

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA**

**DINÂMICA DE BIOMOLÉCULAS EM ÁGUA:
SIMULAÇÕES COM REPRESENTAÇÃO
EXPLÍCITA DO SOLVENTE**

Marta Maria Cassiano

**Tese de Doutorado
submetida ao Instituto de Física
da Universidade de São Paulo**

**Orientador: Prof. Paulo Mascarello Bisch
Co-orientador: Prof. Amando Siuti Ito**

Sylvio Canuto
Prof. Sylvio Roberto Accioly Canuto
Presidente da Comissão de Pós-Graduação

**São Paulo
1998**



572.8
C345d
D
ex 2

FICHA CATALOGRÁFICA

Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Cassiano, Marta Maria

Dinâmica de Biomoléculas em Água: Simulações com
Representação Explícita do Solvente. São Paulo, 1998.

Tese (Doutoramento) Universidade de São Paulo.
Instituto de Física - Departamento de Física Geral

Área de Concentração: Física do Estado Sólido
Orientador: Prof. Dr. Paulo Mascarello Bisch
Co-orientador: Prof. Dr. Amando Siuiti Ito

Unitermos: 1. Dinâmica Molecular; 2. Solvente Explícito;
3. Interações Biomoléculas/Solvente;
4. Função de Distribuição Radial de Pares;
5. Conformações de Biomoléculas.

USP/IF/SBI-014/98

SBI-IFUSP



305M810T2794

**DINÂMICA DE BIOMOLÉCULAS EM ÁGUA:
SIMULAÇÕES COM REPRESENTAÇÃO
EXPLÍCITA DO SOLVENTE**

Marta Maria Cassiano

Orientador: Prof. Paulo Mascarello Bisch

Co-orientador: Prof. Amando Siuiti Ito

Amando Siuiti Ito

Banca examinadora:

Prof^a. Isabelle Ortmans

Prof. José Alfredo Gomes Arêas

Prof. José Roberto Ruggiero

Prof. Luiz Carlos Gomide de Freitas

Prof. Paulo Mascarello Bisch



Agradecimentos

Aos professores Paulo M. Bisch e Amando S. Ito, pela orientação e apoio.

Ao Eduardo L. A. Macchione, pelo completo envolvimento e inestimável contribuição.

Ao Pedro G. Pascutti, Michel Loos, André T. Ota, Kaline Coutinho, Cláudio S. Shida e Carla Goldman, pelas discussões e sugestões.

A todos os amigos e colegas que, direta ou indiretamente, contribuíram para a realização deste trabalho.

Ao Laboratório de Computação Científica Avançada, onde parte das simulações foram realizadas.

Ao CNPq e à FAPESP, pelo auxílio financeiro.

Resumo

O programa THOR de dinâmica e mecânica molecular foi aprimorado com a introdução de técnicas que permitem a modelagem de sistemas biomoleculares na presença do solvente representado explicitamente. Essas técnicas incluem o emprego de condições de contorno periódicas, raio de corte, potencial eletrostático modificado e método *shake* para vínculos de distância. Além disso, a técnica da lista de vizinhos foi implementada devido ao grande aumento no número de átomos do sistema, pela inclusão explícita do solvente. Alguns sistemas foram modelados e a influência do solvente sobre suas propriedades conformacionais foi analisada, respondendo às expectativas baseadas em previsões teóricas e evidências experimentais.

Abstract

The computational package THOR for molecular mechanic and dynamic simulation was improved by the introduction of techniques for the modelling of biomolecular systems with explicit representation of solvent molecules. Those techniques include the application of periodic boundary conditions, the use of a cut-off sphere for the potential, with modifications in the electrostatic potential, and the utilization of the method *shake* to impose distance constraints. Furthermore, a neighbours list was implemented, due to the high number of atoms to be dealt with when the solvent is explicitly represented. Simulations were performed on some systems of biological interest, and the influence of the solvent on their conformational properties was analysed. The results are in good agreement with theoretical previsions and experimental data.

Índice

Introdução : O programa THOR	1
1 Campo de Força	4
1.1 Potencial harmônico linear	4
1.2 Potencial harmônico angular	5
1.3 Potencial diédrico impróprio	6
1.4 Potencial torcional próprio	7
1.5 Potencial de Lennard-Jones	8
1.6 Potencial eletrostático	8
1.7 Potencial total	9
2 Minimização da energia	11
2.1 Método <i>steepest descent</i>	12
2.2 Método dos gradientes conjugados	13
3 Dinâmica Molecular	15
3.1 Algoritmo <i>leap-frog</i>	16
3.2 Temperatura do sistema	17
4 Modelagem do solvente	21
4.1 Condições de contorno periódicas	21
4.2 Raio de corte	23
4.3 Potencial eletrostático modificado	24
4.4 Lista de vizinhos	27
4.5 Método <i>shake</i> para vínculos de distância	28

5	Modelagem de biomoléculas em água	32
5.1	Modelos para a água	32
5.2	A caixa d'água	34
5.3	Biomoléculas em água	35
6	Resultados	37
6.1	Água	37
6.1.1	Função de distribuição radial de pares	39
6.1.2	Resultados	40
6.2	Polialanina	45
6.2.1	Simulação	46
6.2.2	Resultados	50
6.3	Triptofano	62
6.3.1	Simulação	64
6.3.2	Resultados	66
7	Conclusão	78
	Apêndice A: Arquivos do programa THOR	80
A.1	Arquivos IN e NEW	81
A.2	Arquivo TOP	84
A.3	Arquivo DINAMICA.DAT	89
A.4	Arquivo OUT	90
A.5	Arquivo DHA	90
A.6	Arquivo DBK	91
A.7	Arquivos PDB, DMM e XYZ	92
	Apêndice B: Computadores utilizados	94
	Referências	96

Introdução: O Programa THOR

O programa THOR é um programa computacional que vem sendo desenvolvido desde 1992 com o objetivo de modelar sistemas moleculares de interesse biológico.

Utilizando um campo de força clássico para descrever as interações intra- e intermoleculares, o programa THOR permite o estudo de conformações de mínima energia, possibilitando a previsão e análise de estruturas mais prováveis, e o acompanhamento da evolução temporal dessas estruturas.

Idealizado por K. C. Mundim e P. M. Bisch, que traçaram os primeiros algoritmos baseados no campo de força clássico do programa GROMOS (*GR*oningen *MO*lecular *SI*mulation) para proteínas [1], o programa THOR foi posto em funcionamento, sendo capaz de realizar as primeiras simulações de alguns aminoácidos e pequenas moléculas [2].

Em 1993 surgiu o interesse no estudo da dinâmica e mudanças conformacionais de sistemas moleculares na presença de membranas biológicas. No entanto, as limitações dos recursos computacionais disponíveis, que permitiam a modelagem de sistemas com pouco mais de uma centena de átomos e o acompanhamento da dinâmica por alguns picossegundos, levaram ao desenvolvimento de um modelo simples, onde o meio aquoso e o interior da membrana são representados como contínuos, caracterizados por uma constante dielétrica de valor $\epsilon = 80$ e $\epsilon = 2$, respectivamente [3]. A definição de uma interface de descontinuidade dielétrica para representar a separação entre os dois meios possibilita a utilização do método das imagens eletrostáticas para o cálculo das interações interatômicas [3,4].

Embora este modelo venha sendo utilizado com relativo sucesso em vários estudos [5,6], o tratamento do solvente como um contínuo apresenta algumas limitações. Como a dependência da estruturação local do solvente em função do comportamento hidrofóbico ou hidrofílico de diferentes regiões em biomoléculas não é representada explicitamente, este modelo pode levar a distorções estruturais dos sistemas simulados [3,7-11]. Simulações de dinâmica de uma polialanina composta por vinte resíduos em conformação hélice- α , realizadas com o programa THOR, usando o modelo de meio contínuo, resultaram em uma conformação final estendida [4], contrariando previsões teóricas e evidências experimentais [12-15].

A realização deste trabalho visou ao aprimoramento das técnicas de simulação do programa THOR, ampliando sua capacidade de modelar biomoléculas em solução, com a possibilidade de representação explícita das moléculas do solvente. Para isto, vários algoritmos foram desenvolvidos, adaptados e introduzidos no programa.

Além das modificações diretamente relacionadas com a modelagem do solvente, modificações suplementares foram necessárias, devido ao grande aumento no número total de átomos do sistema, pela inclusão explícita do solvente, o qual poderia inviabilizar simulações pelo tempo de processamento requerido.

Paralelamente, alguns programas básicos foram desenvolvidos para a construção automática dos arquivos de entrada do programa THOR, com informações conformacionais dos sistemas constituídos pelas biomoléculas de interesse e moléculas do solvente.

Esta tese inicia-se com uma apresentação geral do programa THOR para modelagem molecular: o campo de força utilizado (capítulo 1), as técnicas de minimização de energia (capítulo 2) e dinâmica molecular

(capítulo 3). Em seguida, são apresentadas as modificações efetuadas no programa para a realização de simulações com o solvente representado explicitamente (capítulo 4) e a metodologia desenvolvida para a geração das conformações iniciais, incluindo as moléculas do solvente (capítulo 5). Alguns sistemas foram estudados com esta nova versão do THOR. A modelagem e os resultados são apresentados e discutidos na seqüência (capítulo 6). Uma avaliação do trabalho desenvolvido é então exposta (capítulo 7). No apêndice A é feita uma apresentação operacional do programa, com descrição dos arquivos de entrada e saída. O apêndice B contém informações sobre os computadores utilizados e tempo de processamento requerido para as simulações realizadas.

1 Campo de Força

O campo de força molecular ou potencial efetivo utilizado no programa THOR para representar as interações entre os átomos de um sistema é baseado no campo de força do programa GROMOS [1]. Ele consiste em uma superposição de vários termos, representando as interações entre átomos covalentemente ligados (potencial harmônico linear, potencial harmônico angular, potencial diédrico impróprio e potencial torcional próprio) e as interações entre átomos não-ligados (potencial de Lennard-Jones e de Coulomb). Em interações intramoleculares, são considerados não-ligados os átomos separados por três ou mais ligações covalentes (ver apêndice A). Os parâmetros empregados na definição de cada termo do potencial são determinados a partir de resultados de cálculos quânticos *ab-initio* de pequenas moléculas e por ajustes de dados experimentais (cristalografia de raios-X, espectroscopia de infravermelho, ressonância magnética nuclear, etc.) [1,16].

1.1 Potencial harmônico linear

A interação entre dois átomos *i* e *j* ligados covalentemente é representada por

$$V_b = \frac{1}{2} K_b (b - b_0)^2 \quad (1-1)$$

onde b é o comprimento da ligação $i - j$, b_0 é o comprimento de equilíbrio e K_b é a constante elástica da força. O potencial V_b mede a distorção da ligação em relação ao equilíbrio ($V_b = 0$ quando $b = b_0$).

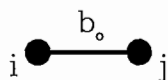


Fig. 1-1 : b_0 é o comprimento de equilíbrio da ligação covalente $i - j$.

1.2 Potencial harmônico angular

O termo do potencial que mantém o ângulo de ligação entre três átomos $i - j - k$, covalente e consecutivamente ligados, próximo à geometria de equilíbrio é representado por

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

onde θ é o ângulo definido pela ligação $i - j - k$, θ_0 é o ângulo de equilíbrio ($V_\theta = 0$ quando $\theta = \theta_0$) e K_θ é a constante elástica da força.

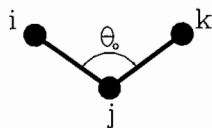


Fig. 1-2 : θ_0 é o ângulo de equilíbrio da ligação $i - j - k$.

1.3 Potencial diédrico impróprio

É o termo do potencial que descreve a interação entre quatro átomos. Se o átomo i está ligado simultaneamente aos átomos j , k e l e ξ é o ângulo entre os planos definidos por $i - j - k$ e por $j - k - l$ [17] (fig. 1-3), então o potencial de interação é descrito por

$$V_{\xi} = \frac{1}{2} K_{\xi} (\xi - \xi_0)^2 \quad (1-3)$$

onde ξ_0 é o ângulo de equilíbrio ($V_{\xi} = 0$ quando $\xi = \xi_0$) e K_{ξ} é a constante elástica da força. O potencial V_{ξ} é definido de modo a manter a configuração tridimensional do conjunto $i - j - k - l$ planar ($\xi_0 = 0^\circ$) ou tetraédrica ($\xi_0 = 35,25^\circ$).

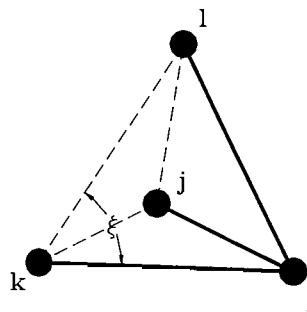


Fig. 1-3 : ξ é o ângulo entre os planos $i - j - k$ e $j - k - l$.

1.4 Potencial torcional próprio

Normalmente é introduzido na expressão do potencial um termo representando explicitamente a torção própria de ligações covalentes entre dois átomos. A técnica mais usada consiste em fazer uma expansão em cosseno e usar somente o termo de mais baixa ordem para cada torção [7].

Para um conjunto de átomos i , j , k e l unidos seqüencialmente por ligações covalentes, se φ é o ângulo entre os planos definidos pelas ligações $i - j - k$ e $j - k - l$ [17] (fig. 1-4), o potencial torcional próprio da ligação $j - k$ é dado por

$$V_{\varphi} = K_{\varphi} [1 + \cos(p\varphi + \delta)] \quad (1-4)$$

onde K_{φ} é a constante de força, p é a periodicidade do potencial e δ é a fase.

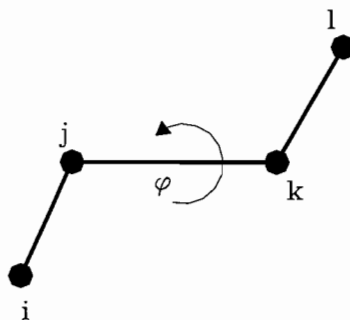


Fig. 1-4 : φ é o ângulo entre os planos $i - j - k$ e $j - k - l$.

1.5 Potencial de Lennard-Jones

O potencial de Lennard-Jones, representando a interação entre dois átomos i e j não-ligados, é dado por

$$V_{LJ} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \quad (1-5)$$

onde r_{ij} é a distância entre os átomos i e j e A_{ij} e B_{ij} são parâmetros que dependem dos tipos de átomos envolvidos na interação.

O primeiro termo do potencial descreve a repulsão das nuvens eletrônicas que sobrepõem-se quando átomos são colocados próximos entre si, enquanto o termo atrativo representa as interações de dipolos instantâneos provocadas por flutuações na densidade de carga eletrônica dos átomos [3].

1.6 Potencial eletrostático

A interação eletrostática entre dois átomos i e j é representada pelo potencial de Coulomb

$$V_c = \frac{q_i q_j}{4\pi\epsilon_0\epsilon r_{ij}} \quad (1-6)$$

onde r_{ij} é a distância entre os átomos i e j , q_i e q_j são suas cargas parciais, respectivamente, ϵ é a constante dielétrica do meio e ϵ_0 é a permissividade do vácuo, igual a $8,854 \cdot 10^{-12}$ (em unidades do S.I.) [18].

Várias propostas têm sido feitas para a representação do comportamento dielétrico de sistemas modelados quando o solvente não é considerado explicitamente. Normalmente um valor fixo é designado para a constante dielétrica ϵ , mas, em alguns casos, uma dependência linear com a distância ($\epsilon = r_{ij}$) é usada para levar em conta o efeito do solvente conforme a separação entre dois átomos aumenta [16,19-21]. Quando todos os átomos e cargas do sistema são modelados explicitamente, deve-se usar $\epsilon = 1$ [1,16].

1.7 Potencial total

O potencial total é dado pela soma de todos os termos representando as interações entre todos os átomos ligados e não-ligados do sistema. Se o sistema é constituído por N átomos e possui N_b pares ligados covalentemente, N_θ conjuntos de três átomos ligados covalente e consecutivamente, N_ξ conjuntos de quatro átomos definindo ligações diédricas impróprias e N_ϕ conjuntos de quatro átomos definindo torções próprias, o potencial total do sistema é representado por

$$\begin{aligned}
V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = & \sum_{n=1}^{N_b} \frac{1}{2} K_{b_n} (b_n - b_{0_n})^2 + \sum_{n=1}^{N_\theta} \frac{1}{2} K_{\theta_n} (\theta_n - \theta_{0_n})^2 + \\
& \sum_{n=1}^{N_\xi} \frac{1}{2} K_{\xi_n} (\xi_n - \xi_{0_n})^2 + \sum_{n=1}^{N_\varphi} \frac{1}{2} K_{\varphi_n} [1 + \cos(p_n \varphi_n + \delta_n)] + \quad (1-7) \\
& \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left[\left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \frac{q_i q_j}{4\pi\epsilon_0 \epsilon r_{ij}} \right]
\end{aligned}$$

Essa forma de potencial tem sido amplamente usada em vários estudos de modelagem molecular [7,10,16,19-22] e é, particularmente, a expressão adotada para o campo de força do programa THOR. Para mais detalhes sobre os parâmetros e utilização do programa, ver apêndice A.

2 Minimização da energia

A minimização de energia consiste em uma técnica que procura, a partir de uma conformação arbitrária de um sistema molecular, a obtenção de uma conformação de mais baixa energia. Esta técnica é bastante utilizada para a obtenção do mapeamento de energia de pequenos peptídios ou para a preparação de conformações para o início de uma simulação de dinâmica molecular.

Existem vários algoritmos para a efetivação desta técnica. No programa THOR foram adaptados dois métodos bastante conhecidos e denominados *steepest descent* e dos gradientes conjugados. Em ambos, a conformação minimizada é obtida por meio de um processo iterativo de ajustes das coordenadas atômicas [7,9,23-25].

A diferença entre os dois métodos reside no modo como as coordenadas atômicas são modificadas. Esta diferença faz com que o método *steepest descent* apresente uma convergência bastante rápida em regiões distantes do mínimo, mas lenta em regiões próximas ao mínimo, sendo o método dos gradientes conjugados mais eficiente nessas regiões [7,9,25]. Deste modo, um procedimento de minimização é normalmente efetuado em duas etapas no programa THOR: inicialmente é feita uma minimização com o emprego do método *steepest descent*, para um relaxamento das estruturas iniciais, e então é utilizado o método dos gradientes conjugados, para a obtenção da conformação final minimizada.

Métodos como o *steepest descent* e dos gradientes conjugados permitem somente a obtenção de conformações de mínimo local, as quais normalmente não estão muito distantes da conformação inicial [7,9].

2.1 Método *steepest descent*

O método segue basicamente o seguinte esquema:

No n -ésimo passo iterativo do processo de minimização, a conformação do sistema, constituído por N átomos, é definida pelo conjunto das coordenadas atômicas $\{\mathbf{r}_1^n, \mathbf{r}_2^n, \dots, \mathbf{r}_N^n\}$ e o potencial total e seu gradiente são calculados a partir do campo de força adotado (eq. 1-7). A conformação do sistema é então modificada, a partir de pequenos deslocamentos das coordenadas atômicas:

$$\mathbf{r}_i^{n+1} = \mathbf{r}_i^n + k^n \mathbf{g}_i^n \quad (2-1)$$

com

$$\mathbf{g}_i^n = - \frac{\partial V^n}{\partial \mathbf{r}_i} \left| \frac{\partial V^n}{\partial \mathbf{r}_i} \right|^{-1} = \frac{\mathbf{F}_i}{|\mathbf{F}_i|} \quad (2-2)$$

O coeficiente k^n define a amplitude dos deslocamentos, enquanto os versores \mathbf{g}_i^n determinam suas direções. A equação (2-2) indica que as coordenadas do átomo i devem ser deslocadas no sentido da força atuando sobre ele.

O valor da amplitude de deslocamento inicial k^0 é um parâmetro de entrada do programa, com valor sugerido de $0,1 \text{ \AA}$. Durante a minimização, o valor de k^n varia, dependendo da variação do potencial em passos consecutivos. Se V^{n+1} é o potencial da nova conformação e V^n é o potencial da conformação anterior, então

se $V^{n+1} < V^n \Rightarrow$ as novas posições são aceitas e $k^{n+1} = 1,2 k^n$

se $V^{n+1} \geq V^n \Rightarrow$ as novas posições são rejeitadas e $k^{n+1} = \frac{k^n}{2}$

Em cada passo iterativo, a diferença entre os valores V^{n+1} e V^n é calculada. Se esta diferença for menor do que um certo valor de convergência ΔV , então o processo de minimização é interrompido. ΔV é um parâmetro de entrada, com um valor sugerido de 10^{-4} kcal mol⁻¹.

2.2 Método dos gradientes conjugados

Este método segue basicamente o mesmo esquema do *steepest descent*. A diferença entre ambos é que, neste caso, as direções dos deslocamentos das coordenadas atômicas no (n+1)-ésimo passo iterativo, ao invés de serem dadas por \mathbf{g}^n (eq. 2-1 e 2-2), são dadas por

$$\mathbf{h}_i^n = \mathbf{g}_i^n + \gamma_i^n \mathbf{h}_i^{n-1} \quad (2-3)$$

onde

$$\gamma_i^n = \frac{(\mathbf{g}_i^n - \mathbf{g}_i^{n-1}) \cdot \mathbf{g}_i^n}{\mathbf{g}_i^{n-1} \cdot \mathbf{g}_i^{n-1}} \quad (2-4)$$

ou seja, as direções \mathbf{h}_i^n dependem das direções do passo anterior \mathbf{h}_i^{n-1} e dos versores \mathbf{g}_i^{n-1} e \mathbf{g}_i^n . As direções iniciais \mathbf{h}_i^0 são definidas como $\mathbf{h}_i^0 \equiv \mathbf{g}_i^0$.

Esta definição das direções de deslocamento faz com que cada nova direção seja conjugada com todas as direções anteriores¹. A utilização de um conjunto de direções mutuamente conjugadas torna este método mais eficiente do que o método *steepest descent* [7,9,25]. Uma explicação mais aprofundada dos dois métodos pode ser encontrada na referência [25].

¹ Dois vetores \mathbf{u} e \mathbf{v} são ditos serem conjugados se $\mathbf{u} \cdot \mathbf{H} \cdot \mathbf{v} = 0$, onde $(H)_{ij} = \frac{\partial^2 V}{\partial x_i \partial x_j}$.

Com a utilização do método dos gradientes conjugados, as direções conjugadas são obtidas sem a necessidade de cálculo das derivadas de segunda ordem do potencial [25].

3 Dinâmica Molecular

A dinâmica molecular consiste em uma técnica que permite o acompanhamento temporal do movimento de todos os átomos de um sistema por meio da integração simultânea das equações de movimento clássicas.

Para um sistema constituído por N átomos com posições $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$, se o potencial total de interação $V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ é uma função diferenciável das coordenadas atômicas, então

$$\frac{d^2 \mathbf{r}_i(t)}{dt^2} = \frac{\mathbf{F}_i}{m_i} \quad (3-1)$$

$$\mathbf{F}_i = - \frac{\partial V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)}{\partial \mathbf{r}_i} \quad (3-2)$$

onde t representa o tempo, \mathbf{F}_i a força agindo sobre o átomo i e m_i a sua massa.

