

Lucas Carvalho Trindade

**Simulação computacional do efeito da pressão  
sobre a enzima pectina metilesterase do tomate**

Vitória

2018

Lucas Carvalho Trindade

**Simulação computacional do efeito da pressão  
sobre a enzima pectina metilesterase do tomate**

Dissertação apresentada ao Programa de Pós-Graduação em Física do Centro de Ciências Exatas da Universidade da Federal do Espírito Santo como requisito parcial para a obtenção do grau de mestre em Física, na área de concentração de Física Aplicada

Universidade Federal do Espírito Santo – UFES

Centro de Ciências Exatas

Programa de Pós-Graduação em Física

Orientador: Prof. Dr. Marcos Tadeu D’Azeredo Orlando

Vitória

2018

Lucas Carvalho Trindade

Simulação computacional do efeito da pressão sobre a enzima pectina metilesterase do tomate/ Lucas Carvalho Trindade. – Vitória, 2018-

120 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Marcos Tadeu D'Azeredo Orlando

Dissertação (Mestrado) – Universidade Federal do Espírito Santo – UFES

Centro de Ciências Exatas

Programa de Pós-Graduação em Física, 2018.

1. Proteínas 2. Gromacs 3. Raio de giro I. Marcos Tadeu D'Azeredo Orlando. II. Universidade Federal do Espírito Santo. III. Centro de Ciências Exatas. IV. Efeitos da Pressão sobre a enzima pectina metilesterase do tomate

Lucas Carvalho Trindade

## **Simulação computacional do efeito da pressão sobre a enzima pectina metilesterase do tomate**

Dissertação apresentada ao Programa de Pós-Graduação em Física do Centro de Ciências Exatas da Universidade da Federal do Espírito Santo como requisito parcial para a obtenção do grau de mestre em Física, na área de concentração de Física Aplicada

Trabalho aprovado. Vitória, 23 de fevereiro de 2018:

---

**Prof. Dr. Vinícius Candido Mota**

---

**Prof. Dra. Márcia Carvalho de Abreu  
Fantini**

---

**Prof. Dr. José Alexandre Nogueira**

---

**Prof. Dr. José Luis Passamai Junior**

Vitória  
2018

*Este trabalho é dedicado ao meu pai Alziro Pereira Trindade, a minha mãe Sônia Carvalho dos Santos e a minha irmã Luana Carvalho Trindade, por terem me dado um imenso apoio e condições necessárias para me adentrar numa jornada científica na Física.*

# Agradecimentos

A DEUS, que sempre esteve ao meu lado.  
Ao professor Marcos Tadeu D'Azeredo Orlando pela sua orientação.  
Ao apoio financeiro da CAPES e aos meus amigos.

*“Não vos amoldeis às estruturas deste mundo,  
mas transformai-vos pela renovação da mente,  
a fim de distinguir qual é a vontade de Deus:  
o que é bom, o que Lhe é agradável, o que é perfeito.  
(Bíblia Sagrada, Romanos 12, 2)*

# Resumo

Este trabalho descreve o desenvolvimento de uma simulação computacional, desenvolvida através do programa GROMACS. Mais especificamente, estudamos os resultados da simulação computacional referente a evolução do raio de giro da proteína Pectina Metilesterase (PME) do tomate em função da pressão aplicada, mantendo a temperatura fixa. Foi realizada uma descrição sobre os modelos físicos utilizados pelo programa. Também foi descrito, de forma sucinta, características sobre sistemas e estruturas dos aminoácidos e proteínas. As simulações computacionais consideraram uma temperatura de 26,85°C e pressões aplicadas de 1 bar, 1kbar, 3 kbar, 5 kbar, 7 kbar, 9 kbar e 10 kbar. As simulações trabalharam com um tempo de ação da pressão de até 100 pico segundos. Os resultados da simulação computacional indicaram uma redução não linear do raio de giro da enzima Pectina Metilesterase (PME) do tomate de acordo com o aumento da pressão aplicada, mantendo temperatura fixa.

**Palavras-chave:** Proteínas. Gromacs. Raio de giro.

# Abstract

This work describes a computer simulation developed by means of the GROMACS suit of program. More specifically, we studied how the computational simulation evaluated the evolution of the radius of gyrate of the Pectin Methylesterase (PME) protein of tomato as a function of the applied pressure under a constant temperature. It was made a description of the physical models used by the program. In addition, it were briefly described characteristics of amino acid, protein systems, and structures. The computational simulations considered a temperature of 26.85°C and pressures applied of 1 bar, 1 kbar, 3 kbar, 5 kbar, 7 kbar, 9 kbar and 10 kbar. The simulations worked a pressure action time of 100 peak seconds. The results of the simulation indicated a nonlinear reduction of the radius of gyrate of the Pectin Methylesterase (PME) enzyme according to the rise in pressure applied, considering a fixed temperature.

**Keywords:** Protein. Gromacs. Radius of gyrate.

# Lista de ilustrações

Figura 1 – Estrutura do aminoácido. Figura adaptada da referência (2). . . . .	18
Figura 2 – Geometria tetraédrica dos aminoácidos. Figura adaptada da referência (3). . . . .	18
Figura 3 – Isomeria espacial ótica. Figura adaptada da referência (1). . . . .	19
Figura 4 – Aminoácidos não ligados. Figura adaptada da referência (27). . . . .	20
Figura 5 – Reação química. Figura adaptada da referência (27). . . . .	20
Figura 6 – Ligação amida realizada com sucesso. Figura adaptada da referência (27). . . . .	20
Figura 7 – Classificação dos aminoácidos. Figura adaptada da referência (28). . . . .	22
Figura 8 – Plano amida do grupo peptídico. Figura adaptada da referência (29). . . . .	23
Figura 9 – Gráfico de Ramachandran. Figura adaptada da referência (29). . . . .	24
Figura 10 – Aspectos das ligações. Figura adaptada da referência (1) . . . . .	25
Figura 11 – Cristalografia de raios X da proteína mioglobina. Figura adaptada da referência (2). . . . .	26
Figura 12 – As diferentes formas de visualizar as proteínas. Figura adaptada da referência (1) . . . . .	27
Figura 13 – Configuração trans. Figura adaptada da referência (29) . . . . .	28
Figura 14 – Tipos de estruturas. Figura adaptada da referência (2). . . . .	29
Figura 15 – Insulina: duas estruturas primárias conectadas por ligações dissulfeto. Figura adaptada da referência (3). . . . .	30
Figura 16 – Hélice- $\alpha$ : levógira (canhota) e dextrógira (desta). Figura adaptada da referência (1). . . . .	31
Figura 17 – a) diagrama idealizado, b) pontes de hidrogênio, c) desenho esquemático, d) cadeira lateral R é projetada para fora da proteína. Figura adaptada da referência (2). . . . .	32
Figura 18 – Três hélices espirais representando as proteínas citrate synthase, alcohol dehydrogenase e troponin-C. Figura adaptada da referência (2). . . . .	32
Figura 19 – a) folha $\beta$ antiparalela, b) folha $\beta$ paralela. Figura adaptada da referência (1). . . . .	33
Figura 20 – Tipos de folhas $\beta$ . Figura adaptada da referência (29). . . . .	34
Figura 21 – Orientação dos terminais (N,C) da folha $\beta$ . Figura adaptada da referência (29). . . . .	34
Figura 22 – Proteína PDB 1rlb. Figura adaptada da referência (29). . . . .	35
Figura 23 – Dois tipos regiões volta $\beta$ (tipo 1,tipo 2) por quantidade de aminoácidos. Figura adaptada da referência (2) . . . . .	35
Figura 24 – Domínios da estrutura terciária. Figura adaptada da referência (3). . . . .	36

Figura 25 – Ligações que estabilizam a estrutura terciária. Figura adaptada da referência (3). . . . .	37
Figura 26 – Estrutura quaternária da hemoglobina. Figura adaptada da referência (3). . . . .	38
Figura 27 – Etapas do enovelamento da estrutura quaternária. Figura adaptada da referência (3). . . . .	39
Figura 28 – Modo de ação das enzimas. Figura adaptada da referência (1). . . . .	40
Figura 29 – Sítio de ligação das proteínas, Figura adaptada da referência (1). . . . .	40
Figura 30 – Sítio de ligação com substrato (1). . . . .	41
Figura 31 – União de proteínas. Figura adaptada da referência (1). . . . .	41
Figura 32 – Anticorpos são enzimas. Figura adaptada da referência (1). . . . .	42
Figura 33 – Velocidade da reação em função da concentração de enzima. Figura adaptada da referência (3). . . . .	44
Figura 34 – Velocidade da reação em função da concentração de substrato. Figura adaptada da referência (3). . . . .	44
Figura 35 – Velocidade de reação em função da temperatura. Figura adaptada da referência (3). . . . .	45
Figura 36 – Velocidade de reação em função do pH. Figura adaptada da referência (3). . . . .	45
Figura 37 – Modelo encaixe induzido. Figura adaptada da referência (3). . . . .	46
Figura 38 – Energia de ativação para reações catalizadas e não-catalizadas. Figura adaptada da referência (3). . . . .	46
Figura 39 – Estratégias de atuação das enzimas. Figura adaptada da referência (1) . . . . .	47
Figura 40 – Visualização em fitas de três proteínas interligadas por um túnel. Figura adaptada da referência (1) . . . . .	47
Figura 41 – Inibição competitivo (esquerda) e não competitivo (direita). Figura adaptada da referência (3). . . . .	48
Figura 42 – A cinética enzimática. Figura adaptada da referência (3). . . . .	48
Figura 43 – Moléculas podem ativar enzimas. Figura adaptada da referência (1). . . . .	49
Figura 44 – Moléculas ativam ou desativam enzimas. Figura adaptada da referência (1). . . . .	49
Figura 45 – Controle por feedback. Figura adaptada da referência (3). . . . .	50
Figura 46 – Célula vegetal. Figura adaptada da referência (33) . . . . .	52
Figura 47 – Pectina. Figura adaptada da referência (38). . . . .	52
Figura 48 – Ácidos galacturônicos conectados com grupos metoxilados (COOHCH <sub>3</sub> ), destacados em azul, formam a pectina. Figura adaptada da referência (41). . . . .	53
Figura 49 – Formas de atuação das pectinases. Figura adaptada da referência (43). . . . .	54
Figura 50 – Estrutura cristalográfica da mioglobina. Figura adaptada da referência (2). . . . .	56
Figura 51 – Visualização da enzima pectina metilesterase no Pymol. . . . .	56

Figura 52 – A composição da enzima pectina metilesterase do tomate, por variados tipos diferentes de estruturas. Figura adaptada da referência (45). . . . .	57
Figura 53 – Etapas do programa GROMACS. Figura adaptada da referência (46). . . . .	60
Figura 54 – Representação do sistema. Figura adaptada da referência (46). . . . .	62
Figura 55 – Caixa de simulação (fcc). Figura adaptada da referência (46). . . . .	62
Figura 56 – Condições de contorno. Figura adaptada da referência (49). . . . .	63
Figura 57 – Comportamento do potencial de Coulomb, plotado no Gnuplot. . . . .	65
Figura 58 – Comportamento do potencial de Lennard-Jones, plotado no Gnuplot. . . . .	65
Figura 59 – Variáveis dos potenciais intra-moleculares. Figura adaptada da referência (50). . . . .	66
Figura 60 – Comportamento do potencial de ligação, plotado no Gnuplot. . . . .	67
Figura 61 – Comportamento do potencial de vibração, plotado no Gnuplot. . . . .	67
Figura 62 – Comportamento do potencial de torção, plotado no Gnuplot. . . . .	68
Figura 63 – Raio de giro da proteína. Figura adaptada da referência (62). . . . .	79
Figura 64 – Raio de giro da PME com pressão de 1 bar. . . . .	80
Figura 65 – Raio de giro da PME com pressão de 1 kbar. . . . .	81
Figura 66 – Raio de giro da PME com pressão de 3 kbar. . . . .	82
Figura 67 – Raio de giro da PME com pressão de 5 kbar. . . . .	83
Figura 68 – Raio de giro da PME com pressão de 7 kbar. . . . .	83
Figura 69 – Raio de giro da PME com pressão de 9 kbar. . . . .	84
Figura 70 – Raio de giro da PME com pressão de 10 kbar. . . . .	85
Figura 71 – Evolução do raio de giro da PME de acordo com os aumentos da pressão. . . . .	86
Figura 72 – Espalhamento de raios X do Laboratório Nacional de Luz Síncroton. Figura adaptada da referência (64). . . . .	88
Figura 73 – Linhas de dipolo elétrico. Figura adaptada da referência (66). . . . .	93
Figura 74 – Dipolo elétrico do grupo peptídico $\vec{\mu}$ . Figura adaptada da referência (67). . . . .	94
Figura 75 – Diagrama da atuação do campo elétrico. Figura adaptada da referência (71). . . . .	96

# Lista de tabelas

Tabela 1 – Funções dos aminoácidos . . . . .	16
Tabela 2 – Aminoácidos essenciais e não-essenciais (1). . . . .	17
Tabela 3 – Funções das enzimas (11, 1) . . . . .	43
Tabela 4 – Porcentagem de pectina nos tecidos vegetais (36). . . . .	52
Tabela 5 – Classificação das enzimas pectolíticas ou pectinases (42). . . . .	54
Tabela 6 – Matriz com todos os pares de interações. Figura adaptada da referência (46) . . . . .	72
Tabela 7 – Matriz com todos os pares de interações. Figura adaptada da referência (46) . . . . .	72

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
1.1	Os aminoácidos compõem as proteínas	16
1.2	Os aminoácidos são encontrados em alimentos	16
1.3	Cadeia lateral R caracteriza o aminoácido	17
1.4	Ligações peptídicas dos aminoácidos	19
1.5	Classificação dos aminoácidos	22
<b>2</b>	<b>PRINCÍPIOS DAS PROTEÍNAS</b>	<b>23</b>
2.1	Plote de Ramachadran	23
2.2	Ligações não covalentes podem estabilizar proteínas	25
2.3	Características conformacionais das proteínas	26
2.4	Proteínas podem ser representadas de diferentes formas	27
2.5	Dipolos favorecem a configuração trans das proteínas	28
<b>3</b>	<b>ESTRUTURAS DAS PROTEÍNAS</b>	<b>29</b>
3.1	Estrutura primária	29
3.2	Estrutura secundária	31
3.2.1	Hélice $\alpha$	31
3.2.2	Folha $\beta$	33
3.2.3	Volta $\beta$	35
3.3	Estrutura terciária	36
3.4	Estruturas quaternária	38
<b>4</b>	<b>ENZIMAS</b>	<b>40</b>
4.1	Modo de ação das proteínas	40
4.2	Características	42
<b>5</b>	<b>A ENZIMA: PECTINA METILESTERASE</b>	<b>51</b>
5.1	Substrato: Pectina	51
5.2	Pectinases	53
5.3	Aplicações da pectina metil esterase	55
5.4	Protein Data Bank (PDB)	56
5.5	Influência da pressão nas atividade enzimática	57
<b>6</b>	<b>SIMULAÇÃO COMPUTACIONAL</b>	<b>59</b>
6.1	Fundamentos	59
6.2	Configuração inicial	61

6.3	Campo de força OPLS/AA . . . . .	64
6.4	Equações de movimento . . . . .	68
6.5	Dinâmica molecular do Potencial de Lennard-Jones . . . . .	70
6.6	Ensembles . . . . .	72
6.7	Método de Monte Carlo . . . . .	74
6.8	Ensemble NVE . . . . .	75
6.9	Ensemble NVT . . . . .	75
6.10	Ensemble NPT . . . . .	77
6.11	Raio de giro . . . . .	78
7	RESULTADOS DAS SIMULAÇÕES . . . . .	80
8	CONCLUSÃO . . . . .	86
9	PERSPECTIVAS FUTURAS . . . . .	88
	<b>APÊNDICES</b> . . . . .	<b>91</b>
	<b>APÊNDICE A – CONFIGURAÇÃO DAS PROTEÍNAS</b> . . . . .	<b>92</b>
A.1	Momento do dipolo elétrico do grupo peptídico . . . . .	92
A.2	Energia do dipolo elétrico . . . . .	95
	<b>APÊNDICE B – ENSEMBLES</b> . . . . .	<b>99</b>
B.1	Parâmetros extensivos e intensivos . . . . .	99
B.2	Ensemble Microcanônico . . . . .	100
B.3	Ensemble Canônico . . . . .	100
	<b>APÊNDICE C – PRESSÃO DO SISTEMA</b> . . . . .	<b>102</b>
	<b>APÊNDICE D – ALGORITMO DA SIMULAÇÃO</b> . . . . .	<b>105</b>
D.0.1	Etapas . . . . .	105
	<b>APÊNDICE E –</b> . . . . .	<b>109</b>
E.1	ions.mdp . . . . .	109
E.2	minim.mdp . . . . .	109
E.3	nvt.mdp . . . . .	110
E.4	npt.mdp . . . . .	111
E.5	md.mdp . . . . .	112
	<b>REFERÊNCIAS</b> . . . . .	<b>114</b>
	<b>Referências Bibliográficas</b> . . . . .	<b>114</b>

# 1 Introdução

Robert Hooke, em 1665, descobriu a célula através de um microscópio. Através de estudos posteriores, revelou-se que a célula é a menor parte do ser vivo correspondendo a unidade estrutural da matéria viva. Com o corpo humano possuindo cerca de 10 trilhões de células (1). Para que a célula consiga realizar suas funções como nutrição, produção de energia e reprodução é necessário fontes, transformação, armazenamento e eficiência de energia. Os carboidratos, obtidos através dos alimentos, é a primeira fonte de energia, cessando os carboidratos a próxima fonte de energia é a energia dos lipídeos, e a terceira fonte de energia é as proteínas fornecendo algo em torno de 4kcal/grama. As atividades celulares, cessam sem as proteínas, demonstrando sua importância (1).

Sabe-se também que o componente mais abundante da célula é a água ( $H_2O$ ) correspondendo aproximadamente a 70%, e o segundo mais abundante é a proteína representando cerca de 15%. De fato a teoria da evolução de Charles Darwin sugere que células que se desenvolveram com suporte de proteínas foram selecionadas e conseguiram crescer, desenvolver e reproduzir (1).

O século XXI, nos revela que existe uma convergência interdisciplinar da física com a biologia molecular devido ao sucesso inicial da descoberta da estrutura do DNA feita pelo cientistas Crick, Watson e Wilkins a partir de difração de raios X (2, 1). Os sistemas biológicos dependem de interações de átomos e moléculas fazendo da física o "coração da pesquisa". A pesquisa busca entender e desenvolver um banco de dados com cadeias de aminoácidos e proteínas prevendo padrões de comportamento. Além disso, faz-se uso de simulação computacional para analisar suas características e testar novos modelos protéicos (2, 1).

Existem muitos diferentes tipos de proteínas, cada uma especializada em particulares mecanismos biológicos. Essas proteínas apresentam características complexas e sofisticadas, sendo que suas funções, incluem aumento da velocidade das reações químicas catalisadas pelas enzimas (3). Participam no transporte de oxigênio, vitaminas, fármacos, lipídeos, ferro e cobre. Por exemplo, a proteína hemoglobina, que está contida no sangue, transporta oxigênio para os pulmões (3). Os anticorpos formados por proteínas trabalham na proteção contra substâncias estranhas (antígenos). Hormônios, como por exemplo, a eritropoietina é uma proteína que atua no crescimento humano (3). Além disso, as proteínas podem possuir características estruturais constituindo a matéria-prima para a construção dos tecidos da pele e cabelos, como por exemplo, as proteínas colágeno e a queratina, respectivamente (3). Outras proteínas também participam na expansão e contração dos músculos, como por exemplo, a proteína do sistema actina-miosina (3). Elementos químicos podem ser

armazenados em órgãos através de proteínas, como por exemplo, o ferro que é armazenado no fígado através da proteína ferritina (3).

## 1.1 Os aminoácidos compõem as proteínas

Os responsáveis por fornecer todo esse poder às proteínas são as funcionalidades de vinte e três aminoácidos. Suas funções particulares biológicas são listadas a seguir:

<b>Aminoácidos</b>	<b>Funções</b>
Alanina (A)	Reconhecimento das substâncias (4).
Fenilalanina (F)	Síntese de dopamina, noradrenalina e adrenalina (5).
Valina (V)	Correto dobramento das proteínas (6).
Leucina (L)	Correto dobramento das proteínas (6).
Isoleucina (I)	Correto dobramento das proteínas (6).
Prolina (P)	Importante na síntese e estruturas de proteínas (7).
Metionina (M)	Lipotrópica (participa da quebra de gordura) (8).
Serina (S)	Importante na aceleração das reações das enzimas (9).
Treonina (T)	Importante no funcionamento do sistema digestivo (10).
Tirosina (Y)	Poder de absorvência da luz ultravioleta (11).
Triptofano (W)	Poder de absorvência máxima da luz ultravioleta (11).
Cisteína (C)	Importante antioxidante (12).
Asparagina (N)	Importante no desenvolvimento e nas funções do cérebro (13).
Arginina (R)	Manutenção e cicatrização de feridas (14).
Glutamina (Q)	Manutenção do sistema imunológico (15).
Lisina (K)	Síntese de hormônios, enzimas e anticorpos (16).
Histidina (H)	Participa da catálise ácido/base (17).
Glicina (G)	Neurotransmissor inibitório (18).
Ácido aspartático (D)	Neurotransmissor excitatório (19).
Ácido glutâmico (E)	Neurotransmissor do aprendizado e memória (20).

Tabela 1 – Funções dos aminoácidos

Em todas as células do organismo estes aminoácidos participam da síntese proteica (formação de proteínas), sendo mais preciso, nas organelas encontradas no citoplasma e no retículo endoplasmático rugoso chamado de ribossomos (1, 21).

## 1.2 Os aminoácidos são encontrados em alimentos

Os aminoácidos não essenciais são sintetizados pelo organismo humano, e os aminoácidos essenciais não são sintetizados, mas sim obtido dos alimentos de origem animal ou vegetal (22). As proteínas podem ser metabolizadas em aminoácidos que permitem a sintetização de outras proteínas (22). Segundo o conceito de pontuação de

aminoácidos corrigido pela digestibilidade protéica (PDCAAS), as proteínas que possuem escore acima de 0.8 são consideradas de alto valor biológico, ou seja, possuem aminoácidos essenciais capazes de suprir as necessidades da dieta humana. A soja possui valor de 1.0, mais alto que a carne de boi, que possui escore 0.92 (22). Portanto, os aminoácidos ditos essenciais podem proceder de 1) origem vegetal: frutas, verduras, legumes, leguminosas, em geral possuem baixo valor biológico, com exceção da soja, mas em compensação são ricos em fibra, vitamina C, ácido fólico (vitamina B9), magnésio, cobre e manganês (23). Também podem proceder de 2) origem animal: obtidos em carnes, peixes, ovos, leite; em geral possuem alto valor biológico, ou seja muita variedade de aminoácidos essenciais, e também são ricos em ferro, cálcio, zinco, riboflavina (vitamina B<sub>2</sub>), vitamina A, vitamina B<sub>12</sub> e possui em geral baixa quantidade de fibras (23).

Essenciais	Não-Essenciais
Fenilalanina (F)	Ácido aspartático (D).
Histidina (H)	Ácido glutâmico (E)
Isoleucina (I)	Alanina (A).
Leucina (L)	Arginina (R).
Lisina (K)	Asparagina (N).
Metionina (M)	Cisteína (C).
Treonina (T)	Glicina (G).
Triptofano (W)	Glutamina (Q).
Valina (V)	Prolina (P).
	Serina (S).
	Tirosina (Y).

Tabela 2 – Aminoácidos essenciais e não-essenciais (1).

É importante uma boa dieta alimentar para se evitar doenças relacionadas à deficiência de proteínas ou aminoácidos. Em uma dieta diária a ingestão de proteína deve representar cerca de (10 – 15%) do consumo total (24). Embora, o corpo humano necessite apenas de aminoácidos dentro desta classe específica para um bom funcionamento, são conhecidos centenas de outras variedades encontrados na natureza em cultura de microorganismo, plantas, frutas e outros (25).

### 1.3 Cadeia lateral R caracteriza o aminoácido

Na natureza, encontramos os aminoácidos formados por substâncias orgânicas como carbono, hidrogênio, oxigênio, nitrogênio, enxofre, entre outros. Composto sua estrutura, que inclui um grupo amina (NH<sub>2</sub>) e um grupo carboxila (COOH) conectados por um carbono central C<sub>α</sub>. Todos os aminoácidos possuem uma identidade que confere sua

especificidade e é chamada de cadeia lateral R. A especificidade confere sua identificação com as suas respectivas funções biológicas (3).

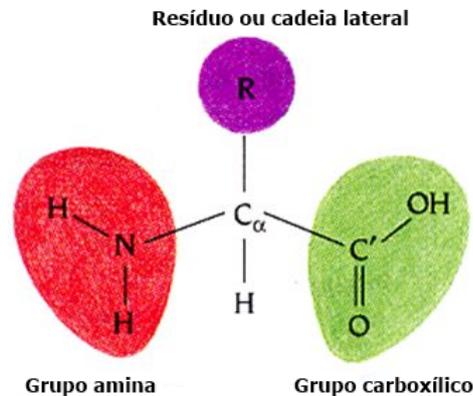


Figura 1 – Estrutura do aminoácido. Figura adaptada da referência (2).

Portanto, os aminoácidos podem ser representados em uma configuração tetraédrica estando em geral diluídos no organismo em meio aquoso. Com isso, sofrem alterações na sua estrutura química. Esta dependência com a característica do meio é chamado de pH. O pH mede a concentração de íon hidrogênio ( $H^+$ ) refletindo na acidez, neutralidade ou alcalinidade da água. Uma solução com  $pH < 7$  é ácida,  $pH = 7$  é neutra ou  $pH > 7$  é básica. Portanto, quando o valor do  $pH = 7$ , a concentração de íon de hidrogênio ( $H^+$ ) e íon de hidroxila ( $OH^-$ ), possuem quantidades iguais. Logo, a carga líquida da solução é nula, por consequência a proteína tende a não sofrer influências elétricas com o sítio de ligação, ou seja, as proteínas se ligam com substâncias químicas para diversas finalidades, sendo assim em pH neutro a água tenderia a não competir por meio de pontes de hidrogênio pelo sítio de ligação (1). A cadeia lateral ou resíduo caracteriza os vinte tipos diferentes de aminoácidos do qual nosso organismo necessita. Em um meio sanguíneo ( $pH = 7$ ), os grupos carboxila e amina estão ionizados formando um dipolo elétrico (1).

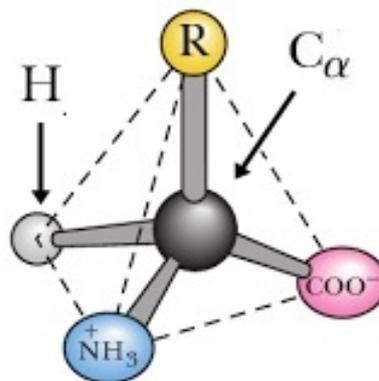


Figura 2 – Geometria tetraédrica dos aminoácidos. Figura adaptada da referência (3).

O carbono  $\alpha$  assimétrico ligado a quatro diferentes ligantes, chama-se carbono quiral. Louis Pasteur, em 1848, descobriu a isomeria espacial ótica na qual compostos quirais apresentam atividade ótica (1). Compostos quirais são a imagem especular um do outro desviam com um mesmo ângulo o feixe de luz polarizado para a esquerda ou direita, compostos quirais chamados por levógiros L ou dextrógiros D, respectivamente (1).

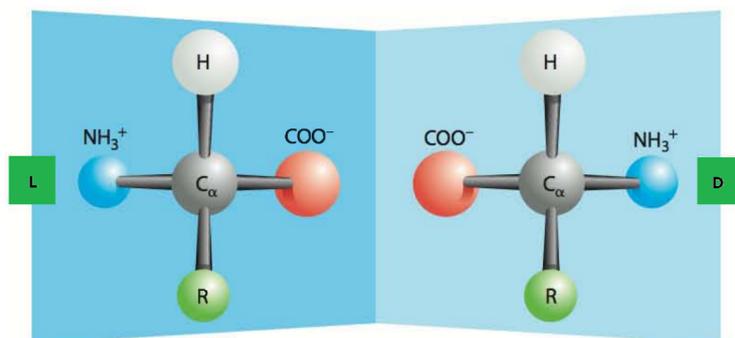


Figura 3 – Isomeria espacial ótica. Figura adaptada da referência (1).

Logo, os aminoácidos levógiros (L-aminoácidos) desviam a luz polarizada para a esquerda, enquanto os aminoácidos dextrógiros (D-aminoácidos) desviam a luz polarizada para a direita (1).

## 1.4 Ligações peptídicas dos aminoácidos

Na natureza, em geral, encontra-se em minoria aminoácidos eletricamente neutros, com grupo amina desprotonado ( $-\text{NH}_2$ ) e o grupo carboxila protonado ( $-\text{COOH}$ ). A maioria, assumirá uma forma ionizada, em que o grupo amina terá a forma protonada ( $-\text{NH}^{+3}$ ) e o grupo carboxílico desprotonado ( $-\text{COO}^-$ ), este formato é denominado de zwitteriônico (palavra de origem alemã zwitter, significando "híbrido") (26).

Nas proteínas, os aminoácidos ficam conectados em sequência por ligações peptídicas. Isto ocorre, por meio de reações químicas da ligação entre o grupo carboxila protonado ( $\text{C}'\text{OO}^-$ ) com o grupo amina protonada ( $\text{NH}^{+3}$ ), é uma reação de desidratação molecular, liberando uma molécula de água ( $\text{H}_2\text{O}$ ) e resultando na ligação peptídica ou amida entre o carbono  $\text{C}'$  com o nitrogênio  $\text{N}$ , esta ligação é abreviada por ( $\text{C}'\text{-N}$ ) (2, 1, 11, 26). Desta forma, terá um grupo amina em uma extremidade chamado de N-terminal e na extremidade oposta o grupo carboxílico denominado de C-terminal .

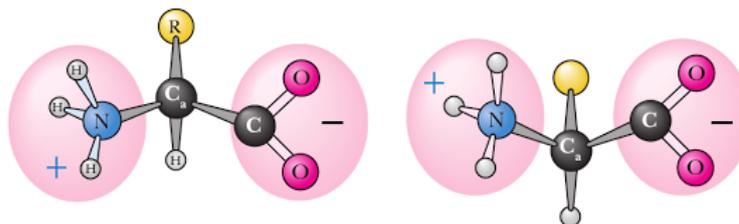


Figura 4 – Aminoácidos não ligados. Figura adaptada da referência (27).

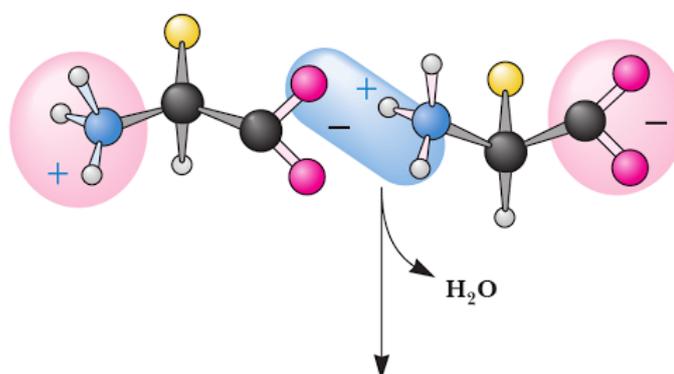


Figura 5 – Reação química. Figura adaptada da referência (27).

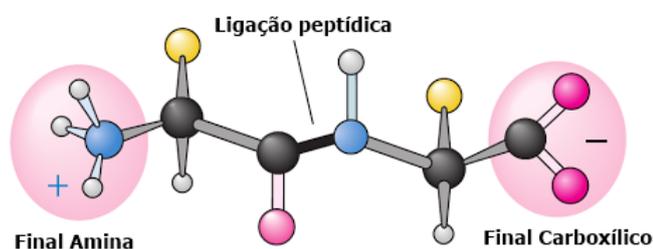


Figura 6 – Ligação amida realizada com sucesso. Figura adaptada da referência (27).

Além disso, existe a isomeria espacial geométrica. Isso ocorre quando os ligantes de maior massa em nosso caso a cadeia lateral R, podem se situar lado a lado ou não, classificados como cis (Z) ou trans (E), respectivamente. As ligações peptídicas na configuração trans (E) são predominantes nas proteínas, a razão cis/trans é aproximadamente 1/1000 a cada mil configurações trans uma delas pode ocorrer na configuração cis (Z) (2, 1, 11).

A configuração cis (Z) não é abundante, pois podem ocorrer interações entre as cadeias laterais (R), fazendo com que possam perder suas propriedades ou também não favorecendo no dobramento das proteínas. É importante ressaltar que uma cadeia de aminoácidos unidos por ligações peptídicas, somente forma um proteína se exercer uma função biológica (2, 1, 11).

Existe uma classificação dos aminoácidos, indicando a polaridade da sua cadeia lateral (R). Estão incluídos em grupos polares, conhecidos por possuírem afinidade com a água ( $\text{H}_2\text{O}$ ), enquanto os apolares (não polares) não possuem esta afinidade (2, 1, 11).

Logo, a consequência de existirem estes dois grupos distintos permite que as proteínas exerçam suas atividades em meio aquoso, afinal como somos compostos em maioria por água ( $\text{H}_2\text{O}$ ). Portanto as estruturas das proteínas vão estar inteiramente relacionados às características das polaridades dos seus aminoácidos (2, 1, 11).

## 1.5 Classificação dos aminoácidos

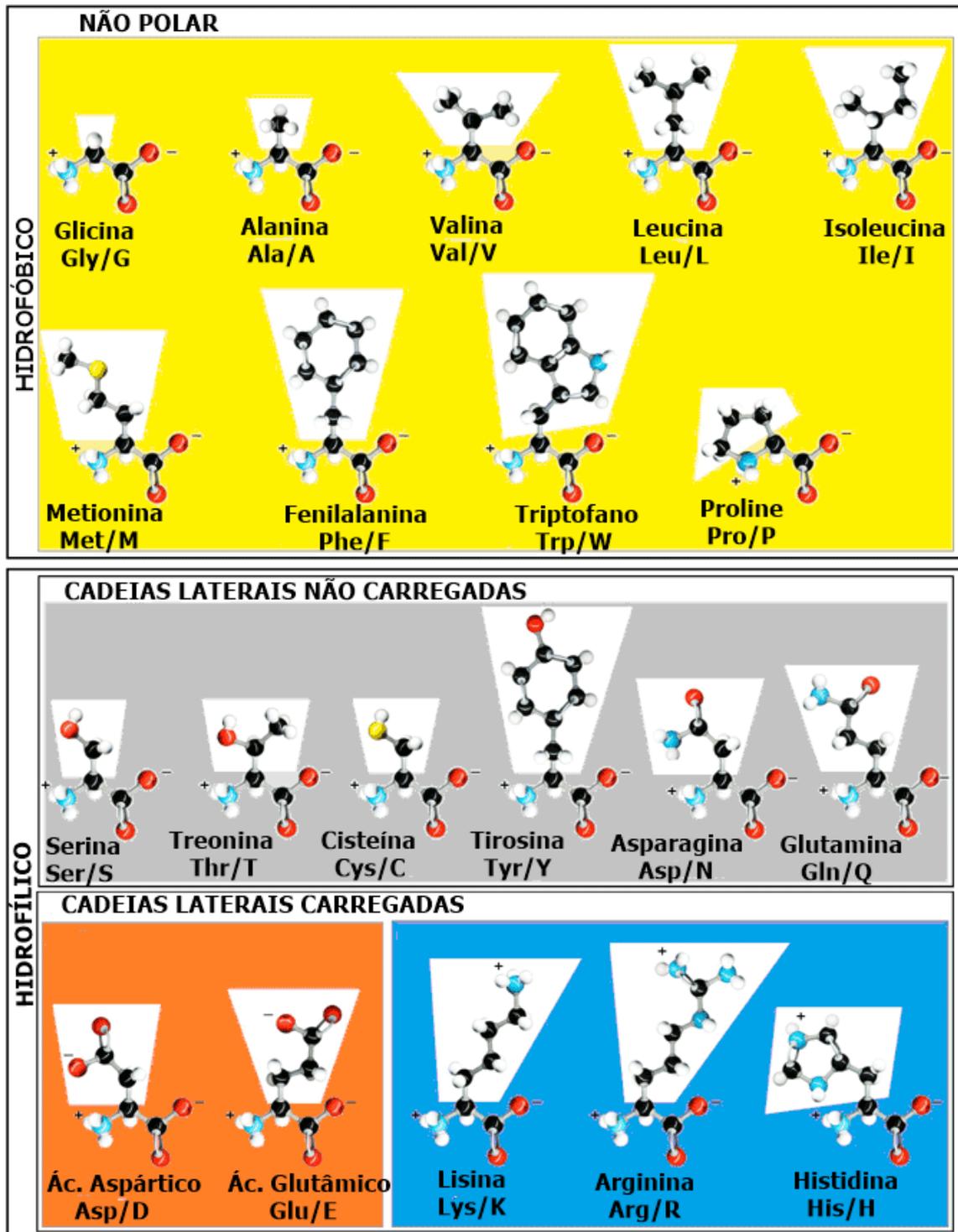


Figura 7 – Classificação dos aminoácidos. Figura adaptada da referência (28).

