



Rayla Kelly Magalhães Costa

A reconciliação entre o coeficiente de partição octanol-água experimental e calculado do 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina usando dinâmica molecular atomística: uma questão em aberto

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Química, do Departamento de Química da PUC-Rio.

Orientador: Prof. Dr. André Silva Pimentel

Rio de Janeiro

Março de 2023



Rayla Kelly Magalhães Costa

A reconciliação entre o coeficiente de partição octanol-água experimental e calculado do 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina usando dinâmica molecular atomística: uma questão em aberto

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Química da PUC-Rio. Aprovada pela Comissão Examinadora abaixo:

Prof. Dr. André Silva Pimentel

Orientador

Departamento de Química da PUC-Rio

Profa. Dr^a Thereza Amélia Soares da Silva

Universidade de São Paulo (USP- Ribeirão Preto)

Profa. Dr^a Iolanda Midea Cuccovia

Universidade de São Paulo (USP)

Rio de Janeiro, 1 de Março de 2023

Todos os direitos reservados. A reprodução, total ou parcial do trabalho, é proibida sem a autorização da universidade, da autora e do orientador.

Rayla Kelly Magalhães Costa

Graduou-se em Química pela Universidade Estadual do Piauí (UESPI) em 2021. Atua na área de físico-química com ênfase em Química computacional.

Ficha Catalográfica

Costa, Rayla Kelly Magalhães

A reconciliação entre o coeficiente de partição octanol-água experimental e calculado do 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina usando dinâmica molecular atomística : uma questão em aberto / Rayla Kelly Magalhães Costa ; orientador: André Silva Pimentel. – 2023.

124 f. : il. color. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Química, 2023.

Inclui bibliografia

1. Química – Teses. 2. Métodos alquímicos. 3. Amostragem guarda-chuva. 4. Cálculos de energia livre. 5. Coeficientes de partição. 6. Moléculas anfifílicas. I. Pimentel, André Silva. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Química. III. Título.

CDD: 540

Dedico esta dissertação aos meus pais Josefa Campelo e Miguel Costa pelo empenho em minha educação, ao meu orientar Dr. André Pimentel, por todos os ensinamentos e a todos os que me ajudaram ao longo desta caminhada.

Agradecimentos

Primeiramente a Deus pelo dom da vida, pelo seu amor que me manteve forte e com saúde, persistente em meus objetivos, conduzindo-me nesta jornada com muita fé. Sem Ele nada seria.

Aos meus pais Miguel Costa e Josefa Magalhães pelo amor estratosférico, pelo apoio e por entenderem meus momentos de ausência. Ao meu irmão Raimundo José por todo o incentivo e por acreditar que tudo isso seria possível. Aos meus afilhados Victor Miguel e Maria Rita, vocês são o motivo do meu empenho. Tudo que eu faço é para orgulhar vocês!

Ao meu namorado Michel Alencar, meu companheiro que me apoiou, me deu forças e me acalentou em momentos de desabafos, sempre me dizendo que, para se conseguir algo, precisamos de muito empenho e dedicação. Obrigada por todo apoio, por toda motivação, por toda força, amo você!

Ao meu orientador Dr. André Silva Pimentel por ter me acolhido em seu Laboratório e pela confiança depositada em mim para a realização deste trabalho, e por todos os ensinamentos e oportunidades. Sempre gentil, ao fim de cada orientação, deixava-me mais confiante para os próximos caminhos.

Aos colegas de laboratório Lucas Miguel, Felipe Souza e Rudielson Santos por toda ajuda e troca de conhecimento.

Ao meu grupo de amigas (Alana, Belisa, Camila, Yuri, Lisa, Luciana, Nathália, Nislene, Wellivanda e Clara) que mesmo distante tornam meus dias mais divertidos. Vocês são incríveis!

À PUC-Rio e o Departamento de Química da PUC-Rio, pelo acolhimento e seus préstimos que me foram de muita valia.

A CAPES e PUC-Rio pelo auxílio concedido para a realização deste trabalho, sem os quais este trabalho não poderia ter sido realizado.

A cidade do Rio de Janeiro que é maravilhosa.

A meu orientador de iniciação científica Dr. Francisco das Chagas Alves Lima por todos os ensinamentos e pelo apoio que me deu quando decidi realizar o mestrado em outra universidade e estado.

A meus amigos de graduação Alexandre Diógenes e Jaqueline Fernanda por todo conhecimento trocado e apoio.

A todos os professores que tive ao longo da minha jornada que sempre estiveram dispostos a ajudar e contribuir para um melhor aprendizado.

Enfim, gratidão a todas as pessoas que contribuíram, direta ou indiretamente, para a concretização desse sonho.

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

Resumo

Costa, Rayla Kelly Magalhães; Pimentel, André Silva. **A reconciliação entre o coeficiente de partição octanol-água experimental e calculado do 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina usando dinâmica molecular atomística: uma questão em aberto.** Rio de Janeiro, 2023. 124p. Dissertação de Mestrado – Departamento de Química, Pontifícia Universidade Católica do Rio de Janeiro.

O coeficiente de partição octanol-água do composto 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina (DPPC) foi investigado utilizando os métodos de integração termodinâmica e amostragem guarda-chuva através de simulações de dinâmica molecular atomística. Os campos de força AMBER/GAFF e CHARMM/CGenFF foram usados com seis modelos de água (SPC, TIP3P, TIP4P, TIP5P, OPC3 e OPC4) amplamente utilizados em simulações de dinâmica molecular. Dentre os modelos utilizados, o modelo de água OPC4 com os dois campos de força em estudo forneceu a melhor concordância com o coeficiente de partição experimental octanol-água do DPPC. No entanto, ainda existe muito espaço para melhorias nos modelos de água que estimam a tensão superficial de forma apropriada. Usando o modelo de água OPC4, a energia livre de Gibbs de transferência do DPPC do octanol para a fase aquosa foi calculada em $19,8 \pm 0,3$ e $20,2 \pm 0,3$ kcal mol⁻¹, estimando um coeficiente de partição octanol-água de $14,5 \pm 0,4$ e $14,8 \pm 0,3$ para os campos de força AMBER/GAFF e CHARMM/CGenFF, respectivamente. A amostragem guarda-chuva apresentou problemas de arrastes de moléculas de uma fase para outra, gerando artefatos e consequentemente subestimando os valores de energia livre e de coeficiente de partição octanol-água. Este estudo mostra a importância do desenvolvimento de novos modelos de água que reproduzam com precisão todas as suas características experimentais. A conciliação entre medições experimentais e cálculos teóricos do coeficiente de partição de moléculas anfifílicas poderia ser

resolvida através do ajuste dos parâmetros do modelo de água. Este estudo possui grande importância na simulação de propriedades moleculares de importância em muitas áreas de aplicações científicas e industriais, tais como biofísica, surfactante, coloides, membranas, medicina, nanotecnologia, e indústrias alimentícias e farmacêuticas.

Palavras-chaves:

Métodos Alquímicos; Amostragem Guarda-chuva; Cálculos de Energia Livre; Coeficientes de Partição; Moléculas Anfifílicas.

Abstract

Costa, Rayla Kelly Magalhães; Pimentel, André Silva (Advisor). **The reconciliation between the experimental and calculated octanol-water partition coefficient of 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine using atomistic molecular dynamics: an open question.** Rio de Janeiro, 2023. 124p. Masters Dissertation – Department of Chemistry, Pontifical Catholic University of Rio de Janeiro.

The octanol-water partition coefficient of the compound 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine (DPPC) molecule was investigated using the methods of thermodynamic integration and umbrella sampling through atomistic molecular dynamics simulations. The AMBER/GAFF and CHARMM/CGenFF force fields were used with six water models (SPC, TIP3P, TIP4P, TIP5P, OPC3, and OPC4) widely used in molecular dynamics simulations. Among the models used, the OPC4 water model with the two force fields provided the best agreement with the experimental octanol-water partition coefficient of the DPPC. However, there is still much room for improvement in water models that correctly estimate the surface tension. Using the OPC4 water model, the Gibbs free energy of transferring DPPC from octanol to the aqueous phase was calculated to be 19.8 ± 0.3 and 20.2 ± 0.3 kcal mol⁻¹, estimating an octanol-water partition coefficient of 14.5 ± 0.4 and 14.8 ± 0.3 for the AMBER/GAFF and CHARMM/CGenFF force fields, respectively. Umbrella sampling presented issues of molecules being dragged between the two phases, generating artifacts, and consequently underestimating the values of free energy and octanol-water partition coefficient. This study shows the importance of developing new models of water that accurately reproduce all its experimental characteristics. The reconciliation between experimental measurements and theoretical calculations of partition coefficients of amphiphilic molecules. This study may have great importance in many areas of scientific and industrial applications, such as

biophysics, surfactant, colloids, membranes, medicine, nanotechnology, and food and pharmaceutical industries.

Keywords:

Alchemical Methods; Umbrella Sampling; Free Energy Calculations; Partition Coefficients; Amphiphilic Molecules.

Sumário

1. Introdução.....	19
2. Objetivos.....	22
2.1. Objetivo Geral	22
2.2. Objetivos específicos.....	22
3. Revisão bibliográfica	23
3.1 Fosfolipídios	23
3.2. 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina (DPPC)	25
3.3. Coeficiente de partição octanol água (<i>logK_{ow}</i>)	27
3.3.1. Métodos experimentais para calcular o <i>logK_{ow}</i> de surfactantes.....	28
3.3.2. Métodos teóricos para estimar o <i>logK_{ow}</i> de surfactantes	30
3.4. Simulação molecular	31
3.4.1. Dinâmica molecular clássica.....	31
3.4.2. Ensembles	32
3.4.3 Campos de força.....	34
3.4.3.1 CHARMM/CGenFF	35
3.4.3.2 AMBER/GAFF.....	37
3.4.4 Cálculo do potencial	38
3.4.4.1 Comprimento, ângulo e torção de Ligação	39
3.4.4.2 Interações de van der Waals	40
3.4.4.3 Interações Eletrostáticas ou de Coulomb	41
3.4.5 Algoritmos na dinâmica molecular.....	42
3.4.5.1 Algoritmo de Verlet.....	43
3.4.5.2 Algoritmo de minimização	44
3.4.6 Condições Periódicas de Contorno	45
3.4.7 Termostatos	45
3.4.7.1 Termostato Berendsen.....	46
3.4.7.2 Termostato de reescalonamento da velocidade	46
3.4.8 Barostatos	47
3.4.8.1. Barostato de Berendsen.....	47
3.4.8.2. Barostato Parrinello–Rahman	47
3.4.10. Algoritmo de restrição: LINCS	48
3.4.11. Soma de Ewald (PME)	49

3.4.12. Modelos de água.....	50
3.4.13. Energias Livres (ΔG).....	52
3.4.14. Integração termodinâmica (TI).....	53
3.4.14.1. Análise alquímica.....	55
3.4.15. Amostragem guarda-Chuva	58
3.4.15.1. Método da análise do histograma ponderado (WHAM)	60
3.4.16. Gromacs	61
4. Metodologia.....	62
4.1. Análises de quimio(bio)informática	62
4.2. Dinâmica molecular.....	62
4.2.1. Cálculo de energia livre Gibbs usando integração termodinâmica (TI)	63
4.3. Simulação de tendência	66
4.3.1. Cálculo de energia livre de Gibbs usando amostragem guarda-chuva	68
5. Resultados e discussões.....	69
5.1. Ferramentas de quimio(bio)informática.....	69
5.2. Integração termodinâmica.....	70
5.2.1. Estimador MBAR	71
5.2.2. Estimador TI	77
5.3. Amostragem guarda-chuva	80
6. Implicações	86
7. Conclusões.....	87
8. Perspectivas Futuras.....	89
9. Referências Bibliográficas	90
10. Anexos	105
ANEXO A.....	105
ANEXO B	107
ANEXO C	109
ANEXO D.....	110
ANEXO E	111
ANEXO F	112
ANEXO G.....	124

Lista de figuras

Figura 01 - Representação esquemática dos fosfolipídios. Os ácidos graxos nas caudas podem ser saturados (S) ou insaturados (I).

Figura 02 - (A) Ilustração de uma micela formada por fosfolipídios. (B) Ilustração de uma bicamada lipídica composta por fosfolipídios.

Figura 03 - Estrutura em 2D e 3D da molécula de DPPC.

Figura 04 - Ilustração do ciclo termodinâmico utilizado no sistema em estudo. A energia livre de Gibbs (ΔG) é calculada contabilizando a solvatação do DPPC pela água (ΔG_w) e a solvatação do DPPC pelo octanol (ΔG_o). As siglas "elec" e "vdW" significam contribuições eletrostáticas e de van der Waals, respectivamente. A representação gráfica do ciclo termodinâmico não físico mostra a molécula "transparente" que não está interagindo com o ambiente (ou seja, está desacoplada). O fundo azul é um lembrete de que a água está presente e o fundo laranja significa que o octanol está presente. O fundo cinza representa o vácuo.

Figura 05 - (A) Sistema inicial de tendência direta onde o DPPC está na fase octanol (em laranja), e (B) sistema final de tendência direta onde o DPPC está na fase aquosa (em azul).

Figura 06 - (A) Sistema inicial de tendência inversa onde o DPPC está na fase aquosa (em azul), e (B) sistema final de tendência inversa onde o DPPC está na fase octanol (em laranja).

Figura 07 - Gráfico do perfil de energia livre de Gibbs da transferência (ΔG_{ow}) de DPPC do octanol para a fase aquosa usando amostragem guarda-chuva do sistema direto, que corresponde a uma caixa de água entre duas caixas de octanol.

Figura 08 - Arraste de moléculas de octanol interagindo fortemente com a cauda hidrofóbica do DPPC no sistema direto. Neste sistema o DPPC é transferido da fase octanol (em laranja) para a fase aquosa (em azul).

Figura 09 - Histograma de janelas de amostragem de guarda-chuva se sobrepõem na coordenada de reação ξ para o sistema direto, que corresponde a uma caixa de água entre duas caixas de octanol. Neste sistema o DPPC é transferido do octanol para a água.

Figura 10 - Gráfico do perfil de energia livre de Gibbs da transferência (ΔG_{ow}) de DPPC da fase aquosa para a fase do octanol usando amostragem guarda-chuva do sistema inverso, que corresponde a uma caixa de octanol entre duas caixas de água.

Figura 11 - Arraste de moléculas de água interagindo fortemente com a cabeça hidrofílica do DPPC no sistema inverso. Neste sistema o DPPC é transferido da fase aquosa (em azul) para a fase octanol (em laranja).

Figura 12 - Histograma de janelas de amostragem de guarda-chuva se sobrepõem na coordenada de reação ξ para o sistema inverso, que corresponde a uma caixa de octanol entre duas caixas de água. Neste sistema o DPPC é transferido da fase aquosa para a fase octanol.

Lista de Tabelas

Tabela 01 - Os coeficientes de partição octanol-água $\log K_{ow}$ previstos por ferramentas de quimioinformática em comparação com o valor experimental obtido na literatura (SOUZA et al., 2020a).

Tabela 02 – Energia livre de Gibbs de transferência (ΔG_{ow}) do DPPC da fase de octanol (puro e úmido) para a fase de água e coeficiente de partição octanol-água ($\log K_{ow}$) utilizando a integração termodinâmica (estimador MBAR) com campos de força AMBER/GAFF e CHARMM/CGenFF.

Tabela 03 - Energia livre de Gibbs de transferência (ΔG_{ow}) do DPPC da fase de octanol (puro e úmido) para a fase de água e coeficiente de partição octanol-água ($\log K_{ow}$) utilizando a integração termodinâmica (estimador TI) com campos de força AMBER/GAFF e CHARMM/CGenFF.

Tabela 04 - Energia livre de Gibbs de transferência ΔG_{ow} e coeficiente de partição octanol-água ($\log K_{ow}$) utilizando a amostragem guarda-chuva com campo de força CHARMM/CGenFF.

Lista de abreviações

DPPC - 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina

AMBER - Assisted Model Building and Energy Refinement

CHARMM - Chemistry at HARvard Macromolecular Mechanics

GAFF – General Amber Force Field

CGenFF – CHARMM General Force Field

$\log K_{ow}$ - Coeficiente de partição octanol-água

ADMET - Absorção, Distribuição, Metabolismo, Excreção e Toxicidade

ΔG_{ow} - Energia livre de Gibbs de transferência

ΔG_w - Energia livre de solvatação da água

ΔG_o - Energia livres de solvatação do octanol

US - Amostragem guarda-chuva

TI - Integração Termodinâmica

SPC - Carga de Ponto Único

TIP3P - Potencial intermolecular transferível com três pontos

TIP4P - Potencial intermolecular transferível com quatro pontos

TIP5P - Potencial intermolecular transferível com cinco pontos

OPC3 - Carga de ponto ótimo com três pontos

OPC4 - Carga de ponto ótimo com quatro pontos

SP-A - Proteína Surfactante A

SP-B - Proteína Surfactante B

SP-C - Proteína Surfactante C

SP-D - Proteína Surfactante D

λ - Janela de amostragem (Lambda)

dU - Variação de energia interna

MBAR - Multistate Bennett Acceptance Ratio

ξ - Coordenada de reação

WHAM - Método de análise de histograma ponderado

CLAE - Cromatografia Líquida de Alta Eficiência

FR - Fase Reversa

CMC – Concentração Micelar Crítica

QSPR - Quantitative Structure-Property Relationship

QSAR - Quantitative Structure-Activity Relationship

DM - Dinâmica Molecular

AA - All-Atom

UA - United-Atom

CG - Coarse Grained

V-Rescale - Termostato de reescalonamento da velocidade

LINCS - LINear Constraint Solver

FEP - Método de perturbação da energia livre

vdW - van der Waals

PMF - Potencial de força média

Devemos acreditar que somos dotados de alguma coisa e que essa coisa deve ser alcançada a qualquer custo.

Marie Curie

1. Introdução

Os fosfolipídios são os principais constituintes das membranas biológicas e possuem papéis importantes para as propriedades estruturais e funcionais das membranas celulares (URAN et al., 2001). O 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina (DPPC) é um dos mais importantes componentes lipídicos zwitteriônicos, que pode ser encontrado em todos os organismos vivos e compreendem cerca de 40-50% do total de fosfolipídios celulares (LEEKUMJORN; SUM, 2006; URAN et al., 2001). O DPPC atua como parte vital na fusão e transporte de membrana, na endocitose e na atividade catalítica de enzimas, diminui os níveis de lipídios e colesterol no sangue prevenindo doenças cardiovasculares e no transporte e absorção de gorduras ajudando na inibição da obesidade (ZHOU et al., 2016). Além disto, os surfactantes pulmonares consistem em uma mistura complexa de proteínas e fosfolipídios, sendo o DPPC o mais abundante já que corresponde a 45% da massa dos surfactantes pulmonares (BAER et al., 2019). O DPPC tem como função reduzir a tensão superficial alveolar para evitar o colapso durante a expiração (SOUZA et al., 2020a).

A estrutura do DPPC é composta por uma cabeça polar, contendo um grupo fosfato e um grupo colina, e por duas caudas derivadas de ácidos graxos, unidas por uma molécula de glicerol (SOUZA et al., 2020a). Devido a sua organização estrutural, o DPPC possui característica hidrofóbica, sendo considerado um composto anfifílico pois apresenta na mesma molécula tanto parte hidrofílica quanto hidrofóbica (COSTA et al., 2023). Além disso, apresenta a capacidade de modificar a tensão de superficial na interface e influencia na superfície de contato onde estão inseridos (PARRA; PÉREZ-GIL, 2015).

DPPC é largamente utilizado tanto em produtos industriais quanto de consumo. Dessa forma, o destino ambiental e a biodisponibilidade destes compostos tornam-se muito relevante (GARRIDO et al., 2009). O coeficiente de partição octanol-água ($\log K_{ow}$) é um parâmetro chave para

avaliar as propriedades mencionadas anteriormente (CHENG et al., 2007). Existem diversos métodos experimentais capazes de determinar $\log K_{ow}$, mas o método de agitação lenta é o mais adequado para surfactantes (FINIZIO; VIGHI; SANDRONI, 1997; MILLER; MCWILLIAMS, 2010; SHORT et al., 2010). No entanto, é um desafio técnico determinar o $\log K_{ow}$ para moléculas anfifílicas. Devido as suas propriedades, os surfactantes formam agregados na solução e podem emulsionar na mistura octanol-água (SHORT et al., 2010).

Além dos métodos experimentais, existem as ferramentas de quimio(bio)informática que são capazes de prever características relacionadas a absorção, distribuição, metabolismo, excreção e toxicidade (ADMET) com base em banco de dados de moléculas (FERREIRA; ANDRICOPULO, 2019; VAN DE WATERBEEMD; GIFFORD, 2003). Geralmente, elas são baseadas em modelos de aprendizado de máquina de grandes conjuntos de dados de pequenas moléculas, podendo não representar com precisão os dados relacionados ao $\log K_{ow}$ de moléculas anfifílicas (SARKAR; KELLOGG, 2010).

A termodinâmica pode ser utilizada para realizar cálculos de $\log K_{ow}$ utilizando as energias livres de hidratação de uma substância na fase octanol e na fase aquosa sem enfrentar tantos desafios quando comparados aos estudos experimentais (FORNASIER et al., 2020; SOUZA et al., 2020a; COSTA et al., 2023). Pode-se calcular o coeficiente de partição correspondente, de acordo com a seguinte equação:

$$\log K_{ow} = \frac{-\Delta G_{ow}}{2.303RT}$$

Eq. 01

onde ΔG_{ow} é a energia livre de transferência da molécula em estudo da fase octanol para a fase aquosa, a temperatura é geralmente medida em 298 K sob pressão de 1 bar e R é a constante universal dos gases ($8.314 \text{ JK}^{-1} \text{ mol}^{-1}$). Dentre as diversas formas de calcular $\log K_{ow}$ através da termodinâmica, este trabalho utiliza dinâmica molecular atomística através de cálculos de amostragem guarda-chuva (US) (TORRIE; VALLEAU, 1977) e integração termodinâmica (TI) (BOYCE et al., 2009).

Recentemente estudos realizados por *Temptra et al* demonstram que as simulações de monocamadas de DPPC usando dinâmica molecular dependem da tensão superficial do modelo de água utilizado e também do método de cálculo das interações de longo alcance (TEMPRA; OLLILA; JAVANAINEN, 2022). Com base neste estudo, os cálculos de $\log K_{ow}$ para as moléculas anfifílicas dependem dessa mesma metodologia. Estudos realizados por *Fan et al* apresentam uma metodologia de cálculo do $\log K_{ow}$ de pequenas moléculas que estão presentes no desafio SAMPL6 utilizando dinâmica molecular atomística e métodos de amostragem alquímica. Observa-se que a maioria das moléculas utilizadas são hidrofóbicas. Dessa forma, essa metodologia pode ser uma alternativa para ser aplicada a molécula de DPPC (FAN; IORGA; BECKSTEIN, 2020; IŞIK et al., 2020a, 2020b; IZADI; ANANDAKRISHNAN; ONUFRIEV, 2014).

Este trabalho busca determinar o $\log K_{ow}$ do DPPC e comparar os dados encontrados com o único valor experimental para a molécula de DPPC existente na literatura. O valor experimental do $\log K_{ow}$ foi obtido utilizando o método de agitação lenta realizado em nosso laboratório e publicado recentemente (SOUZA et al., 2020a). Faz-se necessário utilizar diversos modelos de água existentes já que o comportamento da molécula em estudo depende dos parâmetros de cada modelo. Neste estudo, os campos de força AMBER/GAFF (DUAN et al., 2003) e CHARMM/CGenFF (HUANG; MACKERELL, 2013) são utilizados com os seguintes modelos de água: carga de ponto único (SPC) (BERENDSEN; GRIGERA; STRAATSMA, 1987), potencial intermolecular transferível com três pontos (TIP3P) (PRICE; BROOKS, 2004), quatro pontos (TIP4P) (JORGENSEN et al., 1983) e cinco pontos (TIP5P) (KHALAK; BAUMEIER; KARTTUNEN, 2018), cargas de ponto ótimo com três pontos (OPC3) e quatro pontos (OPC4) (IZADI; ANANDAKRISHNAN; ONUFRIEV, 2014; JAVANAINEN et al., 2018; TEMPRA; OLLILA; JAVANAINEN, 2022). A importância deste estudo está principalmente na tentativa de conciliação dos valores experimentais com os valores encontrados por métodos computacionais, e apontar possíveis origens do porquê da falta de reconciliação. Como este parâmetro é difícil de medir

empiricamente, é importante que se tenham métodos computacionais que possam medir com precisão esta propriedade.

2. Objetivos

2.1. Objetivo Geral

O objetivo deste trabalho é calcular o coeficiente de partição octanol-água ($\log K_{ow}$) do fosfolípido 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina (DPPC) utilizando dinâmica molecular atomística através da energia livre de Gibbs de transferência entre as fases aquosa e octanol (ΔG_{ow}). Para isto, utilizaram-se métodos alquímicos com a integração termodinâmica e métodos de amostragem guarda-chuva.

2.2. Objetivos específicos

- Verificar a influência dos diferentes modelos de água nos cálculos de energia livres de Gibbs de transferência entre as fases aquosa e octanol.
- Identificar as diferenças nos cálculos de energia livre entre os resultados obtidos utilizando os campos de força AMBER/GAFF e CHARMM/CGenFF.
- Avaliar o cálculo da energia livre de Gibbs de transferência entre as fases aquosa e octanol utilizando dois métodos distintos: Método alquímico com a integração termodinâmica e o método da amostragem guarda-chuva. Desta forma, pode-se determinar o melhor método para o sistema em estudo.
- Comparar os resultados encontrados com o único valor experimental disponível na literatura.

3. Revisão bibliográfica

3.1 Fosfolipídios

Os fosfolipídios são os principais componentes lipídicos estruturais das membranas celulares. Trata-se da estrutura que media o que pode entrar e sair das nossas células (WODTKE, 1978). A estrutura destas moléculas consiste de duas "caudas hidrofóbicas" de ácidos graxos, ligadas por meio de um glicerol a uma "cabeça hidrofílica" de um grupo fosfato e outros grupos carregados ou polares (PICHOT; WATSON; NORTON, 2013). Geralmente, o grupo fosfato pode ser modificado com outras moléculas orgânicas simples. A Figura 01 mostra a sua estrutura genérica.

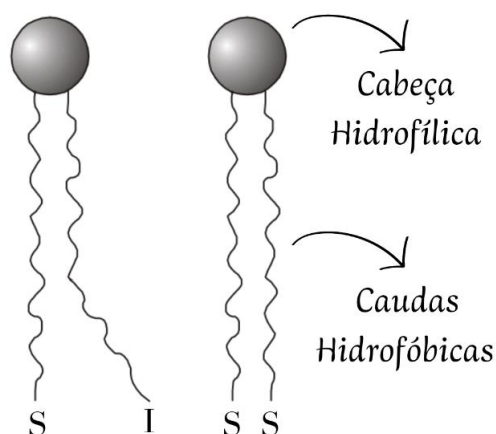


Figura 01 - Representação esquemática dos fosfolipídios. Os ácidos graxos nas caudas podem ser saturados (S) ou insaturados (I).

Vários fosfolipídios são utilizados como agentes emulsificantes, promovendo a dispersão coloidal de um líquido em outro, e como agentes tensoativos, reduzindo a tensão superficial de uma solução (AGASSANDIAN; MALLAMPALLI, 2013). Os fosfolipídios podem exercer essas funções pois são moléculas anfifílicas (DEAMER, 2017).

As propriedades dos fosfolipídios e de outras moléculas lipídicas anfifílicas são importantes para a formação das estruturas das membranas biológicas (DOWHAN, 1997; KHANDWALA; KASPER, 1971). Uma vez que suspensos em água, eles formam espontaneamente agregados esféricos ou cilíndricos. Neste caso, as caudas hidrofóbicas

ficam voltadas para o interior evitando o contato com a água, e as cabeças hidrofílicas ficam em contato com a água, como mostrado na Figura 02(A) (HWANG; RAMSEY; KABANOV, 2020; MENDER, 1979). Em concentrações apropriadas eles formam espontaneamente bicamadas lipídicas, que se rearranjam em estruturas ordenadas, de forma que as caudas hidrofóbicas são orientadas no interior para excluir a água. Simultaneamente, os grupos das cabeças polares são orientados para a água, como mostrado na Figura 02(B), estabelecendo assim a estrutura básica das membranas (LIPOWSKY, 1995; VEATCH; KELLER, 2002).

