UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA

Estudo teórico da estabilidade estrutural e solvatação hidrofílica da α-ciclodextrina

Herbert de Castro Georg

Dissertação de Mestrado submetida
ao Instituto de Física da
Universidade de São Paulo

Orientadora:
Prof[a]. Dra. Kaline Coutinho

Banca examinadora:
Prof[a]. Dra. Kaline Coutinho (UMC)
Prof. Dr. Luiz Carlos Gomide Freitas (DQ/UFSCar)
Prof[a]. Dra. Vera Bohomoletz Henriques (IF/USP)

Financiado pela FAPESP
São Paulo
2002

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

Georg, Herbert de Castro

Estudo Teórico da Estabilidade Estrutural e Solvatação Hidrofílica da α-ciclodextrina.
São Paulo 2002.

Dissertação (Mestrado) - Universidade de São Paulo
Instituto de Física – Departamento de Física dos Materiais e Mecânica

Orientador: Prof. Dr. Kaline Coutinho
Área de Concentração: Física

Unitermos: 1. Efeitos de Solvente;
2. Modelagem Molecular;
3. Mecânica Quântica;
4. Simulação Computacional de Líquidos;
5. Ciclodextrinas.

USP/IF/SBI-039/2002

A meus pais Lauro e Eleusa

e

à Versanna.

Agradecimentos

Agradeço primeiramente aos meus pais Eleusa e Lauro pelo apoio irrestrito, carinho e dedicação ao longo de toda a minha vida. Não há, absolutamente, palavras para expressar minha gratidão. Agradeço também aos meus irmãos Reinvald, Raphaela e Rowena pela contribuição particular de cada um no desenvolvimento da pessoa que sou hoje. Em especial, agradeço à Raphaela pelo companheirismo destacado nos últimos tempos.

À minha namorada, Versanna, que me acompanha já há algum tempo, tenho muito o que agradecer pelo seu amor e também pela compreensão de minhas escolhas e por escolher continuar comigo nessa jornada. O futuro saberá nos recompensar.

Aos amigos e colegas de grupo Eudes, Thaciana, Valdemir, Rivelino, Luís, Puspito, Ivam, Eduardo e Daniel agradeço pela convivência amigável e conversas mil, teorias malucas e discussões proveitosas. Aos demais amigos e colegas do DFMT agradeço pelo ambiente agradável e descontraído e em especial ao Renato conterrâneo e contemporâneo de graduação assim como Eudes e Thaciana.

Agradeço ao Prof. Sylvio e principalmente à Prof^{a}. Kaline pela acolhida no grupo e pela oportunidade de trabalhar num ramo da ciência fascinante como é a Física Molecular e por me propiciar o aprendizado que tive durante o mestrado.

Agradeço ao Prof. Hélio dos Santos da UFJF e ao Prof. Wagner Almeida da UFMG pelas discussões e pelas geometrias experimentais gentilmente cedidas.

Agradeço, por fim, à FAPESP pelo apoio financeiro durante o mestrado.

Abstract

We have made a systematic and careful theoretical study to understand the role of water molecules in α-CD geometry stabilization in aqueous solution. In addition, we have analysed the reliability of largely used classical force fields and the applicability of quantum mechanicaly optimized geometries through comparison with experimetal results.

Initially, we have performed complete geometry optimizatons with α-CD in vacuum using AMBER, OPLS and MM2 as classical methods and with Hartree-Fock, AM1 and PM3 as quantum mechanics methods. Among the 42 geometries found, we have identified the most stable ones obtained from classical (AMBER) and quantum (HF) methods. By comparing structural properties of these two geometries with experimental data, we have noticed that they are very close and differ from each other essentially by their cavity size, conical aperture and primary hydroxyl groups conformation. In AMBER geometry, these hydroxyl groups are all in (-)*gauche* conformation while in HF geometry there are two hydroxyl groups in *trans* conformation caused by formation of two intramolecular hydrogen bonds that make this geometry stabler by around 2 kcal/mol.

Through computer simulations with Monte Carlo method, we have analysed the solvation of these two geometries in water. We have found that AMBER geometry, although having a smaller dipole moment, interacts stronglier with water by around 21 kcal/mol. This is because of 6 hydrogen bonds that AMBER geometry establishes with water more than the the HF geometry. We have also analysed the hydrophobic character of the cavity and calculated the internal density as 0.33 g/cm^3 and 0.46 g/cm^3 in HF and

AMBER respectively. This hydrophobic effect inside the cavity in both geometries is quite strong, however this effect is a little stronger in HF.

Finishing our analysis, we have performed Thermodynamic Perturbation simulations and calculated the free energy diference between these two geometries in aqueous solution. We have found that AMBER geometry is stabler by around 20 kcal/mol and that entropy variation is negligible.

Thus, we have observed that, among the theoretical methods we have used for geometry optimization, AMBER force field gave the best results in describing α-CD geometry.

Resumo

Fizemos um estudo teórico sistemático e criterioso para entender o papel das moléculas de água na estabilização da geometria da α-ciclodextrina (α-CD) em solução aquosa. Adicionalmente, analisamos a confiabilidade dos campos de força clássicos amplamente utilizados e a aplicabilidade de geometrias otimizadas com cálculos quânticos, através de comparação com resultados experimentais.

Inicialmente realizamos otimizações de geometria da α-CD isolada utilizando os métodos clássicos AMBER, OPLS e MM2 e os métodos quânticos Hartree-Fock (HF), AM1 e PM3. Dentre as 42 geometrias encontradas, identificamos a mais estável obtida com métodos clássicos (AMBER) e com métodos quânticos (HF). Através de comparação de propriedades estruturais dessas duas geometrias com os dados experimentais, observamos que elas são muito próximas e que diferem basicamente pelo tamanho da cavidade, abertura do cone e conformação das hidroxilas primárias. Na geometria AMBER essas hidroxilas são todas (-)*gauche* enquanto que na HF há duas hidroxilas na conformação *trans* provocadas pela formação de ligações de hidrogênio intramoleculares e com isso dando uma estabilidade maior para essa geometria de cerca de 2 kcal/mol.

Utilizando simulações computacionais com método de Monte Carlo, analisamos a solvatação dessas duas geometrias em água. Verificamos que a geometria AMBER, embora tendo um dipolo menor, interage mais fortemente com a água por cerca de 21 kcal/mol. Isso é decorrência de 6 ligações de hidrogênio a mais que essa geometria faz com a água em comparação com a geometria HF. Analisamos também o caráter hidrofóbico

da cavidade e calculamos a densidade interna na geometria HF de 0.33 g/cm^3 enquanto que na cavidade da geometria AMBER essa densidade é 0.46 g/cm^3.O efeito hidrofóbico provocado pela cavidade em ambas as geometrias é bastante forte, sendo na geometria HF um pouco maior.

Para finalizar nossa análise, realizamos simulações com Teoria de Perturbação Termodinâmica (PT) e calculamos a variação de energia livre entre essas duas geometrias em solução aquosa. Encontramos que a geometria AMBER é mais estável em água por cerca de 20 kcal/mol. Como a entropia é dada por $(\Delta F-\Delta E)/T$, temos que a variação da entropia entre os dois estados é desprezível.

Assim verificamos que dentre os métodos teóricos utilizados para otimização de geometria, o campo de força clássico AMBER foi o que melhor descreveu a geometria da α-CD.

Índice

1. Introdução Geral 1

1.1. Sistemas moleculares 1

1.2. Carboidratos 3

1.2.1. Monossacarídeos 3

1.2.2. Oligossacarídeos 9

1.2.3. Polissacarídeos 11

1.2.4. Ciclodextrinas 13

2. α-Ciclodextrina Isolada 17

2.1. Introdução 17

2.2. Mecânica molecular 18

2.2.1. Energia de estiramento 19

2.2.2. Energia de deformação angular 20

2.2.3. Energia de torção 20

2.2.4. Energia de deformação fora do plano 21

2.2.5. Energia de van der Waals 22

2.2.6. Energia eletrostática 24

2.2.7. Termos cruzados 25

2.2.8. Tipo de átomo 25

2.2.9. Parametrização 26

2.2.10. Principais campos de força empíricos 26

2.3. Grandezas geométricas da α-CD 28

2.4. Resultados com campos de força 33

2.5. Métodos de química quântica 40

2.5.1. Separação de Born-Oppenheimer 41

2.5.2. Aproximação de Hartree-Fock 43

2.5.3. Aproximação LCAO 47

2.5.4. Métodos semiempíricos 49

2.6. Resultados com métodos quânticos 51

2.7. Comparação das geometrias 53

2.8. Conclusão parcial 58

3. Solvatação da α-Ciclodextrina 59

3.1. Simulação computacional 59

3.2. Método de Monte Carlo 63

3.2.1. Implementação 69

3.2.2. Considerações adicionais 71

3.2.3. Potencial de interação molecular 72

3.2.4. Função de distribuição radial 75

3.2.5. Ergodicidade 77

3.3. Resultados 78

3.3.1. Interação da α-CD com a água 78

3.3.2. Análise das RDFs 83

3.3.3. Ligações de hidrogênio 86

3.3.4. Cavidade da α-CD 90

3.3.5. Conclusão parcial 93

4. Estabilidade Conformacional 95

4.1. Introdução 95

4.2. Perturbação termodinâmica 96

4.2.1. Realização prática 99

4.3. Resultados 101

5. Conclusões e Perspectivas 109

Referências Bibliográficas 115

Apêndice 121

Capítulo 1

Introdução Geral

Neste capítulo faremos a exposição do problema estudado. Faremos também uma breve introdução sobre carboidratos, e em especial sobre as ciclodextrinas, das quais a α-ciclodextrina é o objeto do nosso estudo.

1.1. Sistemas moleculares

Do ponto de vista teórico, para se estudar moléculas em solução é necessário ter informações sobre a geometria e sobre o potencial de interação entre as moléculas. O procedimento comum para se obter a geometria é através dos dados experimentais, quando acessíveis, e através de cálculos de minimização de energia que geram geometrias otimizadas. Dependendo do tamanho da molécula, ou seja da quantidade de átomos e/ou elétrons, diversos níveis de cálculo podem ser realizados: cálculos quânticos [*Szabo*, 1996; *Atkins*, 1997; *Levine*, 1991] podendo ser *ab initio*, semiempírico e funcional da densidade; e cálculos clássicos [*Jensen*, 1999; *Leach*, 2001] em geral através do método de mecânica molecular utilizando campos de força empíricos.

O procedimento comum para incluir as interações com o meio é através dos modelos contínuo ou discretizado de solvente. O método mais popular dentro dos modelos contínuos é o de Campo de Reação Auto-

Consistente (SCRF)[1] [*Tapia*, 1975; *Rivail* 1976; *Karelson*, 1992; *Tomas* 2000] implementado nos cálculos quânticos. Já os modelos discretizados são os comumente implantados nas simulações computacionais [*Allen*, 1989] utilizando os campos de força empíricos [*Allinger*, 1977; *Weiner*, 1986; *Jorgensen*, 1990].

Neste trabalho estamos interessados em estudar propriedades estruturais e eletrônicas de moléculas em solução que tenham papel importante em processos biológicos. Em particular escolhemos a α-ciclodextrina (α-CD) como objeto do nosso estudo. Essa molécula pertence à família das ciclodextrinas que é de grande interesse em pesquisas aplicadas e tecnológicas (como referência geral ver o volume 98 do Chemical Reviews que é inteiramente dedicado a essa família de moléculas).

Na próxima seção iremos falar um pouco sobre carboidratos, que é um dos grupos de moléculas mais importantes nos sistemas biológicos e no qual se enquadra a α-ciclodextrina.

O objetivo do nosso estudo é entender o papel das moléculas de água na estabilização da geometria da α-ciclodextrina em solução aquosa. Adicionalmente através de um estudo teórico sistemático e criterioso e com uma comparação dos resultados obtidos com os resultados experimentais disponíveis, analisaremos a confiabilidade dos campos de força empíricos atualmente utilizados e a aplicabilidade das geometrias otimizadas com cálculos quânticos. Para isso dividimos nosso estudo em três etapas:

1. Utilizamos vários métodos clássicos (AMBER, OPLS e MM2) e métodos quânticos (Hartree-Fock, AM1 e PM3) para determinar a geometria mais estável da α-ciclodextrina isolada. Analisamos as propriedades estruturais como distâncias e ângulos de ligação, ângulos de torção, distâncias e ângulos que definem o formato cônico da α-CD e suas ligações de hidrogênio intramoleculares formadas com as hidroxilas primárias e secundárias. Analisamos também o momento de dipolo que reflete a distribuição eletrônica na molécula (capítulo 2).

2. Através da simulação computacional da α-ciclodextrina em água, analisamos a solvatação hidrofílica a partir da formação das ligações de

[1] As siglas serão utilizadas aqui na forma em que são mais conhecidas, o que significa que pode-se ter siglas que representam termos em inglês.

hidrogênio intermoleculares entre as moléculas de água e as hidroxilas primárias e secundárias da α-CD e analisamos também o caráter hidrofóbico da cavidade interna da molécula. Adicionalmente analisamos a competição das ligações de hidrogênio intra e intermoleculares (capítulo 3).

3. Utilizamos a teoria de Perturbação Termodinâmica aliada à simulação computacional para calcular a variação de energia livre e determinar a geometria mais estável em solução aquosa dentre as obtidas na primeira etapa (capítulo 4).

Ao final apresentaremos um sumário de nossos resultados e as conclusões no capítulo 5.

1.2. Carboidratos

Os carboidratos são a classe de moléculas orgânicas mais abundantemente encontrada na natureza (graças à abundância da celulose, presente nas células de todas as plantas, e do amido usado como sistema de armazenagem de energia pelas mesmas). Estas moléculas estão envolvidas em inúmeros processos biológicos, sozinhas ou em glicoconjugados (glicolipídios e glicoproteínas). Os glicoconjugados são componentes importantes das membranas celulares e estão envolvidas em processos de reconhecimento que determinam, dentre outras coisas, o crescimento e a morte celular e são responsáveis pela adesão celular e pelo transporte de substâncias para dentro das células [*Garret*, 1995].

Os carboidratos estão geralmente classificados em três grupos: monossacarídeos, oligossacarídeos e polissacarídeos. Descreveremos resumidamente estes grupos e suas propriedades gerais nas próximas subseções. Para uma revisão geral sobre carboidratos pode-se consultar os livros texto de Garrett e Grisham [*Garrett*, 1995] e de Nelson e Cox [*Nelson*, 2000].

1.2.1. Monossacarídeos

Os monossacarídeos, também chamados de açúcares simples, são moléculas que têm fórmula $C_nH_{2n}O_n$ com $n \geq 3$. Eles são as estruturas básicas de todos os carboidratos. A fórmula acima pode ser escrita como $(CH_2O)_n$

de onde vem o nome carboidrato dado aos compostos formados por estas unidades básicas e seus derivados. Entre os monossacarídeos temos as *aldoses* e as *cetoses*, dependendo se este for um aldeído ou uma cetona. Na Figura 1.1 mostramos os menores monossacarídeos existentes que são o Gliceraldeído e a Diidroxiacetona que são *trioses*, ou seja, aldoses e cetoses com 3 átomos de carbono. Os carbonos nos açúcares simples são numerados a partir daquele que se encontra na ponta mais perto do grupo carbonila.

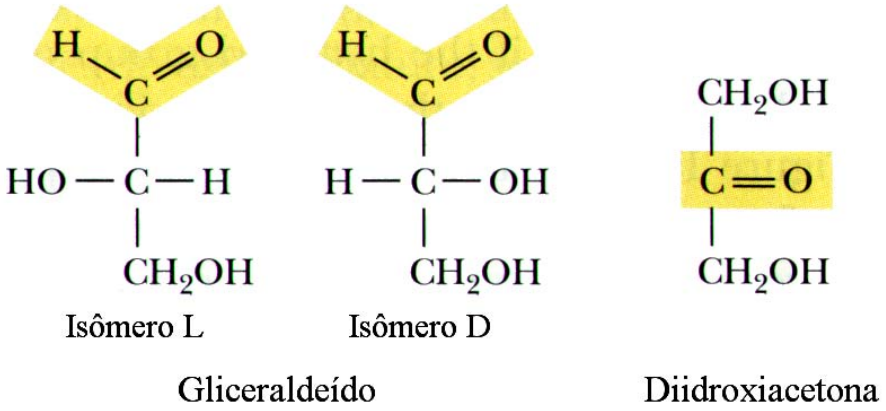


Figura 1.1: As trioses são os monossacarídeos mais simples. No caso do gliceraldeído, existem duas possibilidades D- e L-. Para a diidroxiacetona há apenas uma forma pois não há nenhum carbono assimétrico.

Existem vários tipos de monossacarídeos que compõem os carboidratos em geral, mas os mais comuns são as *hexoses*. Nas Figura 1.2 e na mostramos os monossacarídeos que podem ser formados com 3, 4, 5 e 6 carbonos. Note que todos estão na forma D-, ou seja, têm o grupo OH ligado ao carbono assimétrico mais distante do grupo carbonila, na mesma posição do D-gliceraldeído. Isso porque, assim como os aminoácidos (que ocorrem na natureza predominantemente na forma L- e apenas nessa forma são ativos) também os monossacarídeos ocorrem predominantemente na forma D- apesar de alguns monossacarídeos na forma L- (principalmente L-galactose e L-arabinose) desempenharem algumas funções biológicas específicas.

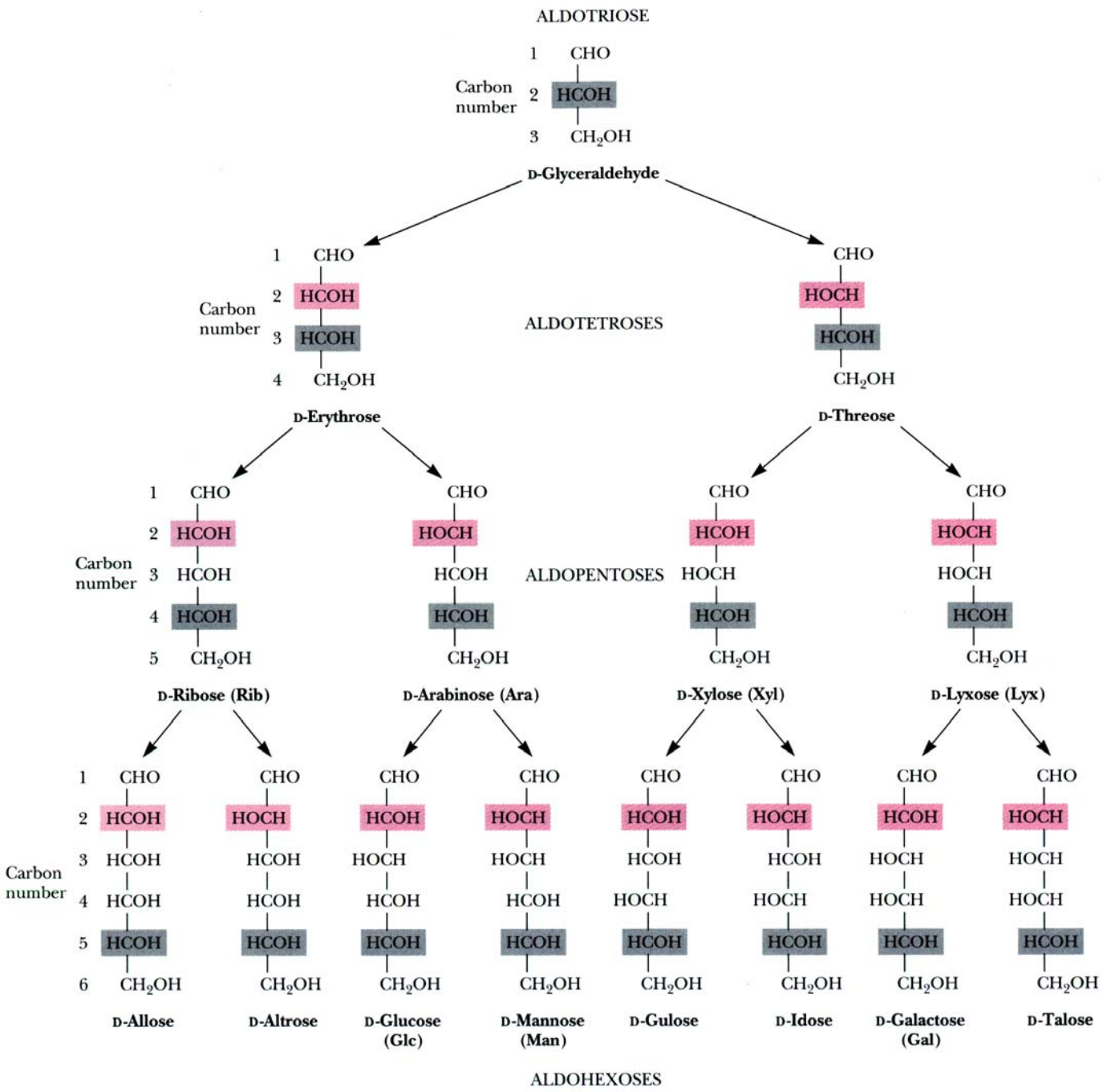


Figura 1.2: Aldoses com 3 a 6 carbonos. Dessas as mais comuns são as hexoses. As aldoses mais importantes possuem abreviaturas conhecidas dadas entre parênteses. O nome dos monossacarídeos aqui estão em inglês. Figura tirada de [*Garret*, 1995].

Pouco se sabe sobre os motivos que levaram a esta predominância de uma e outra forma nos dois casos, mas aparentemente essas preferências se estabeleceram de forma aleatória no inicio da evolução e se mantiveram graças à estereoespecificidade das enzimas que sintetizam e metabolizam estas moléculas.

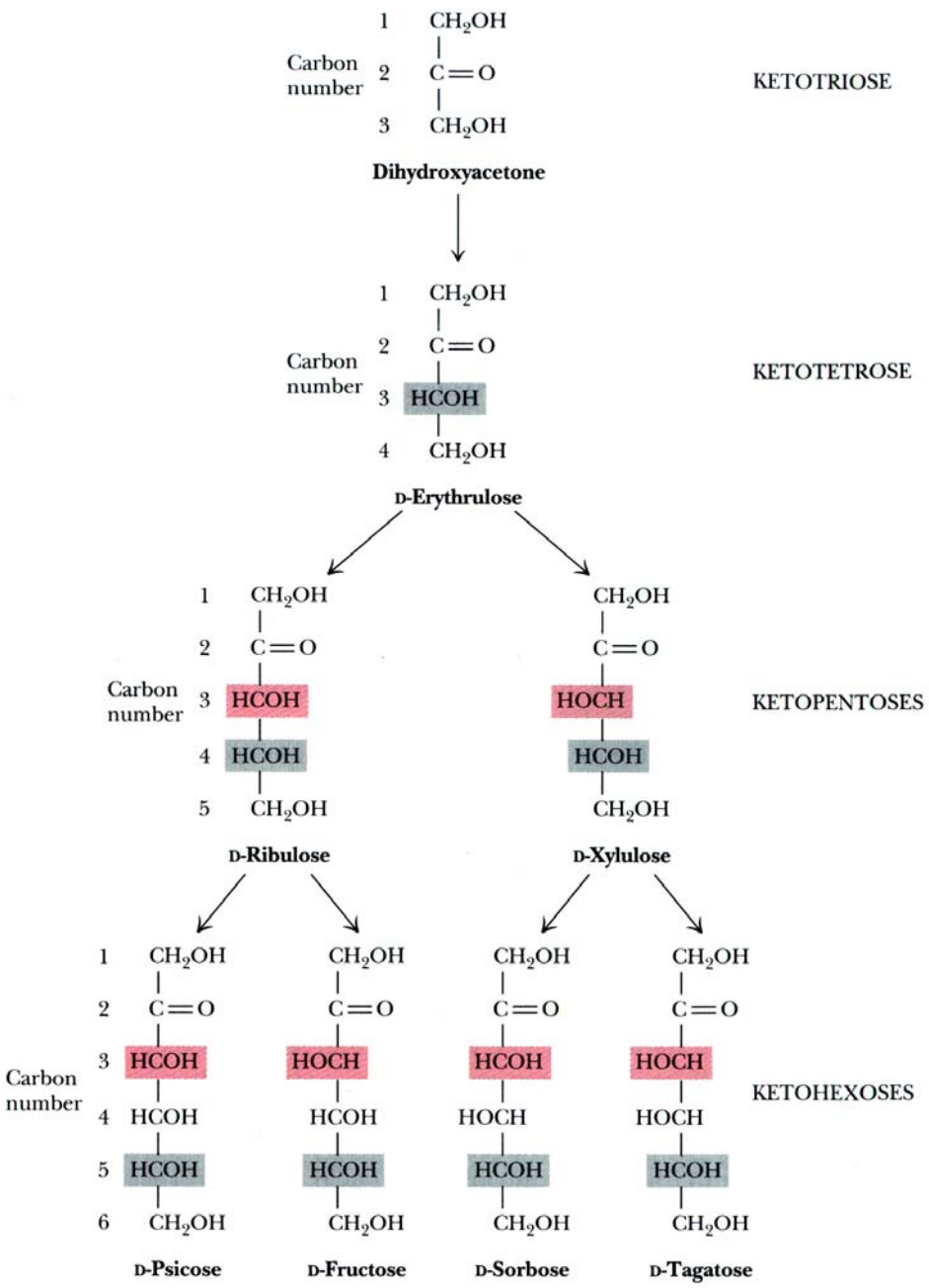


Figura 1.3: Cetoses com 3 a 6 carbonos. As cetoses têm um carbono assimétrico a menos que as aldoses, por isso existem em menor número. A principal cetose é a frutose cuja abreviatura é Fru. Os nomes estão em inglês. Figura retirada de [*Garret*, 1995].

Por simplicidade, costuma-se representar as estruturas desses açúcares na forma linear, mas as aldotetroses e todos os monossacarídeos com mais de quatro carbonos, quando em solução aquosa, ocorrem predominantemente em uma forma cíclica onde o carbono do grupo carbonila forma ligação covalente com o oxigênio de uma das hidroxilas ligadas aos carbonos mais distantes. A formação dessa estrutura cíclica resulta de uma rea-

ção entre aldeídos ou cetonas e álcoois formando hemiacetais ou hemicetais. Na Figura 1.4 mostramos esse processo de ciclização envolvendo uma aldose (D-glicose) e na Figura 1.5 mostramos o processo com uma cetose (D-frutose). Nesse processo pode-se formar uma estrutura com cinco ou seis átomos no anel (um deles sendo o oxigênio que forma a nova ligação covalente). Anéis com menos ou mais componentes que isto são desfavoráveis por serem estruturas tencionadas.

Essas formas cíclicas com anéis de cinco e seis componentes se assemelham a duas moléculas orgânicas cíclicas que são o furano (C_4H_4O) e o pirano (C_5H_6O). Por isso as formas cíclicas dos monossacarídeos são chamadas de furanoses ou piranoses. A favorabilidade de uma ou outra forma depende de fatores estruturais, mas em geral a forma piranosídica é mais favorável às aldoses enquanto a forma furanosídica é mais adotada pelas cetoses.

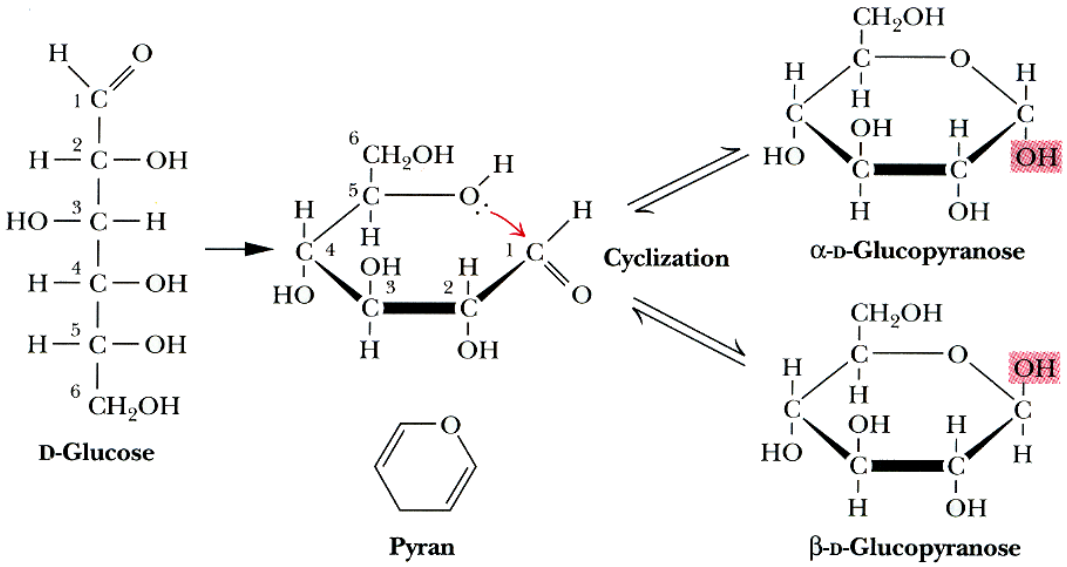


Figura 1.4: Processo de ciclização da D-glicose. Os monossacarídeos reagem internamente formando estruturas cíclicas principalmente com 5 e 6 átomos no anel. Assumem assim forma semelhante à do pirano ou do furano, respectivamente e passam a se chamar piranoses ou furanoses. Tirada de [*Garret*, 1995].

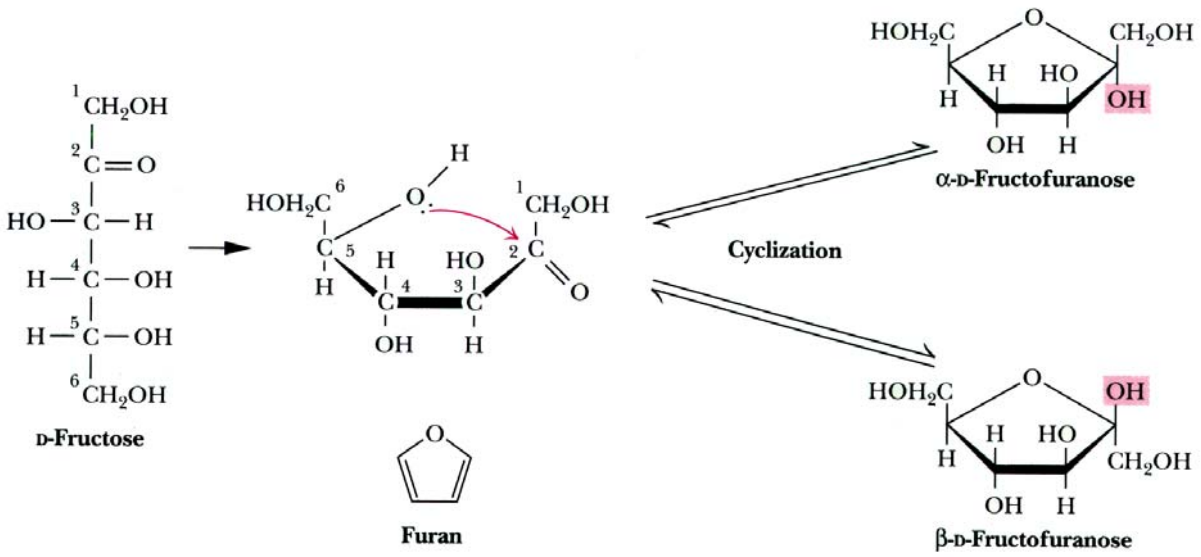


Figura 1.5: Processo de ciclização da D-frutose. Aqui o carbono assimétrico de maior numeração é o quatro e o carbono anomérico é o dois. Tirada de [*Garret*, 1995].

Após a reação de ciclização, o carbono outrora pertencente ao grupo carbonila passa a ser também um carbono assimétrico e duas formas estruturais cíclicas podem ser formadas . Esses isômeros são chamados *anômeros*, e o novo carbono assimétrico é chamado carbono *anomérico*. Quando a hidroxila ligada ao carbono anomérico se projeta no mesmo lado, em relação ao anel, da hidroxila ligada ao carbono assimétrico de maior numeração, o isômero é chamado α. Caso contrário é chamado β. Na Figura 1.4 mostramos as duas possibilidades para o caso da D-glicopiranose. Nesse caso o mais alto carbono assimétrico ligado a hidroxila na estrutura cíclica é o carbono 4. Na Figura 1.5 mostramos os dois isômeros para a D-frutofuranose. Nessa estrutura, o carbono assimétrico mais alto ligado a hidroxila também é o de número 4. Contudo, como a frutose é uma cetose, o carbono anomérico é o de número 2, ao contrário das aldoses onde o carbono anomérico é sempre o primeiro carbono da cadeia.

Tanto os anéis de piranoses quanto de furanoses não assumem forma plana, graças às ligações tetraédricas formadas pelos carbonos. No caso das piranoses, esses açúcares adotam dois tipos de conformações preferenciais que são a conformação bote e cadeira. A conformação bote é vista somente quando há grupos muito extensos ligados ao anel. Normalmente, a conformação cadeira é mais estável principalmente quando os grupos maiores li-

gados ao anel estão nas posições equatoriais. De todas as D-aldohexoses, a β-D-glicose é a única que pode adotar uma conformação com todos os grupos maiores (no caso grupos OH e CH_2OH) em posições equatoriais quando na forma cíclica. Por isso a β-D-glicopiranose é a molécula orgânica de maior ocorrência na natureza e é também a principal hexose envolvida no metabolismo de carboidratos.

Uma experiência pode ilustrar a estabilidade destas estruturas em geral. Se fizermos uma solução aquosa de D-glicose, iremos detectar que, após algum tempo, cerca de um terço da solução é composta de α-D-glicopiranoses, dois terços de β-D-glicopiranoses e uma porção muito pequena de D-glicofuranose e D-glicose não cíclica.

Os derivados de monossacarídeos também desempenham papéis biológicos importantes. A 2-desoxi-D-ribose, que é uma D-ribose com a hidroxila do carbono 2 substituída por hidrogênio, é constituinte do DNA de todos os seres vivos. Alguns açúcares alcoólicos como o glicerol, derivado do gliceraldeído, são componentes de lipídios. Vários outros derivados como ácido-, éster- e amino-açúcares são também produzidos por uma variedade de reações enzimáticas e estão presentes em polissacarídeos estruturais e em glicoconjugados.

1.2.2. Oligossacarídeos

Oligossacarídeos são carboidratos formados pela união de um pequeno grupo de monossacarídeos. Dada a enorme variedade de oligo e polissacarídeos, é surpreendente que eles sejam formados por um número pequeno de unidades básicas (como ocorre também no caso das proteínas) principalmente de glicose, manose, galactose, frutose, ribose e xilose.

Os oligossacarídeos mais simples e mais comuns são os dissacarídeos, dentre os quais os mais conhecidos são: a sacarose (conhecida também como açúcar de cana) produzida por várias plantas e largamente usada na culinária; a lactose que está presente em grande quantidade no leite e é de grande importância nutricional para os mamíferos nos primeiros anos de vida; e a maltose que está presente em bebidas fermentadas como a cerveja.

Os dissacarídeos são formados através da formação de uma ligação covalente entre o carbono anomérico de um açúcar e o oxigênio de alguma hidroxila de outra molécula de açúcar. Essa reação é um tipo de síntese desidratativa (ou condensação) que acontece entre hemiacetais ou hemicetais e álcool formando acetais ou cetais e água. A ligação formada dessa forma entre o carbono anomérico de piranoses e furanoses e o oxigênio do álcool é chamada ligação *glicosídica*. De certa forma essa é uma reação semelhante àquela da formação da ligação peptídica, tanto na formação da água como produto, quanto no fato de ambas as ligações não se estabelecerem diretamente.

Os oligossacarídeos são então perfeitamente discriminados pelas unidades que o compõem e pelas ligações glicosídicas estabelecidas entre eles. A maltose, por exemplo, é formada por uma molécula de α-D-glicose e outra molécula de D-glicose através de uma ligação 1→4, ou seja, o carbono anomérico da α-D-glicose forma ligação covalente com o oxigênio ligado ao carbono 4 da outra glicose. Daí escrevemos Glc-α-1,4-Glc como sendo a forma da maltose, podendo ser α-maltose ou β-maltose segundo a outra molécula de glicose seja α- ou β-. A sacarose é Glc-α-1,2-β-Fru e é única pois a ligação glicosídica envolve os dois carbonos anoméricos. A galactose, por sua vez, é Gal-β-1,4-Glc podendo ser α- ou β- segundo a glicose seja α- ou β-. Na Figura 1.6 mostramos os principais dissacarídeos como ilustração das ligações glicosídicas que formam os oligossacarídeos.

