



**UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE QUÍMICA**

VINICIUS PICCOLI

**ESTUDO DA SOLVATAÇÃO DE PROTEÍNAS POR LÍQUIDOS IÔNICOS USANDO
FUNÇÕES DE DISTRIBUIÇÃO DE MÍNIMA-DISTÂNCIA**

CAMPINAS

2020

VINICIUS PICCOLI

**ESTUDO DA SOLVATAÇÃO DE PROTEÍNAS POR LÍQUIDOS IÔNICOS USANDO
FUNÇÕES DE DISTRIBUIÇÃO DE MÍNIMA-DISTÂNCIA**

**Dissertação de Mestrado apresentada ao Instituto de
Química da Universidade Estadual de Campinas como
parte dos requisitos exigidos para a obtenção do título
de Mestre em Química na área de Físico-Química**

Orientador: Prof. Dr. Leandro Martínez

O arquivo digital corresponde à versão final da Dissertação defendida pelo aluno Vinicius Piccoli e orientada pelo Prof. Dr. Leandro Martínez.

CAMPINAS

2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Química
Simone Luiz Alves - CRB 8/9094

P581e Piccoli, Vinicius, 1996-
Estudo da solvatação de proteínas por líquidos iônicos usando funções de distribuição de mínima-distância / Vinicius Piccoli. – Campinas, SP : [s.n.], 2020.

Orientador: Leandro Martínez.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Química.

1. Solvatação. 2. Líquidos iônicos. 3. Proteínas. 4. Dinâmica molecular. I. Martínez, Leandro, 1979-. II. Universidade Estadual de Campinas. Instituto de Química. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Study of protein solvation by ionic liquids using minimum-distance distribution functions

Palavras-chave em inglês:

Solvation
Ionic liquids
Proteins
Molecular dynamics

Área de concentração: Físico-Química

Titulação: Mestre em Química na área de Físico-Química

Banca examinadora:

Leandro Martínez [Orientador]
Mauro Carlos Costa Ribeiro
Rene Alfonso Nome Silva

Data de defesa: 28-02-2020

Programa de Pós-Graduação: Química

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0003-3379-8760>
- Currículo Lattes do autor: <http://lattes.cnpq.br/8358573667119731>

BANCA EXAMINADORA

Prof. Dr. Leandro Martínez (Orientador)

Prof. Dr. Mauro Carlos Costa Ribeiro (Instituto de Química - USP)

Prof. Dr. René Alfonso Nome Silva (Instituto de Química - UNICAMP)

A Ata da defesa assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Este exemplar corresponde à redação final da Dissertação de Mestrado defendida pelo aluno Vinicius Piccoli, aprovada pela Comissão Julgadora em 28/02/2020.

***Os que se encantam com a prática sem a ciência são como os timoneiros
que entram no navio sem timão nem bússola, nunca tendo certeza do
seu destino.***

Leonardo da Vinci

Agradecimentos

Antes de qualquer coisa, eu agradeço a Deus. Apesar de muitas situações que passamos na vida parecerem incompreensíveis ou difíceis de aceitar, crer em um Deus justo e amoroso me dá forças para seguir em frente. Por isso, eu agradeço a Deus por tudo de bom que aconteceu comigo durante o período de elaboração deste trabalho.

Agradeço, em especial, ao professor Leandro Martínez. O leandro, com toda certeza, foi extremamente importante para a realização deste trabalho. Em todas as etapas, ele sempre foi acessível e me ajudou muito (algumas vezes praticamente pegando na minha mão para me ensinar). Eu admiro muito o Leandro como pesquisador e espero continuar contando com a sua orientação durante o doutorado.

Agradeço a todos os membros do nosso grupo de pesquisa. Àqueles que já foram embora (Gabriel, Luciano, Antonio, Alvaro e Rafael) e, principalmente, àqueles que ainda estão aqui e com quem convivi a maior parte do tempo durante do mestrado (Diego, Guilherme, Brenda, Tayane e Adriano).

Agradeço também à minha família (Valdecir Piccoli e Maria Anete de Souza Piccoli). Em casa somos apenas em três pessoas, mas são três pessoas muito unidas. Meus pais são meu porto seguro e minha fonte de amor e carinho. Sempre me apoiaram e me ajudaram da melhor forma que puderam, talvez até mais do que eu realmente merecia. É muito difícil estar longe de casa, mas com todo amor, carinho e confiança que depositaram em mim, tenho certeza que tudo dará certo.

Agradeço aos meus amigos Letícia, Marcos Vinícius, Michel, Stepherson, Victor e Vinicius por sempre proporcionarem conversas divertidas. Mesmo que eu veja alguns apenas uma vez ao ano, quando nos encontramos sempre é um momento muito bom.

Agradeço também à Maria Eduarda por toda a companhia e carinho. Ela é uma pessoa muito alegre, divertida e amável. Com toda certeza, foi a pessoa que tornou meus dias mais leves e felizes. Obrigado por todo amor e carinho.

Agradeço também à comunidade que desenvolve e distribui programas gratuitamente. No começo do mestrado eu fui inserido (mesmo que levemente) ao mundo da programação. Nesta transição, eu pude entender o quanto é importante a comunidade que constantemente faz inúmeros programas para variadas aplicações de maneira totalmente grátis. Quase tudo nesta dissertação foi feito em programas livres.

Agradeço ao CNPq pelo financiamento deste trabalho (processo 130558/2018-4) e de muitos outros em todo o Brasil.

Agradeço à CAPES pelo apoio ao Programa de Pós-Graduação em Química do Instituto de Química da Unicamp (código de financiamento 001).

Por fim, agradeço ao CEPID (Centro de Pesquisa, Inovação e Difusão) CCE&S - *Center for Computing in Engineering and Sciences*, apoiado financeiramente pela FAPESP (Processos 2010/16947-9, 2013/05475-7, 2013/08293-7). Os cálculos realizados nesta dissertação foram feitos utilizando a infra-estrutura fornecida pelo CCE&S.

Resumo

Os líquidos iônicos são solventes, formados por íons, com notáveis características físico-químicas (baixa pressão de vapor, alta estabilidade térmica, inúmera variabilidade de compostos possíveis. etc.). As diversificadas propriedades que os líquidos iônicos detêm proporcionam sua utilização em aplicações industriais e em variados campos de pesquisa, como a área biotecnológica. Alguns trabalhos tratam da estabilização/desestabilização de proteínas em meios contendo líquidos iônicos e água. Contudo, ainda são limitados os trabalhos experimentais e teóricos que descrevem os mecanismos pelos quais líquidos iônicos e proteínas interagem.

Neste trabalho, a estrutura de solvatação da ubiquitina (em sua conformação nativa e em diversas conformações desenoveladas) por líquidos iônicos imidazólicos foi estudada computacionalmente por meio de simulações de dinâmica molecular e funções de distribuição de mínima-distância. A teoria de soluções de Kirkwood-Buff foi empregada para, a partir das funções de distribuição, calcular quantidades que representam o acúmulo ou a exclusão das moléculas do solvente nos domínios da proteína.

Os líquidos iônicos apresentaram maior acumulação, em relação à água, sobre a superfície das conformações simuladas. Isto, principalmente para o intervalo de 0,5 e 1,5 mol L⁻¹. Observou-se que nas estruturas mais abertas o acúmulo era particularmente maior do que comparado com a conformação nativa da ubiquitina. O ânion DCA apresentou maior acúmulo sobre a proteína em comparação ao BF₄. Além disso, os sistemas com o cátion BMIM demonstram maior acumulação do líquido iônico do que àqueles com o cátion EMIM. Constatou-se que os líquidos iônicos (principalmente àquelas que contém o ânion DCA) solvatam a ubiquitina preferencialmente.

Abstract

Ionic liquids are molten salts (solvents only formed by ions that have melting points lower than 100 °C) with incredible physicochemical properties (low vapor pressure, high thermal stability, variable composition, etc.). Ionic liquids have been used in industry and academy as solvents, catalysts, and cosolvents in systems with proteins. Researchers are using ionic liquids as refolding/denaturing agents of proteins in a large number of systems. Nevertheless, the number of theoretical and experimental studies that describe the interaction between proteins and ionic liquids are still small.

In this thesis, the structure of ubiquitin solvation by ionic liquids was computationally studied by molecular dynamics simulation and minimum-distance distribution functions. The Kirkwood-Buff theory of solutions was applied for the calculation of thermodynamic quantities from minimum-distance distribution functions calculated from simulations.

The protein was preferentially solvated by the ions from the ionic liquids, especially at the lower concentrations simulated. It was noted that the ions had a greater preferential solvation parameters when solvating denaturated and extended ubiquitin conformations. Besides that, we observed that the presence of different cations, for example, causes changes in the anion distribution functions. This fact suggests that there exists a correlation of the distribution of the ions on protein solvation by ionic liquids.

Lista de Figuras

- 1.1 Carga deslocalizada no anel imidazólico. Este cátion é frequentemente escolhido para compor diversos LIs. R1 e R2 são grupos alquílicos, podendo variar de um metil até octil. 21
- 2.1 Configuração que representa o cálculo da função de distribuição radial. A partícula vermelha é a partícula de referência, a partir desta partícula tomamos uma distância r . Esta distância é entre centros de massa, especificamente, entre o centro de massa da bola vermelha e da bola azul apontada pelas setas. 28
- 2.2 Funções de distribuição radiais para a água obtida a partir das técnicas de difração de raios X e espalhamento de nêutrons [36, 37]. A seção de choque do hidrogênio é maior que a do oxigênio, por conta disto o espalhamento de Nêutrons observa o átomo de hidrogênio preferencialmente, e o pico fica mais deslocado à esquerda. . 30
- 2.3 A - Representação da particularidade existente no cálculo das RDFs para solutos que possuem forma muito diferente de uma esfera. B - Exemplificação da ideia básica contida nas funções de distribuição de mínima distância. Contando as mínimas-distâncias r podemos obter a função de distribuição do solvente independentemente do formato do soluto. 31
- 2.4 Esquema simplificado apresentando como as MMDFs são calculadas. Em vermelho temos o soluto, as demais moléculas são águas. As distâncias assinaladas no desenho representam as menores distância entre a superfície do soluto e um átomo das moléculas do solvente. 32

2.5	Representação das funções de distribuição radial e de mínima distância para o ânion DCA ao redor da ubiquitina na conformação nativa. As funções de distribuição foram calculadas a partir da mesma simulação. Podemos notar que RDF apresenta um comportamento que é muito difícil de interpretar. Contudo, a MDFF apresenta picos com formatos que possibilitam uma descrição mais intuitiva do sistema.	33
2.6	Esquema simplificado de um experimento de diálise no equilíbrio. O fundo azul representa a água, os polígonos vermelhos são as moléculas de cossolvente e os desenhos amarelos são as biomoléculas.	36
2.7	Representação do sistema que descreve a adição de LI em uma solução aquosa com proteína. Em (A) temos um sistema hipotético que contém a água e a proteína em solução, nesta configuração a proteína tem um determinado potencial químico (μ_P^0). Em (B) o LI é colocado no sistema e se espalha pela solução, passando a interagir com a proteína. Aqui o potencial químico da proteína foi perturbado, de maneira que acarretou em acúmulo de LI na superfície proteica. Nesta situação, Γ_{uc} seria maior que zero e o novo potencial químico da proteína seria menor do que no sistema (A), levando em conta a equação 2.2.8.	38
4.1	A imagem apresenta, na cor vermelha, os três aminoácidos prolina que existem na ubiquitina nativa. Na cor azul, estão representadas as três alaninas que foram colocadas no lugar das prolinas.	48
4.2	Fórmulas estruturais dos cátions e ânions que compõem os líquidos iônicos utilizados.	49
4.3	Integral de Kirkwood-Buff para um sistema contendo o líquido iônico EMIMDCA. A curva foi calculada integrando a função de distribuição da água ao redor da proteína ao longo de toda a solução. O valor utilizado para os cálculos dos parâmetros de solvatação preferencial é aquele onde há uma constância na curva da integral de Kirkwood-Buff, representado por G_{uw} , sendo que u corresponde ao soluto e w à água. Para este caso, a integral de KB começa a ser constante a partir de 20 Å.	52
5.1	Função de distribuição de mínima distância para a água em relação à ubiquitina. O maior pico em 1,8 Å indica que a densidade da água nessa distância é aproximadamente duas vezes maior que a de bulk	53

5.2	Decomposição das funções de distribuição de mínima distância para a água em suas contribuições atômicas. A decomposição consiste em separar a MDDF total em contribuições atômicas, ou seja, retiramos a informação do quanto determinado átomo contribui para a MDDF.	54
5.3	Decomposição da MDDF que descreve a distribuição água-proteína em contribuições do backbone e de cadeias laterais específicas.	55
5.4	Integral de Kirkwood-Buff para a água no sistema apenas com a ubiquitina no estado nativo.	56
6.1	Parâmetros de solvatação preferencial para o ânion (A), o cátion (B) e a água (C) em sistemas com os LIs EMIMDCA e EMIMBF4. Podemos observar um crescimento inicial de Γ para os íons (até cerca de $1,0 \text{ mol L}^{-1}$), para o DCA temos uma tendência que não segue exatamente um aumento contínuo, isto é, muito provavelmente, devido à variabilidade da medida. A água diminui sua exclusão com o aumento da concentração de LI. Para o LI EMIMBF4, Γ_{uc} apresenta um leve aumento entre $0,5$ e $1,5 \text{ mol L}^{-1}$	58
6.2	MDDFs da água para os sistemas com a proteína no estado nativo e os cossolventes EMIMDCA (A) e EMIMBF4 (B).	59
6.3	Parâmetro de solvatação preferencial para o ânion (A), cátion (B) e água (C) em sistemas contendo BMIMDCA e BMIMBF4.	59
6.4	Funções de distribuição para o cátion nos sistemas contendo o ânion DCA (A) e BF4 (B). O pico em $2,5 \text{ \AA}$ é maior conforme a concentração do líquido iônico é menor, indicando que a densidade relativa ao bulk do cátion nesta distância é maior nas menores concentrações.	61
6.5	MDDFs para o cátion BMIM nos sistemas contendo os ânions DCA (A) e BF4 (B). O pico em $2,5 \text{ \AA}$ aumenta com a diminuição da concentração do LI, indicando que há um acúmulo do cátion na superfície da proteína.	61

6.6	MDDFs para o DCA em sistemas contendo diferentes concentrações do LI EMIMDCA. A ordem de altura dos picos segue a mesma tendência daquela observada para as MDDFs dos cátions. A altura do pico aumenta com a diminuição da concentração de LI no sistema. Aqui, temos a presença de dois picos pronunciados, ambos tem formatos semelhantes, sendo o segundo (2,6 Å) ligeiramente maior do que o primeiro (1,9 Å).	62
6.7	MDDFs para o BF4 em sistemas contendo diferentes concentrações do LI EMIMBF4. Similarmente ao que temos na Figura 6.6, os picos são mais altos nas menores concentrações de LI. Além disso, o pico referente às ligações de hidrogênio é bem menor do que àquele devido às interações não-específicas. . . .	63
6.8	Função de distribuição decomposta nas contribuições parciais para os átomos. A figura apresenta a MDDF para o DCA em um sistema com 0,5 mol L ⁻¹ de EMIMDCA.	63
6.9	Função de distribuição decomposta nas contribuições parciais dos átomos do cossolvente. A figura apresenta a MDDF para o BF4 num sistema com 0,5 mol L ⁻¹ de EMIMBF4. Como a única contribuição para a MDDF total advém dos átomos de Flúor, a MDDf do flúor sobrepõe à total.	64
6.10	Integrais de Kirkwood-Buff para o cátion (A), o ânion (B) e a água (C) em sistemas contendo variadas concentrações de EMIMDCA.	65
6.11	Integrais de Kirkwood-Buff para o cátion (A), o ânion (B) e a água (C) em sistemas contendo variadas concentrações de EMIMBF4	66
7.1	Diferentes conformações simuladas para a ubiquitina. (A) corresponde à estrutura nativa da ubiquitina, (B) é uma das 30 conformações parcialmente desenoveladas utilizadas no trabalho e (C) é um corte da conformação estendida.	68
7.2	Γ para os líquidos iônicos na presença da proteína no estado de enovelamento desnaturado. A e B são os parâmetros de solvatação preferencial para os íons (Γ_{uc}) em sistemas contendo EMIMDCA e EMIMBF4. C e D contém Γ_{uc} para os íons nos sistemas com BMIMDCA e BMIMBF4. É possível notar que em todos os sistemas apresentados, Γ_{uc} aumenta de 0,5 mol L ⁻¹ até 1,5 mol L ⁻¹ . Contudo, o aumento é menos pronunciado para o sistemas contendo o ânion BF4	69

7.3	Parâmetro de solvatação preferencial da água para sistemas contendo EMIMDCA e EMIMBF4 (A), BMIMDCA e BMIMBF4 (B) e a Ubiquitina em seu estado de enovelamento desnaturado.	70
7.4	MDDFs para o cátion BMIM, na presença do ânion DCA, em sistemas contendo os estados de enovelamento nativo (A), desnaturado (B) e estendido (C).	71
7.5	MDDFs para o ânion DCA, na presença do cátion BMIM, em sistemas contendo os estados de enovelamento nativo (A), desnaturado (B) e estendido (C).	72
7.6	Contribuição de cada tipo de resíduo e da cadeia principal da proteína para a função de distribuição total do cátion (sistema com 0,5 mol L ⁻¹ de EMIMDCA). Para o cálculo da contribuição da cada tipo de resíduo para a distribuição total tomou-se a distância de 2,6 Å (distância em que a MDDF para o EMIM apresenta o pico mais representativo). Como exemplo de interpretação, o gráfico indica que para a proteína nativa a função de distribuição do cátion ao redor do backbone corresponde a 15,8 % da função de distribuição ao redor da proteína toda.	72
7.7	A figura (A) contém as MDDFs para o cátion EMIM em sistemas com o 0,5 mol L ⁻¹ do LI EMIMDCA nas conformações desnaturada e estendida. (B) contém as IKBs para EMIM nos mesmos sistemas.	73
7.8	A figura contém Γ_{uc} os íons dos LIs EMIMDCA e EMIMBF (A e B) e dos LIs BMIMDCA e BMIMBF4 (C e D) nos sistemas que contém a conformação estendida da proteína.	74
7.9	A figura contém a MDDF em 3D. Aqui, temos a representação, até uma distância de 2,0 Å da proteína, dos átomos que contribuem para a MDDF. Em verde temos a representação da distribuição do DCA na proteína nativa (A) e desnaturada (B). Os pontos vermelhos representam a distribuição do EMIM. Em função da distância que está sendo representada, observamos nas figuras as ligações de hidrogênio (regiões verdes).	75
7.10	MDDFs para o ânion DCA em diferentes concentrações para os sistemas contendo a ubiquitina em sua conformação estendida (A). Decomposição da MDDF total em contribuições dos resíduos da proteína (B).	76

8.1	Parâmetro de solvatação preferencial para o ânion DCA na presença da ubiquitina nativa, em sistemas com 0,5 mol L ⁻¹ de EMIMDCA e BMIMDCA.	79
8.2	Parâmetro de solvatação preferencial em sistemas contendo a proteína em seu estado nativo e 0,5 mol L ⁻¹ dos líquidos iônicos EMIMBF ₄ e BMIMBF ₄ . O aumento da cadeia carbônica do cátion (EMIM para BMIM) é acompanhada da diminuição Γ_{uw} para concentrações menores do que 2,0 mol L ⁻¹	80
8.3	Funções de distribuição de mínima distância para o DCA nos sistemas com EMIMDCA (A) e BMIMDCA (B). Nota-se que que MDFF no sistema (B) apresenta picos maiores, ou seja, no sistema (B) é maior a densidade relativa do ânion DCA ao redor da proteína em comparação com o sistema (A).	81
8.4	Funções de distribuição de mínima distância para o BF ₄ nos sistemas com EMIMBF ₄ (A) e BMIMBF ₄ (B). A mudança, na altura dos picos, resultante da troca do cátion EMIM para BMIM é muito pequena	81
8.5	IKBs para o DCA nos sistemas contendo o EMIMDCA (A) e o BMIMDCA (B). O ânion DCA apresenta uma IKB muito maior no sistema com o cátion BMIM.	82
8.6	IKBs para o BF ₄ nos sistemas contendo EMIMBF ₄ (A) e BMIMBF ₄ (B). O aumento do tamanho da cadeia carbônica do cátion culminou num pequeno aumento da IKB para o BF ₄ no sistema com BMIM.	82
8.7	MMDFs para o cátion EMIM nos sistemas contendo BF ₄ (A) e DCA (B).	83
8.8	MMDFs para o cátion BMIM nos sistemas contendo BF ₄ (A) e DCA(B).	83
8.9	Em (A) temos as MDFFs para o cátion EMIM e o ânion DCA, (B) contém as IKBs para os mesmos componentes.	85

Lista de Tabelas

- 7.1 Área acessível ao solvente para os diferentes estados de enovelamento simulados da ubiquitina. A área calculada para o estado de enovelamento desnaturado compreende uma média das 30 conformações (diferentes) usadas para representar o estado desnaturado. 68
- 7.2 Variação do parâmetro de solvatação preferencial para cada componente dos sistemas contendo EMIMBF₄. A variação foi calculada utilizando os sistemas com as conformações parcialmente desenovelada e nativa. Os dados indicam forte interação do líquido iônico com a conformação parcialmente desenovelada se comparada à proteína nativa. 77

Lista de Abreviaturas

MDDF	Em inglês, <i>Minimum-Distance Distribution Function</i> , ou funções de Distribuição de Mínima Distância.
IKB	Integral de Kirkwood-Buff.
TKB	Teoria de Kirkwood-Buff.
RDF	Em inglês, <i>Radial Distribution Function</i> , ou Função de Distribuição Radial.
LI	Líquido iônico.
TMAO	O N-óxido de trimetilamina, é um composto orgânico com fórmula química $(\text{CH}_3)_3\text{NO}$.
DCA	Dicianamida.
BF ₄	Tetrafluorborato.
EMIM	1-etil-3-metilimidazol.
BMIM	1-butil-3-metilimidazol.
DM	Dinâmica molecular.
PME	Em inglês, <i>Particle Mesh Ewald Sum</i> .
P^3E	Em inglês, <i>Particle-particle-particle Ewald..</i>
RMN	Ressonância magnética nuclear.
SASA	Em inglês, <i>Solvent Accessible Surface Area</i> ou Área Superficial Acessível ao Solvente.
OPLS	Em inglês, <i>Optimized Potentials for Liquid Simulations</i> .
OPLS-VSIL	Em inglês, <i>Optimized Potentials for Liquid Simulations - Virtual Site Ionic Liquid</i> .

Sumário

1	Líquidos iônicos	20
1.1	Propriedades e aplicações gerais	20
1.2	Líquidos iônicos como cossolventes em sistemas com biomoléculas	22
1.2.1	Líquidos iônicos e proteínas	24
2	Teoria de Soluções	26
2.1	Estrutura de solvatação ao redor de biomoléculas	26
2.1.1	Funções de distribuição de mínima distância	31
2.2	Integrais de Kirkwood-Buff e Parâmetro de Solvatação Preferencial	33
2.2.1	Teoria de soluções de Kirkwood-Buff	33
2.2.2	Parâmetro de Solvatação Preferencial	35
3	Dinâmica molecular	39
3.1	Visão geral do método	39
3.2	Etapas básicas de uma simulação de dinâmica molecular	41
3.2.1	Montagem da caixa de simulação	41
3.2.2	Cálculo das forças	41
3.2.3	Movimentação das partículas e ensembles estatísticos	42
3.2.4	Condições iniciais e etapas da simulação	43
4	Objetivos e Metodologia	46
4.1	Abordagem e objetivos	46
4.1.1	Visão geral da abordagem do sistema	46
4.2	Metodologia Desenvolvida	47

4.2.1	Montagem das caixas de simulação	47
4.2.2	Dinâmica molecular	50
4.2.3	Cálculo das funções de distribuição de mínima-distância	51
5	Distribuição da água	53
5.1	Solvatação da Ubiquitina pela água	53
6	Variação da concentração de líquido iônico	57
6.1	Efeito relativo à variação da concentração	57
7	Alteração do estado de enovelamento da ubiquitina	67
7.1	Efeito na acumulação de líquido iônico sobre a proteína relativo ao estado de enovelamento	67
8	Efeito da troca de ânion e de cátion	78
8.1	Parâmetros de solvatação preferencial dos íons	78
8.2	Funções de distribuição de mínima-distância para os íons	80
9	Considerações Finais	86
9.1	Variação da concentração	86
9.2	Estado de enovelamento da Ubiquitina	87
9.3	Variação do tipo de cátion/ânion	88
9.4	Perspectivas futuras	89
	Bibliografia	90

Capítulo 1

Líquidos iônicos

Os líquidos iônicos são compostos estudados desde a segunda metade no século XX. Em 1970, Brockis e Reddy publicaram um livro intitulado *Modern Electrochemistry*, o capítulo 6 deste livro tinha o título de "Ionic Liquids". Este capítulo tratava de sais de tetraalquilamônio com silicatos e haletos [1]. Além disso, foi um dos primeiros capítulos de livro que tratavam, especificamente, de sais no estado líquido. No entanto, a era moderna dos líquidos iônicos começou com estudos sobre sais de alquilpiridínio e dialquilimidazol. Neste capítulo, apresentaremos o que são os líquidos iônicos. Faremos uma breve revisão sobre as suas características físico-químicas e as aplicações mais comuns na indústria e academia. Uma atenção particular será dada às aplicações dos líquidos iônicos, especialmente imidazólicos, em sistemas biológicos, como, sua utilização como cossolvente em sistemas contendo proteínas. Apresentaremos algumas razões que impulsionaram a utilização, em pesquisas, dos líquidos iônicos como cossolventes em sistemas com proteínas.

1.1 Propriedades e aplicações gerais

Uma das substâncias iônicas líquidas mais simples que podemos pensar é o NaCl fundido. O NaCl precisa ser aquecido até altas temperaturas (801°C) para ser fundido, devido à energia eletrostática de rede do sal [2]. Os líquidos iônicos (LIs), por sua vez, são substâncias iônicas no estado líquido em temperaturas em torno de 25°C. Eles são formados por cátions orgânicos, e ânions que podem ser inorgânicos ou orgânicos. Em função do tamanho dos íons, as interações coulômbicas (responsáveis, nos compostos iônicos comuns, por formar uma forte rede de ligações culminando em um composto de alto ponto de fusão) tornam-se mais fracas. O enfraquecimento do retículo cristalino tornam os LIs, de fato, líquidos em temperaturas onde um composto iônico normalmente

seria sólido [3].

A energia potencial entre duas cargas cresce com o inverso da distância entre as cargas. Os LIs são formados por cátions orgânicos que são bem maiores do que o íon Na^+ , por exemplo. Além disso, a carga do cátion é, em alguns casos, deslocalizada. Tal fato diminui ainda mais a intensidade da interação eletrostática. A Figura 1.1 apresenta um cátion encontrado nos LIs imidazólicos. Além de ser uma molécula grande, a sua carga está deslocalizada no anel [4]. A carga deslocalizada contribui muito menos para a interação eletrostática do que uma carga localizada em um íon pequeno, como ocorre no NaCl. Portanto, o enfraquecimento das ligações iônicas torna o retículo cristalino menos estável, viabilizando a existência desses compostos no estado líquido na temperatura ambiente [3].



Figura 1.1: Carga deslocalizada no anel imidazólico. Este cátion é frequentemente escolhido para compor diversos LIs. R1 e R2 são grupos alquílicos, podendo variar de um metil até octil.

Ainda que enfraquecidas, as interações iônicas em LIs são relevantes. Esta relevância é tal que os LIs possuem propriedades atípicas (se comparados com a maioria dos solventes orgânicos) [5]: alta estabilidade térmica, baixa pressão de vapor, alta capacidade de solubilizar diversos compostos, elevada polaridade, baixa toxicidade e alta viscosidade. Estas propriedades fazem dos LIs objetos de estudo em diversos tipos de aplicações industriais e acadêmicas [6, 7, 8].

