

Guilherme Kilpp Gonzatti

# **Investigação Computacional da Vitrificação da Água de Solvatação do DNA**

Porto Alegre, Brasil

29 de Maio de 2015

Guilherme Kilpp Gonzatti

# **Investigação Computacional da Vitrificação da Água de Solvatação do DNA**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Química

Universidade Federal do Rio Grande do Sul

Instituto de Química

Programa de Pós-Graduação em Química – PPGQ

Orientador: Prof. Dr. Paulo Augusto Netz

Porto Alegre, Brasil

29 de Maio de 2015

*Dedico este trabalho a meus pais,  
que sempre me incentivaram a perseguir meus sonhos.*

## **AGRADECIMENTOS**

Um agradecimento especial ao MSc. William Kelbert Nitschke (`william`) pelas discussões cafeinadas, o pelo apoio técnico e, principalmente, emocional;

Aos colegas do Grupo de Química Teórica, GQT, em especial, `petry`, `jotoldo`, `sandro`, `helen`, `paulo`, `livotto` e `msegala`. Sua companhia, bom humor, e discussões, acadêmicas ou não, foram essenciais;

*"Don't get set into one form, adapt it and build your own, and let it grow, be like water.  
Empty your mind, be formless, shapeless — like water.  
Now, you put water in a cup, it becomes the cup;  
You put water into a bottle it becomes the bottle;  
You put it in a teapot, it becomes the teapot.  
Now, water can flow or it can crash.  
Be water, my friend."*

Bruce Lee

## RESUMO

Motivado por evidências experimentais, foram realizadas extensivas simulações clássicas de dinâmica molecular de soluções aquosas de DNA visando investigar a transição vítrea. Estudou-se o dodecâmero de Dickerson solvatado por moléculas de água do modelo TIP5P e TIP3P em temperaturas de 300 à 240 K e de 300 à 200 K, respectivamente. Análise de propriedades dinâmicas e estruturais mostram que o modelo TIP3P não é capaz de transição vítrea, mas que sistemas compostos pelo modelo TIP5P é capaz de apresentar a transição vítrea dinâmica.

## ABSTRACT

Motivated by experimental evidences, extensive classical molecular dynamics were performed to investigate the glass transition of aqueous solutions of DNA. The Dickerson dodecamer solvated by water molecules of models TIP3P and TIP5P were studied in the range of 300 to 240 K and 300 to 200 K, respectively. Analysis of dynamical and structural properties reveal that the model TIP3P is unable of performing glass transition, but systems with TIP5P can present dynamic glass transition.

# Lista de ilustrações

Figura 1 – Bases nitrogenadas. . . . .	18
Figura 2 – Pareamento dos pares de bases canônico no DNA. . . . .	18
Figura 3 – A-, B- e Z-DNA. . . . .	19
Figura 4 – Espinha de hidratação. . . . .	20
Figura 5 – Representação do modelo de água TIP5P . . . . .	23
Figura 6 – Parâmetros estruturais do DNA. . . . .	25
Figura 7 – Identificação dos átomos que definem os sulcos do DNA. . . . .	36
Figura 8 – Coeficiente de autodifusão em função da temperatura para os dois modelos de água. . . . .	40
Figura 9 – Água <i>bulk</i> próxima ao DNA. . . . .	41
Figura 10 – Espinha de hidratação do DNA . . . . .	42
Figura 11 – Evolução temporal da temperatura no <i>ensemble</i> NVE. . . . .	43
Figura 12 – Função de distribuição radial para água pura. . . . .	44
Figura 13 – Fator de estrutura translacional para água pura. . . . .	45
Figura 14 – Função de distribuição radial de água <i>bulk</i> do sistema com água <i>tip3p</i> . . . . .	46
Figura 15 – Função de distribuição radial de água <i>bulk</i> dos sistemas com água <i>tip5p</i> . . . . .	46
Figura 16 – Função de distribuição radial da espinha de hidratação do sistema TIP3P-DNA. . . . .	47
Figura 17 – Função de distribuição radial da espinha de hidratação dos sistemas com água <i>tip5p</i> . . . . .	48
Figura 18 – Fator de estrutura translacional para todos os sistemas. . . . .	49
Figura 19 – Densidade dos sistemas. . . . .	50
Figura 20 – Função de autocorrelação de velocidades para água <i>bulk</i> do modelo <i>tip3p</i> . . . . .	51
Figura 21 – Função de autocorrelação de velocidades para água nos sulcos do DNA do sistema TIP3P-DNA . . . . .	51
Figura 22 – Função de autocorrelação de velocidades para água <i>bulk</i> do modelo <i>tip5p</i> . . . . .	52
Figura 23 – Função de autocorrelação de velocidades para água nos sulcos do DNA do sistemas com água <i>tip5p</i> . . . . .	52
Figura 24 – Transformada de Fourier da VACF de água <i>tip3p</i> no <i>bulk</i> dos sistemas. . . . .	53
Figura 25 – Transformada de Fourier da VACF de água <i>tip5p</i> no <i>bulk</i> dos sistemas. . . . .	53
Figura 26 – Transformada de Fourier da VACF de água nos sulcos do DNA no sistema TIP3P-DNA. . . . .	54
Figura 27 – Transformada de Fourier da VACF de água nos sulcos do DNA no sistema TIP5P-FISIO e TIP5P-DNA. . . . .	54
Figura 28 – Coeficiente de autodifusão para o sistema TIP3P-DNA e TIP3P-BULK. . . . .	55
Figura 29 – Coeficiente de autodifusão para os sistemas com água <i>tip5p</i> . . . . .	56

Figura 30 – Diagrama de Arrhenius do coeficiente de autodifusão dos sistemas. . . . .	57
Figura 31 – Ângulo das ligações de hidrogênio para <i>bulk</i> com água <i>tip3p</i> . . . . .	58
Figura 32 – Ângulo das ligações de hidrogênio na espinha de hidratação do DNA no sistema TIP3P-DNA. . . . .	59
Figura 33 – Ângulo das ligações de hidrogênio para <i>bulk</i> com água <i>tip5p</i> . . . . .	60
Figura 34 – Ângulo das ligações de hidrogênio na espinha de hidratação do DNA nos sistemas TIP5P-FISIO e TIP5P-DNA. . . . .	61
Figura 35 – Função de autocorrelação contínua das ligações de hidrogênio de águas <i>bulk</i> nos sistemas com água <i>tip3p</i> . . . . .	61
Figura 36 – Função de autocorrelação contínua das ligações de hidrogênio na espinha de hidratação no sistema TIP3P-DNA. . . . .	62
Figura 37 – Função de autocorrelação contínua das ligações de hidrogênio de águas <i>bulk</i> nos sistemas com água <i>tip5p</i> . . . . .	62
Figura 38 – Função de autocorrelação contínua das ligações de hidrogênio na espinha de hidratação nos sistemas TIP5PP-FISIO e TIP5P-DNA. . . . .	63
Figura 39 – Tempo de vida das ligações de hidrogênio em função da temperatura. . . . .	63
Figura 40 – Arrhenius <i>plot</i> to tempo de vida das ligações de hidrogênio. . . . .	64
Figura 41 – Análises comparativa do sistema TIP5P-DNA sob temperatura de 240K. . . . .	66

# Lista de tabelas

Tabela 1 – Tempo de vida das ligações de hidrogênio, $\tau_{HB}$ , em ps, para diferentes configurações iniciais para a água <i>bulk</i> e nos sulcos do DNA. . . . .	65
---	----

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>12</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>14</b>
<b>2.1</b>	<b>Água</b>	<b>14</b>
2.1.1	Água Super-resfriada	14
<b>2.2</b>	<b>Materiais Vítreos e Transição Vítreas</b>	<b>15</b>
2.2.1	Poliamorfismo da Água	15
2.2.2	Transição Dinâmica	16
<b>2.3</b>	<b>DNA</b>	<b>17</b>
2.3.1	Água de solvatação do DNA	19
<b>2.4</b>	<b>Dinâmica Molecular</b>	<b>20</b>
2.4.1	Campo de Forças	21
2.4.1.1	Campo de Forças <i>parmbsc0</i>	22
2.4.1.2	Modelos de Águas	23
2.4.2	Temperatura e Pressão	23
2.4.2.1	Termostato de Berendsen	24
2.4.2.2	Barostato de Parrinello-Rahman	24
2.4.3	Análises	25
2.4.3.1	Estrutura do DNA	25
2.4.3.2	Desvio Quadrático Médio	25
2.4.3.3	Distribuição Radial	26
2.4.3.4	Auto-Correlação de Velocidades	27
2.4.3.5	Ligações de Hidrogênio	27
2.4.3.5.1	Intermitente	28
2.4.3.5.2	Contínua	28
2.4.3.6	Coefficiente de Difusão	28
2.4.3.7	Fator de Estrutura Translacional	29
<b>3</b>	<b>OBJETIVOS</b>	<b>31</b>
<b>3.1</b>	<b>Objetivos Gerais</b>	<b>31</b>
<b>3.2</b>	<b>Objetivos Específicos</b>	<b>31</b>
<b>4</b>	<b>METODOLOGIA</b>	<b>32</b>
<b>4.1</b>	<b>Dinâmica Molecular</b>	<b>32</b>

4.1.1	Caixas de Simulação . . . . .	32
4.1.2	Simulações <i>NpT</i> . . . . .	33
4.1.3	Simulações <i>NVE</i> . . . . .	34
4.1.4	Análises . . . . .	35
4.1.4.1	Análise Estrutural do DNA . . . . .	35
4.1.4.2	Água nos Sulcos e Água . . . . .	35
4.1.4.2.1	Água <i>bulk</i> . . . . .	36
4.1.4.2.2	Água nos Sulcos e Água <i>Bulk</i> . . . . .	36
4.1.4.3	Análises com GROMACS . . . . .	36
4.1.4.3.1	Análises RMSD . . . . .	37
4.1.4.3.2	Análises de Ligação de Hidrogênio . . . . .	37
4.1.4.3.3	Coefficiente de Autodifusão da Água . . . . .	37
4.1.4.3.4	Fator de Estrutura Translacional . . . . .	38
<b>5</b>	<b>RESULTADOS . . . . .</b>	<b>39</b>
<b>5.1</b>	<b>Inspeção Visual . . . . .</b>	<b>40</b>
<b>5.2</b>	<b>Temperatura . . . . .</b>	<b>41</b>
<b>5.3</b>	<b>Estrutura . . . . .</b>	<b>43</b>
5.3.1	Água <i>Bulk</i> . . . . .	43
5.3.2	Espinha de Hidratação . . . . .	47
<b>5.4</b>	<b>Funções de autocorrelação de velocidades . . . . .</b>	<b>50</b>
<b>5.5</b>	<b>Autodifusão . . . . .</b>	<b>55</b>
<b>5.6</b>	<b>Análise de Ligações de Hidrogênio . . . . .</b>	<b>57</b>
5.6.1	Análise do Ângulo Doador-Hidrogênio-Aceptor . . . . .	57
5.6.2	Análise Contínua . . . . .	59
5.6.3	Análise Intermitente . . . . .	59
5.6.4	Influência da Configuração Inicial . . . . .	64
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>68</b>
<b>7</b>	<b>PERSPECTIVAS . . . . .</b>	<b>69</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>70</b>

# 1 Introdução

A água é, sem dúvida, o líquido mais importante na Terra. O fascínio exercido por esta substância sobre o homem faz que tenha sido a mais estudada, desde os gregos, como um dos quatro elementos básicos e, de forma científica há pelo menos 300 anos. (1), com *review* científico datando de 1773. (2) Apesar disso, pode-se dizer que seja talvez um dos líquidos menos compreendidos. Embora a estrutura química de uma molécula de água,  $H_2O$ , seja muito simples, suas fortes e altamente direcionadas interações intermoleculares aliadas a seu baixo peso molecular (e baixos momentos de inércia) fazem com que a matéria condensada resultante apresente propriedades físico-químicas bastante complexas. (3) Como exemplo comum deste comportamento, pode-se citar a expansão da água quando resfriada abaixo de  $4^\circ C$ , sob pressão atmosférica.

O comportamento da água é muitas vezes classificado como anômalo, uma vez que é, por muitas vezes, contra-intuitivo ao que se poderia esperar em uma primeira observação e também por ser diferente da maiorias das outras substâncias. Estes outros fluidos, ditos normais, se comportam como seria esperado, embora muitas outras substâncias também apresentem anomalias similares às da água, as ditas *waterlike anomalies*(4), como sílica, óxido de germânio e outros. (5, 6)

Atualmente, o número de ditas anomalias da água é de 71, classificadas entre dinâmicas, estruturais e termodinâmicas. Como exemplos, podemos citar o simples fato de que, nas condições de pressão atmosférica e temperatura deste planeta podemos encontrar a água no estado líquido, sólido e gasoso; sua alta capacidade calorífica; o aumento de sua mobilidade (expressa pelo coeficiente de autodifusão) sob compressão; o fato de se tornar um líquido menos estruturado sob compressão; e o fato de apresentar diversas formas sólidas cristalinas distintas. (8) Curiosamente, também, as anomalias apresentam uma hierarquia bem definida, isto é, em um diagrama de fases de pressão por temperatura, a região de anomalias estruturais engloba a região de anomalias dinâmicas que, por sua vez, engloba a região de anomalias termodinâmicas.(9)

Interessantemente, a água pode solidificar sem adotar uma forma cristalina, formando assim um sólido amorfo, ou seja, a água apresenta transição vítrea. Mais interessante, no entanto, é o fato de que pode adotar não apenas uma forma amorfa metaestável, a água apresenta poliamorfismo. O estudo destas formas vítreas da água é relativamente recente e apresenta grandes desafios para a pesquisadores, tanto experimentalmente quanto em tratamento teórico e computacional. A compreensão da água vítrea (também denominada vítrosa) é de grande relevância não apenas do ponto de vista puramente científico, como também em campos como biologia e em aplicações tecnológicas. (1, 10)

Recentemente, constatou-se experimentalmente (11) que as moléculas água de solvatação de amostras de DNA resfriado se encontram no estado vítreo, com contribuições de duas formas amorfas, uma de baixa densidade, denominada LDA (do inglês, *Low-Density Amorphous water*), e uma de alta densidade, denominada HDA (do inglês, *High-Density Amorphous water*).

## 2 Fundamentação Teórica

### 2.1 Água

O estudo da água por técnicas computacionais foi instigado pelo grande desafio de compreender o comportamento anômalo da água, fazendo com que com o passar do anos diversas abordagens computacionais fossem desenvolvidas. Assim, atualmente, existe uma vasta diversidade de diferentes modelos de água, desde modelos mais simples à modelos mais realistas, além de fluidos modelo, que não podem verdadeiramente ser classificados como água, mas que são amplamente empregadas no estudo do comportamento anômalo. (12)

A necessidade da água para a existência de vida é refletida também nos estudos computacionais de biomoléculas. No início dos estudos computacionais de biomoléculas não havia recursos computacionais para o uso de solventes explícitos. No entanto, os resultados obtidos sem o uso de solvente em muito divergiam do verificado experimentalmente. Solventes implícitos, ou contínuos, onde o efeito estabilizante do solvente é aproximado como um meio dielétrico, também representavam pobremente a realidade, e certos fenômenos como o enovelamento de proteínas não podiam ser descritos.

Então para o estudo de sistemas de interesse biológico, uma descrição adequada do meio também deve ser realizada. Dadas as dimensões das biomoléculas e a quantidade necessária de águas para solvatá-las completamente, a maior parte do esforço computacional é realizado para o cálculo das trajetória do solvente. Com o avanço dos recursos computacionais, estudos por dinâmica molecular fazem uso de solventes explícitos rotineiramente. (13)

#### 2.1.1 Água Super-resfriada

Muitas das propriedades anômalas da água se tornam mais evidentes no regime super-resfriado, i.e., quando a água líquida é resfriada abaixo da sua temperatura de congelamento sem sofrer uma transição de fase, formando gelo. É, portanto, uma situação metaestável do ponto de vista termodinâmico, o que impõe algumas dificuldades experimentais. (6) Do ponto de vista experimental, há limitações nas taxas máximas de compressão ( $6000 \text{MPa} \cdot \text{min}^{-1}$ ) e resfriamento ( $-10^6 \text{K} \cdot \text{s}^{-1}$ ) que podem ser obtidas. (1)

Talvez a maior limitação em se estudar a água super-resfriada seja a nucleação espontânea, de forma que há regiões inacessíveis experimentalmente. O uso de simulações computacionais oferece uma alternativa para o estudo do regime super-resfriado.

## 2.2 Materiais Vítreos e Transição Vítreo

### 2.2.1 Poliamorfismo da Água

Dependendo da forma com que o resfriamento da água é realizado, pode-se evitar a formação de sólidos cristalinos (fases termodinamicamente estáveis) e se obter uma estrutura não-cristalina, amorfa, denominada vítrea. (14) Hoje em dia reconhece-se que a água apresenta poliamorfismo, isto é, apresenta mais do que uma forma amorfa, com propriedades distintas. Há o LDA (do inglês, *Low Density Amorphous Water*) e o HDA (*High Density Amorphous Water*), que diferem entre si em densidade e espectro vibracional, por exemplo. Mais recente foi também verificado uma forma de água vítrea cujas propriedades diferem do HDA e do LDA. Devido à sua alta densidade foi denominado vHDA.

O LDA foi identificado em 1934 como *Amorphous Solid Water*, obtido pela primeira vez por meio da deposição de vapor de água em placas frias, à 90 K. (15) O LDA pode também ser obtido através de *quenching* rápido, empregando uma fase fluida apolar em baixa temperatura, como demonstrado em 1985. O sólido assim obtido foi originalmente denominado *Hyper Quenched Glass*. (16) Estes dois sólidos amorfos quando aquecidos entre 77 e 113 Kelvin relaxam para uma mesma estrutura amorfa (ainda metaestável), de densidade relativa 0,94, conforme identificado por experimentos de calorimetria, difração de raios-X, espectroscopia Raman e espalhamento de nêutrons. (17) Salienta-se que esta forma de água é a mais abundante no universo. (2) O LDA apresenta, além de estrutura tridimensional amorfa, características vibracionais cooperativas e algumas propriedades termodinâmicas que se assemelham à sólidos cristalino, o que leva alguns autores a dizer que este sólido se assemelha ao que seria um “estado vítreo ideal”. (2)

O HDA, foi obtido experimentalmente pela primeira vez em 1984. No experimento realizado, aplicou-se elevadas pressões (10 kbar) sobre o gelo cristalino resfriado (77 K) e verificou-se a formação de um sólido distinto cuja densidade é 1,31. (18) Este sólido foi caracterizado e é distinto do LDA. (6)

Sugere-se que a presença de duas formas amorfas distintas está ligada de forma íntima às demais anomalias apresentadas pela água. Três principais hipóteses são colocadas: a hipótese do limite da estabilidade, de Speedy (19); a hipótese de que não há singularidades das funções dos potenciais termodinâmicos, (20, 21); a hipótese de que, no estado super-resfriado há uma transição líquido-líquido. (21, 22). As duas fases líquidas teriam densidades distintas, e cada uma delas estaria relacionada à uma das fases sólidas amorfas.

Algumas questões ainda não são bem resolvidas sobre as formas amorfas da água. Como, por exemplo, qual seria a temperatura de transição vítrea e se, mais fundamentalmente, se grandeza é de fato acessível experimentalmente; Se de fato o HDA e o vHDA são formas distintas ou se o vHDA seria uma forma relaxada do HDA. Questões como estas podem estar intimamente ligadas ao comportamento anômalo da água. Uma das possíveis

explicações seria uma transição líquido-líquido, um de baixa e outro de alta densidade, que seriam as formas líquidas responsáveis pelo LDA e HDA, respectivamente.

O uso de ferramentas de dinâmica molecular permite estudar com relativa facilidade o estado super-resfriado, e pode-se identificar a transição líquido-líquido, bem como o seu referente ponto crítico. (23) Experimentos recentes em microporos de xerogéis de sílica permitem estudar água confinada em dimensões nanométricas. Nestas condições a nucleação da água pode ser postergada. Análise dos estados vibracionais sugerem uma transição de fase líquido-líquido na temperatura de 225 K. (24)

A análise deste tipo de evidencia deve ser feita com cuidado. Experimentos (e até mesmo simulações) em regimes tão distantes da situação de equilíbrio podem atingir um ambiente “não-termodinâmico”, como questionado por Kiselev (25, 26, 27, 28)

De qualquer forma, as formas amorfas da água (e de soluções aquosas) apresentam grande importância biológica e aplicação tecnológica. Como exemplos do primeiro pode-se mencionar a formação de água vítrea em árvores e insetos para a preservação em ambientes de baixa temperatura, ou ainda a produção de proteínas anticongelantes. Como exemplo de aplicação tecnológica é a crio-preservação, pois um dos grandes desafios nesta área é evitar a formação de cristais de gelo, que acabam por danificar estruturas celulares. (1)

Desde 1980 a preservação de órgão para transplantes é de em torno de 12 horas. Períodos superiores a este diminuem drasticamente as chances de o transplante ocorrer com sucesso. Naturalmente, esta limitação limita grandemente as possibilidades de realizar transplantes quando o doador e o paciente alvo do transplante estão em distâncias muito grandes. O uso de anti-congelantes atóxicos, como polímeros similares ao polietilenoglicol, porém derivados de glicose, pode aumentar este período. Já foi possível, desta forma, manter o fígado de um rato “vivo” por 3 meses. Recentemente, a realização de um transplante com um órgão assim preservado ocorreu com sucesso mesmo após 4 dias da remoção do órgão. (29)

Diversos estudos recentes também buscam compreender a ação de proteínas anti-congelamento. Estas glicoproteínas podem afetar a estrutura do solvente mesmo a longas distâncias (de 10 à 12 ângstrons), impedindo transição para a fase cristalina. (30)

### 2.2.2 Transição Dinâmica

Certos fluidos tem suas propriedades dinâmicas coletivas, como viscosidade e tempo de relaxação característico influenciada pela temperatura em um comportamento dito Arrhenius-*like*:

$$\tau(T) = \tau_0 \exp\left(\frac{A}{T}\right), \quad (2.1)$$

onde  $\tau$  é propriedades em questão,  $\tau_0$  é o limite desta propriedade para altas temperaturas. Outros materiais, por outro lado, seguem uma dependência denominada Vogel-Fulcher-Tammann-Hesse (ou, simplesmente, VF):

$$\tau(T) = \tau_0 \exp\left(\frac{B}{T - T_0}\right), \quad (2.2)$$

onde  $B$  e  $T_0$  são parâmetros que dependem da natureza do material.  $B$  pode ser interpretado como uma constante que mede a fragilidade do fluido.  $T_0$  é a temperatura na qual ocorre a divergência do parâmetro  $\tau$  e pode ser denominada a temperatura ideal de transição vítrea. Na prática, porém, devido à própria natureza de a transição vítrea não ser um fenômeno cinético (cineticamente frustrado), este parâmetro depende de diversos fatores experimentais, como temperaturas na quais a análise é realizada, como um primeiro exemplo.

Existem diversos processos envolvidos no rearranjo estrutural um sistema. Denomina-se  $\alpha$  os processos cooperativos globais. Por serem processos que dependem de uma elevada cooperação das espécies, são processos mais lentos; Os processos  $\beta$ , por outro lado, envolvem um número muito menor de espécies, ocorrendo em uma região microscópica do espaço, como por exemplo o movimento de *cage rattling*, no qual uma molécula se move dentro de sua esfera (ou “gaiola”) de solvatação.(31)

O resfriamento de um fluido já superresfriado leva à separação das escalas de tempo com quem ocorrem processos envolvidos no rearranjo global do sistema e temos uma separação entre os processos  $\alpha$  e  $\beta$ . Temos então a denominada transição vítrea dinâmica, que é caracterizada pela mudança de um comportamento do tipo Arrhenius, como descrito pela Equação 2.1 para um comportamento do tipo VF, descrito por (2.2). (32, 31, 33)

## 2.3 DNA

O ácido desóxi-ribonucleico, DNA (da sigla em inglês), é responsável pelo armazenamento da informação genética, estando presente em todas as formas vida conhecidas e em muitos vírus. São biopolímeros constituídos por sequências de nucleotídeos. Cada unidade de nucleotídeo é formada por uma unidade de fosfato, um açúcar ( $\beta$ -2'-desoxi-D-ribose) e uma base nitrogenada. Esta base nitrogenada pode ser derivada da purina, como é o caso da guanina (G) e adenina (A), ou da pirimidina, como a timina (T) e citosina (C), (34) como mostra a Figura 1.

Em sistemas biológicos o DNA se encontra usualmente sob a forma de dupla fita, isto é, formado por duas fitas orientadas antiparalelas. Esta conformação é resultante do pareamento das bases nitrogenadas. (36) e das interações  $\pi$  *stacking* entre os pares de bases. Quando em solução aquosa, os pares de bases minimizam suas interações hidrofóbicas

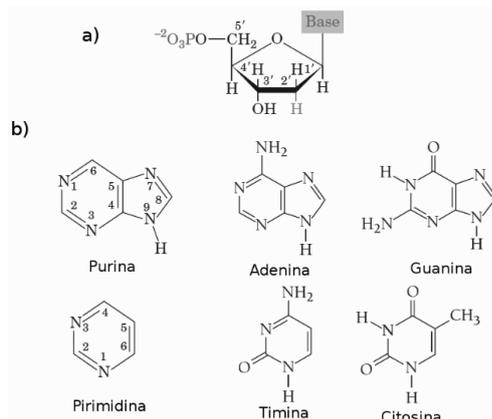


Figura 1 – a) Estrutura geral de um desoxirribonucleotídeo. b) Estrutura das bases nitrogenadas e dos compostos dos quais derivam. Mostra-se a convenção da numeração dos átomos. Adaptado de (35)

fazendo ligações de hidrogênio entre si, como mostra a Figura 2, a passo que o “esqueleto” de açúcar-fosfato fica exposto, interagindo com as moléculas de água.