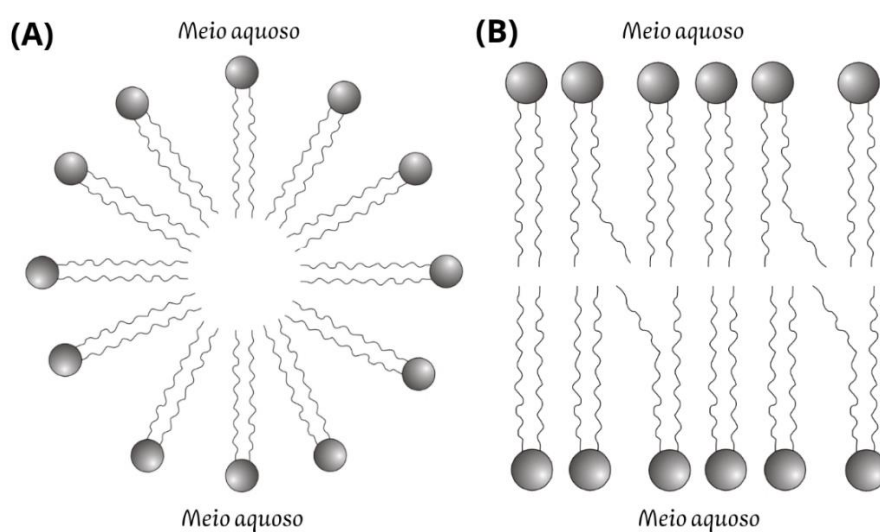


Figura 02 - (A) Ilustração de uma micela formada por fosfolípidios. **(B)** Ilustração de uma bicamada lipídica composta por fosfolípidios.

Além disso, os fosfolípidios compõem os surfactantes pulmonares, cuja função é reduzir de forma significativa a tensão superficial dentro do alvéolo pulmonar. É composto por uma mistura de ~80-85% de fosfolípidios, 5-10% de lípidios neutros e 8-10% de proteínas específicas. Eles formam uma camada entre o líquido aquoso das vias aéreas e o ar inspirado no interior dos pulmões (CHOI; JANG; PARK, 2020). Essas misturas são montadas, transportadas, secretadas e recicladas no espaço alveolar. São moléculas dinâmicas, por se tratar de uma superfície exposta à constante dinâmica de compressão-expansão do ciclo respiratório (NKADI; MERRITT; PILLERS, 2009). O funcionamento anormal do surfactante acontece devido à oxidação, degradação proteolítica e inibição levando ao desconforto respiratório com morbidade

e mortalidade concomitantes (LOPEZ-RODRIGUEZ et al., 2017). O surfactante pulmonar é absolutamente necessário, e a sua ausência, deficiência ou inativação está associada a doenças pulmonares graves. (BLANCO; PÉREZ-GIL, 2007) Além de contribuir para o aumento significativo das respostas imunes alérgicas (BRANDT et al., 2008).

Algumas das aplicações terapêuticas de surfactante pulmonares exógenos é na síndrome do desconforto respiratório em prematuros, levando à redução da mortalidade (HENTSCHEL et al., 2020), em lesão pulmonar aguda e na síndrome do desconforto respiratório agudo (WILLSON; NOTTER, 2011).

3.2. 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina (DPPC)

O 1,2-dipalmitoil-sn-glicero-3-fosfatidilcolina (DPPC) pode ser encontrado em todos os organismos vivos. São predominantemente encontrados nas membranas celulares de animais, e está envolvida em uma série de processos que ajudam a manter o funcionamento do organismo (URAN et al., 2001). Além de ser o componente lipídico predominante nos surfactantes pulmonares (representa aproximadamente 40% da massa total do surfactante) (BLANCO; PÉREZ-GIL, 2007), o DPPC sozinho tem as propriedades de reduzir a tensão superficial alveolar. Porém, necessita das proteínas de surfactante específicas (SP-A, SP-B, SP-C e SP-D) e outros lipídios para facilitar sua absorção na interface ar-líquido (PARRA; PÉREZ-GIL, 2015; SOUZA et al., 2020a).

Possuem diversas funções no organismo: atua mantendo a integridade das membranas celulares (LI et al., 2006), protege e auxilia na recuperação das células do fígado quando expostas à ingestão excessiva de álcool ou ingestão de substâncias tóxicas e de intoxicação medicamentosa (ALICE; MELLO, 2009; AMORIM et al., 2011), exerce função no transporte de gordura de uma célula para outra, mantendo as estruturas celulares saudáveis (SOUZA et al., 2020a), desempenha papel importante em funções cognitivas, como aprendizagem e memória (ZHOU et al., 2016), e também auxilia na diminuição da presença de

homocisteína no sangue, diminuindo o risco de doenças cardiovasculares (PAOLETTI et al., 2011).

A sua estrutura possui uma cabeça polar (hidrofílica) com um grupo de fosfato, um grupo de colina e duas caudas apolares (hidrofóbicas) derivadas de ácidos graxos, unidas por uma molécula de glicerol, como mostra a Figura 03.

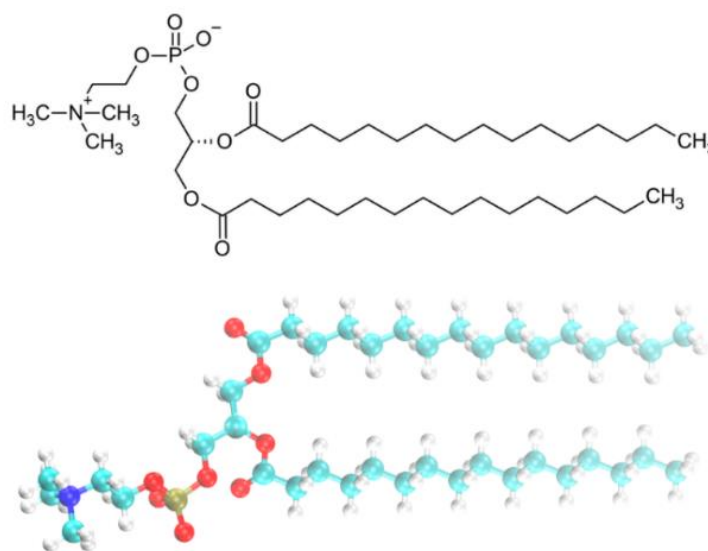


Figura 03 - Estrutura em 2D e 3D da molécula de DPPC.

Devido a sua estrutura, o DPPC tende a agregar-se, organizando-se em estruturas nanométricas. Estas estruturas permitem maximizar as interações hidrofóbicas entre as cadeias hidrocarbônicas, isolando-se da água, e deixam os grupos polares em contato com o solvente, com o qual podem interagir conforme mostrado na Figura 02. Tais arranjos moleculares constituem o estado de menor energia livre (ALLOUCHE, 2012; RUDOLPHI-SKÓRSKA; FILEK; ZEMBALA, 2017; SOUZA et al., 2020b).

O DPPC é amplamente utilizado em todo o mundo, tanto em produtos industriais quanto de consumo, de modo que as propriedades que determinam sua distribuição no meio ambiente são de particular importância (COSTA et al., 2023).

3.3. Coeficiente de partição octanol água ($\log K_{ow}$)

O coeficiente de partição de um composto entre água e octanol foi introduzido pela primeira vez por Hansch e Fujita em 1964 (GARRIDO et al., 2009). Ao longo dos anos, muitas abordagens e métodos foram desenvolvidos para estimar esta propriedade (HODGES et al., 2019). É uma medida da lipofilicidade de um composto e pode ser definido como a razão da concentração do composto, no equilíbrio, após dissolução em um sistema contendo duas fases imiscíveis, neste caso água e octanol (Equação 02) (FINIZIO; VIGHI; SANDRONI, 1997; RIBEIRO; FERREIRA, 2003).

$$P_{ow} = \frac{C_{octanol}}{C_{água}}$$

Eq. 02

O coeficiente de partição é adimensional e é geralmente dado na forma de seu logaritmo decimal.

Nos últimos anos, houve um crescente número de publicações demonstrando a sua correlação com propriedades físicas, químicas e biológicas dos compostos tais como bioconcentração, solubilidade, toxicidade, absorção e transporte (COSTA et al., 2023; IŞIK et al., 2020b; VILAS-BOAS et al., 2022; WANG et al., 2019). O $\log K_{ow}$ se destaca também por ser um dos parâmetros-chave na avaliação de risco ambiental de produtos químicos, devido ser frequentemente usado para estimar o destino ambiental e a biodisponibilidade (FAN; IORGA; BECKSTEIN, 2020).

Esta propriedade pode ser medida utilizando técnicas experimentais: (I) método do frasco de agitação (HEYMAN et al., 2019); (II) técnica geradora de coluna (GC) (PEREIRA; AQUINO NETO, 2000) e (III) medidas de cromatografia líquida de alta eficiência de fase reversa (CLAE-FR) e de camada fina de fase reversa (TLC-FR) (PEREIRA; AQUINO NETO, 2000). Também pode ser calculada teoricamente utilizando diversas metodologias: (I) métodos baseados em fragmentos da molécula (FABIANA MEIJON FADUL, 2019; SUZUKI; KUDO, 1990); (II) métodos baseados na contribuição atômica (PARAMETERS, 1987) e (III)

métodos baseados em modelos que utilizam parâmetros obtidos por meio de cálculos quânticos (CENSE et al., 1994).

3.3.1. Métodos experimentais para calcular o $\log K_{ow}$ de surfactantes

Atualmente, existem vários métodos experimentais que estão disponíveis para medição ou previsão de $\log K_{ow}$ de surfactantes. (OECD, 1989) Os mais adequados são: o método de agitação lenta (OECD 123) (GUIDELINES, 2022), o método de cromatografia líquida de alta eficiência (CLAE) (OECD 117) (OECD, 1989) e o método de razão de solubilidade (OECD 107) (OECD/OCDE, 1995), que faz o uso da razão da solubilidade química em água e octanol.

O método de agitação lenta é o mais utilizado e adequado para o uso em surfactantes devido a sua principal característica, que busca minimizar a agitação e aumentar a troca entre a água e o octanol evitando a formação de gotículas e espuma. Este método é aplicado na determinação de valores de $\log K_{ow}$ de compostos altamente hidrofóbicos variando de 5,0 até 8,2 (GUIDELINES, 2022). É importante ressaltar que este método deve ser aplicado abaixo da concentração micelar crítica (CMC) para evitar a formação de agregados (HODGES et al., 2019).

O método CLAE é realizado em colunas analíticas empacotadas com sílica quimicamente modificada com cadeias longas de hidrocarbonetos (por exemplo, C8 ou C18). Este método permite que os coeficientes de partição sejam estimados na faixa entre 0 e 6. Não é aplicável a ácidos e bases fortes, complexos metálicos, substâncias que reagem com o eluente ou surfactantes ionizáveis (OECD, 1989).

O método da razão das solubilidades é calculado através do logaritmo da razão entre as solubilidades do octanol e a solubilidade da água, que são determinadas experimentalmente. No entanto, a solubilidade do surfactante em água não é propriamente definida nem fácil de medir (HODGES et al., 2019; OECD/OCDE, 1995).

Existe uma preocupação de que esses métodos disponíveis não tenham sido totalmente validados para o uso em surfactantes e possam não ser aplicáveis devido ao comportamento dos surfactantes. Dessa

forma, determinar os dados de $\log K_{ow}$ para surfactantes é um desafio devido às suas propriedades (FAN; IORGA; BECKSTEIN, 2020; FINIZIO; VIGHI; SANDRONI, 1997; MILLER; MCWILLIAMS, 2010; SHORT et al., 2010).

A necessidade de medir um coeficiente de partição para surfactantes com precisão é de importância ambiental significativa (BANAT et al., 2010; OLIVEIRA et al., 2005; OLIVEIRA; MEI, 2009; RAIGER; LÓPEZ, 2009; SILVA et al., 2015). Um dos motivos pelo qual existe uma grande busca por métodos teóricos capazes de estimar com precisão valores de $\log K_{ow}$, é devido as inúmeras dificuldades existentes entre métodos experimentais para estimar esta propriedade para compostos hidrofóbicos e para os surfactantes (HODGES et al., 2019). Algumas das dificuldades serão listadas a seguir.

A tensão superficial de equilíbrio (γ_{eq}) de uma solução de surfactante depende do tipo e da magnitude das forças intermoleculares atuando na região superficial e não é alcançada instantaneamente. Ela depende da taxa de difusão das moléculas do soluto para a superfície e a alteração da tensão superficial ocorre devido à variação da composição da superfície (BENEDETTO; ZOLOTUCHO; PRADO, 2015; EASTOE et al., 1997).

Os surfactantes possuem a tendência de formar agregados e se acumular na interface água/octanol. Portanto, a quantidade do soluto total que reside na interface é significativa (ALAMI; HOLMBERG, 2003). Isso é frequentemente considerado uma razão pela qual valores precisos para $\log K_{ow}$ pode ser difícil de medir. Além disso, pode-se formar uma emulsão ou espuma na interface octanol-água, tornando a medição um desafio (FRIBERG; JANSSON; CEDERBERG, 1976).

Outra dificuldade mais séria é a tendência dos surfactantes de formar agregados em solução aquosa. Os surfactantes se dissolvem não apenas como moléculas únicas, mas em concentrações mais altas também formam diferentes tipos de agregados solúveis. Os agregados constituem uma pseudofase pois o surfactante no agregado não está verdadeiramente em solução aquosa (NIU; LITMAN, 2002). Sua presença pode levar a uma superestimação grosseira da quantidade em solução

aquosa e, por consequência, o valor de $\log K_{ow}$ é subestimado (BERKOWITZ; BOSTICK; PANDIT, 2006; BERKOWITZ; VÁCHA, 2012; HODGES et al., 2019).

3.3.2. Métodos teóricos para estimar o $\log K_{ow}$ de surfactantes

Em princípio, é possível prever valores de $\log K_{ow}$ utilizando métodos teóricos de fácil acesso (COSTA et al., 2023). O mais comum são modelos quantitativos de estrutura-propriedade (QSPR) (MARTINS; FERREIRA, 2013; TAVARES, 2004), ferramentas de quimio(bio)informática baseados em bancos de dados e aprendizado de máquina (ALVES et al., 2017; S.K.LEE, I.H.LEE, H.J.KIM, G.S.CHANG, J.E.CHUNG, 2003), e cálculos termodinâmicos (STOLTZ, 2015a).

Os modelos de QSPR são amplamente aplicados para descrever quantitativamente as relações entre a estrutura química de moléculas e a propriedade físico-química por elas desempenhadas, buscando a identificação de valores ótimos para determinadas propriedades físico-químicas (TAVARES, 2004). No entanto, assume-se um estado neutro do composto. Além disto, o algoritmo exato não é relatado nem publicado, tornando difícil determinar a precisão ou aplicabilidade destes métodos aos surfactantes (MARTINS; FERREIRA, 2013; TAVARES, 2004; HODGES et al., 2019).

As ferramentas de quimio(bio)informática utilizam recursos de aprendizado de máquina e redes neurais para treinar modelos a partir de bases de dados. No entanto, essas ferramentas são em sua grande maioria construídas utilizando dados experimentais de moléculas pequenas, hidrofílicas e moléculas não anfífilas (ALVES et al., 2017). Portanto, a maioria dessas ferramentas não produz estimativas precisas para surfactantes (COSTA et al., 2023).

Outra forma de prever o $\log K_{ow}$ é utilizando cálculos termodinâmicos, calculando as energias livres de solvatação do soluto nos dois solventes (ΔG_w em água e ΔG_o em octanol). Os métodos mais comumente utilizados são: os métodos alquímicos de integração termodinâmica usando um ciclo termodinâmico de estados não-físico

(BOYCE et al., 2009) e o método de amostragem guarda-chuva que é usado para estimar o perfil de energia livre ao longo de uma ou mais variáveis coletivas (TORRIE; VALLEAU, 1977). Estes métodos serão discutidos com detalhes nas seções 3.4.14 e 3.4.15.

3.4. Simulação molecular

Simulação computacional tem sido uma ferramenta poderosa para a compreensão dos fenômenos físicos e químicos. Houve uma crescente busca por esta abrangente área de pesquisa, que pode ser justificada por 3 principais motivos: (I) as simulações moleculares possibilitam uma estreita ligação entre a teoria e a experiência, e quando bem elaboradas podem encaminhar a novas hipóteses e descobertas experimentais; (II) existe um grande número de graus de liberdade, quebras de simetrias e efeitos não-lineares que são entendidos mais facilmente em simulações computacionais do que através de experimentos; (III) as simulações moleculares podem alcançar um universo amplo e que não está limitado a processos que ocorrem na natureza (ADCOCK; MCCAMMON, 2006; HANSSON; OOSTENBRINK; VAN GUNSTEREN, 2002; NAMBA; DA SILVA; DA SILVA, 2008). Além disto, as técnicas de simulação fornecem imagens detalhadas do sistema, permitindo alcançar informações importantes, que são dificilmente obtidas por métodos experimentais (LYUBARTSEV; RABINOVICH, 2016). Elas podem ser usadas para quantificar as propriedades de um sistema com precisão e em uma escala de tempo inacessível de outra forma. Simulações de dinâmica molecular têm sido usadas para estudar propriedades de membranas lipídicas, pois descrevem e predizem detalhes estruturais inacessíveis a medições experimentais (LEE et al., 2014). Portanto, a simulação é uma ferramenta muito valiosa.

3.4.1. Dinâmica molecular clássica

A simulação de dinâmica molecular (DM) é uma das técnicas computacionais mais versáteis (HANSSON; OOSTENBRINK; VAN GUNSTEREN, 2002). A metodologia da DM é fundamentada nos

princípios da mecânica clássica e fornece informações sobre o comportamento dinâmico microscópico dos átomos individuais que compõem o sistema (ADCOCK; MCCAMMON, 2006; NAMBA; DA SILVA; DA SILVA, 2008). É amplamente utilizado para obter informações sobre a evolução temporal de conformações de proteínas, de outras macromoléculas biológicas e informações cinéticas e termodinâmicas.

A simulação de DM consiste da solução numérica da equação de movimento, que pode ser descrita para um sistema molecular simples pela Equação 03:

$$F_i(t) = m_i a_i \quad \text{Eq. 03}$$

$$F_i(t) = - \frac{\partial V(r_i)}{\partial r_i} \quad \text{Eq. 04}$$

onde F_i é a força que atua sobre a partícula i no sistema em um instante de tempo (t), $V(r_i)$ é o potencial que atua na partícula i na posição r_i , e a_i é a aceleração da partícula i de massa m_i . Após definir o campo de força, é possível calcular as forças que atuam sobre cada átomo, calculando-se a derivada primeira da energia potencial, obtida do campo de força escolhido, em relação às posições desses átomos como mostra a Equação 04. A Equação 04 gera de modo direto a aceleração da partícula. A partir desta, integrando-se as equações de movimento, pode-se obter as velocidades, que proporciona a mudança de posição do átomo. Através das novas posições e velocidades de cada partícula, obtêm-se as energias potencial e cinética do sistema. Aplicando-se continuamente esse procedimento, obtém-se a trajetória do sistema, que é o conjunto de posições e velocidades de cada partícula ao longo do tempo.

3.4.2. Ensembles

Em termodinâmica, ensemble é o conjunto de configurações em certas propriedades mantidas constantes durante a integração das equações de Newton que representam o estado de um sistema. Os

ensembles da termodinâmica estatística (NVE, NVT, NPT, μVT , $N\gamma T$) podem ser definidos através da mecânica quântica ou clássica (PARANJAPE, 2015).

O ensemble microcanônico (NVE) possui o número de partículas (N), volume (V) e energia total (E) constantes durante a simulação. Neste caso, o sistema deve permanecer totalmente isolado (sem trocar energia ou partículas com a sua vizinhança), a fim de atingir em equilíbrio estatístico (TOBIAS; MARTYNA; KLEIN, 1993).

No ensemble canônico (NVT), o número de partículas (N), volume (V) e temperatura (T) são mantidos constantes. É apropriado para descrever um sistema fechado em contato térmico com um banho de calor (termostato). Para alcançar o equilíbrio estatístico, o sistema deve se manter totalmente fechado sem trocar partículas com a sua vizinhança. O sistema pode entrar em contato térmico com outros sistemas que possuem a mesma temperatura (FRENKEL; SMIT; RATNER, 1997; LEACH, 2001).

O ensemble isotérmico-isobárico (NPT) mantém o número de partículas (N), pressão (P) e temperatura (T) constantes. Este conjunto desempenha um papel fundamental na química e biologia, visto que as transformações químicas e os processos biológicos ocorrem geralmente sob condições de pressão e temperaturas constantes (BUSSI; DONADIO; PARRINELLO, 2007; ZHOU et al., 2016; ZWANZIG, 1954).

O ensemble grand-canônico (μVT) usa o potencial químico (μ), volume (V) e temperatura (T) constantes. Ele é usado na descrição de um sistema que não está isolado em contato com um reservatório de calor (termostato) e de partículas (potencial químico fixo) (BUSSI; DONADIO; PARRINELLO, 2007; ZHOU et al., 2016; ZWANZIG, 1954).

No ensemble com acoplamento de tensão superficial ($N\gamma T$), são mantidos constantes o número de partículas (N), a tensão superficial (γ) e a temperatura (T) (FRENKEL; SMIT; RATNER, 1997).

3.4.3 Campos de força

Como mostrado anteriormente, para descrever a mudança da posição x de um determinado átomo i como função do tempo é necessário conhecer a massa de cada átomo (m_i) e a força (F_i) sobre cada átomo i em uma dada posição r_i . No entanto, como os átomos não estão isolados, mas ligados a outros átomos formando moléculas que, por sua vez, interagem com outras moléculas, eles estão sujeitos a forças interatômicas e intermoleculares. O cálculo destas forças é realizado através do campo de força (KAMINSKI et al., 1994; LIU et al., 2013).

Cada campo de força determina um conjunto de equações matemáticas dedicadas a descrever aspectos do comportamento molecular, como o estiramento de ligações químicas, a deformação de um ângulo de ligação ou a torção de um diedro. Estas equações são parametrizadas para reproduzir o comportamento dos compostos de interesse (KAMINSKI et al., 1994). Estas equações e parametrizações podem ser alteradas gerando diferentes campos de forças.

O avanço computacional permitiu o desenvolvimento de diversos campos de força. Os campos de força não polarizáveis disponíveis atualmente são descritos através dos modelos: atomístico ou All-Atom (AA), United-Atom (UA) e coarse grained (CG) (FUENTES-AZCATL; ALEJANDRE, 2014; KAMENIK et al., 2020; XIA; CHAI; WANG, 2017).

Os campos de forças AA descrevem todos os átomos do sistema, incluindo os hidrogênios e permitem calcular propriedades a partir de escalas nanométricas e escalas de tempo de nanossegundos, possibilitando obter informações a nível atômico. Isto implica em um maior custo computacional, mas fornecem mais detalhes em comparação com os campos de força UA e CG (MCDONALD; JORGENSEN, 1998; RIZZO; JORGENSEN, 2016).

Os campos de força UA utilizam a simplificação que os átomos de hidrogênio ligados a átomos de carbono (hidrogênios apolares) sejam unidos a este elemento, dando origem a um pseudoátomo (SMONDYREV; BERKOWITZ, 1999). Por exemplo, os grupos CH, CH₂ e

CH₃ são considerados como três partículas únicas (DUBBELDAM et al., 2004).

No campo de força CG, vários átomos são agrupados a uma única partícula chamada de “grãos”. Esse agrupamento ocorre através de um mapeamento de quatro átomos pesados para um grão, que representa o centro de interação (MARRINK et al., 2007). O campo de força CG possui um custo computacional muito menor quando comparado aos campos de força AA e UA. No entanto, estas simplificações trazem consigo algumas limitações (DE JONG et al., 2013) tais como falta de detalhamento das interações e um excessivo grude artificial entre as moléculas (FORNASIER et al., 2020; SOUZA et al., 2020a).

Na maioria dos campos de força, a energia potencial das moléculas é representada pela soma de várias contribuições em função das coordenadas, como apresentada na Equação 05:

$$\Delta E = \sum_{El} + \sum_{Ea} + \sum_{Et} + \sum_{Enl}$$

Eq. 05

onde

ΣEl : Energia em função do comprimento de ligação

ΣEa : Energia em função do ângulo de ligação

ΣEt : Energia em função da torção própria dos ângulos

ΣEnl : Energia em função das interações de átomos não ligados (Coulombianas e de van der Waals).

3.4.3.1 CHARMM/CGenFF

CHARMM (Chemistry at HARvard Macromolecular Mechanics) é um programa de simulação molecular que possui um conjunto de campos de força AA. O campo de força geral do CHARMM é conhecido como CGenFF, e tem sido usado para uma variedade de propósitos e se expandiu rapidamente para lipídios e surfactantes. O CGenFF usa a função de energia potencial aditiva para calcular a energia em função das coordenadas cartesianas do sistema (ALLOUCHE, 2012;

VANOMMESLAEGHE; MACKERELL, 2012), conforme mostrado na Equação 06:

Intramolecular (termos ligados)

$$\begin{aligned}
 E_{total} = & \sum_{\text{ligações}} K_b (b - b_o)^2 + \sum_{\text{ângulos}} K_\theta (\theta - \theta_o)^2 \\
 & + \sum_{\text{diedros}} K_\phi (1 + \cos(n\phi - \delta)) + \sum_{\substack{\text{diedros} \\ \text{impróprios}}} K_\phi (\phi - \phi_o)^2 \\
 & + \sum_{\text{Urey-Bradley}} K_{UB} (r_{1,3} - r_{1,3;o})^2
 \end{aligned}$$

Intermolecular (termos não ligados)

$$E_{total} = \sum_{\text{não ligados}} \frac{q_i q_j}{4\pi D r_{ij}} + \varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^6 \right]$$

Eq. 06

onde, r corresponde as coordenadas cartesianas do sistema. A porção intramolecular da função de energia potencial inclui termos para as ligações, ângulos de ligação, ângulos diedros, ângulos diedros impróprios e um termo Urey-Bradley. Os termos intermoleculares abrangem as interações eletrostáticas e van der Waals (vdW), onde q_i e q_j são as cargas atômicas parciais dos átomos i e j . O termo ε_{ij} é a profundidade do poço, $R_{min,ij}$ é a distância mínima entre os átomos i e j no termo Lennard-Jones (LJ) usado para tratar as interações vdW, e r_{ij} é a distância entre os átomos i e j (ALLOUCHE, 2012; VANOMMESLAEGHE; MACKERELL, 2012, 2012). A função de energia apresentada na Equação 06 foi estendida para incluir um mapa de correção de energia diédrica 2D, conhecido como CMAP, que tem sido usado para melhorar o tratamento das propriedades conformacionais dos termos ϕ e ψ (MACKERELL; FEIG; BROOKS, 2004).

As simulações moleculares de lipídios e surfactantes necessitam de parâmetros precisos capazes de reproduzir e prever propriedades experimentais. Um modelo de cadeia de átomos unidos (UA) CHARMM27/27r foi desenvolvido para os lipídios (HÉNIN; SHINODA; KLEIN, 2008). No entanto, ele sofre da falha de que as simulações de bicamada fosfolipídicas requer um conjunto de área de superfície imposta, o que limita seu uso a sistemas de bicamada puros. Um novo modelo foi desenvolvido por Lee e colaboradores (LEE et al., 2014) com base nos parâmetros lipídicos de todos os átomos CHARMM36 (C36). Este modelo concorda bem com os resultados de membrana lipídica e formação de micela de um surfactante. As simulações resultam em propriedades precisas (área de superfície por lipídio, fatores de forma de raios-X e nêutrons e parâmetros de ordem de cadeia) de várias bicamadas de cadeia saturada e insaturada (LEE et al., 2014).

3.4.3.2 AMBER/GAFF

AMBER (Assisted Model Building and Energy Refinement) refere-se a um conjunto de campos de força mecânica molecular para a simulação de computacional. Foi desenvolvido pelo grupo do pesquisador Peter Kollman da Universidade da Califórnia. Ele recebe o mesmo nome do software de mecânica molecular que originalmente simula esse campo de força (J. WANG, 2000). Ele pode também ser simulado no pacote computacional GROMACS. Como foi mencionado anteriormente cada campo de força segue uma função particular para a descrição de cada um dos termos da equação geral de potencial. A energia potencial dos campos AMBER/GAFF obedece a função mostrada na Equação 07 abaixo:

$$E_{total} = \sum_{\text{ligações}} K_r (r - r_{eq})^2 + \sum_{\text{ângulos}} K_{\theta} (\theta - \theta_{eq})^2 + \sum_{\text{diédros}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

Eq. 07

onde r_{eq} e θ_{eq} são parâmetros estruturais de equilíbrio; K_r , K_θ , V_n são constantes de força; n é a multiplicidade e γ é o ângulo de fase para os parâmetros do ângulo de torção. Os parâmetros A , B e q caracterizam os potenciais não ligados e q_i e q_j são as cargas atômicas parciais dos átomos i e j , respectivamente. O termo K_r é referente à constante de mola da ligação. Há também os termos da distância da ligação e a distância de equilíbrio. A somatória dos ângulos simboliza a energia angular entre três átomos covalentemente ligados e a somatória do diedral corresponde a energia da torção de uma ligação sobre sua ordem (BAYLY et al., 1995; DICKSON et al., 2012; J. WANG, 2000; WANG et al., 2004). As três primeiras somatórias (ligações, ângulos e diedral) são relacionadas à energia potencial da ligação covalente, enquanto o último termo é referente às interações não-covalentes (eletrostáticas e de van der Walls).

