvolvimento de metodologias para diagnosticar a NMO, correlacionando resultados experimentais com modelos teóricos desenvolvidos.

Figura3 – Esquema representativo do mecanismo de ação da neuromielite óptica.
 Fonte: Adaptada de WINGERCHUK et al. (22)

Experimentos adicionais realizados tiveram como base a detecção do anticorpo antiGlicoproteína oligodendrocítica da mielina (MOG, do inglês *myelinoligodendrocyteglycoprotein*) MOG, pois se acredita que os mesmos podem induzir desmielinização na matéria branca e cinzenta, causando inflamação no SNC.(61,67) O papel dos anticorpos contra a MOG na EM também é um modelo instrutivo para a avaliação de anticorpos em qualquer doença autoimune humana.(67,91-92) Segundo Lampasona e colaboradores (93), o anticorpo anti-MOG está ligado à extensa desmielinização, tanto na EM quanto em outras doenças, e outros estudos relataram a presença do anti-MOG no soro e LCR de pacientes com EM.(94) O papel da anti-MOG no início da EM ainda é desconhecido. Ainda não se sabe se eles estão presentes desde o início da doença ou desenvolvida mais tarde durante o curso da EM.(95) Neste contexto, o experimento foi conduzido de forma que o

epítopo MOG₉₂₋₁₀₆ foi utilizado como sensor para detectar o anticorpo anti-MOG, agindo como biomarcador de desmielinização em pacientes com EM.

2.4 Simulações de Dinâmica Molecular (DM)

2.4.1 DM e potenciais envolvidos

A DM é uma poderosa e versátil ferramenta da MMC, cuja aplicabilidade se mostra adequada ao estudo de macromoléculas biológicas. Os métodos experimentais em geral são capazes de mensurar propriedades resultantes da soma de todos os possíveis microestados entrópicos e energéticos que um conjunto de moléculas pode ocupar.⁽⁹⁶⁾ Entretanto, mesmo com a constante evolução de metodologias experimentais, há a necessidade da utilização de modelos teóricos para previsão de propriedades e compreensão, em nível molecular, dos experimentos.⁽³⁾ Sendo assim, o comportamento dos parâmetros experimentais pode ser usado para validar as médias obtidas pelas metodologias teóricas, sendo desenvolvidas para realizar previsões a cada sistema relacionado.⁽⁹⁷⁻⁹⁹⁾

A DM aproxima os átomos para esferas dotadas de carga com posições definidas por um conjunto de potenciais. Sendo assim, a estrutura molecular é descrita por potenciais lineares (V_d), angulares (V_θ) e torcionais (V_ϕ), que visam a reprodução das interações ligadas das moléculas a serem simuladas. As interações não ligadas, ou intermoleculares, são descritas pelos potenciais de Lennard-Jones (V_{LJ}), ou van der Waals, e de Coulomb (V_C), o que permite obter uma expressão da energia potencial total do sistema (V_{total}). O somatório dos potenciais ligados e não ligados compõe o campo de força do sistema (Equação 1), que depende somente das posições atômicas.^(98,100) Os parâmetros que descrevem as energias potenciais podem ser obtidos por cálculos teóricos ou técnicas experimentais, como cristalografia de raios-X, espectroscopia de infravermelho, dentre outras.⁽¹⁰¹⁾

$$V_{total} = V_{ligados} + V_{n\acute{o}o\ ligados} = V_d + V_\theta + V_\phi + V_{LJ} + V_C \quad (1)$$

Em outras palavras, as moléculas são descritas a partir de um modelo de esferas dotadas de cargas ligadas por “molas” devidamente calibradas, de modo que estas possam descrever as vibrações de ligações químicas e as barreiras energéticas torcionais e angulares que simulam o comportamento da molécula a ser mimetizada.(98) O somatório dos potenciais ligados e não ligados compõe o que é denominado campo de força do sistema. Cada função de energia potencial visa a descrever um comportamento estrutural característico da molécula.

O estudo mais aprofundado desta etapa resultou na publicação do capítulo de livro intitulado Modelagem Molecular Aplicada a Nanobiosistemas, do livro Nanociência & Nanotecnologia: Princípios e Aplicações (Grandes Áreas da Nanociência e suas Aplicações), pela editora Elsevier, disponível no **ANEXO C**. Este capítulo aborda a aplicação da DM e a MMC na modelagem de biomoléculas e suas devidas aplicações.

2.4.2 Minimização de energia

Um passo importante que faz parte dos cálculos de DM é a minimização de energia dos sistemas. Em um sistema real a configuração e conformação das moléculas levam ao estado de menor energia. Na DM a conformação das moléculas é descrita pelos parâmetros de campo de força. Seu conhecimento é importante, pois a estrutura, principalmente as proteínas, tem grande efeito em suas interações e função na natureza. Tais interações incluem o alongamento ou compressão das ligações (interações ligadas) além das interações não ligadas. Desta forma as contribuições das energias potenciais descritas no campo de forças possuem um ponto de equilíbrio (estado de menor energia), aproximadas por um poço potencial harmônico, levando-se em consideração as interações ligadas (FIG. 4).

Figura 4 – (A) Estados de compressão, equilíbrio e alongamento para moléculas ligadas. (B) Potencial harmônico aproximado para interações ligadas.

Fonte: SHATTUCK. (102)

O programa GROMACS 4.5.4 (103-104) possui três algoritmos de minimização: *steepestdescent*, *conjugategradients* e *bfgs* (*limited-memoryBroyden-Fletcher-Goldfard-Shannoquasi-Newtonianminimizer*). (104-105)

2.4.3 O RMSD como uma forma de análise

Alguns dos principais parâmetros utilizados para avaliar a evolução do sistema são energias de interação, gráfico de Ramachandran (que avalia a conformação das proteínas em termos da variação conformacional das cadeias principais) ligações de hidrogênio, pontes salinas e o RMSD (raiz quadrada do desvio quadrático médio, do inglês *Root Mean Square Deviation*) (FIG. 5). Destes o RMSD é o parâmetro mais importante indicador da flutuação média dos resíduos que compõem a macromolécula, sendo fundamental na análise da estabilidade dos sistemas citados nesta pesquisa. O RMSD mostra a flutuação ocorrida na posição dos átomos da estrutura em relação à estrutura inicial, sendo a medida do quanto os átomos de macromolécula variaram no tempo da simulação (97), gerando um gráfico que, quando o sistema se estabiliza, aproxima-se de uma reta horizontal.

Figura 5 – RMSD de duas diferentes estruturas, para uma DM de 10000 ps. A curva na cor vermelha indica uma estrutura mais estável, dentro do tempo de simulação, enquanto a curva na cor preta indica uma estrutura menos estável, para o mesmo tempo de simulação.

Fonte: Adaptada de FRANCA et al. (3)

2.5 Dinâmica Molecular Direcionada (SMD)

Introduzido por Izrailev em 1997 (106), a SMD foi inspirada pelo uso da técnica de AFM, usada para estudar propriedades de ligação de biomoléculas e suas respostas à engenharia externa.(107-108) Com esta metodologia estuda-se a dinâmica de eventos de ligações não covalentes em sistemas biomoleculares e suas propriedades elásticas, em escalas de tempo abrangidas por DM, tipicamente em nanossegundos. A SMD aliada às metodologias mencionadas anteriormente possibilitará a correlação entre dados computacionais obtidos para nanossistemas com as curvas de força dos experimentos de AFM.

A aplicação da SMD pode ser mediada por dois protocolos: no primeiro, uma força constante de extensão ou deslocamento é aplicada a um ou mais átomos, e esta é então monitorada ao longo da trajetória dinâmica; no segundo, uma velocidade constante é aplicada aos átomos mencionados anteriormente e um potencial harmônico é utilizado para induzir o movimento ao longo de uma determinada coordenada.(106) A FIG 6A mostra um típico

evento experimental de estiramento molecular com velocidade constante. Este método vem sendo usado pelo nosso grupo de pesquisa e pode ser reproduzido, com suas devidas limitações, utilizando o segundo protocolo de aplicação da SMD, conforme citado anteriormente. Ressalta-se que os sistemas a serem estudados estão relacionados com o mecanismo de interação entre enzima-inibidores e antígeno-anticorpos, mediadas por pontas de AFM funcionalizadas (FIG 6B).

Figura6 – (A) Representação esquemática sem escala do *cantilever* de AFM exercendo uma força em uma molécula alvo. A SMD é capaz de simular experimentos de AFM. (B) Esquema do sistema para investigações das interações enzima-inibidores e antígeno-anticorpos.

Fonte: Adaptada de FRANCA et al. (14).

No caso de complexos proteína-ligante, segundo Izrailev(106,107), uma forma de se aplicar forças externas é direcionar o ligante para um ponto no espaço aplicando um potencial harmônico para que o mesmo seja removido do complexo. Por exemplo, assumindo uma única coordenada de reação x (Equação 2), e um potencial externo, sendo k a rigidez do sistema de retenção, x_0 a posição inicial do ponto de apoio e velocidade v , a força externa exercida sobre o sistema pode ser expressa com a Equação 3.(14,106-107)

$$V = k \cdot (x - x_0 - vt)^2 / 2 \quad (2)$$

$$F = k \cdot (x_0 + vt - x) \quad (3)$$

Esta força medida corresponde a uma molécula sendo “puxada” por uma mola com constante elástica k . A FIG. 7 mostra um exemplo de aplicação de uma força externa em uma molécula numa direção pré-estabelecida, provocando sua remoção do complexo.

Figura 7 – Remoção de uma molécula ligante a partir da cavidade de ligação de uma proteína. A força (representada por uma seta) aplicada ao ligante (representada em esferas de van der Waals) leva a sua dissociação do complexo proteína-ligante.(14,106)

Fonte: Adaptada de FRANCAet. al.(14)

2.6 Docking Molecular

A metodologia de *docking* molecular consiste em prever a orientação preferencial de uma molécula em relação à outra quando estas formam um complexo estável.(109) Na maior parte das vezes, uma molécula maior é escolhida como receptora e uma molécula menor como ligante. A molécula ligante pode ser um fármaco, um antígeno ou alguma molécula específica na qual se deseja estudar sua interação com determinada molécula receptora.(109) Este processo de reconhecimento molecular é regido por uma série de combinações e efeitos entrópicos ΔS e entálpicos ΔH . Esses efeitos são previstos pela da energia livre de ligação de Gibbs (ΔG_{lig}) (Equação 4).(110)

$$\Delta G_{lig} = \Delta H - T\Delta S = RT \ln k_i \quad (4)$$

sendo ΔH a variação da entalpia, T a temperatura, ΔS a variação da entropia e k_i um parâmetro experimental.

Entre os algoritmos utilizados para os cálculos de docking molecular destacam-se os algoritmos genéticos. Estes são algoritmos evolucionários pertencentes a uma classe de

métodos estocásticos de otimização global. Esses algoritmos são algoritmos evolucionários, inspirados no processo biológico de evolução das espécies e na teoria da sobrevivência do mais apto de Darwin.(111-112) Em outras palavras, são algoritmos inspirados no processo de evolução das espécies e na teoria da sobrevivência do mais apto.(113) Os algoritmos genéticos são aplicados principalmente a simulações com ligantes contendo graus de liberdade e rotação. Cada geração de uma nova população é obtida por troca de “genes” (no caso informações) entre dois indivíduos geradores, proporcionando mudanças aleatórias nos valores dos genes transmitidos, ou seja, transmitindo informações. O processo vai se repetindo ao modo que a população evolua para soluções (ou configurações) mais favoráveis, até ser parado por intermédio de um critério pré-estabelecido.(109) Atualmente, o programa *AutoDock* utiliza um algoritmo evolucionário chamado Lamarckiano.(114) Neste algoritmo, cada geração de uma porção da população é escolhida ao acaso para a busca local, assim o indivíduo escolhido substitui o anterior, fazendo uma alusão à teoria da hereditariedade de características de Lamarck.(114-115)

2.7 Cálculos de potenciais eletrostáticos e energia de ligação

A energia de ligação está envolvida na quebra ou na formação de ligações químicas provenientes das interações moleculares. No contexto da DM, as contribuições majoritárias das interações são feitas por cargas eletrostáticas.(116,117) Com isso, a avaliação do potencial eletrostático, ou seja, a distribuição de cargas de um sistema é importante para mostrar as regiões interativas do mesmo. Este parâmetro varia conforme as condições termodinâmicas do meio, podendo ser representado de maneira similar à densidade eletrônica.(118) Em simulações de sistemas moleculares e imobilização de biomoléculas em superfícies funcionalizadas, o potencial eletrostático é útil no entendimento de interações

intermoleculares (3) e permite a representação visual das regiões com cargas elétricas diferentes ao longo da superfície de biomoléculas.

Uma das formas de se calcular a energia livre de ligação em MMC é utilizando um ciclo de energias, como mostrado na FIG 8.

Figura 8 – Ciclo de energia livre de ligação.
Fonte:BAKER.(119).

O ciclo de energia livre de ligação $\Delta_b G$ (FIG. 8) mostra a ligação em termos de energias livres de transferência, de um ambiente dielétrico homogêneo (interações são descritas pela Lei de Coulomb) para um ambiente dielétrico não homogêneo, com constantes dielétricas internas (representado na cor verde) e externas (representado na cor ciano) diferentes. A energia livre de ligação (ou dissociação) é calculada na Etapa 3 (FIG. 8), de acordo com a Equação 5:(119)

$$\Delta_b G = -\Delta_3 G = \Delta_4 G - \Delta_1 G - \Delta_2 G \quad (5)$$

O método mais geral para calcular as energias livres de ligação $\Delta_b G$ consiste em dividir as energias em dois componentes: energia de solvatação $\Delta_s G$ e energia coulombiana (eletrostática) $\Delta_c G$ (Equação 6):

$$\Delta\Delta_b G = \Delta\Delta_s G + \Delta\Delta_c G \quad (6)$$

Para calcular a contribuição da energia total de solvatação $\Delta\Delta_s G$ (etapas 4 e 2 no ciclo de energia livre da FIG 8), basta calcular as energias de solvação para o complexo e os componentes isolados (proteína e ligante). A contribuição da energia de solvatação para a energia de ligação é calculada então de acordo com a Equação 7.

$$\Delta\Delta_s G = \Delta_4 G - \Delta_2 G = \Delta_s G_{complexo} - \Delta_s G_{molécula1} - \Delta_s G_{molécula2} \quad (7)$$

Completando o ciclo de energia livre (Equação 222), a contribuição da energia coulombiana é calculada de forma análoga, calculando as energias do complexo e dos componentes isolados (Equação 8).

$$\Delta\Delta_c G = -\Delta_1 G = \Delta_c G_{complexo} - \Delta_c G_{molécula1} - \Delta_c G_{molécula2} \quad (8)$$

Cada uma das grandezas na equação é a soma de interações coulombicas entre todos os átomos da molécula (ou complexo) para um dielétrico em particular. A fim de combinar estas energias de ligação coulombiana com as energias de solvatação descritas acima, precisamos assegurar que constantes dielétricas consistentes sejam usadas.(119,120).

3 OBJETIVOS

3.1 Objetivo geral

Caracterizar o recobrimento biológico de nanossuperfícies envolvidas em experimentos de AFM e desenvolver um novo protocolo para modelar anticorpos, auxiliando no desenvolvimento de nanoimunossensores para detectar biomarcadores de doenças desmielinizantes, em especial a NMO.

3.2 Objetivos específicos

- i. Aprimorar e desenvolver modelos para o recobrimento das nanossuperfícies funcionalizadas (pontas de AFM e substratos);
- ii. Desenvolver modelos de engenharia molecular e DM para modelar anticorpos;
- iii. Desenvolver modelos para o recobrimento das nanossuperfícies funcionalizadas com biomoléculas relacionadas a doenças autoimunes;
- iv. Mensurar as forças de interação (adesão) antígeno-anticorpo utilizando SMD e correlacionar os resultados da simulação com resultados experimentais de AFM.

4 METODOLOGIA

4.1 AFM parte experimental

O fornecimento de IgGs referentes aos pacientes soropositivos para NMO foram purificadas e fornecidas pelo Dr. Luís Antonio Peroni, especialista na área de imunologia e sócio da empresa P & D da Rheabiotech Desenvolvimento, Produção e Comercialização de Produtos de Biotecnologia Ltda (Campinas, SP).

Para os peptídeos, as pontas de AFM foram funcionalizadas com uma solução de 1-etil-3-(3-dimetilaminopropil) carbodiimida (EDC, 0,4M), e para a ativação dos grupamentos carboxílicos terminais que interagem com os grupamentos de amina do APTES, foi utilizada solução de N-Hidroxi-succinimida (NHS, 0,1M). Em seguida foi adicionado o polietilenoglicol (NH₂ – PEG – COOH 2000 u e com concentração 0,01 µg.mL⁻¹) (121–123), Sigma-Aldrich®, St. Louis, Missouri, EUA). Durante o processo as moléculas foram lavadas com água MiliQ e adicionadas para imobilização em solução com peptídeos (0,1 µg.mL⁻¹).

Para a imobilização das IgGs foi utilizada uma solução de 30 µL de proteína e água MiliQ (0,01 µg.mL⁻¹), mantida por 2h em temperatura ambiente e as soluções foram limpas por três vezes com água e com 30 µL de solução de IgGs (56 µg.mL⁻¹).

4.2 MMC das biomoléculas envolvidas no imunonanobiossensor

As estruturas das biomoléculas (proteínas e autoantígenos) foram obtidas com base em modelos experimentais de cristalografia de raios-X provenientes do repositório de estruturas tridimensionais de moléculas Protein Data Bank (PDB) (24) (3GD8 (124), 1IGY (125), 2R0Z (126), 1KC5 (127) e 3BKY (128)). Átomos de hidrogênio foram adicionados para criar um modelo atômico fidedigno ao real, contendo todos os átomos explícitos para realizar cálculos de DM.

Em condições ambientes, as nanossuperfícies (pontas de AFM e substrato), que podem ser compostas de silício (Si) ou nitreto de silício (Si_3N_4), sofrem oxidação completa ou parcial, formando grupos SiO_2 na superfície e SiOH nas camadas mais externas, respectivamente. Estes grupos reagem fortemente com grupos silanos que são utilizados para funcionalização, como o 3-aminopropiltietoxissilano (APTES) (FIG. 9A). O APTES (FIG. 9) é um *linker* comumente utilizado para fins de funcionalização de superfícies.(4-5,11,57) A extremidade $-\text{CH}_3$ da molécula de APTES se liga ao óxido de silício que é formado nas nanossuperfícies do nitreto de silício (FIG. 9) (57,129), material que constitui a ponta do AFM e o substrato de mica (material predominante de silício, filossilicato, (130)). A outra extremidade da molécula de APTES se liga a um dos grupos $-\text{COOH}$ distribuídos ao longo da superfície das proteínas ou nas moléculas a serem funcionalizadas.

Figura 9 – Descrição do processo de funcionalização da superfície de Si_3N_4 da ponta do AFM. (A) Molécula de APTES: uma extremidade $-\text{CH}_3$ se liga ao óxido de silício e, na outra extremidade, uma amina primária ($-\text{NH}_2$) permanece disponível para interação. (B) Superfície de Si_3N_4 funcionalizada com APTES.

Fonte: Elaborada pelo autor.

Os parâmetros dos ligantes APTES e da nanossuperfície de Si_3N_4 , tanto para a ponta quanto para a superfície de mica (estrutura predominante de silício (material filossilicato) (130)), e a síntese *in silico* do peptídeo da AQP4 foram obtidos utilizando o programa ORCA(131). As cargas atômicas parciais dos ligantes foram determinadas usando o método de RESP (do inglês, *Restrained electrostatic Potential*)(132) no

programa NWChem.(133) Para a visualização e produção de imagens referentes a todos os sistemas foi utilizado o programa VMD (do inglês, *Visual Molecular Dynamics*) 1.9.1 (134) e todas as medições foram realizadas utilizando o programa Jmol 10.0.31.(135)

4.3 Cálculos de *docking* Molecular

Os cálculos de *docking* foram conduzidos utilizando os epítopos de interesse (AKFRHD, HQLDPAFG, EPANPSEKNSPSTQYC e GTEKPLPVD) como estrutura ligante acopladas às estruturas das IgGs modeladas 2R0Z (126), 1KC5 (127) e 3BKY (128) e 3GD8 (124)), respectivamente, conforme será descrito detalhadamente no item 5.4, a fim de prever orientações conformacionais favoráveis a serem utilizadas em simulações de DM. O protocolo de *docking* molecular empregado no presente trabalho pode ser descrito como: i) aplicação do método de Gasteiger-Huckel(136) implementado nas ferramentas do programa *Autodocking*(114), para atribuição das cargas parciais do sistema; ii) avaliação da propensão de ligação entre os ligantes estudados e a IgG por intermédio do programa *AutoGrid* dentro das prováveis regiões de interação antígeno-anticorpo; e iii) aplicação do algoritmo genético Lamarckiano (LGA) para pesquisa do espaço conformacional das células unitárias. As posições de encaixe foram cuidadosamente avaliadas segundo um desvio quadrático médio (RMSD) de 0,2 nm, utilizando as energias mais baixas totais.

4.4 Cálculos de potenciais eletrostáticos e energias de ligação

O protocolo empregado, que engloba as distribuições de cargas e energias de ligação, foram realizados com a aplicação do programa APBS (do inglês, *Adaptive Poisson-Boltzmann Solver*). (120,137-139) Os dados sobre as estruturas das proteínas em solução aquosa foram obtidos usando constante dielétrica $\epsilon = 78,54$ para o solvente e $\epsilon = 2,00$ para o soluto.(3) As especificações do solvente envolvem três parâmetros: tensão superficial 0,105

N/m, raio de corte de 1,4 nm e força iônica de 100 mM.(3) Os potenciais tridimensionais resultam de 129 pontos de rede nas direções de x, y e z.

4.5 Modelagem da estrutura de anticorpos por MMC (Protocolo desenvolvido)

4.5.1 MMC completa das alças dos anticorpos

A estrutura da IgG subtipo 1 foi modelada de forma genérica, sendo construída a partir da estrutura PDB ID: 2R0Z (126) disponível no repositório de estruturas tridimensionais de moléculas PDB. As alças CDRs que compõe a região Fab de ligação foram divididas em pares de resíduos separados por aminoácidos glicina, para conferir espaço físico para conformação e modelagem do sítio de ligação. A FIG. 10A mostra a estrutura inicial modelada de acordo com o comprimento e número de resíduos, sendo que a alça H3 foi modelada com 2 divisões (representadas na cor azul), cada uma com um par de aminoácidos cada, separadas por glicinas (as quais serão denominadas divisões glicina, representadas na cor vermelho). A alça H2 foi modelada com 4 divisões, contendo cada uma um par de aminoácidos separados por divisões glicina, enquanto a alça L3 foi modelada contendo 3 divisões. As demais alças, H1, L1 e L2 foram modeladas com 3 divisões. Na base de cada alça também foram atribuídas divisões glicina (representadas na cor vermelho). Cada divisão, contendo um par de aminoácidos, possuem 20x20 combinações possíveis (correspondentes aos 20 aminoácidos), totalizando 400 mutações cada. Em sequência o epítipo alvo, no qual se deseja modelar o anticorpo específico, é posicionado por cálculos de *docking* molecular. A estrutura inicial da FIG. 1x2A foi o ponto de partida para ser utilizada como *input* padrão do protocolo desenvolvido. O epítipo alvo, no qual se deseja modelar o anticorpo específico, é posicionado por cálculos de *docking* molecular (FIG. 10B).

Figura 10 – (A) Estrutura inicial da porção Fab da IgG, destacando as alças mutadas. Na cor azul são representadas as divisões das alças, cada divisão contendo 2 aminoácidos a serem mutados. São possíveis 400 mutações em cada divisão das alças (em azul). Na cor vermelho, são representados os aminoácidos glicina, inseridos por mutações pré-estabelecidas. (B) Complexo inicial com epítipo alvo nposicionado por cálculos de *docking* molecular.

Fonte: Elaborada pelo autor.

A primeira modelagem consiste na formação da melhor região de ligação com o epítipo, segundo ΔG de ligação, sendo modelada uma a uma cada alça, como descrito nos passos seguintes:

Passo 1) Executar cálculos de *docking* e obter o melhor complexo ligando o epítipo.

Passo 2) Mutar as 400 combinações para cada uma das alças, gerando 7200 novas estruturas (complexos).

Passo 3) Executar 500 ps de cálculos de DM para cada uma das 7200 estruturas e obter o melhor complexo, analisando o menor valor da energia de ligação ΔG .

A segunda modelagem tem o objetivo de formar o melhor complexo antígeno-anticorpo segundo ΔG de ligação. Para isso é necessário a nova divisão dos resíduos que compõem as alças de interação. A nova estrutura dividida de acordo com o comprimento e número de resíduos escolhidos, contendo um par de aminoácidos cada. As alças L1, L2, L3,

H1, H2 e H3 foram divididas da mesma maneira, contendo 5, 4, 5, 4, 7 e 4 divisões respectivamente. Cada divisão, contendo um par de aminoácidos, possuem 20x20 combinações possíveis (correspondentes aos 20 aminoácidos), totalizando 400 mutações cada, totalizando 11600 estruturas. Desta forma as alças foram modeladas uma a uma como descrito nos passos seguintes:

Passo 4) Mutar as 400 combinações para cada uma das alças, gerando 11600 novas estruturas (complexos).

Passo 5) Executar 500 ps de cálculos de DM para cada uma das 11600 estruturas e obter o melhor complexo, analisando o menor valor da energia de ligação ΔG .

A partir dos passos apresentados, espera-se que o modelo obtido leve à formação do melhor complexo, segundo ΔG de ligação. Após a obtenção do melhor complexo, deve-se executar cálculos de DM a fim de refinar o modelo e eliminar possíveis repulsões estéricas no sistema (3,4) e novamente calcular ΔG de ligação. Vale ressaltar que os passos descritos podem ser repetidos livremente, observando também a funcionalidade da estrutura, dependendo do estudo realizado.

4.5.2MMC do anticorpo anti-Aquaporina-4 (anti-AQP4)

Na ponta do AFM são funcionalizados os peptídeos da AQP4. Tal estrutura foi obtida a partir da modelagem e adequação do arquivo PDB ID: 3GD8 (124) para o estudo computacional do recobrimento da superfície da ponta. Na FIG. 11, a sequência peptídica candidata a epítipo-alvo do anti-AQP4 também é destacada, formada pela sequência GTEKPLPVDM, de suma importância no funcionamento do imunonanosensor.

A AQP4 possui papel crucial nas simulações, pois contém os epítopos-alvo da anti-AQP4 com conformação nativa. Os peptídeos localizados em sua grande maioria no meio

extracelular, ou seja, no domínio externo da proteína AQP4 (124,140), têm maior chance de funcionar como um epítipo-alvo (FIG. 11) dos anticorpos anti-AQP4. Segundo testes em nosso grupo de pesquisa (resultados ainda não publicados, obtidos em parceria com a Prof. Dra. Doralina Guimarães Brum Souza), o epítipo formado pelos resíduos GTEKPLPVDM obteve maior afinidade com anticorpos anti-AQP4 encontrados no soro de pacientes com NMO.

Figura 11 – Modelagem molecular da proteína AQP4, destacando um possível epítipo-alvo do anticorpo anti-AQP4.

Fonte: Modelagem a partir do Arquivo PDB 3GD8.(124)

4.6 Simulações de DM e SMD

A estrutura do complexo peptídeo da AQP4 e IgG anti-AQP4 foi tomada como ponto de partida para os cálculos de DM e SMD. O sistema inicial para a simulação por SMD foi obtido a partir da trajetória de um sistema equilibrado numa simulação de DM contendo o complexo mencionado, realizada para o tempo de 10 nanossegundos, pela aplicação do seguinte protocolo: (i) nanossistema solvatado numa caixa com modelos de água do tipo SPC (do inglês *single point charge*) (141) com íons sódio e cloreto sendo adicionados para atingir a neutralidade de cargas; (ii) minimização de energia do sistema utilizando 10.000 passos do algoritmo *steepest decent* (104); (iii) equilibração do sistema por meio de uma simulação de DM de 100 picossegundo com termalização gradativa de 0 a 310 K; (iv) simulação por DM de 10 nanossegundos, realizada em um *ensemble* isotérmico-isobárico (NPT) utilizando o algoritmo *leapfrog* (142) com um tempo de integração de 1 femtosegundo; (v) controle de

temperatura em 310 K e pressão em 1 bar por intermédio do acoplamento do sistema ao termostato de Berendsen(143) e ao barostato de Berendsen(143); (vi) aplicação do algoritmo LINCS (144) para fixar os movimentos vibracionais do sistema; (vii) aplicação do raio de corte de 1,4 nm(4,14) para a caracterização das interações curto alcance de van der Waals e eletrostáticas; (viii) tratamento das contribuições eletrostáticas de longo alcance via campo de reação generalizada (145), com uma constante dielétrica $\epsilon = 66$. Todas as simulações foram realizadas utilizando os campos de força OPLS-AA e o programa GROMACS 4.5.4.

Os cálculos envolvendo a abordagem de SMD foram obtidos em parceria com a aluna de doutorado Jéssica Cristiane Magalhães Ierich (Processo FAPESP 2014/12082-4). As simulações foram realizadas para o sistema modelado considerando a porção variável do anticorpo (Fv) e o peptídeo imunogênico da AQP4. Os cálculos computacionais foram conduzidos no programa NAMD 2.12 (146) em campo de força CHARMM36 (147). O complexo Fv-peptídeo obtido foi inserido na caixa de simulação contendo moléculas de água TIP3P (148), com neutralização de cargas do meio pela adição de 11 contra-íons Cl⁻ utilizando o programa Visual Molecular Dynamics – VMD (134). Após a preparação do sistema de simulação, este foi minimizado em ensemble NVT considerando: (i) 2fs de intervalo de tempo; (ii) 1,2 nm de distância de corte para cálculo das interações de curto alcance; e (iii) o formalismo de *particle-meshEwald* (PME) (149) para tratamento das contribuições eletrostáticas de longo alcance. O sistema minimizado teve sua temperatura gradativamente aumentada para 310 K em ensemble NPT com controle de pressão (1 bar) e temperatura (intervalos de adição de 62 K até atingir 310 K) pela aplicação do pistão e termostato de Langevin.(150)O trabalho mecânico realizado durante o processo de desacoplamento foi obtido integrando-se a força exercida sobre o sistema ao longo da coordenada z utilizada para definir a via de extração do peptídeo da AQP4. A descrição básica de um nanossistema utilizado neste estudo é dada na Tabela 1.

Tabela 1 – Descrição dos sistemas simulados

Nanossistema	nº de átomos de soluto	nº de íons	nº de moléculas de solvente	nº de resíduos de aminoácidos	Força Iônica (mol/L)
Peptídeo	182	6	1422	11	0,1
IgG parte Fab e peptídeo	6432	119	29683	440	0,1

Fonte: Nanossistemas modelados a partir dos Arquivos PDB: 3GD8 (124), 1IGY (125), 2R0Z (126), 1KC5 (127) e 3BKY (128).

4.7 A distribuição gama na interpretação dos experimentos

A distribuição gama é uma família de distribuições contínuas de probabilidade que são dependentes de dois parâmetros.(151) É uma das mais gerais distribuições, pois diversas outras são casos particulares dela, como a exponencial e a qui-quadrado. Tem aplicações meteorológicas e na análise de tempo de vida de produtos.(152) A distribuição gama já é utilizada na interpretação dos experimentos, como os de AFM, pela natureza estatística de alta dispersão dos dados.(153–156) Além da distribuição gama, outras foram testadas sem ter sido obtida a mesma representabilidade, como a Normal, Log-normal, F de Snedecor, Cauch, Gumbel, Beta, Exponencial, Weibull, Rayleigh, Rice, entre outras.

A Função Distribuição (f) e a Densidade de Probabilidade (F) são dadas nas Equações 7 e 8 respectivamente:

$$f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad \text{for } x > 0 \text{ and } \alpha, \beta > 0, \quad (7)$$

$$F(x; \alpha, \beta) = \int_0^x f(u; \alpha, \beta) du \quad (8)$$

sendo que a Função Distribuição tem o aspecto geral mostrado na FIG. 12, Dependente dos parâmetros α e β .

Figura12 – Curvas características da distribuição gama, dependentes dos parâmetros α e β .

Fonte: Elaborada pelo autor.

5 RESULTADOS E DISCUSSÃO

5.1 Recobrimento das nanossuperfícies (nanobiossensor desenvolvido)

5.1.1 MMC da Funcionalização da ponta do AFM

A nanossuperfície da ponta do AFM, representada na cor marrom na FIG. 13, foi modelada computacionalmente e funcionalizada com os peptídeos sintetizados GTEKPLPVDM representados na cor verde. As moléculas de APTES foram ligadas de forma covalente pelos grupos -COOH dos aminoácidos serina terminais dos peptídeos sintetizados, conforme discutido nos parágrafos anteriores. Desta forma, foi obtida uma representação computacional da ponta de AFM funcionalizada com peptídeos da AQP4, conforme evidenciado na FIGURA 13, sugerindo uma possível configuração para o recobrimento da ponta do AFM.

Figura 13 – Ponta de AFM (em marrom) funcionalizada com APTES (em ciano) e peptídeos da AQP4 (em verde).

Fonte: Elaborada pelo autor.

5.1.2 MMC da nanossuperfície do substrato (mica)

O protocolo utilizado para a funcionalização das nanossuperfícies pode afetar o funcionamento do nanoimunossensor, podendo comprometer em 50% ou 100% as regiões de

interação da IgG com o antígeno (157) e proporcionando a detecção de forças interferentes não específicas do sistema. Para subsidiar a metodologia de funcionalização descrita, e a imobilização orientada das moléculas de IgG nas nanossuperfícies, foram realizadas medidas de curvas de força utilizando a imobilização não orientada e orientada das IgGs no substrato.(158) Neste contexto, a análise da rugosidade das superfícies é uma maneira de caracterizar a homogeneidade da funcionalização, devido à irregularidade topográfica da amostra.(159) A imobilização orientada das IgG apresenta uma superfície uniforme e consequentemente mais homogênea topograficamente.(159) Sendo assim a MMC das nanossuperfícies foi realizada de forma orientada utilizando a proteína A.

Para a funcionalização do substrato foi necessário fazer uma estimativa da área ocupada pela proteína A íntegra e de como as moléculas da mesma se distribuem e se ligam na região Fc (constante) da IgG. A Fig. 14 mostra a MMC da IgG anti-fenobarbital interagindo com fragmentos da proteína A (representados na cor verde), de acordo com Deis et. al. 2015.(160) Tendo em vista a indisponibilidade de dados na literatura a respeito da estrutura cristalográfica íntegra da referida proteína, houve a necessidade de predizer um modelo para a proteína A íntegra com base em dados provenientes da literatura e de bancos de dados específicos.(160) De acordo com Erickson (161), é possível estimar o volume de proteínas globulares tendo como ponto de partida sua respectiva massa molecular. Sendo a massa molecular da proteína A íntegra equivalente a 42 kDa, o volume estimado pode ser aproximado em $50,9 \text{ nm}^3$.(161) Dado o volume estimado e sabendo que uma proteína A íntegra imobilizada em superfície pode se ligar a, no máximo, duas moléculas de IgG(160), uma possível estrutura para as proteínas ligadas à região Fc da IgG pode ser construída conforme representado na Fig. 14, na cor verde.

Figura 14 – MMC da IgG tipo 1. As cadeias leves estão representadas na cor púrpura e as cadeias pesadas, na cor rosa. Em verde estão os monômeros de proteínas A. (A) Interação entre monômeros de proteína A e IgG segundo dados cristalográficos. (B) Interação entre moléculas de proteína A íntegras e IgG segundo MMC.

Fonte: Modelagem realizada a partir do Arquivo PDB 1IGY (125); 4ZNC (160) ; 4WWI.(160)

As moléculas de IgG tipo 1 foram imobilizadas de maneira orientada na nanossuperfície modelada (FIG. 15). Considerando que a funcionalização do substrato é idêntica àquela realizada na ponta do AFM, as moléculas de APTES foram ligadas de forma covalente nos grupos -COOH presentes na nanossuperfície das proteínas A (162–165), auxiliando, assim, na construção do substrato a ser detectado, conforme representado na Figura 15.