Com o cálculo do gradiente do potencial, \mathbf{F}_i é obtida (eq. 3-2) e a integração numérica da equação de Newton (3-1) pode ser realizada em pequenos intervalos de tempo Δt , normalmente da ordem de fentossegundos [16,22].

3.1 Algoritmo *leap-frog*

A equação (3-1) é uma equação diferencial de segunda ordem, que pode ser escrita como um conjunto de duas equações diferenciais de primeira ordem para as posições $\mathbf{r}_i(t)$ e velocidades $\mathbf{v}_i(t)$ atômicas

$$\frac{d\mathbf{v}_i(t)}{dt} = \frac{\mathbf{F}_i(\mathbf{r}_i(t))}{m_i} \quad (3-3)$$

$$\frac{d\mathbf{r}_i(t)}{dt} = \mathbf{v}_i(t) \quad (3-4)$$

Um algoritmo simples para a integração numérica dessas equações pode ser obtido por meio da expansão de Taylor de $\mathbf{v}_i(t)$ em torno de $\left(t + \frac{\Delta t}{2}\right)$ e $\left(t - \frac{\Delta t}{2}\right)$, considerando termos até segunda ordem

$$\mathbf{v}_i\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}_i(t) + \frac{d\mathbf{v}_i(t)}{dt} \left(\frac{\Delta t}{2}\right) + \frac{d^2\mathbf{v}_i(t)}{dt^2} \left(\frac{\Delta t}{2}\right)^2 \frac{1}{2!} \quad (3-5a)$$

$$\mathbf{v}_i\left(t - \frac{\Delta t}{2}\right) = \mathbf{v}_i(t) - \frac{d\mathbf{v}_i(t)}{dt} \left(\frac{\Delta t}{2}\right) + \frac{d^2\mathbf{v}_i(t)}{dt^2} \left(\frac{\Delta t}{2}\right)^2 \frac{1}{2!} \quad (3-5b)$$

ou ainda,

$$\mathbf{v}_i\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}_i\left(t - \frac{\Delta t}{2}\right) + \frac{\mathbf{F}_i(\mathbf{r}_i(t))}{m_i} \Delta t \quad (3-6)$$

Para $\mathbf{r}_i(t)$, a expansão de Taylor em torno de $(t + \Delta t)$, até segunda ordem, produz

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \frac{d\mathbf{r}_i(t)}{dt} \Delta t + \frac{d^2\mathbf{r}_i(t)}{dt^2} \frac{(\Delta t)^2}{2!} \quad (3-7)$$

ou ainda, substituindo as equações (3-4) e (3-5a) na equação (3-7) e desprezando termos de ordem superior,

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i \left(t + \frac{\Delta t}{2} \right) \Delta t \quad (3-8)$$

As equações (3-6) e (3-8) podem ser resolvidas numericamente de forma iterativa, formando o conhecido esquema *leap-frog*, um dos algoritmos mais simples, eficientes e estáveis utilizados em dinâmica molecular [16,26].

3.2 Temperatura do sistema

A temperatura T do sistema, em um certo instante de integração, é definida em termos da energia cinética

$$\frac{1}{2} N_{gl} k_B T = \left\langle \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 \right\rangle_{n_p} \quad (3-9)$$

onde N_{gl} é o número de graus de liberdade do sistema, k_B é a constante de Boltzmann ($k_B = 1,38 \cdot 10^{-23} \text{ J K}^{-1}$ [18]) e $\left\langle \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 \right\rangle_{n_p}$ representa a média da energia cinética sobre um número n_p de passos de integração [7,16,19].

Uma simulação de dinâmica molecular normalmente é iniciada com o sistema em uma conformação otimizada pela técnica de minimização de energia. No entanto, para a utilização do algoritmo *leap-frog* são necessárias ainda informações sobre as velocidades atômicas. Um procedimento usualmente empregado no programa THOR consiste em atribuir valores nulos às velocidades iniciais, ou então obtê-las a partir de uma distribuição de Maxwell-Boltzmann para uma temperatura inicial T_0 , com valor próximo a 0 K. Neste caso, aos módulos das velocidades atômicas são atribuídos valores

$$v_i = \sqrt{\frac{2 k_B T_0}{m_i}} \quad (3-10)$$

que correspondem ao valor mais provável da distribuição [27], enquanto as direções são definidas por sorteios aleatórios.

Inicia-se então o processo de termalização, onde o sistema é aquecido de uma temperatura inicial T_0 até uma temperatura final T_F . Neste processo, em cada passo de integração Δt o sistema sofre um pequeno incremento na temperatura. Se ΔT_{aq} é a taxa de aquecimento por picossegundo, espera-se que no n -ésimo passo de integração a temperatura do sistema seja dada por

$$T_n^d = T_0 + (n \Delta t) \Delta T_{aq} \quad (3-11)$$

O valor de T_n^d é comparado com o valor da temperatura T_n calculado naquele passo (eq. 3-9). Caso T_n^d não tenha ainda sido alcançada, as velocidades atômicas são reescaladas :

$$\mathbf{v}_i \rightarrow \alpha \mathbf{v}_i \quad (3-12)$$

onde

$$\alpha = \left(\frac{T_n^d}{T_n} \right)^{\frac{1}{2}} \quad (3-13)$$

O processo é repetido a cada passo de integração, até a obtenção da temperatura final desejada. A taxa de aquecimento ΔT_{aq} é estipulada como parâmetro de entrada do programa e pode assumir qualquer valor, determinando o tempo total que o sistema levará para ir de T_0 a T_F . Em princípio, o aquecimento pode ser realizado em um único passo, no entanto, um aquecimento mais lento é aconselhável, fazendo com que o sistema necessite de menos tempo para atingir o equilíbrio após o alcance da temperatura T_F [19].

A fase seguinte da termalização, a estabilização, é a fase em que ocorre o relaxamento das perturbações introduzidas no sistema durante o aquecimento. O tempo de duração desta fase varia de sistema para sistema e é normalmente determinado pelo acompanhamento das flutuações na temperatura e energias cinética, potencial e total [7]. Durante esta fase, espera-se que a temperatura T_F alcançada no aquecimento seja mantida constante e, para isto, as velocidades

atômicas continuam a ter seus valores ajustados (eq. 3-9) com certa frequência. Isto deve acontecer sempre que a diferença entre a temperatura do sistema e a temperatura desejada for maior do que um desvio máximo permitido para a temperatura $\Delta T_{\text{máx}}$. Neste caso, o fator de reescalonamento das velocidades é dado por [28]

$$\alpha = \left(\frac{T_F}{T_n} \right)^{\frac{1}{2}} \quad (3-14)$$

Mesmo quando atingido o equilíbrio, devido a aproximações nos cálculos numéricos e truncamentos no cálculo das interações entre átomos não-ligados, é necessário fazer um controle da temperatura para que esta permaneça constante. Para esta fase, um reescalonamento mais suave é proposto, com o acoplamento do sistema a um banho térmico [29], sendo o desvio da temperatura em relação à T_F corrigido em um tempo característico τ . Neste caso, o fator de reescalonamento das velocidades em cada passo de integração é dado por

$$\alpha = \left[1 + \frac{\Delta t}{\tau} \left(\frac{T_F}{T_n} - 1 \right) \right]^{\frac{1}{2}} \quad (3-15)$$

A intensidade do acoplamento do sistema ao banho térmico é definido pelo parâmetro τ . Se $\tau = \Delta t$, então a equação (3-15) é reduzida à expressão (3-14). Valores de τ da ordem de 0,1 ps resultarão em um acoplamento fraco, introduzindo, portanto, pouca perturbação ao sistema [29].

4 Modelagem do solvente

Neste capítulo são descritas as técnicas, a partir das quais algoritmos foram desenvolvidos e adaptados no programa THOR, permitindo-o realizar simulações de sistemas biomoleculares com representação explícita do solvente.

Primeiramente é introduzido o conceito de condições de contorno periódicas, utilizadas para minimizar os efeitos de borda de sistemas finitos. O efeito de periodicidade, indesejado em simulações de sistemas líquidos e sistemas constituídos por moléculas em solução, é eliminado com a definição de um raio de corte para o truncamento das interações entre átomos não-ligados do sistema. Truncamentos abruptos dessas interações são evitados com a modificação do potencial eletrostático, tornando-o contínuo na região do corte. Devido ao grande aumento no número total de átomos do sistema pela simulação explícita do solvente, a técnica da lista de vizinhos é utilizada, reduzindo o tempo gasto com os cálculos das interações entre átomos não-ligados. O método *shake*, também adaptado no programa THOR, permite a manutenção de vínculos de distância entre alguns átomos do sistema durante as simulações, possibilitando um aumento na velocidade dos cálculos de dinâmica molecular, com o emprego de passos de integração maiores.

4.1 Condições de contorno periódicas

Em simulações computacionais de sistemas moleculares finitos, condições de contorno periódicas são comumente utilizadas para minimizar os efeitos de borda [16,30]. Os átomos do sistema que está

sendo simulado são colocados em uma caixa retangular e, ao redor desta, são colocadas caixas-imagens, isto é, cópias da caixa original transladadas nas três direções espaciais, formando uma rede. Assim, cada átomo estará sempre rodeado por outros átomos ou por suas imagens.

No decorrer da simulação, quando um átomo move-se em sua caixa original, sua imagem periódica em cada uma das caixas-imagens move-se da mesma maneira. Além disso, quando um átomo sai da caixa central, atravessando uma de suas seis faces, uma de suas imagens entra pela face oposta, com velocidade idêntica. Deste modo, o número total de átomos na caixa central, e portanto do sistema como um todo, é conservado.

Uma versão bidimensional de um sistema periódico é mostrada na figura 4-1.

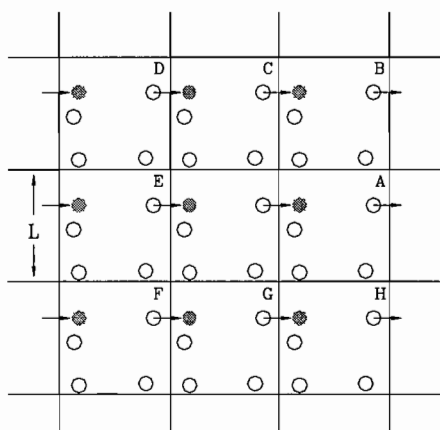


Fig. 4-1 : Sistema periódico bidimensional quadrado. A caixa central representa o sistema original e as demais são suas imagens. O movimento de um átomo na caixa central é repetido em cada uma das caixas-imagens.

4.2 Raio de corte

Aplicações de condições de contorno periódicas significam que, de fato, uma rede periódica infinita está sendo simulada e para o cálculo das interações devem ser considerados todos os átomos de todas as caixas. Esta técnica é ideal para simulações de sistemas cristalinos, onde as posições atômicas devem ser mantidas próximas às posições médias, regularmente espaçadas. Para sistemas líquidos ou sistemas constituídos por moléculas em solução, a periodicidade representa apenas um artifício computacional, representando sistemas contínuos e evitando efeitos de borda. Os efeitos da periodicidade não devem, portanto, afetar os cálculos das forças interatômicas nesses sistemas. Isto significa que um átomo não deve interagir com sua própria imagem ou também com um outro átomo e, simultaneamente, com imagens periódicas deste átomo. Deste modo, um raio de corte R_c , menor ou igual à metade do menor comprimento lateral da caixa retangular, é normalmente definido e somente são consideradas as interações entre átomos separados entre si por uma distância menor ou igual à R_c [30].

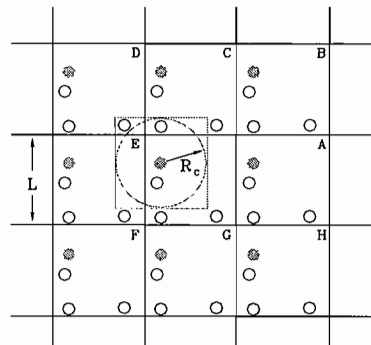


Fig. 4-2 : Sistema periódico bidimensional quadrado. São consideradas somente as interações entre o átomo “cinza” e aqueles situados dentro do círculo de raio R_c .

Com a definição de um raio de corte, interações entre átomos mais distantes do que R_c são desprezadas. As interações eletrostáticas, no entanto, proporcionais a r_{ij}^{-1} (eq. 1-6), apresentam um comportamento de longo-alcance. Para que o truncamento dessas interações, não provoque grandes perturbações no sistema, a caixa que contém o sistema deve ser considerada a maior possível [10,11,30]. O tamanho da caixa e do raio de corte são, entretanto, limitados pela disponibilidade computacional.

4.3 Potencial eletrostático modificado

Quando o raio de corte é utilizado para o truncamento das interações entre átomos não-ligados, a descontinuidade do potencial em $r = R_c$ introduz perturbações no sistema durante simulações de dinâmica molecular [16]. Essas perturbações podem ser reduzidas se um truncamento mais suave for empregado, obtido por meio de uma modificação do potencial coulombiano, de modo que $V \rightarrow 0$ quando $r \rightarrow R_c$.

Existem várias formas propostas para esta modificação [16,19,31-35]. No THOR, quando condições de contorno periódicas e raio de corte são usados, a interação eletrostática entre dois átomos i e j , representada pelo potencial coulombiano V_c (eq. 1-6), é modificada pela soma de um termo V_c^s , resultando no potencial V_c^{mod} [35]:

$$V_c^{\text{mod}} = V_c + V_c^s \quad (4-1)$$

com

$$V_c^s = \frac{q_i q_j}{4\pi\epsilon_0\epsilon} \left[-\frac{c_1}{3}(r_{ij} - r_1)^3 - \frac{c_2}{4}(r_{ij} - r_1)^4 - c_3 \right] \quad (4-2a)$$

$$c_1 = -\frac{5R_c - 2r_1}{R_c^3(R_c - r_1)^2} \quad (4-2b)$$

$$c_2 = \frac{4R_c - 2r_1}{R_c^3(R_c - r_1)^3} \quad (4-2c)$$

$$c_3 = \frac{1}{R_c} - \frac{c_1}{3}(R_c - r_1)^3 - \frac{c_2}{4}(R_c - r_1)^4 \quad (4-2d)$$

Os parâmetros r_1 , c_1 , c_2 e c_3 estão relacionados com o deslocamento do potencial V_c^{mod} em relação ao potencial normal V_c . Enquanto $V_c(r_1) \propto \frac{1}{r_1}$ e $V_c(R_c) \propto \frac{1}{R_c}$, $V_c^{\text{mod}}(r_1) \propto \frac{1}{r_1} - c_3$ e $V_c^{\text{mod}}(R_c) = 0$.

Do mesmo modo, o módulo da força coulombiana atuando sobre o átomo i devido ao potencial V_c^{mod} é modificado por

$$F_{c_i}^{\text{mod}} = F_{c_i} + F_{c_i}^s \quad (4-3)$$

com

$$F_{c_i}^s = \frac{q_i q_j}{4\pi\epsilon_0\epsilon} \left[c_1 (r_{ij} - r_1)^2 + c_2 (r_{ij} - r_1)^3 \right] \quad (4-4)$$

A definição dos parâmetros c_1 e c_2 faz com que $F_{c_i}^{\text{mod}}(R_c) = 0$.

As figuras 4-3 e 4-4 mostram o comportamento do potencial e força eletrostáticos normais e modificados.

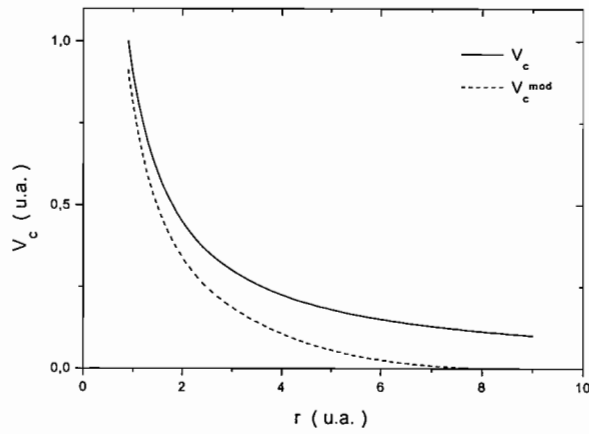


Fig. 4-3 : Comportamento do potencial eletrostático normal e modificado, em unidades arbitrárias, com $\frac{q_i q_j}{4\pi\epsilon_0\epsilon} = 1$; $R_c = 9$ e $r_1 = 4,5$.

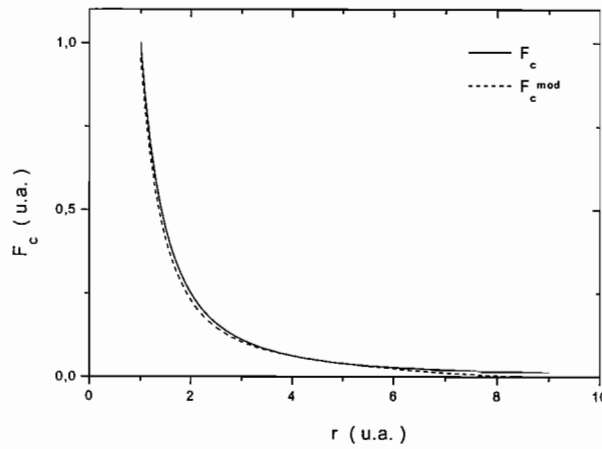


Fig. 4-4 : Comportamento da força eletrostática normal e modificada, em unidades arbitrárias, com $\frac{q_i q_j}{4\pi\epsilon_0\epsilon} = 1$; $R_c = 9$ e $r_1 = 4,5$.

Embora pouca diferença seja observada entre a força modificada e a força normal (fig. 4-4), o mesmo não acontece para o potencial (fig. 4-3). Algumas correções de longo-alcance são sugeridas em cálculos de propriedades termodinâmicas de sistemas simulados quando raio de corte e modificação do potencial eletrostático são empregados [30,35].

4.4 Lista de vizinhos

Em simulações de dinâmica molecular, a maior parte do tempo computacional é gasto com cálculos das interações entre os átomos não-ligados. Mesmo com a introdução de um raio de corte, que reduz o número de interações e causa um decréscimo do tempo utilizado, os cálculos de todas as distâncias interatômicas e posterior comparação com R_c exigem muito tempo.

A lista de vizinhos é uma técnica bastante utilizada para diminuir o tempo gasto com esses cálculos [30,36]. Ela consiste na definição de uma esfera de raio R_L , um pouco maior do que R_c , traçada ao redor de cada átomo do sistema. São considerados vizinhos do átomo central todos os átomos localizados no interior dessa esfera (fig. 4-5).

Uma lista é então construída, contendo todos os vizinhos de cada átomo, a qual é atualizada a certos intervalos de tempo. Entre as atualizações, o programa não calcula as distâncias entre todos os pares de átomos do sistema, mas somente as distâncias entre cada átomo e seus vizinhos.

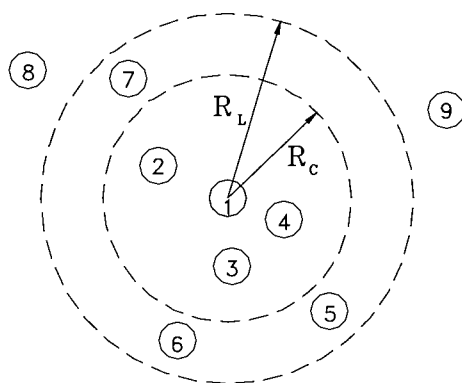


Fig. 4-5 : Definição dos parâmetros R_c e R_L usados na técnica da lista de vizinhos. São vizinhos do átomo 1 os átomos 2, 3, 4, 5, 6 e 7. No entanto, somente as interações com os átomos 2, 3 e 4 são computadas. Com o emprego desta técnica, cálculos das distâncias r_{1-8} e r_{1-9} são evitados.

O valor de R_L deve ser escolhido com certo critério. Valores grandes de R_L geram listas grandes, havendo pouca economia no tempo gasto com os cálculos das distâncias. Por outro lado, se R_L for muito pequeno, a lista deve ser atualizada com uma maior frequência. Um valor razoável para R_L é entre 10% e 15% maior do que o valor de R_c [30]. Neste caso, a atualização da lista de vizinhos deve ser efetuada com uma frequência de 10 a 20 passos de integração [30].

4.5 Método *shake* para vínculos de distância

Muitas vezes, quando sistemas moleculares são estudados por simulações computacionais, deseja-se que as distâncias entre alguns átomos sejam mantidas próximas a valores previamente conhecidos.

Nessas situações, os movimentos atômicos devem ser vinculados, de modo a preservar as distâncias de referência.

Por outro lado, a aplicação de vínculos de distância em cálculos de dinâmica molecular é também muitas vezes utilizada para reduzir o tempo computacional. De fato, o valor máximo do passo Δt empregado nas integrações numéricas das equações de movimento é definido de modo que os movimentos de mais alta frequência que ocorrem no sistema possam ser acompanhados durante a dinâmica. Como esses movimentos são devido às vibrações nas ligações entre pares de átomos, por imposição de determinados vínculos de distância, o valor de Δt pode ser aumentado de duas a quatro vezes [16].

Vários métodos têm sido desenvolvidos para que as equações de movimento possam ser integradas respeitando os vínculos impostos. No programa THOR foi adaptado o *shake*, um dos mais conhecidos métodos desenvolvidos para esta finalidade [37-40].

A resolução do problema pelo método *shake* exige que os N_v vínculos impostos sejam funções somente das coordenadas dos N átomos do sistema e possam ser escritos na forma

$$\sigma_k(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = 0 \quad k = 1, 2, \dots, N_v \quad (4-5)$$

Mais especificamente, os vínculos devem ser representados pelas distâncias interatômicas fixadas. Por exemplo, se o k -ésimo vínculo é imposto sobre a distância entre os átomos i e j , então ele deve ser expresso por

$$\mathbf{r}_{ij}^2 - d_{ij}^2 = 0 \quad (4-6)$$

onde \mathbf{r}_{ij} é o vetor distância entre os átomos i e j e d_{ij} é o módulo da distância de referência, que deve ser mantida no decorrer da simulação.

Formalmente, as equações de movimento sujeitas a um conjunto de vínculos podem ser resolvidas pela aplicação do método dos multiplicadores de Lagrange [41]. *Shake* é um método alternativo, que corrige iterativamente as posições atômicas após cada passo de integração, satisfazendo os vínculos impostos.

Se em um determinado instante $\mathbf{r}_i^*(t+\Delta t)$ for a posição do átomo i calculada sem a consideração do vínculo, então a posição corrigida $\mathbf{r}_i(t+\Delta t)$ será dada por

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i^*(t + \Delta t) + \delta \mathbf{r}_i \quad (4-7)$$

onde $\delta \mathbf{r}_i$ representa a correção devido a todos os vínculos envolvendo o átomo i .

