

Luis Andre Baptista dos Santos

**Dinâmica Molecular Dirigida aplicada ao estudo  
de desenovelamento de um aptâmero de RNA**

Porto Alegre

2019

Luis Andre Baptista dos Santos

**Dinâmica Molecular Dirigida aplicada ao estudo de  
desenvolvimento de um aptâmero de RNA**

Tese apresentada como requisito parcial para  
a obtenção do grau de Doutor em Química.

Universidade Federal do Rio Grande do Sul - UFRGS

Instituto de Química - IQ

Programa de Pós-Graduação em Química - PPGQ

Orientador: Prof. Dr. Paulo Augusto Netz

Porto Alegre

2019

*À minha família, por todo afeto e apoio  
fornecidos durante todos estes anos.*

# Agradecimentos

Gostaria de expressar minha profunda gratidão ao Prof. Paulo A. Netz pela oportunidade de aprendizado e amizade concedidas nestes seis anos. Pela confiança em meu trabalho, pelo incentivo e tutoria que me deram o suporte necessário para a realização não só dos trabalhos de mestrado de doutorado, mas também para me tornar um profissional melhor.

Gostaria de agradecer ao Prof. Hermes Luís Neubauer de Amorim, pois foi sua amizade e orientações que me trouxeram para a área teórica da química.

Agradeço imensamente aos professores do Programa de Pós-Graduação em Química da UFRGS que tive contato e especialmente aos professores do Grupo de Química Teórica, Hubert Stassen, Paulo Gonçalves e Paolo Livotto pelos ensinamentos concedidos durante as disciplinas e conversas no GQT.

Ao estimado amigo Tiago Espinosa pela amizade e frutíferas discussões ao longo desta jornada.

Aos colegas do Grupo de Química Teórica pelas discussões, dicas e cafés.

A todos aqueles que, embora não nominados aqui, tenham contribuído para a execução deste trabalho.

Ao povo brasileiro que, através da CAPES, forneceu o apoio financeiro necessário para a realização deste trabalho.

*"Não como o caos que tanto esmaga quanto machuca.  
Mas como o mundo, harmoniosamente confuso:  
Onde a ordem na variedade nós vemos.  
E onde, embora todas as coisas difiram, todas concordam."  
Alexander Pope, 1713.*

*"Nós somos o que repetidas vezes fazemos.  
Excelência então, não é um ato, mas um hábito."  
Aristóteles*

# Resumo

Neste trabalho foi aplicada a técnica de simulação por dinâmica molecular dirigida com solvente implícito, de forma a estudar o processo de desnaturação de um aptâmero de RNA de fita dupla e de dois complexos com as moléculas de estreptomicina e bluensomicina. De forma a modelar o solvente, as equações de movimento das partículas foram modeladas segundo uma dinâmica de Langevin, como valor da constante de fricção  $\gamma$  sendo calibrado contra simulações por dinâmica molecular de equilíbrio. A dinâmica molecular dirigida foi realizada se aplicando um potencial harmônico entre o átomo O5' do resíduo C101 e um átomo virtual móvel. O átomo O5' do resíduo G1 foi mantido fixo pela aplicação de um potencial de restrição. Os parâmetros do potencial harmônico foram calibrados de forma que a curva de força-extensão gerada fosse a mais suave possível. O valor ajustado para a constante de força do potencial harmônico aplicado foi de  $k = 200,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  e a velocidade de afastamento entre os átomos foi de  $v = 0,05 \text{ nm ps}^{-1}$ . As simulações de Dinâmica Molecular Dirigida, DMD, (500 réplicas por sistema) mostraram que o processo de desnaturação da molécula de RNA por aplicação de força acontece de maneira hierárquica. A aplicação da identidade de Jarzynski aos perfis de força obtidos com as simulações de DMD mostraram que a diferença de energia livre (de não-equilíbrio) para a complexação das moléculas de estreptomicina e bluensomicina pelo RNA corresponde a  $-5,26 \pm 0,15 \text{ kJ mol}^{-1}$  e  $+5,78 \pm 0,34 \text{ kJ mol}^{-1}$ , respectivamente. As simulações de DMD e a decomposição da energia livre em energia interna e entropia evidenciam o padrão de ligações de hidrogênio formadas pela molécula de estreptomicina com o sítio de ligação que conferem estabilidade ao sítio de ligação.

**Palavras-chaves:** Dinâmica molecular Dirigida, Estreptomicina, Bluensomicina, Energia Livre.

# Abstract

In this work we used steered molecular dynamics simulations with implicit solvent model in order to study the unfolding process of a double strand RNA aptamer and the complexes with streptomycin and bluensomycin. In order to model the solvent, the equations of motion of the particles were modelled following a Langevin dynamics, with a friction constant  $\gamma$  calibrated against equilibrium molecular dynamics simulations. The steered molecular dynamics simulations were carried out applying a harmonic potential between the O5' atom of residue C101 and a virtual atom. The O5' atom of residue G1 was restrained by a harmonic potential. The parameters of the harmonic potential were calibrated in such a way that the force-extension curve obtained was the most smooth as possible. The adjusted value for the force constant of the harmonic potential applied to move the O5' atom of residue C101 was  $k = 200.0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , and the pulling velocity constant between the atoms was  $v = 0.05 \text{ nm ps}^{-1}$ . The steered molecular simulations (500 independent runs) showed that the unfolding process of the RNA aptamer under application of external force occurs in a hierarchical way. The application of the Jarzynski's equality to the force profiles obtained by steered dynamics simulations showed that the (nonequilibrium) free energy difference for the binding of the streptomycin and bluensomycin to the RNA aptamer is  $-5.26 \pm 0.15 \text{ kJ mol}^{-1}$  and  $+5.78 \pm 0.34 \text{ kJ mol}^{-1}$ , respectively. The steered molecular dynamics simulations carried out and the free energy partitioning in internal energy and entropy showed that the pattern of hydrogen bonds between the streptomycin molecule and RNA aptamer make the binding pocket stable.

**Key-words:** Steered molecular dynamics, streptomycin, bluensomycin, free energy.

# Lista de ilustrações

Figura 1 – Esquema estrutural do RNA . . . . .	25
Figura 2 – Conformações permitidas para a D-ribose . . . . .	26
Figura 3 – Padrões de estrutura secundária para o RNA . . . . .	27
Figura 4 – Aptâmero de RNA para a molécula de estreptomicina . . . . .	28
Figura 5 – Esquema de um microscópio de força atômica . . . . .	30
Figura 6 – Estrutura dos ligantes estreptomicina e bluensomicina. . . . .	52
Figura 7 – Representação do sistema para dinâmica molecular dirigida . . . . .	56
Figura 8 – RMSD para os três sistemas estudados em solvente explícito . . . . .	60
Figura 9 – Número de PB para os três sistemas estudados em solvente explícito . . . . .	61
Figura 10 – Número de multipletos para os três sistemas estudados em solvente explícito . . . . .	62
Figura 11 – Estrutura representativa do sistema <i>complexo-estreptomicina</i> com tripletos G6•G13•C113 e U16•C18•G110 em evidência. . . . .	63
Figura 12 – Estrutura do sistema <i>aptâmero</i> com tripletos G6•G13•C113, U16•C18•G110, U19•A105•C106 e A8•A14•U12 em evidência. . . . .	65
Figura 13 – Número de HBONDS para os três sistemas estudados em solvente explícito . . . . .	66
Figura 14 – Número de HBONDS RNA fita1-ligante em solvente explícito . . . . .	66
Figura 15 – Número de HBONDS RNA fita2-ligante em solvente explícito . . . . .	67
Figura 16 – Estrutura representativa do sistema <i>complexo-estreptomicina</i> e resíduos que fazem HBOND com ligante. . . . .	68
Figura 17 – RMSD para o sistema <i>aptâmero</i> com solvente explícito . . . . .	69
Figura 18 – RMSD para o sistema <i>complexo-estreptomicina</i> com solvente explícito . . . . .	70
Figura 19 – Parâmetros estruturais para o sistema <i>aptâmero</i> com solvente explícito . . . . .	71
Figura 20 – Parâmetros estruturais para o sistema <i>complexo-estreptomicina</i> com solvente explícito . . . . .	72
Figura 21 – RMSD para o sistema <i>aptâmero</i> com solvente implícito e $\gamma = 2,0 \text{ ps}^{-1}$ . . . . .	72
Figura 22 – RMSD para o sistema <i>aptâmero</i> com solvente implícito e $\gamma = 3,0 \text{ ps}^{-1}$ . . . . .	73
Figura 23 – RMSD para o sistema <i>complexo-estreptomicina</i> com solvente implícito e $\gamma = 2,0 \text{ ps}^{-1}$ . . . . .	73
Figura 24 – RMSD para o sistema <i>complexo-estreptomicina</i> com solvente implícito e $\gamma = 3,0 \text{ ps}^{-1}$ . . . . .	74
Figura 25 – Parâmetros estruturais do sistema <i>aptâmero</i> com solvente implícito e $\gamma = 2,0 \text{ ps}^{-1}$ . . . . .	74
Figura 26 – Parâmetros estruturais do sistema <i>complexo-estreptomicina</i> com solvente implícito e $\gamma = 2,0 \text{ ps}^{-1}$ . . . . .	75

Figura 27 – CFE para o sistema <i>aptâmero</i> para calibração de $v$ e $k$ . . . . .	77
Figura 28 – CFE para o sistema <i>complexo-estreptomicina</i> para calibração de $v$ e $k$ .	77
Figura 29 – Diferença entre $d_{COM}$ e $dpp$ para o sistema <i>aptâmero</i> para calibração de $v$ e $k$ . . . . .	78
Figura 30 – Diferença entre $d_{COM}$ e $dpp$ para o sistema <i>complexo-estreptomicina</i> para calibração de $v$ e $k$ . . . . .	79
Figura 31 – Energia livre em função da velocidade de afastamento . . . . .	81
Figura 32 – Diferença de energia livre experimental e calculada em função da velocidade de afastamento . . . . .	82
Figura 33 – FEC para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i> . . . . .	84
Figura 34 – Estruturas representativas ao longo da CFE para o sistema <i>aptâmero</i> .	85
Figura 35 – Estruturas representativas ao longo da CFE para o sistema <i>complexo-estreptomicina</i> . . . . .	85
Figura 36 – Número de pares de base em função da $dpp$ para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i> . . . . .	86
Figura 37 – Padrão de ligações de hidrogênio entre estreptomicina e aptâmero no sistema <i>complexo-estreptomicina</i> . . . . .	87
Figura 38 – Número de multipletos em função da $dpp$ para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i> . . . . .	88
Figura 39 – Número médio de hélices em função da $dpp$ para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i> . . . . .	89
Figura 40 – Número de ligações de hidrogênio em função da $dpp$ para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i> . . . . .	90
Figura 41 – BSA em função da $dpp$ para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i>	92
Figura 42 – Energia livre $\Delta F$ em função da distância de ponta a ponta para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i> . . . . .	94
Figura 43 – Extrapolação linear de energia livre $\Delta F$ , em função da distância de ponta a ponta, para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i> . .	96
Figura 44 – Diferença de energia livre $\Delta\Delta F$ , em função da distância de ponta a ponta, para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i> . . . . .	97
Figura 45 – $\Delta U$ em função da $dpp$ para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i>	101
Figura 46 – Estrutura representativa do sistema <i>complex-estreptomicina</i> com molécula ligante em ponte . . . . .	103
Figura 47 – $T\Delta S$ em função da $dpp$ para os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i>	103
Figura 48 – FEC para os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i> . . . . .	107
Figura 49 – Estruturas representativas ao longo da CFE para o sistema <i>complexo-bluensomicina</i> . . . . .	108
Figura 50 – Número de pares de base em função da $dpp$ para os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i> . . . . .	109

Figura 51 – Número de multipletos em função da dpp para os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i> . . . . .	109
Figura 52 – Número médio de hélices em função da dpp para os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i> . . . . .	110
Figura 53 – Estruturas representativas ao final da região II para os sistemas <i>aptâmero</i> , <i>complexo-estreptomicina</i> e <i>complexo-bluensomicina</i> . . . . .	110
Figura 54 – Estrutura dos ligantes estreptomicina e bluensomicina . . . . .	111
Figura 55 – Padrão de ligações de hidrogênio entre bluensomicina e aptâmero no sistema <i>complexo-bluensomicina</i> . . . . .	112
Figura 56 – Número de ligações de hidrogênio em função da dpp para os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i> . . . . .	113
Figura 57 – Energia livre $\Delta F$ em função da distância de ponta a ponta para os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i> . . . . .	116
Figura 58 – Diferença de energia livre $\Delta\Delta F$ , em função da distância de ponta a ponta, para os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i> . . . . .	119
Figura 59 – $\Delta U$ em função da dpp para os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i>	122
Figura 60 – $T\Delta S$ em função da dpp para os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i>	124

# Lista de abreviaturas e siglas

A	Adenina
AFM	Microscopia de força atômica - do inglês <i>atomic force microscopy</i>
AMBER	Assisted Model Building with Energy Refinement, família de campos de forças para dinâmica molecular
BSA	Área de superfície de contato - do inglês <i>buried surface area</i>
C	Citosina
CFE	Curva(s) de força e extensão
ddp	distância de ponta a ponta
DL	Dinâmica de Langevin
DM	Dinâmica Molecular
DMD	Dinâmica Molecular Dirigida
DNA	ácido desoxirribonucleico - do inglês <i>deoxyribonucleic acid</i>
dsRNA	RNA de duas cadeias
G	Guanina
GB/SA	Método <i>Generalized Born solvent-accessible surface area</i>
GQT	Grupo de Química Teórica da Universidade Federal do Rio Grande do Sul
GROMACS	GROningen MACHine for Chemical Simulations, pacote de dinâmica molecular para biomoléculas
MC	Dinâmica de Monte Carlo
mRNA	RNA mensageiro
MTTK	Barostato de Martyna-Tuckerman-Tobias-Klein
NPT	<i>Ensemble</i> isotérmico-isobárico
NVE	<i>Ensemble</i> microcanônico
NVT	<i>Ensemble</i> canônico

PDB	<i>Protein Data Bank</i>
RESP	<i>Restrained electrostatic potential</i>
RMN	Ressonância magnética nuclear
RMSD	Raiz quadrada do desvio médio
RNA	ácido ribonucleico - do inglês <i>ribonucleic acid</i>
rRNA	RNA ribossomal
SPME	<i>Smooth Particle-Mesh Ewald</i>
ssRNA	RNA de cadeia única
T	Timina
tRNA	RNA de transportador
U	Uracila
UFRGS	Universidade Federal do Rio Grande do Sul

# Lista de símbolos

$d_{COM}$	Posição do centro de massas
$\mathbf{F}$	Vetor, $3N$ -dimensional, com as forças atuando nas partículas do sistema
$H(\mathbf{p}, \mathbf{q})$	Hamiltoniano do sistema em função dos momentos, $\mathbf{p}$ , e posições, $\mathbf{q}$
$K(\mathbf{p})$	Energia cinética em função dos momentos, $\mathbf{p}$
$k_B$	Constante de Boltzmann
$\mathbf{k}$	Constante de força para o potencial harmônico
$k$	Constante de força para o potencial harmônico
$\mathbf{M}$	Matriz diagonal positiva, $3N$ -dimensional, com as massas atômicas
$\mathbf{p}$	Momentos conjugados
$p(X)$	Probabilidade de distribuição da propriedade $X$
$p(\mathbf{p}_1^i, \dots, \mathbf{p}_N^i, \mathbf{r}_1^i, \dots, \mathbf{r}_N^i)$	Probabilidade do sistema se encontrar em um estado $\mathbf{p}_1^i, \dots, \mathbf{p}_N^i, \mathbf{r}_1^i, \dots$
$\mathbf{q}$	Vetor de coordenadas generalizadas
$Q$	Calor trocado com reservatório térmico
$\mathbb{R}$	Conjunto dos números Reais
$\mathbf{r}$	Vetor com as coordenadas cartesianas de uma partícula, $\mathbf{r} = (x, y, z)$
$\mathbf{R}$	Vetor, $3N$ -dimensional, com as coordenadas cartesianas de todas as partículas do sistema
$\mathbf{R}(t)$	Vetor de forças aleatórias para DL
$t$	Tempo
$T$	Temperatura
$\mathbf{v}$	Velocidade de afastamento entre o átomo de referência e o átomo móvel
$v$	Velocidade de afastamento entre o átomo de referência e o átomo móvel
$V_{ij}(\mathbf{q})$	Função de interação potencial entre as partículas $i$ e $j$
$V(\mathbf{X})$	Potencial em função do espaço de coordenadas $\mathbf{X}$

$\mathbf{v}_i$	Vetor velocidade da $i$ -ésima partícula
$W$	Trabalho realizado sobre o sistema
$\Gamma(\mathbf{Q})$	Espaço de fases em função das coordenadas $\mathbf{Q}$
$\gamma$	Constante (coeficiente) de fricção
$\Delta E$	Energia total do sistema
$\Delta F$	Energia livre de Helmholtz
$\Delta\Delta F$	Diferença de energia livre de Helmholtz
$\Delta S$	Variação de entropia
$\Delta U$	Variação de energia interna
$\lambda$	Coordenada de reação
$\tau_x$	Constante de acoplamento da propriedade $x$

# Sumário

	<b>Introdução</b> . . . . .	<b>17</b>
<b>1</b>	<b>OBJETIVOS</b> . . . . .	<b>21</b>
<b>1.1</b>	<b>Objetivos específicos</b> . . . . .	<b>21</b>
<b>I</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>22</b>
<b>2</b>	<b>ÁCIDOS NUCLEICOS</b> . . . . .	<b>23</b>
<b>2.1</b>	<b>RNA</b> . . . . .	<b>24</b>
<b>2.2</b>	<b>Aptâmero de RNA seletivo à estreptomicina</b> . . . . .	<b>28</b>
<b>3</b>	<b>MICROSCOPIA DE FORÇA ATÔMICA</b> . . . . .	<b>30</b>
<b>4</b>	<b>SIMULAÇÃO COMPUTACIONAL DE SISTEMAS MOLECULARES</b>	<b>32</b>
<b>4.1</b>	<b>Dinâmica Molecular de Equilíbrio</b> . . . . .	<b>32</b>
<b>4.2</b>	<b>Dinâmica de Langevin</b> . . . . .	<b>36</b>
<b>4.3</b>	<b>Dinâmica molecular fora do equilíbrio</b> . . . . .	<b>38</b>
<b>4.3.1</b>	<b>Dinâmica molecular dirigida</b> . . . . .	<b>38</b>
<b>5</b>	<b>ENERGIA LIVRE</b> . . . . .	<b>40</b>
<b>5.1</b>	<b>Igualdade de Jarzynski</b> . . . . .	<b>42</b>
<b>5.1.0.1</b>	<b>Derivação da igualdade de Jarzynski</b> . . . . .	<b>43</b>
<b>5.1.1</b>	<b>Decomposição da energia livre em energia interna e entropia</b> . . . . .	<b>47</b>
<b>5.1.1.1</b>	<b>Energia interna</b> . . . . .	<b>48</b>
<b>5.1.1.2</b>	<b>Entropia do sistema</b> . . . . .	<b>49</b>
<b>II</b>	<b>METODOLOGIA</b>	<b>50</b>
<b>6</b>	<b>METODOLOGIA</b> . . . . .	<b>51</b>
<b>6.1</b>	<b>Dinâmica molecular</b> . . . . .	<b>51</b>
<b>6.1.1</b>	<b>Campo de forças para os ligantes</b> . . . . .	<b>52</b>
<b>6.1.2</b>	<b>Simulação do solvente explícito</b> . . . . .	<b>53</b>
<b>6.1.3</b>	<b>Simulação com solvente implícito</b> . . . . .	<b>54</b>
<b>6.1.4</b>	<b>Dinâmica molecular dirigida</b> . . . . .	<b>55</b>
<b>6.2</b>	<b>Cálculo da Energia Livre</b> . . . . .	<b>55</b>

<b>III</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>58</b>
<b>7</b>	<b>SIMULAÇÕES COM SOLVENTE EXPLÍCITO</b>	<b>59</b>
<b>8</b>	<b>COEFICIENTE DE FRICÇÃO</b>	<b>69</b>
<b>9</b>	<b>VELOCIDADE DE AFASTAMENTO E CONSTANTE DE FORÇA</b>	<b>76</b>
<b>9.1</b>	<b>Constante de força</b>	<b>76</b>
<b>9.2</b>	<b>Velocidade de afastamento</b>	<b>78</b>
<b>10</b>	<b>DESENOVELAMENTO DOS SISTEMAS <i>APTÂMERO</i> E <i>COMPLEXO-ESTREPTOMICINA</i></b>	<b>83</b>
<b>10.1</b>	<b>Análise estrutural</b>	<b>83</b>
10.1.1	Modelo estrutural para desenovelamento induzido por força	87
<b>10.2</b>	<b>Análise energética</b>	<b>92</b>
10.2.1	Análise energética do processo de desenovelamento	93
10.2.2	Diferença de energia livre entre os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i>	95
10.2.2.1	Diferença de energia livre entre os sistemas <i>aptâmero</i> e <i>complexo-estreptomicina</i> no ponto de ruptura	98
10.2.3	Decomposição da energia livre em energia interna e entropia	100
10.2.3.1	Energia interna	100
10.2.3.2	Entropia do sistema	102
<b>11</b>	<b>DESENOVELAMENTO DO SISTEMA <i>COMPLEXO-BLUENSOMICINA</i></b>	<b>106</b>
<b>11.1</b>	<b>Análise estrutural</b>	<b>106</b>
<b>11.2</b>	<b>Análise energética</b>	<b>115</b>
11.2.1	Análise energética do processo de desenovelamento	115
11.2.2	Diferença de energia livre entre os sistemas <i>aptâmero</i> e <i>complexo-bluensomicina</i>	118
11.2.3	Comparação das energias livres de ligação das moléculas <i>estreptomicina</i> e <i>bluensomicina</i> ao sítio de ligação do <i>aptâmero</i>	120
11.2.4	Decomposição da energia livre em energia interna e entropia	121
11.2.4.1	Energia interna	121
11.2.4.2	Entropia do sistema	123
	<b>Conclusão</b>	<b>126</b>
	<b>REFERÊNCIAS</b>	<b>128</b>

<b>APÊNDICES</b>	<b>139</b>
<b>APÊNDICE A – CÁLCULO DO TRABALHO MICROSCÓPICO . .</b>	<b>140</b>
<b>APÊNDICE B – ARTIGO PUBLICADO NO JOURNAL OF CHE- MICAL PHYSICS . . . . .</b>	<b>142</b>

# Introdução

Os ácidos nucleicos foram descobertos em 1869 por Johann Friedrich Miescher.(1) Um estudo posterior destas moléculas mostrou que estas eram formadas por, entre outros componentes, dois tipos distintos de pentoses, a ribose e a desoxirribose de onde vem os nomes ácido ribonucleico (RNA) e ácido desoxirribonucleico (DNA), respectivamente. Inicialmente haviam dúvidas sobre a função biológica destas moléculas. Entretanto, posteriormente foi demonstrado que a molécula de DNA era a responsável por portar a informação genética.(2, 3, 4, 5) Entretanto, por muito tempo se pensou que as moléculas de RNA serviam apenas para transcrever o código do DNA durante a síntese proteica.

Atualmente, se conhecem diversas funções biológicas para a molécula de RNA. Sabe-se agora que esta molécula, antes considerada apenas como uma mera coadjuvante, possui um papel importante para toda a biologia molecular.(5, 6) Além de seu papel na síntese de proteínas (na forma de mRNA, tRNA e rRNA), o RNA atua como catalisador em diversas reações biológicas (riboenzimas), controlando a expressão gênica e, também, funcionam como sensores de sinais celulares.(5) Moléculas de oligonucleotídeos e peptídeos que se ligam a moléculas alvo específicas são conhecidas como aptâmeros. Desta forma, moléculas de RNA, denominadas *riboswitches*, atuam como receptores para metabólitos específico, ou seja, como aptâmeros. Com isto, uma das metas no campo da biologia molecular atual é entender a estrutura, dinâmica e interações que as moléculas de RNA possuem, considerando suas interações com outras moléculas nos diversos processos celulares em que ela atua.(7).

Certas moléculas de RNA em ambiente biológico se enovelam em estruturas tridimensionais nativas estáveis para assim realizar suas funções biológicas.(8, 9, 10) Entretanto, as diversas atividades celulares desempenhadas pelas moléculas de RNA requerem que esta sofra mudanças da estrutura.(11) A compreensão destas mudanças estruturais em um nível molecular se torna fundamental para elucidar a relação entre a estrutura e função em moléculas de RNA biologicamente importantes.(8) O estudo das propriedades físicas e químicas de moléculas de RNA era usualmente realizado em soluções contendo um grande número de moléculas.(12) Entretanto, o estudo em solução dos mecanismos de enovelamento e desnaturação apresentam respostas com grandes variações devido a presença de múltiplas espécies.(9)

A observação direta do processo de desnaturação de uma molécula de RNA em um ambiente de importância biológica não é fácil e, devido à estabilidade da estrutura secundária do RNA, só é possível mediante a aplicação de força à estrutura.(11) Entretanto, na década de 1980 houve o desenvolvimento de diversas técnicas experimentais que

possibilitaram a precisa aplicação e mensuração de força aplicada aos sistemas moleculares em resoluções temporais altas.(13) O desenvolvimento de técnicas como a microscopia de força atômica (AFM - do inglês *atomic force microscopy*) e pinças ópticas, possibilitaram a manipulação de uma única molécula, tornando assim possível o estudo da energia envolvida nos mecanismos de desnaturação de moléculas de ácidos nucleicos sob aplicação de força.(9, 10, 14) Atualmente, medidas envolvendo uma molécula por vez são rotineiramente usados para estudar sistemas e processos biológicos(15).

Além da possibilidade de se estudar o processo mecanicamente induzido de desnaturação e enovelamento de uma única molécula, estas técnicas tem como benefício o uso de uma única e bem definida coordenada de reação, ou seja a distância entre as extremidades da molécula de RNA.(9) Outras vantagens destes métodos de manipulação de uma única molécula são a possibilidade de usar a força como uma variável controlada e a habilidade de aplicar esta força sem perturbar outras espécies moleculares na amostra.(10)

Nos anos 2000, diversos pesquisadores aplicaram as técnicas de microscopia de força atômica e pinças ópticas desenvolvidas na década de 1980 ao estudo das propriedades de diversas moléculas de RNA.(9, 16, 17, 18, 19, 20) Entre outros marcos do estudo de moléculas de RNA alcançados, estes trabalhos demonstraram os efeitos detalhados da força aplicada à molécula nas curvas de força e extensão (CFE), bem como a histerese entre o processo de desnaturação e enovelamento.(9) Outro marco na aplicação de técnicas de medidas sobre uma única molécula foi a aplicabilidade e confirmação da igualdade de Jarzynski(16, 21) e teorema de flutuação de Crooks(19, 22) ao processo de desenovelamento de uma molécula de RNA e com isto se tornou possível a obtenção dos valores de energia livre para este processo.

O estudo do mecanismo de desnaturação e enovelamento de moléculas de RNA por métodos de manipulação de uma única molécula revela as múltiplas possibilidades de caminhos possíveis pelos quais estes processos ocorrem, bem como detalha as magnitudes das forças necessárias para que estes mesmos eventos ocorram.(10) Entretanto, estes métodos nada revelam sobre o que acontece à estrutura do RNA, em um nível de detalhamento atômico, durante o desenrolar destes processos. Em contrapartida, técnicas de simulação computacionais se transformaram em uma ferramenta padrão para facilitar a interpretação em nível atômico ou molecular de resultados gerados em bancada.(23) Técnicas de simulação como dinâmica molecular (DM) e Monte Carlo (MC) são atualmente ferramentas indispensáveis para o entendimento do mundo molecular de sistemas em equilíbrio termoestático<sup>1</sup>. A aplicação destas ferramentas foi expandida para sistemas fora do equilíbrio com o desenvolvimento de técnicas como a dinâmica molecular dirigida

<sup>1</sup> Termoestática (termodinâmica de ordem zero) refere-se ao estudo dos sistemas focados no equilíbrio (uma abordagem estática), onde as forças motrizes do sistema são nulas e não há a variável tempo. Termodinâmica (de primeira ordem) refere-se ao estudo dos sistemas onde as forças motrizes do sistema não são nulas, portanto, o sistema está evoluindo temporalmente para novos estados.

(DMD)(24, 25, 26, 27) e a metadinâmica.(28)

Do ponto de vista computacional, os experimentos de microscopia de força atômica foram simulados por Naserian-Nik, Tahani e Karttunen(26, 27). Estes pesquisadores aplicaram o método de dinâmica molecular dirigida para investigar a nível atômico o processo de desnaturação por aplicação de força de oligonucleotídeos de DNA. Para oligonucleotídeos de RNA, o processo de desnaturação induzido por força externa possui ainda poucos exemplos na literatura. Harlepp et al.(18) e Gerland, Bundschuh e Hwa(29) simularam a desnaturação de moléculas de RNA de fita simples utilizando métodos estocásticos baseados em estrutura secundária. Deng e Cieplak(30) aplicou a DMD em nível atômico para o estudo de desenovelamento de pequenos *hairpins* compostos de dez pares de base e Gupta e Bansal(31) aplicaram a mesma técnicas para investigar o efeito da troca de uma base no desenovelamento de um pseudonó de RNA de cadeia única. Bell et al.(32) desenvolveram um modelo *coarse-grained* de RNA e aplicaram a DMD para obterem a descrição do processo desnaturação de *hairpins* de RNA compostos de dez a 18 bases. Entretanto, até o momento estas metodologias não foram testadas em sistemas de RNA com duas fitas ou na presença de ligantes.

A molécula de estreptomicina, um antibiótico aminoglicosídeo, é conhecida por ter como alvo e formar um complexo com um domínio estrutural do RNA bacteriano.(5, 33) Diversas técnicas foram usadas para entender como os aminoglicosídeos se ligam à conformações específicas de RNA, dentre elas a espectroscopia de ressonância magnética nuclear, as técnicas cristalográficas e a microscopia de força atômica de molécula única.(34, 35) Recentemente, Nick et al.(35) utilizaram simulações computacionais de equilíbrio e microscopia de força atômica de molécula única para estudarem as interações da estreptomicina com um fragmento de RNA (aptâmero). Utilizando os resultados destas duas técnicas, Nick et al.(35) chegaram a conclusão que a molécula de estreptomicina estabiliza a estrutura do RNA devido a ligações de hidrogênio adicionais entre as fitas do aptâmero e a estreptomicina em um padrão muito específico. Entretanto, os autores não conseguiram, devido as simulações de equilíbrio, obter informações a nível atômico sobre o processo de desnaturação ou desenovelamento da estrutura do RNA.

Desta forma, são objetivos deste trabalho adaptar a metodologia utilizada em simulações fora do equilíbrio para oligômeros de DNA, empregada por Naserian-Nik, Tahani e Karttunen(27), para um aptâmero de RNA e o complexo do aptâmero com a estreptomicina, de forma a simular o processo de desnaturação destes sistemas por aplicação de força externa e comparar os resultados obtidos com os valores experimentais e de simulação de equilíbrio obtidos por Nick et al.(35). De forma a avaliar a seletividade da metodologia fora do equilíbrio, será também simulado o complexo do aptâmero de RNA com a molécula de bluensomicina, uma molécula análoga da estreptomicina, porém de baixa afinidade por este aptâmero.(34, 36, 37)

# 1 Objetivos

Este trabalho tem por objetivo estabelecer um protocolo de simulação por dinâmica molecular fora do equilíbrio que permita investigar, em nível atômico, o processo de desnaturação por aplicação de força externa de moléculas de ácido ribonucleico, de forma a se obter uma melhor compreensão dos mecanismos energéticos e estruturais envolvidos durante este processo.

## 1.1 Objetivos específicos

- Estabelecer um protocolo de simulação por dinâmica molecular para um aptâmero de ácido ribonucleico, descrito por Nick et al.(35), em solvente implícito;
- Adaptar o método empregado por Naserian-Nik, Tahani e Karttunen(26) ao processo de desnaturação induzida por força externa de moléculas de RNA de forma a simular a desnaturação destas moléculas durante o experimento de microscopia de força atômica de uma única molécula;
- Estudar as mudanças conformacionais e estruturais do processo de desnaturação, identificando padrões nas curvas de força-extensão.
- Utilizar a aproximação mola rígida(38) (do inglês *stiff-spring*) para o cálculo do trabalho realizado durante a desnaturação induzida por força externa de moléculas de RNA;
- Aplicar a igualdade de Jarzynski por meio do método de aproximação linear, proposto por Ytreberg e Zuckerman(39), de forma a obter a energia livre do processo de desnaturação e compará-las com as energias obtidas experimentalmente por Nick et al.(35);

Parte I

Fundamentação Teórica

## 2 Ácidos Nucleicos

No inverno de 1868, Dr. Johann Friedrich Miescher conduziu uma série de experimentos para caracterizar a composição química de leucócitos que o levaram a descoberta da molécula que posteriormente veio a ser denominada ácido desoxiribonucleico ou DNA. Uma vez que esta molécula foi encontrada no núcleo das células, Dr. Johann Friedrich Miescher a denominou “nucleína”.(1) No início dos anos 1900, pouco se conhecia sobre a diferença biológica e química entre os tipos de ácidos nucleicos, sendo eles classificados com o nome da fonte da qual eles eram obtidos como “ácido nucleico da levedura” e “ácido nucleico do timo”.(40) Estudos bioquímicos subsequentes demonstraram que estes ácidos nucleicos eram compostos de dois tipos diferentes de carboidratos, ribose e desoxiribose, de onde vem sua denominação atual ácido ribonucleico ou RNA e ácido desoxiribonucleico ou DNA.

Entre as décadas de 1940 e 1950, os trabalhos de Avery, MacLeod e MacCarthy(2) e Ephussi-Taylor(3) demonstraram que aqueles ácidos nucleicos conhecidos como DNA eram as biomoléculas responsáveis por portar a informação genética e, portanto, as moléculas responsáveis pela hereditariedade.(2, 3, 4, 5) O ácido desoxirribonucleico, se sabe agora, possui somente duas funções, que é produzir sua própria replicação durante a divisão celular e produzir a transcrição de moléculas de ácido ribonucleico complementar.(5)

Diferentemente da molécula de DNA, o ácido ribonucleico, RNA, possui uma gama diversificada de funções biológicas.(5, 6, 41) Desde 1939 já haviam suspeitas que a molécula de RNA possuía papel na síntese de proteínas.(42) Entretanto, se acreditava que esta molécula possuía apenas um papel passivo na síntese proteica, apenas carregando a informação genética contida no DNA.(6) Hoje se sabe que, o papel na síntese de proteínas é mais amplo, não só como mRNA, que carrega a mensagem transcrita do DNA até o ribossomo, esta molécula está presente em diversas etapas chave da síntese proteica. O próprio ribossomo é um agregado molecular de proteínas e sequências de ácido ribonucleico, denominadas rRNA. Uma terceira forma de RNA, denominada RNA de transportador (tRNA) é importante na síntese proteica, carregando os resíduos de proteínas especificados pelo mRNA até o sítio de crescimento de cadeia no ribossomo.(5, 40, 41)

Atualmente, além de seu papel na síntese de proteínas, atuando como mRNA, tRNA e rRNA, a molécula do ácido ribonucleico atua como catalisador em diversas reações biológicas (riboenzimas), controlando a expressão gênica ou funcionando como sensores dos sinais celulares.(5, 6, 41) Dentro da gama de funções realizadas por esta molécula, o RNA também atua como material genético responsável pela hereditariedade em alguns vírus.(5, 40) E, atualmente já se usam moléculas de RNA como sensores químicos para

análise de moléculas.(43, 44)

## 2.1 RNA

O ácido ribonucleico, RNA, consiste em uma cadeia biopolimérica linear, formada por unidades monoméricas denominadas **ribonucleotídeos** ou simplesmente **nucleotídeos**, Figura 1 (página 25). O RNA pode se apresentar em cadeia simples, ssRNA, ou em arranjos de duas cadeias, dsRNA. Estes nucleotídeos são ésteres de fosfato da D-ribose que possui uma base nitrogenada covalente ligada, pela denominada ligação glicosídea, à posição  $\beta$  de seu carbono C1'. A D-ribose, em ácidos nucleicos, ocorre mais frequentemente com seu anel dobrado na conformação *C2'-endo*, sendo que as conformações *C3'-endo* e *C3'-exo* também ocorrem com certa frequência. A cadeia é formada através de sucessivas ligações 3',5'-fosfodiéster entre um grupo 3'-hidroxil do *i*-ésimo nucleotídeo ao grupo 5'-fosfato do (*i* + 1)-ésimo nucleotídeo.(5, 40, 41, 45, 46) A Figura 2 (página 26) mostra as possíveis conformações para a ribose discutidas neste parágrafo. É interessante notar que a presença do grupo 2'-hidroxil em cada D-ribose torna a ligação 3',5'-fosfodiéster suscetível a hidrólise catalisada por base.(5, 45)

Os nucleotídeos são diferenciados de acordo com suas bases nitrogenadas constituintes. Naturalmente estão presentes cinco tipos de bases nitrogenadas nos ácidos nucleicos, Figura 1 (página 25). São elas: Adenina (A), guanina (G), citosina (C), timina (T) e uracila (U). Entretanto, destas somente A, G, C e U estão presentes na molécula de RNA. Estas bases são classificadas de acordo com seu arcabouço estrutural em bases purínicas e pirimidínicas, sendo A e G bases purínicas e C, T e U bases pirimidínicas.(5, 45, 46) Para bases purínicas, duas orientações são estericamente permitidas, **syn** (do grego com) e **anti** (do grego contra), Figura 2D (página 26). Entretanto, para as bases pirimidínicas só a conformação **anti** é facilmente formada e, portanto, é a conformação permitida para estas bases.(5, 47)

De forma similar aos outros dois biopolímeros mais conhecidos, proteínas e ácido desoxiribonucleico (DNA), os polímeros do ácido ribonucleico também se organizam de forma hierárquica. (5, 41, 45, 46) A sequência como as bases estão dispostas na cadeia polimérica determina a **estrutura primária** da molécula de RNA. Esta estrutura é mantida devido as ligações 3',5'-fosfodiéster entre os nucleotídeos. É importante salientar que a sequência de uma molécula de RNA pode se apresentar em cadeia única, ssRNA, em arranjos de duas cadeias, dsRNA, e até arranjados em três cadeias.

A **estrutura secundária** de uma molécula de RNA é determinada pelas interações entre os nucleotídeos dentro de uma única cadeia, para o caso de ssRNA, ou entre duas cadeias, para dsRNA. A estrutura secundária é formada geralmente pela interação entre duas ou três bases, mas em uma estrutura complexa pode ocorrer a interação entre

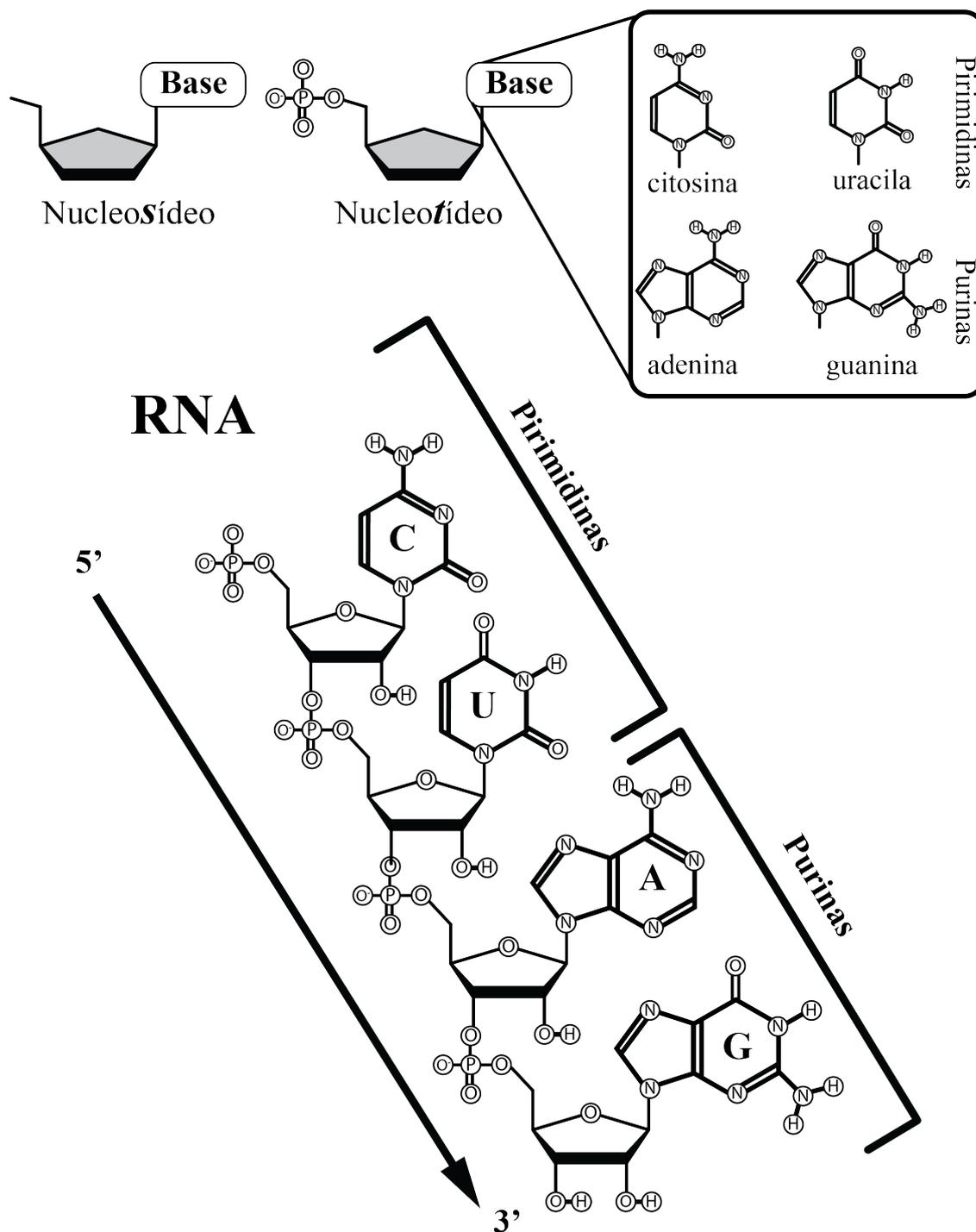


Figura 1 – Esquema estrutural do biopolímero do ácido ribonucleico (RNA), composição química e classificação das bases.

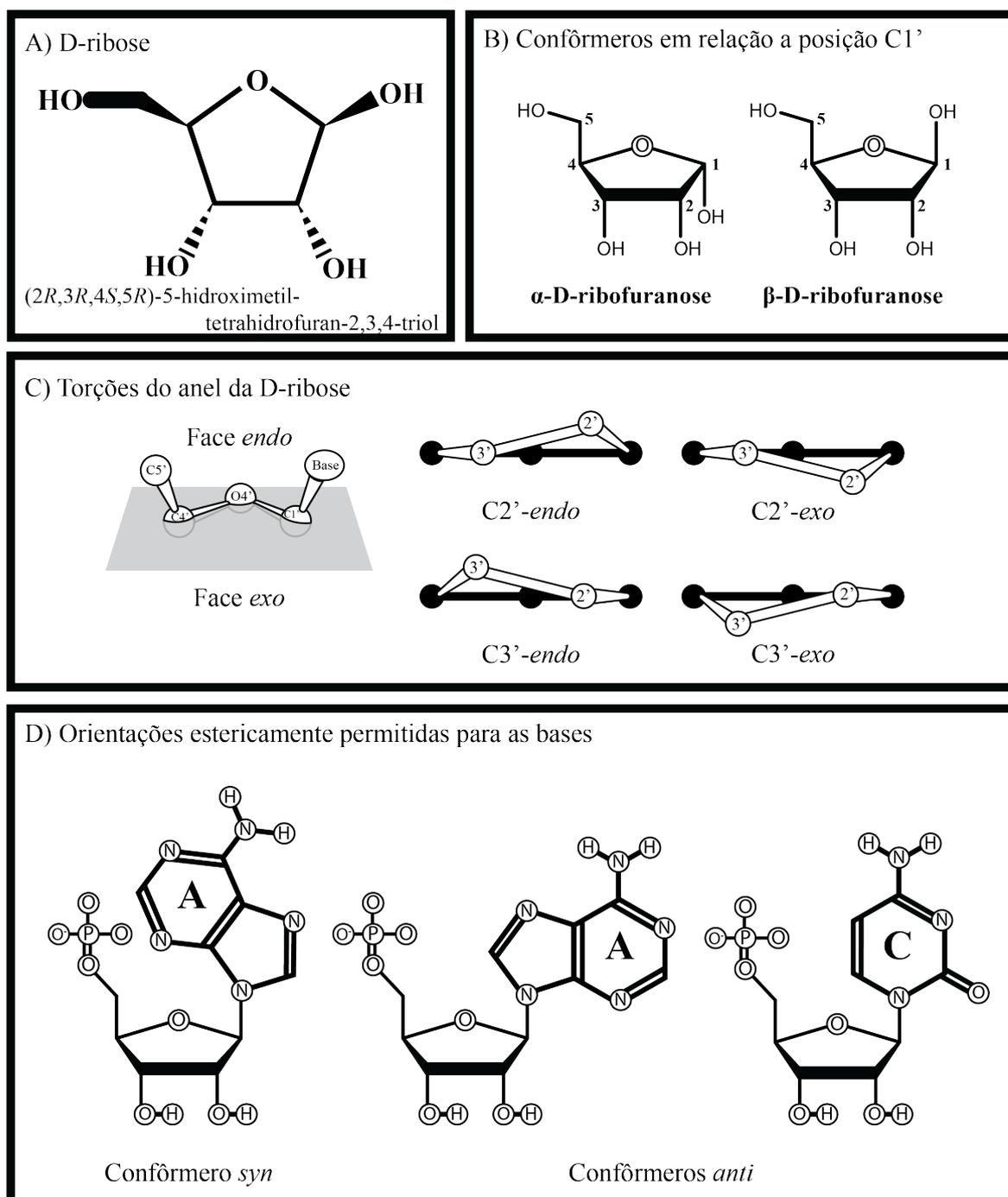


Figura 2 – D-Ribose

mais bases.(5, 48) Estas interações entre as bases de uma mesma cadeia, ou duas cadeias diferentes, pode gerar padrões específicos que ditarão a forma tridimensional da molécula de ácido ribonucleico.

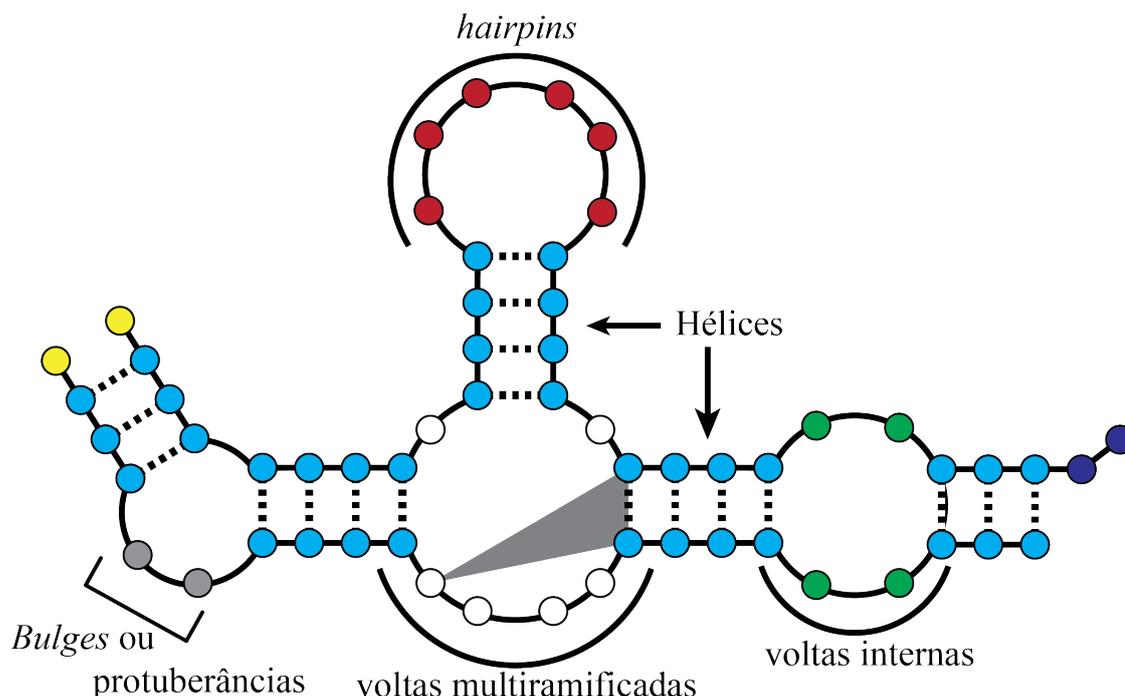


Figura 3 – Padrões de estrutura secundária para o RNA. Cada círculo representa um nucleotídeo. As linhas sólidas representam as ligações 3',5'-fosfodiéster entre os nucleotídeos. Linhas pontilhadas representam interações entre pares de base e a triângulo cinza representa uma interação entre três bases (triplete).

A Figura 3 mostra alguns dos padrões de estrutura secundária que ocorrem frequentemente em RNA. Dentre estas estruturas, se pode citar:

- Hélices, representadas por círculos azuis claros na Figura 3, são estruturas formadas por pares de base empilhados;
- Voltas do tipo *hairpin*, representadas por círculos vermelhos na Figura 3, são voltas da cadeia de RNA sobre si constituídas por no mínimo três nucleotídeos desemparelhados;
- Protuberâncias ou *bulges*, círculos cinzas, são padrões formados por nucleotídeos desemparelhados em uma cadeia do RNA;
- Voltas internas, círculos verdes, são padrões formados por nucleotídeos desemparelhados em ambas as cadeias do RNA;
- Voltas ramificadas, círculos brancos, são voltas que estão entre no mínimo três estruturas do tipo hélice.

A Figura 3 mostra ainda as duas situações que podem ocorrer no final da estrutura secundária de uma molécula de RNA. Os círculos amarelos da Figura 3 representam pares

de base desemparelhados ao final de uma hélice. Os círculos azuis escuros representam bases desemparelhadas ao final da cadeia.

A **estrutura terciária** do RNA compreende todas as interações entre os nucleotídeos e é representada pelas coordenadas de todos os átomos, na sequência de RNA, em um espaço de referência.(5, 45, 46) Esta estrutura é determinada pelo arranjo espacial dos padrões encontrados na estrutura secundária do RNA. Entretanto, a estrutura terciária é influenciada tanto pela estrutura secundária como pela estrutura primária do RNA.(49)

## 2.2 Aptâmero de RNA seletivo à estreptomicina

Esta trabalho tem por objeto de estudo o aptâmero de RNA seletivo à molécula de estreptomicina isolado por Wallace e Schroeder(37). Este aptâmero foi desenvolvido pela combinação das técnicas de seleção *in vitro* e cromatografia por afinidade, de forma que a sequência de RNA isolada pudesse discriminar as moléculas de estreptomicina e bluensomicina, Figura 4a.(37)

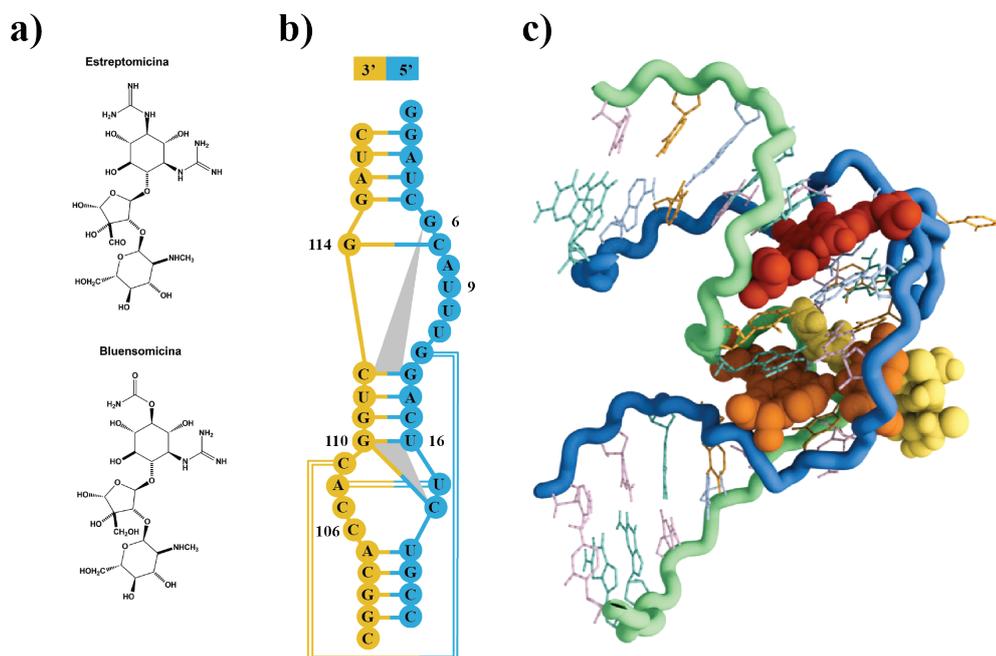


Figura 4 – Aptâmero de RNA para a molécula de estreptomicina. A) Estrutura secundária do complexo. Pares de base de Watson-Crick canônicos estão representados por linhas sólidas e pares não canônicos estão representados por linhas duplas. Tripletos estão representados por triângulos cinza. B) Estrutura da molécula de estreptomicina. C) Complexo do aptâmero de RNA com a estreptomicina e tripletos em evidência. Estreptomicina (amarelo) representada em esferas de van der Waals. Tripletos G6•G13•C113 (laranja escuro) e U16•C18•G110 (laranja) representados em esferas de van der Waals.

A estrutura do complexo entre a sequência isolada por Wallace e Schroeder(37) e a molécula de estreptomicina foi elucidada por cristalografia de raios X, em uma resolução de 2,9 Å, por Tereshko, Skripkin e Patel(34) em 2003. Tereshko, Skripkin e Patel(34) depositaram a estrutura cristalográfica deste complexo no servidor *Protein Data Bank* sob o código PDBid 1NTB.

A estrutura cristalográfica do aptâmero de RNA seletivo à estreptomicina é composta por duas cadeias, sendo uma cadeia composta por 22 nucleotídeos e a segunda cadeia é composta por 18 nucleotídeos, Figura 4b. Sua estrutura secundária, Figura 4b, conta com treze pares de base de Watson-Crick canônicos e mais dois pares de base não canônicos. Também esta presente na estrutura secundária do complexo dois tripletos: G6•G13•C113 e U16•C18•G110, Figura 4c. Há também um átomo de magnésio na estrutura ligado aos resíduos U10, U11 e C109. A estrutura terciária, Figura 4c, apresenta uma forma aproximadamente em L, formada por duas hastes.

### 3 Microscopia de Força Atômica

O microscópio de força atômica foi inventado em 1986 por Gerd Binnig, Calvin Quate e Christoph Gerber.(50) Uma das primeiras aplicações biológicas deste equipamento foi a visualização de moléculas de DNA.(51) Um microscópio de força atômica realiza o mapeamento topológico de uma superfície passando sobre esta uma agulha presa a um *cantilever*, Figura 5a. Quando a ponta da agulha encontra alguma imperfeição na superfície, o *cantilever* sofre uma deflexão. Esta deflexão é quantificada por um laser acoplado ao *cantilever*.

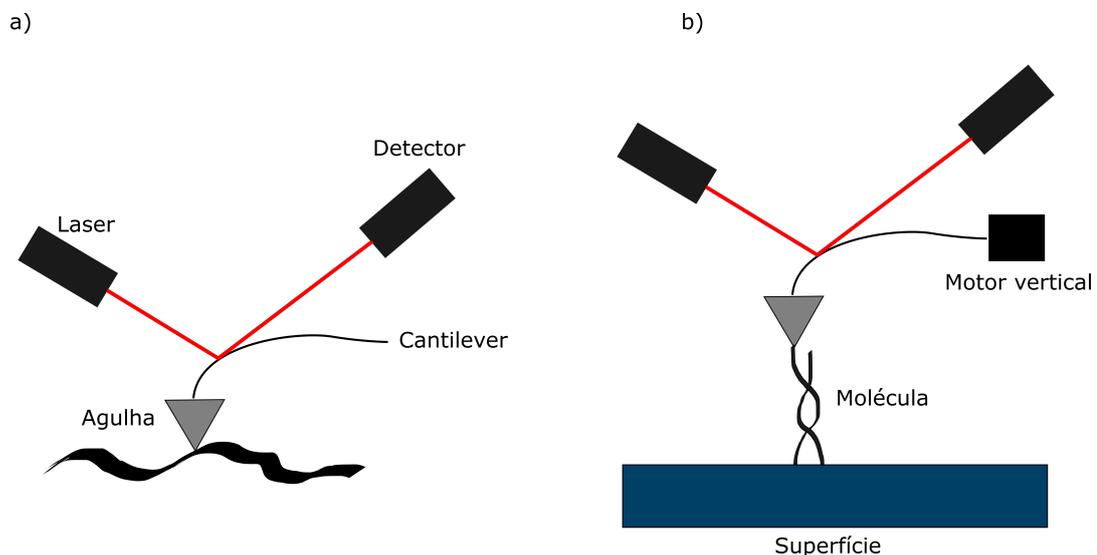


Figura 5 – Esquema de um microscópio de força atômica (a) comum e (b) com configuração para mensuração de força.

