

UNIVERSIDADE ESTADUAL PAULISTA
Faculdade de Engenharia de Ilha Solteira
Departamento de Física e Química

ESTUDO DE PONTES DE HIDROGÊNIO EM MEIO LÍQUIDO
POR SIMULAÇÃO MONTE CARLO

NRD 004101 - Proc. 040/01

UNESP - "CAMPUS DE ILHA SOLTEIRA"	
SERVIÇO TÉCNICO DE BIBLIOTECA E DOCUMENTAÇÃO	
DATA DE CHEGADA	DATA DE TOMBO 31/3/01
REGISTRADO POR Celia	TOMBO 1171
AQUISIÇÃO Doação R\$ 1000 Autogr.	CLASSIFICAÇÃO C744e



Tese apresentada por
João Manuel Marques Cordeiro
visando obtenção do título de livre-docente

1050104- Física estatística e termodinâmica

Auxílio Financeiro: FAPESP

Ilha Solteira - SP
janeiro/2000

Às minhas filhas
Flávia, Larissa e Paula
e seus amigos



INDICE

Resumo	v
1. Introdução	1
2. Teoria	3
2.1 Considerações Gerais	3
2.2 Método Monte Carlo com Algoritmo de Metropolis	5
2.2.1 Processos markovianos	6
2.2.2 O algoritmo de Metropolis	6
2.3. Função Distribuição Radial de Pares	7
2.3.1 Número de Coordenação	8
2.4. Potencial de Interação Intermolecular	9
2.5. Raio de Corte	9
3. Pontes de Hidrogênio	10
3.1 Considerações Gerais	10
3.2 As Pontes de Hidrogênio Fracas	11
3.2.1 Comprimento da ligação	13
3.2.2 Direcionalidade	13
3.2.3 Cooperatividade	14
3.2.4 Interrelação entre pontes de hidrogênio fortes e fracas	15
4. Metodologia	16
4.1 O Potencial de Interação Intermolecular	16
4.2 Os Modelos Moleculares	16
4.3 As Simulações	17
5. Resultados Obtidos e Discussão	19
5.1 Cargas, Parâmetros de Lennard-Jones e Propriedades Termodinâmicas	19
5.2 Correlação Radial de Pares e estruturas representativas dos líquidos puros	25
5.2.1 Acetaldeído	25
5.2.2 Formamida	29
5.2.3 Metilformamida	30
5.2.4 Dimetilformamida	32
5.3 Correlação Radial de Pares e estruturas representativas das soluções	34



5.3.1 A mistura formamida-água	34
5.3.2 A mistura metilformamida-água	37
5.3.3 A mistura acetaldeído-água	39
6. Conclusões	42
7. Tabelas	44
8. Referências	50
9. Proposta de Trabalho para os Próximos 3 Anos	54
Trabalho aceito para publicação	70



RESUMO

Calcularam-se propriedades termodinâmicas e funções de distribuição radial para acetaldeído e suas soluções aquosas e para as soluções aquosas de formamida, metilformamida e dimetilformamida por simulações Monte Carlo, no ensemble isotérmico e isobárico. Um campo de potencial efetivo clássico 6-12 Lennard-Jones mais Coulomb foi otimizado para acetaldeído, no qual as cargas atômicas foram calculadas por análise populacional de Mulliken a nível MP2 com uma base 6-31G*. Os parâmetros de potencial foram otimizados para reproduzir os valores de densidade e calor de vaporização do líquido puro a 18 °C e 1 atm. As funções de distribuição radial obtidas indicam que o líquido possui um bom nível de organização molecular. As funções de distribuição radial são compatíveis com a existência de um trímero representativo das moléculas no volume do líquido. Otimização de geometria desse trímero a nível MM2 mostrou que o mesmo é constituído de duas moléculas paralelas e uma antiparalela. As distâncias entre hidrogênio ligado à carbonila e oxigênio da carbonila de moléculas vizinhas nesse trímero e os cálculos MM realizados, mostram que pontes de hidrogênio fracas desempenham um importante papel na estabilização do mesmo. Da mesma forma, os resultados apontam no sentido de que pontes de hidrogênio fracas entre o hidrogênio ligado à carbonila e oxigênio da água, são responsáveis por uma significativa parcela na interação aldeído-água. Resultados idênticos foram observados presentemente e também em estudos anteriores com dimetilformamida. Isso já não ocorre entre as outras duas amidas. Nestas, as fortes pontes de hidrogênio existentes no sistema, tanto nos líquidos puros, quanto nas soluções aquosas, ditam o comportamento da interação entre as moléculas constituintes dos sistemas. Em todos os três casos de amins estudados, observa-se a ausência de pontes de hidrogênio entre água e o nitrogênio amínico.



RESUMO

Calcularam-se propriedades termodinâmicas e funções de distribuição radial para acetaldeído e suas soluções aquosas e para as soluções aquosas de formamida, metilformamida e dimetilformamida por simulações Monte Carlo, no ensemble isotérmico e isobárico. Um campo de potencial efetivo clássico 6-12 Lennard-Jones mais Coulomb foi otimizado para acetaldeído, no qual as cargas atômicas foram calculadas por análise populacional de Mulliken a nível MP2 com uma base 6-31G*. Os parâmetros de potencial foram otimizados para reproduzir os valores de densidade e calor de vaporização do líquido puro a 18 °C e 1 atm. As funções de distribuição radial obtidas indicam que o líquido possui um bom nível de organização molecular. As funções de distribuição radial são compatíveis com a existência de um trímero que representativo das moléculas no volume do líquido. Otimização de geometria desse trímero a nível MM2 mostrou que o mesmo é constituído de duas moléculas paralelas e uma antiparalela. As distâncias entre hidrogênio ligado à carboníla e oxigênio da carboníla de moléculas vizinhas nesse trímero e os cálculos MM realizados, mostram que pontes de hidrogênio fracas desempenham um importante papel na estabilização do mesmo. Da mesma forma, os resultados apontam no sentido de que pontes de hidrogênio fracas entre o hidrogênio ligado à carboníla e oxigênio da água, são responsáveis por uma significativa parcela na interação aldeído-água. Resultados idênticos foram observados presentemente e também em estudos anteriores com dimetilformamida. Isso já não ocorre entre as outras duas amidas. Nestas, as fortes pontes de hidrogênio existentes no sistema, tanto nos líquidos puros, quanto nas soluções aquosas, ditam o comportamento da interação entre as moléculas constituintes dos sistemas. Em todos os três casos de aminas estudados, assim como em estudos prévios de tetrametilurea, observa-se a ausência de pontes de hidrogênio entre água e o nitrogênio amínico.

1. INTRODUÇÃO

Neste trabalho a investigação se centraliza em dois aspectos: estudar o comportamento termodinâmico e estrutural de líquidos, principalmente soluções binárias, e estudar o papel das pontes de hidrogênio fracas na interação intermolecular de moléculas pequenas. Em um e outro caso, é possível obter uma visão do papel de efeitos de solvente sobre o comportamento molecular do soluto. Se esse soluto serve como modelo para resíduos de moléculas biológicas, pode-se fazer inferências sobre o papel do solvente na conformação e no comportamento reacional dessas moléculas nesse solvente.

Na metodologia aqui utilizada, é fundamental uma boa descrição do estado líquido. Como é de conhecimento de todos, a compreensão do comportamento desse estado da matéria apresenta particular dificuldade, principalmente em razão da complexidade das forças intermoleculares e da ausência de simetria translacional [1]. Nos últimos anos porém, tem havido considerável progresso no tratamento de sistemas em fase líquida, graças sobretudo ao desenvolvimento de novos conceitos e formalismos, com ênfase em teorias de equações integrais e do funcional de densidade, que permitiram a exploração de muitos sistemas e fenômenos não tratáveis classicamente [2]. De outro lado, métodos de simulação computacional, como Monte Carlo (MC) e dinâmica molecular (MD), que permitem de certa forma contornar as dificuldades inerentes de métodos estritamente teóricos, vem se tornando cada vez mais populares na investigação de suas propriedades [3-6]. Estas metodologias, propostas por Metropolis *et al.* em 1953 [7] e Alder e Wainwright em 1959 [8], que a princípio tiveram tímidas aplicações em razão das limitações computacionais, vêm, a medida que computadores mais rápidos e com maior capacidade de memória se tornam acessíveis, adquirindo significativa importância no estudo desses sistemas.

O sucesso da investigação por estas técnicas depende, porém, da otimização de funções de potencial hábeis em reproduzir a interação intermolecular. Tendo-as em mãos, pode-se então utilizar o formalismo da mecânica estatística para calcular propriedades termodinâmicas, dinâmicas e estruturais dos líquidos estudados. Todavia, dada a complexidade das interações intermoleculares, o desenvolvimento de funções de potencial intermolecular não é trivial e certos critérios para simplificação do

potencial de interação, como aditividade aos pares e consideração apenas de interações dipolares, são ainda utilizados na investigação de líquidos polares e outros fluidos, para tornar o trabalho exequível [9-11].

É claro que, operacionalmente, outros problemas relacionados ao cálculo de propriedades macroscópicas surgem, quando se pretende fazer investigação teórica de um determinado sistema. É impraticável o cálculo de médias de observáveis sobre sistemas de tamanho macroscópico, em consequência do enorme número de configurações que se teriam que amostrar em um sistema desse tamanho. Na prática, somente sistemas com algumas centenas de partículas podem ser considerados para investigação. Mas, se um número pequeno de partículas viabiliza o cálculo, introduz a dificuldade de se fazer estimativas sobre propriedades observáveis a partir de sistemas de tamanho tão pequeno. Dentre as metodologias desenvolvidas para superar essa dificuldade, o método Monte Carlo, implementado com o algoritmo de Metropolis e condições de contorno periódicas [7], tem sido das mais utilizadas.

Eu tenho me dedicado nos últimos anos, à investigação teórica de propriedades de líquidos, enfatizando aspectos energéticos, estruturais e de interação soluto-solvente, em líquidos puros e misturas binárias, utilizando o método Monte Carlo [12-15]. Em particular, tenho me debruçado sobre o comportamento de solventes orgânicos polares apróticos e suas soluções aquosas. A compreensão dos processos de solvatação nestes solventes desperta grande interesse em virtude de sua ampla aplicação como meio reacional em química orgânica e por seus acentuados efeitos na velocidade das reações [16].

Em trabalho recente [13] discuti aspectos relacionados à estabilidade das interações $C-H\cdots O$, chamadas pontes de hidrogênio fracas, em amidas. Estas interações, com energias em geral menores que 5 kcal/mol, têm recebido muita atenção nos últimos anos [17-20], em função do papel extremamente importante que desempenham na estrutura de proteínas, especialmente quando ligações de hidrogênio clássicas estão ausentes [21,22]. As peculiaridades dessas interações entre moléculas de dimetilformamida (DMF) em solução aquosa, foram objeto de outro trabalho [15]. Porém, no sistema DMF-água não há a presença de pontes de hidrogênio fortes entre



as moléculas de DMF. No presente trabalho continua-se a aprofundar o estudo das interações de hidrogênio fracas. Para tanto, investiga-se o comportamento das moléculas de formamida e metilformamida (MF), que possuem pontes de hidrogênio tradicionais, em água. Os resultados serão comparados entre si e com os da DMF, já publicados. A título de motivação adicional, vale a pena citar que enquanto formamida e DMF são miscíveis em água, MF é apenas parcialmente solúvel.

Por outro lado, no caso das amidas, as pontes de hidrogênio estão associadas a ressonância peptídica existente nessas moléculas, o que influe no comportamento dessas correlações. Estudos *ab initio* anteriores mostraram que a carga localizada sobre o nitrogênio diminui quando se caminha no sentido da formamida para a DMF, não se verificando portanto o chamado efeito doador de elétrons atribuído ao grupo metil [13]. Tal resultado foi interpretado como consequência de variação na “intensidade” da ressonância nessas moléculas. Para fugir desse caso, analisam-se os resultados obtidos na investigação do acetaldeído, líquido que contempla a possibilidade de interações do tipo $C-H\cdots O$, sem a presença de ressonância. A investigação caminhará no sentido de estudar o comportamento do acetaldeído em água. Este trabalho não pretenderá ser conclusivo, mesmo porque esta pesquisa enseja a realização de uma quantidade apreciável de cálculos e simulações à medida que o trabalho prossegue. Procurar-se-á apresentar e discutir aqui alguns dos resultados até agora obtidos, que estejam de alguma forma relacionados e com os quais se possa avançar posteriormente na análise.

2. TEORIA

2.1 Considerações Gerais

As propriedades termodinâmicas de uma molécula se relacionam com seus níveis energéticos através da função de partição [23]. A utilização de métodos *ab initio* ou semi empíricos permite determinar os níveis de energia molecular, mas o cálculo torna-se mais e mais dispendioso computacionalmente, à medida que cresce o número de partículas do sistema. Quando, em função disso, tais cálculos tornam-se impraticáveis, tem-se lançado mão da aproximação clássica [9,10,23], calculando-se a energia do sistema em termos do hamiltoniano clássico



$$H_c(p^N, r^N) = \sum_{i=1}^{\infty} \frac{p_i^2}{2m} + U_N(X^N), \quad (1)$$

onde $U_N(X^N)$ é a superfície de energia potencial do sistema com N partículas, e p_i os momenta das partículas que constituem o sistema. A função de partição é então transformada de uma soma sobre níveis de energia para uma sobre as possíveis posições e momenta das partículas. Assim sendo, energia interna, entalpia, e demais propriedades relacionadas às suas flutuações, que dependem diretamente da função de partição, poderiam ser calculadas, empregando-se a abordagem mecânico-estatística. Neste caso, o valor de uma propriedade termodinâmica observável é uma média no *ensemble* de estados microscópicos do sistema. A escolha do *ensemble* é definida pelas condições “experimentais” do sistema em estudo. Em química, o *ensemble* mais utilizado (inclusive no presente trabalho), é o isotérmico-isobárico (NpT).

No *ensemble* NpT, a média de um observável F , função da configuração X^N , é dada por [9]

$$\langle F \rangle = \int F(X^N) P(X^N) dX^N \quad (2)$$

onde $P(X^N)$ é a função distribuição de probabilidade. Nesse *ensemble*, a função de distribuição pode ser expressa por [9,23]:

$$P(X^N, V) = \frac{\exp[-\beta PV - \beta U_N(X^N)]}{\int_0^\infty dV \int \exp[-\beta PV - \beta U_N(X^N)] dX^N} \quad (3)$$

onde se omitiu os termos de energia cinética no hamiltoniano. Nessa equação $\beta = 1/k_B T$, sendo k_B a constante de Boltzmann e T a temperatura. O denominador da equação acima é a função de partição configuracional L_{NpT} do *ensemble* isotérmico-isobárico. Cada configuração i tem probabilidade $P(X_i^N, V)$ de ser observada.

A questão agora é: como calcular as integrais configuracionais da equação acima num sistema de muitos corpos, em que o número de possíveis

configurações espaciais do sistema é extremamente grande, para se poder obter as médias dos observáveis?

2.2 Método Monte Carlo com Algoritmo de Metrópolis

Esse problema foi resolvido por Metropolis e col. [7], que implementaram no método Monte Carlo clássico, uma rotina combinando condições periódicas de contorno (o que permite simular um sistema macroscópico com um número pequeno de partículas) e um procedimento para a seleção das configurações mais prováveis, sendo que as posições moleculares são geradas por um processo de Markov [24]. Por condições periódicas entende-se: o líquido modelado consiste de uma caixa contendo um número N de moléculas especificado, usualmente algumas centenas, e um procedimento que “reflete” uma molécula dentro da caixa, quando ela se projeta para fora da mesma. Por outro lado, quando uma molécula próxima à fronteira da caixa é selecionada, a caixa é “replicada” próximo a ela, de tal forma que uma molécula qualquer, quando selecionada, nunca se encontre próximo à fronteira do sistema. Com este procedimento, reduz-se um problema de alta dimensionalidade a um problema tratável computacionalmente, ao mesmo tempo que se simula o sistema em equilíbrio termodinâmico.

Condições como aqui descritas são aplicáveis somente a sistemas homogêneos. Uma vez resolvida a questão de se poder trabalhar com um número viável de partículas, o passo seguinte é calcular o valor médio de F , $\langle F \rangle$. O método MC calcula numericamente as integrais envolvidas nas equações 2 e 3, aproximando-as por somas sobre um número n de pontos escolhidos aleatoriamente no espaço de fase [25,26].

Entretanto, esse método de amostragem aleatória é inaplicável em muitos dos problemas de interesse, porque o número n de configurações que se deveriam amostrar para se obter os valores médios de um observável termodinâmico seria extremamente grande num espaço de muitas dimensões. Além disso, em uma célula com N moléculas, uma amostragem aleatória de estados, produziria, na maioria das vezes, uma pequena separação entre ou mesmo superposição de duas moléculas, o

que implicaria em um grande potencial positivo. Dessa forma, se estaria amostrando um grande número de configurações que são energeticamente inacessíveis [27]. Novamente entra em cena o trabalho de Metropolis e col. [7], que propõem um procedimento de seleção das configurações, que ficam restritas à região mais densa do espaço de fase e fisicamente mais prováveis (*importance sampling*).

2.2.1 Processos markovianos

Para isso Metropolis e col. lançaram mão dos fundamentos de processos markovianos [24]. Num processo de Markov, uma sequência aleatória de pontos do espaço de fase é gerada, em que a probabilidade de um ponto gerado depende do, e somente do, ponto anterior. Especificada uma probabilidade de transição entre dois pontos consecutivos, que pode ser um número ou um intervalo, pontos posteriores a um dado ponto, que não satisfaçam essa probabilidade estão automaticamente excluídos. Então a probabilidade de transição depende somente da mudança de energia $\Delta U_N(X^N)$ entre os dois estados. Um processo de Markov tem a propriedade de “levar” um sistema para o equilíbrio, definido como o estado em que, para qualquer par de pontos gerado a probabilidade de transição num sentido e em sentido inverso é a mesma e igual a $\exp[-\beta \Delta U_N(X^N)]$.

2.2.2 O algoritmo de Metropolis

São então distribuídas N moléculas em uma configuração arbitrária, numa célula cúbica de referência, e uma nova configuração é gerada pela translação, rotação e relaxação de uma molécula escolhida aleatoriamente. Dadas então duas configurações consecutivas, que diferem uma da outra pela posição de uma molécula, pode-se calcular a diferença de energia ΔU entre essas duas configurações por:

$$\Delta U_N = U_{N+1}(X^N) - U_N(X^N) \quad (4)$$

Se a energia da configuração posterior é menor que a da anterior, isto é, $\Delta U_N < 0$, a configuração posterior é mais provável que a anterior e passa a ser a nova configuração do sistema. Caso contrário, gera-se um número aleatório ϵ entre 0 e 1 e compara-se esse número com o valor $\exp[-\beta \Delta U_N(X^N)]$. Se $\epsilon < \exp[-\beta \Delta U_N(X^N)]$ a

nova configuração é aceita, senão o movimento é recusado, e persiste a configuração anterior. Este procedimento faz com que, configurações desfavoráveis do ponto de vista da energia potencial, mas que poderiam ser acessadas graças a energia cinética da molécula, sejam consideradas. Quando um movimento é recusado, a configuração anterior é novamente contada para o cálculo da média de um observável. A repetição desse procedimento constituirá uma sequência markoviana de configurações distribuídas na região de maior probabilidade do espaço de fase configuracional.

Construída a cadeia de Markov com M configurações, a média de um observável será calculada de acordo com:

$$\langle F \rangle = \frac{1}{M} \sum_{i=1}^M F_i \quad (5)$$

onde F_i é o valor da propriedade F para o sistema na i -ésima configuração.

Resumidamente, o algoritmo de Metropolis pode ser escrito como:

- 1- toma-se uma dada configuração designada pelo índice i e com energia potencial U_i
- 2- gera-se uma nova configuração $i+1$ (movendo-se aleatoriamente uma partícula) e calcula-se a diferença de energia potencial $\Delta U = U_{i+1} - U_i$.
- 3- se $\Delta U < 0$, esta nova configuração é aceita. Volta-se à etapa 2.
- 4- se $\Delta U > 0$ gera-se um número aleatório ξ ($0 < \xi < 1$) e compara-se com o valor de $\tau = \exp[-\beta \Delta U]$. Se $\tau > \xi$, a nova configuração é aceita. Volta-se à etapa 2. Caso contrário, a nova configuração é rejeitada e a configuração i é contada novamente. Volta-se à etapa 2.

Quando no *ensemble* de interesse o volume é uma variável, como no caso do isotérmico e isobárico, uma nova configuração pode ser gerada a partir de uma variação no volume do sistema.

2.3 Função Distribuição Radial de Pares

As funções de distribuição desempenham um importante papel na compreensão da estrutura dos líquidos. Por estrutura de um líquido entende-se: a



disposição média relativa entre as partículas que o constituem. Experimentalmente, informações sobre a função de distribuição radial, podem se obter por espalhamento de neutrons ou difração de raios X. Porém, o fato de haverem distâncias de correlação que se sobrepõem, o que significa que há diferentes correlações que contribuem experimentalmente para uma banda, torna o levantamento experimental da função de distribuição radial de um líquido extraordinariamente difícil. Em consequência disto, o emprego de modelos moleculares e simulação computacional para obtenção das funções de correlação é extraordinariamente interessante como ferramenta para “visualizar” a estrutura do líquido. A função de distribuição radial de pares $g_{ij}(r)$, entre os átomos i e j de um par de moléculas distintas, pode ser definida por: [27]

$$g_{ij}(r) = \frac{N_{ij}(r, r + \Delta r)}{4\pi r^2 \Delta r \rho_j} \quad (6)$$

onde r é a separação entre os átomos, $N_{ij}(r, r + \Delta r)$ é o número médio de átomos j encontrados na camada esférica entre r e $r + \Delta r$ centrada no átomo i , $4\pi r^2 \Delta r$ é o volume da camada esférica de espessura Δr e ρ_j é a densidade numérica média de átomos j no líquido. A distribuição radial de pares pode ser interpretada como sendo a razão entre a densidade local de átomos j e a densidade média desse átomo no líquido, o que confere à distribuição radial de pares um sentido probabilístico. Se a distribuição local de átomos j for a mesma da distribuição média, então $\frac{N_{ij}(r, r + \Delta r)}{4\pi r^2 \Delta r} \rightarrow \rho_j$, de modo que $g_{ij}(r) \rightarrow 1$, o que se verifica para distâncias de separação r grandes.

O cálculo de $g_{ij}(r)$ envolve basicamente o registro das distâncias r entre os átomos i e j , que é armazenado em um histograma. Esse histograma é incrementado ao longo da simulação e por fim é normalizado para o número N de moléculas utilizadas na simulação e para o número de configurações do histograma.