Figura 15 – Nanossuperfície de mica (em marrom) funcionalizada com IgGs tipo 1 (em rosa e púrpura). Em verde estão os monômeros de proteínas A, utilizados para orientar a imobilização das IgGs.

Fonte: Modelagem realizada a partir dos arquivos PDB 1IGY (125); 4ZNC (160); 4WWI.(160)

Desta forma, a imobilização das IgGs de forma orientada também foi caracterizada experimentalmente, mantendo as regiões variáveis e disponíveis para interações com os epítomos.(165)

5.1.3 Análise experimental das nanossuperfícies funcionalizadas

As medidas experimentais descritas neste item foram realizadas pela B. Sc. Ariana de Souza Moraes (Projeto FAPESP nº 2015/05283-6). A análise experimental da imobilização orientada e não orientada das IgGs foi realizada com o anticorpo anti-atrazina. A imobilização orientada mostrou-se mais homogênea do que a imobilização não orientada, quando comparadas suas médias e os desvios das medidas (Fig. 16).

Figura 16 – Homogeneidade da imobilização orientada e não orientada. A imobilização orientada (vermelho) apresentou valores de rugosidade menores, ou seja, apresentou a superfície mais homogênea do que a imobilização não orientada (ciano escuro). Valores de rugosidade da imobilização aleatória: mica limpa, $(0,23 \pm 0,01)$ nm; mica com APTES, $(0,24 \pm 0,01)$ nm; mica com APTES e glutaraldeído, $(0,25 \pm 0,01)$ nm; IgG imobilizada, $(0,43 \pm 0,01)$ nm. Imobilização orientada: mica limpa, $(0,24 \pm 0,01)$ nm; com APTES, $(0,24 \pm 0,01)$ nm; proteína A, $(0,13 \pm 0,01)$ nm; IgG, $(0,26 \pm 0,01)$ nm.

Fonte: Elaborada pela Srta. Ariana de Souza Moraes (Projeto FAPESP nº 2015/05283-6).

Na imobilização não orientada a superfície permaneceu com a rugosidade inicial, em torno de $(0,25 \pm 0,01)$ nm, até a etapa de funcionalização com glutaraldeído. Na imobilização da IgG, a rugosidade aumentou para $(0,43 \pm 0,01)$ nm. Na imobilização orientada, a rugosidade inicial da superfície diminuiu para $(0,13 \pm 0,01)$ nm evidenciando uma superfície de proteínas A sem sítios remanescentes de APTES; após a imobilização da IgG, a rugosidade apresentou $(0,27 \pm 0,02)$ nm, tornando-se mais homogênea do que a imobilização aleatória.

5.2 Modelagem da estrutura de anticorpos por MMC (Protocolo desenvolvido)

5.2.1 A importância dos modelos preliminares

Os primeiros modelos desenvolvidos, baseados em cálculos de docking molecular, tiveram como ponto de partida a idéia de engenharia molecular (33–35,39-41) e a ausência na literatura de programas computacionais que fossem capazes de fazer a previsão da estrutura de anticorpos partindo apenas da estrutura de epítomos de interesse.(42–46,166) Neste contexto, a execução do protocolo de *docking* permitiu a obtenção de uma estrutura prévia da IgG anti-AQP4, mostrada no **APÊNDICE A**.

5.2.2 Estudo e validação do protocolo final desenvolvido

Com base nos resultados dos modelos preliminares, o protocolo final desenvolvido, descrito no item 4.4.2, funciona de modo a modelar a estrutura da IgG inicial para modelar a estrutura específica do anticorpo, partindo apenas do *input* da sequência previamente fornecida de aminoácidos que compõe o epítomo. Sua estrutura genérica foi sintetizada *in silico* a partir da sequência antigênica fornecida. Após o fornecimento das sequências antigênica (FIG 17) foram realizadas as síntese e tais estruturas foram submetidas a 2 ns de DM. Os epítomo, nos quais se deseja modelar o anticorpo específico, foram posicionados nas estruturas genéricas de cada IGg por cálculos de *docking* molecular.

Figura 17 – Estrutura dos peptídeos AKFRHD, HQLDPAFG e EPANPSEKNSPSTQYC sintetizados *in silico* e representados em bastões (esquerda) e *new cartoon* (direita).

Fonte: Elaborada pelo autor.

Após a execução dos cálculos de *docking* foi aplicado, a cada estrutura, o protocolo descrito no item 4.4.2. Para a reprodução do modelo, o protocolo foi aplicado três vezes para a estrutura AKFRHD da FIG. 17. Foram comparadas com suas respectivas estruturas e energias de ligação com a estrutura cristalográfica real, para qual foram realizadas as previsões mostradas nos Resultados 1, 2 e 3. Nota-se que as energias dos complexos previstos se assemelham à energia de ligação do cristal e outras energias descritas na literatura. (166)

Resultado 1

Antígeno

AKFRHD semelhança 64% e 75% com cristal 2r0z; energia do complexo -68,5 kJ/mol

Cadeia leve

DVLMQTPLSLPVSLGDQASISCRSSQSI**RPYRKWYWGP**WYLQKPGQSPKLL**ARHKRVPV**SGVPDRFSGSGSGTDFTLKISRVEAEDLGVYYC**VAQWWDAGAV**GAGTKLEIKRADAAPTVISIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDYSTYSMSSTLTTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRECE

Cadeia pesada

QVTLKESGPGILKPSQTLSTCSLSGFSRLRT**KGLGVGWD**RQPSGKGLEWL**VAWDAKRTYIRSEK**LKSQLTISKDTSRNQVFLKITSVDTADTATYYCVR**FRVDQRAYWV**FAYWGQGLTVTVSAAKTAPSVYPLAPVSVTLGCLVKGYFPEPVTLTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVTSSTWPSQITCNVAHPASSTKVDDKIEPR

Resultado 2

Antígeno

AKFRHD semelhança 38% e 64% com cristal 2r0z; energia do complexo -46,2 kJ/mol

Cadeia leve

DVLMQTPLSLPVSLGDQASISCRSSQSI**HDFCPPSWAD**WYLQKPGQSPKLL**WPXGGTYD**SGVPDRFSGSGSGTDFTLKISRVEAEDLGVYYC**AYDGCKKLWS**GAGTKLEIKRADAAPTVISIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDYSTYSMSSTLTTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRECE

Cadeia pesada

QVTLKESGPGILKPSQTLSTCSLSGFSRLRT**RSKNVTGG**RQPSGKGLEWL**GTIKDDFGTDERGS**LKSQLTISKDTSRNQVFLKITSVDTADTATYYCVR**AYYELSGSTC**FAYWGQGLTVTVSAAKTAPSVYPLAPVSVTLGCLVKGYFPEPVTLTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVTSSTWPSQITCNVAHPASSTKVDDKIEPR

Resultado 3

Antígeno

AKFRHD semelhança 62% e 37% com cristal 2r0z; energia do complexo -38,7 kJ/mol

Cadeia leve

DVLMQTPLSLPVSLGDQASISCRSSQSI**NSSTGWHQIR**WYLQKPGQSPKLL**KDFYCWLE**SGVPDRFSGSGSGTDFTLKISRVEAEDLGVYYC**TGSWGSTKSG**GAGTKLEIKRADAAPTVISIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDYSTYSMSSTLTTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRECE

Cadeia pesada

QVTLKESGPGILKPSQTLSTCSLSGFSRLRT**KDGDGTGYT**RQPSGKGLEWL**CCGDEKENGAPTTEL**LKSQLTISKDTSRNQVFLKITSVDTADTATYYCVR**GTPQQYTSWT**FAYWGQGLTVTVSAAKTAPSVYPLAPVSVTLGCLVKGYFPEPVTLTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVTSSTWPSQITCNVAHPASSTKVDDKIEPR

Cristalografia 2r0z

Cadeia leve

DVLMQTPLSLPVSLGDQASISCRSSQSI**VHSNGNTYLE**WYLQKPGQSPKLL**IYKVSNRF**SGVPDRFSGSGSGTDFTLKISRVEAEDLGVYYC**FQGSHPVPLTF**GAGTKLEIKRADAAPTVISIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVLNSWTDQDSKDYSTYSMSSTLTTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRECE

Cadeia pesada

QVTLKESGPGILKPSQTLSTCSLSGFSRLST**SGMGVGVWIR**QPSGKGLEWL**AHIWWDDDRSYNPS**LKSQLTISKDTSRNQVFLRITSVDTADTATYYCVR**RAHTTVLGDW**FAYWGQGLTVTVSAAKTAPSVYPLAPVCGDITGSSVTLGCLVKGYFPEPVTLTWNSGSLSSGVHTFPAVLQSDLYTLSSSVTVTSSTWPSQITCNVAHPASSTKVDDKIEPR

A média da reprodução dos três protocolos é descrita na FIG. 18.

Figura 18 – Estrutura prevista do complexo antígeno anticorpo com epítomos AKFRHD. Apresentando $(55 \pm 14) \%$ e $(55 \pm 14) \%$ de identidade nas cadeias leve e pesada, respectivamente, comparada com a estrutura cristalográfica real, que possui ΔG calculado $-48,7 \text{ kJ/mol}$.

Fonte: Elaborada pelo autor.

Da mesma forma a estrutura o protocolo foi reproduzido três vezes para as estruturas HQLDPAFG e EPANPSEKNSPSTQYC, com suas respectivas estruturas e energias de ligação comparadas com as estruturas cristalográficas:

Resultado 1

Antígeno

HQLDPAFG semelhança 65% e 61% com cristal 1kc5; energia do complexo $-36,7 \text{ kJ/mol}$

Cadeia leve

DIVLTQSPKSMMSVGEKVTLCKASENV**IVYNWRYRRR**PEQPPALLIYGAS**GRYPVRPF**RFTGSGSATDFTLTISVQAED
LADYHCGQSYS**KKRTFKGKTG**LELKRADAAPTVISIFPPSEQLTSGGASVVCFLNFPKDINVKWKIDGSRQNGVANSWT
AQDSKDYSTYMSSTLTLTCKEYERHNSYTCEATHKTSTSPVVKSFNRNEC

Cadeia pesada

QVKLQQSGPGLVKPSQSLTCTVTGYSITS**TDARKWRV**QFPGNKLEWMA**RIFANS**GFT**KKPYR**KSRISITRDTSKNQFF
LQLNSVTTEDTAIYYCAR**GQRTKDWXTG**TTLTVSAAATPPSVYPLAPGSATAAASMTVLGCLVKGYFPEPVTVTWNSGALS
SGVHTFPAVLQSDLYTLSSSVTPSSPWPSETVTCNVAPASSTKVDKIVPRD

Resultado 2

Antígeno

HQLDPAFG semelhança 28% e 51% com cristal 1kc5; energia do complexo $-20,7 \text{ kJ/mol}$

Cadeia leve

DIVLTQSPKSMMSVGEVVTLSCKASENV**NKKS**NMARYWPEQPPALLIYGAS**GQRHEKA**FRTGSGSATDFTLTISVQA
EDLADYXCGQSYS**RKRGVKGRQG**LELKRADAAPTVISIFPPSEQLTSGGASVVCFLNFPKDINVKWKIDGSRQNGVANS
WTAQDSKDYSTYMSSTLTLTCKEYERXNSYTCEATXKTSTSPVVKSFNRNEC

Cadeia pesada

QVKLQQSGPGLVKPSQSLTCTVTGYSITS**ERGAWYRR**QFPGNKLEWMA**MIYPGYRTYFHRRY**KSRISITRDTSKNQF
 FLQLNSVTTEDTAIYYCAR**GQRSDHWHTG**TTLTVSAAATPPSVYPLAPGSATAAASMTLGLVKGYFPEPVTVTWNSGAL
 SSGVXTFPAVLQSDLYTLSSSVTPSSPWPSETVTCNVAXPASSTKVDKKIVPRD

Resultado 3

Antígeno

HQLDPAFG semelhança 28% e 63% com cristal 1kc5; energia do complexo -24,6 kJ/mol

Cadeia leve

DIVLTQSPKMSMSVGEVVTLSCKASENV**NKKSNAQYW**PEQPALLIYGAS**GQRHEPAF**RFTGSGSATDFTLTISVQA
 EDLADYXCGQSYS**RKTGVKGRQG**LELKRADAAPTVSIFPPSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVANS
 WTAQDSKDYSTYMSSTLTTLTKDEYERXNSYTCEATXKTSTSPVVKSFNRNEC

Cadeia pesada

QVKLQQSGPGLVKPSQSLTCTVTGYSITS**ERAKHWRR**QFPGNKLEWMA**MKYYSYRTYFHRRY**KSRISITRDTSKNQ
 FFLQLNSVTTEDTAIYYCAR**GGTSRHWHTG**TTLTVSAAATPPSVYPLAPGSATAAASMTLGLVKGYFPEPVTVTWNSGAL
 LSSGVXTFPAVLQSDLYTLSSSVTPSSPWPSETVTCNVAXPASSTKVDKKIVPRD

Cristalografia 1kc5

Cadeia leve

DIVLTQSPKMSMSVGEKVTLSCKASENV**DTYVSWYQQR**PEQPALLIYGAS**NRYTGVPD**RFTGSGSATDFTLTISVQAE
 DLADYHCGQSYS**YPLTFGGGTK**LELKRADAAPTVSIFPPSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSERQNGVANSW
 TAQDSKDYSTYMSSTLTTLTKDEYERHNSYTCEATHKTSTSPVVKSFNRNEC

Cadeia pesada

QVKLQQSGPGLVKPSQSLTCTVTGYSITS**DYAWNWR**QFPGNKLEWMA**YISYSGSTTYNPSL**KSRISITRDTSKNQFFL
 QLNSVTTEDTAIYYCAR**GGTGFDYWGAG**TTLTVSAAATPPSVYPLAPGSATAAASMTLGLVKGYFPEPVTVTWNSGAL
 SSGVHTFPAVLQSDLYTLSSSVTPSSPWPSETVTCNVAXPASSTKVDKKIVPRD

Resultado 1

Antígeno

EPANPSEKNSPSTQYC semelhança 25% e 81% com cristal 3bky; energia do complexo -
 38,1 kJ/mol

Cadeia leve

QIVLSQSPAILSASPGEKVTMTCRASSSV**RYAEHYDQAP**GSSSPKPWIYAPSN**VAKGRRAR**FSGSGSGTSYSLTISRVEAED
 AATYYCQQWSFN**DPTVGAVTGLE**LKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESV
 EQDSKDYSTYLSSTLTLSKADYEKXKVYACEVTXQGLSSPVTKSFNRGEC

Cadeia pesada

QAYLQQSGAELVRPGASVKMSCKASGYTFTS**RNVPWYKP**TPRQGLEWIGA**HMRGHGKTSWAKKW**KGKATLTVDKS
 SSTAYMQLSSLTSEDSAVYFCARV**VYLNNKRWYL**DVWGTGTTVTVAASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPE
 PVTVSWNSGALTSGVXTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNXXKPSNTKVDKKVEPKSCD

Resultado 2

Antígeno

EPANPSEKNSPSTQYC semelhança 64% e 78% com cristal 3bky; energia do complexo -
 30,4 kJ/mol

Cadeia leve

QIVLSQSPAILSASPGEKVTMTCRASSSV**SYQRWREWKP**GSSSPKPWIYAPSN**GARRPQAF**FSGSGSGTSYSLTISRVEAED
 AATYYCQQWSFN**GPTGAKGAAG**ELKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESV
 TEQDSKDYSTYLSSTLTLSKADYEKXKVYACEVTXQGLSSPVTKSFNRGEC

Cadeia pesada

QAYLQQSGAELVRPGASVKMSCKASGYTFTS**NRMXWGTY**TPRQGLEWIGA**FYGKYGGWSYQFR**KGKATLTVDKSS
 STAYMQLSSLTSEDSAVYFCARV**GYEANYDWYG**DVWGTGTTVTVSAASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEP
 VTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNKKPSNTKVDKKVEPKSCD

Resultado 3***Antígeno***

EPANPSEKNSPSTQYC semelhança 70% e 71% com cristal 3bky; energia do complexo -
 29,4 kJ/mol

Cadeia leve

QIVLSQSPAILSASPGEKVTMTCRASSSV**IWN**DW**NKYQG**GSSPKPWIYAPSN**GNXYKXIF**FSGSGSGTSYSLTISRVEAEDA
 ATYYCQQWSFN**WKSGAKGGS**GELKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVT
 EQDSKDYSLSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Cadeia pesada

QAYLQQSGAELVRPGASVKMSCKASGYTFTS**RGYGWQIV**TPRQGLEWIGA**YRVVQRPKFSRXRS**KGKATLTVDKSSST
 AYMQSSLTSEDSAVYFCARV**GDELPR**SDEGDVWGTGTTVTVSAASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVT
 VSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD

Cristalografia 3bky***Cadeia leve***

QIVLSQSPAILSASPGEKVTMTCRASSSV**SYM**H**WYQQK**PSSPKPWIYAPSN**LASGV**PARFSGSGSGTSYSLTISRVEAEDA
 ATYYCQQWSFN**PPTFGAGTK**LELKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTE
 QDSKDYSLSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

Cadeia pesada

QAYLQQSGAELVRPGASVKMSCKASGYTFTS**YNLH**W**VYQ**TPRQGLEWIGA**IYPGNGDTSYNQKF**KGKATLT
 VDKSSSTAYMQLSSLTSEDSAVYFCARV**VYY**S**NSYWYF**DVWGTGTTVTVSAASTKGPSVFPLAPSSKSTSGGTAALGCLVKD
 YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD

As médias das reproduções dos três protocolos para a previsão das duas estruturas

HQLDPAFG e EPANPSEKNSPSTQYC, são descritas nas FIG. 19 e FIG. 20.

Figura 19 – Estrutura prevista do complexo antígeno anticorpo com epítopo HQLDPAFG. Apresentando $(40 \pm 21) \%$ e $(58 \pm 6) \%$ de identidade nas cadeias leve e pesada, respectivamente, comparada com a estrutura cristalográfica real, que possui ΔG calculado $-20,4 \text{ kJ/mol}$.

Fonte: Elaborada pelo autor.

Figura 20 – Estrutura prevista do complexo antígeno anticorpo com epítopo EPANPSEKNSPSTQYC. Apresentando $(53 \pm 24) \%$ e $(77 \pm 5) \%$ de identidade nas cadeias leve e pesada, respectivamente, comparada com a estrutura cristalográfica real, que possui ΔG calculado $-44,4 \text{ kJ/mol}$.

Fonte: Elaborada pelo autor.

Os resíduos de cada alça foram discriminados de acordo com a sequência prevista pela metodologia e confrontada com os resultados cristalográficos, sendo as alças de interação modeladas destacadas em **negrito**. Embora alguns resíduos não sejam os mesmos da previsão, suas propriedades físico-químicas, semelhança nas cadeias (de acordo com o servidor PIGS (39)) e funcionalidade demonstrada posteriormente no item 5.3 validam o modelo. Cada sequência do anticorpo possui uma média de $(48 \pm 18) \%$ e $(65 \pm 14) \%$ de identidade nas cadeias leve e pesada, respectivamente, com as estruturas 2r0z, 1kc5 e 3bkya qual o protocolo previu, de acordo com o servidor PIGS.(39) Vale destacar que o modelo desenvolvido e

testado foi criado a partir de estruturas de anticorpos presentes no SNC abrindo assim uma gama especial de possibilidades para aplicações e estudo de doenças do SNC em geral.

5.2.2 Aplicação do protocolo e a obtenção do anticorpo anti-AQP4

Com a finalidade de obter comparações futuras para diferentes processos de funcionalização, a sequência do peptídeo alvo foi modelada de duas maneiras: (a) utilizando o peptídeo retirado da estrutura nativa da AQP4, da qual a conformação descrita na cristalografia estava preservada, denominado peptídeo da AQP4; e (b) utilizando o peptídeo linear modelado (síntese *in silico*), referido como peptídeo linear (FIG. 21), o qual representa o peptídeo sintetizado experimentalmente empregado nos estudos de AFM.

Figura 21 – MMC dos epítomos da proteína AQP4. Peptídeo linear sintetizado *in silico*.
Fonte: Modelagem a partir do Arquivo PDB 3GD8.(124)

Este estudo fornece a base para as simulações de SMD em diferentes frentes, tanto para obtenção da força de adesão teórica entre a ponta e o substrato funcionalizado com

peptídeos com estrutura nativa e preservada, quanto para peptídeos sintetizados experimentalmente e disponíveis comercialmente.

O protocolo desenvolvido foi agora aplicado na previsão da estrutura e das propriedades do anticorpo anti-AQP4. A estrutura prevista mostrada na FIG. 22A indica a formação do complexo antígeno (sequência GTEKPLPVDM), representado na cor verde, anticorpo anti-AQP4 representado na cor branca. Para estabilização do complexo e da estrutura foram realizados cálculos de DM por 16 nm e analisada sua estabilidade no RMSD da FIG 22C.(167)

Cadeia leve

DVLMQTPLSLPVSLGDAQISCRSSQSIRRRKIRYFNEWYLQKPGQSPKLLAVPIIGAVSGVPDRFSGSGSGTDFTLKISRVEAEDLG
VYYCAVFKKQAVAVGAGTKLEIKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSRQNGVLNSWTDQDSK
DSTYSMSSTLTTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNE

Cadeia pesada

QVTLKESGPGILKPSQTLSTCSLSGFSRLRTREPDWPKERQPSGKGLEWLKMCKSQPRWDTWRILKSQLTISKDTSRNQVFLKITSV
DTADTATYYCVRDFMPEDMDAVFAYWGQGLTVTSAAKTTAPSVYPLAPVSVTLGCLVKGYFPEPVTLTWNSGSLSSGVHTFPA
VLQSDLYTLSSSVTVTSSTWPSQSITCNVAHPASSTKVDKKIEPR

Antígeno

GTEKPLPVDM

A FIG. 22B destaca o potencial eletrostático das estruturas, destacando prevalência de cargas positivas na região ativa das alças de interação, resultado que acompanha o caráter elétrico predominantemente negativo do epítipo.

Figura 22 – (A) Estrutura prevista do complexo antígeno anticorpo anti-AQP4. (B) Potencial eletrostático do epítipo GTEKPLPVD e do anticorpo anti-AQP4. (C) RMSD do complexo formado.

Nota: k_B é a constante de Boltzmann, T a temperatura e e a carga elementar.

Fonte: Elaborada pelo autor.

A estrutura 3D do anticorpo anti-AQP4 obtida abre possibilidades para as inúmeras ferramentas de MMC, como cálculos de homologia, novos cálculos de docking, DM, SMD, simulações de Monte Carlo, entre outras(168). A partir da estrutura do anticorpo, estudos como a interação com outros epítopos e sequências antigênicas podem ser realizados, além de novos estudos do mecanismo de ação da NMO e outras doenças desmielinizantes.

5.3 O nanoimunossensor desenvolvido para detectar o biomarcador da NMO

5.3.1 Cálculos de SMD

Dispondo da estrutura do complexo antígeno anticorpo anti-AQP4, foi possível realizar as simulações de SMD em parceria com a aluna de doutorado Jéssica Cristiane Magalhães Ierich (Processo FAPESP 2014/12082-4). Tal simulação ocorreu em um tempo total de 3ns. Neste contexto, o átomo $C\alpha$ dos resíduos ALA123 (cadeia leve) e ARG113 (cadeia pesada) foram considerados fixos na estrutura do anticorpo. A força externa foi aplicada no átomo $C\alpha$ do resíduo MET10 do peptídeo imunogênico da AQP4 (FIG. 23). A simulação foi realizada usando uma constante de mola (k) de $0,2 \text{ kcal (mol } \text{\AA}^2)^{-1}$ em uma velocidade constante de $0,00005 \text{ \AA timestep}^{-1}$ ($0,025 \text{ \AA ps}^{-1}$). O valor de k foi escolhido visando aproximar as simulações computacionais dos experimentos de AFM conduzidos pela aluna Ariana de Souza Moraes (Processo FAPESP 2015/05283-6).

Figura 23 - Modelo esquemático da aplicação de uma força externa para remoção do peptídeo imunogênico do complexo da porção Fv do anticorpo específico por SMD.

Fonte: Elaborada pela aluna de doutorado Jéssica Cristiane Magalhães Ierich (Processo FAPESP 2014/12082-4).

Após a realização da simulação descrita, a dinâmica de desligamento do peptídeo do complexo com o anticorpo, foi possível obter a força de interação ilustrada na FIG. 24. Na situação a força de desprendimento foi 224 pN. Desta forma, valores similares de força foram observados nos resultados experimentais do item 5.3.2.

Figura 24 – Curva de força de SMD obtida para a interação peptídeo imunogênico da AQP4 e anticorpo modelado, com força de desprendimento foi 224 pN.

Fonte: Elaborada pela aluna de doutorado Jéssica Cristiane Magalhães Ierich (Processo FAPESP 2014/12082-4).

Desta forma o nanoimunossensor foi desenvolvido teoricamente e a metodologia de comparação, entre resultados teóricos e experimentais, está de acordo com importantes trabalhos realizados recentemente. (4,11-12,158,169-170)

5.3.2 Comprovação experimental do nanoimunossensor para NMO

As medidas experimentais descritas neste item foram realizadas pela aluna de doutorado Ariana de Souza Moraes (Projeto FAPESP nº 2015/05283-6). Para a interpretação e tratamento dos dados experimentais foi utilizada a distribuição gama, tomando como principal parâmetro de comparação a moda (medida mais frequente) dos valores medidos. Os soros pacientes NMO positivos foram obtidos pela Prof^a Dra. Doralina G. Brum Souza. Foi utilizado como elemento sensor o peptídeo específico GTEKPLPVDVM com o intuito de testar a especificidade na detecção do biomarcador da NMO (anti-AQP4). Tal peptídeo foi sintetizado

no laboratório de Química dos Peptídeos, pela Prof^a Dra. Maria Teresa Machini, do Departamento de Bioquímica do Instituto de Química da Universidade de São Paulo (USP).

Os ensaios constataram uma média de força de despreendimento de (296 ± 106) pN, como mostrado na FIG. 25.

Figura25 – Moda dos valores de força de adesão para o soro purificado de pacientes soro positivos para anti-AQP4.

Fonte: Elaborada pelo autor.

Desta forma, a sequência peptídica GTEKPLPVDM utilizada nos experimentos se mostrou eficiente na detecção do anticorpo anti-AQP4. Para uma ponta com raio nominal 7nm (utilizada nos experimentos) e de acordo com as dimensões calculadas no item 5.1.2, espera-se que a interação entre a ponta funcionalizada e a amostra seja entre uma ou duas

porções Fab do anticorpo, ou seja, a força de adesão entre a ponta e a amostra encontrada experimentalmente está de acordo com o resultado teórico de 224 pN de força de despreendimento para uma porção Fab.

6 CONSIDERAÇÕES E PERSPECTIVAS

Atendendo aos objetivos propostos, foi elaborado um protocolo computacional baseado em técnicas de MMC, para modelar as nanossuperfícies de interação envolvidas nos experimentos de AFM, pontas e substratos recobertos. Para as moléculas de antígenos e anticorpos foi desenvolvido de um protocolo inovador para modelar a estrutura de anticorpos, tendo como ponto de partida apenas estrutura do antígeno. O protocolo consistiu em utilizar a mutagênese computacional dos resíduos da região Fab da estrutura de IgG genérica, a fim de se obter o modelo mais fidedigno quando comparado aos resultados experimentais, com base na energia de ligação mais favorável e dinâmica estrutural dos complexos formados. Tal modelo foi resultado de vários outros modelos iniciais que mostraram pontos fundamentais na construção da nova metodologia. A inovadora metodologia desenvolvida neste trabalho foi capaz de prever com $(48 \pm 18) \%$ e $(65 \pm 14) \%$ de identidade das cadeias leve e pesada, respectivamente, da estruturas dos anticorpos reais presentes na literatura. A metodologia foi aplicada no recobrimento das nanossuperfícies por anticorpos e na previsão de estruturas ainda não descritas da NMO, como o anticorpo anti-AQP4, tendo grande importância no direcionamento dos resultados experimentais e envolvendo o biomarcador anti-AQP4.

Experimentalmente, foram realizadas medidas de AFM, com imobilização orientada das IgGs provenientes de soros de pacientes soropositivos para NMO e soropositivos e soronegativos para a EM. Para o tratamento dos dados experimentais foi desenvolvido um modelo estatístico, baseado na distribuição gama, que difere os pacientes soropositivos e soronegativos para os biomarcadores relacionados à EM, bem como a especificidade dos peptídeos antigênicos da AQP4 que interagem com o biomarcador da doença. O modelo estatístico forneceu a comprovação do funcionamento do nanoimunossensor, fornecendo assim, um novo e mais preciso meio para detecção de biomarcadores e até mesmo diagnóstico de doenças desmielinizantes. Desta forma a interpretação e tratamento dos dados

experimentais feitos com a distribuição gama se provou o método mais eficiente. Espera-se que as curvas teóricas levantadas sejam aperfeiçoadas no futuro à medida que os experimentos forem sendo realizados com um número maior de pacientes e que estas se tornem um padrão para diagnóstico de desmielinização e auxilie no diagnóstico das doenças autoimunes como a EM e a NMO.

Por fim, o uso das técnicas computacionais forneceu os modelos descritivos dos nanossistemas envolvendo doenças autoimunes e interações antígeno anticorpo, o que permite traçar as perspectivas futuras:

- i. Efetuar maior número de reproduções do protocolo desenvolvido de modelagem da estrutura de anticorpos e consequentemente o desenvolvimento de um programa computacional em linguagem específica;
- ii. Aplicação do protocolo no estudo de outras doenças desmielinizantes;
- iii. Correlacionar forças de adesão teóricas com resultados experimentais relacionados a outras doenças desmielinizantes.

7 PRODUÇÕES DECORRENTES DESTE TRABALHO

Artigos publicados em periódicos:

AMARANTE, A. M.; OLIVEIRA, G. S.; BUENO, C. C.; CUNHA, R. A.; FREITAS, L. C. G.; FRANCA, E. F.; OLIVEIRA, O. N.; LEITE, F. L.. A new rigid enzyme model and its application in nanobiosensor development. **Journal of Molecular Graphics & Modelling**, v.53, p. 100–104, 2014. doi:10.1016/j.jmgm.2014.07.009.

IERICH, JÉSSICA C. M.; OLIVEIRA, GUEDMILLER S.; VIG, ANA C. A.; **AMARANTE, ADRIANO M.**; FRANCA, EDUARDO F.; LEITE, FÁBIO L.; MASCARENHAS, YVONNE P.. A Computational Protein Structure Refinement of the Yeast Acetohydroxyacid Synthase. **Journal of the Brazilian Chemical Society** (Impresso), v. 26, p. 1702-1709, 2015.

Capítulo de livro publicado:

AMARANTE, A. M.; OLIVEIRA, G. S.; IERICH, J. C. M.; CUNHA, R. A.; FREITAS, L. C. G.; FRANCA, E. F.; LEITE, F. L.. Modelagem Molecular Aplicada a Nanobiosistemas. In: Alessandra Luzia Da Róz, Fabio de Lima Leite, Marystela Ferreira & Osvaldo Novais de Oliveira Jr. (Org.). Nanociência & Nanotecnologia: princípios e aplicações (Grandes Áreas da Nanociência e suas Aplicações). São Paulo: Elsevier, 2014. v. 2, p. 40-60.

REFERÊNCIAS

- 1 RUMAYOR, V. G.; IGLESIAS, O.R. CABEZAS, L.G. *Aplicaciones de biosensores en la industria agroalimentaria*. 2005. Disponível em: <https://www.madrimasd.org/uploads/informacionidi/biblioteca/publicacion/doc/VT/VT1_Aplicaciones_de_biosensores_en_la_industria_agroalimentaria.pdf>. Acesso em: 23 jan. 2018.
- 2 DEDA, D. K. et al. The use of functionalized AFM tips as molecular sensors in the detection of pesticides. *Materials Research*, v. 16, n. 3, 2013. doi:10.1590/S1516-14392013005000043.
- 3 FRANCA, E. F. et al. Designing an enzyme-based nanobiosensor using molecular modeling techniques. *Physical Chemistry Chemical Physics*, v. 13, p. 8894–8899, 2011. doi:10.1039/C1CP20393B.
- 4 OLIVEIRA, G. S. et al. Molecular modeling of enzyme attachment on AFM probes. *Journal of Molecular Graphics and Modelling*, v. 45, p. 128–136, 2013. doi:10.1016/j.jmgm.2013.08.007.
- 5 SILVA, A. et al. Nanobiosensors based on chemically modified AFM probes: a useful tool for metsulfuron-methyl detection. *Sensors*, v. 13 p. 1477–1489, 2013. doi:10.3390/s130201477.
- 6 STEFFENS, C. et al. Herrmann, atomic force microscopy as a tool applied to nano/biosensors. *Sensors*, v. 12, p. 8278–8300, 2012. doi:10.3390/s120608278.
- 7 SMITH, D. P. E. Limits of force microscopy. *Review of Scientific Instruments*, v. 66, p. 3191–3195, 1995. doi:10.1063/1.1145550.
- 8 SMITH, D.A. et al. Chemical force microscopy: applications in surface characterisation of natural hydroxyapatite. *Analytica Chimica Acta*, v. 479, p. 39–57, 2003. doi:10.1016/S0003-2670(02)01374-0.
- 9 SMITH, A. E. et al. Fate of Diclofop-methyl after application to a wheat field 1. *Journal of Environmental Quality*, v. 15, p. 234–238, 1986. doi:10.2134/jeq1986.00472425001500030006x.
- 10 ZHANG, H.; TWEEL, B.; TONG, L. Molecular basis for the inhibition of the carboxyltransferase domain of acetyl-coenzyme-A carboxylase by haloxyfop and diclofop. *Proceedings of the National Academy of Sciences of the United States of America*, v. 101, p. 5910–5915, 2004. doi:10.1073/pnas.0400891101.
- 11 AMARANTE, A. M. et al. Modeling the coverage of an AFM tip by enzymes and its application in nanobiosensors. *Journal of Molecular Graphics & Modelling*, v. 53, p. 100–104, 2014. doi:10.1016/j.jmgm.2014.07.009.
- 12 BUENO, C. C. et al. Nanobiosensor for diclofop detection based on chemically modified AFM probes. *IEEE Sensors Journal*, v. 14, p. 1467–1475, 2014. doi:10.1109/JSEN.2014.2301997.
- 13 LEITE, F. L. et al. Theoretical models for surface forces and adhesion and their measurement using atomic force microscopy. *International Journal of Molecular Sciences*, v. 13, p. 12773–12856, 2012. doi:10.3390/ijms131012773.

