



**UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE QUÍMICA**

**VINICIUS PICCOLI**

**Efeitos de íon-comum e múltiplos eletrólitos na solvatação de proteínas por  
líquidos iônicos**

**CAMPINAS  
2025**

**VINICIUS PICCOLI**

**Efeitos de Íon-comum e múltiplos eletrólitos na solvatação de proteínas por  
líquidos iônicos**

**Tese apresentada ao Instituto de Química da  
Universidade Estadual de Campinas como parte dos  
requisitos exigidos para a obtenção do título de  
Doutor em ciências**

**Orientador: Prof. Dr. Leandro Martínez**

**ESTE TRABALHO CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA  
PELO ALUNO VINICIUS PICCOLI E ORIENTADA PELO PROF. DR. LEANDRO  
MARTÍNEZ.**

**CAMPINAS**

**2025**

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

P581e Piccoli, Vinicius, 1996-  
Efeitos de íon-comum e múltiplos eletrólitos na solvatação de proteínas por líquidos iônicos / Vinicius Piccoli. – Campinas, SP : [s.n.], 2025.

Orientador: Leandro Martínez.  
Tese (doutorado) – Universidade Estadual de Campinas (UNICAMP), Instituto de Química.

1. Líquidos iônicos. 2. Solvatação. 3. Dinâmica molecular. 4. Teoria de soluções de Kirkwood-Buff. 5. Funções de distribuição de mínima distância. I. Martínez, Leandro, 1979-. II. Universidade Estadual de Campinas (UNICAMP). Instituto de Química. III. Título.

Informações complementares

**Título em outro idioma:** Common-ion and multiple-electrolyte effects on protein solvation by ionic liquids

**Palavras-chave em inglês:**

Ionic liquids

Solvation

Molecular dynamics

Kirkwood–Buff solution theory

Minimum-distance distribution functions

**Área de concentração:** Físico-Química

**Titulação:** Doutor em Ciências

**Banca examinadora:**

Leandro Martínez [Orientador]

Agílio Alexandre Henriques Pádua

Kalil Bernardino

Luiz Fernando Mercier Franco

Nelson Henrique Morgon

**Data de defesa:** 27-03-2025

**Programa de Pós-Graduação:** Química

**Objetivos de Desenvolvimento Sustentável (ODS)**

Não se aplica

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

- ORCID do autor: <https://orcid.org/0000-0003-3379-8760>

- Currículo Lattes do autor: <https://lattes.cnpq.br/8358573667119731>

## **BANCA EXAMINADORA**

**Prof. Dr. Leandro Martínez (Orientador)**

**Prof. Dr. Agílio Alexandre Henriques Pádua (École Normale Supérieure de Lyon - França)**

**Prof. Dr. Kalil Bernardino (Universidade Federal de São Carlos)**

**Prof. Dr. Luís Fernando Mercier Franco (Faculdade de Engenharia Química - Unicamp)**

**Prof. Dr. Nelson Henrique Morgon (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.**

***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

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brasil, Processos n°: 2020/04916-3 (Doutorado) e 2023/02167-1 (BEPE).

Expresso minha profunda gratidão ao professor Leandro Martínez. Desde o início do mestrado, tenho o privilégio de contar com sua orientação. Nossas discussões foram sempre enriquecedoras, contribuindo de maneira significativa para minha formação acadêmica. Sou imensamente grato pela oportunidade de realizar meu doutorado em um ambiente de pesquisa bem equipado, com acesso a centros de computação de alto desempenho, essenciais para a realização das simulações deste trabalho. Foi uma honra ser orientado por um pesquisador tão competente e inspirador, cuja trajetória acadêmica serve de modelo para minha própria carreira científica. Tenho enorme respeito e admiração pelo Leandro e agradeço por tudo que aprendi com ele e o agradeço por toda a paciência e boa vontade em sempre me ajudar a melhorar nos mais variados aspectos científicos.

Agradeço especialmente à Isadora. Sua presença, mesmo à distância em alguns momentos, foi fundamental para que eu enfrentasse os desafios do doutorado com mais leveza e alegria. Os momentos que compartilhamos sempre trouxeram equilíbrio à minha rotina, e tê-la ao meu lado é um privilégio que sempre buscarei honrar. O amor da Isa torna meus dias mais leves e felizes, e não tenho dúvidas de que desejo compartilhar minha vida com ela enquanto estiver neste mundo. Estar com a Isa é, sem dúvidas, a melhor parte de todos os dias.

Aos meus pais, Valdecir e Maria Anete, meu mais profundo agradecimento por todo apoio, compreensão e incentivo ao longo desta jornada. Seu carinho e confiança foram pilares fundamentais na minha trajetória acadêmica e pessoal, e sem eles nada disso seria possível.

Sou grato também a todos os membros do nosso grupo de pesquisa, em especial aos amigos Ander e Pamella, cujo companheirismo tornou o dia a dia do doutorado mais leve. As conversas durante o almoço e o café da tarde foram preciosas, ajudando a tornar essa jornada mais descontraída e agradável.

Gostaria de expressar meu agradecimento aos professores Agílio Pádua e Margarida Costa Gomes, por me receberem tão bem em seu grupo de pesquisa durante meu tempo na França. O período em Lyon foi enriquecedor de inúmeras formas, e todas as conversas que tive, especialmente com o Agílio, foram extremamente frutíferas, tanto em âmbito científico quanto em temas variados. Ambos foram muito gentis e atenciosos comigo, e sou profundamente grato pelo ano em que trabalhei com eles.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio contínuo à ciência brasileira e, em particular, pelo financiamento dos primeiros meses da minha bolsa de doutorado (Processo nº 140846/2020-4).

Minha gratidão também ao CEPID (Centro de Pesquisa, Inovação e Difusão) CCE&S – *Center for Computing in Engineering and Sciences*, apoiado financeiramente pela FAPESP (Processos 2013/08293-7, 2018/24293-0 e 2019/17874-0, referentes ao supercomputador Coaraci) e pelo CNPq (Processo 301909/2022-9). Os cálculos realizados nesta tese foram possíveis graças à infraestrutura fornecida pelo CCE&S.

Por fim, agradeço a Deus, por me conceder saúde, força e a graça de percorrer este caminho desafiador, sempre amparado por minha família e por pessoas tão especiais ao longo dessa trajetória.

# Resumo

Os líquidos iônicos (LIs) são sais líquidos a temperaturas inferiores a 100°C, caracterizados por baixa pressão de vapor, alta estabilidade térmica, viscosidade elevada (em comparação com solventes orgânicos regulares e água) e ampla capacidade de solubilização de compostos diversos. No contexto biológico, os LIs são comumente aplicados para aumentar a estabilidade e a solubilidade de proteínas em soluções aquosas. Neste contexto, uma das principais lacunas na área é a compreensão detalhada das interações entre LIs e proteínas, essencial para o desenvolvimento de LIs seletivos com aplicações biotecnológicas.

Este trabalho investiga o papel dos LIs na solvatação de proteínas, tendo a ubiquitina como modelo, por meio de simulações de dinâmica molecular. A estrutura e a termodinâmica de solvatação da ubiquitina em soluções aquosas de LIs foram analisadas por funções de distribuição de mínima-distância e pela teoria de soluções de Kirkwood–Buff. Foram estudados os cátions 1-etil-3-metilimidazólio ( $[EMIM]^+$ ) e 1-butil-3-metilimidazólio ( $[BMIM]^+$ ), combinados com os ânions tetrafluoroborato ( $[BF_4]^-$ ), dicianamida ( $[DCA]^-$ ), cloreto ( $Cl^-$ ) e nitrato ( $NO_3^-$ ).

Os resultados demonstram que os LIs formam interações específicas com proteínas, influenciando significativamente a solubilidade e a desnaturação. O ânion  $[DCA]^-$  destacou-se pela formação de ligações de hidrogênio, desempenhando um papel desnaturante mais acentuado em combinação com o cátion  $[BMIM]^+$ . Contudo, o aumento da concentração de LIs reverte esse efeito, tornando  $[EMIM]^+$  mais eficaz em concentrações elevadas. Este estudo evidencia a natureza não aditiva das combinações de ânions, o impacto da hidrofobicidade dos cátions e os mecanismos cooperativos que governam a solvatação proteica. Esses achados oferecem subsídios para o desenvolvimento de LIs com propriedades ajustáveis para aplicações biotecnológicas específicas.

# Abstract

Ionic liquids (ILs) are salts that remain in the liquid state at temperatures below 100°C, characterized by low vapor pressure, high thermal stability, high viscosity, and broad solvation capabilities for a wide range of compounds. In biological contexts, ILs are commonly used to enhance the stability and solubility of proteins in aqueous solutions. However, a major gap in the field is the detailed understanding of IL–protein interactions, which is essential for the design of selective ILs with biotechnological relevance.

This work investigates the role of ILs in protein solvation using ubiquitin as a model system through molecular dynamics simulations. The structural and thermodynamic aspects of ubiquitin solvation in aqueous IL solutions were analyzed using minimum distance distribution functions and the Kirkwood–Buff theory of solutions. The cations 1-ethyl-3-methylimidazolium ( $[\text{EMIM}]^+$ ) and 1-butyl-3-methylimidazolium ( $[\text{BMIM}]^+$ ) were studied in combination with the anions tetrafluoroborate ( $[\text{BF}_4]^-$ ), dicyanamide ( $[\text{DCA}]^-$ ), chloride ( $\text{Cl}^-$ ), and nitrate ( $\text{NO}_3^-$ ).

The results demonstrate that ILs establish specific protein interactions, significantly affecting solubility and denaturation. The  $[\text{DCA}]^-$  anion was particularly notable for forming strong hydrogen bonds, exhibiting a more pronounced denaturing effect when paired with  $[\text{BMIM}]^+$ . However, increasing IL concentration reverses this trend, rendering  $[\text{EMIM}]^+$  more effective at higher concentrations. This study underscores the non-additive nature of anion combinations, the influence of cation hydrophobicity, and the cooperative mechanisms that govern protein solvation. These findings offer valuable insights for the design of ILs with tunable properties for targeted biotechnological applications.

# Lista de Figuras

- 1.1 Estruturas químicas dos íons mais encontrados em LIs. Os cátions, à direita, são, em sua maioria, compostos orgânicos, frequentemente apresentando cadeias carbônicas extensas. Em contraste, os ânions, à esquerda, podem ser tanto de natureza inorgânica quanto orgânica, variando amplamente em tamanho e na presença de diferentes grupos funcionais. . . . . 29
- 1.2 Modelo comparativo entre íons regulares e líquidos iônicos, destacando as diferenças estruturais e de simetria. . . . . 30
- 1.3 Modelo esquemático ilustrando a variabilidade das interações entre LIs e proteínas. As interações podem incluir forças dispersivas em regiões apolares (representadas em cinza), ligações de hidrogênio com moléculas de água (triângulos azuis) e interações eletrostáticas em regiões carregadas, demonstrando a complexidade das interações preferenciais em diferentes superfícies proteicas. Figura inspirada a partir da discussão sobre interações de LIs com proteínas disponível em [58]. . . . . 35
- 2.1 (A) Representação esquemática da adição de uma proteína a uma solução aquosa. (B) Ciclo termodinâmico utilizado para avaliar a energia livre de solvatação de uma proteína em água. . . . . 42
- 2.2 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 partícula vermelha e da partícula azul apontada pelas setas. Figura reproduzida com autorização do autor a partir de [89]. . . . . 44

2.3	Função de distribuição radial ( $g(r)$ ) para um líquido genérico. $g(r) > 1$ , indicando a presença de interações efetivas e destacando as camadas de solvatação, indicadas pelos picos em $r_1$ e $r_2$ ; e em longas distâncias, maiores que 7 Å, correspondente à região <i>bulk</i> , onde a estrutura local do líquido torna-se menos pronunciada. . . . .	46
3.1	Sistema hipotético com duas partículas ilustrando o efeito de forças na velocidade das partículas ao longo do tempo. . . . .	48
3.2	Exemplificação do efeito de passos de tempo grandes, inadequados e pequenos, da ordem de femtossegundos, na variação da posição de duas partículas em uma simulação de DM. . . . .	49
3.3	Potencial de Lennard-Jones para dois átomos de argônio em interação. A curva preta representa o potencial total, enquanto as curvas pontilhadas destacam os termos atrativo (em azul) e repulsivo (em vermelho). . . . .	54
3.4	Representação esquemática de diferentes ensembles, são eles: A) canônico (NVT), B) microcanônico (NVE), C) isotérmico-isobárico (NPT) e D) $\mu VT$ (grande-canônico). A caixa <b>A</b> representa um sistema em que a energia pode fluir para a vizinhança, com o sistema contringido em número de partículas e volume. A caixa <b>B</b> representa um sistema constringido por uma barreira adiabática, resultando em número de partículas, volume e energia constantes. A caixa C representa um sistema em uma força externa ajustada ao tamanho dele para manter a pressão constante por meio da variação do volume, resultando na pressão constante. Por fim, a caixa <b>D</b> apresenta um sistema aberto, que permite a passagem de matéria e energia. . . . .	60
3.5	Pseudo-algoritmo simplificado para DM, ilustrando o fluxo principal de uma simulação. . . . .	67
5.1	Estrutura tridimensional da ubiquitina com seus elementos de estrutura secundária destacados em cores. Os segmentos na sequência primária (à direita) representam as regiões específicas das fitas $\beta$ (em vermelho) e das hélices $\alpha$ (em azul). As regiões de loop e outras partes não estruturadas estão marcadas em verde e cinza, respectivamente, na estrutura 3D. . . . .	74

5.2	Estruturas químicas dos cátions e ânions dos líquidos iônicos utilizados neste estudo: 1-etil-3-metilimidazólio ([EMIM] <sup>+</sup> ) e 1-butil-3-metilimidazólio ([BMIM] <sup>+</sup> ) como cátions; e diacianamida ([DCA] <sup>-</sup> ), tetrafluoroborato ([BF <sub>4</sub> ] <sup>-</sup> ), nitrato ([NO <sub>3</sub> ] <sup>-</sup> ) e cloreto ([Cl] <sup>-</sup> ) como ânions. Isso mostra a diversidade estrutural e funcional dos íons, essenciais para estudar o efeito da variação de cadeia carbônica do cátion e das propriedades específicas dos ânions na solvatação da ubiquitina. . . . .	75
5.3	A - Limitações do cálculo de $g(r)$ para solutos com forma não esférica. B — Conceito básico das funções de distribuição de mínima distância, que, ao considerar as mínimas distâncias $r$ , permite a obtenção de uma distribuição do solvente que mantém significado estrutural independentemente do formato das espécies. Figura reproduzida com autorização do autor a partir de [89]. . . . .	77
5.4	Representação das funções de distribuição radial e de mínima distância para o ânion [DCA] <sup>-</sup> 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 MDDF apresenta picos com formatos que possibilitam uma descrição mais intuitiva do sistema. Figura reproduzida com autorização do autor a partir de [89]. . . . .	78
5.5	Esquema simplificado de um experimento de diálise no equilíbrio. O fundo azul representa a água, os polígonos vermelhos são as moléculas de cossolvente, e as estruturas amarelas representam as biomoléculas. Figura reproduzida com autorização do autor a partir de [89]. . . . .	83
5.6	Representação do sistema descrevendo a adição de um ligante a uma solução aquosa com proteína. Em (A), temos um sistema inicial com água e proteína em solução, onde a proteína possui um determinado potencial químico ( $\mu_p^0$ ). Em (B), o ligante é introduzido e se distribui na solução, interagindo com a proteína. Esse processo perturba o potencial químico da proteína, causando acúmulo de ligante em sua superfície. Nesta situação, $\Gamma_{uc}$ é positivo, e o potencial químico da proteína é reduzido em relação ao sistema (A), conforme indicado na Equação 5.2.24. . . . .	88

6.1	<p>A) MDDFs (representadas por <math>g_{md}(r)</math>) e B) integrais de KB (<math>G_{cp}(r)</math>) do cátion <math>[\text{EMIM}]^+</math> e do ânion <math>[\text{DCA}]^-</math> relativos à ubiquitina em uma solução de aproximadamente <math>2,0 \text{ mol L}^{-1}</math> de <math>[\text{EMIM}][\text{DCA}]</math> em água. Esta figura ilustra como funções de distribuição marcadamente diferentes podem dar origem a integrais de KB idênticas, implicando que as funções de distribuição estão correlacionadas em grandes distâncias a partir da superfície do soluto.</p>	91
6.2	<p>Diferença nas densidades dos íons próximos de cada resíduo da proteína em <math>2,0 \text{ mol L}^{-1}</math>. As regiões em azul indicam maior densidade de <math>[\text{DCA}]^-</math>, enquanto as regiões em vermelho indicam maior densidade de <math>[\text{EMIM}]^+</math>. A acumulação de <math>[\text{DCA}]^-</math> na faixa de <math>2,5\text{--}3,0 \text{ \AA}</math> ocorre complementarmente à acumulação de <math>[\text{EMIM}]^+</math> entre <math>2,0\text{--}2,5 \text{ \AA}</math>, sugerindo que o segundo pico da distribuição de <math>[\text{DCA}]^-</math> na Figura 6.1 corresponde a uma camada de solvatação do ânion mediada pelo cátion.</p>	94
6.3	<p>Decomposição das MDDFs de cátions e ânions em contribuições de resíduos da ubiquitina em sistemas com <math>2,0 \text{ mol L}^{-1}</math>. A) Decomposição da MDDF do <math>[\text{EMIM}]^+</math> na presença de <math>[\text{DCA}]^-</math>, e B) do <math>[\text{EMIM}]^+</math> na presença de <math>[\text{BF}_4]^-</math>.</p>	95
6.4	<p>(A) Parâmetros de interação preferencial, <math>\Gamma_{cp}</math> dos LIs com a proteína nas concentrações de <math>0,5</math> e <math>3,0 \text{ mol L}^{-1}</math> de LI. (B) Parâmetros de hidratação preferencial, <math>\Gamma_{wp}</math>, nas mesmas concentrações de <math>0,5</math> e <math>3,0 \text{ mol L}^{-1}</math> de LI. Observa-se que todos os LIs possuem valores positivos de <math>\Gamma_{cp}</math> em baixas concentrações, enquanto <math>\Gamma_{wp}</math> é negativo, indicando uma desidratação preferencial em concentrações menores. No entanto, em concentrações mais altas, ocorre um aumento expressivo de <math>\Gamma_{wp}</math> e uma diminuição de <math>\Gamma_{cp}</math>, revelando o comportamento inverso em comparação com as concentrações mais baixas.</p>	97
7.1	<p>MDDFs entre ubiquitina e íons para <math>[\text{DCA}]^-</math>, <math>[\text{Cl}]^-</math>, <math>[\text{EMIM}]^+</math> e água na solução aquosa contendo <math>[\text{EMIM}][\text{DCA}]</math> e <math>[\text{EMIM}][\text{Cl}]</math>. As interações de ligação de hidrogênio correspondem a picos nas distribuições em aproximadamente <math>1,9 \text{ \AA}</math>, observadas para <math>[\text{DCA}]^-</math> e água. Interações eletrostáticas fortes, mais evidentes no pico do <math>[\text{Cl}]^-</math> em torno de <math>2,0 \text{ \AA}</math>, também são identificadas para os demais íons. Interações menos específicas se refletem em picos localizados a distâncias maiores.</p>	100

- 7.2 Decomposição das MDDFs de (A)  $[DCA]^-$ , (B)  $[BF_4]^-$ , (C)  $[Cl]^-$  e (D)  $[NO_3]^-$  nas contribuições dos tipos de resíduos de aminoácidos da proteína em sistemas com LI composto pelo cátion  $[EMIM]^+$  e um único tipo de ânion, a uma concentração de  $2,0 \text{ mol L}^{-1}$ . A MDDF total é a soma das contribuições de cada tipo de resíduo. . . . . 102
- 7.3 (A) Diferença nas densidades das MDDFs dos íons nas proximidades da proteína a aproximadamente  $2,0 \text{ mol L}^{-1}$  de  $[EMIM][DCA] + [EMIM][Cl]$ . Às regiões em azul-púrpura representam áreas onde a densidade de  $[Cl]^-$  é maior do que a de  $[DCA]^-$ , enquanto laranja indica regiões onde a densidade de  $[DCA]^-$  é maior do que a de  $[Cl]^-$ . (B) Perfil similar para a mistura  $[EMIM][DCA] + [EMIM][NO_3]$ , com laranja representando, de maneira semelhante, uma maior densidade de  $[DCA]^-$ . . . 104
- 7.4 (A) integrais de KB da água,  $[EMIM]^+$ ,  $[Cl]^-$  e  $[DCA]^-$  em relação à proteína em sistemas com  $2,0 \text{ mol L}^{-1}$  de mistura de LIs. As linhas sólidas representam os valores médios de 20 réplicas de simulação, enquanto as áreas sombreadas indicam o erro padrão da média. (B) Integrais KB dos íons ( $[EMIM]^+$  e  $[DCA]^- + [Cl]^-$  sobrepostos) em comparação com os integrais da água. Os maiores valores de integrais KB para os íons em relação à água indicam que a proteína é preferencialmente solvatada pelos íons dos líquidos iônicos. . . . . 105
- 7.5 (A) Integrais de KB para ânions tratados como uma única entidade e (B) integrais de KB da água em sistemas com  $2,0 \text{ mol L}^{-1}$  de misturas de LIs. A presença de  $[DCA]^-$  eleva significativamente as integrais de KB em comparação aos demais sistemas. A combinação de  $[DCA]^-$  com ânions de menor afinidade resulta nos maiores valores de integrais de KB, destacando o efeito de acúmulo. . . . . 106

- 7.6 Parâmetros de solvatação preferencial do LI em relação à proteína em todas as concentrações simuladas com diferentes composições aniônicas: em soluções de (A) 0,5, (B) 1,0, (C) 1,5, (D) 2,0, (E) 2,5 e (F) 3,0 mol L<sup>-1</sup>, fornecendo uma análise comparativa do comportamento do LI em diferentes concentrações do soluto. O gráfico destaca as diferentes tendências dos íons em solvatar preferencialmente a proteína em vez da água na presença de íons concorrentes. Devido à eletroneutralidade da solução, esses parâmetros são idênticos aos calculados para o cátion ou para o conjunto de ânions em cada solução. A notação da mistura de ânions aparece simplificada de modo que: DCA é [DCA]<sup>-</sup>, BF<sub>4</sub> é [BF<sub>4</sub>]<sup>-</sup>, NO<sub>3</sub> é [NO<sub>3</sub>]<sup>-</sup> e Cl é [Cl]<sup>-</sup> . . . . . 107
- 8.1 Comparação das MDDFs proteína-cátion para soluções de LIs contendo A) [EMIM][Cl] e [BMIM][Cl], B) [EMIM][DCA] e [BMIM][DCA], C) [EMIM][BF<sub>4</sub>] e [BMIM][BF<sub>4</sub>], e D) [EMIM][NO<sub>3</sub>] e [BMIM][NO<sub>3</sub>]. Cada painel compara as MDDFs dos cátions [EMIM]<sup>+</sup> (azul) e [BMIM]<sup>+</sup> (verde) com diferentes ânions. Os dados representam valores médios de 20 simulações independentes em uma concentração de referência de 1,0 mol L<sup>-1</sup>.111
- 8.2 Comparação das MDDFs proteína-ânion para soluções de LIs contendo A) [EMIM][Cl] e [BMIM][Cl], B) [EMIM][DCA] e [BMIM][DCA], C) [EMIM][BF<sub>4</sub>] e [BMIM][BF<sub>4</sub>], e D) [EMIM][NO<sub>3</sub>] e [BMIM][NO<sub>3</sub>]. Cada painel apresenta as MDDFs dos ânions quando emparelhados com [EMIM]<sup>+</sup> (azul) ou [BMIM]<sup>+</sup> (vermelho). Dados obtidos de 20 simulações independentes em concentração de 1,0 mol L<sup>-1</sup>. . . . . 112
- 8.3 MDDFs proteína-cátion de A) [EMIM]<sup>+</sup> e B) [BMIM]<sup>+</sup> em soluções com [DCA]<sup>-</sup> a 1,0 mol L<sup>-1</sup>, decompostos em contribuições de resíduos básicos, ácidos, neutros e polares. C) Mapa de densidade diferencial de [BMIM]<sup>+</sup> menos [EMIM]<sup>+</sup> em até 3,5 Å da superfície proteica. Vermelho indica maior densidade de [BMIM]<sup>+</sup>, e azul, de [EMIM]<sup>+</sup>. . . . . 113

8.4	Integrais de KB para ânions em relação à proteína em soluções aquosas de LIs com [EMIM] <sup>+</sup> ou [BMIM] <sup>+</sup> a 1,0 mol L <sup>-1</sup> . A) [Cl] <sup>-</sup> , B) [DCA] <sup>-</sup> , C) [BF <sub>4</sub> ] <sup>-</sup> , e D) [NO <sub>3</sub> ] <sup>-</sup> . As linhas sólidas representam a média das integrais KB de 20 simulações, enquanto as áreas sombreadas indicam o erro padrão da média. . . . .	115
8.5	Hidratação proteica em sistemas de líquidos iônicos contendo [EMIM][Cl], [EMIM][DCA], [BMIM][Cl] e [BMIM][DCA] a 1,0 mol L <sup>-1</sup> . A) MDDF proteína-água em soluções com cloreto e B) integrais de KB correspondentes. C) MDDF proteína-água em soluções com [DCA] <sup>-</sup> e D) integrais de KB correspondentes. . . . .	116
8.6	MDDFs proteína-cátion para soluções aquosas de [EMIM][DCA] em A) 1,0 mol L <sup>-1</sup> e B) 3,0 mol L <sup>-1</sup> . C) e D) Integrais de KB correspondentes às concentrações respectivas. Dados representam médias de 20 réplicas de simulação. . . . .	117
8.7	Parâmetros de solvatação preferencial ( $\Gamma_{cp}$ ) para diferentes composições e concentrações de LIs. LIs contendo [BMIM] <sup>+</sup> apresentam maior solvatação preferencial em concentrações baixas, enquanto em concentrações elevadas [EMIM] <sup>+</sup> predomina. Barras de erro representam o desvio padrão de 20 réplicas de simulação. . . . .	118
8.8	Números de coordenação de íons na primeira camada de solvatação em função da concentração de LI. A) e B) Números de coordenação de cátions em 3,5 Å em soluções com [DCA] <sup>-</sup> e [BF <sub>4</sub> ] <sup>-</sup> . C) e D) Números de coordenação de ânions em 2,5 Å em soluções com [DCA] <sup>-</sup> e [BF <sub>4</sub> ] <sup>-</sup> , respectivamente. Barras de erro representam o erro padrão da média. . . . .	119
8.9	Números de coordenação de cátions até 5 Å da superfície proteica em diferentes concentrações para sistemas com [EMIM][DCA] ou [BMIM][DCA]. Painéis A-F apresentam os dados para concentrações de referência de 0,5 a 3,0 mol L <sup>-1</sup> . . . . .	120
9.1	Modelos representativos das conformações A) nativa, B) perturbada, C) desnaturada e D) estendida. As estruturas progressivamente perdem características nativas, com aumento da exposição do núcleo hidrofóbico ao solvente. . . . .	124

- 9.2 (A) MDDFs e (B) integrais de KB para  $[DCA]^-$  em relação à ubiquitina em diferentes estados conformacionais em soluções de  $1,5 \text{ mol L}^{-1}$  de  $[EMIM][DCA]$ . (C) MDDFs e (D) integrais de KB para  $[EMIM]^+$  na mesma solução. Em (A), observa-se que a densidade local de  $[DCA]^-$  em distâncias características de ligações de hidrogênio diminui à medida que a proteína se desenovela, enquanto a densidade local em interações não específicas aumenta. Em (C), as MDDFs de  $[EMIM]^+$  exibem picos mais altos para os estados desnaturados; a curva MDDF para o estado estendido (roxa) sobrepõe-se quase completamente à do estado perturbado (laranja). As Figuras (B) e (D) apresentam as integrais de KB para  $[DCA]^-$  e  $[EMIM]^+$ , respectivamente, que aumentam com o desenovelamento da estrutura proteica. . . . . 127
- 9.3 Efeito do estado de enovelamento nas funções de distribuição do DCA, destacando as contribuições de cada classe de resíduos em uma solução com  $1,5 \text{ mol L}^{-1}$  de LI. (A), (B), (C) e (D) correspondem, respectivamente, às distribuições do DCA com conjuntos de estruturas nativa, perturbada, desnaturada e estendida. O pico ( $1,9 \text{ \AA}$ ) reflete maior acúmulo em resíduos polares e básicos, enquanto o segundo pico ( $2,6 \text{ \AA}$ ) está associado ao acúmulo em resíduos neutros. . . . . 128
- 9.4 Efeito do estado de enovelamento nas funções de distribuição do  $[EMIM]^+$ , destacando as contribuições de cada classe de resíduos em uma solução com  $1,5 \text{ mol L}^{-1}$  de LI. (A), (B), (C) e (D) correspondem, respectivamente, às distribuições do  $[EMIM]^+$  com conjuntos de estruturas nativa, perturbada, desnaturada e estendida. O pico ( $2,4 \text{ \AA}$ ) reflete maior acúmulo em resíduos neutros, proporcional ao desenovelamento da proteína. . . . . 129
- 9.5 Mapa de densidade dos íons ao redor dos resíduos proteicos. Regiões em laranja indicam maior densidade de  $[DCA]^-$  em relação ao  $[EMIM]^+$ , enquanto regiões em azul representam maior densidade de  $[EMIM]^+$  em comparação ao  $[DCA]^-$ . Resíduos polares e básicos estão associados ao pico de ligação de hidrogênio ( $1,9 \text{ \AA}$ ), enquanto resíduos hidrofóbicos e carregados negativamente predominam em regiões de maior concentração de  $[EMIM]^+$ . . . . . 130

9.6	MDDFs e integrais de KB para [EMIM] <sup>+</sup> em sistemas com estruturas nativas (A e C) e desnaturadas (B e D) com soluções de 1,5 mol L <sup>-1</sup> de [EMIM][DCA] (curva verde) e [EMIM][BF <sub>4</sub> ] (curva cinza). As MDDFs mostram que o [DCA] <sup>-</sup> induz um aumento maior na concentração de [EMIM] <sup>+</sup> nas proximidades da proteína em comparação ao [BF <sub>4</sub> ] <sup>-</sup> . . . . .	131
9.7	(A) MDDFs e (B) integrais de KB para água em relação à ubiquitina em diferentes estados conformacionais em soluções de 1,5 mol L <sup>-1</sup> de [EMIM][DCA]. Em (A), observa-se que o aumento da densidade local de água em distâncias características de ligações de hidrogênio diminui conforme a proteína se desenrola, o que está associado ao aumento da área superficial hidrofóbica. O efeito da maior área hidrofóbica é mais evidente em (B), onde a exclusão progressiva da água do domínio proteico é representada pelas integrais KB. . . . .	132
9.8	Coeficientes de solvatação preferencial ( $\Gamma_{cp}$ ) para os líquidos iônicos: (A) [EMIM][DCA], (B) [EMIM][BF <sub>4</sub> ], (C) [BMIM][DCA] e (D) [BMIM][BF <sub>4</sub> ]. Os conjuntos conformacionais da ubiquitina analisados incluem as conformações nativas, perturbada, desnaturada e estendida. As concentrações no eixo <i>x</i> correspondem às concentrações dos LIs no <i>bulk</i> , recalculadas após as simulações. . . . .	134
10.1	Diagrama de fases da mistura de dois sólidos A e B. Representação esquemática da diferença no ponto de fusão ( $\Delta T_f$ ) na composição eutética de uma mistura binária. Interações mais fortes entre os componentes A e B resultam em valores maiores de $\Delta T_f$ . $T_f(A)$ e $T_f(B)$ representam, respectivamente, as temperaturas de fusão para <b>A</b> e <b>B</b> . . . . .	140
10.2	Estruturas dos compostos que formam os solventes eutéticos utilizados para a simulação. . . . .	146
10.3	MDDFs dos componentes do DES BET:EG na proporção, respectivamente, de 1:3 em quantidade de matéria. As MDDFs são mostradas até 8 Å para melhor visualização dos picos importantes. A curva laranja (BET) exibe dois picos principais (1,8 Å e 2,4 Å), enquanto a curva azul (EG) apresenta um pico em 2,5 Å. . . . .	149

10.4	MDDFs dos componentes dos DESs: A) BET:EG, B) BET:PRP e C) BET:BTO, na proporção de 1:3 em quantidade de matéria, respectivamente. D) Comparação das MDDFs das espécies doadoras de ligações de hidrogênio. . . . .	150
10.5	Decomposição dos MDDFs para os componentes do DES BET:EG. (A) Contribuição dos átomos do EG: grupos CH <sub>2</sub> estão mais próximos do ibuprofeno. (B) Contribuição dos átomos da betaína: o pico em 1,8 Å é formado pelo grupo COO <sup>-</sup> , enquanto o pico em 2,4 Å envolve predominantemente o grupo N(CH <sub>3</sub> ) <sub>3</sub> <sup>+</sup> . . . . .	151
10.6	MDDFs decompostas em termos das partes do ibuprofeno. (A) Decomposição da MDDF do EG em termos das partes do ibuprofeno. Os sítios prováveis de interação com o EG são aqueles com características mais hidrofóbicas. (B) O primeiro pico da MDDF de BET apresenta contribuições dos átomos de betaína ao redor do grupo COOH do ibuprofeno. O segundo pico tem contribuições do grupo COOH e contribuições significativas do grupo isopropila (cinza), enquanto os grupos benzeno e alquila (verde) contribuem igualmente para a MDDF total. . . . .	152
10.7	Decomposição das MDDFs em termos de átomos do solvente para o sistema BET-EG. A curva vermelha representa a distribuição dos átomos de carbono ligados ao oxigênio na betaína, indicando interações de ligação de hidrogênio com o ibuprofeno. A curva azul representa os átomos de nitrogênio na betaína, também indicando interações de ligação de hidrogênio com o ibuprofeno. O pequeno pico em 6 Å na distribuição de C e o pico em 8,5 Å na distribuição de N sugerem interações entre os grupos COO <sup>-</sup> e N(CH <sub>3</sub> ) <sub>3</sub> <sup>+</sup> . . . . .	153
10.8	Integrais de KB para BET (curva laranja) e EG (curva azul). A queda inicial na curva laranja indica repulsão estérica, enquanto o aumento subsequente é devido às ligações de hidrogênio. A curva azul representa o comportamento do EG. . . . .	154

10.9 Parâmetros de solvatação preferencial ( $\Gamma$ ) para os sistemas BET-EG, BET-PG e BET-BTO. A barra laranja representa a solvatação preferencial da betaína (BET), enquanto a barra azul refere-se à solvatação pelos doadores de ligação de hidrogênio (HBDs). Valores positivos de  $\Gamma$  indicam acumulação preferencial do componente ao redor do ibuprofeno, enquanto valores negativos indicam exclusão preferencial. Observa-se que o sistema BET-EG apresenta a maior solvatação preferencial por BET, enquanto o aumento do comprimento da cadeia alquila do HBD resulta em mudanças no comportamento de solvatação preferencial, especialmente no caso do BTO, onde a solvatação do HBD torna-se dominante. . . . . 156

# Lista de Tabelas

- 6.1 Concentrações de referência e efetivas das soluções de LI simuladas, juntamente com as integrais de KB para todos os componentes do solvente em relação à proteína. Os erros padrão da média foram calculados a partir de pelo menos 20 simulações independentes. As diferenças entre as concentrações de referência e efetivas devem-se ao fato de que a concentração de referência representa o valor pretendido durante a montagem inicial da caixa de simulação, enquanto a concentração efetiva corresponde à concentração real do LI no *bulk* após a equilibração do sistema. 93
- 7.1 Número de coordenação de [EMIM]<sup>+</sup> e dos ânions ao redor da proteína, contando o número de íons até 5 Å da superfície da proteína, em sistemas com 2,0 mol L<sup>-1</sup> de mistura de LI. Os sistemas possuem 1004 cátions e 502 ânions de cada tipo nesta concentração. As flutuações relatadas são o erro padrão da média das 20 réplicas realizadas para cada sistema . . . . . 103
- 8.1 Concentrações em solução (mol L<sup>-1</sup>) das espécies químicas em sistemas contendo [EMIM][DCA] e [BMIM][DCA] preparados com concentrações de referência ( $C_{IL}$ ) de 0,5 e 3,0 mol L<sup>-1</sup>. Os subscritos *E* e *B* indicam, respectivamente, que a espécie em destaque está num sistema com [EMIM]<sup>+</sup> ou [BMIM]<sup>+</sup>. . . . . 121
- 9.1 Propriedades estruturais médias e desvios padrão das conformações utilizadas: área de superfície acessível ao solvente (SASA), raio de giro, RMSD em relação ao modelo cristalográfico e fração de contatos nativos (NC). . . . . 123

9.2	Número de ligações de hidrogênio (HB, do inglês <i>hydrogen bonds</i> ) entre a proteína e a água, proteína-proteína (intramolecular) e proteína-ânion em sistemas com $1,5 \text{ mol L}^{-1}$ de [EMIM][DCA]. . . . .	133
9.3	Valores de $\Delta\Gamma_{cp}$ para diferentes líquidos iônicos em concentrações de 0,5 e $3,0 \text{ mol L}^{-1}$ . . . . .	136
10.1	Tamanhos da caixa, densidade experimental, densidade calculada após a produção em NPT, e número de componentes usados para criar as caixas de simulação. As densidades e proporções foram referenciadas do trabalho experimental disponível em [229]. . . . .	147

# Lista de Abreviaturas

MDDF	Em inglês, <i>Minimum-Distance Distribution Function</i> , ou Funções de Distribuição de Mínima Distância.
KB	Kirkwood-Buff.
TKB	Teoria de Kirkwood-Buff.
RDF	Em inglês, <i>Radial Distribution Function</i> , ou Função de Distribuição Radial.
LI	Líquido Iônico.
DES	Em inglês, <i>Deep Eutectic Solvent</i> , ou Solvente Eutético Profundo.
TMAO	O N-óxido de trimetilamina, é um composto orgânico com fórmula química $(\text{CH}_3)_3\text{NO}$ .
$[\text{DCA}]^-$	Dicianamida.
$[\text{BF}_4]^-$	Tetrafluorborato.
$[\text{Cl}]^-$	Cloreto.
$[\text{NO}_3]^-$	Nitrato.
$[\text{EMIM}]^+$	1-etil-3-metilimidazol.
$[\text{BMIM}]^+$	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> .
VMD	Em inglês, <i>Visual Molecular Dynamics</i> , software usado para análise e visualização de sistemas moleculares.
HBD	Em inglês, <i>Hydrogen Bond Donor</i> , ou Doador de Ligações de Hidrogênio.
HBA	Em inglês, <i>Hydrogen Bond Acceptor</i> , ou Aceptor de Ligações de Hidrogênio.
<i>NPT</i>	Ensemble termodinâmico isotérmico-isobárico, com Número de partículas ( $N$ ), Pressão ( $P$ ) e Temperatura ( $T$ ) constantes, respectivamente.
<i>NVE</i>	Ensemble termodinâmico micro-canônico, com Número de partículas ( $N$ ), Volume ( $V$ ) e Energia ( $E$ ) constantes, respectivamente.
<i>NVT</i>	Ensemble termodinâmico canônico, com Número de partículas ( $N$ ), Volume ( $V$ ) e Temperatura ( $T$ ) constantes, respectivamente.
$\mu VT$	Ensemble termodinâmico grande-canônico, com potencial químico ( $\mu$ ), volume ( $V$ ) e temperatura ( $T$ ) constantes, respectivamente.

# Sumário

<b>1</b>	<b>Líquidos iônicos e proteínas</b>	<b>28</b>
1.1	Visão Geral sobre Líquidos Iônicos . . . . .	28
1.2	Principais propriedades dos LIs e aplicações mais comuns . . . . .	31
1.3	Líquidos iônicos e proteínas . . . . .	33
1.3.1	Aplicações mais modernas de LIs com proteínas . . . . .	37
<b>2</b>	<b>Soluções com proteínas</b>	<b>39</b>
2.1	Aspectos gerais da solvatação . . . . .	39
2.1.1	Solvatação de biomoléculas . . . . .	41
2.2	Estrutura de solvatação . . . . .	43
<b>3</b>	<b>Simulações de Dinâmica Molecular</b>	<b>47</b>
3.1	Visão Geral do Método . . . . .	47
3.2	Cálculo das forças . . . . .	51
3.3	Etapas básicas de uma simulação de dinâmica molecular . . . . .	57
3.3.1	Montagem da caixa de simulação . . . . .	57
3.3.2	Movimentação das partículas . . . . .	57
3.3.3	Ensembles estatísticos . . . . .	59
3.3.4	Controle de temperatura . . . . .	60
3.3.5	Controle da pressão . . . . .	63
3.3.6	Condições iniciais e etapas da simulação . . . . .	64
3.3.7	Fluxograma simplificado de uma simulação de DM . . . . .	66
<b>4</b>	<b>Objetivos</b>	<b>70</b>
4.1	Objetivos . . . . .	70

4.2	Trabalhos publicados ou em publicação/preparação. . . . .	71
<b>5</b>	<b>Metodologia Geral</b>	<b>73</b>
5.1	Sistemas simulados . . . . .	73
5.1.1	Breve discussão sobre a ubiquitina . . . . .	73
5.1.2	Composição dos sistemas estudados . . . . .	75
5.2	Análise de solvatação . . . . .	76
5.2.1	Funções de distribuição . . . . .	76
5.2.2	Teoria de Soluções de Kirkwood-Buff . . . . .	80
5.2.3	Solvatação Preferencial . . . . .	82
5.2.4	Montagem dos sistemas e detalhes técnicos das simulações de dinâmica molecular . . . . .	88
<b>6</b>	<b>Solvatação de proteínas por LIs: visão geral</b>	<b>90</b>
6.1	Efeitos correlacionados na solvatação . . . . .	90
6.1.1	Natureza da interação entre solvente e ubiquitina. . . . .	94
<b>7</b>	<b>Efeitos não-aditivos</b>	<b>99</b>
7.1	Solvatação da ubiquitina por múltiplos íons . . . . .	99
<b>8</b>	<b>Efeitos da hidrofobicidade do cátion</b>	<b>110</b>
8.0.1	Hidrofobicidade dos cátions e estrutura local da solução . . .	110
8.0.2	Competição com água e solvatação preferencial . . . . .	114
8.0.3	Dependência da concentração na solvatação preferencial . . .	116
8.0.4	Número de coordenação e conteúdo de água . . . . .	118
<b>9</b>	<b>Solvatação de diferentes estados de enovelamento</b>	<b>122</b>
9.1	Detalhes específicos da metodologia . . . . .	122
9.1.1	Estados de enovelamento estudados . . . . .	122
9.1.2	Configurações da dinâmica molecular dos diferentes estados .	124
9.2	Estrutura de solvatação . . . . .	125
9.3	Solvatação por água em soluções de líquidos iônicos . . . . .	132
9.4	Solvatação preferencial em diferentes estados . . . . .	134
<b>10</b>	<b>Solvatação em DES</b>	<b>137</b>

10.1	Visão geral sobre DESs . . . . .	138
10.2	Campos de força polarizáveis e simulações de DESs . . . . .	140
10.3	Aspectos gerais do campo de força CL&Pol . . . . .	143
10.4	Metodologia das simulações de DM para os DESs . . . . .	144
10.4.1	Sistemas estudados . . . . .	146
10.5	Solvatação do ibuprofeno em DESs Baseados em Betaína . . . . .	148
10.6	Integrais de KB . . . . .	153
10.7	Solvatação Preferencial do Ibuprofeno em DESs . . . . .	155
<b>11</b>	<b>Considerações Finais</b>	<b>157</b>
11.1	Interações específicas e acúmulo de íons . . . . .	157
11.2	Interações competitivas e efeitos não-aditivos . . . . .	158
11.3	Acúmulo preferencial de LIs em estados desenovelados . . . . .	158
11.4	Efeitos da hidrofobicidade do cátion . . . . .	159
11.5	Solvatação do ibuprofeno em DESs . . . . .	160
	<b>Bibliografia</b>	<b>162</b>

## Capítulo 1

# Líquidos iônicos e proteínas

Os líquidos iônicos (LIs) formam uma classe de compostos que têm sido amplamente utilizados em diversas áreas devido às suas propriedades variadas e customizáveis como solventes. Este capítulo apresenta uma introdução aos LIs e às suas aplicações em sistemas envolvendo proteínas. Inicialmente, serão apresentados os aspectos gerais da natureza química dos LIs, destacando suas principais propriedades e como estas podem ser correlacionadas com suas estruturas químicas. Uma atenção especial será dedicada às aplicações dos LIs, com ênfase nos compostos imidazólicos, em sistemas biológicos. Será abordada, em particular, sua utilização como cossolventes em sistemas contendo proteínas, evidenciando os avanços e desafios nesse campo de pesquisa.

### 1.1 Visão Geral sobre Líquidos Iônicos

Os LIs são sais que permanecem no estado líquido em temperaturas inferiores a 100°C [1, 2]. São geralmente formados por cátions orgânicos volumosos, como tetraalquilamônios ou compostos imidazólicos, combinados com ânions de estruturas variadas: desde íons simples, como cloreto, até carboxilatos com cadeias químicas mais complexas [3]. A Figura 1.1 ilustra exemplos de cátions e ânions comumente encontrados em LIs.

Os íons presentes nos LIs, especialmente os cátions, possuem características específicas, como o grande tamanho molecular e a carga frequentemente deslocalizada [4]. A deslocalização de carga, juntamente com a assimetria dos íons reduz significativamente a intensidade das interações eletrostáticas, em comparação com cargas localizadas em íons menores, como no caso do NaCl. Consequentemente, o enfraquecimento das ligações iônicas

diminui a estabilidade do retículo cristalino, permitindo que esses compostos existam no estado líquido em temperatura ambiente [5].

Os ânions, por sua vez, apresentam maior diversidade estrutural e química. Enquanto os cátions frequentemente possuem cadeias carbônicas como característica predominante, os ânions podem incluir grupos químicos com propriedades variadas, como doadores e aceptores de ligações de hidrogênio, além de grupos polares e apolares, conforme ilustrado na Figura 1.1(B) [6, 7].

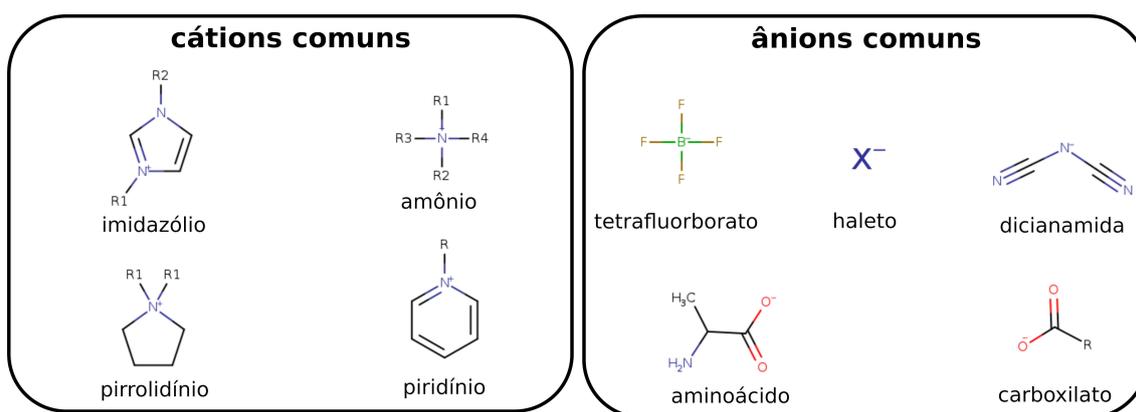


Figura 1.1: Estruturas químicas dos íons mais encontrados em LIs. Os cátions, à direita, são, em sua maioria, compostos orgânicos, frequentemente apresentando cadeias carbônicas extensas. Em contraste, os ânions, à esquerda, podem ser tanto de natureza inorgânica quanto orgânica, variando amplamente em tamanho e na presença de diferentes grupos funcionais.

A configuração assimétrica e variada dos íons nos LIs reduz significativamente a eficácia das interações coulômbicas em comparação com compostos iônicos convencionais, como o NaCl. Esta redução na intensidade das interações eletrostáticas resulta em um ponto de fusão consideravelmente menor para os LIs, o que permite que muitos deles sejam líquidos à temperatura ambiente (25 °C) [8, 9]. A Figura 1.2 ilustra simplificada as diferenças estruturais entre um sal convencional, como o NaCl, e um LI, caracterizado por um cátion assimétrico e significativamente maior que o ânion.

A energia potencial entre duas cargas é inversamente proporcional à distância que as separa. Nos LIs, os cátions orgânicos são substancialmente maiores que íons como o  $\text{Na}^+$  e, em muitos casos, apresentam cargas deslocalizadas. Essa característica reduz ainda mais a intensidade das interações eletrostáticas. Apesar do enfraquecimento dessas interações, elas continuam sendo relevantes nos LIs, contribuindo para suas propriedades únicas. Tais propriedades são atípicas se comparadas com a maioria dos solventes orgânicos, incluindo

alta estabilidade térmica, baixa pressão de vapor, elevada capacidade de solubilizar diferentes compostos, alta polaridade, baixa toxicidade e alta viscosidade [10]. Estas características fazem dos LIs um tema de grande interesse tanto em aplicações industriais quanto em pesquisas acadêmicas [11–13].

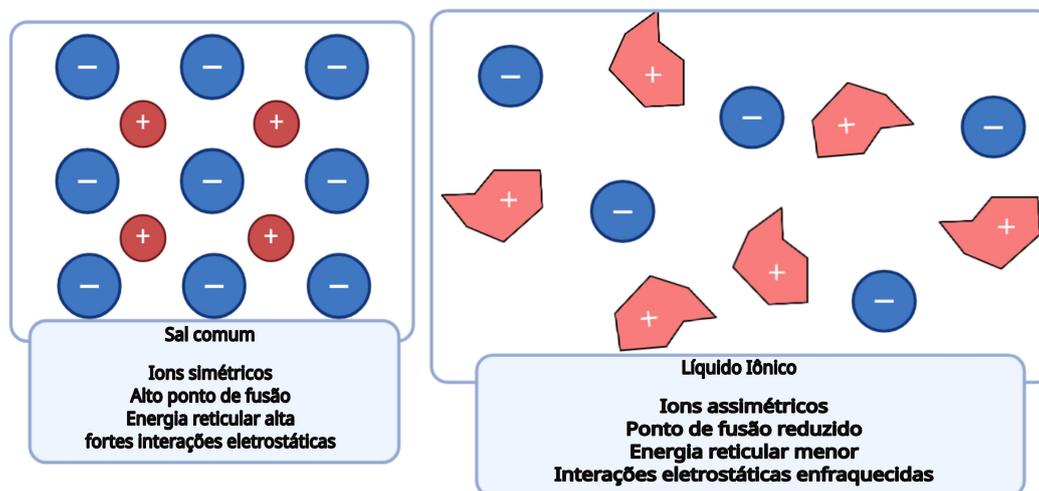


Figura 1.2: Modelo comparativo entre íons regulares e líquidos iônicos, destacando as diferenças estruturais e de simetria.

LIs interagem com outras substâncias de acordo com a natureza de seus íons e das propriedades das substâncias envolvidas. Por exemplo, um LI composto por íons hidrofóbicos será naturalmente excluído de uma fase aquosa [3]. No caso do nitrato de 1-butil-3-metilimidazol ( $[\text{BMIM}][\text{NO}_3]$ ), o cátion é um composto orgânico volumoso com carga deslocalizada no anel imidazólico, além de apresentar baixa polaridade. O grupo imidazólico pode interagir com biomoléculas, ou outros solutos, por meio de interações não específicas, como forças de van der Waals ou íon-dipolo [14]. O ânion  $[\text{NO}_3]^-$ , por sua vez, é altamente polar e pode participar de interações específicas, como ligações de hidrogênio (devido à presença de átomos de oxigênio), bem como interações não específicas, como forças dipolo-dipolo ou íon-dipolo [3].

Além do enfraquecimento geral das interações entre os íons, um aspecto importante que contribui para o baixo ponto de fusão dos LIs é o fator entrópico. O cátion imidazólico, em particular, costuma apresentar dois grupos alquil ligados a cada nitrogênio do anel, aumentando a flexibilidade molecular. Estas cadeias alquílicas conferem mais graus de liberdade para o sistema [3]. Quando o LI passa do estado sólido para o líquido, há um ganho entrópico significativo devido ao aumento do movimento, especialmente dos cátions, que são

maiores e possuem cadeias laterais volumosas [14].

As cadeias alquílicas presentes nos cátions imidazólicos também introduzem um caráter hidrofóbico significativo ao sistema. À medida que o comprimento das cadeias aumenta, cresce a tendência de segregação espacial entre as regiões polares (compostas por íons e grupos funcionais carregados) e as regiões apolares (formadas pelas porções alquiladas dos cátions). Tal comportamento está relacionado à baixa miscibilidade de certos líquidos iônicos com solventes polares, como a água, e à afinidade relativamente maior por espécies orgânicas apolares. O aumento da hidrofobicidade também influencia propriedades macroscópicas como a tensão superficial, a viscosidade e o coeficiente de partição octanol/água, sendo um parâmetro crítico no desenho racional de LIs para aplicações específicas.

Devido à heterogeneidade estrutural, muitos LIs podem ser descritos como fluidos auto-organizados, nos quais coexistem domínios hidrofílicos e hidrofóbicos em escala nanométrica [15]. As regiões hidrofílicas, compostas predominantemente pelos centros carregados dos cátions e ânions, formam redes fortemente correlacionadas por interações coulômbicas e pontes de hidrogênio. Por outro lado, as porções alquiladas dos cátions tendem a se aglomerar, gerando domínios hidrofóbicos segregados, similares a microambientes orgânicos apolares [16]. Essa heterogeneidade microscópica confere aos líquidos iônicos propriedades únicas de solvatação, permitindo, por exemplo, a dissolução seletiva de solutos hidrofílicos e hidrofóbicos em regiões distintas do líquido. Esta organização supramolecular tem implicações profundas em processos como catálise, extração e estabilização de biomoléculas.

## 1.2 Principais propriedades dos LIs e aplicações mais comuns

Propriedades físico-químicas, como estabilidade térmica, capacidade calorífica, viscosidade e temperaturas da fase líquida, são aspectos cruciais para compostos utilizados como solventes. Os LIs apresentam propriedades físico-químicas incomuns quando comparados a solventes orgânicos convencionais, como o clorofórmio. Entre essas características destacam-se: baixo ponto de fusão, baixa pressão de vapor, alta polaridade, elevada estabilidade química e térmica, e baixa toxicidade. Essas propriedades tornam os LIs altamente promissores para uma ampla gama de aplicações [17–20].

LIs têm demonstrado grande potencial em áreas como síntese orgânica, catálise,

eletroquímica, separação de metais, separação de gases, processamento de biomassa, desenvolvimento de produtos farmacêuticos e dispositivos de armazenamento de energia, incluindo baterias, supercapacitores e células de combustível [20–25]. Apesar da diversidade de aplicações, os sistemas biológicos, especialmente aqueles envolvendo proteínas, têm atraído atenção crescente. Isso se deve às inúmeras potencialidades apresentadas pelos Lis em relação às interações que podem ser estabelecidas e às propriedades físico-químicas que podem ser moduladas [6, 26–30].