2.3.1 Número de coordenação

Os picos das funções de distribuição radial representam as distâncias médias mais prováveis entre vizinhos e, no caso de líquidos, são associados às camadas

de solvatação. A integração da equação (6) até os mínimos correspondentes fornece estimativa do número de vizinhos, ou número de coordenação C_{ij} . Por exemplo, para a primeira camada de solvatação com um mínimo em R_c temos,

$$C_{ij}(R_c) = \int_0^{R_c} g_{ij}(r) \rho_j 4\pi r^2 dr \quad (7)$$

2.4 Potencial de Interação Intermolecular

Como deve ter ficado claro do que se discutiu acima, o processo de amostragem utilizado no método Monte Carlo é fundamentalmente dependente da energia configuracional. Portanto é necessário dispor de uma função de potencial a partir da qual se calcule essa energia. Na maioria dos trabalhos relatados, (só recentemente, com disponibilidade de melhores facilidades computacionais, rotinas para relaxação molecular tem sido implementadas), relaxações moleculares associadas aos movimentos translacionais e rotacionais moleculares não são consideradas (a geometria molecular é mantida fixa durante a simulação). Assim sendo, a energia configuracional consiste exclusivamente de energia de interação intermolecular.

Contribuições intermoleculares para a energia podem ser incluídas de diversas formas, sem alterar a estrutura geral apresentada acima, na simulação. Uma aproximação comumente utilizada neste tipo de trabalho, inclusive aqui, é representar a energia intermolecular por uma soma sobre pares de moléculas (*pairwise additivity*). Neste caso, o potencial configuracional é dado por $U = \sum_{a < b} E_{ab}$, onde a e b são duas moléculas quaisquer do sistema e E_{ab} a energia de interação entre elas.

2.5 Raio de Corte

Porém, o cálculo da energia potencial considerando-se pura e simplesmente todos os pares de moléculas do sistema, demandaria muito tempo computacional com contas desnecessárias, uma vez que moléculas muito distantes entre si não contribuiriam para a energia total. Para simplificar o cálculo então, usa-se

a metodologia de raio de corte r_c [28], na qual a interação de pares é calculada efetivamente dentro de um certo raio, utilizando-se $U = \sum_{a < b} E_{ab}$, e depois adiciona-se um termo médio calculado pela equação, [25]

$$E_{LRC} = (N/2) \sum_i^{\text{sobre } a} \sum_j^{\text{sobre } b} \int_{r_c}^{\infty} 4\pi r^2 \rho_j g_{ij}(r) u_{ij}(r) dr \quad (8)$$

para incorporar a parte de energia configuracional que se “perde” com o tratamento de raio de corte. Na implementação da equação acima, assume-se que $g_{ij}(r) = 1$ além do raio de corte, o que se mostra adequado pelos gráficos de $g(r)$ que se obtêm. Melhor estabilidade numérica é conseguida multiplicando-se o potencial de interação por um função *smoothing* no intervalo $[r_c - 0,5\text{\AA}; r_c]$ [29]. Com essa metodologia o potencial de interação é reduzido a zero nos limites da esfera definida pelo raio de corte, de uma forma contínua, evitando-se assim que a energia potencial salte de um valor dado pelo número de moléculas dentro do raio de corte para zero em r_c .

3. PONTES DE HIDROGÊNIO

3.1 Considerações Gerais

Pontes de hidrogênio são, indubitavelmente, as formas de interação não ligante mais estudadas em química. Seu papel em uma série de fenômenos relacionados com o comportamento termodinâmico e estrutural de líquidos, quer sejam puros quer sejam misturas, foi realçado a partir do momento que se teve conhecimento da importância dessas interações na estrutura de proteínas, ácidos nucleicos e outras moléculas bioquimicamente fundamentais. Embora critérios geométricos (juntamente com energéticos) tenham sido largamente empregados para caracterizar ligações de hidrogênio, significativas variações na distância $H \cdots O$ são observadas, em consequência dessas interações serem de cunho eminentemente eletrostático [22], contrariamente ao que ocorre com interações do tipo van der Waals. Por outro lado, ligações de hidrogênio $X-H \cdots A$ são basicamente lineares, enquanto que interações de

van der Waals são isotrópicas [18]. O aceptor A deve apresentar carácter básico (alternativamente, deve carregar uma carga parcial negativa), enquanto o doador X-H deve ser pelo menos levemente ácido [17].

3.2 As Pontes de Hidrogênio Fracas

Muita discussão tem havido a respeito da possibilidade de interações C-H...O poderem ser consideradas pontes de hidrogênio [21]. Recentemente ficou claro que essas interações, direcionais e coesivas, representam interações do tipo de hidrogênio [17,20], chamadas pontes de hidrogênio fracas. Está presentemente bem documentado que este tipo de interação desempenha um papel fundamental em estrutura cristalina [30], conformação molecular [31], estabilização de complexos de inclusão [32], e, do ponto de vista que interessa a este trabalho, na estabilidade e atividade de macromoléculas biológicas [33].

Grupos C-H apresentam carácter doador de hidrogênio [20,21]. A dependência entre a acidez do hidrogênio e a hibridação do carbono foi estabelecida à bastante tempo [34], verificando-se a seguinte ordem $C(sp)-H > C(sp^2)-H > C(sp^3)-H$. Algum efeito doador foi encontrado mesmo em grupos metil fracamente polarizados [35]. Recentemente, análise estatística de bancos de dados cristalográficos mostrou que a distância média H...O em interações C-H...O, está irrefutavelmente correlacionada com a acidez verificada nos grupos C-H [36] o que depende da natureza do grupo R ligado ao carbono [37]. Alcinos terminais apresentam separação H...O entre 2.1 e 2.6 Å [38], muito semelhante portanto ao comprimento de ligação das pontes de hidrogênio normais.

Assim como em toda interação intermolecular, há diferentes tipos de contribuição para a interação de hidrogênio [39]. Há uma contribuição de van der Waals não direcional, que é fracamente ligante a longas distâncias e acentuadamente não ligante a pequenas distâncias (nas melhores condições, forças de van der Waals contribuem com alguns décimos de kcal/mol para a energia da ponte de hidrogênio) e um componente eletrostático direcional e ligante em qualquer distância, que aumenta com o aumento do momento de dipolo ou a carga envolvida e diminui com o aumento da distância. As cargas parciais tipicamente atribuídas ao átomo H variam entre 0 e



+0,2 elétron [20], dependendo do sistema em estudo. Para os doadores N-H, O-H ou grupos C-H fortemente polarizados, o componente eletrostático é o dominante na ligação, enquanto que no caso de C-H fracamente polarizado esta contribuição é de magnitude similar à contribuição de van der Waals. Somente nos tipos mais fortes de pontes de hidrogênio a transferência de carga tem um papel relevante, sendo inexpressivo nas fracas interações C-H...O [39]. Em consequência destas observações pode-se inferir o seguinte:

- i) forças eletrostáticas variam com a geometria e são levemente dependentes da distância. Essa dependência é função da acidez do grupo C-H, o que leve a uma acentuada flexibilidade das fracas interações C-H...O.
- ii) o carácter ligante é fortemente dependente da polarização do doador e aceptor, o que é ditado pelas vizinhanças do grupo C-H participante da ligação.
- iii) com a diminuição da polarização a componente eletrostática se reduz com conseqüente perda de direcionalidade.

As energias atribuídas à interação C-H...O são pequenas, difilmente determináveis experimentalmente, de onde segue que a maioria dos resultados relatados estão baseados em cálculos teóricos. Correntemente há um consenso que as energias da interação C-H...O são tipicamente menores que 2 kcal/mol e diminuem gradualmente com o aumento da separação H...O (as energias de pontes de hidrogênio O-H...O se encontram na faixa de 3-8 kcal/mol) [39]. Porém em casos de grande acidez do grupo C-H, como em $\text{N}\equiv\text{C}-\text{H}$ ou com carácter iônico, esses valores podem chegar aos dos das pontes de hidrogênio "normais" [40].

Todavia, como já salientado, e porque interessa neste trabalho, pontes de hidrogênio são fortemente influenciadas pelas vizinhanças, como solvente ou a existência de outras pontes, tal que a força da interação varia. As pontes de hidrogênio alteram a ligação covalente dos grupos envolvidos. Numa interação $\text{R}_1-\text{X}-\text{H}\cdots\text{O}-\text{R}_2$ as ligações X-H e O-R₂ são enfraquecidas, com conseqüente estiramento, enquanto a ligação R₁-X torna-se mais forte e mais curta. Estes efeitos são, logicamente, muito pronunciados em ligações de hidrogênio normais, mas efeitos análogos ocorrem em

menor intensidade em interações $C-H\cdots O$ [41]. Por outro lado, pontes de hidrogênio reduzem a vibração dos grupos envolvidos, enrijecendo a estrutura.

3.2.1 Comprimento da ligação

As menores distâncias de separação $C\cdots O$ e $H\cdots O$ observadas em estruturas cristalinas são aproximadamente 3,0 Å e 2,0 Å respectivamente. Tais distâncias ocorrem com doadores $C-H$ fortemente ácidos, mas ainda assim, não são as distâncias “normais”. Os valores típicos das distâncias $H\cdots O$ se encontram no intervalo 2,1 e 2,3 Å [20]. No caso de doadores fracamente ácidos, as distâncias de interação são maiores e vão se alongando à medida que a acidez do grupo $C-H$ diminui. Além disso, a variação observada nas distâncias de interação $H\cdots O$ se deve não só a diferença na acidez do hidrogênio doador como também ao fato que, em competição com outras forças intermoleculares, ligações de hidrogênio de menor energia serão frequentemente perdedoras, ficando a geometria da ligação subordinada ao jogo das forças interatômicas e à existência ou não de impedimentos estéricos. Em consequência disto também, é errôneo associar energia da ligação com seu comprimento, como se faz com certa frequência, ou decidir pela existência ou não de interação de hidrogênio entre um par doador-aceptor, em função do comprimento da interação $H\cdots O$. Menor distância de ligação não significa necessariamente maior energia de interação, ou vice-versa.

3.2.2 Direcionalidade

Ligações de hidrogênio são em geral interações direcionais com forte preferência pela linearidade. No caso das interações $C-H\cdots O$, este fato foi primeiramente verificado por Taylor e Kennard [42]. Todavia, neste caso, devido a força da interação, a linearidade pode ser facilmente perdida em consequência de competição com outras pontes de hidrogênio e impedimentos estéricos. Isto é claramente observado em contactos $C-H\cdots O$ de carboidratos, onde ligações de hidrogênio $O-H\cdots O$ são as interações dominantes formando uma estrutura onde as interações $C-H\cdots O$, mais fracas, devem se ajustar [43]. Apesar disto, a maioria das



interações $C-H\cdots O$ tem distâncias de separação $H\cdots O$ menores que 2,25 Å, verificando-se que quanto mais curta essa correlação, mais acentuada a linearidade. Do ponto de vista do aceptor, a direcionalidade é mais complexa e depende da sua natureza química. Para carbonil-aceptores, o caso que interessa diretamente neste trabalho, os vetores $C-H$ tendem a apontar para os pares de elétrons do oxigênio no plano do grupo $C=O$, teoricamente formando ângulos de 120° [44]. Em geral, a geometria do grupo $C-H\cdots O=C$ é planar [45].

3.2.3 Cooperatividade

Arranjos de n pontes de hidrogênio interconectadas têm energia maior que a soma de n pontes de hidrogênio isoladas. Isto se deve basicamente a dois mecanismos:

- i) grupos funcionais que podem agir como doadores e aceptores, formando cadeias. Nestas cadeias, devido a polarização múltipla dos grupos envolvidos, cada ponte de hidrogênio individual aumenta a força das pontes vizinhas em cerca de 20% [39].
- ii) ressonância assistida por pontes de hidrogênio [46].

No caso do primeiro mecanismo, tem sido mais estudado o grupo $C\equiv C-H$ [36-38], embora neste trabalho interesse mais de perto o que ocorre como o grupo aldeídico. Como nestes casos são difíceis de obter evidências experimentais de aumento de energia, cálculos teóricos desempenham importante papel a fim de se entender propriedades observadas em substâncias que apresentam estes grupamentos. No caso em particular deste trabalho, tem-se ainda pontes de hidrogênio assistidas por ressonância, devido a existência de dímeros entre moléculas de amidas [13], donde se deve esperar uma interação mais forte do que uma ponte de hidrogênio simples. Cálculos semiempíricos foram feitos por Sharma e Desiraju para N,N -dimetilnitroamina, os quais encontraram um aumento de cerca de 20% na energia das pontes de hidrogênio das cadeias formadas neste material [47].

3.2.4 Interrelação entre pontes de hidrogênio fortes e fracas.

Em estruturas peptídicas em geral, interações $C-H\cdots O$ são subalternas às clássicas pontes de hidrogênio. A despeito disto porém, pontes de hidrogênio fracas têm papel similar àquelas, na formação estrutural e estabilização dessas macromoléculas [20], em alguns casos, contribuindo para estabilizar ainda mais as pontes de hidrogênio clássicas formadas [13]. Sem dúvida, interações $C-H\cdots O$ são formadas como último recurso se pontes de hidrogênio fortes estão presentes, mas, se elas se formam, doadores $C-H$ desempenham o mesmo papel de $O-H$ ou $N-H$ [48]. Vários exemplos de estruturas isomorfas tem sido publicados onde pontes de hidrogênio clássicas em uma estrutura são funcionalmente substituídas por interação $C-H\cdots O$ na outra [49]. Em todos os casos estudados, ambas as interações tem a mesma função na estabilização estrutural, isto é, a interação $C-H\cdots O$ tem o mesmo papel de uma ponte de hidrogênio forte.

Abordagem teórica tem muito a contribuir no estudo de interações de hidrogênio fracas, uma vez que um grande número de questões são difíceis de serem abordadas experimentalmente. Entre elas podemos citar: interação de longo alcance das pontes de hidrogênio fracas, efeitos cooperativos e de não aditividade, competição com outras forças de interação intermolecular, energia de interação, propriedades termodinâmicas em geral e efeitos de solvente.



4. METODOLOGIA

4.1 O Potencial de Interação Intermolecular

Neste trabalho, como se faz classicamente nesta área, a energia de interação intermolecular foi calculada utilizando-se a equação:

$$E_{ab} = \sum_{ij} \left[A_{ij} / (r_{ij})^{12} - B_{ij} / (r_{ij})^6 + q_i q_j / r_{ij} \right] \quad (9)$$

onde i é um átomo na molécula a , j é um átomo na molécula b e r_{ij} a distância entre i e j . Para cada átomo k , os parâmetros A_{kk} e B_{kk} são definidos por $A_{kk} = 4\epsilon_k \sigma_k^{12}$ e $B_{kk} = 4\epsilon_k \sigma_k^6$, onde ϵ_k e σ_k são os parâmetros de Lennard-Jones para o átomo k . Os parâmetros A_{ij} e B_{ij} para a interação não-diagonal $[ij]$ foram obtidos usando as regras de combinação $A_{ij} = (A_{ii} A_{jj})^{1/2}$ e $B_{ij} = (B_{ii} B_{jj})^{1/2}$ [25]. Em trabalho anterior fez-se uma comparação entre resultados obtidos com essa regra de combinação (média geométrica) e com a também conhecida regra de Lorentz-Berthelot [12]. Os resultados mostraram que ambas são equivalentes para o cálculo dos potencial cruzados.

Os parâmetros q_i e q_j são frações de carga eletrônica localizadas sobre os átomos. Tais cargas procuram reproduzir a assimetria da densidade de carga das moléculas e assume-se que sejam uma forma adequada de representar a contribuição eletrostática ao potencial intermolecular. Em geral, os parâmetros ϵ , σ e q são determinados inicialmente para dímeros em fase gasosa e depois ajustados para reproduzir resultados energéticos e estruturais dos líquidos, medidos experimentalmente.

4.2 Os Modelos Moleculares

Na Fig. 1 representa-se a molécula da MF, a partir da qual se pode imaginar as duas outras moléculas aqui estudadas, ou seja, a formamida e o acetaldeído.

Enquanto na MF tem-se um nitrogênio secundário, na formamida o nitrogênio é primário. No aldeído, o grupo amínico estará substituído por um grupo metil. Só para lembrar, porque é também citado neste trabalho, na DMF o nitrogênio é terciário.

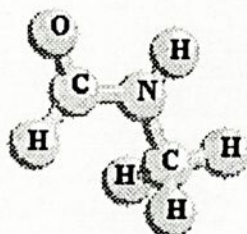


Figura 1. A molécula da metilformamida

As moléculas são rígidas sendo que na MF o grupo metil é representado por um sítio localizado sobre o átomo de carbono. A cada átomo/sítio é atribuído um conjunto de parâmetros de Lennard-Jones e uma carga fracionária. Os parâmetros de potencial para as amidas foram determinados anteriormente [13] enquanto que no estudo envolvendo acetaldeído se utilizam parâmetros otimizados neste trabalho.

4.3 As Simulações

As simulações foram realizadas com Metropolis *importance sampling* e condições de contorno periódicas no *ensemble* NPT a 298 K e 1 atm usando-se o programa DIADORIM [50]. Todos os sistemas investigados eram constituídos de 500 moléculas numa caixa cúbica. O número de moléculas dos componentes nas soluções aquosas foram adequadamente ajustados de forma a se compor as diferentes concentrações estudadas. Para a água se utilizou o modelo SPCE de 3 sítios otimizado por Berendsen *et al* [53]. Em trabalho anterior foi mostrado que, entre os modelos simples disponíveis de água, o modelo SPCE é o que melhor reproduz as propriedades do líquido em diversas condições termodinâmicas [54]. As coordenadas moleculares iniciais são obtidas distribuindo-se as moléculas segundo um retículo de face centrada [51]. O volume inicial da caixa é calculado para reproduzir a densidade ideal da mistura (a densidade experimental no caso de líquidos puros).

No cálculo da energia total configuracional, utiliza-se a equação (9) entre todas as moléculas que tenham pelo menos um sítio dentro de um raio de corte de 10 Å. A contribuição da interação Lennard-Jones além do raio de corte é considerada usando-se o formalismo apresentado por Allen e Tildesley [25], equação (8). Contribuições da interação eletrostática além do raio de corte não foram consideradas, ou seja, na implementação da equação (8), u_{ij} contém apenas o termo de Lennard-Jones. Esta aproximação tem sido sistematicamente utilizada [5] para minimizar o custo computacional, e se baseia na idéia que, em se tratando de moléculas neutras, a principal contribuição de longo alcance do potencial eletrostático para a energia de interação vem da interação dipolo-dipolo. Esta contribuição decresce assintoticamente com $[1/r^3]$, onde r é aproximadamente a distância entre o centro de massa das duas moléculas consideradas. Como tem sido demonstrado, nestes casos a média angular da interação dipolo-dipolo decresce assintoticamente com $[1/r^6]$ [52]. Consequentemente, a contribuição assintótica do potencial eletrostático é significativamente menor do que se esperaria para o caso puramente coulômbico. Tal aproximação é ainda mais justificada pelo fato de que os parâmetros de potencial molecular são otimizados nas mesmas condições computacionais das simulações propriamente ditas, de tal forma que já trazem incorporados em seus valores esta aproximação. Começando da distribuição de moléculas inicial na caixa, novas configurações são geradas por translação e rotação aleatória de uma molécula também escolhida ao acaso. Além disso, novas configurações podem ainda ser geradas permitindo-se alterações no volume do sistema, de tal forma a manter a densidade do líquido constante. Após uma alteração no volume as coordenadas do centro de massa das moléculas do sistema são reescaladas da maneira usual [25]. Os valores dos passos de translação (Δr) e rotação ($\Delta \theta$) são ajustados de forma a produzir uma razão entre o número de configurações aceitas e recusadas da ordem de 0,45. Movimentos no volume foram tentados a cada 500 configurações e o passo de volume também ajustado de forma a que a razão entre o número de movimentos aceitos e recusados seja dessa mesma ordem. Cada simulação de um sistema novo começa com uma fase de equilíbrio de 8×10^6 configurações, seguindo de uma fase para cálculo das médias das propriedades de $1,6 \times 10^7$ configurações. Desvios foram calculados de médias sobre blocos de 2×10^5

configurações. As simulações foram realizadas em microcomputadores Pentium Pro de 300 MHz com sistema operacional FreeBSD.

5. Resultados Obtidos e Discussão

5.1 Cargas, Parâmetros de Lennard-Jones e Propriedades Termodinâmicas

Um dos pontos cruciais na otimização de parâmetros de potencial é a determinação das cargas sobre os sítios moleculares. De alguns anos para cá, tem sido recomendado a utilização do método CHELPG do GAUSSIAN para se estimar a carga sobre os sítios moleculares [55]. Esse método têm sido utilizado também em nosso trabalho nos últimos dois a três anos. Até então, as cargas utilizadas nos modelos moleculares parametrizadas por nós [13], assim como em trabalhos de outros pesquisadores, eram transferidas de outras moléculas com grupamentos semelhantes reportadas e ajustadas por uma série de simulações Monte Carlo, de forma a se reproduzir valores de propriedades experimentais do líquido em questão. Simultaneamente, e com a mesma metodologia, são feitos ajustes nos parâmetros de Lennard-Jones da molécula. Em nosso caso, seguindo metodologia usado no grupo do Dr. Willian Jorgensen, são “fitados” os valores de entalpia de vaporização e densidade. Todavia, observamos, depois de tabelar um número significativo dessas cargas, que elas (sobretudo em alguns casos) são significativamente diferentes das cargas CHELPG. Por outro lado, tentativas feitas neste trabalho de reparametrizar algumas das moléculas anteriormente parametrizadas, usando-se as cargas CHELPG, mostraram-se totalmente infrutíferas, ou seja, não se conseguiu “fitar” os resultados experimentais de densidade e entalpia de vaporização dos líquidos usando-se cargas CHELPG.

Aqui, parece importante a informação dada por Luciene Borges, do Laboratório de Química Teórica da UFSCar, que esteve recentemente trabalhando no grupo do Dr. Jorgensen. Segundo ela, o primeiro parâmetro que lá se “ajusta” para tentar reproduzir os resultados experimentais é a carga e só depois, no sentido de se obter um melhor acordo entre valores teóricos e experimentais, se ajustam os parâmetros L-J. Em outras palavras, o grupo do Dr. Jorgensen usa e recomenda a metodologia CHELPG para



obtenção das cargas pontuais [55], mas essas cargas são depois manipuladas para obtenção dos parâmetros moleculares do líquido. O que parece, de um levantamento feito neste trabalho de potenciais OPLS de várias moléculas, é que nessa otimização as cargas finais sobre os sítios são, pelo menos em alguns casos, mais parecidas às da densidade populacional de Mulliken do que às CHELPG.

Ao tentar, reparametrizar as cargas das amidas que haviam sido anteriormente parametrizadas [13], partindo-se de cargas CHELPG, verificou-se que para se reproduzir os valores experimentais de densidade e entalpia de vaporização, dever-se-ia fazer as cargas CHELPG tenderem às de Mulliken. Efetivamente, como se observou depois, as cargas sobre os sítios das moléculas das amidas anteriormente parametrizadas [13], são praticamente cargas de Mulliken obtidas no cálculo com aproximação MP2. Isto posto, fez-se um estudo mais pormenorizado sobre o cálculo das cargas pontuais para as moléculas aqui estudadas, incluindo a DMF que complementa a série das amidas com respeito à substituição do nitrogênio amínico. As cargas foram calculadas com aproximação HF e MP2, base 6-31G* e otimização de geometria, usando-se o programa GAUSSIAN 94 [56], e os resultados obtidos estão mostrados nas Tabelas 1 a 4. Uma comparação é feita entre as cargas de análise populacional de Mulliken e CHELPG.