- 14 FRANCA, E. F.; AMARANTE, A. M.; LEITE, F. L. Introduction to atomic force microscopy simulation. In: MÉNDEZ-VILAZ, A.; DÍAZ, J. (Ed.). *Microscopy: science, technology, applications and education*. Badajoz: Formatex Research Center, 2010. p. 1338–1349.
- 15 LEITE, F. L. et al. The atomic force spectroscopy as a tool to investigate surface forces: basic principles and applications. In: MÉNDEZ-VILAZ, A.; DÍAZ, J. (Ed.). *Modern research and educational topics in microscopy: applications in physical/chemical sciences, techniques*. Badajoz: Formatex Research Center, 2007. p. 747–757.
- 16 LEITE, F. L. et al. Nanoneurobiophysics: new challenges for diagnosis and therapy of neurologic disorders. *Nanomedicine*, v. 10, p. 3417–3419, 2015. doi:10.2217/nnm.15.164.
- 17 HINSON, S. R. et al. Prediction of neuromyelitis optica attack severity by quantitation of complement-mediated injury to aquaporin-4-expressing cells. *Archives of Neurology*, v. 66, p. 1164–1167, 2009. doi:10.1001/archneurol.2009.188.
- 18 WATERS, P. J. et al. Serologic diagnosis of NMO: a multicenter comparison of aquaporin-4-IgG assays. *Neurology*, v. 78, p. 665–671, 2012. doi:10.1212/WNL.0b013e318248dec1.
- 19 AKMAN-DEMIR, G. et al. Eraksoy, prognostic implications of aquaporin-4 antibody status in neuromyelitis optica patients. *Journal of Neurology*, v. 258, p. 464–470, 2011. doi:10.1007/s00415-010-5780-4.
- 20 WINGERCHUK, D. M. et al. The clinical course of neuromyelitis optica (Devic's syndrome). *Neurology*, v. 53, p. 1107–1107, 1999. doi:10.1212/WNL.53.5.1107.
- 21 MCDONALD, W. I. et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Annals of Neurology*, v. 50, n. 1, p. 121–127, 2001.
- 22 WINGERCHUK, D. M. et al. The spectrum of neuromyelitis optica. *Lancet Neurology*, v. 6, p. 805–815, 2007. doi:10.1016/S1474-4422(07)70216-8.
- 23 WINGERCHUK, D. M. et al. Revised diagnostic criteria for neuromyelitis optica. *Neurology*, v. 66, p. 1485–1489, 2006. doi:10.1212/01.wnl.0000216139.44259.74.
- 24 LENNON, V. A. et al. A serum autoantibody marker of neuromyelitis optica: distinction from multiple sclerosis. *Lancet*, v. 364, p. 2106–2112, 2004. doi:10.1016/S0140-6736(04)17551-X.
- 25 WATERS, P. et al. Aquaporin-4 antibodies in neuromyelitis optica and longitudinally extensive transverse myelitis. *Archives of Neurology*, v. 65, p. 913–919, 2008. doi:10.1001/archneur.65.7.913.
- 26 JARIUS, S. et al. NMO-IgG in the diagnosis of neuromyelitis optica. *Neurology*, v. 68, p. 1076–1077, 2007. doi:10.1212/01.wnl.0000256822.01222.bd.
- 27 BERMAN, H. M. et al. The protein data bank. *Nucleic Acids Research*, v. 28, p. 235–242, 2000. doi:10.1093/nar/28.1.235.

- 28 BENSON, D. A. GenBank. *Nucleic Acids Research*, v. 41, p. D36–42, 2013. doi:10.1093/nar/gks1195.
- 29 WU, C. H. et al. The universal protein resource (UniProt): an expanding universe of protein information. *Nucleic Acids Research*, v. 34 p. D187–D191, 2006. doi:10.1093/nar/gkj161.
- 30 JOHNSON, M. NCBI BLAST: a better web interface. *Nucleic Acids Research*, v. 36, p. W5–9, 2008. doi: 10.1093/nar/gkn201.
- 31 BLECH, M. et al. One target—two different binding modes: structural insights into Gevokizumab and Canakinumab interactions to interleukin-1 β . *Journal of Molecular Biology*, v. 425, p. 94–111, 2013. doi:10.1016/j.jmb.2012.09.021.
- 32 ULTSCH, M. et al. Structural basis of signaling blockade by anti-IL-13 antibody lebrikizumab. *Journal of Molecular Biology*, v. 425, p. 1330–1339, 2013. doi:10.1016/j.jmb.2013.01.024.
- 33 KURODA, D. et al. Computer-aided antibody design. *Protein Engineering, Design and Selection*, v. 25, n. 10, p. 507–522, 2012.
- 34 SEVY, A. M.; MEILER, J. Antibodies: computer-aided prediction of structure and design of function. *Microbiology Spectrum*, v. 2, n. 6, 2014.
- 35 LOYAU, J. et al. Robust antibody-antigen complexes prediction generated by combining sequence analyses, mutagenesis, in vitro evolution, x-ray crystallography and in silico docking. *Journal of Molecular Biology*, v. 427, n. 16, p. 2647–2662, 2015.
- 36 PATTABHI, V.; GAUTHAM, N. *Biophysics*. New York: Springer, 2002.
- 37 BORDOLI, L.; SCHWEDE, T. Automated protein structure modeling with SWISS-MODEL workspace and the protein model portal. In: ORRY, A. J. W.; ABAGYAN, R. *Homology modelling: methods and protocols*. New York: Humana Press, 2012. p. 107–136.
- 38 ARNOLD, K. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, v. 22, n. 2, p. 195–201, 2006.
- 39 CHOI, Y.; DEANE, C. M. Predicting antibody complementarity determining region structures without classification. *Molecular BioSystems*, v. 7, n. 12, p. 3327–3334, 2011.
- 40 MARKS, C.; DEANE, C. M. Antibody H3 structure prediction. *Computational and Structural Biotechnology Journal*, v. 15, p. 222–231, 2017.
- 41 ZHAO, J. et al. In silico methods in antibody design. *Antibodies*, v. 7, n. 3, p. 22, 2018.
- 42 MARCATILI, P. et al. Antibody modeling using the prediction of immunoglobulin structure (PIGS) web server. *Nature Protocols*, v. 9, n. 12, p. 2771–2783, 2014.
- 43 LEPORE, R. et al. PIGSPro: prediction of immunoglobulin structures v2. *Nucleic Acids Research*, v. 45, n. W1, p. W17–W23, 2017.

44 *ROSIE*: the Rosetta online server that includes everyone. Disponível em: <<http://antibody.graylab.jhu.edu/>>. Acesso em: 21 set. 2018.

45 KEMMISH, H.; FASNACHT, M.; YAN, L. Fully automated antibody structure prediction using BIOVIA tools: validation study. *PLoS ONE*, v. 12, p. e0177923, 2017. doi:10.1371/journal.pone.0177923.

46 DUNBAR, J. et al. SAbPred: a structure-based antibody prediction server. *Nucleic Acids Research*, v. 44, n. W1, p. W474-478, 2016.

47 GARCIA, P. S. et al. A Nanobiosensor based on 4-Hydroxyphenylpyruvate dioxygenase enzyme for mesotrione detection. *IEEE Sensors Journal*, v. 15, n. 4, p. 2106, 2014.

48 LEITE, F. L.; HERRMANN, P. S. P. Application of atomic force spectroscopy (AFS) to studies of adhesion phenomena: a review. *Journal of Adhesion Science and Technology*, v. 19, n. 3–5, p. 365–405, 2005.

49 SANTOLIN, M.A. *Nanolitografia do silício utilizando o microscópio de força atômica*. 2009. 97 p. Dissertação (Magister Scientiae) -Universidade Federal de Viçosa, Viçosa, 2009.

50 MIZES, H. A. et al. Submicron probe of polymer adhesion with atomic force microscopy: Dependence on topography and material inhomogeneities. *Applied Physics Letters*, v. 59, n. 22, p. 2901–2903, 1991.

51 BUTT, H.-J. Force measurements with the atomic force microscope: technique, interpretation and applications. *Surface Science Reports*, v. 59, p. 1–152, 2005.

52 CAPPELLA, B.; DIETLER, G. Force-distance curves by atomic force microscopy. *Surface Science Reports*, v. 34, n. 1–3, p. 1–104, 1999.

53 FLORIN, E. L.; MOY, V. T.; GAUB, H. E. Adhesion forces between individual ligand-receptor pairs. *Science*, v. 264, n. 5157, p. 415–417, 1994.

54 CAPPELLA, B. et al. Force-distance curves by AFM. a powerful technique for studying surface interactions. *IEEE Engineering in Medicine and Biology Magazine: the quarterly magazine of the engineering in medicine & biology society*, v. 16, n. 2, p. 58–65, 1997.

55 DUWEZ, A.-S.; NYSTEN, B. Mapping aging effects on polymer surfaces: specific detection of additives by chemical force microscopy. *Langmuir*, v. 17, n. 26, p. 8287–8292, 2001.

56 SATO, F. et al. A study of topographic effects on chemical force microscopy using adhesive force mapping. *Ultramicroscopy*, v. 97, n. 1–4, p. 303–314, 2003.

57 WEETALL, H. H. Covalent coupling methods for inorganic support materials. *Methods in Enzymology*, v. 44, p. 134–148, 1976.

58 MARQUES, P. R. B. DE O.; YAMANAKA, H. Biosensors based on the enzymatic inhibition process. *Química Nova*, v. 31, n. 7, p. 1791–1799, 2008.

- 59 AMARANTE JÚNIOR, O. P. DE et al. Breve revisão de métodos de determinação de resíduos do herbicida ácido 2,4-diclorofenoxiacético (2,4-D). *Química Nova*, v. 26, n. 2, p. 223–229, 2003.
- 60 ETCHEGARAY, A.; BUENO, C. DE C.; TESCHKE, O. Identification of microcystin LR at the molecular level using atomic force microscopy. *Química Nova*, v. 33, n. 9, p. 1843–1848, 2010.
- 61 FERREIRA, A. A. P.; YAMANAKA, H. Atomic force microscopy applied to immunoassays. *Química Nova*, v. 29, n. 1, p. 137–142, 2006.
- 62 KAUR, J. et al. Atomic force spectroscopy-based study of antibody pesticide interactions for characterization of immunosensor surface. *Biosensors & Bioelectronics*, v. 20, n. 2, p. 284–293, 2004.
- 63 RAITERI, R. et al. Sensing of biological substances based on the bending of microfabricated cantilevers. *Sensors and Actuators B*, v. 61, n. 1, 1999.
- 64 BAUMANN, N.; PHAM-DINH, D. Biology of oligodendrocyte and myelin in the mammalian central nervous system. *Physiological Reviews*, v. 81, n. 2, p. 871–927, 2001.
- 65 WOOLSEY, T. A.; HANAWAY, J.; GADO, M. H. The brain atlas: a visual guide to the human central nervous system. *American Journal of Neuroradiology*, v. 29, n. 5, p. e38–e38, 2008.
- 66 SILVER, J.; SCHWAB, M. E.; POPOVICH, P. G. Central nervous system regenerative failure: role of oligodendrocytes, astrocytes, and microglia. *Cold Spring Harbor Perspectives in Biology*, v. 7, n. 3, p. a020602, 2014.
- 67 FUCHS, K. L. *Relapsing–remitting multiple sclerosis*. New York: Springer, 2011.
- 68 BRUM, D. G. et al. European ancestry predominates in neuromyelitis optica and multiple sclerosis patients from Brazil. *PLoS ONE*, v. 8, n. 3, p. e58925, 2013.
- 69 REINDL, M. et al. The spectrum of MOG autoantibody-associated demyelinating diseases. *Nature Reviews Neurology*, v. 9, n. 8, p. 455–461, 2013.
- 70 RAMANATHAN, S.; DALE, R. C.; BRILOT, F. Anti-MOG antibody: the history, clinical phenotype, and pathogenicity of a serum biomarker for demyelination. *Autoimmunity Reviews*, v. 15, n. 4, p. 307–324, 2016.
- 71 GOLDENBERG, M. M. Multiple sclerosis review. *Pharmacy and Therapeutics*, v. 37, n. 3, p. 175–184, 2012.
- 72 EGG, R. et al. Anti-MOG and anti-MBP antibody subclasses in multiple sclerosis. *Multiple Sclerosis*, v. 7, n. 5, p. 285–289, 2001.
- 73 WARSI, N. et al. Distinct immune cell subsets influence astrocytes to express factors relevant for myelin repair. *Neurology*, v. 80, n. 7 Suppl., p. P05169–P05169, 2013.
- 74 KINZEL, S.; WEBER, M. S. The role of peripheral CNS-directed antibodies in promoting inflammatory CNS demyelination. *Brain Sciences*, v. 7, n. 7, 2017.

- 75 DOS PASSOS, G. R. et al. MOG-IgG-associated optic neuritis, encephalitis, and myelitis: lessons learned from neuromyelitisoptica spectrum disorder. *Frontiers in Neurology*, v. 9, n. 4, 2018.
- 76 BENJAMINE ARELLANO, R. H. Human aquaporin 4281-300 Is the immunodominant linear determinant in the context of HLA-DRB1*03:01: relevance for diagnosing and monitoring patients with neuromyelitisoptica. *Archives of Neurology*, v. 69, n. 9, p. 1125–1131, 2012.
- 77 LANA-PEIXOTO, M. A.; CALLEGARO, D. The expanded spectrum of neuromyelitisoptica: evidences for a new definition. *Arquivos de Neuro-Psiquiatria*, v. 70, n. 10, p. 807–813, 2012.
- 78 JARIUS, S. et al. Aquaporin-4 antibody testing: direct comparison of M1-AQP4-DNA-transfected cells with leaky scanning versus M23-AQP4-DNA-transfected cells as antigenic substrate. *Journal of Neuroinflammation*, v. 11, p. 129, 2014.
- 79 JARIUS, S. et al. Testing for antibodies to human aquaporin-4 by ELISA: sensitivity, specificity, and direct comparison with immunohistochemistry. *Journal of the Neurological Sciences*, v. 320, n. 1–2, p. 32–37, 15 2012.
- 80 ZHONG, X.-N. et al. Relationship between neuromyelitisoptica-IgG status and spinal cord magnetic resonance imaging in patients with neuromyelitisoptica. *Chinese Medical Journal*, v. 125, n. 2, p. 270–274, 2012.
- 81 CREE, B. A. C. et al. An open label study of the effects of rituximab in neuromyelitisoptica. *Neurology*, v. 64, n. 7, p. 1270–1272, 2005.
- 82 RESKE, D.; HAUPT, W. F. Use of rituximab in multiple sclerosis: current progress and future perspectives. *Expert Review of Clinical Immunology*, v. 4, n. 5, p. 573–582, 2008.
- 83 BICHUETTI, D. B. et al. Neuromyelitisoptica treatment: analysis of 36 patients. *Archives of Neurology*, v. 67, n. 9, p. 1131–1136, 2010.
- 84 CASETTA, I.; IULIANO, G.; FILIPPINI, G. Azathioprine for multiple sclerosis. *Cochrane Database of Systematic Reviews*, n. 4, p. CD003982, 2007.
- 85 MEUTH, S. G.; GÖBEL, K.; WIENDL, H. Immune therapy of multiple sclerosis--future strategies. *Current Pharmaceutical Design*, v. 18, n. 29, p. 4489–4497, 2012.
- 86 WARABI, Y.; MATSUMOTO, Y.; HAYASHI, H. Interferon beta-1b exacerbates multiple sclerosis with severe optic nerve and spinal cord demyelination. *Journal of the Neurological Sciences*, v. 252, n. 1, p. 57–61, 2007.
- 87 HAUSSEN, S. R. *Esclerose múltipla – informações científicas para o leigo*. Porto Alegre: Conceito, 2004.
- 88 TILBERY, C. P. *Esclerose múltipla no Brasil: aspectos clínicos e terapêuticos*. Rio de Janeiro: Atheneu, 2005.
- 89 HINSON, S. R. et al. Pathogenic potential of IgG binding to water channel extracellular domain in neuromyelitisoptica. *Neurology*, v. 69, n. 24, p. 2221–2231, 2007.

- 90 VINCENT, T. et al. Functional consequences of neuromyelitisoptica-IgG astrocyte interactions on blood-brain barrier permeability and granulocyte recruitment. *Journal of Immunology*, v. 181, n. 8, p. 5730–5737, 2008.
- 91 TANAKA, M.; TANAKA, K. Anti-MOG antibodies in adult patients with demyelinating disorders of the central nervous system. *Journal of Neuroimmunology*, v. 270, n. 1–2, p. 98–99, 2014.
- 92 HOHLFELD, R. et al. The search for the target antigens of multiple sclerosis, part 2: CD8+ T cells, B cells, and antibodies in the focus of reverse-translational research. *Lancet Neurology*, v. 15, n. 3, p. 317–331, 2016.
- 93 LAMPASONA, V. et al. Similar low frequency of anti-MOG IgG and IgM in MS patients and healthy subjects. *Neurology*, v. 62, n. 11, p. 2092–2094, 2004.
- 94 MANTEGAZZA, R. et al. Anti-MOG autoantibodies in Italian multiple sclerosis patients: specificity, sensitivity and clinical association. *International Immunology*, v. 16, n. 4, p. 559–565, 2004.
- 95 SPADARO, M. et al. Autoantibodies to MOG in a distinct subgroup of adult multiple sclerosis. *Neurology® Neuroimmunology&Neuroinflammation*, v. 3, n. 5, 2016.
- 96 CASE, D. A. Perspective on “Dynamics of folded proteins”. *Theoretical Chemistry Accounts*, v. 103, n. 3–4, p. 332–334, 2000.
- 97 CRAMER, C. J. *Essentials of computational chemistry: theories and models*. 2nd ed. New York: Wiley, 2004.
- 98 JENSEN, F. *Introduction to computational chemistry*. 2nd ed. Chichester, Hoboken, NJ: Wiley, 2007.
- 99 ORRY, A. J. W.; ABAGYAN, R. *Homology modeling: methods and protocols*. New York: Humana Press, 2012.
- 100 CRAMER, C. J. *Essentials of computational chemistry: theories and models*. 2nd ed. New York: Wiley, 2004.
- 101 CORNELL, W. et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, v. 117, n. 3, p. 5179–5197, 1995.
- 102 SHATTUCK, T. W. *Molecular mechanics tutorial 2008*. Colby College Molecular Mechanics Tutorial Introduction. 2008.
- 103 VAN DER SPOEL, D. et al. GROMACS: fast, flexible, and free. *Journal of Computational Chemistry*, v. 26, n. 16, p. 1701–18, 2006.
- 104 PRONK, S. et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, v. 29, n. 7, p. 845–854, 2013.
- 105 ABRAHAM, M. J. et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, v. 1–2, p. 19–25, 2015.

- 106 IZRAILEV, S. et al. Steered molecular dynamics. In: DEUFLHARD, P. et al. (Ed.). *Computational molecular dynamics: challenges, methods, ideas*. Berlin: Springer, 1999. p. 39–65. (Lecture notes in computational science and engineering)
- 107 ISRALEWITZ, B.; GAO, M.; SCHULTEN, K. Steered molecular dynamics and mechanical functions of proteins. *Current Opinion in Structural Biology*, v. 11, n. 2, p. 224–230, 2001.
- 108 ISRALEWITZ, B. et al. Steered molecular dynamics investigations of protein function. *Journal of Molecular Graphics & Modelling*, v. 19, n. 1, p. 13–25, 2001.
- 109 MORGON, N. H.; COUTINHO, K. *Métodos de química teórica e modelagem molecular*. São Paulo: Livraria da Física, 2007.
- 110 BROOIJMANS, N.; KUNTZ, I. D. Molecular recognition and docking algorithms. *Annual Review of Biophysics and Biomolecular Structure*, v. 32, p. 335–373, 2003.
- 111 GUIZADO, T. R. C. *Estudos computacionais da interação de porfirinas e seus complexos de ferro com albumina sérica humana*. 2008. 141 p. Dissertação (Mestrado) - Departamento de Física, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008.
- 112 MORRIS, G. M. et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, v. 19, n. 14, p. 1639–1662, 1998.
- 113 MORRIS, G. M. et al. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry*, v. 30, n. 16, p. 2785–2791, 2009.
- 114 ATKINS, P.; JONES, L. *Princípios de química: questionando a vida moderna e o meio ambiente*. 3. ed. Porto Alegre: BOOKMAN, 2006.
- 115 BROWN, T. L.; LEMAY, H. E.; BURSTEN, B. E. *Química: a ciência central*. 9. ed. São Paulo: Pearson Prentice Hall, 2005.
- 116 DAVIS, M. E.; MCCAMMON, J. A. Electrostatics in biomolecular structure and dynamics. *Chemical Reviews*, v. 90, n. 3, p. 509–521, 1990.
- 117 BAKER, N. A. *APBS and PDB2PQR: electrostatic and solvation properties for complex molecules*. Disponível em: <<http://www.poissonboltzmann.org/>>. Acesso em: 4 set. 2018.
- 118 JURRUS, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Science*, v. 27, n. 1, p. 112–128, 2018.
- 119 HO, J. D. et al. Crystal structure of human aquaporin 4 at 1.8 Å and its mechanism of conductance. *Proceedings of the National Academy of Sciences*, v. 106, n. 18, p. 7437–7442, 2009.

- 120 HARRIS, L. J.; SKALETISKY, E.; MCPHERSON, A. Crystallographic structure of an intact IgG1 monoclonal antibody. *Journal of Molecular Biology*, v. 275, n. 5, p. 861–872, 1998.
- 121 GARDBERG, A. S. et al. Molecular basis for passive immunotherapy of Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America*, v. 104, n. 40, p. 15659–15664, 2007.
- 122 NAIR, D. T. et al. Epitope recognition by diverse antibodies suggests conformational convergence in an antibody response. *Journal of Immunology*, v. 168, n. 5, p. 2371–2382, 2002.
- 123 DU, J. et al. Crystal structure of chimeric antibody C2H7 Fab in complex with a CD20 peptide. *Molecular Immunology*, v. 45, n. 10, p. 2861–2868, 2008.
- 124 GUEX, N.; PEITSCH, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, v. 18, n. 15, p. 2714–2723, 1997.
- 125 GUEX, N.; PEITSCH, M. C.; SCHWEDE, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis*, v. 30, p. S162-73, 2009.
- 126 WANG, H. et al. Glutaraldehyde modified mica: a new surface for atomic force microscopy of chromatin. *Biophysical Journal*, v. 83, n. 6, p. 3619–3625, 2002.
- 127 MICA- Infopédia. Disponível em: <[http://www.infopedia.pt/\\$mica](http://www.infopedia.pt/$mica)>. Acesso em: 1 mar. 2015.
- 128 NEESE, F. The ORCA program system. *WileyInterdisciplinaryReviews: computational molecular science*, v. 2, n. 1, p. 73–78, 2012.
- 129 DUPRADEAU, F.-Y. et al. The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Physical Chemistry Chemical Physics*, v. 12, n. 28, p. 7821–7839, 2010.
- 130 VALIEV, M. et al. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, v. 181, n. 9, p. 1477–1489, 2010.
- 131 HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: visual molecular dynamics. *Journal of Molecular Graphics*, v. 14, n. 1, p. 33–38, 1996.
- 132 HERRAEZ, A. Biomolecules in the Computer: Jmol to the Rescue. *Biochemistry and Molecular Biology Education*, v. 34, n. 4, p. 255–261, 2006.
- 133 GASTEIGER, J.; MARSILI, M. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron*, v. 36, n. 22, p. 3219–3228, 1980.
- 134 DAVIS, M. E.; MCCAMMON, J. A. Electrostatics in biomolecular structure and dynamics. *Chemical Reviews*, v. 90, n. 3, p. 509–521, 1990.

135 NIELSEN, J. E. et al. Improving macromolecular electrostatics calculations. *Protein Engineering*, v. 12, n. 8, p. 657–662, 1999.

136 HOLST, M.; BAKER, N.; WANG, F. Adaptive multilevel finite element solution of the Poisson–Boltzmann equation I. algorithms and examples. *Journal of Computational Chemistry*, v. 21, n. 15, p. 1319–1342, 2000.

137 AL-LAZIKANI, B.; LESK, A. M.; CHOTHIA, C. Standard conformations for the canonical structures of immunoglobulins. *Journal of Molecular Biology*, v. 273, n. 4, p. 927–948, 1997.

138 SHIRAI, H.; KIDERA, A.; NAKAMURA, H. H3-rules: identification of CDR-H3 structures in antibodies. *FEBS Letters*, v. 455, n. 1–2, p. 188–197, 1999.

139 CHOTHIA, C.; LESK, A. M. Canonical structures for the hypervariable regions of immunoglobulins. *Journal of Molecular Biology*, v. 196, n. 4, p. 901–917, 1987.

140 VERKMAN, A. S. et al. Aquaporin-4: orthogonal array assembly, CNS functions, and role in neuromyelitisoptica. *ActaPharmacologicaSinica*, v. 32, n. 6, p. 702–710, 2011.

141 BERENDSEN, H. J. C.; GRIGERA, J. R.; STRAATSMA, T. P. The missing term in effective pair potentials. *Journal of Physical Chemistry*, v. 91, n. 24, p. 6269–6271, 1987.

142 HOCKNEY, R. W. The potential calculation and some applications. In: *Methods in computational physics*. New York/London: Academic Press, 1970.

143 BERENDSEN, H. J. C. et al. Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics*, v. 81, n. 8, p. 3684, 1984.

144 HESS, B. et al. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, v. 18, n. 12, p. 1463–1472, 1997.

145 TIRONI, I. G. et al. A generalized reaction field method for molecular dynamics simulations. *Journal of Chemical Physics*, v. 102, n. 13, p. 5451–5459, 1995.

146 PHILLIPS, J. C. et al. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, v. 26, n. 16, p. 1781–1802, 2005.

147 HUANG, J.; MACKERELL, A. D. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of Computational Chemistry*, v. 34, n. 25, p. 2135–2145, 2013.

148 JORGENSEN, W. L. et al. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, v. 79, n. 2, p. 926–935, 1983.

149 DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems. *Journal of Chemical Physics*, v. 98, n. 12, p. 10089–10092, 1993.

150 FELLER, S. E. et al. Constant pressure molecular dynamics simulation: the Langevin piston method. *Journal of Chemical Physics*, v. 103, n. 11, p. 4613–4621, 1995.

- 151 WEISSTEIN, E. W. *Gamma distribution MathWorld*--a Wolfram web resource. Disponível em: <<http://mathworld.wolfram.com/GammaDistribution.html>>. Acesso em: 19 abr. 2018
- 152 THOM, H. C. S. A note on the gamma distribution. *Monthly Weather Review*, v. 86, n. 4, p. 117–122, 1958.
- 153 BORLINHAS, F. et al. Gamma distribution model in breast cancer diffusion-weighted imaging. In: PORTUGUESE MEETING ON BIOENGINEERING, 4, 2015, Porto. *Proceedings ... Porto*: IEEE, 2015.
- 154 GIRARDIN, G. et al. Correlation between dislocation organization and slip bands: TEM and AFM investigations in hydrogen-containing nickel and nickel–chromium. *Acta Materialia*, v. 91, p. 141–151, 2015.
- 155 GONCHAROV, V. K.; KOZADAEV, K. V.; MIKITCHUK, A. P. Piagnostics of the monolayer silver nanostructures on a solid substrate using the bifactorial analysis of the spr band. *High Temperature Material Processes: an international quarterly of high-technology plasma processes*, v. 18, n. 3, 2014.
- 156 ANGELEROU, M. G. F. et al. Surface-directed modulation of supramolecular gel properties. *Chemical Communications*, v. 52, n. 23, p. 4298–4300, 2016.
- 157 LU, B.; SMYTH, M. R.; O’KENNEDY, R. Oriented immobilization of antibodies and its applications in immunoassays and immunosensors. *Analyst*, v. 121, n. 3, p. 29R–32R, 1996.
- 158 OLIVEIRA, G. S. et al. Immobilization and unbinding investigation of the antigen-antibody complex using theoretical and experimental techniques. *Journal of Molecular Graphics and Modelling*, v. 86, p. 219–227, 2019.
- 159 WANG, S. et al. Efficient on-chip isolation of HIV subtypes. *Lab on a Chip*, v. 12, n. 8, p. 1508–1515, 2012.
- 160 DEIS, L. N. et al. Suppression of conformational heterogeneity at a protein-protein interface. *Proceedings of the National Academy of Sciences of the United States of America*, v. 112, n. 29, p. 9028–9033, 2015.
- 161 ERICKSON, H. P. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biological Procedures Online*, v. 11, p. 32–51, 2009.
- 162 ANDERSON, G. P. et al. Effectiveness of protein A for antibody immobilization for a fiber optic biosensor. *Biosensors & Bioelectronics*, v. 12, n. 4, p. 329–336, 1997.
- 163 BABACAN, S. et al. Evaluation of antibody immobilization methods for piezoelectric biosensor application. *Biosensors & Bioelectronics*, v. 15, n. 11–12, p. 615–621, 2000.
- 164 COEN, N. et al. Adsorption and bioactivity of protein A on silicon surfaces studied by AFM and XPS. *Journal of Colloid and Interface Science*, v. 233, n. 2, p. 180–189, 2001.

- 165 KAUSAITE-MINKSTIMIENE, A. et al. Comparative study of Random and oriented antibody immobilization techniques on the binding capacity of immunosensor. *Analytical Chemistry*, v. 82, p. 6401–6408, 2010. doi:10.1021/ac100468k.
- 166 IEONG, P.; AMARO, R. E.; LI, W. W. Molecular dynamics analysis of antibody recognition and escape by human H1N1 influenza hemagglutinin. *Biophysical Journal*, v. 108 p. 2704–2712, 2017. doi:10.1016/j.bpj.2015.04.025.
- 167 KNAPP, B. et al. Variable regions of antibodies and T-cell receptors may not be sufficient in molecular simulations investigating binding. *Journal of Chemical Theory and Computation*, v. 13 p. 3097–3105, 2017. doi:10.1021/acs.jctc.7b00080.
- 168 O'BRIEN, C. J. et al. Biophysical characterization and molecular simulation of electrostatically driven self-association of a single-chain antibody. *Protein Science*, v. 27, p. 1275–1285, 2018. doi:10.1002/pro.3415.
- 169 IERICH, J. C. M. et al. Antibody-mediated biorecognition of myelin oligodendrocyte glycoprotein: computational evidence of demyelination-related epitopes. *Scientific Reports*, v. 9, n. 2033, 2019. doi: 10.1038/s41598-018-36578-8.
- 170 YAMASHITA, T. Toward rational antibody design: recent advancements in molecular dynamics simulations. *International Immunology*, v. 30, p. 133–140, 2018. doi:10.1093/intimm/dxx077.

APÊNDICE A

Modelo preliminar da estrutura de anticorpos por cálculos de docking

Nesta etapa cálculos de *docking* molecular foram conduzidos utilizando a ideia de engenharia molecular (1–6) para o peptídeo da AQP4 como estrutura ligante acoplado às estruturas das IgGs modificadas, a fim de prever orientações conformacionais mais favoráveis, segundo ΔG de ligação. O princípio de modelagem da estrutura da anti-AQP4 consistiu na mutação da estrutura da IgG antifenobarbital, disponível no PDB sob o código PDB ID: 1IGY (13), devido a ausência de programas computacionais que fossem capazes de fazer a previsão da estrutura de anticorpos partindo apenas da estrutura de epítomos de interesse (7–12). A IgG descrita na estrutura 1IGY resolvida em sua totalidade consiste no subtipo 1, o mesmo utilizado experimentalmente para a anti-AQP4 (14). Neste contexto, a estrutura PDB 1IGY foi escolhida como ponto de partida para os cálculos de engenharia molecular baseados em *docking* molecular. Para que a estrutura da IgG subtipo 1 antifenobarbital se torne a mais próxima possível da IgG anti-AQP4, foram realizadas inicialmente mutações nos 2 aminoácidos que mais interagem com o antígeno das alças L1, L2, L3 e H3 da FIG. A1. Estas alças foram inicialmente escolhidas devido ao tamanho do antígeno (10 aminoácidos), bem como a hipervariabilidade da alça H3 (7,15–17).

Figura A1 – Alças que foram mutadas inicialmente da IgG anti-fenobarbital, com o objetivo de obter uma estrutura mais próxima da anti-AQP4.
Fonte: Elaborada pelo autor.

O protocolo foi definido para a realização da engenharia molecular das alças da IgG conforme etapas apresentadas na Tabela A1.

Tabela A1 – Etapas de desenvolvimento do protocolo de engenharia molecular baseada em *docking*

Etapas	Objetivos	Procedimentos
1 ^a	Determinar as possíveis combinações de aminoácidos que compõem as alças e gerar as estruturas a partir da substituição dos resíduos.	MMC e mutação da região de ligação direcionada
2 ^a	Determinar as estruturas mais favoráveis para as alças de interação em relação à afinidade do complexo antígeno-anticorpo.	cálculos de <i>docking</i>
3 ^a	Caracterizar por DM os melhores complexos selecionados e obter a estrutura mais adequada, ou seja, aquela que melhor reproduz a IgG anti-AQP4.	Simulações de DM

Fonte: Elaborada pelo autor.

Pela aplicação da primeira etapa do protocolo descrito na TABELA A1, para cada alça a ser mutada, deve-se considerar 20 aminoácidos que podem ser arranjados inicialmente em duas posições-alvo das mutações (aminoácidos interagentes), o que resulta em 20x20

possibilidades de alocação de novos aminoácidos em substituição àqueles presentes na estrutura cristalográfica PDB 1IGY. Tratando cada possibilidade individualmente, tem-se $20 \times 20 \times 4 = 1600$ estruturas iniciais para os cálculos de *docking* molecular, sendo 20 aminoácidos para cada uma das 2 posições-alvo nas 4 alças. Neste cenário, o passo seguinte, referente à segunda etapa do protocolo, consiste em executar os cálculos de *docking* molecular e analisar os complexos obtidos, a fim de selecionar as melhores estruturas, ou seja, aquelas com a menor energia livre de ligação do complexo antígeno-anticorpo.

A configuração inicial e os aminoácidos a serem substituídos estão representados em diferentes cores na FIG. A2. Para cada aminoácido substituído, como exemplificado na FIG. A2, foi necessária uma nova etapa de minimização de energia do sistema, para permitir a eliminação de possíveis repulsões estéricas entre os átomos que possam interferir na simulação da estrutura do complexo antígeno-anticorpo.

Figura A2 – Mutação de um aminoácido interagente da IgG, representado em (A) azul e em (B) vermelho, após a mutação. (C) Para cada mutação realizada, é necessária uma nova DM.

Fonte: Elaborada pelo autor.

Na terceira etapa do protocolo, as conformações antígeno-anticorpo descritas por *docking* molecular com valores de energia de ligação mais favoráveis foram simuladas computacionalmente por DM de forma a verificar a estabilidade estrutural em água e os valores de energia de interação que descrevem o complexo antígeno-anticorpo.

A Tabela A2 mostra uma das 1600 configurações possíveis que estão sendo simuladas pelos cálculos de *docking* molecular. Esta pode ser considerada uma pré-estrutura da IgG anti-AQP4.

Tabela A2 – Configuração 1 dos aminoácidos mutados centrais da IgG tipo 1.

Alça	Aminoácido 1	Aminoácido 2
L1	31GLU	32LYS
L2	50SER	51TYR
L3	92TYR	93ARG
H3	97CYS	98HIS

Fonte: Elaborada pelo autor.

Para esta configuração, foram selecionados os 6 melhores complexos de interação antígeno-anticorpo (segundo *ranking* de energia) gerados nos cálculos de *docking* (FIG. A3).

Figura A3 – Epítipo da AQP4 e IgG antifenobarbital mutada em complexos. A imagem apresenta os 6 melhores resultados obtidos pelos cálculos de *docking* molecular para configuração 1 dos aminoácidos mutados das alças da IgG. $\Delta G = -4$ kJ/mol.

Fonte: Elaborada pelo autor.

O modelo preliminar com *docking* mostrou o funcionamento e eficiência das mutações, o que serviu de base para o aperfeiçoamento das novas técnicas apresentadas e foi fundamental na execução da formação inicial do complexo antígeno-anticorpo do protocolo final, dada a eficiência de se realizar mutações pontuais em tais estruturas (18,19).

Para a elaboração do protocolo final e a modelagem completa das alças de ligação foi realizado um teste correspondente à modelagem das 400 combinações para a alça H3,

utilizando como *input* o epítipo AQP4. O intuito do teste inicial foi identificar o comportamento da alça H3 na formação do complexo, dado que sua hipervariabilidade e hiperflexibilidade possuem contribuição fundamental na modelagem de anticorpos (20). A simulação de DM foi executada por 4 ns aponta ainda instabilidade no sistema, de acordo com o RMSD mostrado na FIG. A4, com a energia de ligação ΔG do complexo flutuando entre -7 kJ/mol e -10,5 kJ/mol.

Figura A4 – Flutuação conformacional e flutuação da energia de ligação ΔG do complexo.
Fonte: Elaborada pelo autor.

Os cálculos iniciais de SMD do complexo foram realizados com intuito de testar e tirar informações úteis da primeira estrutura gerada (para 1 divisão da alça H3) para auxiliar no refinamento final do modelo. A Fig. A5 mostra o cálculo de SMD realizado em um tempo de

2000 ps. Tempo que foi suficiente para que, levando-se em conta os parâmetros experimentais, ocorresse a reprodução computacional do desacoplamento do ligante, revertendo a formação do complexo e simulando o funcionamento do AFM. A força de adesão, nesta situação, engloba a média de várias curvas de SMD, realizadas em vetores $\vec{F}$ em diferentes direções.