Devido à natureza orgânica de seus íons, especialmente os cátions, os Lis apresentam a combinação de funções orgânicas e cargas que permite a formação de uma ampla gama de interações com outros compostos, diferenciando-os dos sais convencionais, onde predominam interações coulômbicas [31]. Além disso, essas propriedades podem ser ajustadas para atender a demandas específicas por meio da seleção de diferentes cátions e ânions [17, 18]. A ampla aplicabilidade dos Lis destaca a necessidade de uma compreensão mais profunda de suas características, que permita racionalizar e prever seus comportamentos. A possibilidade de ajustar suas propriedades reforça a relevância de métodos preditivos para determinar características termo-físicas e motiva esforços para explorar detalhes moleculares relacionados à estrutura, dinâmica e interações nos Lis [26–30].

Diversas revisões sobre a relação entre estrutura e propriedades estão disponíveis, abordando aspectos fundamentais, como propriedades térmicas e físico-químicas dos Lis puros [32–34]. As metodologias preditivas empregadas no estudo de Lis geralmente adotam abordagens empíricas ou teóricas para representar propriedades como densidade, tensão superficial, capacidade térmica, pontos de fusão e condutividade elétrica [35, 36]. As abordagens empíricas buscam correlacionar as propriedades dos Lis com suas estruturas moleculares [18]. Dentre as propriedades de transporte, a viscosidade destaca-se como uma das mais estudadas, devido à sua importância em aplicações como eletroquímica e separação. Os valores de viscosidade dos Lis variam amplamente, geralmente entre 20 e 40.000 cP, ordens de magnitude maiores que a viscosidade de solventes convencionais, como a água, que é, por definição, 1 cP.

Um comportamento comum observado nos Lis é a redução da viscosidade com o aumento do tamanho dos íons, sendo esse efeito mais pronunciado no caso dos ânions. Para os cátions, a viscosidade tende a aumentar com o comprimento das cadeias alquílicas,

devido, em parte, ao fortalecimento das interações de van der Waals (VDW) [18]. Além disso, a viscosidade é influenciada pela simetria dos cátions, com cátions assimétricos apresentando comportamentos distintos em relação a cátions menores e mais similares aos ânions em tamanho. Em resumo, a viscosidade está diretamente relacionada às interações intermoleculares. Interações fortes, como forças eletrostáticas e de VDW entre ânions e cátions, resultam em viscosidades mais elevadas.

A condutividade elétrica, outra propriedade frequentemente investigada, apresenta valores na faixa de 0,1 a 30 mS/cm. Em temperaturas mais baixas, as estruturas dos cátions e ânions exercem uma influência significativa sobre a condutividade dos LIs. Geralmente, a condutividade diminui com o aumento do comprimento das cadeias alquílicas dos cátions. O efeito dos ânions também é considerável, especialmente devido à distribuição de carga e ao número de portadores de carga [37].

O grupo de pesquisa liderado por Masayoshi Watanabe, da Universidade Nacional de Yokohama, contribuiu significativamente para o campo com estudos que correlacionam estrutura e propriedades de LIs [38, 39]. Em uma série de trabalhos pioneiros, investigaram as relações entre difusividade iônica, viscosidade e condutividade molar para diferentes estruturas catiônicas e aniônicas. Em um estudo publicado em 2004 [38], os pesquisadores analisaram o cátion 1-butil-3-metilimidazólio ( $[BMIM]^+$ ) combinado com diversos ânions fluorados, incluindo  $[NTF_2]^-$ ,  $[OTF]^-$ ,  $[PF_6]^-$ ,  $[CF_3CO_2]^-$  e  $[BF_4]^-$ . Por meio de experimentos de auto diluição e condutividade, foi observado que a contribuição dos íons para a condutividade iônica está relacionada à interação fraca entre os ânions e o cátion.

### 1.3 Líquidos iônicos e proteínas

As aplicações mais comuns de LIs em sistemas envolvendo proteínas incluem a dissolução, estabilização, extração e purificação de proteínas, tanto usando LIs anidros quanto em misturas com água [40–43]. A capacidade dos LIs de solubilizar e manter a integridade das proteínas está diretamente relacionada às suas propriedades físico-químicas, como viscosidade, polaridade, habilidade de doação ou aceitação de ligações de hidrogênio e concentração [44, 45]. Adicionalmente, fatores inerentes às proteínas, como estrutura, peso molecular, ponto isoelétrico e conformação, também desempenham papéis importantes na solubilidade e estabilidade em diferentes LIs. Proteínas possuem uma grande complexidade

estrutural, com diferentes grupos químicos expostos ao solvente, enquanto os LIs, formados por íons de estruturas diversificadas, interagem com as proteínas via múltiplos mecanismos moleculares.

Em solução, proteínas existem em equilíbrio conformacional entre estados nativos, desenovelados e parcialmente enovelados. Moléculas individuais podem estocasticamente transitar entre esses estados. Contudo, em condições de equilíbrio, a distribuição total entre os estados permanece constante. Sob condições fisiológicas e ambientais, conformações mais enoveladas são favorecidas. No entanto, alterações no estado termodinâmico do sistema, como mudanças de temperatura, pressão, pH ou composição do solvente, podem perturbar o equilíbrio entre os estados enovelados e desenovelados [46, 47]. A diferença de energia livre de Gibbs entre esses estados varia entre -5 e -15 kcal mol<sup>-1</sup>, com estados enovelados sendo energeticamente preferidos. Pequenas perturbações, no entanto, podem alterar esse equilíbrio [48].

O processo de enovelamento e desenovelamento de proteínas pode ser modulado por osmólitos, que são pequenas moléculas abundantes em organismos vivos. Osmólitos contrabalanceiam o estresse causado por fatores externos por meio de interações não específicas e geralmente fracas [49, 50]. Em muitos casos, essas condições de estresse comprometem a estabilidade conformacional das proteínas [51, 52].

A estabilidade de proteínas é fortemente influenciada pela composição da solução em que estão inseridas. Cossolventes como glicose e alguns polióis possuem a capacidade de estabilizar proteínas, favorecendo conformações enoveladas. Por outro lado, substâncias como a ureia tendem a desestabilizar proteínas, promovendo a desnaturação [53]. Nesse contexto, LIs têm emergido como cossolventes promissores em sistemas biotecnológicos [26, 28, 29]. É relevante observar que, em sistemas biológicos, os LIs perdem muitas das características físico-químicas observadas em estado puro, tornando-se sais dissolvidos em meios aquosos. Diferentemente de cossolventes tradicionais, como ureia, TMAO e polióis, os LIs são formados por duas espécies químicas iônicas. Essa natureza dual permite que os LIs modulem processos químicos de maneira distinta dos cossolventes convencionais [54].

Os íons interagem com o soluto de maneira dependente da natureza química de ambos. A água, sendo altamente polar, estabelece interações fortes com outras moléculas polares e com aquelas capazes de formar ligações de hidrogênio. Assim, há uma competição

constante entre os íons e a água pelas interações com as superfícies proteicas acessíveis ao solvente. O problema do desenovelamento (ou sua ausência) de proteínas em solução pode ser compreendido considerando as interações preferenciais de cada solvente com diferentes conformações proteicas.

Em organismos vivos, a conformação nativa de uma proteína é geralmente a mais estável [55–57]. No entanto, em soluções capazes de estabilizar outras conformações mais eficientemente, a proteína pode adotar estados diferentes do nativo, resultando na desnaturação. Além disso, como discutido anteriormente, as proteínas são marginalmente estáveis, tornando o desenovelamento um processo provável.

Devido às suas características hidrofóbicas, hidrofílicas e anfífilas, a água e os íons dos LIs interagem de maneiras distintas com as regiões polares, não polares e carregadas das superfícies proteicas. Essas interações podem causar acúmulo ou depleção em áreas específicas da superfície, como ilustrado na Figura 1.3.

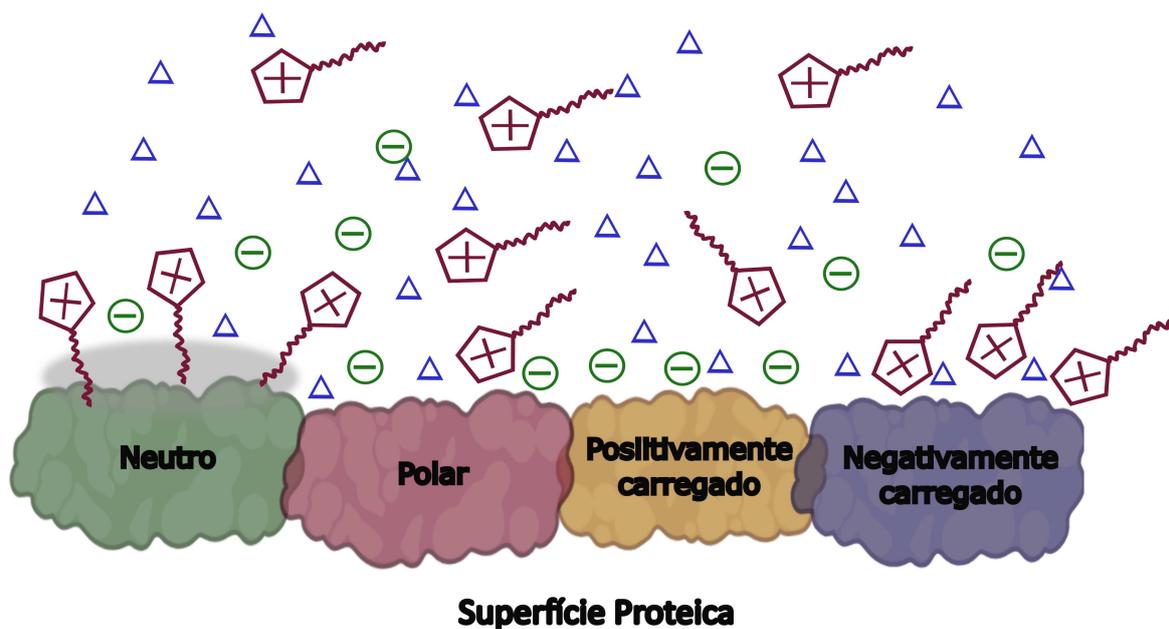


Figura 1.3: Modelo esquemático ilustrando a variabilidade das interações entre LIs e proteínas. As interações podem incluir forças dispersivas em regiões apolares (representadas em cinza), ligações de hidrogênio com moléculas de água (triângulos azuis) e interações eletrostáticas em regiões carregadas, demonstrando a complexidade das interações preferenciais em diferentes superfícies proteicas. Figura inspirada a partir da discussão sobre interações de LIs com proteínas disponível em [58].

Um aspecto relevante na solvatação de proteínas por LIs é o caráter anfílico de muitos cátions. As longas cadeias alquílicas associadas ao centro de carga proporcionam

regiões apolares que interagem por forças dispersivas (representadas em cinza na Figura 1.3) [59, 60]. Essas interações são particularmente importantes em proteínas contendo aminoácidos apolares, que encontram nos cátions parceiros favoráveis para interações hidrofóbicas. Estudos conduzidos por Klahn e colaboradores demonstraram que as cadeias alquílicas dos cátions tendem a se orientar em direção às superfícies não polares das proteínas [61].

Além disso, os cátions, por possuírem carga, podem competir com os ânions pelos sítios de solvatação nas regiões polares das proteínas, estabelecendo interações com resíduos de carga oposta. Em muitos casos, no entanto, as cadeias alquílicas dos cátions se orientam para longe da superfície proteica, especialmente em áreas hidrofílicas [61]. Aminoácidos carregados negativamente atraem preferencialmente cátions, embora também interajam com moléculas de água (representadas por triângulos azuis na Figura 1.3), capazes de formar ligações de hidrogênio com os resíduos superficiais.

Por outro lado, a aproximação de cátions às regiões carregadas positivamente das proteínas não pode ser completamente descartada. Devido à baixa densidade de carga dos cátions e à presença de ânions próximos na superfície proteica, interações favoráveis podem ocorrer, mesmo nessas áreas [29].

Os ânions hidrofílicos tendem a se concentrar nas superfícies positivamente carregadas das proteínas, buscando estabelecer ligações de hidrogênio com resíduos como lisina, arginina ou histidina (quando protonada) [62]. Além disso, esses ânions exibem fortes interações coulômbicas com esses aminoácidos devido à sua alta densidade de carga. De maneira geral, as interações eletrostáticas entre ânions e proteínas são mais intensas em comparação com as interações envolvendo cátions, resultando em tempos de permanência mais longos para os ânions nas proximidades dos resíduos proteicos correspondentes [63, 64].

A remoção de moléculas de água pode impactar positiva ou negativamente a solubilidade das proteínas [65]. Por exemplo, líquidos iônicos hidrofóbicos podem formar sistemas bifásicos com a água, reduzindo a remoção de moléculas de água da superfície proteica. Isso, por sua vez, pode aumentar a solubilidade da proteína [59]. Líquidos iônicos com alta capacidade de formar ligações de hidrogênio ou interações não específicas com proteínas frequentemente exercem efeitos desestabilizantes na estrutura proteica [66, 67]. Por outro lado, ânions como  $[\text{BF}_4]^-$ , que possuem menor capacidade de formar ligações de hidrogênio e baixa nucleofilicidade, demonstram compatibilidade com enzimas em

determinadas situações [68].

### 1.3.1 Aplicações mais modernas de LIs com proteínas

Os LIs têm demonstrado grande relevância como cossolventes em sistemas envolvendo proteínas, devido a dois fatores principais: 1) Os LIs exercem efeitos significativos sobre a estrutura de proteínas mesmo em concentrações menores do que as usualmente requeridas por outros cossolventes [11, 54]. 2) A composição diversificada dos LIs permite a síntese de compostos capazes de interagir seletivamente com sítios específicos das proteínas. Por esse motivo, compreender a distribuição dos íons em solução e seu comportamento em relação às proteínas é essencial [27, 28].

LIs têm demonstrado potencial na preservação da funcionalidade de biomoléculas durante o armazenamento e em aplicações práticas, sendo cruciais para o desenvolvimento de medicamentos e outros produtos biológicos [6, 69]. A capacidade dos LIs de melhorar a estabilidade e a solubilidade de proteínas terapêuticas é particularmente relevante, abrindo caminho para novas formulações de medicamentos. Esses avanços têm impacto direto no desenvolvimento de sistemas de entrega mais eficientes e na eficácia clínica de proteínas destinadas ao tratamento de doenças como a diabetes e o câncer [70, 71].

LIs também favorecem a catálise enzimática em reações orgânicas, aumentando taxas de reação e seletividade [70]. Esses efeitos são fundamentais para o avanço de técnicas bioquímicas e para o desenvolvimento de proteínas terapêuticas mais estáveis e funcionais. Além disso, os LIs vêm sendo explorados no armazenamento de proteínas, promovendo a preservação de sua bioatividade e prolongando sua vida útil, representando uma abordagem inovadora para armazenamento e transporte de biomoléculas sensíveis [72, 73].

Em aplicações estruturais, os LIs têm sido empregados na cristalização de proteínas, aprimorando esse processo e contribuindo significativamente para estudos em biologia molecular [71]. LIs à base de imidazólio, por exemplo, demonstraram maior eficiência de cristalização em comparação a sais inorgânicos tradicionais em proteínas como a lisozima [74], ribonuclease A e glicose isomerase [75, 76].

Análises cristalográficas de proteínas cristalizadas na presença de LIs revelam interações direcionais entre os íons do LI e a superfície proteica. Em um estudo representativo, observou-se que o ânion nitrato do etilamônio nitrato (EAN) se liga

preferencialmente a resíduos carregados positivamente na superfície da lisozima, promovendo a nucleação cristalina [77].

O efeito de LIs em proteínas em solução difere dos mecanismos clássicos induzidos por sais inorgânicos, como os efeitos de *salting-in* e *salting-out* [77]. Enquanto sais convencionais atuam predominantemente por modulação da atividade da água e efeitos de exclusão volumétrica, os LIs estabelecem ligações químicas específicas com sítios da superfície proteica, alterando seletivamente a energia interfacial e favorecendo a formação de arranjos ordenados durante a nucleação [77].

Os LIs também têm aplicação em sistemas de microemulsão, melhorando a estabilidade de proteínas em formulações farmacêuticas complexas [31, 78]. Além disso, os LIs são úteis na eletroforese em gel, onde melhoram a resolução das bandas proteicas, facilitando análises bioquímicas de alta precisão [79].

As interações entre LIs e proteínas constituem, assim, um campo de estudo promissor, com potencial significativo para avanços científicos e tecnológicos. A exploração contínua dessas interações promete expandir o conhecimento fundamental sobre esses sistemas e possibilitar o desenvolvimento de aplicações práticas inovadoras em biomedicina, biotecnologia e setores industriais diversos.

## Capítulo 2

# Soluções com proteínas

Este capítulo apresenta os fundamentos teóricos necessários para compreender a solvatação em torno de biomoléculas, com ênfase nas proteínas. A distribuição das moléculas do solvente na solução será descrita utilizando funções de distribuição, destacando sua relevância para a análise das interações soluto-solvente e para o entendimento do comportamento estrutural e funcional das proteínas.

### 2.1 Aspectos gerais da solvatação

A solvatação pode ser compreendida como as interações intermoleculares entre uma molécula ou íon em solução e o solvente ao seu redor. Em meio aquoso, este processo influencia uma ampla variedade de fenômenos que ocorrem em solução, como a transferência e estabilização de carga, reatividade química e enzimática, o efeito hidrofóbico, solubilidade, separação de fases e precipitação [80–82]. Desta forma, uma compreensão profunda de como ocorre a solvatação de um soluto em um solvente é essencial para o entendimento das propriedades químicas da solução, permitindo uma melhor utilização dessas propriedades em diferentes aplicações.

Os termos *soluto* e *solvente* são geralmente aplicados a misturas diluídas na fase líquida, nas quais o soluto (componente minoritário) está disperso no solvente (componente majoritário). Por esse motivo, o conceito de solvatação é, por vezes, ampliado para incluir a influência de qualquer ambiente circundante no qual uma molécula está inserida. Diversos tipos de interações e efeitos dinâmicos desempenham papéis cruciais no processo de solvatação. Normalmente, as interações soluto-solvente são dominadas por forças eletrostáticas (como interações de cargas, dipolos e dipolos induzidos), bem como por ligações de hidrogênio e

forças repulsivas, ambas com componentes eletrostáticos.

Um aspecto fundamental na avaliação da solvatação em termos energéticos é compreender a energia livre de solvatação. Este conceito pode ser elucidado por meio de um processo hipotético que descreve o fenômeno da solvatação: considerando a transferência do soluto de um estado de referência (no vácuo) para uma solução aquosa diluída, a energia livre de solvatação é, então, definida como a variação na energia livre associada a esse processo, representando a contribuição termodinâmica das interações entre o soluto e o solvente:



onde a transferência de fase representada pela Equação 2.1.1 pode ser dividida em dois processos fundamentais: (a) a energia necessária para criar uma cavidade no líquido e (b) a energia requerida para acomodar o soluto na cavidade, considerando as interações entre soluto e solvente que serão formadas, com eventual reorganização do solvente ao redor do soluto. Assim, a energia livre de solvatação ( $\Delta G_{\text{sol}}$ ) é composta por contribuições entálpicas e entrópicas, podendo ser expressa como:

$$\Delta G_{\text{sol}} = \Delta H_{\text{sol}} - T\Delta S_{\text{sol}} \quad (2.1.2)$$

assim, os termos da Equação 2.1.2 podem ser decompostos para incluir os processos de abertura da cavidade ( $\Delta G_1$ ) e inserção do soluto ( $\Delta G_2$ ), resultando na seguinte expressão:

$$\Delta G_{\text{sol}} = \Delta G_1 + \Delta G_2 = \Delta H_1 - T\Delta S_1 + \Delta H_2 - T\Delta S_2 \quad (2.1.3)$$

onde a equação representa um processo envolvendo a quebra de interações intermoleculares coesivas no solvente ( $\Delta H_1$ ), a formação de uma vacância sob pressão constante e a redução da entropia configuracional da rede de ligações de hidrogênio ( $\Delta S_1$ ). Por essas razões,  $\Delta G_1$  é geralmente grande e positivo, desempenhando um papel central no efeito hidrofóbico. Por exemplo, em modelos atomísticos usados em simulações de dinâmica molecular, as cavidades formadas para biomoléculas são frequentemente definidas com base na área acessível ao solvente (SASA, do inglês *Solvent Accessible Surface Area*).

Por sua vez,  $\Delta G_2$  é a energia livre necessária para inserir o soluto na cavidade e ativar as interações entre soluto e solvente. Esse processo inclui interações eletrostáticas e ligações de hidrogênio favoráveis ( $\Delta H_2$ ), além de outras possíveis interações.

A combinação dos termos que compõem  $\Delta G_1$  e  $\Delta G_2$  fornece uma visão abrangente dos processos de solvatação, ainda que na realidade as contribuições de cada termo sejam difíceis de serem separadas na realidade. Enquanto  $\Delta H_1$  inclui os custos associados à ruptura da rede de ligações de hidrogênio do solvente,  $\Delta H_2$  representa os benefícios energéticos decorrentes das interações soluto-solvente. A análise desses termos é fundamental para compreender os mecanismos moleculares subjacentes a fenômenos como o efeito hidrofóbico, a estabilidade de biomoléculas e suas interações com o meio solvente.

### 2.1.1 Solvatação de biomoléculas

Proteínas têm suas funções biológicas intrinsecamente vinculadas à sua estrutura tridimensional, e a compreensão dessas funcionalidades depende de sua identificação e análise em diferentes meios [83]. O estudo de uma proteína específica, a partir de uma amostra biológica, geralmente requer seu isolamento. Para isso, são empregados processos de separação baseados na solubilidade da proteína em diversos solventes [80]. Dessa forma, compreender os mecanismos que governam a dissolução e estabilização de proteínas em solução é essencial.

A conformação enovelada, frequentemente globular, é caracterizada pela exposição de resíduos hidrofílicos à solução e pela proteção de resíduos hidrofóbicos da água, sendo a conformação preferida pela maioria das proteínas, principalmente devido ao efeito hidrofóbico [84]. Tal efeito consiste em uma força efetiva que minimiza a exposição de superfícies apolares ao solvente aquoso, resultado das ligações de hidrogênio altamente favoráveis entre moléculas de água, interrompidas pela inserção de um soluto apolar na solução.

O custo de energia livre para solvatar moléculas apolares é elevado devido ao aumento da entalpia associado à quebra de ligações de hidrogênio da água, bem como à diminuição da entropia causada pela reorganização da rede de ligações de hidrogênio do solvente ao redor do soluto. No entanto, se os grupos apolares do soluto puderem ser protegidos da água, o custo de energia livre é reduzido. Tal mecanismo é fundamental para o colapso de cadeias polipeptídicas em estruturas globulares, onde os grupos apolares tendem a se localizar no interior do glóbulo.

Vários fatores podem estabilizar ou desestabilizar proteínas em solução, especialmente considerando que estas são apenas marginalmente estáveis [85]. O efeito hidrofóbico é um dos principais responsáveis pela estabilização do estado enovelado das proteínas. Outros fatores incluem interações de van der Waals entre os átomos das cadeias laterais, ligações de hidrogênio entre grupos laterais ou com átomos do *backbone*, e pontes salinas entre cadeias laterais carregadas.

Ao adicionar cossolventes a uma solução, as interações intermoleculares entre as moléculas do solvente e o soluto podem ser alteradas. Para solutos complexos, como proteínas, essas alterações influenciam o  $\Delta H$  e  $\Delta S$  da dissolução. Em particular, a interação de proteínas com o solvente é melhor quantificada por meio do excesso ou déficit de osmólitos (e água) ao redor da macromolécula em relação à solução bulk. O volume afetado pela presença da macromolécula, conhecido como “volume de correlação” [86], não se limita à vizinhança imediata da proteína, podendo se estender por diversas camadas de solvatação até o *bulk*.

O ciclo termodinâmico apresentado na Figura 2.1 ajuda a ilustrar a solvatação de proteínas em meios aquosos. Considerando possíveis mudanças conformacionais na Equação 2.1.1, o termo  $\Delta G_1$  pode ser expandido em dois componentes:  $\Delta G_{11}$ , que representa a energia livre necessária para criar uma cavidade na água capaz de acomodar a molécula solvatada, e  $\Delta G_{12}$ , que corresponde à energia livre necessária para induzir mudanças conformacionais no soluto, preparando-o para o estado solvatado final. O termo  $\Delta G_2$ , por sua vez, descreve a energia livre requerida para inserir o soluto na cavidade e ativar as interações entre soluto e solvente. Tal processo inclui a ativação de interações eletrostáticas e ligações de hidrogênio, bem como a reorganização do solvente ao redor do soluto.

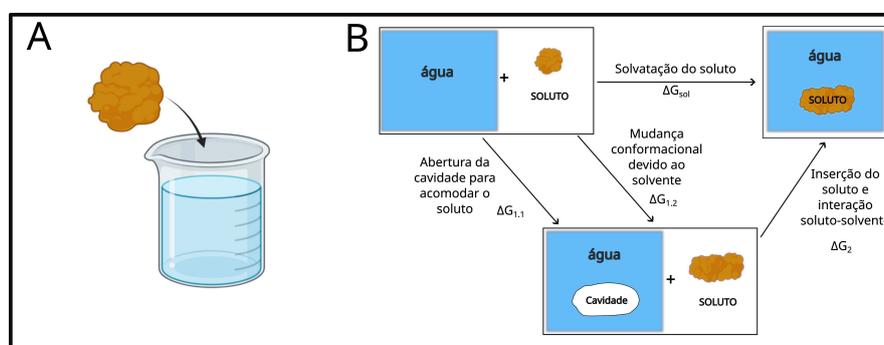


Figura 2.1: (A) Representação esquemática da adição de uma proteína a uma solução aquosa. (B) Ciclo termodinâmico utilizado para avaliar a energia livre de solvatação de uma proteína em água.

No passo 1 ( $\Delta G_1$ ), é possível considerar a média sobre a entropia conformacional do solvente associada à forma da cavidade e à conformação do soluto. Já no passo 2 ( $\Delta G_2$ ), a reorganização configuracional do solvente e as variações correspondentes na força das interações são contempladas, refletindo a complexidade do processo de solvatação. Uma análise detalhada deste fenômeno para proteínas revela múltiplos aspectos cruciais para sua compreensão. A introdução do soluto em uma solução composta por um cossolvente e água perturba as interações preexistentes, podendo estabelecer novas dinâmicas no sistema. Estas mudanças podem, por exemplo, influenciar diretamente a conformação da proteína, destacando a relevância do meio solvente na estabilidade estrutural das biomoléculas.

## 2.2 Estrutura de solvatação

Os estudos computacionais são ferramentas importantes para entender, 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. As funções de distribuição descrevem a probabilidade de encontrar uma partícula em 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 do solvente [87, 88].

Funções de distribuição radiais (RDFs, em inglês *Radial Distribution Functions*) são geralmente calculadas levando em conta a distância entre centros de massa (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 2.2.

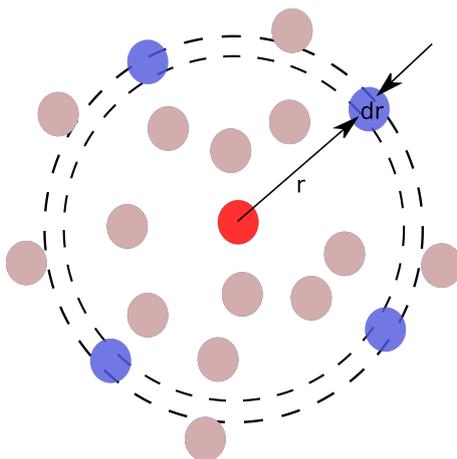


Figura 2.2: Configuração que representa o cálculo da função de distribuição radial. A partícula vermelha é a partícula de referência. A partir desta partícula, tomamos uma distância  $r$ . Esta distância é entre centros de massa, especificamente, entre o centro de massa da partícula vermelha e da partícula azul apontada pelas setas. Figura reproduzida com autorização do autor a partir de [89].

As RDFs podem ser melhor compreendidas pensando na seguinte abordagem qualitativa: supondo que estivéssemos em 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 em 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 em uma casca esférica de raio  $r$ ). O que é observado, de cima da partícula, corresponde a uma densidade média para cada distância. Esta densidade é um desvio da densidade esperada para a solução ( $\rho$ ).

$g(r)$ , portanto, pode ser avaliada considerando a densidade média de moléculas em uma casca de raio  $r$ . A densidade de moléculas na casca será igual a  $\rho g(r)$ . Então, o número médio de moléculas que realmente existem (ao redor da molécula que estamos sentados) na 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 a

uma distribuição aleatória).

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

$$g(r) = \frac{\rho(r)}{\rho^{(bulk)}} \quad (2.2.1)$$

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 [87]. Sendo assim, na função de distribuição calculada a partir das posições dos átomos do sólido, devem existir alguns picos em posições bem específicas. Esses picos possuem uma pequena largura em função da vibração dos átomos em torno da posição no retículo. Os gases e líquidos, por sua vez, possuem mobilidade muito maior em comparação aos sólidos, não apresentando picos em distâncias periódicas.

A Figura 2.3 ilustra uma função de distribuição radial típica para um líquido, que pode ser melhor compreendida por meio de sua divisão em regimes distintos: o primeiro regime corresponde à região em que  $g(r) = 0$  para distâncias  $r$  menores que o raio  $\sigma$ , onde  $\sigma$  representa o raio aproximado da molécula, indicado na Figura 2.3 por aproximadamente 1,5 Å. O segundo regime ocorre para  $\sigma < r < 3\sigma$ , onde  $g(r) > 1$ , indicando a presença de interações efetivas entre as moléculas. Esse regime é caracterizado por duas regiões distintas: a primeira camada de solvatação, representada pela cor lilás, e as camadas adicionais de solvatação, representadas pela cor vermelha, que refletem uma ordem estrutural mais prolongada no líquido. O terceiro regime, para  $r > 3\sigma$ , corresponde à região *bulk*, representada pela cor verde na Figura 2.3, onde a influência das interações intermoleculares locais torna-se menos significativa [87, 88].

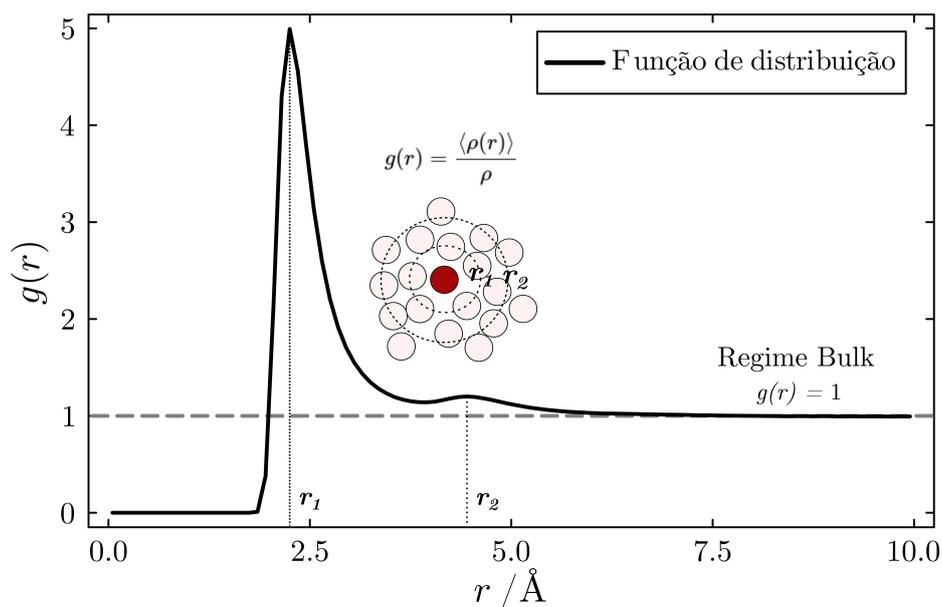


Figura 2.3: Função de distribuição radial ( $g(r)$ ) para um líquido genérico.  $g(r) > 1$ , indicando a presença de interações efetivas e destacando as camadas de solvatação, indicadas pelos picos em  $r_1$  e  $r_2$ ; e em longas distâncias, maiores que 7 Å, correspondente à região *bulk*, onde a estrutura local do líquido torna-se menos pronunciada.

Em comparação com os líquidos, os gases exibem funções de distribuição radial com menor estruturação, onde os picos não apresentam periodicidade. Por outro lado, os sólidos demonstram picos bem definidos em distâncias específicas, relacionados à organização do retículo cristalino da substância. No estado líquido, as moléculas possuem uma mobilidade moderada e apresentam certa organização ao longo da solução, especialmente em substâncias altamente polares ou capazes de formar ligações de hidrogênio, como a água. Essa organização é refletida na função de distribuição radial, que exhibe picos associados às camadas de solvatação e à ordem local.

As funções de distribuição radial fornecem informações microscópicas valiosas sobre o sistema. No Capítulo 5, a análise da estrutura de solvatação de proteínas será explorada em maior profundidade, com uma explicação detalhada dos métodos empregados neste trabalho, em especial sobre as funções de distribuição de mínima distância [90].

## Capítulo 3

# Simulações de Dinâmica Molecular

A dinâmica molecular (DM) é uma ferramenta computacional amplamente utilizada na simulação de sistemas químicos, especialmente aqueles que envolvem biomoléculas. Este capítulo apresenta uma visão geral da técnica, destacando seus fundamentos mecânico-estatísticos e as etapas essenciais para a realização de uma simulação.

### 3.1 Visão Geral do Método

A DM clássica é uma das ferramentas computacionais mais amplamente utilizadas para investigar sistemas líquidos e sólidos ao nível molecular, utilizando os princípios da física newtoniana. Por meio de simulações de DM, é possível acessar o comportamento dinâmico microscópico e dependente do tempo das partículas que compõem um sistema molecular.

Os métodos de simulação por DM geram uma trajetória, ou seja, uma série de configurações espaciais e velocidades do sistema no tempo obtidas pela propagação de um conjunto inicial de coordenadas e velocidades [91]. Essa propagação, no caso de simulações clássicas, é realizada pela aplicação da segunda lei de Newton em uma série finita de passos de tempo. Assim, sucessivas configurações do sistema são geradas por meio da integração das equações de movimento de Newton [92].

O princípio fundamental da DM é que forças aplicadas a partículas causam aceleração, conforme ilustrado na Figura 3.1. Esta figura apresenta um sistema hipotético com duas partículas,  $i$  e  $j$ , que, em um instante inicial  $t$ , possuem posições indicadas pelas bolas pretas e velocidades  $v_t$ , representadas pelas setas pretas. As partículas exercem forças mútuas de intensidade  $F_t$ . A presença dessas forças altera as velocidades no instante  $t + 1$ , fazendo com que passem de  $v_t$  (setas pretas) para  $v_{t+1}$  (setas vermelhas) e,

consequentemente, modificando suas trajetórias.

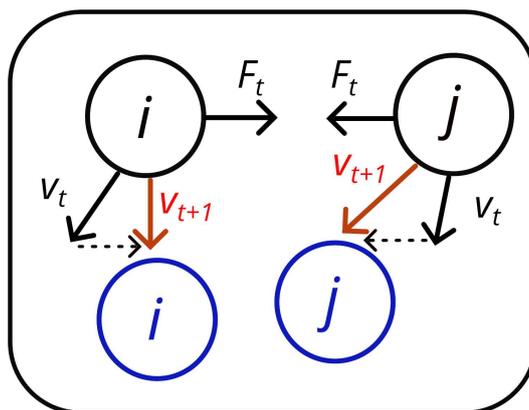


Figura 3.1: Sistema hipotético com duas partículas ilustrando o efeito de forças na velocidade das partículas ao longo do tempo.

De forma simplificada, o exemplo da Figura 3.1 ilustra como, em um sistema com duas partículas, as posições e velocidades são alteradas devido à presença de forças. Por exemplo, a Equação 3.1.1, que corresponde à segunda lei de Newton, mostra que, a partir das forças atuando nas partículas em determinadas posições, obtém-se as acelerações que permitem o cálculo de posições e velocidades.

$$a(r) = \frac{1}{m}F(r) \quad (3.1.1)$$

onde  $a(r)$  representa a aceleração das partículas em função da posição  $r$ ,  $F(r)$  é a força atuando sobre as partículas em função da posição  $r$ ,  $m$  é a massa das partículas.

Sistemas químicos e biológicos, sendo os principais focos da DM, envolvem muitos átomos, tornando inviável a solução analítica das equações de Newton. Para estes sistemas, é necessário recorrer a métodos numéricos e, além disso, para obter aproximações precisas do deslocamento das partículas, as forças devem ser avaliadas em intervalos de tempo muito pequenos. A Figura 3.2 mostra, simplificada, como o intervalo/passo de tempo influencia na posição das partículas. Se o tempo é muito grande, as forças (representadas pelas setas) irão aumentar gradualmente até que, devido ao intervalo de tempo ser grande, haja um deslocamento muito expressivo entre um passo (representado pelos retângulos da Figura 3.2) e outro, inviabilizando a acurácia física da simulação. Contudo, com um intervalo de tempo menor, o movimento das partículas entre *frames* não sofre variações abruptas, permitindo a melhor descrição do sistema.

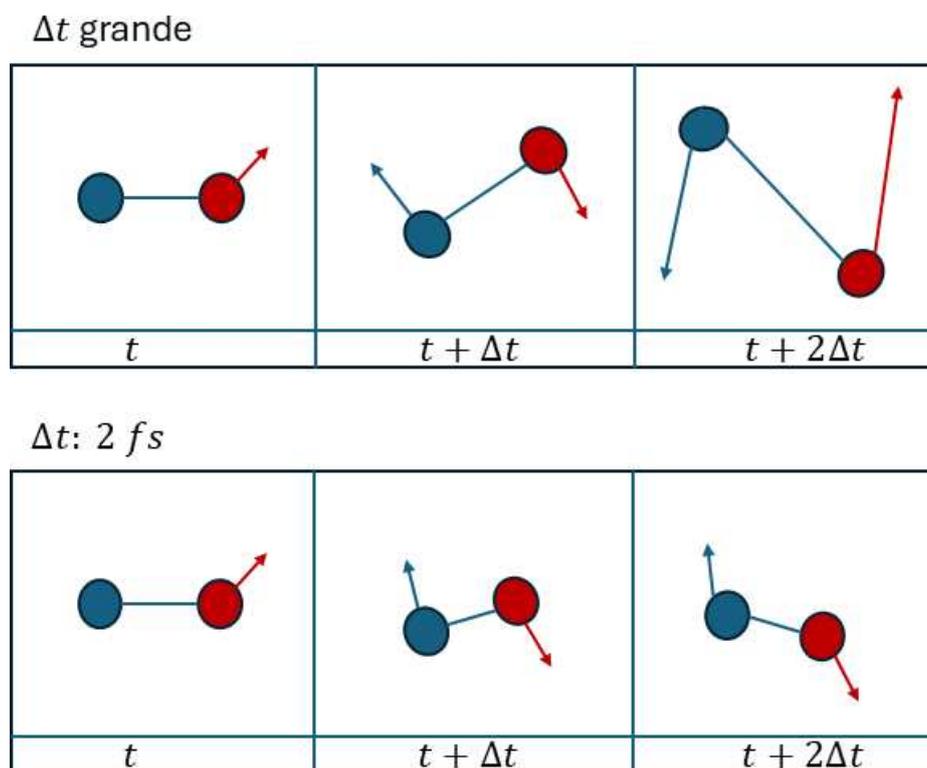


Figura 3.2: Exemplificação do efeito de passos de tempo grandes, inadequados e pequenos, da ordem de femtossegundos, na variação da posição de duas partículas em uma simulação de DM.

Na escolha do algoritmo de integração e da magnitude do passo de tempo em simulações de dinâmica molecular, a conservação da energia total do sistema é um aspecto fundamental. Algoritmos de integração, como o de Verlet ou Velocity Verlet, são projetados para preservar a energia em sistemas isolados ao longo do tempo, desde que o passo de tempo seja suficientemente pequeno para resolver adequadamente os movimentos atômicos mais rápidos, como as vibrações de ligações envolvendo hidrogênios [92, 93].

A definição do passo de tempo ideal precisa da monitoração da estabilidade da energia total ao longo da simulação: flutuações pequenas e simétricas são aceitáveis, mas tendências sistemáticas de aumento ou diminuição indicam a necessidade de ajustes no esquema de integração, ou na escolha do passo de tempo [93]. Este controle é especialmente importante em simulações de equilíbrio, onde a conservação da energia indica a fidelidade física do modelo empregado.

Embora os fundamentos físicos das interações intermoleculares sejam bem estabelecidos, há uma mistura extremamente complexa de fenômenos quânticos atuando em distâncias muito pequenas que acontecem na realidade. As forças entre átomos e moléculas

resultam de interações dinâmicas entre elétrons orbitando os núcleos atômicos. Devido à complexidade dessas interações entre nuvens eletrônicas, elas não podem ser descritas analiticamente ou calculadas numericamente com rapidez suficiente para viabilizar simulações dinâmicas em escalas temporais necessárias para o estudo de sistemas condensados. Um cálculo de DM regular envolve, comumente, centenas de milhares de átomos em sistemas com biomoléculas, sendo o custo computacional para simular tal sistema usando mecânica quântica inviável. Apesar disso, vale destacar que atualmente existem métodos de DM que representam apenas algumas partes do sistema usando o formalismo de química quântica [94].

Para tornar as simulações de DM viáveis, é essencial avaliar as interações moleculares eficientemente. Nesse contexto, as interações são aproximadas por meio de uma função empírica simples de energia potencial. Essa função de energia potencial,  $U$ , desempenha um papel central nas simulações de DM, pois permite o cálculo das forças. A força que atua sobre um objeto é:

$$\mathbf{F} = -\nabla U$$

onde  $\mathbf{F}$  representa a força e  $\nabla U$  denota o gradiente da energia potencial.

Conceitos da mecânica estatística são essenciais para extrair propriedades macroscópicas observáveis, como pressão, energia interna, volume, entropia e energia livre, a partir das propriedades microscópicas obtidas das trajetórias [92]. Um conceito fundamental na mecânica estatística é que, após tempo suficiente, o sistema percorrerá todos os microestados consistentes com as restrições do seu estado termodinâmico. Realizando  $N$  medidas independentes, o valor observável de uma propriedade  $P$  pode ser calculado como:

$$P_{\text{obs}} = \frac{1}{N} \sum_{a=1}^N P_a, \quad (3.1.2)$$

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

A soma na Equação 3.1.2 pode ser reescrita como:

$$P_{\text{obs}} = \sum_{\nu} \left[ \frac{N_{\nu}}{N} \right] P_{\nu}, \quad (3.1.3)$$

onde  $N_{\nu}$  é o número de vezes que o estado  $\nu$  é observado, e  $P_{\nu}$  é o valor esperado de  $P$  quando o sistema está no estado  $\nu$ . Após um longo tempo, assumindo que todos os estados relevantes foram visitados na proporção correta, a fração de tempo no estado  $\nu$ ,  $p_{\nu}$  é igual a  $\left[ \frac{N_{\nu}}{N} \right]$ , resultando em:

$$P_{\text{obs}} = \sum_{\nu} p_{\nu} P_{\nu} \equiv \langle P \rangle. \quad (3.1.4)$$

Onde  $\langle P \rangle$  é a média no ensemble da propriedade  $P$ . Essa equivalência baseia-se na hipótese ergódica, que implica que a média temporal é igual à média do ensemble [87, 88].

## 3.2 Cálculo das forças

Os campos de força consistem em um conjunto de equações que descrevem as interações relevantes em um sistema de interesse, como, por exemplo, os átomos de uma proteína e o solvente ao seu redor. As forças, conforme mencionado anteriormente, representam as interações entre os átomos individuais que compõem o sistema.

Na DM, as moléculas são tratadas como uma coleção de átomos descrita por forças clássicas. Isso implica que, em um sistema simulado por DM, as partículas são mantidas unidas por forças elásticas e harmônicas. Um conjunto completo de potenciais de interação entre as partículas (átomos, moléculas, etc.) é denominado “campo de força” [95, 96].

A energia potencial de uma molécula é descrita como uma função que considera a resistência ao estiramento de ligações, torção de ângulos de ligação, variação de ângulos diedros e repulsão ou atração entre os átomos. Essa equação de energia é utilizada para determinar comprimentos de ligação, ângulos e diedros [97]. Como a DM utiliza uma abordagem clássica, fenômenos quânticos, como a quebra de ligações químicas, não são representados [92].

Portanto, a qualidade dos campos de força precisos é essencial para que as simulações de DM produzam resultados consistentes com os experimentos, permitindo não apenas a validação de dados experimentais, mas também a investigação de aspectos que podem ser inacessíveis por métodos experimentais tradicionais. Podemos representar um

campo de força por:

$$U_{\text{total}} = U_{\text{ligação}} + U_{\text{não-ligação}}. \quad (3.2.1)$$

Onde  $U_{\text{total}}$  representa a energia total do sistema,  $U_{\text{ligação}}$  corresponde às energias intramoleculares e  $U_{\text{não-ligação}}$  às energias intermoleculares e entre átomos distantes na mesma molécula. As interações ligadas ( $U_{\text{ligação}}$ ) descrevem as forças que atuam entre átomos próximos dentro da mesma molécula, como dois átomos conectados por uma ligação, três átomos formando um ângulo e quatro átomos associados a um ângulo diedro:

$$U_{\text{ligação}} = U_{\text{ligações}} + U_{\text{ângulos}} + U_{\text{diedros}}. \quad (3.2.2)$$

Por outro lado, as interações não-ligadas ( $U_{\text{não-ligação}}$ ) incluem interações de Coulomb e forças de van der Waals:

$$U_{\text{não-ligação}} = U_{\text{Coulomb}} + U_{\text{van der Waals}}. \quad (3.2.3)$$

Combinando as Equações 3.2.1, 3.2.2 e 3.2.3, a energia total do sistema pode ser expressa como:

$$U_{\text{total}} = U_{\text{ligações}} + U_{\text{ângulos}} + U_{\text{diedros}} + U_{\text{Coulomb}} + U_{\text{van der Waals}}. \quad (3.2.4)$$

Assim, para as interações não-ligadas, consideram-se as forças de Coulomb e o potencial de Lennard-Jones. No campo de força OPLS, utilizado neste trabalho, as interações ligadas e os respectivos termos de energia são representados por:

$$U_{\text{ligações}}(r) = k_b(r - r_0)^2, \quad (3.2.5)$$

$$U_{\text{ângulos}}(\theta) = k_\theta(\theta - \theta_0)^2, \quad (3.2.6)$$

$$U_{\text{diedros}}(\omega) = \sum_j \frac{1}{2} V_j [1 + \cos(j\omega - \phi)]. \quad (3.2.7)$$

Onde  $k_b$  é a constante de força da ligação, determinando a rigidez da ligação;  $k_\theta$  é a constante de força do ângulo, que determina a rigidez do ângulo;  $V_j$  é a amplitude do potencial para o termo harmônico de ordem  $j$ ;  $\omega$  é o ângulo diedro (ou torsional), que define a rotação ao longo de uma ligação química, e  $\phi$  é a fase do potencial, que determina o deslocamento do mínimo da energia no espaço angular.

Os ângulos diedros são expressos como funções periódicas de cossenos, utilizando a soma dos múltiplos harmônicos para representar as interações rotacionais. As interações não-ligadas são modeladas pela lei de Coulomb, para forças eletrostáticas, e pelo potencial de Lennard-Jones, que descreve as forças de van der Waals:

$$U_{\text{Coulomb}}(r_{ij}) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}, \quad (3.2.8)$$

$$U_{\text{Lennard-Jones}}(r_{ij}) = 4\epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right]. \quad (3.2.9)$$

Onde  $q_i$  e  $q_j$  representam as cargas elétricas das partículas  $i$  e  $j$ , respectivamente,  $r_{ij}$  é a distância entre essas partículas, e  $\epsilon_0$  é a permissividade elétrica do vácuo ( $8.854 \times 10^{-12} \text{ C}^2 \text{ N}^{-1} \text{ m}^{-2}$ ). No potencial de Lennard-Jones,  $\epsilon$  corresponde à profundidade do poço de energia, que determina a intensidade da interação atrativa, enquanto  $\sigma$  é o parâmetro que define a distância na qual o potencial se anula.

O potencial de Lennard-Jones, amplamente utilizado para descrever interações de van der Waals, é um dos modelos para interação intermolecular entre partículas mais fundamentais em DM. Este potencial captura as características essenciais das interações entre átomos e moléculas, descrevendo a repulsão entre partículas em distâncias muito curtas, a atração em distâncias intermediárias e a insignificância da interação em grandes distâncias,

como ilustrado na Figura 3.3. Trata-se de um potencial de pares, ou seja, considera apenas interações entre duas partículas, sem incluir efeitos de três ou mais corpos.

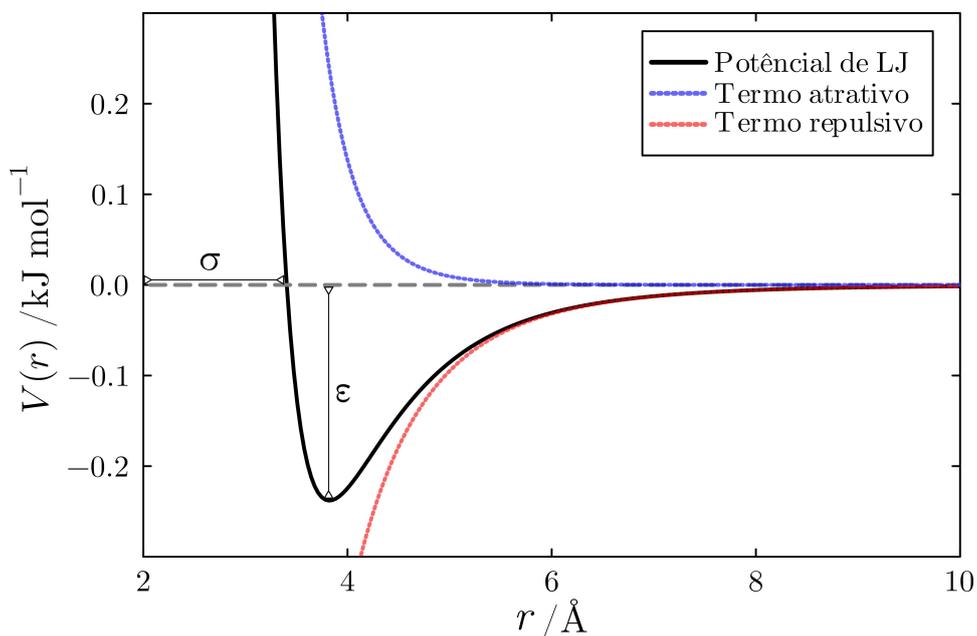


Figura 3.3: Potencial de Lennard-Jones para dois átomos de argônio em interação. A curva preta representa o potencial total, enquanto as curvas pontilhadas destacam os termos atrativo (em azul) e repulsivo (em vermelho).

Com base na suposição de interações aditivas entre pares, é possível juntar as Equações 3.2.7 e 3.2.9 na energia total como:

$$\begin{aligned}
 U(r) = & \sum_b k_b (r - r_0)^2 + \sum_\theta k_\theta (\theta - \theta_0)^2 \\
 & + \sum_\omega \frac{1}{2} V_j [1 + \cos(j\omega - \phi)] \\
 & + \sum_i \sum_{j>i} \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \\
 & + 4\epsilon \sum_i \sum_{j>i} \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right].
 \end{aligned} \tag{3.2.10}$$

Onde os primeiros três termos descrevem as interações ligadas (ligações, ângulos e diedros), enquanto os dois últimos representam as interações não-ligadas. No primeiro termo,  $k_b$  é a constante de força associada à deformação da ligação,  $r$  é a distância atual entre os átomos ligados, e  $r_0$  é a distância de equilíbrio da ligação. No segundo termo,  $k_\theta$  é a constante de

força associada à deformação angular,  $\theta$  é o ângulo atual entre as ligações adjacentes, e  $\theta_0$  é o ângulo de equilíbrio. O terceiro termo descreve as interações de torção ou rotação em torno de ligações, onde  $V_j$  é a amplitude da barreira de torção,  $\omega$  é o ângulo de torção,  $j$  é a periodicidade, e  $\phi$  é a fase.

Os dois últimos termos correspondem às interações não ligadas. O quarto termo descreve as interações eletrostáticas, em que  $q_i$  e  $q_j$  representam as cargas parciais dos átomos  $i$  e  $j$ , respectivamente,  $r_{ij}$  é a distância entre esses átomos, e  $\epsilon_0$  denota a permissividade elétrica do vácuo. O quinto termo, por sua vez, descreve as interações de van der Waals por meio do potencial de Lennard-Jones, onde  $\epsilon$  corresponde à profundidade do poço de energia (indicando a intensidade da atração),  $\sigma$  é o parâmetro de distância relacionado ao diâmetro efetivo dos átomos, e  $r_{ij}$  é a distância entre os átomos  $i$  e  $j$ . Esse termo combina uma contribuição repulsiva, proporcional a  $\left(\frac{\sigma}{r_{ij}}\right)^{12}$ , e uma contribuição atrativa, proporcional a  $\left(\frac{\sigma}{r_{ij}}\right)^6$ .

Os campos de força são parametrizados, ou seja, os parâmetros utilizados para descrever as interações são definidos com base em dados empíricos e/ou cálculos quânticos. Esses parâmetros são ajustados para aplicações específicas. Por exemplo, o campo de força OPLS (*Optimized Potentials for Liquid Simulations*) é especialmente parametrizado para reproduzir comportamentos experimentais em simulações de líquidos de forma geral [98, 99]. Outros campos de força amplamente utilizados são CHARMM, GROMOS e GAFF [100–102].

O campo de força utilizado neste trabalho é uma modificação do OPLS [98, 99], especialmente parametrizado para líquidos iônicos imidazólicos [103]. Este campo de força inclui um sítio virtual localizado no centro do anel imidazólico, cuja função é representar a carga positiva deslocalizada característica desses anéis. A parametrização das cargas foi realizada com base em cálculos de energia livre de hidratação do cátion imidazólico em solução aquosa. Os termos do potencial foram refinados para reproduzir dados experimentais de densidade, entalpia de vaporização, viscosidade, coeficiente de difusão, tensão superficial e capacidade calorífica. Devido ao uso do sítio virtual, o campo de força foi denominado OPLS-VSIL (*Optimized Potentials for Liquid Simulations - Virtual Site Ionic Liquid*) [103].

Em sistemas caracterizados por forte correlação iônica, alta viscosidade e estruturas moleculares complexas, como líquidos iônicos, a representação precisa das interações eletrostáticas e da polarizabilidade torna-se essencial. Nesses casos, campos de força clássicos com cargas fixas podem reproduzir com razoável acurácia propriedades

termodinâmicas como densidade e entalpia de vaporização. No entanto, tais modelos frequentemente falham na predição de propriedades dinâmicas, como coeficientes de autodifusão e viscosidade, que costumam ser, respectivamente, subestimados e superestimados em relação aos dados experimentais.