Aplicações para os LIs surgem conforme o entendimento sobre esse grupo de compostos aumenta. Atualmente, os LIs são usados como lubrificantes e aditivos de combustíveis [9]. Há aplicações em dispositivos eletrolíticos e até mesmo em displays (como cristais líquidos) [10, 11]. A incrível variabilidade de compostos possíveis, aliada ao fato da possível modelagem do cátion e do ânion para aplicações específicas, fazem de seus usos tão amplos e diversos que os LIs têm espaço até em sistemas biotecnológicos [12, 13, 14, 15, 16].

O LI interage com outras substâncias de acordo com natureza dos seus íons e das outras substâncias. Por exemplo, um LI formado por íons muito hidrofóbicos será excluído de uma eventual fase aquosa. No nitrato de 1-butil-3-Metilimidazol (BMIMNO₃), o cátion é um composto orgânico grande que apresenta uma carga deslocalizada no anel imidazólico, além disso, este cátion é pouco

polar. O grupo imidazólico pode interagir com o soluto (que aqui pode ser uma biomolécula) através de interações não-específicas, como interações de van der Waals ou íon-dipolo. O ânion (NO_3^-), que é muito polar, claramente poderia interagir por interações específicas (ligações de hidrogênio) em função da existência dos átomos de oxigênio e também por interações não-específicas (dipolo-dipolo, íon-dipolo, dipolo-dipolo induzido).

É interessante mencionar o fato de que, além de um enfraquecimento geral das interações entre os íons, o fator entrópico nos líquidos iônicos é presente na diminuição do ponto de fusão destes compostos. O cátion imidazol costuma vir acompanhado de dois grupos alquil que são ligados em cada nitrogênio. As duas cadeias alquílicas conferem mais graus de liberdade para a molécula. Sendo assim, há uma vantagem entrópica, relacionada ao ganho de movimento, principalmente para os cátions (que são mais volumosos e possuem cadeias laterais), quando o sistema passa do estado sólido para o líquido.

1.2 Líquidos iônicos como cossolventes em sistemas com biomoléculas

Quando pensamos em biomoléculas (exceto lipídeos) nos referimos a, basicamente, três grupos de polímeros biológicos: polissacarídeos, ácidos nucleicos e proteínas. Os polissacarídeos têm a função de reserva energética em nosso organismo, são marcadores celulares e possuem função estrutural em plantas. Os ácidos nucleicos, por sua vez, possuem a função de armazenamento e transporte de informações. Por último, as proteínas possuem variadas funções nos organismos vivos, e são as moléculas de principal interesse neste trabalho [17, 18].

As proteínas exibem inúmeras propriedades biológicas, como no transporte de oxigênio (hemoglobina), na proteção do corpo contra organismos patogênicos (anticorpos), catalisadores de reações químicas (enzimas), receptoras de membrana, atuação na contração muscular (actina e miosina), além de serem fundamentais para o crescimento corpóreo e formação dos hormônios. Pesquisas envolvendo proteínas contribuem para a medicina, agricultura e ciência de alimentos [19, 20]. Cada proteína tem sua ação específica dependente da sua forma tridimensional e, sendo assim, é importante entender como as proteínas interagem com outras moléculas em soluções, principalmente aquosas. Um dos aspectos fundamentais deste entendimento são os mecanismos que fazem com que as proteínas mantenham ou não sua estrutura tridimensional.

As proteínas são biopolímeros que possuem diversos níveis de complexidade estrutural, levando em conta as variadas conformações espaciais que elas podem assumir. As proteínas são formadas por unidades básicas, denominadas de aminoácidos. Os aminoácidos podem ser carregados, neutros, polares ou apolares. A sequência de aminoácidos que compõe a proteína é a estrutura primária (inclui as ligações peptídicas entre resíduos adjacentes). Elementos espaciais locais como α -hélice e folha- β são a estrutura secundária. O agrupamento de elementos da estrutura secundária forma a estrutura terciária e uma aglomeração de diferentes cadeias forma a estrutura quaternária [17].

O processo de enovelamento e desenovelamento de proteínas pode ser modulado por osmólitos. Estes são moléculas pequenas e abundantes em organismos vivos, empregados no contrabalanceamento de stress causado por fatores externos, por meio de interações não-específicas e fracas [21, 22]. Em muitos casos, o stress compromete a estabilidade conformacional das proteínas, por exemplo [23, 24].

A estabilidade das proteínas pode ser perturbada pela variação da composição dos componentes da solução em que ela se encontra. Cossolventes como a glicose e alguns polióis têm a capacidade de estabilizar proteínas (mantê-las na conformação enovelada), enquanto a ureia, por exemplo, causa desestabilização (levando à desnaturação) [25]. Neste contexto, LIs vêm tornando-se potenciais cossolventes em sistemas biotecnológicos [12, 14, 15]. É interessante notar que LIs aplicados em sistemas biológicos perdem as características físico-químicas que apresentam quando estão puros (apenas com os íons) e passam a ser sais dissolvidos em meios aquosos. Diferentemente de cossolventes comuns (uréia, TMAO, polióis, etc.), os LIs são intrinsecamente formados por duas espécies químicas iônicas. Portanto, LIs modulam processos químicos diferentemente do que normalmente ocorre para cossolventes mais comuns [26].

Basicamente, o problema que buscamos entender melhor é o seguinte: temos uma solução composta por água e LI, adicionaremos à solução uma proteína. A proteína preservará a estrutura que possuía antes de ser adicionada à solução ou irá desenovelar-se assumindo outras conformações? Como discutido anteriormente, os íons estabelecem interações intermoleculares com o soluto que são dependentes da natureza de ambos. A água é muito polar e irá interagir fortemente com outras moléculas polares e que possam estabelecer ligações de hidrogênio. Desta maneira, existe uma competição entre os íons e a água pelas interações com a superfície proteica acessível

ao solvente. O problema de desenovelamento (ou não) da proteína em solução pode ser entendido por meio da relação das interações preferenciais de cada solvente com as diferentes conformações da proteína. A conformação que uma proteína assume em um organismo vivo (estrutura nativa), costuma ser a conformação mais estável. Porém, se uma solução consegue estabilizar melhor estados de enovelamento diferentes do nativo, a proteína poderá assumir outras conformações nesta solução. Esta situação culminaria na desnaturação de uma proteína.

Quanto maior a capacidade que um líquido iônico tem de estabelecer ligações de hidrogênio ou interações não-específicas com a proteína, mais provável é que exerça um efeito desestabilizante na estrutura [27, 28]. Ânions como BF_4^- , com menor capacidade de formar ligações de hidrogênio e menor nucleofilicidade, são compatíveis com enzimas, em algumas situações [29]. Por exemplo, Lozano e colaboradores estudaram a estabilidade da α -crimotripsina em reações de transesterificação na presença de LIs, neste trabalho foram usados 5 LIs imidazólicos e 4 de alquilamônio. O LI contendo o cátion BF_4^- promoveu uma das maiores estabilizações da proteína, medida em função da constante de primeira ordem de desativação da enzima e o seu tempo de meia-vida exibidos nas soluções de LIs.

1.2.1 Líquidos iônicos e proteínas

Os LIs possuem diversas propriedades físico-químicas que fazem deles compostos interessantes para diversas aplicações [30]. Em especial, a sua utilização como cossolventes em sistemas com proteínas tem relevante aplicabilidade por conta de, em resumo, dois motivos: 1) Os LIs apresentam efeitos relevantes sobre a estrutura de proteínas em concentrações menores do que os cossolventes usuais [6, 26]. 2) Os LIs são formados por um cátion e um ânion, de forma que é possível sintetizar LIs que possuam um íon cuja ação acontece em algum sítio específico da proteína. Portanto, é importante entender a distribuição dos íons em solução e também seu comportamento com respeito à proteína. Os cátions, por exemplo, podem apresentar maior interação com domínios apolares. Isto deve-se a baixa polaridade que a maioria dos cátions apresenta [14, 13].

Diversos estudos tratam de sistemas onde proteínas são solubilizadas em soluções aquosas, com ou sem cossolventes. Devido às características hidrofóbicas, hidrofílicas e anfifílicas,

a água e os cossolventes (no caso deste trabalho, líquidos iônicos), interagem com sítios não-polares, polares e carregados das proteínas, resultando em acumulação ou exclusão desses componentes de diferentes regiões em torno da proteína [31, 32, 33].

Capítulo 2

Teoria de Soluções

Neste capítulo buscaremos compreender os fundamentos teóricos do estudo da estrutura de solvatação ao redor de biomoléculas, com especial destaque para as proteínas. Mostraremos que a forma com que as moléculas do solvente distribuem-se na solução pode ser descrita por meio de funções de distribuição. Apresentaremos dois tipos de funções de distribuição: radial e de mínima-distância. Vamos discorrer sobre a Teoria de Soluções de Kirkwood-Buff, que permite calcular diretamente o acúmulo ou a exclusão dos componentes da solução sobre o soluto. Apresentaremos o parâmetro preferencial de solvatação, que é uma forma de medir experimentalmente acumulação de osmólitos sobre macromoléculas. Além disso, mostraremos como o parâmetro de solvatação preferencial pode ser calculado utilizando as integrais de Kirkwood-Buff.

2.1 Estrutura de solvatação ao redor de biomoléculas

Como previamente discutido, as proteínas têm suas funções biológicas vinculadas à sua estrutura tridimensional. Além disso, a compreensão das funcionalidades de uma proteína dá-se pela sua identificação e análises de seu comportamento em variados meios. Para estudar uma proteína específica a partir de uma amostra biológica é necessário separá-la do meio que se encontra. As separações são baseadas na solubilidade que as proteínas podem ter nos mais variados meios. Desta forma, é muito importante entender as razões e os mecanismos que fazem com que as proteínas fiquem dissolvidas.

Em função da grande complexidade estrutural exibida pelas proteínas, faz-se necessário compreender como esses biopolímeros interagem com outras moléculas do solvente em solução.

Vamos analisar como ocorre a dissolução de um soluto genérico em uma solução aquosa. As substâncias dissolvem-se espontaneamente em um solvente quando há uma diminuição na energia livre de Gibbs associada ao processo. Portanto, a dissolução deve resultar em aumento da entropia total ou diminuição da entalpia (liberação de energia na forma de calor), ou ambos. Por exemplo, o NaCl é dissolvido na água pelo balanço entalpia e entropia. A dissolução, neste caso, é um processo endotérmico. Portanto, a dissolução só ocorre por conta do aumento da entropia. A variação da energia livre de Gibbs é dada pela seguinte equação:

$$\Delta G = \Delta H - T\Delta S \quad (2.1.1)$$

onde ΔH é a variação da entalpia, ΔS é a variação da entropia e T é a temperatura.

Um processo é espontâneo quando variação da energia livre de Gibbs é negativa. Logo, devemos avaliar ΔH e ΔS . Se o NaCl é dissolvido em água, sua estrutura, que antes era organizada, passa a ter um certo grau de desordem em função da dispersão dos íons no solvente. Nesta situação, o processo de solubilização leva a $\Delta S > 0$. A variação da entropia irá depender do que acontecerá com a estrutura do solvente posteriormente a adição de um soluto (podendo ser muito complexo de avaliar). A contribuição das interações intermoleculares está contida na ΔH . Se a solubilização for um processo exotérmico, teremos $\Delta H < 0$ (não é o caso do NaCl). A adição de um cossolvente em uma solução altera a forma com que as moléculas interagem entre si.

Podemos pensar, de maneira geral, num raciocínio parecido sendo aplicado para solutos complexos como as proteínas. Sendo assim, ΔH e ΔS poderiam ser alterados na dissolução de uma proteína pela adição de cossolventes. Os estudos computacionais são uma forma de observar microscopicamente o comportamento de proteínas em solução. A estrutura de solvatação ao redor das proteínas pode ser estudada calculando a densidade de um determinado componente da solução em função da distância a partir da proteína. Para isto, calculamos as funções de distribuição.

As funções de distribuição descrevem como a densidade numérica das partículas da solução varia em função da distância, a partir de uma partícula de referência. Dizemos que a função de distribuição (representada por $g(r)$) descreve a probabilidade de encontrar uma partícula dentro de um elemento de volume dr e a uma distância r do sítio de referência, em relação à probabilidade em uma distribuição aleatória com a mesma densidade total do solvente [34, 35].

As funções de distribuição radiais (RDFs) são geralmente calculadas levando em conta a distância entre centros de massa (na verdade, se o cálculo for feito utilizando qualquer sítio de referência que seja único, também será uma função de distribuição radial) de uma partícula de referência e o centro de massa de uma partícula contida no elemento de volume dr , como apresentado na Figura 2.1.

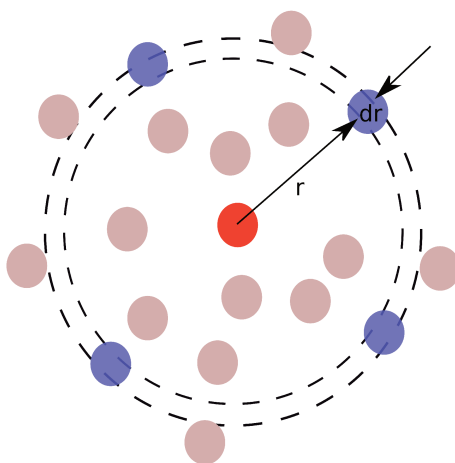


Figura 2.1: Configuração que representa o cálculo da função de distribuição radial. A partícula vermelha é a partícula de referência, a partir desta partícula tomamos uma distância r . Esta distância é entre centros de massa, especificamente, entre o centro de massa da bola vermelha e da bola azul apontada pelas setas.

Podemos entender melhor as RDFs pensando na seguinte abordagem qualitativa. Supondo que pudéssemos estar dentro de um sistema com diversas moléculas de água (a água está no estado líquido) e que, nesta situação, escolhêssemos aleatoriamente o centro de uma molécula para ficar em cima. Apenas olhando as vizinhanças, provavelmente não seria possível observar nenhum padrão na distribuição das moléculas de água. Contudo, se contarmos o número de centros das moléculas que aparecem dentro de uma casca esférica de largura dR , a uma distância R do centro da partícula onde estamos, poderemos encontrar regularidades. Cada casca esférica possui um volume dV que é igual a $4\pi R^2 dR$. Considerando que a densidade do líquido é $\rho = N/V$ (N é o número de partículas no sistema e V é o volume das partículas), devemos esperar um número médio de partículas igual a $\rho 4\pi R^2 dR$ (neste caso, partículas com centro de massa dentro de uma casca esférica de raio R). No entanto, estar em cima de uma partícula equivale em adicionar outra partícula no sistema. O que é observado, de cima da partícula, corresponde a uma densidade média. Esta densidade é um desvio da densidade que é esperada para a solução (ρ). Podemos, dessa forma, pensar em $g(R)$ levando em conta a densidade média

de moléculas dentro de uma casca de raio R . A densidade de moléculas dentro da casca será igual a $\rho g(R)$. Então, o número médio de moléculas que realmente existem (ao redor da molécula que estamos sentados) dentro da casca esférica é $\rho g(R)4\pi R^2 dR$. Além da definição de $g(R)$ que foi dada anteriormente no texto, podemos pensar que $g(R)$ é uma medida do desvio da contagem que fizemos (ou seja, contando de cima de uma molécula aleatória que é centro de uma casca esférica de raio R) em relação ao número médio de partículas que esperaríamos encontrar em uma casca esférica centrada em uma partícula aleatória no líquido.

Resumidamente, a densidade média de um líquido em qualquer ponto da solução é chamada densidade bulk (ρ). A função de distribuição, $g(r)$, é, portanto, dada pela densidade local e pela densidade bulk da solução de acordo com:

$$g(r) = \frac{\rho(r)}{\rho^{(bulk)}} \quad (2.1.2)$$

A função de distribuição radial (RDF) contém a informação do empacotamento das moléculas, por exemplo, na água líquida. RDFs podem se experimentalmente obtidas por meio de dados provenientes de técnicas como difração de raios X (DRX) e espalhamento de nêutrons. A Figura 2.2 contém dados experimentais para a função de distribuição da água na temperatura de 25°C [36, 37].

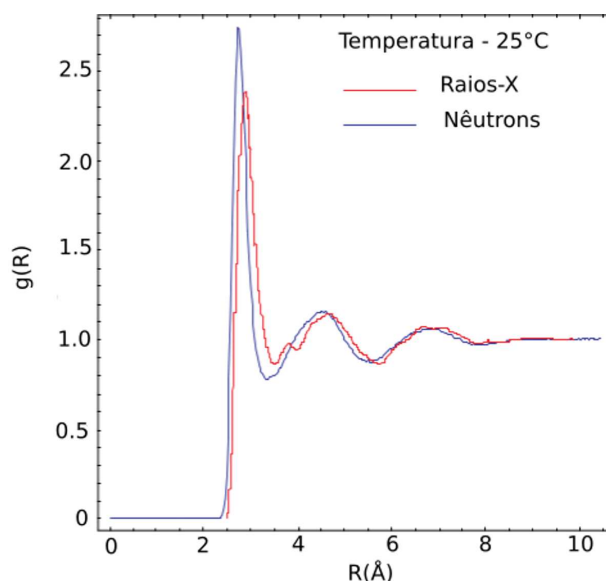


Figura 2.2: Funções de distribuição radiais para a água obtida a partir das técnicas de difração de raios X e espalhamento de nêutrons [36, 37]. A seção de choque do hidrogênio é maior que a do oxigênio, por conta disto o espalhamento de Nêutrons observa o átomo de hidrogênio preferencialmente, e o pico fica mais deslocado à esquerda.

Funções de distribuição podem ser calculadas para sólidos, líquidos e gases. Os sólidos costumam possuir estruturas periódicas e regulares, desta forma, os átomos têm apenas flutuações próximas das posições no retículo cristalino que cada átomo ocupa [34]. Sendo assim, na função de distribuição calculada a partir das posições dos átomos do sólido, devem existir alguns picos em posições bem específicas, estes picos possuirão uma pequena largura em função da vibração dos átomos em torno da posição no retículo. Os gases, por sua vez, possuem mobilidade muito maior em comparação aos sólidos. Desta forma, a função de distribuição para um gás não apresentará picos em distâncias periódicas.

Podemos separar em 3 regimes a função de distribuição de um gás. Primeiramente, teríamos $g(r) = 0$ se a distância r for menor que o raio σ (raio aproximado da molécula de gás). O ponto onde $g(r) > 1$ representa algum tipo de interação e é localizado em $\sigma < r < 2\sigma$. Quando $r > 2\sigma$ teremos regime bulk [34, 35]. Os líquidos terão funções de distribuição com mais informação do que os gases, porém a distância entre os picos não será periódica quanto ao que ocorre nos sólidos. As substâncias no estado líquidos apresentam moléculas que possuem certa mobilidade e alguma organização ao longo da solução (principalmente tratando-se de substâncias muito polares, ou formadoras de ligações de hidrogênio, como a água).

2.1.1 Funções de distribuição de mínima distância

Em razão das RDFs tomarem o centro de massa das espécies como sítios de referência, elas não consideram o formato do soluto e não representam bem interações entre moléculas da solução e a superfície do soluto [38]. Se um soluto possui um formato muito diferente de uma esfera, a função de distribuição radial não representará a camada de solvatação de forma precisa. A Figura 2.3A apresenta como seria a problemática existente no cálculo de RDFs em solutos com forma muito diferente de uma esfera. Podemos observar que, na situação representada por 2.3A, parte do volume que estamos contando as partículas do sistema está sendo ocupado pelo próprio soluto. Esta situação culmina em uma função de distribuição sem muita utilidade física, isto pensando em retirar informação qualitativa sobre as interações entre solvente e soluto a partir da RDF.

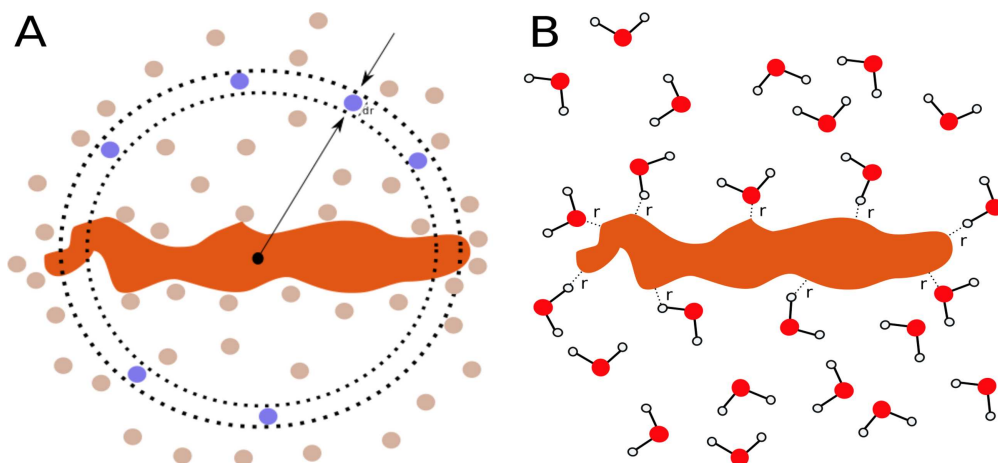


Figura 2.3: A - Representação da particularidade existente no cálculo das RDFs para solutos que possuem forma muito diferente de uma esfera. B - Exemplificação da ideia básica contida nas funções de distribuição de mínima distância. Contando as mínimas-distâncias r podemos obter a função de distribuição do solvente independentemente do formato do soluto.

Logo, uma função de distribuição que leva em conta a forma do soluto deve proporcionar uma caracterização da estrutura de solvatação mais intuitiva. As funções de distribuição de mínima distância (MDDFs) levam em conta a forma do soluto [38]. Essas funções são melhor ajustadas à caracterização da estrutura da solução de acordo com a distância entre a superfície do soluto e do solvente, que está em linha com a visão clássica da superfície de solvatação da proteína. Na função de distribuição de mínima-distância definimos que a "mínima-distância" é a menor distância entre qualquer átomo do solvente e qualquer átomo do soluto. A Figura 2.3B apresenta um conjunto de moléculas de água que estão a uma mínima-distância r da superfície proteica. Neste novo arranjo,

se contarmos as mínimas-distâncias existentes a uma distância r da proteína teremos uma função de distribuição cuja interpretação em termos de interações moleculares faz sentido independentemente das por diferentes formas do soluto.

Podemos definir a função de distribuição de mínima-distância (MDDF) como:

$$g_{us}^{md}(r) \equiv \frac{n_{us}(r)}{n_{us}^*(r)} \quad (2.1.3)$$

onde $n_{us}(r)$ é a densidade numérica média de átomos do solvente que estão a uma distância mínima r do soluto. $n_{us}^*(r)$ é a densidade numérica média que deveria ser observada se não houvesse interação do soluto com o solvente [38]. É importante observar que, se o soluto e o solvente têm somente um átomo, a função de distribuição de mínima distância reduz-se a função de distribuição radial, pois $n_{us}(r)$ passa a ser a densidade de moléculas de solvente a uma distância r , e $n_{us}^*(r)$ passa a ser constante e igual a densidade do solvente no bulk. As densidades $n_{us}(r)$ e $n_{us}^*(r)$ diferem da densidade do solvente, pois elas são associadas com a contagem de qualquer átomo da molécula de solvente a cada distância [38].

A Figura 2.4 apresenta uma representação simplificada de como seria a contagem de mínimas distâncias. Na representação o soluto hipotético está em vermelho. O átomo do solvente que está mais próximo da proteína é contado (isto é equivalente a contar mínimas distâncias). Por fim, é estabelecida uma razão com o número de átomos que seriam contados caso existisse a proteína, mas não houvessem interações entre as moléculas. Desta maneira, distribuições do solvente ao redor de solutos com variadas formas são calculadas com precisão.

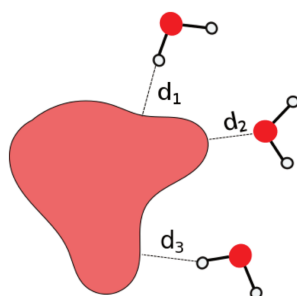


Figura 2.4: Esquema simplificado apresentando como as MMDFs são calculadas. Em vermelho temos o soluto, as demais moléculas são águas. As distâncias assinaladas no desenho representam as menores distância entre a superfície do soluto e um átomo das moléculas do solvente.

A Figura 2.5 contém uma comparação (que serve como exemplo) entre uma $RDF(g(r))$

e uma MDDF ($g_{md}(r)$) para um mesmo sistema. Nela, vemos estamos observando a distribuição do ânion (DCA) do líquido iônico EMIMDCA em um sistema cujo soluto é a ubiquitina em sua conformação nativa. Sem entrar em detalhes sobre o significado de cada pico, podemos notar que a MDDF apresenta uma distribuição com picos com um formato bem definido, além disso pode-se notar a convergência para 1 de maneira clara. A RDF, por sua vez, não apresenta picos com forma definida e a convergência para 1 não acontece até a distância máxima representada. Sendo assim, a MDDF apresenta melhor a distribuição do ânion em torno da proteína.

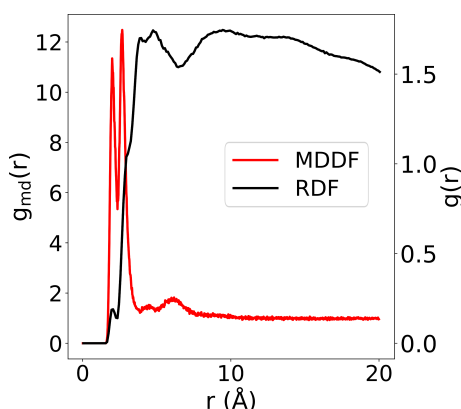


Figura 2.5: Representação das funções de distribuição radial e de mínima distância para o ânion DCA ao redor da ubiquitina na conformação nativa. As funções de distribuição foram calculadas a partir da mesma simulação. Podemos notar que a RDF apresenta um comportamento que é muito difícil de interpretar. Contudo, a MDDF apresenta picos com formatos que possibilitam uma descrição mais intuitiva do sistema.

2.2 Integrais de Kirkwood-Buff e Parâmetro de Solvatação Preferencial

2.2.1 Teoria de soluções de Kirkwood-Buff

O estudo de misturas é fundamentado em quantidades termodinâmicas de excesso (energia livre, entropia, entalpia e volume de excesso). A teoria de Kirkwood-Buff (TKB) conecta propriedades macroscópicas com propriedades microscópicas de soluções [39, 35]. Essa teoria torna possível o cálculo de propriedades termodinâmicas (por exemplo, coeficiente de atividade e volume molar parcial) a partir de integrais de funções de correlação de pares sobre o espaço. As funções de correlação de pares (aqui, funções de distribuição de mínima distância) podem ser calculadas a partir de métodos computacionais, como as simulações de dinâmica molecular (DM). A TKB pode ser aplicada a qualquer solução em equilíbrio envolvendo variado número de componentes [40, 41].

Além disso, a TKB estabelece relação de propriedades como compressibilidade, volume molar parcial, potenciais químicos e do parâmetro preferencial de solvatação, com funções de distribuição, com as integrais de Kirkwood-Buff (IKB):

$$G_{ij} = 4\pi \int_0^{\infty} r^2 [g_{ij}(r) - 1] dr \quad (2.2.1)$$

onde $g_{ij}(r)$ é a função de distribuição radial entre duas espécies i e j . r é a distância da partícula de referência.

Fisicamente, as IKBs refletem a afinidade entre o soluto e as moléculas do solvente [42, 43, 40, 44]. Além disso, podemos obter o número de moléculas que se acumulam ou são excluídas (do soluto) multiplicando a IKB pela concentração da moléculas em questão. Portanto, $\rho_i G_{ji}$ pode ser entendido como o número de partículas i em toda a região integrada (Equação 2.2.1) após a alocação de uma partícula do tipo j (ρ_i é a densidade da partícula i).

Assim, a TKB permite o cálculo de propriedades termodinâmicas a partir das funções de distribuição. Grande parte das funções de distribuição são obtidas por meio de simulações computacionais [39]. O ensemble μ, V, T corresponde ao sistema aberto e as simulações computacionais normalmente utilizam ensembles que correspondem a sistemas fechados (as simulações nos sistemas abertos são mais difíceis de serem realizadas), como N, V, T e N, P, T , onde N é o número de partículas no sistema, V , P e T são, respectivamente, volume, pressão e temperatura do sistema. No entanto, para sistemas suficientemente grandes, não há uma diferença importante em simulações realizadas em diferentes ensembles. Logo, podemos usar funções de distribuição obtidas com outros ensembles na TKB. Para sistemas suficientemente grandes as quantidades termodinâmicas convergem para valores equivalentes, independentemente do ensemble utilizado.