Figura A5 – SMD inicial do completo antígeno-anticorpo genérico. Peptídeo ligante sendo extraído da macromolécula, desfazendo o complexo. Dividindo o tempo total de simulação (2 ns) em quatro etapas iguais A, B, C e D.

Fonte: Elaborada pelo autor.

De acordo com a energia de ligação calculada para o complexo, a força de adesão encontra-se, inicialmente, em valores próximos a 1500 kJ/mol/nm (FIG. A6) e, após cerca de 10 ps, diminui bruscamente para zero, sugerindo interação forte do antígeno com solvente.

Figura A6 – Gráfico da SMD inicial do completo antígeno-anticorpo genérico. Mostra o gráfico da força calculada para reverter o complexo, até o tempo de 2000 ps.

Fonte: Elaborada pelo autor.

A partir dos dados apresentados, o modelo obtido levou à formação de um complexo com cavidade mais apropriada à interação. Na FIG. A7(A), o mapa de potencial eletrostático da superfície da parte Fab da IgG, mostra a região construída em sua maioria por resíduos carregados positivamente e neutros, adaptando-se às características da distribuição de cargas predominantemente neutro do epítipo FIG A7(B) (21).

Figura A7 – Potencial eletrostático da IgG genérica com alça H3 mutada. O epítipo é representado na cor verde.

Nota: k_B é a constante de Boltzmann, T a temperatura e e a carga elementar.

Fonte: Elaborada pelo autor.

Apesar da formação do complexo ter apresentado bom resultado, estruturalmente (FIG. A7) e de acordo com a energia de ligação mostrada na FIG. A6, a estrutura não se mostrou estável na presença das moléculas de água, fazendo com que o complexo antígeno anticorpo seja desfeito, como mostrado nos cálculos de SMD, evidenciando a necessidade da modelagem completa das alças (20).

Referências

- (1) D. Kuroda, H. Shirai, M.P. Jacobson, H. Nakamura, Computer-aided antibody design, *Protein Eng Des Sel.* 25 (2012) 507–522. doi:10.1093/protein/gzs024.
- (2) A.M. Sevy, J. Meiler, Antibodies: Computer-Aided Prediction of Structure and Design of Function, *Microbiol Spectr.* 2 (2014). doi:10.1128/microbiolspec.AID-0024-2014.
- (3) J. Loyau, G. Didelot, P. Malinge, U. Ravn, G. Magistrelli, J.-F. Depoisier, G. Pontini, Y. Poitevin, M. Kosco-Vilbois, N. Fischer, S. Thore, F. Rousseau, Robust Antibody-Antigen Complexes Prediction Generated by Combining Sequence Analyses, Mutagenesis, In Vitro Evolution, X-ray Crystallography and In Silico Docking, *J. Mol. Biol.* 427 (2015) 2647–2662. doi:10.1016/j.jmb.2015.05.016.

- (4) Y. Choi, C.M. Deane, Predicting antibody complementarity determining region structures without classification, *Mol Biosyst.* 7 (2011) 3327–3334. doi:10.1039/c1mb05223c.
- (5) C. Marks, C.M. Deane, Antibody H3 Structure Prediction, *Comput Struct Biotechnol J.* 15 (2017) 222–231. doi:10.1016/j.csbj.2017.01.010.
- (6) J. Zhao, R. Nussinov, W.-J. Wu, B. Ma, J. Zhao, R. Nussinov, W.-J. Wu, B. Ma, In Silico Methods in Antibody Design, *Antibodies.* 7 (2018) 22. doi:10.3390/antib7030022.
- (7) P. Marcatili, P.P. Olimpieri, A. Chailyan, A. Tramontano, Antibody modeling using the Prediction of ImmunoGlobulin Structure (PIGS) web server, *Nature Protocols.* 9 (2014) 2771–2783. doi:10.1038/nprot.2014.189.
- (8) R. Lepore, P.P. Olimpieri, M.A. Messih, A. Tramontano, PIGSPro: prediction of immunoGlobulin structures v2, *Nucleic Acids Res.* 45 (2017) W17–W23. doi:10.1093/nar/gkx334.
- (9) ROSIE: The Rosetta Online Server that Includes Everyone, (n.d.). <http://antibody.graylab.jhu.edu/> (accessed September 21, 2018).
- (10) H. Kemmish, M. Fasnacht, L. Yan, Fully automated antibody structure prediction using BIOVIA tools: Validation study, *PLOS ONE.* 12 (2017) e0177923. doi:10.1371/journal.pone.0177923.
- (11) J. Dunbar, K. Krawczyk, J. Leem, C. Marks, J. Nowak, C. Regep, G. Georges, S. Kelm, B. Popovic, C.M. Deane, SAbPred: a structure-based antibody prediction server, *Nucleic Acids Res.* 44 (2016) W474–478. doi:10.1093/nar/gkw361.
- (12) A. Koivuniemi, K. Takkinen, T. Nevanen, A computational approach for studying antibody–antigen interactions without prior structural information: the anti-testosterone binding antibody as a case study, *Proteins: Structure, Function, and Bioinformatics.* 85 (2017) 322–331. doi:10.1002/prot.25226.
- (13) L.J. Harris, E. Skaletsky, A. McPherson, Crystallographic structure of an intact IgG1 monoclonal antibody, *J. Mol. Biol.* 275 (1998) 861–872. doi:10.1006/jmbi.1997.1508.
- (14) P. Waters, S. Jarius, E. Littleton, M.I. Leite, S. Jacob, B. Gray, R. Geraldles, T. Vale, A. Jacob, J. Palace, S. Maxwell, D. Beeson, A. Vincent, Aquaporin-4 antibodies in neuromyelitis optica and longitudinally extensive transverse myelitis, *Arch. Neurol.* 65 (2008) 913–919. doi:10.1001/archneur.65.7.913.
- (15) C. Chothia, A.M. Lesk, Canonical structures for the hypervariable regions of immunoglobulins, *J. Mol. Biol.* 196 (1987) 901–917.
- (16) H. Shirai, A. Kidera, H. Nakamura, H3-rules: identification of CDR-H3 structures in antibodies, *FEBS Lett.* 455 (1999) 188–197.
- (17) B. Al-Lazikani, A.M. Lesk, C. Chothia, Standard conformations for the canonical structures of immunoglobulins, *J. Mol. Biol.* 273 (1997) 927–948. doi:10.1006/jmbi.1997.1354.
- (18) S. Yanaka, Y. Moriwaki, K. Tsumoto, K. Sugase, Elucidation of potential sites for antibody engineering by fluctuation editing, *Scientific Reports.* 7 (2017) 9597. doi:10.1038/s41598-017-10246-9.
- (19) Y. Zhang, J. Guo, L. Huang, J. Tian, X. Yao, H. Liu, The molecular mechanism of two coreceptor binding site antibodies X5 and 17b neutralizing HIV-1: Insights from

- molecular dynamics simulation, *Chemical Biology & Drug Design*. 92 (2018) 1357–1365. doi:10.1111/cbdd.13201.
- (20) J.R. Jeliazkov, A. Sljoka, D. Kuroda, N. Tsuchimura, N. Katoh, K. Tsumoto, J.J. Gray, Repertoire Analysis of Antibody CDR-H3 Loops Suggests Affinity Maturation Does Not Typically Result in Rigidification, *Front. Immunol.* 9 (2018). doi:10.3389/fimmu.2018.00413.
- (21) M.M. Al Qaraghuli, K. Kubiak-Ossowska, P.A. Mulheran, Thinking outside the Laboratory: Analyses of Antibody Structure and Dynamics within Different Solvent Environments in Molecular Dynamics (MD) Simulations, *Antibodies*. 7 (2018) 21. doi:10.3390/antib7030021.

APÊNDICE B

Distribuição Gama Interpretação dos resultados experimentais relacionados à EM

A interpretação dos experimentos de AFM pela distribuição gama também foi aplicada nos experimentos envolvendo a EM. Tais resultados experimentais foram obtidos pela Dra. Pâmela Soto Garcia (Projeto FAPESP nº 2014/21540-0) e consistiram em detectar o biomarcador que pode indicar desmielinização na EM.

Tendo como elemento sensor o epítipo MOG₉₂₋₁₀₆ foram levantadas as curvas controle apresentadas na FIG B1. Com base nos soros e nos LCRs dos pacientes (FIG. B1), e utilizando a distribuição gama, foi possível fazer o levantamento teórico das curvas padrão de comportamento para o que se espera de pacientes EM positivos (FIG. B2) e EM negativos (FIG. B3).

Figura B1 – Moda dos valores de força de adesão para as medidas controles para anti-MOG e curvas padrão teóricas levantadas a partir de todos os pacientes soro positivos e soro negativos para anti-MOG.

Fonte: Elaborada em conjunto com a Dra. Pâmela Soto Garcia (Projeto FAPESP nº 2014/21540-0).

Figura B2 – Moda dos valores de força de adesão para os soros e LCRs purificados dos pacientes soropositivos para anti-MOG e EM.

Fonte: Elaborada em conjunto com a Dra. Pâmela Soto Garcia (Projeto FAPESP nº 2014/21540-0).

Figura B3 – Moda dos valores de força de adesão para o soros e LCRs purificados dos pacientes soronegativos para anti-MOG e EM.

Fonte: Elaborada em conjunto com a Dra. Pâmela Soto Garcia (Projeto FAPESP nº 2014/21540-0).

O anticorpo anti-MOG foi detectado com eficiência nos LCRs dos pacientes EM positivos, tomando como base as medidas controle mostradas na FIG. 31. Houve também detecção, em menor quantidade, nos soros destes pacientes. Nos pacientes negativos para EM pode-se dizer que não houve detecção, tanto nos LCRs quanto nos soros, indicando que o anticorpo anti-MOG é um bom indicador de desmielinização.

ANEXO A

Universidade de São Paulo

Biblioteca Digital da Produção Intelectual - BDPI

Departamento de Física e Ciências Materiais - IFSC/FCM

Artigos e Materiais de Revistas Científicas - IFSC/FCM

2014-09

Modeling the coverage of an AFM tip by enzymes and its application in nanobiosensors

Journal of Molecular Graphics and Modelling, Philadelphia : Elsevier, v. 53, p. 100-104, Sept. 2014
<http://www.producao.usp.br/handle/BDPI/50304>

Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Modeling the coverage of an AFM tip by enzymes and its application in nanobiosensors

Adriano M. Amarante^a, Guedmiller S. Oliveira^{a,*}, Carolina C. Bueno^a, Richard A. Cunha^b,
Jéssica C.M. Ierich^a, Luiz C.G. Freitas^c, Eduardo F. Franca^b, Osvaldo N. Oliveira Jr.^d,
Fábio L. Leite^{a,*}

^a Nanoneurobiophysics Research Group, Department of Physics, Chemistry and Mathematics, Federal University of São Carlos, 13052-780 Sorocaba, SP, Brazil

^b Chemistry Institute, Federal University of Uberlândia, 38400-902 Uberlândia, MG, Brazil

^c Chemistry Department, Federal University of São Carlos, 13565-905 São Carlos, SP, Brazil

^d São Carlos Institute of Physics, University of São Paulo, 13566-590 São Carlos, SP, Brazil

ARTICLE INFO

Article history:

Accepted 14 July 2014

Available online 23 July 2014

Keywords:

Computational rigid model

Stochastic adsorption

Enzyme modeling

ACC enzyme

SMD

AFM

ABSTRACT

A stochastic simulation of adsorption processes was developed to simulate the coverage of an atomic force microscope (AFM) tip with enzymes represented as rigid polyhedrons. From geometric considerations of the enzyme structure and AFM tip, we could estimate the average number of active sites available to interact with substrate molecules in the bulk. The procedure was exploited to determine the interaction force between acetyl-CoA carboxylase enzyme (ACC enzyme) and its substrate diclofop, for which steered molecular dynamics (SMD) was used. The theoretical force of (1.6 ± 0.5) nN per enzyme led to a total force in remarkable agreement with the experimentally measured force with AFM, thus demonstrating the usefulness of the procedure proposed here to assist in the interpretation of nanobiosensors experiments.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Nanobiosensors are highly sensitive and selective sensors with nanoscale resolution (1–100 nm) [1], which can be based on atomic force microscopy (AFM). Since with AFM one is capable of measuring forces of the order of magnitude of a chemical bond [2], specific molecular detection can be achieved in the so-called Chemical Force Microscopy (CFM). Various combinations of sensing molecules (e.g. enzymes, antibody) and substrates (e.g. inhibitors, antigens) can be used to build nanobiosensors with AFM. The design of an effective sensor, however, requires understanding of the specific interactions leading to detection signals, such as those generated by enzyme-inhibitor pairs [3–6]. Owing to the complexity of the enzyme–substrate interaction, *in silico* studies are now employed, including molecular dynamics (MD) simulations [7–9]. With the present computational capabilities, even using enhanced sampling techniques, the simulation of these interactions cannot be made at the atomistic level due to the size of the system and

the time scale of the phenomena taking place during operation of a nanobiosensor [10–12]. Therefore, the development of approximated methodologies is needed.

In this study, a coarse grained model is proposed to determine the interaction between acetyl-CoA carboxylase (ACC) and diclofop, a commercial herbicide known to block ACC activity by occupying its binding pocket [13–17] in a non-covalent, reversible interaction. The coarse-grained model for the AFM tip surface covered by enzymes uses atomistic information obtained from MD simulations and technical specifications from the manufacturer of the AFM tip. The model was used to determine the average interaction energy per active site while the number of active sites available to interact with the herbicide substrate was obtained using a stochastic model, where the enzymes were represented by polyhedrons adsorbing on the tip. The theoretical forces are compared to experimentally obtained values.

2. Methodology

2.1. ACC as dimeric structure

The initial structure of the ACC enzyme, shown in Fig. 1, was obtained from the Protein Data Bank online repository [18] under

* Corresponding authors. Tel.: +55 15 3229 6014.

E-mail addresses: guedmuller@gmail.com (G.S. Oliveira), fabioleite@ufscar.br (F.L. Leite).

Fig. 1. Artwork of the ACC enzymes arrangement on the AFM tip and substrate surface with functionalized diclofop molecules for detection.

PDB ID 1UYR. The enzyme was assumed to adopt a dimeric conformation in solution, according to molecular dynamics simulation with an all-atom model [19,20]. The coverage of the AFM tip with the enzymes was simulated with a stochastic process in MD simulations, where a coarse-grained model had to be used. The enzyme dimers were represented as rigid structures, as already applied in molecular docking calculations of inhibitor–enzyme interactions [21,22]. The use of a rigid model is justified based on the stability of the dimeric structures. Indeed, a stable dimeric ACC structure was reached within 10 ns of MD trajectory with a RMSD of 0.3 nm [19], whereas the dimeric structure was attached to a functionalized surface, it remained stable within a RMSD of 0.35 nm after a 50 ns MD trajectory [20].

2.2. Stochastic coverage of the AFM tip

The dimeric form of ACC was represented by a polyhedron with six faces, which dimensions A through F, depicted in Fig. 2, were obtained using the procedures described in [23,24] with the MD data by Franca et al. [19]. Because the charge distribution on the ACC enzyme dimer surface is homogeneous [19], there is no preferential region for enzyme adsorption on the AFM tip. Therefore, the enzyme–surface energy interaction will be assumed proportional to the contact area, and so will be the probability of each of the six faces (A–F) of the polyhedron to be in contact with the AFM tip. The adsorption process is then assumed to be stochastic with adsorption probability proportional to the contact area, which is defined in Fig. 2. As a consequence of the ACC dimer symmetry a twofold degeneracy is found between faces {A, B}, {C, D} and {E, F}. Hence, there are only three energetically distinct attachment positions.

2.3. Number of functionalized enzymes on the AFM tip surface

For a given configuration of polyhedrons covering the tip surface, geometric considerations can be used to determine the availability of the enzyme cavity to interact with substrates from the bulk. The number of ACC dimers attached to the AFM tip, which are effective for interacting with substrates, is required to normalize the signal measured in an AFM experiment. This number depends on geometrical parameters of the substrates as well as on the density of enzyme dimers on the AFM tip surface. The latter density depends on the tip radius and on the average area occupied by a single enzyme dimer. Fig. 3 shows a monolayer of adsorbed enzyme dimers which will effectively interact with the substrate molecules from the bulk.

The ACC dimer hydration energy calculated by Franca et al. [19] is $-23.36 \cdot 10^3$ kJ/mol, and its isoelectric point is 6.23 [25]. As

Table 1
Number of active sites available according to the polyhedron faces orientations.

Faces	Contact area (nm ²)	Covered area of the AFM tip (nm ²)	Active sites available
1	42.4	33.9	0
2	42.4	33.9	0
3	30.2	36.7	1
4	30.2	36.7	1
5	17.0	33.9	1
6	17.0	33.9	1

the hydration energy is small and electrostatic repulsion can be neglected, we assume that the AFM tip surface is uniformly covered by ACC dimers represented by rigid polyhedrons. The spherical AFM tip area A is calculated using $A = 2\pi\delta R$, where R is the radius given by the manufacturer [26] and δ is a parameter depending on the substrate geometry as indicated in Fig. 3. The calculated area for the AFM tip available for adsorption is $A = (220 \pm 31)$ nm².

A modeling software [23,24] was employed to obtain distributions of polyhedron conformations that could be adsorbed on the tip surface. The contact area for each face, and the area of the AFM tip covered by a given face were determined in Table 1. Also included in the table, the availability of the active site is showed. The stochastic program [27] to obtain sets of configurations for the adsorption of polyhedrons has a 4-step procedure, as follows:

- Step 1: choose the number N of dimers on the surface.
- Step 2: sort a given polyhedron face. For a given polyhedron face its Boltzmann probability is calculated assuming that the interaction energy is proportional to the contact area given in Table 1.
- Step 3: using information from Table 1, the number of active sites available for interaction is obtained.
- Step 4: return to Step 2 until N is reached.

For a given number N of dimers, random sets of arrangements for the polyhedrons on the AFM tip surface were generated. The average number of active sites available for interacting with the substrates is obtained in Step 3 at the end of the iterative stochastic procedure. Note that the number of active sites is smaller than the number of adsorbed enzymes because some sites are not exposed to the substrates, as will be discussed later on.

2.4. Steered molecular dynamics

Steered molecular dynamics (SMD) [10,28] calculations were carried out to simulate the average force required to detach the herbicide diclofop from the ACC active site cavity. As shown in Fig. 4, a vector toward the entrance of the active site was imposed to represent the pathway for detachment of the diclofop molecule. The calculations were performed using GROMACS computational package version 4.5.4 [8,29] with the following protocol: (i) the enzyme structure reported by Franca et al. was solvated by SPC [30] water molecules and ions were added to balance/equilibrate charge distribution in the system; (ii) the calculations were performed in the NpT ensemble at $p = 1$ bar and $T = 298$ K. Standard Berendsen barostat and Langevin thermostat procedures implemented in GROMACS were used; (iii) a cutoff radius of 1.4 nm was used and long range corrections for electrostatic potential were considered using the Particle Mesh Ewald (PME) formalism; (iv) using a 1 fs time step for integration, equilibrium was achieved within a 50 ps NVT calculations and a further 1 ns NpT trajectory was used for averaging; (v) a constant force of 3670 pN nm⁻¹ and a speed rate of 0.001 nm/ps were used in the detachment process.

Fig. 2. Proposed rigid model of the ACC enzyme according to its dimensions.

Fig. 3. Schematic monolayer section of both substrate surface and AFM tip radius covered by diclofop molecules and ACC enzymes respectively.

Fig. 4. Steered molecular dynamics representation of the herbicide diclofop decoupling process from the active site of the ACC enzyme.

Table 2

For a given number (N) of enzyme dimers covering the AFM tip, estimations were made of the area of the tip covered and the average number of active sites available for interaction, in the second and third columns, respectively. Using the force obtained from SMD simulations per enzyme dimer, the theoretical force was obtained. For full coverage of the tip, $N=6$, the predicted force agrees with the experimentally measured average force from Ref. [26]. These values are highlighted in bold in the third line. Because the experimental results led to a histogram of force values, from 3 to 8 nN, it is possible to predict how many active sites are responsible for the interaction. Here, N would vary from 4 to 7, as indicated in lines from 1 to 4. Results with confidence of 95%.

Number of enzyme (N)	AFM tip area covered by enzymes (nm^2)	Average number of active sites	Theoretical force (nN)	Experimental force (nN)
4	139.4	2.53	4.0 ± 1.3	3.9 ± 1.8
5	174.5	3.20	5.1 ± 1.6	4.9 ± 1.9
6	208.8	3.86	6.2 ± 1.9	6.1 ± 2.0
7	243.5	4.54	7.2 ± 2.3	7.0 ± 2.2

Fig. 5. Integrated force curve of the diclofop decoupling from the ACC active site. The calculated average force was estimated from 0 to 1500 ps time step.

3. Results

The force curve along the detachment pathway in Fig. 4, obtained with SMD simulations, is presented in Fig. 5. The average force was calculated in the time interval from $t=0$ to $t_f=1000$ ps, since the diclofop molecule is outside the cavity of the enzyme active site after a time t_f . An average force $F_S = (1.6 \pm 0.5)$ nN was obtained per enzyme, with a confidence level of 95%. Using the technical specifications from the manufacturer, the AFM tip radius (R) was assumed to be (17.5 ± 2.5) nm. With $A = 2\pi\delta R$ and $\delta = 2.0$ nm, the AFM available area is (220 ± 31) nm².

Table 2 shows the AFM tip area covered by enzymes for a given N of ACC dimers, obtained from averaging thousands of combinations of polyhedron faces generated by the stochastic program. Full coverage of the AFM tip corresponds roughly to adsorption of 6 enzyme dimers ($N=6$), to which the average number of active sites would lead to a force of (6.2 ± 1.9) nN, using the average force per enzyme of (1.6 ± 0.5) nN. The agreement with the experimental data is remarkable. For a fully covered AFM tip, Bueno et al. [26] measured forces varying from 3 to 8 nN, leading to a histogram of forces which most probable value is (6.1 ± 2.0) nN. These values are highlighted in line 3 of Table 2. It is also shown that the average number of active sites is smaller than N , as one should expect since some active sites are not exposed to the substrates and this was taken into account in our stochastic model. Moreover, with the theoretical results we may predict the number of enzymes covering the tip to account for the measured forces in the histogram, which was also indicated in Table 2.

4. Conclusions

In this paper we introduced a procedure that allows the estimation of the interaction force between the ACC enzyme and its substrate diclofop, as well as compare this value with experimental results from AFM experiments. This was achieved with steered

molecular dynamics (SMD), where the enzyme dimers were represented as rigid polyhedrons and the coverage of the AFM tip with enzymes was simulated using a stochastic process. The dimension of each polyhedron face was calculated using routines provided by the programs [23,24,31] and the interaction energy between the enzyme and AFM tip surface was assumed to be proportional to the contact area. From geometry considerations it was possible to determine the number (N) of enzymes attached to the AFM tip and the average number of active sites effectively interacting with the substrate. The remarkable agreement between the theoretical and experimental results demonstrates that the procedure proposed is useful to interpret AFM experiments.

Acknowledgements

This work was supported by CNPq – National Council for Scientific and Technological Development (CNPq – INCT, 573742/2008-1), FAPESP – São Paulo Research Foundation (FAPESP – INCT, 2007/05089-9, 2008/57859-5, 2010/00463-2, 2010/04599-6, 2013/09746-5), CAPES, FAPEMIG (CEX-APQ-02176-11) and nBioNet.

References

- [1] V.G. Rumayor, O.R. Iglesias, L.G. Cabezas, Aplicaciones de biosensores en la industria agroalimentaria, CEIM, Mdri, 2005.
- [2] D.P.E. Smith, Limits of force microscopy, Rev. Sci. Instrum. 66 (1995) 3191–3195.
- [3] F.L. Leite, C.C. Bueno, A.L. Da Róz, E.C. Ziemath, O.N. Oliveira, Theoretical models for surface forces and adhesion and their measurement using atomic force microscopy, Int. J. Mol. Sci. 13 (2012) 12773–12856.
- [4] A. Etchegaray, C. de, C. Bueno, O. Teschke, Identification of microcistin LR at the molecular level using atomic force microscopy, Quím. Nova 33 (2010) 1843–1848.
- [5] D.K. Deda, B.B.S. Pereira, C.C. Bueno, A.N. da Silva, G.A. Ribeiro, A.M. Amarante, et al., The use of functionalized AFM tips as molecular sensors in the detection of pesticides, Mater. Res. (2013) 683–687.
- [6] A. da Silva, D. Deda, A. da Róz, R. Prado, C. Carvalho, V. Viviani, et al., Nanobiosensors based on chemically modified AFM probes: a useful tool for metsulfuron-methyl detection, Sensors 13 (2013) 1477–1489.
- [7] D. Van Der Spoel, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2006) 1701–1718.
- [8] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, et al., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics 29 (2013) 845–854.
- [9] L.V. Kalé, M. Bhandarkar, M. Bh. R. Brunner, N. Krawetz, A. Shinozaki, et al., NAMD: a Case Study in Multilingual Parallel Programming, 1997, pp. 367–381.
- [10] B. Isralewitz, J. Baudry, J. Gullingsrud, D. Kosztin, K. Schulten, Steered molecular dynamics investigations of protein function, J. Mol. Graph. Model. 19 (2001) 13–25.
- [11] Y.M. Rhee, V.S. Pande, Multiplexed-replica exchange molecular dynamics method for protein folding simulation, Biophys. J. 84 (2003) 775–786.
- [12] A.V. Vargiu, P. Ruggerone, A. Magistrato, P. Carloni, Dissociation of minor groove binders from DNA: insights from metadynamics simulations, Nucl. Acids Res. 36 (2008) 5910–5921.
- [13] H. Zhang, B. Tweel, J. Li, L. Tong, Crystal structure of the carboxyltransferase domain of acetyl-coenzyme A carboxylase in complex with CP-640186, Structure 12 (2004) 1683–1691.
- [14] M.D. Devine, A. Shukla, Altered target sites as a mechanism of herbicide resistance, Crop Prot. 19 (2000) 881–889.
- [15] C. Delye, X.-Q. Zhang, C. Chalopin, S. Michel, S.B. Powles, An isoleucine residue within the carboxyl-transferase domain of multidomain acetyl-coenzyme A

- carboxylase is a major determinant of sensitivity to aryloxyphenoxypropionate but not to cyclohexanedione inhibitors, *Plant Physiol.* 132 (2003) 1716–1723.
- [16] H. Zhang, Z. Yang, Y. Shen, L. Tong, Crystal structure of the carboxyltransferase domain of acetyl-coenzyme A carboxylase, *Science* 299 (2003) 2064–2067.
- [17] Y. Sasaki, Y. Nagano, Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding, *Biosci. Biotechnol. Biochem.* 68 (2004) 1175–1184.
- [18] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, et al., The protein data bank, *Nucl. Acids Res.* 28 (2000) 235–242.
- [19] E.F. Franca, F.L. Leite, R.A. Cunha, O.N. Oliveira Jr., L.C.G. Freitas, Designing an enzyme-based nanobiosensor using molecular modeling techniques, *Phys. Chem. Chem. Phys.* 13 (2011) 8894–8899.
- [20] G.S. Oliveira, F.L. Leite, A.M. Amarante, E.F. Franca, R.A. Cunha, J.M. Briggs, et al., Molecular modeling of enzyme attachment on AFM probes, *J. Mol. Graph. Model.* 45 (2013) 128–136.
- [21] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, et al., Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Comput. Chem.* 19 (1998) 1639–1662.
- [22] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, et al., AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, *J. Comput. Chem.* 30 (2009) 2785–2791.
- [23] A. Herráez, Biomolecules in the computer: Jmol to the rescue, *Biochem. Mol. Biol. Educ.* 34 (2006) 255–261.
- [24] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph.* 14 (1996) 33–38.
- [25] H. Cheng, N. Ji, Y. Peng, X. Shen, J. Xu, Z. Dong, et al., Molecular characterization and tissue-specific expression of the acetyl-CoA carboxylase α gene from Grass carp, *Ctenopharyngodon idella*, *Gene* 487 (2011) 46–51.
- [26] C. Castro Bueno, A. Moraes Amarante, G.S. Oliveira, D. Kotra Deda, O. Teschke, E. de Faria Franca, et al., Nanobiosensor for diclofop detection based on chemically modified AFM probes, *IEEE Sens. J.* 14 (2014) 1467–1475.
- [27] L.C.G. Freitas, Program Recover, Department of Chemistry, UFSCar, Fortran Code, 2013.
- [28] S. Izrailev, S. Stepaniants, B. Isralewitz, D. Kosztin, H. Lu, F. Molnar, et al., Steered molecular dynamics, in: P. Deuffhard, J. Hermans, B. Leimkuhler, A.E. Mark, S. Reich, R.D. Skeel (Eds.), *Computational Molecular Dynamics: Challenges, Methods, Ideas*, Springer, Berlin/Heidelberg, 1999, pp. 39–65.
- [29] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, *J. Chem. Theory Comput.* 4 (2008) 435–447.
- [30] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, The missing term in effective pair potentials, *J. Phys. Chem.* 91 (1987) 6269–6271.
- [31] F. Provenza, PRO-TEC-Desenhista de Máquinas, 46th ed., F. Provenza, 1991.

ANEXO B

Article

doi number

J. Braz. Chem. Soc., Vol. 00, No. 00, 1-8, 2015.
 Printed in Brazil - ©2015 Sociedade Brasileira de Química
 0103 - 5053 \$6.00+0.00

A Computational Protein Structure Refinement of the Yeast Acetohydroxyacid Synthase

Jéssica C. M. Ierich,^a Guedmiller S. Oliveira,^{*,a} Ana C. A. Vig,^a Adriano M. Amarante,^a
 Eduardo F. Franca,^b Fábio L. Leite^{*,a} and Yvonne P. Mascarenhas^c

^aGrupo de Pesquisa em Nanoneurobiofísica, Departamento de Física, Química e Matemática,
 Universidade Federal de São Carlos, 13052-780 Sorocaba-SP, Brazil

^bInstituto de Química, Universidade Federal de Uberlândia, 38400-902 Uberlândia-MG, Brazil

^cDepartamento de Física e Informática, Instituto de Física de São Carlos,
 Universidade de São Paulo, 13566-590 São Carlos-SP, Brazil

A combined molecular modeling and molecular dynamics simulation was carried out to obtain an improved description of the yeast acetohydroxyacid synthase (AHAS) in aqueous solution. After a thorough homology modeling, the AHAS catalytic dimer was subjected to a molecular dynamics (MD) simulation to analyze its behavior and optimize its geometry. The AHAS 3D molecular structure was analyzed according to the number of salt bridges and hydrogen bonds formed. During 20 ns of MD simulation, an average fluctuation of 3.9 Å was obtained. The cofactor thiamine diphosphate makes a relevant contribution to the system stability; this hypothesis was confirmed by the decrease in the average fluctuation of 0.3 Å. Moreover, the Ramachandran plot revealed no denaturation framework during the time of the simulation.

Keywords: acetohydroxyacid synthase, biomolecules, molecular structure description, computational modeling, molecular dynamics

Introduction

The experimental methods for determining the three-dimensional structure of biomolecules are essential for the study of their structure and, therefore, their functions.^{1,2} Presently, due to the technological development and improvement of computers, the amount of structural data is increasing, especially for proteins.^{1,3,4} However, several problems in experimental methods are still obstacles to the experimental determination of structures; this is particularly true in the application of X-ray diffraction crystallography (XRD), one of the main techniques for the determination of macromolecular structures. Problems related to protein purification, crystallogenesis, vitrification and harvesting of the crystal, obtaining phase information, structure refinement, flexibility and/or disorder of parts of the molecules in the crystal can directly affect the 3D structure description process.^{1,3} The last problem is particularly important in the determination of the positions of all residues in the protein, as in the case of the loops;

in the analysis of loops, a diffuse scattering is generated that is useless to the high resolution calculation of the 3D structure.⁵⁻⁷ In this context, the advances in computing and programming areas in recent decades have resulted in the creation of tools that make the generation of high quality structural models possible.⁸

The acetohydroxyacid synthase (AHAS) is of great relevance as a target in some of the main groups of enzyme-inhibiting herbicides, the sulfonylureas and imidazolinones.⁹⁻¹¹ AHAS is the first enzyme in the biosynthetic pathway of branched chain amino acids (valine, leucine and isoleucine) present in plants, seaweed, fungi, bacteria and archaea but not in animals. For this reason, the enzymes involved with the biosynthesis of such amino acids, especially AHAS, are potential targets in the development of agrochemicals (herbicides, fungicides and antimicrobial compounds).^{9,12} The molecular structure of AHAS involves two distinct subunits: (i) the catalytic subunit, which has all the apparatus necessary for the catalysis; and (ii) the regulatory subunit responsible for modulating the activity of the catalytic subunit.¹⁰ During its activity, AHAS demands the following cofactors: thiamine

*e-mail: fabioleite@ufscar.br, guedmuller@gmail.com

diphosphate (ThDP), flavin adenine dinucleotide (FAD), a magnesium ion (Mg^{2+}) in the active site^{10,12-14} and, generally, a potassium ion (K^+) to guarantee structural stability.^{11,15}

The great interest in this enzyme has resulted in a series of structural studies which have already described its three-dimensional structure.^{12,13,15-19} The smallest AHAS unit with catalytic activity is a dimer composed of two catalytic subunits, with the active site located at the interface between these subunits.⁹⁻¹¹ The dimeric structure described for the AHAS of *Saccharomyces cerevisiae* (protein data bank (PDB) 1JSC)¹³ includes some missing residues, namely, loop regions of undetermined position, giving a disordered conformation to these regions. Thus, the present study involves the use of molecular modeling and molecular dynamics (MD), as well as available experimental data for the enzyme to propose a full atomic three-dimensional structure description of the yeast AHAS enzyme functional dimer.

Methodology

Describing the molecular structure of the AHAS enzyme

The amino acid sequence of the AHAS (*S. cerevisiae*) was obtained online from the protein database of the National Center for Biotechnology Information (NCBI). The missing residues were added based on homology modeling protocols: (i) identification of proper templates by searching the structural database of protein sequences in the PDB⁴ using the basic local alignment search tool (BLAST) server²⁰ and protein BLAST (BLAST-P) algorithm;²¹ (ii) selection of templates based on the following criteria: highest percentage of sequence identity and similarity, cofactor integrity and proportion of missing residues on the disordered loops of AHAS molecular structure; (iii) alignment of the selected templates and target amino acid sequences as well as determination of three-dimensional AHAS structure by applying the Swiss-PDB viewer program;²² (iv) obtaining the AHAS missing loop structure by the SWISS-MODEL server;²²⁻²⁴ and (v) insertion of the cofactors according to the atomic coordinates of the templates.

Molecular dynamics simulation

To evaluate and to refine the AHAS molecular structure, after all adjustments, an MD simulation was carried out to provide some preliminary insights about AHAS behavior in aqueous solution. As a result, three input files were prepared in a box of water transferable intermolecular potential three point (TIP3P)²⁵ to evaluate their behavior

and interaction energies. The systems are summarized in Table 1 and shown in Figure 1.

Table 1. Description of the simulated systems

System ^a	No. of solute atoms	Counterion	Water molecule (TIP3P)	No. of atoms
1	17327	9	38368	132440
2	17497	5	38354	132564
3	17581	9	38350	132640

^aSystem 1: AHAS only; system 2: AHAS + FAD + Mg^{2+} and K^+ ions; system 3: AHAS + FAD + ThDP + Mg^{2+} and K^+ ions.

Figure 1. Simulated systems. (a) System 1: AHAS enzyme only. (b) System 2: AHAS and cofactors FAD, Mg^{2+} and K^+ . (c) System 3: AHAS and cofactors ThDP, FAD, Mg^{2+} and K^+ . Chains A and B are shown in green and orange, respectively, and cofactors ThDP, FAD, Mg^{2+} and K^+ are shown in red, blue, black and purple, respectively.