AMBER é uma família de campos de força para mecânica molecular frequentemente utilizada para modelagem de biomoléculas. Os parâmetros lipídicos reproduzem corretamente as características estruturais das bicamadas de fosfatidilcolina em uma variedade de sistemas, incluindo íons, colesterol, peptídeos e proteínas (CORDOMÍ; CALTABIANO; PARDO, 2012).

3.4.4 Cálculo do potencial

O somatório das contribuições das interações ligadas (L) e não ligadas (NL) de um conjunto de partículas define o potencial dessas interações na dinâmica molecular clássica, que está representado na Equação 08:

$$E_{total} = \sum_{i=1}^N (E_{L,i} + E_{NL,i})$$

Eq. 08

A Equação 08 mostrada acima pode ser dividida em dois termos, no primeiro termo estão as interações ligadas representada por (E_L), que são as contribuições covalentes que preserva as partículas unidas: potencial de comprimento de ligação, potencial de ângulo de ligação e o

potencial de torção de ligação. E no segundo termo tem-se as interações não ligadas representada por (E_{NL}), que correspondem a contribuição das interações de van der Waals e de Coulomb. Isto é, as interações eletrostáticas são calculadas por termos de Coulomb e as interações de van der Waals são calculadas pelo potencial de Lennard-Jones (ALLEN, 2017).

É importante ressaltar que normalmente os campos de força são agnósticos no que diz respeito aos tipos de ligações, consequentemente, não há diferença visual entre uma ligação simples, dupla ou tripla. A diferença é definida pelo comprimento da ligação e a constante de força.

3.4.4.1 Comprimento, ângulo e torção de Ligação

É possível descrever a energia da ligação covalente de duas partículas i e j em função do comprimento de ligação no campo de força, através do potencial harmônico (ALLEN, 2017) mostrado na Equação 09

$$E_{ij} = \frac{1}{2} K_{ij} (r_{ij} - r_{ij}^o)^2$$

Eq. 09

onde $r_{ij} - r_{ij}^o$ é a distância entre as partículas i e j a partir de uma referência padrão, r_{ij}^o . A constante de força de ligação específica para esse par de partículas é dada por K_{ij} . Essa aproximação é baseada na lei de Hooke em função dos termos de coordenadas internas moleculares que dão origem ao potencial, sendo válida para tratar moléculas com pequenas flutuações (CRAMER, 2004).

A deformação do ângulo é descrita através da aproximação ou afastamento que ocorre entre duas partículas. Sejam as partículas i , j e l , unidas por duas ligações, seu potencial será gerado pela deformação das mesmas e pode ser calculado pela aproximação do potencial harmônico (Equação 10)

$$E_a = \frac{1}{2} K_a (\theta_{ijl} - \theta_{ijl}^o)^2$$

Eq. 10

onde θ_{ijl} é a variação do ângulo, θ_{ijl}^o é o ângulo de equilíbrio e K_a é a constante de deformação angular das três partículas (ALLEN;, 2017; CRAMER, 2004).

O campo de força CHARMM, por exemplo, emprega uma equação adicional na descrição dos ângulos de ligação, chamada Urey-Bradley (Equação 06), que busca preservar a distância entre o primeiro e o terceiro átomos de um ângulo.

E o potencial de torção é definido pelo ângulo diedro a partir de quatro partículas i, j, k e l , onde ϕ_{ijkl} é o ângulo formado entre os planos ijk e jkl como é mostrado na Equação 11.

$$E_d = K_d(1 + \cos(n\phi_{ijkl} - \phi_{ijkl}^o))$$

Eq. 11

$$E_{di} = \frac{1}{2}K_{di}(\xi_{ijkl} - \xi_{ijkl}^o)^2$$

Eq. 12

onde ϕ_{ijkl}^o é ângulo do diedro de referência, K_d é constante do diedro e n é a multiplicidade da função. Na Equação 12 está descrito os ângulos diedros impróprios que também são representados pelos campos de forças. Neste caso, o ângulo diedro impróprio de referência é descrito por ξ_{ijkl}^o e K_{di} é a constante de ângulo diedro impróprio (ALLEN;, 2017; CRAMER, 2004).

3.4.4.2 Interações de van der Waals

Antes de iniciar a discussão é importante ressaltar que o potencial de van der Waals está relacionado com o campo de força. Consequentemente, os parâmetros de vdW em diferentes campos de forças são distintos (CRAMER, 2004).

Para o campo de força CHARMM as interações de van der Waals são calculadas com o potencial de Lennard-Jones 12-6 (Equação 13):

$$E_{vdW} = \sum_{excl(i,j)=1} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) sw(r_{ij}^2, r_{on}^2, r_{off}^2)$$

Eq. 13

onde sw é uma função de comutação, o $excl(i, j)$ é zero (0) se os átomos estiverem conectados por ângulos ou ligações ou $i \geq j$, e r_{ij} é a distância entre o centro de geometria dos grupos contendo átomos i e j . Os termos r_{ij}^{12} e r_{ij}^6 descrevem a repulsão e a atração das partículas, respectivamente (BROOKS et al., 1983).

Para o campo de força AMBER, existem duas maneiras comuns de calcular as interações do potencial de Lennard-Jones, sendo elas representadas nas Equações 14 e 15.

$$E_{vdW} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad \text{Eq. 14}$$

$$E_{vdW} = 4\epsilon \left[\frac{1}{4} \left(\frac{r_0}{r} \right)^{12} - \frac{1}{2} \left(\frac{r_0}{r} \right)^6 \right] \quad \text{Eq. 15}$$

Ambas descrevem a interação não ligada entre duas partículas na distância r como a soma de um atrativo r^6 e um r repulsivo r^{12} . Neste caso, dois parâmetros são necessários, a profundidade do poço ϵ e uma distância característica σ ou r_0 , que é a distância energeticamente mais favorável de dois átomos (BAYLY et al., 1995).

3.4.4.3 Interações Eletrostáticas ou de Coulomb

Quando se tem duas partículas carregadas, o potencial eletrostático é dado a partir da lei de Coulomb, que geralmente é calculado pela Equação 16 abaixo:

$$E_{el} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}} \quad \text{Eq. 16}$$

onde q_i e q_j descrevem a magnitude das cargas pontuais de cada partícula, r_{ij} é a distância entre as cargas, ϵ_0 é a permissividade do vácuo e ϵ_r a constante dielétrica relativa do meio (CRAMER, 2004).

3.4.5 Algoritmos na dinâmica molecular

Nos campos de força clássicos, o Hamiltoniano H que descreve o sistema pode ser dividido em energia potencial V e energia cinética K (Equação 17).

$$H(\mathbf{p}, \mathbf{r}; m, s) = K(\mathbf{p}; m) + V(\mathbf{r}; s)$$

Eq. 17

Em geral, a energia cinética não depende das coordenadas e a energia potencial não depende do momento. Dessa forma, é necessário somente as coordenadas e as funções que descrevem as interações para calcular V , e das velocidades e das massas das partículas para calcular K (ALLEN, 2017).

A configuração do sistema depois de um tempo determinado é calculada através da equação diferencial (Equação 18), que vem da segunda lei de Newton mostrada na Equação 18:

$$\frac{d^2 X_I}{dt^2} = \frac{F_{X_I}}{M_I}$$

Eq. 18

A equação descreve o movimento de uma partícula I de massa M_I com a influência de uma força F_{X_I} na dimensão X . O movimento das partículas está acoplado as suas coordenadas. Dessa forma, quando uma partícula se movimenta ocorre uma mudança na magnitude das interações com as outras partículas. Como resultado, mudará a força que atua sobre ela e a força que atua sobre outras partículas. Para resolver este problema, métodos numéricos de integração são utilizados (CRAMER, 2004).

A integração deste sistema é executada dividindo o tempo em uma série de intervalos de tempo, até que os valores de força e velocidade sejam considerados constantes. A aceleração das partículas é obtida através das coordenadas, velocidades e interações no tempo t . Paralelamente, as coordenadas e velocidades nesse tempo t são calculadas e as novas coordenadas e velocidades são medidas no tempo $t + \Delta t$. Posteriormente, com os novos valores de coordenada e

velocidade, a aceleração é calculada no tempo $t + \Delta t$ e as coordenadas e velocidades no tempo $t + 2\Delta t$.

As coordenadas e velocidades do sistema são obtidas utilizando as séries de Taylor (Equação 19):

$$\begin{aligned} R(t + \Delta t) &= R(t) + \Delta t V(t) + \frac{1}{2} \Delta t^2 A(t) + \frac{1}{6} \Delta t^3 B(t) + \frac{1}{24} \Delta t^4 C(t) + \dots \\ V(t + \Delta t) &= V(t) + \Delta t A(t) + \frac{1}{2} \Delta t^2 B(t) + \frac{1}{6} \Delta t^3 C(t) + \dots \\ A(t + \Delta t) &= A(t) + \Delta t B(t) + \frac{1}{2} \Delta t^2 C(t) + \dots \end{aligned}$$

Eq. 19

onde R é a coordenada, V é a velocidade e A é a aceleração no tempo $t + \Delta t$, B e C são a terceira e quarta derivadas da posição com respeito ao tempo (t). Para não utilizar os termos B e C uma variedade de algoritmos são usados. Os mais comuns são o algoritmo de Verlet (OLIVEI, 1976; VERLET, 1967), o algoritmo Leap-Frog (VAN GUNSTEREN; BERENDSEN, 1988) e o algoritmo de velocidade de Verlet (VERLET, 1967).

Os algoritmos envolvidos na dinâmica molecular e utilizados neste trabalho serão descritos a seguir.

3.4.5.1 Algoritmo de Verlet

O Algoritmo de Verlet (OLIVEI, 1976) é um dos mais utilizados nas simulações de DM por ser considerado simples e robusto. Este algoritmo utiliza a posição no tempo t e no tempo $t - \Delta t$ e usa aceleração para determinar a posição no tempo $t + \Delta t$. O cálculo pode ser expresso da seguinte maneira (Equação 20).

$$R(t + \Delta t) = 2R(t) - R(t - \Delta t) + \Delta t^2 A(t)$$

Eq. 20

Uma das vantagens desse algoritmo está na omissão de alguns termos da série de Taylor (VERLET, 1967; NAMBA; DA SILVA; DA SILVA, 2008). No entanto, a velocidade não é calculada explicitamente. Em geral, ela é calculada através da diferença de posições entre os

passos $t - \Delta t$ e $t + \Delta t$ dividido pelo intervalo de tempo passado $2\Delta t$, como pode ser observado na Equação 21.

$$V(t) = \frac{|R(t + \Delta t) - R(t - \Delta t)|}{2\Delta t}$$

Eq. 21

3.4.5.2 Algoritmo de minimização

A minimização de um sistema é muito importante para que ele não entre em colapso caso a configuração inicial tenha uma energia muito alta. Ele é utilizado também para eliminar as sobreposições de partículas, suavizando as distorções nas ligações químicas, nos ângulos entre as ligações e nas interações de van der Waals. Isso diminui o tempo usado para equilibrar o mesmo (BINDER et al., 2004; HANSSON; OOSTENBRINK; VAN GUNSTEREN, 2002). A minimização de energia, também conhecida como otimização da geometria, é uma técnica que busca encontrar um conjunto de coordenadas que minimizam a energia potencial do sistema de interesse.

Um dos algoritmos de minimização mais conhecidos é o método de descida mais íngreme (do termo em inglês *steepest descent*). Ele usa a derivada primeira para determinar a direção para o mínimo. Esta técnica é robusta e é usada para minimizar inicialmente uma estrutura que está longe de um ponto de mínimo. Ela consiste em alterar as coordenadas das partículas na direção de gradiente negativo da energia resultante das interações das partículas com as coordenadas iniciais (Equação 22).

$$d_n = - \frac{\partial E(C_n)}{\partial C}$$

Eq. 22

$$C_{n+1} = C_n + \Delta_n d_n$$

Eq. 23

onde ∂E é a energia em função das coordenadas C , C_n é a coordenada no passo n , Δ_n é sempre positivo para a energia sempre diminuir. O mínimo local é sempre garantido de ser encontrado. No entanto, o valor de d_n

diminui consideravelmente próximo ao mínimo e a convergência é lenta (BINDER et al., 2004).

3.4.6 Condições Periódicas de Contorno

Normalmente, os sistemas possuem entre 10^2 - 10^6 partículas em simulações de DM, sendo insignificante quando comparado com o número de átomos presentes em sistemas macroscópicos. Deste modo, o número relativo de partículas na parede do sistema nas simulações de DM é muito maior do que para um sistema macroscópico, provocando alguns efeitos de superfície indesejáveis. A forma mais usada para minimizar estes efeitos e obter as propriedades macroscópicas a partir das trajetórias das simulações de DM é a partir da utilização de condições periódicas de contorno (FRENKEL; SMIT; RATNER, 1997; LEACH, 2001).

Neste caso, os efeitos de superfície são eliminados e as partículas consideradas sofrem o efeito das forças como se estivessem no interior do sistema. Nesta técnica, os átomos são colocados em uma caixa, e a caixa original é reproduzida em todas as direções no espaço. Quando um átomo se move na caixa original durante a simulação, a sua imagem periódica se move da mesma maneira. Caso um átomo saia da caixa original, sua imagem entra pela face oposta da caixa imagem com a mesma velocidade. Desta forma, o número total de átomos na caixa central e no sistema como um todo é conservado e os efeitos de superfície desaparecem. As posições são dadas por (Equação 24)

$$r_{in} = r_{i0} + n$$

Eq. 24

com r_{i0} sendo as posições primárias das partículas, $n = n_1a_1 + n_2a_2 + n_3a_3$, onde n_1, n_2 e n_3 são números inteiros e a_i representam os vetores de base da caixa de simulação (CRAMER, 2004).

3.4.7 Termostatos

A energia cinética do sistema pode variar, modificando assim sua temperatura. No entanto, a temperatura deve ser mantida constante durante toda a simulações de DM. Neste caso, os termostatos foram

desenvolvidos com o objetivo de controlar a temperatura do sistema. Os termostatos mais conhecidos são: Berendsen (BERENDSEN et al., 1984) e reescalonamento da velocidade (BUSSI; DONADIO; PARRINELLO, 2007).

3.4.7.1 Termostato Berendsen

O controle da temperatura pode ser atingido através do termostato de Berendsen, que acopla ao sistema um banho térmico com uma temperatura fixa T_0 e uma cinética de primeira ordem como mostra a Equação 25.

$$\frac{dT}{dt} = \frac{T_0 - T(t)}{\tau_t}$$

Eq. 25

onde τ_t é uma constante de tempo e $T(t)$ corresponde a temperatura no tempo t . Durante as simulações as velocidades são reescaladas a cada passo de integração, com objetivo de ajustar a energia cinética do sistema até a temperatura escolhida (BERENDSEN et al., 1984).

3.4.7.2 Termostato de reescalonamento da velocidade

O termostato de reescalonamento da velocidade (v-rescale) é um termostato de Berendsen com adição de um termo estocástico que assegura uma distribuição de energia cinética correta modificando-o de acordo (Equação 26)

$$dK = (K_0 - K) \frac{dt}{\tau_t} + \sqrt{\frac{K K_0}{N_f}} \frac{dW}{\sqrt{\tau_t}}$$

Eq. 26

onde K é a energia cinética, N_f é o número de graus de liberdade e dW é o processo estocástico associado ao movimento Browniano, chamado processo de Wiener. O termo, τ_t é a constante de tempo para o acoplamento de temperatura. Este termostato produz um ensemble canônico correto e apresenta vantagem do termostato Berendsen devido ao decaimento de primeira ordem de desvios de temperatura e ausência de oscilações (BUSSI; DONADIO; PARRINELLO, 2007).

3.4.8 Barostatos

Além da temperatura outra variável que normalmente precisa ser controlada é a pressão. Para manter a pressão constante no decorrer da simulação utilizamos os barostatos. Os mais conhecidos são: barostatos de Berendsen (BERENDSEN et al., 1984) e o Parrinello-Rahman (PARRINELLO; RAHMAN, 1981).

Da mesma forma que no termostato, é necessário no barostato um maior controle sobre a pressão na etapa inicial da simulação, e posteriormente, pode-se diminuir a influência do barostato.

3.4.8.1. Barostato de Berendsen

A taxa de flutuação da pressão é comparada com a diferença entre a pressão no tempo t e a pressão de referência P_0 , utilizando τ_p como constante do tempo no barostato que ao ser integrada resulta na Equação 27.

$$\frac{dP}{dt} = \frac{P_0 - P(t)}{\tau_p}$$

Eq. 27

A cada n passos, o barostato de Berendsen faz um reescalamento das coordenadas e dos vetores de uma caixa, com o coeficiente de escalonamento μ , tem-se que:

$$\mu = 1 - \frac{n_{pc}\Delta t}{3\tau_p} \beta(P_0 - P(t))$$

Eq. 28

onde β é a compressibilidade isotérmica para a água a pressão de 1 atm a 300 K que equivale a $4,6 \times 10^{-5} \text{ bar}^{-1}$ e n_{pc} é o número de graus de liberdade (BERENDSEN et al., 1984).

3.4.8.2. Barostato Parrinello–Rahman

Nos casos em que as flutuações de pressão ou volume são importantes, é necessário um barostato com acoplamento mais forte como o Parrinello-Rahman (PARRINELLO; RAHMAN, 1981). Com o

barostato de Parrinello-Rahman, os vetores de caixa representados pela matriz obedecem à equação da matriz de movimento (Equação 29)

$$\frac{d^2b}{dt^2} = VW^{-1}b^{-1}(P - P_0)$$

Eq. 29

O volume da caixa é dado por V e W^{-1} é um parâmetro de matriz que determina a força do acoplamento. Os vetores da caixa são representados por b . As pressões em um dado tempo t e a pressão de referência são dadas por P e P_0 , respectivamente. Como a força de acoplamento depende do tamanho da caixa, basta fornecer as compressibilidades isotérmicas aproximadas β , a constante de tempo de pressão τ_p , e o maior comprimento da caixa L . A matriz inversa dos parâmetros de massa pode ser definida pela Equação 30.

$$(W^{-1}) = \frac{4\pi^2\beta_{ij}}{3\tau_p^2L}$$

Eq. 30

3.4.10. Algoritmo de restrição: LINCS

O algoritmo de restrição LINCS (do termo em inglês *LINear Constraint Solver*) foi criado com objetivo de solucionar problemas de vibrações de ligações, como por exemplo, para que as ligações retornem aos seus comprimentos corretos após uma atualização evitando o aquecimento do sistema. Este fenômeno costuma ser mais comum em partículas de pouca massa pois o momento médio para o sistema resulta em um grande movimento das partículas.

Uma possível correção é utilizar intervalos de tempo menores. No entanto, isto aumenta significativamente o custo computacional, inviabilizando a simulação. O método de LINCS restabelece o comprimento de ligação em cada novo cálculo de posição e corrige a modificação na rotação que resulta do restabelecimento do comprimento de ligação.

O método de LINCS introduz ao potencial $V(r)$ para uma restrição k , um multiplicador de Langrange $\lambda k(t)$ com uma função de restrição $g_k = (r_a - r_b)^2 - d_k^2$, onde r_a e r_b são as posições das partículas participantes da restrição e d_k é a distância ideal da ligação. O uso de um multiplicador de Langrange não modifica a energia total do sistema (HESS et al., 1997). A Equação 31 demonstra o método de LINCS:

$$-m_i \frac{dr_i^2}{dt^2} = \frac{d}{dr} (V(r_i(t)) - k(t)g_k)$$

Eq. 31

3.4.11. Soma de Ewald (PME)

Na DM, um dos principais problemas é o tratamento das interações de longo alcance. Normalmente, elimina-se as interações a partir de uma certa distância que pode ser o suficiente quando estamos tratando das interações de vdW. Por outro lado, as interações eletrostáticas possuem grande influência, e realizar o mesmo procedimento afastaria o resultado da realidade, principalmente quando existem muitos momentos dipolos e se usam condições periódicas de contorno (ESSMANN et al., 1995).

O PME computa as interações de dipolo em células unitárias de cristais. A energia resultante das imagens periódicas das partículas é calculada a partir da Equação 32:

$$E = \frac{f}{2} \sum_{n_x} \sum_{n_y} \sum_{n_z} \sum_i^N \sum_j^N \frac{q_i q_j}{r_{ij,n}}, (n_x, n_y, n_z) = n \neq 0$$

Eq. 32

onde n é o número de imagens da célula unitária para cada coordenada. Uma desvantagem é que ela converge lentamente. Para esta função convergir mais rápido, é necessário dividir a função em três partes independentes: a primeira parte (Equação 33) é de curto alcance (onde é calculada de forma direta), a segunda parte (Equação 34) é onde se presume uma distribuição gaussiana sobre a energia (utilizando uma transformada de Fourier), e uma terceira (Equação 35) que é inserida para manter a neutralidade de carga da célula unitária.

$$E_{short} = \frac{f}{2} \sum_{i,j}^N \sum_{n_x} \sum_{n_y} \sum_{n_z} \frac{q_i q_j}{r_{ij,n}} \operatorname{erfc}(\beta r_{ij,n})$$

Eq. 33

$$E_{away} = \frac{f}{2\pi V} \sum_{i,j}^N q_i q_j \sum_{m_x} \sum_{m_z} \sum_{m_y} \frac{\exp\left(-\left(\frac{\pi m}{\beta}\right)^2 + 2\pi i m(r_i - r_j)\right)}{m^2}$$

Eq. 34

$$E_{own} = \frac{f\beta}{\sqrt{\pi}} \sum_i^N q_i^2$$

Eq. 35

onde β é o peso do parâmetro direto e indireto da soma, V corresponde ao volume da célula unitária, m é o número de células unitárias, e q é a carga da partícula i (ESSMANN et al., 1995).

3.4.12. Modelos de água

A água é a molécula mais amplamente estudada e uma das substâncias mais importantes da Terra. A molécula da água possui uma estrutura enganosamente simples, H_2O . No entanto, apesar de décadas de estudos, muitas anomalias ainda permanecem inexplicáveis. Uma das principais características da molécula de água é a sua polarizabilidade, e capacidade de formar ligações de hidrogênio com caráter doador ou aceptor (LUNDBORG; LINDAHL, 2015).

Os modelos computacionais de água são úteis, porém sabe-se que é difícil de simular com precisão usando modelos de água atomísticos, que tendem a ser os mais eficientes computacionalmente. Utilizar um bom modelo de água é fundamental para obter resultados de alta qualidade (TEMPRA; OLLILA; JAVANAINEN, 2022).

Os modelos mais utilizados em simulações MD são: carga de ponto único (SPC) (BERENDSEN; GRIGERA; STRAATSMA, 1987), potencial intermolecular transferível com três pontos (TIP3P) (PRICE; BROOKS, 2004), potencial intermolecular transferível com quatro pontos (TIP4P)

(JORGENSEN et al., 1983) e potencial intermolecular transferível com cinco pontos (KHALAK; BAUMEIER; KARTTUNEN, 2018). Existem novos modelos com cargas de ponto ótimo com três pontos (OPC3) e quatro pontos (OPC4) (HODGES et al., 2019; HUANG; MACKERELL, 2013; TEMPRA; OLLILA; JAVANAINEN, 2022), que estão ganhando espaço nas simulações.

O modelo SPC (BERENDSEN; GRIGERA; STRAATSMA, 1987) é bastante utilizado em campos de força do GROMOS (OOSTENBRINK et al., 2004). Já o modelo TIP3P (PRICE; BROOKS, 2004) é usado principalmente com os campos de força CHARMM (HUANG; MACKERELL, 2013) e AMBER (DUAN et al., 2003). Apesar dos estudos direcionados ao desenvolvimento de novos modelos de água nenhum deles reproduz fielmente todas as características da água. Do um ponto de vista prático, a escolha do modelo de água deve ser baseada no custo computacional e a fidelidade do modelo.

Existem várias dificuldades para obter modelos de água precisos. As transições de fase da água são uma das propriedades de interesse, porém a água tem um diagrama de fases bastante complexo. Os pontos de fusão e ebulição são difíceis de serem alcançados e podem ser um grande problema para quem deseja modelar o ambiente aquoso fora da “faixa biológica” das condições ambientais (LUNDBORG; LINDAHL, 2015). Existem várias propriedades do modelo que são sensíveis a extraordinária complexidade das interações água-água reais e as redes de ligações de hidrogênio na fase líquida (IZADI; ANANDAKRISHNAN; ONUFRIEV, 2014).

Os modelos de água OPC4 e OPC3 foram desenvolvidos de forma diferente da maioria dos outros modelos de água comumente utilizados. Utilizando o espaço de parâmetros do momento multipolar elétrico, o modelo de água OPC4 foi construído como um ótimo global na reprodução de seis propriedades importantes da água bruta. Otimização global semelhante foi usada para desenvolver o modelo OPC3, que possui três locais de interação em vez de quatro para o modelo OPC4. Pouco se sabe sobre o desempenho desses modelos fora da fase líquida. Recentemente, um estudo comparou os modelos OPC4 e OPC3 em

cálculos de tensão superficial (GABRIELI et al., 2018), e os resultados mostram boa concordância com o resultado do experimento (IZADI; ANANDAKRISHNAN; ONUFRIEV, 2014; JAVANAINEN et al., 2018; TEMPRA; OLLILA; JAVANAINEN, 2022).

Estudos realizados por Tempa e colaboradores sugerem que um modelo de água com tensão superficial correta seja necessário para reproduzir as isotermas experimentais de pressão-área de superfície e o comportamento de fase de monocamadas lipídicas (TEMPRA; OLLILA; JAVANAINEN, 2022).

3.4.13. Energias Livres (ΔG)

A energia livre (G) é de suma importância na química. Quase todas as propriedades experimentais tradicionalmente interpretadas, podem ser bem descritas com conceitos de energia livre. Por exemplo, concentração de equilíbrio, taxas de reação, estabilidade, enovelamento de macromoléculas, formação de complexos, catálise, afinidade de ligação, e solubilidade são propriedades que podem ser entendidas utilizando o conceito de energia livre (LUNDBORG; LINDAHL, 2015).

Os cálculos das mudanças na energia livre de Gibbs após a transferência da fase gasosa para a fase de outro solvente amadureceu rapidamente pelo fato de a energia livre ser uma função de estado (STRAATSMA TP, BERENDSEN HJC, 1986). Dessa forma, para calcularmos sua variação, precisamos apenas calcular os estados inicial e final, sem nos preocupar com o caminho entre os dois estados. Sendo assim, pode-se calcular a variação da energia livre de forma muito mais rápida computacionalmente utilizando ciclos termodinâmicos (BOYCE et al., 2009).

Existem algumas preocupações para cálculos de energia livre de hidratação ou solvatação. Por exemplo, se a sua precisão é suficiente para uso prático (LUNDBORG; LINDAHL, 2015), se os modelos de água largamente utilizados não são tão confiáveis para modelar as energias livres de hidratação (TEMPRA; OLLILA; JAVANAINEN, 2022) e se os

campos de força moleculares descrevem corretamente as moléculas a serem estudadas (IZADI; ANANDAKRISHNAN; ONUFRIEV, 2014).

Para se obter computacionalmente a variação da energia livre de solvatação de um processo podemos utilizar vários métodos, os dois métodos utilizados neste trabalho serão apresentados nas seções a seguir.

3.4.14. Integração termodinâmica (TI)

Transformações alquímicas, na qual uma espécie química é transformada em outra através de um caminho de estados não-físicos (alquímicos), desempenham um papel importante na química computacional moderna. Por meio dos métodos alquímicos é possível desligar ou ligar gradualmente as interação soluto-solvente, utilizando um parâmetro alquímico intermolecular de desacoplamento/reacoplamento (JORGENSEN et al., 1983).