A forma e tamanho das moléculas não têm influência no formalismo da TKB, logo, pode-se aplicá-la para moléculas com formas e tamanhos variados. As IKBs vêm sendo utilizadas para gerar expressões exatas para as interações preferenciais de cossolventes com proteínas. Além disso, quantificam as interações entre soluto e solvente em termos do excesso ou déficit de água e cossolvente nas proximidades do soluto em relação às suas concentrações no seio da solução. As IKBs podem ser escritas em termos da densidade numérica média de átomos do solvente que estão a

uma distância mínima r do soluto ($n_{us}(r)$) e da densidade numérica média que deveria ser observada se não houvesse interação do soluto com o solvente ($n_{us}^*(r)$):

$$G_{US}(R) = (1/\rho_s) \int_0^R S(r)[n_{us}(r) - n_{us}^*(r)] dr \quad (2.2.2)$$

onde $S(r)$ é a superfície definida pela mínima distância r , no caso das MDDFs, em relação a qualquer átomo do soluto, como estamos integrando as funções de distribuição ao longo de todo o volume, o resultado final será a correlação geral entre duas partículas que é dada pela integração de todos os desvios da distribuição bulk. De maneira simples e direta, as IKBs representam o volume que o componente ocupa na solução em relação ao que ocuparia se não houvessem interações. Portanto, são propriedades de excesso.

2.2.2 Parâmetro de Solvatação Preferencial

O parâmetro de solvatação preferencial (Γ) é uma quantidade termodinâmica que ajuda a entender como componentes de uma solução interagem com solutos complexos, como é o caso das proteínas [25, 45]. Γ pode ser experimentalmente mensurado por densimetria envolvendo diálise no equilíbrio, assim como por centrifugação analítica [25]. O experimento de diálise determina Γ por meio da relação:

$$\Gamma_{\mu_1, \mu_3} \equiv \left(\frac{\partial m_3}{\partial m_2} \right)_{T, \mu_1, \mu_3} \quad (2.2.3)$$

onde m_2 e m_3 são, respectivamente, a molalidade da proteína e do cossolvente.

O objetivo da diálise no equilíbrio é medir a variação na concentração de cossolventes devida à adição da proteína na solução. Isto é feito, experimentalmente, através de um processo indireto. Em qualquer mistura de cossolvente e biomolécula é difícil distinguir o quanto do cossolvente interage efetivamente com o soluto. Contudo, caso o cossolvente livre possa ser dialisado até que a concentração através da membrana entre em equilíbrio, será possível mensurar a concentração do cossolvente que está ligado [25]. Portanto, com a concentração do cossolvente conhecida, a concentração do cossolvente ligado à biomolécula pode ser calculada. Sendo assim, Γ_{μ_1, μ_3} é interpretado num contexto de domínio local-bulk como uma medida de acúmulo/exclusão,

que é experimentalmente dado em função das molalidades do soluto pequeno (cossolvente) em duas soluções na diálise no equilíbrio. O experimento é simplificadaamente representado na Figura 2.6 [46].

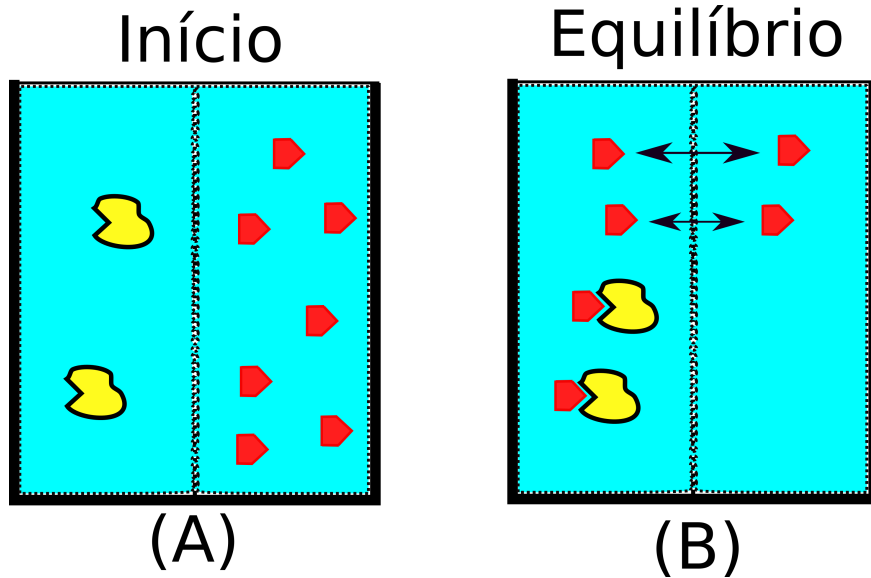


Figura 2.6: Esquema simplificado de um experimento de diálise no equilíbrio. O fundo azul representa a água, os polígonos vermelhos são as moléculas de cossolvente e os desenhos amarelos são as biomoléculas.

Como Γ é uma quantidade física, pode ser obtida a partir de teorias gerais para soluções, como a teoria de KB, pela relação contida na equação 2.2.4 [40, 44].

$$\Gamma_{uc} = \rho_c(G_{uc} - G_{wc}) \quad (2.2.4)$$

onde G_{wc} e G_{uc} são as integrais de Kirkwood-Buff (KBIs), ρ_c é concentração do cossolvente no bulk, u corresponde ao soluto e w ao solvente principal (geralmente água).

Havendo um equilíbrio químico entre conformações desnaturada e nativa de uma proteína, podemos estudar como um cossolvente interage com cada estado. Um cossolvente que apresenta um parâmetro de solvatação preferencial maior no estado desnaturado da proteína (Γ_{uc}^D) irá deslocar o equilíbrio químico em favor do estado desnaturado. Este tipo de cossolvente será tratado como desnaturante. Como regra, se $\Gamma_{uc} < 0$, a proteína é preferencialmente hidratada. Porém, se $\Gamma_{uc} > 0$, a proteína é preferencialmente solvatada pelo cossolvente. Esta situação leva ao desenovelamento da estrutura [31, 33]. O parâmetro de hidratação preferencial (Γ_{uw}) indica se

a água é excluída ou acumulada em torno da proteína.

Este fato pode ser observado de maneira mais clara relacionando Γ com o potencial químico. Quando um cossolvente é adicionado a uma solução aquosa que contém uma proteína, há uma alteração no potencial químico da proteína devido à presença do cossolvente [47, 48]. Resumidamente, este processo é a mudança nas interações da proteína com o solvente quando é transferida da água pura para um sistema com cossolvente:

$$\Delta\mu_u^{tr} = \mu_u(\text{cossolvente}) - \mu_u(\text{água}) \quad (2.2.5)$$

onde μ_i é o potencial químico do componente i . É definido como $\mu_i = \mu_i^\circ(P, T) + RT \ln a_i$, aqui a_i é a atividade do coeficiente i , i.e. $a_i = m_i \gamma_i$, onde m_i é a molalidade do componente i e γ_i é o coeficiente de atividade de i . Quando um cossolvente é adicionado em um sistema com água, há uma perturbação mútua no potencial químico da proteína e do cossolvente [25, 49]. Esta "perturbação" é dada pela seguinte igualdade de derivadas:

$$\left(\frac{\partial \mu_u}{\partial m_c} \right)_{T, P, m_u} = \left(\frac{\partial \mu_c}{\partial m_u} \right)_{T, P, m_c} \quad (2.2.6)$$

A relação 2.2.6 é o gradiente de transferência de energia livre em uma concentração de cossolvente (medida termodinâmica de interação em uma dada concentração). Podemos relacionar 2.2.6 com 2.2.5 usando a seguinte relação:

$$\Delta\mu_u^{tr} = \int_0^{m_c^*} \left(\frac{\partial \mu_u}{\partial m_c} \right)_{m_u} dm_c \quad (2.2.7)$$

A Equação 2.2.7 relaciona a variação do potencial químico com o efeito decorrente da adição do cossolvente (que é a derivada). O parâmetro de solvatação preferencial é uma forma de definir termodinamicamente a ligação entre uma componente da solução e o soluto. Se a concentração de um cossolvente é maior nas vizinhanças da proteína do que no bulk, Γ_{uc} é maior do que zero, e μ_u é menor na presença no cossolvente do que na sua ausência [47]. Podemos relacionar a variação do potencial químico da proteína (na adição de um cossolvente à uma solução de proteína

e água) com o parâmetro de solvatação preferencial por meio da seguinte relação [47, 48]:

$$\Delta\mu_u^{tr} = -RT\Gamma_{uc} \quad (2.2.8)$$

A equação 2.2.8 mostra uma relação direta entre Γ_{uc} e o potencial químico. A solvatação preferencial, e a perturbação no potencial químico, referem-se à uma dada composição do solvente. O efeito do cossolvente, relativo à água pura como estado de referência, é expresso por $\Delta\mu_{tr}$, a energia livre de transferência da proteína da água para o solvente de composição. A Figura 2.7 mostra uma situação onde o Γ_{uc} do seria positivo. O resultado seria a diminuição do potencial químico da proteína na presença de um LI, levando em conta um sistema hipotético anterior que teria LI confinado em uma caixa dentro de um sistema com a proteína e a água.

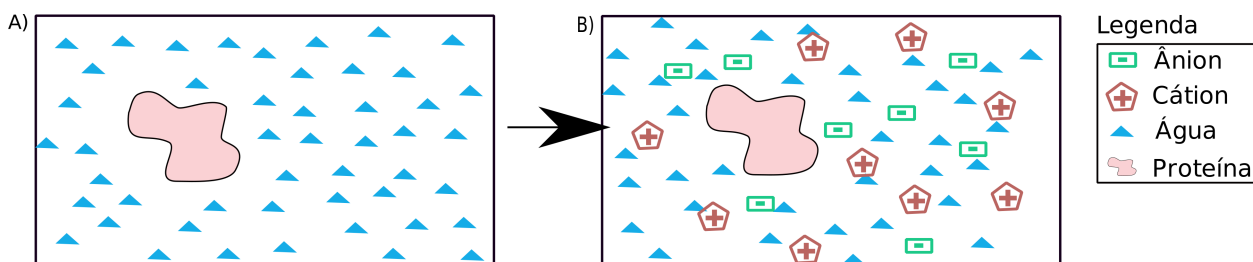


Figura 2.7: Representação do sistema que descreve a adição de LI em uma solução aquosa com proteína. Em (A) temos um sistema hipotético que contém a água e a proteína em solução, nesta configuração a proteína tem um determinado potencial químico (μ_p^0). Em (B) o LI é colocado no sistema e se espalha pela solução, passando a interagir com a proteína. Aqui o potencial químico da proteína foi perturbado, de maneira que acarretou em acúmulo de LI na superfície proteica. Nesta situação, Γ_{uc} seria maior que zero e o novo potencial químico da proteína seria menor do que no sistema (A), levando em conta a equação 2.2.8.

Capítulo 3

Dinâmica molecular

Diversos sistemas físicos e químicos possuem enorme complexidade, portanto frequentemente não é possível compreender todos os mecanismos por trás de um fenômeno apenas com técnicas experimentais. Porém, se houvesse algum microscópio que possibilitasse a observação da movimentação das partículas (em tempo real), inúmeras propriedades e fenômenos poderiam ser entendidos mais profundamente. Apesar deste aparato experimental ser, atualmente, utópico, diversas técnicas computacionais aproximam-se de um "microscópio em tempo real". Uma das técnicas computacionais mais amplamente utilizadas é a dinâmica molecular. A dinâmica molecular é uma das principais técnicas computacionais para a simulação de sistemas químicos, em especial sistemas contendo biomoléculas. Este capítulo será dedicado à apresentação da dinâmica molecular. Inicialmente apresentaremos uma visão geral do método, focando nos princípios mecânico-estatísticos que fazem com que a dinâmica molecular seja uma técnica confiável. Além disso, apresentaremos as etapas básicas de uma simulação de dinâmica molecular e uma breve introdução a respeito do campo de força utilizado nesta dissertação

3.1 Visão geral do método

A dinâmica molecular (MD) fornece informações a respeito do comportamento dinâmico microscópico, dependente do tempo, das partículas que compõem um sistema molecular. Os métodos de simulação por dinâmica molecular (MD) geram uma trajetória, i.e., uma série de configurações espaciais do sistema geradas pela propagação de um conjunto inicial de coordenadas e velocidades [50]. Essa propagação dá-se, no caso de simulações clássicas, pela aplicação da segunda lei de Newton por uma série finita de passos de tempo. Sendo assim, sucessivas

configurações dos sistemas são geradas por meio da integração das equações do movimento de Newton. O resultado final é uma trajetória que compreende posições e velocidades das partículas pelo tempo [51].

A trajetória é obtida pela resolução da equação diferencial existente na segunda lei de Newton, ou seja, resolvemos as equações de movimento para cada partícula que compõe o sistema. O número de partículas do sistema simulado depende do tipo de propriedade que se deseja calcular e da capacidade computacional disponível [52]. Além disso, é necessário que a razão do número de partículas (N) com o volume da caixa de simulação (V) represente a concentração de um sistema real que se deseja simular.

São necessários conceitos da mecânica estatística para, a partir de propriedades microscópicas obtidas da trajetória, extrair grandezas macroscópicas observáveis (pressão, energia interna, volume, entropia, energia livre, etc.) [51]. Um conceito fundamental na mecânica estatística é que, se esperarmos tempo suficiente, o sistema irá, eventualmente, fluir através de todos os microestados consistentes com as restrições que definem seu estado termodinâmico. Realizando N medidas independentes no sistema simulado, o valor observável dessas medidas para a propriedade P é:

$$P_{obs} = \frac{1}{N} \sum_{a=1}^N P_a \quad (3.1.1)$$

onde P_a é o valor numérico da a -ésima medida, e N é o número de medidas.

Podemos tratar a soma em 3.1.1 como:

$$P_{obs} = \sum_v \left[\frac{1}{N} N_v \right] P_v \quad (3.1.2)$$

onde N_v é o número de vezes que o estado v é observado, P_v é o valor esperado para P quando o sistema está no estado v . Se assumirmos que, após um longo tempo, todos os possíveis estados do sistema foram visitados, podemos atribuir à fração de tempo (probabilidade) no estado v o símbolo p_v e, conseqüentemente, teremos:

$$P_{obs} = \sum_v p_v P_v \equiv \langle P \rangle \quad (3.1.3)$$

A equação 3.1.3 é a média no ensemble da propriedade P . Considera-se que a média

temporal é a mesma da média do ensemble (hipótese conhecida como hipótese Ergódica na mecânica estatística) [34, 35].

3.2 Etapas básicas de uma simulação de dinâmica molecular

3.2.1 Montagem da caixa de simulação

A etapa inicial de uma simulação de DM consiste na definição da configuração inicial, ou seja, precisamos montar a caixa de simulação. No caso dos sistemas que serão abordados nesta dissertação, a caixa inicial contém a proteína alocada no centro da caixa, e as moléculas de água e de LI aleatoriamente distribuídas pelo espaço restante de uma caixa, que frequentemente é cúbica. As dimensões da caixa são escolhidas de forma que a densidade numérica (razão entre o número de moléculas dentro da caixa pelo seu volume) seja correspondente àquela desejada para o sistema real. As posições iniciais que cada partícula irá ocupar na caixa de simulação normalmente é organizada levando em conta a disposição que as partículas ocupam em redes cristalinas [52]. Esta forma de dispor as partículas evita sobreposições e forças de interações intermoleculares muito fortes na montagem da caixa. A estrutura da proteína é obtida a partir do banco de dados de proteína (PDB - *Protein Data Bank*).

3.2.2 Cálculo das forças

Na DM as moléculas são tratadas como uma coleção de átomos que pode ser descrita por forças clássicas, ou seja, consideramos que em um sistema tratado na DM as partículas são mantidas unidas por forças elásticas e harmônicas. Um conjunto completo de potências de interação entre as partículas (átomos, moléculas...) é denominado "campo de força" [53, 54]. A energia de uma molécula é expressa como sendo uma função da resistência ao estiramento da ligação, torção da ligação, ângulos diedros e aglomeração dos átomos. A equação para a energia é usada para encontrar o comprimento das ligações, ângulos e diedros [55]. Na DM as partículas são tratadas de maneira clássica. Portanto, fenômenos quânticos como quebra de ligações químicas não são observados [51]. Bons campos de força são imprescindíveis para que a simulação de DM possa gerar resultados reprodutíveis e concordantes com os experimentos.

Podemos apresentar uma forma geral para as equações que regem os campos de forças por meio das equações abaixo. Desta forma, um campo de força tem a seguinte equação como geral :

$$E = \sum_{\text{ligações}} E_{\text{estiramento}} + \sum_{\text{ângulos}} E_{\text{dobramento}} + \sum_{\text{dihedros}} E_{\text{torção}} + \sum_{\text{pares}} E_{\text{não-ligação}} \quad (3.2.1)$$

onde as energias ($E_{\text{estiramento}}$, $E_{\text{dobramento}}$, $E_{\text{torção}}$ e $E_{\text{não-ligação}}$) são contribuições do estiramento de ligações, dobramento de ângulos, movimento torcional ao redor de ligações simples, e interações entre átomos ou grupos que são de não-ligação. De maneira geral, os campos de força tratam basicamente de interações de ligação e não-ligação. Normalmente, os campos de força são parametrizados, i.e. definição dos parâmetros que serão utilizados para descrever as interações, estes parâmetros podem ser obtidos por meio de dados empíricos e/ou cálculos quânticos, para uma determinada aplicação, por exemplo o campo de força OPLS (*Optimized Potentials for Liquid Simulations*) é especialmente parametrizado para reproduzir comportamentos experimentais em simulações de líquidos de uma maneira Geral [56, 57]. Existem diversos campos de forças diferentes como o CHARMM, GROMOS, GAFF, entre outros [58, 59, 60].

O campo de força utilizado neste trabalho consiste em uma alteração dos potenciais existentes no OPLS, especialmente parametrizado para líquidos iônicos imidazólicos. Este campo de força consiste de parâmetros que utilizam um sítio virtual no centro do anel imidazólico. A função do sítio virtual é de representar a carga positiva deslocalizada que os anéis imidazólicos possuem. A parametrização das cargas foi realizada por meio de cálculos de energia livre de hidratação do cátion imidazólico em solução aquosa. Todos os termos dos potenciais deste campo de força foram refinados de forma a reproduzir dados experimentais de densidades, entalpia de vaporização, viscosidade, coeficiente de difusão, tensão superficial e capacidade calorífica. Devido ao uso de um sítio virtual, o campo de força recebe o nome de OPLS-VSIL (*Optimized Potentials for Liquid Simulations - Virtual Site Ionic Liquid*) [61].

3.2.3 Movimentação das partículas e ensembles estatísticos

Como já mencionado, a trajetória resultante de uma simulação de DM advém da integração das equações de movimento para cada partícula do sistema. O método das diferenças finitas é aplicado em diversos algoritmos que são utilizado na integração das equações de

movimento. Neste método a integração é dividida em intervalos temporais pequenos, Δt , permitindo a simulação das movimentações de maior frequência que acontecem no sistema (normalmente são as vibrações das ligações). O algoritmo de Verlet é frequentemente usado no cálculo da trajetórias de partículas dinâmicas em simulações computacionais [62]. Resumidamente, o algoritmo utiliza posições e acelerações dos átomos no tempo t e as posições do passo anterior, $r(t + \Delta t)$, para determinar as novas posições no tempo $t + \Delta t$, como apresentado da equação 3.2.2:

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + a(t)\Delta t^2 \quad (3.2.2)$$

O conjunto de configurações é o ensemble. As propriedades que são mantidas constantes definem o tipo de ensemble [34]. Um possível ensemble que pode ser utilizado pode ser o NVE , onde são as variáveis que são mantidas fixas são o número de partículas, volume e energia. Outras opções são os ensembles μVT , NVT e NPT . As variáveis que dão nome ao ensemble são àquelas mantidas fixas. A escolha do ensemble que será escolhido para simulação depende das condições que queremos para o sistema.

Para realizar uma simulação de DM é necessário que sejam definidas algumas variáveis de controle. O controle da temperatura é alcançado por meio de termostatos. O termostato de Berendsen, por exemplo, funciona da seguinte forma: durante as simulações, as velocidades são reescaladas a cada passo de integração, de forma que a energia cinética do sistema é ajustada até chegar na temperatura requerida [51, 52]. O controle da pressão dá-se pela utilização dos barostatos. Alguns barostatos que são muito utilizados são os de Berendsen e o de Parrinello-Rahman [63, 64, 65].

3.2.4 Condições iniciais e etapas da simulação

Em uma simulação de DM as posições iniciais das partículas que constituem o sistema devem ser especificadas [51, 52, 55]. É importante que as posições iniciais não tenham sobreposição das partículas. Softwares como o *Packmol* desempenham o papel de construção de configurações iniciais com eficiência [66, 67]. Em sistemas com biomoléculas, por exemplo, estruturas obtidas de técnicas experimentais como difração de raios X (DRX) e ressonância magnética nuclear (RMN) são utilizadas como configurações iniciais. Apesar das posições iniciais

não contemplarem situações onde há sobreposições (quando feitas corretamente), elas, muito provavelmente, possuem configurações que resultam em energia potenciais muito altas. Assim, a primeira etapa da DM deve ser uma minimização da energia potencial do sistema [51, 55, 52]. As configurações iniciais de uma DM que obtemos após a minimização não correspondem às estruturas adequadas nas condições termodinâmicas desejadas, sendo necessárias etapas adicionais, chamadas de equilibração. A equilibração é uma simulação que ocorre até que propriedades como pressão, temperatura ou energia fiquem constantes ao longo do tempo (as propriedades que devem atingir valores constantes durante a equilibração dependem do ensemble utilizado). Na equilibração as propriedades termodinâmicas não se mantêm constantes, o tempo de equilibração é dependente do tipo e do tamanho do sistema.

Outro aspecto importante nas simulações de DM são as condições periódicas de contorno (PBC) [68]. As PBC viabilizam a realização das simulações com um número relativamente pequeno de partículas, de forma que as partículas experimentem forças semelhantes ao que ocorreriam se elas estivessem no regime bulk da solução. As PBC possuem a função de minimizar os efeitos de superfície em uma simulação de DM. Isto é feito fazendo com que os átomos existentes dentro da caixa que será simulada sejam replicados em todas as direções no espaço. Durante as simulações, quando um átomo se movimenta na caixa inicial, sua imagem periódica (em uma das caixas que foram replicadas) move-se da mesma forma. Se um átomo sair da caixa inicial, sua imagem entrará pela face oposta (contendo a mesma velocidade).

Com as PBC aplicadas ao sistema, temos uma situação em que as N partículas iniciais (por exemplo) são replicadas em todas as direções. Desta forma, se o sistema contém N partículas e o potencial se expressa como a soma das interações entre pares de partículas, teríamos que fazer essa soma para as outras caixas replicadas [68]. A soma, portanto, contaria com infinitos termos, sendo necessária a realização de aproximações. Para as interações de curto alcance (van der Waals) utilizamos a convenção da imagem mínima [69]. Basicamente, é realizado um truncamento do potencial em um raio de corte esférico, costumeiramente da ordem de metade do comprimento da caixa de simulação. Isto evita a interação da partícula com sua própria imagem ou com a mesma partícula duas vezes.

O tratamento das interações de longo alcance (potencial de Coulomb), em que o potencial decai com a distância que é menor do que r^d (onde d é a dimensionalidade do sistema) é

diferente do que vimos para as interações de curto alcance [51]. O truncamento abrupto do potencial pode comprometer fortemente os resultados da simulações. Este problema vem sendo tratado usando método derivados das somas de Ewald, como P^3M (Particle-Particle-Particle Mesh Ewald) e PME (Particle Mesh Ewald) que produzem trajetórias estáveis [70, 51, 71, 72, 73].

Diversos textos tratam de maneira aprofundada e completa os mais diversos aspectos das simulações por dinâmica molecular [51, 52, 55, 74, 75]. Desde a parametrização de campos de força até todos os algoritmos usados para atribuição de velocidades iniciais, controle das variáveis do sistema (ensembles), cálculo das forças sobre as partículas, propagação das partículas ao longo da simulação, manutenção da geometria molecular, cálculo de forças de longo alcance e determinação de propriedades físico químicas a partir das trajetórias.

Capítulo 4

Objetivos e Metodologia

Este capítulo apresentará os principais objetivos do trabalho. Também será descrito o protocolo utilizado para realizar as simulações de DM e todas as análises.

4.1 Abordagem e objetivos

4.1.1 Visão geral da abordagem do sistema

Os sistemas que serão estudados nesta dissertação são formados por uma proteína, água e LIs. Nesta configuração, vamos tratar o LI como um cossolvente. Por meio das MMDFs será almejado o entendimento da distribuição das moléculas na solução (sempre tomando como ponto de referência a superfície da proteína). Isto possibilitará a determinação da forma com que cada composto se distribui e, conseqüentemente, a natureza da interação entre as moléculas da solução com a proteína. Portanto, a ideia geral do trabalho é estudar a estrutura de solvatação de proteínas por líquidos iônicos. A informação inicial reportada pelas funções de distribuição de mínima-distância diz respeito às interações que ocorrem entre os componentes da solução e a superfície da proteína, ou seja, é uma informação microscópica.

Além da distribuição dos componentes ao longo da solução, precisamos obter quantidades macroscópicas. Isto pode ser alcançado pela teoria de soluções de Kirkwood-Buff (TKB). Com ela podemos calcular volumes molares de excesso tendo uma noção de acúmulo ou exclusão efetivos de uma componente da solução dos arredores da proteína. Com isto podemos calcular o parâmetro preferencial de solvatação. A ideia principal consiste em calcular o parâmetro de

solvatação preferencial (Γ) para sistemas com diferentes tipos de líquidos iônicos e em diversas concentrações. Além disso, buscamos entender o que acontece no sistema quando a solução de LI e água é colocada em contato com a proteína em conformações que possuem uma maior área superficial d . Se Γ apresenta um valor maior para um determinado componente numa conformação mais aberta da proteína deve ser um indicativo que, ao longo do tempo, a proteína tenderá a aumentar sua área superficial para maximizar a interação com este componente.

De maneira geral, a abordagem segue as seguintes etapas:

- Simulação dos sistemas
- Cálculo das funções de distribuição
- Cálculo das integrais de Kirkwood-Buff
- Cálculo do Γ

Portanto, o trabalho possui dois grandes objetivos, o primeiro deles é entender como os LIs distribuem-se em uma solução que contém uma proteína. O segundo deles, que é consequência do primeiro, consiste em entender o efeito do LI sobre a estabilidade da proteína. Além disso, por meio das funções de distribuição e teorias de soluções pertinentes, determinar propriedades que permitam o entendimento físico-químico das interações estabelecidas entre água-proteína e LI-proteína.

4.2 Metodologia Desenvolvida

4.2.1 Montagem das caixas de simulação

Os sistemas simulados consistiram de 4 componentes: água, proteína, ânion e cátion (provenientes do LI). A ubiquitina é uma proteína que possui 76 resíduos e é uma das proteínas mais amplamente estudadas [76, 77], também desempenha uma importante função reguladora nas células eucariotas, onde é encontrada. Ademais, detém a função de marcar proteínas que não estão na conformação tridimensional adequada para que sejam degradadas [78]. Sendo assim, a ubiquitina por si só é uma proteína que desempenha papel muito importante. Também é conveniente porque possui uma carga total nula. Utilizou-se a estrutura da ubiquitina proveniente de difração de raios X. A estrutura foi retirada do "Protein Data Bank", com código PDB *1UBQ* [79, 80].

