# UNIVERSIDADE FEDERAL DE GOIÁS PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

LEONARDO MARTINS DE ABREU

VARIAÇÃO DA ENERGIA LIVRE NA HIDRATAÇÃO DE SÉRIES HOMÓLOGAS DE POLIOL, GLICINA E PEPTÍDEOS  $A_NK$  USANDO MODELOS CONTÍNUO E DISCRETO DE SOLVENTE

GOIÂNIA

2017





#### TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: [] Dissertação [x] Tese

#### 2. Identificação da Tese ou Dissertação

Nome completo do autor: Leonardo Martins de Abreu

Título do trabalho: Variação da energia livre na hidratação de séries homólogas de poliol, glicina e peptídeos A<sub>N</sub>K usando modelos contínuo e discreto de solvente.

#### 3. Informações de acesso ao documento:

Concorda com a liberação total do documento [ x ] SIM [ ] NÃO<sup>1</sup>

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.

(apparally) Assinatura do (a) autor (a)

Data: <u>12</u> / <u>05</u> / <u>2017</u>

<sup>&</sup>lt;sup>1</sup> Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

# **LEONARDO MARTINS DE ABREU**

# VARIAÇÃO DA ENERGIA LIVRE NA HIDRATAÇÃO DE SÉRIES HOMÓLOGAS DE POLIOL, GLICINA E PEPTÍDEOS $A_NK$ USANDO MODELOS CONTÍNUO E DISCRETO DE SOLVENTE

Tese apresentada ao Programa de Pós-Graduação em Física da Universidade Federal de Goiás como requisito para obtenção do título de Doutor em Física.

Prof. Dr. Tertius Lima Fonseca – Orientador Prof. Dr. Eudes Eterno Fileti – Coorientador

GOIÂNIA

2017

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Martins de Abreu, Leonardo VARIAÇÃO DA ENERGIA LIVRE NA HIDRATAÇÃO DE SÉRIES HOMÓLOGAS DE POLIOL, GLICINA E PEPTÍDEOS ANK USANDO MODELOS CONTÍNUO E DISCRETO DE SOLVENTE [manuscrito] / Leonardo Martins de Abreu 2017. cvii, 107 f.: il.
Orientador: Prof. Dr. Tertius Lima da Fonseca; co-orientador Dr. Eudes Eterno Fileti. Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Física (IF), Programa de Pós-Graduação em Física, Goiânia, 2017. Bibliografia. Apêndice. Inclui gráfico, tabelas, lista de figuras, lista de tabelas.
1. energia livre de hidratação. 2. poliol. 3. glicina. 4. peptídeos ANK. I. Lima da Fonseca, Tertius , orient. II. Título.
CDU 539



# Universidade Federal de Goiás Instituto de Física Programa de Pós-Graduação em Física

Ata Nº 30 de defesa de tese de Leonardo Martins de Abreu para obtenção do título de Doutor em Física.

Aos 12 dias do mês de maio de 2017, às 10h00min, no Miniauditório do Instituto de Física da UFG, reuniu-se a Banca Examinadora indicada pela Coordenadoria do Programa de Pós-Graduação, aprovada pelo Conselho Diretor e designada pela Diretoria do Instituto de Física da Universidade Federal de Goiás, composta pelo Prof. Dr. Tertius Lima da Fonseca (orientador e presidente da Banca - IF/UFG), Prof. Dr. Roberto Rivelino de Melo Moreno (IF/UFBA), Prof. Dr. Milton Massumi Fujimoto (DF/UFPR), Prof. Dr. Ladir Cândido da Silva (IF/UFG) e Prof. Dr. Renato Pessoa Vale (IF/UFG) para julgar a tese de doutorado de Leonardo Martins de Abreu, intitulada: "Variação da energia livre na hidratação de séries homólogas de poliois, glicinas e peptídeos ANK: um estudo via dinâmica molecular". O Presidente abriu os trabalhos agradecendo a presença dos membros da Banca Examinadora e concedeu a palavra a Leonardo Martins de Abreu, que expôs detalhadamente seu trabalho. Em seguida, os membros da Banca fizeram suas considerações e procederam à arguição do candidato. Concluída essa etapa, a Banca, em sessão fechada, deu prosseguimento ao julgamento do trabalho, atribuindo os seguintes conceitos:

Prof. Dr. Tertius Lima da Fonseca (orientador)

Prof. Dr. Roberto Rivelino de Melo Moreno (IF/UFBA)

Prof. Dr. Milton Massumi Fujimoto (DF/UFPR)

Prof. Dr. Ladir Cândido da Silva (IF/UFG)

Prof. Dr. Renato Pessoa Vale (IF/UFG)

APROVADO
APROVADO
Athymout APROVADO
Abrovento
Aprovado

Novamente em sessão aberta, o presidente da Banca anunciou o resultado final do julgamento, declarando o candidato Leonardo Martins de Abreu APROVADO pela Banca Examinadora. Nada mais havendo a tratar, a sessão foi encerrada e lavrou-se a presente ata que segue assinada pelos membros da Banca Examinadora.

Goiânia, 12 de maio de 2017.

tertins L. Fensera	Robert Justi	#thomas
Prof. Dr Tertius L. da Fonseca	Prof. Dr. Roberto R. de M. Moreno	Prof. Ør. Milton M. Fujimoto
Dodie Sch	at of Fi Dunto	PERSOA VAL

Prof. Dr. Ladir Cândido da Silva

Prof. Dr. Renato Pessoa Vale

À minha esposa, Adriana Pereira de Sousa Abreu e ao amigo João Batista de Macêdo.

#### **AGRADECIMENTOS**

A Deus, sem Ele nada seria possível.

À minha esposa, Adriana, por cuidar de mim, de nós, de nossos filhos, de nossa casa, de nosso ministério com tanto zelo. Por me ensinar o é o amor de Deus, o que é o amor de esposa e o que é o amor de ser pai. Por tornar essa caminha muito mais segura e tranquila. Você é a melhor parte de mim. Amo você! Aos meus pais, João Antônio Gonçalves de Abreu e Hamilta Martins Pereira Gonçalves de Abreu, por todo investimento de amor incondicional e por todo o empenho empregado em minha formação, sobretudo em minha formação educacional e pelo.

Ao professor Dr. Tertius Lima da Fonseca por todo o respeito e atenção dedicados. Sobretudo por toda a significativa contribuição em minha formação acadêmica nos últimos anos sob sua orientação que permitiu a realização deste trabalho.

Ao professor Eudes Eterno Fileti por todo o ensinamento, pela valiosa colaboração e contribuição na coorientação deste trabalho. Ao professor Dr. Guilherme Colherinhas por todas as conversas, sugestões e orientações que ajudaram a construir este trabalho.

Aos familiares, amigos, em especial os companheiros de curso, aos professores e demais funcionários do Instituto de Física da Universidade Federal de Goiás.

Em tempo agradeço ainda à Lílian Maria Cintra Campos Batista pelos valiosos ensinamentos, pela confiança que depositou em mim e principalmente por sua amizade. Ao amigo João Batista de Macêdo por todo o cuidado, atenção e investimento dedicados quando eu era apenas um aluno de ensino médio. Pelo exemplo de pessoa e de profissional que é. Pela amizade e pelos valiosos ensinamentos.

Ao CNPq pelo suporte financeiro

Se cheguei até aqui foi porque me apoiei no ombro dos gigantes.

Isaac Newton

#### RESUMO

Utilizando modelos contínuo e discreto de descrição de solventes investigamos o comportamento da variação da energia livre na hidratação de series homólogas de poliol, glicina e peptídeos A<sub>N</sub>K. Neste trabalho os valores da energia de hidratação foram obtidos por integração termodinâmica utilizando simulações de dinâmica molecular clássica em sistemas com uma molécula de soluto para aproximadamente 1000 moléculas de solvente. Adicionalmente, os resultados de ΔG foram obtidos utilizando o modelo contínuo polarizável combinado com o modelo de solvatação universal. Com o modelo PCM estudamos ainda o momento de dipolo e a polarizabilidade linear destes sistemas em questão e analisamos as contribuições que estas propriedades tem sobre a energia livre de hidratação e sua evolução com o crescimento de sistemas. Os resultados obtidos mostram que a variação da energia livre em função do número de resíduos cresce em valor absoluto para os três sistemas estudados. Para as moléculas de glicina e de peptídeos A<sub>N</sub>K observa-se um comportamento linear de ΔG com o número N de subunidades do sistema, semelhante ao observado para sistemas moleculares apresentados na literatura. Diferentemente, para as moléculas de poliol, os resultamos de ΔG exibem um comportamento não linear com o crescimento do sistema. Para esse grupo de moléculas inclusão do solvente tem grande impacto sobre o momento de dipolo e sobre a polarizabilidade linear. Este resultado mostra que para as moléculas de poliol não apenas a polarizabilidade, mas também variações no momento de dipolo podem influenciar o comportamento da energia livre de hidratação. Todos os cálculos foram realizados utilizando o programa GROMACS 4.5.4 e GAUSSIAN09.

#### ABSTRACT

Using continuum and discrete solvent description models, we have investigated the chain length dependence of hydration free energy of polyol, glycine and  $A_NK$  peptides. In this work this property values were calculated using thermodynamics integration with classical molecular dynamics simulation of one solute molecule to around 1000 solvent molecules. In addition, hydration free energy of polyol, glycine and  $A_NK$  peptides were also calculated using polarizable continuum model in combination with the universal solvation model. Dipole moment and linear polarizability of the three systems had been also calculate using the PCM method. Our MD results show that  $\Delta G$  exhibit a linear pattern with the size of glycine and  $A_NK$  chain as well as observed for organic compounds in literature. In opposition, polyol's  $\Delta G$  shows a nonlinear behavior with increasing chain length and for this group of molecules it were found that water solvating effects have an important role over the polyol's electric properties. These results suggest that in the case of polyols not only variation of polarizability but also variations of dipole moment in going from vacuum to water can influence the behavior of the free energy of hydration. All DM and PCM calculations were performed using the GROMACS 4.5 and the GAUSSIAN09 programs, respectively.

### LISTA DE FIGURAS

Figura 1. Ilustração das sete primeiras moléculas de poliol representadas pela série C<sub>N</sub>H<sub>2N+2</sub>O<sub>N</sub>. Em cinza átomos de carbono, em vermelho átomos de oxigênio e em branco átomos de hidrogênio. 19

Figura 6. Representação esquemática do raio de corte aplicado para interações nas simulações de dinâmica molecular. Em destaque a condição de mínima imagem obtida para a circunferência de raio menor ou igual à metade da menor aresta da caixa de simulação (menor raio) na interação da partícula centrada na circunferência com as demais da região delimitada pelo raio de corte...... 26

Figura 9. Evolução da polarizabilidade linear ( $\alpha$ , em u.a.) para moléculas de poliol obtida no vácuo e em solução utilizando o método PCM-MP2 com conjunto de funções base 6-311+G(2d). ..... 59

Figura 11. Energia de polarização do soluto (U<sub>POL</sub>, em kJ/mol) obtida para moléculas de poliol com os modelos B3LYP e MP2 e conjunto de funções base 6-311+G(2d)......63

#### LISTA DE TABELAS

Tabela 2. Comparação entre os valores de energia livre de hidratação ( $\Delta G$ , em kJ/mol) para moléculas de poliol obtidos com os modelos MD-CHARMM, PCM-B3LYP/6-311+G(2d) e PCM-MP2/6-311+G(2d)......60

Tabela 6. Momento de dipolo (D) e polarizabilidade linear (u.a.) no vácuo e em solução para peptídeos A<sub>N</sub>K obtidos no nível B3LYP/6-311+G(2d)......76

# SUMÁRIO

AGRADECIMENTOSV				
RESUMOVII				
ABSTRACTVIII				
LISTA DE FIGURASIX				
LISTA	DE TABELAS	XI		
1	INTRODUÇÃO	.14		
2	METODOLOGIA	. 23		
2.1	DINÂMICA MOLECULAR	23		
2.1.1	CAIXA DE SIMULAÇÃO E CONDIÇÕES DE CONTORNO PERIÓDICAS	24		
2.1.2	A SOMA DE EWALD	27		
2.1.3	AS EQUAÇÕES DE MOVIMENTO	29		
2.1.4	A INTEGRAÇÃO DAS EQUAÇÕES DE MOVIMENTO	30		
2.1.5	AS PROPRIEDADES TERMODINÂMICAS	32		
2.1.6	ACOPLAMENTO DE TEMPERATURA E PRESSÃO	35		
2.1.7	A ENERGIA LIVRE DE HIDRATAÇÃO ( $\Delta G$ )	39		
2.1.8	MODELO DE CAROÇO MACIO (SOFT-CORE)	41		
2.2	MODELOS CONTÍNUO POLARIZAVEL E DE SOLVATAÇÃO UNIVERSAL.	42		
2.3	O MÉTODO DE HARTREE-FOCK	45		
2.4	A TEORIA DE PERTURBAÇÃO DE MØLLER-PLESSET	50		
2.5	O PROCEDIMENTO DE CAMPO FINITO	52		
2.6	DETALHES COMPUTACIONAIS	. 52		

3	RESULTADOS E DISCUSSÕES	56
3.1	POLIOL	56
3.2	GLICINA	65
3.3	PEPTÍDEOS A <sub>N</sub> K	76
4	CONCLUSÕES E PERSPECTIVAS	. 80
REFE	RÊNCIAS	. 82
APÊNI	DICE A – MODELO DE ARQUIVO DE PARÂMETROS DE SIMULAÇÃO DE	1
DINÂMICA MOLECULAR9		
APÊNI	DICE B – PARAMETRIZAÇÃO DO CAMPO DE FORÇA CHARMM PARA	
MOLÉ	CULAS DE POLIOL	. 97

# 1 INTRODUÇÃO

O interesse no estudo teórico de sistemas líquidos e de soluções tem crescido em virtude do grande número de processos físicos, químicos e biológicos que ocorrem nestes meios [1-4]. Se olharmos para o funcionamento do corpo humano, por exemplo, veremos que a maior parte dos processos celulares ocorre em meio aquoso e, portanto, as reações químicas associadas a eles acontecem em solução. A compreensão dos aspectos estruturais e eletrônicos de sistemas moleculares bem como das suas interações com o ambiente podem ser determinantes no desenvolvimento de novos materiais com aplicações tecnológicas [5-9]. Esta compreensão abre espaço ainda para melhorar propriedades e processos que trazem benefícios como, por exemplo, sínteses mais rápidas, eficientes e menos dispendiosos do ponto de vista energético ou mesmo que tornem tratamentos terapêuticos medicamentosos mais eficientes, menos debilitantes e mais baratos.

A descrição de sistemas físicos no estado líquido é um grande desafio em razão da grande complexidade envolvida. Do ponto de vista microscópico substâncias em fase líquida e em solução apresentam alta densidade, grande mobilidade dos constituintes e um alto grau de desordem estrutural. Além do grande número de configurações diferentes que os sistemas líquidos podem apresentar as interações entre moléculas precisam ser descritas de forma adequada e esse problema se torna ainda mais complexo uma vez que não existe uma equação simples capaz de descrever as forças intermoleculares. No caso de soluções a mudança de ambiente para o soluto pode mudar de forma bastante significativa suas propriedades estruturais e eletrônicas [10,11]. Neste sentido, as simulações computacionais clássicas têm um papel importante no estudo destes sistemas porque permitem que tanto teorias e modelos sejam testados quanto permitem predizer que materiais são mais adequados para uma determinada aplicação ou quais processos podem ocorrer para sistemas específicos estudados. A evolução dos computadores tem permitido que sistemas cada vez maiores e mais complexos possam ser investigados [12,13]. Com as novas capacidades computacionais uma série de abordagens tem sido utilizada para estudar as propriedades termodinâmicas de sistemas líquidos e tratar os efeitos do solvente sobre as propriedades do soluto no caso de soluções.

Atualmente, o uso do método de Monte Carlo, da Dinâmica Molecular e do Modelo Contínuo Polarizável estão entre as principais formas de estudar os sistemas mencionados [14-15].

No método de Monte Carlo (MC), os valores esperados para as propriedades do sistema simulado são determinados através de médias estatísticas sobre um conjunto de configurações descorrelacionadas. A partir de uma configuração inicial aleatória que obedece a distribuição de Boltzmann cada nova configuração é obtida por alterações nos graus de liberdade do sistema, onde para cada modificação o valor da energia é calculado para aceitação ou rejeição da configuração com modificações. Se a energia da nova configuração for menor que a da anterior ela é então incluída na amostra e passa a ser a nova configuração inicial, se a energia da nova configuração for maior, um número aleatório é sorteado e, caso esse número seja menor que o valor para aceitação aceita-se a nova configuração, caso contrário rejeita-se ela. Cada uma dessas repetições é dita ser um passo Monte Carlo e este processo é repetido até que o critério de parada seja satisfeito. Este modelo tem sido utilizado com sucesso pelo grupo de Física Atômica e Molecular do Instituto de Física da Universidade Federal de Goiás através do programa DICE [16]. Algumas restrições deste modelo são o elevado tempo computacional e a utilização de uma geometria rígida para o soluto. Uma alternativa para contornar as limitações mencionadas, aparece com a utilização da dinâmica molecular.

A dinâmica molecular (DM) clássica modela o sistema em estudo utilizando uma caixa de simulação que contém os átomos e moléculas do soluto e do solvente em uma configuração inicial aleatória. Cada passo de simulação na DM consiste na determinação das novas posições e momentos dos constituintes do sistema. Para isso primeiramente é determinado o valor do potencial efetivo sobre cada partícula, a partir dele é obtida a força resultante sobre cada um deles. Com o valor da força é determinada então a aceleração a que estão submetidos cada um dos constituintes do sistema simulado. Como os passos de simulação possuem um tempo definido, com os valores de aceleração são obtidas as novas posições e momentos que estabelecem a nova configuração do sistema. Este processo é repetido de forma que se obtenha um sistema equilibrado, capaz de representar o sistema real e que permita determinar as propriedades desejadas. Uma descrição mais detalhada da DM será apresentada na seção 2.1.

Para uma descrição mais completa dos sistemas de fase líquida faz-se necessária a inclusão dos efeitos quânticos, entretanto, o custo computacional de incluir esses efeitos para todas as

moléculas do liquido é extremamente elevado. Desta forma, uma alternativa para levar em consideração, ao menos em parte, a estrutura eletrônica das moléculas em solução é a utilização do modelo contínuo polarizável (do inglês polarizable continuum model - PCM). Na aproximação PCM o soluto é tratado de forma quântica por primeiros princípios enquanto o solvente é tratado como um meio uniforme representado por uma constante dielétrica, que representa a média das interações da molécula de soluto com cada molécula de solvente. Diferentemente do método de Monte Carlo e da Dinâmica Molecular, neste caso o solvente é tratado como um único elemento, que pode ser polarizado em consequência da distribuição de cargas do soluto. A polarização do solvente é capaz de induzir uma redistribuição de cargas no soluto, definida como potencial de reação. Esse ajuste de cargas no soluto e no solvente pode ser tratado de forma autoconsistente e é a chave para o modelo PCM que se baseia na solução da equação de Poisson do sistema.

Utilizando os modelos mencionados muitos trabalhos têm se dedicado ao estudo de propriedades como a energia livre de Gibbs, entropia e entalpia [17] para descrever e entender o comportamento de sistemas orgânicos afim de elucidar processos físicos, químicos e biológicos. Estimar a variação na energia livre de Gibbs é de extrema importância para a caracterização de mecanismos de reação e equilíbrio químico. Esta variação da energia livre, chamada energia livre de hidratação ( $\Delta G$ ), é a diferença entre a energia livre para um soluto em água e em vácuo. É importante ressaltar que esta variação está associada a espontaneidade de reações químicas, aos efeitos hidrofóbicos e hidrofílicos, além de poder ser utilizada para avaliar a solvatação ou não de um soluto em um dado solvente e para explicar uma série de fenômenos físico-químicos em meio aquoso [18-25]. Do ponto de vista computacional esta propriedade pode ser utilizada para parametrização de campos de forças usados em DM.

Neste trabalho temos por objetivo estudar à variação da energia livre de Gibbs na hidratação de sistemas moleculares (poliol, glicina e peptídeos  $A_N K$ ) e o seu comportamento com o aumento do número de átomos no sistema. É objetivo deste trabalho ainda analisar o impacto que as variações nas propriedades elétricas destas moléculas têm sobre a variação da energia livre de hidratação. Em particular deseja-se obter domínio no conhecimento e utilização do programa GROMACS, que possibilitará estudar de novas formas os sistemas estudados hoje e adicionalmente investigar novos sistemas e diferentes propriedades físico-químicas de sistemas líquidos. Aqui os resultados de  $\Delta G$  foram obtidos via dinâmica molecular utilizando o módelo contínuo polarizável combinado com o modelo

de solvatação universal (do inglês universal solvente model - SMD). O modelo continuo polarizável com formalismo de equações integrais (do inglês integral equation formalism, IEF-PCM) permitiu estudar o momento de dipolo e a polarizabilidade linear destes sistemas em questão e estimar as contribuições que estas propriedades tem sobre a energia livre de hidratação e sua evolução com o crescimento de sistemas. Aqui analisamos qual o impacto das mudanças no momento de dipolo na presença de solvente sobre a variação da energia livre de hidratação e sua relação com a fuga da linearidade que é a base dos métodos de aditividade de grupo utilizados para descrever os dados disponíveis e prever o comportamento das funções termodinâmicas de compostos orgânicos extensos constituídos de unidades menores repetidas. É importante ressaltar que este trabalho se limita ao estudo de um único confôrmero de cada molécula, no caso dos cálculos quânticos, o que significa não incluir os efeitos de temperatura no sistema e também não incluir flexibilidade na geometria do soluto. Outra limitação deste trabalho é a polarização do soluto nos cálculos de dinâmica molecular considerada apenas na distribuição de cargas que representa o soluto e não inclui a sua modificação com a evolução da simulação. Todos os cálculos clássicos e quânticos foram realizados com o programa GROMACS e GAUSSIAN09, respectivamente.

As moléculas de poliol são compostos orgânicos da família dos álcoois que contém um ou mais grupos hidroxila conectados à cadeia carbônica. Estes compostos são bastante utilizados na indústria alimentar e na química de polímeros. Na indústria de alimentos são utilizados como edulcorantes, substituindo os hidratos de carbono, por terem um valor energético mais baixo. Estas moléculas também estão presentes na composição de compostos utilizados nas áreas médica, farmacêutica e de higiene pessoal na forma de produtos como: xaropes expectorantes e para tosse; cremes dentais e enxaguantes bucais; sabonetes; e produtos para pele e cabelo, e são ainda base para a produção de outros compostos químicos como formaldeídos, plásticos, tintas e explosivos. Nos últimos anos um grande número de trabalhos tem se dedicado a estudar as moléculas de poliol e suas aplicações. Estudos recentes mostram a sua utilização na seleção de hidrofluorcarbonetos eficientes para produção de poliuretano [26], no estudo da estabilidade de proteínas nestes solventes a baixas temperaturas [27] e na produção de dendrímeros com estruturas hiperramificadas para encapsulamento fotoativo, núcleo-catalíticos e oxiredutivos [28,29].

Consideramos neste estudo a série de poliol estudada é composta por moléculas saturadas, alifáticas com átomos de carbono mono hidroxilados, representados pela série  $C_N H_{2N+2} O_N$ , onde

N varia de 1 a 7. Compreender as propriedades de solvatação de moléculas em diferentes ambientes é de grande importância para a compreensão de vários processos físicos, químicos e biológicos. Sistemas polares, como o poliol, podem apresentar para a energia livre e para outras propriedades termodinâmicas um padrão linear de dependência com o número de átomos de carbono. Esta dependência linear constitui a base dos métodos de aditividade de grupo utilizados para descrever os dados disponíveis e prever o comportamento das funções termodinâmicas de compostos orgânicos extensos constituídos de unidades menores repetidas [30]. Entretanto, para alguns sistemas esta linearidade com o número de átomos de carbono pode não ser observada em razão de diferenças geométricas ou energéticas envolvidas na adição de cada nova unidade. Os poliol são sistemas com muitos sítios polares, para os quais se podem esperar um desvio na linearidade na dependência da energia livre com o número de átomos de carbono em razão das interações entre grupos vizinhos (hidroxilas) que podem promover mudanças geométricas e/ou energéticas na adição de novas unidades H-C-OH. Para investigar esta hipótese as cadeias da série  $C_N H_{2N+2} O_N$ , são atraentes pois, além de possuírem múltiplos grupos hidroxila também possuem papel importante em bioquímica como precursores metabólicos e fazendo parte do processo de regulação osmótico [31]. A Figura 1 representa de forma esquemática as sete primeiras moléculas da série de poliol estudada.