Outra forma possível de utilização do microscópio de força atômica é na mensuração de propriedades elásticas ou de adesão de uma superfície, gerando curvas de força.(52) Estas curvas são geradas realizando o deslocamento vertical da agulha enquanto se mensura a deflexão do *cantilever*, de forma que a força vertical à superfície seja mensurada em uma escala de nanonewtons. Esta configuração do microscópio de força atômica é interessante do ponto de vista teórico, pois permite que se obtenham valores a respeito das energias necessárias para a manipulação, desnaturação e renaturação de polímeros e biopolímeros.(53)

A obtenção de curvas de força-extensão por microscopia de força atômica de processos como a desnaturação e renaturação de moléculas de DNA e RNA vem contribuindo para o entendimento de como ocorrem estes processos em ambiente biológico. Nestes estudos, uma das extremidades da molécula é presa à uma superfície enquanto que a agulha do microscópio fixa a outra extremidade, Figura 5b, da molécula. Conforme a agulha é

deslocada verticalmente, a molécula em questão é desenovelada e a força necessária para este processo é registrada. O processo inverso também é possível, ou seja, após realizar o estiramento da molécula, a agulha pode se deslocar no sentido inverso e, com isso, se obter a força que é requerida para o enovelamento da molécula.

## 4 Simulação computacional de sistemas moleculares

A Dinâmica Molecular (DM) é o termo usado para descrever a solução numérica das equações clássicas do movimento, e.g. as equações de Newton, para um conjunto de partículas.(54) Devido às dificuldades técnicas, a primeira simulação de dinâmica molecular só foi obtida por volta de 1957 por Alder e Wainwright(55), que simulou um sistema de esferas rígidas.(55, 56) A primeira simulação de DM que teve sucesso em utilizar um potencial de Lennard-Jones só foi realizada em 1964 por Rahman(57), que utilizou a técnica da DM para simular argônio líquido.(57)

Em 1971, Rahman e Stillinger(58) simularam um modelo de água rígida por dinâmica molecular. Já a primeira simulação de dinâmica molecular de um sistema realístico, ou seja se considerando os graus de liberdade internos da molécula (e.g. deformações angulares ou variação nos comprimentos de ligação), foi a simulação de água líquida também realizada por Stillinger e Rahman(59) em 1974. Se tratando de sistemas biológicos, a primeira simulação de DM de uma proteína foi reportada em 1977, quando McCammon, Gelin e Karplus(60) publicaram os resultados de uma simulação de DM do inibidor da tripsina pancreática bovina. Seis anos mais tarde, em 1983, Levitt(61) publicou os dados referentes à primeira simulação de dinâmica molecular de um ácido nucleico. Levitt(61) realizou simulações do dodecamero complementar (5'→3')CGCGAATTCGCG e de um fragmento com 24 pares de base, ambos no vácuo. Entretanto, estas primeiras simulações de MD de ácidos nucleicos produziram estruturas altamente distorcidas, e portanto, não físicas, para o DNA.(62) A aplicação do método de dinâmica molecular a estruturas de ácidos nucleicos apresenta um dos maiores desafios ainda hoje devido ao caráter polieletrólítico destes, bem como devido as suas interações com o solvente. Somente por volta de 1990 é que modelos de água e tratamentos para os potenciais eletrostáticos adequados, bem como o desenvolvimento de campos de forças específicos para ácidos nucleicos, possibilitaram a obtenção de trajetórias de DM estáveis.(62)

A seção 4.1 descreve o método da dinâmica molecular.

### 4.1 Dinâmica Molecular de Equilíbrio

Nas simulações por dinâmica molecular (DM) clássica as equações da segunda lei do movimento de Newton são numericamente integradas para reproduzir, passo a passo, a evolução temporal dos átomos do sistema em estudo, ou seja, produzir sucessivas configurações deste sistema. Para um sistema atômico simples, estas podem ser escritas

como:

$$\mathbf{M} \frac{d^2 \mathbf{r}(t)}{dt^2} = \mathbf{M} \frac{d \mathbf{v}}{dt} = \mathbf{F} \quad (4.1)$$

e

$$\mathbf{F} = -\nabla_{\mathbf{r}} V(\mathbf{r}) \quad (4.2)$$

onde  $\mathbf{r}(t) \in \mathbb{R}^{3N}$  é o vetor com as posições cartesianas das partículas evoluindo sob um potencial  $V(\mathbf{r})$ , e  $\mathbf{M}$  é uma matriz diagonal positiva,  $3N$ -dimensional, com as massas atômicas:

$$\mathbf{M} = \text{diag}(m_1, m_1, m_1, \dots, m_N, m_N, m_N)$$

Esta aproximação leva em conta somente as posições cartesianas,  $\mathbf{r}$ , das moléculas. Porém, é mais conveniente expressar as equações do movimento em termos de coordenadas generalizadas,  $\mathbf{q} \in \mathbb{R}^{dN}$ , onde  $d$  é o número de graus de liberdade, e de seus momentos conjugados,  $\mathbf{p} \in \mathbb{R}^{dN}$ . (54, 63) Isto pode ser feito utilizando-se a formulação de Hamilton da dinâmica clássica. Nesta formulação, o sistema é descrito completamente por uma função Hamiltoniana,  $H$ , que é função das variáveis nucleares (lentas), sendo que o movimento (rápido) dos elétrons é tido como uma interação média. (54) Esta função pode ser representada como a soma das energias cinética,  $K$ , e potencial,  $V$ , do sistema (64):

$$H(\mathbf{P}, \mathbf{Q}) = K(\mathbf{P}) + V(\mathbf{Q}) \quad (4.3)$$

onde:  $\mathbf{P}$  é o conjunto com os momentos,  $\mathbf{p}_i = m_i \mathbf{v}_i \forall i \in \{1, \dots, N\}$ , associado às  $N$  partículas do sistema;  $\mathbf{Q}$  é o conjunto com as posições das  $N$  partículas no sistema.

As equações do movimento de Hamilton são:

$$\dot{\mathbf{p}}_i = \frac{d\mathbf{p}_i}{dt} = -\frac{\partial H}{\partial \mathbf{q}_i} \quad (4.4)$$

e

$$\dot{\mathbf{q}}_i = \frac{d\mathbf{q}_i}{dt} = \frac{\partial H}{\partial \mathbf{p}_i} \quad (4.5)$$

A energia cinética total é, então, descrita pela soma das energias cinéticas de todas as partículas do sistema e assume a seguinte forma:

$$K(\mathbf{P}, \mathbf{M}) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} \quad \forall i \in \{1, \dots, N\} \quad (4.6)$$

onde:  $N$  é o número total de partículas do sistema;  $m_i$  é a massa da partícula  $i$ ; e  $\mathbf{p}_i$  é o momento da  $i$ -ésima partícula. Desse modo, a energia cinética é uma função apenas das massas e dos momentos, sendo assim, independente das coordenadas das partículas que compõem o sistema. Deste modo:

$$\dot{\mathbf{q}}_j = \frac{\partial}{\partial \mathbf{p}_j} K(\mathbf{P}, \mathbf{M}) = \frac{\partial}{\partial \mathbf{p}_j} \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} = \frac{\mathbf{p}_j}{m_j} \quad \forall j \in \{1, \dots, N\} \quad (4.7)$$

A descrição da energia potencial, entretanto, é uma função composta, formada pela descrição de diversas formas de interações entre as partículas contidas num dado sistema, estando relacionada às forças que atuam sobre estas partículas conforme:

$$\mathbf{F} = -\nabla V(\mathbf{q}) \quad (4.8)$$

Uma vez que para sistemas de muitos corpos não se tem a descrição analítica da função  $V(\mathbf{q})$ , ela é aproximada como uma soma de potenciais entre pares de partículas (64):

$$V(\mathbf{q}) = \sum_{i=1}^{N-1} \sum_{j>i}^N V_{ij}(\mathbf{q}) \quad (4.9)$$

onde  $V(\mathbf{q})$  é a função potencial total descrita pelo campo de forças;  $(V_{ij}(\mathbf{q}))_{i \in \{1, \dots, N-1\}, j \in \{i, \dots, N\}}$  é o potencial atuando entre a  $i$ -ésima partícula e a partícula  $j$ ; e  $N$  é o número total de partículas do sistema. De modo que:

$$\dot{\mathbf{p}}_i = -\nabla_{\mathbf{q}_i} V(\mathbf{q}) = \mathbf{F}_i \quad \forall i \in \{1, \dots, N\} \quad (4.10)$$

Desta forma temos:

$$\frac{d\mathbf{q}_i}{dt} = \frac{\mathbf{p}_i}{m_i} \quad \forall i \in \{1, \dots, N\} \quad (4.11)$$

e

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i \quad \forall i \in \{1, \dots, N\} \quad (4.12)$$

Podemos utilizar estas equações para gerar um esquema de integração numérica. Como resultado deste processo, se obtém uma trajetória através do espaço de fases,  $\Gamma(\mathbf{p}, \mathbf{q}) \in \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ , (i.e. o espaço  $6N$ -dimensional gerado pelas coordenadas e momentos de todas as partículas) que especifica as posições e os momentos (ou velocidades) das partículas do sistema num dado tempo,  $t$ .(63, 65) Dentre os algoritmos dedicados a este processo de integração, um dos mais utilizados é o *Leap frog*.(63) O algoritmo *Leap frog* trabalha atualizando as velocidades em tempos intermediários em relação aos tempos nos quais as posições de cada partícula são atualizadas:(63, 65)

$$\mathbf{v}_i\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}_i\left(t - \frac{\Delta t}{2}\right) + \frac{\Delta t}{m_i} \mathbf{F}_i(t) \quad \forall i \in \{1, \dots, N\} \quad (4.13)$$

$$\mathbf{q}_i(t + \Delta t) = \mathbf{q}_i(t) + \Delta t \mathbf{v}_i\left(t + \frac{\Delta t}{2}\right) \quad \forall i \in \{1, \dots, N\} \quad (4.14)$$

onde  $(\mathbf{q}_i)_{i \in \{1, \dots, N\}}$ , é o vetor posição e  $(\mathbf{v}_i)_{i \in \{1, \dots, N\}}$  é o vetor velocidade da  $i$ -ésima partícula;  $(\mathbf{F}_i)_{i \in \{1, \dots, N\}}$  é o vetor força atuando na  $i$ -ésima partícula;  $(m_i)_{i \in \{1, \dots, N\}}$  é a massa desta partícula  $i$ ;  $\Delta t$  é o passo de integração. Para a iniciação deste algoritmo, são necessários os valores de posições e velocidades em duas etapas consecutivas. As velocidades num meio passo temporal futuro,  $\mathbf{v}(t + \frac{\Delta t}{2})$ , são calculadas a partir das velocidades num meio passo temporal passado,  $\mathbf{v}(t - \frac{\Delta t}{2})$ , e das forças que atuam nesta partícula no instante atual,  $\mathbf{F}(t)$ . De posse de  $\mathbf{v}(t + \frac{\Delta t}{2})$ , então podemos deduzir as posições em um passo de integração futuro,  $\mathbf{q}(t + \Delta t)$ .(63, 65) Em outras palavras, utiliza-se a força atuante sobre uma partícula no instante  $t$  e a velocidade num instante  $t - \frac{\Delta t}{2}$  para obter-se a velocidade no instante  $t + \frac{\Delta t}{2}$ . Em seguida, esta velocidade no instante  $t + \frac{\Delta t}{2}$  é utilizada com a posição no instante  $t$  pra atualizar a posição no instante  $t + \Delta t$ . Nesta nova posição é calculada a força atuante nesta partícula e então esta nova força é utilizada conjuntamente com a velocidade em  $t + \frac{\Delta t}{2}$  para atualizar novamente a velocidade. Este processo é repetido até que um número de passos desejados seja alcançado.

Os valores estruturais para a inicialização do algoritmo, ou seja, as posições iniciais das partículas no sistema podem ser obtidas por resolução de estruturas moleculares por técnicas como RMN, cristalografia de raios X ou modelagem por homologia. Atualmente, existem grandes bancos de dados, e.g. Protein Data Bank (68), com milhares de estruturas experimentais depositadas. As velocidades atômicas, cuja média estatística está relacionada à temperatura do sistema, podem ser obtidas a partir da distribuição de Maxwell-Boltzmann.(63, 65)

Entretanto, as funções de energia potencial,  $V(\mathbf{Q})$ , que atuam em cada átomo, não podem ser obtidas experimentalmente, devendo ser obtidas através de uma função potencial-energética empírica.(69) Como consequência, foram desenvolvidos modelos de

campos de força, baseados na Mecânica Estatística de forma a aproximar satisfatoriamente a função de energia potencial.

A grande parte dos campos de forças utilizados nas simulações de sistemas biológicos é expressa em termos da soma de potenciais de interações ligantes,  $V_{lig}$ , e não-ligantes,  $V_{n\tilde{a}o\ lig}$ :(23, 70)

$$V(\mathbf{Q}) = V_{lig} + V_{n\tilde{a}o\ lig} \quad (4.15)$$

onde a forma de  $V_{lig}$  e  $V_{n\tilde{a}o\ lig}$  é definida pelo campo de forças utilizado. Nestas funções de energia potencial, as interações ligantes,  $V_{lig}$ , ou seja, as ligações químicas, usualmente são descritas pelas interações entre dois (estiramento das ligações), três (ângulo de ligação) e quatro corpos (ângulos diedros). Existem dois tipos de interações em função dos ângulos diedros: diedros próprios, compreendendo o ângulo de torção entre quatro átomos consecutivos, e os diedros impróprios, que são usados para que os átomos se mantenham em um plano ou para prevenir que haja uma transição de configuração. Estas interações, geralmente, são modeladas utilizando-se funções senoidais que aproximam o perfil energético destas, entretanto, os diedros impróprios são modelados segundo um potencial harmônico. As interações não ligantes,  $V_{n\tilde{a}o\ lig}$ , são geralmente tratadas como interações de van der Waals, e modeladas como potenciais 6-12 de Lennard-Jones (71) e interações eletrostáticas, modeladas com base na lei de Coulomb.(72, 73)

O comportamento médio do sistema pode ser calculado em diferentes *ensembles* estatísticos, onde se pode realizar a média sobre todos os estados microscópicos do sistema que estão de acordo com certas propriedades microscópicas. Um *ensemble* é uma coleção de sistemas representativos de todos os possíveis estados microscópicos do sistema no espaço de fases que satisfazem certas distribuições de probabilidade.(63, 74, 75) Entretanto, a dinâmica molecular, na sua forma original, realiza a amostragem apenas do *ensemble* microcanônico (NVE), que corresponde a um sistema isolado, ou seja, são mantidos constantes o número de moléculas, o volume e a energia total do sistema.(63) Para se realizar a amostragem em outros *ensembles*, como por exemplo o canônico (NVT), onde a temperatura é mantida constante ao invés da energia total, pode-se utilizar um termostato, como o termostato de Nosé-Hoover.(76, 77) Para amostrarmos o *ensemble* isotérmico-isobárico (NPT), onde a pressão também é mantida constante, se pode acoplar ao sistema um barostato como o de Martyna-Tuckerman-Tobias-Klein (MTTK).(78)

## 4.2 Dinâmica de Langevin

A dinâmica molecular tal qual é descrita na seção 4.1 pode ser dita dinâmica clássica ou Newtoniana, uma vez que ela resolve as equações de movimento de Newton.

A resolução das equações de Newton levam a trajetórias determinísticas no espaço de fases levado em conta, ou seja o espaço  $6N$ -dimensional  $\Gamma(\mathbf{p}, \mathbf{q}) \in \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ . Entretanto, muitos problemas em química, biologia e física envolvem sistemas não homogêneos ou sistemas no qual a descrição do sistema deve levar em conta o comportamento global em longas escalas de tempo em relação a escala de tempo da DM.(79, 80) Infelizmente, o tamanho dos sistemas e as escalas de tempo necessárias para o estudo destes problemas geralmente estão além do que se pode calcular por dinâmica molecular com o poder de processamento computacional atual.(79)

Uma das possíveis abordagens para solucionar o problema da capacidade de processamento é distinguir entre os graus de liberdade relevantes daqueles graus de liberdade irrelevantes. Deve se ter em mente que estes graus de liberdades irrelevantes são importantes para o estudo do comportamento dos graus de liberdade relevantes, entretanto eles não demandam um conhecimento detalhado dos mesmos.(79)

Uma alternativa estocástica à dinâmica clássica é a dinâmica de Langevin (DL).(81) A dinâmica de Langevin tem por característica o uso de modelos simplificados que substituem os graus de liberdade irrelevantes, que descrevem o sistema, por equações diferenciais estocásticas. Em simulações de sistemas de interesse biológico, a DL tem grande aplicação em simulações com solvente implícito, onde a DL é usada para representar o efeito estocástico médio das moléculas de água explícita.(80, 82)

O modelo empregado na dinâmica de Langevin é fenomenológico, onde se adicionam um termo dissipativo e outro aleatório às forças do sistema.(74) Desta forma, em uma simulação de dinâmica de Langevin, as equações do movimento podem ser escritas como:(74, 81, 82)

$$\mathbf{M} \frac{d^2}{dt^2} \mathbf{q}(t) = -\nabla V(\mathbf{q}) - \gamma \mathbf{M} \frac{d}{dt} \mathbf{q}(t) + \mathbf{R}(t) \quad (4.16)$$

onde  $\mathbf{q}(t) \in \mathbb{R}^{3N}$  é o vetor com as posições cartesianas das partículas evoluindo sob um potencial  $V(\mathbf{q})$ , e  $\mathbf{M}$  é uma matriz diagonal positiva,  $3N$ -dimensional, com as massas atômicas.  $\gamma$  é o parâmetro que controla as colisões, também chamado de coeficiente de fricção, e  $\mathbf{R}(t)$  é o vetor com forças aleatórias. Desta forma, a dinâmica de Langevin trata-se de uma dinâmica clássica com o acréscimo de um termo aleatório e um termo de fricção.

O vetor de forças aleatórias,  $\mathbf{R}(t)$ , corresponde a um termo descrito como um processo Gaussiano com as seguintes propriedades:

$$\langle \mathbf{R}(t) \rangle = 0, \quad \langle \mathbf{R}(t) \mathbf{R}(t')^T \rangle = 2\gamma k_B T \mathbf{M} \delta(t - t') \quad (4.17)$$

onde  $k_B$  é a constante de Boltzmann,  $T$  é a temperatura e  $\delta$  é o delta de Dirac.

A magnitude de  $\gamma$  determina a intensidade relativa entre as forças inerciais e as forças randômicas, ou seja externas. Desta forma, com o aumento de  $\gamma$  o sistema passa de um regime inercial para o regime difusivo.(81) Quando  $\gamma \rightarrow 0$ , as equações de movimento de Langevin se tornam as equações de movimento de Newton.(82) Comparando uma simulação por dinâmica molecular com solvente explícito com uma onde o solvente é simulado por dinâmica de Langevin se pode observar que as frequências vibracionais características do soluto no vácuo são amortecidas.(83) Particularmente, todos os modos vibracionais de baixa frequência são mais amortecidos e várias funções de correlação são suavizadas.(81) A magnitude destas perturbações em relação ao comportamento Newtoniano se expressa em função da intensidade de  $\gamma$ .(84)

### 4.3 Dinâmica molecular fora do equilíbrio

O método de dinâmica molecular apresentado na Seção 4.1 considera um sistema mecânico conservativo de  $N$  partículas mantido a uma temperatura constante  $T$ . Uma vez que o sistema é conservativo, temos as seguintes restrições:

$$H(\mathbf{q}, \mathbf{p}) = E \quad \text{e} \quad \frac{dH}{dt} = 0 \quad (4.18)$$

Além disto, o Teorema de Liouville estabelece que a densidade de pontos no espaço de fases  $6N$ -dimensional é constante ao longo do tempo e o sistema é dito estar em *equilíbrio estatístico*.(85) Portanto, a dinâmica molecular tal qual esta descrita na Seção 4.1 possibilita o cálculo de trajetórias para um sistema apenas nas imediações de uma dada posição de equilíbrio no correspondente espaço de fases  $6N$ -dimensional,  $\Gamma(\mathbf{p}, \mathbf{q}) \in \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ , do sistema.

De forma a realizarmos simulações por dinâmica molecular onde queremos que o sistema evolua para posições longe da posição de equilíbrio inicial ou seja simulações de não-equilíbrio, devemos modificar o esquema clássico da dinâmica molecular. Uma das possíveis modificações será descrita na próxima seção e é a denominada **Dinâmica Molecular Dirigida** (DMD).

#### 4.3.1 Dinâmica molecular dirigida

A Dinâmica Molecular Dirigida (DMD)(24, 25, 26, 27) é uma extensão do método da dinâmica molecular clássica que possibilita a simulação de sistemas fora do equilíbrio. A DMD opera segundo a aplicação de uma força externa às partículas do sistema de forma a conduzi-las em uma determinada direção. Esta força é aplicada ao sistema na forma

de uma restrição harmônica, de forma que este é um método que pode ser utilizado para emular a ideia de um experimento de microscopia de força atômica, descrita na seção 3 a seguir.(26)

Se considerarmos apenas uma coordenada de reação  $\xi$  dependente da posição dos átomos  $\xi(\mathbf{q})$ , sendo restringida por um potencial harmônico

$$V[\xi(\mathbf{q}); \lambda(t)] = \frac{1}{2}k \left( \xi(\mathbf{q}) - \lambda(t) \right)^2 \quad (4.19)$$

onde  $k$  é a constante de força da restrição e  $\lambda(t)$  é o valor do parâmetro externo de controle dependente do tempo para esta coordenada de reação, a força atuando no sistema será dada por:

$$\begin{aligned} \mathbf{F}(t) &= -\nabla_{\xi(\mathbf{r})} V(\xi(\mathbf{q}); \lambda(t)) \\ &= -k \left( \xi(\mathbf{q}) - \lambda(t) \right) \end{aligned} \quad (4.20)$$

Esta força é adicionada às forças devidas à dinâmica do sistema segundo os métodos padrões de restrição de dinâmica molecular.(54, 86) Para a  $i$ -ésima partícula, sujeita a  $g$  restrições holonômicas, a força total atuando sobre esta partícula será:

$$\mathbf{F}_{\text{total}}^i = \mathbf{F}_V^i - \sum_{k=1}^g \Lambda_k \nabla_{\xi(\mathbf{q})} V_k[\xi(\mathbf{q}); \lambda(t)] \quad \forall i \in 1, \dots, N \quad (4.21)$$

onde  $\mathbf{F}_V^i$  é a força resultante do campo de forças do sistema e  $\Lambda_k$  são multiplicadores de Lagrange.(54)

É importante observar que, uma vez que o potencial adicionado ao sistema, Equação 4.19, é uma função dependente do tempo, devido ao parâmetro externo de controle  $\lambda(t)$  que é uma função dependente do tempo, o sistema não é mais conservativo, e consequentemente as Equações 4.18 não são mais válidas, sendo que agora o Hamiltoniano do sistema pode ser descrito como:

$$\tilde{H}(\mathbf{q}, \mathbf{p}; \lambda(t)) = H(\mathbf{q}, \mathbf{p}) + V(\mathbf{q}; \lambda(t)) \quad (4.22)$$

Desta forma, temos as seguintes restrições para o sistema:

$$\tilde{H}(\mathbf{q}, \mathbf{p}; \lambda(t)) = E(t) \quad \text{e} \quad \frac{d\tilde{H}}{dt} = f(\mathbf{q}; t) \quad (4.23)$$

E, consequentemente, as trajetórias obtidas desta forma podem evoluir para longe de uma dada posição de equilíbrio no correspondente espaço de fases  $6N$ -dimensional,  $\Gamma(\mathbf{p}, \mathbf{q}) \in \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ , do sistema.

## 5 Energia Livre

Em 1944, o grande físico austríaco Ervin Schrödinger defendeu em seu livro "*What is life*" (87) que os organismos permanecem vivos se mantendo afastados do equilíbrio e o fazem trocando energia livre com o meio ambiente. (88) De fato, a energia livre é de suma importância para os processos biológicos. A energia livre representa a maior quantidade de energia que pode ser recuperada na forma de trabalho útil. (5) Desta forma, nas últimas décadas, o cálculo da energia livre se tornou uma ferramenta indispensável e inestimável para ajudar a explicar profundas questões biológicas, que apenas com a simples observação dos experimentos continuariam sem resolução. (89) De fato, diversas propriedades químicas (e.g. constantes de equilíbrio e coeficientes de partição e adsorção) podem ser relacionadas à diferença de energia livre entre dois estados de um sistema. (79, 90)

O comportamento dos sistemas biomoleculares pode ser descrito pelas leis da mecânica estatística. (23, 79, 91) Desta forma, para um sistema clássico onde o número de partículas  $N$ , o volume  $V$  e a temperatura  $T$  são mantidos constantes (ou seja no denominado *ensemble* canônico) onde a dinâmica do sistema é determinada pelo Hamiltoniano  $H(\mathbf{p}_1, \dots, \mathbf{p}_N, \mathbf{r}_1, \dots, \mathbf{r}_N)$ , c.f. Capítulo 4 e Equação 4.3, a probabilidade  $p(\mathbf{p}_1^i, \dots, \mathbf{p}_N^i, \mathbf{r}_1^i, \dots, \mathbf{r}_N^i)$  do sistema se encontrar em um estado  $\mathbf{p}_1^i, \dots, \mathbf{p}_N^i, \mathbf{r}_1^i, \dots, \mathbf{r}_N^i$  é dada por:

$$p(\mathbf{p}_1^i, \dots, \mathbf{p}_N^i, \mathbf{r}_1^i, \dots, \mathbf{r}_N^i) = \frac{e^{-\beta H(\mathbf{p}_1^i, \dots, \mathbf{p}_N^i, \mathbf{r}_1^i, \dots, \mathbf{r}_N^i)}}{\int_{\Gamma} \dots \int_{\Gamma} e^{-\beta H(\mathbf{p}_1, \dots, \mathbf{p}_N, \mathbf{r}_1, \dots, \mathbf{r}_N)} d\mathbf{p}_1 \dots d\mathbf{p}_N d\mathbf{r}_1 \dots d\mathbf{r}_N} \quad (5.1)$$

onde  $\beta^{-1} = k_B T$ ,  $k_B$  é a constante de Boltzmann e  $\Gamma(\mathbf{p}_1, \dots, \mathbf{p}_N, \mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ , é o espaço de fases.

Desta forma, a energia livre de Helmholtz  $F$  do sistema é dada por: (74, 79, 90)

$$F = -k_B T \ln \left\{ c \int_{\Gamma} \dots \int_{\Gamma} e^{-\beta H(\mathbf{p}_1, \dots, \mathbf{p}_N, \mathbf{r}_1, \dots, \mathbf{r}_N)} d\mathbf{p}_1 \dots d\mathbf{p}_N d\mathbf{r}_1 \dots d\mathbf{r}_N \right\} \quad (5.2)$$

onde  $c$  é uma constante que está relacionada à indistinguibilidade das partículas tratadas classicamente. Deve-se notar que a integração deve ser realizada sobre todo o espaço  $\Gamma(\mathbf{p}_1, \dots, \mathbf{p}_N, \mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ , ou seja o espaço  $6N$ -dimensional, gerado pelas coordenadas e momentos de todas as partículas.

Entretanto, para a maioria das aplicações práticas, estamos interessados em calcular apenas a diferença de energia livre relativa  $\Delta F_{BA} = F_B - F_A$  entre dois estados,  $A$  e  $B$ , do sistema. Também se pode calcular a diferença de energia livre relativa entre

dois estados do sistema conectados por um parâmetro de acoplamento  $\lambda$ , cuja dinâmica é descrita pelo  $H(\mathbf{p}_1, \dots, \mathbf{p}_N, \mathbf{r}_1, \dots, \mathbf{r}_N; \lambda)$ , e portanto(90, 92):

$$p(\mathbf{p}_1^i, \dots, \mathbf{p}_N^i, \mathbf{r}_1^i, \dots, \mathbf{r}_N^i | \lambda_j) = \frac{e^{-\beta H(\mathbf{p}_1^i, \dots, \mathbf{p}_N^i, \mathbf{r}_1^i, \dots, \mathbf{r}_N^i; \lambda_j)}}{\int_{\Gamma} \dots \int e^{-\beta H(\mathbf{p}_1, \dots, \mathbf{p}_N, \mathbf{r}_1, \dots, \mathbf{r}_N; \lambda_j)} d\mathbf{p}_1 \dots d\mathbf{p}_N d\mathbf{r}_1 \dots d\mathbf{r}_N} \quad (5.3)$$

e, consequentemente:

$$F(\lambda_j) = -k_B T \ln \left\{ c \int_{\Gamma} \dots \int e^{-\beta H(\mathbf{p}_1, \dots, \mathbf{p}_N, \mathbf{r}_1, \dots, \mathbf{r}_N; \lambda_j)} d\mathbf{p}_1 \dots d\mathbf{p}_N d\mathbf{r}_1 \dots d\mathbf{r}_N \right\} \quad (5.4)$$

Desta forma, a diferença de energia livre entre dois estados,  $\lambda_A$  e  $\lambda_B$ , é uma medida logarítmica da relação entre as funções de partição referentes a cada estado de um mesmo Hamiltoniano, mas integrado sobre diferentes regiões do espaço de fases do sistema.(90, 92)

Embora diversos esquemas de simulação computacional tenha sido desenvolvidos para a aplicação das Equações 5.2 e 5.4, geramente não é possível realizar o cálculo de tais integrais.(79, 90) Os únicos casos tratáveis são daqueles sistemas com distribuições homogêneas ou aqueles cuja função de distribuição possa ser aproximada por uma distribuição gaussiana, e.g. potenciais harmônicos.(79) Devido ao elevado número de átomos, complexidade matemática do campo de forças, i.e. a forma funcional do Hamiltoniano, e limite da capacidade de calculo computacional, o uso direto das Equações 5.2 e 5.4 se torna impraticável para o uso direto em sistemas biológicos, levando a longas e proibitivas simulações computacionais.

O progresso na teoria por de trás dos cálculos de energia livre e o aumento da capacidade dos processadores, tornaram a determinação da energia livre por simulações computacionais um processo factível.(89, 92, 93) Em termos de aplicação, as metodologias para o cálculo da energia livre para sistemas biológicos podem ser divididas em duas categorias(79, 89): transformações geométricas, e.g. restrição de posições ou araste de alguns grupos atômicos através de uma coordenada de reação, e transformações "alquímic", na qual uma molécula é transformada em outra pelo aparecimento ou desaparecimento de um ou mais átomos durante a simulação.

Do ponto de vista metológico, a aplicação destas transformações geraram uma miríade de métodos computacionais dedicados a implantação destes.(90) Estes métodos podem ser classificados em quatro categorias(79, 89):

1. Métodos baseados em probabilidades de distribuição e histogramas;
2. Métodos perturbativos;

3. Métodos baseados em gradientes;
4. Métodos de não-equilíbrio.

Dentre estes métodos, os métodos de não-equilíbrio são uma opção viável para o estudo de processos onde a escala de tempo necessária para que este ocorra não está dentro do factível de uma simulação de dinâmica molecular em equilíbrio.(89) Nos métodos de não-equilíbrio, como a dinâmica molecular dirigida descrita na Seção 4.3.1, uma força externa é imposta ao sistema, paralela a uma determinada coordenada de reação, fazendo com que o sistema evolua de um estado a outro.

Quando um parâmetro do sistema for variado ao longo de uma coordenada de reação com uma velocidade *infinitamente lenta* de um ponto inicial  $\lambda_A$  até um ponto final  $\lambda_B$ , a termoestática determina que o trabalho total  $W$  realizado sobre o sistema é igual a energia livre de Helmholtz(94):

$$W = \Delta F = F_{\lambda_B} - F_{\lambda_A} \quad (5.5)$$

Entretanto, devido às velocidades de afastamento relativamente grandes utilizadas nas simulações de dinâmica molecular dirigida, necessárias para que os eventos estudados fiquem acessíveis a escala temporal desta metodologia, que chegam a ser várias ordens de grandeza maior que as utilizadas experimentalmente(95), geram uma quantidade de trabalho dissipativo, de forma que a Equação 5.5 não é mais válida.(89, 95) Neste caso, temos que a energia livre é, em média, menor ou igual ao trabalho realizado sobre o sistema, de forma que este trabalho se torna um limite superior para a energia livre(96):

$$\langle W \rangle \equiv \int \rho(W)W \, dW \geq \Delta F \quad (5.6)$$

onde  $\langle \rangle$  denotam a média sobre um *ensemble* de medidas de  $W$ .

Esta dificuldade veio por terra em 1997 quando Christopher Jarzynski publicou a equação de leva seu nome relacionando a energia livre de equilíbrio à média de *ensemble* de processos realizados em não-equilíbrio.(21) A Seção 5.1, a seguir, tratará desta igualdade.

## 5.1 Igualdade de Jarzynski

Ao realizarmos trabalho sobre um sistema clássico, de forma a mover algum parâmetro de controle  $\lambda$  ao longo de um determinado caminho entre dois estados de equilíbrio, a termoestática nos diz que só iremos obter a igualdade entre o trabalho realizado sobre o sistema e energia livre entre estes dois estados,  $W = \Delta F$ , se e somente se a mudança de  $\lambda$  ocorrer com uma velocidade *infinitamente lenta*. Devido à necessidade de longas simulações de equilíbrio entre os passos em um determinado caminho, aliado ao número

de passos em si, do ponto de vista computacional, esta limitação torna impraticável o cálculo da energia livre para muitos sistemas de interesse químico, físico e biológico. Também, o estudo de muitos destes processos, por ocorrerem em regime de não-equilíbrio ou por ocorrerem em escalas de tempo inacessíveis do ponto de vista da atual capacidade computacional necessitam da aplicação de uma força externa que direcione o fenômeno estudado, acelerando o processo, e com isso a diferença de energia livre será menor ou igual ao trabalho aplicado ao sistema:  $W \geq \Delta F$ .

Entretanto, no ano de 1997, Christopher Jarzynski publicou um artigo seminal na revista *Physical Review Letters* que iria contornar esta limitação de forma marcante.(21) Jarzynski(21) chegou a seguinte fórmula, agora denominada igualdade de Jarzynski:

$$\Delta F = -k_B T \ln \left\langle e^{-\frac{W}{k_B T}} \right\rangle \quad (5.7)$$

onde  $\Delta F$  é a energia livre entre dois estados,  $A$  e  $B$ , em equilíbrio,  $k_B$  é a constante de Boltzmann e  $T$  a temperatura do sistema.  $\langle \rangle$  denota a média sobre um *ensemble* de caminhos de não-equilíbrio entre os estados  $A$  e  $B$ .

A Equação 5.7 mostra que, apesar da diferença de energia livre entre dois estados possuir como limite superior o trabalho realizado em regime de não-equilíbrio entre estes dois sistemas, a diferença de energia livre entre dois estados é igual a média da exponencial do trabalho realizado em regime de não-equilíbrio. Ou seja, a Equação 5.7 possibilitou que uma desigualdade termodinâmica fosse reescrita como uma igualdade, fazendo a ligação entre os regimes de equilíbrio e não-equilíbrio. De fato, esta é a consequência mais notável da igualdade derivada por Jarzynski(21). A Equação 5.7 permite extrair, independente do caminho tomado entre dois estados ou da velocidade de evolução do parâmetro externo  $\lambda$ , uma informação de equilíbrio,  $\Delta F$  a partir de um *ensemble* de medidas realizadas em regime de não-equilíbrio.(21)

A igualdade de Jarzynski foi posta à prova experimentalmente, primeiramente, no início dos anos 2000, onde a Equação 5.7 foi utilizada na obtenção da diferença de energia livre do processo de desenovelamento de ácidos nucleicos.(9, 11, 12, 16, 17, 19, 20) Além de permitir o melhor entendimento dos experimentos em regime de não-equilíbrio, o estabelecimento da Equação 5.7 permitiu o desenvolvimento de métodos numéricos para o cálculo da diferença de energia livre.(38, 39, 96)

A Seção 5.1.0.1, a seguir, demonstra uma derivação da Equação 5.7

### 5.1.0.1 Derivação da igualdade de Jarzynski

Diversas provas para a igualdade de Jarzynski foram dadas na literatura, em diferentes graus de complexidade ou facilidade de compreensão.(21, 79, 93, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106) Crooks(98), Crooks(99), Hummer e Szabo(100), Athènes(101),

Schurr e Fujimoto(102), Berendsen(79), Liang et al.(93) e Sun(103) demonstraram a igualdade de Jarzynski para sistemas evoluindo segundo uma dinâmica estocástica. Page(104), Gittes(105) demonstraram a igualdade para sistemas quânticos e Cuendet(106) realizou a demonstração desta igualdade para sistemas não Hamiltonianos. Crooks(98) escreveu uma das provas mais compreensíveis e, uma vez que sua prova possibilita estabelecer alguns critérios para a validade da igualdade de Jarzynski, esta será seguida na demonstração a seguir.

Primeiramente devemos caracterizar o sistema e postular algumas propriedades deste. Seja um sistema com volume  $V$  e número de partículas  $N$  constantes, cuja evolução temporal é governada estocasticamente pela dinâmica de Langevin, em contato com um grande reservatório térmico mantido a temperatura  $T$ . O estado do sistema em um determinado tempo  $t$  é especificado pelo estado interno  $i_t$  do sistema e pelo parâmetro externamente controlado  $\lambda_t$ . Desta forma, a energia do sistema, no tempo  $t$  é uma função  $E(i_t, \lambda_t)$ .

Uma vez que estamos tratando o sistema como pertencente ao *ensemble* canônico,  $NVT$  constantes, a probabilidade de equilíbrio de um determinado estado  $A$  ocorrer, dado um valor fixo do parâmetro de controle  $\lambda$  é dada por:

$$P(A|\lambda) = \frac{e^{-\beta E(A,\lambda)}}{\sum_i e^{-\beta E(i,\lambda)}} \quad (5.8)$$

onde  $\beta = 1/k_B T$ ,  $k_B$  a constante de Boltzmann. Neste ponto, se deve notar que a igualdade de Jarzynski requer que o processo de evolução temporal conserve a distribuição canônica para cada valor de  $\lambda_t$  assumido.(79)

Por conveniência, será assumido que o tempo e o espaço de fases são discretos e que a evolução temporal do sistema ocorre com o parâmetro externo se movendo em uma sequência fixa  $\lambda_i \in \{\lambda_0, \lambda_1, \dots, \lambda_\tau\}$ . Inicialmente, tempo  $t = 0$ , o sistema se encontra no estado  $i_0$  com o parametro de controle  $\lambda_0$  e, desta forma, um caminho direto no espaço de fases pode ser descrito como:

$$i_0 \xrightarrow{\lambda_1} i_1 \xrightarrow{\lambda_2} i_2 \xrightarrow{\lambda_3} \dots \xrightarrow{\lambda_\tau} i_\tau \quad (5.9)$$

O caminho reverso, correspondendo a reversão do tempo, pode ser descrito como:

$$i_0 \xleftarrow{\lambda_1} i_1 \xleftarrow{\lambda_2} i_2 \xleftarrow{\lambda_3} \dots \xleftarrow{\lambda_\tau} i_\tau \quad (5.10)$$

Assumiremos que a evolução do sistema ocorre em duas etapas. Primeiro o parâmetro de controle  $\lambda_t$  evolui para um novo valor  $\lambda_{t+1}$ . Nesta etapa é realizado um trabalho

igual a

$$W_t = E(i_t, \lambda_{t+1}) - E(i_t, \lambda_t)$$

então o estado interno do sistema  $i_t$  evolui para um novo estado  $i_{t+1}$ . Esta segunda etapa faz com que o sistema troque uma quantidade

$$Q_t = E(i_{t+1}, \lambda_{t+1}) - E(i_t, \lambda_{t+1})$$

com o reservatório térmico. Esta evolução no espaço de fases é repetida por  $\tau$  espaços de tempo.

Desta forma, o trabalho total realizado sobre o sistema  $W$  e o calor total trocado com o reservatório térmico  $Q$  são dados pela soma sobre todos os passos realizados:

$$W = \sum_{t=0}^{\tau-1} E(i_t, \lambda_{t+1}) - E(i_t, \lambda_t) \quad (5.11)$$

$$Q = \sum_{t=1}^{\tau} E(i_t, \lambda_t) - E(i_{t-1}, \lambda_t) \quad (5.12)$$

A energia livre entre dois estados  $A$  e  $B$  em equilíbrio é dado pelo trabalho reversível

$$W_r = \Delta F = F(\beta, \lambda_B) - F(\beta, \lambda_A) \quad (5.13)$$

O trabalho dissipativo é definido como a diferença entre o trabalho total realizado sobre o sistema, Equação 5.11 e o trabalho reversível, Equação 5.13.

$$W_d = W - W_r \quad (5.14)$$

É interessante notar que o trabalho dissipativo depende do caminho percorrido pelo sistema entre dois estados, enquanto que o trabalho reversível só depende dos estados inicial e final.

Utilizando as Equações 5.11 e 5.12 podemos obter a variação na energia total  $\Delta E$ :

$$\Delta E = Q + W \quad (5.15)$$

$$\begin{aligned} &= \sum_{t=1}^{\tau} E(i_t, \lambda_t) - E(i_{t-1}, \lambda_t) + \sum_{t=0}^{\tau-1} E(i_t, \lambda_{t+1}) - E(i_t, \lambda_t) \\ &= E(i_1, \lambda_1) - E(i_0, \lambda_1) + E(i_2, \lambda_2) \\ &\quad - E(i_1, \lambda_2) + \cdots + E(i_\tau, \lambda_\tau) - E(i_{\tau-1}, \lambda_\tau) \\ &\quad + E(i_0, \lambda_1) - E(i_0, \lambda_0) + E(i_1, \lambda_2) \\ &\quad - E(i_1, \lambda_1) + \cdots + E(i_{\tau-1}, \lambda_\tau) - E(i_\tau, \lambda_\tau) \\ \therefore \Delta E &= E(i_\tau, \lambda_\tau) - E(i_0, \lambda_0) \end{aligned} \quad (5.16)$$

Por fim, definimos que a evolução do sistema seja um processo de Markov, ou seja, para o caminho direto representado pela Equação 5.9, onde o parâmetro externo evolui em uma sequência fixa e sucessiva  $\lambda_i \in \{\lambda_0, \lambda_1, \dots, \lambda_\tau\}$ , a probabilidade condicional de se encontrar o sistema a um tempo  $\tau$  é dado por(98, 107):

$$P(i_0 \xrightarrow{\lambda_1} i_1 \xrightarrow{\lambda_2} i_2 \xrightarrow{\lambda_3} \dots \xrightarrow{\lambda_\tau} i_\tau) = P(i_0 \xrightarrow{\lambda_1} i_1) P(i_1 \xrightarrow{\lambda_2} i_2) \dots P(i_{\tau-1} \xrightarrow{\lambda_\tau} i_\tau) \quad (5.17)$$

onde  $P(A \xrightarrow{\lambda} B)$  é denominada probabilidade de transição.

Ao assumirmos que o sistema é microscopicamente reversível para todo tempo e, portanto, obedecem a condição do balanço detalhado para todos os valores do parâmetro externo de controle  $\lambda$ , que utilizando a Equação 5.8:

$$\frac{P(A \xrightarrow{\lambda} B)}{P(A \xleftarrow{\lambda} B)} = \frac{P(B|\lambda)}{P(A|\lambda)} = \frac{e^{-\beta E(B,\lambda)}}{e^{-\beta E(A,\lambda)}} \quad (5.18)$$

Considerando os caminhos direto, Equação 5.9, e inverso, Equação 5.10, aplicando a condição do balanço detalhado, Equação 5.18, para múltiplos intervalos de tempo e utilizando a propriedade do processo de Markov, Equação 5.17 obtemos:

$$\begin{aligned} \frac{P(i_0|\lambda_0)P(i_0 \xrightarrow{\lambda_1} i_1 \xrightarrow{\lambda_2} i_2 \dots \xrightarrow{\lambda_\tau} i_\tau)}{P(i_\tau|\lambda_\tau)P(i_0 \xleftarrow{\lambda_1} i_1 \xleftarrow{\lambda_2} i_2 \dots \xleftarrow{\lambda_\tau} i_\tau)} &= \frac{P(i_0|\lambda_0)P(i_0 \xrightarrow{\lambda_1} i_1)P(i_1 \xrightarrow{\lambda_2} i_2) \dots P(i_{\tau-1} \xrightarrow{\lambda_\tau} i_\tau)}{P(i_\tau|\lambda_\tau)P(i_0 \xleftarrow{\lambda_1} i_1)P(i_1 \xleftarrow{\lambda_2} i_2) \dots P(i_{\tau-1} \xleftarrow{\lambda_\tau} i_\tau)} \\ &= \frac{P(i_0|\lambda_0)P(i_1|\lambda_1)P(i_2|\lambda_2) \dots P(i_\tau|\lambda_\tau)}{P(i_\tau|\lambda_\tau)P(i_0|\lambda_1)P(i_1|\lambda_2) \dots P(i_{\tau-1}|\lambda_\tau)} \\ &= \frac{P(i_0|\lambda_0)e^{-\beta E(i_0,\lambda_0)}e^{-\beta E(i_1,\lambda_1)} \dots e^{-\beta E(i_\tau,\lambda_\tau)}}{P(i_\tau|\lambda_\tau)e^{-\beta E(i_0,\lambda_1)}e^{-\beta E(i_1,\lambda_2)} \dots e^{-\beta E(i_{\tau-1},\lambda_\tau)}} \\ &= \frac{P(i_0|\lambda_0)}{P(i_\tau|\lambda_\tau)} e^{-\beta Q} \end{aligned} \quad (5.19)$$

Rearranjando a Equação 5.8, obtemos:

$$P(A|\lambda) = \frac{e^{-\beta E(A,\lambda)}}{\sum_i e^{-\beta E(i,\lambda)}} = e^{\beta F(\beta,\lambda) - \beta E(A,\lambda)} \quad (5.20)$$

onde  $F(\beta, \lambda) = \beta^{-1} \ln \sum_i e^{-\beta E(i,\lambda)}$ , e utilizando as Equações 5.13, 5.14 e 5.15, a Equação 5.19 se torna:

$$\begin{aligned} \frac{P(i_0|\lambda_0)P(i_0 \xrightarrow{\lambda_1} i_1 \xrightarrow{\lambda_2} i_2 \dots \xrightarrow{\lambda_\tau} i_\tau)}{P(i_\tau|\lambda_\tau)P(i_0 \xleftarrow{\lambda_1} i_1 \xleftarrow{\lambda_2} i_2 \dots \xleftarrow{\lambda_\tau} i_\tau)} &= \frac{P(i_0|\lambda_0)}{P(i_\tau|\lambda_\tau)} e^{-\beta Q} \\ &= e^{\beta \Delta E - \beta \Delta F} e^{-\beta Q} \\ &= e^{+\beta W_d} \end{aligned} \quad (5.21)$$

Por fim, definimos a média da exponencial do trabalho  $e^{-\beta W}$  sobre todos os caminhos diretos  $i_0 \xrightarrow{\lambda_1} i_1 \xrightarrow{\lambda_2} i_2 \xrightarrow{\lambda_3} \dots \xrightarrow{\lambda_\tau} i_\tau$  pelo espaço de fases:

$$\langle e^{-\beta W} \rangle = \sum_{\{i_0, i_1, \dots, i_\tau\}} P(i_0 | \lambda_0) P(i_0 \xrightarrow{\lambda_1} i_1 \xrightarrow{\lambda_2} i_2 \dots \xrightarrow{\lambda_\tau} i_\tau) e^{-\beta W} \quad (5.22)$$

Entretanto, a média sobre os caminhos diretos podem, usando a Equação 5.21, pode ser trocada para a média sobre os caminhos inversos  $i_0 \xleftarrow{\lambda_1} i_1 \xleftarrow{\lambda_2} i_2 \xleftarrow{\lambda_3} \dots \xleftarrow{\lambda_\tau} i_\tau$ :

$$\langle e^{-\beta W} \rangle = \sum_{\{i_0, i_1, \dots, i_\tau\}} P(i_\tau | \lambda_\tau) P(i_0 \xleftarrow{\lambda_1} i_1 \xleftarrow{\lambda_2} i_2 \dots \xleftarrow{\lambda_\tau} i_\tau) e^{+\beta W_d} e^{-\beta W} \quad (5.23)$$

Usando as definições de trabalho dissipativo  $W_d = W - W_r$  e trabalho reversível  $W_r = \Delta F$ , bem como a propriedade de que o trabalho reversível é independente do caminho percorrido pelo sistema e a propriedade de que as propabilidades de transição são normalizadas, a Equação 5.23 se torna:

$$\begin{aligned} \langle e^{-\beta W} \rangle &= \sum_{\{i_0, i_1, \dots, i_\tau\}} P(i_\tau | \lambda_\tau) P(i_0 \xleftarrow{\lambda_1} i_1 \xleftarrow{\lambda_2} i_2 \dots \xleftarrow{\lambda_\tau} i_\tau) e^{+\beta W_d} e^{-\beta W} \\ &= \sum_{\{i_0, i_1, \dots, i_\tau\}} P(i_\tau | \lambda_\tau) P(i_0 \xleftarrow{\lambda_1} i_1 \xleftarrow{\lambda_2} i_2 \dots \xleftarrow{\lambda_\tau} i_\tau) e^{-\beta \Delta F} \\ &= e^{-\beta \Delta F} \sum_{\{i_0, i_1, \dots, i_\tau\}} P(i_\tau | \lambda_\tau) P(i_0 \xleftarrow{\lambda_1} i_1 \xleftarrow{\lambda_2} i_2 \dots \xleftarrow{\lambda_\tau} i_\tau) \\ &= e^{-\beta \Delta F} \end{aligned} \quad (5.24)$$

Extraindo o logarítmo natural de ambos os lados da Equação 5.24 e multiplicando ambos os lados da equação resultante por  $-\beta^{-1}$  obtemos a Equação 5.7

$$\Delta F = -k_b T \ln \left\langle e^{-\frac{W}{k_B T}} \right\rangle \quad (5.7)$$

### 5.1.1 Decomposição da energia livre em energia interna e entropia

A energia livre de um sistema reflete a competição entre energia e entropia.(157) Em outras palavras, do ponto de vista microscópico é um balanço entre a energia interna deste sistema e o número de microestados acessíveis ao mesmo.(74, 158) De fato, a possibilidade de acessar separadamente as contribuições da energia interna e as entrópicas da energia livre de um processo é de suma importância para a compreensão da força motriz que rege este.(159) Em química e química-física, a energia interna de um sistema é, geralmente, considerada como aqueles tipos de energia que podem sofrer alteração por um processo químico, como energia translacional, vibracional e rotacional das moléculas,

bem como as parcelas de energia envolvidas nas ligações químicas e nas interações não ligantes.(47)

### 5.1.1.1 Energia interna

Para se obter a energia interna de um sistema de não-equilíbrio, podemos utilizar o seguinte teorema:(99)

$$\langle \mathcal{F} \rangle_R = \langle \hat{\mathcal{F}} e^{-\beta \mathcal{W}_d} \rangle_D \quad (5.25)$$

Onde  $\mathcal{F}$  e  $\hat{\mathcal{F}}$  é uma função de caminho e sua inversa no tempo, respectivamente.  $\mathcal{W}_d = W - \Delta F$  é o trabalho dissipativo e os índices  $D$  e  $R$  denotam a média sobre os caminhos direto e inverso, respectivamente.

Realizando a média de uma função  $\mathcal{F}$  de estado do sistema, no estado final do processo direto, i.e.  $\mathcal{F} = f[x(\tau)]$ , e a média desta função no estado inicial do processo reverso, i.e.  $\hat{\mathcal{F}} = f[\hat{x}(0)]$ , temos:

$$\langle f[x(\tau)] e^{-\beta \mathcal{W}_d} \rangle_D = \langle f[\hat{x}(\tau)] \rangle_R \quad (5.26)$$

Desta forma, um vez que no processo reverso a média é realizada sobre o *ensemble* inicial, de equilíbrio, do sistema e sua dinâmica subsequente é irrelevante, se pode trocar os índices  $D$  e  $R$  do processo reverso e inverso pelos índices que indicam as médias de não-equilíbrio e equilíbrio, respectivamente.(99) Portanto, energia interna para um sistema de não-equilíbrio pode ser obtida pelo seguinte teorema de flutuação<sup>1</sup>:

$$\langle f[x(\tau)] e^{-\beta \mathcal{W}_d} \rangle_{neq} = \langle f[x(\tau)] \rangle_{eq} \quad (5.27)$$

Uma vez que estamos interessado na energia interna,  $f[x(\tau)] = U(x)$ . Utilizando  $\lambda$  como coordenada de reação, a energia interna do estado final  $\lambda = 1$ , pode ser obtida por:(99, 159)

$$\langle U(x) \rangle_{\lambda=1} = \langle U(x) e^{-\beta \mathcal{W}_d} \rangle = \frac{\langle U(x) e^{-\beta W} \rangle}{\langle e^{-\beta W} \rangle} \quad (5.28)$$

<sup>1</sup> É interessante notar que, se a função  $f[x]$  for definida como uma constante  $c$  para todo o domínio de tempo, a Equação 5.27 retorna diretamente a Igualdade de Jarzynski, Equação 10.1:

$$\begin{aligned} \langle c e^{-\beta \mathcal{W}_d} \rangle_{neq} &= \langle c \rangle_{eq} \\ c \langle e^{-\beta \mathcal{W}_d} \rangle_{neq} &= c \end{aligned}$$

Uma vez que  $\mathcal{W}_d = W - \Delta F$ , obtemos a igualdade de Jarzynski:

$$\langle e^{-\beta w} \rangle_{neq} = e^{-\beta \Delta F}$$

Portanto, usando a Equação 5.28, a diferença de energia interna entre um estado inicial ( $\lambda = 0$ ) e um outro estado ( $\lambda$ ) é dada pela equação:

$$\Delta U = \frac{\langle U(x)e^{-\beta W} \rangle}{\langle e^{-\beta W} \rangle} - \langle U(x=0) \rangle \quad (5.29)$$

onde a energia  $U$  pode ser obtida do Hamiltoniano  $H(\mathbf{r}, \mathbf{p})$  do sistema.

### 5.1.1.2 Entropia do sistema

Como mencionado no parágrafo inicial desta Seção, Seção 5.1.1, a energia livre de um sistema é composta por duas parcelas: energia interna e entropia ponderada pela temperatura do sistema. Esta relação, considerando como variáveis independentes a temperatura e volume, é dada pela transformada de Legendre da energia interna:(74, 94)

$$F(V, T) = U - TS \quad (5.30)$$

Onde  $F$  é a energia livre de Helmholtz,  $U$  a energia interna e  $S$  a entropia.

Do ponto de vista da simulação computacional de sistemas biomoleculares, o comportamento destes sistemas são regidos pelas leis da mecânica estatística.(23, 79) Entretanto, da perspectiva da mecânica estatística, a entropia para um sistema que evolui com temperatura, volume e quantidade de matéria constantes, ou seja no *ensemble* NVT, é obtida pela soma entre um termo correspondente a uma integral sobre todo o espaço de fases e um termo correspondente a variação desta integral sobre todo o espaço de fases:(74, 75, 80)

$$S_{NVT} = k_B T \left( \frac{\partial \ln Q}{\partial T} \right)_{N,V} + k_B \ln Q \quad (5.31)$$

Onde  $k_B$  é a constante de Boltzman,  $T$  a temperatura do sistema e

$$Q(N, V, T) = \sum_j e^{-\beta E_j(N,T)}$$

é a função de partição canônica.

A utilização da Equação 5.31 para a obtenção da entropia se torna proibitiva para o calculo desta propriedade para qualquer sistema biomolecular, devido a necessidade de se conhecer todo o espaço de fases.(23) Entretanto, se pode rearranjar a Equação 5.30 de forma que a entropia seja uma função de  $F$  e  $U$ :

$$T \Delta S_{NVT} = \Delta U - \Delta F \quad (5.32)$$

e utilizarmos os valores de  $\Delta F$  e  $\Delta U$ , obtido das equações 5.25 e 5.29, respectivamente.