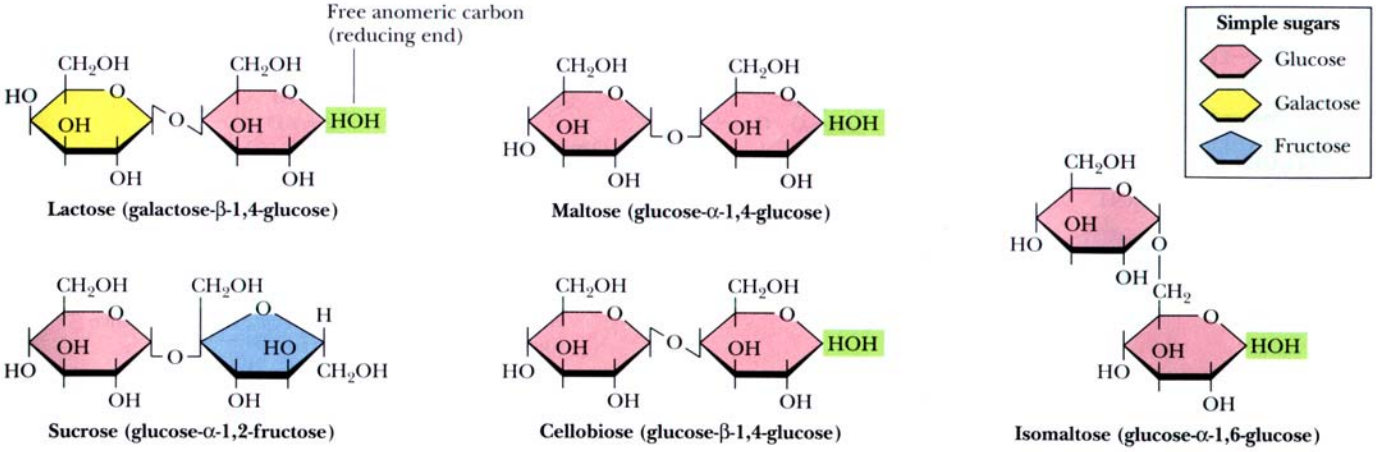


Figura 1.6: Ilustração de alguns dissacarídeos. A ponta redutora é aquela que possui um carbono anomérico livre. Figura tirada de [*Garret*, 1995].

Outros oligossacarídeos maiores estão presentes naturalmente nos seres vivos e desempenhas inúmeros papéis. Adicionalmente uma gama de oligossacarídeos está presente em antibióticos e alguns até mesmo mostram atividade anti-tumoral.

1.2.3. Polissacarídeos

A grande maioria dos carboidratos ocorre na natureza em forma de polissacarídeos. Polissacarídeos são estruturas formadas por dezenas ou centenas de monossacarídeos e derivados. Essas estruturas têm inúmeras funções biológicas sendo as principais de armazenagem de energia, de formação de estruturas de sustentação e como substâncias protetoras. Esses polímeros são geralmente homopolímeros ou polímeros com poucas unidades básicas diferentes. Oligômeros ou polímeros com muitas unidades diferentes aparentemente estão envolvidos em processos de reconhecimento celular e eventos de comunicação intercelular [*Nelson*, 2000, Cap. 9 e 13].

Como componentes em estruturas de sustentação os principais polissacarídeos são a celulose - que é a forma mais abundante dos carboidratos - na parede celular das plantas e o quitino no exoesqueleto dos artrópodes. A celulose é um homopolímero linear formado por grandes cadeias de β-D-glicose unidas por ligações 1→4. Em sua situação mais estável as unidades de glicose ficam rotacionadas em torno da ligação glicosídica de 180º umas em relação às outras. Com isso forma-se uma cadeia linear com o mesmo número de hidroxilas dos dois lados. Através dessas hidroxilas, essas cadeias formam uma extensa rede de ligações de hidrogênio com outras cadeias formando uma espécie de folha de celulose. As folhas de celulose podem ainda interagir através de uma rede de ligações de hidrogênio interplanos formando assim uma estrutura muito resistente que dá a sustentação às plantas em geral.

O quitino também é um homopolímero com estrutura primária, secundária e terciária idênticas à da celulose. A diferença é que a unidade básica do quitino é um derivado da β-D-glicose onde a hidroxila ligada ao carbono 2 é substituída por um grupo $NHCOCH_3$. Este derivado é chamado N-acetil-D-glicosamina.

Dentro da função de armazenagem, uma das formas mais comuns de polissacarídeos é o amido, formado por uma conjunção de amilose e amilopectin presentes nos cloroplastos e amiloplastos. Amilose é um homopolímero formado por unidades de α-D-glicopiranose unidas por ligações 1→4. Amilopectin é também um homopolímero formado por resíduos de α-D-glicopiranose com ligações 1→4 que formam várias cadeias ramificadas através de ligações 1→6, ou seja, são várias amiloses unidas por ligações 1→6. A maior parte do amido (70 - 90%) é formada por cadeias ramificadas.

Estas ligações α(1→4) são tais que a conformação mais estável adotada por esses polissacarídeos não é linear como na celulose, mas sim circular de modo que em água essas longas cadeias formam uma estrutura helicoidal. Esta estrutura helicoidal da amilose e amilopectin, ilustrada na Figura 1.7, é precursora da forma das ciclodextrinas.

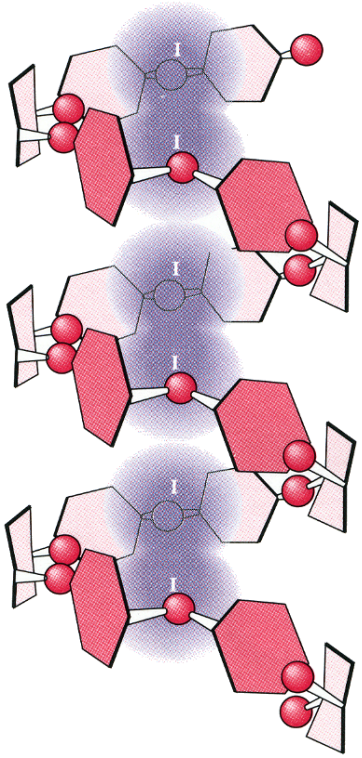


Figura 1.7: Amilose em água adota conformação helicoidal. Essa conformação está intimamente ligada à natureza das ligações glicosídicas entre as unidades de α-D-glicose. Essa conformação circular natural provocada por tais ligações é precursora da geometria das ciclodextrinas. Tirada de [*Garret*, 1995].

Em animais o polissacarídeo de armazenagem é o glicogênio presente principalmente no fígado e nos músculos. É formado por unidades de α-D-glicose e possui estrutura semelhante ao amilopectin, mas com mais ramificações e mais compacto.

1.2.4. Ciclodextrinas

As ciclodextrinas (CD), também chamadas de cicloamiloses, são oligossacarídeos cíclicos formados por unidades (ou resíduos como se costuma dizer) de α-D-glicopiranose ligadas através de ligações glicosídicas 1→4. Essas moléculas são formadas a partir da degradação de amido por enzimas (glicosiltransferases) presentes em uma variedade de bactérias principalmente do gênero *bacilus*.

Na degradação, essas enzimas quebram várias ligações glicosídicas entre os resíduos de glicopiranose na estrutura helicoidal do amido e provocam a formação de novas ligações entre as extremidades dos oligossacarídeos resultantes tornando-os cíclicos e dando origem assim às ciclodextrinas.

As CDs são moléculas tão sedutoras que despertam o interesse tanto nas pesquisas aplicadas como nas pesquisas tecnológicas [*Szejtli*, 1998]. Essa família de moléculas é tão incomum que transcende as barreiras tradicionais que separam as subáreas da ciência como biologia, química e física. Alguns dos aspectos mais intrigantes das CDs são sua forma circular gerada por ligações entre monossacarídeos, a possibilidade de hospedar outras moléculas em sua cavidade circular e sua especificidade estereoquímica. Estes três aspectos dão origem a várias aplicações. Entre elas temos:

- Catalisam reações orgânicas e reações biomiméticas (atuando como enzimas artificiais) [*Breslow*, 1998];

- Sua cavidade hidrofóbica serve como um campo de reação através do qual mediam reações orgânicas [*Takahashi*, 1998].

- Podem formar complexos de inclusão com uma variedade de moléculas inclusive fármacos, servindo como mecanismo de transporte e aumentando sua solubilidade e estabilidade [*Uekama*, 1998].

Estas propriedades podem ser melhoradas através da utilização de derivados específicos das ciclodextrinas.

Um volume de Chemical Reviews de 1998 foi especialmente dedicado às CDs. Lá pode-se encontrar várias informações sobre essa família de moléculas como: espectro NMR [*Schneider*, 1998]; cristalografia de raio-X [*Saenger*, 1998]; estrutura cristalina [*Harata*, 1998]; cálculos teóricos [*Lipkowitz*, 1998]; propriedades termodinâmicas [*Rekharsky*, 1998]; rotas de síntese [*Gattuso*, 1998]; aceitação e posicionamento de moléculas hospedeiras [*Nepogodiev*, 1998]; modificações substitucionais [*Khan*, 1998]; reações catalíticas [*Breslow*, 1998]; mediação de reações [*Takahashi*, 1998]; aplicações industriais [*Hedges*, 1998] e transporte de drogas [*Uekama*, 1998].

Dentre todas as moléculas que compõem a família das CDs, três são de grande interesse nas aplicações citadas acima. São elas a α-ciclodextrina, β-ciclodextrina e γ-ciclodextrina, que diferem no número de monossacarídeos (α-D-glicopiranoses). A α-CD é formada por 6 moléculas de glicose num total de 126 átomos; a β-CD tem 7 glicoses e soma 147 átomos e a γ-CD tem 8 glicoses com um total de 168 átomos. Na Figura 1.8 mostramos uma ilustração da α-CD e suas dimensões principais.

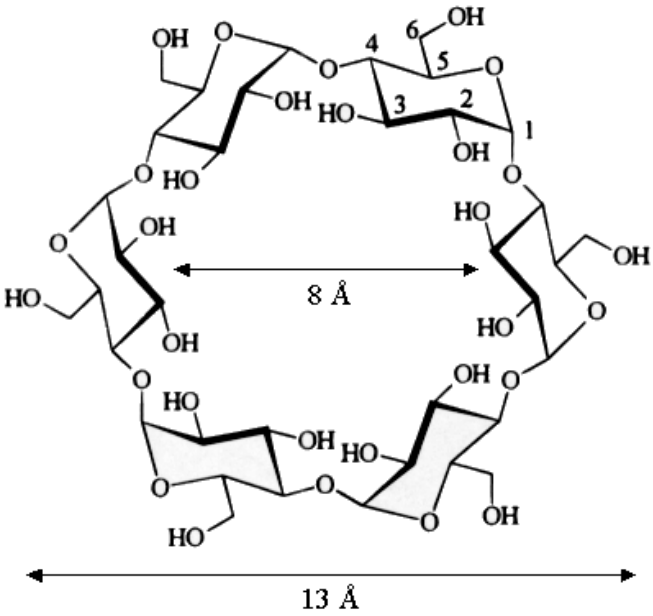


Figura 1.8: Esquema ilustrativo da α-ciclodextrina. O diâmetro externo é de aproximadamente 13 Å, o diâmetro interno é em torno de 8 Å. A altura do tronco de cone formado pela molécula não é mostrado mas é de aproximadamente 8 Å.

Por serem moléculas com muitos átomos, as CDs foram pouco estudadas teoricamente. Os cálculos teóricos das CDs estão divididos em duas categorias: cálculos quânticos de otimização de geometria em nível semiempírico, realizada para determinar a estrutura mais estável da molécula isolada [*Li*, 2000] e no estado complexado com moléculas hospedeiras [*Santos*, 2000]; e cálculos com campos de força empíricos, também realizados para determinar a estrutura mais estável das CDs no estado complexado com moléculas hospedeiras e para realizar simulações computacionais em solução aquosa [*Momany*, 2000a; 2000b]

Em ambas as categorias, foram analisadas as mudanças conformacionais das CDs com métodos clássicos e quânticos em nível semiempírico. Em nenhum momento foram realizados cálculos de otimização de geometria em nível *ab initio*, como também não foi analisado o papel da água como fator de estabilização conformacional e a competitividade das ligações de hidrogênio intra e intermoleculares.

Capítulo 2

α-Ciclodextrina Isolada

Neste capítulo discorreremos a respeito dos métodos utilizados na determinação de geometria de moléculas isoladas. Descreveremos as metodologias teóricas e apresentaremos os resultados correspondentes. Ao final faremos uma comparação entre as geometrias da α-ciclodextrina obtidas com métodos clássicos e quânticos.

2.1. Introdução

Uma das propriedades de maior interesse no estudo de moléculas é a energia. Dessa propriedade decorrem inúmeros conceitos e propriedades derivadas como estabilidade, geometria de equilíbrio, reatividade, estados de transição, barreiras de ativação, constantes de equilíbrio, etc.

Como dissemos, existem vários métodos teóricos de se calcular a energia de um sistema molecular. Dentre eles estão os métodos clássicos e os métodos quânticos. Descreveremos brevemente os métodos clássicos utilizados bem como os quânticos e apresentaremos os resultados obtidos. Posteriormente iremos analisar os resultados obtidos com os dois métodos.

2.2. Mecânica Molecular

O método clássico usado para cálculo de energia de moléculas é chamado Mecânica Molecular e utiliza campos de força empíricos que mostraremos adiante. Esses campos de força se baseiam na representação dos elétrons através do seu efeito coligativo entre os núcleos, ou seja, através da modelagem das ligações químicas e físicas que são formadas por seu intermédio. Essas ligações são representadas por interações mecânicas (daí o nome mecânica molecular) do tipo massa-mola, por razões que iremos mostrar, onde os elétrons fariam então o papel de molas unindo as massas que são os núcleos. É claro que não só esse efeito de ligar os núcleos é modelado como também todos os efeitos eletrônicos devem ser levados em conta.

O principal ganho com tal modelagem das interações entre núcleos e elétrons é a utilização de um funcional analítico para a energia do sistema o que, de longe, é de menor custo computacional do que qualquer cálculo quântico. Como conseqüência o seu uso para cálculo de energia em macromoléculas tem sido muito difundido.

Um dos fatos que mais encorajou a criação de tais campos de força é a observação de que diferentes moléculas tendem a ser formadas por unidades estruturalmente similares. Por exemplo, o comprimento da ligação C–H varia pouco para as várias moléculas conhecidas (entre 1.06 e 1.10 Å) e pode ser considerado constante. A freqüência de vibração dessa ligação também fica numa faixa estreita entre 2900 e 3300 cm^{-1}. Se separamos essas ligações em grupos diferentes em que o carbono forma ligação simples, dupla e tripla esses intervalos são menores ainda.

Os campos de força são baseados em três conceitos principais que são: a forma funcional usada para modelar as interações mediadas pelos elétrons; os tipos de átomos que servem para diferenciar o mesmo átomo em situações diferentes; e o conjunto de parâmetros associados aos tipos de átomos.

A forma funcional de cada campo de força é um conjunto de termos que descrevem movimentos específicos da molécula. Em geral esta forma funcional é dada por

$$E = E_{\text{str}} + E_{\text{bend}} + E_{\text{tors}} + E_{\text{oop}} + E_{\text{vdW}} + E_{\text{el}} + E_{\text{cross}} \tag{2.1}$$

onde os termos componentes são, na ordem, energia de estiramento (stretching), de deformação angular (bending), de torção, a energia de deformação fora do plano, os termos não ligantes que são o de van der Waals e o eletrostático, e finalmente os termos cruzados reunidos no último termo. Descreveremos estes termos a seguir

2.2.1. Energia de estiramento

A energia de estiramento é a energia necessária para deformar uma ligação entre dois tipos de átomos A e B do seu comprimento de equilíbrio r_0. Variando o comprimento de ligação entre esses átomos obtemos uma curva de potencial que é bem descrita por uma exponencial como no potencial de Morse [*Morse*, 1929].

$$E_{\text{Morse}}(r) = D[1 - e^{\alpha(r-r_0)}]^2 \tag{2.2}$$

onde D é a energia de dissociação e α esta relacionado à rigidez no comprimento da ligação.

Um método mais usual é tomar uma expansão em série de Taylor em torno da distância de equilíbrio para a ligação r_0

$$E_{\text{str}}(r) = E(r_0) + \frac{dE(r_0)}{dr}\left(r - r_0\right) + \frac{1}{2}\frac{d^2E(r_0)}{d^2r}\left(r - r_0\right)^2 + \text{K} \tag{2.3}$$

$E(r_0)$ é tomado como zero e a primeira derivada é zero porque estamos expandindo em torno da distância de equilíbrio. Assim a energia de estiramento pode ser modelada da forma

$$E_{\text{str}}(r) = k_2^{AB}\left(r - r_0\right)^2 + k_3^{AB}\left(r - r_0\right)^3 + k_4^{AB}\left(r - r_0\right)^4 + \text{L} \tag{2.4}$$

onde os k_i são os parâmetros a serem ajustados. Em geral trunca-se até o termo em quarta ordem. O mais comum é tomar apenas o termo de segunda ordem, o que se denomina aproximação harmônica. Neste caso $k_2^{AB} = 2D\alpha^2$ é a constante de mola ou constante de força.

O uso mais amplo dos potenciais polinomiais harmônico (P2), cúbico (P3) e quártico (P4) se deve ao fato de serem esses mais rápidos de se computar e mais fáceis de se parametrizar, além de alguns inconvenientes a respeito do potencial de Morse, como dificuldades de convergência em algumas circunstâncias [*Jensen*, 1999]. Além disso, como se verifica que os comprimentos de ligação se modificam muito pouco em moléculas diferentes, uma aproximação harmônica ou cúbica, é muitas vezes suficiente para reproduzir geometrias experimentais.

2.2.2. Energia de deformação angular

A energia de deformação angular é a energia requerida para tirar um ângulo de ligação formado por três tipos de átomos A, B e C de seu valor de equilíbrio θ_0. Assim como na energia de estiramento, usualmente se expande o potencial relacionado ao ângulo de ligação em uma série de Taylor. Assim temos

$$E_{\text{bend}}(\theta) = k_2^{ABC}(\theta-\theta_0)^2 + k_3^{ABC}(\theta-\theta_0)^3 + k_4^{ABC}(\theta-\theta_0)^4 + \text{L} \qquad (2.5)$$

onde novamente os k_i são os parâmetros a serem ajustados. Em geral, como no caso da energia de estiramento, usa-se uma aproximação harmônica ou cúbica, o que descreve suficientemente bem os ângulos de ligação numa grande faixa de valores.

2.2.3. Energia de torção

A energia de torção é a energia associada à rotação em torno de uma ligação B–C. O ângulo de torção ou ângulo diedral é definido como aquele formado entre a ligação C–D e a ligação B–A.

Nesse caso o potencial é periódico, com período de no máximo 360º, e isso deve ser levado em conta porque as barreiras de rotação são bem pequenas em comparação com os poços de potencial para a distância de ligação e para o ângulo de ligação. Por isso não é adequado fazer uma expansão em série de Taylor. O que se faz então é expandir em uma série de Fourier

$$E_{\text{tor}}(\omega) = \sum_n V_n \cos(n\omega). \tag{2.6}$$

onde o termo $n = 1$ é responsável pela periodicidade em 360º, $n = 2$ é periódico em 180º e assim por diante. As constantes V_n determinam o valor das barreiras de rotação em torno da ligação B–C. Nos casos em que não houver periodicidade para alguns ângulos de rotação, as constantes correspondentes serão nulas.

2.2.4. Energia de deformação fora do plano

Quando um átomo B tem hibridização sp^2, sua conformação mais estável em relação aos átomos ao qual é ligado é planar e há um acréscimo de energia quando o arranjo formado por esses átomos sai dessa conformação.

Há duas maneiras de se computar isso. Uma é através do ângulo χ formado entre a ligação AB e sua projeção no plano ACD (Figura 2.1). Através desse ângulo faz-se geralmente uma aproximação harmônica da expansão em série de Taylor em torno de $\chi = 0$. Outra maneira de se levar isso em conta é fazendo uma expansão em série de Fourier como na equação (2.6) para o ângulo diedral formado pela seqüência ABCD, apesar de não haver a ligação C–D. Esse ângulo diedral é chamado de *impróprio* por esse motivo. É claro que a seqüência ABDC é equivalente a ABCD e qualquer desses ângulos pode ser usado. Há também a possibilidade de se tomar uma aproximação harmônica no ângulo diedral impróprio.

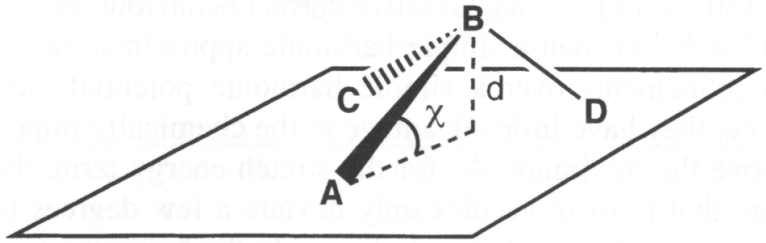


Figura 2.1: O átomo B tem hibridização sp^2 e portanto tende a ser planar com os átomos adjacentes. A não coplanaridade leva a um acréscimo de energia que é modelada através da energia fora do plano. Tirada de [*Jensen*, 1999]

2.2.5. Energia de van der Waals

Para dar conta de interações entre tipos de átomos que não estão ligados por nenhuma das interações descritas acima, usa-se os termos chamados não ligantes, que são os termos de van der Waals e eletrostático.

Em teoria de forças intermoleculares sabe-se que a interação entre duas espécies não quimicamente ligadas se dá através de diferentes componentes provenientes da separação dos fenômenos que lhes dão origem. Essas componentes estão separadas em termos de longo alcance, que decrescem mais lentamente com a distância entre as espécies interagentes, e termos de curto alcance, que decrescem rapidamente com a distância.

Os principais termos de longo alcance são a energia eletrostática, a energia de indução e a energia de dispersão. A energia eletrostática é proveniente da interação entre os momentos de multipolo permanentes das espécies. Essa energia pode ser atrativa ou repulsiva dependendo da orientação relativa entre as moléculas. A principal componente na energia eletrostática é a interação carga-carga. Para moléculas neutras e assimétricas, o termo dominante se torna a interação dipolo-dipolo. A energia de indução vem da interação entre os momentos de multipolo permanentes de uma molécula e os momentos induzidos de outra e é sempre atrativa. Finalmente a energia de dispersão vem da interação entre os multipolos induzidos das duas moléculas e também é sempre atrativa além de ser ubíqua.

A energia eletrostática será levada em conta através do termo homônimo que temos na nossa energia clássica. Já as energias de indução e de dispersão são levadas em conta através da energia de vdW.

Para moléculas não carregadas os termos dominantes da energia de indução e de dispersão decrescem com a sexta potência da distância. Isso dá origem ao termo atrativo da energia de vdW que tem a forma funcional r^{-6}.

A parte repulsiva vem das interações de curto alcance. A principal interação desta categoria é a energia de troca [*Stone*, 1996]. A forma desse potencial é aproximadamente exponencial, mas é comum utilizar uma forma

r^{-n} que apresente a característica principal desta interação que é a forte repulsão a pequenas distâncias.

Assim há duas maneiras de se constituir a energia de van der Waals. Uma delas é o potencial de Buckingham ou de Hill dado por [*Hill*, 1948]

$$E_{\mathrm{Hill}}(r) = Ae^{-Br} - \frac{C}{r^6} \tag{2.7}$$

onde temos então três parâmetros para ajustar. A outra maneira é através do potencial de Lennard-Jones dado por [*Lennard-Jones*, 1937]

$$E_{\mathrm{LJ}}(r) = \frac{C_1}{r^{12}} - \frac{C_2}{r^6} \tag{2.8}$$

que também pode ser escrito da forma

$$E_{\mathrm{LJ}}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]. \tag{2.9}$$

De qualquer forma há dois parâmetros a serem ajustados no potencial de Lennard-Jones. Na equação (2.9) ε é a profundidade do poço de potencial e σ é a distância em que a interação se torna zero.

Um tratamento especial deve ser dado, entretanto, às ligações de hidrogênio. Como essas interações estão numa faixa de energia entre as interações de van der Waals e as interações químicas, uma forma diferente de interação não ligante deve ser adotada para representá-las. Apesar desse tipo de interação poder ser bem descrita através de interações eletrostáticas, a utilização dos funcionais usuais para a energia de vdW sem a adoção de parâmetros específicos não permitirá a maior proximidade entre o hidrogênio e os átomos eletronegativos não ligados a ele. Por isto substitui-se, em alguns campos de força, para esse fim específico, o funcional de LJ por um mais apropriado que geralmente é dado por

$$E_{\mathrm{HBond}}(r) = \frac{D_1}{r^{12}} - \frac{D_2}{r^{10}} \tag{2.10}$$

ou ainda por

$$E_{\mathrm{HBond}}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{10}\right]. \tag{2.11}$$

2.2.6. Energia eletrostática

O outro termo usado para tipos de átomos não ligados é o da energia eletrostática. Em uma molécula, em geral haverá partes que concentram mais elétrons que outras e isso pode ser representado classicamente por cargas parciais sobre os átomos.

A energia eletrostática tem papel fundamental nas interações intermoleculares. Ligações de hidrogênio são descritas classicamente principalmente através das cargas parciais nos átomos presentes nas moléculas, com uma ajuda do potencial de LJ modificado para dar um ajuste fino para essa interação como vimos na subseção anterior.

Como dissemos, o termo dominante na energia eletrostática é a interação carga-carga. A forma funcional portanto mais utilizada para esse termo é

$$E_{\mathrm{el}}(r) = \frac{q_A q_B}{r}. \tag{2.12}$$

As cargas devem em princípio ser encaradas como parâmetros a serem ajustados. Entretanto é mais comum o ajuste desses parâmetros ser feito separadamente através de ajuste de cargas pontuais aos momentos de multipolo calculados por métodos quânticos bastante acurados. Em alguns casos essas cargas são bem conhecidas como no caso da água.

Outra forma funcional menos utilizada é a interação dipolo-dipolo que na verdade é apenas um termo na expansão multipolar de (2.12). Nesse caso a energia eletrostática fica

$$E_{\mathrm{el}}(r) = \frac{\mu_A \mu_B}{r^3}\left(\cos\phi - 3\cos\alpha_A \cos\alpha_B\right) \tag{2.13}$$

onde r agora é a distância entre os dois dipolos (geralmente tomada como a distância entre os centros de massa dos mesmos), ϕ é o ângulo de torção entre os dois dipolos e α_A é o ângulo formado entre $\vec{\mu}_A$ e o vetor $\vec{r}$. Os

dipolos são geralmente formados por dois átomos que formam uma ligação química.

2.2.7. Termos cruzados

Os diferentes termos componentes da energia não são independentes. Por exemplo, quando mudamos o valor do ângulo de ligação da água, o comprimento de equilíbrio da ligação O–H muda. Quando diminuímos o ângulo de ligação o comprimento de equilíbrio aumenta para distanciar os hidrogênios. Já quando aumentamos o ângulo de ligação o comprimento de equilíbrio diminui. Também o aumento ou diminuição na distância de uma ligação influi na ligação adjacente. Essa é a origem dos termos cruzados que são adotados por alguns campos de força para aumentar a precisão na descrição geométrica das moléculas. Os valores mais afetados nessa correlação são das distâncias e dos ângulos de ligação. Por isto, o termo cruzado mais importante é o que combina estiramento com deformação angular. Os termos cruzados são obtidos em geral através do produto dos termos individuais envolvidos. No caso de acoplamento entre estiramento e deformação angular esse termo é dado por

$$E_{\text{str/bend}}(r^{AB}, r^{BC}, \theta) = k^{ABC}\left(\theta - \theta_0\right)\left[\left(r^{AB} - r_0^{AB}\right) + \left(r^{BC} - r_0^{BC}\right)\right]. \qquad (2.14)$$

Outros termos podem ser escritos de forma semelhante.

2.2.8. Tipo de átomo

O tipo de átomo é um conceito que tenta definir o ambiente químico do mesmo na molécula. As principais informações usadas na definição dos tipos de átomos são o número atômico dos átomos que se ligam a este átomo e o tipo das ligações (simples, duplas ou triplas). Assim um átomo de carbono, por exemplo, pode ser classificado em vários tipos segundo sua hibridização, se está ligado a outros carbonos ou se está ligado a oxigênio. O oxigênio também tem propriedades diferentes se está entre dois carbonos (tipo éster), se está ligado a um hidrogênio (álcool) ou se faz parte de um grupo carboxila.

2.2.9. Parametrização

Os conjuntos de parâmetros associados a esses tipos de átomos são obtidos através de ajuste das constantes do funcional através de inúmeros dados experimentais. Entretanto esta não é uma tarefa muito simples. Em geral é necessário pelo menos 3 a 4 dados experimentais independentes para cada parâmetro. Além disso, é difícil encontrar dados experimentais sobre barreiras rotacionais pois, em geral, os experimentos são feitos apenas perto da geometria de equilíbrio.

Por isto os campos de força mais recentes têm se aproveitado de cálculos de estrutura eletrônica de alto nível para ajustar seus parâmetros. Estes são chamados de campos de força Classe II. Outra observação importante a se fazer é sobre o conjunto de moléculas utilizado para a parametrização. Alguns campos de força são mais abrangentes tendo sua parametrização baseada numa grande variedade de moléculas e por isto podem ser usados com boa precisão para cálculos de energia em várias moléculas. Outros campos de força baseiam sua parametrização em tipos específicos de moléculas como é o caso de proteínas e ácidos nucléicos principalmente. Assim esses campos de força apresentam resultados mais precisos para moléculas desses tipos e menos precisos para outras moléculas.

2.2.10. Principais campos de força empíricos

Existem vários campos e força empíricos disponíveis. Os mais conhecidos são os métodos: MM2 desenvolvido por Allinger e colaboradores [*Allinger*, 1977] inicialmente para uma gama de moléculas orgânicas pequenas e posteriormente estendido para peptídios [*Lii*, 1982] e outras moléculas [*Lipkowitz*, 1992]; AMBER desenvolvido pelo grupo de pesquisa de Peter Kollman para proteínas e ácidos nucléicos inicialmente do tipo *átomos unidos*[1] [*Weiner*, 1984] com uma versão *todos os átomos*[2] posterior [*Weiner*, 1986] e depois estendido também para carboidratos [*Hohmans*, 1990]; OPLS desenvolvido pelo grupo de Jorgensen também

[1] Em inglês "united atom".
[2] Em inglês "all atom".

para proteínas [*Jorgensen*, 1988] e para ácidos nucléicos [*Pranata*, 1991] e outros sistemas [*Jorgensen*, 1990]; e GROMOS desenvolvido por van Gunsterenm e colaboradores também para proteínas, ácidos nucléicos e carboidratos.

Desses utilizamos os três primeiros através do programa Hyperchem [*Hyperchem*, 1996]. Os funcionais que cada um utiliza podem ser resumidos na Tabela 2.1. O método MM2 implementado no programa tem forma funcional original [*Allinger*, 1977] mas utiliza os tipos de átomos e parâmetros mais recentes MM2(1991). Essa versão também é conhecida como MM+. As versões dos campos de força implementados no programa Hyperchem são: o MM2 parametrizado para moléculas orgânicas pequenas em geral, principalmente hidrocarbonetos; o AMBER possui várias versões sendo que a utilizada foi o AMBER94 parametrizado para proteínas, ácidos nucléicos e polissacarídeos; o OPLS parametrizado principalmente para proteínas e ácidos nucléicos.

Tabela 2.1: Resumo dos termos pertencentes aos funcionais de cada campo de força.

	E_{str}	E_{bend}	E_{tor}	E_{oop}	E_{vdW}	E_{HBond}	E_{el}	E_{cross}
MM2	P2+3	P2+6	F123	P2	Exp-6	-	dd	sb
AMBER	P2	P2	F1-6	Imp	12-6	12-10	qq	-
OPLS	P2	P2	F1-6	Imp	12-6	-	qq	-

Como podemos ver na Tabela 2.1 o método MM2 utiliza aproximação cúbica para a energia de estiramento com um termo harmônico também. A energia de deformação angular possui um termo harmônico e um sêxtico, a energia fora do plano é computada como um termo harmônico no ângulo χ. A energia de torção possui os termos $n = 1,2,3$. A energia de vdW é dada pelo potencial de Hill e a energia eletrostática pela interação dipolo-dipolo tomando como centro do dipolo o ponto médio da ligação ao invés do centro de massa dos dois átomos. Além disso somente esse campo de força possui termo cruzado, nesse caso referente ao acoplamento entre distância e ângulo de ligação.

Já os métodos AMBER e OPLS possuem quase a mesma forma funcional. Ambos usam aproximação harmônica para o estiramento e para a deformação angular. Para a energia de torção é empregada uma série de Fourier com termos $n = 1$ até 6 dependendo dos tipos de átomos envolvidos. A energia fora do plano é calculada através de ângulo diedral impróprio com série de Fourier. Em ambos os casos a energia de vdW é dada pelo potencial de Lennard-Jones. A energia de ligação de hidrogênio no AMBER é dado pelo potencial de LJ com expoente 10 para o termo atrativo. Também ambos usam a interação carga-carga para a energia eletrostática.

2.3. Grandezas geometricas da α-CD

Antes de analisarmos os resultados, precisamos falar um pouco da geometria da α-CD. A numeração tradicional da glicose pode ser vista na Figura 2.2. Os oxigênios O2 e O3 pertencem às hidroxilas secundárias e o oxigênio O6 faz parte da hidroxila primária.

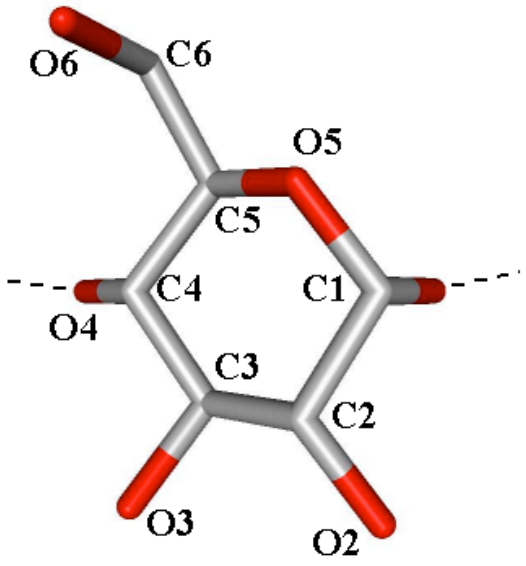


Figura 2.2: Numeração da unidade de glicopiranose. Os hidrogênios foram suprimidos em todas as figuras por clareza.

A partir dessa numeração podemos definir novas grandezas geométricas na molécula de α-CD. Todos os comprimentos de ligação na α-CD e os ângulos de ligação internos das unidades de glicopiranose mudam

muito pouco, de modo que essas unidades podem ser consideradas como blocos bastante rígidos [*Saenger*, 1998], a menos das hidroxilas primárias (O6 e H6) que têm certa liberdade torsional. Assim, algumas poucas grandezas geométricas podem caracterizar a estrutura da α-CD.

Um dos graus de liberdade restantes para a α-CD é a torção relativa entre as unidades de glicose intermediada pela ligação glicosídica. Essa torção relativa pode ser medida por dois ângulos diedrais Φ e Ψ. O ângulo Φ é aquele formado pelos átomos C4-O4-C1′-O5′, onde C1′ é o carbono anomérico da glicose adjacente e O5′ também pertence à glicose adjacente (ver Figura 2.3)

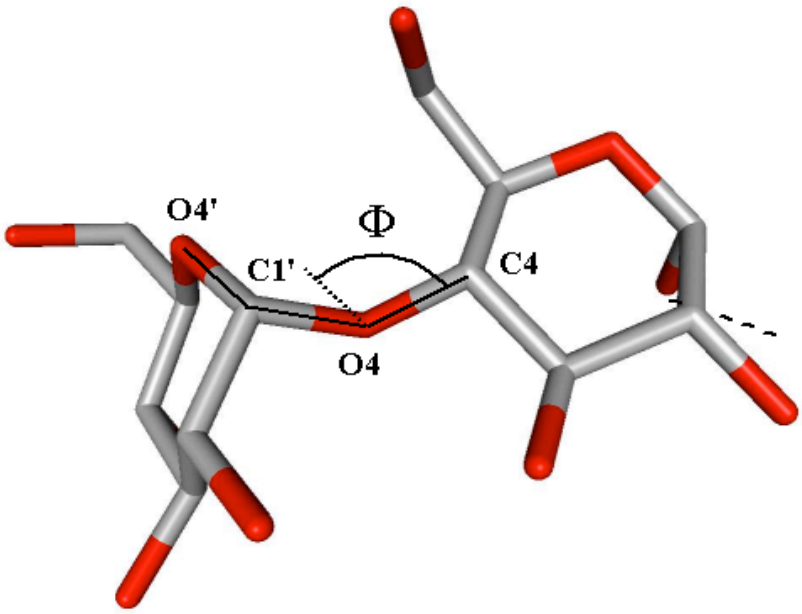


Figura 2.3: Definição do ângulo torsional interglicoses Φ. O oxigênio intersacarídeo faz parte da ligação em torno da qual há a torção.

O ângulo Ψ é formado pelos átomos C3-C4-O4-C1′ (ver Figura 2.4). Através desses dois ângulos podemos analisar a estrutura macrocíclica da α-CD no que tange à abertura do cone.

Há uma grandeza geométrica, entretanto, que também pode informar sobre a abertura da α-CD e torção relativa entre as unidades de glicose de modo equivalente aos ângulos Φ e Ψ. Este é o ângulo τ definido como sendo o ângulo entre o plano médio da α-CD (dado pelos oxigênios

intersacarídeos) e o plano médio de cada glicose (dado pelos átomos do anel de piranose) como ilustrado na Figura 2.5.