Existe uma variedade de métodos para obter resultados de cálculos alquímicos de energia livre e produzir diferenças de energia livre (KLIMOVICH; SHIRTS; MOBLEY, 2015). Conceitualmente, esses métodos podem ser divididos em duas categorias: métodos de integração termodinâmica (TI) e métodos de perturbação de energia livre (FEP) (ZWANZIG, 1954).

A integração termodinâmica (TI) utiliza uma via alquímica. Consiste em adicionar um composto X ao sistema de forma gradativa incluindo o potencial de interação de seus constituintes com o sistema e determinar a diferença de energia livre entre dois estados A e B . Esta transição entre estados é realizada usando um parâmetro λ que varia entre 0 e 1, onde o estado A se refere ao sistema sem o composto X e o estado B ao sistema com o composto X (STOLTZ, 2015a).

Para realizar a TI, utiliza-se o λ para acoplar o estado final ao estado inicial, transformando o sistema com solvente somente em soluto solvatado com o solvente (SHIRTS et al., 2003; SHIRTS; PANDE, 2005a; STOLTZ, 2015a). Para a energia livre de Helmholtz (F), podemos escrever (Equação 36):

$$F(\lambda) = -k_B T \ln(Z)$$

Eq. 36

onde Z é uma função de partição dada pela seguinte equação (Equação 37):

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

Eq. 37

onde c é a constante de normalização, que pode ser escrita como (Equação 38) e β corresponde a Equação 39:

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

Eq. 38

$$\beta = \frac{1}{k_B T}$$

Eq. 39

Se a derivada da energia livre de Helmholtz for calculada em relação ao parâmetro λ , pode-se observar que ela corresponde a Equação 40:

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

Eq. 40

Ao integrar esta expressão em λ nos limites que levam o sistema do estado A até o estado B , a variação da energia livre de Helmholtz pode ser descrita na Equação 41:

$$F^B(V, T) - F^A(V, T) = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_{NVT; \lambda} d\lambda$$

Eq. 41

E então pode-se obter para o ensemble NPT, uma expressão equivalente para encontrar a variação da energia livre de Gibbs de acordo com a Equação 42:

$$G^B(P, T) - G^A(P, T) = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_{NPT; \lambda} d\lambda$$

Eq. 42

E para conectar as energias livre de Helmholtz e energia livre de Gibbs pode-se utilizar a Equação 43

$$G = F + PV$$

Eq. 43

Portanto:

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

Eq. 44

Já o método de perturbação de energia livre (FEP), baseia-se na seguinte equação:

$$\Delta G_{10} = -RT \ln \left(e^{\frac{-\Delta H_{10}}{RT}} \right)_0$$

Eq. 45

onde R é a constante dos gases ideais, T é a temperatura, ΔH_{10} é a diferença de energia entre os estados 0 e 1, e $(\dots)_0$ indica a média de ensemble para o estado 0. Um estado é sempre tomado como referência e o próximo estado é considerado perturbado. Este método FEP produz a média exponencial de deleção (DEXP), a média exponencial de inserção (IEXP), a taxa de aceitação de Bennett (BAR) e a razão de aceitação de multiestados de Bennett (MBAR) que será discutida na próxima seção (ZWANZIG, 1954).

3.4.14.1. Análise química

Esta ferramenta é uma biblioteca Python que estima energias livres para realizar análises de cálculos químicos. A análise de energia livre

conduzida pelo GROMACS é realizada por meio dos métodos de cálculo de energia livre: a razão de aceitação de Bennett (BAR) (BENNETT, 1976), a razão de aceitação de multiestados de Bennett (MBAR) (BENNETT, 1976), a integração termodinâmica (TI) (SHIRTS et al., 2003) e a relação de Zwanzig (média exponencial). Estes métodos fornecem uma boa análise das energias livres computadas (SHIRTS et al., 2003; SHIRTS; PANDE, 2005b).

O TI é um dos métodos mais comuns para calcular as diferenças de energia livre. A energia livre pode ser calculada pela derivada da diferença de energia livre em relação a λ (BOYCE et al., 2009) através das seguintes equações:

$$G = \beta^{-1} \ln Q$$

Eq. 46

$$\frac{dG}{d\lambda} = -\beta^{-1} \frac{d}{d\lambda} \ln \int e^{-\beta U(\lambda, q)} dq = -\beta^{-1} \frac{\frac{1}{d\lambda} \int e^{-\beta U(\lambda, q)} dq}{Q}$$

Eq. 47

que pode ser reescrita pela seguinte equação:

$$\frac{dG}{d\lambda} = -\beta^{-1} \frac{-\beta \int \frac{dU(\lambda, q)}{d\lambda} e^{-\beta U(\lambda, q)} dq}{Q} = \left\langle \frac{dU(\lambda, q)}{d\lambda} \right\rangle_{\lambda}$$

Eq. 48

finalmente, realiza-se a integração em todos os parâmetros λ (Equação 49):

$$\Delta G = \int_0^1 \left\langle \frac{dU(\lambda, q)}{d\lambda} \right\rangle d\lambda$$

Eq. 49

A derivação torna bastante simples o processo de estimar as energias livres a partir do método TI. Isto porque há a necessidade de informações de apenas um único estado são necessárias para calcular a derivada, e não há necessidade de solução iterativa e (STOLTZ, 2015b).

O método BAR é um dos primeiros métodos de energia livre que se baseia em dados de vários estados para estimar a diferença de energia livre, ou seja, requer informações de configuração de pelo menos dois

estados. O BAR requer uma solução iterativa, mas todas as amostras são correlacionadas, pois os resultados são obtidos de estados adjacentes. Ao contrário de TI, os resultados não são tão claros para obter uma estimativa de erro (BENNETT, 1976).

O método MBAR é uma extensão do método BAR. O primeiro possibilita avaliar dados de todos os estados e prever a energia livre até mesmo de um estado não amostrado. As energias livres fornecidas por este método são apenas utilizadas por um estimador estatístico. Entretanto, o MBAR demonstrou ter a menor variância até o momento. O MBAR é derivado de um conjunto $K \times K$ de funções ponderadas $\alpha_{i,j}(q)$, que minimiza a variância durante a preponderação geral (ZHOU et al., 2016). O método MBAR origina-se a partir da equação de energia livre central (Equação 50):

$$\Delta G_{ij} = -\beta^{-1} \ln \frac{Q_j}{Q_i} \quad \text{Eq. 50}$$

Pode-se manipular a mesma identidade que o BAR usa ao iniciar a derivação e então pode-se escrever as Equações 51 e 52:

$$Q_i \langle \alpha_{i,j} \exp(-\beta U_j) \rangle_i = Q_j \langle \alpha_{i,j} \exp(-\beta U_i) \rangle_j \quad \text{Eq. 51}$$

$$\sum_{j=1}^K \frac{\hat{Q}_i}{N_i} \sum_{n=1}^{N_i} \alpha_{i,j} \exp(-\beta U_j(q_{i,n})) = \sum_{j=1}^K \frac{\hat{Q}_j}{N_j} \sum_{n=1}^{N_j} \alpha_{i,j} \exp(-\beta U_i(q_{j,n})) \quad \text{Eq. 52}$$

assumindo que o estimador é usado e escolhendo os parâmetros $\alpha_{i,j}$ ideais, obtém-se:

$$\alpha_{i,j} = \frac{N_j \hat{c}_j^{-1}}{\sum_{k=1}^K N_j \hat{c}_j^{-1} \exp(-\beta U_k)} \quad \text{Eq. 53}$$

Após realizar todas as substituições necessárias, obtém-se uma expressão para uma energia livre estimada de:

$$\hat{G}_l = -\beta^{-1} \ln \sum_{j=1}^K \sum_{n=1}^{N_j} \frac{\exp[-\beta U_l]}{\sum_{k=1}^K N_k \exp[\beta \hat{G}_k - \beta U_k]}$$

Eq. 54

Observe que se tem uma única energia livre (Equação 54), não uma diferença. Neste caso, as energias livres para um determinado conjunto de estados são determinadas exclusivamente até uma constante aditiva. Devido a isto, uma energia livre deve ser tomada como referência (BOYCE et al., 2009).

3.4.15. Amostragem guarda-Chuva

A amostragem guarda-chuva foi desenvolvida por Torrie e Valleau em 1977 (TORRIE; VALLEAU, 1977). É uma técnica utilizada para obter perfis de energia livre e outras informações termodinâmicas, por meio de amostragem conformacional aprimorada ao longo de uma coordenada de reação (ξ). Logo após o término da simulação, uma série de configurações são geradas entre as duas espécies que devem ser totalmente conectadas através da sobreposição das janelas ao longo da ξ . Uma dessas espécies é mantida como referência, enquanto a outra é fixada a uma certa distância do centro de massa da espécie de referência. Em geral, a ξ é definida em termos geométricos, como distâncias, torsões e RMSD de dois estados de referências. Após a definição da ξ é possível calcular a energia livre (G) de um determinado processo através da Equação 55:

$$G(\xi) = -\frac{1}{\beta} \ln Q(\xi)$$

Eq. 55

onde $G(\xi)$ é o potencial de força média (PMF). O termo β é o inverso da constante de Boltzmann (k_b) multiplicada pela temperatura absoluta T , e Q é a função de partição canônica. Para determinar $G(\xi)$ em simulações computacionais, é necessário capturar configurações suficientes para obter espaçamentos regulares entre as janelas, cujas distribuições se

sobrepõem, diminuindo assim o erro estatístico. Com a suposição de ter um sistema ergódico, a $Q(\xi)$ é igual a função de distribuição $P(\xi)$, que é calculada pela Equação 56:

$$P(\xi) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^1 \rho[\xi(t')] dt'$$

Eq. 56

onde t corresponde ao tempo, e ρ é um termo que conta a ocorrência dos estados em um intervalo de ξ .

Este método utiliza a aplicação de potenciais harmônicos ao longo de uma coordenada para que o sistema seja conduzido de um estado termodinâmico para outro (Equação 57) (TORRIE; VALLEAU, 1977). A amostragem deste processo é geralmente realizada através de uma série de janelas para mostrar o intervalo de cada ξ .

$$\omega_i(\xi) = \frac{K}{2} (\xi - \xi_{ref})^2$$

Eq. 57

onde ω_i é o potencial de polarização da janela i , K é a constante de força harmônica e ξ_{ref} é a coordenada de reação de referência. A aplicação do potencial em cada janela é usada para manter o sistema próximo ao ponto de referência ξ_{ref} da respectiva janela i . Neste método, a energia livre (G) é determinada indiretamente através da função de distribuição não enviesada (P_i^u), conforme mostra as Equações 58 e 59:

$$P_i^u = P_i^b(\xi) \exp[\beta \omega_i(\xi)] \langle \exp[-\beta \omega_i(\xi)] \rangle$$

Eq. 58

$$G(\xi) = -\frac{1}{\beta} \ln P_i^b(\xi) - \omega_i(\xi) + F_i$$

Eq. 59

onde P_i^b é a função de distribuição enviesada e F_i é uma ponderação que é calculado por $F_i = -1/\beta \ln \langle \exp[-\beta \omega(\xi)] \rangle$.

As janelas de amostragem podem ser utilizadas através do método da análise do histograma ponderado (WHAM) (KUMAR et al., 1992) para estimar o erro da energia livre obtida.

3.4.15.1. Método da análise do histograma ponderado (WHAM)

A análise do histograma ponderado (WHAM) (do termo em inglês *Weighted Histogram Analysis Method*) (KUMAR et al., 1992) é um método para estimar a incerteza estatística da distribuição de probabilidade parcial através dos histogramas guarda-chuva. Este método minimiza o erro estatístico associado a uma função de distribuição global $P^u(\xi)$ (KUMAR et al., 1992). A determinação da distribuição $P^u(\xi)$ é calculada através de uma média ponderada das distribuições $P_i^u(\xi)$ das janelas individuais i , mostrada a Equação 60:

$$P^u(\xi) = \sum_i^{Janelas} p_i(\xi) p_i^u(\xi) \quad \text{Eq. 60}$$

onde $p_i(\xi)$ é um peso usado para minimizar o erro estatístico da distribuição global, representado pelas Equações 61 e 62:

$$p_i = \frac{a_i}{\sum_i a_i} \quad \text{Eq. 61}$$

$$a_i(\xi) = N_i \exp[-\beta \omega_i(\xi) + \beta F_i] \quad \text{Eq. 62}$$

no qual N_i corresponde ao número de amostras para cada janela i . Dessa forma, é possível observar a associação da ponderação F_i com a distribuição global $P^u(\xi)$ mostrada na Equação 63 e o peso $p_i(\xi)$ pelas Equações 61 e 62. No WHAM, a ponderação F_i pode ser calculada através da Equação 63:

$$\exp(-\beta F_i) = \langle \exp(-\beta \omega(\xi)) \rangle = \int p^u(\xi) \exp[-\beta \omega_i(\xi)] d\xi \quad \text{Eq. 63}$$

Associando a Equação 60 onde tem-se $P^u(\xi)$ e as Equações 61 e 62 que mostram a associação de F_i com $P^u(\xi)$, a presença dos dois termos na Equação 63 sugere que estes devem ser iterados até a convergência para obtenção do PMF (KUMAR et al., 1992).

3.4.16. Gromacs

O GROMACS (GRONingen MACHine for Chemical Simulations) é um pacote de DM projetado principalmente para simulações de proteínas, lipídios e ácidos nucleicos, mas funciona para outros tipos de sistemas. Foi desenvolvido pelo grupo de Berendsen na Universidade de Groningen na Holanda. É um dos pacotes mais rápidos e mais populares de química computacional disponíveis até o presente momento. O pacote computacional GROMACS não possui um campo de força próprio. No entanto, tem objetivo de ser versátil e robusto sendo compatível com vários campos de força GROMOS, OPLS, AMBER, CHARMM e outros (VAN DER SPOEL et al., 2005).

4. Metodologia

4.1. Análises de quimio(bio)informática

As predições de quimio(bio)informática para o $\log K_{ow}$ foram obtidas através de ferramentas de aprendizado de máquina de grandes conjuntos de dados de pequenas moléculas. Os programas Molinspiration Cheminformatics (KUMAR et al., 2018), SwissADME (DAINA; MICHIELIN; ZOETE, 2014, 2017), PreADMET (BROWN, 2001), FAF-drugs4 (LAGORCE et al., 2008), ALOGPS (TETKO; TANCHUK, 2002) e RDKit (GREG et al., 2022) foram utilizados com o código SMILES do DPPC (CCCCCCCCCCCCCCCC(=O)OCC(COP(=O)([O-])OCC[N+](C)(C)C)OC(=O)CCCCCCCCCCCCCCC). O SwissADME (DAINA; MICHIELIN; ZOETE, 2014, 2017) utiliza o $\log K_{ow}$ em consenso, que consiste na média aritmética dos valores previstos por cinco métodos (XLOGP3 (CHENG et al., 2007), WLOGP (WILDMAN; CRIPPEN, 1999), MLOGP (MORIGUCHI et al., 1992, 1994), SILICOS-IT (DAINA; MICHIELIN; ZOETE, 2017) e iLOGP (DAINA; MICHIELIN; ZOETE, 2014)).

4.2. Dinâmica molecular

Simulações atomísticas foram realizadas com o software Gromacs 5.0.6 (VAN DER SPOEL et al., 2005). Foram utilizados os campos de força AMBER/GAFF (DUAN et al., 2003) e CHARMM/CGenFF (HUANG; MACKERELL, 2013), e os modelos de água SPC216 (BERENDSEN; GRIGERA; STRAATSMA, 1987), TIP3P (PRICE; BROOKS, 2004), TIP4P (JORGENSEN et al., 1983), TIP5P (KHALAK; BAUMEIER; KARTTUNEN, 2018), OPC3 e OPC4 (IZADI; ANANDAKRISHNAN; ONUFRIEV, 2014; JAVANAINEN et al., 2018; TEMPRA; OLLILA; JAVANAINEN, 2022). Todos os parâmetros de interação do DPPC e do octanol para o campo de força AMBER/GAFF (DUAN et al., 2003) foram obtidos usando o código ACPYPE (SOUSA DA SILVA; VRANKEN, 2012) e os do campo de força CHARMM/CGenFF (HUANG; MACKERELL, 2013) foram obtidos através do CHARMM-GUI (ALLOUCHE, 2012; LEE et al., 2016).

Duas caixas com dimensões 5 nm x 5 nm x 5 nm foram geradas para simulações independentes, uma contendo água e outra contendo octanol puro. Um caixa com mistura água/octanol também foi utilizada com 27% (mol/mol) de água no octanol que será denominado de octanol úmido (FAN; IORGA; BECKSTEIN, 2020; LANG, 2012). A caixa de água foi preenchida com 4000 moléculas de água em torno de uma molécula de DPPC no centro da caixa. A caixa de octanol puro foi preenchida com 200 moléculas de octanol em torno de uma molécula de DPPC no centro da caixa. A caixa de octanol úmido foi preenchida com 146 moléculas de octanol, 54 moléculas de água e uma molécula de DPPC no centro da caixa.

A etapa de minimização de energia foi realizada usando o algoritmo de descida mais íngreme até que o gradiente de energia não ultrapasse o critério de $100 \text{ kJ mol}^{-1} \text{ nm}^{-1}$. Uma etapa de equilíbrio de 10 ns foi realizada utilizando-se o ensemble NVT com o termostato de reescalonamento da velocidade (BUSSI; DONADIO; PARRINELLO, 2007) a 298 K ($\tau = 0,1 \text{ ps}$). Em seguida, foi realizada uma etapa de equilíbrio com ensemble NPT de 50 ns seguida de uma simulação de dinâmica molecular de produção de 50 ns utilizando o termostato de reescalonamento da velocidade (BUSSI; DONADIO; PARRINELLO, 2007) a 298 K ($\tau = 2 \text{ ps}$) e o barostato de Berendsen (BERENDSEN et al., 1984) a 1 bar ($\tau = 2 \text{ ps}$). Em todas as simulações neste estudo, os critérios de corte para interações de Lennard Jones e de Coulomb foram de 1,0 nm. Exceto para as simulações usando o modelo de água OPC4, onde os critérios de Lennard Jones e de Coulomb foram de 1,2 nm (TEMPRA; OLLILA; JAVANAINEN, 2022).

4.2.1. Cálculo de energia livre Gibbs usando integração termodinâmica (TI)

Posteriormente, as trajetórias foram amostradas para cálculo de ΔG_{ow} usando o método TI (BOYCE et al., 2009). Este método separa o sistema numa série de estados intermediários não físicos (ou alquímicos) chamados de lambdas (λ). Depois, ele calcula a

variação da energia interna dU em função destes λ . É possível estimar a energia livre alquímica com base em estados alquímicos com $dU/d\lambda$ (SHIRTS; CHODERA, 2008b). Para estimar a energia livre alquímica de solvatação foi aplicado o estimador MBAR (SHIRTS; CHODERA, 2008a). O método MBAR é uma extensão direta do método BAR (KIM; ALLEN, 2012), pois prevê a energia livre em um estado não amostrado avaliando dados de todos os estados para obter a energia livre. Além disso, ele também usa o estimador de integração termodinâmica (TI) por ser um método mais comum e de fácil uso.

Para calcular a energia livre desse sistema alquímico, as caixas de água foram simuladas por desacoplamento das interações utilizando 26 λ em intervalos de lambdas de Coulomb:

0.00 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00 1.00 1.00
1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00

E para as interações de van der Waals foram utilizados 26 lambdas:

0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.05 0.10
0.20 0.30 0.40 0.50 0.60 0.65 0.70 0.75 0.80 0.85 0.90 0.95 1.00.

O cálculo da energia livre desse sistema alquímico foi feito com as caixas de octanol puro/úmido simuladas por desacoplamento das interações utilizando 41 lambdas em intervalos de Coulomb:

0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60
0.65 0.70 0.75 0.80 0.85 0.90 0.95 1.00 1.00 1.00 1.00 1.00 1.00
1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
1.00 1.00

e para as interações de van der Waals utilizou-se 41 lambdas:

0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.05 0.10 0.15 0.20 0.25
0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70 0.75 0.80 0.85 0.90
0.95 1.00.

Em cada λ , as etapas de minimização de energia, equilíbrio NVT de 10 ps, equilíbrio NPT de 100 ps, e dinâmica molecular de produção de 5 ns foram realizadas com o termostato de

reescalonamento da velocidade (BUSSI; DONADIO; PARRINELLO, 2007) a 298 K ($\tau = 0,5$ ps) e com o barostato de Berendsen (BERENDSEN et al., 1984) a 1 bar ($\tau = 2$ ps). Os demais parâmetros do protocolo de simulação foram os mesmos aplicados na etapa de dinâmica molecular descrita acima. Uma vez concluídas todas as simulações, os resultados foram analisados com o código de python `alchemical_analysis.py` (KLIMOVICH; SHIRTS; MOBLEY, 2015). O ciclo termodinâmico utilizado para obter as energias livres de solvatação com água (ΔG_w) e com octanol (ΔG_o) está na Figura 04.

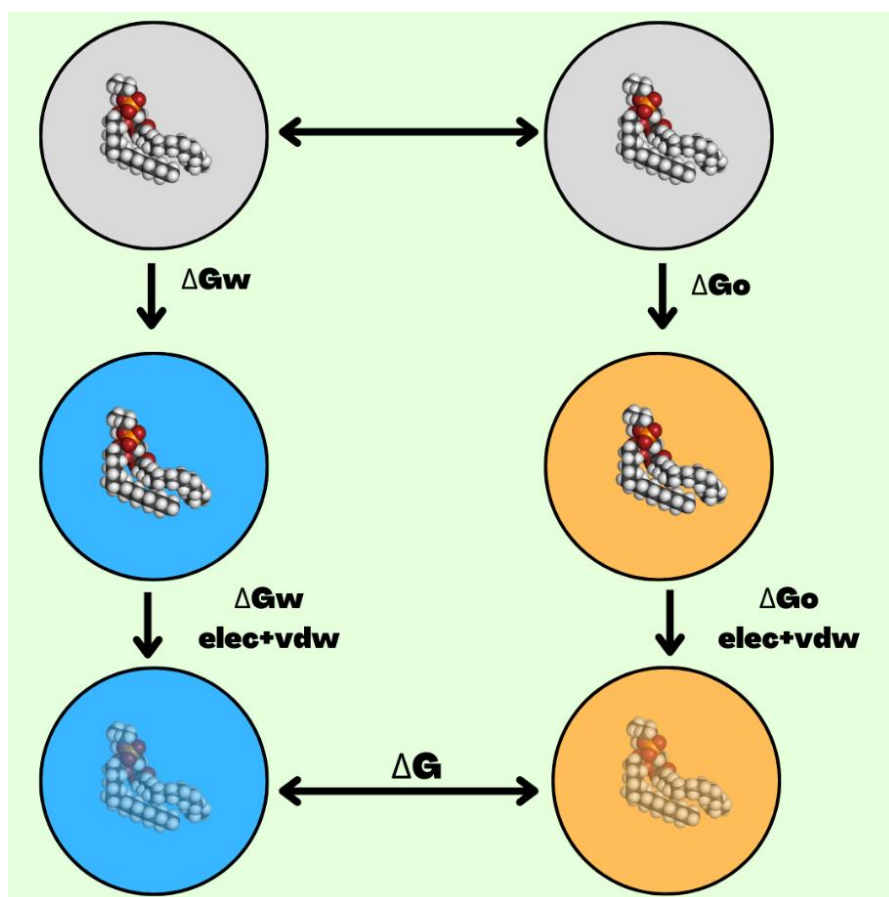


Figura 04 - Ilustração do ciclo termodinâmico utilizado no sistema em estudo. A energia livre de Gibbs (ΔG) é calculada contabilizando a solvatação do DPPC pela água (ΔG_w) e a solvatação do DPPC pelo octanol (ΔG_o). As siglas "elec" e "vdW" significam contribuições eletrostáticas e de van der Waals, respectivamente. A representação gráfica do ciclo termodinâmico

não físico mostra a molécula “transparente” que não está interagindo com o ambiente (ou seja, está desacoplada). O fundo azul é um lembrete de que a água está presente e o fundo laranja significa que o octanol está presente. O fundo cinza representa o vácuo.

4.3. Simulação de tendência

A simulação de tendência foi realizada utilizando o campo de força CHARMM/CGenFF (HUANG; MACKERELL, 2013) com modelo de água TIP3P (PRICE; BROOKS, 2004). Nestas simulações, o modelo octanol foi obtido no CHARMM-GUI (ALLOUCHE, 2012; LEE et al., 2016). Em seguida, foi construído um sistema bifásico contendo água e octanol em uma caixa de dimensões 5 nm × 5 nm × 24 nm. O sistema de tendência direta é apresentado na Figura 05 que mostra duas caixas pré-equilibradas de octanol separadas por uma caixa de água pré-equilibrada. O sistema de tendência inversa é mostrado na Figura 06 que apresenta duas caixas pré-equilibradas de água separadas por uma caixa pré-equilibrada de octanol.

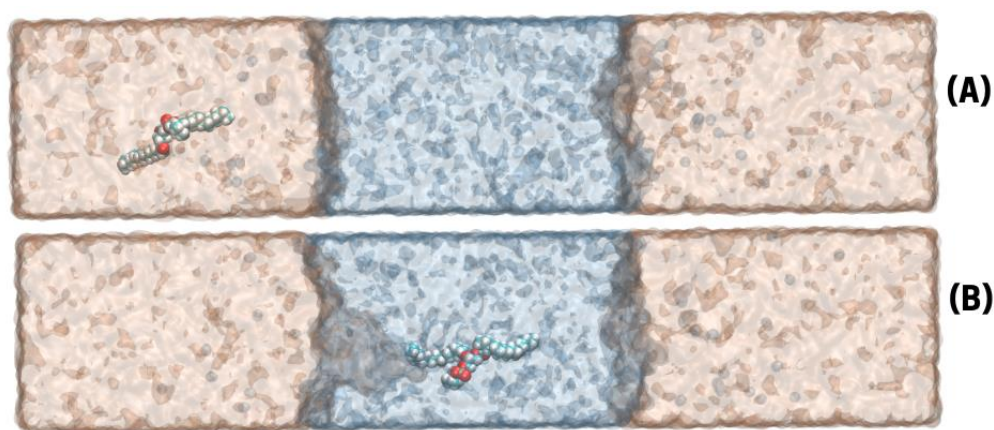


Figura 05 – (A) Sistema inicial de tendência direta onde o DPPC está na fase octanol (em laranja), e (B) sistema final de tendência direta onde o DPPC está na fase aquosa (em azul).

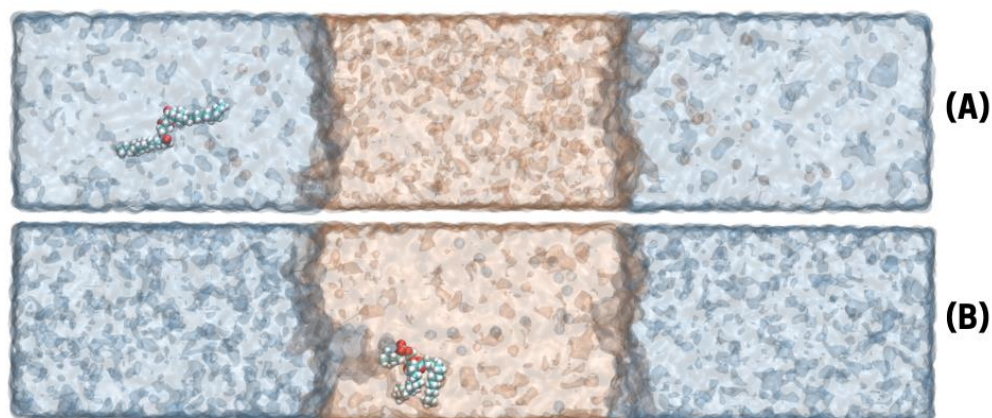


Figura 06 - (A) Sistema inicial de tendência inversa onde o DPPC está na fase aquosa (em azul), e (B) sistema final de tendência inversa onde o DPPC está na fase octanol (em laranja).

O sistema de tendência direta foi construído com dimensões de 5 nm × 5 nm × 8 nm contendo cerca de 6793 moléculas de água, 1605 moléculas de octanol e 1 molécula de DPPC inserida no centro de uma caixa de octanol. O sistema de tendência inversa possui as mesmas dimensões e contém cerca de 13593 moléculas de água, 804 moléculas de octanol e 1 molécula de DPPC inserida no centro de uma caixa de água.