## 2 Princípios das proteínas

### 2.1 Plote de Ramachadran

Foi convenicionado atribuir o conceito de blocos rígidos formando uma estrutura no plano amida. A esse bloco deu-se o nome de unidade peptídica. Cada unidade peptídica pode rotacionar sobre duas ligações: 1) ligação entre o carbono alfa e o carbono do grupo carboxila ( $C_{\alpha}$ -C'); rotação por um ângulo ( $\Psi$ ), e 2) ligação entre o nitrogênio da amina com o carbono alfa (N- $C_{\alpha}$ ); rotação por um ângulo ( $\Phi$ ). Cada unidade peptídica está associada a estes dois ângulos de torção representando os graus de liberdade das proteínas (29).

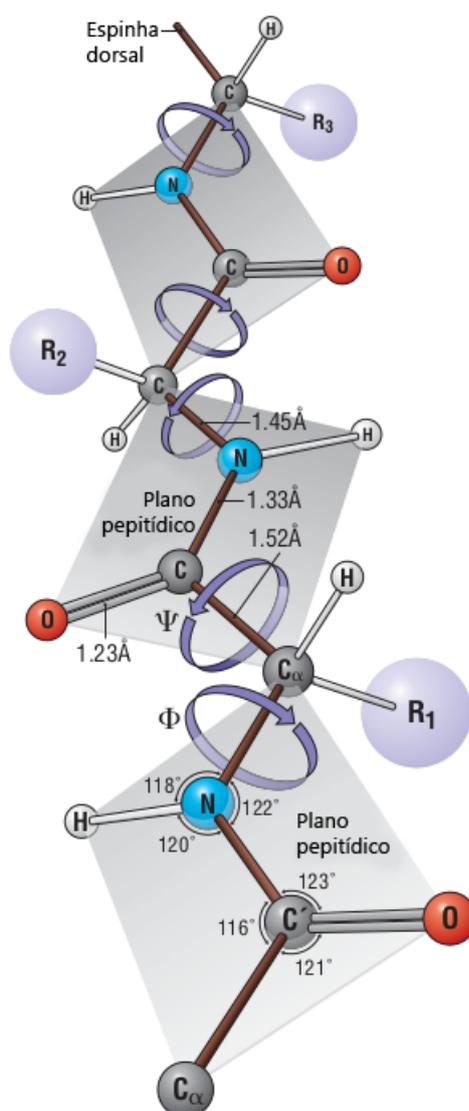


Figura 8 – Plano amida do grupo peptídico. Figura adaptada da referência (29).

Portanto os valores dos ângulos de torção nos fornecem a “conformação” da proteína. As unidades peptídicas estão rotacionadas em torno do carbono  $\alpha$ . G.N Ramachandran, em 1963 foi o primeiro a descrever arranjos proteicos estáveis, descobriu que apenas certos valores dos ângulos ( $\Psi, \Phi$ ) são permitidos para cada proteína, e estes valores dos ângulos são apresentados no gráfico a seguir (29):

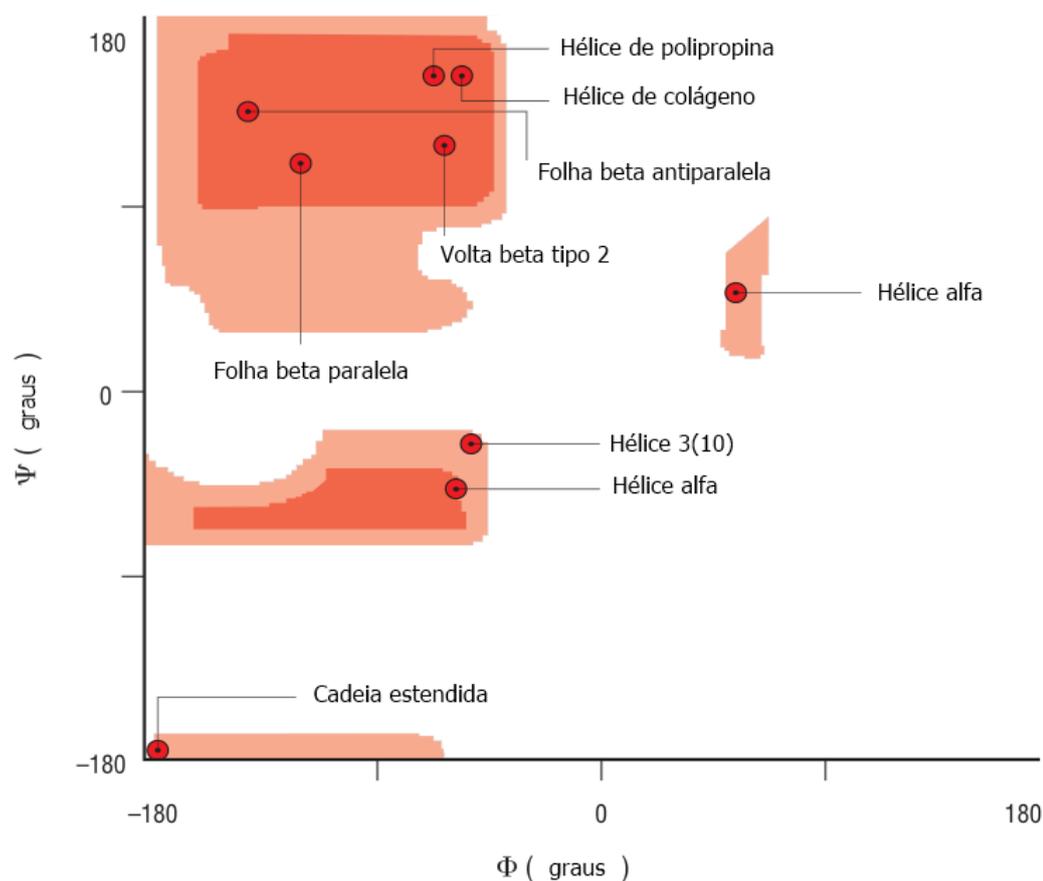


Figura 9 – Gráfico de Ramachandran. Figura adaptada da referência (29).

Certas rotações são mais estáveis que outras; a região branca é instável. Nas regiões vermelhas são permitidas rotações, por não implicar numa interferência estérica entre os resíduos R (cadeias laterais). Regiões alaranjadas representam formação de proteínas sob condições de efeito estérico (pode ser de resistência ou atração a novas ligações). Os valores zero de phi ( $\Phi$ ) e psi ( $\Psi$ ) definem a configuração trans (29).

## 2.2 Ligações não covalentes podem estabilizar proteínas

Embora, ligações não-covalentes sejam fracas, muitas delas favorecem o enovelamento criando um forte vínculo de ligação entre as cadeias laterais dos aminoácidos (1).

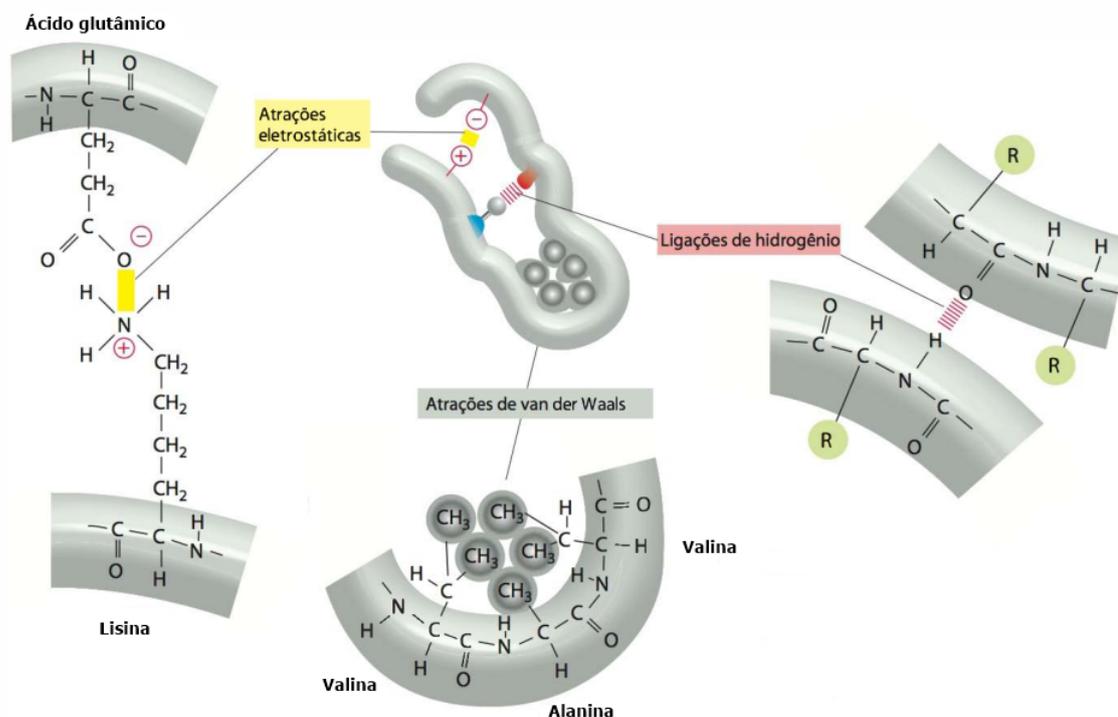


Figura 10 – Aspectos das ligações. Figura adaptada da referência (1)

As atrações eletrostáticas existem entre cargas elétricas opostas. As ligações de hidrogênio são ligações entre o hidrogênio (H) com flúor (F), nitrogênio (N) ou oxigênio (O). Enquanto, as atrações de van der Waals são caracterizadas por ligações entre dipolos elétricos. Estes tipos de ligações não covalentes são as mais encontradas em proteínas (1).

## 2.3 Características conformacionais das proteínas

John Kendrew, em 1958, apresentou os primeiros resultados em cristalografia de raios X da proteína mioglobina, revelando uma estrutura irregular. Posteriores estudos mostraram que o interior das proteínas continha aminoácidos com cadeias laterais hidrofóbicas. Isso implica que a proteína enovela em um conformação compacta.

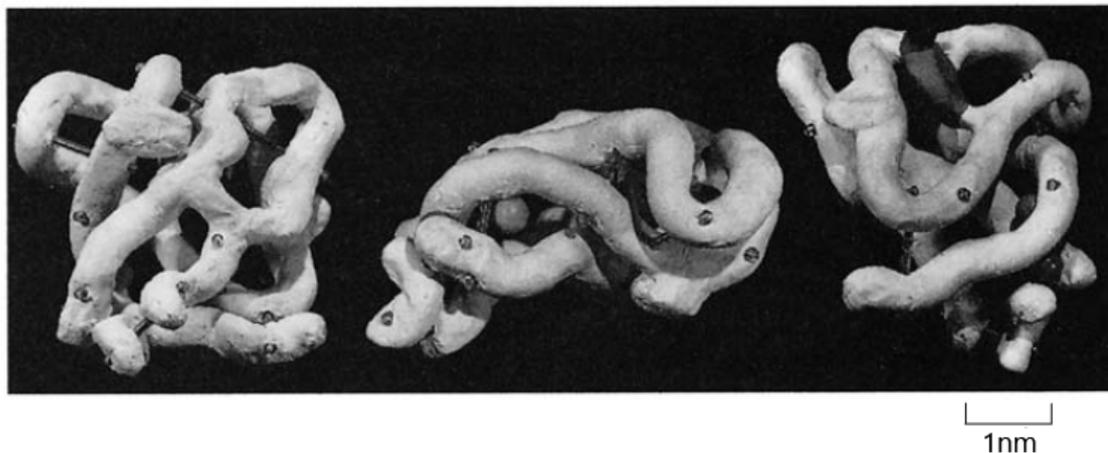


Figura 11 – Cristalografia de raios X da proteína mioglobina. Figura adaptada da referência (2).

Através de posteriores estudos, descobriu-se alguns princípios do enovelamento das proteínas listados a seguir (1, 2):

1) Compactar: a estabilidade está ligada ao tamanho das proteínas, quanto menor, então mais átomos estarão próximos um dos outros tornando a estrutura mais densa reduzindo os espaços (voids), impedindo o efeito indesejado de moléculas de água no interior (1, 2).

2) Encobrir os hidrofóbicos: a grande maioria das proteínas flutuam na água, a tendência é manter o máximo de hidrofílicos na superfície enquanto os hidrofóbicos estão cobertos (1, 2).

3) Eliminar os atritos (clashes): cadeias laterais muito próximas alteram a conformação das proteínas, e por consequência, podendo se tornar inativa ou alterar sua função (1, 2).

## 2.4 Proteínas podem ser representadas de diferentes formas

Maneiras diferentes do desenho esquemático das proteínas podem ser úteis para analisar diferentes aspectos das estruturas; permitindo a análise sobre diferentes pontos de vista, formando uma visão mais completa. É apresentado a seguir algumas dessas maneiras tais como cadeia principal, fitas, esqueleto e preenchimento espacial:

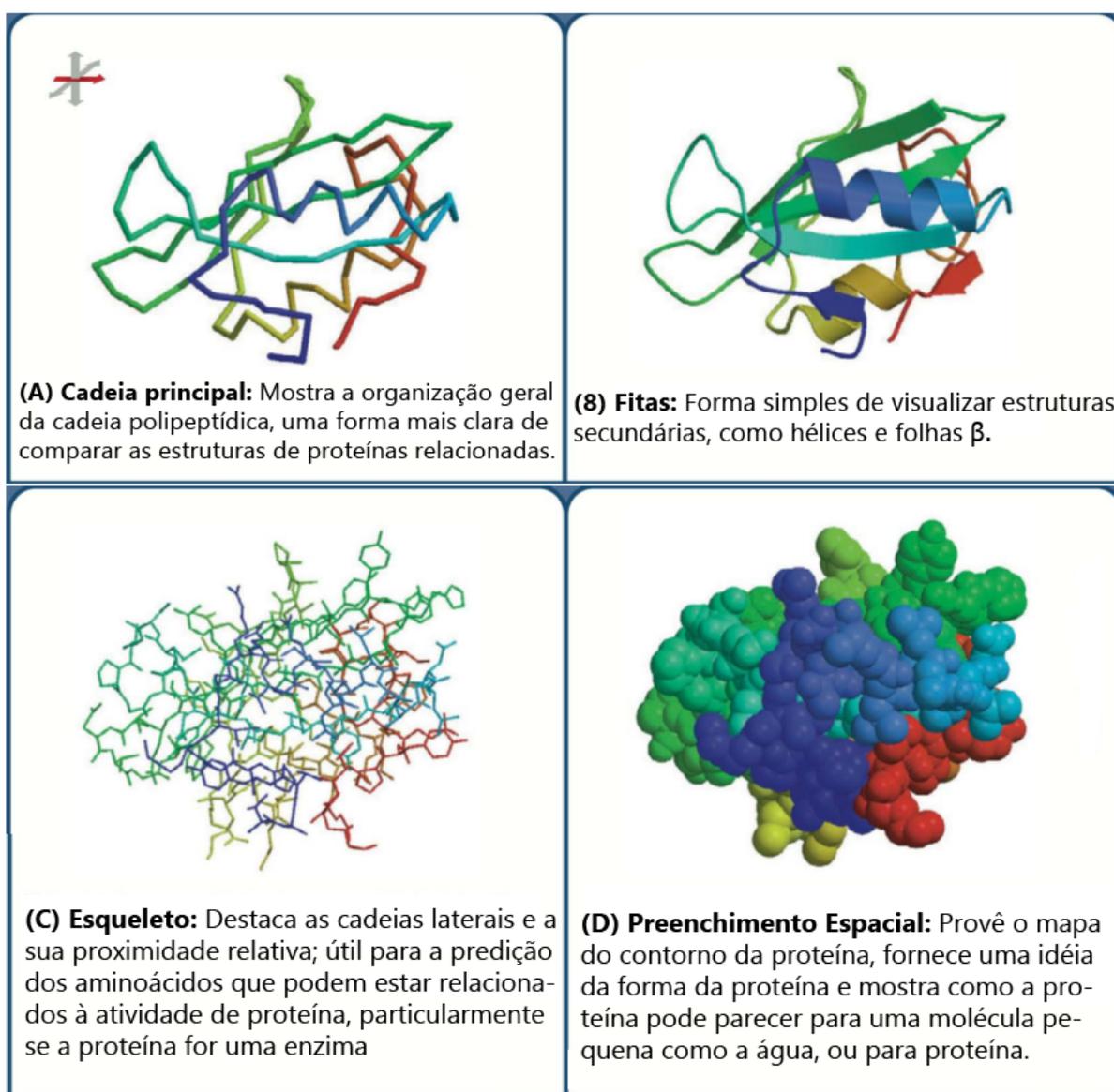


Figura 12 – As diferentes formas de visualizar as proteínas. Figura adaptada da referência (1)

A cadeia principal fornece uma idéia geral da cadeia polipeptídica podendo comparar com outras estruturas, enquanto as fitas revelam estruturas secundárias como hélice  $\alpha$  e folha  $\beta$ . Além disso, o esqueleto apresenta as cadeias laterais e suas distâncias, utilizada para prever a atividade da proteína, especialmente no caso das enzimas. O mapa de

contorno da proteína é mostrada no preenchimento espacial e com isso pode-se saber quais cadeias de aminoácidos estão expostas na sua superfície (1).

## 2.5 Dipolos favorecem a configuração trans das proteínas

Conforme explicado no "Apêndice A", a energia potencial do dipolo elétrico  $U$  quando está sobre influência de um campo elétrico  $\vec{E}$  pode ser escrito na forma (68):

$$U = \vec{\mu} \vec{\nabla} \psi = -\vec{\mu} \vec{E} = -E\mu \cos\theta, \quad (2.1)$$

onde no caso do ângulo  $\theta = \pi$  com  $U = \mu E$ , ou seja,  $\vec{\mu}$  e  $\vec{E}$  antiparalelos representado um sistema de energia máxima. O grupo peptídico  $n$  configura um dipolo  $\vec{\mu}_n$ . Considera-se a análise do campo elétrico  $\vec{E}_n$  na direção  $(-y)$  no eixo  $(+x)$  do grupo peptídico  $n$ , gerado pelo dipolo  $\vec{\mu}_n$ . De forma, que atua um torque  $\tau$  sobre o dipolo  $\vec{\mu}_{n+1}$  do grupo peptídico  $n+1$ , fazendo-o alinhar-se ao campo elétrico  $\vec{E}_n$ . Isto acontece, pois a natureza tende a minimizar a energia.

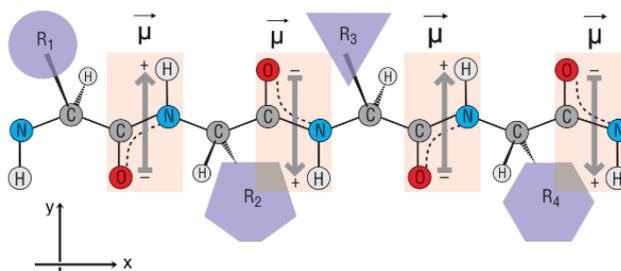


Figura 13 – Configuração trans. Figura adaptada da referência (29)

Então, agora o grupo peptídico  $n+1$  com momento de dipolo  $\vec{\mu}_{n+1}$  gera um outro campo elétrico  $\vec{E}_{n+1}$  sobre a direção  $(y)$  no eixo  $(+x)$  do grupo peptídico  $n+2$  (com momento de dipolo  $\vec{\mu}_{n+2}$ ); de forma a alinhar o momento de dipolo  $\vec{\mu}_{n+2}$  com o campo elétrico  $\vec{E}_{n+1}$ . Estes sucessivos alinhamentos dos grupos peptídicos resultarão na configuração trans predominante em proteínas (29).

## 3 Estruturas das proteínas

Apenas algumas cadeias polipeptídicas são úteis para a célula, para uma proteína com quatro aminoácidos existem  $20 \times 20 \times 20 \times 20 = 160.000$  possíveis cadeias polipeptídicas. Para uma proteína com  $N$  aminoácidos existem  $20^N$  (1). Em geral, as proteínas possuem cerca de 300 aminoácidos. Isto representa  $20^{300}$  possíveis cadeias polipeptídicas, apenas uma pequena fração em estimativas de um bilhão apresentam enovelam em uma estrutura estável (1). Então, o enovelamento final da proteína em um processo natural é denominado conformação nativa, e podem assumir quatro possíveis estruturas: primárias, secundárias, terciárias e quaternárias.

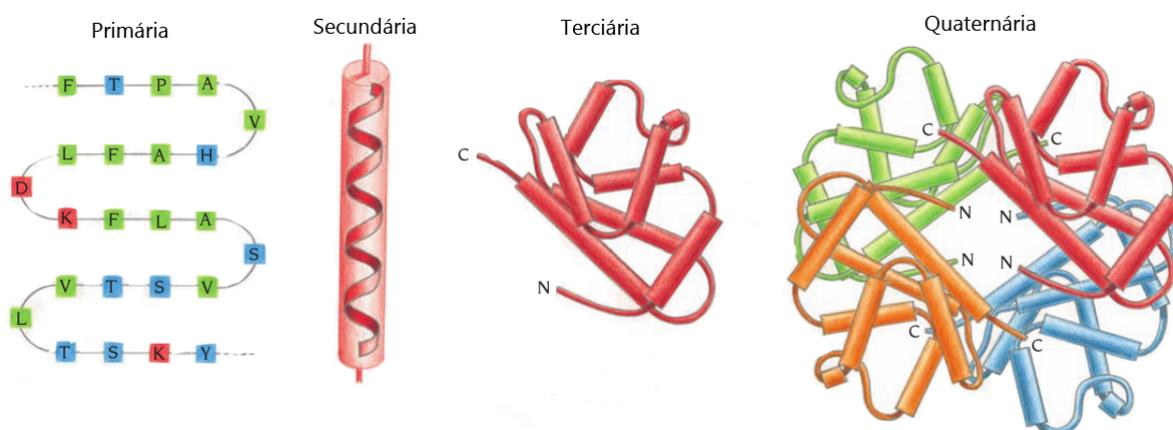


Figura 14 – Tipos de estruturas. Figura adaptada da referência (2).

De acordo com a seleção natural baseada na teoria da evolução de Charles Darwin, pode-se afirmar que apenas as proteínas que colaboraram para a sobrevivência da célula foram selecionadas por tentativa e erro ao longo da evolução biológica (1). Logo, o fato de assumir uma determinada estrutura dependerá de três fatores: tipos de aminoácidos, tamanho da cadeia polipeptídica e configuração espacial (2).

### 3.1 Estrutura primária

A cadeia polipeptídica composta de aminoácidos em sequência conectados por ligações peptídicas (C'-N) formam a estrutura primária da proteína. É a estrutura mais simples, sendo preservada em caso de enovelamento e posterior desenovelamento.

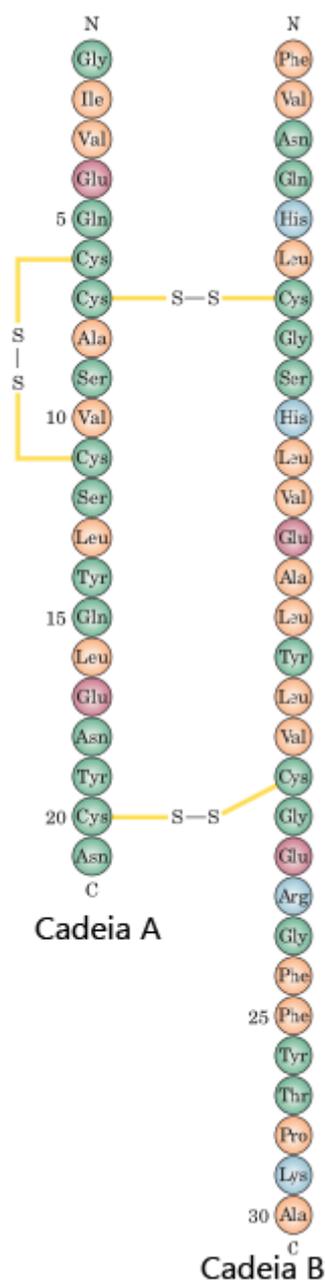


Figura 15 – Insulina: duas estruturas primárias conectadas por ligações dissulfeto. Figura adaptada da referência (3).

Cadeias laterais da cadeia polipeptídica que possuem átomos de enxofre (S), se estiverem direcionadas formam ligação covalente entre os átomos de enxofre (S). Esta reação libera dois átomos de hidrogênio (H<sub>2</sub>) ao meio, conhecida como ligação dissulfeto (S-S). É uma reação de oxidação, com perda de elétrons. Caso ocorra a redução, com ganho de elétrons, a ligação é desfeita. A orientação da cadeia polipeptídica, é convencionalmente iniciando no terminal amina (terminal N) e terminando no terminal carboxila (terminal C) (3).

## 3.2 Estrutura secundária

Cadeias polipeptídicas representando estruturas primárias diferentes, podem enovelar ou não, podem ter função ou não. Todavia, caso se enovolem e possuam diferentes funções, todas possuirão diferentes estruturas tridimensionais (2). No entanto, a estrutura primária pode conter um setor favorável e enovelar para estruturas mais estáveis, como por exemplo as estruturas secundárias hélice- $\alpha$  ou folha- $\beta$ .

### 3.2.1 Hélice $\alpha$

Em 1951, Linus Pauling e Robert Corey propuseram uma estrutura secundária, em que a proteína se enovela em um espiral sustentada por ligações de hidrogênio paralelas ao eixo da espiral. Chama-se hélice- $\alpha$ , e pode assumir duas conformações nativas, destra ou canhota, de acordo com a orientação do parafuso da cadeia polipeptídica. Exemplos: é a estrutura dominante na proteína queratina- $\alpha$ , abundante na pele e tecidos e seu derivados, como cabelo e unha (1).

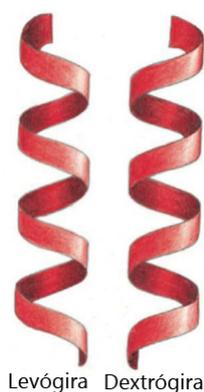


Figura 16 – Hélice- $\alpha$  : levógira (canhota) e dextrógira (destra). Figura adaptada da referência (1).

A conformação dextrógira é predominante nas proteínas, possuem 3.6 cadeias laterais R por volta de  $360^\circ$  localizados no exterior da estrutura, as pontes de hidrogênio são responsáveis pela estabilização com plote de Ramachandran ( $\Phi = -57^\circ$ ,  $\psi = -47^\circ$ ) (1, 2, 29).

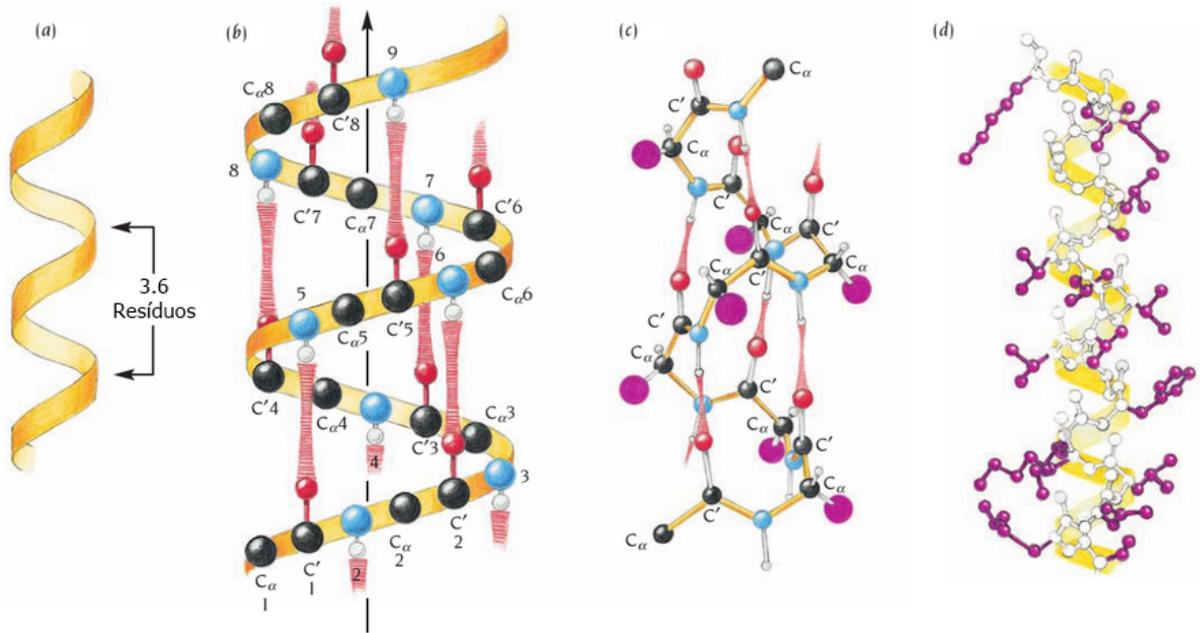


Figura 17 – a) diagrama idealizado, b) pontes de hidrogênio, c) desenho esquemático, d) cadeia lateral R é projetada para fora da proteína. Figura adaptada da referência (2).

Portanto, liga-se o grupo peptídico  $n$ , a partir de pontes de hidrogênio, com o grupo peptídico  $n+4$ . Mais especificamente, o oxigênio do carbono ( $C=O$ ) do grupo peptídico  $n$ , liga-se com o hidrogênio do nitrogênio ( $N-H$ ) do grupo peptídico  $n+4$ . Então, no processo de enovelamento, múltiplas pontes de hidrogênio se formam até que sua estrutura final esteja estabilizada (1). Existem vários tipos de hélice- $\alpha$ , a seguir é apresentada algumas delas com as cadeias laterais dos aminoácidos representados a cada  $100^\circ$ :

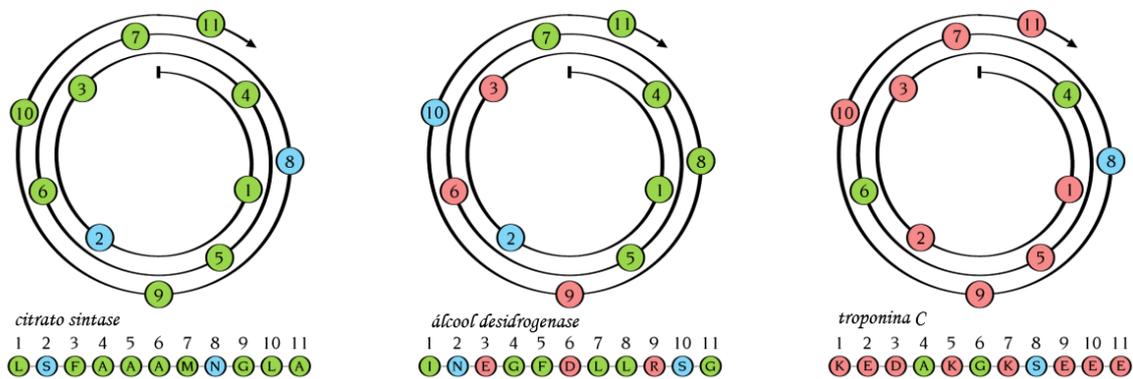


Figura 18 – Três hélices espirais representando as proteínas citrate synthase, alcohol dehydrogenase e troponin-C. Figura adaptada da referência (2).

onde a cor verde representa um aminoácido hidrofóbico, a cor azul a cadeia lateral hidrofílica e representa a cor vermelho para a cadeia lateral carregada. Então, a primeira

hélice é hidrofóbica. A segunda é polar em um lado e hidrofóbica em outro lado, e a terceira é polar (1).

### 3.2.2 Folha $\beta$

Um segundo padrão de enovelamento foi encontrado no principal componente da seda, a proteína fibroína. Este padrão é caracterizado por representar uma estrutura rígida, compondo o centro hidrofóbico de várias proteínas e é formado por cadeias polipeptídicas paralelas e relativamente próximas uma das outras. Representando um conjunto de estruturas com orientações paralelas (todas as finalizações N-terminais de um mesmo lado) ou antiparalelas (todas as finalizações N-terminais em lados opostos). Tais estruturas são chamadas de folhas- $\beta$  (1).

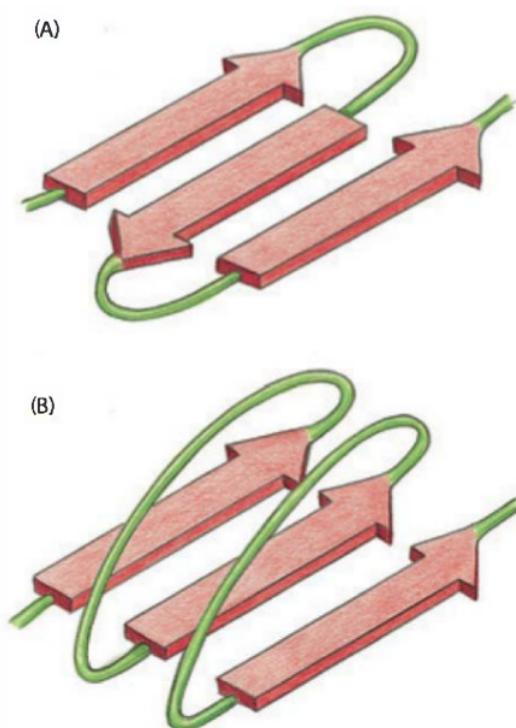


Figura 19 – a) folha  $\beta$  antiparalela, b) folha  $\beta$  paralela. Figura adaptada da referência (1).

As folhas- $\beta$  paralelas e folhas- $\beta$  antiparalelas são estáveis e podem unir-se através de pontes de hidrogênio (1).

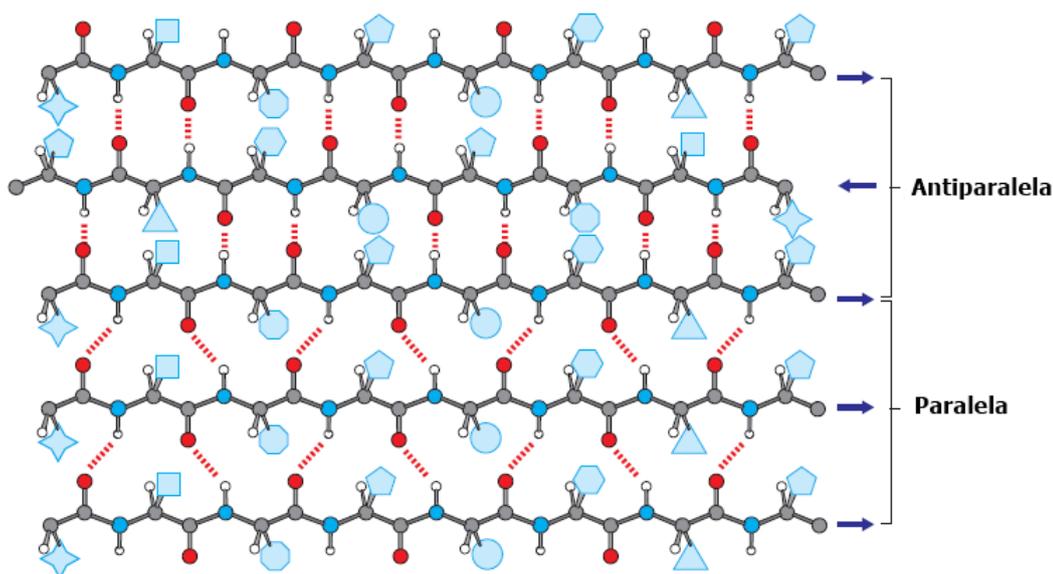


Figura 20 – Tipos de folhas  $\beta$ . Figura adaptada da referência (29).