De qualquer forma as unidades de glicose não possuem muita liberdade para rotacionar em torno da ligação glicosídica de modo que esses três ângulos podem variar, em geral, apenas dentro de um pequeno intervalo. Isso por causa da rede de ligações de hidrogênio entre as hidroxilas secundárias que impede as unidades de glicose de girarem entre si. Assim há uma correlação entre esses ângulos e uma outra grandeza interessante de se observar que é a distância O3-O2′ entre os oxigênios das hidroxilas secundárias. Essa distância pode nos dizer sobre o caráter dessas ligações de hidrogênio.

Dados experimentais [*Manor*, 1974; *Klar*, 1980; *Chacko*, 1981; *Lindner*, 1982] demonstram que essa rede de ligações só é rompida quando há a interação entre uma hidroxila primária e uma molécula hospedeira ou com a água provocando uma maior torção da glicose em relação às adjacentes.

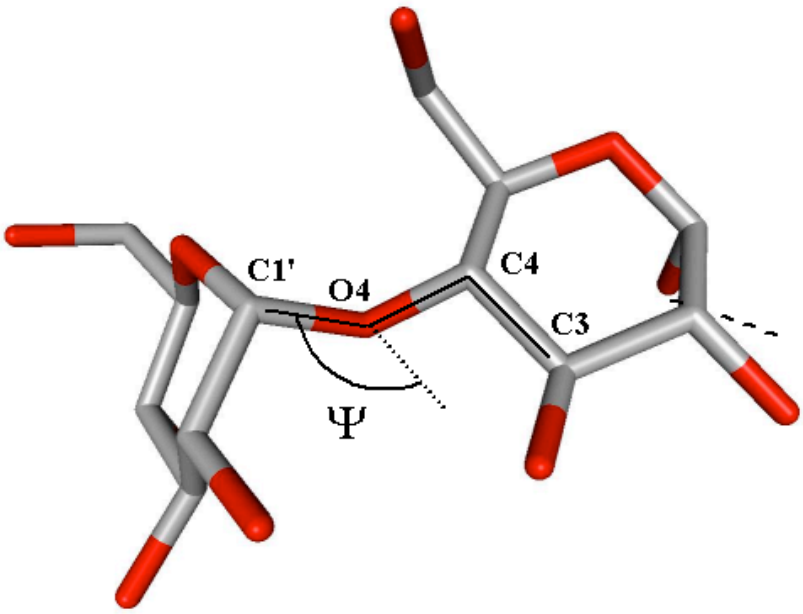


Figura 2.4: Definição do ângulo torsional Ψ. Assim como o ângulo Φ, esse ângulo informa sobre a torção relativa entre as unidades de glicose.

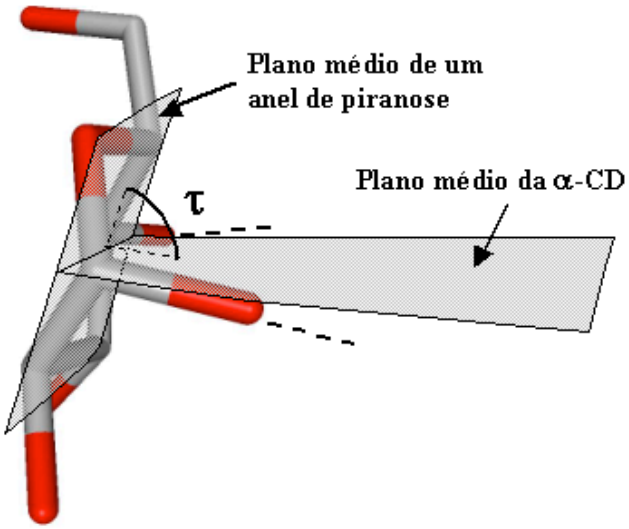


Figura 2.5: Definição do ângulo τ.

Duas grandezas geométricas interessantes para analisar a estrutura da cavidade da α-CD são a distância O4-O4′ e o ângulo δ formado por três oxigênios intersacarídeos consecutivos O4-O4′-O4′′ (Figura 2.6). Um aumento nessas grandezas acarreta em aumento no volume da cavidade. Entretanto, essas grandezas também variam muito pouco e estão relacionadas com os ângulos que dão a estrutura cônica da molécula τ, Φ e Ψ.

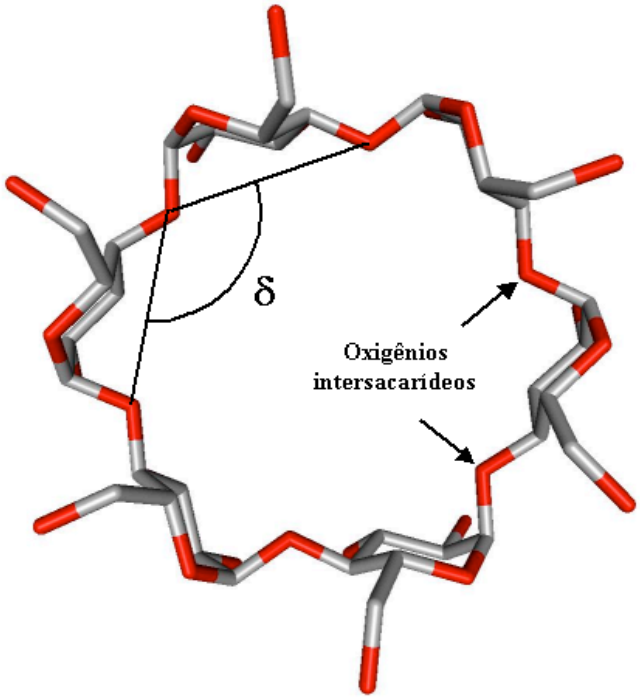


Figura 2.6: Definição do ângulo δ.

Uma grandeza também comumente usada na literatura é o ângulo glicosídico φ formado pelos átomos C4-O4-C1′, como ilustra a Figura 2.7.

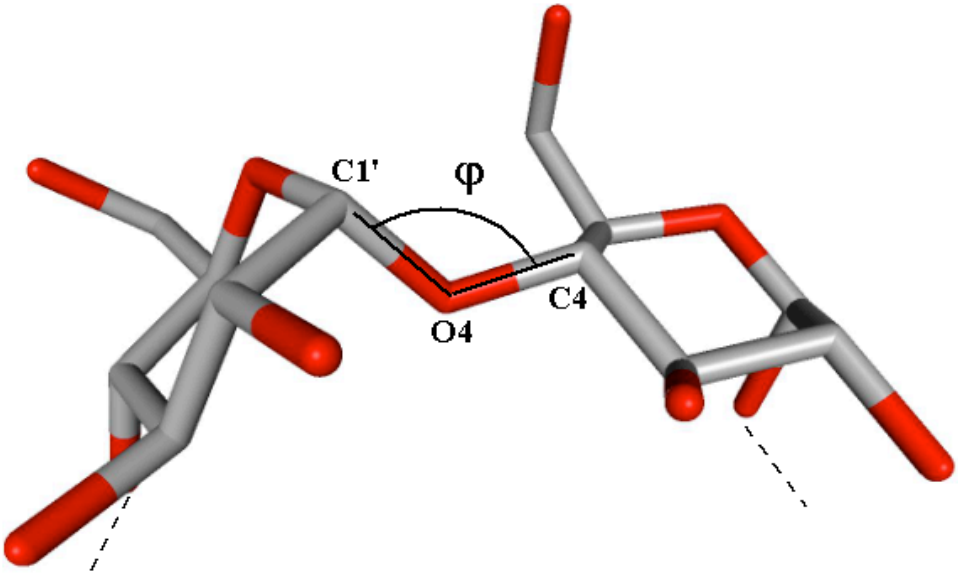


Figura 2.7: Definição do ângulo glicosídico φ.

Finalmente, para analisar as hidroxilas primárias que possuem uma liberdade maior para se moverem que os demais grupos da molécula, existe o ângulo ω que é o ângulo diedral O5-C5-C6-O6 mostrado na Figura 2.8. A

posição das OH primárias pode ser (-)*gauche* (G(-)), (+)*gauche* (G(+)) ou *trans* (T). A conformação (-)*gauche* é a que ocorre com maior freqüência seguida de (+)*gauche* e a conformação *trans* é a menos favorável segundo os dados experimentais em cristais, onde até hoje não se detectou tal conformação para uma hidroxila primária [*Saenger*, 1998] e em meio solvente onde ocorre também com menor freqüência [*Schneider*, 1998].

Observe na Figura 2.6 que todas as 6 hidroxilas primárias estão na conformação (-)*gauche*. Isso provoca uma orientação dessas hidroxilas para fora da cavidade interna da α-CD e as deixa livres para uma possível interação com solventes. A conformação (+)*gauche* provoca uma orientação dessas hidroxilas para dentro da cavidade e a conformação *trans* promove uma posição eclipsada de O6 com o oxigênio intersacarídeo O4 seguinte.

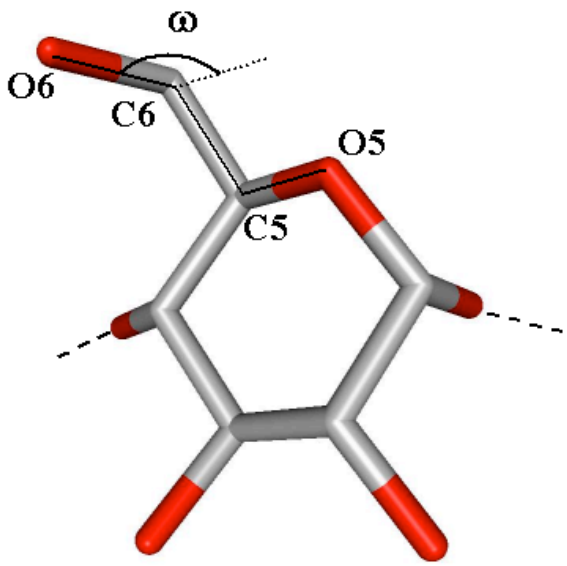


Figura 2.8: Definição do ângulo ω.

2.4. Resultados com campos de força

Para estudar as possíveis geometrias da α-CD geradas pelos diferentes campos de força empíricos fizemos várias otimizações envolvendo todos os graus de liberdade da α-CD utilizando o método de mecânica molecular partindo de várias geometrias diferentes. Partimos de uma determinada geometria a e otimizamos com os campos de força. A partir daí tomamos geometrias otimizadas com um campo de força, e até

mesmo com métodos quânticos (veja Seção 2.5), como ponto de partida para otimização com outro campo de força. Com isso pudemos variar bastante o ponto de partida das diversas otimizações. Também, no intuito de encontrar outros mínimos locais na superfície de energia potencial de cada campo de força, utilizamos em algumas otimizações critérios de convergência mais apertados para sair do ponto de mínimo recém encontrado.

Como era esperado para uma molécula com tantos graus de liberdade, observamos a existência de vários mínimos na superfície de potencial de todos os campos de força.

Obtivemos assim um total de 13 mínimos na superfície de potencial da α-CD com campo de força OPLS, 13 mínimos com o campo de força AMBER e 9 mínimos com o campo de força MM2. Como cada campo de força utilizado tem uma forma funcional diferente e como cada um tem seu nível zero de energia, as energias clássicas somente são comparáveis dentro do mesmo campo de força. Assim não podemos comparar as geometrias obtidas com os três campos de força para encontrar qual é mais estável. Para solucionar este problema submetemos todas as 35 geometrias encontradas com os campos de força empíricos a um mesmo nível de cálculo de energia.

Assim calculamos para todas as geometrias obtidas a energia em nível HF/6-31G (método/base, vide seção 2.5) e pudemos observar a estabilidade relativa entre elas e posteriormente compará-las com as geometrias obtidas quanticamente,que também foram submetidas ao mesmo nível de cálculo para a energia.

Para definir as ligações de hidrogênio formadas entre as hidroxilas adotamos o seguinte critério geométrico [*Mezei*, 1981; *Stillinger*, 1980]: distância entre os oxigênios das hidroxilas $R_{OO} \leq 3.5$ Å e ângulo formado entre oxigênio aceitador, oxigênio doador e hidrogênio doado $\theta_{OOH} \leq 50°$. Esse é um critério bastante abrangente de modo que é interessante também analisar, através de um critério mais rigoroso, quantas das ligações formadas com o critério acima são ligações fortes. Para isto adotamos o critério: $R_{OO} \leq 3.3$ Å e $\theta_{OOH} \leq 30°$.

Como subproduto dos cálculos quânticos para determinação das energias relativas entre as geometrias obtidas com campos de força e com métodos quânticos, obtivemos o momento de dipolo para todas elas. Adicionalmente, fizemos uma análise preliminar dessas geometrias através da identificação das conformações das hidroxilas primárias (G(-), G(+) e T) e da formação de ligações de hidrogênio entre as hidroxilas secundárias.

Os resultados estão resumidos na Tabela 2.2 para o campo de força AMBER, na Tabela 2.3, para o OPLS e na Tabela 2.4 para o MM2.

Tabela 2.2: Geometrias obtidas com o método AMBER. Na primeira coluna mostramos as energias relativas calculadas com HF/6-31G e na segunda coluna a energia relativa obtida classicamente, ambas em kcal/mol. Na terceira coluna mostramos o número de ligações de hidrogênio formadas entre as OH secundárias segundo o critério mais rigoroso e na quarta coluna segundo o critério mais abrangente. Na quinta coluna o dipolo em Debye e na última a conformação das hidroxilas primárias.

Estruturas	ΔE (Qua)	ΔE (Cla)	HB sec	HB sec	μ	OH prim
AMBER1	0.00	0.34	4	6	13.11	G (-)
AMBER2	0.19	2.70	4	6	10.99	T
AMBER3	1.70	0.29	4	6	13.01	G (-)
AMBER4	1.73	0.00	4	6	13.11	G (+)
AMBER5	2.21	0.07	5	5	13.03	G (-)
AMBER6	2.52	2.68	4	6	10.85	T
AMBER7	3.70	0.10	4	5	13.21	G (-)
AMBER8	3.91	0.07	5	5	13.02	G (-)
AMBER9	4.73	2.78	4	6	10.68	T
AMBER10	9.05	0.19	4	6	12.95	G (-)
AMBER11	10.97	0.31	3	4	7.32	G (-)
AMBER12	11.28	2.02	3	5	10.86	T
AMBER13	12.28	2.02	4	5	10.93	T

Nas 3 tabelas a energia relativa quântica ΔE(Qua) está numa escala em que a geometria clássica mais estável tem energia zero, ou seja a

geometria AMBER1. Já a energia relativa clássica ΔE(Cla) só pode ser alvo de comparação dentro do mesmo campo de força e cada um tem seu nível zero de energia de acordo com a geometria mais favorável classicamente de cada um (AMBER4, OPLS12 e MM1 respectivamente).

Tabela 2.3: Geometrias obtidas com o método OPLS. As colunas seguem o mesmo padrão da Tabela 2.2. Energias em kcal/mol e dipolo em Debye.

Estruturas	ΔE (Qua)	ΔE (Cla)	HB sec	HB sec	μ	OH prim
OPLS1	5.54	3.70	2	6	13.06	G (-)
OPLS2	6.97	6.68	2	6	10.87	T
OPLS3	7.22	6.67	4	6	10.75	T
OPLS4	7.29	3.56	2	6	12.93	G (-)
OPLS5	7.65	3.56	2	6	13.08	G (-)
OPLS6	7.76	3.55	2	6	13.05	G (-)
OPLS7	9.12	6.58	2	6	10.55	T
OPLS8	11.83	4.18	0	2	9.69	T
OPLS9	13.64	1.81	0	1	13.18	G (+)
OPLS10	14.52	1.48	0	0	10.90	G (+)
OPLS11	18.63	0.60	1	2	6.59	G (-)
OPLS12	22.80	0.00	0	2	12.65	G (-)
OPLS13	23.03	0.00	0	2	12.64	G (-)

Tabela 2.4: Geometrias obtidas com o método MM2. Energias em kcal/mol e dipolo em Debye.

Estruturas	ΔE (Qua)	ΔE (Cla)	HB sec	HB sec	μ	OH prim
MM1	26.58	0.00	2	3	12.93	G (+)
MM2	26.77	0.00	2	3	12.91	G (+)
MM3	30.30	23.64	0	3	8.40	G (-)

Estruturas	ΔE (Qua)	ΔE (Cla)	HB sec	HB sec	μ	OH prim
MM4	36.01	25.21	0	4	10.95	T
MM5	36.13	3.25	2	3	15.70	T
MM6	52.72	7.14	2	2	13.32	G (+)
MM7	57.27	6.60	1	2	14.34	G (+)
MM8	59.93	7.35	1	2	13.36	G (-)
MM9	59.95	7.09	2	2	13.02	G (-)

Analisando os resultados apresentados nas 3 tabelas acima, observamos que, de forma geral, o conjunto de geometrias geradas pelo AMBER apresenta energias relativas menores que as geradas pelo OPLS que por sua vez também apresentam energias relativas menores que as geometrias geradas pelo MM2.

Adicionalmente observamos que as energias relativas clássicas não seguem a mesma tendência das quânticas, sendo que elas não concordam em geral sobre a seqüência de estabilidade entre as geometrias de um mesmo campo de força. Entretanto, nos campos de força AMBER e OPLS notamos uma relação clara de desfavorabilidade em relação à conformação trans adotada pelas hidroxilas primárias, o que não acontece para o MM2.

Em todas as geometrias obtidas as hidroxilas primárias assumem preferencialmente a conformação (-)*gauche* sendo no mínimo 5 das 6 hidroxilas primárias nessa conformação. Portanto, a última coluna das três tabelas se refere a uma única hidroxila. Nesse ponto todos os campos de força clássicos estão de acordo com os dados experimentais, tanto cristalográficos [*Saenger*, 1998; *Harata*, 1998] quanto de NMR [*Schneider*, 1998] que mostram ser mais favorável a conformação (-)*gauche* que a (+)*gauche*, que por sua vez se mostra mais favorável que a conformação *trans*.

Quanto às ligações de hidrogênio entre as hidroxilas secundárias, os resultados experimentais [*Manor*, 1974; *Klar*, 1980; *Chacko*, 1981; *Lindner*, 1982; *Puliti*, 1998] indicam que devem existir no mínimo 4 ligações de hidrogênio. Assim observamos que o método MM2 não é capaz

de descrever a geometria da α-CD corretamente. Esse campo de força é ineficiente em descrever as ligações de hidrogênio formadas entre as hidroxilas secun-dárias como podemos ver na Tabela 2.4. Nessa tabela podemos ver que mesmo sob o critério mais abrangente, a geometria mais favorável só estabelece três ligações de hidrogênio entre as hidroxilas secundárias. Como conseqüência as energias quânticas das geometrias MM2 são bem mais altas que as energias das demais geometrias. Também como conseqüência essas geometrias são altamente irregulares com grandes variações no ângulo τ.

Desta forma, analisaremos mais profundamente apenas as geometrias geradas com os campos de força AMBER e OPLS. Esses se mostraram eficazes em descrever a α-CD. Isso pode ser visto primeiramente através da formação da rede de ligações de hidrogênio no anel secundário em um número significativo de geometrias nos dois métodos. Isso é interessante porque, como dissemos, se espera que uma molécula com vários graus de liberdade como a α-CD tenha várias geometrias de equilíbrio e certamente muitas delas mantendo a rede de ligações no anel secundário.

Na Tabela 2.3, por exemplo, podemos ver que nas geometrias mais favoráveis do OPLS há a formação das seis ligações de hidrogênio possíveis entre as OH secundárias. Entretanto, dessas, em geral apenas duas são formadas segundo o critério mais rigoroso, de modo que não são formadas ligações muito fortes.

Enquanto isso, a maioria das geometrias obtidas com o AMBER conseguem formar cinco ou seis ligações entre as OH secundárias no critério abrangente e quatro ou cinco no critério forte. Isso provavelmente se deve ao fato de o campo de força AMBER tratar de modo especial as ligações de hidrogênio como dissemos na seção 2.2.10.

Analisando ainda as geometrias obtidas com o AMBER e OPLS, podemos notar que nove geometrias AMBER tem energia quântica abaixo de 5 kcal/mol enquanto que a geometria mais favorável obtida com o método OPLS tem energia de 5.54 kcal/mol. Juntamente com o fato de as geometrias AMBER formarem mais ligações entre as hidroxilas secundárias, isso demonstra um desempenho diferenciado do campo de

força AMBER para descrever a α-CD. Isso se deve certamente também ao fato de o método AMBER ser parametrizado para polissacarídeos (carboidratos) e portanto descreve melhor as ligações glicosídicas envolvidas na estrutura das ciclodextrinas e que as influencia em grande medida.

É interessante observar que há claramente uma correlação entre a conformação adotada pelas hidroxilas primárias e o dipolo da molécula. As conformações *gauche* têm dipolo aproximadamente igual a 13 Debye (valor obtido também por outros cálculos da literatura [*Kitagawa*, 1987]) enquanto que as conformações *trans* tem dipolo aproximadamente igual a 11 Debye com algumas exceções.

Ainda com relação ao dipolo da α-CD encontramos que em todas as geometrias, sem exceção, a maior contribuição está no plano médio da molécula formado pelos oxigênios intersacarídeos. Seria de se imaginar que a maior contribuição para o dipolo estivesse no eixo do tronco de cone formado pela molécula por ser esse um eixo com bastante simetria. Contudo, a α-CD não possui uma simetria tão elevada e, por ter um diâmetro aproximadamente igual ao dobro de sua altura, qualquer deslocamento de carga dentro da molécula fará com que a contribuição do dipolo no plano médio da molécula seja mais potencializada do que a contribuição ao longo do eixo do cone. Portanto esse resultado não é absolutamente estranho. Na Figura 2.9 ilustramos o momento de dipolo da α-CD.

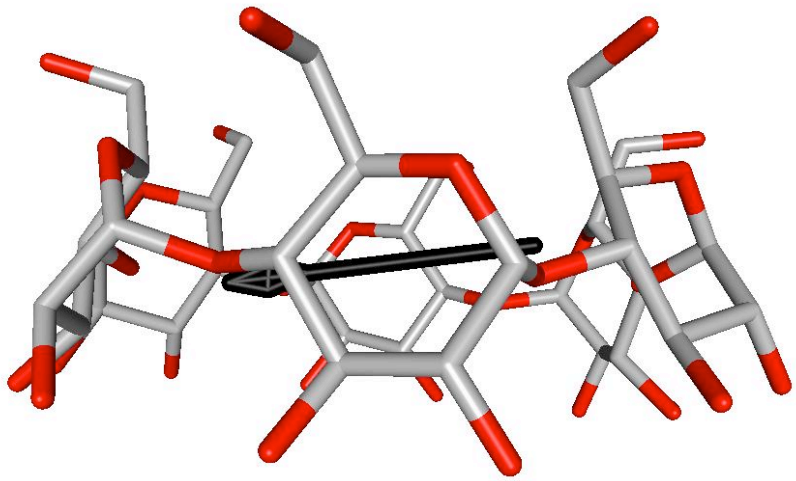


Figura 2.9: Ilustração do dipolo da α-CD representado pela seta escura no interior da molécula. A maior contribuição está no plano médio da molécula como pode-se ver.

2.5. Métodos de Química Quântica

Em mecânica quântica estamos sempre interessados em resolver a equação de Schrödinger de um sistema, de modo a obter suas autofunções e seus autovalores (energias). Em se tratando de moléculas, o hamiltoniano envolvido é dado pela energia cinética dos elétrons e núcleos e pela energia de interação entre núcleos, entre elétrons e entre núcleos e elétrons. As duas primeiras interações são repulsivas e a última é atrativa. Desprezando os termos relativísticos, o hamiltoniano que descreve um sistema de M núcleos e N elétrons é escrito da forma

$$\mathrm{H}(\boldsymbol{r},\boldsymbol{R}) = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{r_{AB}}. \quad (2.15)$$

onde $\boldsymbol{r}$ é o conjunto das coordenadas dos elétrons, $\boldsymbol{R}$ é o conjunto das coordenadas dos núcleos, i e j se referem a elétrons e A e B se referem a núcleos. Estas equações são apresentadas em unidades atômicas para facilitar o seu manejo. Nesse sistema a unidade de comprimento é o *Bohr* (a_0), a de carga é a carga elementar (e), a de massa é a massa do elétron (m_e) e a de momento angular é a constante de Planck dividida por 2π ($\hbar$).

A equação de Schrödinger é então escrita como

$$\mathrm{H}(\boldsymbol{r},\boldsymbol{R})\Psi(\boldsymbol{r},\boldsymbol{R}) = E\Psi(\boldsymbol{r},\boldsymbol{R}) \tag{2.16}$$

A rigor apenas átomos hidrogenóides podem ser resolvidos sem a utilização de aproximações. Para qualquer outro sistema a equação de Schrödinger não é rigorosamente separável. A primeira aproximação a se fazer é a separação dos movimentos nuclear e eletrônico, chamada separação de Born-Oppenheimer.

2.5.1. Separação de Born-Oppenheimer

A aproximação ou separação de Born-Oppenheimer está em tomar a função de onda do sistema dada pela equação (2.16) como o produto de uma função que depende apenas parametricamente de $\boldsymbol{R}$ e outra que depende explicitamente das coordenadas nucleares:

$$\Psi(\boldsymbol{r},\boldsymbol{R}) = \Psi_{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R})\Psi_{\mathrm{nuc}}(\boldsymbol{R}) \tag{2.17}$$

Numa visão mais mecanicista, esta aproximação se baseia no fato de os núcleos serem bem mais pesados que os elétrons. Portanto esses últimos têm inércia muito menor podendo se adaptar a qualquer arranjo dos núcleos, quase instantaneamente.

O hamiltoniano também é separado em uma parte eletrônica e outra parte nuclear:

$$\mathrm{H}(\boldsymbol{r},\boldsymbol{R}) = \mathrm{H}_{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R}) + \mathrm{H}_{\mathrm{nuc}}(\boldsymbol{R}) \tag{2.18}$$

Os hamiltonianos e as respectivas equações são

$$\mathrm{H}_{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R}) = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{r_{ij}} \tag{2.19}$$

$$\mathrm{H}_{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R})\Psi_{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R}) = E_{\mathrm{el}}(\boldsymbol{R})\Psi_{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R}) \tag{2.20}$$

que são as equações do movimento eletrônico. Costuma-se nomear uma grandeza que é a energia total da molécula para uma posição fixa dos núcleos E_{tot} dada por

$$E_{\mathrm{tot}} = E_{\mathrm{el}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{r_{AB}} \tag{2.21}$$

que é o potencial sentido pelos núcleos. As equações de movimento nuclear são então as seguintes

$$\mathrm{H}_{\mathrm{nuc}}(\boldsymbol{R}) = -\sum_{A=1}^{M} \frac{1}{2}\nabla_A^2 + E_{\mathrm{tot}}(\boldsymbol{R}) \tag{2.22}$$

$$\mathrm{H}_{\mathrm{nuc}}(\boldsymbol{R})\Psi_{\mathrm{nuc}}(\boldsymbol{R}) = E\Psi_{\mathrm{nuc}}(\boldsymbol{R}) \tag{2.23}$$

Para escrever as equações (2.20) e (2.23) utilizamos a aproximação [*Canuto*, 2001]

$$\mathrm{T}_{\mathrm{nuc}}\Psi(\boldsymbol{r},\boldsymbol{R})\,;\;\; \Psi_{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R})\mathrm{T}_{\mathrm{nuc}}\Psi_{\mathrm{nuc}}(\boldsymbol{R}) \tag{2.24}$$

com $\mathrm{T}_{\mathrm{nuc}}$ sendo o operador de energia cinética dos núcleos. Através da separação dos movimentos eletrônico e nuclear podemos agora fixar os núcleos em uma determinada posição e calcular $\Psi_{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R})$ através da equação (2.20).

É interessante notar que o potencial de interação entre os núcleos pode ser levado em conta na energia eletrônica E_{el}, o que não acarreta diferença alguma para $\Psi_{\mathrm{el}}(\boldsymbol{r};\boldsymbol{R})$ pois a repulsão nuclear é constante para uma dada configuração dos núcleos. Neste caso a energia potencial utilizada na equação (2.22) seria a própria energia eletrônica E_{el}.

Resolvendo-se o problema eletrônico para um $\boldsymbol{R}$ qualquer, obtemos uma superfície de potencial $E_{\mathrm{tot}}(\boldsymbol{R})$ através da qual podemos calcular $\Psi_{\mathrm{nuc}}(\boldsymbol{R})$ que descreve o movimento nuclear, ou seja, a rotação, vibração e translação da molécula.

Em algumas circunstâncias, entretanto, a aproximação (2.24) não é válida e não é possível separar os movimentos eletrônico e nuclear. Nessas circunstâncias há um acoplamento eletrônico-vibracional gerando o efeito Jahn-Teller [*Englman*, 1972].

Consideraremos aqui apenas o problema eletrônico e a equação que queremos resolver é a equação (2.20). Assim não usaremos mais os subescritos para o hamiltoniano, função de onda e energia eletrônica.

2.5.2. Aproximação de Hartree-Fock

Outra aproximação comumente utilizada para resolver sistemas de muitos elétrons é a aproximação de partículas independentes que consiste em considerar a função de onda eletrônica como um produto de funções de onda de um elétron (chamados de orbitais)

$$\Psi(\boldsymbol{r}) = \psi_1(\boldsymbol{r}_1)\psi_2(\boldsymbol{r}_2)\cdots\psi_N(\boldsymbol{r}_N). \tag{2.25}$$

Os ψ_i são os orbitais espaciais e $\boldsymbol{r}_i$ representa as coordenadas espaciais do elétron i.

Tal produto - chamado produto de Hartree - todavia, não satisfaz o princípio de antissimetria ou princípio de exclusão de Pauli. Esse princípio enuncia na sua forma forte que a função de onda eletrônica deve ser antissimétrica ante à troca das coordenadas espaciais e de spin de dois elétrons quaisquer, ou seja

$$\Psi(\boldsymbol{x}_1,...,\boldsymbol{x}_i,...,\boldsymbol{x}_j,...,\boldsymbol{x}_N) = -\Psi(\boldsymbol{x}_1,...,\boldsymbol{x}_j,...,\boldsymbol{x}_i,...,\boldsymbol{x}_N) \tag{2.26}$$

onde $\boldsymbol{x}_i = (\boldsymbol{r}_i, \omega_i)$ e ω_i é a coordenada de spin do elétron i.

Portanto, além de incluir funções de onda de spin para os elétrons, devemos achar uma forma para a função de onda eletrônica que seja antissimétrica. A solução natural para este problema é a utilização de determinantes de Slater como função de onda. O determinante de Slater para um sistema de N elétrons é dado por

$$\Phi(\boldsymbol{x}) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \varphi_1(\boldsymbol{x}_1) & \varphi_2(\boldsymbol{x}_1) & \cdots & \varphi_N(\boldsymbol{x}_1) \\ \varphi_1(\boldsymbol{x}_2) & \varphi_2(\boldsymbol{x}_2) & \mathrm{L} & \varphi_N(\boldsymbol{x}_2) \\ \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} \\ \varphi_1(\boldsymbol{x}_N) & \varphi_2(\boldsymbol{x}_N) & \mathrm{L} & \varphi_N(\boldsymbol{x}_N) \end{vmatrix} \tag{2.27}$$

ou pelas formas simplificadas

$$\Phi(\boldsymbol{x}) = \frac{1}{\sqrt{N!}}\det(\varphi_1\varphi_2\mathrm{L}\ \varphi_N) \tag{2.28}$$

$$\Phi(\boldsymbol{x}) = \left|\varphi_1\varphi_2\mathrm{L}\ \varphi_N\right\rangle \tag{2.29}$$

onde agora $\varphi_i(\boldsymbol{x}_1) = \psi_i(\boldsymbol{r}_1)\alpha(\omega_1)$ ou $\varphi_i(\boldsymbol{x}_1) = \psi_i(\boldsymbol{r}_1)\beta(\omega_1)$ são os spin-orbitais moleculares. Na última notação o fator de normalização é suprimido por simplicidade já que está notação será mais utilizada.

Essa forma para a função de onda garante automaticamente sua antissimetria por propriedade de determinantes. Note também que se dois elétrons ocuparem o mesmo spin orbital, a função de onda se anula.

Apesar de não sabermos ainda quem são os spin-orbitais φ_i, podemos pensar na seqüência que eles seguem segundo a sua energia característica. Assim podemos pensar num determinante que represente o estado fundamental Φ_0 composto pelos orbitais de mais baixa energia ocupados. Os estados excitados seriam determinantes Φ_a^k onde um elétron sairia do orbital ocupado *a* para o orbital virtual *k* (excitações simples), Φ_{ab}^{kl} onde dois elétrons sairiam dos orbitais *a* e *b* para *k* e *l* respectivamente (excitações duplas) e assim por diante. A expansão completa da função de onda eletrônica é então dada por

$$\Psi(x) = C_0\Phi_0 + \sum_{a,k} C_a^k \Phi_a^k + \sum_{a,k}\sum_{b,l} C_{ab}^{kl}\Phi_{ab}^{kl} + \mathrm{L} \tag{2.30}$$

A aproximação de Hartree-Fock (HF) se baseia na utilização de um único determinante de Slater (Φ_0) para representar o estado fundamental de um sistema molecular qualquer. Além disso, o método de Hartree-Fock envolve a utilização do princípio variacional para a obtenção dos spin-orbitais φ_i de modo a resultar no melhor determinante que possa descrever esse estado eletrônico.

A energia do estado fundamental do sistema é dada por

$$E_0 = \langle \Phi_0 | \mathrm{H} | \Phi_0 \rangle \tag{2.31}$$

Podemos ver que E_0 é um funcional dos spin-orbitais que queremos obter. Pode-se mostrar facilmente [*Castro*, 1999; *Szabo*, 1996] que esta dependência dos spin-orbitais se dá pela seguinte forma

$$E_0 = \sum_{i=1}^{N} \langle i | h | i \rangle + \sum_{i=1}^{N}\sum_{j>1}^{N} \langle ij || ij \rangle \tag{2.32}$$

onde h é o operador de um elétron também chamado de *hamiltoniano de caroço*

$$h(\boldsymbol{x}_1) = -\frac{1}{2}\nabla_1^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{1A}} \tag{2.33}$$

e $|i\rangle$ é o spin-orbital $\varphi_i(\boldsymbol{x}_1)$. O segundo termo em (2.32) engloba as integrais de dois elétrons

$$\langle ij||ij\rangle = \langle ij|ij\rangle - \langle ij|ji\rangle \tag{2.34}$$

$$\langle ij|kl\rangle = \int \varphi_i^*(\boldsymbol{x}_1)\varphi_j^*(\boldsymbol{x}_2)\frac{1}{r_{12}}\varphi_k(\boldsymbol{x}_1)\varphi_l(\boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2. \tag{2.35}$$

Pela equação (2.35) podemos perceber que o primeiro termo em (2.34) é simplesmente a repulsão coulombiana entre dois elétrons nos orbitais i e j, enquanto que o segundo termo é a *interação de troca* entre os elétrons nesses orbitais que se origina da antissimetria da função de onda e está relacionado à correlação de troca, ou seja, à correlação no movimento de elétrons com spins paralelos [*Szabo*, 1996, Seção 2.2.3].

A idéia agora é minimizar esse funcional dos spin-orbitais mantendo o vínculo de ortonormalidade para os mesmos. Isso evoca naturalmente o método dos multiplicadores de Lagrange. O funcional a ser minimizado é então

$$\mathrm{L} = \sum_{i=1}^{N}\langle i|h|i\rangle + \sum_{i=1}^{N}\sum_{j>1}^{N}\langle ij||ij\rangle + \sum_{ij}^{N}\lambda_{ij}\left(\langle i|j\rangle - \delta_{ij}\right) \tag{2.36}$$

onde os λ_{ij} são os multiplicadores de Lagrange. Minimizando esse funcional obtemos finalmente

$$f(\boldsymbol{x}_1)\varphi_i(\boldsymbol{x}_1) = \sum_{j=1}^{N}\lambda_{ij}\varphi_j(\boldsymbol{x}_1). \tag{2.37}$$

f é o operador de Fock que apresentaremos adiante. Se diagonalizarmos a matriz $\boldsymbol{\lambda}$, que é hermiteana, através de uma transformação unitária

obteremos um conjunto novo, porém ainda ortonormal, de spin-orbitais[3]. Esse novo conjunto de spin-orbitais são os autovetores do operador de Fock. Os autovalores λ_{ii} podem ser interpretados como sendo as energias associadas aos orbitais moleculares, já que f representa a energia sentida por um elétron. Esses autovalores possuem outras interpretações também em termos de energias de ionização e afinidade eletrônica através do teorema de Koopman [*Wittel*, 1977; *Szabo*, 1996]. A equação (2.37) fica então da forma

$$f(\boldsymbol{x}_1)\varphi_i(\boldsymbol{x}_1) = \varepsilon_i\varphi_i(\boldsymbol{x}_1) \tag{2.38}$$

Essa é a equação canônica de Hartree-Fock e é a equação de trabalho para a obtenção dos φ_i. f é dado por

$$f(\boldsymbol{x}_1) = h(\boldsymbol{x}_1) + \sum_{j=1}^{N}\left[\mathrm{J}_j(\boldsymbol{x}_1) - \mathrm{K}_j(\boldsymbol{x}_1)\right] \tag{2.39}$$

onde h é o hamiltoniano de caroço e no somatório temos os operadores de Coulomb e de troca. Esses são os operadores que produzem as integrais de dois elétrons da equação (2.35) e podem ser escritos da forma

$$\mathrm{J}_j(\boldsymbol{x}_1)\varphi_i(\boldsymbol{x}_1) = \left[\int \varphi_j^*(\boldsymbol{x}_2)\frac{1}{r_{12}}\varphi_j(\boldsymbol{x}_2)d\boldsymbol{x}_2\right]\varphi_i(\boldsymbol{x}_1) \tag{2.40}$$

$$\mathrm{K}_j(\boldsymbol{x}_1)\varphi_i(\boldsymbol{x}_1) = \left[\int \varphi_j^*(\boldsymbol{x}_2)\frac{1}{r_{12}}\varphi_i(\boldsymbol{x}_2)d\boldsymbol{x}_2\right]\varphi_j(\boldsymbol{x}_1). \tag{2.41}$$

Mais uma vez podemos ver que o operador de Coulomb J_j representa o potencial médio gerado por um elétron no orbital j. O operador de troca K_j, como o próprio nome diz, faz a troca entre os elétrons do orbital j e do orbital em que o operador atua. Esse operador não possui análogo clássico e, como dissemos, surge da antissimetria da função de onda. Somando sobre todos os orbitais ocupados, podemos encarar o segundo termo em (2.39) como um potencial efetivo gerado por todos os outros elétrons do sistema.