A etapa de minimização de energia foi realizada usando o algoritmo de descida mais íngreme até que o gradiente de energia não exceda $100 \text{ kJ mol}^{-1} \text{ nm}^{-1}$. A etapa de equilíbrio com ensemble NPT de 50 ns foi executada utilizando o termostato de reescalonamento da velocidade (BUSSI; DONADIO; PARRINELLO, 2007) a 298 K ($\tau = 0,1 \text{ ps}$) e o barostato de Berendsen (BERENDSEN et al., 1984) a 1 bar ($\tau = 2 \text{ ps}$). Para transferir o DPPC entre as fases, foi aplicada uma constante de força de $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. O acoplamento de pressão isotrópica foi utilizado com a compressibilidade ajustada para $4,5 \times 10^{-5} \text{ bar}^{-1}$ em todas as direções. Os critérios de corte para as interações de Lennard Jones e para interações de Coulomb foram de 1,4 nm.

4.3.1. Cálculo de energia livre de Gibbs usando amostragem guarda-chuva

As trajetórias da dinâmica molecular foram amostradas para a realização do cálculo de ΔG_{ow} utilizando o método US (TORRIE; VALLEAU, 1977) com janelas igualmente espaçadas em 0,2 nm. Em cada janela, uma etapa de equilíbrio de 20 ns foi realizada utilizando o ensemble NPT com o termostato de reescalonamento da velocidade (BUSSI; DONADIO; PARRINELLO, 2007) a 298 K ($\tau = 0,1$ ps) e o barostato de Berendsen (BERENDSEN et al., 1984) a 1 bar ($\tau = 1$ ps). Em seguida, foi realizada uma etapa de dinâmica molecular de produção de 1 ns com o termostato de reescalonamento da velocidade (BUSSI; DONADIO; PARRINELLO, 2007) a 298 K ($\tau = 0,1$ ps) e barostato de Parrinello-Rahman (PARRINELLO; RAHMAN, 1981) a 1 bar ($\tau = 2$ ps). A distância entre o centro de massa da molécula de DPPC e o centro de massa da caixa de água ou de octanol foi utilizada como referência para amostragem. Nessas trajetórias, o potencial de força média foi calculado em função da coordenada de reação (ξ). Os perfis de energia de livre de Gibbs foram obtidos pelo método de análise de histograma ponderado (KUMAR et al., 1992). Todos os cálculos aqui mencionados foram realizados sob condições periódicas de contorno.

5. Resultados e discussões

5.1. Ferramentas de quimio(bio)informática

Existe uma grande variedade de ferramentas de quimio(bio)informática destinadas a prever parâmetros de ADMET a partir de estruturas moleculares. São amplamente utilizadas com o intuito de tomar decisões melhores e mais rápidas na área de identificação e otimização dos mais diversos compostos. Os dados $\log K_{ow}$ são descritores clássicos para a lipofilicidade de compostos e podem ser previstos por estas ferramentas (ARNOTT; PLANEY, 2012).

Os resultados que estão apresentados na Tabela 01 foram obtidos após uma previsão dos dados de $\log K_{ow}$ utilizando as ferramentas de quimio(bio)informática. É possível observar que os valores variam de 5,39 a 14,86. O resultado mais próximo do valor experimental encontrado na literatura ($11,50 \pm 1,43$) (SOUZA et al., 2020a) foi obtido pelo PreADMET com $\log K_{ow}$ igual a 10,02. O PreADMET produz estimativas de $\log K_{ow}$ razoavelmente precisas dos valores. Este método é aplicável em diversas áreas relacionadas da química computacional e medicinal, incluindo aplicações de 3D-QSAR, geração de superfícies lipofílicas de moléculas e avaliação de interações hidrofóbicas em complexos (GHOSE; VISWANADHAN; WENDOLOSKI, 1998).

Tabela 01 - Os coeficientes de partição octanol-água $\log K_{ow}$ previstos para o DPPC por ferramentas de quimio(bio)informática em comparação com o valor experimental obtido na literatura (SOUZA et al., 2020a).

Ferramenta de quimi(bio)informática	$\log K_{ow}$
Molinspiration	6,38
SwissADME	8,20*
PreADMET	10,02
FAF-drugs4	14,86
ALOGPS	5,39
RDKit	6,98
Experimental	11,50±1,43

Apesar dos diferentes métodos utilizados para prever $\log K_{ow}$, este estudo mostra a maioria das ferramentas ainda falham para moléculas hidrofóbicas e anfifílicas (FORNASIER et al., 2020; SOUZA et al., 2020a). Alguns estudos mostram que a predição desta propriedade para compostos zwitteriônicos e tautoméricos não é prevista corretamente por ferramentas de contribuição de grupo (EDWARD; CHUBB; SANGSTER, 1997; MACHATHA; YALKOWSKY, 2005; COSTA et al., 2023).

No geral, todas as ferramentas de quimio(bio)informática são construídas utilizando dados experimentais de moléculas pequenas e polares, e não moléculas anfifílicas ou hidrofóbicas. Além disso, a maioria dessas ferramentas não produz estimativas precisas para fosfolipídios porque são baseadas em modelos de aprendizado de máquina de grandes conjuntos de dados de moléculas pequenas (COSTA et al., 2023). Com base nos dados apresentados, conclui-se que as ferramentas quimio(bio)informática podem ser inadequadas para surfactantes. As ferramentas de quimio(bio)informática foram usadas apenas como referência e para comparação de diferentes estimadores de $\log K_{ow}$. Assim, os resultados utilizando cálculos termodinâmicos serão apresentados nas seções a seguir.

5.2. Integração termodinâmica

Além dos métodos de estimação mencionado acima, o $\log K_{ow}$ também pode ser obtido através das energias livres de Gibbs de

solvatação ou da energia livre de transferência (ΔG_{ow}) entre duas fases diferentes na mesma temperatura por meio de simulação de DM.

A energia livre de transferência da molécula de DPPC do octanol (puro ou úmido) para a fase aquosa é obtida através de cálculos de solvatação em água e em octanol utilizando os métodos TI (BOYCE et al., 2009) e MBAR (SHIRTS; CHODERA, 2008b). As energias livres são obtidas através do ciclo termodinâmico mostrado na Figura 04. Desta forma, os $\log K_{ow}$ foram calculados a partir da Equação 01.

Além do octanol puro, optou-se por utilizar uma mistura de água e octanol (octanol úmido) para obter uma melhor aproximação ao sistema real. Desta forma, pode-se comparar o cálculo teórico com maior confiabilidade com o valor de $\log K_{ow}$ experimental obtido através do método de agitação lenta disponível na literatura (SOUZA et al., 2020a). A mistura de água e octanol foi utilizada para mimetizar a solubilidade experimental de água em octanol que é de 27% (mol/mol) na temperatura de 298 K (CUMMING; RÜCKER, 2017; FAN; IORGA; BECKSTEIN, 2020). O octanol também possui uma certa solubilidade em água. No entanto, sua solubilidade é muito baixa em água a 298K. Portanto, a simulação de octanol como impureza na fase aquosa não foi realizada (FAN; IORGA; BECKSTEIN, 2020).

5.2.1. Estimador MBAR

A Tabela 02 apresenta os resultados para o estimador MBAR utilizando o sistema com octanol puro e úmido com seis modelos de água e com os campos de força AMBER/GAFF e CHARMM/CGenFF.

Tabela 02 - Energia livre de Gibbs de transferência (ΔG_{ow}) do DPPC da fase de octanol (puro e úmido) para a fase aquosa e coeficiente de partição octanol-água ($\log K_{ow}$) utilizando a integração termodinâmica (estimador MBAR) com campos de força AMBER/GAFF e CHARMM/CGenFF.

Método	Campo de força	Fase do octanol	Modelo de água	ΔG_{ow} (kcal mol ⁻¹)	$\log K_{ow}$
MBAR	AMBER	Puro	TIP3P	-20,0±0,6	14,6±0,6
			TIP4P	-19,0±0,6	13,9±0,6
			TIP5P	-30,2±0,6	22,1±0,6
			SPC	-21,5±0,7	15,7±0,7
			OPC4	-10,4±0,6	7,6±0,6
		Úmido	TIP3P	-21,5±0,4	15,7±0,4
			TIP4P	-27,3±0,5	19,9±0,5
			TIP5P	-32,5±0,5	23,8±0,5
			SPC	-28,2±0,4	20,6±0,4
			OPC4	-19,8±0,3	14,5±0,3
	CHARMM	Puro	OPC3	-26,1±0,4	19,1±0,4
			OPC4	-22,9±0,1	16,4 ±0,1
		Úmido	OPC3	-24,1±0,5	17,7±0,5
			OPC4	-20,2±0,3	14,8±0,3

Para o sistema octanol puro utilizando o campo de força AMBER/GAFF e o modelo TIP4P, o ΔG_{ow} é de 19,0±0,6 kcal mol⁻¹, o que fornece um $\log K_{ow}$ de 13,9±0,6. Além disso, o ΔG_{ow} é estimado ser 10,4±0,6 kcal mol⁻¹ usando o campo de força AMBER/GAFF e o modelo de água OPC4, fornecendo um $\log K_{ow}$ igual 7,6±0,6. Esses resultados são os mais próximos para o sistema simulado com octanol puro, embora ambos diferem em cerca de duas ou três ordens de magnitude do valor experimental de 11,50±1,43 como mostra a Tabela 01. No entanto, é importante enfatizar que o octanol puro não representa a realidade.

O sistema é também calculado com octanol úmido e outros cinco modelos de água usando o campo de força AMBER/GAFF. Para este sistema, o ΔG_{ow} calculado com o modelo de água TIP4P tem um valor consideravelmente superior ao do sistema com octanol puro, que é estimado em 27,3±0,5 kcal mol⁻¹, fornecendo um $\log K_{ow}$ igual a 19,9±0,5. O ΔG_{ow} calculado com o modelo de água OPC4 é estimado ser 19,8±0,3 kcal mol⁻¹, que fornece um

valor mais elevado de $\log K_{ow}$ igual a $14,5 \pm 0,4$. Além disso, vale a pena mencionar que os calculados para o sistema octanol úmido usando os modelos TIP3P, TIP5P e SPC são: $21,5 \pm 0,4$, $32,5 \pm 0,5$ e $28,2 \pm 0,4$ kcal mol⁻¹, fornecendo $\log K_{ow}$ igual a $15,7 \pm 0,4$, $23,8 \pm 0,5$ e $20,6 \pm 0,4$, respectivamente. Com exceção do modelo de água OPC4, o sistema octanol puro também apresenta valores mais elevados do que a medição experimental.

Depois desta análise é possível observar que os valores de $\log K_{ow}$ para octanol puro estão mais próximos do valor experimental e em geral são menores em comparação com o octanol úmido para o campo de força AMBER/GAFF. Todavia, como mencionado anteriormente, o octanol puro não representa o que acontece experimentalmente, se tratando apenas uma situação idealizada. Com isso, os resultados para o sistema com octanol puro serviram apenas para comparação e não podem ser comparadas aos dados experimentais. Por fim, os valores calculados do $\log K_{ow}$ utilizando o octanol puro tendem a ser menores, mas menos precisos e não são realistas. Essa característica não é somente do DPPC. O estudo realizado por Fan e colaboradores (FAN; IORGA; BECKSTEIN, 2020), demonstra que os $\log K_{ow}$ para sistemas com octanol puro e com octanol úmido são diferentes para muitos outros compostos.

Os resultados para o sistema com octanol puro utilizando dois modelos de água (OPC3 e OPC4) e o campo de força CHARMM/CGenFF acontece o oposto. O $\log K_{ow}$ para o sistema octanol úmido é menor quando comparado com o octanol puro. O ΔG_{ow} calculado com o modelo de água OPC3 no sistema com octanol puro é $26,1 \pm 0,4$ kcal mol⁻¹, fornecendo um $\log K_{ow}$ igual a $19,1 \pm 0,4$. Para o modelo de água OPC4, o ΔG_{ow} é calculado como sendo $22,9 \pm 0,1$ kcal mol⁻¹, o que dá um $\log K_{ow}$ igual a $16,4 \pm 0,1$. O ΔG_{ow} para o octanol úmido e modelo de água OPC3 é estimado em $24,1 \pm 0,5$ kcal mol⁻¹, dando $\log K_{ow}$ igual a $17,7 \pm 0,5$. Para o modelo OPC4, o ΔG_{ow} é de $20,2 \pm 0,3$ kcal mol⁻¹, o que dá um $\log K_{ow}$ de $14,8 \pm 0,3$.

É notável que os modelos de água influenciam nos resultados de $\log K_{ow}$ calculados. Um dos principais objetivos do desenvolvimento de novos modelos de água é para melhorar a precisão dos efeitos de hidratação em simulações de DM (COSTA et al., 2023). O modelo de água TIP3P é utilizado como padrão no desenvolvimento do campo de força AMBER/GAFF. É possível observar na Tabela 02 que o valor de $\log K_{ow}$ para este modelo de água é consideravelmente menor quando comparado com os modelos TIP4P, TIP5P e SPC para o sistema octanol úmido. Isso ocorre porque o modelo TIP3P tem uma constante dielétrica menor do que o modelo TIP4P, causando energias livres de hidratação mais baixas e consequentemente $\log K_{ow}$ menores quando comparado ao modelo TIP4P (BRINI et al., 2017; FAN; IORGA; BECKSTEIN, 2020; PARANAHEWAGE; GIERHART; FENNELL, 2016).

Vale ressaltar que os modelos de água comumente utilizados nas simulações de DM são bons em reproduzir as propriedades da água pura. No entanto, eles não são confiáveis em prever as energias livres da hidratação (LUNDBORG; LINDAHL, 2015). O Anexo G traz as propriedades experimental da água em comparação com os modelos teóricos mencionados neste trabalho.

Os modelos de água OPC3 e OPC4 apresentam semelhanças, mas possuem muitas diferenças (ver Anexo G). Por exemplo, o modelo OPC3 tem três locais de interação em vez dos quatro presentes no modelo OPC4. Os comprimentos e ângulos de ligação e as cargas atômicas são diferentes. Todos esses parâmetros influenciam as temperaturas de fusão, a tensão superficial e, consequentemente, a energia livre de hidratação. Portanto, os resultados para esses dois modelos de água diferem significativamente. Uma possível explicação pode estar relacionada à diferença das interações intermoleculares que produzem diferentes tensões superficiais para os dois modelos. Estudos realizados por Tempra *et al.* mostram que o modelo de água OPC4

tem uma tensão superficial mais próxima da experimental do que o modelo de água OPC3 (TEMPRA; OLLILA; JAVANAINEN, 2022).

Para o modelo de água OPC4, o critério de corte de Lennard-Jones usado foi de 1,2 nm, que é o recomendado para modelos lipídicos (TEMPRA; OLLILA; JAVANAINEN, 2022). Os resultados para este modelo têm valores muito semelhantes nos dois campos de força deste estudo (AMBER/GAFF e CHARMM/CGenFF). De fato, os valores de $\log K_{ow}$ encontrados para o modelo de água OPC4 estão em melhor concordância com o valor experimental quando comparados com os demais modelos. Alguns estudos tem demonstrado que o modelo OPC4 apresenta energias de hidratação com maior precisão em comparação com os modelos TIP3P e TIP4P, considerando os valores experimentais (FAN; IORGA; BECKSTEIN, 2020; IZADI; ANANDAKRISHNAN; ONUFRIEV, 2014; TEMPRA; OLLILA; JAVANAINEN, 2022). Na Tabela 02, os valores encontrados com o modelo OPC4 e os demais modelos são consideravelmente diferentes. No entanto, o modelo OPC4 mostra-se mais preciso quando comparado com o valor experimental.

O campo de força AMBER/GAFF e, até certo ponto, o campo de força CHARMM/CGenFF são subpolarizados, o que pode resultar em energias livres de hidratação positivas para pequenas moléculas (SWOPE; HORN; RICE, 2010). Estes campos de força utilizam átomos com cargas fixas em simulações clássicas para calcular propriedades termodinâmicas associadas a processos que envolvem uma mudança no estado de polarização. Por exemplo, pode-se citar entalpias de vaporização e energias livres de solvatação. Fan *et al.* afirmam que o campo de força AMBER/GAFF mostrou alguns erros sistemáticos para moléculas hidrofóbicas, enquanto CHARMM/CGenFF demonstra ser mais preciso e equilibrado. Neste estudo, não foi possível realizar cálculos com os mesmos modelos de água para os dois campos de força. Desta forma, não se pode fazer essa comparação. O único modelo de água em comum nos dois campos de força é a OPC4. Contudo, os resultados possuem energias livres próximas e

coerentes, sendo $19,8 \pm 0,3$ e $20,2 \pm 0,3$ kcal mol⁻¹ para AMBER/GAFF e CHARMM/CGenFF, respectivamente.

O campo de força CHARMM/CGenFF foi desenvolvido com o objetivo de corrigir algumas falhas. Uma destas falhas é o cálculo da tensão superficial, fator crucial para o sistema octanol-água e $\log K_{ow}$ (KLAUDA et al., 2010). No entanto, este campo de força possui uma limitação que é o tratamento aproximado das forças de Lennard-Jones de longo alcance (COSTA et al., 2023).

É interessante mencionar que o cálculo é feito através do desacoplamento da molécula de DPPC do solvente (água, octanol puro ou octanol úmido) em uma caixa de simulação pelo método alquímico. Nesta situação, tensão superficial não influencia, pois não há interface entre fases, existe apenas uma caixa com limites periódicos. De fato, a correção para o coeficiente de partição não vem da tensão superficial, mas sim das interações água-água e água-DPPC não ligadas. Desta forma, a explicação da discrepância entre experimento e simulação também pode vir da energia livre de solvatação, que também afeta a tensão superficial da água. As energias livres da hidratação podem ser calculadas com precisão usando vários métodos de energia livre alquímica. No entanto, a sua precisão depende de diversos fatores, tais como o campo de força, cargas atômicas parciais, parâmetros de Lennard-Jones e a escolha do modelo de água (DENG; ROUX, 2004; MOBLEY et al., 2007). Em síntese, o modelo da água OPC4 estimam com precisão a tensão superficial e as energias livres da solvatação. Neste estudo, o modelo de água OPC4 é aquele que fornece um $\log K_{ow}$ mais próximo do valor experimental (Tabela 01) (SOUZA et al., 2020b). Contudo, é importante destacar que o cálculo obtido através do modelo de água OPC4 ainda é três unidades superior ao valor experimental.

A sobreposição entre as distribuições é indispensável para cálculos precisos de energia livre, e pode ser quantificada pelo cálculo da matriz de sobreposição que pode ser visualizada nos Anexos A e B. De acordo com isso, pode-se observar a

sobreposição nas Figuras A.1 e A.2 para o campo de força AMBER/GAFF e nas Figuras B.1 e B.2 para o campo de força CHARMM/CGenFF ambos para os modelos de água em destaque OPC4.

Para que os estados terminais (0 e 26 para as caixas de água e de 0 a 40 para as caixas de octanol úmido) sejam interligados, a matriz de sobreposição deve ser no mínimo tridiagonal, ou seja, todos os pares de estados adjacentes devem ter sobreposição substancial e nenhum elemento deve ser zero na diagonal principal e nas primeiras diagonais acima e abaixo da principal. Esses elementos diferentes de zero devem ser apreciavelmente acima de 0,03. Qualquer número abaixo de 0,03 deve servir como um sinal de alerta, pois os métodos baseados em FEP tendem a subestimar a variação quando a sobreposição do espaço de fase é tão baixa.

A probabilidade de observar, por exemplo (ANEXO A – Figura A.1), uma amostra coletada do estado 11 vinda de um estado de simulação 10 é de 0,37. Neste caso, a probabilidade de encontrar um microestado amostrado do estado 11 no estado 10 é 0,37. As matrizes de sobreposição do campo de força CHARMM/CGenFF apresentam algumas falhas, entretanto, não afetaram de forma significativa a amostragem, já que temos valores muito próximos para o modelo de água em comum (OPC4) em ambos os campos de força.

5.2.2. Estimador TI

A Tabela 03 mostra os resultados para o estimador TI usando o sistema com octanol puro e úmido com seis modelos de água e com os campos de força AMBER/GAFF e CHARMM/CGenFF.

Tabela 03 - Energia livre de Gibbs de transferência (ΔG_{ow}) do DPPC da fase de octanol (puro e úmido) para a fase aquosa e o coeficiente de partição octanol-água ($\log K_{ow}$) utilizando a integração termodinâmica (estimador TI) com campos de força AMBER/GAFF e CHARMM/CGenFF.

Método	Campo de força	Fase do octanol	Modelo de água	ΔG_{ow} (kcal mol ⁻¹)	$\log K_{ow}$
TI	AMBER	Puro	TIP3P	-20,8±0,6	15,2±0,6
			TIP4P	-19,8±0,6	14,5±0,6
			TIP5P	-29,8±0,6	21,8±0,6
			SPC216	-21,7±0,7	15,9±0,7
			OPC4	-11,1±0,9	8,1±0,9
		Úmido	TIP3P	-23,2±0,6	17,3±0,6
			TIP4P	-27,8±0,5	20,4±0,5
			TIP5P	-31,2±0,5	22,9±0,5
			SPC216	-28,0±0,6	20,5±0,6
			OPC4	-23,2±0,7	17,0±0,7
	CHARMM	Puro	OPC3	-26,4±0,6	19,4±0,6
			OPC4	-23,2±0,6	16,9±0,6
		Úmido	OPC3	-24,7±0,6	18,1±0,6
			OPC4	-21,9±0,5	16,0±0,5

O método TI é um dos poucos estimadores que requerem o cálculo de $\frac{\partial U(\lambda, q)}{\partial \lambda}$. No entanto, calcular esta derivada pode trazer alguns problemas e limitações. Além disso, alguns códigos de simulação não suportam a avaliação dessa quantidade.

Os resultados do estimador TI para ΔG_{ow} e de $\log K_{ow}$ com os dois campos de força e os seis modelos de água são maiores e se distanciam bastante dos resultados do resultado experimental quando comparados aos dados utilizando o estimador MBAR (Tabela 02). Embora o TI seja um dos métodos de energia livre mais simples de utilizar e analisar, ele apresenta alguns problemas que precisam ser

cuidadosamente evitados. Por exemplo, se a curvatura de $\frac{dU}{d\lambda}$ for muito grande, o viés introduzido pelos estados λ discretos torna-se relevante. Em vista disso, ao fazer o uso do estimador TI, é importante verificar se dados suficientes de um número de estados foram coletados para que a energia livre se torne independente do número de estados com precisão estatística.

Outro problema se dá pela avaliação de $\frac{dU}{d\lambda}$ no $\lambda=0$ quando são utilizados os potenciais de Lennard-Jones. Para evitar o problema pode-se usar um caminho linear. No entanto, a transformação linear sempre terá um potencial infinito em $\lambda=0$ o que leva instabilidade ao calcular $\frac{dU}{d\lambda}$.

Mesmo diante dos problemas citados acima o modelo de água OPC4 ainda apresenta o menor valor de $\log K_{ow}$ quando comparado aos demais resultados apresentados na Tabela 03. O ΔG_{ow} calculado é $-23,2 \pm 0,7$ e $\log K_{ow}$ igual a $17,0 \pm 0,7$ para o campo de força AMBER/GAFF. Para o campo de força CHARMM/CGenFF, o ΔG_{ow} é $-21,9 \pm 0,5$, e conseqüentemente, o $\log K_{ow}$ calculado é $16,0 \pm 0,5$. No entanto, existem alguns problemas relacionados à falta de forças de longo alcance atraentes de van der Waals causadas pelo truncamento do potencial de Lennard Jones.

Nos anexos C e D, estão dispostos os gráficos de barras das diferenças de energia livre avaliadas entre pares de estados adjacentes por meio de vários estimadores disponíveis na ferramenta `alchemical_analyses.py`. Além disso, as estimativas de erro correspondentes para cada um dos estimadores são apresentadas. Os resultados de TI estão representados pela cor roxa. Em todos os gráficos os erros do estimador TI são pequenos, assim como os erros do estimador MBAR (em azul). No entanto, eles diferem consideravelmente nas barras das diferenças de energia livre avaliadas entre pares de estados adjacentes. Este fato ocasionou a diferença entre os valores calculados pelos dois estimadores.

5.3. Amostragem guarda-chuva

A energia livre de Gibbs para transferir DPPC do octanol para a fase aquosa (ΔG_{ow}), e vice-versa, foi avaliada usando amostragem guarda-chuva. O método de amostragem guarda-chuva é um padrão ouro de amostragem aprimorada utilizado para calcular ΔG e, consequentemente, $\log K_{ow}$. Para este método, apenas o campo de força CHARMM/CGenFF e o modelo de água TIP3P são usados para efeito de demonstração. Os resultados utilizando este método estão expostos na Tabela 04 a seguir.

Tabela 04 - Energia livre de Gibbs de transferência (ΔG_{ow}) e coeficiente de partição octanol-água ($\log K_{ow}$) utilizando a amostragem guarda-chuva com campo de força CHARMM/CGenFF.

Método	Campo de força	Modelo de água	Sistema	ΔG (kcal mol ⁻¹)	$\log K_{ow}$
US	CHARMM	TIP3P	Direto	27,0±0,4	20,0±0,4
			Inverso	-21,4±0,5	15,0±0,5

O ΔG do sistema direto foi calculado em $27,03 \pm 0,4$ kcal mol⁻¹ usando o perfil de energia mostrado na Figura 07. Aplicando este valor de ΔG na Equação 01, o $\log K_{ow}$ calculado é de $20,01 \pm 0,4$ que está muito distante do valor experimental apresentado na Tabela 01.

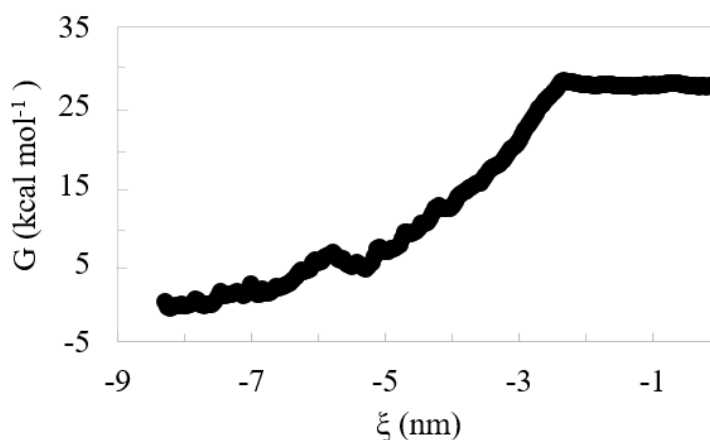


Figura 07 - Gráfico do perfil de energia livre de Gibbs da transferência (ΔG_{ow}) de DPPC do octanol para a fase aquosa usando amostragem guarda-chuva do sistema direto, que corresponde a uma caixa de água entre duas caixas de octanol.

É importante mencionar que a molécula DPPC no processo de transferência entre as fases arrasta muitas moléculas de octanol ligadas a ele como mostra a Figura 08.

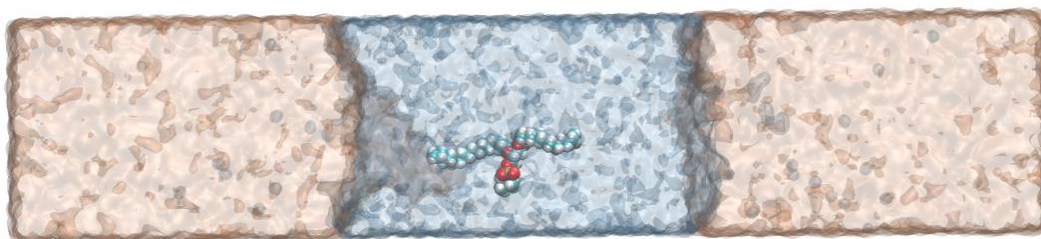


Figura 08 - Arraste de moléculas de octanol presas na cauda hidrofóbica do DPPC no sistema direto. Neste sistema o DPPC é transferido da fase octanol (em laranja) para a fase aquosa (em azul).

A Figura 08 mostra que as moléculas de octanol aderem à molécula de DPPC mesmo depois de ser transferido para a fase aquosa. Supõe-se que este artefato tenha influenciado no cálculo da energia livre, e consequentemente, causa a superestimação do $\log K_{ow}$ encontrado. Tentou-se reduzir esse artefato aumentando o tempo de simulação, mas é inviável obter resultados razoáveis diante do alto custo computacional. Visto que, os tempos de equilíbrio e amostragem necessários para calcular uma energia livre razoável são muito grandes ($\gg 200$ ns), de modo que a amostragem guarda-chuva não é viável para produzir cálculos de energia livre rápidos e confiáveis sem artefatos para o sistema em estudo.