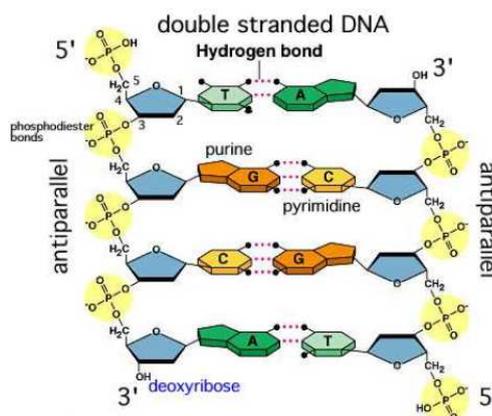


Figura 2 – Pareamento dos pares de bases canônico no DNA. A timina faz duas ligações de hidrogênio com adenina e guanina faz três ligações de hidrogênio com citosina. Adaptado de (35)

O DNA naturalmente ocorrente é formado por uma grande quantidade de pares de bases, 192 milhões de pares em cromossomos de ratos, por exemplo, que, se estendida, mediria mais do que 6 cm (37). A presença do DNA no interior do núcleo celular se dá por meio da formação de supraestruturas.

A estrutura tridimensional de uma fita de DNA é helicoidal e há a formação de dois sulcos de larguras diferentes, o sulco maior e o menor, mostrados Figura 4. Esta estrutura do DNA, denominada B-DNA, é que ocorre em maior frequência nos seres vivos, embora outras formas conformações também possam ser adotadas, como o A-DNA e o Z-DNA, ilustradas na Figura 3. (34) Vários fatores influenciam na estabilidade relativa destas diferentes formas, como a concentração e a natureza de solutos no meio, bem como

a sequência dos pares de base. Outro fator que apresenta grande influência é a quantidade de água disponível para solvatar e estabilizar o DNA, de forma que quando há pouca água ocorre preferência pela forma A e, com maior quantidade de água, pela forma B.

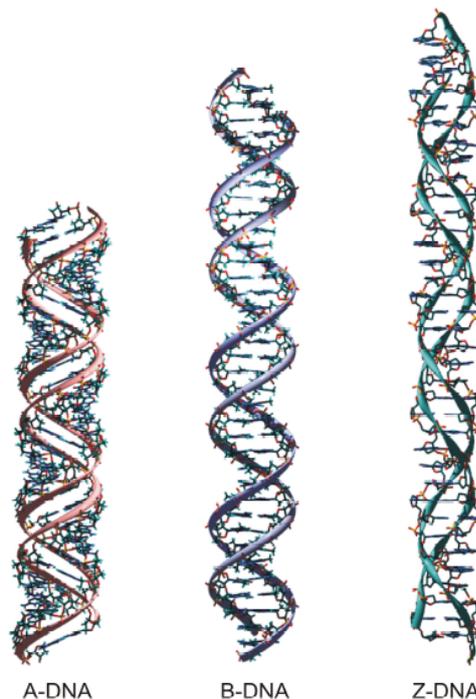


Figura 3 – Estrutura das fitas de DNA, nas formas A-, B- e Z-. Cada estrutura é composta por 36 pares de bases. Adaptado de (38)

### 2.3.1 Água de solvatação do DNA

A atividade da água é um dos fatores mais importantes para a estabilidade conformacional do DNA. Além disso, pode desempenhar um papel significativo mediando a interação específica dos ácidos nucleicos com proteínas e outros ligantes dependentes de conformações e sequências específicas. (39)

As moléculas de água que solvata o DNA são sujeitas à um ambiente químico diferente daquele experimentado no *bulk* do solvente,, como ilustrado na Figura 4. Consequentemente, as propriedades físico-químicas destas molécula de água serão alteradas. Por exemplo, a água de hidratação apresenta mobilidade muito mais baixa do que no *bulk* (40).

Uma interessante consequência desta interação é que com o resfriamento (abaixo de, aproximadamente 247 Kelvin), ocorre uma transição dinâmica macromolecular, caracterizada pelo uma diminuição dramática das flutuações quadráticas médias do DNA, sendo uma transição tipo vítrea. (41) Além de afetar o DNA, também é notada uma mudança significativa nas propriedades dinâmicas, como tempo de relaxação translacional

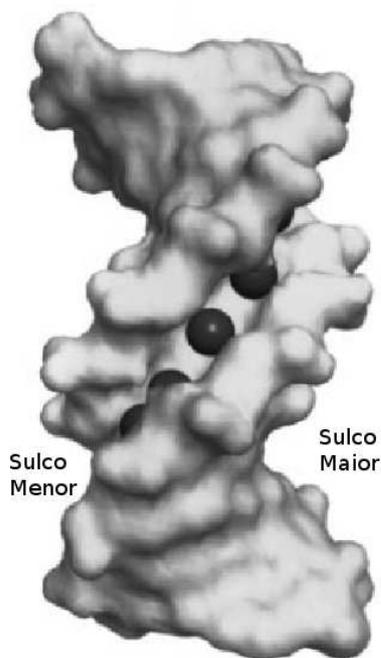


Figura 4 – Estrutura tridimensional do DNA mostrando moléculas de água formando a “espinha de hidratação”. Devido à forte interação com o DNA estas moléculas tem mobilidade muito menor do que no *bulk*. Também são identificados os sulcos maior e menor. Adaptado de 40

e rotacional, bem como estruturais, como também sobre a distribuição de ângulo O-O-O (entre átomos de oxigênio de três moléculas próximas). Este último parâmetro permite inferir sobre o número de vizinhos que uma molécula central têm em sua primeira esfera de solvatação. Por fim, esta influência sobre o comportamento da água é diferente conforme o ambiente em que se encontra: moléculas no sulco menor são mais influenciadas do que as no sulco maior e moléculas no *bulk* são pouco afetadas.

Em um recente experimento de espalhamento de nêutrons, foi publicado uma observação de que, em baixas temperaturas (100 K), moléculas de água envolvidas nas primeiras camadas de solvatação de DNA (na forma B-DNA) se encontram em forma amorfa, principalmente LDA e, em menor escala por uma contribuição de HDA (coexistindo também com uma fase cristalina), demonstrado pelo fator de estrutura estático. (11)

## 2.4 Dinâmica Molecular

Simulações de MD são técnicas empregadas para calcular propriedades de equilíbrio e de transporte de sistemas clássicos de muitos corpos, cujas equações de movimento não possuem solução analítica. Diz-se clássicos pois considera-se que o movimento nuclear das partículas que constituem o sistema são descritas pelas leis clássicas da mecânica, isto é, são descritas pela física newtoniana. Para o estudo de uma vasta gama de sistemas de

interesse esta é uma excelente aproximação. (42, 13)

Devido às interações com o sistema, um átomo  $i$  é sujeito à uma força,  $\mathbf{F}$ , que depende de sua posição, que por sua vez varia no tempo,  $\mathbf{r}(\mathbf{t})$ , Na notação de Newton:

$$\mathbf{F}_i = \frac{d\mathbf{p}_i}{dt} = -\nabla_{\mathbf{r}_i} V(\mathbf{r}) \quad (2.3)$$

onde  $\mathbf{p}_i$  é o *momentum* de  $i$  e  $\nabla_{\mathbf{r}_i} V(\mathbf{r})$  é o gradiente da energia potencial de  $i$ .

A evolução temporal do sistema é determinada numericamente resolvendo a Equação 2.3 considerando a passagem de um pequeno intervalo de tempo,  $\Delta t$ . As posições dos átomos são atualizadas, e o processo é repetido sucessivamente.

Empregando ferramentas da termodinâmica e da mecânica estatísticas pode-se então extrair propriedades estruturais e dinâmicas do sistema a partir das trajetórias atômicas.

### 2.4.1 Campo de Forças

Não há uma solução analítica para a superfície potencial,  $V(\mathbf{r})$ , e ela deve ser aproximada. A aproximação mais empregada é a aditividade do potencial aos pares:

$$V(\mathbf{r}) = \sum_{i=1}^{N-1} \sum_{j>i}^N V_{ij}(\mathbf{r}) \quad (2.4)$$

onde  $N$  é o número total de átomos. Considerando que os termos energéticos são aditivos a energia pode ser decomposta em componentes que contribuem para a energia total:

$$V(\mathbf{r}) = V_l + V_\theta + V_\phi + V_{vdW} + V_{Coul} \quad (2.5)$$

Onde o termo  $V_l$  é o potencial desenvolvido devido ao estiramento de uma ligação química, conforme os átomos envolvidos saem da posição de energia mínima;  $V_\theta$  corresponde às deformações angulares;  $V_\phi$  é potencial associado à deformação dos ângulo diedros próprios. A repulsão de van der Waals é representada pelo termo  $V_{vdW}$ , esta interação descreve a interação intermolecular de átomos ou sítios interagentes. Por fim, o termo  $V_{Coul}$  representa as interações eletrostáticas, descrito pelo potencial de Coulomb.

Então, para um dado sistema em uma determina configuração atômica, para a avaliação do potencial assumido deve-se definir quais os valores dos parâmetros mencionados. Estes valores são parametrizados e o seu conjunto (bem como as funções de potencial) formam o campo de forças.

Existem diversos campos de forças distintos, cada qual parametrizado de uma forma diferente. Conseqüentemente, cada um destes pode apresentar comportamento que melhor

se aproxima da realidade para um determinado tipo de estudo (ou grupo de substâncias), porém não ser tão adequada para estudos de outros tipos. (43) Os principais campos de força empregados para o estudo de biomoléculas são CHARMM (44), AMBER (45), GROMOS (46) e OPLS (47).

#### 2.4.1.1 Campo de Forças *parmbsc0*

Embora as simulações de dinâmica molecular de proteínas tenha tido início do final da década de 1970, (48) a simulação de ácidos nucleicos teve um início mais tardio, cujas primeiras tentativas com sucesso datando nos anos 90. Esta apreciável diferença de tempo é reflexo da necessidade de desenvolvimento de métodos e campos de força adequados para representar as interações eletrostáticas de longo alcance, necessárias para simulações fidedignas de polímeros flexíveis eletricamente carregados.

Dos diversos campos de forças existentes, para o estudo de biomoléculas, em especial de DNA, se destacam o CHARMM e o AMBER. Destes dois campos de força, o AMBER mostrou-se capaz de reproduzir com acurácia propriedades dinâmicas e estruturais de diversas estruturas canônicas e não-canônicas de ácidos nucleicos em soluções aquosas, (45) além de descrever satisfatoriamente a transição entre a forma A- e B- do DNA. (49)

Os campos de força CHARMM apresentam uma conhecida tendência em favorecer a estabilidade do A-DNA, a passo que o AMBER favorece a forma B. (43) Para estudos que envolvam a transição entre as formas do DNA e/ou interação desta biomolécula com outras substâncias, estas características dos campos de força deve ser levada em consideração muito cuidadosamente. Para o presente estudo, no entanto, há maior atenção à influência que o DNA exerce sobre o solvente explícito. Como se sabe que, em água pura e em meios biológicos de forma geral, o DNA se apresenta na forma B, optou-se pelos campos de força da família AMBER.

Os campos de força AMBER parm94 e parm99 foram os principais empregados em simulações de DNA. Sua parametrização foi realizada com cálculos mecânico-quânticos e com simulações de tempos da ordem de 1 ns. Os resultados obtidos foram surpreendentemente bons na escala de uma dezena de nanosegundos, porém em simulações mais longas, foram notadas fortes distorções na geometria.

Estas distorções são, felizmente, pouco significativas para simulações curtas, e não invalidam resultados até então publicados. Para simulações mais longas, como no caso do presente trabalho, os campos de força previamente mencionados não são adequados. O campo de forças AMBER parmbsc0 (49) é uma completa reparametrização baseada no AMBER-parm99. Ao contrário de trabalhos prévios, esta parametrização foi realizada sem fazer uso de refinamento iterativo baseado em simulações de dinâmica molecular. Esta reparametrização foi realizada apenas por meio de cálculos mecânico-quânticos empregando metodologia DFT, ou teoria do funcional da densidade (do inglês, *Density Functional*

Theory), e posterior validação por metodologia de alto nível de teoria como *coupled clusters*, CCSD(T). , por fim, por simulações de dinâmicas clássicas muito longas. Desta forma, o AMBER parmbsc0 é uma melhoria dos campos de forças anteriores, ao menos no que diz respeito à simulações de DNA e RNA.

#### 2.4.1.2 Modelos de Águas

Assim como com os campos de força, diferentes modelos de água são parametrizados de forma diferente e, em função disso, reproduzem o real comportamento da água com maior ou menor fidelidade frente diferentes condições. Os modelos de água mais empregados são o SPC/E (50), ST2 (51), TIP3P, TIP4P, (52) e TIP5P (53), este último representado na Figura 5.

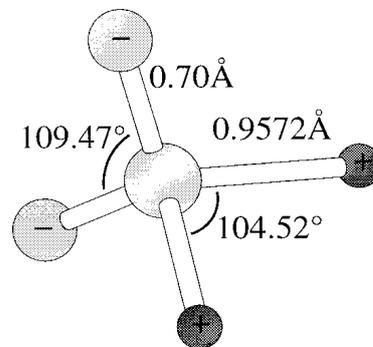


Figura 5 – Representação esquemática do modelo de água TIP5P, mostrando ângulos e comprimentos das ligações. Adaptado de (53).

#### 2.4.2 Temperatura e Pressão

Como descrito até o momento, a dinâmica molecular lida com partículas discretas, determinando suas posições e momentos. Esta metodologia permite obter diretamente informações a nível atômico do sistema que muitas vezes não podem ser obtidas experimentalmente com facilidade. Para determinar propriedades macroscópicas a partir de posições e velocidades de um sistema composto por partículas discretas é necessário usar a termodinâmica estatística.

A temperatura é calculada de acordo com a Equipartição da Energia Cinética do Sistema,  $E_{kin}$ , por meio das equações:

$$E_{kin} = \frac{1}{2} \sum_{i=1}^N m_i v_i^2 \quad (2.6)$$

$$= \frac{1}{2} N_{df} kT \quad (2.7)$$

onde  $k$  é a constante de Boltzmann e  $N_{df}$  é o número de graus de liberdade do sistema.

A pressão escalar,  $p$ , é calculada por meio do tensor pressão,  $\mathbf{P}$ :

$$p = \frac{\text{traço}(\mathbf{P})}{3} \quad (2.8)$$

que por sua vez é calculado pela diferença entre o virial,  $\Xi$  e  $\mathbf{E}_{kin}$ .

$$\mathbf{P} = \frac{2}{V}(\mathbf{E}_{kin} - \Xi) \quad (2.9)$$

$$\Xi = -\frac{1}{2} \sum_{i < j} \mathbf{r}_{ij} \otimes \mathbf{F}_{ij} \quad (2.10)$$

Em princípio, a dinâmica molecular é realizada sem imposições quanto à temperatura e pressão do sistema. O sistema é composto de uma caixa de simulação de volume definido,  $V$ , de composição definida,  $N$ . Além disso, a caixa de simulação é tratada como sendo isolada, isto é, é incapaz de trocar energia ou matéria. Isto caracteriza o *ensemble* NVE, ou microcanônico.

Em experimentos reais, usualmente trabalha-se com condições em que há troca de energia com o meio e que se tem temperatura e pressões constante. Há várias metodologias e algoritmos para a imposição de valores definidos de pressão e temperatura para uma simulação, i.e., a realização de dinâmica molecular no *ensemble* NPT.

#### 2.4.2.1 Termostato de Berendsen

O termostato de Berendsen é um dos algoritmos mais empregados para fazer o acoplamento de uma caixa de simulação a um “banho” externo de temperatura constante de referência,  $T_0$  (54). De forma sucinta, podemos dizer que este termostato re-escala a velocidade das partículas do sistema, de forma a tornar a energia cinética condizente com a temperatura do banho externo.

Matematicamente, em cada passo de tempo no algoritmo,  $\Delta t$ , a velocidade  $v_i$ , é re-escalada à  $\lambda v_i$ , com  $\lambda$  definido por:

$$\lambda^2 = 1 + \frac{\Delta t}{2\tau_T} \left( \frac{T_0}{T} - 1 \right) \quad (2.11)$$

onde  $\tau_T$  é a constante de tempo para o acoplamento ( $\tau_T = (2\gamma)^{-1}$ ;  $\gamma$  é uma constante de fricção).

#### 2.4.2.2 Barostato de Parrinello-Rahman

O barostato de Parrinello-Rahman (55) permite que o tamanho da caixa de simulação varie com o tempo, de forma a atingir a pressão de referência,  $p_0$ . Representando os

vetores da caixa de simulação pela matriz  $\mathbf{b}$ , a seguinte condição é imposta:

$$\frac{d\mathbf{b}^2}{dt^2} = \mathbf{V}\mathbf{W}^{-1}\mathbf{b}'^{-1}(\mathbf{P} - \mathbf{P}_0) \quad (2.12)$$

onde  $\mathbf{W}$  é matriz de parâmetros que define a intensidade do acoplamento deste barostato. Seus termos,  $(\mathbf{W}^{-1})_{ij}$ , são proporcionais à  $\tau_p^{-2}$ . Analogamente ao termostato de Berendsen,  $\tau_p$  é a constante de tempo empregada neste algoritmo.

## 2.4.3 Análises

### 2.4.3.1 Estrutura do DNA

Avaliou-se a estrutura do DNA ao longo das simulações por meio de parâmetros que relacionam consecutivos pares de base e parâmetros que descrevem as bases nitrogenadas em cada par.

Como exemplo de análises entre os pares de bases temos os parâmetros: *shift*, *tilt*, *slide*, *roll*, *rise*, *twist*; parâmetros entre as bases nitrogenadas de um mesmo par são: *shear*, *buckle*, *stretch*, *propeller*, *stagger* e *opening*. (56) Estes parâmetros são ilustrados na Figura 6.

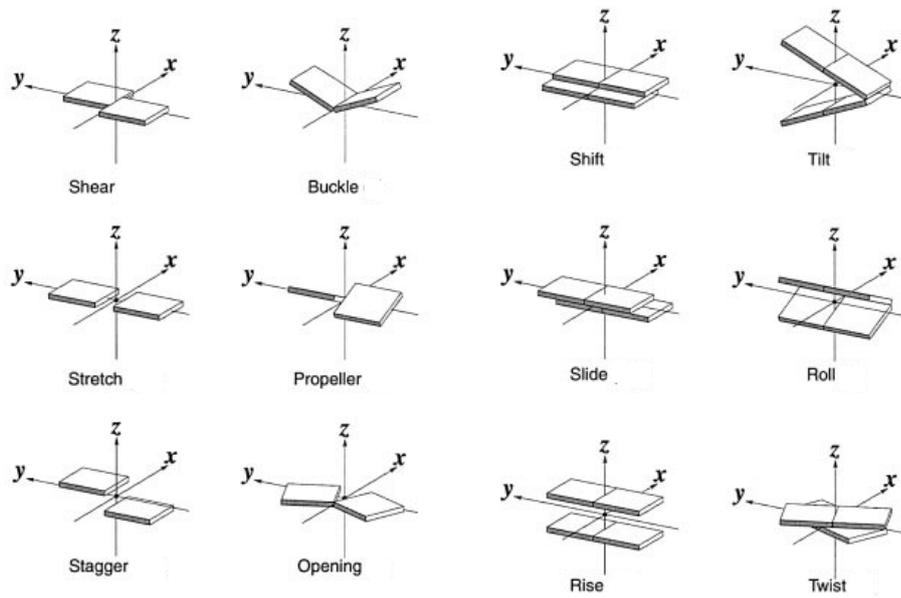


Figura 6 – Definição pictorial dos parâmetros estruturais das bases nitrogenadas do que compõe a fita do *duplex* do DNA. Adaptado de (56).

### 2.4.3.2 Desvio Quadrático Médio

O desvio quadrático médio, ou RMSD (do inglês, *Root Mean Square Deviation*), de uma molécula é calculado com respeito à uma estrutura de referência, usualmente a da

própria molécula em um tempo anterior. É calculado pelo programa *g\_rms* através da Equação 2.13:

$$RMSD(t_1, t_2) = \left[ \frac{1}{M} \sum_{i=1}^N m_i \|\mathbf{r}_i(t_1) - \mathbf{r}_i(t_2)\|^2 \right]^{\frac{1}{2}} \quad (2.13)$$

onde  $M$  é a massa total e  $\mathbf{r}_i(t)$  é o vetor posição do  $i$ -ésimo átomo no tempo  $t$ .

Ao se realizar uma simulação de dinâmica molecular algumas condições e propriedades são impostas e, usualmente, o sistema não se encontra em uma configuração energeticamente favorável no instante inicial, mas evolui para tal configuração com o passar do tempo. Assim a função RMSD cresce em função do tempo, passando a oscilar em torno de um determinado valor. Isto é um indicativo de que o sistema se encontra em um mínimo de sua hypersuperfície de energia. Para a obtenção de dados confiáveis, a simulação deve ser longa o suficiente para que o sistema chegue suficientemente próximo de um mínimo de energia e então se estender por tempo suficiente para que o sistema visite as vizinhanças suficientemente próximas a cada ponto acessível, num tempo finito, ou então as vizinhanças arbitrariamente próximas a cada ponto acessível, de forma a garantir amostragem grande para o cálculo das propriedades de interesse.

### 2.4.3.3 Distribuição Radial

A função de distribuição radial ou RDF (do inglês, *Radial Distribution Function*, comumente denotada por  $g_{\alpha\beta}(r)$ ) é calculada entre duas partículas  $\alpha$  e  $\beta$  usando o program *g\_rdf*, do pacote GROMACS. Seu cálculo é realizado através de:

$$\begin{aligned} g_{\alpha\beta}(r) &= \frac{\langle \rho_{\beta}(r) \rangle}{\langle \rho_{\beta} \rangle_{local}} \\ &= \frac{1}{\langle \rho_{\beta} \rangle_{local}} \frac{1}{N_{\alpha}} \sum_{i \in \alpha} \sum_{j \in \beta} \frac{\delta(r_{ij} - r)}{4\pi r^2} \end{aligned} \quad (2.14)$$

onde  $\langle \rho_{\beta}(r) \rangle$  é a densidade de partículas  $\beta$  em uma distância  $r$  de  $\alpha$ , e  $\langle \rho_{\beta}(r) \rangle_{local}$  é a média da densidade de partículas de  $\beta$  em todas as esferas em torno de  $\alpha$  de raio máximo,  $r_{max}$ , cujo valor normalmente é igual à metade do comprimento da caixa de simulação. Este processo também é feito em diferentes tempo, de forma a obter estatísticas mais confiáveis.

A Equação 2.14 pode ser interpretada como um histograma de partículas em “fatias” esféricas entre  $r$  e  $r + \Delta r$ . Desta forma, altos valores de RDF para um dado valor de  $r$  é indicativo de uma maior densidade de partículas  $\beta$  àquela distância de  $\alpha$ , ou seja, esta função permite inferir sobre a estrutura do sistema, por exemplo, para um gás ideal,

$g(r) = 1, r > 0$ . Por esta razão, é uma das formas mais comuns de representar e estudar a estrutura de líquidos.

Esta análise se faz importante por duas principais razões. Primeiramente, por que dados experimentais de difração de raios-X e de neutrons fornecem informações sobre a RDF da fase fluida. Em segundo lugar, pois  $g(r)$  desempenha um fundamental na descrição teórica de fluidos, principalmente por meio de ferramentas de termodinâmica estatística. (42)

#### 2.4.3.4 Auto-Correlação de Velocidades

A função de auto correlação de velocidades ou VACF (do inglês *Velocity Autocorrelation Function*),  $C_v$ , é calculada pelo produto escalar dos vetores velocidade da partícula de interesse em dois dados tempos, o tempo inicial e  $\tau$ . Uma média é feita de *ensemble* é realizada sobre todas as  $i$  as partículas de  $\alpha$ . Esta função é calculada pelo programa `g_velacc`.

$$C_v(\tau) = \langle \mathbf{v}_i(\tau) \cdot \mathbf{v}_i(0) \rangle_{i \in \alpha} \quad (2.15)$$

Normalmente emprega-se a VACF normalizada,

$$c_v(\tau) = \frac{\langle \mathbf{v}_i(\tau) \cdot \mathbf{v}_i(0) \rangle}{\langle \mathbf{v}_i(0) \cdot \mathbf{v}_i(0) \rangle} \quad (2.16)$$

#### 2.4.3.5 Ligações de Hidrogênio

Em um dinâmica molecular clássica, a análise de ligações de hidrogênio é realizada empregando critérios geométricos. (57) De acordo com este critério, considera-se a existência de uma ligação de hidrogênio quando a distância entre os átomos aceptores e doadores da ligação for igual ou menor a 0,35 nm e se o ângulo Doador-Aceptor-Hidrogênio é entre 0 e 30 graus.