Note que as sequências de aminoácidos das cadeias principais apresentam uma orientação definida dos terminais (N,C) :

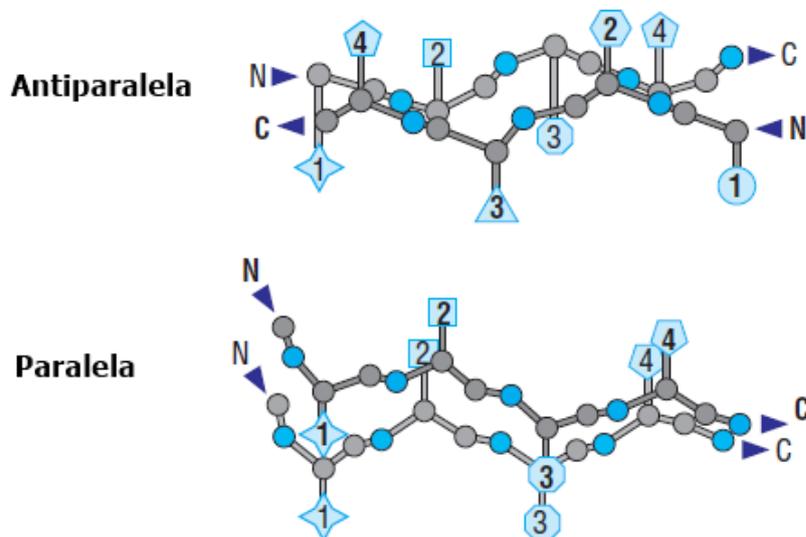


Figura 21 – Orientação dos terminais (N,C) da folha  $\beta$ . Figura adaptada da referência (29).

Como exemplo, temos o complexo protéico com a sigla PDB 1rlb pois é encontrado no banco de dados de proteínas conhecido como Protein Data Bank (PDB) e a sua estrutura é uma folha- $\beta$  antiparalela. Sua funcionalidade consiste em empacotar moléculas apolares tal como o retinol no centro hidrofóbico.

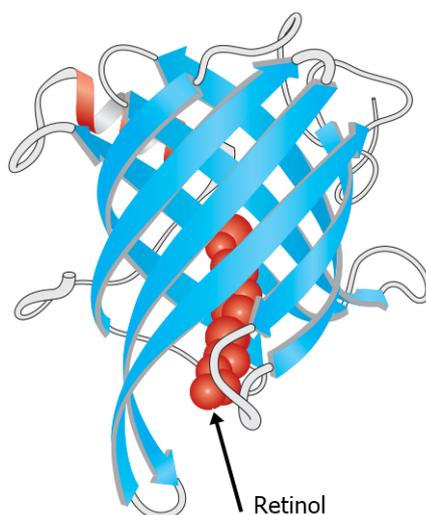


Figura 22 – Proteína PDB 1rlb. Figura adaptada da referência (29).

A barreira beta forma um cilindro fechado tendendo a não permitir a entrada de água ( $H_2O$ ), com a cadeia lateral hidrofóbica alinhada ao interior da proteína.

### 3.2.3 Volta $\beta$

Para que a folha- $\beta$  fique conectada, é necessário um arranjo chamado de voltas  $\beta$ . Regiões de voltas nas folhas- $\beta$  possuem similares estruturas. Em um conjunto de 72 proteínas mostrou-se que a estrutura volta  $\beta$  ocorria com mais frequência quando é formada apenas por duas cadeias laterais e podem ser encontradas em dois formatos diferentes.

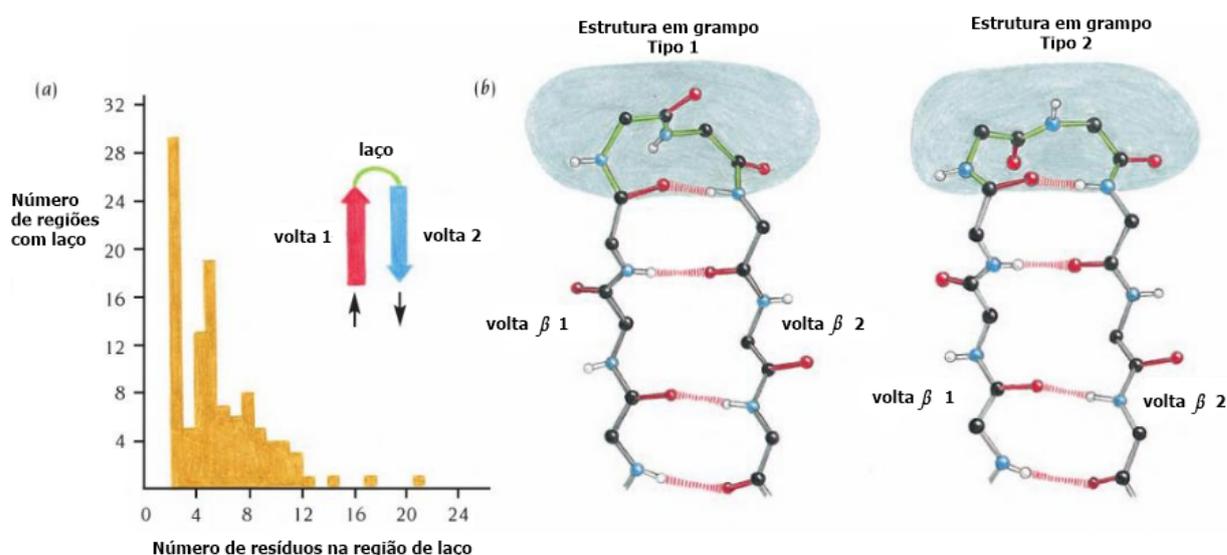


Figura 23 – Dois tipos regiões volta  $\beta$  (tipo 1, tipo 2) por quantidade de aminoácidos. Figura adaptada da referência (2)

Em geral, são encontradas em dois tipos: tipo 1) o carbono representado na cor vermelha projetado para fora e tipo 2) carbono projetado para dentro (2).

### 3.3 Estrutura terciária

O dobramento e conformação dos domínios está associado à estrutura terciária. Os domínios enovelam independentemente na cadeia polipeptídica, em geral cada particular domínio está relacionado às funções, as quais podem ser encontradas em mais de uma proteína (3).



Figura 24 – Domínios da estrutura terciária. Figura adaptada da referência (3).

Sendo assim, a estrutura secundária enovela em domínios, formando assim uma unidade (monômero) protéica (30). Cada domínio possui funções específicas, e pode ser encontrado em diferentes proteínas. Os domínios encontrados em muitas proteínas diferentes são conhecidos como módulos (2). Portanto, a seguir são apresentados os tipos de interações que cooperam para estabilizar os domínios:

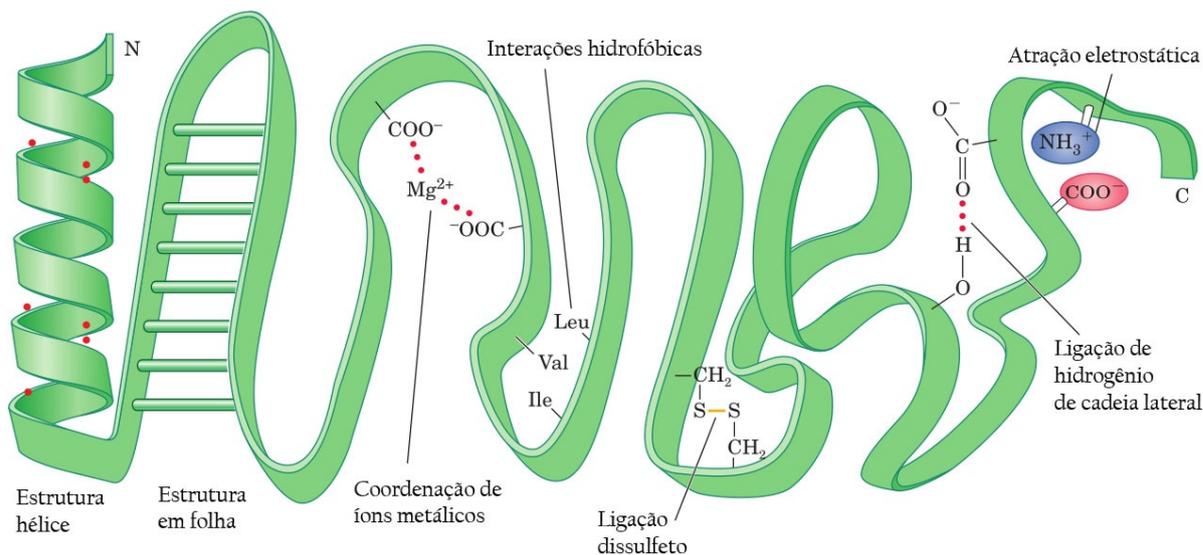


Figura 25 – Ligações que estabilizam a estrutura terciária. Figura adaptada da referência (3).

As características fundamentais da estrutura terciária não inclui apenas as interações entre as cadeias principais. Ela também inclui as interações entre as cadeias laterais R dos aminoácidos, tais como: pontes dissulfeto, pontes de hidrogênio, interações hidrofóbicas, atração eletrostática, ligações coordenadas de íons metálicos. Então, as estruturas terciárias são formadas a partir de estruturas secundárias, que possuem a característica de dobrarem sobre si mesmas.

### 3.4 Estruturas quaternária

As estruturas das proteínas com mais de uma cadeia polipeptídica, que apresentam um nível mais alto de organização proteica, são denominadas quaternárias. Dois ou mais monômeros (unidades) de domínios caracterizam uma estrutura quaternária. Caso sejam duas unidades é denominada “dimérica”, se forem três, “trimétrica”, em caso de várias, “multimérica”. Em geral, as subunidades quaternárias podem trabalhar independentemente ou cooperativamente: Por exemplo: a hemoglobina, forma um “tetrâmero” (quatro monômeros de domínios) trabalhando cooperativamente (3).

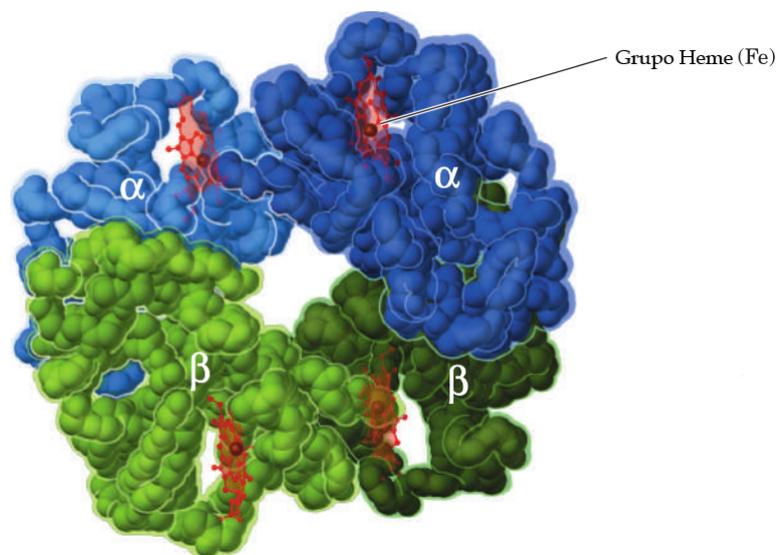


Figura 26 – Estrutura quaternária da hemoglobina. Figura adaptada da referência (3).

A hemoglobina tem duas hélice- $\alpha$  idênticas subunidades quaternárias, e duas idênticas subunidades quaternárias de folha- $\beta$ , sendo sua estrutura global ( $\alpha_2 - \beta_2$ ). As subunidades quaternárias são mantidas compactadas pelas mesmas ligações que favorecem a estrutura terciária: interações hidrofóbicas, pontes de hidrogênio, pontes salinas e atrações eletrostáticas (3).

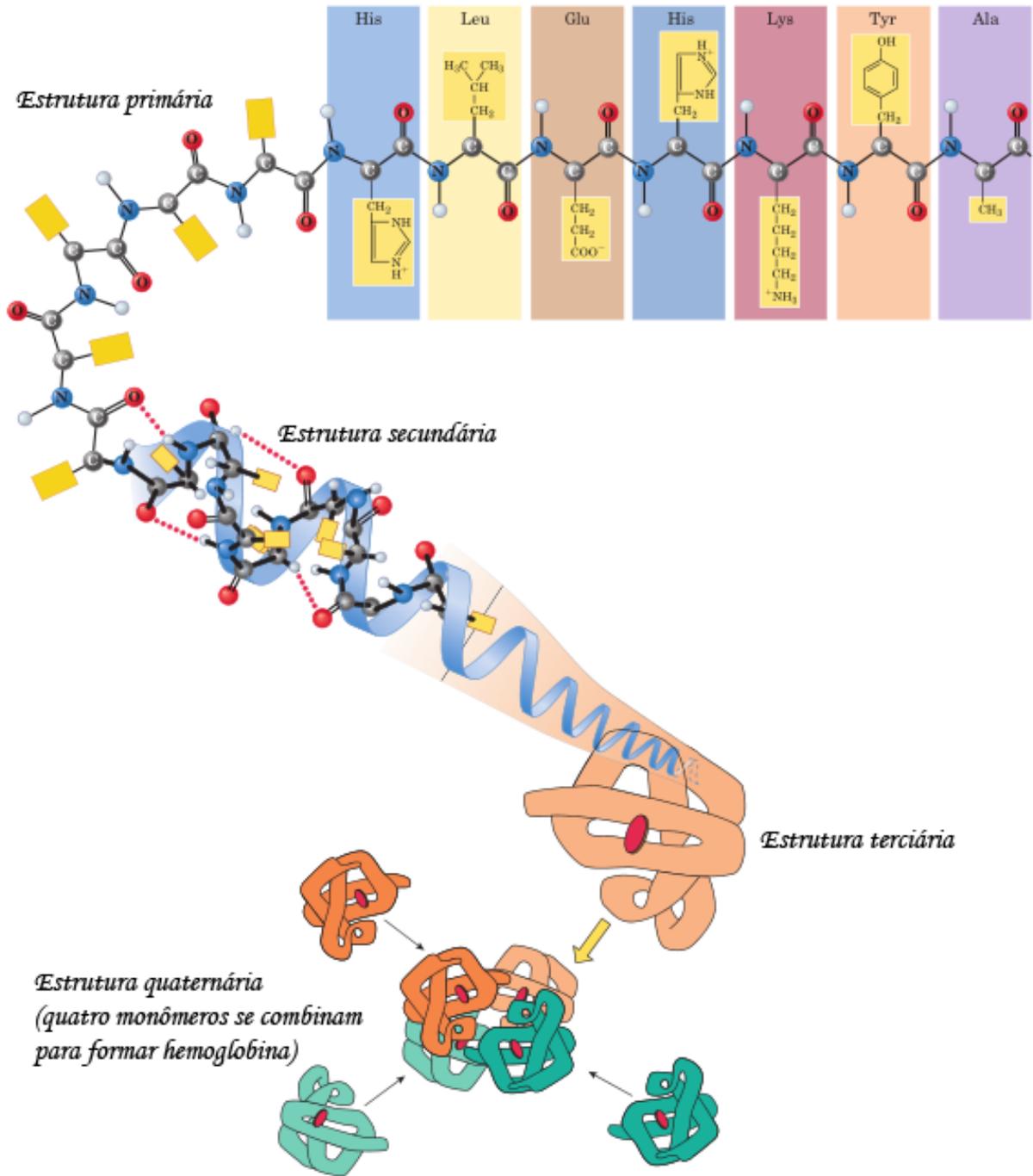


Figura 27 – Etapas do enovelamento da estrutura quaternária. Figura adaptada da referência (3).

Neste caso, a proteína é composta de sete aminoácidos alanina, tirosina, lisina, histina, ácido glutâmico, leucina, histina unidos por ligações peptídicas representando a estrutura primária, a partir de então enovela-se formando a estrutura secundária e depois na sua forma mais evoluída terciária (monômeros). Por final, a junção de monômeros caracteriza uma estrutura quaternária (3).

## 4 Enzimas

### 4.1 Modo de ação das proteínas

Todas as proteínas ligam-se a outras moléculas. A substância que se liga à proteína é chamada de ligante. A região aonde ocorre este processo, é chamada de sítio de ligação (1).

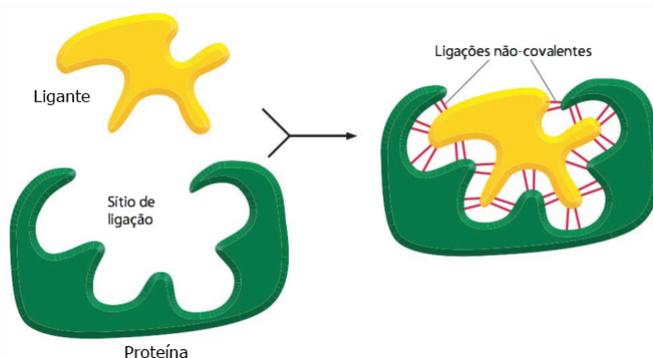


Figura 28 – Modo de ação das enzimas. Figura adaptada da referência (1).

As ligações são do tipo não covalente – ligações de hidrogênio, atrações eletrostáticas (Coulomb) e de van der Waals (ligações entre dipolos elétricos).

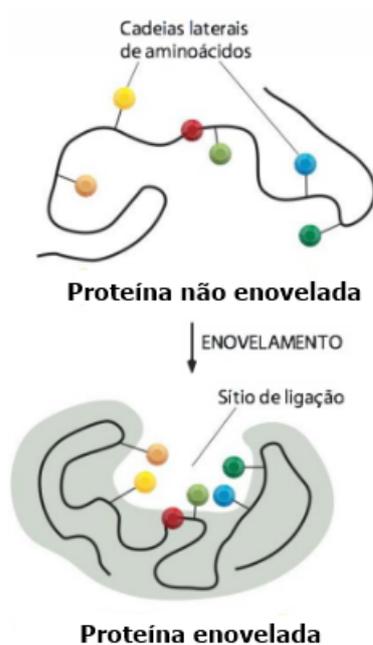


Figura 29 – Sítio de ligação das proteínas, Figura adaptada da referência (1).

A proteína enovelada cria uma fenda (cavidade) na sua superfície com as cadeias laterais (radicais), fazendo a interação com seus ligantes. Este agrupamento de cadeias laterais pode alterar sua reatividade.

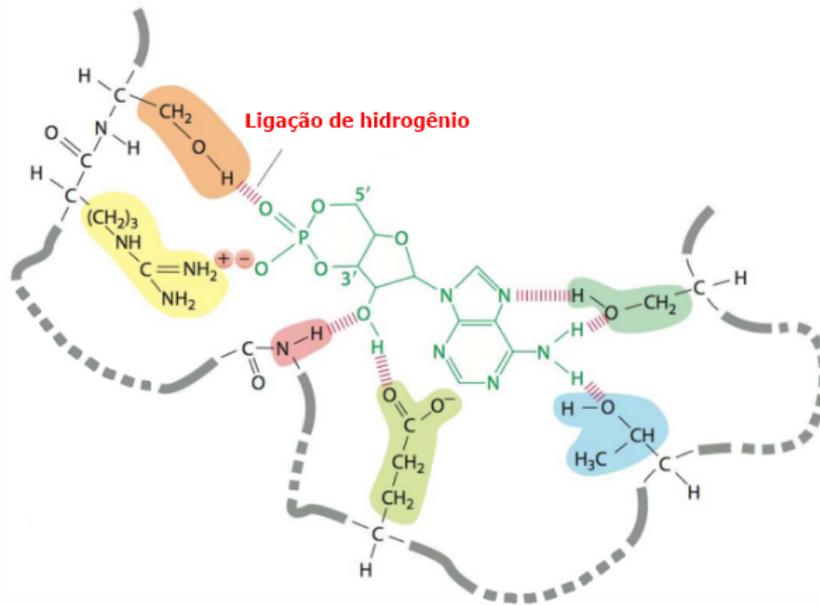


Figura 30 – Sítio de ligação com substrato (1).

As interações das proteínas com seus ligantes formam fortes ligações de hidrogênio e interações eletrostáticas, apenas se a proteína puder excluir as moléculas de água do seu sítio de ligação. É difícil de se imaginar tal mecanismo, entretanto para a molécula de água individual é energeticamente desfavorável afastar-se da sua rede de ligação (1). As proteínas podem-se ligar para formar outras proteínas de três maneiras.

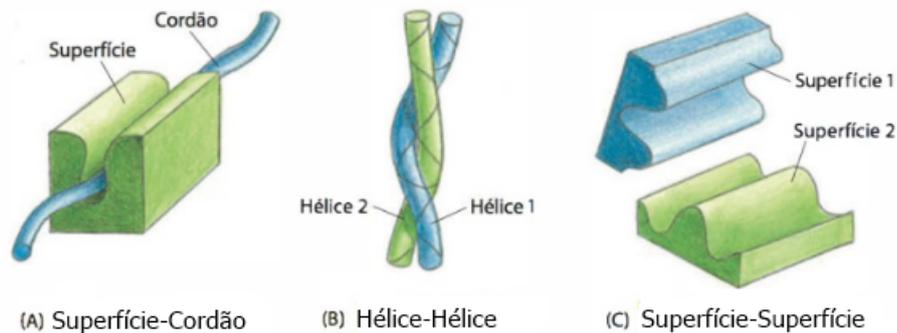


Figura 31 – União de proteínas. Figura adaptada da referência (1).

Estas possibilidades de ligação permitem criar proteínas altamente evoluídas, como por exemplo, anticorpos ou imunoglobulinas que são proteínas produzidas pelo sistema imunológico em resposta a moléculas estranhas presentes na superfície de micro-organismos invasores (1).

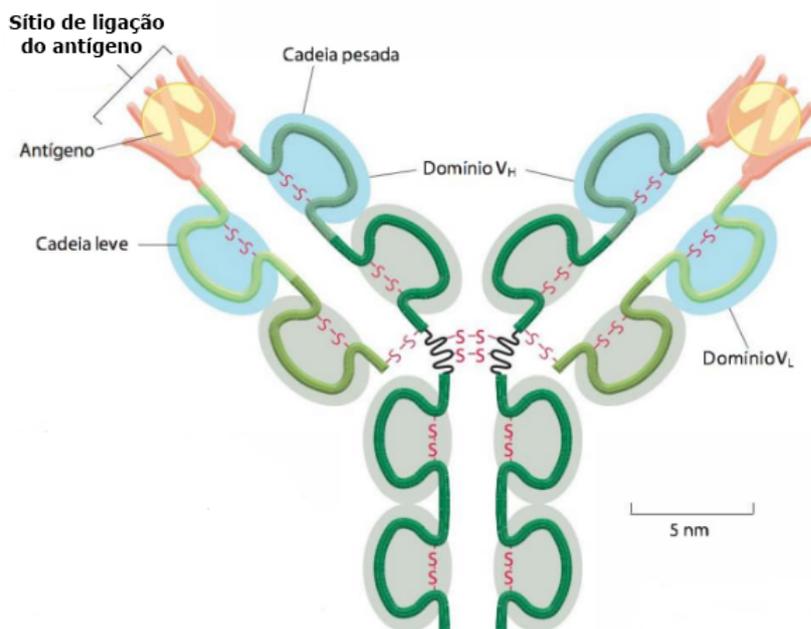


Figura 32 – Anticorpos são enzimas. Figura adaptada da referência (1).

A molécula alvo (antígeno) no caso é o invasor. O anticorpo reconhece e liga-se ao invasor ou à molécula alvo (antígeno). Isto ocorre, devido a alta especificidade das enzimas. O resultado, é o antígeno sendo inativado ou marcado para ser aniquilado (1).

## 4.2 Características

Nas células ocorrem reações as quais devido à sua complexidade, supõem-se serem muito lentas nas temperaturas que se processam (em torno de 37°). Para surpresa, essas reações são rápidas, o que leva à concluir que existem substância catalisadoras. Estas substâncias são as conhecidas enzimas, formadas no interior das células, porém também com capacidade de ação fora delas (11, 1).

Louis Pasteur (1822-1895) foi quem descobriu as enzimas, concluindo que a fermentação do açúcar em álcool por levedura é catalisada por fermento e era inseparável das estruturas das células do levedo. Wilhelm Kühne (1837-1900) foi quem denominou o termo "enzima" para descrever esse fermento. Este termo passou a ser utilizado para as proteínas com capacidade catalítica, enquanto o termo "fermento" referia-se à atividade exercida por organismos vivos (11, 1).

O prêmio Nobel de química de 1907 foi para Eduard Buchner, por descobrir que os extratos de levedo tinham capacidade de fermentar o açúcar em álcool e provou que as enzimas envolvidas na fermentação continuavam funcionando mesmo quando removidas das células (11, 1). Havia dúvidas em relação à natureza das enzimas. A resposta veio com James Batcheller Sumner (1887-1955) que purificando e cristalizando as enzimas urease e

catalase descobriu tratar-se de uma proteína. Este processo abriu a possibilidade de que as estruturas moleculares pudessem ser analisadas por cristalografia de raios X, revelando as estruturas atômicas. A primeira a ser examinada desta maneira foi a lisozima, encontrada na saliva, lágrimas e na clara do ovo, e com a capacidade de destruir a parede celular de bactérias (11, 1).

As enzimas ligam-se a um ou mais ligantes, chamados substratos, e os convertem em um ou mais produtos quimicamente modificados, fazendo isso muitas vezes e de uma maneira mais rápida. Aceleram reações, frequentemente por fatores de milhões de vezes ou mais, sem que elas próprias sejam modificadas. As enzimas modificam moléculas de acordo com a necessidade do organismo. A alta especificidade significa que as proteínas reagem com moléculas específicas ou grupos de moléculas específicas (11, 1).

<b>Enzima</b>	<b>Reação Catalisada</b>
Hidrolases	Clívagem hidrolítica
Nucleases	Clívagem de ácidos nucleicos .
Proteases	Clívagem de proteínas.
Sintases	Síntese de proteínas.
Isomerases	Rearranjo das ligações de uma molécula.
Polimerases	Reações de polimerização como a síntese de DNA e RNA.
Cinases	Adição de grupos fosfato à moléculas.
Fosfatases	Remoção hidrolítica de grupos fosfato de uma molécula.
Óxido-redutases	Reações para oxidar e reduzir moléculas..
ATPases	Hidrólise de ATP.

Tabela 3 – Funções das enzimas (11, 1)

A catálise faz acelerar as reações químicas. Algumas das mais importantes é a hidrólise, ou seja, a quebra da ligação química entre moléculas com adição de água. Já na clivagem a quebra da ligação química faz divisão da molécula, dando origem a outras moléculas. Quando a clivagem é hidrolítica, a quebra acontece somente na presença de água e no caso de polissacarídeos se denomina amilase (11, 1). A concentração de substratos em enzimas tem um influência na sua velocidade de reação com “cinética do tipo michaeliano ou hiperbólico”. A reação enzimática tem sua velocidade aumentada até um valor máximo, quando isso ocorre todos os sítios de ligação estão totalmente ocupados pelos substratos e por isso o processo é limitado por um valor de velocidade máxima  $V_{máx}$  (11, 1).

A parte protéica de uma enzima é chamada de apoenzima, mas a parte não protéica é chamada de cofator e caso seja um composto orgânico é conhecido como coenzima. Um importante exemplo é a vitamina B, vital para atividade de várias enzimas por ligar-se junto com o substrato no sítio ativo proporcionando um melhor funcionamento (11, 1).

Os fatores que influenciam na atividade enzimática são a concentração da enzima e

do substrato, temperatura e pH. Para o caso da concentração de substratos, temperatura e pH constantes, temos uma taxa de reação linear de acordo com o aumento de enzimas.

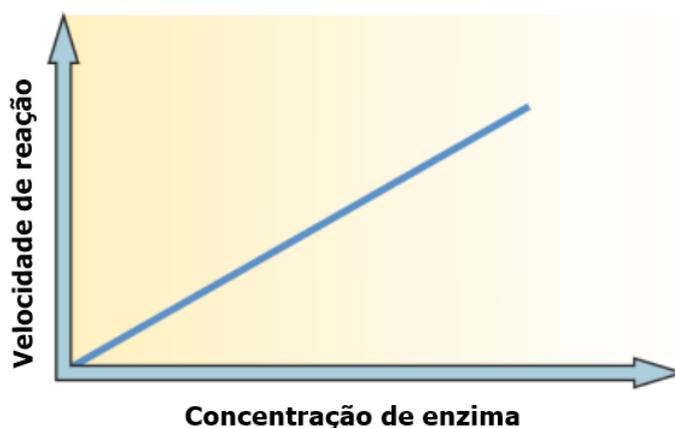


Figura 33 – Velocidade da reação em função da concentração de enzima. Figura adaptada da referência (3).

No caso da concentração de enzimas, temperatura e pH constantes: temos um comportamento hiperbólico da taxa de reação em função da concentração de substratos.

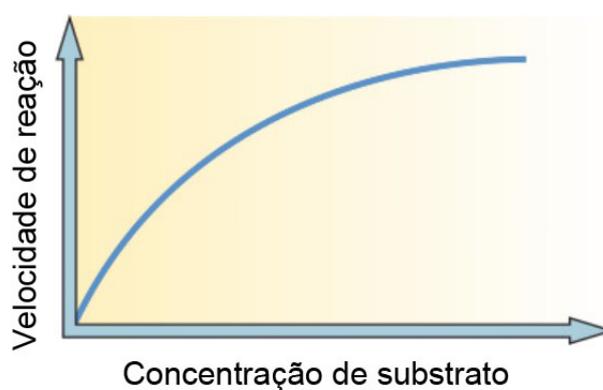


Figura 34 – Velocidade da reação em função da concentração de substrato. Figura adaptada da referência (3).

No caso dos substratos, enzimas e pH constantes, o efeito da temperatura na taxa de reação é uma atividade máxima para os valores de temperatura em torno de 37°C:

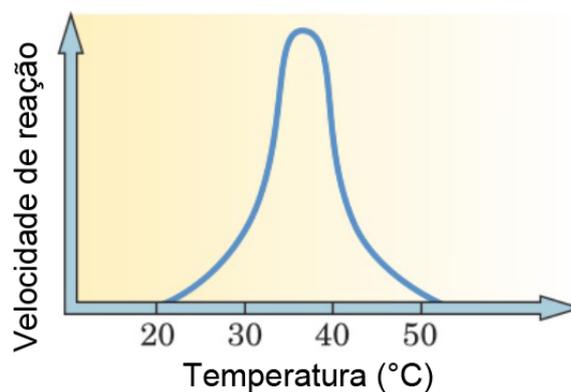


Figura 35 – Velocidade de reação em função da temperatura. Figura adaptada da referência (3).

E por último para a situação da concentração de substratos, enzimas e temperatura constantes. A taxa de reação química possui um valor máximo, próximo da neutralidade com pH em torno de 7:

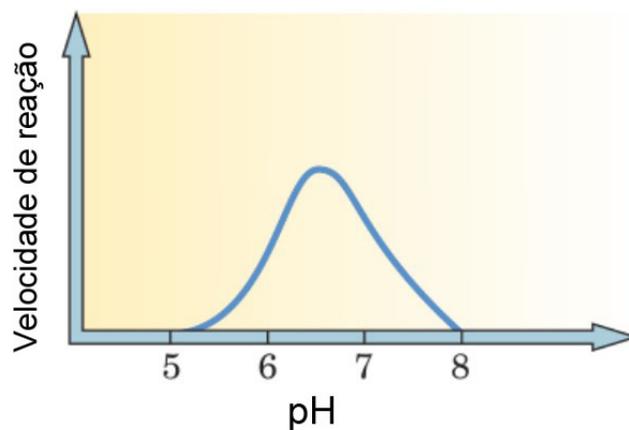


Figura 36 – Velocidade de reação em função do pH. Figura adaptada da referência (3).

O modo de ação das enzimas pode também alterar sua conformação e mudar os parâmetros da cinética enzimática. No modelo mais atual conhecido como encaixe induzido, a enzima muda sua geometria de modo que o substrato possa encaixar-se no sítio ativo.

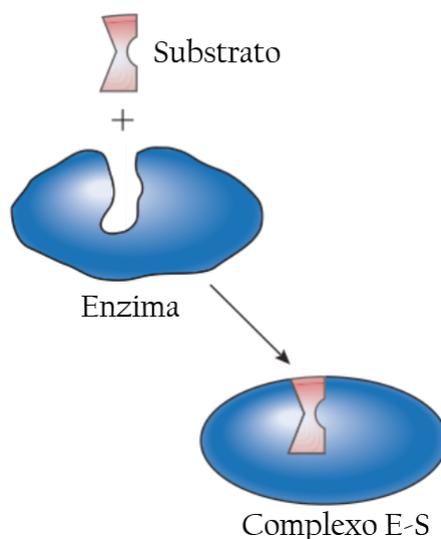


Figura 37 – Modelo encaixe induzido. Figura adaptada da referência (3).

Para que as enzimas possam reagir quimicamente com as moléculas da solução, é necessário energia para que favoreça o encontro e a colisão entre elas, a qual é chamada de energia de ativação (3, 1).

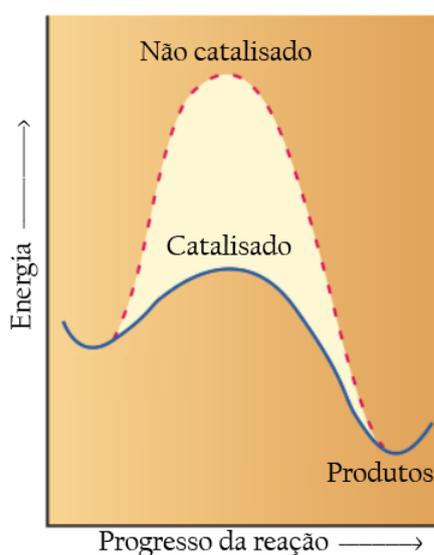


Figura 38 – Energia de ativação para reações catalizadas e não-catalizadas. Figura adaptada da referência (3).

Como pode-se perceber há uma diminuição da energia de ativação na presença de enzimas (catálise) e isso favorece a ocorrência da reação química (3, 1). As enzimas atuam de maneira geral de três modos:

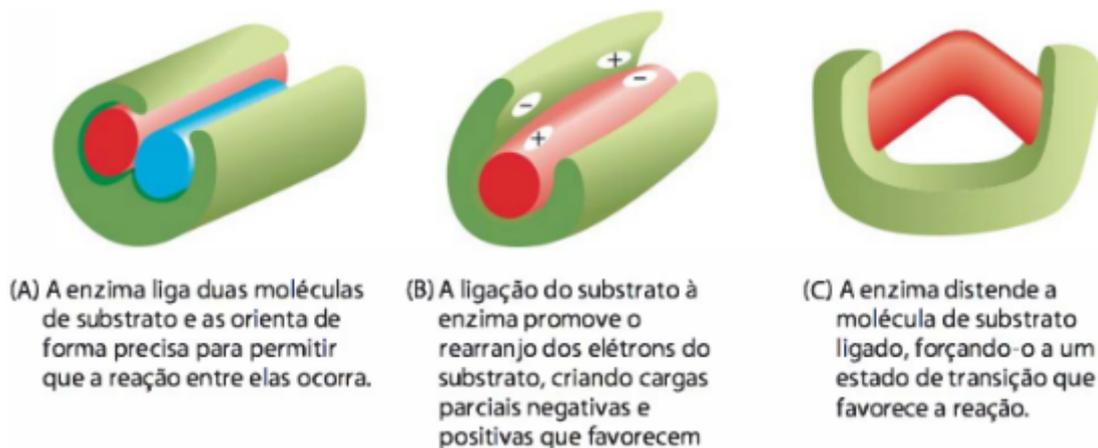


Figura 39 – Estratégias de atuação das enzimas. Figura adaptada da referência (1)

Em relação a figura acima, temos a) alinhando os substratos de forma precisa, b) estabilizando as cargas elétricas de tal forma a favorecer a reação, c) aplicando forças que distorcem as ligações do substrato fazendo aumentar a cinemática enzimática. As enzimas evoluíram de forma a direcionar seus substratos por túneis moleculares em múltiplos sítios de ligação indicada pela faixa vermelha (1).

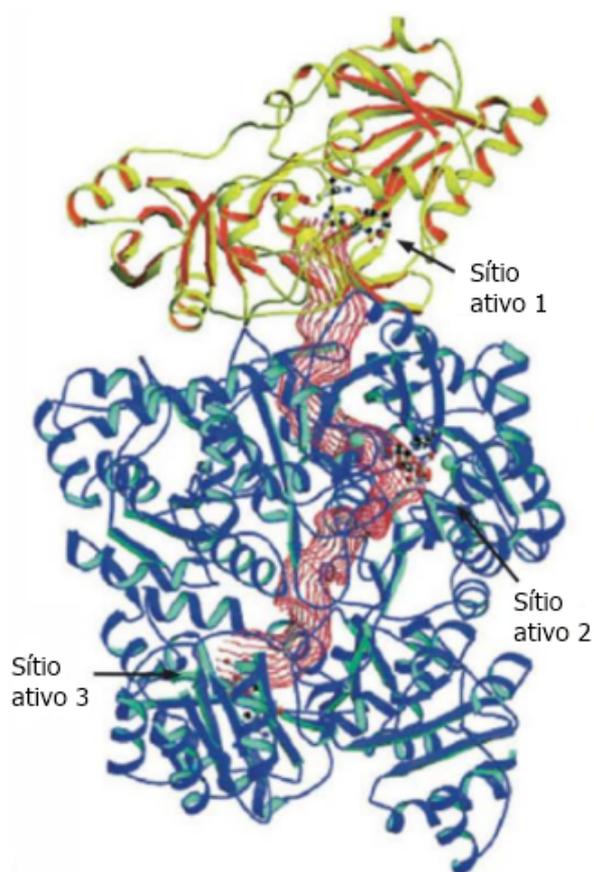


Figura 40 – Visualização em fitas de três proteínas interligadas por um túnel. Figura adaptada da referência (1)

É importante lembrar que ocorre difusão, ou seja, o transporte de matéria de uma região de maior concentração de substrato para uma região de menor concentração até atingir o equilíbrio (1). Como tem uma grande quantidade de diferentes reações na célula, há um limite de metabólitos (produtos do metabolismo de alguma molécula ou substância) que está presente em concentrações micromolares ( $10^{-6} \mu M$ ) e as enzimas estão presentes em concentrações mais baixas (1). Com o objetivo de manter uma taxa metabólica rápida as células desenvolveram-se de modo a aproximar grupos de enzimas (complexo multienzimático), cada grupo realizando funções específicas, objetivando maximizar o número de reações, podemos inferir que a natureza sempre utilizou o caminho que nós denominamos de "revolução industrial"(1). As células podem inibir as atividades das enzimas. A molécula inibidora pode-se ligar à enzima de modo a diminuir a taxa de reação.