Foram utilizadas diversas conformações da ubiquitina, divididas em 3 tipos. O primeiro tipo de ubiquitina usada corresponde à estrutura nativa da proteína, que é a conformação da ubiquitina retirada diretamente do experimento de difração de raios X. O segundo tipo compreende um grupo de 30 diferentes conformações parcialmente desenoveladas, obtidas simulando a ubiquitina no vácuo em 600 K (as condições são as mesmas da etapa de termalização que será descrita na seção "Dinâmica molecular", exceto pela proteína estar fixa). As conformações foram obtidas tomando 30 frames, igualmente intervalados, da trajetória final. O último tipo é uma representação de uma proteína totalmente desenovelada. Utilizando a sequência de aminoácidos que compõem a ubiquitina criou-se uma estrutura estendida de forma linear (foram substituídas 3 prolina por 3 alaninas, fazendo com que a estrutura tridimensional resultando da ligação entre cada aminoácido assumisse uma conformação reta) .

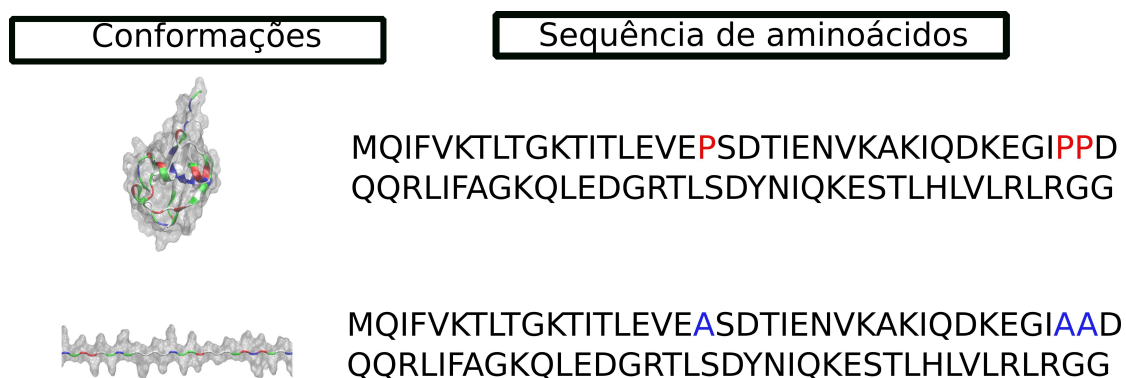


Figura 4.1: A imagem apresenta, na cor vermelha, os três aminoácidos prolina que existem na ubiquitina nativa. Na cor azul, estão representadas as três alaninas que foram colocadas no lugar das prolina.

Quatro LIs foram utilizados como cossolventes no sistema, compostos por dois cátions diferentes: 1-alquil-3-metilimidazol, onde o alquil é um etil (EMIM) ou butil (BMIM), e dois ânions: dicianamida (DCA) e tetrafluoreto de boro (BF₄). Os líquidos iônicos, portanto, são resultado da combinação dos dois cátions e dois ânions citados. A partir de agora, usaremos os seguintes nomes para nos referir aos líquidos iônicos: EMIMDCA, BMIMDCA, EMIMBF₄ e BMIMBF₄. A fórmula estrutural de cada componente está representada na figura 4.2.

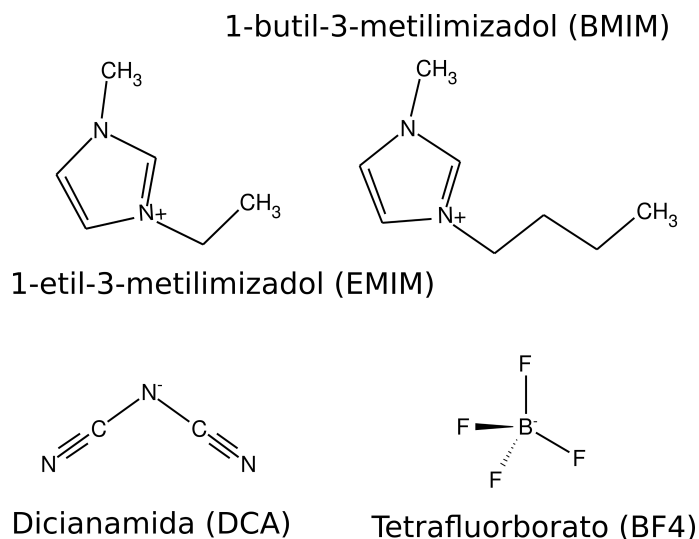


Figura 4.2: Fórmulas estruturais dos cátions e ânions que compõem os líquidos iônicos utilizados.

Para as caixas de simulação definiu-se um padrão de forma que houvesse distâncias mínimas iguais entre as bordas das proteínas e as faces da caixa. Sendo assim, independentemente da conformação que estava sendo utilizada, a mínima distância entre a borda da proteína e a face da caixa foi a mesma. Cada caixa teria um tamanho de 30 Å a mais em cada direção, além das dimensões da proteína.

Para calcular a concentração, o volume destinado para a solução era a diferença entre o volume total da caixa e o volume ocupado pela proteína. O número de moléculas de água e de LI necessários para chegar em uma concentração específica foi definido utilizando a aproximação em que cada molécula teria um volume proporcional à sua massa. Assim, foi possível calcular o número de moléculas necessário para solvatar a proteína e ainda respeitar a concentração inicial. Os sistemas foram montados com as seguintes concentrações de LI: 0,5; 1,0; 1,5; 2,0; 2,5 e 3,0 mol L⁻¹. É importante ressaltar que, no momento que fizemos a simulação usando o ensemble *NPT* o volume das caixas mudou, se adequando às condições termodinâmicas impostas. Dessa forma, as concentrações aqui informadas não passam de estimativas. Elas serão recalculadas e as concentrações que serão reportadas nos resultados corresponderão àquelas correspondentes ao sistema simulado.

As configurações iniciais foram criadas contendo a proteína, água e LI. Todas as caixas contendo as configurações de partida foram criadas utilizando o software Packmol [66]. A disposição das moléculas tinha a proteína ocupando o centro da caixa enquanto os componentes da solução

eram dispostos de maneira aleatória. O cálculo das funções de distribuição de mínima-distância e das integrais de Kirkwood-Buff foi feito utilizando o software MDAnalysis com o módulo gmd [38].

4.2.2 Dinâmica molecular

Os potenciais (ligação e não-ligação) utilizados para a ubiquitina são provenientes do campo de força OPLS, enquanto que para os líquidos iônicos são provenientes do OPLS-VSIL [61]. Para a água utilizou-se o campo de força TIP3P [81]. O campo de força OPLS-VSIL foi parametrizado utilizando um sítio virtual. Este artifício foi utilizado para fixar uma carga no centro do anel imidazólico e, assim, simular a carga deslocalizada que existe nos imidazóis presentes nos cátions dos LIs.

Para a proteína no estado nativo a simulação ocorreu como descrito adiante. As simulações foram realizadas no software GROMACS.2018.3.CUDA [50]. A etapa inicial foi uma minimização utilizando o método *Steepest Descent* com no máximo 50000 passos. Nesta etapa foram aplicadas restrições de movimento à proteína, deixando-a fixa [82].

A segunda etapa consistiu numa equilibração usando o ensemble NVT por 1 ns usando o termostato modificado de Berendsen, ainda com a proteína fixa [65]. Esta etapa tem como objetivo a termalização do solvente, ou seja, estamos equilibrando a estrutura do solvente em torno da estrutura proteica. A temperatura foi mantida fixa em 300 K, o raio de corte utilizado para as interações (eletrostáticas e van der Waals) de curta distância foi 10 Å. O passo de tempo utilizado foi de 2 fs. Particle Mesh Ewald foi utilizado para o cálculo das interações eletrostáticas de longa distância [83].

Após a equilibração da temperatura, equilibrou-se a pressão com uma simulação usando o ensemble NPT (300 K e 1.0 bar), continuamos equilibrando a temperatura aqui utilizando o termostato modificado de Berendsen [65] e o barostato de Parinello-Rhman [64, 63] por 5 ns (ainda com a proteína fixa). Foi feita uma dinâmica molecular de 1 ns com o sistema livre e posteriormente uma simulação de produção por 10 ns. Para cada líquido iônico foram simuladas 6 concentrações diferentes. Para cada concentração de LI, 20 simulações foram realizadas. A difícil convergência das integrais de Kirkwood-Buff requer que muitas simulações sejam feitas para cada concentração.

As simulações da proteína na forma estendida e das conformações provenientes da simulação no vácuo foram realizadas da mesma forma. Os sistemas foram minimizados e equilibrados no ensemble NVT da mesma forma do que foi feito para a proteína no estado nativo, contudo, a proteína sempre foi mantida fixa no centro da caixa. Somente era possível a movimentação das cadeias laterais. Simulações de 5 e 10 ns foram feitas usando o ensemble NPT. A simulação utilizada para a produção foi de 10 ns.

A razão de terem sido feitas diversas simulações com tempo de 10 ns ao invés de uma simulação é que estamos interessados em estudar a estrutura de solvatação da conformação do início da simulação. Portanto, com pouco tempo de simulação poderemos entender a estrutura de solvatação ao redor das conformações que propusemos inicialmente e não de possíveis estruturas que possam vir a aparecer ao longo de simulações mais longas.

4.2.3 Cálculo das funções de distribuição de mínima-distância

As funções de distribuição de mínima distância (MDDFs) foram calculadas até 20Å distante da superfície da proteína, consideramos que a estruturação da solução sofre influência da proteína. Esta condição surge do fato de que, em 20Å as KBIs convergem para um valor aproximadamente constante. Sendo assim, tudo que está neste raio de distância da superfície da proteína será considerado como sofrendo influência da presença da mesma. As demais moléculas que compõem a solução são constituintes do bulk. Portanto, as concentrações foram re-calculadas levando em conta a quantidade de moléculas presentes no bulk de cada sistema simulado. As funções de distribuição de mínima distância foram calculadas até uma distância de 20 Å.

A função de distribuição total pode ser decomposta em suas contribuições parciais para cada átomo do solvente. Além disso, também podemos decompor a função de distribuição total das contribuições parciais levando em conta o soluto. Com a decomposição é possível saber com quais partes da proteína o solvente interage mais, ou qual átomo no solvente é responsável por algum comportamento observado na curva.

O cálculo do parâmetro preferencial de solvatação (Equação 2.2.4) utiliza as integrais de Kirkwood-Buff (Equação 2.2.2). Estas são calculadas integrando todos os desvios da distribuição das componentes do solvente, em relação ao bulk, sob todo o volume. A Figura 4.3 é um exemplo

de integral de Kirkwood-Buff. O parâmetro preferencial de solvatação é obtido pela multiplicação da concentração bulk da espécie em questão pela diferença entre a KBI da substância e da água.

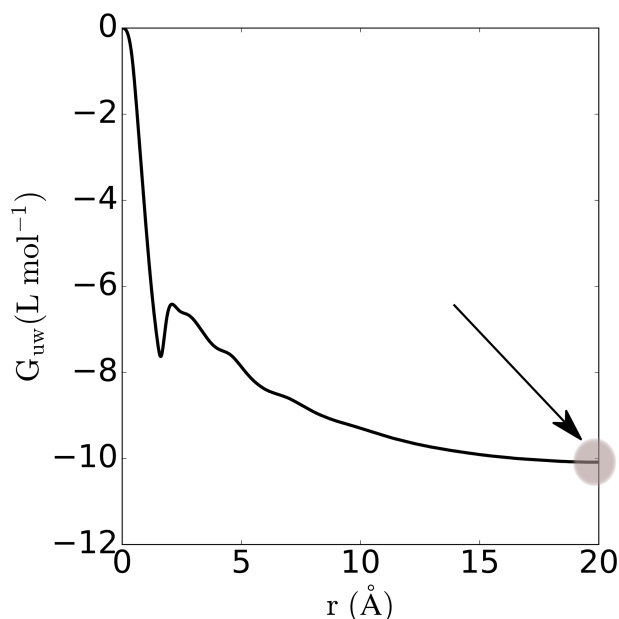


Figura 4.3: Integral de Kirkwood-Buff para um sistema contendo o líquido iônico EMIMDCA. A curva foi calculada integrando a função de distribuição da água ao redor da proteína ao longo de toda a solução. O valor utilizado para os cálculos dos parâmetros de solvatação preferencial é aquele onde há uma constância na curva da integral de Kirkwood-Buff, representado por G_{uw} , sendo que u corresponde ao soluto e w à água. Para este caso, a integral de KB começa a ser constante a partir de 20 Å.

Na introdução foi apresentado que Γ_{uc} é calculado por meio da equação $\Gamma_{uc} = \rho_c(G_{uc} - G_{wc})$. O cálculo realizado para a água é simetricamente oposto, corrigido pela concentração de água no bulk, que é muito maior do que qualquer outra espécie em solução ($\Gamma_{uw} = \rho_w(G_{uw} - G_{uc})$). A magnitude de Γ_{uw} será maior do que para Γ_{uc} , visto que a água possui uma concentração maior do que as outras componentes do solvente.

Vale mencionar que, o presente trabalho utiliza as simulações de dinâmica molecular como técnica computacional para obter as trajetórias do sistema. As funções de distribuição de mínima-distâncias poderiam ter sido calculadas por meio de simulações de Monte-Carlo. Contudo, na prática, acaba não sendo vantajoso utilizar simulações de Monte-Carlo pois as configurações que são geradas na simulação acabam por possuir muitas sobreposições de átomos. De maneira geral, a dinâmica molecular é mais eficiente do que os métodos por simulações de Monte-Carlo em sistemas com moléculas de grande complexidade estrutural [84].

Capítulo 5

Distribuição da água

Neste capítulo vamos estudar um sistema que consiste de água e da ubiquitina em sua conformação nativa. Observaremos como é a distribuição da água sem a presença dos LIs.

5.1 Solvatação da Ubiquitina pela água

A Figura 5.1 contém a MDDF para a água ao redor da ubiquitina no estado nativo. Há um pico fino e pronunciado em 1,8 Å, correspondente à distância típica na qual as ligações de hidrogênio (interações específicas) são estabelecidas [85]. No sistema em estudo, as moléculas de água estão interagindo por ligações de hidrogênio com a superfície da proteína.

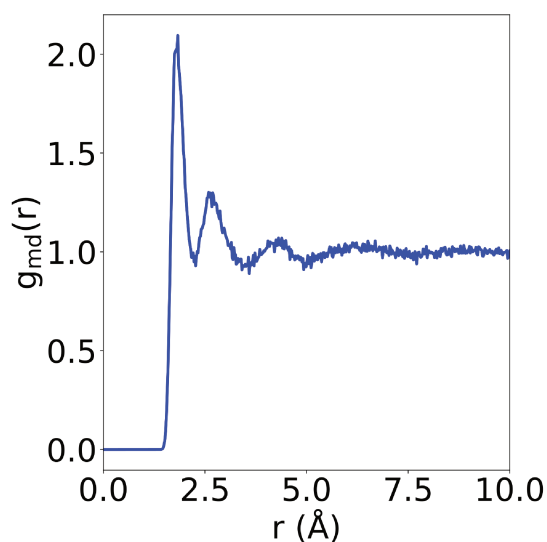


Figura 5.1: Função de distribuição de mínima distância para a água em relação à ubiquitina. O maior pico em 1,8 Å indica que a densidade da água nessa distância é aproximadamente duas vezes maior que a de bulk

As moléculas da água podem estabelecer as ligações de hidrogênio através dos dois átomos de hidrogênio ou pelo oxigênio (neste caso, realizando a ligação com os hidrogênios, ligados em átomos muito eletronegativos, presentes na superfície da proteína). A Figura 5.2 mostra a contribuição dos átomos de oxigênio e hidrogênio para a função de distribuição total. É evidente que a contribuição dos hidrogênios é maior.

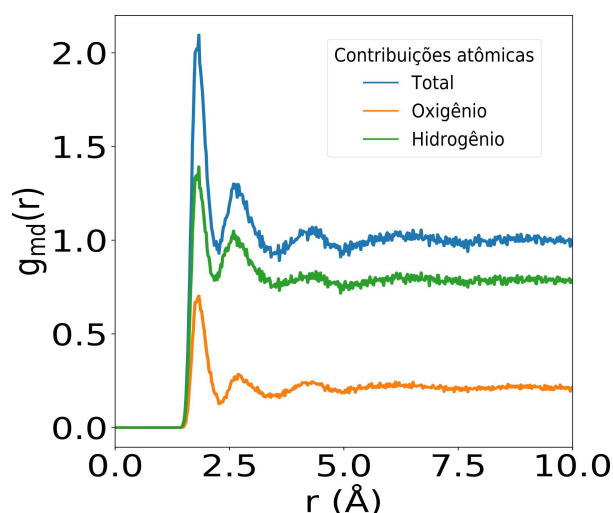


Figura 5.2: Decomposição das funções de distribuição de mínima distância para a água em suas contribuições atômicas. A decomposição consiste em separar a MDDF total em contribuições atômicas, ou seja, retiramos a informação do quanto determinado átomo contribui para a MDDF.

É interessante observar que as MDDFs provenientes das decomposições não convergem para 1. Na verdade, a decomposição representa a probabilidade de que o átomo em questão seja aquele que está mais próximo da superfície da proteína. Desta forma, as MDDFs dos átomos não convergem para 1, e sim para um valor associado com a probabilidade deste átomo estar mais perto da proteína quando a molécula está distante da superfície desta [martinez2017].

A segunda banda (na figura 5.1), em 2,7 Å, corresponde à segunda camada de solvatação, e a terceira banda, 4,3 Å, corresponde à terceira camada de solvatação. O primeiro pico é maior e mais fino que outros. Portanto, ele é devido às ligações de hidrogênio entre as moléculas de água e a superfície da proteína (como já mencionado).

A Figura 5.3 contém a MDDF total da água ao redor da proteína, assim como a contribuição da distribuição ao redor de cada resíduo para a função de distribuição total. Podemos observar que, para o pico referente às ligações de hidrogênio, as maiores contribuições são de resíduos polares e carregados, a contribuição existente do backbone deve-se que, neste caso, os

átomo de hidrogênio que estão ligados ao backbone também foram considerados. A maior contribuição dos resíduos polares é totalmente esperada, visto que a polaridade da água faz com que ocorram interações entre a água e os sítios polares da proteína. Quando observamos os demais picos (2,7 e 4,5 Å) podemos notar um crescimento da contribuição proveniente dos resíduos alifáticos. As interações entre a água e os resíduos alifáticos ocorreram com menor frequência do que as outras.

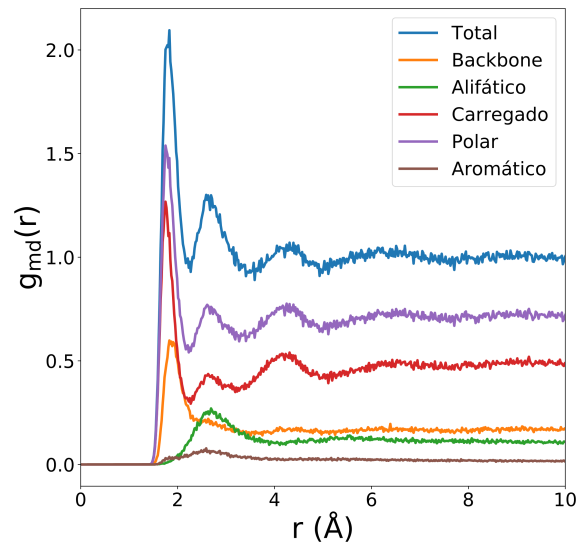


Figura 5.3: Decomposição da MDDF que descreve a distribuição água-proteína em contribuições do backbone e de cadeias laterais específicas.

Em longas distâncias a distribuição é aleatória, portanto a ordem observada em 10Å na Figura 5.3 reflete a proporção de cada tipo de resíduo que forma a superfície da proteína. Ou seja, não são apenas as interações que definem as proporções iniciais (picos iniciais na Figura 5.3), mas também a composição da superfície.

A Figura 5.4 contém a IKB para a água. A curva converge para um valor negativo. Isto significa que a água é efetivamente excluída da região contendo a proteína. A equação 5.1.1 apresenta de maneira direta como a IKB é calculada.

$$G_{uw}(R) = \frac{1}{\rho_s} [N_{uw}(R) - N_{uw}^*(R)] \quad (5.1.1)$$

onde $N_{uw}(R)$ é o número de moléculas de água que possuem pelo menos um átomo a uma distância R da proteína, $N_{uw}^*(R)$ é o número de moléculas de água que possui pelo menos um átomo a uma distância R da proteína em uma distribuição aleatória com a mesma densidade de bulk. w

corresponde à água e u é o soluto (proteína). De agora em diante, o termo domínio da proteína irá se referir às distâncias onde $r < R$. Nestas regiões, consideramos que o soluto têm influência sobre a distribuição dos outros compostos ao longo da solução. Ou seja, as interações entre soluto e solução não podem ser negligenciadas. A partir de R a distribuição da solução é descorrelacionada com o soluto, portanto temos o regime bulk.

A IKB ser negativa ($N_{uw}(R) < N_{uw}^*(R)$), indica que na simulação existe um número menor de moléculas dentro da distância R do que existiria numa distribuição aleatória. O comportamento geral é a exclusão das moléculas de água da região que compreende o domínio de proteína.

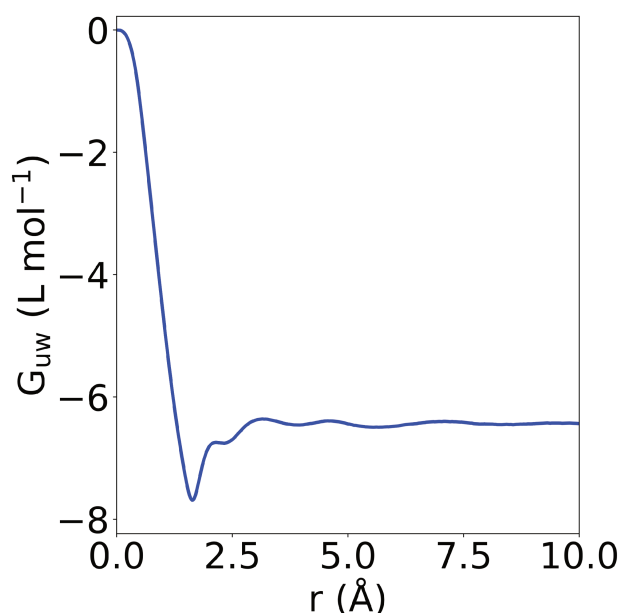


Figura 5.4: Integral de Kirkwood-Buff para a água no sistema apenas com a ubiquitina no estado nativo.

A proteína possui um volume de exclusão, isto faz com que não existam moléculas do solvente muito próximas aos átomos da proteína. Porém, conforme podemos observar na Figura 5.4, entre 1,5 e 3,0 Å há um aumento na curva. Como o solvente é excluído em distâncias próxima da proteína, posteriormente há uma acumulação (que, como vimos na curva, ocorre ao longo da solução). O resultado final da IKB é proveniente do balanço entre a exclusão inicial e o acúmulo posterior. As análises apresentadas aqui servem como exemplo (e base) para a discussão que será desenvolvida para sistemas mais complexos durante os próximos capítulos.

Capítulo 6

Variação da concentração de líquido iônico

Este capítulo apresentará os resultados dos sistemas simulados com os LIs EMIMDCA, BMIMDCA, EMIMBF4 e BMIMBF4 em sistemas cujo soluto é a Ubiquitina em sua conformação nativa. Discutiremos os resultados focando no efeito que a variação da concentração do LI tem sobre a acumulação (ou exclusão) do solvente sobre a proteína. O parâmetro de solvatação preferencial para os íons do LI (Γ_{uc}) apresentou crescimento de 0,5 até 1,5 mol L⁻¹ (para a maioria dos líquidos iônicos). Após esta concentração, Γ_{uc} passou a diminuir. Estas observações indicam que, em baixas concentrações, os LIs solvatam preferencialmente a proteína. Em maiores concentrações o LI passa a ser excluído. Vamos apresentar que os íons distribuem-se de maneira distinta ao longo da solução. Os ânions ocupam regiões mais próximas da superfície proteica, estabelecendo ligações de hidrogênio, enquanto os cátions apresentam maior densidade relativa em distâncias em torno de 2,5 Å, estabelecendo interações não-específicas.

6.1 Efeito relativo à variação da concentração

O parâmetro de solvatação preferencial (Γ_{uc} para os íons e Γ_{uw} para a água) é uma medida relativa entre duas componentes da solução. Valores positivos de Γ indicam que há acúmulo relativo de uma componente do solvente sobre a superfície do soluto [13]. A Figura 6.1 contém Γ_{uc} para os íons dos LIs EMIMDCA e EMIMBF4. Para o EMIMDCA, Γ_{uc} é positivo em todas as concentrações simuladas, enquanto Γ_{uc} do sistema com EMIMBF4 só é positivo no intervalo de 0,5 até 2,0 mol L⁻¹. Para o LI EMIMDCA, existe um aumento do Γ_{uc} que ocorre de 0,5 até 1,0 mol L⁻¹ e de 1,5 até 2,0 mol L⁻¹. Para o EMIMBF4 há um leve aumento de 0,5 até 1,5 mol L⁻¹. Este comportamento pode estar relacionado com a partição do LI entre a solução

(majoritariamente composta por água em baixas concentrações) e a superfície proteica. A existência de uma concentração em que Γ_{uc} foi maior em todos os íons remete à uma situação em que naquela quantidade disponível de moléculas de LI no sistema, a acumulação sobre a proteína é máxima. O LI, principalmente o cátion, possui pouca afinidade com a água, o que favorece sua interação com as regiões hidrofóbicas da proteína.

O parâmetro de solvatação preferencial da água (Γ_{uw}) é menor quando a concentração de LI é $0,5 \text{ mol L}^{-1}$. Para a água, conforme a concentração de LI cresce, Γ_{uw} passa a ficar menos negativo. A água tem sua exclusão atenuada com o aumento da concentração de LI.

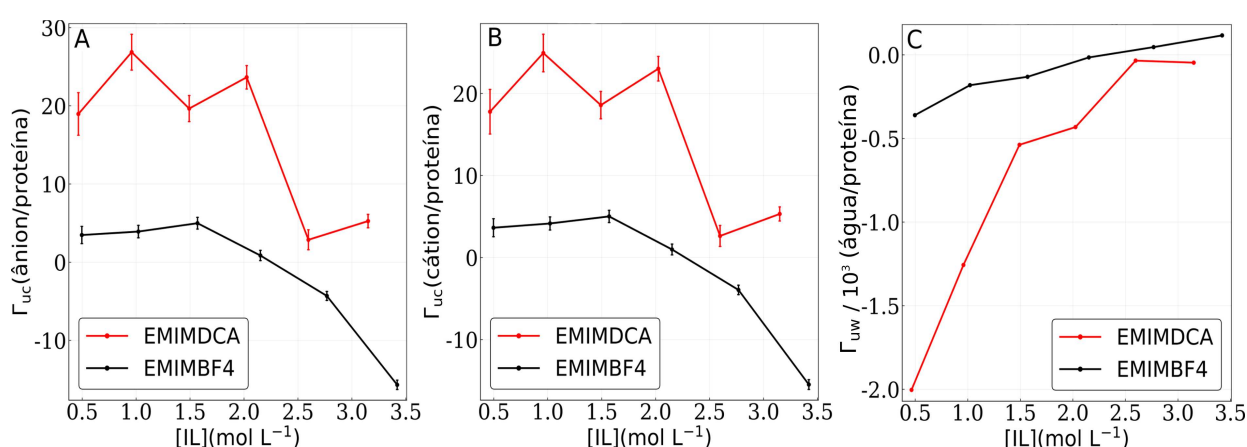


Figura 6.1: Parâmetros de solvatação preferencial para o ânion (A), o cátion (B) e a água (C) em sistemas com os LIs EMIMDCA e EMIMBF4. Podemos observar um crescimento inicial de Γ para os íons (até cerca de $1,0 \text{ mol L}^{-1}$), para o DCA temos uma tendência que não segue exatamente um aumento contínuo, isto é, muito provavelmente, devido à variabilidade da medida. A água diminui sua exclusão com o aumento da concentração de LI. Para o LI EMIMBF4, Γ_{uc} apresenta um leve aumento entre $0,5$ e $1,5 \text{ mol L}^{-1}$.

As MDDFs dão uma visão mais intuitiva da distribuição das componentes do solvente na solução. As MDDFs para as moléculas de água, apresentadas na Figura 6.2, aumentam com o aumento da concentração dos LIs. Existem picos bem definidos nos sistemas contendo os dois LIs (EMIMDCA e EMIMBF4).