Para o k -ésimo vínculo

$$\delta \mathbf{r}_i^k = -\frac{g_{ij}}{m_i} \mathbf{r}_{ij}(t) \quad (4-8a)$$

$$\delta \mathbf{r}_j^k = +\frac{g_{ij}}{m_j} \mathbf{r}_{ij}(t) \quad (4-8b)$$

onde

$$g_{ij} = \frac{[d_{ij}^2 - (\mathbf{r}_{ij}^*(t + \Delta t) \cdot \mathbf{r}_{ij}^*(t + \Delta t))]}{2M_{ij}(\mathbf{r}_{ij}(t) \cdot \mathbf{r}_{ij}^*(t + \Delta t))} \quad (4-9a)$$

$$M_{ij} = \frac{1}{m_i} + \frac{1}{m_j} \quad (4-9b)$$

O parâmetro g_{ij} é proporcional à distorção da distância ideal d_{ij} : $g_{ij} < 0$ significa que a ligação $i - j$ está esticada, enquanto $g_{ij} > 0$ significa que ela está comprimida. Em ambos os casos, as correções nas posições dos átomos i e j são feitas em sentidos opostos entre si, aproximando-os ou afastando-os (eq. 4-8a e 4-8b).

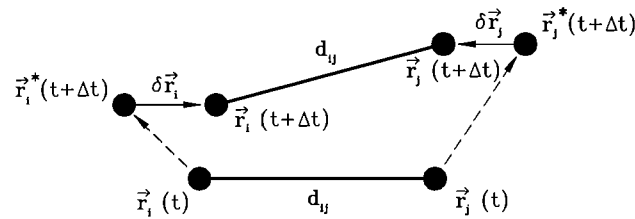


Fig. 4-6 : Representação esquemática do método shake para correção das posições atômicas e manutenção de vínculos de distância.

Ao fazer com que o k -ésimo vínculo seja satisfeito, em geral outros vínculos previamente satisfeitos são prejudicados. O procedimento é repetido iterativamente, até que todos os vínculos sejam adequadamente

satisfeitos, dentro de uma certa tolerância η , tal que $\frac{\mathbf{r}_{ij}^2 - d_{ij}^2}{d_{ij}^2} < \eta$,

com η tipicamente da ordem de 10^{-4} [37].

5 Modelagem de biomoléculas em água

Neste capítulo são apresentados os modelos para as moléculas de água incorporados ao banco de dados do programa THOR. São também apresentados os procedimentos básicos empregados nos algoritmos para a geração dos sistemas constituídos pelas biomoléculas de interesse e moléculas do solvente. As informações conformacionais dos sistemas gerados são armazenadas, em formatos apropriados, em arquivos para serem utilizados em simulações realizadas com o programa THOR.

5.1 Modelos para a água

Devido à fundamental importância da água em vários processos químicos e biológicos, esta tem sido, desde há muitos anos, intensivamente estudada por diferentes técnicas, entre as quais, a modelagem molecular [7,30,42]. Com o emprego desta técnica, o sucesso das investigações depende, essencialmente, da adoção de um modelo molecular razoável e um campo de força adequadamente parametrizado, de modo a reproduzir propriedades estruturais e dinâmicas observadas em experimentos.

Vários modelos para a representação de moléculas de água na fase líquida têm sido propostos, diferindo entre si basicamente em relação à geometria molecular e parâmetros envolvendo as interações intermoleculares, obtidos por ajustes de dados experimentais e resultados de cálculos quânticos [43-55].

Para a escolha do modelo a ser empregado, além da sua capacidade de reprodução de determinadas propriedades de interesse, a

simplicidade é um importante atributo a ser considerado, representando, em geral, vantagens quanto à economia do tempo computacional gasto nas simulações. Os modelos mais simples apresentados na literatura são os de três sítios, com as cargas localizadas nas posições das massas atômicas (fig. 5-1). Nesses modelos, as interações intermoleculares requerem o cálculo de somente nove distâncias para cada par de moléculas, enquanto este número pode ser muito maior em modelos mais sofisticados [51,55].

No banco de dados do programa THOR foram incluídos parâmetros de três modelos de três sítios: SPC (*simple point charge model*) [1], SPC/E (*extended simple point charge model*) [54] e TIP3P (*transferable intermolecular potential functions*) [55]. A diferença entre eles reside na distância b_0 entre os átomos de oxigênio e hidrogênio da molécula, no ângulo θ_0 definido pelas ligações dos átomos H - O - H (fig. 5-1), nos valores das cargas q_O e q_H e nos parâmetros de Lennard-Jones A_{OO} e B_{OO} (eq. 1-5). A geometria e os parâmetros que definem cada modelo são apresentados na tabela (5-1).

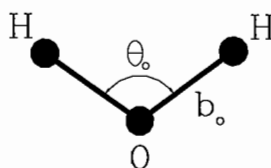


Fig. 5-1 : Modelo de três sítios para água, com as cargas localizadas nas posições das massas atômicas.

		SPC	SPC/E	TIP3P
b_0	(Å)	1,00	1,00	0,9572
θ_0	(grau)	109,47	109,47	104,52
q_H	(carga eletrônica)	+ 0,41	+ 0,4238	+ 0,417
q_O	(carga eletrônica)	- 0,82	- 0,8476	- 0,834
A_{OO}	(kcal mol ⁻¹ Å ¹²)	629320	629320	582000
B_{OO}	(kcal mol ⁻¹ Å ⁶)	625,50	625,50	595,00

Tabela 5-1: Geometria e parâmetros dos modelos SPC, SPC/E e TIP3P.

5.2 A caixa d'água

Foi desenvolvido um programa para a construção de sistemas formados por moléculas de água do modelo SPC, SPC/E ou TIP3P contidas em uma caixa retangular. As moléculas são posicionadas dentro da caixa de lados (L_x, L_y, L_z) por sorteios aleatórios, realizados de modo que o centro da caixa coincida com a origem do sistema de coordenadas, ou seja, as coordenadas atômicas cartesianas (x_i, y_i, z_i) são sorteadas no intervalo $\left(\pm \frac{L_x}{2}, \pm \frac{L_y}{2}, \pm \frac{L_z}{2}\right)$, respectivamente.

O número total de moléculas e a dimensão da caixa estão relacionados entre si pela densidade do sistema, condizente com a temperatura a ser simulada. Por exemplo, a uma temperatura de 25°C a densidade da água é 0,99567 g cm⁻³ [18].

Durante o sorteio, as sobreposições atômicas são evitadas com a definição de um parâmetro d_{\min} , que determina a menor distância interatômica permitida. Assim, uma molécula sorteada só é aceita se seus átomos estiverem a uma distância maior do que d_{\min} dos átomos de

todas as outras moléculas previamente sorteadas. Como os sistemas gerados pelo programa terão suas geometrias otimizadas pela técnica de minimização de energia, o valor de $d_{\text{mín}}$ parece não ser muito crítico. Algumas simulações foram realizadas e valores de $d_{\text{mín}}$ em torno de 2 Å mostraram-se razoáveis, com os sistemas convergindo rapidamente para um mínimo.

Concluídos os sorteios, a conformação gerada é armazenada em arquivos usados como entrada no programa THOR.

5.3 Biomoléculas em água

Para a construção de sistemas constituídos por biomoléculas em água, foi desenvolvido um programa que faz a junção de dois arquivos: um contendo a biomolécula e outro as moléculas de água na caixa. Uma confrontação é realizada entre as posições atômicas da biomolécula e das moléculas de água, sendo excluídas do sistema final todas as moléculas de água cujos átomos estão separados da biomolécula por uma distância menor do que um valor mínimo permitido $d_{\text{mín}}$. Conforme discutido no item anterior, um valor de $d_{\text{mín}}$ em torno de 2 Å parece ser razoável ².

A conformação do sistema final é armazenada em arquivos usados como entrada no programa THOR.

Uma apresentação esquemática do procedimento acima descrito pode ser apreciada na figura 5-2.

² Procedimento semelhante foi empregado em simulações de biomoléculas em água [56] realizadas com o GROMOS [1], com mesmo valor adotado para $d_{\text{mín}}$.

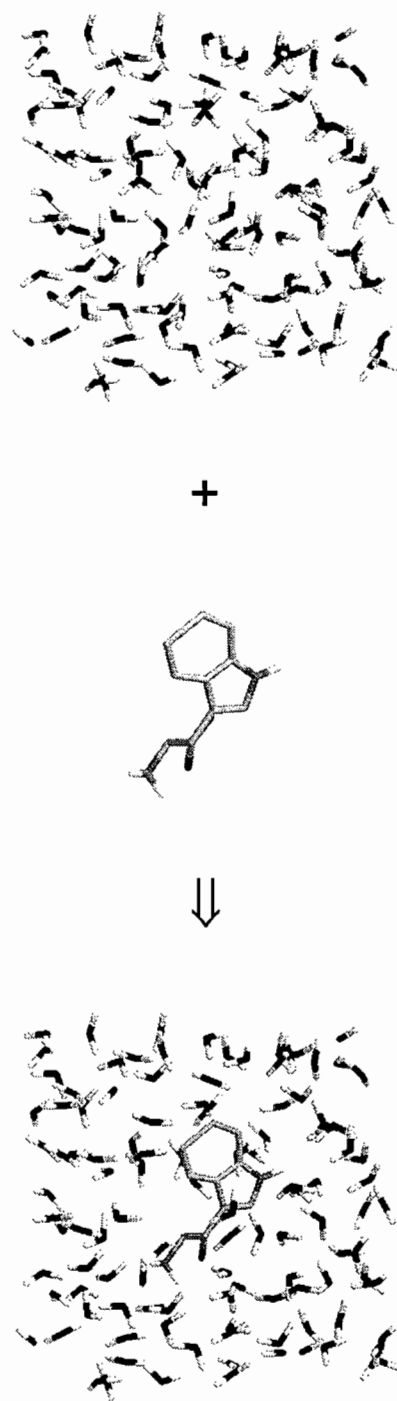


Fig. 5-2 : Esquema utilizado para a geração de sistemas constituídos por biomoléculas em água.

6 Resultados

Neste capítulo são apresentados resultados obtidos em estudos de dinâmica molecular realizados com o programa THOR modificado, com representação explícita das moléculas do solvente. Propriedades estruturais da água na fase líquida, bem como sua influência sobre propriedades conformacionais dos sistemas simulados são analisadas.

6.1 Água

Um sistema constituído por 216 moléculas de água do modelo SPC em uma caixa cúbica de lado 18,64 Å (fig. 6-1) foi gerado conforme procedimento descrito no capítulo 5. A geometria do sistema foi otimizada pela técnica de minimização de energia, com o emprego consecutivo dos métodos *steepest descent* e dos gradientes conjugados. Foram definidos valores nulos para as velocidades atômicas iniciais e a dinâmica molecular foi iniciada com um aquecimento do sistema até 300 K nos primeiros 10 ps. Alcançada esta temperatura, sua manutenção foi obtida com um reescalonamento direto das velocidades (eq. 3-12 e 3-14) até 400 ps. A partir daí, o sistema foi acoplado fracamente a um banho térmico, com tempo de relaxação característico $\tau = 0,1$ ps (eq. 3-12 e 3-15), e a dinâmica prosseguiu até que um tempo total de 1000 ps fosse completado. O método *shake* foi empregado e as distâncias entre os átomos de oxigênio e hidrogênio das moléculas de água ($O_w - H_w$ e $H_w - H_w$) foram mantidas constantes, com vínculos impostos de acordo com o modelo assumido. As equações de movimento (3-6) e (3-8) foram integradas com um passo de tempo de 2 fs.

Um raio de corte de 9 Å foi definido para os cálculos das interações entre átomos não-ligados. Variações da temperatura e energias cinética, potencial e total foram registradas a cada 0,1 ps, enquanto as conformações do sistema a cada 1 ps.

O mesmo procedimento foi utilizado para a simulação de sistemas constituídos por moléculas de água dos modelos SPC/E e TIP3P.

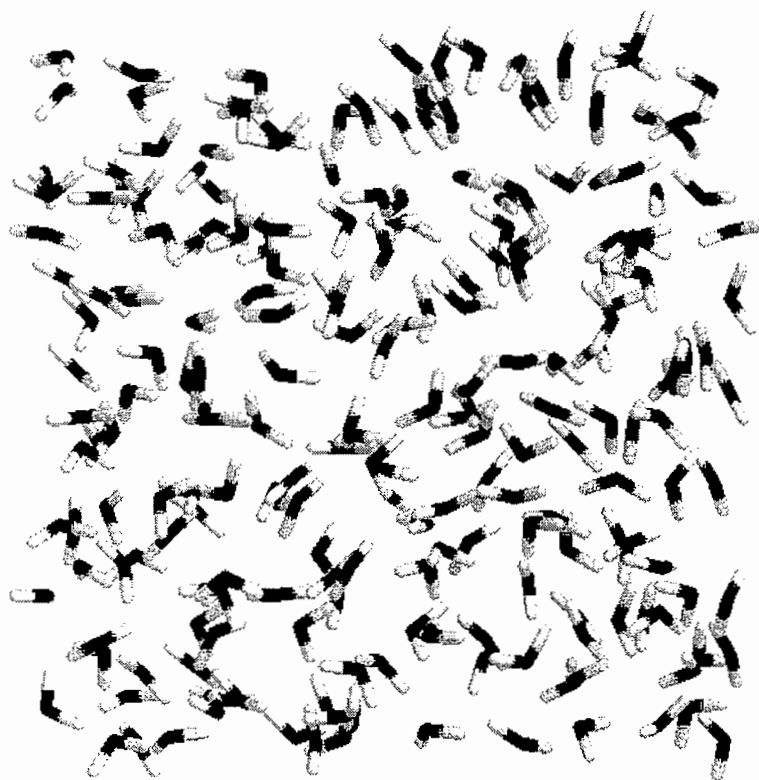


Fig. 6-1 : Sistema gerado por sorteios aleatórios das posições de 216 moléculas de água do modelo SPC em uma caixa cúbica de lado 18,64 Å.

6.1.1 Função de distribuição radial de pares

As propriedades estruturais de sistemas líquidos são usualmente caracterizadas pela função de distribuição radial de pares $g(r)$.

A função $g(r)$ dá a probabilidade de encontrar um par de átomos separados por uma distância r , relativa à probabilidade esperada para uma distribuição totalmente aleatória e com a mesma densidade [30].

Para dois tipos de átomos A e B do sistema, a função $g_{AB}(r)$ é definida como [7,30,42]

$$g_{AB}(r) = \frac{\rho_{AB}(r)}{\rho_B^0} \quad (6-1)$$

onde $\rho_{AB}(r)$ é a densidade média de átomos do tipo B a uma distância entre r e $r + dr$ de átomos do tipo A e $\rho_B^0 = \frac{N_B}{V}$ é a densidade média de átomos do tipo B no sistema, sendo N_B o número de átomos do tipo B e V o volume do sistema. A função $g_{AB}(r)$ evidencia, portanto, as flutuações na densidade do sistema devido à estruturação local dos átomos B em torno dos átomos A.

A integral de $g_{AB}(r)$ calculada para os limites de integração r_1 e r_2 resulta no denominado número de coordenação, ou seja, o número médio de átomos do tipo B ao redor de átomos do tipo A a uma distância entre r_1 e r_2 :

$$n_{AB} = 4\pi \rho_B^0 \int_{r_1}^{r_2} g_{AB}(r) r^2 dr \quad (6-2)$$

Experimentalmente a função $g(r)$ e o número de coordenação são obtidos por técnicas de difração de raios-X, nêutrons e elétrons [57-62].

A análise estrutural dos sistemas simulados foi realizada com a utilização de um programa desenvolvido para a determinação das funções de distribuição radial de pares.

6.1.2 Resultados

Um acompanhamento temporal das energias cinética, potencial e total foi realizado durante a dinâmica dos três sistemas simulados. A figura 6-2 mostra o comportamento energético do sistema SPC. Comportamento semelhante foi apresentado pelos sistemas SPC/E e TIP3P.

As fases de aquecimento, reescalamento direto das velocidades atômicas e acoplamento fraco do sistema a um banho térmico podem ser distinguidas visualmente na figura 6-2. A coleta de informações para os cálculos das funções de distribuição radial foi feita com gravações das conformações dos sistemas obtidas entre 500 e 1000 ps de dinâmica.

A figura 6-3 mostra as funções de distribuição radial calculadas para os pares $O_w - O_w$, $O_w - H_w$ e $H_w - H_w$ nos três sistemas estudados. Os três modelos apresentam um comportamento estrutural bastante similar quando simulados com o programa THOR, com pequenas diferenças quanto às posições e alturas dos picos observados nas distribuições.

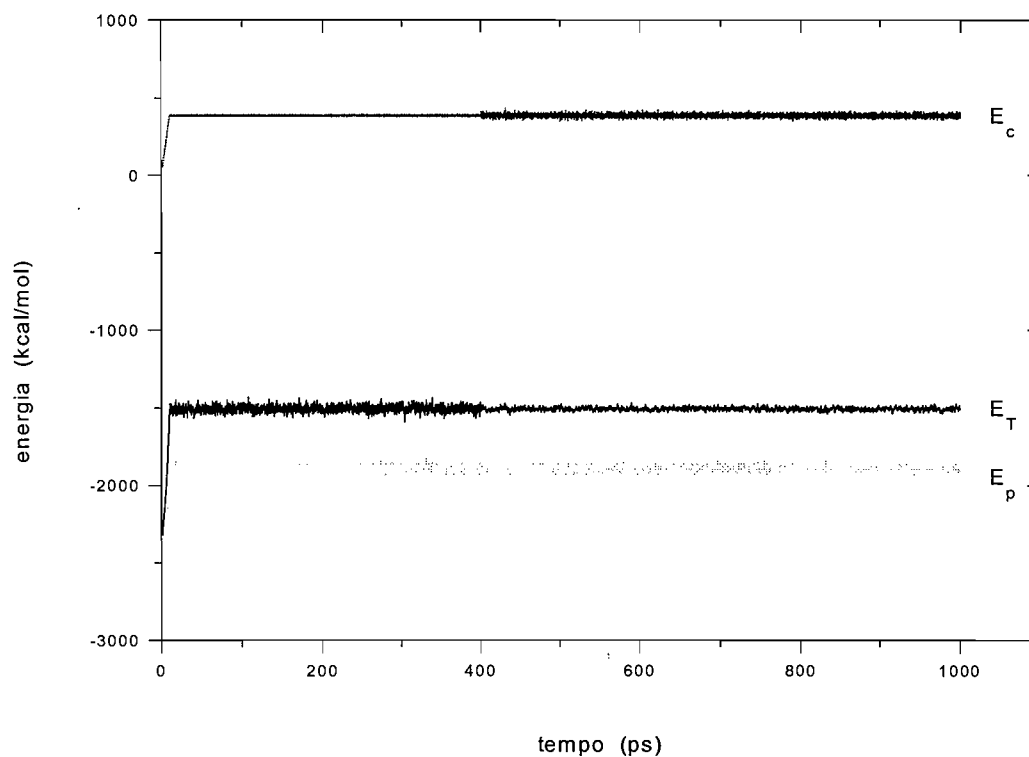


Fig. 6-2 : Acompanhamento temporal das energias cinética (E_c), potencial (E_p) e total (E_T) durante a dinâmica do sistema simulado com água do modelo SPC.

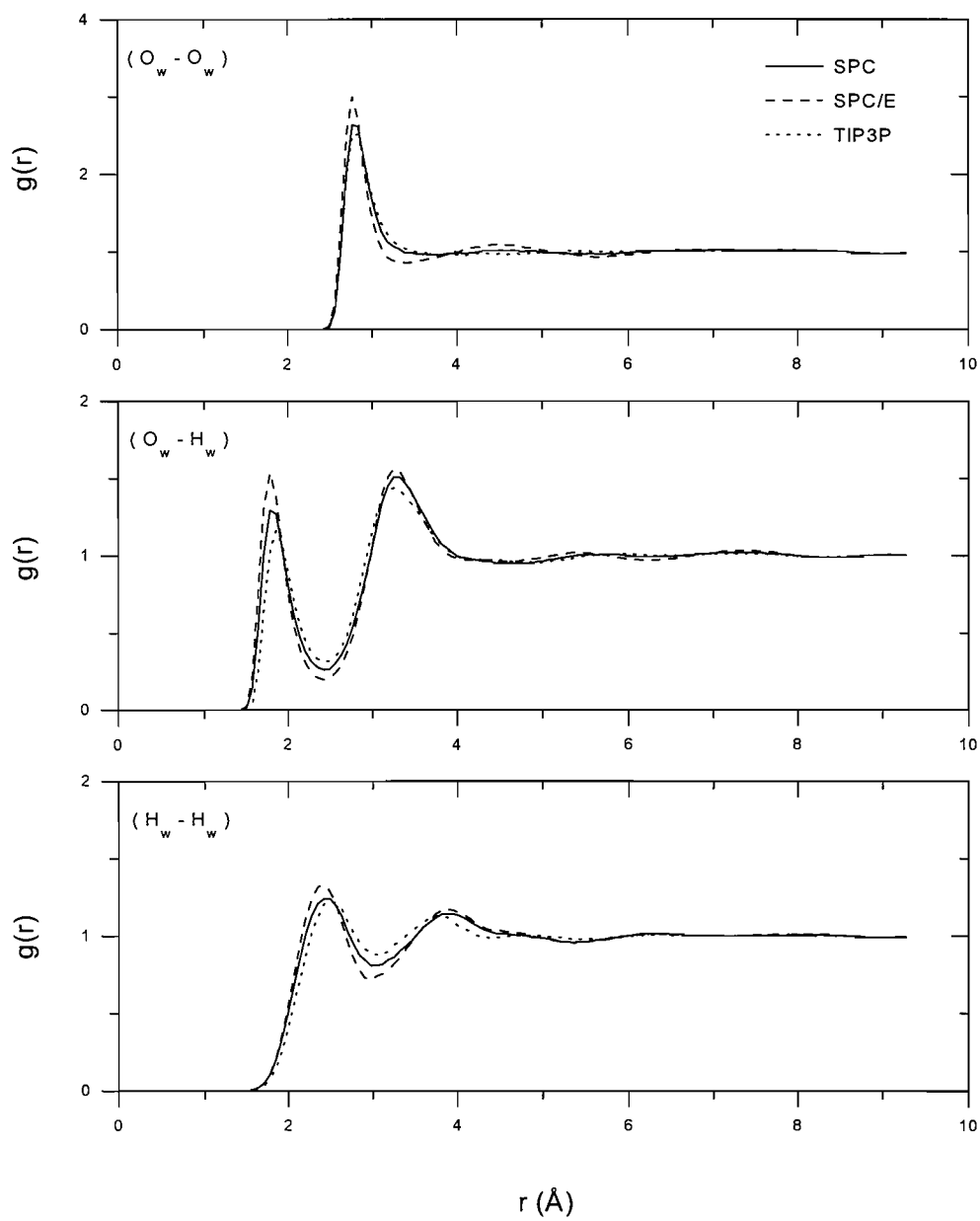


Fig. 6-3 : Funções de distribuição de pares para os átomos de oxigênio (O_w) e hidrogênio (H_w) dos modelos SPC (—), SPC/E(---) e TIP3P (.....).