All MD simulations were performed by means of the nanoscale molecular dynamics (NAMD) 2.7 program^{26,27} using the following protocol: (i) minimization step at constant number of particles, system volume and temperature (NVT ensemble), where the system temperature was gradually increased in steps of 100, 200 and 298 K, until the energy stabilization; (ii) a cutoff distance of 12 Å was used, and long-range corrections were considered using the Ewald sum formalism;²⁸ (iii) in the equilibration step,

a Langevin piston²⁹ and thermostat were applied at 1.0 bar and 310 K, respectively, using the isothermal-isobaric with constant number of particles (NpT) ensemble for 20 ns. The force field parameters for all simulated systems were taken from the SwissParam web server³⁰ and added to the “chemistry at Harvard macromolecular mechanics” 27 (CHARMM27)³¹ force field. The results were analyzed using visual molecular dynamics (VMD)³² and Grace³³ programs.

System evaluation

To evaluate the structural quality of the simulated systems, a Ramachandran diagram³⁴ was plotted. A Ramachandran diagram assessment is based on permitted values of dihedral angles ϕ (phi) and ψ (psi) of residues in the protein structure. Three parameters were analyzed: (i) accuracy of the atomic coordinates of the AHAS molecular structure; (ii) denaturation of the AHAS during the MD simulation timescale; and (iii) impact of the inserted loops on the AHAS molecular structure. Ramachandran diagrams were calculated using the torsion angle plot program³⁵ at the StrucTools online server.³⁶

Results and Discussion

For a preliminary evaluation and validation of the AHAS refined model as well as to characterize the AHAS refined molecular structure behavior in aqueous solution, the structure and three systems trajectories, obtained as detailed in Methodology, were analyzed and are presented below.

Molecular modeling

Most studies reported in the literature involving AHAS three-dimensional structures refer to the dimeric form.^{9,11,37} The crystallographic structure of chain A from *S. cerevisiae* AHAS (PDB 1JSC)¹³ and also chain A from *S. cerevisiae* AHAS in complex with the herbicide chlorimuron-ethyl (PDB 1N0H)³⁸ were 100% identical with the target sequence in BLAST, being highly rated according to this algorithm. The chain A structure from *Arabidopsis thaliana* AHAS (PDB 1Z8N),³⁹ considered one of the most representative structures of the enzyme,⁹ exhibited 43% identity with the target sequence. Nevertheless, the template selection was not only envisioned as the results presented by BLAST but also as the criteria concerning the integrity of cofactors and the lost portion of the residues in the crystallographic structure. Therefore, the structures chosen for modeling were PDB 1JSC, the only structure free from herbicides and with intact cofactors, and PDB 1Z8N, which shows a

minimal proportion of lost residues in the XRD structure determination. The *S. cerevisiae* AHAS dimer was built by considering that the geometry of the association of its two monomers is similar to the dimeric structure presented in the PDB 1JSC template.

The alignment between the selected template sequences and the *S. cerevisiae* AHAS was important to the construction of atomic coordinates of the conserved regions in the AHAS model. Figure 2a shows the superposition of the templates (structural alignment) in which the PDBs 1JSC and 1Z8N are in green and purple, respectively. By using this consideration, the enzyme core region is more conserved, and as a result, the superposition was attained successfully. Then, this result was used to build a preliminary model of the molecular structure of the AHAS (Figure 2b). However, as highlighted in Figure 2b, several interruptions in the polypeptide chain are present, and these interruptions correspond to external variable regions from the templates. Consequently, no structural superposition for these regions was feasible.

Figure 2. (a) Superposition of the template structures PDB 1JSC (green) and PDB 1Z8N (purple). (b) Preliminary model obtained from the alignment of the template and the target sequences (chain A shown in green and chain B shown in orange). The interruptions in the polypeptide chain are highlighted, indicating the absence of some loops.

The variable regions were modeled with the server SWISS-MODEL that completed the missing loops, which in turn had the anchor residues: serine (SER) 188-alanine (ALA) 198 (α), glutamic acid (GLU) 497-SER514 (β), asparagine (ASN) 795-threonine (THR) 801 (γ), ALA823-valine (VAL) 827 (δ), ALA919-isoleucine (ILE) 927 (ϵ), proline (PRO) 964-arginine (ARG) 968 (ζ) and phenylalanine (PHE) 1114-histidine (HIS) 1117 (η) (Figure 3b). In the end, the coordinates of the complete AHAS model were obtained with a .pdb extension. Based on the coordinates of the template structure PDB 1JSC, the cofactors were directly inserted into the .pdb file. As a result, the complete model obtained is shown in Figure 3a.

Figure 3. AHAS model after homology modeling. (a) Complete model obtained and cofactors (FAD shown in blue, ThDP shown in red, Mg^{2+} in black and K^+ in violet). (b) Seven modeled loops (protein in white and loops in purple).

Molecular dynamics simulation

After obtaining the simulated trajectories for the three systems described (Figure 1), some major parameters were analyzed to elucidate the behavior of the systems in solution. First, the evolution of the interaction energy (Coulomb and van der Waals) between the protein and water were monitored as a function of time. Table 2 shows that no significant disparities were observed in the energy average values between the systems. The energy values reached their equilibrium at a temperature of 310 K and pressure of 1 bar. The temperature used in the simulations is in the range in which the enzyme is active, according to data available in the enzymatic repository, Braunschweig enzyme database (BRENDA),^{40,41} from 30 to 40 °C.⁴²

Table 2. Coulomb, van der Waals and total energies involved during 20 ns of MD simulation

System	Interaction energy / (kJ mol ⁻¹)		
	Coulomb	van der Waals	Total
1	$(-1.15 \pm 0.01) \times 10^5$	$(-6.98 \pm 0.2) \times 10^3$	$(-1.23 \pm 0.01) \times 10^5$
2	$(-1.14 \pm 0.01) \times 10^5$	$(-7.09 \pm 0.2) \times 10^3$	$(-1.21 \pm 0.01) \times 10^5$
3	$(-1.17 \pm 0.01) \times 10^5$	$(-7.08 \pm 0.2) \times 10^3$	$(-1.24 \pm 0.01) \times 10^5$

To characterize the stability of the system, both hydrogen bonds between AHAS and water molecules and electrostatic interactions between oppositely charged amino acids (salt bridges) were analyzed. For the hydrogen bond analysis, a cutoff distance of 3.0 Å and a cutoff angle of 20° between donor and acceptor atoms were used.^{43,44} For the salt bridge analysis, a cutoff distance of 3.2 Å between the oxygen atoms of acidic residues and the nitrogen atoms of basic residues was employed.^{45,46} Both hydrogen bonds and the formation of salt bridges are presented in Table 3.

Table 3. Hydrogen bonds and salt bridges formed during 20 ns of MD simulation for the three analyzed systems

System	time / ns	No. of hydrogen bonds between AHAS and water molecules	No. of salt bridges
1	0	778	111
	20	840	110
2	0	768	123
	20	822	110
3	0	740	114
	20	813	100

Concerning the hydrogen bonds, an average of 809, 794 and 777 bonds were formed between the enzyme and water molecules in systems 1, 2 and 3, respectively. The initial value for hydrogen bonds was smaller compared with the final value for all systems analyzed because the systems do not interact effectively with water molecules at the beginning of the simulation. As the simulation proceeds, the number of hydrogen bonds increases as a consequence of the system solvation.

As was expected, the aspartic acid (ASP), GLU and lysine (LYS) residues were the main contributors to the hydrogen bond formation in the three simulated systems because ASP and GLU are negatively charged, and LYS is positively charged. As a result, the side chains having the amino acids mentioned can be stabilized by forming hydrogen bonds in aqueous solution. Cysteine (CYS), ILE and tryptophan (TRP) amino acids contribute much less to the formation of hydrogen bonds due to their hydrophobic character and their being more internalized in the structure of the enzyme.

An important observation in systems 2 and 3 was the number of hydrogen bonds between AHAS and the FAD cofactor. Although the function of this cofactor in the structure and the enzymatic catalysis are still not clear for the scientific community,^{9,11,37} the cofactor's atomic coordinates and interactions with the enzyme residues are well defined. In simulation, an average value of 10 hydrogen bonds was observed between the FAD and the AHAS for both systems 2 and 3. In the literature, Pang *et al.*¹³ determined the position of the FAD cofactor in the crystallographic structure of AHAS from *S. cerevisiae*, which is strongly associated with the enzyme through 12 hydrogen bonds, established with specific residues that form the binding site of this cofactor. McCourt and Duggleby⁹ revised all structural information available on the AHAS enzyme (bacterial, fungal and plant) and reported that FAD is linked to the enzyme through 7 hydrogen bonds and another 42 non-bonded interactions. This way, the

average number of hydrogen bonds between AHAS and FAD obtained in the simulation trajectory is in agreement with literature data, which can be used as a validation parameter for the final model.

Concerning the salt bridge formation, the initial and final values are close to each other during the entire trajectory, although, in the initial nanoseconds of the simulations of systems 2 and 3, a higher number of salt bridges were observed, stabilizing the initial structure of the protein. A decrease in the number of these salt bridges were observed in the simulation, which could be attributed to the interactions between water molecules and the charged amino acids (ASP, GLU and LYS). Observing the progression in time of the average number of hydrogen bonds to these systems (Table 3), the number of salt bridges was inversely proportional to the number of hydrogen bonds, which could indicate the influence of the solvation of the molecule. Franca *et al.*⁴⁶ also had similar results with simulations of the acetyl-CoA carboxylase, in which a decrease in the number of salt bridges was observed during the simulation. This decrease could be attributed to the formation of new hydrogen bonds by charged amino acids, causing a reduction in the number of salt bridges and fluctuations in the three-dimensional structure. These fluctuations in all systems were evaluated along the timeline using the root mean square deviation (RMSD) of the positions of the α carbons in the initial and final structures presented in Figure 4.

Figure 4. Time evolution of the RMSD of AHAS atoms from the initial structure.

According to Figure 4, system 2 has the highest conformational fluctuation. In system 2, the RMSD varies considerably, with a fluctuation of approximately 3.9 ± 0.3 Å. In contrast, systems 1 and 3 have a smaller conformational variation than the variation shown in system 2, reaching a value of 3.6 ± 0.2 Å after 6 ns.

The smaller stability observed in system 2 may be attributed to two factors: (i) insufficient simulation time

for the system to reach a required stability, and (ii) the absence of the ThDP cofactor in the simulations, which plays a role in the stabilization. These factors, especially the factor presented in (ii), can be studied better through the system 3 simulation, which involves the polypeptide chain of AHAS, with all cofactors (FAD, ThDP and K^+ and Mg^{2+} ions). This system presents smaller conformational fluctuations; after 6 ns of simulation, some variations occurred at approximately 3.5 ± 0.2 Å. Thus, the data presented by this system suggest the importance of the ThDP cofactor in the AHAS enzymatic stability. In the absence of ThDP and the presence of Mg^{2+} ion, a greater variation was verified, probably because the Mg^{2+} has an anchor function towards the ThDP.

For the modeled loops, an average fluctuation of 11.4 ± 0.6 Å was observed for the three evaluated systems in the protein structure. Because the loops are externally located, the flexibility can be attributed to the solvation effects when compared with the protein core regions. In this context, the RMSD is in agreement with the expected values.

System evaluation

The initial and final conformations (2 and 20 ns) of the three simulated systems were analyzed in terms of the ϕ and ψ torsion angles in the Ramachandran diagram (Figure 5). The evolution of this graphic indicated that considerable changes between the torsion angles ϕ and ψ did not occur after 20 ns of simulation. Despite the presence or absence of the cofactors, no denaturation was observed in the AHAS model.

Regarding the quality of the atomic coordinates, it was verified using the evaluation of the 1136 enzyme residues that a great portion of its residues concentrated in the most favorable region (core) and additionally allowed regions for all systems, whereas few residues were located in the generously allowed and disallowed regions. In the diagrams presented, the MD simulation can be observed to change the position of several residues from generously allowed and disallowed regions to allowed regions. However, given the motion in the enzymatic structure in the aqueous environment, other residues were able to move to allowed and disallowed regions over time. By analyzing the interrupting regions after inclusion of missing residues and 20 ns of MD minimization and equilibration, it is possible to infer that the fluctuation range within acceptable values refers to external protein regions. Therefore, the model presented is in agreement with a full atomic model description for the proteins, and the model presented is suggested as an addition to an online protein repository.

Figure 5. Ramachandran diagram for: (a) the initial (2 ns) and (b) the final (20 ns) instances of MD simulation.

The quality of a model in homology modeling depends on several factors, including the quality of the template structure used.⁴⁷⁻⁴⁹ Thus, the templates used were submitted for evaluation by the server StrucTools, aiming to obtain data for comparison. The residues from the 1JSC and 1Z8N templates concentrated almost completely into the most favorable and additionally allowed regions, presenting only two residues in the generously allowed regions and none in disallowed regions. Therefore, the residues in the most favorable and allowed regions of all three systems are presumed to be in agreement with the values of the templates used. Nevertheless, the three systems showed an average of three residues in the disallowed regions, which is not justified by the quality of the templates used and deserves further investigation.

Conclusions

An improvement in the determination of the AHAS atomic coordinates was obtained by choosing accurate template files; as a result, the position of all missing

crystallographic residues could be determined. The model quality was in agreement with the crystallographic template structures, therefore indicating that no damage to the AHAS three-dimensional structure occurred after loop modeling. Despite this agreement, some residues of the AHAS model occupied unfavorable regions in the Ramachandran plot. This result can be explained by the presence of variable regions in the template files. In both the presence and the absence of the cofactors, no denaturation evidence was verified. The presence of all cofactors confers stability onto the AHAS molecular structure, whereas the absence of the cofactors leaves empty space in an important AHAS region, consequently conferring instability onto the model.

In summary, the structural refinement of the AHAS enzyme was obtained according to its accuracy, and a well-coordinated atomic position was obtained. The molecular modeling used in the study provided structural characterization of specific and highly disordered regions of protein loops. This study is a prerequisite to the evaluation of the AHAS molecular structure as an enzyme-sensor to probe its possible inhibitors because the enzyme stability is fundamental to the development of a precise and accurate nanobiosensor.⁵⁰⁻⁵²

Acknowledgments

This work was supported by the Higher Education Improvement Coordination Agency (CAPES); São Paulo Research Foundation (FAPESP) under grants FAPESP-INCT 2007/05089-9, 2008/57859-5, 2010/00463-2, 2010/04599-6, 2013/09746-5, 2013/21958-8, 2014/12082-4, 2014/12466-7 and 2014/26369-3; National Council for Scientific and Technological Development (CNPq) under grant CNPq-INCT 573742/2008-1; Minas Gerais Research Foundation (FAPEMIG) under grant CEX-APQ-02176-11; Rede Mineira de Química (RQ-MG), nBioNet; and CNPq YPM fellowship. The authors would like to thank Prof James M. Briggs from the University of Houston, TX, USA, for use of the Texas Learning & Computation Center (TLC2) supercomputer cluster facilities.

References

1. Garman, E. F.; *Science* **2014**, *343*, 1102.
2. Rhodes, G.; *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models*, 3rd ed.; Elsevier: San Diego, 2006.
3. Yonath, A.; *Curr. Opin. Struct. Biol.* **2011**, *21*, 622.
4. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E.; *Nucleic Acids Res.* **2000**, *28*, 235.

5. Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; Ausio, J.; Nissen, M. S.; Reeves, R.; Kang, C.; Kissinger, C. R.; Bailey, R. W.; Griswold, M. D.; Chiu, W.; Garner, E. C.; Obradovic, Z.; *J. Mol. Graphics Modell.* **2001**, *19*, 26.
6. Linding, R.; Jensen, L. J.; Diella, F.; Bork, P.; Gibson, T. J.; Russell, R. B.; *Structure* **2003**, *11*, 1453.
7. Dong, G. Q.; Fan, H.; Schneidman-Duhovny, D.; Webb, B.; Sali, A.; *Bioinformatics* **2013**, *29*, 3158.
8. Bordoli, L.; Schwede, T. In *Homology Modelling: Methods and Protocols*; Orry, A. J. W.; Abagyan, R., eds.; Humana Press: New York, 2012, pp. 107-136.
9. McCourt, J. A.; Duggleby, R. G.; *Amino Acids* **2006**, *31*, 173.
10. Duggleby, R. G.; McCourt, J. A.; Guddat, L. W.; *Plant Physiol. Biochem.* **2008**, *46*, 309.
11. Duggleby, R. G.; Pang, S. S.; *J. Biochem. Mol. Biol.* **2000**, *33*, 1.
12. McCourt, J. A.; Pang, S. S.; King-Scott, J.; Guddat, L. W.; Duggleby, R. G.; *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 569.
13. Pang, S. S.; Duggleby, R. G.; Guddat, L. W.; *J. Mol. Biol.* **2002**, *317*, 249.
14. Tittmann, K.; Schröder, K.; Golbik, R.; McCourt, J.; Kaplun, A.; Duggleby, R. G.; Barak, Z.; Chipman, D. M.; Hübner, G.; *Biochemistry (Moscow)* **2004**, *43*, 8652.
15. Wang, J.-G.; Lee, P. K.-M.; Dong, Y.-H.; Pang, S. S.; Duggleby, R. G.; Li, Z.-M.; Guddat, L. W.; *FEBS J.* **2009**, *276*, 1282.
16. Pang, S. S.; Duggleby, R. G.; Schowen, R. L.; Guddat, L. W.; *J. Biol. Chem.* **2004**, *279*, 2242.
17. McCourt, J. A.; Pang, S. S.; Guddat, L. W.; Duggleby, R. G.; *Biochemistry (Moscow)* **2005**, *44*, 2330.
18. Kaplun, A.; Vyazmensky, M.; Zherdev, Y.; Belenky, I.; Slutzker, A.; Mendel, S.; Barak, Z.; Chipman, D. M.; Shaanan, B.; *J. Mol. Biol.* **2006**, *357*, 951.
19. Karanth, N. M.; Sarma, S. P.; *Biochemistry (Moscow)* **2013**, *52*, 70.
20. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T. L.; *BMC Bioinf.* **2009**, *10*, 421.
21. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J.; *J. Mol. Biol.* **1990**, *215*, 403.
22. Guex, N.; Peitsch, M.; *Electrophoresis* **1997**, *18*, 2714.
23. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T.; *Bioinformatics* **2006**, *22*, 195.
24. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C.; *Nucleic Acids Res.* **2003**, *31*, 3381.
25. Wu, Y.; Tepper, H. L.; Voth, G. A.; *J. Chem. Phys.* **2006**, *124*, 24503.
26. Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K.; *J. Comput. Phys.* **1999**, *151*, 283.
27. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K.; *J. Comput. Chem.* **2005**, *26*, 1781.
28. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G.; *J. Chem. Phys.* **1995**, *103*, 8577.
29. Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R.; *J. Chem. Phys.* **1995**, *103*, 4613.
30. Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O.; *J. Comput. Chem.* **2011**, *32*, 2359.
31. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caffisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodosscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M.; *J. Comput. Chem.* **2009**, *30*, 1545.
32. Humphrey, W.; Dalke, A.; Schulten, K.; *J. Mol. Graphics Modell.* **1996**, *14*, 33.
33. <http://plasma-gate.weizmann.ac.il/Grace/devell/> accessed in May 2015.
34. Ramachandran, G. N.; Venkatachalam, C. M.; Krimm, S.; *Biophys. J.* **1966**, *6*, 849.
35. Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M.; *Proteins: Struct., Funct., Bioinf.* **1992**, *12*, 345.
36. <http://helixweb.nih.gov/structbio/basic.html> accessed in May 2015.
37. Duggleby, R. G.; McCourt, J. A.; Guddat, L. W.; *Plant Physiol. Biochem.* **2008**, *46*, 309.
38. Pang, S. S.; Duggleby, R. G.; Schowen, R. L.; Guddat, L. W.; *J. Biol. Chem.* **2004**, *279*, 2242.
39. McCourt, J. A.; Pang, S. S.; King-Scott, J.; Guddat, L. W.; Duggleby, R. G.; *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 569.
40. Schomburg, I.; Chang, A.; Hofmann, O.; Ebeling, C.; Ehrentreich, F.; Schomburg, D.; *Trends Biochem. Sci.* **2002**, *27*, 54.
41. Schomburg, I.; Chang, A.; Placzek, S.; Sohngen, C.; Rother, M.; Lang, M.; Munaretto, C.; Ulas, S.; Stelzer, M.; Grote, A.; Scheer, M.; Schomburg, D.; *Nucleic Acids Res.* **2013**, *41*, D764.
42. <http://www.brenda-enzymes.org/enzyme.php?ecno=2.2.1.6> accessed in May 2015.
43. <http://www.ks.uiuc.edu/Research/vmd/plugins/hbonds/> accessed in May 2015.
44. Torshin, I. Y.; Weber, I. T.; Harrison, R. W.; *Protein Eng.* **2002**, *15*, 359.
45. <http://www.ks.uiuc.edu/Research/vmd/plugins/saltbr/> accessed in May 2015.
46. Franca, E. F.; Leite, F. L.; Cunha, R. A.; Oliveira, O. N.; Freitas, L. C.; *Phys. Chem. Chem. Phys.* **2011**, *13*, 8894.
47. Sánchez, R.; Sali, A. In *Protein Structure Prediction: Methods and Protocols*; Webster, D. M., ed.; Kluwer: New York, 2000, pp. 97-130.

48. Santos-Filho, O. A.; Alencastro, R. B.; *Quim. Nova* **2003**, *26*, 1.
49. Wallner, B.; Elofsson, A.; *Protein Sci.* **2005**, *14*, 1315.
50. Oliveira, G. S.; Leite, F. L.; Amarante, A. M.; Franca, E. F.; Cunha, R. A.; Briggs, J. M.; Freitas, L. C. G.; *J. Mol. Graphics Modell.* **2013**, *45*, 128.
51. Amarante, A. M.; Oliveira, G. S.; Bueno, C. C.; Cunha, R. A.; Ierich, J. C. M.; Freitas, L. C. G.; Franca, E. F.; Oliveira Jr., O. N.; Leite, F. L.; *J. Mol. Graphics Modell.* **2014**, *53*, 100.
52. Bueno, C. C.; Amarante, A. M.; Oliveira, G. S.; Deda, D. K.; Teschke, O.; Franca, E. F.; Leite, F. L.; *IEEE Sens. J.* **2014**.

Submitted: March 14, 2015

Published online: June 3, 2015

FAPESP has sponsored the publication of this article.

ANEXO C

MODELAGEM MOLECULAR APLICADA A NANOBIOSSISTEMAS

Adriano Moraes Amarante

*Grupo de Pesquisa em Nanoneurobiofísica, Universidade Federal de São Carlos, Campus Sorocaba,
Departamento de Física, Química e Matemática*

Guedmiller Souza de Oliveira

*Grupo de Pesquisa em Nanoneurobiofísica, Universidade Federal de São Carlos, Campus Sorocaba,
Departamento de Física, Química e Matemática*

Jéssica Cristiane Magalhães Ierich

*Grupo de Pesquisa em Nanoneurobiofísica, Universidade Federal de São Carlos, Campus Sorocaba,
Departamento de Física, Química e Matemática*

Richard André Cunha

Universidade Federal de Uberlândia, Instituto de Química

Luiz Carlos Gomide Freitas

Universidade Federal de São Carlos, Departamento de Química

Eduardo de Faria Franca

Universidade Federal de Uberlândia, Instituto de Química

Fábio de Lima Leite

*Grupo de Pesquisa em Nanoneurobiofísica, Universidade Federal de São Carlos, Campus Sorocaba,
Departamento de Física, Química e Matemática*

PALAVRAS-CHAVE: biomoléculas; estrutura tridimensional de proteínas; modelagem molecular computacional; difração de raios x; ressonância magnética nuclear; modelagem por homologia; dinâmica molecular; docking molecular; nanobiotecnologia; métodos híbridos de mecânica quântica/mecânica molecular, termodinâmica molecular.

RESUMO: A Modelagem Molecular é uma importante ferramenta para o estudo e o desenvolvimento de novas nanobiotecnologias. É amplamente aplicada na descrição e caracterização de sistemas biológicos, auxiliando a pesquisa e o desenvolvimento de novas tecnologias relacionadas com a saúde. O presente capítulo fornece ao leitor os princípios básicos, porém fundamentais, de simulação computacional e do uso do repositório de dados de biomoléculas online. Também, são apresentados os principais programas para manipulação e visualização de moléculas, além da teoria que fundamenta os principais métodos computacionais abordados.

1. INTRODUÇÃO

1.1. Modelagem Molecular

Dadas as propriedades físico-químicas da matéria e às leis da Física, é possível criar modelos que facilitem a compreensão e podem prever características intrínsecas de sistemas químicos e biológicos.¹ Assim, a modelagem molecular é um tipo de estudo que envolve a aplicação de modelos teóricos específicos a determinadas propriedades do sistema de interesse. Por intermédio do uso dos conceitos de átomo e molécula é possível prever o comportamento de sistemas reais aplicados nas áreas de Biologia, Química, Farmácia e Ciência dos Materiais.^{2; 3} Também é possível descrever as estruturas moleculares e suas propriedades físico-químicas e termodinâmicas utilizando a modelagem molecular.^{2; 4}

Com o uso de modelos teóricos na descrição das propriedades da matéria é possível a idealização do mundo real, porém, existem limitações no uso destes modelos. A modelagem molecular é uma ferramenta na qual é possível representar e manipular as formas estruturais de moléculas, propiciando o entendimento de reações químicas e fenômenos físicos, bem como o estabelecimento de relações entre a estrutura e as propriedades da matéria.¹ Para cada sistema de interesse, é possível modelar algo representativo, com suas características de interesse e extrair informações relevantes do sistema real. Para cada situação de estudo há um determinado modelo, adequado somente àquela situação. Assim, para uma determinada propriedade da molécula,

adequa-se um determinado modelo que evidencie essa propriedade, podendo existir diversos modelos para uma mesma molécula e diversos tipos de representações para um mesmo modelo. A ferramenta de simulação computacional também visa a construção de modelos advindos das propriedades físicas do objeto de estudo e acabam por definir o nível do conhecimento da estrutura da matéria.^{3; 5; 6; 7}

1.2. Modelagem Molecular Computacional

Com o objetivo de entender e prever o comportamento de sistemas reais, a modelagem molecular engloba um número de ferramentas e métodos computacionais que são usados para descrever e prever estruturas moleculares, propriedades de reações, propriedades termodinâmicas, entre outras.^{1; 8} Sendo assim, a modelagem molecular computacional é aquela que estuda as propriedades atômicas com o uso da química computacional e de técnicas de visualização gráfica. Este é um ramo no qual se estuda sistemas químicos e biológicos mais complexos, com um número elevado de partículas, da ordem de 10^{23} ; podendo realizar cálculos específicos e criar modelos que seriam impossíveis de serem obtidos de outra maneira.^{9; 10}

Com o avanço dos recursos computacionais, tratando-se de software e hardware, a modelagem molecular computacional se desenvolveu de maneira surpreendente nas últimas décadas. No passado, esta era restrita ao grupo de pesquisadores e cientistas responsáveis pelo desenvolvimento dos componentes de software utilizados. Nos dias atuais, diversos recursos e programas de modelagem estão disponíveis gratuitamente online, permitindo a comunidade científica explorar os dados físico-químicos das estruturas moleculares de sistemas de interesse.⁷ Alguns dos principais componentes de software disponíveis gratuitamente serão tratados posteriormente neste capítulo.

Cabe salientar que o campo da Química Computacional e da modelagem molecular computacional foi, de fato, reconhecido em 1998, acontecimento que acabou laureando John Pople e Walter Kohn com o prêmio Nobel em Química¹¹ pelas contribuições à química quântica e mais recentemente, em 2013, com a premiação a Martin Karplus, Michael Levitt e Arieh Warshel pela contribuição ao desenvolvimento de novos métodos para modelagem de sistemas contendo muitos átomos.¹²

2. TIPOS BÁSICOS DE REPRESENTAÇÃO

A descoberta do arranjo tetraédrico de carbono de compostos orgânicos em meados de 1874 foi o marco inicial da modelagem molecular, deixando de lado outras estruturas imagináveis das

moléculas e da matéria que se diferenciavam umas das outras basicamente por seu tamanho.¹³ A estrutura de Lewis, ou diagrama estrutural de pontos, surgiu em 1916 e foi proposta por Gilbert Lewis. Esta é utilizada até hoje como a mais básica forma de modelagem, enfatizando a sequência de átomos e tipos de ligação presentes em uma estrutura química ou biológica.^{14; 15}

Com a proposta da estrutura de Lewis e a descoberta de arranjos geométricos estruturais de determinadas moléculas, a modelagem molecular começa a tomar forma e diferentes tipos de representação começam a surgir. Barton, aproximadamente 60 anos após Gilbert Lewis, formula o conceito de análise conformacional, o qual se refere a moléculas sendo representadas por arranjos atômicos com energias específicas dependendo de sua localização espacial.^{3; 4} Com a cristalografia por difração de raios X (XRD) moderna (segunda metade do Século XX) e outras técnicas de análise estrutural, foi possível uma melhor compreensão das estruturas tridimensionais das moléculas, chegando a padrões de ligação, ângulos e outras propriedades intrínsecas de cada molécula.¹⁶

Todas essas informações sobre a estrutura das moléculas levaram ao desenvolvimento de tipos básicos de representação, utilizados na modelagem molecular. Três dos principais tipos básicos de representação são mostrados na Figura 7.1, na qual está presente a estrutura do butano, representado em bastões, bolas-e-bastões e volume de van der Waals, respectivamente.

Figura 7.1. Diferentes formas de representação molécula do butano: bastões, bolas-e-bastões e volume de van der Waals, respectivamente.

Os três modelos apresentados possuem vantagens na simplicidade, no desenho e na forma de apresentação. Destacam informações importantes de como a geometria molecular está disposta, assim como os átomos envolvidos. Os tipos de ligação entre os átomos são dados que permitem previsões teóricas sobre as propriedades físicas e químicas da matéria.¹⁴ Em algumas situações, os modelos podem possuir limitações por apresentarem um caráter qualitativo.^{9; 17} Outras representações que serão estudadas posteriormente fornecerão informações de caráter mais detalhadas, como tipo de enovelamento de proteínas, propriedades eletrônicas e de superfície.

3. BIOMOLÉCULAS E MODELAGEM DE PROTEÍNAS

3.1. Biomoléculas

As biomoléculas são compostos químicos sintetizados por seres vivos, dentre elas é possível destacar as proteínas, que apresentam estruturas amplamente estudadas por intermédio da modelagem molecular.^{18; 19} Devido ao seu alto grau de complexidade, são exigidos modelos mais dinâmicos que retratem suas propriedades químicas e físicas. Neste contexto, a modelagem molecular conta com poderosas ferramentas computacionais específicas para esse tipo de estrutura.²⁰

3.2. Proteínas

As proteínas são as macromoléculas biológicas mais abundantes em todas as partes das células.¹⁹ Elas são encontradas em uma enorme variedade de espécies e massa molecular, o que resulta em uma grande diversidade de funções biológicas. Uma proteína é definida como uma sequência ou polímeros de aminoácidos, sendo que cada resíduo se une por um tipo específico de ligação covalente chamada de ligação peptídica. As proteínas são compostas, em sua grande maioria, por apenas quatro elementos químicos: C, H, O e N.^{19; 21; 22}

Os aminoácidos são subunidades monoméricas relativamente simples que fornecem a estrutura básica das proteínas. Todas as proteínas, desde aquelas provenientes das mais simples e antigas cepas bacterianas, até aquelas componentes das mais complexas formas de vidas, são construídas a partir de um mesmo conjunto de 20 aminoácidos principais.^{17; 22} As células podem produzir proteínas com diversas propriedades e atividades, apenas pela união dos mesmos 20 aminoácidos em diferentes combinações e sequências.²² Consequentemente, os organismos podem produzir vários produtos como enzimas, hormônios, anticorpos, transportadores, fibras e elementos como penas, teias, chifres, entre outros. Entre estes produtos proteicos, as enzimas são as mais variadas e especializadas.^{19; 22}

Todos os 20 aminoácidos comuns são α -aminoácidos, ou seja, possuem um grupo carboxila e um grupo amino ligados ao mesmo átomo de carbono, determinado carbono α . Eles diferem uns dos outros em suas cadeias laterais ou grupos R, que variam em estrutura, tamanho e carga elétrica, além de influírem em sua solubilidade em água.^{19; 22} Modelos simples e de grande funcionalidade podem ser usados para representar estas estruturas, e neste caso, o modelo bolas-e-bastões é o mais indicado (Figura 7.2).

Figura 7.2. Representação da forma geral de um aminoácido.

A ligação entre o grupo amino e o grupo carboxila entre dois aminoácidos é chamada de ligação peptídica (Figura 7.3). Este processo também resulta na perda de uma molécula de água, gerando o que chamamos de resíduo dos dois aminoácidos iniciais.¹⁷

Figura 7.3. Representação esquemática da união entre dois aminoácidos por intermédio da ligação peptídica.

3.3. Estrutura das Proteínas

Dada a complexidade da estrutura proteica⁴, esta é comumente descrita por quatro níveis de estrutura básica. (Figura 7.4):¹⁸

- *Estrutura primária*: consiste de uma sequência de aminoácidos unidos por ligações peptídicas e inclui quaisquer ligações dissulfeto.²³
- *Estrutura secundária*: é o polipeptídeo resultante da estrutura primária e pode ser espiralado ou ter a estrutura de uma hélice.²³

- *Estrutura terciária*: várias hélices ou espirais que descrevem todos os aspectos do enovelamento tridimensional da cadeia polipeptídica.¹⁹
- *Estrutura quaternária*: ocorre quando duas cadeias polipeptídicas ficam unidas entre si.¹⁹

Figura 7.4.Tipos de modelagem de estruturas básicas formados pelas proteínas. Modelagem realizada a partir do arquivo PDBreferência 3HHB.²⁴ Adaptado de Nelson e Cox(2012).²²

A proteína representada na Figura 7.4 é a hemoglobina, uma proteína transportadora de oxigênio no sistema circulatório.²⁵ As junções das cadeias polipeptídicas nas proteínas ocorrem, em sua maioria, por ligações dissulfeto e ligações de hidrogênio. A estrutura tridimensional, em conjunto com a sequência de aminoácidos na molécula, conferem à proteína sua função específica.¹⁸ Algumas dessas funções específicas, muitas vezes, consistem em catalisar reações químicas. Desse modo, a proteína passa a pertencer a uma classe denominada enzima.^{19; 22}

3.4. Tipos de Representação de Biomoléculas e Proteínas

Como mencionado anteriormente, a modelagem molecular consiste em criar modelos teóricos que representem as estruturas e certas propriedades das moléculas. Sendo assim, diferentes propriedades, tais como, arranjo atômico, disposição espacial da molécula, grupos funcionais, entre outros, podem ser destacadas em diferentes modelos de representação. Os modelos têm como objetivo salientar certas propriedades importantes das moléculas, podendo, muitas vezes, tais propriedades serem criadas de maneira distante da realidade, mas que possam representar o objeto de estudo em si. Para elucidar melhor esses conceitos, diferentes formas de representação de estruturas de enzimas obtidas do repositório de estruturas cristalográficas, serão exemplificadas neste capítulo. Primeiramente, serão observadas representações para a enzima acetil-coenzima A carboxilase (ACCase), a qual participa da síntese de lipídeos em plantas e animais e é amplamente estudada pela indústria farmacêutica e em pesquisas tecnológicas.²⁶

O modelo representado na Figura 7.5 mostra a estrutura secundária da proteína ACCase. Este tipo de representação destaca as propriedades estruturais de uma proteína, arranjos e enovelamentos secundários, representados principalmente pelos enovelamentos hélices α e folhas β de forma espiral e setas planas, respectivamente.²³ Cabe salientar que as hélices α também são comumente representadas na forma de cilindros. O tipo de representação de estrutura secundária é um dos mais comuns quando se trata de representar proteínas, pois há uma grande importância em se estudar o seu arranjo tridimensional já que o mesmo fornece informações sobre o seu comportamento estrutural.