As ligações de hidrogênio são interações que ocorrem à uma distância ótima de $1,8 \text{ Å}$. Desta forma, o pico em $1,8 \text{ Å}$ existente para a MDDF da água representa uma densidade relativa de moléculas de água que estão estabelecendo ligações de hidrogênio com átomos da superfície da proteína. Além disso, este pico apresenta o maior tamanho, indicando que nesta região a água possui densidade maior que a densidade de água no bulk. Os demais picos ($2,8$, $4,5$, $6,4 \text{ Å}$) são

devidos às demais camadas de solvatação.

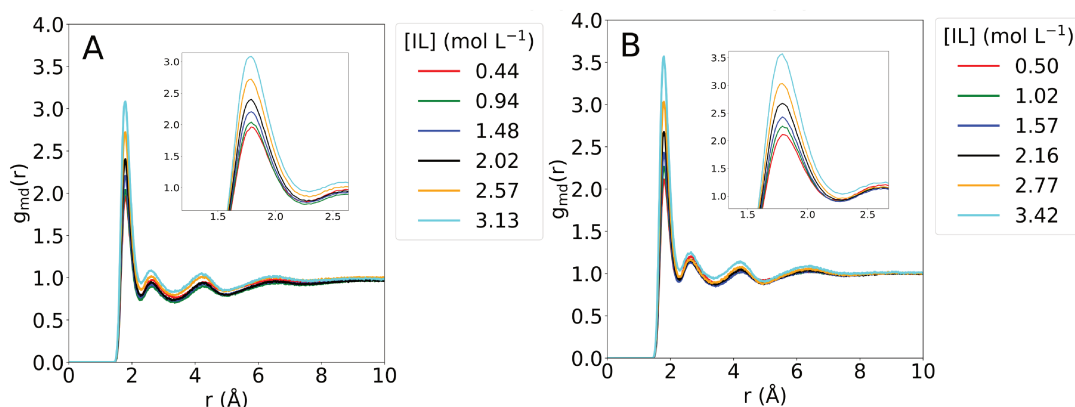


Figura 6.2: MDDFs da água para os sistemas com a proteína no estado nativo e os cossolventes EMIMDCA (A) e EMIMBF4 (B).

A Figura 6.3 contém os parâmetros de solvatação para as componentes da solução em sistemas com BMIMDCA e BMIMBF4. O comportamento geral aqui é parecido ao que temos no cálculo dos mesmos parâmetros para EMIMDCA e EMIMBF4. Contudo, nota-se que Γ_{uc} para os íons é maior. O LI com cátion mais volumoso tende a ficar mais acumulado sobre a superfície da proteína, do que o LI com cátion menos volumoso (como mostrado na Figura 6.1). Além disso, Γ_{uw} é mais negativo, demonstrando que a água é ainda mais excluída quando está em sistemas com LIs mais volumosos.

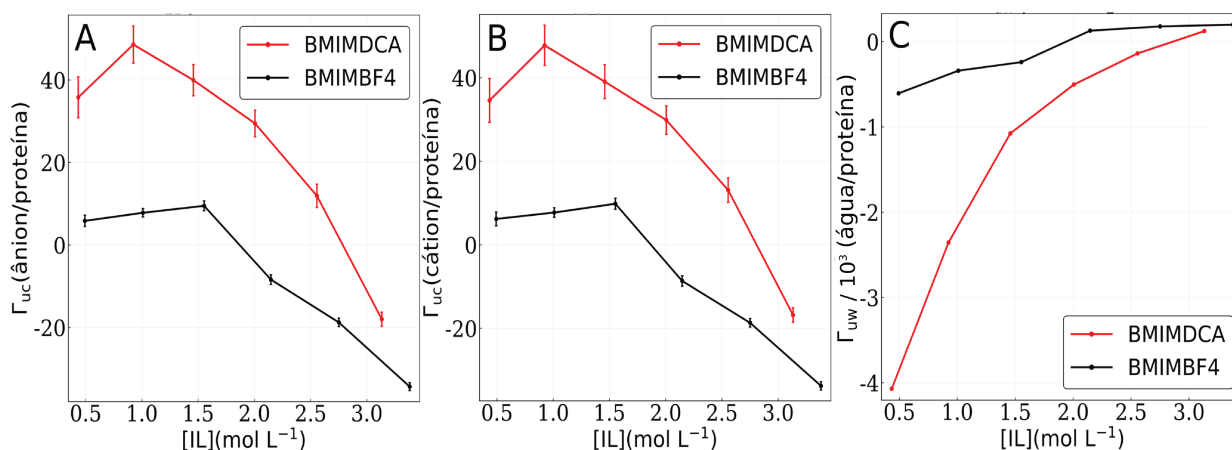


Figura 6.3: Parâmetro de solvatação preferencial para o ânion (A), cátion (B) e água (C) em sistemas contendo BMIMDCA e BMIMBF4.

As MDDFs para os cátions (Figura 6.4) apresentam menor número de bandas em comparação com a água. Há um pico principal em 2,5 Å e uma pequena banda em torno de 5,0

Å. Duas características são interessantes nas MDDFs dos cátions: A primeira delas é a altura dos picos, que segue tendência oposta ao observado para a água. A segunda é que o pico em 2,5 Å apresenta uma largura maior se comparado ao pico principal observado nas MDDFs da água.

O fato de que a MDDF para o cátion apresenta um pico muito pronunciado em 2,5 Å indica que, nesta região da solução, a densidade relativa do EMIM é maior que a densidade relativa do EMIM na distribuição bulk. A distância em que o pico aparece é maior do que a distância média para as ligações de hidrogênio. É importante notar que a altura dos picos também está relacionada com a quantidade de moléculas no sistema. Por exemplo, quando voltamos à Equação 2.1.3 notamos que a MDDF é resultado da razão do número de moléculas em um determinado volume no sistema pelo número de moléculas no mesmo volume na distribuição bulk. Se existem poucas moléculas no bulk, o pico será grande, pois o número de moléculas no bulk é o denominador.

Quanto à largura do pico, podemos pensar da seguinte forma: uma interação específica, altamente energética, ocorreria com pouca variabilidade de possíveis distâncias entre os átomos. Vamos exemplificar usando uma mola. Se temos uma mola que possui uma alta rigidez e aplicamos à ela uma força F , haverá uma deformação determinada pela relação F/k . Contudo, ao aplicarmos a mesma força em uma mola com $k' < k$, teremos uma deformação maior, dada por F/k' . Analogamente, se o pico é mais largo, ele representa que os átomos das moléculas que estão distribuídas na região onde o pico aparece podem interagir com maior variabilidade de distâncias (uma interação mais fraca permitiria maiores "estiramentos"). Portanto, o pico largo contido nas MDDFs para os cátions corresponde a interações não-específicas. As interações não-específicas (no sistema discutido) podem ser de natureza dispersiva (forças de London) ou até mesmo íon-dipolo (por conta da carga positiva nos cátions).

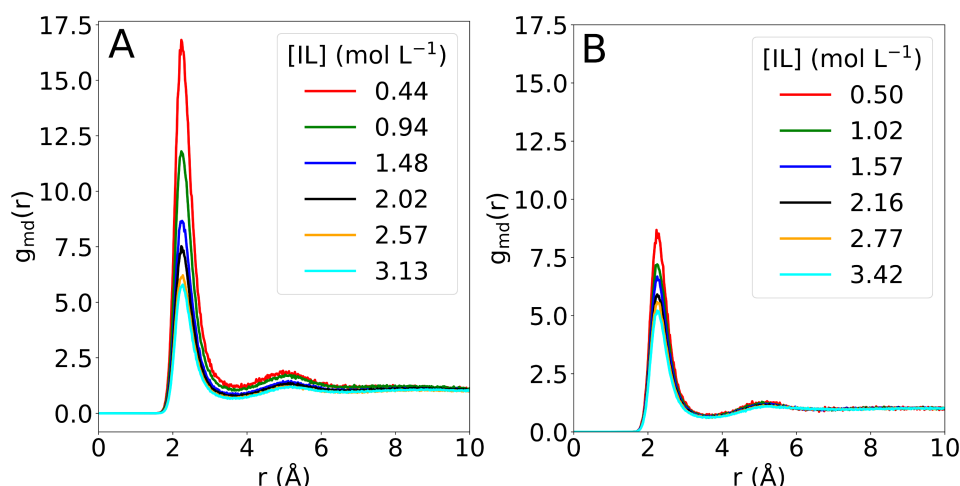


Figura 6.4: Funções de distribuição para o cátion nos sistemas contendo o ânion DCA (A) e BF₄ (B). O pico em 2,5 Å é maior conforme a concentração do líquido iônico é menor, indicando que a densidade relativa ao bulk do cátion nesta distância é maior nas menores concentrações.

A Figura 6.5 contém as MDDFs para o cátion BMIM em sistemas com diferentes ânions. É perceptível o aumento na altura dos picos e, além disso, uma maior distância necessária para convergência da MDDF para os cátions mais volumosos. Logo, maior é a distância requerida para que a correlação com a proteína seja perdida.

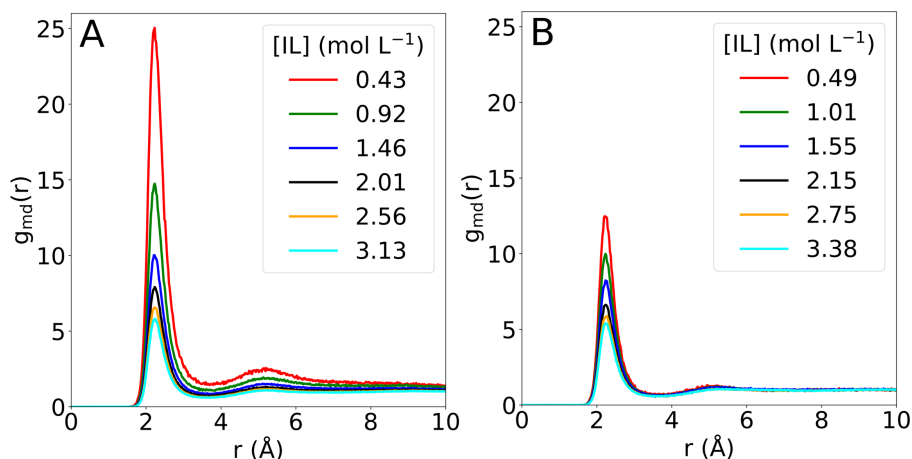


Figura 6.5: MDDFs para o cátion BMIM nos sistemas contendo os ânions DCA (A) e BF₄ (B). O pico em 2,5 Å aumenta com a diminuição da concentração do LI, indicando que há um acúmulo do cátion na superfície da proteína.

As MDDFs para o ânion DCA em diferentes concentrações estão apresentadas na Figura 6.6. Há um pico em torno de 1,8 Å, proveniente de interações específicas que são de natureza semelhante ao que temos para água (DCA e BF₄ (Figura 6.7) podem atuar como

aceptores de ligações de hidrogênio). Em aproximadamente 2,6 Å há outro pico pronunciado, como esta distância é um pouco maior daquela existente para os cátions. Claramente, as interações existentes em 2,6 Å não são ligações de hidrogênio, desta forma podemos pensar em ligações não-específicas como Van der Waals e dipolo-dipolo.

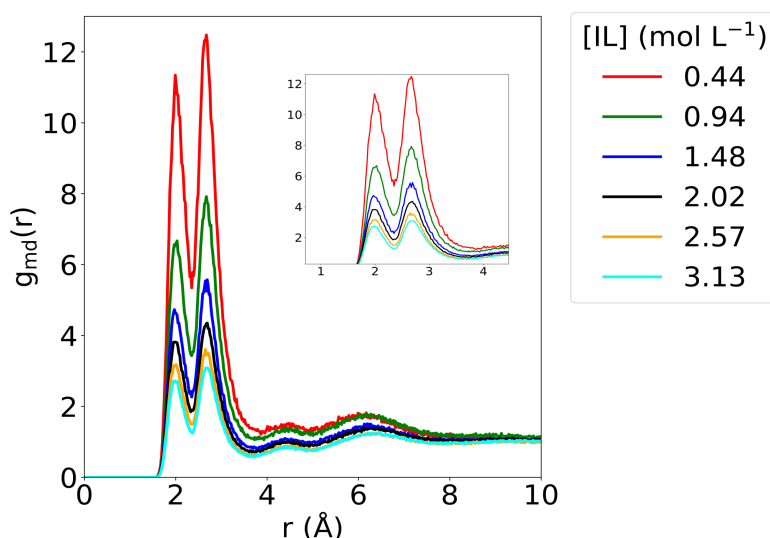


Figura 6.6: MDDFs para o DCA em sistemas contendo diferentes concentrações do LI EMIMDCA. A ordem de altura dos picos segue a mesma tendência daquela observada para as MDDFs dos cátions. A altura do pico aumenta com a diminuição da concentração de LI no sistema. Aqui, temos a presença de dois picos pronunciados, ambos tem formatos semelhantes, sendo o segundo (2,6 Å) ligeiramente maior do que o primeiro (1,9 Å).

A Figura 6.7 contém as MDDFs para o ânion BF₄ em diferentes concentrações. Notam-se picos pronunciados nas distâncias de 1,9 e 2,6 Å (estas distâncias são semelhantes às existentes para as MDDFs do DCA). Contudo, a altura dos picos não apresenta comportamento semelhante ao observado para o DCA. O pico referente às ligações de hidrogênio possui altura bem menor em comparação ao pico em 2,6 Å, mostrando que o BF₄ possui maior densidade na distância de 2,6 Å (interações não-específicas) do que em 1,9 Å (ligações de hidrogênio).

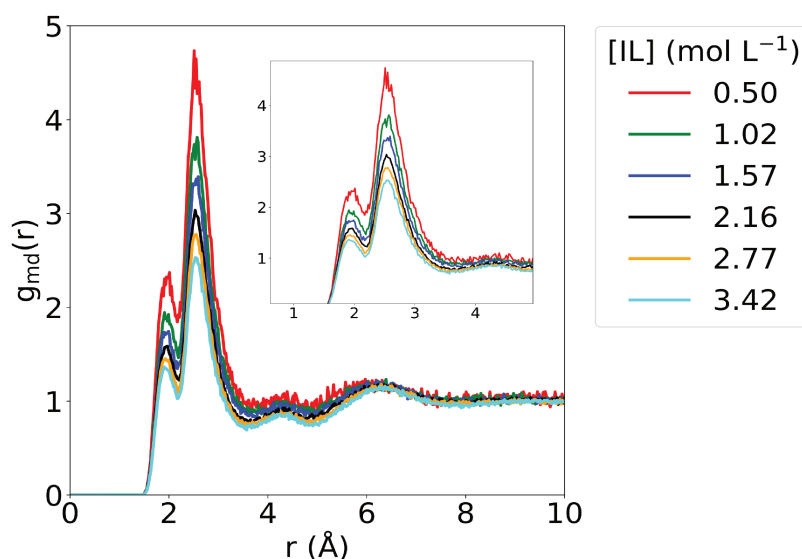


Figura 6.7: MDDFs para o BF_4 em sistemas contendo diferentes concentrações do LI EMIMBF4. Similarmente ao que temos na Figura 6.6, os picos são mais altos nas menores concentrações de LI. Além disso, o pico referente às ligações de hidrogênio é bem menor do que àquele devido às interações não-específicas.

As Figuras 6.8 e 6.9 mostram a decomposição da função de distribuição total em suas contribuições atômicas. Para o DCA (Figura 6.8) a maior parte da função de distribuição é proveniente dos dois nitrogênios neutros que DCA possui. Estes são especialmente importantes para o primeiro pico que aparece em uma distância próxima de $1,8 \text{ \AA}$.

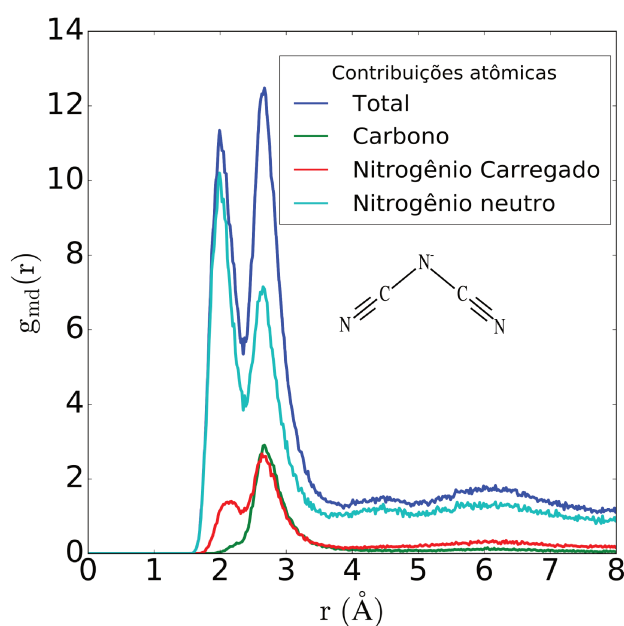


Figura 6.8: Função de distribuição decomposta nas contribuições parciais para os átomos. A figura apresenta a MDDF para o DCA em um sistema com $0,5 \text{ mol L}^{-1}$ de EMIMDCA.

Na Figura 6.8 o carbono e o nitrogênio carregado contribuem de maneira mais efetiva para o segundo pico apresentado na MDDF (Figura 6.8). Os carbonos não poderiam estabelecer interações específicas com os outros componentes do sistema. Desta forma, o fato do segundo pico ser formado com grande participação de átomo de carbono indica que parte das moléculas de DCA apresentam interações que são necessariamente não-específicas.

As contribuições para o BF_4 são todas devidas aos átomos de flúor. Isto é claramente observado na Figura 6.9. A contribuição proveniente do átomo de Boro é nula (representada por uma reta em $y = 0$). Como o Boro está num centro de um arranjo tetraédrico, ele nunca será o átomo que estará mais próximo da proteína, dessa maneira sua contribuição é zerada.

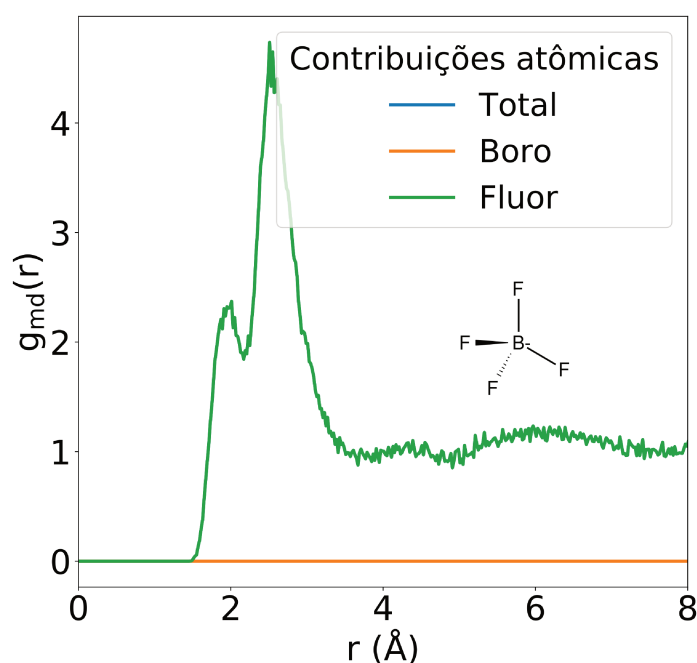


Figura 6.9: Função de distribuição decomposta nas contribuições parciais dos átomos do cossolvente. A figura apresenta a MDDF para o BF_4 num sistema com $0,5 \text{ mol L}^{-1}$ de EMIMBF₄. Como a única contribuição para a MDDF total advém dos átomos de Flúor, a MDDf do flúor sobrepõe à total.

Integrais de Kirkwood-Buff

As integrais de Kirkwood-Buff (IKBs) para os componentes das soluções contendo EMIMDCA e EMMBF₄ são, respectivamente, apresentadas pelas Figuras 6.10 e 6.11. Primeiramente, a curva que representa a IKB começa com um vale, que tem profundidade

diferente para cada componente da solução. A integral de Kirkwood-Buff representa o volume de excesso de cada substância no sistema, os volumes são diferentes para cada componente do sistema em função de um efeito relacionado ao tamanho da molécula. Por exemplo, nas conformações aleatórias, quando parte de uma molécula do solvente está dentro da proteína, e parte está fora, ainda conta-se a menor distância e isto conta para o volume aparente da proteína.

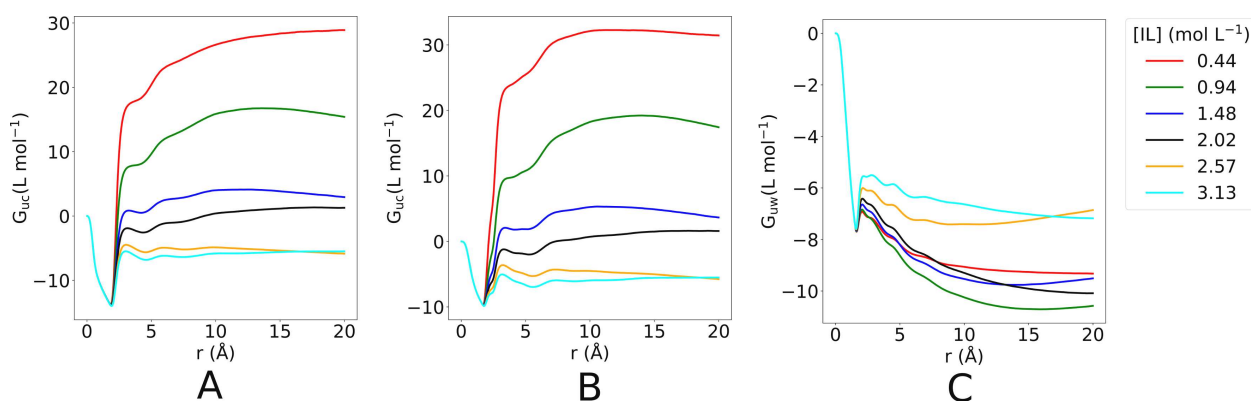


Figura 6.10: Integrales de Kirkwood-Buff para o cátion (A), o ânion (B) e a água (C) em sistemas contendo variadas concentrações de EMIMDCA.

As IKBs para os íons têm maiores valores nas concentrações mais baixas. A IKB mede diretamente o acúmulo/exclusão de uma componente do solvente sobre o soluto, observamos diretamente que nas menores concentrações de LI, há maior acúmulo dos íons do LI. Podemos observar que as IKBs para os íons do EMIMDCA são maiores em comparação aos íons que estão no sistema com EMIMBF4. Este fato, juntamente com o que foi observado nos parâmetros preferenciais de solvatação e MDDFs, demonstram que o EMIMDCA acumula mais sobre a superfície da proteína do que o EMIMBF4.

A menor concentração resulta na maior IKB para ambos os LIs. De fato, os resultados apresentados no cálculo de Γ_{uc} para os íons é reflexo da IKB. A IKB para a água, por sua vez têm valores mais negativos em concentrações abaixo de 2,0 mol L⁻¹. As IKBs para a água, no sistema com EMIMDCA, não apresentam boa correlação com aumento/diminuição da concentração. Como observamos na Figura 6.1, o parâmetro de solvatação preferencial dos íons não é máximo na mesma concentração em que a IKB é máxima. Isto decorre do fato do parâmetro de solvatação preferencial ser uma medida de acúmulo/exclusão relativo, portanto, nesta concentração o acúmulo do íon em relação à água é máximo.

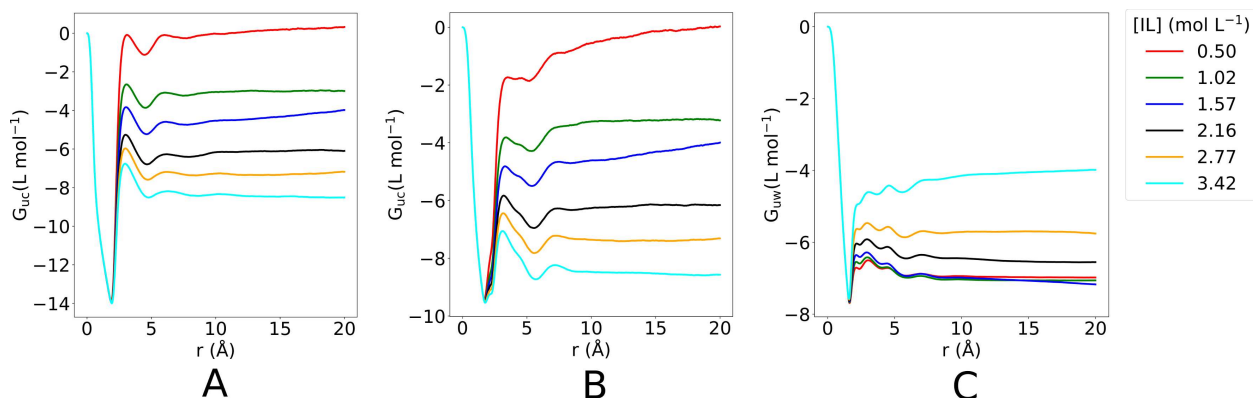


Figura 6.11: Integrais de Kirkwood-Buff para o cátion (A), o ânion (B) e a água (C) em sistemas contendo variadas concentrações de EMIMBF4

Os sistemas estudados nesta seção possuem a ubiquitina em seu estado nativo como soluto, a água como solvente e o LI como cossolvente. Γ_{uc} positivo para os íons em concentrações menores indica acúmulo do LI ao redor da proteína. Consequentemente, a água é inicialmente excluída ($\Gamma_{uw} < 0$). Todavia, nas maiores concentrações do LI na solução, os íons passam a ser globalmente excluídos, enquanto a água solvata preferencialmente a proteína. Podemos observar que o LI apresenta uma afinidade relativa maior com a superfície da proteína. As densidades relativas, assim como acúmulo dos íons, são maiores nas menores concentrações do LI nos sistemas. O parâmetro de solvatação preferencial, que é uma medida relativa entre duas substâncias, aumenta até uma determinada concentração (que fica em torno de $2,0 \text{ mol L}^{-1}$ para a maioria dos LIs). O comportamento não linear mostra que há uma concentração na qual o LI solvata a proteína com máxima preferência. Sendo assim, o LI apresenta uma interação relativa maior com a superfície da proteína.

Capítulo 7

Alteração do estado de enovelamento da ubiquitina

A modulação da estabilidade de proteínas por cossolventes pode ser quantificada pela solvatação preferencial, ou seja, por uma competição existente entre interações proteína-água e proteína-cossolvente para estados nativos e desnaturados. Neste capítulo, apresentaremos resultados referentes às simulações de sistema contendo conformações nativa, desnaturada e estendida da Ubiquitina. Veremos que a MDFF para os LIs apresenta picos muito maiores nas conformações mais abertas da ubiquitina. O parâmetro de solvatação preferencial dos LIs aumenta conjuntamente com o grau de desenovelamento das conformações da proteína. Isto implica a uma indução de desenovelamento da proteína por parte do LI.

7.1 Efeito na acumulação de líquido iônico sobre a proteína relativo ao estado de enovelamento

A Figura 7.1 apresenta um exemplo ilustrativo dos diferentes tipos de conformações simuladas da ubiquitina. Recapitulando o que foi apresentado no capítulo 4, quando nos referimos ao tipo "nativa" falamos apenas da conformação cristalográfica da Ubiquitina. O tipo "desnaturada" corresponde a um conjunto de 30 conformações parcialmente desenoveladas. O último tipo da Ubiquitina, "estendida", corresponde a somente uma conformação. Nesta situação, a proteína assume uma configuração onde todos os resíduos possuem os mesmos ângulos de diedro na cadeia principal.