[3] Pode-se mostrar que qualquer conjunto de spin-orbitais que sofra uma transformação unitária formará um determinante de Slater que diferirá do determinante inicial apenas por um fator de fase que não influi no valor das médias dos observáveis [*Szabo*, 1996, Seção 3.2.3].

Podemos escrever, neste sentido e de maneira ilustrativa o operador de Fock da forma

$$f(\boldsymbol{x}_i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \upsilon^{\mathrm{HF}}(\boldsymbol{x}_i) \tag{2.42}$$

Esse potencial υ^{HF} é um potencial médio pois é dado pela integração de todos os outros spin-orbitais ocupados. Assim a aproximação de Hartree-Fock é uma aproximação de campo médio. Por esse motivo perde-se nessa aproximação a correlação eletrônica que deve ser restaurada com outros métodos chamados de pós HF.

Deve-se notar que o potencial efetivo υ^{HF} depende dos spin-orbitais ocupados e portanto a equação de Hartree-Fock não é linear, devendo ser resolvida iterativamente. O processo usado para se resolver a equação de HF é chamado de método do Campo Autoconsistente (SCF). O método consiste em tomar spin-orbitais iniciais por tentativa e a partir daí construir o potencial efetivo υ^{HF}, possibilitando a obtenção de um novo conjunto de spin-orbitais. Esse processo se auto-alimenta até que não haja nenhuma mudança nos orbitais e então é tido como convergido dentro de um critério adotado.

2.5.3. Aproximação LCAO

Na forma mostrada na equação (2.38), a equação de HF deve ser resolvida numericamente o que é bastante dispendioso. Considerando isso, em 1951 Roothaan apresentou uma abordagem matricial para a resolução da equação de HF [*Roothaan*, 1951]. A idéia é expandir os spin-orbitais moleculares em um conjunto de funções conhecidas. Os candidatos naturais para essas funções são os orbitais atômicos e por isso dizemos que aproximamos os orbitais moleculares por uma combinação linear de orbitais atômicos conhecidos (aproximação LCAO).

Antes de se proceder a expansão entretanto, é necessário eliminar as coordenadas de spin do problema, voltando aos orbitais espaciais. Para moléculas de camada fechada, fazendo isso a equação de HF em termos dos orbitais espaciais fica

$$f(\boldsymbol{r}_1)\psi_i(\boldsymbol{r}_1) = \varepsilon_i\psi_i(\boldsymbol{r}_1) \tag{2.43}$$

com o operador de Fock ligeiramente diferente:

$$f(\boldsymbol{r}_1) = h(\boldsymbol{r}_1) + \sum_{j=1}^{N/2}\left[2\mathrm{J}_j(\boldsymbol{r}_1) - \mathrm{K}_j(\boldsymbol{r}_1)\right]. \tag{2.44}$$

O fator 2 no termo do operador de Coulomb se deve ao fato de em cada orbital espacial coexistirem dois elétrons. O termo de troca não é dobrado porque só há interação de troca entre elétrons de spins paralelos. A forma dos operadores de Coulomb e de troca e do hamiltoniano de caroço são idênticas às mostradas nas equações (2.40) e (2.41) com a única diferença de atuarem sobre coordenadas espaciais apenas agora.

Usando um conjunto de orbitais $\{\phi_\mu\}$ podemos então expandir os orbitais moleculares ψ_i

$$\psi_i = \sum_{\mu=1}^{K} C_{\mu i}\phi_\mu \tag{2.45}$$

onde K é o tamanho do conjunto. Substituindo a expansão (2.45) na equação de HF para orbitais espaciais temos

$$\mathbf{FC} = \mathbf{SC\varepsilon} \tag{2.46}$$

que é a equação de Roothaan. A matriz de Fock **F** é dada por

$$F_{\mu\nu} = \int \phi_\mu^*(\boldsymbol{r}_1) f(\boldsymbol{r}_1)\phi_\nu(\boldsymbol{r}_1)d\boldsymbol{r}_1 \tag{2.47}$$

a matriz dos coeficientes de expansão **C** tem os elementos dados por (2.45), a matriz de overlap **S** é dada por

$$S_{\mu\nu} = \int \phi_\mu^*(\boldsymbol{r}_1)\phi_\nu(\boldsymbol{r}_1)d\boldsymbol{r}_1 \tag{2.48}$$

e $\boldsymbol{\varepsilon}$ é a matriz diagonal contendo as energias dos orbitais. O trabalho agora está em ortogonalizar a matriz **S** de modo que a equação de Roothaan se torna uma equação matricial de autovetores. Como **S** é hermiteana, pode ser diagonalizada por uma matriz unitária. Essa matriz unitária é usada então para obter uma nova matriz de Fock e uma nova matriz de coeficientes e

obtemos assim a equação de autovetores. Essa nova equação é finalmente resolvida ortogonalizando-se a matriz de Fock resultante.

Para moléculas de camada aberta devemos generalizar a equação de Roothaan. Na verdade, o que se faz é separar os spin-orbitais com spins paralelos em dois conjuntos. Sabendo quais os orbitais espaciais tem elétrons com spin α e quais tem spin β, elimina-se novamente os spins e obtem-se duas equações equivelentes à de Roothaan. Isto foi feito por Pople e Nesbet [*Pople*, 1954], três anos depois que Roothaan estabeleceu as equações matriciais de Hartree-Fock.

Como dissemos os orbitais utilizados dentro da aproximação LCAO são orbitais atômicos, ou seja, são orbitais centrados em um núcleo. Se o conjunto de funções $\{\phi_\mu\}$ - também chamado de conjunto-base, ou simplesmente base - fosse completo então teríamos o resultado exato para a função de onda molecular dentro da aproximação de Hartree-Fock. Como um conjunto completo (como o do átomo de hidrogênio por exemplo) é em geral infinito, ficamos limitados a uma base incompleta e a energia obtida dentro dessa base finita é chamada energia SCF. Felizmente a energia SCF se aproxima rapidamente da energia HF com o número de funções-base. Quando o acréscimo de funções-base não resulta mais em diferença na energia diz-se que a base está saturada.

2.5.4. Métodos Semiempíricos

O custo computacional em se realizar um cálculo de Hartree-Fock é da ordem de K^4, sendo K o número de funções-base. Esta escala decorre do número de integrais de dois elétrons distintas necessárias para compor a matriz de Fock. A essência dos métodos semi-empíricos é diminuir o custo computacional reduzindo o número dessas integrais.

Todas as aproximações semi-empíricas começam por considerar apenas os elétrons de valência. Os outros elétrons são considerados apenas através do efeito de redução da carga nuclear efetiva[4] que provocam. Além disso, apenas uma base mínima é usada para representar os elétrons de valência, ou seja, um orbital atômico para cada elétron de valência. A

[4] Efeito conhecido como shielding ou screening.

grande maioria dos métodos semiempíricos utiliza apenas funções base do tipo Slater [*Szabo*, 1996; *Jensen*, 1999] e apenas do tipo s e p.

Outra aproximação adotada por todos os métodos semi-empíricos é a aproximação de Overlap Diferencial Zero (ZDO) que despreza todos os produtos de funções que dependam das mesmas coordenadas eletrônicas e que estejam centradas em núcleos diferentes, ou seja

$$\phi_A^t(\boldsymbol{r}_i)\cdot\phi_B^u(\boldsymbol{r}_i)=0. \tag{2.49}$$

onde $t,u=s,p_x,p_y,p_z$ são os tipos de funções base. Como conseqüência, a matriz de overlap **S** se torna uma matriz unidade, todas as integrais de dois elétrons envolvendo três ou quatro centros são desprezadas e as integrais de um elétrons envolvendo três centros também são nulas.

Para substituir essas integrais desprezadas, utilizam-se parâmetros ajustados de acordo com dados experimentais ou através de cálculos bastante precisos. Para uma revisão geral sobre métodos semiempíricos ver o livro texto de Pople [*Pople*, 1970].

A aproximação NDDO (Neglect of Diatomic Diferential Overlap) não toma nenhuma aproximação além das já mencionadas. Assim esse é o menor nível de aproximação dentro dos métodos semiempíricos.

Como os orbitais centrados no mesmo núcleo são ortonormais, as integrais de overlap podem ser escritas da forma

$$S_{\mu\nu}=\left\langle\phi_A^t\middle|\phi_B^u\right\rangle=\delta_{tu}\delta_{AB} \tag{2.50}$$

onde os índices μ e ν são agora dados pelo centro ao qual a função está relacionada e pelo tipo de função. Os operadores de um elétron são escritos como

$$h(\boldsymbol{r}_1)=-\frac{1}{2}\nabla_1^2-\sum_{A=1}^{M}\frac{Z'_A}{r_{1A}} \tag{2.51}$$

onde Z'_A é a carga nuclear efetiva devida aos elétrons de caroço. As integrais de um elétron ficam da forma

$$\left\langle \phi_A^t(1) \middle| \frac{Z'_C}{r_{1C}} \middle| \phi_B^u(1) \right\rangle = 0 \qquad \text{para} \quad A \neq C \ \text{ e } \ B \neq C \tag{2.52}$$

$$\left\langle \phi_A^t(1) \middle| h(1) \middle| \phi_B^u(1) \right\rangle = \left\langle \phi_A^t(1) \middle| -\frac{1}{2}\nabla_1^2 - \frac{Z'_A}{r_{1A}} - \frac{Z'_B}{r_{1B}} \middle| \phi_B^u(1) \right\rangle \tag{2.53}$$

$$\left\langle \phi_A^t(1) \middle| -\frac{1}{2}\nabla_1^2 - \frac{Z'_A}{r_{1A}} \middle| \phi_A^u(1) \right\rangle = \delta_{ut} \left\langle \phi_A^t(1) \middle| -\frac{1}{2}\nabla_1^2 - \frac{Z'_A}{r_{1A}} \middle| \phi_A^t(1) \right\rangle. \tag{2.54}$$

As integrais de dois elétrons dentro da aproximação ZDO ficam

$$\left\langle \phi_A^t \phi_B^u \middle| \phi_C^v \phi_D^w \right\rangle = \delta_{AC}\delta_{BD} \left\langle \phi_A^t \phi_B^u \middle| \phi_A^v \phi_B^w \right\rangle \tag{2.55}$$

Para substituir as integrais desprezadas, utiliza-se de valores obtidos através de dados experimentais ou obtidos de cálculos quânticos de alto nível. Maiores detalhes sobre o processo de parametrização dos métodos semiempíricos podem ser vistos em [*Jensen*, 1999].

Os métodos AM1 [*Dewar*, 1985] e PM3 [*Stewart*, 1989; 1990] são modificações da aproximação NDDO. Essas modificações estão no potencial de repulsão caroço-caroço.

No caso do método AM1 foi feita a modificação do potencial de repulsão caroço-caroço e posteriormente a reparametrização de forma manual. Segundo os autores, com essas modificações o método AM1 melhorou a descrição de ligações de hidrogênio intermoleculares em relação ao método MNDO (Modified Neglect of Diferential Overlap) e também melhorou as energias de ativação entre outras propriedades.

O método PM3 consiste numa reparametrização utilizando o hamiltoniando do método AM1 só que através de um método automatizado, possibilitando uma otimização total e simultânea de todos os parâmetros e também a inclusão de um número maior de dados utilizados para o ajuste.

2.6. Resultados com métodos quânticos

Fizemos otimizações completas de geometria da α-ciclodextrina com os métodos quânticos HF, AM1 e PM3. Como dissemos os métodos AM1 e PM3 só utilizam base mínima e apenas para os elétrons de valência. Com o

método HF usamos base mínima em duas otimizações e base 6-21G em uma otimização.

Com base mínima STO-3G a α-CD tem 390 funções base e uma otimização nesse nível demora 1 dia ou 2 dias e 7 horas dependendo do ponto de partida. Os cálculos quânticos foram feitos através do programa Gaussian 98 [*Gaussian*, 1998] numa servidora Alpha 5000 com 2GB de memória RAM e os tempos mencionados aqui correspondem ao tempo de CPU nessa máquina. Com base 6-21G a α-CD tem 714 funções e sua otimização levou 5 dias e 5 horas de CPU. Isso reflete a dificuldade em se estudar macromoléculas via métodos *ab initio*.

Fizemos também duas otimizações com métodos AM1 e PM3. Esses levam cerca de 25 minutos de CPU, confirmando a principal vantagem dos métodos semi-empíricos.

Para todas as geometrias calculamos também o dipolo, analisamos a formação de ligações de hidrogênio entre as hidroxilas secundárias através dos mesmos critérios adotados na seção 2.4. e analisamos as conformaçõesdas hidroxilas primárias. Esses resultados estão na Tabela 2.5. O nível zero para ΔE(Qua) calculado em nível HF/6-31G continua sendo a geometria AMBER1.

Tabela 2.5: Resultados obtidos com os métodos quânticos. A energia relativa quântica é calculada também aqui com método HF/6-31G e é dada em kcal/mol e portanto é comparável às energias quânticas ΔE(Qua) das geometrias obtidas classicamente sendo que o nível zero de energia continua sendo a energia da geometria AMBER1. Dipolo dado em Debye.

Estruturas	ΔE (Qua)	HB sec	HB sec	μ	OH prim
HF/STO-3G(1)	33.42	4	5	9.35	2T
HF/STO-3G(2)	22.94	6	6	16.52	2T
HF/6-21G	-56.51	6	6	14.82	2T
AM1(1)	17.10	5	6	11.85	2T
AM1(2)	17.10	5	6	11.84	2T
PM3(1)	31.66	6	6	12.95	2T

Estruturas	ΔE (Qua)	HB sec	HB sec	μ	OH prim
PM3(2)	45.81	3	5	10.88	G(-)

As geometrias obtidas com os métodos quânticos são eficazes em manter a rede de ligações de hidrogênio nas hidroxilas secundárias. Na verdade, esses métodos formam ligações fortes como podemos ver na terceira coluna da Tabela 2.5. Entretanto, observando a energia relativa vemos que essas geometrias são desfavoráveis em relação às geometrias obtidas com os campos de força clássicos. Com exceção da geometria obtida com método HF/6-21G que é a única mais estável que as geometrias obtidas anteriormente.

Na Tabela 2.5 não pudemos observar relação entre o dipolo e as posições das OH primárias, mesmo porque nessas as hidroxilas primárias assumem várias conformações tendo em geral duas hidroxilas envolvidas em ligações de hidrogênio internas e portanto tendendo a conformações *trans*. A única exceção foi a geometria PM3(2). É importante salientar que não é observado experimentalmente ligações de hidrogênio intramoleculares envolvendo as hidroxilas primárias [*Saenger*, 1998].

As geometrias AM1(2) e PM3(2) foram obtidas de otimizações feitas partindo-se da geometria AMBER1. O AM1 fortaleceu as ligações de hidrogênio secundárias e restabeleceu as ligações primárias dando origem à mesma geometria obtida anteriormente que partiu da geometria HF/6-21G. Por isso, as duas geometrias AM1 são idênticas e representam o mesmo ponto na superfície de potencial deste método. Segundo o método AM1 não há barreiras energéticas entre a geometria AMBER1 e HF/6-21G. Já o PM3 encontrou uma geometria bem diferente e acabou por desfazer uma ligação de hidrogênio secundária e das cinco ligações restantes apenas 3 permaneceram fortes.

2.7. Comparação das geometrias

Tomamos então as geometrias mais estáveis obtidas com os campos de força clássicos (AMBER1) e com os métodos quânticos (HF/6-21G) e

promovemos uma comparação mais minuciosa entre essas duas geometrias e geometrias experimentais obtidas da cristalização de α-CD em meio aquoso [*Lindner*, 1982; *Puliti*, 1998]. Uma das geometrias experimentais possui todas as 6 ligações entre hidroxilas secundárias e a outra possui 5 ligações estabelecidas graças a interação entre algumas hidroxilas primárias e a água o que forçou essas a adotarem a conformação (+)*gauche*. Estas geometrias experimentais foram escolhidas porque são as menos interagentes com a água das geometrias experimentais das quais possuímos as coordenadas cartesianas.

Tomamos então as duas geometrias otimizadas clássica e quanticamente e as duas geometrias experimentais e analisamos seus valores médios para algumas grandezas geométricas mais relevantes. O resultado dessa comparação está na Tabela 2.6. Para as geometrias experimentais tomamos a média global das grandezas.

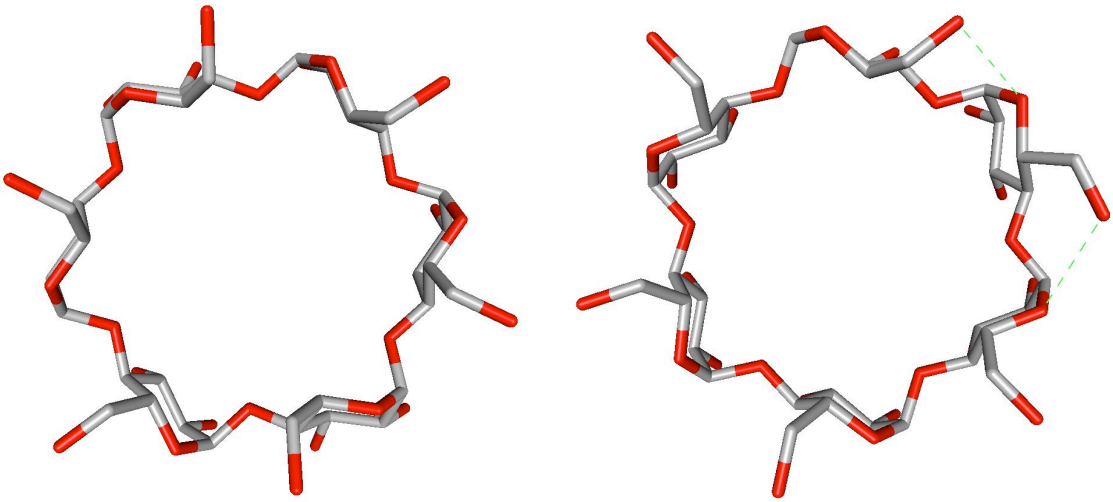
Tabela 2.6: Comparação entre grandezas geométricas relevantes das geometrias AMBER, HF e duas geometrias experimentais. Os valores indicados são a média e o desvio padrão das respectivas grandezas. No número de ligações de hidrogênio formadas entre as OH secundárias temos: ligações formadas segundo critério forte/ligações formadas segundo critério abrangente. ΔE(Qua) em kcal/mol e dipolo em Debye.

Grandeza	AMBER1	HF/6-21G	Experim.
Φ	106±6	113±4	109±7
Ψ	130±8	128±2	128±14
τ	76±7	83±5	80±9
ω	-71±1	2T	3G(+)
δ	120±1	120±2	120±5
$R_{O4\text{-}O4'}$	4.30±0.03	4.28±0.04	4.24±0.09
φ	117.8±0.3	121.2±0.8	119±1
$R_{O3\text{-}O2'}$	3.22±0.03	2.86±0.02	3.0±0.4
HB sec	4/6	6/6	5.5/5.5
ΔE (Qua)	0.00	-56.51	-
μ	13.11	14.82	-

Podemos ver que as duas geometrias calculadas são muito próximas e também tem grandezas geométricas bastante semelhantes às das geometrias experimentais. Algumas pequenas diferenças contudo merecem algum comentário e podem ser vistas claramente na Tabela 2.6. Uma delas é o ângulo de cone τ. Através dele podemos ver que as glicoses nas estruturas experimentais estão inclinadas, em média, de 80º em relação ao plano médio da molécula, a meio caminho portanto da geometria AMBER e HF. Essa diferença no ângulo de cone da geometria HF e da AMBER é reflexo da formação de ligações de hidrogênio secundárias. Através da distância O3-O2′ podemos ver que a geometria HF efetivamente forma ligações mais fortes entre suas hidroxilas secundárias forçando assim uma forma cônica mais aberta que na geometria AMBER. O mesmo pode ser visto no número de ligações de hidrogênio secundárias formadas pelas duas geometrias mostradas na última linha da Tabela 2.6.

Podemos ver que o ângulo φ varia muito pouco dentro de cada geometria e mesmo nas diferentes geometrias. O ângulo ω não é comparável porque na geometria HF há uma OH primária na conformação *trans* e uma quase eclipsada na direção da conformação *trans* e ambas formam ligações de hidrogênio internas. Já nas geometrias experimentais há uma com três OH primárias em conformação (+)*gauche* por causa da interação com a água enquanto a outra tem todas as hidroxilas primárias na conformação (-)*gauche*.

Por último é interessante analisar que o ângulo δ médio é o mesmo para todas as geometrias mas a distância O4-O4′ é ligeiramente maior para a geometria AMBER indicando uma cavidade ligeiramente maior. Na Figura 2.10 ilustramos as duas geometrias analisadas e pode-se ver claramente as hidroxilas primárias apontando para fora da cavidade na geometria AMBER enquanto duas hidroxilas primárias estão em conformação trans formando ligações de hidrogênio com o oxigênio piranosídico da glicose subseqüente.



(a) (b)

Figura 2.10: Ilustração das duas geometrias mais estáveis obtidas (a) com o campo de força AMBER e (b) com método *ab initio* HF/6-21G. Observe a diferença nas hidroxilas secundárias que assumem conformação (-)*gauche* na geometria AMBER enquanto duas tem conformação *trans* na HF.

Analisando as diferenças entre as geometrias calculadas, concentradas principalmente nas ligações de hidrogênio secundárias mais fortes e nas ligações formadas com as hidroxilas primárias, não encontramos justificativas para uma diferença tão grande de energia entre as duas geometrias. Por isso, resolvemos analisar a diferença de energia e o momento de dipolo com o aumento de funções base. Os resultados obtidos estão na Tabela 2.7.

Tabela 2.7: Progressão da diferença de energia interna entre a geometria AMBER e a geometria HF com o aumento do conjunto-base. ΔE = E(HF) – E(AMBER). Os dipolos também são apresentados para as duas geometrias de acordo com o conjunto base.

Base	Funções	ΔE (kcal/mol)	μ_{AMBER} (D)	μ_{HF} (D)
STO-3G	390	0.99	9.17	10.89
6-31G	714	-56.51	13.11	14.82
6-311G	1038	-50.31	12.93	14.63
6-311+G	1302	-46.97	12.69	14.40
6-311G*	1368	-10.40	11.68	13.37
6-311+G*	1632	-7.17	11.58	13.25

Base	Funções	ΔE (kcal/mol)	μ_{AMBER} (D)	μ_{HF} (D)
6-311++G**	1872	-2.25	11.01	12.71

Esses cálculos *single point* realizados com a α-CD variam drasticamente de 4 horas em média com base 6-31G para cerca de 15 dias com base 6-311++G**. Percebemos que com o aumento do número de funções de base o valor da diferença de energia entre as duas geometrias diminui e fica mais condizente com suas diferenças estruturais. Adicionalmente, podemos perceber também que a inclusão de funções de polarização (sinal * acompanhando o nome da base), que são funções com momento angular maior como funções do tipo p, d e f, é mais importante para a diferença de energia do que funções difusas (sinal + no nome da base) que são funções que tem uma abrangência espacial maior como o próprio nome diz.

Como a α-CD é uma molécula muito grande, é esperado que as funções difusas seja menos importantes pois o conjunto total de funções de todos os átomos já é difuso.

A diferença estrutural entre a geometria AMBER e HF é essencial mente a conformação das hidroxilas secundárias que são todas (-)*gauche* na geometria AMBER, ou seja, apontam para fora da cavidade interna, e soa *trans* em duas hidroxilas na geometria HF, ou seja, apontam aproximadamente no sentido da circunferência formada pelas glicoses formando ligações de hidrogênio com os oxigênios piranosídicos (ver Figura 2.10). Dessa forma, a geometria HF perde em energia devido à aproximação desses oxigênios com os oxigênios anoméricos e piranosídicos adjacentes, porém ganha energia com a formação das ligações de hidrogênio. Assim é razoável esperar uma pequena diferença de energia entre as duas geometrias. Para nosso melhor nível de cálculo com 1872 funções base (ver Tabela 2.7) encontramos que a geometria HF é mais estável que a AMBER por 2.25 kcal/mol.

Quanto ao dipolo, também apresentado na Tabela 2.7, observamos que o aumento das funções base não tem efeito tão drástico como na energia. Com o aumento das funções o dipolo diminui e o da geometria HF

permanece maior que o da AMBER. Em nosso melhor nível de cálculo temos que o dipolo da geometria AMBER é 11.01 D e que para a geometria HF é 12.71 D

2.8. Conclusão parcial

Observamos que o campo de força AMBER, dentre os estudados foi o que melhor descreveu a geometria da α-CD. Dos métodos quânticos utilizados, o HF foi o que melhor descreveu a geometria da molécula. Fazendo uma análise detalhada destas duas geometrias, verificamos que a HF é mais estável que a AMBER por cerca de 2 kcal/mol. As duas geometrias, entretanto, são próximas das geometrias experimentais obtidas de cristais hidratados [*Saenger*, 1998] sendo que a AMBER se aproxima mais das geometrias experimentais no que diz respeito à conformação das hidroxilas primárias. É interessante salientar que a conformação destas hidroxilas e a formação das ligações de hidrogênio intramoleculares na geometria HF constitui a principal diferença entre esta e a AMBER, juntamente com a abertura do cone um pouco menor na geometria HF com um maior fortalecimento das ligações de hidrogênio entre as hidroxilas secundárias.

A geometria AMBER tem assim mais hidroxilas livres para uma possível interação com solventes que formem ligações de hidrogênio. Entretanto como a geometria HF tem um dipolo maior, a interação eletrostática será mais forte com um solvente polar como a água.

Portanto, em meio solvente, a formação das ligações intra e intermoleculares e a interação eletrostática irão competir entre si de modo a estabilizar o sistema. Assim será interessante observar o comportamento dessas geometrias em meio solvente (no caso a água). No próximo capítulo estudaremos a solvatação hidrofílica da α-CD em água, fazendo uma análise detalhada das ligações de hidrogênio intermoleculares formadas pelas hidroxilas e na cavidade interna.

Capítulo 3

Solvatação da α-Ciclodextrina.

Neste capítulo versaremos sobre os métodos de simulação computacional comumente utilizados para estudar líquidos moleculares. Apresentaremos o método de amostragem de Monte Carlo e alguns detalhes computacionais envolvidos. Apresentaremos os resultados obtidos da aplicação desse método no estudo da α-CD em solução aquosa e faremos uma comparação da solvatação nas duas geometrias da α-CD (AMBER e HF). Também discutiremos o caráter hidrofílico da parte externa e o caráter hidrofóbico da cavidade da molécula.

3.1. Simulação Computacional

Estudar sistemas solvatados ou sistemas em fase líquida é um problema que envolve muitas partículas e, por isso, qualquer tratamento dado a este tipo de problema deve passar pelos conceitos de mecânica estatística. Por outro lado, para tratar de sistemas grandes, torna se indispensável o uso de ferramentas computacionais aliadas às técnicas mecânico-estatísticas. Essas técnicas e ferramentas que iremos utilizar

neste estudo serão descritas neste capítulo juntamente com os respectivos resultados.

Em mecânica estatística de equilíbrio [*Lage*, 1995] estamos freqüentemente interessados em calcular médias de propriedades da forma:

$$\langle f \rangle_{\text{ens}} = \int f(\boldsymbol{x}) P_{\text{ens}}(\boldsymbol{x}) d\boldsymbol{x} \tag{3.1}$$

onde $P_{\text{ens}}(\boldsymbol{x})$ é o peso normalizado atribuído ao ponto $\boldsymbol{x}$ do espaço de configurações, referente ao ensemble escolhido.

Para nossos propósitos iremos considerar aqui apenas os ensembles canônico (NVT) e isobárico-isotérmico (NPT). Iremos também considerar apenas sistemas cuja hamiltoniana seja da forma

$$\mathrm{H} = \mathrm{H}(\boldsymbol{R}, \boldsymbol{P}) = \sum_{i=1}^{N} \left(\frac{\vec{p}_i^{\,2}}{2m_i} + \frac{\vec{l}_i^{\,2}}{2I_i} \right) + U(\boldsymbol{R}) \tag{3.2}$$

$\vec{p}_i$ e $\vec{l}_i$ são os momentos linear e angular da partícula i enquanto m_i e I_i são a massa e o momento de inércia dessa partícula. $\boldsymbol{P}$ é o vetor que representa os momentos linear e angular de todas as partículas podendo então ser escrito $\boldsymbol{P} = (\vec{P}_1, \vec{P}_2, ..., \vec{P}_N)$ sendo $\vec{P}_i = (\vec{p}_i, \vec{l}_i)$ o conjunto dos momentos da partícula i. Seguindo esta notação temos também o conjunto de momentos lineares $\boldsymbol{p} = (\vec{p}_1, \vec{p}_2, ..., \vec{p}_N)$ e o conjunto de momentos angulares $\boldsymbol{l} = (\vec{l}_1, \vec{l}_2, ..., \vec{l}_N)$ de todas as partículas de modo que podemos escrever $\boldsymbol{P} = (\boldsymbol{p}, \boldsymbol{l})$. $\boldsymbol{R}$ é o vetor que representa as coordenadas referentes a todos os graus de liberdade translacionais e rotacionais do sistema, ou seja, $\boldsymbol{R} = (\boldsymbol{r}, \boldsymbol{\theta})$ onde $\boldsymbol{r}$ é o conjunto das coordenadas de centro de massa de todas as partículas [$\boldsymbol{r} = (\vec{r}_1, \vec{r}_2, ..., \vec{r}_N)$] e $\boldsymbol{\theta}$ é o conjunto de todas as coordenadas angulares [$\boldsymbol{\theta} = (\vec{\theta}_1, \vec{\theta}_2, ..., \vec{\theta}_N)$]. Também podemos escrever $\boldsymbol{R} = (\vec{R}_1, \vec{R}_2, ..., \vec{R}_N)$ com $\vec{R}_i$ sendo as coordenadas generalizadas da partícula i e, portanto, $\vec{R}_i = (\vec{r}_i, \vec{\theta}_i)$.

Essa notação generalizada será bastante útil para englobar translação e rotação quando necessário e pode ser reduzida nos casos em que houver apenas graus de liberdade translacionais ($\boldsymbol{P} = \boldsymbol{p}$ e $\boldsymbol{R} = \boldsymbol{r}$) ou apenas graus

rotacionais ($\boldsymbol{P} = \boldsymbol{l}$ e $\boldsymbol{R} = \boldsymbol{\theta}$). Nos desenvolvimentos a seguir consideraremos que ambos existem, mas pode-se facilmente deduzir os resultados na ausência de um ou outro tipo de grau de liberdade.

Em especial iremos tratar de sistemas moleculares. Além disso, nos restringiremos ao modelo de moléculas rígidas – e por isso a ausência do termo de vibração na energia cinética - de modo que a energia potencial $U(\boldsymbol{R})$ é simplesmente a energia de interação intermolecular.

A distribuição de probabilidades no ensemble NVT é dada por

$$P_{\mathrm{NVT}}(\boldsymbol{R},\boldsymbol{P}) = \frac{1}{Z} e^{-\beta \mathrm{H}(\boldsymbol{R},\boldsymbol{P})} \tag{3.3}$$

onde $\beta = 1/k_B T$, k_B é a constante de Boltzmann ($k_B = 1.38 \cdot 10^{-23}\,\mathrm{J/K}$ = $1.986 \cdot 10^{-3}$ kcal/mol·K) e

$$Z = Z(N,V,T) = \int e^{-\beta \mathrm{H}(\boldsymbol{R},\boldsymbol{P})} d\boldsymbol{R} d\boldsymbol{P} \tag{3.4}$$

é a função de partição canônica. A dependência da função de partição com N e T é imediata. A dependência em V está implícita na região de integração das coordenadas dos centros de massa das partículas do sistema.

Como a hamiltoniana é separável nas suas partes cinética e potencial, a função de partição também é separável em uma parte cinética e outra parte potencial, também chamada de configuracional:

$$Z(N,V,T) = Z_K \cdot Z_C$$

$$Z_K = \int \exp\left(-\beta \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m_i} + \frac{\vec{l}_i^{\,2}}{2I_i} \right) d\boldsymbol{P} \tag{3.5}$$

$$Z_C = \int e^{-\beta U(\boldsymbol{R})} d\boldsymbol{R}.$$

Uma grandeza fundamental e cujo valor médio e flutuação são muito importantes no cálculo de outras propriedades termodinâmicas é a energia E do sistema. Pela equação (3.1) temos

$$\langle E \rangle = \frac{1}{Z} \int \mathrm{H}\, e^{-\beta \mathrm{H}} d\boldsymbol{R} d\boldsymbol{P} \tag{3.6}$$

Por uma propriedade das gaussianas [*Lage*, 1995, Capítulo 3], esta integral se reduz a:

$$\langle E \rangle = 3Nk_B T + \langle U \rangle \tag{3.7}$$

de modo que cada termo na hamiltoniana em ax^2 (com a positivo e x uma variável de integração) contribui para a média com $k_B T/2$. Este resultado geral é conhecido como princípio da equipartição da energia. Considerando que todas as moléculas possuem os graus de liberdade translacionais, temos três casos principais: as moléculas monoatômicas, sem graus rotacionais, terão sua energia cinética média igual a $3k_B T/2$; as moléculas diatômicas, com dois graus de liberdade rotacionais, terão energia média de $5k_B T/2$; e as moléculas poliatômicas terão energia cinética média igual à dada pela equação (3.7).

Por causa do princípio da equipartição, se alguma grandeza depender das velocidades das partículas apenas através de sua energia cinética, então pode ser facilmente reduzida a termos em $k_B T$. A menos de alguma propriedade que possa depender de outra forma das velocidades das partículas, podemos considerar que as grandezas termodinâmicas dependem apenas das coordenadas do sistema. Sendo assim teremos

$$\langle f \rangle_{\mathrm{NVT}} = \frac{1}{Z} \int f(\boldsymbol{R}) e^{-\beta \mathrm{H}(\boldsymbol{R},\boldsymbol{P})} d\boldsymbol{R} d\boldsymbol{P} \tag{3.8}$$

$$\langle f \rangle_{\mathrm{NVT}} = \frac{1}{Z_K Z_C} \int f(\boldsymbol{R}) e^{-\beta U(\boldsymbol{R})} d\boldsymbol{R} \int \exp\left(-\beta \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m_i} + \frac{\vec{l}_i^{\,2}}{2I_i} \right) d\boldsymbol{P} \tag{3.9}$$

$$\langle f \rangle_{\mathrm{NVT}} = \frac{1}{Z_C} \int f(\boldsymbol{R}) e^{-\beta U(\boldsymbol{R})} d\boldsymbol{R} \tag{3.10}$$

Deste ponto em diante consideraremos que o espaço de configurações está suficientemente definido pelo espaço das coordenadas $\boldsymbol{R}$ e que as grandezas a serem calculadas são dependentes apenas das coordenadas do sistema. Por conseguinte a distribuição de probabilidades do ensemble se restringe a

$$P_{\mathrm{NVT}}(\boldsymbol{R}) = \frac{e^{-\beta U(\boldsymbol{R})}}{\int e^{-\beta U(\boldsymbol{R})} d\boldsymbol{R}} \tag{3.11}$$

Há duas maneiras de se reproduzir a distribuição de probabilidades do ensemble. Uma delas é através do método de Dinâmica Molecular [*Allen*, 1989] que gera configurações deterministicamente resolvendo a segunda equação de Newton ($\partial \boldsymbol{P}/\partial t = \boldsymbol{F}$), usando $\boldsymbol{F} = -\partial U(\boldsymbol{R})/\partial \boldsymbol{R}$ e tomando um intervalo de tempo Δt. As propriedades são calculadas como médias temporais sobre a trajetória do sistema. A outra maneira é através do método de Monte Carlo que é um processo estocástico de se resolver as integrais envolvidas no cálculo das médias das propriedades.