A análise da dinâmica das ligações de hidrogênio foi estudada por duas abordagens distintas. (58, 59, 60) Para ambas análises utilizou-se uma função de auto-correlação da existência de ligações de hidrogênio entre os sítios  $i$  e  $j$ :

$$C_{HB}(t) = \frac{\sum_{ij} s_{ij}(t_0) s_{ij}(t + t_0)}{\sum_{ij} s_{ij}(t_0)} \quad (2.17)$$

onde  $s_{ij}$  assume valor unitário quando houver uma ligação de hidrogênio entre os sítios  $i$  e  $j$  e nulo nos demais casos.

#### 2.4.3.5.1 Intermitente

Para esta análise, considera-se que o sistema “não possui memória”. Isto é, para um dado par de sítios  $i$  e  $j$  a função  $s_{ij}(t + t_0) = 1$  se há uma ligação de hidrogênio entre estes sítios, independentemente da ausência ou presença de ligações de hidrogênio entre estes sítios em tempos entre  $t$  e  $t_0$ .

#### 2.4.3.5.2 Contínua

Nesta análise, por outro lado,  $s_{ij}$  é uma função cujo valor não pode aumentar com o passar do tempo. Isto é, para um determinado par de sítios  $i$  e  $j$  em que há uma ligação de hidrogênio em  $t_0$ , temos que  $s_{ij}(t_0) = 1$ . Com a evolução temporal, caso a ligação de hidrogênio seja desfeita em um tempo  $t' + t_0$ , temos que  $s_{ij}(t' + t_0) = 0$ . Mesmo que a uma nova ligação de hidrogênio seja estabelecida entre estes sítios em um tempo  $t'' + t_0$  (com  $t'' > t'$ ), a função não recupera seu valor unitário, ou seja  $s_{ij}(t'' + t_0) = 0$ .

#### 2.4.3.6 Coeficiente de Difusão

O coeficiente de difusão,  $D$ , de uma dada substância é o fator de proporcionalidade da Lei de Fick de difusão, e está diretamente relacionado com a mobilidade de suas moléculas. O coeficiente de difusão é uma constante física que depende, entre outros fatores, da pressão, da temperatura, da massa da substância e do ambiente em que a substância se encontra. Pode-se falar de dois tipos de coeficiente de difusão, o de interdifusão e o de autodifusão. O primeiro trata da difusão em misturas, e pode ser mais facilmente determinado experimentalmente. Já o segundo refere-se à difusão de uma molécula de referência em outras moléculas de mesma espécie. Experimentalmente, sua determinação apresenta maiores dificuldades, e destaca-se o uso de experimentos de ressonância magnética nuclear. Computacionalmente, sua determinação não apresenta grandes desafios.

A determinação do coeficiente de autodifusão da água é de grande interesse em diversos aspectos. Das diversas anomalias que este líquido apresenta, a família de anomalias ditas dinâmicas tem talvez como melhor exemplo o fato de que o coeficiente de autodifusão aumenta com a compressão do líquido, de forma contrária à intuição físico-química e ao comportamento da maioria dos demais fluidos.

Além disso, uma vez que este solvente atua na interação de ligantes com sítio ativos de proteínas e com os sulcos do DNA, o coeficiente de autodifusão é um parâmetro de interesse para avaliar a mobilidade da água nestes processos. Em particular, sabe-se que moléculas de água que compõe a espinha de hidratação do DNA apresentam mobilidade muito reduzida se compara às demais. (40)

O coeficiente de autodifusão pode ser determinado por uma das relações de Green-Kubo. (61, 62) Especificamente, o coeficiente de autodifusão se relaciona com a função de auto-correlação de velocidade do fluido através da Equação 2.18:

$$D = \frac{1}{3} \int_0^{\infty} \langle \mathbf{v}(t = \mathbf{0}) \cdot \mathbf{v}(t) \rangle dt \quad (2.18)$$

O coeficiente de autodifusão também pode ser calculado empregando a fórmula de Einstein.

$$\langle \Delta \mathbf{r}(t)^2 \rangle = A + 6Dt + \text{flutuações} \quad (2.19)$$

Assim, D pode ser determinado pelo coeficiente angular do desvio quadrático médio das posições do solvente. Esta abordagem é mais adequada para sistemas com temperaturas não muito baixas, quando  $D > 10^{12} m^2 \cdot s^{-1}$ . Além disso, não pode ser calculado com tempos muito longos, pois D tende a zero quando o valor do deslocamento se aproxima das dimensões do sistema.(42)

#### 2.4.3.7 Fator de Estrutura Translacional

Um dos comportamentos anômalos da água mais facilmente observado, o mínimo de densidade em função da temperatura em 4°C e 1 atm, evidencia que sua estrutura tridimensional se reorganiza de maneiras bem características; Sua estrutura se torna menos organizada com o aumento de pressão. Debenedetti e Errington (9) introduziram duas métricas extremamente úteis para avaliar o grau de estruturação de um fluido.

O primeiro, fator de estrutura translacional,  $\tau_{trans}$ , é uma medida do quanto a estrutura difere daquela que seria a de um fluido “*idealmente desorganizado*”, isto é, o quanto difere estruturalmente de um gás perfeito. Intuitivamente, este parâmetro pode ser calculado por:

$$\tau_{trans} = \int_0^{\infty} |g(\mathbf{r}) - 1| d\mathbf{r} \quad (2.20)$$

A segunda métrica é o fator de estrutura tetraédrico,  $q_{tet}$ . Este fator orientacional avalia o quanto o ângulo formado entre o átomo de oxigênio de uma molécula de água de referência se aproxima (ou difere) de um arranjo tetraédrico com as moléculas vizinhas. É calculador por:

$$q = 1 - \frac{3}{8} \sum_{j=1}^3 \sum_{k=j+1}^4 \left( \cos \phi_{jk} + \frac{1}{3} \right)^2 \quad (2.21)$$

onde  $\phi_{jk}$  é o ângulo formado entre o átomo de oxigênio e o de suas quatro vizinhas mais próximas,  $j$  e  $k$ . Este parâmetro varia entre 0, para um gás ideal, e 1, para um sistema de arranjo perfeitamente tetraédrico.

Esta segunda métrica, embora extremamente útil para o estudo da fase condensada, não foi empregado neste estudo, pois perde seu sentido para estudar moléculas de água nos sulcos do DNA, uma vez que lá o solvente não é circundado por demais moléculas da mesma espécie.

## 3 Objetivos

### 3.1 Objetivos Gerais

Verificar se modelos de água e metodologia empregadas em simulações de biomoléculas reproduzem a vitrificação da água de solvatação do DNA, como observado experimentalmente. Fornecer uma visão a nível molecular deste fenômeno.

### 3.2 Objetivos Específicos

1. Realizar estudos de dinâmica molecular modelos explícitos de água, na presença ou não de DNA e de íons de concentração fisiológica.
2. Avaliar detalhadamente o comportamento estrutural e dinâmico de moléculas de água, fazendo distinção entre o solvente em cada sulco do DNA (maior e menor) e do solvente *bulk*.
3. Interpretação dos resultados obtidos, comparação entre os diferentes sistemas e o solvente em diferentes ambientes químicos;
4. Comparação e interpretação com resultados computacionais e experimentais disponíveis na literatura.

## 4 Metodologia e Equipamentos

Todas as simulações foram realizadas no supercomputador *cluster SGI:Altix (Gauss)* do CESUP (Centro Nacional de Supercomputação) UFRGS. Este *cluster* possui performance de pico teórica de 15,97 Tflops.

Todos as análises foram realizados no *cluster Vader*, localizado no laboratório do Grupo de Química Teórica da UFRGS. O *cluster* é composto de 12 nodos e possui performance de pico teórica de 1,06 Tflops.

### 4.1 Dinâmica Molecular

As etapas de dinâmica molecular realizadas neste trabalho podem ser divididas em três etapas: definição e montagem do sistema em caixas de simulação; simulações longas no *ensemble* NpT e simulações mais curtas no *ensemble* NVE.

#### 4.1.1 Caixas de Simulação

Primeiramente, elaborou-se um um arquivo `.pdb` com a estrutura do DNA, no caso, com uma sequência do *duplex* de dodecâmero de Dickerson. Para isto, usou-se a programa `fiber` do X3DNA (56) e usou-se a sequência CGCGAATTCGCG de bases nitrogenadas. Este arquivo foi modificado manualmente com um editor de texto para remover a unidade de fosfato da referente à primeira unidade de base nitrogenada de cada fita do DNA.

Esta estrutura é composta apenas por átomos pesados, i.e., não possui hidrogênios. A adição destes é feita com o programa `pdb2gmx`, onde é definido também o campo de forças a ser empregado, no caso, o *AMBER parmbsc0* (49). Nota-se que este campo de forças não é nativo à versão empregada do GROMACS, sendo incorporada manualmente empregando o flag `-ff`. Os arquivos são disponibilizados no site do próprio GROMACS (63) e material de suporte é disponibilizado pelos autores em seu site.(64).

O programa `genbox` é empregado para colocar a estrutura no centro uma caixa de simulação cúbica. As dimensões da caixa foram escolhidas de forma que a menor distância entre algum átomo do DNA os limites da caixa fosse de 1 nm.

A estrutura do DNA gerado pelo X3DNA não está relaxada em uma conformação realista. Para relaxar a estrutura é realizada uma rápida minimização de energia empregando o algoritmo *steepest descent*.

Em seguida, uma vez que a molécula de DNA esta ionizada negativamente, são adicionados 22 cátions de sódio através do programa `genbox`. Estes íons não são dispostos

pelo programa em posições próximas dos sítios negativos do DNA, como ocorre na realidade. Portanto, procedeu-se uma curta dinâmica molecular com o algoritmo *md* no vácuo, com o uso de *restraints*.

Em seguida, à caixa adicionou-se moléculas do solvente. Isto é feito novamente pelo programa *genbox*. Esta adição é feita através diversas cópias de caixas constituídas por moléculas de água do modelo de interesse que já foram previamente equilibradas. Por esta razão, o sistema resultante apresenta simetria artificialmente elevada. Portanto é feita uma simulação de dinâmica molecular para que o solvente relaxe para uma configuração mais realista.

Enfatiza-se que a ordem na qual estas duas últimas etapas (adição de contra-íons e adição de solvente) deve ser respeitada. Ao se adicionar os cátions em uma caixa hidratada, estes estariam distantes da macromolécula negativamente carregada e solvatados. Embora o sistema assim obtido seja eletricamente neutro, não condiz com a realidade. Além disso, desta forma, seria necessário desperdiçar recurso computacional para atingir a situação adequada, uma vez que a mobilidade iônica é muito maior para íons no vácuo do que para íons hidratados.

O sistema resultante, constituído pelo DNA, solvente explícito e contra-íons, pode ser empregado para simulações posteriores simulações com temperatura e pressão definidas. Porém, para o sistema no qual o DNA não está no solvente puro, e sim em uma solução de cloreto de sódio de concentração fisiológica ( $C_{NaCl} = 0,154 mol.dm^{-3}$ ). Para tanto, por meio do programa *genion*, um total de 54 moléculas de água foram removidas do sistema e substituídas por 27 pares de íons. Este sistema é então submetido à uma dinâmica molecular para que o sistema relaxe, em especial o solvente em torno dos íons que foram colocados de forma artificial.

Também preparou-se caixas de simulação de tamanho similar contendo apenas o solvente puro.

### 4.1.2 Simulações $NpT$

Realizou-se simulações de dinâmica molecular com o sistema obtido em 4.1.1 no *ensemble*  $NpT$ , isto é, mantendo a composição, a pressão e a temperatura constantes ao longo da simulação. A temperatura foi controlada por meio do termostato de Berendsen (54), utilizando constante de tempo,  $\tau_T$ , de 0,5 ps. A pressão foi controlada por meio do barostato de Parrinello-Raman (55), utilizando uma constante de tempo,  $\tau_p$  de 2 ps. A pressão foi mantida em uma atmosfera em todas as simulações.

A interação eletrostática foi avaliada pelo método *Particle Mesh Ewald* (65, 66). Utilizou-se um passo de integração,  $\Delta t$ , de 0,002 ps e as coordenadas de todos os átomos do sistema foram registradas em um arquivo de trajetórias (*.xtc*) a cada 500 passos

de simulação. Ressalta-se que, nestas simulações, as velocidades dos átomos não foram registradas.

Primeiramente, empregou-se uma temperatura de 300 Kelvin, gerada por uma distribuição de Maxwell-Boltzmann de velocidades. Após uma simulação relativamente longa, chegava-se a uma situação de equilíbrio e o espaço de fases era amostrado. Desprezando-se tempos iniciais da simulação, pôde-se calcular propriedades estruturais e dinâmicas do sistema (desde que não dependessem explicitamente das velocidades dos átomos), como RMSD e funções de distribuição radial.

A posição e velocidade dos átomos do sistema eram então empregados como dados de entrada para uma nova simulação de dinâmica molecular na qual a temperatura era 290 K. Novamente procedia-se uma simulação longa para que se atingisse um equilíbrio compatível a nova temperatura. Desprezando os tempos de pré-equilíbrio procedia-se a análise.

O processo foi repetido, baixando-se a temperatura em 10 Kelvin a cada iteração. Deve-se ressaltar que, baixando-se a temperatura, diminui-se a velocidade média das partículas que compõe o sistema. As consequências prática disto são que conforme temperaturas mais baixas eram atingidas, simulações mais longas eram necessárias para atingir o equilíbrio termodinâmico. E, além disso, uma vez atingido o equilíbrio, também se era necessário simulações mais longas para amostrar adequadamente o espaço de fases.

Para temperaturas mais altas (entre 300 e 270 K) estas simulações foram de 30 ns. Com o resfriamento, o tempo de simulação foi aumentado conforme a necessidade de estabilizar cada sistema, avaliado por meio do RMSD.

### 4.1.3 Simulações *NVE*

Para a análise de fenômenos dependentes da velocidade dos átomos, como a função de auto-correlação de velocidades, e/ou de fenômenos que ocorrem em escalas de tempo muito curtas, a metodologia descrita em 4.1.2 não é adequada. Primeiramente por um caráter prático: não foram registradas as velocidades, e as coordenadas foram salva com intervalo longo. Tecnicamente, poder-se-ia registrar os dados de forma a evitar esta limitação, porém a quantidade de dados gerado exigiria uma capacidade de armazenamento indisponível. Em segundo lugar, o uso do barostato e, principalmente, do termostato, influencia de forma artificial o sistema.

Para a grande maioria das análises realizadas isto conduz a um erro pouco significativo, uma vez que o acoplamento é baixo. Porém pode ser importante para algumas análises. Assim, para estas análises realizou-se simulações sem o uso de termostato ou barostato, portanto, no *ensemble* NVE, onde, além da composição, o volume e a energia do sistema são mantidas constantes.

Uma simulação curta neste *ensemble* foi feita para cada temperatura. Empregou-se como estrutura de entrada a estrutura resultante no final da simulação NpT de mesma temperatura. O passo de integração empregado foi de 0,001 ps e a cada 5 passos de simulação registrou-se a posição e a velocidade de cada átomo do sistema em um arquivo de trajetórias (`.trr`). Estas simulações tiveram duração de 25 ps e todas as análises foram realizadas sobre todo este espaço temporal.

Usando esta metodologia estudou-se cinco sistemas:

- **TIP3P-DNA**, composto de um dodecâmero de Dickerson e seus contra-íons, o sistema é solvatado por água explícita do modelo *tip3p*;
- **TIP5P-DNA**, como o modelo TIP3P-DNA, porém empregando o modelo de água *tip5p*;
- **TIP5P-FISIO**, sistema composto pelo dodecâmero de Dickerson e seus contra-íons e hidratado com água do modelo TIP5P, há também a substituição de algumas moléculas de água por íons cloreto e sódio para que se tenha concentração fisiológica dos mesmos;
- **TIP3P-BULK**, sistema de água pura *tip3p*, sem a adição de nenhum soluto;
- **TIP5P-BULK**, um sistema de água pura, como o anterior, porém empregando solvente do modelo *tip5p*.

#### 4.1.4 Análises

##### 4.1.4.1 Análise Estrutural do DNA

A estrutura do DNA foi analisada empregando o pacote X3DNA, usando o programa `analyze`. A cada 100 ps das trajetórias foi obtida uma *snapshot* do sistema sob a forma de um arquivo `.gro` que foi analisado. Desta forma, pôde-se avaliar a evolução temporal dos parâmetros estudados.

##### 4.1.4.2 Água nos Sulcos e Água

Para diversas análises realizadas neste trabalho é importante distinguir quais moléculas se encontram nos sulcos maior e menor do DNA, bem como de moléculas de água que se encontrem suficientemente afastadas da biomoléculas, i.e, água *bulk*.

Por estarem em constante movimento ao longo das simulações, a lista, ou *index*, de moléculas de água que se encontram em uma determinada região de interesse sofre constante modificação. O programa `g_select` do pacote GROMACS permite a indexação dinâmica de átomos segundo critérios definidos pelo usuários.

Este processo foi empregado em uma única *frame*, ou configuração instantânea, gerando uma lista de átomos de oxigênio pertencentes às moléculas de água no instante inicial,  $t_0$  da simulação no *ensemble* NVE, segundo critérios geométricos de interesse esclarecidos abaixo.

#### 4.1.4.2.1 Água *bulk*

Para estabelecer quais moléculas de água não são influenciadas pelo DNA, utilizou-se o `g_select` para identificar as moléculas de água cujo oxigênio se encontrasse no instante  $t_0$  a uma distância superior à 1 nm. A Figura 9 mostra esta seleção.

#### 4.1.4.2.2 Água nos Sulcos e Água *Bulk*

Para identificar as moléculas de água localizadas nos sulcos do DNA, e ainda distinguir entre as que estão no sulco maior e menor, utilizou-se como critério a distância entre o átomo de oxigênio da água e alguns átomos específicos dos pares de bases nitrogenadas do DNA. Considera-se que está em um sulco moléculas cujos oxigênio estejam situados até 0,3 nm de átomos específicos do DNA, identificados na Figura 7.

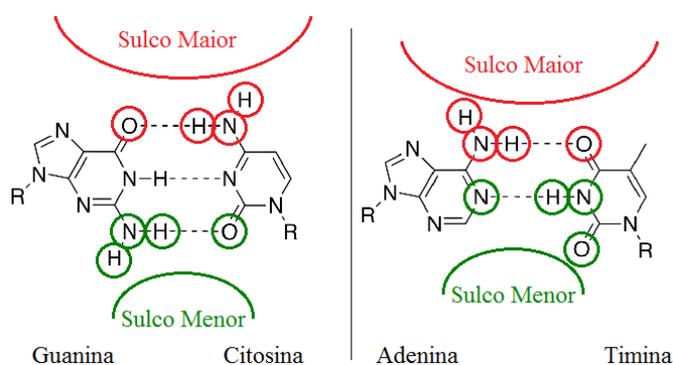


Figura 7 – Pareamento de bases GC e AT. São representadas as ligações de hidrogênio características dos pareamentos. São identificados os átomos de Nitrogênio, Oxigênio e Hidrogênio que ficam expostos ao solvente nos sulcos maior, em vermelho, e menor, em verde. Figura adaptada de (67).

O pareamento canônico das bases nitrogenadas não ocorre com geometria arbitrária e sua orientação na macromolécula é bem definida com relação aos sulcos. Assim, é possível mapear manualmente a estrutura DNA e gerar um índice com os nitrogênios, oxigênios e hidrogênios que ficam expostos a cada sulco.

#### 4.1.4.3 Análises com GROMACS

Análises de RMSD, energia (total, cinética e potencial), pressão, temperatura e densidade foram realizadas nas simulações do *ensemble* NpT para avaliar a evolução do

sistema para uma situação de estabilidade (ou meta-estabilidade) termodinâmica.

As principais análises do trabalho foram realizadas com as trajetórias geradas nas simulações do *ensemble* NVE. Estas análises foram realizadas considerando todo o tempo de simulação neste *ensemble*.

#### 4.1.4.3.1 Análises RMSD

Para as análises de RMSD, torna-se necessário fazer uma correção devido à periodicidade do sistema. Ao chegar a extremidade de uma caixa de simulação, um átomo é “transportado” para o outro lado da caixa, o que leva à uma grande distância entre a posição no instante  $t_1$  e no instante  $t_2$  (Equação 2.13), causando um “salto” artificial no gráfico de  $\text{RMSD}(t)$ . Esta correção das trajetórias foi realizada com o programa *trjconv* e as opções *-pb* *nojump*.

#### 4.1.4.3.2 Análises de Ligação de Hidrogênio

As ligações de hidrogênio têm cinética relativamente rápida, o que leva a tempos de vida curtos. Desta forma, para analisar este fenômeno, não são necessárias simulações muito longas, porém, prescinde-se de simulações mais detalhadas. Isto é, o passo de integração da dinâmica molecular deve ser menor. Além disso, deve-se registrar as posições dos átomos com maior frequência.

Para realizar esta análise foram realizadas simulações no *ensemble* NVE, como descrito acima, porém utilizando-se passo de integração de 0,0001 ps e salvando a posição dos átomos a cada 5 passos de integração.

As ligações de hidrogênio do sistema foram analisadas empregando o programa *g\_hbonds* para analisar as trajetórias geradas com este procedimento. Analisou-se as ligações de hidrogênio entre moléculas de água e átomos que compõe o sulco maior, entre moléculas de água e o sulco menor e entre moléculas de água que constituem o *bulk*.

#### 4.1.4.3.3 Coeficiente de Autodifusão da Água

Determinou-se o coeficiente de auto-difusão da água por meio da relação de Green-Kubo, Equação 2.18. Para isto, é necessário o cálculo da função de autocorrelação de velocidades não-normalizada e empregando moléculas inteiras. Isto foi realizado por meio do programa *g\_velacc*. A integração da função assim obtida foi feita por meio do programa *g\_analyze* e sua correção pelo fator de um terço (bem como a conversão para unidades do Sistema Internacional) por meio de um *script* empregando o programa *awk*, nativo ao Linux.

Esta análise foi realizada para moléculas que estavam no sulco menor do DNA, no sulco maior do DNA e no *bulk* dos sistemas.

O cálculo do coeficiente de autodifusão por meio da fórmula de Einstein não foi realizado. É adequado fazer esta análise e comparar os resultados obtidos pelas duas abordagens, visando uma melhor distinção entre o coeficiente de autodifusão propriamente dito e o uma “aparente” difusão devido ao movimento de uma molécula dentro de sua “jaula” de hidratação (movimentos rápidos,  $\beta$  do tipo *rattling in the cage*).

#### 4.1.4.3.4 Fator de Estrutura Translacional

O fator de estrutura translacional é calculado pela Equação 2.20. Uma vez que a função de distribuição radial,  $g(\mathbf{r})$  já é calculada, pelo programa `g_rdf`, basta apenas realizar a subtração de uma unidade e posterior integração do valor absoluto do resultado. A subtração foi realizada por meio de um *script* usando o programa `awk` e a posterior integração numérica utilizando o programa `g_analyze`.

Esta análise também foi realizada fazendo distinção entre águas nos sulcos maior e menor do DNA e da água *bulk*.

## 5 Resultados

Aplicando a metodologia descrita, foi possível baixar a temperatura do sistema TIP3P-DNA até a temperatura de 200 K e os sistemas TIP5P-FISIO e TIP5P-DNA até a temperatura de 240 K. Obteve-se as mesmas respectivas temperaturas para os sistemas de água pura, TIP3P-BULK e TIP5P-BULK.

A simulação de sistemas em baixas temperatura tem conhecidamente um elevado custo computacional. Com a diminuição da temperatura do sistema e redução da velocidade média de seus constituintes, o sistema precisa de mais tempo para poder adotar diferentes conformações condizentes com seu estado termodinâmico, isto é, amostrar seu espaço de fases. Computacionalmente isso leva à necessidade de um maior número de passos de integração.

Além disso, como descrito em (53), o modelo de água *tip5p* possui dois átomos *dummy*, que não existem no modelo *tip3p*. Por esta razão, os sistemas TIP5P-FISIO e TIP5P-H2O possuem cerca de 18 mil átomos a mais, que devem percorrer todas as etapas do algoritmos de dinâmica molecular, acarretando em simulações que exigem maior número de horas de cálculo.

A diferença entre as temperaturas mais baixas nas quais se pôde chegar entre os modelos de água se deve também à outra diferença, cuja natureza também é, essencialmente, fruto do caráter artificial dos modelos, mas que apresenta uma interpretação física: o coeficiente de auto-difusão.

A Figura 8 mostra como varia o coeficiente de autodifusão dos dois modelos de água com a temperatura. Nota-se claramente que o modelo *tip5p* tem locomoção muito mais baixa que o modelo *tip3p*. Portanto, não obstante os cálculos em si serem mais lentos com o modelo *tip5p*, fisicamente, estes sistemas precisam também de cálculos mais longos para percorrer o espaço de fases. A temperatura de 240 K se torna um limite de resfriamento que se pode obter para o modelo *tip5p*. Nesta temperatura o coeficiente de autodifusão do sistema assume o valor de  $2,333 \cdot 10^{-11} m^2 \cdot s^{-1}$ . Nesta condição o fluido está tão super-resfriado que pode-se dizer que é um sólido, ou, dependendo do seu grau de estruturação e de suas propriedades de transporte, um vidro.