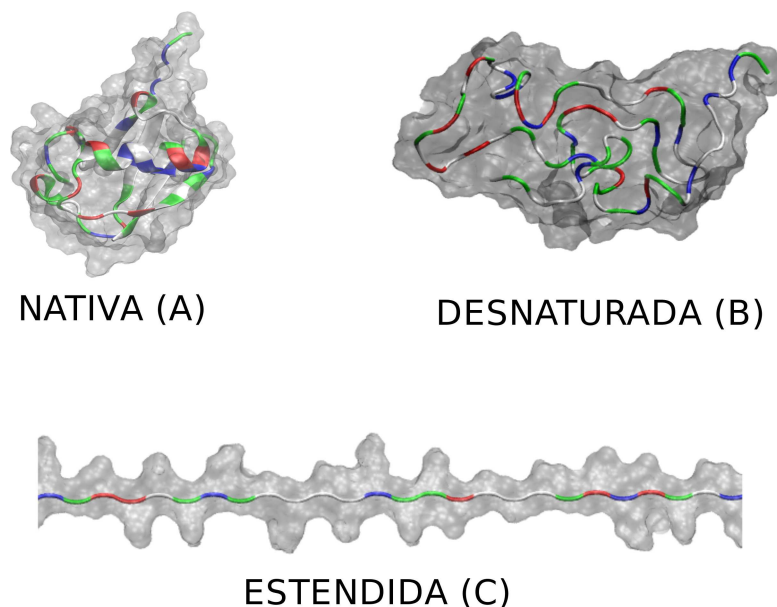


Figura 7.1: Diferentes conformações simuladas para a ubiquitina. (A) corresponde à estrutura nativa da ubiquitina, (B) é uma das 30 conformações parcialmente desenoveladas utilizadas no trabalho e (C) é um corte da conformação estendida.

A Tabela 7.1 contém a área superficial acessível ao solvente (SASA) para os diferentes estados de enovelamento da ubiquitina. Esta área superficial foi calculada por meio do software *GROMACS* [50]. As conformações mais desenoveladas possuem maior área superficial. Sendo assim, se uma determinada conformação possui uma maior área acessível ao solvente, é possível que existam mais moléculas do solvente interagindo com a superfície da proteína.

Tabela 7.1: Área acessível ao solvente para os diferentes estados de enovelamento simulados da ubiquitina. A área calculada para o estado de enovelamento desnaturado compreende uma média das 30 conformações (diferentes) usadas para representar o estado desnaturado.

Estado de enovelamento	SASA (nm^2)
Nativa	49,601
Desnaturada	$73,972 \pm 8,889$
Estendida	120,152

A Figura 7.2 contém o parâmetro de solvatação preferencial para os íons nos sistemas com a Ubiquitina na conformação desnaturada. O perfil de aumento e diminuição de Γ_{uc} é muito parecido com o que foi calculado para os sistemas com a ubiquitina nativa (Capítulo 6). Além disso,

Γ_{uc} é, em geral, maior em todas as concentrações de LI nos sistemas com a proteína desnaturada em comparação com o calculado na proteína nativa (Figura 6.1). A conformação desnaturada possui maior área superficial, à vista disso, mais resíduos (principalmente apolares) estão expostos ao solvente. Com a abertura da proteína, tornam-se possíveis mais interações entre a superfície da ubiquitina e mais moléculas do solvente. Além disso, o aumento do parâmetro preferencial de solvatação está relacionado com as interações com a água. O LI apresenta interação preferencial muito maior com a superfície da proteína do que com a água. Logo, o fato de água e LI não interagirem tão bem, ajuda a entender o aumento do parâmetro preferencial de solvatação para o LI.

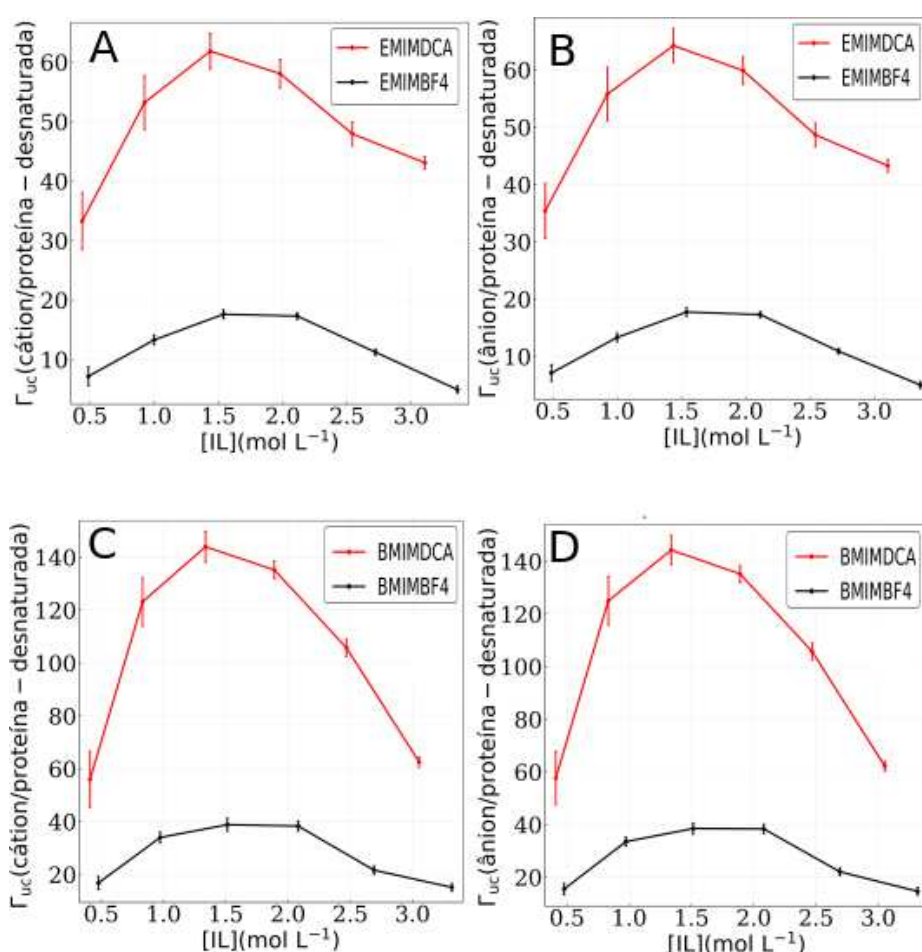


Figura 7.2: Γ para os líquidos iônicos na presença da proteína no estado de enovelamento desnaturado. A e B são os parâmetros de solvatação preferencial para os íons (Γ_{uc}) em sistemas contendo EMIMDCA e EMIMBF4. C e D contém Γ_{uc} para os íons nos sistemas com BMIMDCA e BMIMBF4. É possível notar que em todos os sistemas apresentados, Γ_{uc} aumenta de 0,5 mol L⁻¹ até 1,5 mol L⁻¹. Contudo, o aumento é menos pronunciado para os sistemas contendo o ânion BF4

A Figura 7.3 contém os parâmetros preferenciais de solvatação para a água (Γ_{uw}).

Podemos observar que a água é mais excluída nos sistemas que contém os cátions mais volumosos. É possível notar que conforme a concentração do LI aumenta, a água passa a ser menos excluída. Observando os dados, temos a seguinte situação: o cossolvente está solvatando preferencialmente a proteína enquanto a água é excluída. Este cenário indicaria uma provável desnaturação da ubiquitina. Se a interação proteína-solvente é muito favorável, não haverá um custo energético alto da proteína em explorar outras conformações na solução. Portanto, a proteína tenderá a se desenovelar, visto que a interação com o solvente é favorável. O efeito desnaturante ou estabilizante depende do efeito da estabilidade relativa do cossolvente em interações com a proteína nativa ou desenovelada. Em nosso sistema, claramente há um maior acúmulo sobre as conformações que compõem o tipo desenovelada.

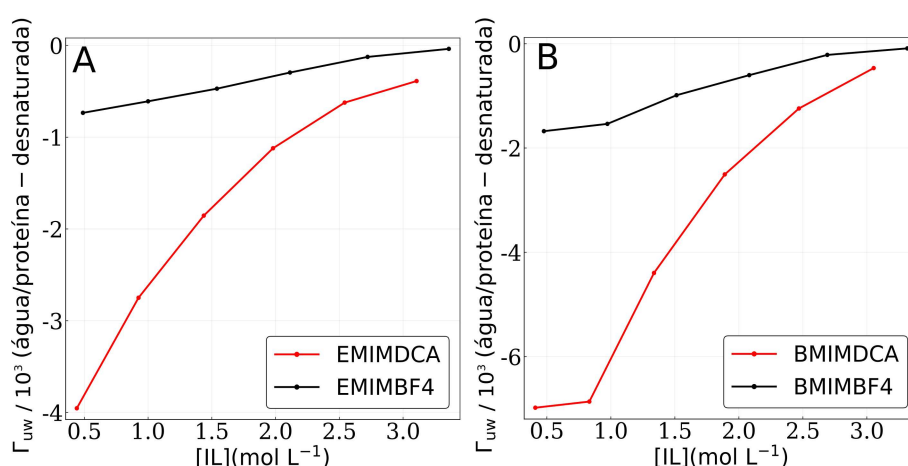


Figura 7.3: Parâmetro de solvatação preferencial da água para sistemas contendo EMIMDCA e EMIMBF4 (A), BMIMDCA e BMIMBF4 (B) e a Ubiquitina em seu estado de enovelamento desnaturado.

Conforme desenovelamos uma proteína, fazemos com que resíduos que antes estavam em seu interior tornem-se expostos à solução. Os resíduos apolares costumam estar localizados no centro da proteína (na sua conformação nativa) e com o desenovelamento passam a estar acessíveis à solução. Os LIs que estão sendo estudados nesta dissertação possuem cadeias apolares (especialmente pensando nos cátions EMIM e BMIM) e a interação entre componentes apolares (aliada à hidrofobicidade do LI) leva a um maior acúmulo dos íons para as conformações mais abertas.

Podemos observar na Figura 7.4 as MDDFs para o cátion BMIM nos três estados de enovelamento distintos. Nota-se que as MDFFs para o BMIM nos sistemas com os estados

desnaturado e estendido apresentam picos maiores, isto em comparação ao sistema com o estado nativo. Particularmente, a MDDF do BMIM no estado desnaturado apresentou o maior pico de todos (considerando o maior pico de cada gráfico (cor vermelha)).

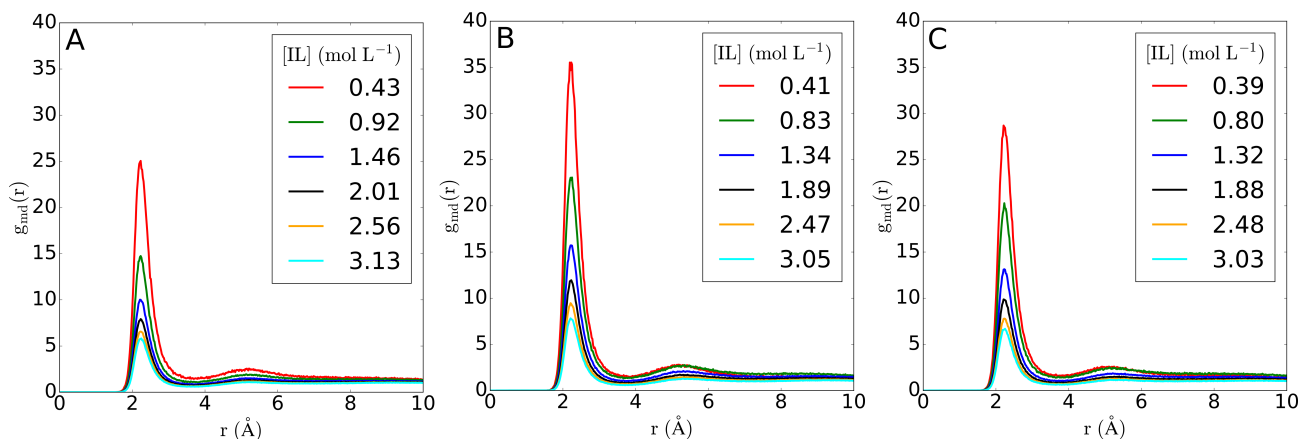


Figura 7.4: MDDFs para o cátion BMIM, na presença do ânion DCA, em sistemas contendo os estados de enovelamento nativo (A), desnaturado (B) e estendido (C).

A Figura 7.5 mostra as MDDFs para o ânion DCA. Similarmente ao que há na Figura 7.4, temos aqui três estados de enovelamento distintos da Ubiquitina. Para o cátion, vimos que conforme alteramos o estado de enovelamento da proteína há mudanças, principalmente, na altura dos picos das MDDFs. Entretanto, olhando para as MDDFs do ânion DCA, temos mais diferenças aparecendo conforme alteramos o grau de enovelamento. Primeiramente, as MDDFs provenientes dos sistemas com a Ubiquitina desnaturada (Figura 7.5B) e estendida (Figura 7.5C) possuem a altura do pico principal muito semelhante (no quadro C). O pico principal que acabamos de nos referir corresponde às interações não-específicas. Contudo, quando olhamos para o que acontece no pico referente à ligação de hidrogênio temos um mudança interessante. Na Figura 7.5A temos MDDFs para o DCA em sistemas com a Ubiquitina em seu estado nativo. Nestas MDDFs, existem dois picos bem definidos (suas alturas são semelhantes). O primeiro deles, em torno de 1,8 Å, corresponde à ligação de hidrogênio (como já mencionado algumas vezes). Conforme passamos para estados de enovelamento mais abertos, o pico referente as ligações de hidrogênio passa a se tornar menor do que o outro pico existente nas MDFFs. Além disso, ele é menor nos sistemas com a Ubiquitina desnaturada e estendida do que em comparação com o sistema com a Ubiquitina em seu estado de enovelamento nativo.

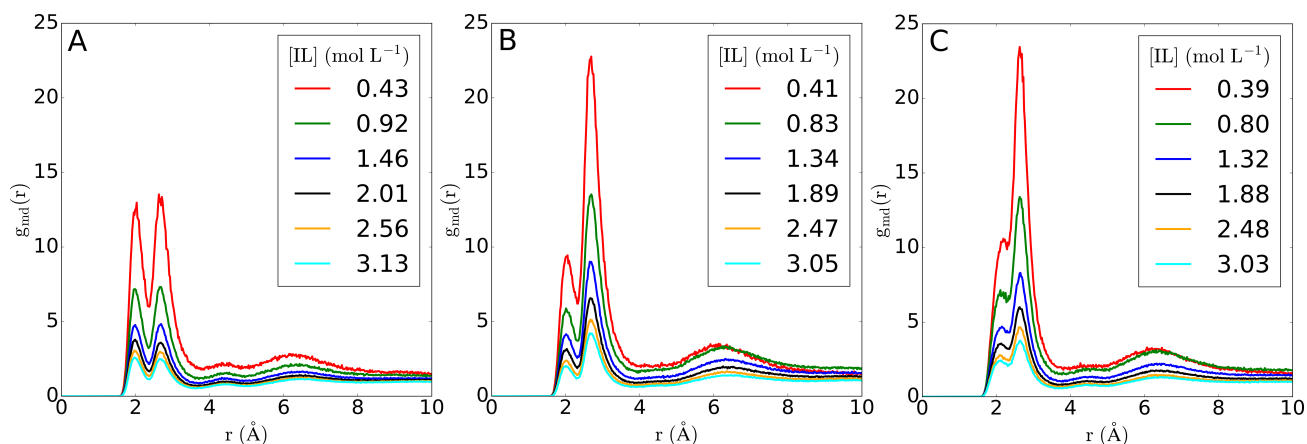


Figura 7.5: MDDFs para o ânion DCA, na presença do cátion BMIM, em sistemas contendo os estados de enovelamento nativo (A), desnaturado (B) e estendido (C).

A Figura 7.6 contém a contribuição, em gráfico de pizza, do backbone e cadeias laterais da proteína para a MDDF do cátion. Tomou-se a decomposição da MDDF dos cátions na distância de 2,5 Å, que corresponde ao pico principal da Figura 7.4 (a única diferença é que na Figura 7.6 estamos avaliando o cátion EMIM).

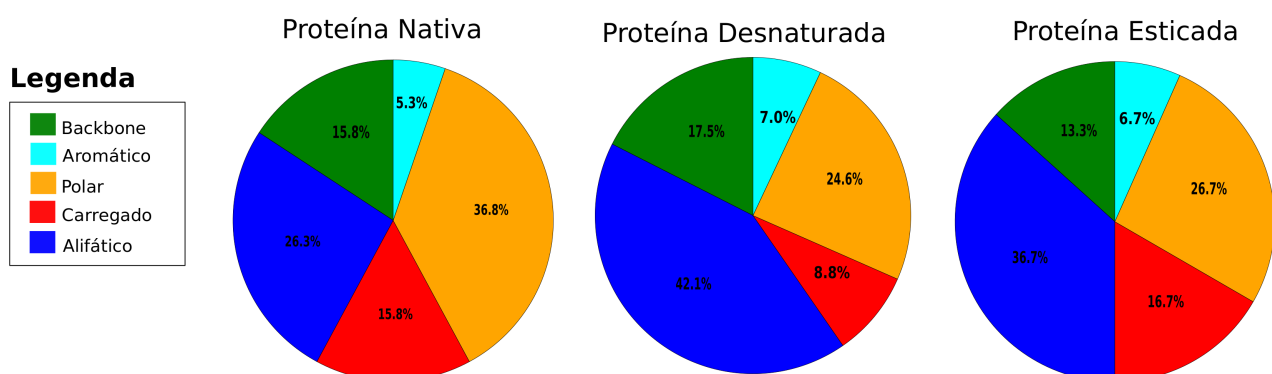


Figura 7.6: Contribuição de cada tipo de resíduo e da cadeia principal da proteína para a função de distribuição total do cátion (sistema com 0,5 mol L⁻¹ de EMIMDCA). Para o cálculo da contribuição de cada tipo de resíduo para a distribuição total tomou-se a distância de 2,6 Å (distância em que a MDDF para o EMIM apresenta o pico mais representativo). Como exemplo de interpretação, o gráfico indica que para a proteína nativa a função de distribuição do cátion ao redor do backbone corresponde a 15,8 % da função de distribuição ao redor da proteína toda.

Vemos que no sistema onde a proteína está parcialmente desenovelada a função de distribuição total possui maior contribuição por parte dos resíduos alifáticos. É interessante observar que, para a conformação nativa, as maiores contribuições para o primeiro pico da função de distribuição cátion-proteína devem-se aos resíduos polares e carregados. As conformações

parcialmente desenovelada e estendida apresentam aumento substancial da contribuição dos resíduos aromáticos e alifáticos da proteína.

Na Figura 7.4 pudemos ver que o pico em 2,5 Å da MDDF para o cátion, no sistema com a Ubiquitina estendida, foi menor em comparação ao mesmo pico no sistema com a Ubiquitina desnaturada. Contudo, o cátion tem maior acumulação sobre a superfície estendida. Para entender isto, vamos tomar o seguinte exemplo: a Equação 2.2.2 mostra que a IKB calculada para um componente do sistema depende da área superficial da soluto ($S(r)$). Portanto, se o soluto possui uma área superficial maior, a IKB também será maior. Na Figura 7.7 (A) temos que a MDDF para o cátion no sistema com a proteína desnaturada tem um pico maior do que o da MDDF na proteína estendida. Aparentemente, a interação do EMIM com a superfície da ubiquitina estendida é mais fraca, evidenciado pelo pico menor da MDDF. Contudo, a interação mais fraca está ocorrendo ao longo de uma superfície muito maior. O resultado disso é que, mesmo com a MDDF sendo menor, a IKB para o cátion no sistema com a proteína estendida é maior, indicando que a acumulação será efetivamente maior nesta conformação.

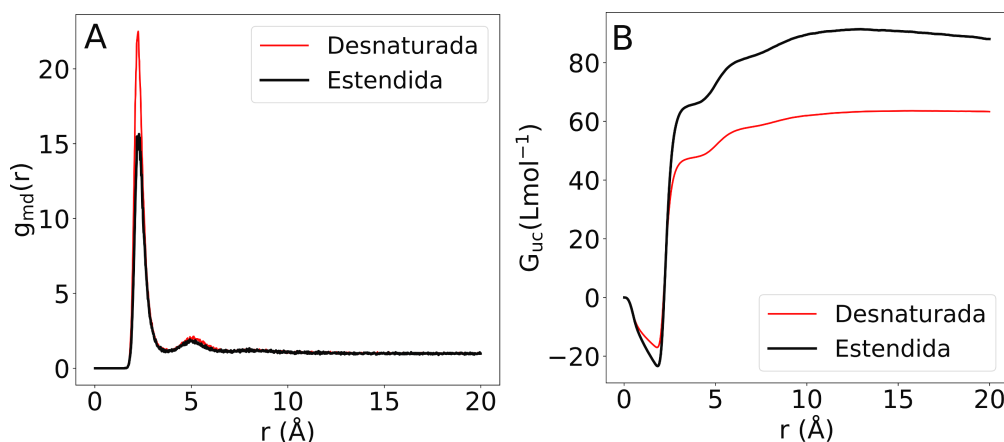


Figura 7.7: A figura (A) contém as MDDFs para o cátion EMIM em sistemas com 0,5 mol L⁻¹ do LI EMIMDCA nas conformações desnaturada e estendida. (B) contém as IKBs para EMIM nos mesmos sistemas.

A Figura 7.8 apresenta Γ_{uc} nos sistemas contendo a ubiquitina em sua conformação estendida. Γ_{uc} é ainda maior nos sistemas com a conformação estendida. A conformação estendida representa uma situação extrema em que a proteína possui todos os seus resíduos expostos. Todos os resíduos são expostos ao solvente e os íons do LI podem interagir com mais grupos apolares da proteína. Isto aumenta muito Γ_{uc} para os íons em todos os sistemas, principalmente aqueles que

possuem o ânion DCA.

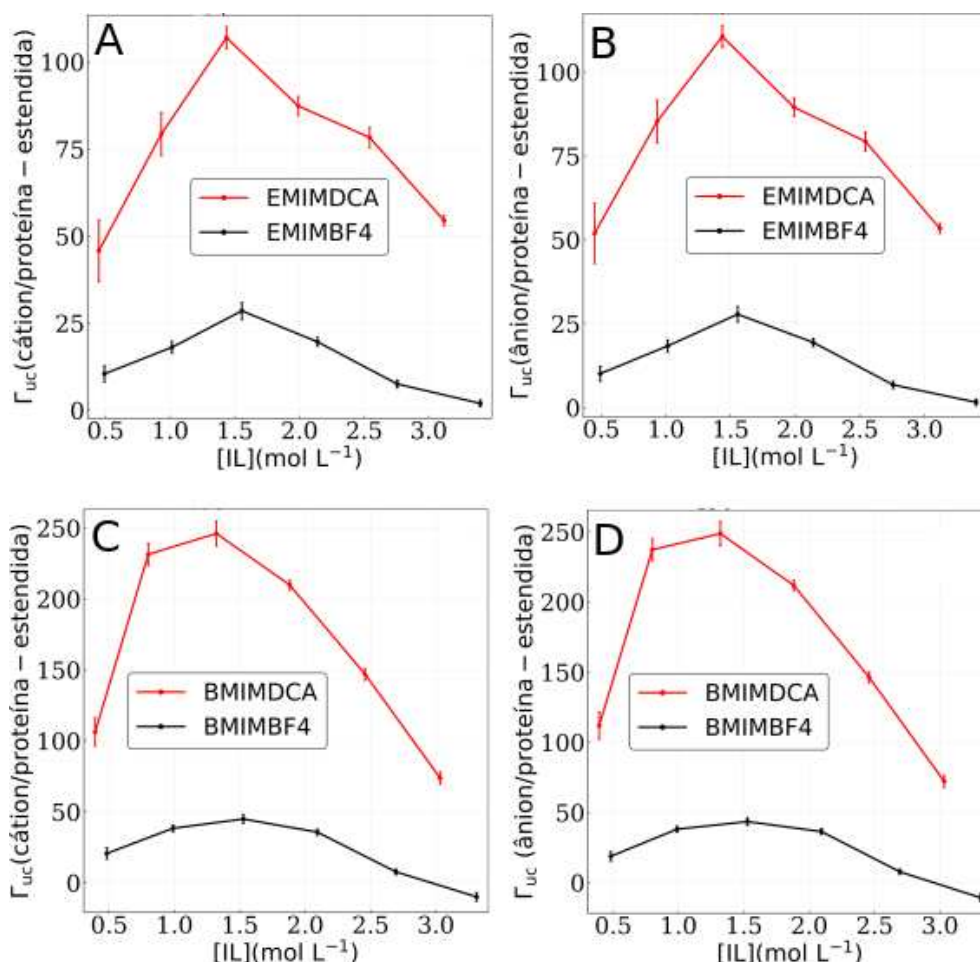


Figura 7.8: A figura contém Γ_{uc} os íons dos LIs EMIMDCA e EMIMBF (A e B) e dos LIs BMIMDCA e BMIMBF4 (C e D) nos sistemas que contém a conformação estendida da proteína.

Levando em conta sistemas com a proteína estendida, Γ_{uc} calculado sugere acúmulo ainda maior do LI nesta conformação do que se comparado às outras. O acúmulo é particularmente maior para os LIs que contém DCA. A água, por sua vez, é muito mais excluída se comparada com as conformações parcialmente desenovelada e nativa.

A Figura 7.9 contém uma representação da MDDF em 3D para sistemas com as conformações nativa e estendida da ubiquitina, além de 0,5 mol L⁻¹ do LI EMIMDCA. É possível notar que na proteína estendida a distribuição do DCA em maiores distâncias estende-se por toda sua extensão. As regiões que representam as ligações de hidrogênio (verde) são muito mais espalhadas ao longo da proteína estendida em comparação com a proteína nativa. O cátion EMIM está representado como pontos vermelhos. Em 7.9A podemos observar que a distribuição do

cátion ocorre mais distante da proteína em comparação ao ânion. Na 7.9B podemos ver que o cátion passa a se distribuir ao longo de toda a área superfície acessível da proteína. Desta forma, podemos observar que interações não-específicas passam a ocorrer em uma maior área na Ubiquitina estendida do que na Ubiquitina nativa.

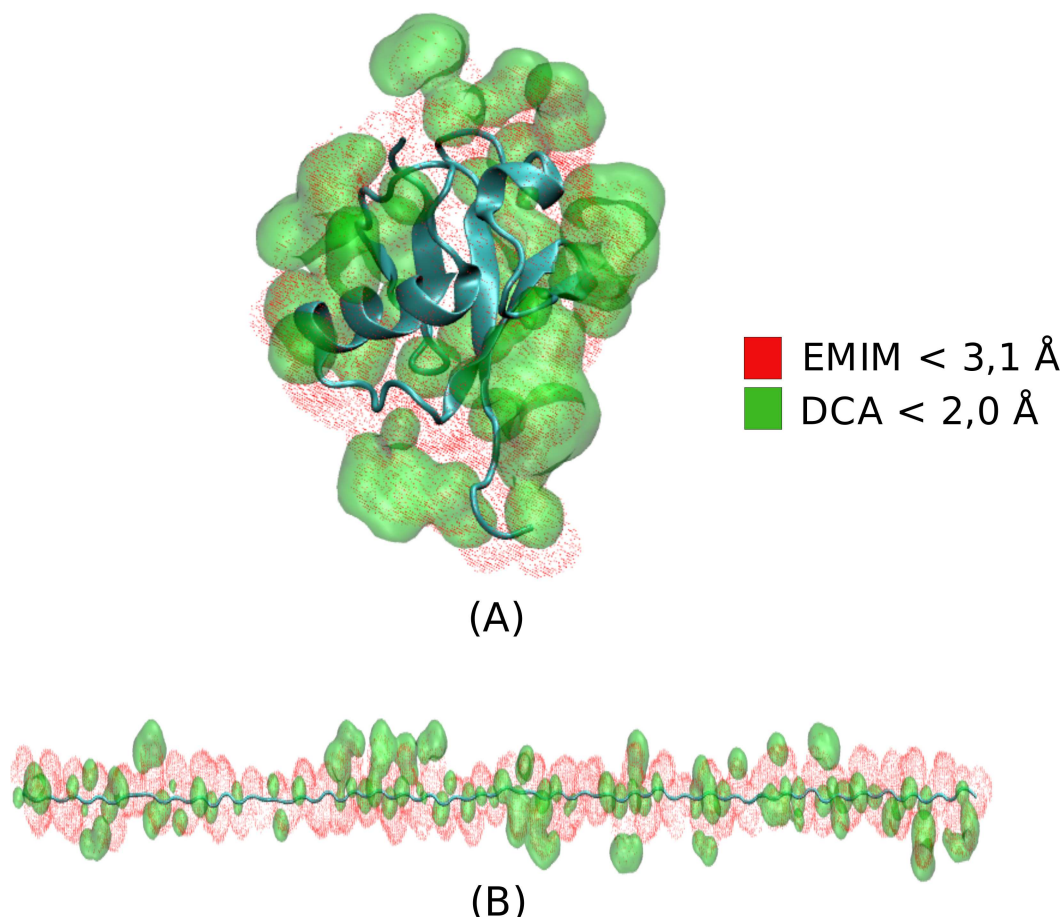


Figura 7.9: A figura contém a MDFF em 3D. Aqui, temos a representação, até uma distância de 2,0 Å da proteína, dos átomos que contribuem para a MDFF. Em verde temos a representação da distribuição do DCA na proteína nativa (A) e desnaturada (B). Os pontos vermelhos representam a distribuição do EMIM. Em função da distância que está sendo representada, observamos nas figuras as ligações de hidrogênio (regiões verdes).