A Figura 09 apresenta a análise WHAM e a sobreposição entre as janelas guarda-chuva do sistema em estudo referente à transferência de DPPC da fase octanol para a fase aquosa. O histograma é razoavelmente sobreposto e mostra dois picos duplos entre a região de -7 a -5 nm (interface octanol-água). A amostragem na região da interface costuma ser mais difícil do que em outras regiões, porém há uma boa sobreposição entre as janelas.

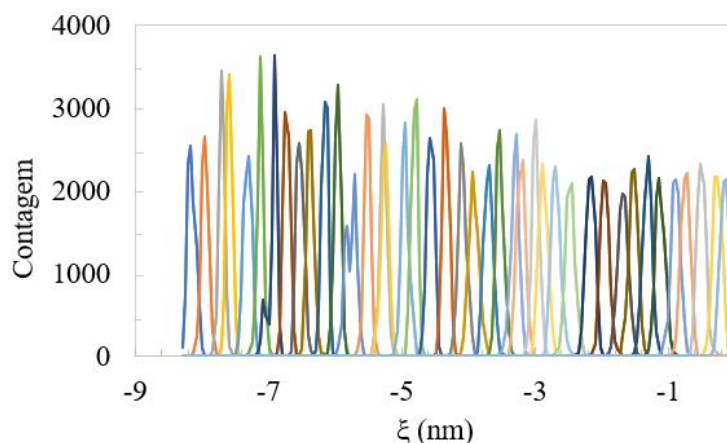


Figura 09 - Histograma de janelas de amostragem de guarda-chuva se sobrepõem na coordenada de reação ξ para o sistema direto, que corresponde a uma caixa de água entre duas caixas de octanol. Neste sistema o DPPC é transferido do octanol para a água.

Para tentar reduzir ou até mesmo evitar o arraste as moléculas, tentou-se realizar o inverso do sistema mencionado anteriormente. Neste caso, o sistema contém uma caixa de octanol entre duas caixas de água como mostra a Figura 06. Para o sistema inverso, o ΔG calculado é $-21,4 \pm 0,5$ kcal mol⁻¹, utilizando o perfil de energia mostrado na Figura 10. Logo, o $\log K_{ow}$ encontrado é de $15,0 \pm 0,5$ (Tabela 04). É possível observar que o sistema inverso possui um $\log K_{ow}$ mais próximo do experimental (Tabela 01). No entanto, ainda continua muito distante do valor experimental. Pode-se observar novamente o arraste de moléculas de uma fase para outra, que pode explicar o desacordo com o valor experimental. As moléculas de água foram arrastadas pela molécula de DPPC até a caixa de octanol. Desta vez, presas na cabeça hidrofílica da molécula de DPPC (Figura 11). Dessa forma, o valor do cálculo da energia livre, e consequentemente, a estimativa de $\log K_{ow}$ podem ter sido afetados novamente.

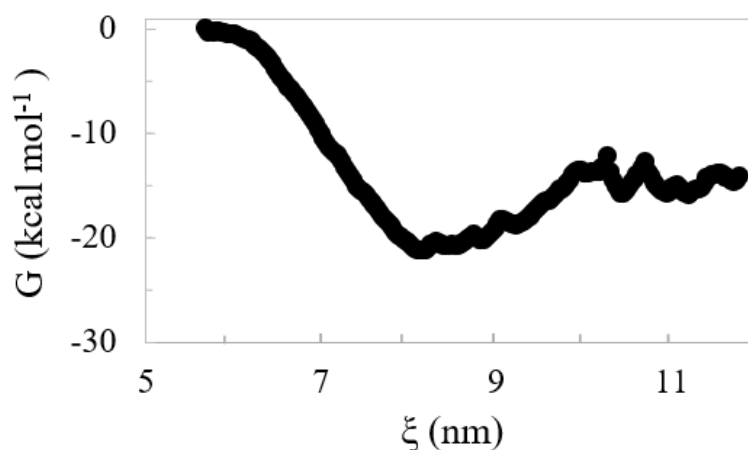


Figura 10 - Gráfico do perfil de energia livre de Gibbs da transferência (ΔG_{ow}) de DPPC da fase aquosa para o octanol usando amostragem guarda-chuva do sistema inverso, que corresponde a uma caixa de octanol entre duas caixas de água.

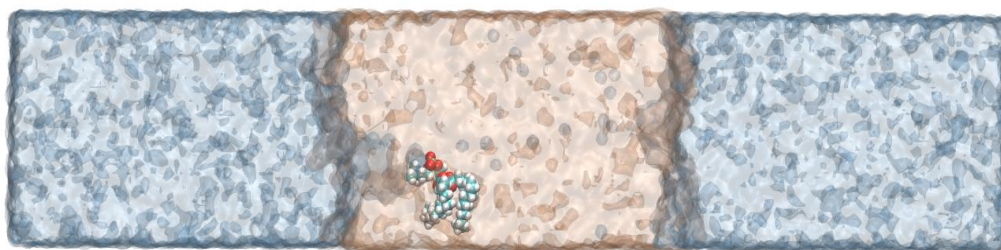


Figura 11 - Arraste de moléculas de água presas na cabeça hidrofílica do DPPC no sistema inverso. Neste sistema o DPPC é transferido da fase aquosa (em azul) para a fase octanol (em laranja).

Isso mostra uma desvantagem para aplicar a amostragem guarda-chuva em sistemas como estes, embora seja bastante conhecido e aplicado por ser um dos melhores métodos de amostragem aprimorados. Figura 12 mostra as sobreposições entre as janelas guarda-chuva do sistema em estudo referente à transferência de DPPC octanol. É observada uma razoavelmente sobreposição mesmo na interface octanol-água.

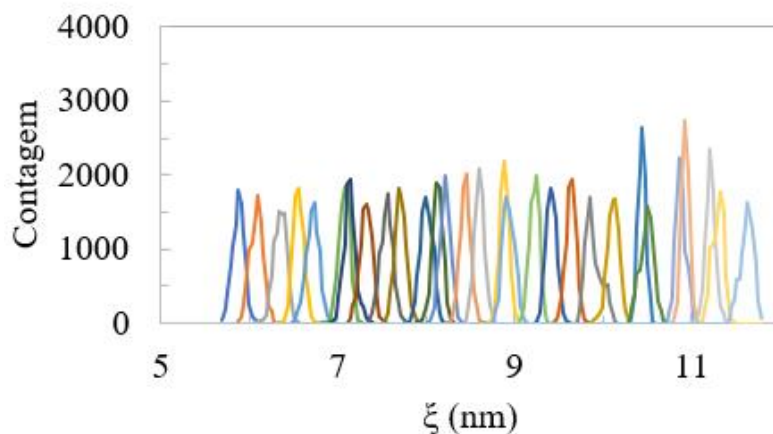


Figura 12 - Histograma de janelas de amostragem de guarda-chuva se sobrepõem na coordenada de reação ξ para o sistema inverso, que corresponde a uma caixa de octanol entre duas caixas de água. Neste sistema o DPPC é transferido da fase aquosa para a fase octanol.

É muito importante também discutir se o método de agitação lenta é confiável para calcular o $\log K_{ow}$ de moléculas anfífilas e zwitteriônicas. A parte experimental dessa propriedade é muito delicada, principalmente porque os fosfolípidios podem ser hidrolisados quando permanecem nas fases octanol-água por um longo período, como no método de agitação lenta. As moléculas de DPPC também podem permanecer na interface, conforme foi observado nas simulações de amostragem guarda-chuva. Dessa forma, se a tensão superficial do lipídio for medida após adição à água, a tensão superficial diminuirá com a adsorção na superfície. E quando a superfície estiver saturada, o monômero também satura a fase aquosa. Outro problema evidente é que o DPPC pode formar agregados de tamanhos diferentes na fase aquosa. Neste contexto, é um grande problema definir/prever a solubilidade da molécula de DPPC em água.

Outro desafio pode estar relacionado ao método analítico utilizados nessas medições. Como o DPPC é muito solúvel em octanol, a quantidade de DPPC solúvel em água será tão pequena que a concentração de DPPC deve estar abaixo do limite de detecção para a maioria das análises de espectroscopia de massa. Portanto, haverá a necessidade de aumentar a concentração da amostra ou concentrar a solução aquosa antes da análise por espectroscopia de massa. A

formação de agregados na água pode ocorrer a curto prazo, e pode haver agregados na fase aquosa ao concentrar a amostra para medir a concentração via espectroscopia de massa, por exemplo. Além disso, haverá uma concentração de DPPC não desprezível na interface octanol-água ou na superfície da vidraria que não é mensurável com precisão com metodologias analíticas confiáveis.

6. Implicações

Este trabalho demonstra que os modelos de água que são largamente utilizados nas simulações de DM falham em calcular com precisão energias livres de hidratação. Embora, as energias livres de hidratação possam ser calculadas através diversos métodos, a sua precisão depende de diversos fatores, como o campo de força, cargas atômicas parciais, parâmetros de Lennard-Jones e a escolha do modelo de água. O estudo reforça a importância do desenvolvimento de novos modelos de água que reproduzam com precisão todas as suas características da água experimental para conciliar a discrepância que existe entre medições experimentais e teóricas de moléculas anfifílicas.

7. Conclusões

Nesta dissertação, as simulações de dinâmica molecular atomística utilizando dois campos de força (AMBER/GAFF e CHARMM/CGenFF) e seis modelos de água (SPC, TIP3P, TIP4P, TIP5P, OPC3 e OPC4) foram realizadas com sucesso. Dois métodos de cálculo de energia livre foram utilizados: o método alquímico e o método de amostragem guarda-chuva.

Para o método alquímico utilizaram-se dois estimadores de energia livre: TI e MBAR. Também, octanol foi utilizado de duas formas diferentes: puro e úmido. Os resultados para a fase de octanol puro não foram analisados por ser uma situação idealizada, que não condiz com o que acontece experimentalmente. Dentre os cálculos de $\log K_{ow}$ feitos com diferentes modelos de água utilizados, quase todos os $\log K_{ow}$ calculados se distanciaram do valor experimental, $11,50 \pm 1,43$, que foi obtido através do método de agitação lenta e usado como referência neste estudo.

O estimador MBAR tem energias livres, e por consequência, coeficientes de partição octanol-água menores, quando comparados ao estimador TI. Este fato pode ser explicado devido ao estimador TI possuir alguns problemas e limitações. No entanto, os coeficientes de partição octanol-água calculado com o modelo de água OPC4 (que estima com precisão a tensão superficial) apresenta os melhores resultados nos dois estimadores. Para o estimador MBAR, o $\log K_{ow}$ é calculado em $14,5 \pm 0,3$ e $14,8 \pm 0,3$ para os campos de força AMBER/GAFF e CHARMM/CGenFF, respectivamente. Para o estimador TI, o $\log K_{ow}$ é de $17,0 \pm 0,7$ e $16,0 \pm 0,5$ para os campos de força AMBER/GAFF e CHARMM/CGenFF, respectivamente. Embora sejam valores muito distantes do experimental, foram os menores valores quando comparados aos demais para estes estimadores.

O método de amostragem guarda-chuva mostrou-se ineficaz para o sistema em estudo. Isto se deve ao arraste de moléculas de água ou octanol pelo DPPC durante as simulações gerando artefatos. O sistema inverso (sistema com uma caixa de octanol entre duas caixas de água, onde o DPPC é transferido da fase aquosa para a fase octanol)

apresentou menos arraste, fornecendo um $\log K_{ow}$ igual a $15,0 \pm 0,5$, porém apresenta um valor calculado ainda distante do valor experimental de $(11,50 \pm 1,43)$.

Por fim, é de grande importância o desenvolvimento de novos modelos de água que reproduzam com precisão todas as suas propriedades. A diminuição da discrepância entre os valores teóricos e experimentais torna o cálculo do $\log K_{ow}$ mais confiável, que é uma propriedade de grande importância em diversas áreas de aplicações científicas e industriais.

8. Perspectivas Futuras

Para uma possível continuação deste trabalho sugere-se:

- Realizar um novo estudo experimental para determinar o coeficiente de partição octanol-água do DPPC.
- Também realizar um estudo similar para outros fosfolipídios importantes tais como o DPPA, DPPE, DPPG, DPPI e DPPS;
- Utilizar outros métodos para prever energias livres tal como o método replica exchange;
- Testar outros campos de força;
- Usar ou desenvolver outros modelos de água.

9. Referências Bibliográficas

- ADCOCK, S. A.; MCCAMMON, J. A. Molecular dynamics: Survey of methods for simulating the activity of proteins. **Chemical Reviews**, v. 106, n. 5, p. 1589–1615, 2006.
- ALAMI, E. O.; HOLMBERG, K. Heterogemini surfactants. **Advances in Colloid and Interface Science**, v. 100–102, n. SUPPL., p. 13–46, 2003.
- ALICE, M.; MELLO, R. DE. Revisão Uso da fosfatidilcolina no tratamento de depósitos localizados de gordura The effect of phosphatidylcholine in the treatment of localized fat pads. v. 8, n. 19, p. 34–37, 2009.
- ALLEN, M. P. . D. J. T. **Computer Simulation of Liquids**. 2. ed. New York: [s.n.].
- ALLOUCHE, A. Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. **Journal of computational chemistry**, v. 32, p. 174–182, 2012.
- ALVES, V. et al. Quimioinformática: Uma Introdução. **Química Nova**, v. 41, n. 2, p. 202–212, 2017.
- AMORIM, F. G. et al. Bioquímica Clínica Da Aterosclerose Provocada Por Hiperhomocisteinemia. **Revista Eletrônica de Farmácia**, v. 8, n. 1, 2011.
- ARNOTT, J. A.; PLANEY, S. L. The influence of lipophilicity in drug discovery and design. **Expert Opinion on Drug Discovery**, v. 7, n. 10, p. 863–875, 2012.
- BAER, B. et al. New insights into exogenous surfactant as a carrier of pulmonary therapeutics. **Biochemical Pharmacology**, v. 164, n. 2, p. 64–73, 2019.
- BANAT, I. M. et al. Microbial biosurfactants production, applications and future potential. **Applied Microbiology and Biotechnology**, v. 87, n. 2, p. 427–444, 2010.
- BAYLY, C. I. et al. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. **Journal of the American Chemical Society**, v. 117, n. 19, p. 5179–5197, 1995.
- BENEDETTO, F. E.; ZOLOTUCHO, H.; PRADO, M. O. Critical assessment of the surface tension determined by the maximum pressure

- bubble method. **Materials Research**, v. 18, n. 1, p. 9–14, 2015.
- BENNETT, C. H. Efficient estimation of free energy differences from Monte Carlo data. **Journal of Computational Physics**, v. 22, n. 2, p. 245–268, 1976.
- BERENDSEN, H. J. C. et al. Molecular dynamics with coupling to an external bath. **The Journal of Chemical Physics**, v. 81, n. 8, p. 3684–3690, 1984.
- BERENDSEN, H. J. C.; GRIGERA, J. R.; STRAATSMA, T. P. The missing term in effective pair potentials. **Journal of Physical Chemistry**, v. 91, n. 24, p. 6269–6271, 1987.
- BERKOWITZ, M. L.; BOSTICK, D. L.; PANDIT, S. Aqueous solutions next to phospholipid membrane surfaces: Insights from simulations. **Chemical Reviews**, v. 106, n. 4, p. 1527–1539, 2006.
- BERKOWITZ, M. L.; VÁCHA, R. Aqueous solutions at the interface with phospholipid bilayers. **Accounts of Chemical Research**, v. 45, n. 1, p. 74–82, 2012.
- BINDER, K. et al. Molecular dynamics simulations. **Journal of Physics Condensed Matter**, v. 16, n. 5, 2004.
- BLANCO, O.; PÉREZ-GIL, J. Biochemical and pharmacological differences between preparations of exogenous natural surfactant used to treat Respiratory Distress Syndrome: Role of the different components in an efficient pulmonary surfactant. **European Journal of Pharmacology**, v. 568, n. 1–3, p. 1–15, 2007.
- BOYCE, S. E. et al. Predicting Ligand Binding Affinity with Alchemical Free Energy Methods in a Polar Model Binding Site. **Journal of Molecular Biology**, v. 394, n. 4, p. 747–763, dez. 2009.
- BRANDT, E. B. et al. Surfactant protein D alters allergic lung responses in mice and human subjects. **Journal of Allergy and Clinical Immunology**, v. 121, n. 5, 2008.
- BRINI, E. et al. How Water's Properties Are Encoded in Its Molecular Structure and Energies. **Chemical Reviews**, v. 117, n. 19, p. 12385–12414, 2017.
- BROOKS, B. R. et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. **Journal of Computational**

Chemistry, v. 4, n. 2, p. 187–217, 1983.

BROWN, R. D. ET AL. Tools for designing diverse, drug-like, cost-effective combinatorial libraries”; in *Combinatorial Library Design and Evaluation*. **Marcel Dekker**, p. 328, 2001.

BUSSI, G.; DONADIO, D.; PARRINELLO, M. Canonical sampling through velocity rescaling. **Journal of Chemical Physics**, v. 126, n. 1, 2007.

CENSE, J. M. et al. Neural networks prediction of partition coefficients. **Chemometrics and Intelligent Laboratory Systems**, v. 23, n. 2, p. 301–308, 1994.

CHENG, T. et al. Computation of octanol-water partition coefficients by guiding an additive model with knowledge. **Journal of Chemical Information and Modeling**, v. 47, n. 6, p. 2140–2148, 2007.

CHOI, Y.; JANG, J.; PARK, H. S. Pulmonary Surfactants: a New Therapeutic Target in Asthma. **Current Allergy and Asthma Reports**, v. 20, n. 11, 2020.

CORDOMÍ, A.; CALTABIANO, G.; PARDO, L. Membrane protein simulations using AMBER force field and Berger lipid parameters. **Journal of Chemical Theory and Computation**, v. 8, n. 3, p. 948–958, 2012.

COSTA, R. K. M. et al. The reconciliation between the experimental and calculated octanol-water partition coefficient of 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine using atomistic molecular dynamics: an open question. **Journal of Biomolecular Structure and Dynamics**, v. 1, n. 0, p. 1–8, 2023.

CRAMER, C. J. **Essentials of Computational Chemistry: Theories and Models**. 2. ed. USA: John Wiley & Sons Ltd, 2004.

CUMMING, H.; RÜCKER, C. Octanol-Water Partition Coefficient Measurement by a Simple ¹H NMR Method. **ACS Omega**, v. 2, n. 9, p. 6244–6249, 2017.

DAINA, A.; MICHIELIN, O.; ZOETE, V. iLOGP: A Simple, Robust, and Efficient Description of n -Octanol/Water Partition Coefficient for Drug Design Using the GB/SA Approach. **Journal of Chemical Information and Modeling**, v. 54, n. 12, p. 3284–3301, 22 dez. 2014.

DAINA, A.; MICHIELIN, O.; ZOETE, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry

friendliness of small molecules. **Scientific Reports**, v. 7, n. January, p. 1–13, 2017.

DE JONG, D. H. et al. Improved parameters for the martini coarse-grained protein force field. **Journal of Chemical Theory and Computation**, v. 9, n. 1, p. 687–697, 2013.

DENG, Y.; ROUX, B. Hydration of amino acid side chains: Nonpolar and electrostatic contributions calculated from staged molecular dynamics free energy simulations with explicit water molecules. **Journal of Physical Chemistry B**, v. 108, n. 42, p. 16567–16576, 2004.

DICKSON, C. J. et al. GAFFlipid: A General Amber Force Field for the accurate molecular dynamics simulation of phospholipid. **Soft Matter**, v. 8, n. 37, p. 9617–9627, 2012.

DOWHAN, W. Molecular basis for membrane phospholipid diversity: Why are there so many lipids? **Annual Review of Biochemistry**, v. 66, n. 12, p. 199–232, 1997.

DUAN, Y. et al. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. **Journal of Computational Chemistry**, v. 24, n. 16, p. 1999–2012, 2003.

DUBBELDAM, D. et al. United atom force field for alkanes in nanoporous materials. **Journal of Physical Chemistry B**, v. 108, n. 33, p. 12301–12313, 2004.

EASTOE, J. et al. Dynamic surface tensions of nonionic surfactant solutions. **Journal of Colloid and Interface Science**, v. 188, n. 2, p. 423–430, 1997.

EDWARD, J. T.; CHUBB, F. L.; SANGSTER, J. Iron chelators of the pyridoxal isonicotinoyl hydrazone class. Relationship of the lipophilicity of the apochelator to its ability to mobilize iron from reticulocytes in vitro: Reappraisal of reported partition coefficients. **Canadian Journal of Physiology and Pharmacology**, v. 75, n. 12, p. 1362–1368, 1997.

ESSMANN, U. et al. A smooth particle mesh Ewald method. **The Journal of Chemical Physics**, v. 103, n. 19, p. 8577–8593, 1995.

FABIANA MEIJON FADUL. Calculation of Hydrophobic Constant (Log P) from π and f Constantst. **Journal of Medicinal Chemistry**, v. 18, n. 9, p.

865–868, 2019.

FAN, S.; IORGA, B. I.; BECKSTEIN, O. Prediction of octanol-water partition coefficients for the SAMPL6- log P molecules using molecular dynamics simulations with OPLS-AA, AMBER and CHARMM force fields. **Journal of Computer-Aided Molecular Design**, v. 34, n. 5, p. 543–560, 2020.

FERREIRA, L. L. G.; ANDRICOPULO, A. D. ADMET modeling approaches in drug discovery. **Drug Discovery Today**, v. 24, n. 5, p. 1157–1165, 2019.

FINIZIO, A.; VIGHI, M.; SANDRONI, D. Determination of n-octanol/water partition coefficient (Kow) of pesticide critical review and comparison of methods. **Chemosphere**, v. 34, n. 1, p. 131–161, 1997.

FORNASIER, F. et al. Lipophilicity of Coarse-Grained Cholesterol Models. **Journal of Chemical Information and Modeling**, v. 60, n. 2, p. 569–577, 2020.

FRENKEL, D.; SMIT, B.; RATNER, M. A. Understanding Molecular Simulation: From Algorithms to Applications. **Physics Today**, v. 50, n. 7, p. 66–66, jul. 1997.

FRIBERG, S.; JANSSON, P. O.; CEDERBERG, E. Surfactant association structure and emulsion stability. **Journal of Colloid And Interface Science**, v. 55, n. 3, p. 614–623, 1976.

FUENTES-AZCATL, R.; ALEJANDRE, J. Non-polarizable force field of water based on the dielectric constant: TIP4P/ε. **Journal of Physical Chemistry B**, v. 118, n. 5, p. 1263–1272, 2014.

GABRIELI, A. et al. High-temperature dynamic behavior in bulk liquid water: A molecular dynamics simulation study using the OPC and TIP4P-Ew potentials. **Frontiers of Physics**, v. 13, n. 1, p. 1–15, 2018.

GARRIDO, N. M. et al. 1-Octanol/Water Partition Coefficients of n - Alkanes from Molecular Simulations of Absolute Solvation Free Energies. **Journal of Chemical Theory and Computation**, v. 5, n. 9, p. 2436–2446, 8 set. 2009.

GHOSE, A. K.; VISWANADHAN, V. N.; WENDOLOSKI, J. J. Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: An analysis of ALOGP and CLOGP methods.

Journal of Physical Chemistry A, v. 102, n. 21, p. 3762–3772, 1998.

GREG LANDRUM, PAOLO TOSCO, BRIAN KELLEY, RIC, SRINIKER, GEDECK, RICCARDO VIANELLO, NADINESCHNEIDER, EISUKE KAWASHIMA, DAVID COSGROVE, ANDREW DALKE, DAN N, GARETH JONES, BRIAN COLE, MATT SWAIN, SAMO TURK, ALEXANDERSAVELYEV, ALAIN VAUCHER, M. W. rdkit/rdkit. v. rdkit/rdki, 2022.

GUIDELINES, O. **Test Guideline No . 123 Test No . 123: Partition Coefficient (1-Octanol / Water): Slow-Stirring Method.**

HANSSON, T.; OOSTENBRINK, C.; VAN GUNSTEREN, W. F. Molecular dynamics simulations. **Current Opinion in Structural Biology**, v. 12, n. 2, p. 190–196, 2002.

HÉNIN, J.; SHINODA, W.; KLEIN, M. L. United-atom acyl chains for CHARMM phospholipids. **Journal of Physical Chemistry B**, v. 112, n. 23, p. 7008–7015, 2008.

HENTSCHEL, R. et al. Surfactant replacement therapy: from biological basis to current clinical practice. **Pediatric Research**, v. 88, n. 2, p. 176–183, 2020.

HESS, B. et al. LINCS: A Linear Constraint Solver for molecular simulations. **Journal of Computational Chemistry**, v. 18, n. 12, p. 1463–1472, 1997.

HEYMAN, B. et al. Shake flask methodology for assessing the influence of the maximum oxygen transfer capacity on 2,3-butanediol production. **Microbial Cell Factories**, v. 18, n. 1, p. 1–16, 2019.

HODGES, G. et al. A comparison of log K_{ow} (n-octanol–water partition coefficient) values for non-ionic, anionic, cationic and amphoteric surfactants determined using predictions and experimental methods. **Environmental Sciences Europe**, v. 31, n. 1, p. 1–18, 2019.

HUANG, J.; MACKERELL, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. **Journal of Computational Chemistry**, v. 34, n. 25, p. 2135–2145, 2013.

HWANG, D.; RAMSEY, J. D.; KABANOV, A. V. Polymeric micelles for the delivery of poorly soluble drugs: From nanoformulation to clinical approval. **Advanced Drug Delivery Reviews**, v. 156, n. 5, p. 80–118, 2020.

- IŞIK, M. et al. **Assessing the accuracy of octanol–water partition coefficient predictions in the SAMPL6 Part II log P Challenge.** [s.l.] Springer International Publishing, 2020a. v. 34
- IŞIK, M. et al. Octanol–water partition coefficient measurements for the SAMPL6 blind prediction challenge. **Journal of Computer-Aided Molecular Design**, v. 34, n. 4, p. 405–420, 2020b.
- IZADI, S.; ANANDAKRISHNAN, R.; ONUFRIEV, A. V. Building Water Models: A Different Approach. **The Journal of Physical Chemistry Letters**, v. 5, n. 21, p. 3863–3871, 6 nov. 2014.
- J. WANG, P. C. AND P. A. K. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? **J. Comput. Chem**, v. 21, n. 2, p. 1049–1074, 2000.
- JAVANAINEN, M. et al. Atomistic Model for Nearly Quantitative Simulations of Langmuir Monolayers. **Langmuir**, v. 34, n. 7, p. 2565–2572, 2018.
- JORGENSEN, W. L. et al. Comparison of simple potential functions for simulating liquid water. **The Journal of Chemical Physics**, v. 79, n. 2, p. 926–935, 1983.
- KAMENIK, A. S. et al. Polarizable and non-polarizable force fields: Protein folding, unfolding, and misfolding. **The Journal of Chemical Physics**, v. 153, n. 18, p. 185102, 14 nov. 2020.
- KAMINSKI, G. et al. = $\beta + \sqrt{2}$. p. 13077–13082, 1994.
- KHALAK, Y.; BAUMEIER, B.; KARTTUNEN, M. Improved general-purpose five-point model for water: TIP5P/2018. **Journal of Chemical Physics**, v. 149, n. 22, 2018.
- KHANDWALA, A. S.; KASPER, C. B. The fatty acid composition of individual phospholipids from rat liver nuclear membrane and nuclei. **Journal of Biological Chemistry**, v. 246, n. 20, p. 6242–6246, 1971.
- KIM, I.; ALLEN, T. W. Bennetts acceptance ratio and histogram analysis methods enhanced by umbrella sampling along a reaction coordinate in configurational space. **Journal of Chemical Physics**, v. 136, n. 16, 2012.
- KLAUDA, J. B. et al. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. **Journal of Physical**

Chemistry B, v. 114, n. 23, p. 7830–7843, 2010.

KLIMOVICH, P. V.; SHIRTS, M. R.; MOBLEY, D. L. Guidelines for the analysis of free energy calculations. **Journal of Computer-Aided Molecular Design**, v. 29, n. 5, p. 397–411, 2015.

KUMAR, M. M. et al. International Journal of Advances in Pharmacy and Biotechnology ADMET PREDICTORS ARE THE TOOLS FOR THE ENHANCEMENT OF DRUG DESIGN AND DEVELOPMENT: A SYSTEMATIC REVIEW. **I J a P B**, v. 4, n. 4, p. 6–13, 2018.

KUMAR, S. et al. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. **Journal of Computational Chemistry**, v. 13, n. 8, p. 1011–1021, out. 1992.

LAGORCE, D. et al. FAF-Drugs2: Free ADME/tox filtering tool to assist drug discovery and chemical biology projects. **BMC Bioinformatics**, v. 9, p. 1–9, 2008.