Continuar o resfriamento dos sistemas envolveria um custo computacional muito elevado. Uma alternativa para diminuir o custo computacional seria o uso de algoritmos estocásticos do tipo Monte-Carlo(42, 68) para resfriar o sistema. Neste âmbito, porém, ocorreria a perda de informações dinâmicas, uma vez que a variável temporal é extinta. Pode-se, no entanto, ainda inferir em propriedades dinâmicas do sistema a partir do cálculo da entropia de excesso do sistema, que pode, por exemplo, ser relacionada ao coeficiente

de autodifusão do sistema. (69, 70)

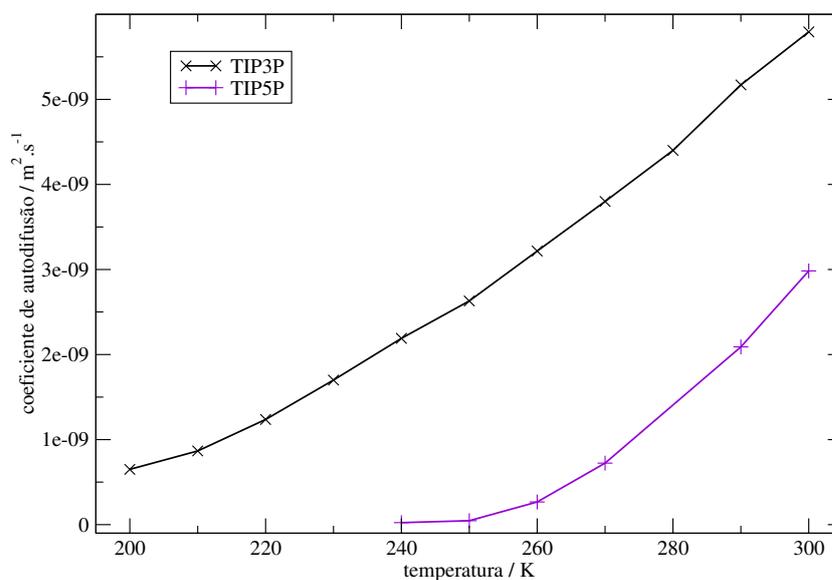


Figura 8 – Coeficiente de autodifusão em função da temperatura para os modelos de água *tip3p* e *tip5p* simulados na ausência de DNA.

## 5.1 Inspeção Visual

A Figura 9 mostra um *snapshot*, ou configuração instantânea, de um dos sistemas obtidos, gerada empregando o programa de visualização *VMD*. (71)

Pouca informação pode ser extraída pela inspeção visual desta Figura. De qualquer forma, nota-se que o DNA permanece em sua forma característica de dupla hélice. A Figura 9 mostra uma *snapshot* equivalente. Porém, nesta representação é evidenciada a espinha de hidratação do DNA, por meio dos átomos de oxigênios próximos dos átomos que compõem os sulcos maior e menor, como descrito em na metodologia.

Podemos verificar que o número de moléculas se encontram no sulco é muito pequeno: em torno de 25 (27 para o sistema exposto na Figura 9) no sulco menor e 30 (31) no maior, uma fração muito pequena frente o número total, que é da ordem de 9000 moléculas. Desta forma, constatamos que, embora os sulcos apresentem apreciavelmente diferente, o número de moléculas do solvente pouco diferente entre eles. Isto é condizente com o que a Figura nos mostra: qualitativamente, as moléculas de água no sulco menor se encontram muito próximas umas das outras; já no sulco maior elas conseguem se acomodar melhor no sulco, estando mais afastadas umas das outras.

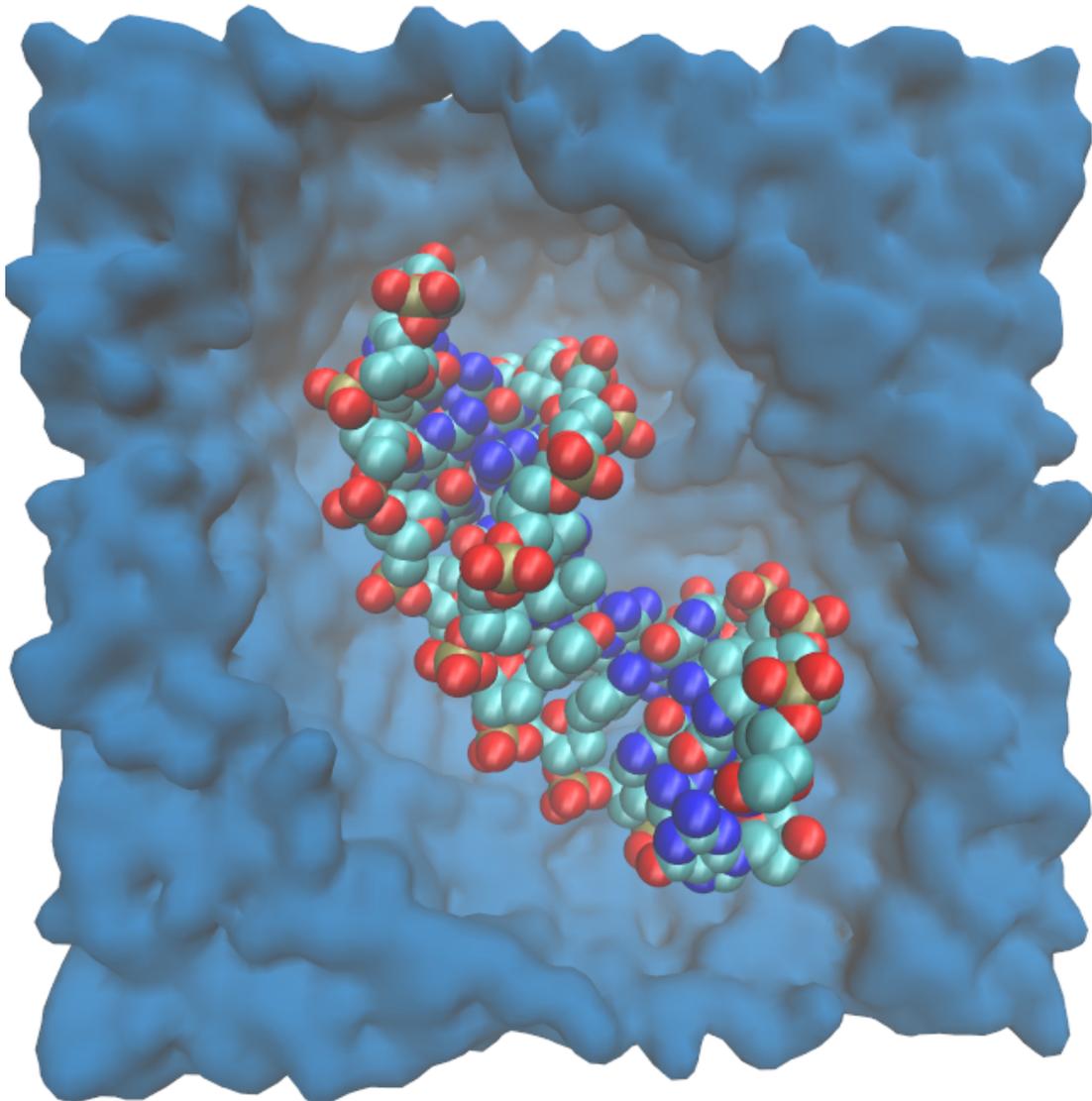


Figura 9 – Representação da caixa de simulação, evidenciando o DNA (omitindo átomos de hidrogênio para maior clareza) e, em azul, a superfície azul escura foi gerada pelos átomos de oxigênio de moléculas de água distando pelo menos 1 nm do DNA e omitindo as demais moléculas do solvente. Figura obtida para o DNA após simulação no *ensemble* NVE, para o sistema TIP5P-DNA sob temperatura de 240K. Em vermelho são representados átomos de oxigênio, em marrom átomos de fósforo, em azul claro átomos de carbono e em azul escuro átomos de nitrogênio.

## 5.2 Temperatura

No *ensemble* NpT a temperatura é mantida no valor desejado por meio do termostato, sofrendo flutuações. No *ensemble* NVE, no entanto, a temperatura não é necessariamente mantida fixa. Espera-se que ao término de simulações NpT suficientemente longas,

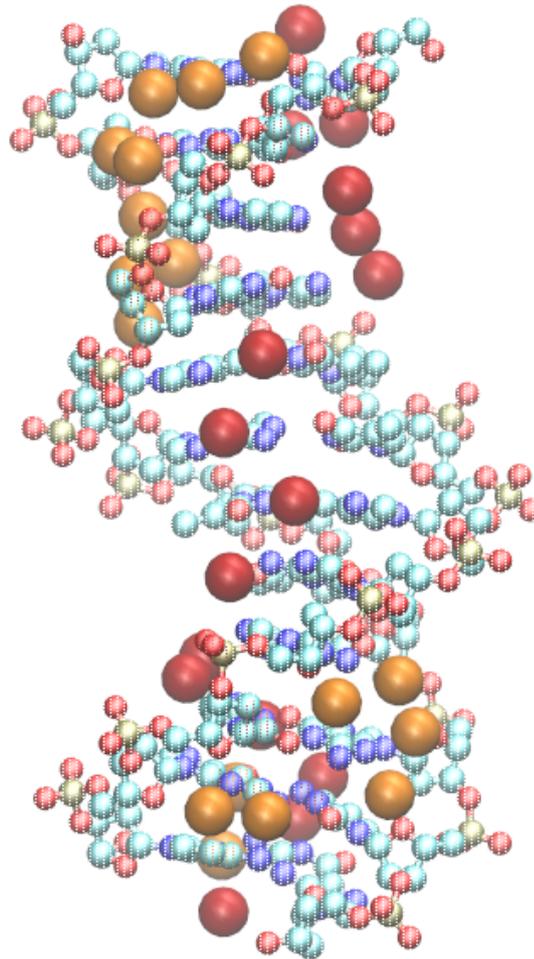


Figura 10 – Representação da caixa de simulação, evidenciando o DNA (omitindo átomos de hidrogênio para maior clareza) e átomos de oxigênio de moléculas de água que solvatam o sulco maior (vermelho) e menor (laranja), como descrito em 4.1.4.2.2. Figura obtida para o DNA após simulação no *ensemble* NVE, empregando o sistema TIP5P-DNA sob temperatura de 240K. As esferas menos representam em azul claro átomos de carbono, em vermelhos de oxigênio, em azul escuro nitrogênio e em laranja fósforo.

após atingir uma situação de equilíbrio, o sistema se mantenha com temperatura e pressão constantes, mesmo com a desativação do termostato e barostato, assim como a energia. Porém, sabe-se que para que a energia seja verdadeiramente conservada, é necessário um tratamento mais rigoroso das interações de longo alcance. (54).

Nas simulações NVE realizadas houve um aumento de energia cinética com um

atenuado aumento da energia total do sistema, isto é, rigorosamente ela não foi conservada. Sobretudo com passos de integração muito pequenos, onde simulações sobre períodos curtos necessitam de várias iterações, e assim, sendo mais propensas a erros numéricos cumulativos. O pacote GROMACS apresenta inúmeros pontos vantajosos frente outros programas de dinâmica molecular. Infelizmente, porém, a adequada conservação de energia no *ensemble* NVE não é uma delas. Para estes sistemas o aumento de energia cinética leva a um aumento da temperatura dos sistemas, como ilustrado na Figura 11.

Isto ocorre mais pronunciadamente nos sistemas abaixo do ponto de congelamento da água, i.e., sistemas super-resfriados. Estando super-resfriado e incapaz de solidificar, o sistema tende a relaxar para conformações mais estáveis, com diminuição de sua energia potencial e aumento da temperatura

Felizmente, porém para a análise de fenômenos como ligações de hidrogênio, que possuem tempos característicos muito curtos (da ordem de décimos de picosegundos) este aumento da temperatura não é capaz de levar à erros significativos.

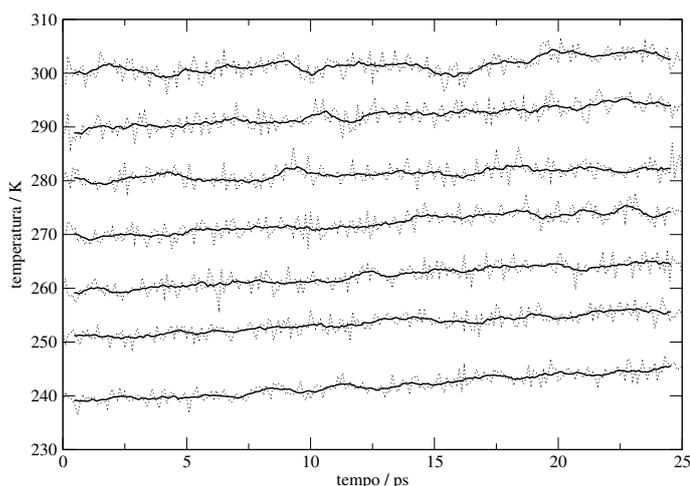


Figura 11 – Temperatura em função do tempo para o sistema TIP5P-FISIO após simulação no *ensemble* NVE, empregando o sistema TIP5P-FISIO.

## 5.3 Estrutura

### 5.3.1 Água *Bulk*

A estrutura dos sistemas é estudada por meio de sua função de distribuição radial,  $g(r)$  e por meio do fator de estrutura translacional,  $\tau_{trans}$ . Estas análises foram realizadas empregando o átomo de oxigênio das moléculas de água, OW, como referência. Na Figura 12 é mostradas a função  $g(r)$  para os sistemas puros, i.e., sem o soluto DNA, em diferentes temperaturas.

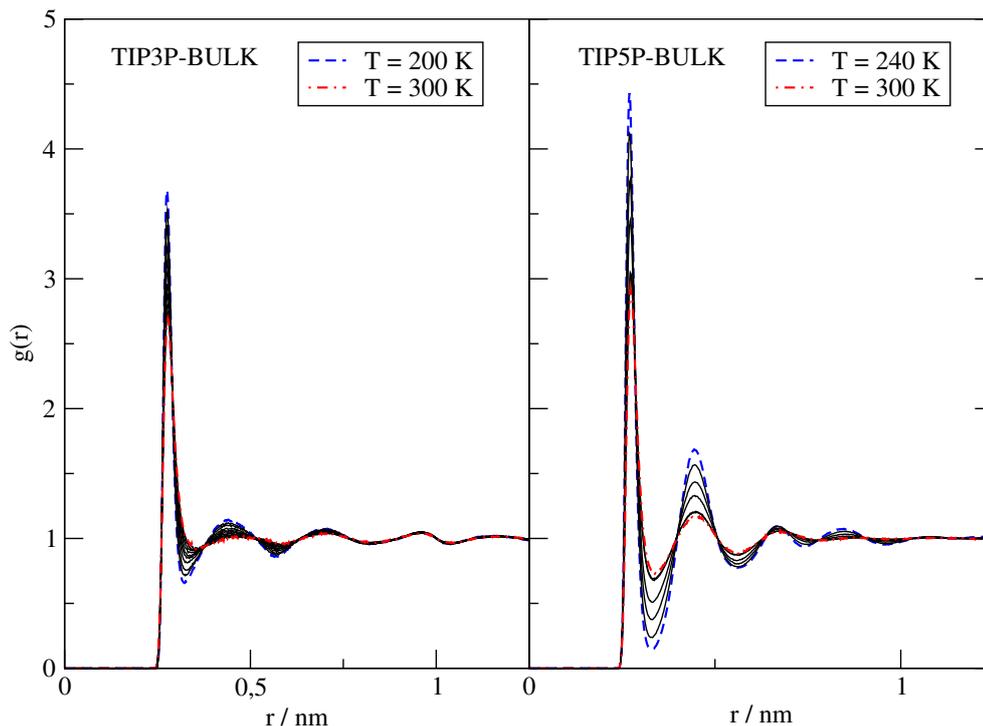


Figura 12 – Função de distribuição radial OW-OW para sistema TIP3P-BULK e TIP5P-BULK em diferentes temperaturas. A linha pontilhada azul representa temperatura mais baixa (200 K para o Sistema TIP3P-BULK e 240K para o sistema TIP5P-BULK.), a linha azul a temperatura mais alta (300 K em ambos sistemas), as demais temperaturas intermediárias são representadas em preto.

Nota-se claramente que a  $g(r)$  assume valores muito diferentes da unidade em distâncias curtas, indicando regiões com maior ou menor quantidade de moléculas. Isto mostra as diferentes camadas de hidratação do solvente. A estruturação do solvente não se limita à primeira camada de hidratação, sendo possível identificar que a estrutura se mantém até a terceira camada de solvatação.

O abaixamento da temperatura torna os máximos e mínimos mais pronunciados, indicando que o sistema torna-se mais estruturado. Podemos observar que com o resfriamento o número de moléculas na primeira camada de solvatação aumenta. Comparando as funções entre os dois modelos, verificamos que o modelo *tip5p* apresenta estrutura mais bem definida que o outro modelo, além de ter sua influência mais intensamente influenciada pela mudança de temperatura.

Notamos então que em ambos os casos o fluido apresenta-se com alto grau de estrutura, mesmo em elevadas temperaturas. Este comportamento é mais transparente por meio do fator de estrutura translacional, mostrado na Figura 13. Notamos que para o modelo *tip5p*, entre 280 e 300 K,  $\tau_{trans}$  sofre uma variação muito menor do que nas demais temperaturas. Esta observação é explicada pelo fato de que, nesta faixa de temperatura,

ocorre o máximo de densidade deste fluido, o que não ocorre no modelo mais simples.

Com ambas as métricas é fácil de constatar que não só o modelo *tip5p* se encontra com maior grau de estruturação em todas as temperaturas, mas como também seu grau de estruturação é mais sensível à variações de temperatura.

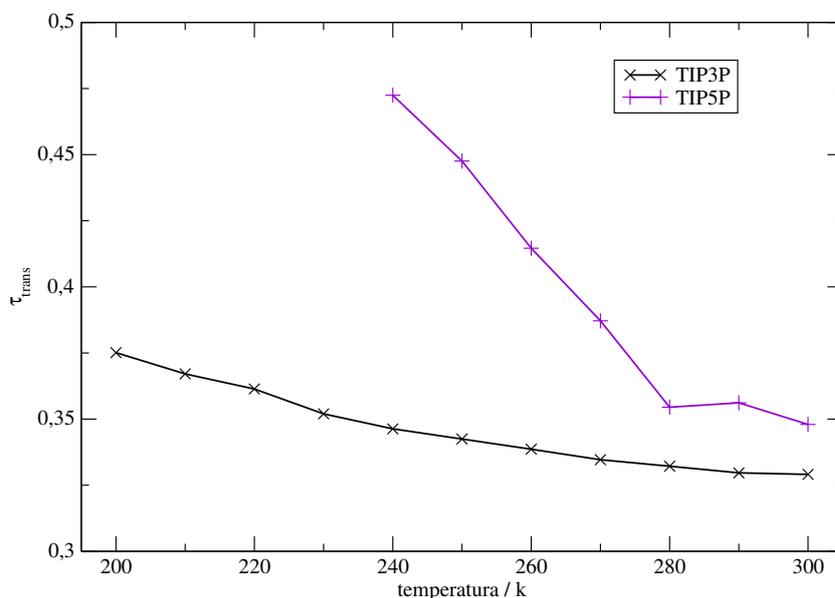


Figura 13 – Influência da temperatura sobre o fator de estrutura translaconal,  $\tau_{trans}$ , para os sistemas TIP3P-BULK e TIP5P-BULK.

Nas Figuras 14 e 15 são mostradas as funções de distribuição radial da água *bulk* dos sistemas TIP3P-DNA e, TIP5P-FISIO e TIP5P-DNA, respectivamente, comparando-os com os sistemas sem soluto.

Analisando a Figura 14, verificamos que a presença do DNA é capaz de influenciar levemente o número de coordenação do solvente *bulk* de água TIP3P, tornando o sistema mais estruturado. Este efeito torna-se mais pronunciado em temperaturas mais baixas.

Um comportamento similar é observado para o sistema TIP5P-FISIO em relação ao sistema TIP5P-BULK. Para o sistema TIP5P-DNA, no entanto, este efeito é mais pronunciado.

Assim, interpretamos que o a estrutura do solvente *tip3p* sofre pouca influência pela macromolécula. O modelo de água *tip5p*, por outro lado, pode ter sua estrutura fortemente influenciada pelo DNA, mesmo em distância de 1 nm. Porém, este fenômeno só é constatado na ausência de outros solutos, indicando que em soluções eletrolíticas a formação de esferas de solvatação em torno de íons domina a estrutura do solvente, o que sugere que a influência causada pelo DNA se deva à interações eletrostáticas de longo alcance.

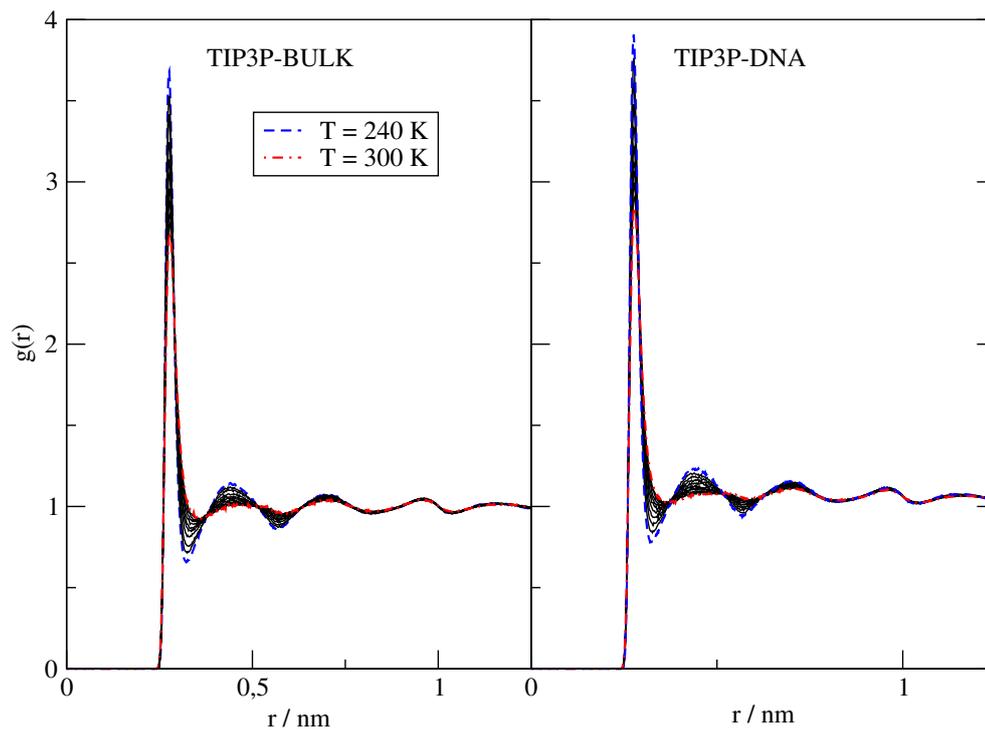


Figura 14 – Função de distribuição radial OW-OW dos sistemas TIP3P-DNA e TIP3P-BULK em diferentes temperaturas.

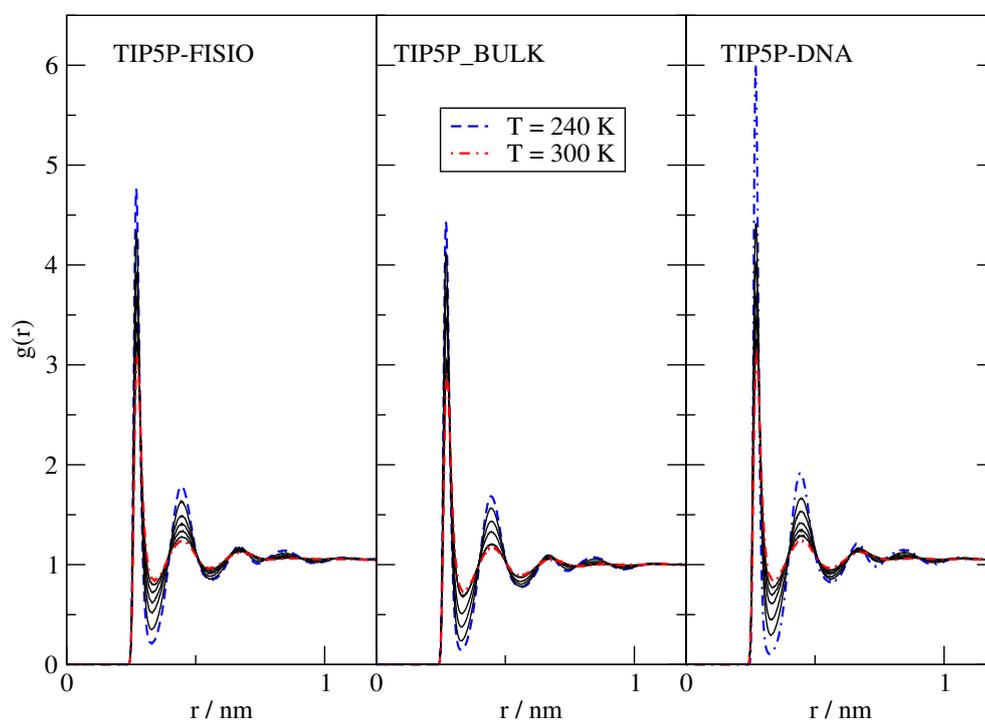


Figura 15 – Função de distribuição radial OW-OW dos sistemas TIP5P-FISIO, TIP5P-DNA e TIP5P-BULK em diferentes temperaturas.