Isso ocorre porque os modelos de carga fixa tendem a superestimar as interações coulômbicas, promovendo uma dinâmica excessivamente lenta e dificultando a difusão dos íons. Uma solução comum para mitigar esses efeitos consiste na adoção de modelos com cargas escalonadas, também conhecidos como *scaled-charge models*. Nessa abordagem, as cargas atômicas são reduzidas tipicamente para cerca de 80% de seus valores formais, o que suaviza as interações eletrostáticas e resulta em melhorias significativas nas propriedades de transporte, sem acréscimo no custo computacional quando comparado aos modelos tradicionais [104].

Apesar das vantagens, os modelos com cargas escalonadas apresentam limitações importantes. Estudos demonstram que essa simplificação pode comprometer a acurácia de propriedades estruturais e termodinâmicas, levando à subestimação de grandezas como densidade líquida e entalpia de vaporização. Além disso, tais modelos não corrigem adequadamente o comportamento de triagem eletrostática nem a resposta dielétrica do sistema, aspectos críticos para a descrição realista de líquidos iônicos [105].

Por outro lado, os campos de força explicitamente polarizáveis oferecem melhorias substanciais na descrição das propriedades termodinâmicas e dinâmicas, ao incorporar diretamente os efeitos de interação de muitos corpos. Diferentes estratégias estão disponíveis para representar a polarizabilidade eletrônica, incluindo modelos com dipolos induzidos, osciladores de Drude e cargas flutuantes. Embora essas abordagens apresentem maior custo computacional, sua capacidade de reproduzir com fidelidade tanto a estrutura local quanto as propriedades coletivas dos líquidos iônicos é amplamente reconhecida [106–108].

Nesta tese, optou-se pela utilização de cargas escalonadas devido à inviabilidade prática de empregar campos de força polarizáveis em sistemas compostos simultaneamente por água, proteínas e líquidos iônicos. Além do custo computacional elevado, há ainda desafios relacionados à compatibilidade entre os diversos componentes do sistema, dada a ausência de campos polarizáveis amplamente validados para proteínas e solventes mistos contendo líquidos

iônicos. Apesar das limitações dos campos de força com carga escalonada, os parâmetros OPLS-VSIL foram cuidadosamente ajustados para refinar os parâmetros de Lennard-Jones dos íons, a fim de reproduzir com precisão as interações intermoleculares locais [103].

### 3.3 Etapas básicas de uma simulação de dinâmica molecular

#### 3.3.1 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 frequentemente cúbica, e as moléculas de água e de LI aleatoriamente distribuídas pelo espaço restante. As dimensões da caixa são escolhidas de forma que a densidade numérica (razão entre o número de moléculas na caixa pelo seu volume) seja correspondente àquela desejada para o sistema real.

#### 3.3.2 Movimentação das partículas

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 utilizados na integração das equações de movimento. Neste método, a integração é dividida em intervalos temporais pequenos ( $\Delta t$ ), permitindo a simulação das movimentações de maior frequência que acontecem no sistema (são normalmente as vibrações das ligações).

O algoritmo de *Velocity-Verlet* é frequentemente usado no cálculo das trajetórias de partículas dinâmicas em simulações computacionais [109]. Resumidamente, o algoritmo calcula as posições usando uma expansão de Taylor de segunda ordem, como apresentado pela Equação 3.3.1. As velocidades no instante  $t + \Delta t$  são calculadas por meio das velocidades no instante  $t$  e a média das acelerações nos instantes  $t$  e  $t + \Delta t$ , conforme apresentado pela Equação 3.3.2:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 \quad (3.3.1)$$

$$v(t + \Delta t) = v(t) + \frac{a(t) + a(t + \Delta t)}{2}\Delta t. \quad (3.3.2)$$

onde  $a(t)$  é a aceleração da partícula no tempo  $t$ , determinada pelas forças agindo sobre a partícula divididas pela sua massa, e  $r(t)$  é a posição da partícula no tempo atual  $t$ .

No algoritmo *Velocity-Verlet*, dois vetores principais são sempre monitorados: o vetor contendo as posições e o vetor contendo as velocidades. O algoritmo a ser implementado pode seguir a seguinte sequência de passos:

- Inicialmente, as posições no instante  $t + \Delta t$  são calculadas por meio da equação:

$$\vec{x}(t + \Delta t) = \vec{x}(t) + \vec{v}(t)\Delta t + \frac{1}{2}\vec{a}(t)\Delta t^2.$$

- Novas acelerações são calculadas a partir das posições atualizadas,  $\vec{x}(t + \Delta t)$ . Estas acelerações são obtidas a partir da aplicação do campo de forças derivado do potencial ao novo conjunto de posições. Em termos gerais, as forças, para  $N$  partículas, são dadas por:

$$\vec{F}_i(t + \Delta t) = -\nabla_{\vec{x}_i} U(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N),$$

onde  $U$  é o potencial total do sistema, e  $\vec{F}_i$  representa a força atuando sobre o átomo  $i$ . As novas acelerações são então calculadas por meio da segunda lei de Newton:

$$\vec{a}_i(t + \Delta t) = \frac{\vec{F}_i(t + \Delta t)}{m_i}.$$

- Com as novas acelerações, as velocidades são atualizadas por:

$$\vec{v}(t + \Delta t) = \vec{v}(t) + \frac{1}{2}[\vec{a}(t) + \vec{a}(t + \Delta t)]\Delta t.$$

O erro global do método *Velocity-Verlet* é da ordem  $\mathcal{O}(\Delta t^2)$ , o que significa que o erro acumulado nas variáveis (posição e velocidade) ao longo de toda a simulação cresce proporcionalmente ao quadrado do tamanho do passo de tempo. Por outro lado, o erro local — ou seja, introduzido a cada passo individual — é da ordem  $\mathcal{O}(\Delta t^3)$ . Isso caracteriza o método como um integrador de segunda ordem, o que oferece um bom equilíbrio entre precisão e custo computacional [92, 93].

Além disso, como a aceleração é obtida a partir de forças derivadas de um potencial conservativo, a energia total do sistema (soma de energia cinética e potencial) tende a oscilar

ao redor de um valor médio constante ao longo da simulação. Esta conservação aproximada da energia reflete a boa estabilidade numérica do algoritmo e é uma das razões pelas quais o *Velocity-Verlet* é amplamente utilizado em simulações de dinâmica molecular.

Há uma grande variedade de integradores adicionais de dinâmica molecular, referências como [92] e [93] apresentam aspectos detalhados de cada um, entrando em detalhes sobre benefícios e limitações que vão além do escopo principal desta tese.

### 3.3.3 Ensembles estatísticos

O conjunto de configurações acessadas por um sistema em equilíbrio térmico define um *ensemble*. As propriedades mantidas constantes caracterizam o tipo de *ensemble* adotado [87].Dentre os mais comuns, destaca-se o *ensemble NVE* (microcanônico), no qual o número de partículas ( $N$ ), o volume ( $V$ ) e a energia ( $E$ ) são mantidos fixos. Outras opções incluem os *ensembles*  $\mu VT$  (grande-canônico),  $NVT$  (canônico) e  $NPT$  (isotérmico-isobárico), nos quais diferentes combinações de variáveis termodinâmicas são controladas. A Figura 3.4 apresenta uma visão ilustrativa de como seria cada sistema considerando os ensembles discutidos anteriormente. A escolha do *ensemble* adequado para uma simulação depende das condições experimentais que se deseja reproduzir e das propriedades termodinâmicas de interesse.

Em dinâmica molecular, as equações de movimento de Newton são resolvidas numericamente, garantindo a conservação da energia total do sistema no *ensemble* microcanônico ( $NVE$ ). No entanto, para simular condições mais realistas, como aquelas em que a temperatura e a pressão são mantidas constantes, são empregados algoritmos adicionais, conhecidos como termostatos e barostatos.

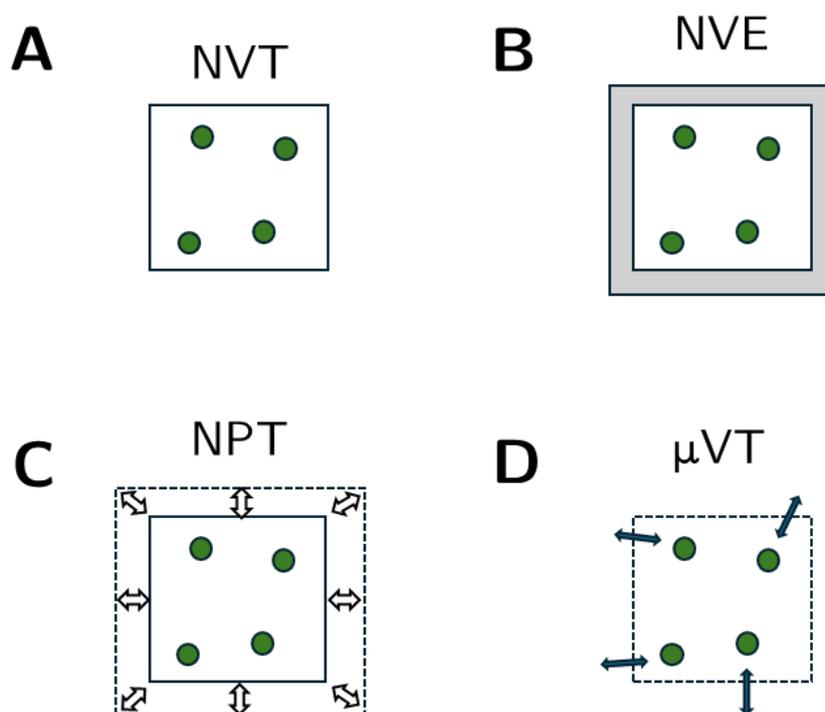


Figura 3.4: Representação esquemática de diferentes ensembles, são eles: A) canônico (NVT), B) microcanônico (NVE), C) isotérmico-isobárico (NPT) e D)  $\mu VT$  (grande-canônico). A caixa **A** representa um sistema em que a energia pode fluir para a vizinhança, com o sistema contringido em número de partículas e volume. A caixa **B** representa um sistema contringido por uma barreira adiabática, resultando em número de partículas, volume e energia constantes. A caixa **C** representa um sistema em uma força externa ajustada ao tamanho dele para manter a pressão constante por meio da variação do volume, resultando na pressão constante. Por fim, a caixa **D** apresenta um sistema aberto, que permite a passagem de matéria e energia.

### 3.3.4 Controle de temperatura

O controle da temperatura em simulações de dinâmica molecular pode ser realizado por meio de diferentes termostatos, que ajustam a energia cinética do sistema ao longo da simulação. Do ponto de vista da mecânica estatística, a temperatura ( $T$ ) de um sistema sem restrições está, essencialmente, relacionada com a energia cinética média ( $\langle K \rangle$ ) [110]. Para um sistema composto por  $N$  partículas, a energia cinética média é dada por:

$$\langle K \rangle = \frac{3}{2} N k_B T, \quad (3.3.3)$$

onde  $k_B$  é a constante de Boltzmann. A partir da Equação 3.3.3, nota-se que, para manter a temperatura fixa, é necessário que  $\langle K \rangle$  permaneça constante, embora a energia cinética instantânea ( $K$ ) em cada passo de simulação possa flutuar de acordo com:

$$K = \frac{1}{2} \sum_i m_i v_i^2, \quad (3.3.4)$$

onde  $m_i$  e  $v_i$  são, respectivamente, a massa e a velocidade da partícula  $i$  do sistema.

A Equação 3.3.4 sugere uma das formas mais diretas de controle de temperatura: o reescalonamento de velocidades. Neste método, comumente denominado *velocity rescaling*, ajusta-se a velocidade de cada partícula com base em um fator de escala  $\lambda$ , de modo a controlar  $\langle K \rangle$ . Esse fator é definido como:

$$v_i^{\text{nova}} = v_i^{\text{anterior}} \cdot \lambda, \quad (3.3.5)$$

com  $\lambda$  igual a:

$$\lambda = \sqrt{\frac{T_{\text{des}}}{T(t)}}, \quad (3.3.6)$$

onde  $T_{\text{des}}$  é a temperatura desejada e  $T(t)$  é a temperatura instantânea no tempo  $t$ . Assim, quando  $T(t) < T_{\text{des}}$ ,  $\lambda > 1$ , aumentando a velocidade das partículas. Quando  $T(t) > T_{\text{des}}$ ,  $\lambda < 1$ , reduzindo a velocidade. Esse procedimento é aplicado a cada passo da simulação.

Embora didático e simples de implementar, o método de reescalonamento direto das velocidades não reproduz corretamente o ensemble canônico ( $NVT$ ), pois suprime as flutuações naturais de temperatura. Como consequência, propriedades termodinâmicas como capacidade calorífica e distribuições de energia podem ser incorretas. Apesar disso, esse método é útil em etapas iniciais de termalização.

O termostato de Berendsen representa uma sofisticação do método anterior, realizando o controle térmico de maneira mais suave ao acoplar o sistema a um "banho térmico" externo com temperatura fixa [92, 111]. A taxa de variação da temperatura do sistema é proporcional à diferença entre a temperatura do banho ( $T_{\text{bath}}$ ) e a temperatura instantânea do sistema,  $T(t)$ :

$$\frac{\partial T(t)}{\partial t} = \frac{1}{\tau} (T_{\text{bath}} - T(t)), \quad (3.3.7)$$

onde  $\tau$  é o tempo de relaxação térmica, que determina a intensidade do acoplamento ao banho.

O fator de reescalonamento correspondente é dado por:

$$\lambda = \sqrt{1 + \frac{\delta t}{\tau} \left( \frac{T_{\text{bath}}}{T(t)} - 1 \right)}, \quad (3.3.8)$$

sendo  $\delta t$  o passo de tempo da simulação. Para  $\tau \gg \delta t$ , o acoplamento é fraco e  $\lambda \approx 1$ . Para  $\tau \ll \delta t$ , o acoplamento é forte, levando a equilíbrios mais rápidos, porém com flutuações artificiais.

O termostato de Nosé-Hoover, por sua vez, permite uma amostragem correta do ensemble  $NVT$  e é adequado para etapas de produção [92, 93]. A diferença central em relação aos métodos anteriores está na forma como o banho térmico é tratado: ele é incorporado ao sistema via um grau de liberdade adicional no Hamiltoniano, com uma massa fictícia  $M_f$ , associada a uma energia potencial e cinética específicas. A transferência de energia entre sistema e banho térmico é mediada por este parâmetro.

Um valor elevado de  $M_f$  implica acoplamento fraco (variações térmicas lentas), enquanto valores baixos conduzem a forte acoplamento (variações rápidas). Ajustes adequados de  $M_f$  permitem controlar a velocidade de equilibração do sistema. Em simulações complexas, múltiplos termostatos podem ser acoplados a diferentes regiões para distribuir mais eficientemente a energia térmica. O comportamento oscilatório característico do Nosé-Hoover pode ser interpretado como uma “reação” do termostato às variações de temperatura: quando as partículas se movem muito rapidamente, o banho atua para desacelerá-las, e vice-versa.

Outra abordagem amplamente utilizada é o termostato de Langevin, baseado na adição de ruído estocástico à equação de movimento de Newton [92, 93]. O método introduz dois termos adicionais: um dissipativo (proporcional à velocidade) e outro aleatório (representando colisões com o meio), conforme:

$$m\ddot{\mathbf{r}} = -\nabla U(\mathbf{r}) - \gamma m\dot{\mathbf{r}} + \sqrt{2\gamma mk_B T} \mathbf{R}(t), \quad (3.3.9)$$

onde  $m$  é a massa da partícula,  $U(\mathbf{r})$  é o potencial,  $\gamma$  é o coeficiente de atrito e  $\mathbf{R}(t)$  é um ruído branco gaussiano de média nula e autocorrelação  $\delta(t - t')$ . O termo dissipativo remove

energia, enquanto o termo aleatório a reinjeta, assegurando o equilíbrio térmico no ensemble *NVT*.

As principais vantagens do termostato de Langevin são sua robustez, simplicidade e capacidade de controle local da temperatura. É particularmente eficaz na equilibrção de sistemas com solvente implícito, baixa dimensionalidade ou poucos graus de liberdade. Em sua forma discreta, métodos numéricos como o esquema VRORV (também conhecido como BAOAB) proporcionam estabilidade e precisão estatística.

Entretanto, o método também possui limitações. O uso de forças estocásticas quebra a conservação do momento linear, inviabilizando o estudo de fenômenos que dependem dessa conservação, como transporte coletivo e hidrodinâmica. Além disso, o atrito artificial pode distorcer escalas de tempo características, afetando propriedades dinâmicas como coeficientes de difusão. A escolha do coeficiente de atrito  $\gamma$  é crítica: valores muito elevados produzem dinâmica superamortecida; valores baixos podem não estabilizar a temperatura. Ademais, o modelo Langevin padrão ignora interações hidrodinâmicas entre partículas, tratando dissipação e flutuações de forma puramente local [93].

Cada método possui vantagens e desvantagens específicas, sendo fundamental uma leitura atenta e criteriosa para a seleção do termostato mais adequado à finalidade da simulação [92, 93].

### 3.3.5 Controle da pressão

O controle da pressão em simulações de dinâmica molecular é realizado por algoritmos denominados barostatos, que ajustam dinamicamente o volume da célula simulada para manter a pressão em um valor alvo. Entre os barostatos mais utilizados, destacam-se os métodos de Berendsen, Andersen e Parrinello–Rahman [112–115], cada um com características distintas em termos de precisão estatística, tipo de deformação da caixa de simulação e estabilidade numérica.

O barostato de Berendsen emprega uma abordagem de relaxação exponencial do volume em direção à pressão desejada, modificando suavemente as dimensões da caixa de simulação. Embora seja numericamente estável e eficiente na fase inicial da simulação, esse método não reproduz corretamente as flutuações de volume características do ensemble isobárico-isotérmico (NPT), suprimindo a compressibilidade isotérmica do sistema. Por essa

razão, seu uso é recomendado apenas na etapa de equilíbrio, visando ajustar o volume de forma rápida e sem oscilações significativas.

O barostato de Andersen introduz um pistão fictício acoplado ao volume do sistema, com graus de liberdade adicionais, analogamente ao formalismo de Nosé–Hoover para controle de temperatura. Essa abordagem pertence à classe de métodos de sistema estendido, permitindo a amostragem rigorosa do ensemble NPT. No entanto, sua formulação original considera apenas variações isotrópicas de volume, em que todas as direções espaciais são escaladas uniformemente. Apesar dessa limitação, o método de Andersen é adequado para líquidos e sistemas em que a simetria cúbica ou esférica é preservada.

O método mais abrangente e estatisticamente rigoroso é o barostato de Parrinello–Rahman, que estende a abordagem de Andersen ao permitir flutuações anisotrópicas da caixa de simulação [93, 114, 116, 117]. Nesse formalismo, o tensor da matriz da caixa pode evoluir livremente, possibilitando a deformação independente das direções espaciais. Essa característica é essencial para simulações de sólidos, materiais anisotrópicos, transições de fases estruturais e biomoléculas em condições de pressão variável. Apesar de sua precisão e generalidade, o barostato de Parrinello–Rahman apresenta maior complexidade numérica, sendo sensível à escolha de parâmetros como a massa fictícia do pistão e o tempo de acoplamento à pressão.

De modo geral, recomenda-se utilizar o barostato de Berendsen apenas durante a fase de equilibração inicial. Para as etapas de produção, o método de Andersen é apropriado quando se deseja preservar a simetria isotrópica, enquanto o de Parrinello–Rahman deve ser preferido em casos que exigem liberdade de deformação da caixa, como em simulações de sólidos ou interfaces sob tensão. A escolha criteriosa do barostato, em conjunto com um termostato compatível, é fundamental para garantir a validade estatística e a relevância física dos resultados obtidos [92, 93, 118].

### 3.3.6 Condições iniciais e etapas da simulação

Em uma simulação de DM as posições iniciais das partículas que constituem o sistema devem ser especificadas [92, 97, 111]. 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 [119, 120]. 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 energias potenciais altas. Assim, a primeira etapa da DM deve ser uma minimização da energia potencial do sistema [92, 97, 111].

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 equilíbrio. A equilíbrio é uma simulação que ocorre até que propriedades como pressão, temperatura e energia atinjam os valores correspondentes aos das condições termodinâmicas de interesse. Na equilíbrio, as propriedades termodinâmicas não se mantêm constantes, o tempo de equilíbrio é dependente do tipo e do tamanho do sistema.

A etapa de equilíbrio em dinâmica molecular é considerada concluída quando as propriedades termodinâmicas do sistema, como temperatura, pressão e energia potencial, atingem valores estáveis ao longo do tempo, oscilando apenas dentro das flutuações estatísticas esperadas para o *ensemble* adotado. Somente após essa fase é possível iniciar a coleta de trajetórias produtivas e calcular propriedades de interesse com maior confiabilidade [92, 93].

A DM é, em essência, um método determinístico: ao iniciar duas simulações com as mesmas condições iniciais—isto é, idênticas coordenadas e velocidades das partículas—, ambas deveriam, em princípio, gerar trajetórias idênticas. No entanto, na prática, o comportamento do sistema é fortemente influenciado pela sensibilidade às condições iniciais. Pequenas variações nas coordenadas ou velocidades podem levar a trajetórias completamente divergentes após alguns milhares de etapas de integração, devido à natureza caótica do sistema e à amplificação exponencial dessas diferenças [92, 93].

Além disso, a integração numérica das equações de movimento introduz pequenos erros de arredondamento a cada passo de tempo. Embora individuais, essas imprecisões se acumulam ao longo da simulação, podendo levar a desvios significativos na energia total e outras grandezas. O controle desses erros é essencial para garantir a estabilidade e a confiabilidade das simulações, sendo normalmente abordado por meio da escolha adequada de algoritmos de integração e do tamanho do passo de tempo (*timestep*),

equilibrando precisão e eficiência computacional.

Outro aspecto importante nas simulações de DM são as condições periódicas de contorno (PBC, do inglês *periodic boundary conditions*) [121]. 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 fronteira em uma simulação de DM. Isto é feito fazendo com que os átomos existentes dentro da caixa que será simulada sejam virtualmente replicados em todas as direções no espaço. Durante as simulações, os átomos interagem com as réplicas mais próximas, conforme a imagem mínima.

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 [121]. A soma, portanto, contaria com infinitos termos, sendo necessária a realização de aproximações. Para as interações de curto alcance (van der Waals), é realizado um truncamento do potencial em um raio de corte esférico [93]. 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 [92]. O truncamento abrupto do potencial pode comprometer fortemente os resultados das simulações. Este problema vem sendo tratado usando métodos derivados das somas de Ewald, como  $P^3M$  (Particle-Particle-Particle Mesh Ewald) e PME (Particle Mesh Ewald) que produzem trajetórias estáveis [92, 118, 122–124].

### 3.3.7 Fluxograma simplificado de uma simulação de DM

Um algoritmo típico de DM para simulação de proteínas, por exemplo, envolve diversos passos interdependentes, ajustados para garantir precisão e eficiência. Isso torna códigos modernos bastante complexos de entender. Contudo, é possível estudar a DM usando algoritmos mais simples. A Figura 3.5, apresenta um pseudo-algoritmo simplificado que captura o fluxo principal da DM.

```
function molecular_dynamics(r, num_steps, T, box_size)
    # Gerar velocidades iniciais (Maxwell-Boltzmann)
    v = initial_velocities(T)
    for i in 1:num_steps
        # Aplicar condições periódicas de contorno (PBCs) nas coordenadas
        r = apply_pbc(r, box_size)
        # Calcular forças (depende do campo de força)
        F = calc_forces(r, box_size)
        # Aplicar controle de temperatura e pressão
        v = temp_scaling(v, T)
        r = pressure_scaling(r, F)
        # Atualizar velocidades e coordenadas
        v += F ./ m * Δt
        r += v .* Δt
        # Aplica PBCs novamente após a atualização das coordenadas
        r = apply_pbc(r, box_size)
        # Registrar saídas periódicas
        if i % nst_output == 0
            write_output(r, v, F)
        end
    end
end
```

Figura 3.5: Pseudo-algoritmo simplificado para DM, ilustrando o fluxo principal de uma simulação.

O algoritmo segue as seguintes etapas fundamentais: 1) As velocidades iniciais são geradas pela função `initial_velocities`, utilizando a distribuição de Maxwell-Boltzmann [92]. 2) As forças entre átomos, calculadas por `calc_forces`, consideram interações ligadas (ligações, ângulos e diedros) e não-ligadas (Coulomb e Van der Waals). A implementação dessa função varia conforme o campo de força usado [92, 93]. 3) Métodos como Berendsen ou Nosé-Hoover são utilizados para controle de temperatura e pressão, garantindo condições termodinâmicas adequadas. 4) A integração das equações de movimento, com esquemas como Verlet, Velocity-Verlet ou leap-frog, atualiza velocidades e coordenadas. A precisão depende

do passo de tempo ( $\Delta t$ ) e do algoritmo de integração empregado [92]. 5) Saídas periódicas permitem monitorar o progresso da simulação, garantindo controle de qualidade e possibilitando análises posteriores.

A função `apply_pbc` garante que as coordenadas dos átomos permaneçam dentro da caixa de simulação, aplicando condições periódicas de contorno (PBCs). Se um átomo ultrapassa uma das faces da caixa, ele reaparece na face oposta, mantendo a continuidade espacial do sistema. Já a função `calc_forces` incorpora a **convenção da imagem mínima** (MIC, do inglês *minimum image convention*), que assegura o cálculo correto das distâncias entre os átomos considerando as PBCs. Nessa abordagem, cada partícula interage apenas com a imagem mais próxima das demais partículas do sistema, evitando interações artificiais com imagens distantes.

As PBCs são essenciais em simulações de DM, pois permitem representar um sistema extenso (como uma solução real) utilizando apenas uma célula unitária—neste caso, a caixa de simulação contendo a proteína e a solução de LI com água. Assim, a caixa original funciona como a célula unitária do sistema, replicada em todas as direções para formar um ambiente periódico e evitar efeitos de superfície artificiais.

É importante compreender as escalas de tempo dos fenômenos observados e verificar se essas escalas podem ser capturadas nas trajetórias geradas. Flutuações individuais de átomos ocorrem em frações de *ångstroms* (Å,  $10^{-10}$  m), com tempos na ordem de dezenas a centenas de femtossegundos ( $10^{-15}$  s). Movimentos de cadeias laterais de aminoácidos, compostas por vários átomos, acontecem em tempos na ordem de picossegundos ( $10^{-12}$  s), envolvendo deslocamentos da ordem de *ångstroms*. Mudanças de rotâmeros, que correspondem a rotações em torno de ligações, produzem deslocamentos maiores, na ordem de Å, em escalas de dezenas a centenas de picossegundos.

Considerando proteínas, regiões terminais e também alças, compostas por vários resíduos, podem apresentar movimentos mais amplos. Estes movimentos podem acontecer por meio de deslocamentos de até dezenas de *ångstroms*, em tempos que variam de nanossegundos ( $10^{-9}$  s) a microssegundos ( $10^{-6}$  s).

Os movimentos das maiores partes de proteínas, como domínios ou subunidades, ocorrem em escalas ainda mais lentas, alcançando dezenas de *ångstroms* em deslocamentos e, muitas vezes, escalas de segundos para acontecer. Transições conformacionais e interações

proteína-proteína frequentemente acontecem no mesmo intervalo de tempo que os movimentos de domínios. Já o enovelamento de proteínas ocorre tipicamente na escala de segundos, embora proteínas que dobram rapidamente possam levar apenas milissegundos [125, 126].

Para que as simulações sejam precisas, o passo de tempo deve ser menor do que os movimentos mais rápidos observados. Vibrações rápidas, como as de átomos de hidrogênio e ângulos rígidos, têm períodos vibracionais em torno de 10 femtossegundos. Por isso, passos de tempo típicos de 1 a 2 fs ( $10^{-15}$  s) são utilizados. Para atingir escalas de tempo biologicamente relevantes, como milissegundos, seriam necessários cerca de  $10^{12}$  passos de integração. Embora o poder computacional moderno esteja avançando, simular fenômenos como o dobramento completo de proteínas ainda não é rotineiramente viável devido às limitações no tempo total de simulação.

Diversos textos tratam de maneira aprofundada e completa os mais diversos aspectos das simulações por dinâmica molecular [92, 111, 127, 128], 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

Neste capítulo, serão apresentados os principais objetivos específicos desta tese, bem como os trabalhos científicos produzidos ao longo do doutorado. Os capítulos dedicados à discussão dos resultados são baseados nas publicações realizadas durante esse período, devidamente adaptadas para o formato tradicional de tese.

### 4.1 Objetivos

Este trabalho busca caracterizar a estrutura de solvatação de uma proteína modelo (ubiquitina) em soluções aquosas de líquidos iônicos imidazólicos, utilizando simulações de dinâmica molecular e a teoria de soluções de Kirkwood-Buff. Para alcançar esse objetivo geral, foram definidos os seguintes objetivos específicos:

- Compreender os princípios da solvatação da ubiquitina por líquidos iônicos, calculando funções de distribuição de mínima distância.
- Aplicar a teoria de Kirkwood-Buff para obter os parâmetros de solvatação preferencial e analisar a afinidade dos líquidos iônicos pela proteína.
- Realizar simulações em diferentes concentrações e com várias combinações de íons, avaliando tendências relacionadas à concentração e à natureza dos íons.
- Investigar efeitos aditivos e não aditivos na solvatação da proteína em misturas de líquidos iônicos com ânions variados, mantendo o cátion fixo.
- Analisar o impacto do aumento da cadeia carbônica do cátion na interação total do líquido iônico com a proteína.

## 4.2 Trabalhos publicados ou em publicação/preparação.

Os seguintes trabalhos foram publicados, submetidos ou estão em preparação no contexto desta tese:

1. **Piccoli, V.; Martínez, L.** *Correlated counterion effects on the solvation of proteins by ionic liquids*. J. Mol. Liq. **320**, 114347, 2020. Este estudo investiga os efeitos correlacionados de contraíons na solvatação de proteínas por líquidos iônicos, destacando como as interações eletrostáticas e específicas entre íons influenciam a estabilidade e a estrutura proteica.
2. **Piccoli, V.; Martínez, L.** *Ionic liquid solvation of proteins in native and denatured states*. J. Mol. Liq. **363**, 119953, 2022. Neste trabalho, exploramos a solvatação de proteínas em seus estados nativo e desnaturado por líquidos iônicos, elucidando os mecanismos moleculares que governam a estabilização ou desestabilização de proteínas em diferentes condições.
3. **Piccoli, V.; Martínez, L.** *Competitive Effects of Anions on Protein Solvation by Aqueous Ionic Liquids*. J. Phys. Chem. B, 2024. Este estudo recente aborda os efeitos competitivos de ânions na solvatação de proteínas por líquidos iônicos aquosos.
4. **Nakasu, P.; Piccoli, V.; Ovejero-Pérez, A.; Kumar, P.; Al Ghatta, A.; Melanie, S.; Polesca, C.; Martínez, L.; Hallett, J.** *Fractionation of Squid Pens with Ionic Liquids – An upgraded  $\beta$ -Chitin and Shellfish Protein Production*. ACS Sustainable Chemistry & Engineering (doi:10.1021/acssuschemeng.4c04217)
5. **Piccoli, V.; Martínez, L.** *Cation Hydrophobicity Effects on Protein Solvation in Aqueous Ionic Liquids*. Submetido ao *Journal of Physical Chemistry B* (Manuscript ID:jp-2025-00779w). Status: Aceito para publicação em 19 de Maio de 2025.
6. **Piccoli, V.; Pádua, A.; Martínez, L.** *Exploring Ibuprofen Solvation in Bio-Compatible Deep Eutectic Solvents: Insights from Polarizable Force Fields and Kirkwood-Buffer Theory*. Em preparação. Este trabalho, desenvolvido durante o estágio de BEPE na França, explora a solvatação do ibuprofeno em solventes eutéticos profundos biocompatíveis, utilizando campos de força polarizáveis e a teoria de Kirkwood-Buffer.

**Trabalhos não relacionados à tese**

1. **Pereira, A. F.; Piccoli, V.; Martínez, L.** *Trifluoroethanol direct interactions with protein backbones destabilize alpha-helices*. J. Mol. Liq. **365**, 120209, 2022. Este estudo investiga as interações diretas do trifluoretano com cadeias principais de proteínas, elucidando seu efeito desestabilizador em hélices alfa.
2. **Pereira, A. F.; Piccoli, V.; Martínez, L.** *Amyloid beta peptide (A $\beta$ ) clusterization in glucose solutions*. Em preparação.

## Capítulo 5

# Metodologia Geral

Este capítulo apresenta a metodologia geral utilizada ao longo dos estudos descritos nesta tese. Inicialmente, serão apresentados os sistemas simulados, assim como o arranjo empregado na montagem das configurações iniciais. Em seguida, uma breve discussão sobre a Ubiquitina será realizada no intuito de apresentar a composição química dos resíduos expostos ao solvente, assim como a distribuição de carga na superfície proteica. O formalismo teórico utilizado será apresentado para discutir a análise da estrutura de solvatação e a aplicação da teoria de soluções de Kirkwood-Buff para a investigação do aspecto termodinâmico da solvatação da Ubiquitina em LIs.

## 5.1 Sistemas simulados

### 5.1.1 Breve discussão sobre a ubiquitina

A ubiquitina é uma proteína globular de 8,6 kDa amplamente distribuída em células eucarióticas, desempenhando funções regulatórias essenciais no citoplasma, núcleo e membrana plasmática [129, 130]. Descoberta por Gideon Goldstein em 1975, foi inicialmente chamada de “polipeptídeo imunopoiético ubíquo” (UBIP) [129]. Aaron Ciechanover, Avram Herskho e Irwin Rose descreveram os mecanismos e componentes da via de ubiquitinação, trabalho que lhes rendeu o Prêmio Nobel de Química em 2004 [131–133]. Sua função primária é marcar proteínas mal enoveladas ou danificadas para degradação pelo proteassoma, mas também pode alterar localização celular, conformação estrutural e função biológica das proteínas-alvo.

Composta por 76 resíduos de aminoácidos, a ubiquitina é uma das proteínas mais conservadas evolutivamente, apresentando, por exemplo, sequência muito parecida entre humanos e invertebrados (moluscos como *Aplysia*) [130, 134]. Seus sete resíduos de lisina

(Lys6, Lys11, Lys27, Lys29, Lys33, Lys48, Lys63) e o resíduo terminal Gly76 desempenham papéis cruciais na formação de cadeias de poli ubiquitina por ligação covalente a substratos-alvo [130, 134].

A estrutura secundária da ubiquitina inclui três hélices ( $\alpha$ -hélice e hélice 310, totalizando 23,7%) e uma folha- $\beta$  formada por cinco fitas  $\beta$ , conferindo alta estabilidade estrutural, mesmo sob condições adversas de pH e temperatura. Estudos de RMN demonstraram que a proteína mantém sua conformação entre pH 1,18 e 8,48 e temperaturas de até 80°C [135]. A estabilidade estrutural da ubiquitina deve-se à sua rede de ligações de hidrogênio intrínsecas, mesmo na ausência de pontes de dissulfeto, íons metálicos coordenados ou cofatores. Essa característica permite que a proteína se enovele espontaneamente após desnaturação *in vitro* [136].

Uma região hidrofóbica centrada no resíduo Ile44, composta por Leu8, Ile44 e Val70, associada a resíduos carregados positivamente (Lys6, Arg42, Lys48), facilita interações não covalentes com domínios de reconhecimento de ubiquitina, que assim desempenha também papéis essenciais em vias de sinalização celular [137].

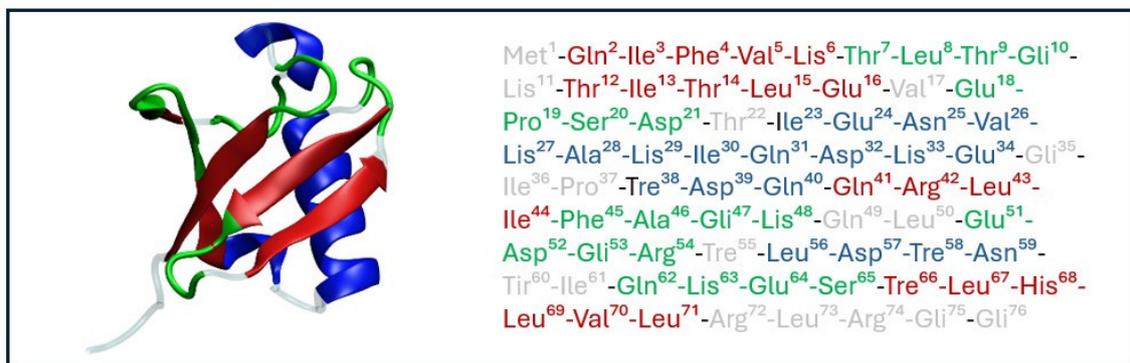


Figura 5.1: Estrutura tridimensional da ubiquitina com seus elementos de estrutura secundária destacados em cores. Os segmentos na sequência primária (à direita) representam as regiões específicas das fitas  $\beta$  (em vermelho) e das hélices  $\alpha$  (em azul). As regiões de loop e outras partes não estruturadas estão marcadas em verde e cinza, respectivamente, na estrutura 3D.

Assim, a ubiquitina é um modelo de estabilidade estrutural e funcionalidade versátil, desempenhando papéis fundamentais em diversos processos celulares. Sua conservação evolutiva reflete sua importância biológica, enquanto suas propriedades físico-químicas a tornam uma ferramenta valiosa em estudos bioquímicos e biotecnológicos.

### 5.1.2 Composição dos sistemas estudados

Os LIs utilizados são combinações do cátion 1-etil-3-metilimidazólio ( $[\text{EMIM}]^+$ ) com os ânions diacianamida ( $[\text{DCA}]^-$ ), tetrafluoroborato ( $[\text{BF}_4]^-$ ), nitrato ( $[\text{NO}_3]^-$ ) e cloreto ( $[\text{Cl}]^-$ ). As estruturas químicas dos íons estão representadas na Figura 5.2. Os cátions imidazólicos foram escolhidos devido ao seu amplo uso em sistemas com ou sem biomoléculas, e por possuírem cadeias de diferentes comprimentos, proporcionando uma variabilidade útil para estudar o efeito de características dos cátions maiores na solvatação da ubiquitina. Os ânions, por sua vez, apresentam características químicas distintas, incluindo diferentes átomos, capacidades de aceitar ligações de hidrogênio e variadas estruturas, sendo predominantemente inorgânicos.

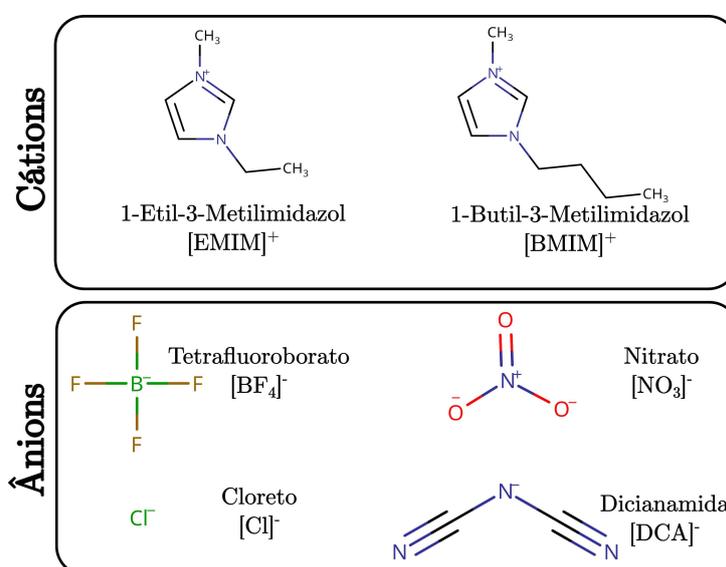


Figura 5.2: Estruturas químicas dos cátions e ânions dos líquidos iônicos utilizados neste estudo: 1-etil-3-metilimidazólio ( $[\text{EMIM}]^+$ ) e 1-butil-3-metilimidazólio ( $[\text{BMIM}]^+$ ) como cátions; e diacianamida ( $[\text{DCA}]^-$ ), tetrafluoroborato ( $[\text{BF}_4]^-$ ), nitrato ( $[\text{NO}_3]^-$ ) e cloreto ( $[\text{Cl}]^-$ ) como ânions. Isso mostra a diversidade estrutural e funcional dos íons, essenciais para estudar o efeito da variação de cadeia carbônica do cátion e das propriedades específicas dos ânions na solvatação da ubiquitina.

Na configuração da caixa de dinâmica molecular, a proteína é posicionada no centro e, em seguida, solvatada por soluções aquosas de LIs em concentrações de 0,5, 1,0, 1,5, 2,0, 2,5 e 3,0 mol L<sup>-1</sup>. Dois arranjos básicos de sistemas foram montados para realizar as simulações:

- O primeiro tipo consiste na simulação da proteína em uma solução aquosa contendo apenas um tipo de LI; por exemplo, ubiquitina, água e  $[\text{EMIM}][\text{NO}_3]$ . Estes sistemas servem como base para a compreensão de sistemas mais complexos, nos quais vários

parâmetros são alterados simultaneamente, e ajudam a entender a mudança na solvatação em função da concentração do LI e da identidade química do ânion.

- O segundo tipo de sistema introduz uma variação intencional para estudar efeitos aditivos e não aditivos. Esses sistemas contêm a proteína, água e uma mistura de LIs com o mesmo cátion, por exemplo, [EMIM][Cl] + [EMIM][NO<sub>3</sub>]. Eles foram projetados para manter a mesma concentração total de LIs que os sistemas contendo apenas um LI, permitindo uma comparação direta. Por exemplo, em um sistema com 0,5 mol L<sup>-1</sup> de [EMIM][DCA], há 0,5 mol L<sup>-1</sup> de [EMIM]<sup>+</sup> e [DCA]<sup>-</sup>. Em um sistema misto de, por exemplo, 0,5 mol L<sup>-1</sup> de [EMIM][DCA] + [EMIM][Cl], há 0,5 mol L<sup>-1</sup> de [EMIM]<sup>+</sup> e 0,25 mol L<sup>-1</sup> de [DCA]<sup>-</sup> e [Cl]<sup>-</sup>.

## 5.2 Análise de solvatação

Esta seção apresenta os métodos e teorias fundamentais utilizados para a análise de solvatação, abrangendo tanto aspectos estruturais quanto termodinâmicos dos sistemas simulados. As abordagens descritas fornecem as bases para investigar como os solventes interagem com as proteínas e como essas interações afetam a dinâmica molecular e as propriedades macroscópicas dos sistemas estudados.

### 5.2.1 Funções de distribuição

Funções de distribuição radial ( $g(r)$ ) baseiam-se em um único sítio de referência, tipicamente o centro de massa, ignorando o formato das espécies. Isso limita sua precisão ao representar as interações entre moléculas da solução e a superfície do soluto [90, 118]. Para solutos com forma não esférica, como mostrado na Figura 5.3A, a  $g(r)$  não descreve adequadamente a camada de solvatação, ao partir do volume usado para o cálculo da densidade das partículas de interesse, incluindo o próprio soluto, resultando em uma distribuição de utilidade física limitada para analisar as interações soluto-solvente.

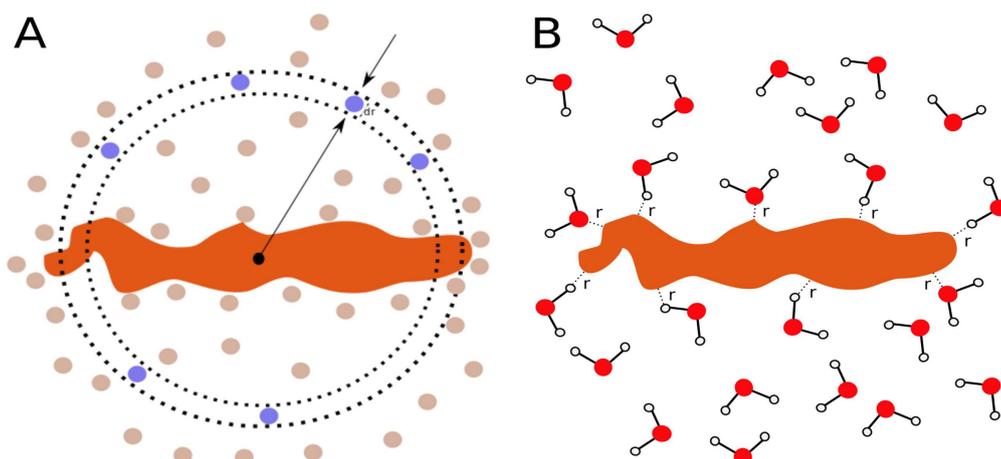


Figura 5.3: A - Limitações do cálculo de  $g(r)$  para solutos com forma não esférica. B — Conceito básico das funções de distribuição de mínima distância, que, ao considerar as mínimas distâncias  $r$ , permite a obtenção de uma distribuição do solvente que mantém significado estrutural independentemente do formato das espécies. Figura reproduzida com autorização do autor a partir de [89].

As funções de distribuição de mínima distância (MDDF, do inglês *minimum-distance distribution function*) consideram o formato das moléculas do soluto e do solvente [90], calculando a distribuição do solvente com base na menor distância entre quaisquer átomos do soluto e do solvente. A Figura 5.3B ilustra moléculas de água posicionadas a uma mínima distância  $r$  da superfície proteica. Contando essas mínimas distâncias em torno da proteína, obtém-se uma função de distribuição cuja interpretação molecular se mantém útil, independentemente do formato do soluto e do solvente. A MDDF é definida como:

$$g_{us}^{md}(r) \equiv \frac{n_{us}(r)}{n_{us}^*(r)} \quad (5.2.1)$$

onde  $n_{us}(r)$  representa a densidade numérica média de átomos do solvente ( $s$ ) a uma distância mínima  $r$  do soluto ( $u$ ), e  $n_{us}^*(r)$  corresponde à densidade numérica média esperada na ausência de interação entre soluto e solvente [90]. Se o soluto e o solvente possuírem apenas um átomo, a MDDF se reduz à  $g(r)$ , pois  $n_{us}(r)$  descreve a densidade de moléculas do solvente a uma distância  $r$ , enquanto  $n_{us}^*(r)$  se torna constante, igual à densidade do solvente no *bulk*. As densidades  $n_{us}(r)$  e  $n_{us}^*(r)$  diferem da densidade *bulk* por representarem a contagem de qualquer átomo do solvente a cada distância [90].

A Figura 5.4 contém uma comparação (que serve como exemplo) entre uma RDF

( $g(r)$ ) e uma MDDF ( $g_{md}(r)$ ) para um mesmo sistema. Nela, observa-se a distribuição do ânion ( $[DCA]^-$ ) de LI [EMIM][DCA] 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 representa melhor a distribuição do ânion em torno da proteína.

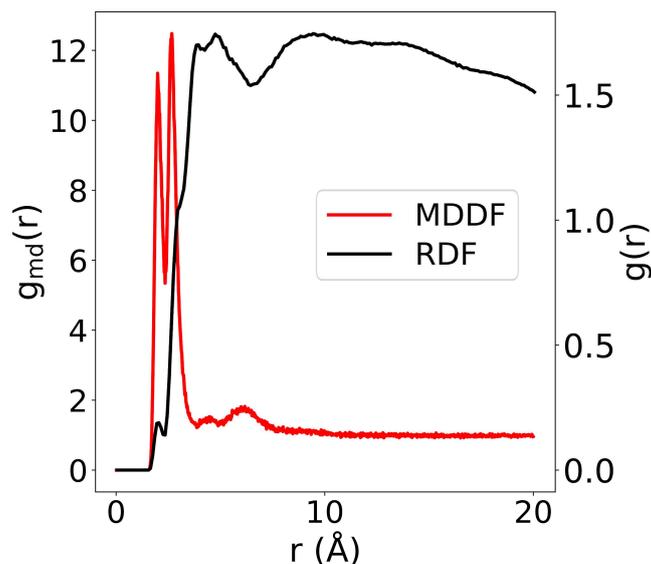


Figura 5.4: Representação das funções de distribuição radial e de mínima distância para o ânion  $[DCA]^-$  ao redor da ubiquitina na conformação nativa. As funções de distribuição foram calculadas a partir da mesma simulação. Podemos notar que RDF apresenta um comportamento que é muito difícil de interpretar. Contudo, a MDDF apresenta picos com formatos que possibilitam uma descrição mais intuitiva do sistema. Figura reproduzida com autorização do autor a partir de [89].

Um aspecto muito importante das MDDFs, e fundamental para as suas potencialidades em avaliar a estrutura de solvatação, diz respeito à decomposição das MDDFs em termos de contribuição de resíduos ou átomos do soluto e de átomos do solvente. Para compreender a decomposição da MDDF, considere uma molécula de soluto e um único componente do solvente. Aqui, definimos  $r$  como a menor distância entre quaisquer átomos do soluto e do solvente. A função  $g^{md}(r)$  é definida em termos da densidade de distâncias mínimas em  $r$ ,  $n(r)$ , e da densidade de distâncias mínimas em  $r$  na ausência de interações entre soluto e solvente,  $n^*(r)$ .

Seja  $i$  um tipo específico de átomo do solvente ( $i = 1, \dots, N$ , onde  $N$  representa

o número total de átomos na molécula de solvente). A densidade numérica de átomos  $i$  a uma distância  $r$  do soluto é indicada por  $n_i(r)$ , de modo que  $n_i(r)dV(r)$  representa o número de átomos do tipo  $i$  em um volume entre  $r$  e  $r + dr$ . A probabilidade de que um átomo  $i$  no elemento de volume  $dV(r)$  seja o mais próximo do soluto é representada por  $w_i(r)$ . Esta probabilidade depende tanto da forma do volume definido quanto da estrutura do solvente [90].

Desta forma, a contribuição dos átomos do tipo  $i$  para a densidade de distâncias mínimas em  $r$  é dada por  $n_i(r)w_i(r)$ . A densidade total de distâncias mínimas em  $r$  é, então, expressa por:

$$n(r) = \sum_{i=1}^N n_i(r)w_i(r). \quad (5.2.2)$$

A normalização da função  $g^{\text{md}}(r)$  depende do cálculo da densidade  $n^*(r)$ , sendo obtida considerando a ausência de interações entre soluto e solvente. Para cada valor de  $r$ , o volume associado depende da estrutura do soluto. Podemos imaginar um modelo teórico de uma molécula de soluto “fantasma”, imersa em um solvente com propriedades de densidade *bulk*. Esse modelo auxilia na definição da forma das superfícies de distâncias mínimas para cada valor de  $r$  [90].

Caso o solvente apresentasse densidade uniforme e não houvesse interações entre soluto e solvente, a densidade de átomos  $i$  em qualquer distância  $r$  seria igual à densidade *bulk*  $\rho_s$ . A probabilidade  $w_i^*(r)$  de que um átomo  $i$  seja o mais próximo do soluto dependeria apenas da forma da superfície definida pelo soluto à distância  $r$  [90]. Assim, a densidade  $n^*(r)$  no *bulk* é dada por:

$$n^*(r) = \rho_s \sum_{i=1}^N w_i^*(r) \quad (5.2.3)$$

A densidade  $n^*(r)$  geralmente é maior que a densidade *bulk* do solvente, especialmente para valores baixos de  $r$ . Isso ocorre porque  $w_i^*(r) = 1$  para todos os átomos em distâncias muito curtas.

Portanto, a MDDF  $g^{\text{md}}(r)$  pode ser escrita como:

$$g^{\text{md}}(r) = \sum_{i=1}^N \frac{n_i(r)}{\rho_{\text{bulk}}} \left( \frac{w_i(r)}{\sum_{i=1}^N w_i^*(r)} \right), \quad (5.2.4)$$

onde a expressão permite decompor  $g^{\text{md}}(r)$  em termos de contribuições atômicas, considerando tanto a densidade relativa de cada tipo de átomo em relação à densidade *bulk* quanto os pesos orientacionais definidos por  $w_i(r)$ . Esta abordagem é particularmente útil para analisar as contribuições de diferentes átomos do solvente na interação com o soluto em função da distância  $r$  [90].

Em termos práticos, o cálculo da MDDF é realizado utilizando a trajetória obtida por DM. Durante o processo, identifica-se o átomo de cada molécula do solvente que está mais próximo da proteína. Esse procedimento é repetido para todas as moléculas e íons de todos os componentes do solvente, resultando, por exemplo, na decomposição da MDDF em termos dos átomos que estão mais próximos do soluto no sistema.

### 5.2.2 Teoria de Soluções de Kirkwood-Buff

A teoria de soluções de Kirkwood-Buff (KB) estabelece uma conexão entre as propriedades macroscópicas e microscópicas das soluções [88, 138, 139]. Essa teoria permite o cálculo de propriedades termodinâmicas — como coeficientes de atividade, volumes molares parciais, parâmetros de solvatação preferencial, etc. — por meio de integrais das funções de correlação de pares ao longo do espaço, definidas como integrais de KB (quantidades centrais na teoria de KB). As funções de correlação de pares, representadas aqui pelas MDDFs, podem ser determinadas via métodos computacionais, como simulações de DM [139, 140].

A integral de KB de um solvente  $i$  com respeito a um soluto  $j$  ( $G_{ij}$ ) tem como definição clássica:

$$G_{ij} = G_{ji} = 4\pi \int_0^{\infty} [g_{ij}^{\mu,V,T}(r) - 1] r^2 dr \quad (5.2.5)$$

onde  $g_{ij}^{\mu,V,T}(r)$  é a função de distribuição radial no ensemble grande-canônico entre duas espécies  $i$  e  $j$ , e  $r$  é a distância a partir da partícula de referência.

A teoria de KB permite o cálculo de propriedades termodinâmicas diretamente das funções de distribuição, muitas das quais são obtidas via simulações computacionais [138]. O ensemble  $\mu, V, T$  é o ensemble onde a integral de KB foi definida, sendo adequado para

sistemas abertos, onde  $\mu$  corresponde ao potencial químico, indicando variação no número de partículas [138, 139]. Assim, considerando apenas a teoria desenvolvida, a integral de KB só seria definida e, portanto, possível de ser calculada no ensemble  $\mu, V, T$ , que representa sistemas abertos.

Simulações computacionais, em geral, utilizam ensembles que representam sistemas fechados, como, por exemplo, os ensembles canônico  $(N, V, T)$  e isotérmico-isobático  $(N, P, T)$ , onde  $N$  é o número de partículas,  $V$ ,  $P$  e  $T$  são o volume, pressão e temperatura do sistema, respectivamente. No entanto, em sistemas suficientemente grandes, as propriedades termodinâmicas obtidas convergem para valores equivalentes, independentemente do ensemble escolhido, permitindo o uso de funções de distribuição calculadas em ensembles fechados na teoria de soluções de KB [86].

$G_{ij}$  reflete a afinidade entre o soluto e as moléculas do solvente [139–142]. Ao multiplicar  $G_{ij}$  pela concentração da espécie  $i$ , é possível determinar o número efetivo de moléculas que se acumula ou é excluído ao redor do soluto. Assim,  $\rho_i G_{ij}$  representa o número médio de partículas  $i$  menos o número observado em uma distribuição homogênea em toda a região de integração (Equação 5.2.5) ao redor de uma partícula  $j$ , onde  $\rho_i$  é a densidade da partícula  $i$ .