As funções obtidas para os pares $O_w - O_w$ apresentam pouca estruturação além do primeiro pico. Para o modelo SPC/E, um segundo pico, embora débil, pode ser observado. Esta parece mesmo ser uma deficiência dos modelos de três sítios para a água, uma vez que o segundo e terceiro picos são claramente observados em dados de raios-X [48,55]. Funções mais estruturadas podem ser obtidas com o emprego de modelos mais sofisticados [42,55]. Como esperado pela análise de resultados experimentais, as funções para os pares $O_w - H_w$ apresentam dois picos, sendo o primeiro menos intenso e mais estreito que o segundo. O primeiro pico corresponde à separação atômica devido às interações de ponte de hidrogênio $O_w \cdots H_w$ entre moléculas vizinhas. As funções de distribuição radial para os pares $H_w - H_w$ também apresentam os dois picos observados experimentalmente. Os resultados encontrados estão em razoável acordo com aqueles obtidos experimentalmente [57-62] e em simulações de Monte Carlo [46-48,51,52,55] e dinâmica molecular [32,49,50,53,54].

A tabela 6-1 apresenta os valores obtidos para as posições dos primeiros e segundos picos das funções de distribuição radial calculadas. Nesta tabela também são apresentados os valores estimados em experimentos de difração de nêutrons e raios-X [58,62].

A integral do primeiro pico da função $g(r)$ para os pares $O_w - O_w$ até a distância de 3,5 Å resulta em um número de coordenação (eq. 6-2) de 5,19 para o modelo SPC, 5,02 para o modelo SPC/E e 5,28 para o modelo TIP3P, enquanto o valor experimental é 5 átomos de oxigênio [48,55]. Para os pares $O_w - H_w$, a integral do primeiro pico até a distância de 2,5 Å resulta em 1,99 para os modelos SPC e SPC/E e 1,98 para o modelo TIP3P, enquanto o valor experimental é 2 átomos de hidrogênio [55].

$g(r)$	SPC	SPC/E	TIP3P	ref. [58]	ref. [62]
$O_w - O_w$	2,77	2,77	2,81	2,85	2,85
$H_w - O_w$	1,83 3,29	1,79 3,29	1,85 3,25	1,92 3,28	1,86 3,20
$H_w - H_w$	2,49 3,91	2,37 3,89	2,49 3,73	2,40 3,58	2,29 3,90

Tabela 6-1: Posições (Å) dos primeiros e segundos picos das funções de distribuição radial obtidas para os modelos SPC, SPC/E e TIP3P e estimadas experimentalmente por difração de nêutrons e raios-X [58,62].

A análise dos resultados obtidos com as simulações dos três modelos para água mostrou que eles são equivalentes quando simulados com o programa THOR. Os três modelos conseguiram reproduzir, razoavelmente e de maneira bastante similar, a forma e as posições dos picos das funções de distribuição radial de pares esperadas. Observando as três funções para os pares $O_w - O_w$, $O_w - H_w$ e $H_w - H_w$ (fig. 6-3), aquelas calculadas para o modelo SPC mostraram, de maneira sistemática, um comportamento intermediário em relação às posições e alturas dos picos, quando comparadas com aquelas obtidas para os modelos SPC/E e TIP3P. O modelo SPC foi então o escolhido para ser adotado como solvente em simulações de biomoléculas em água representada explicitamente, apresentadas a seguir. No entanto, pela similaridade dos resultados, acredita-se que, se outro modelo fosse empregado, a essência dos resultados não seria modificada.

6.2 Polialanina

A alanina é um aminoácido fundamental, de cadeia lateral curta, apolar, sem preferência conformacional isoladamente (fig. 6-4). A alanina apresenta, no entanto, uma propensão helicoidal intrínseca [13,63-66], de modo que cadeias de alanina comumente são encontradas formando estruturas em hélice- α [14,15,67]. Em estudos de predição da estrutura secundária de polialaninas, conformações em hélice- α são encontradas para cadeias compostas por vinte resíduos [68]. Além disso, hélices em proteínas apresentam, em geral, uma alta concentração de alanina [14,67].

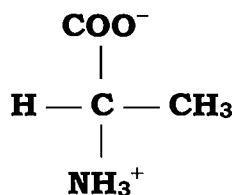


Fig. 6-4 : Fórmula estrutural do aminoácido alanina.

A hélice- α constitui-se no elemento mais abundante presente na estrutura secundária de proteínas [14,66]. Mesmo peptídios relativamente pequenos podem conter substancial caráter helicoidal em soluções aquosas, conforme evidências experimentais obtidas por meio das técnicas de dicroísmo circular e ressonância magnética nuclear [64,65,69]. Cálculos de energia de interação eletrostática entre resíduos

em hélices- α indicam, no entanto, que hélices isoladas só são estáveis se forem formadas por mais do que quatorze resíduos [70].

Simulações de dinâmica molecular, realizadas com o programa THOR, com solvente representado como um contínuo, mostram que a estrutura helicoidal de uma cadeia formada por vinte alaninas, com terminais não carregados, é rapidamente desfeita quando $\epsilon = 80$ e permanece estável quando $\epsilon = 2$. Simulações da mesma cadeia, mas em conformação inicial *trans*, mostram que esta não evolui para uma conformação que indique uma possível formação de hélice em um nanossegundo de dinâmica [4]. Ambos os resultados contradizem as expectativas teóricas e experimentais expostas acima.

6.2.1 Simulação

Estruturas em hélice- α são definidas por padrões de repetição envolvendo pontes de hidrogênio entre o grupo C = O do resíduo n e o grupo N - H do resíduo $n + 4$ (fig. 6-5) e valores dos ângulos diedros ϕ e ψ , definidos respectivamente em torno das ligações C - N - C $_{\alpha}$ - C e N - C $_{\alpha}$ - C - N (fig. 6-6), mantidos próximos a -60° e -40° , para todos os resíduos envolvidos [14]. O passo da hélice- α mede 5,4 Å em média e cada volta completa contém cerca de 3,6 resíduos. Deste modo, os resíduos estão deslocados entre si por uma translação de aproximadamente 1,5 Å ao longo do eixo da hélice [12-14].



Fig. 6-5 : Pontes de hidrogênio entre o grupo C = O do resíduo n e o grupo N - H do resíduo n + 4, mantendo a estrutura em hélice- α da polialanina.

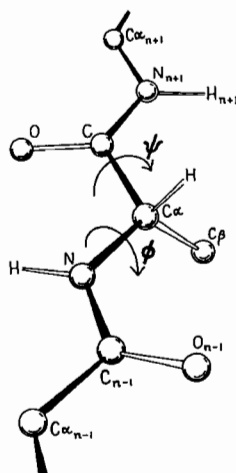
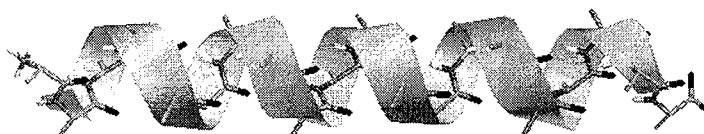
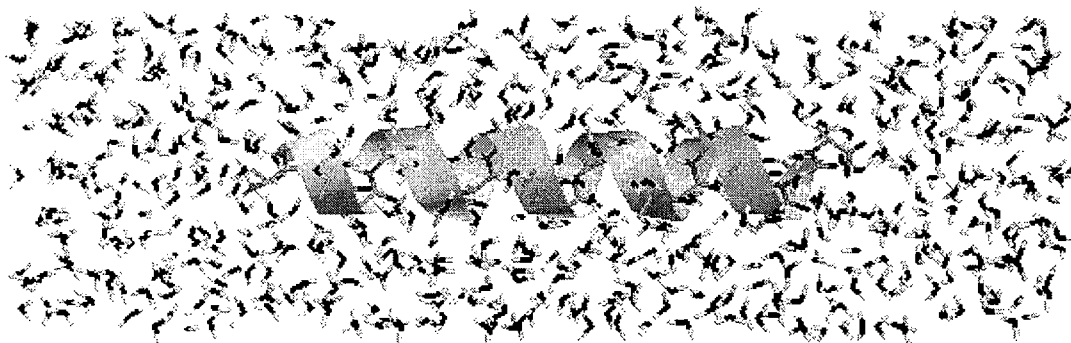


Fig. 6-6 : Definição dos ângulos diedros ϕ e ψ do esqueleto peptídico, em torno das ligações C - N - C_{α} - C e N - C_{α} - C - N, respectivamente. (Fig. adaptada da ref. [14]).

Um sistema *zwitteriônico*, com terminais NH_3^+ e COO^- , formado por vinte alaninas em conformação hélice- α , foi construído com a imposição dos valores apropriados aos ângulos ϕ e ψ do esqueleto peptídico. O sistema foi então colocado em uma caixa retangular de lados 18,64 Å, 18,64 Å e 62,16 Å, conforme procedimento descrito no capítulo 5, com 645 moléculas de água do modelo SPC (fig. 6-7).



(a)



(b)

Fig. 6-7 : Polialanina em conformação hélice- α (a) em água (b).

A geometria do sistema foi otimizada pela técnica de minimização de energia, com o emprego consecutivo dos métodos *steepest descent* e dos gradientes conjugados. Foram definidos valores nulos para as velocidades atômicas iniciais e a dinâmica molecular foi iniciada com um aquecimento do sistema até 300 K nos primeiros 10 ps. Alcançada esta temperatura, sua manutenção foi obtida com um reescalonamento direto das velocidades (eq. 3-12 e 3-14) até 50 ps. A partir daí, o sistema foi acoplado fracamente a um banho térmico, com tempo de relaxação característico $\tau = 0,1$ ps (eq. 3-12 e 3-15), e a dinâmica prosseguiu até que um tempo total de 300 ps fosse completado. O método *shake* foi empregado e as distâncias $O_w - H_w$ e $H_w - H_w$ das moléculas de água e $N - H$ da polialanina foram mantidas constantes com a imposição de vínculos apropriados. O modelo de átomos unidos (ver apêndice A) foi utilizado para os átomos de hidrogênio ligados aos átomos de carbono C_α e C_β , do grupo metila. As equações de movimento (3-6) e (3-8) foram integradas com um passo de tempo de 1 fs. Um raio de corte de 9 Å foi definido para os cálculos das interações entre átomos não-ligados. Variações da temperatura e energias cinética, potencial e total foram registradas a cada 0,1 ps. As conformações do sistema e os ângulos diedros ϕ e ψ foram acompanhados e registrados a cada 0,1 ps.

Para uma avaliação do efeito do solvente explícito sobre a conformação helicoidal da polialanina *zwitteriônica*, a mesma conformação inicial foi simulada em solvente representado como um contínuo, caracterizado pela constante dielétrica $\epsilon = 80$. Complementando o estudo, uma terceira simulação foi realizada, com o sistema em um meio contínuo com $\epsilon = 1$, representando o vácuo.

6.2.2 Resultados

O comportamento energético dos três sistemas simulados foi acompanhado durante a dinâmica. A figura 6-8 apresenta o resultado obtido para o sistema polialanina em solvente explícito. As fases de aquecimento, reescalonamento direto das velocidades atômicas e acoplamento fraco do sistema a um banho térmico podem ser identificadas visualmente nesta figura.

Na figura 6-9 são apresentados os diagramas de Ramachandran [17,71] obtidos para a polialanina modelada em solvente explícito (a), em solvente representado como um meio contínuo, com $\epsilon = 80$ (b) e no vácuo, com $\epsilon = 1$ (c), em diferentes fases da dinâmica: de 0 a 100 ps, de 100 a 200 ps e de 200 a 300 ps. A evolução temporal dos ângulos diedros ϕ e ψ indica uma forte dependência com meio. Na simulação em solvente explícito, esses ângulos permanecem, durante toda a dinâmica, com valores próximos aos iniciais. Valores médios de $(-60 \pm 7)^\circ$ e $(-46 \pm 7)^\circ$ são estimados, respectivamente, para ϕ e ψ , estando em acordo com os valores $(-57^\circ, -47^\circ)$, que definem as estruturas em hélice- α [17]. Nos outros dois sistemas, o mesmo não é observado. Logo nas primeiras dezenas de picossegundos de dinâmica a hélice já é desfeita, com as conformações ocupando também o quadrante superior esquerdo do diagrama, com alto nível de dispersão, no meio contínuo com $\epsilon = 80$, e os quadrantes da direita, no vácuo.

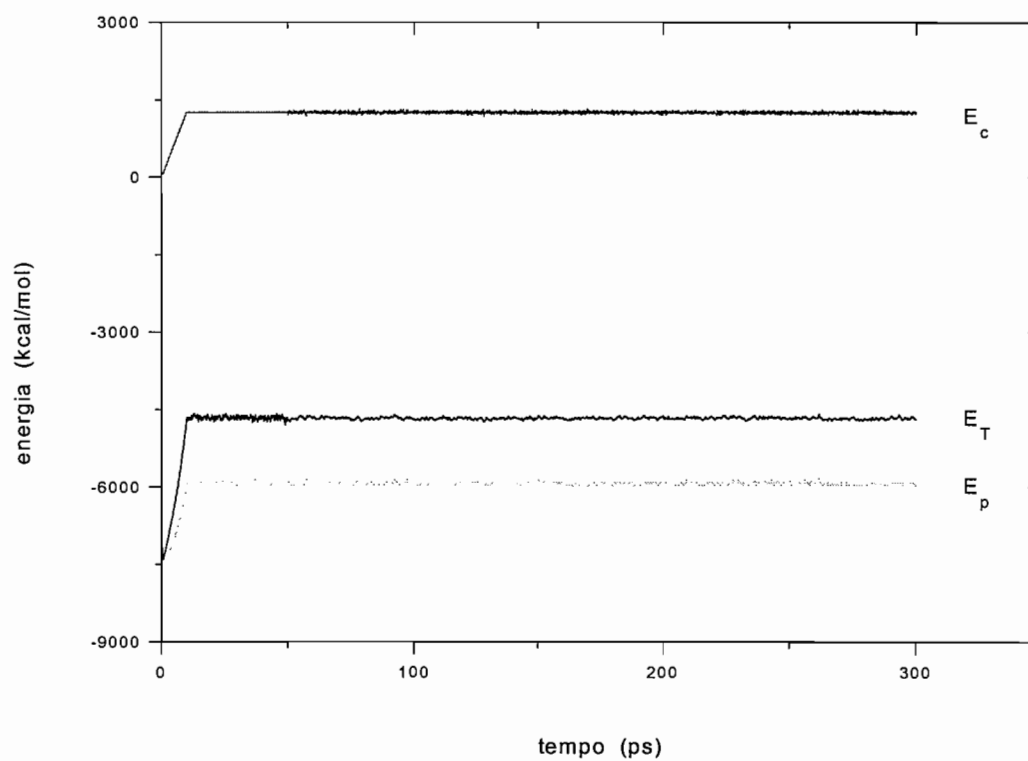


Fig. 6-8 : Acompanhamento temporal das energias cinética (E_c), potencial (E_p) e total (E_T) para o sistema composto pela polialanina e moléculas de água.

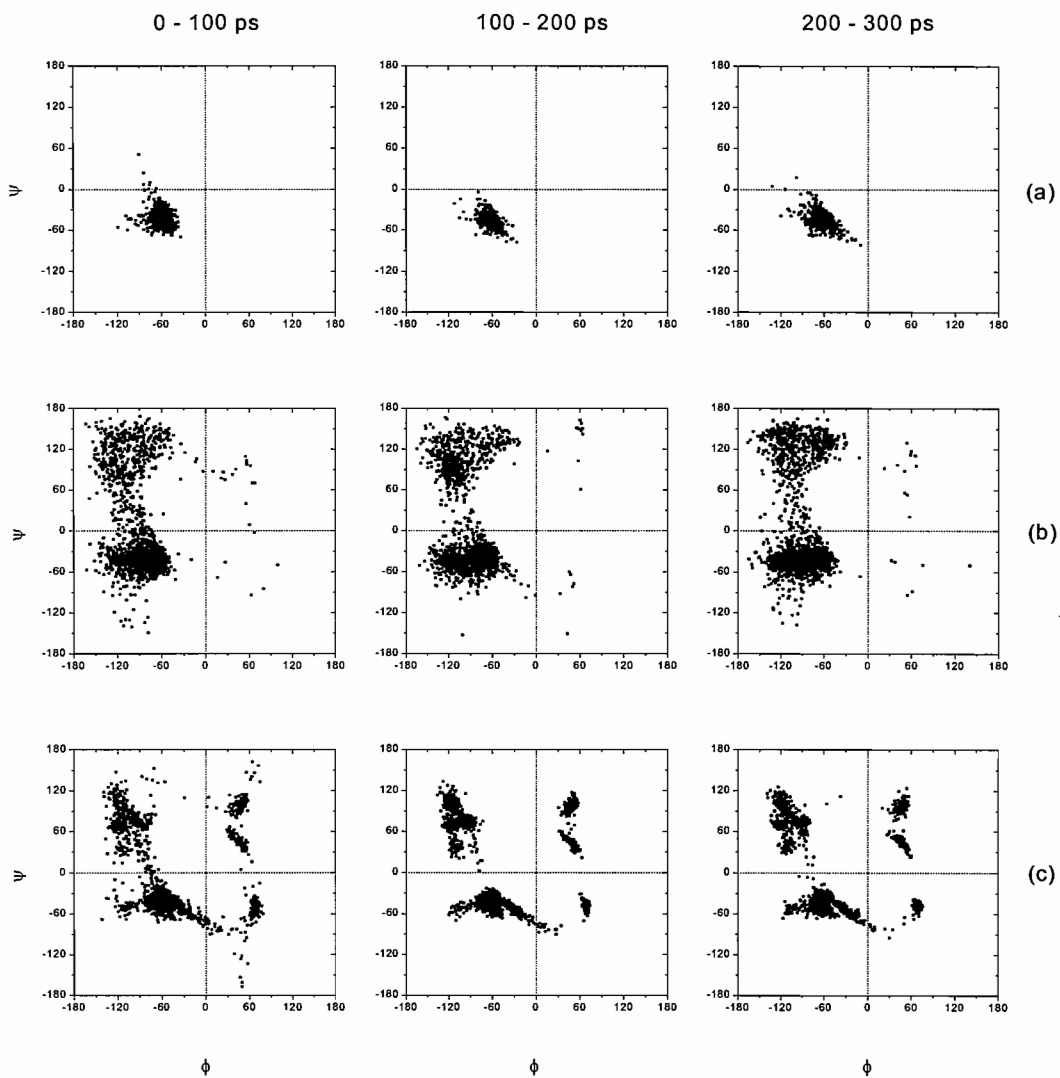


Fig. 6-9 : Diagramas de Ramachandran obtidos em diferentes estágios da dinâmica dos sistemas: polialanina em solvente explícito (a), em meio contínuo, com $\epsilon = 80$ (b) e no vácuo , com $\epsilon = 1$ (c).

A figura 6-10 apresenta o resultado da monitorização da distância entre os átomos de nitrogênio do terminal amina e carbono do terminal carboxila, aqui identificados por N_T e C_T , durante a dinâmica da polialanina nos três meios. Para uma hélice- α ideal, composta por vinte resíduos, o valor esperado para essa distância é 30 Å [12-14]. Porém, a figura 6-10 mostra que, enquanto a distância $N_T - C_T$ manteve-se próxima da distância ideal na simulação em solvente explícito, nos outros dois casos os terminais sofreram uma atração, de menor ou maior intensidade, nos meios com $\epsilon = 80$ e $\epsilon = 1$, respectivamente. Distâncias médias $N_T - C_T$ foram calculadas para os três sistemas, no intervalo de dinâmica entre 100 e 300 ps, sendo obtidos os valores (30 ± 1) Å, (8 ± 3) Å e $(3,1 \pm 0,1)$ Å, respectivamente para as simulações no solvente explícito, no meio contínuo e no vácuo.

A figura 6-11 mostra a conformação da polialanina após 300 ps de dinâmica nos três meios, evidenciando a desestabilização da hélice, quando simulada nos meios contínuos, e a sua manutenção, quando simulada em solvente explícito.

A estruturação do solvente ao redor da polialanina pode ser avaliada por meio das funções de distribuição radial de pares. A figura 6-12 apresenta as distribuições das moléculas de água ao redor dos átomos de carbono C_β , hidrogênio H_T do terminal amina e oxigênio O_T do terminal carboxila, calculadas a partir das conformações obtidas entre 100 ps e 300 ps de dinâmica. É clara a diferença entre as funções $g(r)$ apresentadas, no que diz respeito à forma e posições dos picos. Para os pares $C_\beta - O_w$, a distribuição é larga, com pico em 3,40 Å, característica da solvatação de grupos apolares [7,72,73], enquanto uma forte estruturação, caracterizada por picos estreitos e bem definidos, é evidenciada para as distribuições da água ao redor dos terminais.

As posições dos picos em 1,93 Å e 1,86 Å, respectivamente para os pares $H_T - O_W$ e $O_T - H_W$, indicam interação por pontes de hidrogênio entre os terminais e moléculas de água vizinhas.

Portanto, a análise dos resultados apresentados indica que, na simulação da polialanina em solvente explícito, a estruturação da água ao redor dos terminais atua no sentido de blindar a interação de caráter atrativo entre eles, enquanto a hélice, circundada pela cavidade de exclusão hidrofóbica, permanece protegida e estável. A presença dessa cavidade, protegendo o esqueleto peptídico, é confirmada pelo cálculo das funções de distribuição radial de pares $H - O_W$ e $O - H_W$.

A função de distribuição radial foi calculada para os átomos de oxigênio do grupo $C = O$ e hidrogênio do grupo $N - H$ da polialanina em solvente explícito, apresentando um pico em 1,92 Å, correspondente à ponte de hidrogênio $O \cdots H$ entre os resíduos n e $n + 4$. Conformações registradas durante a dinâmica foram monitoradas com o programa RASMOL [74], que revelou uma variação entre 15 e 17 pontes de hidrogênio envolvendo os resíduos n e $n + 4$.

A figura 6-13 apresenta as funções de distribuição radial para as moléculas de água no sistema polialanina em água. Pouca diferença é observada quando essas funções são comparadas com as funções apresentadas na figura 6-3, para sistemas compostos por água pura. Para os pares $O_W - O_W$ o pico da distribuição encontra-se posicionado em 2,77 Å, para os pares $O_W - H_W$ as posições do primeiro e segundo picos são 1,79 Å e 3,27 Å, respectivamente, enquanto 2,43 Å e 3,89 Å são as posições do primeiro e segundo picos para os pares $H_W - H_W$.