### 5.3.2 Espinha de Hidratação

Nas Figuras 16 e 17 são mostradas as  $g(r)$  da espinha de hidratação nos sistemas TIP3P-DNA e, TIP5P-FISIO e TIP5P-DNA, respectivamente. As funções foram calculadas entre átomos OW e átomos que compõem os sulcos.

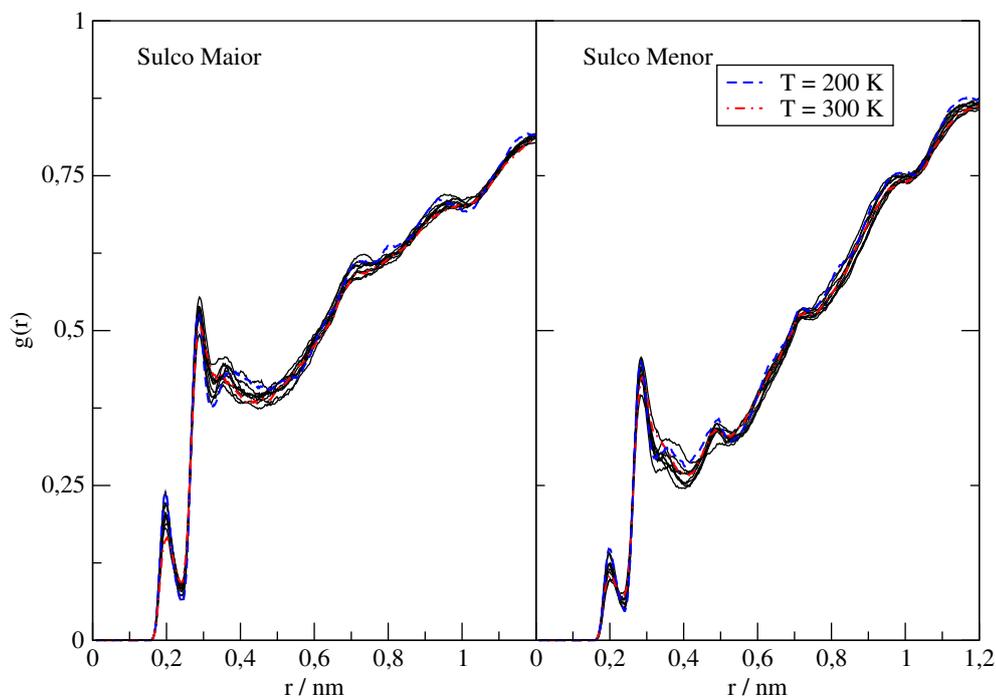


Figura 16 – Função de distribuição radial entre átomos OW e átomos dos sulcos do DNA para o sistema TIP3P-DNA em diferentes temperaturas.

Para o modelo TIP3P-DNA podemos observar em ambos os sulcos um pequeno máximo de  $g(r)$  para pequenas distâncias, que corresponde às moléculas que se encontram nos sulcos propriamente ditos. Este número é muito pequeno, indicando a dificuldade de acomodar as moléculas no espaço confinado sulco. De qualquer forma, constata-se o que poderia ser esperado: o sulco maior apresentar valores de  $g(r)$  maiores e menos estreitos que o sulco menor. Além disso, após o primeiro mínimo ocorre um segundo máximo, indicando uma preferência configuracional para as moléculas do solvente nas proximidades das moléculas que estão no interior do sulco, formando assim a espinha de hidratação do DNA. Após o segundo mínimo  $g(r)$  cresce quase que linearmente até atingir o valor unitário, mantendo-se essencialmente constante. Assim, os sulcos não têm forte influência além da segunda esfera de hidratação. Por fim, a temperatura parece ter pequena influência sobre a estrutura da água nos sulcos. Isto é facilmente compreendido, pois como o DNA mantém sua estrutura com o resfriamento, o tamanho dos sulcos em si sofre pequena variação, não alterando o número de moléculas de solvente neles contidos. Estes resultados são condizentes com o que poderia ser esperado observando a Figura 10.

Nos sistemas TIP5P-FISIO e TIP5P-DNA a função  $g(r)$  comporta-se qualitativa-

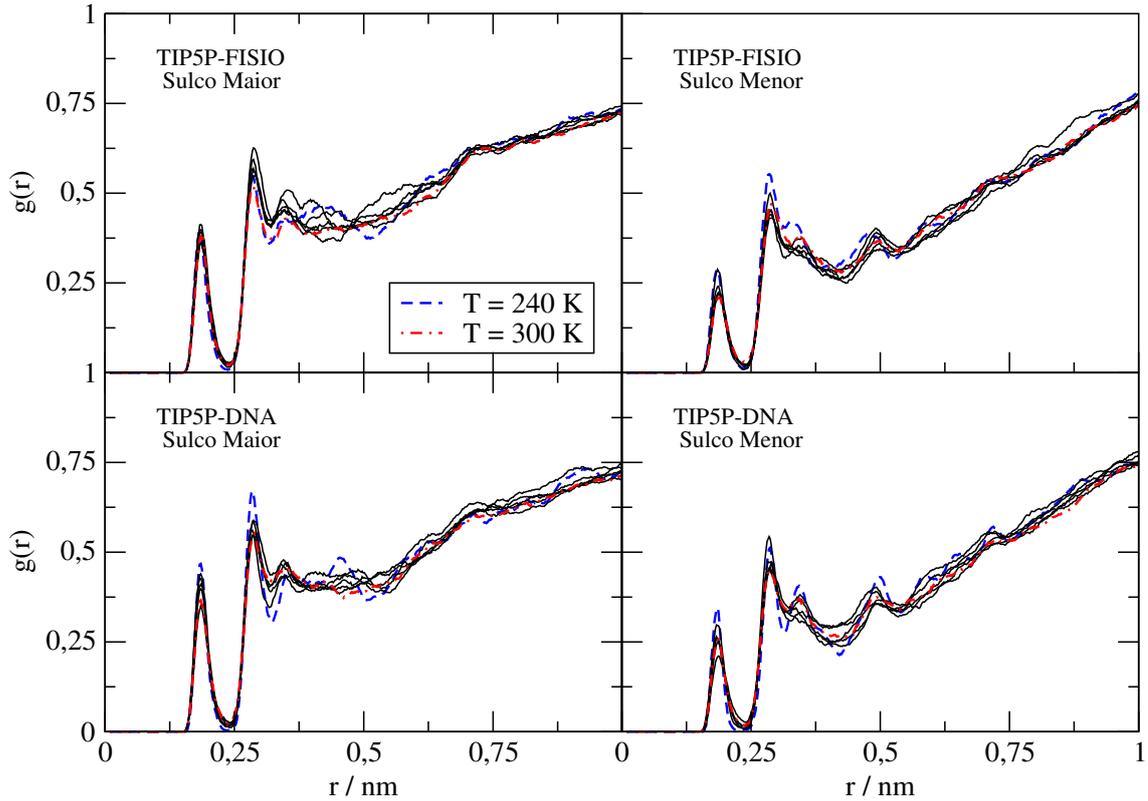


Figura 17 – Função de distribuição radial entre átomos OW e átomos dos sulcos do DNA para os sistemas TIP5P-FISIO TIP5P-DNA em diferentes temperaturas.

mente de mesma forma que no sistema TIP3P-DNA, diferindo apenas em alguns aspectos. Primeiramente, uma maior quantidade de moléculas do modelo *tip5p* consegue se acomodar dentro dos sulcos, pois a presença dos dois sítios de interação eletrostática permite uma melhor interação com os átomos da periferia dos pares de bases nitrogenadas do DNA. Em segundo lugar, no primeiro mínimo, que define as primeiras regiões externas aos sulcos,  $g(r)$  assume valores nulos para o modelo de água *tip5p*, a passo que no sistema TIP3P-DNA é não nulo. Interpretamos este resultado como consequência da melhor hidratação dos sulcos nos sistemas TIP5P-; há uma região melhor definida para a hidratação preferencial. Por fim, a função  $g(r)$  assume valores nulos após a primeira camada de hidratação nos sulcos. A ausência de moléculas nesta região revela que, em média, o solvente que se encontra dentro do sulco não o deixa, i.e., durante o tempo de simulação mantém-se na mesma distância em relação ao DNA.

Também não notamos grandes diferenças entre os resultados obtidos para os sistemas TIP5P-FISIO e TIP5P-DNA, o que mostra que os íons dissolvidos não são capazes de influenciar o interior dos sulcos; eles são um ambiente químico à parte do restante do sistema.

A Figura 18 mostra a análise estrutural por meio do fator de estrutura translacional. Notamos claramente que  $\tau_{trans}$  dos sulcos é muito pouco sensível à temperatura. Curiosa-

mente, porém, para os sistemas TIP5P- esta métrica diminui com com a diminuição da temperatura, até 250 K, passando então a aumentar, a passo que para o *bulk* do sistema ocorre um aumento monotônico. Este comportamento pode ser um indicio de facilitação de comportamento anômalo causado pelo ambiente químico diferente dos sulcos, uma vez que neste as anomalias da água apresentam uma hierarquia bem definida, na qual a manifestação das anomalias estruturais cobre maior parte dos diagramas de fases. Este comportamento ocorre consistentemente nos sistemas TIP5P-FISIO e TIP5P-DNA, sendo mais pronunciado no sulco menor e praticamente inexistente no sistema TIP3P-DNA.

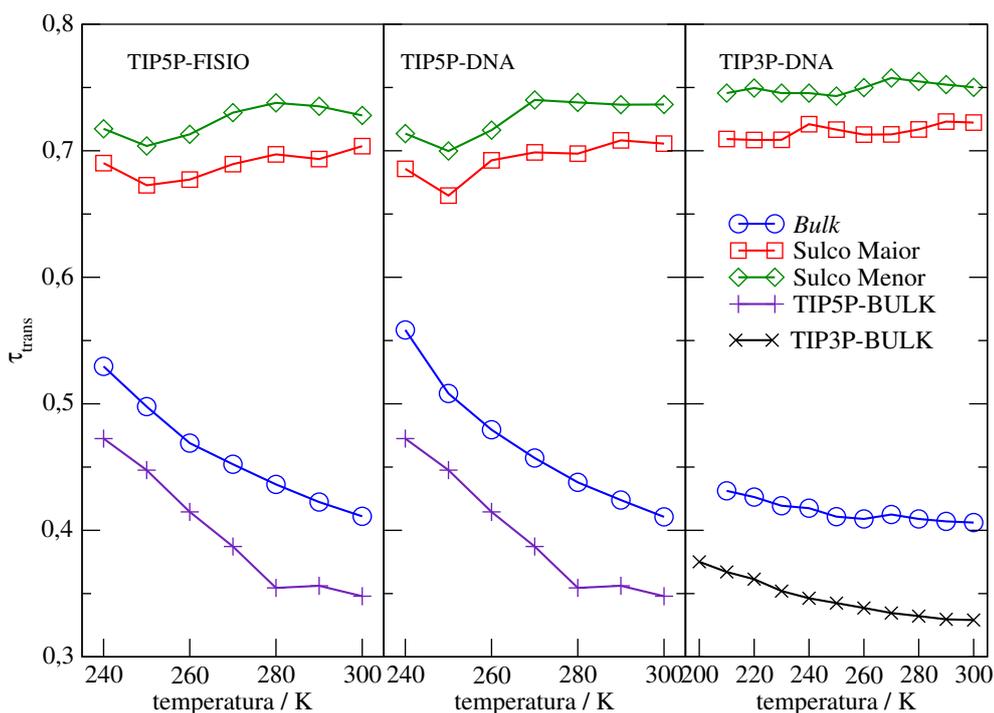


Figura 18 – Influência da temperatura sobre o fator de estrutura translacional para água *bulk* e espinha de hidratação para todos os sistemas estudados.

Verificamos que o DNA influencia muito pouco o solvente em longas distâncias e que esta influência, quando presente, é devida essencialmente às interações eletrostáticas e facilmente eclipsado na presença de outros eletrólitos, indicando que o DNA não altera significativamente propriedades estruturais da água a longo alcance, ao contrário das *anti-freezing proteins*.

Na Figura 19 é mostrada a influência da temperatura na densidade dos diversos sistemas estudados, bem como dos modelos de solvente puros como referência. Primeiramente, podemos notar que no sistema TIP3P-BULK e TIP3P-DNA ocorre um aumento monotônico da densidade com a diminuição da temperatura. Isto acontece pois este modelo de água não foi parametrizado para reproduzir o máximo de densidade apresentada pelo fluido real. Os sistemas TIP5P-FISIO e TIP5P-DNA pode-se notar que ocorre um máximo de densidade entre 280 e 290 K. Este modelo de água é parametrizado para reproduzir

esta anomalia termodinâmica em 277,15 K.

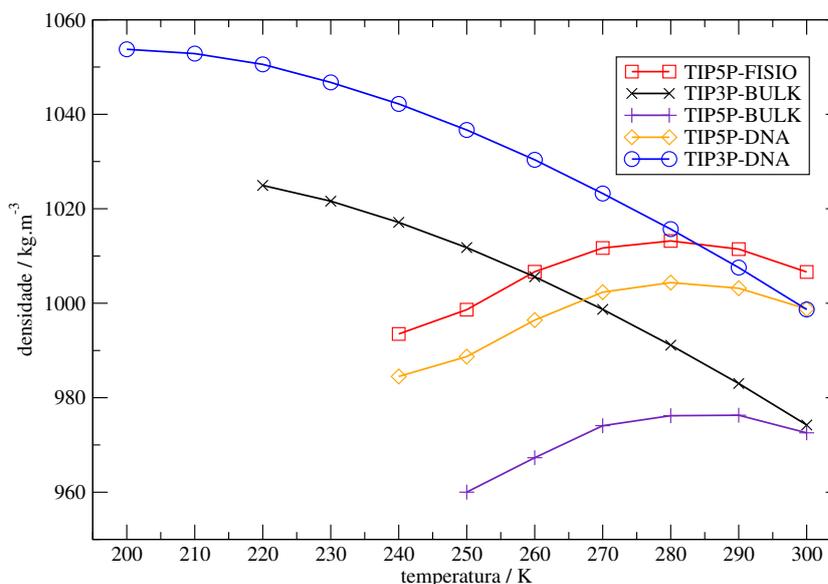


Figura 19 – Densidades de equilíbrio em função da temperatura para os diferentes sistemas.

Em todos os casos a variação da densidade com a temperatura não é influenciada pela presença da biomolécula por íons. Isto também é explicado pelo fato de que não há grandes mudanças no *bulk* do sistema, e, desta forma, seu volume é pouco influenciado pela presença destes solutos. As diferenças em absoluto entre as densidades acaba sendo devido apenas a diferente composição e massa entre os sistemas.

## 5.4 Funções de autocorrelação de velocidades

Nas Figuras 20 a 23 são mostradas as funções de autocorrelação de velocidades normalizadas para os átomos OW dos sistemas. Analisando qualitativamente os gráficos mencionados, verificamos que para a água *bulk* a correlação é perdida rapidamente, atingindo valores nulos em apenas cerca de meio picossegundo, independente do modelo de água empregado. Nota-se também que a correlação é mais intensa para temperaturas mais baixas. Por fim, verificamos que a função de autocorrelação das velocidades de moléculas de água modelo *tip3p* que se encontram nos sulcos são influenciadas pela temperatura, o que não acontece com o modelo *tip5p*, pois mantem-se mais próximas dos sulcos, em média.

A autocorrelação de velocidades pode ser mais claramente estudada por meio de sua transformadas de Fourier. Nas Figuras 24 a 27 são mostradas estas transformadas.

Para a água *bulk*, a dinâmica do solvente é dominada por duas frequências características uma em  $2 \text{ ps}^{-1}$  e outra em  $8,5 \text{ ps}^{-1}$ . A separação entre estes dois modos é acentuada com o resfriamento do sistema. Para temperatura de 300 K, nem o modelo *tip3p* nem

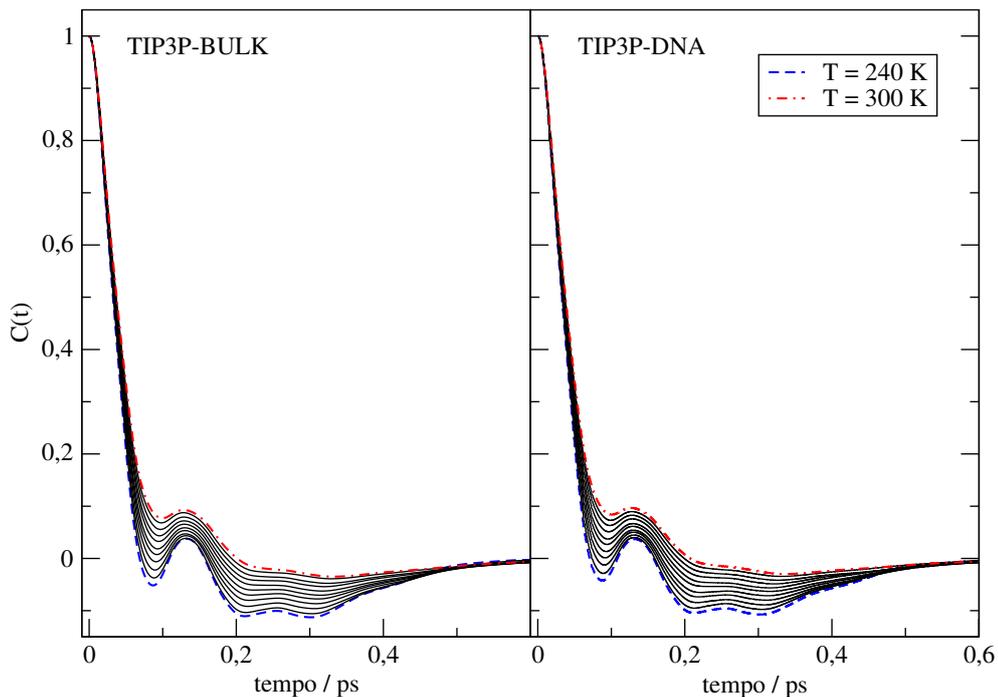


Figura 20 – Funções de autocorrelação dos átomos OW em diferentes temperaturas para o *bulk* do sistema TIP3P-DNA e TIP3P-BULK.

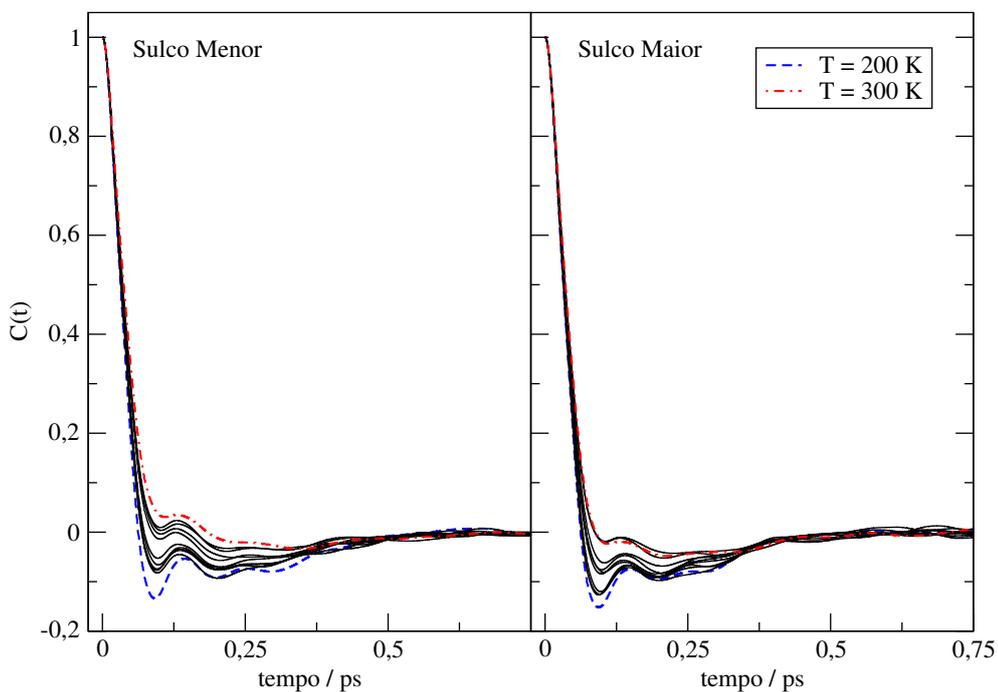


Figura 21 – Funções de autocorrelação dos átomos OW da espinha de hidratação no sistema TIP3P-DNA em diferentes temperaturas.

o *tip5p* possui uma contribuição significativa da alta frequência. O resfriamento acentua esta dinâmica, mas ela permanece insignificante no modelo *tip3p*. Podemos interpretar este movimento de alta frequência como sendo característico de um sistema com dinâmica

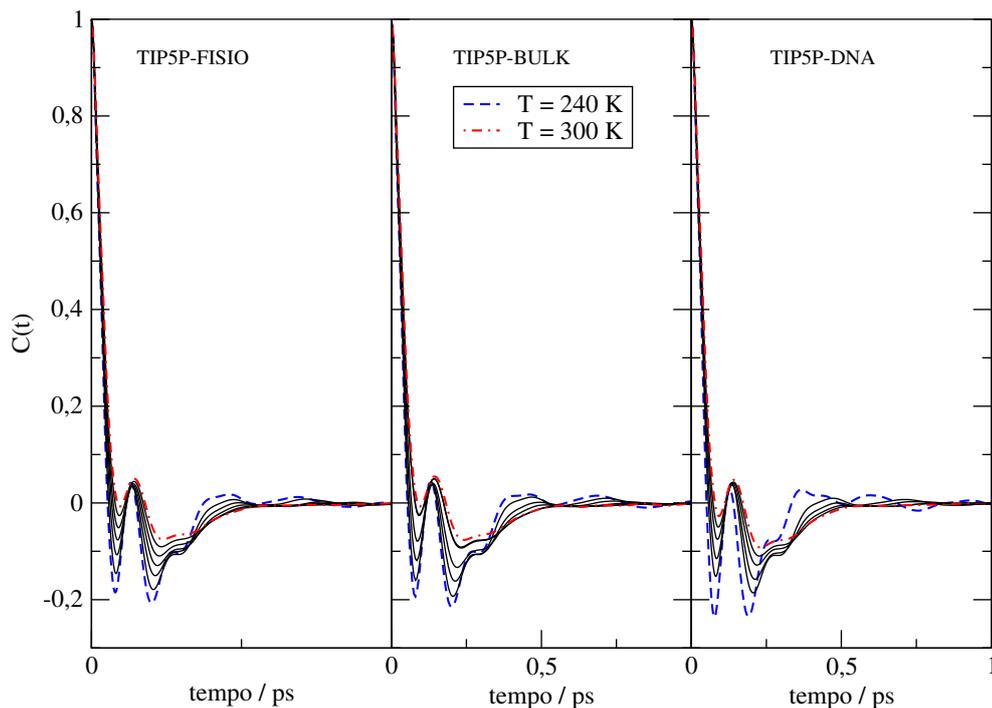


Figura 22 – Funções de autocorrelação dos átomos OW em diferentes temperaturas para o *bulk* do sistema TIP5P-DNA, TIP5P-FISIO e TIP5P-BULK.

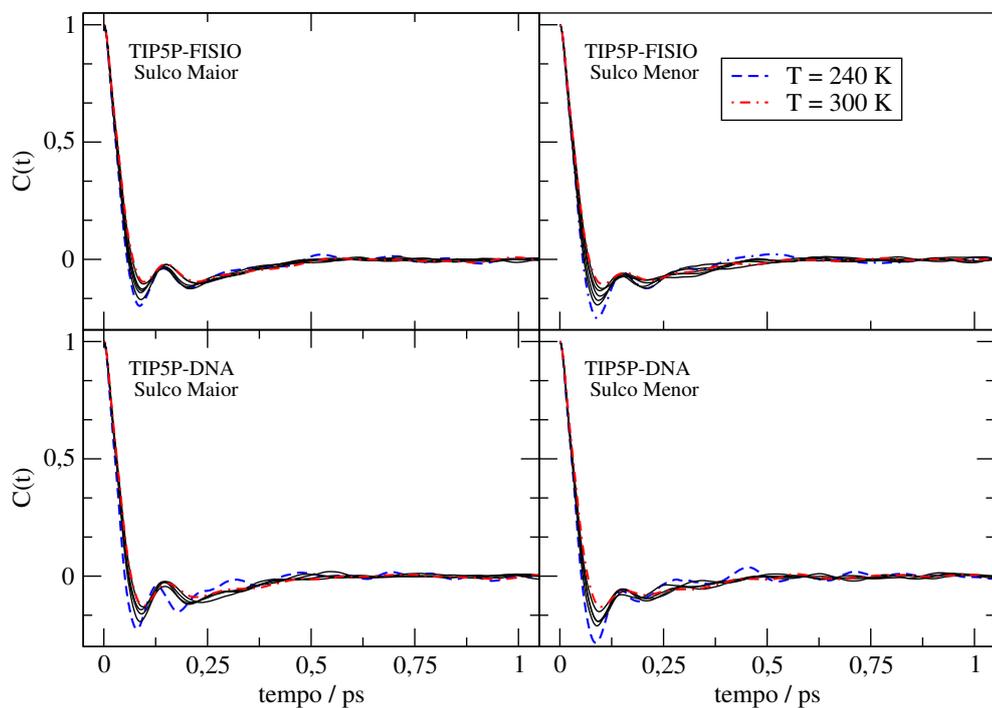


Figura 23 – Funções de autocorrelação dos átomos OW da espinha de hidratação no sistemas TIP5P-DNA e TIP5P-FISIO em diferentes temperaturas.

atenuada e envolvida na manifestação de uma transição vítrea dinâmica.