Via de regra, as cargas de Mulliken e CHELPG tem valores diferentes, como seria esperado, uma vez que são calculadas baseadas em princípios diferentes. Chama a atenção todavia que, embora as cargas sejam diferentes, os momentos de dipolo fornecidos por ambos os conjuntos de cargas sejam iguais. Pareceria lógico que isso pudesse acontecer eventualmente mas não fosse regra geral, como parece ser (não se tratou até agora de entender esta verificação). Os momentos obtidos por HF são em torno de 10% maiores do que os obtidos por MP2. Neste caso também parece haver alguma contradição com o que se tem difundido [ref. 55 e alguns trabalhos depois desse] como sendo uma razão em prol da utilização de cargas CHELPG, quer seja, a de que o momento de dipolo da molécula com tais cargas é de 15 a 20% maior do que o experimental, o que é desejável na simulação de líquidos para compensar a falta de relaxação molecular. Como se vê, as cargas de Mulliken apresentam basicamente o mesmo resultado, e portanto, também satisfariam esse critério. Como acima antecipado,

comparação entre as cargas dos modelos de amidas reportadas anteriormente [13] e as obtidas neste estudo mostra que as cargas utilizadas naqueles modelos são praticamente as de Mulliken com aproximação MP2. Portanto, análise populacional de Mulliken com cálculos a nível MP2, parece fornecer (pelo menos em certos casos) valores de cargas mais adequados para simulação de líquidos do que CHELPG.

Outro dado importante é que a diferença entre os valores das cargas obtidos com e sem correlação eletrônica, pode chegar a 30% (observar por exemplo os valores das cargas do C carbonílico em algumas das amidas), o que nos parece significativo. Ficaria então colocada a questão de qual metodologia seria mais adequada para se obter essas cargas. Como acima citado, nas moléculas por nós estudadas, aproximação MP2 é a que melhores resultados fornece. Contrariamente, tem sido divulgado que as cargas obtidas seriam independentes do tipo de cálculo feito [57]. Uma possibilidade que foi levantada por Gomide e Borges (Laboratório de Química Teórica - DQ - UFSCar), para explicar o que está acontecendo, é que a semelhança entre as cargas calculadas com um e outro método se deva ao fato de não se incluir a *keyword* DENSITY= "MP2" ou "CHELPG" na linha de instruções do Gaussian. Se isso não for feito, o cálculo das cargas será HF, independente de se estar usando MP2 para o cálculo da energia. Isto é uma coisa a ser investigada. Resultados preliminares do acetaldeído são mostrados na Tabela 5, em que as cargas foram calculadas numa rodada com MP2 sem inclusão dessa *keyword*. Como se observa, os resultados são muito semelhantes aos de HF da Tabela 1.

Outro resultado igualmente interessante que aflora dos valores mostrados na Tabelas 1 a 4, é que as cargas dos grupos aldeídico e amínico (metil para o caso do acetaldeído) praticamente se anulam, o que mostra que a aproximação por nós utilizada de fazer a carga líquida desses dois grupos igual a zero para cada um, é uma boa aproximação.

Além disto, observa-se que as cargas de Mulliken são menos suscetíveis às características particulares das moléculas, como isomeria e substituição por exemplo, que é o que se têm em nosso caso. Então as cargas sobre o hidrogênio e os grupos metil amínicos são, correspondentemente, as mesmas, independente do grupo ou átomo ser cis



ou trans e independente de haver um ou dois grupos iguais ligados ao nitrogênio. Isto a nosso ver é mais um ponto à favor das cargas de Mulliken, quando se adota a filosofia de transferência de parâmetros, o que resulta em um menor número de parâmetros representativos de um maior número de moléculas. Diferentemente, os valores obtidos de cargas CHELPG são afetados por estas características moleculares, o que deve estar mais próximo da realidade. Estes resultados parecem indicar que quando se necessitar reproduzir a carga molecular com mais precisão, a metodologia CHELPG é mais adequada. Mas, a se adotar cargas CHELPG, ter-se-iam conjuntos de parâmetros individuais para cada molécula, o que aumentaria enormemente o número de parâmetros moleculares com os quais se teria que lidar. Neste sentido portanto, a utilização de um ou outro conjunto de cargas depende dos objetivos e da filosofia do trabalho

Considerando isto, numa das ramificações do presente estudo, que será abordada futuramente, que é analisar o efeito da ressonância sobre as cargas dos átomos que participam da ligação peptídica, deverá se utilizar metodologia CHELPG. Relativamente a este ponto, e repetindo-se o que se verificou acima, os resultados apresentado nas Tabelas 2 a 4 mostram uma constância nas cargas de Mulliken do carbono e oxigênio carbonílicos, independente da substituição no nitrogênio, enquanto as cargas CHELPG são significativamente diferentes. Este é um assunto para ser explorado.

Outra questão que se levanta aqui e que igualmente deverá ser investigada, é o fato que à medida que se aumenta a substituição no nitrogênio, não se verifica aumento de carga resultante do chamado efeito doador de elétrons dos grupos metil. Pelo contrário, o nitrogênio se torna menos negativo com a substituição. Como se verifica um aumento de carga negativa no carbono da carbonila, pode estar acontecendo que este fato seja uma particularidade da ligação peptídica, cuja ressonância "joga" a carga para cima do carbono que é o átomo central. Nesta linha de pensamento será calculado brevemente a distribuição de carga das moléculas H_3N , H_2NCH_3 , $HN(CH_3)_2$ e $N(CH_3)_3$.

Para encerrar esta discussão, o C e O carbonílicos do aldeído apresentam carga intermediária à das amidas (é inclusive semelhante à da metilformamida). Parece então que o papel desempenhado pela ressonância peptídica sobre a carga desses átomos, se houver, deve ser razoavelmente complexo.

Como se disse anteriormente, este trabalho tem os seguintes objetivos: estudar as pontes de hidrogênio fracas nos líquidos acetaldeído, formamida e metilformamida e em suas soluções aquosas, as interações entre as pontes de hidrogênio fortes e fracas presentes nesses sistemas, e a influência da ressonância existente nas amidas nessas pontes de hidrogênio. Esse é, como se pode depreender, um trabalho longo, que não se restringe portanto, aos resultados aqui apresentados. Por outro lado, os próprios resultados aqui discutidos, ensejarão a realização de outros cálculos e simulações, os quais corroborarão ou refutarão as conclusões a que se chegar. Neste tópico, iniciar-se-á pela apresentação de resultados obtidos para o acetaldeído, cujas simulações foram realizadas a 18 °C e pressão ambiente, lembrando que esse líquido tem ponto de ebulição de 20 °C nessa pressão. As amidas já haviam sido investigadas e foram objeto de trabalho publicado anteriormente [13].

Para iniciar a simulação do acetaldeído, dispunha-se dos parâmetros de potencial all-atoms OPLS, os quais foram fornecidos por W. Jorgensen via email. Na época, a cerca de 1 ano atrás, o Dr. Jorgensen tinha um aluno terminando de parametrizar esse líquido, e tais parâmetros não haviam ainda sido publicados. Não é de conhecimento deste autor se tal já ocorreu. Tais parâmetros de potencial estão listados na Tabela 6. Na Tabela 7 estão apresentados os resultados termodinâmicos obtidos com esse conjunto de parâmetros. Como se pode observar, há um erro de aproximadamente 3% no valor de entalpia de vaporização fornecido por esse modelo, o que, para um estudo sobre pontes de hidrogênio fracas, pode ser um valor elevado, considerando-se as pequenas energias envolvidas nessas interações. Além disso, as cargas atribuídas aos átomos nesse modelo, parecem ter sido demasiadamente simplificadas, principalmente quando se atribui carga zero ao hidrogênio aldeídico (que é um critério que vem sendo adotado por Jorgensen ao longo do tempo de forma um tanto quanto indiscriminada), mas que para o caso específico de hidrogênios aldeídicos, parece não ser adequado [20]. Some-se a isto que tais cargas são totalmente discordantes das de análise populacional de Mulliken a nível MP2 (Tab. 1), o que, conforme se discutiu anteriormente, parece ser uma metodologia adequada para se estabelecer valores de densidade de cargas atômicas em moléculas. Note-se que a análise populacional de Mulliken, fornece praticamente o mesmo valor de carga para o hidrogênio



aldeídico (0,12) em todas as moléculas estudadas (Tabs. 1, 2, 3 e 4), o que concorda com os valores citados em outras fontes [20,59]. Tendo isto em mente, partiu-se para a parametrização de novos parâmetros de potencial para o acetaldeído, usando-se a distribuição de cargas de Mulliken com aproximação MP2 e os parâmetros de Lennard-Jones do modelo de Jorgensen como um conjunto de parâmetros-tentativa inicial. Novamente, como já anteriormente explicado, fez-se uma série de simulações usando Monte Carlo, nos quais os parâmetros de potencial são sucessivamente ajustados, no sentido de se obter os valores de densidade e entalpia de vaporização experimentais do líquido. Os parâmetros finais a que se chegou estão listados na Tabela 6, e os respectivos valores para as propriedades termodinâmicas calculadas estão na Tabela 7. Como se pode observar, houve uma melhora significativa no valor de entalpia de vaporização teórico do líquido, sem perda na qualidade das outras grandezas. Mas o que chama a atenção sobretudo é a praticamente insignificante mudança que se necessita fazer nos parâmetros Lennard-Jones do modelo de Jorgensen para o presente modelo, a despeito da enorme diferença que há nas cargas utilizadas em um e outro caso. Daí chega-se às seguintes conclusões, a primeira das quais continua sendo uma proposição. i) análise populacional de Mulliken a nível MP2, é um bom método para se obter valores de densidade de carga atômica para moléculas polares, neste campo de estudo. ii) Os parâmetros Lennard-Jones do modelo OPLS all-atom de ethanal, e provavelmente, por extensão, para todas as outras moléculas OPLS, estão muito bem parametrizados.

Na Tabela 6 estão listados ainda os parâmetros de potencial usados neste trabalho para a simulação das amidas, os quais foram relatados anteriormente [13]. Na Tabela 7 listam-se os valores de entalpia de vaporização e densidade obtidos em cada caso e os respectivos valores experimentais. Observa-se então que os modelos parametrizados por nós reproduzem satisfatoriamente a termodinâmica subjacente à entalpia de vaporização e densidade.

5.2 Correlação Radial de Pares e estruturas representativas dos líquidos puros

5.2.1 Acetaldeído

Embora o interesse do trabalho esteja centralizado nas correlações H/O, no caso do acetaldeído será feito um estudo mais completo, uma vez que esse líquido teve seus parâmetros de potencial otimizados presentemente. No caso dos outros líquidos serão selecionadas as correlações sítio-sítio que interessam para a discussão. Então, na Figura 2 se apresenta o $g(r)$ obtido para acetaldeído.

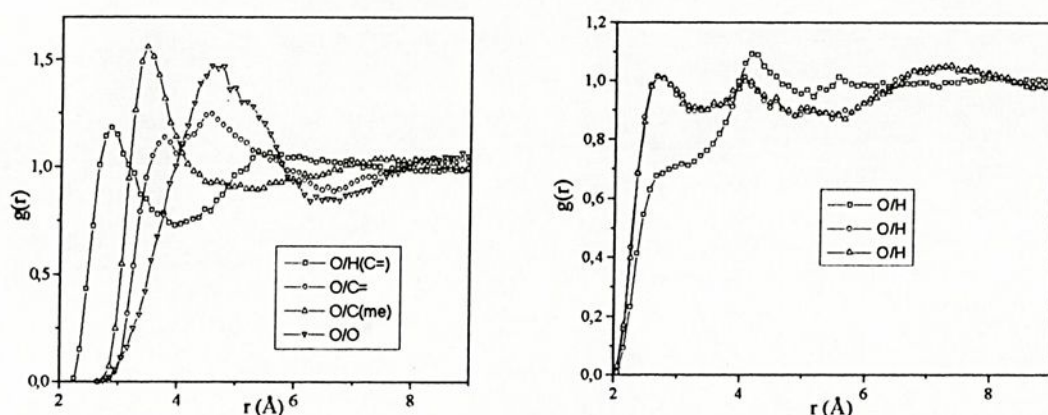


Figura 2. Funções de distribuição radial do acetaldeído

Merecem destaque os seguintes aspectos: i) o acetaldeído é um líquido relativamente bem estruturado, quando comparado a outros líquidos orgânicos como tetrahydrofurano e acetona, que apresentam todas as correlações entre sítios acima dos 3 Å e bandas largas [27] ao invés dos picos relativamente bem definidos que se verificam aqui; ii) a correlação de pares mais curta é a O/H(C=), em torno de 2,7 Å, que pode ser um indício de que essa correlação tenha papel importante na estrutura do líquido. A distância O–H é característica de ponte de hidrogênio fraca [20]; iii) a correlação entre oxigênio e os hidrogênios do grupo metil mostra que dois hidrogênios se correlacionam mais estreitamente com o oxigênio e à mesma distância (aproximadamente 2,7Å), enquanto um terceiro está mais distante e é mais label. Novamente neste caso a distância

de correlação é característica de ponte de hidrogênio fraca e esta interação pode também contribuir para a estabilidade estrutural do líquido, e iv) a correlação de pares O/C= apresenta dois picos, um que se sobrepõem ao da correlação O/C(Me), o que significa que o oxigênio de uma molécula está centrado aproximadamente na bissetriz do ângulo COC, onde os dois carbonos, um da carbonila e outro do grupo metil, são da molécula vizinha, e um segundo pico em aproximadamente 4,4 Å, o que significa que a estrutura do líquido não pode ser explicada em termos de uma correlação simples entre moléculas como num dímero. Chama atenção ainda alguns pequenos ombros que se verificam nas correlações O/H(C=) e O/O e a assimetria dos picos, que levam a supor a existência de uma estrutura um pouco mais complicada no líquido, de tal forma a se ter correlações entre os mesmos sítios em distâncias levemente diferentes, contribuindo para cada uma das funções de distribuição radial. Após um extensivo estudo de mecânica molecular [62] procurando otimizar uma estrutura que explicasse a distribuição de sítios mostrada nos gráfico das funções de distribuição radial, obteve-se o trímero mostrado na Figura 3. Nesse trímero as moléculas não são coplanares e é notável a concordância entre a geometria de tal trímero e o gráfico de $g(r)$ mostrado na Figura 2. Há pequenas variações nas distâncias entre os mesmos sítios considerando-se moléculas duas a duas (da ordem de décimo de ångstroms), o que explica o padrão verificado nos gráficos de distribuição radial, mas, o que chama realmente a atenção é o fato de haver, duas diferentes distâncias O/C=, que concordam perfeitamente com o gráfico de $g(r)$.

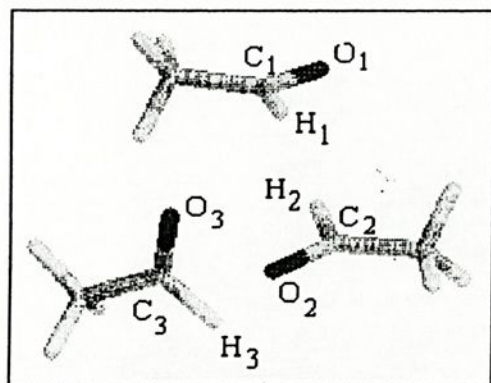


Figura 3. Trímero de etanal obtido a nível MM2, o qual reproduz as correlações verificadas no gráfico de distribuição radial de pares [63].

$$O_1C_2 \approx O_2C_3 \approx O_3C_1 \approx O_2C_1 \approx 3,3 \text{ Å}$$

$$O_1C_3 \approx O_3C_2 \approx 4,5 \text{ Å}$$

Como se vê na figura, o trímero é constituído de duas moléculas paralelas e uma antiparalela, o que origina as características acima apontadas das diversas curvas em $g(r)$. O cálculo do número de pontes de hidrogênio O/H(C=), obtido da integral sob a respectiva curva de distribuição radial fornece como resultado 2. Este resultado vem corroborar a proposição de que a espécie representativa da estrutura do líquido é um trímero. Na estrutura mostrada na Figura 3, portanto, cada oxigênio tem 2 átomos de hidrogênio aldeídico mais próximos e igualmente equidistantes. Desta otimização de geometria para a espécie preponderante na estrutura do líquido, fica claro que as pontes de hidrogênio entre os hidrogênios e oxigênios da carbonila é determinante na estabilidade da estrutura do aldeído em fase líquida. Este resultado caracterizaria bem a importância das pontes de hidrogênio fracas na estrutura de espécies associadas. Porém, para além deste resultado, a posição de uma molécula com respeito à outra, particularmente a distância entre o oxigênio e os dois carbonos de uma mesma molécula vizinha, mostra claramente que está havendo uma contribuição significativa da interação entre oxigênio e os hidrogênios do grupo metil, na estrutura dimérica. Isto mostra até onde pode ser importante a competitividade e a interação hidrogeniônica fraca numa estrutura protéica, desde que pontes de hidrogênio normais não estejam envolvidas. O trímero aqui mostrado para acetaldeído apresenta uma energia de estabilização a nível MM2 de -10,23 kcal/mol, enquanto essa mesma energia para o dímero representativo de acetona líquida é de -2,91 kcal/mol. Este assunto merecerá um estudo mais aprofundado, mas o trímero representativo de acetaldeído parece ser desproporcionalmente mais estável do que o dímero de acetona, o que concorda com o fato, discutido na introdução deste trabalho, que as pontes de hidrogênio são cooperativas, no sentido de que duas pontes de hidrogênio associadas, contribuem com uma energia de estabilização maior do que a soma de duas pontes de hidrogênio separadas [20]. Por outro lado, essa maior energia relativa do trímero poderia ser responsável pelo que já se mencionou acima sobre as diferenças nas funções de distribuição radial de um e outro líquidos. Uma estrutura mais estável deveria originar funções de distribuição radial características de maior rigidez. Só a título de complementar o estudo, é mostrado na Figura 4 o gráfico das funções de distribuição

radial do acetaldeído obtidas com o modelo OPLS. Como se pode observar, os gráficos de um e outro modelo são praticamente idênticos, o que significa que as diferentes cargas atômicas em um e outro modelo não alteram significativamente a estrutura do líquido. Isso chama a atenção levando-se em conta a enorme diferença entre as cargas atômicas nos dois modelos e suscita uma dúvida, considerando-se a semelhança entre os momentos de dipolo de um e outro modelo. Seria a estrutura de um líquido preponderantemente determinada pelas cargas de seus átomos ou pelo momento de dipolo de suas moléculas? A princípio parece razoável pensar que o dipolo molecular seja responsável pela orientação de uma molécula com respeito à outra e as cargas atômicas responsáveis pelas interações particulares entre átomos de moléculas vizinhas, já orientadas segundo o momento de dipolo. Parece-nos que seria importante a investigação teórica no sentido de quantificar o papel de um e outro aspecto na correlação molecular. Com respeito ainda a esta questão, observa-se que a correlação O/H(C=) é praticamente idêntica nos dois modelos, embora no modelo OPLS a carga sobre o H ligado à carbonila seja zero, o que inviabiliza falar em ponte de hidrogênio. Isto lembra uma questão levantada anteriormente por Luzar e outros [60], que mostraram que algumas correlações nos líquidos se devem a efeitos de empacotamento e não a interações específicas entre os sítios em questão. É possível então que uma fração significativa das interações tidas como pontes de hidrogênio fracas a partir de análise estatística de bancos de dados de estruturas proteicas [20], sejam meramente consequência destes efeitos. É preciso neste caso, mais do que meramente estabelecer a distância e os ângulos entre tais correlações, desenvolver metodologias que permitam calcular a energia de cada interação e qual a contribuição dessa energia para a estabilização estrutural.

Para encerrar, o gráfico das funções de distribuição radial do acetaldeído exemplifica bem a poderosa ferramenta que pode ser a simulação computacional no auxílio às técnicas de determinação da estrutura de líquidos e materiais amorfos, difração de raios-X e neutrons. Por essas técnicas seria impossível visualizar a correlação do oxigênio com os carbonos, uma vez que essas bandas seriam somadas e vistas como uma só.

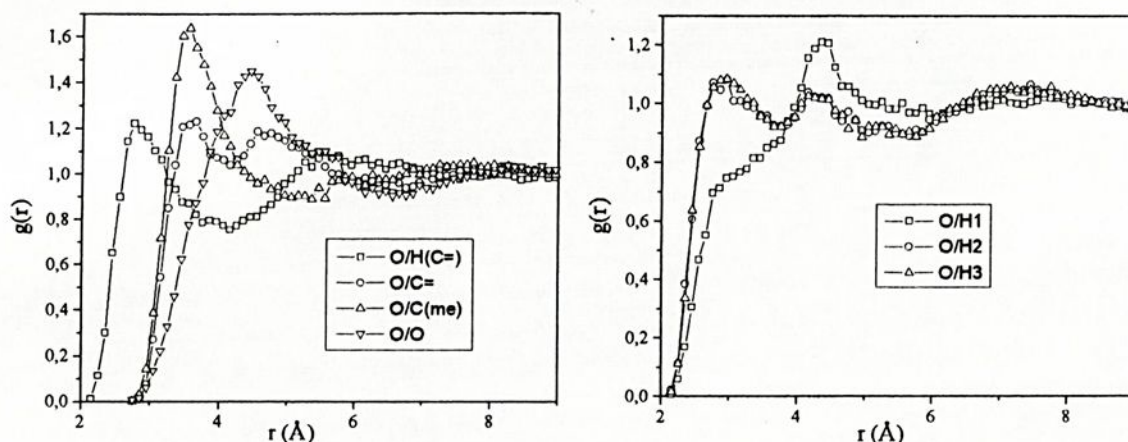


Figura 4. Funções de distribuição radial do acetaldéido obtidas com o modelo OPLS.

5.2.2 Formamida

Várias correlações radiais da formamida foram reportadas em trabalho anterior e a estrutura do líquido foi discutida com base nessas correlações [13]. Serão reproduzidas neste trabalho, algumas funções de distribuição, que são as que interessam para a discussão que se pretende aqui fazer (Figura 5).

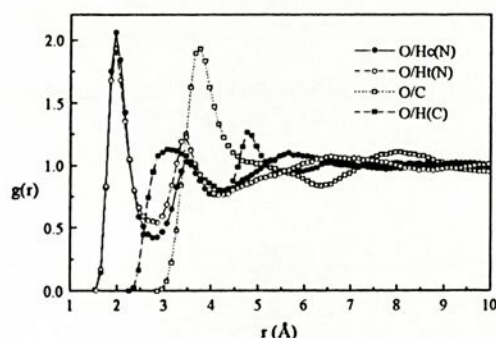


Figura 5. Funções de distribuição radial da formamida

Neste caso foram encontrados dois dímeros que concordam com a correlação sítio-sítio mostrada nos gráficos de $g(r)$ da Figura 5, os quais estão representados na Figura 6.