As integrais de KB podem ser calculadas a partir da densidade numérica média de átomos do solvente que estão a uma mínima distância  $r$  do soluto ( $n_{us}(r)$ ) e da densidade numérica média esperada na ausência de interação entre soluto e solvente ( $n_{us}^*(r)$ ):

$$G_{ij} = \frac{1}{\rho_s} \int_0^{\infty} [n_{us}(r) - n_{us}^*(r)] S(r) dr = \frac{1}{\rho_s} \int_0^{\infty} [g_{us}(r) - 1] S(r) dr \quad (5.2.6)$$

onde  $S(r)$  é a área superficial associada à mínima distância  $r$  em relação a qualquer átomo do soluto, como descrito nas MDDFs [90].  $g_{us}(r)$  corresponde à MDDF. Ao integrar as funções de distribuição em todo o volume, obtemos a correlação geral entre duas espécies, representada pela soma dos desvios em relação à distribuição ideal. Em termos simples, as integrais de KB representam o volume efetivo que um componente ocupa na solução em comparação ao volume esperado na ausência de interações específicas — isto é, uma propriedade de excesso.

### 5.2.3 Solvatação Preferencial

O parâmetro de solvatação preferencial (comumente representado por  $\Gamma$ ) é uma quantidade termodinâmica que ajuda a compreender a competição entre componentes do solvente de uma solução pela solvatação de solutos complexos, como proteínas [53, 143]. A quantidade definida como  $\Gamma$  pode ser mensurada experimentalmente por diálise no equilíbrio, ou por centrifugação analítica [53]. De acordo com Timasheff em seus trabalhos [53, 144, 145], o parâmetro de solvatação preferencial é definido por:

$$\Gamma_{\mu_w, \mu_c} \equiv \left( \frac{\partial m_c}{\partial m_p} \right)_{T, \mu_w, \mu_c} \quad (5.2.7)$$

onde  $c$ ,  $w$  e  $p$  correspondem, respectivamente, ao cossolvente, à água e à proteína.  $m_p$  e  $m_c$  representam, respectivamente, a molalidade da proteína e do cossolvente.  $\mu_w$  e  $\mu_c$  são, respectivamente, o potencial químico da água e do cossolvente. Assim, a equação é a mudança na concentração de cossolvente causada pela adição da proteína em equilíbrio químico, assim, com potencial químico constante para a água e o cossolvente.

A diálise no equilíbrio mede a variação na concentração de cossolventes resultante da adição de proteína à solução. Experimentalmente, esse processo é indireto: em misturas de cossolvente e biomolécula, distinguir a fração de cossolvente interagindo efetivamente com o soluto é difícil. No entanto, ao dialisar o cossolvente livre até que sua concentração atinja equilíbrio através da membrana, possibilita-se medir a diferença de concentração na presença e ausência da proteína [53]. Com a concentração conhecida, pode-se então calcular a quantidade de cossolvente associada à biomolécula. Assim,  $\Gamma_{\mu_1, \mu_3}$  é interpretado como uma medida de acúmulo ou exclusão do cossolvente na vizinhança do soluto, dada experimentalmente em função das molalidades do cossolvente em duas soluções em equilíbrio, como mostra a Figura 5.5 [146].

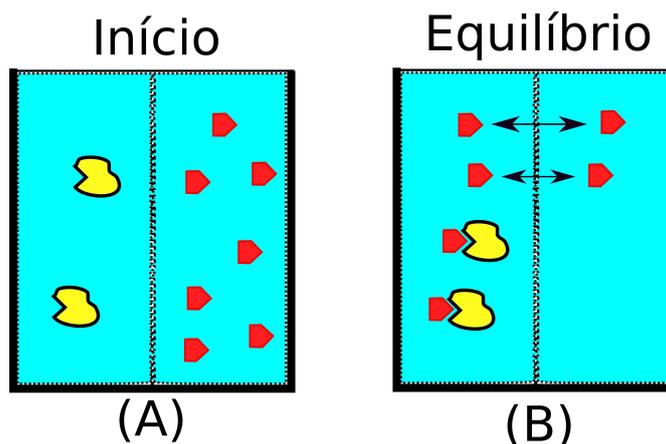


Figura 5.5: Esquema simplificado de um experimento de diálise no equilíbrio. O fundo azul representa a água, os polígonos vermelhos são as moléculas de cossolvente, e as estruturas amarelas representam as biomoléculas. Figura reproduzida com autorização do autor a partir de [89].

$\Gamma$  é uma quantidade física, que pode ser obtida por meio de teorias gerais de soluções, como a teoria de KB, pela relação na Equação 5.2.8 [139, 142].

$$\Gamma_{cp} = \rho_c(G_{cp} - G_{wp}). \quad (5.2.8)$$

onde  $G_{wp}$  e  $G_{cp}$  são as integrais de KB,  $\rho_c$  é a concentração do cossolvente no *bulk*,  $p$  representa o soluto e  $w$  o solvente principal (geralmente água). Componentes com  $\Gamma_{cp}$  positivo geralmente solvatam preferencialmente o soluto com relação à água.

Quando há um equilíbrio químico entre as conformações desnaturada e nativa de uma proteína, é possível estudar como um cossolvente interage com cada estado. Um cossolvente com parâmetro de solvatação preferencial maior no estado desnaturado ( $\Gamma_{cp}^D$ ) irá deslocar o equilíbrio em favor da conformação desnaturada, caracterizando-o como um desnaturante [8, 58]. Equivalentemente, se  $\Gamma_{cp} < 0$ , a proteína é preferencialmente hidratada; quando a solvatação é avaliada em termos da água em relação ao soluto, há a utilização do parâmetro preferencial de hidratação  $\Gamma_{wp}$ .

A interação de uma proteína com um cossolvente pode ser analisada a partir de um sistema com dois estados de referência: a proteína em água pura e a proteína em uma solução de água com cossolvente em uma determinada concentração. O parâmetro  $\Gamma$ , que descreve a solvatação preferencial, está relacionado à energia livre de transferência  $\Delta\mu_p^{tr}$ , definida como a energia livre necessária para transferir a proteína da água pura para uma solução contendo

cossolvente. A adição de um cossolvente a uma solução aquosa contendo uma proteína altera o potencial químico da proteína devido às interações com o cossolvente [57, 147]. Esta mudança reflete a variação nas interações da proteína com o solvente ao ser transferida da água pura para uma solução com cossolvente e pode ser expressa por:

$$\Delta\mu_p^{tr} = \mu_p(\text{cossolvente}) - \mu_p(\text{água}). \quad (5.2.9)$$

Onde  $\mu_i$  é o potencial químico do componente  $i$ , definido como  $\mu_i = \mu_i^\circ(P, T) + RT \ln a_i$ . Aqui,  $a_i$  representa a atividade do componente  $i$ , dada por  $a_i = m_i \gamma_i$ , em que  $m_i$  é a molalidade e  $\gamma_i$  o coeficiente de atividade. A adição de um cossolvente ao sistema aquoso induz uma perturbação mútua no potencial químico da proteína e do cossolvente [144, 145]. Essa perturbação é descrita pela igualdade das derivadas parciais [144, 145, 148]:

$$\left( \frac{\partial \mu_p}{\partial m_c} \right)_{T, P, m_p} = \left( \frac{\partial \mu_c}{\partial m_p} \right)_{T, P, m_c}. \quad (5.2.10)$$

A Equação 5.2.10 descreve a variação do potencial químico das espécies em função da concentração do cossolvente. Por exemplo, se  $\left( \frac{\partial \mu_c}{\partial m_p} \right)_{T, P, m_p}$  diferir de zero, a adição da proteína ao sistema contendo cossolvente perturba o equilíbrio termodinâmico. Para restaurar o equilíbrio, é necessário ajustar a concentração do solvente nas proximidades da proteína, seja por adição ou remoção de cossolvente, ou água.

O potencial químico da proteína,  $\mu_p$ , depende da concentração do cossolvente,  $m_c$ , no sistema. Assim,  $\mu_p$  pode ser expresso como uma derivada total em relação à molalidade do cossolvente:

$$d\mu_p = \left( \frac{\partial \mu_p}{\partial m_c} \right)_{T, P, m_p} dm_c. \quad (5.2.11)$$

A Equação 5.2.10 é amplamente utilizada para explicar a energia livre de transferência, conforme discutido em [144, 145, 148]. Integrando a Equação 5.2.11 de 0 até a molalidade do cossolvente ( $m_c$ ) na solução de destino, obtém-se:

$$\Delta\mu_p^{tr} = \int_0^{m_c} \left( \frac{\partial \mu_p}{\partial m_c} \right)_{m_p} dm_c. \quad (5.2.12)$$

As Equações 5.2.12 e 5.2.7 podem ser correlacionadas para estabelecer uma relação entre o parâmetro de solvatação preferencial e a energia livre de transferência.

A **regra cíclica** estabelece uma relação geral entre derivadas parciais de três variáveis relacionadas,  $x$ ,  $y$  e  $z$ , em uma função implícita  $f(x,y,z) = 0$ . Temos:

$$\left(\frac{\partial x}{\partial y}\right)_z \cdot \left(\frac{\partial y}{\partial z}\right)_x \cdot \left(\frac{\partial z}{\partial x}\right)_y = -1. \quad (5.2.13)$$

Considerando esta expressão, sendo uma consequência da dependência funcional das três variáveis, e as variáveis  $u_c$ ,  $m_c$  e  $m_p$ , que estão relacionadas por uma função implícita  $f(u_c, m_c, m_p) = 0$ , a aplicação da regra cíclica para  $\left(\frac{\partial u_c}{\partial m_p}\right)_{m_c}$  resulta em:

$$\left(\frac{\partial u_c}{\partial m_p}\right)_{m_c} = - \left(\frac{\partial m_c}{\partial m_p}\right)_{u_c} \left(\frac{\partial u_c}{\partial m_c}\right)_{m_p}. \quad (5.2.14)$$

A Equação 5.2.14 pode, então, ser relacionada com a integral construída em 5.2.12 por meio da relação apresentada na Equação 5.2.10. Portanto, substituindo 5.2.10 em 5.2.14, obtém-se:

$$\left(\frac{\partial u_p}{\partial m_c}\right)_{T,P,m_p} = - \left(\frac{\partial m_c}{\partial m_p}\right)_{\mu_c} \left(\frac{\partial \mu_c}{\partial m_c}\right)_{m_p}. \quad (5.2.15)$$

Com esta relação, a derivada  $\left(\frac{\partial u_p}{\partial m_c}\right)_{T,P,m_p}$  pode ser substituída na Equação 5.2.12, resultando em:

$$\Delta\mu_p^{tr} = - \int_0^{m_c} \left(\frac{\partial m_c}{\partial m_p}\right)_{T,P,\mu_c} \left(\frac{\partial \mu_c}{\partial m_c}\right)_{T,P,m_p} dm_c. \quad (5.2.16)$$

Esta é uma forma expandida para a variação da energia livre de transferência.

A definição do parâmetro preferencial de solvatação pode ser introduzida na Equação 5.2.16 usando a aproximação proposta por Stigter em 1960 [148, 149], onde:

$$\left(\frac{\partial m_c}{\partial m_p}\right)_{T,P,\mu_c} = \left(\frac{\partial m_c}{\partial m_p}\right)_{T,\mu_w,\mu_c} = \Gamma_{cp}. \quad (5.2.17)$$

Logo, a expressão pode ser usada para introduzir explicitamente o parâmetro preferencial

de solvatação na Equação 5.2.16. Além disso, a derivada  $\left(\frac{\partial \mu_c}{\partial m_c}\right)_{T,P,m_p}$  pode ser interpretada como uma medida de como o potencial químico do cossolvente responde a uma variação na molalidade/concentração do cossolvente. Tal derivada pode ser usada para simplificar ainda mais a Equação 5.2.16. O potencial químico do cossolvente  $\mu_c$  pode ser dado por:

$$\mu_c = \mu_c^\circ + RT \ln a_c, \quad (5.2.18)$$

onde  $\mu_c^\circ$  é o potencial químico padrão do cossolvente e  $a_c$  o coeficiente de atividade para o cossolvente.

A derivada  $\left(\frac{\partial \mu_c}{\partial m_c}\right)_{T,P,m_p}$  é obtida derivando a Equação 5.2.18 em relação à molalidade,  $m_c$ , do cossolvente:

$$\left(\frac{\partial \mu_c}{\partial m_c}\right)_{T,P,m_p} = \left(\frac{\partial \mu_c^\circ}{\partial m_c}\right) + RT \left(\frac{\partial (\ln a_c)}{\partial m_c}\right), \quad (5.2.19)$$

onde o potencial químico padrão,  $\mu_c^\circ$ , corresponde ao potencial químico de uma substância em uma fase específica a uma temperatura  $T$  e pressão  $P$  definidas, não variando com a molalidade do cossolvente. Portanto,  $\left(\frac{\partial \mu_c^\circ}{\partial m_c}\right) = 0$ . Além disso, considerando que a atividade do cossolvente é dada por  $a_c = m_c \gamma_c$  e que o coeficiente de atividade  $\gamma_c$  depende da molalidade, a Equação 5.2.19 pode ser desenvolvida como:

$$\left(\frac{\partial \mu_c}{\partial m_c}\right)_{T,P,m_p} = RT \left(\frac{1}{m_c} + \frac{\partial \ln \gamma_c}{\partial m_c}\right)_{T,P,m_p}. \quad (5.2.20)$$

Assim, combinando as Equações 5.2.17 e 5.2.20, a variação da energia livre de solvatação (Equação 5.2.16) pode ser reescrita como:

$$\Delta \mu_p^{tr} = -RT \int_0^{m_c} \Gamma_{cp} \left(\frac{1}{m_c} + \frac{\partial \ln \gamma_c}{\partial m_c}\right)_{T,P,m_p} dm_c. \quad (5.2.21)$$

A Equação 5.2.21 mostra como a energia livre de transferência para transportar uma proteína de um sistema com água pura para uma solução de molalidade  $m_c$  está relacionada ao coeficiente

de solvatação preferencial do cossolvente em relação à proteína ( $\Gamma_{cp}$ ). Além disso, o cálculo requer a consideração da variação do coeficiente de atividade do cossolvente com a molalidade, conforme evidenciado pelo termo  $\left(\frac{\partial \ln \gamma_c}{\partial m_c}\right)_{T,P,m_p}$ .

Para simplificar o desenvolvimento, é possível assumir, como uma aproximação válida para baixas concentrações e em uma faixa estreita de molalidade, que o coeficiente de atividade não varia com a molalidade. A Equação 5.2.21 pode ser frequentemente encontrada em uma forma alternativa, como em [57, 147], onde a razão  $\left(\frac{\Gamma_{cp}}{m_c}\right)$  é isolada e tratada como uma média ao longo do intervalo de integração. Isto dá-se pela aproximação em que, em intervalos de molalidade pequenos,  $\Gamma_{cp}$  é proporcional à  $m_c$ . Baynes e Trout, em [147], aplicam esta ideia para reduzir a integral da Equação 5.2.21 fazendo:

$$\Delta\mu_p^{tr} = -RT \int_0^{m_c} \Gamma_{cp} \left(\frac{1}{m_c}\right) dm_c, \quad (5.2.22)$$

$$\Delta\mu_p^{tr} = -RT \left(\frac{\Gamma_{cp}}{m_c}\right) \int_0^{m_c} dm_c, \quad (5.2.23)$$

$$\Delta\mu_p^{tr} = -RT\Gamma_{cp}. \quad (5.2.24)$$

O desenvolvimento acima mostra a conexão entre o parâmetro preferencial de solvatação e a energia de transferência, exemplificando como ele pode ser usado para estudar a termodinâmica da solvatação. Vale destacar que, por conveniência, a escala de concentração foi expressa em molalidade. Trabalhos que calculam a energia livre de transferência de proteínas costumam adotar escalas em concentração molar, sendo necessárias transformações adequadas, conforme apresentado em diversos trabalhos de Timasheff [53, 144, 145, 148, 150].

O parâmetro de solvatação preferencial,  $\Gamma_{cp}$ , define termodinamicamente a associação entre um componente da solução e o soluto. Se a concentração do cossolvente nas proximidades da proteína for maior que no bulk, então  $\Gamma_{cp} > 0$ , e o potencial químico da proteína,  $\mu_p$ , será menor na presença do cossolvente do que na sua ausência [147].

A Equação 5.2.24 mostra a relação direta entre  $\Gamma_{cp}$  e o potencial químico. A solvatação preferencial e a perturbação no potencial químico referem-se a uma composição específica do solvente. O efeito do cossolvente, tomando a água pura como estado de referência, é expresso por  $\Delta\mu_p^{tr}$ , que representa a energia livre de transferência da proteína da água para o

solvente de composição. A Figura 5.6 ilustra uma situação em que  $\Gamma_{cp}$  é positivo, resultando na diminuição do potencial químico da proteína na presença de um ligante. Neste exemplo hipotético, o ligante está inicialmente confinado em uma caixa no sistema que contém proteína e água.

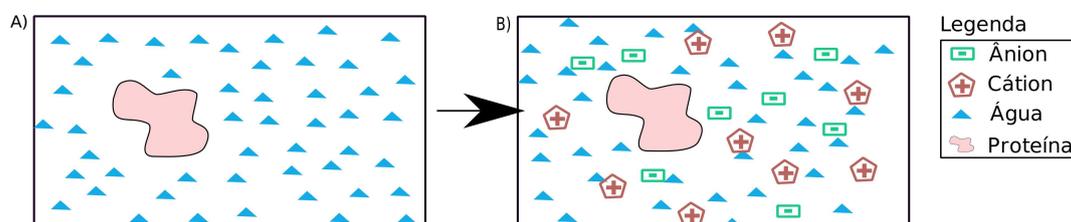


Figura 5.6: Representação do sistema descrevendo a adição de um ligante a uma solução aquosa com proteína. Em (A), temos um sistema inicial com água e proteína em solução, onde a proteína possui um determinado potencial químico ( $\mu_p^0$ ). Em (B), o ligante é introduzido e se distribui na solução, interagindo com a proteína. Esse processo perturba o potencial químico da proteína, causando acúmulo de ligante em sua superfície. Nesta situação,  $\Gamma_{uc}$  é positivo, e o potencial químico da proteína é reduzido em relação ao sistema (A), conforme indicado na Equação 5.2.24.

#### 5.2.4 Montagem dos sistemas e detalhes técnicos das simulações de dinâmica molecular

As simulações de dinâmica molecular de equilíbrio para as estruturas solvatadas por água e com diversas combinações de pares de LIs e concentrações foram realizadas usando o software GROMACS 2018.3 CUDA [151, 152]. As configurações iniciais dos sistemas foram construídas usando o software Packmol [153, 154]. Os parâmetros para descrever os líquidos iônicos e a proteína foram obtidos, respectivamente, dos campos de força OPLS de sítio virtual [155] e OPLS-AA [156], utilizando o modelo de água TIP3P [157]. A integração numérica das equações de movimento foi realizada com o algoritmo de Verlet leapfrog e um passo de 2 fs. Interações eletrostáticas de curto alcance e de Lennard-Jones foram cortadas a 1,0 nm, enquanto as interações eletrostáticas de longo alcance foram tratadas pela soma de Ewald com malha de partículas [123], usando uma interpolação de quarta ordem e espaçamento de grade de 0,16 nm. A temperatura foi mantida em 300 K com o termostato Berendsen modificado e um tempo de relaxamento de 0,1 ps [116, 158].

A pressão foi mantida constante em 1 bar usando o barostato de Parrinello-Rahman, com tempo de relaxamento de 2 ps e compressibilidade isotérmica de  $4.5 \times 10^{-5} \text{ bar}^{-1}$  [116, 117]. A energia de cada sistema foi inicialmente minimizada por

50.000 passos do método do gradiente [159], mantendo as coordenadas da proteína fixas. O processo incluiu 1 ns de equilíbrio térmico em condições *NVT*, seguido por 5 ns de MD no ensemble *NPT*, com a *backbone* da proteína fixada. Após essas etapas, as restrições foram removidas, e simulações adicionais de 1 ns em condições *NPT* foram realizadas. As simulações de produção duraram 10 ns também no conjunto *NPT*. Foram simuladas 20 conformações independentes usando o procedimento descrito, obtendo-se uma boa amostra da estrutura do solvente, dado que a estrutura da ubiquitina manteve-se próxima ao estado inicial. Assim, os dados apresentados neste trabalho correspondem às médias de 20 simulações independentes para cada sistema, com o erro padrão da média incluído. Restrições harmônicas suaves com uma constante de força de  $10 \text{ kJ mol}^{-1} \text{ \AA}^{-2}$  foram aplicadas aos átomos  $C\alpha$  para impedir rotações durante as etapas de equilíbrio.

As MDDFs, integrais de KB e análises das contribuições atômicas de soluto e solvente foram calculadas usando o pacote *ComplexMixtures.jl* [160, 161]. A densidade foi obtida pelo número médio de distâncias mínimas em cada bin de  $0,1 \text{ \AA}$ , e as funções de distribuição de distâncias mínimas foram construídas usando uma versão discretizada da Eq. 5.2.1 [90]. As Eqs. 5.2.6 e 5.2.8 foram empregadas para calcular as integrais de KB e os parâmetros de solvatação preferencial.

Para determinar a convergência da integral de KB, utilizamos  $R = 20 \text{ \AA}$  em todos os sistemas. Este valor, apesar de incomumente grande [162, 163], foi necessário devido ao tamanho e natureza eletrostática dos íons dos LIs, que possuem correlações de longo alcance. O volume de solução a uma distância menor que  $R$  do soluto foi considerado *domínio da proteína* — a região onde a estrutura da solução é influenciada pela presença da proteína. O volume externo a essa região representa a mistura de cossolventes, sendo utilizado para deduzir a estrutura e propriedades termodinâmicas da solução sem a proteína. A concentração efetiva do *bulk* na solução foi obtida das simulações, calculando a densidade de cada solvente em uma região entre 20 e 30  $\text{\AA}$  da superfície da proteína. Ou seja, calculando de um sub volume aberto do sistema, provendo uma correção de tamanho finito na computação das integrais de KB.

## Capítulo 6

# Solvatação de proteínas por Lis: visão geral

Este capítulo aborda a análise estrutural e termodinâmica da solvatação em sistemas contendo soluções de um único tipo de LI em água. A distribuição dos íons nas camadas de solvatação próximas à proteína será investigada por meio das MDDFs. Além disso, a decomposição das MDDFs em contribuições individuais do soluto e do solvente permitirá uma compreensão detalhada da natureza das interações estabelecidas entre a ubiquitina e os Lis.

### 6.1 Efeitos correlacionados na solvatação

A Figura 6.1 apresenta as MDDFs e integrais de KB para o cátion  $[\text{EMIM}]^+$  e o ânion  $[\text{DCA}]^-$  relativamente à ubiquitina em uma solução de LI com concentração aproximada de  $2,0 \text{ mol L}^{-1}$ . Embora tenham sido simulados sistemas com concentrações de LI variando de  $0,5$  a  $3,0 \text{ mol L}^{-1}$ , os perfis das MDDFs nesta concentração fornecem características que ajudam a entender sobre a natureza das interações e a formação das camadas de solvatação, que podem ser consideradas para as demais concentrações.

Observa-se que o  $[\text{EMIM}]^+$  se concentra significativamente na superfície da proteína, com uma densidade que chega a ser cerca de sete vezes maior em uma distância de  $2,3 \text{ \AA}$  da proteína, em comparação à densidade de *bulk*. Essa distância é típica de interações não polares ou dipolares de baixa especificidade. Esse comportamento é corroborado pelo fato de que as interações diretas do  $[\text{EMIM}]^+$  com a proteína, descritas pela interação de Lennard-Jones ( $-601 \text{ kcal mol}^{-1}$ ), são substancialmente mais fortes do que as interações eletrostáticas ( $-134 \text{ kcal mol}^{-1}$ ), como reportamos em [164]. Além disso, há um segundo pico difuso em aproximadamente  $5 \text{ \AA}$ , característico de uma camada de solvatação

secundária, na qual as interações com a proteína podem ser mediadas por água, pelo contra-íon ou pelo próprio  $[\text{EMIM}]^+$ . Em distâncias entre 8 e 10 Å, a MDDF converge para 1.

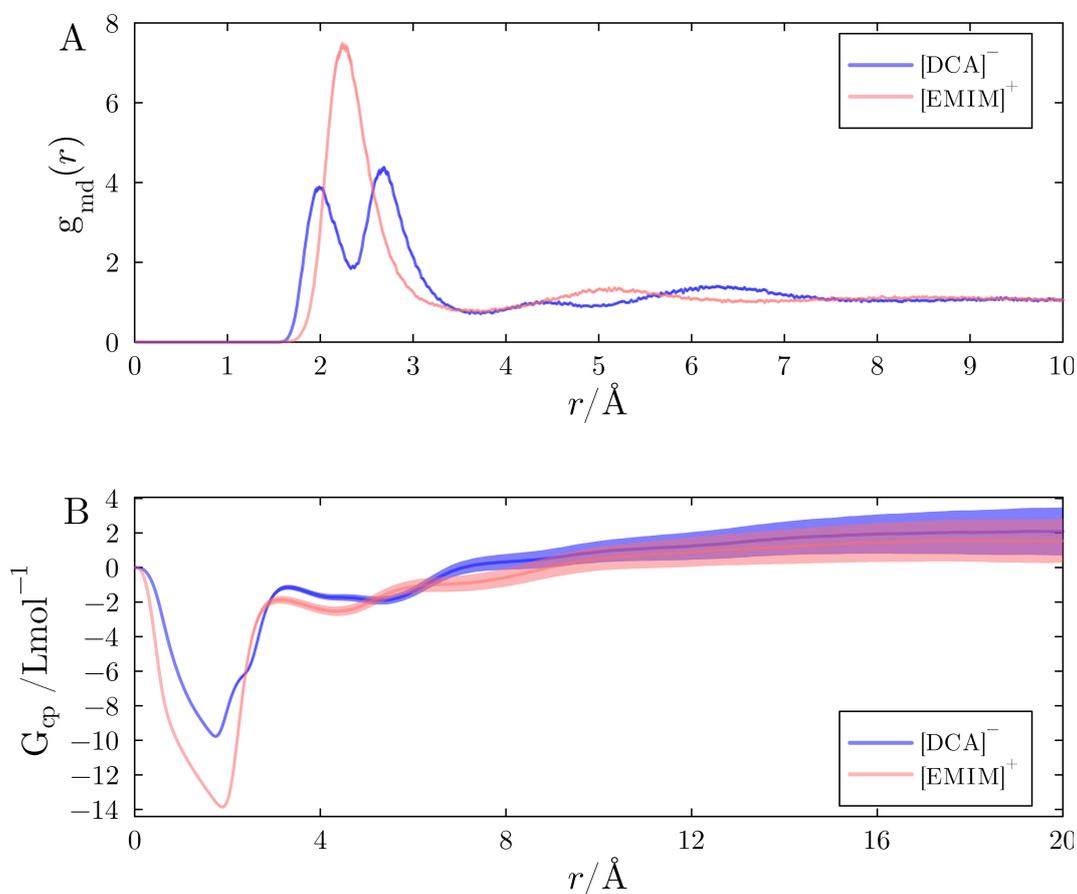


Figura 6.1: A) MDDFs (representadas por  $g_{md}(r)$ ) e B) integrais de KB ( $G_{cp}(r)$ ) do cátion  $[\text{EMIM}]^+$  e do ânion  $[\text{DCA}]^-$  relativos à ubiquitina em uma solução de aproximadamente  $2,0 \text{ mol L}^{-1}$  de  $[\text{EMIM}][\text{DCA}]$  em água. Esta figura ilustra como funções de distribuição marcadamente diferentes podem dar origem a integrais de KB idênticas, implicando que as funções de distribuição estão correlacionadas em grandes distâncias a partir da superfície do soluto.

A distribuição dos ânions na solução difere da dos cátions devido à capacidade de formação de ligações de hidrogênio. O ânion  $[\text{DCA}]^-$  também se concentra significativamente na superfície da proteína, apresentando picos nas MDDFs a 1,9 Å e 2,7 Å, com um aumento de densidade aproximadamente quatro vezes maior em relação à distribuição de referência. O pico em 1,9 Å é característico de ligações de hidrogênio, enquanto interações não específicas com a proteína podem levar ao acúmulo de  $[\text{DCA}]^-$  em torno de 2,7 Å. Nessa concentração, o  $[\text{DCA}]^-$  forma cerca de 12 ligações de hidrogênio e exibe fortes interações dispersivas e eletrostáticas com a proteína [164].

Observam-se flutuações difusas entre 3 e 8 Å, e, além dessa distância, a distribuição de  $[DCA]^-$  não parece estar correlacionada com a proteína (convergência para 1, indicando comportamento *bulk*). As MDDFs de  $[EMIM]^+$  e  $[DCA]^-$  refletem interações diretas entre a proteína e os íons, levando a aumentos de densidade localizados em distâncias curtas. No entanto, a natureza das interações difere, com o  $[DCA]^-$  formando ligações de hidrogênio e múltiplas camadas de solvatação bem definidas, enquanto o  $[EMIM]^+$  estabelece principalmente interações não-específicas (diferentes de ligações de hidrogênio) com a superfície da proteína.

A diferença notável entre as funções de distribuição de  $[EMIM]^+$  e  $[DCA]^-$  na Figura 6.1A sugere que, em distâncias curtas, um dos íons apresenta afinidade preferencial pela proteína em relação ao outro. No entanto, em uma solução contendo apenas um tipo de cátion e de ânion, esse comportamento não ocorre isoladamente. A Figura 6.1B apresenta as integrais de KB dos íons, calculadas a partir das contagens de distâncias mínimas e utilizando a Eq. 5.2.6. Em distâncias muito curtas, ou seja,  $r < 1,5$  Å, as integrais de KB obtidas a partir das MDDFs estão associadas ao volume de exclusão da proteína e do solvente, assumindo valores negativos, pois a densidade dos íons nessas distâncias é nula, resultando na exclusão efetiva de B) 3,0 mol Las moléculas de solvente do domínio. A queda da integral de KB para  $[EMIM]^+$  é mais acentuada do que para  $[DCA]^-$ , devido ao maior volume do cátion.

Entre aproximadamente 1,5 Å e 8 Å, as integrais KB consideram o número de moléculas de solvente efetivamente correlacionadas com a superfície da proteína. Para cada íon, a integral de KB varia em função da MDDF, com o  $[EMIM]^+$  apresentando um aumento acentuado em torno de 2,7 Å, enquanto o  $[DCA]^-$  exibe dois patamares de acumulação associados aos picos de sua MDDF. Em distâncias maiores, as correlações indiretas com a superfície da proteína levam ao acúmulo final de íons no domínio da proteína, que, neste exemplo, começam a convergir em 10 Å (para a maioria das integrais de KB, a convergência foi observada em 20 Å).

Para esses íons (no sistema com  $2,0 \text{ mol L}^{-1}$ ), os valores finais das integrais de KB em distâncias grandes aproximam-se de zero, sugerindo que o aumento de densidade dos íons no domínio da proteína compensa o volume excluído pela proteína. A integral de KB da água neste sistema é negativa, conforme mostrado na Tabela 6.1, indicando que os íons se ligam preferencialmente à proteína em comparação com a água.

Tabela 6.1: Concentrações de referência e efetivas das soluções de LI simuladas, juntamente com as integrais de KB para todos os componentes do solvente em relação à proteína. Os erros padrão da média foram calculados a partir de pelo menos 20 simulações independentes. As diferenças entre as concentrações de referência e efetivas devem-se ao fato de que a concentração de referência representa o valor pretendido durante a montagem inicial da caixa de simulação, enquanto a concentração efetiva corresponde à concentração real do LI no *bulk* após a equilibração do sistema.

Sistema	Concentração de LI (referência/efetiva) mol L <sup>-1</sup>	Integral de KB do cátion / L mol <sup>-1</sup>	Integral de KB do ânion / L mol <sup>-1</sup>	Integral de KB da água / L mol <sup>-1</sup>
	0,5/0,465 ± 0,002	30,89 ± 2,64	31,40 ± 2,78	-9,32 ± 0,19
	1,0/0,960 ± 0,004	15,41 ± 2,26	17,43 ± 2,22	-10,58 ± 0,37
	1,5/1,493 ± 0,005	2,94 ± 1,65	3,66 ± 1,63	-9,51 ± 0,46
	2,0/2,024 ± 0,005	1,28 ± 1,38	1,58 ± 1,41	-10,09 ± 0,60
	2,5/2,597 ± 0,005	-5,85 ± 1,09	-5,75 ± 1,08	-6,86 ± 0,72
	3,0/3,151 ± 0,004	-5,49 ± 0,64	-5,50 ± 0,68	-7,18 ± 0,60

Valores equivalentes (de fato iguais em longas distâncias) implicam que o número de excesso de íons é o mesmo para o cátion e o ânion ao considerar toda a região da solução definida como domínio da proteína. Este é um resultado fundamental que decorre do fato de que as forças eletrostáticas eventualmente levam à acumulação ou depleção correlacionada dos íons, de modo que a solução *bulk* permaneça eletricamente neutra. Portanto, compreender os detalhes moleculares de como essas correlações emergem a partir das interações e das funções de distribuição proteína-íon, que diferem notavelmente entre si, é essencial para entender as interações de LIs com solutos complexos.

A Figura 6.2 mostra um mapa da diferença de densidade entre os íons [DCA]<sup>-</sup> e [EMIM]<sup>+</sup> a uma distância de até 3,5 Å da superfície proteica. Regiões em azul indicam que a densidade de [DCA]<sup>-</sup> é maior do que a de [EMIM]<sup>+</sup>, enquanto regiões em vermelho indicam que a densidade de [EMIM]<sup>+</sup> é superior à de [DCA]<sup>-</sup>. O primeiro pico na distribuição de mínima distância do [DCA]<sup>-</sup> pode agora ser associado a resíduos específicos da proteína, notadamente os resíduos carregados positivamente: K6, K11, K33, R42, K48, K63 e R64. Contribuições relevantes de alguns resíduos de Glutamina (Q) também podem ser identificadas. Regiões hidrofóbicas e carregadas negativamente são responsáveis pelos aumentos mais significativos de densidade de [EMIM]<sup>+</sup>, particularmente nos segmentos proteicos T7-L8-G9-T10, V17-E18-P19-S20, E34-G35-I36-P37, L43-I44-F45-A46 e L56-S57-D58.

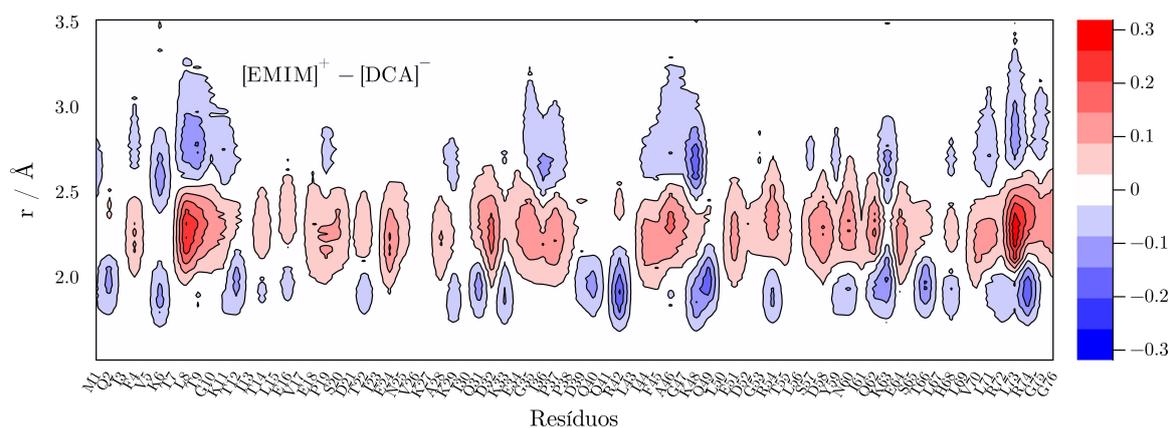


Figura 6.2: Diferença nas densidades dos íons próximos de cada resíduo da proteína em  $2,0 \text{ mol L}^{-1}$ . As regiões em azul indicam maior densidade de  $[\text{DCA}]^-$ , enquanto as regiões em vermelho indicam maior densidade de  $[\text{EMIM}]^+$ . A acumulação de  $[\text{DCA}]^-$  na faixa de  $2,5\text{--}3,0 \text{ \AA}$  ocorre complementarmente à acumulação de  $[\text{EMIM}]^+$  entre  $2,0\text{--}2,5 \text{ \AA}$ , sugerindo que o segundo pico da distribuição de  $[\text{DCA}]^-$  na Figura 6.1 corresponde a uma camada de solvatação do ânion mediada pelo cátion.

As interações da ubiquitina com os componentes de um LI em água resultam em perfis de densidade específicos. No entanto, independentemente da natureza química variável dos dois íons envolvidos, o acúmulo ou exclusão global excessiva dos íons no domínio da proteína é o mesmo, como esperado para garantir a eletro neutralidade da solução a longas distâncias da proteína. Este fenômeno implica uma alta correlação entre as distribuições do cátion e do ânion, evidenciando que não é possível projetar ou prever as interações entre LI e solutos apenas a partir da natureza química de cada íon isoladamente.

### 6.1.1 Natureza da interação entre solvente e ubiquitina.

A MDDF pode ser interpretada em termos da natureza das interações soluto-solvente. Cada pico no perfil de uma MDDF pode ser decomposto em contribuições dos grupos do LI e da proteína com maior probabilidade de estabelecer interações entre si. Essa decomposição permite analisar cada pico individualmente, identificando os tipos de resíduos que mais contribuem para o acúmulo de íons ao redor da proteína. A Figura 6.3 apresenta a decomposição das MDDFs dos cátions e ânions com respeito aos resíduos polares (Gln, Tyr, etc. — verde), neutros (Ala, Phe, etc. — cinza), ácidos (Glu, Asp — vermelho) e básicos (Lys, Arg, His — azul), seguindo a classificação e o esquema de cores do VMD [165].

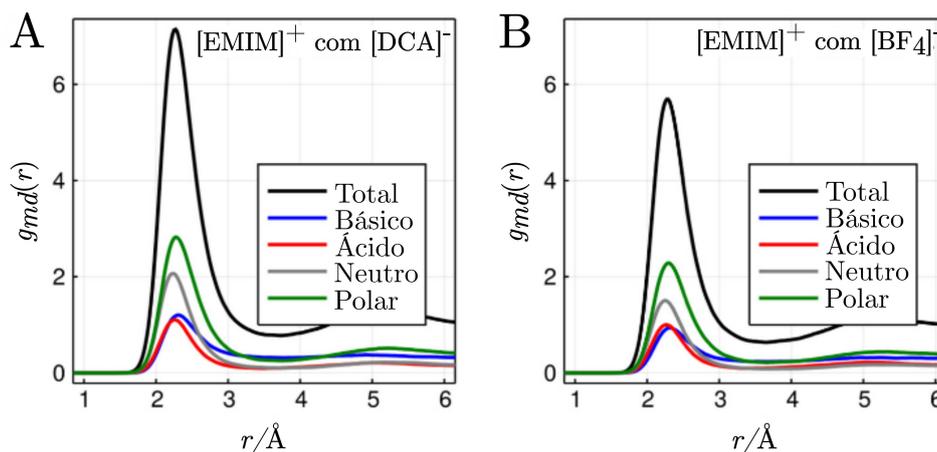


Figura 6.3: Decomposição das MDDFs de cátions e ânions em contribuições de resíduos da ubiquitina em sistemas com  $2,0 \text{ mol L}^{-1}$ . A) Decomposição da MDDF do  $[\text{EMIM}]^+$  na presença de  $[\text{DCA}]^-$ , e B) do  $[\text{EMIM}]^+$  na presença de  $[\text{BF}_4]^-$ .

A MDDF do  $[\text{EMIM}]^+$  na Figura 6.3A mostra um aumento de aproximadamente 7,1 vezes na densidade do cátion a  $2,3 \text{ Å}$ , comparado à densidade no *bulk*. Este pico ( $2,3 \text{ Å}$ ) é composto por contribuições de 2,8 de aumento de densidade associada a resíduos polares (verde), 2,0 a resíduos neutros (cinza), e 1,2 em torno de resíduos básicos e ácidos (vermelho e azul). Assim, o  $[\text{EMIM}]^+$  interage preferencialmente com resíduos polares e neutros, em comparação aos resíduos carregados, na distância do pico mais importante da MDDF.

Ao substituir o ânion  $[\text{DCA}]^-$  pelo  $[\text{BF}_4]^-$ , observa-se uma diminuição do pico total da MDDF do cátion (de 7,1 para 5,7), pois o  $[\text{DCA}]^-$  se liga mais fortemente à superfície da proteína, forçando o  $[\text{EMIM}]^+$  a se redistribuir para contrabalancear as cargas, conforme discutido anteriormente. A contribuição absoluta dos resíduos carregados é reduzida de 1,2 para 0,9 ( $-25\%$ ) para resíduos básicos e de 1,1 para 1,0 ( $-9\%$ ) para resíduos ácidos. Resíduos polares e neutros também mostram uma diminuição relativa de densidade: 0,5 para cada um (de 2,8 para 2,3, ou  $-18\%$ , para polares; e de 2,0 para 1,5, ou  $-25\%$ , para resíduos neutros). Dessa forma, a troca de ânion tem um impacto menor próximo a resíduos ácidos, que formam interações eletrostáticas atrativas com o cátion. A substituição do ânion afeta especialmente as interações mais fracas do cátion com a superfície da proteína, particularmente as interações dispersivas com resíduos neutros e, possivelmente, as interações indiretas com resíduos carregados positivamente, mediadas pelo ânion.

As interações de proteínas com cossolventes em solução aquosa podem ser

estudas pelo parâmetro de solvatação preferencial ( $\Gamma_{cp}$ ). Qualitativamente,  $\Gamma_{cp} > 0$  indica que o cossolvente interage favoravelmente com a superfície da proteína em relação à água, favorecendo estruturas com maior área superficial, podendo atuar como desnaturante. A Figura 6.4 apresenta parâmetros de solvatação preferencial ( $\Gamma_{cp}$ ) e hidratação preferencial ( $\Gamma_{wp}$ ) para as concentrações de 0,5 e 3,0 mol L<sup>-1</sup> dos LIs [EMIM][DCA], [EMIM][BF<sub>4</sub>], [EMIM][NO<sub>3</sub>] e [EMIM][Cl]. Como apresentado na Figura 6.4A,  $\Gamma_{cp}$  é positivo para todos os LIs na menor concentração, enquanto na Figura 6.4B é possível ver que  $\Gamma_{wp}$  é negativo para os sistemas correspondentes. Isto indica que na concentração de 0,5 mol L<sup>-1</sup>, a proteína é preferencialmente solvatada pelo LI, em todos os diferentes sistemas simulados.

Apesar da análise em uma única concentração indicar a solvatação preferencial da proteína pelo LI em 0,5 mol L<sup>-1</sup>, nota-se haver um importante efeito de concentração. Analisando as barras laranjas nas Figuras 6.4A e 6.4B temos que os valores de  $\Gamma_{cp}$  diminuem consideravelmente enquanto os valores de  $\Gamma_{wp}$  aumentam para sistemas com 3,0 mol L<sup>-1</sup> de LI (sendo uma exceção o sistema com [EMIM][Cl]). A dependência da concentração dos coeficientes de ligação preferencial e de hidratação preferencial pode ser explicada por uma saturação progressiva da camada de solvatação com os cátions do LI, acompanhada pela menor concentração de água no volume para o LI com [DCA]<sup>-</sup>.

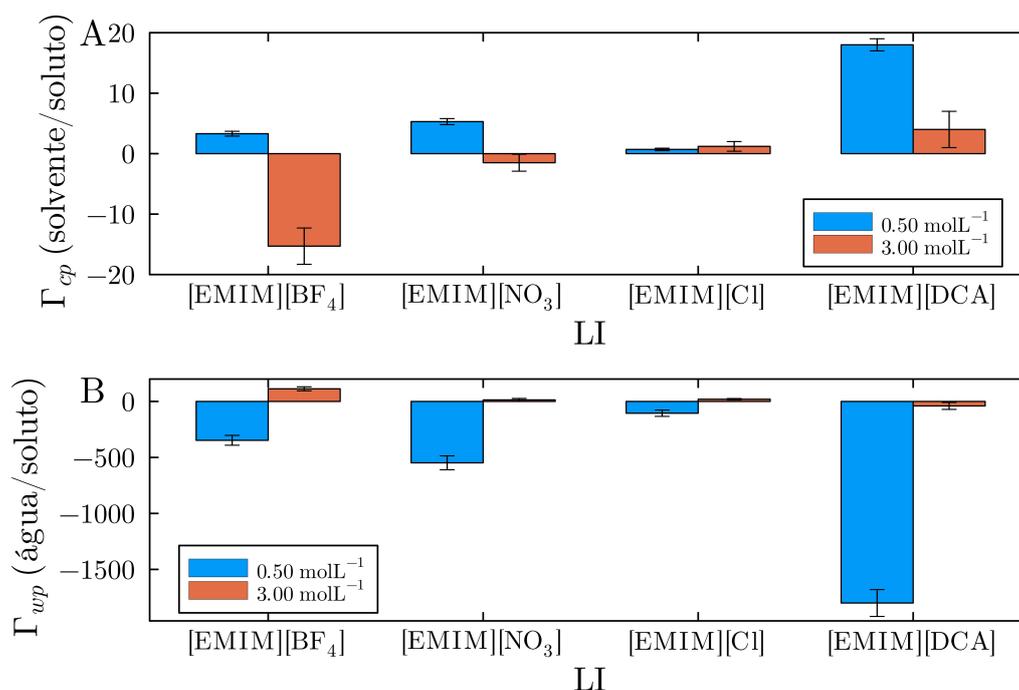


Figura 6.4: (A) Parâmetros de interação preferencial,  $\Gamma_{cp}$  dos LIs com a proteína nas concentrações de 0,5 e 3,0 mol L<sup>-1</sup> de LI. (B) Parâmetros de hidratação preferencial,  $\Gamma_{wp}$ , nas mesmas concentrações de 0,5 e 3,0 mol L<sup>-1</sup> de LI. Observa-se que todos os LIs possuem valores positivos de  $\Gamma_{cp}$  em baixas concentrações, enquanto  $\Gamma_{wp}$  é negativo, indicando uma desidratação preferencial em concentrações menores. No entanto, em concentrações mais altas, ocorre um aumento expressivo de  $\Gamma_{wp}$  e uma diminuição de  $\Gamma_{cp}$ , revelando o comportamento inverso em comparação com as concentrações mais baixas.

O efeito da troca de ânion pode ser analisado pela comparação entre as Figuras 6.4A e 6.4B. O [DCA]<sup>-</sup> é um ligante mais forte à superfície da proteína do que [BF<sub>4</sub>]<sup>-</sup>, [NO<sub>3</sub>]<sup>-</sup> e [Cl]<sup>-</sup>, resultando em parâmetros de interação preferencial maiores em todos os casos. Em 0,5 mol L<sup>-1</sup>, as interações da Ubiquitina com os LIs são sempre favorecidas em relação à água, e a proteína é preferencialmente desidratada. A exceção é o sistema com [EMIM][Cl], onde a desidratação é menor em comparação aos outros ânions. Em 3,0 mol L<sup>-1</sup>, os LIs tendem a preferir à solução em *bulk* em relação à superfície da proteína. O ânion desempenha um papel crucial nessa ligação preferencial, com o [DCA]<sup>-</sup> promovendo maior desidratação da proteína quando comparado a outros ânions. Em concentrações altas, os LIs tornam-se preferencialmente excluídos do domínio da proteína em todos os casos, exceto na solução com [EMIM][DCA]. [EMIM][Cl] também possui valor positivo para  $\Gamma_{cp}$ , contudo não há grandes variações de  $\Gamma_{cp}$  do LI com [Cl]<sup>-</sup> ao longo das concentrações, conforme discutido em [166].

A análise das interações de proteínas com LIs revela como a troca de contra íons influencia as distribuições dos cossolventes em relação à superfície proteica. Em função da

eletroneutralidade das soluções em *bulk*, as distribuições dos cátions e ânions dos LIs estão correlacionadas. O ânion  $[\text{DCA}]^-$ , ao formar ligações de hidrogênio fortes com a superfície proteica, provoca um acúmulo substancial de ambos os íons na superfície da proteína, efeito que pode ser atenuado pela presença de cátions hidrofóbicos em concentrações mais altas. A extensão do acúmulo dos cátions próximos à proteína varia com a natureza do ânion, onde interações específicas entre íons e superfícies proteicas se mantêm mesmo após a troca de contra íons. Por outro lado, interações menos específicas nas segundas e terceiras camadas de solvatação são moduladas pela variação do contra íon e da concentração. Desta forma, a composição e concentração do LI podem ser ajustadas para desempenhar papéis estabilizadores ou desestabilizadores nas estruturas proteicas, permitindo um desenho direcionado de solventes.

## Capítulo 7

# Efeitos não-aditivos

O presente capítulo estuda a solvatação da ubiquitina em sistemas com misturas de LIs. Estas misturas mantêm como cátion apenas o íon  $[EMIM]^+$  mas contendo duas espécies aniônicas diferentes. A discussão desenvolvida neste capítulo irá apresentar que a combinação de ânions com diferentes afinidades pela superfície proteica modula o acúmulo iônico de maneira não-aditiva, destacando a relevância de fatores como competição por sítios de interação específicos, ligação cooperativa, afinidade em solução a granel e efeitos de compensação de carga na capacidade geral de solvatação.

### 7.1 Solvatação da ubiquitina por múltiplos íons

A influência dos íons na estabilidade das proteínas tem sido o foco de extensas pesquisas desde os estudos pioneiros de Hofmeister no século XIX [167, 168]. No entanto, apesar de décadas de investigação, a natureza precisa das interações entre íons e proteínas continua a ser um tema de debate, especialmente para íons complexos e misturas de eletrólitos [169, 170]. Essa complexidade emerge da interação entre íons, moléculas de água e a superfície da proteína.

Por exemplo, o conhecimento das cargas e concentrações de íons é insuficiente para entender as interações entre íons e proteínas: a identidade específica dos íons desempenha também um papel fundamental [171]. Esse fenômeno, conhecido como efeito específico dos íons, foi explorado sistematicamente por pesquisadores como Lewith e Hofmeister. Eles observaram que diferentes íons possuem eficácias variadas na precipitação de proteínas de soluções aquosas de soro sanguíneo e clara de ovo [172, 173]. Essa ordenação dos íons, conhecida como série de Hofmeister, foi também verificada em sistemas biológicos

[55], poliméricos [174] e não aquosos [175], destacando a importância da identidade dos íons na modelagem do comportamento de solutos, especialmente proteínas [176].

A Figura 7.1 apresenta MDDFs de cada componente de um sistema com  $2,0 \text{ mol L}^{-1}$  da mistura dos LIs [EMIM][Cl] e [EMIM][DCA]. O ânion  $[\text{DCA}]^{-}$  como já discutido anteriormente, é, de todos os ânions estudados nesta tese, o com a maior capacidade de se distribuir próximo da superfície da proteína. A distribuição do  $[\text{DCA}]^{-}$  possui o mesmo perfil do apresentado na Figura 6.1, com dois picos significativos de acúmulo a uma distância de até  $4 \text{ \AA}$ . O pico inicial, a aproximadamente  $1,9 \text{ \AA}$ , surge das ligações de hidrogênio formadas pelos átomos de nitrogênio do  $[\text{DCA}]^{-}$ . Em contraste, o segundo pico decorre de uma variedade de interações, incluindo interações diretas o  $[\text{DCA}]^{-}$  e a proteína, bem como interações mediadas por outros íons ou moléculas posicionadas entre o  $[\text{DCA}]^{-}$  e os átomos da superfície da proteína. O amplo pico que se estende de  $5$  a  $7,5 \text{ \AA}$  indica que a correlação entre a distribuição do  $[\text{DCA}]^{-}$  e a presença da proteína persiste até grandes distâncias.

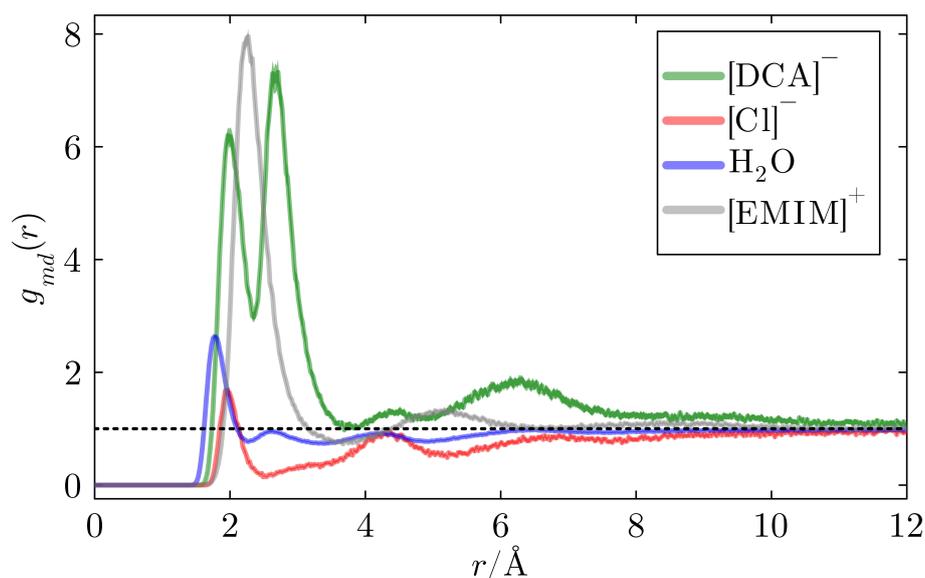


Figura 7.1: MDDFs entre ubiquitina e íons para  $[\text{DCA}]^{-}$ ,  $[\text{Cl}]^{-}$ ,  $[\text{EMIM}]^{+}$  e água na solução aquosa contendo [EMIM][DCA] e [EMIM][Cl]. As interações de ligação de hidrogênio correspondem a picos nas distribuições em aproximadamente  $1,9 \text{ \AA}$ , observadas para  $[\text{DCA}]^{-}$  e água. Interações eletrostáticas fortes, mais evidentes no pico do  $[\text{Cl}]^{-}$  em torno de  $2,0 \text{ \AA}$ , também são identificadas para os demais íons. Interações menos específicas se refletem em picos localizados a distâncias maiores.

O ânion  $[\text{DCA}]^{-}$  ocupa o volume próximo à superfície da proteína devido às ligações de hidrogênio estabelecidas com a proteína, enquanto os ânions  $[\text{Cl}]^{-}$  não interagem na mesma intensidade, distribuindo-se relativamente menos do que os íons  $[\text{DCA}]^{-}$ . Além

disso, a MDDF que representa a distribuição dos cátions  $[\text{EMIM}]^+$  exibe um pico proeminente em aproximadamente 2,4 Å, indicando interações dispersivas estabelecidas pelo  $[\text{EMIM}]^+$  com a proteína. Ademais, o acúmulo de  $[\text{EMIM}]^+$  ocorre em distâncias complementares com as do  $[\text{DCA}]^-$ , pois os íons contrapõem localmente suas cargas de forma cooperativa, conforme reportamos em [164, 166, 177]. Assim, o  $[\text{DCA}]^-$  apresenta interações notavelmente mais favoráveis com a superfície da proteína do que o  $[\text{Cl}]^-$ , tanto especificamente em distâncias de ligações de hidrogênio quanto não especificamente, formando uma segunda camada de solvatação mediada pelo acúmulo do cátion.