Figura 1. Ilustração das sete primeiras moléculas de poliol representadas pela série  $C_NH_{2N+2}O_N$ . Em cinza átomos de carbono, em vermelho átomos de oxigênio e em branco átomos de hidrogênio.

Nos últimos anos um grande número de peptídeos tem sido estudado em função da sua versatilidade, biocompatibilidade e possibilidade de aplicação medicinal [32-42]. A atividade microbiótica dos peptídeos compreende um largo espectro bactericida e antifúngicos, além de atividade antiviral e anticâncer [41,43-46]. Neste sentido, a atividade bactericida de vários peptídeos tem sido exaustivamente investigada nas últimas décadas [9,47-54] e os peptídeos anfifílicos constituem uma importante classe de estudos de compostos com esta propriedade. Por apresentarem uma região hidrofílica e uma hidrofóbica mostram uma capacidade de auto-organização em água que origina diferentes nanoestruturas como fibras, nanohastes, nanotubos e bicamadas [55]. O tipo de estrutura formada depende de fatores da conformação do peptídeo e de propriedades físico-químicas como pH, temperatura, força iônica, pela concentração dos grupos anfifílicos.

Trabalhos recentes examinaram a relação antibacteriana para os peptídeos A<sub>3</sub>K, A<sub>6</sub>K e A<sub>9</sub>K [8,9]. Estes estudos publicaram a síntese destes compostos e a caracterização de suas atividades antibacterianas observando a formação de nano estruturas. Diferentemente do que ocorre para a maior parte dos compostos com atividade antibacteriana que apresentam estrutura grande e complexa, as moléculas A<sub>N</sub>K são sintetizadas com sequências curtas de moléculas estruturalmente simples (lisina e alanina). Estas moléculas apresentam um comportamento anfotérico, que deriva da presença, na sua estrutura, de dois grupos funcionais de diferentes polaridades (o grupo amino, NH2 e o grupo carboxila COOH), estes grupos funcionais tem uma participação importante no número de ligações de hidrogênio. Por apresentarem uma região hidrofílica e uma hidrofóbica são capazes de mimetizar a base estrutural de membranas lipídicas e mostram uma capacidade de autoorganização em água que origina diferentes nanoestruturas [55]. Para compreender sistemas maiores e mais complexos como as moléculas A<sub>N</sub>K estudamos neste trabalho primeiro a variação da energia livre de hidratação em moléculas de glicina considerando sua geometria neutra em solução.

A glicina é o aminoácido mais simples e compreender como esse sistema se comporta pode ser uma importante chave nos estudos de aspectos estruturais, dinâmicos e de formação de sistemas mais complexos. Nas últimas décadas vários trabalhos têm se dedicado a estudar de forma teórica e experimentalmente as propriedades estruturais e eletrônicas da glicina [56-61]. Neste trabalho apresentamos a evolução da energia livre de hidratação com o crescimento do sistema para as sete primeiras moléculas de glicina e para os peptídeos A<sub>3</sub>K, A<sub>6</sub>K e A<sub>9</sub>K. As Figuras 2 e 3 apresentam de forma esquemática, respectivamente, as sete primeiras moléculas da série de glicina e os peptídeos A<sub>N</sub>K estudados.



Figura 2. Ilustração das sete primeiras moléculas de glicina. Em cinza átomos de carbono, em vermelho átomos de oxigênio, em azul átomos de nitrogênio e em branco átomos de hidrogênio.



Figura 3. Ilustração dos peptídeos A<sub>3</sub>K, A<sub>6</sub>K e A<sub>9</sub>K. Em cinza átomos de carbono, em vermelho átomos de oxigênio, em azul átomos de nitrogênio e em branco átomos de hidrogênio.

#### 2 METODOLOGIA

# 2.1 DINÂMICA MOLECULAR

A dinâmica molecular (DM) é uma técnica computacional utilizada para descrever propriedades de equilíbrio e de transporte em sistemas físicos de muitos corpos a partir da descrição do movimento microscópico dos seus constituintes [62]. Na década de 1950 Alder e Wainwright introduziram o conceito de DM no estudo de sistemas moleculares aplicando o modelo de esferas rígidas e utilizando mecânica estatística para investigar as propriedades de líquidos simples [63]. A eficiência da DM se tornou maior com a substituição do modelo de esferas rígidas pelo modelo de potenciais contínuos para simular líquidos de argônio [64] e água [65]. A abordagem utilizando potenciais contínuos realísticos tornou o método mais abrangente, e a evolução dos procedimentos de simulação juntamente com a elevação da capacidade de processamento dos computadores permitiu aplicar a DM ao estudo de sistemas maiores e mais complexos. Atualmente diversos trabalhos utilizando DM têm sido realizados em áreas de pesquisa como a ciência dos materiais [66], a biofísica e a bioquímica molecular [67-68], no estudo de estruturas de proteínas [69], em nanotecnologia [70] e em física atômica e molecular [71-72].

Nas simulações de dinâmica molecular o primeiro passo é definir o tamanho da caixa de simulação e a composição do sistema simulado com as coordenadas e momentos lineares de cada uma das partículas. Com estas informações inicia-se um procedimento cíclico repetido primeiramente para minimizar a energia do sistema (fase de minimização) e posteriormente até que o sistema alcance o equilíbrio<sup>1</sup>. Cada ciclo consiste em: para uma dada configuração do sistema, obter-se o potencial total sobre cada partícula, a partir do potencial obtido determinar os valores de força que atuam sobre cada componente, a partir dos valores de força determinar os valores de aceleração e a partir destes as novas coordenadas e momentos de cada constituinte do sistema. Esta nova configuração obtida é a configuração de entrada para o próximo ciclo. Uma vez alcançadas às condições de equilíbrio desejadas, a simulação de dinâmica molecular propriamente dita, fase de produção, é realizada durante um tempo suficientemente longo (número de passos grande) para

<sup>&</sup>lt;sup>1</sup> Configuração inicial para a fase de produção que possui as condições desejadas (temperatura, pressão, etc.) em que a simulação deve ocorrer.

se obter um conjunto de configurações que represente adequadamente o sistema onde então as propriedades físicas e químicas são obtidas.

# 2.1.1 CAIXA DE SIMULAÇÃO E CONDIÇÕES DE CONTORNO PERIÓDICAS

Como o número de partículas de um sistema real é, em geral, maior que 10<sup>23</sup> e a capacidade de processamento atual dos computadores limitam as simulações a um número de partículas da ordem de 10<sup>6</sup>, para contornar o problema associado ao tamanho do sistema simulado aplicamos as condições de contorno periódicas (CCP). Este modelo consiste em definir uma caixa de simulação<sup>2</sup> finita com um número N de partículas e, a partir da translação desta caixa nas três dimensões espaciais, construir o sistema macroscópico ou infinito (ver na Figura 2 o caso bidimensional).



Figura 4. Representação esquemática de um sistema periódico bidimensional construído pela translação de réplicas da caixa de simulação (unidade central) a partir da aplicação das condições de contorno periódicas na simulação de dinâmica molecular.

<sup>&</sup>lt;sup>2</sup>Caixa de simulação é a região do espaço delimitada por um contorno geométrico, geralmente cúbico, dentro da qual estão as partículas do sistema simulado.

Desta forma mesmo simulando uma única caixa com número reduzido de partículas obtêm-se resultados correspondentes ao de um sistema físico substancialmente maior que o efetivamente simulado, garantindo a representatividade do sistema real nas simulações, minimizando os erros associados ao tamanho do sistema e reduzindo o tempo computacional das simulações uma vez que apenas as partículas da caixa de simulação são simuladas. Como, neste caso, o sistema é periódico as interações das partículas da caixa de simulação com as replicadas são realizadas fazendo a translação espacial das partículas para fora da caixa de simulação até o limite desejado na simulação. Outra vantagem deste método é garantir a conservação do número de partículas na caixa de simulação durante a DM (quando uma partícula entra/sai por uma das paredes da caixa de simulação outra equivalente a ela sai/entra pela parede oposta). Adicionalmente aplicar as CCP remove os efeitos de superfície (efeitos de borda ou efeitos de fronteira) que podem ser significativamente importantes e devem ser levados em consideração no caso de simulações de sistemas pequenos como, por exemplo, os sistemas nano-estruturados.

As CCP são uma aproximação bastante poderosa na descrição de sistemas físicos sólidos, em particular para sistemas cristalinos que possuem a simetria que as CCP carregam. Para sistemas fluidos, no entanto, esta periodicidade introduzida pelas CCP é um fator que introduz erro na obtenção das propriedades do sistema, uma vez que, cada partícula possui a mesma vizinhança em todas as caixas, situação que não ocorre em um fluido real devido à mobilidade de seus constituintes. Outro fator importante é o esforço e tempo computacional necessários para resolver as interações de um sistema muito grande como os que as CPP permitem gerar, assim, para reduzir o tempo computacional e a periodicidade introduzida pelas CCP utiliza-se juntamente com elas o método da mínima imagem. Este modelo introduz um raio de corte (Rc) menor ou igual à metade do valor da menor aresta da caixa de simulação (L) para definir a região de interação de uma partícula com a sua vizinhança. Desta forma cada partícula do sistema só interage com aquelas que

se encontram dentro da região delimitada pelo raio de corte e interage apenas uma vez com cada partícula ou imagem de partícula que se encontra em uma caixa de simulação vizinha, esta condição é atendida através do limite Rc < L/2 (ver Figura 4 e Figura 5).



Figura 5. Representação esquemática de uma caixa de simulação de lado L e das caixas de simulação adjacentes obtidas pela aplicação das condições de contorno periódicas. Em destaque o raio de corte para interação de curto alcance (círculo pontilhado) e as partículas da região de interação (sombreado amarelo) dentro da caixa simulada.



Figura 6. Representação esquemática do raio de corte aplicado para interações nas simulações de dinâmica molecular. Em destaque a condição de mínima imagem obtida para a circunferência de raio menor ou igual à metade da menor aresta da caixa de simulação (menor raio) na interação da partícula centrada na circunferência com as demais da região delimitada pelo raio de corte.

Uma descrição mais detalhada sobre CCP pode ser encontrada em [73-75]. A exclusão das interações de uma partícula com aquelas que se encontram além do raio de corte definido pelo método da mínima imagem pode incluir um erro por interromper de forma abrupta interações de maior alcance como, por exemplo, a interação de Coulomb. Para incluir na simulação estas interações de longo alcance de forma eficiente uma alternativa bastante utilizada é aplicar a soma de Ewald.

#### 2.1.2 A SOMA DE EWALD

A soma de Ewald é um método para computar a energia de interação de sistemas periódicos trocando a soma das energias no espaço real por uma soma equivalente no espaço recíproco (espaço de Fourier). A vantagem desta aproximação é a rápida convergência da série de somas para as interações de longo alcance. No caso da energia eletrostática, por exemplo, é mais eficiente do ponto de vista computacional decompô-la nas contribuições de curto alcance (somadas no espaço real) e nas contribuições de longo alcance (somadas no espaço de Fourier). Neste modelo o potencial de interação é escrito como uma soma de dois termos, um referente às interações de curto e outro referente às interações de longo alcance:

$$V_T(\mathbf{r}) = V_{ca}(\mathbf{r}) + V_{la}(\mathbf{r}). \tag{1}$$

Na expressão acima  $V_{ca}(\mathbf{r})$  representa o potencial de interação de curto alcance, somado no espaço real e  $V_{la}(\mathbf{r})$  é o potencial de interação de longo alcance somado no espaço de fase. Para a interação de Coulomb, a energia de interação de longo alcance,  $E_{LA}$ , entre cargas da célula central com as demais células da rede pode ser representada por uma integral dupla sobre duas densidades de carga que representam os campos da célula central e da rede periódica.

$$E_{LA} = \iint dr dr' \,\rho_{TOT}(\mathbf{r})\rho_{cs}(\mathbf{r}')V_{la}(\mathbf{r}-\mathbf{r}'),\tag{2}$$

onde  $\rho_{cs}$  é a densidade de cargas na caixa de simulação e é dada pela soma sobre todas as posições *r* das cargas *q* dentro da caixa,

$$\rho_{cs}(\mathbf{r}) = \sum_{cargas\,k} q_k \delta(\mathbf{r} - \mathbf{r}_k) , \qquad (3)$$

e  $\rho_{TOT}$  é a densidade total de cargas e é dada pela soma sobre todas as posições r das cargas q na caixa mais as imagens da rede:

$$\rho_{TOT}(\mathbf{r}) = \sum_{n_1, n_2, n_3} \sum_{\text{cargas } k} q_k \delta(\mathbf{r} - \mathbf{r_k} - n_1 \mathbf{a_1} - n_2 \mathbf{a_2} - n_3 \mathbf{a_3}).$$
(4)

 $\delta$  é a função delta de Dirac,  $a_1$ ,  $a_2$  e  $a_3$ , são os vetores da rede e  $n_1$ ,  $n_2$  e  $n_3$  são números inteiros. A densidade total de carga pode ser escrita como uma convolução<sup>3</sup> da função densidade com a função da rede. Desta forma podemos escrever a transformada de Fourier da função densidade total de carga como:

$$\tilde{\rho}_{TOT}(\boldsymbol{k}) = \tilde{L}\tilde{\rho}_{cs}(\boldsymbol{k}) , \qquad (5)$$

onde L é a função de rede, dada por:

$$L(\mathbf{r}) = \delta(\mathbf{r} - n_1 \mathbf{a_1} - n_2 \mathbf{a_2} - n_3 \mathbf{a_3})$$
(6)

e a transformada de Fourier de *L* é dada por:

$$\tilde{L}(\boldsymbol{k}) = \frac{(2\pi)^3}{\Omega} \sum_{m_1, m_2, m_3} \delta(\boldsymbol{r} - m_1 \boldsymbol{b_1} - m_2 \boldsymbol{b_2} - m_3 \boldsymbol{b_3}).$$
(7)

Nesta última expressão,  $\Omega$  é o volume da caixa central,  $\delta$  é a função delta de Dirac;  $\boldsymbol{b}_1$ ,  $\boldsymbol{b}_2$  e  $\boldsymbol{b}_3$ são os vetores da rede recíproca; e  $m_1$ ,  $m_2$  e  $m_3$  são números inteiros.

Podemos definir o potencial efetivo sobre uma partícula como:

$$\Phi(\mathbf{r}) = \int dr' \rho_{cs}(\mathbf{r}') \,\varphi_{la}(\mathbf{r}'). \tag{8}$$

Escrevendo desta forma, sua transformada de Fourier será dada pela expressão:

$$\widetilde{\Phi}(\mathbf{k}) = \int d\mathbf{r} \, \Phi(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} \tag{9}$$

e a energia de interação pode ser escrita como uma integral simples:

<sup>&</sup>lt;sup>3</sup> Convolução é a operação matemática linear que produz a partir de duas funções uma terceira função que representa a área encerrada pela superposição das duas funções iniciais em função do deslocamento entre elas.

$$E_{la} = \int d\mathbf{r} \,\rho_{TOT}(\mathbf{r}) \Phi(\mathbf{r}). \tag{10}$$

Usando o teorema de Parseval [76] chegamos à expressão final para a energia de interação de longo alcance:

$$E_{la} = \int \frac{d\boldsymbol{k}}{2\pi} \tilde{\rho}^*_{TOT}(\boldsymbol{k}) \tilde{V}(\boldsymbol{k}) = \int \frac{d\boldsymbol{k}}{2\pi} \tilde{L}^*(\boldsymbol{k}) |\tilde{\rho}_{cs}(\boldsymbol{k})|^2 \tilde{\Phi}(\boldsymbol{k})$$
(11)

$$E_{la} = \frac{1}{\Omega} \sum_{m1,m2,m3} |\tilde{\rho}_{cs}(\boldsymbol{k})|^2 \tilde{\Phi}(\boldsymbol{k}).$$
(12)

Uma descrição mais detalhada da soma de Ewald pode ser encontrada em [62, 77-78]. Para cada configuração da dinâmica molecular é necessário obter o potencial sobre cada partícula a fim de se obter a força sobre elas e assim os novos valores de coordenadas e momentos que constituem a nova configuração.

Uma vez obtida a configuração da caixa de simulação e o potencial que cada uma das partículas do sistema experimenta é necessário evoluir o sistema e determinar as novas sucessivas configurações de cada átomo integrando suas equações de movimento.

# 2.1.3 AS EQUAÇÕES DE MOVIMENTO

Para cada constituinte do sistema podemos escrever as sucessivas coordenadas e momentos se resolvermos as equações:

$$\boldsymbol{F}_{i} = -\nabla V_{i}. \tag{13}$$

onde,  $F_i$  é o valor da força resultante devido ao potencial total  $V_i$  de interação da partícula *i* com todas as partículas do sistema. Uma vez conhecida a força que atua sobre cada partícula e conhecendo a massa de cada uma delas podemos então obter a partir da segunda lei de Newton o valor da aceleração instantânea,  $a_i$ , de cada componente do sistema:

$$a_i = \frac{F_i}{m_i} \,. \tag{14}$$

A partir dos valores de aceleração podemos obter os novos valores de posição e velocidade integrando as equações:

$$a_i = \frac{dv_i}{dt} \tag{15}$$

e

$$\boldsymbol{\nu}_i = \frac{d\boldsymbol{R}_i}{dt}.$$
(16)

Nestas equações  $v_i \in R_i$  representam, respectivamente, a velocidade e a posição de cada partícula *i*.

# 2.1.4 A INTEGRAÇÃO DAS EQUAÇÕES DE MOVIMENTO

Uma série de algoritmos de integração pode ser utilizada para resolver as expressões acima. Neste trabalho utilizamos os algoritmos Steepest Descent e L-BFGS nos processos de minimização de energia e o algoritmo de Dinâmica Estocástica (SD – do inglês Stocastic Dynamics) para equilibração e produção no estudo dos poliol.

#### STEEPEST DESCENT

O método de integração Steepest Descent consiste em primeiramente calcular as forças e potenciais sobre cada constituinte do sistema, então as novas posições são determinadas através da equação:

$$\boldsymbol{R_{n+1}} = \boldsymbol{R_n} + \frac{\boldsymbol{F_n}}{\max(|\boldsymbol{F_n}|)} \boldsymbol{h_n} \quad , \tag{17}$$

onde,  $h_n$  é o afastamento máximo,  $F_n$  a força e max $(|F_n|)$  o maior valor absoluto de força. As forças e a energia são então novamente determinadas para as novas posições obedecendo às condições:

- Se  $V_{n+1} < V_n$  a posição é aceita e  $h_{n+1} = 1, 2h_n$
- Se  $V_{n+1} \ge V_n$  a posição é rejeitada e  $h_n = 0,2h_n$

Neste modelo, o algoritmo é interrompido se o número máximo de iterações é executado ou se o valor absoluto da força é menor que o parâmetro de convergência,  $\epsilon$ , aproximado pela raiz quadrada da força de oscilador, f, dada pela equação:

$$f = 2\pi\nu\sqrt{2mk_BT}.$$
(18)

Na equação (18),  $\nu$  é a freqüência de oscilador, m é a massa reduzida,  $k_B$  é a constante de Boltzmann e T é a temperatura absoluta do sistema.

# DINÂMICA ESTOCÁSTICA

Na dinâmica estocástica ou dinâmica de Langevin [76] um termo de fricção e um termo de ruído são adicionadas às equações de movimento de Newton que passa a ser representada pela equação:

$$m_i \frac{d^2 \mathbf{R}_i}{dt^2} = -m_i \varepsilon_i \frac{d\mathbf{R}}{dt} + \mathbf{F}_i(\mathbf{R}) + \dot{\mathbf{R}}_i , \qquad (19)$$

nela,  $\varepsilon_i$  é a constante de fricção e  $\dot{R}_i$  é o termo de ruído, onde:

$$\langle \dot{\mathbf{R}}_{i}(t)\dot{\mathbf{R}}_{j}(t+s)\rangle = 2m_{i}\varepsilon_{i}k_{B}T\delta(s)\delta_{ij}.$$
(20)

Uma descrição mais detalhada dos métodos de integração pode ser obtida em [75]. Uma vez integradas às equações de movimento, por um ou outro método, podemos evoluir o sistema temporalmente obtendo sucessivas configurações em equilíbrio e em acordo com as condições iniciais da simulação. Desta forma através de médias estatísticas o valor da grandeza física verificada experimentalmente pode ser obtido se o observável puder ser escrito em termos das coordenadas e momentos do sistema.

### 2.1.5 AS PROPRIEDADES TERMODINÂMICAS

ara obter o valor de um observável por dinâmica molecular primeiro é necessário expressá-lo em termos das coordenadas e dos momentos de cada partícula do sistema. Para a temperatura, por exemplo, no caso de sistemas clássicos de muitos corpos o teorema da equipartição é escrito como:

$$\langle \frac{1}{2}m\boldsymbol{v}_i^2 \rangle = \frac{1}{2}k_B T. \tag{21}$$

Desta forma podemos obter a temperatura instantânea do sistema pela expressão:

$$T(t) = \sum_{i=1}^{N} \frac{m_i v_i^2}{k_B N_f} \,.$$
(22)

Nela  $m_i \in v_i$  são, respectivamente, a massa e a velocidade da partícula *i*;  $k_B$  é a constante de Boltzmann e  $N_f$  é o número de graus de liberdade do sistema de N partículas com momento linear fixo.

Um importante fator na determinação das propriedades termodinâmicas de um sistema físico consiste na escolha do ensemble utilizado para descrevê-lo.

### ENSEMBLE CANÔNICO

Quando um sistema físico é colocado em contato com um reservatório térmico, a probabilidade de encontrar este sistema em um dado estado microscópico acessível a ele é dada pela expressão:

$$P_j = c\Omega_R (E_0 - E_j), \tag{23}$$

onde c é a constante de normalização,  $\Omega_R(E_0 - E_j)$  é o número de estados acessíveis ao reservatório, e  $E_0$ e  $E_j$  são a energia total e a energia do sistema, respectivamente. Para um

reservatório térmico infinito e  $E_0$  muito maior que  $E_j$ , podemos expandir  $lnP_j$  em torno de  $E_0$  até segunda ordem na forma:

$$lnP_{j} = lnc + ln\Omega_{R}(E) + \left[\frac{\partial ln_{R}(E)}{\partial E}\right]_{E=E_{j}}E .$$
(24)

Desta forma a expansão se reduz a:

$$lnP_j = constante - \frac{1}{k_B T} E_j, \tag{25}$$

ou seja,

$$P_{j} = \frac{\exp\left(-\frac{1}{k_{B}T}E_{j}\right)}{\sum_{k}\exp\left(-\frac{1}{k_{B}T}E_{k}\right)} \quad .$$
(26)

A função de partição, Z, é a soma sobre os microestados acessíveis de um sistema que normaliza a probabilidade  $P_j$ . Levando em consideração a degenerescência do sistema podemos escrevê-la como:

$$Z = \sum_{j} \exp\left(-\frac{1}{k_B T} E_j\right) = \sum_{E} \Omega_R(E) \exp\left(-\frac{1}{k_B T} E\right).$$
 (27)

A conexão com a termodinâmica é feita através da relação:

$$Z \to \exp\left(-\frac{1}{k_B T}F\right). \tag{28}$$

Substituindo a soma pelo seu termo máximo e utilizando a definição de entropia:

$$S(E) = k_B ln \Omega(E) , \qquad (29)$$

podemos escrever então:

$$F = -k_B T \ln Z \,. \tag{30}$$

onde *F* é a energia livre de Helmholtz,  $k_B$  é a constante de Boltzmann, *T* é a temperatura absoluta do sistema e *Z* a função de partição. Uma análise mais detalhada do ensemble canônico pode ser encontrada na referência [79].

#### ENSEMBLE DAS PRESSÕES

Outro ensemble importante e bastante utilizado em simulações de dinâmica molecular é o ensemble das pressões. Este ensemble está associado a um sistema em contato com um reservatório de forma que se mantenham fixos os valores de pressão e temperatura. Para um sistema com pressão e temperatura constantes podemos escrever a probabilidade de encontrá-lo em um estado microscópico j na forma:

$$P_{j} = c\Omega_{R} (E_{0} - E_{j}; V_{0} - V_{j}).$$
(31).