LANG, B. E. Solubility of water in octan-1-ol from (275 to 369) K. **Journal of Chemical and Engineering Data**, v. 57, n. 8, p. 2221–2226, 2012.

LEACH, A. R. L. A. **Molecular Modelling: Principles and Applications**. [s.l: s.n.].

LEE, J. et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. **Journal of Chemical Theory and Computation**, v. 12, n. 1, p. 405–413, 2016.

LEE, S. et al. CHARMM36 united atom chain model for lipids and surfactants. **Journal of Physical Chemistry B**, v. 118, n. 2, p. 547–556, 2014.

LEEKUMJORN, S.; SUM, A. K. Molecular simulation study of structural and dynamic properties of mixed DPPC/DPPE bilayers. **Biophysical Journal**, v. 90, n. 11, p. 3951–3965, 2006.

LI, Z. et al. The ratio of phosphatidylcholine to phosphatidylethanolamine influences membrane integrity and steatohepatitis. **Cell Metabolism**, v. 3, n. 5, p. 321–331, 2006.

LIPOWSKY, R. The morphology of lipid membranes. **Current Opinion in Structural Biology**, v. 5, n. 4, p. 531–540, 1995.

LIU, Y. et al. Constructing ab initio force fields for molecular dynamics

simulations Constructing ab initio force fields for molecular dynamics simulations. v. 4739, n. 1998, 2013.

LOPEZ-RODRIGUEZ, E. et al. Lung surfactant metabolism: early in life, early in disease and target in cell therapy. **Cell and Tissue Research**, v. 367, n. 3, p. 721–735, 2017.

LUNDBORG, M.; LINDAHL, E. Automatic GROMACS topology generation and comparisons of force fields for solvation free energy calculations. **Journal of Physical Chemistry B**, v. 119, n. 3, p. 810–823, 2015.

LYUBARTSEV, A. P.; RABINOVICH, A. L. Force Field Development for Lipid Membrane Simulations. **Biochimica et Biophysica Acta - Biomembranes**, v. 1858, n. 10, p. 2483–2497, 2016.

MACHATHA, S. G.; YALKOWSKY, S. H. Comparison of the octanol/water partition coefficients calculated by ClogP®, ACDlogP and KowWin® to experimentally determined values. **International Journal of Pharmaceutics**, v. 294, n. 1–2, p. 185–192, 2005.

MACKERELL, A. D.; FEIG, M.; BROOKS, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation. **Journal of Computational Chemistry**, v. 25, n. 11, p. 1400–1415, 2004.

MARRINK, S. J. et al. The MARTINI force field: Coarse grained model for biomolecular simulations. **Journal of Physical Chemistry B**, v. 111, n. 27, p. 7812–7824, 2007.

MARTINS, J. P. A.; FERREIRA, M. M. C. QSAR modeling: um novo pacote computacional open source para gerar e validar modelos QSAR. **Química Nova**, v. 36, n. 4, p. 554–560, 2013.

MCDONALD, N. A.; JORGENSEN, W. L. Development of an All-Atom Force Field for Heterocycles. Properties of Liquid Pyrrole, Furan, Diazoles, and Oxazoles. v. 5647, n. 98, p. 8049–8059, 1998.

MENGER, F. M. ACCOUNTS OF CHEMICAL RESEARCH On the Structure of Micelles. **Accounts of Chemical Research**, v. 12, n. 4, p. 111–117, 1979.

MILLER, D. J.; MCWILLIAMS, P. Octanol-water partition coefficients of surfactants: Slow stirring/surface tension method. **Tenside, Surfactants**,

Detergents, v. 47, n. 1, p. 28–33, 2010.

MOBLEY, D. L. et al. Comparison of charge models for fixed-charge force fields: Small-molecule hydration free energies in explicit solvent. **Journal of Physical Chemistry B**, v. 111, n. 9, p. 2242–2254, 2007.

MORIGUCHI, I. et al. Simple Method of Calculating Octanol/Water Partition Coefficient. **Chemical and Pharmaceutical Bulletin**, v. 40, n. 1, p. 127–130, 1992.

MORIGUCHI, I. et al. Comparison of Reliability of log P Values for Drugs Calculated by Several Methods. **Chemical and Pharmaceutical Bulletin**, v. 42, n. 4, p. 976–978, 1994.

NAMBA, A. M.; DA SILVA, V. B.; DA SILVA, C. H. T. P. Dinâmica molecular: Teoria e aplicações em planejamento de fármacos. **Ecletica Quimica**, v. 33, n. 4, p. 13–24, 2008.

NIU, S. L.; LITMAN, B. J. Determination of membrane cholesterol partition coefficient using a lipid vesicle-cyclodextrin binary system: Effect of phospholipid acyl chain unsaturation and headgroup composition. **Biophysical Journal**, v. 83, n. 6, p. 3408–3415, 2002.

NKADI, P. O.; MERRITT, T. A.; PILLERS, D. A. M. An overview of pulmonary surfactant in the neonate: Genetics, metabolism, and the role of surfactant in health and disease. **Molecular Genetics and Metabolism**, v. 97, n. 2, p. 95–101, 2009.

OECD/OCDE. Test No. 107: Partition Coefficient (n-octanol/water): Shake Flask Method. **Guidance**, v. 107, n. July, p. 1–4, 1995.

OECD. **OECD117 guideline for testing of chemicals: Partitioning coefficient (n-octanol/water) HPLC method.**

OLIVEI, A. Surface imaging by scanning-laser-induced voltage. **Journal of Physics D: Applied Physics**, v. 9, n. 2, p. 183–195, 1976.

OLIVEIRA, C. H. et al. Surfactantes derivados do fruto de coco (*Cocos nucifera* L .) e sensibilidade cutânea Coconut (*Cocos nucifera* L .) surfactants and skin sensitivity. **Rev. bras. alerg. imunopatol.**, v. 28, n. 3, p. 155–160, 2005.

OLIVEIRA, J. M. DE; MEI, L. H. I. Surfactantes reativos não-iônicos em polimerização em emulsão de látices de acetato de vinila - vinil neodecanoato: influência nas propriedades de barreira à água.

Polímeros, v. 19, n. 1, p. 22–30, 2009.

OOSTENBRINK, C. et al. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1656–1676, 2004.

PAOLETTI, L. et al. Role of Phosphatidylcholine during Neuronal differentiation. **IUBMB Life**, v. 63, n. 9, p. 714–720, 2011.

PARAMETERS, P. Atomic for Structure-Activity Modeling Interactions. **J. Chem. Inf. Comput. Sci.**, v. 35, n. 3, p. 21–35, 1987.

PARANAHEWAGE, S. S.; GIERHART, C. S.; FENNELL, C. J. Predicting water-to-cyclohexane partitioning of the SAMPL5 molecules using dielectric balancing of force fields. **Journal of Computer-Aided Molecular Design**, v. 30, n. 11, p. 1059–1065, 2016.

PARANJAPE, V. V. Randomization of Energy and Momentum in Statistical Mechanics. **Journal of Modern Physics**, v. 06, n. 15, p. 2198–2201, 2015.

PARRA, E.; PÉREZ-GIL, J. Composition, structure and mechanical properties define performance of pulmonary surfactant membranes and films. **Chemistry and Physics of Lipids**, v. 185, n. 4, p. 153–175, jan. 2015.

PARRINELLO, M.; RAHMAN, A. Polymorphic transitions in single crystals: A new molecular dynamics method. **Journal of Applied Physics**, v. 52, n. 12, p. 7182–7190, 1981.

PEREIRA, A. DOS S.; AQUINO NETO, F. R. DE. Estado da arte da cromatografia gasosa de alta resolução e alta temperatura. **Química Nova**, v. 23, n. 3, p. 370–379, 2000.

PICHOT, R.; WATSON, R. L.; NORTON, I. T. Phospholipids at the interface: Current trends and challenges. **International Journal of Molecular Sciences**, v. 14, n. 6, p. 11767–11794, 2013.

PRICE, D. J.; BROOKS, C. L. A modified TIP3P water potential for simulation with Ewald summation. **Journal of Chemical Physics**, v. 121, n. 20, p. 10096–10103, 2004.

RAIGER, L.; LÓPEZ, N. Los biosurfactantes y la industria petrolera. **QuímicaViva**, v. 3, p. 145–209, 2009.

- RIBEIRO, L.; FERREIRA, M. M. C. Artigo. v. 26, n. 3, p. 312–318, 2003.
- RIZZO, R. C.; JORGENSEN, W. L. OPLS All-Atom Model for Amines: Resolution of the Amine Hydration Problem. n. 11, p. 4827–4836, 2016.
- RUDOLPHI-SKÓRSKA, E.; FILEK, M.; ZEMBALA, M. The Effects of the Structure and Composition of the Hydrophobic Parts of Phosphatidylcholine-Containing Systems on Phosphatidylcholine Oxidation by Ozone. **Journal of Membrane Biology**, v. 250, n. 5, p. 493–505, 2017.
- S.K.LEE, I.H.LEE,H.J.KIM, G.S.CHANG, J.E.CHUNG, K. T. N. The PreADME Approach: Web-based program for rapid prediction of physico-chemical, drug absorption and drug-like properties. **EuroQSAR Designing Drugs and Crop Protectants: processes, problems and solutions**, v. 2, n. 03, p. 418–420, 2003.
- SARKAR, A.; KELLOGG, G. Hydrophobicity - Shake Flasks, Protein Folding and Drug Discovery. **Current Topics in Medicinal Chemistry**, v. 10, n. 1, p. 67–83, 2010.
- SHIRTS, M. R. et al. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. **Journal of Chemical Physics**, v. 119, n. 11, p. 5740–5761, 2003.
- SHIRTS, M. R.; CHODERA, J. D. Statistically optimal analysis of samples from multiple equilibrium states. **Journal of Chemical Physics**, v. 129, n. 12, p. 1–10, 2008a.
- SHIRTS, M. R.; CHODERA, J. D. Statistically optimal analysis of samples from multiple equilibrium states. **The Journal of Chemical Physics**, v. 129, n. 12, p. 124105, 28 set. 2008b.
- SHIRTS, M. R.; PANDE, V. S. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. **Journal of Chemical Physics**, v. 122, n. 13, p. 541–558, 2005a.
- SHIRTS, M. R.; PANDE, V. S. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. **Journal of Chemical Physics**, v. 122, n. 13, 2005b.
- SHORT, J. et al. Practical methods for the measurement of logP for surfactants. **Ecotoxicology and Environmental Safety**, v. 73, n. 6, p.

1484–1489, 2010.

SILVA, J. D. F. DA et al. Microemulsions: Components, Characteristics, Potentialities in Food Chemistry and Other Applications. **Química Nova**, v. 38, n. 9, p. 1196–1206, 2015.

SMONDYREV, A. M.; BERKOWITZ, M. L. United atom force field for phospholipid membranes: Constant pressure molecular dynamics simulation of dipalmitoylphosphatidicholine/water system. **Journal of Computational Chemistry**, v. 20, n. 5, p. 531–545, 1999.

SOUSA DA SILVA, A. W.; VRANKEN, W. F. ACPYPE - AnteChamber PYthon Parser interface. **BMC Research Notes**, v. 5, p. 1–8, 2012.

SOUZA, L. M. P. et al. Tuning the hydrophobicity of a coarse grained model of 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine using the experimental octanol-water partition coefficient. **Journal of Molecular Liquids**, v. 319, n. 6, p. 114132, 2020a.

SOUZA, L. M. P. et al. Tuning the hydrophobicity of a coarse grained model of 1,2-dipalmitoyl-sn-glycero-3-phosphatidylcholine using the experimental octanol-water partition coefficient. **Journal of Molecular Liquids**, v. 319, p. 114132, 2020b.

STOLTZ, G. Computation of Free Energy Differences. **Encyclopedia of Applied and Computational Mathematics**, v. 85, n. 2, p. 254–260, 2015a.

STOLTZ, G. Computation of Free Energy Differences. **Encyclopedia of Applied and Computational Mathematics**, v. 85, p. 254–260, 2015b.

STRAATSMA TP, BERENDSEN HJC, P. J. Free energy of hydrophobic hydration: a molecular dynamics study of noble gases in water. **J Chem Phys**, v. 85, n. 11, p. 248–253, 1986.

SUZUKI, T.; KUDO, Y. Automatic log P estimation based on combined additive modeling methods. **Journal of Computer-Aided Molecular Design**, v. 4, n. 2, p. 155–198, 1990.

SWOPE, W. C.; HORN, H. W.; RICE, J. E. Accounting for polarization cost when using fixed charge force fields. I. Method for computing energy. **Journal of Physical Chemistry B**, v. 114, n. 26, p. 8621–8630, 2010.

TAVARES, L. C. QSAR: a abordagem de Hansch. **Química Nova**, v. 27, n. 4, p. 631–639, 2004.

TEMPRA, C.; OLLILA, O. H. S.; JAVANAINEN, M. Accurate Simulations of Lipid Monolayers Require a Water Model with Correct Surface Tension. **Journal of Chemical Theory and Computation**, v. 18, n. 3, p. 1862–1869, 2022.

TETKO, I. V.; TANCHUK, V. Y. Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program. **Journal of Chemical Information and Computer Sciences**, v. 42, n. 5, p. 1136–1145, 2002.

TOBIAS, D. J.; MARTYNA, G. J.; KLEIN, M. L. Molecular dynamics simulations of a protein in the canonical ensemble. **The Journal of Physical Chemistry**, v. 97, n. 49, p. 12959–12966, 1 dez. 1993.

TORRIE, G. M.; VALLEAU, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. **Journal of Computational Physics**, v. 23, n. 2, p. 187–199, fev. 1977.

URAN, S. et al. Analysis of phospholipid species in human blood using normal-phase liquid chromatography coupled with electrospray ionization ion-trap tandem mass spectrometry. **Journal of Chromatography B: Biomedical Sciences and Applications**, v. 758, n. 2, p. 265–275, 2001.

VAN DE WATERBEEMD, H.; GIFFORD, E. ADMET in silico modelling: Towards prediction paradise? **Nature Reviews Drug Discovery**, v. 2, n. 3, p. 192–204, 2003.

VAN DER SPOEL, D. et al. GROMACS: Fast, flexible, and free. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1701–1718, 2005.

VAN GUNSTEREN, W. F.; BERENDSEN, H. J. C. A Leap-frog Algorithm for Stochastic Dynamics. **Molecular Simulation**, v. 1, n. 3, p. 173–185, 3 mar. 1988.

VANOMMESLAEGHE, K.; MACKERELL, A. D. Automation of the CHARMM general force field (CGenFF) I: Bond perception and atom typing. **Journal of Chemical Information and Modeling**, v. 52, n. 12, p. 3144–3154, 2012.

VEATCH, S. L.; KELLER, S. L. Organization in Lipid Membranes Containing Cholesterol. **Physical Review Letters**, v. 89, n. 26, p. 1–4, 2002.

VERLET, L. Computer “Experiments” on Classical Fluids. I.

Thermodynamical Properties of Lennard-Jones Molecules. **Physical Review**, [s.l.], v. 159, n. 1, p. 98–103, 1967.

VILAS-BOAS, S. M. et al. Octanol-Water Partition Coefficients and Aqueous Solubility Data of Monoterpenoids: Experimental, Modeling, and Environmental Distribution. **Industrial and Engineering Chemistry Research**, v. 61, n. 8, p. 3154–3167, 2022.

WANG, J. et al. Development and testing of a general Amber force field. **Journal of Computational Chemistry**, v. 25, n. 9, p. 1157–1174, 2004.

WANG, Z. et al. Predictive deep learning models for environmental properties: The direct calculation of octanol-water partition coefficients from molecular graphs. **Green Chemistry**, v. 21, n. 16, p. 4555–4565, 2019.

WILDMAN, S. A.; CRIPPEN, G. M. Prediction of physicochemical parameters by atomic contributions. **Journal of Chemical Information and Computer Sciences**, v. 39, n. 5, p. 868–873, 1999.

WILLSON, D. F.; NOTTER, R. H. The future of exogenous surfactant therapy. **Respiratory Care**, v. 56, n. 9, p. 1369–1386, 2011.

WODTKE, E. Lipid adaptation in liver mitochondrial membranes of carp acclimated to different environmental temperatures. Phospholipid composition, fatty acid pattern, and cholesterol content. **Biochimica et Biophysica Acta (BBA)/Lipids and Lipid Metabolism**, v. 529, n. 2, p. 280–291, 1978.

XIA, M.; CHAI, Z.; WANG, D. Polarizable and Non-Polarizable Force Field Representations of Ferric Cation and Validations. **Journal of Physical Chemistry B**, v. 121, n. 23, p. 5718–5729, 2017.

ZHOU, M. M. et al. Effects of different fatty acids composition of phosphatidylcholine on brain function of dementia mice induced by scopolamine. **Lipids in Health and Disease**, v. 15, n. 1, p. 1–10, 2016.

ZWANZIG, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. **Journal of Chemical Physics**, v. 22, n. 8, p. 1420–1426, 1954.

10. Anexos

ANEXO A

Matrizes de sobreposições dos λ nos cálculos de energia livre utilizando o método de integração termodinâmica para o modelo de água OPC4 no campo de força AMBER/GAFF.

λ	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0	.91	.08	.01																							
1	.29	.32	.38																							
2	.01	.06	.81	.10	.01																					
3			.37	.35	.26	.03																				
4			.02	.13	.51	.32	.02																			
5				.13	.58	.27	.02																			
6				.01	.20	.53	.23	.03																		
7					.01	.21	.47	.28	.03																	
8						.02	.22	.49	.16	.07	.02	.02														
9							.03	.23	.25	.34	.09	.05														
10								.04	.12	.62	.14	.07														
11								.03	.08	.37	.23	.27	.02													
12								.01	.03	.12	.19	.49	.14													
13										.01	.15	.62	.21													
14											.13	.67	.19													
15												.13	.75	.12												
16													.17	.71	.11	.01										
17														.10	.58	.31	.01									
18															.01	.26	.66	.07								
19																.03	.45	.44	.08							
20																	.01	.15	.66	.18						
21																		.03	.70	.25	.03					
22																			.13	.55	.31					
23																				.32	.62	.02	.03			
24																					.13	.66	.04	.17		
25																						.03	.39	.06	.52	

Figura A.1 – Matriz de sobreposição para a caixa de água com o modelo OPC4 no campo de força AMBER/GAFF.

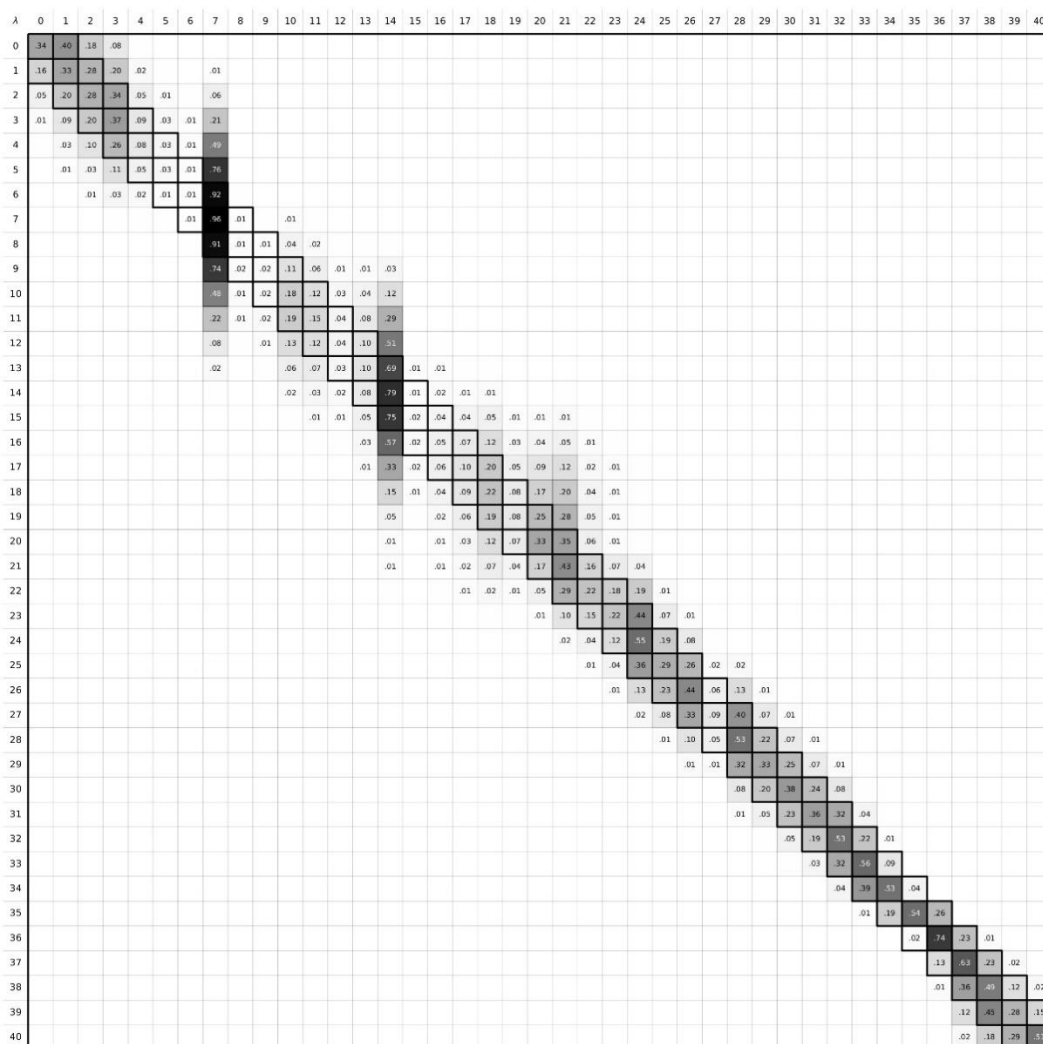


Figura A.2 – Matriz de sobreposição para a caixa de octanol úmido com o modelo OPC4 no campo de força AMBER/GAFF.

ANEXO B

Matrizes de sobreposições dos λ nos cálculos de energia livre utilizando o método de integração termodinâmica para o modelo de água OPC4 no campo de força CHARMM/CGenFF.

λ	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0	.05	.93	.03																							
1		.74	.25																							
2		.25	.72		.02																					
3		.02	.46	.01	.52																					
4			.02		.91		.06																			
5					.53		.45	.01																		
6				.06			.71	.20	.02																	
7						.20	.50	.25	.05																	
8						.02	.25	.45	.28																	
9							.05	.28	.65	.01	.02															
10								.07	.81	.03	.08	.01														
11								.06	.60	.02	.24	.06	.01													
12								.03	.25	.01	.31	.25	.14													
13									.02	.07	.68	.23														
14										.04	.93	.03														
15											.26	.59	.15													
16											.08	.74	.17	.01												
17											.07	.61	.29	.03												
18											.27	.50	.20	.02												
19											.04	.28	.50	.18	.01											
20											.03	.26	.56	.14												
21												.02	.25	.64	.10											
22													.15	.65	.19											
23														.09	.70	.20										
24															.17	.69	.13									
25																.01	.34	.65								

Figura B.1 – Matriz de sobreposição para a caixa de água com o modelo OPC4 no campo de força CHARMM/CGenFF.

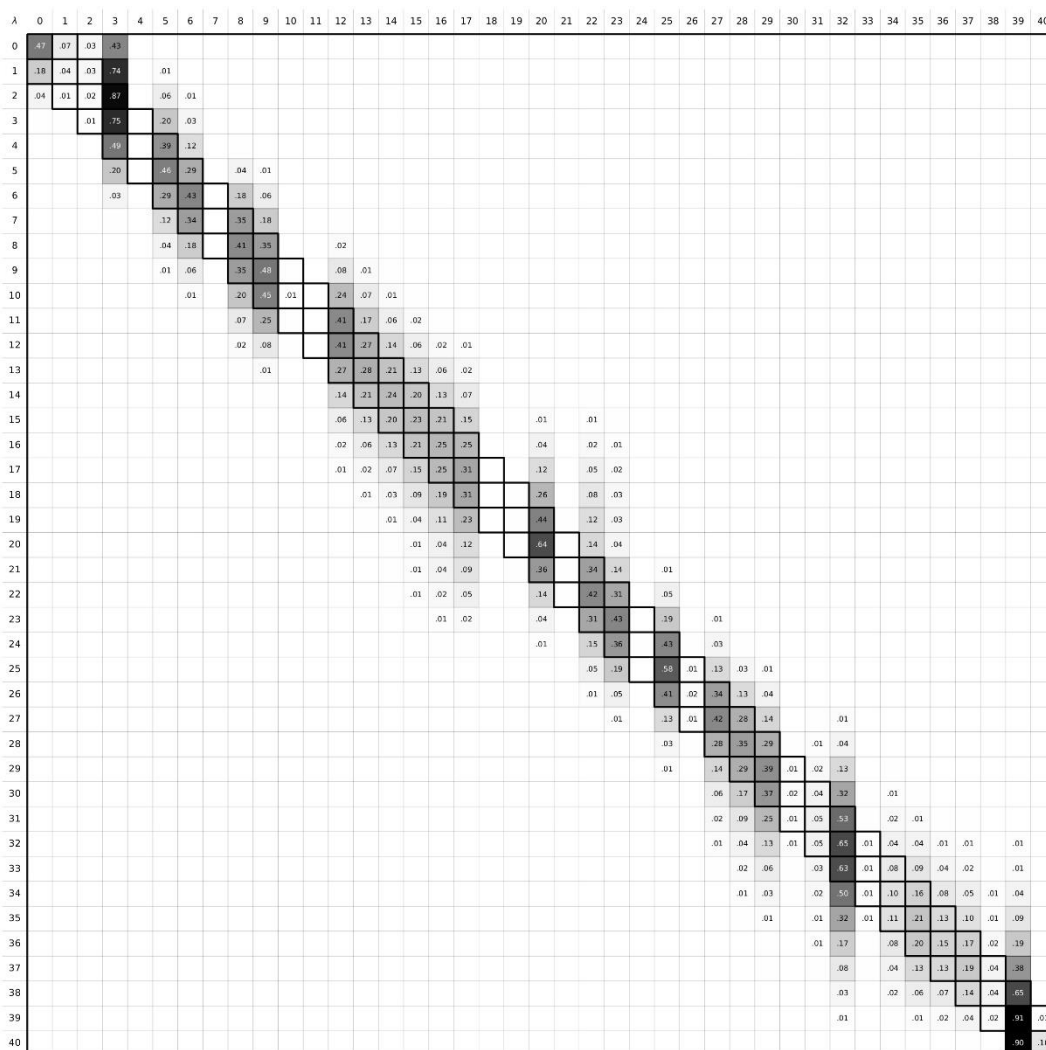


Figura B.2 – Matriz de sobreposição para a caixa de octanol úmido com o modelo OPC4 no campo de força CHARMM/CGenFF.

ANEXO C

Gráfico de barras das diferenças de energia livre avaliadas entre pares de estados adjacentes por meio de vários métodos, com estimativas de erro correspondentes para cada método, para o modelo de água OPC4 no campo de força AMBER/GAFF.

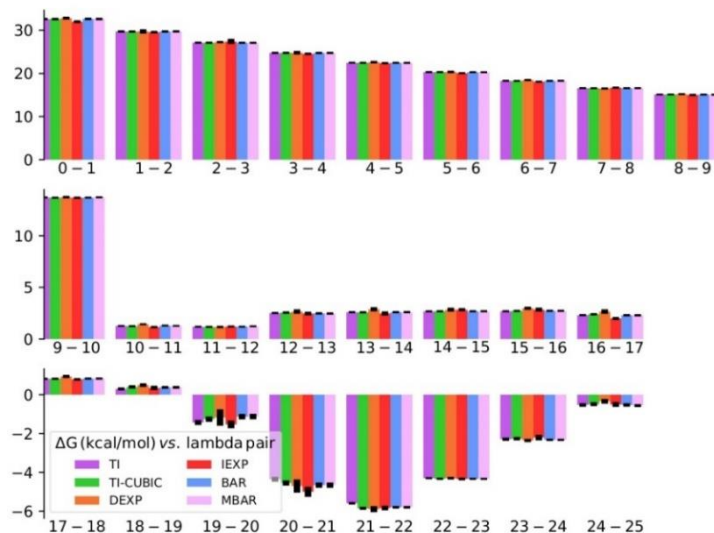


Figura C.1 – Diferenças de energia livre avaliadas entre pares de estados adjacentes por meio de vários métodos com seus respectivos erros, para a caixa de água com o modelo OPC4 no campo de força AMBER/GAFF.