A Figura 7.2 exibe as MDDFs dos ânions em relação à proteína nas soluções de IL com um único tipo de ânion, decompostas nas contribuições dos diferentes tipos de resíduos de aminoácidos da proteína. Para  $[\text{DCA}]^-$ ,  $[\text{BF}_4]^-$  e  $[\text{NO}_3]^-$ , há picos distintivos centrados em torno de 1,8 e 2,0 Å, que correspondem às ligações de hidrogênio que esses íons estabelecem com a proteína. O  $[\text{Cl}]^-$  exibe interações fortes com resíduos básicos, como esperado, em distâncias ligeiramente maiores. O primeiro pico para  $[\text{DCA}]^-$ ,  $[\text{NO}_3]^-$  e  $[\text{BF}_4]^-$  é principalmente devido a interações com resíduos polares e básicos (que podem ser, respectivamente, glutamina e lisina). O segundo pico, que aparece entre 2,4 e 2,6 Å para esses três ânions, tem contribuições adicionais significativas originadas de resíduos neutros (glicina, leucina e valina). Esse segundo pico decorre de interações mediadas pelo cátion do LI, que exibe interações hidrofóbicas favoráveis e atrai o ânion para a proximidade de resíduos neutros.

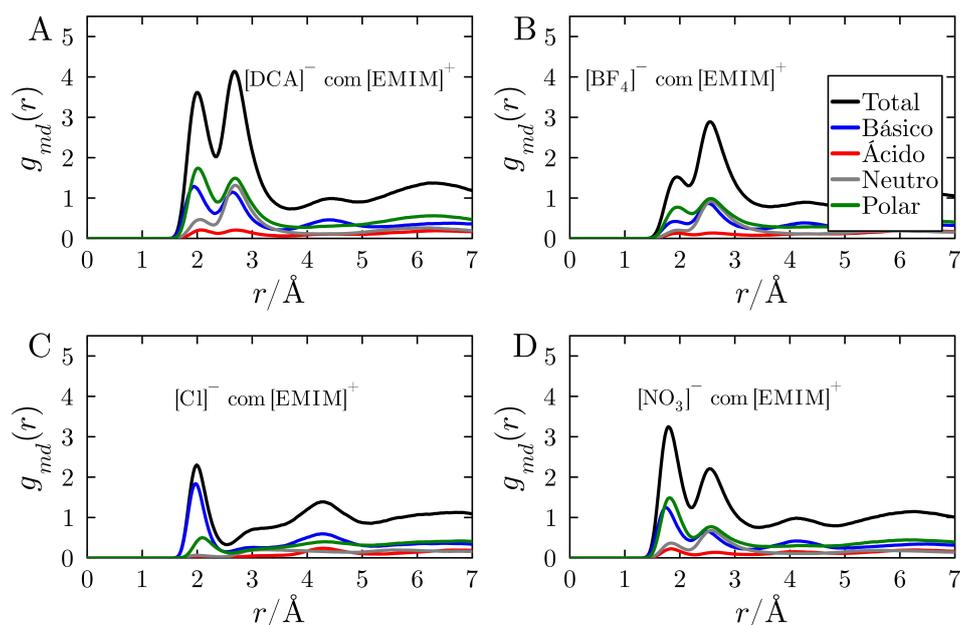


Figura 7.2: Decomposição das MDDFs de (A)  $[\text{DCA}]^-$ , (B)  $[\text{BF}_4]^-$ , (C)  $[\text{Cl}]^-$  e (D)  $[\text{NO}_3]^-$  nas contribuições dos tipos de resíduos de aminoácidos da proteína em sistemas com LI composto pelo cátion  $[\text{EMIM}]^+$  e um único tipo de ânion, a uma concentração de  $2,0 \text{ mol L}^{-1}$ . A MDDF total é a soma das contribuições de cada tipo de resíduo.

A Tabela 7.1 apresenta o número de íons até uma distância de  $5,0 \text{ \AA}$  da superfície da ubiquitina. Observa-se que os íons  $[\text{DCA}]^-$  são mais abundantes na proximidade da proteína em comparação com os demais ânions, especialmente em relação aos íons  $[\text{Cl}]^-$ . O ânion  $[\text{NO}_3]^-$  também aparece em maior número, embora menor que de  $[\text{DCA}]^-$ . Entre os ânions simulados,  $[\text{DCA}]^-$  e  $[\text{NO}_3]^-$  são aqueles com maior afinidade pela superfície da ubiquitina, considerando as análises das MDDFs, integrais de KB, e o número de ligações de hidrogênio estabelecidas, como demonstramos em [164, 166]. Quando esses ânions são misturados com  $[\text{BF}_4]^-$  e  $[\text{Cl}]^-$ , eles prevalecem em número próximo à superfície da proteína.

Tabela 7.1: Número de coordenação de  $[\text{EMIM}]^+$  e dos ânions ao redor da proteína, contando o número de íons até 5 Å da superfície da proteína, em sistemas com  $2,0 \text{ mol L}^{-1}$  de mistura de LI. Os sistemas possuem 1004 cátions e 502 ânions de cada tipo nesta concentração. As flutuações relatadas são o erro padrão da média das 20 réplicas realizadas para cada sistema

Mistura de Líquido Iônico ( $[\text{EMIM}]^+$ + Ânions)	$[\text{EMIM}]^+$	Ânion 1	Ânion 2
$[\text{DCA}]^-$ e $[\text{BF}_4]^-$	$47 \pm 1$	$25 \pm 1$ ( $[\text{DCA}]^-$ )	$11 \pm 1$ ( $[\text{BF}_4]^-$ )
$[\text{DCA}]^-$ e $[\text{NO}_3]^-$	$47 \pm 1$	$23 \pm 1$ ( $[\text{DCA}]^-$ )	$14 \pm 1$ ( $[\text{NO}_3]^-$ )
$[\text{DCA}]^-$ e $[\text{Cl}]^-$	$51 \pm 1$	$30 \pm 1$ ( $[\text{DCA}]^-$ )	$7 \pm 1$ ( $[\text{Cl}]^-$ )
$[\text{BF}_4]^-$ e $[\text{NO}_3]^-$	$43 \pm 1$	$14 \pm 1$ ( $[\text{BF}_4]^-$ )	$18 \pm 1$ ( $[\text{NO}_3]^-$ )
$[\text{BF}_4]^-$ e $[\text{Cl}]^-$	$42 \pm 1$	$18 \pm 1$ ( $[\text{BF}_4]^-$ )	$12 \pm 1$ ( $[\text{Cl}]^-$ )
$[\text{NO}_3]^-$ e $[\text{Cl}]^-$	$44 \pm 1$	$20 \pm 1$ ( $[\text{NO}_3]^-$ )	$9 \pm 1$ ( $[\text{Cl}]^-$ )

O  $[\text{DCA}]^-$  é o ânion que mais fortemente promove a aproximação do cátion  $[\text{EMIM}]^+$  à superfície proteica. Em soluções contendo  $[\text{DCA}]^-$ , encontram-se entre 47 e 51 íons  $[\text{EMIM}]^+$  a até 5,0 Å da proteína, enquanto, na ausência de  $[\text{DCA}]^-$ , esse número não ultrapassa 43. As menores concentrações de  $[\text{EMIM}]^+$  nessa faixa de distância ocorrem nas soluções com  $[\text{BF}_4]^-$  e  $[\text{Cl}]^-$ , os ânions com menor afinidade pela superfície proteica. Ânions com interações mais fortes tendem a se acumular próximos à superfície, criando uma carga líquida negativa que, por compensação eletrostática, resulta em um acúmulo de cátions. Tanto o  $[\text{DCA}]^-$  quanto o  $[\text{NO}_3]^-$  formam ligações de hidrogênio com a superfície, mas o  $[\text{DCA}]^-$  demonstra maior afinidade, evidenciada pelo maior número de ânions  $[\text{DCA}]^-$  em contato próximo com a proteína.

A Figura 7.3 apresenta mapas de diferença de densidade dos ânions ao redor dos resíduos proteicos em sistemas com mistura de LIs. Na Figura 7.3A, regiões em laranja indicam uma maior densidade de  $[\text{DCA}]^-$  em comparação com o  $[\text{Cl}]^-$ , enquanto áreas em azul indicam maior densidade de  $[\text{Cl}]^-$ . O  $[\text{DCA}]^-$  está mais concentrado na proximidade da proteína em relação ao  $[\text{Cl}]^-$ .

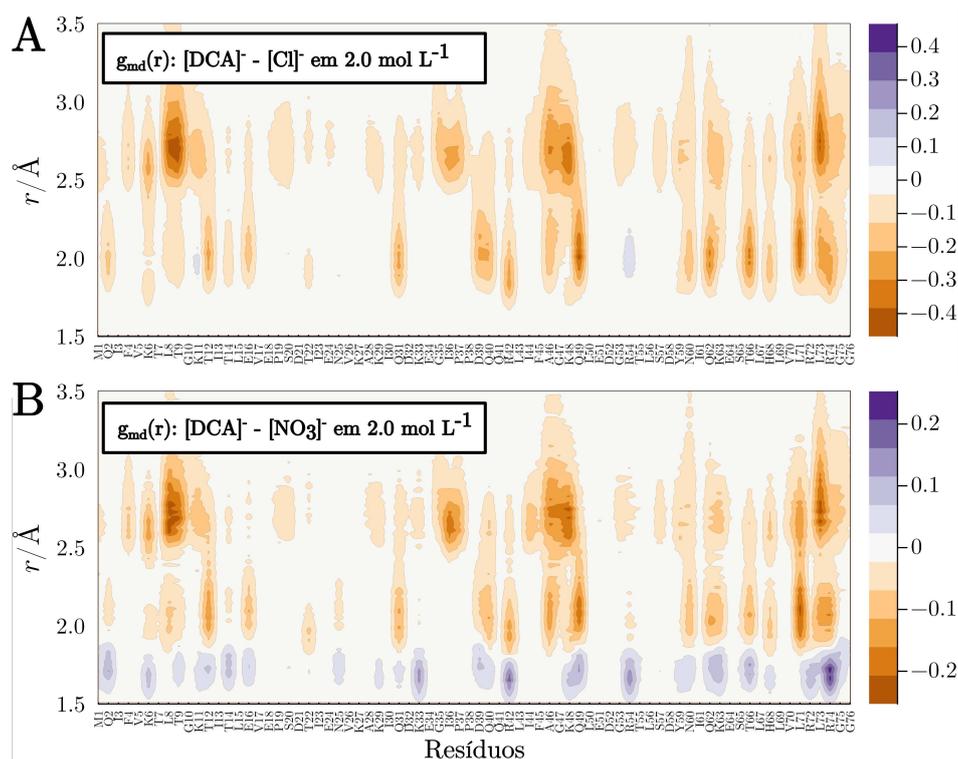


Figura 7.3: (A) Diferença nas densidades das MDDFs dos íons nas proximidades da proteína a aproximadamente  $2,0 \text{ mol L}^{-1}$  de  $[\text{EMIM}][\text{DCA}] + [\text{EMIM}][\text{Cl}]$ . As regiões em azul-púrpura representam áreas onde a densidade de  $[\text{Cl}]^{-}$  é maior do que a de  $[\text{DCA}]^{-}$ , enquanto laranja indica regiões onde a densidade de  $[\text{DCA}]^{-}$  é maior do que a de  $[\text{Cl}]^{-}$ . (B) Perfil similar para a mistura  $[\text{EMIM}][\text{DCA}] + [\text{EMIM}][\text{NO}_3]$ , com laranja representando, de maneira semelhante, uma maior densidade de  $[\text{DCA}]^{-}$ .

A análise evidencia a competição entre ânions por interações com os resíduos proteicos, com características distintas influenciando a afinidade de solvatação. O  $[\text{Cl}]^{-}$  interage com resíduos básicos carregados positivamente, mas de forma mais fraca que o  $[\text{DCA}]^{-}$ , o qual se acumula mais devido à sua capacidade de formar ligações de hidrogênio com a proteína. Na Figura 7.3B, que compara  $[\text{DCA}]^{-}$  e  $[\text{NO}_3]^{-}$ , observa-se que o  $[\text{NO}_3]^{-}$  tende a se concentrar em regiões mais próximas à proteína para formação de ligações de hidrogênio, enquanto o  $[\text{DCA}]^{-}$  apresenta maior concentração em distâncias de solvatação mais amplas.

A competição por locais de ligação de hidrogênio é evidente, e embora a distribuição de densidade sugira que o  $[\text{NO}_3]^{-}$  possua maior afinidade, dados adicionais, como as integrais de KB e os parâmetros de solvatação preferencial, indicam que o  $[\text{DCA}]^{-}$  apresenta uma afinidade de solvatação significativamente superior ao  $[\text{NO}_3]^{-}$ . Esses resultados destacam a complexidade das interações específicas e a importância de múltiplas métricas na avaliação da

preferência de solvatação em sistemas com diferentes ânions.

A Figura 7.4 apresenta as integrais KB para íons e água em um sistema com concentração  $2,0 \text{ mol L}^{-1}$ . A rápida queda a curtas distâncias ( $r < 1,5 \text{ \AA}$ ) está associada ao volume de exclusão imposto pela proteína. Essa descida inicial é seguida por uma fase de acúmulo, envolvendo interações específicas e inespecíficas entre soluto e solvente, que ocorrem na faixa de  $1,9\text{--}5,0 \text{ \AA}$ . A integral KB para o  $[\text{DCA}]^{-}$  converge para um valor positivo, indicando as fortes interações desse íon com a proteína, implicando em um acúmulo no domínio proteico suficiente para contrabalançar o volume de exclusão. Por sua vez, a integral KB para o  $[\text{EMIM}]^{+}$  neste sistema está próxima de zero, enquanto as integrais para  $[\text{Cl}]^{-}$  e água são negativas.

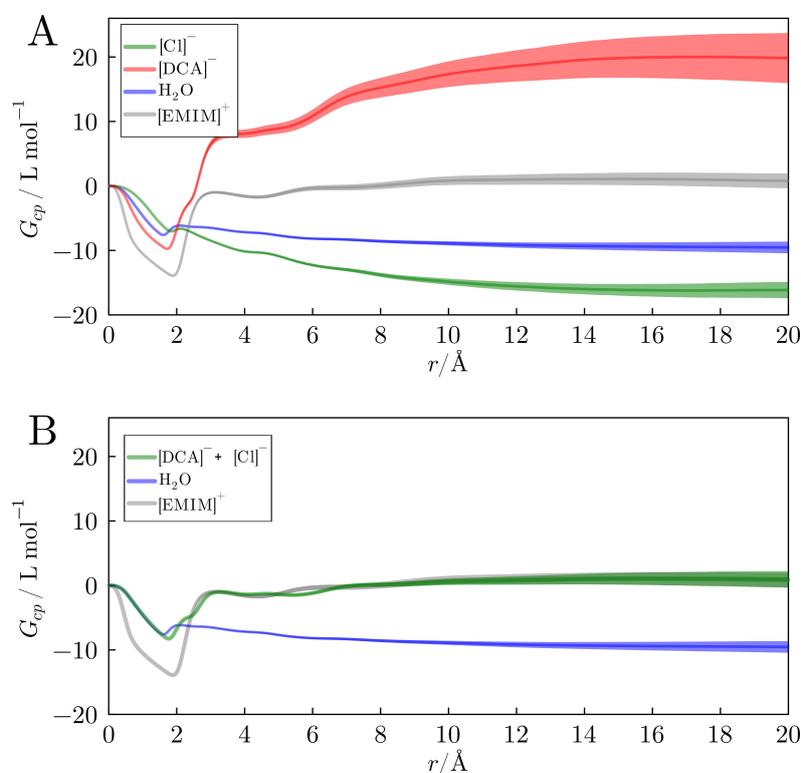


Figura 7.4: (A) integrais de KB da água,  $[\text{EMIM}]^{+}$ ,  $[\text{Cl}]^{-}$  e  $[\text{DCA}]^{-}$  em relação à proteína em sistemas com  $2,0 \text{ mol L}^{-1}$  de mistura de LIs. As linhas sólidas representam os valores médios de 20 réplicas de simulação, enquanto as áreas sombreadas indicam o erro padrão da média. (B) Integrais KB dos íons ( $[\text{EMIM}]^{+}$  e  $[\text{DCA}]^{-} + [\text{Cl}]^{-}$  sobrepostos) em comparação com os integrais da água. Os maiores valores de integrais KB para os íons em relação à água indicam que a proteína é preferencialmente solvatada pelos íons dos líquidos iônicos.

Como antecipado pela depleção dos íons  $[\text{Cl}]^{-}$  no MDDF da Figura 7.1, a integral de KB do  $[\text{Cl}]^{-}$  é negativa, indicando a sua exclusão do domínio da proteína na presença do  $[\text{DCA}]^{-}$ . O  $[\text{EMIM}]^{+}$ , sendo a única fonte de carga positiva, apresenta uma integral KB

intermediária entre as do  $[\text{DCA}]^-$  e do  $[\text{Cl}]^-$ . Um valor de integral KB igual a zero significa que o composto não está nem acumulado, nem depletado no domínio da proteína. Na Figura 7.4B, as integrais de KB consideram ambos os ânions como entidades indistinguíveis e, como esperado, as integrais de KB para cátions e ânions convergem para valores iguais.

A Figura 7.5A mostra as integrais de KB para os ânions em relação à proteína nas misturas de LI a  $2,0 \text{ mol L}^{-1}$ , enquanto a Figura 7.5B apresenta as integrais para a água nos mesmos sistemas. O  $[\text{DCA}]^-$  acumula-se mais nas proximidades da proteína, resultando nas maiores integrais de KB para os íons e nas menores para a água. A combinação de ânions com alta ( $[\text{DCA}]^-$ ) e baixa afinidade ( $[\text{Cl}]^-$ ) gera um acúmulo predominante de  $[\text{DCA}]^-$ , aumentando os parâmetros de interação preferencial do LI sobre a água. As integrais de KB da água, negativas nos sistemas da Figura 7.5B, indicam uma depleção substancial de moléculas de água em relação ao bulk. Em sistemas com maiores integrais de KB para ânions, a exclusão de água é mais acentuada do que naqueles com menores integrais de KB para ânions.

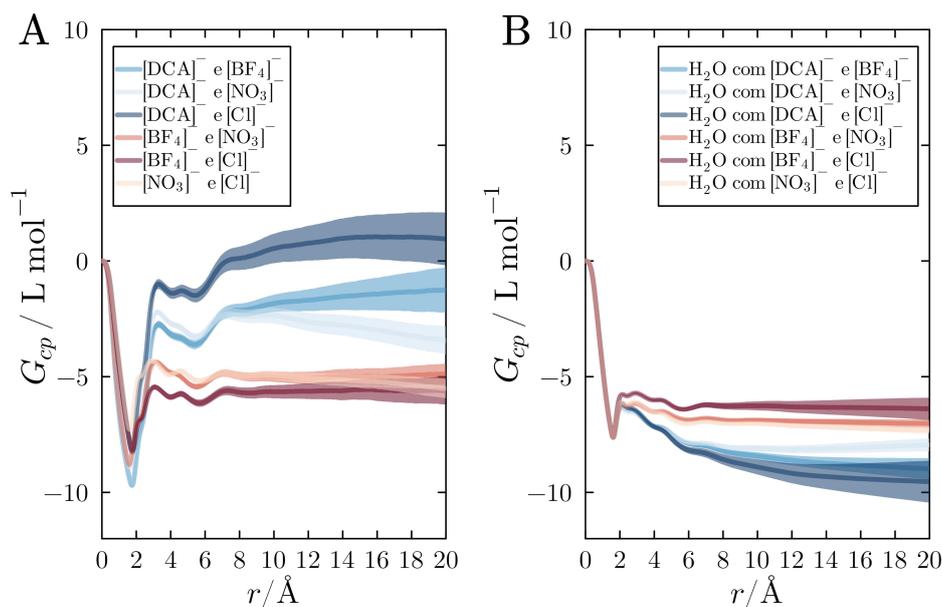


Figura 7.5: (A) Integrais de KB para ânions tratados como uma única entidade e (B) integrais de KB da água em sistemas com  $2,0 \text{ mol L}^{-1}$  de misturas de LIs. A presença de  $[\text{DCA}]^-$  eleva significativamente as integrais de KB em comparação aos demais sistemas. A combinação de  $[\text{DCA}]^-$  com ânions de menor afinidade resulta nos maiores valores de integrais de KB, destacando o efeito de acúmulo.

A Figura 7.6 ilustra  $\Gamma_{cp}$  para os íons em relação à proteína ao longo da faixa de concentrações simuladas, variando de  $0,5$  a  $3,0 \text{ mol L}^{-1}$  de LIs, com valores predominantemente positivos para sistemas com combinações de ânions  $[\text{DCA}]^- + [\text{Cl}]^-$ ,

$[\text{DCA}]^- + [\text{NO}_3]^-$  e  $[\text{DCA}]^- + [\text{BF}_4]^-$ , especialmente entre 0,5 e 2,0 mol L<sup>-1</sup>. As combinações  $[\text{BF}_4]^- + [\text{NO}_3]^-$  e  $[\text{NO}_3]^- + [\text{Cl}]^-$  também exibem parâmetros de solvatação preferencial positivo. Consequentemente, espera-se que esses LIs atuem como desnaturantes, promovendo conformações proteicas com maiores áreas de superfície exposta. Os sistemas com  $[\text{DCA}]^-$  mostram a maior redução na interação com a água, com a sequência  $[\text{DCA}]^- + [\text{Cl}]^-$ ,  $[\text{DCA}]^- + [\text{BF}_4]^-$  e  $[\text{DCA}]^- + [\text{NO}_3]^-$  apresentando, respectivamente, os maiores valores de  $\Gamma_{cp}$ , conforme ilustrado na Figura 7.6. No entanto, a 0,5 mol L<sup>-1</sup>, os valores de  $\Gamma_{cp}$  para as combinações com  $[\text{DCA}]^-$  são estatisticamente equivalentes, considerando o erro padrão calculado a partir das 20 simulações. Embora os efeitos específicos de diferentes combinações de ânions não sejam completamente claros, as que envolvem  $[\text{DCA}]^-$  resultam, em particular, em uma maior desidratação da proteína.

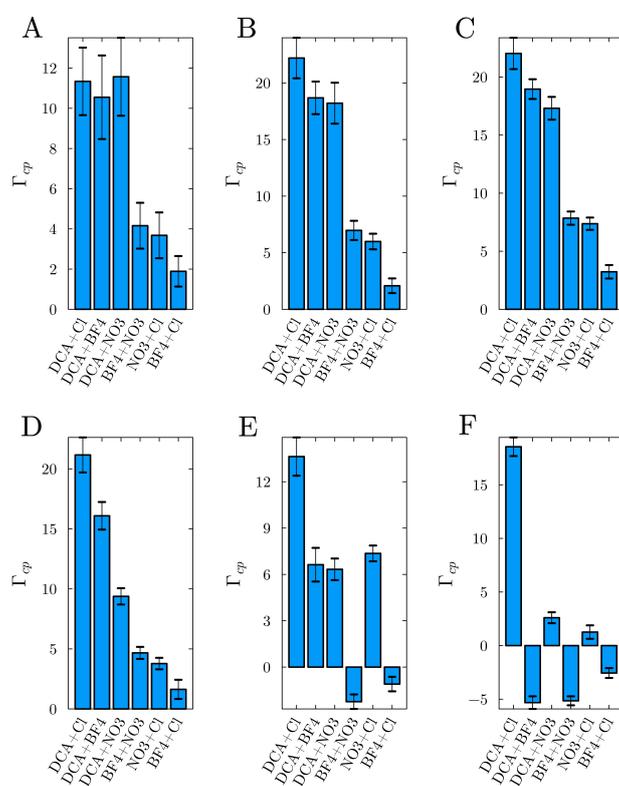


Figura 7.6: Parâmetros de solvatação preferencial do LI em relação à proteína em todas as concentrações simuladas com diferentes composições aniônicas: em soluções de (A) 0,5, (B) 1,0, (C) 1,5, (D) 2,0, (E) 2,5 e (F) 3,0 mol L<sup>-1</sup>, fornecendo uma análise comparativa do comportamento do LI em diferentes concentrações do soluto. O gráfico destaca as diferentes tendências dos íons em solvatar preferencialmente a proteína em vez da água na presença de íons concorrentes. Devido à eletroneutralidade da solução, esses parâmetros são idênticos aos calculados para o cátion ou para o conjunto de ânions em cada solução. A notação da mistura de ânions aparece simplificada de modo que: DCA é  $[\text{DCA}]^-$ , BF4 é  $[\text{BF}_4]^-$ , NO3 é  $[\text{NO}_3]^-$  e Cl é  $[\text{Cl}]^-$ .

Combinações como  $[\text{DCA}]^- + [\text{Cl}]^-$ ,  $[\text{DCA}]^- + [\text{NO}_3]^-$  e  $[\text{DCA}]^- + [\text{BF}_4]^-$  apresentam um acúmulo significativo de íons em comparação com outros sistemas mistos, embora esses níveis de acúmulo sejam geralmente menores do que os observados em sistemas contendo exclusivamente  $[\text{DCA}]^-$  em várias concentrações.

A análise da Figura 7.6 indica tendências gerais de solvatação preferencial nas menores concentrações (0,5 a 2,0 mol L<sup>-1</sup>), apesar de não serem triviais de elucidar completamente. A afinidade dos ânions pela proteína segue a sequência:  $[\text{DCA}]^- > [\text{NO}_3]^- > [\text{BF}_4]^- > [\text{Cl}]^-$ . Se as contribuições de afinidade dos ânions fossem aditivas, combinações de íons teriam afinidades que seguem tendência similar à da soma dos ânions individuais.

A natureza aditiva da afinidade é observada nas sequências de parâmetros de solvatação preferencial quando os ânions comuns são  $[\text{NO}_3]^-$ ,  $[\text{BF}_4]^-$ , e  $[\text{Cl}]^-$  até 2,0 mol L<sup>-1</sup>. Por exemplo, para  $[\text{NO}_3]^-$ , a sequência de solvatação preferencial é  $[\text{NO}_3]^- + [\text{DCA}]^- > [\text{NO}_3]^- + [\text{BF}_4]^- > [\text{NO}_3]^- + [\text{Cl}]^-$  (Figuras 7.6D–F, especialmente a 2,0 mol L<sup>-1</sup>). O mesmo padrão é observado para  $[\text{BF}_4]^-$  e  $[\text{Cl}]^-$ .

Contudo, para  $[\text{DCA}]^-$ , a solvatação preferencial exhibe uma sequência contrastante:  $[\text{DCA}]^- + [\text{Cl}]^- > [\text{DCA}]^- + [\text{BF}_4]^- > [\text{DCA}]^- + [\text{NO}_3]^-$ . Essa inversão, comparada a  $[\text{NO}_3]^-$ ,  $[\text{BF}_4]^-$ , e  $[\text{Cl}]^-$ , sugere um comportamento não aditivo de afinidade ao envolver  $[\text{DCA}]^-$ . Para entender o impacto das combinações de  $[\text{DCA}]^-$  com ânions de menor afinidade na ligação preferencial do LI à proteína, especulamos que um modelo de adsorção que inclua efeitos cooperativos e competitivos, bem como múltiplos tipos de sítios de ligação, seja necessário.

Em resumo, os resultados demonstram que as interações preferenciais e as propriedades moleculares da solvatação de solutos complexos, como proteínas, não podem ser deduzidas trivialmente a partir das características de ligação individuais de cada componente da mistura. Isso evidencia a complexidade na interpretação da termodinâmica de solvatação com base apenas na estrutura das espécies envolvidas [65, 178]. Modelar interações entre íons e proteínas requer mais do que suposições simplificadas sobre a natureza dos íons, já que as interações específicas entre os íons e o soluto desempenham um papel fundamental na determinação das propriedades globais da solução [179]. Por exemplo, a série de Hofmeister, embora útil para prever o efeito de sais na precipitação da lisozima,

mostra-se imprecisa em condições neutras e ácidas, refletindo um padrão de interação mais complexo que o proposto originalmente [179–182].

## Capítulo 8

# Efeitos da hidrofobicidade do cátion

Este capítulo apresenta um aspecto complementar às análises discutidas nos Capítulos 6 e 7, focando no efeito que a troca do cátion  $[\text{EMIM}]^+$  pelo cátion  $[\text{BMIM}]^+$  em sistemas contendo soluções aquosas de LIs formados pelos ânions  $[\text{DCA}]^-$ ,  $[\text{BF}_4]^-$ ,  $[\text{Cl}]^-$  e  $[\text{NO}_3]^-$  possui sobre a estrutura e termodinâmica de solvatação da ubiquitina. A análise é realizada visando investigar como o cátion  $[\text{BMIM}]^+$  influencia a estrutura de solvatação local e as interações entre os componentes do sistema. Os resultados obtidos revelam que o cátion  $[\text{BMIM}]^+$  promove uma maior exclusão de moléculas de água e apresenta um acúmulo mais significativo próximo à superfície proteica, especialmente em sistemas contendo ânions com forte capacidade de interação, como  $[\text{DCA}]^-$  e  $[\text{NO}_3]^-$ . Em concentrações mais baixas ( $\leq 1,5 \text{ mol L}^{-1}$ ), os LIs baseados em  $[\text{BMIM}]^+$  exibem uma solvatação preferencial mais acentuada quando comparados aos sistemas baseados em  $[\text{EMIM}]^+$ .

### 8.0.1 Hidrofobicidade dos cátions e estrutura local da solução

A análise da Figura 8.1 mostra que a densidade de  $[\text{BMIM}]^+$  é maior do que  $[\text{EMIM}]^+$ , independente do contra-íon a  $1,0 \text{ mol L}^{-1}$  de concentração de referência. A maior densidade é significativamente influenciada pelas correlações entre os ânions e cátions, que determinam as distribuições de ambos no sistema, como previamente discutido no capítulo 6 e apresentado em [164]. Por exemplo, na Figura 8.1A, onde o contra-íon é o cloreto ( $[\text{Cl}]^-$ ), as MDDFs dos cátions são bastante similares a partir de distâncias maiores que as da primeira camada de solvatação ( $r > 3,5 \text{ \AA}$ ). Em contraste, na Figura 8.1B, a MDDF de  $[\text{BMIM}]^+$  é maior que o de  $[\text{EMIM}]^+$  em um grande intervalo de distâncias. Tal comportamento é atribuído à forte interação entre  $[\text{DCA}]^-$  e os átomos da superfície da proteína, que aumenta a densidade de carga negativa próxima à proteína e promove o acúmulo de cátions na interface

proteína-solução, conforme discutido em [164, 166, 177].

As MDDFs presentes nas Figuras 8.1C e 8.1D mostram que a maior densidade de  $[\text{BMIM}]^+$  é perceptível apenas em distâncias curtas, com a grande diferença no pico das curvas vermelha e azul, comportamento qualitativamente semelhante ao observado com  $[\text{Cl}]^-$ . Tais resultados destacam a influência do contra-íon (neste caso, do ânion) nas interações entre os cátions e a superfície proteica, reforçando a relevância da composição do líquido iônico na solvatação local. Neste caso, a diferença entre as MDDFs dos cátions pareados com  $[\text{DCA}]^-$  é maior ao longo de longas distâncias.

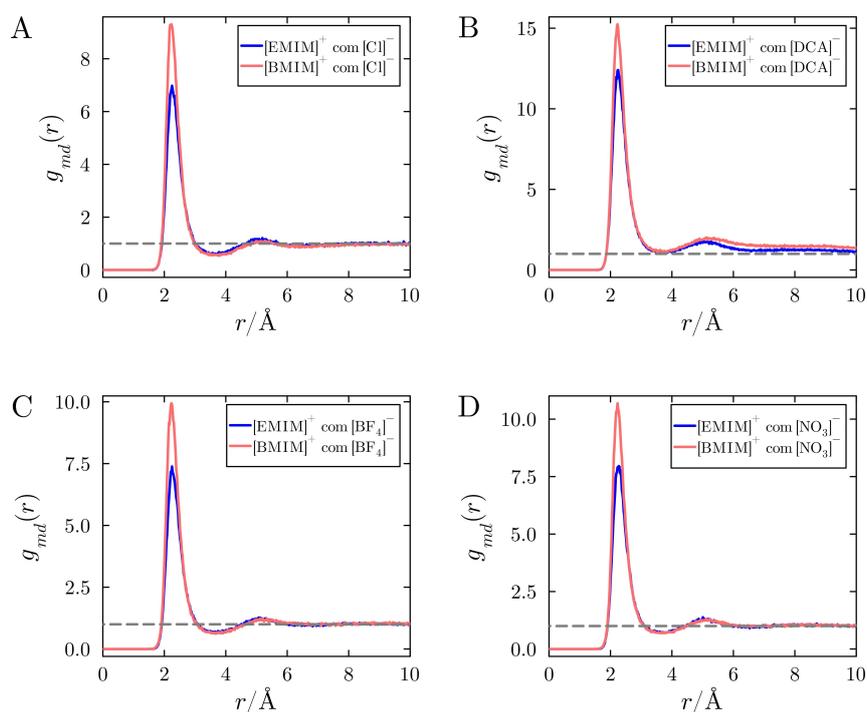


Figura 8.1: Comparação das MDDFs proteína-cátion para soluções de LIs contendo A)  $[\text{EMIM}][\text{Cl}]$  e  $[\text{BMIM}][\text{Cl}]$ , B)  $[\text{EMIM}][\text{DCA}]$  e  $[\text{BMIM}][\text{DCA}]$ , C)  $[\text{EMIM}][\text{BF}_4]$  e  $[\text{BMIM}][\text{BF}_4]$ , e D)  $[\text{EMIM}][\text{NO}_3]$  e  $[\text{BMIM}][\text{NO}_3]$ . Cada painel compara as MDDFs dos cátions  $[\text{EMIM}]^+$  (azul) e  $[\text{BMIM}]^+$  (vermelha) com diferentes ânions. Os dados representam valores médios de 20 simulações independentes em uma concentração de referência de  $1,0 \text{ mol L}^{-1}$ .

A Figura 8.2 apresenta as MDDFs dos ânions em relação à proteína, obtidas em soluções aquosas de  $1,0 \text{ mol L}^{-1}$  de LIs contendo  $[\text{EMIM}]^+$  ou  $[\text{BMIM}]^+$ . Os perfis observados para as MDDFs são consistentes com estudos anteriores [164, 166, 177]. Por exemplo, no caso de  $[\text{Cl}]^-$ , ilustrado na Figura 8.2A, o primeiro pico é menor em comparação aos outros ânions e deslocado para maiores distâncias, correspondendo às interações iônicas com resíduos básicos [144]. Por outro lado, ânions como  $[\text{DCA}]^-$ ,  $[\text{NO}_3]^-$  e até  $[\text{BF}_4]^-$  exibem interações mais

intensas com a superfície proteica, mediadas por ligações de hidrogênio, conforme indicado pelos picos em torno de 1,9 Å nas Figuras 8.2B-D.

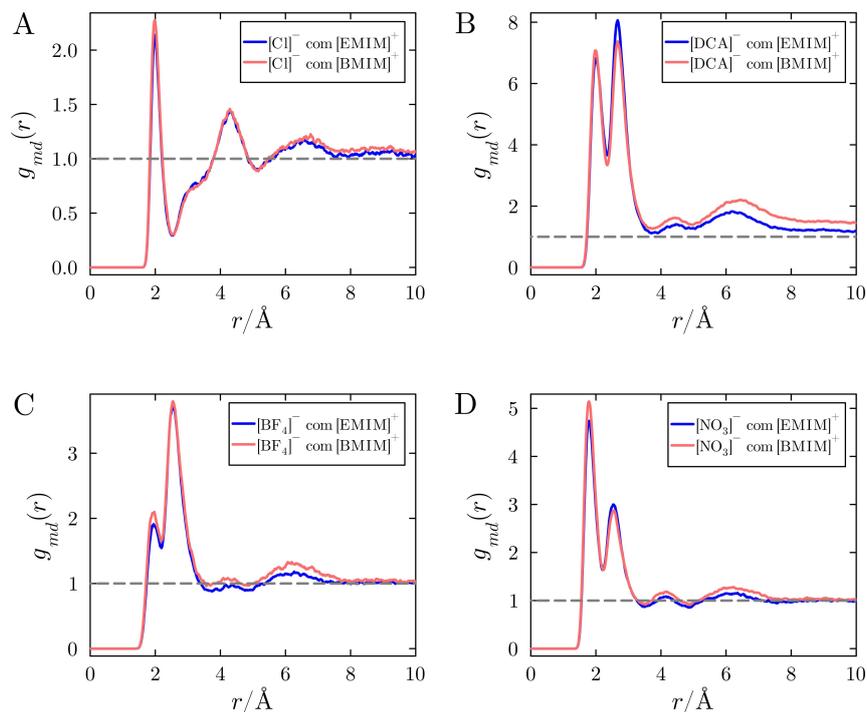


Figura 8.2: Comparação das MDDFs proteína-ânion para soluções de LIs contendo A) [EMIM][Cl] e [BMIM][Cl], B) [EMIM][DCA] e [BMIM][DCA], C) [EMIM][BF<sub>4</sub>] e [BMIM][BF<sub>4</sub>], e D) [EMIM][NO<sub>3</sub>] e [BMIM][NO<sub>3</sub>]. Cada painel apresenta as MDDFs dos ânions quando emparelhados com [EMIM]<sup>+</sup> (azul) ou [BMIM]<sup>+</sup> (vermelho). Dados obtidos de 20 simulações independentes em concentração de 1,0 mol L<sup>-1</sup>.

Embora a escolha do cátion não altere qualitativamente os formatos das MDDFs proteína-ânion, a densidade dos ânions nas proximidades da proteína é influenciada pelo cátion. Em soluções contendo [BMIM]<sup>+</sup>, as densidades dos ânions, mostradas na Figura 8.2, são geralmente maiores do que em soluções contendo [EMIM]<sup>+</sup>. Exceções notáveis ocorrem no segundo pico das distribuições de [DCA]<sup>-</sup> e [NO<sub>3</sub>]<sup>-</sup>. Assim, o maior comprimento da cadeia alquílica de [BMIM]<sup>+</sup> promove, indiretamente, aumento na densidade dos ânions. Este efeito, em sua maioria, independe da natureza química do ânion, sendo perceptível, na concentração de referência de 1,0 mol L<sup>-1</sup>, para os quatro ânions estudados.

As interações entre os cátions e a proteína são investigadas em maior detalhe por meio da decomposição das MDDFs (proteína-cátion) em contribuições dos diferentes tipos de resíduos da proteína. Esta decomposição foi realizada utilizando o recurso ResidueContributions do pacote ComplexMixtures.jl [161]. A Figura 8.3 compara os MDDFs dos cátions [EMIM]<sup>+</sup> e [BMIM]<sup>+</sup>, ambos com [DCA]<sup>-</sup>, decompostos por tipo de

resíduo, juntamente com um mapa de densidade mostrando as contribuições de cada resíduo específico (Figura 8.3C). Observa-se que os cátions interagem predominantemente com resíduos neutros, embora interações com resíduos polares, ácidos e básicos também ocorram, por meio de interações iônicas com resíduos ácidos e interações dispersivas com resíduos neutros devido às cadeias alifáticas.

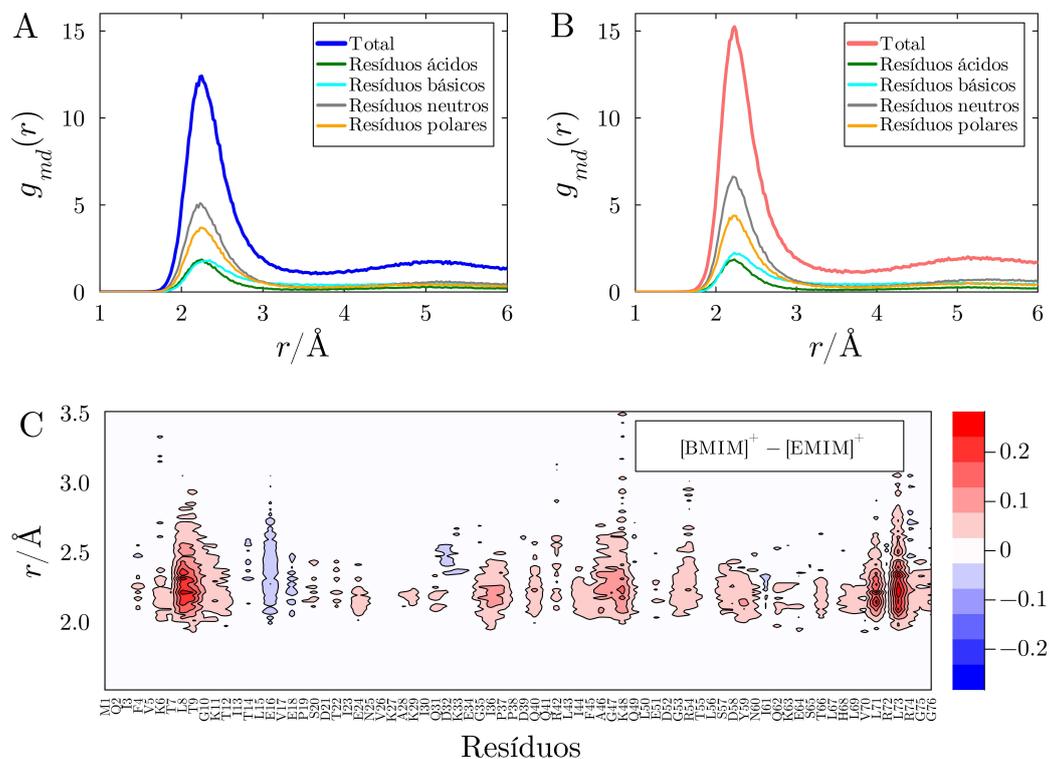


Figura 8.3: MDDFs proteína-cátion de A) [EMIM]<sup>+</sup> e B) [BMIM]<sup>+</sup> em soluções com [DCA]<sup>-</sup> a 1,0 mol L<sup>-1</sup>, decompostos em contribuições de resíduos básicos, ácidos, neutros e polares. C) Mapa de densidade diferencial de [BMIM]<sup>+</sup> menos [EMIM]<sup>+</sup> em até 3,5 Å da superfície proteica. Vermelho indica maior densidade de [BMIM]<sup>+</sup>, e azul, de [EMIM]<sup>+</sup>.

A Figura 8.3C mostra que [BMIM]<sup>+</sup> possui maior densidade relativa do que [EMIM]<sup>+</sup>, particularmente, próximo de resíduos hidrofóbicos e em algumas áreas básicas específicas. Este comportamento é atribuído à maior hidrofobicidade do cátion [BMIM]<sup>+</sup> e também à sua correlação com ânions, que aumenta a densidade relativa de [BMIM]<sup>+</sup> próximo de outros tipos de resíduos. A contribuição predominante de certos grupos de resíduos indica uma maior probabilidade de encontrar os cátions distribuídos próximos a esses resíduos, refletindo preferências estruturais que influenciam a solvatação.

### 8.0.2 Competição com água e solvatação preferencial

As MDDFs demonstram que, embora o perfil das curvas permaneça consistente, há variação na densidade dos ânions ao longo da solução quando estão emparelhados com  $[\text{EMIM}]^+$  ou  $[\text{BMIM}]^+$ . A avaliação do acúmulo ou depleção de cossolventes ao redor de um soluto é realizada por meio da integração das funções de distribuição, conforme discutido ao longo dos capítulos desta tese.

A Figura 8.4 apresenta as integrais de KB para os ânions em relação à proteína, emparelhados com dois cátions distintos, em soluções de  $1,0 \text{ mol L}^{-1}$ . Esta análise complementa os resultados apresentados nas MDDFs da Figura 8.2, permitindo interpretar os dados em termos de acumulação ou depleção dos ânions no domínio proteico. Na Figura 8.4A, observa-se que as integrais de KB para  $[\text{Cl}]^-$  em soluções com  $[\text{EMIM}]^+$  e  $[\text{BMIM}]^+$  convergem para valores negativos. Contudo, quando o contra-íon é  $[\text{BMIM}]^+$ , o valor da integral de KB é maior em comparação ao observado com  $[\text{EMIM}]^+$ .

O comportamento descrito acima, qualitativamente, também é observado para os demais ânions analisados: os ânions pareados com  $[\text{BMIM}]^+$  apresentam as integrais de KB calculadas em relação à proteína que convergem para valores mais elevados do que as integrais associadas a  $[\text{EMIM}]^+$ . Estes resultados reforçam a influência do contra-íon  $[\text{BMIM}]^+$  no aumento da acumulação dos ânions próximos à superfície proteica, em comparação ao cátion  $[\text{EMIM}]^+$ .

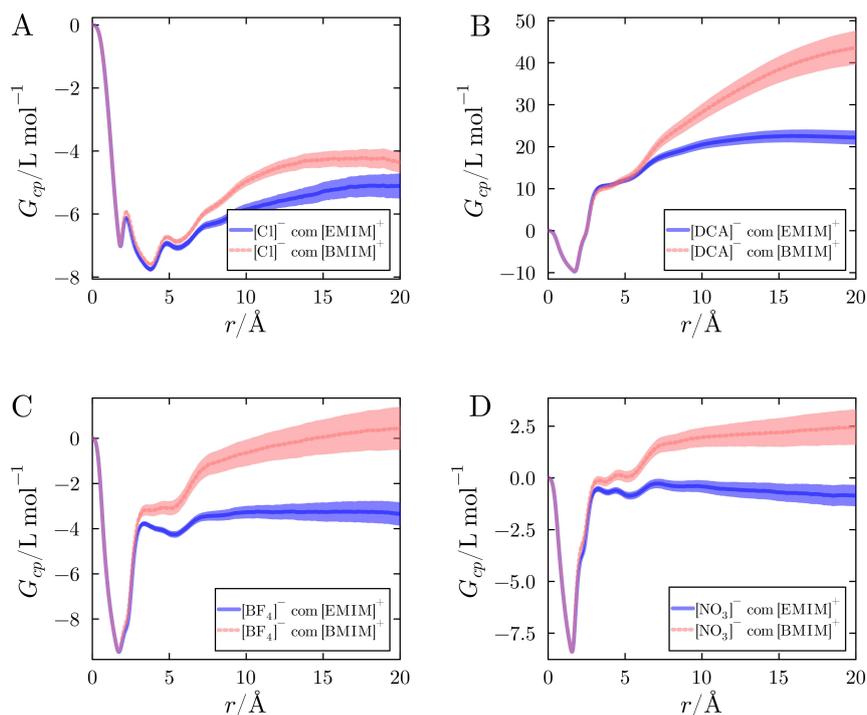


Figura 8.4: Integrais de KB para ânions em relação à proteína em soluções aquosas de LIs com  $[\text{EMIM}]^+$  ou  $[\text{BMIM}]^+$  a  $1,0 \text{ mol L}^{-1}$ . A)  $[\text{Cl}]^-$ , B)  $[\text{DCA}]^-$ , C)  $[\text{BF}_4]^-$ , e D)  $[\text{NO}_3]^-$ . As linhas sólidas representam a média das integrais KB de 20 simulações, enquanto as áreas sombreadas indicam o erro padrão da média.

As MDDFs e integrais de KB para água em sistemas contendo  $[\text{Cl}]^-$  e  $[\text{DCA}]^-$ , ambos na concentração de  $1,0 \text{ mol L}^{-1}$ , são apresentadas na Figura 8.5. Nos sistemas com  $[\text{Cl}]^-$ , a MDDF de água é maior quando o cátion é  $[\text{EMIM}]^+$ , em comparação ao sistema com  $[\text{BMIM}]^+$ , até aproximadamente  $8 \text{ \AA}$ . Após o pico inicial, correspondente às ligações de hidrogênio, ambos os sistemas mostram uma depleção de água nas proximidades da superfície proteica, sendo essa depleção mais acentuada no sistema contendo  $[\text{BMIM}]^+$ . A diferença reflete o maior acúmulo de  $[\text{BMIM}]^+$  próximo à superfície proteica, resultado de sua maior hidrofobicidade, reduzindo a densidade relativa de água nos sistemas com  $[\text{BMIM}]^+$ .

Além disso, os resultados observados são influenciados por uma combinação de fatores. Por exemplo, a troca do cátion afeta as MDDFs da água de maneira distinta nos sistemas com  $[\text{Cl}]^-$  e  $[\text{DCA}]^-$ , evidenciando que o comportamento da água depende não apenas do cátion, mas também da natureza do ânion presente no sistema.

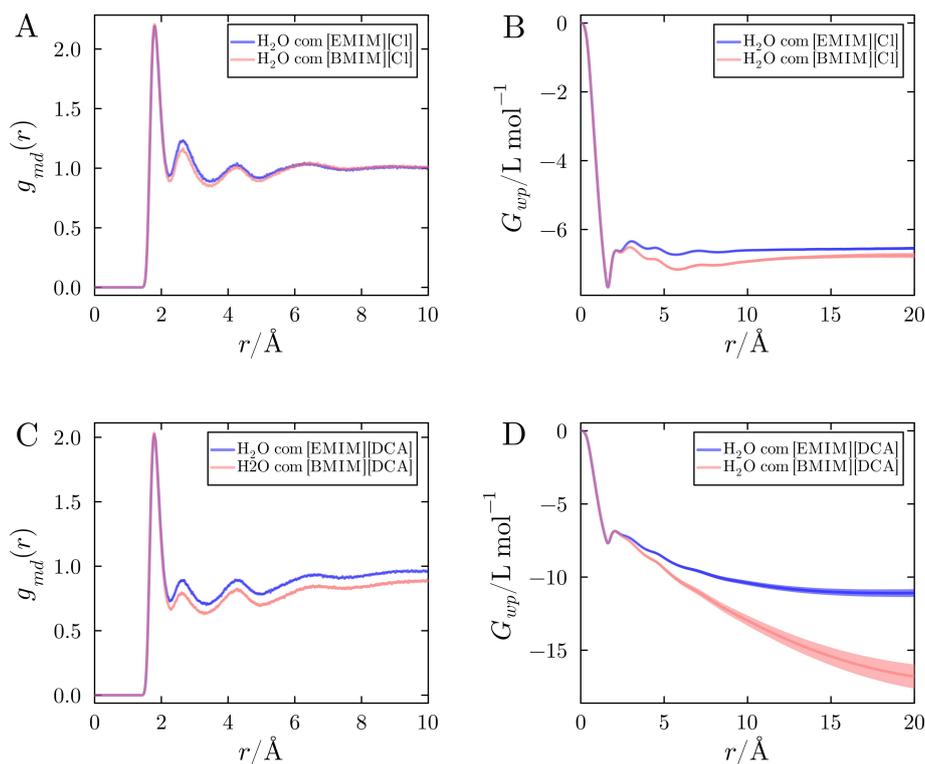


Figura 8.5: Hidratação proteica em sistemas de líquidos iônicos contendo [EMIM][Cl], [EMIM][DCA], [BMIM][Cl] e [BMIM][DCA] a  $1,0 \text{ mol L}^{-1}$ . A) MDDF proteína-água em soluções com cloreto e B) integrais de KB correspondentes. C) MDDF proteína-água em soluções com  $[\text{DCA}]^-$  e D) integrais de KB correspondentes.

### 8.0.3 Dependência da concentração na solvatação preferencial

A Figura 8.6 ilustra a dependência da concentração na distribuição de cátions em relação à proteína, usando como exemplo [EMIM][DCA] nas concentrações de  $1,0$  e  $3,0 \text{ mol L}^{-1}$ . Em  $1,0 \text{ mol L}^{-1}$  (Figura 8.6A), observa-se que o pico da MDDF em  $2,4 \text{ Å}$  para o cátion [EMIM]<sup>+</sup> é menor do que o de [BMIM]<sup>+</sup>, indicando maior probabilidade de interações estabelecidas por [BMIM]<sup>+</sup> com a superfície proteica. Essa tendência é corroborada pelas integrais de KB (Figura 8.6C), que mostram maior acumulação efetiva de [BMIM]<sup>+</sup> na primeira camada de solvatação. Em contraste, a  $3,0 \text{ mol L}^{-1}$  (Figura 8.6B), essa tendência se inverte: o pico da MDDF para [EMIM]<sup>+</sup> supera ligeiramente o de [BMIM]<sup>+</sup>, levando a maiores valores de integrais de KB para [EMIM]<sup>+</sup> (Figura 8.6D).

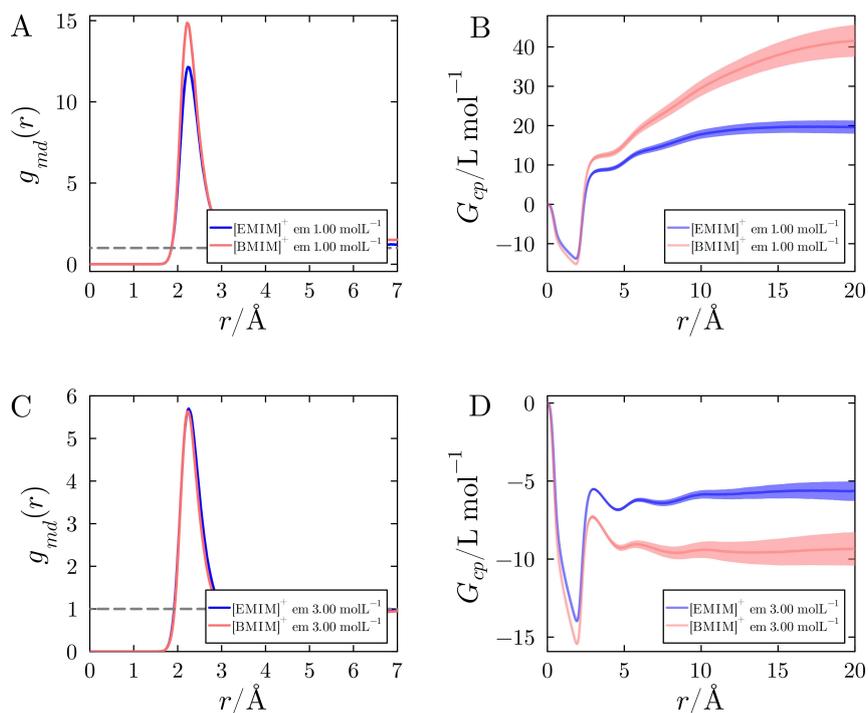


Figura 8.6: MDDFs proteína-cátion para soluções aquosas de [EMIM][DCA] em A) 1,0 mol L<sup>-1</sup> e B) 3,0 mol L<sup>-1</sup>. C) e D) Integrais de KB correspondentes às concentrações respectivas. Dados representam médias de 20 réplicas de simulação.

A Figura 8.7 apresenta os parâmetros de solvatação preferencial ( $\Gamma_{cp}$ ) para diferentes composições e concentrações de LIs. Em baixas concentrações ( $\leq 1,5$  mol L<sup>-1</sup>), os  $\Gamma_{cp}$  de [BMIM]<sup>+</sup> excedem os de [EMIM]<sup>+</sup>, especialmente em sistemas com ânions [DCA]<sup>-</sup> e [NO<sub>3</sub>]<sup>-</sup>. Contudo, essa tendência se reverte em concentrações mais altas ( $\geq 2,5$  mol L<sup>-1</sup>), onde os  $\Gamma_{cp}$  de [EMIM]<sup>+</sup> tornam-se predominantes.