Figura 7.5. Representação da estrutura secundária da enzima ACCase. Modelagem realizada a partir do arquivo PDB referência 3FF6.²⁷

Diferentes tipos de representação podem criar um modelo bem diferente para uma mesma molécula. Na Figura 7.6, a mesma proteína é representada em forma de bastões. Este modelo permite o destaque de certos resíduos de aminoácidos específicos que compõem a proteína, e se pode atribuir ocasionalmente nome, número, ou ainda, diferentes cores a cada um destes. Esta representação é muito utilizada, principalmente quando se pretende destacar aminoácidos que participam de uma determinada ligação ou reação química.

Figura 7.6. Representação em forma de bastões da enzima ACCase. Os resíduos mais escuros indicam a entrada do sítio ativo enzimático. Modelagem realizada a partir do arquivo PDB referência 3FF6.²⁷

O modelo de bastões apresentado anteriormente é muito utilizado, quando se trata do estudo dos elementos químicos ou até mesmo alguns tipos de ligação em uma molécula. Na Figura 7.7 é possível observar, na enzima ACCase, o detalhe dos átomos que a constituem.

Figura 7.7. Representação na forma de bolas-e-bastões dos átomos constituintes da enzima ACCase. As diferentes cores indicam os elementos presentes (do mais claro ao mais escuro seguem hidrogênio, carbono, oxigênio e nitrogênio). Modelagem realizada a partir do arquivo PDB referência 3FF6.²⁷

Quando se trata de estudar efeitos de superfície, as representações anteriores não destacam propriedades como distribuição de cargas e massa, geometria espacial e possíveis sítios reacionais. A Figura 7.8 mostra uma destas possíveis representações de superfície. Nela foi utilizada a representação da distribuição de massa atômica (em unidade de massa atômica – u. a. m.) na superfície da enzima ACCase.

Figura 7.8. Representação da enzima ACCase com relação a distribuição de cargas superficiais. Com gradiente (1 a 16) u. a. m. Modelagem realizada a partir do arquivo PDB referência 3FF6.²⁷

Muitas enzimas possuem partes não proteicas, ou seja, partes que não são aminoácidos. Essas partes, denominadas cofatores, podem ser de caráter orgânico, no caso das coenzimas, ou

inorgânico. Dada a necessidade de representar em um mesmo modelo de estruturas com propriedades distintas, é comum o uso de representação mista, que é uma combinação das representações mostradas anteriormente. A Figura 7.9 mostra um exemplo de diferentes tipos de representação em um mesmo modelo para a enzima acetolactatosintase, conhecida como ALS. A ALS é uma proteína que participa da síntese de aminoácidos nas plantas, fungos e bactérias.²⁸ Neste modelo, foi representada a estrutura secundária para cada uma das duas estruturas terciárias da proteína; a representação de bastões para as coenzimas e a representação de bolas (sem escala) para os íons Mg^{2+} e K^+ .

Figura 7.9. Representação mista da enzima ALS. Modelagem realizada a partir do arquivo PDB referência 1T9D.²⁹

3.5. Bancos de Dados de Estruturas Biológicas

A descrição da estrutura de biomoléculas pode fundamentar o entendimento de suas propriedades, comportamento em sistemas biológicos e, conseqüentemente, a função específica desempenhada.²² Além disso, a configuração e o funcionamento molecular formam a base para o desenvolvimento de novas tecnologias como inibidores, medicamentos, vacinas, entre outras.³⁰ Assim, é de suma importância tornar a estrutura molecular da biomolécula acessível à comunidade científica, colaborando e incentivando estudos subjacentes.

Atualmente, são numerosas as informações de estruturas biomoleculares produzidas, reunidas, organizadas e disponibilizadas em bancos de dados específicos.

Em especial, as estruturas proteicas estão disponíveis numa gama de bancos de dados, os quais podem armazenar a estrutura primária (sequência de aminoácidos que constitui a proteína), informações sobre a função ou, ainda, conter as estruturas tridimensionais resolvidas experimentalmente. Dentre estes bancos de dados podem ser citados (i) o Universal Protein Resource (UniProt),^{31;32} que contém as sequências de aminoácidos que formam a proteína, a descrição de sua função, a caracterização dos domínios proteicos e as modificações pós-

traducionais; (ii) o Protein do National Center for Biotechnology Information,³³ que armazenam sequenciamento de proteínas; (iii) o sistema de informações enzimáticas BRENDA (Braunschweig EnzymeDatabase),^{3; 34; 35} que armazena um conjunto de dados, majoritariamente retirados da literatura, sobre função, estrutura, obtenção, preparação e aplicação de enzimas, assim como caracterização de mutações e variantes; (iv) o PROSITE,^{36; 37} que armazena dados proteicos os quais descrevem as famílias de proteínas, domínios e sítios funcionais (integrado ao UniProt); e (v) o Protein Data Bank (PDB),³⁸ o maior repositório de estruturas tridimensionais de biomoléculas obtidas por métodos experimentais.^{8; 39; 40}

Em especial, o PDB possui grande importância nos métodos de predição de estruturas proteicas por modelagem computacional, principalmente a modelagem comparativa ou por homologia, a qual fará uso de dados estruturais de proteínas armazenados no repositório em questão como referência para prever a estrutura de uma proteína-alvo. Após a determinação da estrutura cristalográfica, por XRD ou Ressonância Magnética Nuclear (NMR), o pesquisador disponibiliza os dados de sua pesquisa ao PDB, fornecendo as coordenadas tridimensionais obtidas, informações sobre a estrutura para que outros pesquisadores possam utilizar as coordenadas geradas pelo arquivo em formato PDB na obtenção de informações estáticas e dinâmicas da molécula.⁴¹ As buscas por modelos são realizadas num campo específico, podendo ser conduzida pelo nome da macromolécula, autor da pesquisa, código do arquivo PDB ou até mesmo pela entrada da sequência de aminoácidos da macromolécula de interesse. A estrutura molecular pode ser visualizada na própria página do banco de dados e os arquivos de estrutura PDB são abertos em programas específicos de edição ou processamento de texto. Os arquivos PDB podem ser lidos pelos programas biomoleculares e caracterizam a estrutura da macromolécula pelo conjunto de coordenadas tridimensionais de cada átomo constituinte da mesma.³⁰

3.6. Ferramentas de Visualização, Manipulação e Análise de Estruturas de Proteínas

Com o grande avanço tecnológico, acompanhado do aperfeiçoamento e sofisticação dos computadores, componentes de software foram aprimorados e tornaram-se importantes ferramentas para o estudo e resolução de problemas em pesquisas científicas.^{42; 43} No entanto, assim como é grande o número de dados de estruturas de biomoléculas disponíveis, também são numerosos os programas para leitura, exibição e/ou edição de dados estruturais. Estes podem não apenas editar e exibir a estrutura de macromoléculas, mas também auxiliar no estudo do comportamento da macromolécula no sistema biológico, suas interações com o meio, tal como com demais moléculas, como substratos enzimáticos e inibidores, por exemplo.

Todavia, além da aplicabilidade do programa na resolução do problema, é fundamental que este apresente uma interface simples para o usuário, de fácil entendimento, com clara exibição das funções e dinâmica, a fim de garantir uma comunicação eficiente computador-usuário. É muito importante que os programas tornem esta interação a mais facilitada possível, utilizando sequências de comandos simples para que o usuário tenha êxito na execução das funções. O papel computacional na edição, obtenção e visualização de estruturas proteicas é crescente, principalmente no emprego de métodos de predição estrutural de proteínas, os quais são realizados pela aplicação de um conjunto de etapas baseados na utilização de programas específicos.⁴³ Assim, dada a sua importância na visualização, manipulação e interpretação de estruturas proteicas, alguns dos principais programas serão abordados mais detalhadamente.

O Visual Molecular Dynamics (VMD),⁴⁴ desenvolvido e distribuído pelo Grupo de Biofísica Teórica e Computacional (Theoretical and Computational Biophysics Group), da Universidade de Illinois em Urbana e Champaign, e o Instituto Beckman, é um programa gratuito de visualização, modelagem e análise de estruturas tridimensionais de macromoléculas, como proteínas e ácidos nucleicos. É codificado em linguagem C e C++, com versões disponíveis para a maioria dos Sistemas Operacionais (Mac OS X, Unix e MS-DOS). Dispõe de um pacote de edição que inclui vários métodos de renderização e coloração de moléculas. Dada a capacidade do software em gerar imagens e filmes em alta definição, o mesmo permite analisar o comportamento da estrutura macromolecular ao longo do tempo em simulações por dinâmica molecular (MD), destacando sua vasta aplicabilidade em pesquisa científica. Ademais, apresenta uma interface gráfica com o usuário bastante intuitiva, podendo ser utilizado também como ferramenta de ensino.^{45; 46; 47}

Outro programa interessante para visualização e manipulação de estruturas tridimensionais de macromoléculas é o PyMol.⁴⁸ Desenvolvido por Warren Lyford DeLano, é um sistema multiplataforma escrito em linguagem Python e em C. A partir de 2006, algumas versões deixaram de ser distribuídas gratuitamente, exceto para uso acadêmico. Possui bom desempenho, em especial durante tarefas complexas, como representação de superfícies, mas pode requerer melhoria do hardware, como aceleração gráfica. Este software oferece recursos necessários para a geração de imagens de alta resolução, com ótima qualidade e riqueza de detalhes, e é muito utilizado em publicações científicas. Além disso, o programa fornece aplicabilidade em análises de MD, permitindo ao usuário a elaboração de vídeos sobre o comportamento molecular durante o tempo. No entanto, apresenta uma interface gráfica complexa, o que pode dificultar a utilização dos recursos disponíveis, considerando que algumas funções não são tão intuitivas.^{49; 50}

O software Jmol^{51; 52} também é muito utilizado em pesquisas científicas para visualização, análise e edição de modelos de biomoléculas. Codificado em linguagem Java, ele é compatível com grande parte dos Sistemas Operacionais. As ferramentas disponibilizadas ao usuário não exigem conhecimentos avançados na área de informática para sua utilização, já que sua interface gráfica é simples, permitindo o emprego descomplicado dos recursos e tornando o software uma poderosa ferramenta de apoio ao ensino.⁵³ A renderização de modelos tridimensionais dispensa a utilização de recursos avançados de hardware. Possui componentes de integração com outras aplicações Java, a qual garante maior funcionalidade ao software, como o componente JmolApplet, o qual integra o Jmol a páginas da rede, permitindo, por exemplo, a visualização direta de estruturas em bancos de dados. Além disso, é possível executar animações e vibrações das moléculas, ambas úteis em simulações por MD.⁵⁴

Outra opção para análise de estruturas moleculares, amplamente utilizada pela comunidade científica, é o software UCSF Chimera,⁵⁵ ou apenas Chimera. Foi desenvolvido pelo Grupo de Pesquisas em Biocomputação, Visualização e Informática (Resource for Biocomputing, Visualization, and Informatics) da Universidade da Califórnia, São Francisco. Este programa é gratuito para fins não comerciais, codificado em linguagem Python, com alguns cabeçalhos escritos em C e C++ e compatível com os Sistemas Operacionais MS-DOS, Mac OS X e Linux. Executa renderização de alta qualidade, apesar de, em alguns casos, exigir processamento gráfico mais intenso, dependente de aceleradores gráficos. Este fato, em partes, pode ser consequente da utilização da linguagem Python, a qual pode impactar negativamente a velocidade e também o uso da memória dos computadores. Todavia, tais problemas podem ser contornados com melhorias no hardware, além de mudanças algorítmicas na linguagem de programação do software.⁵⁵ Ademais, apresenta uma interface gráfica simples, com funções básicas evidentes, garantindo a utilização facilitada para iniciantes assim como praticidade aos usuários avançados.³³

O Swiss-PDBViewer, ou DeepViewer,⁵⁷ também é muito utilizado em tarefas de visualização e manipulação de estruturas de biomoléculas. Desenvolvido e fornecido gratuitamente pelo Instituto Suíço de Bioinformática (Swiss Institute of Bioinformatics), apresenta extrema versatilidade, reunindo várias ferramentas de estudo, comparação e busca de erros em modelos tridimensionais, com grande impacto na técnica de Modelagem Molecular por Homologia. Possui ampla integração com servidores da rede mundial de computadores, trabalhando em associação com o ExPASy (Expert Protein Analysis System) e fazendo buscas de homologia em bancos de dados internacionais.⁵⁷ Além disso, todas as funções são disponibilizadas numa interface gráfica intuitiva e simples, o que dispensa a necessidade de conhecimentos prévios em programação.^{30; 58}

4. MÉTODOS DE MODELAGEM MOLECULAR COMPUTACIONAL APLICADO A BIOMOLÉCULAS

4.1. Métodos Clássicos

4.1.1. Dinâmica Molecular (MD)

AMD é uma poderosa e versátil técnica para estudos de macromoléculas biológicas tendo como base a Mecânica Clássica. Experimentalmente é possível mensurar propriedades que são resultado da soma de todos os possíveis microestados entrópicos e energéticos que um conjunto de moléculas pode ocupar. Entretanto, mesmo com a constante evolução de metodologias experimentais, há a necessidade da utilização de modelos teóricos para previsão de propriedades e compreensão, em nível molecular, dos experimentos.^{59; 60; 61} Desta forma, dados experimentais podem ser utilizados na validação e no melhoramento das propriedades que podem ser medidas através de metodologias teóricas desenvolvidas para realizar previsões relacionadas com aquele sistema.

Para os sistemas moleculares compostos por centenas, milhares e até milhões de átomos, como proteínas em meio aquoso, não há a possibilidade de caracterizá-los com precisão em nível quântico. A física clássica, no entanto, oferece uma boa aproximação para a modelagem desses sistemas. Isto é devido ao fato de os orbitais moleculares serem bem definidos e, conseqüentemente, a diferença de energia entre o orbital molecular ocupado mais alto (em inglês *Highest Occupied Molecular Orbital* ou HOMO) e o orbital molecular não ocupado mais baixo (em inglês *Lowest Unoccupied Molecular Orbital* ou LUMO) é muitas vezes maior que $k_B T$ (sendo k_B constante de Boltzmann e T temperatura dada em Kelvin).⁶⁰

4.1.1.1 Dinâmica Molecular e Potenciais Relacionados

A primeira aproximação da Mecânica Molecular é considerar os átomos como esferas dotadas de carga com posições definidas no espaço por um conjunto de potenciais harmônicos. Desta forma, a evolução temporal da estrutura molecular depende dos termos lineares (V_d), angulares (V_θ) e torcionais (V_ϕ) que descrevem as interações ligadas da molécula a ser simulada.^{62;}

⁶³Da mesma forma, as interações não ligadas, ou interações intermoleculares, são descritas pelos potenciais de Lennard-Jones (V_L) ou van der Waals, e de Coulomb (V_C). A soma dos potenciais das interações ligadas e não ligadas resulta em uma expressão completa da energia potencial (V_{total}) do sistema que depende somente das posições atômicas (Equação 1).⁷

$$V_{total} = V_d + V_\theta + V_\phi + V_{LJ} + V_C \quad (1)$$

Em outras palavras, as moléculas são descritas a partir de um modelo de esferas indeformáveis dotadas de cargas ligadas por “molas” (potenciais harmônicos) devidamente calibradas, de modo que estas possam descrever as vibrações de ligações químicas e as barreiras energéticas associadas às transições conformacionais torcionais e angulares que reconstróem o comportamento da molécula a ser mimetizada. O somatório dos potenciais ligados e não ligados compõe o que é denominado campo de força do sistema.⁶³ Cada função de energia potencial, portanto, visa descrever um comportamento estrutural característico da molécula.

Potencial Harmônico de Ligação (ou Linear): Em métodos de mecânica molecular as ligações covalentes são descritas como um oscilador harmônico clássico. O estiramento entre dois átomos é definido pelo Potencial Harmônico Linear, V_d , o qual descreve a energia associada ao desvio da distância de equilíbrio d_0 (Equação 2).⁶⁴

$$V_d = \frac{1}{2} k_d (d - d_0)^2 \quad (2)$$

sendo d o comprimento da ligação entre dois átomos i e j (Figura 7.10) e k_d é a constante elástica da força.

A distância d_0 e a constante de força são os parâmetros livres (a serem determinados) do campo de força. Tais parâmetros podem ser obtidos experimentalmente por intermédio dos espectros de infravermelho ou pela mecânica quântica. A constante k_d varia de acordo com a rigidez da ligação química. Como exemplo, a constante k_d de uma ligação simples entre carbonos C–C é 317 kcal/mol Å², enquanto que a ligação dupla C=C é 570 kcal/mol Å².⁵⁹

Potencial Harmônico Angular: Este descreve a energia potencial relacionada aos desvios relativos ao ângulo de referência. Analogamente ao Potencial Harmônico Linear, o Potencial Harmônico Angular pode ser definido pela Equação 3.⁵⁹

$$V_\theta = \frac{1}{2} k_\theta (\theta - \theta_0)^2 \quad (3)$$

Potencial Torcional (ou Diedral): Considerando um conjunto de quatro átomos ligados, i , j , k e l , como mostrado na Figura 7.10, o ângulo diedro ϕ é definido como o ângulo entre os planos formados pelos átomos ijk e jkl .

Pode-se entender o potencial torcional V_ϕ como a representação explícita da barreira energética atribuída à rotação dos átomos i e l ao redor da ligação $j-k$, sendo o mesmo definido pela Equação 4.⁵⁹

$$V_\phi = k_\phi [1 + \cos(n\phi + \delta)] \quad (4)$$

sendo k_φ a constante que define a altura da barreira energética para uma possível torção, n o número de mínimos na curva de energia potencial pelo ângulo de torção, φ o ângulo diedral da ligação central e em relação a quatro átomos e δ a diferença de fase do ângulo diedral φ , que pode ser 0° e 180° , dependendo de este valor ser um ponto de máximo ou de mínimo.

Figura 7.10. Representação geométrica do ângulo de torção entre i, j, k e l . Adaptado de Cornell et al. (1995).⁵⁹

A determinação experimental de potenciais torcionais não é trivial. Na definição de tais parâmetros para novas moléculas, é comum se utilizar de potenciais diedrais semelhantes aos de uma molécula conhecida e previamente descrita. Outra estratégia para se determinar os parâmetros k_φ, n , e δ é otimizar os parâmetros para moléculas simples e extrapolar para moléculas complexas com grupamentos de átomos semelhantes. A curva de energia potencial pode ser obtida via cálculos mecânico-quânticos, o que é uma grande vantagem pela independência de valores experimentais.^{7; 64}

Potencial de Lennard-Jones (ou van der Waals): O potencial de Lennard-Jones é utilizado para descrever interações intra e intermoleculares do tipo van der Waals entre átomos ligados e não ligados. Este potencial é, geralmente, utilizado para descrever interações de átomos de uma mesma molécula se estiverem separados por três ou mais ligações covalentes.⁶³ A expressão da energia potencial que descreve este tipo de interação entre dois átomos (i e j) é dada pela Equação 5.⁶⁵

$$V_{LJ} = 4\varepsilon \left[\frac{\sigma^{12}}{r_{ij}^{12}} - \frac{\sigma^6}{r_{ij}^6} \right] \quad (5)$$

sendo r_{ij} a distância entre os átomos i e j (Figura 7.10), ε é a profundidade do poço de energia potencial e σ o diâmetro de Lennard-Jones, o qual depende das espécies químicas ligadas. Existem outras formas de abordagem para a descrição deste tipo específico de interação, no entanto, o formato da Equação 5 é o mais simples e geral.^{59; 65}

Potencial de Coulomb: As interações de carga que são puramente eletrostáticas entre dois átomos i e j (Figura 7.10) são descritas pela equação de energia potencial Coulombiana (Equação 6).⁵⁹

$$V_C = \frac{q_i q_j}{4\pi\epsilon_0\epsilon r_{ij}} \quad (6)$$

sendo r_{ij} a distância entre os átomos i e j , q_i e q_j as cargas dos átomos, ϵ_0 a permissividade no vácuo e ϵ a constante dielétrica do meio.

Faz-se necessário determinar vários parâmetros para a aplicação das equações que descrevem as energias potenciais para as interações presentes em sistemas moleculares. Como citado anteriormente, os parâmetros podem ser obtidos via cálculos ou técnicas experimentais como pela XRD, espectroscopia de infravermelho, NMR entre outros.⁵⁹

4.1.1.2 Métodos de Simulação por Dinâmica Molecular

As metodologias de MD são baseadas nas leis da mecânica clássica.²² As forças que atuam em cada átomo são baseadas no conjunto de parâmetros do campo de força, nas coordenadas e nas velocidades atômicas que são mapeadas durante todo o processo. Em uma simulação por MD, a força F atuante em determinado átomo é expressa em função do tempo, e é igual ao negativo do gradiente da energia potencial V em relação à posição r_i de cada átomo, tal como expressa na Lei de Newton (Equação 7).⁶⁶

$$-\frac{dV}{dr_i} = m \frac{d^2 r_i}{dt^2} \quad (7)$$

Em linhas gerais, nos métodos de MD, as equações clássicas de movimento são iterativa e numericamente integradas para cada átomo do sistema a um determinado tempo, e a cada passo, o processo de integração se repete considerando as posições atômicas e obedecendo a um determinado incremento no tempo (Δt) o qual é determinado pelo usuário e, em geral, está na escala de femtossegundos (10^{-15} s). Essas escalas de tempo tão pequenas são necessárias para que seja possível amostrar os modos normais mais rápidos, ou seja, a frequência de oscilação de ligações químicas. De modo simplificado, é possível entender que se o valor de Δt for grande, alguns eventos moleculares importantes, que ocorrem em escalas de tempo pequenas, não seriam amostrados, criando artifícios computacionais que resultam erros comprometedores para considerar qualquer resultado do modelo simulado. Tais artifícios podem gerar erros significativos no cálculo das forças, consequentemente na amostragem da distribuição canônica dos microestados do sistema.⁶⁷

Outro princípio básico da MD é a relação entre os dados teóricos e experimentais que pode ser feita partir do uso da Hipótese Ergódica de Boltzmann.⁶⁸ Com esta hipótese é possível mostrar que o valor médio de uma propriedade medida em um pequeno número de partículas por um longo tempo é equivalente às médias obtidas para um sistema de muitas partículas por um curto intervalo

de tempo. Baseado nesta observação, as simulações de MD que possuem poucas partículas (comparado a qualquer sistema real) devem possuir trajetórias suficientemente longas para que o sistema visite um número suficiente de conformações representativas para satisfazer este princípio. Esta justificativa teórica valida as médias obtidas de propriedades termodinâmicas de sistemas simulados por tempo suficiente num sentido ergódico.⁶⁷

Atualmente, existem vários algoritmos disponíveis que realizam a integração numérica das equações de movimento, sendo os mais comuns os do tipo *Verlet*.⁶⁵ Os algoritmos de *Verlet*, *Velocity-Verlet*⁶⁹ e *leap-frog*^{70; 71} são os mais utilizados, pois apresentam um custo computacional menor em termos de consumo de memória RAM e tempo de processamento.^{65; 72} Em geral, a escolha do algoritmo a ser usado se baseia em duas importantes variáveis: custo computacional e precisão na integração numérica.⁷³ O algoritmo *Velocity-Verlet* utiliza a velocidade das partículas para calcular a força e, com isso, atualiza a posição dos átomos. Deste modo, as forças inter e intramoleculares são calculadas a cada passo de MD, tornando computacionalmente mais conveniente o cálculo da energia cinética e potencial, em relação ao algoritmo de *Verlet*, por exemplo, que apenas depende das posições atômicas anteriores e atuais para determinar as próximas. Outros algoritmos conseguem ser bem mais precisos quanto à conservação da energia livre total do sistema, entretanto apresentam um custo computacional bem mais elevado, como o algoritmo de Beeman.⁷⁴ A condição para que seja escolhido um algoritmo mais leve (no sentido de processamento) a um mais preciso é a de que com menos custo computacional as simulações podem atingir escalas de tempo maiores, representando mais microestados entrópicos e energéticos do sistema, permitindo-se obter ergodicamente médias de propriedade que podem, seguramente, ser confrontadas com valores experimentais.⁷³

Atualmente, já é possível realizar simulações por MD que variam entre centenas de nanossegundos a alguns microsegundos, dependendo do número de átomos do sistema.⁷⁵ Entretanto, mesmo com a evolução temporal obtida pelos cálculos computacionais, as escalas de tempo simuláveis não são suficientes para obter informações completas de muitos fenômenos importantes, como *folding* de proteínas e enovelamento de DNA.⁷⁵ Algumas estratégias têm sido usadas visando aumentar o tempo de cada passo (Δt) obtendo-se, assim, simulações mais longas. Dentre os processos mais rápidos que ocorrem em sistemas moleculares estão as vibrações de ligação, em especial as ligações hidrogênio.⁷⁶ Como estas vibrações têm relativa influência em várias propriedades moleculares, alguns algoritmos mantêm estas ligações fixas, como o SHAKE,⁷⁷ RATTLE⁷⁸ e o LINCS.⁷⁹

Em paralelo a estas estratégias, há metodologias alternativas de MD que visam acelerar os processos moleculares modificando a distribuição de probabilidade dos estados do sistema, introduzindo um termo extra ao termo de energia potencial. Nos casos em que as informações do comportamento atômico podem ser descartadas, podem-se utilizar modelos *coarse-grained* que consideram grupos de átomos como partículas.⁸⁰ Estes grupos são apontados, via criteriosa análise do comportamento químico da molécula, respeitando suas características moleculares mais determinantes para dada propriedade. Além disso, é possível agrupar em uma única partícula um aminoácido ou até mesmo unidades monoméricas.⁸¹ Isso pode ser feito a partir de uma melhor amostragem dos estados conformacionais do sistema em questão, direcionando-o a outros mínimos de energia potencial alterando a probabilidade relativa de alguns estados que são definidos por variáveis que descrevam bem uma coordenada reacional. Algumas destas metodologias são *local elevation search*, *replica-exchange*, *parallel-tempering*, *simulated annealing*, *metadynamics*.^{82; 83; 84}

Outras problemáticas envolvidas nas simulações de MD são a utilização de solvente e do modelo a ser utilizado. O uso de solvente(s) em uma dada simulação aumenta o custo computacional desta; no entanto, permite a descrição precisa de várias propriedades que são diretamente dependentes da interação com solventes.⁷⁶ Os modelos de água mais utilizados são o SPC, TIP3P, TIP4P e TIP5P.^{85; 86; 87} O aumento do número de partículas em uma simulação pode levar a vários problemas na estimativa das forças entre as mesmas, devido a possíveis efeitos de superfície,⁶⁷ em outras palavras, efeitos de bordas do sistema devido à tendência das moléculas de colidirem com as paredes da caixa de simulação. Uma aproximação que contorna este problema é o uso de condições periódicas de contorno,⁸⁷ no qual o sistema é periódico em representação bidimensional, as partículas podem entrar ou sair da caixa de simulação atravessando qualquer uma das seis faces do cubo. Desta forma os efeitos de superfície são cancelados e a simulação se dá mimetizando o sistema em solução.

Em suma, simulações computacionais de sistemas com milhares de átomos podem ser realizadas com o uso de MD considerando fenômenos que ocorrem em alguns nanossegundos, sem a quebra ou formação de ligações químicas. Qualquer propriedade termodinâmica do sistema em questão pode ser obtida se a simulação considerar um número de estados energéticos suficientemente grandes para satisfazer a Hipótese Ergódica. A partir desta poderosa técnica podem-se analisar processos de adsorção, interação e o comportamento molecular conformacional de sistemas biológicos em geral.^{87; 88; 89; 90; 91; 92}

4.1.2. Docking Molecular

Docking molecular é um método que prevê a orientação preferencial de uma molécula com relação à outra quando ambas formam complexos estáveis.^{93; 94} Geralmente, escolhe-se uma molécula maior como receptora, por exemplo, uma enzima, e uma molécula menor como ligante. Essa molécula ligante pode ser um fármaco, um herbicida ou alguma molécula específica a qual se deseja estudar sua interação com a proteína.^{95; 96} Nos itens seguintes são apresentados as características deste tipo de cálculo.

4.1.2.1. Docking Molecular Receptor-Ligante

A metodologia de *docking* molecular receptor-ligante baseia-se na predição do modo de ligação de uma molécula menor, denominada ligante, na região de ligação uma macromolécula, denominada receptor. Com isso, é possível quantificar a energiaafinidade de ligação receptor-ligante.⁹⁷ A interação molecular proteína-ligante envolve combinações de efeitos entrópicos e entálpicos, previstos na Equação 8 por intermédio da energia livre de ligação de Gibbs (ΔG_{lig}), que por sua vez está relacionada com a uma constante de inibição experimental k_i .^{96; 98}

$$\Delta G_{lig} = \Delta H - T\Delta S = RT \ln k_i \quad (8)$$

sendo ΔH a variação de entalpia, T a temperatura absoluta, ΔS a variação de entropia e R a constante universal dos gases.

A metodologia de docking molecular tem grande aplicação no reconhecimento de fármacos, no estudo de doenças, no desenvolvimento de novos pesticidas e no estudo de sensores de detecção.^{10;}
⁹⁹ A Figura 7.11 mostra o exemplo de receptor-ligante com a enzima transcriptase reversa do vírus HIV-1 da Imunodeficiência Humana (HIV – Human Immunodeficiency Virus) complexada com fármaco DMP-266 (Efavirenz).⁹⁹

Figura 7.11. Enzima transcriptase reversa do vírus HIV-1 complexada com fármaco Efavirenz. Modelagem realizada a partir do arquivo PDB referência 1FK9.⁹⁹

A descrição do sistema receptor-ligante utilizando técnicas de modelagem molecular computacionais como o docking molecular tem como principal motivação a redução de tempo e dos altos custos envolvidos no desenvolvimento de novos medicamentos, sensores, herbicidas e outros fármacos e tecnologias.^{95; 96; 100}

4.1.2.2. Algoritmos Genéticos

O método de docking molecular requer técnicas computacionais com potencial para investigar um grande número de soluções possíveis, com o intuito de encontrar a solução ótima, ou seja, aquela que melhor atende as necessidades do sistema.^{96; 101} Os primeiros algoritmos de docking começaram a ser desenvolvidos no início da década de 1980, constituindo uma área de pesquisa bastante ativa.¹⁰¹ Dentre os algoritmos utilizados para a construção de modelos utilizando docking destacam-se os algoritmos genéticos. Algoritmos genéticos são algoritmos evolucionários pertencentes a uma classe de métodos estocásticos de otimização global que são inspirados no processo biológico de evolução das espécies e na teoria da sobrevivência do mais apto de Darwin.^{102;}¹⁰³ Os algoritmos são aplicados, principalmente, ao docking com ligantes flexíveis, ou seja, aqueles aos quais são atribuídos graus de liberdade de rotação no próprio ligante.¹⁰⁴ Novas configurações são geradas num processo denominado troca de “genes”, em que dois indivíduos geradores proporcionam mudanças aleatórias no valor dos genes transmitidos. O processo é iterativo, no qual a população evolui para estados energéticos favoráveis, até que um critério de parada seja satisfeito.^{102; 103} Atualmente, esses algoritmos são a ferramenta matemática mais utilizada pelos programas de *docking* molecular para executar cálculos de otimização de sistemas envolvendo receptor-ligante, de maneira a evoluir múltiplas subpopulações de ligantes.¹⁰⁵ No caso do programa AutoDock, um algoritmo genético adicional híbrido chamado Lamarckiano, realiza buscas locais.¹⁰⁰ A cada geração uma porção da população é escolhida ao acaso para a busca local; assim, o indivíduo escolhido substitui o anterior, fazendo uma alusão à teoria da hereditariedade de Lamarck.^{96; 104; 106; 107}

4.2. Métodos Híbridos (Mecânica Quântica/Mecânica Molecular)

4.2.1 Introdução aos Métodos Híbridos

A descrição quântica de biomoléculas pode ser feita através da combinação de métodos de modelagem molecular e mecânica quântica. Estes métodos híbridos são aplicados para descrição de interações específicas proteína-ligante e ligante-proteína e são comumente usados em sistemas muito grandes, tais como proteínas (mais de 10.000 átomos), fornecendo resultados mais consistentes.⁶³ O termo que se refere à combinação dos métodos de mecânica quântica e mecânica molecular é o QM/MM (Quantum Mechanics/Molecular Mechanics) que foi introduzido em 1976 por

Washel e Levitt, que apresentaram o primeiro modelo semiempírico aplicado a uma reação enzimática.¹⁰⁸ Com isso, ao longo dos anos este modelo foi utilizado como ponto de partida para os mais refinados métodos existentes nos dias atuais. Enquanto que métodos QM são usados para descrever o sítio ativo, onde as reações químicas ou excitações eletrônicas ocorrem, os métodos MM são empregados para capturar os efeitos do ambiente no sítio ativo.⁶³

A aproximação QM/MM é uma ferramenta valiosa para investigar sistemas grandes, como compostos inorgânicos, orgânicos, organometálicos e sistemas sólidos.^{109; 110; 111} Algumas práticas metodológicas e aplicações irão ser discutidas nas seções 4.2.2. e 4.2.6 respectivamente.

4.2.2 Simulação de sistemas biomoleculares

Estudos de sistemas grandes, tais como enzimas, requerem estruturas realísticas como ponto de partida. Isso é possível quando se usa estruturas moleculares advindas de técnicas experimentais de XRD e NMR.¹¹² Algumas pequenas modificações nas estruturas experimentais são frequentes, como, por exemplo, na troca de um inibidor por um substrato ou na substituição de resíduos específicos para gerar uma estrutura de partida mutante de interesse.¹¹³ Comumente, as estruturas geradas pelos experimentos são armazenadas em websites, que são banco de dados em 3D de proteínas e ácidos nucleicos (seção 3.5), e estas estruturas, na maioria das vezes, não estão completas, podendo causar erros quando usadas como ponto de partida em programas de simulação computacional. Nesse contexto, a adequação química da estrutura molecular se faz necessária através de programas que inferem, computacionalmente, a possível sequência de aminoácidos incompleta.

Critérios específicos são utilizados para a escolha dos métodos para descrição das regiões QM e MM. A região do QM é quanticamente tratada, enquanto que a região MM é descrita por campos de força.⁶² A interface destas regiões pode ser fixada numa região definida (Figura 7.12) ou se pode permitir que a fronteira se mova durante o curso da simulação, a fim de descrever processos como mudanças no sítio ativo e/ou movimento de íons solvatados.¹¹⁴ A escolha da região QM leva em conta a precisão e confiabilidade *versus* o desprendimento computacional.¹¹⁵ Apesar do custo computacional, os métodos *ab-initio* são os mais usados e continuam sendo os mais importantes em uma caracterização quântica.^{116; 117} Entretanto, o método que usa a Teoria do Funcional de Densidade (DFT)^{67; 118} é considerado um método alternativo e sofisticado em muitos estudos contemporâneos por fornecer resultados com alta precisão em um tempo computacional menor comparado aos métodos *ab-initio*. Na região do MM existem campos de força disponíveis para aplicações biomoleculares, como, por exemplo, o CHARMM^{119; 120}, AMBER^{121; 122}, GROMOS^{123; 124}, e OPLS^{125; 126}; e

para o tratamento de solventes explícitos, como a água, têm-se os modelos TIP3P ou SPC conforme descrito no item 4.1.1. Na Figura 7.12 é possível diferenciar as regiões QM e MM.

Figura 7.12. Descrição do sistema molecular com as interfaces das regiões QM e MM.

Uma das dificuldades no tratamento QM/MM é que os métodos MM não possuem parâmetros moleculares para a descrição de compostos organometálicos e sólidos contendo metais pesados. Esta limitação se deve principalmente pelo fato da região MM não ser polarizável, o que pode resultar em uma perda de precisão nos resultados do QM/MM.¹²⁷ Por outro lado, alguns campos de força polarizáveis vêm sendo desenvolvidos usando diferentes modelos clássicos, como dipolos, flutuações de carga, ou cargas em modelos de massa-mola.¹¹⁰

4.2.3 Programas para uso da metodologia QM/MM

As aplicações computacionais que empregam a metodologia QM/MM nos dias atuais estão disponíveis em diferentes pacotes computacionais e incluem os seguintes programas: MOPAC7,¹²⁸ GROMACS¹²⁹ compilado com código fonte do Gaussian03,¹³⁰ NWChem,⁷ TURBOMOLE,¹³¹ ORCA,¹³² e GAMESS-UK,¹³³ que usam campos de força como CHARMM,^{119; 120} GROMOS,^{64; 123; 124} AMBER,^{59; 121; 122} bem como os programas GULP,⁴⁹ DIADORIM⁵ e DICE^{134; 135}.