Os dois métodos podem produzir a distribuição do ensemble, mas para se obter informações temporais do sistema, apenas a Dinâmica Molecular pode ser empregada. Como estamos interessados em propriedades estruturais e de equilíbrio, o método de Monte Carlo pode responder aos nossos problemas. Este é o método que utilizaremos através da técnica de amostragem de Metropolis e passaremos a descrevê-lo na seção seguinte.

3.2. Método de Monte Carlo.

O método de Monte Carlo, desenvolvido no início da década de 50 [*Metropolis*, 1953], é conhecido em várias áreas do conhecimento por ser um método de resolução de integrais utilizando amostragem aleatória. É recorrentemente usado para resolver integrais do tipo da equação (3.1) onde o espaço de integração é de alta dimensionalidade, tornando inviáveis os métodos numéricos disponíveis.

Consideremos então o cálculo de médias como na equação (3.10). Podemos aproximar as integrais no numerador e no denominador por somas, de modo que podemos escrever

$$\langle f \rangle_{\mathrm{NVT}} = \frac{\sum_{i=1}^{M} f(\boldsymbol{R}_i) e^{-\beta U(\boldsymbol{R}_i)}}{\sum_{i=1}^{M} e^{-\beta U(\boldsymbol{R}_i)}} \tag{3.12}$$

onde $\boldsymbol{R}_i$ é a i-ésima configuração do sistema. No limite $M \to \infty$ restabelece-se a integral.

Num procedimento de cálculo numérico padrão, dividiríamos o intervalo de cada uma das $6N$ variáveis em h pequenos intervalos e teríamos ao todo $M = h^{6N}$ pontos de integração.

Para se ter uma idéia do que significa isso, consideremos um sistema de 10 partículas (bastante pequeno) que não possuem graus de liberdade rotacionais, pra simplificar. Teríamos assim 30 variáveis. Se dividirmos o intervalo de todas as variáveis por 10 (o que daria um valor bastante grosseiro para o resultado da integral) teremos então uma grade de 10^{30} pontos no somatório. Consideremos ainda que um computador possa realizar o cálculo de f e $e^{-\beta U}$ em um ponto da grade a cada nanossegundo (10^{-9}s). Nestas condições teríamos o resultado da integração em 10^{14} anos. É impensável, como se pode constatar, a utilização de métodos numéricos desta natureza para tal problema.

A solução prescrita pelo método de Monte Carlo consiste em tomar pontos do espaço de configurações aleatoriamente ao invés de seqüencialmente, como seria feito acima. Assim nos aproximaríamos mais rapidamente da média mostrada na equação (3.10). Para esta finalidade é necessário que se tenha um bom gerador de números aleatórios.

Antes de prosseguir na descrição do método, devemos generalizar a equação (3.12). A escolha das configurações que irão participar do cálculo da média pode ser governada por uma probabilidade arbitrária $P'(\boldsymbol{R})$. Neste caso podemos reescrever a equação (3.12) como [*Binder*, 1997]

$$\langle f \rangle = \frac{\sum_{i=1}^{M} f(\boldsymbol{R}_i) e^{-\beta U(\boldsymbol{R}_i)} / P'(\boldsymbol{R}_i)}{\sum_{i=1}^{M} e^{-\beta U(\boldsymbol{R}_i)} / P'(\boldsymbol{R}_i)}. \tag{3.13}$$

Se $P'(\boldsymbol{R}) = cte$, então voltamos aos somatórios em (3.12) e temos a chamada amostragem simples. A amostragem simples, como o próprio nome indica, é equivalente a escolher de forma totalmente aleatória as configurações que participarão da média. Esta técnica de amostragem tem um problema: para sistemas compactos, como líquidos por exemplo, este

processo escolherá muitas configurações com baixo peso para a média. Sendo assim, a convergência no valor da média será extremamente lenta nestes sistemas.

Em 1953, Metropolis sugeriu que se tomasse $P'(\boldsymbol{R}) = P_{\text{ens}}(\boldsymbol{R})$ o que significa, no caso do ensemble NVT, $P'(\boldsymbol{R}) = e^{-\beta U(\boldsymbol{R})}$ [*Metropolis*, 1953], de modo que a média na equação (3.13) se torna uma média aritmética simples:

$$\langle f \rangle = \frac{1}{M} \sum_{i=1}^{M} f(\boldsymbol{R}_i). \tag{3.14}$$

Esta técnica é chamada de amostragem relevante[1]. Através dessa técnica, seguindo a prescrição de Metropolis, a convergência no cálculo das médias das grandezas deve ser atingida mais rapidamente.

A idéia de Metropolis envolve também que não se tomem configurações independentemente umas das outras. Este procedimento ainda seria bastante ineficiente pois ainda com toda probabilidade seriam geradas configurações de alta energia e portanto baixa probabilidade [$P'(\boldsymbol{R})$] de serem escolhidas para participar da média da equação (3.14). Estaríamos assim recaindo no mesmo problema da amostragem simples. Ao invés disso, Metropolis sugeriu a construção de uma cadeia de Markov[2] como método de geração de configurações. Esta cadeia é construída de modo que cada configuração $\boldsymbol{R}_{i+1}$ é atingida a partir de uma configuração $\boldsymbol{R}_i$ através de uma taxa de transição (probabilidade de transição por número de transições) adequada $w(\boldsymbol{R}_i \rightarrow \boldsymbol{R}_{i+1})$.

As taxas $w_{ij} = w(\boldsymbol{R}_i \rightarrow \boldsymbol{R}_j)$ podem ser claramente identificadas como elementos de uma matriz $\boldsymbol{W}$. A matriz $\boldsymbol{W}$, chamada matriz de transição, deve satisfazer certos requisitos. O primeiro é a condição de normalização[3]:

$$\sum_j w_{ij} = 1 \tag{3.15}$$

[1] O termo usado em inglês é "importance sampling".
[2] Uma cadeia de Markov é um conjunto seqüencial de configurações em que a probabilidade de se tomar cada configuração depende apenas desta e da configuração anterior. Costuma-se dizer que as configurações geradas não guardam informação sobre os estados visitados anteriormente.
[3] As matrizes que obedecem esta condição são chamadas estocásticas.

que significa que a probabilidade do sistema transitar da configuração i para qualquer outra, mais a probabilidade dele permanecer onde está, é igual a 1. O segundo é a ergodicidade:

$$\forall i, j \quad \exists n > 0, \quad \left(\boldsymbol{W}^n\right)_{ij} > 0 \tag{3.16}$$

ou seja, para qualquer par de estados, existe um intervalo de transições n que permite a passagem de um estado para outro. Por fim, a matriz $\boldsymbol{W}$ deve satisfazer a uma equação de autovalores com autovalor 1. O autovetor correspondente é a distribuição limite, que deve ser igual à distribuição do ensemble.

Uma maneira equivalente de se chegar à distribuição limite é através da equação mestra, que é a equação que governa a evolução da distribuição de probabilidade $P(\boldsymbol{R}_i)$ com o número de transições:

$$\frac{\partial}{\partial n} P(\boldsymbol{R}_i, n) = \sum_j w_{ji} P(\boldsymbol{R}_j, n) - \sum_j w_{ij} P(\boldsymbol{R}_i, n). \tag{3.17}$$

O significado desta equação é bastante simples: a variação da probabilidade de se encontrar o sistema no estado i é dada pela diferença entre a taxa de entrada [$\sum_j w_{ji} P(\boldsymbol{R}_j, n)$] e a taxa de saída [$\sum_j w_{ij} P(\boldsymbol{R}_i, n)$] do estado i. Aqui pode-se ver claramente a dependência com o inverso do número de transições nas taxas.

Quando o sistema está em equilíbrio, a distribuição de probabilidades deve ser estacionária e portanto não pode depender do número de transições, de modo que temos:

$$\sum_j \left[w_{ji} P_{\text{est}}(\boldsymbol{R}_j) - w_{ij} P_{\text{est}}(\boldsymbol{R}_i) \right] = 0. \tag{3.18}$$

Portanto, se houver distribuição limite (P_{est}), ela deve satisfazer a equação (3.18). No nosso caso, como queremos gerar uma amostragem que reproduza a distribuição canônica, devemos tomar $P_{\text{est}}(\boldsymbol{R}_i) = e^{-\beta U(\boldsymbol{R}_i)}$.

A maneira mais comum de se satisfazer à equação (3.18) é através do princípio da reversibilidade microscópica (também chamado de balanço detalhado):

$$w_{ji}e^{-\beta U(\boldsymbol{R}_j)} = w_{ij}e^{-\beta U(\boldsymbol{R}_i)} \tag{3.19}$$

de onde se tem a razão entre as taxas de transição

$$\frac{w_{ij}}{w_{ji}} = \frac{e^{-\beta U(\boldsymbol{R}_j)}}{e^{-\beta U(\boldsymbol{R}_i)}} = e^{-\beta \Delta U} \tag{3.20}$$

com $\Delta U = U(\boldsymbol{R}_j) - U(\boldsymbol{R}_i)$. Portanto, para se obter uma amostragem que reproduza a distribuição canônica, devemos escolher as taxas de transição de modo a satisfazer a equação (3.20).

Existe mais de uma maneira de se satisfazer a equação (3.20). As mais conhecidas são as proposições de Glauber [*Salinas*, 1997]:

$$w_{ij} = \frac{1}{2\tau}\left[1 - \tanh\left(\frac{\beta \Delta U}{2}\right)\right] \tag{3.21}$$

e de Metropolis:

$$w_{ij} = \begin{cases} \dfrac{1}{\tau}e^{-\beta \Delta U} & \text{se } \Delta U > 0 \\ \dfrac{1}{\tau} & \text{se } \Delta U \leq 0 \end{cases}. \tag{3.22}$$

O número de transições característico τ é uma constante que pode ser tomada igual a 1. Alguns autores tomam τ como o volume ao qual as transições são limitadas [*Allen*, 1989]. Este volume está presente no método de geração da cadeia markoviana como veremos na subseção 3.2.1.

Apesar de Metropolis ter desenvolvido o método para o ensemble NVT, o mesmo se aplica perfeitamente a qualquer ensemble. Devemos apenas igualar a razão entre as taxas de transição à distribuição de probabilidades do ensemble desejado como na equação (3.20). No caso do NPT esta distribuição é dada por

$$P_{\text{NPT}}(\boldsymbol{R}, V) = e^{-\beta[U(\boldsymbol{R}) + PV]}. \tag{3.23}$$

Esta forma, entretanto, é inadequada porque o volume, além de estar explícito na probabilidade do ensemble, está também presente implicitamente nas coordenadas de centro de massa ($\boldsymbol{r}$) do sistema ou, mais

precisamente, no domínio de integração destas coordenadas. Isso é ruim porque torna difícil o controle ou monitoração do volume. Através da equação (3.23), a média de uma grandeza $f(\boldsymbol{R})$ neste ensemble será dada por

$$\langle f \rangle_{\mathrm{NPT}} = \frac{\int f(\boldsymbol{R}) e^{-\beta[U(\boldsymbol{R})+PV]} d\boldsymbol{R} dV}{\int e^{-\beta[U(\boldsymbol{R})+PV]} d\boldsymbol{R} dV}. \tag{3.24}$$

É mais interessante, portanto, desvincular o domínio de $U(\boldsymbol{R})$ do volume. E isto se faz tornando as coordenadas dos centros de massa das partículas do sistema adimensionais e relativas ao volume ocupado pelo mesmo, através da transformação

$$\vec{r}_i \rightarrow \left(L_x q_{ix}, L_y q_{iy}, L_z q_{iz} \right) \tag{3.25}$$

onde $\boldsymbol{q} = (\vec{q}_1, \vec{q}_2, ..., \vec{q}_N)$ é o conjunto de coordenadas de centro de massa adimensionais de todas as partículas do sistema. Agora, o domínio de cada coordenada q_{ij}, com $i = 1,2,...,N$ e $j = x,y,z$, é o intervalo [0,1][4]. Com isto a equação (3.24) passa a ser escrita da forma

$$\langle f \rangle_{\mathrm{NPT}} = \frac{\int f(\boldsymbol{Q}) e^{-\beta[U(\boldsymbol{Q})+PV]} V^N d\boldsymbol{Q} dV}{\int e^{-\beta[U(\boldsymbol{Q})+PV]} V^N d\boldsymbol{Q} dV} \tag{3.26}$$

sendo $\boldsymbol{Q} = (\boldsymbol{q}, \boldsymbol{\theta})$, cujo significado é imediato. A distribuição de probabilidades do ensemble NPT pode ser escrita do seguinte modo:

$$P_{\mathrm{NPT}}(\boldsymbol{Q}, V) = e^{-\beta[U(\boldsymbol{Q})+PV]+N\ln V}. \tag{3.27}$$

Fazendo agora $\Delta U = U(\boldsymbol{Q}_j) - U(\boldsymbol{Q}_i)$, $\Delta V = V_j - V_i$ e reconhecendo a entalpia configuracional $H^c = U + PV$, temos que

$$\frac{w_{ij}}{w_{ji}} = \frac{\exp\left(-\beta H_j^c + N \ln V_j\right)}{\exp\left(-\beta H_i^c + N \ln V_i\right)} \tag{3.28}$$

[4] Na verdade este intervalo vai depender das coordenadas dos vértices da caixa que contém o sistema. Se tivermos os vértices nos pontos (0,0,0), $(L_x,0,0)$, $(0,L_y,0)$ e $(0,0,L_z)$ então o intervalo [0,1] se aplicará a todos os q_{ij}.

$$\frac{w_{ij}}{w_{ji}} = \exp\left\{-\beta\Delta H^c + N\ln\left(V_j/V_i\right)\right\} \tag{3.29}$$

Analogamente à proposição de Metropolis teremos as taxas de transição dadas por

$$w_{ij} = \begin{cases} \dfrac{1}{\tau}e^{-\beta\Delta H^c + N\ln\left(V_j/V_i\right)} & \text{se } \Delta H^c - \dfrac{N}{\beta}\ln\left(V_j/V_i\right) > 0 \\ \dfrac{1}{\tau} & \text{se } \Delta H^c - \dfrac{N}{\beta}\ln\left(V_j/V_i\right) \leq 0 \end{cases} . \tag{3.30}$$

Mais uma vez podemos tomar $\tau = 1$. Esta forma da distribuição de probabilidades para o ensemble NPT é mais interessante porque torna a mudança no volume independente da mudança nas coordenadas do sistema que por sua vez são relativas ao volume corrente da simulação. Ao mesmo tempo, o fato de as coordenadas agora serem relativas às medidas da caixa faz com que ao mudar-se o volume do sistema, não se tenham partículas fora deste novo volume nem haja espaço vazio dentro do mesmo, nos casos em que se aumenta ou diminui o volume, respectivamente.

3.2.1. Implementação

No estudo de líquidos moleculares utiliza-se, em geral, o algoritmo de Metropolis na realização do método de Monte Carlo. Em nosso trabalho utilizamos o programa DICE [*DICE*, 1997] cuja implementação segue o seguinte algorítmo:

- 1. Construção da configuração inicial.

  Gera-se uma configuração inicial, ordenada ou aleatória, o que for mais conveniente em cada caso, acondicionada em uma caixa, com lados fornecidos a priori ou de acordo com uma densidade dada. A partir daí replica-se a caixa em todas as direções, impondo condições de contorno periódicas, para eliminar os efeitos de borda.

- 2. Construção de uma nova configuração.

Partindo da configuração anterior, gera-se uma nova configuração transladando-se e rotacionando-se uma partícula sorteada.

A translação é feita calculando-se um novo valor para as coordenadas (x, y, z) do centro de massa da molécula através das equações

$$\begin{aligned} x_{i+1} &= x_i + a\xi_1 \\ y_{i+1} &= y_i + a\xi_2 \\ z_{i+1} &= z_i + a\xi_3 \end{aligned} \tag{3.31}$$

onde os ξ_i são números aleatórios entre -1 e 1 e a é o deslocamento máximo, que pode ser constantemente ajustado ao longo da simulação em prol da rapidez na convergência, da eficiência e da ergodicidade da mesma.

A rotação é promovida sorteando-se um eixo aleatório ξ_4 dentre os eixos x, y, z e um ângulo para rotacionar a molécula através da equação

$$\alpha_{i+1} = \alpha_i + b\xi_5 \tag{3.32}$$

onde b é o ângulo máximo de rotação.

No caso do ensemble NPT há também a mudança no volume que deve ser promovido segundo as equações

$$\begin{aligned} L_{i+1,x} &= L_{ix} + c\xi_6 \\ L_{i+1,y} &= L_{iy} + c\xi_6 \\ L_{i+1,z} &= L_{iz} + c\xi_6 \end{aligned} \tag{3.33}$$

com c (variação máxima dos lados da caixa) também sendo um valor que pode ser ajustado durante a simulação.

- 3. Aceitação ou rejeição da nova configuração.

Dada uma nova configuração, calcula-se a energia $U(\boldsymbol{R})$ da mesma e a diferença de energia $\Delta U = U(\boldsymbol{R}_{i+1}) - U(\boldsymbol{R}_i)$. Se $\Delta U \leq 0$, então a

configuração é aceita e torna-se parte de nossa cadeia markoviana. Se, por outro lado, $\Delta U > 0$, então sorteia-se um número aleatório ξ_7 entre 0 e 1 e compara-se com $e^{-\beta\Delta U}$. Se $\xi_7 > e^{-\beta\Delta U}$, então rejeita-se a nova configuração e a configuração i passa a ser a configuração i+1. Se $\xi_7 \leq e^{-\beta\Delta U}$, então a configuração nova é aceita e passa a ser a configuração i+1.

Para simulações em NPT calcula-se, além de $\Delta U = U(\boldsymbol{Q}_{i+1}) - U(\boldsymbol{Q}_i)$, a diferença de volume $\Delta V = V_{i+1} - V_i$, a razão entre os volumes V_{i+1}/V_i e, finalmente, a quantidade $-\beta\Delta H^c + N\ln\left(V_{i+1}/V_i\right)$. Faze-se então o mesmo teste de aceite e rejeição.

- 4. Cálculo das propriedades termodinâmicas.

Calculam-se então algumas grandezas como energia interna (U_i), primeiro virial e o segundo virial [*Coutinho*, 1997] com a configuração resultante do processo de aceite/rejeição. Acumula-se esses valores para a média e volta-se para a etapa 2 do processo. Cada vez que se chega a este ponto, costuma-se dizer que foi realizado um passo de Monte Carlo.

Ao final da simulação essas médias serão utilizadas para calcular as propriedades termodinâmicas como calor específico a volume ou pressão constante, compressibilidade isotérmica, coeficiente de expansão térmica, etc.

3.2.2. Considerações adicionais

Como foi dito, é necessária a construção de replicas da caixa em todas as direções de modo a eliminar os efeitos de borda, fazendo com que as moléculas próximas ao lado da caixa sintam o mesmo ambiente que as demais. A utilização de réplicas da caixa leva às condições de contorno periódicas. Assim, quando um movimento leva uma molécula para fora da caixa, sua réplica na caixa na direção oposta entra pelo outro lado.

A periodicidade introduzida é indesejável mas pode ser contornada com a adoção de um raio de corte r_c (raio além do qual não se computa mais nenhuma interação intermolecular) que pode assumir um valor máximo igual à metade do lado da caixa. Desta forma nenhuma molécula irá interagir simultaneamente com uma molécula original e com sua réplica.

Também implementamos um raio de *overlap* r_o (distância mínima entre dois átomos) com o intuito de eliminar o desperdício computacional. Isto porque quando dois sítios (vide subseção 3.2.3) de moléculas diferentes estão muito próximos, a energia potencial intermolecular será fortemente repulsiva e portanto tal configuração será rejeitada. Portanto podemos economizar milhares de cálculos fazendo com que a configuração seja rejeitada assim que distâncias menores que r_o sejam detectadas. Esse raio, é claro, depende da temperatura, pois quanto maior a temperatura, mais perto uma molécula pode chegar da outra.

Como vimos, o princípio da reversibilidade microscópica leva a uma distribuição limite que deve ser igual à distribuição do ensemble. Nos primeiros passos, entretanto, temos um transiente em que o sistema está caminhando para o equilíbrio. As configurações geradas nesta etapa - chamada geralmente de *termalização* ou *equilibração* - não constituem configurações típicas de equilíbrio e seu uso na composição das médias a serem calculadas deve ser evitado. Por isso é costumeiro se separar as configurações da termalização das configurações efetivamente de equilíbrio e usar somente essas últimas no cálculo das médias. Ainda durante a termalização, e para acelerar esse estágio, utilizamos outro procedimento que chamamos de *cooling* onde não é feito o teste de aceitação de Metropolis, mas sim são aceitas apenas configurações de energia mais baixa que a anterior. Esse processo dura, em geral, apenas algumas centenas de passos.

3.2.3. Potencial de Interação Molecular

Como pode-se perceber, o potencial de interação $U(\boldsymbol{R})$ tem um papel muito importante no processo de geração do ensemble de configurações.

Como dissemos, este potencial pode ser tratado como uma soma da energia interna de cada molécula e a energia intermolecular. No caso do modelo de moléculas rígidas, a energia interna das moléculas é constante durante todo o processo e, portanto, não influencia na geração do ensemble de configurações.

A energia de interação intermolecular, por outro lado, surge como uma soma de inúmeras contribuições que podem, para efeitos práticos, ser agrupadas em interações de dois, três, quatro corpos, etc - onde os corpos em questão são moléculas.

$$U^{inter}(N) = U_2(N) + U_3(N) + U_4(N) + \cdots + U_N(N). \tag{3.34}$$

A contribuição de pares para a energia é dada por

$$U_2(N) = \sum_{A<B} U_{AB} \tag{3.35}$$

com $U_{AB} = U(AB) - (U_A + U_B)$, sendo que $U(AB)$ é a energia do complexo formado pelas moléculas A e B e U_A é a energia interna da molécula A. A contribuição de três corpos é

$$U_3(N) = \sum_{A<B<C} U_{ABC} \tag{3.36}$$

onde $U_{ABC} = U(ABC) - (U_{AB} + U_{AC} + U_{BC}) - (U_A + U_B + U_C)$. As contribuições de quatro e mais corpos são escritas analogamente.

A maior contribuição para a energia intermolecular é dada pela interação de pares $U_2(N)$. Em muitos sistemas a contribuição de três corpos é relevante mas em sistemas sem orientações preferenciais, como acontece em geral nos líquidos, e a altas temperaturas, a contribuição de três corpos já é bastante pequena [*Xantheas*, 1994; *Cruzan*, 1996], de modo que se torna intuitivo, para modelar potenciais para tais sistemas, truncar a série (3.34) apenas com o primeiro termo. Em verdade, o maior erro em se fazer isto se dá pela não aditividade da energia de indução e de troca[5] [*Stone*, 1996; *Sheykhet*, 1990].

Escrevemos então a energia intermolecular da forma

[5] Em inglês "exchange".

$$U^{inter}(N) = \sum_{A=1}^{N-1} \sum_{B=A+1}^{N} U_{AB}^{eff} \tag{3.37}$$

onde U_{AB}^{eff} agora deve ser um potencial efetivo que inclua a energia de interação de pares e as contribuições de muitos corpos. Como conseqüência, este potencial efetivo de pares deve ser capaz de reproduzir dados experimentais. Com efeito, é através destes dados e também de resultados de estudos *ab initio* que se constrói o potencial efetivo U_{AB}^{eff}. A aproximação (3.37) é comumente chamada de aproximação de aditividade de pares, mas é melhor interpretada como uma forma matemática conveniente de se reproduzir a interação intermolecular real.

Usualmente, este potencial efetivo é escrito como uma soma de potenciais de pares entre os sítios das moléculas, de modo a dar maior generalidade ao modelo de interação de pares. Estes sítios são as unidades básicas das moléculas, podendo ser os próprios átomos ou até certos grupos como o metil (CH_3) por exemplo. Assim podemos definir

$$U_{AB}^{eff} = \sum_{i \in A}^{n_A} \sum_{j \in B}^{n_B} U_{ij} \tag{3.38}$$

onde i e j correm sobre os sítios das moléculas A e B respectivamente e n_A e n_B são os números de sítios nestas moléculas. U_{ij} é o potencial de interação sítio a sítio. Em sua maioria, os sítios são representados pelos próprios átomos.

A energia potencial intermolecular se vê então reduzida a uma grande soma de potenciais de interação de pares sítio-sítio. Como não se conhece a expressão formal deste potencial, há várias formas de se escrever U_{ij} analiticamente. Todas estas formas apresentam características comuns: uma parte fortemente repulsiva a pequenas distâncias; e uma parte atrativa dominante a médio e longo alcance que se anula quando $r_{ij} \to \infty$. Normalmente a parte atrativa assume a forma funcional r_{ij}^{-6}, ao contrário da parte repulsiva que assume diversas formas funcionais. Uma discussão detalhada sobre a origem destes termos de curto e longo alcance pode ser encontrada em [*Stone*, 1996].

A combinação dos termo atrativo e repulsivo gera um poço de potencial de profundidade ε e uma distância σ aquém da qual o potencial total é repulsivo. Estes são justamente os parâmetros usados no potencial Lennard-Jones (LJ) [*Lennard-Jones*, 1937] que é o mais comumente usado em simulações de líquidos. Esse é o modelo de potencial de interação que adotamos em nossa implementação do método de Monte Carlo.

A forma usual do potencial LJ é a seguinte:

$$U_{ij}^{\mathrm{LJ}} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{3.39}$$

onde r_{ij} é a distância entre os dois sítios e ε_{ij} e σ_{ij} são obtidos a partir dos parâmetros individuais dos sítios através das médias geométricas

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \quad \text{e} \quad \sigma_{ij} = \sqrt{\sigma_i \sigma_j}. \tag{3.40}$$

Adicionalmente ao potencial LJ, é comum também usar o potencial de Coulomb. Esta parte do potencial é responsável pela descrição mais detalhada das interações eletrostáticas, dentre elas interações específicas como ligações de hidrogênio, por exemplo. Sendo assim, nosso potencial U_{ij} finalmente é dado por

$$U_{ij} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{r_{ij}} \tag{3.41}$$

Outros modelos de potenciais de interação existem e são bastante utilizados como o de Buckingham, onde a parte repulsiva do potencial LJ é substituída pela forma exponencial ($B_{ij}e^{-C_{ij}r_{ij}}$).

3.2.4. Função de Distribuição Radial

Para se estudar propriedades estruturais de sistemas moleculares é comum usar a função de distribuição radial de pares (RDF ou $G(r)$). Como o próprio nome indica, esta é uma função que descreve como está disposto radialmente no espaço um sítio em relação a outro. Essa função dá a

probabilidade de se encontrar um par de sítios a uma distância r relativa à mesma probabilidade para um gás ideal de mesma densidade.

Essa função pode ser obtida experimentalmente através de experiências de difração de raios-X e de nêutrons [*Elliott*, 1990; *Gingrich*, 1943; *Guinier*, 1955]. Teoricamente a RDF pode ser calculada através da integração da função de distribuição do ensemble em todas as coordenadas menos as dos dois sítios envolvidos. No ensemble NVT será da forma:

$$G_{ij}(r) = \frac{N(N-1)}{\rho^2 Z} \int e^{-\beta U(\boldsymbol{R})} d\vec{r}_1 \cdots d\vec{r}_{i-1} d\vec{r}_{i+1} \cdots d\vec{r}_{j-1} d\vec{r}_{j+1} \cdots d\vec{r}_N \tag{3.42}$$

com ρ sendo a densidade do sistema.

Para a simulação é usada uma forma mais simples dada por

$$G_{ij}(r + \frac{1}{2}dr) = \frac{n_{ij}(r, r+dr)}{n^{id}(r, r+dr)} \tag{3.43}$$

onde $n_{ij}(r, r+dr)$ é o número de pares (i, j) que estão separados por uma distância entre r e $r+dr$ e n^{id} é o número equivalente de pares num gás ideal de mesma densidade dado por:

$$n^{id}(r, r+dr) = \frac{4\pi}{3} \rho \left[(r+dr)^3 - r^3 \right]. \tag{3.44}$$

Na Figura 3.1 ilustramos a estrutura de um líquido atômico e a representação dessa estrutura pela função de distribuição radial.

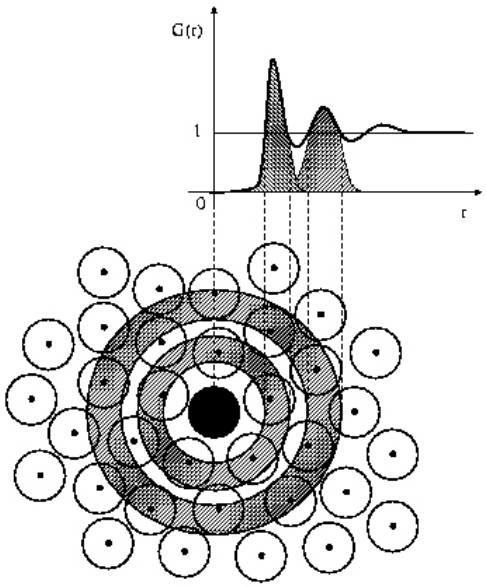


Figura 3.1: Esquema da representação da estrutura de um líquido pela RDF. Tirada de [*Coutinho*, 1997].

Como podemos observar, os picos da RDF definem as camadas de solvatação. É através da integração desses picos em cascas esféricas que podemos calcular o número de moléculas de solvente que se distribuem nas camadas de solvatação ao redor do soluto. Essa integração é dada por

$$N_S(r) = 4\pi\rho\int_0^r G(r)r^2 dr \tag{3.45}$$

O número de moléculas que resulta da integração do primeiro pico da RDF entre centros de massa de soluto e solvente é conhecido como número de coordenação. Podemos com a RDF apropriada contar também o número de ligações de hidrogênio que podem estar sendo formadas.

3.2.5. Ergodicidade

A ergodicidade, ou seja, a capacidade de se atingir qualquer estado durante a simulação, pode ser verificada para o Monte Carlo Metropolis como mostraremos a seguir.

Do ponto de vista da matriz de transição $\boldsymbol{W}$ podemos ver que ela satisfaz à condição (3.16). Tomando os elementos de matriz em (3.22) podemos ver que $w_{ij} \to 0$ somente quando $\Delta U \to \infty$ o que ocorre, segundo a

equação (3.41), apenas quando $r_{ij} \to 0$ para algum par de sítios (i, j). Mas tal situação não corresponde a uma configuração física, já que a sobreposição excessiva das nuvens eletrônicas é altamente desfavorável. Portanto temos que todas as configurações físicas do sistema podem ser alcançadas a partir de qualquer outra através de uma taxa de transição $w_{ij} > 0$.

Do ponto de vista prático da simulação, em cada passo uma mudança limitada ocorre com a configuração inicial. Assim é impossível, em apenas um passo, passar-se de uma dada configuração para uma outra em que as posições das moléculas estejam fora do alcance dado pelas equações (3.31) e (3.32). Entretanto, após um número suficiente de passos, pode-se alcançar qualquer estado partindo-se de qualquer outro. Portanto, o método é ergódico dado um número suficiente de passos.

3.3. Resultados

3.3.1. Interação da α-CD com a água

Utilizando o método de Monte Carlo Metropolis, geramos dois conjuntos de configurações de α-CD em água. Um com a α-CD na geometria AMBER e outro na geometria HF.

Os parâmetros de Lennard-Jones (ε e σ) usados para as duas geometrias da α-CD foram retirados do campo de força AMBER, que se mostrou mais eficiente em descrever a geometria da α-CD. As cargas parciais para as duas geometrias foram obtidas através de cálculos quânticos *ab initio* (método HF/6-31G com técnica CHelpG de ajuste de cargas para reprodução do potencial eletrostático).

Para a água foram utilizados tanto os parâmetros LJ e cargas parciais como também a geometria do potencial SPC [*Berendsen*, 1981]. Esse potencial para a água constitui-se de três sítios que são os próprios átomos. Todos os parâmetros utilizados nas simulações estão no apêndice.

Todas as simulações foram feitas à temperatura ambiente (298.15 K) e à pressão de uma atmosfera. Utilizamos uma caixa cúbica com uma

molécula de α-CD e 1000 moléculas de água. Em todas as simulações geramos a configuração inicial aleatoriamente e utilizamos $150 \cdot 10^3$ passos de *cooling* para abaixar a energia do sistema antes da termalização.

Começamos por realizar a termalização da α-CD em água no ensemble NPT para determinar qual a alteração na densidade do sistema devido à inclusão dessa molécula e conseqüentemente qual o volume da caixa a ser utilizado na simulação no ensemble NVT. Na Figura 3.2 mostramos o gráfico da termalização com a geometria AMBER.

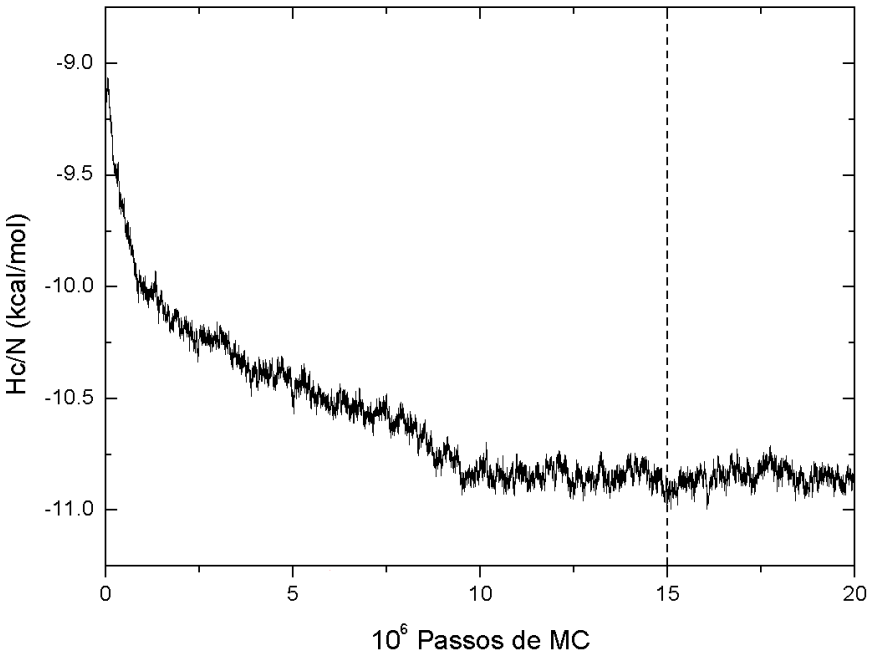


Figura 3.2: Termalização da α-CD em água com a geometria AMBER.

Nessa primeira simulação observamos, ao final de $20 \cdot 10^6$ passos, que o sistema se ajustou de modo a ter densidade igual a 1.03 g/cm^3. Isso significa que o lado da caixa deve ser de 31.28 Å. Considerando as dimensões típicas da α-CD, temos então um intervalo mínimo de 9 Å entre esta molécula e os limites da caixa, ocupado por moléculas de água.

Na Tabela 3.1 resumimos os detalhes computacionais envolvidos nessas duas simulações.

Tabela 3.1: Detalhes computacionais das simulações realizadas para as duas geometrias. A densidade foi determinada através de termalização do sistema com a geometria AMBER no ensemble NPT.

Simulações de α-CD em água	
α-CD:água	1:1000
Temperatura	298.15 K
Pressão	1 atm
Densidade	1.03 g/cm^3
L (caixa cúbica)	31.28 Å
Passos na termalização	$20 \cdot 10^6$
Passos em equilíbrio	$40 \cdot 10^6$

Considerando a densidade do sistema de 1.03 g/cm^3, realizamos as simulações no ensemble NVT com as geometrias AMBER e HF. Para a geometria HF realizamos outra termalização também com o ensemble NVT. Na Figura 3.3 mostramos a evolução da energia na simulação da geometria HF.

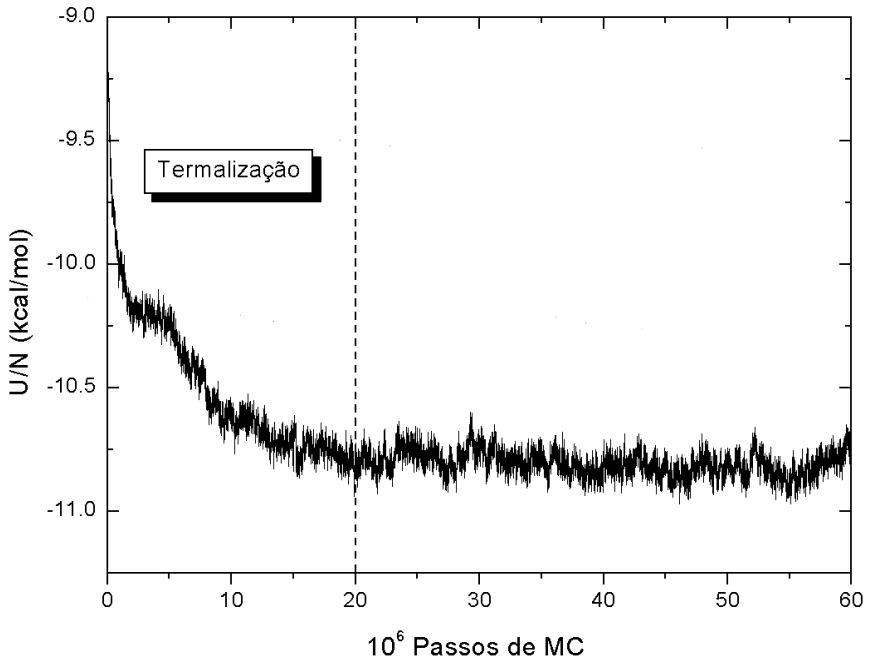


Figura 3.3: Evolução da energia potencial total por molécula U/N durante a simulação de α-CD (geometria HF) em água. Os primeiros $20 \cdot 10^6$ passos são de termalização. Após esse estágio podemos ver que a energia flutua em torno da média.