Expandindo o logaritmo da probabilidade  $P_j$  e usando a definição de entropia da segunda lei da termodinâmica, podemos reescrever esta expressão como:

$$\frac{\partial ln\Omega_R}{\partial E} = \frac{1}{k_B T}; \frac{\partial ln\Omega_R}{\partial V} = \frac{p}{k_B T}.$$
(32)

No limite de um reservatório muito grande podemos abandonar os termos de segunda ordem e de ordem maior que 2 ficando com a expressão:

$$lnP_j = constante - \frac{E_j}{k_B T} - \frac{pV_j}{k_B T} .$$
(33)

Assim podemos escrever que:

$$P_j = \frac{1}{Y} exp \left[ -\frac{E_j}{k_B T} - \frac{pV_j}{k_B T} \right], \qquad (34)$$
de onde obtemos a função de partição:

$$Y = \sum_{j} exp \left[ -\frac{E_j}{k_B T} - \frac{pV_j}{k_B T} \right]$$
(35)

A soma dos microestados da equação anterior pode ser feita em duas etapas, fixando os valores de volume e somando sobre todos os valores de volume. Desta forma o somatório será reescrito:

$$Y = \sum_{V} exp \left[ -\frac{pV}{k_B T} \right] \sum_{j} exp \left[ -\frac{E_j(V)}{k_B T} \right].$$
(36)

Aproximando esta soma pelo seu termo máximo, podemos escrever,

$$Y = \sum_{V} exp \left[ -\frac{pV}{k_{B}T} + lnZ \right] \sim exp \left[ -\frac{1}{k_{B}T} \min(-k_{B}T lnZ + pV) \right]$$
(37)

$$Y = \sim exp \left[ -\frac{1}{k_B T} \min(F + pV) \right] \rightarrow exp \left[ -\frac{1}{k_B T} G \right].$$
(38)

Através da energia livre de Gibbs, G, fazemos a conexão desse ensemble com a termodinâmica. Uma descrição mais detalhada do ensemble das pressões pode ser vista na referência [79].

### 2.1.6 ACOPLAMENTO DE TEMPERATURA E PRESSÃO

### TERMOSTATO DE BERENDSEN

Nas simulações de dinâmica molecular para manter a temperatura constante podemos escolher alguns tipos de termostato, um deles, o de Berendsen é um algoritmo que corrige a diferença de temperatura de forma lenta através da expressão, que aproxima a derivada da temperatura absoluta em relação ao tempo pela razão entre a diferença entre as temperaturas instantânea (T) e de referência ( $T_0$ ) e o tempo  $\tau$  característico:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \ . \tag{39}$$

Isto significa que o desvio em relação à temperatura de referência cai exponencialmente e o calor que entra ou sai do sistema para manter a temperatura constante é controlado a partir de um parâmetro  $\theta$  dependente do tempo, dado pela expressão:

$$\theta = \left[1 + \frac{n_{TC}\Delta t}{\tau_T} \left\{ \frac{T_0}{T\left(t - \frac{1}{2}\Delta t\right)} - 1 \right\} \right]^{\frac{1}{2}}$$
(40)

Na expressão acima,  $n_{TC}$  é o número de passos para reescalonamento dos valores de velocidade (fornecido pelo usuário) para ajuste na temperatura,  $\Delta T$  é a diferença entre a temperatura de referência e a temperatura obtida no passo de simulação,  $\tau_T$  é próximo, mas não exatamente a constante de acoplamento de temperatura ( $\tau$ ) dada pela expressão:

$$\tau = \frac{2C_V \tau_T}{N_{df} k_B} \quad , \tag{41}$$

onde,  $C_V$  é a capacidade térmica a volume constante,  $N_{df}$  é o número de graus de liberdade do sistema e  $k_B$  é a constante de Boltzmann. O termostato de Berendsen funciona muito bem para acoplamentos fracos de temperatura, caso dos processos da fase de equilibração do sistema, para a fase de produção que corresponde à simulação propriamente dita, precisamos utilizar um termostato que garanta de forma mais eficiente ensemble escolhido para a simulação. Este termostato, entretanto, não leva em consideração flutuações na energia cinética, isto significa que o ensemble gerado por ele não é rigorosamente canônico.

#### TERMOSTATO DE VELOCIDADE REESCALADA

O termostato de velocidade reescalada é semelhante ao termostato de Berendsen diferindo pela adição de um termo estocástico que corrige a distribuição de energia cinética e, portanto, é capaz de fornecer um ensemble realmente canônico. Neste termostato a correção na temperatura é feita a partir da expressão:

$$dK = (K_0 - K)\frac{dt}{\tau_T} + 2\sqrt{\frac{KK_0}{N_{df}}}\frac{dW}{\sqrt{\tau_T}} , \qquad (42)$$

onde, K é a energia cinética do sistema,  $N_{df}$  é o número de graus de liberdade do sistema e dW é um processo de Wiener<sup>4</sup>.

## BARÔMETRO DE BERENDSEN

De forma semelhante ao modelo utilizado para acoplamento de temperatura, o acoplamento de pressão no modelo de Berendsen corrige a diferença entre a pressão instantânea e a pressão de referência de forma exponencial através da expressão:

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_P} \,. \tag{43}$$

O reescalonamento das dimensões da caixa de simulação é dado pela matriz  $\mu$ , cujos elementos são dados por:

$$\mu_{ij} = \delta_{ij} - \frac{n_{PC}\Delta t}{3\tau_P} \beta_{ij} \{ P_{0ij} - P_{ij}(t) \}.$$
(44)

<sup>&</sup>lt;sup>4</sup> O processo de Wiener é um processo estocástico contínuo no tempo.

Na expressão acima,  $\delta_{ij}$  é a função delta de Dirac,  $n_{PC}$  é o número de graus de liberdade,  $\Delta t$  é o intervalo de tempo,  $\tau_P$  é o tempo característico  $P_{0i} \in P_{ij}(t)$  são a pressão de referência e a pressão instantânea.

## BARÔMETRO DE PARRINELLO-RAHMAN

Outra forma de fazer o acoplamento de pressão nas simulações de dinâmica molecular é utilizar o modelo de Parrinello-Rahman, que consiste em representar os vetores da caixa por uma matriz de elementos, b, dados pela expressão:

$$\frac{d^2b}{dt^2} = V W^{-1} b'^{-1} (P - P_{ref}).$$
(45)

Na equação acima, V é o volume da caixa de simulação, W é a matriz que determina a força do acoplamento e P e  $P_{ref}$  são as pressões instantânea e de referência. Nesta aproximação as equações de movimento são escritas como:

$$\frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i} - M \frac{dr_i}{dt} , \qquad (46)$$

onde:

$$M = b^{-1} \left[ b \frac{db'}{dx} + \frac{db}{dx} b' \right] {b'}^{-1} .$$
 (47)

Uma análise mais detalhada dos modelos de acoplamento de temperatura e pressão pode ser vista na referência [75, 78].

# 2.1.7 A ENERGIA LIVRE DE HIDRATAÇÃO ( $\Delta G$ )

## MÉTODO DO CRESCIMENTO LENTO

Para se obter computacionalmente a variação da energia livre de Gibbs de um processo utilizamos o método do crescimento lento, que consiste em adicionar um componente X ao sistema de forma gradativa incluindo o potencial de interação de seus constituintes com o sistema e determinar a diferença de energia livre entre dois estados A e B [80, 81]. Esta transição entre estados é feita utilizando um parâmetro lambda que varia lentamente entre 0 e 1, onde o estado A se refere ao sistema sem o componente X e o estado B ao sistema com o componente X. A Figura 7 ilustra a inclusão do componente X (triângulo azul) em um sistema (bolas vermelhas) variando o parâmetro  $\lambda$  entre 0 e 1.



Figura 7. Representação esquemática da inclusão de um elemento X (triângulo azul) em um meio solvente (bolas vermelhas) por meio da variação do parâmetro  $\lambda$  de 0 a 1.

## INTEGRAÇÃO TERMODINÂMICA

Para realizarmos a integração termodinâmica, lançamos mão de um parâmetro  $\lambda$  que acopla o estado final ao estado inicial de forma lenta, transformando o sistema de solvente simples em soluto solvatado. Para a energia livre de Helmholtz, *F*, podemos escrever:

$$F(\lambda) = -k_B T ln(Z), \tag{48}$$

onde Z é a função de partição dada por:

$$Z = c \iint \exp[-\beta H(p, q, \lambda)] \, dp dq. \tag{49}$$

Na equação acima, c é a constante de normalização:

$$c = (N! h^{3N})^{-1} (50)$$

e

$$\beta = \frac{1}{k_B T}.$$
(51)

Se tomarmos a derivada da energia livre de Helmholtz em relação ao parâmetro  $\lambda$ , vemos que ela corresponde a:

$$\frac{dF}{d\lambda} = \frac{\iint \left(\frac{\partial H}{\partial \lambda}\right) \exp[-\beta H(\boldsymbol{p}, \boldsymbol{q}, \lambda)] \, dp \, dq}{\iint \exp[-\beta H(\boldsymbol{p}, \boldsymbol{q}, \lambda)] \, dp \, dq} = \langle \frac{\partial H}{\partial \lambda} \rangle_{NVT;\lambda}$$
(52)

Integrando esta expressão em  $\lambda$  nos limites que levam o sistema do estado A para o estado B, ficamos com a variação da energia livre de Helmholtz escrita como:

$$F^{B}(V,T) - F^{A}(V,T) = \int_{0}^{1} \langle \frac{\partial H}{\partial \lambda} \rangle_{NVT;\lambda} \, d\lambda \,.$$
(53)

De forma análoga, podemos obter para o ensemble NPT, uma expressão equivalente para encontrarmos a variação da energia livre de Gibbs:

$$G^{B}(P,T) - G^{A}(P,T) = \int_{0}^{1} \langle \frac{\partial H}{\partial \lambda} \rangle_{NPT;\lambda} \, d\lambda.$$
(54)

A conexão entre as energias livres de Gibbs e Helmholtz é feita por meio da expressão:

$$G = F + PV. \tag{55}$$

Portanto,

$$G^{B}(p) - G^{A}(p) = F^{B}(V) - F^{A}(V) - \int_{p^{A}}^{p^{B}} [V^{B}(P') - V] dP'$$
(56)

Uma descrição mais detalhada do método de integração termodinâmica pode ser obtida nas referências [69, 80, 82-83].

#### 2.1.8 MODELO DE CAROÇO MACIO (SOFT-CORE)

Se nos cálculos de energia livre consideramos que uma partícula aparece no sistema (incluímos gradativamente os potenciais de interação de Coulomb e de Lennard-Jones dela com as demais partículas da caixa de simulação), quando a partícula começa a aparecer no sistema as interações são muito fracas e podem ocorrer flutuações nos valores de  $\frac{dv}{d\lambda}$ . Para contornar este problema precisamos remover esta singularidade, este procedimento pode ser feito modificando as interações de Coulomb e de Lennard-Jones por potenciais de soft-core dentro do limite de  $0 < \lambda < 1$ . No programa GROMACS o potencial que descreve a interação de soft-core é representado pela equação:

$$V_{sc} = (1 - \lambda)V^A(r_A) + \lambda V^B(r_B)$$
(57)

onde,

$$r_A = (\alpha \sigma_A^6 \lambda^p + r^6)^{\frac{1}{6}}$$
(58)

e

$$r_B = (\alpha \sigma_B^6 (1 - \lambda)^p + r^6)^{\frac{1}{6}} .$$
 (59)

Nas expressões acima,  $\alpha$  é o parâmetro de soft-core, p é a potência de lambda para softcore e  $\sigma$  é o raio de interação de soft-core. Para o potencial de soft-core a força pode ser escrita como:

$$F_{SC}(r) = -\frac{\partial V_{SC}(r)}{\partial r} = (1 - \lambda)F^A(r_A)\left(\frac{r}{r_A}\right)^5 + \lambda F^B(r_B)\left(\frac{r}{r_B}\right)^5$$
(60)

onde a contribuição para a derivada da energia livre é dada por:

$$\frac{\partial V_{SC}(r)}{\partial \lambda} = V^B(r_B) - V^A(r_A) + (1-\lambda)\frac{\partial V^A(r_A)}{\partial r_A}\frac{\partial r_A}{\partial \lambda} + \lambda \frac{\partial V^B(r_B)}{\partial r_B}\frac{\partial r_B}{\partial \lambda}$$
(61)

A forma dos potenciais que descrevem as interações ligadas e os potenciais de Coulomb e Lennard-Jonnes na aproximação de soft-core podem ser vistas na referência [75].

Apesar de ser uma poderosa ferramenta no estudo de sistemas moleculares a Dinâmica Molecular clássica não inclui efeitos da estrutura eletrônica sobre as propriedades investigadas neste trabalho. Neste sentido para incorporar tais efeitos à energia livre de hidratação lançamos mão do modelo contínuo polarizável combinado com o modelo de solvatação universal para descrição do solvente.

## 2.2 MODELOS CONTÍNUO POLARIZAVEL E DE SOLVATAÇÃO UNIVERSAL

Em razão do grande número de constituintes e da mobilidade do sistema, do ponto de vista computacional, torna-se inviável tratar quanticamente todos os componentes de um sistema de fase líquida. Para que se possa levar em consideração ao menos em parte os efeitos de estrutura eletrônica em sistemas líquidos a proposta dos modelos contínuos consiste em tratar o solvente como um material dielétrico contínuo caracterizado por uma constante dielétrica ɛ. Neste modelo, o soluto é descrito em nível quântico e as interações soluto-solvente são tratadas de forma aproximada apenas pelas interações de natureza eletrostática. Assim o sistema pode ser resolvido a partir da solução da equação de Poisson:

$$-\varepsilon \nabla^2 V_{(r)} = 4\pi \rho_{(r)} \tag{62}$$

Tomando dentro do solvente de constante dielétrica  $\varepsilon$ , uma cavidade ( $\varepsilon = 1$ ) que contém o soluto em estudo, tem-se duas regiões distintas onde a equação de Poisson deve ser resolvida, dentro e fora da cavidade, assim temos:

$$-\nabla^2 V_{(r)} = 4\pi \rho_{(r)}, dentro \ da \ cavidade \tag{63}$$

e

$$-\varepsilon \nabla^2 V_{(r)} = 0, for a \ da \ cavidade. \tag{64}$$

Nas expressões acima,  $\rho_{(r)}$  representa a distribuição de cargas,  $\varepsilon$  é a constante dielétrica do solvente e  $V_{(r)}$  é a soma dos potenciais gerado pela distribuição de cargas e do potencial de reação que surge devido à interação soluto-solvente. Para garantir a continuidade do potencial e a descontinuidade da componente perpendicular do campo elétrico na superfície da cavidade, temos que:

$$V_{dentro} - V_{fora} = 0 \tag{65}$$

e

$$\left(\frac{\partial V}{\partial n}\right)_{dentro} - \varepsilon \left(\frac{\partial V}{\partial n}\right)_{fora} = 0 , \qquad (66)^5$$

onde n é um vetor que aponta para fora da superfície da cavidade e o campo elétrico é representado pelo gradiente do potencial. A distribuição de cargas do soluto induz no solvente uma polarização que dá origem a um campo elétrico que atuará sobre o soluto. Esse campo de reação possui dependência direta com o momento de dipolo do soluto, com a constante dielétrica do solvente e com o tamanho e forma da cavidade [83,84].

O Modelo Contínuo Polarizável (PCM) é um dos métodos de carga aparente na superfície (ASC) [84], em que se define uma densidade superfícial de carga,  $\sigma(\mathbf{r}')$ , distribuída sobre a superfície da cavidade que dará origem ao potencial elétrico oriundo do campo de reação. Se  $\sigma(\mathbf{r}')$ for definida a partir de uma equação eletrostática apropriada, o valor encontrado para o potencial de reação será exato. O potencial eletrostático de uma distribuição de cargas na cavidade é dado pela equação:

$$V_{\sigma}(\mathbf{r}) = \int_{s} \frac{\sigma(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^{2}r, \qquad (67)$$

onde a integral é calculada sobre toda a superfície da cavidade, representada por S. A grande simplificação trazida pelos métodos ASC consiste em reduzir a fonte do potencial de reação a uma limitada distribuição de cargas na superfície da cavidade. Contudo, o cálculo desta integral nem sempre é simples, de modo que na maioria dos casos utiliza-se processos de discretização da integral em um número finito de elementos [84, 85] que divide a superfície da cavidade em pedaços suficientemente pequenos de modo que  $\sigma(r'_i)$  seja constante, que a integral possa ser escrita em termos de um somatório e o potencial seja escrito como:

$$V_{\sigma}(\mathbf{r}) \approx \sum_{i} \frac{\sigma(\mathbf{r}'_{i})A_{i}}{|\mathbf{r} - \mathbf{r}'_{i}|} = \sum_{i} \frac{q_{i}}{|\mathbf{r} - \mathbf{r}'_{i}|} \quad ,$$
(68)

Na expressão acima  $A_i$  e  $q_i$  representam, respectivamente, a área e a carga de cada elemento da superfície da cavidade. A partir do campo de reação é possível calcular a energia livre associada ao processo de solvatação.

<sup>&</sup>lt;sup>5</sup> Equação para a condição de cavidade de  $\varepsilon = 1$  e o meio externo de constante finita e maior que a da cavidade.

Na aproximação do modelo de solvatação universal [86] a energia livre de solvatação padrão,  $\Delta G$ , que corresponde a energia para transferir um sistema de fase gasosa para o estado em solução, é escrita como a soma de três contribuições conforme a expressão:

$$\Delta G = \Delta G_{ENP} + \Delta G_{CDS} + \Delta G_{CONC}. \tag{69}$$

Na equação acima, o primeiro termo corresponde as contribuições eletrônica, nuclear e de polarização ( $\Delta G_{ENP}$ ). Como o termo de relaxação nuclear é igual a diferença entre a energia total da fase de vácuo calculada com a geometria de equilíbrio de fase gasosa e a energia total calculada de fase gasosa para a geometria de equilíbrio em solução, se aproximarmos as geometrias de fase gasosa e em solução como a mesma, a contribuição  $\Delta G_{ENP}$  torna-se apenas a contribuição eletrônica e de polarização  $\Delta G_{EP}$ . A contribuição de polarização é determinada por meio de cálculos dos orbitais moleculares de campo de reação autoconsistentes, onde os campos reativos são definidos por soluções de equações de Poisson não lineares. Neste caso a equação não linear pode ser resolvida com uma descrição quântica do soluto para se obter o potencial elétrico devido ao meio dielétrico contínuo e a polarização do soluto. Este potencial,  $\varphi$ , chamado potencial de campo de reação é igual ao potencial total ( $\Phi$ ) menos o potencial eletrostático da molécula de fase gasosa,  $\Phi^{(0)}$ . A partir do potencial de reação calcula-se a mudança na energia livre que corresponde ao processo de solvatação. Se considerarmos o soluto rígido a contribuição ENP para a energia livre torna-se apenas eletrostática e de polarização (EN) dada pela equação:

$$\Delta G_{EP} = \langle \Psi | H^{(0)} - \frac{e}{2} \varphi | \Psi \rangle + \frac{e}{2} \sum_{k} Z_k \varphi_k - \langle \Psi^{(0)} | H^{(0)} | \Psi^{(0)} \rangle.$$
<sup>(70)</sup>

O segundo termo da equação 69,  $\Delta G_{CDS}$ , leva em consideração a mudança de energia livre associadas a cavidade do solvente, a mudanças de dispersão e a possíveis mudanças locais na estrutura do solvente. A contribuição  $\Delta G_{CDS}$  para a energia livre é dada pela equação:

$$\Delta G_{CDS} = \sum_{k}^{\text{átomos}} \sigma_k A_k(\boldsymbol{R}, \{R_{Z_k} + r_S\} + \sigma^{[M]} \sum_{k}^{\text{átomos}} A_k(\boldsymbol{R}, \{R_{Z_k} + r_S\},$$
(71)

onde  $\sigma_k$  e  $\sigma^{[M]}$  são a tensão superficial atômica do átomo k e a tensão superficial molecular, respectivamente,  $A_k$  é a superficie acessível ao solvente do átomo k. Este último termo depende da geometria  $\mathbf{R}$ , do conjunto  $\{R_{Z_k}\}$  de todos os raios de Van der Walls e do raio do solvente ( $r_S$ ) que é adicionado a cada átomo. Esse modelo pode ser aplicado para qualquer solvente com relevantes

descrições macroscópicas conhecidas ou estimadas. A água é tratada como um caso especial de solvente que tem seu próprio conjunto de coeficientes de tensão superficial. Para a água a tensão superficial molecular,  $\sigma^{[M]}$ , é tomada como zero, assim quando o modelo SMD é utilizado para calcular a energia livre de hidratação a equação 71 se reduz ao primeiro termo.

Por fim, o último termo da equação 69,  $\Delta G_{CONC}$ , leva em consideração as mudanças de concentração entre a fase gasosa e a fase líquida. Se as concentrações são as mesmas nos dois estados  $\Delta G_{CONC} = 0$ . A descrição mais detalhada do modelo contínuo polarizável e do modelo de solvatação universal pode ser encontrada em [86-88].

# 2.3 O MÉTODO DE HARTREE-FOCK

A aproximação de Hartree-Fock é o ponto de partida padrão para o tratamento de propriedades moleculares com métodos de função de onda mais sofisticados. Como os efeitos de correlação eletrônica não são considerados nesta aproximação, para muitas propriedades, os cálculos HF estão, em geral, em concordância qualitativa com medidas experimentais.

A equação de Schrödinger independente do tempo que descreve um sistema quântico é:

$$H\phi = E\phi , \qquad (72)$$

onde H é o Hamiltoniano total do sistema,  $\phi$  é a função de onda total e E é a energia total do sistema. Para um sistema molecular constituído por M núcleos atômicos e N elétrons o Hamiltoniano não relativístico total, escrito em unidades atômicas (u.a.), é dado por:

$$H = -\sum_{i=1}^{N} \frac{1}{2} \nabla_{i}^{2} - \sum_{A=1}^{M} \frac{1}{2M_{A}} \nabla_{A}^{2} - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_{A}}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_{A}Z_{B}}{R_{AB}},$$
(73)

nele  $\nabla_i^2 e \nabla_A^2$  são, respectivamente, operadores diferenciais nas coordenadas do elétron *i* e do núcleo *A*; *Z*<sub>A</sub>é o número atômico do núcleo *A*; *M*<sub>A</sub>é a massa do núcleo *A*; *r*<sub>*ij*</sub> é a distância entre os elétrons *i* e *j*; *r*<sub>*i*A</sub> é a distância entre o elétron *i* e o núcleo *A*; e *R*<sub>AB</sub> é a distância entre os núcleos *A* e *B*. Na expressão acima o primeiro e o segundo termos se referem, respectivamente, à energia cinética dos elétrons e à energia cinética dos núcleos atômicos; o terceiro termo se refere à energia potencial de interação atrativa entre núcleos atômicos e elétrons, o quarto termo se refere à energia potencial de interação repulsiva entre os elétrons e o último termo se refere a energia potencial de interação repulsiva entre os núcleos atômicos.

Na aproximação de Born-Oppenheimer supõe-se que o movimento dos núcleos atômicos é pequeno quando comparados com o movimento dos elétrons em razão da diferença de massa entre prótons e elétrons ( $m_{próton} \sim 10^3 m_{elétron}$ ). Assumindo a condição estática para os núcleos atômicos, podemos separar o movimento nuclear do movimento eletrônico e tratar o sistema molecular como uma rede de núcleos fixos onde os elétrons se movimentam. Analisando desta forma, o Hamiltoniano total do sistema pode ser separado em dois termos: um que trata da energia cinética dos elétrons, da interação elétron-elétron e da interação dos elétrons com os núcleos atômicos (Hamiltoniano eletrônico); e outro que trata da energia cinética dos núcleos atômicos e da interação entre os núcleos atômicos (Hamiltoniano nuclear). Dentro da aproximação de Born-Oppenheimer como os núcleos estão fixos a contribuição do segundo termo do Hamiltoniano total contribui com uma constante aditiva na energia considerada no cálculo da energia total e não tendo efeito sobre as autofunções do sistema. Desta forma, partindo da aproximação de Born-Oppenheimer podemos escrever a equação de Schrödinger de um sistema multieletrônico em termos somente das coordenadas dos elétrons na forma:

$$H_{ele}\phi_{ele} = E_{ele}\phi_{ele} , \qquad (74)$$

onde  $H_{ele}$  é o Hamiltoniano eletrônico, em unidades atômicas, que descreve o movimento de N elétrons em um campo médio produzido por Mnúcleos fixos,

$$H_{ele} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{r_{ij}}.$$
 (75)

Nele o primeiro termo se refere à energia cinética dos elétrons, o segundo refere-se à interação elétron-núcleo e o terceiro à interação elétron-elétron;  $\phi_{ele}$  é a função de onda que depende somente das coordenadas (espaciais e de spin) dos elétrons; e  $E_{ele}$  é a energia eletrônica do sistema.