4.2.4 Fundamentação teórica e descrição matemática

Hamiltonianos clássicos e quânticos levam em conta diferentes propriedades do sistema, enquanto que o hamiltoniano clássico do sistema molecular clássico pode ser escrito como função da posição e do momento.⁶² Sendo H a energia total do sistema, as equações de movimento de Hamilton conduzem às equações de movimento de Newton; os hamiltonianos quânticos levam em conta as contribuições nucleares e de densidade eletrônica do sistema e na comparação, os hamiltonianos dos dois tipos de metodologia, são bem diferentes as contribuições de cada um, de acordo com as Equações 9 e 10.^{62; 63}

$$\hat{H} = \sum_{\text{ligações}} k_r (r - r_{eq})^2 + \sum_{\text{ângulos}} k_\theta (\theta - \theta_{eq})^2 + \sum_{\text{diedros}} \frac{V_\phi}{2} (1 + \cos(n\phi - \phi)) + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right] + \sum_{i < j} \frac{q_i q_j}{\epsilon R_{ij}} \quad (9)$$

O hamiltoniano clássico descreve: $H = E_{\text{covalente}} + E_{\text{não covalente}}$

Da esquerda para direita, na Equação 9, os termos se referem a ligações, ângulos, ângulos diedros, potencial de Lennard-Jones (forças de van der Waals) e forças eletrostáticas, respectivamente (item 4.1). Em contrapartida, o hamiltoniano quântico é descrito pela Equação 10.^{62; 63}

$$\hat{H} = - \sum_N \frac{\hbar^2}{2m_N} \nabla_N^2 - \sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_i \sum_j \frac{Z_i e^2}{4\pi\epsilon_0 |R_i - r_j|} + \frac{1}{4\pi\epsilon_0} \sum_{i < j} \frac{e^2}{|r_i - r_j|} + \frac{1}{4\pi\epsilon_0} \sum_{A < B} \frac{Z_A Z_B e^2}{R_A R_B} \quad (10)$$

No hamiltoniano quântico desta equação estão escritos, da esquerda para direita, as contribuições cinética (núcleo), cinética (elétrons), atração (núcleo-elétron), repulsão (elétron-elétron) e repulsão (núcleo-núcleo).

4.2.5 Interações Eletrostáticas

A descrição das interações eletrostáticas em sistemas biomoleculares pode ser feita através da mecânica quântica (resolução da equação de *Schrödinger* – Equação 10) ou, pela resolução da equação de Poisson-Boltzmann, a (Equação 11). Nesta equação temos informações a respeito da distribuição de carga em diferentes regiões das proteínas e dos aminoácidos que resultam nas interações de forças de longo alcance, as quais são essencialmente de natureza eletrostática.^{62; 63; 110}

Cabe ressaltar que para sistemas com mais de 10.000 átomos, o custo computacional do uso do método quântico nos cálculos se torna muito elevado e, por esse motivo, é estratégica a utilização desta abordagem somente para avaliação das interações eletrostáticas no sítio ativo e, no restante do sistema, usa-se o cálculo do potencial eletrostático através da resolução numérica da equação de Poisson-Boltzmann.¹³⁶

$$\vec{\nabla} \cdot [\epsilon(\vec{r}) \vec{\nabla} \psi(\vec{r})] = -4\pi\rho^f(\vec{r}) - 4\pi \sum_i c_i^\infty z_i q \lambda(\vec{r}) \exp \left[\frac{-z_i q \psi(\vec{r})}{k_B T} \right] \quad (11)$$

Sendo $\vec{\nabla}$ o operador divergente, $\epsilon(\vec{r})$ a dependência da posição dielétrica, $\vec{\nabla}\psi(\vec{r})$ o gradiente do potencial eletrostático, $\rho^f(\vec{r})$ a densidade de carga do soluto, c_i^∞ a concentração de íon i a uma distancia do infinito do soluto, z_i a carga do íon, q a carga do próton, k_B a constante de Boltzmann, T a temperatura e $\lambda(\vec{r})$ um fator para a acessibilidade da posição-dependente da posição de r acima dos íons em solução. Se o potencial não é tão grande se comparado ao valor de kT , a equação pode ser linearizada sendo resolvida mais eficientemente, levando a equação de Debye-Huckel, a qual tem a forma mais simplificada e pode ser descrita pela Equação 12.^{62; 63}

$$\vec{\nabla} \cdot [\epsilon(\vec{r}) \vec{\nabla} \psi(\vec{r})] = \epsilon k^2 \varphi(\vec{r}) - 4\pi\rho(\vec{r}) \quad (12)$$

sendo k a constante de Debye-Huckel ou blindagem iônica do meio. A forma da interface dielétrica e a exclusão de íons contidos no interior da proteína são introduzidos no cálculo através da dependência das constantes ϵ e k do vetor $\vec{r}$. Em condições onde $k = 0$, íons estão ausentes e a Equação de Debye-Huckel pode ser reduzida a equação de Poisson.^{62; 63}

4.2.6 Aplicações da QM/MM em sistemas moleculares

Os métodos híbridos de QM/MM não são válidos apenas para investigação de reações químicas no sítio ativo de sistemas grandes, mas também podem ser empregados na investigação de processos eletrônicos, tais como excitação eletrônica.¹¹⁰ Estudos biomoleculares típicos podem empregar cálculos de DFT/MM com um campo de força padrão, forças eletrostáticas e uma abordagem na qual são incluídos átomos de fronteira com a possibilidade de mudança de carga atômica.^{110; 137} O esquema QM/MM dependerá do nível de tratamento dado ao sistema e também do tipo de sistema. Com as metodologias combinadas QM/MM e com os dados obtidos das interações observadas na técnica experimental de XRD, é possível definir as contribuições por resíduo das interações intermoleculares, energia de interação do sistema e a obtenção da trajetória do sistema em 3D ao longo do tempo. Neste contexto, o planejamento racional de drogas por intermédio do uso de técnicas computacionais surge como uma alternativa para a síntese de novos fármacos, nas quais os métodos tradicionais de ataque não vêm alcançando resultados satisfatórios, incentivando novos estudos promissores.

A combinação dos métodos de QM/MM foi desenvolvida para cálculos com proteínas e vem sendo usado para o estudo de reações enzimáticas, especialmente em catálise enzimática, a qual

concerne na mudança da energia livre de ligação dos reagentes, estado de transição e produtos.²⁶ Neste método, o ligante é explicitamente tratado por um modelo QM,²⁶ enquanto que o ambiente proteico e o solvente são representados por campo de força MM.

As equações de dinâmica clássica são adequadas para muitos sistemas, entre eles os biomoleculares, pois através de um estudo de MD pode-se determinar as distintas conformações de um sistema, além de avaliar sua evolução no tempo.⁶³ A MD é uma técnica na qual se estudam os movimentos em um sistema de partículas por simulação e, com isso, pode ser empregada em sistemas de elétrons, átomos ou moléculas.¹³⁸ A primeira simulação de MD envolvendo proteínas foi feita em 1987.¹³⁹ Atualmente são encontrados muitos trabalhos na literatura envolvendo esse tipo de simulação com proteínas, principalmente quando se pretende confirmar dados experimentais tais como XRD e NMR. O potencial de interação entre as partículas e as equações de movimento que governam a dinâmica das mesmas são pré-requisitos para realização da MD. O potencial pode variar do simples, como o gravitacional para interações entre estrelas, ao complexo, composto por vários termos como o que descreve as interações entre átomos e moléculas.¹¹⁵ Para alguns problemas, como reações químicas envolvendo tunelamento, transferência de elétrons, ruptura e formação de ligações, são necessárias correções quânticas.¹³⁸ Dessa forma, o uso da metodologia QM/MM combinado com MD torna-se apropriado para o estudo de reações envolvendo sistemas enzimáticos.¹¹⁵

Um típico procedimento no emprego da metodologia QM/MM é o refinamento estrutural de sistemas biomoleculares grandes. A ideia básica é o uso do QM/MM em vez de somente o MM, no qual o modelo que é refinado contra os dados experimentais.¹¹⁴ Esta particularidade é vantajosa, pois ao redor do sítio ativo, os campos de força padrão são menos confiáveis para os inibidores e substratos.⁶² Essa aproximação tem sido aplicada para o refinamento dos dados de XRD e NMR.¹⁴⁰

4.3. Caracterização Estrutural

A determinação e divulgação da estrutura tridimensional de biomoléculas são essenciais para impulsionar o desenvolvimento de novos estudos e novas tecnologias. No entanto, o processo de determinação estrutural é extremamente complexo, o qual pode necessitar de métodos computacionais, escopo deste capítulo, e métodos experimentais avançados, intermediados por tecnologias sofisticadas, como a XRD e a NMR. O estudo para determinação estrutural de biomoléculas é realizado de forma a analisar a nanoestrutura do material, a qual envolve todos os átomos e arranjos atômicos que descrevem a conformação do sistema molecular.¹⁴¹ A aplicação dos métodos de determinação estrutural, tanto os experimentais quanto os de predição computacional,

culmina na geração de aproximações gráficas computacionais da estrutura de macromoléculas, tornando possível a construção de modelos computacionais em três dimensões pelo tratamento dos dados obtidos e/ou coletados.³⁰

4.3.1 Difração de raios X (XRD)

No Capítulo 5: Sistemas de Baixa Dimensionalidade de Nanopartículas do Volume 1: Nanoestruturas – Princípios e Aplicações, também da coleção Nanociência e Nanotecnologia, foi introduzido ao leitor alguns conceitos de XRD voltados à aplicações em nanomateriais. Para a descrição de estruturas de biomoléculas, a técnica de XRD também apresenta vasta aplicabilidade, sendo uma das principais técnicas de descrição da estrutura tridimensional, especialmente de proteínas³⁰. A aplicação eficaz da XRD para estudo estrutural de biomoléculas está fundamentada na alta organização atômica característica de um sólido cristalino. Na natureza, é possível que os átomos se organizem de tal forma a caracterizar sólidos amorfos, descritos pela ordenação atômica em curtas distâncias, não apresentando periodicidade; e os sólidos cristalinos, os quais são altamente organizados e caracterizados pela formação de padrões tridimensionais de ordenação atômica que se repetem periodicamente ao longo da estrutura do material.¹⁴¹ Estes últimos são de grande relevância científica para o estudo de macromoléculas por XRD, já que os raios X podem descrever especificamente o arranjo atômico que forma uma estrutura biológica.¹¹¹

Desde o início da utilização do fenômeno de XRD, que ocorreu com o alemão Max Von Laue, em 1912, o qual observou a capacidade de cristais em difratar os feixes de raios X segundo um padrão específico de organização atômica do material,¹⁴³ esta técnica vem sendo empregada no estudo e descrição estrutural de biomoléculas. Atualmente, a modernização e sofisticação dos difratômetros, aparelhos específicos para a aplicação da XRD (Figura 7.13A), assim como o desenvolvimento de novos métodos computacionais de interpretação e tratamento de dados cristalográficos, impactaram na cristalografia de forma a aumentar número de modelos macromoleculares obtidos. Nos dias de hoje, o repositório PDB conta com aproximadamente 103,265 estruturas resolvidas pela aplicação de XRD, sendo que, no ano de 2015, cerca de 8,680 estruturas haviam sido depositadas.¹⁴⁴ Alguns exemplos de estruturas recentemente disponibilizadas envolvem uma enzima transferase de bactérias em complexo com antibacterianos da classe dos aminoglicosídeos (PDB ID 5C4K), determinada por Kaplan e colaboradores (2016)¹⁴⁵; e também a estrutura da enzima metiltransferase do vírus da dengue em complexo com potenciais inibidores (PDB ID 5CUQ), resolvida por Brecher e colaboradores (2015)¹⁴⁶.

De uma forma breve, na estrutura cristalina, sendo o comprimento de onda dos raios X comparável ao espaçamento entre os átomos, estes podem ser difratados pelos mesmos, de forma a

descrever o posicionamento atômico.¹⁶ No difratômetro, os feixes de raios X incidentes são espalhados pela estrutura da amostra e a relação entre as fases dos feixes espalhados é que define a ocorrência da difração.¹⁴⁷ Se as ondas, após se dispersarem, ficarem fora de fase, estas se interferem de maneira destrutiva, se anulando. Em contrapartida se, posteriormente ao espalhamento as ondas permanecerem em fase, estas se interferem construtivamente, ou seja, as ondas se reforçam, caracterizando o fenômeno da difração (Figura 7.13B).^{16; 148}

Figura 7.13. Diagrama esquemático da XRD no difratômetro.

O diagrama mostrado na Figura 7.13 descreve a disposição dos componentes num difratômetro (A) e quando, após o evento de dispersão, os feixes de raios X permanecerem em fase, suas amplitudes são somadas e eles se reforçam, ocorrendo, assim, a difração (B). O feixe de raios X, gerado na fonte, é direcionado à amostra e os ângulos onde ocorreu a difração são registrados. A amostra e o detector são acoplados, de forma que uma rotação α da amostra equivalha a uma rotação 2α do detector.¹⁴⁹ Grande parte da radiação espalhada pela amostra se anula e a minoria dos feixes de raios X desviados se reforça, sofrendo a difração. O reforço mútuo dos feixes ocorre quando as condições satisfazem a Lei de Bragg, conforme descrito no Volume 1: Nanoestruturas – Princípios e Aplicações, Capítulo 5: Sistemas de Baixa Dimensionalidade de Nanopartículas desta coleção.

Quando difratados, os feixes informam algumas características estruturais da amostra (tipos de átomos que a constituem, arranjo dos mesmos e a geometria cristalina). Ao serem registrados pelo detector, estes geram o padrão de difração, ou difratograma, o qual informa os ângulos onde ocorreu a difração e a intensidade dos feixes difratados.¹⁵⁰

Os dados do difratograma são base para a determinação da estrutura do material. Com o tratamento destes dados obtém-se o modelo estrutural da molécula. Esta parte final de estudo

cristalográfico é descrita por três etapas principais: a resolução da estrutura, intermediada pelo emprego de métodos matemáticos para a determinação do espaço real na estrutura e reconstrução computacional da imagem da ordenação atômica na macromolécula; o refinamento do modelo, com base no mapa de densidade eletrônica da estrutura cristalina; e, finalmente, obtenção do modelo final contendo coordenadas atômicas da estrutura estudada.³⁰ Para uma revisão mais profunda dos conceitos de XRD, o leitor é convidado à leitura do Volume 3: Técnicas de Nanocaracterização – Princípios e Aplicações da coleção Nanociência e Nanotecnologia, Capítulo 5 – Difração e Espalhamento de Raios X por Nanomateriais.

4.3.2 Ressonância Magnética Nuclear (NMR)

Outra forma de determinação estrutural de biomoléculas envolve a aplicação da técnica de NMR. Nesta, as amostras geralmente são analisadas em solução e, por isso, o mais próximo possível das condições biológicas. Além disso, tal técnica dispensa a cristalização prévia da amostra, que é um passo crítico na técnica de XRD. Atualmente, a base de dados PDB já conta com cerca de 11,260 estruturas tridimensionais de proteínas e demais biomoléculas resolvidas por NMR e este número é crescente.¹⁵¹ No ano de 2015, cerca de 440 estruturas resolvidas por NMR foram depositadas no PDB, sendo algumas das mais recentes a estrutura de uma metaloprotease zinco dependente de *Clostridium difficile* (PDB ID 2N6J), determinada por Rubino e colaboradores (2015)¹⁵²; bem como a estrutura de uma chaperona bacteriana periplasmática ácido-resistente (PDB ID 2MYJ), determinada por Ding e colaboradores (2015)¹⁵³.

A NMR fundamenta-se na interação da radiação eletromagnética com a amostra. Mais precisamente, esta técnica utiliza a propriedade magnética apresentada por alguns núcleos atômicos da molécula para o estudo estrutural do material.^{154; 155} Determinados núcleos, como ^1H , ^{13}C e ^{15}N , apresentam o spin nuclear, uma propriedade descrita pela Mecânica Quântica cuja a orientação nuclear é dada por um momento angular característico.^{22; 156} Na ausência de campo magnético, o núcleo estará orientado de forma aleatória. No entanto, na presença de um campo magnético externo, o campo magnético gerado ao redor do núcleo faz com que este assuma uma orientação específica.¹⁵⁷ Quando num estado de menor energia, denominado estado α , a orientação nuclear é a favor do campo magnético aplicado. Entretanto, quando num estado de maior energia, denominado estado β , a orientação nuclear é contrária ao campo magnético externo.¹⁵⁷ A diferença de energia entre os estados α e β é proporcional à energia necessária para que um núcleo mude seu momento angular, ou seja, passe de um estado de menor energia para um estado de maior energia.¹⁵⁶ Essa diferença entre os estados α e β é proporcional à intensidade do campo magnético externo.¹⁵⁸ Sob a ação de um campo magnético intenso, a diferença de energia entre os estados será maior e,

consequentemente, a diferença entre o número de núcleos no estado α e no estado β também será maior.¹⁵⁶ Isso resulta num aumento da sensibilidade do espectrômetro e da resolução espectral, possibilitando aos aparelhos que empregam campos magnéticos intensos a descrição de sistemas moleculares mais complexos.¹⁵⁸

Num espectrômetro de NMR, a amostra em solução é submetida a um campo magnético de alta intensidade (B_h), o qual fará com que diferentes níveis energéticos sejam gerados, dadas as diferentes orientações dos núcleos alinhados ao campo. Em seguida, uma série de pulsos de energia eletromagnética na faixa da radiofrequência é aplicada, sendo estes pulsos capazes de tirar o sistema do equilíbrio. A radiação eletromagnética na radiofrequência fornecerá aos núcleos energia, o que levará à mudança de orientação dos mesmos. Na ausência de radiação na radiofrequência, o sistema irá se reestabelecer em equilíbrio, os núcleos voltarão à posição original e, para isso, emitirão radiação característica (Figura 7.14).^{159; 160; 161} Estas radiações emitidas são registradas pelo aparelho, originando o sinal de decaimento de indução livre (FID – Free Induction Decay).¹⁶² O sinal FID, dada a sua complexidade, sofre tratamento matemático adequado pela aplicação da transformada de Fourier, obtendo-se, assim, dados que gerarão um gráfico de intensidade da radiação ressonante *versus* frequência, característico do material.^{161; 163}

A frequência de emissão do núcleo atômico é dada pela relação entre o momento angular nuclear e o momento magnético nuclear: a denominada razão giromagnética. Contudo, esta frequência pode sofrer interferência do ambiente químico da biomolécula. Assim, para interpretar o espectro de NMR é adicionada à amostra uma substância de referência. A posição dos espectros da amostra é dada em relação à sua distância ao sinal de referência, caracterizando o deslocamento químico (valor dado em picômetros).¹⁵⁶ Pela interação entre o momento magnético nuclear e o campo magnético externo é possível caracterizar o ambiente ao redor do núcleo atômico de interesse e, assim, caracterizar a estrutura e a dinâmica da biomolécula em estudo. Portanto, do espectro de NMR é possível obter: (i) o ambiente químico do átomo estudado, via deslocamento químico; (ii) a quantificação do átomo em cada tipo de ambiente, pela área de cada sinal; e (iii) o número de átomos em ambientes químicos diferentes, pelo número de sinais diferentes quando comparados ao padrão (referência).^{160; 161}

A sequência da Figura 7.14 resume a técnica de NMR: (A) alguns núcleos atômicos possuem a propriedade spin que, na ausência de campo magnético externo, orientam-se aleatoriamente; (B) na aplicação de um campo magnético externo, os núcleos com $spin \neq 0$ assumem orientações específicas com diferentes níveis energéticos; (C) a radiação de radiofrequência é utilizada para perturbar o sistema, mudando a orientação dos núcleos; e, (D) ao retornar à posição original, o

núcleo libera o excesso de energia num sinal característico que dará origem ao espectro de NMR da amostra.^{159; 162}

Figura 7.14. Sequência de eventos dos núcleos atômicos em espectroscopia por NMR.

4.3.3 Modelagem de proteínas por homologia

Apesar da grande participação na determinação de estruturas tridimensionais de biomoléculas, os métodos experimentais, como XRD e NMR, podem apresentar alguns problemas que dificultam e, em certos casos, impossibilitam a sua utilização, como as dificuldades no isolamento e purificação da biomolécula de interesse, assim como na cristalização e obtenção de cristais de qualidade adequada à análise.^{142; 164} Assim, outros métodos de predição estrutural têm ganhado muita importância como ferramenta alternativa para descrever estruturas de biomoléculas, especialmente a modelagem por homologia, ou também modelagem comparativa, considerada a mais confiável dentre as técnicas de predição.^{164; 165; 166}

O grande objetivo das técnicas de modelagem de proteínas é a predição da estrutura tridimensional da mesma a partir de sua estrutura primária (sequência de aminoácidos formadora da proteína) de forma a obter modelos proteicos de qualidade comparável àqueles gerados pelas técnicas experimentais.¹⁶⁷ A modelagem por homologia baseia-se na premissa de que estruturas de proteínas homólogas mantêm alto grau de identidade, o que pode fundamentar a construção de modelos tridimensionais.¹⁶⁸ Isso se deve ao fato das proteínas homólogas evoluírem de um ancestral comum e que podem ter sofrido num dado momento alguma mutação, culminando em diferenças entre si.¹⁶⁹ Tais proteínas conservam numerosos trechos semelhantes ou idênticos, os quais podem ser categorizados como padrões. As regiões mais internas da estrutura proteica são mais conservadas. Já as regiões próximas às superfícies, as dobras ou “*loops*” que conectam elementos da estrutura secundária, são mais susceptíveis a alterações.¹⁷⁰ Desta forma, tais padrões de identidades formam a base para a construção da estrutura de uma proteína desconhecida a partir de outra homóloga cuja estrutura já fora determinada.¹⁶⁸

Em comparação com os métodos experimentais, a construção de um modelo tridimensional pela aplicação da modelagem por homologia dispensa rotinas laboratoriais de preparação e análise da amostra; todavia, exige programas e ferramentas computacionais especializados, assim como consultas a várias bases de dados disponíveis, já que utiliza dados estruturais pré-existentes para a modelagem.¹⁷¹ As etapas envolvidas no processo de modelagem por homologia, de uma forma geral, são: (i) busca e seleção de proteínas que servirão de molde; (ii) alinhamento das sequências da proteína-molde e da proteína-alvo; (iii) construção do modelo da proteína-alvo; e (iv) validação do modelo.¹⁷² Cada etapa da modelagem por homologia é baseada num método computacional específico.¹⁷²

A identificação e seleção dos moldes a serem utilizados no processo são auxiliadas por ferramentas que realizam buscas em bancos de dados estruturais por estruturas homólogas àquela de interesse. Um exemplo de ferramenta aplicada neste processo é o BLAST (Basic Local Alignment Search Tool), o qual encontra estruturas proteicas homólogas, identificando regiões similares entre as estruturas catalogadas, calculando o grau de identidade entre as mesmas.^{45; 173; 174} O programa Swiss-PDB Viewer, por possuir integração com servidores específicos, também pode auxiliar na busca por moldes. As estruturas mais adequadas (geralmente aquelas com maior grau de identidade) são escolhidas para compor os moldes da modelagem.¹⁷⁵

O modelo é construído com base no alinhamento realizado entre as sequências-molde e a sequência-alvo. As sequências são sobrepostas e, por meio do deslocamento e reposicionamento dos resíduos (com adição de espaços ou “*gap*”, representando deleções de resíduos) obtém-se simetria entre estas. Tal processo evidencia a identidade entre as sequências, indicando regiões conservadas e variáveis. A etapa de alinhamento das sequências é crucial na modelagem por homologia, interferindo diretamente na qualidade do modelo construído.¹⁷⁶ Com as sequências alinhadas, é iniciada a etapa de construção do modelo propriamente dito, a qual se inicia pelas regiões mais conservadas, com base nas informações estruturais da proteína-molde. Após, as regiões variáveis da proteína alvo, como as alças e as cadeias laterais, são modeladas de acordo com predições baseadas nas sequências de aminoácidos destas regiões ou por comparação com modelos de alças e cadeias laterais disponíveis no PDB.¹⁷⁷ Esta etapa pode ser subsidiada pelos programas abordados no item 3.5 deste capítulo, como o Swiss-PDB Viewer.

Com a obtenção do modelo para a proteína-alvo, é essencial a verificação de possíveis erros durante o processo de modelagem. A etapa de avaliação das características do modelo, a validação, envolve a busca por erros estruturais, assim como a avaliação da qualidade estereoquímica do modelo.¹⁷⁸ Além disso, a avaliação das interações entre o modelo obtido e o meio deve ser realizada

por simulações por MD, a qual acompanha o comportamento do sistema molecular ao longo do tempo.^{179; 180; 181} Tais avaliações podem ser intermediadas pelos mesmos programas utilizados para construção do modelo. Ao final, obtêm-se o modelo tridimensional predito da proteína.

5. ALGUMAS APLICAÇÕES RECENTES

5.1 Aplicações na Indústria Farmacológica

5.1.1. Desenvolvimento de Fármacos

A modelagem molecular computacional é uma importante ferramenta para o detalhamento dos mecanismos de reconhecimento molecular entre proteínas e inibidores, que podem ser a chave para o entendimento e o desenvolvimento de novos fármacos no tratamento de várias doenças. Assim, a modelagem molecular computacional tem um papel fundamental na farmacologia, assim como na medicina, por descrever estruturas de biomoléculas e agentes patológicos.^{182; 183} Ela tem sido um dos avanços mais importantes no planejamento e descoberta de novos fármacos, permitindo a obtenção de propriedades específicas de compostos que podem influenciar na interação proteína-ligante, no caso proteína-inibidor.

O exemplo do método de docking molecular citado no item 4.1 ilustra o local de inibição da enzima transcriptase reversa do vírus HIV-1 da Imunodeficiência Humana. O vírus causador é passível de sofrer mutações, por isso, até os dias atuais ainda é desconhecida a cura ou imunização profilática para o HIV, tornando essa doença um potencial alvo para o estudo de novas técnicas e concepção de novas drogas.^{184; 185; 186} Como visto na Figura 7.11, a enzima do Vírus HIV-1 da Imunodeficiência Humana foi estudada por intermédio do complexo proteína-ligante (fármaco Efavirenz). Neste nível de estudo, a modelagem molecular computacional é essencial, pois utiliza os métodos mencionados no item 4 para o desenvolvimento de novas drogas no combate a AIDS e outras doenças.^{99; 187}

5.1.2. Desenho Racional de Fármacos

A metodologia de triagem utilizada até a década de 1980 consistia em testar aleatoriamente fármacos em microcélulas, em uma espécie de “tentativa e erro” pouco eficiente na descoberta de novos fármacos.¹⁸⁸ Sendo assim, o conhecimento do mecanismo de ação proteína-inibidor é a uma das maneiras mais eficientes, e de menor custo, de se desenvolver novos fármacos.^{187; 189; 190} É possível evidenciar, dentre as metodologias de modelagem molecular computacionais mais eficientes, a combinação de técnicas de docking molecular, MD e a QM. Os resultados dessas combinações permitem a elaboração de índices de identidade que podem ser relacionados com a atividade farmacológica.¹⁹¹

Por meio da realização de diversos testes envolvendo um grande número de ligantes inibidores com tempo e custo muito inferior aos métodos práticos de triagem cega, a modelagem computacional auxilia no método de triagem de novos fármacos, acelerando consideravelmente o desenvolvimento de novas drogas.⁹⁹ Todos os métodos de modelagem moleculares computacionais combinados, juntamente com testes finais *in vitro* e *in vivo*, e testes clínicos dão origem ao desenvolvimento do fármaco de acordo com as etapas do desenho racional mostrado no diagrama da Figura 7.15.^{192; 193}

Figura 7.15. Fluxograma do Desenho Racional de Fármacos baseado em estrutura.

5.2. Aplicações Tecnológicas

5.2.1. *Nanoneurobiofísica: novas perspectivas para a aplicação da nanotecnologia nos estudos biomédicos*

Recentemente foi delineada uma nova área de conhecimento e pesquisa, denominada Nanoneurobiofísica, a qual visa a investigação de mecanismos moleculares importantes em distúrbios desmielinizantes e neurodegenerativos bem como o desenvolvimento de novos métodos de diagnóstico e tratamento para estas distúrbios com base na nanotecnologia.¹⁹⁴ Neste contexto, o detalhamento dos mecanismos que fundamentam as neurodistúrbios é realizado em escala molecular, propiciado pela aplicação combinada de métodos experimentais e teórico-computacionais. Mais precisamente, a aplicação de tal combinação metodológica para fins de diagnóstico e terapêutica é delineada pelo desenvolvimento de nanobiossensores, os quais são dispositivos sensíveis em escala nanométrica cujo sensoriamento é mediado por biomoléculas de alta especificidade, como enzimas e anticorpos.^{194; 195}

A modificação química das pontas e hastes de AFM bem como da superfície de nanopartículas com biomoléculas especializadas é um poderoso método o qual permite obtenção de um dispositivo de sensoriamento de alta especificidade^{195; 196; 197; 198; 199; 200; 201; 202}. Atualmente, vem sendo desenvolvido nanobiossensores para fins de detecção de contaminantes ambientais, bem como ensaios envolvendo a funcionalização de nanopartículas, abrindo novas possibilidades de aplicação de tais métodos no detalhamento molecular de doenças complexas, como a Esclerose Múltipla e a Neuromielite Óptica.^{91; 195; 198; 199; 200; 201; 202} Neste contexto, a aplicação da modelagem molecular computacional mostra-se promissora, especialmente nas etapas de desenvolvimento dos referidos biossensores. Estes nanodispositivos requerem uma visão atômica para que os resultados possam ser interpretados corretamente, o que é possibilitado pela aplicação de abordagens computacionais. Recentemente, estudos apontam para uma grande aplicabilidade de aproximações e abordagens computacionais como suporte para o entendimento de questões que cingem tanto a funcionalização de superfícies das pontas de AFM e de nanopartículas quanto as interações biomoleculares envolvidas no processo de detecção.^{89; 91; 195; 198; 199; 201; 203} Assim, dada a elevada sensibilidade característica dos nanobiossensores desenvolvidos é possível detectar com alta precisão herbicidas inibidores de enzimas, antígenos, toxinas e outros agentes patogênicos. Portanto, o uso combinado das técnicas computacionais e métodos experimentais torna possível o entendimento dos mecanismos moleculares envolvidos em um sistema de interesse.

5.2.2. Métodos computacionais aplicados ao desenvolvimento de nanobiossensores

Conforme abordado no item anterior, a aplicação de nanotecnologias de sensoriamento é um dos propósitos da emergente área da Nanoneurobiofísica.¹⁹⁴ Neste contexto, o desenvolvimento de nanobiossensores baseados na funcionalização de superfícies tanto do AFM quanto de nanopartículas para o estudo e detecção de doenças complexas pode ser embasado por métodos computacionais de DM e Modelagem.^{89; 91; 203}

De uma forma mais detalhada, num contexto de utilização de pontas e hastes de AFM bem como nanopartículas funcionalizadas e, conseqüentemente, construção de um nanobiossensor, o conhecimento e descrição específica dos eventos moleculares envolvidos mostram-se cruciais.^{89; 91; 200} Esses têm sido possibilitados pela aplicação de métodos computacionais de modelagem molecular, docking molecular, DM e cálculos híbridos de QM/MM desenvolvidos para fornecer um background dos tipos de interações envolvidas neste tipo de experimento.^{91; 195; 198; 199; 203} Desta forma, técnicas computacionais específicas estão sendo refinadas com o avanço das pesquisas, das técnicas experimentais bem como o desenvolvimento dos computadores, de modo a tornar possível o desenvolvimento racional de sensores específicos baseados em superfícies funcionalizadas.^{198; 204}

As aplicações computacionais no desenvolvimento de nanobiossensores vêm ocorrendo de acordo com uma série de vertentes, conforme segue: (i) aplicação de técnicas de modelagem por homologia no refinamento de estruturas descritas experimentalmente para estudos posteriores de imobilização e funcionalização de superfícies;^{89; 181} (ii) utilização de abordagens de docking molecular para o estudo de ligações receptor-ligante (enzima-substrato e/ou anticorpo-antígeno) durante o processo de sensoriamento;^{89; 195; 199; 203} (iii) parametrização de campos de força para a simulação de DM de moléculas envolvidas na imobilização e funcionalização de superfícies com compostos biológicos;⁹¹ (iv) simulações computacionais para estudo da estabilidade estrutural de sistemas biomoleculares proteicos^{89; 181; 195} bem como para a análise das orientações das biomoléculas durante a funcionalização de superfícies e conseqüente obtenção de nanobiossensores;^{91; 203} e (v) estabelecimento de protocolos para estudo da imobilização de biomoléculas em superfícies.^{91; 199; 203}

Desta forma, os métodos computacionais mostram-se poderosas alternativas para o detalhamento em alta resolução dos eventos moleculares envolvidos tanto processo de desenvolvimento de nanobiossensores quanto durante o processo de sensoriamento mediado por tais dispositivos. Para exemplificar, a Figura 7.16 mostra a modelagem molecular computacional da superfície de uma ponta do AFM com raio aproximado de 20 nanômetros, funcionalizada com enzimas ALS (enzima apresentada na Figura 7.9). As enzimas são modeladas destacando suas respectivas superfícies e fixadas na superfície da ponta por tióis orgânicos representados por

bastões. As enzimas ALS permanecem ativas tornando a superfície da ponta, como um todo, um nanobiossensor específico para detecção de herbicidas (*e.g.* metsulfuron-metil e imazaquin, utilizados mundialmente).^{205; 206} A modelagem molecular computacional, assim, permite a descrição aproximada da orientação que as biomoléculas assumem quando imobilizadas em superfície, formando o nanobiossensor propriamente dito.

Figura 7.16. Representação da superfície da ponta do AFM funcionalizada com enzimas ALS. Modelagem realizada a partir do arquivo PDB referência 1T9D.²⁹

6. COSIDERAÇÕES FINAIS

Neste capítulo foram apresentados conceitos básicos e históricos da modelagem molecular computacional. Foram abordados conceitos desde definições de biomoléculas e proteínas, passando por algumas das principais formas de representação de modelos proteicos, até exemplos de aplicação destes modelos pelo uso de técnicas computacionais avançadas. Também foram apresentados softwares de manipulação e visualização de biomoléculas, além de indicados bancos de dados onlinenos quais o leitor poderá consultar e obter estruturas previamente descritas de biomoléculas e sistemas biológicos.

As técnicas de modelagem molecular computacional aplicadas a biomoléculas englobam três principais métodos: Clássicos (MD e Docking Molecular), Híbridos (QM/MM) e métodos de Caracterização Estrutural (cálculo de homologia). Todas estas ferramentas e conhecimentos permitem ao leitor desenvolver o seu próprio estudo e modelagem de sistemas biológicos através de diferentes aproximações teóricas.

Por fim, as perspectivas futuras para a Modelagem Molecular Computacional são as mais promissoras possíveis. Com o aumento da capacidade de processamento computacional e o

desenvolvimento contínuo de softwares ligados a área, os modelos desenvolvidos para as biomoléculas serão mais fidedignos aos sistemas reais e a gama de aplicações será bastante extensa.