Para ambas as geometrias realizamos $40 \cdot 10^6$ passos de simulação em estágio de equilíbrio. Podemos ver que as configurações geradas perfazem uma distribuição limite através dos histogramas de energia potencial total por molécula (U/N). Nos gráficos da Figura 3.4 e da Figura 3.5 mostramos essas distribuições.

Para a geometria AMBER obtivemos que a energia potencial total média por molécula é de[6] -10.86 ± 0.05 kcal/mol enquanto que para a geometria HF obtivemos -10.81 ± 0.05 kcal/mol. Num sistema composto por uma molécula de soluto e $N - 1$ moléculas de solvente, a energia potencial total engloba a interação solvente-solvente num total de $(N - 1)(N - 2)$ termos e solvente-soluto num total de $N - 1$ termos. Assim, é difícil se extrair precisamente a diferença entre a interação das duas geometrias com a água. Por isto, é mais adequado analisar diretamente a energia de interação soluto-solvente para verificar qual geometria interage mais com a água.

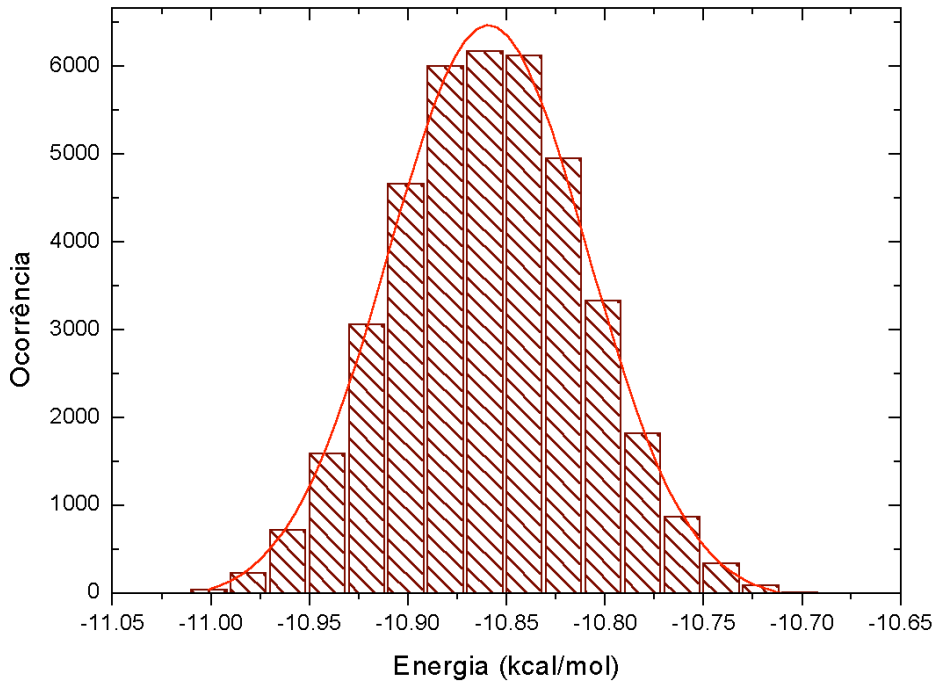


Figura 3.4: Distribuição de energia potencial total por molécula (U/N) para a geometria AMBER. Podemos ver que a distribuição é gaussiana e portanto reflete uma distribuição estacionária, ou seja, formada por configurações de equilíbrio.

[6] Os erros apresentados na forma $\pm e$ se referem ao desvio padrão da grandeza.

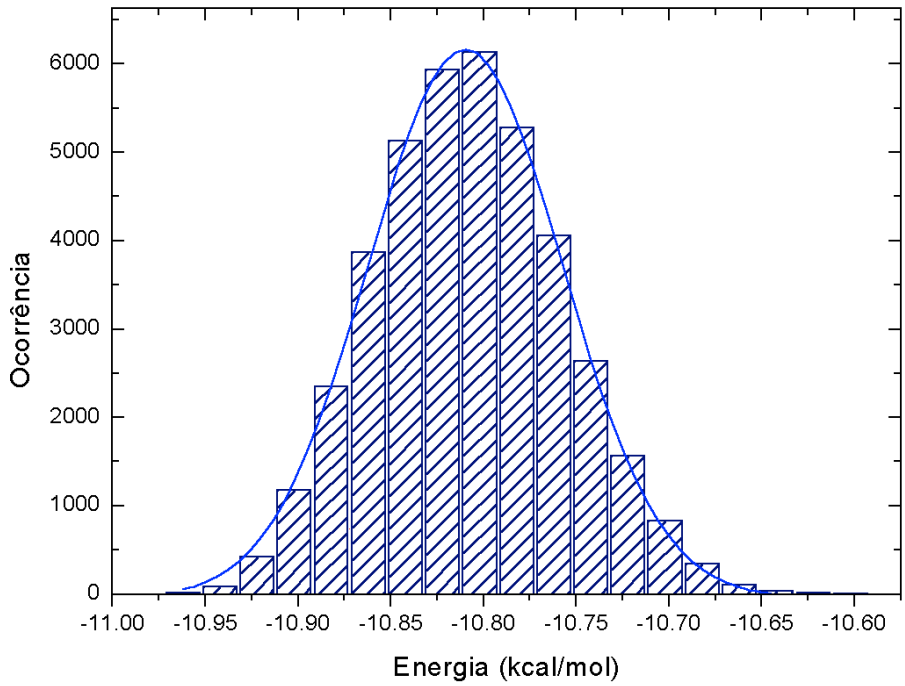


Figura 3.5: Distribuição de energia potencial total por molécula (U/N) para a geometria HF. Essa também é uma distribuição gaussiana mostrando que as configurações que a compõem são configurações de equilíbrio.

A energia de interação soluto-solvente U_{SS} deve ser fortemente influenciada pela formação de ligações de hidrogênio entre a α-CD e a água e deve também nos fornecer informações antecipadas sobre a competição entre as ligações intra e intermoleculares.

Para a geometria AMBER obtivemos uma energia de interação soluto-solvente média de –168 ± 5 kcal/mol enquanto que para a geometria HF obtivemos –147 ± 5 kcal/mol. Uma diferença de energia de interação α-CD-água média de cerca de 21 ± 10 kcal/mol.

Esse resultado indica que a geometria AMBER interage mais fortemente com a água do que a geometria HF. A principal diferença entre essas duas geometrias é a formação de ligações de hidrogênio internas com duas hidroxilas primárias na geometria HF que não ocorrem na geometria AMBER. Por isso, observamos que a formação de ligações de hidrogênio internas com os oxigênios primários em detrimento de ligações intermoleculares é energeticamente desfavorável em solução aquosa.

É interessante mostrar os histogramas de energia de interação soluto-solvente (Figura 3.6) porque a pequena sobreposição entre eles ilustra bem

a diferença de energia de interação da α-CD com a água nas geometrias AMBER e HF. Note que essa diferença marcante não é observada na energia potencial total mostrada na Figura 3.4 e na Figura 3.5.

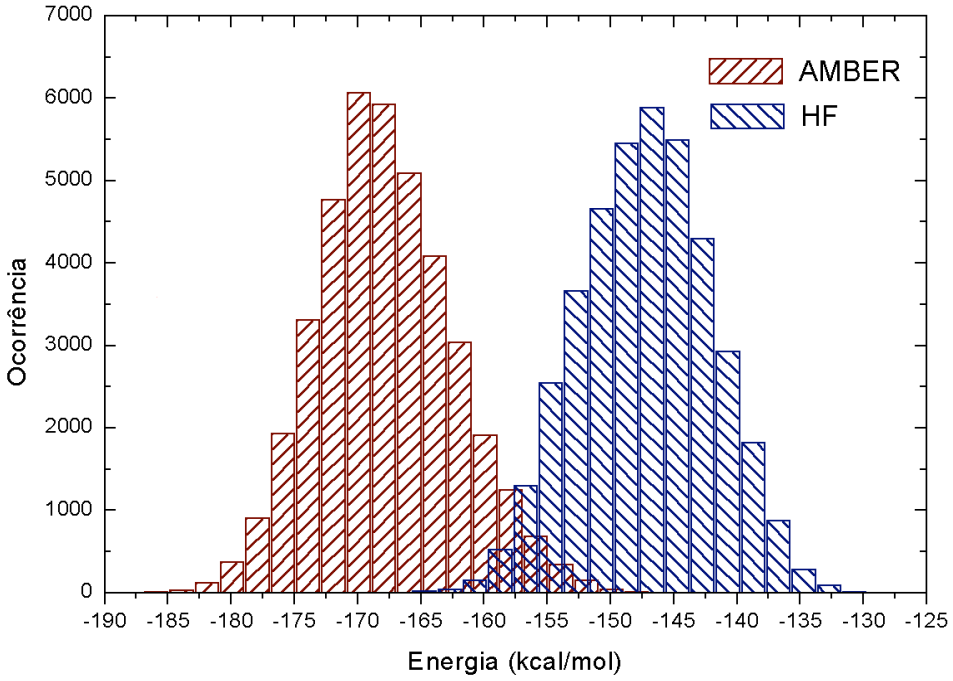


Figura 3.6: Histogramas de energia de interação soluto-solvente U_{SS}.

3.3.2. Análise das RDFs

A diferença de energia de interação soluto-solvente pode ser analisada através das ligações de hidrogênio formadas entre a α-CD e a água. Essa análise pode ser feita inicialmente através da integração esférica das RDFs entre os oxigênios da α-CD e os oxigênios das moléculas de água. Os gráficos das RDFs para as duas geometrias são mostrados na Figura 3.7 e Figura 3.8. Cada RDF se refere a um tipo de oxigênio diferente na molécula de α-CD.

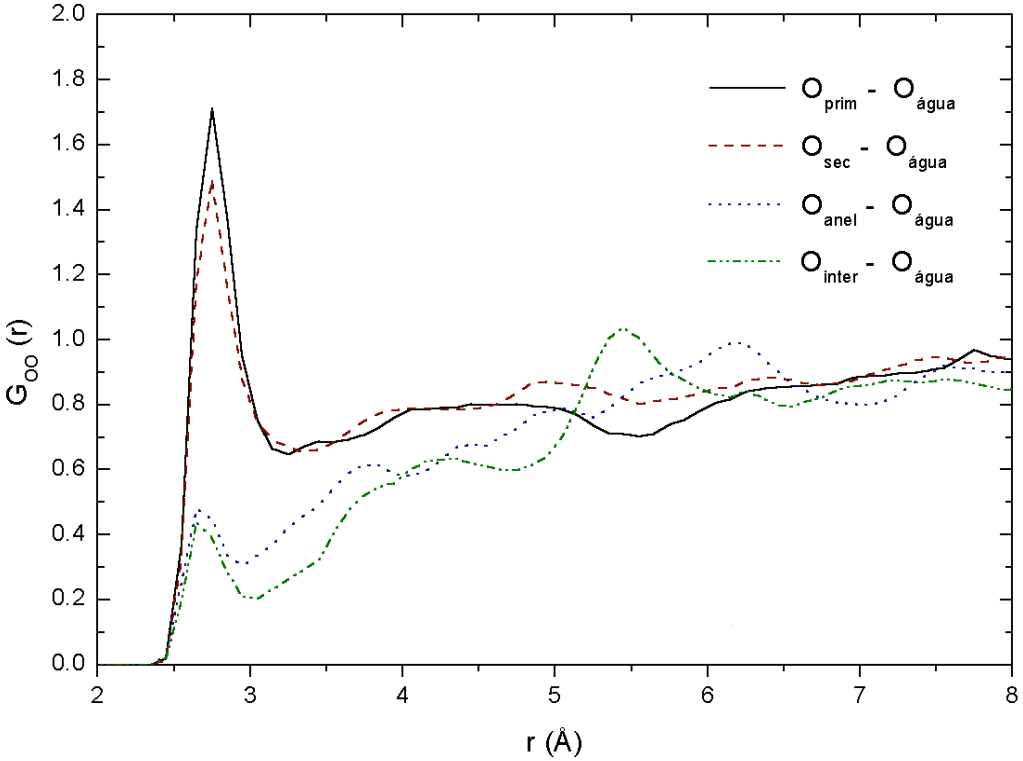


Figura 3.7: RDFs entre os oxigênios da α-CD (AMBER) e das moléculas de água.

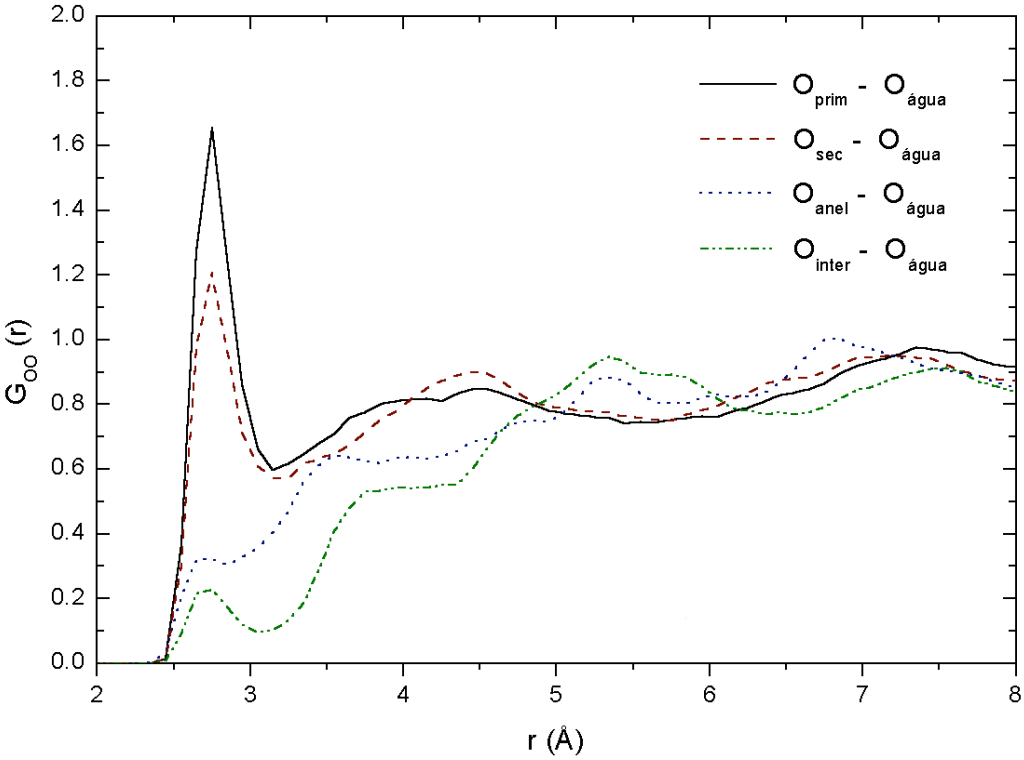


Figura 3.8: RDFs entre os oxigênios da α-CD (HF) e das moléculas de água.

Notamos que as RDFs da geometria AMBER apresentam o primeiro pico mais pronunciado do que na geometria HF. Assim, integrando o primeiro pico de cada RDF calculamos o número de moléculas de água ($N_{água}$) ligadas aos oxigênios da α-CD nas duas geometrias. Esses valores estão na Tabela 3.2.

Tabela 3.2: Resumo da análise das RDFs para as duas geometrias. R_{OO} é a distância do mínimo da primeira camada para os respectivos tipos de oxigênios. $N_{água}$ é o total de moléculas de água próximas a cada tipo de oxigênio. $N_{água}/N_{oxig}$ é o número médio de moléculas de água por oxigênio.

	AMBER			HF		
	R_{OO}	$N_{água}/N_{oxig}$	$N_{água}$	R_{OO}	$N_{água}/N_{oxig}$	$N_{água}$
O_{prim} - $O_{água}$	3.25	2.66	15.96	3.15	2.21	13.24
O_{sec} - $O_{água}$	3.35	2.76	33.08	3.25	2.03	24.40
O_{anel} - $O_{água}$	2.95	0.57	3.44	2.85	0.36	2.15
O_{inter} - $O_{água}$	3.05	0.55	3.31	3.05	0.30	1.78
TOTAL	-	-	55.79	-	-	41.57

Nessa tabela observamos que na primeira camada de solvatação existem cerca de 16 moléculas de água nas proximidades dos 6 oxigênios primários (O6) na geometria AMBER, cerca de 33 águas nos 12 oxigênios secundários (O2 e O3), cerca de 3 águas nos 6 oxigênios dos anéis de piranose (O5) e também cerca de 3 águas nos 6 oxigênios intersacarídeos (O4). Assim temos cerca de 55 moléculas de água na cercania da α-CD na geometria AMBER. Para a geometria HF, esse número é reduzido para 41, onde cerca de 13 águas estão próximas aos oxigênios primários, 24 nos secundarios, 2 nos oxigênios piranosídicos e 2 nos intersacarídeos. Dessa forma, temos uma diferença de 14 moléculas de água na primeira camada de solvatação a mais na geometria AMBER, com cerca de 9 águas a mais nas hidroxilas secundárias.

Considerando que essa diferença de 14 moléculas de água na primeira camada de moléculas de água em torno dos oxigênios da α-CD será a maior responsável pela diferença de 21 kcal/mol de energia de interação com a água nas duas geometrias, temos então uma energia de ligação média entre essas águas e a α-CD de 1.5 kcal/mol por molécula de

água. Esse valor é muito pequeno quando comparado com valores típicos de ligação de hidrogênio [*Mezei*, 1981; *Stillinger*, 1980; *Malaspina*, 2000]. Assim concluímos que nem todas as moléculas de água encontradas na integração das RDFs devem estar formando ligações de hidrogênio com a α-CD.

3.3.3. Ligações de hidrogênio

Para analisarmos a formação de ligações de hidrogênio entre α-ciclodextrina e a água utilizamos o critério geométrico-energético [*Mezei*, 1981; *Stillinger*, 1980]. Esse critério consiste de: um valor máximo para a distância entre os oxigênios (R_{OO}); um valor máximo para o ângulo formado pelo oxigênio aceitador, oxigênio doador e o hidrogênio doado (θ_{OOH}); e um valor mínimo para a energia de ligação entre a α-CD e a água (E_{lig}). É claro que uma ligação de hidrogênio pode ter valores para estas grandezas dentro de um grande intervalo - um limite superior de 70.53º pode ser atribuído à grandeza θ_{OOH} por razões geométricas e qualitativas [*Mezei*, 1981]. Entretanto um valor intuitivo para se tomar para R_{OO} máximo é aquele que abrange o número de coordenação dado pelas funções de distribuição radial, mostrados na Tabela 3.2. Quanto às outras duas grandezas tomamos valores que podem ser considerados de uma ligação de hidrogênio forte, que são $\theta_{OOH} \leq 30º$ e $E_{lig} \geq 3.5$ kcal/mol.

Através desse critério analisamos o comportamento das duas geometrias em relação à formação de ligações de hidrogênio. Na Tabela 3.3 mostramos o número médio de ligações formadas com a água para cada tipo de oxigênio e o número de ligações por oxigênio. Junto com o tipo de oxigênio envolvido na ligação, discriminamos se a α-CD é aceitadora ou doadora na ligação. Relacionamos também as cargas parciais sobre os oxigênios e hidrogênios a eles ligados na α-CD.

Tabela 3.3: Número médio de ligações de hidrogênio $N_{água}$ formadas com cada tipo de oxigênio e de ligação com a água através do critério $\theta_{OOH} \leq 30°$, $E_{lig} \geq 3.5$ kcal/mol e R_{OO} dado pela primeira camada definida na respectiva RDF. Aceitadora/doadora se refere à α-CD. ⟨q⟩ é a carga parcial média sobre o átomo.

	AMBER			HF		
	$\langle N_{água} \rangle$	$\langle N_{água}/N_{ox} \rangle$	$\langle q \rangle$	$\langle N_{água} \rangle$	$\langle N_{água}/N_{ox} \rangle$	$\langle q \rangle$
O_{prim}/ac	6.14	1.02	-0.88 (O6)	7.28	1.21	-0.83 (O6)
O_{prim}/do	4.30	0.72	0.48 (H6)	2.29	0.38	0.49 (H6)
Total O_{prim}	10.44	1.74	-	9.57	1.60	-
O_{sec}/ac	9.11	1.52	-0.82 (O2) -0.85 (O3)	8.62	1.44	-0.86 (O2) -0.88 (O3)
O_{sec}/do	6.73	1.12	0.41 (H2) 0.50 (H3)	4.67	0.78	0.52 (H2) 0.52 (H3)
Total O_{sec}	15.84	1.32	-	13.29	1.11	-
O_{anel}/ac	2.32	0.39	-0.71 (O5)	1.35	0.23	-0.66 (O5)
O_{inter}/ac	2.38	0.40	-0.75 (O4)	1.09	0.18	-0.65 (O4)
TOTAL	30.98	-	-	25.30	-	

Comparando a Tabela 3.2 e a Tabela 3.3, verificamos que das 55 águas presentes na primeira camada de solvatação da α-CD na geometria AMBER, cerca de 31 estão formando ligações de hidrogênio com a mesma. Para a geometria HF, das 44 águas, cerca de 25 estão formando essas ligações. Nesse caso a geometria AMBER tem cerca de 6 ligações de hidrogênio com a água a mais que a geometria HF. Assim considerando que essa diferença de 6 ligações será a maior responsável pela diferença de energia de interação da α-CD com a água entre as duas geometrias, temos uma interação média de –3.8 kcal/mol por ligação de hidrogênio, o que é um valor razoável para essa energia quando comparado aos valores típicos de ligações de hidrogênio.

Podemos ver que as hidroxilas primárias formam duas ligações a mais como doadoras na geometria AMBER do que na geometria HF. Isso porque dois hidrogênios H6 (nos grupos primários) na geometria HF estão

envolvidos em ligações de hidrogênio internas e portanto não podem participar de nenhuma outra ligação, como já era esperado.

Observamos também, para as duas geometrias que, embora as hidroxilas secundárias formem mais ligações de hidrogênio, as OHs primárias formam mais ligações por hidroxila. Isto também é natural já que as OHs secundárias estão envolvidas em ligações de hidrogênio intramoleculares. No caso da geometria HF há proporcionalmente menos ligações com as OHs secundárias em relação à geometria AMBER do que com as OHs primárias, confirmando o que as RDFs indicavam. Isso se deve principalmente ao fato de na geometria HF as ligações entre as OHs secundárias serem mais fortes. Com o anel secundário mais estreito na geometria HF, há menos espaço para formar ligações com as hidroxilas secundárias.

De um modo geral os oxigênios piranosídicos e intersacarídeos participam de um número bem menor de ligações. Isso ocorre por serem estes oxigênios tipo éster e portanto com carga parcial diminuta em relação aos oxigênios de hidroxilas. No caso da geometria HF, dois oxigênios piranosídicos estão envolvidos em interação interna com OHs primárias o que diminui o número de possíveis ligações e contribuindo para o menor número médio dessas em relação à geometria AMBER.

Quanto à diferença no número de ligações com os oxigênios intersacarídeos entre as duas geometrias se deve principalmente ao fato de na geometria AMBER a carga parcial sobre este oxigênio ser um pouco maior que na geometria HF (ver Tabela 3.3 ou apêndice, átomo tipo 3).

Comparando o número de ligações de hidrogênio ($N_{água}$) e as cargas parciais médias ($\langle q \rangle$) nos átomos de oxigênio e hidrogênio da α-CD envolvidos nessas ligações, observamos que só nos oxigênios intersacarídeos (O4) e dos anéis de piranose (O5) mostram a mesma tendência, ou seja, a maior carga parcial resulta em maior quantidade de ligações. Nos outros oxigênios não há essa relação. Assim, vemos que o potencial eletrostático não é dominante na definição das ligações de hidrogênio. Acreditamos que, nesse caso, as características geométricas são decisivas na definição dessas ligações.

É interessante notar que eventualmente existem moléculas de água que fazem duas ligações com a α-CD, ora fazendo duas ligações com hidroxilas secundárias consecutivas, com uma hidroxila secundária e um oxigênio intersacarídeo, ou fazendo uma ligação com um oxigênio piranosídico e uma hidroxila primária. Na Figura 3.9 e na Figura 3.10 ilustramos essas ocorrências.

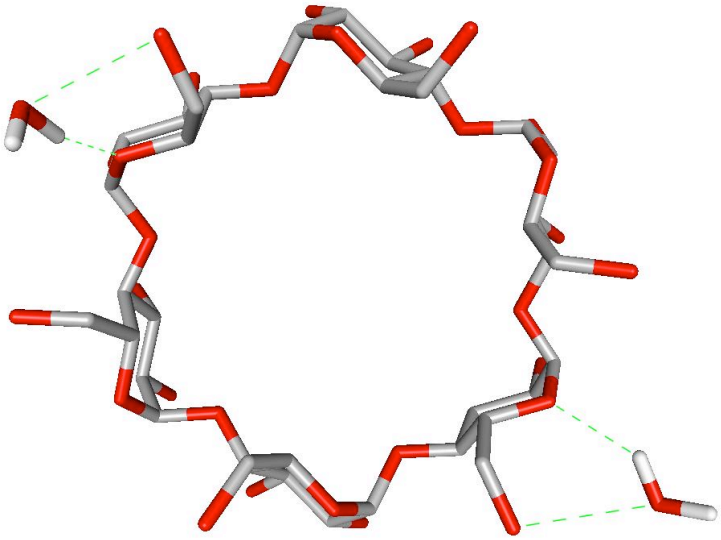


Figura 3.9: Ligações de hidrogênio formadas entre a água e a α-CD. Nesta configuração podemos observar a formação de duas ligações com cada água. Uma através da hidroxila primária e outra através do oxigênio piranosídico.

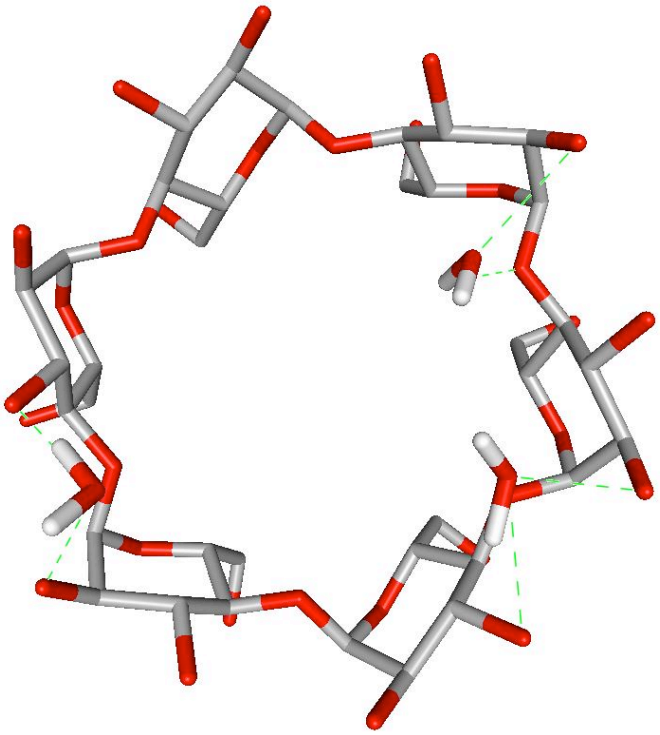


Figura 3.10: Ligações de hidrogênio formadas entre a água e a α-CD. Mais uma vez podemos observar a formação de duas ligações com cada água. Neste caso temos as duas ligações formadas com as hidroxilas secundárias em duas moléculas de água e a outra molécula de água forma uma ligação com uma hidroxila secundária e outra com um oxigênio piranosídico.

3.3.4. Cavidade da α-CD

Para analisar a cavidade da α-CD, recorremos às RDFs entre os centros de massa dessa e das águas. A Figura 3.11 mostra essas funções de distribuição.

Pelas RDFs podemos ver que o pico da $G_{cm\text{-}cm}(r)$ para a geometria HF é mais pronunciado. Entretanto, a primeira camada na $G_{cm\text{-}cm}(r)$ da geometria AMBER se estende até 3.7 Å enquanto que a HF se estende até 3 Å. Com a integração das respectivas primeiras camadas obtivemos que há em média 5 moléculas de água na cavidade da α-CD geometria AMBER e 4 moléculas de água na geometria HF. Isso ocorre por dois motivos: primeiro pela formação de mais ligações de hidrogênio com os oxigênios intersacarídeos na geometria AMBER; o segundo é o fato de a cavidade na geometria AMBER ser um pouco maior por ter uma forma cônica mais

aberta que a da geometria HF e distância média entre os oxigênios intersacarídeos ligeiramente maior.

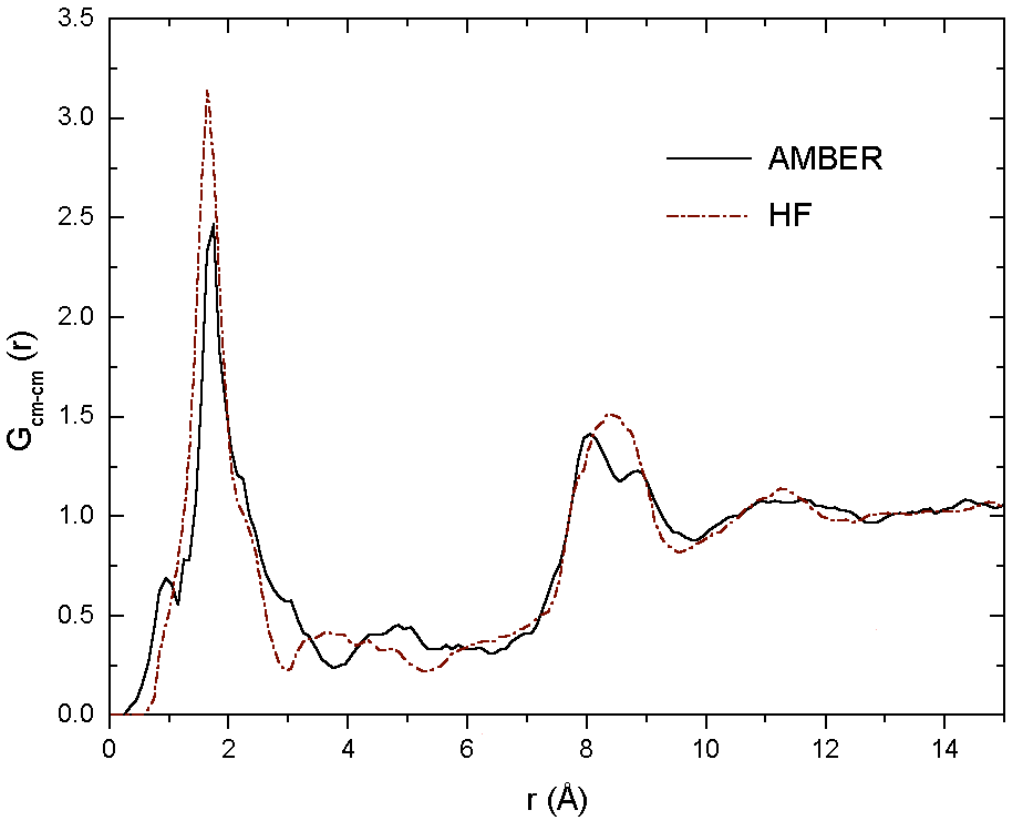


Figura 3.11: RDFs entre os centros de massa da α-CD e das águas para as duas geometrias (AMBER e HF).

Dados experimentais mostram que na ausência de moléculas hospedeiras e em solução aquosa, a cavidade da α-CD abriga cerca de 3 [*Chacko*, 1981] ou mais [*Puliti*, 2000] moléculas de água apesar dessa cavidade ter caráter apolar [*Szejtli*, 1998]. Essas moléculas de água em geral ocupam a cavidade sem orientação preferencial e eventualmente formam ligações de hidrogênio com oxigênios intersacarídeos ou hidroxilas secundárias [*Manor*, 1974]. É interessante salientar que nossos resultados estão em concordância com essas observações experimentais.

Adicionalmente, analisamos o caráter hidrofóbico da cavidade da molécula como sendo a capacidade de expulsar moléculas de água para fora da cavidade. Para isso calculamos o volume efetivo da cavidade da α-CD. Vimos que a α-CD é bastante regular (Tabela 2.6). Assim, consideramos o volume efetivo da cavidade como o volume de um tronco de cone de altura

7.9 Å e raio médio dado pela circunferência inscrita ao hexágono regular formados pelos oxigênios intersacarídeos (O4). Esse raio também é equivalente à distância $R_{O4\text{-}O4}$ subtraída da metade da distância mínima entre os oxigênios intersacarídeos e os hidrogênios da água. Esse valor subtraído de $R_{O4\text{-}O4}$ funciona como um raio de van der Waals entre oxigênio e hidrogênio e foi retirado da $G_{O\text{-}H}(r)$.

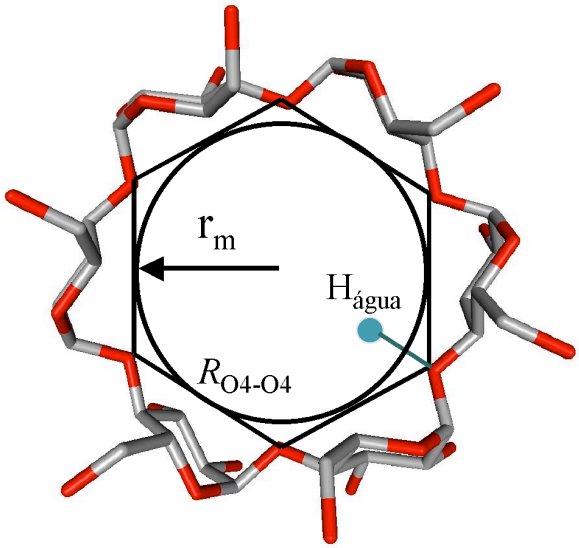


Figura 3.12: Ilustração do raio médio do tronco de cone formado na cavidade da α-CD. Note que os oxigênios intersacarídeos formam um hexágono bastante regular.

Tendo o raio médio e o ângulo de inclinação do cone τ para as duas geometrias, calculamos o volume efetivo da cavidade na geometria AMBER como 326.03 $Å^3$ e na geometria HF 316.45 $Å^3$. Conhecendo a densidade da água (1.0 g/cm^3 = $3.4 \cdot 10^{-2}$ águas/$Å^3$) calculamos que nesses volumes deveriam existir 10.8 e 11.1 moléculas de água nas geometrias HF e AMBER, respectivamente. Portanto, como a geometria HF tem em média 3.6 águas em sua cavidade, ela expulsa 7.2 moléculas de água que caberiam no volume livre da cavidade mas não foram aceitas devido ao potencial intermolecular que gera um efeito hidrofóbico. Para a geometria AMBER observamos que 5.2 moléculas foram expulsas da cavidade. Por esta aproximação, o efeito hidrofóbico da geometria AMBER é menor, proporcionalmente, que o da geometria HF. A densidade da água na cavidade da geometria HF é 0.33 g/cm^3 enquanto na geometria AMBER a densidade da água é 0.46 g/cm^3. Isso está de acordo com o fato de a carga

parcial sobre os oxigênios intersacarídeos ser menor na geometria HF do que na AMBER o que torna, em princípio, a cavidade da geometria HF mais hidrofóbica.

Podemos ver também que há uma região após os primeiro pico que se estende até 7 Å em que as RDFs assumem um patamar bastante baixo. Esta região compreende o tronco de cone formado pela molécula de α-ciclodextrina e portanto não há moléculas de água nesta região a menos das que podem estar acima ou abaixo do cone. Logo após esta região temos outra, entre 7 Å e 10 Å, que mostra uma camada de água que está solvatando a molécula inteiramente. Ainda podemos detectar um terceiro pico até cerca de 13 Å e a partir daí as RDFs vão a 1 assumindo um comportamento de gás ideal de mesma densidade.

3.3.5. Conclusão parcial

A partir dos resultados obtidos das simulações da α-CD em solução aquosa com as geometrias AMBER e HF, observamos que a geometria AMBER interage mais fortemente com a água por cerca de 21 kcal/mol em comparação com a geometria HF. Analisando a solvatação dos oxigênios da α-CD encontramos para a geometria AMBER que cerca de 55 águas estão presentes na primeira camada porém só 31 águas estão efetivamente formando ligações de hidrogênio. Para a geometria HF, 44 águas estão na primeira camada, porém só 25 estão formando ligações de hidrogênio. Assim identificamos que a maior interação da geometria AMBER com a água é provocada por essa maior quantidade de ligações de hidrogênio (cerca de 6).

Analisando as características estruturais dessas duas geometrias e o potencial eletrostático provocado pelas cargas parciais nos átomos, vimos que o efeito das características estruturais tem papel importante para a definição das ligações de hidrogênio.

Adicionalmente analisamos a hidrofobicidade da cavidade da α-CD. Verificamos ambas as geometrias têm uma cavidade com caráter hidrofóbico bastante pronunciado reduzindo a densidade da água em menos que a metade. Assim, verificamos que apesar de a geometria HF ser mais

estável isoladamente, quando solvatada em água, é a geometria AMBER que apresenta maior interação hidrofílica na parte mais externa.