Na teoria de Hartree-Fock a função de onda,  $\phi_{ele}$ , é aproximada pelo produto das funções de onda de um elétron, escrita na forma de um único determinante de Slater com *N* spin-orbitais<sup>6</sup>:

<sup>&</sup>lt;sup>6</sup> Spin-orbital é a função de onda de um elétron que descreve tanto a distribuição espacial quanto seu estado de spin.

$$\phi_{ele}(\mathbf{x_1}, \mathbf{x_2}, \dots, \mathbf{x_n}) = (N!)^{-1/2} \begin{vmatrix} \varphi_i(\mathbf{x_1}) & \varphi_j(\mathbf{x_1}) & \cdots & \varphi_k(\mathbf{x_1}) \\ \varphi_i(\mathbf{x_2}) & \varphi_j(\mathbf{x_2}) & \cdots & \varphi_k(\mathbf{x_2}) \\ \vdots & \vdots & & \vdots \\ \varphi_i(\mathbf{x_n}) & \varphi_j(\mathbf{x_n}) & \cdots & \varphi_k(\mathbf{x_n}) \end{vmatrix},$$
(76)

nele as funções  $\varphi_i(x)$  são chamadas spin-orbitais moleculares, x denota as coordenadas espaciais e de spin e o termo  $(N!)^{-1/2}$  é um fator de normalização. Escrita desta forma, a função de onda total do sistema multieletrônico é antissimétrica com respeito à troca de coordenadas de dois elétrons, obedecendo, portanto, o princípio de exclusão de Pauli<sup>7</sup>.

A energia total do sistema é escrita como a soma da energia eletrônica e da energia de repulsão nuclear:

$$E = E_{ele} + E_{nuc} \,. \tag{77}$$

As soluções aproximadas para a energia eletrônica são obtidas, no método de Hartree-Fock, por meio do princípio variacional. Este princípio baseia-se no fato de que o valor esperado da energia,

$$E = \langle \phi | H | \phi \rangle, \tag{78}$$

calculado com uma função tentativa  $\phi$  é sempre maior ou igual ao valor da energia do estado fundamental (*E*<sub>0</sub>). Desta forma, a melhor função de onda é aquela que minimiza o valor esperado da energia eletrônica. Uma análise mais detalhada do princípio variacional pode ser vista nas referências [89,90]

Em linhas gerais o método de Hartree-Fock consiste na conversão de uma equação de Schrödinger de um sistema de N elétrons em N equações de um elétron impondo o vínculo de ortonormalidade dos spin-orbitais moleculares, de modo que suas partes espaciais (orbitais moleculares,  $\Psi$ ) satisfaçam a equação de Hartree-Fock, escrita na forma restrita e canônica como:

$$f\Psi = \varepsilon\Psi.$$
(79)

Nesta equação, f é o operador de Fock, escrito para sistemas de camada fechada<sup>8</sup> como:

<sup>&</sup>lt;sup>7</sup>O princípio de Pauli afirma que dois férmions idênticos não podem ocupar o mesmo estado quântico simultaneamente, desta forma, a função de onda total de um sistema multieletrônico deve ser antissimétrica com relação à troca de coordenadas de dois elétrons.

<sup>&</sup>lt;sup>8</sup> Sistemas de camada fechada são aqueles onde cada orbital molecular está ocupado por dois elétrons com spins opostos.

$$f = h + \sum_{j=1}^{N/2} 2J_j - K_j;$$
(80)

$$J_{i}\psi_{j}(\boldsymbol{r_{1}}) = \left[\int d\boldsymbol{r}\,\psi_{i}^{*}(\boldsymbol{r_{2}})r_{12}^{-1}\psi_{i}(\boldsymbol{r_{2}})\right]\psi_{j}(\boldsymbol{r_{1}})$$
 (81)

$$K_{i}\psi_{j}(\boldsymbol{r_{1}}) = \left[\int d\boldsymbol{r}\,\psi_{i}^{*}(\boldsymbol{r_{2}})r_{12}^{-1}\psi_{j}(\boldsymbol{r_{2}})\right]\psi_{i}(\boldsymbol{r_{1}}),\tag{82}$$

onde h é o operador Hamiltoniano de caroço que contém o termo de energia cinética dos elétrons e o termo de interação elétron-núcleo. J e K são, respectivamente, os operadores de Coloumb e de troca. O primeiro operador se refere à interação coloumbiana entre os elétrons9, enquanto que o segundo operador leva em conta a indistinguibilidade dos elétrons. Como os operadores J e K dependem dos orbitais moleculares, a equação de Hartree-Fock pode ser resolvida de forma iterativa. A solução iterativa da equação de Hartree-Fock fornece os orbitais moleculares que melhor descrevem o sistema. Com eles é possível obter os spin-orbitais, e posteriormente, a função de onda total na forma de um determinante de Slater. Neste procedimento, um conjunto de spinorbitais teste é utilizado para formar o operador de Fock. Este é utilizado para resolver as equações de HF e obter um novo conjunto de spin-orbitais que são utilizados por sua vez para construir um novo operador de Fock corrigido. O ciclo de cálculo é reformulado e repetido até que se obtenha o critério de convergência previamente estabelecido<sup>10</sup>. A solução da equação de Hartree-Fock fornece um conjunto de spin-orbitais Hartree-Fock ortonormais  $\{\varphi_i\}$  com energias orbitais  $\{\varepsilon_i\}$ . O determinante de Slater formado pelos N spin-orbitais de menor energia, denominados de spinorbitais ocupados e identificados como  $\varphi_a, \varphi_b, (...)$ , representa a função de onda de Hartree-Fock com a melhor aproximação variacional para o estado fundamental do sistema. Os outros spinorbitais restantes pertencentes ao conjunto são chamados spin-orbitais virtuais e são identificados por  $\varphi_r, \varphi_s, (...)$ . Os estados indicados por índices superiores e inferiores, indicam a substituição dos spin-orbitais ocupados pelos spin-orbitais virtuais. Portanto, indica-se um estado de simples excitação por  $|\Psi_a^r\rangle$ , correspondendo a um determinante em que o spin orbital ocupado  $\varphi_i$  foi substituído pelo spin-orbital virtual  $\varphi_r$ , e um estado de dupla excitação por  $|\Psi_{ab}^{rs}\rangle$ , um determinante onde os spin-orbitais ocupados  $\varphi_a$  e  $\varphi_b$  foram substituídos pelos spin-orbitais virtuais  $\varphi_r$  e  $\varphi_s$  e

48

<sup>&</sup>lt;sup>9</sup> Tratada como a interação de um elétron com o potencial local médio gerado pelos demais elétrons.

<sup>&</sup>lt;sup>10</sup> Os critérios de convergência utilizados são a energia e a matriz densidade.

assim sucessivamente. No entanto, para sistemas moleculares a solução numérica torna-se inviável do ponto de vista computacional. Nestes sistemas com grande número de elétrons uma solução viável para a equação de Hartree-Fock, proposta por Roothaan em 1951 [91], consiste em escrever os orbitais em termos de um conjunto de funções base ( $g_n$ ) conhecidas:

$$\Psi_i(\mathbf{r}) = \sum_n C_{ni} g_n(\mathbf{r}) , \qquad (84)$$

onde os  $C_{ni}$ 's são os coeficientes da expansão. Nos cálculos HF as funções mais utilizadas são as funções do tipo Gaussianas. A substituição de (84) na equação (79) resulta na equação de Hartree-Fock-Roothaan, escrita na forma matricial como:

$$FC = SCE, \tag{85}$$

onde F é a matriz que representa o operador de Fock, cujos elementos são dados por:

$$F_{ij} = \langle g_i(\mathbf{r}) | f(\mathbf{r}) | g_j(\mathbf{r}) \rangle, \tag{86}$$

Cé a matriz dos coeficientes de expansão, Sé uma matriz de superposição:

$$S_{ij} = \langle g_i(\mathbf{r}) | g_j(\mathbf{r}) \rangle, \tag{87}$$

e *E* é a matriz das energias orbitais.

Na aproximação de HF, a energia total  $(E_{tot})$  é obtida a partir da expressão:

$$E_{tot} = E_0 + \sum_A \sum_{B>A} \frac{Z_A Z_B}{R_{AB}},$$
 (88)

onde

$$E_0 = 2\sum_{a}^{N/2} h_{aa} + \sum_{a}^{N/2} \sum_{b}^{N/2} (2J_{ab} - K_{ab}), \qquad (89)$$

onde  $h_{aa}$  é a energia cinética média e a energia de atração nuclear de um elétron descrito pela função de onda  $\phi_a$ :

$$h_{aa} = \int d\boldsymbol{r_1} \, \phi_a^*(\boldsymbol{r_1}) \left( -\frac{1}{2} \nabla^2 - \sum_{A} \frac{Z_A}{r_{1A}} \right) \phi_a(\boldsymbol{r_1}) \tag{90}$$

E os operadores  $J_{ab}$  e  $K_{ab}$  são, respectivamente, as integrais de Coulomb:

$$J_{ab} = \int d\mathbf{r_1} d\mathbf{r_2} \phi_a^*(\mathbf{r_1}) \phi_a(\mathbf{r_1}) \frac{1}{r_{12}} \phi_b^*(\mathbf{r_2}) \phi_b(\mathbf{r_2})$$
(91)

e de troca:

$$K_{ab} = \int d\mathbf{r_1} d\mathbf{r_2} \phi_a^*(\mathbf{r_1}) \phi_b(\mathbf{r_1}) \frac{1}{r_{12}} \phi_b^*(\mathbf{r_2}) \phi_a(\mathbf{r_2})$$
(92)

para dois elétrons descritos por funções de onda  $\phi_a e \phi_b$ . Uma descrição mais detalhada do método de Hartree-Fock pode ser encontrada em [89, 90]

A representação da função de onda por um único determinante de Slater é uma limitação da aproximação de HF. Como neste modelo os elétrons estão sujeitos a um campo local médio, ou seja, o movimento dos elétrons não é correlacionado, a energia HF encontrada possui um erro intrínseco chamado erro de correlação. A energia de correlação ( $E_{corr}$ ) é definida como a diferença entre a energia exata não relativística e a energia limite de HF [90]:

$$E_{corr} = E - E_{HF} \,. \tag{93}$$

## 2.4 A TEORIA DE PERTURBAÇÃO DE MØLLER-PLESSET

A ideia central por trás das teorias perturbativas consiste em escrever o Hamiltoniano completo (*H*) de um sistema como a soma de um Hamiltoniano de referência (*H*<sub>0</sub>) - não perturbado - mais uma perturbação ( $\Delta H$ ),

$$H = H_0 + \lambda \Delta H, \tag{94}$$

onde  $\lambda$  é um parâmetro associado à intensidade da perturbação. Desta forma parte-se da solução já conhecida do Hamiltoniano não perturbado para encontrar as correções para o sistema perturbado. Para construir as correções geradas por  $\Delta H$ , os auto-estados e as auto-energias são escritos como expansões em potências de  $\lambda$ :

$$|\Psi_n\rangle = \sum_m \lambda_n^m |\Psi_n^m\rangle \ e \ E_n = \sum_m \lambda_n^m |E_n\rangle.$$
(95)

Substituindo as expansões do auto-estado e da auto-energia na equação de Schrödinger e comparando os termos de mesma ordem em  $\lambda$ , podemos encontrar as correções na energia e nas autofunções do sistema perturbado. Em particular, as correções de primeira e segunda ordem na energia são dadas por [92]:

$$E_n^{(1)} = \langle \phi_n | \Delta H | \phi_n \rangle \qquad e \qquad E_n^{(2)} = \sum_{m \neq n} \frac{|\langle \phi_n | \Delta H | \phi_m \rangle|^2}{E_n^0 - E_m^0}.$$
 (96)

Na teoria de Møller-Plesset, o Hamiltoniano de ordem zero (não perturbado) é dado pela soma dos operadores de Fock de um elétron:

$$H_0 = \sum_i f(i). \tag{97}$$

Com essa escolha para  $H_0$ , a perturbação  $\Delta H$  é escrita na forma,

$$\Delta H = H - H_0 = \left(\sum_i h(i) + \sum_{i < j} \frac{1}{r_{ij}}\right) - \left(\sum_i f(i)\right).$$
(98)

A energia HF,  $E_{HF}$ , associada à função de onda HF do estado fundamental é dada pelo valor esperado:

$$E_{HF} = \langle \phi_0 | H | \phi_0 \rangle$$
$$= \langle \phi_0 | H_0 + \Delta H | \phi_0 \rangle = \langle \phi_0 | H_0 | \phi_0 \rangle + \langle \phi_0 | \Delta H | \phi_0 \rangle = E_0^{(0)} + E_0^{(1)}, \tag{99}$$

onde  $E_0^{(0)}$  e  $E_0^{(1)}$ , são respectivamente a energia de ordem zero e a correção em primeira ordem da teoria de perturbação. Desta forma a energia de correlação é obtida a partir da correção de segunda ordem, no método MP2, escrita como:

$$E_{corr} = \sum_{\substack{a < b \\ r < s}} \frac{|\langle \phi_0 | \Delta H | \phi_{ab}^{r_s} \rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s},$$
(100)

onde a e r são índices dos spin-orbitais que são funções das coordenadas do elétron 1 enquanto b e s são índices dos spin-orbitais dependentes das coordenadas do elétron 2. A descrição mais detalhada do método MP2 pode ser encontrada em [89, 90].

#### 2.5 O PROCEDIMENTO DE CAMPO FINITO

O procedimento de campo finito baseia-se na perturbação da energia eletrônica total de um sistema multieletrônico usando pequenos valores de campo elétrico externo, desta forma, a energia eletrônica total pode ser escrita em termos do momento de dipolo ( $\mu$ ), da polarizabilidade linear ( $\alpha$ ), da primeira hiperpolarizabilidade ( $\beta$ ) por meio de uma série de potências na intensidade do campo [93],

$$E = E_0 - \sum \mu_i F_i - \frac{1}{2!} \sum \alpha_{ij} F_{ij} - \frac{1}{3!} \sum \beta_{ijk} F_{ijk} - \dots , \qquad (101)$$

onde os coeficientes da expansão são as componentes destas propriedades elétricas e podem ser obtidos por meio, por exemplo, do ajuste de mínimos quadrados. Na expansão acima *F* representa o campo elétrico externo aplicado, *Eo* representa a energia da molécula na ausência de campo,  $\mu_i$ representa o momento de dipolo permanente,  $\alpha_{ij}$  e  $\beta_{ijk}$  representam, respectivamente, as componentes da polarizabilidade linear e da primeira hiperpolarizabilidade.

#### 2.6 DETALHES COMPUTACIONAIS

Para estudar a variação da energia livre de hidratação por dinâmica molecular, primeiramente, foram construídas caixas de simulação cúbicas preenchidas com 1 molécula de soluto para, aproximadamente, 1000 moléculas de solvente e aplicadas as condições de contorno periódicas utilizando o método de mínima imagem.

Neste trabalho a fase de produção das configurações para obtenção da variação da energia livre de hidratação no estudo das moléculas de poliol, glicina e peptídeos A<sub>N</sub>K foi precedida de duas fases de minimização de energia (utilizando os algoritmos Steepest Descent e L-BFGS para

integração das equações de movimento); e de duas fases de equilíbrio para estabilização de temperatura e pressão em ensemble NVT e NPT (utilizando o algoritmo dinâmica estocástica). A fase de produção foi realizada em ensemble NPT e, assim como as fases anteriores de equilibração, as equações de movimento foram integradas com algoritmo de dinâmica estocástica. Para cada simulação, o sistema foi equilibrado por 1 ns seguido por uma produção de 5 ns. As propriedades apresentadas foram calculadas de simulações de dinâmica molecular considerando passos de simulação de 2 fs com aquisição de dados a cada 0.01 ps. Todas as simulações foram realizadas com valores de temperatura de 300 K e pressão de 1 atm. O controle de temperatura foi realizado utilizando o termostato de velocidade reescalada e o controle de pressão com o barômetro de Parrinello-Rahman com constantes de acoplamento de 0,1 e 1,0, respectivamente. Todos os comprimentos de ligação foram restringidos utilizando o algoritmo LINCS, como implementado no GROMACS. As interações de Coulomb foram tratadas utilizando o modelo PME com o raio de corte de 1,2 nm

Neste trabalho a variação na energia livre de hidratação foi obtida utilizando dinâmica molecular estocástica associada com a integração termodinâmica pela inserção do soluto no meio solvente variando um parâmetro lambda de 0 a 1, correspondentes aos estados não acoplado: de simulação sem o soluto e acoplado: com o soluto completamente interagente. O acoplamento das interações de Van der Walls e Coulomb pode levar a instabilidades durante o processo de integração termodinâmica. Para remover as singularidades deste processo foi utilizado o modelo de caroço macio (soft-core) para as interações de Lennard-Jones. Aqui a inserção da molécula de soluto na solução foi realizada primeiramente com a inclusão das interações de Van der Walls variando o parâmetro  $\lambda$  no intervalo de 0 a 1 com incrementos de 0,04. Para incluir as interações de Coulomb entre soluto e solvente o parâmetro  $\lambda$  foi variando: de 0 a 0.20 com intervalos de 0.04; de 0,22 a 0,38 com intervalos de 0,02; e de 0,40 a 1 com intervalos de 0,04. Estes intervalos utilizados garantem a suavidade na curva da derivada  $\frac{\partial H}{\partial \lambda}$ , posteriormente integrada nos limites de lambda para obter a variação da energia livre de Gibbs entre os estados A (sem soluto) e B (com soluto). A variação da energia livre total de hidratação foi determinada pela soma das energias obtidas nestes dois procedimentos: inclusão das interações de VdW e inclusão das interações de Coulomb. Para as interações de Van der Waals o modelo de soft-core foi utilizado com os parâmetros  $\alpha = 0.5$ ;  $\sigma = 1.0$  e  $\beta = 0.3$ . As interações intra e intermoleculares de poliol, glicina e peptídeos A<sub>N</sub>K foram descritas utilizando o campo de força CHARMM e o potencial TIP3P foi utilizado para descrever a água. Para descrever as moléculas de poliol foi utilizada a parametrização proposta por Hatcher e colaboradores [94]. Um modelo de arquivo com os parâmetros da simulação para as moléculas de poliol e a parametrização utilizada para descrever este sistema são apresentados nos apêndices A e B, respectivamente. Todas as simulações de dinâmica molecular foram realizadas com o programa GROMACS 4.5.4 [95, 96].

Para acessar as propriedades elétricas e a variação da energia livre de hidratação com o modelo SMD-PCM, primeiramente, foram determinadas as geometrias de equilíbrio no vácuo e em meio solvente para os sistemas em estudo. Neste trabalho as moléculas de poliol foram otimizadas de forma completa utilizando a teoria de perturbação de Møller-Plesset de segunda ordem com o conjunto de funções base 6-311+G(2d) em vácuo e em água. Para as moléculas de glicina as otimizações de geometria foram realizadas utilizando a teoria de perturbação de Møller-Plesset e o funcional B3LYP com conjunto de funções base 6-311+G(2d) e para os peptídeos A<sub>N</sub>K foram realizadas com o funcional B3LYP e conjunto de funções base cc-pVTZ. Para todos os sistemas a dependência da geometria do soluto com o solvente foi obtida aplicando o modelo continuo polarizável com formalismo de equações integrais (IEF-PCM).

Os valores de momento de dipolo e polarizabilidade linear foram obtidos por cálculos numéricos de campo finito, com comprimento de campo de 0,001 u.a., como implementado no programa Gaussian09. Neste trabalho apresentamos o valor médio do momento de dipolo,  $\mu$ , e da polarizabilidade linear,  $\alpha$  que são definidos como:

$$\mu = \sqrt{\mu_x^2 + \mu_y^2 + \mu_z^2}$$
 ( 102 )

e

$$\alpha = \frac{\alpha_{xx} + \alpha_{yy} + \alpha_{zz}}{3} \tag{103}$$

onde $\mu_x$ ,  $\mu_y$  e  $\mu_z$  são as componentes longitudinais do momento permanente de dipolo e  $\alpha_{xx}$ ,  $\alpha_{yy}$  e  $\alpha_{zz}$  são as componentes longitudinais da polarizabilidade linear médias cujas expressões podem ser encontradas na referência [97]. A energia livre de hidratação foi obtida nos níveis MP2/6-

311+G(2d) (para poliol e glicina) e B3LYP/6-311+G(2d) (para glicina e peptídeos A<sub>N</sub>K) utilizando PCM combinado com o modelo de solvatação SMD. Todos os cálculos quânticos foram realizados com o programa GAUSSIAN09 [98].

#### **3 RESULTADOS E DISCUSSÕES**

## 3.1 POLIOL

A Tabela 1 apresenta a comparação entre os resultados de momento de dipolo permanente ( $\mu$ ) e polarizabilidade linear ( $\alpha$ ) obtidos em vácuo e em solução para as moléculas de poliol utilizando modelo de cálculo MP2/6-311+G(2d)//MP2/6-311+G(2d) - notação usual que representa: modelo de cálculo das propriedades elétricas//método de otimização de geometria.

Tabela 1. Resultados PCM-MP2 com conjunto de funções base 6-311+G(2d) obtidos para o momento de dipolo ( $\mu$ , em D) e para a polarizabilidade linear ( $\alpha$ , em u.a.) de moléculas de poliol no vácuo e em solução.

poliol	$\mu^{VACUO}$	μ <sup>solução</sup>	Δμ	$\alpha^{VACUO}$	$\alpha^{SOLUÇÃO}$	Δα
metanol	1,92	2,30	0,38	18,65	22,12	3,47
etileno glicol	0,00	0,00	0,00	34,32	42,02	7,70
glicerol	3,68	3,56	-0,12	50,62	63,27	12,65
D-treitol	0,00	0,00	0,00	66,57	83,92	17,35
D-arabitol	3,35	5,35	2,00	82,31	104,61	22,30
D-glucitol	2,63	4,39	1,76	98,90	125,76	26,86
D-volemitol	2,70	4,99	2,29	114,32	145,87	31,55

De forma geral, a presença de solvente tem grande influência sobre a geometria e sobre as propriedades elétricas de solutos moleculares polares. Uma estimativa para o efeito de solvente sobre estas propriedades pode ser encontrada por meio da diferença entre os resultados obtidos em solução e no vácuo ( $\Delta \mu = \mu^{\text{SOLU} \subset AO} - \mu^{\text{V} \land \text{CUO}}$ ,  $\Delta \alpha = \alpha^{\text{SOLU} \subset AO} - \alpha^{\text{V} \land \text{CUO}}$ ). As diferenças entre os resultados obtidos em solução e no vácuo para o momento de dipolo,  $\Delta \mu$ , e para a polarizabilidade linear,  $\Delta \alpha$ , também são apresentadas na Tabela 1.

#### MOMENTO DE DIPOLO

Embora o momento de dipolo das moléculas de poliol não exiba um padrão evidente com o crescimento do sistema é possível observar a relação de u com a geometria molecular. Veja que, para os cálculos no vácuo e em solução, as moléculas etileno glicol e D-treitol não exibem momento de dipolo permanente. Para essas moléculas, N = 2 e N = 4, a simetria na estrutura e consequentemente na distribuição de cargas faz com que o momento de dipolo permanente total seja nulo. Observe que, para a molécula D-glucitol (N = 6), apesar de um resultado semelhante ser esperado, o valor de µ obtido no vácuo é diferente de zero. É importante observar ainda que, para esta molécula, o resultado obtido para o momento de dipolo em solução (4,39 D) é da ordem de 70% maior que o valor obtido em vácuo (2,63 D). Estes resultados mostram que o crescimento do sistema leva a estruturas moleculares de menor energia que apresentam momento de dipolo não nulo e que a presença de água tem efeitos sobre a geometria de equilíbrio e sobre o momento de dipolo deste sistema. Os resultados apresentados na Tabela 1 mostram que o momento de dipolo deste sistema cresce inicialmente em função do número N de unidades, mas que a taxa de crescimento diminui a medida que o número de átomos no sistema aumenta o que sugere um padrão de saturação do valor desta propriedade para moléculas maiores. A Figura 8 apresenta a evolução do momento de dipolo em função do crescimento do sistema e mostra de forma mais evidente o padrão de saturação desta propriedade.