Como foi observado na Figura 7.9, as ligações entre as moléculas de DCA com a superfície da proteína diminuíram na conformação estendida. Podemos observar este efeito de maneira mais efetiva na Figura 7.10. Em 7.10A vemos as MDFFs para o ânion DCA em variadas concentrações nos sistemas contendo a ubiquitina estendida. Quando observamos a decomposição da MDFF total do DCA em contribuições parciais dos tipos de resíduos da proteína (Figura 7.10B), vemos que o segundo pico possui grande contribuição dos resíduos alifáticos. Resumidamente, isto indica maior

presença de interações não-específicas.

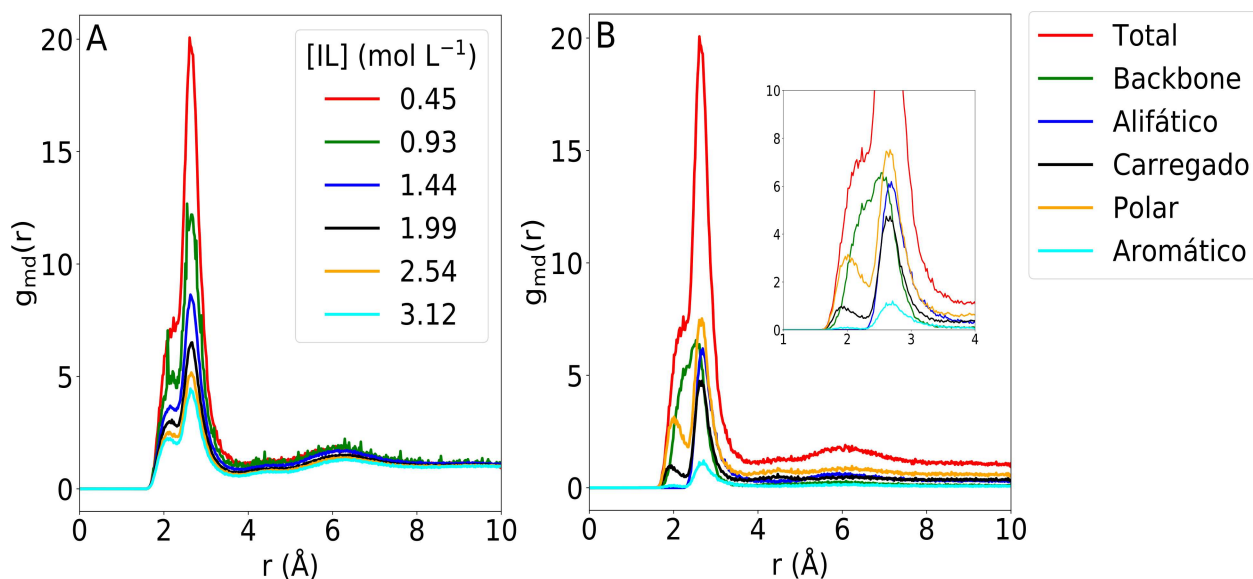


Figura 7.10: MDDFs para o ânion DCA em diferentes concentrações para os sistemas contendo a ubiquitina em sua conformação estendida (A). Decomposição da MDDF total em contribuições dos resíduos da proteína (B).

Após a observação da distribuição dos componentes da solução ao redor das diferentes conformações podemos pensar em uma abordagem que relaciona duas dessas diferentes estruturas. Vamos pensar num sistema bem simples. Este consiste de uma solução de água e LI semelhante ao que estudamos. Neste sistema, a ubiquitina pode assumir duas possíveis conformações, uma conformação nativa (N) e uma conformação desnaturada (D). A variação do parâmetro de solvatação preferencial ($\Delta\Gamma$) indica se o cossolvente apresenta maior acumulação sobre a estrutura nativa ou desnaturada da proteína. A variação do parâmetro de solvatação preferencial é calculada pela seguinte relação:

$$\Delta\Gamma_{uc} = \Gamma_{uc}^D - \Gamma_{uc}^N = \rho_c (\Delta G_{uc} - \Delta G_{uw}) \quad (7.1.1)$$

onde $\Delta\Gamma_{uc}$ é a variação do parâmetro de solvatação preferencial do estado nativo (Γ_{uc}^N) para o desnaturado (Γ_{uc}^D). ρ_c é a densidade do cossolvente no bulk, ΔG_{uc} e ΔG_{uw} são, respectivamente, a variação da integral de Kirkwood-Buff para a co-soluto/proteína e água/proteína dos sistemas com as conformações parcialmente desenovelada e nativa.

A Tabela 7.2 contém $\Delta\Gamma_{uc}$ levando em conta os sistema com a ubiquitina nativa e desnaturada para o cossolvente EMIMBF4. O cálculo foi realizado por meio da Equação 7.1.1.

Conforme a concentração do LI aumenta, há um aumento de $\Delta\Gamma_{uc}$, indicando maior acumulação sobre proteína desnaturada. O conteúdo da Tabela 7.2 indica solvatação preferencial da conformação parcialmente desenovelada pelos íons e, consequentemente, exclusão para a água.

Tabela 7.2: Variação do parâmetro de solvatação preferencial para cada componente dos sistemas contendo EMIMBF₄. A variação foi calculada utilizando os sistemas com as conformações parcialmente desenovelada e nativa. Os dados indicam forte interação do líquido iônico com a conformação parcialmente desenovelada se comparada à proteína nativa.

Concentração			
inicial de LI	$\Delta\Gamma_{uc}(EMIM)$	$\Delta\Gamma_{uc}(BF_4)$	$\Delta\Gamma_{uw}(H_2O)$
(mol L⁻¹)			
0,5	6 ± 4	6 ± 3	-11 ± 5
1,0	16 ± 3	16 ± 2	-29 ± 4
1,5	18 ± 2	18 ± 2	-31 ± 4
2,0	25 ± 2	25 ± 2	-42 ± 2
2,5	21 ± 1	21 ± 1	-33 ± 2
3,0	29 ± 1	29 ± 1	-43 ± 2

Apesar de, em todos os sistemas, Γ_{uc} diminuir com o aumento da concentração a partir de 1,5 mol L⁻¹ temos que $\Delta\Gamma_{uc}$ (de maneira geral) continuou aumentando. Tal fato evidencia que a acumulação dos LIs, em grandes concentrações, para as conformações desnaturada e estendida é maior, de forma que a diferença entre ambas e a proteína nativa resulta em $\Delta\Gamma_{uc} > 0$.

Capítulo 8

Efeito da troca de ânion e de cátion

Até o momento, vimos que os sistemas que continham as menores concentrações de líquido iônico apresentam solvatação preferencial da ubiquitina por parte dos íons. Além disso, observamos que há maior acúmulo do líquido iônico para as conformações desnaturada e estendida da proteína. Neste capítulo vamos estudar o efeito relativo sobre a distribuição ao longo da solução do cátion/ânion quando alteramos o seu par iônico no sistema. Demonstraremos que o ânion DCA faz com que a densidade relativa do cátion aumente muito em distâncias próximas da proteína. Além disso, os cátions mais volumosos fizeram com que os ânions possuíssem maior densidade relativa próxima da proteína. Apresentaremos que, apesar de se distribuírem diferentemente ao longo da solução, os íons dos LIs possuem integrais de Kirkwood-Buff com mesmo valor. Este comportamento trata-se de uma consequência da eletroneutralidade da solução.

8.1 Parâmetros de solvatação preferencial dos íons

Vamos exemplificar a discussão que segue com os sistemas contendo $0,5 \text{ mol L}^{-1}$ de LI e com o soluto sendo a ubiquitina nativa. Na Figura 8.1 temos o parâmetro de solvatação preferencial calculado para o ânion na presença do EMIM e do BMIM. Qualitativamente, temos agora (em relação ao que tínhamos com EMIM) uma molécula maior e mais apolar. Consequentemente, sua afinidade pela água é menor do que a do EMIM. Podemos notar que o Γ_{uc} para o ânion no sistema que contém o cátion BMIM é maior do que para o cátion EMIM, para as concentrações menores. Conforme a concentração de LI aumenta, Γ_{uc} do ânion diminui, mostrando que pelo fato de usarmos um cátion maior, o ânion passa a ser preferencialmente excluído (nas maiores concentrações) e acumula mais em concentrações abaixo de $2,0 \text{ mol L}^{-1}$.

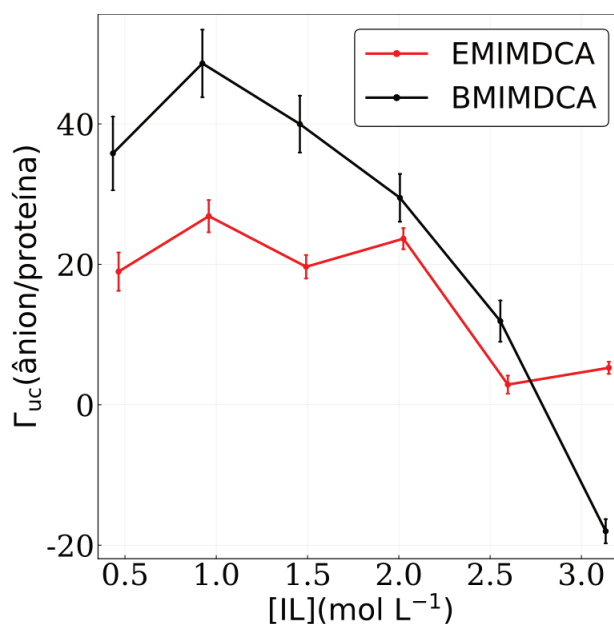


Figura 8.1: Parâmetro de solvatação preferencial para o ânion DCA na presença da ubiquitina nativa, em sistemas com 0,5 mol L⁻¹ de EMIMDCA e BMIMDCA.

Observando a Figura 8.1 vemos que, na concentração inicial (0,5 mol L⁻¹), a troca do cátion EMIM para BMIM ocasiona em um aumento em Γ_{uc} em torno de 17. Contudo, a mesma mudança para o sistema com BF₄ (Figura 8.2) culminou em um aumento claramente menor. Porém, se olharmos para as maiores concentrações de LI, temos a situação oposta. A diferença de Γ_{uc} para o sistema com EMIM para com BMIM na presença do DCA é em torno de 25 unidades de Γ , o mesmo tipo de variação para os sistemas com BF₄ é em torno de 20 unidades de Γ . Além disso, a concentração bulk de LI nos sistemas com DCA fica em torno de 3,2 mol L⁻¹, enquanto que para os sistemas com BF₄ temos quase 3,5 mol L⁻¹. Isto mostra que o LI foi mais excluído nos sistemas com BF₄ ao ponto da concentração de LI no bulk exceder a concentração inicial em quase 17% seu valor inicial.

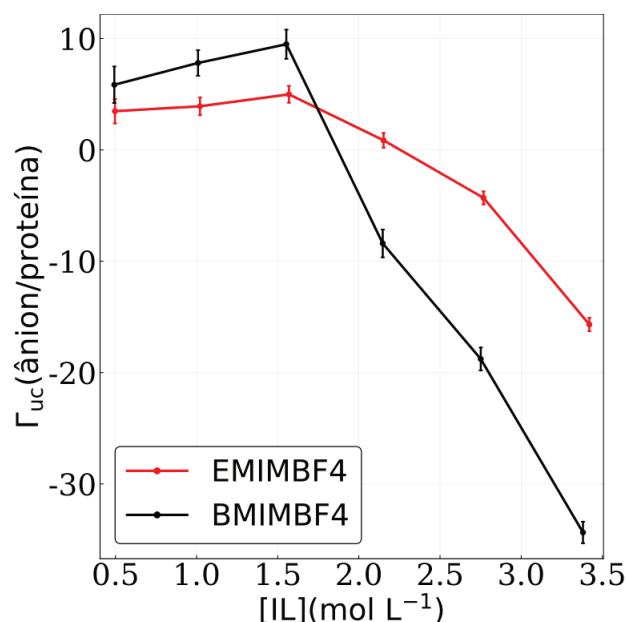


Figura 8.2: Parâmetro de solvatação preferencial em sistemas contendo a proteína em seu estado nativo e 0,5 mol L⁻¹ dos líquidos iônicos EMIMBF4 e BMIMBF4. O aumento da cadeia carbônica do cátion (EMIM para BMIM) é acompanhada da diminuição Γ_{uw} para concentrações menores do que 2,0 mol L⁻¹.

8.2 Funções de distribuição de mínima-distância para os íons

A Figura 8.3 contém as MDDFs para o ânion DCA em dois sistemas diferentes (um contendo EMIMDCA e o outro BMIMDCA), podemos notar a existência de dois picos, o primeiro deles (tanto para DCA quanto BF4) aparece por volta de 1,9 Å e é referente às ligações de hidrogênio. Os picos no sistema com EMIMDCA são bem definidos e, a partir do segundo pico, a função decresce rapidamente, com pequenas bandas (4,3 Å e 6,0 Å) indicando pequeno acúmulo em distâncias maiores a partir da superfície proteica. O resultado disto é que em 10 Å a função converge para o regime bulk ($g_{md}(r) = 1$). No sistema com o BMIMDCA, o ânion apresenta picos maiores. Sendo assim, a densidade relativa do DCA, nas distâncias 1,9 e 2,6 Å, é maior quando o DCA forma par iônico com o BMIM do que com o EMIM.

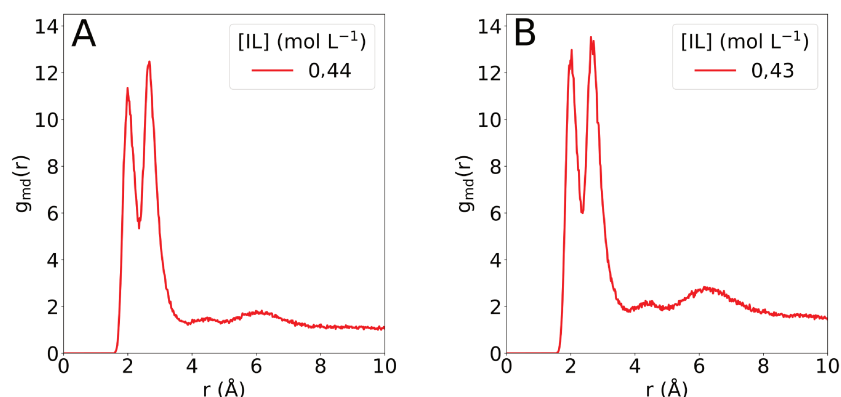


Figura 8.3: Funções de distribuição de mínima distância para o DCA nos sistemas com EMIMDCA (A) e BMIMDCA (B). Nota-se que que MDDF no sistema (B) apresenta picos maiores, ou seja, no sistema (B) é maior a densidade relativa do ânion DCA ao redor da proteína em comparação com o sistema (A).

A Figura 8.4 contém a MDDF para o ânion BF₄ nos sistemas com EMIMBF₄ e BMIMBF₄. Pode-se observar que a MDDF apresentou uma sensibilidade bem menor (levando em consideração a mudança na altura do pico resultante da troca de EMIM para BMIM) se comparada à do DCA. Os picos em 1,9 e 2,6 Å têm um ligeiro aumento quando o cátion passa de EMIM para BMIM.

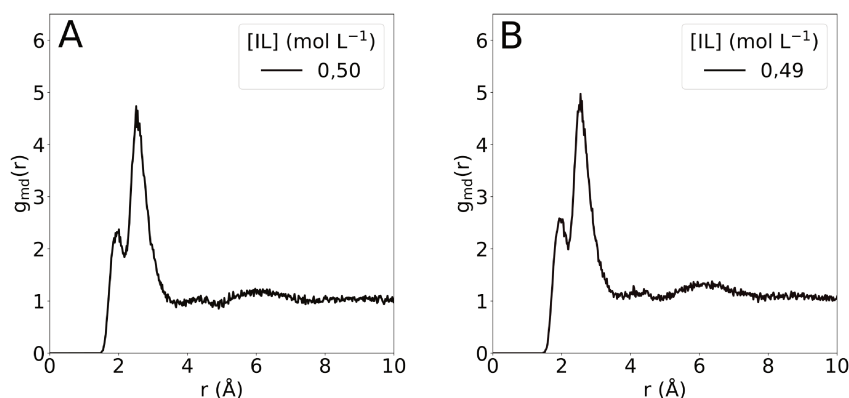


Figura 8.4: Funções de distribuição de mínima distância para o BF₄ nos sistemas com EMIMBF₄ (A) e BMIMBF₄ (B). A mudança, na altura dos picos, resultante da troca do cátion EMIM para BMIM é muito pequena.

Nas figuras 8.5 e 8.6 vemos que a diferença nas funções de distribuição reflete-se em um aumento muito grande da IKB para o ânion DCA. Para o ânion BF₄ o aumento é bem menos pronunciado. Como a IKB mede diretamente acúmulo ou exclusão de uma componente do solvente sob o soluto, temos que há maior acúmulo de DCA sobre a proteína em comparação ao BF₄.

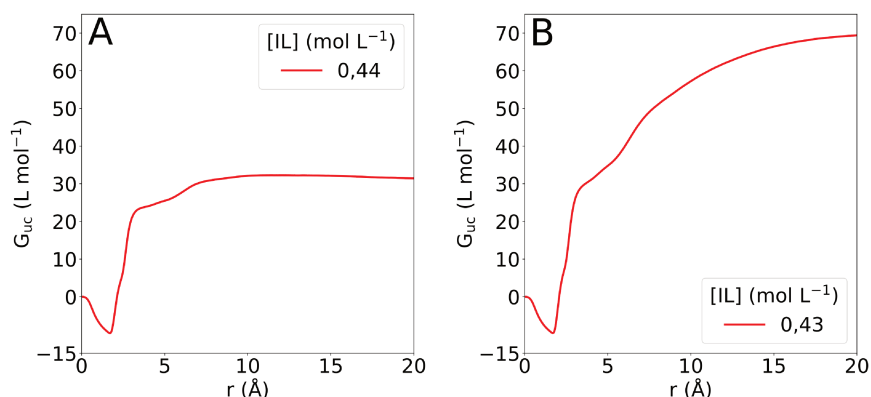


Figura 8.5: IKBs para o DCA nos sistemas contendo o EMIMDCA (A) e o BMIMDCA (B). O ânion DCA apresenta uma IKB muito maior no sistema com o cátion BMIM.

Pela observação das Figuras 8.5B e 8.6B é possível notar que para distâncias menores as curvas são parecidas, contudo, conforme a distância aumenta, as diferenças tornam-se maiores (BMIMDCA apresenta IKB muito maior do que BMIMBF₄). Isto indica que as principais diferenças não ocorrem na primeira camada de solvatação, mas sim em distâncias maiores.

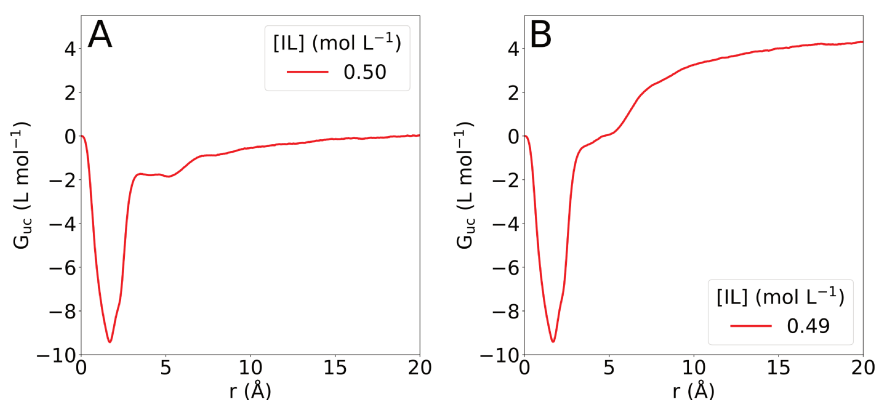


Figura 8.6: IKBs para o BF₄ nos sistemas contendo EMIMBF₄(A) e BMIMBF₄(B). O aumento do tamanho da cadeia carbônica do cátion culminou num pequeno aumento da IKB para o BF₄ no sistema com BMIM.

Pelos resultados apresentados até agora vemos que a troca do cátion causa alteração da distribuição e aumento da IKB maior para o do que o BF₄. Vamos agora mudar o ânion e observar o que acontece com a função de distribuição do cátion. Nas figuras 8.7 e 8.8 estamos observando mudanças na MDDF do cátion resultantes da presença de DCA ou BF₄. É importante lembrar que, em cada figura, estamos observando a função de distribuição do mesmo cátion, que faz par iônico com ânions diferentes. Na Figura 8.7 temos que as funções de distribuição apresentam a mesma forma, isto é, o pico surge na mesma distância e a função converge de maneira semelhante.

Contudo, a altura do pico é muito diferente nos dois sistemas. Para o ânion BF_4 o máximo da distribuição é em torno de 8,75 enquanto que para o DCA é 17,0.

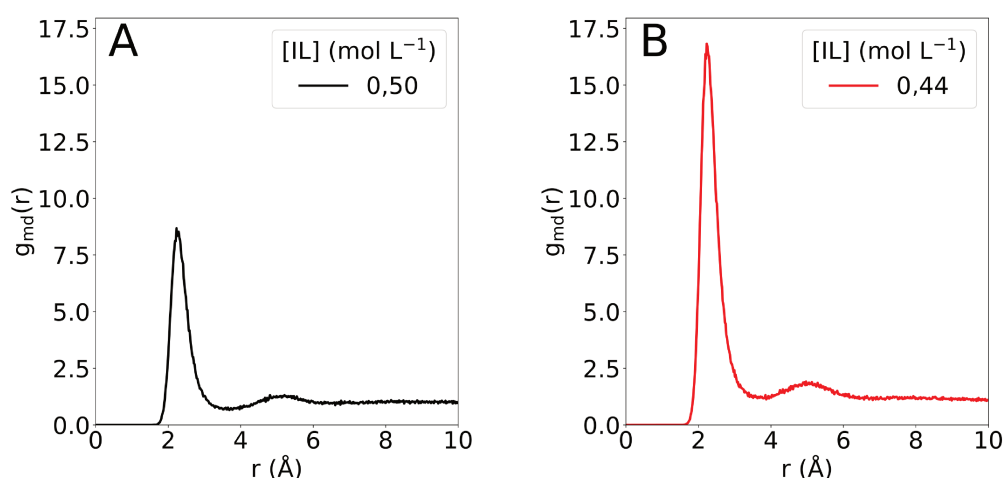


Figura 8.7: MMDFs para o cátion EMIM nos sistemas contendo BF_4 (A) e DCA (B).

É observado o mesmo comportamento da Figura 8.7 na Figura 8.8. O sistema com o ânion BF_4 apresenta máximo da distribuição igual a 12,5 enquanto que para o DCA é aproximadamente 25,0.

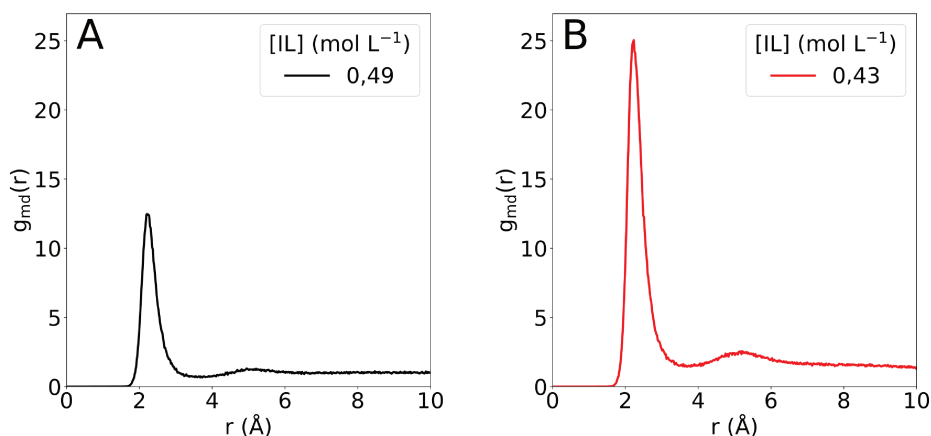


Figura 8.8: MMDFs para o cátion BMIM nos sistemas contendo BF_4 (A) e DCA(B).

O ânion DCA levou a um aumento na MMDF muito maior para os cátions se comparado ao BF_4 . Os ânions possuem a capacidade de formar ligações de hidrogênio com a superfície da proteína, ficando a uma distância de $1,9 \text{ Å}$ da proteína. A primeira camada de solvatação, nos sistemas aqui mencionados, é composta por água e os íons de carga negativa. Tal fato faz com que os cátions se agrupem nas proximidades dessa camada, ainda que de maneira difusa. Tão maior será

o número de moléculas de cátion próximos da proteína quanto maior for o número de moléculas do ânion ocupando regiões próximas da proteína, até que ocorra saturação de íons ao redor da proteína. Os dados indicam que o DCA solvata preferencialmente a proteína de maneira mais intensa do que o BF₄. Isto explica porque o EMIM que faz par iônico com DCA acumula muito mais que o EMIM que faz par iônico com BF₄.

Os resultados indicam se o ânion apresenta maior acumulação sobre a superfície da proteína, o seu par iônico também apresentará um aumento na sua acumulação em comparação ao que ocorreria se, por exemplo, o cátion estivesse em contato com um ânion com menor Γ . A análise das MDDFs para os ânions DCA e BF₄, mostra que o ânion DCA apresenta maior capacidade de estabelecer ligações de hidrogênio com os átomos da superfície da proteína do que o BF₄. Esta afirmação vem do fato de que, o pico em 1,9 Å da MDDF do DCA tem forma mais definida e é mais alto. Além disso, no Capítulo 7 vimos que o pico em 1,9 Å tem maior contribuição sendo proveniente dos nitrogênios, sendo estes aceptores de ligações de hidrogênio (sendo assim, uma evidência muito forte da importância das ligações de hidrogênio na interação entre DCA e a proteína).

A Figura 8.9 contém as MDDFs (A) e IKBs (B) para os íons EMIM e DCA (o sistema em questão contém 3,0 mol L⁻¹ de LI e a proteína no estado nativo). Os íons distribuem-se ao longo da solução distintamente, por exemplo, a MDDF do DCA possui picos em distâncias mais próximas da proteína. Tal fato indica que o DCA tem densidade maior próximo da ubiquitina, indicando interação entre a ubiquitina e o ânion DCA. O EMIM, por sua vez, estabelece interações que ocorrem em distâncias maiores. Quando calculamos a IKB, levando em conta os domínios da proteína ($r < 20,0$ Å), temos o mesmo resultado para ambos os íons. A eletroneutralidade dos íons estabelece que em uma solução iônica, a soma de cargas elétricas negativas atraia uma soma igual de cargas elétricas positivas. Isto faz com que a IKB para os íons sejam iguais, pois será requerido que, o bulk da solução seja eletricamente neutro.