Agora, para julgar a respeito da estabilidade relativa das duas geometrias em solução devemos avaliar também a contribuição entrópica, ou seja, devemos olhar para a diferença de energia livre entre essas duas geometrias em água.

Capítulo 4

Estabilidade Conformacional

Neste capítulo versaremos sobre o método de Perturbação Termodinâmica para cálculos de variação de energia livre e alguns detalhes operacionais intrínsecos do método. Apresentaremos os resultados e ao final faremos uma análise da estabilidade relativa das duas geometrias (AMBER e HF) quando solvatadas em água.

4.1. Introdução

Uma das quantidades termodinâmicas que se pode medir ou calcular das quais se tem grande interesse em química é a diferença de energia livre que um sistema sofre ao transitar entre dois estados. Para um sistema em contato térmico com um reservatório (e, portanto, com temperatura fixa) essa quantidade é relacionada à energia livre de Helmholtz (F) ou potencial canônico. Para sistemas cuja temperatura e pressão são constantes por estarem em contato com um reservatório térmico-barométrico, a energia em questão é a energia livre de Gibbs (G).

A energia livre de Gibbs está bastante envolvida em processos biológicos. Isso porque os meios biológicos são constituídos de tal forma que sofrem pequeníssimas variações de temperatura e pressão. As células

do nosso corpo, por exemplo, sempre atuam de modo a manter a temperatura no pequeno intervalo em que os ácidos nucléicos e as proteínas desempenham suas funções corretamente, e a pressão de modo a compensar a pressão externa. Quando em equilíbrio, nosso corpo - como a grande maioria dos sistemas biológicos - tem, portanto, temperatura e pressão constantes. Inúmeras propriedades de interesse físico-químico e farmacológico estão relacionadas à energia livre de Gibbs [*King*, 1990].

Aqui, entretanto, trataremos apenas da energia livre de Helmholtz. A extensão para a energia livre de Gibbs é imediata, trocando-se apenas o ensemble canônico pelo ensemble NPT.

A transição entre dois estados pode envolver inúmeros processos, inclusive mudanças na molécula de soluto ou até mesmo de solvente. Em casos mais simples, como o nosso, pode envolver apenas uma mudança estrutural e/ou eletrônica na molécula de soluto.

Aqui utilizaremos o cálculo da variação de energia livre de Helmholtz para analisar a estabilidade das geometrias AMBER e HF em solução aquosa.

4.2. Perturbação Termodinâmica

A energia livre pode ser dada pela equação

$$F(N,V,T) = -\frac{1}{\beta}\ln Z(N,V,T). \tag{4.1}$$

Quando um sistema sofre uma transição qualquer, sua função de partição sofre uma mudança resultando numa variação de energia livre dada por

$$\Delta F = -\frac{1}{\beta}\ln\left(\frac{Z_f}{Z_i}\right) \tag{4.2}$$

onde os índices f e i se referem aos estados final e inicial da transição. A função de partição canônica, lembremos, é dada por

$$Z(N,V,T) = \int e^{-\beta \mathrm{H}(\boldsymbol{R},\boldsymbol{P})} d\boldsymbol{R}d\boldsymbol{P} \tag{4.3}$$

sendo a hamiltoniana dada pela equação

$$\mathrm{H} = \mathrm{H}(\boldsymbol{R},\boldsymbol{P}) = \sum_{i=1}^{N}\left(\frac{\vec{p}_i^{\,2}}{2m_i} + \frac{\vec{l}_i^{\,2}}{2I_i}\right) + U(\boldsymbol{R}) \tag{4.4}$$

Agora, Z_f e Z_i são dados por

$$\begin{aligned} Z_f &= \int \exp\left(-\beta \sum_{j=1}^{N} \frac{\vec{p}_j^{\,2}}{2m_j} + \frac{\vec{l}_j^{\,2}}{2I_j} + U_f(\boldsymbol{R})\right) d\boldsymbol{R}d\boldsymbol{P} \\ Z_i &= \int \exp\left(-\beta \sum_{k=1}^{N} \frac{\vec{p}_k^{\,2}}{2m_k} + \frac{\vec{l}_k^{\,2}}{2I_k} + U_i(\boldsymbol{R})\right) d\boldsymbol{R}d\boldsymbol{P}. \end{aligned} \tag{4.5}$$

$U_f(\boldsymbol{R})$ é a função energia potencial para o estado final e $U_i(\boldsymbol{R})$ é a função energia potencial para o estado inicial. A razão entre Z_f e Z_i pode ser escrita da forma

$$\frac{Z_f}{Z_i} = \frac{\int \exp\left(-\beta \sum_{j=1}^{N} \frac{\vec{p}_j^{\,2}}{2m_j} + \frac{\vec{l}_j^{\,2}}{2I_j}\right) d\boldsymbol{P} \int e^{-\beta U_f(\boldsymbol{R})} d\boldsymbol{R}}{\int \exp\left(-\beta \sum_{k=1}^{N} \frac{\vec{p}_k^{\,2}}{2m_k} + \frac{\vec{l}_k^{\,2}}{2I_k}\right) d\boldsymbol{P} \int e^{-\beta U_i(\boldsymbol{R})} d\boldsymbol{R}}. \tag{4.6}$$

A razão entre as integrais em $\boldsymbol{P}$ pode ser facilmente reduzida a termos que dependem apenas das massas e dos momentos de inércia das moléculas em cada estado, de modo que temos

$$\frac{Z_f}{Z_i} = \frac{\prod_j (2\pi m_j)^{3N_j/2} \prod_j (2\pi I_j)^{3N_j/2} \int e^{-\beta U_f(\boldsymbol{R})} d\boldsymbol{R}}{\prod_k (2\pi m_k)^{3N_k/2} \prod_k (2\pi I_k)^{3N_k/2} \int e^{-\beta U_i(\boldsymbol{R})} d\boldsymbol{R}} = R_m R_I \frac{\int e^{-\beta U_f(\boldsymbol{R})} d\boldsymbol{R}}{\int e^{-\beta U_i(\boldsymbol{R})} d\boldsymbol{R}} \tag{4.7}$$

onde N_j é o número de moléculas do tipo j. As razões R_m e R_I entre os produtórios são constantes dependendo parametricamente das massas e dos momentos de inércia das moléculas, respectivamente, nos dois estados. Se tivermos a mesma quantidade de cada tipo de molécula nos dois estados então $R_m = 1$. R_I é um pouco mais complicado. No entanto, para sistemas em que apenas uma ou pouquíssimas moléculas sofrem mudança geométrica, $R_I \approx 1$.

Mais uma vez o espaço de configurações pode ser restrito ao espaço de coordenadas $\boldsymbol{R}$. Tomando agora $U_f = U_i + \Delta U$, temos que a variação de energia livre será

$$\Delta F = -\frac{1}{\beta}\ln\left(R_m R_I \frac{\int e^{-\beta\Delta U} e^{-\beta U_i} d\boldsymbol{R}}{\int e^{-\beta U_i} d\boldsymbol{R}}\right). \tag{4.8}$$

O logaritmando pode ser reconhecido facilmente como as constantes R_m e R_I multiplicadas pela média da exponencial de $-\beta\Delta U$. Assim temos finalmente:

$$\Delta F = -\frac{1}{\beta}\left[\ln R_m + \ln R_I + \ln\left\langle e^{-\beta\Delta U}\right\rangle_i\right] \tag{4.9}$$

onde o índice i lembra que a média deve ser feita no ensemble do estado inicial. Os termos $-(1/\beta)\ln R_m$ e $-(1/\beta)\ln R_I$ são responsáveis pela variação de energia livre devida à mudança na massa e no momento de inércia do sistema. Se uma parcela muito pequena das moléculas do sistema sofre mudança de massa ou momento de inércia, esses termos serão desprezíveis e teremos então

$$\Delta F = -\frac{1}{\beta}\ln\left\langle e^{-\beta\Delta U}\right\rangle_i \tag{4.10}$$

As funções $U_f(\boldsymbol{R})$ e $U_i(\boldsymbol{R})$ compreendem a energia potencial intermolecular e a energia interna das moléculas, de modo que poderíamos escrever

$$U_f(\boldsymbol{r}) = U_f^{inter}(\boldsymbol{R}) + \sum_{j=1}^{N} \varepsilon_f^j \tag{4.11}$$

onde ε_f^j é a energia interna da molécula j no estado final. Analogamente construímos uma expressão para $U_i(\boldsymbol{R})$ e temos então que a diferença de energia potencial será

$$\Delta U = \Delta U^{inter} + \sum_{j=1}^{N} \Delta\varepsilon^j \tag{4.12}$$

onde as variações têm significados imediatos. As diferenças de energia interna $\Delta\varepsilon^j$ são fixas e portanto contribuem multiplicativamente para a média na equação (4.10), de modo que temos enfim

$$\Delta F = -\frac{1}{\beta}\ln\left\langle e^{-\beta\Delta U}\right\rangle_i + \sum_{j=1}^{N}\Delta\varepsilon^j \tag{4.13}$$

com ΔU representando apenas a energia potencial intermolecular de agora em diante.

O primeiro termo do lado direito da equação (4.13) é chamado energia livre intermolecular (ΔF^{inter}) e pode ser calculado através de métodos computacionais tradicionais para cálculos de médias como o método de Monte Carlo Metropolis. O segundo termo pode ser calculado com bastante acurácia através de métodos quânticos já bem estabelecidos. Esse método é conhecido como Teoria de Perturbação Termodinâmica [*Zwanzig*, 1954].

4.2.1. Realização prática

O cálculo de ΔF^{inter} através da técnica de Monte Carlo Metropolis envolve algumas dificuldades numéricas. Se a transição de um estado para outro resulta em uma variação de energia livre intermolecular muito grande, então, com toda probabilidade, a diferença de energia potencial intermolecular entre um estado e outro, para uma configuração qualquer do ensemble do estado inicial, será grande e a contribuição de $e^{-\beta\Delta U}$ para a média será pequena. Assim serão necessárias muitas configurações para se obter convergência na média da equação (4.13) e portanto um número maior de passos será necessário.

A maneira comumente utilizada para contornar esse problema é a técnica de "janelas termodinâmicas" [*Sheykhet*, 1990; *Jorgensen*, 1985]. Essa técnica consiste em dividir o caminho que una os dois estados em vários pedaços (ou janelas). Esse caminho pode ser físico, ou seja, pode constituir-se de estados intermediários em que o sistema pode existir, caso em que podemos obter, uma curva de energia livre, ou curva de reação. Entretanto, não é necessário que o caminho seja físico para que possamos

obter uma diferença total de energia livre entre os estados final e inicial. Isso porque a energia livre é uma função de estado e portanto não depende do caminho que liga os dois estados em questão.

Quando se está interessado apenas na variação de energia livre entre dois estados, é mais fácil utilizar um caminho não físico onde os estados intermediários são misturas dos estados inicial e final, constituindo assim uma trajetória linear de variação dos parâmetros que levam um estado em outro. Dessa forma podemos escrever

$$\Delta F = -\frac{1}{\beta}\sum_m \ln\left\langle e^{-\beta\Delta U}\right\rangle_m + \Delta\varepsilon \tag{4.14}$$

onde $\Delta U = U_{m+1}^{inter} - U_m^{inter}$ é a diferença de energia potencial intermolecular entre dois estados quaisquer e $\Delta\varepsilon$ é a variação total de energia interna das moléculas entre os estados inicial e final. Agora m passa a ser o estado de referência que irá transitar entre os diversos estados intermediários.

Os estados intermediários são construídos misturando-se os estados 0 e 1 (inicial e final) da forma

$$\begin{aligned} \boldsymbol{s}_\lambda &= \boldsymbol{s}_0 + \lambda(\boldsymbol{s}_1 - \boldsymbol{s}_0) \\ q_\lambda &= q_0 + \lambda(q_1 - q_0) \\ \sigma_\lambda &= \sigma_0 + \lambda(\sigma_1 - \sigma_0) \\ \varepsilon_\lambda &= \varepsilon_0 + \lambda(\varepsilon_1 - \varepsilon_0) \end{aligned} \tag{4.15}$$

onde $\boldsymbol{s}$ é o conjunto das coordenadas internas das moléculas do sistema e q, σ e ε são os parâmetros do potencial de interação intermolecular. λ varia entre 0 e 1 em intervalos $\Delta\lambda = 1/n$ com $n-1$ sendo o número de estados entre o inicial e o final. λ é dado então por $\lambda = m\Delta\lambda$. Quando $m = 0$ ($\lambda = 0$) temos os parâmetros do estado inicial. Quando $m = n$ ($\lambda = 1$) temos os parâmetros do estado final. E para m qualquer temos o m-ésimo estado intermediário.

O valor de $\Delta\lambda$ está intimamente ligado à velocidade da convergência do valor de ΔF^{inter} em cada etapa intermediária, pois a variação dos parâmetros é que dita a variação na energia intermolecular de um estado para outro. Em geral escolhe-se $\Delta\lambda$ de modo que a variação de energia

livre intermolecular seja em torno de 4 $k_B T$ ou menos, ou seja, $\Delta F^{inter}_{m\to m+1} \leq 4$ $k_B T$ [*Sheykhet*,1990].

Uma outra técnica comumente usada em perturbação termodinâmica e que facilita o trabalho computacional em se calcular a média da exponencial de $-\beta\Delta U$ é a técnica de *double-wide* [*Jorgensen*, 1989]. Essa técnica consiste em calcular duas médias com cada estado de referência: a média referente à passagem do estado de referência para o estado anterior; e do estado de referência para o posterior. Assim, simulando o sistema no estado m, podemos calcular $\Delta F^{inter}_{m\to m-1}$ e $\Delta F^{inter}_{m\to m+1}$. Essa técnica reduz o número de simulações necessárias pela metade o que a torna bastante atraente. O esquema dessa técnica é ilustrado na Figura 4.1. Com o uso do *double-wide*, o cálculo de ΔF^{inter} se torna mais eficiente se usarmos um número ímpar de estados intermediários. Assim, se um número par de estados intermediários for suficiente dentro de uma variação máxima desejada dos parâmetros, é vantajoso incluir-se um estado intermediário a mais, já que o custo computacional será praticamente o mesmo e a mudança de um estado para outro será mais suave tornando as simulações de mais fácil convergência.

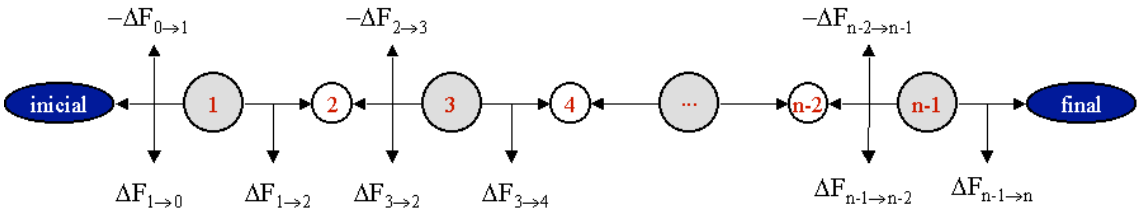


Figura 4.1: Esquema ilustrativo da técnica *double-wide* com a teoria de perturbação termodinâmica. Os ensembles dos estados ímpares (representados pelos círculos maiores) são usados para calcular as diferenças de energia livre entre esses e os estados imediatamente anteriores e também os estados imediatamente posteriores.

4.3. Resultados

A diferença entre as geometrias AMBER e HF reside somente nas suas coordenadas e nas cargas parciais dos átomos. Fizemos a variação linear desses parâmetros simultaneamente e construímos um caminho com

19 estados intermediários para a α-CD. Esse número de estados intermediários é decorrente da escolha em se provocar deslocamentos máximos de 1.0 Å nas coordenadas dos átomos nos estados intermediários. A mudança nas cargas parciais é pequena da geometria AMBER para a HF de modo que o fator determinante no número de estados intermediários é dado pelas coordenadas. É interessante salientar também que o maior deslocamento na geometria da α-CD ocorre para as hidroxilas primárias que formam ligações de hidrogênio internas na geometria HF e que estão na conformação (-)*gauche* na geometria AMBER. Essa mudança das duas hidroxilas primárias de (-) *gauche* para *trans* é que torna grande o número necessário de simulações com estados intermediários.

Utilizando o *double-wide*, reduzimos o número de simulações necessárias de 20 para 10. Realizamos as 10 simulações partindo da geometria AMBER (que portanto chamaremos de geometria 0) até a geometria HF (que será a geometria número 20). As simulações dessa etapa foram realizadas com um número ligeiramente maior de passos tanto para a termalização ($30 \cdot 10^6$) quanto para a equilibração ($60 \cdot 10^6$). Um resumo dos detalhes técnicos das simulações de perturbação termodinâmica (PT) é apresentado na Tabela 4.1.

Tabela 4.1: Detalhes computacionais das simulações de perturbação termodinâmica.

Simulações de PT	
α-CD:água	1:1000
Temperatura	298.15 K
Pressão	1 atm
Densidade	1.03 g/cm^3
L (caixa cúbica)	31.28 Å
Passos na termalização	$30 \cdot 10^6$
Passos em equilíbrio	$60 \cdot 10^6$

Na Figura 4.2 e na Figura 4.3 mostramos as distribuições de energia potencial de interação soluto-solvente para as simulações realizadas no primeiro e terceiro estados intermediários.

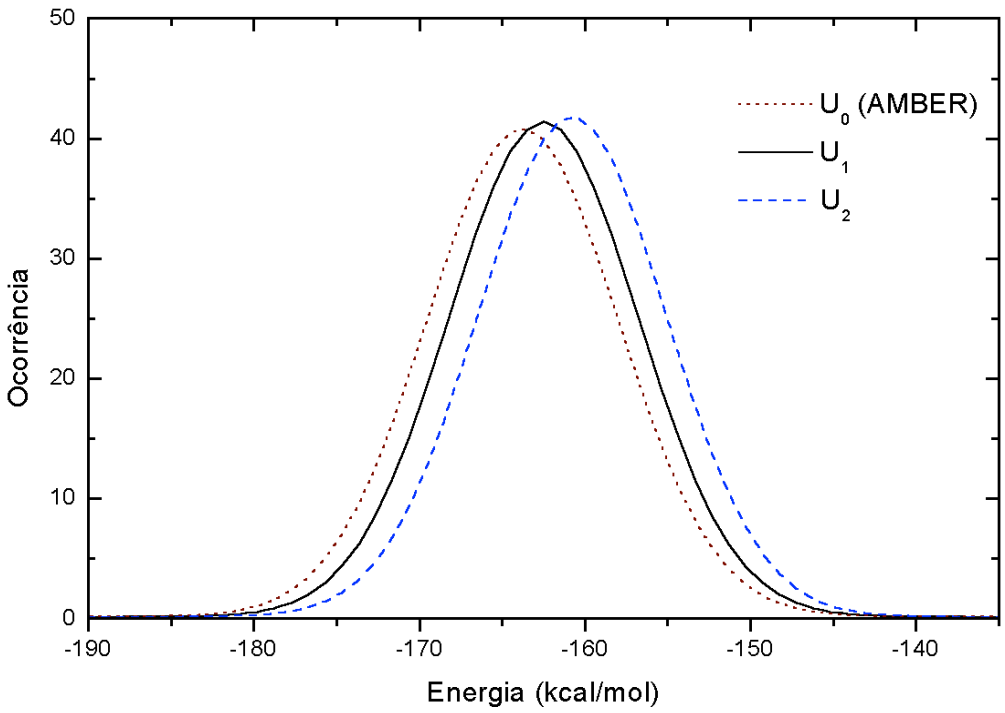


Figura 4.2: Distribuições de energia potencial soluto-solvente para as três primeiras geometrias do caminho linear. As energias de interação foram calculadas no ensemble de configurações geradas com a geometria 1. O *overlap* menor ocorre para a transição 1→2 e é de 97.73%.

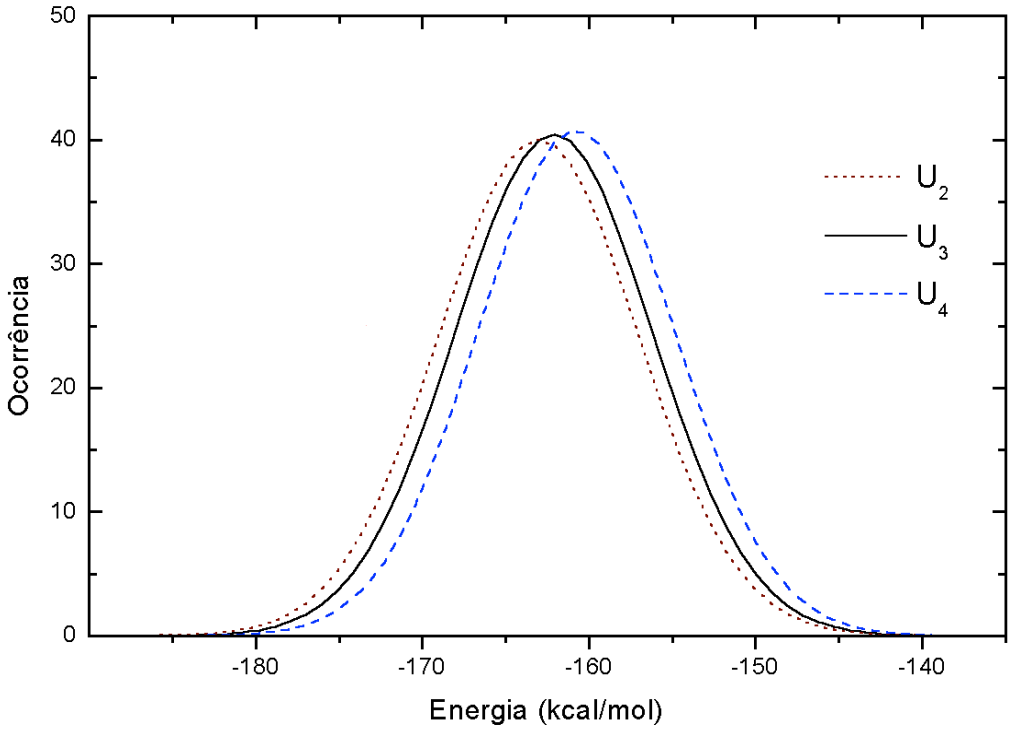


Figura 4.3: Distribuições de energia potencial soluto-solvente para os estados 2, 3 e 4. A simulação foi realizada com o estado intermediário 3. Aqui o *overlap* mínimo ocorre para a transição 3→4 e é de 98.69%.

Nestes gráficos podemos ver que há uma grande sobreposição das distribuições o que deve refletir na convergência das simulações. Podemos perceber também que o menor *overlap*, de 97.73%, ocorre para a transição do estado 1 para o estado 2. Essa é a transição de menor *overlap* dentre todas entre o estado inicial e final como podemos observar também nos gráficos da Figura 4.4 e da Figura 4.5. A sobreposição das distribuições nos estados subseqüentes aumenta gradualmente até atingir overlap praticamente total, indicando menor variação de energia livre e convergência mais facilmente atingida. Na Figura 4.4 mostramos as distribuições referentes à simulação no estado 11, no meio do caminho entre as duas geometrias, e podemos ver que a diferença entre as mesmas é desprezível. Na Figura 4.5 temos as distribuições referentes às últimas transições e podemos ver que a diferença entre as mesmas também é desprezível. Os diferentes valores para as sobreposições nos diferentes estágios estão de acordo com as diferenças de energia livre intermolecular referentes às respectivas transições como mostraremos adiante.

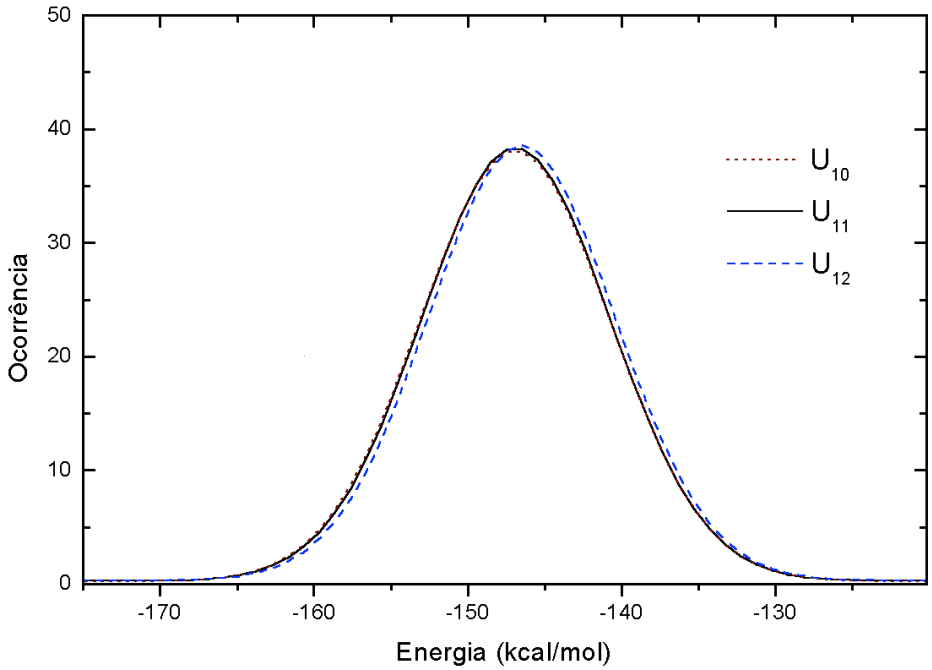


Figura 4.4: Distribuições de energia potencial soluto-solvente. Estamos no meio do caminho entre as duas geometrias e a diferença entre as distribuições é desprezível.

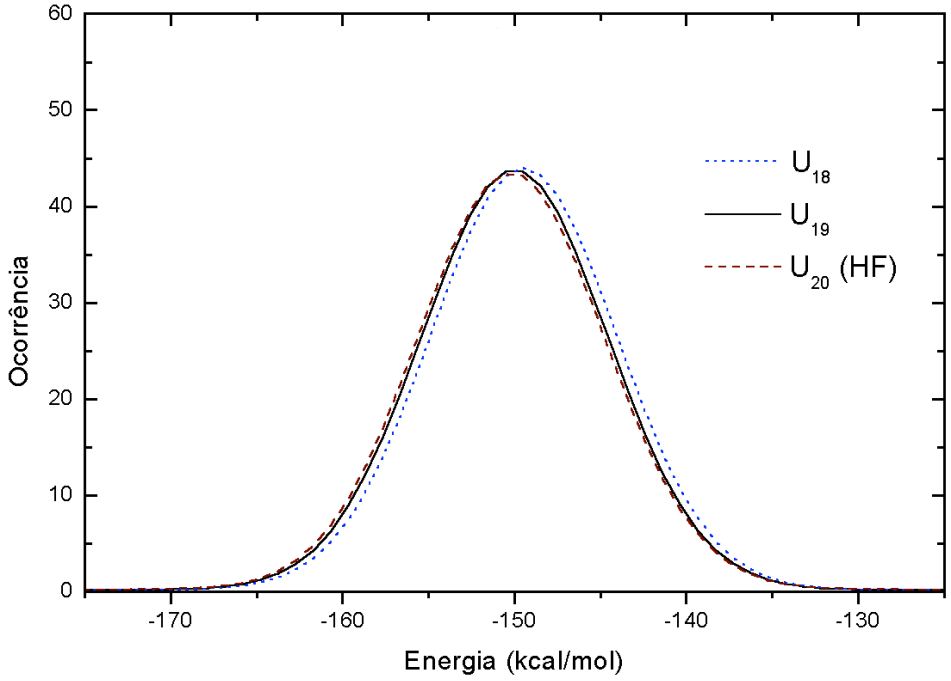


Figura 4.5: Distribuições de energia potencial soluto-solvente. Este é o último estado intermediário e a diferença entre as distribuições também é desprezível.

Na Tabela 4.2 mostramos as diferenças de energia livre calculadas em cada estágio do caminho entre a geometria AMBER e HF. Podemos ver que os valores começam da ordem de 5 $k_B T$ e diminuem gradualmente até ficarem dentro de 3 $k_B T$ já para as simulações no sétimo estado intermédiario. As primeiras simulações devem, portanto, ser mais sucetíveis a problemas de convergência. No entanto, como mostraremos, $60 \cdot 10^6$ passos de Monte Carlo parecem ter sido suficientes para alcançar convergência.

Tabela 4.2: Diferença de energia livre de Helmholtz em cada transição intermediária. *i* representa os estados nos quais se realizaram as respectivas simulações. As energias são dadas em kcal/mol.

i	$-\Delta F_{i\to i\text{-}1}$	$\Delta F_{i\to i+1}$
1	3.14	2.80
3	2.44	2.14
5	1.97	1.89
7	1.83	1.60
9	1.10	0.94
11	0.56	0.40
13	0.75	0.58
15	0.06	-0.11
17	-0.34	-0.55
19	-0.64	-0.89
TOTAL	10.87	8.80

ΔF^{inter}(AMBER→HF) = 19.67 kcal/mol

Na Figura 4.6 mostramos a evolução da média da variação de energia livre entre os estados 0 e 1 e entre os estados 1 e 2. Estas transições são as de maior variação de energia como mostra a Tabela 4.2 e portanto devem apresentar uma lentidão maior na convergência do que as outras transições. Entretanto, podemos ver na evolução de $-\Delta F_{1\to 0}$ e $\Delta F_{1\to 2}$ que a partir de $30\cdot 10^6$ passos de MC já se observa uma estabilidade no valor destas grandezas. Podemos ver que a oscilação nestes valores a partir daí fica dentro de 0.1 kcal/mol o que representa um erro em torno de 3%.

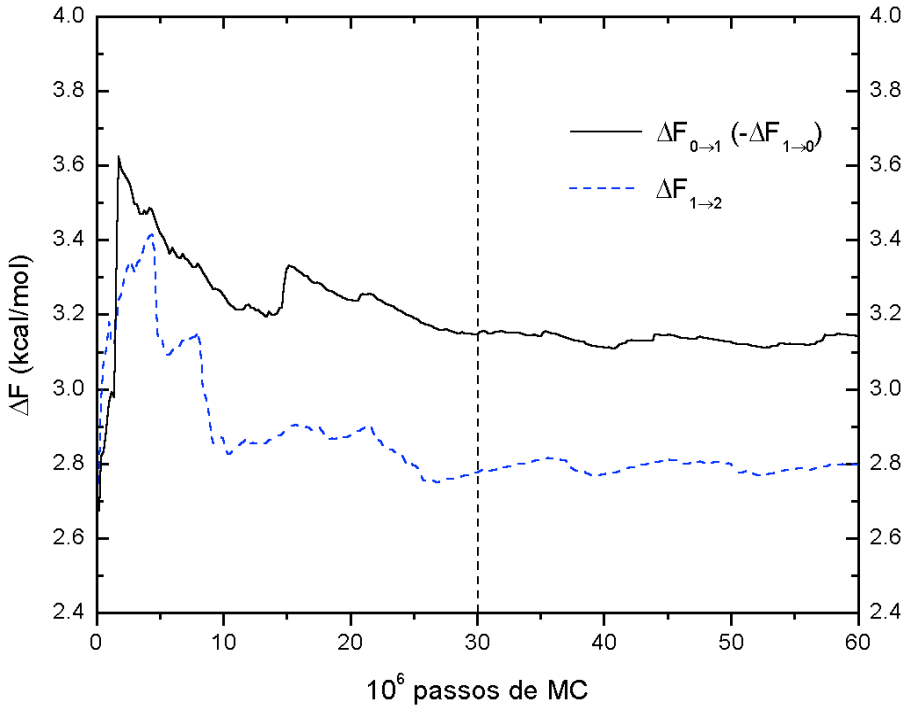


Figura 4.6: Evolução de ⟨ΔF⟩. A partir de $30 \cdot 10^6$ passos de simulação os valores de ΔF já estão oscilando com amplitude menor que 0.1 kcal/mol.

Adotando esse valor como erro para todas as etapas intermediárias, temos que o erro total em ΔF^{inter} é de 2 kcal/mol. Dessa forma, temos que ΔF^{inter}(AMBER→HF) = 20 ± 2 kcal/mol

Calculamos também a variação de energia livre na transição do estado 0 para o 1 através do ensemble gerado no estado 0, para analisar a histerese entre os caminhos 0→1 e 1→0. Obtivemos para $\Delta F_{0\to1}$ 3.28 kcal/mol o que significa um acréscimo de 0.14 kcal/mol em relação a $-\Delta F_{1\to0}$ representando uma diferença de aproximadamente 4%. Na Figura 4.7 ilustramos a evolução da variação de energia livre nas duas simulações.

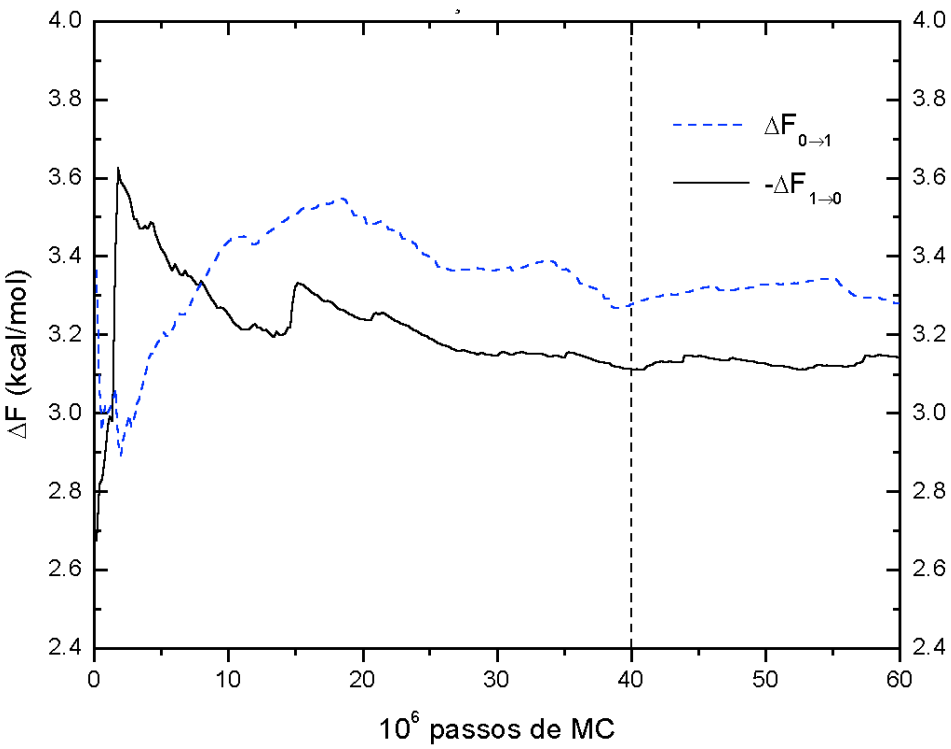


Figura 4.7: Evolução de ⟨ΔF⟩ para os caminhos 1→0 e 0→1. A partir de $40 \cdot 10^6$ passos de simulação os valores de ΔF já estão flutuando com amplitude de cerca de 0.1 kcal/mol. A diferença entre os valores de ΔF ao final de $60 \cdot 10^6$ passos é de 0.14 kcal/mol, dentro portanto da flutuação térmica.

Como vimos na seção 3.3, a diferença de energia potencial intermolecular entre as duas geometrias e a água é de cerca de 21 kcal/mol. Como a variação de energia livre intermolecular em solução aquosa é de 20 kcal/mol, e sabendo que $\Delta F = \Delta U + T\Delta S$, podemos ver que a contribuição entrópica ($T\Delta S$) para a estabilidade relativa entre as duas geometrias é pequena.

Somando-se este valor de ΔF^{inter} à diferença de energia interna entre as duas geometrias $\Delta\varepsilon$ = 2.25 kcal/mol, temos que a variação de energia livre total entre os dos estados é de 17 ± 2 kcal/mol. Portanto a geometria AMBER é de fato mais estável em solução aquosa.

Capítulo 5

Conclusão e Perspectivas

Neste capítulo resumiremos os resultados obtidos, as conclusões acerca desses resultados e as perspectivas também em torno do estudo das ciclodextrinas.

Nosso principal interesse é analisar a atuação dos métodos clássicos, em especial os métodos de mecânica molecular, na descrição de macromoléculas comparando com os métodos quânticos. Como a utilização de métodos de química quântica para esses sistemas ainda é bastante restrita por limitações computacionais, é interessante verificar a eficácia desses métodos clássicos na otimização de geometrias. Em especial nosso interesse é verificar essa eficácia em obter geometrias para estudo em meio solvente dentro do modelo de moléculas rígidas.

Inicialmente pudemos observar que os métodos clássicos AMBER e OPLS fornecem uma boa descrição da α-CD sendo capazes de manter a estrutura macrocíclica da molécula formada através da rede de ligações de hidrogênio secundárias. Esses métodos também descreveram corretamente os posicionamentos das hidroxilas primárias mostrando uma favorabilidade das conformações *gauche* em detrimento das *trans* o que está em acordo com os resultados experimentais tanto em cristais quanto em solução. Dentre esses métodos, notamos uma eficácia maior do método AMBER na obtenção das geometrias isoladas.

Os métodos quânticos mostraram uma favorabilidade em manter algumas hidroxilas primárias em conformação *trans* nas otimizações com a α-CD isolada. Isso por causa da formação de ligações de hidrogênio entre essas hidroxilas primárias e o oxigênio piranosídico da unidade de glicose consecutiva.