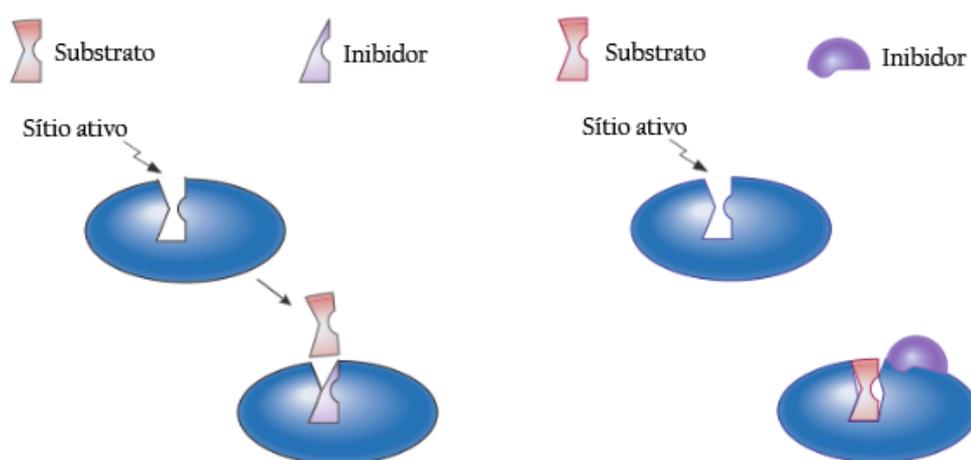


Figura 41 – Inibição competitivo (esquerda) e não competitivo (direita). Figura adaptada da referência (3).

A taxa é máxima quando a enzima não possui inibidor. Seu valor é mínimo, para o caso de ligar a um inibidor não competitivo.

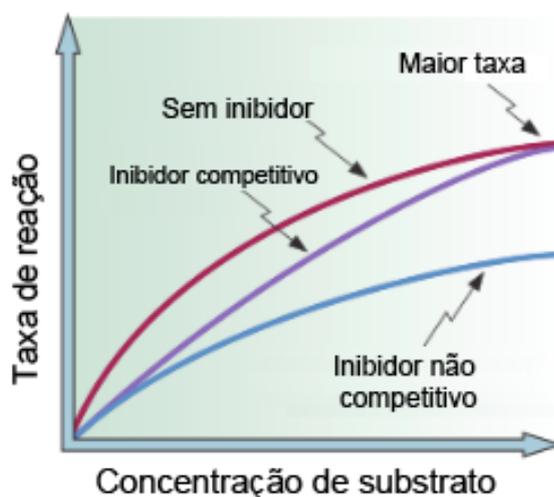


Figura 42 – A cinética enzimática. Figura adaptada da referência (3).

Existem formas de regular atividades enzimáticas, uma molécula pode ativar ou desativar uma determinada enzima.

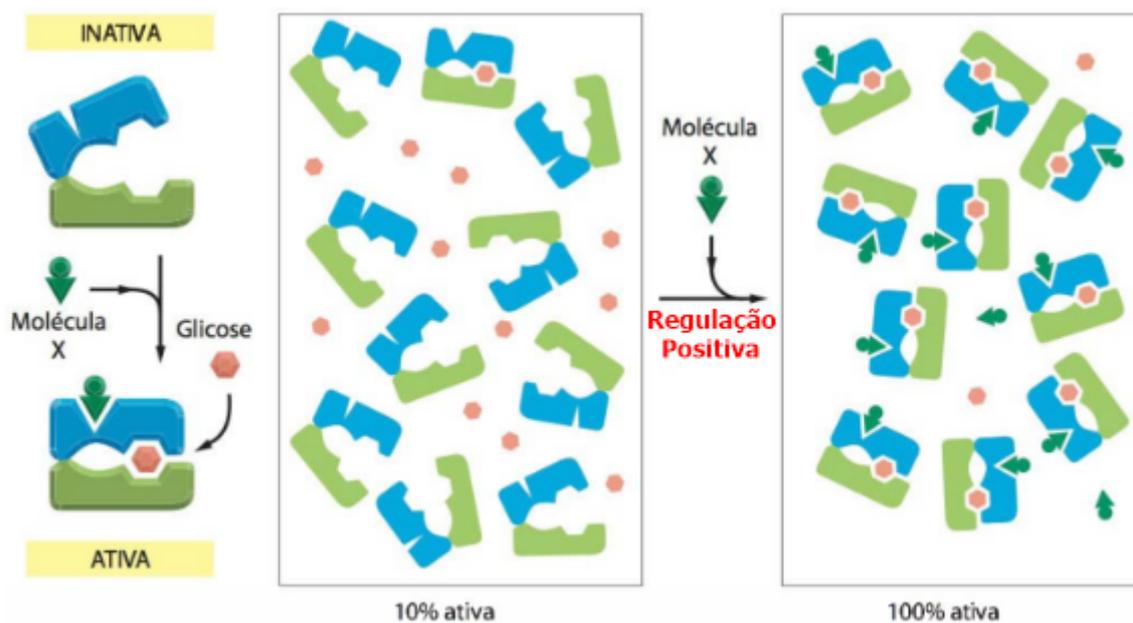


Figura 43 – Moléculas podem ativar enzimas. Figura adaptada da referência (1).

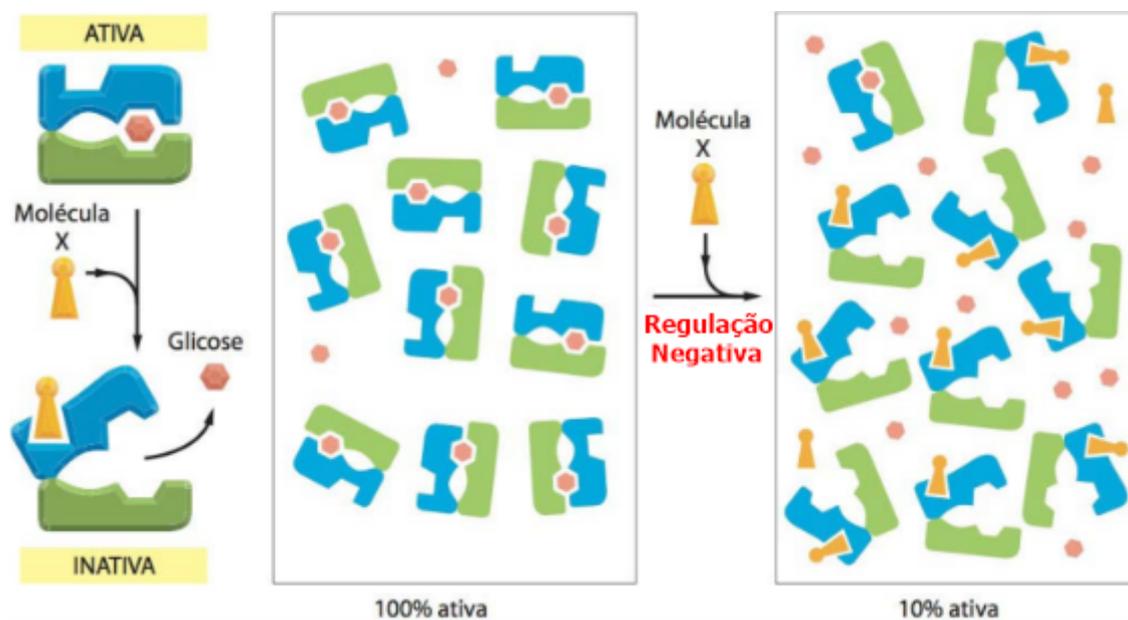


Figura 44 – Moléculas ativam ou desativam enzimas. Figura adaptada da referência (1).

Outro exemplo, é o controle por retroação, ou conhecida por feedback:

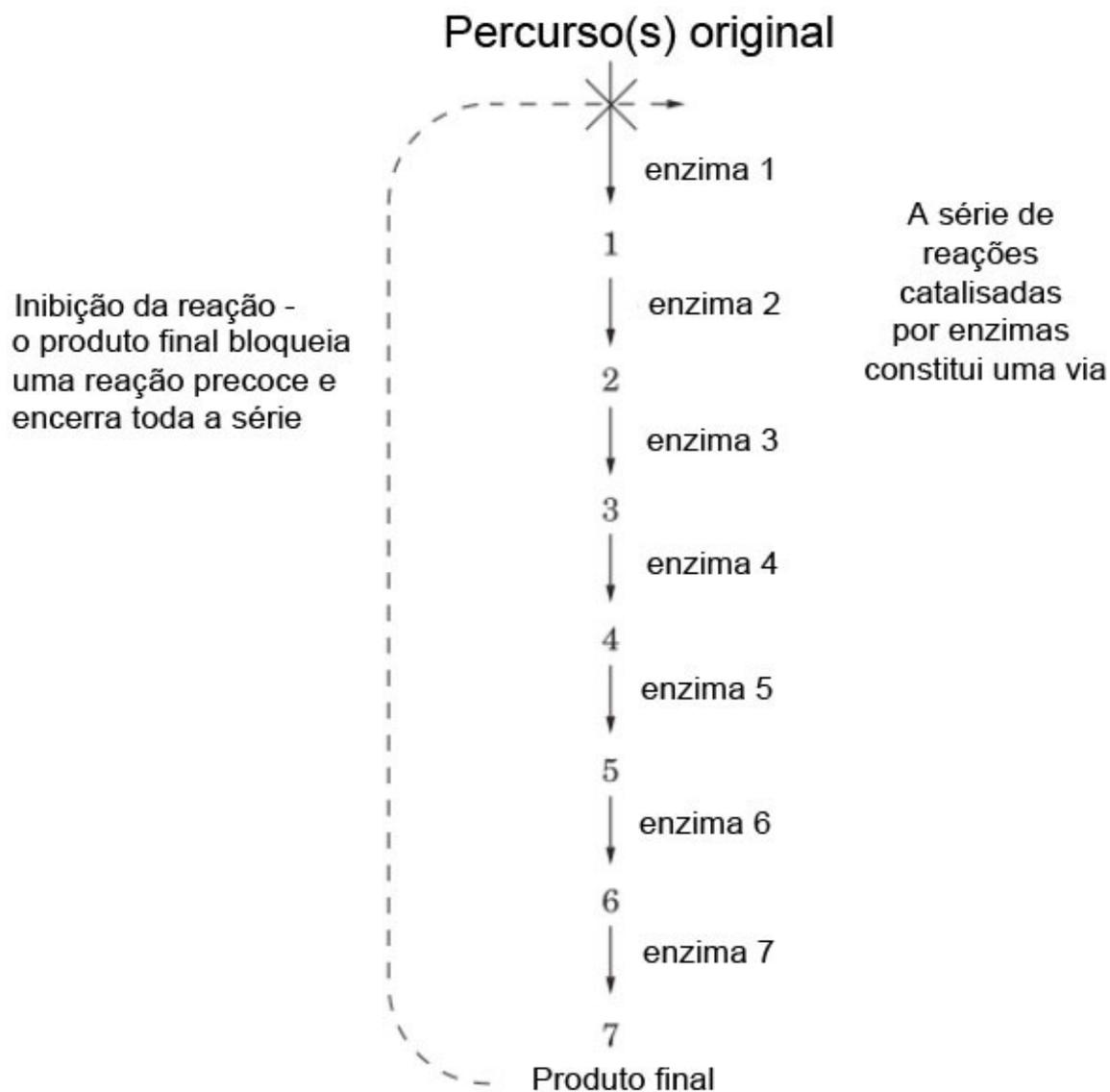


Figura 45 – Controle por feedback. Figura adaptada da referência (3).

Sendo assim, o produto final bloqueia a reação inicial, interrompendo todo o processo. É importante para a célula ter o controle das reações químicas e poder interferir no funcionamento das enzimas.

## Capítulo 5

# A enzima: Pectina metilesterase

### 5.1 Substrato: Pectina

As paredes celulares, que envolvem as células vegetais, conferem características particulares a esses organismos. Essas paredes são encarregadas de fornecer a forma e a proteção dessas células (31, 32). Os constituintes da parede celular são polímeros de celulose (constituídos por uma grande quantidade de moléculas de glicose), hemicelulose (constituído por um grande quantidade de hexoses, pentoses e ácidos urônicos) e pectina (constituídos em maioria por ácidos galacturônicos metoxilados) (32). Em geral, os polisacarídeos (constituídos por carboidratos com função de fornecer energia ao seres vivos) são heterogêneos, como por exemplo a pectina, apresentando diversos açúcares em sua composição: galactose, arabinose e xilose (34).

A pectina é formada por uma cadeia central de 100 monômeros a 1.000 monômeros de ácido galacturônico (32). É importante destacar que a pectina possui caráter hidrofílico, devido à presença de grupos polares. Portanto, tem característica de atrair as moléculas de água ( $H_2O$ ), produzindo uma solução viscosa (35). Nas frutas verdes está contida a protopectina que é a pectina na forma insolúvel e sua ação é similar a de uma "cola", conectando as paredes celulares das diferentes células vegetais (34). Na fase de pleno desenvolvimento das frutas a protopectina é abundante, contudo no processo é transformada em pectina. A função da pectina é dar estrutura aos tecidos vegetais, firmeza e resistência à compressão (32).

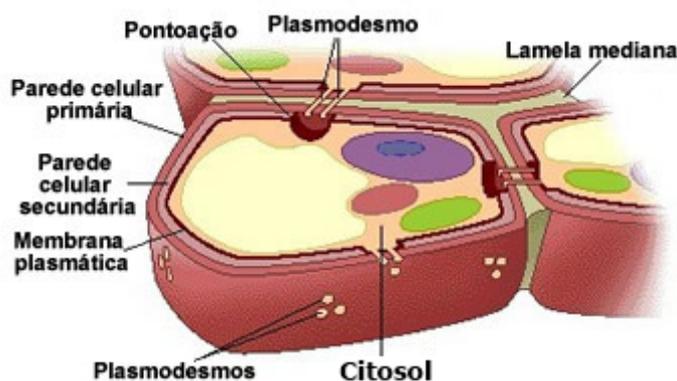


Figura 46 – Célula vegetal. Figura adaptada da referência (33) .

As pectinas estão no grupo das fibras dietéticas. As fibras alimentares ou também conhecida como fibras dietéticas são classificadas em dois grupos: (1) insolúveis: celulose, lignina, e a maioria das hemiceluloses- (2) solúveis: pectina, gomas, mucilagens e algumas hemiceluloses (32). Comercialmente é produzido a pectina como um pó branco de cor castanho claro, com aplicações em enchimentos de sobremesas, medicamentos, doces, bebidas com leite e como fonte de fibra alimentar (39).



Figura 47 – Pectina. Figura adaptada da referência (38).

Principalmente extraída de cascas de frutas cítricas, é usada em alimentos para se obter ação geleificante, em especial geléias. Em geral, a pectina é manuseado como geleificante e estabilizante na indústria de alimentos (40).

Origem	Pectina %
Batata	2,5
Tomate	3,0 .
Maçã	5,0 a 7,0.
Beterraba	15,0 a 20,0.
Frutas cítricas	30,0 a 35,0.

Tabela 4 – Porcentagem de pectina nos tecidos vegetais (36).

A pectina é formada principalmente por uma estrutura central de ácidos galacturônicos, altamente metoxilado (ATM), com no mínimo 65% de grupos metoxilados ( $\text{COOHCH}_3$ ). Aproximadamente 80 % da produção mundial desta modalidade é usada em geleias e compotas (conservas) devido a seu poder geleificante (40, 36).

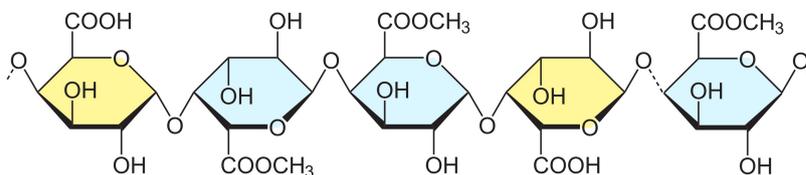


Figura 48 – Ácidos galacturônicos conectados com grupos metoxilados ( $\text{COOHCH}_3$ ), destacados em azul, formam a pectina. Figura adaptada da referência (41).

Os ácidos galacturônicos ligados em maioria ao grupo carboxílico ( $\text{COOH}$ ), fica conhecido por ácido péctico (42). Enquanto, se os ácidos galacturônicos estiverem ligados em maioria ao grupo metoxilado ( $\text{COOHCH}_3$ ) ficam conhecidos por ácidos pectínicos (42). A cadeia principal fica interligada entre os monômeros por meio de ligações glicosídicas  $\alpha$  (1 $\rightarrow$ 4)

Em relação a suas aplicações, em um sentido químico, a pectina está incluída no grupo das substâncias pécticas: pectinas, ácido péctico, ácido pectínico e protopectina. Além disso, as pectinas tem como característica formar gel em meio ácido e na presença de açúcar, entretanto, para a pectina de baixa metoxilação (BTM), ou seja, abaixo de 65% de ácidos galacturônicos ligados ao grupo metoxilado, este efeito só ocorre na presença de íons metálicos divalentes, como o  $\text{Ca}^{2+}$  (42). Agora em relação a fatores físicos, o efeito da diminuição da temperatura até um ponto crítico, abaixo desta temperatura as pectinas de baixa metoxilação (BTM), irão geleificar quase que instantaneamente, enquanto as de alta metoxilação (ATM), ainda dependerá do fator do tempo necessário para chegar a tal condição, além de que os géis serem termorervíseis (36).

## 5.2 Pectinases

Os termos enzimas pectinolíticas ou pectinases referem-se a ação sobre as substâncias pécticas presentes nos tecidos vegetais (37). As pectinases têm aplicações industriais na extração e clarificação de sucos de frutas, na maceração de vegetais e frutas, e também na extração de óleos essenciais (37). Alguns seres vivos possuem a capacidade de produzir pectinases, incluem plantas, bactérias, leveduras e fungos (42). A classificação das pectinases é feita sobre a forma de atuação nas substâncias pécticas constituintes da parede celular e da lamela média da célula vegetal.

Existem dois principais grupos os desesterificantes ou desmetoxilantes: como principal representante temos a pectina metilesterase agindo nos grupos metil éster

(metoxilados), desesterificando ou desmetoxilando a pectina convertendo-a em ácido péctico e liberando metanol ( $\text{CH}_3\text{OH}$ .) através de um mecanismo de hidrólise, ou seja a quebra de ligação por meio de adição de água ( $\text{H}_2\text{O}$ ). O outro grupo é das despolimerases: podemos citar a pectina liase (PL) e poligalacturonase com atividade sobre a cadeia de ácidos galacturônicos, despolimerizando as substâncias pécticas por meio da quebra das ligações glicosídicas  $\alpha$  (1 $\rightarrow$ 4) da cadeia principal, este procedimento pode ser por hidrólise ou por transeliminação ou conhecida também como  $\beta$ -eliminação. Além disso as despolimerases são subdivididas conforme sua especificidade em relação ao substrato de preferência (pectina, ácido péctico) e segundo padrões de locais de atuação na cadeia principal poligalacturônica que compõem as pectinas (37).

Enzima	Modo de ação	Substrato preferencial
Poligalacturonase (PG)	Hidrólise de ligações glicosídicas	Ácido péctico
Pectina liase (PL)	Trans-eliminação ou $\beta$ -eliminação	Ácido péctico
Pectina metil esterase (PME)	Desmetoxilação de grupos metil	Pectina.

Tabela 5 – Classificação das enzimas pecnolíticas ou pectinases (42).

Por causa da imensa diversidade de pectinas nas plantas, existem diferentes pectinases com diversas maneiras de atuação: as enzimas poligalacturonase (PG) e pectina liase (PL) atuam na diminuição da viscosidade, enquanto a enzima pectina metil esterase (PME) tem pouco efeito sobre a viscosidade das soluções, com ação sobre a pectina de alta metoxilação (ATM) produzindo uma pectina de baixa metoxilação (BTM) a qual sofre clivagem por meio da enzima poligalacturonase (PG) pois é seu substrato preferencial. A forma da atuação das principais pectinases na pectina podem ser divididos em duas etapas:

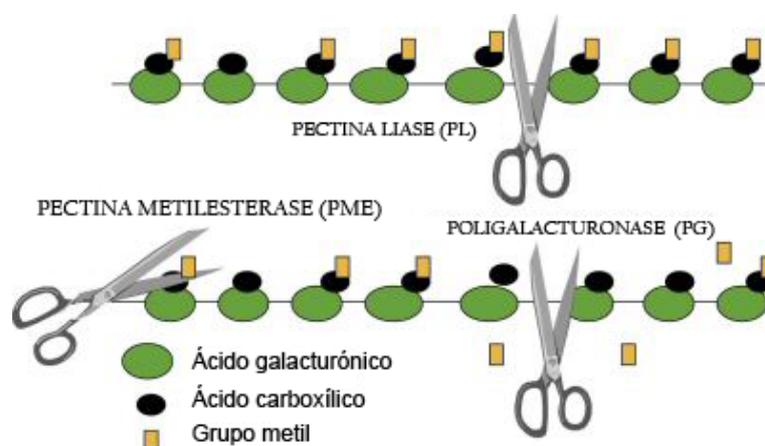


Figura 49 – Formas de atuação das pectinases. Figura adaptada da referência (43).

Em resumo a pectina liase (PL), tem a função de quebrar a ligação dos ácidos pécticos resultando na quebra da pectina, enquanto, no caso da pectina metil esterase (PME) possui a função de quebrar as ligações do grupos metílicos ( $\text{CH}_3$ ) gerando o substrato preferencial da enzima poligacturonase (PG). A quebra em torno de 2-3% de ligações glicosídicas obtém cerca de 50% de redução de viscosidade (32, 42). Algumas das principais funções das enzimas pectinases está na degradação do seu substrato pectina onde induz uma característica visual de clarificação e age na diminuição da viscosidade e também atua na perda da turbidez (31, 32).

### 5.3 Aplicações da pectina metil esterase

As aplicações da enzima pectina metilesterase (PME) incluem amadurecimento de frutas, clarificação e diminuição da viscosidade e melhora da eficiência de filtração em suco de frutas, tratamento preliminar de suco de uva utilizado na preparação de vinhos, e também na extração da polpa do tomate, fermentação de chá e chocolate, tratamento de resíduos vegetais, degomagem (a remoção de fosfolipídeos) de fibras na indústria têxtil e de papel, nutrição de animais e enriquecimento de proteínas nos alimentos infantis e de óleo (31). Na fase de compressão da fruta, as pectinas, por serem solúveis, passam para o suco fazendo com que se torne mais viscoso, com pouca cor e aroma. As fases de processamento de clarificação e filtração são mais difíceis, dificultando o rendimento (34).

Para contornar isso, antes da prensagem adicionam-se preparações enzimáticas especiais no mosto facilitando a futura extração, aumentando consideravelmente o rendimento do suco. O processo completo de despectinização (degradação da pectina) com o uso de enzimas pectinases resulta em uma boa clarificação e filtração do suco (34). Após a prensagem, é necessária a despectinização dos sucos para se conseguir baixa viscosidade. No caso dos sucos concentrados, a despectinização é obrigatória para evitar a geleificação durante a concentração (34).

As pesquisas da produção de pectinases estão direcionadas para os fungos, uma vez que as enzimas comerciais provêm destes organismos com pH ideal dentro da faixa da maioria dos sucos de frutas, em torno de 3,0-5,5 (42). Em geral, o fungo *Aspergillus niger* é utilizado na indústria, na qual é classificado como Generally Recognized as Safe (GRAS) pelo órgão do governo dos Estados Unidos conhecido como Food and Drug Administration (FDA) atuando no controle de qualidade de alimentos e medicamentos (42).

## 5.4 Protein Data Bank (PDB)

Os primeiros resultados da estrutura cristalográfica de raios X foram relatados sobre a proteína mioglobina em 1958 por John Kendrew, e veio como um choque porque esperava-se uma estrutura simples e bonita como a cadeia dupla da estrutura do DNA, descoberta cinco anos antes por James Watson e Francis Crick. No entanto, mostrou-se ser uma estrutura complexa e irregular, para o desapontando da comunidade científica (2).

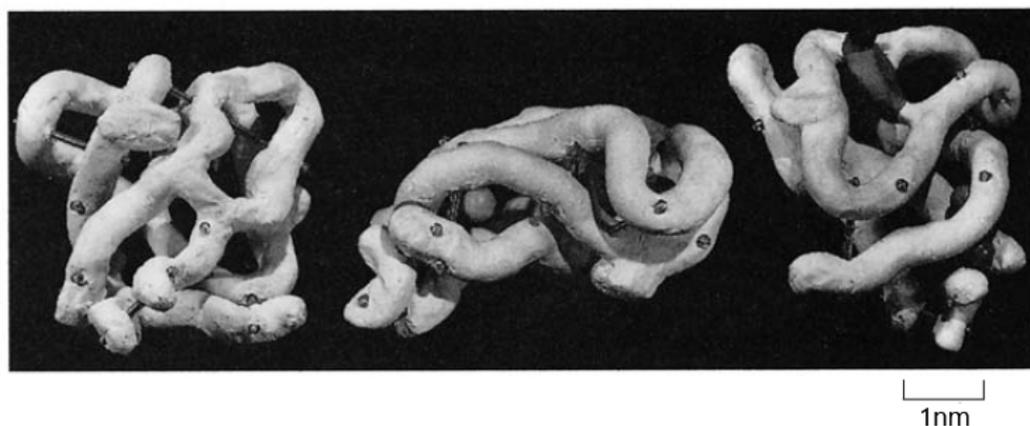


Figura 50 – Estrutura cristalográfica da mioglobina. Figura adaptada da referência (2).

Por meio desta técnica foi possível criar bancos de dados em 3D de proteínas como o Protein Data Bank (PDB), disponível no endereço eletrônico: <<http://www.rcsb.org>>. Estes arquivos no formato PDB contém informações sobre as posições de todos os aminoácidos e de seus respectivos átomos (44, 45).

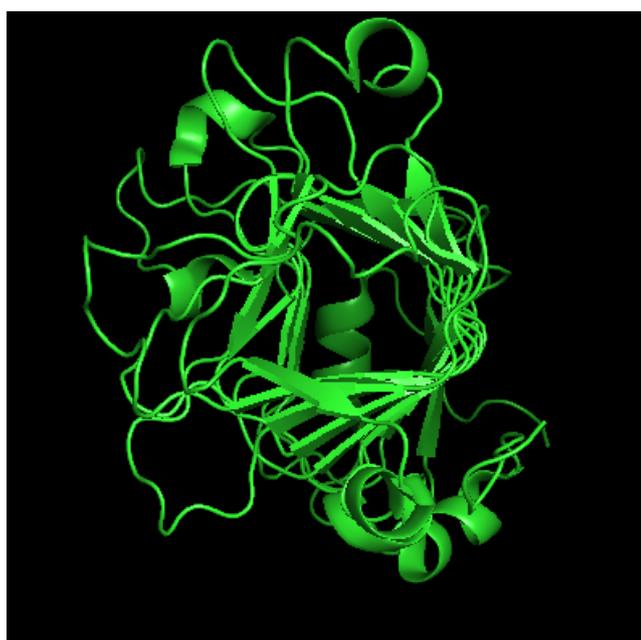


Figura 51 – Visualização da enzima pectina metilesterase no Pymol.

As observações dos arquivos cristalográficos do Protein Data Bank (PDB) em forma de fitas, permite obter informações sobre as estruturas secundárias da proteína de estudo.

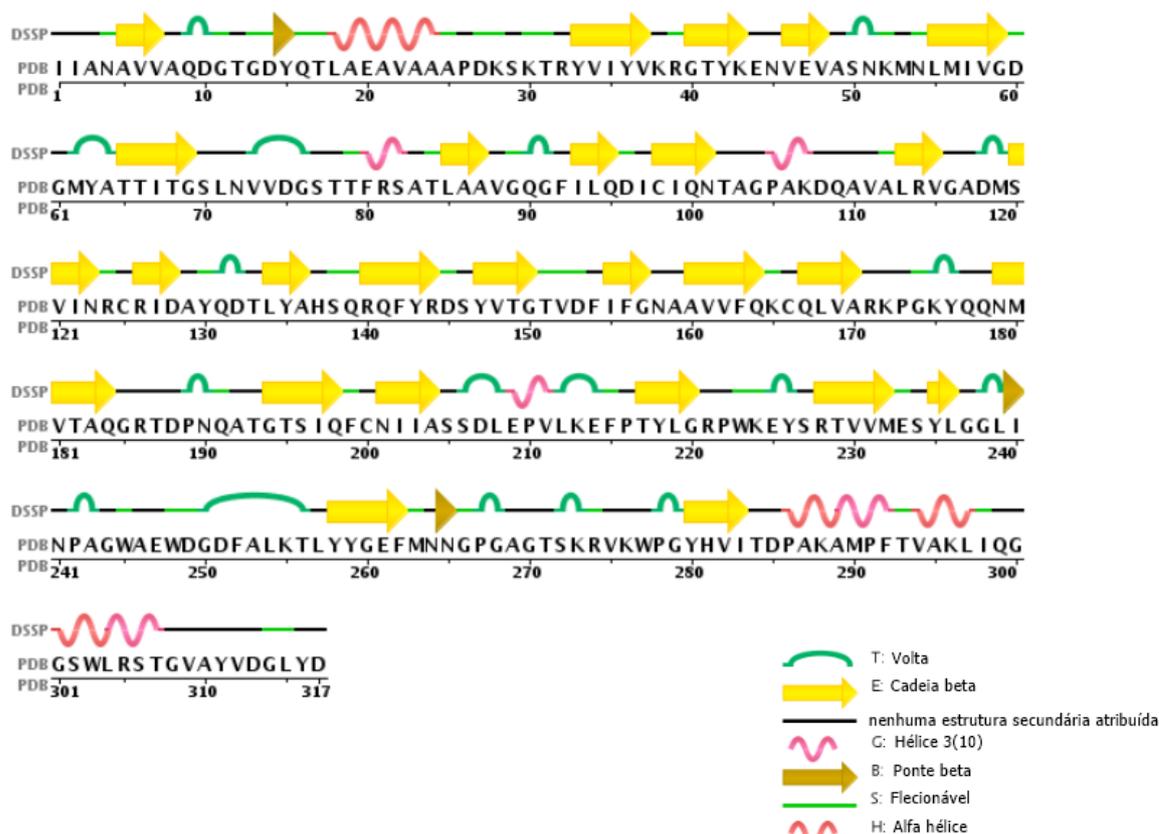


Figura 52 – A composição da enzima pectina metilesterase do tomate, por variados tipos diferentes de estruturas. Figura adaptada da referência (45).

Com as informações das estruturas de proteínas obtidos por meio do Protein Data Bank (PDB) é possível extrair informações e descobrir uma função biológica, pode-se também analisar suas propriedades por meio de simulações computacionais (44, 46, 1, 11, 2).

## 5.5 Influência da pressão nas atividade enzimática

Na indústria de alimentos a qualidade dos alimentos, em geral, é baseada nas características sensoriais e nutricionais, de forma a proporcionar benefício a saúde e vida útil dos produtos (47). As etapas de conservação de alimentos mais utilizados para que satisfaça esas condições incluem: pasteurização, esterilização, branqueamento, desidratação e secagem como forma de diminuir a atividade da água, e também a refrigeração e congelamento (47). Os processos térmicos produzem alterações na qualidade dos alimentos, são exemplos: componentes vitamínicos, modificação da textura e da cor, podendo adquirir sabores ou odores indesejáveis ("off-flavours") (47, 48).

Para superar essas dificuldades foi necessário desenvolver novos métodos tecnológicos. De modo que complementam ou substituem a manutenção das características dos alimentos, de forma a reduzir a utilização da energia térmica na conservação e processamento alimentar. Destaca-se em especial o mecanismo a alta pressões hidrostáticas designado pelas nomenclaturas internacionais como high pressure processing (HHP) (47, 48).

Significa em sujeitar o alimento sólido ou líquido, pré-embalado ou pós-embalado a pressões que podem variar entre 1 kbar-10 kbar. As pesquisas nesta área aponta uma preservação no valor nutricional dos alimentos em relação à aplicada à temperatura ambiente possuindo um efeito mínimo sob os compostos bioativos (nutrientes não essenciais) de vários frutos e vegetais comparado às técnicas usuais.

De uma maneira geral, pode-se resumir os diversos efeitos e objetivos: inativação microbiana, ativação/inativação enzimática, prolongamento da vida útil do produto e permanência das características de sabor, valor nutricional e cor (47, 48). Com o intuito de investigar esta questão de altas pressões sobre proteínas, foi realizada uma análise do raio de giro para diferentes pressões.

## Capítulo 6

# Simulação computacional

### 6.1 Fundamentos

As simulações computacionais a partir da Dinâmica Molecular (DM) e do Método de Monte Carlo (MC) baseiam-se respectivamente nos princípios da mecânica clássica e da física estatística (44). A simulação ajuda a compreender as propriedades de conjuntos atômicos-moleculares representando sistemas físicos, químicos e biológicos formado por aproximadamente  $10^{23}$  partículas e não envolvem reações químicas ou transições entre diferentes estados eletrônicos no átomo (44). Estes sistemas são termodinâmicos e sua compreensão atômico molecular precisa de uma abordagem mecânica estatística com temperatura ambiente entre 100K e 800K. A ordem de grandeza da energia envolvida nos fenômenos de interesse são típicas de interações entre átomos, íons e moléculas (44).

A simulação computacional possui um número de partículas que depende da natureza do sistema, ou seja, cada proteína possui sua respectiva quantidade de aminoácidos compostos por diversos átomos. São tratados sistemas contendo entre  $10^3$  partículas e  $10^6$  partículas, e as condições de equilíbrio são obtidos em médias temporais sobre um grande intervalo de tempo na escala atômica aproximadamente entre  $10^{-11}$  segundos e  $10^{-8}$  segundos. Algumas propriedades de energia de interação e distribuição radial de pares convergem rapidamente, entretanto, a pressão, tensão superficial e os coeficientes de transporte requerem longas trajetórias de simulação para convergir (44).

O interessante da técnica de Dinâmica Molecular (MD) e de Monte Carlo (MC) é compatibilidade com os dados experimentais. Os resultados das investigações computacionais produzem novos conhecimentos do nível microscópico dos fenômenos observados em laboratório. Outra vantagem é a diversidade de sistemas que podem ser tratados por este método, sistemas homogêneos: gases, fluidos supercríticos, líquidos, soluções e misturas, sistemas pouco ou nada homogêneos como interfaces, filmes de Langmuir-Blodgett, biomembranas, polímeros orgânicos e inorgânicos, polissacarídeos, lipídios, proteínas, ácidos nucleicos e seus complexos, passando por sistemas como zeólitos, argilas, sólidos cristalinos

e vítreos, e nanomateriais, entre outros (44).

Os átomos que formam as moléculas permanecem unidos por forças associadas aos potenciais de interação. Com o conjunto desses potenciais defini-se um "Campo de Força", descrevendo propriedades dinâmicas e de equilíbrio de um sistema molecular. Por meio das equações de movimento de Newton calcula-se as propriedades do sistema, fornecendo informações a nível microscópico: posições e velocidades atômicas, a partir destes dados, então, a conexão a nível macroscópico é feita pela mecânica estatística, gerando as propriedades observáveis (pressão, temperatura, energia, entre outras) (44).