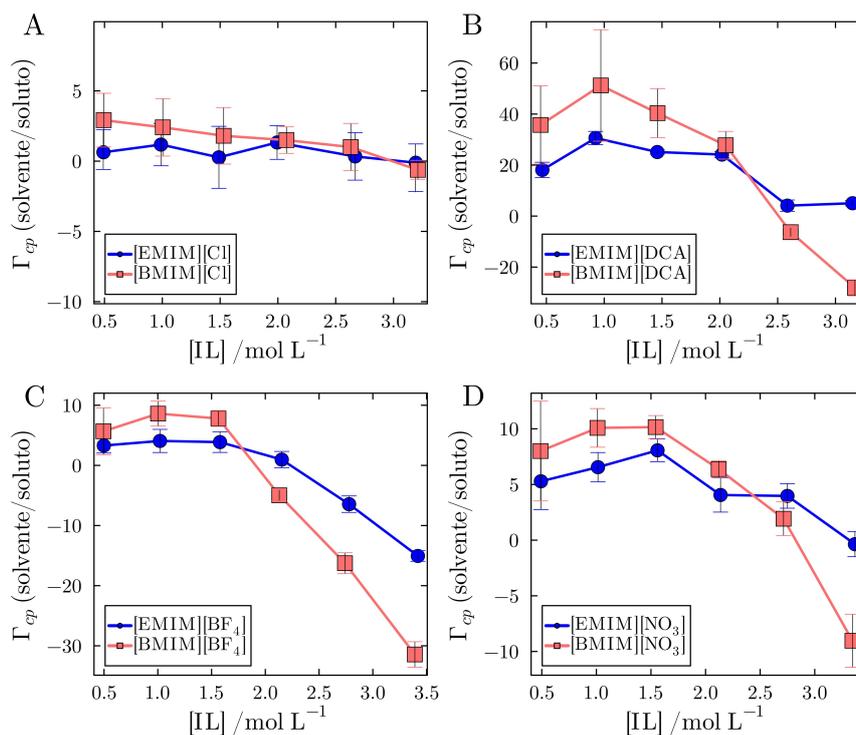


Figura 8.7: Parâmetros de solvatação preferencial ( $\Gamma_{cp}$ ) para diferentes composições e concentrações de LIs. LIs contendo [BMIM]<sup>+</sup> apresentam maior solvatação preferencial em concentrações baixas, enquanto em concentrações elevadas [EMIM]<sup>+</sup> predomina. Barras de erro representam o desvio padrão de 20 réplicas de simulação.

Em concentrações baixas, os LIs desidratam a proteína, atuando, numa primeira aproximação, como agentes desnaturantes. Entretanto, à medida que a concentração aumenta, os valores de  $\Gamma_{cp}$  diminuem, indicando uma maior hidratação preferencial da proteína, especialmente em sistemas com [BMIM]<sup>+</sup>. Apesar dessa mudança, o efeito agregado dos cossolventes sobre a estrutura proteica depende fortemente das interações com os estados desnaturados, como mostrado em capítulos anteriores. Assim, os LIs estudados favorecem, em diferentes graus, a desnaturação proteica em todas as concentrações avaliadas, mas com uma dinâmica de solvatação que varia com a concentração e a composição dos cátions e ânions.

#### 8.0.4 Número de coordenação e conteúdo de água

Em baixas concentrações, os LIs solvata preferencialmente a proteína, impulsionados por interações de curto alcance entre os íons e a superfície proteica. Contudo, em concentrações mais altas, ocorre uma inversão na solvatação preferencial: a proteína torna-se predominantemente hidratada, e [EMIM]<sup>+</sup> supera [BMIM]<sup>+</sup> em capacidade de

solvatação. Esses fenômenos são explicados pela análise dos números de coordenação dos cátions e pela concentração de água em cada sistema.

A Figura 8.8 ilustra a variação dos números de coordenação ( $CN$ , do inglês *coordination number*) de  $[EMIM]^+$  e  $[BMIM]^+$  na primeira camada de solvatação com a concentração de LI. Os dados destacados referem-se aos sistemas contendo  $[DCA]^-$  e  $[BF_4]^-$ , onde os efeitos da concentração sobre a solvatação preferencial são mais pronunciados. Em concentrações mais baixas,  $[BMIM]^+$  apresenta maior número de coordenação com a proteína, explicando os maiores parâmetros de interação preferencial e maior desidratação. Com o aumento da concentração, as diferenças nos números de coordenação diminuem, tornando-se semelhantes em torno de  $2 \text{ mol L}^{-1}$ . Para os ânions, observa-se uma tendência sutil: em concentrações baixas, o número de coordenação é maior com  $[BMIM]^+$ , enquanto em concentrações mais altas ocorre o inverso.

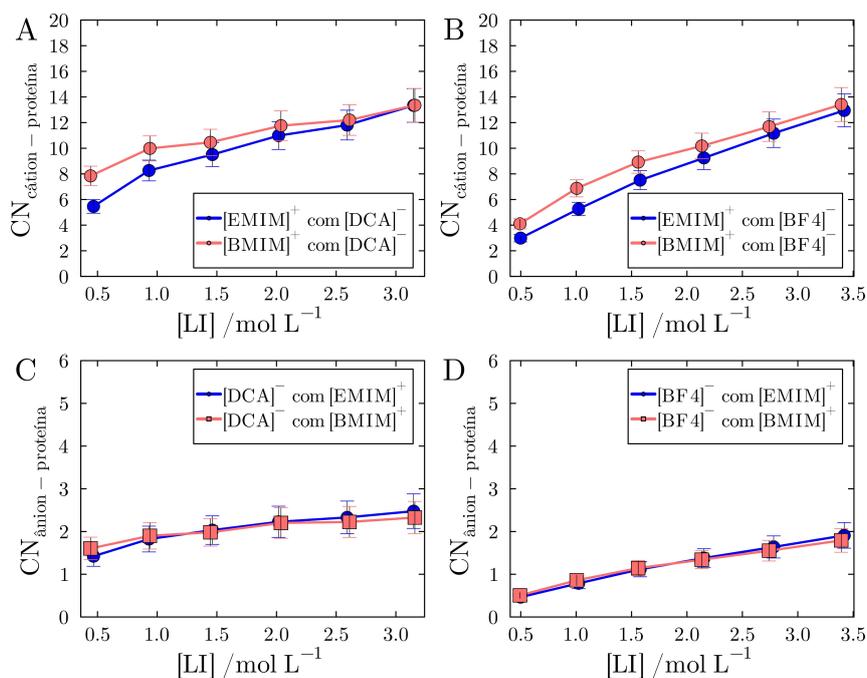


Figura 8.8: Números de coordenação de íons na primeira camada de solvatação em função da concentração de LI. A) e B) Números de coordenação de cátions em 3,5 Å em soluções com  $[DCA]^-$  e  $[BF_4]^-$ . C) e D) Números de coordenação de ânions em 2,5 Å em soluções com  $[DCA]^-$  e  $[BF_4]^-$ , respectivamente. Barras de erro representam o erro padrão da média.

A Figura 8.9 detalha os números de coordenação dependentes da distância para os cátions até 5 Å da superfície proteica, avaliados para sistemas com  $[DCA]^-$  em diferentes concentrações. Em concentrações mais baixas (Figuras 8.9A–C),  $[BMIM]^+$  apresenta números de coordenação maiores que  $[EMIM]^+$ , consistentes com sua maior afinidade pela

superfície proteica devido à hidrofobicidade. No entanto, à medida que a concentração aumenta (Figuras 8.9D–F), as curvas tornam-se semelhantes, indicando saturação dos sítios de interação específica para cátions. Assim, o número total de cátions no domínio proteico não determina a solvatação preferencial de  $[\text{EMIM}]^+$  em concentrações mais altas.

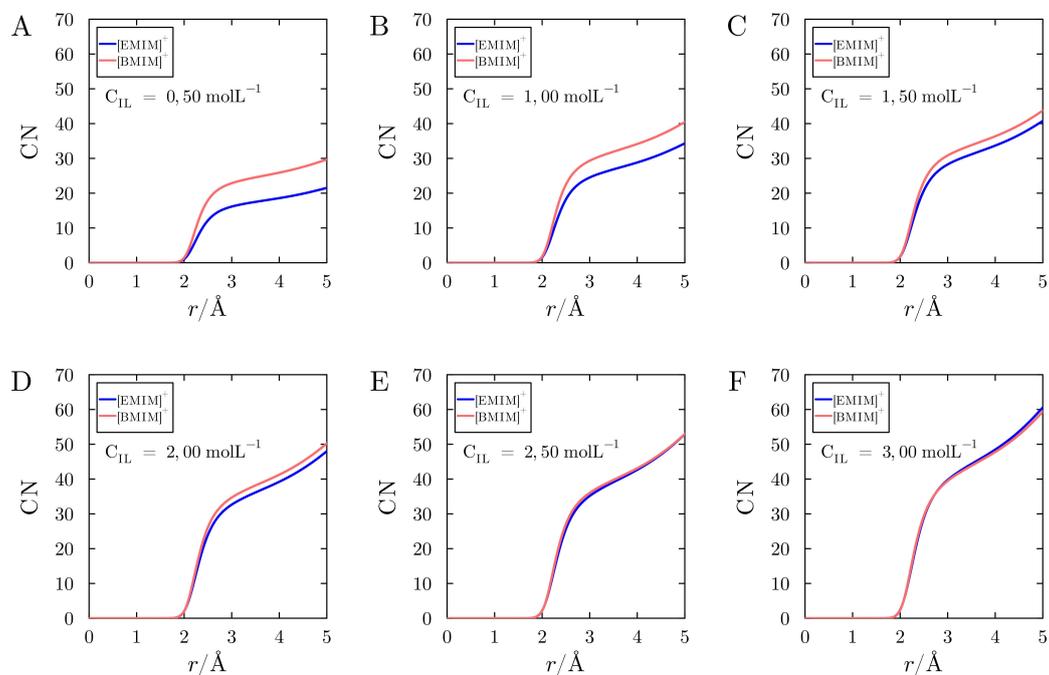


Figura 8.9: Números de coordenação de cátions até 5 Å da superfície proteica em diferentes concentrações para sistemas com  $[\text{EMIM}][\text{DCA}]$  ou  $[\text{BMIM}][\text{DCA}]$ . Painéis A–F apresentam os dados para concentrações de referência de 0,5 a 3,0 mol L<sup>-1</sup>.

A Tabela 8.1 mostra as concentrações das espécies químicas em sistemas com  $[\text{EMIM}][\text{DCA}]$  e  $[\text{BMIM}][\text{DCA}]$ . Em concentrações de 0,5 mol L<sup>-1</sup>, o conteúdo de água é similar entre os cátions. Entretanto, a 3,0 mol L<sup>-1</sup>, a quantidade de água é significativamente menor nos sistemas com  $[\text{BMIM}]^+$  de cerca de 27 mol L<sup>-1</sup> com  $[\text{EMIM}]^+$  para aproximadamente 22 mol L<sup>-1</sup> com  $[\text{BMIM}]^+$ . Essa menor concentração de água nos sistemas com  $[\text{BMIM}]^+$  implica menor competição com a proteína, resultando em uma solvatação preferencial inferior para  $[\text{BMIM}]^+$  em concentrações mais altas.

Tabela 8.1: Concentrações em solução ( $\text{mol L}^{-1}$ ) das espécies químicas em sistemas contendo [EMIM][DCA] e [BMIM][DCA] preparados com concentrações de referência ( $C_{IL}$ ) de 0,5 e 3,0  $\text{mol L}^{-1}$ . Os subscritos *E* e *B* indicam, respectivamente, que a espécie em destaque está num com sistema com [EMIM]<sup>+</sup> ou [BMIM]<sup>+</sup>.

$C_{IL}$	[EMIM] <sup>+</sup>	[BMIM] <sup>+</sup>	[DCA] <sup>-</sup> <sub>E</sub>	[DCA] <sup>-</sup> <sub>B</sub>	Água <sub>E</sub>	Água <sub>B</sub>
0.50	0.469±0.003	0.444±0.006	0.464±0.003	0.439±0.006	50.70±0.02	50.09±0.06
3.00	3.150±0.006	3.16±0.02	3.148±0.007	3.16±0.02	27.86±0.05	22.1±0.2

Em concentrações mais altas, a semelhança nos números de coordenação para [EMIM]<sup>+</sup> e [BMIM]<sup>+</sup> indica saturação dos sítios de ligação da proteína específicos para cátions. Assim, o fator determinante para a solvatação preferencial torna-se a natureza volumosa de [BMIM]<sup>+</sup>, que reduz a concentração de água em sistemas com alta molaridade de LI.

## Capítulo 9

# Solvatação de diferentes estados de enovelamento

Neste capítulo, a solvatação de diferentes estados de enovelamento da ubiquitina é investigada em soluções aquosas contendo quatro LIs, formados pela combinação dos cátions [EMIM]<sup>+</sup> e [BMIM]<sup>+</sup> com os ânions [BF<sub>4</sub>]<sup>-</sup> e [DCA]<sup>-</sup>. Na maioria dos sistemas, os LIs apresentam solvatação preferencial das estruturas proteicas, caracterizando-os como agentes desnaturantes. Entretanto, mesmo em cenários onde a hidratação preferencial favorece a estrutura nativa, a desnaturação é predominante devido à forte interação preferencial dos LIs com estados desnaturados.

### 9.1 Detalhes específicos da metodologia

Para as simulações discutidas aqui, são considerados quatro estados de enovelamento distintos, cada um composto por um conjunto de 20 estruturas com características estruturais semelhantes, com exceção do estado estendido, composto por apenas uma estrutura.

#### 9.1.1 Estados de enovelamento estudados

- **Estado Nativo:** Corresponde às conformações obtidas a partir da estrutura cristalográfica (PDB: 1UBQ [178]), obtidas por meio de equilibrações da temperatura e pressão. Estas estruturas possuem um RMSD médio de 1,76 Å em relação ao modelo cristalográfico, preservando 95% dos contatos nativos, como apresentado pela Tabela 9.1. O software *VMD* foi utilizado para identificar o núcleo proteico [165], composto exclusivamente por resíduos hidrofóbicos: metionina (1 resíduo), isoleucina (8 resíduos), fenilalanina (2 resíduos), valina (5 resíduos), leucina (9 resíduos) e alanina

(2 resíduos). A desnaturação expõe predominantemente esses resíduos apolares ao solvente.

- **Estado Perturbado:** Gerado por simulações com duração de 50 ns da proteína solvatada em água a 500 K, resultando em conformações com RMSD médio de 3,6 Å em relação à estrutura cristalográfica e preservação de aproximadamente 75% dos contatos nativos. Estas estruturas representam estados parcialmente desnaturados, mantendo o formato globular e parte das estruturas secundária e terciária nativas.
- **Estado Desnaturado:** Composto por estruturas obtidas através simulação da ubiquitina no vácuo a 600 K por 100 ns. Conformações espaçadas em intervalos de 5 ns foram extraídas e submetidas a simulações adicionais, para relaxar as estruturas obtidas em água. Estas estruturas apresentaram RMSD médio de 12 Å e preservaram apenas 37% dos contatos nativos.
- **Estado Estendido:** Representa o limite superior de exposição superficial e desnaturação. Construído como uma cadeia linear com ângulos de torção  $\phi = -180^\circ$  e  $\psi = 180^\circ$ , utilizando o plugin *Molefactory* do *VMD*. Para esta conformação, os resíduos de prolina foram substituídos por alaninas para evitar restrições estruturais e possibilitar a estrutura final sendo uma reta.

A Tabela 9.1 contém propriedades estruturais médias para o conjunto de estruturas que forma cada estado de enovelamento simulado.

Tabela 9.1: Propriedades estruturais médias e desvios padrão das conformações utilizadas: área de superfície acessível ao solvente (SASA), raio de giro, RMSD em relação ao modelo cristalográfico e fração de contatos nativos (NC).

Estado	SASA ( $10^2 \text{ \AA}^2$ )	Raio de Giro (nm)	RMSD (Å)	NC
Nativo	$49.70 \pm 0.05$	$1.179 \pm 0.002$	$1.76 \pm 0.02$	$0.95 \pm 0.01$
Perturbado	$53.0 \pm 0.4$	$1.221 \pm 0.008$	$3.6 \pm 0.02$	$0.75 \pm 0.03$
Desnaturado	$74 \pm 9$	$1.292 \pm 0.007$	$12.15 \pm 0.06$	$0.37 \pm 0.01$
Estendido	117.4	8.01	76.7	0

A Figura 9.1 apresenta exemplos representativos das estruturas nativas, perturbadas, desnaturadas e estendida. As estruturas do estado nativo (Figura 9.1A) preservam as características do modelo cristalográfico, enquanto as estruturas perturbadas (Figura 9.1B) mostram perda parcial dos elementos estrutura secundária. As estruturas desnaturadas, por sua vez, (Figura 9.1C) são caracterizadas pela perda quase total das

estruturas secundárias, mantendo formas globulares. Já a conformação estendida (Figura 9.1D) representa um estado extremo de desnaturação, sem similaridade estrutural com os outros estados.

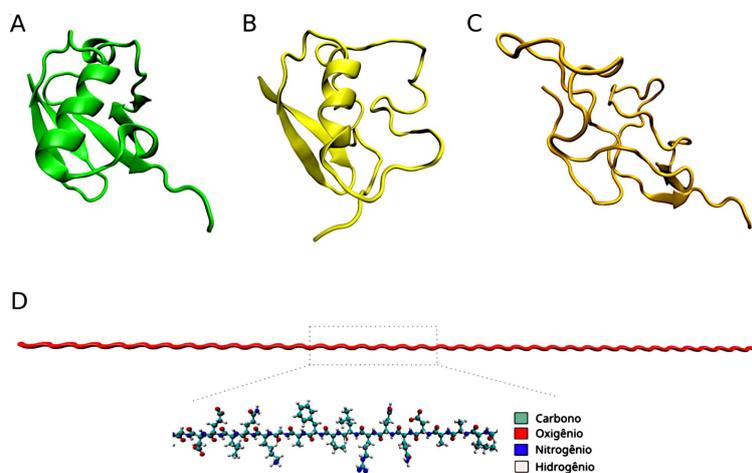


Figura 9.1: Modelos representativos das conformações A) nativa, B) perturbada, C) desnaturada e D) estendida. As estruturas progressivamente perdem características nativas, com aumento da exposição do núcleo hidrofóbico ao solvente.

### 9.1.2 Configurações da dinâmica molecular dos diferentes estados

Os potenciais utilizados para os cálculos de ligação e não-ligação foram derivados do campo de força OPLS para a ubiquitina e do campo de força OPLS-VSIL para os líquidos iônicos (LIs) [103]. O modelo TIP3P foi empregado para representar as moléculas de água [183]. O campo de força OPLS-VSIL foi parametrizado utilizando um sítio virtual para fixar uma carga no centro do anel imidazólico dos cátions dos LIs, representando a deslocalização de carga característica destes compostos. Assim, mantendo a abordagem com respeito ao campo de força igual a dos capítulos anteriores.

Para as simulações no estado nativo, o software GROMACS 2018.3.CUDA foi utilizado [91]. Inicialmente, foi realizada uma minimização de energia utilizando o método de *Steepest Descent*, limitado a um máximo de 50.000 passos. Durante esta etapa, a proteína foi mantida fixa por meio de restrições de movimento [184].

Na etapa subsequente, foi conduzida uma equilibração no ensemble NVT durante 1 ns, empregando o termostato modificado de Berendsen [112]. Nesta etapa, a temperatura foi mantida em 300 K, e o raio de corte para interações de curta distância (eletrostáticas e de van der Waals) foi fixado em 10 Å. O cálculo das interações eletrostáticas de longa distância

foi realizado utilizando o método *Particle Mesh Ewald* [123], com um passo de tempo de 2 fs. Esta fase teve como objetivo a termalização do solvente ao redor da proteína.

Após a equilibração térmica, a pressão foi equilibrada utilizando o ensemble NPT (300 K e 1 bar) por 5 ns, novamente com o uso do termostato modificado de Berendsen [112] e do barostato de Parinello-Rhman [114, 115]. Durante essa etapa, a proteína permaneceu fixa. Em seguida, uma dinâmica molecular de 1 ns foi realizada com o sistema livre, culminando em uma simulação de produção de 10 ns. Para cada combinação de LI e concentração, foram simuladas 20 réplicas, abrangendo seis concentrações diferentes. Este procedimento foi necessário para assegurar a convergência das integrais de Kirkwood-Buff, que requerem amostragem extensa e estatisticamente robusta.

As simulações da proteína na conformação estendida e das conformações derivadas das simulações no vácuo seguiram protocolos similares aos descritos para a proteína no estado nativo. Os sistemas foram submetidos a minimização e equilibração no ensemble NVT, com a proteína mantida fixa no centro da caixa e apenas as cadeias laterais permitidas a se moverem. Simulações adicionais no ensemble NPT foram realizadas por períodos de 5 e 10 ns, seguidas por uma fase de produção de 10 ns.

## 9.2 Estrutura de solvatação

As Figuras 9.2A e 9.2B apresentam, respectivamente, as MDDFs e as integrais de KB do ânion  $[\text{DCA}]^-$  em soluções contendo aproximadamente  $1,5 \text{ mol L}^{-1}$  de  $[\text{EMIM}][\text{DCA}]$ , considerando diferentes estados conformacionais da ubiquitina. Na Figura 9.2A, observa-se o pico referente às ligações de hidrogênio em  $1,9 \text{ \AA}$ , onde a densidade local do  $[\text{DCA}]^-$  é até cinco vezes maior em relação ao estado de referência, para a conformação nativa (curva azul). Este pico torna-se gradualmente menos pronunciado com o desenovelamento da proteína: para os estados perturbado, desnaturado e estendido, os valores máximos dos picos de ligação de hidrogênio são 4,1, 3,8 e 3,6, respectivamente. Por sua vez, o pico em  $2,5 \text{ \AA}$ , que está relacionado às interações não específicas, exibe um comportamento oposto: a densidade local do  $[\text{DCA}]^-$  aumenta em estruturas com maior área de superfície exposta.

A Figura 9.2B mostra as integrais de KB para o ânion  $[\text{DCA}]^-$ . Em curtas distâncias ( $r < 1,5 \text{ \AA}$ ), observa-se uma queda atribuída ao volume excluído pela proteína, assim, os estados com estruturas de maior SASA apresentam maior exclusão inicial. Após

esta região, ocorre um acúmulo progressivo devido às interações específicas e não específicas no intervalo de 1,9 a 5,0 Å. O pico difuso das MDDFs em 6 Å também contribui significativamente para os valores finais dos integrais de KB. Os valores convergidos indicam que o acúmulo de  $[DCA]^-$  no domínio proteico aumenta com a área de superfície exposta. Por exemplo, a integral de KB relativa ao conjunto nativo (curva azul) é negativo, sugerindo que o acúmulo em longas distâncias não compensa a exclusão inicial. Por outro lado, para os conjuntos desnaturado, perturbado e estendido, as integrais são positivas, indicando maior concentração de  $[DCA]^-$  no domínio proteico em comparação ao *bulk*.

Na Figura 9.2C, as MDDFs para o cátion  $[EMIM]^+$  apresentam um único pico em 2,4 Å, característico do cátion  $[EMIM]^+$  nos sistemas sendo estudados nesta tese. Para a conformação estendida, o pico é menos pronunciado em comparação aos estados desnaturados. Além disso, as curvas MDDF para os estados perturbado e estendido são praticamente idênticas, sugerindo que as interações cátions-proteína são semelhantes para essas conformações.

As Figuras 9.2B e 9.2D ilustram as integrais de KB para  $[DCA]^-$  e  $[EMIM]^+$ , respectivamente. Em ambos os casos, as integrais aumentam à medida que a estrutura proteica se desenrola, devido à maior área de superfície exposta (Eq. 5.2.6). Ao comparar as MDDFs e as integrais de KB dos sistemas nativo e estendido (curvas azul e roxa), verifica-se que, embora as interações médias dos íons com a superfície das conformações estendidas sejam menos favoráveis, o maior acúmulo total de íons no domínio proteico se deve à ampliação da área de superfície disponível.

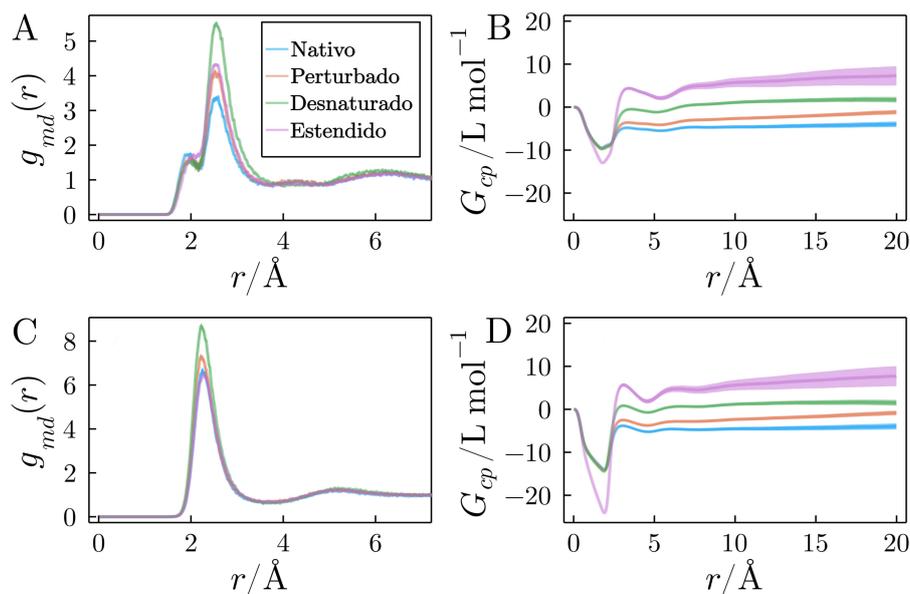


Figura 9.2: (A) MDDFs e (B) integrais de KB para  $[\text{DCA}]^-$  em relação à ubiquitina em diferentes estados conformacionais em soluções de  $1,5 \text{ mol L}^{-1}$  de  $[\text{EMIM}][\text{DCA}]$ . (C) MDDFs e (D) integrais de KB para  $[\text{EMIM}]^+$  na mesma solução. Em (A), observa-se que a densidade local de  $[\text{DCA}]^-$  em distâncias características de ligações de hidrogênio diminui à medida que a proteína se desenrola, enquanto a densidade local em interações não específicas aumenta. Em (C), as MDDFs de  $[\text{EMIM}]^+$  exibem picos mais altos para os estados desnaturados; a curva MDDF para o estado estendido (roxa) sobrepõe-se quase completamente à do estado perturbado (laranja). As Figuras (B) e (D) apresentam as integrais de KB para  $[\text{DCA}]^-$  e  $[\text{EMIM}]^+$ , respectivamente, que aumentam com o desenovelamento da estrutura proteica.

As MDDFs foram decompostas em contribuições dos tipos de resíduos que compõem a ubiquitina, conforme apresentado na Figura 9.3. Resíduos polares e básicos contribuem significativamente para o pico associado às ligações de hidrogênio ( $1,9 \text{ Å}$ ). O segundo pico ( $2,6 \text{ Å}$ ) também é influenciado por tais tipos de resíduos, mas a contribuição dos resíduos neutros, predominantemente alifáticos, é substancial. À medida que o desenovelamento da proteína avança, observa-se uma redução no pico de ligação de hidrogênio e um aumento no pico de interações não específicas. Como consequência, a participação de resíduos neutros na MDDF total aumenta consideravelmente com o progresso da desnaturação.

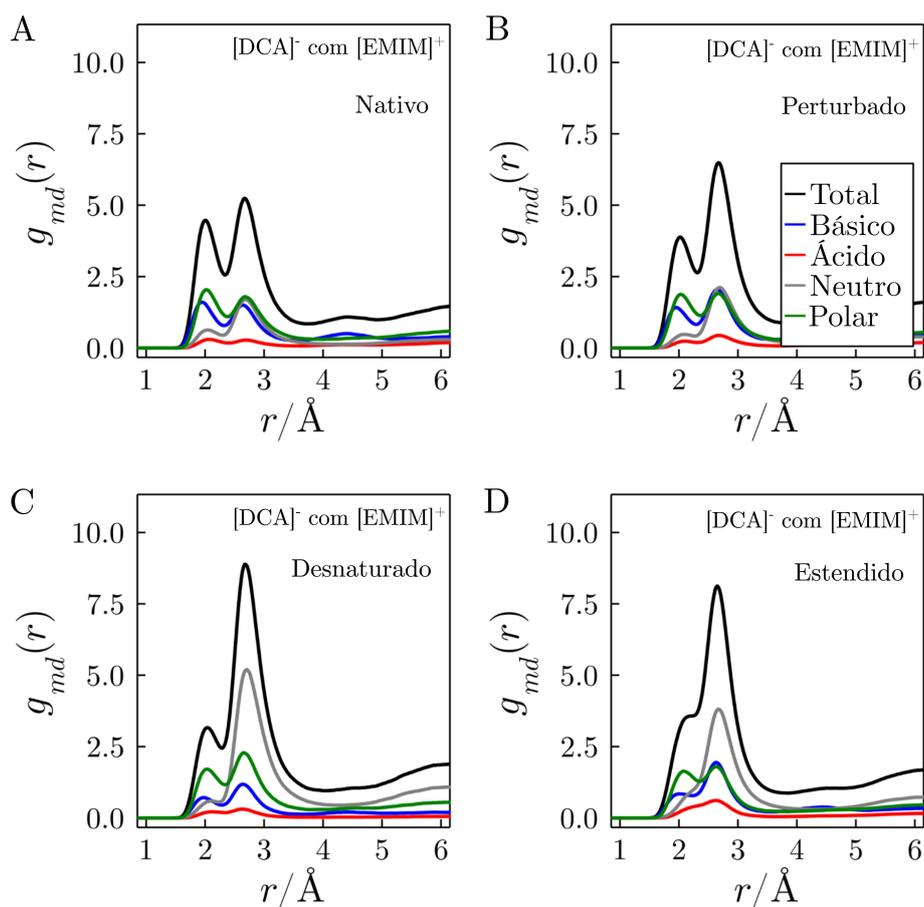


Figura 9.3: Efeito do estado de enovelamento nas funções de distribuição do DCA, destacando as contribuições de cada classe de resíduos em uma solução com  $1,5 \text{ mol L}^{-1}$  de Li. (A), (B), (C) e (D) correspondem, respectivamente, às distribuições do DCA com conjuntos de estruturas nativa, perturbada, desnaturada e estendida. O pico ( $1,9 \text{ \AA}$ ) reflete maior acúmulo em resíduos polares e básicos, enquanto o segundo pico ( $2,6 \text{ \AA}$ ) está associado ao acúmulo em resíduos neutros.

A Figura 9.4 apresenta as contribuições dos tipos de resíduos de aminoácidos para as MDDFs do cátion  $[\text{EMIM}]^+$ . Nas conformações nativas, resíduos polares e neutros têm participação predominante. Com o desenovelamento progressivo da proteína, resíduos neutros, especialmente alifáticos, tornam-se os principais contribuidores. Estruturas com maior área de superfície exposta disponibilizam esses resíduos para interações não específicas com o cátion  $[\text{EMIM}]^+$ , conforme evidenciado pelo aumento da contribuição de resíduos neutros da Figura 9.4A até 9.4D.

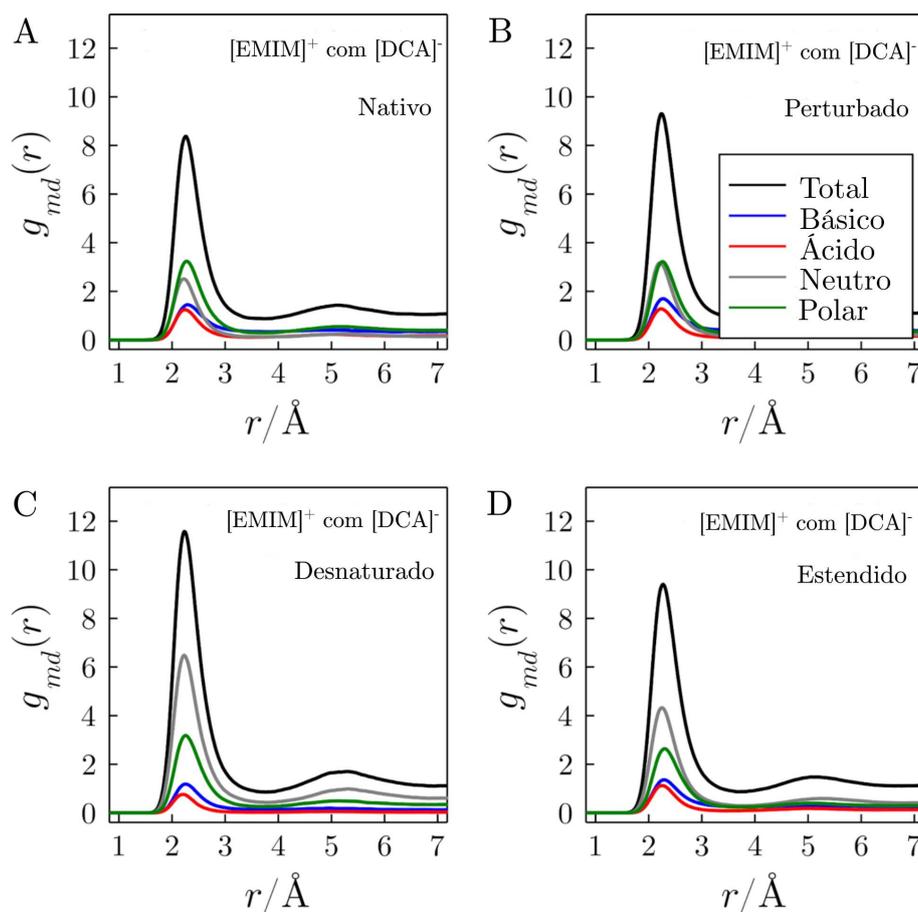


Figura 9.4: Efeito do estado de enovelamento nas funções de distribuição do  $[\text{EMIM}]^+$ , destacando as contribuições de cada classe de resíduos em uma solução com  $1,5 \text{ mol L}^{-1}$  de LI. (A), (B), (C) e (D) correspondem, respectivamente, às distribuições do  $[\text{EMIM}]^+$  com conjuntos de estruturas nativa, perturbada, desnaturada e estendida. O pico ( $2,4 \text{ \AA}$ ) reflete maior acúmulo em resíduos neutros, proporcional ao desenovelamento da proteína.

As análises das MDDFs (Figuras 9.2A e 9.2C) já demonstraram que os íons interagem de maneira distinta com a proteína. A Figura 9.5 apresenta um mapa de densidade ilustrando a distribuição dos íons ao redor dos resíduos proteicos. Regiões em laranja indicam maior densidade de  $[\text{DCA}]^-$  em relação a  $[\text{EMIM}]^+$ , enquanto as regiões em azul representam maior densidade de  $[\text{EMIM}]^+$  em comparação a  $[\text{DCA}]^-$ . As áreas em laranja próximas a  $1,9 \text{ \AA}$  estão associadas ao pico de ligação de hidrogênio do  $[\text{DCA}]^-$  (Figura 9.2A), concentrando-se principalmente em resíduos como M1, K6, K11, Q31, R42, K48, Q49 e R74, sendo majoritariamente polares e carregados positivamente. Por outro lado, resíduos hidrofóbicos e carregados negativamente, como I3, T7, L8, D32, I44 e L73, predominam em regiões onde a concentração de  $[\text{EMIM}]^+$  é maior.

Um ponto de destaque é a representação ocasional de concentrações mais elevadas

de  $[\text{DCA}]^-$  próximas a resíduos negativos, como E16. Essa distribuição reflete uma proximidade espacial dos resíduos no sistema e não interações predominantes.

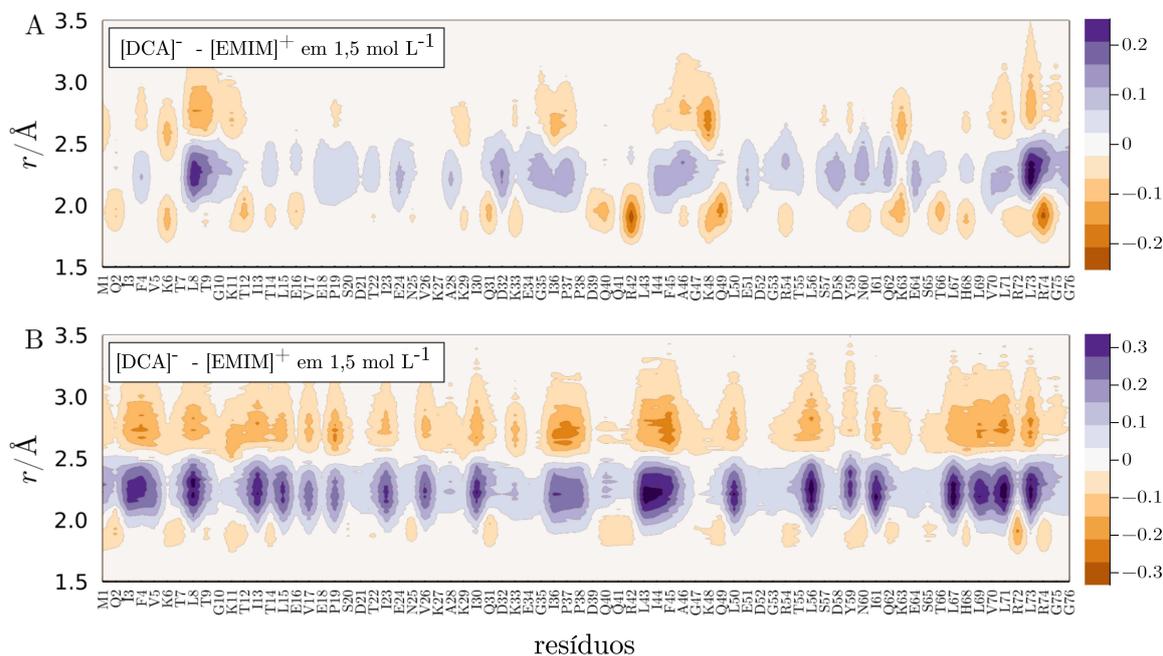


Figura 9.5: Mapa de densidade dos íons ao redor dos resíduos proteicos. Regiões em laranja indicam maior densidade de  $[\text{DCA}]^-$  em relação ao  $[\text{EMIM}]^+$ , enquanto regiões em azul representam maior densidade de  $[\text{EMIM}]^+$  em comparação ao  $[\text{DCA}]^-$ . Resíduos polares e básicos estão associados ao pico de ligação de hidrogênio (1,9 Å), enquanto resíduos hidrofóbicos e carregados negativamente predominam em regiões de maior concentração de  $[\text{EMIM}]^+$ .

A Figura 9.6 ilustra o efeito da troca do ânion de  $[\text{DCA}]^-$  para  $[\text{BF}_4]^-$  nas MDDFs e nas integrais de KB para o  $[\text{EMIM}]^+$  em sistemas com soluções de  $1,5 \text{ mol L}^{-1}$  de LI. As MDDFs do  $[\text{EMIM}]^+$  mantêm a mesma forma, independentemente do ânion (Figura 9.6A) ou da conformação da proteína (comparar Figuras 9.6A e 9.6B). No entanto, a altura do pico em  $2,4 \text{ Å}$  é maior para  $[\text{EMIM}]^+$  com  $[\text{DCA}]^-$  do que para  $[\text{EMIM}]^+$  com  $[\text{BF}_4]^-$ . A densidade de  $[\text{EMIM}]^+$  com  $[\text{DCA}]^-$  é 8,5 vezes maior do que no *bulk* a  $2,4 \text{ Å}$  da superfície proteica nativa (Figura 9.6A) e 12 vezes maior próximo a estruturas desnaturadas (Figura 9.6B).

As integrais de KB revelam dois pontos principais: primeiro, que o  $[\text{EMIM}]^+$  se acumula mais quando o sistema contém  $[\text{DCA}]^-$ ; segundo, que a acumulação de  $[\text{EMIM}]^+$  no domínio proteico é maior para estruturas com maior área de superfície.

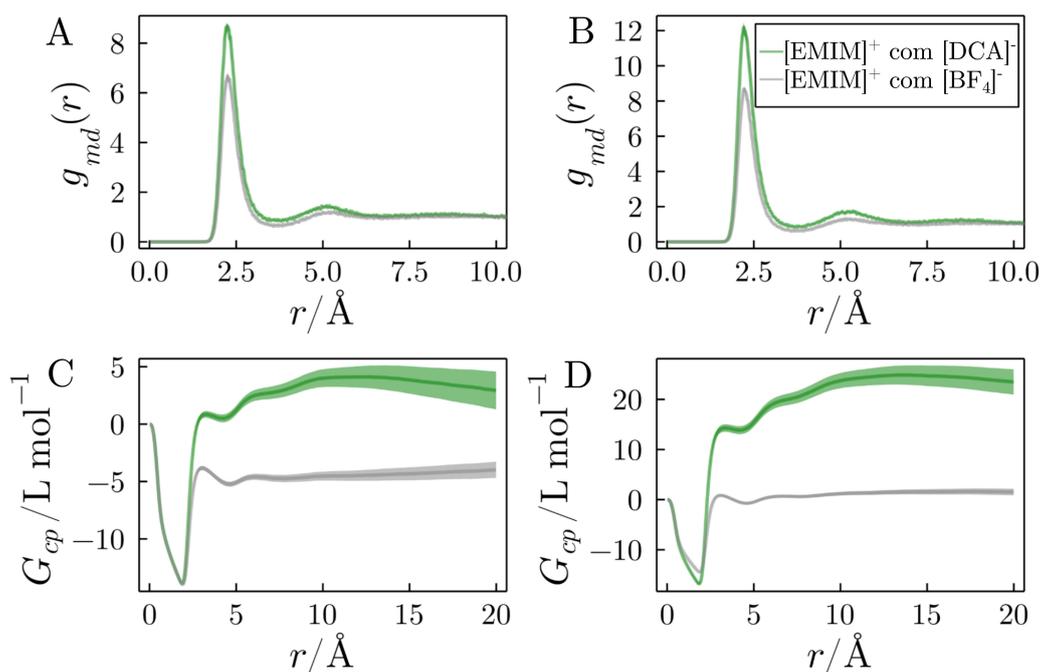


Figura 9.6: MDDFs e integrais de KB para  $[\text{EMIM}]^+$  em sistemas com estruturas nativas (A e C) e desnaturadas (B e D) com soluções de  $1,5 \text{ mol L}^{-1}$  de  $[\text{EMIM}][\text{DCA}]$  (curva verde) e  $[\text{EMIM}][\text{BF}_4]$  (curva cinza). As MDDFs mostram que o  $[\text{DCA}]^-$  induz um aumento maior na concentração de  $[\text{EMIM}]^+$  nas proximidades da proteína em comparação ao  $[\text{BF}_4]^-$ .

Como mostrado na Figura 9.6, o  $[\text{DCA}]^-$  promove um aumento na concentração de  $[\text{EMIM}]^+$  próximo à proteína. O  $[\text{DCA}]^-$ , um ânion que forma contatos mais fortes com a superfície proteica, promove claramente uma maior concentração de  $[\text{EMIM}]^+$  nas regiões proximais à proteína. Como ilustrado na Figura 9.5, a concentração de LI é maior perto de resíduos neutros. Em estruturas mais abertas, ocorrem interações não específicas entre proteína e LI, cuja magnitude varia entre  $[\text{DCA}]^-$  e  $[\text{BF}_4]^-$ , resultando em maior concentração de  $[\text{EMIM}]^+$  ao redor de resíduos proteicos com  $[\text{DCA}]^-$  em relação a  $[\text{BF}_4]^-$ .

O tipo de cátion também influencia a distribuição do ânion. A presença do cátion  $[\text{BMIM}]^+$  aumenta as MDDFs e as integrais de KB. De fato, as integrais de KB para o  $[\text{DCA}]^-$  no sistema com  $[\text{BMIM}]^+$  são substancialmente maiores do que com  $[\text{EMIM}]^+$ , tanto para conformações nativas quanto desnaturadas. O cátion  $[\text{BMIM}]^+$  é maior e mais hidrofóbico do que o  $[\text{EMIM}]^+$ , aumentando as interações não específicas entre a proteína e o cátion. Devido à correlação entre a distribuição do cátion e do ânion, o ânion se concentra mais ao redor da proteína na presença do cátion  $[\text{BMIM}]^+$ . Como resultado, o  $[\text{BMIM}]^+$  causa uma acumulação mais efetiva do ânion  $[\text{DCA}]^-$  nos domínios proteicos em comparação ao  $[\text{EMIM}]^+$ . Parâmetros de solvatação preferencial maiores são obtidos,

particularmente para estruturas mais desnaturadas, conferindo ao líquido iônico [BMIM][DCA] o caráter desnaturante mais forte entre as soluções de LI estudadas [166].

### 9.3 Solvatação por água em soluções de líquidos iônicos

A Figura 9.7A apresenta as MDDFs da água em relação à ubiquitina em soluções de  $1,5 \text{ mol L}^{-1}$  de [EMIM][DCA] para cada conjunto estrutural. O primeiro pico, em  $1,8 \text{ \AA}$ , apresenta um aumento em mais de duas vezes na densidade de água em relação ao estado de referência, por exemplo, para as conformações nativas (curva azul). A altura dos picos, assim, varia conforme o estado conformacional da ubiquitina: o primeiro pico torna-se menos pronunciado conforme a proteína se desenovela, devido ao aumento da área superficial da proteína, que expõe resíduos menos polares do que aqueles presentes na superfície da proteína nativa, em média. Para as conformações perturbadas, estendidas e desnaturadas, o pico da MDDF correspondente às ligações de hidrogênio apresenta valores máximos de 1,7, 1,4 e 1,2, respectivamente. Já o pico em  $2,5 \text{ \AA}$ , associado à segunda camada de hidratação da proteína, não varia significativamente entre os diferentes conjuntos estruturais.

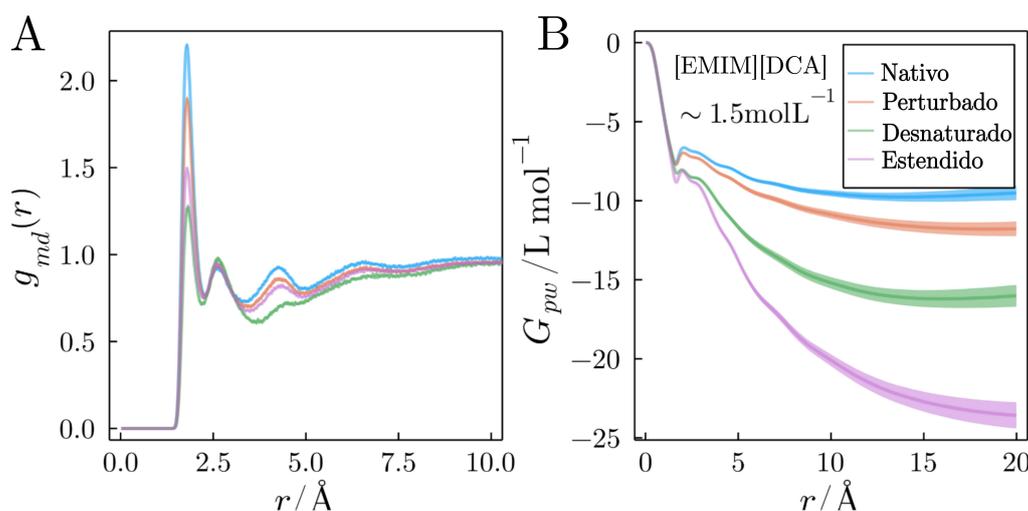


Figura 9.7: (A) MDDFs e (B) integrais de KB para água em relação à ubiquitina em diferentes estados conformacionais em soluções de  $1,5 \text{ mol L}^{-1}$  de [EMIM][DCA]. Em (A), observa-se que o aumento da densidade local de água em distâncias características de ligações de hidrogênio diminui conforme a proteína se desenovela, o que está associado ao aumento da área superficial hidrofóbica. O efeito da maior área hidrofóbica é mais evidente em (B), onde a exclusão progressiva da água do domínio proteico é representada pelas integrais KB.

A Tabela 9.2 apresenta a proporção de ligações de hidrogênio entre a proteína e a água em soluções de  $1,5 \text{ mol L}^{-1}$  de [EMIM][DCA], bem como a área superficial acessível ao

solvente (SASA) média das conformações e o número de ligações de hidrogênio por unidade de área. Observa-se que há mais ligações de hidrogênio por unidade de SASA nas conformações mais enoveladas da proteína. Por exemplo, o número de ligações de hidrogênio por unidade de área é de 2,81 e 1,55 para os modelos nativo e desnaturado, respectivamente, em soluções de [EMIM][DCA]. Assim, as moléculas de água interagem relativamente mais por meio de ligações de hidrogênio com as estruturas proteicas mais enoveladas.

Tabela 9.2: Número de ligações de hidrogênio (HB, do inglês *hydrogen bonds*) entre a proteína e a água, proteína-proteína (intramolecular) e proteína-ânion em sistemas com 1,5 mol L<sup>-1</sup> de [EMIM][DCA].

Estrutura	HB água/SASA	HB água	HB proteína-proteína	HB proteína-ânion
Nativa	2.81	140.2	56.9	11.4
Perturbada	2.40	124.8	59.7	9.7
Desnaturada	1.55	115.3	60.0	11.0
Estendida	1.70	195.4	6.9	20.8

A Figura 9.7B apresenta as integrais de KB para água em soluções de 1,5 mol L<sup>-1</sup> de [EMIM][DCA]. Observa-se que quanto maior a área superficial da proteína, menor é a integral de KB para a água. Isso ocorre porque as ligações de hidrogênio na camada de solvatação inicial são insuficientes para compensar os volumes excluídos e a depleção de água em distâncias maiores. Assim, à medida que a estrutura se afasta da forma nativa, a proteína é progressivamente desidratada.

A análise adicional mostra que a água é mais excluída do domínio proteico em sistemas com [DCA]<sup>-</sup> do que com [BF<sub>4</sub>]<sup>-</sup>, como evidenciado pelas integrais de KB e pelo número de ligações de hidrogênio (Tabela 9.2). Isso é atribuído às interações mais fortes entre o ânion [DCA]<sup>-</sup> e a proteína. Comparativamente, os valores de HB/SASA para água em soluções de [EMIM][BF<sub>4</sub>] são maiores do que em soluções de [EMIM][DCA], indicando que a água interage mais favoravelmente com proteínas em presença de [BF<sub>4</sub>]<sup>-</sup>.

Os resultados apresentados são consistentes com estudos prévios, como o realizado por Diddens et al. [185], que investigaram a estabilidade de peptídeos em soluções aquosas de LIs. Esses autores demonstraram que interações dispersivas de [EMIM]<sup>+</sup> com peptídeos favorecem a acumulação de cátions em torno de conformações nativas, enquanto ânions interagem com a camada de [EMIM]<sup>+</sup> por correlações eletrostáticas. Esses achados

complementam a presente análise, onde as interações específicas entre  $[\text{DCA}]^-/[\text{BF}_4]^-$  e a proteína emergem como forças motrizes principais para o aumento da solvatação preferencial dos LIs. A maior capacidade do  $[\text{DCA}]^-$  de formar ligações de hidrogênio com a proteína, em comparação ao  $[\text{BF}_4]^-$ , explica sua solvatação mais forte. Como  $[\text{DCA}]^-$  e  $[\text{BF}_4]^-$  possuem tamanhos similares, a diferença reside predominantemente nas interações específicas com a proteína.

## 9.4 Solvatação preferencial em diferentes estados

A Figura 9.8 apresenta  $\Gamma_{cp}$  para os LIs nos sistemas com os diferentes estados de enovelamento estudados. Para praticamente todos os sistemas, o valor de  $\Gamma_{cp}$  aumenta de 0,5 para 1,5 mol L<sup>-1</sup> e, em seguida, diminui a partir de 1,5 mol L<sup>-1</sup>. Quando  $\Gamma_{cp} > 0$ , os cossolventes (líquidos iônicos) solvatam preferencialmente a proteína. Exceções notáveis são algumas estruturas nativas em concentrações mais altas de LIs: em B, C e D, as estruturas nativas exibem parâmetros de solvatação negativos para LIs em concentrações mais elevadas, indicando que a proteína é preferencialmente hidratada.

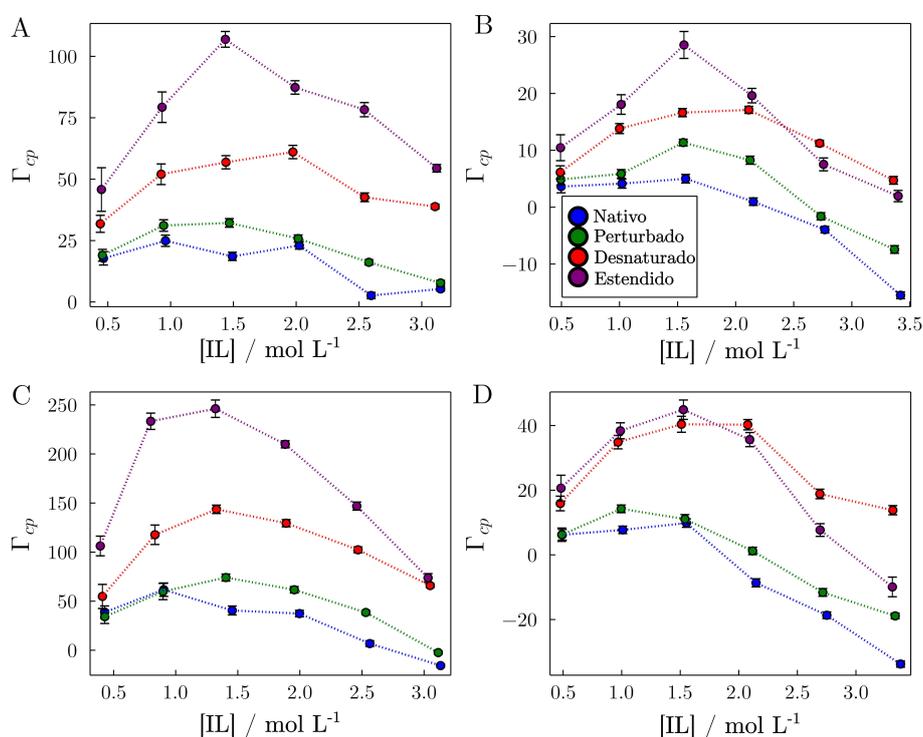


Figura 9.8: Coeficientes de solvatação preferencial ( $\Gamma_{cp}$ ) para os líquidos iônicos: (A)  $[\text{EMIM}][\text{DCA}]$ , (B)  $[\text{EMIM}][\text{BF}_4]$ , (C)  $[\text{BMIM}][\text{DCA}]$  e (D)  $[\text{BMIM}][\text{BF}_4]$ . Os conjuntos conformacionais da ubiquitina analisados incluem as conformações nativas, perturbada, desnaturada e estendida. As concentrações no eixo  $x$  correspondem às concentrações dos LIs no *bulk*, recalculadas após as simulações.