Comparando a geometria mais estável dos métodos clássicos (AMBER1) e dos métodos quânticos (HF/6-21G), pudemos observar que as diferenças se concentram principalmente na formação de ligações de hidrogênio secundárias mais fortes, na formação de duas ligações de hidrogênio com as hidroxilas primárias e no tamanho da cavidade. Como conseqüência da formação de ligações mais fortes entre as hidroxilas secundárias, a geometria clássica tem uma forma cônica mais aberta do que a geometria quântica.

A temática principal na interação da α-CD com a água é justamente a disputa entre a formação de ligações de hidrogênio internas na molécula e ligações intermoleculares com a água. No caso das hidroxilas secundárias, os dados experimentais indicam que a formação da rede de ligações é mais interessante do que o seu rompimento para aumentar o número de ligações intermoleculares envolvendo essas hidroxilas. O rompimento dessa rede, na verdade, se dá apenas em condições específicas. Já as hidroxilas primárias, sabe-se que são mais livres e portanto mais susceptíveis a essa disputa.

Fizemos então duas simulações de Monte Carlo para analisar a diferença de interação entre a geometria HF e a geometria AMBER. Através da simulação dessas duas geometrias em água pudemos ver que a geometria AMBER interage mais com a água do que a geometria HF com uma diferença de 21 kcal/mol na energia de interação soluto-solvente. Isto por causa da formação de 6 ligações de hidrogênio intermoleculares a mais entre a geometria AMBER e a água.

Analisamos a cavidade da α-CD nas duas geometrias e observamos que na geometria AMBER a densidade da água dentro da cavidade é de 0.46 g/cm^3 enquanto que na geometria HF a densidade da água dentro da cavidade é de 0.33 g/cm^3. Esses resultados indicam um caráter bastante

hidrofóbico na cavidade da α-CD sendo ligeiramente maior na geometria HF.

Para analisar a estabilidade das duas moléculas em solução entretanto, fizemos um cálculo de perturbação termodinâmica saindo da geometria AMBER até chegar na geometria HF. Utilizamos 19 janelas termodinâmicas para diminuir os erros no cálculo da energia livre de Helmholtz intermolecular. Obtivemos então que a diferença de energia livre entre as duas geometrias em água é de 19.7 kcal/mol em favor da geometria AMBER.

Isso reforça os dados experimentais que mostram que a formação de ligações de hidrogênio internas é desfavorável. Em solução aquosa pudemos ter uma boa idéia dessa desfavorabilidade através do cálculo da energia livre de Helmholtz.

Podemos resumir nossos principais resultados na forma da Tabela 5.1. Nessa tabela relacionamos, na ordem: o ângulo de cone τ; o ângulo formado pelos oxigênios intersacarídeos δ; a distância entre os oxigênios intersacarídeos O4-O4'; a distância entre o oxigênio O3 e o oxigênio O2' adjacente denotando a possível formação de ligação de hidrogênio entre hidroxilas secundárias; o número de ligações entre as hidroxilas secundarias segundo critério forte/fraco; conformação das hidroxilas primárias; energia relativa entre as geometrias isoladas (ΔE) em nível HF/6-311++G**; dipolo também em nível HF/6-311++G**; número de moléculas de água na primeira camada de solvatação dos oxigênios da α-CD ($N_{água}$); número de ligações de hidrogênio formadas com a água ($HB_{água}$); densidade da água dentro da cavidade; energia relativa de interação com a água; e finalmente diferença de energia livre de Helmholtz em solução aquosa.

Tabela 5.1: Resumo dos principais resultados obtidos no presente trabalho.

Grandeza	AMBER	HF	Experim.
τ (º)	76±7	83±5	80±9
δ (º)	120±1	120±2	120±5
$R_{O4\text{-}O4'}$ (Å)	4.30±0.03	4.28±0.04	4.24±0.09
$R_{O3\text{-}O2'}$ (Å)	3.22±0.03	2.86±0.02	3.0±0.4
HB sec	4/6	6/6	5.5/5.5
OH prim	100% G(-)	30% T, 70% G(-)	25% G(+), 75%G(-)
ΔE (kcal/mol)	0.00	-2.25	-
μ (D)	11.01	12.71	-
$N_{água}$	55	44	-
$HB_{água}$	31	25	-
$\rho_{cavidade}$	0.46	0.33	-
ΔU^{inter} (kcal/mol)	0	21	-
ΔF^{inter} (kcal/mol)	0	20	-

Concluímos que os métodos clássicos AMBER e OPLS são adequados para a obtenção de geometrias da α-CD e certamente de outras dextrinas e até mesmo outros oligo e polissacarídeos. Na verdade, para a obtenção de geometrias para simulação em meio solvente dentro do modelo de moléculas rígidas, podemos dizer até que esses métodos são mais adequados do que os métodos quânticos. Isso se deve ao fato de esses métodos serem parametrizados para fase líquida e portanto devem possuir intrinsecamente em seus parâmetros a interação com o solvente. Em especial o método AMBER se mostrou excelente para a descrição desse sistema e isso com certeza é conseqüência da sua parametrização especial envolvendo polissacarídeos.

Considerando que o dipolo da geometria HF é maior que o da geometria AMBER, nossos resultados indicam que campo de reação autoconsistente pode não ser um método adequado para tentar obter a

geometria de ciclodextrinas e até polissacarídeos em solução. Como vimos, apesar de a geometria HF ter dipolo maior, ela interage menos fortemente com a água do que a geometria AMBER.

Estamos interessados agora em estudar processos de inclusão em ciclodextrinas, utilizando os conhecimentos obtidos com esse tipo de molécula. Em especial a inclusão de compostos de interesse biológico é uma grande motivação para tal estudo em virtude até mesmo do sucesso que as ciclodextrinas e seus derivados têm obtido nesse campo.

Apêndice

Neste apêndice mostramos os parâmetros usados nas simulações de Monte Carlo e também usados como estado inicial e final da perturbação termodinâmica. Abaixo estão os parâmetros da geometria AMBER:

Tipo	N° At	X	Y	Z	Carga	Epslon	Sigma
1	6	-3.91667	3.61666	0.20877	0.561420	0.1094	3.4000
1	6	-3.55369	4.62455	-0.89374	0.258430	0.1094	3.4000
1	6	-2.05619	4.52833	-1.23851	0.278031	0.1094	3.4000
1	6	-1.18823	4.66322	0.03085	0.751260	0.1094	3.4000
1	6	-1.71172	3.73613	1.15608	-0.169962	0.1094	3.4000
1	6	-1.02929	3.93916	2.51917	0.861849	0.1094	3.4000
2	8	-4.35662	4.38677	-2.03350	-0.808138	0.2104	3.0665
2	8	-1.71952	5.56898	-2.13574	-0.865029	0.2104	3.0665
3	8	0.13347	4.27303	-0.34037	-0.789072	0.1700	3.0000
4	8	-3.11704	3.90474	1.35013	-0.671392	0.1700	3.0000
5	8	-1.04368	5.29007	2.93039	-0.868477	0.2104	3.0665
6	1	-5.27856	4.47814	-1.78193	0.486108	0.0000	0.0000
6	1	-2.25893	5.47743	-2.92487	0.533582	0.0000	0.0000
7	1	-1.95700	5.57780	2.99868	0.466313	0.0000	0.0000
1	6	1.20375	5.17198	-0.03408	0.569972	0.1094	3.4000
1	6	2.18577	5.26840	-1.21322	0.339064	0.1094	3.4000
1	6	2.84645	3.90391	-1.47787	0.288579	0.1094	3.4000
1	6	3.48763	3.34587	-0.19006	0.654390	0.1094	3.4000
1	6	2.49708	3.42896	0.99818	0.020502	0.1094	3.4000
1	6	3.11304	3.07436	2.36217	0.831034	0.1094	3.4000
2	8	1.50254	5.73597	-2.36010	-0.839395	0.2104	3.0665
2	8	3.85205	4.05774	-2.46096	-0.860824	0.2104	3.0665
3	8	3.80211	1.98001	-0.45746	-0.766010	0.1700	3.0000
4	8	1.92430	4.73348	1.11207	-0.731197	0.1700	3.0000
5	8	4.28489	3.81535	2.63218	-0.862978	0.2104	3.0665
6	1	0.81608	5.10338	-2.58448	0.498611	0.0000	0.0000
6	1	3.44560	4.39597	-3.26217	0.518632	0.0000	0.0000
7	1	4.06208	4.74912	2.62933	0.458924	0.0000	0.0000
1	6	5.14291	1.54447	-0.21428	0.567269	0.1094	3.4000
1	6	5.63038	0.63405	-1.35306	0.320194	0.1094	3.4000
1	6	4.76682	-0.63704	-1.42904	0.309042	0.1094	3.4000
1	6	4.71271	-1.34862	-0.06075	0.523802	0.1094	3.4000
1	6	4.39147	-0.34209	1.07378	0.260157	0.1094	3.4000
1	6	4.51790	-0.91962	2.49403	0.462955	0.1094	3.4000
2	8	5.59835	1.34609	-2.57474	-0.836996	0.2104	3.0665
2	8	5.32171	-1.52130	-2.38419	-0.845808	0.2104	3.0665
3	8	3.66385	-2.30814	-0.16431	-0.729604	0.1700	3.0000

4	8	5.22791	0.81528	1.00488	-0.743109	0.1700	3.0000
5	8	5.77208	-1.53061	2.71381	-0.830287	0.2104	3.0665
6	1	4.69086	1.60333	-2.75384	0.499245	0.0000	0.0000
6	1	5.34497	-1.07577	-3.23416	0.511905	0.0000	0.0000
7	1	6.45837	-0.87100	2.58942	0.470431	0.0000	0.0000
1	6	3.87890	-3.60518	0.39718	0.567988	0.1094	3.4000
1	6	3.58597	-4.70201	-0.63891	0.395625	0.1094	3.4000
1	6	2.11606	-4.62726	-1.09106	0.501963	0.1094	3.4000
1	6	1.16066	-4.64859	0.12206	0.352794	0.1094	3.4000
1	6	1.62445	-3.67065	1.23024	-0.045590	0.1094	3.4000
1	6	0.86720	-3.82901	2.55999	1.066980	0.1094	3.4000
2	8	4.46857	-4.56439	-1.73530	-0.836612	0.2104	3.0665
2	8	1.82938	-5.73686	-1.92104	-0.862292	0.2104	3.0665
3	8	-0.11740	-4.22917	-0.35250	-0.711496	0.1700	3.0000
4	8	3.01752	-3.81456	1.51023	-0.692773	0.1700	3.0000
5	8	0.92429	-5.15036	3.05649	-0.960427	0.2104	3.0665
6	1	4.32648	-3.70297	-2.13464	0.486997	0.0000	0.0000
6	1	0.91623	-5.67316	-2.21058	0.485678	0.0000	0.0000
7	1	0.52582	-5.73675	2.40908	0.532220	0.0000	0.0000
1	6	-1.17641	-5.18805	-0.31064	0.521751	0.1094	3.4000
1	6	-2.05853	-5.08145	-1.56499	0.230181	0.1094	3.4000
1	6	-2.73733	-3.70246	-1.62735	0.434675	0.1094	3.4000
1	6	-3.50036	-3.40321	-0.31960	0.445570	0.1094	3.4000
1	6	-2.61853	-3.70525	0.91829	0.279209	0.1094	3.4000
1	6	-3.36989	-3.64295	2.25896	0.530081	0.1094	3.4000
2	8	-1.27042	-5.31189	-2.71699	-0.793881	0.2104	3.0665
2	8	-3.65193	-3.68299	-2.70665	-0.842256	0.2104	3.0665
3	8	-3.80336	-2.01002	-0.34893	-0.742637	0.1700	3.0000
4	8	-1.99318	-4.98861	0.83725	-0.730159	0.1700	3.0000
5	8	-4.50903	-4.47808	2.26898	-0.882055	0.2104	3.0665
6	1	-1.83560	-5.25249	-3.49074	0.505175	0.0000	0.0000
6	1	-4.32942	-4.34331	-2.54442	0.471164	0.0000	0.0000
7	1	-5.10635	-4.18708	1.57594	0.500129	0.0000	0.0000
1	6	-5.11310	-1.59383	0.04759	0.470590	0.1094	3.4000
1	6	-5.76268	-0.73422	-1.04871	0.421496	0.1094	3.4000
1	6	-4.93207	0.54037	-1.28741	0.365933	0.1094	3.4000
1	6	-4.70600	1.30583	0.03405	0.542829	0.1094	3.4000
1	6	-4.22810	0.34992	1.15731	0.087489	0.1094	3.4000
1	6	-4.19357	0.97809	2.56116	0.726657	0.1094	3.4000
2	8	-5.88387	-1.49465	-2.23443	-0.830896	0.2104	3.0665
2	8	-5.62038	1.38407	-2.19115	-0.845847	0.2104	3.0665
3	8	-3.69861	2.27702	-0.24273	-0.735220	0.1700	3.0000
4	8	-5.05332	-0.81328	1.23631	-0.698714	0.1700	3.0000
5	8	-5.41357	1.60465	2.90143	-0.855414	0.2104	3.0665
6	1	-5.00427	-1.74794	-2.52470	0.498466	0.0000	0.0000
6	1	-5.08534	2.16631	-2.34450	0.485640	0.0000	0.0000
7	1	-6.11246	0.94703	2.87294	0.464133	0.0000	0.0000
8	1	-4.96176	3.78193	0.47491	0.054623	0.0000	0.0000
8	1	-3.77043	5.62953	-0.52723	0.034875	0.0000	0.0000
8	1	-1.85501	3.56752	-1.71625	0.048911	0.0000	0.0000
8	1	-1.24259	5.70070	0.36530	-0.032597	0.0000	0.0000
8	1	-1.51694	2.70221	0.86508	0.075553	0.0000	0.0000
8	1	0.00113	3.59838	2.46369	-0.137873	0.0000	0.0000
8	1	-1.53770	3.33557	3.27141	-0.092600	0.0000	0.0000
8	1	0.83158	6.17873	0.16148	0.067708	0.0000	0.0000
8	1	2.96018	5.99404	-0.95754	0.046957	0.0000	0.0000
8	1	2.09447	3.20241	-1.84438	0.004317	0.0000	0.0000
8	1	4.37142	3.94598	0.03393	-0.033881	0.0000	0.0000
8	1	1.69604	2.70845	0.82326	0.042733	0.0000	0.0000
8	1	3.35289	2.01447	2.38871	-0.163247	0.0000	0.0000
8	1	2.38416	3.26914	3.14943	-0.082621	0.0000	0.0000

8	1	5.83005	2.38992	-0.15655	0.072753	0.0000	0.0000
8	1	6.66461	0.34997	-1.15000	0.048978	0.0000	0.0000
8	1	3.75528	-0.36851	-1.74026	0.005345	0.0000	0.0000
8	1	5.68609	-1.81130	0.11293	-0.019407	0.0000	0.0000
8	1	3.35243	-0.02776	0.96005	-0.033025	0.0000	0.0000
8	1	3.73309	-1.65345	2.66266	0.013489	0.0000	0.0000
8	1	4.38434	-0.11927	3.22217	0.001287	0.0000	0.0000
8	1	4.90744	-3.72765	0.74007	0.059868	0.0000	0.0000
8	1	3.76688	-5.67309	-0.17461	0.027740	0.0000	0.0000
8	1	1.95974	-3.70884	-1.66049	-0.017810	0.0000	0.0000
8	1	1.14482	-5.66000	0.53150	-0.033481	0.0000	0.0000
8	1	1.43945	-2.65109	0.88665	0.047396	0.0000	0.0000
8	1	-0.17545	-3.55368	2.42349	-0.314636	0.0000	0.0000
8	1	1.29757	-3.15437	3.30108	-0.102810	0.0000	0.0000
8	1	-0.78153	-6.20502	-0.28004	0.103209	0.0000	0.0000
8	1	-2.82674	-5.85556	-1.51429	0.015504	0.0000	0.0000
8	1	-1.97866	-2.93453	-1.79137	0.061629	0.0000	0.0000
8	1	-4.39597	-4.02689	-0.30351	-0.071836	0.0000	0.0000
8	1	-1.84088	-2.94194	0.96195	0.025579	0.0000	0.0000
8	1	-3.68017	-2.61926	2.45861	-0.057317	0.0000	0.0000
8	1	-2.70040	-3.95150	3.06307	0.057008	0.0000	0.0000
8	1	-5.76330	-2.44792	0.24275	0.054878	0.0000	0.0000
8	1	-6.76443	-0.45169	-0.71942	0.027628	0.0000	0.0000
8	1	-3.96803	0.26815	-1.72249	0.012559	0.0000	0.0000
8	1	-5.65691	1.75845	0.31943	-0.052570	0.0000	0.0000
8	1	-3.20612	0.04021	0.93151	0.009304	0.0000	0.0000
8	1	-3.39347	1.71209	2.61393	-0.077267	0.0000	0.0000
8	1	-3.97858	0.20253	3.29719	-0.085420	0.0000	0.0000

Abaixo os parâmetros da geometria HF:

Tipo	N° At	X	Y	Z	Carga	Epslon	Sigma
1	6	1.45669	-5.0389	0.08697	0.509870	0.1094	3.4000
1	6	0.64018	-5.46847	-1.12878	0.191960	0.1094	3.4000
1	6	-0.53082	-4.52038	-1.29378	0.369810	0.1094	3.4000
1	6	-1.38612	-4.54687	-0.04531	0.526968	0.1094	3.4000
1	6	-0.56854	-4.21932	1.19868	0.062921	0.1094	3.4000
1	6	-1.3358	-4.56501	2.46552	0.468259	0.1094	3.4000
2	8	1.40548	-5.3633	-2.35353	-0.811977	0.2104	3.0665
2	8	-1.37462	-4.94108	-2.38618	-0.866220	0.2104	3.0665
3	8	-2.38324	-3.52588	-0.21431	-0.665758	0.1700	3.0000
4	8	0.6058	-5.08901	1.23163	-0.650799	0.1700	3.0000
5	8	-1.64105	-5.97386	2.4544	-0.830145	0.2104	3.0665
6	1	1.93584	-6.15403	-2.52539	0.501418	0.0000	0.0000
6	1	-0.85116	-5.00094	-3.19873	0.519107	0.0000	0.0000
7	1	-0.80397	-6.46046	2.44874	0.475320	0.0000	0.0000
1	6	-3.76983	-3.80312	-0.02942	0.385565	0.1094	3.4000
1	6	-4.47604	-3.26832	-1.26778	0.333521	0.1094	3.4000
1	6	-4.23377	-1.77153	-1.35263	0.375470	0.1094	3.4000
1	6	-4.73441	-1.08948	-0.10173	0.418103	0.1094	3.4000
1	6	-4.09381	-1.67961	1.14944	0.060543	0.1094	3.4000
1	6	-4.823	-1.23315	2.407	0.440088	0.1094	3.4000
2	8	-4.03706	-3.89527	-2.47629	-0.882327	0.2104	3.0665
2	8	-4.96468	-1.20929	-2.46323	-0.872144	0.2104	3.0665
3	8	-4.32739	0.2863	-0.19922	-0.610869	0.1700	3.0000
4	8	-4.24949	-3.13381	1.1316	-0.609548	0.1700	3.0000
5	8	-6.19307	-1.67461	2.3346	-0.819392	0.2104	3.0665
6	1	-3.09569	-4.15725	-2.43749	0.541312	0.0000	0.0000
6	1	-4.84797	-1.77689	-3.24055	0.524946	0.0000	0.0000

7	1	-6.18997	-2.63997	2.25746	0.470639	0.0000	0.0000
1	6	-5.2479	1.35595	-0.00596	0.401065	0.1094	3.4000
1	6	-5.09482	2.26884	-1.21344	0.299572	0.1094	3.4000
1	6	-3.66783	2.79049	-1.25728	0.378423	0.1094	3.4000
1	6	-3.33123	3.50713	0.03103	0.399373	0.1094	3.4000
1	6	-3.5585	2.60574	1.24037	0.033418	0.1094	3.4000
1	6	-3.51042	3.38437	2.54474	0.524647	0.1094	3.4000
2	8	-5.40951	1.61745	-2.44807	-0.872171	0.2104	3.0665
2	8	-3.52448	3.74592	-2.32669	-0.837334	0.2104	3.0665
3	8	-1.93297	3.82965	-0.04161	-0.605079	0.1700	3.0000
4	8	-4.9167	2.06495	1.1847	-0.622952	0.1700	3.0000
5	8	-4.51383	4.41906	2.51512	-0.817995	0.2104	3.0665
6	1	-5.15411	0.67268	-2.44926	0.535009	0.0000	0.0000
6	1	-3.9474	3.39434	-3.12474	0.510165	0.0000	0.0000
7	1	-5.37351	3.99355	2.38035	0.464625	0.0000	0.0000
1	6	-1.44065	5.09981	0.3962	0.165619	0.1094	3.4000
1	6	-0.67866	5.68405	-0.78888	0.419045	0.1094	3.4000
1	6	0.51454	4.78985	-1.11481	0.537684	0.1094	3.4000
1	6	1.31577	4.53425	0.15282	0.203297	0.1094	3.4000
1	6	0.49942	3.97893	1.29863	0.229816	0.1094	3.4000
1	6	1.30717	3.87442	2.60313	0.531209	0.1094	3.4000
2	8	-1.57716	5.86566	-1.87978	-0.847212	0.2104	3.0665
2	8	1.34938	5.48909	-2.04516	-0.912695	0.2104	3.0665
3	8	2.34229	3.54515	-0.16823	-0.746198	0.1700	3.0000
4	8	-0.59861	4.90517	1.5307	-0.577279	0.1700	3.0000
5	8	2.45111	4.74758	2.64123	-0.865633	0.2104	3.0665
6	1	-2.08598	5.05634	-2.07185	0.503915	0.0000	0.0000
6	1	2.07313	4.91745	-2.35525	0.531320	0.0000	0.0000
7	1	3.23443	4.30312	2.28157	0.518896	0.0000	0.0000
1	6	3.69263	3.92548	-0.17613	0.636701	0.1094	3.4000
1	6	4.34092	3.34998	-1.4236	0.267454	0.1094	3.4000
1	6	4.18401	1.84213	-1.40089	0.464953	0.1094	3.4000
1	6	4.82475	1.22342	-0.16805	0.026731	0.1094	3.4000
1	6	4.3385	1.91995	1.10851	0.478544	0.1094	3.4000
1	6	5.25701	1.65921	2.31705	0.359947	0.1094	3.4000
2	8	3.6628	3.89049	-2.56859	-0.834341	0.2104	3.0665
2	8	4.69955	1.2797	-2.63884	-0.865350	0.2104	3.0665
3	8	4.34529	-0.13861	-0.14276	-0.599718	0.1700	3.0000
4	8	4.37429	3.37107	0.96258	-0.877785	0.1700	3.0000
5	8	5.93016	0.3941	2.25457	-0.866993	0.2104	3.0665
6	1	3.80172	3.31159	-3.3331	0.515434	0.0000	0.0000
6	1	5.66683	1.34326	-2.68231	0.485434	0.0000	0.0000
7	1	5.33702	-0.36467	2.36297	0.548921	0.0000	0.0000
1	6	5.21472	-1.24807	0.07912	0.183227	0.1094	3.4000
1	6	5.12273	-2.13702	-1.15415	0.607413	0.1094	3.4000
1	6	3.70767	-2.67924	-1.29392	0.320396	0.1094	3.4000
1	6	3.29311	-3.35376	0.00506	0.437391	0.1094	3.4000
1	6	3.41281	-2.45547	1.22304	0.167082	0.1094	3.4000
1	6	3.27543	-3.24697	2.51536	0.360227	0.1094	3.4000
2	8	5.56455	-1.37574	-2.27816	-0.892532	0.2104	3.0665
2	8	3.70346	-3.64939	-2.3446	-0.899833	0.2104	3.0665
3	8	1.89684	-3.71805	-0.15558	-0.666876	0.1700	3.0000
4	8	4.77551	-1.93161	1.25934	-0.653336	0.1700	3.0000
5	8	4.28265	-4.27775	2.54548	-0.806738	0.2104	3.0665
6	1	4.9298	-0.67226	-2.49182	0.537198	0.0000	0.0000
6	1	2.81042	-4.01825	-2.45274	0.537168	0.0000	0.0000
7	1	5.15015	-3.85086	2.48692	0.472395	0.0000	0.0000
8	1	2.28138	-5.701	0.29719	0.082430	0.0000	0.0000
8	1	0.27604	-6.47661	-0.98714	0.048921	0.0000	0.0000
8	1	-0.15283	-3.52481	-1.46063	0.043637	0.0000	0.0000
8	1	-1.8427	-5.51783	0.07562	0.007815	0.0000	0.0000

8	1	-0.26671	-3.18511	1.17204	0.032204	0.0000	0.0000
8	1	-2.27827	-4.04344	2.4819	0.040720	0.0000	0.0000
8	1	-0.74952	-4.29214	3.33386	-0.026896	0.0000	0.0000
8	1	-3.94173	-4.85542	0.12455	0.111519	0.0000	0.0000
8	1	-5.53578	-3.45332	-1.17687	0.039568	0.0000	0.0000
8	1	-3.17642	-1.5867	-1.46366	0.013791	0.0000	0.0000
8	1	-5.80842	-1.17317	-0.02983	0.021770	0.0000	0.0000
8	1	-3.04805	-1.41893	1.17233	0.046292	0.0000	0.0000
8	1	-4.84399	-0.15716	2.45952	0.052530	0.0000	0.0000
8	1	-4.32321	-1.63345	3.28054	-0.025334	0.0000	0.0000
8	1	-6.25465	0.99204	0.1156	0.113052	0.0000	0.0000
8	1	-5.77197	3.10358	-1.1124	0.048511	0.0000	0.0000
8	1	-2.98945	1.96186	-1.39291	0.010252	0.0000	0.0000
8	1	-3.92987	4.39998	0.128	0.029833	0.0000	0.0000
8	1	-2.84054	1.80159	1.2269	0.046329	0.0000	0.0000
8	1	-2.55646	3.87318	2.65075	-0.003821	0.0000	0.0000
8	1	-3.6679	2.70588	3.37473	-0.046849	0.0000	0.0000
8	1	-2.24496	5.74191	0.71267	0.181667	0.0000	0.0000
8	1	-0.30228	6.66057	-0.52886	0.040775	0.0000	0.0000
8	1	0.15854	3.84675	-1.50782	-0.019434	0.0000	0.0000
8	1	1.78201	5.45283	0.46857	0.058004	0.0000	0.0000
8	1	0.09292	3.01715	1.02848	0.023147	0.0000	0.0000
8	1	1.60408	2.84247	2.75305	-0.112826	0.0000	0.0000
8	1	0.65516	4.16578	3.41037	0.003611	0.0000	0.0000
8	1	3.81823	4.99225	-0.11128	0.109363	0.0000	0.0000
8	1	5.38821	3.62834	-1.42925	0.012125	0.0000	0.0000
8	1	3.13418	1.61342	-1.40238	0.069253	0.0000	0.0000
8	1	5.9049	1.25152	-0.23108	0.109777	0.0000	0.0000
8	1	3.32252	1.59924	1.27201	0.051819	0.0000	0.0000
8	1	4.70092	1.7819	3.23863	-0.013661	0.0000	0.0000
8	1	6.02928	2.40897	2.28408	0.055836	0.0000	0.0000
8	1	6.21902	-0.92533	0.27268	0.161122	0.0000	0.0000
8	1	5.80171	-2.96863	-1.05549	-0.001102	0.0000	0.0000
8	1	3.02676	-1.86095	-1.49157	0.011123	0.0000	0.0000
8	1	3.89198	-4.23796	0.16143	0.001668	0.0000	0.0000
8	1	2.70308	-1.64603	1.1617	0.027254	0.0000	0.0000
8	1	2.32287	-3.74814	2.5468	0.070291	0.0000	0.0000
8	1	3.36449	-2.58017	3.36464	-0.008389	0.0000	0.0000

E finalmente os parâmetros utilizados para a água:

Tipo	N° At	X	Y	Z	Carga	Epslon	Sigma
1	8	0.0000	0.0000	0.0000	-0.820	0.1550	3.1650
2	1	0.5774	0.8165	0.0000	0.410	0.0000	0.0000
2	1	0.5774	-0.8165	0.0000	0.410	0.0000	0.0000

Referências Bibliográficas

Allen, M. P. and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford, New York, 1989.

Allinger, N. L., *J. Am. Chem. Soc.*, **99**, 8127 (1977).

Atkins, P. W. and R. S. Friedman, *Molecular Quantum Chemistry*, 3rd Edition, Oxford, New York, 1997.

Berendsen, H. J. C. et al., *Intermolecular Forces*, edited by B. Pullmann, Reidel, Dordreicht, 1981, p. 331.

Binder, K. and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics: An Introduction*, 3rd edition, Springer, Heidelberg, 1997.

Breslow, R. and S. D. Dong, *Chem Rev.*, **98**, 1997 (1998).

Canuto, S., *Métodos Clássicos e Quânticos em Estrutura Eletrônica*, Notas de Aula, Universidade de São Paulo, 2001.

Castro, M. A., *A Aproximação de Hartree-Fock*, Notas de Aula, Universidade Federal de Goiás, 1999.

Chacko, K. K. and W. Saenger, *J. Am. Chem. Soc.*, **103**, 1708 (1981).

Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson., D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, **117**, 5179 (1995).

Coutinho, K., *Modelo Discreto de Solvente. Solvatocromismo no Espectro de Absorção Molecular*, Tese de Doutoramento, Universidade de São Paulo, 1997.

Cruzan, J. D., L. B. Braly, K. Liu, M. G. Brown, J. G. Loeser, R. J. Saykally , *Science*, **271**, 59 (1996).

Dewar, M. J. S., E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *J. Am. Chem. Soc.*, **107**, 3902 (1985).

DICE: A Monte Carlo Program for Molecular Liquid Simulation, K. Coutinho, and S. Canuto, Universidade de São Paulo, 1997.

Elliott, S. R., *Physics of Amorphous Materials*, 2nd Edition, Longman Scientific, New York, 1990.

Englman, R., *The Jahn-Teller Efect in Molecules and Crystals*, Wiley, New York, 1972.

Garrett, R. H. and C. M. Grisham, *Biochemistry*, Saunders College Publishing, 1995.

Gattuso, G., S. A. Nepogodiev and J. F. Stoddart, *Chem. Rev.*, **98**, 1919 (1998).

Gaussian 98 Revision A.7, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian Inc., Pittsburgh, 1998.

Gingrich, N. S., *Rev. Mod. Phys.*, **15**, 90 (1943).

Guinier, A., G. Fournet, C. B. Walker and K. L. Yudowitch, *Small Angle Scattering of X-Ray*, Wiley, New York, 1955.

Harata, K., *Chem Rev.*, **98**, 1803 (1998).

Hedges, A. R., *Chem. Rev.*, **98**, 2035 (1998).

Hill, T. L., *J. Chem. Phys.*, **16**, 399 (1948).

Hohmans, S. W., *Biochemistry*, **29**, 9110 (1990).

Hyperchem 5.0 for Windows, Hypercube Inc., Gainesvile, 1996.

Jensen, F., *Introduction to Computational Chemistry*, Wiley, Chichester, 1999.

Jorgensen, W. L. and C. Ravimohan, *J. Chem. Phys.*, **83**, 3050 (1985).

Jorgensen, W. L. and J. Tirado-Rives, *J. Am. Chem. Soc.*, **110**, 1657 (1988).

Jorgensen, W. L., *Acc. Chem. Res.*, **22**, 184 (1989).

Jorgensen, W. L. and D. L. Severance, *J. Am. Chem. Soc.*, **112**, 4768 (1990).

Karelson, M. M. and M. C. Zerner, *J. Phys. Chem.*, **96**, 6949 (1992).

Khan, A. R., P. Forgo, K. J. Stinne and V. T. D'Souza, *Chem. Rev.*, **98**, 1977 (1998).

King, P. M., C. A. Reynold, J. W. Essex, G. A. Worth and W. G. Richards, *Mol. Simul.*, **5**, 265, (1990).

Kitagawa, M., H Hoshi, M. Sakurai, Y. Inoue and K. Chûjô, *Carbohyd. Res.*, **163**, c1 (1987).

Klar, B, B. Hingerty and W. Saenger, *Acta Cryst. B*, **36**, 1154 (1980).

Lage, E. J. S., *Física Estatística*, Fundação Calouste Gulbenkian, Lisboa, 1995.

Leach, A. R., *Molecular Modeling. Principles and Aplications*, 2nd Edition, Pearson Education, Harlow, 2001.

Lennard-Jones, J. E., *Physica*, **4**, 957 (1937).

Levine, I. N., *Quantum Chemistry*, 4th Edition, Prentice Hall, Englewood Cliffs, 1991.

Li, X., L. Liu, T. Um and Q. Guo, *Monatshefte Chem.*, **131**, 849 (2000).

Lii, J. H., S. Gallion, C. Bender, H. Wikstrom, N. L. Allinger, K. M. Flurchick and M. M. Teeter, *J. Comp. Chem.*, **10**, 503 (1989).

Lindner, K. and W. Saenger, *Acta Cryst. B*, **38**, 203 (1982).

Lipkowitz, K. B., *Chem Rev.*, **98**, 1829 (1998).

Lipkowitz, K. B., *QCPE Bulletin - Indiana University*, **12**, 1 (1992).

Malaspina, T., K. Coutinho, S. Canuto, *J. Chem. Phys.*, **117**, 000 (2002).

Manor, P. C. and Saenger, *J. Am. Chem. Soc.*, **96**, 3630 (1974).

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.*, **21**, 1087 (1953).

Mezei, M. and D. Beveridge, *J. Chem. Phys.*, 74, 622 (1981).

Momany, F. A. and J. L. Willet, *Carbohyd. Res.*, **326**, 194 (2000)a.

Momany, F. A. and J. L. Willet, *Carbohyd. Res.*, **326**, 210 (2000)b.

Morse, P. M., *Phys. Rev.*, **34**, 57 (1929).

Nelson, D. L. and M. M. Cox, *Lehninger Principles of Biochemistry*, 3rd Edition, Worth Publishers, New York, 2000.

Nepogodiev, S. A. and J. F. Stoddart, *Chem. Rev.*, **98**, 1959 (1998).

Pople, J. A. and D. L. Beveridge, *Aproximate Molecular Orbital Theory*, McGraw Hill, New York, 1970.

Pople, J. A. and R. K. Nesbet, *J. Chem. Phys.*, **22**, 571 (1954).

Pranata, J., S. G. Wierschke and W. L. Jorgensen, *J. Am. Chem. Soc.*, **113**, 2810 (1991).

Puliti, R., C. A. Mattia, L. Paduano, *Carbohyd. Res.*, **310**, 1 (1998).

Rekharsky, M. V. and Y. Inoue, *Chem Rev.*, **98**, 1875 (1998).

Rivail, J. L. and D. Rinaldi, *Chem. Phys.*, **18**, 233 (1976).

Roothaan, C. C. J., *Rev. Mod. Phys.*, **23**, 69 (1951).

Saenger, W., J. Jacob, K. Gessler, T. Steiner, D. Hoffman, H. Sanbe, K. Koizumi, S. M. Smith and T. Takaha, *Chem Rev.*, **98**, 1787 (1998).

Salinas, S. R. A., *Introdução à Física Estatística*, EDUSP, São Paulo, 1997.

Santos, H. F., H. A. Duarte, R. D. Sinisterra, S. V. M. Mattos, L. F. C. Oliveira and W. B. Almeida, *Chem. Phys Lett.*, **319**, 569 (2000).

Schneider, H-J., F. Hacket and V. Rüdiger, *Chem Rev.*, **98**, 1755 (1998).

Sheykhet, I. I. and B. Ya. Simkin, *Comp. Phys. Rep.*, **12**, 67 (1990).

Stewart, J. J. P., *J. Comp. Chem.*, **10**, 209 (1989).

Stewart, J. J. P., *J. Comp. Chem.*, **11**, 543 (1990).

Stillinger, F. H., *Science*, **209**, 451 (1980).

Stone, A. J., *The Theory of Intermolecular Forces*, Oxford, New York, 1996.

Szabo, A. and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Eletronic Structure Theory*, Dover, Mineola, 1996.

Szejtli, J., *Chem. Rev.*, **98**, 1743 (1998).

Takahashi, K., *Chem Rev.*, **98**, 2013 (1998).

Tapia, O. and O. Goscinski, *Molec. Phys.*, **29**, 1653 (1975).

Tomasi, J., *Theor. Chem. Acc.*, **103**, 196 (2000).

Uekama, K., F. Hirayama and T. Irie, *Chem Rev.*, **98**, 2045 (1998).

Weiner, S. J., P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, P. Weiner, *J. Am. Chem. Soc.*, **106**, 765 (1984).

Weiner, S. J., P. A. Kollman, D. T. Nguyen, D. A. Case, *J. Comp. Chem.*, **7**, 230 (1986).

Wittel, K. and S. P. McGlynn, *Chem. Rev.*, **77**, 745 (1977).

Xantheas, S. S., *J. Chem. Phys.*, **100**, 7523 (1994).

Zwanzig, R. W., *J. Chem. Phys.*, **22**, 1420 (1954).