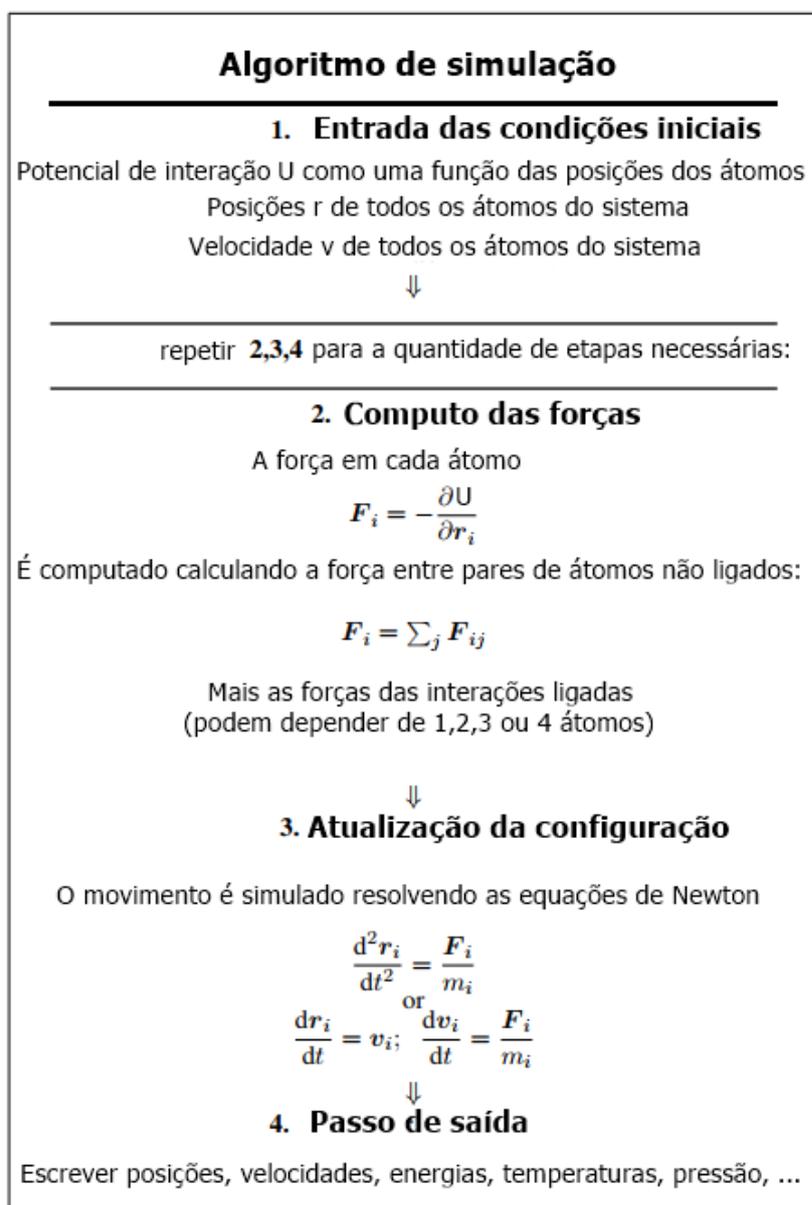


Figura 53 – Etapas do programa GROMACS. Figura adaptada da referência (46).

As posições dos átomos em uma dada proteína são determinadas por técnicas de medida estrutural e conformacional, como por exemplo a técnica de espalhamento de raios X a baixos ângulos (Small-angle X-ray scattering). Essa informação estrutural é catalogada num banco de dados sob forma de arquivos de texto. As informações estruturais desses arquivos de proteínas são os arquivos de entrada do GROMACS. Esse arquivo de dados é encontrado na base de dados Protein Data Bank (PDB), disponível no endereço eletrônico: <<http://www.rcsb.org>>. A estrutura foi montada com base em experiências de cristalografia de raios X, ressonância magnética e no espalhamento dinâmico de luz (DLS) (45). Existe um limite de validade para utilizar as forças de Newton clássicas, para retratar o comportamento atômico. Por meio de um "teste de validade", fundamentado no comprimento de onda térmico de Broglie ( $\Lambda$ ):

$$\Lambda = \left[ \frac{2\pi\hbar^2}{mk_B T} \right]^{1/2}, \quad (6.1)$$

onde  $m$  é a massa atômica,  $T$  a temperatura do sistema, e  $h$  é a constante de Planck ( $\hbar = h/2\pi$ ), é possível aplicar a dinâmica molecular quando  $\Lambda \ll a$ , sendo  $a$  a separação média entre os átomos vizinhos (44, 49). Em temperaturas suficientemente baixas, as determinações da dinâmica molecular falham (49). Por conta desse limite, nossas simulações foram realizadas com temperatura de 300 K.

## 6.2 Configuração inicial

Numa configuração inicial de partículas interagentes as posições e velocidades são dispostas em uma "caixa de simulação" e passam a movimentar-se sob ação de potenciais de interação intramoleculares e intermoleculares (44). Considerando um sistema físico constituído de  $N$  partículas (átomos ou moléculas) identificado por  $3N$  coordenadas de posição e  $3N$  coordenadas de velocidade (46). A configuração inicial é dada a partir destas  $6N$  variáveis iniciais das moléculas, posicionando-se em princípio em sítios pré estabelecidos.

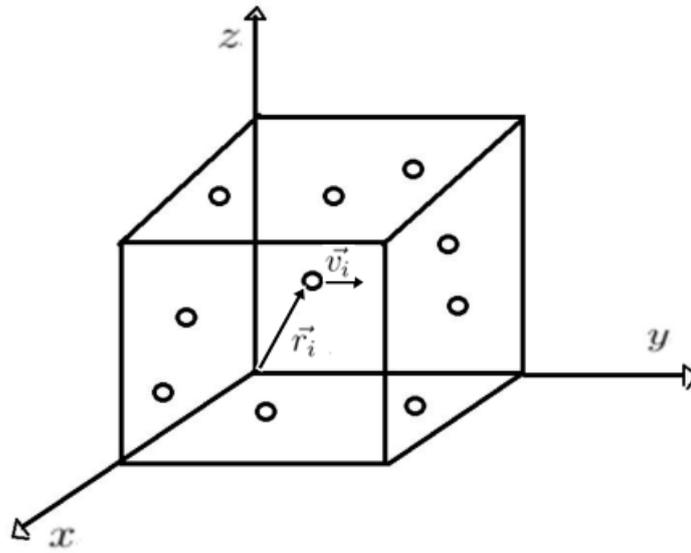


Figura 54 – Representação do sistema. Figura adaptada da referência (46).

$$\vec{r}_i = [x_i, y_i, z_i] \quad ; \quad \vec{v}_i = [v_{xi}, v_{yi}, v_{zi}]. \quad (6.2)$$

Adota-se a rede cúbica de face centrada (fcc) como a caixa de simulação de comprimento  $L$ . Assim, as partículas ocupam os vértices e o meio das faces de um cubo chamado de célula unitária, cada partícula no vértice é compartilhada por oito células unitárias, enquanto a partícula da face é compartilhada por duas células unitárias:

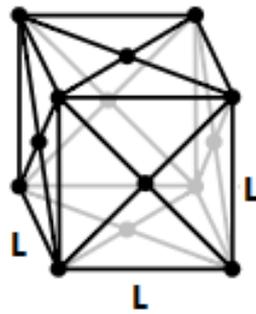


Figura 55 – Caixa de simulação (fcc). Figura adaptada da referência (46).

$$(1/8 \text{ átomos}).8 + (1/2 \text{ átomos}).6 = 4 \text{ átomos}. \quad (6.3)$$

Então, a densidade de partículas é  $\rho = \frac{N}{\Omega}$ . Sabendo que o número de partículas é quatro e  $\Omega$  é o volume dado pelo comprimento das arestas  $L$  elevado ao cubo, temos  $\rho = \frac{4}{L^3}$  (46). As posições das quatro partículas são :

$$(0, 0, 0); (0, 1/2, 1/2); (1/2, 0, 1/2); (1/2, 1/2, 0). \quad (6.4)$$

Para a construção da rede cristalina, as posições das demais partículas são obtidas pela translação das células unitárias ao longo das três direções  $(x, y, z)$  (46). Portanto, o número total de partículas na simulação é:

$$N_{\text{partículas}} = n_{pcu} \cdot \eta_x \cdot \eta_y \cdot \eta_z, \quad (6.5)$$

onde  $n_{pcu}$ ,  $\eta_x$ ,  $\eta_y$ ,  $\eta_z$ , são o número de partículas na célula unitária, e a quantidade de repetições das células nas três direções, respectivamente (46). A cada passo de integração das equações de movimento teremos que ter as coordenadas das partículas dentro da caixa, isto significa dizer:

$$0 \leq x_i \leq L_x, \quad 0 \leq y_i \leq L_y, \quad 0 \leq z_i \leq L_z. \quad (6.6)$$

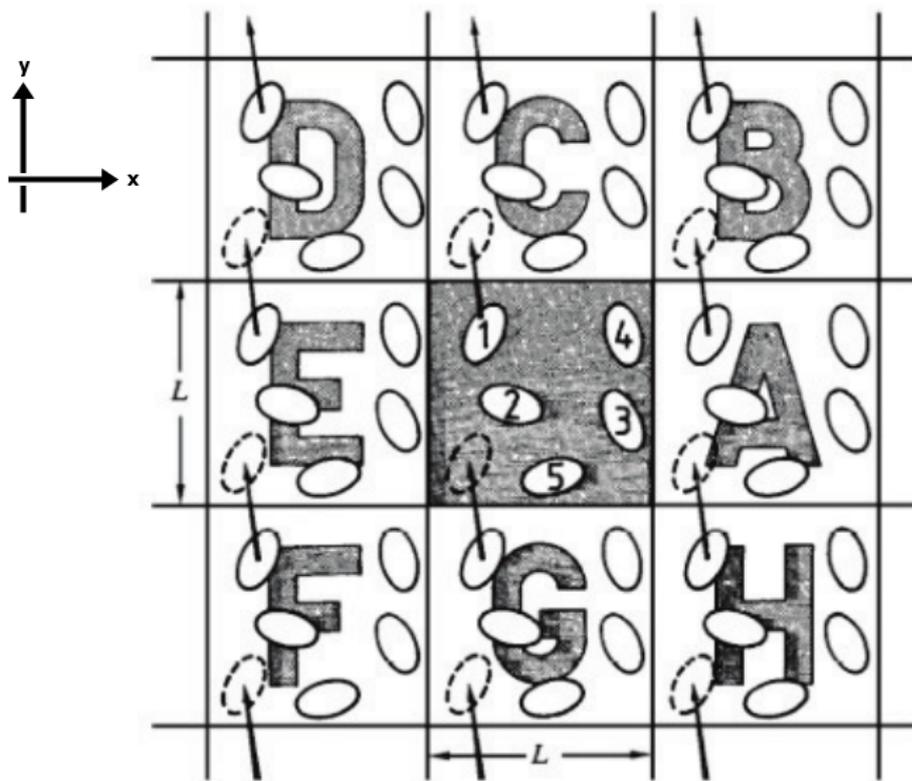


Figura 56 – Condições de contorno. Figura adaptada da referência (49).

Dado as posições e velocidades de todas as partículas em um instante  $t_0$ , pode-se obter as forças resultantes em cada partícula, devido às interações com as demais, e então determinar posições e velocidades em um instante posterior  $t_0 + \delta t$ , através das equações de movimentos Newtonianas. As novas posições são utilizadas para o cálculo das novas

forças e daí obter as posições e velocidades em  $t_0+2\delta t$ . Este procedimento é realizado recursivamente, gerando trajetórias moleculares para todo o sistema (44).

A escolha do passo de integração  $\delta t$  precisa ser feita de tal forma que a energia total do sistema seja uma constante durante todo o intervalo de tempo da simulação. Esta escolha é baseada em um conhecimento já estabelecido, característico de algum sistema físico. A escolha deve ser feita resultando em estabilidade do sistema. É importante ressaltar, como as equações de movimento são integradas via discretização em formas de série de Taylor, a energia total não será exatamente constante, no entanto, terá uma pequena flutuação a qual deve ser menor ou da ordem de  $10^{-4}$  durante todo o tempo de simulação (44).

### 6.3 Campo de força OPLS/AA

Na descrição das interações dos sistemas moleculares, assume-se dois tipos de potenciais ou campos de força, importantes para a descrição correta do sistema: intermoleculares e intramoleculares. Entre os mais conhecidos campos de força, destacam-se o Optimized Potentials for Liquid Simulation (OPLS), com métodos apropriados para proteínas, lipídios, sacarídeos e diversas outras moléculas orgânicas. Seus parâmetros foram divididos em átomos agrupados OPLS-UA (United-Atom), considerando, por exemplo o grupo metil ( $\text{CH}_3$ ) como sendo um único sítio. Agora o OPLS-ALL (All atom) descreve todos os átomos do sistema (44).

Em um modelo de potencial intermolecular que representa a interação entre os átomos  $i$  e  $j$ , consiste da soma das interações entre par de átomos de moléculas distintas descritas por um termo de Lennard-Jones acrescido ao termo correspondente às interações eletrostáticas (44):

$$U_{\text{inter-molecular}} = \sum_{i=1} \sum_{j<i} U_{\text{Couloumb}}(r_{ij}) + \sum_{i=1} \sum_{j<i} U_{\text{Lennard-Jones}}(r_{ij}). \quad (6.7)$$

O potencial de Coulomb é

$$U_{\text{Couloumb}}(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}, \quad (6.8)$$

onde  $q_i$  e  $q_j$  correspondem às cargas parciais localizadas nos seus respectivos sítios, e  $r_{ij}$  é a separação entre os átomos  $i$  e  $j$  de moléculas distintas.

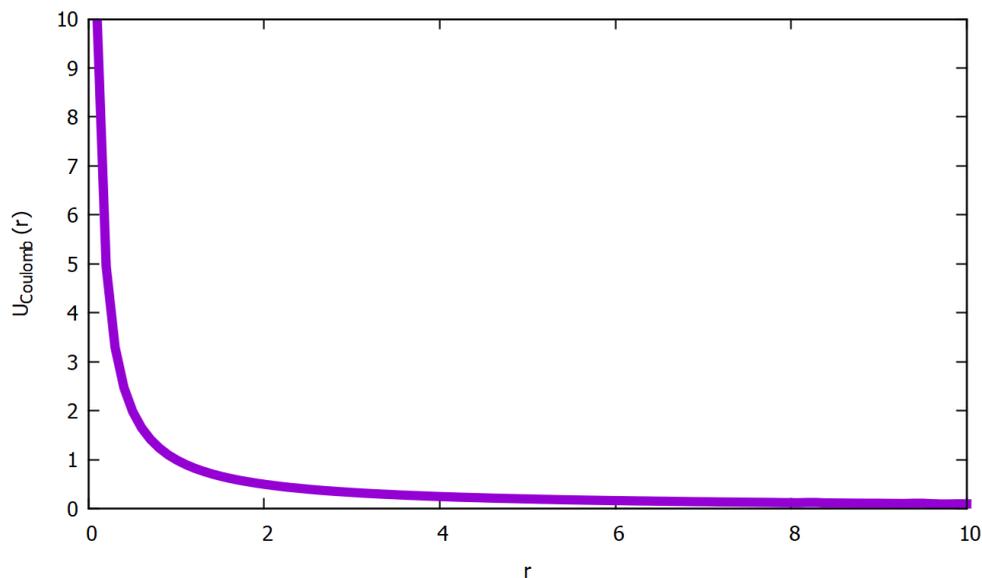


Figura 57 – Comportamento do potencial de Coulomb, plotado no Gnuplot.

O potencial de Lennard-Jones é

$$U_{\text{Lennard-Jones}}(r_{ij}) = 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]. \quad (6.9)$$

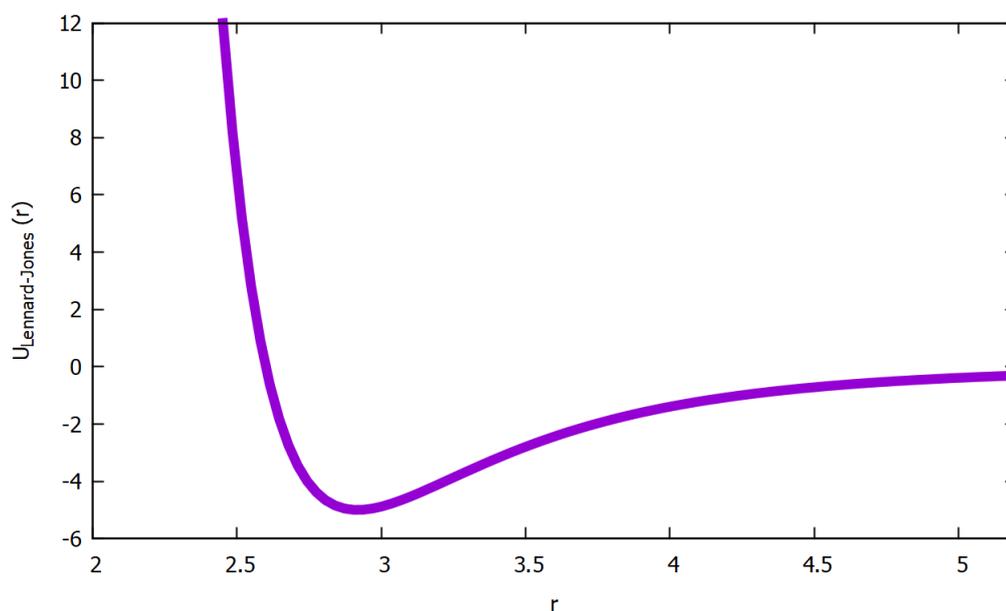


Figura 58 – Comportamento do potencial de Lennard-Jones, plotado no Gnuplot.

Para o potencial de Lennard-Jones, o mínimo de energia corresponde a  $\epsilon$  e o valor de sigma  $\sigma$  representa a distância entre os átomos em que o potencial de Lennard-Jones é nulo  $\Phi_{\text{Lennard-Jones}}(\sigma) = 0$ . Para calcular a interação entre os átomos de moléculas distintas é

necessário o uso de alguma regra de combinação para os coeficientes de Lennard-Jones. O campo de força OPLS adota o critério geométrico para ambos os parâmetros (44):

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \quad ; \quad \sigma_{ij} = \sqrt{\sigma_{ii}\sigma_{jj}}, \quad (6.10)$$

onde  $\epsilon_{ii}$ ,  $\epsilon_{jj}$ ,  $\sigma_{ii}$ ,  $\sigma_{jj}$  são os parâmetros de energia e diâmetro, respectivamente (44). As interações de caráter intramoleculares descrevem as deformações moleculares, e podem ser subdivididas nas seguintes contribuições (44):

$$U_{\text{intra-molecular}} = \sum_{\text{ligações}} U_{\text{ligação}} + \sum_{\text{ângulos}} U_{\text{vibração}} + \sum_{\text{diedros}} U_{\text{torção}}; \quad (6.11)$$

O potencial de ligação ( $U_{\text{ligação}}$ ) descreve estiramentos de uma ligação química, enquanto, o potencial de vibração ( $U_{\text{vibração}}$ ) e o potencial de torção ( $U_{\text{torção}}$ ) correspondem, respectivamente, às descrições das deformações angulares e das deformações dos ângulos diedros, obtidos por meio das seguintes variáveis:

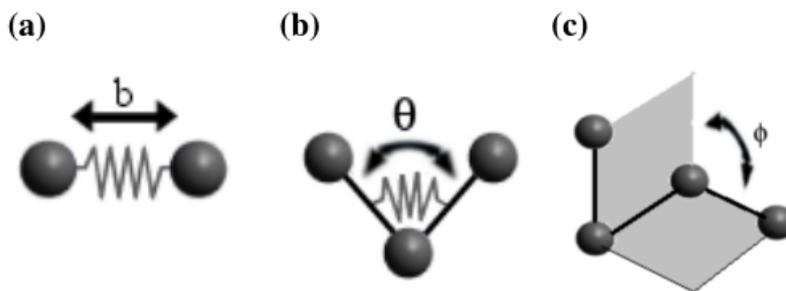


Figura 59 – Variáveis dos potenciais intra-moleculares. Figura adaptada da referência (50).

O termo de estiramento é descrito por um potencial harmônico. A distância entre os dois átomos é  $b$ , agora a distância de equilíbrio da ligação é  $b_0$  e a constante de força é  $K_b$ ; o potencial de ligação é (44).

$$U_{\text{ligação}}(b) = \frac{1}{2}K_b(b - b_0)^2. \quad (6.12)$$

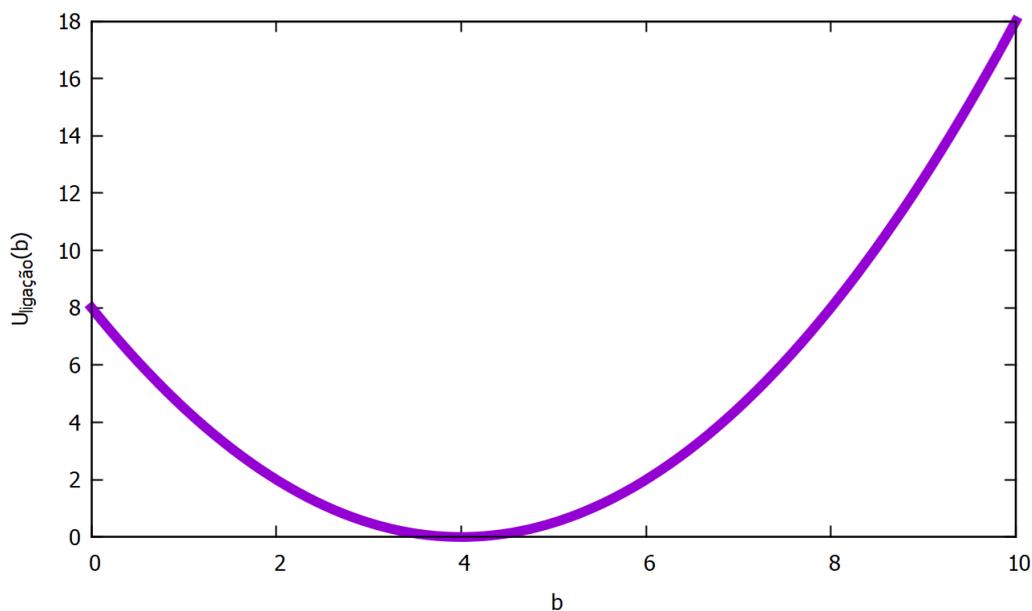


Figura 60 – Comportamento do potencial de ligação, plotado no Gnuplot.

O termo referente à deformação está relacionado à vibração, e também é um potencial harmônico: onde  $\theta$  é o ângulo entre três átomos ligados, e  $\theta_0$  o ângulo de equilíbrio da ligação, e  $K_\theta$  uma constante de força(44):

$$U_{\text{vibração}}(\theta) = \frac{1}{2}K_\theta(\theta - \theta_0)^2, \quad (6.13)$$

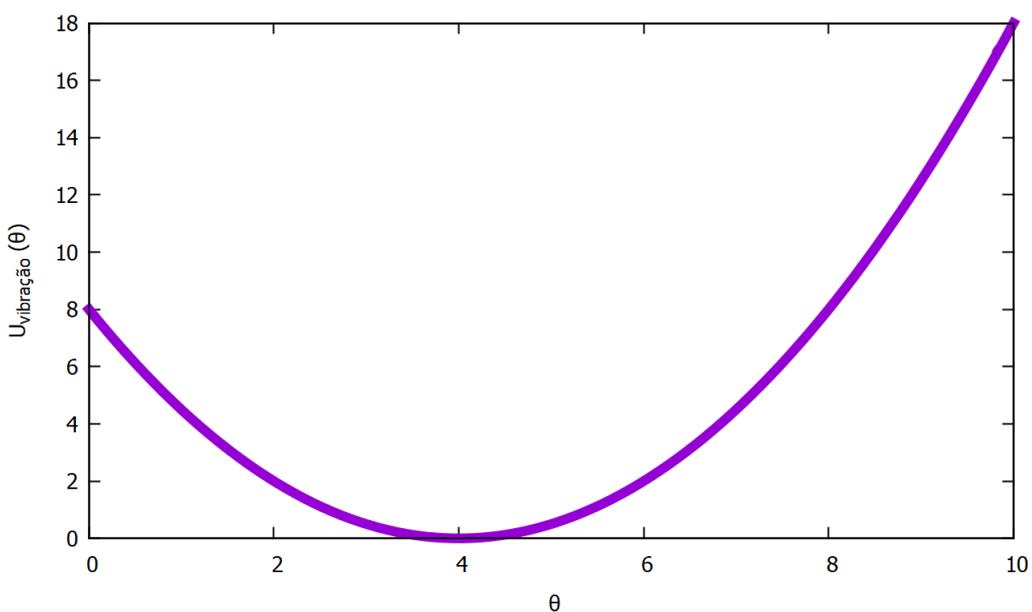


Figura 61 – Comportamento do potencial de vibração, plotado no Gnuplot.

O potencial de torção é contínuo em todo o intervalo  $[0, 2\pi]$ . Pode ser representado

por uma expansão em série de Fourier truncada. A expansão comumente empregada nos campos de forças é (44):

$$U_{\text{torção}}(\phi) = U_0 + \frac{U_1}{2}[1 + \cos(\phi)] + \frac{U_2}{2}[1 + \cos(2\phi)] + \frac{U_3}{2}[1 + \cos(3\phi)], \quad (6.14)$$

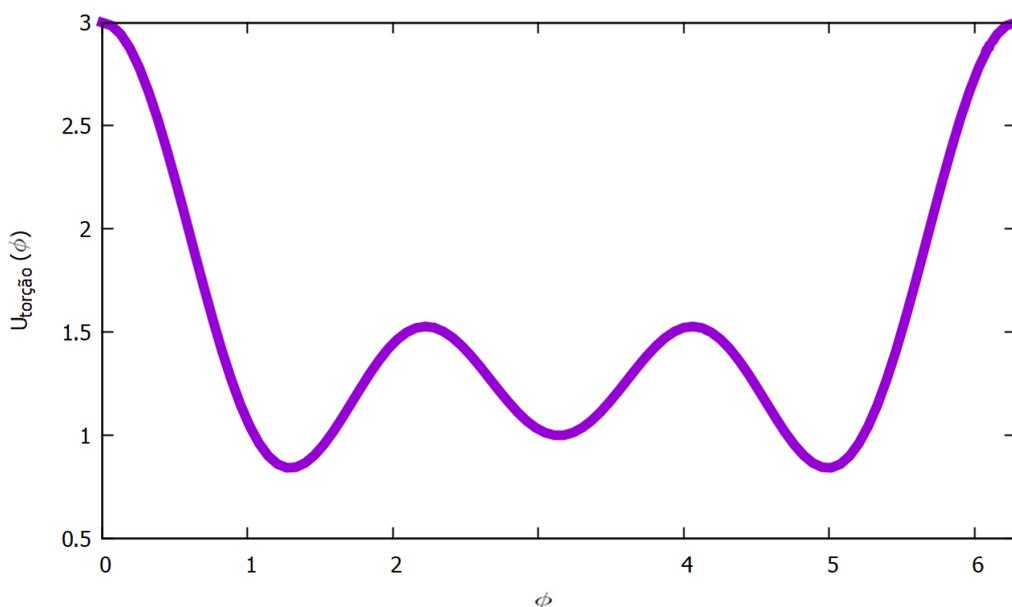


Figura 62 – Comportamento do potencial de torção, plotado no Gnuplot.

onde  $U_0, U_1, U_2, U_3$  são os coeficientes de Fourier da série. O valor do ângulo de torção ou diedral é  $\phi$ . Para descrever as interações em sistemas moleculares é descrita como:

$$U = U_{\text{inter-molecular}} + U_{\text{intra-molecular}} \quad (6.15)$$

O potencial entre duas moléculas é descrito como uma soma de interações entre cada par de átomos ou sítios de interação, e que estas interações dependem apenas da distância que separa estes sítios. Nos casos em que os movimentos internos das moléculas são importantes, adiciona-se a este potencial as interações intramoleculares, relacionadas à deformação da geometria molecular (44). Os potenciais intermoleculares e intramoleculares fornecem a descrição do estado do sistema (44).

## 6.4 Equações de movimento

Devido à complexidade da energia potencial ser função das posições de todas as partículas do sistema, não há solução analítica para o problema. Uma forma de contornar

este problema é por aproximações das posições e velocidades das partículas em expansão de Taylor da função  $x(t + \Delta t)$  em torno de  $t$  (51):

$$x(t + \Delta t) = x(t) + \frac{dx}{dt} \Delta t + \frac{d^2x}{dt^2} \frac{\Delta t^2}{2} + \frac{d^3x}{dt^3} \frac{\Delta t^3}{6} + \dots \quad (6.16)$$

Para gerar as trajetórias das partículas, é importante ressaltar que a configuração é decomposta em estágios separados em intervalos de tempo de integração  $\delta t$  (44). A força total de cada partícula é computada como uma soma vetorial de suas interações com outras partículas. Então, é obtida as acelerações na qual são combinadas com as posições e velocidades no tempo  $t$ , obtendo-as posteriormente em  $t + \delta t, t + 2\delta t, t + 3\delta t$  e assim por diante (44, 51):

$$\vec{r}_i(t + \delta t) = \vec{r}_i(t) + \vec{v}_i(t)\delta t + \frac{1}{2}\vec{a}_i(t)\delta t^2 + \frac{1}{6}\vec{b}_i(t)\delta t^3 + \dots \quad (6.17)$$

$$\vec{r}_i(t - \delta t) = \vec{r}_i(t) - \vec{v}_i(t)\delta t + \frac{1}{2}\vec{a}_i(t)\delta t^2 - \frac{1}{6}\vec{b}_i(t)\delta t^3 + \dots \quad (6.18)$$

Para encontrar as posições das partículas, fazemos a soma das equações (6.17) e (6.18). Temos (44, 51):

$$\vec{r}_i(t + \delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \delta t) + \vec{a}(t)\delta t^2 = 2\vec{r}_i(t) - \vec{r}_i(t - \delta t) + \frac{F_i(t)}{m_i}\delta t^2. \quad (6.19)$$

Agora para determinar a velocidades das partículas devemos subtrairmos as equações (6.17) e (6.18). Sendo assim (44, 51):

$$\vec{v}_i(t) = \frac{\vec{r}_i(t + \delta t) - \vec{r}_i(t - \delta t)}{2\delta t}. \quad (6.20)$$

A partir da equação (6.20), temos  $\vec{r}_i(t - \delta t) = \vec{r}_i(t + \delta t) - 2\vec{v}_i(t)\delta t$ . Logo, substituindo na equação (6.19), chegamos na equação da evolução da posição (52, 53):

$$\vec{r}_i(t + \delta t) = \vec{r}_i(t) + \vec{v}_i(t)\delta t + \frac{1}{2}\vec{a}_i(t)\delta t^2. \quad (6.21)$$

Para determinar a evolução da velocidade, temos  $t \rightarrow t + dt$  na equação (6.20):

$$\vec{v}_i(t) = \frac{\vec{r}_i(t + 2\delta t) - \vec{r}_i(t)}{2\delta t}. \quad (6.22)$$

Precisamos determinar  $\vec{r}_i(t + 2\delta t)$ . Implementando então  $t \rightarrow t + \delta t$  na equação (6.19). Teremos  $\vec{r}_i(t + 2\delta t) = 2\vec{r}_i(t + \delta t) - \vec{r}_i(t) + \vec{a}_i(t + \delta t)\delta t^2$ . Agora substituindo na equação (6.22):

$$\vec{v}_i(t + \delta t) = \frac{\vec{r}_i(t + \delta t) - \vec{r}_i(t)}{\delta t} + \frac{1}{2}\vec{a}_i(t + \delta t)\delta t. \quad (6.23)$$

Com base na evolução da posição, equação (6.21). Substituímos na equação (6.23), temos a equação da evolução temporal da velocidade (52, 53):

$$\vec{v}_i(t + \delta t) = \vec{v}_i(t) + \frac{1}{2}[\vec{a}_i(t + \delta t) + \vec{a}_i(t)]\delta t. \quad (6.24)$$

O valor de  $\delta t$  pode ser aproximado em um valor menor que a metade do tempo de colisão entre partículas e a determinação deste valor é por meio da conservação da energia (44). As vantagens são: 1) É possível utilizar tempo igual a zero, 2) a velocidade não obitda a partir da razão entre números pequenos, 3) posição e velocidade são avaliados no mesmo tempo. As desvantagens, é que exige o cálculo das acelerações duas vezes (52).

## 6.5 Dinâmica molecular do Potencial de Lennard-Jones

As simulações de dinâmica molecular clássicas se concentram nas equações de movimento de Newton para partículas (átomos) que compõem o sistema em estudo (46):

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N), \quad i = 1, 2, \dots, N \quad (6.25)$$

$\vec{r}_i$  é o vetor posição da partícula  $i$ , e as forças atuando sobre cada partícula  $i$  é  $\vec{F}_i$ . Nos casos em que o trabalho é independente da trajetória, as forças, ditas conservativas, derivam da energia potencial  $U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$ (46):

$$\vec{F}_i(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = -\nabla_{\vec{r}_i} U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N). \quad (6.26)$$

Esta representação acarreta na conservação da energia total, ou seja, a quantidade de energia cinética mais o potencial mantém-se  $E_{total} = E_{cinética} + U$ . O potencial pode ser apresentado, no caso mais simples, como uma soma de pares de interação (46):

$$U = \sum_{i=1}^N \sum_{i < j}^N U(r_{ij}), \quad (6.27)$$

$r_{ij}$  é distância entre os dois átomos da ligação, e a condição  $i < j$  é para evitar contagem dupla das energias  $U(\vec{r}_{ij}) = U(\vec{r}_{ji})$ . As forças atuando nas partículas do sistema são (46):

$$\vec{F}_i = \sum_{j \neq i}^N \vec{f}_{ij}, \quad \vec{f}_{ij} = -\frac{dU(r_{ij})}{dr_{ij}} \frac{\vec{r}_{ij}}{r_{ij}}, \quad (6.28)$$

$\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$  é o vetor posição ligando as partículas  $i$  e  $j$ . As funções energias potenciais ou potenciais de interação intra-moleculares e inter-moleculares constituem um passo para a descrição do sistema de interesse. Estes potenciais determinam as forças que atuam nos átomos e como será sua evolução no tempo, gerando trajetórias para análise. Por exemplo a partir do potencial de Lennard-Jones (46):

$$\phi(r_{ij}) = 4\epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right]. \quad (6.29)$$

Agora derivando o potencial em relação à coordenada radial para encontrar a força:

$$\vec{f}_{ij} = -\frac{\partial}{\partial r} \Phi(r_{ij}) \hat{r}_{ij} = -4\epsilon \left[ -12\sigma^{12} \left( \frac{1}{r_{ij}} \right)^{13} + 6\sigma^6 \left( \frac{1}{r_{ij}} \right)^7 \right] \hat{r}_{ij}, \quad (6.30)$$

$$\vec{f}_{ij} = 48\epsilon \left[ \sigma^{12} \left( \frac{1}{r_{ij}} \right)^{13} - \frac{1}{2}\sigma^6 \left( \frac{1}{r_{ij}} \right)^7 \right] \hat{r}_{ij}, \quad (6.31)$$

lembrando que o vetor unitário é escrito na forma  $\hat{r}_{ij} = \frac{\vec{r}_{ij}}{r_{ij}}$ , segue-se (54):

$$\vec{f}_{ij} = \frac{48\epsilon}{\sigma^2} \left[ \left( \frac{\sigma}{r_{ij}} \right)^{14} - \frac{1}{2} \left( \frac{\sigma}{r_{ij}} \right)^8 \right] \vec{r}_{ij}. \quad (6.32)$$

Para o argônio os parâmetros do potencial são :

$$m = 6,6 \cdot 10^{23} g; \quad \epsilon = 1,66 \cdot 10^{-14} erg; \quad \sigma = 3,4 \cdot 10^{-8} cm \quad (6.33)$$

Em simulação computacional com o potencial de Lennard-Jones, é conveniente trabalhar com grandezas adimensionais (46):

$$\vec{r}'_{ij} \rightarrow \frac{\vec{r}_{ij}}{\sigma}; \quad \Phi \rightarrow \frac{\phi}{\epsilon}; \quad t' \rightarrow t\sigma \left( \frac{m}{\epsilon} \right)^{1/2} \quad (6.34)$$

A partir da força em cada átomo dada pela equação 6.32, é possível calcular a força total sobre todos os átomos do sistema fazendo a soma sobre os índices  $i$  e  $j$ , com  $i \neq j$ .

$$\frac{d^2 \vec{r}'_i}{dt'^2} = \sum_i \sum_{i \neq j} \vec{f}'_{ij} = 48 \sum_i \sum_{i \neq j} \left[ \left( \frac{1}{r'_{ij}} \right)^{14} - \left( \frac{1}{2r'_{ij}} \right)^8 \right] \vec{r}'_{ij}, \quad (6.35)$$

Para poupar tempo computacional, o potencial de Lennard-Jones é truncado na distância limite de  $r_c = 2,5$ , já que as contribuições são pequenas refletindo em um potencial de curto alcance (46). As tabelas abaixo mostram o esquema para o cálculo de forças. A parte mais trabalhosa para o programa de dinâmica molecular é o cômputo da

forças e por isso é a parte que mais demanda tempo computacional, para minimizar esse efeito é a criação de tabelas ou matrizes (46):

i/k	1	2	3	...	N-1	N
1	(1,1)	(1,2)	(1,3)	...	(1,N-1)	(1,N)
2	(2,1)	(2,2)	(2,3)	...	(2,N-1)	(2,N)
3	(3,1)	(3,2)	(3,3)	...	(3,N-1)	(3,N)
...	...	...	...	...	...	...
N-1	(N-1,1)	(N-1,2)	(N-1,3)	...	(N-1,N-1)	(N-1,N)
N	(N,1)	(N,2)	(N,3)	...	(N,N-1)	(N,N)

Tabela 6 – Matriz com todos os pares de interações. Figura adaptada da referência (46)

Por meio da terceira lei de Newton, a força que a partícula  $i$  exerce na partícula  $k$  é igual em intensidade e direção, mas de sentido contrário, à força que  $k$  alica em  $i$ , logo, será necessário calcular somente uma parte, reduzindo o esforço da simulação (46):

i/k	1	2	3	...	N-1	N
1	(1,1)	(1,2)	(1,3)	...	(1,N-1)	(1,N)
2	-(1,2)	(2,2)	(2,3)	...	(2,N-1)	(2,N)
3	-(1,3)	-(2,3)	(3,3)	...	(3,N-1)	(3,N)
...	...	...	...	...	...	...
4	-(1,N-1)	-(2,N-1)	-(3,N-1)	...	(N-1,N-1)	(N-1,N)
5	-(1,N)	-(2,N)	-(3,N)	...	-(N-1,N)	(N,N)

Tabela 7 – Matriz com todos os pares de interações. Figura adaptada da referência (46)

Portanto, com isso é feito uma otimização, resultando na redução do esforço computacional, pois somente o triângulo superior será calculado. Com os valores da força em cada partícula pode-se calcular a evolução das posições e velocidades do sistema estudado (46).