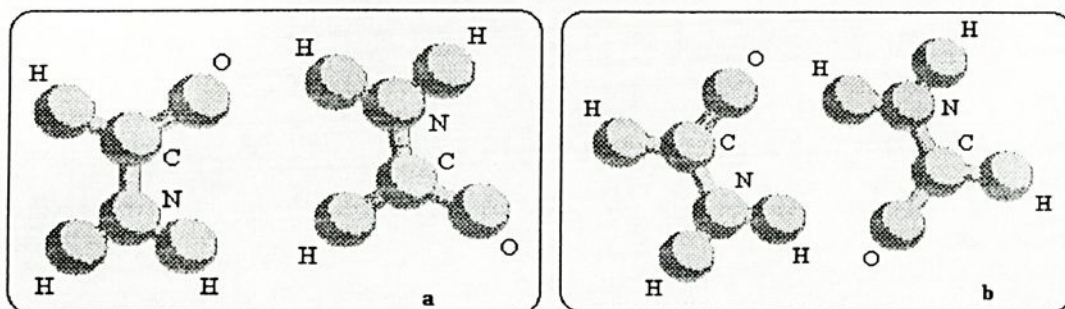


Figura 6. Dímeros otimizados a nível MM2 que reproduzem a distribuição radial de pares da formamida

No caso do dímero **b** mostrado, o mais estável, não há o que discutir em termos de pontes de hidrogênio fracas, uma vez que a interação se dá devido às duas pontes de hidrogênio tradicionais. Todavia, cabe aqui ressaltar que esse dímero é o responsável pelo pico $O/H(C=)$ em aproximadamente $4,8\text{\AA}$. Essa é a distância entre o oxigênio e o H aldeídico dessas duas moléculas. No caso do dímero **a**, tem-se uma ponte de hidrogênio normal, que é a principal responsável pela estabilidade do dímero e tem-se uma distância $O-H(C=)$ de aproximadamente $2,9\text{\AA}$, o que concorda com o pico mostrado no gráfico de $g(r)$. Ainda não foi possível verificar se essa correlação contribui ou não para a estabilidade desse dímero, mas é interessante notar que a os picos de correlação $O/H(N)$ relativos aos dois hidrogênios amínicos (cis e trans) são de mesma altura, o que indica que os dímeros **a** e **b** participam de forma igualitária da fase líquida. Como o dímero **b** é estabilizado por duas pontes de hidrogênio, pode-se pensar então que esteja havendo participação do hidrogênio da carbonila na estabilidade do dímero **a**, via ponte de hidrogênio fraca com o oxigênio, o que aumentaria a estabilidade desse dímero.

5.2.3 Metilformamida

Em primeiro lugar, diferentemente da formamida, a metilformamida apresenta um único dímero, mostrado na Figura 7, representativo da distribuição radial de pares.

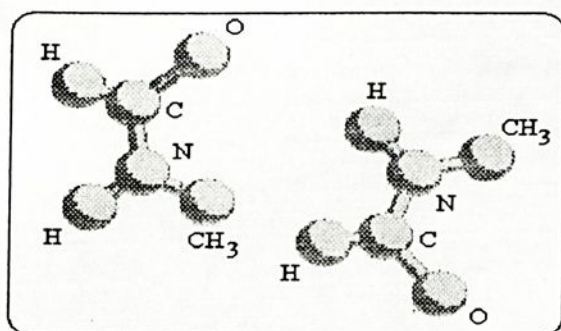


Figura 7. Dimero da metilformamida otimizado a nível MM2 que reproduz a distribuição radial de pares do líquido.

É interessante notar a posição do hidrogênio amínico neste dímero, colocado numa posição central no triângulo NCO da molécula da esquerda. Em consequência disto o hidrogênio amínico deve estar formando uma ponte de hidrogênio altamente estável. Esta interação necessita ser melhor investigada, o fato é que a metilformamida parece ser um líquido mais estruturado do que seus semelhantes formamida e dimetilformamida [61]. Convém lembrar aqui que enquanto formamida e dimetilformamida são miscíveis em água, metilformamida é apenas parcialmente solúvel. O gráfico da distribuição radial de pares é mostrado na Figura 8.

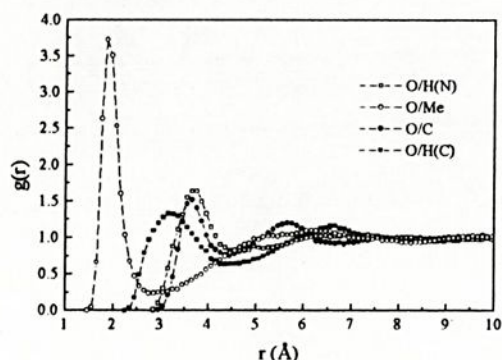


Figura 8. Funções de distribuição radial da metilformamida

É de se notar a alta definição do pico O/H em 1,9 Å, da ligação de hidrogênio normal, o que praticamente determina a posição das demais correlações moleculares. É interessante também comparar a amplitude desta interação com as correspondentes da formamida para se ter em mente a diferença que deve haver na energia de ligação. Esta verificação merecerá oportunamente investigação mais detalhada.

5.2.4 Dimetilformamida

O gráfico de $g(r)$ da DMF mostra que a correlação sítio-sítio mais curta é a O/H (hidrogênio da carbonila) em aproximadamente 2,9 Å (Figura 9). Trabalhos de difração de raios-X das amidas aqui estudadas [61], mostram que DMF é a mais desestruturada entre elas, não se verificando nenhum pico de difração, o que já não acontece com as outras duas. Em função disto, vale a pena investigar do ponto de vista teórico, a origem das forças que mantém unidas as moléculas em fase líquida. A interação O/H(C), deve desempenhar papel importante nessa estabilização.

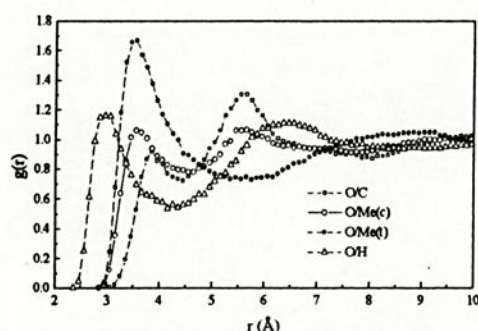


Figura 9. Funções de distribuição radial da DMF. Me(c) e Me(t) são os grupos metil cis e trans ao oxigênio respectivamente

Os dois picos correspondentes a interação O/Me atestam a existência de dois dímeros em solução. A forma desses dímeros como obtidos de otimização MM2 podem ser vistos na Figura 10

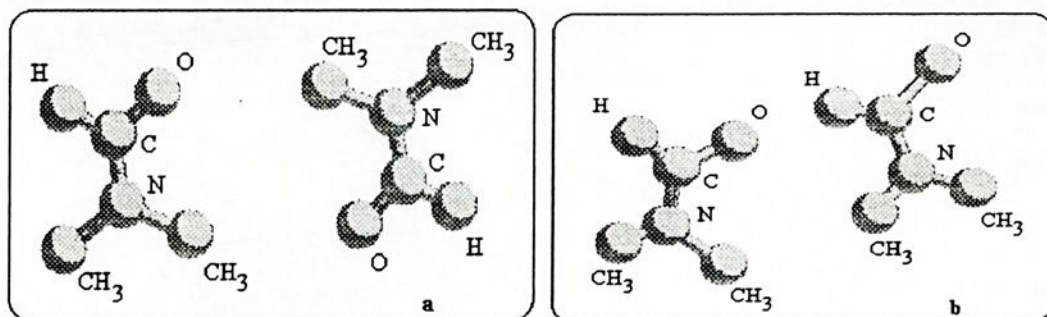


Figura 10. Os dois dímeros de DMF otimizados a nível MM2, que concordam com os gráficos de $g(r)$.

É interessante notar que um dos dímeros é paralelo enquanto o outro é antiparalelo. Neste, onde há a possibilidade de formação de ponte de hidrogênio fraca com

o hidrogênio carbonílico, é justamente no qual a interação O/Me é mais intensa. Isto parece claramente ser indício de contribuição da ponte de hidrogênio para a estabilidade do dímero. Como o modelo usado de dmf não explicita os hidrogênios dos grupos metil, não se pode estudar a influência destes hidrogênios para a estabilidade dos dímeros. Mas é bem provável que eles tenham papel importante, inclusive, sendo a distância O/Me em torno de 3,8 Å, a distância O/H(Me) deverá ser próxima de 2,8 Å, o que permitiria falar em ponte de hidrogênio fraca.

Do que se discutiu aqui, pode-se inferir que as pontes de hidrogênio fracas desempenham papel importante na estrutura dos líquidos acetaldeído, formamida e dimetilformamida. Esta interação tanto é importante para assegurar maior estabilidade das estruturas presentes, quanto influi na geometria das pontes de hidrogênio normal e de espécies associadas. Os exemplos aqui analisados confirmam que pontes de hidrogênio fracas tem papel importante na estrutura de macromoléculas biológicas, onde muitas resíduos possuem grupamentos semelhantes aos aqui estudados. Sobretudo pode-se inferir que, se nos casos estudados verifica-se influência das pontes de hidrogênio fracas na geometria e estabilidade das espécies em fase líquida, isso deve ser verdadeiro para outros sistemas onde seja possível a existência desse tipo de interação. Todavia, como os aspectos aqui analisados, o foram quase que exclusivamente baseado em simulação da fase líquida, muitas questões necessitam ser mais detalhadamente analisadas, para o que se necessita lançar mão de outras metodologias.



5.3 Correlação Radial de Pares e estruturas representativas das soluções

Os principais dados que se obtêm diretamente da simulação são os valores de densidade e entalpia de vaporização das misturas nas composições estudadas e os gráficos de distribuição radial de pares, além de dados sobre a energética de interação entre as moléculas constituintes do sistema. Com os resultados termodinâmicos pode-se fazer inferências sobre a interação intermolecular soluto-solvente e sobre mudanças nas interações intermoleculares de um dos líquidos devido a presença do outro. Exemplos desses estudos podem ser vistos em trabalhos anteriores [12,14,15]. Nesses trabalhos mostrou-se que a metodologia de simulação empregada por nós e os modelos moleculares teóricos desenvolvidos são suficientemente hábeis para fornecer teoricamente as curvas de entalpia de excesso de soluções aquosas e análise energética da interação molecular soluto-solvente em função da composição. Portanto, estudos semelhantes podem se fazer para as soluções amidas-água aqui estudadas, e, talvez mais interessante, para a solução acetaldeído-água, que, até onde foi possível verificar, não está estudada experimentalmente. Na Tabela 8 estão listados os dados de propriedades termodinâmicas obtidas neste trabalho e que são usados nesses estudos. Essa análise para os sistemas em questão será feita oportunamente.

Da distribuição radial de pares obtêm-se informações sobre as correlações sítio-sítio em particular e sobre a estrutura do líquido em geral. Iniciar-se-á pela análise dos resultados obtidos para as soluções aquosas das amidas. Isso possibilitará que se obtenha uma visão razoavelmente completa da forma da interação amida-água e das diferentes pontes de hidrogênio formadas, inclusive por comparação com outros resultados publicados [15,62]. Os resultados apresentados foram obtidos a partir da simulação das respectivas soluções a 50% tendo sido utilizado o modelo SPCE para a água [53], como já mencionado anteriormente.

5.3.1 A mistura formamida-água

Na Figura 11 apresenta-se o gráfico de várias funções de distribuição radial para a interação formamida-água. Os gráficos deixam claro a formação de 3 pontes de hidrogênio entre a amida e água, mas mostram também que há algumas diferenças entre essas pontes, que merecem ser analisadas. As distâncias de correlação



entre Of e Hw e entre Hc ou Ht e Ow e o cálculo do número de moléculas de água coordenadas a cada um desses sítios (calculado pela área sobre a respectiva curva) mostram claramente que cada um dos átomos da amida está coordenado a uma molécula de água. Tem-se então a formação de um tetrâmero, o que foi confirmado pela otimização a nível MM2 da estrutura que reproduziria a distribuição radial de pares mostrada na Figura 11. Essa estrutura é apresentada na Figura 12.

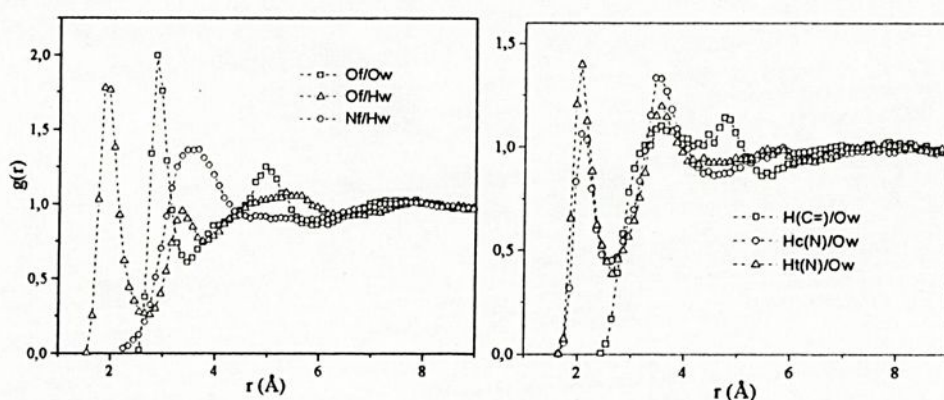


Figura 11. Correlação sítio-sítio na mistura formamida-água a 50%. Hc e Ht são os hidrogênio amínicos cis e trans ao oxigênio da amida, respectivamente.

As distâncias inter-atômicas obtidas na estrutura tetramérica mostrada, concordam perfeitamente com as posições dos picos das funções de distribuição radial da Figura 11. Ademais permite explicar porque as amplitudes dos picos das correlações Hc/Ow e Ht/Ow são diferentes. Esses picos são o resultado da correlação entre o hidrogênio amínico e oxigênios de moléculas diferentes. Como se pode observar, o primeiro pico Ht/Ow é mais alto do que o de Hc/Ow, enquanto no segundo pico verifica-se o contrário. Isso ocorre porque há uma competitividade do hidrogênio cis pelo oxigênio da molécula de água diretamente coordenada e pelo oxigênio da molécula de água coordenada ao oxigênio da carbonila. Então a correlação com a molécula de água diretamente solvatante é mais fraca relativamente ao hidrogênio trans. Por outro lado, o segundo pico da correlação é mais alto, porque o hidrogênio cis está rodeado por um número maior de oxigênios.

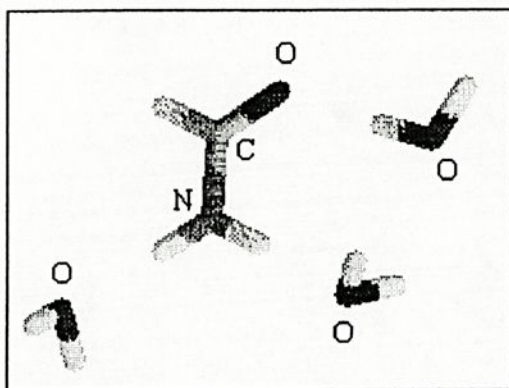


Figura 12. Complexo formamida-água que mostra a estrutura das espécies que geram a distribuição radial de pares

Esse tetrâmero esclarece ainda outros pontos observados nos gráficos de $g(r)$. Primeiro, os dois picos verificados no gráfico da correlação radial $H(C=)/O_w$, se devem ao fato de que há duas diferentes distâncias entre o hidrogênio aldeídico e o oxigênio das moléculas de água que solvatam a amida. Daqui decorre que não há formação de ponte de hidrogênio entre esse hidrogênio e água. Portanto, pontes de hidrogênio fracas podem contribuir para a estabilização da formamida pura, mas não desempenham nenhum papel na estabilização da solução aquosa.

Segundo, há uma concordância perfeita entre a estrutura desse complexo e o gráfico da correlação O_f/O_w , onde se verifica um pico em $2,8 \text{ \AA}$, um ombro em $3,7 \text{ \AA}$ e outro pico em $4,8 \text{ \AA}$. Essas distâncias são exatamente reproduzidas pelas distâncias entre o oxigênio da formamida e das 3 moléculas de água que a estão hidratando. O complexo mostrado reproduz ainda as distâncias observadas na função de distribuição N_f/O_w (não inserida no gráfico da Figura 11 para não sobrecarregar o gráfico), que, assim como a correlação O_f/O_w , mostra a distribuição dos 3 oxigênios vizinhos, em torno do nitrogênio. Para finalizar, os dados mostram que não há correlação entre o nitrogênio e hidrogênio de água, sendo a curva mostrada no gráfico da Figura 11, meramente consequência dos hidrogênios das 3 moléculas de água hidratantes da formamida a que se fez referência até agora. Quando da discussão dos resultados obtidos para a mistura metilformamida-água, se falará mais sobre esta correlação.

5.3.2 A mistura metilformamida-água

Na Figura 13 mostra-se o gráfico de várias funções de distribuição radial para a interação metilformamida-água.

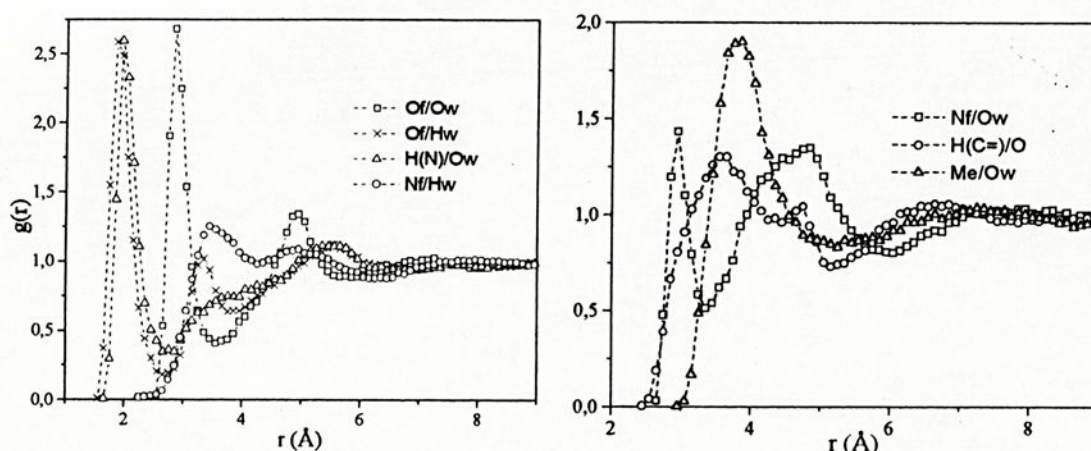


Figura 13. Algumas correlações metilformamida-água na mistura 50%.

Os gráficos de $g(r)$ mostram claramente a formação de duas pontes de hidrogênio entre a metilformamida e água. As posições relativas dos picos de correlação dos demais sítios e a integração das áreas sobre as curvas Of/Hw e H(N)/Ow, concordam em que duas moléculas de água distintas são responsáveis por essas duas pontes de hidrogênio, ou seja, há uma molécula de água coordenando o hidrogênio do grupo amina e outra coordenando o oxigênio da carbonila. Na Figura 14 é mostrado o trímero metilformamida-água que concorda com a distribuição radial de pares mostrada na Figura 13. Aqui, como nos caso anteriores, é expressiva a concordância entre as várias distâncias sítio-sítio nesse trímero e a localização dos picos de correlação sítio-sítio nos gráficos de $g(r)$. Deve-se novamente salientar que as correlações H(C=)/Ow e Nf/Hw observadas nos gráficos de distribuição radial, se devem às posições relativas das moléculas de água no complexo, e não a correlação propriamente dita entre esses sítios. É interessante mencionar que, tanto no estudo da solução aquosa de tetrametilurea [62], quanto no de DMF [15], foi verificada a completa ausência da correlação Nf/Hw, o que significa, por comparação com os resultados aqui obtidos, que essa falta de correlação se deve a ausência nessas duas moléculas, de hidrogênio ligado ao nitrogênio, por onde ocorre a formação de ponte de hidrogênio com água. Fica claro também, que a não existência de ponte de

hidrogênio entre nitrogênio e hidrogênio de água se deve as características particulares das amidas, e não a qualquer problema relacionado com impedimento estérico. Belletato e outros [62], atribuíram a ausência de ponte de hidrogênio entre N e água na solução aquosa de tetrametilureia à pequena carga do nitrogênio naquela amida ($q_N = -0,2961$), o que, segundo os autores, não seria suficiente para estabilizar a ponte com água na solução. Os resultados aqui apresentados mostram que essa explicação não é satisfatória no momento em que as cargas do nitrogênio em todas as amidas aqui simuladas são maiores que as do oxigênio carbonílico (Tabela 6), todavia não se verificam pontes de hidrogênio com o nitrogênio mas sim com o oxigênio. Por outro lado, como já acima mencionado, os presentes estudos com formamida e MF mostram que esse fato também não se deve a impedimento estérico, o que seria difícil de se esperar, pelo menos na formamida. Minimização de energia feita a nível MM2 com uma molécula de MF e quatro de água, igualmente não fornece qualquer complexo com água coordenando o nitrogênio via ponte N-HOH. Poder-se-ia concluir que a ausência de coordenação N/Hw seja consequência do fato de que os dois elétrons “livres” do nitrogênio estão sendo usados na ressonância existente nessas moléculas e não se encontram disponíveis para formação de ponte com o hidrogênio da água. Por outro lado, no oxigênio da carbonila, que também participa dessa coordenação, como se têm dois elétrons livres, o átomo mantém sua capacidade de coordenação. Obviamente, é surpreendente que a simulação MC consiga captar estes detalhes, uma vez que as cargas localizadas sobre os sítios são pontuais, de onde se conclui, que este assunto merece uma maior investigação tanto a nível de simulação propriamente dita quanto de metodologia, inclusive com a realização de cálculos de nível superior. De qualquer forma e para já, acreditamos que a falta de correlação verificada entre o nitrogênio e hidrogênio da água com a metodologia aqui empregada, se deve à repulsão entre o hidrogênio da água e os sítios ligados ao nitrogênio aldeídicos, esses sim, com uma acentuada carga positiva. No caso de tal correlação efetivamente não ocorrer, cabe a métodos de mais alto nível, descreverem por qual das duas razões acima apontadas, isso ocorre. Esses estudos serão feitos em breve, e estão relacionados com nosso trabalho de pos-doc.



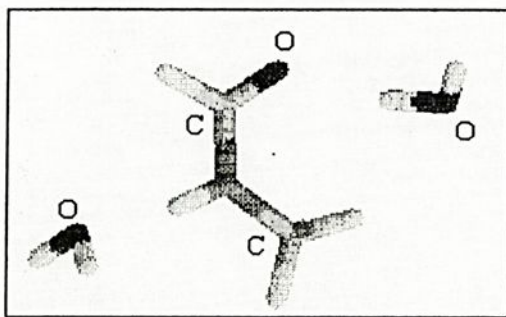


Figura 14. Complexo metilformamida-água que mostra a estrutura das espécies que geram a distribuição radial de pares

5.3.3 A mistura acetaldeído - água

Os gráficos de várias funções de distribuição radial para a interação metilformamida-água são vistos na Figura 15

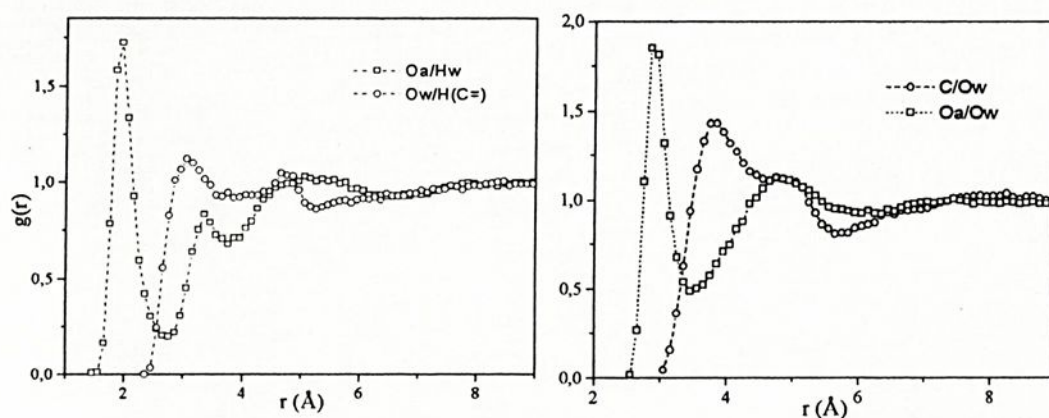


Figura 15. Gráficos para algumas correlações acetaldeído-água na mistura 50%.