Parte II

Metodologia

## 6 METODOLOGIA

Neste capítulo estão descritos os métodos experimentais executados pelo autor para a elaboração desta Tese. Todas as simulações por dinâmica molecular foram realizadas no cluster local do laboratório do Grupo de Química Teórica (GQT) vinculado ao Instituto de Química da Universidade Federal do Rio Grande do Sul (UFRGS). As análises das trajetórias, a salvo programas de livre acesso referenciados no texto, foram realizadas com o emprego de *softwares* escritos pelo autor em linguagem Python 2.7, com auxílio das bibliotecas `numpy` 1.14.3 e `scipy` 1.1.0, e executados em computador pessoal. Os gráficos foram gerados por programa escrito em linguagem Python 2.7 com auxílio da biblioteca gráfica `matplotlib` 2.2.3. As imagens estruturais foram geradas com os softwares livres Visual Molecular Dynamics - VMD - 1.9.4(108), e renderizadas utilizando o programa Tachyon(109), e Blender 2.79b, renderizadas no próprio Blender.

### 6.1 Dinâmica molecular

As simulações de dinâmica molecular foram realizadas no cluster de computadores local do laboratório do Grupo de Química Teórica (GQT) empregando o pacote de programas para simulações por dinâmica molecular GROMACS 4.5.5 (72, 110, 111, 112, 113). Para a descrição das interações intramoleculares presentes no ácido ribonucleico (RNA) e deste com o solvente foi empregando o campo de forças AMBER parmbsc0.(114, 115) A obtenção dos parâmetros para a descrição das interações intramoleculares intramoleculares para os ligantes será descrita na próxima seção, 6.1.1.

Foram estudados três sistemas distintos: RNA, *ex nunc* sistema *aptâmero*, complexo entre RNA e estreptomicina, *ex nunc* sistema *complexo-estreptomicina*, e complexo entre RNA e bluensomicina, *ex nunc* sistema *complexo-bluensomicina*.

A orientação inicial do complexo RNA com a estreptomicina foi obtida diretamente do servidor *Protein Data Bank* (PDB) depositado em 2003 por Tereshko, Skripkin e Patel sob o código PDBid 1NTB.(68) As coordenadas atômicas referentes ao RNA foram obtidas realizando a remoção manual das coordenadas referentes a estreptomicina do complexo. Uma vez que a bluensomicina possui semelhança estrutural com a estreptomicina, as coordenadas atômicas do complexo-bluensomicina foram obtidas pela modificação de um grupo guanidino e um grupo aldeído por um grupo carbamido e um grupo hidroximetil, respectivamente, diretamente na estrutura 1NTB. Para os três sistemas, durante as simulações com solvente explícito, os íons  $Mg^{2+}$  foram mantidos na estrutura do aptâmero.

As estruturas iniciais dos três sistemas foram submetidas a processos de minimização de energia, no vácuo, empregando o método *steepest descent*. A cada passo do algoritmo, o movimento de cada átomo foi considerado como aceitável se, e somente se, o passo dado for menor ou igual a 0,01 nm. O critério de finalização do algoritmo estabelecido foi que o valor absoluto máximo do conjunto das componentes das forças fosse menor que o valor de tolerância de  $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$ . Após esta etapa, os sistemas foram preparados para as simulações com solvente explícito e implícito, descritas nas seções 6.1.2 e 6.1.3.

A próxima seção, seção 6.1.1, descreve o procedimento utilizado para a obtenção do campo de forças utilizado para descrever a dinâmica da estreptomicina e da bluensomicina condizentes com a filosofia do campo de forças AMBER utilizado para descrever o RNA.

### 6.1.1 Campo de forças para os ligantes

As estruturas correspondentes a estreptomicina, Figura 6A, e a bluensomicina, Figura 6B, foram otimizadas para o cálculo das cargas com o programa Gaussian 09, revisão A1(116), com a interface do programa GaussView, versão 5.(117) Devido a estas moléculas possuírem átomos de nitrogênio passíveis de serem protonados, a forma de protonação dos nitrogênios predominante nas condições experimentais foram asseguradas com os *plugins* pKa e Major microspecies da versão *on line* do programa Marvin, disponível em: [www.chemaxon.com/marvin/sketch/index.php](http://www.chemaxon.com/marvin/sketch/index.php) (Acessado em 02 de setembro de 2013).

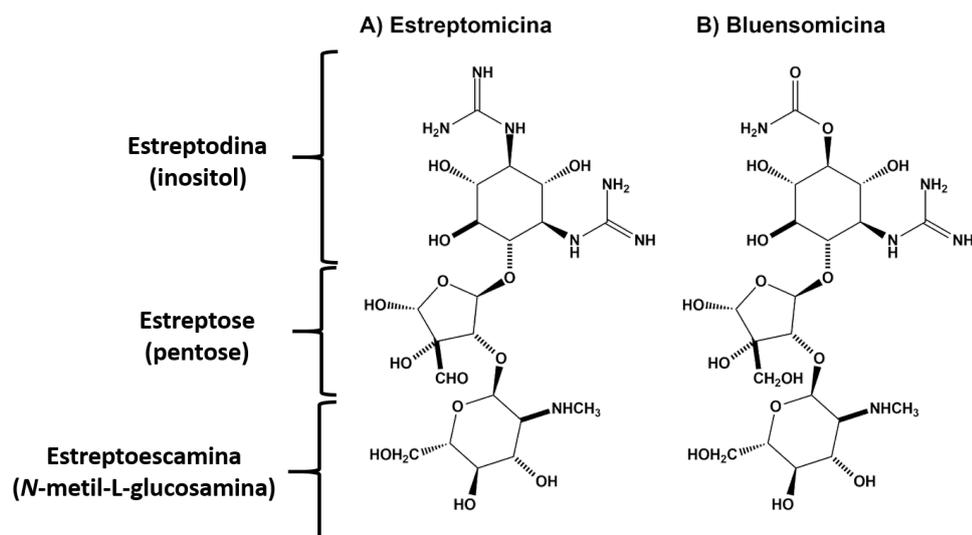


Figura 6 – Estrutura dos ligantes (A) Estreptomicina e (B) Bluensomicina.

As estruturas foram previamente otimizadas utilizando o método semi-empírico restrito, utilizando o hamiltoniano PM3(118, 119), e posteriormente as estruturas foram otimizadas pelo método de Hartree-Fock restrito, com o conjunto de base 6-31G(120, 121),

com auxílio de funções de polarização  $d$ . As cargas foram geradas pelo método de população de Merz-Singh-Kollman(122, 123), sinônimo das cargas calculadas pelo esquema de potencial eletrostático (ESP, do inglês *electrostatic potential*), com 10 camadas de pontos por átomo (*keyword* do **Gaussian**: IOp(6/41)=10) e tendo 17 pontos por unidade de área em cada camada (*keyword* do **Gaussian**: IOp(6/42)=17). Então, foi gerado pelo programa **Gaussian** (*keyword* do **Gaussian**: IOp(6/50)=1) um arquivo de extensão *gesp*, com as cargas ESP, que serviu de *input* para o programa **Antechamber** (124, 125) gerar as cargas calculadas pelo método AM1BCC, que foram utilizadas na topologia para a dinâmica molecular.

As topologias da estreptomicina e bluensomicina foram geradas com o auxílio do programa **Acpype** (129), utilizando parâmetros estruturais do *Generalized Amber Force Field*.(125)

### 6.1.2 Simulação do solvente explícito

Nesta seção está descrito o procedimento utilizado na construção dos sistemas com solvente explícito, bem como o procedimento e parâmetros utilizados para rodar estas simulações por dinâmica molecular, que posteriormente foram utilizadas como simulações de referência para calibrar os parâmetros necessários para executar as simulações com solvente implícito.

Após a etapa de minimização de energia, descrita na Seção 6.1, da estrutura do RNA ser realizada em todos os sistemas (aptâmero, complexo-estreptomicina e complexo-bluensomicina), foi então realizada a etapa de eletroneutralização destes sistemas. Uma vez que o RNA é um biopolímero aniônico, foram adicionados íons sódio(1+), de forma aleatória no vácuo, para neutralizar os sistemas. Os sistemas então foram submetidos a um novo processo de minimização de energia, empregando os mesmos algoritmos e parâmetros, e, em seguida, submetidos a 0,02 ns de dinâmica molecular com aplicação do método de restrição de posições da estrutura do ácido nucleico e ligantes, de forma a relaxar as fortes interações entre íons e íons-molécula. Para restringir as posições atômicas foi utilizada uma constante de força de 1000 kJ mol<sup>-1</sup> nm<sup>-2</sup> por átomo.

Os sistemas neutralizados e relaxados foram, então, solvatados com moléculas de água, modelo TIP3P(130) e concentração fisiológica de íons cloreto e sódio(1+) (0,154 mol L<sup>-1</sup> de NaCl), com aplicação de condições periódicas de contorno cúbicas. A seguir, os sistemas foram submetidos a um curto período, 0,02 ns, de dinâmica molecular com aplicação de restrição de posições nas estruturas dos solutos de forma a relaxar as fortes interações entre as moléculas de solvente e entre as moléculas de solvente e solutos. Estas etapas foram realizadas no *ensemble* NPT, empregando o termostato de Nosé-Hoover(76, 77), com constante de acoplamento  $\tau_T=0.1$  ps e temperatura alvo de 50 K, e o barostato de Parrinello-Rahman(131), constante de acoplamento  $\tau_P=2.0$  ps e pressão alvo de 1,0 atm.

A etapa posterior consistiu em quatro etapas de termalização, cada uma com 500 ps de extensão, em que os sistemas foram levados à temperatura final de 300 K. Cada degrau desta rampa foi simulado na temperatura de 50 K, 150 K, 250 K e 300 K, respectivamente.

Finalmente, foram rodadas três simulações de dinâmica molecular por sistema, com extensão de 100 ns, para equilíbrio do sistema no *ensemble* NPT, com temperatura constante de 310 K (37 °C), utilizando o termostato de Nosé-Hoover(76, 77), com pressão de 1,0 atm, mantida constante utilizando o barostato Parrinello-Rahman.(131) Se utilizou o algoritmo Smooth Particle-Mesh Ewald (SPME)(132, 133, 134), com um grid recíproco de  $108^3$  pontos com interpolação do tipo B-spline de quarta ordem, para o cálculo das interações eletrostáticas. A parte real do SPME foi calculada dentro de um raio de corte de 1,30 nm. As interações de van der Waals foram calculadas até um raio de corte de 1,30 nm.

Em todas as etapas de dinâmica molecular, a integração das equações do movimento foram realizadas com o algoritmo *leap-frog*(135), com passo de integração de 2 fs. A busca de vizinhos foi realizada a cada 50 passos, 100 fs.

### 6.1.3 Simulação com solvente implícito

Nesta seção está descrito o procedimento utilizado para rodar as simulações por dinâmica molecular dos sistemas com solvente implícito. Estas simulações foram rodadas com o pacote de programas GROMACS 4.5.5 (72, 110, 111, 112, 113) instalados no cluster local do laboratório do Grupo de Química Teórica (GQT). Esta versão foi utilizada devido a esta ser a versão nativa em nosso cluster e também devido ao fato que o método para solvente implícito foi retirado das versões mais atuais.

Na simulação por dinâmica molecular com solvente implícito dos três sistemas estudados, o solvente foi modelado usando dinâmica de Langevin onde a integração das equações do movimento foi realizada com o integrador Sd(136). As médias das interações instantâneas entre soluto e solvente foram aproximadas pelo modelo *Generalized Born solvent-accessible surface area* (GB/SA) com base no método de Hawkins-Cramer-Truhlar(137, 138) e parâmetros publicados por Tsui e Case(139). A máxima distância entre dois átomos para o cálculo do raio de Born foi estipulada em menor ou igual a 2,5 nm. A constante dielétrica do solvente foi mantida igual a 79,0. As interações não ligantes foram avaliadas dentro de um raio de corte de 15,0 nm. Como estes parâmetros já haviam sido testados por outros pesquisadores para um sistema de ácido desoxiribonucleico (DNA)(26, 27, 140), apenas o valor do coeficiente de fricção utilizado para a integração das equações de movimento,  $\gamma$ , foi calibrado. Este coeficiente foi, então, calibrado em relação ao comportamento do sistema nas simulações de dinâmica molecular com solvente explícito conforme a seção 6.1.2.

Após obter as estruturas iniciais das moléculas para os três sistemas, estas estruturas foram submetidas a processos de minimização de energia, empregando o método *steepest descent* enquanto seus átomos pesados foram mantidos (restringidos) em suas posições pela aplicação de um potencial harmônico com constante de  $1300,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ . Concluída a fase de relaxamento da estrutura inicial, foram geradas velocidades iniciais para cada átomo condizentes com a distribuição de Maxwell Boltzmann e temperatura alvo de 300 K, então, os sistemas foram submetidos a cinco simulações de 1,0 ns de extensão onde a restrição do movimento dos átomos foram progressivamente relaxados. As constantes utilizadas nestas cinco etapas foram  $420,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ ,  $310,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ ,  $200,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ ,  $100,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  e  $50,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , respectivamente.

#### 6.1.4 Dinâmica molecular dirigida

Nesta seção está descrito o procedimento para as simulações de não-equilíbrio, com solvente implícito, empregadas para o estudo estrutural e energético da desnaturação por força externa da molécula de RNA e complexos com estreptomicina e bluensomicina.

A estrutura inicial para as simulações de desnaturação foi obtida conforme a seção 6.1.3. Durante estas simulações o átomo O5' do resíduo C101 das sequências de RNA foi mantido fixo mediante a aplicação de uma restrição harmônica com constante de força de  $5000,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ . O átomo O5' do resíduo G1 foi conectado, por um potencial harmônico, a um átomo virtual que foi, então, puxado com uma velocidade constante na direção oposta ao átomo O5' do resíduo C101. Primeiramente, a constante de força e a velocidade de afastamento foram calibradas de forma que a curva de força-extensão produzida seja uma curva bem comportada e com discontinuidades claras. Após esta primeira verificação do comportamento das curvas em relação à constante de força e velocidade de afastamento, a velocidade de afastamento foi calibrada em busca da menor diferença entre a energia livre de desenovelamento experimental, obtido por Nick et al. (35), e a obtida pelo método descrito na próxima Seção 6.2.

## 6.2 Cálculo da Energia Livre

Nesta seção serão abordados o método e os parâmetros utilizados para o cálculo da energia livre de desenovelamento da molécula de RNA em função da distância de ponta-a-ponta. A distância de ponta-a-ponta é definida, deste momento em diante, como a distância entre o átomo O5' do resíduo G1 e o átomo O5' do resíduo C101.

O cálculo da energia livre para cada sistema foi realizado em duas etapas:

- 1) Cálculo do trabalho em função da distância de ponta-a-ponta;

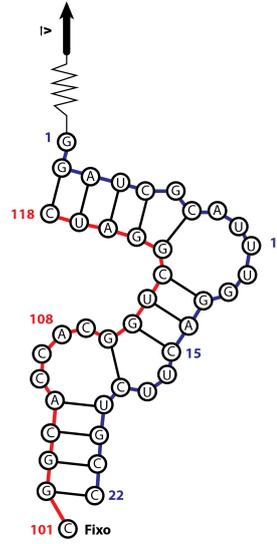


Figura 7 – Representação esquemática do sistema de referência empregado nas as simulações de dinâmica molecular dirigida

- 2) Cálculo da energia livre em função da distância de ponta-a-ponta utilizando a igualdade de Jarzynski.

Para cada uma replicata de cada um dos sistemas, o trabalho foi calculado conforme a aproximação de mola rígida.(38) Durante a simulação por dinâmica molecular dirigida, foi aplicado ao átomo móvel um potencial externo na forma  $V(\mathbf{q}; \lambda) = \frac{k}{2}(\xi(\mathbf{q}) - \lambda)^2$ , onde  $\xi(\mathbf{q})$  é a coordenada de reação e  $\lambda = \lambda(t)$  é o parâmetro externo de controle dependente do tempo (distância de ponta-a-ponta).

Desta forma, a função hamiltoniana do sistema é descrita por:

$$\tilde{H}(\mathbf{q}, \mathbf{p}; \lambda) = H(\mathbf{q}, \mathbf{p}) + V(\mathbf{q}; \lambda) \quad (6.1)$$

e desta forma, o trabalho externo  $W_{0 \rightarrow t}$  realizado no sistema entre o tempo zero e  $t$  é dado por:

$$W_{0 \rightarrow t} = \int_0^t \frac{\partial \lambda(t')}{\partial t'} \left[ \frac{\partial \tilde{H}(\mathbf{q}, \mathbf{p}; \lambda)}{\partial \lambda} \right]_{(\mathbf{q}, \mathbf{p}; \lambda(t)) = (\mathbf{q}_t, \mathbf{p}_t; \lambda(t))} dt' \quad (6.2)$$

que usando as equações para  $V(\mathbf{q}; \lambda)$ ,  $\tilde{H}(\mathbf{q}, \mathbf{p}; \lambda)$  e admitindo um avanço linear para o

parâmetro externo de controle  $\lambda(t) = \lambda_0 + vt$  nos dá o trabalho externo<sup>1</sup>:

$$W_{0 \rightarrow t} = -kv \int_0^t (\xi(\mathbf{q}) - \lambda(t')) dt' \quad (6.3)$$

De posse do trabalho em função da distância de ponta-a-ponta, a energia livre em função desta distância foi obtida pelo método de extrapolação linear da média de blocagem do estimador de não-equilíbrio padrão da igualdade de Jarzynski(21),  $e^{-\beta\Delta F} = \langle e^{-\beta W} \rangle$ , como proposto por Ytreberg e Zuckerman(39). A energia livre foi plotada como uma função de  $\chi = 1/n^{0.5}$ , onde  $n$  é o tamanho dos blocos, e extrapolada para  $\chi = 0$  utilizando a ajuste por mínimos quadrados. Os blocos foram compostos por 500 valores de trabalho escolhidos aleatoriamente e sem substituição.

---

<sup>1</sup> O Apêndice A decreve passo a passo a obtenção da Equação 6.3 a partir da Equação 6.2.

## Parte III

### Resultados e discussão

## 7 Simulações com solvente explícito

Nesta seção serão apresentadas as simulações por dinâmica molecular em equilíbrio com solvente explícito para os três sistemas: *aptâmero*, *complexo-estreptomicina* e *complexo-bluensomicina*. Será abordado a estabilidade estrutural para os três sistemas. O oligonucleotídeo, aptâmero, aqui estudado possui grande afinidade à molécula de estreptomicina(36, 37), enquanto que possui uma baixa afinidade à molécula de bluensomicina.(37) Desta forma, devido à grande semelhança estrutural entre estreptomicina e bluensomicina e à sua diferença de afinidade (a estreptomicina possui uma afinidade de ligação com o aptâmero quatro ordens de grandeza maior que a bluensomicina(34)), o estudo computacional em equilíbrio dos complexos entre o aptâmero e estreptomicina e bluensomicina é importante como ponto de partida para a aplicação da validação do método para o cálculo de energia livre proposto nesta tese.

Na Figura 8 se encontram os gráficos da raiz quadrada do deslocamento quadrático médio (RMSD) da estrutura do RNA, em relação a estrutura inicial da etapa de produção, para os três sistemas estudados. Analisando os valores de RMSD se pode ver que este, para todos os sistemas, se encontra abaixo de 0,6 nm. Levando em consideração que a estrutura do RNA compreende 1277 átomos, podemos inferir que os sistemas se encontram, dentro do período simulado, estruturalmente equilibrados. É interessante notar que, para o sistema *aptâmero* o comportamento da RMSD mostra que a ausência de uma molécula ligada à estrutura torna esta mais flexível, de forma que a estrutura transita entre dois estados de equilíbrio. Este fato fica claro tanto no valor do RMSD em função do tempo como na distribuição bimodal deste, apresentada nos histogramas para o sistema aptâmero. Esta menor flexibilidade da estrutura, quando ligada a molécula de estreptomicina está de acordo com o observado experimentalmente, onde a presença da estreptomicina amarra o *backbone* do aptâmero por quatro diferentes partes do seu sistema de anéis por ligações salinas ou de hidrogênio.(34, 141) Também, é interessante notar que a presença da molécula de bluensomicina ligada ao sítio de ligação não reduz a flexibilidade do aptâmero. Este comportamento está claro no padrão de RMSD para o sistema *complexo-bluensomicina*, Figura 8, onde se pode ver que as distribuições do RMSD ao longo das trajetórias possuem médias diferentes e, portanto, demonstram que este sistema também transita entre dois estados de equilíbrio.

É interessante ainda notar na Figura 8 que, para o sistema onde o RNA está ligado à molécula de estreptomicina, o RMSD reduz levemente sua intensidade com o avanço do tempo de simulação. Entretanto, quando o RNA está ligado à molécula de bluensomicina, o RMSD aumenta levemente sua intensidade conforme a simulação evolui temporalmente. Este comportamento pode ser melhor observado nos histogramas da

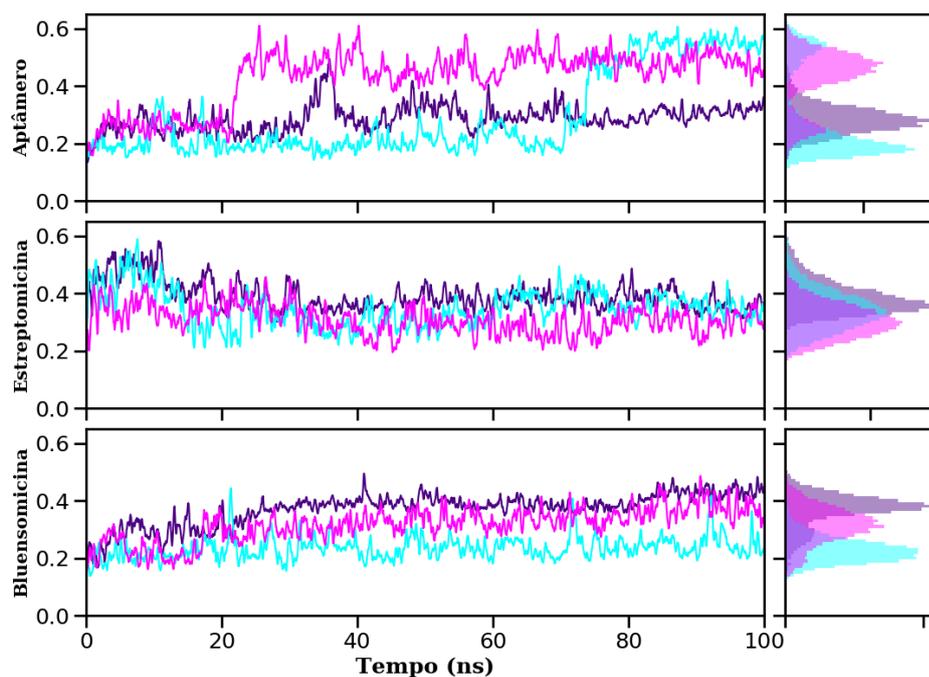


Figura 8 – Raiz quadrada do deslocamento quadrático médio da estrutura do RNA para as simulações com solvente explícito para os três sistemas estudados. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3.

Figura 8 para o sistemas *complexo-estreptomicina*, que possui uma cauda à esquerda, e para o sistema *complexo-bluensomicina*, que possui uma cauda à direita. Este fato é uma indicação adicional de que a molécula de estreptomicina, ao se ligar à molécula de RNA, confere uma menor flexibilidade ao RNA. Em contrapartida, a molécula de bluensomicina está aumentando a flexibilidade do RNA ao se ligar a esta, embora a quantificação deste efeito em simulações clássicas irrestritas demandaria uma janela temporal de várias centenas de nanossegundos.

A Figura 9 mostra o número total de pares de base na estrutura de RNA para os três sistemas estudados. Como esperado, a introdução de uma molécula no sítio de ligação, entre as duas fitas de RNA, leva a diminuição do número de pares de base formados, conforme é visto se comparando o sistema *aptâmero* com os sistemas *complexo-estreptomicina* e *complexo-bluensomicina*. Comparando os histogramas da Figura 9, podemos notar também que a presença de uma molécula, tanto a estreptomicina como a bluensomicina, ligada ao sítio de ligação leva a uma menor flutuação no número de pares de base na estrutura do RNA. Este fator atenuador na flutuação da interação entre bases é mais evidente quando comparamos os números de multipletos, interações por ligações de hidrogênio entre três ou mais bases coplanares(48), ao longo do tempo de simulação conforme a Figura 10. Pode-se ver na Figura 10 que a presença de uma molécula ligada ao sítio de ligação estabiliza o número de multipletos na estrutura de RNA. Este fato é marcadamente pronunciado pelos histogramas referentes ao número de multipletos nos sistemas, já que há uma considerável

diminuição da largura das curvas de distribuição. Como o número de multipletos é um indicativo da presença de estrutura terciária(48), se pode inferir que a presença da estreptomicina e da bluensomicina ligadas ao sítio de ligação, uma vez que estas diminuem a flutuação, estabiliza a estrutura terciária do RNA.

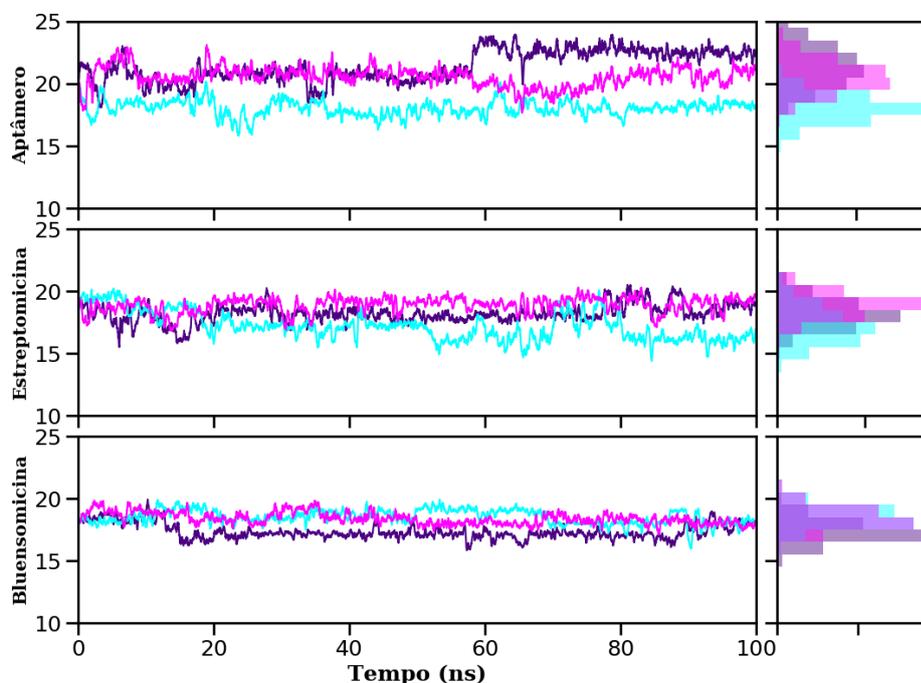


Figura 9 – Número de pares de base formados pelo RNA para as simulações com solvente explícito para os três sistemas estudados. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3.

Segundo a estrutura cristalográfica depositada por Tereshko, Skripkin e Patel(34) para o sistema *complexo-estreptomicina* (PDBid: 1NTB (34)), este complexo possui dois tripletos, interações por ligações de hidrogênio entre três bases coplanares, entre os resíduos G6•G13•C113 e U16•C18•G110, Figura 11, de forma que os resultados da Figura 10 estão de acordo com os experimentos. Entretanto, a ausência de uma molécula ligada à estrutura possibilita a formação de mais dois novos tripletos: U19•A105•C106 e A8•A14•U12, conforme a Figura 12. A ausência de um ligante, e conseqüentemente a maior flexibilidade da estrutura, possibilita a alteração do triplete U16•C18•G110 para C15•C18•G110.

O número de ligações de hidrogênio entre as duas fitas, Figura 13, para as três simulações do sistema *aptâmero* se mantém constante durante o período simulado. As distribuições para as três simulações do sistema *aptâmero* possuem grande sobreposição. Em contrapartida, para o sistema *complexo-estreptomicina* as distribuições não se sobrepõem completamente. De fato, se pode ver nas evoluções temporais que cada uma das três simulações para esse sistema possuem, em média, um número diferente de ligações de hidrogênio entre as duas fitas. Sendo que em duas das simulações, o número de ligações

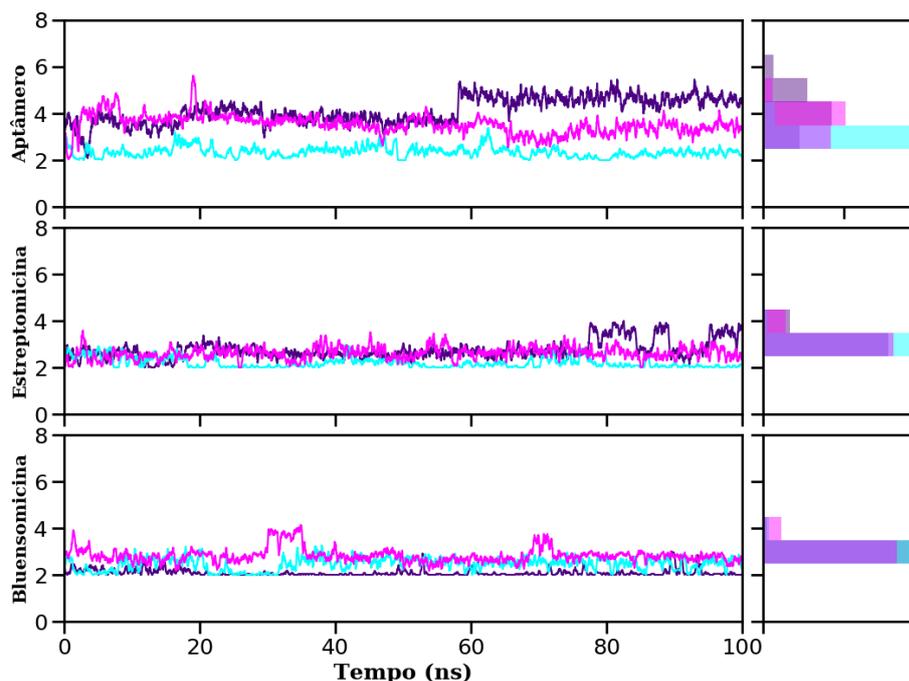


Figura 10 – Número de multipletos formados para as simulações com solvente explícito para os três sistemas estudados. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3.

de hidrogênio entre as fitas foi maior e menor, respectivamente, que para o sistema *aptâmero*. Esta diferença entre o número de ligações de hidrogênio para os sistemas *aptâmero* e *complexo-estreptomicina* é compensada pelas ligações de hidrogênio que as fitas fazem com a molécula de estreptomicina. Conforme a Figura 15, podemos ver que embora a Figura 13 mostre que o número de ligações de hidrogênio entre as fitas diminua e a Figura 14 mostre que o número de ligações de hidrogênio entre a fita 1 e a molécula de estreptomicina se mantém constante, a Figura 15 mostra que, para esta mesma simulação, o número de ligações de hidrogênio entre a fita 2 e a estreptomicina aumenta, em média, de dois para nove.

O número de ligações de hidrogênio entre as fitas do RNA para o sistema *complexo-bluensomicina*, Figura 13, aumentam em relação ao sistema *aptâmero*, indicando que a presença da bluensomicina promove uma melhor interação direta entre as fitas, em contra partida a estreptomicina promove uma melhor interação mediada entre ambas as fitas. Pode-se ver que a interação por ligação de hidrogênio da bluensomicina com a fita 2, Figura 15, se mantém constante e em valor menor do que a estreptomicina durante o período simulado. Este fato está de acordo com o obtido experimentalmente por Tereshko, Skripkin e Patel(34), uma vez que o fato do grupo para-guanidino (doador de ligação de hidrogênio) na estreptomicina é substituído pelo grupo para-carbamido (aceptor de ligação de hidrogênio) na bluensomicina, faz com que a bluensomicina não possua uma forte interação por ligação de hidrogênio como as guaninas nas posições

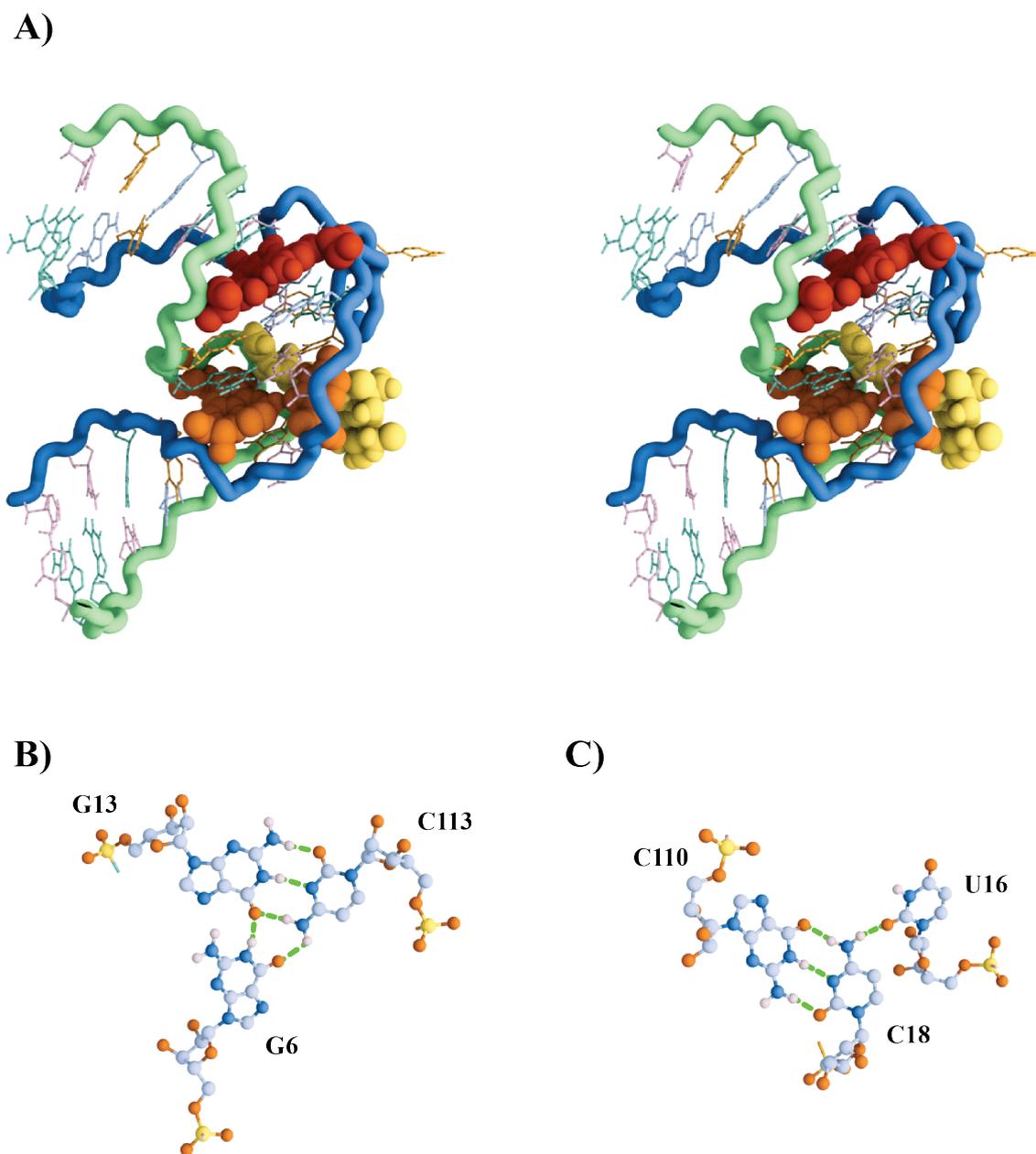


Figura 11 – A) Modelo estereo da estrutura do sistema *complexo-estreptomicina* com os tripletos G6•G13•C113 (laranja escuro) e U16•C18•G110 (laranja) em esferas de van der Waals. Estreptomicina (amarelo) em esferas de van der Waals. B) Tripletos G6•G13•C113. C) Tripletos U16•C18•G110.

110 e 111 da fita 2. De fato, a Tabela 2 mostra que a bluensomicina realiza ligações de hidrogênio entre 25 % à 99 % e 11 % à 62 % das simulações com as guaninas 110 e 111, respectivamente. Em contrapartida, a estreptomicina realiza ligações de hidrogênio mais de 96 % e entre 26 % à 95 % das simulações (conforme Tabela 1) com as guaninas 110 e 111, respectivamente, de forma que é claro que a bluensomicina interage menos com a Fita 2 do que a estreptocimicina o faz.

Tabela 1 – Ligações de hidrogênio entre a estreptomicina e o aptâmero.

Fita	Resíduo	Porcentagem de Ligações de hidrogênio											
		Total*			$n_{HB=1^*}$			$n_{HB=2^*}$			$n_{HB=3^*}$		
1	G6	0,0	10,5	41,1	0,0	10,5	40,8	0,0	0,0	0,3	0,0	0,0	0,0
	A8	5,0	59,8	29,4	5,0	58,9	28,9	0,1	0,9	0,5	0,0	0,0	0,0
	U11	69,9	96,1	84,3	50,9	86,8	64,6	16,2	9,1	16,9	2,4	0,1	2,7
	G12	98,2	96,7	97,7	21,9	36,9	27,5	53,4	58,9	65,3	22,2	0,9	4,9
	G13	5,9	0,0	0,0	3,9	0,0	0,0	2,1	0,0	0,0	0,0	0,0	0,0
	A14	0,2	0,5	0,5	0,2	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0
	C15	29,8	0,9	0,1	29,8	0,9	0,1	0,0	0,0	0,0	0,0	0,0	0,0
	U16	96,3	92,7	94,6	84,1	91,6	94,3	11,5	1,2	0,3	0,7	0,0	0,0
	U17	40,4	5,5	77,0	33,7	4,5	76,5	6,1	1,0	0,5	0,6	0,0	0,0
	C18	0,2	0,7	0,4	0,2	0,7	0,4	0,0	0,0	0,0	0,0	0,0	0,0
2	C107	11,2	85,8	65,5	8,8	8,3	60,2	2,3	38,6	5,2	0,1	35,8	0,1
	A108	26,6	97,0	64,0	14,8	62,7	42,4	8,5	34,0	18,6	3,3	0,2	2,8
	C109	2,3	85,6	0,8	2,1	77,9	0,6	0,2	7,5	0,2	0,0	0,2	0,0
	G110	97,2	98,5	96,5	20,5	13,2	43,9	31,7	36,8	34,8	42,7	41,0	17,3
	G111	26,9	95,0	84,9	26,3	33,2	69,0	0,6	38,1	15,9	0,0	23,7	0,0
	U112	1,0	2,8	2,6	1,0	2,8	2,6	0,0	0,0	0,0	0,0	0,0	0,0

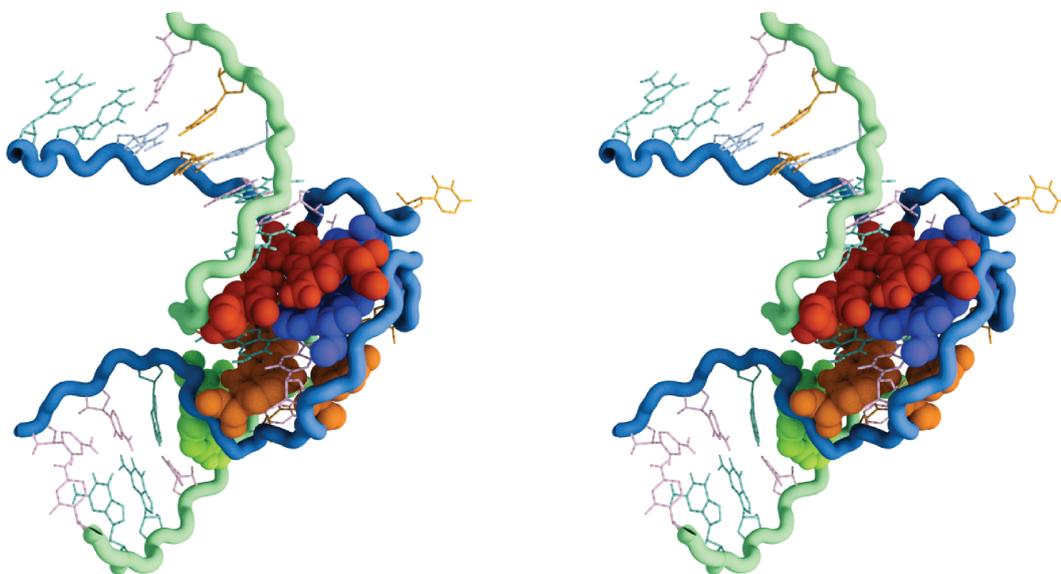
\* Cada subcoluna corresponde ao valores correspondente às 3 réplicas.

Tabela 2 – Ligações de hidrogênio entre a bluensomicina e o aptâmero.

Fita	Resíduo	Porcentagem de Ligações de hidrogênio											
		Total*			$n_{HB=1^*}$			$n_{HB=2^*}$			$n_{HB=3^*}$		
1	G6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	A8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	U11	49,9	61,2	82,8	43,0	44,3	53,9	6,5	13,3	25,9	0,3	3,3	2,8
	G12	56,2	94,9	90,2	45,1	32,2	35,4	10,8	42,5	45,8	0,3	18,7	8,4
	G13	0,4	0,2	0,0	0,4	0,1	0,0	0,0	0,1	0,0	0,0	0,0	0,0
	A14	0,9	2,4	0,0	0,6	2,3	0,0	0,4	0,2	0,0	0,0	0,0	0,0
	C15	13,0	61,9	1,6	12,9	61,3	1,6	0,1	0,6	0,0	0,0	0,0	0,0
	U16	68,7	92,0	52,4	51,7	65,2	39,5	15,8	23,4	11,9	1,1	3,3	1,0
	U17	49,4	25,5	65,3	47,9	23,6	64,1	1,4	1,8	1,2	0,1	0,1	0,0
	C18	0,3	0,3	0,0	0,3	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0
2	C107	49,3	2,7	34,4	33,6	2,5	31,7	14,3	0,2	2,6	1,3	0,0	0,1
	A108	51,7	3,0	44,2	34,7	2,5	43,2	9,6	0,4	1,0	4,8	0,0	0,0
	C109	0,4	8,3	0,5	0,4	7,9	0,5	0,0	0,4	0,0	0,0	0,0	0,0
	G110	53,8	99,0	25,5	51,0	35,2	17,0	2,9	60,7	8,4	0,0	3,2	0,1
	G111	20,5	61,9	11,5	19,9	60,3	11,5	0,7	1,6	0,0	0,0	0,0	0,0
	U112	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

\* Cada subcoluna corresponde ao valores correspondente às 3 réplicas.

A)



B)

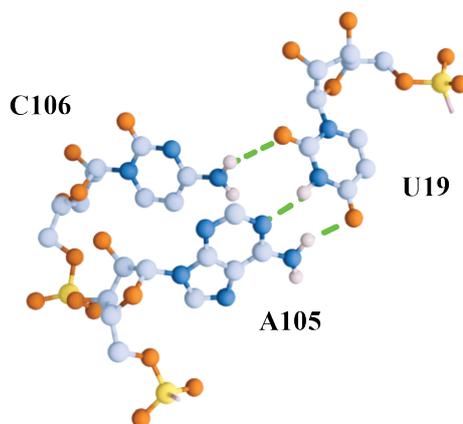


Figura 12 – A) Modelo estereo da estrutura do sistema *aptâmero* com os tripletos G6•G13•C113 (laranja escuro) e U16•C18•G110 (laranja), U19•A105•C106 (verde) e A8•A14•U12 (azul) em esferas de van der Waals. B) Tripletto U19•A105•C106.

A interação da bluensomicina por ligações de hidrogênio com a fita 1, Figura 14, mostra uma grande flutuação. Uma vez que a molécula de bluensomicina é ancorada ao aptâmero por ligação de hidrogênio através de seu grupo guanidino na posição orto do anel, este acaba interagindo com diversos aceptores de ligação de hidrogênio, principalmente com os nucleotídeos U11, G12 e U17, pelo fato da molécula de bluensomicina se encontrar mais solta no sítio de ligação, de acordo com o que foi encontrado experimentalmente por Tereshko, Skripkin e Patel(34).

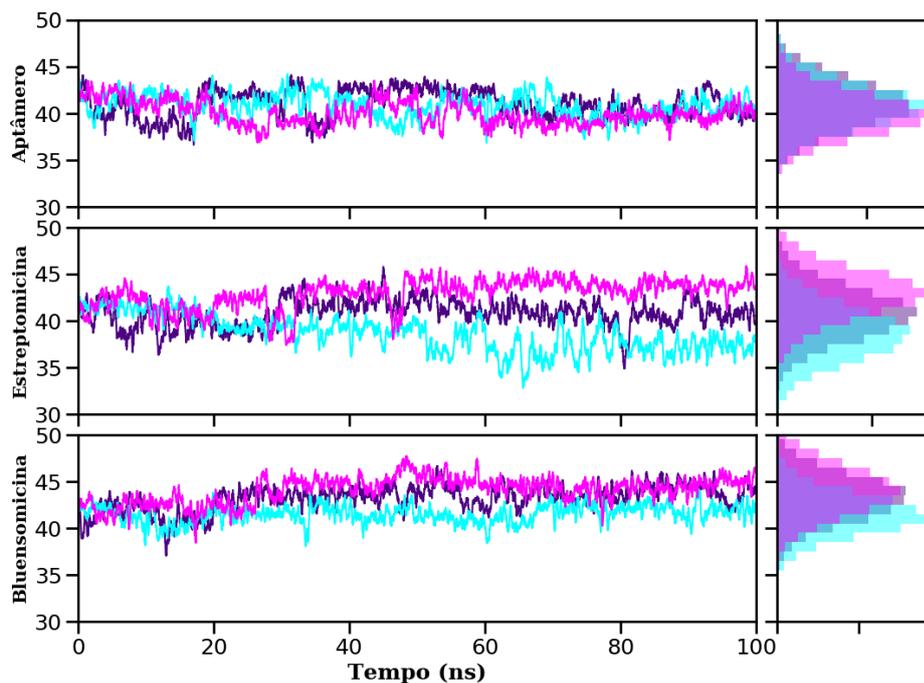


Figura 13 – Número de ligações de hidrogênio entre as fitas de RNA para as simulações com solvente explícito para os três sistemas estudados. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3.

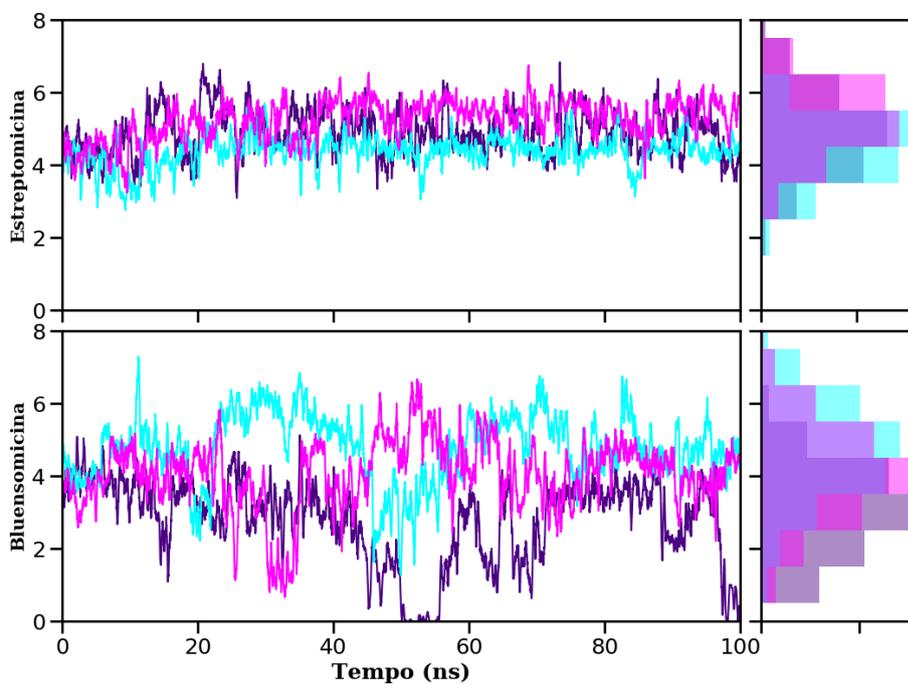


Figura 14 – Número de ligações de hidrogênio entre a fita 1 de RNA e ligante para as simulações com solvente explícito. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3.

Comparando os valores referentes ao número de ligações de hidrogênio entre ligante e aptâmero para os sistemas *complexo-estreptomicina*, Tabela 1, e *complexo-bluensomicina*,

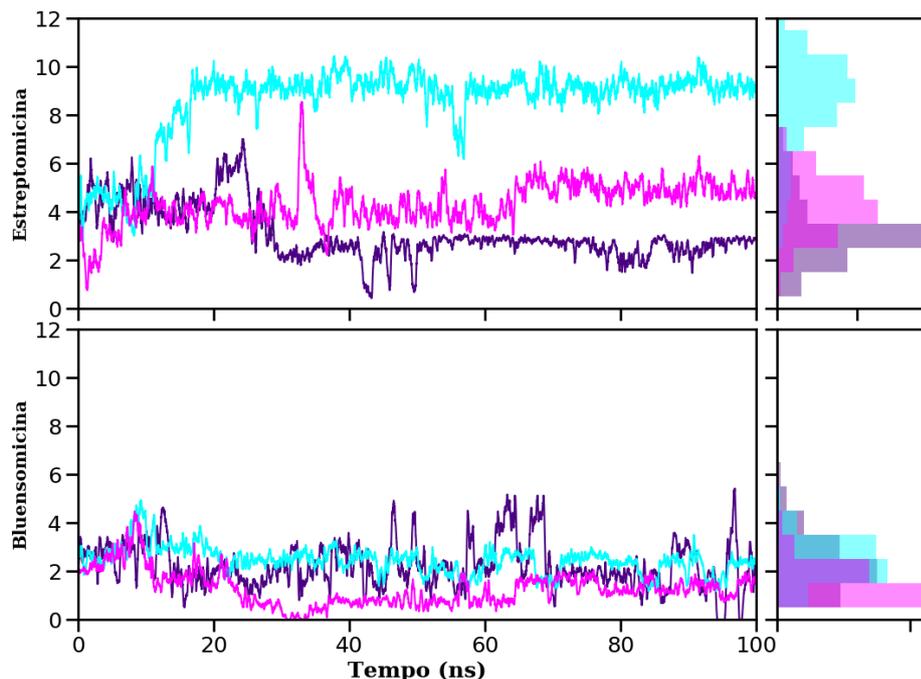


Figura 15 – Número de ligações de hidrogênio entre a fita 2 de RNA e ligante para as simulações com solvente explícito. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3.

Tabela 2, podemos ver que, a maioria das ligações de hidrogênio realizadas são com os resíduos U11, G12, U16 e U17 da fita 1 e C107, A108, G110 e G111 da fita 2. Desta interações, podemos ver que os resíduos U11 e G12 da fita 1 tendem a formar mais de uma ligação de hidrogênio com ambos os ligantes ao longo das simulações, principalmente o resíduo G12 que realiza duas ou três ligações de hidrogênio simultâneas com ambos os ligantes. Na fita 2, os resíduos C107, A108, G110 e G111 tendem a formar mais de uma ligação de hidrogênio com ambos os ligantes, sendo que a guanina na posição 110 realiza significativamente ao longo da simulação duas ligações de hidrogênio com ambos os ligantes.

Além das interações constantes com os resíduos U11, G12, U16, U17, C107, A108, G110 e G111, as Tabelas 1 e 2 mostram que a estreptomicina também realiza ligações de hidrogênio com os resíduos G6 e A8 da fita 1 e U112 da fita 2 ao longo da simulação, enquanto que a molécula de bluensomicina não interage com estes resíduos por este tipo de ligação. Essas ligações de hidrogênio adicionais contribuem para a seletividade do aptâmero pela estreptomicina frente a bluensomicina, conforme encontrado na literatura.(34, 36, 37, 141) Como pode ser visto na Figura 16, este conjunto de resíduos estão dispostos em torno da molécula de estreptomicina, onde esta forma ligações de hidrogênio com ambas as fitas do aptâmero, prendendo-as e tornando a estrutura menos flexível e menos propensa ao desenovelamento.

Neste capítulo foi mostrado que a presença das moléculas de estreptomicina e blu-

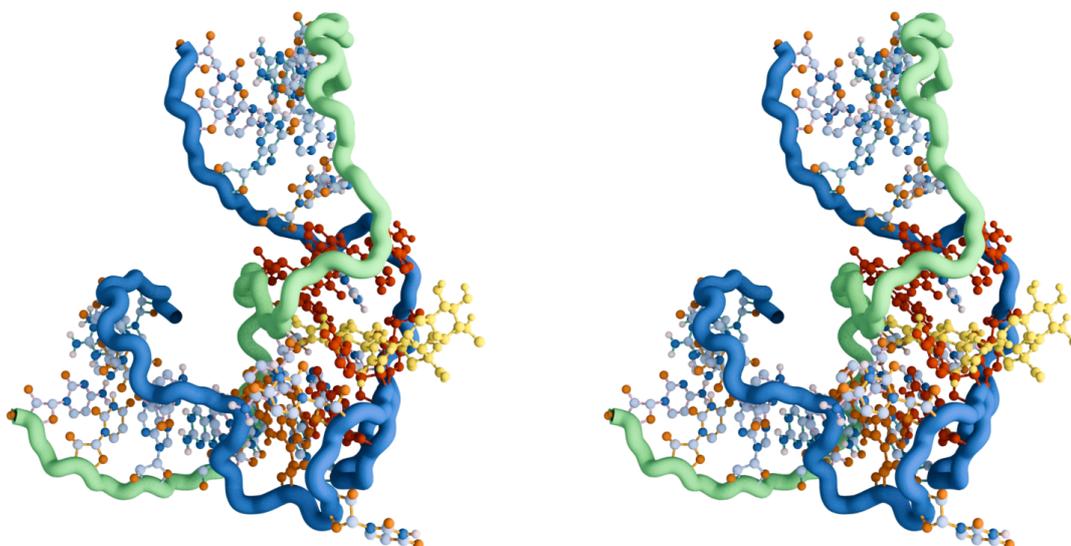


Figura 16 – Modelo estereo da estrutura do sistema *complexo-estreptomicina* com resíduos que fazem ligação de hidrogênio com o ligante em destaque. Legenda: Fita 1 em azul, fita 2 em verde. Átomos representados no modelo bola e vareta. Oxigênio em laranja, Nitrogênio em azul real, carbono em azul claro e hidrogênio em branco. Estreptomicina representada em amarelo. Resíduos U11, G12, U16, U17, C107, A108, G110 e G111 em vermelho e resíduos G6, A8, G13, A14, C15, C109 e U112.

ensomicina ligadas ao sítio de ligação do aptâmero leva a uma diminuição da flexibilidade deste, conforme visto pelos valores de RMSD. Por esta mesma análise se pode constatar que o sistema *complexo-streptomicina* possui um estado conformacional, ou seja, em média o valor de RMSD convergiu para uma mesma média, e que o sistema *complexo-bluensomicina* parece ter dois estados conformacionais distintos. Pode-se inferir, também, o aumento da estabilidade da estrutura, devido à presença dos ligantes, com a diminuição da flutuação do número de pares de base e de multipletos estabelecidos ao longo das simulações. Por fim, a análise do número de ligações de hidrogênio formados revelou que, conforme o que já se tinha conhecimento experimentalmente, a estreptomicina possui uma maior interação com o aptâmero do que a bluensomicina, razão esta que é apontada como a causa da seletividade da estreptomicina frente a bluensomicina.

## 8 Coeficiente de fricção

De forma a calibrarmos o coeficiente de fricção  $\gamma$  utilizado para a integração das equações de movimento nas simulações com solvente implícito, foram consideradas como referência as três simulações com solvente explícito dos sistemas *aptâmero* e *complexo-estreptomicina*.

Nas Figuras 17 e 18 estão plotados a sobreposição das curvas suavizada da raiz quadrada do deslocamento quadrático médio (RMSD) em função do tempo, para cada réplica, para os sistemas *aptâmero* e *complexo-estreptomicina*, respectivamente. Pode-se ver que o sistema onde não há a presença da estreptomicina, ou seja o sistema *aptâmero*, apresenta maior desvio em relação a estrutura inicial e que este desvio ocorre principalmente devido a um maior rearranjo da fita 1. Esta mobilidade da fita 1 decorre do fato desta se rearranjar devido a remoção da estreptomicina da estrutura cristalográfica para montar o sistema *aptâmero*, conforme foi descrito na seção 6.1.