## 6.6 Ensembles

A identificação do valor dos parâmetros, como por exemplo o número de partículas, volume, e energia, definem o macroestado do sistema. No âmbito molecular, existe um grande número de possibilidades que o macroestado  $(N, V, E)$  pode ser obtido. J. Willard Gibbs (1839 – 1903) definiu o ensemble como o conjunto de cópias do sistema, cada uma representando uma possível configuração que o sistema pode estar com os mesmos estados macroscópicos, i.e., energia  $E$ , pressão  $P$  e temperatura  $T$ , mas com diferentes valores de estados microscópicos (55, 56). Resumindo, podemos também dizer que o ensemble é a

distribuição de probabilidade do sistema (55, 56). Para determinarmos as propriedades mensuráveis de um fenômeno, corpo ou substância, devemos computar o valor médio ou valor esperado de uma certa grandeza física  $f$  medida em laboratório:

$$\langle f \rangle_{lab} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} f(t) dt. \quad (6.36)$$

É uma média temporal sobre um tempo grande  $\tau$  (55, 46). Para sistemas em equilíbrio termodinâmico (equilíbrio térmico, equilíbrio mecânico, equilíbrio radioativo e equilíbrio químico) pode-se usar a hipótese ergódica proposta por Boltzmann supondo que a média temporal das funções termodinâmicas são substituídas por médias de ensemble (44, 55). Seja então uma variável aleatória  $x$  assumindo valores entre  $a$  e  $b$ . O termo da diferencial  $\rho(x)dx$  pode ser entendida como a probabilidade de que a variável  $x$  esteja entre valores  $x$  e  $x + dx$ :

$$\int_a^b \rho(x) dx = 1. \quad (6.37)$$

A função  $f(x)$  referente a um conjunto variáveis aleatórias representando a evolução de um sistema com valor esperado dado a seguir:

$$\langle f(x) \rangle = \int_a^b f(x) \rho(x) dx, \quad (6.38)$$

onde a densidade probabilidade  $\rho(x)$  é conhecida por distribuição de probabilidades (55, 46). Em um sistema físico, parâmetros macroscópicos são mantidos constantes, como por exemplo o número de partículas  $N$ , a pressão  $P$ , o volume  $V$ , a temperatura  $T$  ou o potencial químico da substância  $\mu$  (55, 44). O conjunto destas grandezas físicas constantes numa amostra são conhecidos como ensemble microcanônico (NVE): o número de partículas, volume e energia são constantes (55, 44). No ensemble canônico (NVT): o número de partículas, volume e temperatura são constantes (55, 44). No ensemble das pressões (NPT): o número de partículas, temperatura e pressão são constantes (55, 44). No ensemble microcanônico (NVE) o número de partículas, volume e energia são constantes (55, 44). Conforme explicado no "Apêndice B", o conjunto dos microestados do sistema é representado por  $\Omega$  ocupando regiões do espaço de fase (55, 44). A equação de estado é uma função que conecta possíveis variáveis do sistema com propriedades termodinâmicas, e podem ser definidos através de potenciais termodinâmicos (energia livre disponível para realizar trabalho), por exemplo: energia interna  $U=U(S,V,N)$ , energia interna magnética  $U=U(S,M,N)$ , energia livre de Helmholtz  $F=F(T,V,N)$ , energia livre de Helmholtz magnética  $F=F(T,M,N)$ , energia livre de Gibbs  $G=G(T,P,N)$ , energia livre de Gibbs magnética  $G=G(T,H,N)$  (57, 55). A conexão das equações de estado (parâmetros) com a física estatística é feita no limite termodinâmico por meio da função de partição do sistema (55).

## 6.7 Método de Monte Carlo

O método de Monte Carlo (MC) está sendo amplamente utilizado nas simulações computacionais para o cálculo de propriedades termodinâmicas do sistema físico, possibilitando obter o valor esperado das média dos ensembles de uma grandeza física, fazendo médias aritméticas sobre as configurações selecionadas, com a vantagem de redução do tempo de simulação. Mais detalhadamente, isso se dá por meio da amostragem seletiva (importance sampling), este método consiste em amostrar o espaço de configurações baseado na distribuição de Boltzmann. Considere que temos a função integranda  $f(x)$ , sobre o intervalo de integração:

$$I = \int_a^b f(x)dx = (b - a) \langle f(x) \rangle_{[a,b]}. \quad (6.39)$$

O valor esperado em um dado intervalo  $[a, b]$  é (46, 55):

$$\langle f(x) \rangle = \int_a^b p(x)f(x)dx, \quad 1 = \int_a^b p(x)dx, \quad (6.40)$$

$p(x)$  é uma distribuição de probabilidade arbitrária, portanto a área sob a curva é chamada de quadratura, e pode ser estimada fazendo médias sobre amostras aleatórias de pontos no intervalo  $[x, y]$  com  $x \in [a, b]$  e  $y \in [f(a), f(b)]$ . O valor médio  $\langle f(x) \rangle$  é aferido como média aritmética (46, 55),

$$\langle f(x) \rangle_{[a,b]} \approx \frac{(b - a)}{N} \sum_{i=1}^N f(x_i)p(x_i). \quad (6.41)$$

No caso de amostrarmos a função  $f(x)$  de acordo com a distribuição de probabilidade  $p(x)$ , seu valor esperado se torna uma média aritmética dos valores amostrados (46, 55):

$$\langle f(x) \rangle_{[a,b]} \approx \frac{1}{N} \sum_{i=1}^N f(x_i), \quad (6.42)$$

N configurações geradas no equilíbrio  $\{x\} = x_1, x_2, x_3, \dots, x_N$ . O Hamiltoniano  $H$  é um função destas coordenadas  $H(\{x\})$  (46, 55). Então,

$$\langle f(x) \rangle = \frac{\int f(x)e^{-\beta H(\{x\})}dx}{\int e^{-\beta H(\{x\})}dx} = \frac{1}{N} \sum_{i=1}^N f(x_i). \quad (6.43)$$

Em física estatística, amostragem seletiva quer dizer escolher configurações do sistema que satisfaçam à distribuição de probabilidade de Gibbs, a qual é proporcional ao fator de Boltzmann  $e^{-H(\{x\})/k_b T}$ . Isto significa dizer, a probabilidade que um sistema

esteja em equilíbrio à temperatura  $T$ , com energia  $H(\{x\})$ . O resultado é válido para funções de mais uma variável (46, 55).

## 6.8 Ensemble NVE

No ensemble microcânico ou NVE, o número de partículas ( $N$ ), o volume ( $V$ ) e a energia ( $E$ ) são mantidos constantes no sistema, correspondendo a um processo adiabático, ou seja, sem trocas de calor. As trajetórias da dinâmica molecular microcanônica é vista como as trocas de energia cinética e potencial, mantendo a energia total constante (44). Com isso podemos escrever as equações de Newton na forma:

$$\vec{F}_i = -\vec{\nabla}_{\vec{r}_i} U(r_1, r_2, \dots, r_N) = m_i \frac{d\vec{v}_i}{dt} \quad (6.44)$$

$$\vec{v}_i = \frac{d\vec{x}_i}{dt}. \quad (6.45)$$

A força  $\vec{F}_i$  que atua em cada partícula do sistema pode ser calculada como o gradiente negativo de  $U(r)$ . Para cada passo de simulação  $\delta t$ , cada posição e velocidade da partícula pode ser determinado por um método simplético, como por exemplo, o algoritmo de Verlet. A evolução temporal da equação da posição representa uma trajetória. Dadas as posições iniciais (ex.: conhecimento teórico) e as velocidades (ex.: aleatórias Gaussianas) ou nulas (deixar o sistema evoluir), pode-se calcular todas as posições e velocidade futuras (58).

## 6.9 Ensemble NVT

No ensemble canônico ou NVT, o número de partículas  $N$ , o volume  $V$  e a temperatura  $T$  são mantidos constantes. Esta é a etapa de termalização, é a fase de ajustes para que corresponda ao estado termodinâmico de interesse. Neste ensemble, a energia total não é conservada (44). Há um banho térmico de energia térmico capaz de fornecer ou retirar energia cinética das moléculas do sistema. A temperatura pode permanecer inalterável por vários métodos, o mais simples é pelo "reescalonamento de velocidades". O método é multiplicar as velocidades atômicas por um fator  $\gamma$  (44):

$$\gamma = \sqrt{T^{desejada}/T(t)}, \quad (6.46)$$

a cada passo de simulação ou em intervalos periódicos, de maneira que a temperatura  $T$  é aproximada da temperatura desejada  $T^{desejada}$  (44). É possível encontrar a função temperatura fazendo uso do ensemble canônico, calculando o valor esperado da energia

cinética do sistema com  $N$  partículas (46). Pode-se definir a densidade de probabilidade por meio da distribuição canônica de Gibbs (46, 55):

$$\rho(p, q) = \frac{1}{Z} \exp\{-\beta H(q, p)\} \quad ; \quad Z = \int \exp\{-\beta H(q, p)\} dq dp, \quad (6.47)$$

a função de partição  $Z$  é o fator de normalização. É importante ressaltar que o hamiltoniano  $H$  do sistema é igual a soma da energia cinética  $E_{cinética}$  e da energia potencial  $U$ , portanto  $H = E_{cinética} + U$  (46, 55). A energia cinética de uma partícula é  $E_{cinética}$ :

$$E_{cinética} = \frac{1}{2} m \vec{v}^2 = \frac{\vec{p}^2}{2m}. \quad (6.48)$$

Para calcular o valor esperado da energia sem nenhum potencial de interação, o momento linear é vetorialmente escrito no espaço tridimensional como:

$$\vec{p} = m\vec{v} = m(v_x \hat{i} + v_y \hat{j} + v_z \hat{k}). \quad (6.49)$$

O valor esperado é:

$$\langle f(x) \rangle = \int_a^b f(x) \rho(x) dx. \quad (6.50)$$

A função de interesse é a energia cinética  $f = E_{cinética}$ :

$$\langle E_{cinética} \rangle = \frac{\int E_{cinética} \exp(-\beta E_{cinética}) d\vec{q} d\vec{p}}{\int \exp(-\beta E_{cinética}) d\vec{q} d\vec{p}}. \quad (6.51)$$

Como não temos o Hamiltoniano  $H$  em função das coordenadas da posição  $\vec{q}$ , as integrais anulam-se. Portanto, há três integrações iguais relacionadas às três direções do espaço tridimensional  $(x, y, z)$ . E como o valor de cada uma delas é igual, neste caso o resultado é o mesmo de multiplicar por três o valor da integral escolhido, por exemplo, na direção  $x$  (46, 55):

$$\langle E_{cinética} \rangle = 3 \frac{\int \left( \frac{\vec{p}^2}{2m} \right) \exp \left[ -\beta \left( \frac{\vec{p}^2}{2m} \right) \right] dp_x}{\int \exp \left[ -\beta \left( \frac{\vec{p}^2}{2m} \right) \right] dp_x}, \quad \beta = \frac{1}{k_B T} \quad (6.52)$$

logo,

$$\langle E_{cinética} \rangle = \frac{3}{2} k_B T, \quad (6.53)$$

pois, as integrais valem:

$$\int_{-\infty}^{\infty} e^{-[\frac{\beta}{2m}]p^2} dx = \sqrt{\frac{\pi}{[\frac{\beta}{2m}]}} \quad ; \quad \int_{-\infty}^{\infty} p^2 e^{-[\frac{\beta}{2m}]p^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{[\frac{\beta}{2m}]^3}}. \quad (6.54)$$

Por meio da teoria de Monte Carlo (MC), a energia cinética é (46, 55):

$$\langle E_{cinética} \rangle = \frac{1}{N} \sum_i E_{cinética_i}. \quad (6.55)$$

Logo, reescrevendo a temperatura na seguinte maneira:

$$\frac{1}{2} \sum_{i=1}^N m_i \vec{v}_i^2 = \frac{3}{2} N k_B T, \quad (6.56)$$

obtém-se (44):

$$T = \frac{1}{3Nk_B} \sum_{i=1}^N m_i \vec{v}_i^2, \quad (6.57)$$

onde  $k_b$  é a constante de Boltzmann, e  $N$  é o número de partículas, e  $m_i$  e  $v_i$  são respectivamente a massa e a velocidade da partícula  $i$  (44). Conclui-se então, que a temperatura do sistema está associado a velocidades das partículas (44). E a soma vetorial das velocidades deve ser um vetor nulo, evitando assim o deslocamento da caixa (44).

## 6.10 Ensemble NPT

No ensemble das pressões ou NPT, o número de partículas  $N$ , a pressão  $P$  e a temperatura  $T$  são mantidas constantes. A energia total não é conservada por causa da troca de calor com o reservatório térmico e há a realização de trabalho por pistão externo (ambos fictícios) (44). A temperatura é mantida pelo reescalonamento das velocidades, e em relação à pressão o método é bem parecido, o processo é multiplicar às dimensões da caixa de simulação por um fator  $\gamma$ :

$$\gamma = \sqrt[3]{P(t)/P^{desejada}}, \quad (6.58)$$

a cada passo de simulação ou em intervalos periódicos, de maneira que a pressão  $P(t)$  aproxime-se da pressão desejada  $P^{desejada}$ , e alcançando-a, mantenha-se com este valor durante toda a simulação. A dedução da pressão do sistema em um determinado instante de tempo pode ser acompanhada no "Apêndice C", e é determinada por meio do teorema do virial (59, 60):

$$P(t) = \frac{Nk_bT}{V} + \frac{1}{3V} \left\langle \sum_{i=1}^N \sum_{j<i}^N \vec{r}_{ij}(t) \vec{F}_{ij}(t) \right\rangle. \quad (6.59)$$

Para manter este ensemble é necessário o uso de dois acoplamentos temperatura e pressão. Sendo assim, a pressão está interligada à temperatura, resultando em uma acumulação de erro ou incerteza nesta propriedade (44). Alterando o volume da caixa de simulação, alteram-se as distâncias, e conseqüentemente as forças, mantendo o controle sobre a pressão exercida (44).

## 6.11 Raio de giro

A determinação do raio de giro é baseado na análise do momento de inércia  $I$ , representando a inércia de rotação, ou seja, quanto maior for seu valor mais difícil será girá-lo ou alterar sua rotação (53).

$$I = \int r^2 dm. \quad (6.60)$$

Em simulação computacional a integral é aproximada por uma soma (53):

$$I = \sum_i r_i^2 m_i. \quad (6.61)$$

Considerando a massa da proteína concentrada a uma distância  $R_g$  (53):

$$I = R_g^2 \sum_i m_i = R_g^2 M \rightarrow R_g = \sqrt{\frac{I}{M}}. \quad (6.62)$$

O raio de giro ou raio de giração é escrito como (53):

$$R_g = \sqrt{\frac{\sum_i r_i^2 m_i}{\sum_i m_i}}. \quad (6.63)$$

Com sua componente na direção x (53):

$$R_{g,x} = \sqrt{\frac{\sum_i (r_{i,y}^2 + r_{i,z}^2) m_i}{m_i}}. \quad (6.64)$$

O cálculo do raio de giro ( $R_g$ ) nos fornece uma informação aproximada do raio de massa ( $R_m$ ) da proteína em estudo.

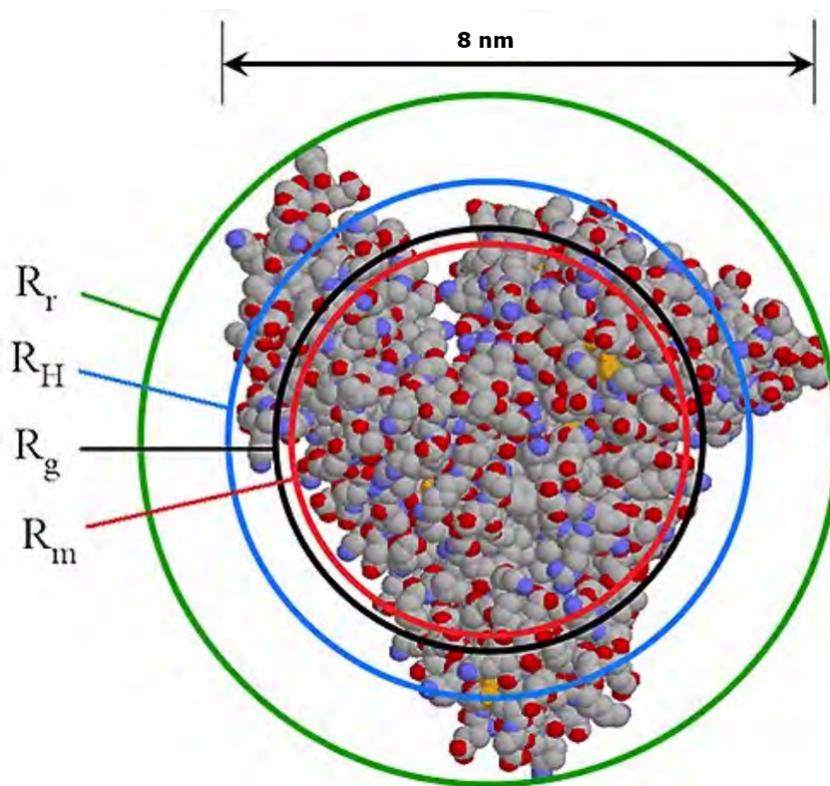


Figura 63 – Raio de giro da proteína. Figura adaptada da referência (62).

Na figura 63, algumas medidas descritivas de diferentes proteínas são observadas: raio de rotação ( $R_r$ ), raio hidrodinâmico ( $R_H$ ), raio de giro ( $R_g$ ) e o raio de massa ( $R_m$ ) (62).

## Capítulo 7

# Resultados das simulações

A simulação no GROMACS está descrita de forma detalhada nos "Apêncies D e E". Os gráficos abaixo refletem os resultados da simulação na PME do tomate para uma temperatura constante de 300 K. As pressões assumem diferentes valores a partir de 1 bar a 10 kbar, o tempo de atuação da pressão de até 100 ps.

A figura 64 mostra o gráfico da evolução do raio de giro da PME do tomate em função do tempo para uma pressão aplicada de 1 bar.

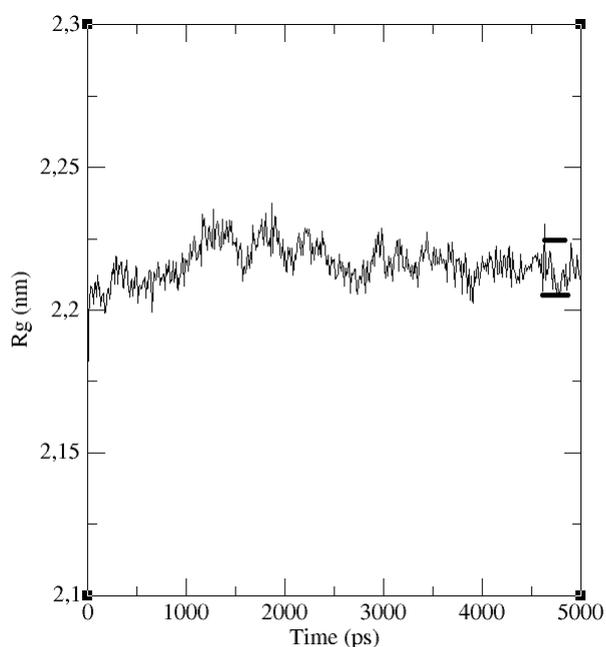


Figura 64 – Raio de giro da PME com pressão de 1 bar.

A região de tempos maiores que 4500 ps representa uma região onde a simulação computacional apresenta respostas envolvendo tempos maiores que os envolvidos em processos atômicos (da ordem de 100 ps). Portanto, analisando graficamente a envoltória que limita as oscilações das simulações, podemos definir um intervalo gráfico de confiança

associado ao valor mais provável do raio de giro. No gráfico sinalizamos essa região para auxiliar o leitor. Considerando essa envoltória, definimos graficamente o valor médio e a incerteza associada a essa simulação, a saber: 2.208(8) nm.

A figura 65 mostra o gráfico da evolução do raio de giro da PME do tomate em função do tempo para uma pressão aplicada de 1 kbar.

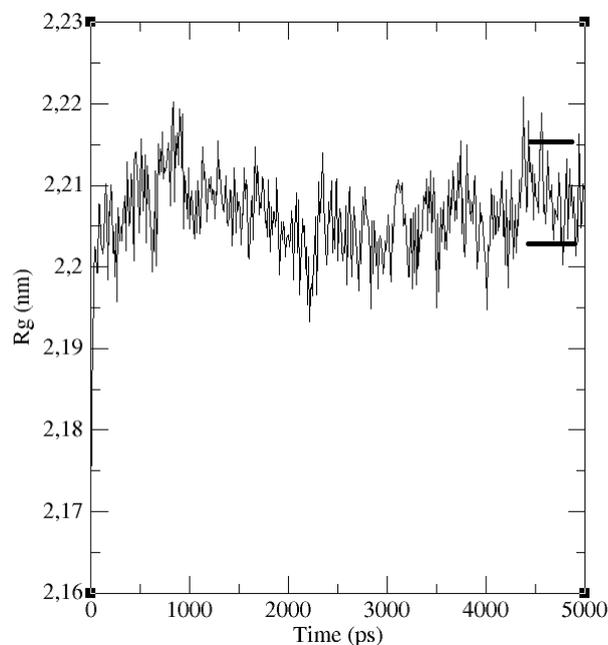


Figura 65 – Raio de giro da PME com pressão de 1 kbar.

Novamente, neste caso, a região de tempos maiores que 4500 ps representa uma região onde a simulação computacional apresenta respostas envolvendo tempos maiores que os envolvidos em processos atômicos (da ordem de 100ps). Analisando graficamente a envoltória que limita as oscilações das simulações, podemos definir um intervalo gráfico de confiança associado ao valor mais provável do raio de giro. No gráfico sinalizamos essa região para auxiliar o leitor. Considerando essa envoltória, definimos graficamente o valor médio e a incerteza associada a essa simulação, a saber: 2.210(5) nm.

A figura 66 mostra o gráfico da evolução do raio de giro da PME do tomate em função do tempo para uma pressão aplicada de 3 kbar.

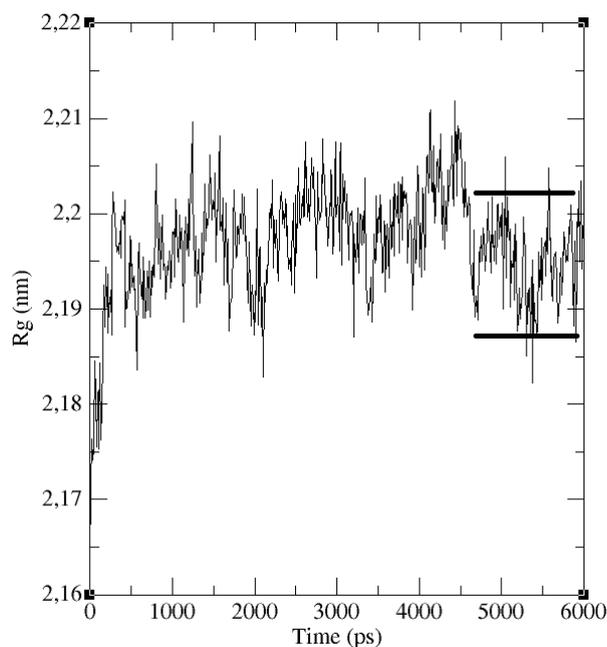


Figura 66 – Raio de giro da PME com pressão de 3 kbar.

Como já descrito, a região de tempos maiores que 4500 ps representa a região de interesse. Portanto, analisando graficamente a envoltória que limita as oscilações das simulações, podemos definir um intervalo gráfico de confiança associado ao valor mais provável do raio de giro. No gráfico sinalizamos essa região para auxiliar o leitor. Considerando essa envoltória, definimos graficamente o valor médio e a incerteza associada a essa simulação, a saber: 2.195(8) nm.

A figura 67 mostra o gráfico da evolução do raio de giro da PME do tomate em função do tempo para uma pressão aplicada de 5 kbar.

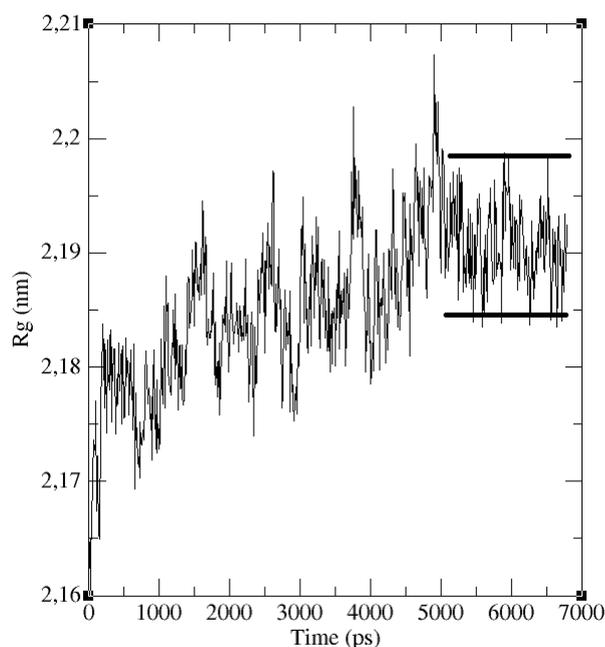


Figura 67 – Raio de giro da PME com pressão de 5 kbar.

Analisando graficamente a envoltória que limita as oscilações das simulações (para tempos maiores que 4500 ps), podemos definir um intervalo gráfico de confiança associado ao valor mais provável do raio de giro. Considerando essa envoltória, definimos graficamente o valor médio e a incerteza associada a essa simulação, a saber: 2.190(8) nm.

A figura 68 mostra o gráfico da evolução do raio de giro da PME do tomate em função do tempo para uma pressão aplicada de 7 kbar.

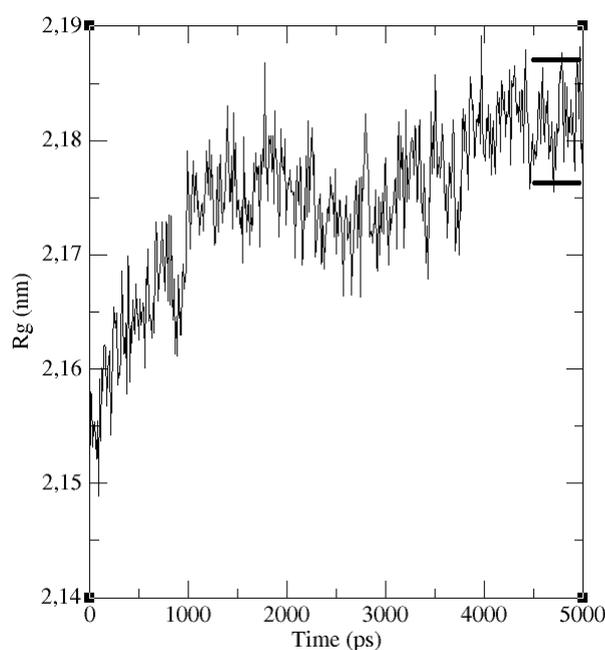


Figura 68 – Raio de giro da PME com pressão de 7 kbar.

A região de interesse é a de tempos maiores que 4500 ps. Analisando graficamente a envoltória que limita as oscilações das simulações, podemos definir um intervalo gráfico de confiança associado ao valor mais provável do raio de giro. No gráfico sinalizamos essa região para auxiliar o leitor. Considerando essa envoltória, definimos graficamente o valor médio e a incerteza associada a essa simulação, a saber: 2.182(6) nm.

A figura 69 mostra o gráfico da evolução do raio de giro da PME do tomate em função do tempo para uma pressão aplicada de 9 kbar.

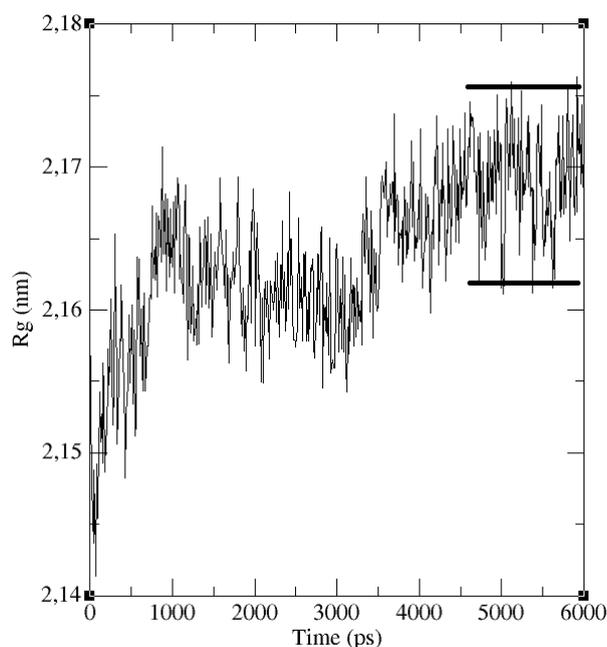


Figura 69 – Raio de giro da PME com pressão de 9 kbar.

A região de interesse é a de tempos maiores que 4500 ps. Analisando graficamente a envoltória que limita as oscilações das simulações, podemos definir um intervalo gráfico de confiança associado ao valor mais provável do raio de giro. No gráfico sinalizamos essa região para auxiliar o leitor. Considerando essa envoltória, definimos graficamente o valor médio e a incerteza associada a essa simulação, a saber: 2.171(6) nm.

A figura 70 mostra o gráfico da evolução do raio de giro da PME do tomate em função do tempo para uma pressão aplicada de 10 kbar.

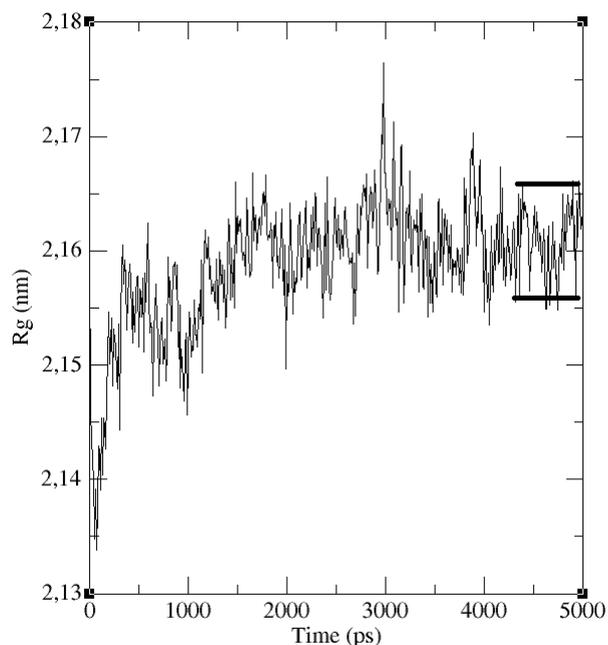


Figura 70 – Raio de giro da PME com pressão de 10 kbar.

A região de interesse é a de tempos maiores que 4500 ps. Analisando graficamente a envoltória que limita as oscilações das simulações, podemos definir um intervalo gráfico de confiança associado ao valor mais provável do raio de giro. No gráfico sinalizamos essa região para auxiliar o leitor. Considerando essa envoltória, definimos graficamente o valor médio e a incerteza associada a essa simulação, a saber: 2.161(5) nm.

Como comentario geral em relação às pressões em proteínas podemos dizer que elas apresentam diferentes níveis de resistência. Em geral, pressões relativamente baixas em torno 1 kbar-2 kbar possibilitam estimular a ativação de algumas enzimas, enquanto em pressões maiores em torno de 4 kbar-10 kbar tem como característica inativação (47). A perda e inativação da atividade enzimática podem estar associadas aos grupos: 1) completa e irreversivelmente inativada; 2) completa e reversivelmente inativada; (3) incompleta e irreversivelmente inativada ; (4) incompleta e reversivelmente inativada (47). No processamento a alta pressão (HHP), o alimento é armazenado em uma resistente embalagem plástica a qual não se rompe e é lacrada a vácuo (48). O alimento embalado é disposto dentro de um vaso de pressão preenchido por um fluido permitindo a aplicação da pressão que pode variar em torno de 1 kbar-10 kbar. Em geral, os equipamentos de HPP utilizam como fluido de trabalho a água H<sub>2</sub>O (48). A pressurização leva poucos minutos e causa pouco aquecimento térmico no alimento. Terminado o tratamento é de imediato despressurizado e o alimento pode ser removido do equipamento (48).

## Capítulo 8

## Conclusão

A partir dos resultados do raio de giro da Pectina Metilesterase (PME) do tomate gerados pelo GROMACS para as pressões: 1 bar, 1 kbar, 3 kbar, 5kbar, 7 kbar, 9kbar e 10 kbar foi realizada uma avaliação gráfica da incerteza e do valor médio entre as barras no tempo de simulação  $t > 4500$  ps.

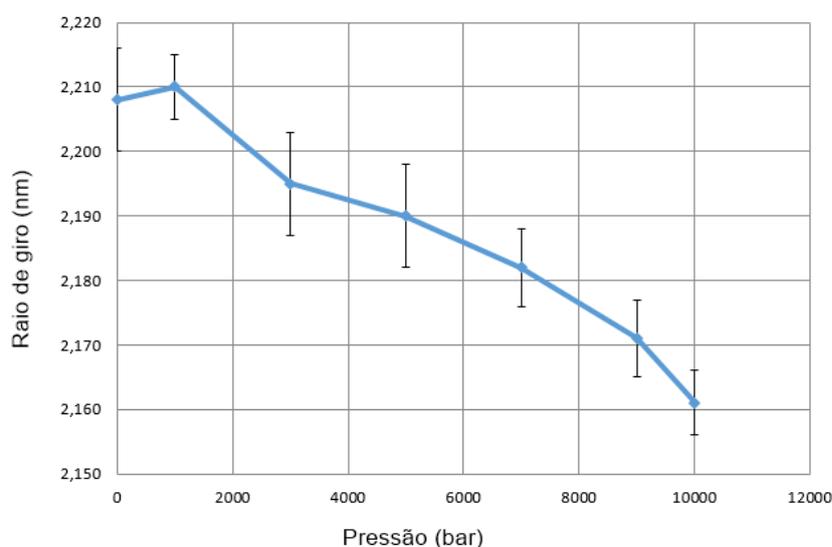


Figura 71 – Evolução do raio de giro da PME de acordo com os aumentos da pressão.