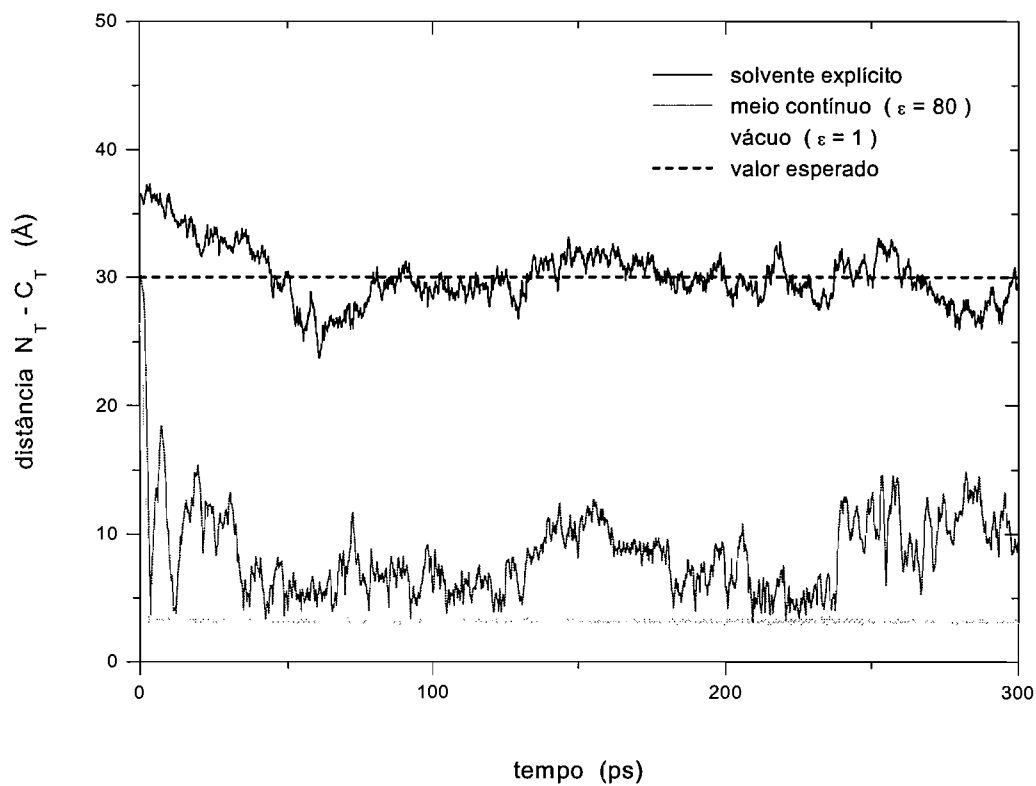
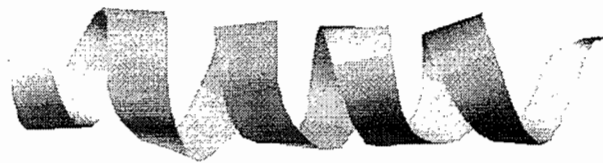
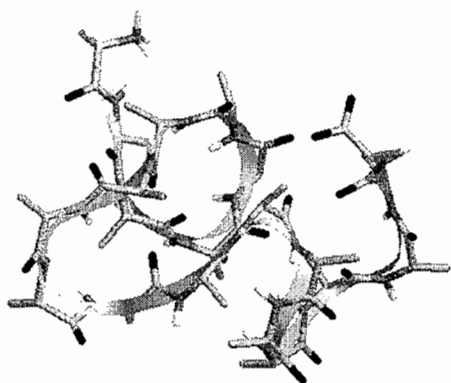


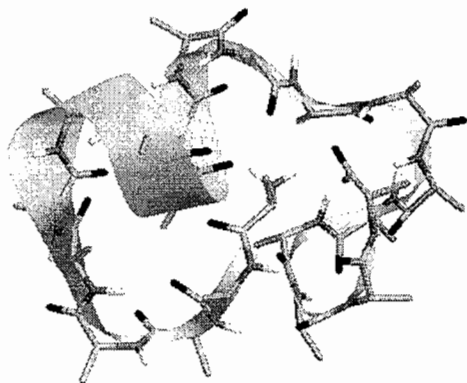
Fig. 6-10 : Distância entre os átomos de nitrogênio do terminal amina (N_T) e carbono do terminal carboxila (C_T) da polialanina em solvente explícito, em meio contínuo ($\epsilon = 80$) e no vácuo ($\epsilon = 1$). A reta tracejada representa o valor esperado para uma hélice- α ideal, composta por vinte resíduos.



(a)



(b)



(c)

Fig. 6-11 : Conformação da polialanina após 300 ps de dinâmica em solvente explícito (a), em meio contínuo, com $\epsilon = 80$ (b) e no vácuo, com $\epsilon = 1$ (c).

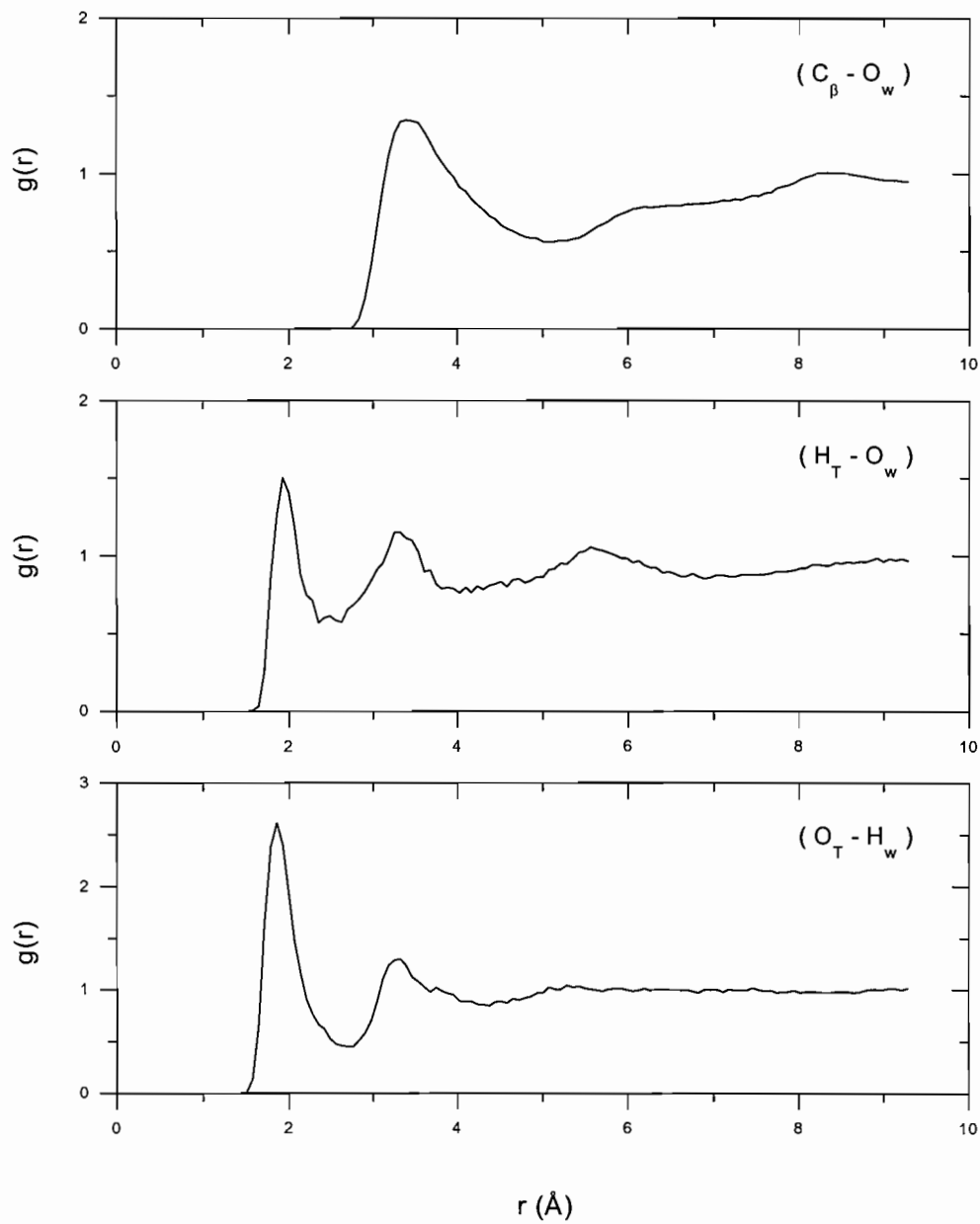


Fig. 6-12 : Funções de distribuição radial de pares para moléculas de água (átomos O_w e H_w) ao redor dos átomos C_{β} (grupo metila), H_T (terminal amina) e O_T (terminal hidroxila) da polialanina.

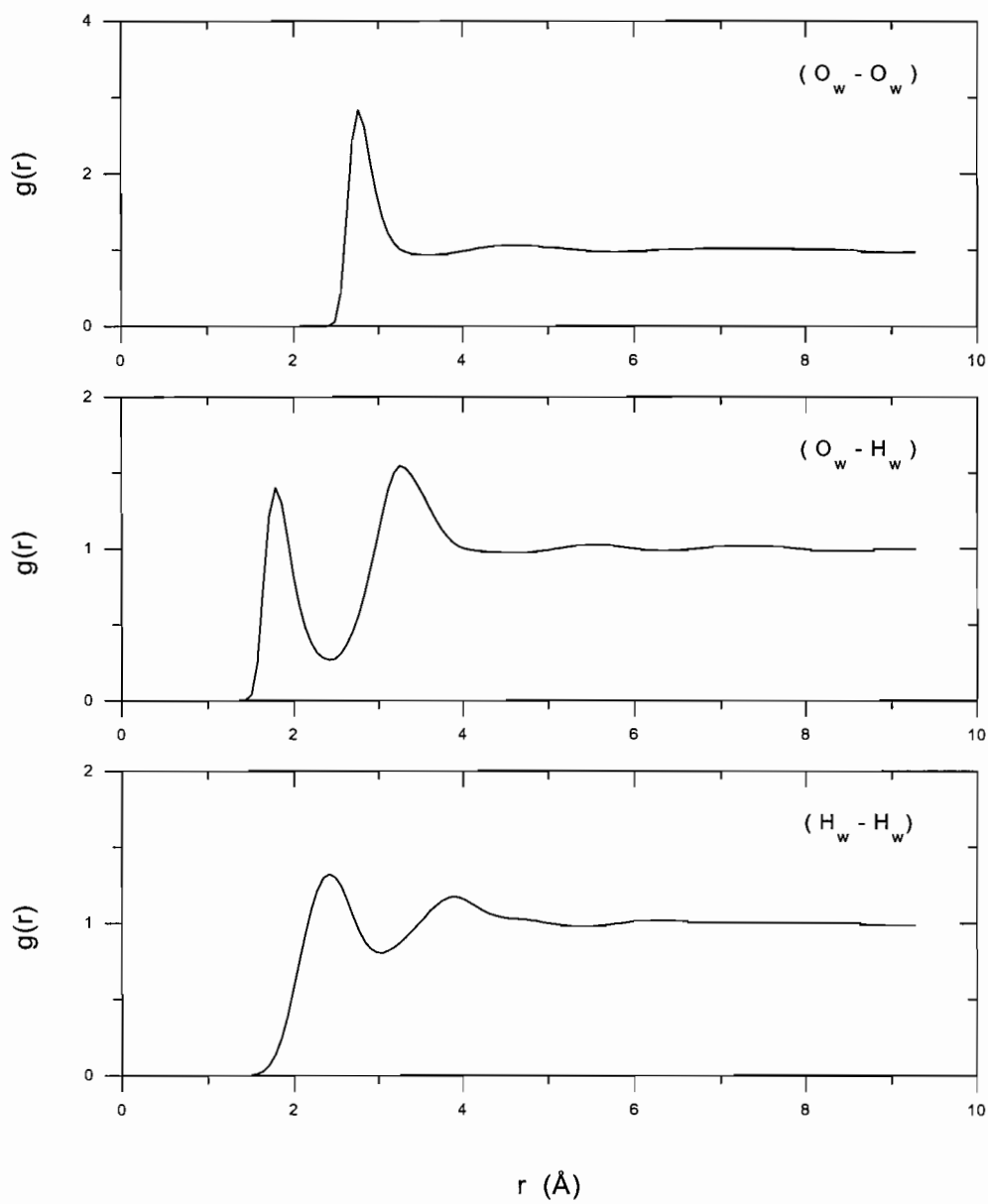


Fig. 6-13 : Funções de distribuição radial de pares para moléculas de água no sistema polialanina em água.

Apesar de todos os resultados apresentados e discutidos apontarem para a estruturação da água ao redor da polialanina como a principal responsável pela diferença entre o comportamento conformacional dos sistemas estudados, simulações em $\epsilon = 80$ e $\epsilon = 1$ foram refeitas, agora com o emprego de um raio de corte, utilizado, até então, somente na simulação em solvente explícito. O objetivo foi esclarecer se o emprego de um corte nas interações entre átomos não-ligados nos sistemas modelados em meio contínuo também não levaria a um bloqueio nas interações atrativas entre os terminais e, conseqüentemente, a uma manutenção da estrutura helicoidal desses sistemas.

O mesmo procedimento descrito anteriormente foi utilizado para a otimização da geometria e dinâmica de 300 ps dos dois sistemas. As conformações e os ângulos diedros ϕ e ψ foram acompanhados durante a dinâmica e registrados a cada 0,1 ps.

A figura 6-14 apresenta os diagramas de Ramachandran obtidos, em diferentes estágios da dinâmica, para a polialanina simulada em meio contínuo, com $\epsilon = 80$ (a) e no vácuo, com $\epsilon = 1$ (b). Os diagramas mostram o mesmo comportamento daqueles apresentados na figura 6-9, obtidos a partir das simulações sem a adoção de um raio de corte. A hélice é igualmente desestabilizada nas primeiras dezenas de picossegundos de dinâmica. No meio aquoso contínuo a ocupação adicional do quadrante superior esquerdo, com alto nível de dispersão, é imediata, enquanto no vácuo os quadrantes da direita também são ocupados.

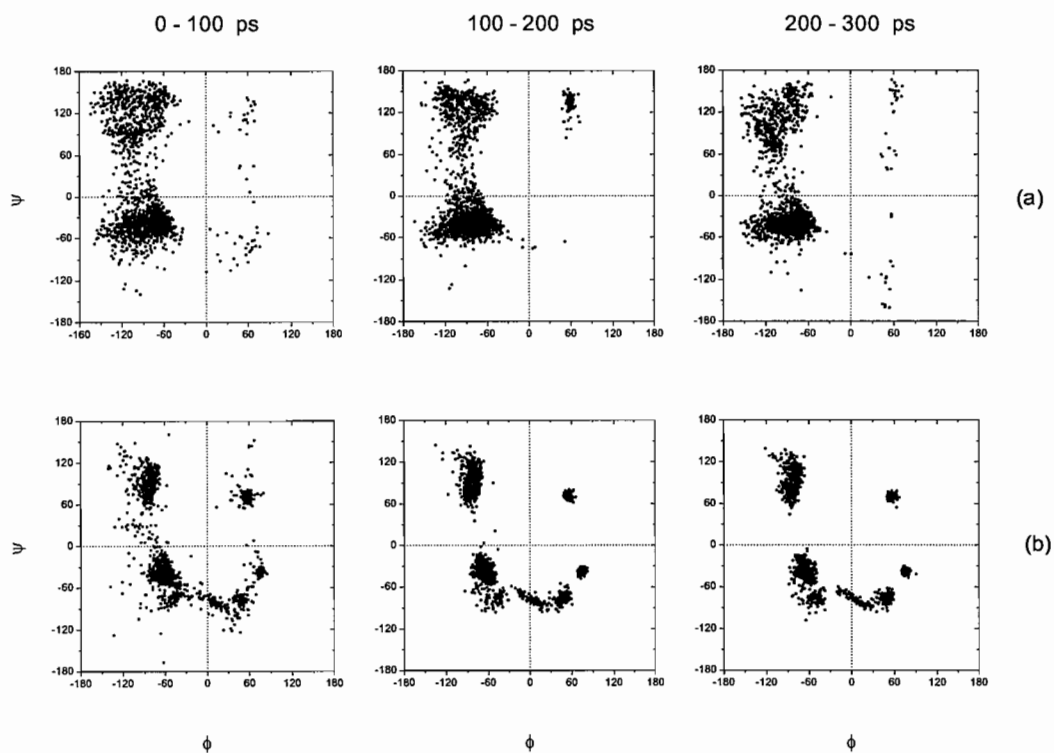


Fig. 6-14 : Diagramas de Ramachandran obtidos em diferentes estgios da dinmica realizada com o emprego de um raio de corte para as interaces entre tomos no-ligados dos sistemas: polialanina em meio contnuo, com $\epsilon = 80$ (a) e no vcuo , com $\epsilon = 1$ (b).

Conforme exposto no início do capítulo e de acordo com referências citadas, espera-se que conformações helicoidais de pequenas cadeias de alanina não sejam estáveis em soluções aquosas. Com o objetivo de verificar a habilidade do programa THOR em reproduzir a relação entre estabilidade helicoidal e tamanho da cadeia, uma cadeia composta por oito alaninas, em conformação inicial hélice- α foi gerada e simulada.

Seguindo o mesmo procedimento descrito anteriormente, um sistema *zwitteriônico* foi construído e colocado em uma caixa retangular de lados 18,64 Å, 18,64 Å e 27,97 Å, contendo 291 moléculas de água. A geometria do sistema foi otimizada e uma dinâmica de 300 ps foi realizada. As conformações do sistema e os ângulos diedros ϕ e ψ foram acompanhados durante a dinâmica e registrados a cada 0,1 ps.

A figura 6-15 apresenta os diagramas de Ramachandran obtidos para o sistema, em diferentes fases da dinâmica. A figura mostra que, ao contrário do sistema formado por vinte alaninas em água explícita, o sistema helicoidal formado por oito alaninas não permanece estável durante a dinâmica.

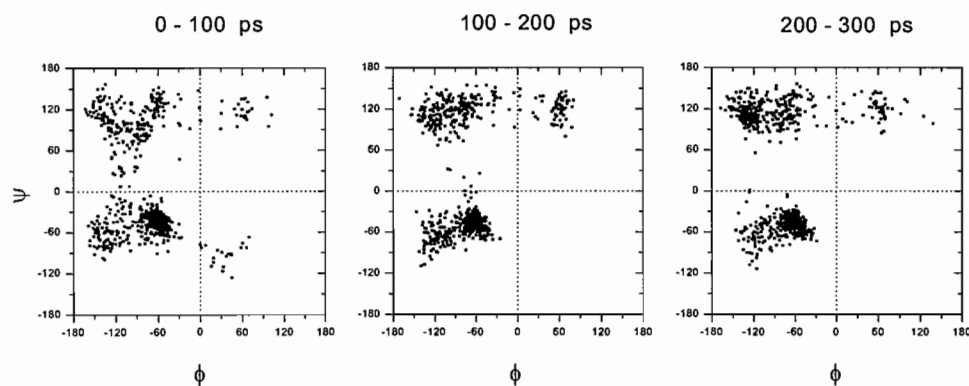


Fig. 6-15 : Diagramas de Ramachandran obtidos em diferentes estágios da dinâmica do sistema composto por oito alaninas em solvente explícito.

6.3 Triptofano

O triptofano é uma importante sonda fluorescente, amplamente utilizada para monitorar propriedades estruturais e dinâmicas de proteínas e peptídios contendo este aminoácido [75-78]. Um entendimento detalhado de seu comportamento fluorescente é, portanto, requerido para uma interpretação adequada dos resultados experimentais [79]. Decaimentos multiexponenciais caracterizam os espectros de fluorescência do triptofano, em diferentes soluções e pH's, estando ele na forma de monômeros ou presente na estrutura de proteínas e peptídios [78,80-84]. Para o triptofano *zwitteriônico* em água, em pH 7, foi observado que uma função do tipo biexponencial ajusta-se adequadamente à curva de decaimento fluorescente [80]. Com base nesse comportamento, foi criado um modelo propondo a presença de múltiplas conformações do triptofano no estado fundamental, que não se interconvertem durante o tempo de vida da fluorescência [80]. São sugeridas seis conformações, obtidas por rotações dos ângulos diedros χ_1 e χ_2 em torno das ligações N - C $_{\alpha}$ - C $_{\beta}$ - C $_{\gamma}$ e C $_{\alpha}$ - C $_{\beta}$ - C $_{\gamma}$ - C $_{\delta 2}$, respectivamente (fig. 6-16). Para χ_1 são sugeridos valores próximos a 60°, 180° e 300°, definindo as conformações *gauche*⁺ (*g*⁺), *trans* (*t*) e *gauche*⁻ (*g*⁻), e para χ_2 valores próximos a 90° e 270°, definindo as conformações *perpendicular* (*perp* χ_2) e *antiperpendicular* (*anti* χ_2).

Embora a heterogeneidade conformacional do triptofano seja confirmada por análise de resultados experimentais obtidos por técnicas de ressonância magnética nuclear [78,85] e espalhamento de raios-X [86], uma correlação direta entre as seis conformações propostas e o seu espectro de fluorescência resolvido no tempo não é ainda clara.

Estudos de dinâmica molecular procuram por estados conformacionais estáveis, analisando interconversões dos ângulos χ_1 e χ_2 e tentando encontrar alguma relação com as características fluorescentes do triptofano, porém discrepâncias têm sido encontradas [79,87]. Resultados obtidos em simulações têm mostrado forte dependência com o modelo utilizado para o solvente, representado explicitamente ou como um contínuo, bem como com o modelo atômico adotado, com os átomos de hidrogênio apolar representados um a um ou unidos aos átomos de carbono (ver apêndice A) [87]. A frequência de interconversões de χ_1 e χ_2 é maior para um ou outro ângulo, dependendo do modelo adotado. Resultados de dinâmica do triptofano, realizada com o programa THOR, em solvente representado como um contínuo [4], mostram mudanças conformacionais por interconversões somente do ângulo χ_2 , não sendo observadas interconversões do ângulo χ_1 .

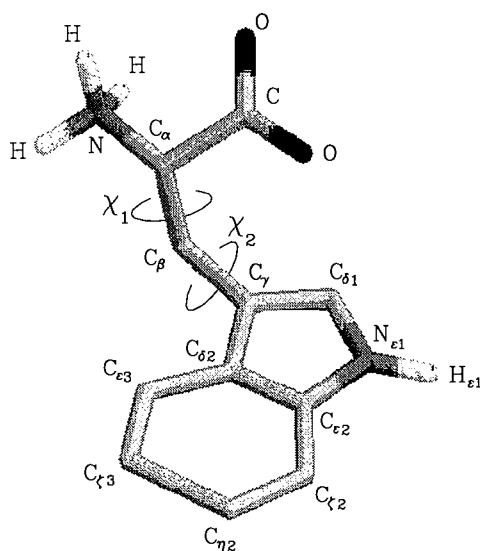


Fig. 6-16 : Aminoácido triptofano, com representação dos ângulos χ_1 e χ_2 definidos em torno das ligações N - C_α - C_β - C_γ e C_α - C_β - C_γ - C_{δ2}, respectivamente.

6.3.1 Simulação

Seis modelos para o triptofano *zwitteriônico* foram gerados a partir das seis conformações sugeridas na literatura [79,80], diferenciadas por valores dos ângulos diedros χ_1 e χ_2 (tabela 6-1).

conformação inicial	χ_1 (grau)	χ_2 (grau)
(a) anti χ_2 g ⁺	60	270
(b) anti χ_2 t	180	270
(c) anti χ_2 g ⁻	300	270
(d) perp χ_2 g ⁺	60	90
(e) perp χ_2 t	180	90
(f) perp χ_2 g ⁻	300	90

Tabela 6-1: Conformações iniciais do triptofano modelado, definidas por valores de χ_1 e χ_2 , sugeridos na literatura [79,80].

Cada triptofano foi colocado em uma caixa cúbica de lado 15,54 Å, conforme procedimento descrito no capítulo 5, com 110 moléculas de água do modelo SPC (fig. 6-17).