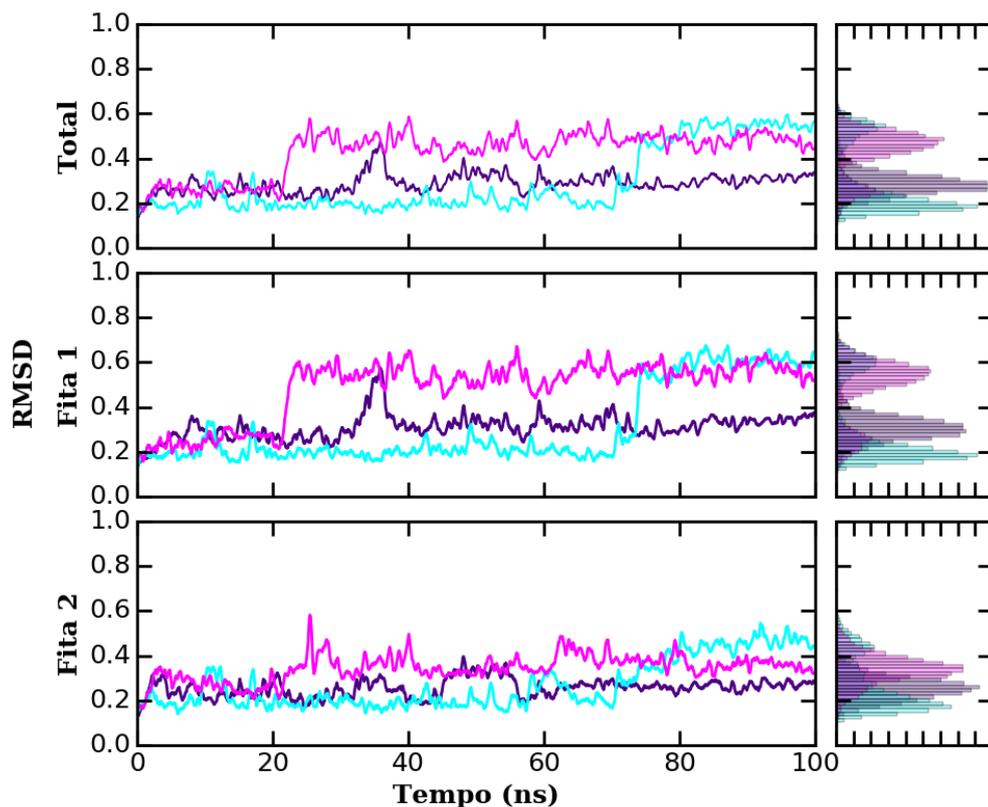


Figura 17 – Raiz quadrada do deslocamento quadrático médio da estrutura do RNA para as simulações com solvente explícito do sistema *aptâmero*. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3. Linha sólida representa a média sobre uma janela de 50 ps

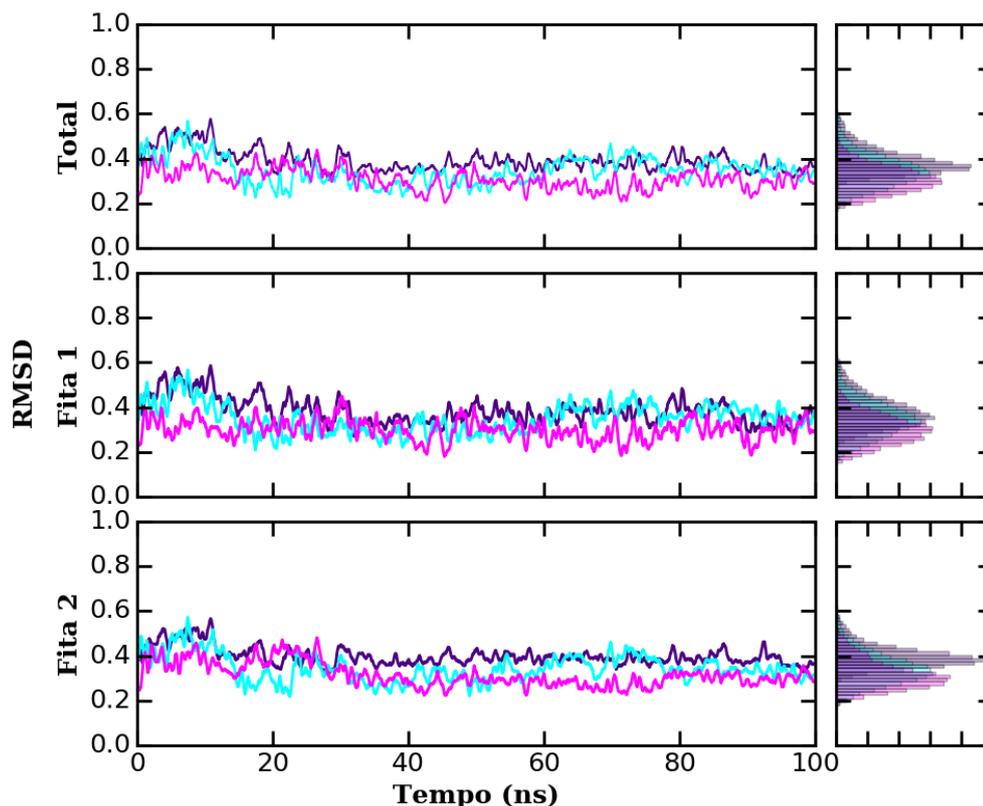


Figura 18 – Raiz quadrada do deslocamento quadrático médio da estrutura do RNA para as simulações com solvente explícito do sistema *complexo-estreptomicina*. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3. Linha sólida representa a média sobre uma janela de 50 ps

Os parâmetros estruturais número de ligações de hidrogênio total, número de ligações de hidrogênio entre fitas, o número de pares de base e número de multipletos foram monitorados para os sistemas *aptâmero* (Figura 19) e *complexo-estreptomicina* (Figura 20). Pode-se ver que em ambos os sistemas os parâmetros estão se comportando com distribuição normal, gaussiana, e que sua flutuação é pequena comparada à média. É possível notar a formação de mais um multipletos para a primeira réplica do sistema *aptâmero* após aproximadamente 60 ns de simulação. Esta formação provavelmente está ligada ao rearranjo da fita 1 devido a exclusão da estreptomicina do modelo durante a construção do sistema.

Foram realizadas cinco simulações com solvente implícito para os sistemas *aptâmero* e *complexo-estreptomicina*, respectivamente, com os valores de coeficiente de fricção  $\gamma$  de  $2,0 \text{ ps}^{-1}$  e  $3,0 \text{ ps}^{-1}$ . Estes valores foram os utilizados por Naserian-Nik, Tahani e Karttunen(26) e Naserian-Nik, Tahani e Karttunen(27) em seus respectivos trabalhos. Nas Figuras 21-24 estão a sobreposição das curvas suavizadas da raiz quadrada do deslocamento quadrático médio em função do tempo para as cinco réplicas para os valores de  $\gamma$  de  $2,0 \text{ ps}^{-1}$  e  $3,0 \text{ ps}^{-1}$  para ambos os sistemas. Embora ambos os sistemas com

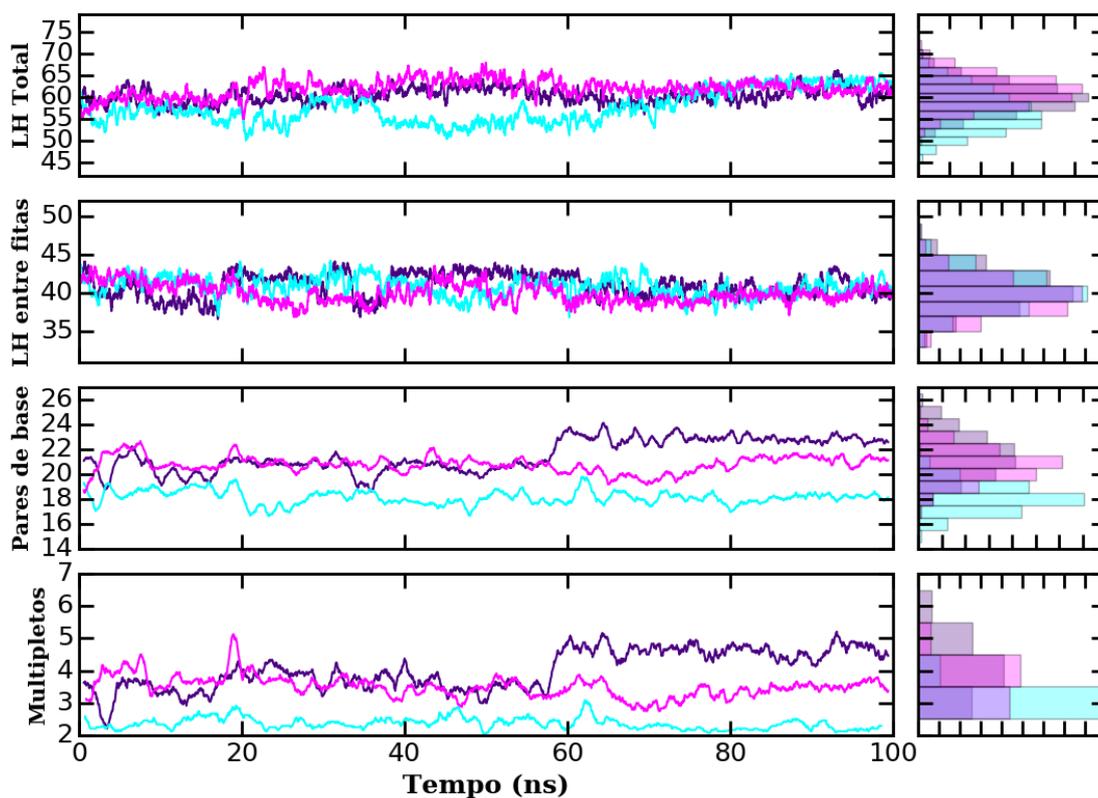


Figura 19 – Parâmetros estruturais para as simulações com solvente explícito do sistema *aptâmero*. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3.

solvente implícito se mostrem estáveis em relação aos valores de RMSD, podemos ver que para o valor de  $\gamma$  de  $3,0 \text{ ps}^{-1}$ , Figuras 22 e 24, ambos os sistemas mostram que os valores de RMSD se apresentam distribuídos de forma mais gaussiana que as respectivas distribuições com solvente explícito. Em contrapartida, os gráficos de RMSD para as simulações com valores do coeficiente de fricção igual a  $2,0 \text{ ps}^{-1}$ , Figuras 21 e 23, mostram uma flexibilidade mais coerente com aqueles encontrados para os sistemas com solvente explícito.

Os parâmetros estruturais analisados para as simulações com solvente implícito cujo coeficiente de fricção igual a  $2,0 \text{ ps}^{-1}$  estão representados nas Figuras 25 e 26 para os sistemas *aptâmero* e *complexo-estreptomicina*, respectivamente. Se pode ver que, para ambos os sistemas, os parâmetros estruturais apresentam uma distribuição muito próxima a gaussiana, indicando que os sistemas estão estáveis, sem que aconteça desnaturação da estrutura da molécula de RNA.

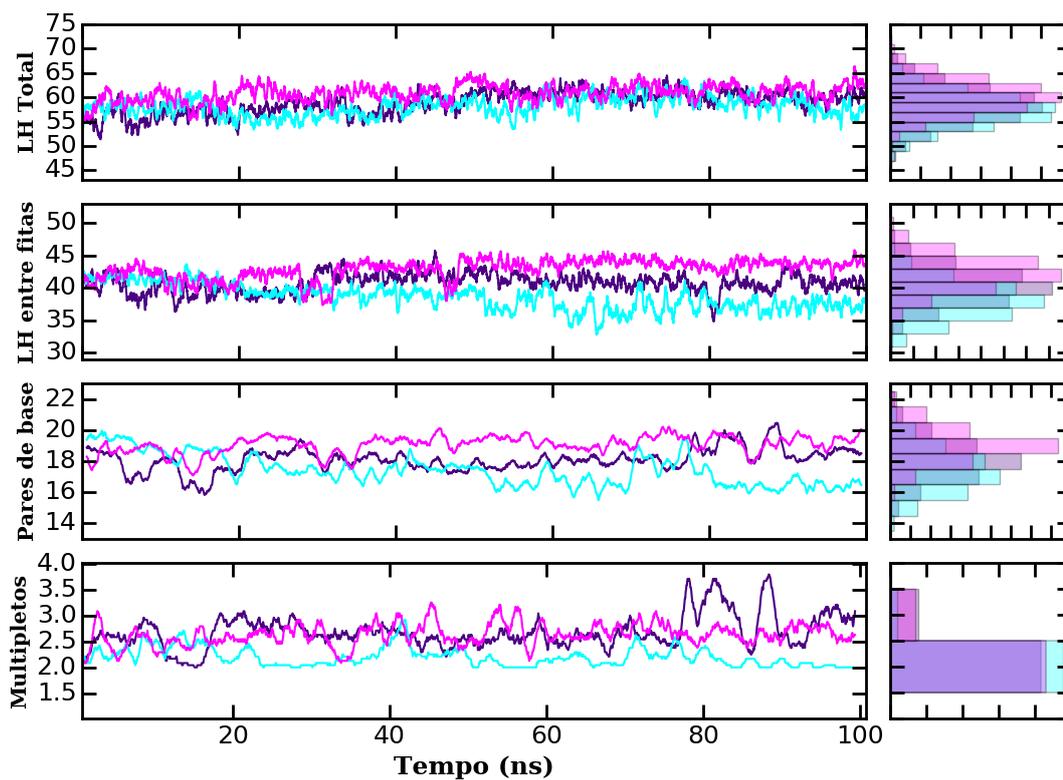


Figura 20 – Parâmetros estruturais para as simulações com solvente explícito do sistema *complexo-streptomicina*. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3.

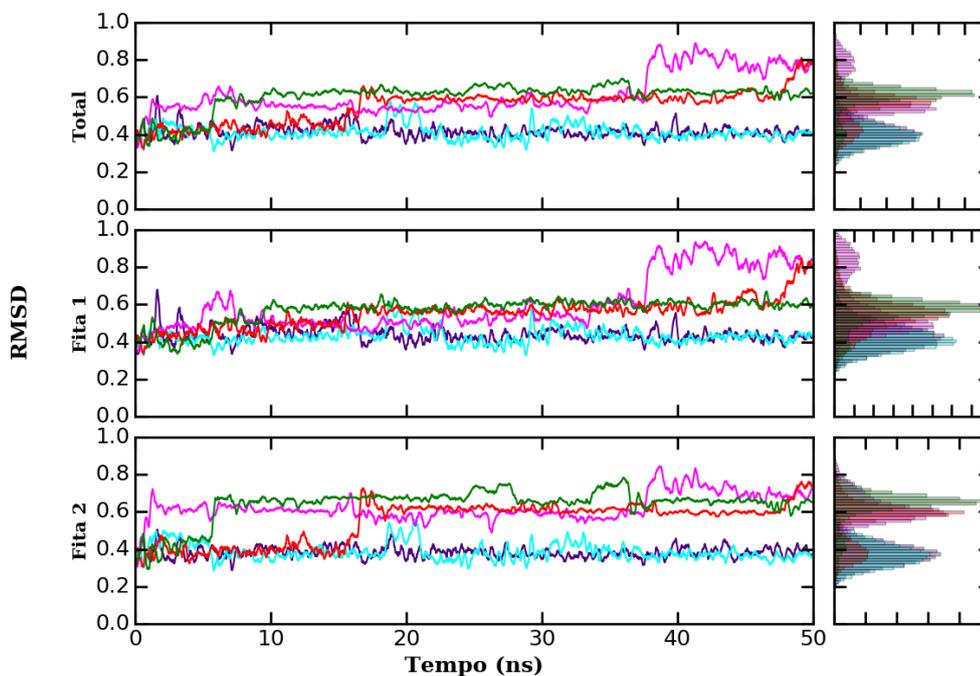


Figura 21 – Raiz quadrada do deslocamento quadrático médio da estrutura do RNA para as simulações com solvente implícito e  $\gamma = 2,0 \text{ ps}^{-1}$  do sistema *aptâmero*. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3; Fúcsia: Réplica 4; Verde: Réplica 5. Linha sólida representa a média sobre um janela de 50 ps

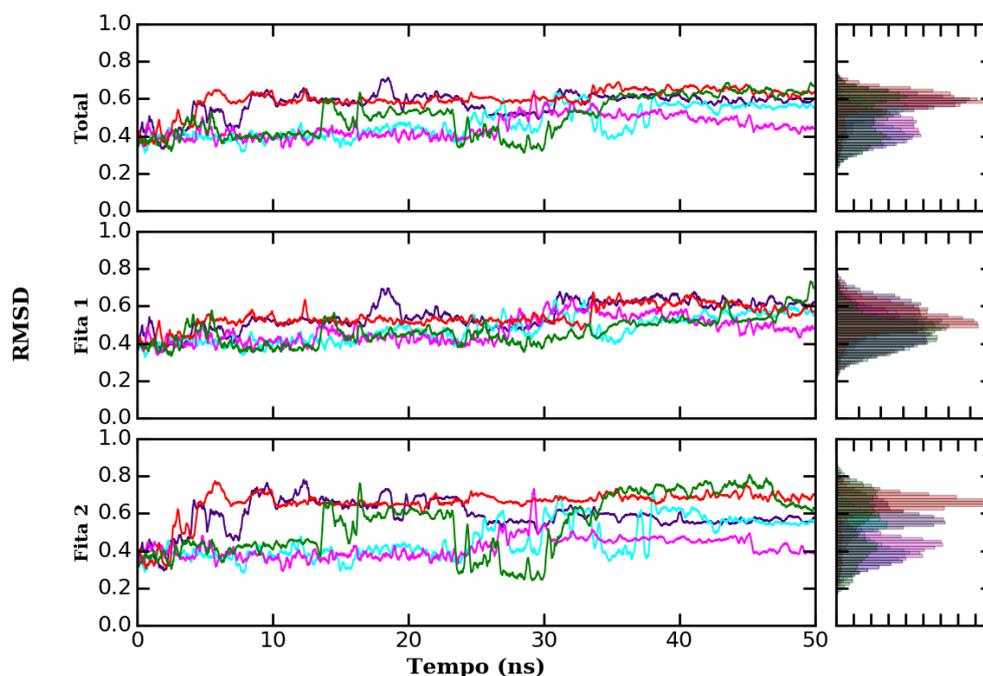


Figura 22 – Raiz quadrada do deslocamento quadrático médio da estrutura do RNA para as simulações com solvente implícito e  $\gamma = 3,0 \text{ ps}^{-1}$  do sistema *aptâmero*. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3; Fúcsia: Réplica 4; Verde: Réplica 5. Linha sólida representa a média sobre uma janela de 50 ps

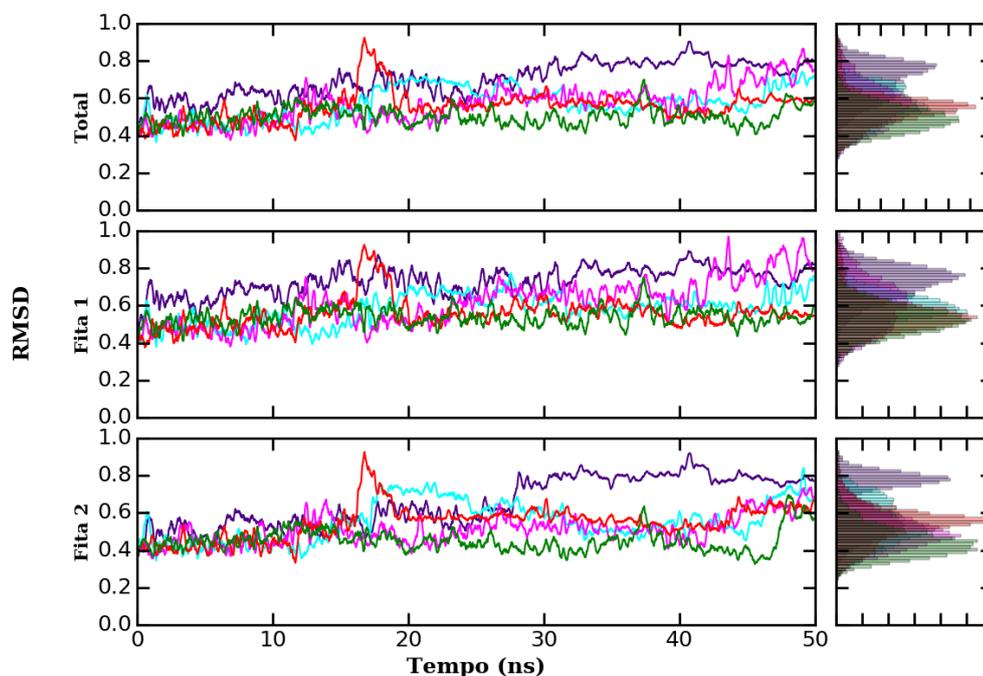


Figura 23 – Raiz quadrada do deslocamento quadrático médio da estrutura do RNA para as simulações com solvente implícito e  $\gamma = 2,0 \text{ ps}^{-1}$  do sistema *complexo-estreptomicina*. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3; Fúcsia: Réplica 4; Verde: Réplica 5. Linha sólida representa a média sobre uma janela de 50 ps

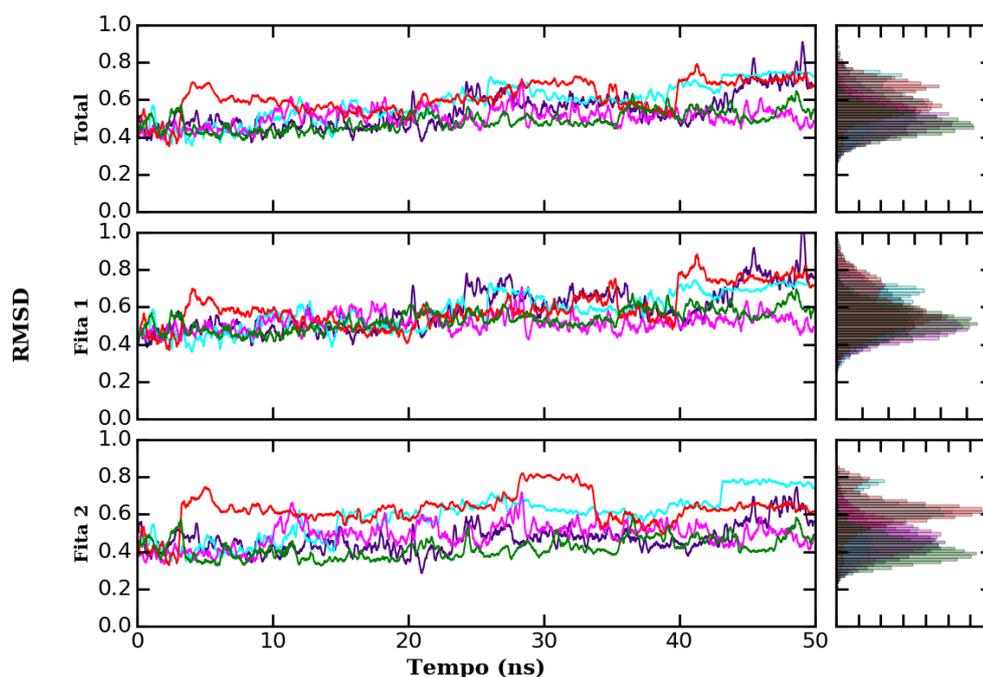


Figura 24 – Raiz quadrada do deslocamento quadrático médio da estrutura do RNA para as simulações do sistema *complexo-estreptomicina* com solvente implícito e  $\gamma = 3,0 \text{ ps}^{-1}$ . Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3; Fúcsia: Réplica 4; Verde: Réplica 5. Linha sólida representa a média sobre uma janela de 50 ps

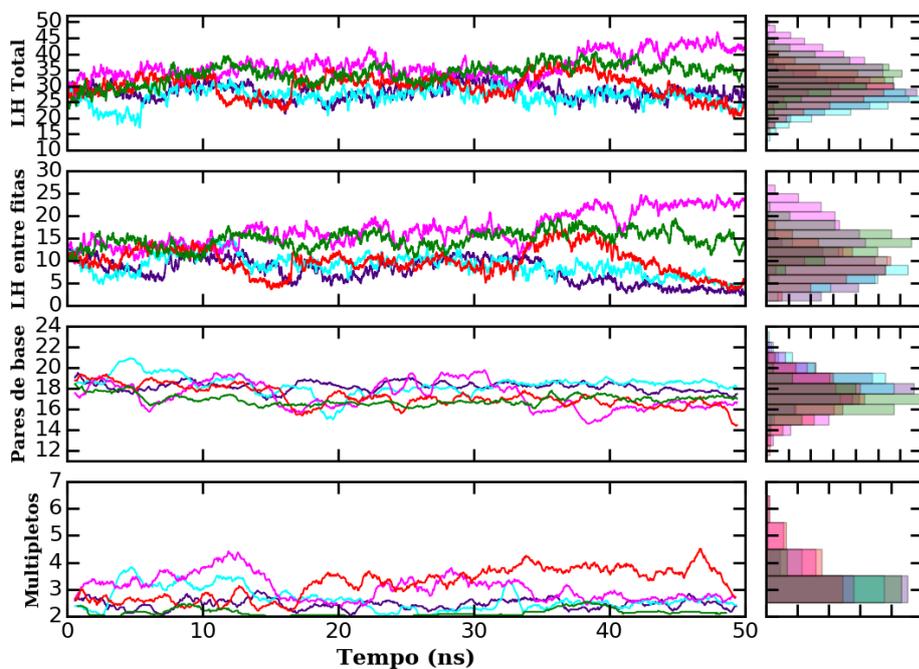


Figura 25 – Parâmetros estruturais para as simulações do sistema *aptâmero* com solvente implícito e  $\gamma = 2,0 \text{ ps}^{-1}$ . Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3; Fúcsia: Réplica 4; Verde: Réplica 5.

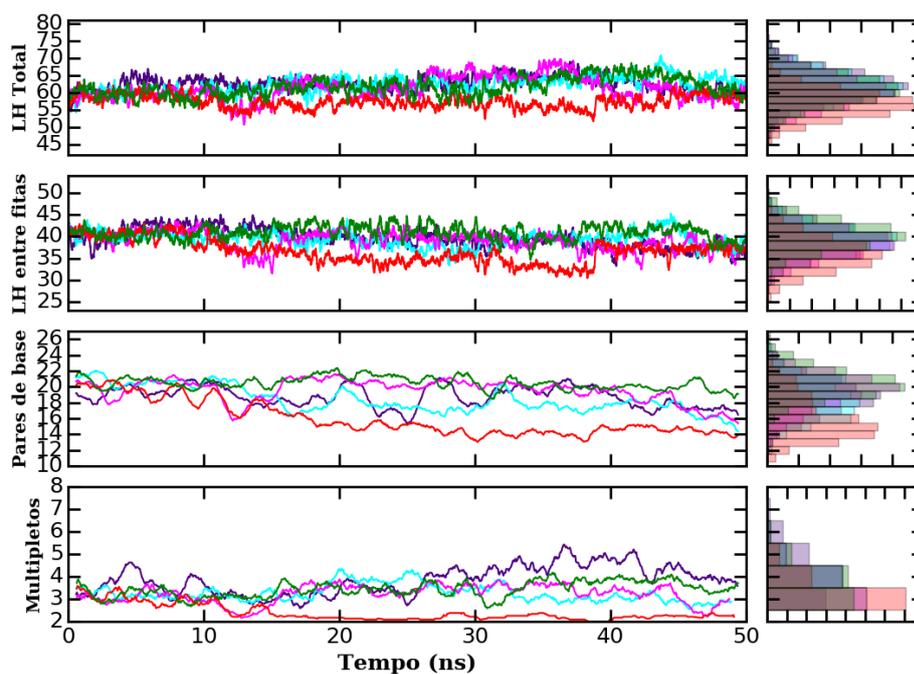


Figura 26 – Parâmetros estruturais para as simulações com solvente implícito e  $\gamma = 2,0 \text{ ps}^{-1}$  do sistema *complexo-streptomicina*. Legenda: Indigo: Réplica 1; Ciano: Réplica 2; Magenta: Réplica 3; Fúscia: Réplica 4; Verde: Réplica 5.

## 9 Velocidade de afastamento e constante de força

A escolha da constante de força,  $k$ , e da velocidade de afastamento do átomo móvel,  $v$ , utilizadas para as simulações de dinâmica molecular dirigida, foi realizada primeiramente através da observação da flutuação das curvas de força e extensão (CFE) em relação aos valores de  $k$  e  $v$  e, posteriormente se comparou os valores de energia livre em função de  $v$ , obtida para os sistemas *aptâmero* e *complexo-estreptomicina*, com os valores experimentalmente obtidos por Nick et al.(35).

### 9.1 Constante de força

De forma a avaliar o comportamento das CFE em relação a constante de força e a velocidade de afastamento do átomo móvel, foram realizadas simulações de dinâmica molecular dirigida com as constantes de força iguais a 100, 200, 300 e 400 kJ mol<sup>-1</sup> nm<sup>-2</sup> e velocidades de afastamento iguais a 0,001, 0,005, 0,010, 0,050 nm ps<sup>-1</sup>, Figuras 27 e 28 para os sistemas *aptâmero* e *complexo-estreptomicina*, respectivamente. Pode-se notar que para ambos os sistemas, a redução do módulo da constante de força leva à melhora da resolução da curva de força e extensão (CFE), ou seja, quanto menor a intensidade de  $k$ , menor a flutuação em seus valores. O mesmo padrão é visto nas CFE dos dois sistemas com o aumento do módulo de  $v$ .

Devido à natureza de não-equilíbrio das simulações por dinâmica molecular dirigida, o valor da distância de ponta a ponta, dpp, flutua entre as replicatas. De forma a calcular a energia livre, é necessário combinar os valores de trabalho de diferentes valores de tempo ou da distância do centro de massas do átomo virtual que sofre o deslocamento.(38, 100) Uma forma de contornar a tarefa de combinar estes valores de uma maneira rápida e eficiente é utilizar uma constante de força suficientemente grande, de forma que o valor de dpp seja o mais próximo possível do valor da posição centro de massas,  $d_{COM}$ .(38) As Figuras 29 e 30 mostram a diferença entre o valor da posição do centro de massas e a distância de ponta a ponta para os sistemas *aptâmero* e *complexo-estreptomicina*, respectivamente. Para ambos os sistemas, se pode ver que a redução da velocidade de afastamento leva a um aumento da diferença entre  $d_{COM}$  e dpp. Conforme a velocidade de afastamento aumenta, a diferença entre  $d_{COM}$  e dpp é reduzida, principalmente para os maiores valores de distância de ponta a ponta. Analisando as Figuras 29 e 30, se pode ver que, para uma mesma velocidade de afastamento, a diferença entre  $d_{COM}$  e dpp não é significativamente afetada pela variação da constante de força,  $k$ . Desta

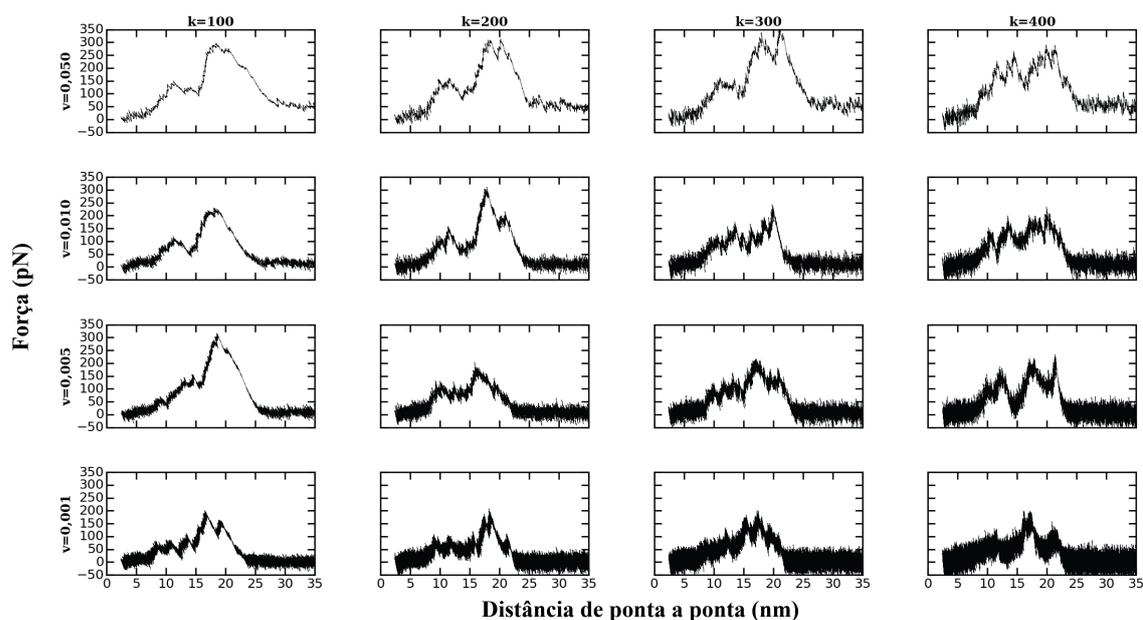


Figura 27 – Curvas de força-extensão para o sistema *aptâmero* para as simulações de calibração da velocidade de afastamento e constante de força para a dinâmica molecular dirigida.

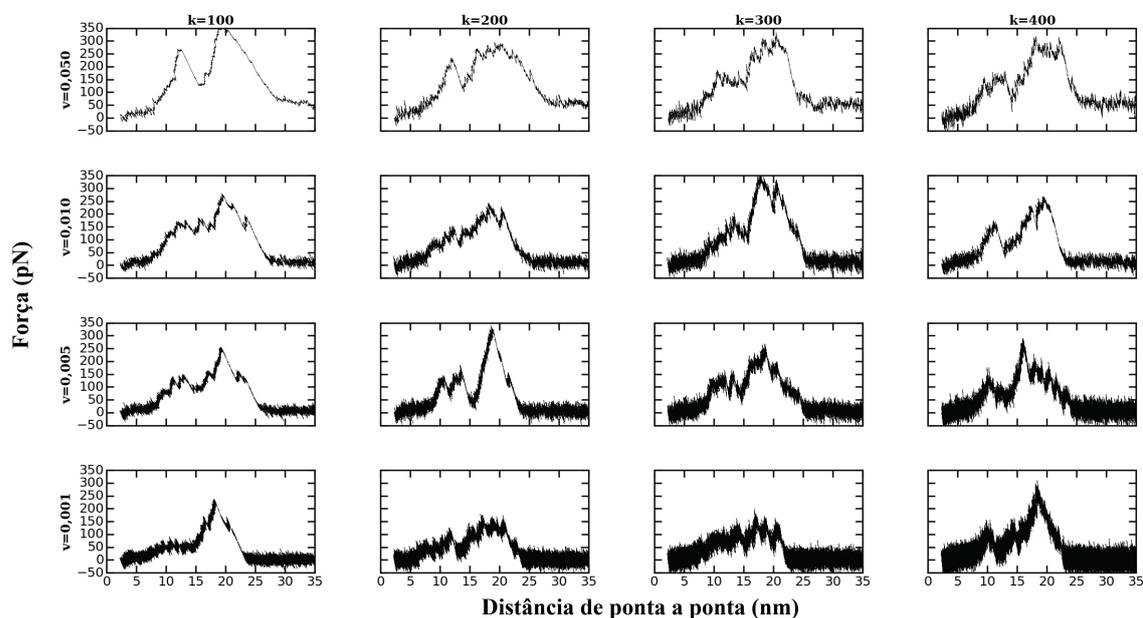


Figura 28 – Curvas de força-extensão para o sistema *complexo-estreptomicina* para as simulações de calibração da velocidade de afastamento e constante de força para a dinâmica molecular dirigida.

forma, se optou utilizar, para a constante de força o valor de  $200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ . Este valor, além de ser o que apresenta menor variação no valor da diferença entre  $d_{COM}$  e  $d_{pp}$ , também foi o valor utilizado nos trabalhos realizados por Naserian-Nik, Tahani e Karttu-

nen(26) e Naserian-Nik, Tahani e Karttunen(27) em simulações de dinâmica dirigida de desenovelamento de ácidos desoxirribonucleicos.

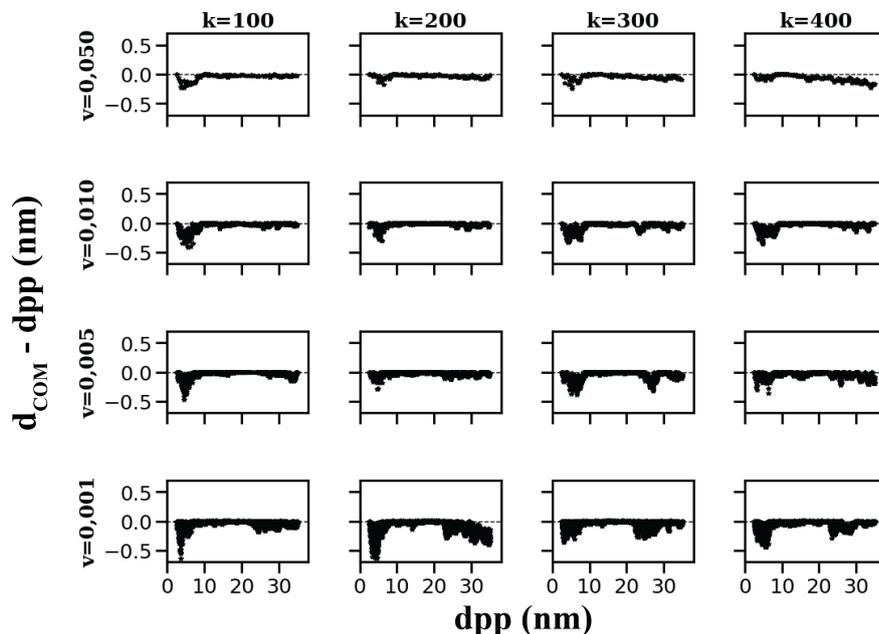


Figura 29 – Diferença entre a posição do centro de massas do átomo virtual,  $d_{COM}$ , e a distância de ponta a ponta,  $dpp$ , para o sistema *aptâmero* para as simulações de calibração da velocidade de afastamento e constante de força para a dinâmica molecular dirigida.

## 9.2 Velocidade de afastamento

Após estabelecer o melhor valor para a constante de força,  $k$ , devemos estabelecer o valor da constante de afastamento,  $v$ . Apesar da igualdade de Jarzynski, derivada por Jarzynski(21), ser, em teoria, independente da velocidade de afastamento, é bem conhecido que a cinética e o comportamento energético dos experimentos de dinâmica molecular de não-equilíbrio serem, de fato, dependentes da velocidade de afastamento.(142, 143, 144, 145, 146, 147) Desta forma, como última etapa para a obtenção do melhor conjunto de valores dos parâmetros, foram realizadas simulações de dinâmica molecular dirigida com velocidades de afastamento iguais a 0,02, 0,03, 0,04, 0,05, 0,06 and 0,07  $\text{nm ps}^{-1}$ , 500 réplicas por  $v$ , para os sistemas *aptâmero* e *complexo-estreptomicina*.

Para a escolha dos melhores parâmetros, foram comparados os valores de energia livre obtidos através das simulações de não-equilíbrio, calculados no ponto de ruptura, com os valores experimentais, obtidos pela técnica de microscopia de força atômica de molécula única por Nick et al.(35). O valor da energia livre no ponto de ruptura em função da velocidade de afastamento para os sistemas *aptâmero* e *complexo-estreptomicina* estão

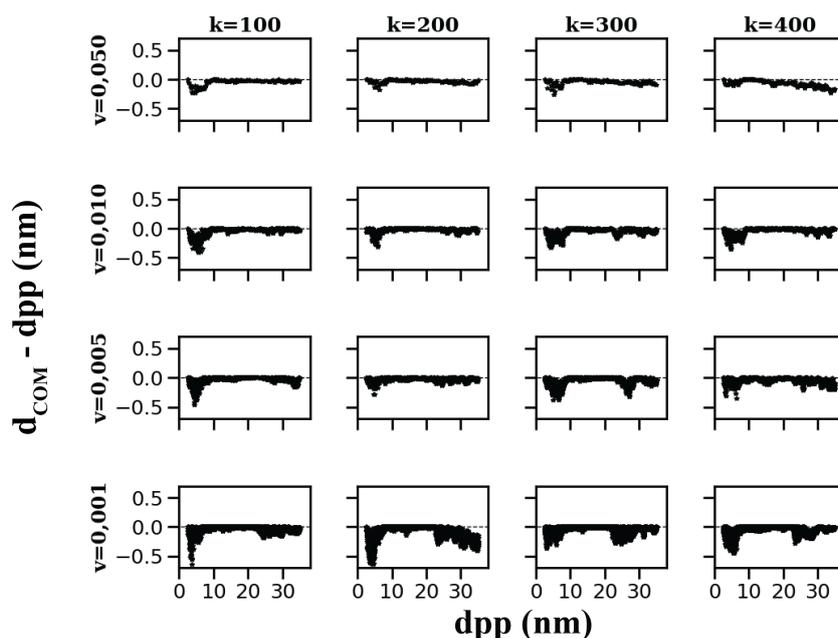


Figura 30 – Diferença entre a posição do centro de massas do átomo virtual,  $d_{COM}$ , e a distância de ponta a ponta,  $dpp$ , para o sistema *complexo-estreptomicina* para as simulações de calibração da velocidade de afastamento e constante de força para a dinâmica molecular dirigida.

apresentados na Tabela 3 e mostrados graficamente na Figura 31. Se pode ver que a energia livre possui um comportamento aproximadamente polinomial de ordem 3 (cúbico) com a velocidade de afastamento, para no domínio da velocidade de afastamento considerada. Também se pode ver que a energia livre possui uma tendência inversamente proporcional a  $v$ , ou seja, conforme a velocidade de afastamento aumenta, a energia livre obtida pelo método se torna menos positiva. Outro ponto a ser observado é a região entre  $0,03 \text{ nm ps}^{-1}$  e  $0,05 \text{ nm ps}^{-1}$ . Nesta região de  $v$  a energia livre varia com pouca intensidade, sendo que para o sistema *aptâmero* esta diferença fica abaixo de  $0,2 \text{ kJ mol}^{-1}$ .

Nick et al.(35) obtiveram, por espectroscopia de força atômica de molécula única, os valores de energia livre de não-equilíbrio para o processo de desenovelamento do aptâmero e de seu complexo com a molécula de estreptomina. Segundo Nick et al.(35), a energia livre para o desenovelamento, medida no ponto de ruptura, é de  $75,1 \pm 0,5 \text{ kJ mol}^{-1}$  e de  $78,5 \pm 1,7 \text{ kJ mol}^{-1}$  para os sistemas *aptâmero* e *complexo-estreptomina*, respectivamente. Comparando os valores de energia livre da Tabela 3, e demonstrados graficamente na Figura 31, com os valores experimentais, se pode ver que os valores calculados para ambos os sistemas se encontram, em módulo, dentro de um intervalo de aproximadamente  $10,0 \text{ kJ mol}^{-1}$  dos valores experimentais para o intervalo de  $v$  de  $0,02 \text{ nm ps}^{-1}$  até  $0,06 \text{ nm ps}^{-1}$ , sendo esta diferença um pouco maior para a velocidade de  $0,07 \text{ nm ps}^{-1}$ .

O comportamento do valor calculado, obtido pela aplicação da equação de Jarzynski

Tabela 3 – Energia livre calculada no ponto de ruptura em função da velocidade de afastamento.

$v$ (nm ps <sup>-1</sup> )	Energia livre <sup>†</sup> (kJ mol <sup>-1</sup> )		
	<i>aptâmero</i>	<i>complexo-estreptomicina</i>	$\Delta\Delta F^*$
0,02	78,04(0,02)	87,96(0,14)	-9,91(0,14)
0,03	71,96(0,07)	83,65(0,10)	-11,69(0,13)
0,04	72,12(0,07)	77,38(0,14)	-5,26(0,15)
0,05	72,02(0,23)	76,63(0,11)	-4,61(0,25)
0,06	68,84(0,05)	70,64(0,35)	-1,80(0,35)
0,07	62,37(0,35)	60,18(0,12)	2,18(0,37)

\*  $\Delta\Delta F = \Delta F_{\text{aptâmero}} - \Delta F_{\text{complexo}}$

† Desvio-padrão entre parênteses.

ao trabalho obtido pelas simulações de não-equilíbrio, em relação aos valores experimentais obtidos por Nick et al.(35), para ambos os sistemas, pode ser melhor observada na Figura 32, onde é mostrado o módulo da diferença entre os valores experimentais e os valores calculados. Se pode ver que, dentro do intervalo de velocidades de afastamento aqui estudado, o módulo da diferença entre os valores de energia livre experimentais e os valores calculados para os sistemas *aptâmero* e *complexo-estreptomicina* possuem comportamentos diferentes. Enquanto que para o sistema *aptâmero*, esta diferença possui uma relação aproximadamente exponencial com a velocidade de afastamento, para o sistema *complexo-estreptomicina* esta diferença é aproximadamente quadrática com um ponto de extremo (ponto de mínimo) em aproximadamente  $v = 0,04$  nm ps<sup>-1</sup>. É interessante notar que, para o sistema *aptâmero*, o módulo desta diferença é praticamente constante no intervalo  $v \in [0,05 ; 0,07]$  nm ps<sup>-1</sup> e inferior a 5,0 kJ mol<sup>-1</sup>.

Pode-se definir, segundo Nick et al.(35), como a energia livre (de não-equilíbrio) para o processo de ligação da estreptomicina à estrutura do aptâmero como a diferença entre a energia livre do sistema *aptâmero* e *complexo-estreptomicina*, no ponto de ruptura:

$$\Delta\Delta F \doteq \Delta F_{\text{aptâmero}} - \Delta F_{\text{complexo}} \quad (9.1)$$

A última coluna da Tabela 3 mostra os valores calculados de  $\Delta\Delta F$  para o intervalo de velocidades de afastamento em estudo. Se pode ver que a diferença de energia livre se torna mais positiva com o aumento da velocidade de afastamento, variando de aproximadamente -10,0 kJ mol<sup>-1</sup> até aproximadamente +2,0 kJ mol<sup>-1</sup> conforme  $v$  varia de 0,02 nm ps<sup>-1</sup> até 0,07 nm ps<sup>-1</sup>.

Utilizando microscopia de força atômica de molécula única, Nick et al.(35) obtiveram  $\Delta\Delta F = -3,4 \pm 1,8$  kJ mol<sup>-1</sup>, no ponto de ruptura. Comparando os valores calculados, Tabela 3, com o valor experimental, se pode ver que a diferença entre estes, em módulo, é menor que 10,0 kJ mol<sup>-1</sup>. Notavelmente, para o intervalo  $v \in [0,04 ; 0,06]$  • (nm ps<sup>-1</sup>), a diferença, em módulo, entre o valor de  $\Delta\Delta F$  calculado e experimental é menor que

1,9 kJ mol<sup>-1</sup>, ou seja, a diferença entre os valores é menor que o desvio padrão do valor medido por microscopia de força atômica de molécula única, indicando uma excelente concordância do valor calculado com o valor experimental, no intervalo de velocidades de afastamento estudado neste trabalho.

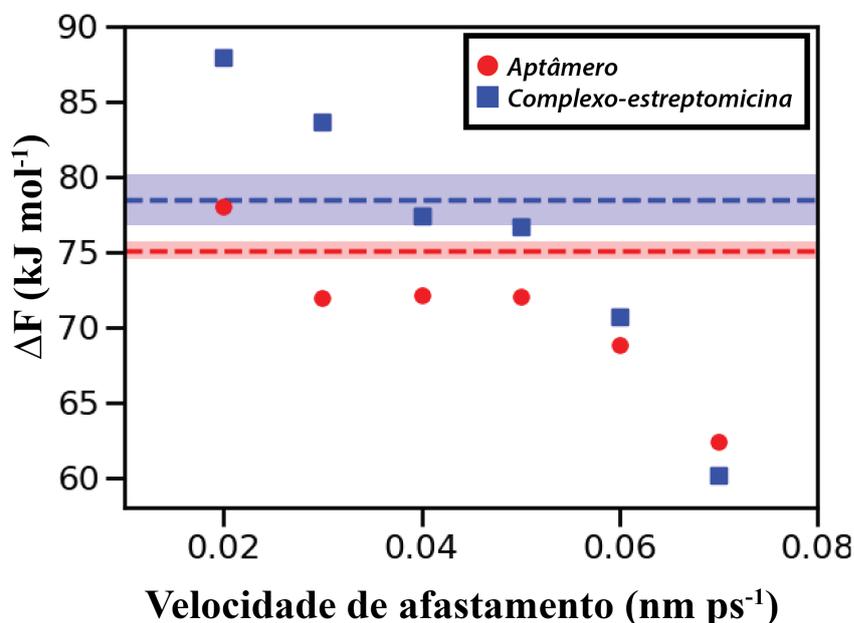


Figura 31 – Energia livre no ponto de ruptura em função da velocidade de afastamento para os sistemas *aptâmero* e *complexo-estreptomicina*. As linhas tracejadas e áreas sombreadas correspondem aos valores de energia livre e intervalo de desvio padrão, respectivamente, obtidos experimentalmente por Nick et al.(35).

Considerando que a diferença entre a energia livre de não-equilíbrio calculada e a obtida experimentalmente por Nick et al.(35) é em módulo menor do que 10,0 kJ mol<sup>-1</sup> para o intervalo de velocidades de afastamento aqui estudado, e, considerando que a energia livre de não-equilíbrio para o processo de ligação da estreptomicina à estrutura do aptâmero, calculada pela diferença das energias livres entre os sistemas no ponto de ruptura, se manteve dentro da distância de 10,0 kJ mol<sup>-1</sup> do valor experimental, podemos inferir que, para o valor da constante de força  $k$  escolhido na Seção 9.1, todo o intervalo de velocidades de afastamento utilizado nesta seção, possibilita a realização de simulações por dinâmica molecular dirigida que reproduzem os valores de energia livre de não-equilíbrio obtidos experimentalmente.

Levando em consideração o que foi exposto neste capítulo, de forma a simplificar a análise do mecanismo de desenovelamento do aptâmero de RNA induzido por força, deste ponto em diante as simulações por dinâmica molecular dirigida serão realizadas com apenas uma velocidade de afastamento. Uma vez que as menores diferenças, em módulo, entre os valores calculados e os obtidos experimentalmente, para os sistemas

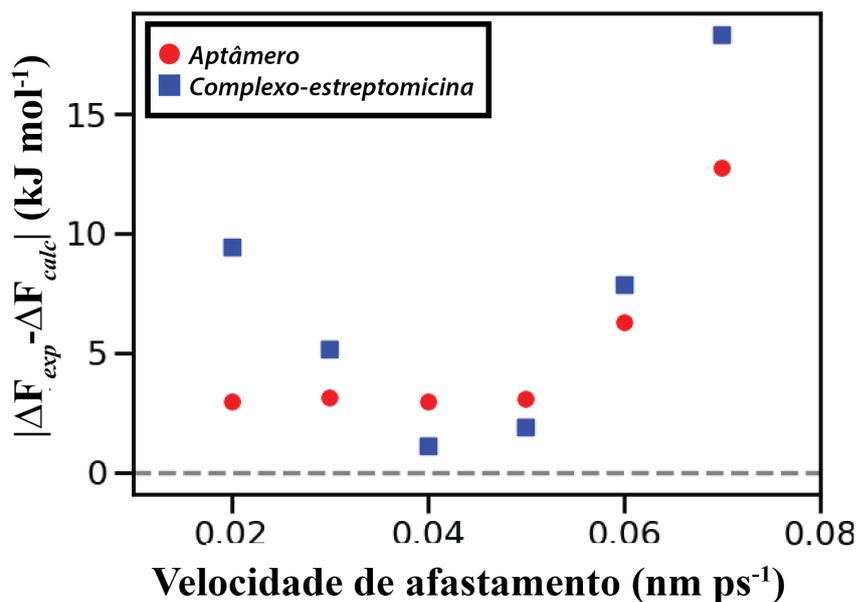


Figura 32 – Módulo da diferença de energia livre, no ponto de ruptura, entre os valores calculados e obtidos experimentalmente(35), em função da velocidade de afastamento para os sistemas *aptâmero* e *complexo-estreptomicina*.

*aptâmero* e *complexo-estreptomicina*, foram encontrados para a velocidade de afastamento igual a  $v = 0,04 \text{ nm ps}^{-1}$ , este valor será adotado, juntamente com a constante de força  $k = 200,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ , como representativo para as simulações. Entretanto, quando necessário, as outras velocidades serão abordadas.

# 10 Desenovelamento dos sistemas *aptâmero* e *complexo-estreptomicina*

Neste capítulo estão descritos os resultados das simulações por dinâmica molecular dirigida (DMD), aplicada no estudo do processo de desenovelamento induzido por força de um aptâmero de RNA e do complexo formado pelo aptâmero e estreptomicina. Conforme a Seção 6.1.4 e os parâmetros obtidos nos Capítulos 8 e 9, foram realizadas 500 simulações de DMD de forma a emular o desenovelamento induzido por força da molécula de RNA, conforme ocorre no experimento de microscopia de força atômica de molécula única. O capítulo está dividido em duas seções: Uma primeira seção, com a descrição das curvas de força-extensão obtidas durante as simulações de desenovelamento do RNA e a análise estrutural deste processo; e uma segunda seção com a análise da energia livre do processo de desenovelamento do RNA.

## 10.1 Análise estrutural

Nesta seção serão discutidos os processos de desenovelamento induzido por força dos sistemas *aptâmero* e *complexo-estreptomicina*, simulados por dinâmica molecular dirigida. A Figura 33 mostra o comportamento da força média de extensão em função da distância de ponta a ponta (dpp), ou curva de força-extensão (CFE), para os sistemas *aptâmero* e *complexo-estreptomicina*. Se pode ver que as CFE de ambos os sistemas exibem o mesmo comportamento, sendo que ambas as curvas de força-extensão exibem dois máximos ou *force jumps*. O segundo máximo nestas CFE correspondem a força de ruptura nos experimentos de espectroscopia de força atômica de molécula única, e portanto é, deste ponto em diante, chamado de ponto de ruptura. Analisando o padrão das CFE, se pode dividir as CFE em quatro regiões distintas, indicadas pelos numerais romanos, I a IV, na Figura 33:

- I) Nesta região, a força de extensão aumenta rapidamente com o aumento da dpp até aproximadamente 12 nm. No fim desta região se encontra o primeiro máximo da CFE;
- II) Região dentro do intervalo de 12 nm a 15 nm, onde a força de extensão se mantém aproximadamente constante;
- III) Região entre 15 nm e 18 nm onde a força de extensão aumenta linearmente até que a força atinge o segundo máximo da CFE ou ponto de ruptura;

IV) Nesta região, após o ponto de ruptura, a força de extensão decresce linearmente com a dpp até atingir um valor constante.

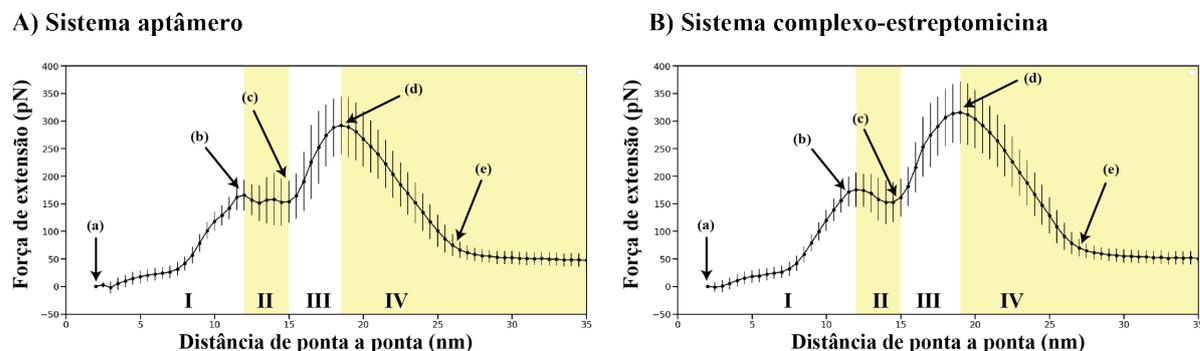


Figura 33 – Curvas de força-extensão médias para os sistemas A) *aptâmero* e B) *complexo-estreptomicina*. As médias foram calculadas sobre 500 simulações.

De forma a compreender o que ocorre com a estrutura do RNA durante o desenovelamento induzido por força, algumas estruturas representativas foram extraídas ao longo das CFE, marcadas com as letras (a) a (e) na Figura 33. As estruturas referentes aos pontos (a) a (e) são mostradas nas Figuras 34 e 35 para os sistemas *aptâmero* e *complexo-estreptomicina*, respectivamente. As estruturas iniciais, de ambos os sistemas, em seu estado de equilíbrio se encontram nas Figuras 34(a) e 35(a). As Figuras 34(b) e 35(b) mostram as estruturas para os sistemas ao final da região I. É possível ver para o sistema *aptâmero* que a estrutura se encontra consideravelmente alongada em relação à estrutura inicial, onde o sítio de ligação já apresenta grande desestruturação. Em contrapartida, para o sistema *complexo-estreptomicina* o centro da estrutura ainda permanece estruturado em torno da molécula de estreptomicina ao final da região I.