REFERÊNCIAS

1. Schlick, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide: An Interdisciplinary Guide*. 2nd ed. New York: Springer, 2010.
2. Casadesús, R. *Molecular Modeling*. In: Runehov, A.L.C.; Oviedo, L. *Encyclopedia of Sciences and Religions*. Dordrecht: Springer Netherlands, 2013. p. 1344–1346.
3. Berg, J.M.; Tymoczko, J.L.; Stryer, L. *Biochemistry*. 5th ed. New York: W H Freeman, 2002. 1100 p.
4. Rausch, F. *et al.* Protein modeling and molecular dynamic studies of two new surfactant proteins. *J Cheminform*, v. 5, n. 1, p. 1-11, 1 mar. 2013.
5. Freitas, L.C.G. DIADORIM: a Monte Carlo Program for liquid simulations including quantum mechanics and molecular mechanics (QM/MM) facilities: applications to liquid ethanol. *J. Braz. Chem. Soc.*, v. 20, n. 8, 2009.
6. National Research Council (US) Committee on Challenges for the Chemical Sciences in the 21st Century. *Beyond the Molecular Frontier: Challenges for Chemistry and Chemical Engineering*. 1st ed. Washington (DC): National Academic Press, 2003. 224 p.
7. Valiev, M. *et al.* NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, v. 181, n. 9, p. 1477-1489, set. 2010.
8. Scheer, M. *et al.* BRENDA, the enzyme information system in 2011. *Nucleic Acids Res*, v. 39, n. suppl 1, p. D670-D676, Jan 2011.
9. Ramachandran, K.I.; Deepa, G.; Namboori, K. *Computational Chemistry and Molecular Modeling: Principles and Applications*. 1st ed. Heidelberg: Springer, 2008. 398 p.
10. Bortolato, A. *et al.* *Molecular Docking Methodologies*. Monticelli, L.; Salonen, E. *Biomolecular Simulations: Methods and Protocols*. New York: Springer, 2013. p. 339-360.
11. Nobel Media AB. Press Release: The 1998 Nobel Prize in Chemistry. Disponível em: <http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1998/press.html>. Acesso em 22 jan. 2016.
12. Nobel Media AB. The Nobel Prize in Chemistry 2013 – Press Release. Disponível em: <http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2013/press.html>. Acesso em 23 ago. 2014.

13. Sudik, A.C. *et al.* Design, Synthesis, Structure, and Gas (N₂, Ar, CO₂, CH₄, and H₂) Sorption Properties of Porous Metal-Organic Tetrahedral and Heterocuboidal Polyhedra. *Journal of the American Chemical Society*, v. 127, n. 19, p. 7110-7118, 1 maio. 2005.
14. Lee, M.S. *et al.* New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations. *J ComputChem*, v. 24, n. 11, p. 1348-1356, 2003.
15. Young, D. *Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems*. 1st ed. New York: John Wiley & Sons, 2004. 408 p.
16. Callister, W.D.; Rethwisch, D.G. *Materials Science and Engineering: An Introduction*. 9th ed. Hoboken: Wiley, 2014. 984 p.
17. Brown, T.E. *et al.* *Chemistry: The Central Science*. 13th ed. Upper Saddle River: Prentice Hall, 2014. 1248 p.
18. Atkins, P.; Jones, L. *Chemical Principles: The Quest For Insight*. 5th ed. New York: W. H. Freeman, 2010.
19. Voet, D.; Voet, J.G. *Biochemistry*. 4th ed. Hoboken: Wiley, 2010. 1248 p.
20. Schueler-Furman, O. *et al.* Progress in Modeling of Protein Structures and Interactions. *Science*, v. 310, n. 5748, p. 638-642, Oct 2005.
21. Doig, A.J.; Baldwin, R.L. N- and C-capping preferences for all 20 amino acids in α -helical peptides. *Protein Sci.*, v. 4, n. 7, p. 1325-1336, 1995.
22. Nelson, D.L.; Cox, M.M. *Lehninger Principles of Biochemistry*. 6th ed. New York: W.H. Freeman, 2012. 1340 p.
23. Zvelebil, M.J.; Baum, J.O. *Understanding Bioinformatics*. 1st ed. New York: Garland Science, 2008. 772 p.
24. Fermi, G. *et al.* The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution. *Journal of Molecular Biology*, v. 175, n. 2, p. 159-174, May 1984.
25. Yokoyama, T. *et al.* Novel mechanisms of pH sensitivity in tuna hemoglobin: a structural explanation of the root effect. *The Journal of biological chemistry*, v. 279, n. 27, p. 28632-28640, July 2004.
26. Zhang, H.; Tweel, B.; Tong, L. Molecular basis for the inhibition of the carboxyltransferase domain of acetyl-coenzyme-A carboxylase by haloxyfop and diclofop. *Proceedings of the National Academy of Sciences of the United States of America*, v. 101, n. 16, p. 5910-5915, April 2004.

27. Madauss, K.P. *et al.* The human ACC2 CT-domain C-terminus is required for full functionality and has a novel twist. *Acta Crystallographica. Section D, Biological Crystallography*, v. 65, n. Pt 5, p. 449-461, May 2009.
28. Mccourt, J.A.; Duggleby, R.G. Acetohydroxyacid synthase and its role in the biosynthetic pathway for branched-chain amino acids. *Amino acids*, v. 31, n. 2, p. 173-210, Sept 2006.
29. Mccourt, J.A. *et al.* Elucidating the specificity of binding of sulfonylurea herbicides to acetohydroxyacid synthase. *Biochemistry*, v. 44, n. 7, p. 2330-2338, Feb 2005.
30. Rhodes, G. *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models*. 3rd ed. San Diego: Academic Press, 2010.
31. Apweiler, R. *et al.* Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nucleic Acids Res*, v. 41, n. D 1, p. D43-D47, 2013.
32. Wu, C.H. *et al.* The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Research*, v. 34, n. Database issue, p. D187-D191, Jan 2006.
33. Pruitt, K.D. *et al.* RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res*, v. 42, n. Database issue, p. D756-63, Jan 2014.
34. Schomburg, I. *et al.* BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. *Nucleic Acids Res*, v. 41, n. Database issue, p. D764-72, Jan 2013.
35. Schomburg, I. *et al.* BRENDA: a resource for enzyme data and metabolic information. *Trends Biochem Sci*, v. 27, n. 1, p. 54-6, Jan 2002.
36. Sigrist, C.J. *et al.* New and continuing developments at PROSITE. *Nucleic Acids Res*, v. 41, n. Database issue, p. D344-7, Jan 2013.
37. Sigrist, C.J. *et al.* PROSITE: a documented database using patterns and profiles as motif descriptors. *Brief Bioinform*, v. 3, n. 3, p. 265-74, Sept 2002.
38. Berman, H.M. *et al.* The Protein Data Bank. *Nucleic Acids Res*, v. 28, n. 1, p. 235-242, Jan 2000.
39. Laskowski, R.A. Protein Structure Database. *Mol Biotechnol*, v. 48, n. 2, p. 183-198, Jun 2008.
40. SIB – Swiss Institute Of Bioinformatics. UniProtKB/Swiss-Prot. Lausanne. 2013. Disponível em: <http://web.expasy.org/docs/swiss-prot_guideline.html>. Acesso em 20 mai. 2013.
41. Rose, P.W. *et al.* The RCSB Protein Data Bank: redesigned web site and web services. *Nucleic Acids Res.*, v. 39, n. Database issue, p. D392-D401, Jan 2011.

42. Küppers, G.; Lenhard, J.; Shinn, T. Computer Simulation: Practice, Epistemology, and Social Dynamics. In: Lenhard, J.; Küppers, G.; Shinn, T. (eds.). *Simulation: Pragmatic Construction of Reality*. 1. ed. Dordrecht: Springer, 2006. p. 3-22.
43. Winsberg, E. Computer Simulation and the Philosophy of Science. *Philosophy Compass*, v. 4, n. 5, p. 835-845, Sept 2009.
44. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J Mol Graph*, v. 14, n. 1, p. 33-38, Feb 1996.
45. Johnson, M. *et al.* NCBI BLAST: a better web interface. *Nucleic Acids Res*, v. 36, n. suppl 2, p. W5-W9, July 2008.
46. Stone, J. *et al.* Using VMD. Urbana e Champaign. 2012. Disponível em: <<http://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf>>. Acesso em: 19 mai. 2013.
47. Theoretical And Computational Biophysics Group. What is VMD? Urbana e Champaign. 2008. Disponível em: <http://www.ks.uiuc.edu/Research/vmd/allversions/what_is_vmd.html>. Acesso em: 19 mai. 2013.
48. De Lano, W.L. The PyMol Molecular Graphics System. San Carlos. 2002. Disponível em: <<http://www.pymol.org/>>. Acesso em 02 jul. 2013.
49. Gale, J.D. Gulp: A computer program for the symmetry-adapted simulation of solids. *J. Chem. Soc., Faraday Trans.*, v. 93, p. 629-637, 1997.
50. Stockwell, G. PyMOL tutorial. Cambridge. 2003. Disponível em: <<http://www.ebi.ac.uk/~gareth/pymol/>>. Acesso em: 19 mai. 2013.
51. Hanson, R. M. *et al.* JSmol and the Next-Generation Web-Based Representation of 3D Molecular Structure as Applied to Proteopedia. *Isr. J. Chem*, v. 53, n. 3-4, p. 207-216, 2013
52. Willighagen, E.L. Processing CML conventions in Java. *Internet Journal of Chemistry*, v. 4, p. 4, 2001.
53. Hanson, R.M. Jmol – a paradigm shift in crystallographic visualization. *J. Appl. Cryst.*, v. 43, n. 5, p. 1250-1260, Sept 2010.
54. Herraiez, A. Biomolecules in the Computer: Jmol to the Rescue. *Biochem Mol Biol Educ*, v. 34, n. 4, p. 255-261, July 2006.
55. Pettersen, E.F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, v. 25, n. 13, p. 1605-1612, Oct 2004.

56. Couch, G.S.; Hendrix, D.K.; Ferrin, T.E. Nucleic acid visualization with UCSF Chimera. *Nucleic Acids Res.*, v. 34, n. 4, p. e29, 2006.
57. Guex, N.; Peitsch, M.C. Swiss-Model and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, v. 18, n. 15, p. 2714-2723, Dec 1997.
58. Kaplan, W.; Littlejohn, T.G. Swiss-PDB Viewer (Deep View). *Brief Bioinform*, v. 2, n. 2, p. 195-197, May 2001.
59. Cornell, W. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.*, v. 117, n. 3, p. 5179-5197, Jan 1995.
60. Kalé, L.V. *et al.* NAMD: A case study in multilingual parallel programming. Li, Z. *et al.* (ed.). Languages and compilers for parallel computing. Heidelberg: Springer, 1998. cap. 26, p. 367-381. (Lecture Notes in Computer Science) ISBN 978-3-540-64472-9.
61. Woods, R.J. *et al.* Molecular Mechanical and Molecular Dynamic Simulations of Glycoproteins and Oligosaccharides. 1. GLYCAM_93 Parameter Development. *The Journal of Physical Chemistry*, v. 99, n. 11, p. 3832-3846, 1 mar. 1995.
62. Cramer, C.J. *Essentials of computational chemistry: theories and models*. 2nd ed. Chichester: John Wiley & Sons, 2004. 596 p.
63. Jensen, F. *Introduction to computational chemistry*. 2nd ed. Chichester: John Wiley & Sons, 2007. 599 p.
64. Koehler, J.E.; Saenger, W.; Van Gunsteren, W.F. A molecular dynamics simulation of crystalline alpha-cyclodextrin hexahydrate. *Eur. Biophys. J.*, v. 15, n. 4, p. 197-210, 1987.
65. Verlet, L. Computer "experiments" on Lennard-jones Molecules. I. Thermodynamical Properties. Defense Technical Information Center, 1966.
66. Brooks, C.L.; Karplus, M.; Pettitt, B.M. *Advances in Chemical Physics, Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*. New York: John Wiley & Sons, 1990.
67. Frenkel, D. Statistical mechanics for computer simulators. *Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, v. 49, p. 3-42, 1996.
68. Plato, J. Boltzmann's Ergodic Hypothesis. *Archive for History of Exact Sciences*, v. 42, n. 1, p. 71-89, 1991.
69. Swope, W.C. *et al.* A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters. *The Journal of Chemical Physics*, v. 76, n. 1, p. 637-649, 1982.

70. Snyder, J.A. A new and dynamic method for unconstrained minimization. *Applied Mathematical Modelling*, v. 6, n. 6, p. 449-462, Dec. 1982.
71. Snyder, J.A. An improved version of the original leap-frog dynamic method for unconstrained minimization: LFOP1(b). *Applied Mathematical Modelling*, v. 7, n. 3, p. 216-218, June 1983.
72. Alder, B. *Methods in Computational Physics*. New York: Academic Press, 1977.
73. van Gunsteren, W. F.; Berendsen, H.J.C. Algorithms for Macromolecular Dynamics and Constraint Dynamics. *Molecular Physics*, v. 34, n. 5, p. 1311-1327, 1977.
74. Beeman, D. Some multistep methods for use in molecular dynamics calculations. *Journal of Computational Physics*, v. 20, n. 2, p. 130-139, Feb 1976.
75. van Gunsteren, W.F. *et al.* Biomolecular modeling: Goals, problems, perspectives. *Angewandte Chemie – International Edition*, v. 45, n. 25, p. 4064-4092, 2006.
76. Berendsen, H.J.C. *et al.* Interaction models for water in relation to protein hydration. *Intermolecular Forces*, p. 331-342, 1981.
77. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H.J. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, v. 23, n. 3, p. 327-341, Mar 1977.
78. Vannucchi, G. *et al.* Rituximab treatment in patients with active Graves' orbitopathy: effects on proinflammatory and humoral immune reactions. *Clinical & Experimental Immunology*, v. 161, n. 3, p. 436-443, 2010.
79. Hess, B. P-Lincs: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation*, v. 4, n. 1, p. 116-122, Jan 2008.
80. Shelley, J. C.; Shelley, M. Y. Computer simulation of surfactant solutions. *Current Opinion in Colloid & Interface Science*, v. 5, n. 1-2, p. 101-110, 2000.
81. Modesto Orozco, L.O. *et al.* Coarse-grained representation of protein flexibility. *Foundations, successes, and shortcomings*. *Adv Protein Chem Struct Biol*, v. 85, p. 183-215, 2011.
82. Barducci, A.; Bonomi, M.; Parrinello, M. *Metadynamics*. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, v. 1, n. 5, p. 826-843, 2011.
83. Medarde Agustín, A. *et al.* Value of the detection of Australia antigen in blood donors. *Rev Sanid Hig Publica (Madr)*, v. 47, n. 6, p. 529-536, June 1973.

84. Trzesniak, D.; Lins, R.D.; Van Gunsteren, W.F. Protein under pressure: Molecular dynamics simulation of the arc repressor. *Proteins: Structure, Function, and Bioinformatics*, v. 65, n. 1, p. 136-144, 2006.
85. Berendsen, H.J.C.; Grigera, J.R.; Straatsma, T.P. The missing term in effective pair potentials. *J. Phys. Chem.*, v. 91, n. 24, p. 6269-6271, Nov 1987.
86. Jorgensen, W.L. *et al.* Comparison of simple potential functions for simulating liquid water. *J. Mol. Biol.*, v. 79, n. 2, p. 926-935, July 1983.
87. Mahoney, M.W.; Jorgensen, W.L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.*, v. 112, n. 20, p. 8910-8922, May 2000.
88. Degreve, L.; Fuzo, C.A. Structure and dynamics of the monomer of protein E of dengue virus type 2 with unprotonated histidine residues. *Genet Mol Res*, v. 12, n. 1, p. 348-59, Feb 2013.
89. Franca, E.F. *et al.* Designing an enzyme-based nanobiosensor using molecular modeling techniques. *Phys Chem Chem Phys*, v. 13, n. 19, p. 8894-9, mai. 2011.
90. Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*. 2nd ed. v. 1. Florida: Academic Press, 2001.
91. Oliveira, G.S. *et al.* Molecular modeling of enzyme attachment on AFM probes. *J Mol GraphModel*, v. 45, p. 128-36, Sept 2013.
92. Oliveira, O.V.D.; Santos, J.D.D.; Freitas, L.C.G. Molecular dynamics simulation of the GAPDH–NAD⁺ complex from *Trypanosoma cruzi*. *Molecular Simulation*, New York, v. 38, n. 13, p. 1124-1131, ago. 2012.
93. Morris, G.M.; Lim-Wilby, M. Molecular Docking. In: Kukol, A. (ed.) *Molecular Modeling of Proteins*. 1st ed. New York: Human Press, 2008. p. 363-382.
94. Garcia-Godoy, M.J. *et al.* Solving molecular docking problems with multi-objective metaheuristics. *Molecules*, v. 20, n. 6, p. 10154-10183, Jun 2015.
95. Lengauer, T.; Rarey, M. Computational methods for biomolecular docking. *Curr Opin Struct Biol*, v. 6, n. 3, Jun. 1996.
96. Morris, G.M. *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, v. 19, n. 4, p. 1639–1662, 1998.

97. Ng, P.C.; Henikoff, S. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*, v. 31, n. 13, p. 3812-3814, July 2003.
98. Brooijmans, N.; Kuntz, I.D. Molecular recognition and docking algorithms. *Annu Rev Biophys Biomol Struct*, v. 32, p. 335-373, 2003.
99. Ren, J. *et al.* Structural basis for the resilience of efavirenz (DMP-266) to drug resistance mutations in HIV-1 reverse transcriptase. *Structure*, v. 8, n. 10, p. 1089-1094, Oct 2000.
100. Morris, G.M. *et al.* Automated Docking with Protein Flexibility in the Design of Femtomolar "Click Chemistry" Inhibitors of Acetylcholinesterase. *J Chem Inf Model*, v. 53, n. 4, April 2013.
101. Kitchen, D.B. *et al.* Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov*, v. 3, n. 11, p. 935-949, Nov 2004.
102. de Azevedo; W.F. Jr. MolDock applied to structure-based virtual screening. *Curr Drug Targets*, v. 11, n. 3, 327-334, Mar 2010.
103. Heberlé, G.; de Azevedo; W.F. Jr. Bio-inspired algorithms applied to molecular docking simulations. *Curr Med Chem*, v. 18, n. 9, p. 1339-1352, 2011.
104. Morris, G.M. *et al.* AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of computational chemistry*, v. 30, n. 16, p. 2785-2791, Dec 2009.
105. Jones, G. *et al.* Development and validation of a genetic algorithm for flexible docking. *Journal of molecular biology*, v. 267, n. 3, p. 727-748, April 1997.
106. Melo, M.C. *et al.* GSAFold: a new application of GSA to protein structure prediction. *Proteins*, v. 80, n. 9, p. 2305-10, ago. 2012.
107. Halperin, I. *et al.* Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins*, v. 47, n. 4, p. 409-443, Jun 2002.
108. Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of molecular biology*, v. 103, n. 2, p. 227-249, 15 maio. 1976.
109. Freitas, L.C.G.; Longo, R.L.; Simas, A.M. Reaction-field–supermolecule approach to calculation of solvent effects. *J. Chem. Soc., Faraday Trans.*, v. 88, n. 2, p. 189-193, 1992.
110. Senn, H.M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed. Engl.*, v. 48, n. 7, p. 1198-1229, 2009.

111. Tchougréeff, A.L. Hybrid Methods of Molecular Modeling. 1st ed. New York: Springer, 2008. 344 p.
112. Feng, W.; Pan, L.; Zhang, M. Combination of NMR spectroscopy and X-ray crystallography offers unique advantages for elucidation of the structural basis of protein complex assembly. *Sci China Life Sci*, v. 54, n. 2, Feb 2011.
113. Grotendorst, J. *et al.* Hierarchical Methods for Dynamics in Complex Molecular Systems. Forschungszentrum Jülich, Germany: IAS Series, 2012. v. 10.
114. Heyden, A.; Lin, H.; Truhlar, D.G. Adaptive Partitioning in Combined Quantum Mechanical and Molecular Mechanical Calculations of Potential Energy Functions for Multi-scale Simulations. *J PhysChem B*, v. 111, n. 9, p. 2231-2241, Mar 2007.
115. Leach, A.R. Molecular modelling: principles and applications. 2. ed. Londres: Pearson Education, 2002. 744 p.
116. Friesner, R.A. Ab initio quantum chemistry: methodology and applications. *Proc Natl Acad Sci U S A*, v. 102, n. 19, p. 6648-6653, May 2005.
117. Friesner, R.A.; Guallar, V. Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu Rev Phys Chem*, v. 56, p. 389-427, 2005.
118. Sholl, D.; Steckel, J.A. Density Functional Theory: A Practical Introduction. 1st ed. Hoboken: Wiley, 2009. 252 p.
119. Brooks, B.R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem*, v. 4, n. 2, 187-217, 1983.
120. Brooks, B.R. *et al.* CHARMM: The Biomolecular Simulation Program. *J Comput Chem*, v. 30, n. 10, p. 1545-1614, Jul 2009.
121. Case, D.A. *et al.* The Amber biomolecular simulation programs. *J Comput Chem*, v. 26, n. 16, p. 1668-1688, Dec 2005.
122. Case, D.A. *et al.* AMBER 2015, University of California, San Francisco, 2015.
123. van Gunsteren, W.F.; Berendsen, H. J. C. Groningen Molecular Simulation (GROMOS) Library Manual, Biomos, Groningen, The Netherlands, 1987, p. 1-221.
124. Scott, W.R.P. *et al.* The GROMOS Biomolecular Simulation Program Package. *J Phys Chem A*, v. 103, p. 3596-3607, 1999.

125. Jorgensen, W.L.; Tirado-Rives, J. The OPLS Force Field for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J Am Chem Soc*, v. 110, n. 6, p. 1657-1666, 1988.
126. Jorgensen, W.L.; Maxwell, D.S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc*, v. 118, n. 45, p. 11225-11236, 1996.
127. Reuter, N. *et al.* Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches. *J. Phys. Chem. A*, v. 104, p. 1720-1735, 2000.
128. Stewart, J. MOPAC: A semiempirical molecular orbital program. *Journal of Computer-Aided Molecular Design*, v. 4, n. 1, p. 1-103, mar. 1990.
129. van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *J ComputChem*, v. 26, n. 16, p. 1701-18, 2006.
130. Frisch, M. J. *et al.* Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford CT, 2004.
131. Furche, F. *et al.* Turbomole. *WIREs Comput Mol Sci*, v. 4, n. 2, p. 91–100, March/April 2014.
132. Neese, F. The ORCA program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, v. 2, n. 1, p. 73-78, 2012.
133. Guest, M.F. *et al.* The GAMESS-UK structure package: algorithms, developments and applications. *Molecular Physics*, v. 103, n. 6-8, p. 719-747, March/April 2005.
134. Coutinho, K.; Canuto, S. DICE: A Monte Carlo program for molecular liquid simulation, University of São Paulo, Brasil, version 2.9 (2003).
135. Coutinho, K.; Canuto, S. DICE Manual: Version 2.9. São Paulo, Brasil. Disponível em <<http://fig.if.usp.br/~kaline/dice/dicemanual.pdf>>. Acesso em 22 jan. 2016.
136. Fogolari, F.; Brigo, A.; Molinari, H. The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology. *J. Mol. Recognit.*, v. 15, n. 6, p. 377-392, 2002.
137. Elstner, M.; Frauenheim, T.; Suha., S. An approximate DFT method for QM/MM simulations of biological structures and processes. *Journal of Molecular Structure: THEOCHEM*, v. 632, n. 1-3, p. 29-41, Aug 2003.
138. Karplus, M.; Petsko, G.A. Molecular dynamics simulations in biology. *Nature*, v. 347, Oct 1990.
139. Koehler, J.E.; Saenger, W.; van Gunsteren, W.F. A molecular dynamics simulation of crystalline alpha-cyclodextrin hexahydrate. *Eur. Biophys. J.*, v. 15, n. 4, p. 197-210, 1987.

140. Stewart, J.J.P. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. *J Mol Model. D*, v. 13, n.12, p. 1173-1213, dez. 2007.
141. Askeland, D.R.; Phulé, P.P. *The science and engineering of materials*. 5th ed. London: Thomson, 2006. 790 p.
142. Pattabhi, V.; Gautham, N. *Biophysics*. New York: Kluwer, 2002. 253 p.
143. Serdyuk, I.N.; Zaccai, N.R.; Zaccai, J. *Methods in Molecular Biophysics: Structure, Dynamics, Function*. New York: Cambridge University Press, 2007.
144. RCSB – Research Collaboratory For Structural Bioinformatics. Yearly growth of structures solved by X-ray. Camden. 2016. Disponível em:
<<http://www.rcsb.org/pdb/statistics/contentGrowthChart.do?content=explMethod-xray&seqid=100>>. Acesso em: 6 Fev. 2016.
145. Kaplan, E. *et al.* Aminoglycoside binding and catalysis specificity of aminoglycoside 2''-phosphotransferase IVa: A thermodynamic, structural and kinetic study. *Biochim Biophys Acta*, v. 1860, n. 4, p. 802-8013, Jan 2016.
146. Brecher, M. *et al.* Identification and Characterization of Novel Broad-Spectrum Inhibitors of the Flavivirus Methyltransferase. *ACS Infect Dis*, v. 1, n. 8, p. 340-349, 2015.
147. Suryanarayana, C.; Norton, M.G. *X-ray Diffraction: A Practical Approach*. New York: Plenum Press, 1998.
148. Snyder, R.L. X-Ray Diffraction. In: Lifshin, E. (ed.) *X-ray Characterization of Materials*. New York: Wiley-VCH, 1999. p. 1-103.
149. Birkholz, M. Principles of X-ray Diffraction. In: Birkholz, M. *Thin Film Analysis by X-Ray Scattering*. 1st ed. Weinheim: Wiley, 2006. Chapter 1. p. 1-40.
150. Messerschmidt, A. *X-Ray Crystallography of Biomacromolecules*. 1^a ed. Weinheim: Wiley-VCH, 2007.
151. RCSB – Research Collaboratory For Structural Bioinformatics. Yearly growth of structures solved by NMR. Camden. 2016. Disponível em:
<<http://www.rcsb.org/pdb/statistics/contentGrowthChart.do?content=explMethod-nmr&seqid=100>>. Acesso em 6 Fev. 2016.
152. Rubino, J.T. *et al.* Structural characterization of zinc-bound Zmp1, a zinc-dependent metalloprotease secreted by *Clostridium difficile*. *J Biol Inorg Chem*, p. 1-12, Dec 2015.

153. Ding, J. *et al.* HdeB chaperone activity is coupled to its intrinsic dynamic properties. *Sci Rep*, v. 5, n. 16856, p. 1-12, Dec 2015.
154. Duer, M.J. *Solid State NMR Spectroscopy*. 1^a ed. Londres: Blackwell Science, 2002.
155. Graaf, R.A. *In vivo NMR Spectroscopy: Principles and Techniques*. 2nd ed. Chichester: John Wiley & Sons, 2007.
156. Rule, G.S.; Hitchens, T.K. *Fundamentals of Protein NMR Spectroscopy*. Dordrecht: Springer, 2006.
157. Pavia, D.L. *et al.* *Introduction to Spectroscopy*. 4th ed. Stamford: Cengage Learning, 2009. 752 p.
158. Renault, M.; Cukkemane, A.; Baldus, M. Solid-state NMR spectroscopy on complex biomolecules. *Angew Chem Int Ed Engl*, v. 49, n. 45, p. 8346-57, Nov 2010.
159. Balci, M. *Basic ¹H- and ¹³C- NMR Spectroscopy*. 1. ed. Amsterdam: Elsevier, 2005.
160. Gerothanassis, I.P. *et al.* Nuclear magnetic resonance (NMR) spectroscopy: basic principles and phenomena, and their applications to chemistry, biology and medicine. *Chem. Educ. Res. Pract.*, v. 3, n. 2, p. 229-252, May 2002.
161. Keeler, J. *Understanding NMR Spectroscopy*. 2nd ed. Hoboken: Wiley, 2010. 526 p.
162. Cavanagh, J. *Protein NMR Spectroscopy: Principles and Practice*. 2nd ed. San Diego: Elsevier, 2007. 912 p.
163. Wider, G. Technical aspects of NMR Spectroscopy with biological macromolecules and studies of hydration in solution. *Progress in Nuclear Magnetic Resonance Spectroscopy*, v. 32, n. 4, p. 193-275, May 1998.
164. Bordoli, L.; Schwede, T. Automated Protein Structure Modeling with SWISS-MODEL Workspace and the Protein Model Portal. In: *Homology Modelling: Methods and proto-cols*. 1. ed. New York: Humana Press, 2012. p. 107-136.
165. Moulton, J. Predicting protein three-dimensional structure. *Curr. Opin. Biotechnol.*, v. 10, n. 6, p. 583-588, Dez. 1999.
166. Shortle, D. Prediction of protein structure. *Current Biology*, v. 10, n. 2, 2000.
167. Venselaar, H.; Krieger, E.; Vriend, G. Homology Modelling. In: Gu, J.; Bourne, P.E. (ed.) *Structural Bioinformatics*. 2. ed. Hoboken: Wiley-Blackwell, 2009. p. 715-736.
168. Suhrer, S.J. *et al.* Effective Techniques for Protein Structure Mining. In: Orry, A. J. W.; Abagyan, R. (ed.) *Homology Modeling: Methods and Protocols*. New York: Springer, 2012. cap. 2, p. 33-54.

169. Bhattacharya, A. *et al.* Assessing model accuracy using the homology modeling automatically (HOMA) software. *Proteins*, v. 70, p. 105-118, Jan 2008.
170. Alberts, B. *et al.* *Molecular Biology of the Cell*. 6th ed. New York: Garland Science, 2015. 1464 p.
171. Arnold, K. The Swiss-Model workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, v. 22, p. 195-201, Jan 2006.
172. Wallner, B.; Elofsson, A. All are not equal: A benchmark of different homology modeling programs. *Protein Science*, v. 14, n. 5, p. 1315-1327, May 2005.
173. Altschul, S.F. *et al.* Basic local alignment search tool. *J Mol Biol*, v. 215, n. 3, p. 403-10, Oct 1990.
174. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics*, v. 10, p. 421, 2009.
175. Guex, N.; Peitsch, M.C.; Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-Pdb Viewer: a historical perspective. *Electrophoresis*, v. 30, p. S162-73, 2009.
176. Venclovas, C. Methods for Sequence–Structure Alignment. In: In: Orry, A. J. W.; Abagyan, R. (ed.) *Homology Modeling: Methods and Protocols*. New York: Springer, 2012. Chapter3. p. 55-82.
177. Ramachandran, S.; Dokholyan, N.V. Homology Modeling: Generating Structural Models to Understand Protein Function and Mechanism. In: Dokholyan, N. (ed.) *Computational Modeling of Biological Systems: From Molecules to Pathways*. New York: Springer, 2012. p. 97-116.
178. Xiang, Z. Advances in homology protein structure modeling. *Curr Protein PeptSci*, v. 7, n. 3, p. 217-27, Jun 2006.
179. Rodriguez, R. Homology modeling, model and software evaluation: three related resources. *Bioinformatics*, v. 14, n. 6, p. 523-528, 1998.
180. Sanchez, R.; Sali, A. Comparative protein structure modeling. In: Webster, D.M. (ed.). *Protein structure prediction: methods and protocols*. New York: Kluwer, 2000. Chapter 6, p. 97-130.
181. Ierich, J.C.M. *et al.* A Computational Protein Structure Refinement of the Yeast Acetohydroxyacid Synthase. *J. Braz. Chem. Soc.*, v. 26, n. 8, p. 1702-1709, 2015.
182. Gonzalez, T.; Díaz-Herrera, J. *Computing handbook: Computer science and software engineering*. 3rd ed. Boca Raton: CRC Press, 2014. 2326 p.
183. Palermo, G.; De Vivo, M. *Computational Chemistry for Drug Discovery*. Bhushan, B. *Encyclopedia of Nanotechnology*. 1st ed. New York: Springer, 2015. p 1-15.

184. Fauci, A.S.; Marston, H.D. Ending AIDS - is an HIV vaccine necessary? *N. Engl. J. Med.*, v. 370, n. 6, p. 495-498, Feb 2014.
185. Kent, S.J. *et al.* The search for an HIV cure: tackling latent infection. *Lancet Infect. Dis.*, v. 13, n. 7, p. 614-621, Jul 2013.
186. Katlama, C. Barriers to a cure for HIV: new ways to target and eradicate HIV-1 reservoirs. *Lancet*, v. 381, n. 9883, p. 2109-2117, Jun 2013.
187. Jorgensen, W.L. The many roles of computation in drug discovery. *Science*, v. 303, n. 5665, p. 1813-8, Mar 2004.
188. Lombardino, J.G.; Lowe, J.A. The role of the medicinal chemist in drug discovery - then and now. *Nat Rev Drug Discov*, v. 3, n. 10, p. 853-862, 2004.
189. Babine, R.E.; Bender, S.L. Molecular recognition of protein-ligand complexes: applications to drug design. *Chem Rev*, v. 97, n. 5, p. 1359-1472, Aug 1997.
190. Lybrand, T.P. Ligand-protein docking and rational drug design. *Curr Opin Struct Biol*, v. 5, n. 2, p. 224-228, Apr 1995.
191. Carvalho, I.; Borges A.D.L.; Bernardes L.S.C. Medicinal Chemistry and Molecular Modeling: An Integration To Teach Drug Structure–Activity Relationship and the Molecular Basis of Drug Action. *J. Chem. Educ.*, v. 84, n. 4, p.588-596, April 2005.
192. Anderson, A.C. The process of structure-based drug design. *Chem Biol*, v. 10, n. 9, p. 787-797, Sep 2003.
193. Singh, S.; Malik, B.K.; Sharma, D.K. Molecular drug targets and structure based drug design: A holistic approach. *Bioinformation*, v. 1, n. 8, p. 314-320, Dec 2006.
194. Leite, F.L. *et al.* Nanoneurobiophysics: new challenges for diagnosis and therapy of neurologic disorders. *Nanomedicine (Lond)*, v. 10, n. 23, p. 3417-3419, Nov 2015.
195. Garcia, P.S. *et al.* A Nanobiosensor Based on 4-Hydroxyphenylpyruvate Dioxygenase Enzyme for Mesotrione Detection. *IEEE Sensors Journal*, v. 15, n. 4, p. 2106 - 2113, April 2015.
196. Steffens, C. *et al.* Atomic force microscope microcantilevers used as sensors for monitoring humidity. *Microelectronic Engineering*, v. 113, p. 80-85, 2014.
197. Steffens C. *et al.* Microcantilever Sensors Coated with a Sensitive Polyaniline Layer for Detecting Volatile Organic Compounds. *Journal of Nanoscience and Nanotechnology (Print)*, v. 14, p. 6718-6722, 2014.

198. Da Silva, A. *et al.* Nanobiosensors Based on Chemically Modified AFM Probes: A Useful Tool for Metsulfuron-Methyl Detection. *Sensors*, v. 13, n. 2, p. 1477-1489, Jan 2013.
199. Bueno, C.C. *et al.* Nanobiosensor for Diclofop Detection Based on Chemically Modified AFM Probes. *IEEE Sensors Journal*, v. 14, n. 5, p. 1467-1475, May 2014.
200. Deda, D.K. *et al.* The use of functionalized AFM tips as molecular sensors in the detection of pesticides. *Mat. Res.*, v. 16, n. 3, p. 683-687, Jun 2013.
201. Moraes, A.S. *et al.* Evidences of detection of atrazine herbicide by atomic force spectroscopy: a promising tool for environmental sensing. *Acta Microscopica*, v. 24, n. 1, p. 53-63, 2015.
202. Higa, A.M. *et al.* Ag-nanoparticle-based nano-immunosensor for anti-glutathione S-transferase detection. *Biointerface Research in Applied Chemistry*, v. 6, n. 1, p. 1053-1058, 2016.
203. Amarante, A.M. *et al.* Modeling the coverage of an AFM tip by enzymes and its application in nanobiosensors. *J Mol Graph Model.*, v. 53, p. 100-104, Sept 2014.
204. Leite, F.L. *et al.* Theoretical Models for Surface Forces and Adhesion and Their Measurement Using Atomic Force Microscopy. *IJMS*, v. 13, n. 12, p. 12773–12856, 2012.
205. Cai, X.; Liu, W.; Chen, S. Environmental effects of inclusion complexation between methylated beta-cyclodextrin and diclofop-methyl. *J. Agric. FoodChem.*, v. 53, n. 17, p. 6744-6749.
206. Lu, Y. *et al.* Stereo selective behaviour of diclofop-methyl and diclofop during cabbage pickling. *Food Chemistry*, v. 129, n. 4, p. 1690-1694, 2011.