A geometria de cada sistema foi otimizada pela técnica de minimização de energia, com o emprego consecutivo dos métodos *steepest descent* e dos gradientes conjugados. Foram definidos valores nulos para as velocidades atômicas iniciais e a dinâmica molecular foi iniciada com um aquecimento do sistema até 300 K nos primeiros 10 ps.

Alcançada esta temperatura, sua manutenção foi obtida com um reescalamento direto das velocidades (eq. 3-12 e 3-14) até 100 ps. A partir daí, o sistema foi acoplado fracamente a um banho térmico, com tempo de relaxação característico $\tau = 0,1$ ps (eq. 3-12 e 3-15), e a dinâmica prosseguiu até que um tempo total de 4 ns fosse completado. O método *shake* foi empregado e as distâncias $O_w - H_w$ e $H_w - H_w$ das moléculas de água e N - H do triptofano foram mantidas constantes com a imposição de vínculos apropriados. O modelo de átomos unidos (ver apêndice A) foi utilizado para os átomos de hidrogênio ligados aos átomos de carbono alifáticos e aromáticos. As equações de movimento (3-6) e (3-8) foram integradas com um passo de tempo de 1 fs. Um raio de corte de 7,77 Å foi definido para os cálculos das interações entre átomos não-ligados. Variações da temperatura e energias cinética, potencial e total foram registradas a cada 0,1 ps. As conformações dos sistemas e os ângulos χ_1 e χ_2 foram monitorados durante a dinâmica e registrados a cada 0,1 ps.

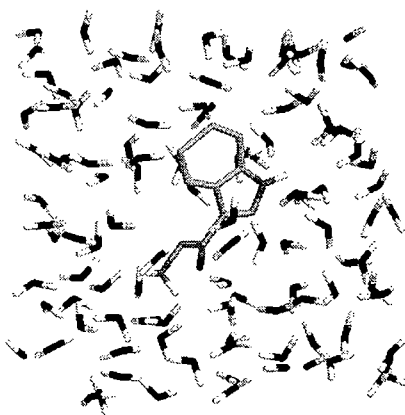


Fig. 6-17 : Aminoácido triptofano em água.

6.3.2 Resultados

O comportamento energético dos seis sistemas simulados foi acompanhado durante a dinâmica. A figura 6-18 mostra o resultado obtido para o sistema $perp\chi_2 g^-$. As fases de aquecimento, reescalonamento direto das velocidades atômicas e acoplamento fraco do sistema a um banho térmico podem ser identificadas visualmente na figura apresentada.

A estruturação das moléculas de água ao redor do triptofano foi avaliada por meio das funções de distribuição radial de pares, calculadas a partir de conformações registradas entre 0,3 e 2 ns de dinâmica. Para todos os sistemas, as funções analisadas apresentaram um comportamento bastante similar. Como será visto a seguir, isto é devido ao fato que, durante a maior parte do tempo da dinâmica, os seis sistemas modelados assumem as mesmas conformações. Deste modo, funções médias foram calculadas a partir das distribuições obtidas para cada sistema e são apresentadas na figura 6-19. A função para os pares C - O_w, com C representando os átomos de carbono dos anéis aromáticos, mostra uma distribuição larga, com pico em 3,78 Å, caracterizando o comportamento hidrofóbico da região, enquanto uma forte estruturação da água é observada ao redor dos átomos de hidrogênio H_T do terminal amina e oxigênio O_T do terminal carboxila, com picos estreitos e bem definidos em 1,97 Å e 1,86 Å, respectivamente, caracterizando interações por pontes de hidrogênio entre o solvente e os terminais.

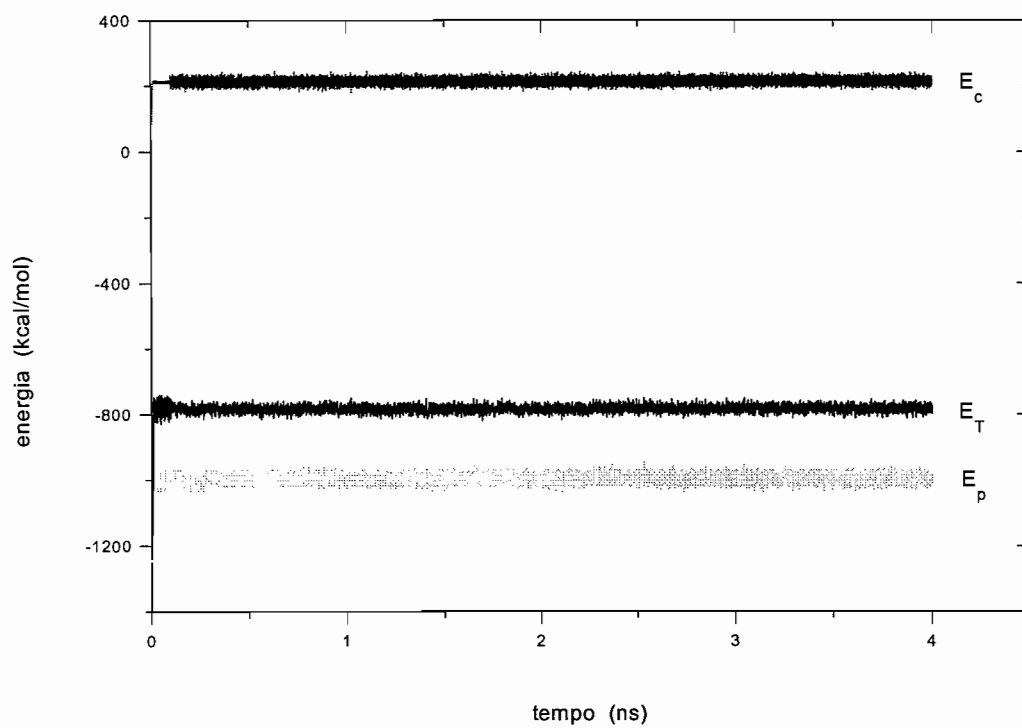


Fig. 6-18 : Acompanhamento temporal das energias cinética (E_c), potencial (E_p) e total (E_T) durante a dinâmica do sistema (f), definido na tabela 6-1.

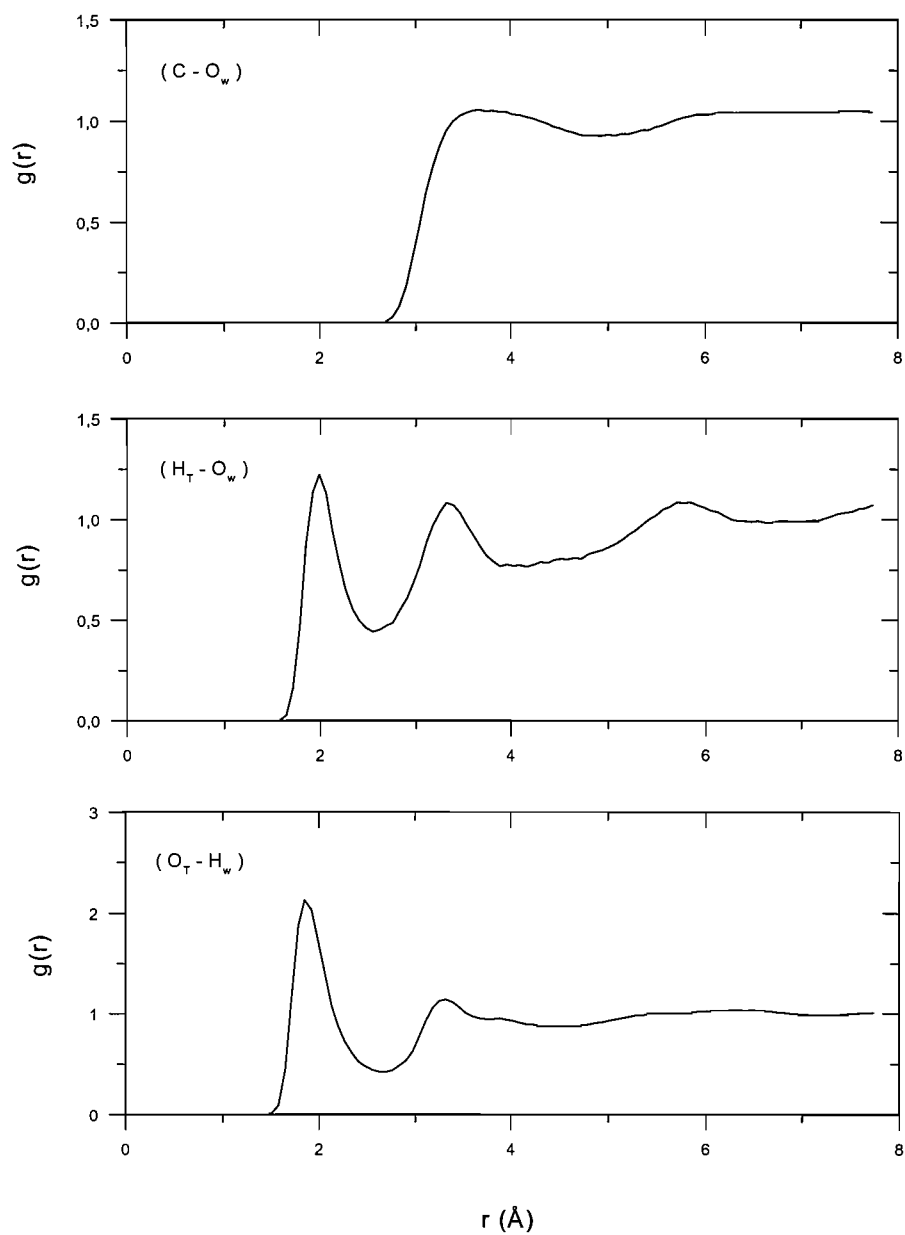


Fig. 6-19 : Funções de distribuição radial de pares para moléculas de água (átomos O_w e H_w) ao redor dos átomos de carbono dos anéis aromáticos (C), hidrogênio do terminal amina (H_T) e oxigênio do terminal hidroxila (O_T).

As figuras 6-20 e 6-21 apresentam a evolução temporal dos ângulos χ_1 e χ_2 dos seis sistemas modelados, monitorada durante 4 ns de dinâmica. A figura 6-20 mostra que, embora os sistemas (a) e (d) tenham sido iniciados na conformação g^+ ($\chi_1 = 60^\circ$), o sistema (d) abandona-a definitivamente após poucas dezenas de picossegundos, enquanto o sistema (a) se mantém nessa conformação por mais de 400 ps, voltando a assumi-la após 3 ns de dinâmica e aí permanecendo por cerca de uma centena de picossegundos. Os sistemas (c) e (f), iniciados na conformação g^- ($\chi_1 = 300^\circ$), também assumem a conformação g^+ , mas com pouco tempo de permanência. O sistema (c) a assume pouco antes de completar 2 ns e 4 ns de dinâmica e o sistema (f) pouco antes de 4 ns. Os sistemas (b) e (e), iniciados na conformação t ($\chi_1 = 180^\circ$), não assumem a conformação g^+ em momento algum. A figura 6-20 também mostra que, independentemente da conformação inicial adotada, durante a dinâmica todos os sistemas assumem, na maior parte do tempo, a conformação g^- , com rápidas transições, da ordem de dezenas de picossegundos, para a conformação t . Em relação ao ângulo χ_2 , a figura 6-21 mostra um maior equilíbrio entre as conformações $perp\chi_2$ ($\chi_2 = 90^\circ$) e $anti\chi_2$ ($\chi_2 = 270^\circ$) assumidas, com transições ocorrendo na escala de centenas de picossegundos.

A partir de um levantamento da frequência de visitação dos ângulos χ_1 e χ_2 durante os intervalos de tempo de dinâmica de 0 a 2 ns e de 0 a 4 ns, as distribuições angulares parciais e totais foram construídas para os seis sistemas simulados e estão apresentadas nas figuras 6-22 e 6-23. Pouca diferença é observada entre as distribuições parciais e totais obtidas para o ângulo χ_1 (fig. 6-22). A relação de proporcionalidade entre as três conformações definidas por esse ângulo fica, naturalmente, melhor delineada, quanto maior for o tempo de

dinâmica. Para o ângulo χ_2 (fig. 6-23), uma maior diferença é observada quando as distribuições parciais e totais são comparadas, mas, de um modo geral, elas indicam que todos sistemas caminham para um equilíbrio entre as duas conformações definidas por esse ângulo, com uma leve tendência para a conformação *perp* χ_2 .

As tabelas 6-2 e 6-3 apresentam os valores médios e desvios padrão das distribuições totais dos ângulos χ_1 e χ_2 , respectivamente, e a porcentagem do tempo em que os sistemas permaneceram em cada conformação, durante 4 ns de dinâmica. A tabela 6-2 mostra que em todas as simulações os sistemas assumiram a conformação g^- em mais do que 70% do tempo de dinâmica, com valores de χ_1 mantidos em torno de 294° , entre 10% e 20% do tempo os sistemas assumiram a conformação t , com valores de χ_1 mantidos em torno de 194° , e no tempo restante, menos que 10%, exceto para o sistema (a), os sistemas assumiram a conformação g^+ , com valores de χ_1 mantidos em torno de 55° . De acordo com a tabela 6-3, um maior equilíbrio foi observado entre as conformações definidas pelo ângulo χ_2 , mantido em torno de 112° , em cerca de 60% do tempo total de dinâmica, e em torno de 265° , no tempo restante.

Portanto, os resultados das simulações de dinâmica dos seis sistemas compostos por triptofano em água representada explicitamente, gerados a partir das conformações propostas para explicar o comportamento fluorescente do triptofano, com base na possível heterogeneidade do estado fundamental, indicam que em todos os casos o triptofano assume, durante 70% a 80% do tempo, apenas duas conformações. Deste modo, o decaimento biexponencial observado pode estar relacionado com as populações possuindo as conformações *perp* χ_2 g^- e *anti* χ_2 g^- .

Os resultados encontrados nas simulações encontram-se em acordo com os obtidos por experimentos de ressonância magnética nuclear, que indicam que o triptofano em solução assume uma conformação predominantemente do tipo *perp* χ_2 g^- [85], sendo que outras conformações contribuem com pelo menos 20% ou possivelmente 40% da população. Além disso, medidas feitas por espalhamento de raios-X de proteínas mostram que 57% do triptofano da amostra apresenta-se na conformação com χ_1 em torno de 290° , 30% em torno de 190° e 13% em torno de 61° [86].

Os resultados obtidos nas simulações são também coerentes com estudos teóricos sobre a desexcitação do triptofano por transferência de elétrons [88]. Estes estudos indicam que tal processo de desexcitação não radiativo é dominante para as conformações g^+ e t , enquanto para as conformações g^- a transferência de elétrons está competindo com a fluorescência, e portanto somente essas conformações podem estar relacionadas com o decaimento do estado excitado por emissão fluorescente.

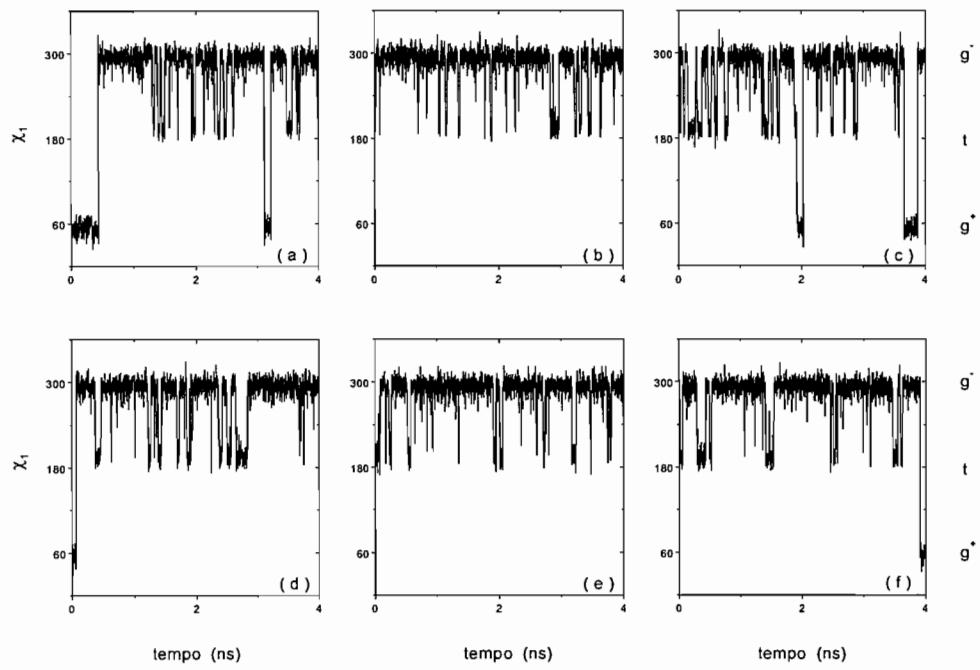


Fig. 6-20 : Evolução temporal do ângulo χ_1 dos seis sistemas apresentados na tabela 6-1, durante 4 ns de dinâmica.

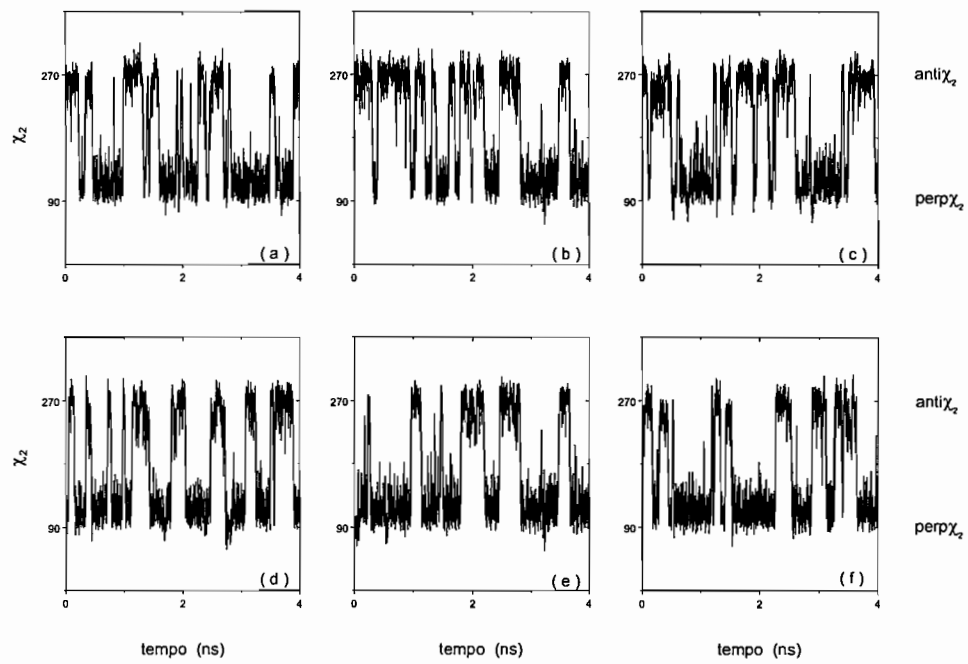


Fig. 6-21 : Evolução temporal do ângulo χ_2 dos seis sistemas apresentados na tabela 6-1, durante 4 ns de dinâmica.

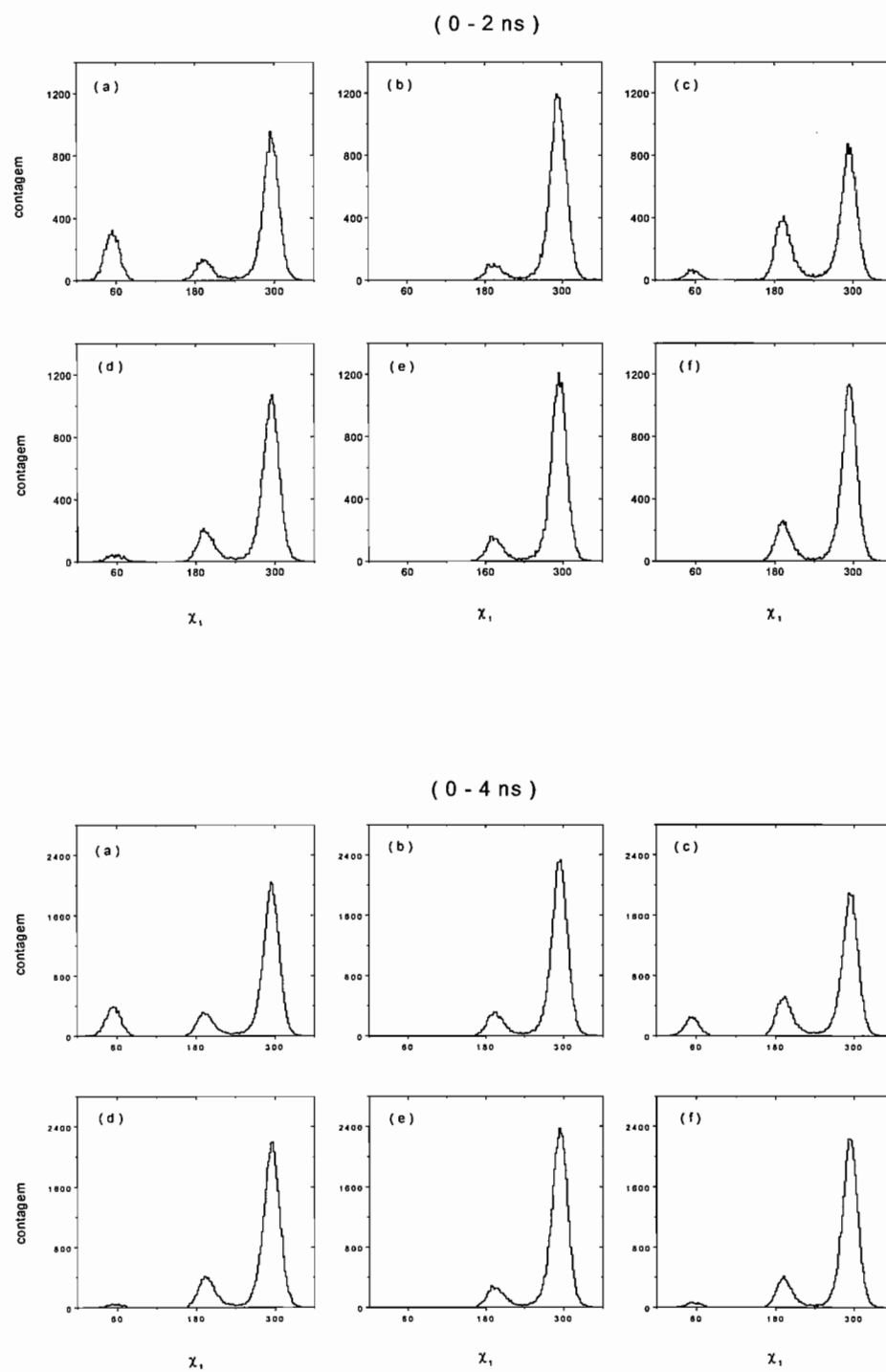


Fig. 6-22 : Distribuições do ângulo χ_1 dos seis sistemas modelados, obtidas durante os intervalos de tempo de dinâmica de 0 a 2 ns e 0 a 4 ns.