A Figura 36 mostra o número médio de pares de base em função da distância de ponta a ponta para ambos os sistemas. Pode-se ver que até aproximadamente 10 nm de extensão, o número de pares de base se mantém praticamente constante, diminuindo linearmente após este valor de extensão, conforme a distância de ponta a ponta atinge o fim da região I. O comportamento do número de pares de base na região I indica que, embora a molécula esteja sendo alongada, ou seja, a estrutura terciária da molécula de RNA esteja sendo modificada, a estrutura secundária do aptâmero se mantém praticamente constante.

O final da região II está representado nas Figuras 34(c) e 35(c) para os sistemas *aptâmero* e *complexo-estreptomicina*, respectivamente. Se pode ver que para o sistema *aptâmero* que o sítio de ligação está completamente desestruturado, onde a estrutura apresenta uma forma alongada com grande separação entre as duas fitas. Entretanto, se pode ver na Figura 35(c) que ao final da região II da CFE para o sistema *complexo-estreptomicina* que o sítio de ligação ainda se encontra parcialmente estruturado em torno

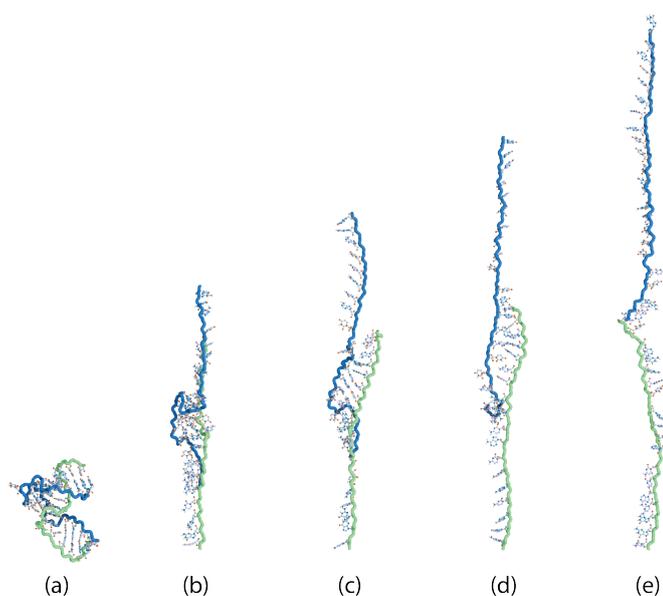


Figura 34 – Estruturas representativas ao longo da CFE para o sistema *aptâmero*. a) Estrutura inicial. b) Estrutura ao final da região I. c) Estrutura ao final da região II. d) Estrutura no ponto de ruptura. e) Estrutura no último ponto de contato.

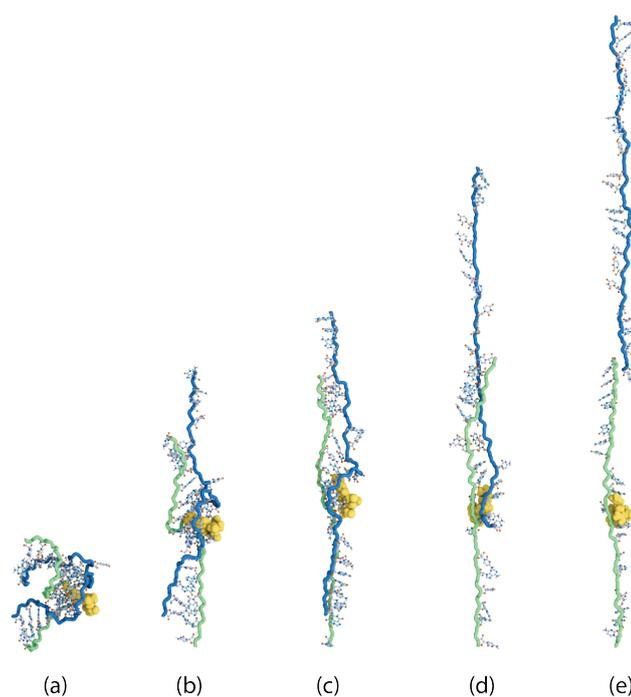


Figura 35 – Estruturas representativas ao longo da CFE para o sistema *complexo-estreptomicina*. a) Estrutura inicial. b) Estrutura ao final da região I. c) Estrutura ao final da região II. d) Estrutura no ponto de ruptura. e) Estrutura no último ponto de contato.

da molécula de estreptomicina. Este fato demonstra o efeito de estabilização da organização estrutural que a molécula de estreptomicina causa ao centro da estrutura do aptâmero. Com isso, se pode inferir qualitativamente, que a molécula de estreptomicina confere estabilidade à estrutura terciária do aptâmero. De fato, o efeito estabilizador da molécula de estreptomicina ao centro da estrutura do aptâmero se mostra evidente ao compararmos a região II na CFE, Figura 33, para ambos os sistemas, onde podemos notar que após o primeiro evento de *force jump* na CFE a força possui um comportamento mais definido para a sistema *complexo-estreptomicina*, como esperado para um sistema mais estável.

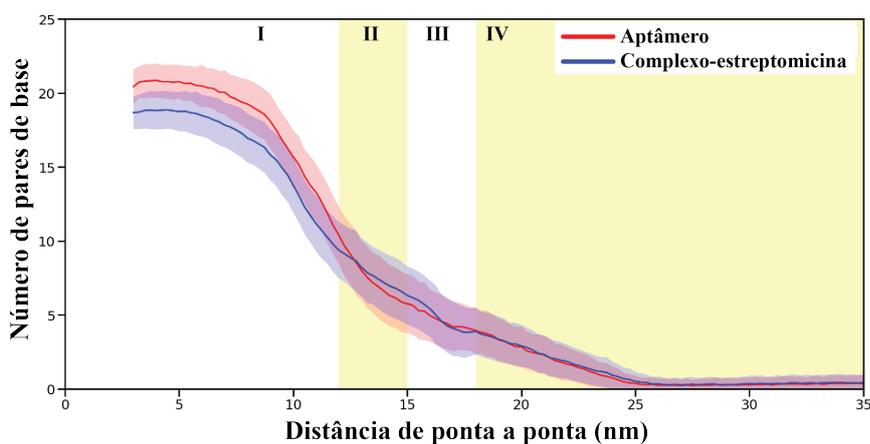


Figura 36 – Número de pares de base em função da dpp para os sistemas *aptâmero* e *complexo-estreptomicina*.

Uma forma de inferir a estabilização do sítio de ligação é mapear com quais nucleotídeos do aptâmero a molécula de estreptomicina realiza ligações de hidrogênio em função da distância de ponta a ponta. A Figura 37 mostra a frequência, em função da dpp, com que a molécula de estreptomicina realizou ligações de hidrogênio com o aptâmero nas 500 simulações por dinâmica molecular dirigida. Se pode ver claramente que a molécula de estreptomicina realiza tais ligações com diversos nucleotídeos do sítio de ligação. Notadamente, os nucleotídeos U16, U17, C109 e G110 interagem com a estreptomicina até que o aptâmero atinja a dpp de aproximadamente 15 nm. Portanto, se pode ver que até que o final da região II e início da região III a molécula de estreptomicina realiza no mínimo 4 ligações de hidrogênio com os nucleotídeos do sítio de ligação e, desta forma, estabilizando a estrutura deste.

Após o término da região II, as CFE de ambos os sistemas mostram um aumento linear da força de extensão com a distância de ponta a ponta até que o segundo máximo seja alcançado. Nesta região, denominada região III, a estrutura do aptâmero continua a ser alongada e a desenovelar até atingir uma estrutura em forma de escada no ponto de ruptura, Figuras 34(d) e 35(d). Depois de o sistema atingir o ponto de ruptura da CFE, o sistema entra na região IV onde a força de extensão decresce linearmente até

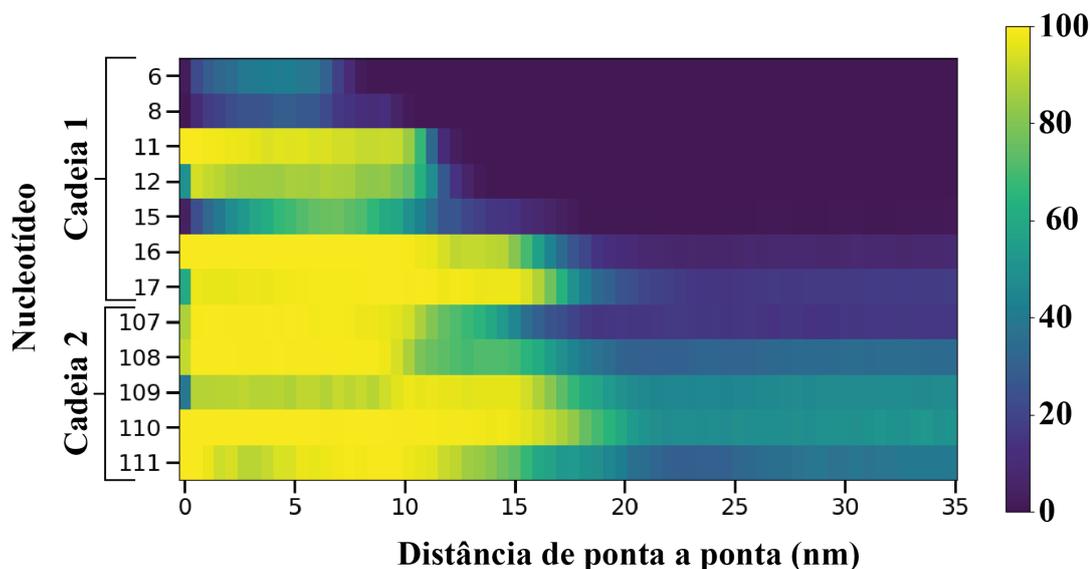


Figura 37 – Padrão de ligações de hidrogênio (%) entre a molécula de estreptomicina e aptâmero no sistema *complexo-estreptomicina*.

atingir um valor constante. O decréscimo linear da força de extensão pode ser explicado pela conformação de escada atingida no ponto de ruptura. Após o sistema atingir esta conformação, uma das cadeias do aptâmero desliza pela outra cadeia com um mínimo de interação. Com o aumento da distância de ponta a ponta, nesta conformação de escada, menos interação existe entre as cadeias do aptâmero e, conseqüentemente, menor a força necessária para puxar a cadeia móvel. Desta forma, a força de extensão diminui linearmente até que as cadeias se encontrem em um estado de mínimo contato, mostrados nas Figuras 34(e) e 35(e). Uma vez atingido este ponto de mínimo contato, ou último contato, a força de extensão atinge um valor final constante. Este valor constante é devido ao arraste hidrodinâmico da cadeia móvel através do solvente. Este movimento onde a cadeia móvel desliza em relação à cadeia fixa é corroborado pelo decréscimo linear no número de pares de base, Figura 36, nas regiões III e IV.

### 10.1.1 Modelo estrutural para desenovelamento induzido por força

Racionalizando as estruturas representativas extraídas ao longo das CFE, mostradas nas Figuras 34 e 35, podemos construir um modelo qualitativo para o processo de desenovelamento induzido por força externa de um aptâmero de RNA de duas cadeias. Conforme as estruturas para ambos os sistemas, se pode ver que há um processo hierárquico em duas etapas. Primeiramente, a estrutura terciária é desfeita devido ao alongamento da estrutura (regiões I e II), seguida de uma segunda etapa onde a estrutura secundária do aptâmero é desfeita devido ao deslizamento da cadeia móvel em relação a

cadeia fixa (regiões III e IV). Adicionalmente ao comportamento bifásico das CFE (dois máximos nas curvas de força-extensão) e às estruturas representativas extraídas ao longo desta, podemos calcular alguns parâmetros estruturais de forma a dar mais evidências a este comportamento.

A Figura 38 mostra o número médio de multipletos em função da distância de ponta a ponta. Multipletos são interações nas quais três ou mais bases coplanares interagem por meio de ligações de hidrogênio e são indicativos da presença de estrutura terciária.(48) Conforme mencionado acima no Capítulo 7, a estrutura do sistema *complexo-estreptomicina* possui dois tripletos G6•G13•C113 e U16•C18•G110 e, para o sistema *aptâmero*, há ainda a formação do triplete U19•A105•C106, de forma que os multipletos presentes em ambos os sistemas são, de fato, tripletos. Se deve notar que estes tripletos são formados por dois nucleotídeos de uma cadeia e o terceiro nucleotídeo pertence a outra cadeia, de forma que esta interação, formada por estes três nucleotídeos, são parte da estrutura terciária do aptâmero. Se pode ver na Figura 38 que o número de multipletos, ou tripletos, se mantém constante por grande parte da região I, caindo drasticamente entre as regiões I e II e passando a ser virtualmente zero da região III em diante. Em outras palavras, a estrutura terciária é desfeita ao longo do final da região I e região II.

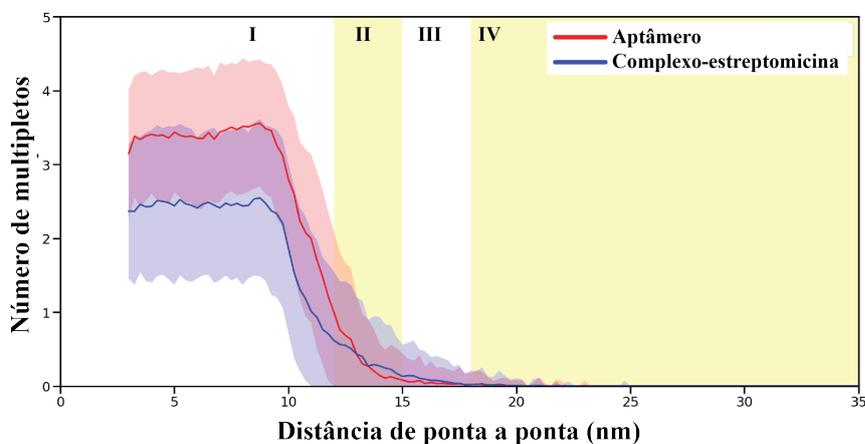


Figura 38 – Número de multipletos em função da dpp para os sistemas *aptâmero* e *complexo-estreptomicina*.

Outro indicativo de estrutura terciária é o número de hélices. Uma hélice é definida como interações entre dois ou mais pares de bases empilhados.(148) A Figura 39 mostra o número médio de hélices em função da distância de ponta a ponta. Igualmente ao número médio de multipletos, o número médio de hélices se mantém constante até aproximadamente o final da região I e, então, começa a decrescer até atingir zero na região III. Diferentemente do número de multipletos, que decresce abruptamente, o número médio de hélices possui um decaimento linear. Este comportamento se deve ao fato de que, devido

ao empilhamento de pares de base, as hélices são estruturas terciárias mais estáveis frente ao alongamento do que as voltas ou dobras, indicadas pelo número de multipletos.

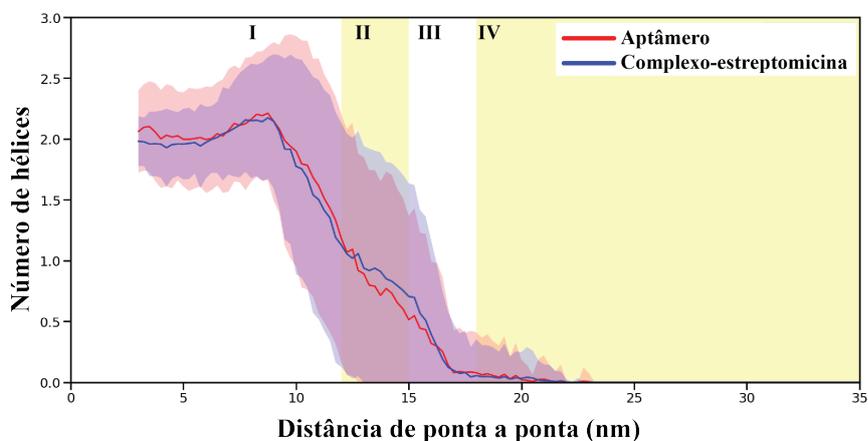


Figura 39 – Número médio de hélices em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-estreptomicina*.

Na seção anterior se comentou que o número médio de pares de base em função da distância de ponta a ponta, Figura 36 (página 86), indica que a estrutura secundária se mantém praticamente constante na região I. Entre o final da região I e o início da região II há uma diminuição acentuada no número de pares de base. Após o início da região II, o número médio de pares de base diminui linearmente até que não exista mais pares de base na estrutura. Este decréscimo linear é devido ao deslizamento de uma cadeia sobre a outra. Desta forma, embora o último nucleotídeo da cadeia móvel possa formar um par de base adequado com o próximo nucleotídeo da cadeia móvel, os nucleotídeos superiores da cadeia móvel vão se afastando da cadeia fixa e não podem mais interagir com a outra cadeia. Este comportamento pode ser bem visualizado, como já mencionado na seção anterior, nas Figuras 34(c-d) e 35(c-d).

A Figura 40 mostra o número de ligações de hidrogênio em função da distância de ponta a ponta. Entretanto é necessário ter cuidado com a correta definição deste parâmetro para que não sejam realizadas interpretações equivocadas. Seja  $A = (a_i)_{i \in \{1, \dots, n_A\}}$  e  $B = (b_j)_{j \in \{1, \dots, n_B\}}$  os conjuntos de átomos capazes de realizar ligações de hidrogênio das cadeias 1 e 2, respectivamente. O número total de ligações de hidrogênio é computado para os pares  $(i, j)_{i \in A; j \in B}$ . Desta forma, para o número total de ligações de hidrogênio, serão computadas estas interações **intra** e **entre** as cadeias. Embora importante do ponto de vista energético, e conseqüentemente será discutida na Seção 10.2.3 mais adiante, não acrescenta muita informação ao nosso modelo de desenovelamento.

A Figura 40a mostra o número total de ligações de hidrogênio. Nela se pode ver que esta quantidade se mantém aproximadamente constante durante quase toda a região I e depois seu comportamento muda para um decréscimo linear. É importante notar que

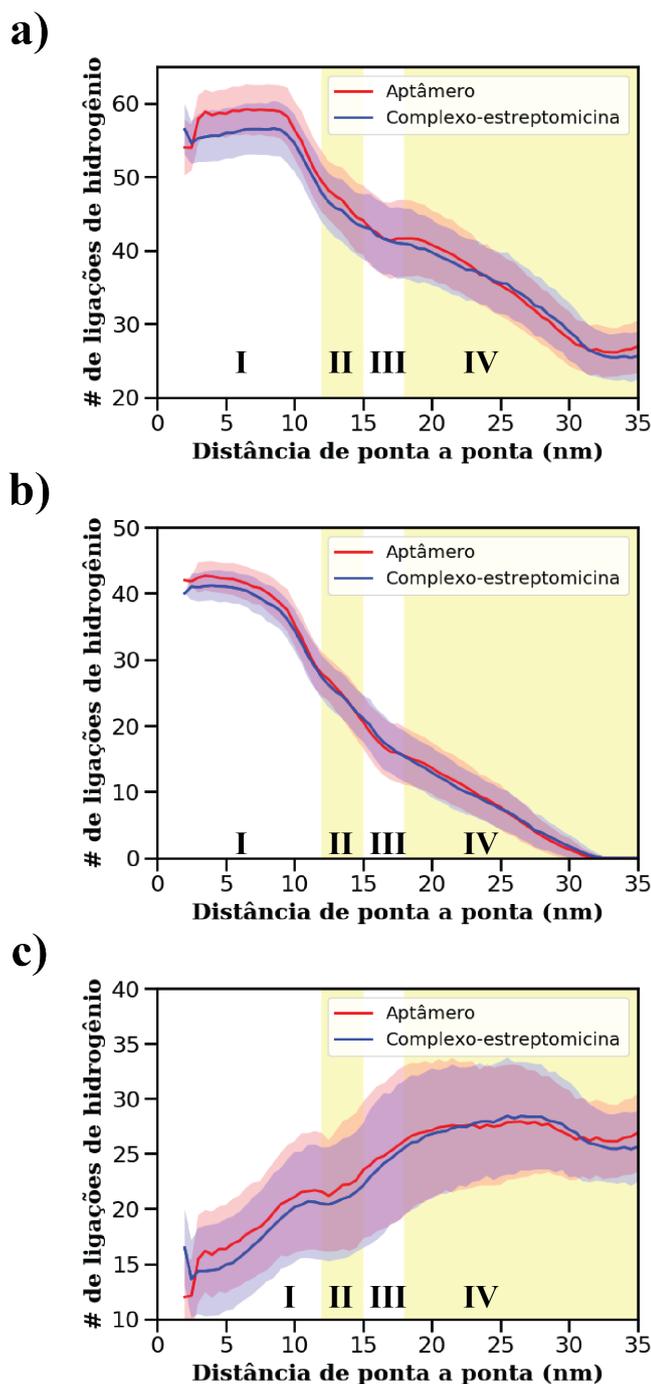


Figura 40 – Número de ligações de hidrogênio em função da dpp para os sistemas *aptâmero* e *complexo-estreptomicina*. a) Número total de ligações de hidrogênio; b) Número de ligações de hidrogênio entre as cadeias 1 e 2; c) Número total de ligações de hidrogênio intra cadeia.

o número total de ligações de hidrogênio no início das simulações de desenvolvimento é maior para o sistema *complexo-estreptomicina* do que para o sistema *aptâmero*, de acordo com as simulações de equilíbrio realizadas neste trabalho e por Nick et al.(35). Entretanto, uma vez que o processo de desenvolvimento começa, este padrão se inverte. Ou seja, o

número total de ligações de hidrogênio é maior para o sistema *aptâmero* do que para o sistema *complexo-estreptomicina* durante seu comportamento constante na região I. Este fato decorre da maior possibilidade de contato entre as duas cadeias ocorrerem no sistema onde não há a presença da molécula de estreptomicina. Também se deve notar que o número total de ligações de hidrogênio não tende a zero conforme a distância de ponta a ponta cresce. Este comportamento decorre da tendência das cadeias realizarem ligações de hidrogênio intra cadeias, sendo abordado a seguir.

O número total de ligações de hidrogênio pode ser dividido em dois tipos: Ligações de hidrogênio entre cadeias e ligações de hidrogênio intra cadeias. Se computarmos as ligações de hidrogênio apenas para os pares de átomos  $(i, j)_{i \in A; j \in B}$ , estaremos computando o número de ligações entre fitas. Esta interação é diretamente relacionada a estrutura secundária do aptâmero de RNA e está diretamente relacionada com o número de pares de base e, em menor grau (devido a proporção do número de pares de base em relação ao número de multipletos), com o número de multipletos. A Figura 40b mostra o número de ligações de hidrogênio entre as fitas para os sistemas *aptâmero* e *complexo-estreptomicina*. Se pode ver que esta propriedade apresenta o mesmo padrão decrescente que o número de pares de base, Figura 36 (página 86). Diferente do número total de ligações de hidrogênio, o número de ligações de hidrogênio entre cadeias tende a zero com o aumento da distância de ponta a ponta, pois, de fato, há a separação entre as cadeias no processo de desenovelamento.

Ao se analisar a região I da Figura 40b, se pode ver que o número de ligações de hidrogênio entre as cadeias para ambos os sistemas se mantém constante durante praticamente toda esta região. Uma vez que o número de ligações entre cadeias reflete a integridade da estrutura secundária da molécula de RNA, se pode inferir que o modelo de desenovelamento hierárquico em duas etapas, aqui postulado, está de acordo com o observado nas simulações computacionais. Também se pode observar que este número decresce linearmente após a região I, de modo reforçar a ideia de que uma cadeia desliza sobre a outra, diminuindo progressivamente a interação entre as cadeias.

Ao computarmos as ligações de hidrogênio apenas para os pares de átomos de uma mesma cadeia, ou seja,  $(i, j)_{i \in A; j \in A} \cup (i, j)_{i \in B; j \in B}$ , estaremos computando o número total de ligações intra fitas. Embora esta interação esteja relacionada com a estrutura terciária da molécula de RNA, ela não é específica como parâmetro indicador de organização. A Figura 40c mostra o número total de ligações de hidrogênio intra fitas. Diferente dos comportamentos em função da distância de ponta a ponta até aqui observados, o número total de ligações de hidrogênio intra fitas cresce conforme a distância de ponta a ponta cresce. Este comportamento é devido ao fato as extremidades das cadeias, ao serem separadas na segunda etapa do modelo de desenovelamento induzido por força aqui proposto, se dobram sobre si formando ligações de hidrogênio intra cadeias.

A segunda etapa no desenvolvimento induzido por força do aptâmero de RNA fica ainda mais evidente no comportamento da área de superfície de contato(149, 150, 151), calculada pela diferença da área acessível ao solvente(152), mostrada na Figura 41. Se pode ver que a área de superfície de contato (BSA, do inglês *buried surface area*) fica constante até o final da região I, passando para um segundo patamar do início da região II até aproximadamente o meio da região III, onde as estruturas de ambos os sistemas, *aptâmero* e *complexo-estreptomicina* atingem uma conformação de escada. Após esta conformação de escada ser atingida, as estruturas secundárias começam a se desfazer e a área de contato começa a diminuir linearmente com o aumento da distância de ponta a ponta, até não haver mais contato entre as fitas.

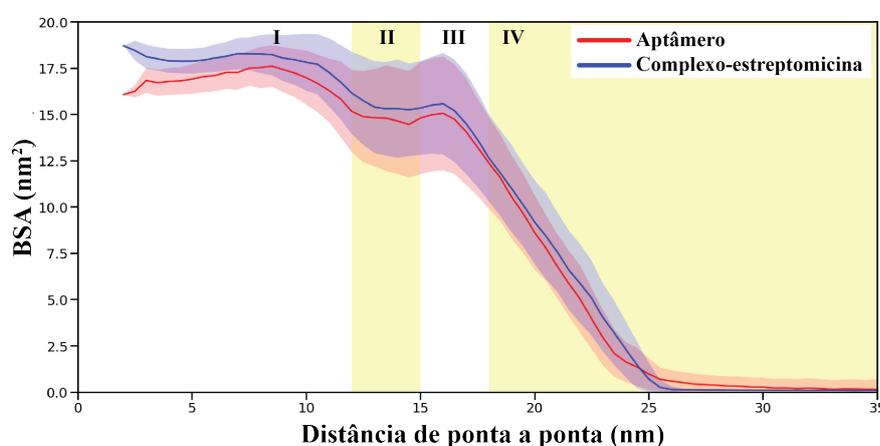


Figura 41 – Área de contato (BSA, do inglês *buried surface area*) entre as cadeias do aptâmero em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-estreptomicina*.

## 10.2 Análise energética

Experimentos de desenvolvimento utilizando microscopia de força atômica são cada vez mais utilizados para se avaliar eventos moleculares raros.(9, 12, 144, 142) A seção anterior, Seção 10.1, tratou da análise estrutural das simulações de dinâmica molecular dirigida, realizadas com o intuito de simular o desenvolvimento induzido por força de um aptâmero de RNA como o ocorrido nos experimentos de microscopia de força atômica realizados por Nick et al.(35). Aliada à análise estrutural, o cálculo da energia livre  $\Delta F$  possui um papel fundamental para a melhor entendimento do mecanismo e dos determinantes estruturais dos processos biomoleculares.(39, 38) Uma vez que estes experimentos são realizados em regime de não-equilíbrio, o cálculo da energia livre de equilíbrio, através do trabalho de não-equilíbrio realizado sobre o sistema, utilizando o método apresentado no Capítulo 5 e seção 5.1, é importante para se compreender o comportamento nos experimentos de microscopia de força atômica de molécula única.(38, 153)

A Seção 9.2 tratou da comparação entre o valor de energia livre calculada utilizando a metodologia de não-equilíbrio no ponto de ruptura, descrita no Capítulo 6, com o valor obtido experimentalmente por Nick et al.(35). Entretanto, nesta seção será abordada a energia livre ao longo do processo de desenovelamento induzido por força externa, ou seja o perfil de energia livre, de forma a obtermos mais informações a cerca deste processo. Esta seção esta dividida em três partes. Na primeira parte será discutida a forma das curvas de energia livre em função da distância de ponta a ponta, dos sistemas *aptâmero* e *complexo-estreptomicina*, e sua correlação com o modelo de desenovelamento proposto na seção anterior. Após a descrição dos perfis de  $\Delta F$ , será realizada a discussão da diferença destes perfis entre os sistemas *aptâmero* e *complexo-estreptomicina*. Por fim, esta seção será finalizada com uma discussão sobre a decomposição dos perfis de energia livre em suas componentes energéticas e entrópicas.

### 10.2.1 Análise energética do processo de desenovelamento

Os valores de energia livre foram calculados utilizando o método de extrapolação linear proposto por Ytreberg e Zuckerman(39), conforme Seção 6.2, onde se usou a igualdade de Jarzynski(21) como estimador para  $\Delta F$ :

$$\Delta F = -\frac{1}{\beta} \ln \langle e^{-\beta W} \rangle \quad (10.1)$$

A média da exponencial do trabalho  $\langle e^{-\beta W} \rangle$ , em função da distância de ponta a ponta, foi calculada em relação as 500 simulações de dinâmica molecular dirigida, ou seja, de não-equilíbrio, utilizando os parâmetros escolhidos nos Capítulos 8 e 9.

A Figura 42 mostra a energia livre, em função da distância de ponta a ponta, para os sistemas *aptâmero* e *complexo-estreptomicina*. Como o desenovelamento do RNA de ambos os sistemas é realizado mediante aplicação de força externa, a energia livre em função da distância de ponta a ponta (dpp) é sempre positiva, uma vez que é necessário fornecer energia ao sistema para que ele desnature. O perfil de energia livre para ambos os sistemas mostram um comportamento crescente com a distância de ponta a ponta, com dois pontos de inflexão, o primeiro em aproximadamente 8,5 nm e o segundo, menos pronunciado, em 25,0 nm. Se pode ver que em ambos os sistemas a energia possui uma taxa de crescimento elevada até que a dpp seja de aproximadamente 8,0 nm. Esta taxa de crescimento do valor de  $\Delta F$  diminui após este ponto de inflexão em ambos os sistemas e fica ainda menor após o segundo ponto de inflexão.

Conforme a seção anterior, Seção 10.1, as curvas de força e extensão resultantes do desenovelamento induzido por força apresentam quatro regiões distintas, indicadas por numerais romanos ao fundo dos perfis de energia livre da Figura 42. A análise de estruturas representativas ao longo das curvas de força e extensão, Figuras 34 e 35, bem

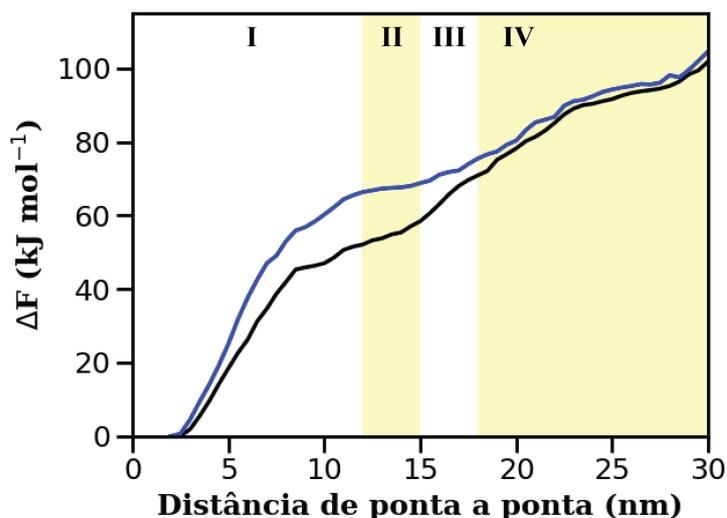


Figura 42 – Energia livre  $\Delta F$  em função da distância de ponta a ponta para os sistemas *aptâmero*, linha preta, e *complexo-estreptomicina*, linha azul.  $v = 0,04 \text{ nm ps}^{-1}$  e  $k = 200,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$

como a análise de parâmetros estruturais levou à formulação de um modelo em duas etapas para o desenovelamento induzido por força externa da estrutura do RNA.

De acordo com o modelo em duas etapas da Seção 10.1.1, foi postulado que na primeira etapa, que acontece nas regiões I e II, ocorre a desorganização da estrutura terciária do RNA. Se relacionarmos as curvas de energia livre, Figura 42, com o número de multipletos, Figura 38 (página 88), e o número de hélices, Figura 39 (página 89), podemos ver claramente que o primeiro ponto de inflexão nos perfis de energia livre coincidem com a transição da estrutura completamente formada, estrutura nativa, do aptâmero para a estrutura em conformação de escada.

Na segunda etapa do processo de desenovelamento, o modelo proposto na Seção 10.1.1 prevê que, após a estrutura sofrer a transição da conformação nativa para a conformação de escada, a cadeia móvel deslize sobre a cadeia fixa, desfazendo e formando novos pares de base até que a estrutura secundária seja completamente desfeita. Pode-se ver nos perfis de energia livre da Figura 42 que após a primeira inflexão nos perfis, em aproximadamente 8,0 nm, há uma redução na taxa de crescimento da energia livre em função da distância de ponta a ponta, redução esta que segue a mesma tendência que a do número de pares de base formados, Figura 36 (página 86), indicando que a energia livre do processo durante esta etapa é fortemente influenciada pela formação e ruptura dos pares de base. Com isto, podemos ver que os perfis de energia livre corroboram o modelo em duas etapas, introduzido na Seção 10.1.1, para o desenovelamento induzido por força.

O segundo ponto de inflexão ocorre aproximadamente na distância de ponta a ponta de 25,0 nm. Observando as curvas de força e extensão, Figura 33 (página 84), bem

como as estruturas representativas ao longo das CFE para ambos os sistemas, podemos ver que este segundo ponto de inflexão nas curvas de  $\Delta F$  ocorre no mesmo ponto que ocorre o último contato entre as duas cadeias do aptâmero. Após este ponto, não há mais interação entre as cadeias, de forma que o taxa decrescimento da  $\Delta F$  em função da distância de ponta a ponta é muito pequena, porém não é nula. O aumento da energia livre após o último contato entre as cadeias se deve à natureza de não-equilíbrio das simulações. Uma vez que as cadeias são continuamente afastadas, há sempre trabalho sendo realizado sobre o sistema e, portanto, a energia livre não pode ser constante, como aconteceria em uma simulação em equilíbrio.

### 10.2.2 Diferença de energia livre entre os sistemas *aptâmero* e *complexo-estreptomicina*

A Seção 10.2.1 tratou da forma das curvas de energia livre obtidas para as simulações de desenovelamento do aptâmero de RNA induzida por força para os sistemas *aptâmero* e *complexo-estreptomicina*. Nesta seção será tratada a diferença destes perfis entre os sistemas e também a diferença de energia livre entre os dois sistemas no ponto de ruptura.

Ao compararmos as curvas de energia livre da Figura 42 (página 94), podemos ver que o sistema *complexo-estreptomicina* demanda mais energia para desenovelar que o sistema *aptâmero*, ou seja, para cada valor de distância de ponta a ponta, o valor de  $\Delta F$  para o sistema *complexo-estreptomicina* é mais positivo que o valor de  $\Delta F$  para o sistema *aptâmero*. Esta diferença de energia livre é claramente pronunciada nas regiões I e II e, em menor intensidade na região III. Uma vez que a única diferença entre os dois sistemas é a presença da molécula de estreptomicina, se pode inferir que a molécula de estreptomicina contribui significativamente para a estabilização do estado enovelado da molécula de RNA.

Conforme mencionado na Seção 10.2.1, as curvas de energia livre dos sistemas *aptâmero* e *complexo-estreptomicina*, Figura 42 (página 94), apresentam um comportamento linear em relação a distância de ponta a ponta, com dois pontos de inflexão em aproximadamente 8,5 nm e 25,0 nm. A Figura 43 mostra as curvas de energia livre em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-estreptomicina* e as três curvas de  $\Delta F$  lineares em dpp entre os pontos de inflexão descritos na Seção 10.2.1 e a Tabela 4 mostra os coeficientes obtidos para as curvas de  $\Delta F$  lineares em dpp da Figura 43.

Analisando as curvas na Figura 43, se pode ver que na região I, a energia livre, em função da distância de ponta a ponta, para o sistema *complexo-estreptomicina* cresce a uma maior taxa do que para o sistema *aptâmero*. De fato, a extrapolação linear, reta

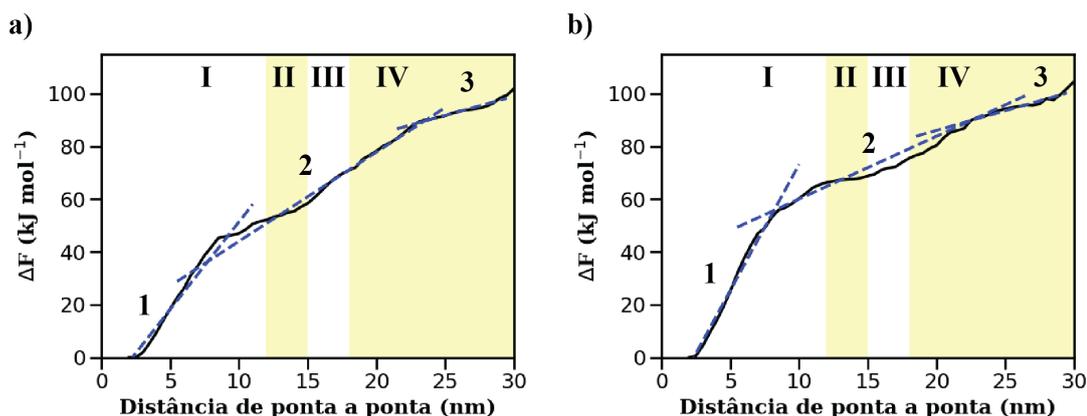


Figura 43 – Extrapolação linear de energia livre  $\Delta F$  (linha tracejada azul), em função da distância de ponta a ponta, para os sistemas a) *aptâmero* e b) *complexo-estreptomicina*.

Tabela 4 – Coeficientes para as extrapolações lineares  $\Delta F = mx + b$  da Figura 43.

Equação	(m*; b†)	
	<i>aptâmero</i>	<i>complexo-estreptomicina</i>
1	(6,63; -14,86)	(9,55; -22,26)
2	(3,38; 10,25)	(2,38; 36,30)
3	(1,45; 55,58)	(1,49; 56,26)

\* Coeficiente angular em  $\text{kJ mol}^{-1} \text{nm}^{-1}$ .

† Constante em  $\text{kJ mol}^{-1}$ .

1 na Figura 43, da região I leva a um coeficiente angular igual a  $6,63 \text{ kJ mol}^{-1} \text{nm}^{-1}$  e  $9,55 \text{ kJ mol}^{-1} \text{nm}^{-1}$  para os sistemas *aptâmero* e *complexo-estreptomicina*, respectivamente. Esta maior taxa de crescimento resulta em valores de energia livre no primeiro ponto de inflexão iguais a  $45,30 \text{ kJ mol}^{-1}$  e  $56,00 \text{ kJ mol}^{-1}$  para os sistemas *aptâmero* e *complexo-estreptomicina*, respectivamente.

Segundo o mecanismo estrutural para o desenovelamento descrito na Seção 10.1, nas regiões I e II a estrutura terciária é desfeita, este aumento de aproximadamente 44 % na taxa de crescimento de  $\Delta F$ , na região I, para o sistema *complexo-estreptomicina* em relação ao sistema *aptâmero* e a diferença de  $10,70 \text{ kJ mol}^{-1}$  no valor de  $\Delta F$  no primeiro ponto de inflexão, indicam que se deve fornecer mais energia para que a estrutura terciária do aptâmero realize a transição da sua conformação nativa para a conformação de escada. Desta forma, esta maior taxa de crescimento para a energia livre do sistema *complexo-estreptomicina* na região I é reflexo da importância da presença da molécula de estreptomicina para a conservação da estrutura terciária do aptâmero, contribuindo de forma a tornar o estado de equilíbrio do sistema *complexo-estreptomicina* mais favorável energeticamente do que o do sistema *aptâmero*.

Esta diferença significativa tanto no aumento da energia livre, como na taxa de

aumento de  $\Delta F$  com a distância de ponta a ponta, na região I pode ser explicada pela forma como a molécula de estreptomicina está ligada ao aptâmero. Foi visto no Capítulo 7 que a molécula de estreptomicina realiza ligações de hidrogênio com nucleotídeos de ambas as cadeias do aptâmero, notavelmente com os nucleotídeos U11, G12, U16 e U17 da cadeia 1 e A108, G110 e G111 da cadeia 2. Desta forma, a molécula de estreptomicina serve de amarra entre as duas cadeias, dificultando assim a transição da conformação nativa para a conformação de escada conforme a força é aplicada ao sistema. Esta interação por ligações de hidrogênio é tão significativa que existe até que a distância de ponta a ponta seja de aproximadamente 15,0 nm, conforme mostrado pelo padrão de ligações de hidrogênio entre a estreptomicina e o aptâmero, Figura 37 (página 87).

Para poder comparar os perfis de energia livre, em função da distância de ponta a ponta, entre os dois sistemas, se deve definir a diferença de energia livre, de não-equilíbrio, em função da distância de ponta a ponta:

$$\Delta\Delta F(x) \doteq \Delta F_{\text{aptâmero}}(x) - \Delta F_{\text{complexo-estreptomicina}}(x) \quad (10.2)$$

onde  $x$  é a distância de ponta a ponta. A Figura 44 mostra a diferença de energia livre entre os sistemas *aptâmero* e *complexo-estreptomicina* em função da distância de ponta a ponta. Pode-se ver que esta diferença atinge um ponto de mínimo, ou a maior diferença de energia livre entre os sistemas, entre as regiões I e II e depois esta diferença entre os valores de energia livre decrescem até um valor limite.

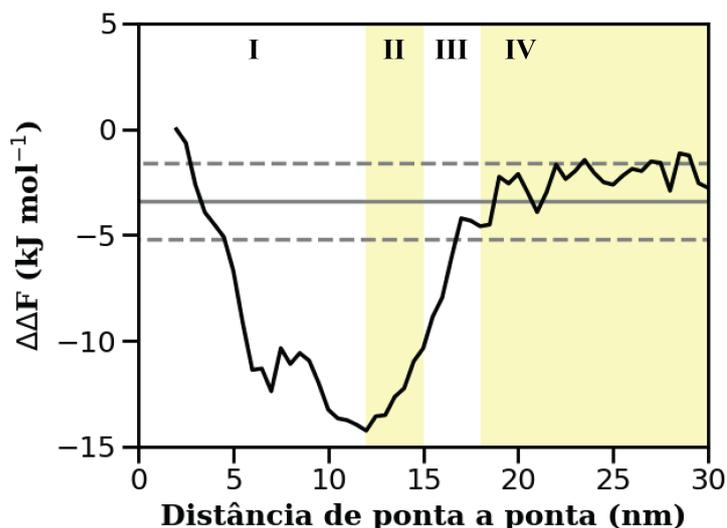


Figura 44 – Diferença de energia livre  $\Delta\Delta F = \Delta F_{\text{aptâmero}} - \Delta F_{\text{complexo}}$  em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-estreptomicina*. A linha sólida cinza representa o valor de  $\Delta\Delta F$  obtido experimentalmente por Nick et al.(35). As linhas tracejadas cinza representam os limites superior e inferior de  $\Delta\Delta F$  obtido experimentalmente, considerando o erro experimental.

Levando em consideração o modelo proposto na seção 10.1 e o fato de que a única diferença entre os dois sistemas é a presença da molécula de estreptomicina, se pode notar que, devido à presença desta molécula no sistema *complexo-estreptomicina*, o processo de desenovelamento induzido por força externa leva a uma maior diferença no valor de  $\Delta\Delta F$  nas regiões que são governadas pela ruptura das interações que mantém a estrutura terciária, ou seja, regiões I e II.

Como explicado anteriormente, podemos atribuir este maior valor de  $\Delta\Delta F$ , nas regiões onde a estrutura terciária sofre maior variação, à característica da molécula de estreptomicina engessar a estrutura terciária do aptâmero ao realizar ligações de hidrogênio com ambas as cadeias. Por conseguinte, uma vez que a estrutura terciária tenha sido desfeita, o valor da diferença de energia livre entre os dois sistemas decresce, já que a estrutura secundária em conformação de escada para ambos os sistemas possuem grande semelhança, diferindo em si apenas pela presença da molécula de estreptomicina. A Tabela 4 nos mostra que a diferença entre os coeficientes angulares para a extrapolação linear para as regiões II e III é de apenas  $1,0 \text{ kJ mol}^{-1} \text{ nm}^{-1}$ , mostrando que a diferença de energia livre entre estes sistemas é linear em distância de ponta a ponta.

Após o ponto de ruptura, ou seja a partir do ponto onde há o segundo máximo na curva de força e extensão, região III, a Figura 44 mostra que a diferença de energia livre entre os sistemas *aptâmero* e *complexo-estreptomicina* flutua em torno de um valor constante. Este comportamento se mantém, então, conforme a distância de ponta a ponta aumenta. Embora os perfís de energia livre para ambos os sistemas, Figura 42, mostrem um aumento de energia livre, mesmo após o último contato, devido a natureza de não-equilíbrio das simulações, a flutuação do valor de  $\Delta\Delta F$  em torno de um valor limite indica que nosso modelo é capaz de eliminar o efeito de não-equilíbrio do trabalho extra realizado sobre o sistema e, assim fornecer a energia livre de equilíbrio.

#### 10.2.2.1 Diferença de energia livre entre os sistemas *aptâmero* e *complexo-estreptomicina* no ponto de ruptura

No trabalho experimental realizado por Nick et al.(35), os autores estabeleceram que a diferença de energia livre de não-equilíbrio entre os sistemas *aptâmero* e *complexo-estreptomicina*,  $\Delta\Delta F$ , obtidas no ponto de ruptura é, por definição, a energia livre de ligação da molécula de estreptomicina ao sítio de ligação do aptâmero. Embora a concordância entre os valores calculados e experimentais de  $\Delta\Delta F$  já tenha sido discutida na Seção 9.2, nesta seção serão abordados alguns dos fatores teóricos que contribuem ou tem influência para esta concordância e como ela pode ser utilizada como um qualificador do modelo de cálculo empregado neste trabalho ou mesmo dar indicativo de como melhorar o modelo.

Existem muitos fatores que podem contribuir para a diferença entre os valores cal-

culados de energia livre daqueles obtidos experimentalmente, como a amostragem correta ou número de simulações(23) e a velocidade de afastamento(142, 143, 144, 145, 146, 147), como mencionada na Seção 9.2. Um destes fatores importantes é o fato já estabelecido de que o estimador de Jarzynski, Equação 10.1, pode levar a desvios indesejados nos valores de energia livre.(154, 153, 155, 156) Embora existam melhores estimadores para a energia livre de não-equilíbrio, como o método de Bennett para o cálculo da média dos valores de trabalho(156), estes métodos, de forma a se obterem melhores estimativas para a energia livre, levam em consideração a amostragem dos caminhos direto e inverso de não-equilíbrio entre dois estados.(153, 155)

Entretanto, devido a natureza dos sistemas de RNA de dupla fita, o caminho inverso, que corresponde à aproximação e enovelamento das duas cadeias de RNA, possui probabilidade muito baixa de ocorrer em uma simulação computacional. De forma a contornar os desvios indesejados nos valores de energia livre causados pelo uso direto do estimador de Jarzynski, Equação 10.1 em conjunto com o uso apenas do caminho direto, foi utilizado neste trabalho o esquema de extrapolação linear proposto por Ytreberg e Zuckerman(39). Neste método a energia livre é plotada em função do número de simulações e, então, é estimada no limite de um conjunto infinito de simulações. Este esquema explora o comportamento sistemático do desvio do estimador de Jarzynski de modo a se obterem melhores estimativas para a energia livre.(39)

Além dos fatores teóricos já discutidos, se deve considerar que o aparato experimental possui elementos adicionais que não são modelados explicitamente no modelo simplificado deste trabalho, como solvente, os *handles* e as esferas de vidro, das quais uma pode ser mantidas fixa e a outra movimentada no microscópio de força atômica. Uma vez que estes elementos são parte inseparável de qualquer experimento de microscopia de força atômica, estes possuem influência na partição de energia dos processos por ele medido.(142) Estes fatores são levados em conta implicitamente no potencial de restrição, utilizado para se manter fixa uma das cadeias, e no potencial harmônico, utilizado para puxar a outra cadeia.

Conforme já discutido na Seção 9.2, o valor de  $\Delta\Delta F$  calculado no ponto de ruptura para a velocidade de afastamento de  $0,04 \text{ nm ps}^{-1}$  é igual a  $-5,26 \pm 0,15 \text{ kJ mol}^{-1}$ . Este valor se encontra dentro do valor experimental de  $-3,4 \pm 1,8 \text{ kJ mol}^{-1}$ , medido por microscopia de força atômica de molécula única(35), sendo que a diferença entre estes dois valores ( $1,86 \text{ kJ mol}^{-1}$ ) é equivalente ao desvio padrão experimental ( $1,8 \text{ kJ mol}^{-1}$ ), de forma que o método empregado para o cálculo da diferença de energia livre é capaz de reproduzir os valores experimentais dentro da acurácia experimental. De fato, todos os valores de  $\Delta\Delta F$  calculados dentro do domínio de velocidades estudado na Seção 9.2, calculados no ponto de ruptura, levam a justa concordância entre experimento e cálculo teórico, sendo que em média os valores preditos diferem do valor obtido experimentalmente

por aproximadamente  $4,0 \text{ kJ mol}^{-1}$ .

Embora se possa discutir que o valor de energia livre calculado neste trabalho, sendo calculado via processos ou simulações de não-equilíbrio, não seja, a rigor, o mesmo valor de energia livre de equilíbrio de ligação da molécula de estreptomicina ao sítio de ligação do aptâmero, o valor aqui calculado corresponde à quantidade medida nos experimentos de microscopia de força atômica de molécula única, ou seja, experimentos em regime de não equilíbrio. De fato, em artigo seminal, Jarzynski(21) relata que sua igualdade possibilita extrair informações de equilíbrio de um conjunto de medidas de não equilíbrio.(21) Por conseguinte, o método de cálculo de energia livre utilizado neste trabalho tem por base a metodologia empregada nos trabalhos experimentais pioneiros de validação da igualdade de Jarzynski.(6, 9, 10, 12, 16, 17, 19, 20)

Desta forma, ao analisarmos o comportamento da diferença de energia livre entre os sistemas *aptâmero* e *complexo-estreptomicina*, em função da distância de ponta a ponta para a velocidade de afastamento  $v = 0,04 \text{ nm ps}^{-1}$ , Figura 44 (página 97), podemos ver que seu valor, após a distância de ponta a ponta correspondente ao ponto de ruptura, flutua dentro dos limites de energia livre obtidos experimentalmente. Desta forma, reafirmam que o modelo de cálculo de energia livre utilizando simulações de não-equilíbrio, empregado nesta tese pode fornecer valores de energia livre de equilíbrio.

### 10.2.3 Decomposição da energia livre em energia interna e entropia

Conforme apresentado na Seção 5.1.1, a decomposição da energia livre em suas componentes de energia interna e entrópica é importante para a compreensão da força motriz que atua no sistema. Desta, nas seções seguintes o desenovelamento dos sistemas *aptâmero* e *complexo-estreptomicina* serão descritas com relação a variação de energia interna,  $\Delta U$ , e variação de entropia,  $\Delta S$ , em função da distância de ponta a ponta.

#### 10.2.3.1 Energia interna

A Figura 45 mostra a diferença de energia interna  $\Delta U$ , de não-equilíbrio, em função da distância de ponta a ponta para o processo de desenovelamento da molécula de RNA induzido por força externa para os sistemas *aptâmero* e *complexo-estreptomicina*. As curvas de  $\Delta U$ , assim como as curvas de força e extensão, possuem um comportamento bifásico com dois valores de máximo.

Se pode ver na Figura 45 que, até aproximadamente o meio da região I, o valor da diferença de energia interna  $\Delta U$  é aproximadamente zero para o sistema *complexo-estreptomicina*. Entretanto, para o sistema *aptâmero*,  $\Delta U$  é negativo nesta primeira parte da região I, ou seja, o sistema está liberando energia. Este fato pode ser explicado devido a formação de ligações de hidrogênio adicionais, conforme explicado na Seção 10.1, devido

a ausência da molécula de estreptomicina neste sistema.

Após este ligeiro rearranjo da estrutura do sistema *aptâmero* na primeira metade da região I, a energia interna cresce em função da distância de ponta a ponta até a fronteira entre as regiões I e II, onde o valor de  $\Delta U$  atinge o primeiro máximo local. Este crescimento do valor de  $\Delta U$  com a dpp até o máximo entre as regiões I e II também é observado para o sistema *complexo-estreptomicina*. Em outras palavras, após o rearranjo inicial das estruturas, ambos os sistemas começam a absorver energia, aumentando sua energia interna. Com base na definição de energia interna do início desta seção e na descrição do número total de ligações de hidrogênio, bem como do número de pares de base e multipletos, da Seção 10.1, se pode atribuir este aumento no valor de  $\Delta U$  à quebra das ligações de hidrogênio intra fitas, necessárias à manutenção da estrutura terciária.

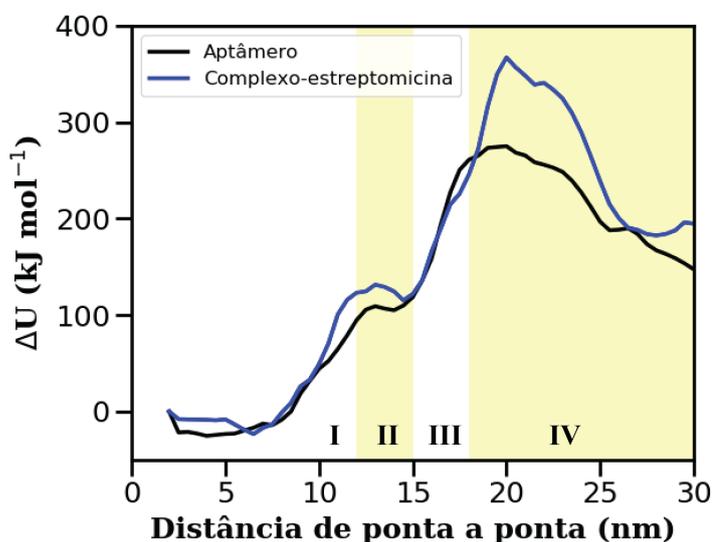


Figura 45 – Diferença de energia interna  $\Delta U$  em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-estreptomicina*.

Os parâmetros estruturais calculados na Seção 10.1 nos levaram a inferir que, até a região II a estrutura terciária em torno do sítio de ligação, presente na conformação inicial, é desfeita e que a molécula de estreptomicina possui um papel fundamental na manutenção da estrutura terciária deste sítio de ligação, uma vez que esta molécula serve de amarra entre as duas cadeias e evita o desenovelamento. Comparando o valor da energia interna entre os sistemas *aptâmero* e *complexo-estreptomicina* entre as regiões I e II da Figura 45, região esta que compreende o final do desenovelamento do sítio de ligação no modelo proposto, se pode ver que o valor de  $\Delta U$  é mais positivo para o sistema *complexo-estreptomicina*. Em outras palavras, o sistemas *complexo-estreptomicina* necessita absorver mais energia para romper as interações internas e chegar neste estado do processo de desenovelamento. Esta energia adicional se pode atribuir à energia necessária

para romper as interações via ligação de hidrogênio que a molécula de estreptomicina realiza com ambas as cadeias.

Após as interações que compõem a estrutura do sítio de ligação serem rompidas e a estrutura, pelo efeito da aplicação de força externa, se encontrar na forma de escada, pode-se ver, na região III da Figura 45 que a energia interna de ambos os sistemas crescem linearmente, conforme as ligações de hidrogênio entre fitas são desfeitas, até que os sistemas alcancem o ponto de ruptura, fronteira entre as regiões III e IV, onde o valor de  $\Delta U$  é aproximadamente igual para os sistemas *aptâmero* e *complexo-estreptomicina*.