Fica claro da observação dos gráficos mostrados a formação de uma ponte de hidrogênio entre o oxigênio da carbonila aldeídica e água. Indubitavelmente, essa ponte de hidrogênio será a responsável pela estabilidade da solução aquosa do aldeído. A questão agora reside na análise da interação, claramente notada no gráfico de $g(r)$, entre o hidrogênio ligado à carbonila e o oxigênio da água. Para fundamentar mais a discussão, é mostrado na Figura 16, a espécie otimizada a nível MM2, cuja geometria concorda com as distâncias de pico nos gráficos de $g(r)$ (Figura 15). Nessa espécie, a distância verificada entre o hidrogênio da carbonila e uma molécula de água que se encontra “coordenando” esse hidrogênio (molécula de água da esquerda), é de aproximadamente 2,9 Å, enquanto que a distância entre esse mesmo hidrogênio e o

oxigênio da molécula de água que está formando ponte com o oxigênio carbonílico (molécula de água da direita) é de aproximadamente 4,7 Å. Estes resultados deixam claro que uma molécula de água se encontra hidratando o acetaldeído por interação de seu oxigênio com o hidrogênio da carbonila. É interessante notar ainda, que, segundo a geometria da espécie mostrada na Figura a distância entre o oxigênio da carbonila e os hidrogênios da molécula de água coordenada ao hidrogênio da carbonila é de aproximadamente 4,9 Å, que é a região onde no gráfico de $g(r)$ se nota o início de uma banda de correlação Oa/Hw. Este é mais um dado indicando que a interação entre o hidrogênio carbonílico e a molécula de água da esquerda tem uma certa estabilidade, o que permite seguramente falar em ponte de hidrogênio fraca. A forma dessa banda no gráfico de $g(r)$ Oa/Hw mostra claramente a labilidade da molécula de água, isto é, enquanto o oxigênio coordenado ao hidrogênio carbonílico está relativamente fixo, o que é reflexo da interação entre os dois átomos em questão, a molécula de água em si realiza um movimento semelhante ao de um leque, o que gera a distribuição dos hidrogênios com respeito ao oxigênio carbonílico. Portanto, neste caso, parece não haver dúvida sobre o papel da ponte de hidrogênio fraca na estabilidade do aldeído em solução aquosa e na interação entre essa molécula e água.

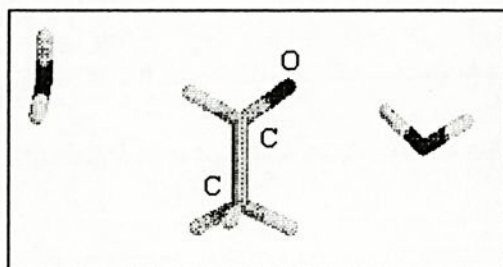


Figura 16. O trímero acetaldeído-água, que explica o gráfico de distribuição radial de pares para a correlação aldeído-água.

A energia de cada uma das pontes de hidrogênio do trímero mostrado na Figura 16 foi calculada por MM. Os valores obtidos foram de 3,89 kcal/mol para a ponte com o oxigênio da carbonila e de 1,85 kcal/mol para a ponte com o hidrogênio ligado à carbonila. Segundo Jeffrey e Saenger [39], a energia das pontes de hidrogênio normais variam entre 3 e 8 kcal/mol. Stillinger, em seu clássico artigo sobre estudos teóricos de água [69], mostra uma tabela com 6 valores de energia para a ponte de

hidrogênio em um dímero de água, calculada por métodos *ab initio* com diferentes conjuntos de bases. Os valores relatados variam entre 4,6 e 6,09 kcal/mol, dependendo do conjunto de bases utilizado (a citada tabela mostra um valor de 12.6 kcal/mol - o mais antigo de todos- obtido com uma base simples. Como esse valor se encontra desproporcionalmente fora do intervalo dos demais, o próprio Stillinger atribui essa discrepância à precariedade do cálculo feito em tal caso. Então vamos nos deter no intervalo acima citado). Os valores decrescem à medida que o conjunto de bases utilizado no cálculo se expande. Stillinger sugere que o valor de energia da ponte de hidrogênio deve tender para o valor limite de 4,5 kcal/mol com uma base muito expandida. Nesse sentido, o valor calculado para a ponte de hidrogênio com o oxigênio carbonílico neste trabalho parece bastante razoável. Por outro lado, de acordo com Steiner [20], as pontes de hidrogênio fracas apresentam uma energia de até 2 kcal/mol. Assim sendo, o valor acima citado para a energia da ponte de hidrogênio fraca aldeído-água, obtido por cálculo de MM, encontra-se na faixa esperada, e mais, sugere uma ponte de hidrogênio relativamente forte, o que concorda com a boa correlação entre as moléculas de aldeído e água verificada nos gráficos das funções de distribuição radial mostradas na Figura 15. Cálculos *ab initio* serão brevemente realizados, para comparação.

E por fim, chama-se a atenção para o fato de que o trímero mostrado na Figura 16, explica perfeitamente os gráficos das correlações sítio-sítio C/Ow e Oa/Ow mostradas na Figura 15.



6. CONCLUSÕES

Tendo em vista os resultados analisados no presente trabalho pode-se concluir que a metodologia utilizada apresenta bons recursos para fornecer informações relevantes no estudo da interação intermolecular do tipo pontes de hidrogênio, tanto das clássicas quanto das chamadas pontes de hidrogênio fracas. Cálculos de mais alto nível são desejados para detalhar o tipo de interações aqui estudadas. O modelo de 7 sítios parametrizado para acetaldeído mostrou-se hábil em reproduzir propriedades termodinâmicas e estruturais do líquido puro e em solução aquosa. Os resultados obtidos para essa substância e para as amidas indicam que o cálculo de cargas de Mulliken a nível MP2 é adequado como metodologia para se obter cargas pontuais para modelagem do potencial de interação coulômbico entre moléculas polares neutras. A análise dos resultados obtidos para os sistemas aqui estudados mostram que em se tratando de acetaldeído e DMF a pontes de hidrogênio $C=O\cdots H-C=$ devem desempenhar um papel de extrema importância na estabilidade do líquido. Por extensão, resíduos com esse tipo de estrutura podem ser responsáveis por importantes efeitos configuracionais em moléculas que os apresentem. Esse tipo de interação parece ser menos importante em formamida e inexistente em metilformamida. No caso da formamida, a geometria das pontes de hidrogênio normais parece ser distorcida pela presença de ponte de hidrogênio fraca. Os resultados deixam claro portanto, que pontes de hidrogênio fracas são menos importantes em sistemas que contenham pontes de hidrogênio tradicionais. Os resultados obtidos com acetaldeído mostraram ainda que efeitos de empacotamento podem desempenhar um papel importante nas diversas funções de distribuição radial de um líquido e portanto, cuidado deve ser tomado quando se fizer inferências sobre correlações sítio-sítio em líquidos, a partir das funções de distribuição radial. Não se pode esquecer do papel da interação dipolo-dipolo na estrutura molecular. Nas misturas formamida-água e metilformamida-água, não se observa qualquer sinal de que pontes de hidrogênio fracas contribuam para a interação intermolecular soluto- solvente. Já no caso do sistema aldeído-água, os resultados mostram que esse tipo de interação desempenha papel fundamental na interação entre essas moléculas. A comparação desses resultados com os obtidos para a interação DMF-água, anteriormente estudada, mostram que



resíduos como esses apresentam uma acentuada interação com água, via pontes de hidrogênio fracas. Estas interações devem ser fundamentais na hidratação e consequentemente nas propriedades das moléculas que apresentem resíduos com esse padrão, que sejam dependentes da interação com o solvente. Por outro lado, os resultados aqui mostrados indicam que o nitrogênio não contribui para a hidratação molecular em amino-ácidos. Cálculos de mecânica molecular fornecem resultados surpreendentemente interessantes quando se trata de estudar as espécies representativas da estrutura em fase líquida. Em particular, o campo de força MM mostra-se extremamente hábil para otimizar a geometria das espécies que representam a estrutura da substância na fase líquida, obtida pelas funções de distribuição radial. Observa-se que tal campo de potencial produz, em fase gasosa, complexos de interação intermolecular extremamente parecidos com os que ocorrem na fase líquida.



7. TABELAS

Tabela 1. Cargas fornecidas pelo programa Gaussian para **acetaldeído**

ÁTOMOS	RHF ($\mu = 3,0$)		MP2 ($\mu = 2,6$)	
	MULLIKEN	CHELPG	MULLIKEN	CHELPG
C	-0,586	-0,246	-0,576	-0,246
C=	0,332	0,619	0,239	0,528
O	-0,466	-0,538	-0,363	-0,449
H	0,188	0,065	0,188	0,070
H	0,208	0,084	0,199	0,079
H	0,188	0,065	0,188	0,070
H(C=)	0,137	-0,049	0,124	-0,052

E(HF)= -152,91; E(MP2) = -153,35

Tabela 2. Cargas fornecidas pelo programa Gaussian para **formamida**

ÁTOMOS	RHF ($\mu = 4,1$)		MP2 ($\mu = 3,8$)	
	MULLIKEN	CHELPG	MULLIKEN	CHELPG
N	-0,884	-0,995	-0,782	-0,908
C=	0,508	0,785	0,352	0,622
O	-0,555	-0,617	-0,440	-0,511
H(C=)	0,143	-0,020	0,126	-0,015
H(N)	0,389	0,408	0,369	0,396
H(N)	0,398	0,439	0,375	0,417

E(HF)= -168,93; E(MP2)= -169,39

Tabela 3. Cargas fornecidas pelo programa Gaussian para **metilformamida**

ÁTOMOS	RHF ($\mu = 4,3$)		MP2 ($\mu = 4,0$)	
	MULLIKEN	CHELPG	MULLIKEN	CHELPG
C=	0,521	0,590	0,349	0,418
N	-0,742	-0,551	-0,603	-0,460
C	-0,281	0,080	-0,325	0,029
O	-0,569	-0,602	-0,452	-0,494
H(C=)	0,137	0,032	0,119	0,038
H(N)	0,391	0,345	0,367	0,327
H	0,180	0,032	0,180	0,042
H	0,180	0,032	0,180	0,042
H	0,182	0,045	0,185	0,057

E(HF)= -207,96; E(MP2)= -208,55

Tabela 4. Cargas fornecidas pelo programa Gaussian para dimetilformamida

ÁTOMOS	RHF ($\mu = 4,1$)		MP2 ($\mu = 3,9$)	
	MULLIKEN	CHELPG	MULLIKEN	CHELPG
C=	0,522	0,483	0,337	0,322
N	-0,606	-0,154	-0,427	-0,083
C	-0,281	-0,188	-0,329	-0,208
C	-0,294	-0,037	-0,342	-0,067
O	-0,579	-0,574	-0,459	-0,469
H(C=)	0,142	0,030	0,123	0,033
H	0,175	0,089	0,175	0,093
H	0,181	0,085	0,183	0,092
H	0,175	0,089	0,175	0,093
H	0,167	0,043	0,170	0,052
H	0,167	0,043	0,170	0,052
H	0,232	0,089	0,223	0,088

E(HF)= -246,99; E(MP2)= -247,72

Tabela 5. Cargas fornecidas pelo programa Gaussian para acetaldeído (cálculo MP2 sem keyword *DENSITY*)

ÁTOMOS	MP2 ($\mu = 3,2$)	
	MULLIKAN	CHELPG
C	-0,582	-0,224
C=	0,335	0,589
O	-0,490	-0,538
H	0,190	0,063
H	0,210	0,082
H	0,190	0,064
H(C=)	0,146	-0,035

E(HF)= -152,91; E(MP2) = -153,35

Tabela 6. Parâmetros de potencial para os líquidos simulados neste trabalho

sites	q	$\sigma(\text{\AA})$	$\epsilon(\text{kcal/mol})$
ACETALDEÍDO OPLS ($\mu=2,8$)			
C=	0,45	3,75	0,105
O	-0,45	2,96	0,210
H(C=)	0,00	2,42	0,015
C	-0,18	3,50	0,066
H	0,06	2,42	0,015
ACETALDEÍDO ESTE TRABALHO ($\mu=2,6$)			
C=	0,24	3,75	0,105
O	-0,36	2,90	0,210
H(C=)	0,12	2,75	0,015
C	-0,54	3,50	0,058
H	0,18	2,31	0,010
FORMAMIDA ($\mu=4,2$)			
C=	0,34	3,75	0,105
O	-0,46	2,96	0,21
H(C=)	0,12	2,75	0,038
N	-0,830	3,25	0,17
H(N)	0,415	0,00	0,00
METILFORMAMIDA ($\mu=4,3$)			
C=	0,34	3,75	0,105
O	-0,46	2,96	0,21
H(C=)	0,12	2,75	0,038
N	-0,70	3,25	0,14
Me	0,285	3,80	0,13
H(N)	0,415	0,00	0,00
DIMETILFORMAMIDA ($\mu=4,4$)			
C=	0,34	3,75	0,105
O	-0,46	2,96	0,21
H(C=)	0,12	2,75	0,038
N	-0,57	3,25	0,14
Me	0,285	3,80	0,13

Tabela 7 - Valores de Propriedades termodinâmicas dos líquidos puros este trabalho

<i>líquido</i>	ΔH_v	$-E_i$	$\bar{V}$	ρ
acetald OPLS	6,36±0,01	5,78±0,01	57,03±0,08	0,77±0,005
acetald este trab.	6,16±0,02	5,58±0,02	57,42±0,09	0,77±0,003
formamida	14,79±0,02	14,20±0,02	40,93±0,08	1,10±0,004
MF	13,77±0,02	13,18±0,02	59,32±0,08	1,00±0,003
DMF	11,40±0,03	10,81±0,03	78,54±0,09	0,93±0,003
valores experimentais [58]				
<i>líquido</i>	ΔH_v	$-E_i$	$\bar{V}$	ρ
acetaldéido	6,16	5,58	56,23	0,78
formamida	14,70	14,11	39,89	1,13
MF	13,77	13,18	57,56	1,02
DMF	11,10	10,51	77,66	0,94

Calor de vaporização (ΔH_v) e energia interna (E_i) em kcal/mol,

calor de vaporização é estimado por $\Delta H_v = -E_i + RT$

volume molar ($\bar{V}$) em cm³/mol (o valor experimental é calculado com a densidade e massa molar)

densidade (ρ) em g/cm³

Tabela 8. Valores de Propriedades Termodinâmicas das Soluções em Função da Fração Molar da Água

χ_{H_2O}	ΔH_V	$-E_i$	$\bar{V}$	ρ
Formamida				
0,1	14,37±0,01	13,78±0,01	38,63±0,02	1,095±0,001
0,2	13,99±0,01	13,40±0,01	36,04±0,02	1,099±0,001
0,3	13,40±0,01	12,81±0,01	31,44±0,03	1,090±0,001
0,4	12,94±0,01	12,35±0,01	31,44±0,03	1,088±0,001
0,5	12,44±0,01	11,85±0,01	29,33±0,03	1,074±0,001
0,6	12,00±0,01	11,41±0,01	26,85±0,02	1,072±0,001
0,7	11,57±0,01	10,98±0,01	24,67±0,02	1,058±0,001
0,8	11,20±0,01	10,61±0,01	22,41±0,02	1,044±0,001
0,9	10,80±0,01	10,21±0,01	20,20±0,01	1,024±0,001
Metilformamida				
0,1	13,82±0,01	13,23±0,01	56,62±0,03	0,972±0,001
0,2	13,55±0,01	12,96±0,01	53,58±0,02	0,973±0,001
0,3	13,10±0,01	12,51±0,01	48,70±0,02	0,975±0,001
0,4	12,70±0,01	12,11±0,01	44,75±0,03	0,978±0,001
0,5	12,31±0,01	11,72±0,01	39,06±0,03	0,985±0,001
0,6	11,92±0,01	11,33±0,01	34,87±0,03	0,986±0,001
0,7	11,53±0,01	10,94±0,01	30,50±0,02	0,993±0,001
0,8	11,19±0,01	10,60±0,01	26,17±0,02	1,001±0,001
0,9	10,86±0,01	10,27±0,01	21,97±0,02	1,005±0,001

Tabela 8. continuação

	Acetaldeído			
0,1	7,42 ±0,03	6,83±0,03	52,56±0,21	0,79±0,01
0,2	7,71± 0,02	7,12±0,02	48,69±0,17	0,80±0,01
0,3	8,01± 0,02	7,42±0,02	44,82±0,11	0,81±0,01
0,4	8,37± 0,01	7,78±0,01	40,77±0,09	0,83±0,01
0,5	8,73±0,01	8,14±0,01	36,73±0,07	0,84±0,01
0,6	9,21±0,02	8,62±0,02	32,27±0,07	0,88±0,01
0,7	9,96±0,01	9,37±0,01	27,76±0,04	0,93±0,01
0,8	10,35±0,01	9,78±0,01	24,11±0,04	0,96±0,01
0,9	10,79±0,01	10,11±0,01	20,79±0,03	0,99±0,01

8. REFERÊNCIAS

1. Ver os artigos de Finley, Luzar, Soper e outros em *Faraday Discussions*, **103** (1996) e referências lá citadas.
2. Caccamo, C., Hansen, J-P., Stell, G., *New Approaches to Problems in Liquid State Theory*; NATO ASI series, Series C: Mathematical and Physics Sciences, vol. 529, Kluwer Academic (1999)
3. Strastsma, T. P., McCammon, J. A., *Annu. Rev. Phys. Chem.*, **43** (1992) 407
4. E. R. Davidson, *Chem. Rev.*, **93** (1993) 2337; C. E. Dykstra, idem 2339; R. S. Berry, idem 2379
5. B. M. Ladanyi, M. S. Skaf, *Ann. Rev. Phys. Chem.*, **44** (1993) 335
6. Y. M. Kessler, Y. P. Puhovski, M. G. Kiselev, I. I. Vaisman, *Chemistry of NonAqueous Solutions: Current Progress*, G. Mamantov e A. I. Popov eds, VCH, N.Y. (1994) 307
7. Metropolis, N. A., Rosenbluth, W., Rosenbluth, M. N., Teller, A. H.; Teller, E.; *J. Chem. Phys.*, **21** (1953) 108
8. Alder, B.J., Wainwright, T.E., *J. Chem. Phys.* **31** (1959) 459; idem, (1960) 33, 1439
9. Ben Naim, A., *Water and Aqueous Solution*, Plenum Press, New York, (1986)
10. Hansen, J., McDonald, I. R., *Theory of Simple Liquids*, Academic Press, London, (1986)
11. Teixeira-Dias, J.J. (ed.), *Molecular Liquids: New Perspectives in Physics and Chemistry*, NATO ASI Series, Kluwer Academic, (1991)
12. Freitas L.C.G., Cordeiro, J.M.M., *J. Mol. Struct. (Theochem)*, **333** (1995) 189
13. Cordeiro, J.M.M., *J. Int. Quant. Chem*, **65** (1997) 709
14. Freitas, L.C.G., Cordeiro J.M.M., Garbujo, F.L.L., *J. Mol. Liq.*, **79** (1999) 1
15. Cordeiro, J.M.M., Freitas, L.C.G., *Z. Naturforsch*, **54a** (1999) 110
16. Parker, A.J., *Chem. Rev.*, **69** (1969) 1; Abraham, M.H., *Prog. Phys. Org. Chem.*, **11** (1974) 185
17. Steiner, T., *Chem. Commun.*, (1999) 313
18. Steiner, T., Desiraju, G.R., *Chem. Commun.*, (1998) 891
19. Steiner, T., Lutz, B., van der Maas, J., Schreurs, A.M.M., Kroon, J., Tamm, M., *Chem. Commun.*, (1998) 171
20. Steiner, T., *Chem. Commun.*, (1997) 727



21. Desiraju, G.R., *Acc. Chem. Res.*, **24** (1991) 290; **29** (1996) 441
22. Desiraju, G.R., *Angew. Chem. Int. Ed. Engl.*, **34** (1995) 2311
23. Hill, T.E. *An Introduction to Statistical Thermodynamics*, Dover, New York, (1986)
24. Ripley, B.D., *Stochastic Simulation*, Wiley, New York, (1987)
25. Allen, M., Tildesley, D.J., *Computer Simulations of Liquids*, Oxford University Press, (1987)
26. Hermann, D.W., *Computer Simulation Methods*, Springer-Verlag, Berlin (1986)
27. Cordeiro, J.M.M., Tese de Doutorado, UFSCar, São Carlos, (1994)
28. Aqvist, J., Hansson, T., *J. Phys. Chem. B*, **102** (1998) 3837
29. Andrea, T.A., Swope, W. C., Andersen, H. C., *J. Chem. Phys.* (1983) 4579
30. Berkovitch-Yellin, Z., Leiserowitz, L., *Acta Cryst.*, **B40** (1984) 159
31. Chao, I., Chen, J.C., *Angew. Chem.*, **108** (1996) 200; Müller, G., Lutz, M. Harder, S., *Acta Cryst.*, **B52** (1996) 1024
32. Steiner, T., Saenger, J., *J. Chem. Soc., Chem. Commun.*, (1995) 2087; Braga, D. Grepioni, F. Byrne, J.J., Wolf, A., *J. Chem. Soc., Chem. Commun.*, (1995) 125; Behrens, P., van de Goor, G., Freyhardt, C.C., *Angew. Chem., Int. Ed. Engl.*, **34** (1995) 2895
33. Derewenda, Z.S., Derewenda, U., Kobos, P.M., *J. Mol. Biol.*, **241** (1994) 83; Derewenda, Z.S., Lee, L., Derewenda, U., *J. Mol. Biol.*, **252** (1995) 248, Leonard, G.A., McAuley-Hecht, K., Brown, T., Hunter, W.N., *Acta Cryst.*, **D51** (1995) 136; Wahl, M.C., Rao, S.T., Sundaralingam, M., *Nature Struct. Biol.* **3** (1996) 24
34. Allerhand, A., Schleyer, P. von P., *J. Am. Chem. Soc.*, **85** (1963) 1715
35. Sutor, D.J., *Nature*, **195** (1962) 68; *J. Chem. Soc.*, (1963) 1105
36. Pedireddi, V. R., Desiraju, G. R., *J. Chem. Soc., Chem. Commun.*, (1992) 988
37. T. Steiner, *J. Chem. Soc. Chem. Commun.*, (1994) 2341
38. Desiraju, G. R., *J. Chem. Soc. Chem. Commun.*, (1990) 454
39. Jeffrey, G.A., Saenger, W., *Hydrogen Bonding in Biological Structures*, Springer, Berlin, (1991)
40. Turi, L., Dannenberg, J.J., *J. Phys. Chem.*, **97** (1993) 7899; Novoa, J.J., Tarron, B., Whangbo, M.H., Williams, J.M., *J. Chem. Phys.*, **95** (1991) 5179; Mourik, T. van, van Duijneveldt, F.B., *J. Mol. Struct. (Theochem)* **341** (1995) 63