Na água dos sulcos do sistema TIP3P-DNA não é possível identificar modos bem

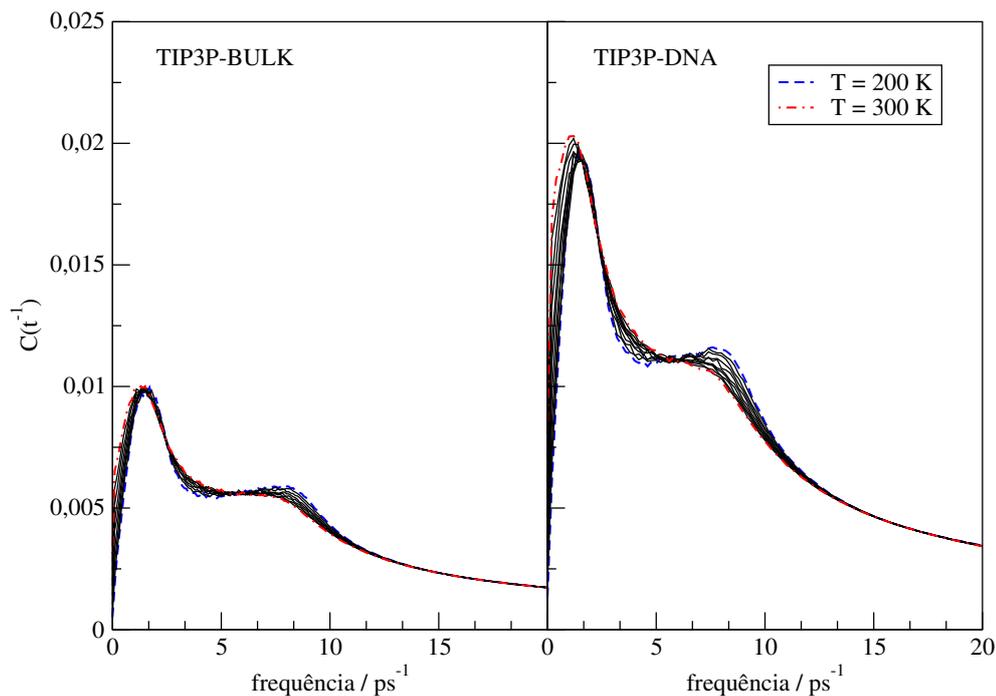


Figura 24 – Transformada de Fourier da VACF de água *tip3p* no *bulk* dos sistemas.

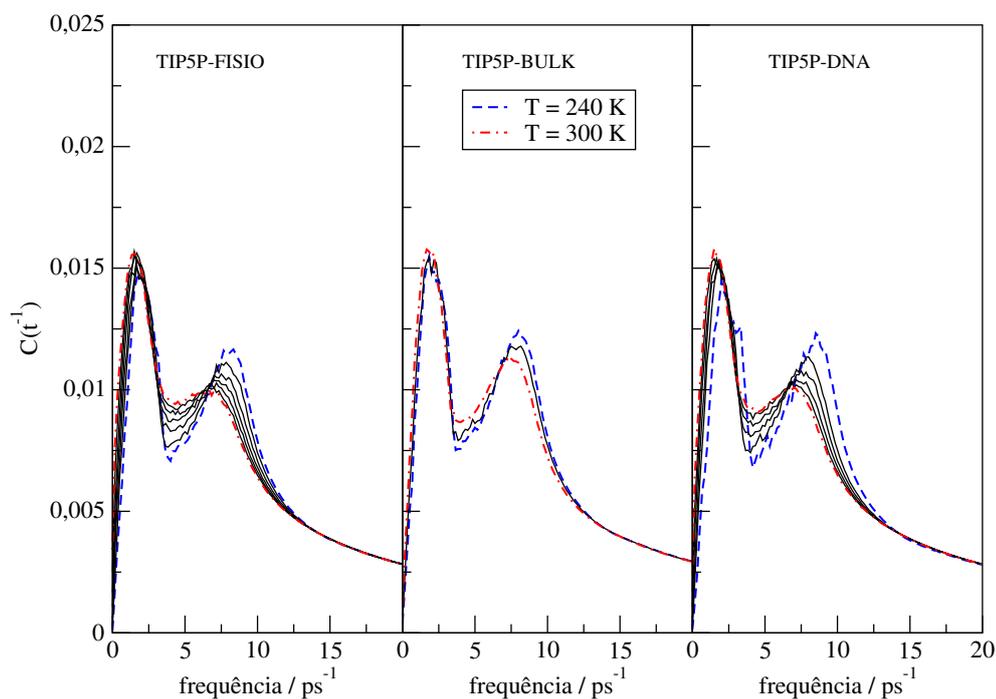


Figura 25 – Transformada de Fourier da VACF de água *tip5p* no *bulk* dos sistemas.

definidos.

Nos sistemas TIP5P-FISIO e -DNA, é possível identificar que a dinâmica do sistema é composto por duas frequências principais, embora estas não sejam bem definidas.

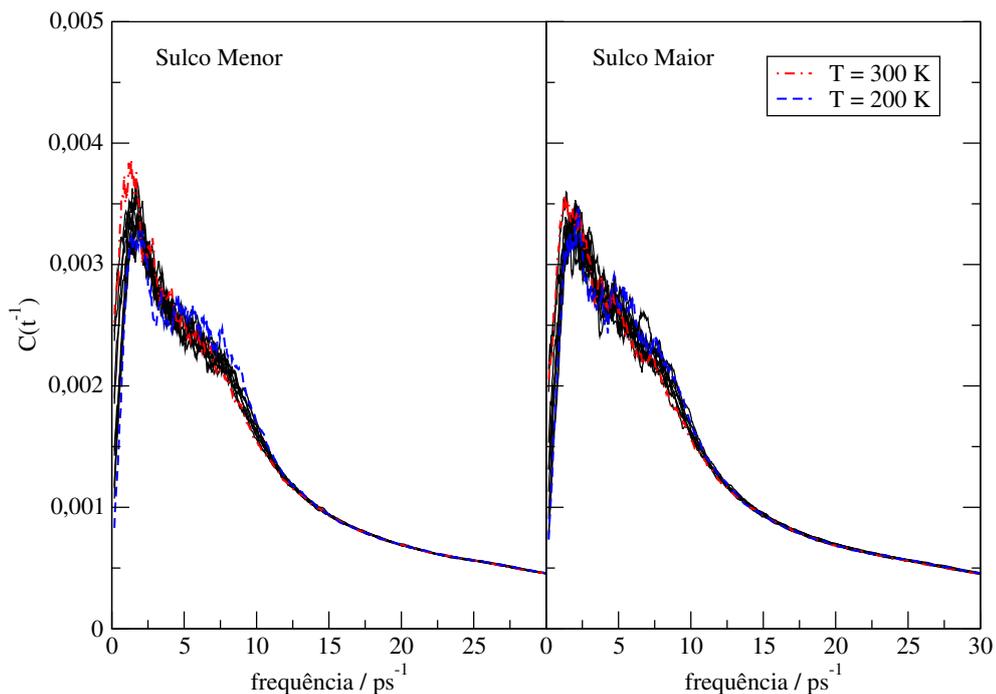


Figura 26 – Transformada de Fourier da VACF de água nos sulcos do DNA no sistema TIP3P-DNA.

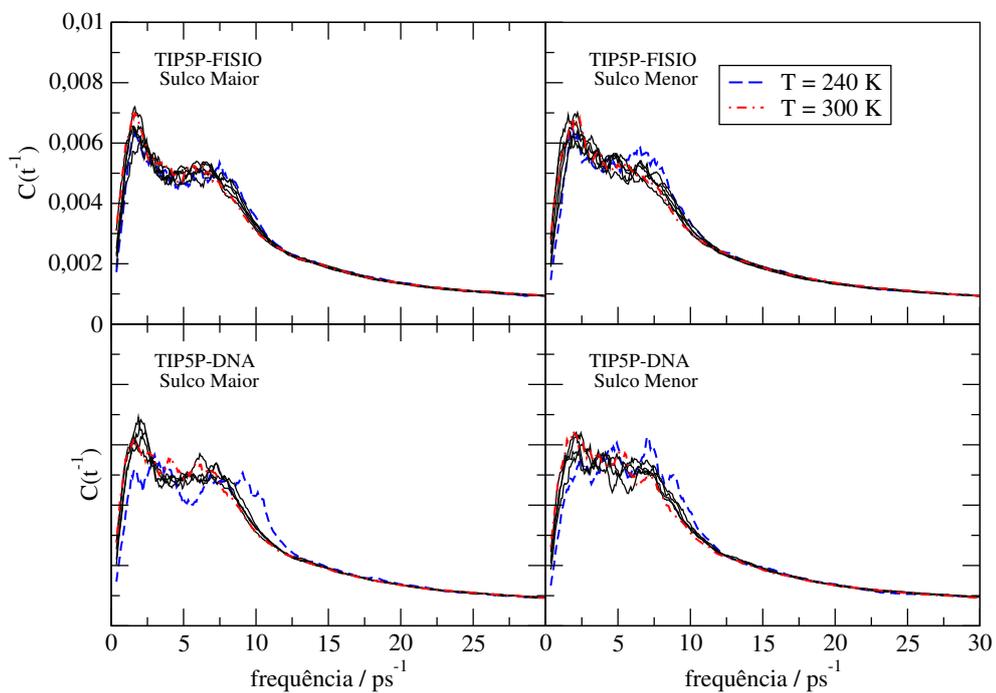


Figura 27 – Transformada de Fourier da VACF de água nos sulcos do DNA no sistema TIP5P-FISIO e TIP5P-DNA.

## 5.5 Autodifusão

A autodifusão está intimamente relacionada com a mobilidade das moléculas no sistema. Na Figura 28 é mostrado o efeito da temperatura sobre o valor do coeficiente de autodifusão para os sistemas TIP3P-BULK e TIP3P-DNA.

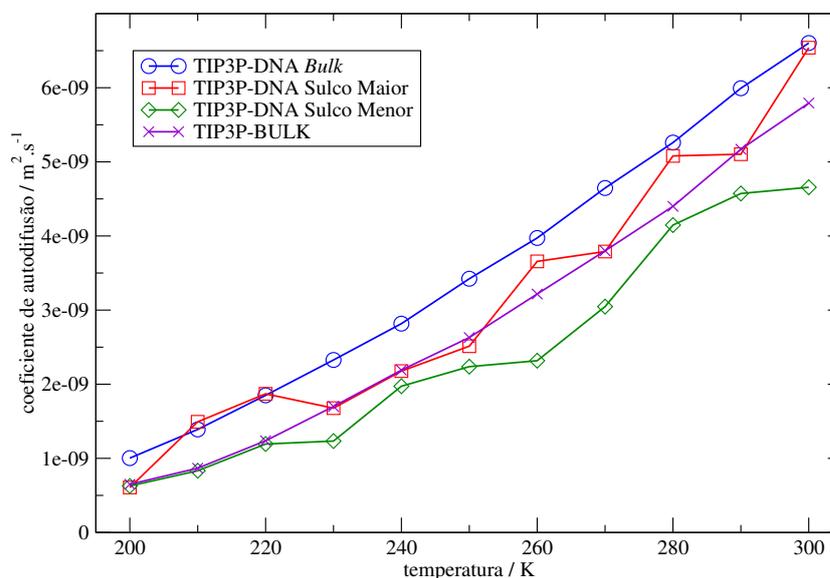


Figura 28 – Coeficiente de autodifusão da água em diferentes ambientes para diferentes temperaturas no sistema TIP3P-DNA e TIP3P-BULK

A água nos sulcos tem menor mobilidade que a água *bulk* e o sulco menor tem menor mobilidade do que no sulco maior. Por fim, sob resfriamento, as diferenças de mobilidade entre os ambientes químicos se torna cada vez menor.

Na Figura 29 mostra-se os valores obtidos para o coeficiente de autodifusão da água *tip5p* nos diferentes ambientes químicos dos sistemas estudados.

Assim como o que é constatado para o modelo *tip3p*, o coeficiente de autodifusão para água *bulk* nos sistemas TIP5P-FISIO e TIP5P-DNA, é ligeiramente maior do que para o sistema de referência sem o DNA, o sistema TIP5P-BULK. Desta forma, assim, como o DNA não afeta a estrutura do solvente, também não afeta as suas propriedades dinâmicas em longas distâncias, sendo esta propriedade dinâmica indiferente à presença dos eletrólitos do sistema TIP5P-FISIO.

Constata-se que a mobilidade das moléculas de água nos sulcos é muito inferior à da água *bulk* para temperatura elevadas maiores que 260 K. Assim, o modelo *tip5p* representa melhor a observação experimental de que a água nos sulcos tem menor mobilidade.

Em temperaturas mais altas, verificamos que a água *bulk* tem sua difusão mais influenciada pela temperatura do que a água nos sulcos. Em temperaturas abaixo de 260 K, porém, a dinâmica do solvente é influenciada pela temperatura independente de estar

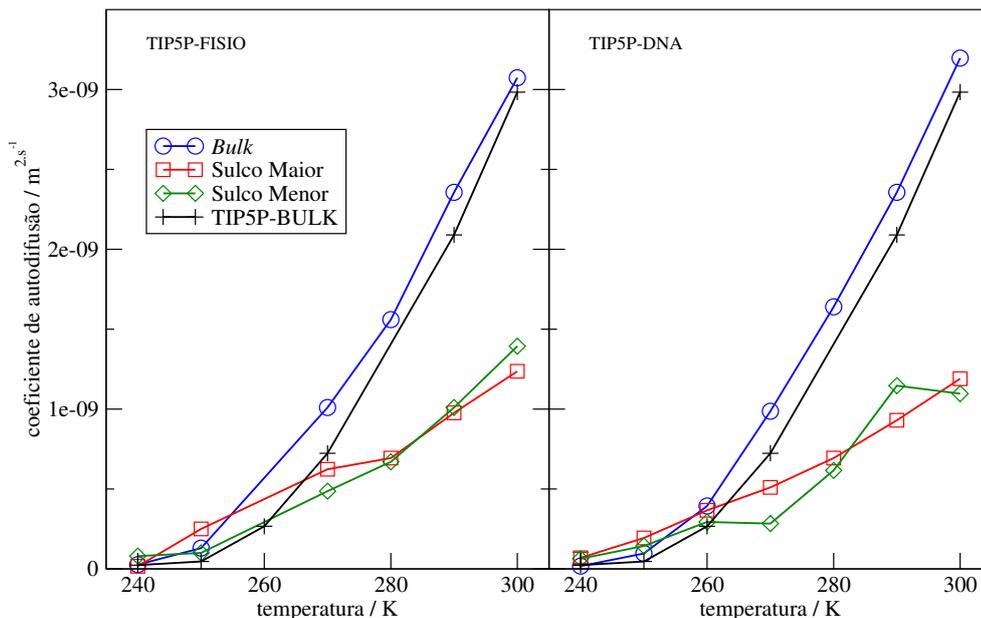


Figura 29 – Coeficiente de autodifusão da água em diferentes ambientes para diferentes temperaturas para os sistemas TIP5P-FISIO, TIP5P-DNA e TIP5P-BULK.

em um dos sulcos ou estar no *bulk*.

Os valores encontrados para o sistema TIP5P-FISIO e TIP5P-H2O são essencialmente os mesmos, de forma que podemos considerar que o solvente que se encontra nos sulcos não sofre influência dos demais solutos iônicos no sistema também no que diz respeito à propriedades de transporte também nos sulcos.

Comparando a mobilidade do solvente nos diferentes sulcos, não é verificada grandes diferenças entre os valores do coeficiente de autodifusão entre os sulcos maior e menor. Nas altas temperaturas a diferença parece ser um pouco menor, com a mobilidade no sulco menor sendo não expressivamente mais rápida, mas a diferença não é expressiva, principalmente considerando a baixa estatística com que a análise de águas no sulcos é realizada.

O comportamento dinâmico do sistema pode ser mais facilmente visualizado em um diagrama de Arrhenius, isto é, em um gráfico que relaciona o logaritmo da propriedade dinâmica (neste caso, o coeficiente de autodifusão) como uma função do inverso da temperatura (31). A Figura 30 mostra este gráfico para os sistemas estudados.

Analisando a Figura 30 verificamos que os sistemas TIP3P-DNA e TIP3P-BULK apresentam um comportamento do tipo Arrhenius para todas as temperaturas estudadas. Não é verificada uma transição vítrea dinâmica. O que corresponde a uma baixa energia de ativação, ou parâmetro A, da Equação 2.1.

Para a água *bulk* com modelos *tip5p*, no entanto, verifica-se essa transição em uma temperatura de aproximadamente 260 K. Íons dissolvidos não influenciam a transição,

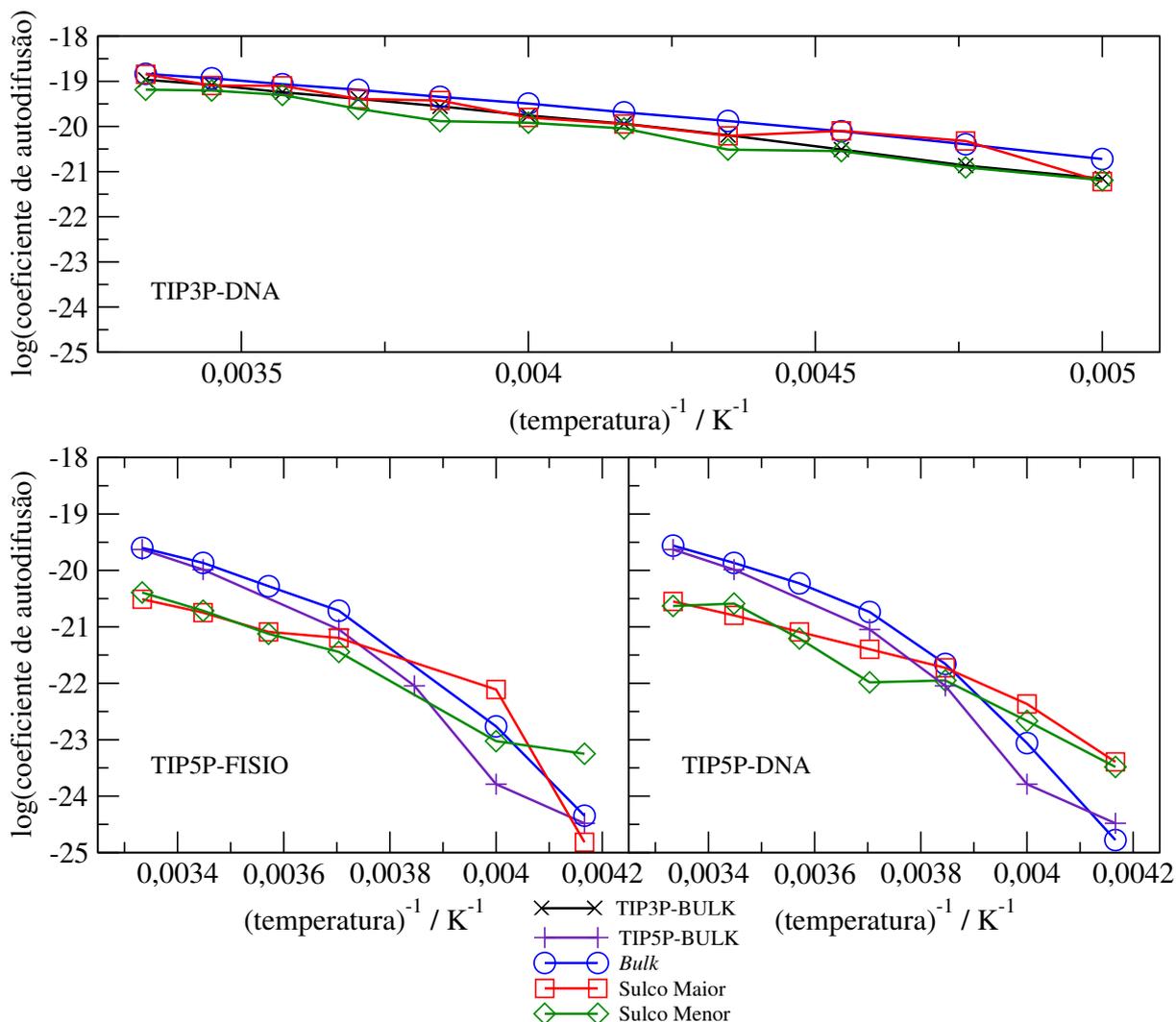


Figura 30 – Diagrama de Arrhenius do coeficiente de autodifusão da água em diferentes ambientes químicos para os sistemas estudados.

porém a presença do DNA em si causa uma pequena influência, ela torna a transição um pouco mais abrupta do que no sistema TIP5P-BULK. Notamos também que, para estes sistemas, a água nos sulcos a transição é menos abrupta e deslocada para temperatura um pouco mais baixas (em torno de 250 K).

## 5.6 Análise de Ligações de Hidrogênio

### 5.6.1 Análise do Ângulo Doador-Hidrogênio-Aceptor

Avaliou-se a distribuição dos ângulos da ligações de hidrogênio em função da temperatura. Os resultados são mostrados nas Figuras 31 a 34. Nota-se que para a água *bulk* o máximo da distribuição se desloca, com o resfriamento, para ângulos mais baixos, de 11 para 8 graus. Com o abaixamento da temperatura as moléculas envolvidas nesta

interação podem se manter em posições mais estáveis, com ângulos mais baixos. Do ponto de vista da interação entre o hidrogênio e o átomo aceptor, conforme a ligação tende a se tornar linear, pode ocorrer uma melhor interação entre a densidade eletrônica do aceptor e o orbital  $\sigma$  aintiligante do hidrogênio. Embora simulações clássicas não possam representar esta interação eletrônica, os modelos clássicos foram parametrizados visando reproduzir propriedades estruturais reais da água. Assim, a dinâmica clássica é capaz de reproduzir este efeito.

Para o modelo *tip5p* esta mudança é mais acentuada do que no modelo mais simples. Isto ocorre pois, quando o oxigênio de uma molécula de água *tip5p* é aceptor de uma ligação de hidrogênio, o par de elétrons (representada pelo átomo *dummy*), tem sua posição bem definida no espaço, favorecendo uma interação eletrostática linear.

A água que se encontra nos sulcos do DNA sofre pequena influência da temperatura. Embora se esperasse que com o abaixamento da temperatura as ligações de hidrogênio se tornassem mais próximas de lineares, mas há imposição geométrica pelos sulcos. Consistentemente com esta explicação, constata-se que, no sulco menor esta variação é menor do que no sulco maior.

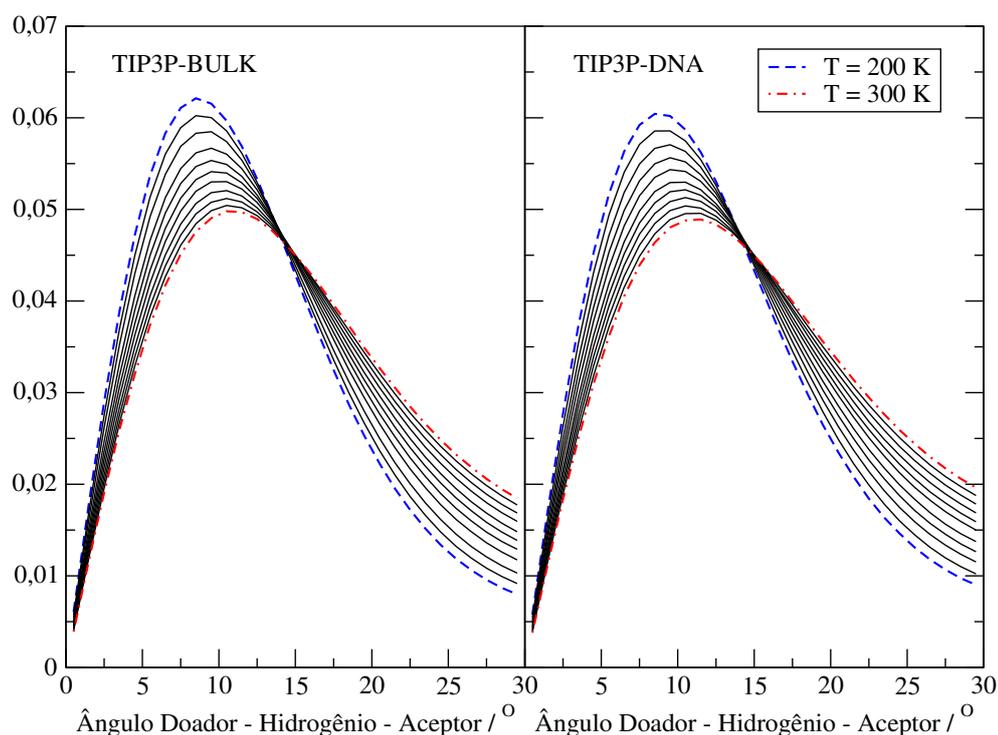


Figura 31 – Efeito da temperatura sobre a distribuição de ângulo Doador-Hidrogênio-Aceptor no *bulk* do sistema TIP3P-DNA e no sistema TIP3P-BULK.