Após o ponto de ruptura, a Figura 45 mostra uma grande diferença de energia interna entre os sistemas. A energia interna, Equação 5.29, assim como a energia livre, Equação 10.1, são calculadas pela média de *ensemble* de trajetórias de não-equilíbrio. A presença da molécula de estreptomicina muda de forma significativa a forma como a molécula do aptâmero de RNA se apresenta no estado de último contato entre as cadeias. Para o sistema *aptâmero*, no ponto de último contato há apenas o estado onde os nucleotídeos das extremidades das cadeias fazem contato entre si. Porém, para o sistema *complexo-estreptomicina* existem três possibilidades de estado no ponto de ruptura: A molécula de estreptomicina pode ficar ligada mais ao centro da cadeia fixa; A molécula de estreptomicina pode ficar ligada mais ao centro da cadeia móvel; Ou a molécula de estreptomicina pode rolar entre as cadeias e ficar entre as fitas, agindo como uma ponte, no ponto de último contato, como mostrado na Figura 46

A diferença de energia interna entre os sistemas observada no início da região IV advém do fato da molécula de estreptomicina se mover ao longo das cadeias, entre o ponto de ruptura e o ponto de último contato entre as cadeias, faz com que ligações de hidrogênio, e outras interações não ligantes, sejam desfeitas e refeitas. Este processo faz com que o sistema *complexo-estreptomicina* necessite absorver mais energia do que o sistema *aptâmero* para alcançar o ponto de último contato.

Pode-se ver na Figura 45 que ambas as curvas convergem para aproximadamente o mesmo valor de  $\Delta U$  após o ponto de último contato, onde a energia interna do *complexo-estreptomicina* sendo mais positiva do que para o sistema *aptâmero*.

### 10.2.3.2 Entropia do sistema

Nesta seção será tratada a componente entrópica da energia livre, conforme a Equação 5.32 da Seção 5.1.1.2.

A Figura 47 mostra a entropia  $\Delta S$  em função da distância de ponta a ponta, calculadas pela Equação 5.32, para os sistemas *aptâmero* e *complexo-estreptomicina*. Se pode ver que as curvas de  $\Delta S$  apresentam dois pontos de máximo locais, como as curvas de força e extensão (Figura 33, página 84) e de energia interna (Figura 45, página 101).

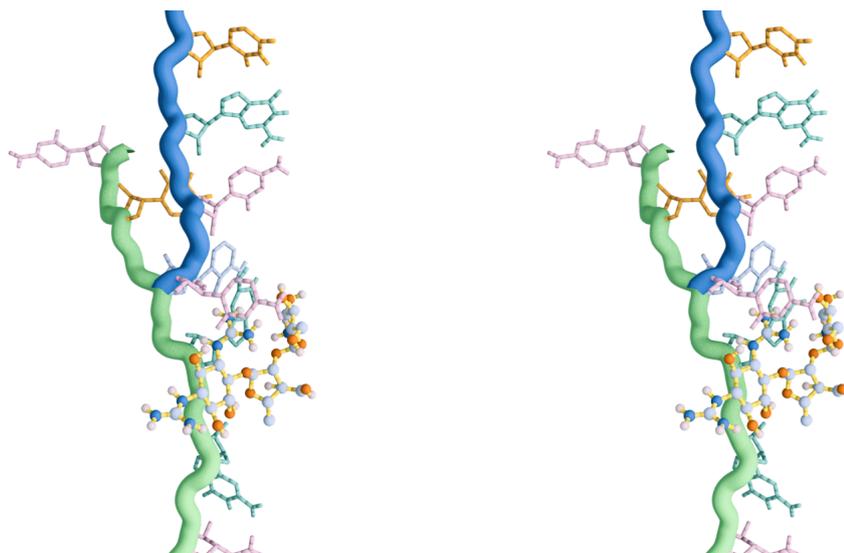


Figura 46 – Modelo estereo da estrutura do sistema *complexo-estreptomicina*, no ponto de último contato entre as fitas, como molécula de estreptomicina servindo como ponte entre ambas as fitas. Cadeia móvel representada em azul e cadeia fixa representada em verde. Nucleotídeos representados na forma licorice: Adenina em azul; Guanina em verde; Citosina em magenta; e Uracil em laranja. Molécula de estreptomicina representada no modelo CPK com ligações em amarelo.

Entretanto, na região I, as curvas de  $\Delta S$  em função da dpp apresentam um mínimo local bem pronunciado.

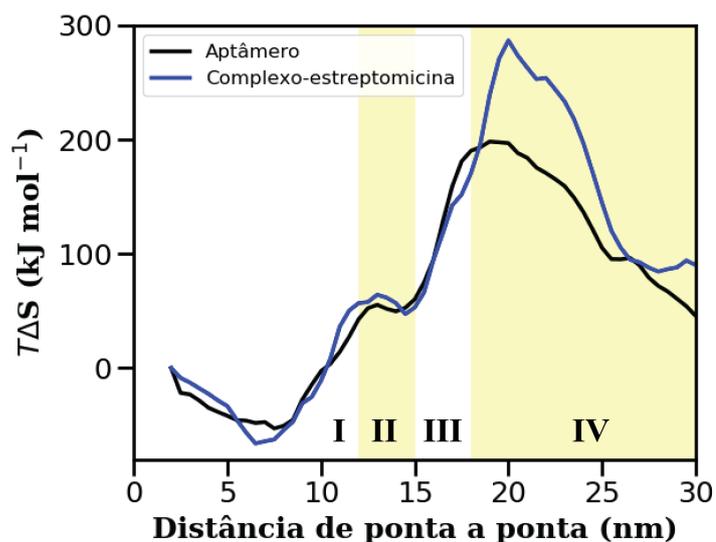


Figura 47 – Produto entre a temperatura e diferença de entropia  $T\Delta S$  em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-estreptomicina*.

Analisando a região I das curvas de entropia, se pode ver que, conforme a distância de ponta a ponta aumenta da distância inicial até aproximadamente 7,0 nm o valor de

$\Delta S$  decresce. Esta redução no valor de  $\Delta S$  pode ser atribuída as restrições de movimento impostas pelas ligações de hidrogênio, tanto intra como entre cadeias, à estrutura da molécula do aptâmero de RNA. Racionalizando a relação entre entropia e restrição de movimento, quanto menor forem as restrições no movimento de partículas independentes e ou na direção do movimento, maior será o valor da entropia. (160, 161, 158) Embora a estrutura do aptâmero de RNA esteja sendo alongada a região I, as interações que mantêm a estrutura terciária coesa ainda estão presentes, conforme se pode ver no comportamento do número total de ligações de hidrogênio na Figura 40a. Com o aumento da distância de ponta a ponta na região I, a estrutura se torna mais tensionada e com isso a amplitude dos movimentos dos grupos atômicos se tornam menores e, com isso, o número de estados acessíveis ao sistema se torna menor, têm-se assim também uma redução da entropia em relação ao sistema inicial.

Se pode ver que na região I, o valor da entropia é mais negativo no ponto de mínimo local para o sistema *complexo-estreptomicina* em relação ao sistema *aptâmero*. Esta diferença reflete o papel que a molécula de estreptomicina desempenha em manter a estrutura terciária da molécula de aptâmero de RNA coesa. Ou seja, conforme o aptâmero é alongado como consequência da aplicação de força externa na região I, a presença da molécula de estreptomicina no sistema, devido a maior restrição de movimento, reduz o número de estados acessíveis ao sistema em relação ao sistema onde esta molécula está ausente.

Uma vez que a força externa aplicada ao sistema forneça energia suficiente para que se rompam as interações que estão mantendo coesa a estrutura terciária, as restrições que faziam com que a entropia descesse com a dpp são liberadas, o número de estados acessíveis ao sistema volta a crescer, há a dispersão da energia nestes níveis e, consequentemente, o valor de  $\Delta S$  torna a crescer.

Entre as regiões I e II, a Figura 47 mostra que a entropia possui um ponto de máximo local. Ponto este que coincide com o desenovelamento da estrutura terciária do sítio de ligação. Podemos ver que neste ponto, o valor de  $\Delta S$  é levemente mais positivo para o sistema *complexo-aptâmero* em relação ao sistema *aptâmero*. Com o desenovelamento do sítio de ligação, mesmo que parcial (conforme Figura 35b), a molécula de estreptomicina é liberada de algumas restrições de movimento existentes, devido a sua interação por ligações de hidrogênio com as cadeias do aptâmero de RNA e, com isso é disponibilizado ao sistema *complexo-estreptomicina* acesso a níveis de energia adicionais em relação ao sistema *aptâmero*. Desta forma, entre as regiões I e II, há este leve aumento do valor de  $\Delta S$  para o sistema *complexo-estreptomicina* quando comparado ao sistema *aptâmero*.

O segundo ponto de máximo nas curvas de entropia, Figura 47, ocorre logo após a fronteira entre as regiões III e IV. Assim como foi constatado para a energia interna na Figura 45 (Página 101), existe uma grande diferença no valor de  $\Delta S$  entre os sistemas

aqui estudados. Como mencionado na Seção 10.2.3.1, no ponto de último contato, para o sistema *complexo-estreptomicina*, a molécula de estreptomicina pode estar presente em uma destas três possibilidades:

- Ficar ligada ao longo da cadeia fixa;
- Ficar ligada ao longo cadeia móvel;
- Ficar ligada entre as cadeias móvel e fixa, formando uma ponte.

Aliada ao aumento do número de estados finais acessíveis para o sistema *complexo-estreptomicina* em relação ao sistema *aptâmero*, ainda devemos notar que a molécula de estreptomicina tem a liberdade de se encontrar em qualquer ponto ao longo tanto da cadeia móvel como da cadeia fixa. Desta forma, podemos inferir que o sistema *complexo-estreptomicina* possui, de fato, um maior número de estados acessíveis no ponto de último contato. Consequentemente, a entropia medida neste ponto é mais positiva em relação a estrutura inicial para o sistema *complexo-estreptomicina* do que para o sistema *aptâmero*.

# 11 Desenovelamento do sistema *complexo-bluensomicina*

Neste capítulo estão descritos os resultados das simulações por dinâmica molecular dirigida (DMD), aplicada no estudo do processo de desenovelamento induzido por força do complexo formado pelo aptâmero de RNA e a molécula de bluensomicina. Uma vez que o aptâmero de RNA estudado neste trabalho possui afinidade à molécula de estreptomicina e baixa afinidade a molécula de bluensomicina(36, 37, 34), os resultados das simulações por dinâmica molecular dirigida aqui descritos serão comparados aos das simulações descritas no Capítulo 10.

As simulações de desenovelamento do complexo entre o aptâmero de RNA e a molécula de bluensomicina foram realizadas conforme a Seção 6.1.4 e os parâmetros obtidos nos Capítulos 8 e 9. Foram realizadas 500 simulações de DMD de forma a emular o desenovelamento induzido por força da molécula de RNA. O capítulo está dividido em duas seções: Uma primeira seção, com a descrição das curvas de força-extensão obtidas durante as simulações de desenovelamento do RNA, análise estrutural deste processo e comparação destas com os resultados obtidos no Capítulo 10; e uma seção com a análise da energia livre do processo de desenovelamento do RNA, bem como sua comparação com os respectivos valores de energia livre para os sistemas *aptâmero* e *complexo-estreptomicina*.

## 11.1 Análise estrutural

Nesta seção será discutido o processo de desenovelamento induzido por força do sistema *complexo-bluensomicina*, simulados por dinâmica molecular dirigida e sua comparação com os sistemas *aptâmero* e *complexo-estreptomicina*, discutidos no Capítulo 10.

A Figura 48 mostra a curva média de força de extensão (CFE), em função da distância de ponta a ponta (dpp) para os sistemas *aptâmero* e *complexo-bluensomicina*. Se pode ver que a CFE para o sistema *complexo-bluensomicina* possui qualitativamente o mesmo comportamento com a distância de ponta a ponta que as CFE para os sistemas *aptâmero* e *complexo-estreptomicina* (Figura 33, página 84). A CFE para o sistema *complexo-bluensomicina* apresenta dois *force jumps*. Desta forma, se pode dividir as CFE em quatro regiões distintas, indicadas pelos numerais romanos, I a IV, na Figura 48. Estas regiões foram descritas no o Capítulo 10 e com mais detalhes na Seção 10.1.1.

Os três sistemas, *aptâmero*, *complexo-estreptomicina* e *complexo-bluensomicina*, apresentam o mesmo comportamento na região correspondente ao ponto de ruptura, ou seja entre as regiões III e IV, com diferença apenas na intensidade do valor da força de

extensão no ponto de máximo.

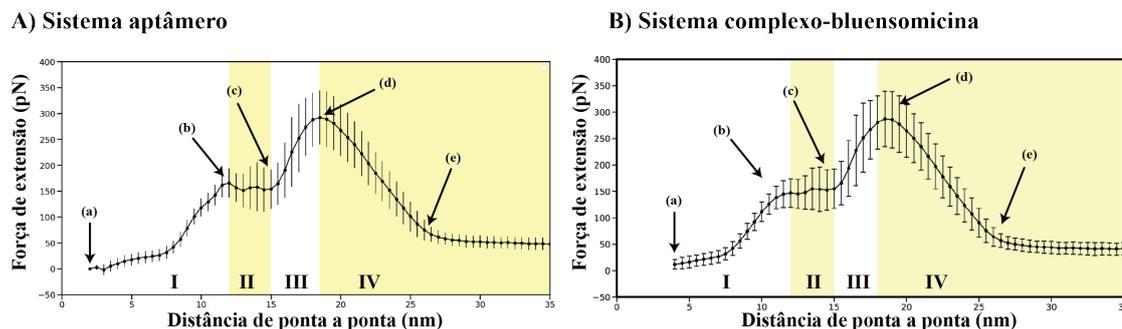


Figura 48 – Curvas de força-extensão médias para os sistemas A) *aptâmero* e B) *complexo-bluensomicina*. As médias foram calculadas sobre 500 simulações.

Ao compararmos as CFE médias entre as regiões I e II para os sistemas *aptâmero* e *complexo-estreptomicina*, Figura 33, se pode notar a presença de um máximo local. Este sendo mais pronunciado para o sistema *complexo-estreptomicina* do que para o sistema *aptâmero*. Entretanto, na CFE média para o sistema *complexo-bluensomicina* este máximo local não está presente, existindo apenas uma inflexão na curva de força e extensão média. Uma vez que na seção 10.1.1, sobre o modelo para o desenovelamento induzido por força de um aptâmero de RNA, esta região foi identificada como a região onde acontece o desenovelamento do sítio de ligação, o comportamento da CFE do sistema *complexo-bluensomicina* nesta região pode indicar que, neste sistema, a desestruturação do sítio de ligação é menos impedida, por restrições internas à molécula, neste sistema.

A Figura 49 mostra as estruturas representativas do sistema *complexo-bluensomicina* ao longo da curva de força e extensão média nos pontos indicados com as letras (a) a (e) na Figura 48B. Se pode ver que o desenovelamento do sistema *complexo-bluensomicina* segue qualitativamente o mesmo padrão de desenovelamento estrutural apresentado pelos sistemas *aptâmero* (Figura 34, página 85) e *complexo-estreptomicina* (Figura 35, página 85)

A Figura 49b mostra a estrutura representativa para o sistema *complexo-bluensomicina* ao final da região I da CFE, Figura 48B. Se pode ver que a estrutura possui um expressivo alongamento, paralelo a direção de afastamento da coordenada de reação, em relação a estrutura representativa inicial, Figura 49a. Entretanto, ao final da região I, assim como ocorreu para o sistema *complexo-estreptomicina*, o centro da estrutura permanece presente envolvendo a molécula de bluensomicina. Utilizando o número de pares de base, Figura 50, como um indicativo da estrutura secundária do aptâmero e o número de multipletos, Figura 51, e número de hélices, Figura 52, como indicativo da estrutura terciária, como foi utilizado na Seção 10.1, se pode ver que as estruturas secundária e terciária se mantêm aproximadamente constantes até o sistema evoluir para uma distância

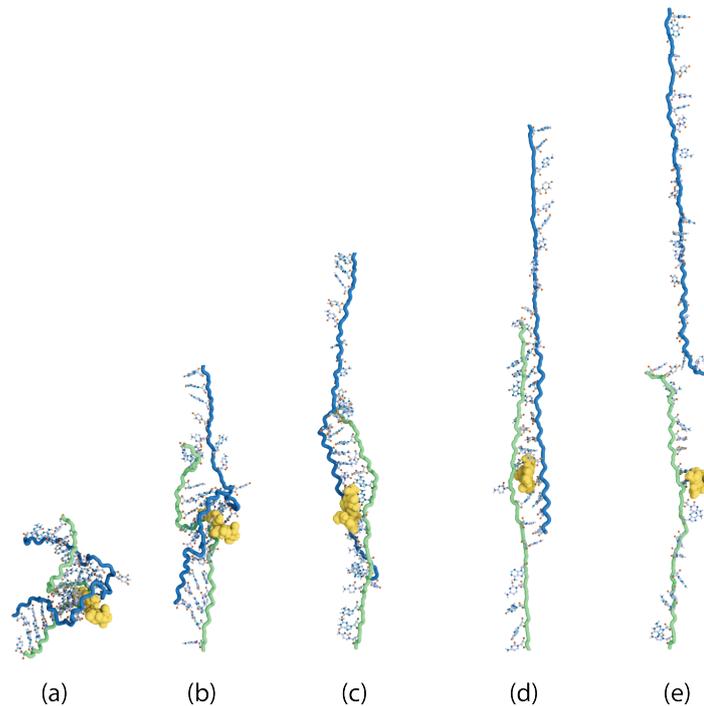


Figura 49 – Estruturas representativas ao longo da CFE para o sistema *complexo-bluensomicina* inicial. b) Estrutura ao final da região I. c) Estrutura ao final da região II. d) Estrutura no ponto de ruptura. e) Estrutura no último ponto de contato.

de ponta a ponta de aproximadamente 10,0 nm.

De fato, realizando a análise combinada das estruturas representativas ao longo da CFE com o número de pares de base, Figura 50, número de multipletos, Figura 51, e número de hélices, Figura 52, para o sistema *complexo-bluensomicina*, podemos concluir que este sistema é primeiramente alongado, sob a atuação da força externa, até que organização estrutural do sítio de ligação seja desfeita. Após esta quebra no enovelamento de equilíbrio, as Figuras 49c-d nos levam a concluir que a estrutura assume uma conformação de escada e, então, a estrutura secundária é desfeita conforme uma cadeia desliza ao longo da outra. Este comportamento é o esperado para um sistema que desenovela de acordo com o modelo de desenovelamento, induzido por aplicação de força externa, proposto na Seção 10.1.1.

Uma vez que a análise da curva de força e extensão, bem como das estruturas representativas ao longo desta para o sistema *complexo-bluensomicina* indicam que este sistema segue aproximadamente o mesmo modelo hierárquico proposto na Seção 10.1.1, é interessante comparar a sua estrutura representativa ao final da região II para este sistema com a estrutura representativa do sistema *complexo-estreptomycinina*, Figura 53. Em

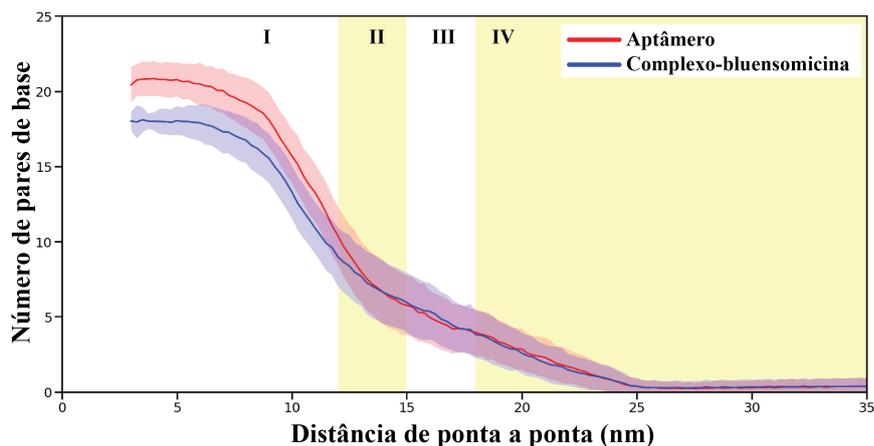


Figura 50 – Número de pares de base em função da dpp para os sistemas *aptâmero* e *complexo-bluensomicina*.

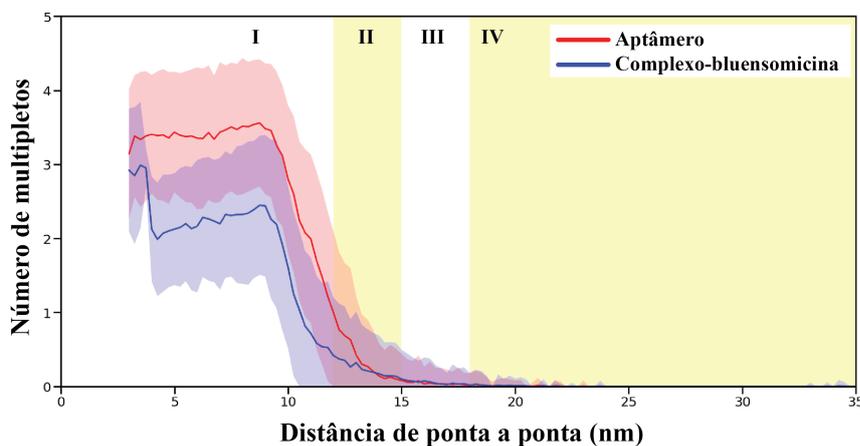


Figura 51 – Número de multipletos em função da dpp para os sistemas *aptâmero* e *complexo-bluensomicina*.

ambos os sistemas, a estrutura se apresenta em uma forma alongada com os pares de base nas extremidades das cadeias desfeitos. Para o sistema *complexo-estreptomycina* o sítio de ligação ainda se encontra parcialmente estruturado em torno da molécula de estreptomycina, ou seja de sua molécula ligante. Entretanto, o mesmo não ocorre para o sistema *complexo-bluensomicina*, em que ambas as fitas do aptâmero se encontram desenoveladas em torno da molécula de bluensomicina. Este fato demonstra o maior efeito da estabilização da organização estrutural que a molécula de estreptomycina causa à estrutura do aptâmero em relação ao efeito causado pela molécula de bluensomicina.

O comportamento descrito no parágrafo anterior infere, qualitativamente, que a molécula de estreptomycina confere maior estabilidade à estrutura terciária do aptâmero do que a molécula de bluensomicina o faz. Este efeito estabilizador ao centro da estrutura do aptâmero se mostra evidente ao compararmos a região II na CFE, Figuras 33B

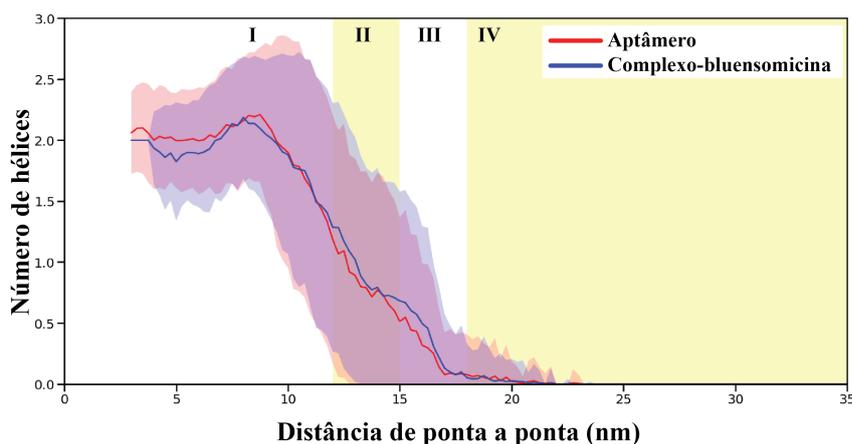


Figura 52 – Número médio de hélices em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-bluensomicina*.

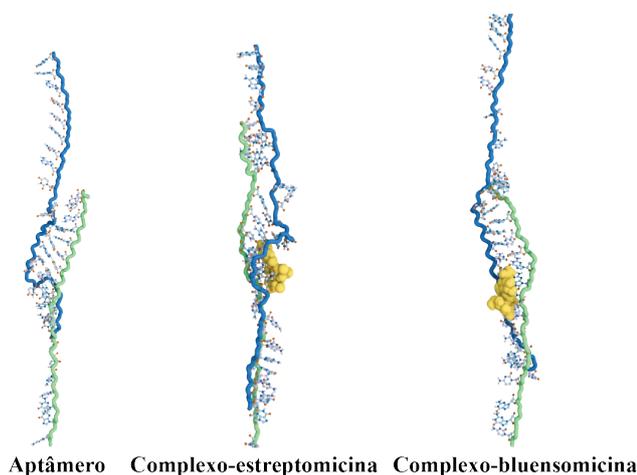


Figura 53 – Estruturas representativas ao final da região II para os sistemas *aptâmero* (esquerda), *complexo-estreptomicina* (centro) e *complexo-bluensomicina* (direita).

(página 84) e 48B (página 107), para os sistemas *complexo-estreptomicina* e *complexo-bluensomicina*, respectivamente. Nesta região, podemos notar que após a primeira descontinuidade na CFE, a força possui um comportamento definido para a sistema *complexo-estreptomicina*, como esperado para um sistema mais estável, do que para o sistema *complexo-bluensomicina*, que apresenta apenas uma descontinuidade, não um pico definido.

Do ponto de vista da estrutura química, a molécula de bluensomicina difere da molécula de estreptomicina em dois pontos, Figura 54. 1) Um dos grupos guanidino para-substituídos no anel inositol na estreptomicina é substituído por um grupo carbamido na bluensomicina. Ou seja, onde há dois grupos doadores de ligação de hidrogênio na

estreptomicina, há dois grupos aceptores de ligação de hidrogênio na bluensomicina. 2) O grupo aldeído no anel metil-pentose da estreptomicina é substituído por um grupo hidroximetil na bluensomicina. Agora, temos a substituição de um grupo aceptor de ligação de hidrogênio na estreptomicina por um grupo doador na bluensomicina. Desta forma, como discutido no Capítulo 7, a molécula de bluensomicina possui uma interação mais fraca com o sítio de ligação do aptâmero do que a molécula de estreptomicina, principalmente com os resíduos da 110 e 111 da cadeia 2.(34)

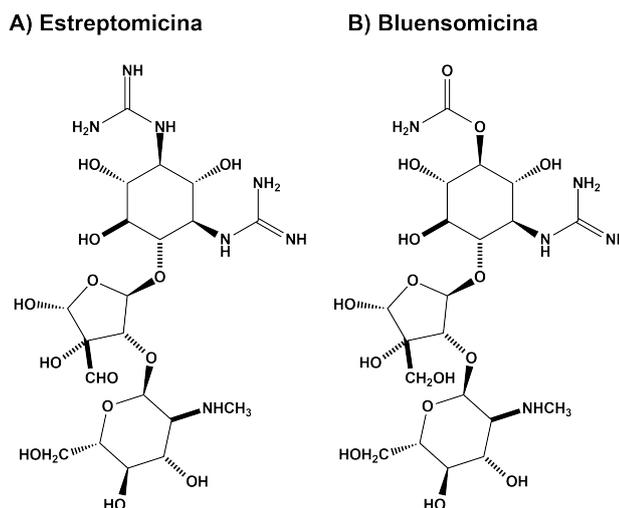


Figura 54 – Estrutura dos ligantes (A) estreptomicina e (B) Bluensomicina.

Para obtermos uma estimativa qualitativa da estabilização do sítio ativo pela molécula de bluensomicina, foi realizado o mapeamento dos nucleotídeos do aptâmero com os quais a molécula ligante está realizando ligações de hidrogênio em função da distância de ponta a ponta. A Figura 55 mostra a frequência, por nucleotídeo, de ligações de hidrogênio com a molécula de bluensomicina como função da distância de ponta a ponta. Se pode ver que a bluensomicina realiza ligações de hidrogênio com alguns nucleotídeos que compõem o sítio de ligação, assim como a molécula de estreptomicina o faz. Entretanto, se deve notar que a frequência, dentro do conjunto de simulações realizadas, é muito inferior ao registrado para a molécula de estreptomicina (Figura 37, página 87). De fato, podemos notar que a frequência de interação da bluensomicina com os nucleotídeos da cadeia 2 não são tão intensos como ocorre no sistema onde a molécula ligante é a estreptomicina. Ao compararmos os dois padrões de frequência, podemos notar, também, que as interações significativas (interação maior que 80 %) por ligações de hidrogênio entre a bluensomicina e as cadeias do aptâmero terminam com uma distância de ponta a ponta de aproximadamente 10,0 nm, enquanto que para o sistema onde a molécula ligante é a estreptomicina, existem interações significativas até aproximadamente a distância de ponta a ponta de 15,0 nm.

Os resultados apresentados no parágrafo anterior demonstram que o sítio de liga-

ção em torno da molécula de bluensomicina desenovela mais facilmente do que ao redor da molécula de estreptomicina. Este fato está de acordo com o observado na estrutura cristalográfica obtida experimentalmente por Tereshko, Skripkin e Patel(34), uma vez que há a substituição de grupos doadores de ligação de hidrogênio por grupos aceptores, fazendo com que a bluensomicina não possua uma interação de ligações de hidrogênio considerável como as guaninas 110 e 111, da cadeia 2, podendo ser este o motivo da grande diferença de afinidade destas moléculas com a molécula de aptâmero de RNA.(37)

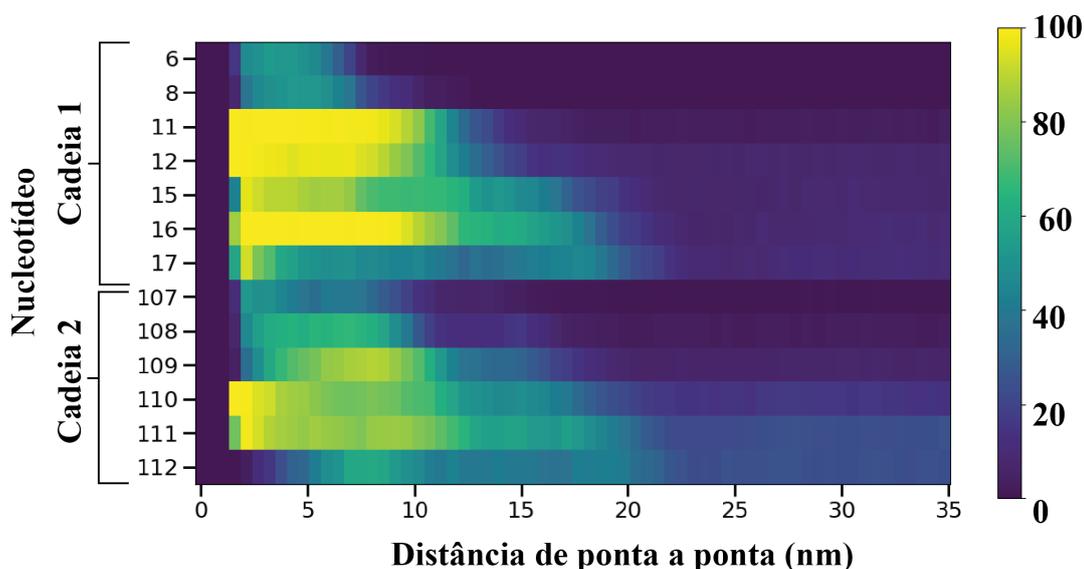


Figura 55 – Padrão de ligações de hidrogênio (%) entre a molécula de bluensomicina e aptâmero no sistema *complexo-bluensomicina*.

De fato, a substituição de um grupo guanidino e aldeído na molécula de estreptomicina pelos grupos carbamido e álcool na molécula de bluensomicina causam um efeito muito grande na estabilidade do padrão de ligações de hidrogênio do aptâmero de RNA. A Figura 56 mostra o número de ligações de hidrogênio em três definições: Total, entre-cadeias e intra-cadeias. A Figura 56a mostra o número total de ligações de hidrogênio. Nela se pode notar que este descritor mantém um valor aproximadamente constante durante quase toda a região I e depois seu comportamento muda, decrescendo conforme a distância de ponta a ponta aumenta. É importante notar que o número total de ligações de hidrogênio no início das simulações de desenovelamento é maior para o sistema *complexo-bluensomicina* do que para o sistema *aptâmero*. Este fato já havia sido observado para o sistema *complexo-estreptomicina*. Uma vez iniciado o processo de desenovelamento, o padrão inverte. Ou seja, o número total de ligações de hidrogênio, durante seu comportamento constante na região I, para o sistema *aptâmero* se torna maior que o número total de ligações de hidrogênio do sistema *complexo-bluensomicina*. Conforme explicado na Seção 10.1 para o sistema *complexo-estreptomicina*, este fato decorre da maior possibilidade

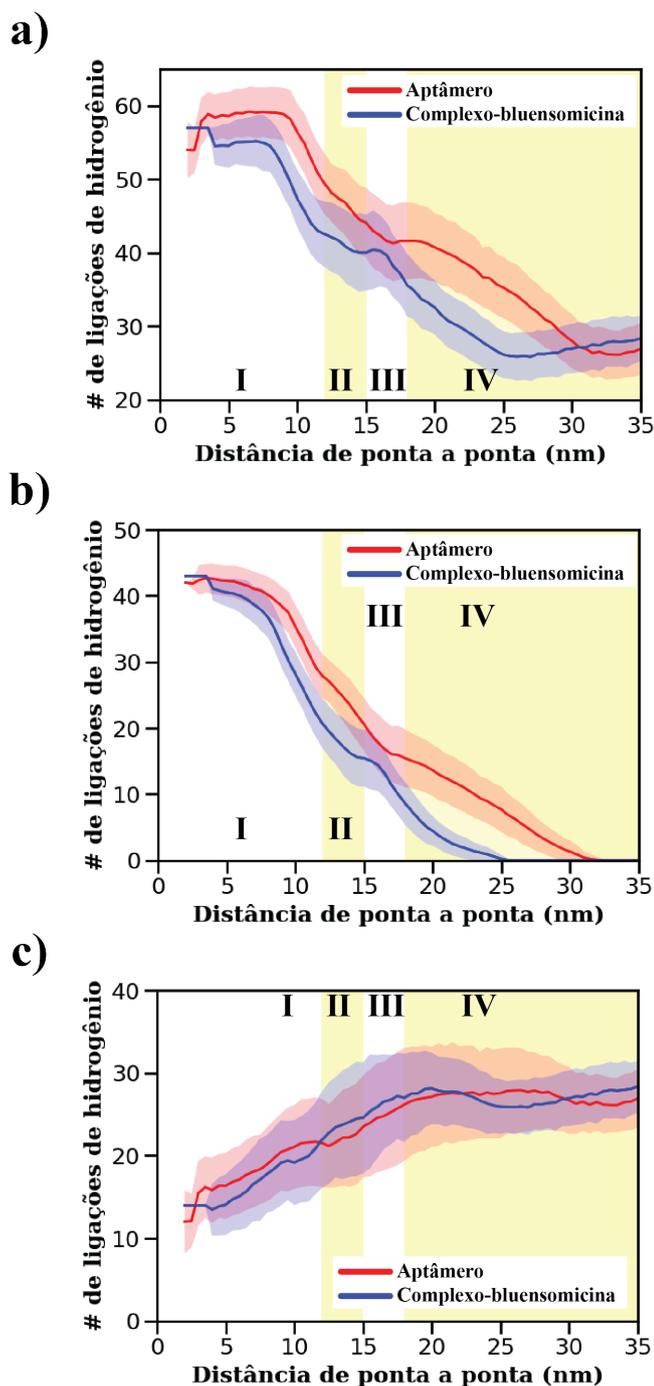


Figura 56 – Número de ligações de hidrogênio em função da dpp para os sistemas *aptâmero* e *complexo-bluensomicina*. a) Número total de ligações de hidrogênio; b) Número de ligações de hidrogênio entre as cadeias 1 e 2; c) Número total de ligações de hidrogênio intra cadeia.

de contato entre as duas cadeias ocorrerem no sistema onde não há a presença de uma molécula ligante.

Deve-se notar que, após a inversão inicial no valor do número total de ligações de hidrogênio ( $HB_{tot}$ ), o valor deste parâmetro para o sistema *complexo-bluensomicina*

é significativamente menor que o valor de  $HB_{tot}$  para o sistema *aptâmero* até aproximadamente a distância de ponta a ponta de 30,0 nm, onde novamente o padrão se inverte. Porém, ao analisarmos este mesmo descritor para o sistema *complexo-estreptomicina* (Figura 40a, página 90), vemos que o valor de  $HB_{tot}$  para o sistema *complexo-estreptomicina* é virtualmente indistinguível do valor de  $HB_{tot}$  para o sistema *aptâmero*. Desta forma, se pode inferir que a presença da molécula de bluensomicina contribui com o processo de desnaturação induzida por força externa do aptâmero.

Ao analisarmos o número de ligações de hidrogênio entre as cadeias do aptâmero ( $HB_e$ ) para o sistema *complexo-bluensomicina*, Figura 56b, a influência da molécula de bluensomicina no processo de desnaturação induzida por força se torna mais evidente. Se pode notar que o número  $HB_e$  para o sistema *complexo-bluensomicina* é sistematicamente menor que para o sistema *aptâmero* durante todo o processo de desenovelamento. É importante notar que, embora esta propriedade decresça para ambos os sistemas, o número de ligações de hidrogênio entre as cadeias do aptâmero para o sistema *complexo-bluensomicina* se torna nulo para uma distância de ponta a ponta menor, em aproximadamente 5,0 nm, do que para o sistema *aptâmero*. Este fato indica que a molécula de bluensomicina contribui para que a estrutura secundária do aptâmero de RNA seja desfeita, e portanto contribui de forma significativa com o processo de desnaturação do aptâmero.

A Figura 56c mostra o número total de ligações de hidrogênio intra fitas ( $HB_i$ ) para os sistemas *aptâmero* e *complexo-bluensomicina*. Se pode ver que o valor de  $HB_i$  não é significativamente diferente entre estes dois sistemas, assim como já se havia descrito entre os sistemas *aptâmero* e *complexo-estreptomicina*. Embora esta interação esteja relacionada com a estrutura terciária da molécula de RNA, ela não é específica como parâmetro indicador de organização. Entretanto, seu aumento com o aumento da distância de ponta a ponta indica que, após serem separadas, as cadeias interagem consigo mesmas e formando ligações de hidrogênio intra cadeias.

Racionalizando os resultados dos parâmetros estruturais apresentados nesta seção, podemos concluir que o desenovelamento do sistema *complexo-bluensomicina* segue o modelo proposto na Seção 10.1.1 e que a molécula de bluensomicina contribui de forma significativa para o processo de desenovelamento do aptâmero de RNA. Os resultados mostram que, uma vez que a molécula de bluensomicina não é capaz de interagir de forma eficiente com a cadeia 2, via ligações de hidrogênio, esta contribui de forma efetiva para que a estrutura secundária do aptâmero seja desfeita conforme a força externa é aplicada ao sistema, promovendo assim o desenovelamento do aptâmero em um valor de dpp menor que o que ocorre nos sistemas *aptâmero* e *complexo-estreptomicina*.

## 11.2 Análise energética

A Seção 11.1 realizou a análise estrutural do processo de desenvolvimento induzido por força externa do sistema *complexo-bluensomicina*, onde foi identificado que a molécula de bluensomicina contribui para que a estrutura secundária seja desfeita. Nesta seção, entretanto, o processo de desenvolvimento induzido por força externa deste sistema será tratado do ponto de vista energético. Esta análise energética será realizada utilizando a Igualdade de Jarzynski e comparada à simulação do sistema *aptâmero*, assim como foi realizada na Seção 10.2.

### 11.2.1 Análise energética do processo de desenvolvimento

Os valores de energia livre, em função da distância de ponta a ponta para o sistema *complexo-bluensomicina*, foram calculados utilizando o método de extrapolação linear proposto por Ytreberg e Zuckerman(39), conforme Seção 6.2, utilizando como estimador da energia livre a igualdade de Jarzynski.(21) Os valores referentes à média da exponencial do trabalho  $\langle e^{-\beta W} \rangle$ , em função da distância de ponta a ponta, foram calculados em relação às 500 simulações de dinâmica molecular dirigida, conforme a metodologia descrita no Capítulo 6, utilizando os parâmetros escolhidos nos Capítulos 8 e 9.

A Figura 57 mostra o valor da energia livre, em função da distância de ponta a ponta, para os sistemas *aptâmero* e *complexo-bluensomicina*. O perfil de energia livre para o sistema *aptâmero* já foi analisado e explicado neste trabalho, na Seção 10.2.1. Portanto, nesta seção este perfil será utilizado apenas como referência. Uma vez que é aplicado uma força externa de forma a forçar a molécula de RNA, ou seu complexo com a molécula ligante, a desenvolver, a energia livre em função da distância de ponta a ponta (dpp) será sempre positiva.

O perfil de energia livre para o sistema *aptâmero*, conforme descrito na Seção 10.2.1, se comporta como uma função composta por funções lineares, ou seja, pode ser dividida em três regiões lineares, divididas por dois pontos de inflexão. O primeiro ponto de inflexão em aproximadamente 8,5 nm e o segundo, menos pronunciado, em 25,0 nm. Entretanto, o perfil de energia livre para o sistema *complexo-bluensomicina* apresenta um comportamento diferente em relação ao apresentado pelo sistema *aptâmero*, sendo o perfil para este sistema diferente, também, do perfil de energia livre para o sistema *complexo-estreptomicina* (Figura 42, página 94), que possui o mesmo comportamento do perfil exibido pelo sistema *aptâmero*. Portanto, não podemos dividir o perfil de energia livre para o sistema *complexo-bluensomicina* nas mesmas três regiões em que foram estratificados as curvas de  $\Delta F$  dos sistemas *aptâmero* e *complexo-estreptomicina*.

A Seção 11.1 determinou que o sistema *complexo-bluensomicina* segue o modelo em duas etapas para o desenvolvimento induzido por força externa da estrutura do RNA,

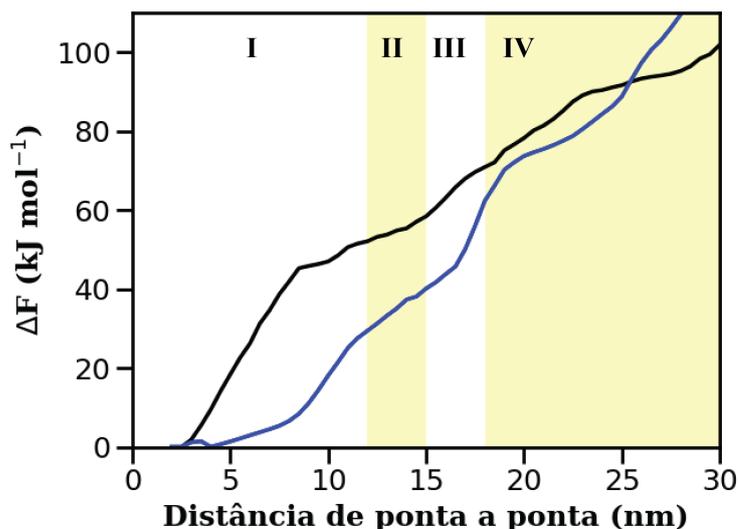


Figura 57 – Energia livre  $\Delta F$  em função da distância de ponta a ponta para os sistemas *aptâmero*, linha preta, e *complexo-bluensomicina*, linha azul.  $v = 0,04 \text{ nm ps}^{-1}$  e  $k = 200,0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ .

proposto na Seção 10.1.1. De acordo com este modelo, a primeira etapa compreende a desorganização da estrutura terciária do RNA. Conforme descrito na Seção 10.2.1, o ponto de inflexão em aproximadamente 8,5 nm corresponde a transição entre a estrutura completamente formada, estrutura nativa, do aptâmero para a estrutura em conformação de escada. Ao observarmos o perfil de energia livre para o sistema *complexo-bluensomicina*, Figura 57, se pode ver que há um ponto de inflexão em aproximadamente 8,5 nm.

Comparando as estruturas (a) e (b) da Figura 49, bem como o número de multipletos (Figura 51, página 109) e o padrão de ligações de hidrogênio entre a molécula de bluensomicina e o aptâmero (Figura 55, página 112), se pode ver que, de fato, este ponto de inflexão, em aproximadamente 8,5 nm, na curva de  $\Delta F$  corresponde a transição entre a estrutura de equilíbrio do aptâmero para a estrutura em conformação de escada. Entretanto, neste ponto a energia livre para o sistema *complexo-bluensomicina* é muito inferior à energia livre para o sistema *aptâmero*, indicando que o sistema, onde há a presença da molécula de bluensomicina, demanda uma menor energia livre para que a estrutura faça a transição da estrutura nativa para a estrutura em forma de escada. Também é visível na Figura 57 que a taxa de crescimento da energia livre para o sistema *aptâmero* é maior que para o sistema *complexo-bluensomicina*.

O modelo em duas etapas para o desenovelamento induzido por força externa da estrutura do RNA, bem como a análise estrutural do desenovelamento dos sistemas *aptâmero*, *complexo-estreptomicina* e *complexo-bluensomicina* nos mostraram que na região I ocorre a desestruturação do sítio de ligação em torno da molécula ligante. Uma vez que a taxa de crescimento da energia livre na região I para o sistema *complexo-bluensomicina*

é menor que para o sistema *aptâmero*, podemos inferir que presença da molécula de bluensomicina no sítio de ligação favorece energeticamente a desestruturação deste.

Segundo o modelo de desenovelamento em duas etapas. após sofrer a transição da estrutura nativa para a estrutura em forma de escada, o sistema começa a desfazer a estrutura secundária e conseqüentemente, a romper as ligações de hidrogênio entre as cadeias de RNA. Os perfis temporais de ligações de hidrogênio entre as cadeias (Figura 56b, página 113) para o sistema *complexo-bluensomicina* deixa claro que este comportamento ocorre neste sistema. Se pode observar no perfil de energia livre para o sistema *complexo-bluensomicina*, Figura 57, que após o sistema ultrapassar a barreira do primeiro ponto de inflexão no perfil de energia livre, a taxa de crescimento da energia livre com a distância de ponta a ponta aumenta. Uma vez que a molécula de bluensomicina não mais interage significativamente com as cadeias 1 e 2 do aptâmero (vide Figura 55, página 112), este comportamento pode ser atribuído à quebra das ligações de hidrogênio entre as cadeias, de forma semelhante ao processo que ocorre no sistema *aptâmero*. De fato, na Seção 10.2.1, o estudo dos perfis de energia livre em função da dpp para os sistemas *aptâmero* e *complexo-estreptomycina* também indicaram que a energia livre do processo durante esta etapa é fortemente correlacionada com a formação e ruptura dos pares de base, e conseqüentemente como o número de ligações de hidrogênio entre as cadeias.

É interessante notar que o número de ligações entre as cadeias (Figura 56b, página 113) é significativamente menor para o sistema *complexo-bluensomicina* do que para o sistema *aptâmero* em um mesmo valor da dpp. Em outras palavras, as ligações de hidrogênio entre as cadeias, e portanto também o número de pares de base, são rompidas mais rápido para sistema *complexo-bluensomicina*. Comparando a curva de  $\Delta F$  para o sistema *complexo-bluensomicina* com o número de ligações de hidrogênio entre as cadeias, se pode ver uma forte correlação, mostrando que a energia livre requerida para o desenovelamento deste sistema é fortemente dependente do número de ligações de hidrogênio entre cadeias. Estes fatos justificam a maior taxa de crescimento de energia livre para o sistema *complexo-bluensomicina*, em relação ao sistema *aptâmero*, no intervalo entre 8,5 nm e 25,0 nm.

De acordo com a Seção 10.2.1, um segundo ponto de inflexão ocorre na curva de energia livre em uma distância de ponta a ponta de aproximadamente 25,0 nm. Este ponto foi identificado como a dpp em que ocorre o último contato entre as duas cadeias do aptâmero. Este fato é corroborado para o sistema *complexo-bluensomicina* pela estrutura representativa da Figura 49e (página 108). Portanto, após este ponto não existe mais interação entre as cadeias, sendo as únicas interações possíveis aquelas intra cadeias e aquelas interações entre a molécula de bluensomicina e a cadeia que continua em contato com esta. Após este ponto, a taxa decréscimo da  $\Delta F$  em função da distância de ponta a ponta para o sistema *aptâmero* é muito pequena, porém não é nula. Por outro lado,

este comportamento não é observado para o sistema *complexo-bluensomicina*. Depois de transpor a dpp correspondente ao ponto de último contato entre as cadeias, o valor de  $\Delta F$  continua a aumentar significativamente para este sistema. Conforme explicado na Seção 10.2.1, o aumento da energia livre após o último contato entre as cadeias se deve à natureza de não-equilíbrio das simulações. Uma vez que há continuamente trabalho sendo realizado sobre o sistema, a energia livre não pode ser constante.

### 11.2.2 Diferença de energia livre entre os sistemas *aptâmero* e *complexo-bluensomicina*

A Seção 11.2.1 comparou a forma dos perfis de energia livre em função da distância de ponta a ponta, para o processo de desenovelamento induzido por força externa, entre os sistemas *aptâmero* e *complexo-bluensomicina*. Dando seguimento a análise energética do processo de desenovelamento induzido por força externa do sistema *complexo-bluensomicina*, esta seção tratará da diferença entre os valores em função da dpp.

Comparando os perfis de energia livre em função da dpp para os sistemas *aptâmero* e *complexo-bluensomicina* da Figura 57 (página 116), se pode notar que o sistema *complexo-bluensomicina* demanda menos energia para desenovelar que o sistema *aptâmero*. Este comportamento só inverte após o ponto de último contato entre as cadeias, em uma dpp de aproximadamente 25,0 nm. Em outras palavras, até a distância de ponta a ponta atingir o marco de 25,0 nm o valor de  $\Delta F$ , em cada valor da dpp, para o sistema *complexo-bluensomicina* é menor do que para o sistema *aptâmero*. Este comportamento é oposto do comportamento apresentado pelo sistema *complexo-estreptomicina* (Figura 42, página 94), em que é necessário mais energia para desnaturar o complexo.

Em analogia a Equação 10.2, definimos a diferença de energia livre, de não-equilíbrio, em função da distância de ponta a ponta:

$$\Delta\Delta F(x) \doteq \Delta F_{\text{aptâmero}}(x) - \Delta F_{\text{complexo-bluensomicina}}(x) \quad (11.1)$$

onde  $x$  é a distância de ponta a ponta. A Figura 58 mostra a diferença de energia livre entre os sistemas *aptâmero* e *complexo-bluensomicina* em função da distância de ponta a ponta. Se pode ver que esta diferença é positiva até a distância de ponta a ponta igual a aproximadamente 25,0 nm, sendo negativa após este ponto. Este comportamento nos indica que o sistema *complexo-bluensomicina* é menos estável que o sistema *aptâmero* para todos os valores de dpp inferiores a 25,0 nm. É interessante notar que, para valores de dpp maiores que 25,0 nm, ou seja para os pontos onde o sistema se encontra desenovelado, o valor de  $\Delta\Delta F$  é negativo, indicando que a presença da molécula de bluensomicina faz com que o estado desenovelado seja energeticamente mais estável que o estado enovelado.

De modo geral, a diferença de energia livre entre os dois sistemas, apresentada na Figura 58, é notavelmente divergente nas regiões I, II e III. Uma vez que a única diferença

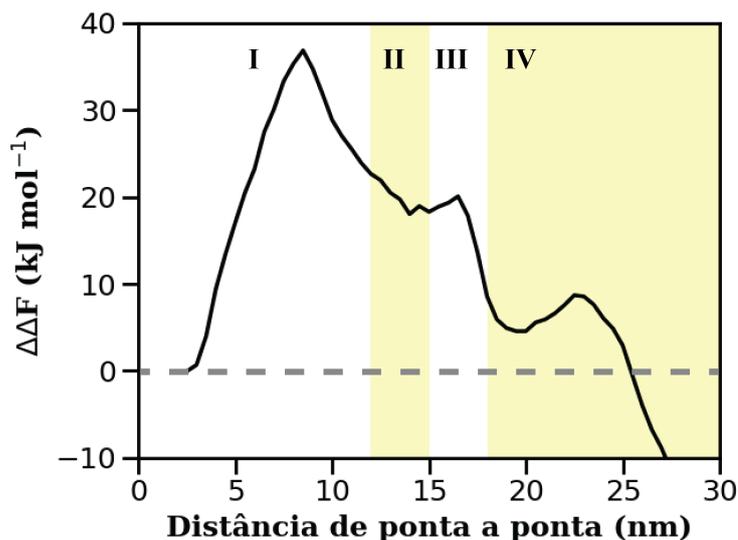


Figura 58 – Diferença de energia livre  $\Delta\Delta F = \Delta F_{\text{aptâmero}} - \Delta F_{\text{complexo}}$  em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-bluensomicina*.

entre estes dois sistemas é a presença da molécula de bluensomicina, a interpretação deste resultado leva a inferir que a molécula de bluensomicina contribui significativamente de forma negativa para a estabilização do estado enovelado da molécula de RNA.

Levando em consideração o modelo de desenovelamento em duas etapas, para o desenovelamento induzido por força externa da estrutura do RNA (Seção 10.1.1), energia livre para os sistemas na região I corresponde à energia necessária para a desorganização da estrutura terciária do RNA. A diferença de energia livre entre os dois sistemas  $\Delta\Delta F$  na região I é positiva e atinge um máximo em uma dpp de aproximadamente 8,5 nm. Uma vez que este ponto está relacionado a transição da estrutura nativa, de equilíbrio, do aptâmero para a estrutura em forma de escada e a desestruturação do sítio de ligação, podemos inferir que a presença da molécula de bluensomicina ligada ao sítio de ligação torna o desenovelamento do sítio de ligação do aptâmero de RNA energeticamente favorável.

Nas Seções 9.2 e 10.2.2 foi utilizada a definição da energia livre de ligação da molécula de estreptomicina ao sítio de ligação dada em trabalho experimental por Nick et al.(35). Analogamente, definimos a energia livre de ligação da molécula de bluensomicina ao sítio de ligação como sendo a diferença de energia livre de não-equilíbrio entre os sistemas *aptâmero* e *complexo-bluensomicina*,  $\Delta\Delta F$ , obtidas no ponto de ruptura.

No ponto de ruptura, o valor para a energia livre de ligação da molécula de bluensomicina ao sítio de ligação  $\Delta\Delta F$  é igual a  $+5,78 \pm 0,34 \text{ kJ mol}^{-1}$ . Desta forma, podemos argumentar que a ligação da molécula de bluensomicina ao sítio de ligação do aptâmero não é, de fato, favorável. O aptâmero de RNA estudado neste trabalho foi desenhado de forma a possuir afinidade à molécula de estreptomicina e baixa afinidade a outras moléculas.(37) Consequentemente, o valor positivo de  $\Delta\Delta F$  para a molécula de bluen-

somicina, aliado ao valor negativo encontrado para o valor de  $\Delta\Delta F$  para a molécula de estreptomicina na Seção 10.2.2.1, estão de acordo com o comportamento seletivo esperado do aptâmero de RNA encontrado experimentalmente por Wallace e Schroeder(37), corroborando, assim, o modelo de simulação empregado neste trabalho e também o modelo de desenovelamento em duas etapas proposto.

### 11.2.3 Comparação das energias livres de ligação das moléculas estreptomicina e bluensomicina ao sítio de ligação do aptâmero

Conforme já mencionado no início do Capítulo 11, o aptâmero estudado nesta tese possui grande afinidade à molécula de estreptomicina, enquanto que, apesar da semelhança de seu arcabouço estrutural, possui uma baixa afinidade à molécula de bluensomicina.(34, 36, 37) De fato, a análise da seletividade *in vitro* deste aptâmero foi estudado por Wallace e Schroeder(37). Esta análise revelou que o aptâmero favorece a ligação da molécula de estreptomicina em relação à ligação da molécula de bluensomicina por quatro ordens de grandeza.(37, 34)

Se assumirmos que a constante de equilíbrio  $K$  de uma determinada reação é proporcional a exponencial da energia livre:

$$K \propto e^{\frac{-\Delta F}{RT}} \quad (11.2)$$

onde  $R = 8,3145 \text{ J K}^{-1} \text{ mol}^{-1}$  é a constante dos gases e  $T$  a temperatura do sistema. Podemos inferir aproximadamente a seletividade como:

$$\frac{K_{\text{STR}}}{K_{\text{BLU}}} \approx e^{\frac{-1}{RT}[\Delta\Delta F_{\text{STR}} - \Delta\Delta F_{\text{BLU}}]} \quad (11.3)$$

Desta forma, utilizando os valores de energia livre de ligação, de não-equilíbrio, destes ligantes ao sítio de ligação  $\Delta\Delta F$  para calcularmos a seletividade. Conforme Seções 10.2.2.1 e 11.2.2, temos  $\Delta\Delta F_{\text{STR}} = -5,26 \text{ kJ mol}^{-1}$  e  $\Delta\Delta F_{\text{BLU}} = +5,78 \text{ kJ mol}^{-1}$  para a ligação das moléculas de estreptomicina e bluensomicina, respectivamente, ao sítio de ligação do aptâmero. Desta forma, utilizando a Equação 11.3 e utilizando a temperatura do sistema  $T = 300 \text{ K}$ , temos:

$$\begin{aligned} \frac{K_{\text{STR}}}{K_{\text{BLU}}} &\approx e^{\frac{-1}{RT}[\Delta\Delta F_{\text{STR}} - \Delta\Delta F_{\text{BLU}}]} \\ &\approx \frac{K_{\text{STR}}}{K_{\text{BLU}}} \propto e^{\frac{-1}{8,3145 \cdot 300}[-5260,0 - 5780,0]} \\ &\approx e^{+4,4260} \\ &\approx 83,6 \end{aligned} \quad (11.4)$$

Desta forma, a comparação entre as energias livres, de não-equilíbrio, de ligação para os ligantes estreptomicina e bluensomicina, obtidas por dinâmica molecular dirigida nas Seções 10.2.2.1 e 11.2.2, respectivamente, resulta que o aptâmero favorece a ligação da molécula de estreptomicina em relação à ligação da molécula de bluensomicina por duas ordens de grandeza.