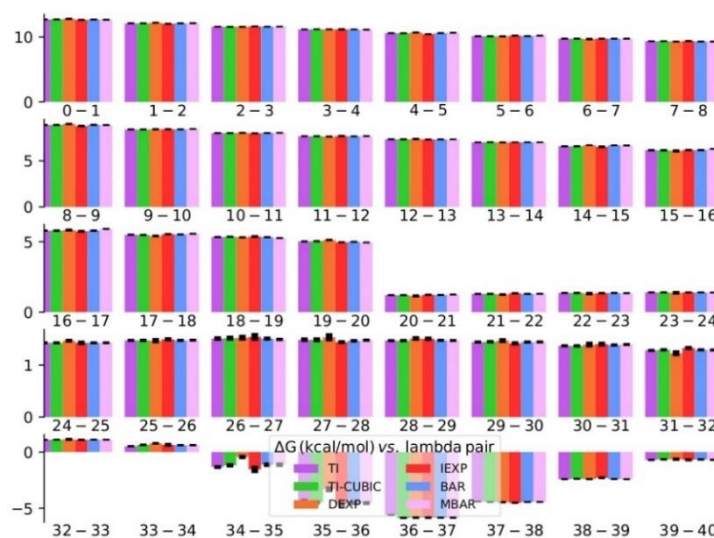


Figura C.2 – Diferenças de energia livre avaliadas entre pares de estados adjacentes por meio de vários métodos com seus respectivos erros, para a caixa de octanol úmido com o modelo OPC4 no campo de força AMBER/GAFF.

ANEXO D

Gráfico de barras das diferenças de energia livre avaliadas entre pares de estados adjacentes por meio de vários métodos, com estimativas de erro correspondentes para cada método, para o modelo de água OPC4 no campo de força CHARMM/CGenFF.

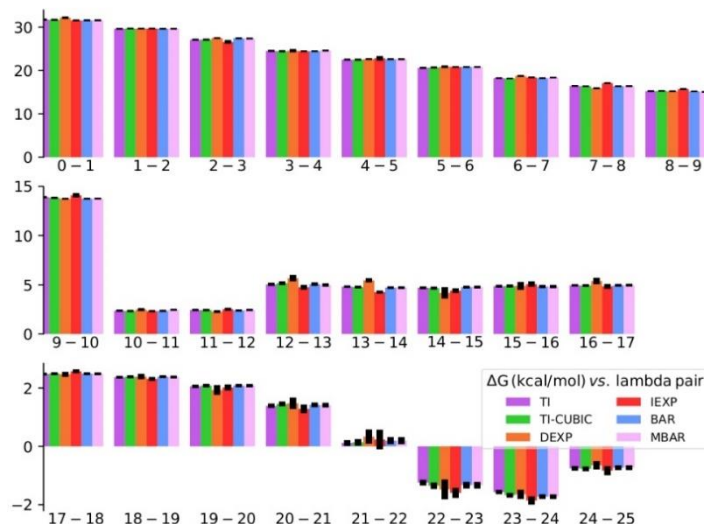


Figura D.1 –Diferenças de energia livre avaliadas entre pares de estados adjacentes por meio de vários métodos com seus respectivos erros, para a caixa de água com o modelo OPC4 no campo de força CHARMM/CGenFF.

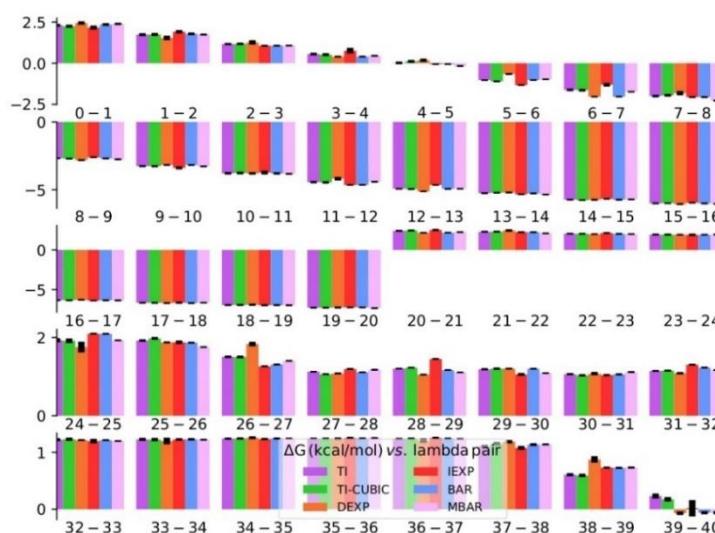


Figura D.1 – Diferenças de energia livre avaliadas entre pares de estados adjacentes por meio de vários métodos com seus respectivos erros, para a caixa de octanol úmido com o modelo OPC4 no campo de força CHARMM/CGenFF.

ANEXO E

DIVULGAÇÃO DO TRABALHO

Essa dissertação foi apresentada na forma de pôster durante o XVIII Encontro Regional da Sociedade Brasileira de Química do Rio de Janeiro (SBQ-Rio). Foi premiado como melhor pôster da sessão de química teórica.



Figura E.1 - Certificado de premiação de pôster do XVIII Encontro Regional da SBQ-Rio, em outubro de 2022.

Este trabalho de dissertação foi publicado na revista “Journal of Biomolecular Structure and Dynamics”. Cujo fator de impacto é de 5.235. Disponível no link: <https://doi.org/10.1080/07391102.2023.2173298>

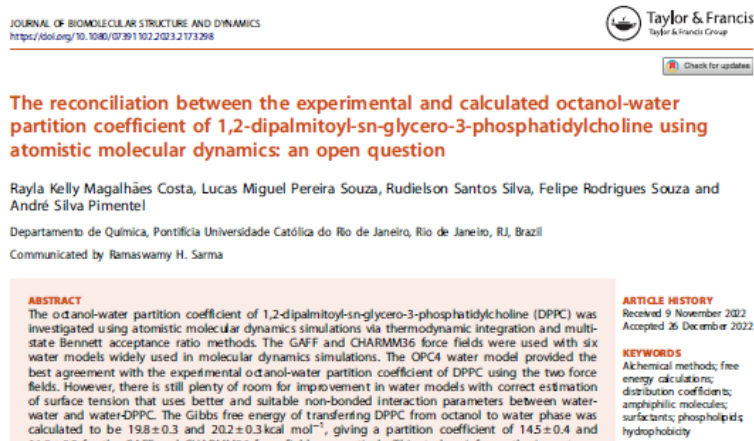


Figura E.2 – Artigo publicado

ANEXO F

ARQUIVOS DE ENTRADA DA INTEGRAÇÃO TERMODINÂMICA.

➤ Parâmetros de dinâmica molecular – minim.mdp

```

=====
; Energy minimization
=====
;-----
; RUN CONTROL & MINIMIZATION
;-----
define                = -DFLEXIBLE
integrator             = steep
nsteps                = 5000
emtol                 = 100
emstep                = 0.01
nstcomm               = 100
;-----
; OUTPUT CONTROL
;-----
nstxout               = 250      ; save coordinates to .trr every 250 steps
nstvout               = 0        ; don't save velocities to .trr
nstfout               = 0        ; don't save forces to .trr
nstxout-compressed    = 500      ; xtc compressed trajectory output every
500 steps
compressed-x-precision = 1000
nstlog                = 500      ; update log file every 500 steps
nstenergy             = 500      ; save energies every 500 steps
nstcalcenergy         = 100
;-----
; NEIGHBOR SEARCHING
;-----
cutoff-scheme         = Verlet
ns-type               = grid
nstlist               = 1
rlist                 = 1.0
;-----
; BONDS
;-----
constraints           = none
;-----
; ELECTROSTATICS
;-----
coulombtype           = PME
rcoulomb              = 1.0
pme-order             = 6
fourierspacing        = 0.10
ewald-rtol            = 1e-6
;-----

```



```

; VDW
;-----
vdw-type          = PME
rvdw              = 1.0
vdw-modifier      = Potential-Shift
ewald-rtol-lj     = 1e-3
lj-pme-comb-rule  = Geometric
DispCorr          = EnerPres
;-----
; TEMPERATURE & PRESSURE COUPL
;-----
Tcoupl            = no
Pcoupl            = no
gen_vel           = no
;-----
; FREE ENERGY CALCULATIONS
;-----
free-energy        = yes
couple-moltype    = ligand
couple-lambda0    = vdw-q
couple-lambda1    = none
couple-intramol   = yes
separate-dhdl-file = yes
sc-alpha          = 0.5
sc-power          = 1
sc-sigma          = 0.3
init-lambda-state = 0
coul-lambdas      = 0.0 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90
1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0
vdw-lambdas       = 0.0 0.00 0.0 0.00 0.0 0.00 0.0 0.00 0.0 0.00 0.0
0.05 0.1 0.2 0.3 0.4 0.5 0.6 0.65 0.7 0.75 0.8 0.85 0.9 0.95 1.0
nstdhdl           = 100
calc-lambda-neighbors = -1

```

➤ Parâmetros de dinâmica molecular - NVT.mdp

```

;=====
; NVT equilibration
;=====
;-----
; RUN CONTROL
;-----
define          = -DPOSRES
integrator      = sd          ; stochastic leap-frog integrator
nsteps         = 5000         ; 2 * 5,000 fs = 10 ps
dt             = 0.002        ; 2 fs
comm-mode      = Linear       ; remove center of mass translation
nstcomm        = 100          ; frequency for center of mass motion removal
;-----
; OUTPUT CONTROL

```

```

;-----
nstxout      = 0      ; don't save coordinates to .trr
nstvout      = 0      ; don't save velocities to .trr
nstfout      = 0      ; don't save forces to .trr
nstxout-compressed = 5000 ; xtc compressed trajectory output
every 5000 steps
compressed-x-precision = 1000 ; precision with which to write to the
compressed trajectory file
nstlog       = 5000    ; update log file every 10 ps
nstenergy    = 5000    ; save energies every 10 ps
nstcalcenergy = 100    ; calculate energies every 100 steps
;-----
; BONDS
;-----
constraint_algorithm = lincs ; holonomic constraints
constraints          = all-bonds ; hydrogens only are constrained
lincs_iter           = 1      ; accuracy of LINCS (1 is default)
lincs_order          = 4      ; also related to accuracy (4 is default)
lincs-warnangle      = 30     ; maximum angle that a bond can rotate
before LINCS will complain (30 is default)
continuation         = no     ; formerly known as 'unconstrained-start' -
useful for exact continuations and reruns
;-----
; NEIGHBOR SEARCHING
;-----
cutoff-scheme = Verlet
ns-type       = grid ; search neighboring grid cells
nstlist       = 10   ; 20 fs (default is 10)
rlist         = 1.0   ; short-range neighborlist cutoff (in nm)
pbc           = xyz   ; 3D PBC
;-----
; ELECTROSTATICS
;-----
coulombtype    = PME      ; Particle Mesh Ewald for long-range
electrostatics
rcoulomb       = 1.0      ; short-range electrostatic cutoff (in nm)
ewald_geometry = 3d       ; Ewald sum is performed in all three
dimensions
pme-order      = 6        ; interpolation order for PME (default is 4)
fourierspacing = 0.10     ; grid spacing for FFT
ewald-rtol     = 1e-6     ; relative strength of the Ewald-shifted direct
potential at rcoulomb
;-----
; VDW
;-----
vdw-type       = PME
rvdw           = 1.0
vdw-modifier    = Potential-Shift
ewald-rtol-lj   = 1e-3
lj-pme-comb-rule = Geometric

```

```

DispCorr          = EnerPres
;-----
; TEMPERATURE & PRESSURE COUPL
;-----
tc_grps   = System
tau_t     = 1.0
ref_t     = 298
pcoupl    = no
;-----
; VELOCITY GENERATION
;-----
gen_vel    = yes      ; Velocity generation is on (if gen_vel is 'yes',
continuation should be 'no')
gen_seed   = -1       ; Use random seed
gen_temp   = 298
;-----
; FREE ENERGY CALCULATIONS
;-----
free-energy      = yes
couple-moltype   = ligand
couple-lambda0   = vdw-q
couple-lambda1   = none
couple-intramol  = yes
separate-dhdl-file = yes
sc-alpha        = 0.5
sc-power        = 1
sc-sigma        = 0.3
init-lambda-state = 0
coul-lambdas     = 0.0 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90
1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0
vdw-lambdas      = 0.0 0.00 0.0 0.00 0.0 0.00 0.0 0.00 0.0 0.00 0.0
0.05 0.1 0.2 0.3 0.4 0.5 0.6 0.65 0.7 0.75 0.8 0.85 0.9 0.95 1.0
nstdhdl         = 100
calc-lambda-neighbors = -1

```

➤ Parâmetros de dinâmica molecular - NPT.mdp

```

;=====
; NVT equilibration
;=====
;-----
; RUN CONTROL
;-----
define      = -DPOSRES
integrator  = sd      ; stochastic leap-frog integrator
nsteps     = 50000    ; 2 * 50,000 fs = 100 ps
dt         = 0.002    ; 2 fs
comm-mode   = Linear   ; remove center of mass translation
nstcomm    = 100      ; frequency for center of mass motion removal
;-----
; OUTPUT CONTROL

```

```

;-----
nstxout      = 0      ; don't save coordinates to .trr
nstvout      = 0      ; don't save velocities to .trr
nstfout      = 0      ; don't save forces to .trr
nstxout-compressed = 5000 ; xtc compressed trajectory output
every 5000 steps
compressed-x-precision = 1000 ; precision with which to write to the
compressed trajectory file
nstlog       = 5000    ; update log file every 10 ps
nstenergy    = 5000    ; save energies every 10 ps
nstcalcenergy = 100    ; calculate energies every 100 steps
;-----
; BONDS
;-----
constraint_algorithm = lincs ; holonomic constraints
constraints          = all-bonds ; hydrogens only are constrained
lincs_iter           = 1      ; accuracy of LINCS (1 is default)
lincs_order          = 4      ; also related to accuracy (4 is default)
lincs-warnangle      = 30     ; maximum angle that a bond can rotate
before LINCS will complain (30 is default)
continuation         = yes    ; formerly known as 'unconstrained-start' -
useful for exact continuations and reruns
;-----
; NEIGHBOR SEARCHING
;-----
cutoff-scheme = Verlet
ns-type       = grid ; search neighboring grid cells
nstlist       = 10   ; 20 fs (default is 10)
rlist         = 1.0  ; short-range neighborlist cutoff (in nm)
pbc           = xyz  ; 3D PBC
;-----
; ELECTROSTATICS
;-----
coulombtype    = PME      ; Particle Mesh Ewald for long-range
electrostatics
rcoulomb       = 1.0      ; short-range electrostatic cutoff (in nm)
ewald_geometry = 3d       ; Ewald sum is performed in all three
dimensions
pme-order      = 6        ; interpolation order for PME (default is 4)
fourierspacing = 0.10     ; grid spacing for FFT
ewald-rtol     = 1e-6     ; relative strength of the Ewald-shifted direct
potential at rcoulomb
;-----
; VDW
;-----
vdw-type       = PME
rvdw           = 1.0
vdw-modifier    = Potential-Shift
ewald-rtol-lj   = 1e-3
lj-pme-comb-rule = Geometric

```

```

DispCorr          = EnerPres
;-----
; TEMPERATURE & PRESSURE COUPL
;-----
tc_grps          = System
tau_t            = 1.0
ref_t            = 298
pcoupl           = Berendsen
pcoupltype       = isotropic
tau_p            = 0.5          ; time constant (ps)
ref_p            = 1.0          ; reference pressure (bar)
compressibility   = 4.5e-05      ; isothermal compressibility of water
(bar^-1)
refcoord-scaling = all
;-----
; VELOCITY GENERATION
;-----
gen_vel          = no
;-----
; FREE ENERGY CALCULATIONS
;-----
free-energy       = yes
couple-moltype    = ligand
couple-lambda0    = vdw-q
couple-lambda1    = none
couple-intramol   = yes
separate-dhdl-file = yes
sc-alpha          = 0.5
sc-power          = 1
sc-sigma          = 0.3
init-lambda-state = 0
coul-lambdas      = 0.0 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90
1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0
vdw-lambdas       = 0.0 0.00 0.0 0.00 0.0 0.00 0.0 0.00 0.0 0.00 0.0 0.00 0.0
0.05 0.1 0.2 0.3 0.4 0.5 0.6 0.65 0.7 0.75 0.8 0.85 0.9 0.95 1.0
nstdhdl           = 100
calc-lambda-neighbors = -1

```

➤ Parâmetros de dinâmica molecular - PROD.mdp

```

=====
; Production simulation
=====
;-----
; RUN CONTROL
;-----
integrator        = sd          ; stochastic leap-frog integrator
nsteps           = 2500000      ; 2 * 2,500,000 fs = 5000 ps
dt               = 0.002        ; 2 fs
comm-mode        = Linear       ; remove center of mass translation
nstcomm          = 100          ; frequency for center of mass motion removal

```

```

;-----
; OUTPUT CONTROL
;-----
nstxout      = 0      ; don't save coordinates to .trr
nstvout      = 0      ; don't save velocities to .trr
nstfout      = 0      ; don't save forces to .trr
nstxout-compressed = 1000 ; xtc compressed trajectory output
every 1000 steps (2 ps)
compressed-x-precision = 1000 ; precision with which to write to the
compressed trajectory file
nstlog       = 1000   ; update log file every 2 ps
nstenergy    = 1000   ; save energies every 2 ps
nstcalcenergy = 100   ; calculate energies every 100 steps
;-----
; BONDS
;-----
constraint_algorithm = lincs ; holonomic constraints
constraints         = all-bonds ; hydrogens only are constrained
lincs_iter          = 1      ; accuracy of LINCS (1 is default)
lincs_order         = 4      ; also related to accuracy (4 is default)
lincs-warnangle     = 30     ; maximum angle that a bond can rotate
before LINCS will complain (30 is default)
continuation        = yes
;-----
; NEIGHBOR SEARCHING
;-----
cutoff-scheme = Verlet
ns-type       = grid ; search neighboring grid cells
nstlist       = 10   ; 20 fs (default is 10)
rlist         = 1.0   ; short-range neighborlist cutoff (in nm)
pbc           = xyz   ; 3D PBC
;-----
; ELECTROSTATICS
;-----
coulombtype    = PME      ; Particle Mesh Ewald for long-range
electrostatics
rcoulomb       = 1.0      ; short-range electrostatic cutoff (in nm)
ewald_geometry = 3d       ; Ewald sum is performed in all three
dimensions
pme-order      = 6        ; interpolation order for PME (default is 4)
fourierspacing = 0.10     ; grid spacing for FFT
ewald-rtol     = 1e-6     ; relative strength of the Ewald-shifted direct
potential at rcoulomb
;-----
; VDW
;-----
vdw-type       = PME
rvdw           = 1.0
vdw-modifier    = Potential-Shift
ewald-rtol-lj  = 1e-3

```

```

lj-pme-comb-rule      = Geometric
DispCorr              = EnerPres
;-----
; TEMPERATURE & PRESSURE COUPL
;-----
tc_grps               = System
tau_t                 = 1.0
ref_t                 = 298
pcoupl                = Berendsen
pcoupltype            = isotropic      ; uniform scaling of box vectors
tau_p                 = 2              ; time constant (ps)
ref_p                 = 1.0            ; reference pressure (bar)
compressibility        = 4.5e-05        ; isothermal compressibility of water
(bar^-1)
;-----
; VELOCITY GENERATION
;-----
gen_vel               = no            ; Velocity generation is off (if gen_vel is 'yes',
continuation should be 'no')
gen_seed              = -1            ; Use random seed
gen_temp              = 298
;-----
; FREE ENERGY CALCULATIONS
;-----
free-energy            = yes
couple-moltype         = ligand
couple-lambda0         = vdw-q
couple-lambda1         = none
couple-intramol        = yes
separate-dhdl-file     = yes
sc-alpha               = 0.5
sc-power               = 1
sc-sigma               = 0.3
init-lambda-state      = 0
coul-lambdas           = 0.0 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90
1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0 1.00 1.0
vdw-lambdas            = 0.0 0.00 0.0 0.00 0.0 0.00 0.0 0.00 0.0 0.00 0.0
0.05 0.1 0.2 0.3 0.4 0.5 0.6 0.65 0.7 0.75 0.8 0.85 0.9 0.95 1.0
nstdhdl                = 100
calc-lambda-neighbors  = -1

```

ARQUIVOS DE ENTRADA AMOSTRAGEM GUARDA CHUVA

➤ Parâmetros de dinâmica molecular – md_pull.mdp

```

title      = Umbrella pulling simulation
;define    = -DPOSRES
integrator = md
dt         = 0.002
tinit      = 0
nsteps     = 1000000 ; 2 ns

```

```

nstcomm      = 10
; Output parameters
nstxout      = 5000      ; every 10 ps
nstvout      = 5000
nstfout      = 500
nstxtcout    = 500      ; every 1 ps
nstenergy    = 500
; Bond parameters
constraint_algorithm = lincs
constraints   = h-bonds
continuation  = yes      ; continuing from NPT
; Single-range cutoff scheme
cutoff-scheme = Verlet
nstlist      = 20
ns_type      = grid
rlist        = 1.4
rcoulomb     = 1.4
rvdw         = 1.4
; PME electrostatics parameters
coulombtype  = PME
fourierspacing = 0.12
fourier_nx   = 0
fourier_ny   = 0
fourier_nz   = 0
pme_order    = 4
ewald_rtol   = 1e-5
optimize_fft = yes
; Berendsen temperature coupling is on in two groups
Tcoupl       = V-rescale
tc_grps      = DPPC SOL OCT
tau_t        = 0.1 0.1 0.1
ref_t        = 298 298 298
; Pressure coupling is on
Pcoupl       = Parrinello-Rahman
pcoupltype   = isotropic
tau_p        = 2.0
compressibility = 4.5e-5
ref_p        = 1.0
;refcoord_scaling = com
; Generate velocities is off
gen_vel      = no
; Periodic boundary conditions are on in all directions
pbc          = xyz
; Long-range dispersion correction
DispCorr     = EnerPres
; Pull code
pull         = yes
pull_ncoords  = 1      ; only one reaction coordinate
pull_ngroups  = 2      ; two groups defining one reaction coordinate
pull_group1_name = DPPC

```



```

pull_group2_name      = OCT
pull_group2_pbcatom   = -1
pull_coord1_type      = umbrella ; harmonic potential
pull_coord1_geometry  = direction ; simple distance increase
pull_coord1_dim       = N N Y
pull_coord1_groups    = 1 2
pull_coord1_vec       = 0.0 0.0 -1.0
pull_coord1_start     = yes      ; define initial COM distance > 0
pull_coord1_rate      = 0.004   ; 0.004 nm per ps = 4 nm per ns
pull_coord1_k         = 1000    ; kJ mol-1 nm-2

```

➤ **Parâmetros de dinâmica molecular – NPT_umbrella.mdp**

```

title      = Umbrella pulling simulation
;define    = -DPOSRES_P
; Run parameters
integrator = md
dt         = 0.002
tinit      = 0
nsteps     = 10000000 ; 20ns
nstcomm    = 10
; Output parameters
nstvout     = 25000 ; every 50 ps
nstfout     = 25000
nstxout-compressed = 25000
nstenergy   = 25000
; Bond parameters
constraint_algorithm = lincs
constraints    = h-bonds
continuation   = no
; Single-range cutoff scheme
cutoff-scheme = Verlet
nstlist       = 20
ns_type       = grid
rlist         = 1.4
rcoulomb      = 1.4
rvdw          = 1.4
; PME electrostatics parameters
coulombtype   = PME
fourierspacing = 0.12
fourier_nx    = 0
fourier_ny    = 0
fourier_nz    = 0
pme_order     = 4
ewald_rtol    = 1e-5
optimize_fft  = yes
; Berendsen temperature coupling is on in two groups
Tcoupl       = Berendsen
tc_grps      = DPPC SOL OCT
tau_t        = 0.1 0.1 0.1
ref_t        = 298 298 298

```

```

; Pressure coupling is on
Pcoupl      = Berendsen
pcoupltype   = isotropic
tau_p       = 1.0
compressibility = 4.5e-5
ref_p       = 1.0
;refcoord_scaling = com
; Generate velocities is on
gen_vel     = yes
gen_temp    = 298
; Periodic boundary conditions are on in all directions
pbc        = xyz
; Long-range dispersion correction
DispCorr    = EnerPres
; Pull code
pull        = yes
pull_ncoords      = 1      ; only one reaction coordinate
pull_ngroups      = 2      ; two groups defining one reaction coordinate
pull_group1_name   = DPPC
pull_group2_name   = SOL
pull_group2_pbcatom = -1
pull_coord1_type    = umbrella ; harmonic potential
pull_coord1_geometry = direction ; simple distance increase
pull_coord1_dim     = N N Y
pull_coord1_groups  = 1 2
pull_coord1_vec     = 0.0 0.0 -1.0
pull_coord1_start   = yes      ; define initial COM distance > 0
pull_coord1_rate    = 0.0      ; restrain in place
pull_coord1_k       = 2000     ; kJ mol-1 nm-2

```

➤ **Parâmetros de dinâmica molecular – md_umbrella.mdp**

```

title      = Umbrella pulling simulation
;define    = -DPOSRES_P
; Run parameters
integrator = md
dt         = 0.002
tinit      = 0
nsteps     = 500000 ; 1 ns
nstcomm    = 10
; Output parameters
nstxout-compressed = 500 ; every 1 ps
nstenergy      = 500
; Bond parameters
constraint_algorithm = lincs
constraints         = h-bonds
continuation        = yes
; Single-range cutoff scheme
cutoff-scheme       = Verlet
nstlist            = 20

```

```

ns_type      = grid
rlist        = 1.4
rcoulomb     = 1.4
rvdw         = 1.4
; PME electrostatics parameters
coulombtype  = PME
fourierspacing = 0.12
fourier_nx   = 0
fourier_ny   = 0
fourier_nz   = 0
pme_order    = 4
ewald_rtol   = 1e-5
optimize_fft = yes
; Berendsen temperature coupling is on in two groups
Tcoupl       = V-Rescale
tc_grps      = DPPC SOL OCT
tau_t        = 0.1 0.1 0.1
ref_t        = 298 298 298
; Pressure coupling is on
Pcoupl       = Parrinello-Rahman
pcoupltype   = isotropic
tau_p        = 2.0
compressibility = 4.5e-5
ref_p        = 1.0
; refcoord_scaling = com
; Generate velocities is off
gen_vel      = no
; Periodic boundary conditions are on in all directions
pbc          = xyz
; Long-range dispersion correction
DispCorr     = EnerPres
; Pull code
pull          = yes
pull_ncoords  = 1      ; only one reaction coordinate
pull_ngroups  = 2      ; two groups defining one reaction coordinate
pull_group1_name = DPPC
pull_group2_name = SOL
pull_group2_pbcatom = -1
pull_coord1_type = umbrella ; harmonic potential
pull_coord1_geometry = direction ; simple distance increase
pull_coord1_dim = N N Y
pull_coord1_groups = 1 2
pull_coord1_vec = 0.0 0.0 -1.0
pull_coord1_start = yes ; define initial COM distance > 0
pull_coord1_rate = 0.0 ; restrain in place
pull_coord1_k = 500 ; kJ mol-1 nm-2

```

ANEXO G

Modelo vs. propriedades experimentais da água em condições ambientais (298,16 K, 1 bar): O ângulo entre as ligações H-O, raio da ligação H-O, carga de cada átomo, momento dipolo, Densidade, constante dielétrica e Calor de vaporização.

	SPC	TIP3P	TIP4P	TIP5P	OPC3	OPC4	EXP
Θ_{HOH}	109,47	104,52	104,52	104,52	109,47	103,60	104,52
$r_{\text{O-H}}$	0,1000	0,0957	0,0957	0,0957	0,0978	0,0872	0,09584
q_{H}	+0,410	+0,417	+0,520	+0,241	+0,447	0,679	-
q_{O}	-0,820	-0,834	-1,040	-0,241	-0,8951	-1,358	-
Dipolo	2,352	2,348	2,320	2,290	2,43	2,48	2,5–3
ρ	0,994	0,980	0,995	0,979	0,996	0,997	0,997
ϵ_0	68	94	63,90	92	78,4	78,4	78,4
ΔH_{vap}	10,43	10,26	10,58	10,46	10,73	10,57	10,52

Θ_{HOH} (graus): Ângulo entre as ligações H-O

$r_{\text{O-H}}$ (nm): Raio da ligação H-O

q_{H} e q_{O} (elétrons): Carga de cada átomo

Dipolo (D): Momento dipolo

ρ [g/cm³]: Densidade

ϵ_0 : Constante dielétrica

ΔH_{vap} [kcal/mol]: Calor de vaporização