Verificamos na figura 71 uma tendência de redução do raio de giro da PME do tomate com o aumento da pressão. Como esse é um resultado oriundo da simulação computacional, considerando o campo de força escolhido, faz-se necessário averiguar se esse comportamento se verifica experimentalmente. Não encontramos na literatura resultados associados à PME do tomate sob pressão. Em uma avaliação preliminar, podemos dizer que o grupo de física aplicada da UFES teve sucesso em implementar e utilizar o programa GROMACS. Nossa proposta é dar continuidade a este estudo aumentando o tempo de simulação computacional para simular tempos de pressão aplicada da ordem de milisegundos. Além

disso, temos trabalhado no sentido de implementar uma célula de pressão capaz de realizar os experimentos na PME do tomate.

## Capítulo 9

# Perspectivas Futuras

Como expectativas futuras na continuação desta pesquisa experimental, um aspecto fundamental da pesquisa foi a obtenção do raio de giro da enzima Pectina Metiltransferase (PME) por meio dos resultados das simulações computacionais com Gromacs, e a partir dos resultados computacionais nossa missão é poder compará-las aos resultados obtidos em laboratório, mais precisamente, por meio da técnica de espalhamento a baixo ângulo (SASX) é possível obter o valor do raio de giro ( $Rg$ ). No espalhamento a baixo ângulo (SAXS), feixes de raios X com comprimentos de onda com valores entre ( $\lambda : 0.5 - 2.0$  Angstrom ) atravessam a amostra e interagem com os elétrons do material, fazendo-os alterarem seus níveis enérgicos. Ao voltar ao seu estado de equilíbrio emitem radiação secundária. A radiação espalhada e as ondas eletromagnéticas secundárias espalhadas interferem umas com as outras, cancelando-se completamente em algumas direções (interferência destrutiva). Para ângulos pequenos a intensidade alcança seu máximo quando  $2\theta \approx 0^\circ$ , onde todas as ondas estão em fase (interferência construtiva) (63, 64).

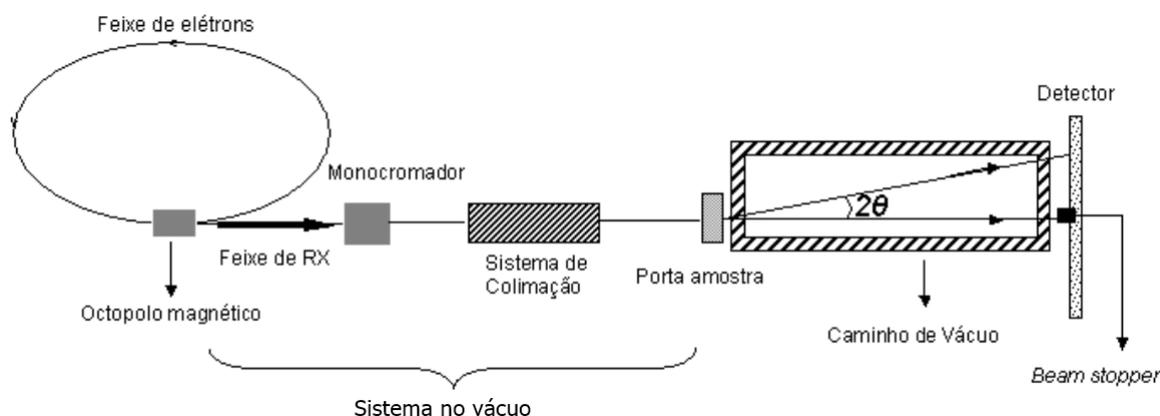


Figura 72 – Espalhamento de raios X do Laboratório Nacional de Luz Síncrotron. Figura adaptada da referência (64).

No espalhamento de raios X, emprega-se técnicas analíticas e não destrutivas para

analisar a forma e o tamanho de partículas tão pequenas que não podem ser vistas a olho nu (dimensões nanométricas;  $1nm = 10^{-9}m$ ), como por exemplo, as proteínas.(63, 64).

Em um sistema de partícula idênticas, baixa polidispersão e baixa anisometria e esfericamente simétrico, temos a intensidade da curva de SAXS, dada por (64):

$$I(q) = kn_p \langle F^2(q) \rangle \{1 + 4\pi n_p \int [g(r) - 1] \frac{\text{sen}(qr)}{qr} r^2 dt\} = kn_p P(q) S(q), \quad (9.1)$$

onde  $k$  é uma constante relacionada ao experimento que pode ser obtida através dos resultados do SAXS de uma amostra, e a densidade numérica das partículas é  $n_p$ . Agora  $P(q) = \langle F^2(q) \rangle$ , é conhecido como fator de forma da partícula espalhadora na qual  $F(q)$  é a amplitude de espalhamento com  $q = \frac{4\pi}{\lambda} \text{sen}\theta$ , sendo  $\lambda$  o comprimento de onda da radiação que incide e o ângulo de espalhamento é  $2\theta$ . A função de distribuição radial é  $g(r)$ , que é entendida como a probabilidade de encontrar uma partícula distante de um raio  $r$  do centro de uma partícula de referência e chama-se de função de interferência entre partículas espalhadoras  $S(q) = 1 + 4\pi n_p \int [g(r) - 1] \frac{\text{sen}(qr)}{qr} r^2 dr$  (64).

No caso de sistemas não interagentes  $S(q) \approx 1$ , pode-se obter o raio de giro através da lei de Guinier dado por um função exponencial (64):

$$I(q \rightarrow 0) = I(0) e^{-\frac{q^2 Rg^2}{3}}. \quad (9.2)$$

Esta relação é válida quando  $qRg \leq 1$  e  $I(0)$  pode ser escrito como (64):

$$I(0) = kn_p \Delta\rho^2 V^2, \quad (9.3)$$

onde  $\Delta\rho = \rho_{particula} - \rho_{meio}$ , é a diferença da densidade eletrônica das partículas em relação ao meio e o volume das partículas espalhadoras é  $V$ . Uma outra maneira de escrevermos a intensidade de espalhamento, é através da função de distribuição de distâncias,  $p(r)$  (64). Esta função está relacionada com a probabilidade de encontrarmos dois centros espalhadores distantes de  $r$  da partícula em estudo. Esta função pode ser escrita como (64):

$$p(r) = \frac{1}{2\pi^2} \int I(q) qr \text{sen}(qr) dq; P(q) = 4\pi \int p(r) \frac{\text{sen}(qr)/qr}{d} r. \quad (9.4)$$

Através da função distribuição de distâncias  $p(r)$ , é possível obter o valor do raio de giro ( $Rg$ ) da proteína (64):

$$Rg^2 = \frac{\int p(r) r^2 dr}{2 \int p(r) dr} \quad (9.5)$$

A função de distribuição espacial  $p(r)$ , é obtida através da transformada inversa de Fourier da curva de espalhamento experimental da intensidade  $I(q)$  (64).

A partir da análise da figura 71, vemos que o desafio é encontrar a resposta para questão da inativação, ou seja, a possibilidade da proteína desnovelar perdendo funcionalidade ou não desnovelar e manter as suas propriedades biológicas (47). Será necessário medidas experimentais do raio de giro da Pectina Metilesterase (PME) por meio de técnicas de espalhamento a baixo ângulos (SAXS), possibilitando comparar com os dados teóricos da simulação afim de averiguar a confiabilidade no programa GROMACS (64).

# Apêndices

# APÊNDICE A

## Configuração das Proteínas

### A.1 Momento do dipolo elétrico do grupo peptídico

No século VI a.C, Tales de Mileto foi quem descobriu os fenômenos elétricos, através de uma seiva de árvore solidificada que representa uma resina amarela conhecida como âmbar, quando atritado com pele de animais possui a capacidade de atrair sementes ou fragmentos de palha (65). Em 1600, William Gilbert menciona outros corpos que podem ser eletrizados por atrito citando os materiais como vidro, enxofre e o lacre (65). Em 1729, Stephen Gray, descobriu que cargas elétricas podia se deslocar em diferentes tipos de materiais, os bons isolantes âmbar, quartzo, vidro, água destilada, ar seco, borracha e plástico tendem suas cargas a permanecer retidas, enquanto bons condutores como metal, água em solução com ácidos, bases ou sais, corpo humano, terra tendem suas cargas a ser transmitidas com mais facilidade que os isolantes (65).

Em 1733, Charles du Fay, estudando objetos eletrizados a partir do atrito do tecido como o âmbar e vidro, ao ser aproximados âmbar com âmbar se repeliam e âmbar com vidro se atraíam. Posteriormente, Benjamin Franklin classificou os corpos eletrizados possuindo “carga positiva” ou “carga negativa”, então interpretou a experiência de Charles du Fay como a conhecido princípio da atração e repulsão das cargas: cargas de mesmo sinal se repelem, enquanto de sinais contrários se atraem (65).

Por meio dos seus estudo com balança de torção. No ano de 1775, Charles-Augustin Coulomb, pode enunciar a força eletrostática (lei das cargas em repouso). Expressando a lei de Coulomb em uma linguagem matemática de vetores, representando uma força proporcional ao produto das cargas e inversamente proporcional ao quadrado da distância entre elas, então a força elétrica entre as cargas  $q$  e  $q'$  é (65),

$$\vec{F} = q' \frac{q}{4\pi\epsilon_0 r^2} \hat{r}, \quad (\text{A.1})$$

onde  $\epsilon_0$  é a constante chamada de permissividade do espaço livre, e  $r$  é a distância

entre as duas cargas com vetor unitário  $\hat{r} = \vec{r}/r$  representando a direção da força  $\vec{F}$  e o sentido é dado pelo produto das cargas  $q$  e  $q'$ . Logo, uma força elétrica sobre uma carga possui a direção do vetor unitário  $\hat{r}$ . Podemos determinar sua reação. E se dá em sentido oposto, e com a mesma magnitude, sobre a outra a carga de acordo com a lei da ação e reação de Newton. Então, um corpo que aplica uma força receberá de volta a mesma força, contudo em sentido contrário. Cada carga possui intrinsecamente um campo elétrico, que sinaliza a possibilidade de interação elétrica com as cargas da vizinhança:

$$\vec{F} = q'\vec{E} \quad ; \quad \vec{E} = \frac{q}{4\pi\epsilon_0 r^2} \hat{r}, \quad (\text{A.2})$$

onde o campo elétrico  $\vec{E}$  está na mesma direção do campo elétrico e o sentido é determinado pelo sinal da carga elétrica, portanto para ter uma força elétrica necessita ter duas cargas (65). Então, Michael Faraday (1791-1867), para facilitar no entendimento dos fenômenos eletromagnéticos sugeriu o conceito de linhas de força do campo elétrico e magnético indicando sua direção sendo tangente aos pontos da linha, com sentido da orientação das linhas e com magnitude proporcional a concentração das linhas (65).

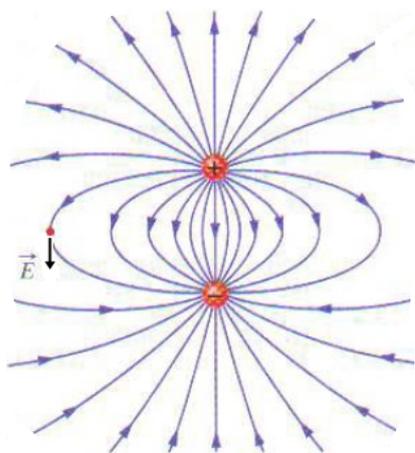


Figura 73 – Linhas de dipolo elétrico. Figura adaptada da referência (66).

Em 1909, houve uma importante descoberta da carga elementar realizada por Robert A. Milikan e Harvey Fletcher, a partir do experimento da gotícula de óleo. Foi verificado que cargas elétricas são quantizadas. Ou seja, sempre com valores múltiplos inteiros elementar, interpretada como a carga do elétron  $e$ , dada por:

$$e = 1,602 \cdot 10^{-19} C, \quad (\text{A.3})$$

onde  $C$  é a grandeza de carga Coulomb no sistema internacional (65). O momento de dipolo elétrico  $\vec{\mu}$ , é o responsável pela estabilização da hélice- $\alpha$  e folha- $\beta$ . Isto permite que ocorra as pontes de hidrogênio. O dipolo é formado por duas cargas sinais opostos. Em

um sistema de cargas, os dipolos representam a direção e o sentido em que a molécula está orientada e sua magnitude o poder de interação com a vizinhança (2). O grupo peptídico localizado no plano amida da cadeia polipeptídica, possui um dipolo resultante responsável pela estabilização de algumas proteínas. Portanto, é de muita importância entendermos seu comportamento, então o grupo peptídico está localizado no sistema de coordenadas cartesianas (x,y) com o sentido positivo (x) para a esquerda e o sentido positivo (y) para baixo, numa escala de picômetros  $10^{-12}$  m:

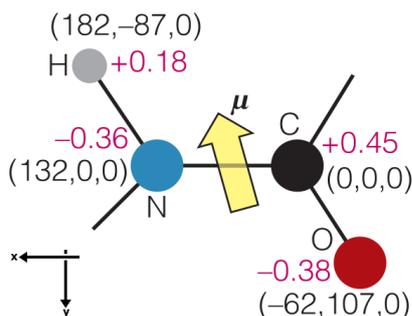


Figura 74 – Dipolo elétrico do grupo peptídico  $\vec{\mu}$ . Figura adaptada da referência (67).

onde a molécula de hidrogênio com carga  $(0,18e)$ , fica localizada no plano com as coordenadas  $(182,-87)$ . Para a molécula de nitrogênio com carga  $(-0,36e)$ , está localizada em  $(132,0)$ . No caso da molécula de carbono  $(0,45e)$ , se encontra na origem do sistema de coordenadas  $(0,0)$ . Finalmente, a molécula de oxigênio com carga  $(-0,38e)$ , é localizada em  $(-62,107)$ . Em relação ao par de cargas elétricas  $q$  e  $-q$ , embora sua carga total seja nula, formam um sistema chamado de dipolo responsáveis por interações elétricas com a vizinhança fornecendo o mecanismo de enovelamento ou estabilização das proteínas. O dipolo  $\vec{\mu}$ , é a soma dos dipolos individuais do sistema (67):

$$\vec{\mu} = q \vec{d}, \quad (\text{A.4})$$

onde o vetor posição  $\vec{d}$ , com origem na carga negativa e apontando para a carga positiva, onde  $q$  é carga líquida (65). A magnitude do dipolo elétrico  $\vec{\mu}$  na direção (x):

$$\mu_x = (-0,36e) \cdot (132pm) + (0,45e) \cdot (0pm) + (0,18e) \cdot (182pm) + (-0,38e) \cdot (-62,0pm) \quad (\text{A.5})$$

logo,

$$\mu_x = 8.8e pm = 8.8 \cdot (1,602 \cdot 10^{-19} C) (10^{-12} m), \quad (\text{A.6})$$

portanto a magnitude do momento de dipolo elétrico  $\mu_x$  :

$$\mu_x = 1,4 \cdot 10^{-30} Cm. \quad (\text{A.7})$$

Agora em relação ao dipolo, sua magnitude na direção (y), é dado por:

$$\mu_y = (-0,36e).(0pm) + (0,45e).(0pm) + (0,18e).(-87pm) + (-0,38e).(107pm) \quad (\text{A.8})$$

portanto,

$$\mu_y = -56e pm = -56.(1,602.10^{-19}C)(10^{-12}m), \quad (\text{A.9})$$

Então a magnitude do momento de dipolo elétrico  $\mu_y$  :

$$\mu_y = -9,0.10^{-30}Cm. \quad (\text{A.10})$$

A grandeza física para representar os dipolos é o Debye, com uma unidade correspondendo  $1D=3,335.10^{-30}Cm$ . O momento de dipolo pode ser escrito como:

$$\vec{\mu} = (\mu_x, \mu_y) = \mu_x \hat{i} + \mu_y \hat{j} = [(0,42D), (-2,7D)] = 0,42D \hat{i} + (-2,7D) \hat{j}. \quad (\text{A.11})$$

Cálculo da magnitude do dipolo elétrico  $\mu$  é:

$$\mu = \sqrt{\mu_x^2 + \mu_y^2} = \sqrt{(0,42D)^2 + (-2,7D)^2} \quad (\text{A.12})$$

a magnitude do momento de dipolo resultante  $\mu$  pode ser aproximado para o valor:

$$\mu = 2,7D. \quad (\text{A.13})$$

Note que o dipolo ( $\mu_x$ ) na direção (x) tem magnitude bem inferior ao dipolo ( $\mu_y$ ) em relação a direção (y). Temos uma influência bem fraca do dipolo ( $\mu_x$ ), na magnitude do dipolo resultante  $\mu$ . Pois, o valor do momento do dipolo é  $\mu = \sqrt{0 + \mu_y^2} = 2,7D$ . Logo, por este motivo muitas vezes  $\mu_x$  é ignorado em muitos livros textos, sendo o momento de dipolo resultante representando por  $\vec{\mu} = \mu_y \hat{j}$  (67).

## A.2 Energia do dipolo elétrico

Note que o tamanho do grupo peptídico e as distâncias entre os grupos peptídicos são da ordem de Angstrom ( $10^{-10}m$ ). Sendo assim, não é apropriado usar a aproximação por meio da expansão de Taylor do tipo  $r \gg l$  (distâncias do braço do dipolo muito maior que as distâncias de estudo), pois neste caso as distâncias inter-moleculares dos aminoácidos são aproximadamente as distâncias intra-moleculares (ver as figuras 8 e 13). Para descobrir como o dipolo interage com outras cargas, precisamos conhecer o seu campo elétrico:

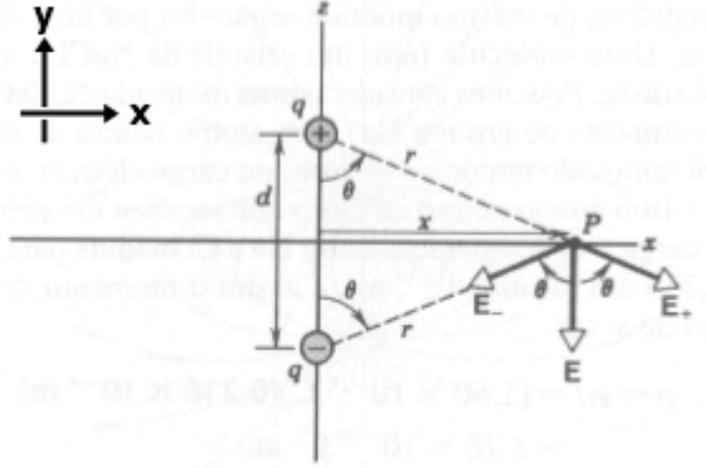


Figura 75 – Diagrama da atuação do campo elétrico. Figura adaptada da referência (71).

Se o campo elétrico indica a presença de cargas, então para o caso do dipolo o campo elétrico resultante é a soma do campo elétrico da carga  $q$  ( $\vec{E}_+$ ), com o campo elétrico da carga  $-q$  ( $\vec{E}_-$ ). Então, o campo elétrico resultante ( $\vec{E}_r$ ), é dado por:

$$\vec{E}_r = \vec{E}_+ + \vec{E}_- = (E_x, E_y) + (E'_x, E'_y) = (E_x \hat{i} + E_y \hat{j}) + (E'_x \hat{i} + E'_y \hat{j}). \quad (\text{A.14})$$

Portanto, de acordo com o sistema de coordenadas cartesianas (x,y):

$$\vec{E}_r = (E \sin \theta, -E \cos \theta, 0) + (-E' \sin \theta, -E' \cos \theta, 0) = (E \sin \theta \hat{i} - E \cos \theta \hat{j}) + (-E' \sin \theta \hat{i} - E' \cos \theta \hat{j}), \quad (\text{A.15})$$

onde os módulos dos vetores unitários ( $|\hat{r}| = |\hat{r}'| = 1$ ) e os módulos das cargas ( $|q| = |-q| = q$ ). Como as cargas estão a mesma distância do ponto de estudo situado no eixo x ( $r = r'$ ), o módulo do vetor campo elétrico é:

$$|\vec{E}_+| = |\vec{E}_-| = E_+ = E_- = \frac{q}{4\pi\epsilon_0 \left[ x^2 + \left(\frac{d}{2}\right)^2 \right]}. \quad (\text{A.16})$$

O campo elétrico resultante  $\vec{E}_r$  se soma na direção (y) pois ( $\vec{E}'_y, \vec{E}_y$ ) possuem o mesmo módulo e sentido, todavia devido ( $\vec{E}'_x, \vec{E}_x$ ) terem o mesmo módulo e sentidos diferentes na direção (x) cancelam-se por questões de simetria, sua representação vetorial é:

$$\vec{E}_r = -E \cos \theta \hat{j} - E' \cos \theta \hat{j} = -2E \cos \theta \hat{j}, = (0, -2E \cos \theta, 0), \quad (\text{A.17})$$

onde,

$$\cos\theta = \frac{d/2}{\sqrt{x^2 + (\frac{d}{2})^2}}. \quad (\text{A.18})$$

A magnitude do momento de dipolo é  $\mu = qd$ ,

$$\vec{E}_r = -\frac{\mu}{4\pi\epsilon_0 \left[ x^2 + (\frac{d}{2})^2 \right]^{\frac{3}{2}}} \hat{j}, \quad (\text{A.19})$$

onde o campo resultante  $\vec{E}_r$ , é dado em termo do módulo do dipolo  $\vec{\mu}$  e está no sentido para baixo na direção ( $y$ ) (65). Agora iremos analisar a situação de mínima energia do dipolo. Para isso, vamos levar em consideração alguns conceitos da mecânica. Primeiramente, a energia representa uma capacidade de realizar trabalho. Para entendermos como a força elétrica é dita conservativa, ou seja, seu trabalho independe do caminho escolhido, lançamos mão do teorema trabalho-energia potencial:

$$\Delta U = -W_{A \rightarrow B}^{\vec{F}} = -\int_A^B \vec{F} \cdot d\vec{r}. \quad (\text{A.20})$$

O trabalho de uma força  $\vec{F}$  que atua sobre um corpo da posição inicial A até a posição final B. A força elétrica atuando sobre o deslocamento diferencial  $d\vec{r} = dr\hat{r}$ :

$$\Delta U = -W_{A \rightarrow P}^{\vec{F}} = -\int_A^P \vec{F} \cdot d\vec{r} = -\int_A^P q' \frac{q}{4\pi\epsilon_0 r^2} dr = \left[ q' \frac{q}{4\pi\epsilon_0 r_P} \right] - \left[ q' \frac{q}{4\pi\epsilon_0 r_A} \right] = q'\psi_P - q\psi_A, \quad (\text{A.21})$$

Implicitamente as energias são calculas em relação ao ponto de referência A no infinito  $PR(\infty)$ . Logo, a energia no ponto P é:

$$\Delta U = -W_{PR(\infty) \rightarrow P}^{\vec{F}} = U_P - U_{PR(\infty)} = \left[ q' \frac{q}{4\pi\epsilon_0 r_P} \right] - 0, \quad (\text{A.22})$$

Como  $\Delta U = U_P$ , temos que  $U_{PR(\infty)} = 0$ . Portanto:

$$U_P = -W_{PR(\infty) \rightarrow P}^{\vec{F}} = q' \frac{q}{4\pi\epsilon_0 r_P} = q'\psi_P, \quad (\text{A.23})$$

onde a energia potencial  $U_P$ , pode ser decomposta no produto de dois termos. O termo ligado a fatores do campo elétrico e outro que depende da carga de teste  $q'$ . Para o termo que depende exclusivamente de fatores ligados ao campo elétrico, chama-se de diferença de potencial elétrico  $\psi$  (d.d.p). Sendo assim, para o caso do dipolo temos duas cargas. Para a carga de teste positiva  $q' = q$  ( $U = q\psi_+$ ) e para a carga de teste negativa

$q' = -q$  ( $U_- = -q\psi_-$ ). Então, a energia total do sistema é a soma das energias dos pares de cargas (68):

$$U = U_+ + U_- = q\psi_+ - q\psi_- = q[\psi_+ - \psi_-] = q\Delta\psi, \quad (\text{A.24})$$

onde a partir do sistema cartesiano de eixos pode-se escrever  $\psi_- = \psi(x, y, z)$  e  $\psi_+ = \psi(x + \Delta x, y + \Delta y, z + \Delta z)$ . Deve-se notar que as componentes do vetor  $\vec{d}$ , são  $(\Delta x, \Delta y, \Delta z)$ . Pode-se escrever esta igualdade do "cálculo diferencial e integral" (68):

$$\Delta\psi = \frac{\partial\psi}{\partial x}\Delta x + \frac{\partial\psi}{\partial y}\Delta y + \frac{\partial\psi}{\partial z}\Delta z. \quad (\text{A.25})$$

O gradiente de uma função é (68):

$$\vec{\nabla}\psi = \frac{\partial\psi}{\partial x}\hat{i} + \frac{\partial\psi}{\partial y}\hat{j} + \frac{\partial\psi}{\partial z}\hat{k} \quad ; \quad \vec{d} = \Delta x\hat{i} + \Delta y\hat{j} + \Delta z\hat{k}, \quad (\text{A.26})$$

logo,

$$U = q\Delta\psi = q(\vec{\nabla}\psi \cdot \vec{d}) = \vec{\mu} \cdot \vec{\nabla}\psi. \quad (\text{A.27})$$

É importante lembrar agora teorema do trabalho-energia infinitesimal  $dU = -dW$

$$q'd\psi = -q'\vec{E} \cdot \vec{ds}, \quad (\text{A.28})$$

$\vec{ds}$  é o vetor infinitesimal da trajetória, logo temos (68):

$$E_x = -\frac{\partial\psi}{\partial x}; E_y = -\frac{\partial\psi}{\partial y}; E_z = -\frac{\partial\psi}{\partial z}. \quad (\text{A.29})$$

Podemos então concluir que o campo elétrico nas três direções é dado igual a menos o gradiente do potencial elétrico,  $\vec{E} = -\vec{\nabla}\psi$ . Com isso, a energia potencial  $U$  em função do campo elétrico  $\vec{E}$  (68):

$$U = \vec{\mu} \cdot \vec{\nabla}\psi = -\vec{\mu} \cdot \vec{E} = -E\mu\cos\theta, \quad (\text{A.30})$$

onde no caso do ângulo  $\theta = \pi$  com  $U = \mu E$ , ou seja,  $\vec{\mu}$  e  $\vec{E}$  antiparalelos representado um sistema de energia máxima. No entanto, a natureza tende a estados de menor energia possível, sobre o ponto de vista do dipolo buscamos um torque que em que o sistema de menor energia é  $\theta = 0^\circ$  com  $U = -\mu E$ , ou seja,  $\vec{\mu}$  e  $\vec{E}$  paralelos  $\vec{E}$  (68).

# APÊNDICE B

## Ensembles

### B.1 Parâmetros extensivos e intensivos

*Grandezas extensivas:* São variáveis de estado que dependem do tamanho ou da quantidade de massa do sistema. Se o sistema é subdividido em várias partes, então o valor total da grandeza extensiva é igual a soma dos valores das partes individuais. Exemplos: Energia, volume, entropia (55, 69).

*Grandezas intensivas:* São variáveis de estado que independem do tamanho ou da quantidade de massa do sistema. Se o sistema é subdividido em várias partes então o valor total da grandeza intensiva é igual os valores das partes individuais. Exemplos: Temperatura, pressão, densidade, potencial químico, calor específico, coeficiente de expansão térmica, compressibilidade isotérmica (55, 69).

## B.2 Ensemble Microcanônico

Um sistema físico formado por dois fluidos simples (ausente de campos externos: elétricos, magnéticos ou gravitacional) separados por uma parede adiabática, fixa e impermeável. Sendo seu estado microscópico representado por uma energia  $E$ , volume  $V$  e número de partículas  $N$ , e podendo estar vinculado à um conjunto de  $X_i$  de vínculos internos. Isto é, uma caracterização das propriedades termodinâmicas. O primeiro postulado ou o postulado fundamental da mecânica estatística de equilíbrio (valem os postulados da termodinâmica), define o ensemble microcanônico: *os estados microscópicos acessíveis a um sistema fechado em equilíbrio são igualmente prováveis*. Como consequência do primeiro postulado é de a probabilidade de achar o sistema aos vínculos é proporcional ao número de estados microscópicos  $\Omega(E, V, N; X_i)$  (55, 69):

$$P(X_i) \propto \Omega(E, V, N; X_i). \quad (\text{B.1})$$

Sendo assim, o segundo postulado da mecânica estatística de equilíbrio é definição da entropia reproduzindo a conexão do ensemble microcanônico com a termodinâmica:

$$S(E, V, N) = k_B \ln \Omega(E, V, N). \quad (\text{B.2})$$

A entropia  $S$  é dada então pelo logaritmo do número de microestados (55, 69). No entanto, essa conexão deve ser feita no limite termodinâmico, ou seja, para  $E, V, N \rightarrow \infty$ , com densidades fixas,  $E/N = u$ ,  $V/N = v$ , onde  $u$  e  $v$  são constantes (55, 69).

## B.3 Ensemble Canônico

No caso de um sistema termodinâmico simples em contato com um reservatório térmico através de uma parede diatérmica, mas fixa e impermeável. No caso do sistema composto isolado, com energia total  $E_0$ , os postulados fundamentais da mecânica estatística de equilíbrio podem ser utilizados. A probabilidade  $P_j$  de encontrar o sistema  $S$  num determinado estado microscópico  $j$  é dado por (55, 69),

$$P_j = c \Omega_R(E_0 - E_j), \quad (\text{B.3})$$

$c$  é uma constante de normalização,  $E_j$  é a energia do sistema  $S$  no particular estado microscópico  $j$  e  $\Omega_R(E)$  é o número de estados microscópicos acessíveis ao reservatório  $R$  com energia  $E$ , onde pode-se mostrar:

$$P_j = \frac{\exp(-\beta E_j)}{\sum_j \exp(-\beta E_j)}, \quad (\text{B.4})$$

onde  $\beta = \frac{1}{k_b T}$ . O conjunto de microestados  $j$  associados a esta distribuição de probabilidades acessíveis ao sistema  $S$ , em contato com um reservatório térmico a temperatura  $T$ , definem o ensemble canônico. Logo, se define a função de partição a partir da normalização da probabilidade  $P_j$

$$Z = \sum_j \exp(-\beta E_j), \quad (\text{B.5})$$

é uma soma dos fatores de Boltzmann  $\exp(-\beta E_j)$  dos estados microscópicos acessíveis ao sistema. A partir de então podemos determinar a energia livre de Helmholtz  $F$ ,

$$F = F(T, V, N) \rightarrow -\frac{1}{\beta} \ln Z(T, V, N), \quad (\text{B.6})$$

se fizermos o limite termodinâmico temos,

$$-\beta f(T, v) = \lim_{V, N \rightarrow \infty; \frac{V}{N} = v} \frac{1}{N} \ln Z(T, V, N), \quad (\text{B.7})$$

onde a energia livre de Helmholtz por partícula é  $f$ . Entropia por partícula  $s$  e a energia por partícula  $u$ , vão ser dados por:

$$s = -\frac{\partial f}{\partial T} \quad ; \quad u = -\frac{1}{N} \frac{\partial}{\partial \beta} \ln Z. \quad (\text{B.8})$$

A conexão entre o ensemble canônico e a termodinâmica é realizado no limite termodinâmico, na qual é exprimido pela energia livre de Helmholtz por partícula (55, 69).

# APÊNDICE C

## Pressão do sistema

Em relação a pressão do sistema, faremos uso do teorema do virial. Definimos primeiro a quantidade física virial  $G$ . Esta expressão tem origem do Latim, vis significa "força" ou "energia" e foi denominada por Rudolf Clausius (1822-1888) (70):

$$G = \sum_{i=1}^N \vec{q}_i \vec{p}_i, \quad (\text{C.1})$$

onde  $q_i$  e  $p_i$  são funções limitadas no tempo, logo, o virial  $G$  também está delimitado. O numerador é limitado mas o denominador cresce indefinidamente, então (70):

$$\left\langle \frac{dG}{dt} \right\rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} \frac{dG}{dt} dt = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} (G(\tau) - G(0)) = 0. \quad (\text{C.2})$$

A partir deste resultado do teorema do virial, podemos trabalhar o valor esperado do virial (59, 60):

$$\left\langle \frac{dG}{dt} \right\rangle = \left\langle \frac{d}{dt} \left[ \sum_{i=1}^N \vec{q}_i \vec{p}_i \right] \right\rangle = 2 \left\langle \sum_{i=1}^N \frac{\vec{p}_i^2}{2m} \right\rangle + \left\langle \sum_{i=1}^N \vec{q}_i \vec{F}_i(q) \right\rangle = 2 \langle K(p) \rangle + \left\langle \sum_{i=1}^N \vec{q}_i \vec{F}_i(q) \right\rangle = 0. \quad (\text{C.3})$$

$K(p)$  é a energia cinética em função do momento e  $f_i(q)$  é a força total aginda na partícula  $i$ , logo:

$$\vec{F}_i(q) = \vec{F}_{int}(q_i) + \vec{F}_{ext}(q_i). \quad (\text{C.4})$$

A soma das forças internas  $\vec{F}_{int}(q_i)$  com as forças externas  $\vec{F}_{ext}(q_i)$ , e a derivada parcial é  $\partial_x = \partial/\partial x$  (59, 60). Com suporte do teorema do virial podemos escrever:

$$\left\langle \frac{dG}{dt} \right\rangle = 2 \langle K(p) \rangle + \left\langle \sum_{i=1}^N \vec{q}_i \cdot \vec{F}_i(q) \right\rangle = 2 \langle K(p) \rangle + \left\langle \sum_{i=1}^N q_i [\vec{F}_{int}(q_i) + \vec{F}_{ext}(q_i)] \right\rangle = 0. \quad (\text{C.5})$$

Aplicando a propriedade de valor esperado da soma, é a soma dos valores esperados:

$$\left\langle \frac{dG}{dt} \right\rangle = 2 \langle K(p) \rangle + \left\langle \sum_{i=1}^N \vec{q}_i \cdot \vec{F}_{int}(q_i) \right\rangle + \left\langle \sum_{i=1}^N \vec{q}_i \cdot \vec{F}_{ext}(q_i) \right\rangle = 0. \quad (\text{C.6})$$

A força externa, está associada há uma pressão sobre o sistema (59, 60):

$$\left\langle \sum_{i=1}^N \vec{q}_i \cdot \vec{F}_{ext}(q_i) \right\rangle = N \langle \vec{q}_1 \cdot \vec{F}_{ext}(q_1) \rangle = 3N \langle \vec{q}_{1_x} \cdot \vec{F}_{ext}(q_{1_x}) \rangle = 3N \langle W_{ext} \rangle. \quad (\text{C.7})$$

O valor esperado do trabalho externo, pode ser obtido fazendo uso do método de Monte Carlo, portanto (46):

$$\langle W_{ext} \rangle = \frac{1}{N} \sum_i W_{ext(i)} = \frac{1}{N} W_{ext}. \quad (\text{C.8})$$

E o trabalho sobre o sistema é representado por  $W_{ext} = -pV$  (21), temos:

$$\left\langle \sum_{i=1}^N \vec{q}_i \cdot \vec{F}_{ext}(q_i) \right\rangle = 3N \langle W_{ext} \rangle = 3N(W_{ext}/N) = -3pV. \quad (\text{C.9})$$

Podemos relacionar o valor esperado da energia cinética do sistema com sua respectiva temperatura (59, 60):

$$\langle K \rangle = \frac{3}{2} N k_B T. \quad (\text{C.10})$$

Portanto, segue-se:

$$\left\langle \frac{dG}{dt} \right\rangle = 2 \left( \frac{3}{2} N k_B T \right) + \left\langle \sum_{i=1}^N \vec{q}_i \cdot \vec{F}_{int}(q_i) \right\rangle - 3pV = 0. \quad (\text{C.11})$$

Em um dado instante da simulação, a pressão do sistema é (59, 60, 61):

$$P = \frac{N k_B T}{V} + \frac{1}{3V} \left\langle \sum_{i=1}^N \vec{q}_i \cdot \vec{F}_{int}(q_i) \right\rangle. \quad (\text{C.12})$$

Sendo assim,

$$\sum_{i=1}^N \vec{q}_i \vec{F}_{int}(q_i) = \sum_{i=1}^N \sum_{j \neq i}^N \vec{q}_i \vec{F}_{ij} = \left( \sum_{i=1}^N \sum_{j < i}^N \vec{q}_i \vec{F}_{ij} + \sum_{i=1}^N \sum_{j > i}^N \vec{q}_i \vec{F}_{ji} \right). \quad (\text{C.13})$$

Mas, pela terceira lei de Newton  $\vec{F}_{ij} = -\vec{F}_{ji}$ :

$$\sum_{i=1}^N \vec{q}_i \vec{F}_{int}(q_i) = \left( \sum_{i=1}^N \sum_{j < i}^N \vec{q}_i \vec{F}_{ij} - \sum_{i=1}^N \sum_{j < i}^N \vec{q}_j \vec{F}_{ij} \right) = \left( \sum_{i=1}^N \sum_{j < i}^N (\vec{q}_i - \vec{q}_j) \vec{F}_{ij} \right), \quad (\text{C.14})$$

onde a partícula  $i$  em relação a partícula  $j$ , possui vetor posição  $\vec{q}_{ij} = \vec{q}_i - \vec{q}_j$ . Portanto, em um dado instante da simulação a pressão pode ser obtida através do teorema do virial. Lembrando, uma força conservativa possui trabalho independente do caminho. A coordenada generalizada é dada em função da distância  $\vec{q}_{ij} = \vec{r}_{ij}$  (59, 60, 61):

$$p(t) = \frac{Nk_b T}{V} + \frac{1}{3V} \left\langle \sum_{i=1}^N \sum_{j < i}^N \vec{r}_{ij}(t) \vec{F}_{ij}(t) \right\rangle, \quad (\text{C.15})$$

onde  $V$  é o volume,  $N$  o número de partículas,  $T$  a temperatura do sistema,  $k_B$  é a constante de Boltzmann. Os vetores posição e força, são  $\vec{r}_{ij}$  e  $\vec{F}_{ij}(t)$ , respectivamente (44).