Embora a seletividade relativa encontrada relacionando-se os valores de  $\Delta\Delta F$ , provenientes das simulações de não-equilíbrio, para as moléculas estreptomicina e bluensomicina ser aproximadamente duas ordens de grandeza menor que a seletividade relativa descrita experimentalmente por Wallace e Schroeder(37), este resultado, aliado aos valores de  $\Delta\Delta F$  calculados nas Seções 10.2.2.1 e 11.2.2, demonstram que o modelo de simulação aqui empregado pode ser utilizado como um indicativo de seletividade relativa. Estando estes qualitativamente de acordo com o comportamento seletivo esperado do aptâmero de RNA encontrado experimentalmente por Wallace e Schroeder(37)

#### 11.2.4 Decomposição da energia livre em energia interna e entropia

Conforme realizado na Seção 10.2.3, nesta Seção será realizada a decomposição da energia livre nas suas contribuições de energia interna e entropia. Para isto, serão utilizadas a Equação 5.29, para a energia interna, e a Equação 5.32, para a entropia.

##### 11.2.4.1 Energia interna

A diferença de energia interna entre um estado inicial ( $\lambda = 0$ ) e um outro estado ( $\lambda$ ) é dada, conforme desenvolvida na Seção 5.29, pela equação:

$$\Delta U = \frac{\langle U(x)e^{-\beta W} \rangle}{\langle e^{-\beta W} \rangle} - \langle U(x=0) \rangle \quad (5.29)$$

onde a energia  $U$  pode ser obtida do Hamiltoniano  $H(\mathbf{r}, \mathbf{p})$  do sistema.

A Figura 59 mostra a diferença de energia interna  $\Delta U$ , de não-equilíbrio, em função da distância de ponta a ponta para o processo de desenovelamento da molécula de RNA induzido por força externa para os sistemas *aptâmero* e *complexo-bluensomicina*. As curvas de  $\Delta U$ , assim como as curvas de força e extensão, possuem um comportamento bifásico com dois valores de máximo. Entretanto, a curva de  $\Delta U$  para o sistema *complexo-bluensomicina* apresenta um comportamento diferente daquele apresentado para os sistemas *aptâmero* e *complexo-estreptomicina* (Figura 45, página 101). A curva de  $\Delta U$  para o sistema *complexo-bluensomicina* apresenta um comportamento crescente, com um ponto de máximo local em aproximadamente 12,0 nm, até o ponto de último contato em 25,0 nm.

Se pode ver na Figura 59 que, até aproximadamente o meio da região I ( $\approx 7,0$  nm), o valor da diferença de energia interna  $\Delta U$  é aproximadamente zero para o sistema

*complexo-bluensomicina*. Entretanto, para o sistema *aptâmero*,  $\Delta U$  é negativo nesta primeira parte da região I, ou seja, o sistema está liberando energia. Este fato foi explicado na Seção 10.2.3.1, onde se inferiu a formação de ligações de hidrogênio adicionais devido a ausência da molécula ligante.

Após este ligeiro rearranjo da estrutura do sistema *aptâmero* na primeira metade da região I, a energia interna cresce em função da distância de ponta a ponta até a fronteira entre as regiões I e II, onde o valor de  $\Delta U$  atinge o primeiro máximo local. Entretanto, a curva de  $\Delta U$  para o sistema *complexo-bluensomicina* não atinge um máximo local. Para este sistema, há apenas uma redução na taxa de crescimento de  $\Delta U$ . Deve-se notar, que o sistema *complexo-bluensomicina* absorve mais energia que o sistema *aptâmero* nas regiões I e II. A Seção 11.1 mostrou que a estrutura do sistema *complexo-bluensomicina* sofre um rearranjo muito maior que a estrutura do sistema *aptâmero* nas regiões I e II, justificando o maior valor de sua energia interna. Utilizando a definição de energia interna dada no início da Seção 10.2.3.1, o comportamento da curva de  $\Delta U$  para o sistema *complexo-bluensomicina* pode ser atribuído principalmente à necessidade do sistema absorver energia para promover a quebra das ligações de hidrogênio entre as cadeias do aptâmero, Figura 56b (página 113), influenciadas pela presença da molécula de bluensomicina ligada à estrutura.

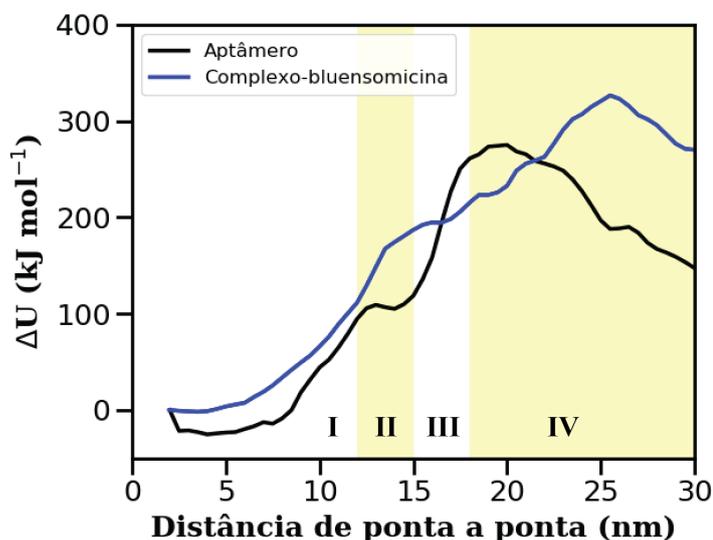


Figura 59 – Diferença de energia interna  $\Delta U$  em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-bluensomicina*.

De acordo com a Figura 56b, a estrutura secundária do aptâmero inicia a se desfazer em aproximadamente 5,0 nm, de forma que a transição da estrutura inicial para a estrutura em forma de escada ocorre de forma menos abrupta para o sistema *complexo-bluensomicina* do que ocorre para o sistema *aptâmero*, justificando assim a falta de um máximo local no perfil de  $\Delta U$  para o sistema *complexo-bluensomicina*, como ocorre nas

curvas de  $\Delta U$  para os sistemas *aptâmero* e *complexo-estreptomicina*.

Conforme os parâmetros estruturais calculados na Seção 11.1, as interações que mantêm a estrutura coesa em torno do sítio de ligação são rompidas, mediante a aplicação de força externa, nas regiões I e II da CFE. A análise das estruturas representativas da Figura 49 mostram que, após as interações que compõem a estrutura do sítio de ligação serem rompidas a estrutura passa para uma forma de escada. Pode-se ver, na região III da Figura 59 que a energia interna do sistema *aptâmero* cresce linearmente, conforme as ligações de hidrogênio entre fitas são desfeitas, até que o sistema alcance o ponto de ruptura, fronteira entre as regiões III e IV. Entretanto, para o sistema *complexo-bluensomicina* a energia interna aumenta linearmente até o ponto de último contato, não apresentando máximo no ponto de ruptura entre as regiões III e IV.

A falta de um máximo no perfil de  $\Delta U$  para o sistema *complexo-bluensomicina* pode ser explicado mediante a avançada separação das duas cadeias de RNA do aptâmero. O número de ligações de hidrogênio entre as cadeias para o sistema *complexo-bluensomicina*, Figura 56b (página 113), mostra que as duas cadeias possuem significativamente menos interação do que ocorre no sistema *aptâmero*. Aliado ao fato da molécula de bluensomicina não possuir uma interação forte com as cadeias via ligações de hidrogênio, Figura 55 (página 112), a separação final entre as cadeias para o sistema *complexo-bluensomicina* ocorre de forma linear e sem impedimento.

#### 11.2.4.2 Entropia do sistema

Utilizando os resultados de energia livre em função da distância de ponta a ponta da Seção 11.2.1 e os resultados de energia interna da Seção 11.2.4.1, podemos obter, com auxílio da Equação 5.32, o valor da entropia do sistema em função da distância de ponta a ponta.

A Figura 60 mostra a entropia  $\Delta S$  em função da distância de ponta a ponta, calculadas pela Equação 5.32, para os sistemas *aptâmero* e *complexo-bluensomicina*. Se pode ver que a curva de  $\Delta S$  para o sistema *aptâmero* apresenta dois pontos de máximo locais que coincidem com os máximos locais presentes nas curvas de força e extensão da Figura 48A (página 107) e de energia interna (Figura 59, página 122). Entretanto, a curva de  $\Delta S$  para o sistema *complexo-bluensomicina* apresenta apenas uma inflexão em seu crescimento no valor de dpp correspondente ao primeiro ponto de máximo da curva de força e extensão da Figura 48B (página 107) e um ponto de máximo na dpp correspondente ao ponto de último contato entre as cadeias.

Segundo a análise da curva de entropia do sistema para o sistema *aptâmero* realizada na Seção 10.2.3.2, na região I o valor de  $T\Delta S$  para este sistema decresce até aproximadamente o meio desta região devido as restrições de movimento impostas pelas ligações de hidrogênio, tanto intra como entre cadeias, à estrutura da molécula do ap-

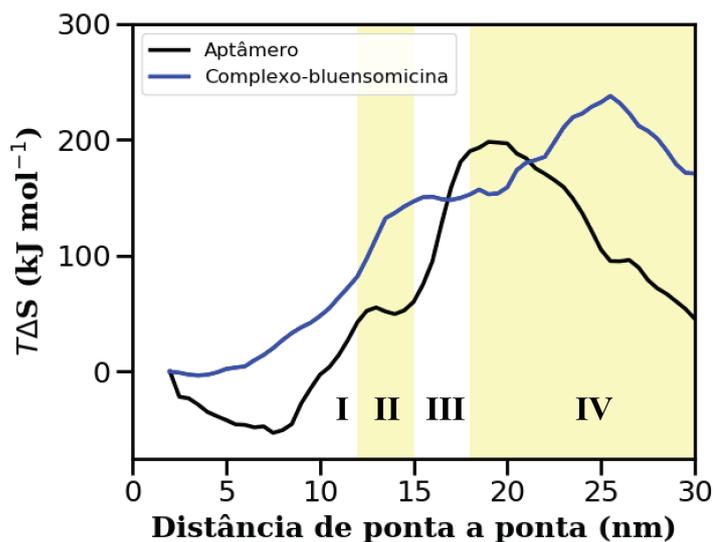


Figura 60 – Produto entre a temperatura e diferença de entropia  $T\Delta S$  em função da distância de ponta a ponta para os sistemas *aptâmero* e *complexo-bluensomicina*.

tâmero de RNA, e uma vez que estas restrições são removidas, o valor de  $T\Delta S$  volta a crescer até o sistema evolua o valor da dpp até o primeiro ponto de máximo local. Porém, podemos ver na Figura 60 que para o sistema *complexo-bluensomicina*, o valor de  $T\Delta S$  cresce até atingir o ponto de inflexão em uma dpp de aproximadamente 12,0 nm. Este comportamento indica que o sistema *complexo-bluensomicina* não apresenta restrições de movimento significativas em sua estrutura. De fato, o número de multipletos, Figura 51 (página 109) e o número de ligações de hidrogênio entre as cadeias, Figura 56b (página 113), para o sistema *complexo-bluensomicina* mostram que a estrutura terciária e secundária do aptâmero começam a se desestruturar tão logo a força externa começa a ser aplicada ao sistema.

A Figura 60 mostra que a entropia do sistema é mais positiva para o sistema *complexo-bluensomicina* entre as regiões I e II do que para o sistema *aptâmero*. Este comportamento é explicado pela diminuição das ligações de hidrogênio entre as cadeias no sistema *complexo-bluensomicina*. Este menor número de ligações de hidrogênio faz com que as cadeias possuam menor restrição de movimento e, portanto maior número de estados acessíveis ao sistema. Consequentemente, a entropia do sistema para o sistema *complexo-bluensomicina* é mais positiva do que para o sistema *aptâmero* nestas duas regiões.

Uma vez que a única diferença entre estes sistemas é a presença da molécula de bluensomicina, podemos inferir que a presença desta molécula no sítio de ligação torna o desenovelamento do sítio ativo entropicamente favorável mediante a aplicação de força externa ao sistema. A presença da molécula de bluensomicina induz a desestruturação do sítio ativo, estrutura terciária, e também das ligações de hidrogênio entre as cadeias,

estrutura secundária, de forma que o número de estados acessíveis ao sistema *complexo-bluensomicina* aumenta significativamente com a dpp.

Podemos ver, na Figura 60, que o valor de  $T\Delta S$  para o sistema *complexo-bluensomicina* cresce aproximadamente monotonicamente até que o sistema se encontre no ponto de último contato entre as cadeias. Este aumento linear no valor da entropia do sistema se deve a separação linear dps pares de base das cadeias do aptâmero, uma vez que a molécula de bluensomicina interage apenas com uma cadeia do aptâmero de RNA.

## Conclusão

A aplicação da metodologia de dinâmica molecular dirigida, uma técnica de simulação fora do equilíbrio, permitiu aumentar o conhecimento sobre o comportamento da molécula de RNA, aptâmero, e de seus complexos com as moléculas estreptomicina e bluensomicina sobre a aplicação de força, como o que ocorre experimento de microscopia de força atômica de molécula única. A aplicação desta técnica possibilitou o entendimento em nível atômico do processo de desenovelamento do aptâmero de RNA e seus complexos.

As análises estruturais realizadas sobre as trajetórias de desnaturação mostraram que existe uma pequena, porém não desprezível, diferença entre o processo de desnaturação do aptâmero e seu complexo com a molécula de estreptomicina. Os resultados aqui apresentados mostram que existe uma ordem hierárquica no processo de desnaturação, onde o sítio de ligação permanece estruturado durante a primeira etapa deste processo. Estes resultados permitiram a construção de um modelo em duas etapas para o desenovelamento induzido por aplicação de força externa. As análises estruturais realizadas sobre as trajetórias de dinâmica molecular dirigida, do sistema composto pelo complexo entre o aptâmero de RNA e a molécula de bluensomicina, mostraram que este sistema também segue o modelo proposto de desnaturação em duas etapas.

A aplicação da igualdade de Jarzynski nos valores de trabalhos obtidos para o processo de desnaturação dos sistemas *aptâmero* e *complexo-estreptomicina* forneceu valores de energia livre de acordo com os valores experimentais. A decomposição da energia livre, utilizando o Teorema de Flutuação, permitiu um entendimento das contribuições de energia interna e entropia relacionadas ao processo de desnaturação. O modelo de simulação aqui proposto, ao ser aplicado ao complexo entre o aptâmero e a molécula de bluensomicina, forneceu valores de energia livre compatíveis com os da literatura, mostrando corretamente que o complexo formado entre estas duas moléculas não é favorável. A inclusão ao modelo aqui proposto de parâmetros que sejam dependentes de variáveis experimentais possibilitará uma maior concordância entre os valores de energia livre calculados e experimentais.

Desta forma, o modelo de simulação por dinâmica molecular dirigida e a análise dos dados por este obtidos permitiram a correta ordem de seletividade do aptâmero de RNA pelas moléculas de estreptomicina e bluensomicina, bem como aumentou o conhecimento, estrutural e energético, relativo ao processo de desenovelamento induzido por aplicação de força externa dos sistemas *aptâmero*, *complexo-estreptomicina* e *complexo-bluensomicina*, podendo ser aplicado ao estudo do complexo entre o aptâmero de RNA e outras moléculas ligantes.

As simulações de dinâmica molecular dirigida realizadas neste trabalho levaram a identificar as interações por ligação de hidrogênio entre a molécula de estreptomicina e o aptâmero de RNA como o principal responsável pela estabilidade do sítio de ligação, e portanto do complexo entre aptâmero e estreptomicina. A aplicação desta metodologia ao sistema *complexo-bluensomicina* permitiu identificar que a presença desta molécula no sítio de ligação do aptâmero contribui energeticamente e entropicamente para que o aptâmero seja desenovelado. Desta forma, a metodologia de simulação empregada e o cálculo dos valores de  $\Delta F$ ,  $\Delta U$  e  $\Delta S$  permitiram um maior entendimento do processo de desenovelamento induzido por força dos complexos entre o aptâmero de RNA e as moléculas de estreptomicina e bluensomicina.

# Referências

- 1 DAHM, R. Friedrich Miescher and the discovery of DNA. *Developmental Biology*, v. 278, n. 2, p. 274–288, 2005. Citado 2 vezes nas páginas 17 e 23.
- 2 AVERY, O. T.; MACLEOD, C. M.; MCCARTHY, M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *The Journal of Experimental Medicine*, v. 79, n. 2, p. 137–158, 1944. Citado 2 vezes nas páginas 17 e 23.
- 3 EPHUSSI-TAYLOR, H. Genetic aspects of transformations of pneumococci. *Cold Springs Harbor Symposia on Quantitative Biology*, v. 16, p. 445–456, 1951. Citado 2 vezes nas páginas 17 e 23.
- 4 ŠPONER, J.; LANKAŠ, F. *Computational Studies of RNA and DNA*. [S.l.]: Springer, 2006. Citado 2 vezes nas páginas 17 e 23.
- 5 VOET, D.; VOET, J. G. *Bioquímica*. 3. ed. [S.l.]: Artmed, 2006. Citado 7 vezes nas páginas 17, 19, 23, 24, 26, 28 e 40.
- 6 TINOCO Jr., I.; BUSTAMANTE, C. How RNA folds. *J. Mol. Biol.*, v. 293, p. 271–281, 1999. Citado 3 vezes nas páginas 17, 23 e 100.
- 7 WINDBICHLER, N.; SCHROEDER, R. Isolation of specific rna-binding proteins using the streptomycin-binding rna aptamer. *Nature Protocols*, v. 1, p. 637–640, 2006. Citado na página 17.
- 8 ALEMÁN, E. A.; LAMICHHANE, R.; RUEDA, D. Exploring RNA folding one molecule at a time. *Current Opinion in Chemical Biology*, v. 12, p. 647–654, 2008. Citado na página 17.
- 9 LIPHARDT, J. et al. Reversible unfolding of single RNA molecules by mechanical force. *Science*, v. 292, p. 733–737, 2001. Citado 5 vezes nas páginas 17, 18, 43, 92 e 100.
- 10 TINOCO Jr., I.; LI, P. T. X.; BUSTAMANTE, C. Determination of thermodynamics and kinetics of RNA reactions by force. *Quarterly Reviews of Biophysics*, v. 39, p. 325–360, 2006. Citado 3 vezes nas páginas 17, 18 e 100.
- 11 ONOA, B.; TINOCO Jr., I. RNA folding and unfolding. *Current Opinion in Structural Biology*, v. 14, p. 374–379, 2004. Citado 2 vezes nas páginas 17 e 43.
- 12 BUSTAMANTE, C. et al. Single-molecule studies of DNA mechanics. *Current Opinion in Structural Biology*, v. 10, p. 279–285, 2000. Citado 4 vezes nas páginas 17, 43, 92 e 100.
- 13 CLAUSEN-SCHUMANN, H. et al. Force spectroscopy with single bio-molecules. *Current Opinion in Chemical Biology*, v. 4, p. 524–530, 2000. Citado na página 18.
- 14 ZHUANG, X. et al. A single-molecule study of RNA catalysis and folding. *Science*, v. 288, p. 2048–2051, 2000. Citado na página 18.

- 15 MAKAROV, D. E.; SCHULER, B. Preface: Special topic on single-molecule biophysics. *The Journal of Chemical Physics*, v. 148, p. 123001, 2018. Citado na página 18.
- 16 LIPHARDT, J. et al. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science*, v. 296, p. 1832–1835, 2002. Citado 3 vezes nas páginas 18, 43 e 100.
- 17 ONOA, B. et al. Identifying kinetic barriers to mechanical unfolding of the *T. thermophila* ribozyme. *Science*, v. 299, p. 1892–1895, 2003. Citado 3 vezes nas páginas 18, 43 e 100.
- 18 HARLEPP, S. et al. Probing complex RNA structures by mechanical force. *The European Physical Journal E*, v. 12, p. 605–615, 2003. Citado 2 vezes nas páginas 18 e 19.
- 19 COLLIN, D. et al. Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature Letters*, v. 437, p. 231–234, 2005. Citado 3 vezes nas páginas 18, 43 e 100.
- 20 LI, P. T. X. et al. Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods. *Biophysical Journal*, v. 90, p. 250–260, 2006. Citado 3 vezes nas páginas 18, 43 e 100.
- 21 JARZYNSKI, C. Nonequilibrium equality for free energy differences. *Physical Review Letters*, v. 78, p. 2690–2693, 1997. Citado 9 vezes nas páginas 18, 42, 43, 57, 78, 93, 100, 115 e 140.
- 22 CROOKS, G. E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, v. 60, p. 2721–2726, 1999. Citado na página 18.
- 23 VAN GUNSTEREN, W. F. et al. Biomolecular modeling: Goals, problems, perspectives. *Angewandte Chemie International Edition*, v. 45, n. 25, p. 4064–4092, 2006. Citado 5 vezes nas páginas 18, 36, 40, 49 e 99.
- 24 LI, M. S.; MAI, B. K. Steered molecular dynamics - a promising tool for drug design. *Current Bioinformatics*, v. 7, p. 342–351, 2012. Citado 2 vezes nas páginas 19 e 38.
- 25 ALLEN, W. J.; BEVAN, D. R. Steered molecular dynamics simulations reveal important mechanisms in reversible monoamine oxidase b inhibition. *Biochemistry*, v. 50, p. 6441–6454, 2011. Citado 2 vezes nas páginas 19 e 38.
- 26 NASERIAN-NIK, A. M.; TAHANI, M.; KARTTUNEN, M. Pulling of double-stranded DNA by atomic force microscopy: a simulation in atomistic details. *RSC Advances*, v. 3, p. 10516–10528, 2013. Citado 7 vezes nas páginas 19, 21, 38, 39, 54, 70 e 78.
- 27 NASERIAN-NIK, A. M.; TAHANI, M.; KARTTUNEN, M. Molecular dynamics study of DNA oligomers under angled pulling. *RSC Advances*, v. 4, p. 10751–10760, 2014. Citado 5 vezes nas páginas 19, 38, 54, 70 e 78.
- 28 LAIO, A.; PARRINELLO, M. Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, v. 99, n. 20, p. 12562–12566, 2002. Citado na página 19.

- 29 GERLAND, U.; BUNDSCHUH, R.; HWA, T. Force-induced denaturation of RNA. *Biophysical Journal*, v. 81, p. 1324–1332, 2001. Citado na página 19.
- 30 DENG, N.-J.; CIEPLAK, P. Free energy profile of rna hairpins: A molecular dynamics simulation study. *Biophys Journal*, v. 98, p. 627–636, 2010. Citado na página 19.
- 31 GUPTA, A.; BANSAL, M. The role of sequence in altering the unfolding pathway of an rna pseudoknot: a steered molecular dynamics study. *Phys. Chem. Chem. Phys.*, v. 18, p. 28767–28780, 2016. Citado na página 19.
- 32 BELL, D. R. et al. Capturing rna folding free energy with coarse-grained molecular dynamics simulations. *Scientific Reports*, v. 7, p. 45812, 2017. Citado na página 19.
- 33 SHARMA, D. et al. Mutational analysis of s12 protein and implications for the accuracy of decoding by the ribosome. *Journal of Molecular Biology*, v. 374, p. 1065–1076, 2007. Citado na página 19.
- 34 TERESHKO, V.; SKRIPKIN, E.; PATEL, D. J. Encapsulating Streptomycin within a Small 40-mer RNA. *Chemistry & Biology*, v. 10, p. 175–187, 2003. Citado 12 vezes nas páginas 19, 20, 29, 59, 61, 62, 65, 67, 106, 111, 112 e 120.
- 35 NICK, T. A. et al. Stability of a split streptomycin binding aptamer. *The Journal of Pyhysical Chemistry B*, v. 120, p. 6479–6489, 2016. Citado 16 vezes nas páginas 19, 21, 55, 76, 78, 79, 80, 81, 82, 90, 92, 93, 97, 98, 99 e 119.
- 36 VON AHSEN, U.; SCHROEDER, R. Streptomycin inhibits splicing of group I introns by competition with the guanosine substrate. *Nucleic Acids Research*, v. 19, p. 2261–2265, 1991. Citado 5 vezes nas páginas 20, 59, 67, 106 e 120.
- 37 WALLACE, S. T.; SCHROEDER, R. In vitro selection and characterization of streptomycin-binding RNAs: Recognition discrimination between antibiotics. *RNA*, v. 4, p. 112–123, 1998. Citado 10 vezes nas páginas 20, 28, 29, 59, 67, 106, 112, 119, 120 e 121.
- 38 PARK, S. et al. Free energy calculation from steered molecular dynamics simulations using jarzynski's equality. *Journal of Chemical Physics*, v. 119, n. 6, p. 3559–3566, 2003. Citado 6 vezes nas páginas 21, 43, 56, 76, 92 e 140.
- 39 YTREBERG, F. M.; ZUCKERMAN, D. M. Efficient use of nonequilibrium measurement to estimate free energy differences for molecular systems. *J. computational Chemistry*, v. 25, p. 1749–1759, 2004. Citado 7 vezes nas páginas 21, 43, 57, 92, 93, 99 e 115.
- 40 ATTWOOD, T. K. et al. *OXFORD dictionary of Biochemistry and Molecular Biology*. [S.l.]: Oxford University Press, 2006. Citado 2 vezes nas páginas 23 e 24.
- 41 ALBERTS, B. et al. *Molecular Biology of the Cell*. 3. ed. [S.l.]: Garland Publishing, Inc., 1994. Citado 2 vezes nas páginas 23 e 24.
- 42 CASPERSSON, T.; SCHULTZ, J. Pentose nucleotides in the cytoplasm of growing tissues. *Nature*, v. 143, p. 602–603, 1939. Citado na página 23.

- 43 LIU, J.; CAO, Z.; LU, Y. Functional nucleic acid sensors. *Chemical Reviews*, v. 109, p. 1948–1998, 2009. Citado na página 24.
- 44 BLANK, M.; BLIND, M. Aptamers as tools for target validation. *Current Opinion in Chemical Biology*, p. 336. Citado na página 24.
- 45 WATSON, J. D. et al. *Molecular Biology of the Gene*. 5. ed. [S.l.]: Pearson Benjamin Cummings, 2004. Citado 2 vezes nas páginas 24 e 28.
- 46 DOONAN, S. *Nucleic Acids*. [S.l.]: Royal Society of Chemistry, 2004. Citado 2 vezes nas páginas 24 e 28.
- 47 HOLDE, K. E. van; JOHNSON, W. C.; HO, P. S. *Principles of Physical Biochemistry*. [S.l.]: Prentice Hall, 1998. Citado 2 vezes nas páginas 24 e 48.
- 48 LU, X.-J. *Detection of multiplets in DSSR*. Acessado em 7 de julho de 2018. Disponível em: <<http://x3dna.org/highlights/detection-of-multiplets-in-dssr>> Citado 4 vezes nas páginas 26, 60, 61 e 88.
- 49 CANTOR, C. R.; SCHIMMEL, P. R. *Biophysical Chemistry*. [S.l.]: W.H. Freeman and Company, 1984. Citado na página 28.
- 50 BINNING, G.; QUATE, C. F.; GERBER, C. Atomic force microscope. *Physical Review Letters*, v. 56, p. 930–933, 1986. Citado na página 30.
- 51 HANSMA, H. G.; KASUYA, K.; OROUDJEV, E. Atomic force microscopy imaging and pulling of nucleic acids. *Current Opinion in Structural Biology*, v. 14, p. 380–385, 2004. Citado na página 30.
- 52 ALONSO, J. L.; GOLDMANN, W. H. Feeling the forces: atomic force microscopy in cell biology. *Life Sciences*, v. 72, p. 2553–2560, 2003. Citado na página 30.
- 53 WILIAMS, M. C.; ROUZINA, L. Force spectroscopy of single DNA and RNA molecules. *Current Opinion in Structural Biology*, v. 12, p. 330–336, 2002. Citado na página 30.
- 54 ALLEN, M. P.; TILDESLEY, D. J. *Computer Simulation of Liquids*. [S.l.]: Clarendon Press, 1987. Citado 3 vezes nas páginas 32, 33 e 39.
- 55 ALDER, B. J.; WAINWRIGHT, T. E. Phase Transition for a Hard Sphere System. *The Journal of Chemical Physics*, v. 27, p. 1208–1209, 1957. Citado na página 32.
- 56 ALDER, B. J.; WAINWRIGHT, T. E. Studies in Molecular Dynamics. I. General Method. *The Journal of Chemical Physics*, v. 31, p. 459–466, 1959. Citado na página 32.
- 57 RAHMAN, A. Correlations in the motion of atoms in liquid argon. *Physical Review*, v. 136, p. A405–A411, 1964. Citado na página 32.
- 58 RAHMAN, A.; STILLINGER, F. H. Molecular dynamics study of liquid water. *TBE Journal of Chemical Physics*, v. 55, n. 7, p. 3336–3359, 1971. Citado na página 32.
- 59 STILLINGER, F. H.; RAHMAN, A. J. Improved simulation of liquid water by molecular dynamics. *The Journal of Chemical Physics*, v. 60, p. 1445–1557, 1974. Citado na página 32.

- 60 MCCAMMON, J. A.; GELIN, B. R.; KARPLUS, M. Dynamics of folded proteins. *Nature*, v. 267, p. 585–590, 1977. Citado na página 32.
- 61 LEVITT, M. Computer simulation of DNA double-helix dynamics. *Cold Spring Harbor Symposia on Quantitative Biology*, v. 47, p. 251–262, 1983. Citado na página 32.
- 62 BEVERIDGE, D. L.; CHEATHAM, T. E.; MEZEL, M. The ABCs of molecular dynamics simulations on B-DNA, circa 2012. *Journal of Biosciences*, v. 37, n. 3, p. 379–397, 2012. Citado na página 32.
- 63 FRENKEL, D.; SMIT, B. *Understanding molecular simulation: from algorithms to applications*. 2. ed. [S.l.]: Academic Press, 1996. Citado 3 vezes nas páginas 33, 35 e 36.
- 64 OOSTENBRINK, C. et al. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, v. 25, n. 13, p. 1656–1676, 2004. Citado 2 vezes nas páginas 33 e 34.
- 65 LEACH, A. *Molecular modelling: principles and applications*. 2. ed. [S.l.]: Pearson Education, 2001. Citado na página 35.
- 66 FRENKEL, D. Simulations: The dark side. *The European Physical Journal Plus*, v. 120, n. 1, 2013. Nenhuma citação no texto.
- 67 LEINES, G. D. *Path-metadynamics: A computational study of conformational transitions in proteins*. Tese (Doutorado) — Universiteit van Amsterdam. Faculteit der Natuurwetenschappen, Wiskunde en Informatica, 2012. Nenhuma citação no texto.
- 68 RCSC. *RCSC Protein Data Bank*. 2013. Disponível em: <<http://www.rcsb.org/pdb/home/home.do>>. Citado 2 vezes nas páginas 35 e 51.
- 69 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. 5179–5197, 1995. Citado na página 35.
- 70 MORGON, N. H.; COUTINHO, K. *Métodos de Química Teórica e Modelagem Molecular*. 1. ed. [S.l.]: Livraria da Física, 2007. Citado na página 36.
- 71 JONES, J. E. On the determination of molecular fields II: From the equation of state of a gas. *Proceedings of the Royal Society of London Series A*, v. 106, n. 738, p. 463–477, Oct 1924. Citado na página 36.
- 72 APOL, E. et al. *GROMACS USER MANUAL, version 4.5.4*. Suécia, 2010. Citado 3 vezes nas páginas 36, 51 e 54.
- 73 DURRANT, J.; MCCAMMON, J. A. Molecular dynamics simulations and drug discovery. *BMC Biology*, v. 9, n. 1, p. 71, 2011. Citado na página 36.
- 74 MCQUARRIE, D. A. *Statistical Mechanics*. [S.l.]: University Science Books, 2000. Citado 5 vezes nas páginas 36, 37, 40, 47 e 49.
- 75 CHANDLER, D. *Introduction to modern statistical mechanics*. [S.l.]: Oxford University Press, 1987. Citado 2 vezes nas páginas 36 e 49.

- 76 NOSÉ, S. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, v. 52, p. 255–268, 1984. Citado 3 vezes nas páginas 36, 53 e 54.
- 77 HOOVER, W. G. Canonical dynamics: equilibrium phase-space distributions. *Physical Review A*, v. 31, p. 1695–1697, 1985. Citado 3 vezes nas páginas 36, 53 e 54.
- 78 MARTYNA, G. J. et al. Explicit reversible integrators for extended systems dynamics. *Molecular Physics*, v. 87, p. 1117–1157, 1996. Citado na página 36.
- 79 BERENDSEN, H. J. *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics*. [S.l.]: Cambridge University Press, 2007. Citado 6 vezes nas páginas 37, 40, 41, 43, 44 e 49.
- 80 TUCKERMAN, M. E. *Statistical Mechanics: Theory and Molecular Simulation*. [S.l.]: Oxford University Press, 2010. Citado 2 vezes nas páginas 37 e 49.
- 81 SCHLINCK, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide*. 2. ed. [S.l.]: Springer, 2010. Citado 2 vezes nas páginas 37 e 38.
- 82 PASTOR, R. W. Techniques and applications of langevin dynamics simulations. In: LUCKHURST, G. R.; VERACINI, C. A. (Ed.). *The Molecular Dynamics of Liquid Crystals*. Dordrecht: Springer Science+Business Media, 1994. p. 85–138. Citado 2 vezes nas páginas 37 e 38.
- 83 CASE, D. A. Normal mode analysis of protein dynamics. *Current Opinion in Structural Biology*, v. 4, p. 285–290, 1994. Citado na página 38.
- 84 BARTH, E.; SCHLICK, T. Overcoming stability limitations in biomolecular dynamics. i. combining force splitting via extrapolation with langevin dynamics in. *The Journal of Chemical Physics*, v. 109, p. 1617–1632, 1998. Citado na página 38.
- 85 YOURGRAU, W.; MERWE, A.; RAW, G. *Treatise on Irreversible and Statistical Thermophysics*. [S.l.]: Dover Publications, Inc, 1966. Citado na página 38.
- 86 RYCKAERT, J.; CICCOTTI, G.; BERENDSEN, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, v. 23, n. 3, p. 327–341, 1977. Citado na página 39.
- 87 SCHRÖDINGER, E. *What is life? The physical Aspect of the living cell*. [S.l.]: Cambridge University Press, 1944. Citado na página 40.
- 88 ORNES, S. How nonequilibrium thermodynamics speaks to the mystery of life. *PNAS - Proceedings of the National Academy of Sciences*, v. 114, n. 3, p. 423–324, 2017. Citado na página 40.
- 89 CHIPOT, C. Frontiers in free-energy calculations of biological systems. *WIREs Computational Molecular Science*, v. 4, p. 71–89, 2014. Citado 3 vezes nas páginas 40, 41 e 42.
- 90 HANSEN, N.; GUNSTEREN, W. F. van. Practical aspects of free-energy calculations: A review. *Journal of Chemical Theory and Computation*, v. 10, p. 2632–2647, 2014. Citado 2 vezes nas páginas 40 e 41.

- 91 BECKER, O. M. et al. *Computational Biochemistry and Biophysics*. [S.l.]: Marcel Dekker, Inc., 2001. Citado na página 40.
- 92 POHORILLE, A.; JARZYNSKI, C.; CHIPOT, C. Good practices in free-energy calculations. *Journal of Physical Chemistry B*, v. 114, p. 10235–10253, 2010. Citado na página 41.
- 93 LIANG, C. et al. Nonequilibrium thermodynamics and fluctuation relations for small systems. *Chinese Physical B*, v. 27, p. 070501, 2014. Citado 3 vezes nas páginas 41, 43 e 44.
- 94 KONDEPUDI, I. P. e D. *Termodinâmica: Dos motores térmicos às estruturas dissipativas*. [S.l.]: Instituto Pger, 1999. Citado 2 vezes nas páginas 42 e 49.
- 95 RICO, F. et al. High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations. *Science*, v. 342, p. 741–743, 2013. Citado na página 42.
- 96 JARZYNSKI, C. Equalities and inequalities: Irreversibility and the second law of the thermodynamics at the nanoscale. *Annu. Rev. Condens. Matter Phys.*, v. 2, p. 329–351, 2010. Citado 3 vezes nas páginas 42, 43 e 140.
- 97 JARZYNSKI, C. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Physical Review E*, v. 56, p. 5018–5035, 1997. Citado na página 43.
- 98 CROOKS, G. E. Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. *Journal of Statistical Physics*, v. 90, n. 5/6, p. 1481–1487, 1998. Citado 3 vezes nas páginas 43, 44 e 46.
- 99 CROOKS, G. E. Path-ensemble averages in systems driven far from equilibrium. *Physical Review E*, v. 61, n. 3, p. 2361–2366, 2000. Citado 2 vezes nas páginas 43 e 48.
- 100 HUMMER, G.; SZABO, A. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. v. 98, n. 7, p. 3658–3661, 2001. Citado 2 vezes nas páginas 43 e 76.
- 101 ATHÉNES, M. A path-sampling scheme for computing thermodynamic properties of a many-body system in a generalized ensemble. *The European Physical Journal B*, v. 38, n. 4, p. 651–663, 2004. Citado na página 43.
- 102 SCHURR, J. M.; FUJIMOTO, B. S. Equalities for the nonequilibrium work transferred from an external potential to a molecular system. analysis of single-molecule extension experiments. *The Journal of Physical Chemistry B*, v. 107, n. 50, p. 14007–14019, 2003. Citado 2 vezes nas páginas 43 e 44.
- 103 SUN, S. X. Equilibrium free energies from path sampling of nonequilibrium trajectories. *Journal of Chemical Physics*, v. 118, p. 5769–5775, 2003. Citado 2 vezes nas páginas 43 e 44.
- 104 PAGE, D. N. *Generalized Jarzynski Equality*. 2012. Citado 2 vezes nas páginas 43 e 44.

- 105 GITTES, F. Two famous results of Einstein derived from the Jarzynski equality. *American Journal of Physics*, v. 86, n. 1, p. 31–35, 2018. Citado 2 vezes nas páginas 43 e 44.
- 106 CUENDET, M. A. Statistical mechanical derivation of Jarzynski's identity for thermostated non-Hamiltonian dynamics. *Physical Review Letters*, v. 96, p. 120602, 2006. Citado 3 vezes nas páginas 43, 44 e 140.
- 107 Van KAMPEN, N. G. *Stochastic Processes in Physics and Chemistry*. 3. ed. [S.l.]: Elsevier, 2007. Citado na página 46.
- 108 HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics*, v. 14, p. 33–38, 1996. Citado na página 51.
- 109 STONE, J. An Efficient Library for Parallel Ray Tracing and Animation. Dissertação (Mestrado) — Computer Science Department, University of Missouri-Rolla, April 1998. Citado na página 51.
- 110 BERENDSEN, H. J. C.; SPOEL, D. v. d.; DRUNEN, R. v. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications*, v. 91, n. 1-3, p. 43–56, 1995. Citado 2 vezes nas páginas 51 e 54.
- 111 SPOEL, D. V. D. et al. GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, v. 26, n. 16, p. 1701–1718, 2005. Citado 2 vezes nas páginas 51 e 54.
- 112 HESS, B. et al. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, v. 4, n. 3, p. 435–447, 2008. Citado 2 vezes nas páginas 51 e 54.
- 113 PRONK, S. et al. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. v. 29, n. 7, p. 845–854, 2013. Citado 2 vezes nas páginas 51 e 54.
- 114 PÉREZ, A. et al. Refinement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers. *Biophysical Journal*, v. 92, n. 11, p. 3817–3829, June 2007. Citado na página 51.
- 115 GUY, A. T.; PIGGOT, T. J.; KHALID, S. Single-stranded DNA within nanopores: Conformational dynamics and implications for sequencing: a molecular dynamics simulation study. *Biophysical journal*, v. 103, n. 5, p. 1028–1036, 2012. Citado na página 51.
- 116 FRISCH, M. J. et al. *Gaussian 09, Revision A1*. Wallingford CT, 2009. Citado na página 52.
- 117 DENNINGTON, R.; KEITH, T.; MILLAM, J. *GaussView, Version 5*. Shawnee Mission, KS: [s.n.], 2009. Citado na página 52.
- 118 STEWART, J. J. P. Optimization of parameters for semiempirical methods I. Method. *Journal of Computational Chemistry*, v. 10, n. 2, p. 209–220, 1989. Citado na página 52.

- 119 STEWART, J. J. P. Optimization of parameters for semiempirical methods II. Applications. *Journal of Computational Chemistry*, v. 10, n. 2, p. 221–264, 1989. Citado na página 52.
- 120 RASSOLOV, V. A. et al. 6-31G\* basis set for atoms K through Zn. *The Journal of Chemical Physics*, v. 109, n. 4, p. 1223–1229, 1998. Citado na página 52.
- 121 RASSOLOV, V. A. et al. 6-31G\* basis set for third-row atoms. *Journal of Computational Chemistry*, v. 22, n. 9, p. 976–984, 2001. Citado na página 52.
- 122 SINGH, U. C.; KOLLMAN, P. A. An approach to computing electrostatic charges for molecules. *Journal of Computational Chemistry*, v. 5, n. 2, p. 129–145, 1984. Citado na página 53.
- 123 BESLER, B. H.; MERZ, K. M.; KOLLMAN, P. A. Atomic charges derived from semiempirical methods. *Journal of Computational Chemistry*, v. 11, n. 4, p. 431–439, 1990. Citado na página 53.
- 124 WANG, J. et al. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*, v. 25, n. 2, p. 247–260, 2006. Citado na página 53.
- 125 WANG, J. et al. Development and testing of a general amber force field. *Journal of Computational Chemistry*, v. 25, p. 1157–1174, 2004. Citado na página 53.
- 126 BAYLY, C. et al. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: The RESP model. *Journal of Physical Chemistry*, v. 97, p. 10269–10280, 1993. Nenhuma citação no texto.
- 127 CORNELL, W. et al. Application of RESP charges to calculate conformational energies, hydrogen bond energies and free energies of solvation. *Journal of the American Chemical Society*, v. 115, p. 9620–9631, 1993. Nenhuma citação no texto.
- 128 CIEPLAK, P. et al. Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA and proteins. *Journal of computational chemistry*, v. 16, p. 1357–1377, 1995. Nenhuma citação no texto.
- 129 SILVA, A. W. S.; VRANKEN, W. F. ACPYPE-AnteChamber PYthon Parser interfacE. *BMC Research Notes*, v. 5, n. 1, p. 1–8, 2012. Citado na página 53.
- 130 JORGENSEN, W. L. et al. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, v. 79, n. 2, p. 926–935, 1983. Citado na página 53.
- 131 PARRINELLO, M.; RAHMAN, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, v. 52, p. 7182–7190, 1981. Citado 2 vezes nas páginas 53 e 54.
- 132 ESSMANN, U.; L., P.; BERKOWITZ, M. L. A smooth particle mesh Ewald method. *Journal of Chemical Physics*, v. 103, n. 19, p. 8577–8593, 1995. Citado na página 54.

- 133 TOUKMAJI, A. Y.; BOARD, J. A. Ewald summation techniques in perspective: a survey. *Computer Physics Communications*, v. 95, p. 73–92, 1996. Citado na página 54.
- 134 DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *Journal of Chemical Physics*, v. 98, n. 12, p. 10089–10092, 1993. Citado na página 54.
- 135 HOCKNEY, R. W.; GOEL, S. P.; EASTWOOD, J. Quiet high resolution computer models of a plasma. *J. Comp. Phys.*, v. 14, p. 148–158, 1974. Citado na página 54.
- 136 GUNSTEREN, W. F. van; BERENDSEN, H. J. C. A leap-frog algorithm for stochastic dynamics. *Molecular Simulation*, v. 1, p. 173–185, 1988. Citado na página 54.
- 137 HAWKINS, G. D.; CRAMER, C. J.; TRUHLAR, D. G. Pairwise solute descreening of solute charges from a dielectric medium. *Chemical Physics Letters*, v. 246, p. 122–129, 1995. Citado na página 54.
- 138 HAWKINS, G. D.; CRAMER, C. J.; TRUHLAR. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *The Journal of Physical Chemistry*, v. 100, p. 19824–19839, 1996. Citado na página 54.
- 139 TSUI, V.; CASE, D. A. Theory and applications of the generalized born solvation models in macromolecular simulations. *Biopolymers*, v. 56, p. 275–291, 2001. Citado na página 54.
- 140 ŘEZÁČ, J.; HOBZA, P.; HARRIS, S. A. Stretched DNA investigated using molecular-dynamics and quantum-mechanical calculations. *Biophysical Journal*, v. 98, p. 101–110, 2010. Citado na página 54.
- 141 CARTER, A. P. et al. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*, v. 407, p. 340–348, 2000. Citado 2 vezes nas páginas 59 e 67.
- 142 MANOSAS, M.; RITORT, F. Thermodynamic and kinetic aspects of rna pulling experiments. *Biophysical Journal*, v. 88, p. 3224–3242, 2005. Citado 3 vezes nas páginas 78, 92 e 99.
- 143 NUNES-ALVES, A.; ARANTES, G. M. Mechanical Unfolding of Macromolecules Coupled to Bond Dissociation. *J. Chem. Theory Comput.*, v. 14, p. 282–290, 2018. Citado 2 vezes nas páginas 78 e 99.
- 144 HUMMER, G.; SZABO, A. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophysical Journal*, v. 85, p. 5–15, 2003. Citado 3 vezes nas páginas 78, 92 e 99.
- 145 HENDRIX, D. A.; JARZYNSKI, C. A "fast growth" method of computing free energy differences. *The Journal of Chemical Physics*, v. 114, n. 14, p. 5974–5981, 2001. Citado 2 vezes nas páginas 78 e 99.
- 146 SHERIDAN, S.; GRÄTER, F.; DADAY, C. How fast is too fast in force-probe molecular dynamics simulations? *J. Phys. Chem. B*, v. 123, p. 3658–3664, 2019. Citado 2 vezes nas páginas 78 e 99.

- 147 WEST, D. K.; OLMSTED, P. D.; PACI, E. Free energy of protein folding from nonequilibrium simulations using the Jarzynski equality. *The Journal of Chemical Physics*, v. 125, p. 204910, 2006. Citado 2 vezes nas páginas 78 e 99.
- 148 LU, X.-J. *Stem, helix, and coaxial stacking in DSSR*. Acessado em 7 de julho de 2018. Disponível em: <<http://x3dna.org/highlights/stem-helix-and-coaxial-stacking-in-dssr>>. Citado na página 88.
- 149 MARSH, J. A. Buried and Accessible Surface Area Control Intrinsic Protein Flexibility. *Journal of Molecular Biology*, v. 425, n. 17, p. 3250–3263, 2013. Citado na página 92.
- 150 CHAKRAVARTY, D. et al. Reassessing buried surface areas in protein–protein complexes. *PROTEIN SCIENCE*, v. 22, p. 1453–1457, 2013. Citado na página 92.
- 151 RASHIN, A. A. Buried surface area, conformational entropy, and protein stability. *Biopolymers*, v. 23, n. 8, p. 1605–1620, 1984. Citado na página 92.
- 152 EISENHABER, F. et al. The double lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comp. Chem.*, v. 16, p. 273–284, 1995. Citado na página 92.
- 153 CHEN, L. Y.; BASTIEN, D. A.; ESPEJEL, H. E. Determination of equilibrium free energy from nonequilibrium work measurements. *Physical Chemistry Chemical Physics*, v. 12, p. 6579–6582, 2010. Citado 2 vezes nas páginas 92 e 99.
- 154 OBERHOFER, H.; DELLAGO, C.; GEISLER, P. L. Biased sampling of Nonequilibrium Trajectories Can Fast Switching Simulations outperform Conventional Free energy Calculation Methods. *J. Phys. Chem. B*, v. 109, p. 6902–6915, 2005. Citado na página 99.
- 155 ARRAR, M. et al. On the Accurate Estimator of Free energies Using the Jarzynski Equality. *J. Comput. Chem.*, v. 40, p. 688–696, 2019. Citado na página 99.
- 156 YTREBERG, F. M.; SWENDSEN, R. H.; ZUCKERMAN, D. M. Comparison of free energy methods for molecular systems. *The Journal of Chemical Physics*, v. 125, p. 184114, 2006. Citado na página 99.
- 157 NELSON, P. *Física Biológica: Energia, Informação, Vida*. [S.l.]: Guanabara Koogan, 2006. Citado na página 47.
- 158 LAMBERT, F. L. Entropy is simple, Qualitatively. *Journal of Chemical Education*, v. 79, n. 10, p. 1241–1246, 2002. Citado 2 vezes nas páginas 47 e 104.
- 159 NUMMELA, J.; YASSIN, F.; ANDRICIOAEI, I. Entropy-energy decomposition from nonequilibrium work trajectories. *The Journal of Chemical Physics*, v. 128, p. 024104, 2008. Citado 2 vezes nas páginas 47 e 48.
- 160 LAMBERT, F. L. *Entropy and Constraint of Motion*. 2004. Disponível em: <[http://http://entropysite.oxy.edu/constraint\\_of\\_motion.html](http://http://entropysite.oxy.edu/constraint_of_motion.html)>. Citado na página 104.
- 161 JENSEN, W. B. Entropy and constraint of motion. *Journal of Chemical Education*, v. 81, n. 5, p. 639–640, 2004. Citado na página 104.

# Apêndices

# APÊNDICE A – Cálculo do trabalho microscópico

O cálculo do trabalho realizado entre dois pontos de uma trajetória é importante para se obter valores de energia livre, utilizando a Equação 5.7.

Seja um sistema com apenas uma coordenada de reação  $\xi$  dependente da posição dos átomos  $\xi(\mathbf{q})$ , sendo restringida por um potencial harmônico externo,

$$V[\xi(\mathbf{q}); \lambda(t)] = \frac{1}{2}k \left( \xi(\mathbf{q}) - \lambda(t) \right)^2 \quad (4.19)$$

onde  $k$  é a constante de força da restrição e  $\lambda(t)$  é o valor do parâmetro externo de controle dependente do tempo para esta coordenada de reação. A função Hamiltoniana do sistema, então, será:

$$\tilde{H}(\mathbf{q}, \mathbf{p}; \lambda) = H(\mathbf{q}, \mathbf{p}) + V[\xi(\mathbf{q}); \lambda(t)] \quad (A.1)$$

Ao pensarmos em um experimento de Dinâmica Molecular Dirigida, c.f. Seção 4.3.1, de forma a emular um experimento de microscopia de força atômica, c.f. Capítulo 3, podemos pensar em um afastamento linear em  $t$ , de forma que:

$$\lambda(t) = \lambda_0 + vt \quad (A.2)$$

onde  $v$  é a velocidade de afastamento do parâmetro externo de controle e  $\lambda_0$  é o valor do parametro externo no tempo  $t = 0$ .

Desta forma, o trabalho aplicado ao sistema no intervalo de tempo de 0 até o tempo  $t$  é dado por:(38, 96, 21, 106)

$$W_{0 \rightarrow t} = \int_0^t \frac{\partial \lambda(t')}{\partial t'} \left[ \frac{\partial \tilde{H}(\mathbf{q}, \mathbf{p}; \lambda)}{\partial \lambda} \right]_{(\mathbf{q}, \mathbf{p}; \lambda(t)) = (\mathbf{q}_t, \mathbf{p}_t; \lambda(t))} dt' \quad (6.2)$$

Utilizando as Equações A.1, 4.19 e A.2, se pode obter diretamente as derivadas do integrando da Equação 6.2:

$$\begin{aligned} \frac{\partial \lambda(t')}{\partial t'} &= \frac{d}{dt'} (\lambda_0 + vt') \\ &= v \end{aligned} \quad (A.3)$$

e

$$\begin{aligned}
 \frac{\partial \tilde{H}(\mathbf{q}, \mathbf{p}; \lambda)}{\partial \lambda} &= \frac{d}{d\lambda} \left[ H(\mathbf{q}, \mathbf{p}) + V[\xi(\mathbf{q}); \lambda(t)] \right] \\
 &= \frac{d}{d\lambda} \left[ V[\xi(\mathbf{q}); \lambda(t)] \right] \\
 &= \frac{d}{d\lambda} \left[ \frac{1}{2} k (\xi(\mathbf{q}) - \lambda(t))^2 \right] \\
 &= -k (\xi(\mathbf{q}) - \lambda(t))
 \end{aligned} \tag{A.4}$$

Substituindo as Equações A.3 e A.4 na Equação 6.2, obtemos:

$$W_{0 \rightarrow t} = \int_0^t -k (\xi(\mathbf{q}) - \lambda(t')) v dt' \tag{A.5}$$

Uma vez que  $k$  é independente do tempo e  $\lambda(t)$  é uma função do tempo, utilizando a Equação A.2, obtemos:

$$W_{0 \rightarrow t} = -k \int_0^t [\xi(\mathbf{q})v - \lambda_0 v - t'v^2] dt' \tag{A.6}$$

Colocando a velocidade de afastamento do parâmetro externo de controle  $v$  em evidência dentro do integrando da Equação A.6, e levando em conta que  $v$  é independente do tempo, obtemos finalmente a equação para o cálculo do trabalho:

$$W_{0 \rightarrow t} = -kv \int_0^t (\xi(\mathbf{q}) - \lambda(t')) dt' \tag{6.3}$$

APÊNDICE B – Artigo publicado no  
Journal of Chemical Physics

# Single molecule force spectroscopy of a streptomycin-binding RNA aptamer: An out-of-equilibrium molecular dynamics study

Cite as: *J. Chem. Phys.* **151**, 195102 (2019); doi: [10.1063/1.5128126](https://doi.org/10.1063/1.5128126)

Submitted: 17 September 2019 • Accepted: 31 October 2019 •

Published Online: 18 November 2019



View Online



Export Citation



CrossMark

Luis A. Baptista<sup>1</sup>  and Paulo A. Netz<sup>2,a)</sup> 

## AFFILIATIONS

<sup>1</sup>Programa de Pós-Graduação em Química, Instituto de Química, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

<sup>2</sup>Department of Physical Chemistry, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

<sup>a)</sup>Electronic mail: [netz@iq.ufrgs.br](mailto:netz@iq.ufrgs.br)

## ABSTRACT

Here, we investigate the unfolding behavior of a streptomycin-binding ribonucleic acid (RNA) aptamer under application of force in shear geometry. Using Langevin out-of-equilibrium simulations to emulate the single-molecule force spectroscopy (SMFS) experiment, we were able to understand the hierarchical unfolding process that occurs in the RNA molecule under application of stretching force and the influence of streptomycin modifying this unfolding. Subsequently, the application of the Jarzynski equality to the force profiles obtained in the pulling simulations shows that the free energies for individual systems and the difference of unfolding free energy upon streptomycin binding to the RNA free aptamer are in fair agreement with the experimental values, obtained through SMFS by Nick *et al.* [*J. Phys. Chem. B* **120**, 6479 (2016)].

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5128126>

## I. INTRODUCTION

Ribonucleic acids (RNAs) are important polynucleotides present in several cellular processes, from cellular replication to synthesis of peptides.<sup>1</sup> One of the major goals in the field of molecular biology is to understand the structure, dynamics, and interactions of RNA molecules, considering their interaction partners in these cellular processes.<sup>2,3</sup>

Streptomycin, an aminoglycoside antibiotic, is known to target the bacterial complex RNA structural domain.<sup>1,4</sup> Several techniques have been used to understand how aminoglycoside molecules are able to bind to specific RNA folding topologies; among them are Nuclear Magnetic Resonance (NMR) spectroscopic and crystallographic approaches.<sup>5</sup> Recently, Nick and co-workers<sup>6</sup> used computer simulations and single-molecule force spectroscopy (SMFS) to study the interaction of streptomycin with a specific

streptomycin-binding RNA fragment (aptamer). Using a combination of these two approaches, they found that streptomycin energetically stabilizes the RNA structure due to the additional hydrogen bonds formed between RNA and streptomycin, in a very specific pattern.