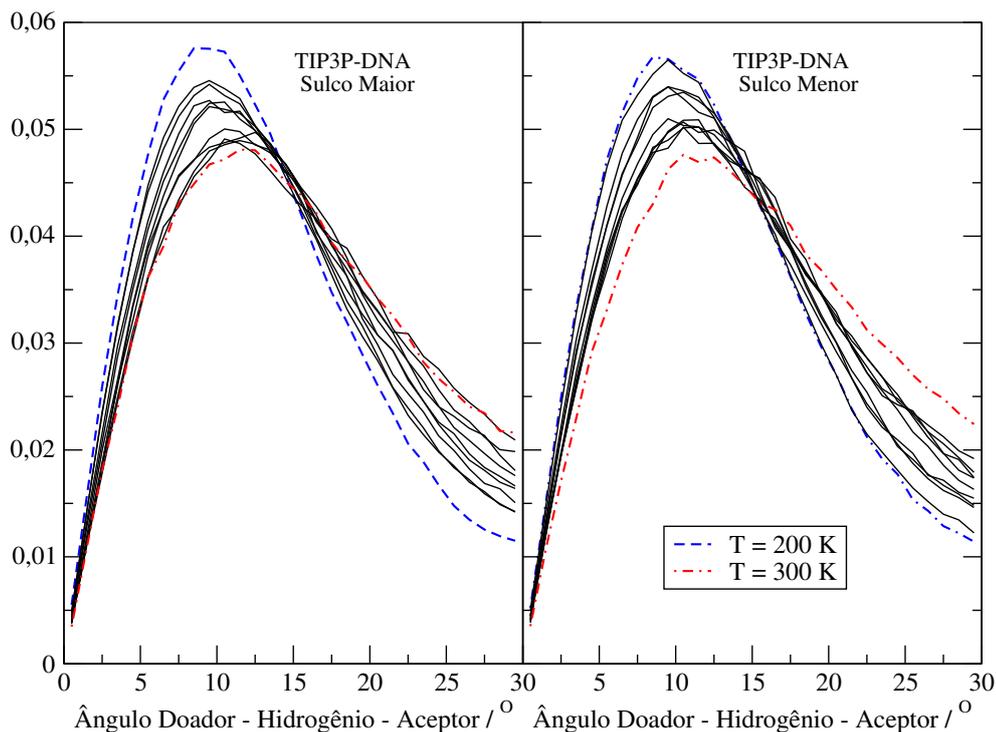


Figura 32 – Efeito da temperatura sobre a distribuição de ângulo Doador-Hidrogênio-Acceptor nas ligações de hidrogênio entre o solvente e átomos dos sulcos do DNA.

### 5.6.2 Análise Contínua

As Figura 35 à 38 mostram as funções de autocorrelação de ligações de hidrogênio para o *bulk* e para os sulcos dos sistemas.

Qualitativamente, verificamos que em temperaturas elevadas a correlação é perdida rapidamente, porém se mantém por períodos maiores conforme o sistema é resfriado. Não há diferença significativa no comportamento destas funções para o sistema TIP3P-BULK e TIP3P-DNA. O decaimento da função de autocorrelação com o tempo é do tipo exponencial, podendo ser descrito como:

$$C(t) = \exp\left(-\frac{t}{\tau_{HB}(T)}\right), \quad (5.1)$$

$\tau_{HB}(T)$  é o tempo de relaxamento do processo. Infelizmente, porém, para temperaturas baixas o tempo de vida do processo se torna muito alto, tornando proibitivo um *fit* para determinar o seu valor.

### 5.6.3 Análise Intermitente

Os tempos de vida das ligações de hidrogênio é mostrada na Figura 39.

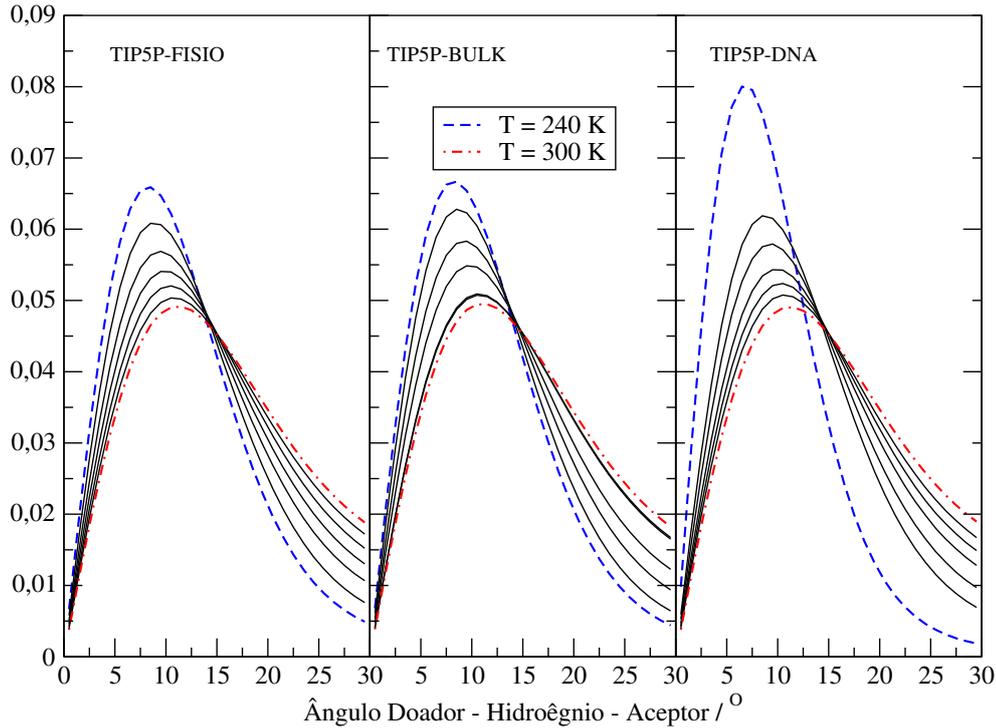


Figura 33 – Efeito da temperatura sobre a distribuição de ângulo Doador-Hidrogênio-Acceptor no *bulk* do sistemas TIP5P-FISIO e TIP5P-DNA e no sistema TIP5P-BULK.

Analisando a Figura 40, vemos que, analogamente à análise feita na Figura 30, há uma mudança no comportamento dinâmico do sistema para os sistemas cuja hidratação é feita com água do tipo *tip5p* - utilizando esta métrica o modelo *tip3p* também não apresenta vitrificação.

Comparando as figuras 40 e ??, vemos que a água *bulk* apresenta mudança de comportamento do tipo Arrhenius para VF em torno de 270 K, assim como quando analisada por meio da constante de autodifusão. Nos sulcos, porém, a mudança de comportamento é ainda mais drástica e ocorre de forma bem definida na temperatura de 250 K. Simulações mais longas são necessárias, porém, aparentemente a mudança da declividade para a água nos sulcos é característica de uma transição vítrea termodinâmica completa ( $T_0$  próximo de 250 K) com  $\log(lifetime)$  tendendo ao infinito, a passo que a água *bulk* apenas tem uma mudança de declividade (transição dita dinâmica).

Curiosamente, esta transição não é tão perceptível na Figura 30. Possivelmente, nas baixas temperaturas as moléculas de água oscilam em torno de posições, sem a cisão da ligação de hidrogênio. Assim, ainda há a dinâmica rápida de relaxamento  $\beta$  (*rattling in cage*) que contribui com o coeficiente de autodifusão, (uma métrica mais sensível), mas sendo interpretada como imóvel pela métrica da ligação de hidrogênio intermitente, uma vez que a ligação não se rompe. Novamente, simulações mais longas e em temperaturas mais baixas seriam interessantes para elucidar estas observações. Esta observação também

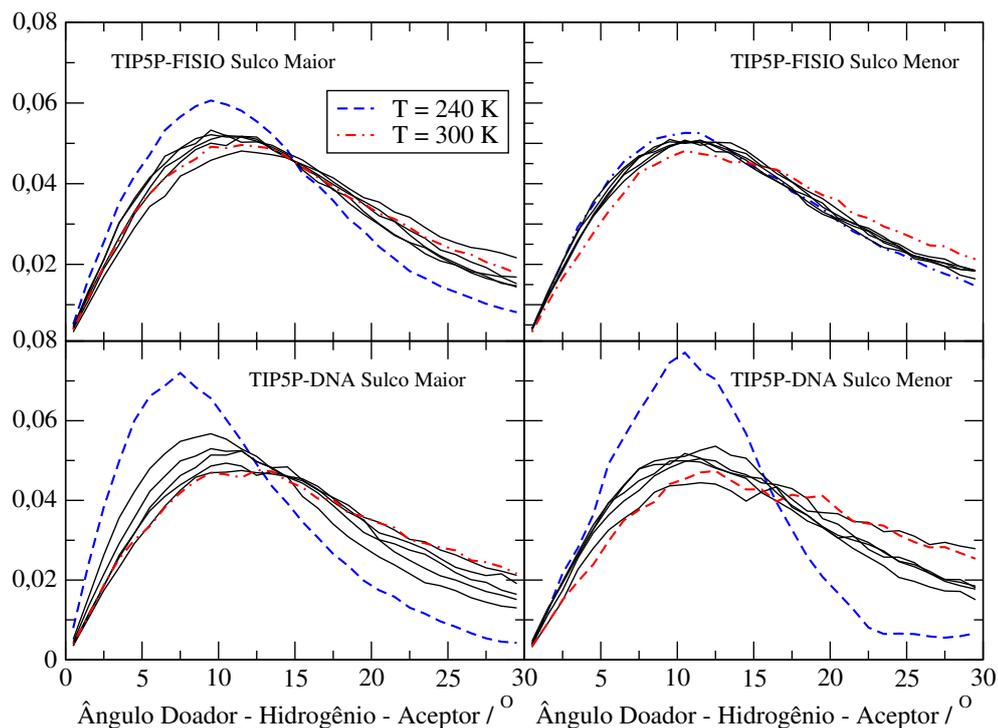


Figura 34 – Efeito da temperatura sobre a distribuição de ângulo Doador-Hidrogênio-Aceptor nas ligações de hidrogênio entre o solvente e átomos dos sulcos do DNA para os sistemas TIP5P-FISIO e TIP5P-DNA.

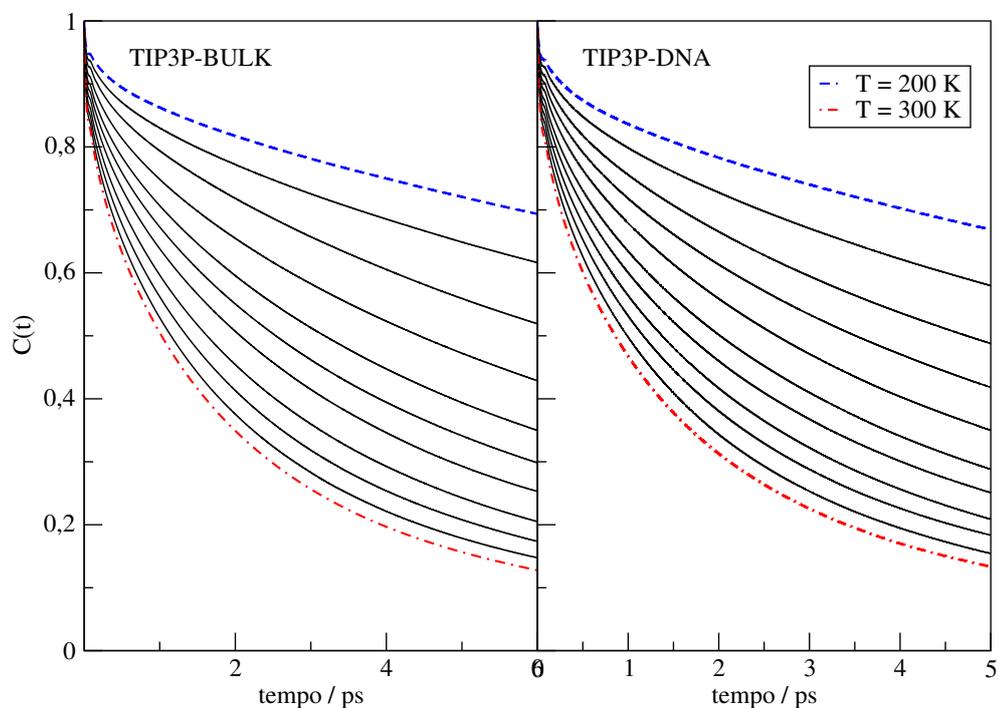


Figura 35 – Função de autocorrelação contínua de ligações de hidrogênio em diferentes temperaturas para moléculas de água no *bulk* dos sistemas TIP3P-DNA e TIP3P-BULK.

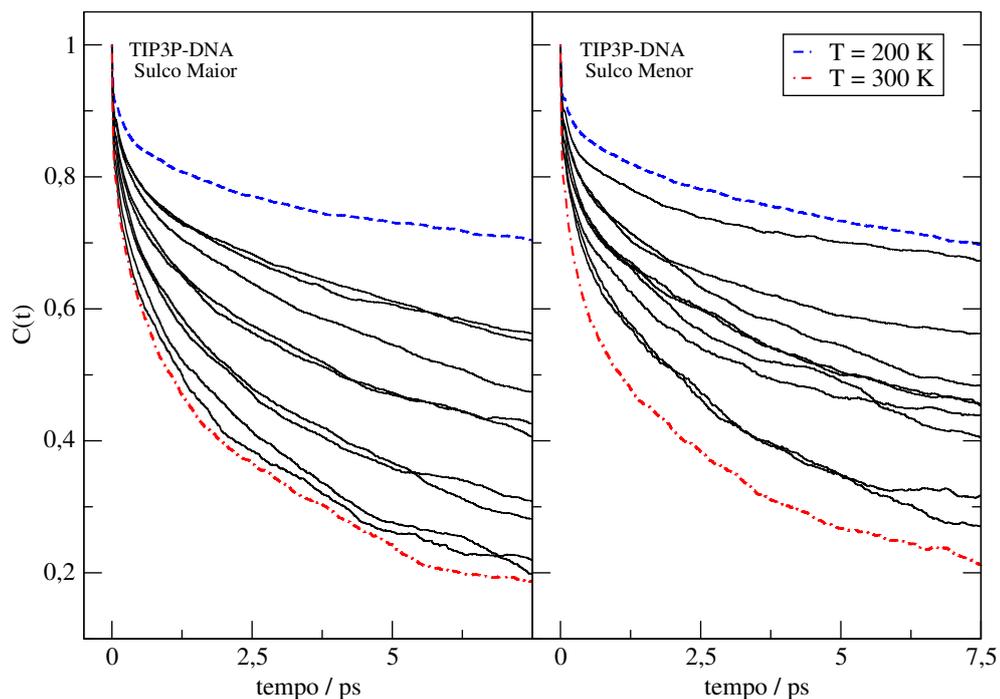


Figura 36 – Função de autocorrelação contínua de ligações de hidrogênio entre átomos dos sulcos do DNA e moléculas de água que se encontravam próximas dos mesmo no início da análise para o sistema TIP3P-DNA em diferentes temperaturas.

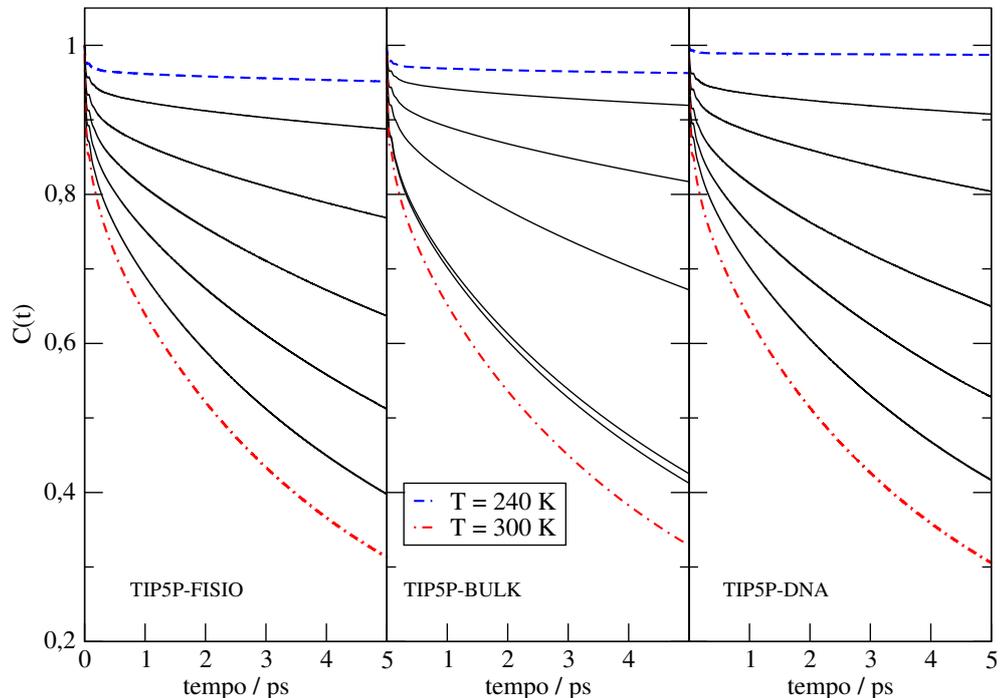


Figura 37 – Função de autocorrelação contínua de ligações de hidrogênio em diferentes temperaturas para moléculas de água no *bulk* dos sistemas TIP5P-FISIO, TIP5P-DNA e TIP5P-BULK

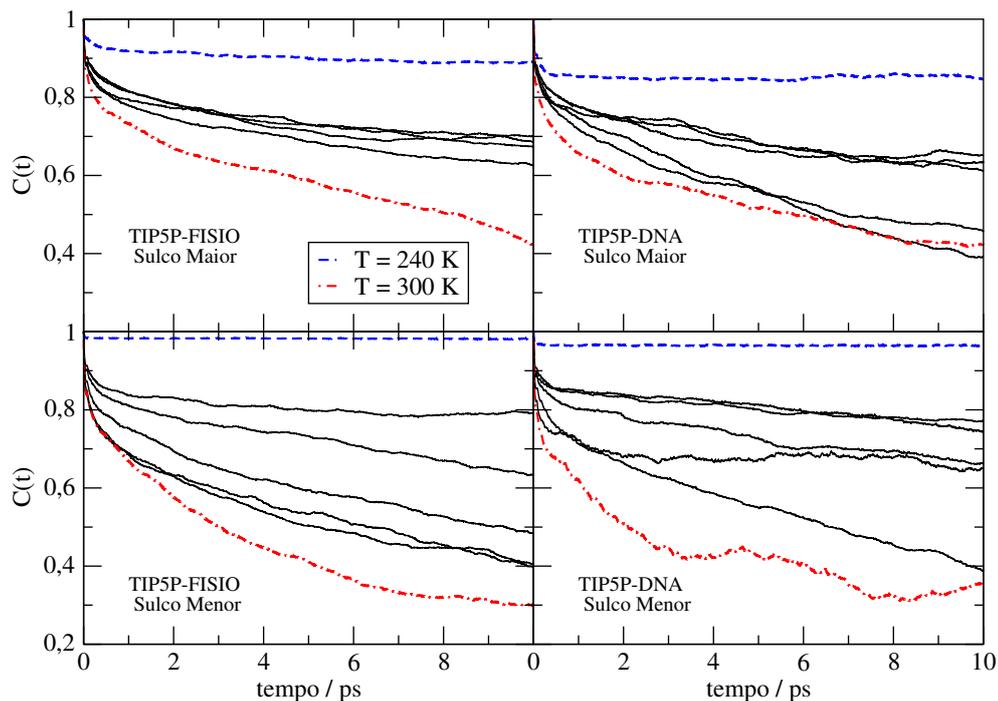


Figura 38 – Função de autocorrelação contínua de ligações de hidrogênio entre átomos dos sulcos do DNA e moléculas de água que se encontravam próximas dos mesmos no início da análise para os sistemas TIP5P-FISIO e TIP5P-DNA em diferentes temperaturas.

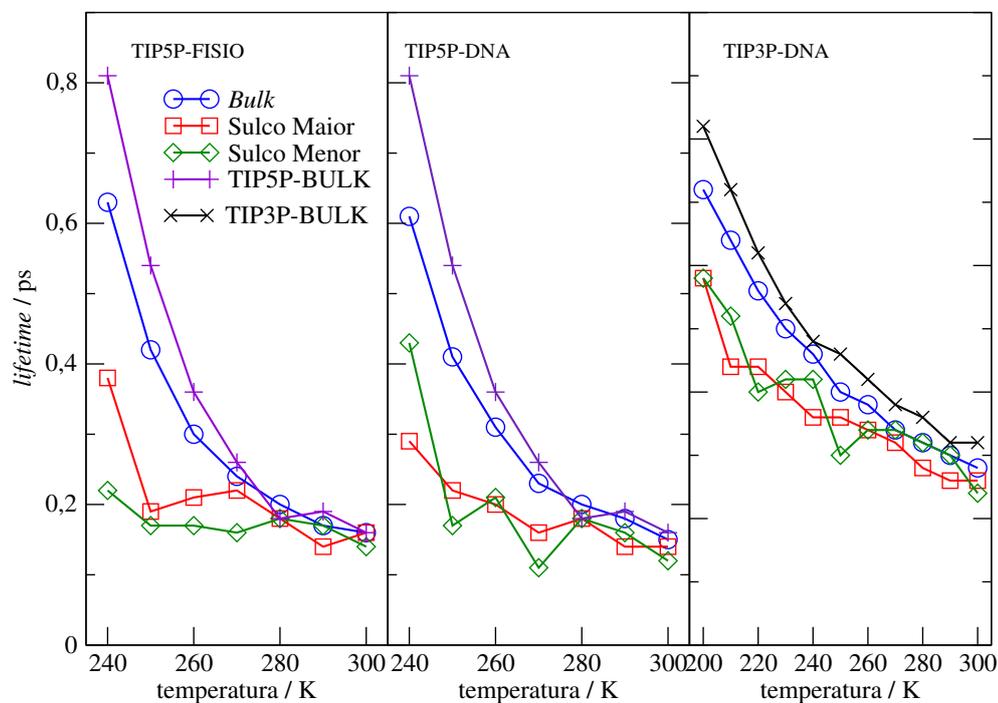


Figura 39 – Influência da temperatura sobre o tempo de vida, *lifetime*, das ligações de hidrogênio para água *bulk* e para a espinha de hidratação para todos os sistemas estudados.

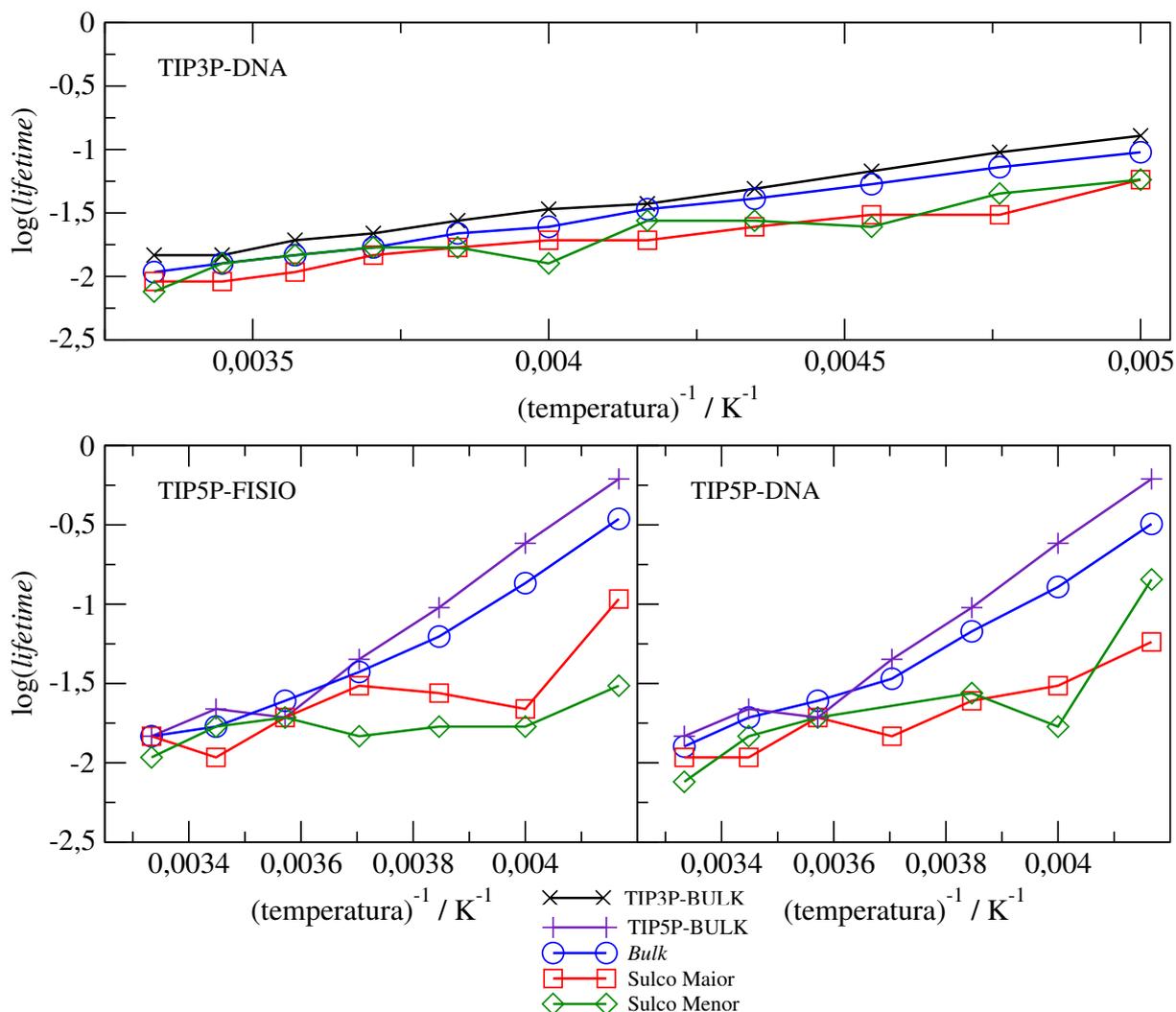


Figura 40 – Arrhenius *plot* do tempo de vida, *lifetime*, das ligações de hidrogênio entre átomos de água *bulk*, e de água nas espinhas de hidratação de DNA para todos os sistemas estudados.

pode ser um artefato do modelo de água empregado, que possui átomos *dummy*, que não contribuem com a difusão explicitamente, (e nem com a definição da temperatura, uma vez que são desprovidos de massa), porém facilitam a formação de ligações de hidrogênio mais duradouras.