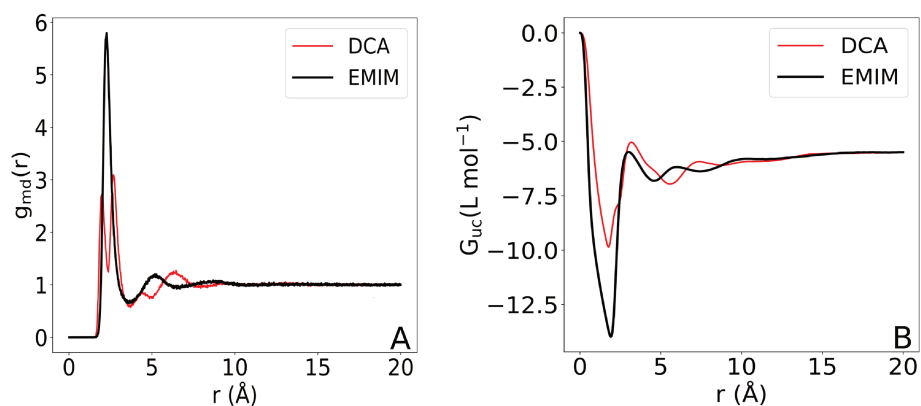


Figura 8.9: Em (A) temos as MDDFs para o cátion EMIM e o ânion DCA, (B) contém as IKBs para os mesmos componentes.

Capítulo 9

Considerações Finais

O entendimento da estabilidade das proteínas em solução é um problema que pode ser estudado por meio da interação preferencial relativa de diferentes conformações com o solvente. Uma forma de estudar a interação de moléculas com a superfície da proteína é por meio do parâmetro preferencial de solvatação. Esta quantidade mede o acúmulo/exclusão sobre a proteína de uma substância relativamente a outra. O parâmetro de solvatação preferencial pode ser experimentalmente obtido através do experimento de diálise no equilíbrio. Além disso, podemos calcular esse parâmetro usando a teoria de soluções de Kirkwood-Buff e as MDDFs. O entendimento da estabilidade de proteína em solução é importante por conta de diversos processos biológicos requerem a proteína solubilizadas. Neste sentido, esta dissertação buscou estudar a estrutura de solvatação de uma proteína amplamente estudada, a ubiquitina, em sistemas contendo líquidos iônicos, que são solventes com forte apelo científico devido às suas características físico-químicas. Neste capítulo, iremos revisar os principais resultados e apresentar as conclusões que surgem a partir de toda a discussão realizada até então na dissertação. Realizamos a caracterização da estrutura de solvatação da Ubiquitina pelos líquidos iônicos EMIMDCA, EMIMBF₄, BMIMDCA, BMIMBF₄ e a água utilizando as funções de distribuição de mínima-distância.

9.1 Variação da concentração

Observamos que os LIs solvatam preferencialmente a ubiquitina nos sistemas contendo as menores concentrações. Vimos que a acumulação de íons sobre a superfície proteica é maior quando o ânion é o DCA e o cátion é o BMIM. Além disso, pudemos observar que em uma concentração

específica Γ_{uc} apresentou valor máximo, indicando que, nesta situação, a acumulação relativa do LI alcança seu maior valor. A água, por sua vez, sempre foi excluída da superfície proteica (apresentando $\Gamma_{uw} < 0$). Entretanto, conforme a concentração do LI aumentava, Γ_{uw} tornava-se menos negativo, o que nos faz supor que, em grandes concentrações de LI (maiores do que às aquelas simuladas) a água passaria a ter $\Gamma_{uw} > 0$, solvatando preferencialmente a proteína.

Pela análise das MDDFs notamos que a água e os ânions estudados interagem com a superfície da proteína mediante ligações de hidrogênio, contudo, o ânions também podem estabelecer interações não-específicas. Os cátions, por sua vez, apresentam apenas um pico significativo em sua MDDF. Este, da mesma forma com o que ocorre com os ânions, também está relacionado às interações de natureza não-específicas.

De maneira geral, vimos que as menores concentrações culminaram em maiores densidades relativas (picos maiores nas MDDFs) dos íons. Além disso, houve a maior acumulação relativa de íons em relação a água com $1,5 \text{ mol L}^{-1}$ de líquido iônico.

9.2 Estado de enovelamento da Ubiquitina

Realizamos simulações em três estados de enovelamento diferentes da Ubiquitina. Pela análise dos parâmetros de solvatação preferencial, pudemos constatar que conforme aumentava o grau de desenovelamento da proteína, maior era o valor de Γ_{uc} . Assim, o LI apresentou maior acumulação relativa sobre as conformações mais abertas da ubiquitina, tendo os sistemas com a conformação estendida apresentando os maiores valores de Γ_{uc} . O perfil das curvas contendo o parâmetro de solvatação preferencial dos íons foi parecido na maioria dos sistemas. Há um crescimento Γ_{uc} no intervalo entre $0,5$ e $1,5 \text{ mol L}^{-1}$ do LI e decréscimo no intervalo entre $1,5$ até $3,0 \text{ mol L}^{-1}$.

Utilizamos como um exemplo sistemas contendo EMIMBF₄ e as conformações nativa e desnaturada para o cálculo de $\Delta\Gamma$. Sempre tivemos $\Delta\Gamma > 0$, indicando que o cosolvente (EMIMBF₄) acumula-se mais sobre a proteína desnaturada. Independente da apresentação de $\Delta\Gamma$ apenas para o EMIMBF₄, ao longo da dissertação, todos os LIs simulados apresentaram Γ maiores nos sistemas com as conformações desnaturada e estendida da Ubiquitina. De forma que, $\Delta\Gamma$ seria positivo para todos os sistemas. Os dados sugerem que os LIs simulados seriam compostos

desnaturantes para a Ubiquitina.

Pela observação das MDDFs, pudemos ver que há um aumento substancial no pico das interações não-específicas (2,6 Å para os ânions e 2,5 Å para os cátions) nos sistemas com a Ubiquitina estendida e desnaturada em relação aos sistemas com a Ubiquitina nativa. Isto indica que há um aumento das interações-não específica entre o LI e as conformações mais abertas da Ubiquitina. De maneira geral, os LIs, possivelmente, levariam a Ubiquitina ao desenovelamento da Ubiquitina em soluções com as concentrações simuladas de LI (sendo este efeito particularmente maior nos sistemas contendo o ânion DCA).

9.3 Variação do tipo de cátion/ânion

Quando discutimos os resultados levando em conta a variação do tipo do cátion/ânion estávamos preocupados em entender se havia alguma mudança na distribuição de um cátion ao trocarmos o par iônico, por exemplo. Analisando as MDDFs para os ânions pudemos notar que para o DCA, um aumento no tamanho do cátion culminava em picos mais altos na MDDF do DCA. O mesmo não foi observado para o BF₄.

Observou-se que ao trocarmos o DCA pelo BF₄, a MDDF para o cátion apresentou diminuição considerável para o pico em 2,5 Å. De maneira geral, os LIs que contém o DCA apresentaram maior acumulação do que aqueles com BF₄. O aumento da cadeia carbônica do cátion, em um primeiro momento, culminou em uma maior acumulação do ânion sobre a superfície da proteína. Porém, a exclusão do LI nas maiores concentrações pode ser indicativo de que em grandes concentrações (maiores do que as que foram simuladas) ele faça com que uma maior quantidade de água acumule sobre a proteína.

Apesar de apresentarem distribuições muito distintas, cátions e ânions possuem integrais de Kirkwood-Buff iguais. Isto porque a solução tem que ser eletricamente neutra. Portanto, a natureza das interações do cátion afeta a distribuição do ânion, e vice-versa.

9.4 Perspectivas futuras

Ao longo desta dissertação estudamos a estrutura de solvatação de variadas conformações da Ubiquitina por soluções aquosas de líquidos iônicos imidazólicos. Apesar de ter sido feito um vasto estudo das funções de distribuição, a quantidade de líquidos iônicos estudados foi limitada. Além disso, não foram feitas análises que reportam informações energéticas à respeito da interação entre as componentes do sistema. Desta forma, possíveis análises poderiam ser realizadas visando a informação energética que falta para o estudo das interações na solução.

Pode muito ser interessante adicionar outros tipos de ânions e cátions ao sistema, com o objetivo de estudar como esses novos íons irão alterar (ou não) a distribuição dos íons provenientes do líquido iônico.

Como foi apresentado na introdução, o parâmetro de solvatação preferencial pode ser experimentalmente medido por meio da técnica de diálise. Para um futuro próximo, podemos buscar na literatura sistemas semelhantes aos trabalhados nessa dissertação, e que tenham medido Γ . Desta forma, poderíamos confrontar resultados teóricos e experimentais.

Além disso, em uma eventual sequência desta linha no doutorado, pode vir a ser interessante a utilização de campos de força polarizáveis na avaliação de interações eletrostáticas, que são fundamentais para estrutura e função das proteínas.

Bibliografia

- [1] J. O'M Brockis e A. K. N. Reddy. *Modern Eletrochemistry*. Plenum, 1970.
- [2] M. D. Joshia e J. L. Anderson. "Recent advances of ionic liquids in separation science and mass spectrometry". Em: *RSC Adv.* 2 (2012), pp. 5470–5484. DOI: 10.1039/C2RA20142A.
- [3] I. Krossing et al. "Why Are Ionic Liquids Liquid? A Simple Explanation Based on Lattice and Solvation Energies". Em: *J. Am. Chem. Soc.* 128 (2006), 1342713434. DOI: 10.1021/ja073579a.
- [4] D. Christen, J. H. Griffiths e J. Sheridan. "The Microwave Spectrum of Imidazole; Complete Structure and the Electron Distribution from Nuclear Quadrupole Coupling Tensors and Dipole Moment Orientation". Em: *A J. Phys. Sci.* 36 (2014), pp. 1378–1385. DOI: 10.1515/zna-1981-1220.
- [5] T. L. Greaves e C. J. Drummond. "Protic Ionic Liquids: Properties and Applications". Em: *Chem. Rev.* 108 (2008), pp. 206–237. DOI: 10.1021/cr068040u.
- [6] H. Zhao. "Protein Stabilization and Enzyme Activation in Ionic Liquids: Specific Ion Effects". Em: *J Chem Techmol Biotechnol* 91 (2016), pp. 25–50. DOI: 10.1002/jctb.4837.
- [7] M. Haberler, C. Schröder e O. Steinhauser. "Hydrated ionic liquids with and without solute: The influence of water content and protein solutes". Em: *J. Chem. Theory Comput.* 8 (2012), pp. 3911–3928. ISSN: 15499618. DOI: 10.1021/ct300191s.
- [8] Z. Lei et al. "Introduction: Ionic Liquids". Em: *Chem. Rev.* 117 (2017), pp. 6633–6635. ISSN: 15206890. DOI: 10.1021/acs.chemrev.7b00246.
- [9] J. Qu et al. "Ionic Liquids as Novel Lubricants and Additives for Diesel Engine Applications". Em: *Tribol. Lett.* 35 (2009), pp. 181–189. ISSN: 1573-2711. DOI: 10.1007/s11249-009-9447-1.

- [10] M. G. Lin, H. Chen e H. Dai. "Ionic Liquids Based Electrolytes for Rechargeable Batteries". Em: *Mat. Matt.* 13 (2018), pp. 31–40.
- [11] K. Goossens et al. "Ionic Liquid Crystals: Versatile Materials". Em: *Chem. Rev.* 116 (2016), pp. 4643–4807. DOI: 10.1021/cr400334b.
- [12] E. J. Maginn. "Molecular simulation of ionic liquids: current status and future opportunities". Em: *J. Phys.: Condens. Matter* 21 (2009), pp. 1–17. DOI: 10.1088/0953-8984/21/37/373101.
- [13] J. Smiatek. "Aqueous ionic liquids and their effects on protein structures: an overview on recent theoretical and experimental results". Em: *J. Phys.: Condens. Matter* 29 (2017), pp. 1–19. DOI: 10.1088/1361-648X/aa6c9d.
- [14] R. Patel, M. Kumari e A. B. Khan. "Recent advances in the applications of ionic liquids in protein stability and activity: A review". Em: *Appl. Biochem. Biotechnol.* 172 (2014), pp. 3701–3720. ISSN: 15590291. DOI: 10.1007/s12010-014-0813-6.
- [15] M. Harbeler, C. Schröder e O. Steinhauser. "Solvation studies of a zinc finger protein in hydrated ionic liquids". Em: *Phys. Chem. Chem. Phys.* 13 (2011), pp. 6955–6969. ISSN: 1463-9076. DOI: 10.1039/c0cp02487b.
- [16] F. van Rantwijk e R. A. Sheldon. "Biocatalysis in ionic liquids". Em: *Chem. Rev.* 107 (2007), pp. 2757–2785. ISSN: 00092665. DOI: 10.1021/cr050946x.
- [17] A. L. Lehninger, D. L. Nelson e M. M. Cox. *Princípios de Bioquímica de Lehninger*. 6. ed. Artmed Editora LTDA, 2014, p. 1250. ISBN: 9781429234146.
- [18] T.W. G. Solomons e C. B. Fryhle. *Química Orgânica - Volume 2*. 10nd. LTC, 2011, pp. 507–550. ISBN: 9788521620341.
- [19] A. Soroushova et al. "The Collagen Suprafamily: From Biosynthesis to Advanced Biomaterial Development". Em: *Adv. Mat.* 31 (2019), pp. 1–39. DOI: 10.1002/adma.201801651.
- [20] Y. Zhao et al. "Engineering plant virus resistance: from RNA silencing to genome editing strategies". Em: *Plant Biotechnol. J.* (2019), pp. 1–9. DOI: 10.1111/pbi.13278.
- [21] A. E. Rydeen, E. M. Brustad e G. J. Pielak. "Osmolytes and Protein-Protein Interactions". Em: *J. Am. Chem. Soc.* 140 (2018), pp. 7441–7444. DOI: 10.1021/jacs.8b03903.
- [22] A. Fasil, M. Usma e A. Mudasser. *Cellular Osmolytes: From Chaperoning Protein Folding to Clinical Perspectives*. Springer Singapore, 2017, pp. 35–53. DOI: 10.1007/978-981-10-3707-8_2.

- [23] P. H. Yancey. "Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses." Em: *J Exp Biol* 208 (2005), pp. 2819–2830. DOI: 10.1242/jeb.01730.
- [24] L. M. F. Holthauzen e D. W. Bolen. "Mixed osmolytes: The degree to which one osmolyte affects the protein stabilizing ability of another". Em: *Protein Sci.* 16 (2007), pp. 293–298. DOI: 10.1110/ps.062610407.
- [25] S. N. Timasheff. "Protein-solvent preferential interactions , protein hydration , and the modulation of biochemical reactions by solvent components". Em: *Biochemistry* 99.15 (2002), pp. 9721–9726.
- [26] J. Smiatek. "Aqueous ionic liquids and their effects on protein structures: an overview on recent theoretical and experimental results". Em: *J. Phys.: Condens. Matter* 29 (2017), pp. 1–20. DOI: 10.1088/1361-648X/aa6c9d.
- [27] D. R. MacFarlane et al. "Lewis base ionic liquids". Em: *Chemical Communications* 18 (2006), pp. 1905–1917.
- [28] P. Lozano et al. "Dynamic structure–function relationships in enzyme stabilization by ionic liquids". Em: *Biocat. Biotrans.* 23 (2009), pp. 169–176.
- [29] P. Lozano et al. "Stabilization α -Chymotrypsin by Ionic Liquids in Transesterification Reactions". Em: *Biotech. Bioeng.* 75 (2001), pp. 563–569.
- [30] R. Tomaš et al. "Interactions in aqueous solutions of imidazolium chloride ionic liquids [Cnmim][Cl] (n=0, 1, 2, 4, 6, 8) from volumetric properties, viscosity B-coefficients and molecular dynamics simulations". Em: *J. Mol. Liq.* 254 (2018), pp. 267–271. ISSN: 01677322. DOI: 10.1016/j.molliq.2018.01.076.
- [31] E. J. Maginn. "Molecular simulation of ionic liquids: Current status and future opportunities". Em: *J. Phys. Condens. Matter* 21 (2009). ISSN: 09538984. DOI: 10.1088/0953-8984/21/37/373101.
- [32] J. Smiatek. "Aqueous ionic liquids and their effects on protein structures: an overview on recent theoretical and experimental results". Em: *J. Phys. Condens. Matters.* 29 (2017), pp. 1–20. DOI: 10.1088/1361-648X/aa6c9d.
- [33] C. Schröder. "Proteins in Ionic Liquids: Current Status of Experiments and Simulations". Em: *Top. Curr. Chem.* 375.2 (2017). ISSN: 03401022. DOI: 10.1007/s41061-017-0110-2.
- [34] D. Chandler. 1st. Vol. 1. Oxford University Press, 1987, p. 288. ISBN: 978-0195042771.

- [35] D. A. McQuarrie. 1st. Vol. 1. University Science Books, 2000, p. 641. ISBN: 1891389157.
- [36] A. H. Narten e H. A. Levy. "Liquid Water: Molecular Correlation Functions from X-Ray Diffraction". Em: *J. Chem. Phys.* 55 (1971), pp. 2263–2269. DOI: 10.1063/1.1676403.
- [37] A. K. Soper. "The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MP". Em: *Chem. Phys.* 258 (2000), pp. 121–137. DOI: 10.1016/S0301-0104(00)00179-8.
- [38] L. Martínez e S. Shimizu. "Molecular Interpretation of Preferential Interactions in Protein Solvation: A Solvent-Shell Perspective by Means of Minimum-Distance Distribution Functions". Em: *J. Chem. Theory Comput.* 13 (2017), pp. 6358–6372. ISSN: 15499626. DOI: 10.1021/acs.jctc.7b00599.
- [39] K. Newman. "Kirkwood-Buff solution theory: Derivation and applications". Em: *Chemical Society Reviews* 23 (1994), pp. 31–40. ISSN: 03060012. DOI: 10.1039/CS9942300031.
- [40] J. G. Kirkwood e F. P. Buff. "The statistical mechanical theory of solutions. i". Em: *J. Chem. Phys.* 19 (1951), pp. 774–777. ISSN: 00219606. DOI: 10.1063/1.1748352.
- [41] C. L. Boon. "The Kirkwood – Buff theory and the effect of cosolvents on biochemical reactions". Em: *J. Chem. Phys.* 121 (2004), pp. 9147–9155. DOI: 10.1063/1.1806402.
- [42] D. G. Hall. "The Kirkwood–Buff theory of solutions". Em: *Trans. Faraday Soc.* 67 (1971), pp. 2516–2524. DOI: 10.1039/TF9716702516.
- [43] C. L. Boon e S. Shimizu. "The Kirkwood – Buff theory and the effect of cosolvents on biochemical reactions". Em: *J. Chem. Phys.* 121 (2004), pp. 9147–9155. DOI: 10.1063/1.1806402.
- [44] J. G. Kirkwood. "Statistical mechanics of fluid mixtures". Em: *J. Chem. Phys.* 3 (1935), pp. 300–313. ISSN: 00219606. DOI: 10.1063/1.1749657.
- [45] S. Shimizu e D. J. Smith. "Preferential hydration and the exclusion of cosolvents from protein surfaces". Em: *J. Chem. Phys.* 121 (2004), pp. 1148–1154. DOI: 10.1063/1.1759615.
- [46] E. S. Courtenay et al. "Vapor Pressure Osmometry Studies of OsmolyteProtein Interactions: Implications for the Action of Osmoprotectants in Vivo and for the Interpretation of "Osmotic Stress" Experiments in Vitro". Em: *Biochem.* 35 (2000), pp. 4455–4471. DOI: 10.1021/bi9928871.
- [47] B. M. Baynes e B. L. Trout. "Proteins in Mixed Solvents: A Molecular-Level Perspective". Em: *J. Phys. Chem. B* 107 (2003), pp. 14058–14067. DOI: 10.1021/jp0363996.

- [48] J. C. Lee e S. N. Timasheff. "The stabilization of proteins by sucrose". Em: *J. Biol. Chem.* 256 (1981), pp. 7193–7201.
- [49] S. N. Timasheff. "Control of protein stability and reactions by weakly interacting cosolvents: The simplicity of the complicated." Em: *Advances in Protein Chemistry* 51 (1998), pp. 355–432.
- [50] M. J. Abraham et al. "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers". Em: *SoftwareX* 2 (2015), pp. 19–25. ISSN: 23527110. DOI: 10.1016/j.softx.2015.06.001.
- [51] A. R. Leach. "Molecular Modelling - Principles and applications". Em: *Mol. Model. - Princ. Appl.* Second. Person Education, 2001. Cap. 7, pp. 353–406. ISBN: 0-582-3821-6.
- [52] N. H. Morgon e K. Coutinho. "Fundamentos de simulação por dinâmica molecular". Em: *Métodos de Química Teórica e Modelagem Molecular*. Livraria da Física, 2007. Cap. 12, pp. 413–451.
- [53] K. G. Sprenger, Vance W. Jaeger e Jim Pfaendtner. "The general AMBER force field (GAFF) can accurately predict thermodynamic and transport properties of many ionic liquids". Em: *J. Phys. Chem. B* 119 (2015), pp. 5882–5895. ISSN: 15205207. DOI: 10.1021/acs.jpcc.5b00689.
- [54] C. Y. Son et al. "First-Principles United Atom Force Field for the Ionic Liquid BMIM+BF₄⁻: An Alternative to Charge Scaling". Em: *J. Phys. Chem. B* 120 (2016), pp. 3560–3568. ISSN: 15205207. DOI: 10.1021/acs.jpcc.5b12371.
- [55] A. M. Namba, V. B. da Silva e C. H. T. P. da Silva. "Dinâmica molecular: Teoria e aplicações em planejamento de fármacos". Em: *Eclet. Quim.* 33 (2008), pp. 13–24. DOI: 10.1590/S0100-46702008000400002.
- [56] W. L. Jorgensen e J. Tirado-Rives. "The OPLS Force Field for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin". Em: *J. Am. Chem. Soc.* 110 (1988), pp. 1657–1666. DOI: 10.1021/ja00214a001.
- [57] W. L. Jorgensen, J. Tirado-Rives e D. S. Maxwell. "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids". Em: *J. Am. Chem. Soc.* 118 (1996), pp. 11225–11236. DOI: 10.1021/ja9621760.

- [58] A. D. MacKerell Jr et al. "All-atom empirical potential for molecular modeling and dynamics studies of proteins". Em: *J. Phys. Chem. B* 102 (1998), pp. 3586–3616. DOI: 10.1021/jp973084f.
- [59] L. D. Schuler, D. Xavier e F. V. G. Wilfred. "An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase". Em: *J. Comput. Chem.* 22 (2001), pp. 1205–1218.
- [60] W. D. Cornell et al. "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules". Em: *J. Am. Chem. Soc.* 117 (1995), pp. 5179–5197. DOI: 10.1021/ja00124a002.
- [61] B. Doherty, X. Zhong e O. Acevedo. "A Virtual Site OPLS Force Field for Imidazolium-Based Ionic Liquids". Em: *J. Phys. Chem. B* 11 (2018), pp. 2962–2974. DOI: 10.1021/acs.jpcc.7b11996.
- [62] L. Verlet. "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules". Em: *Phys. Rev.* 159 (1967), pp. 98–103.
- [63] M. Parrinello e A. Rahman. "Polymorphic transitions in single crystals: A new molecular dynamics method". Em: *J. Appl. Phys.* 52 (1981), pp. 7182–7190. ISSN: 00218979. DOI: 10.1063/1.328693.
- [64] M. Parrinello e A. rahman. "Crystal Structure and Pair Potential". Em: *Phys. Rev. Lett.* 45 (1980), pp. 1196–1199. ISSN: 1365-3075, 0033-4545. DOI: 10.1351/pac198052010233.
- [65] H. J.C. Berendsen et al. "Molecular dynamics with coupling to an external bath". Em: *J. Chem. Phys.* 81.8 (1984), pp. 3684–3690. ISSN: 00219606. DOI: 10.1063/1.448118.
- [66] L. Martínez et al. "PACKMOL: A package for building initial configurations for molecular dynamics simulations". Em: *J. Comput. Chem.* 30 (2009), pp. 2157–2164. DOI: 10.1002/jcc.21224.
- [67] J. M. Martínez e L. Martínez. "Packing Optimization for Automated Generation of Complex System's Initial Configurations for Molecular Dynamics and Docking". Em: *J. Comp. Chem.* 24 (2003), pp. 819–825.
- [68] W. F. van Gunsteren e H. J. C. Berendsen. "Molecular dynamics computer simulations; Methodology, applications and perspectives in chemistry". Em: *Angewedte Chemie* 102 (1990), pp. 1020–1055. DOI: 10.1002/ange.19901020907.
- [69] D. Frenkel e B. Smit. *Understanding Molecular Simulations: from Algorithms to Applications*. 2nd. Academic Press - London, 2002. ISBN: 9780080519982.

- [70] M. P. Allen e D. J. Tildesley. *Computer simulation of liquids*. first edition. Oxford University Press, 1987, p. 385. ISBN: 0198553757.
- [71] I. G. Tironi, R. Sperb e W. F. van Gunsteren P. E. Smith. "A generalized reaction field method for molecular dynamics simulations". Em: *J. Chem. Phys.* 102 (1995), pp. 5451–5458. DOI: 10.1063/1.469273.
- [72] T. Darden, D. York e L. Pedersen. "Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems". Em: *J. Chem. Phys.* 98 (1993), pp. 10089–10092. DOI: 10.1063/1.464397.
- [73] R. W. Hockey e J. W. Eastwood. *Computer Simulation using Particles*. McGraw-Hill, 1981, p. 540. ISBN: 9780852743928.
- [74] Z. Pouramini, A. Mohebbi e M. H. Kowsari. "Atomistic insights into the thermodynamics , structure , and dynamics of ionic liquid 1-hexyl-3-methylimidazolium hexa fl uorophosphate via molecular dynamics study". Em: *J. Mol. Liq.* 246 (2017), pp. 39–47. ISSN: 0167-7322. DOI: 10.1016/j.molliq.2017.09.043.
- [75] J. C. Phillips et al. "Scalable molecular dynamics with NAMD". Em: *J. Comput. Chem.* 26 (2005), pp. 1781–1802. ISSN: 01928651. DOI: 10.1002/jcc.20289.
- [76] L. Leherter e D. P. Vercauteren. "Reduced Point Charge Models of Proteins: Effect of Protein-Water Interactions in Molecular Dynamics Simulations of Ubiquitin Systems". Em: *J. Phys. Chem. B* 121 (2017), pp. 9771–9784. DOI: 10.1021/acs.jpcb.7b06355.
- [77] V. Rogov et al. "Interactions between Autophagy Receptors and Ubiquitin-like Proteins Form the Molecular Basis for Selective Autophagy". Em: *Molecular Cell* 53 (2014), pp. 167–178. DOI: 10.1016/j.molcel.2013.12.014..
- [78] A. Ciechanover. "The ubiquitin–proteasome pathway: on protein death and cell life". Em: *EMBO J.* 17 (1998), pp. 7151–7160. DOI: 10.1093/emboj/17.24.7151.
- [79] S. Kumar, C. E. Bugg e Cook W. J. "Structure of ubiquitin refined at 1.8 Å resolution". Em: *J. mol. Biol.* 194 (1987), pp. 531–544.
- [80] J. D. Watson e F. H. C. Crick. "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid". Em: *Nature* 171 (1953), pp. 737–738.
- [81] W. L. Jorgensen et al. "Comparinon of simple potencial functions for simulating liquid water". Em: *J. Chem. Phys.* 79 (1983), pp. 926–935. ISSN: 09725938. DOI: 10.1063/1.445869.

- [82] E. J. Haug, J. S. Arora e K. Matsui. "A steepest-descent method for optimization of mechanical systems". Em: *J. Optim. Theory Appl.* 19 (1976), pp. 401–424. ISSN: 00223239. DOI: 10.1007/BF00941484.
- [83] T. Darden, D. York e L. Pedersen. "Particle mesh Ewald: An $N\log(N)$ method for Ewald sums in large systems". Em: *J. Chem. Phys.* 98 (1993), pp. 10089–10092. DOI: 10.1063/1.464397.
- [84] S. H. Northrup e J. A. McCammon. "Simulation Methods for Protein Structure Fluctuations". Em: *Biopolymers* 19 (1980), pp. 1001–1016. DOI: 10.1002/bip.1980.360190506.
- [85] P. Subrata e P. Sandip. "Molecular insights into the Role of Aqueous Trehalose Solution on Temperature-Induced Protein Denaturation". Em: *J. Phys. Chem. B* 119 (2015), pp. 1598–1610. ISSN: 15205207. DOI: 10.1021/jp510423n.