Steered molecular dynamics simulations (SMD), an out-of-equilibrium scheme of molecular dynamics simulation, have been used to help the understanding and to interpret the results of single molecule force spectroscopy for proteins.<sup>7–11</sup> Karttunen and co-workers used SMD to investigate the unfolding of a double strand DNA.<sup>12,13</sup> In this way, the use of computational simulations, such as SMD, provides a unique and invaluable tool to understand the unfolding process at the atomic level.<sup>12</sup>

The purpose of this work is to broaden the understanding of the interactions between RNA and streptomycin using Langevin out-of-equilibrium (pulling) computer simulations, in order to study the

unfolding process of RNA molecules as it occurs during dynamic SMFS experiments.

## II. SIMULATION DETAILS

### A. Molecular dynamics

All simulations were carried out with GROMACS version 4.5.5<sup>14,15</sup> using the AMBER PARMBSC0 force field<sup>16,17</sup> for the nucleic acid and the generalized AMBER force field (GAFF)<sup>18</sup> for the streptomycin, in a similar setup as employed by Nick and co-workers.<sup>6</sup> Since a huge number (3000) of simulations in a large simulation box were performed, in the present work, we used an implicit solvation model. With the purpose of reducing the computational effort of solvent interaction calculations, the interactions between the solute and the solvent were considered implicitly and the average of instantaneous interactions between them was modeled using the generalized Born formalism based on the Hawkins-Cramer-Truhlar method<sup>19,20</sup> with parameters of Tsui and Case.<sup>21</sup> Similar models were used by Karttunen and co-workers to model nucleic acid pulling experiments.<sup>12,13</sup>

The initial structure was taken from the crystallographic one deposited at the Protein Data Bank server under PDBid 1NTB<sup>5</sup> (Fig. 1). Two systems were constructed, one system called *complex* which corresponds to a molecule of streptomycin bonded to RNA and another system called *aptamer* which corresponds only to the RNA molecule where the streptomycin coordinates were removed manually.

After the systems were thermalized, pulling simulations were performed to mimic the single-molecule force spectroscopy (SMFS) experiment. In order to emulate the unfolding process of the RNA aptamer by an external force as carried out by Nick and co-workers,<sup>6</sup> the O5' atom of the C101 residue of RNA, the red sphere in Fig. 1,

was maintained fixed by a harmonic force of  $1000.0 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  and the O5' atom of the G1 residue, the green sphere in Fig. 1, was pulled parallel to the vector with origin in the O5' atom of the C101 residue and terminal point in the O5' atom of residue G1, under shear geometry, with a string force constant of  $200 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  and pulling velocities equal to 0.02, 0.03, 0.04, 0.05, 0.06, and  $0.07 \text{ nm ps}^{-1}$ . For each system, *aptamer* and *complex*, respectively, and for each pull rate, 500 pulling simulations were carried out until the end-to-end distance was bigger than 35.0 nm. At the beginning of each new pulling simulation, a set of velocities was calculated to assert the Jarzynski equality condition of independent initial states.

The structural analysis was performed with DSSR version 1.2.6<sup>22-24</sup> and homemade Python software.

### B. Free energy calculations

The work was calculated based on stiff-spring approximation, as described elsewhere.<sup>25</sup> Applying an external potential  $V[\mathbf{r}; \lambda(t)] = \frac{k}{2} [\xi(\mathbf{r}) - \lambda(t)]^2$ , where  $\xi(\mathbf{r})$  is the reaction coordinate and  $\lambda(t)$  is the external parameter of control, the Hamiltonian of the system is then described by  $\hat{H}(\mathbf{r}, \mathbf{p}; \lambda) = H(\mathbf{r}, \mathbf{p}) + V(\mathbf{r}; \lambda)$ . Thus, the work is given by

$$W_{0 \rightarrow t} = -k v \int_0^t dt' [\xi(\mathbf{r}(t')) - \lambda(t')]. \quad (1)$$

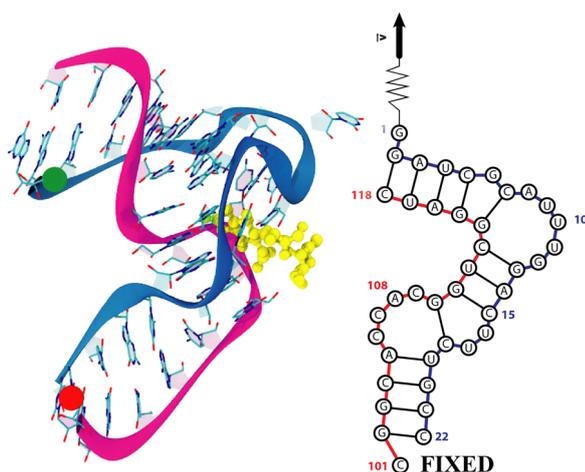
The free energies were obtained through linear extrapolations of block averages, as described elsewhere by Ytreberg and Zuckerman<sup>26</sup> of the standard nonequilibrium estimates of Jarzynski's equality<sup>27,28</sup>  $e^{-\beta \Delta F} = \langle e^{-\beta W} \rangle$ . The average free energies were plotted as a function of  $\chi = 1/n$ ,<sup>0,5</sup> where  $n$  is the size of blocks, and extrapolated to  $\chi = 0$  by linear least squares approximation. The average free energies, for each value of  $n$ , were obtained choosing 500 blocks with  $n$  values at random with replacement.

## III. RESULTS

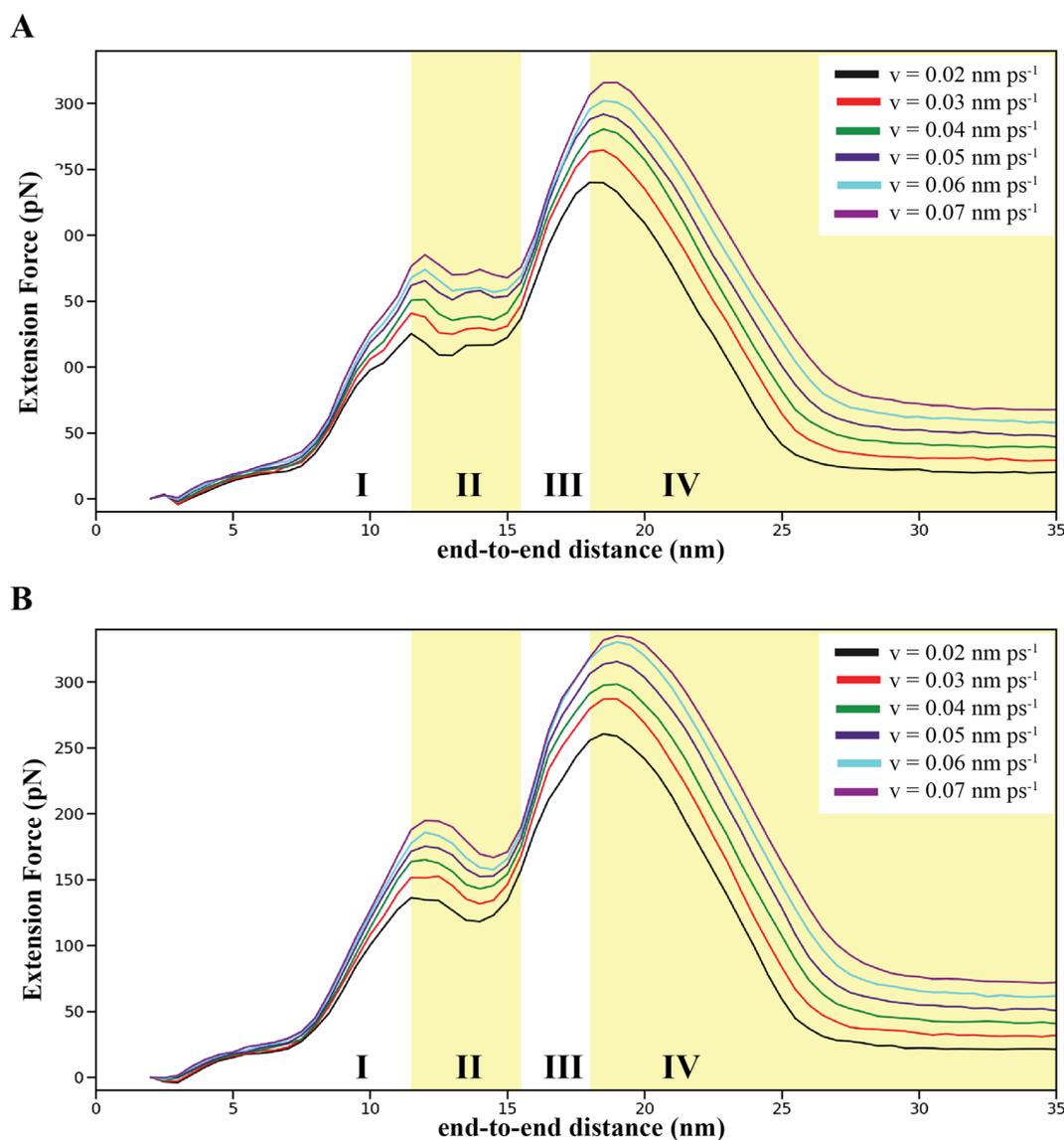
### A. Structural

It is known that the kinetics and the energetic behavior of nonequilibrium molecular dynamics experiments are dependent on the pulling velocity.<sup>29-34</sup> In order to investigate the behavior of the force-extension curves (FEC) and the structural motions with the pulling velocity, we carried out steered molecular dynamics simulations with pulling velocities equal to 0.02, 0.03, 0.04, 0.05, 0.06, and  $0.07 \text{ nm ps}^{-1}$ . Figure 2 shows the average force-extension curves for the pulling velocities for both systems studied. Both systems show FEC with a biphasic behavior that have a plateau between two linear regimes with different slopes. These FEC should be divided into four regions (described below). The last force-jump, between regions III and IV, determines the rupture point.

The differences seem rather small, but they are consistent among the simulation series. As we shall see below, the proper averaging over many simulations could reveal important features. Figure S1, in the [supplementary material](#), shows the value of the force extension at the rupture point as a function of the pulling rate for both systems. One should see that there is a moderate increase in the average force at the rupture point with the increase



**FIG. 1.** Crystallographic structure and diagram of the secondary structure of 1NTB. The streptomycin is shown in the crystallographic structure in yellow and omitted in the diagram. The red sphere corresponds to the O5' atom of the C101 residue (fixed), and the green sphere corresponds to the O5' atom of the G1 residue (pulled).

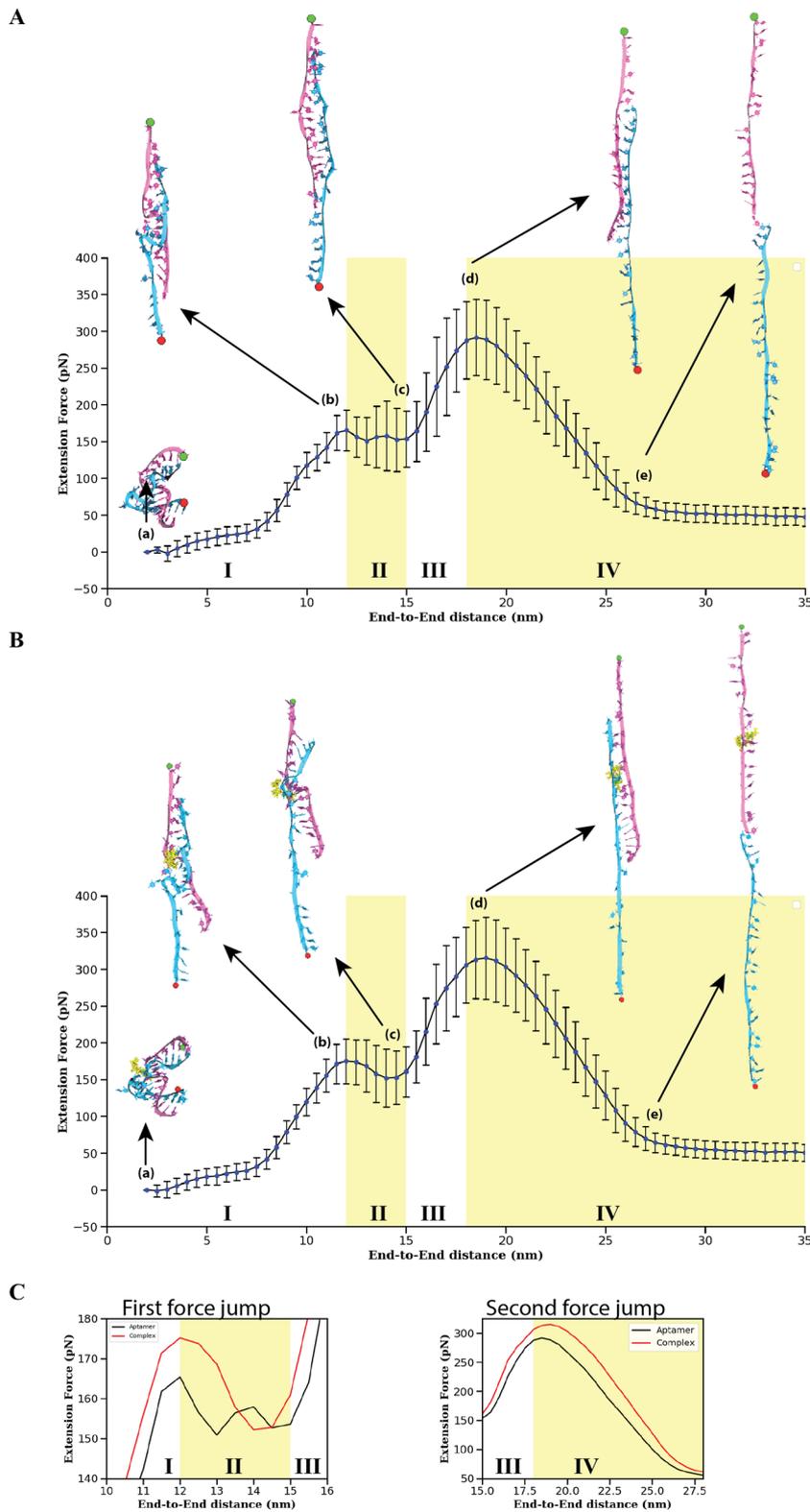


**FIG. 2.** Average force-extension curves for *aptamer* (a) and *complex* (b) systems in different pulling velocities. The white and yellow areas, identified by roman numerals, indicate the regions discussed in the text.

of the pulling velocity. This trend is in agreement with the behavior already reported for the unfolding process by SMFS with other molecules.<sup>7,35,36</sup> Although the direct comparison between the extension forces provided by SMFS and SMD must be done with care,<sup>7</sup> at least a qualitative comparison of the unfolding behavior of a macromolecule and an estimate of the influence of a small molecule on this behavior is possible, considering additionally that the pulling rate effects are roughly of the same magnitude in both systems, with and without the small ligand.

Figure 3 shows the average force-extension curves (FEC) obtained by the average of pulling simulations for shear geometry of RNA strands for both systems, for the velocity  $0.05 \text{ nm ps}^{-1}$ . The

average FEC for both systems show approximately the same behavior with two force-jumps. Figure 3(A) shows the average FEC for pulling of the *aptamer* system. For this system, one should identify four distinct regions [indicated by roman numerals in Fig. 3(A)]: (I) strong increase of force with the increase of end-to-end (*ee*) RNA distance until 12 nm, approximately, which ends in the first force-jump; (II) between 12 nm and 15 nm of end-to-end distance, the force remains almost constant; (III) between 15 nm and 18 nm, there is a linear increase in the force until a maximum extension force is reached and the second force-jump takes place. This maximum extension force would correspond to the rupture force in SMFS experiments; and (IV) after the rupture point, around 18 nm, there



**FIG. 3.** Average FEC ( $v = 0.05 \text{ nm ps}^{-1}$ ) with representative structures for end-to-end pulling of the *aptamer* (A) and *complex* (B) systems. The white and yellow areas in the FEC, identified by roman numerals, indicate the regions discussed in the text. Plates A and B: (a) initial structure; (b) elongated structure with binding pocket still conserved; (c) after binding pocket disruption; (d) ladder form at the rupture point; (e) structure at the last contact between two strands. Plate C: zoomed-in view of the FEC force-jumps.

is a linear decrease in the force with the increase of *ee* distance until the force reaches a constant value.

In order to better understand the behavior of RNA strands being pulled under shear geometry, some representative structures along the pulling path are shown in the average FEC in Fig. 3(A). Figure 3(A-a) shows the initial structure with RNA folded in the equilibrium state. Figure 3(A-b) shows the representative structure for the end of region I where the RNA structure was elongated; however, the core of the structure remains intact. The end of region II is represented in Fig. 3(A-c), where the binding pocket was completely unstructured. After the end of region II, there is a linear increase in the force with the *ee* distance (region III) until the RNA structure assumes an untwisted ladder form [Fig. 3(A-d)] and the last force-jump occurs. After this second force-jump, the force decreases linearly. This behavior can be explained by considering that, in this form of untwisted ladder, one chain can slide over the other with minimal interactions. Figure 3(A-e) shows the structure just before the two strands lose their close contacts.

Figure 3(B) shows the average FEC for the streptomycin/RNA complex and the four distinct regions similar to the previous system. Regions I, III, and IV display a force behavior similar to the *aptamer* system. However, in region II, between 12 nm and 15 nm and after the first force-jump, the force does not remain constant, with the increase of *ee* distance, but decreases, as shown in Fig. 3(C) for the first force-jump. The representative structures for the pulling experiment of the *complex* system are inserted in the average FEC [Fig. 3(B)]. Due to the presence of streptomycin, the structure at the end of region I [Fig. 3(B-b)] shows a still completely conserved binding pocket, which implies the stabilization effect of this ligand. Even after the force-jump at the FEC, between regions I and II, it was possible to see remnants of the binding pocket [Fig. 3(B-c)]. This stabilization effect due to presence of streptomycin is evident in the comparison of region II between the *aptamer* and *complex* systems in Fig. 2. One should see that, for the *complex* system, the first force-jump region has a clear sigmoidal shape, unlike the case of the same region for the *aptamer* system.

This difference can be seen more clearly in Plate C of Fig. 3, which shows the average FEC in the region of the two force-jumps. One should note that the average force increased 9.8 pN, in the first force-jump, and 23.6 pN, in the second force-jump, from the *aptamer* system to the *complex* system. There is some overlap in the error bars, which can decrease with a larger number of simulations; nevertheless, the qualitative differences are noticeable. This behavior is in fair agreement with experimental results which showed that the presence of streptomycin led to an increase of  $9.2 \pm 3.5$  pN in the average rupture force.<sup>6</sup> The relative increase in the force at the first force-jump is due to the strong interaction between the two strands and streptomycin in the structured binding pocket, mainly by several hydrogen bonds made by the streptomycin molecule as detailed below.<sup>6</sup> The increase in the force at the second force-jump indicates that streptomycin induces additional molecular interactions, even after the disruption of the binding pocket, which render the complex less labile upon application of an external force. Nick *et al.*<sup>6</sup> analyzed the hydrogen bond pattern using classical simulations, mapping the key residues making hydrogen bonds with the streptomycin. We analyzed the presence of such hydrogen bonds

along the simulations. Figure 4 shows the pattern of hydrogen bonds between the streptomycin and the nucleotides of the binding pocket as a function of end-to-end distance. Streptomycin makes hydrogen bonds with the nucleotides of the binding pocket, specially with nucleotides U16 and U17 of strand 1 and C109 and G110 of strand 2 until the end-to-end distance of about 15 nm; in other words, the streptomycin maintains hydrogen bonds with up to 4 nucleotides until near the rupture point, stabilizing therefore the complex.

Figure 5 shows the average structural parameters evaluated for the two systems along the pulling simulations. In Fig. 5(a), the number of base pairs is shown. Until the *ee* distance of about 10 nm, the number of base pairs in both systems remains almost constant and is slightly higher in the *aptamer* system than in the *complex* system. In this domain, which corresponds to region I in the FEC of both systems, the RNA molecule is partially elongated [Fig. 3(A-b)]; however, the tertiary structure of the binding pocket around the streptomycin molecule is maintained [Figs. 3(B-a) and 3(B-c)]. This behavior is even more clear for the number of multiplets, the coplanar arrangement of three or more bases interacting via hydrogen bonding interaction,<sup>37</sup> as shown in Fig. 5(b). The reason to the number of base pairs and multiplets being slightly higher in the *aptamer* system than in the *complex* system is the presence of the streptomycin molecule in the binding pocket. One should see in Fig. S2, in the supplementary material, that, besides the two triplets (G6-G13-C113 and U16-C18-G110) described in the crystallographic structure,<sup>5</sup> the absence of the streptomycin molecule allows the two more triplets to be formed (U19-A105-C106 and A8-A14-U12). Also, the absence of the streptomycin molecule enables the two strands to interact directly and form a slightly higher number of base pairs. After the *ee* distance of 10 nm, the base pairs begin to break until the rupture point is reached. This domain corresponds to the total disruption of the tertiary structure of the RNA binding pocket and, consequently, linearization of the two strands, as shown in Fig. 3(A-c).

Figure 5(c) shows the average buried surface area (BSA) for both systems as a function of the end-to-end distance. Until the *ee* distance of about 10 nm, the BSA remains almost constant. This is expected since in this region the RNA molecule is undergoing a process of elongation and the base pairs do not start to break yet.

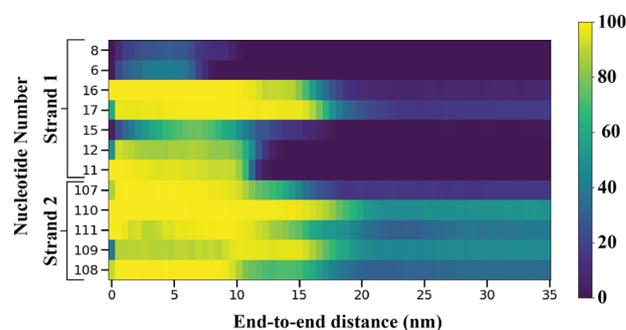
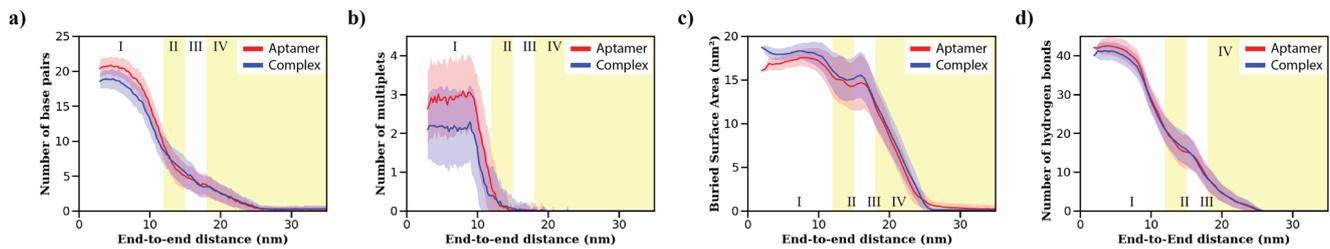


FIG. 4. Pattern of hydrogen bonds between nucleotides and streptomycin as a function of end-to-end distance. The colors represent the percentage of bonds over all simulations.



**FIG. 5.** Average structural parameters for the pulling simulations. The white and yellow areas, identified by roman numerals, indicate the regions discussed in the text. (a) Number of base pairs. (b) Number of multiplets. (c) Buried surface area. (d) Number of hydrogen bonds between the two strands.

However, after the system reaches this  $ee$  distance, the BSA decreases until the point at 26 nm where there is no contact between the two RNA chains.

Figure 5(d) shows the average number of hydrogen bonds between the two strands in the systems. This number decreases linearly with the end-to-end distance until it reaches zero, at the  $ee$  distance of approximately 26 nm for both systems. This decrease is expected since the two ends of the structure were elongated and so the correct geometry for these interactions is no longer maintained.

According to Fig. 5(d), the number of interstrand hydrogen bonds in the *aptamer* system is slightly higher than that in the *complex* system because of the presence of the streptomycin molecule, which hinders the direct interaction between the two strands. However, if we take into account the intrastrand hydrogen bonds as well (i.e., the total number of RNA hydrogen bonds; Fig. S3 in the [supplementary material](#)), the initial number of hydrogen bonds is indeed higher in the *complex* system, in agreement with the equilibrium molecular dynamics simulations performed by Nick *et al.* However, in the very beginning of the disruption of the binding pocket, there is an increase in the number of hydrogen bonds in the *aptamer* system due to the approximation of two RNA strands. Once this approximation is hindered by the presence of the streptomycin molecule in the *complex* system, the number of the total hydrogen bonds remains constant in region I.

**TABLE I.** Free energy calculated at the rupture point in function of pulling velocity.

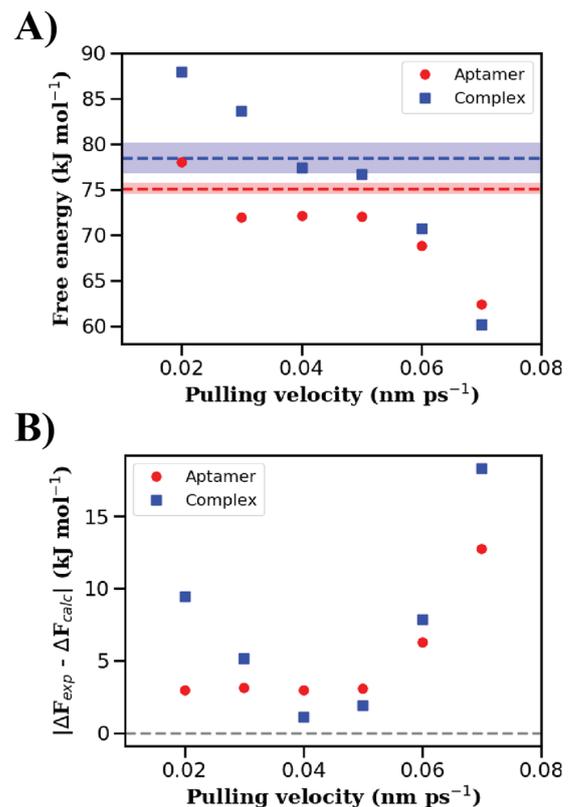
$v$ (nm ps <sup>-1</sup> )	Free-energy (kJ mol <sup>-1</sup> )		
	<i>Aptamer</i> system <sup>a</sup>	<i>Complex</i> system <sup>a</sup>	$\Delta\Delta F^{a,b}$
0.02	78.04(0.02)	87.96(0.14)	-9.91(0.14)
0.03	71.96(0.07)	83.65(0.10)	-11.69(0.13)
0.04	72.12(0.07)	77.38(0.14)	-5.26(0.15)
0.05	72.02(0.23)	76.63(0.11)	-4.61(0.25)
0.06	68.84(0.05)	70.64(0.35)	-1.80(0.35)
0.07	62.37(0.35)	60.18(0.12)	2.18(0.37)

<sup>a</sup>Standard deviations are given within parentheses.

<sup>b</sup> $\Delta\Delta F = \Delta F_{\text{aptamer}} - \Delta F_{\text{complex}}$ .

## B. Energetics

The use of the stiff-spring approximation for the calculation of the work performed upon the system requires that the end-to-end distance  $\xi$  closely follows the constraint center of pulling  $\lambda$ .<sup>25</sup> Figure S4, in the [supplementary material](#), shows the histogram of the difference between  $\xi$  and  $\lambda$  for pulling velocities equal to 0.02, 0.04, and 0.07 nm ps<sup>-1</sup> for both systems. One should see that, for both



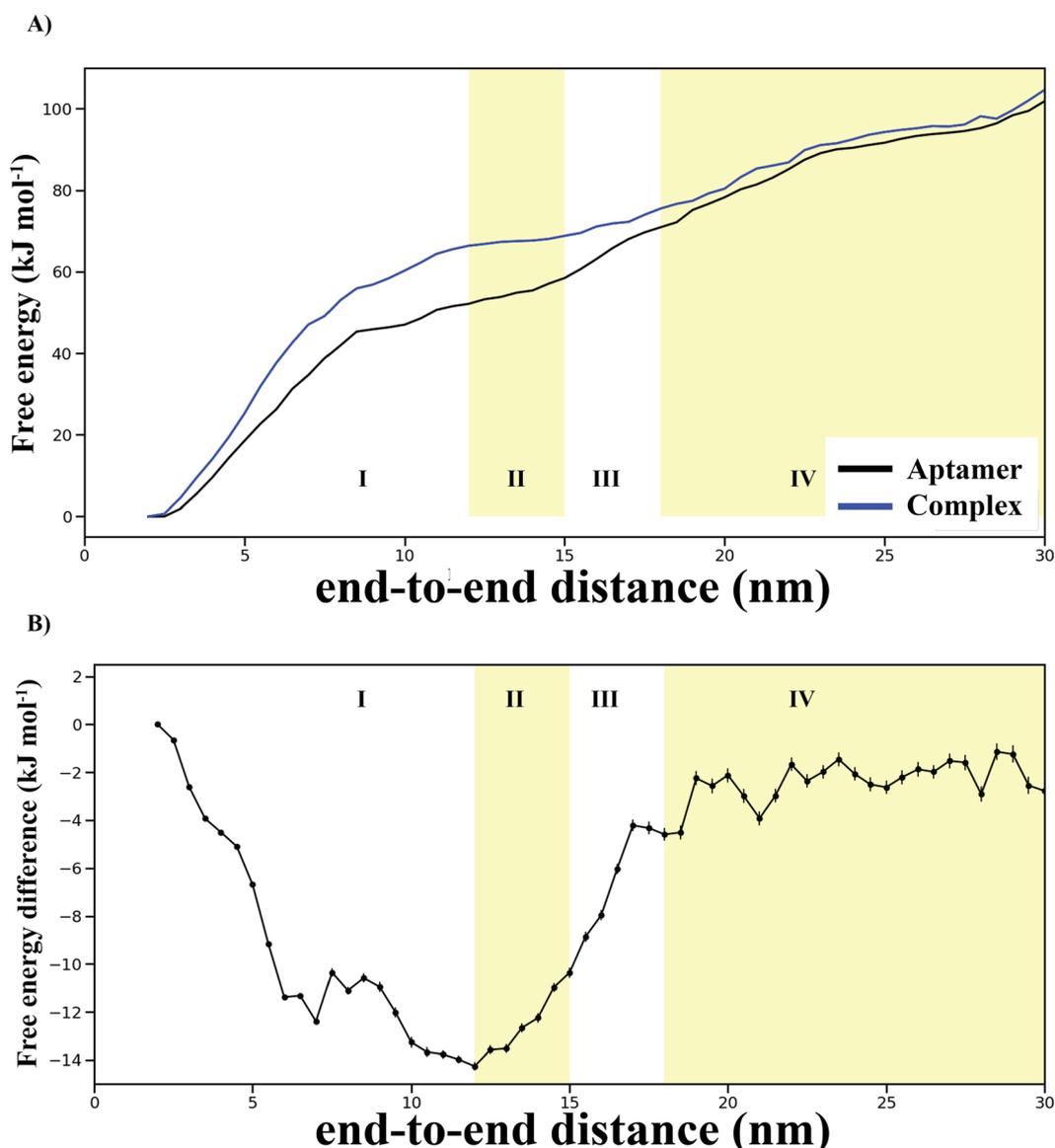
**FIG. 6.** (a) Free energy as a function of pulling velocity. The dashed line corresponds to the experimental values obtained by Nick *et al.*, and the shadow area corresponds to the standard deviation interval. (b) Modulus of the difference between the calculated and experimental free energies.

systems and pulling rate in the range of  $0.02\text{--}0.07\text{ nm ps}^{-1}$ , the distribution of the difference between  $\xi$  and  $\lambda$  is centred below  $0.05\text{ nm}$ , indicating that the end-to-end distance closely follows the constraint center.

Table I and Fig. 6(a) show the free energies for the *aptamer* and *complex* systems calculated at the rupture point. One should see that, for both systems, *aptamer* and *complex*, the free energy had an overall decrease with the increase in the pulling velocity from  $0.02$  to  $0.07\text{ nm ps}^{-1}$ . Nick *et al.* found experimentally, by single-molecule force spectroscopy, free energies equal  $75.1 \pm 0.5\text{ kJ mol}^{-1}$  and  $78.5 \pm 1.7\text{ kJ mol}^{-1}$  for *aptamer* and *complex* systems, respectively.

Comparing these experimental values with the calculated values of the free energies in Table I, we could conclude that the agreement is fair. Moreover, as shown in Fig. 6, for both systems, one should see that, for the *aptamer* system, the difference between simulated and experimental values stays below  $5\text{ kJ mol}^{-1}$  for the pulling velocity in the range of  $0.02\text{--}0.05\text{ nm ps}^{-1}$ , and for the *complex* system, this difference has a parabolic behavior with a minimum value at the pulling velocity of  $0.04\text{ nm ps}^{-1}$ .

The unfolding free energy difference upon streptomycin binding should be defined as  $\Delta\Delta F = \Delta F_{\text{aptamer}} - \Delta F_{\text{complex}}$ .<sup>6</sup> Nick *et al.* found that this difference equals  $-3.4 \pm 1.8\text{ kJ mol}^{-1}$ . Although



**FIG. 7.** (a) Free energy profile for aptamer (black line) and complex (blue line) systems, respectively, with pulling velocity equal to  $0.04\text{ nm ps}^{-1}$ . (b) Free energy upon streptomycin binding  $\Delta\Delta F = \Delta F_{\text{aptamer}} - \Delta F_{\text{complex}}$ . The white and yellow areas, identified by roman numerals, indicate the regions discussed in the text.

the calculated free energy difference of this forced RNA unfolding, being a nonequilibrium process, is not the same as the equilibrium free energy of the streptomycin binding to the aptamer, it corresponds to the experimental quantity measured by the nonequilibrium experiments and also is an estimate in the same spirit of the pioneering works on the experimental verification of the Jarzynski equality.<sup>38–44</sup> The last column of Table I shows the values of  $\Delta\Delta F$  in function of the pulling velocities studied. It is possible to see in Table I that the free energy difference for the slowest pulling is about  $-10 \text{ kJ mol}^{-1}$ , in fair agreement with the experimental value.

Many factors can contribute to the difference between experimental and calculated free energies, such as the pulling rate, as mentioned above in the initial part of this section. Another important fact is that Jarzynski estimator can lead to unwanted bias in the calculated free energies.<sup>45–48</sup> Although there are better estimators, like Bennett averaging of Jarzynski work values,<sup>45</sup> these methods take into consideration the forward and backward nonequilibrium paths between two states. Indeed, it was shown that methods based on the sampling of forward and reverse paths yield better estimates of free energy.<sup>46,47</sup> However, due to the nature of our double strand RNA system, the refolding of the two strands, which corresponds to the backward path, is very unlikely to happen in a computational simulation. In order to minimize the cited bias in the Jarzynski estimator, we used the linear extrapolation scheme proposed by Ytreberg and Zuckerman<sup>26</sup> where the free energy is estimated in the limit of a set of infinite simulations. This scheme explores the systematic behavior in the bias of the estimator, as pointed out by Ytreberg and Zuckerman,<sup>26</sup> in order to improve the free energy obtained. Moreover the experimental apparatus has additional elements (e.g., *beads, handles*, solvent, and others) that are not part of our simplified model. Once these are inseparable part of any pulling experiment, they have an influence on the energetics of this.<sup>30</sup> Nevertheless, our detailed microscopic analysis allowed us to identify some key aspects of the rupture mechanism.

Figure 7(a) shows the free energy profile as a function of the end-to-end distance for the pulling simulations with a pulling velocity of  $0.04 \text{ nm ps}^{-1}$ . The standard deviations of the free energy, calculated by block average, are small for both systems, less than  $0.50 \text{ kJ mol}^{-1}$  or  $0.20 k_B T$ . It should be noted that, in the region from  $7.0 \text{ nm}$  to  $15 \text{ nm}$ , there is a significant difference between the free energy profile of *aptamer* and *complex* systems. This region was precisely where the disruption of the tertiary structure of the RNA binding pocket occurred, and in this region, the differences between both systems are more pronounced, for instance, regarding the hydrogen bond pattern in the binding pocket. The disruption of the binding pocket for the *complex* system requires more energy compared to that for the *aptamer* system. This fact is in agreement with experimental results where an increase of the rupture force of  $9.2 \pm 3.5 \text{ pN}$  was measured at the retract velocity of  $100 \text{ nm s}^{-1}$ .<sup>6</sup> This is expected since the streptomycin molecule stabilizes the binding pocket. From  $15 \text{ nm}$  to the end, both systems behave similarly, where both profiles present almost the same linear pattern. This linear pattern formed, in the free energy profile, after the end-to-end distance of  $15 \text{ nm}$  is due to extra work done in the system by the pulling of the molecular apparatus against friction. However, for the free energy upon streptomycin binding [Fig. 7(b)] after the last force-jump, region IV,  $\Delta\Delta F$  fluctuates around a constant value, indicating

that our model, at large values of end-to-end distance, could average out the nonequilibrium effect of extra work and, thus, recover the correct equilibrium free energy limit.

#### IV. CONCLUSIONS

The application of a large number of Langevin out-of-equilibrium molecular dynamics simulations allowed us to improve the understanding of the unfolding behavior of the streptomycin RNA aptamer under application of force, similar to SMFS experiments. The structural analyses show that there is a small but noticeable difference between the pulling of *aptamer* and *complex* systems. Our results show that there is a hierarchical order in the unfolding process, where the binding pocket remains structured during the first step of this process and four hydrogen bonds between streptomycin and RNA remain conserved until the rupture. Due to the elongation, the RNA strands assume a form of an untwisted ladder and one chain slides over the other until they lose contact. The unfolding free energies calculated for each system as well as the unfolding free energy difference upon streptomycin binding were in fair agreement with the experimental quantities. The inclusion of apparatus-specific parameters in the model or the increase in the number of simulations or further slowing of the pulling speed would possibly improve the agreement of the calculated free energies with the experimental values, but with a much higher computational cost.

#### SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for figures of the average extension force at the rupture point for several pulling rates, representative initial structures for complex and aptamer systems, number of total RNA hydrogen bonds as a function of end-to-end distance, and difference between end-to-end distance and constraint center of pulling.

#### ACKNOWLEDGMENTS

The authors thank Dr. Rüdiger Berger, Thomas Nick, and Tiago E. Oliveira for the critical reading of the manuscript and valuable suggestions. P.A.N. acknowledges the CNPq for financial support. L.A.B. acknowledges the CAPES for financial support (Financial code 001).

#### REFERENCES

- <sup>1</sup>D. Voet and J. G. Voet, *Biochemistry*, 4th ed. (John Wiley & Sons, New York, 2010).
- <sup>2</sup>P. Dua, S. Kim, and D. Lee, "Nucleic acid aptamers targeting cell-surface proteins," *Methods* **54**, 215–225 (2011).
- <sup>3</sup>N. Windbichler and R. Schroeder, "Isolation of specific RNA-binding proteins using the streptomycin-binding RNA aptamer," *Nat. Protoc.* **1**, 637–640 (2006).
- <sup>4</sup>D. Sharma, A. R. Cukras, E. J. Rogers, D. R. Southworth, and R. Gree, "Mutational analysis of S12 protein and implications for the accuracy of decoding by the ribosome," *J. Mol. Biol.* **374**, 1065–1076 (2007).
- <sup>5</sup>V. Tereshko, E. Skripkin, and D. J. Patel, "Encapsulating streptomycin within a small 40-mer RNA," *Chem. Biol.* **10**, 175–187 (2003).

- <sup>6</sup>T. A. Nick, T. E. de Oliveira, D. W. Pilat, F. Spenkuch, H. Butt, M. Helm, P. A. Netz, and R. Berger, "Stability of a split streptomycin binding aptamer," *J. Phys. Chem. B* **120**, 6479–6489 (2016).
- <sup>7</sup>F. Rico, L. Gonzalez, I. Casuso, M. Puig-Vidal, and S. Scheuring, "High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations," *Science* **342**, 741–743 (2013).
- <sup>8</sup>A. V. Glyakina, N. K. Balabaev, and O. V. Galzitskaya, "Mechanical unfolding of proteins L and G with constant force: Similarities and differences," *J. Chem. Phys.* **131**, 045102 (2009).
- <sup>9</sup>E. H. Lee, J. Hsin, M. Sotomayor, G. Comellas, and K. Schulten, "Discovery through the computational microscope," *Structure* **17**, 1295–1306 (1999).
- <sup>10</sup>H. Lu, B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten, "Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation," *Biophys. J.* **75**, 662–671 (1998).
- <sup>11</sup>H. Grubmüller, B. Heymann, and P. Tavan, "Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force," *Science* **271**, 997–999 (1996).
- <sup>12</sup>A. M. Naserian-Nik, M. Tahani, and M. Karttunen, "Pulling of double-stranded DNA by atomic force microscopy: A simulation in atomistic details," *RSC Adv.* **3**, 10516–10528 (2013).
- <sup>13</sup>A. M. Naserian-Nik, M. Tahani, and M. Karttunen, "Molecular dynamics study of DNA oligomers under angled pulling," *RSC Adv.* **4**, 10751–10760 (2014).
- <sup>14</sup>B. Hess, C. Kutzner, D. van der Spoel, B. Hess, and E. Lindahl, "GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation," *J. Chem. Theory Comput.* **4**, 435–447 (2008).
- <sup>15</sup>S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl, "GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit," *Bioinformatics* **29**, 845–854 (2013).
- <sup>16</sup>A. Pérez, I. Marchán, D. Svozil, J. Spöner, T. E. Cheatham, C. A. Laughton, and M. Orozco, "Refinement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers," *Biophys. J.* **92**, 3817–3829 (2007).
- <sup>17</sup>A. T. Guy, T. J. Piggot, and S. Khalid, "Single-stranded DNA within nanopores: Conformational dynamics and implications for sequencing: A molecular dynamics simulation study," *Biophys. J.* **103**, 1028–1036 (2012).
- <sup>18</sup>J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general amber force field," *J. Comput. Chem.* **26**, 114 (2005).
- <sup>19</sup>G. Hawkins, C. Cramer, and D. Truhlar, "Pairwise solute descreening of solute charges from a dielectric medium," *Chem. Phys. Lett.* **246**, 122–129 (1995).
- <sup>20</sup>G. Hawkins, C. Cramer, and D. Truhlar, "Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium," *J. Phys. Chem.* **100**, 19824–19839 (1996).
- <sup>21</sup>V. Tsui and D. Case, "Theory and applications of the generalized born solvation model in macromolecular simulations," *Biopolymers* **56**, 275–291 (2001).
- <sup>22</sup>X. Lu, W. K. Olson, and H. J. Bussemaker, "The RNA backbone plays a crucial role in mediating the intrinsic stability of the GpU dinucleotide platform and the GpU/GpA mini duplex," *Nucleic Acids Res.* **38**, 4868–4876 (2010).
- <sup>23</sup>X. Lu and W. K. Olson, "3DNA: A versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures," *Nat. Protoc.* **3**, 1213–1227 (2008).
- <sup>24</sup>X. Lu and W. K. Olson, "3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures," *Nucleic Acids Res.* **31**, 5108–5121 (2003).
- <sup>25</sup>S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten, "Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality," *J. Chem. Phys.* **119**, 3559–3566 (2003).
- <sup>26</sup>F. M. Ytreberg and D. M. Zuckerman, "Efficient use of nonequilibrium measurement to estimate free energy differences for molecular systems," *J. Comput. Chem.* **25**, 1749–1759 (2004).
- <sup>27</sup>C. Jarzynski, "Nonequilibrium equality for free energy differences," *Phys. Rev. Lett.* **78**, 2690–2693 (1997).
- <sup>28</sup>C. Jarzynski, "Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach," *Phys. Rev. E* **56**, 5018–5035 (1997).
- <sup>29</sup>A. Nunes-Alves and G. M. Arantes, "Mechanical unfolding of macromolecules coupled to bond dissociation," *J. Chem. Theory Comput.* **14**, 282–290 (2018).
- <sup>30</sup>M. Manosas and F. Ritort, "Thermodynamic and kinetic aspects of RNA pulling experiments," *Biophys. J.* **88**, 3224–3242 (2005).
- <sup>31</sup>G. Hummer and A. Szabo, "Kinetics from nonequilibrium single-molecule pulling experiments," *Biophys. J.* **85**, 5–15 (2003).
- <sup>32</sup>D. A. Hendrix and C. Jarzynski, "A 'fast growth' method of computing free energy differences," *J. Chem. Phys.* **114**, 5974–5981 (2001).
- <sup>33</sup>S. Sheridan, F. Gräter, and C. Daday, "How fast is too fast in force-probe molecular dynamics simulations?," *J. Phys. Chem. B* **123**, 3658–3664 (2019).
- <sup>34</sup>D. K. West, P. D. Olmsted, and E. Paci, "Free energy of protein folding from nonequilibrium simulations using the Jarzynski equality," *J. Chem. Phys.* **125**, 204910 (2006).
- <sup>35</sup>M. Carrion-Vazquez, A. F. Oberhauser, S. B. Fowler, P. E. Marszalek, S. E. Broedel, J. Clarke, and J. M. Fernandez, "Mechanical and chemical unfolding of a single protein: A comparison," *Proc. Natl. Acad. Sci. U. S. A.* **96**, 3694–3699 (1999).
- <sup>36</sup>M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, "Reversible unfolding of individual titin immunoglobulin domains by AFM," *Science* **276**, 1109–1112 (1997).
- <sup>37</sup>X.-J. Lu, Detection of multiplets in DSSR; accessed 7 July 2018.
- <sup>38</sup>C. Bustamante, S. B. Smith, J. Liphardt, and D. Smith, "Single-molecule studies of DNA mechanics," *Curr. Opin. Struct. Biol.* **10**, 279–285 (2000).
- <sup>39</sup>J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, "Reversible unfolding of single RNA molecules by mechanical force," *Science* **292**, 733–737 (2001).
- <sup>40</sup>J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, "Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality," *Science* **296**, 1832–1835 (2002).
- <sup>41</sup>B. Onoa, S. Dumont, J. Liphardt, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, "Identifying kinetic barriers to mechanical unfolding of the *T. thermophila* ribozyme," *Science* **299**, 1892–1895 (2003).
- <sup>42</sup>B. Onoa and I. Tinoco, Jr., "RNA folding and unfolding," *Curr. Opin. Struct. Biol.* **14**, 374–379 (2004).
- <sup>43</sup>D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, "Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies," *Nature* **437**, 231–234 (2005).
- <sup>44</sup>P. T. X. Li, D. Collin, S. B. Smith, C. Bustamante, and I. Tinoco, Jr., "Probing the mechanical folding kinetics of TAR RNA by hopping, force-jump, and force-ramp methods," *Biophys. J.* **90**, 250–260 (2006).
- <sup>45</sup>F. M. Ytreberg, R. H. Swendsen, and D. M. Zuckerman, "Comparison of free energy methods for molecular systems," *J. Chem. Phys.* **125**, 184114 (2006).
- <sup>46</sup>M. Arrar, F. M. Boubeta, M. E. Szretter, M. Sued, L. Boechi, and D. Rodriguez, "On the accurate estimator of free energies using the Jarzynski equality," *J. Comput. Chem.* **40**, 688–696 (2019).
- <sup>47</sup>L. Y. Chen, D. A. Bastien, and H. E. Espeje, "Determination of equilibrium free energy from nonequilibrium work measurements," *Phys. Chem. Chem. Phys.* **12**, 6579–6582 (2010).
- <sup>48</sup>H. Oberhofer, C. Dellago, and P. L. Geissler, "Biased sampling of nonequilibrium trajectories: Can fast switching simulations outperform conventional free energy calculation methods?," *J. Phys. Chem. B* **109**, 6902–6915 (2005).

Supporting information of Single Molecule Force  
Spectroscopy of a Streptomycin-Binding RNA  
Aptamer: An Out-of-Equilibrium Molecular  
Dynamics Study.

Luis A. Baptista and Paulo A. Netz

September 9, 2019

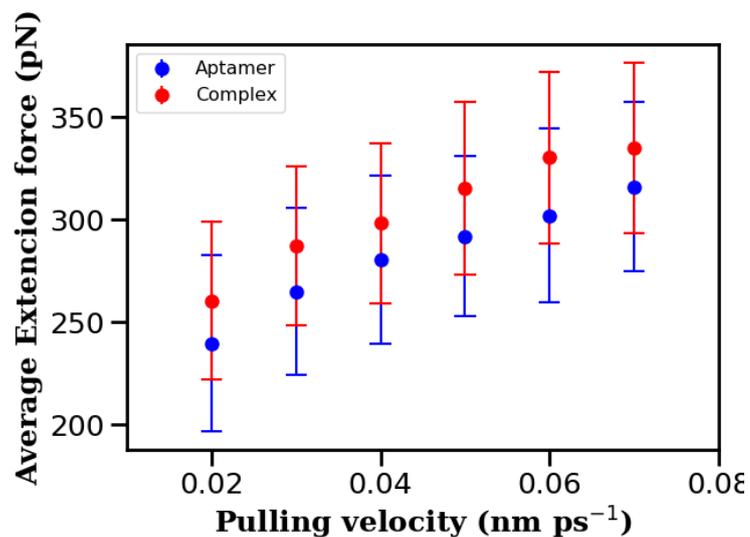


Figure S1: Average extension force at the rupture point in function of the pulling rate for *aptamer* and *complex* systems.

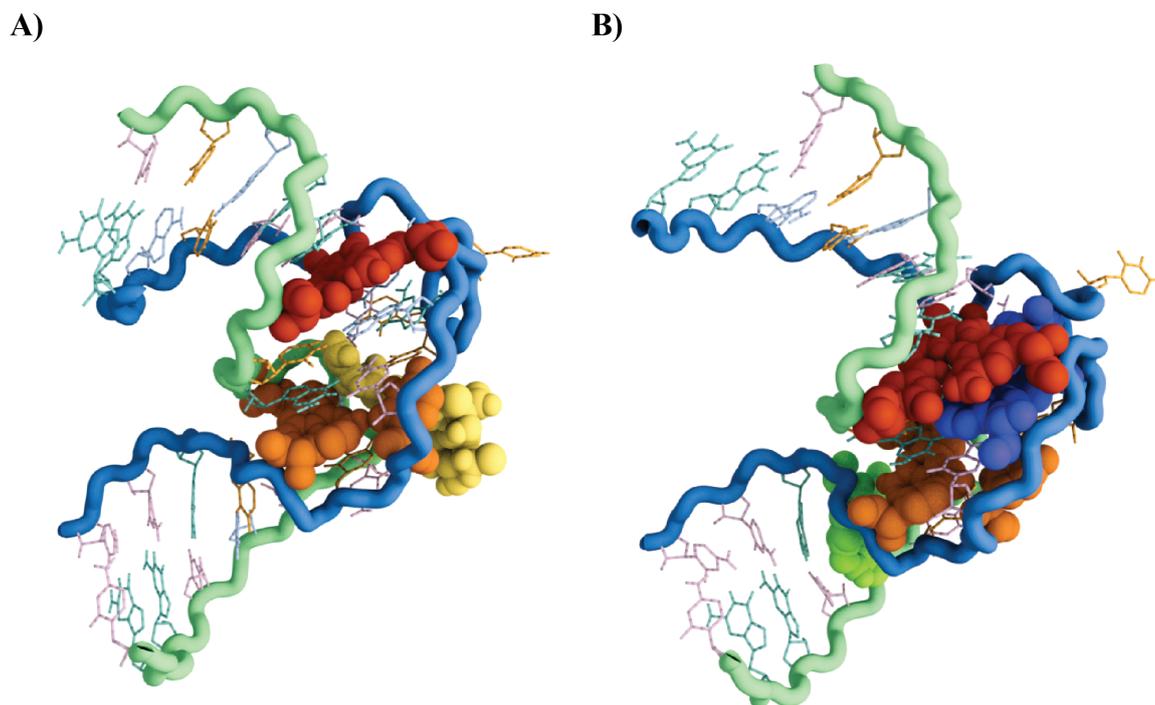


Figure S2: Representative initial structures for A) *complex* and B) *aptamer* systems. In the plate A, the Triplets G6•G13•C113 (dark orange) and U16•C18•G110 (orange) and the streptomycin molecule (yellow) are shown in van der Waals spheres representation. In the plate B, the triplets G6•G13•C113 (dark orange), U16•C18•G110 (orange), U19•A105•C106 (green) and A8•A14•U12 (blue) are shown in van der Waals spheres representation.

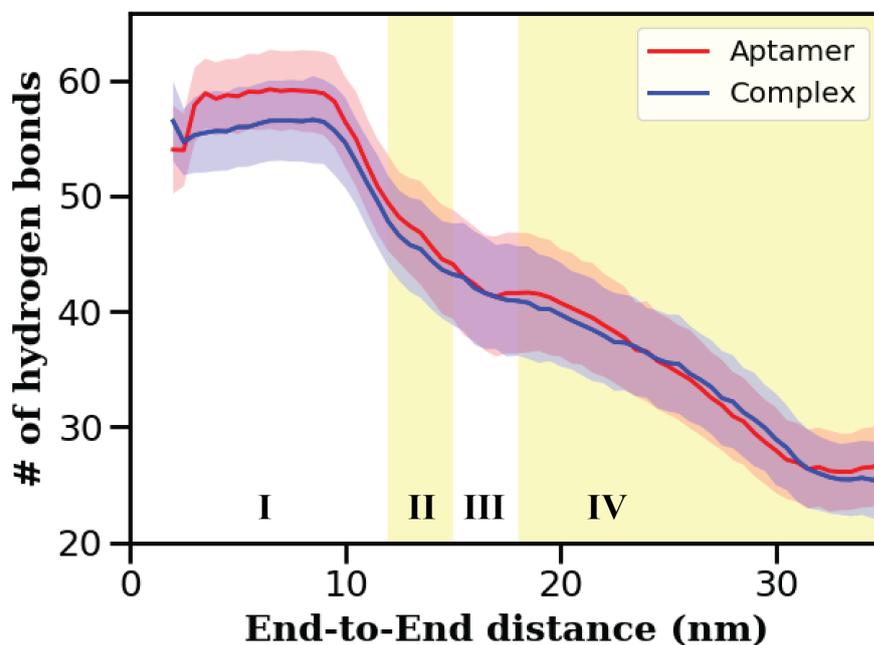


Figure S3: Number of total hydrogen bonds (inter and intra-strand) of the RNA molecule in the *aptamer* and *complex* system.

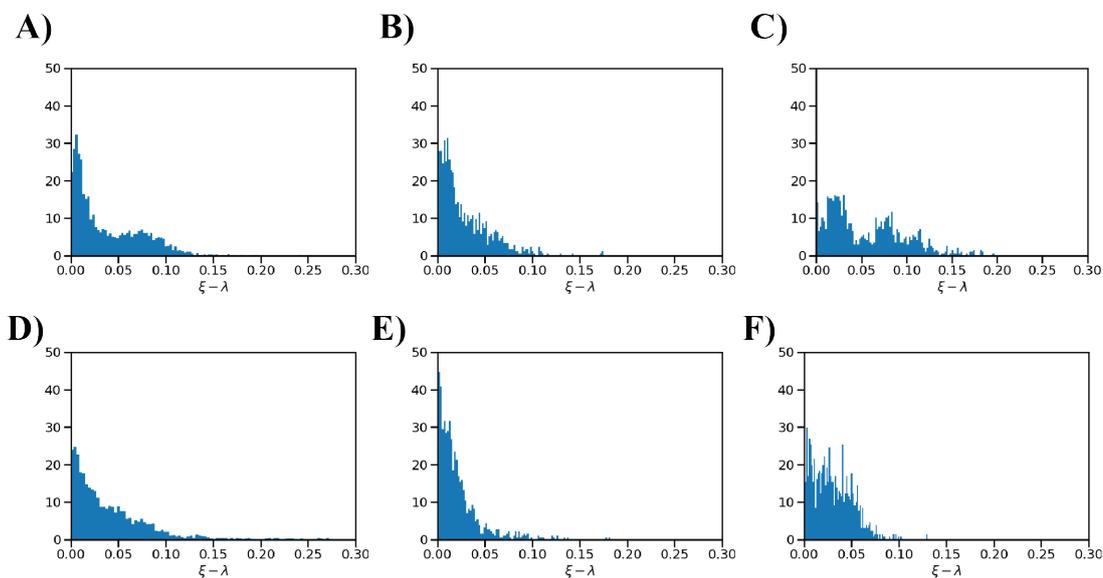


Figure S4: Histogram of the difference between  $\xi$  and  $\lambda$  for pulling velocities equal to 0.02, 0.04, and 0.07 nm ps<sup>-1</sup> (left to right) and *aptamer* (upper) and *complex* (bottom) systems. A) *aptamer* system and  $v = 0.02$  nm ps<sup>-1</sup>. B) *aptamer* system and  $v = 0.04$  nm ps<sup>-1</sup>. C) *aptamer* system and  $v = 0.07$  nm ps<sup>-1</sup>. D) *complex* system and  $v = 0.02$  nm ps<sup>-1</sup>. E) *complex* system and  $v = 0.04$  nm ps<sup>-1</sup>. F) *complex* system and  $v = 0.07$  nm ps<sup>-1</sup>.