41. Lutz, B., van der Maas, J., Kanters, J.A., *J. Mol. Struct.*, **325** (1994) 203; Steiner, T., *J. Chem. Soc., Perkin Trans.*, **2** (1995) 1315; Allen, F.H., Howard, J.A.K., Hoy, V.J., Desiraju, G.R., Reddy, D.S., Wilson, C.C., *J. Am. Chem. Soc.*, **118** (1996) 4081; Brown, I.D., *Acta Cryst.* **B48** (1992) 553
42. Taylor, R., Kennard, O., *J. Am. Chem. Soc.*, **104** (1982) 5063
43. Steiner, T., Saenger, W., *J. Am. Chem. Soc.*, **114** (1992) 10146
44. Taylor, R., Kennard, O., Versichel, W., *J. Am. Chem. Soc.*, **105** (1983) 5761; Murray-Rust, P., Glusker, J.P., *J. Am. Chem. Soc.*, **106** (1984) 1018
45. Steiner, T., Kanters, J.A., Kroon, J., *Chem. Commun.* (1996) 1277
46. Gilli, G., Bellucci, F., Ferreti, V., Bertolasi, V., *J. Am. Chem. Soc.*, **111** (1989) 1023; Bertolasi, V., Gilli, P., Ferreti, V., Gilli, G., *Acta Cryst.*, **B51** (1995) 1004
47. Sharma, C.V., Desiraju, G.R., *J. Chem. Soc., Perkin Trans.*, **2** (1994) 2345
48. Steiner, T., Saenger, W., *J. Am. Chem. Soc.*, **115** (1993) 4540; Steiner, T., Saenger, W., *Acta Cryst.* **D49** (1993) 592; Steiner, T., *Acta Cryst.* **D51** (1995) 93
49. Steiner, T., Koellner, G., Gessler, K., Saenger, W., *J. Chem. Soc. Chem. Commun.*, (1995) 511
50. Freitas, L.C.G., programa DIADORIM, Dept. de Química, UFSCar (1998)
51. Freitas, L.C.G., programa CMIX, Dept. de Química, UFSCar (1991)
52. Israelachvili, J.N., *Intermolecular and Surface Forces*, Academic Press, London (1985)
53. Berendsen, H.J.C., Grigera, J.R., Straatsma, T.P., *J. Phys. Chem.*, **91** (1987) 6269
54. Cordeiro, J.M.M., *Z. Naturforsch.*, **54a** (1999) 311
55. Carlson, H.A., Nguyen, T.B., Orosco, M., Jorgensen, W.L., *J. Comp. Chem.*, **14** (1993) 1240
56. Gaussian 94, Frish, M.J. e uma série de autores, Gaussian Inc, Pittsburg, PA 1995
57. Merz Jr, K.M., *J. Comp. Chem.*, **13** (1992) 749; Luque, F.J., Orosco, M., Illas, F., Rubio, J., *J. Am. Chem. Soc.*, **113** (1991) 5203
58. Lide, D.R. ed.-in-chief, *Handbook of Chemistry and Physics*, 76ª edição, CRC Press Inc, (1995)
59. Hagler, A.T., Huler, E., Lifson, S., *J. Am. Chem. Soc.*, **96** (1975) 5319
60. Luzar, A., Soper, A.K., Chandler, D., *J. Chem. Phys.*, **99** (1993) 6836



61. Ohtaki, H., Itoh, S., Rode, B.M., *Bull. Chem. Soc. Jap.*, **59** (1986) 271; Radnai, T., Itoh, S., Ohtaki, H., *Bull. Chem. Soc. Jap.*, **61** (1988) 3845; Ohtaki, H., Funaki, A., Rode, B.M., Reibnegger, G.J., *Bull. Chem. Soc. Jap.*, **56** (1983) 2116; Kálmán, E., Serke, I., Pálinkás, G., Zeidler, M.D., Wiesmann, F.J., Bertagnolli, H., Chieux, P., *Z. Naturforsch.* **38a** (1983) 231; Schoester, P.C., Zeidler, M.D., Radnai, T., Bopp, P.A., *Z. Naturforsch.* **50a** (1995) 38; Wiesmann, F.-J., Zeidler, M.D., Bertagnolli, H., Chieux, P., *Mol. Phys.* **57** (1986) 275
62. Todas as espécies otimizadas a nível MM2, mostradas neste trabalho, o foram com o programa PCModel, Serena Software, Bloomington, USA.
63. As figuras de moléculas foram obtidas com o programa RASMOL
64. Belletato, P., Freitas, L.C.G., Arêas, E.P.G., Santos, P.S., *Phys. Chem. Chem. Phys.*, **1** (1999) 4769
65. Stillinger, F.H., *Adv. Chem. Phys.*, **31** (1975) 1



9. PROPOSTA DE TRABALHO PARA OS PRÓXIMOS 3 ANOS:

INVESTIGAÇÃO DA CATÁLISE ENZIMÁTICA POR SIMULAÇÃO DE MECANISMOS E DINÂMICA DE REAÇÃO

9.1 Introdução. Enzimas são catalisadores altamente específicos e eficientes, que desempenham papel fundamental na maioria das reações bioquímicas. Em função disto é de enorme importância a compreensão detalhada do processo de catálise, o que será vital no desenvolvimento de drogas inibidoras, na compreensão de mutações e no *design* de enzimas catalisadoras para aplicações práticas, só para citar alguns casos. Está longe ainda de ser atingido o grau de compreensão necessário para isso, a despeito da grande quantidade de investigação experimental e teórica já realizada [1-6]. Certos aspectos do problema como o detalhamento do mecanismo da reação e as diversas contribuições energéticas ao processo de catálise são de difícil abordagem experimental, enquanto podem ser apreendidos com certa facilidade por técnicas de simulação computacional [2-6]. Resultados teóricos podem complementar dados experimentais de estrutura, cinética e espectroscopia, e assim, responder questões que são difíceis de resolver experimentalmente. Por exemplo, o estado de transição e sua estabilidade é central na teoria de catálise. No entanto é difícil de estudar diretamente por técnicas experimentais, devido ao seu pequeníssimo tempo de vida, embora possa ser modelado e analisado relativamente bem usando-se técnicas calculacionais. Simulações podem ainda ser usadas para calcular as diversas contribuições energéticas para a estabilização do estado de transição, bem como identificar efeitos relevantes na aceleração da velocidade [2-6]. Por outro lado, calculada a geometria do estado de transição, estados de transição análogos podem ser desenhados. Estas moléculas podem então se ligar à enzima com altíssima afinidade, tornando-se inibidores extremamente eficientes ou compostos de teste úteis no desenho de novas drogas [23]. Recentemente, dinâmica molecular clássica tem contribuído para demonstrar a natureza dinâmica de proteínas e outras macromoléculas biológicas [7]. Está agora claro que enzimas sofrem mudanças dinâmicas complexas, e que, em alguns casos, a dinâmica está diretamente relacionada com a função [8,9]. Simulações podem prover um quadro detalhado da variabilidade



configuracional e interações em complexos enzimáticos, o que é essencial para aprofundar a compreensão. Propõem-se aqui estudar mecanismos de reações enzimáticas em detalhes e compará-los com reações análogas não catalisadas em solução, para contribuir na compreensão do mecanismo da catálise enzimática. As enzimas que serão estudadas representam importante classe de reações bioquímicas. Para tanto utilizar-se-ão métodos de simulação que incluem sofisticadas técnicas de combinação química quântica/mecânica molecular (QM/MM) que são o foco de interesse atualmente. Estes métodos permitem a simulação de reações químicas em fase condensada e podem fornecer interessantes resultados. Além disto, simulações de dinâmica molecular serão realizadas usando técnicas clássicas para investigar solvatação e comportamento conformacional. Será feito uma análise detalhada de pequenas moléculas, isoladas e em solução, e os resultados comparados com resultados experimentais. Este conhecimento é essencial para uma completa descrição do problema. O trabalho também será útil como uma forma de testar e desenvolver métodos que venham a ser aplicados no estudo de outras reações enzimáticas o que poderá vir a ser de grande importância para o futuro da bioquímica e da pesquisa biotecnológica.

9.2 Teoria. É fundamental em bioquímica entender detalhadamente o processo pelo qual as enzimas aumentam a velocidade das reações. Todavia ainda falta muito para ser atingido uma compreensão satisfatória desse fenômeno, a despeito do enorme esforço científico que tem sido feito. Uma boa compreensão dos princípios da catálise enzimática será essencial para o desenvolvimento de proteínas catalisadoras para aplicações práticas [24]. É difícil diferenciar experimentalmente entre diferentes mecanismos propostos ou avaliar a importância relativa de fatores apresentados como sendo importantes na catálise. Entre esses fatores pode-se citar estiramento de reagentes, redução de energias de solvatação, efeitos dinâmicos, de-solvatação, efeitos estéricos e eletrônicos, interação eletrostática e pontes de hidrogênio [1,2,22,]. Um importante objetivo de investigações teóricas é avaliar o papel de tais contribuições, para elucidar os princípios básicos pelos quais se dá a atividade catalítica enzimática. Com respeito às pontes de hidrogênio em particular, vimos já à dois anos investigando o papel das pontes de hidrogênio na conformação de moléculas modelos de peptídeos [11,12].



Para descrever uma reação catalizada enzimaticamente, seu mecanismo deve ser claramente estabelecido, e os intermediários devem ser identificados. É necessário conhecer o papel de resíduos específicos na estabilização do complexo de transição e na desestabilização dos reagentes. A contribuição das interações eletrostáticas, pontes de hidrogênio, polarização do substrato e possível estiramento, devem ser avaliados. Além disso, a contribuição de mudanças conformacionais e efeitos dinâmicos no aumento da velocidade deve ser determinado. Somente quando a descrição de uma reação catalizada enzimaticamente possa ser fornecida no mesmo nível de detalhamento de uma reação simples, tal como a reação de troca $H + H_2$, estará o problema da catálise enzimática resolvido em qualquer caso dado. Tal descrição de uma reação num sistema mesoscópico, como é uma enzima em solução, é um desafio formidável e requer contribuições tanto de trabalhos experimentais quanto teóricos. Resultados no nível requerido de detalhamento podem ser obtidos teoricamente. Assim, por exemplo, o perfil de energia para a reação enzimática pode ser calculado e o papel de varias interações como pontes de hidrogênio e estabilização eletrostática podem ser quantificados. Cálculos computacionais em reações catalizadas enzimaticamente progrediram ao estágio em que podem ser úteis como fonte deste tipo de informação [2-6, 25].

Enzimas sofrem complexas alterações conformacionais durante sua operação [1] e para que se compreenda perfeitamente o funcionamento enzimático a dinâmica molecular deve ser levada em consideração. Mostrou-se que a dinâmica enzimática é central na atividade de grande número de casos [8]. Flutuações estruturais afetam não só a coordenação, com substratos ou cofatores por exemplo, mas também a velocidade da reação na enzima, efeitos que são ainda pouco compreendidos. Outro aspecto vital são os efeitos de solvente. A maioria das enzimas atuam em meio aquoso e as simulações, para serem realistas devem incluir efeitos de solvente. Modelagem de pequenas moléculas representando grupos de esqueleto proteico em meio líquido permite o estudo detalhado de efeitos de solvatação além de teste de funções de potencial [11,12]. Simulações de dinâmica molecular permitem analisar em detalhes movimentos proteicos. Funções de potencial bem comportadas para proteínas e outras moléculas biológicas tem sido desenvolvidas através de processo intensivo de testes e parametrização [26] e são extensivamente usadas. Com os

computadores modernos, trajetórias no intervalo de nanosegundos podem ser geradas [18]. Como em toda simulação de proteínas, se requer uma estrutura de boa qualidade como ponto de partida. As simulações podem então acrescentar uma dimensão extra à solução do problema, por descrever detalhadamente a dinâmica protéica. Uma aplicação particularmente excitante é a combinação de simulação e cristalografia para estudar mecanismos de reações enzimáticas. Simulações de dinâmica molecular são um meio de validar as estruturas e podem fornecer modelos para posterior refinamento. Esta combinação é particularmente útil para estudos de intermediários reacionais de vida curta usando técnicas de Laue [27]. A estrutura de tais complexos são ideais para simulação do mecanismo de reação [2].

9.3 Métodos. A maior dificuldade de simular reações enzimáticas é o grande tamanho dos sistemas envolvidos (10.000 átomos ou mais em um complexo solvatado). Algumas simplificações podem ser feitas porque as mudanças químicas no sítio ativo usualmente envolvem diretamente apenas o substrato, certos resíduos catalíticos e cofatores. Porém, uma grande porção do restante da enzima e o solvente podem ter papel significativo na catálise e devem ser incluídos num tratamento completo, pelo menos de uma maneira aproximada. Para representar a formação e quebra de ligações que ocorre nas reações enzimáticas, a natureza quanto-mecânica (QM) do processo deve ser levada em consideração. Isto contrasta com o grande número de exitosos cálculos em sistemas macromoleculares para os quais uma função de potencial mecânico-molecular (MM) empírica é adequada. Tais simulações tem fornecido significativos *insights* da dinâmica protéica, por exemplo em estudos de ligação substrato-ligante. Simulações de dinâmica molecular clássica empregando funções de potencial MM recentemente desenvolvidas para proteínas e outras moléculas biológicas, serão realizadas para examinar o comportamento conformacional proteico, e sua solvatação, e comporão uma parte importante da pesquisa. Por outro lado, para modelar a reação química, requer-se o tratamento QM. Métodos de QM geralmente requerem muito tempo computacional e muita memória, mesmo para moléculas pequenas. Métodos *ab initio* de orbital molecular podem calcular precisamente as propriedades de pequenas moléculas isoladas e estados de transição de reações, se uma

teoria de suficiente alto nível é usada [28]. Tratamentos baseados em teorias de funcional de densidade mostram-se altamente promissores para aplicação em alguns sistemas grandes [29]. Técnicas semi-empíricas de orbital molecular (tais como AM1 [30] e PM3[31] baseadas no modelo MNDO) podem tratar sistemas grandes, porque as aproximações e a parametrização para reproduzir dados experimentais tornam-as muito menos exigentes em termos computacionais. Porém, mesmo cálculos semi-empíricos de orbitais moleculares são muito lentos para calcular dinâmica e caminho de reações enzimáticas. É também importante notar que, devido à sua natureza aproximativa, métodos semi-empíricos sofrem de certas limitações que restringem sua aplicação e sua precisão.

Então, como mencionado anteriormente, a maioria das simulações com proteínas são realizadas com funções de potencial MM empíricas, que permitem que se avalie rapidamente a energia potencial do sistema e as forças atuando sobre os átomos. Tipicamente, átomos são representados por cargas parciais pontuais com raio de van der Waals, que dão conta das interações não ligantes e termos simples (por exemplo ângulos e ligações harmônicas) são usados para representar as ligações químicas. Através de extensiva parametrização, funções de potencial MM fornecem resultados precisos de estrutura e dinâmica das macromoléculas [18,28]. Porém, a forma dos termos de energia, por exemplo a forma harmônica para a energia de estiramento ligante e a parametrização para a geometria da molécula no estado de fundamental, faz com que campos de força MM não possam ser aplicados em processos envolvendo reações químicas. Um tratamento atraente para simular reações enzimáticas é o uso de potenciais que combinem uma descrição QM para uma pequena região, incluindo o sítio ativo e o substrato e uma descrição MM para o restante da enzima e o solvente [3]. Tais métodos QM/MM podem tratar alterações químicas e eletrônicas e levar em consideração efeitos do meio. A combinação da versatilidade de um método QM com a simplicidade e rapidez de um campo de força MM permite estudar reações em grandes sistemas. Ela torna possível a determinação da energia e da variação de energia livre ao longo do caminho da reação e de reações em solução. Efeitos dinâmicos podem também ser estudados usando potenciais QM/MM. A maioria das aplicações até agora tem sido feitas a nível semiempírico do tratamento QM [2,13,14,33]. Métodos QM/MM usando tratamentos QM de orbital molecular *ab initio* ou funcional de



densidade tem sido desenvolvidos mais recentemente [34]. Cálculos QM/MM semi-empíricos tem a vantagem de serem mais rápidos e requererem consideravelmente menos memória e disco, permitindo que sistemas maiores sejam estudados e estudos extensivos de caminhos de reação e dinâmica sejam feitos. Cálculos QM/MM *ab initio* são essenciais para algumas aplicações nas quais tratamento semi-empírico tem fraco desempenho e são também muito úteis para testar e calibrar caminhos de reação determinados por métodos mais aproximados. Num cálculo QM/MM o sistema a ser estudado é dividido em três regiões: uma região QM, uma MM e uma região fronteira que leva em consideração o tamanho finito do sistema. Para a última tanto condições de contorno periódicas quanto métodos estocásticos podem ser usados. Os átomos na região QM são influenciados pelas cargas parciais dos átomos MM, incluídos no hamiltoniano uni-eletrônico. A energia de interação de Lennard-Jones entre átomos QM e MM é calculada como num cálculo MM normal (por ex. por um potencial 6-12): este termo é essencial porque não há repulsão de Pauli entre átomos QM e MM. Outras interações QM/MM são incluídas, por ex. átomos MM interagem com a distribuição de cargas QM, e alguns termos MM são selecionados para levar em consideração qualquer interação covalente que envolva pelo menos um átomo MM. Se as regiões QM e MM envolvem diferentes moléculas, que é normalmente o caso quando se estuda reações em solução, a separação QM/MM é particularmente simples. Se a linha que divide as regiões QM e MM separa átomos ligados covalentemente, as ligações dos átomos QM são satisfeitas introduzindo-se átomos "ligantes" (frequentemente átomos de hidrogênio), para completar as ligações incompletas. Muito cuidado deve ser tomado quando se divide o sistema nas regiões QM e MM de tal forma que o sistema seja realisticamente representado e não se provoque ruptura da estrutura eletrônica dos reagentes. Outros fatores que são importantes em cálculos MM tais como incluir efeitos dielétricos de solvente, devem ser considerados.

9.4 Procedimento. O estudo será feito utilizando-se o program CHARMM, que implementa todos os aspectos acima citados [26,35]. O input básico para um cálculo QM/MM requer a escolha de um sistema QM representativo, no qual se realizarão as alterações químicas. É importante testar o nível requerido de tratamento teórico, por ex.



através de cálculos QM em pequenos sistemas modelo [16,20,21]. Os parâmetros que serão usados também devem ser testados. Por ex., os parâmetros LJ para os átomos QM para interação com o sistema MM pode ser otimizado para reproduzir energias de interação e geometrias [4,34]. É importante também testar as cargas atômicas usadas no sistema MM. Para este fim, estudo de pequenas moléculas e suas interações são essenciais [16]. Procedimento similar é seguido no cálculo de parâmetros MM para funções de potencial padrão [26]. Simulações computacionais de reações enzimáticas requerem algum conhecimento estrutural da enzima se se quer calcular precisamente o papel catalítico de resíduos individuais e a importância do restante da proteína. Na prática isto significa que a estrutura cristalina é necessária, com inibidor ou substrato ligado, para mostrar a orientação dos grupos catalíticos em volta do substrato no sítio ativo. Dados sobre o complexo intermediários, por ex. obtidos por técnicas cristalográficas de Laue [27], podem ser extremamente úteis. Estudos cinéticos sobre a velocidade da reação são importantes para comparação. Estudos em enzimas mutantes podem testar conclusões sobre o papel do resíduo catalítico, embora conclusões possam ser dificultadas porque mutações podem causar mudanças na estrutura ou no mecanismo, então simulações na enzima nativa e na mutante são necessárias.

No estudo de uma reação enzimática por métodos QM/MM, uma estrutura representativa do complexo Michaelis é escolhida. Algumas modificações necessárias para converter um inibidor em substrato ou estrutura intermediária são realizadas e átomos de hidrogênio (que não são observados em estruturas protéicas obtidas por raios-X) são adicionados. Cuidado deve ser tomado para assegurar que os estados corretamente protonados são modelados. Cálculos eletrostáticos usando o método de Poisson-Boltzmann podem ser usados para prever pK_a s de grupos ionizáveis [40]. A estrutura é hidratada pela adição e termalização de moléculas de água (usando um modelo simples de água [10,26,41], enquanto moléculas de água observadas cristalograficamente são mantidas. Condições de contorno, por ex. similares às usadas em métodos estocásticos ou dinâmica molecular [7,18], são aplicados para garantir que o sistema se comporte realisticamente em casos onde a enzima como um todo não possa ser considerada e para incluir efeitos do solvente. A energia da estrutura é minimizada para se obter uma boa estrutura de partida. A

minimização da energia ao longo de uma coordenada de reação aproximada fornece o perfil de energia e a geometria aproximada do estado de transição no máximo de energia. Técnicas têm sido desenvolvidas para refinamento da estrutura do estado de transição em cálculos QM/MM [42]. Comparação com estruturas exatas do estado de transição em modelos pequenos representando importantes grupos funcionais também é útil [16]. As reações de substratos alternativos e enzimas mutantes também podem ser estudados. Simulações de dinâmica molecular clássica são importantes para amostragem configuracional: elas fornecem uma série de configurações protéicas ao longo de uma trajetória, as quais são usadas para modelar a reação. Comparação dos perfis de energia resultantes mostram os efeitos da variação da conformação protéica na reação. Sobretudo, estes cálculos podem indicar viabilidade de diferentes mecanismos propostos, identificando prováveis estruturas intermediárias e estados de transição, e os grupos responsáveis por esta estabilização. Para simular reações em solução segue-se um tratamento similar: o sistema reagente é imerso em uma caixa pré equilibrada de moléculas de água e algumas moléculas que ocupam o mesmo espaço da proteína são deletadas. Usa-se dinâmica molecular para equilibrar novamente o sistema soluto/solvente. Simulações podem ser realizadas a diferente temperatura, onde temperatura é definido em termos de uma energia cinética média dos átomos do sistema, para testar os efeitos da temperatura no comportamento da proteína.