Figura 8. Evolução do momento de dipolo ( $\mu$ , em D) em função do crescimento das moléculas de poliol obtida no vácuo e em solução utilizando o método PCM-MP2 e conjunto de funções base 6-311+G(2d).

Os resultados obtidos para o momento de dipolo em água mostram que o efeito de solvente da água tem impacto bastante significativo sobre os valores de  $\mu$ . É importante ressaltar aqui que os maiores valores desta propriedade obtidos em vácuo e em solução ocorrem, respectivamente, para as moléculas de glicerol (3,68 D), e D-arabitol (5,35 D). Perceba ainda, que o efeito da água reduz levemente o momento de dipolo para a molécula N = 3 e aumenta de forma bastante significativa esta propriedade para as moléculas N =4, N = 5 e N = 6. Note que para os cálculos em solução o momento de dipolo do glicerol é inferior ao valor obtido para as moléculas D-glucitol (4,39 D) e D-volemitol (4,99 D). Esta leve redução no momento de dipolo observada para a molécula de glicerol pode estar relacionada com a geometria molecular do confôrmero utilizado. Como apenas uma geometria foi utilizada para os cálculos quânticos esse comportamento pode ser diferente para outras conformações desta molécula. Apesar disto, os resultados apresentados mostram que os efeitos de solvente têm grande importância sobre as propriedades estruturais e eletrônicas do sistema em estudo e que as moléculas de poliol podem ser polarizados quando são transferidas para o meio solvente afetando suas interações com o meio. É importante ressaltar aqui que os efeitos da água sobre o momento de dipolo são menos significativos para as quatro primeiras moléculas de poliol, mas são bastante importantes para as moléculas maiores, com incrementos nos valores de  $\mu$  que vão de 60 a 85% em relação aos valores obtidos em vácuo.

### POLARIZABILIDADE LINEAR

A Figura 9 ilustra o comportamento da polarizabilidade linear em função do crescimento das moléculas de poliol obtidos no nível MP2 com conjunto de funções base 6-311+G(2d). A polarizabilidade linear é uma propriedade aditiva, assim espera-se que o valor obtido para esta propriedade cresça com o tamanho das moléculas.



Figura 9. Evolução da polarizabilidade linear ( $\alpha$ , em u.a.) para moléculas de poliol obtida no vácuo e em solução utilizando o método PCM-MP2 com conjunto de funções base 6-311+G(2d).

Os resultados obtidos mostram que a polarizabilidade linear cresce com o incremento do sistema e que esse crescimento é mais acentuado nos cálculos em solução. Para a molécula Darabitol, por exemplo, os valores obtidos em vácuo e em solução são, respectivamente, 82,31 u.a. e 104, 61 u.a.. Isto representa um aumento da ordem de 25%. Para  $\alpha$ , os resultados obtidos em solução exibem aumentos de 19% a 28%, em relação ao resultado obtido em vácuo, e mostram que esse aumento é maior para cadeias maiores, indicando que as moléculas de poliol são altamente polarizáveis e que esta polarização tem papel importante na descrição correta das interações solutosolvente. Podemos observar que  $\Delta \alpha$  também é proporcional ao volume molecular e que moléculas maiores são mais polarizáveis. Veja que a molécula D-volemitol, N = 7, é da ordem de nove vezes mais polarizável que a molécula de metanol, N = 1. Para moléculas maiores os resultados apresentados na Tabela 1 permitem estimar um crescimento para a polarizabilidade linear da ordem de 16 u.a. e 21 u.a. no vácuo e em solução, respectivamente. Este comportamento tem consequência direta nas propriedades de hidratação das moléculas de poliol, e em particular sobre a energia livre de hidratação.

### ENERGIA LIVRE DE HIDRATAÇÃO

A comparação entre os resultados MD-CHARMM, PCM-B3LYP/6-311+G(2d) e PCM-MP2/6-311+G(2d) obtidos para a energia livre de hidratação,  $\Delta G_{CHARMM}$ ,  $\Delta G_{PCM}^{B3LYP}$  e  $\Delta G_{PCM}^{MP2}$ ; e resultados experimentais [99-101] encontrados na literatura são apresentados na Tabela 2 e na Figura 10.

N	$\Delta G_{CHARMM}$	$\Delta G_{PCM}^{B3LYP}$	$\Delta G_{PCM}^{MP2}$	$\Delta G_{EXP}$
1	-17,4	-20,5	-23,5	-21,2
2	-37,8	-43,6	-49,0	-38,9
3	-51,6	-58,5	-64,0	-56,1
4	-63,4	-70,8	-76,1	
5	-71,3	-107,6	-114,1	
6	-80,9	-119,0	-124,8	
7	-84,2	-128,4	-133,9	

Tabela 2. Comparação entre os valores de energia livre de hidratação ( $\Delta G$ , em kJ/mol) para moléculas de poliol obtidos com os modelos MD-CHARMM, PCM-B3LYP/6-311+G(2d) e PCM-MP2/6-311+G(2d).



Figura 10. Evolução da energia livre de hidratação ( $\Delta G$ , em kJ/mol) em função do crescimento para moléculas de poliol obtida por dinâmica molecular utilizando campo de força CHARMM (A) e pelo método SMD-PCM nos níveis MP2 e B3LYP (B).

A comparação com resultados da literatura mostra que os valores obtidos para a energia livre utilizando dinâmica molecular com potencial CHARMM (MD-CHARMM) e diferem entre 3 e 18%, estando em bom acordo com os resultados experimentais. De forma similar os resultados correspondentes obtidos com modelo PCM e conjunto de funções base B3LYP (PCM-B3LYP) diferem entre 3 e 12% e os resultados PCM com conjunto de funções base MP2 (PCM-MP2) variam entre 10 e 26%. É importante observar que para as quatro menores moléculas todos os modelos apresentam um comportamento aproximadamente linear como o verificado experimentalmente (ver Figura 10). Entretanto os modelos PCM indicam uma inflexão para N = 4 afetando de forma significativa o seu comportamento com o crescimento das moléculas de poliol. Este resultado difere do comportamento que é reportado para a energia de hidratação de n-alcanos [102], para os quais cada etileno adicional aumenta a energia livre de aproximadamente o mesmo valor. Para das moléculas de poliol a energia livre de hidratação aumenta em valor absoluto com o número de grupos hidroxila só que não de maneira linear. Desta forma, as contribuições de unidades moleculares devem ser usadas com cuidado para esta série de homólogos.

A Tabela 3 e a Figura 11 apresentam os resultados obtidos para a energia de polarização do soluto,  $U_{pol}$ , para moléculas de poliol obtidos com o modelo PCM utilizando os métodos B3LYP e MP2 com conjunto de funções base 6-311+G(2d). A energia de polarização do soluto corresponde a diferença entre as energias de equilíbrio obtidas na presença de água e no vácuo.

Tabela 3. Resultados B3LYP e MP2 da energia de polarização do soluto (UPOL, em kJ/mol)
para moléculas de poliol obtida por PCM com conjunto de funções base 6-311+G(2d) e
número de ligações de hidrogênio entre soluto e solvente (#LH) obtido por dinâmica
molecular.

Ν	Upol <sup>B3LYP</sup>	$U_{POL}^{MP2}$	#LH	
1	5,0	6,4	2,4	
2	7,4	10,2	4,3	
3	9,2	14,0	5,3	
4	9,9	14,5	6,7	
5	21,2	31,5	7,5	
6	24,4	35,2	8,5	
7	22,3	33,6	10,6	



Figura 11. Energia de polarização do soluto (U<sub>POL</sub>, em kJ/mol) obtida para moléculas de poliol com os modelos B3LYP e MP2 e conjunto de funções base 6-311+G(2d).

Note que a mudança no comportamento de  $\Delta G$  observado para a energia livre de hidratação para a molécula D-treitol (Figura 9) também é observada para a energia de polarização (Figura 11) afetando de forma significativa o comportamento linear desta propriedade com o tamanho do sistema. Este resultado mostra que distribuição de cargas no soluto é afetada pela presença de água o que modifica o momento de dipolo do sistema, sua interação com o solvente e consequentemente a energia de hidratação do sistema. Note que a mudança de comportamento de  $\Delta \mu$  coincide com a mudança de comportamento da U<sub>pol</sub> (ver Tabela 1 e Tabela 3).

A presença de múltiplos grupos hidroxila na série de poliol permite a formação de ligações de hidrogênio dessas moléculas com a água. O número de ligações de hidrogênio entre soluto e solvente obtidos por DM para cada molécula de poliol é apresentado na coluna #LH da Tabela 3 e na Figura 12. Neste trabalho o número de hidrogênio entre soluto e solvente (#LH) apresentado corresponde ao valor médio no conjunto de configurações obtido na simulação de dinâmica molecular e foi obtido utilizando critérios geométricos para classificar a interação soluto-solvente como está implementado no programa GROMACS.



Figura 12. Evolução do número de ligações de hidrogênio entre soluto e solvente para moléculas de poliol obtida utilizando dinâmica molecular com campo de força CHARMM.

Os resultados de dinâmica molecular mostram que o número de ligações de hidrogênio cresce com o tamanho das moléculas de poliol e que está relacionado de forma muito próxima com a energia livre de hidratação. A partir dos resultados apresentados na Tabela 2 e na Tabela 3 é possível estimar uma contribuição para a energia livre de hidratação para as moléculas de poliol da ordem de -8 a -10 kJ/mol para cada ligação de hidrogênio. É preciso ressaltar que esta estimativa pode ser afetada pelos efeitos de polarização. Relações entre mudanças na energia livre e mudanças no momento de dipolo devido aos efeitos de solvente são observados para esta série de moléculas. A partir da Tabela 2 é possível obter a diferença na variação da energia livre,  $\Delta\Delta G = \Delta G(N+1) - \Delta G(N)$ , para a série de poliol considerada. Para as quatro primeiras moléculas em que os efeitos da água como solvente são sutis a energia livre obtida nos modelos PCM-B3LYP e MD-CHARMM são essencialmente equivalentes com diferenças entre os correspondentes valores que não excedem 13%. Para as três maiores moléculas onde os efeitos de solvente são substancialmente importantes sobre o momento de dipolo, entretanto, as discrepâncias observadas entre os dois modelos são bastante significativas. Para a molécula D-arabitol, por exemplo, a magnitude de  $\Delta\Delta G$  com MD-

CHARMM é 80% menor que o valor obtido com o modelo PCM-B3LYP. Este resultado mostra que no caso das moléculas de poliol não apenas a mudança na polarizabilidade, mas a variação no momento de dipolo do vácuo para a solução influencia o comportamento da energia livre e que a utilização de campo de forças aditivos como os utilizados neste trabalho são limitados para descrever a variação na energia livre de hidratação de poli álcoois. Desta forma, para tratar de forma adequada sistemas maiores faz-se necessária a inclusão de forma explicita dos efeitos de polarização. Os resultados PCM-MP2 apresentam essencialmente o mesmo comportamento descrito para o modelo PCM-B3LYP.

#### 3.2 GLICINA

A Tabela 4 apresenta os resultados de momento de dipolo permanente ( $\mu$ ) e polarizabilidade linear ( $\alpha$ ) para as moléculas de glicina obtidos nos níveis B3LYP/6-311+G(2d)//B3LYP-cc-pVTZ e MP2/6-311+G(2d)//MP2/6-311+G(2d). Assim como para as moléculas de poliol, uma estimativa para o efeito de solvente foi obtida por meio da diferença entre os resultados em solução e em vácuo para  $\mu$  e  $\alpha$ . Os valores de  $\Delta\mu$  e  $\Delta\alpha$  são apresentados também na Tabela 4.

Tabela 4. Resultados PCM com conjunto de funções base 6-311+G(2d) obtidos para o momento de dipolo ( $\mu$ , em D) e para a polarizabilidade linear ( $\alpha$ , em u.a.) de moléculas de glicina no vácuo e em solução.

		-	B3LYP			
Ν	$\mu^{VACUO}$	μ <sup>solução</sup>	Δμ	$\alpha^{VACUO}$	$\alpha^{SOLUÇÃO}$	Δα
1	2,08	3,22	1,14	41,12	50,99	9,87
2	2,90	4,07	1,17	75,90	95,50	19,60
3	4,50	5,77	1,27	111,58	140,11	28,53
4	6,76	8,31	1,55	147,83	184,78	36,95
5	9,05	10,70	1,65	184,39	229,46	45,07
6	11,51	13,37	1,86	221,11	274,14	53,03
7	13,95	15,88	1,93	257,94	318,83	60,89

			MP2			
Ν	$\mu^{VACUO}$	μ <sup>solução</sup>	Δμ	$\alpha^{VACUO}$	α <sup>SOLUÇÃO</sup>	Δα
1	2,11	3,20	1,09	40,37	49,95	9,58
2	2,82	4,09	1,27	74,74	94,17	19,43
3	4,22	5,54	1,32	109,90	138,34	28,44
4	6,46	8,11	1,65	145,58	182,57	36,99
5	8,62	10,32	1,70	181,53	226,80	45,27

Tabela 4. Resultados PCM com conjunto de funções base 6-311+G(2d) obtidos para o momento de dipolo ( $\mu$ , em D) e para a polarizabilidade linear ( $\alpha$ , em u.a.) de moléculas de glicina no vácuo e em solução.

Note que, os modelos B3LYP e MP2 descrevem de maneira bastante semelhante o momento de dipolo e a polarizabilidade linear deste sistema. Os valores apresentados mostram que para o momento de dipolo os resultados obtidos em vácuo nos dois modelos não diferem por mais que 6% e as diferenças em solução são ainda menores, não ultrapassando 4%. Para a polarizabilidade linear a diferença entre os resultados B3LYP e MP2 em vácuo (e em solução) não é superior a 2%.

### MOMENTO DE DIPOLO

As Figuras 13A e 13B apresentam, respectivamente, a evolução do momento de dipolo e do momento de dipolo por subunidade,  $\Delta \mu(N) = |\mu(N) - \mu(N-1)|$  [12], obtidos para moléculas de glicina nos níveis MP2 e B3LYP com conjunto de funções base 6-311+G(2d).



Figura 13. Evolução do momento de dipolo (A) e do momento de dipolo por subunidade (B) obtidos para moléculas de glicina utilizando os modelos B3LYP e MP2 com conjunto de funções base 6-311+G(2d).

Os resultados PCM obtidos nos níveis B3LYP e MP2 para o momento de dipolo em vácuo e em solução mostram que esta propriedade cresce com o incremento de unidades de glicina no sistema. Note que os valores obtidos em solução são maiores que os obtidos em vácuo para os dois modelos. A partir do momento de dipolo por subunidade,  $\Delta\mu(N)$ , podemos determinar para moléculas com N > 7 um incremento da ordem de 2,27 D a 2,54 D por unidade de glicina adicional. A Tabela 4 e a Figura 13 mostram que o efeito de solvente muda de forma significante o momento de dipolo das moléculas de glicina e que seu impacto é mais importante para as moléculas menores. Para as moléculas N = 1 e N = 5, por exemplo, a diferença entre os resultados B3LYP (MP2) obtidos em vácuo e em solução são 1,14 D (1,09 D) e 1,65 D (1,70 D), respectivamente. Isto representa, para o modelo B3LYP (MP2), um aumento em relação ao resultado em vácuo da ordem de 55% (52%) para N = 1 e de 14% (20%) para N = 5. É importante destacar aqui que os resultados de  $\mu$  estão em bom acordo com resultados teóricos apresentados na literatura [103,104], o que assegura que apesar da limitação pela utilização de uma única geometria , esta descreve de forma adequada o sistema em estudo.

#### POLARIZABILIDADE LINEAR

A Figura 14 apresenta o comportamento da polarizabilidade linear e da polarizabilidade linear por subunidade,  $\Delta \alpha(N) = |\alpha(N) - \alpha(N-1)|$  [12], em função do número N de unidades de glicina obtido com modelos B3LYP/6-311+G(2d)//B3LYP-cc-pVTZ e MP2/6-311+G(2d)//MP2/6-311+G(2d).



Figura 14. Evolução da polarizabilidade linear (A) e da polarizabilidade linear por subunidade (B) de moléculas de glicina obtidos com os modelos B3LYP e MP2 e conjunto de funções base 6-311+G(2d)

É importante observar que para ambos os modelos  $\alpha$  apresenta um crescimento linear com o número de unidades de glicina no sistema (ver Figura 14A) o que reflete o caráter aditivo desta propriedade. O padrão de crescimento linear desta propriedade é evidenciado no comportamento da polarizabilidade linear por unidade em função do crescimento do sistema, como apresentado na Figura 14B. Note que os valores de  $\Delta \alpha$  tendem a um valor constante com o aumento de N. Podemos estimar, para o sistema em solução, um crescimento da ordem de 45 u.a. por unidade de glicina adicional. É importante ressaltar que, enquanto o impacto do efeito de solvente sobre o momento de dipolo diminui com o crescimento das moléculas, para a polarizabilidade linear esse efeito é praticamente constante aumentando o valor obtido em vácuo para esta propriedade da ordem de 25%.

## ENERGIA LIVRE DE HIDRATAÇÃO

A comparação entre os resultados MD-CHARMM, PCM-B3LYP/6-311+G(2d) e PCM-MP2/6-311+G(2d) obtidos para a energia livre de hidratação,  $\Delta G_{CHARMM}$ ,  $\Delta G_{PCM}^{(B3LYP)}$  e  $\Delta G_{PCM}^{(MP2)}$  de moléculas de glicina é apresentada na Tabela 5 e na Figura 15. O resultados de energia de polarização do soluto obtida nos modelos PCM-B3LYP/6-311+G(2d) e PCM-MP2/6-311+G(2d) e do número de ligações de hidrogênio entre as moléculas de soluto e solvente obtidos por dinâmica molecular clássica (U<sub>POL</sub><sup>B3LYP</sup>, U<sub>POL</sub><sup>MP2</sup> e #LH, respectivamente) também são apresentados na Tabela 5.

Tabela 5. Energia livre de hidratação obtida com os modelos MD-CHARMM ( $\Delta G_{CHARMM}$ , em kJ/mol), PCM-B3LYP e PCM-MP2 ( $\Delta G_{B3LYP}$  e  $\Delta G_{MP2}$ , em kJ/mol); energia de polarização do soluto obtidas em nível B3LYP e MP2 ( $U_{POL}^{B3LYP}$  e  $U_{POL}^{MP2}$ ); e número de ligações de hidrogênio entre soluto e solvente (#LH) obtido por dinâmica molecular clássica.

Ν	$\Delta G_{CHARMM}$	$\Delta G_{B3LYP}$	$\Delta G_{MP2}$	Upol <sup>B3LYP</sup>	$U_{POL}^{MP2}$	#LH
1	-35,65	-49,3	-44,9	12,2	18,3	4,5
2	-63,44	-75,9	-67,9	19,6	30,7	6,9
3	-95,31	-103,6	-90,1	26,0	42,3	9,2
4	-122,58	-130,9	-111,2	32,5	54,3	11,6
5	-143,09	-157,3	-131,4	38,6	66,0	13,9
6		-183,5		44,7		
7		-209,4		50,8		


Figura 15. Evolução da energia livre de hidratação ( $\Delta G$ , em kJ/mol) em função do crescimento de moléculas de glicina obtida por dinâmica molecular utilizando campo de força CHARMM e pelo método SMD-PCM nos níveis MP2 e B3LYP.

Os resultados apresentados mostram, para todos os cálculos, que a energia de hidratação das moléculas de glicina exibe um comportamento aproximadamente linear com o crescimento do sistema, veja Figura 15. Note ainda que os resultados MP2 são maiores que os B3LYP para todos os valores de N e que a diferença entre os dois resultados aumenta com o crescimento do sistema. Para as moléculas N = 2 e N = 5, por exemplo, as diferenças entre os resultados B3LYP e MP2 são, respectivamente, de 8,0 kJ/mol (11%) e 25,9 kJ/mol (16%). Observe que os resultados de  $\Delta G$  obtidos por dinâmica molecular são, para as duas primeiras moléculas, maiores que os resultados PCM, mas para todas as outras, N > 2, são maiores que os B3LYP e menores que MP2. É importante destacar que esse mesmo comportamento é observado para as moléculas de poliol. Note ainda, que os resultados obtidos para a energia de polarização do soluto e para o número de ligações de hidrogênio entre soluto e solvente (ver Tabela 5) mostram que estas propriedades crescem de forma linear com o crescimento das moléculas de glicina.

A partir dos resultados apresentados na Tabela 5 é possível estimar para moléculas maiores uma contribuição para a energia livre de hidratação da ordem de 10 kJ/mol para cada ligação de hidrogênio.

A Figura 16A apresenta uma comparação entre os resultados obtidos para a energia livre de hidratação das moléculas de glicina e poliol. Uma comparação entre resultados obtidos para o momento de dipolo, para a energia de polarização e para o número de ligações de hidrogênio entre soluto e solvente destes dois sistemas também é apresentada nas Figuras 16B, 16C e 16D, respectivamente.



Figura 16. Comparação entre os resultados de energia livre de hidratação (A) e do momento de dipolo (B) obtidos no vácuo e em solução para moléculas poliol e de glicina.



Figura 16. Comparação entre os resultados de energia de polarização do soluto (C) e do número de ligações de hidrogênio entre soluto e solvente (D) obtidos no vácuo e em solução para moléculas poliol e de glicina.

Note que para todas as propriedades apresentadas na Figura 16 o regime de evolução com o crescimento das moléculas é linear para as moléculas de glicina e em oposição a isso foge da linearidade para as moléculas de poliol. Este resultado evidencia que a energia livre de hidratação é influenciada pelo momento de dipolo do soluto e também por suas ligações intermoleculares com a água. Desta forma, espera-se que sistemas moleculares orgânicos homólogos que apresentem fugas do regime linear, com o crescimento das moléculas, para o momento de dipolo e/ou para o número de ligações de hidrogênio com o solvente, apresentem para a energia livre de hidratação um regime que não seja linear.

Os resultados de polarizabilidade linear obtidos para moléculas de poliol e de glicina mostram que estes sistemas exibem o mesmo comportamento crescente de forma linear com o incremento de unidades num padrão típico de sistemas de cadeia saturada. Por outro lado, o momento de dipolo os dois sistemas apresentam comportamentos distintos. Enquanto a glicina exibem um momento de dipolo linearmente crescente com o tamanho do sistema as moléculas de poliol sofrem alterações no valor desta propriedade para moléculas menores e o incremento no valor de  $\mu$  por unidade adicional tende a zero para moléculas maiores ( $\Delta \Delta \mu \rightarrow 0$ ). De forma semelhante ao padrão observado para o momento de dipolo, para a variação da energia livre de hidratação observa-se no caso da glicina que o valor absoluto desta propriedade cresce num padrão aproximadamente linear com o número de subunidades, enquanto tende para um valor constante para no caso das moléculas de poliol. Estes resultados mostram que variações no momento de dipolo podem ter grande impacto sobre a energia livre de hidratação de sistemas orgânicos e adicionalmente sugerem que para as moléculas de glicina o custo energético de hidratação de moléculas maiores cresce com uma razão constante por unidade adicional estimada com o modelo PCM-B3LYP em 26 kJ/mol e que para as moléculas de poliol, a partir de um determinado limite, o custo energético na hidratação de uma molécula é essencialmente o mesmo obtido para a molécula anterior.

É importante ressaltar que, neste trabalho, os cálculos para as moléculas de glicina foram realizados utilizando a geometria neutra, não sendo consideradas aqui as formas zwitteriônicas deste sistema.