# APÊNDICE D

## Algoritmo da simulação

O Gromacs é um pacote de simulação computacional baseado principalmente em Dinâmica Molecular (MD) e Monte Carlo (MC), foi utilizado neste trabalho a versão 5.1.4. É recomendado trabalhar na plataforma Linux, por questões de melhores performance de processamento e sem problemas de compatibilidade. As proteínas utilizadas no estudo são obtidas através do banco digital "Protein Data Bank (PDB)". Nosso objetivo é obter o "raio de giro", da enzima pectina metilesterase do tomate (1xg2).

Segue-se então o algoritmo de simulação. O download da estrutura protéica, portanto será necessário visitar o depósito oficial de modelos de proteínas na internet, o site do Research Collaboratory for Structural Bioinformatics (RCSB) (<http://www.rcsb.org/>). No nosso caso corresponde ao arquivo 1xg2.pdb e servirá como um arquivo de entrada "input" para o Gromacs. Seguiremos o seguinte tutorial:

<http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/index.html>

Primeiro acessa o diretório de instalação, portanto entre em "build" e agora cria uma pasta por exemplo "PME", neste caso se referindo a proteína pectina metil esterase. Salve o arquivo 1xg2.pdb na pasta "PME", além disso crie cinco arquivos de texto vazios com o seguintes nomes: "ions.mdp", "minim.mdp", "nvt.mdp", "npt.mdp" e "md.mdp". O conteúdo destes arquivos podem ser encontrados no "Apêndice E" desta dissertação.

### D.0.1 Etapas

Através de um programa auxiliar PYMOL, clique na opção "S" na parte inferior direita, isso possibilita apagar os ligantes. Agora clique nas moléculas de água H<sub>2</sub>O estando abreviadas do número zero e pode ser encontrada passando a barra na parte superior, com o botão direito escolha a opção REMOVE e salve as alterações. Por meio deste tutorial vamos dar início a nossa simulação com Gromacs:

<http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/lysozyme/index.html>

Entre na pasta "PME" por meio do terminal do Linux e digite:

```
gmx pdb2gmx -f 1xg2.pdb -o 1xg2_processed.gro -water spce
```

onde -f e -o significam o arquivo entrada e saída, respectivamente. Será apresentado algumas opções de campo de força, digite "15" para selecionar a opção OPLS-AA/L, este campo é um ajuste nos coeficientes de Fourier torsional do campo de força OPLS-AA de forma a otimizar os resultados da simulação. O comando editconf irá criar a "caixa de simulação" para depositarmos a proteína

```
gmx editconf -f 1xg2_processed.gro -o 1xg2_newbox.gro -c -d 1.0 -bt cubic
```

onde se utiliza a caixa cúbica de face centrada. O comando solvate trabalhará no processo de salvatação do sistema

```
gmx solvate -cp 1xg2_newbox.gro -cs spc216.gro -o 1xg2_solv.gro -p topol.top
```

onde é adicionado o modelo de água "Simple Point Charge Water (SPC)" (ângulo da ligação  $H_2O$  é calculado como  $109.47^\circ$ ) à caixa de simulação. Com a descrição atômica realizada, por meio do comando genion na qual aleatoriamente substitui as moléculas de solvente por íons monoatômicos e fará a neutralização das cargas do sistema para isso selecione "13":

```
gmx genion -s ions.tpr -o 1xg2_solv_ions.gro -p topol.top -neutral
```

onde a opção "SOL" é para incorporar os íons sem alterar a estrutura da proteína. No genion, a estrutura/estado de entrada (-s) é o arquivo "ions.tpr" e o arquivo de saída (-o) é representado por "1xg2\_solv\_ions.gro" e processamos a topologia (-p) para a remoção de moléculas de água e a adição de íons. O comando grompp lê o arquivo de topologia molecular (contém informações sobre o números e tipos de moléculas), verifica a sua validade

```
gmx grompp -f minim.mdp -c 1AKI_solv_ions.gro -p topol.top -o em.tpr
```

E importante salientar também que se expande a descrição molecular para uma descrição atômica. O comando mdrun é o programa de dinâmica molecular

```
gmx mdrun -v -deffnm em
```

O usamos neste caso para realizar a minimização da energia (EM). Para analisarmos o resultado usaremos o seguinte comando:

```
gmx energy -f em.edr -o potential.xvg
```

E digite "10 0", isto é, apresenta a média da energia do sistema e pode ser visualizado com um outro programa auxiliar conhecido como Grace. Dinâmica de restrição da temperatura, é o penúltimo passo para a principal simulação, a proteína tem sua movimentação restringida e é feita a otimização da distribuição do solvente envolvido.

```
gmx grompp -f nvt.mdp -c em.gro -p topol.top -o nvt.tpr
```

E tende a entrar em equilíbrio padrão bioquímico de relaxamento de 10ps, portanto é aconselhável rodarmos a simulação com uma ordem de magnitude superior, para o nosso caso (nsteps x dt) = 50000 x 0,002 = 100 ps. Para realizar o processamento do ensemble NVT,

```
gmx mdrun -deffnm nvt
```

É parte da equilibração da temperatura. Agora, para gerarmos o gráfico da temperatura, segue-se:

```
gmx energy -f nvt.edr -o temperature.xvg
```

E digite "15 0", o arquivo pode ser visualizado através do Grace. Seguimos então com a dinâmica de restrição da pressão

```
gmx grompp -f npt.mdp -c nvt.gro -t nvt.cpt -p topol.top -o npt.tpr
```

De maneira análogo ao da temperatura é feita otimização do solvente. Em relação ao processamento do ensemble NPT, fazemos:

```
gmx mdrun -deffnm npt
```

Sendo a última etapa da fase de equilibração do sistema. Podemos então determinar a pressão:

```
gmx energy -f npt.edr -o pressure.xvg
```

E digite "16 0", gera o arquivo que pode ser visualizado pelo Grace. E para determinar a densidade:

```
gmx energy -f npt.edr -o density.xvg
```

E digite "22 0", para gerar o arquivo de visualização. É a parte da dinâmica molecular, na qual gerará os dados que serão utilizados neste trabalho. Antes disso, vamos comentar um pouco sobre este processo. É necessário que a estrutura protéica já esteja e uma lista de vizinhanças é feita, então as forças começam a serem computadas e então as posições e velocidades são atualizadas. Caso necessário, uma agitação é realizada de forma a restringir certos comprimentos e ângulos das ligações.

```
gmx grompp -f md.mdp -c npt.gro -t npt.cpt -p topol.top -o md_0_1.tpr
```

```
gmx mdrun -deffnm md_0_1
```

No final do processo é gerado arquivos que salvam a trajetória contendo as coordenadas da posição e velocidade de todos os átomos ao longo do tempo. Além disso guarda os valores de energia, pressão, temperatura. O comando trjconv é usado como ferramenta de análise pós-processamento com a finalidade de corrigir erros:

```
gmx trjconv -s md_0_1.tpr -f md_0_1.xtc -o md_0_1_noPBC.xtc -pbc mol -ur compact
```

Digite "0" caso seja necessário retirar coordenadas, corrigir a periodicidade ou

alterar a trajetória do sistema. A partir desta "trajetória corrigida", será feita mais análises, segue então:

```
gmx rms -s md_0_1.tpr -f md_0_1_noPBC.xtc -o rmsd.xvg -tu ns
```

```
gmx rms -s em.tpr -f md_0_1_noPBC.xtc -o rmsd_xtal.xvg -tu ns
```

Digite "4" no sentido de avaliar a estabilidade estrutural por meio do pacote RMSD. Obtenção do raio de giro por meio do comando:

```
gmx gyrate -s md_0_1.tpr -f md_0_1_noPBC.xtc -o gyrate.xvg
```

Será gerado um arquivo de saída (output) gyrate.xvg, ou seja, seu formato é ".xvg". Todas as imagens neste padrão podem ser visualizados no programa GRACE. É importante destacar que os resultados podem variar dependendo das escolhas da temperatura e pressão encontrados nos arquivos de parâmetro. Se numa próxima simulação necessitar, por exemplo, mudar o valor da pressão é necessário corrigir este valor nos arquivos "npt.mdp" e "md.mdp".

# APÊNDICE E

Para realizar a simulação computacional deve-se em primeiro lugar deve-se se cria cinco documentos vazios com os nomes: "ions.mdp", "minim.mdp", "nvt.mdp", "npt.mdp" e "md.mdp". Estes documentos acomodarão respectivos valores dos parâmetros da dinâmica molecular, segue-se abaixo as respectivas informações para cada arquivo.

## E.1 ions.mdp

```
; ions.mdp - used as input into grompp to generate ions.tpr
; Parameters describing what to do, when to stop and what to save
integrator = steep      ; Algorithm (steep = steepest descent minimization)
emtol = 1000.0         ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm
emstep = 0.01         ; Energy step size
nsteps = 5000         ; Maximum number of (minimization) steps to perform
; Parameters describing how to find the neighbors of each atom and how to calculate the
interactions
nstlist = 1           ; Frequency to update the neighbor list and long range forces
cutoff-scheme = Verlet
ns_type = grid        ; Method to determine neighbor list (simple, grid)
coulombtype = PME      ; Treatment of long range electrostatic interactions
rcoulomb = 1.0        ; Short-range electrostatic cut-off
rvdw = 1.0           ; Short-range Van der Waals cut-off
pbc = xyz             ; Periodic Boundary Conditions (yes/no)
```

## E.2 minim.mdp

```
; minim.mdp - used as input into grompp to generate em.tpr
integrator = steep      ; Algorithm (steep = steepest descent minimization)
emtol = 1000.0         ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm
```

```

emstep = 0.01      ; Energy step size
nsteps = 50000    ; Maximum number of (minimization) steps to perform
; Parameters describing how to find the neighbors of each atom and how to calculate the
interactions
nstlist = 1      ; Frequency to update the neighbor list and long range forces
cutoff-scheme = Verlet
ns_type = grid   ; Method to determine neighbor list (simple, grid)
coulombtype = PME ; Treatment of long range electrostatic interactions
rcoulomb = 1.0   ; Short-range electrostatic cut-off
rvdw = 1.0      ; Short-range Van der Waals cut-off
pbc = xyz       ; Periodic Boundary Conditions (yes/no)

```

### E.3 nvt.mdp

```

title = OPLS Protein equilibration
define = -DPOSRES ; position restrain the protein
; Run parameters
integrator = md   ; leap-frog integrator
nsteps = 50000   ; 2 * 50000 = 100 ps
dt = 0.002      ; 2 fs
; Output control
nstxout = 500    ; save coordinates every 1.0 ps
nstvout = 500    ; save velocities every 1.0 ps
nstenergy = 500  ; save energies every 1.0 ps
nstlog = 500     ; update log file every 1.0 ps
; Bond parameters
continuation = no ; first dynamics run
constraint_algorithm = lincs ; holonomic constraints
constraints = all-bonds ; all bonds (even heavy atom-H bonds) constrained
lincs_iter = 1    ; accuracy of LINCS
lincs_order = 4   ; also related to accuracy
; Neighborsearching
cutoff-scheme = Verlet ns_type = grid ; search neighboring grid cells
nstlist = 10     ; 20 fs, largely irrelevant with Verlet
rcoulomb = 1.0   ; short-range electrostatic cutoff (in nm)
rvdw = 1.0      ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics

```

```

pme_order = 4      ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT
; Temperature coupling is on
tcoupl = V-rescale ; modified Berendsen thermostat
tc-grps = Protein Non-Protein      ; two coupling groups - more accurate
tau_t = 0.1 0.1      ; time constant, in ps
ref_t = 300 300      ; reference temperature, one for each group, in K
; Pressure coupling is off
pcoupl = no      ; no pressure coupling in NVT
; Periodic boundary conditions
pbc = xyz      ; 3-D PBC
; Dispersion correction
DispCorr = EnerPres      ; account for cut-off vdW scheme
; Velocity generation
gen_vel = yes      ; assign velocities from Maxwell distribution
gen_temp = 300      ; temperature for Maxwell distribution
gen_seed = -1      ; generate a random seed

```

## E.4 npt.mdp

```

title = OPLS Protein NPT equilibration
define = -DPOSRES      ; position restrain the protein
; Run parameters
integrator = md      ; leap-frog integrator
nsteps = 50000      ; 2 * 50000 = 100 ps
dt = 0.002      ; 2 fs
; Output control
nstxout = 500      ; save coordinates every 1.0 ps
nstvout = 500      ; save velocities every 1.0 ps
nstenergy = 500      ; save energies every 1.0 ps
nstlog = 500      ; update log file every 1.0 ps
; Bond parameters
continuation = yes      ; Restarting after NVT
constraint_algorithm = lincs      ; holonomic constraints
constraints = all-bonds      ; all bonds (even heavy atom-H bonds) constrained
lincs_iter = 1      ; accuracy of LINCS
lincs_order = 4      ; also related to accuracy
; Neighborsearching

```

```

cutoff-scheme = Verlet
ns_type = grid      ; search neighboring grid cells
nstlist = 10       ; 20 fs, largely irrelevant with Verlet scheme
rcoulomb = 1.0     ; short-range electrostatic cutoff (in nm)
rvdw = 1.0        ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype = PME   ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4      ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT
; Temperature coupling is on
tcoupl = V-rescale ; modified Berendsen thermostat
tc-grps = Protein Non-Protein ; two coupling groups - more accurate
tau_t = 0.1 0.1    ; time constant, in ps
ref_t = 300 300    ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl = Parrinello-Rahman ; Pressure coupling on in NPT
pcoupltype = isotropic ; uniform scaling of box vectors
tau_p = 2.0       ; time constant, in ps
ref_p = 1000.0    ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility of water, bar^-1
refcoord_scaling = com
; Periodic boundary conditions
pbc = xyz ; 3-D PBC
; Dispersion correction
DispCorr = EnerPres ; account for cut-off vdW scheme
; Velocity generation
gen_vel = no      ; Velocity generation is off

```

## E.5 md.mdp

```

title = OPLS Protein MD simulation ; Run parameters
integrator = md ; leap-frog integrator
nsteps = 2500000 ; 2 * 2500000 = 5000 ps (5 ns)
dt = 0.002 ; 2 fs
; Output control
nstxout = 5000 ; save coordinates every 10.0 ps
nstvout = 5000 ; save velocities every 10.0 ps
nstenergy = 5000 ; save energies every 10.0 ps

```

```

nstlog = 5000      ; update log file every 10.0 ps
nstxout-compressed = 5000      ; save compressed coordinates every 10.0 ps
; nstxout-compressed replaces nstxtcout
compressed-x-grps = System      ; replaces xtc-grps
; Bond parameters
continuation = yes      ; Restarting after NPT
constraint_algorithm = lincs      ; holonomic constraints
constraints = all-bonds      ; all bonds (even heavy atom-H bonds) constrained
lincs_iter = 1      ; accuracy of LINCS
lincs_order = 4      ; also related to accuracy
; Neighborsearching
cutoff-scheme = Verlet
ns_type = grid      ; search neighboring grid cells
nstlist = 10 ; 20 fs, largely irrelevant with Verlet scheme
rcoulomb = 1.0      ; short-range electrostatic cutoff (in nm)
rvdw = 1.0      ; short-range van der Waals cutoff (in nm)
; Electrostatics coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4      ; cubic interpolation
fourierspacing = 0.16      ; grid spacing for FFT
; Temperature coupling is on
tcoupl = V-rescale      ; modified Berendsen thermostat
tc-grps = Protein Non-Protein      ; two coupling groups - more accurate
tau_t = 0.1 0.1      ; time constant, in ps
ref_t = 300 300      ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl = Parrinello-Rahman      ; Pressure coupling on in NPT
pcoupltype = isotropic      ; uniform scaling of box vectors
tau_p = 2.0      ; time constant, in ps
ref_p = 1000.0      ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility of water, bar^-1
; Periodic boundary conditions pbc = xyz      ; 3-D PBC
; Dispersion correction
DispCorr = EnerPres      ; account for cut-off vdW scheme
; Velocity generation
gen_vel = no      ; Velocity generation is off

```

## Referências

- 1 BRUCE, A.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. *Biologia Molecular da Célula*. Artmed Editora. 2009. Citado 27 vezes nas páginas 9, 10, 12, 15, 16, 17, 18, 19, 20, 21, 25, 26, 27, 28, 29, 31, 32, 33, 40, 41, 42, 43, 46, 47, 48, 49 e 57.
- 2 TOOZE, J.; BRANDEN, C. *Introduction to protein structure*. Second edition. Garland Science. Citado 17 vezes nas páginas 9, 10, 15, 18, 19, 20, 21, 26, 29, 31, 32, 35, 36, 37, 56, 57 e 94.
- 3 BETTELHEIM, F.; BROWN, W.H.; CAMPBELL, M.K.; FARRELL, S. *Introdução à bioquímica*. São Paulo. Cengage Learning. 2012 Citado 15 vezes nas páginas 9, 10, 15, 16, 18, 30, 36, 37, 38, 39, 44, 45, 46, 48 e 50.
- 4 VOLLHARDT, K.P.C.; SCORE, N.E. *Organic Chemistry*. Sixth edition. W. H. Freeman, 2010. Citado na página 16.
- 5 CHARNEY, D.S. *Neurobiology of Mental Illness*. Fourth edition. Cram101 Textbook Reviews. Citado na página 16.
- 6 ROGERO, M.M. ; TIRAPÉGUI, J. Aspectos atuais sobre aminoácidos de cadeia ramificada e exercício físico. *Revista Brasileira de Ciências Farmacêuticas*, 2008. Citado na página 16.
- 7 WU, G. ; BAZER, F. W. ; BURGHARDT, ROBERT, C. ; JOHNSON, G. A. ; KIM, S.W. ; KNABE, D.A. ; LI, P. ; LI, X. ; MCKNIGHT, J.R. ; SATTERFIELD, M. C. ; SPENCER, T. E. Proline and hydroxyproline metabolism: implications for animal and human nutrition. *NIH Public Access*, v.4, n.4, p.1053–1063, Apr, 2011. Citado na página 16.
- 8 LONGE, J.L. *The Gale Encyclopedia of Alternative Medicine*. Second edition. Gale, 2005. Citado na página 16.
- 9 LLMAS, E. *Biología para el examen de admisión*. Lulu.com, 2011. Citado na página 16.

- 10 AZZAM, M.M.M; ZOU, X.T ; DONG, X.Y. ; XIE, P. Effect of supplemental L-threonine on mucin 2 gene expression and intestine mucosal immune and digestive enzymes activities of laying hens in environments with high temperature and humidity . *Poultry Science*, vol.90, p.2251–2256, 2011. Citado na página 16.
- 11 NELSON, D. L ; COX, M. M. *Princípios de Bioquímica de Lehninger*. Sarvier, 2002 Citado 8 vezes nas páginas 12, 16, 19, 20, 21, 42, 43 e 57.
- 12 CABALLERO, B. ; FINGLAS, P. ; TOLDRA, F. *Encyclopedia of Food and Health*. Academic Press, 2005. Citado na página 16.
- 13 RUZZO, E. K. ; CHICI, J. M.C. ; ZEEV, B.B. ; CHITAYAT, D. ; MAO, H.; PAPPAS, A.L. ; HITOMI, Y. ; LU, Y. F. ; YAO, X. ; HAMDAN, F. ; PELAK ; WOLF, H. R. ; JOSEPH, I. B. ; LEVI, D. O. ; LEV, D. ; SAGIE, T. L. ; SILVER, E. L. ; ANIKSTER, Y. ; ASHER, E. B. ; OLENDER, T. ; COLLEAUX, L. ; DÉCARIE, J. C. ; BLASER, S. ; BANWELL, B. ; JOSHI, R. B ; HE, X. P. ; PATRY, L. ; SILVER, J. R. ; DOBRZENIECKA, S. ; ISLAM, M. S. ; HASNAT, A. ; SAMUELS, M.E. ; ARYAL, D.K. ; RODRIGUIZ, R.M. ; JIANG, Y. ; WETSEL, W. C. ; MCNAMARA, J. O. ; ROULEAU, G.A. ; SILVER, D. L. ; LANCET, D. ; PRAS, E. ; MITCHELL, G. A. ; MICHAUD, J. L ; GOLDSTEIN, D.B. Deficiency of asparagine synthetase causes congenital microcephaly and a progressive form of encephalopathy. *NIH Public Access*, v.80, n.2, 2013. Citado na página 16.
- 14 NOVAES, M. R. C. G. ; PANTALEÃO, C.M. Arginina: bioquímica, fisiologia e implicações terapêuticas em pacientes com câncer gastrointestinal. *Revista de Ciências Médicas*, Campinas, 2011. Citado na página 16.
- 15 GALE, T. ; LONGE, J.L. *The Gale Encyclopedia Of Cancer: A Guide To Cancer And Its Treatments*. Second edition. Gale, 2005. Citado na página 16.
- 16 LONGE, J. L. *The Gale Encyclopedia of Alternative Medicine*. Second edition. Gale, 2005. Citado na página 16.
- 17 INGLE, R.A. Histidine biosynthesis. *American Society of Plant Biologists*, v.9, Feb, 2011. Citado na página 16.
- 18 SANZ, J.J. ; NUNES, E. ; ZAFRA, F. ; BERROCAL, M. ; CORBACHO, I. ; IBANEZ, I. ; GONZALES, E. A. ; MARCOS, D. ; CORCUERA, B.L. ; MATA, A. M. ; ARAGÓN, C. Presynaptic Control of Glycine Transporter 2 (GlyT2) by Physical and Functional Association with Plasma Membrane Ca<sup>2+</sup> -ATPase (PMCA) and Na<sup>+</sup> -Ca<sup>2+</sup> Exchanger (NCX)\*. *Journal Of Biological Chemistry*, vol.289, n.49, p.34308 –34324, 2017. Citado na página 16.

- 19 CHEN, P.E. ; GEBALLE, M. T. ; STANSFELD, P.J. ; JOHNSTON, A. R. ; YUAN, H. ; JACOB, A. L. ; SNYDER, J.P. ; TRAYNELIS, S.F. ; WYLLIE, D.J.A . Structural Features of the Glutamate Binding Site in Recombinant NR1/NR2A N-Methyl-d-aspartate Receptors Determined by Site-Directed Mutagenesis and Molecular Modeling. *Molecular Pharmacology*, vol.67, n.5, p.1470-1484, May, 2005. Citado na página 16.
- 20 MCENTEE, W. J. ; CROOK, T. H. Glutamate: its role in learning, memory, and the aging brain. *Psychopharmacology*, v.111, n.4, p-391-401, 1993. Citado na página 16.
- 21 FEYNMAN, R. ; LEIGHTON, R.; SANDS, M. *Lições de Física*. Bookman, 2008. Volume 1. Citado 2 vezes nas páginas 16 e 103.
- 22 HOFFMAN, J. R ; FALVO, M. Protein - Which is best?. *Journal of Sports Science and Medicine*, v.3, n.3, p.118-130, 2004. Citado 2 vezes nas páginas 16 e 17.
- 23 MURPHY, S. ; ALLEN, L. Nutritional Importance of Animal Source Foods. *American Society for Nutritional Sciences*, vol.133, Nov, 2003. Citado na página 17.
- 24 WORLD HEALTH ORGANIZATION. Who/fao release independent expert report on diet and chronic disease. *Saudi Medical Journal*, v.24, n.10, 2003. Citado na página 17.
- 25 WAGNER, I. ; MUSSO, H. New Naturally Occurring Amino Acids. *Angewandte Chemie International Edition*, v.22, n.11, p.816–828, Nov, 1983. Citado na página 17.
- 26 DOS SANTOS, E.G. *Desenvolvimento de uma metodologia rápida e de baixo custo para diagnóstico da Anemia Falciforme*. Tese (Doutorado)- Instituto de Pesquisas Energéticas e Nucleares. São Paulo. 2014 Citado na página 19.
- 27 LIGAÇÃO PEPTÍDICA. Disponível em:  
<<https://image.slidesharecdn.com/aminoacidppt-170113172447/95/amino-acid-ppt-6-638.jpg?cb=1484328337>>. Acesso em: 15 de dez. 2017. Citado 2 vezes nas páginas 9 e 20.
- 28 LIFETEIN. Amino acids chart. 2017. Disponível em:  
<[https://www.lifetein.com/images/Knowledge\\_Center/Amino\\_acids\\_chart.png](https://www.lifetein.com/images/Knowledge_Center/Amino_acids_chart.png)>. Acesso em: 20 de dez. 2017. Citado 2 vezes nas páginas 9 e 22.
- 29 PETSKO, G.A.; RINGE, D. *Protein structure and function*. New Science Press, 2004. Citado 7 vezes nas páginas 9, 23, 24, 28, 31, 34 e 35.
- 30 CHAMPE, P.C.; HARVEY, R.; Bolner, A.R. *Bioquímica ilustrada*. Editora Artmed, 2002. Citado na página 37.
- 31 UENOJO, M. ; MARIA, G.P. Pectinases: aplicações industriais e perspectivas. *Quim. Nova*, v.30, n.2, 2007. Citado 2 vezes nas páginas 51 e 55.

- 32 SANTI, L. ; BERGER, M. ; ORLANDO, W. B. S. Pectinases e pectina: aplicação comercial e potencial biológico. *Revista Caderno Pedagógico*, v.11, n.1, 2014. Citado 3 vezes nas páginas 51, 52 e 55.
- 33 SO BIOLOGIA. Paredes primárias e secundárias. 2017. Disponível em: <[http://www.sobiologia.com.br/conteudos/Morfofisiologia\\_vegetal/morfovegetal15.php](http://www.sobiologia.com.br/conteudos/Morfofisiologia_vegetal/morfovegetal15.php)>. Acesso em: 15 de dez. 2017. Citado 2 vezes nas páginas 10 e 52.
- 34 FOOD INGREDIENTS BRASIL. Natureza e função das enzimas nos alimentos. *Revista-Fi*, n.35, 2015. Citado 2 vezes nas páginas 51 e 55.
- 35 COELHO, M.T. *Pectina: Características e Aplicações em Alimentos*. Monografia (Graduação)- Universidade Federal de Pelotas. Pelotas, 2008 Citado na página 51.
- 36 FOOD INGREDIENTS BRASIL. Pectinas propriedades e aplicações. *Revista-Fi*, n.35, 2015. Citado 3 vezes nas páginas 12, 52 e 53.
- 37 MENESES, A. G. *Produção de enzimas pectinolíticas e seletividade de produtos fitossanitários sobre o agente biológico "G088"*. Dissertação (Mestrado)- Universidade Federal de Lavras. Minas Gerais, 2007 Citado 2 vezes nas páginas 53 e 54.
- 38 WIKIPEDIA. Disponível em: <<https://en.wikipedia.org/wiki/Pectin>>. Acesso em: 7 de nov. 2017. Citado 2 vezes nas páginas 10 e 52.
- 39 ADAMU, M.S ; BELLO, U.M ; S, BILAL ; MASHI, A.H. Extraction and Characterization of Pectin from Orange Peels. *International Journal of Biotechnology and Biochemistry*, v.13, n.1, 2007. Citado na página 52.
- 40 CANTERI, M. H.G ; MORENO, L. ; WOSIACKI, G. ; SCHEER, A.P. Pectina: da Matéria prima ao produto final. *Polímeros*, v.22, n.2, 2012. Citado 2 vezes nas páginas 52 e 53.
- 41 SILVATEAM. Disponível em: <<https://www.silvateam.com/es/productos-y-servicios/aditivos-alimentarios/pectina/qu-es-la-pectina.html>>. Acesso em: 8 de dez. 2017. Citado 2 vezes nas páginas 10 e 53.
- 42 SANTI, L. *Produção, caracterização e aplicação de preparados pectinolíticos produzidos por penicillium oxalicum utilizando resíduos agroindustriais*. Dissertação (Mestrado)- Universidade Federal do Rio Grande do Sul. Porto Alegre, 2005 Citado 4 vezes nas páginas 12, 53, 54 e 55.
- 43 WINELAND. Enzymes in Winemaking. Disponível em: <[http://wineland.archive.shapeshift.co.za/archive/index.php?option=com\\_zine&view=article&id=479-in-winemaking](http://wineland.archive.shapeshift.co.za/archive/index.php?option=com_zine&view=article&id=479-in-winemaking)>. Acesso em: 04 de abr. 2017. Citado 2 vezes nas páginas 10 e 54.

- 44 MORGON, N.H.; JOHNSON, E. ; COUTINHO, K. *Métodos de química teórica e modelagem molecular*. Primeira edição. Livraria da física, 2007. Citado 16 vezes nas páginas 56, 57, 59, 60, 61, 64, 66, 67, 68, 69, 70, 73, 75, 77, 78 e 104.
- 45 PROTEIN DATA BANK. 1xg2-Sequence. Disponível em: <<http://www.rcsb.org/pdb/explore/remediatedSequence.do?structureId=1XG2>>. Acesso em: 04 de abr. 2017. Citado 4 vezes nas páginas 11, 56, 57 e 61.
- 46 RINO, J.P ; DA COSTA, B. V. *ABC da Simulação Computacional*. Primeira edição. Editora Livraria da Física, 2013. Citado 16 vezes nas páginas 11, 12, 57, 60, 61, 62, 63, 70, 71, 72, 73, 74, 75, 76, 77 e 103.
- 47 ROLDÃO, M.M.G. *Altas pressões hidrostáticas na conservação de sumo de maçã 100%*. Dissertação (Mestrado)- Universidade de Lisboa. Lisboa, 2014 Citado 4 vezes nas páginas 57, 58, 85 e 90.
- 48 DOS SANTOS, C. D. *Extração, clarificação e estabilização de betalaínas provenientes de talos de beterraba vermelha (Beta vulgaris L.)*. Tese (Doutorado)- Universidade Federal do Rio Grande do Sul. Porto Alegre, 2017 Citado 3 vezes nas páginas 57, 58 e 85.
- 49 BORTOLINI, G. *O oscilador harmônico e o gás de Lennard-Jones na dinâmica de Nosé-Hoover*. Dissertação (Mestrado)- Universidade do Estado de Santa Catarina. Joinville, 2015 Citado 3 vezes nas páginas 11, 61 e 63.
- 50 CARVALHO, A.G. *Dinâmica molecular aplicada à análise estrutural da Hipoxantina-Guanina-Xantina Fosforibosil transferase de Plasmodium Falciparum*. Dissertação (Mestrado)- Universidade Federal de Rondônia. Porto Velho, 2015 Citado 2 vezes nas páginas 11 e 66.
- 51 DIEHL, A. As bases da dinâmica molecular-1 . 2017. Disponível em: <[http://wp.ufpel.edu.br/diehl/files/2017/05/lec3\\_idmsf.pdf](http://wp.ufpel.edu.br/diehl/files/2017/05/lec3_idmsf.pdf)>. Acesso em: 27 de out. 2017. Citado na página 69.
- 52 DIEHL, A. As bases da dinâmica molecular -2 .2017. Disponível em: <[http://wp.ufpel.edu.br/diehl/files/2017/05/lec4\\_idmsf.pdf](http://wp.ufpel.edu.br/diehl/files/2017/05/lec4_idmsf.pdf)>. Acesso em: 27 de out. 2017. Citado 2 vezes nas páginas 69 e 70.
- 53 GROMACS. Reference manual. Disponível em: <<ftp://130.237.25.126/pub/manual/manual-5.1.4.pdf>>. Acesso em: 05 de abr. 2017. Citado 4 vezes nas páginas 69, 70, 78 e 79.
- 54 SCROFANO, R.; PRASANNA, V. K. *Computing Lennard-Jones Potentials and Forces with Reconfigurable Hardware*. 2004 Citado na página 71.

- 55 SALINAS, S. R. *Introdução à Física Estatística*. Segunda edição. Edusp - Editora da Universidade de São Paulo. 2013 Citado 9 vezes nas páginas 72, 73, 74, 75, 76, 77, 99, 100 e 101.
- 56 KUMARAN, V. Josiah Willard Gibbs. *Resonance*, v. 12, n.7, 2007 Citado 2 vezes nas páginas 72 e 73.
- 57 DE OLIVEIRA, M.J. Sistemas magnéticos. Disponível em: <[fig.if.usp.br/oliveira/term132ed.pdf](http://fig.if.usp.br/oliveira/term132ed.pdf)>. Acesso em: 04 de abr. 2017. Citado na página 73.
- 58 WIKIPEDIA. Molecular Dynamics. Disponível em: <[https://en.wikipedia.org/wiki/Molecular\\_dynamics](https://en.wikipedia.org/wiki/Molecular_dynamics)>. Acesso em: 27 de out. 2017. Citado na página 75.
- 59 SCHREIBER, F.; ZANINI, F.; RUNGE, F. R. *Virial expansion - A brief introduction*. 2011. Citado 4 vezes nas páginas 77, 102, 103 e 104.
- 60 TSAI, D. H. The virial theorem and stress calculation in molecular dynamics. *The Journal of Chemical Physics*, v. 70, n.3, 1979. Citado 4 vezes nas páginas 77, 102, 103 e 104.
- 61 MIGUEL, Enrique ; Jackson, George. The nature of the calculation of the pressure in molecular simulations of continuous models from volume perturbations . *The Journal of Chemical Physics*, v.125, n.16, 2006. Citado 2 vezes nas páginas 103 e 104.
- 62 PRIMAVERA, R.; BARBACANE, R.C.; CONGIA, M.; CELIA, C. Laser diffraction and light scattering techniques for the analysis of food matrices. *AFSH*, v.6, n.1, 2014. Citado 2 vezes nas páginas 11 e 79.
- 63 WIKIPEDIA. Espalhamento de raios-X a baixo ângulo. Disponível em: <<https://pt.wikipedia.org/wiki/Pectina>>. Acesso em: 04 de abr. 2017. Citado 2 vezes nas páginas 88 e 89.
- 64 BARBOSA, L.R.S. *Estudos de relevância biológica por espalhamento de raios-X a baixos ângulos*. Tese (Doutorado)- Universidade de São Paulo (USP), 2008. Citado 4 vezes nas páginas 11, 88, 89 e 90.
- 65 NUSSENZVEIG,H. *Curso de Física básica: Eletromagnetismo*. Segunda edição. Editora Blucher, v.3, 2015. Citado 4 vezes nas páginas 92, 93, 94 e 97.
- 66 WORD PRESS. Campo elétrico produzido por um dipolo.2017. Disponível em: <<https://campeletrico.wordpress.com/2014/02/14/campo-eletrico-produzido-por-um-dipolo-eletrico/>>. Acesso em: 15 de dez. 2017. Citado 2 vezes nas páginas 11 e 93.

- 67 ATKINS, P.; DE PAULA, J. *Physical chemistry for the life sciences*. USA. 2011. Citado 3 vezes nas páginas 11, 94 e 95.
- 68 FEYNMAN, R. *Lições de física - Eletricidade e Magnetismo*. Editora Bookman. 2008. Citado 2 vezes nas páginas 28 e 98.
- 69 NUSSENZVEIG, MOYSÉS. *Curso de Física Básica: Fluidos, Oscilações e Ondas, Calor*. Quinta edição. Editora Blucher, v.2, 2014. Citado 3 vezes nas páginas 99, 100 e 101.
- 70 LEMOS, A. N. *Mecânica Analítica*. Segunda edição. Edusp - Editora Livraria da Física, 2013 Citado na página 102.
- 71 HALLIDAY. Física 3. 2017. Disponível em:  
<[http://www.profanderson.net/files/problemasresolvidos/rhk4/c28/rhk4\\_c28\\_p011.pdf](http://www.profanderson.net/files/problemasresolvidos/rhk4/c28/rhk4_c28_p011.pdf)>. Acesso em: 15 de dez. 2017. Citado 2 vezes nas páginas 11 e 96.