Para calcular a energia livre de reação em uma enzima ou em solução, são necessárias técnicas especializadas porque os eventos reativos ocorrem com baixa frequência para serem acessíveis em uma escala de tempo de simulação de dinâmica molecular. Então usa-se uma combinação de dinâmica molecular e teoria de estado de transição (dinâmica ativada [7,43]). Uma coordenada de reação é escolhida (tipicamente um conjunto de n coordenadas atômicas) em que a configuração do reagente se converte na configuração do produto. A energia livre (potencial de força médio) pode ser calculada por métodos “umbrella sampling” [44,45]: uma série de potenciais “umbrella” é aplicado sequencialmente para obrigar a coordenada de reação a permanecer próxima a um conjunto de valores intermediários entre reagentes e produtos. A densidade de probabilidade parcial é calculada para cada valor da coordenada de reação sem restrições sobre as demais

coordenadas do sistema. A densidade de probabilidade real ao longo da coordenada de reação é obtida fazendo-se correções no potencial usado. Da energia livre de ativação a velocidade da reação pode ser calculada usando-se teoria do estado de transição [46]. Em termos de teoria do estado de transição, efeitos dinâmicos tais como barreira de transição, são englobados dentro do coeficiente de transmissão, τ , e este pode ser calculado por trajetórias partindo do estado de transição (máximo de energia livre) e monitorando-se o progresso da reação no sentido dos produtos, ou o retorno à configuração dos reagentes [43]. Onde a coordenada de reação tiver uma forma particularmente simples, por ex. uma distância interatômica como na reação de adição nucleofílica, a coordenada de reação pode ser restringida a um conjunto de valores fixos (por ex. usando o algoritmo SHAKE [47]), e a energia livre ao longo da coordenada de reação (potencial de força médio) pode ser calculada por integração termodinâmica do campo de força.

A contribuição de interações individuais ao longo da coordenada de reação podem ser calculadas, por ex., por uma simples análise de perturbação de primeira ordem na qual uma diferença de energia ou barreira é calculada na presença e ausência de grupos individuais. Isto fornece uma boa indicação daquelas interações diretamente responsáveis pelo estado de transição ou estabilização de intermediários, ou desestabilização de reagentes. Esta técnica tem fornecido resultados úteis para a enzima citrato sintase [13] e outras [3,14,17], identificando em cada caso grupos que são importantes na catálise. Conhecimento de quais resíduos amino ácidos ou outros grupos participam diretamente ou indiretamente na aceleração da velocidade pela enzima é importante para interpretar os efeitos de mutação ou polimorfismo genético. Além de diferenças na energia de interação, mudanças na energia livre devido à mutação podem ser calculadas por teoria de perturbação de energia livre [12,48].

Para simulações de dinâmica reacional e cálculos de energia livre, técnicas QM/MM semi-empíricas serão usadas uma vez que elas são consideravelmente menos exigentes em termos computacionais que tratamentos QM de alto nível. Cálculos QM *ab initio* e funcional de densidade serão usados para cálculo de energia ao longo do caminho de reação e cálculos adiabáticos do perfil de energia da reação. Resultados QM/MM de mais alto nível serão usados para testar e validar os métodos QM/MM semi-empíricos



aproximados. Cálculos sobre modelos pequenos podem também ser muito úteis neste sentido. Eles podem fornecer energias e estruturas completamente otimizadas de estados de transição e intermediários para comparação [16]. Uma abordagem muito interessante para modelar reações enzimáticas é o uso de métodos semi-empíricos adaptados para reproduzir resultados de alto nível para uma reação particular. Os parâmetros que integram o tratamento semiempírico podem ser otimizados para fitar resultados ab initio de alto nível ou experimentais [4]. Esta metodologia tem se mostrado útil para cálculos QM/MM e para simulações de reações orgânicas [49]. Cálculos ab initio ou de funcional de densidade são essenciais para alguns sistemas como muitas metaloenzimas por ex., para as quais tratamentos semi-empíricos são inadequados.

Simulações de dinâmica molecular puramente clássicas tem também um papel importante a desempenhar. A forma funcional simples de campos de força MM possibilitam que se estendam tais simulações ao domínio de tempo do nanosegundo, permitindo que se estude o comportamento conformacional de um grande número de enzimas. Simulações de dinâmica molecular clássica podem acessar escalas de tempo muito maiores do que os próprios métodos QM/MM semi-empíricos e serão usados para termalização e a geração de configurações representativas nas quais a reação possa ser estudada. Simulações de dinâmica molecular permitem calcular o tempo de vida das pontes de hidrogênio e analisar as energias de interação. Flutuações atômicas podem ser calculadas e comparadas com medidas experimentais de mobilidade. A descrição do movimento atômico fornecido pela simulação é consideravelmente mais detalhado do que aquele fornecido por estudos cristalográficos de proteínas [18]. Tipicamente, uma estrutura cristalina representa uma proteína por um modelo atômico simples, com conformações alternativas incluídas para alguns resíduos onde a densidade eletrônica claramente o indique. Movimento atômico é descrito pelos fatores B de Debye-Waller, que são isotrópicos e harmônicos. Isto é uma aproximação todavia, uma vez que se mostrou que flutuações atômicas em proteínas são altamente anisotrópicas e anarmônicas, e leva a uma subestimação da flutuação atômica, particularmente para grupos com alta mobilidade [7]. Outros fatores, como desordem cristalina, são incorporados no fator B e não são facilmente separados. Uma descrição mais



detalhada, do tipo disponível através de simulação molecular, é essencial para uma mais completa compreensão sobre a estrutura protéica.

A extensiva amostragem tornada possível pela dinâmica molecular clássica permite que mudanças de energia livre sejam calculadas por teoria de perturbação termodinâmica e técnicas tais como integração termodinâmica [7,12]. Tais simulações podem ser usadas para calcular a contribuição de efeitos particulares através da dissecação de interações, e testar parâmetros por ex., através de cálculos de energia livre de hidratação. A metodologia de perturbação de energia livre tem sido aplicada para investigar ciclos catalíticos em enzimas e em solução [7,33,48]. A investigação de efeitos de hidratação na estrutura e reatividade de moléculas biologicamente significativas é uma contribuição importante à compreensão do papel da água em sistemas biológicos. Simulações de dinâmica molecular de pequenas moléculas representativas, e reações, em solução, irão identificar efeitos e interações chaves. Em alguns aspectos, o solvente pode ser considerado como um catalisador [22], e é vital comparar reações em solução com a equivalente catalisada enzimaticamente a fim de entender o mecanismo catalítico da enzima.

9.5 O trabalho proposto. O trabalho que se está propondo se centraliza numa investigação detalhada sobre o mecanismo de reação de importantes enzimas e o desenvolvimento paralelo de técnicas de simulação. O objetivo é calcular o perfil de energia da reação dentro da enzima, identificar estruturas intermediárias e estados ativados na reação e examinar a influência de alguns aspectos na velocidade da reação. A reação enzimática será comparada com a reações de referência não catalisadas em solução, no sentido de se aprofundar a compreensão da reação catalisada. A dinâmica, o comportamento conformacional e a hidratação de enzimas e pequenas moléculas representativas será estudado em detalhes. A validade da aproximação linear de energia livre para interpretação da catálise será investigada. As enzimas escolhidas para o estudo, o serão em função da disponibilidade de dados experimentais de boa qualidade, do interesse em seu mecanismo de um ponto de vista aplicado, bem como de sua adequabilidade para tratamento no nível teórico que se pretende. Um objetivo particular é entender as causas fundamentais das diferenças na atividade entre enzimas de organismos normais e aquelas de organismos



adaptados a condições extremas como termofílicos por exemplo. As razões destas diferenças não são claras, uma vez que as estruturas tridimensionais de enzimas termofílicas são muito similares àsquelas de suas equivalentes normais. Tal estudo tem particular interesse em aplicações práticas em biotecnologia por exemplo. Nesse aspecto a enzima citrato sintase é particularmente interessante para estudo porque sua estrutura está resolvida para uma grande variedade de organismos, com um grande intervalo de temperatura ótima de reprodução (desde 10 °C - bactéria DS2-3R antártica - até cerca de 90 °C - *Pirococcus Furiosus*). Citrato sintase é uma enzima universal que catalisa um passo fundamental no ciclo do ácido cítrico. Estudos prévios em citrato sintase de frango identificaram um sítio ativo aspartato como uma base catalítica e o enolato de acetil-CoA como o intermediário formado por desprotonação de acetil-CoA. O enolato é estabilizado por ponte de hidrogênio com uma histidina de cadeia lateral e uma molécula de água e atua como nucleófilo para ataque no carbono carbonílico do oxaloacetato, que está polarizado no sítio ativo. Está se propondo neste projeto estender esse estudo a condições extremas para investigar as razões da mudança de atividade enzimática em função das condições termodinâmicas. Então, os mecanismos de reação nestas enzimas serão calculados por métodos QM/MM e os intermediários e estado de transição serão comparados para examinar onde residem as diferenças na atividade molecular. Citrato sintase sofre uma profunda mudança estrutural no processo de se ligar ao substrato indo de uma forma aberta a uma fechada, e vice-versa no processo de liberação dos produtos. Simulação computacional será usada para investigar a relação entre as mudanças químicas no sítio ativo e a mudança conformacional, como isto é afetado pela temperatura e como afeta a reação, e a razão para as diferenças que porventura hajam entre as diferentes enzimas. Tem se proposto que proteínas termoestáveis são menos flexíveis à temperatura ambiente e que isto está relacionado com redução na sua atividade [9,55]. Não está claro como, ou em que extensão, os diversos tipos de movimento que há numa proteína [7] são afetados, nem como eles estão relacionados com a atividade. Simulações de dinâmica molecular em um intervalo de temperatura usando técnicas MM será vital para identificar qualquer diferença na flexibilidade entre proteínas, e pode indicar se isto afeta diretamente a velocidade da reação. Elas indicarão o efeito da temperatura na estrutura e conformação protéicas.

Resultados recentes indicam que interação eletrostática entre grupos carregados é importante na estabilização de proteínas termofílicas [50]. Isto acontece porque tais interações tornam-se mais favoráveis com o aumento da temperatura, devido a mudanças na solvatação. Estudos das interações eletrostáticas em sistemas modelo e em proteínas a diferentes temperaturas serão realizados para investigar tais efeitos e o efeito na estrutura e dinâmica da enzima. Efeitos de solvente e interação entre grupos funcionais da proteína serão detalhadamente analisados. Outro fator possivelmente importante é a variação do pKa de resíduos básicos com a temperatura [54]. Simulações de pequenas moléculas em água a diferentes temperaturas indicarão a importância deste aspecto no comportamento proteico.

Trabalhos prévios mostraram que cálculos QM/MM semiempíricos fornecem bons resultados para citrato sintase [16,20,21] e em conjunto com cálculos de nível superior [4,39], eles serão estendidos para examinar o progresso da reação nas diferentes enzimas. A energia livre de ativação dos passos importantes da reação serão calculados com simulação de dinâmica molecular usando "umbrella sampling" [7] e comparados com resultados experimentais. Efeitos dinâmicos na velocidade da reação serão calculados [43,46]. Análise detalhada indicará as causas das diferenças de atividade observadas. Reações em solução de moléculas semelhantes ao substrato serão simuladas para investigar a origem do poder catalítico das enzimas. A discussão será centrada em interpretações em termos do formalismo de Marcus, que relaciona a energia de ativação total $\Delta G^\ddagger$, com a diferença de energia livre entre os reagentes e os produtos de um passo da reação, ΔG^o , e uma energia de ativação intrínseca ΔG_{int}^o característica do tipo de reação [13] :

$$\Delta G^\ddagger = \Delta G_{int}^o (1 + \Delta G^o / 4\Delta G_{int}^o)^2$$

Tem sido proposto que enzimas provocam uma redução na energia de ativação através de um abaixamento tanto de ΔG^o (por ex. por estabilização do intermediário) quanto de ΔG_{int}^o (por reduzir a componente de solvatação, abaixando efetivamente a energia dentro da enzima [1,5,25]. Isto é particularmente importante em reações envolvendo desprotonação de ácidos carboxílicos (por ex. desprotonação de acetil-CoA por citrato sintase e outras enzimas como triosefosfato isomerase), para o que ambos

$\Delta G^{\circ}_{\text{int}}$ e ΔG° são grandes [13]. A contribuição relativa destes fatores na eficiência da catálise enzimática não são conhecidos. Simulação da reação na enzima e a reação de referência em solução permitirá que diferentes efeitos sejam distinguidos. Avaliação da estratégia catalítica adotada por citrato sintase neste aspecto, terá implicações importantes para a compreensão de catálise enzimática em geral. A importância de outras estratégias catalíticas, tais como polarização de substrato (que se acredita contribuir na citrato sintase na medida em que oxaloacetato é polarizado na ligação), será investigado. Métodos QM/MM podem modelar mudanças na distribuição eletrônica e também podem ser usados para calcular polarização do substrato. Tais análises irão não somente identificar as contribuições mais importantes para a aceleração da velocidade conseguida por citrato sintase, mas também fornecer indicações dos aspectos gerais do mecanismo enzimático. Para estender as investigações e testar os resultados considerando os princípios da catálise enzimática, serão comparados os resultados obtidos com os de outras enzimas.

Sobretudo, este estudo fornecerá idéias significativas sobre os mecanismos das reações enzimáticas. Espera-se elucidar detalhes mecanísticos cruciais, e entender as linhas gerais dos princípios pelos quais a catálise ocorre. As simulações irão também investigar solvatação e comportamento conformacional e assim contribuir para o esclarecimento de efeitos fundamentais nas enzimas. Comparação das mesmas enzimas adaptadas a diferentes condições físicas ajudará a identificar as causas das diferenças na atividade, o que fornecerá subsídios importantes para a engenharia de enzimas incorporar a estabilidade de enzimas termofílicas com alta estabilidade. Os resultados serão testados e validados por comparação com resultados experimentais e estudos de pequenos modelos. Os métodos que serão aplicados e desenvolvidos neste trabalho serão de importância crescente em pesquisa química e bioquímica e têm potencial para contribuir com muitas áreas no futuro.

9.6 Referências.

1. Fersht, A., *Enzyme Structure and Mechanisms*, 2nd ed., Freeman, New York (1985)
2. Mulholland, A.J., Grant, G.H., Richards, W.G., *Protein Engng*, **6** (1993) 133
3. Mulholland, A.J., Karplus, M., *Biochem. Soc. Trans.*, **24** (1996) 247



4. Mulholland, A.J., Richards, W.G., in *Transition State Modeling for Catalyses*, Eds. Truhlar, D.G., Morokuma, K., American Chemical Society, Washington D.C., (1999) 448
5. Åqvist, J., Warshel, A., *Chem. Rev.*, **93** (1993) 2523
6. Burton, N.A., et al., *Faraday Discuss.*, **110** (1998) 463
7. Brooks, C.L., III, Karplus, M., Petit, B.M., *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*, Wiley, New York (1988)
8. Rasmussen, B.J., et al., *Nature*, **357** (1992) 423
9. Kohen, A., et al., *Nature*, **399** (1999) 496
10. Cordeiro, J.M.M., *Z. Naturforsch.*, **54a** (1999) 311
11. Freitas, L.C.G., Cordeiro J.M.M., Garbujó, F.L.L., *J. Mol. Liq.*, **79** (1999) 1
12. Cordeiro, J.M.M., Freitas, L.C.G., *Z. Naturforsch.*, **54a** (1999) 110
13. Mulholland, A.J., Richards, W.G., *Proteins: Structure, Function and Genetics*, **27** (1997) 9
14. Lyne, P., Mulholland, A.J., Richards, W.G., *J. Am. Chem. Soc.*, **117** (1995) 11345
15. Elcock, A.H., et al., *J. Am. Chem. Soc.*, **117** (1995) 4706
16. Mulholland, A.J., Richards, W.G., *J. Phys. Chem.*, **102** (1998) 6635
17. Rider, L., Mulholland, A.J., *J. Am. Chem. Soc.*, **120** (1998) 7641
18. Hadfield, A.T., Mulholland, A.J., *Int. J. Quant. Chem. (Biophys. Q.)*, **73** (1999) 137
19. Mulholland, A.J., Karplus, M., *FASEB J.*, **11** (1997) A1315
20. Mulholland, A.J., Richards, W.G., *J. Mol. Struct. (TheoChem)*, **429** (1998) 13
21. Mulholland, A.J., Richards, W.G., *J. Mol. Struct. (TheoChem)*, **427** (1998) 175
22. Jenks, W.P., *Catalysis in Chemistry and Enzymology*, Dover, New York (1987)
23. Wolfenden, R., *Bioorganic Med. Chem.*, **7** (1999) 647
24. Govardhan, C.P., Margolin, A.L., *Chemistry and Industry*, (1995) 689
25. Warshel, A., *J. Biol. Chem.* **273** (1998) 27035
26. MacKerell, A.D., et al., *J. Phys. Chem.*, **102** (1998) 3586
27. Stoddard, B.L., Dean, A., Bash, P. A., *Nature Structural Biology*, **3** (1996) 590
28. Hehre, W.J., et al., *Ab Initio Molecular Orbital Theory*, Wiley, New York (1986)



29. Parr, R.G., Yang, W., *Density Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford (1989)
30. Dewar, M.J.S., et al., *J. Am. Chem. Soc.*, **107** (1985) 3902
31. Stewart, J.J.P., *J. Comp. Chem.*, **10** (1989) 209
32. Mulholland, A.J., Richards, W.G., *Int. J. Quant. Chem.*, **51** (1994) 161
33. Cunningham, M. A., et al., *Biochemistry*, **36** (1997) 4800
34. Lyne, P.D., Hodoscek, M. Karplus, M., *J. Phys. Chem. A.*, **103** (1999) 3462
35. Brooks, B.R., et al., *J. Comp. Chem.*, **4** (1983) 187
36. Field, P.J., Bash, P. A., Karplus, M., *J. Comp. Chem.*, **11** (1990) 700
37. Schmidt, M.W. et al., *J. Comp. Chem.*, **14** (1993) 1347
38. Amos, R.D., et al., *CADPAC: The Cambridge Derivatives Package*, Cambridge (1995)
39. Mulholland, A.J., Lyne, P.D., Karplus, M., *J. Am. Chem. Soc.*, in press
40. Antosiewicz, J., et al., *J. Comp. Chem.*, **17** (1996) 1633
41. Jorgensen, J.L., *J. Chem. Phys.*, **79** (1983) 926
42. Turner, A.J., Moliner, V., William, I.H., *Phys. Chem. Chem. Phys.*, **1** (1999) 1323
43. Neria, E., Karplus, M., *Chem. Phys. Lett.*, **267** (1997) 23
44. Bartels, C., *J. Comp. Chem.*, **18** (1997) 1450
45. Bartels, C., Karplus, M., *J. Phys. Chem. B*, **102** (1998) 865
46. Steinfeld, J.I., Francisco, J.S., Hase, W.L., *Chemical Kinetics and Dynamics*, Prentice Hall, Englewood Cliffs, NJ (1989)
47. van Gasteren, W.F., Karplus, M., *J. Comp. Chem.*, **1** (1980) 266
48. Bash, P.A., Field, M.J., Karplus, M., *J. Am. Chem. Soc.*, **109** (1987) 8092
49. Rossi, I., Truhlar, D.G., *Chem. Phys. Lett.*, **233** (1995) 231
50. Elcock, A. H., *J. Mol. Biol.*, **284** (1998) 489
51. Russell, R.J.M., et al., *Structure*, **2** (1994) 1157
52. Russell, R.J.M., et al., *Biochemistry*, **36** (1997) 9983
53. Russell, R.J.M., et al., *Structure*, **6** (1998) 351
54. Danson, M.J., *Protein Engng.*, **9** (1996) 629
55. Wrba, A., et al., *Biochemistry*, **29** (1990) 7584



Trabalho aceito para publicação
(Molecular Engineering)



Study of Liquid Dimethyl Sulfoxide by Computer Simulation

João Manuel Marques Cordeiro

Departamento de Física e Química - FEIS - UNESP

CP 31 15.385 - 000 Ilha Solteira SP Brazil

Abstract

The paper reports Monte Carlo and molecular dynamics results for pure liquid dimethyl sulfoxide (DMSO) at 298K and 1 atm. The classical 6-12 Lennard-Jones plus Coulomb pairwise potential was used to calculate intermolecular interaction energy. Potential parameters for the liquid were optimized in this work. Some thermodynamic and dynamical properties obtained, such as heat of vaporization, density and diffusion coefficient, are in good agreement with the experimental values. The present model is compared with other models for DMSO reported previously. It is shown to be an improvement over earlier potentials. The structure factors and the radial distribution functions (rdf), are compared with experimental results for the liquid. The analysis shows that the structure of DMSO is not completely understood yet and deserves deeper investigation. The geometry of the dimer that corresponds to the rdf plots obtained is reported. The results suggest that the dipole moment of this dimer plays a important role in the structure of the liquid.

key words: liquid structure, pairwise potential, Monte Carlo simulation, molecular dynamics simulation.

1. Introduction

Studying dimethyl sulfoxide (DMSO) and its solutions has relevance to a broad range of topics in chemistry, biochemistry, environmental and medicinal science [1,2]. Allen *et al* and Borin and Skaf [3,4] have recently related a list of the uses of DMSO and the main properties shown for it and its solutions. The unique properties of DMSO give rise to its wide use as a solvent both for polar and nonpolar molecules [5,6]. Its aqueous solutions present very interesting anomalous behavior [3,4]. The electron donating and accepting capacity by the sulfur and oxygen gives DMSO both acid and basic characteristics. The SO polar group indicates the possibility of hydrogen bond formation with water, and the two CH₃ groups can give rise to hydrophobic effects. Due to these characteristics, there has been great interest in parametrizing reliable force fields for DMSO. Five rigid non-polarizable intermolecular potentials have been reported: the Rao and Singh model (RS) [7], the potential of Luzar and Chandler (LC) [8], the model parameterized by van Gunsteren *et al* (vG) [9], and Jorgensen's OPLS potential [10]. All these models use the same number of sites, the same site charge distribution, obtained using the CHELPG procedure with a 6-31G* basis set, and different sets of Lennard-Jones parameters. Several thermodynamic, structural and dynamic aspects have been investigated through computer simulations using them [7-14]. These studies have revealed that, to a higher or lesser degree, they yield reasonable descriptions of the liquid. All the cited models present the same feature for the weighted sum of pair distribution functions ($g_{XX}(r)$) and show a discrepancy, at a higher or lower level depending on the model, from the $g_{XX}(r)$ experimental plot [14],



I have recently undertaken Monte Carlo investigations on pure liquids and binary mixtures emphasizing relationships between intermolecular interaction, structure and thermodynamic properties [15-18]. This research area has a broad significance for chemistry and biochemistry, since the comprehension of these aspects is essential to understand environmental effects on the structure and reactivity of molecules [19-20]. Monte Carlo and molecular dynamics simulations have been shown to be powerful tools in this task. Taking into account that DMSO is a very interesting medium to study protein behavior, but that in this case solute-solvent interaction must be very well described, I have optimized new parameters for the molecule. The Lennard-Jones parameters of RS [7] (the model that presents better agreement with experimental structure, although the worst agreement with thermodynamics properties) have been used as a starting trial set, and the point charges have been adjusted through a series of Monte Carlo simulations, to reproduce the experimental density and heat of vaporization of the liquid.

In this paper, thermodynamic results, radial correlation functions and structure factors obtained with the optimized potential are presented and compared with the earlier ones. Both MC and molecular dynamics techniques were used where appropriate. In the next sections, the potential function used to calculate the intermolecular interaction energy and the molecular model will be described. The results obtained are given afterwards, followed by the discussion, and in conclusion a summary of the main results.