### 3.3 PEPTÍDEOS $A_N K$

A Tabela 6 apresenta os resultados do momento de dipolo permanente ( $\mu$ , em D) e da polarizabilidade linear ( $\alpha$ , em u.a.) obtidos para os peptídeos A<sub>N</sub>K utilizando o modelo B3LYP/6-311+G(2d)//B3LYP-cc-pVTZ. As contribuições dos efeitos de solvente para estas propriedades ( $\Delta\mu e \Delta\alpha$ ) também são apresentadas na Tabela 6.

Tabela 6. Momento de dipolo (D) e polarizabilidade linear (u.a.) no vácuo e em solução para peptídeos A<sub>N</sub>K obtidos no nível B3LYP/6-311+G(2d).

			B3LYP			
Ν	$\mu^{VACUO}$	μ <sup>solução</sup>	Δμ	$\alpha^{VACUO}$	$\alpha^{\text{SOLUÇÃO}}$	$\Delta \alpha$
3	37,69	41,02	3,33	281,67	357,44	75,77
6	69,44	71,83	2,39	428,31	540,31	112,00
9	102,13	106,93	4,80	575,34	723,30	147,96

Observe que para os peptídeos  $A_N K$  os resultados obtidos no vácuo e em solução mostram que o momento de dipolo e a polarizabilidade linear crescem de maneira aproximadamente linear com o tamanho do sistema. A partir dos resultados em solução apresentados na Tabela 6 é possível estimar um incremento da ordem de 10 D e de 60 u.a. nos valores do momento de dipolo e da polarizabilidade linear para cada unidade adicional de alanina na estrutura peptídica. É possível verificar ainda que os efeitos de solvente da água contribuem aumentando o valor das duas propriedades, mas que o aumento mais intenso ocorre para a polarizabilidade linear. Enquanto para o momento de dipolo os incrementos são menores que 10% e diminuem com o crescimento das moléculas, para a polarizabilidade linear, o aumento em relação ao valor obtido em vácuo é da ordem de 25%. É importante ressaltar aqui que o impacto do efeito de solvente sobre o momento de dipolo diminui com o incremento de unidades de alanina no sistema de maneira semelhante ao comportamento observado para as moléculas de glicina, e que, de forma semelhante ao que foi observado para as moléculas de poliol e de glicina os resultados de  $\alpha$  em solução são da ordem de 25% maiores que os obtidos em vácuo. Os resultados obtidos para a energia livre de hidratação das moléculas A<sub>N</sub>K utilizando os modelos MD-CHARMM e PCM-B3LYP/6-311+G(2d) são apresentados na Tabela 7 e na Figura 17A. As Figuras 17B e 17C apresentam a evolução da energia de polarização e o número de ligações de hidrogênio entre soluto e solvente com o crescimento do sistema.

Tabela 7. Energia livre de hidratação obtida para peptídeos  $A_NK$  com os modelos MD-CHARMM ( $\Delta G_{CHARMM}$ , em kJ/mol), PCM-B3LYP e PCM-MP2 ( $\Delta G^{B3LYP}$  and  $\Delta G^{MP2}$ , em kJ/mol); energia de polarização do soluto obtida nos níveis B3LYP e MP2; e número de ligações de hidrogênio entre soluto e solvente (#LH) obtido por dinâmica molecular clássica.

A <sub>N</sub> K	$\Delta G_{CHARMM}$	$\Delta G_{VAC}^{B3LYP}$	$\Delta G_{SOL}^{B3LYP}$	U <sub>VAC</sub> <sup>B3LYP</sup>	U <sub>SOL</sub> <sup>B3LYP</sup>	#LH
3	-222,92	-407,55	-427,56	57,06	63,04	14,3
6	-259,66	-475,24	-499,77	73,09	81,08	20,1
9	-314,58	-538,11	-565,91	88,49	94,94	25,9



Figura 17. Evolução da energia livre de hidratação das moléculas A<sub>N</sub>K utilizando os modelos MD-CHARMM e PCM-B3LYP/6-311+G(2d) (A); da energia de polarização do soluto (B) e do número de ligações de hidrogênio entre soluto e solvente (C) com o crescimento do sistema.



Figura 17. Evolução da energia livre de hidratação das moléculas  $A_N K$  utilizando os modelos MD-CHARMM e PCM-B3LYP/6-311+G(2d) (A); da energia de polarização do soluto (B) e do número de ligações de hidrogênio entre soluto e solvente (C) com o crescimento do sistema.

Note que, qualitativamente, os modelos DM e PCM descrevem de forma bastante parecida o comportamento da energia livre de hidratação, mas que os resultados clássicos são significativamente menores que os resultados quânticos. Esta diferença é para todas as moléculas da ordem de 40%. A diferença entre os resultados clássicos e quânticos pode ser justificada neste caso pela geometria do soluto. Enquanto os cálculos de PCM foram realizados considerando um único confôrmero, a dinâmica molecular permite que diferentes configurações do soluto sejam acessadas durante a simulação. No caso dos peptídeos AnK, a estrutura molecular pode ser modificada por interações intramoleculares que levam a molécula do estado alfa-hélice para o estado folha-beta. Desta forma como o valor de  $\Delta G$  é obtido como média sobre as configurações, o valor desta propriedade difere do calculado pelo método PCM. Embora não tenha sido observada na simulação a transição completa da molécula entre estas duas configurações, estados intermediários entre estes dois extremos são acessados pelo sistema. Para incluir o efeito de temperatura e levar em consideração esta mudança geométrica nos cálculos quânticos são necessários cálculos PCM adicionais que não estão incluídos neste trabalho. Note ainda que assim como o valor absoluto de  $\Delta G$  obtidos por PCM, os valores de U<sub>POL</sub> e do #LH crescem de forma linear com o crescimento das moléculas ANK. Das simulações de DM é possível estimar um incremento da ordem de 13 kJ/mol para cada ligação de hidrogênio. Esse resultado mostra que as moléculas ANK são solúveis em água e que essa solubilidade aumenta com o crescimento do sistema. O aumento no número de ligações de hidrogênio com adição de resíduos de alanina evidencia um comportamento hidrofílico deste sistema e justifica a sua solubilidade em água. É importante ressaltar que este comportamento ocorre para um sistema muito diluídos como o simulado neste trabalho, estudos ressentes mostraram que para peptídeos A<sub>N</sub>K maiores o aumento da concentração reduz a solubilidade deste sistema que passa a formar agregados auto organizados [105].

# 4 CONCLUSÕES E PERSPECTIVAS

Neste trabalho, investigamos o comportamento da variação da energia livre de hidratação para cadeias de poliol, glicina e peptídeos  $A_NK$ . Os valores da energia de hidratação foram obtidos por integração termodinâmica utilizando simulações de dinâmica molecular clássica com descrição das moléculas de soluto feita pelo campo de força CHARMM e com o modelo TIP3P de descrição para água. Adicionalmente, os resultados de  $\Delta G$  foram obtidos utilizando o modelo contínuo polarizável combinado com o modelo de solvatação universal. O método PCM permitiu estudar ainda o momento de dipolo e a polarizabilidade linear destes sistemas em questão e estimar as contribuições que estas propriedades tem sobre a energia livre de hidratação e sua evolução com o crescimento de sistemas.

A parametrização CHARMM proposta por Hatcher [94] e colaboradores se mostrou adequada na descrição da variação da energia livre de hidratação das moléculas de poliol, os resultados obtidos com esta parametrização estão em boa concordância com os resultados experimentais apresentados na literatura, não diferindo para as três primeiras moléculas por mais que 18%.

Os resultados obtidos para a variação da energia livre de hidratação mostram que, para todos os sistemas, ela cresce em valor absoluto com o aumento do número de unidade. Para as moléculas de glicina e  $A_N K \Delta G$  exibe um comportamento linear com o crescimento do sistema semelhante ao observado para outros sistemas orgânicos descritos na literatura. Nestes dois sistemas o momento de dipolo, a polarizabilidade linear, a energia de polarização e o número de ligações de hidrogênio entre soluto e solvente exibem um comportamento linear com o crescimento do sistema. Um resultado diferente é observado para as moléculas de poliol que exibem um incremento não linear nos valores de  $\Delta G$  com o crescimento do sistema. Neste caso o impacto dos efeitos de solvente da água sobre o momento de dipolo afeta a geometria molecular e o comportamento da energia livre de hidratação.

Os resultados de polarizabilidade linear obtidos para moléculas de poliol e de glicina mostram que estes sistemas exibem o mesmo comportamento crescente de forma linear com o incremento de unidades num padrão típico de sistemas de cadeia saturada. Por outro lado, o momento de dipolo os dois sistemas apresentam comportamentos distintos. Enquanto as glicina exibem um momento de dipolo linearmente crescente com o tamanho do sistema os poliol sofrem alterações no valor desta propriedade para moléculas menores e o incremento no valor de  $\mu$  por unidade adicional tende a zero para moléculas maiores ( $\Delta \Delta \mu \rightarrow 0$ ).

De forma semelhante ao padrão observado para o momento de dipolo, para a variação da energia livre de hidratação observa-se no caso das glicinas que o valor absoluto desta propriedade cresce num padrão aproximadamente linear com o número de subunidades, enquanto tende para um valor constante para no caso das moléculas de poliol. Estes resultados mostram que variações no momento de dipolo podem ter grande impacto sobre a energia livre de hidratação de sistemas orgânicos e adicionalmente sugerem que para as moléculas de glicina o custo energético de hidratação de moléculas maiores cresce com uma razão constante por unidade adicional estimada com o modelo PCM-B3LYP em 26 kJ/mol e que para as moléculas de poliol, a partir de um determinado limite, o custo energético na hidratação de uma molécula é essencialmente o mesmo obtido para a molécula anterior. Para os peptídeos A<sub>N</sub>K nossos resultados sugerem que, para sistemas bastante diluídos, estas moléculas se tornam mais solúvel em água com o aumento do número de resíduos de alanina em razão do aumento do número de ligações de hidrogênio entre soluto e solvente.

Esperamos que este trabalho possa contribuir na descrição das propriedades termodinâmica de sistemas moleculares orgânicos e melhorar compreensão dos aspectos associados a solvatação e a influência dos efeitos de polarização e das mudanças nas propriedades eletrônicas destes sistemas sobre a energia livre de hidratação. Embora finalizado enquanto tese de doutorado, este trabalho será continuado investigando o comportamento da energia livre de hidratação, nos cálculos PCM, serão analisados diferentes confôrmeros para moléculas menores de poliol e de glicina para incluir a flexibilidade na geometria do soluto e os efeitos de temperatura no sistema. Para os cálculos clássicos de DM uma extensão possível é analisar o comportamento de  $\Delta G$  para as conformações zwitteriônicas. Outra sequência deste trabalho consiste em estudar, utilizando dinâmica molecular, o comportamento de  $\Delta G$  em função da temperatura e avaliar sistemas nanoestruturados formados pelos peptídeos A<sub>N</sub>K. O presente trabalho abre espaço ainda para investigar o comportamento da energia livre de hidratação de diferentes compostos incluindo uma extensão dos cálculos para diferentes solventes e misturas de solventes.

# REFERÊNCIAS

[1] KARPLUS, M.; MCCAMMON, A. Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, v. 9, p. 646, 2002.

[2] CHEVROT, G.; FILETI, E. E.; CHABAN, V. V., Enhanced Stability of the Model Mini-Protein in Amino Acid Ionic Liquids and Their Aqueous Solutions, *Journal of Computational Chemistry*, vol. 36, p. 2044, 2015.

[3] MENNUCCI, B.; CAMMI, R. Continuum Solvation Models in Chemical Physics: From Theory to Applications, ed., Wiley, London, 2008, 636 p.

[4] VANOMMESLAEGHE, K.; et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.*, v. 31, p. 671, 2010.

[5] MEYERS, F. et al. Eletric Field Modulated Nonlinear Optical Properties of Donor-Acceptor Polyenes: Sum-over-States Investigation of the Relationship between Molecular Polarizabilities (.alpha., .beta., and .gama.) and Bond Length Alternation. *Journal of American Chemical Society*, v. 116, p. 10703, 1994.

[6] KOKUB, H. et al. Solvation Free Energies of Alanine Peptides: The Effect of Flexibility. J. Phys. Chem. B, v. 117, p. 16428, 2013.

[7] COLHERINHAS, G.; FILETI, E. E.. Molecular Dynamics Study of Surfactant-Like Peptide Based Nanostructures, *J. Phys. Chem. B* v. 118, p. 12215, 2014.

[8] XU; H.; et al. Hydrophobic-Region-Induced Transitions in Self-Assembled Peptide. *Nanostructures Langmuir*, v. 25, p. 4115, 2009.

[9] CHEN, C. et al. Antibacterial Activities of Short Designer Peptides: a Link between Propensity for Nanostructuring and Capacity for Membrane Destabilization. *Biomacromolecules*, v. 11, p. 402, 2010.

[10] BOGDAN, E. et al. Solvent Effects on the Second-Order Nonlinear Optical Responses in the Keto-Enol Equilibrium of a 2-Hydroxy-1-naphthaldehyde Derivative. J. *Phys. Chem. C*, v. 114, p. 12760, 2010.

[11] CHAUDHURI, P.; CANUTO, S. Conformational behavior of different possible ways of oligoglycine formation in a solvent-free environment. *Journal of Molecular Structure*, v. 849, p. 25, 2008.

[12] ABREU, L. M.; FONSECA, T. L.; CASTRO, M. A. Electron correlation effects on the electric properties of fluorinated polyacetylene. *The Journal of Chemical Physics*, v. 136, p. 234311, 2012.

[13] CHAVENT, M.; DUNCAN, A. L.; SANSOM, M.S. P. Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. *Structural Biology*, v. 40, p. 8, 2016.

[14] TOMASI, J.; MENNUCCI, B.; CANCE'S, E. The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level. *Theochem*, v. 464, p. 211, 1999.

[15] GONTRANI, L.; MENNUCCI, B.; TOMASI, J. Glycine and alanine: a theoretical study of solvent effects upon energetics and molecular response properties. *Theochem*, v. 500, p. 113, 2000.

[16] COUTINHO, K.; CANUTO, S. DICE: A Monte Carlo program for molecular liquid simulation, University of São Paulo, Brazil, version 2.9 (2003).

[17] REIF, M. M.; HÜNENBERGER, P. H. Computation of methodology-independent single-ion solvation properties from molecular simulations. III. Correction terms for the solvation free energies, enthalpies, entropies, heat capacities, volumes, compressibilities, and expansivities of solvated ions. *The Journal of Chemical Physics*, v. 134, p. 144103, 2011.

[18] MACIEL, C.; MALASPINA, T.; FILETI, E. E. Prediction of the Hydration Properties of Diamondoids from Free Energy and Potential of Mean Force Calculations. *The Journal of Physical Chemistry. B*, v. 116, p. 13467, 2012.

[19] HUANG, N.; KALYANARAMAN, C.; BERNACKI, K.; JACOBSON, M. P. Molecular mechanics methods for predicting protein-ligand binding. *Phys. Chem. Chem. Phys.*, v. 8, p. 5166, 2006.

[20] BOUVIER, B. Decoding the patterns of ubiquitin recognition by ubiquitin-associated domains from free energy simulations. *Phys. Chem. Chem. Phys.*, v. 16, p. 48, 2014.

[21] CHARLIER, L.; et al. Binding free energy prediction in strongly hydrophobic biomolecular systems. *Phys. Chem. Chem. Phys.*, v. 9, p. 5761, 2007.

[22] GUERARD, J. J.; AREY, J. S. J. Critical Evaluation of Implicit Solvent Models for Predicting Aqueous Oxidation Potentials of Neutral Organic Compounds. *Chem. Theory Comput.*, v. 9, p. 5046, 2013.

[23] VEMBANUR, S.; VENKATESHWARAN V.; GARDE, S. Structure and dynamics of single hydrophobic/ionic heteropolymers at the vapor-liquid interface of water. *Langmuir*, v. 30, p. 4654, 2014.

[24] CALERO, C.; MARTI, J.; GUARDIA, E.; MASIA, M. J. Chem. Theory Comput., v. 9, p. 5070, 2013.

[25] CHABAN, V. V.; MACIEL, C.; FILETI, E. E. Solvent polarity considerations are unable to describe fullerene solvation behavior. *J. Phys. Chem. B*, v. 118, p. 3378, 2014.

[26] YELISETTY, S. S.; VISCO JR, D. P.. Solubility of HFC32, HFC125, HFC152a, and HFC143a in Three Polyols. *J. Chem. Eng.*, v. 54, p. 781, 2009.

[27] LIU, X.; ZHOU, P.; TRAN, A.; LABUZA, T. P. Effects of Polyols on the Stability of Whey Proteins in Intermediate-Moisture Food Model Systems. *Agric. Food Chem.* v. 57, p. 2339, 2009.

[28] WILMS, D.; STIRIBA, S.; FREY, H. Hyperbranched Polyglycerols: From the Controlled Synthesis of Biocompatible Polyether Polyols to Multipurpose Applications accounts of chemical research; V. 43, p. 129, 2010.

[29] STIRIBA S. E.; FREY, H.; KILBINGER, A. F. M. et al. Incorporation of a Photosensitizer Core within Hyperbranched Polyether Polyols: Effect of the Branched Shell on the Core Properties. *Macromolecules*, v. 41, p. 1189, 2008.

[30] GARRIDO, N. M.; QUEIMADA, A. J.; JORGE, M. 1-Octanol/Water Partition Coefficients of n-Alkanes from Molecular Simulations of Absolute Solvation Free Energies. *J. Chem. Theory Comput.*, v. 5, p. 2436, 2009.

[31] FONG, D. S. et al. Diabetic retinopathy. Diabetes Care. v. 10, p. 2540, 2004.

[32] SOUZA, M. I. et al. Structural and photophysical properties of peptide micro/ nanotubes functionalized with hypericin. *J. Phys. Chem. B*, v. 117, p. 2605, 2013.

[33] COLHERINHAS, G.; FILETI, E. E. Molecular Description of Surfactantlike Peptide Based Membranes. J. Phys. Chem. C, v. 118, p. 9598, 2014.

[34] DEHSORKHI, A.; CASTELLETTO, V.; HAMLEY, I. W. Self-assembling amphiphilic peptides. *J. Pept Sci.*, v. 20, p. 453, 2014.

[35] CARUSO, M.; PLACIDI, E.; GATTO, E.; MAZZUCA, C.; STELLA, L.; BOCCHINFUSO, G.; PALLESCHI, A.; FORMAGGIO, F.; TONIOLO, C.; VENANZI, M. Fibrils or globules? Tuning the morphology of peptide aggregates from helical building blocks. *J. Phys. Chem. B*, v. 117, p. 5448, 2013.

[36] LEMKUL, J. A.; BEVAN, D. R. Aggregation of Alzheimer's amyloid beta-peptide in biological membranes: a molecular dynamics study. *Biochemistry*, v. 52, p. 4971, 2013.

[37] MUNOZ, E.; SREELATHA, A.; GARRIGA, R.; BAUGHMAN, R. H.; GOUX, W. J. Amyloidogenic peptide/single-walled carbon nanotube composites based on tau-protein-related peptides derived from AcPHF6:preparation and dispersive properties. *J. Phys. Chem. B*, v. 117, p. 7593, 2013.

[38] HOSSEINKHANI, H.; HONG, P.-D.; YU, D.-S. Self-Assembled Proteins and Peptides for Regenerative Medicine. *Chem. Rev.*, v. 113, p. 4837, 2013.

[39] MAZZA, M. et al. Nanofiber-Based Delivery of Therapeutic Peptides to the Brain. *ACS Nano*, v. 7, p. 1016, 2013.

[40] LEMKUL, J. A.; BEVAN, D. R. The role of molecular simulations in the development of inhibitors of amyloid beta-peptide aggregation for the treatment of Alzheimer's disease. *ACS Chem. Neurosci.*, v. 3, p. 845, 2012.

[41] ZHANG, P.; CHEETHAM, A. G.; LIN, Y.-A.; CUI, H. Self-Assembled Tat Nanofibers as Effective Drug Carrier and Transporter. *ACS Nano*, v. 7, p. 5965, 2013.

[42] YU, T.; SCHATZ, G. C. Free energy profile and mechanism of selfassembly of peptide amphiphiles based on a collective assembly coordinate. *J. Phys. Chem. B*, v. 117, p. 9004, 2013.

[43] GAO, B.; ZHU, S. An insect defensin-derived beta-hairpin peptide with enhanced antibacterial activity. *ACS Chem. Biol.* v. 9, p. 405, 2014.

[44] ETAYASH, H. et al. Peptide-bacteria interactions using engineered surface-immobilized peptides from class IIa bacteriocins. *Langmuir*, v. 29, p. 4048, 2013.

[45] XIA, H et al. Activatable cell penetrating peptide-conjugated nanoparticles with enhanced permeability for site-specific targeting delivery of anticancer drug. *Bioconjugate Chem.* v. 24, p. 419, 2013.

[46] PALFFY, R. et al. On the physiology and pathophysiology of antimicrobial peptides. *Mol. Med.*, v. 15, p. 51, 2009.

[47] SCHLAMADINGER, D. E. et al. Spectroscopic and computational study of melittin, cecropin A, and the hybrid peptide CM15. *J. Phys. Chem. B*, v. 116, p. 10600, 2012.

[48] NICHOLS, M. et al. Dynamic turn conformation of a short tryptophan-rich cationic antimicrobial peptide and its interaction with phospholipid membranes. *J. Phys. Chem. B*, v. 117, p. 14697, 2013.

[49] NICOLAS, P.; MOR, A. Peptides as weapons against microorganisms in the chemical defense system of vertebrates. *Annu. Rev. Microbiol.* v. 49, p. 277, 1995.

[50] JAVADPOUR, M. M.; BARKLEY, M. D. Self-assembly of designed antimicrobial peptides in solution and micelles. *Biochemistry*, v. 36, p. 9540, 1997.

[51] ZHANG, L. et al. Electrochemical and Surface Properties of Solid-Supported, Mobile Phospholipid Bilayers on a Polyion/Alkylthiol Layer Pair Used for Detection of Antimicrobial Peptide Insertion. *Langmuir*, v. 18, p. 1318, 2002.

[52] SCHMITT, M. A.; WEISBLUM, B.; GELLMAN, S. H. Unexpected relationships between structure and function in alpha, beta-peptides: antimicrobial foldamers with heterogeneous backbones. *J. Am. Chem. Soc.* v. 126, p. 6848, 2004.

[53] MISHRA, A. et al. Translocation of HIV TAT peptide and analogues induced by multiplexed membrane and cytoskeletal interactions. Proc. *Natl. Acad. Sci. U. S. A.*, v. 108, p. 16883, 2011.

[54] Lind, T. K.; Zielinska, P.; Wacklin, H. P.; Urbanczyk-Lipkowska, Z.; Cardenas, M. Continuous flow atomic force microscopy imaging reveals fluidity and time-dependent interactions of antimicrobial dendrimer with model lipid membranes. *ACS Nano*, v. 8, p. 396, 2014.

[55] HAMLEY, I. W. Self-assembly of amphiphilic peptides. Soft Matter, v. 7, p. 4122, 2011.

[56] RYAN, J.A.; WHITTEN, L.A. Self-consistent field studies of glycine and glycylglycine. Simplest example of a peptide bond. J. Am. Chem. Soc., v. 94, p. 2396, 1972.

[57] SHIPMAN,L. L. ; CHRISTOFFERSEN, R.E. Ab initio calculation on large molecules using molecular fragments. Polypeptides of glycine. J. Am. Chem. Soc., v. 95, p. 4733, 1973.

[58] SCOTT, R.A.; SCHERAGA, H.A. J. Chem. Phys., v. 45, p. 2091, 1966.

[59] IMAMURA, A. M. Oshaku, Polymer, v. 24, p. 1639, 1983.

[60] PICHIERRI, F. RIKEN Review, No. 40, p. 51, 2001.

[61] SCHEINER, S.; KERN, C.W. Energies of polypeptides: theoretical conformational study of polyglycine using quantum mechanical partitioning *Proc. Natl Acad. Sci. USA*, v. 75, p. 2071, 1978.

[62] FRENKEL, D; SMIT, B.. Understanding molecular simulation – From algorithms to applications. 2 ed. San Diego CA, 2002, 444 p.

[63] B. J. ALDER; T. E. WAINWRIGHT. Phase Transition for a Hard Sphere System. J. Chem. Phys., 27 (5): 1208. Phys. Rev. A. 1957.

[64] RAHMAN, A. Correlations in the Motion of Atoms in Liquid Argon. *Phys. Rev.*, v. 136:p. A405, 1964.