#### 5.6.4 Influência da Configuração Inicial

Como exposto na descrição metodológica, ligações de hidrogênio possuem tempos de vida curtos. Uma vez que o número de moléculas água que de fato estão no sulco constituem uma fração muito pequena do sistema (de 20 a 30 unidades), o tempo de vida médio das ligações de hidrogênio determinado pode sofrer efeito da configuração instantânea no qual foi realizada a análise.

Este efeito foi avaliado em um sistema. Para o sistema TIP5P-DNA, na temperatura de 240 K, obteve-se da trajetória equilibrada 5 configurações instantâneas, denominadas 20000, 25000, 30000, 35000 e 40000, conforme o tempo, em ps, em que ocorreram. Cada uma dessas configurações foi empregada como configuração de partida para uma simulação detalhada no *ensemble* NVE.

Nota-se uma significativa influência por parte da configuração inicial para realizar esta análise. Esta influência também pode ser constatada nas funções de auto-correlação das ligações de hidrogênio, como mostrados na Figura 41.

Nesta Figura, vemos, através da análise de RMSD do DNA que no intervalo considerado a biomolécula não sofreu nenhuma transformação estrutural. Configurações desta trajetória foram tomadas como ponto de partida para as simulações NVE. A análise estrutural da trajetória assim gerada mostra que os pontos de partida são equivalentes, como mostrado pelo  $g(r)$  obtido.

As funções de autocorrelação de velocidades do átomos OW apresenta comportamento muito similar para os diferentes pontos de partida. O comportamento da água *bulk* é essencialmente insensível, porém as águas presentes nos sulcos tem a mesma forma qualitativa entre as simulações, mas a posição exata dos máximos e mínimos da função sofre um pequeno deslocamento.

A análise da função de autocorrelação contínua das ligações de hidrogênio, no entanto, é fortemente influenciada. Qualitativamente, em todas ocorre a rápida evolução de um patamar, indicando que, em média as ligações de hidrogênio se desfazem e refazem. Porém, o valor em que o patamar ocorre é bastante variável.

Por fim, avaliando a dinâmica das ligações de hidrogênio por meio da análise intermitente encontramos uma influência muito grande. O tempo de vida para a ligação de hidrogênio foi calculado e obteve-se os mostrados na Tabela 1.

Tabela 1 – Tempo de vida das ligações de hidrogênio,  $\tau_{HB}$ , em ps, para diferentes configurações iniciais para a água *bulk* e nos sulcos do DNA.

Configuração Inicial	<i>Bulk</i>	Sulco Maior	Sulco Menor
20000	0,63	0,22	0,30
25000	0,63	0,16	0,75
30000	0,61	0,31	0,16
35000	0,60	0,50	0,49
40000	0,56	0,24	0,29
Média	0,61	0,29	0,40
Desvio Padrão	0,03	0,13	0,23

Para a obtenção de resultados mais precisos e confiáveis o procedimento adequado seria realizar uma média de *ensemble*. Isto é, tomar um número suficientemente grande

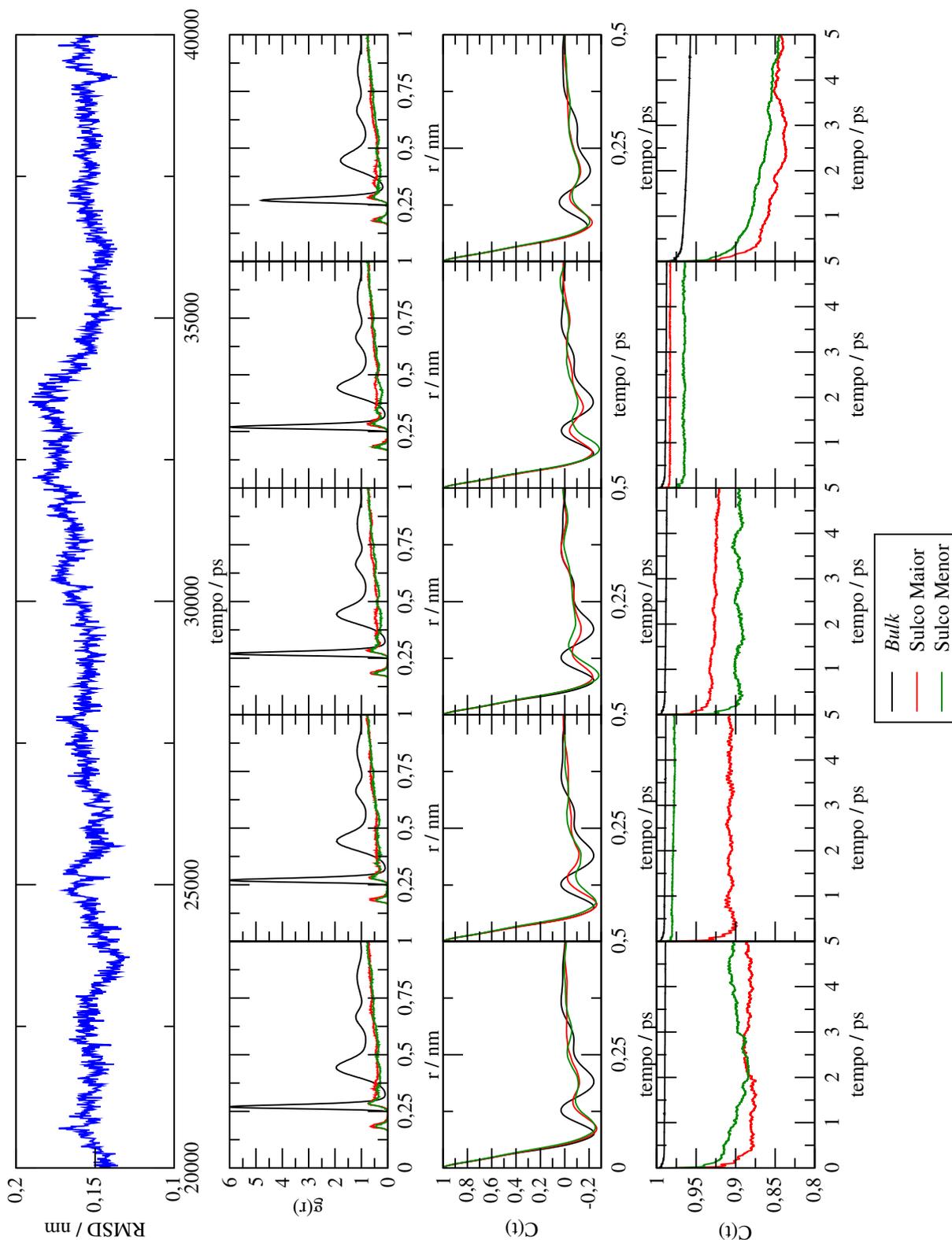


Figura 41 – Diferentes análises realizadas no TIP5P-DNA sob temperatura de 240K para avaliar o efeito da configuração instantânea inicial. Análises, de cima para baixo: RMSD,  $g(r)$  função de autocorrelação de velocidades OW, função de autocorrelação de ligações de hidrogênio.

de pontos de partida a partir de uma trajetória equilibrada nas condições de interesse, e utilizá-las como ponto de partida para a análise das ligações de hidrogênio, e fazer-se uma média sobre todos os resultados.

Avalia-se que para este projeto o custo computacional é proibitivo para a realização da análise da maneira mais adequada.

A análise que foi feita das ligações de hidrogênio para todos os demais sistemas teve apenas uma única configuração instantânea como ponto de partida. Por esta razão, os resultados obtidos não podem ser considerados resultados precisos; servem apenas como guias do comportamento geral dos sistemas. E, mesmo neste âmbito, conclusões assim obtidas devem ser alvo de posteriores estudos e escrutínio, principalmente quanto às análises referentes ao sulco menor, onde o menor volume leva a uma menor quantidade de solvente e, assim, estatística ainda mais baixa.

## 6 Conclusão

Motivado por evidências experimentais, foram realizadas extensivas simulações clássicas de dinâmica molecular de soluções aquosas de DNA visando investigar a transição vítrea. Estudou-se o dodecâmero de Dickerson solvatado por moléculas de água do modelo TIP5P e TIP3P em temperaturas de 300 à 240 K e de 300 à 200 K, respectivamente.

Mostrou-se que o modelo de água TIP3P não é capaz de representar adequadamente as observações experimentais, tanto do que diz respeito à transição vítrea como no comportamento dinâmico do solvente nos sulcos. Uma observação interessante, visto que é o modelo mais empregado nas simulações biológicas e, principalmente, porque o solvente aquoso desempenha um papel fundamental na interação de ligantes com os sulcos do DNA.

O modelo TIP5P foi capaz de mostrar transição vítrea dinâmica. O comportamento dinâmico do solvente, monitorado tanto pelo coeficiente de difusão como pelo tempo de vida das ligações de hidrogênio, é distinto entre o *bulk* do solvente e aquele que fica nos sulcos. O *bulk* sofre transição independente do DNA, de forma que, ao contrário das proteínas anticongelantes, o DNA não afeta o solvente à longas distâncias. De mesma forma, a adição de íons à solução não influencia significativamente o sistema.

A água nos sulcos tem sua transição vítrea deslocadas para temperaturas mais baixas (250 K) e, quando esta ocorre, ela é menos acentuada. Assim, os sulcos do DNA parecem proteger o solvente, tornando-o menos sensível a variação ao resfriamento em baixas temperaturas. Isto pode favorecer a transição vítrea termodinâmica da espinha de hidratação, pois, como a água nos sulcos permanece mais móvel, não permitiria a formação de núcleos adequados para a cristalização.

Concluimos que as análises dinâmicas de eventos rápidos é muito sensível à configuração momentânea tomada como inicial, sendo necessário tomar uma amostragem grande de dados para uma análise mais confiável.

## 7 Perspectivas

Como futura perspectiva para este trabalho, temos a continuada refrigeração dos sistemas TIP5P-DNA mantendo a mesma metodologia, ou, de forma mais viável do ponto de vista computacional, a implementação de uma metodologia Monte Carlo, com cálculo das propriedades dinâmicas por meio da entropia de excesso do sistema e comparação com os dados já obtidos, para posterior extrapolação nas temperaturas mais baixas.

Do trabalho já realizado pretende-se recalculer os coeficiente de auto-difusão por meio da fórmula de Einstein. Pretende-se também obter o fator de estrutura estático a partir das dunções distribuição radial, que pode mais prontamente ser comparado com dados experimentais de espalhamento de baixo ângulo de nêutrons.

Deve-se desenvolver uma rotina de análise das propriedades dinâmicas rápidas, como das ligações de hidrogênio, que permita uma amostragem maior de configurações momentâneas iniciais.

## Referências

- 1 LOERTING, T.; GIOVAMBATTISTA, N. Amorphous ices: experiments and numerical simulations. *Journal of Physics: Condensed Matter*, v. 18, p. R919–R977, 2006. Citado 3 vezes nas páginas 12, 14 e 16.
- 2 ANGELL, C. A. *Annu. Rev. Phys. Chem.*, v. 55, 2004. Citado 2 vezes nas páginas 12 e 15.
- 3 SONODA, M. T. et al. A review on the dynamics of water. *Brazilian Journal of Physics*, v. 34, n. 1, p. 3–16, 2004. Citado na página 12.
- 4 JABES, B. S.; AGARWAL, M.; CHAKRAVARTY, C. Tetrahedral order, pair correlation entropy, and waterlike liquid state anomalies: Comparison of geo2 with bef2, sio2, and h2o. *The Journal of Chemical Physics*, v. 132, p. 234507, 2010. Citado na página 12.
- 5 HEAD-GORDON, S. *J. Phys. Chem.*, v. 98, 1193. Citado na página 12.
- 6 DEBENEDETTI, P. G. *Journal of Physics: Condensed Matter*, v. 15, 2003. Citado 3 vezes nas páginas 12, 14 e 15.
- 7 CHAPLIN, M. *Water Structure and Science* (<http://www.lsbu.ac.uk/water/>). 2013. [Online; accessed 21-Maio-2013]. Disponível em: <<http://www.lsbu.ac.uk/water/>>. Nenhuma citação no texto.
- 8 NETZ, P. A.; BARBOSA, M. C.; STANLEY, H. E. *J. Phys. Chem.*, v. 115, 2001. Citado na página 12.
- 9 ERRINGTON, J. R.; DEBENEDETTI, P. G. *Nature*, v. 409, 2001. Citado 2 vezes nas páginas 12 e 29.
- 10 ANGELL, C. A. Liquid fragility and the glass transition in water and aqueous solutions. *Chemical Reviews*, v. 102, p. 2627–2650, 2002. Citado na página 12.
- 11 PACIARONI, A. et al. Glassy character of dna hydration water. *The Journal of Physical Chemistry B*, v. 117, p. 2026–2031, 2013. Citado 2 vezes nas páginas 13 e 20.
- 12 NETZ, P. A. et al. Phase diagram and waterlike anomalies in core-softened shoulder-dumbbell complex fluids. In: MORENO-PIRAJÁN, J. C. (Ed.). *Thermodynamics – Physical Chemistry Of Aqueous Systems*. Croatia: InTech, 2011. p. 391–406. Citado na página 14.
- 13 GUNSTEREN, W. F. van et al. *Computer Physics Communications*, v. 91, 1995. Citado 2 vezes nas páginas 14 e 21.
- 14 FRANZESE, G.; VILASECA, P. Isotropic soft-core potentials with two characteristic length scales and anomalous behaviour. *Journal of Non-Crystalline Solids*, v. 357, p. 419–426, 2011. Citado na página 15.
- 15 BURTON, E. F.; OLIVER, W. F. *Nature*, v. 135, 1935. Citado na página 15.

- 16 MEYER, E. *J. Appl. Phys.*, v. 58, 1985. Citado na página 15.
- 17 HALBRUNCKER, A.; MEYER, E.; JOHARI, J. P. *J. Phys. Chem.*, v. 93, 1989. Citado na página 15.
- 18 MISHIMA, O.; CALVERT, L. D.; WHALLEY, E. *Nature*, v. 310, 1984. Citado na página 15.
- 19 SPEEDY, J. R. *J. Phys. Chem.*, v. 86, 1982. Citado na página 15.
- 20 SASTRY, S. et al. *Phys. Rev.*, v. 53, 1996. Citado na página 15.
- 21 MISHIMA, O.; STANLEY, E. *Nature*, v. 396, 1998. Citado na página 15.
- 22 POOLE, P. H. *Nature*, v. 360, 1992. Citado na página 15.
- 23 STANLEY E., e. a. *Nature: Scientific Reports*, v. 3, 2013. Citado na página 16.
- 24 CUPANE A., e. a. *J. Phys. Chem.*, v. 141, 2014. Citado na página 16.
- 25 KISELEV, S. B. *Physica A*, v. 269, 1999. Citado na página 16.
- 26 KISELEV, S. B.; ELY, J. F. *Physica A*, v. 299, 2001. Citado na página 16.
- 27 KISELEV, S. B.; ELY, J. F. *J. Chem. Phys.*, v. 116, 2002. Citado na página 16.
- 28 KISELEV, S. B.; ELY, J. F. *J. Chem. Phys.*, v. 118, 2003. Citado na página 16.
- 29 BERENDSEN, T. A. *Nature Medicine*, v. 20, 2014. Citado na página 16.
- 30 MALLAJOSYULA, S.; VANOMMESLAEGHE, K.; MACKERELL, A. *J. Phys. Chem. B*, 2014. Citado na página 16.
- 31 LEUZZI, L.; MIEUWENHUIZEN, T. M. *Thermodynamics of the glassy state*. 1. ed. [S.l.]: Taylor - Francis, 2008. (Series in Condensed Matter Physics). ISBN 978-0-7503-09997-4. Citado 2 vezes nas páginas 17 e 56.
- 32 BINDER, K. Simulations clarify when supercooled water freezes into glassy structures. *PNAS*, v. 111, n. 26, p. 9374–9375, 2014. Citado na página 17.
- 33 ANGELL, G. A. et al. Nonexponential relaxations in strong and fragile glass formers. *J. Chem. Phys.*, v. 99, n. 5, p. 4201–4209, 1993. Citado na página 17.
- 34 VOET, D.; VOET, J. G.; PRATT, C. W. *Fundamentos de Bioquímica*. 1. ed. Porto Alegre: Artes Médicas Sul, 2000. Citado 2 vezes nas páginas 17 e 18.
- 35 NISHIURA, J. *Biology 4 Lectures*. 2013. [Online; accessed 24-Maio-2013]. Disponível em: <<http://academic.brooklyn.cuny.edu/biology/bio4fv/page/molecular%20biology/dna-structure.html>>. Citado na página 18.
- 36 WATSON, J. D.; CRICK, F. H. C. Molecular structure of nucleic acids. *Nature*, v. 171, p. 737–738, 1953. Citado na página 17.
- 37 PHILLIPS, D.; WOOLLINS, J. D.; DAVIES, A. G. *Nucleic Acids*. The Royal Society of Chemistry, 2004. (Tutorial Chemistry Texts). ISBN 978-0-85404-481-8. Disponível em: <<http://dx.doi.org/10.1039/9781847551542>>. Citado na página 18.

- 38 RICCI, C. G.; NETZ, P. A. *A, B e Z-DNA: Caracterização Conformacional e Importância Biológica*. Citado na página 19.
- 39 FEIG, M.; PETTITT, B. M. A molecular simulation picture of dna hydration around a- and b-dna. *Biopolymers*, v. 48, p. 199–209, 1998. Citado na página 19.
- 40 DENISOV, V. P.; CARLSTRÖM, G.; VENU, K. Kinetics of dna hydration. *Journal of Molecular Biology*, v. 268, p. 118–136, 1997. Citado 3 vezes nas páginas 19, 20 e 28.
- 41 JANA, B. et al. Dynamic and thermodynamic anomalies of water at low temperature: from bulk water to reverse micelles and dna hydration layer. *Current Science*, v. 101, n. 7, p. 900–922, 2011. Citado na página 19.
- 42 FRENKEL, D.; SMIT, B. *Understanding Molecular Simulation: From algorithms to applications*. 1. ed. San Diego: Academic Press Limited, 1996. Citado 4 vezes nas páginas 21, 27, 29 e 39.
- 43 MACKERELL, A. D. Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, v. 25, n. 13, p. 1584–1604, 2004. Citado na página 22.
- 44 BROOKS, B. R. et al. Charmm: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, v. 4, p. 187–217, 1983. Citado na página 22.
- 45 CORNELL, W. D. et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, v. 117, n. 19, p. 5169–5428, 1995. Citado na página 22.
- 46 GUNSTEREN, W. F. van; BERENDSEN, J. C. *Groningen Molecular Simulation (GROMOS) Library Manual*. [S.l.]: Biomos, 1987. 1-221 p. Citado na página 22.
- 47 JORGENSEN, W. L.; TIRADO-RIVES, J. The oplis [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, v. 110, n. 6, p. 1657–1666, 1988. Citado na página 22.
- 48 KARPLUS, M.; MACCAMMON, J. A.; GERLIN, B. R. *Nature*, v. 267, 1977. Citado na página 22.
- 49 PÉREZ, A. et al. Refinement of the amber force field for nucleic acids: Improving the description of alpha/gamma conformers. *Biophysical Journal*, v. 92, p. 3817–3829, 2007. Citado 2 vezes nas páginas 22 e 32.
- 50 BERENDSEN, H. J. C.; GRIGERA, J. R.; STRAATSMA, T. P. The missing term in effective pair potentials. *The Journal of Physical Chemistry*, v. 91, n. 24, p. 6269–6271, 1987. Citado na página 23.
- 51 STILLINGER, F. H.; RAHMAN, A. Improved simulation of liquid water by molecular dynamics. *The Journal of Chemical Physics*, v. 60, n. 4, p. 1545–1557, 1974. Citado na página 23.

- 52 JORGENSEN, W. L. et al. Comparison of simple potential functions for simulating liquid water. *Journal Chemical Physics*, v. 79, n. 2, p. 926–935, 1983. Citado na página 23.
- 53 MAHONEY, M. W.; JORGENSEN, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *The Journal of Chemical Physics*, v. 112, n. 20, p. 8910–8922, 2000. Citado 2 vezes nas páginas 23 e 39.
- 54 BERENDSEN, H. J. C. et al. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, v. 81, n. 8, p. 3684–3690, 1984. Citado 3 vezes nas páginas 24, 33 e 42.
- 55 PARRINELLO, M.; RAHMAN, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, v. 52, n. 12, p. 7182–7190, 1981. Citado 2 vezes nas páginas 24 e 33.
- 56 LU, X.-J.; OLSON, W. K. 3dna: a software for the analysis, rebuilding and visualization of three-dimensional nucleic structures. *Nucleic Acids Research*, v. 31, n. 17, p. 5108–5121, 2003. Citado 2 vezes nas páginas 25 e 32.
- 57 PRONK, S. et al. Gromacs 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, v. 29, n. 7, p. 845–854, 2013. Citado na página 27.
- 58 RAPAPORT, D. C. Hydrogen bonds in water network organization and lifetimes. *Molecular Physics*, v. 50, n. 5, p. 1151–1162, 1983. Citado na página 27.
- 59 LUZAR, A. Resolving the hydrogen bonds dynamics conundrum. *J. Phys. Chem.*, v. 113, n. 23, 2000. Citado na página 27.
- 60 SPOEL, D. van der et al. Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media. *J. Phys. Chem. B*, n. 110, p. 4393–4398, 2006. Citado na página 27.
- 61 GREEN, M. S. Markof random process and the statistical mechanics of time-dependent phenomena ii. irreversible process in fluids. *J. Chem. Phys.*, v. 22, p. 398–413, 1954. Citado na página 29.
- 62 KUBO, R. Statistical-mechanical theory of irreversible processes. i. general theory and simple applications to magnetic and conduction problems. *J. Phys. Soc. Jpn.*, v. 12, p. 570–586, 1957. Citado na página 29.
- 63 GROMACS - Force fields. 2013. [Online; accessed 21-Maio-2013]. Disponível em: <[http://www.gromacs.org/Downloads/User\\_contributions/Force\\_fields](http://www.gromacs.org/Downloads/User_contributions/Force_fields)>. Citado na página 32.
- 64 REFINEMENT of the amber force field for nucleic acids simulations: PARMBSC0. 2013. [Online; accessed 21-Maio-2013]. Disponível em: <<http://mmb.pcb.ub.es/PARMBSC0/>>. Citado na página 32.
- 65 DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh ewald: An  $n \log(n)$  method for ewald sums in large systems. *J. Chem. Phys.*, v. 98, p. 10089–10092, 1993. Citado na página 33.

- 66 ESSMANN, U. et al. A smooth particle mesh ewald potential. *J. Chem. Phys.*, v. 103, p. 8577–8592, 1995. Citado na página 33.
- 67 Z., E. *Socratic*. 2015. [Online; accessed 4-Junho-2015]. Disponível em: <<http://socratic.org/questions/what-hydrogen-bonds-join-cytosine-to-guanine>>. Citado na página 36.
- 68 N., M.; ULAM, S. The monte carlo method. *Journal of the American Statistical Association*, v. 44, n. 247, p. 335–341, 1949. Citado na página 39.
- 69 CHAKRAVARTY, C. et al. Relationship between structure, entropy and diffusivity in water and water-like liquids. *J. Phys. Chem. B*, v. 114, n. 20, p. 6995–7001, 2010. Citado na página 40.
- 70 CHAKRAVARTY, C.; CHAKRABOORTY, N. S.; SHARMA, R. Excess entropy, diffusivity and structural order in liquids with water-like anomalies. *J. Chem. Phys.*, v. 125, n. 20, p. 204501, 2006. Citado na página 40.
- 71 THEORETICAL and Computational Biophysics Group: VMD - Visual Molecular Dynamics. [Online; accessed 5-Junho-2015]. Disponível em: <<http://www.ks.uiuc.edu/Research/vmd/>>. Citado na página 40.