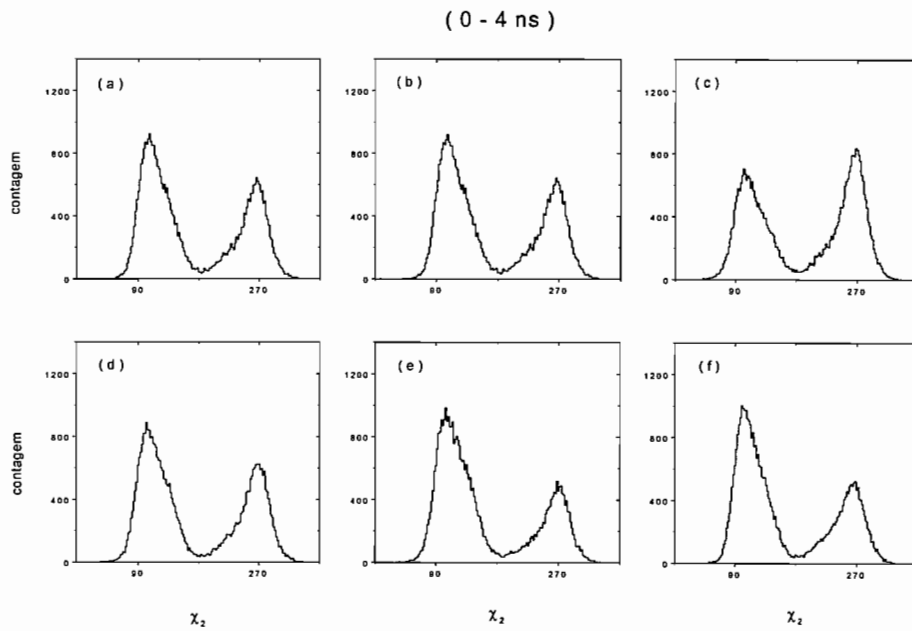
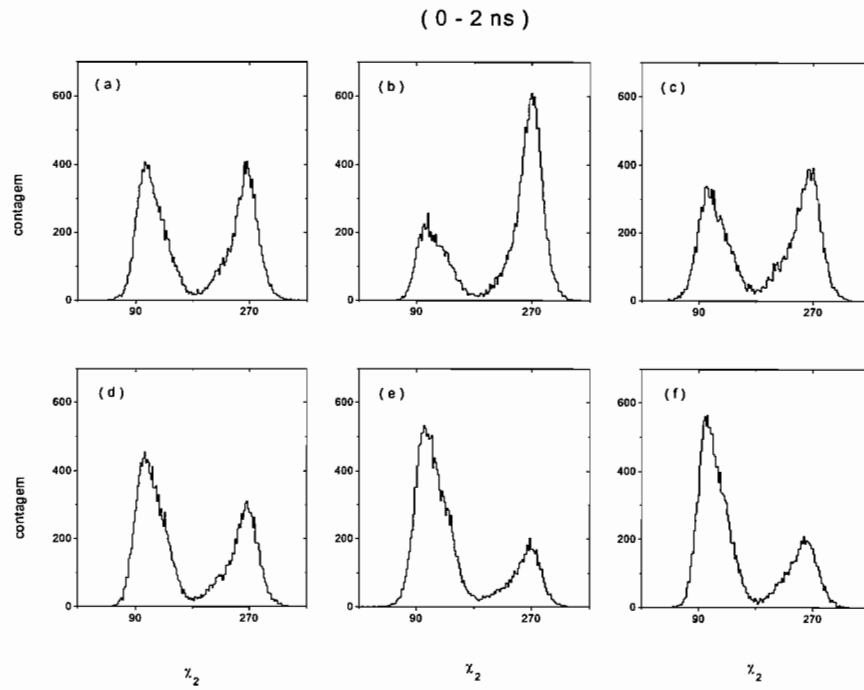


Fig. 6-23 : Distribuições do ângulo χ_2 dos seis sistemas modelados, obtidas durante os intervalos de tempo de dinâmica de 0 a 2 ns e 0 a 4 ns.

sistema	$\langle \chi_1 \rangle$	σ	% permanência
(a)	54,3	11,1	13,74
	193,9	11,9	11,99
	294,4	11,5	74,27
(b)	-----	-----	-----
	194,8	12,4	12,47
	294,0	11,7	87,53
(c)	53,5	10,3	8,08
	192,9	11,7	19,73
	294,0	11,7	72,12
(d)	56,9	12,7	1,67
	195,2	11,9	16,47
	294,3	11,8	81,86
(e)	-----	-----	-----
	195,0	13,3	11,56
	294,0	11,6	88,44
(f)	55,9	11,6	2,29
	192,6	11,6	15,25
	293,9	11,6	82,46

Tabela 6-2: Valores médios e desvios padrão (grau) das distribuições do ângulo χ_1 e porcentagem do tempo de permanência na conformação referida, obtidas para cada sistema, durante 4 ns de dinâmica.

sistema	$\langle \chi_2 \rangle$	σ	% permanência
(a)	112,0	22,1	59,67
	262,8	19,5	40,33
(b)	111,7	23,3	55,57
	267,4	17,9	44,43
(c)	111,1	23,4	47,32
	264,0	18,9	52,68
(d)	111,0	23,1	58,54
	265,6	19,6	41,46
(e)	113,1	24,2	68,29
	267,0	19,1	31,71
(f)	110,5	21,8	63,70
	260,4	22,1	36,30

Tabela 6-3: Valores médios e desvios padrão (grau) das distribuições do ângulo χ_2 e porcentagem do tempo de permanência na conformação referida, obtidas para cada sistema, durante 4 ns de dinâmica.

7 Conclusão

Apesar da existência de diversos programas acadêmicos e comerciais de modelagem molecular, a necessidade de ter em mãos ferramentas próprias, que permitissem uma maior compreensão dos cálculos e resultados envolvidos, além de total liberdade de intervenção em modelagens específicas, levou ao início do desenvolvimento do programa THOR, em 1992, e ao seu contínuo aperfeiçoamento no decorrer dos anos.

Esta tese consiste em uma série de implementações efetivadas no programa, ampliando sua capacidade de modelar biomoléculas em solução, com a possibilidade de representação explícita das moléculas do solvente. Resultados obtidos em simulações de alguns sistemas mostram a habilidade do programa em realizar esse tipo de modelagem.

Para a simulação do solvente composto por moléculas de água, foram incorporados ao banco de dados do THOR parâmetros de três modelos para a água: SPC, SPC/E e TIP3P, propostos na literatura. Os resultados obtidos indicam que os três modelos são equivalentes, quando simulados com o THOR, apresentando propriedades estruturais bastante similares entre si e em acordo com resultados experimentais. No entanto, por apresentar, de maneira sistemática, um comportamento intermediário em relação às posições e alturas dos picos das funções de distribuição radial de pares analisadas, o modelo SPC foi escolhido para ser adotado como solvente nas simulações de biomoléculas em água representada explicitamente.

Uma polialanina *zwitteriônica*, composta por vinte resíduos aminoácidos, formando uma estrutura inicial em hélice- α , foi simulada

em três diferentes meios: solvente aquoso explícito, meio contínuo com constante dielétrica $\epsilon = 80$, representando a água, e no vácuo, com $\epsilon = 1$. Os resultados mostram que a estrutura helicoidal só é mantida estável no primeiro caso, desfazendo-se nos dois últimos. A análise dos resultados aponta para a estruturação das moléculas de água ao redor da polialanina como a principal responsável pela sua manutenção, quando simulada em solvente explícito, com ambos os terminais envolvidos em interações por pontes de hidrogênio com moléculas de água vizinhas e os grupos metila imersos em uma região de exclusão hidrofóbica, protegidos de quaisquer interações com o solvente. Por outro lado, a estrutura em hélice- α de uma cadeia composta por oito alaninas não é mantida durante a simulação em água representada explicitamente, em acordo com as previsões para cadeias helicoidais pequenas. Deste modo, o programa THOR conseguiu reproduzir, para as duas cadeias modeladas, a relação esperada entre estabilidade helicoidal e tamanho da cadeia. Um estudo mais sistemático, no entanto, deverá ser efetuado, para uma verificação mais exata dessa relação.

Foram realizadas também simulações de seis sistemas compostos por um aminoácido triptofano em água representada explicitamente. As conformações iniciais foram construídas de acordo com um modelo que sugere a coexistência de múltiplas conformações do triptofano no estado fundamental para explicar o seu comportamento fluorescente. As conformações são geradas pela combinação de três valores para o ângulo diedro χ_1 , definindo as conformações *gauche*⁺ ($\chi_1 = 60^\circ$), *trans* ($\chi_1 = 180^\circ$) e *gauche*⁻ ($\chi_1 = 300^\circ$), e dois valores para o ângulo diedro χ_2 , definindo as conformações *perpendicular* ($\chi_2 = 90^\circ$) e *antiperpendicular* ($\chi_2 = 270^\circ$). Os resultados das simulações de dinâmica dos seis sistemas mostram que em todos os casos o triptofano assume, durante 70% a 80% do tempo,

apenas duas conformações: *gauche⁻ perpendicular* e *gauche⁻ antiperpendicular*. Deste modo, o decaimento biexponencial observado no espectro de fluorescência do triptofano pode estar relacionado com as populações possuindo essas duas conformações.

De um modo geral, os resultados obtidos nas simulações realizadas indicam que as implementações das técnicas para representação explícita do solvente composto por moléculas de água foram efetivadas com sucesso. Como continuidade deste trabalho, outros solventes e meios mais complexos, como interface água/membrana, deverão ser modelados com o emprego do programa THOR.

Apêndice A: Arquivos do programa THOR

Neste apêndice são apresentados alguns exemplos de arquivos de entrada e saída do programa THOR, seus formatos e as informações neles contidas.

Todos os arquivos do programa THOR ficam armazenados em três diretórios: *data*, *data-in* e *data-out*.

O diretório *data* contém o banco de dados do THOR, baseado no banco de dados do GROMOS, usado para a construção dos arquivos de entrada. O arquivo *thor-1.dat* contém os nomes e as massas dos átomos e o arquivo *thor-13.dat* contém as suas cargas em cada resíduo aminoácido. Esses dois arquivos são usados na construção do arquivo *molécula.IN*. Os arquivos *thor-2.dat*, *thor-3.dat*, *thor-4.dat* e *thor-8.dat* contêm informações sobre os parâmetros dos potenciais harmônicos, diédricos e torcionais do campo de força e são usados na construção do arquivo *molécula.TOP*. Os arquivos *thor-9.dat*, *thor-10.dat* e *thor-12.dat* contêm informações sobre os parâmetros do potencial de Lennard-Jones.

O diretório *data-in* contém os arquivos de entrada *molécula.IN* e *molécula.TOP* e o arquivo de saída *molécula.NEW*, utilizado como arquivo de entrada em simulações subseqüentes. Este diretório contém ainda o arquivo *DINAMICA.DAT*, com os parâmetros necessários para os cálculos de dinâmica molecular.

O diretório *data-out* contém os arquivos de saída *molécula.OUT*, *molécula.DHA*, *molécula.DBK*, *molécula.PDB*, *molécula.DMM* e *molécula.XYZ*, com os resultados obtidos nas simulações.

A.1 Arquivos IN e NEW

A primeira informação que um modelo molecular deve conter é sobre os tipos atômicos do sistema. No THOR, átomos iguais são designados por diferentes nomenclaturas, de acordo com sua localização ou função nos diferentes resíduos. Por exemplo, átomos de oxigênio podem ser designados por O (oxigênio da carbonila C = O), OM (oxigênio da carboxila COO⁻), OA (oxigênio da hidroxila OH) ou OW (oxigênio pertencente a uma molécula de água), os átomos de hidrogênio podem ser designados por H (hidrogênio ligado ao nitrogênio), HO (hidrogênio da hidroxila), HW (hidrogênio pertencente a uma molécula de água) ou HS (hidrogênio ligado ao enxofre), etc.

Além disso, um procedimento comumente utilizado e que leva a uma grande redução no número total de átomos do sistema é o emprego do modelo de átomos unidos. No THOR, os átomos das ligações C - H são considerados como átomos únicos CH1, CH2, CH3, etc.

O arquivo com extensão IN é um arquivo de entrada do programa THOR que contém informações sobre o tipo de modelagem que será feito, os termos do campo de força que serão considerados, os tipos atômicos do sistema e suas posições, as especificações das ligações covalentes de todo o sistema, as cargas e os resíduos aminoácidos dos quais cada átomo faz parte. Este arquivo pode ser gerado manualmente ou com a utilização de programas específicos, como aqueles descritos no capítulo 5 para sistemas constituídos por biomoléculas em água explicitamente representada. As duas primeiras linhas do arquivo são reservadas para comentários. As duas linhas seguintes contêm o “menu” de opções para os cálculos e informações que devem ser guardadas no decorrer da simulação. OTMZ ou OPT3 indicam que cálculos de minimização de

energia serão realizados usando o método *steepest descent* ou dos gradientes conjugados, respectivamente, e MD indica que cálculos de dinâmica serão efetuados. As siglas HOOK, FTHET, FDHI e FDHP indicam que serão incluídos no cálculo do potencial, respectivamente, o potencial harmônico linear, o potencial harmônico angular, o potencial diédrico impróprio e o potencial torcional próprio, enquanto as siglas NBBC e NBCVW determinam a inclusão do cálculo das interações entre átomos não-ligados, com ou sem utilização de condições de contorno periódicas. SHAKE indica que o método *shake* será empregado para a manutenção de vínculos de distância, BEREN=0,1 indica que o sistema será acoplado a um banho térmico com tempo característico $\tau = 0,1$ ps (eq. 3-15), SLIDE indica que as posições dos átomos serão gravadas no arquivo XYZ e WRTDBK as posições e velocidades dos átomos serão gravadas no arquivo DBK. As duas linhas seguintes contêm os parâmetros utilizados na minimização da energia: o valor do passo inicial usado no deslocamento das coordenadas atômicas (Å) e o valor da precisão do potencial (kcal/mol), que determina o término da otimização. As seis linhas consecutivas contêm as seguintes informações: constante dielétrica, dimensão da caixa usada nas condições de contorno periódicas (Å), raio de corte (Å) e raio para a construção da lista de vizinhos (Å).

Exemplo:

Comentário

Comentário

MD HOOK FTHET FDHP FDHI NBBC SHAKE BEREN=0,1 WRTDBK

0.1000000	Passo inicial usado na minimização
0.0001000	Precisão no cálculo do potencial
1.0000000	Constante dielétrica ϵ
18.64954548	Lado X da caixa para condições de contorno periódicas
18.64954548	Lado Y da caixa para condições de contorno periódicas
18.64954548	Lado Z da caixa para condições de contorno periódicas
9.0000000	Raio de corte
10.0000000	Raio para lista de vizinhos

As linhas seguintes contêm informações sobre o sistema modelado: identificação numérica do átomo no sistema, tipo atômico na nomenclatura do GROMOS, posição cartesiana (x,y,z) dos átomos (Å), especificações das ligações covalentes, cargas atômicas (unidade de carga eletrônica), tipo atômico na nomenclatura PDB (*Brookhaven Protein Data Bank*), resíduo aminoácido do qual o átomo faz parte e o número do resíduo no sistema.

Exemplo:

1	H	- 0.9950	1.3920	1.0200	3	0	0	0	0	0	0.248	H	TRP	1
2	H	- 0.9950	1.3920	2.6530	3	0	0	0	0	0	0.248	H	TRP	1
3	NL	- 0.4180	1.3920	1.8370	1	2	4	5	0	0	0.129	N	TRP	1
4	H	0.1590	2.2080	1.8360	3	0	0	0	0	0	0.248	H	TRP	1
5	CH1	0.4310	0.1920	1.8360	3	6	17	0	0	0	0.127	CA	TRP	1
6	CH2	1.3140	0.1920	0.5860	5	7	0	0	0	0	0.000	CB	TRP	1

As posições atômicas também podem ser dadas em coordenadas internas (r_{ij} , θ_{ijk} , ϕ_{ijkl}), em vez de coordenadas cartesianas. Esta facilidade é bastante útil nos casos em que conformações iniciais específicas devem ser consideradas. Neste caso, o “menu” deve conter a opção INTERNAL, e as linhas com informações sobre o sistema devem conter o tipo atômico, as coordenadas internas (Å e grau) e os átomos (i,j,k,l) envolvidos nas especificações das coordenadas internas, as cargas atômicas (unidade de carga eletrônica) e as ligações covalentes.

Exemplo:

H	0.000	0.00	0.00	0	0	0	0	0.248	3	0	0	0	0	0
H	1.633	0.00	0.00	1	0	0	0	0.248	3	0	0	0	0	0
NL	1.000	35.25	0.00	2	1	0	0	0.129	1	2	4	5	0	0
H	1.000	109.50	120.00	3	2	1	0	0.248	3	0	0	0	0	0
CH1	1.470	109.50	240.00	3	2	1	0	0.127	3	6	17	0	0	0
CH2	1.530	109.50	180.00	5	3	2	0	0.000	5	7	0	0	0	0

Nos cálculos de dinâmica, as velocidades atômicas iniciais podem ser geradas a partir de uma distribuição de Maxwell-Boltzmann (eq. 3-10) ou podem ser definidas no arquivo IN. Neste caso, as linhas seguintes desse arquivo devem conter a identificação numérica do átomo no sistema, o tipo atômico e as componentes da velocidade (v_x, v_y, v_z) (\AA ps^{-1}).

Exemplo:

1	H	0.00000	0.00000	0.00000
2	H	0.00000	0.00000	0.00000
3	NL	0.00000	0.00000	0.00000
4	H	0.00000	0.00000	0.00000
5	CH1	0.00000	0.00000	0.00000
6	CH2	0.00000	0.00000	0.00000

Quando uma simulação de minimização de energia ou dinâmica molecular é finalizada, um arquivo com extensão NEW é gerado no mesmo formato que o arquivo IN, contendo as informações das novas posições e velocidades alcançadas. Posteriormente a extensão desse arquivo deverá ser modificada para IN e ele poderá ser usado para dar prosseguimento à simulação.

A.2 Arquivo TOP

O arquivo com extensão TOP contém os parâmetros que definem o campo de força: K_b , b_0 , K_θ , θ_0 , K_ξ , ξ_0 , K_ϕ , p e δ (eq.1-7). Este arquivo é construído automaticamente pelo programa THOR quando um sistema é modelado.

As informações do potencial harmônico linear são apresentadas no formato: numeração dos pares de átomos ligados covalentemente,

identificação do par por meio de número e tipo atômico, constante K_b ($\text{kcal mol}^{-1} \text{Å}^{-2}$) e distância de equilíbrio b_0 (Å) (eq.1-1).

Exemplo:

Parâmetros para o potencial harmônico linear (i - j)

1) --	1	3	H - NL	895.0000	1.0000
2) --	2	3	H - NL	895.0000	1.0000
3) --	3	4	NL - H	895.0000	1.0000
4) --	3	5	NL - CH1	900.0000	1.4700
5) --	5	6	CH1 - CH2	800.0000	1.5300

As informações do potencial harmônico angular são apresentadas no formato: numeração do conjunto de três átomos seqüencialmente ligados, identificação numérica e por tipo atômico dos átomos envolvidos, constante K_θ ($\text{kcal mol}^{-1} \text{rad}^{-2}$) e ângulo de equilíbrio θ_0 (grau) (eq.1-2).

Exemplo:

Parâmetros para o potencial harmônico angular (i - j- k)

1) --	1	3	2	H - NL - H	80.0000	109.5000
2) --	1	3	4	H - NL - H	80.0000	109.5000
3) --	1	3	5	H - NL - CH1	90.0000	109.5000
4) --	2	3	4	H - NL - H	80.0000	109.5000
5) --	2	3	5	H - NL - CH1	90.0000	109.5000

As informações do potencial diédrico impróprio são apresentadas no formato: numeração dos diedros impróprios, identificação numérica e por tipo atômico dos quatro átomos envolvidos, constante K_ξ ($\text{kcal mol}^{-1} \text{rad}^{-2}$) e ângulo ξ_0 (grau) (eq.1-3).

Exemplo:

Parâmetros para o potencial diédrico impróprio (i - j - k - l)

1) --	3	2	5	1	NL - X - Y - H	80.0000	35.26439
2) --	3	4	5	1	NL - X - Y - H	80.0000	35.26439
3) --	17	5	19	18	C - X - Y - OM	40.0000	00.00000
4) --	5	3	17	6	CH1 - X - Y - CH2	80.0000	35.26439

As informações do potencial torcional próprio são apresentadas no formato: numeração dos diedros próprios, identificação numérica e por tipo atômico dos quatro átomos envolvidos, constante K_{ϕ} (kcal mol^{-1}), fase δ (grau) e periodicidade p (eq.1-4).

Exemplo:

Parâmetros para o potencial torcional próprio (i - j - k - l)

1) --	1	3	5	17	X - NL - CH1 - Y	0.9000	0.0000	3 *
2) --	3	5	17	19	X - CH1 - C - Y	0.1000	0.0000	6 *
3) --	3	5	6	7	X - CH1 - CH2 - Y	1.4000	0.0000	3 *

O asterisco na última coluna indica que aquela ligação será acompanhada durante a dinâmica e o valor do ângulo será gravado no arquivo DHA em intervalos de tempo determinados no arquivo DINAMICA.DAT.

As informações usadas na rotina *shake* para vínculos de distância são apresentadas no formato: numeração do vínculo, número total dos átomos envolvidos no vínculo, identificação dos átomos envolvidos e quadrado das distâncias a serem mantidas.

Exemplo:

Tabela de vínculos para a rotina *shake*

1) --	3	1	2	3	1	1.000000000	1.000000000	2,666626495
2) --	3	4	5	6	4	1.000000000	1.000000000	2,666626495
3) --	3	7	8	9	7	1.000000000	1.000000000	2,666626495
4) --	3	10	11	12	10	1.000000000	1.000000000	2,666626495
5) --	3	13	14	15	13	1.000000000	1.000000000	2,666626495

Os parâmetros A_{ij} e B_{ij} referentes ao potencial de Lennard-Jones não são apresentados no arquivo TOP, uma vez que o número de parâmetros, da ordem do quadrado do número de átomos do sistema,

poderia gerar arquivos muito grandes. Durante o cálculo, esses parâmetros são lidos diretamente do banco de dados *thor-9.dat*, *thor-10.dat* e *thor-12.dat*.

Em um sistema molecular, os átomos ligados covalentemente são designados primeiros vizinhos. Por exemplo, na figura A-1 os átomos C e E são os primeiros vizinhos do átomo D. Do mesmo modo, os primeiros vizinhos dos átomos C e E (átomos B e F, respectivamente) são designados segundos vizinhos do átomo D, e os primeiros vizinhos dos átomos B e F (átomos A e G, respectivamente) são os segundos vizinhos dos átomos C e E, respectivamente, e terceiros vizinhos do átomo D.

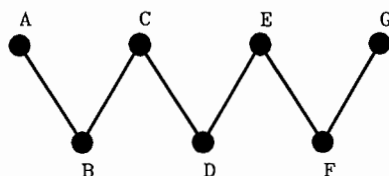


Fig. A-1 : Esquema utilizado para a explicação do conceito de primeiros, segundos e terceiros vizinhos (ver texto acima).