Na Figura 9.8, observa-se que  $\Gamma_{cp}$  aumenta com o desenovelamento das estruturas na maioria dos sistemas simulados. A Equação 9.4.1 relaciona a constante de equilíbrio de um hipotético equilíbrio de enovelamento em dois estados com a diferença nos coeficientes de solvatação preferencial do cossolvente para a proteína em dois estados de enovelamento distintos:

$$\left(\frac{\partial \ln K}{\partial \ln a_3}\right)_{T,P,m_p} = \Delta\Gamma_{cp} = \Gamma_{\text{desnaturado}} - \Gamma_{\text{nativo}} \quad (9.4.1)$$

onde  $K$  é a constante de equilíbrio da reação de enovelamento  $\text{Nativo} \rightleftharpoons \text{Desnaturado}$ , e  $a_3$  representa a atividade do cossolvente [144, 162, 163]. Se a diferença entre  $\Gamma_{cp}$  para os conjuntos de estruturas dos estados nativo e desnaturado for positiva ( $\Gamma_{cp}(\text{desnaturado}) - \Gamma_{cp}(\text{nativo}) > 0$ ), a constante de equilíbrio aumenta com a atividade do cossolvente, favorecendo o estado desnaturado. Assim,  $\Delta\Gamma_{cp} > 0$  indica que o aumento da concentração de cossolvente favorece estruturas desnaturadas em relação às nativas. Por outro lado,  $\Delta\Gamma_{cp} < 0$  implica que o estado nativo é favorecido com o aumento da concentração de cossolvente. Dessa forma, cossolventes com  $\Delta\Gamma_{cp} < 0$  podem ser classificados como protetores, enquanto aqueles com  $\Delta\Gamma_{cp} > 0$  são classificados como desnaturantes.

A Tabela 9.3 apresenta a diferença nos valores de  $\Gamma_{cp}$  entre cada estado simulado e o estado nativo como uma função da concentração de referência de LI, considerando concentrações de 0,5 e 3,0 mol L<sup>-1</sup>. Dados completos e mais abrangentes podem ser obtidos na referência referente a este capítulo [166]. Em geral,  $\Delta\Gamma_{cp} > 0$  para todos os sistemas e concentrações. Assim, os LIs deslocam o equilíbrio de enovelamento para estados desnaturados, mesmo em sistemas onde a proteína nativa é preferencialmente hidratada, sugerindo um papel desnaturante do cossolvente.

Tabela 9.3: Valores de  $\Delta\Gamma_{cp}$  para diferentes líquidos iônicos em concentrações de 0,5 e 3,0 mol L<sup>-1</sup>.

Líquido Iônico	Concentração			
	de referência ( $C_{IL}$ ) / mol L <sup>-1</sup>	Perturbado - Nativo	Desnaturado - Nativo	Estendido - Nativo
[EMIM][DCA]	0,5	6	17	28
	3,0	5	32	44
[EMIM][BF <sub>4</sub> ]	0,5	2	3	7
	3,0	8	20	16
[BMIM][DCA]	0,5	-1	13	63
	3,0	12	80	85
[BMIM][BF <sub>4</sub> ]	0,5	0	10	15
	3,0	13	45	25

Por exemplo, em um sistema com 3,0 mol L<sup>-1</sup> de [BMIM][BF<sub>4</sub>], o  $\Gamma_{cp}$  para a estrutura nativa da ubiquitina é negativo, indicando exclusão preferencial do LI dos domínios proteicos. No entanto, para o mesmo sistema com conformações desnaturadas,  $\Gamma_{cp}$  é maior do que no estado nativo, resultando em  $\Delta\Gamma_{cp} > 0$ , o que implica que o solvente favorece as estruturas desnaturadas. Em geral, quanto mais avançado o estado de desnaturação da proteína, maior o valor de  $\Gamma_{cp}$  em relação à estrutura nativa para uma dada concentração. Isso sugere uma relação cooperativa entre a desnaturação da proteína e suas interações com o LI. À medida que a proteína se desnatura, a natureza química da superfície proteica torna-se mais compatível com o LI, favorecendo a exposição adicional da superfície proteica.

## Capítulo 10

# Solvatação em DES

O estudo da solvatação de proteínas em soluções aquosas de líquidos iônicos ao longo desta tese exigiu o desenvolvimento e aplicação de metodologias específicas para caracterizar efeitos estruturais e termodinâmicos da solvatação. Durante o período de um ano na França, sob a orientação do Prof. Agílio Pádua, expandimos essa abordagem para investigar a solvatação do ibuprofeno em solventes eutéticos profundos (DESs) à base de betaína. A motivação central foi compreender como flutuações locais na densidade do solvente influenciam a organização estrutural e as interações soluto-solvente nesses sistemas altamente viscosos. Assim, a continuidade do trabalho realizado em proteínas permitiu adaptar as ferramentas analíticas para um novo contexto, mantendo a consistência metodológica na caracterização da solvatação.

Para descrever com maior precisão a estrutura e a dinâmica desses sistemas complexos, utilizamos simulações de dinâmica molecular empregando campos de força polarizáveis, desenvolvidos pelo próprio grupo do Prof. Agílio Pádua. Embora esses campos de força sejam computacionalmente mais exigentes, eles são essenciais para representar adequadamente os efeitos eletrônicos presentes em DESs, especialmente devido às fortes interações intermoleculares e à heterogeneidade do meio. Além disso, investigamos a solvatação do ibuprofeno em diluição infinita, considerando solventes formados pela betaína combinada com etilenoglicol, propilenoglicol e 1,2-butanodiol, o que permitiu uma análise comparativa detalhada das influências estruturais e termodinâmicas de diferentes composições do solvente.

Neste capítulo, apresentamos uma análise aprofundada da solvatação do ibuprofeno em DESs, combinando campos de força polarizáveis com ferramentas teóricas

robustas, como as MDDFs e a teoria de soluções de Kirkwood-Buff. Essa abordagem permitiu investigar a afinidade relativa dos componentes do solvente eutético pelo soluto e caracterizar a estrutura de solvatação em diferentes escalas. Além de discutir os resultados obtidos, exploramos as vantagens e desafios metodológicos da aplicação dessas técnicas a sistemas altamente viscosos, fornecendo percepções fundamentais para o desenho racional de DESs para aplicações biomoleculares.

## 10.1 Visão geral sobre DESs

Os solventes eutéticos profundos (DESs, do inglês *Deep Eutectic Solvents*) emergiram como uma nova classe de solventes formados por misturas eutéticas compostas de ácidos e bases de Lewis ou Brønsted, contendo uma variedade de espécies aniônicas e catiônicas [186, 187]. O interesse por esses solventes cresceu significativamente após o estudo pioneiro de Abbott e colaboradores [188], no qual foi observada uma depressão anormalmente grande no ponto de fusão de misturas contendo doadores (HBD, do inglês *hydrogen bond donor*) e aceptores (HBA, do inglês *hydrogen bond acceptor*) de ligação de hidrogênio [187–189].

O primeiro sistema a exibir o comportamento descrito acima foi a mistura 1:2 de cloreto de colina ((2-hidroxi-etil)-trimetilamônio cloreto,  $T_m \approx 302$  °C) com ureia cristalina ( $T_m \approx 133$  °C) [188, 189]. A mistura resultante, denominada "reline", apresentou-se líquida em temperatura ambiente, com um ponto eutético de  $T_{\text{eutético}} = 12$  °C. Abbott e colaboradores cunharam o termo "mistura eutética profunda" para descrever este fenômeno [187, 189], e o termo solvente eutético profundo passou a ser amplamente utilizado para designar líquidos formados próximos à composição eutética da mistura, ou seja, na proporção molar resultante no menor ponto de fusão [187–189].

Os DESs mais frequentemente preparados e estudados podem ser classificados em cinco tipos principais [186]: Tipo I, composto por um sal de amônio quaternário e um cloreto metálico; Tipo II, que contém um sal de amônio quaternário e um cloreto metálico hidratado; Tipo III, formado por um sal de amônio quaternário e um doador de ligação de hidrogênio (HBD, do inglês *hydrogen bond donor*), geralmente uma molécula orgânica como uma amida, um ácido carboxílico ou um polioliol; Tipo IV, que combina um cloreto metálico hidratado e um HBD [190, 191]; e Tipo V, uma classe mais recente composta exclusivamente por espécies

moleculares não iônicas que atuam como HBAs e HBDs [191, 192].

A característica distintiva dos DESs é a significativa depressão do ponto de fusão observada em determinadas composições, especialmente na eutética, quando comparada ao comportamento previsto para uma mistura ideal. Quanto maior for a depressão no ponto de fusão, maior será a faixa de composições e temperaturas que permitem a formação de DESs a partir de sólidos distintos [186, 187, 193]. A Figura 10.1 ilustra um diagrama de fases temperatura-fração molar para uma mistura genérica de dois sólidos, **A** e **B**. A linha tracejada representa a variação do ponto de fusão com a fração molar de **B**, considerando o comportamento ideal da mistura. No entanto, DESs contendo cloreto de colina, por exemplo, apresentam uma diminuição significativa no ponto de fusão em relação à curva ideal, devido a desvios negativos de idealidade. Essa diferença entre o ponto de fusão real e o ideal é indicada por ( $\Delta T_f$ ).

O parâmetro ( $\Delta T_f$ ) é particularmente relevante, pois reflete a magnitude da depressão do ponto de fusão no ponto eutético. Tal redução está frequentemente associada à intensidade das interações intermoleculares entre **A** e **B**. Interações mais favoráveis resultam em uma maior depressão da temperatura de fusão [186, 187]. Por exemplo, em DESs contendo cloreto de colina e ureia, como estudado por Abbott e colaboradores [188, 189], a redução do ponto de fusão é atribuída às fortes ligações de hidrogênio entre a colina e os grupos  $\text{NH}_2$  da ureia, que perturbam a estrutura cristalina dos sólidos puros, conforme discutido nos trabalhos de revisão [187, 189, 194].

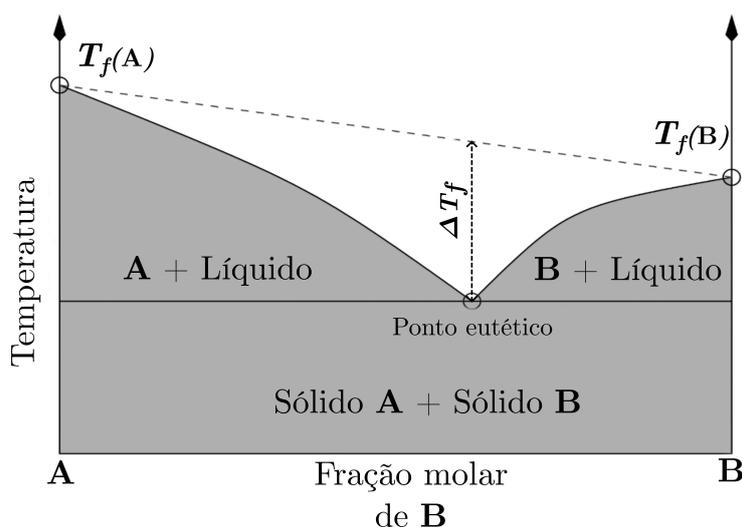


Figura 10.1: Diagrama de fases da mistura de dois sólidos A e B. Representação esquemática da diferença no ponto de fusão ( $\Delta T_f$ ) na composição eutética de uma mistura binária. Interações mais fortes entre os componentes A e B resultam em valores maiores de  $\Delta T_f$ .  $T_f(A)$  e  $T_f(B)$  representam, respectivamente, as temperaturas de fusão para **A** e **B**.

Apesar da explicação baseada exclusivamente nas ligações de hidrogênio para a diminuição do ponto de fusão em DESs ser amplamente aceita, há estudos que sugerem que o fenômeno da depressão do ponto de fusão pode ser atribuído a uma diversidade de interações intermoleculares, que aumentam a desordem local nos DESs [195, 196]. Esta desordem promove uma maximização da entropia [195], contribuindo para a depressão do ponto de fusão ao reduzir a energia livre no ponto eutético, considerando a transição dos sólidos separados para a mistura líquida. Tal perspectiva desafia a visão tradicional de que a deslocalização de carga no ânion ou a formação de clusters coordenados seja o principal mecanismo responsável pelo comportamento eutético profundo [196]. Portanto, a depressão do ponto de fusão é o resultado de um equilíbrio dinâmico entre contribuições entálpicas e entrópicas.

## 10.2 Campos de força polarizáveis e simulações de DESs

Os campos de força são fundamentais para que as simulações de DM possam representar adequadamente sistemas químicos. Mediante isto, um aspecto crucial a ser considerado é que, quanto mais preciso é um campo de força, mais complexo e computacionalmente exigente a simulação se torna. Assim, o desenvolvimento de campos de força para simulações de DM frequentemente envolve um compromisso entre precisão e eficiência computacional [197]. Para que essas simulações sejam aplicáveis em escalas de tempo relevantes para fenômenos químicos, várias aproximações são introduzidas nos campos

de força.

Uma das aproximações mais importantes diz respeito à descrição das propriedades eletrostáticas dos átomos e moléculas [92, 197]. Na maioria dos campos de força aplicados a sistemas biomoleculares, as propriedades eletrostáticas são tratadas posicionando uma carga parcial pontual no núcleo atômico, onde a magnitude da carga e sua posição permanecem fixas ao longo da simulação. No entanto, esta abordagem não captura variações nas propriedades eletrostáticas que dependem do ambiente molecular. Átomos e moléculas podem ser mais ou menos polarizáveis e, em muitos casos, a polarizabilidade é fundamental para uma descrição correta das propriedades de um sistema de interesse.

Por exemplo, o momento de dipolo de uma molécula de água varia significativamente conforme o ambiente: 1,9 D na fase gasosa [198], 2,1 D em pequenos aglomerados [199] e 2,9 D no líquido em massa [200]. Esses efeitos são particularmente relevantes em sistemas biomoleculares. Sabe-se, por exemplo, que a distribuição eletrônica de um peptídeo varia conforme sua conformação [201] e que a distribuição eletrônica de um ligante se altera ao se ligar a uma proteína [202]. Nenhum desses efeitos seria capturado por um campo de força não polarizável.

A inclusão da polarizabilidade em campos de força representa um avanço significativo para descrever interações mais realistas. No entanto, esta inclusão aumenta consideravelmente o custo computacional das simulações, tornando-se um tópico de desenvolvimento ativo na modelagem de sistemas complexos [197, 203, 204].

Os DESs formam fases complexas e altamente estruturadas, dominadas por ligações de hidrogênio e interações coulômbicas, caracterizando um grau de organização de ordem média no líquido [186, 205]. Em particular, as propriedades apresentadas por solventes como DESs são notavelmente interessantes, embora desafiadoras de serem simuladas computacionalmente. Estas dificuldades decorrem das limitações dos métodos, que, dependendo do nível de teoria empregado, podem ter dificuldade em representar adequadamente as características estruturais e dinâmicas desses sistemas [206].

Tipicamente, DESs exibem características de estruturação de longo alcance e dinâmicas lentas, dado que muitos desses sistemas possuem viscosidades dezenas de vezes superiores à da água. Devido a essas características, a representação correta desses sistemas requer que as simulações sejam realizadas em escalas de tempo muito maiores do que

aquelas utilizadas para solventes convencionais, como água ou etanol.

Dada a diversidade estrutural dos DESs, os campos de força moleculares devem possuir ampla aplicabilidade, tanto em relação aos grupos funcionais que podem representar quanto às propriedades que podem ser calculadas. Campos de força tradicionais para compostos orgânicos, como OPLS-AA [98], CHARMM/CGenFF [100] e AMBER/GAFF [102], oferecem vantagens significativas, como conjuntos extensivos de parâmetros cobrindo diversos grupos funcionais e um elevado grau de transferibilidade, permitindo sua aplicação a novas moléculas ou íons.

Apesar dos esforços para aplicar campos de força convencionais a DESs, modelos de carga fixa demonstraram-se, em geral, inadequados para representar corretamente tanto propriedades estruturais quanto dinâmicas. Uma das principais limitações desses modelos é que não foram projetados para descrever fases iônicas [197]. Quando cargas iônicas inteiras ( $\pm 1$ ) são utilizadas, as dinâmicas são frequentemente superestimadas, com discrepâncias nos coeficientes de difusão e viscosidade chegando a uma ordem de grandeza em relação aos valores experimentais. Além disso, propriedades como o calor de vaporização tendem a ser superestimadas, como apresentado em [206].

Uma solução prática proposta para superar essas limitações é a redução das cargas iônicas por um fator de aproximadamente 0,8 [207]. Esta abordagem reduz a energia coesiva iônica, melhorando a representação das dinâmicas do sistema. Justificativas físicas para essa redução incluem considerações sobre transferência de carga e efeitos de polarização efetiva [208]. Embora o escalonamento das cargas melhore as propriedades dinâmicas, ele pode subestimar as densidades dos líquidos (em alguns casos, até 5%) e alterar o equilíbrio entre interações soluto-soluto, soluto-solvente e solvente-solvente [209], dificultando a predição precisa da solvatação.

Outro desafio crítico é que o estudo de diferentes misturas exige o uso de espécies moleculares e/ou iônicas em ambientes distintos daqueles nos quais foram parametrizadas. Assim, a representação correta das cargas torna-se um aspecto crucial nas simulações desses sistemas. Logo, a adaptação de campos de força para DESs exige um equilíbrio cuidadoso entre precisão na descrição estrutural e dinâmica e robustez para aplicações em condições variadas. O desenvolvimento de modelos que incorporem polarização explícita ou ajustem as cargas adaptativamente é uma direção promissora para superar essas limitações.

Embora métodos quânticos, como cálculos de estrutura eletrônica baseados na teoria do funcional da densidade, forneçam descrições detalhadas das interações eletrônicas, seu custo computacional elevado inviabiliza seu uso em escalas relevantes para LIs e DESs. Como alternativa, o uso de campos de força polarizáveis em DM tem se mostrado uma solução viável [197]. Estes modelos incorporam explicitamente efeitos de polarização, representando graus de liberdade eletrônicos (*electronic degrees of freedom*, DOF) de maneira eficiente, mantendo o custo computacional comparável ao de campos de força tradicionais de carga fixa. Avanços recentes nesses modelos ampliaram sua aplicabilidade a sistemas complexos, como líquidos iônicos práticos e DESs, permitindo capturar interações intensas de ligação de hidrogênio e facilitando o projeto de novos materiais com propriedades ajustadas a aplicações específicas [206, 210].

As simulações dos DESs estudados neste capítulo foram realizadas utilizando o campo de força CL&Pol para permitir uma reprodução mais precisa das propriedades físico-químicas dos DESs e, assim, entender como solvatam moléculas como o ibuprofeno mais corretamente [206, 211, 212].

### 10.3 Aspectos gerais do campo de força CL&Pol

O campo de força CL&Pol foi desenvolvido para superar as limitações dos campos de força tradicionais de carga fixa, especialmente em sistemas altamente polares e carregados, como LIs e DESs. Estes sistemas, como discutido anteriormente, apresentam interações eletrostáticas complexas e efeitos de polarização significativos, que não são adequadamente descritos por modelos de carga fixa. O CL&Pol incorpora explicitamente a polarização eletrônica, permitindo uma representação mais realista das propriedades estruturais e dinâmicas desses sistemas [206, 212].

Um dos princípios centrais do CL&Pol é a transferibilidade. Em vez de parametrizar cada molécula ou íon individualmente, o campo de força utiliza uma abordagem baseada em fragmentos: moléculas ou íons complexos são decompostos em unidades funcionais menores, como grupos químicos específicos (por exemplo, grupos metila, hidroxila ou anéis aromáticos). Cada fragmento é parametrizado de forma independente, garantindo que os parâmetros sejam consistentes e aplicáveis em diferentes contextos moleculares. Essa modularidade é particularmente útil para simular a diversidade química de LIs e DESs,

permitindo a extensão do campo de força a novos compostos sem a necessidade de reparametrização extensiva.

O campo de força CL&Pol separa explicitamente as contribuições de polarização das interações de dispersão, modeladas por potenciais de Lennard-Jones. Esta separação é crucial para evitar sobreposição de efeitos e garantir que as forças de dispersão e os efeitos de indução sejam tratados de forma independente. A polarização é modelada por meio de dipolos de Drude, que representam a resposta dinâmica das nuvens eletrônicas ao ambiente eletrostático local, aumentando a confiabilidade dos resultados[206, 212].

No modelo de dipolos de Drude, cada átomo polarizável é representado por duas cargas pontuais: um núcleo Drude, localizado no centro do sítio atômico, e uma partícula Drude, que representa a nuvem eletrônica polarizável. A partícula Drude é conectada ao núcleo por uma mola harmônica, permitindo que o dipolo induzido responda ao campo eletrostático local. Durante uma simulação de dinâmica molecular, os dipolos de Drude ajustam-se dinamicamente ao ambiente eletrostático local, utilizando métodos computacionalmente eficientes para gerenciar este ajuste.

O campo de força CL&Pol apresenta um nível de complexidade significativamente maior em comparação com aqueles tradicionalmente empregados em simulações de dinâmica molecular. Esta complexidade decorre da necessidade de descrever com maior precisão as interações intermoleculares em sistemas envolvendo solventes eutéticos profundos. Muitos dos aspectos técnicos essenciais para a implementação e aplicação de campos de força polarizáveis são discutidos em profundidade nos trabalhos do Prof. Agílio Pádua e colaboradores. Para uma compreensão mais detalhada desses desenvolvimentos, o leitor pode se interessar na leitura de referências como [206, 211, 212].

## 10.4 Metodologia das simulações de DM para os DESS

As simulações foram realizadas utilizando o campo de força CL&Pol, baseado nos parâmetros OPLS-AA para álcoois e outras moléculas [98, 206, 211, 213]. Grande parte dos parâmetros utilizados para os sistemas simulados foi obtida a partir do repositório GitHub `clandpol`, sendo necessário apenas calcular as cargas para o nitrogênio presente na betaína. A construção do zwitterion de betaína foi realizada no software Avogadro [214], e as cargas da molécula foram obtidas por meio de uma otimização geométrica no Gaussian [215] com o

funcional B3LYP e a base 6-31G(d) [216]. Como os demais sistemas já possuíam parâmetros definidos no campo de força ( $\epsilon$ ,  $\sigma$ , cargas, ligações, etc.), os arquivos foram gerados para ibuprofeno, betaína (BET), etileno glicol (EG), propileno glicol (PG) e 1,2-butanodiol (BTO) utilizando o `fftool` [217].

Cada átomo polarizável, principalmente os pesados como carbono, oxigênio e nitrogênio, foi associado a uma partícula Drude de 0,4 Da, conectada por uma mola harmônica com constante de força  $k_D = 4184 \text{ kJ mol}^{-1}$ . A magnitude das cargas opostas das partículas Drude foi determinada com base na polarizabilidade do átomo correspondente, calculada pela fórmula  $\alpha = q_D^2/k_D$  [206, 218]. Hidrogênios foram tratados como não polarizáveis devido à sua massa e densidade eletrônica reduzidas, com suas polarizabilidades atribuídas aos átomos pesados ligados a eles [211, 213, 219].

A implementação do modelo Drude requer um esquema de termostatização duplo, no qual o movimento relativo entre o núcleo Drude (DC) e a partícula Drude (DP) é termostatizada em temperaturas muito baixas ( $T = 1 \text{ K}$ ), enquanto os demais graus de liberdade são mantidos na temperatura desejada [220]. Pacotes de MD amplamente utilizados, como OpenMM e LAMMPS, implementaram recentemente um termostato Nosé–Hoover de tripla temperatura para corrigir desequilíbrios causados pelo fluxo de calor para os graus de liberdade dos Drudes [219, 221].

Interações de curto alcance entre partículas Drude foram atenuadas usando funções de Thole, com o parâmetro universal  $a = 2.6$ . Funções de amortecimento Tang-Toennies foram aplicadas para mitigar interações eletrostáticas entre hidrogênios de grupos hidroxila e dipolos Drude em curtas distâncias [220, 222, 223]. Essas funções evitam instabilidades nas trajetórias dos dipolos Drude (amortecendo o potencial eletrostático em curtas distâncias), conhecidas como "catástrofes de polarização", que ocorrem quando a aproximação de cargas pontuais se torna inadequada [224].

No CL&Pol, os efeitos de polarização são contabilizados explicitamente pelos dipolos Drude, e os parâmetros de Lennard-Jones (LJ) são ajustados para representar apenas dispersão. Esse ajuste é realizado por um fator de escala ( $k_{ij}$ ), definido como a fração da energia de dispersão ( $E_{\text{disp}}$ ) na soma de energias de dispersão e indução ( $E_{\text{ind}}$ ). O fator de escala é derivado por cálculos SAPT e pode ser expresso em termos de polarizabilidades atômicas ( $\alpha$ ), cargas líquidas ( $Q$ ) e momentos de dipolo ( $\mu$ ) das interações entre fragmentos

[212, 225].

### 10.4.1 Sistemas estudados

Foram estudados sistemas contendo DESs na proporção 1:3 (HBA:HBD) com ibuprofeno em diluição infinita utilizando simulações DM, utilizando o software OpenMM [226]. Três sistemas foram considerados: BET:EG, BET:PG e BET:BTO. As configurações iniciais foram criadas com `fftool` e `Packmol` [153, 154, 217]. Utilizou-se um corte de 12 Å para interações não ligadas, e o método PPPM foi empregado para calcular energias eletrostáticas com precisão de  $10^{-5}$  [123]. Ligações terminando em hidrogênios foram restringidas com o algoritmo SHAKE [227, 228].

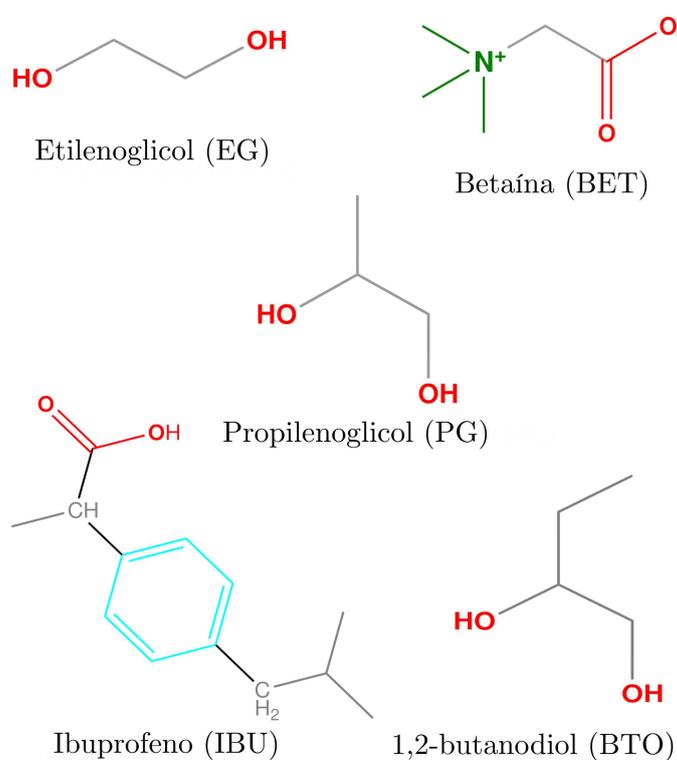


Figura 10.2: Estruturas dos compostos que formam os solventes eutéticos utilizados para a simulação.

Para simular os sistemas, caixas de aproximadamente 85 Å foram preenchidas com um número suficiente de moléculas para atender às proporções desejadas, estimadas com base na densidade experimental dos DESs sem água, utilizando o pacote `MolsimToolkit.jl`.

Tabela 10.1: Tamanhos da caixa, densidade experimental, densidade calculada após a produção em NPT, e número de componentes usados para criar as caixas de simulação. As densidades e proporções foram referenciadas do trabalho experimental disponível em [229].

(1:3)	Tamanho da caixa (Å)	Densidade experimental (g/cm <sup>3</sup> )	Densidade calculada (g/cm <sup>3</sup> )	Número de moléculas HBA	Número de moléculas HBD
BET:EG	81.12	1.13	1.129 ± 0.001	1200	3600
BET:PG	79.44	1.07	1.077 ± 0.002	945	2835
BET:BTO	80.70	–	1.033 ± 0.001	823	2469

Devido à alta viscosidade dos DESSs, as simulações foram realizadas a 353 K utilizando o termostato Drude-Langevin e o barostato Monte Carlo [222]. Os graus de liberdade Drude foram mantidos a 1 K, conforme procedimento desenvolvido para dipolos Drude frios. Cada simulação seguiu um protocolo regular: minimização de energia, fase de equilíbrio e execução de produção. Após 10.000 passos de minimização para eliminar configurações desfavoráveis, os sistemas passaram por uma fase de equilíbrio NPT (número de partículas, pressão e temperatura constantes) de 5 ns a 1 bar e 353 K, garantindo densidade e energia estáveis. As execuções das etapas de produção da DM tiveram duração de 30 ns nas mesmas condições. Ao total, 10 réplicas para cada sistema foram simuladas.

A análise da solvatação do ibuprofeno foi realizada utilizando os mesmos métodos empregados para a ubiquitina em soluções aquosas de LIs. Neste capítulo, o ibuprofeno é tratado como o soluto, simulado em diluição infinita, enquanto as espécies HBA (apenas a betaína, BET) e HBD (os álcoois simulados, como etileno glicol, propileno glicol e 1,2-butanodiol) são os constituintes do solvente. Portanto, as MDDFs são sempre calculadas em relação ao ibuprofeno, considerando as interações com os componentes HBA e HBD do solvente.

O pacote `ComplexMixtures.jl` foi utilizado para calcular as MDDFs e as integrais de KB [139, 161]. A convergência mais adequada possível das integrais de Kirkwood-Buff foi observada com  $R = 20$  Å na maioria dos sistemas. O volume da solução mais próximo do soluto, dentro dessa distância, foi considerado como o “domínio do soluto”, enquanto a região externa a esse domínio foi utilizada para inferir a estrutura e as propriedades termodinâmicas da solução sem a presença da proteína. As concentrações efetivas do bulk foram obtidas a

partir das simulações, calculando-se a densidade de cada solvente na região entre 20 e 30 Å da superfície do ibuprofeno. Esse procedimento permitiu a análise de um subsistema aberto dentro da simulação, minimizando efeitos de tamanho finito.

No entanto, a aplicação dessa metodologia aos DESs apresentou desafios adicionais devido à alta viscosidade desses sistemas. A difusão limitada das moléculas impacta diretamente a convergência das integrais de Kirkwood-Buff, tornando essencial a escolha criteriosa do raio de integração e a análise detalhada da estabilidade das distribuições de solvente ao redor do soluto.

## 10.5 Solvatação do ibuprofeno em DESs Baseados em Betaína

A Figura 10.3 apresenta as MDDFs, em relação ao ibuprofeno, para os componentes do DES formado por BET e EG. Tomando como base toda a discussão realizada nos capítulos anteriores sobre as MDDFs e considerando a natureza química de BET e EG, algumas análises gerais podem ser tiradas das MDDFs da Figura 10.3: os dois picos observados para a BET indicam diferentes tipos de interação. O primeiro pico, em 1,8 Å, está associado às ligações de hidrogênio estabelecidas pelo grupo  $\text{COO}^-$  da betaína, que atua como aceptor de ligação de hidrogênio. Já o segundo pico, em 2,4 Å, reflete interações não específicas ou mediadas por outras moléculas do solvente, caracterizadas por sua forma mais ampla, que sugere variedade nas interações.

A MDDF do EG, representada pela curva azul-claro na Figura 10.3, exibe apenas um pico centrado em aproximadamente 2,5 Å. Tal distância sugere que as moléculas de EG não estabelecem ligações de hidrogênio diretas com o ibuprofeno.

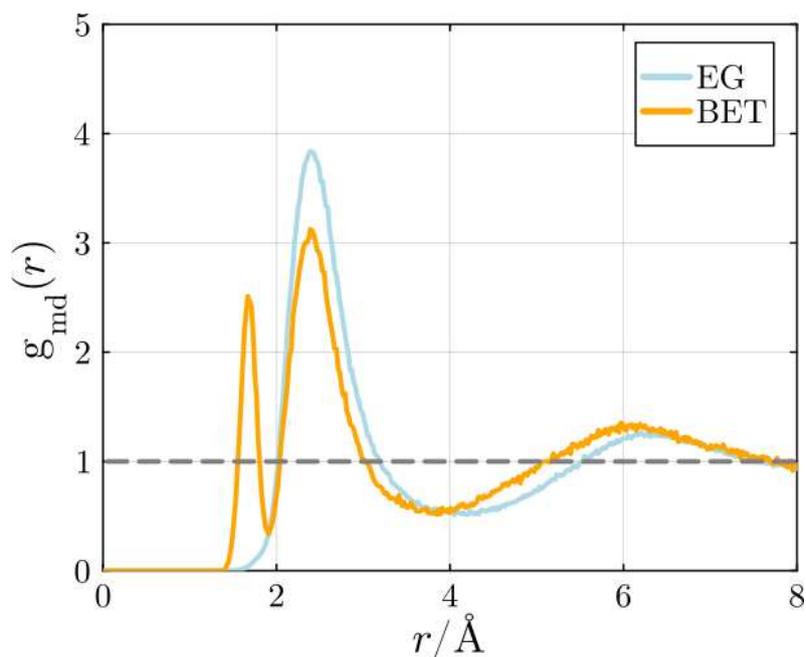


Figura 10.3: MDDFs dos componentes do DES BET:EG na proporção, respectivamente, de 1:3 em quantidade de matéria. As MDDFs são mostradas até 8 Å para melhor visualização dos picos importantes. A curva laranja (BET) exibe dois picos principais (1,8 Å e 2,4 Å), enquanto a curva azul (EG) apresenta um pico em 2,5 Å.

A Figura 10.4 apresenta as MDDFs dos componentes que formam os DESs nos sistemas simulados. Observa-se, na Figura 10.4D, que o pico central da MDDF do HBD é mais pronunciado no sistema com BTO, o HBD mais hidrofóbico dentre os analisados. Este comportamento é acompanhado por uma diminuição do pico em 2,4 Å na MDDF de betaína à medida que o HBD se torna maior e mais hidrofóbico. Assim, as MDDFs indicam que, conforme a espécie HBD se torna mais hidrofóbica, a densidade relativa de betaína na região correspondente à distribuição do HBD diminui, enquanto a densidade relativa na faixa de ligação de hidrogênio aumenta.

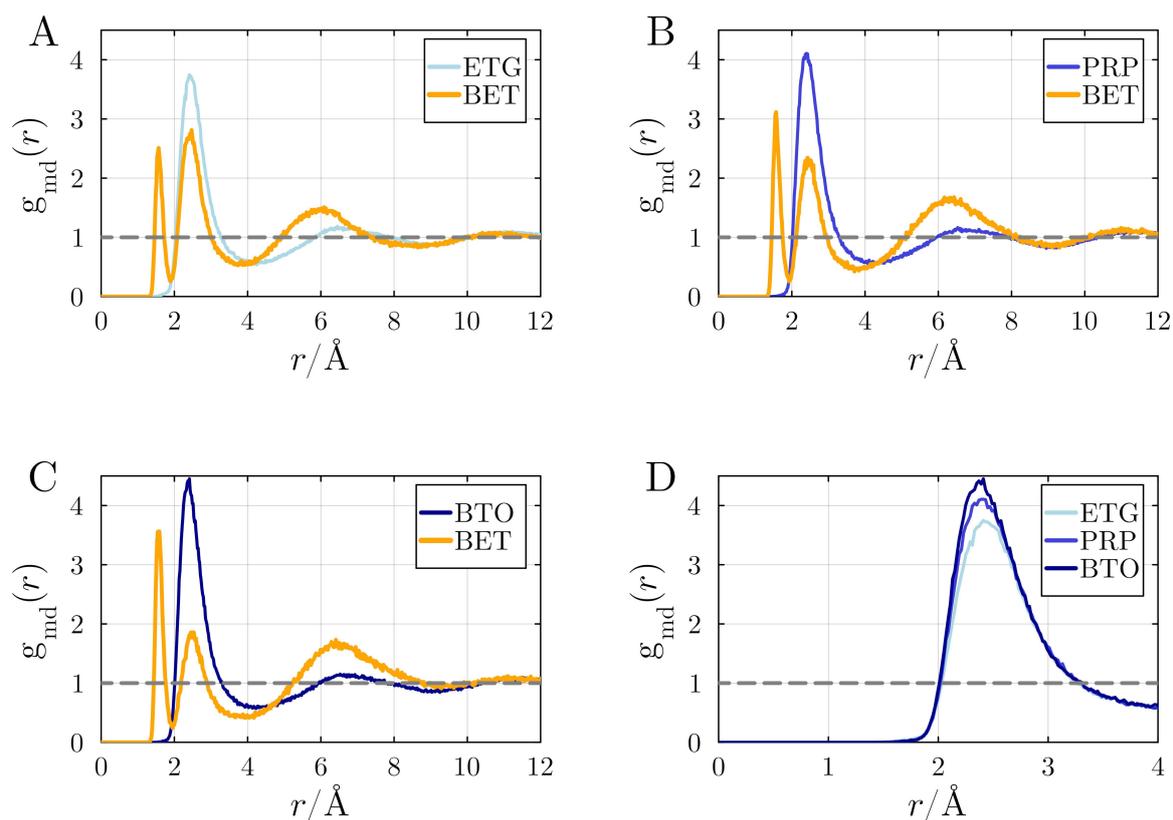


Figura 10.4: MDDFs dos componentes dos DESs: A) BET:EG, B) BET:PRP e C) BET:BTO, na proporção de 1:3 em quantidade de matéria, respectivamente. D) Comparação das MDDFs das espécies doadoras de ligações de hidrogênio.

Assim como apresentado em capítulos passados, é possível investigar os detalhes moleculares ao decompor as MDDFs nas contribuições de cada átomo do DES, conforme ilustrado na Figura 10.5. A Figura 10.5B apresenta a decomposição da MDDF de betaína em termos dos átomos que compõem os grupos  $\text{COO}^-$ ,  $\text{CH}_2$  e  $\text{N}(\text{CH}_3)_3^+$ . O primeiro pico é formado inteiramente por interações estabelecidas através do grupo  $\text{COO}^-$  da betaína, confirmando seu papel como um sítio aceitor de ligações de hidrogênio. O segundo pico, conforme sugerido pela análise geral das MDDFs, está mais provavelmente associado a átomos do grupo  $\text{N}(\text{CH}_3)_3^+$  que estão mais próximos da superfície do ibuprofeno.

A Figura 10.5A apresenta a decomposição das MDDFs para o EG em termos dos átomos do solvente. A análise indica que os átomos mais próximos do ibuprofeno são aqueles pertencentes aos dois grupos  $\text{CH}_2$ . No caso do EG, o grupo  $\text{OH}$  apresenta uma contribuição menor para o pico em comparação aos grupos hidrofóbicos  $\text{C}_2\text{H}_4$ .

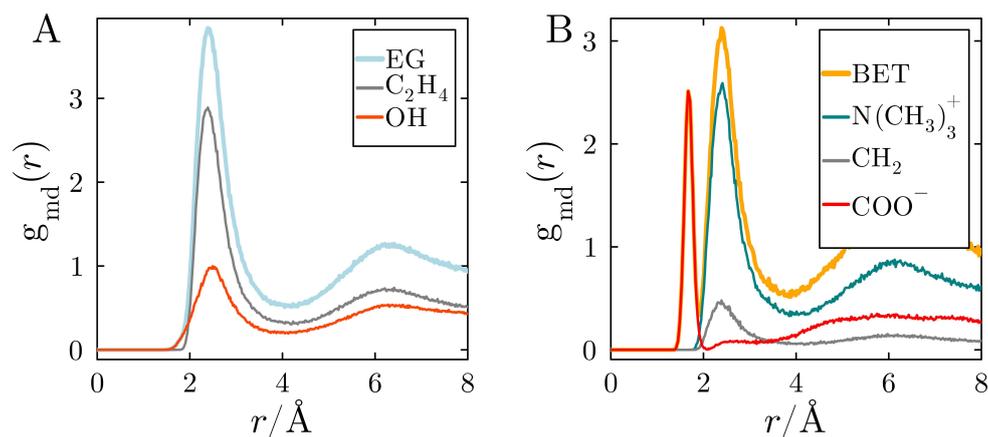


Figura 10.5: Decomposição dos MDDFs para os componentes do DES BET:EG. (A) Contribuição dos átomos do EG: grupos  $\text{CH}_2$  estão mais próximos do ibuprofeno. (B) Contribuição dos átomos da betaína: o pico em  $1,8 \text{ Å}$  é formado pelo grupo  $\text{COO}^-$ , enquanto o pico em  $2,4 \text{ Å}$  envolve predominantemente o grupo  $\text{N}(\text{CH}_3)_3^+$ .

Com a decomposição em termos dos átomos do solvente, é possível identificar quais átomos apresentam maior probabilidade de interagir com o soluto. No caso ilustrativo de BET e EG, observa-se que, em distâncias muito pequenas, a BET estabelece ligações de hidrogênio com o ibuprofeno. Por outro lado, as MDDFs indicam que o EG, devido à ocupação da vizinhança imediata do ibuprofeno por BET, apresenta baixa probabilidade de formar ligações de hidrogênio diretamente com o ibuprofeno.

Para obter uma visão mais detalhada sobre os aspectos moleculares da solvatação do ibuprofeno, é possível analisar a decomposição das MDDFs em termos das contribuições dos grupos funcionais do soluto. A Figura 10.6 apresenta a decomposição das MDDFs com base nas contribuições dos grupos do ibuprofeno. O soluto foi dividido em quatro partes:  $\text{CH}_2\text{CH}(\text{CH}_3)_2$ ,  $\text{C}_6\text{H}_4$ ,  $\text{CHCH}_3$  e  $\text{COOH}$ .

Como destacado na Figura 10.5B, a decomposição do primeiro pico da MDDF de betaína indica que todas as contribuições se originam de átomos de betaína distribuídos ao redor do grupo  $\text{COOH}$  do ibuprofeno. Em contraste, o segundo pico apresenta uma composição mais complexa: há uma contribuição relevante do grupo  $\text{COOH}$ , porém o grupo isopropila, representado em cinza, apresenta a maior contribuição entre todos os grupos do ibuprofeno. Os grupos benzeno e alquila (representados em verde) contribuem igualmente para a MDDF total. Em média, as contribuições dos grupos hidrofóbicos são significativamente relevantes.

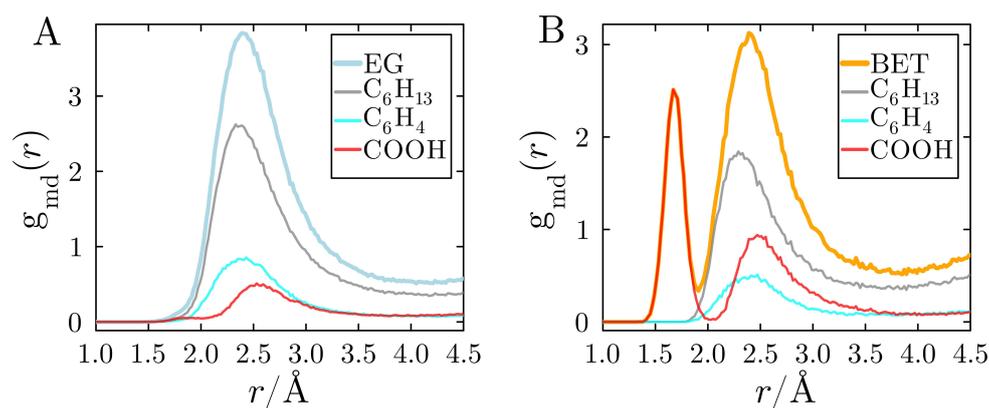


Figura 10.6: MDDFs decompostas em termos das partes do ibuprofeno. (A) Decomposição da MDDF do EG em termos das partes do ibuprofeno. Os sítios prováveis de interação com o EG são aqueles com características mais hidrofóbicas. (B) O primeiro pico da MDDF de BET apresenta contribuições dos átomos de betaína ao redor do grupo COOH do ibuprofeno. O segundo pico tem contribuições do grupo COOH e contribuições significativas do grupo isopropila (cinza), enquanto os grupos benzeno e alquila (verde) contribuem igualmente para a MDDF total.

Considerando a alta viscosidade e as diferentes estruturas dos DESs, a análise da solvatação constitui uma tarefa desafiadora. Inicialmente, ao analisar a MDDF total e sua decomposição em termos de soluto e solvente, os principais pontos a serem destacados são os seguintes: a BET interage por meio de ligações de hidrogênio com o ibuprofeno, e essa interação ocorre entre os grupos  $COO^-$  da BET e COOH do ibuprofeno. O segundo pico pode estar relacionado aos grupos  $COO^-$  da BET interagindo com as porções carregadas positivamente ( $N(CH_3)_3^+$ ) das moléculas de BET mais próximas do ibuprofeno. O EG tem distribuição em distâncias mais longas, com a contribuição mais significativa proveniente dos grupos CH. De fato, muitas interações podem ser estabelecidas pelos grupos OH com outras moléculas de EG e BET. Para investigar mais detalhadamente a estrutura de solvatação, pode-se examinar as funções de distribuição atômica no sistema.

A Figura 10.7 apresenta as funções de distribuição para os átomos de carbono do grupo  $COO^-$  e para o nitrogênio ( $N^+$ ) da BET. Como observado nas Figuras 10.5B e 10.6B, as MDDFs indicam que a distribuição da BET sugere sua interação com o ibuprofeno por meio de ligações de hidrogênio. A distância entre um dos oxigênios do grupo  $COO^-$  e o nitrogênio ( $N^+$ ) é de aproximadamente 3,30 Å, enquanto a distância entre o nitrogênio ( $N^+$ ) e o carbono do grupo  $COO^-$  é de cerca de 2,7 Å.

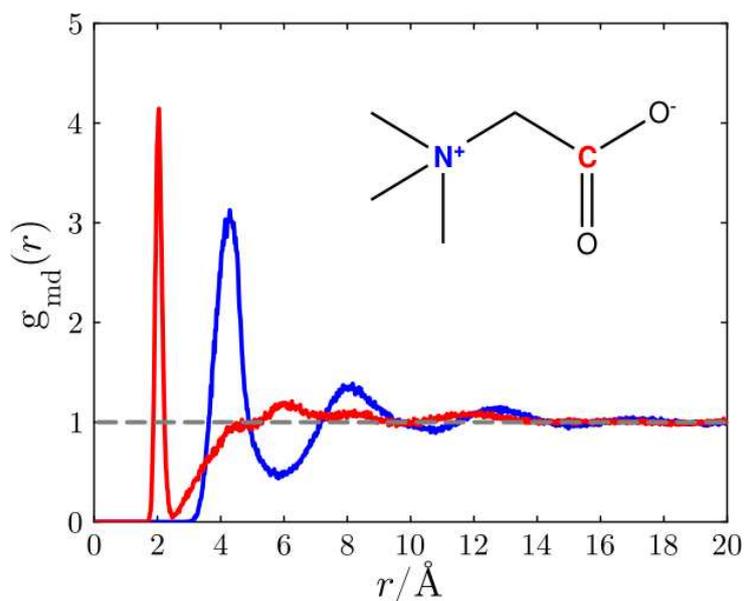


Figura 10.7: Decomposição das MDDFs em termos de átomos do solvente para o sistema BET-EG. A curva vermelha representa a distribuição dos átomos de carbono ligados ao oxigênio na betaína, indicando interações de ligação de hidrogênio com o ibuprofeno. A curva azul representa os átomos de nitrogênio na betaína, também indicando interações de ligação de hidrogênio com o ibuprofeno. O pequeno pico em 6 Å na distribuição de C e o pico em 8,5 Å na distribuição de N sugerem interações entre os grupos  $\text{COO}^-$  e  $\text{N}(\text{CH}_3)_3^+$ .

O primeiro pico na curva vermelha da Figura 10.7 representa uma distribuição que indica uma interação direta entre a BET e o ibuprofeno, especificamente por ligações de hidrogênio. Como a distância mínima para a formação de uma ligação de hidrogênio entre a BET e o ibuprofeno é de aproximadamente 1,8 Å, o primeiro pico da curva vermelha corresponde aos átomos de carbono mais próximos ligados aos átomos de oxigênio que estabelecem tais interações. Neste contexto, o primeiro pico da curva azul é atribuído ao nitrogênio da BET, que também forma ligações de hidrogênio com o ibuprofeno.

Em torno de 6 Å, observa-se um pequeno pico na função de distribuição de C, enquanto o pico correspondente ao nitrogênio (N) na mesma BET aparece em aproximadamente 8,5 Å. Os grupos  $\text{COO}^-$  distribuídos na solução a 6 Å podem estar interagindo com os grupos  $\text{N}(\text{CH}_3)_3^+$ , que aparecem em 4,5 Å.

## 10.6 Integrais de KB

A Figura 10.8 apresenta as integrais de KB para BET, curva laranja, e EG, curva azul). Como mencionado na seção de métodos, os DESs são tipicamente misturas com alta viscosidade, algumas vezes centenas de vezes maiores do que a água. Devido a essa

viscosidade, bem como à natureza das ligações de hidrogênio e das interações eletrostáticas presentes nos DESs, interações de longo alcance podem ocorrer no sistema. Consequentemente, simulações mais longas em temperaturas mais altas são necessárias para alcançar uma amostragem razoável das orientações dos solventes ao redor do soluto.

A curva laranja na Figura 10.8 inicialmente apresenta uma queda, atribuída à exclusão de moléculas de BET das regiões imediatamente próximas ao ibuprofeno, devido à repulsão estérica entre moléculas de solvente ao redor de qualquer soluto. Após essa queda inicial, observa-se um pequeno aumento entre 1,8 e 2,1 Å, associado a ligações de hidrogênio, como discutido na análise das MDDFs na Figura 10.3.

Embora a análise das MDDFs sugira que a característica mais importante para a solvatação seja a ligação de hidrogênio entre a BET e o ibuprofeno, uma análise mais detalhada das integrais de KB revela que o acúmulo total mostra que as ligações de hidrogênio representam apenas um pequeno incremento na integral de KB de BET, quando comparado ao acúmulo que ocorre a distâncias de longo alcance. Isso sugere que, embora as ligações de hidrogênio iniciem o acúmulo inicial, são as interações de longo alcance que contribuem significativamente para o aumento observado na integral de KB de BET.

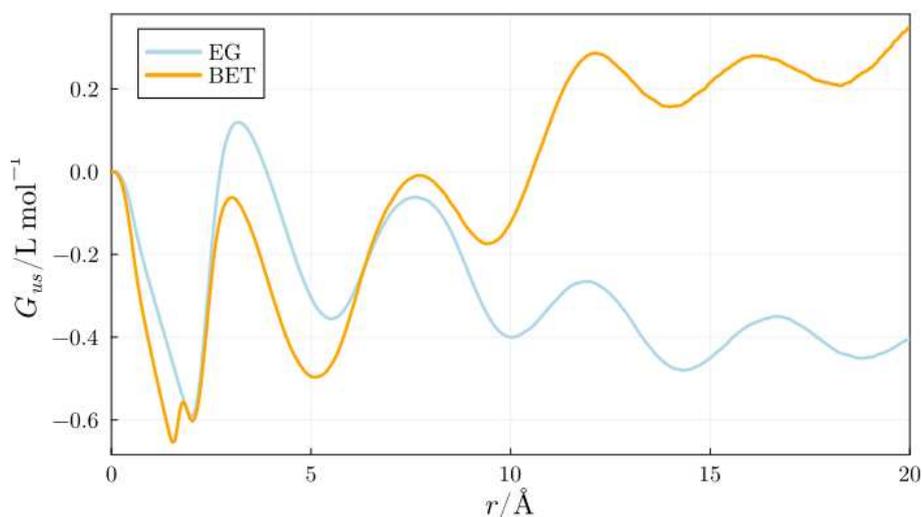


Figura 10.8: Integrais de KB para BET (curva laranja) e EG (curva azul). A queda inicial na curva laranja indica repulsão estérica, enquanto o aumento subsequente é devido às ligações de hidrogênio. A curva azul representa o comportamento do EG.

Antes de discutir o significado termodinâmico dos resultados mostrados na Figura 10.8, é importante destacar alguns detalhes técnicos relacionados às integrais de KB, que medem o acúmulo ou exclusão efetiva de um componente ao redor de outro em um sistema

de dois componentes. Por definição, todas as variações em relação à distribuição no *bulk* são integradas ao longo do volume. O regime de *bulk* começa quando a função de distribuição converge para 1. Para solventes eutéticos profundos, alcançar a convergência total para 1 é bastante desafiador, principalmente devido à alta viscosidade do solvente, que dificulta o movimento de íons e moléculas. Em casos extremos, isso leva a um comportamento vitrificado, tornando os resultados da simulação altamente dependentes da configuração inicial.

As integrais de KB para BET e EG exibem uma queda inicial relacionada ao volume de exclusão ao redor do soluto, com o tamanho da queda correspondente aos tamanhos de BET e EG. Esta queda inicial é compensada tanto para EG quanto para BET. Para BET, há um acúmulo inicial devido às ligações de hidrogênio a 1,8 Å, seguido por outro acúmulo significativo em 2,6 Å. No caso de EG, um acúmulo substancial ocorre após a primeira queda, aproximadamente em 2,7 Å. A curva continua a aumentar com flutuações até 13 Å. Após esse ponto, a curva apresenta flutuações em torno de um valor médio.

Este comportamento ilustra a dificuldade de alcançar integrais de KB convergidas para os DESs. A alta viscosidade do líquido frequentemente exige tempos de simulação mais longos e temperaturas mais altas do que o usual. Além disso, o soluto ibuprofeno é muito menor do que proteínas, como a ubiquitina, que possui 76 resíduos. Essa discrepância de tamanho leva a um problema estatístico relacionado à contagem de partículas, já que solutos maiores resultam em mais partículas no domínio do soluto, melhorando a amostragem e permitindo funções de distribuição mais convergentes e ligações de hidrogênio mais bem definidas.

## 10.7 Solvatação Preferencial do Ibuprofeno em DESs

Os parâmetros de solvatação preferencial  $\Gamma_{cp}$  foram calculados a partir das integrais de KB dos componentes do solvente em relação ao soluto, permitindo avaliar quais espécies apresentam maior afinidade pelo ibuprofeno nos sistemas simulados. A Figura 10.9 apresenta os valores de  $\Gamma$  para os sistemas formados por BET em combinação com os diferentes HBDs: EG, PG e BTO.

Os resultados mostram que o maior valor para o parâmetro de solvatação preferencial da betaína ocorre no sistema BET-EG. O valor positivo de  $\Gamma$  para BET ( 2,5) em comparação ao de EG ( 7,5) implica que o ibuprofeno é preferencialmente solvatado pela BET, sugerindo que as interações entre BET e ibuprofeno são mais favoráveis do que entre

EG e ibuprofeno. Esse comportamento pode ser atribuído à capacidade da BET de formar interações específicas, como ligações de hidrogênio, com o grupo COOH do ibuprofeno.

À medida que o comprimento da cadeia alquila do HBD aumenta, passando de EG para PG e BTO, os valores de  $\Gamma$  indicam uma mudança nas preferências de solvatação. Observa-se que o parâmetro de solvatação da BET diminui gradualmente, enquanto o parâmetro do HBD torna-se menos negativo para PG e positivo para BTO. Isso sugere que a solvatação preferencial do ibuprofeno pelo HBD se intensifica com o aumento da hidrofobicidade do doador de ligação de hidrogênio. No entanto, não há uma tendência clara ou linear nos valores de  $\Gamma$ , indicando que outros fatores, como a estrutura local e as interações cooperativas entre BET e HBD, desempenham papéis significativos na solvatação do ibuprofeno.