[65] RAHMAN, A.; STILLINGER, F. H. J. Molecular Dynamics Study of Liquid Water. *Chem. Phys.*, v. 55, p. 3336, 1971.

[66] HOU Q., LI M., ZHOU Y., et al. Molecular dynamics simulations with many-body potentials on Multiple GPUs – The implementation, package and performance. v. 184, p. 2091, 2013.

[67] P. MOHSEN, K.; MIKKO. Early stages of interactions of cell-penetrating peptide penetratin with a DPPC bilayer. *Chemistry and Physics of Lipids*, v. 169, p. 85, 2013.

[68] BEIERLEIN, F. R.; KNEALE, G. G.; CLARK T. Predicting the Effects of Basepair Mutations in DNA-Protein Complexes by Thermodynamic Integration. *Biophysical Journal*, v. 101, p. 1130, 2011.

[69] SHIRTS, M. R.; PITERA, J. W.; SWOPE, W. C.; PANDE, V. S. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J. Chem. Phys.*, v.119, p. 5740, 2003.

[70] NEDYALKOVA, M,; MADURGA, S.; PISOV, S. et al.. Molecular dynamics simulation of the spherical electrical double layer of a soft nanoparticle: Effect of the surface charge and counterion Valence. *The Journal of Chemical Physics*, v. 137, p. 174701, 2012.

[71] COSTA, L. T.; MALASPINA, T.; FILETI, E. E.; RIBEIRO M. C. C. Molecular dynamics simulation of liquid trimethylphosphine. *The Journal of Chemical Physics*, v. 135, p. 064506, 2011.

[72] MACIEL, C.; FILETI E. E.; RIVELINO, R. Assessing the solvation mechanism of C60(OH)24 in aqueous solution. *Chemical Physics Letters*, v. 507, p. 244, 2011.

[73] ASHCROFT N. W.; DAVID, N. Física do estado sólido. Mermin, São Paulo: Cengage Learning, 2011, 870 p.

[74] KITTEL, C. Introdução à física do estado sólido 8 ed., Rio de Janeiro: LTC, 2006, 578 p.

[75] APOL, E.; APOSTOLOV, R.; BERENDSEN, H. J. C. et al. Gromacs user manual, versão 4.5.4 (online), disponível em: http://www.gromacs.org/Documentation/Manual.

[76] ARFKEN, G. B; WEBER, H. J., Mathematical Methods for Physicists, International edition, sixth edition, Elsevier Academic Press 2005, 1182 p.

[77] ALLEN, M. P.; TILDESLEY, D. J. Computer Simulation of Liquids, Biddles, Oxford, NY, 1989.

[78] RAPAPORT, D. C. The art of molecular dynamics simulation – 2<sup>a</sup> Ed. Cambridge, London, UK, 2005, 549 p.

[79] SALINAS, R. S. Introdução a física estatística – 2<sup>a</sup> ed., São Paulo SP: EDUSP, 1999, 463 p.

[80] SHIRTS, M. R. ; PANDE, V. S. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. The *Journal of chemical physics*, v. 122, p. 134508, 2005.

[81] BEVERIDGE, D. L.; MEZEI, M.; RAVISHANKER, G.; JAYARAM, B. Free energy simulations: Applications to the study of liquid water, hydrophobic interactions and solvent effects on conformational Stability. *Journal of Biosciences*, v. 8, p.167, 1986.

[82] VAN GUNSTEREN, W. F.; DAURA, X.; MARKC, A. E. Computation of Free Energy. *Helvetica Chimica Acta*, v. 85, p. 3113, 2002.

[83] MORGON; N. H.; K. COUTINHO. Métodos de Química Teórica e Modelagem Molecular, São Paulo: Livraria da Física, 2007, 540 p.

[84] TOMASI, J.; MENNUCCI, B.; CAMMI, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* V. 105, p. 2999, 2005.

[85] CAMMI, R.; COSSI, M.; MENNUCCI, B.; TOMASI, J. Analytical derivatives for molecular solutes. III. Hartree-Fock static polarizability and hyperpolarizabilities in the polarizable continuum model. *J. Chem. Phys.* v. 105, p. 10556, 1996.

[[86] MARENICH, A. V.; CRAMER, C. J.; TRUHLAR, D. G.. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B*, v. 113, p. 6378, 2009.

[87] MIERTUS, S.; SCROCCO, E.; TOMASI, J. Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *J. Chem. Phys.* v.55, p. 117, 1981.

[88] MIERTUS, S.; TOMASI, Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes. J. Chem. Phys. v. 65, p. 239, 1982.

[89] SZABO, Attila; OSTLUND, Neil S. Modern quantum chemistry: introduction to advanced electronic structure theory. Mineola. NY: Dover, 1996. 466 p.

[90] LEVINE, Ira N. Quantum chemistry. 4th ed – Englewood Cliffs: Prentice Hall, c 1991. 629 p.

[91] ROOTHAAN, J. C. C. New developments in molecular orbital theory. *Rewiew of Modern Physics*, v. 23, p. 69, 1951.

[92] COHEN-TANNOUDJI, Claude; DIU, Bernard; LALOE, Franck. Quantum mechanics. New York; Paris: J. Wiley: Hermann, c1977. 2v.

[93] MCLEAN, A. D.; YOSHIMINE, N. Theory of Molecular Polarizabilities. *Journal of Chemical Physics*, v. 47, p. 1927, 1967.

[94] HATCHER, E. R.; GUVENCH, O.; MACKERELL, JR. A. D. CHARMM Additive All-Atom Force Field for Acyclic Polyalcohols, Acyclic Carbohydrates, and Inositol *J. Chem. Theory Comput.*, v. 5, p. 1315, 2009.

[95] LINDAHL, E.; HESS, B.; SPOEL, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.*, v. 7, p. 306, 2001.

[96] BERENDSEN, H. J. C.; SPOEL, D.; DRUNEN, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.*, v. 91, p. 43, 1995.

[97] CASTET, F. et al. Reference molecules for nonlinear optics: A joint experimental and theoretical investigation. *The Journal of Chemical Physics* v. 136, p. 024506, 2012.

[98] FRISCH, M. J.; TRUCKS, G. W.; SCHLEGEL, H. B. et al. Gaussian 09, Revision A.1, Gaussian, Inc., Wallingford CT, 2009.

[99] C. P. KELLY, C. J. CRAMER and D. G. TRUHLAR, J. SM6: A Density Functional Theory Continuum Solvation Model for Calculating Aqueous Solvation Free Energies of Neutrals, Ions, and Solute–Water Clusters. *Chem. Theory Comput.*, v.1, p.1133, 2005.

[100] RIBEIRO, R. F.; MARENICH, A. V.; CRAMER, C. J.; TRUHLAR, D. G. Prediction of SAMPL2 aqueous solvation free energies and tautomeric ratios using the SM8, SM8AD, and SMD solvation models. *J. Comput.-Aided Mol. Des.*, v. 24, p. 317, 2010.

[101] WOLFENDEN, R.; ANDERSSON, L.; CULLIS, P. M.; SOUTHGATE, C. C. B. Affinities of amino acid side chains for solvent water. *Biochemistry*, v. 20, p. 849, 1981.

[102] PLYASUNOV, A. V.; SHOCK, E. L. Thermodynamic functions of hydration of hydrocarbons at 298.15 K and 0.1 MPa. *Geochim. Cosmochim. Acta*, v. 64, p. 439, 2000.

[103] ALPARONE, A. Evolution of Electric Dipole (Hyper)polarizabilities of β-Strand Polyglycine Single Chains: An ab Initio and DFT Theoretical Study. *J. Phys. Chem. A*, v. 117, p. 5184, 2013.

[104] CHAUDHURI, P.; CANUTO, S. Rayleigh scattering properties of small polyglycine molecules. *Theochem*, v. 760, p. 15, 2006.

[105] Cenker, Ç. Ç.; Bucak, S.; Olsson, U. Aqueous Self-Assembly within the Homologous Peptide Series AnK. *Langmuir*, v. 30, p. 10072, 2014.

# APÊNDICE A – MODELO DE ARQUIVO DE PARÂMETROS DE SIMULAÇÃO DE

# DINÂMICA MOLECULAR

VARIOUS PREPROCESSING OPTIONS

; Preprocessor information: use cpp syntax. ; e.g.: -I/home/joe/doe -I/home/mary/roe

#### include ; e.g.: -DPOSRES -DFLEXIBLE (note these variable names are case sensitive) define = ; RUN CONTROL PARAMETERS integrator = sd ; Start time and timestep in ps tinit = 0= 0.002dt nsteps = 5000000; For exact run continuation or redoing part of a run init step = 0; Part index is updated automatically on checkpointing (keeps files separate) simulation part = 1 ; mode for center of mass motion removal comm-mode = Linear ; number of steps for center of mass motion removal = 1 nstcomm ; group(s) for center of mass motion removal comm-grps = ; LANGEVIN DYNAMICS OPTIONS

; Friction coefficient (amu/ps) and random seed bd-fric = 0 ld-seed = 993

### ; ENERGY MINIMIZATION OPTIONS ; Force tolerance and initial step-size emtol = 10

emstep = 0.01; Max number of iterations in relax\_shells niter = 20; Step size (ps^2) for minimization of flexible constraints fcstep = 0; Frequency of steepest descents steps when doing CG nstcgsteep = 1000nbfgscorr = 10 ; TEST PARTICLE INSERTION OPTIONS = 0.05rtpi ; OUTPUT CONTROL OPTIONS ; Output frequency for coords (x), velocities (v) and forces (f) nstxout = 0= 0nstvout = 0nstfout ; Output frequency for energies to log file and energy file = 100nstlog nstcalcenergy = -1= 100nstenergy ; Output frequency and precision for .xtc file nstxtcout = 0xtc-precision = 1000; This selects the subset of atoms for the .xtc file. You can ; select multiple groups. By default all atoms will be written. xtc-grps ; Selection of energy groups energygrps = ; NEIGHBORSEARCHING PARAMETERS ; nblist update frequency nstlist = 5 ; ns algorithm (simple or grid) ns type = grid ; Periodic boundary conditions: xyz, no, xy pbc = X Y Zperiodic molecules = no ; nblist cut-off rlist = 1.4; long-range cut-off for switched potentials rlistlong = -1 ; OPTIONS FOR ELECTROSTATICS AND VDW ; Method for doing electrostatics coulombtype = PMErcoulomb-switch = 0rcoulomb = 1.4; Relative dielectric constant for the medium and the reaction field epsilon r = 1 epsilon rf = 1 ; Method for doing Van der Waals vdw-type = switch ; cut-off lengths rvdw-switch = 0.9rvdw = 1.1

; Apply long range dispersion corrections for Energy and Pressure DispCorr = EnerPres ; Extension of the potential lookup tables beyond the cut-off table-extension = 1 ; Seperate tables between energy group pairs energygrp table ; Spacing for the PME/PPPM FFT grid fourierspacing = 0.12; FFT grid size, when a value is 0 fourierspacing will be used fourier nx = 0fourier ny = 0fourier nz = 0; EWALD/PME/PPPM parameters pme order = 6 ewald rtol = 1e-06ewald geometry = 3depsilon surface = 0optimize fft = no

```
; IMPLICIT SOLVENT ALGORITHM
implicit solvent = No
```

```
; GENERALIZED BORN ELECTROSTATICS
; Algorithm for calculating Born radii
                    = Still
gb algorithm
; Frequency of calculating the Born radii inside rlist
nstgbradii
                  = 1
; Cutoff for Born radii calculation; the contribution from atoms
; between rlist and rgbradii is updated every nstlist steps
                 = 1
rgbradii
; Dielectric coefficient of the implicit solvent
gb epsilon solvent
                      = 80
; Salt concentration in M for Generalized Born models
                   = 0
gb saltconc
; Scaling factors used in the OBC GB model. Default values are OBC(II)
gb obc alpha
                     = 1
gb obc beta
                    = 0.8
gb obc gamma
                       = 4.85
gb dielectric offset = 0.009
sa algorithm
                    = Ace-approximation
; Surface tension (kJ/mol/nm<sup>2</sup>) for the SA (nonpolar surface) part of GBSA
; The value -1 will set default value for Still/HCT/OBC GB-models.
sa surface tension
                      = -1
; OPTIONS FOR WEAK COUPLING ALGORITHMS
; Temperature coupling
```

tcoupl = v-rescale

nsttcouple = -1 nh-chain-length = 10; Groups to couple separately tc grps = system ; Time constant (ps) and reference temperature (K) tau t = 1.0= 298ref t ; Pressure coupling Pcoupl = Parrinello-Rahman Pcoupltype = Isotropic nstpcouple = 2; Time constant (ps), compressibility (1/bar) and reference P (bar) = 0.1tau p compressibility = 4.5e-05ref p = 1.0; Scaling of reference coordinates, No, All or COM refcoord scaling = No; Random seed for Andersen thermostat andersen seed = 815131; OPTIONS FOR QMMM calculations QMMM = no ; Groups treated Quantum Mechanically QMMM-grps = ; QM method QMmethod = ; QMMM scheme OMMMscheme = normal ; QM basisset QMbasis = ; QM charge QMcharge = ; QM multiplicity QMmult = ; Surface Hopping SH ; CAS space options CASorbitals = CASelectrons = SAon = SAoff = SAsteps = ; Scale factor for MM charges MMChargeScaleFactor = 1 ; Optimization of QM subsystem bOPT = bTS =

; SIMULATED ANNEALING

; Type of annealing for each temperature group (no/single/periodic) annealing ; Number of time points to use for specifying annealing in each group annealing\_npoints ; List of times at the annealing points for each group annealing time ; Temp. at each annealing point, for each group. annealing temp =

## ; GENERATE VELOCITIES FOR STARTUP RUN

gen_vel	= no
gen-temp	= 300
gen-seed	= 13529

; OPTIONS FOR BONDS

constraints = all-bonds

; Type of constraint algorithm

constraint-algorithm = lincs

; Do not constrain the start configuration

continuation = yes

; Use successive overrelaxation to reduce the number of shake iterations

Shake-SOR = no

; Relative tolerance of shake

shake-tol = 0.0001

; Highest order in the expansion of the constraint coupling matrix lincs-order = 12

; Number of iterations in the final step of LINCS. 1 is fine for

; normal simulations, but use 2 to conserve energy in NVE runs.

; For energy minimization with constraints it should be 4 to 8. lincs-iter = 1

; Lincs will write a warning to the stderr if in one step a bond

; rotates over more degrees than

lincs-warnangle = 30

; Convert harmonic bonds to morse potentials morse

```
= no
```

; ENERGY GROUP EXCLUSIONS

; Pairs of energy groups for which all non-bonded interactions are excluded energygrp excl =

#### ; WALLS

; Number of walls, type, atom types, densities and box-z scale factor for Ewald nwall = 0wall type = 9-3wall r linpot = -1

wall atomtype =wall density = wall ewald zfac = 3 ; COM PULLING ; Pull type: no, umbrella, constraint or constant force = nopull ; NMR refinement stuff ; Distance restraints type: No, Simple or Ensemble disre = No; Force weighting of pairs in one distance restraint: Conservative or Equal disre-weighting = Conservative ; Use sqrt of the time averaged times the instantaneous violation disre-mixed = no disre-fc = 1000disre-tau = 0; Output frequency for pair distances to energy file nstdisreout = 100; Orientation restraints: No or Yes orire = no ; Orientation restraints force constant and tau for time averaging orire-fc = 0= 0orire-tau orire-fitgrp ; Output frequency for trace(SD) and S to energy file = 100nstorireout ; Dihedral angle restraints: No or Yes dihre = no dihre-fc = 1000; Free energy control stuff free energy = yes init lambda = 1.00delta lambda = 0foreign lambda = 0.0sc-alpha = 0.0sc-power = 0.0sc-sigma = 0.0nstdhdl = 10separate dhdl file = yes dhdl-derivatives = yes dh hist size = 0dh hist spacing = 0.1couple-moltype = methanol couple-lambda0 = vdw couple-lambda1 = vdw-q

couple-intramol = no

; Non-equilibrium MD stuff acc-grps =

accelerate	=
freezegrps	=
freezedim	=
cos-acceleration	= 0
deform	=

; Electric fields

; Format is number of terms (int) and for all terms an amplitude (real) ; and a phase angle (real) E-x =

L'-A	_
E-xt	=
E-y	=
E-yt	=
E-z	=
E-zt	=

; User defined thingies

user1-grps	=
user2-grps	=
userint1	= 0
userint2	= 0
userint3	= 0
userint4	= 0
userreal1	= 0
userreal2	= 0
userreal3	= 0
userreal4	= 0

## APÊNDICE B – PARAMETRIZAÇÃO DO CAMPO DE FORÇA CHARMM PARA

## **MOLÉCULAS DE POLIOL**

[ bondedtypes ] ; Col 1: Type of bond

; Col 2: Type of angles

; Col 3: Type of proper dihedrals

, Col 3. Type of proper differences

; Col 4: Type of improper dihedrals

; Col 5: Generate all dihedrals if 1, only heavy atoms of 0.

; Col 6: Number of excluded neighbors for nonbonded interactions

; Col 7: Generate 1,4 interactions between pairs of hydrogens if 1

; Col 8: Remove propers over the same bond as an improper if it is 1

; bonds angles dihedrals impropers all dihedrals nrexcl HH14 RemoveDih

1	. 5	Ģ	92	2 1	3	31	0

[PA1]

; Poly-Alcohol(n=1) - Methanol [ atoms ] CA CC322 0.050 1 HA1 HCA3 0.060 1 HA2 HCA3 0.060 1 HA3 HCA3 0.060 1 OH OC311 -0.650 2 HO HCP1 0.420 2 [bonds] CA HA1 CA HA2 CA HA3 CA OH OH НО

[PA2] ; Poly-Alcohol(n=2) - Glycol [ atoms ] HO1 HCP1 0.420 1 OH1 OC311 -0.650 1 CA CC322 0.050 1 HA1 HCA2 0.090 1 HA2 HCA2 0.090 1 CB CC322 0.050 2 HB1 HCA2 0.090 2 HB2 HCA2 0.090 2 OH2 OC311 -0.650 2 HO2 HCP1 0.420 2 [bonds] OH1 HO1 OH1 CA CA HA1 CA HA2

CB

CB

CB

OH2

HO2

CA

HB1

HB2

CB

OH2

[PA3] ; Poly-Alcohol(n=3) - Glycerol [ atoms ] HCP1 HO1 0.420 1 OH1 OC311 -0.650 1 CA CC322 0.050 1 HA1 HCA2 0.090 1 HA2 HCA2 0.090 1 CB CC312 0.140 2 HCA1 0.090 2 HB OH2 OC311 -0.650 2 HO2 HCP1 0.420 2 0.050 CC CC322 3 HC1 HCA2 0.090 3 HC2 HCA2 0.090 3 OH3 OC311 -0.650 3 HCP1 0.420 3 HO3 [bonds] OH1 HO1 OH1 CA CA HA1 CA HA2 CB CA CB HB CB OH2 OH2 HO2 CC CB CC HC1  $\mathbf{C}\mathbf{C}$ HC2 CC OH3 OH3 HO3 [PA4] ; Poly-Alcohol(n=4) - Threitol [ atoms ] HO1 HCP1 0.420 1

OH1 OC311 -0.650 1 CA CC322 0.050 1 HA1 HCA2 0.090 1 HA2 HCA2 0.090 1 CB CC312 0.140 2 HB HCA1 0.090 2 OH2 OC311 -0.650 2 HO2 HCP1 0.420 2 CC CC312 0.140 3 HC HCA1 0.090 3 OH3 OC311 -0.650 3 HO3 HCP1 0.420 3 CD CC322 0.050 5 HD1 HCA2 0.090 5 HD2 HCA2 0.090 5 OH4 OC311 -0.650 5 HO4 HCP1 0.420 5 [bonds] OH1 HO1 OH1 CA CA HA1 CA HA2 CB CA CB HB CB OH2 OH2 HO2 CC CB CC HC CCOH3 OH3 HO3 CD CCCD HD1 CD HD2 CD OH4 OH4 HO4

98

[PA5] ; Poly-Alcohol(n=5) - Arabitol [ atoms ] HO1 HCP1 0.420 1 OH1 OC311 -0.650 1 CA CC322 0.050 1 0.090 1 HA1 HCA2 HA2 HCA2 0.090 1 CB CC312 0.140 2 HB HCA1 0.090 2 OH2 OC311 -0.650 2 2 HO2 HCP1 0.420 CC CC312 0.140 3 HC HCA1 0.090 3 OH3 OC311 -0.650 3 HO3 HCP1 0.420 3 CD CC312 0.140 4 HD HCA1 4 0.090 OH4 OC311 -0.650 4 HO4 HCP1 0.420 4 0.050 CE CC322 5 HE1 HCA2 0.090 5 HE2 HCA2 0.090 5 OH5 OC311 -0.650 5 HO5 HCP1 0.420 5 [bonds] OH1 HO1 OH1 CA CA HA1 CA HA2 CB CA CB HB CB OH2 OH2 HO2 CCCB CC HC CC OH3 OH3 HO3 CD CC CD HD CD OH4 OH4 HO4 CE CD CE HE1 CE HE2 CE OH5

OH5 HO5

[ PA6 ] ; Poly-Alcohol(n=6) - Glucitol [ atoms ] HO1 HCP1 0.420 1 OH1 OC311 -0.650 1 CA CC322 0.050 1 HA1 HCA2 0.090 1 HA2 HCA2 0.090 1 CB CC312 0.140 2 HB HCA1 0.090 2 OH2 OC311 -0.650 2 HO2 HCP1 0.420 2 CC CC312 0.140 3 HC HCA1 0.090 3 OH3 OC311 -0.650 3 HO3 HCP1 0.420 3 CD CC312 0.140 4 HD HCA1 0.090 4 OH4 OC311 -0.650 4 HO4 HCP1 0.420 4 CE CC312 0.140 5 HCA1 5 HE 0.090 OH5 OC311 -0.650 5 HO5 HCP1 0.420 5 CF CC322 0.050 6 HF1 HCA2 0.090 6 HF2 HCA2 0.090 6 OH6 OC311 -0.650 6

HO6 HCP1 0.420 6 [bonds] OH1 HO1 OH1 CA CA HA1 CA HA2 CB CA CB HB CB OH2 OH2 HO2 CCCB CC HC CC OH3 OH3 HO3 CD CCCD HD CD OH4 OH4 HO4 CE CD CE HE CE OH5 OH5 HO5 CF CE CF HF1 CF HF2 CF OH6 OH6 HO6

[PA7]						
; Poly-Alcohol(n=7) - Volemitol						
[ atom	s ]					
HO1	HCP1	0.420	1			
OH1	OC311	-0.650	1			
CA	CC322	0.050	1			
HA1	HCA2	0.090	1			
HA2	HCA2	0.090	1			
CB	CC312	0.140	2			
HB	HCA1	0.090	2			
OH2	OC311	-0.650	2			
HO2	HCP1	0.420	2			
CC	CC312	0.140	3			
HC	HCA1	0.090	3			
OH3	OC311	-0.650	3			
HO3	HCP1	0.420	3			
CD	CC312	0.140	4			
HD	HCA1	0.090	4			
OH4	OC311	-0.650	4			
HO4	HCP1	0.420	4			
CE	CC312	0.140	5			
HE	HCA1	0.090	5			
OH5	OC311	-0.650	5			
HO5	HCP1	0.420	5			
CF	CC312	0.140	6			
HF	HCA1	0.090	6			
OH6	OC311	-0.650	6			
HO6	HCP1	0.420	6			
CG	CC322	0.050	7			
HG1	HCA2	0.090	7			
HG2	HCA2	0.090	7			
OH7	OC311	-0.650	7			