As interações entre os primeiros e segundos vizinhos são modeladas pelos três termos harmônicos do campo de força e devem, portanto, ser excluídas do cálculo das interações entre átomos não-ligados [1,7,35].

Em relação às interações entre terceiros vizinhos, os parâmetros do termo de repulsão do potencial de Lennard-Jones assumem, em alguns

casos, valores muito elevados, que podem estressar algumas ligações, provocando deformações especialmente no casos de interações carbono-carbono, levando o sistema a conformações irreais [1].

Para evitar este problema, um conjunto de parâmetros menos repulsivos é utilizado nos casos em que os átomos de carbono são terceiros vizinhos de outros átomos do sistema.

O arquivo TOP contém informações sobre os primeiros, segundos e terceiros vizinhos de cada átomo do sistema no formato: número do átomo, número de vizinhos, identificação numérica dos vizinhos.

Exemplo:

Primeiros, segundos e terceiros vizinhos

1	6	2	3	4	5	6	17	
2	5	3	4	5	6	17		
3	7	4	5	6	7	17	18	19
4	3	5	6	17				
5	7	6	7	8	9	17	18	19

Terceiros vizinhos

1	2	6	17	
2	2	6	17	
3	3	7	18	19
4	2	6	17	
5	2	8	9	

O arquivo TOP contém ainda informação sobre as massas atômicas no formato: número do átomo no sistema, número do átomo no banco de dados, tipo e massa atômicos.

Exemplo:

Massas atômicas

1	18	H	1.00800
2	18	H	1.00800
3	7	NL	14.00670
4	18	H	1.00800
5	12	CH1	13.01900

A.3 Arquivo DINAMICA.DAT

O arquivo DINAMICA.DAT, localizado no diretório *data-in*, contém informações utilizadas nos cálculos de dinâmica molecular.

As duas primeiras linhas do arquivo podem ser usadas para comentários. Nas demais linhas seguem as informações: temperatura inicial do sistema (K), taxa de aquecimento por picossegundo ($K \text{ ps}^{-1}$), desvio permitido da temperatura em relação à temperatura de referência (K), que define a necessidade ou não de reescalamento das velocidades, temperatura final de referência (K), valor do passo de integração (ps), número total de passos da dinâmica, número de passos usado para o cálculo da média da energia cinética e potencial, frequência em passos para a gravação dos arquivos XYZ, DBK, OUT e DHA e parâmetro que indica se as velocidades atômicas iniciais são dadas no arquivo de entrada IN ou se elas serão geradas a partir de uma distribuição de Maxwell-Boltzmann.

Exemplo:

Parâmetros referentes ao cálculo da dinâmica molecular

1.0000	Temperatura inicial
30.0000	Taxa de aquecimento por picossegundo
3.0000	Reescala as velocidades quando $T - T_0 >$ que este valor
300.0000	Temperatura final de referência
0.0010	Passo de integração
20000	Número total de passos para o cálculo da dinâmica
1	Número de passos para a média da energia cinética
200	Frequência (em passos) para escrever XYZ e DBK
200	Frequência (em passos) para escrever OUT
200	Frequência (em passos) para escrever DHA
1	0 sorteia velocidades, 1 lê velocidades no arquivo IN

A.4 Arquivo OUT

O arquivo com extensão OUT contém os resultados dos cálculos de minimização de energia e dinâmica molecular. No primeiro caso, o arquivo apresenta, em seqüência, o potencial inicial total (kcal/mol), o número de ciclos necessários para a otimização, o potencial final total (kcal/mol) e as coordenadas cartesianas finais de cada átomo (Å). Em cálculos de dinâmica molecular, o arquivo apresenta valores da temperatura (K), energia cinética (kcal/mol), energia potencial (kcal/mol) e energia total (kcal/mol) em intervalos de tempo especificados no arquivo DINAMICA.DAT.

Exemplo:

Tempo	Temperatura	En. Cinética	En. Potencial	En. Total
0.1000	7,75679	5,50288	- 1200,77044	- 1195,26755
0.2000	10,05152	7,13083	- 1222,00539	- 1214,87456
0.3000	12,34465	8,75763	- 1240,91420	- 1232,15657
0.4000	15,40547	10,92906	- 1249,94452	- 1239,01546
0.5000	13,42521	9,52421	- 1248,74991	- 1239,22570

A.5 Arquivo DHA

O arquivo com extensão DHA guarda informação dos ângulos torcionais próprios, especificados no arquivo TOP, durante a dinâmica molecular.

Na primeira coluna fica registrado o tempo da dinâmica (ps) e em cada coluna subsequente fica registrado o ângulo de uma ligação (grau).

Exemplo:

0.1000	171.72	294.31	306.56	170.58
0.2000	173.19	295.10	305.53	169.02
0.3000	172.59	295.92	307.90	167.40
0.4000	170.44	295.99	312.57	169.26
0.5000	169.30	294.89	317.24	173.57

A.6 Arquivo DBK

O arquivo com extensão DBK guarda informações das posições e velocidades atômicas durante a dinâmica molecular.

Na primeira linha, o arquivo contém o número do passo e o tempo de integração correspondente e, nas linhas seguintes, o número do átomo, as coordenadas cartesianas (x,y,z) (Å) e as componentes da velocidade (v_x, v_y, v_z) (Å ps⁻¹). O formato é repetido com a gravação consecutiva dessas informações em intervalos de tempo especificados no arquivo DINAMICA.DAT.

Exemplo:

PASSO = 0		TEMPO = 0.0000 ps				
1	0.54404	0.25227	-1.07890	-12.86776	10.27376	5.58859
2	2.08911	0.47444	-1.82734	-17.41174	-10.85154	21.91892
3	1.53634	0.44335	-0.98478	1.39805	-5.25211	-7.21711
4	1.65911	-0.57407	-1.14865	14.97235	-11.87848	14.05578
5	2.17299	0.44031	0.33366	-5.17509	-3.39699	4.56353
...						
PASSO = 100		TEMPO = 0.1000 ps				
1	0.73307	0.27358	-1.36906	8.35133	-11.91490	1.63196
2	2.35620	0.79512	-1.72638	-2.31704	-30.59404	-6.10033
3	1.73488	0.43890	-1.06095	-2.63692	0.35045	-0.32706
4	2.03609	-0.48918	-1.39998	4.62534	-4.66542	-9.64228
5	2.22581	0.26352	0.31664	-0.13348	1.39527	6.43125
...						

A.7 Arquivos PDB, DMM e XYZ

Os arquivos com extensão PDB e DMM contêm a conformação final do sistema após o cálculo de minimização de energia ou dinâmica molecular. O arquivo PDB segue o formato do banco de dados *Brookhaven Protein Data Bank* e pode ser lido por programas como o RASMOL para visualização molecular [74], enquanto o arquivo DMM pode ser lido pelo programa DTMM (*Desktop Molecular Modeller*) para visualização em microcomputadores padrão IBM-PC [89]. O arquivo com extensão XYZ contém informações conformacionais guardadas durante a dinâmica e seu formato pode ser lido pelo programa XMOL [90], que permite o acompanhamento visual da dinâmica molecular.

O formato PDB contém o número do átomo, a nomenclatura PDB, o resíduo do qual o átomo faz parte, o número do resíduo no sistema, e as coordenadas cartesianas (x,y,z) (Å).

Exemplo:

ATOM	1	H	TRP	1	- 0.995	1.392	1.020
ATOM	2	H	TRP	1	- 0.995	1.392	2.653
ATOM	3	N	TRP	1	- 0.418	1.392	1.837
ATOM	4	H	TRP	1	0.159	2.208	1.836
ATOM	5	CA	TRP	1	0.431	0.192	1.836
ATOM	6	Cb	TRP	1	1.314	0.192	0.586

O arquivo DMM contém, nas três primeiras linhas informações usadas para o funcionamento do programa e o número total de átomos do sistema. As linhas subsequentes contêm o número do átomo, a nomenclatura DMM, as coordenadas cartesianas (x,y,z) (Å) e as especificações das ligações covalentes.

Exemplo:

		1.000			1.000			1.000		
		90.000	90.000	90.000						
349										
1	H	- 0.99500	1.39200	1.02000	3	0	0	0	0	0
2	H	- 0.99500	1.39200	2.65300	3	0	0	0	0	0
3	N	- 0.41800	1.39200	1.83700	1	2	4	5	0	0
4	H	0.15900	2.20800	1.83600	3	0	0	0	0	0
5	C	0.43100	0.19200	1.83600	3	6	17	0	0	0
6	C	1.31400	0.19200	0.58600	5	7	0	0	0	0

O arquivo XYZ contém, na primeira linha, o número total de átomos do sistema e, nas linhas seguintes, o tipo atômico e as coordenadas cartesianas (x,y,z) (Å). Este bloco de informações é gravado repetidamente em intervalos de tempo especificados no arquivo DINAMICA.DAT.

Exemplo:

349				
H	0.54404	0.25227	-1.07890	
H	2.08911	0.47444	-1.82734	
N	1.53634	0.44335	-0.98478	
...				
349				
H	0.73307	0.27358	-1.36906	
H	2.35620	0.79512	-1.72638	
N	1.73488	0.43890	-1.06095	
...				
349				
H	0.76471	0.07260	-1.14875	
H	2.43519	0.17521	-1.63199	
N	1.72747	0.09695	-0.89281	
...				

Apêndice B: Computadores utilizados

O programa THOR, escrito em linguagem FORTRAN - 77, foi desenvolvido para operar em diferentes plataformas.

As simulações descritas neste trabalho foram realizadas no computador CRAY J916 do Laboratório de Computação Científica Avançada (LCCA/USP), na estação de trabalho HP-9000 série 700 e em um microcomputador pentium II - 266 MHz do Grupo de Biofísica.

A seguir é feito um resumo dos sistemas estudados, os computadores utilizados e o tempo de processamento consumido.

Na dinâmica dos sistemas formados somente por água, as equações de movimento (3-6) e (3-8) foram integradas com o emprego de um passo de 2 fs, enquanto nos outros casos foi utilizado um passo de 1 fs.

- Três modelos para a água - 648 átomos.

Para cada sistema foi realizada uma dinâmica de 1 ns, dividida em 10 etapas de 100 ps. No CRAY, o tempo médio de processamento de cada etapa foi de 7 horas.

- Cadeia de 20 alaninas em água - 2058 átomos.

Foi realizada uma dinâmica de 300 ps, dividida em 30 etapas de 10 ps. No CRAY, o tempo médio de processamento de cada etapa foi de 8 horas.

- Cadeia de 20 alaninas em solvente contínuo e no vácuo - 123 átomos.

Para cada sistema foi realizada uma dinâmica de 300 ps. Na estação da HP, o tempo médio de processamento foi de 6 horas.

- Cadeia de 8 alaninas em água - 924 átomos.

Foi realizada uma dinâmica de 300 ps, dividida em 3 etapas de 100 ps. No CRAY, o tempo médio de processamento de cada etapa foi de 26 horas.

- Seis conformações para o triptofano em água - 349 átomos

Para cada sistema, foi realizada uma dinâmica de 4 ns, dividida em 40 etapas de 100 ps. Cada etapa utilizou um tempo médio de processamento de 6 horas, quando realizada no CRAY, 15 horas na estação da HP e 3 horas no microcomputador pentium.

Referências

- [1] van Gunsteren, W.F., Berendsen, H.J.C., 1987, *GRONingen MOlecular Simulation (GROMOS) Library Manual*, Biomos, Groningen.
- [2] Mundim, K.C., Bisch, P.M., Pascutti, P.G., 1993, XVI Encontro Nacional de Física da Matéria Condensada, Caxambu, Programas e Resumos.
- [3] Israelachvili, J.N., 1991, *Intermolecular and Surface Forces*, 2nd ed., Academic Press, London.
- [4] Pascutti, P.G., 1996, tese de doutoramento, Instituto de Física, Universidade de São Paulo.
- [5] Arêas, E.P.G., Pascutti, P.G., Schreier, S., Mundim, K.C., Bisch, P.M., 1994, *Braz. J. Med. Biol. Res.*, **27**, 527.
- [6] Arêas, E.P.G., Pascutti, P.G., Schreier, S., Mundim, K.C., Bisch, P.M., 1995, *J. Phys. Chem.*, **99**, 14882.
- [7] Brooks III, C.L., Karplus, M., Pettitt, B.M., 1988, *Proteins: A theoretical perspective of dynamics, structure and thermodynamics - Advances in chemical physics - Vol. LXXI*, John Wiley & Sons, New York.
- [8] Moss, D.S., Flores, T.P., 1990, *Supercomputational Science*, ed. by Evans, R.G., Wilson, S., Plenum Press, New York.
- [9] Mackay, D.H.J., Cross, A.J., Hagler, A.T., 1990, *Prediction of protein structure and the principles of protein conformation*, ed. by Fasman, G.D., Plenum Press, New York.
- [10] de Vlieg, J., van Gunsteren, W.F., 1991, *Methods in Enzymology*, **202**, 268, Academic Press, New York.
- [11] van Gunsteren, W.F., Luque, F.J., Timms, D., Torda, A.E., 1994, *Annu. Rev. Biophys. Struct.*, **23**, 847.

- [12] Stryer, L., 1988, *Biochemistry*, 3rd ed., W.H. Freeman and Company, New York.
- [13] Voet, D., Voet, J.G., 1995, *Biochemistry*, 2nd ed., John Wiley & Sons, New York.
- [14] Richardson, J.S., Richardson, D.C., 1990, *Prediction of protein structure and the principles of protein conformation*, ed. by Fasman, G.D., Plenum Press, New York.
- [15] Lotan, N., Berger, A., Katchalski, E., 1972, *Annu. Rev. Biochem.*, **41**, 869.
- [16] van Gunsteren, W.F., Berendsen, H.J.C., 1990, *Angew. Chem. Int. Ed. Engl.*, **29**, 992.
- [17] IUPAC - IUB Commission on Biochemical Nomenclature, 1970, *Biochemistry*, **9**, 3471.
- [18] CRC Handbook of Chemistry and Physics, 1988, 1st Student ed., ed. by Weast, R.C., CRC Press, USA.
- [19] Brooks, B.R., Bruccoreli, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M., 1983, *J. Comput. Chem.*, **4**, 187.
- [20] Weiner, S.J., Kollman, P.A., Nguyen, D.T., Case, D.A., 1986, *J. Comput. Chem.*, **7**, 230.
- [21] Clark, M., Cramer III, R.D., van Opdenbosch, N., 1989, *J. Comput. Chem.*, **10**, 982.
- [22] Karplus, M., Petsko, G.A., 1990, *Nature*, **347**, 632.
- [23] Wiberg, K.B., 1965, *J. Am. Chem. Soc.*, **87**, 1070.
- [24] Williams, J.E., Stang, P.J., 1968, *Annu. Rev. Phys. Chem.*, **19**, 531.
- [25] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992, *Numerical Recipes in Fortran. The Art of Scientific Computing*, 2nd ed., Cambridge University Press.
- [26] Aqvist, J., van Gunsteren, W.F., Leijonmarck, M., Tapia, O., 1985, *J. Mol. Biol.*, **183**, 461.

- [27] Salinas, S.R.A., 1997, *Introdução à Física Estatística*, Edusp - Editora da Universidade de São Paulo.
- [28] Woodcock, L.V., 1971, *Chem. Phys. Lett.*, **10**, 257.
- [29] Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.R., 1984, *J. Chem. Phys.*, **81**, 3684.
- [30] Allen, M.P., Tildesley, D.J., 1989, *Computer Simulation of Liquids*, Clarendon Press, Oxford.
- [31] Brooks III, C.L., Pettitt, B.M., Karplus, M., 1985, *J. Chem. Phys.*, **83**, 5897.
- [32] Prevost, M., van Belle, D., Lippens, G., Wodak, S., 1990, *Mol. Phys.*, **71**, 587.
- [33] Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham III, T.E., DeBolt, S., Ferguson, D., Seibel, G., Kollman, P., 1995, *Comput. Phys. Commun.*, **91**, 1.
- [34] Berendsen, H.J.C., van der Spoel, D., van Drunen, R., 1995, *Comput. Phys. Commun.*, **91**, 43.
- [35] van der Spoel, D., van Buuren, A.R., Apol, E., Meulenhoff, P.J., Tieleman, D.P., Sijbers, A.L.T.M., van Drunen, R., Berendsen, H.J.C., 1996, *Gromacs User Manual*, University of Groningen.
- [36] Finchman, D., Ralston, B.J., 1981, *Comput. Phys. Commun.*, **23**, 127.
- [37] Ryckaert, J.-P., Ciccotti, G., Berendsen, H.J.C., 1977, *J. Comput. Phys.*, **23**, 327.
- [38] van Gunsteren, W.F., Berendsen, H.J.C., 1977, *Mol. Phys.*, **34**, 1311.
- [39] Ciccotti, G., Ferrario, M., Ryckaert, J.-P., 1982, *Mol. Phys.*, **47**, 1253.
- [40] Ciccotti, G., Ryckaert, J.-P., 1986, *Comput. Phys. Reports*, **4**, 348.
- [41] Goldstein, H., 1981, *Classical Mechanics*, 2nd ed., Addison-Wesley, London.
- [42] Landnyi, B.M., Skaf, M.S., 1993, *Annu. Rev. Phys. Chem.*, **44**, 335.
- [43] Stillinger, F.H., Rahman, A., 1974, *J. Chem. Phys.*, **60**, 1545.

- [44] Matsuoka, O., Clementi, E., Yoshimine, M., 1976, J. Chem. Phys., **64**, 1351.
- [45] Jorgensen, W.L., 1979, J. Am. Chem. Soc., **101**, 2011.
- [46] Jorgensen, W.L., 1979, J. Am. Chem. Soc., **101**, 2016.
- [47] Jorgensen, W.L., 1980, Chem. Phys. Lett., **70**, 326.
- [48] Jorgensen, W.L., 1981, J. Am. Chem. Soc., **103**, 335.
- [49] Impey, R.W., Klein, M.L., McDonald, I.R., 1981, J. Chem. Phys., **74**, 647.
- [50] Impey, R.W., Madden, P.A., McDonald, I.R., 1982, Mol. Phys., **46**, 513.
- [51] Jorgensen, W.L., 1982, J. Chem. Phys., **77**, 4156.
- [52] Jorgensen, W.L., 1982, Chem. Phys. Lett., **92**, 405.
- [53] Dang, L.X., Pettitt, B.M., 1987, J. Phys. Chem., **91**, 3349.
- [54] Berendsen, H.J.C., Griegera, J.R., Straatsma, T.P., 1987, J. Phys. Chem., **91**, 6269.
- [55] Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., 1983, J. Chem. Phys., **79**, 926.
- [56] Hu, Y., Fleming, G.R., 1991, J. Chem., Phys., **94**, 3857.
- [57] Narten, A.H., Levy, H.A., 1971, J. Chem. Phys., **55**, 2263.
- [58] Narten, A.H., 1972, J. Chem. Phys., **56**, 5681.
- [59] Pálinkás, G., Kálmán, E., Kovács, P., 1977, Mol. Phys., **34**, 525.
- [60] Soper, A.K., Silver, R.N., 1982, Phys. Pev. Lett., **49**, 471.
- [61] Thiessen, W.E., Narten, A.H., 1982, J. Chem., Phys., **77**, 2656.
- [62] Narten, A.H., Thiessen, W.E., Blum, L., 1982, Science, **217**, 1033.
- [63] Chou, P.Y., Fasman, G.D., 1978, Annu. Rev. Biochem., **47**, 974.
- [64] Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T.M., Baldwin, R.L., 1990, Nature, **344**, 268.
- [65] Lyu, P.C., Liff, M.I., Marky, L.A., Kallenbach, N.R., 1990, Science, **250**, 669.



- [66] Tobias, D.J., Brooks III, C.L., 1991, *Biochemistry*, **30**, 6059.
- [67] Sung, S.-S., 1994, *Biophys. J.*, **66**, 1796.
- [68] Moret, M.A., Pascutti, P.G., Bisch, P.M., Mundim, K.C., 1998, *J. Comput. Chem.*, **19**, 647.
- [69] Wright, P.E., Dyson, H.J., Lerner, R.A., 1988, *Biochemistry*, **27**, 7167.
- [70] Rogers, N.K., 1990, *Prediction of protein structure and the principles of protein conformation*, ed. by Fasman, G.D., Plenum Press, New York.
- [71] Ramachandran, G.N., Ramakrishnan, C., Sasisekharan, V., 1963, *J. Mol. Biol.*, **7**, 95.
- [72] Brooks III, C.L., Karplus, M., 1989, *J. Mol. Biol.*, **208**, 159.
- [73] Wodak, S.J., van Belle, D., Froeyen, M., Prevost, M., 1989, *Proceedings of an International Meeting*, ed. by Rivail, J.-L., Nancy.
- [74] RASMOL - Molecular Visualization Program, Sayle R., V2.6, 1992-1995, Glaxo Wellcome Research and Development - Stevenage, Hertfordshire, UK.
- [75] Sun, S.F., 1994, *Physical Chemistry of Macromolecules*, John Wiley & Sons, New York.
- [76] Cantor, C.R., Schimmel, P.R., 1980, *Biophysical Chemistry*, W. H. Freeman and Company, San Francisco.
- [77] Beechem, J.M., Brand, L., 1985, *Annu. Rev. Biochem.*, **54**, 43.
- [78] Ross, J.B.A., Wyssbrod, H.R., Porter, R.A., Schwartz, G.P., Michaels, C.A., Laws, W.R., 1992, *Biochemistry*, **31**, 1585.
- [79] Engh, R.A., Chen, L.X.-Q, Fleming, G.R., 1986, *Chem. Phys. Lett.* **126**, 365.
- [80] Szabo, A.G., Rayner, D.M., 1980, *J. Am. Chem. Soc.*, **102**, 554.
- [81] Chang, M.C., Petrich, J.W., McDonald, D.B., Fleming, G.R., 1983, *J. Am. Chem. Soc.*, **105**, 3819.
- [82] Petrich, J.W., Chang, M.C., McDonald, D.B., Fleming, G.R., 1983, *J. Am. Chem. Soc.*, **105**, 3824.

- [83] Swaminathan, R., Krishnamoorthy, G., Periasamy, N., 1994, *Biophys. J.*, **67**, 2013.
- [84] Ito, A.S., Castrucci, A.M.L., Hraby, V.J., Hadley, M.E., Krajcarski, D.T., Szabo, A.G., 1993, *Biochemistry*, **32**, 12264.
- [85] Dezube, B., Dobson, C.M., Teague, C.E., 1981, *J. Chem. Soc., Perkin Trans. 2*, 730.
- [86] Janin, J., Wodak, S., Levitt, M., Maigret, B., 1978, *J. Mol. Biol.*, **125**, 357.
- [87] Gordon, H.L., Jarrel, H.C., Szabo, A.G., Willis, K.J., Somorjai, R.L., 1992, *J. Phys. Chem.*, **96**, 1916.
- [88] Goldman, C., Pascutti, P.G., Piquini, P., Ito, A.S., 1995, *J. Chem. Phys.*, **103**, 10614.
- [89] DTMM - Desktop Molecular Modeller , V3.0, 1994, Oxford University Press.
- [90] XMOL Program, V1.3.1, 1994, Minnesota Supercomputer Center, Inc.