2. Computational Details

2.1 The intermolecular potential function



Following usual procedures in force field calculations, the molecules were modeled by collections of interacting sites. The energy E_{ab} between molecules **a** and **b** was given by a sum of Coulomb plus Lennard Jones terms centered on the sites:

$$E_{ab} = \sum_{ij} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{r_{ij}} \right] \quad (1)$$

where r_{ij} is the distance between site *i* in **a** and site *j* in **b** and q_i and q_j are fractional point charges located on the *i* and *j* molecular sites. For each site *k*, the parameters A_{kk} and B_{kk} were given by $A_{kk} = 4\epsilon_k \sigma_k^{12}$ and $B_{kk} = 4\epsilon_k \sigma_k^6$, where ϵ_k and σ_k are the Lennard-Jones parameters for the *k*th site. Parameters A_{ij} and B_{ij} for a nondiagonal interaction [*i,j*] were obtained using the geometric combining rules $A_{ij} = (A_{ii} A_{jj})^{1/2}$ and $B_{ij} = (B_{ii} B_{jj})^{1/2}$ [21]. This is a standard procedure used in this field of study. In an earlier work was shown that the results obtained using this combining rule and the well-known Lorentz-Berthelot rule are very similar [15].

2.2 Molecular model

DMSO has a pyramidal shape with the sulfur at the apex. The molecule was represented in this work as a four site rigid structure with each methyl group as a united atom site. From this point of view, the present model is the same as those reported earlier. Therefore, contributions from intramolecular relaxation effects are not considered, an approximation that has been widely used in liquid simulation [22]. The molecular geometry parameters used were those obtained from liquid X-ray study by Itoh and Ohtaki [23]. Vaismann and Berkowitz present a table with

values of these parameters obtained from several experimental techniques [11]. All the values are quite similar, including those ones reported later by Luzar and coworkers [12].

2.3 Monte Carlo simulations

MC simulations were used to calculate density and heat of vaporization, and the rdf of the liquid. The main aspects involved will be given here. For additional detail one can see the recent paper of Freitas *et al.* [17] that report results for the acetone-water mixture.

The simulations were carried out in the NPT ensemble at 298 K and 1 atm with Metropolis importance sampling [24] and periodic boundary conditions on systems consisting of 500 molecules placed in a cubic box [25]. A full intermolecular interaction was calculated whenever any of the site-to-site distances r_{ij} fell below a cutoff radius of $10\text{\AA}$. The contributions of Lennard-Jones interactions beyond the cut-off radius were considered using the formalism presented by Allen and Tildesley [21]. In this methodology the long range contribution from the Lennard-Jones potential is included by integrating the potential function from the cut-off radius to infinity, assuming that the pair correlation function is unitary. As it has been widely used in this field of study, long range corrections to take into account Coulombic interactions beyond the cut-off radius were not included in the present work. A brief analyses on the aspects that validate this assumption was done recently [17]. On the other hand, as discussed elsewhere, site-site pair correlation functions, thermodynamic properties and single particle dynamics of polar liquids appear to be quite insensitive to the long - range forces in the system [7], which is in accord with results obtained



recently[26] using the generalized Born model [26]. The calculation started with an equilibration phase with 1.6×10^7 configurations and the averages were then obtained from a new segment with others 1.6×10^7 configurations. Statistical uncertainties were calculated from separate averages over blocks of 2×10^5 configurations.

2.4. Molecular dynamics simulations

MD simulations were used to calculate the diffusion coefficient and the structure factor for DMSO. The main aspects involved in the MD simulations will be given here. Additional detail can be found in the recent paper of Skaf [14].

The simulations were run in the NVE ensemble on systems consisting of 864 molecules placed in a cubic box with periodic boundary conditions. The box length was chosen to match the experimental density of DMSO. The Lennard-Jones interactions were cut off at $\frac{1}{2}$ box length and the long-ranged portions of the electrostatic potentials were treated by Ewald sums with conducting boundaries [28]. The equations of motion were integrated using the leap-frog algorithm [21] with a time step of 4 fs, while the molecular geometries were restored using SHAKE [29]. This enabled total energy conservation within 0.1% during unperturbed runs lasting 20 ps. The results reported here are based on averages over about 15 000 time steps (60 ps).

3. Results and Discussions

3.1 Parametrization and Thermodynamics.

The site charges optimized in this work and those obtained by CHELPG calculations are listed in Table 1. The dipole moment is greater than the



experimental one. This criterion has been extensively used to compensate for the neglect of polarization effects in the simulations [7,22]. With these values for the charges, a little final adjusting of the Lennard-Jones parameters of the methyl group was necessary to obtain better agreement with experimental values of density and heat of vaporization. Table 2 shows the Lennard-Jones potential parameters optimized in this work compared with those of the other quoted models.

Table 3 lists some thermodynamic results obtained with the model optimized in this work along with experimental data. There is a good agreement between the theoretical values and the experimental ones.

The values of average potential energy and diffusion coefficients obtained for the other models are listed in Table 4 for comparison. As one can notice by comparing the results listed in Tables 3 and 4, the model parametrized in this work shows better agreement with the experimental results than the earlier models reported, representing an improvement in the intermolecular potentials available for DMSO.

3.2 Structure.

3.2.1 Structure Factors

The weighted sum of heavy atoms (O, S, C) pair distribution obtained with the model optimized in this work can be compared with the neutron diffraction data reported by Luzar *et al.* [12], and the results are shown in Figure 1. For comparison with the structure factors produced by the others reported models see ref. 14.

Comparing the results obtained in this work and the previously related ones [14], it is clear that the present model differs from the experimental $g_{xx}(r)$ in a similar way to the best of the earlier models, the OPLS potential in particular. It is interesting to notice the similarity among the $g_{xx}(r)$ obtained with the various models available,



including the present one, in spite of the significant difference in their potential parameters. Firstly, this agrees with earlier work [12], that showed packing effects play an important role in the structure of liquid DMSO (see below also). Secondly, besides the peak at 4 Å shown by the theoretical plots (only a shoulder in the experimental plot), the peak in 6 Å is higher than the experimental one, which indicates that the theoretical liquid is more structured than the real one. This is not in agreement with other theoretical findings, such as the radial distribution function [7-10 and this work] and Kirkood g factors [14] for the reported models that show the liquid DMSO to be less structured than the real liquid. This lack of structure is mainly a consequence of the molecular rigidity of the models that do not consider appropriately the high polarizability of liquid DMSO. In the particular case of $g_{XX}(r)$, it is possible that the theoretical results may be better than the experimental one (that is, the experimental result should be even higher than the theoretical one), as a consequence of the neutron diffraction methodology used to obtain the experimental $g_{XX}(r)$ [32]. There are many contributions to the experimental $g_{XX}(r)$ function, and the most important of them, SS, OO and SO, are also the weakest. Hence these correlations could have a lot of structure but would be almost invisible to the neutron experiment. Overlap and geometric constraints prevent the dominant CC, CO and CS correlations. So, the experimental $g_{XX}(r)$ may be under-dimensioned.

3.2.2 Distribution Functions

Distribution functions play an important role in the understanding liquid structure. Experimental information about radial distribution functions (rdf) can be obtained by neutron scattering and x-ray diffraction. Nevertheless, the great number of overlapping distances, both within and between molecules, and the great number of



different site-site correlation functions added together in the diffraction pattern, make the experimental investigation of liquid structure exceedingly difficult [17]. Therefore molecular modeling, and computer simulations to acquire the various site-site correlation functions, can be useful in providing improved insight about the liquid structure.

Figure 2 shows the site-site radial distribution function (rdf) obtained with the present model for DMSO. The curves are similar to theoretical ones reported earlier [7,11,12], mainly those obtained with the P2 potential [12]. All correlation distances agree with the intermolecular correlation reported by Itoh and Ohtaki [23]. The shortest site-site correlation amplitude is observed at 3.3 Å in the rdf for methyl to oxygen interaction. Previous works on DMSO have attributed this peak to weak O-Me hydrogen bonds in DMSO [7,11]. As it has been already appointed by Luzar *et al.* [12], this is weak evidence for the existence of hydrogen bonds in DMSO, especially considering what has been discussed in later papers [33]. The same authors have shown that a substantial part of this peak correlation is due to packing effects in DMSO [12] and, hence, cannot be entirely credited to molecular association. It will be discussed later in this work that this correlation is also a consequence of the preferential geometry arising from dipole-dipole orientation, which agrees with studies based on dielectric data [34]. In that work, Amey has emphasized that any form of molecular array in the liquid must be related to dipole-dipole interaction.

Excluding this sharp correlation, one can notice that, as a general trend, the rdfs for DMSO exhibit large bands at mean site-to-site distances ranging from 4 to 7 Å. Radial distribution functions with broad peaks are characteristic of structureless fluids [17] and agree with the small entropy of vaporization of organic liquids [35,36].

Experimental and theoretical studies have shown that organic liquids, capable of hydrogen bonding, such as alcohols, exhibit quite different behavior [37-39].

A standard procedure in this field, is the calculation of the site-site coordination number by integrating the appropriated correlation function over the first maximum up the position of first minimum [21]. Integration throughout the first peak in the S/S rdf gives a coordination number near 12. This number agrees with that found by Luzar *et al.* using neutron diffraction [12] and can be interpreted as the average number of first neighbors for a given DMSO molecule. A coordination number of this magnitude is characteristic of structureless liquids and can also be associated with packing of hard spheres. It is possible, as it has been suggested in some recent analysis for water, that the area of the peak is not a reliable measure of the coordination number, and so the figures may have some intrinsic errors [40-42].

The energy and structure of DMSO dimers were also investigated using molecular mechanics [43], as was done in a earlier work for acetone [17]. Figure 3 shows the structure of the two most stable dimers obtained in gas phase.

The most stable dimer in the gas phase is the anti-parallel one (label **a**), which has a dipole moment equal to 0.0 D, while dimer labeled **b** has a dipole moment of 5.9 D. The site-site distances calculated for the dimers shown in Figure 3 can be compared with rdf shown in Figure 2. As has been found for liquid acetone [17], the dimer with greater dipole moment is that one which site-site distances agree with peak distances in the $g(r)$ plots. Therefore, the structure of the liquids seems to be determined more by the dimer with the greater dipole moment. It is interesting to note that the O-O and Me-Me distances in dimer **b** also agree with the respective $g(r)$ (not shown in the Figure 2). Finally, I would like to draw attention to the form of the S-O $g(r)$ curve in Figure 2. It is interesting that this curve has such a diffuse shape, considering the

agreement between the geometry of the dimer **b** and the rdf plots. This behavior is also observed in the Luzar *et al.*'s report [12]. I can not explain this found fully but, of course, it must be consequence of the disorder in the liquid DMSO. Therefore, the dimer **b** does not fully represent the liquid structure and larger clusters must have a significant role in the DMSO properties.

4. Conclusions

A new model for liquid DMSO has been optimized, and some energetic, dynamical and structural features of it have been calculated using a series of MD and MC simulations. The theoretical density and heat of vaporization obtained with this model have compared well with the experimental values. The diffusion coefficient and $g_{xx}(r)$ function for the model show considerable improvement compared with previous models. Radial distribution functions show broad correlation curves, generally found for non-structured liquids. Molecular mechanics calculations show that the predominant dimer in the liquid is not the most stable found in gas phase, but the dimer with the greater dipole moment. The structural results related in the present work and others compiled in the literature indicate the structure of DMSO is not completely understood yet and deserves deeper investigation.

Acknowledgments

The author is grateful to Dr. A. K. Soper, for useful discussions and Dr. A. J. Mulholland for the interesting suggestions. This work was partially supported by Fundação para o Desenvolvimento da UNESP (FUNDUNESP). The author also acknowledges Dr. Munir Skaf for the MD calculations, and the Centro Nacional de Processamento de Alto Desempenho (CENAPAD) for the computational facilities provided.



References

- 1- S.W. Jacobs, E. Rosenbaum and D. Wood: Dimethyl Sulfoxide, Marcel Dekker, N.Y. (1971).
- 2- D. Martin and H. Hanthal: Dimethyl Sulfoxide, Wiley, N. Y. (1975).
- 3- H.C. Allen, D.E. Gragson and G.L. Richmond: J. Phys. Chem. B, 103, 660 (1999).
- 4- I.A. Borin and M.S. Skaf: J. Chem. Phys., 110, 6412 (1999).
- 5- W.R. Fawcett and A. A. Kloss: J. Chem. Soc., Faraday Trans., 92, 3333 (1996).
- 6- R.H. Figueroa, E. Roig and H.H. Szmant: Spectrochim. Acta, 22, 587 (1996).
- 7- B.G. Rao and U.C. Singh: J. Am. Chem. Soc., 112, 3803 (1990).
- 8- A. Luzar and D. Chandler: J. Chem. Phys., 98, 8160 (1993).
- 9- H. Liu, F. Müller-Plathe and W. F. van Gunsteren: J. Am. Chem. Soc., 117, 4363 (1995).
- 10- Y. Zheng and R. L. Ornstein: J. Am. Chem. Soc., 118, 4175 (1996).
- 11- I.I. Vaismann and M.L. Berkowitz: J. Am. Chem. Soc., 114, 7889 (1992).
- 12- A. Luzar, A.K. Soper and D. Chandler: J. Chem. Phys., 99, 6836 (1993).
- 13- M. S. Skaf: Mol. Phys., 90, 25 (1997).
- 14- M. S. Skaf: J. Chem. Phys., 107, 7996 (1997).
- 15- L.C.G. Freitas and J.M.M. Cordeiro: J. Mol. Struct. (THEOCHEM), 333, 189 (1995).
- 16- J.M.M. Cordeiro: Int. J. Quant. Chem, 65, 709 (1997).
- 17- L.C.G. Freitas, J.M.M. Cordeiro and F.L.L. Garbujo: J. Mol. Liq., 79, 1 (1999).
- 18- J.M.M. Cordeiro and L.C.G. Freitas: Z. Naturforsch, 54a, 110 (1999).
- 19- W.F. van Gusteren and P.K. Weiner (eds.): Computer Simulation of Biomolecular Systems, ESCOM, Leiden (1989).



- 20- A. Warshel: Computer Simulation of Chemical Reactions in Enzymes and Solutions, John Wiley&Sons, New York (1991).
- 21- M. Allen and D. J. Tildesley: Computer Simulation of Liquids, Oxford University Press, Oxford (1987).
- 22- B.M. Ladanyi and M.S. Skaf: Ann. Rev. Phys. Chem., 44, 335 (1993).
- 23- S. Itoh and H. Ohtaki: Z. Naturforsch., 42a, 858 (1987).
- 24- N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller: J. Chem. Phys., 21, 1087 (1953).
- 25- The calculations were performed with the DIADORIM program, a FORTRAN code developed by L. C. Gomide Freitas, Department of Chemistry, Universidade Federal de São Carlos, São Paulo, Brazil.
- 26- P. Belletato, L.C.G.Freitas, E.P.G. Arêas and P.S. Santos: Phys. Chem. Chem. Phys., 1 (1999), to be published
- 27- M. Neumann: J. Chem. Phys., 82, 5663 (1985); H.E. Alper and R.M. Levy: J. Chem. Phys., 91, 1242 (1989)
- 28- S.W. De Leeuw, J.M. Perram and E.R. Smith: Annu. Rev. Phys. Chem., 37, 245 (1986).
- 29- G. Ciccotti and J-P. Ryckaert: Comput. Phys. Rep., 4, 345 (1986).
- 30- T.B. Douglas: J. Am. Chem. Soc. 70, 2001 (1948).
- 31- D. R. Lide, editor-in-chief: Handbook of Chemistry and Physics, 76RD edition, CRC Press INC (1995).
- 32- A. K. Soper, personal e-mail communication.
- 33- see refs. 16 and 18, and references cited therein.
- 34- R.L. Amey: J. Phys. Chem., 72, 3358 (1968)



- 35- Y. Marcus: Ion Solvation, John Wiley and Sons, New York, (1985)
- 36- N. Isaacs: Physical Organic Chemistry, Longman Scientific and Technical, Essex (1995)
- 37- L.C.G. Freitas: J. Mol. Struct. (THEOCHEM), 282, 335 (1993)
- 38- A.L.L. Sinoti, J.R.S. Politi and L.C.G. Freitas: J. Mol. Struct. (THEOCHEM), 333, 249 (1996)
- 39- W.L. Jorgensen, D.S. Maxwell and J. Tirado-Rives: J. Am. Chem. Soc., 118, 11225 (1996)
- 40- J.M.M. Cordeiro: Z. Naturforsch, 54a, 311 (1999)
- 41- A.A. Chialvo and P.T. Cummings: J. Chem. Phys., 101, 4466 (1994)
- 42- Yu. E. Gorbaty and A.G. Kalinishev: J. Phys. Chem., 99, 5336 (1996)
- 43- PCModel, Serena Software, Bloomington, USA



Table I - Site charges parameters and dipole moment for DMSO

site	this work	CHELPG
O	-0.53	-0.459
S	0.17	0.139
CH ₃	0.18	0.16
dipole moment. (D)	5.0	4.4

The experimental dipole moment is 4.1 D [1]



Table II - Nonbonded Lennard-Jones Parameters for DMSO

LJ parameters	LC	RS	VG	OPLS	present model
ϵ_{OO}	72	66	410	280	66
σ_{OO}	2.80	2.94	2.63	2.93	2.94
ϵ_{SS}	239	202	310	395	202
σ_{SS}	3.40	3.56	3.56	3.56	3.56
ϵ_{CC}	294	160	225	160	190
σ_{CC}	3.80	3.60	3.66	3.81	3.80

ϵ in cal/mol, σ in Å

Table III - Values of some theoretical (this work) and experimental thermodynamics properties of DMSO

ΔH_v	$-E_i$	D (10^{-5})	$\bar{V}$	ρ
Simulated				
12.62±0.02	12.06±0.02	0.9	71.06±0.12	1.10±0.03
Experimental				
12.64 [30]	12.08	0.8	71.35	1.095 [31]

Heat of Vaporization (ΔH_v) and intermolecular energy (E_i) in kcal/mol. The experimental value for the average inner energy was estimated by $-E_i = \Delta H_v - RT$, where ΔH_v is the experimental heat of vaporization.

Diffusion coefficient (D) in cm^2/s . Errors have not been estimated for this property.

Molar volume ($\bar{V}$) in cm^3/mol (the experimental ones are calculated using experimental density and molecular weight).

Density (ρ) in g/cm^3 .

Table IV - Average inner energy and diffusion coefficients for the models of DMSO earlier reported. The experimental values are given in Table 2.

property	LC	RS	VG	OPLS
$-E_i$ (kcal/mol)	11.72	8.73	12.80	11.77
D (10^{-5} cm ² /s)	1.2	2.5	1.3	0.9

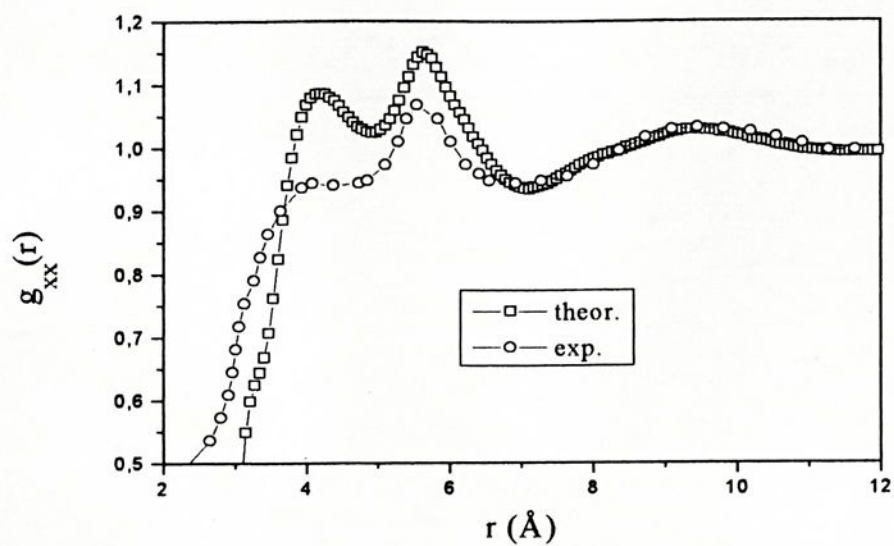
These data were published in the reference 14

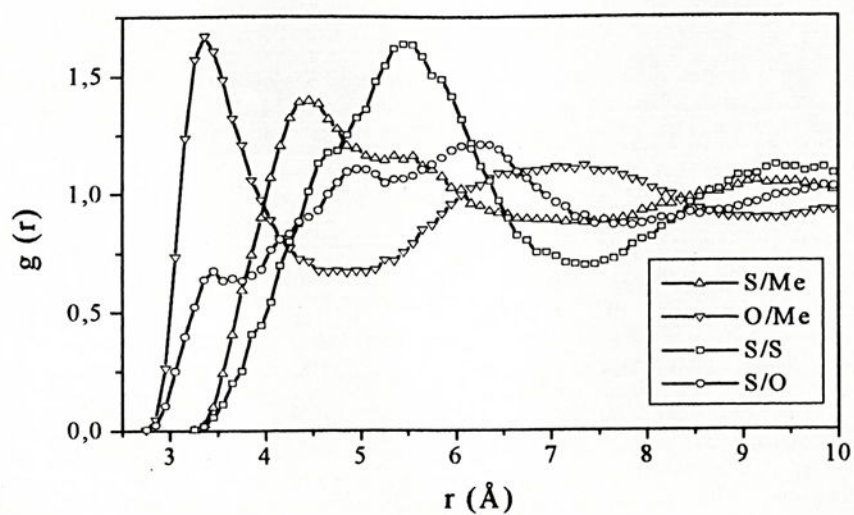
Figure Captions

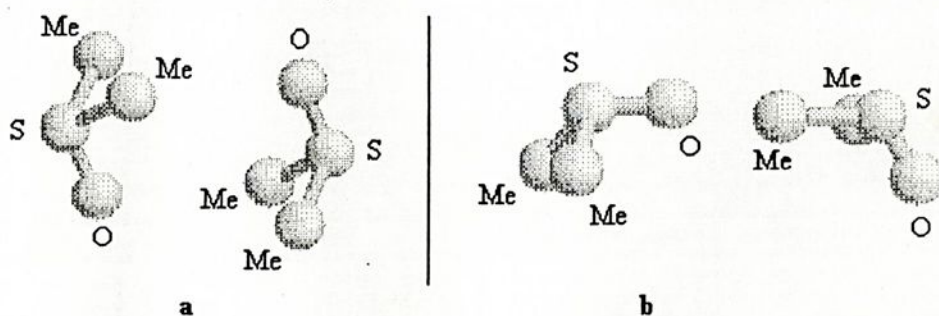
Figure 1. Weighted sum of pair distribution functions of the heavy atoms for DMSO.

Figure 2. Radial distribution functions for pure liquid dimethyl sulfoxide. S = sulfur, O = oxygen, Me = methyl group.

Figure 3. Structures of DMSO dimers optimized using molecular mechanics and values of the most representative site-site distances.







Some site-site distances obtained by molecular mechanics for the two dimers above

dimer/distances (Å)	S-S	S-Me	O-Me	O-S	O-O	Me-Me
a	4.8	4.5	3.7	4.3	4.2	4.0, 4.9
b	5.4	4.2, 5.4	3.4	4.3	5.2	3.9, 5.2