HO7 HCP1 0.420 7 [bonds] OH1 HO1 CA OH1 CA HA1 CA HA2 CA CBCB HB OH2 CB OH2 HO2  $\mathbf{C}\mathbf{C}$ CBHC CC OH3 CCOH3 HO3  $\mathbf{C}\mathbf{C}$ CD HD CDOH4 CDHO4 OH4 CE CD CE HE CE OH5 HO5 OH5 CF CE CF HF CF OH6 OH6 HO6 CF CG CGHG1 CG HG2 CG OH7 OH7 HO7

```
; Adicionado por mim para o campo de forca dos poli-alcoois
                    CH carbon in linear polyols
CC312 12.01100;
                    C2, C3, C4 CH bound to OH
CC3161 12.01100;
                    CH2 carbon in linear polyols erh
CC322 12.01100;
                    generic acyclic CH3 carbon
CC331 12.01100;
                    sp2 carbon in acetone, ketoses
CC2O3 12.01100;
CC2O4 12.01100;
                    c22 CD
                   aliphatic proton, CH
HCA1
       1.00800;
                   aliphatic proton, CH2
HCA2
       1.00800;
                   aliphatic proton, CH3
HCA3
       1.00800;
HCP1
       1.00800;
                   polar H
HCR1
       1.00800;
                   c22 HR1
OC311 15.99940;
                    hydroxyl oxygen
                    sp2 oxygen in acetone, ketoses
OC2D3 15.99940;
OC2D4 15.99940;
                    par22 O
[bondtypes]
                                   func
                                            b0
                                                  kb
;i
      1
;PYRENYL
                0.151 192464.0
CPT
     С
           1
;POLYIOLS
OC2D3 CC2O3 1
                     0.1215 585760.0 ; ketone MP2/6-31g*, CSD geometry
CC2O3 CC331 1
                    0.1500 276144.0 ; ketone MP2/6-31g*, CSD geometry
CC2O3 CC312 1
                    0.1500 276144.0 ; ketone
CC2O3 CC322 1
                    0.1500 276144.0 ; ketone
OC2D4 CC2O4 1
                     0.1215 585760.0 ; acetaldehyde, adm
HCR1 CC2O4 1
                    0.1110 276144.0 ; acetaldehyde, adm
CC2O4 CC331 1
                    0.1500 209200.0 ; acetaldehyde, adm
                    0.1500 209200.0 ; acetaldehyde, adm
CC2O4 CC312 1
CC312 CC312 1
                    0.1485 186188.0 ; adm 11/08, glycerol
                    0.1485 186188.0 ; adm 11/08, glycerol
CC312 CC322 1
                     0.1480 186188.0; sng qm and crystal analysis
CC3161 CC3161 1
CC3161 OC311 1
                    0.1410 343088.0 ; og crystal analysis
CC312 OC311 1
                    0.1420 334720.0 ; adm 11/08, glycerol
                    0.1420 334720.0 ; adm 11/08, glycerol
CC322 OC311
               1
                    0.1420 334720.0 ; adm 11/08, glycerol
CC322 CC322 1
                    0.1111 258571.2 ; par22 HA CT1
CC312 HCA1
               1
CC3161 HCA1
               1
                    0.1111 258571.2 ; par22 HA CT1
CC322 HCA2
                    0.1111 258571.2 ; par22 HA CT2
               1
                    0.1111 269449.6 ; par22 HA CT3
CC331 HCA3
               1
CC322 HCA3
               1
                    0.1111 269449.6 ; par22 HA CT3
CC322 HCP1
               1
                   0.1111 269449.6 ; par22 HA CT3
OC311 HCP1
               1
                    0.0960 456056.0 ; par22 OH1 H
;
```

[ anglety	ypes ]								
;i j					k func	th0	cth	ub0	cub
;POLYC	DLS								
CC322	CC312	CC322	5	111.00	371.16	0.0	0.00	; adm	11/08, glycerol
CC312	CC312	CC322	5	111.00	371.16	0.0	0.00	; adm	11/08, glycerol
CC312	CC2O3	OC2D3	5	121.50	618.60	0.0	0.00	; ket	one, par22 O CD CT2
CC331	CC2O3	OC2D3	5	121.50	618.60	0.0	0.00	; ket	one, par22 aldehydes.st
CC322	CC2O3	OC2D3	5	121.50	618.60	0.0	0.00	; ket	one, par22 O CD CT2
CC331	CC2O3	CC331	5	116.00	288.68	0.0	0.00	; keto	one, C-C-C
CC312	CC2O3	CC322	5	116.00	288.68	0.0	0.00	: keto	one
CC322	CC2O3	CC322	5	116.00	288.68	0.0	0.00	: keto	one
CC312	CC312	CC2O3	5	113.80	494.88	0.0	0.00	; ket	one, par22 str CT3 CT2
CD			-		.,			,	, , , , , , , , , , , , , , , , , , ,
CC2O3	CC331	HCA3	5	109.50	412.40	0.0	0.00	; keto	one
CC2O3	CC322	HCA2	5	109.50	412.40	0.0	0.00	; keto	one, par22 HB CT2 CD
HCA1	CC312	CC2O3	5	109.50	412.40	0.0	0.00	; keto	one, par22 HB CT1 CD
OC311	CC312	CC2O3	5	108.89	453.64	0.0	0.00	; keto	one, og 1/06 EtOH NIST
IR fit									
CC2O3	CC322	OC311	5	108.89	453.64	0.0	0.00	; kete	one, og 1/06 EtOH NIST
IR fit									-
OC2D4	CC2O4	HCR1	5	118.00	536.12	0.0	0.00	; ace	taldehyde, adm
HCR1	CC2O4	CC331	5	116.00	536.12	0.0	0.00	; acet	taldehyde, adm
HCR1	CC2O4	CC312	5	116.00	536.12	0.0	0.00	; acet	taldehyde, adm
CC312	CC2O4	OC2D4	5	126.00	371.16	0.0	0.00	; ace	etaldehyde, adm
CC331	CC2O4	OC2D4	5	126.00	371.16	0.0	0.00	; ace	etaldehyde, adm
CC312	CC312	CC2O4	5	113.80	494.88	0.0	0.00	; alde	ehyde, par22 CT3 CT2
CD									
OC311	CC312	CC2O4	5	108.89	453.64	0.0	0.00	; ald	ehyde, og 1/06 EtOH
NIST IR	L fit								
CC2O4	CC312	HCA1	5	109.50	272.18	0.2163	3 2474	4.00 ;	acetaldehyde, adm
CC2O4	CC331	HCA3	5	109.50	272.18	0.2163	3 2474	4.00 ;	acetaldehyde, adm
CC3161	CC3161	L CC316	15	111.00	440.03	0.256	659	8.40	; par22 CT1 CT1 CT1
CC312	CC312	CC312	5	111.00	371.16	0.0	0.00	; adm	11/08, glycerol
CC312	CC312	HCA1	5	110.10	284.56	0.2179	18582	2.74 ;	par22 HA CT1 CT1
CC3161	CC3161	HCA1	5	110.10	284.56	0.217	9 1858	32.74	; par22 HA CT1 CT1
CC322	CC312	HCA1	5	110.10	284.56	0.2179	18582	2.74 ;	par22 HA CT1 CT2
CC322	CC322	HCA1	5	110.10	284.56	0.2179	18582	2.74 ;	par22 HA CT1 CT2
CC322	CC322	HCA2	5	110.10	284.56	0.2179	18582	2.74 ;	par22 HA CT1 CT2
CC312	CC322	HCA2	5	110.10	275.73	0.2179	18582	2.74 ;	par22 HA CT2 CT1
HCA2	CC322	HCA2	5	109.00	292.80	0.1802	4453.	92 ;]	par22 HA CT2 HA
HCA3	CC331	HCA3	5	108.40	292.80	0.1802	4453.	92 ;]	par22 HA CT3 HA
HCA3	CC322	HCA3	5	108.40	292.80	0.1802	4453.	92 ;]	par22 HA CT3 HA
HCA3	CC322	OC311	5	108.40	412.40	0.1802	4453.	92 ;	par22 HA CT3 HA
CC312	CC312	OC311	5	108.00	624.37	0.0	0.00	; adm	11/08, glycerol
OC311	CC322	CC322	5	108.00	624.37	0.0	0.00	; adm	11/08, glycerol
CC3161	CC3161	l OC311	5	110.10	624.37	0.0	0.00	; pai	22 OH1 CT1 CT1

CC312 CC322 OC311 5 106.00 412.40 0.0 0.00 ; adm 11/08, glycerol CC322 CC312 OC311 5 108.00 624.37 0.0 0.00 ; adm 11/08, glycerol CC3161 OC311 HCP1 5 109.00 412.40 0.00 ; sng mod (qm and crystal 0.0 data) II HCP1 OC311 CC312 5 106.00 412.40 0.0 0.00 ; og 1/06 EtOH IR fit HCP1 OC311 CC322 5 106.00 412.40 0.0 0.00 ; og 1/06 EtOH IR fit 108.89 453.64 HCA1 CC312 OC311 5 0.0 0.00 ; og 1/06 EtOH NIST IR fit 108.89 453.64 HCA1 CC3161 OC311 5 0.0 0.00 ; og 1/06 EtOH NIST IR fit HCA2 CC322 OC311 5 108.89 453.64 0.0 0.00 ; og 1/06 EtOH NIST IR fit ; [ dihedraltypes ] k 1 func phi0 ;i cp mult j ; PYRENYL HP CA CPT CA 9 0.00 12.552 1 HP CA CPT С 9 0.00 12.552 1 180.00 12.9704 1 CPT CA 9 CA CA 9 CPT CPT С 0 180.00 12.9704 2 9 CA CPT С 0 180.00 12.9704 2 CPT С 9 0.00 12.9704 1 :CPT 0 ;CA CPT С 0 9 0.00 12.9704 1 С CPT 9 180.00 12.9704 1 CA CA CPT CA CPT 9 CA 0.00 12.9704 1 ;POLYOLS CC312 CC312 CC312 CC2O3 9 180.0 0.08 1 ; erh ketone, n=6 polyol CC312 CC312 CC312 CC2O3 9 0.0 0.25 2 ; " RIMP2/cc-pVTZ//MP2/631gd 0.0 10.33 3 ; " C2-C3-C4-C5 CC312 CC312 CC312 CC203 9 9 CC312 CC312 CC2O3 CC322 180.0 0.08 1 ; erh ketone, n=6 polyol CC312 CC312 CC2O3 CC322 9 0.0 0.25 2 ; " RIMP2/cc-pVTZ//MP2/631gd 0.0 10.33 3 ; " C2-C3-C4-C5 CC312 CC312 CC2O3 CC322 9 CC2O3 CC312 CC312 OC311 9 180.0 10.96 1 ; erh ketone, fit psicose CC2O3 CC312 CC312 OC311 9 180.0 1.17 2 ; "RIMP2/cc-pVTZ// MP2/631gd CC2O3 CC312 CC312 OC311 9 0.0 9.54 3 ; " C2-C3-C4-O4 CC312 CC312 CC2O3 OC2D3 9 0.0 8.62 1 ; erh ketone, fit psicose CC312 CC312 CC2O3 OC2D3 9 0.0 0.54 2 ; " RIMP2/cc-pVTZ// MP2/631gd CC312 CC312 CC2O3 OC2D3 9 0.0 12.01 3 : "C4-C3-C2-O2 CC312 CC2O3 CC322 OC311 9 0.0 12.30 1 ; erh ketone, fit psicose CC312 CC2O3 CC322 OC311 9 180.0 6.19 2 ; "RIMP2/cc-pVTZ//MP2/631gd CC312 CC2O3 CC322 OC311 180.0 2.09 3 ; "C3-C2-C1-O1 9 CC322 CC2O3 CC322 OC311 0.0 4.31 1 ; erh ketone, fit psicose 9 9 180.0 10.21 2 ; " RIMP2/cc-pVTZ// MP2/631gd CC322 CC2O3 CC322 OC311 CC322 CC2O3 CC322 OC311 180.0 5.19 3 ; " C1-C2-C3-O3 9 CC322 CC2O3 CC312 OC311 9  $0.0 \quad 4.31 \quad 1$ ; erh ketone, fit psicose CC322 CC2O3 CC312 OC311 9 180.0 10.21 2 ; " RIMP2/cc-pVTZ// MP2/631gd 9 180.0 5.19 3; "C1-C2-C3-O3 CC322 CC2O3 CC312 OC311 OC2D3 CC2O3 CC312 OC311 9 1 ; erh ketone, fit psicose 0.0 8.74

OC2D3 CC2O3 CC312 OC311 9 180.0 0.71 2; "RIMP2/cc-pVTZ//MP2/631gd OC2D3 CC2O3 CC312 OC311 3 : " 02-C2-C3-O3 9 180.0 9.00 1 ; erh ketone, fit psicose OC2D3 CC2O3 CC322 OC311 9 0.0 11.72 OC2D3 CC2O3 CC322 OC311 9 180.0 3.14 2 ; " RIMP2/cc-pVTZ// MP2/631gd OC2D3 CC2O3 CC322 OC311 9 180.0 1.34 3 ; " O2-C2-C1-O1 CC331 CC2O3 CC331 HCA3 0.0 0.42 3 ; erh ketone RIMP2/ccpVTZ//MP2/6-9 31G(d) CC2O3 CC312 CC312 HCA1 9 0.0 0.42 3; erh ketone RIMP2/ccpVTZ//MP2/6-31G(d) CC322 CC2O3 CC312 HCA1 9 0.0 0.42 3; erh ketone RIMP2/ccpVTZ//MP2/6-31G(d) CC312 CC2O3 CC322 HCA2 9 0.0 0.42 3; erh ketone RIMP2/ccpVTZ//MP2/6-31G(d) CC322 CC2O3 CC322 HCA2 9 0.0 0.42 3 ; erh ketone RIMP2/ccpVTZ//MP2/6-31G(d) OC2D3 CC2O3 CC331 HCA3 9 0.0 0.00 3; erh ketone RIMP2/ccpVTZ//MP2/6-31G(d) OC2D3 CC2O3 CC322 HCA2 9 0.0 0.00 3; erh ketone RIMP2/ccpVTZ//MP2/6-31G(d) CC2O3 CC312 HCA1 OC2D3 0.0 0.00 3; erh ketone RIMP2/ccpVTZ//MP2/6-9 31G(d) CC2O3 CC312 OC311 HCP1 9 0.0 1.46 1 ; erh ketone, n=6polyol CC2O3 CC312 OC311 HCP1 9 2 ; " RIMP2/cc-pVTZ//MP2/631gd 0.0 1.55 9 CC2O3 CC312 OC311 HCP1 180.0 0.79 3 ; "С-С-О-Н CC2O3 CC322 OC311 HCP1 9 0.0 1.46 1 ; erh ketone, n=6polyol CC2O3 CC322 OC311 HCP1 9 0.0 1.55 2; "RIMP2/cc-pVTZ//MP2/631gd 9 180.0 0.79 3 ; "С-С-О-Н CC2O3 CC322 OC311 HCP1 CC312 CC312 CC312 CC204 9 180.0 0.08 1 ; erh aldehyde, n=6polyol 9 CC312 CC312 CC312 CC204 0.0 0.25 2 ; " RIMP2/cc-pVTZ//MP2/631gd CC312 CC312 CC312 CC204 9 0.0 10.33 3 ; " C2-C3-C4-C5 CC312 CC312 CC2O4 OC2D4 9 180.0 0.71 1 ; erh aldehyde, fitallose CC312 CC312 CC2O4 OC2D4 9 180.0 1.09 2 ; " RIMP2/cc-pVTZ//MP2/631gd 180.0 4.35 3 ; "C3-C2-C1-O1 CC312 CC312 CC2O4 OC2D4 9 CC2O4 CC312 CC312 OC311 9 0.0 4.81 1 ; erh aldehyde, fitallose CC2O4 CC312 CC312 OC311 0.0 12.55 2 ; " RIMP2/cc-pVTZ//MP2/631gd 9 CC2O4 CC312 CC312 OC311 9 180.0 10.00 3 ; " C1-C2-C3-O3 OC2D4 CC2O4 CC312 OC311 9 180.0 2.34 1 ; erh aldehyde, fitallose OC2D4 CC2O4 CC312 OC311 9 180.0 6.69 2 ; " RIMP2/cc-pVTZ//MP2/631gd 0.0 3.89 3 ; " O1-C1-C2-O2 OC2D4 CC2O4 CC312 OC311 9 HCP1 OC311 CC312 CC2O4 9 0.0 1.46 1 ; erh aldehyde, n=6polyol OC311 CC312 CC2O4 9 HCP1 0.0 1.55 2 ; " RIMP2/cc-pVTZ//MP2/631gd 180.0 0.79 3 ; "С-С-О-Н HCP1 OC311 CC322 CC322 9 180.0 0.00 3 ; acetaldehyde, adm HCR1 CC2O4 CC331 HCA3 9 OC2D4 CC2O4 CC331 HCA3 9 180.0 0.84 3 ; acetaldehyde, adm CC312 CC312 CC2O4 HCR1 9 180.0 0.00 6 ; aldehdye par22 X CT1 CD X 0.0 0.84 3 ; aldehyde par22 X CT1 CT1 X CC2O4 CC312 CC312 HCA1 9 OC2D4 CC2O4 CC312 HCA1 180.0 0.00 6 ; aldehdye par22 X CT1 CD X 9

OC311 CC312 CC2O4 HCR1 9 180.0 0.00 6 ; aldehdye par22 X CT1 CD X CC312 CC312 CC312 CC322 9 180.0 0.08 1 ; erh n=6 polyol CC312 CC312 CC312 CC322 9 0.0 0.25 2 ; " RIMP2/cc-pVTZ// MP2/631gd CC312 CC312 CC312 CC322 9 0.0 10.33 3 ; " C2-C3-C4-C5 CC322 CC312 CC312 CC322 9 180.0 0.08 1; erh glycerol, n=6 polyol CC322 CC312 CC312 CC322 9 0.0 0.25 2 ; " RIMP2/cc-pVTZ//MP2/631gd CC322 CC312 CC312 CC322 9 0.0 10.33 3 ; " C2-C3-C4-C5 CC312 CC312 CC312 CC312 9 180.0 0.08 1 ; erh n=6 polyol CC312 CC312 CC312 CC312 9 0.0 0.25 2 ; " RIMP2/cc-pVTZ//MP2/631gd CC312 CC312 CC312 CC312 9 0.0 10.33 3 ; " C2-C3-C4-C5 CC3161 CC3161 CC3161 CC3161 9 180.0 9.67 3 ; erh fit inositol RIMP2/ccpVTZ//MP2/631gd C1-C2-C3-C4 CC3161 CC3161 CC3161 OC311 9 0.0 0.84 3 ; par27 X CTL1 CTL1 X CC322 CC312 CC312 OC311 9 0.0 1.46 1 ; erh n=6 polyol CC322 CC312 CC312 OC311 9 0.0 2.89 2 ; " RIMP2/cc-pVTZ//MP2/631gd CC322 CC312 CC312 OC311 9 180.0 11.67 3 ; "C1-C2-C3-O3 180.0 5.61 1 ; erh n=6 polyol OC311 CC322 CC322 OC311 9 0.0 4.98 2 ; " RIMP2/cc-pVTZ//MP2/631gd OC311 CC322 CC322 OC311 9 OC311 CC322 CC322 OC311 9 0.0 10.79 3 ; " O2-C2-C1-O1 CC312 CC312 CC322 OC311 9 0.0 1.46 1 ; erh n=6 polyol CC312 CC312 CC322 OC311 9 0.0 2.89 2 ; " RIMP2/cc-pVTZ//MP2/631gd CC312 CC312 CC322 OC311 9 180.0 11.67 3 ; "C3-C2-C1-O1 CC312 CC312 CC312 OC311 9 0.0 1.46 1 ; erh n=6 polyol CC312 CC312 CC312 OC311 9 0.0 2.89 2 ; " RIMP2/cc-pVTZ//MP2/631gd CC312 CC312 CC312 OC311 9 180.0 11.67 3 ; "C4-C3-C2-O2 CC322 CC312 CC322 OC311 9  $0.0 \quad 1.46 \quad 1$ ; erh glycerol, n=6 polyol CC322 CC312 CC322 OC311 9 0.0 2.89 2 ; " RIMP2/cc-pVTZ//MP2/631gd CC322 CC312 CC322 OC311 9 180.0 11.67 3 ; "C1-C2-C3-O3 OC311 CC3161 CC3161 OC311 9 180.0 11.09 1 ; og ethylene glycol combo OC311 CC3161 CC3161 OC311 9 0.0 0.00 2 ; " 180.0 0.54 3 ; " OC311 CC3161 CC3161 OC311 9 OC311 CC312 CC322 OC311 9 180.0 5.61 1 ; erh n=6 polyol OC311 CC312 CC322 OC311 9 0.0 4.98 2 ; " RIMP2/cc-pVTZ//MP2/631gd OC311 CC312 CC322 OC311 9 0.0 10.79 3 ; " O2-C2-C1-O1 OC311 CC312 CC312 OC311 9 180.0 5.61 1 ; erh n=6 polyol OC311 CC312 CC312 OC311 9 0.0 4.98 2 ; " RIMP2/cc-pVTZ//MP2/631gd OC311 CC312 CC312 OC311 9 0.0 10.79 3 : " O2-C2-C3-O3 HCP1 OC311 CC3161 HCA1 9  $0.0 \quad 0.75 \quad 3$ ; og methanol 0.0 0.75 3; og methanol HCP1 OC311 CC322 HCA2 9 HCP1 OC311 CC312 HCA1 9 0.0 0.75 3 ; og methanol HCA1 CC3161 CC3161 OC311 9 0.0 0.59 3 ; og ethanol 0.0 0.59 3 ; og ethanol HCA2 CC322 CC312 OC311 9 0.0 0.59 3 ; og ethanol HCA2 CC322 CC322 OC311 9 HCA1 CC312 CC322 OC311 9 0.0 0.59 3; og ethanol HCA1 CC312 CC312 OC311 9 0.0 0.59 3 ; og ethanol HCA1 CC3161 CC3161 HCA1 9 0.0 0.84 3 ; par27 X CTL1 CTL1 X CC312 CC322 HCA2 0.0 0.84 3 ; par27 X CTL1 CTL2 X HCA1 9
HCA1 CC312 CC312 HCA1 9	0.0 0.84 3 ; par27 X CTL1 CTL1 X
HCA2 CC322 CC322 HCA2 9	0.0 0.84 3 ; par27 X CTL1 CTL1 X
HCA1 CC312 CC2O4 HCR1 9	180.0 0.00 6 ; par22 X CT3 CD X
HCA1 CC3161 CC3161 CC3161 9	0.0 0.84 3 ; par27 X CTL1 CTL1 X
HCA1 CC312 CC312 CC322 9	0.0 0.84 3 ; par27 X CTL1 CTL1 X
CC312 CC312 CC312 HCA1 9	0.0 0.84 3 ; par27 X CTL1 CTL1 X
CC312 CC312 CC322 HCA2 9	0.0 0.84 3 ; par27 X CTL1 CTL2 X
CC322 CC312 CC322 HCA2 9	0.0 0.84 3 ; par27 X CTL1 CTL2 X
CC312 CC322 OC311 HCP1 9	0.0 1.46 1 ; erh n=6 polyol
CC312 CC322 OC311 HCP1 9	0.0 1.55 2 ; " RIMP2/cc-pVTZ//MP2/631gd
CC312 CC322 OC311 HCP1 9	180.0 0.79 3 ; " С2-С1-О1-НО1
CC322 CC312 OC311 HCP1 9	0.0 1.46 1 ; erh n=6 polyol
CC322 CC312 OC311 HCP1 9	0.0 1.55 2 ; " RIMP2/cc-pVTZ//MP2/631gd
CC322 CC312 OC311 HCP1 9	180.0 0.79 3 ; " С1-С2-О2-НО2
CC312 CC312 OC311 HCP1 9	0.0 1.46 1 ; erh n=6 polyol
CC312 CC312 OC311 HCP1 9	0.0 1.55 2 ; " RIMP2/cc-pVTZ//MP2/631gd
CC312 CC312 OC311 HCP1 9	180.0 0.79 3 ; " С2-С3-О3-НОЗ
CC3161 CC3161 OC311 HCP1 9	0.0 1.21 1 ;
CC3161 CC3161 OC311 HCP1 9	0.0 2.59 2 ;
CC3161 CC3161 OC311 HCP1 9	0.0 0.21 3 ;
HCA3 CC322 OC311 HCP1 9	180.0 0.79 3 ; "
;	
[ dihedraltypes ]	
;i j	k l func q0 cq
;POLYOLS	
CC2O4 CC331 OC2D4 HCR1 2	0.00 209.20 ; acetaldehyde adm
CC2O4 CC312 OC2D4 HCR1 2	0.00 209.20 ; acetaldehyde adm
OC2D3 CC331 CC331 CC2O3 2	0.00 292.88 ; ketone, acetone
OC2D3 CC312 CC322 CC2O3 2	0.00 292.88 ; "
;	