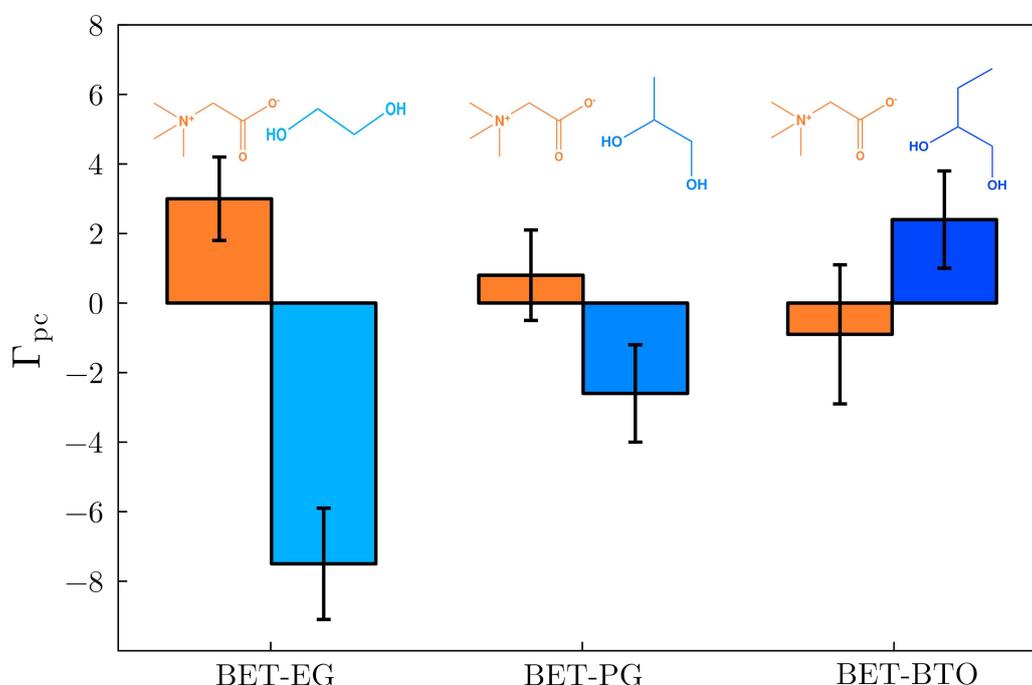


Figura 10.9: Parâmetros de solvatação preferencial ( $\Gamma$ ) para os sistemas BET-EG, BET-PG e BET-BTO. A barra laranja representa a solvatação preferencial da betaína (BET), enquanto a barra azul refere-se à solvatação pelos doadores de ligação de hidrogênio (HBDs). Valores positivos de  $\Gamma$  indicam acumulação preferencial do componente ao redor do ibuprofeno, enquanto valores negativos indicam exclusão preferencial. Observa-se que o sistema BET-EG apresenta a maior solvatação preferencial por BET, enquanto o aumento do comprimento da cadeia alquila do HBD resulta em mudanças no comportamento de solvatação preferencial, especialmente no caso do BTO, onde a solvatação do HBD torna-se dominante.

## Capítulo 11

# Considerações Finais

### 11.1 Interações específicas e acúmulo de íons

A base molecular das interações da ubiquitina com líquidos iônicos foi investigada utilizando funções de distribuição de mínima distância, com foco nos efeitos da troca de contra-íons nas distribuições dos cossolventes em relação à proteína. Como reflexo da eletroneutralidade da solução, observou-se a correlação entre as distribuições de cátions e ânions, bem como a correspondência das integrais de Kirkwood-Buff em longas distâncias, indicando que a acumulação efetiva de cátions e ânions na região ao redor da proteína na solução é equivalente. O ânion  $[\text{DCA}]^-$ , que estabelece fortes ligações de hidrogênio com a superfície da proteína, promove um acúmulo significativo de ambos os íons do líquido iônico na interface proteica.

Dentre os íons analisados, o ânion  $[\text{DCA}]^-$  apresentou a maior afinidade pela proteína, impulsionado pelas ligações de hidrogênio com os átomos da superfície proteica. Desta forma, considerando que as integrais de Kirkwood-Buff indicam a mesma acumulação efetiva para cátions e ânions, um resultado particularmente interessante foi a constatação de que, mesmo  $[\text{BMIM}]^+$  sendo um ânion volumoso e hidrofóbico, sua presença na interface proteica é aumentada quando pareado com  $[\text{DCA}]^-$ , em comparação a quando está associado a ânions como  $[\text{BF}_4]^-$ , por exemplo.

Portanto, a extensão do acúmulo de cátions no domínio proteico depende fortemente da natureza do ânion. Os resultados evidenciam que as interações mais fortes e específicas entre os íons e a superfície da proteína são preservadas após a troca de contra-íons, enquanto interações menos específicas, localizadas nas segunda e terceira camadas de solvatação, são moduladas pela identidade e concentração do contra-íon. Assim,

a composição e a concentração do líquido iônico podem ser ajustadas para obter efeitos estabilizadores ou desestabilizadores nas estruturas proteicas, fornecendo uma justificativa molecular para o design racional de solventes.

## 11.2 Interações competitivas e efeitos não-aditivos

A estrutura de solvatação também foi explorada em sistemas contendo múltiplos líquidos iônicos, com foco específico em compostos à base do cátion [EMIM]<sup>+</sup>. Identificou-se que interações específicas desempenham um papel fundamental na solvatação preferencial observada, especialmente em sistemas que contêm os ânions [DCA]<sup>-</sup> e [NO<sub>3</sub>]<sup>-</sup>.

O ânion [DCA]<sup>-</sup> apresentou maior tendência a formar ligações de hidrogênio com proteínas em comparação a outros ânions, sendo essa forte afinidade determinante para seu comportamento distinto na solução. No entanto, essa propensão é modulada pela presença de íons concorrentes, evidenciando um efeito competitivo nos parâmetros de solvatação preferencial. Observou-se que estes parâmetros geralmente diminuem em misturas, em relação aos sistemas contendo apenas [DCA]<sup>-</sup>, sugerindo que as interações entre ânions podem gerar efeitos competitivos sutis.

O impacto de ânions com menor capacidade de ligação nas propriedades de solvatação ainda não está completamente esclarecido, representando uma oportunidade relevante para investigações futuras. Um estudo mais aprofundado poderia focar na análise detalhada dos papéis individuais dos ânions em soluções mistas de líquidos iônicos, combinando abordagens experimentais e teóricas para elucidar as interações complexas que governam a solvatação nesses sistemas.

## 11.3 Acúmulo preferencial de LIs em estados desenovelados

Foram realizadas simulações considerando quatro estados distintos de enovelamento da ubiquitina. A análise dos parâmetros de solvatação preferencial revelou que, à medida que a proteína se tornava mais desenovelada, os valores de  $\Gamma_{uc}$  aumentavam. Dessa forma, os líquidos iônicos estudados apresentaram maior acúmulo relativo sobre as conformações mais abertas da ubiquitina, sendo que os sistemas contendo a conformação estendida exibiram os maiores valores de  $\Gamma_{uc}$ . O perfil das curvas do parâmetro de solvatação preferencial dos íons mostrou um comportamento semelhante na maioria dos sistemas, com

um crescimento de  $\Gamma_{uc}$  no intervalo entre 0,5 e 1,5 mol L<sup>-1</sup> do líquido iônico, seguido por um decréscimo entre 1,5 e 3,0 mol L<sup>-1</sup>.

Para ilustrar este comportamento, sistemas contendo [EMIM][BF<sub>4</sub>] e as conformações nativa e desnaturada da ubiquitina foram utilizados no cálculo de  $\Delta\Gamma$ . Em todos os casos,  $\Delta\Gamma > 0$ , indicando que o cossolvente [EMIM][BF<sub>4</sub>] acumula-se mais sobre a proteína desnaturada. Embora  $\Delta\Gamma$  tenha sido apresentado apenas para [EMIM][BF<sub>4</sub>], todos os líquidos iônicos simulados exibiram valores de  $\Gamma$  mais elevados nos sistemas com ubiquitina desnaturada e estendida, sugerindo que  $\Delta\Gamma$  seria positivo para todos os sistemas analisados. Esses resultados indicam que os líquidos iônicos testados atuam como agentes desnaturantes para a ubiquitina.

A análise das funções de distribuição de mínima distância revelou um aumento substancial nos picos das interações não específicas (2,6 Å para os ânions e 2,5 Å para os cátions) nos sistemas com ubiquitina estendida e desnaturada, quando comparados aos sistemas com ubiquitina nativa. Esse comportamento evidencia um aumento das interações não específicas entre o líquido iônico e as conformações mais abertas da ubiquitina. De modo geral, os líquidos iônicos, nas concentrações simuladas, parecem favorecer o desenovelamento da ubiquitina, sendo esse efeito particularmente mais pronunciado nos sistemas contendo o ânion [DCA]<sup>-</sup>.

## 11.4 Efeitos da hidrofobicidade do cátion

A comparação entre a estrutura e a termodinâmica de soluções de líquidos iônicos contendo os cátions [EMIM]<sup>+</sup> e [BMIM]<sup>+</sup> revelou que a maior cadeia alquílica do cátion [BMIM]<sup>+</sup> favorece a solvatação da proteína em baixas concentrações, promovendo a exclusão de água e aumentando a acumulação do cátion na superfície proteica. No entanto, em concentrações mais elevadas, as soluções de [BMIM]<sup>+</sup> apresentam menor solvatação preferencial em comparação com aquelas contendo [EMIM]<sup>+</sup>.

Tal efeito ocorre porque, em ambos os casos, os sítios da proteína com afinidade pelos cátions tornam-se saturados. No entanto, nos sistemas baseados em [EMIM]<sup>+</sup>, a solução bulk retém uma maior molaridade de água, influenciando diretamente o comportamento da solvatação. Estes resultados fornecem percepções valiosas sobre o equilíbrio entre composição, concentração e interações proteína-líquido iônico, aspectos

fundamentais para o desenvolvimento de líquidos iônicos voltados para aplicações bioquímicas específicas.

## 11.5 Solvatação do ibuprofeno em DESs

O mecanismo de solvatação em líquidos eutéticos profundos é intrinsecamente complexo devido à alta viscosidade desses sistemas. Esta característica exigiu a realização de simulações a temperaturas elevadas para minimizar a influência da configuração inicial nas funções de distribuição de mínima distância e evitar a vitrificação. Os resultados indicam que a betaína pode estabelecer tanto ligações de hidrogênio quanto interações de natureza dispersivas com o ibuprofeno. Além disso, as funções de distribuição de mínima distância sugerem que os álcoois não interagem diretamente com o ibuprofeno por meio de ligações de hidrogênio. Tal conclusão é reforçada pela decomposição das funções de distribuição em termos dos grupos funcionais dos solventes e do soluto, que revelou que o sítio COOH do ibuprofeno interage predominantemente com a betaína ao longo da simulação.

Apesar de formar ligações de hidrogênio fortes com o ibuprofeno, a betaína também possui grupos químicos capazes de estabelecer interações dispersivas. Os resultados sugerem que este equilíbrio pode ser modulado pela estrutura do doador de ligação de hidrogênio. Como o ibuprofeno é altamente hidrofóbico, um doador de ligação de hidrogênio com grupos hidrofóbicos mais extensos poderia competir de forma mais eficaz com a betaína por interações preferenciais com o soluto.

Embora o cálculo da integral de Kirkwood-Buff para solventes complexos e viscosos como os líquidos eutéticos profundos apresente desafios, esta abordagem demonstrou ser uma ferramenta valiosa para aprofundar o entendimento da estrutura desses sistemas e de suas capacidades de solvatação. A aplicação da teoria de KB permitiu vincular de maneira eficaz a estrutura local do ambiente às propriedades macroscópicas, destacando estratégias para o ajuste da composição dos líquidos eutéticos profundos com vistas a melhorar sua eficiência como solventes.

De forma geral, as simulações poderiam ser aprimoradas com o uso de métodos de amostragem ampliada, de modo que a limitação na amostragem do solvente não representasse um obstáculo para a convergência das integrais de Kirkwood-Buff. Além disso, embora não tenham sido incluídos nas simulações, sistemas contendo líquidos eutéticos

profundos frequentemente apresentam uma certa porcentagem em massa de água, o que poderia alterar a dinâmica da solvatação em algum grau. Nesse contexto, esses sistemas ainda oferecem um amplo campo para investigações no âmbito da dinâmica molecular e das análises exploradas nesta tese.

## Bibliografia

- [1] J. Wells, M. Taylor e R. Stevenson. "Ionic liquids as alternative solvents for biological reactions". Em: *Green Chemistry* 17 (2015), pp. 605–616.
- [2] S. Somers e C. Bryant. "Low-Cost Ionic Liquids with Tunable Properties". Em: *J. Chem. Eng.* 45 (2017), pp. 234–245.
- [3] T. Welton. "Room-Temperature Ionic Liquids. Solvents for Synthesis and Catalysis". Em: *Chem. Rev.* 99 (1999), pp. 2071–2084.
- [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] 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.
- [6] Pankaj Bharmoria et al. "Do Ionic Liquids Exhibit the Required Characteristics to Dissolve, Extract, Stabilize, and Purify Proteins? Past-Present-Future Assessment". Em: *Chem. Rev.* (2024). Published as part of Chem. Rev. virtual special issue "Ionic Liquids for Diverse Applications".
- [7] K. Ghandi. "A Review of Ionic Liquids, Their Limits and Applications". Em: *Green and Sustainable Chemistry* 4.1 (2014), pp. 44–53. DOI: 10.4236/gsc.2014.41008.
- [8] E. Maginn e D. Jiang. "Thermodynamic Properties of Ionic Liquids: Insights from Molecular Simulation". Em: *J. Phys. Chem. B* 116 (2012), pp. 5243–5249.
- [9] P. Wagle e S. Devereux. "Thermodynamic Stability of Ionic Liquids at Room Temperature". Em: *J. Chem. Phys.* 140 (2014), p. 124107.

- [10] 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.
- [11] H. Zhao. "Protein Stabilization and Enzyme Activation in Ionic Liquids: Specific Ion Effects". Em: *J Chem Technol Biotechnol* 91 (2016), pp. 25–50. DOI: 10.1002/jctb.4837.
- [12] 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.
- [13] Z. Lei et al. "Introduction: Ionic Liquids". Em: *Chem. Rev.* 117 (2017), pp. 6633–6635. ISSN: 15206890. DOI: 10.1021/acs.chemrev.7b00246.
- [14] M. Freemantle. "DESIGNER SOLVENTS". Em: *Chem. Eng. News Archive* 76 (1998), pp. 32–37.
- [15] Karina Shimizu et al. "Self-Organization in Ionic Liquids: From Bulk to Interfaces and Films". Em: *Journal of the Brazilian Chemical Society* 27.2 (2016). Printed in Brazil – ©2016 Sociedade Brasileira de Química, pp. 349–362. DOI: 10.5935/0103-5053.20150274.
- [16] R. Hayes, G. G. Warr e R. Atkin. "Structure and Nanostructure in Ionic Liquids". Em: *Chem. Rev.* 115 (2015), pp. 6357–6426. DOI: 10.1021/cr500411q.
- [17] V.I. Pârvulescu e C. Hardacre. "Catalysis in ionic liquids". Em: *Chem. Rev.* 107 (2007), pp. 2615–2665.
- [18] J.P. Hallett e T. Welton. "Room-temperature ionic liquids: solvents for synthesis and catalysis. 2". Em: *Chem. Rev.* 111 (2011), pp. 3508–3576.
- [19] R. F. de Souza et al. "Room temperature dialkylimidazolium ionic liquid-based fuel cells". Em: *Electrochemistry Communications* 5.8 (2003), pp. 728–731. DOI: 10.1016/S1388-2481(03)00173-5.
- [20] J. Dupont, R. F. de Souza e P. A. Z. Suarez. "Ionic Liquid (Molten Salt) Phase Organometallic Catalysis". Em: *Chem. Rev.* 102.10 (2002), pp. 3667–3692. DOI: 10.1021/cr010338r.
- [21] F. U. Shah, O. Gnezdilov e A. Filippov. "Ion dynamics in halogen-free phosphonium bis(salicylato)borate ionic liquid electrolytes for lithium-ion batteries". Em: *Phys. Chem. Chem. Phys.* 19 (2017), pp. 16721–16730. DOI: 10.1039/C7CP02722B.

- [22] D. R. MacFarlane et al. "Energy applications of ionic liquids". Em: *Energy Environ. Sci.* 7 (2014), pp. 232–250. DOI: 10.1039/C3EE42099.
- [23] S. K. Singh e A. W. Savoy. "Ionic liquids synthesis and applications: an overview". Em: *J. Mol. Liq.* 297 (2020), p. 112038. DOI: 10.1016/j.molliq.2019.112038.
- [24] S. Spange et al. "Complementary interpretation of ET(30) polarity parameters of ionic liquids". Em: *Phys. Chem. Chem. Phys.* 22 (2020), pp. 9954–9966. DOI: 10.1039/D0CP01480J.
- [25] J. Dupont e P. A. Z. Suarez. "Physico-chemical processes in imidazolium ionic liquids". Em: *Phys. Chem. Chem. Phys.* 8 (2006). Invited Article, pp. 2441–2452. DOI: 10.1039/B602046A.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] 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.
- [31] F. A. e Silva et al. "Temperature Dependency of Aqueous Biphasic Systems: an Alternative Approach for Exploring the Differences Between Coulombic-Dominated Salts and Ionic Liquids". Em: *Chem. Commun.* 53 (2017), pp. 7298–7301. DOI: 10.1039/c7cc03484h.
- [32] T. Le et al. "Quantitative Structure–Property Relationship Modeling of Diverse Materials Properties". Em: *Chem. Rev.* 112 (2012), pp. 2889–2919. DOI: 10.1021/cr200066h.

- [33] T. L. Greaves e C. J. Drummond. "Protic Ionic Liquids: Evolving Structure-Property Relationships and Expanding Applications". Em: *Chem. Rev.* 115 (2015), pp. 11379–11448. DOI: 10.1021/acs.chemrev.5b00158.
- [34] J. Dupont. "On the solid, liquid and solution structural organization of imidazolium ionic liquids". Em: *Journal of the Brazilian Chemical Society* 15.3 (2004), pp. 341–350. DOI: 10.1590/S0103-50532004000300002.
- [35] B. Sepehri. "A review on created QSPR models for predicting ionic liquids properties and their reliability from chemometric point of view". Em: *J. Mol. Liq.* 297 (2020), p. 112013. DOI: 10.1016/j.molliq.2019.112013.
- [36] G. Saielli. "Bulk Structure Elucidation: Computational Spectroscopy of Ionic Liquids for Bulk Structure Elucidation". Em: *Adv. Theory Simul.* 1 (2018), p. 1870026. DOI: 10.1002/adts.201800026.
- [37] K. Dong et al. "Multiscale Studies on Ionic Liquids". Em: *Chem. Rev.* 117 (2017), pp. 6636–6695. DOI: 10.1021/acs.chemrev.6b00776.
- [38] H. Tokuda et al. "Physicochemical Properties and Structures of Room Temperature Ionic Liquids. 1. Variation of Anionic Species". Em: *J. Phys. Chem. B* 108 (2004), pp. 16593–16600. DOI: 10.1021/jp047640w.
- [39] H. Tokuda et al. "How Ionic Are Room-Temperature Ionic Liquids? An Indicator of the Physicochemical Properties". Em: *J. Phys. Chem. B* 110 (2006), pp. 19593–19600. DOI: 10.1021/jp064159v.
- [40] Zheng Yang e Wen Pan. "Ionic liquids: Green solvents for nonaqueous biocatalysis". Em: *Enzyme and Microbial Technology* 44.5 (2009), pp. 102–111. DOI: 10.1016/j.enzmictec.2008.11.017.
- [41] Mara G. Freire et al. "High-performance biocatalysis in ionic liquids". Em: *Chem. Commun.* 47.9 (2011), pp. 901–913. DOI: 10.1039/C0CC03165A.
- [42] Sang-Eun Park e Romas J. Kazlauskas. "Biocatalysis in ionic liquids - advantages beyond green technology". Em: *Current Opinion in Biotechnology* 18.4 (2007), pp. 1–7. DOI: 10.1016/j.copbio.2007.08.014.
- [43] M. G. Freire et al. "Aqueous Biphasic Systems: A Boost Brought About by Using Ionic Liquids". Em: *Chem. Soc. Rev.* 41 (2012), pp. 4966–4995.

- [44] H. Zhao e S. Xia. "Impact of ions on the conformation and stability of proteins: Insights from molecular dynamics simulations". Em: *Phys. Chem. Chem. Phys.* 18 (2016), pp. 8798–8806. DOI: 10.1039/C5CP06717A.
- [45] J. N. A. Canongia Lopes e A. A. H. Padua. "Nanostructural Organization in Ionic Liquids". Em: *J. Phys. Chem. B* 110 (2006), pp. 3330–3335.
- [46] T. Y. Lin e Serge N. Timasheff. "Why do some organisms use a urea-based metabolism? A thermodynamic analysis of urea effects on proteins". Em: *Biochemistry* 48.29 (2009), pp. 7401–7408. DOI: 10.1021/bi9004707.
- [47] Edson P. Melo, Peter Fojan e Soren B. Petersen. "How co-solvents modulate protein function: a multi-timescale molecular dynamics study of ribonuclease A in alcohol-water mixtures". Em: *Biophysical Journal* 95.11 (2008), pp. 5627–5636. DOI: 10.1529/biophysj.108.138800.
- [48] E. E. Lattman e G. D. Rose. "Protein folding—what's the question?" Em: *PNAS USA* 90.2 (1993), pp. 439–441. DOI: 10.1073/pnas.90.2.439.
- [49] 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.
- [50] 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.
- [51] 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.
- [52] 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.
- [53] 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.
- [54] 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.

- [55] R. L. Baldwin. "How Hofmeister ion interactions affect protein stability". Em: *Biophys. J.* 71 (1996), pp. 2056–2063.
- [56] J. Dunbar et al. "The Effect of Denaturants on Protein Structure". Em: *Protein Sci.* 6.8 (1997), pp. 1727–1733.
- [57] J. C. Lee e S. N. Timasheff. "The stabilization of proteins by sucrose". Em: *J. Biol. Chem.* 256 (1981), pp. 7193–7201.
- [58] C. Schröder. "Proteins in Ionic Liquids: Current Status of Experiments and Simulations". Em: *Top. Curr. Chem.* 375 (2017), p. 25.
- [59] A. Salis e B. W. Ninham. "Hofmeister Effects in Biology: New Insights from Ionic Liquids". Em: *Chem. Soc. Rev.* 43 (2014), p. 7358. DOI: 10.1039/c4cs00144c.
- [60] Y. Zhang e P. S. Cremer. "Interactions between Proteins and Ionic Liquids: Implications for Biotechnology". Em: *Current Opinion in Chemical Biology* 10 (2006), p. 658. DOI: 10.1016/j.cbpa.2006.09.020.
- [61] G. S. Lim et al. "Protein Folding and Unfolding in Ionic Liquids: Structural Stability and Dynamics". Em: *J. Phys. Chem. B.* 118 (2014), p. 10444. DOI: 10.1021/jp504356m.
- [62] J. L. Anderson et al. "Structure and Interactions in Ionic Liquids and Their Potential as Solvents". Em: *J. Am. Chem. Soc.* 124 (2002), p. 14247. DOI: 10.1021/ja027728+.
- [63] M. Klähn et al. "Insights into the Structure and Dynamics of Ionic Liquids from Simulations". Em: *Phys. Chem. Chem. Phys.* 13 (2011), p. 1649. DOI: 10.1039/c0cp01501k.
- [64] M. Haberler e O. Steinhauser. "Dynamic Properties of Proteins in Ionic Liquid Solutions". Em: *Phys. Chem. Chem. Phys.* 13 (2011), p. 17994. DOI: 10.1039/c1cp21016e.
- [65] K. P. Gregory et al. "Understanding Specific Ion Effects and the Hofmeister Series". Em: *Phys. Chem. Chem. Phys.* 24.21 (2022), pp. 12682–12718.
- [66] D. R. MacFarlane et al. "Lewis base ionic liquids". Em: *Chem. Commun.* 18 (2006), pp. 1905–1917.
- [67] P. Lozano et al. "Dynamic structure–function relationships in enzyme stabilization by ionic liquids". Em: *Biocat. Biotrans.* 23 (2009), pp. 169–176.
- [68] P. Lozano et al. "Stabilization  $\alpha$ -Chymotrypsin by Ionic Liquids in Transesterification Reactions". Em: *Biotech. Bioeng.* 75 (2001), pp. 563–569.

- [69] R. Tomaš et al. "Interactions in aqueous solutions of imidazolium chloride ionic liquids [C<sub>n</sub>mim][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.
- [70] M. C. Manning et al. "Stability of Protein Pharmaceuticals: An Update". Em: *Pharmaceutical Research* 27 (2010), pp. 544–575. DOI: 10.1007/s11095-009-0045-6.
- [71] A. Nisthal et al. "Protein Stability Engineering Insights Revealed by Domain-Wide Comprehensive Mutagenesis". Em: *PNAS USA* 116 (2019), pp. 16367–16377. DOI: 10.1073/pnas.1903901116.
- [72] R. de la Rica e H. Matsui. "Applications of Peptide and Protein-Based Materials in Bionanotechnology". Em: *Chem. Soc. Rev.* 39 (2010), pp. 3499–3509. DOI: 10.1039/b921297g.
- [73] P. G. Chandler et al. "Strategies for Increasing Protein Stability". Em: *Methods in Molecular Biology* 2073 (2020), pp. 163–181. DOI: 10.1007/978-1-4939-9869-2\_12.
- [74] Magdalena Kowacz et al. "Hofmeister effects of ionic liquids in protein crystallization: Direct and water-mediated interactions". Em: *CrystEngComm* 14.15 (2012), p. 4912. ISSN: 1466-8033. DOI: 10.1039/c2ce25129a. URL: <http://xlink.rsc.org/?DOI=c2ce25129a>.
- [75] X. Chen, Y. Ji e J. J. A. Wang. "Improvement on the crystallization of lysozyme in the presence of hydrophilic ionic liquid". Em: *Journal of Crystal Growth* 312.9 (2010), pp. 2241–2248. DOI: 10.1016/j.jcrysgro.2010.03.005.
- [76] R. A. Judge et al. "The effect of ionic liquids on protein crystallization and X-ray diffraction resolution". Em: *Crystal Growth & Design* 9.8 (2009), pp. 3463–3469. DOI: 10.1021/cg900367e.
- [77] Qi Han et al. "Protein aggregation and crystallization with ionic liquids: Insights into the influence of solvent properties". Em: *Journal of Colloid and Interface Science* 608.Part 2 (2022), pp. 1173–1190. DOI: 10.1016/j.jcis.2021.10.087.
- [78] S. R. Tomlinson et al. "Solubility of the Corn Protein Zein in Imidazolium-Based Ionic Liquids". Em: *Industrial Engineering Chemistry Research* 53 (2014), pp. 2293–2298. DOI: 10.1021/ie402836n.

- [79] F. Hasan et al. "Ionic Liquids as Buffer Additives in Ionic Liquid Polyacrylamide Gel Electrophoresis Separation of Mixtures of Low and High Molecular Weight Proteins". Em: *RSC Advances* 5 (2015), pp. 69229–69237. DOI: 10.1039/C5RA08883C.
- [80] D. Harries e J. Rösgen. "A practical guide on how osmolytes modulate macromolecular properties". Em: *Methods Cell Biol.* 84 (2008), pp. 679–735.
- [81] J. L. England e G. Haran. "Role of solvation effects in protein denaturation: from thermodynamics to single molecules and back". Em: *Annual Review of Physical Chemistry* 62 (2011), pp. 257–277. DOI: 10.1146/annurev-physchem-032210-103518.
- [82] Ira. N. Levine. *Quantum Chemistry (7th Edition)*. McGraw-Hill, 2009. ISBN: 9780321803450.
- [83] 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.
- [84] C. Nick Pace et al. "Contribution of Hydrophobic Interactions to Protein Stability". Em: *J. Mol. Biol.* 408.3 (mar. de 2011), pp. 514–528. DOI: 10.1016/j.jmb.2011.02.053.
- [85] Darin M. Taverna e Richard A. Goldstein. "Why are proteins marginally stable?" Em: *Proteins* 46.1 (2002), pp. 105–109. DOI: 10.1002/prot.10016.
- [86] A. Ben-Naim. *Statistical thermodynamics for chemists and biochemists*. New York, NY: Springer, 2013. DOI: 10.1007/978-1-4757-1598-9.
- [87] D. Chandler. *Introduction to Modern Statistical Mechanics*. 1st. Vol. 1. Oxford University Press, 1987, p. 288. ISBN: 978-0195042771.
- [88] D. A. McQuarrie. *Statistical Mechanics*. 1st. Vol. 1. University Science Books, 2000, p. 641. ISBN: 1891389157.
- [89] Vinicius Piccoli. "Estudo da solvatação de proteínas por líquidos iônicos usando funções de distribuição de mínima-distância". Diss. de maestr. Universidade Estadual de Campinas, Instituto de Química, 2020. DOI: <https://doi.org/10.47749/T/UNICAMP.2020.1128941>. URL: <https://repositorio.unicamp.br/acervo/detalhe/1128941>.
- [90] 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.

- [91] 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.
- [92] 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.
- [93] D. Frenkel e B. Smit. *Understanding Molecular Simulations: from Algorithms to Applications*. 2nd. Academic Press - London, 2002. ISBN: 9780080519982.
- [94] Martin J. Field, Paul A. Bash e Martin Karplus. "A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations". Em: *J. Comput. Chem.* 11.6 (jul. de 1990), pp. 700–733. DOI: 10.1002/jcc.540110605.
- [95] 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.
- [96] 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.
- [97] A. M. Namba, V. B. da Silva e C. H. T. P. da Silva. "Dinâmica molecular: Teoria e aplicações em planejamento de fármacos". Em: *Eclét. Quím.* 33 (2008), pp. 13–24. DOI: 10.1590/S0100-46702008000400002.
- [98] 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.
- [99] 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.
- [100] 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.

- [101] 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.
- [102] 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.
- [103] 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.
- [104] Christian Schröder. "Comparing Reduced Partial Charge Models with Polarizable Simulations of Ionic Liquids". Em: *Phys. Chem. Chem. Phys.* 14 (2012), pp. 3089–3102. DOI: 10.1039/c2cp23329k.
- [105] Jesse G. McDaniel e Arun Yethiraj. "Influence of Electronic Polarization on the Structure of Ionic Liquids". Em: *J. Phys. Chem. Lett.* 9 (2018), pp. 4765–4770. DOI: 10.1021/acs.jpcclett.8b02120.
- [106] Dmitry Bedrov et al. "Influence of Polarization on Structural, Thermodynamic, and Dynamic Properties of Ionic Liquids Obtained from Molecular Dynamics Simulations". Em: *J. Phys. Chem. B* 114 (2010), pp. 4984–4997. DOI: 10.1021/jp911670f.
- [107] T. Yan et al. "Molecular Dynamics Simulation of Ionic Liquids: The Effect of Electronic Polarizability". Em: *J. Phys. Chem. B* 108 (2004), pp. 11877–11881. DOI: 10.1021/jp047619y.
- [108] Dmitry Bedrov et al. "Molecular Dynamics Simulations of Ionic Liquids and Electrolytes Using Polarizable Force Fields". Em: *Chem. Rev.* 119 (2019), pp. 7940–7995. DOI: 10.1021/acs.chemrev.8b00763.
- [109] L. Verlet. "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules". Em: *Phys. Rev.* 159 (1967), pp. 98–103.
- [110] Terrell L Hill e Dover Publications. *Zyx*. ISBN: 9780486652429.
- [111] 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.
- [112] 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.

- [113] Hans C. Andersen. "Molecular dynamics simulations at constant pressure and/or temperature". Em: *J. Chem. Phys.* 72.4 (1980), pp. 2384–2393. DOI: 10.1063/1.439486.
- [114] 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.
- [115] 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.
- [116] M. Parrinello e A. Rahman. "Polymorphic transitions in single crystals: A new molecular dynamics method". Em: *J. Appl. Phys.* 52 (1981), pp. 7182–7190.
- [117] M. Parrinello e A. Rahman. "Strain fluctuations and elastic constants". Em: *J. Chem. Phys.* 76 (1982), pp. 2662–2666.
- [118] M. P. Allen e D. J. Tildesley. *Computer simulation of liquids*. first edition. Oxford University Press, 1987, p. 385. ISBN: 0198553757.
- [119] 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.
- [120] 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.
- [121] W. F. van Gunsteren e H. J. C. Berendsen. "Molecular dynamics computer simulations; Methodology, applications and perspectives in chemistry". Em: *Angewandte Chemie* 102 (1990), pp. 1020–1055. DOI: 10.1002/ange.19901020907.
- [122] I. G. Tironi et al. "A generalized reaction field method for molecular dynamics simulations". Em: *J. Chem. Phys.* 102 (1995), pp. 5451–5458. DOI: 10.1063/1.469273.
- [123] 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.
- [124] R. W. Hockey e J. W. Eastwood. *Computer Simulation using Particles*. McGraw-Hill, 1981, p. 540. ISBN: 9780852743928.

- [125] J. Kubelka, J. Hofrichter e W. A. Eaton. "The protein folding 'speed limit'". Em: *Curr. Opin. Struct. Biol.* 14.1 (fev. de 2004), pp. 76–88. DOI: 10.1016/j.sbi.2004.01.013.
- [126] Barbara Scalvini, Vahid Sheikhhassani e Alireza Mashaghi. "Topological principles of protein folding". Em: *Phys. Chem. Chem. Phys.* 23.37 (2021), pp. 21316–21328. DOI: 10.1039/D1CP03390E.
- [127] Z. Pouramini, A. Mohebbi e M. H. Kowsari. "Atomistic insights into the thermodynamics, structure, and dynamics of ionic liquid 1-hexyl-3-methylimidazolium hexafluorophosphate via molecular dynamics study". Em: *J. Mol. Liq.* 246 (2017), pp. 39–47. ISSN: 0167-7322. DOI: 10.1016/j.molliq.2017.09.043.
- [128] 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.
- [129] G. Goldstein et al. "Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells". Em: *Proceedings of the National Academy of Sciences USA* 72.1 (1975), pp. 11–15. DOI: 10.1073/pnas.72.1.11.
- [130] S. Vijay-Kumar et al. "Three-dimensional structure of ubiquitin at 2.8 Å resolution". Em: *Proceedings of the National Academy of Sciences USA* 82.11 (1985), pp. 3582–3585. DOI: 10.1073/pnas.82.11.3582.
- [131] Aaron Ciechanover e Kazuhiro Iwai. "The ubiquitin system: from basic mechanisms to the patient bed". Em: *IUBMB Life* 56.4 (abr. de 2004), pp. 193–201. DOI: 10.1080/1521654042000223616.
- [132] Aaron Ciechanover, Ariel Orian e Allen L. Schwartz. "The ubiquitin-mediated proteolytic pathway: mode of action and clinical implications". Em: *J. Cell. Biochem. Suppl.* 34 (2000), pp. 40–51. DOI: 10.1002/(sici)1097-4644(2000)77:34+<40::aid-jcb9>3.0.co;2-6.
- [133] Aaron Ciechanover. "The ubiquitin-mediated proteolytic pathway: mechanisms of action and cellular physiology". Em: *Biological Chemistry Hoppe-Seyler* 375.9 (set. de 1994), pp. 565–581. DOI: 10.1515/bchm3.1994.375.9.565.
- [134] K. D. Wilkinson e T. K. Audhya. "Stimulation of ATP-dependent proteolysis requires ubiquitin with the COOH-terminal sequence Arg-Gly-Gly". Em: *J. Biol. Chem.* 256.17 (set. de 1981), pp. 9235–9241.

- [135] R. E. Lenkinski et al. "Nuclear magnetic resonance studies of the denaturation of ubiquitin". Em: *Biochimica et Biophysica Acta* 494 (1977), pp. 126–130. DOI: 10.1016/0005-2795(77)90140-4.
- [136] P. L. Wintrode, G. I. Makhatadze e P. L. Privalov. "Thermodynamics of ubiquitin unfolding". Em: *Proteins* 18.3 (1994), pp. 246–253. DOI: 10.1002/prot.340180305.
- [137] Ivan Dikic, Satoshi Wakatsuki e Kim J. Walters. "Ubiquitin-binding Domains – from Structures to Functions". Em: *Nature Reviews Molecular Cell Biology* 10.10 (2009), pp. 659–671. DOI: 10.1038/nrm2767. URL: <https://doi.org/10.1038/nrm2767>.
- [138] K. Newman. "Kirkwood-Buff solution theory: Derivation and applications". Em: *Chem. Soc. Rev.* 23 (1994), pp. 31–40. ISSN: 03060012. DOI: 10.1039/CS9942300031.
- [139] 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.
- [140] 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.
- [141] D. G. Hall. "The Kirkwood–Buff theory of solutions". Em: *Trans. Faraday Soc.* 67 (1971), pp. 2516–2524. DOI: 10.1039/TF9716702516.
- [142] J. G. Kirkwood. "Statistical mechanics of fluid mixtures". Em: *J. Chem. Phys.* 3 (1935), pp. 300–313. ISSN: 00219606. DOI: 10.1063/1.1749657.
- [143] 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.
- [144] S.N. Timasheff. "Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components". Em: *Proc. Natl. Acad. Sci. U. S. A.* 99 (2002), pp. 9721–9726.
- [145] S. N. Timasheff. "The control of protein stability and association by weak interactions with water: how do solvents affect these processes?" Em: *Annual Review of Biophysics and Biomolecular Structure* 22 (1993), pp. 67–97. DOI: 10.1146/annurev.bb.22.060193.000435.
- [146] 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.
- [147] 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.
- [148] 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.
- [149] D. Stigter. "Interactions in aqueous solutions. IV. Light scattering of colloidal electrolytes." Em: *Journal of Physical Chemistry* 64 (1960), pp. 842–846.
- [150] Serge N. Timasheff. "Control of Protein Stability and Reactions by Weakly Interacting Cosolvents: The Simplicity of the Complicated". Em: *Department of Biochemistry, Brandeis University, Waltham, Massachusetts* (1993).
- [151] D.V.D. Spoel et al. "GROMACS: Fast, flexible, and free". Em: *J. Comput. Chem.* 26 (2005), pp. 1701–1718.
- [152] B. Kohnke et al. "A Flexible, GPU-Powered Fast Multipole Method for Realistic Biomolecular Simulations in Gromacs". Em: *Biophysical Journal* 112 (2017), 448a.
- [153] L. Martínez et al. "PACKMOL: A package for building initial configurations for molecular dynamics simulations". Em: *J. Comp. Chem.* 30 (2009), pp. 2157–2164.
- [154] 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.
- [155] B. Doherty, X. Zhong e O. Acevedo. "Virtual Site OPLS Force Field for Imidazolium-Based Ionic Liquids". Em: *J. Phys. Chem. B* 122 (2018), pp. 2962–2974.
- [156] M.J. Robertson, J. Tirado-Rives e W.L. Jorgensen. "Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field". Em: *J. Chem. Theory Comput.* 11 (2015), pp. 3499–3509.
- [157] W.L. Jorgensen et al. "Comparison of simple potential functions for simulating liquid water". Em: *J. Chem. Phys.* 79 (1983), pp. 926–935.
- [158] H.J.C. Berendsen et al. "Molecular dynamics with coupling to an external bath". Em: *J. Chem. Phys.* 81 (1984), pp. 3684–3690.

- [159] T. McSherry. "A General Steepest Descent Algorithm". Em: *IEEE Transactions on Aerospace and Electronic Systems* AES-12 (1976), pp. 12–22.
- [160] Lone Madsen et al. "Ion-specific effects of hydrophilic and hydrophobic ionic liquids on protein stability: A comprehensive study". Em: *J. Phys. Chem. B.* 121.16 (2017), pp. 3595–3604. DOI: 10.1021/acs.jpcc.6b10683.
- [161] L. Martínez. "ComplexMixtures.jl: Investigating the structure of solutions of complex-shaped molecules from a solvent-shell perspective". Em: *J. Mol. Liq.* 347 (2022), p. 117945.
- [162] D.R. Canchi et al. "Molecular Mechanism for the Preferential Exclusion of Osmolytes from Protein Surfaces". Em: *Biophys. J.* 104 (2013), 189a.
- [163] C.P. Woodbury Jr. *Introduction to Macromolecular Binding Equilibria*. Boca Raton, FL: CRC Press, 2007. URL: <https://doi.org/10.1201/b12823>.
- [164] V. Piccoli e L. Martínez. "Correlated counterion effects on the solvation of proteins by ionic liquids". Em: *Journal of Molecular Liquids* 320 (2020), p. 114347.
- [165] William Humphrey, Andrew Dalke e Klaus Schulten. "VMD – Visual Molecular Dynamics". Em: *J. Mol. Graph.* 14.1 (1996), pp. 33–38. DOI: 10.1016/0263-7855(96)00018-5.
- [166] V. Piccoli e L. Martínez. "Ionic liquid solvation of proteins in native and denatured states". Em: *Journal of Molecular Liquids* 363 (2022), p. 119953. DOI: 10.1016/j.molliq.2022.119953.
- [167] Franz Hofmeister. "Zur Lehre von der Wirkung der Salze". Em: *Archiv für experimentelle Pathologie und Pharmakologie* 24 (1888), pp. 247–260.
- [168] F. Hofmeister. "Zur Lehre von der Wirkung der Salze (II)". Em: *Arch. Exp. Pathol. Pharmacol.* 25 (1889), pp. 1–30.
- [169] Y. Zhang e P. S. Cremer. "Interactions between macromolecules and ions: The Hofmeister series". Em: *Curr. Opin. Chem. Biol.* 10 (2006), pp. 658–663.
- [170] K. D. Collins e M. W. Washabaugh. "The Hofmeister effect and the behaviour of water at interfaces". Em: *Q. Rev. Biophys.* 18 (1985), pp. 323–422.
- [171] W. Kunz, J. Henle e B. W. Ninham. "Zur Lehre von der Wirkung der Salze: Specific ion effects in colloid and interface science". Em: *Curr. Opin. Colloid Interface Sci.* 9 (2004), pp. 19–37.

- [172] J. Lewith e F. Hofmeister. "Über den Einfluss von Salzen auf die Eiweisslösung". Em: *Z. Physiol. Chem.* 12 (1888), pp. 132–145.
- [173] J. Lewith e F. Hofmeister. "Untersuchungen über den Einfluss verschiedener Ionen auf die Ausfällung von Eiweisslösungen". Em: *Z. Physiol. Chem.* 13 (1889), pp. 99–120.
- [174] G. S. Manning. "Counterion condensation theory of polyelectrolytes: a review". Em: *J. Chem. Phys.* 51 (1969), pp. 924–933.
- [175] Y. Marcus. "Effects of ions on the structure of water: Implications for solvation and ionic effects on biochemical processes". Em: *Chem. Rev.* 109 (2009), pp. 1346–1370.
- [176] M. G. Cacace, E. M. Landau e J. J. Ramsden. "The Hofmeister series: salt and solvent effects on interfacial phenomena". Em: *Q. Rev. Biophys.* 30 (1997), pp. 241–277.
- [177] V. Piccoli e L. Martínez. "Competitive Effects of Anions on Protein Solvation by Aqueous Ionic Liquids". Em: *The J. Phys. Chem. B.* 128 (2024). Published as part of The J. Phys. Chem. B. virtual special issue "COIL-9: 9th Congress on Ionic Liquids", pp. 7792–7802. DOI: 10.1021/acs.jpcc.4c03735.
- [178] R. A. Miranda-Quintana e J. Smiatek. "Specific Ion Effects in Different Media: Current Status and Future Challenges". Em: *J. Phys. Chem. B* 125.51 (2021), pp. 13840–13849.
- [179] H. I. Okur et al. "Beyond the Hofmeister series: Ion-specific effects on proteins and their biological function". Em: *J. Phys. Chem. B.* 121.9 (2017), pp. 1997–2014. DOI: 10.1021/acs.jpcc.6b10797.
- [180] Y. Zhang e P. S. Cremer. "The Inverse and Direct Hofmeister Series for Lysozyme". Em: *Proc. Natl. Acad. Sci. U.S.A.* 106.36 (2009), pp. 15249–15253.
- [181] M. M. Ries-Kautt e A. F. Ducruix. "Relative Effectiveness of Various Ions on the Solubility and Crystal Growth of Lysozyme". Em: *J. Biol. Chem.* 264.2 (1989), pp. 745–748.
- [182] T. B. Robertson. "Contributions to the Theory of the Mode of Action of Inorganic Salts upon Proteins in Solution". Em: *J. Biol. Chem.* 9.3 (1911), pp. 303–326.
- [183] 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.
- [184] 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.

- [185] D. Diddens et al. "Aqueous ionic liquids and their influence on peptide conformations: denaturation and dehydration mechanisms". Em: *Phys. Chem. Chem. Phys.* 19 (2017), pp. 20430–20440.
- [186] E. L. Smith, A. P. Abbott e K. S. Ryder. "Deep eutectic solvents (DESs) and their applications". Em: *Chem. Rev.* 114 (2014), pp. 11060–11082. DOI: 10.1021/cr300162p.
- [187] Benworth B. Hansen et al. "Deep Eutectic Solvents: A Review of Fundamentals and Applications". Em: *Chem. Rev.* 121 (2021), pp. 1232–1285. DOI: 10.1021/acs.chemrev.0c00385.
- [188] A. P. Abbott et al. "Preparation of novel, moisture-stable, Lewis-acidic ionic liquids containing quaternary ammonium salts with functional side chains". Em: *Chem. Commun.* (2001), pp. 2010–2020.
- [189] Andrew P. Abbott et al. "Novel Solvent Properties of Choline Chloride Urea Mixtures". Em: *Chem. Commun.* 0 (2003), pp. 70–71. DOI: 10.1039/B210714G.
- [190] Andrew P. Abbott et al. "Eutectic-based Ionic Liquids with Metal-containing Anions and Cations". Em: *Chemistry - A European Journal* 13 (2007), pp. 6495–6501. DOI: 10.1002/chem.200700091.
- [191] P.E. Smith. "Cosolvent Interactions with Biomolecules: Relating Computer Simulation Data to Experimental Thermodynamic Data". Em: *J. Phys. Chem. B.* 108 (2004), pp. 18716–18724.
- [192] Diana O. Abranches et al. "Phenolic Hydrogen Bond Donors in the Formation of Non-ionic Deep Eutectic Solvents: The Quest for Type V DES". Em: *Chem. Commun.* 55 (2019), pp. 10253–10256. DOI: 10.1039/C9CC04846D.
- [193] Adriaan van den Bruinhorst et al. "Deep eutectic solvents on a tightrope: balancing the entropy and enthalpy of mixing". Em: *Faraday Discussions* 253 (2024). Previous Article Next Article. DOI: 10.1039/D4FD00048J.
- [194] Liliana P. Silva et al. "What a difference a methyl group makes – probing choline–urea molecular interactions through urea structure modification". Em: *Phys. Chem. Chem. Phys.* 21 (2019), pp. 18278–18289. DOI: 10.1039/C9CP03552D.
- [195] C. R. Ashworth et al. "Doubly ionic hydrogen bond interactions within the choline chloride–urea deep eutectic solvent". Em: *Phys. Chem. Chem. Phys.* 18 (2016), pp. 18145–18160. DOI: 10.1039/C6CP03199E.

- [196] O. S. Hammond e K. J. Edler. *Structure and implications in Deep Eutectic Solvents: Synthesis, Properties, and Applications*. Wiley, 2019, pp. 25–42. ISBN: 9781119468738.
- [197] Christopher M. Baker. “Polarizable force fields for molecular dynamics simulations of biomolecules”. Em: *WIREs Comput. Mol. Sci.* 5 (2015), pp. 241–254. DOI: 10.1002/wcms.1215.
- [198] David R. Lide, ed. *CRC Handbook of Chemistry and Physics*. 84th. Boca Raton, FL: CRC Press, 2003.
- [199] J. K. Gregory et al. “The water dipole moment in water clusters”. Em: *Science* 275 (1997), pp. 814–817.
- [200] Y. S. Badyal et al. “Electron distribution in water”. Em: *J. Chem. Phys.* 112 (2000), pp. 9206–9208.
- [201] D. E. Williams. “Alanyl dipeptide potential-derived net atomic charges and bond dipoles, and their variation with molecular conformation”. Em: *Biopolym.* 29 (1990), pp. 1367–1386.
- [202] C. Hensen et al. “A combined QM/MM approach to protein-ligand interactions: polarization effects of the HIV-1 protease on selected high affinity inhibitors”. Em: *J. Med. Chem.* 47 (2004), pp. 6673–6680.
- [203] Junmei Wang et al. “Development of polarizable models for molecular mechanical calculations. 3. Polarizable water models conforming to Thole polarization screening schemes”. Em: *J. Phys. Chem. B.* 116 (2012), pp. 7999–8008. DOI: 10.1021/jp211197f.
- [204] Arieh Warshel, Masahiro Kato e Alexander V. Pisiakov. “Polarizable Force Fields: History, Test Cases, and Prospects”. Em: *J. Chem. Theory Comput.* 3.6 (2007), pp. 2034–2045. DOI: 10.1021/ct700127w. URL: <https://pubs.acs.org/doi/10.1021/ct700127w>.
- [205] Q. Zhang et al. “Deep eutectic solvents: syntheses, properties and applications”. Em: *Chem. Soc. Rev.* 41.21 (2012), pp. 7108–7146. DOI: 10.1039/C2CS35178A.
- [206] Kateryna Goloviznina, Zheng Gong e Agilio A. H. Padua. “The CL&Pol polarizable force field for the simulation of ionic liquids and eutectic solvents”. Em: *WIREs. Comput. Mol. Sci.* 17 (2021). First published on 17 August 2021. DOI: 10.1002/wcms.1572.

- [207] B. L. Bhargava e S. Balasubramanian. "Refined potential model for atomistic simulations of ionic liquid [bmim][PF<sub>6</sub>]" . Em: *Journal of Chemical Physics* 127 (2007), p. 114510. DOI: 10.1063/1.2772253.
- [208] 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.
- [209] Kun Cui, Arun Yethiraj e J. R. Schmidt. "Influence of charge scaling on the solvation properties of ionic liquid solutions". Em: *J. Phys. Chem. B.* 123 (2019), pp. 9222–9229. DOI: 10.1021/acs.jpcc.9b07046.
- [210] B. R. Brooks et al. "CHARMM: The Biomolecular Simulation Program". Em: *J. Comput. Chem.* 30.10 (2009), pp. 1545–1614. ISSN: 1096-987X. DOI: 10.1002/jcc.21287.CHARM.
- [211] José N. Canongia Lopes e Agílio A. H. Pádua. "CL&P: a generic and systematic force field for ionic liquids modeling". Em: *Theor. Chem. Acc.* 131 (2012), p. 1129. DOI: 10.1007/s00214-012-1129-y.
- [212] Ksenia Goloviznina et al. "Transferable, polarizable force field for ionic liquids". Em: *J. Chem. Theory Comput.* 15 (2019), pp. 5858–5871. DOI: 10.1021/acs.jctc.9b00783.
- [213] José N. Canongia Lopes, Julien Deschamps e Agílio A. H. Pádua. "Modeling ionic liquids using a systematic all-atom force field". Em: *J. Phys. Chem. B.* 108 (2004), pp. 2038–47. DOI: 10.1021/jp036506h.
- [214] Marcus D. Hanwell et al. "Avogadro: An advanced semantic chemical editor, visualization, and analysis platform". Em: *J. Cheminformatics* 4 (2012), p. 17. DOI: 10.1186/1758-2946-4-17.
- [215] M. J. Frisch et al. *Gaussian09 Revision E.01*. Gaussian Inc. Wallingford CT 2009.
- [216] Julian Tirado-Rives e William L. Jorgensen. "Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules". Em: *J. Chem. Theory Comput.* 4.2 (2008). Publication Date: January 24, 2008, pp. 297–306. DOI: 10.1021/ct700248k.
- [217] Agilio Padua. *fftool v1.0.0*. Versão v1.0.0. 2015. DOI: 10.5281/zenodo.18618. URL: <https://doi.org/10.5281/zenodo.18618>.

- [218] Justin A. Lemkul et al. "An empirical polarizable force field based on the classical Drude oscillator model: development history and recent applications". Em: *Chem Rev* 116 (2016), pp. 4983–5013. DOI: 10.1021/acs.chemrev.5b00594.
- [219] Ksenia Goloviznina et al. "Extension of the CL&Pol polarizable force field to electrolytes, protic ionic liquids, and deep eutectic solvents". Em: *J. Chem. Theory Comput.* 17 (2021), pp. 1606–1617. DOI: 10.1021/acs.jctc.0c01115.
- [220] Guillaume Lamoureux e Benoît Roux. "Modeling induced polarization with classical Drude oscillators: theory and molecular dynamics simulation algorithm". Em: *J. Chem. Phys.* 119 (2003), pp. 3025–39. DOI: 10.1063/1.1592264.
- [221] Chang-Yun Son et al. "Proper thermal equilibration of simulations with Drude polarizable models: temperature-grouped dual-Nosé–Hoover thermostat". Em: *J. Phys. Chem. Lett.* 10 (2019), pp. 7523–7530. DOI: 10.1021/acs.jpcllett.9b03136.
- [222] Paul Drude. *The Theory of Optics*. Mineola, NY: Dover Publications, 1959.
- [223] Guillaume Lamoureux, Alexander D. MacKerell e Benoît Roux. "A simple polarizable model of water based on classical Drude oscillators". Em: *J. Chem. Phys.* 119 (2003), pp. 5185–97. DOI: 10.1063/1.1591819.
- [224] K. T. Tang e J. Peter Toennies. "An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients". Em: *J. Chem. Phys.* 80 (1984), pp. 3726–41. DOI: 10.1063/1.447689.
- [225] Bogumil Jeziorski, Robert Moszynski e Krzysztof Szalewicz. "Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes". Em: *Chem. Rev.* 94 (1994), pp. 1887–1930. DOI: 10.1021/cr00031a008.
- [226] Jing Huang et al. "Molecular dynamics simulations using the Drude polarizable force field on GPUs with OpenMM: implementation, validation, and benchmarks". Em: *J. Comput. Chem.* 39 (2018), pp. 1682–9. DOI: 10.1002/jcc.25365.
- [227] J.-P. Ryckaert, G. Ciccotti e H. J. C. Berendsen. "Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes". Em: *J. Comput. Phys.* 23.3 (1977), pp. 327–341. DOI: 10.1016/0021-9991(77)90098-5.

- [228] G. Ciccotti e J. P. Ryckaert. "Molecular Dynamics Simulation of Rigid Molecules". Em: *Comput. Phys. Rep.* 4.6 (1986), pp. 345–392. DOI: 10.1016/0167-7977(86)90022-5.
- [229] Liliana A. Rodrigues et al. "Deep eutectic systems from betaine and polyols – Physicochemical and toxicological properties". Em: *J. Mol. Liq.* 335 (2021), p. 116201. DOI: 10.1016/j.molliq.2021.116201.